UGC MODEL CURRICULUM STATISTICS

SUBMITTED TO THE UNIVERSITY GRANTS COMMISSION IN APRIL 2001

Introduction

In September, 2000 the University Grants Commission constituted a new Curriculum Development Committee in Statistics. The *Core Committee* consisted of:—

1. Prof. J. V. Deshpande, University of Pune (Convener)
2. Prof. N. R. Mohan, University of Mysore
3. Prof. G. K. Shukla, Indian Institute of Technology, Kanpur, and
4. Prof. B. K. Sinha, Indian Statistical Institute, Calcutta.

The following were further coopted as PG members and the first meeting of the extended Committee was held on November 13/14, 2000.

(a) **Postgraduate Committee**

1. Prof. R. L. Karandikar, Indian Statistical Institute, New Delhi
2. Prof. S. P. Mukherjee, Calcutta University,
3. Dr A. Sarkar, IBM India Research Lab, New Delhi
4. Prof. T. Srivenkataramana, Bangalore University,

In this meeting it was decided to co-opt the following as UG members and they were invited to attend the next two meetings of the Committee.

(b) **Undergraduate Committee**

1. Dr B. Das, Presidency College, Calcutta,
2. Dr T. M. Durairajan, Loyola College, Chennai,
3. Dr B. R. Gupta, Venkateswara College, New Delhi
4. Dr. K. K. Madhava Rao, Bhandarkar's College, Kundapura,
5. Dr Ram Karan Singh, Lucknow University.

In the first meeting held on November 13/14, 2001, visualizing the vast diversity in different branches of Statistics, it was decided to seek further help from the following colleagues in drafting the syllabi for the various papers. They were:

1. Dr S. Adak, IBM India Research Lab, New Delhi,
2. Dr A. D. Dharmadhikari, University of Pune,
3. Prof. A. P. Gore, University of Pune,

4. Prof. P. C. Joshi, Indian Institute of Technology, Kanpur,
5. Prof. B. K. Kale, University of Pune,
6. Prof. S. Kunte, University of Pune,
7. Dr Chitra Lele, Pfizer (India) Ltd., Mumbai,
8. Prof. P. P. Mazumdar, Indian Statistical Institute, Calcutta,
9. Prof. S. M. Manjunath, Bangalore University,
10. Dr N. K. Mondal, Calcutta University,
11. Dr U. V. Naik-Nimbalkar, University of Pune
12. Dr M. B. Rajarshi, University of Pune,
13. Dr S. S. Sengupta, Calcutta University,
14. Prof. R. Vasudeva, University of Mysore.

The committees met subsequently on 15-16-17 January, 2001 and 28 Feb. -1 - 2 March, 2001. Much of the detailed work was done by the Committee members and the other colleagues at their headquarters. It was discussed and finalized in the committee meetings at New Delhi. The Committee had the benefit of discussions with Prof. A. S. Nigavekar, Vice-Chairman, UGC, during its first meeting in Nov. 2000.

The Convener reports with great pleasure and a sense of fulfillment that the utmost cooperation in this task was received from the core members, the co-opted members as well as the other colleagues listed above. The Convener thanks them heartily for their devotion to the subject and hard work. The syllabi provided here are the result of this joint effort. At the end of this report certain operative recommendations are made to the Commission. They are also the result of the concensus of the members of the Committee.

The Committee had to take into account the changes which have occurred since the publication of the last CDC report in 1989, in (a) the subject matter, (b) the revolution in computational power, (c) the general societal changes in which IT has come to play a very important role and (d) usage of Statistics in industries in the new economic order. The Committee has striven hard to take cognizance of all these aspects and has suggested the enclosed undergraduate and postgraduate syllabi, accordingly.

Now it is well recognized that the days of manual, electrical and electronic calculators have passed and Statistics is being developed and will be applied in the new Century only with the help of computers (both PC and mainframe). With this major transformation in view the Committee has suggested a full integration of computers in the instruction. This is the only way in which Statistics should now be taught. Any University not having adequate laboratory (computation) facility should not be allowed to start instruction in Statistics, at least at the postgraduate level.

Review of the Existing Syllabi

The Committee had the benefit of the CDC report finalized in November 1988 under the Chairmanship of the late Prof. C. G. Khatri. Also, the UGC Panel for Statistics constituted in 1994 had conducted a survey of the Centres of post-graduate teaching in India. It had received information including the current syllabi from about 30 universities. The members of the current CDC (the core group and the co-opted members) were from Centres situated in different parts in India and had intimate personal knowledge of the state of affairs in their own University as well as neighbouring universities. The UGC Panel for Statistics had convened two meetings (one for northern States and the other for the Southern States) of the Conveners of the Postgraduate and Undergraduate Boards of Studies of the Universities in 1999-2000. About 60 Universities were represented. Their comments and suggestions regarding the syllabi were recorded and were available to the CDC.

It is noted that the 1988 CDC report has an exhaustive prefatory statement. We do not repeat the points made in it. Here we only emphasize the changes that have been deemed appropriate to be carried out in the following syllabi.

Undergraduate Syllabi

It is recognized that purpose of undergraduate teaching is primarily to prepare the students for post-graduate instruction and also to prepare them to hold (subordinate) positions dealing with statistical analysis. Hence these courses have to be in tune with the latest developments in the subject and also take cognizance of the newer techniques of statistical analysis which are now totally based on computers. It is taken for granted that the students will offer Mathematics as another subject at least at the General / Ancillary level.

Undergraduate B.A./B.Sc. (General / Ancillary) Course

It is noted that the structure in most of the Universities having the general degree or having Statistics as an ancillary subject with another principal subject continues to be as 2 theory papers of 100 marks each + 1 practical of 100 marks per year contributing 300 marks in each of the three years of the B.A./B.Sc. course, making up a total of 6 theory papers and 3 practical papers.

It is recognized that the students learning Statistics at the Ancillary/General level may be specializing in another subject. To do quantitative work in any subject familiarity with basic statistical methodology is necessary. While drafting this syllabus this need has been kept in view.

The contents of the subject of Statistics have now stabilized to a considerable extent. Hence the basic framework of this course remains the same as included in the 1989

syllabus. However, in view of the experience gained in these years certain modifications have been carried out emphasizing certain areas and deemphasizing certain other areas. Also a much larger role has been given to computer based methods. Familarity with computers, computer programming and software packages is sought to be achieved in this course. This has been done as, now-a-days all the work of statistical nature is carried out through the medium of computers and this familiarity and training has become a necessary part of statistical education.

B.A/B.Sc. (General/Ancillary) Course in Statistics

Paper Number	Title	Full or Half Year	Contact Hours per week	Marks
FIRST YEAR				
101	Probability Theory	Full	3	100
102	Descriptive Statistics	Full	3	100
103	Practical	Full	3	100
SECOND YEAR				
201	Statistical Methods	Full	3	100
202	A : Sample Surveys	Half	3	50
	B : Analysis and Design of Experiments	Half	3	50
203	Practical	Full	3	100
THIRD YEAR				
301	Applied Statistics	Full	3	100
302	A : Statistical Quality Control	Half	3	50
	B : Computational Techniques	Half	3	50
303	Practical	Full	3	100

Undergraduate B.A./B.Sc. (Honours / Special) Course

Here the structure usually is 2 theory papers of 100 marks each + 1 practical paper of 100 marks in the first two years and 4 theory papers of 100 marks each + 2 practical papers of 100 marks each in the third year making up a total of 8 theory papers and 4 practical papers in the three years.

Once again it is noted that due to the stability achieved in the fundamental aspects of the subject, the changes proposed during the first two years are mainly of organizational nature with certain modernization being introduced in terms of use of computers. In the third year it is now suggested that only 2 theory papers out of 4 and 1 practical out of the 2 should be compulsory. 1 theory paper and ½ practical (50 marks) should be offered out a list of several elective papers and the students be required to do a project worth the remaining 150 marks. Only if the project is not at all feasible then they may be allowed to choose another elective theory paper (100 marks) and the corresponding practical (50 marks). These elective courses contain several new and important topics such as Statistical computation, Environmental Statistics, Applied Decision Making, etc. The practical worth 50 marks will be based on the Elective paper offered. The exact syllabus of the practical is left to the respective Boards of Studies. A set of guidelines for the Project Work are provided at the end of the detailed B.A./B.Sc. (Honours) syllabus.

B.A/B.Sc. (Honours) Course in Statistics

Paper Number	Title	Full or Half Year	Contact Hours per week	Marks
FIRST YEAR				
101	Probability I	Half	3	50
	B : Mathematical Methods II	Half	3	50
102	Descriptive Statistics	Full	3	100
103	Practical	Full	3	100
SECOND YEAR				
201	A : Probability II	Half	3	50
	B : Mathematical Methods II	Half	3	50
202	Statistical Inference	Full	3	100
203	Practical	Full	3	100
THIRD YEAR				
301	Applied Statistics	Full	3	100
302	Design of Experiments, Linear Models and Sample Surveys Official Statistics	Full	3	100
303	Practical	Full	3	100
304	Elective Paper	Full	3	100
305	Practical Based on the Elective paper	Full	1½	50
306	Project	Full	4½	150

LIST OF ELECTIVES

1. Elementary Stochastic Processes
2. Statistics for Environment
3. Industrial Statistics and Operational Management
4. Computation Applications
5. Applied Decision Making

N.B. The University Boards of Studies in Statistics may add further elective courses to this list.

Post-Graduate M.A./M.Sc. Syllabi

The new developments in the subject as well as the changes in emphasis are to be reflected to a much larger extent in the PG course as this course has several purposes. One is that it brings the students almost to the threshold of new research in the subject (with the help of pre-Ph.D. or M.Phil. courses) and secondly it prepares the students to hold responsible and superior positions which use advanced statistical methodology.

Expanding role of Statistical usage

The role of statistics as the technology of data analysis and decision aking under uncertainty has expanded vastly in the past few years. Decisions regarding planning, development, spread of services, etc., whether on the national canvass or at a local scale are all based on statistical analysis. Also, developmental industries, such as biopharmaceutics, insurance, investments, market research etc. are major users of statistical methodology. In the post – liberaliazation scenerio, in which we are committed to the WTO and GATT regimes of relatively free international movement in goods and services, it is imperative that the productivity of our processes and quality and reliability of our products match and even surpass the international standards. Statistical methods play an important role in these matters as laid down in the current ISO standards.

Need for advanced computational support

In the 21st century, as already discussed in the section on the UG syllabus, all statistical work will be carried out through the medium of fast computers. There is no role at all for old fashioned mechanical or hand held electronic calculators which have been the mainstay of our laboratories for the last 50 years. Therefore this syllabus lays the necessary stress on the computerization of the whole course. It will help the students to use effectively the tools of the IT in their work.

Major changes in the structure of the syllabus

To meet the above challanges several radical changes have been proposed in the existing syllabi. Firstly, the concept of having seperate papers in practicals has been abondoned. All the numerical / practical work has been integrated with the teaching of theory courses. Secondly, after teaching 12 semester papers in the core areas, it is proposed that6 further papers be selected from within specialized modules. Six such modules have been proposed herewith. Thirdly it is suggested that all the numerical / practical work be carried out on computers. Pentium level PC's (one for every 2 or 3 students admitted to the course) are seen to be the best vehicles for carrying out this major changeover at this time. (April, 2001).

Since the practical and theory work is being amalgamated in many papers the examinations in such papers will have to be appropriately modified. Each such examination should be conducted in two parts, possibly at different times, apportioning the total marks

in proper proportion and giving adequate time to each part. The details will have to be worked out by the respective Board of Studies.

The last component of the course is project work equivalent to two semester papers. Carrying out of projects (they may involve field work or not) is recognized to be the best way of developing the creative, problem solving and innovative faculties of the students. It brings forth the students' abilities to cope with real problems, be they of research (theoretical) or of practical nature. It helps them to become more proficient in communication skills (written as well as oral) and teaches them to work in diverse environments. The tradition of projects has long been established in other technological courses. It is sought to be introduced in Statistics through the proposed syllabus. Thus in summary the structure is as follows.

M.A./M.Sc. Statistics Syllabus

Course No.	Title	Contact hours per week
Semester I		
101	Real Analysis	3 lectures, 1 tutorial
102	Linear Algebra	3 lectures, 1 tutorial, 1 practical
103	Distribution Theory	3 lectures, 1 tutorial
104	Sample Surveys and Statistics for National Development	3 lectures, 2 Practicals
105	Statistical Computing	2 lectures, 3 practicals
Semester II		
201	Probability	3 lectures, 1 tutorial
202	Stochastic Processes	3 lectures, 1 tutorial
203	Inference I	3 lectures, 1 tutorial, 1 practical
204	Linear Models and Regression Analysis	3 lectures, 1-2 practicals
205	Multivariate Analysis	3 lectures, 2 practicals
Semester III		
301	Inference II	3 lectures, 1 tutorial, 1 practical
302	Design and Analysis of of Experiments	3 lectures, 2 practicals
303	Course from the selected module	
304	Course from the selected module	
305	Project	5 hours per week
Semester IV		
401	Course from the selected module	
402	-do-	
403	-do-	
404	-do-	
405	Project	5 Hours per week

The Modular Structure

The following 6 modules have been proposed. Except the first, all the remaining modules include extensive statistical computing.

M-1 Probability

M-2 Advanced Statistical Analysis

M-3 Industrial Statistics and Operations Research

M-4 Biostatistics

M-5 Computational Statistics

M-6 Economic and Financial Statistics

The proposed papers in these modules and their detailed syllabi have been listed later on. The modules are not nonoverlapping. However, it is felt that as far as possible the students should be allowed to offer papers from within a single module. Each Statistics Department will have to decide how many and which of these modules should actually be taught as it may not be feasible to provide instruction in all of them. Such a choice will have to be made keeping in view the specializations of the faculty and considerations of employability of the students.

The first module is meant to go deep in the area of Probability. This will allow the students to proceed towards a research / teaching career. The second module presents advanced statistical analyses of very general applicability. Teaching this module will serve a dual purpose. The students will gain knowledge of contemporary developments in statistical analysis enabling them to make further contributions to these as well as to use the advanced methodology in new and innovative manner to emerging applied problems. Modules 3, 4 and 6 are specializations of statistical methodology as applied to manufacturing industries, biopharmeceuties, and areas such as insurance, investments etc. Departments of Statistics envisaging a demand for personnel trained to undertake such statistical analysis should establish instruction in these modules. The fifth module - Computational Statistics is in fact the interface between Statistics, Computer Science and their applications. The courses in this module are so devised that after proper instruction the students should be capable of (i)carrying out statistical analysis using modern statistical softwares, (ii) developing softwares for use in statistical purposes, (iii) acquiring enough proficiency in computer science and programming so that they can be absorbed in the general software / IT sector. Hence the parts of the country where the IT sector is active will see a demand for students trained in these areas.

As stated earlier the Departments should offer courses from as many of the modules as possible subject to local conditions. While not totally barring students from taking courses across the modules, it should be seen that a group of allied courses is offered as far as possible. The allied courses could also be from other disciplines. For examples, the students

wishing to specialize in the Biostatistics module could be advised to take a single semester introductory paper in 'Life Sciences', similarly those specializing in Economic / Financial Statistics may take a course in 'Introduction to Economics'. These allied courses may be offered by other Departments of the University.

If a student takes at least 4 elective courses from the same module then he may be certified to have specialized in that module. The project work should be related to the module of specialization.

Module specific elective papers for the M.A./M.Sc. (Statistics) Course.
(L : Lecture, T : Tutorial, P : Practical)

Module 1 : Probability

Courses	Total Contact Hours
M.1.1 : Advanced Distribution theory	45L + 15T
M.1.2 : Functional Analysis	45L + 15T
M.1.3 : Martingales	45L + 15T
M.1.4 : Probability Measures on Metric Spaces	45L + 15T
M.1.5 : Inference in Stochastic Processes	45L + 15T
M.1.6 : Aplied Stochastic Models	45L + 15P
M.1.7 : Advanced Stochastic Processes	45L + 15T

Module 2 : Advanced Statistical Analysis

Courses	Total Contact Hours
M.2.1 : Advanced Distribution Theory (same as M.1.1)	45L + 15T
M.2.2 : Statistical Decision Theory	45L + 5T
M.2.3 : Bayesian Inference	45L + 15T
M.2.4 : Nonparametric and Semiparametric Methods	45L + 15P
M.2.5 : Advanced Multivariate Analysis	45L + 20P
M.2.6 : Advanced Sampling Theory	45L + 30P
M.2.7 : Time Series Analysis (same as (M.6.1)	45L + 20P
M.2.8 : Applied Regression Analysis	45L + 20P
M.2.9 : Advanced Design and Analysis of Experiments	45L + 20P
M.2.10 : Applied Stochastic Models (same as M.1.6)	45L + 15P

Module 3 : Industrial Statistis and Operations Research

Courses	Total Contact Hours
M.3.1 : Statistical Process and Quality Control	45L + 20P
M.3.2 : Planning and Analysis of Industrial Experiments	45L + 30P
M.3.3 : Reliability Theory	45L + 16P
M.3.4 : Statistical Methods for total Quality Management	45L + 18P
M.3.5 : Operations Research I	45L + 28P
M.3.6 : Operations Research II	45L + 24P

Module 4 : Biostatistics

Courses	Total Contact Hours
M.4.1 : Bioassay	45L + 10P
M.4.2 : Statistical Genetics	45L + 10P
M.4.3 : Survival Analysis	45L + 15P
M.4.4 : Statistical Ecology	45L + 15P
M.4.5 : Quantitative Epidemiology	45L + 15P
M.4.6 : Clinical Trials	45L + 18P
M.4.7 : Demography	45L + 10P
M.4.8 : Generalized Linear Models	45L + 18P
M.4.9 : Applied Regression Analysis (same as M.2.8)	45L + 20P

Module 5 : Computational Statistics

Course	Total Contact Hours
M.5.1 : Computer Intensive Statistical Methods I	45L + 25P
M.5.2 : Computer Intensive Statistical Methods II	45L + 20P
M.5.3 : Computer Programming	35L + 30P
M.5.4 : Knowledge Discovery and Data Mining	45L + 16P
M.5.5 : Statistical Pattern Recognition	45L + 16P

Module 6 : Economic and Financial Statistics

Course	Total Contact Hours
M.6.1 : Time Series Analysis	45L + 18P
M.6.2 : Econometrics	45L + 24P
M.6.3 : Game Theory	45L + 15T
M.6.4 : Investments under Uncertainty	45L + 15T
M.6.5 : Optimization in Economics	45L + 15T
M.6.6 : Actuarial Statistics	45L + 15T
M.6.7 : Official Statistics	45L + 15T

The University of Boards of Studies in Statistics may add other appropriate courses to the list of module courses.

All the elective courses, except.

M.5.3 : Computer Programming has 35 theory periods over the 15 week semester. This course requires more practicals and less theory instruction as specified. In general

the practical and the tutorial hours are not uniform. The differences reflect the thrust and the nature of the course.

Operative Recommendations

The Committee is acutely aware that a large proportion of Post-graduate teaching Centres in India lack the basic facility of a computer based laboratory. As a one time measure the Committee strongly suggests that the UGC make a grant of Rs. 4.0 lacs to any such Centre which is adopting this syllabus. At current low prices this grant will enable the Centres to set up a statistical laboratory with 10 PC's (at the Pentium III/Celeron level with networking and some softwares). In such a laboratory it will be possible to train 20/ 25 Part I students over 3 days of the week and a similar number of Part II students over the other 3 days of the week. A proper modus operendi for awarding this grant should be established by the Commission. In fact it will be desirable to make such grants to UG colleges also. However, the numbers will be large. It is therefore recommended that an adequate computer laboratory be set up jointly with the Computer Science Departments in those colleges which provide instruction in Computer Science also.

There is also the question of running expenses for the laboratory. This can be solved by charging an extra laboratory fee upto Rs. 2000 per semester from the students by the concerned institutions. The Committee is convinced that unless some such unusual measures are taken, the recommendations of the Committee will not be put into practice. There is a great paucity of teachers at both UG and PG levels with adequate exposure to computers. The Committee strongly recommends that the following two measures be taken to meet this challenge : (i) Retrain the teachers through the medium of refresher courses, workshops, etc. The UGC should make special provision for such Refresher Courses at a place where adequate trained personnel and infrastructure is available, such as the Indian Statistical Institute, the University of Pune, etc. (ii) Wherever available adjunct teachers should be appointed from the Industry, other Institutions, etc. The payment to these adjunct teachers may be made out of the extra fees generated as suggested above.

DETAILED SYLLABI

B.A/B.Sc. (General/Ancillary) Course in Statistics

Paper Number	Title	Full or Half Year	Contact Hours per week	Marks
FIRST YEAR				
101	Probability Theory	Full	3	100
102	Descriptive Statistics	Full	3	100
103	Practical	Full	3	100
SECOND YEAR				
201	Statistical Methods	Full	3	100
202	A : Sample Surveys	Half	3	50
	B : Analysis and Design of Experiments	Half	3	50
203	Practical	Full	3	100
THIRD YEAR				
301	Applied Statistics	Full	3	100
302	A : Statistical Quality Control	Half	3	50
	B : Computational Techniques	Half	3	50
303	Practical	Full	3	100

FIRST YEAR

101 : PROBABILITY THEORY

Important Concepts in Probability: Definition of probability - classical and relative frequency approach to probability, Richard Von Mises, Cramer and Kolmogorov's approaches to probability, merits and demerits of these approaches only general ideas to be given). (5L)

Random Experiment: Trial, sample point and sample space, definition of an event, operation of events, mutually exclusive and exhaustive events. Discrete sample space, properties of probability based on axiomatic approach, conditional probability, independence of events, Bayes' theorem and its applications. (15L)

Random Variables: Definition of discrete random variables, probability mass function, idea of continuous random variable, probability density function, illustrations of random variables and its properties, expectation of a random variable and its properties -moments, measures of location, dispersion, skewness and kurtosis, probability generating function (if it exists), their properties and uses. (20L)

Standard univariate discrete distributions and their properties: Discrete Uniform, Binomial, Poisson, Hypergeometric, and Negative Binomial distributions. (20L)

Continuous univariate distributions- uniform, normal, Cauchy, Laplace, Exponential, Chi-Square, Gamma and Bea distributions. Bivariate normal distribution (including marginal and conditional distributions). (25L)

Chebyshev's inequality and applications, statements and applications of weak law of large numbers and central limit theorems. (5L)

(90L)

REFERENCES

Bhat B.R, Srivenkatramana T and Rao Madhava K.S. (1997):Statistics: A Beginner's Text, Vol. II, New Age International (P) Ltd.

Edward P.J., Ford J.S.and Lin (1974): Probability for Statistical Decision-Making, Prentice Hall.

Goon A.M., Gupta M.K., Das Gupta.B. (1999): Fundamentals of Statistics, Vol.II, World Press, Calcutta.

Mood A.M, Graybill F.A and Boes D.C. (1974): Introduction to the Theory of Statistics, McGraw Hill.

ADDITIONAL REFERENCES

Cooke, Cramer and Clarke (): Basic Statistical Computing, Chapman and Hall.

David S (1996): Elementary Probability, Oxford Press.

Hoel P.G (1971): Introduction to Mathematical Statistics, Asia Publishing House.

Meyer P.L (1970): Introductory Probability and Statistical applications. Addision Wesley.

102 : DESCRIPTIVE STATISTICS

Types of Data: Concepts of a statistical population and sample from a population; qualitative and quantitative data; nominal and ordinal data; cross sectional and time series data; discrete and continuous data; frequency and non- frequency data. Different types of scales - nominal, ordinal, ratio and interval. (5L)

Collection and Scrutiny of Data: Primary data - designing a questionnaire and a schedule; checking their consistency. Secondary data - its major sources including some government publications. Complete enumeration, controlled experiments, observational studies and sample surveys. Scrutiny of data for internal consistency and detection of errors of recording. Ideas of cross-validation. (5L)

Presentation of Data: Construction of tables with one or more factors of classification. Diagrammatic and graphical representation of grouped data. Frequency distributions, cumulative frequency distributions and their graphical representation, histogram, frequency polygon and ogives. Stem and leaf chart. Box plot. (15L)

Analysis of Quantitative Data: Univariate data-Concepts of central tendency or location, dispersion and relative dispersion, skewness and kurtosis, and their measures including those based on quantiles and moments. Sheppard's corrections for moments for grouped data (without derivation). (20L)

Bivariate Data: Scatter diagram. Product moment correlation coefficient and its properties. Coefficient of determination. Correlation ratio. Concepts of error in regression. Principle of least squares. Fitting of linear regression and related results. Fitting of curves reducible to polynomials by transformation. Rank correlation — Spearman's and Kendall's measures. (20L)

Multivariate data: Multiple regression, multiple correlation and partial correlation in three variables. Their measures and related results. (15L)

Analysis of Categorical Data: Consistency of categorical data. Independence and association of attributes. Various measures of association for two -waand three-way classified data. Odds ratio. (10L)

(90L)

REFERENCES

Bhat B.R, Srivenkatramana T and Rao Madhava K.S.(1996): Statistics: A Beginner's Text, Vol. I, New Age International(P) Ltd.

Croxton F.E, Cowden D.J and Kelin S (1973): Applied General Statistics, Prentice Hall of India.

Goon A.M., Gupta M.K., Das Gupta.B. (1991): Fundamentals of Statistics, Vol.I, World Press, Calcutta.

ADDITIONAL REFERENCES

Anderson T.W and Sclove S.L (1978) An Introduction to the Statistical Analysis of Data, Houghton Miffin\Co.

Cooke, Cramer and Clarke (): Basic Statistical Computing, Chapman and Hall.

Mood A.M, Graybill F.A and Boes D.C. (1974): Introduction to the Theory of Statistics, McGraw Hill.

Snedecor G.W and Cochran W. G. (1967): Statistical Methods. Iowa State University Press.

Spiegel, M. R. (1967): Theory & Problems of Statistics, Schaum's Publishing Series.

103 : PRACTICAL

1. Presentation of data by Frequency tables, diagrams and graphs.
2. Calculation of Measures of central tendecy, dispersion, skewness and Kurtosis.
3. Product Moment Correlation and Correlation ratio.
4. Fitting of Curves by the least square method.
5. Regression of two variables.
6. Spearman's Rank correlation and Kendall's tau.
7. Multiple regression of three variables.
8. Multiple correlation and Partial correlation.
9. Evaluation of Probabilities using Addition and Multiplication theorems, conditional probabilities, and Baye's theorems.
10. Exercises on mathematical expectations and finding measures of central tendency, dispersion, skewness and kurtosis of univariate probability distributions.
11. Fitting of standard univariate and continuous distributions. (30 Practicals)

SECOND YEAR

201 : STATISTICAL METHODS

Sampling from a distribution: Definition of a random sample, simulating random sample from standard distributions, concept of a derived distributions of a function of random variables. Concept of a statistic and its sampling distribution, Point estimate of a parameter, Concept of bias and standard error of an estimate. Standard errors of sample mean, sample proportion. Sampling distribution of sum of binomial, Poisson and mean of normal distributions. Independence of sample mean and variance in random sampling from a normal distribution(without derivation). (20L)

Statistical Tests and Interval Estimation: Null and alternative hypotheses, Types of errors, p- values, Statement of chi-square, t, and F statistics. Testing for the mean and variance of univariate normal distribution, testing of quality of two means and testing of equality of two variances of two univariate normal distributions. Related confidence intervals. Testing for the significance of sample correlation coefficient in sampling from bivariate normal distribution and for the equality of means and equality of variances in sampling from bivariate normal distributions. (30L)

Large sample tests: use of central limit theorem for testing and interval estimation of a single mean and a single proportion and difference of two means and two proportions, Fisher's Z transformation and its uses. Pearson's chi-square test for goodness of fit and for homogeneity for standard distributions. Contingency table and test of independence in a contingency table. (20L)

Nonparametric tests: Definition of order statistics and their distributions, Non-parametric tests, Sign test for univariate and bivariate distributions, Wilcoxon-Mann- Whitney test, Run test, median test, and Spearman's rank correlation test. (20L)

(90L)

REFERENCES

Freund J.E (2001): Mathematical Statistics, Prentice Hall of India.

Goon A.M., Gupta M.K., Das Gupta.B. (1991): Fundamentals of Statistics, Vol.I, World Press, Calcutta.

Hodges J.L and Lehman E.L (1964): Basic Concepts of Probability and Statistics, Holden Day.

Mood A.M, Graybill F.A and Boes D.C. (1974): Introduction to the Theory of Statistics, McGraw Hill.

ADDITIONAL REFERENCES

Bhat B.R. Srivenkatramana T and Rao Madhava K.S. (1997): Statistics: A Beginner's Text, Vol. II, New Age International (P) Ltd.

Rohatgi V.K (1967): An Introduction to Probability Theory and Mathematical Statistics, John Wiley & Sons.

Snedecor G.W and Cochran W. G. (1967): Statistical Methods. Iowa State University Press.

202 A- SAMPLE SURVEYS

Sample Surveys, Concepts of population and sample, need for sampling, Census and sample survey, basic concepts in sampling, organizational aspects of survey sampling, sample selection and sample size. (10L)

Some basic sampling methods- simple random sampling (SRS) with and with out replacement. (8L)

Stratified random sampling, Systematic sampling, ratio and regression methods of estimation under SRS. (15L)

Non sampling errors, acquaintance with the working (questionnaires, sampling design, methods followed in field investigation, principal findings etc.) of NSSO, and other agencies undertaking sample surveys. (12L)

B- ANALYSIS AND DESIGN OF EXPERIMENTS

Analysis of variance for one way and two-way classifications. (10L)

Need for design of experiments, fundamental principles of design, basic designs - CRD, RBD, LSD and their analysis. (20L)

Factorial designs - 2 n designs, illustrations, main effects and interaction effects and confounding in 23 design. (15L)

REFERENCES

Cochran W.G and Cox G.M (1957): Experimental Designs, John Wiley and Sons.

Das M.N and Giri (1986): Design and Analysis of Experiments, Springer Verlag

Murthy M.N(1967): Sampling Theory and Methods, Statistical Publishing Society, Calcutta.

Sampath S. (2000): Sampling Theory and Methods, Narosa Publishing House.

Sukhatme B.V(1984) : Sample Survey methods and Its Applications, Indian Society of Agricultural Statistics.

ADDITIONAL REFERENCES

Des Raj (2000) : Sample Survey Theory, Narosa Publishing House.

Goon A.M.,Gupta M.K.,Das Gupta.B. (1986): Fundamentals of Statistics, Vol.II, World Press, Calcutta.

Kempthorne O. (1965): The Design and Analysis of Experiments, Wiley Eastern.

203 : PRACTICAL

1. Drawing random samples from standard univariate discrete and continuous distributions such as binomial, Poisson, Normal, Cauchy and exponential.

2. Tests of significance based on t, chi-square, F.Testing of significance of sample correlation coefficient, use of Z transformation. Testing of equality of means and equality of variances in sampling from bivariate normal.

3. Large sample tests for means and proportions, tests of goodness of fit and independence of attributes in contingency tables.

4. Nonparametric tests- Sign, Run, Median and Wilcoxon- Mann- Whitney tests. Selection of sample and determination of sample size, simple random sampling, Stratified SRS, and systematic sampling, Allocation problems in stratified SRS, Ratio and Regression methods of estimation in SRS.

5. Analysis of variance for one-way and two-way classifications. Analysis of CRD, RBD, and LSD. Analysis of 22 and 23 factorial designs.

(30 Practicals)

THIRD YEAR

301 : APPLIED STATISTICS

Indian Applied Statistical System: Present official statistical system in India, Methods of collection of official statistics, their reliability and limitations, and the principal publications containing such statistics on the topics- population, agriculture, industry, trade, price, labour and employment, transport and communications, banking and finance. (15L)

Demographic Methods : Sources of demographic data - census, register, adhoc survey, hospital records, demographic profiles of Indian census. Measurement of mortality and life tables- crude, death rates, infant mortality rates, death date by cause, standardized death rate, complete life table - its main features, mortality rate and probability of dying, use of survival tables. Measurement of fertility - crude birth rate, general fertility rate, total fertility rate, gross reproduction rate, net reproduction rate. (25L)

Economic Statistics : Index number- its definition, applications of index numbers, price relatives and quantity or volume relatives, link and chain relatives, problems involved in computation of index numbers, use of averages, simple aggregative and weighted average methods, Laspeyre's, Paasche's and Fisher's index numbers, time and factor reversal tests of index numbers. Consumer Price Index. (20L)

Static laws of demand and supply, price elasticity of demand, analysis of income and allied size distribution - Pareto distribution, graphical test, fitting of Pareto's law, log normal distribution and its properties, Lorenz curve and estimation of elasticity from time series data. Gini's coefficient. (15L)

Time Series Analysis : Economic time series, its different components, Illustrations, additive and multiplicative models, determination of trend, growth curves, analysis of seasonal fluctuations, construction of seasonal indices. (15L)

(90L)

REFERENCES

Croxton F.E and Cowden D.J. (1969) : Applied General Statistics, Prentice Hall of India.

Goon A.M.,Gupta M.K.,Das Gupta.B. (1986): Fundamentals of Statistics, Vol.II, World Press,Calcutta.

Guide to Current Indian Official Statistics: Central Statistical Organization, Govt. of India, New Delhi.

Saluja M. P ( ) Indian Official statistical Systems, Statistical Publishing Society, Calcutta.

Srivatava O.S (1983): A Textbook of Demography, Vikas Publishing.

ADDITIONAL REFERENCES

Gupta and Mukhopadhyay P.P ( ) Applied Statistics, Central Book Agency.

Pressat R (1978) : STATISTICAL Demography, Methuen and Co. Ltd.

302 :A - STATISTICAL QUALITY CONTROL

Importance of statistical methods in industrial research and practice, specification of items and lot qualities corresponding to visual gauging, count and measurements, types of inspection, determination of tolerance limits. General theory of control charts, causes of variation in quality, control limits, sub-grouping, summary of out-of control criteria, charts for attributes, np chart, p-chart, c- chart, u-chart, Charts for variables- X- and R charts, design of X and R charts versus p-charts, process capability studies. (30L)

Principle of acceptance sampling- problem of lot acceptance, stipulation of good and bad lots, producer's and consumers risks, single and double sampling plans, their OC functions, concepts of AQL, LTPD, AOQL, average amount of inspection and ASN function, rectifying inspection plans, Sampling inspection plans, Indian Standards Tables Part I (including applications), IS 2500 Part I. (15L)

B - COMPUTATIONAL TECHNIQUES

Computational techniques: Difference tables and methods of interpolation, Newton's and Lagrange's methods of interpolation, Divided differences, numerical differentiation and integration, Trapezoidal rule, Simpson's one-third formula, iterative solution of non-linear equations. (15L)

Linear Programming : Elementary theory of convex sets, definition of general linear programming problems(LPP),formulation problems of LPP, examples of LPP, Problems occurring in various fields, graphical and Simplex method of solving an LPP, artificial variables, duality of LPP.Transportation Problem (non-degenerate and balanced cases only), Assignment Problem. (30L)

(90L)

REFERENCES

Brownlee K.A(1960): Statistical Theory and Methodology in Science and Engineering, John Wiley and Sons.

Grant E.L (1964):Statistical Quality Control, McGraw Hill.

Duncan A.J (1974): Quality Control and Industrial Statistics, Taraporewala and Sons.

Gass S.I (1975): Linear Programming Methods and Applications, McGraw Hill.

Rajaraman, V (1981).Computer Oriented Numerical Methods, Prentice Hall.

Sastry S.S(1987): Introductory Methods of Numerical Analysis, Prentice Hall.

Taha H.A (1989): Operations Research: An Introduction, Macmillan Publishing Company.

ADDITIONAL REFERENCES

Bowker H.A and Liberman G.T (1962): Engineering Statistics, Prentice Hall.

Cowden D.J (1960): Statistical Methods in Quality Control, Asia Publishing Society.

Garvin W.W (1960): Introduction to Linear Programming, McGraw Hill.

Mahajan M (2001): Statistical Quality Control, Dhanpat Rai & Co.(P) Ltd.

Rao S.S (1984) Optimization Theory and Applications, Wiley Eastern.

Krishnamurthy E.V and Sen S.K (1976): Computer Based Numerical Algorithms, Affiliated East-West Press.

303 : PRACTICAL

1. Computing measures of mortality & fertility, Construction of life tables and examples involving use of life tables, Graduation of mortality rates by Gompertz curve, fitting of a logistic curve.
2. Construction of Index Numbers by Laspeyre's, Paasche's, Fisher's method.
3. Determination of trend in a time series,construction of seasonal indices.
4. Fitting of Pareto curve to income data, Lorenz curve of concentration, estimation of price elasticity of demand form time series data.
5. Drawing of X - R, np, p and c- charts. Drawing of OC curve for single and double sampling plans for attributes, AOQ and ATI curves.
6. Construction of difference tables, use of Newton's Lagrange's methods of interpolation and divided difference formulae, numerical evaluation of integrals using Trapezoidal and Simpson's one-third formulae, solution of non-linear equation by Newton-Raphson iterative method.
7. Formulation of LPP's and their duals. Solving LPPs by graphical and simplex methods, transportation and assignment problems.

(30 Practicals)

B.A/B.Sc. (Honours) Course in Statistics

Paper Number	Title	Full or Half Year	Contact Hours per week	Marks
FIRST YEAR				
101	A : Probability I	Half	3	50
	B : Mathematical Methods II	Half	3	50
102	Descriptive Statistics	Full	3	100
103	Practical	Full	3	100
SECOND YEAR				
201	A : Probability II	Half	3	50
	B : Mathematical Methods II	Half	3	50
202	Statistical Inference	Full	3	100
203	Practical	Full	3	100
THIRD YEAR				
301	Applied Statistics	Full	3	100
302	Design of Experiments, Linear Models and Sample Surveys	Full	3	100
303	Practical	Full	3	100
304	Elective Paper	Full	3	100
305	Practical (Elective)	Full	1½	50
306	Project	Full	1½	150

LIST OF ELECTIVES

1. Elementary Stochastic Processes
2. Statistics for Environment
3. Industrial Statistics and Operational Management
4. Computation Applications
5. Applied Decision Making

FIRST YEAR

101 A : PROBABILITY

Random experiment: trial, sample point and sample space, event, Operations of events, concepts of mutually exclusive and exhaustive events. (5L)

Definition of probability : classical and relative frequency approach. Discrete probability space, Properties of probability, Independence of events, Conditional probability, total and compound probability rules, Bayes' theorem and its applications. (10L)

Discrete random variable (rv); its probability mass function (pmf) and cumulative distribution function (cdf). Joint pmf of several discrete rvs. Marginal and conditional pmfs. Independence of rvs. Expectation of a rv and its properties. Moments, measures of location and dispersion of a rv. Probability generating function (pgf) and moment generating function (mgf) of a rv, their properties and uses. (15L)

Standard univariate / bivariate discrete distributions: degenerate, discrete uniform, binomial, hypergeometric, Poisson, geometric and negative binomial distributions. Marginal and conditional distributions, Distributions of functions of discrete rvs, reproductive property of standard distributions. (15L)

(45L)

REFERENCES

Chung, K. L. (1979). Elementary Probability Theory with Stochastic Processes, Springer International Student Edition.

David Stirzaker (1994). Elementary Probability, Cambridge University Press.

Feller, W. (1968). An Introduction to Probability Theory and Its Applications, Wiley.

Mukhopadhyay, P. (1996). Mathematical Statistics, New Central Book Agency, Calcutta.

Parzen, E. (1960). Modern Probability Theory and Its Applications, Wiley Eastern.

Pitman, Jim (1993). Probability, Narosa Publishing House.

101.B: MATHEMATICAL METHODS -1

Vector space with real field; linear combination of vectors; subspaces; Examples in R with geometrical interpretation; Linear dependence and independence of vectors; Dimension and basis of a vector space. (10L)

Matrices; types of matrices; operation on matrices; partitioned matrices; determinants; singular and nonsingular matrices. (8L)

Rank of a matrix; row-rank and column- rank, properties of rank, rank of sum and

product of matrices; Inverse of a matrix; Orthogonal matrix; Idempotent matrix. (7L)

Linear equations, homogeneous and non homogeneous system of equations; solution space; consistency and general solution. (10L)

Sequences and series of real numbers: convergence, Cauchy criterion and simple tests for convergence. (10L)

(45L)

REFERENCES

Apostol,T.M. (1985): Mathematical Analysis, Narosa Publishing House.

Burkill, J. C. (1980): A First Course in Mathematical Analysis, Vikas Publishing House.

Deshpande, J. V. (1981): Text Book of Mathematical Analysis, Tata McGraw Hill.

Goldberg, R. R. (1970): Methods of Real Analysis, Oxford and IBH

Khuri, A. I. (1983): Advanced Calculus with Applications in Statistics, Wiley.

Searle, S. R. (1982): Matrix Algebra Useful for Statistics, Wiley,

Shanti Narayan, (1998): Matrix Algebra, S. Chand & Co.

102 : DESCRIPTIVE STATISTICS

Types of Data : Concepts of a statistical population and sample from a population; qualitative and quantitative data; nominal and ordinal data; cross sectional and time series data; discrete and continuous data; frequency and non-frequency data. Different types of scale - nominal, ordinal, ratio and interval. (5L)

Collection and Scrutiny of Data: Primary data - designing a questionnaire and a schedule; checking their consistency. Secondary data - their major ources including some government publications. Complete enumeration, controlled experiments, observational studies and sample surveys. Scrutiny of data for internal consistency and detection of errors of recording. Ideas of cross-validation. (5L)

Presentation of Data : Construction of tables with one or more factors of classification. Diagrammatic and graphical representation of non-frequency data. Frequency distributions, cumulative frequency distributions and their graphical and diagrammatic representation - column digram, histogram, frequency polygon and ogives. Stem and leaf chart. Box plot. (15L)

Analysis of Quantitative Data : Univariate data: Concepts of central tendency or location, dispersion and relative dispersion, skewness and kurtosis, and their measures including those based on quantiles and moments. Sheppard‘s corrections for moments for

grouped data (without derivation). Measures of inequality - Gini's coefficient and Lorenz Curve. (20L)

Bivariate data: Scatter diagram. Product moment correlation coefficient and its properties. Coefficient of determination. Correlation ratio. Concepts of regression. Principle of least squares. Fitting of linear and quadratic regression and related results. Correlation index. Fitting of curves reducible to polynominals by log and inverse transformation. Fitting of curves by the method of group averages. Intra-class correlation coefficient with equal and unequal group sizes. Rank correlation- Spearman's and Kendall's measures. (15L)

Multivariate data: Multiple regression, multiple correlation and partial correlation in 3 variables. Their measures and related results. (12L)

Analysis of Categorical Data : Consistency of categorical data. Independence and association of attributes. Various measures of association for two-way and three-way classified data. Odds ratio. (8L)

Scaling of Data : Motivation for scaling. Measurement for psychological traits. Scaling of items according to difficulty. Scaling of test scores. Scaling of rates and ranks. Scaling of judgments. (10L)

(90L)

REFERENCES

Goon A.M., Gupta M.K., Dasgupta. B. Fundamentals of Statistics, Vol 1(1991) & Vol. 2(2001). World Press, Calcutta.

Croxton FE, Cowden DJ and Klein S: Applied General Statistics(1973). Prentice Hall of India.

Snedecor GW and Cochran WG: Statistical Methods(1967). Iowa State University Press.

Spiegel, MR: Theory & Problems of Statistics(1967), Schaum's Publishing Series.

ADDITIONAL REFERENCES

Agreshti, A: An Introduction to Categorical Data Analysis(1996), John Wiley & Sons Inc, NY.

Samprit Chatterjee & Bertram Price: Regression analysis by Example (1991). John Wiley & Sons, Inc.

Guilford, J. P. & Fruchter B: Fundamental Statistics in Psychology and Education (1980). McGraw Hill.

103 : COMPUTER APPLICATIONS AND PRACTICAL I

Historical evolution of computers, Generations of Computers, Classification of Computers, Hardware : CPU, I/O Devices, Block diagram. System Software. MS-DOS : Filenames, Creating, Editing and Printing of Files. Other File Management Commands, Disk-Management Commands.

Windows : The user Interface, The Desk Top, The Task Bar, The Control Panel, The Find Features, Properties, Font Management, Systems Tools, Character Map, Note Pad, The My Computer ICON, Folders, Short-Cuts.

Word Processing : Creating and Saving a document, Editing the text; Printing Documents.

Working with Software Packages : MS-Excel, Minitab, SPSS or SYSTAT.

1. Classification, Tabulation and Frequency Tables.
2. Bar Graphs, DOT Diagram and Histogram
3. Stem-and-Leaf Plots.
4. Box Plots.
5. Summary Statistics
6. Two-way tables and plots.
7. Product moment correlation coefficient, Rank Correlation coefficient.
8. Curve setting by method of least squares : Exponential and polynomial (upto three degree) equations.
9. Regression lines
10. Correlation ratios.
11. Multiple and partial correlation coefficients. Regression equations (for three variables only).
12. Gram-Schmidt orthonoralization process.
13. Rank and Inverse of a matrix
14. Solution of set of linear equations.
15. Fitting of Binomial, Poisson, Negative Binomial, Normal and Gamma Distributions.

Programming with FORTRAN (A) :

Introduction to FORTRAN 77 or 90, FORTRAN Character codes, constants, variables, names; arithmetic, logical and relation operators, Expressions : Arithmetic, Character, Relation and Logical expressions. Rules for writing arithmetic expressions.

Statements : Specification Data, Parameter, Format, one dimensional Array statements, Assignment Statements, Arithmetic Statement, Control Statements : DO loops, Do-Continue,

IF-Blocks, Unconditional GOTO, IF (condition) GOTO label, and Computed GOTO statements, STOP, RETURN and END statements.

I/O Statements : Format free Read, Write and Print Statements, Simple Formats, OPEN and CLOSE statements.

Built-in-functions : Names of Generic and non-generic functions. Flow-chart, Body of the program, compiling and running of the programs.

FORTRAN Programmes for :

1. Factorial of a positive Integer
2. Ordering of a given set of observations
3. Finding maximum and minimum of a given set of observations.
4. Mean, variance and quantiles for ungrouped and grouped data.
5. Correlation coefficients for ungrouped data, Intra-class correlation coefficient.
6. Fitting of exponential curve and straight line to the given data.
7. Fitting of Binomial and Poisson distributions.
8. Calculating the correlation coefficient for grouped data.

(90 hours)

REFERENCES

Norton, Peter : Guide to MS-DOS.

Mathur, Rajiv: Learning Windows -98 Step by step. Galgotia

Mathur, Rajiv: Learning Excel-97 for Windows step by step. Galgotia

Sanders, H.D.: Computer Today. Mc Graw Hill.

Rajaraman : Computer Programming in FORTRAN-77, Prentice Hall.

Page, Clive G.: The professional programmers guide to FORTRAN-77. Pitman publishing House, London.

Reddy, R.N. and Ziegler, C.A. :FORTRAN-77-With applications for scientists and engineers. JAICO Publishing House, Bombay, Calcutta & Madras.

SECOND YEAR

201 A : PROBABILITY II

General probability space : Sigma algebra of events; axiomatic definition of probability; random variable (rv), cumulative distribution function (cdf) of a rv and its propeties; continuous rv and its probability density function (pdf); Expectation, moments and quantiles. (12L)

Univariate continuous distributions and their properties: uniform, beta, gamma, exponential Pareto, Weibull, Laplace, normal, Cauchy, logistic, lognormal; Truncated distributions, Moment generating function (mgf): its properties and applications. (12L)

Cdf of rv's; independence of rv's, sampling distribution of sample mean for normal and sample total for binomial, Poisson, negative binomial, normal and Gamma distributions. Chisquare, t and F distributions as distributions of functions of standard normal rv's. (8L)

Continuous bivariate distribution and its pdf; marginal and conditional distributions. Expectation and conditional expectation, covariance, correlation and regression. Bivariate normal distribution and its properties. Distribution of functions of bivariate rv's. Convolution and compound distribution. Distributions of order statistics. (8L)

Tchebycheff's inequality, convergence in distribution and in probability, convergence of binomial to Poisson and of binomial to normal (Demoivre - Laplace limit theorem), statement and applications of WLLN and CLT. (5L)

(45L)

REFERENCES

Chung, K. L. (1979). Elementary Probability Theory with Stochastic Processes, Springer International Student Edition.

David Stirzaker (1994). Elementary Probability, Cambridge University Press.

DeGroot, M. H. (1975). Probability and Statistics, Addison - Wesley.

Feller, W. (1968). An Introduction to Probability Theory and Its Applications, Wiley.

Hoel, P. G., Port, S. C. and Stone, C. T. (1991). Introduction to Probability Theory, Universal Book Stall.

Hogg, R. V. and Craig, A. T. (1971). Introduction to Mathematical Statistics, McMillan.

Parzen, E. (1960). Modern Probability Theory and Its Applications, Wiley Eastern.

Pitman, Jim (1993). Probability, Narosa Publishing House.

201 B : MATHEMATICAL METHODS - 2

Difference table. Methods of interpolation: Newton's forward and backward interpolation formulae. Newton's divided difference interpolation formula. Numerical differentiation and integration. Trapezoidal, Simpson's one - third and three - eighth formulae for numerical integration. (10L)

Function of one variable: limit, continuity and differentiability. Mean value theorems; Maxima and minima. Power series, uniform convergence. (10L)

Riemann integral; integration by parts, change of variables; improper integrals, beta and gamma integrals. (10L)

Function of several variables: partial derivatives; maxima and minima, constrained maxima and minima; applications of Lagrangian multipliers; multiple integrals, transformations and Jacobians. (15L)

(45L)

REFERENCES

Apostol, T. M. (1985). Mathematical Analysis, Narosa Publishing House.

Burkill, J. C. (1980). A First Course in Mathematical Analysis, Vikas publishing House.

Deshpande, J. V. (1981). Text Book of Mathematical Analysis, Tata McGraw Hill.

Goldberg, R. R. (1970). Methods of Real Analysis, Oxford and IBH.

Khuri, A. I. (1983). Advanced Calculus with Applications in Statistics, Wiley.

Rudin, W. (1976). Principles of Mathematical Analysis, 3/e, McGraw Hill.

Shanti Narayan (1993). Mathematical Analysis, S. Chand and Co.

Somasundaram, D. and Choudhary, B. (1996). A First Course in Mathematical Analysis, Narosa Publishing House.

202 : STATISTICAL INFERENCE

Introduction : Parametric models, parameters; random sample and its likelihood; statistic and its sampling distribution; problems of inference. (2L)

Data Reduction: Sufficiency, Factorization Theorem (proof of discrete case only), Illustrations, Concept of Minimal Sufficiency. (3L)

Point Estimation : Properties of estimators: Mean square error (MSE) and minimum MSE estimator,: Unbiasedness and Minimum Variance Unbiased Estimator (MVUE) Rao-Cramer Lower Bound of Variance and related results. Relative efficiency of an estimator.

Amount of information. Consistency of estimators and sufficient conditions for consistency. Asymptotic efficiency. Sufficient Statistic and optimal estimators. Concept of complete sufficient statistic with illustration. Rao-Blackwell Theorem. (20L)

Methods of Estimation : Method of Moments, Method of maximum likelihood. Statement of properties of MLE, Method of Minimum chi-square. (8L)

Testing of Hypotheses : Statistical Hypotheses- Simple and Composite. Statistical tests, Critical region, Errors of Type I and Type II, Size and Power of a test. Definition of Most Powerful (MP),Uniformly Most Powerful (UMP), Unbiased and Uniformly Most Powerful Unbiased (UMPU) tests. Neyman Pearson Lemma and its application in testing hypotheses regarding binomial, Poisson, normal and exponential distributions. Power functions of UMP tests with simple illustrations. Likelihood Ratio tests. Tests for parameters of binomial, Poisson, and univariate and two independent normal populations. Tests for equality of means, significance of correlation coefficient, and equality of variances of a bivariate normal population. (20L)

Interval Estimation : Concepts of Confidence Interval and Confidence Coefficient. Confidence Intervals for the Parameters of univariate normal, two independent normal and one-parameter exponential distribution. (5L)

Sequential Tests : Need for Sequential tests, Wald's SPRT with illustrations. Approximate OC and ASN functions for tests regarding parameters of binomial and normal distributions. (7L)

Large Sample Tests : Use of CLT for deriving large sample tests for binomial proportion, difference of two binomial proportions, mean of a population and difference of means of two independent populations. Related confidence intervals. Variance-stabilizing transformations (arcsine, square-root and z) and their uses in large sample tests and interval estimation. Pearsonian Chi-square tests for independence of attributes, homogeneity of populations and goodness of fit. (20L)

Nonparametric Tests : Need for non-parametric tests. Sign test for location of univariate and bivariate populations; Wilcoxon-Mann-Whitney test; Run Test, Median test and Test based on Spearman's rank correlation. (10L)

(90L)

REFERENCES

Bhattacharya G.K and Johnson R.A.(1977): Statistical Concepts and Methods. John Wiley and Sons.

Dudewicz E.J and Mishra S.N.(1988): Modern Mathematical Statistics, International Student Edition, John Wiley and Sons.

Freund J.E.(2000): Mathematical Statistics, Prentice - Hall of India.

Goon A.M., Gupta M.K. and Dasgupta B.(1980): An outline of Statistical Theory, Vol.2, The World Press Publishers Private Limited, Calcutta.

Hogg R.V and Craig A.T.(1978): Introduction to Mathematical Statistics, Fourth edition, Collier Macmillan Publishers.

Mood A.M., Graybill F.A and Boes D.C.(1974): Introduction to the Theory of Statistics, Third Edition, McGraw Hill.

ADDITIONAL REFERENCES

Rao C.R.(1973): Linear Statistical Inference and Its Applications, Revised Edition, Wiley Eastern.

Rohatgi V.K.(1986): An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern.

203 : COMPUTER APPLICATIONS AND PRACTICAL II

Programming with FORTRAN (B) - Towards Better Programms : REAL, INTEGER, DOUBLE PRECISION, COMPLEX and LOGICAL variables, COMMON Statement, COMMON BLOCKS, BLOCK DATA, EQUIVALENCE, Two Dimensional and Multidimensional Arrays.

Multi-branch Transfer of control : Logical and Arithmetic IF, Computed GOTO, Assigned GOTO, Nested and Implied DO Loops, Counting loops, Simple and nested Blocks:

IF-THEN

IF-THEN-ENDINF

IF-THEN-ELSE-ENDIF

IF-THEN-ELSEIF-THEN-ENDIF

DO WHILE.

Format refinements.

Design of Modular Programs : User defined functions, Function-Sub-programs, Subroutine-Subprograms. Programs using sub-programs : Square-root subroutine, Matrix inversion, Sub-programs for computing mean variance and covariance for n pairs of observations, Main programs for computing correlation-coefficient and regression lines using sub-programs.

Application of packages (SPSS, Minitab or SYSTAT) and FORTRAN Programs for solving the following problems :

1. Drawing of random samples from Binomial, Poisson, Normal and Gamma distributions.

2. Fitting of standard distributions and tests for goodness of fit.

3. Method of Estimation : Moments, MLE, MLE for location parameter of Cauchy distribution.

4. Tests of significance : Drawing of power curve of a test. Test based on Chi-square, t and F statistics and related confidence intervals. Large sample tests.

5. Non-Parametric Tests : Sign Test, Wilcoxon-Signed rank test, Kolmogorov's test. Kolmogovorov Smirnov's test, Wilcoxon-Man-whitney test, Median Test, Mood's Test.

6. Construction of difference table, Forward, Backward and Central difference interpolation formulae. Divided difference table. Newton's divided difference interpolation formula, Lagrange's interpolation formula. Estimation of error.

7. Trapezoidal Rule, Simpson's 1/3rd rule and Weddle' formula for numerical integration. Comparision of values obtained by different formulae.

(90 Hours)

REFERENCES

Reddy, R.N. and Ziegler, C.A.: FORTRAN-77-With Applications for Scientists and engineers, JAICO Publishing House Bombay, Calcutta & Madras.

Rajaraman : Computer Programming in FORTRAN-77, Prentice Hall.

Sastry, S.S. : Introductory Methods of Numerical Analysis, PHI India.

THIRD YEAR

301 : APPLIED STATISTICS

Demographic Methods : Sources of demographic data:census, registrar, ad-hoc surveys, hospital records, demographic profiles of the Indian Census. Measurement of mortality and life table : Crude death rate, age-specific death rates, infant mortality rates,death rate by cause, complete life table and its main features, usesof life table. Measurement of fertility : Crude birth rate, general fertility rate, age-specific birth rate, total fertility rate, gross reproduction rate, net reproduction rate, standardised deathrates, age pyramid of sex composition. raduation of mortality curves : Gompertz and Makeham formulae, logistic curve and its use in the population projection. (25L)

Statistical Process And Product Control : Quality of a product, need for quality control, Basic concepts of Process control, Process capability and Product Control.

General theory of control charts, Causes of variation in quality, control limits, subgrouping, summary of out of control criteria. Charts for attributes - chart, np-chart, c-chart, u-chart. Charts for variables - X & R and X &(Sigma - charts, CUSUM charts. Process capability measurement. Process capability ratios-Cpk and Cp, Six-Sigma capability. (15L)

Principles of accptance sampling : Problem of lot acceptance, good and bad lots, producer's & consumer's risks, single, double and sequential sampling plans for attributes and their OC functions, concepts of AQL, LTPD, AOQL, average amount of inspection and ASN functions. Rectifying inspection plans, sampling plans for variables. Bureau of Indian Standards IS 25000 Table 1 and applications. (15L)

Economic Statistics : Time series analysis - Economic time series, different components, illustrations, additive and multiplicative models, determination of trend, growth curves, analysis of seasonal fluctuations, construction of seasonal indices.

Analysis of income and allied distributions - Pareto distribution, graphical test, fitting of Pareto's law, illustrations, log-normal distribution and its properties, Lorenz curve, Gini's coefficient.

Index Numbers - Price relatives and quantity or volume relatives. Link and chain relatives, computation of index numbers, Laspeyre's, Paasche's, Marshal - Edgeworth's and Fisher's index numbers, chain base index number, consumer price - index numbers. Tests for index numbers : Time and Factor reversal tests. (25L)

Demand Analysis : Theory and analysis of consumer's demand : Law of demand, Price elasticity of demand, Estimation of demand curves; Forms of demand functions, Engel's curves, Income elasticity of demand. (10L)

(90L)

REFERENCES

Mukhopadhyay, P. (1999) : Applied Statistics. New Central Book Agency Pvt. Ltd., Calcutta.

Srivastava, O.S. (1983) : A text book of demography. Vikas Publishing House, New Delhi.

Goon, A.M.; Gupta, M.K. and Dasgupta, B (1986) : Fundamentals of Statistics. Volume - Two. World Press, Calcutta.

Benjamin, B. (1959) : Health and vital statistics. Allen and Unwin.

Hald, A. (1952) - Statistical theory with engineering applications. John Wiley.

Duncan, A.J. (1974) : Quality control and industrial statistics. Fourth edition, Taraporewala & sons.

Montgomery, D. C. (1991) : Introduction to the statistical quality control. econd edition. John Wiley & Sons.

Wetherill G.B. and Brown, D.W. (1991) : Statistical Process Control. Chapman and Hall.

Wadsworth, H.H.; Stephens, K.S. and Godfree, A.B. (1986) : Modern methods for quality control and improvement. Wiley.

Brown, R.G. (1963) : Smoothing, forecasting and prediction of discrete time series. Prentice Hall.

Chatfield C. (1980) : The Analysis of Time Series - An Introduction, Second edition. Chapman and Hall.

302 : DESIGN OF EXPERIMENTS, LINEAR MODELS, SAMPLE SURVEYS AND OFFICIAL STATISTICS

Linear Models and ANOVA : Theory of linear estimation, Estimability of linear parametric function, and BLUE, Method of least squares, Gauss-Markov theorem. Estimation of error variance.

Regression Analysis : Estimation and tests of regression parameters in uni-variate (linear, quadratic and cubic) and multivariate linear regression under usual assumptions and related interval estimation. Violation of usual assumptions concerning normality, homoscedasticity and collinearity. Diagonostics using probability plots.

ANOVA : Decomposition of sum of squares in one-way and two-way orthogonal classifications. ANOVA Table. (25L)

Design of Experiment : Need for design of experiments, Fundamental principles of design of experiments, Basic designs-CRD, RBD, LSD and their analyses, Orthogonality of classification in two-way lay-outs, advantages of orthosonality relation, simple illustrations. Analysis of co-variance, missing plot technique.

Factorial experiments : 2n, 32 factorial experiments, illustrations, main effects and interactions, confounding and illustrations. (25L)

Sample Surveys : Concepts of population and sample, need for sampling, census and sample surveys, Basic concepts in sampling and designing of large scale surveys. Simple random sampling with and without replacement, Stratified random sampling, Allocation problems, Systematic sampling, Cluster sampling, Ratio, Product and Regression methods of estimation. Double sampling. Two-stage and Multi-stage sampling. Non-sampling errors. Randomized response technique (Warner's model only). (25L)

Indian official statistics : Present official statistical system in India, Methods of collection of official statistics, their reliability and limitations. Principal publications containing data on the topics such as population, agriculture, industry, trade, prices, labour and employment, transport and communications, banking and finance. Various official agencies responsible for data collection and their main functions. (15L)

(90L)

REFERENCES

Guide to current Indian Official Statistics. Central Statistical Organisation, Govt. of India, New Delhi.

Saluja, M.P. ( ) : Indian official statistical systems. Statistical Publishing Society, Calcutta.

Draper, N.R. and Smith, H. (1981) : Applied Regression Analysis. John Wiley.

Chatterjee, S. and Price, P. (1991) : Regression Analysis by example. Second edition. John Wiley & sons.

Murthy, M.N. (1967) : Sampling theory and methods. Statistical Publishing Society, Calcutta.

Sukhatme et al. (1984) : Sampling Theory of Surveys with applications. Indian Society of Agricultural Statistics.

Cochran, W.G. (1984) : Sampling Techniques Third edition. Wiley Eastern.

Des Raj and Chandok (199?) : Sampling design. Tata Mc Graw Hill.

Federer, W.T. (1975) : Experimental designs-theory and applications. Oxford & IBH.

Goon, A.M., Gupta, M.K. & Das Gupta, B. (1986) : Fundamentals of Statistics. Vol-II. World Press, Calcutta.

Montgomery, D.C. (1991) : Design and analysis of experiments. John Wiley.

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

Mukhopadhyay, Parimal (1998) : Theory and Methods of Survey Sampling. Prentice Hall.

Mukhopadhyay, Parimal (1996) : Mathematical Statistics. New Central Book Agency Pvt. Ltd., Calcutta.

Sampat, S. (2000) : Sampling Theory. Narosa.

303 : COMPUTER APPLICATIONS AND PRACTICAL III

Use of software packages (SPSS, Minitab or SYSTAT) and FORTRAN Programs for solving the following problems :

1. Multiple regression : Estimation of regression coefficients, Testing of hypotheses concerning regression coefficients, confidence intervals, Residual plots and Model disasonostics.
2. One-way and two-way ANOVA (orthogonal case only).
3. Analyses of CRD, RBD, LSD. Analysis of covariance.
4. Factorial Experiments : 2n, 32 factorial experiments and confounding.
5. SRS, Stratifical random sampling. Estimation of parameters and estimation of standard error of these estimates. Cluster sampling, Two-stage sampling, Ratio and regression method of estimation.
6. Computing measures of mortality and fertility, construction of life-tables.
7. Fitting of Gompertz, Makeham and Logistic curves. Population Projection.
8. Drawing of X-R, X-(, p, np, c and u charts. CUSUM Charts. Process capability. Single and double sampling plans. OC, AOQ and ATI curves.
9. Determination of trend in time series. Construction of seasonal indices. Fitting of Pareto curves to income data. Lorenz curve of concentration, Gini's coeffiecient.
10. Construction of index numbers by Laspeyre's, Paasche's, Fishers's and Marshal-Edqeworth's formulae. Chain Base Index Number, Consumer's Price Index Numbers.

(90 hours)

ELECTIVE 1 : ELEMENTS OF STOCHASTIC PROCESSES

Basic concepts : Definition and examples of stochastic process, classification of general stochastic processes into discrete/continuous time, discrete/ continuous state spaces, types of stochastic processes, elementary problems. (10L)

Markov chains : Definition and examples of Markov chain, Transition Probability Matrix, classification of states, recurrence, simple problems, basic limit theorem of Markov chain (statement only), stationary probability distribution, applications. (40L)

Continuous Time Markov Chain : Pure birth process and Poisson process, Birth and Death process, Problems. (20L)

Branching Process : Definition and examples of discrete time branching process, probability generating function, mean and variance, probability of extinction, problems. (20L)

(90L)

REFERENCES

Karlin, S. and Taylor, H. M. (1975). A first course in Stochastic Processes, Academic Press.

Hoel, P.mG., Port, S. C. and Stone, C. J. (1991). Introduction to Stochastic Processes, Universal Book Stall.

Parzen, E. (1962). Stochastic Processes, Holden-Day.

Cinlar, E. (1975). Introduction to Stochastic Processes, Prentice Hall.

Adke, S. R. and Manjunath, S. M. (1984). An Introduction to Finite Markov Processes, Wiley Eastern.

Medhi, J. (1996). Stochastic Processes, New Age International (P) Ltd.

Ross, S. M. (1983). Stochastic Processes, John Wiley.

Taylor, H. M. and Karlin,S. (1999). Stochastic Modelling, Academic Press.

ELECTIVE 2 : STATISTICS FOR ENVIRONMENT

Statistical Genetics.

Cells, Chromosomes, Gametes, Genes and Gene frequency, Mendel's laws. Single locus with two alleles. Hardy-Weinberg equilibrium.

A-B-O bloodgroup system. Calculation of probabilities of offspring bloodgroup for given parental blood group. Chance of incompatibility. (25L)

Definition of ED50, ED90 etc. Simple method of estimation of the above. Data: Dose levels (Z i), number of individuals exposed (n i), number responding (r i). Simple regression of probit on log dose to estimate parameters of tolerance distribuion.

Introduction to logistic regression with binary response and one independent variable (continuous). (25L)

Exponential and logistic models of population growth, solving the following differential equations :

$dNt/dt = KNt$, $dNt/dt = aNt(K-Nt)$.

Fitting the above growth models to data by linearization and regression.

Capture-recapture method of abundance estimation. One and two recapture occasions. Use of likelihood under binomial distribution.

Concept of biodiversity. Simpson's and Shannon-Wiener indices. (25L)

Study of exponential and Weibull distributions as models for survivorship data. Corresponding hazard functions and interpretation of their shapes. Applications to environmental data. (15L)

(90L)

REFERENCES

D.J.Finney (1978) : Statistical Methods in Biological Assays, Charles Griffics & Co.

A.P. Gore and S.A. Paranjpe (2000) : A Course in Mathematical & Statistical Ecology, Kluwer.

Z.Govindarajulu (2000) : Statistical Techniques in Bioassay, 2nd Edition, S.Karger.

D.W.Hosmer & S.Lemeshaw (1989) : Applied Logistic Regression Wiley.

R.C.Elandt Johnson (1975) : Probability Models & Statistical Methods in Genetics Wiley.

C.C.Li (1976) : First Course in Population Genetics, Boxwood Press.

E.C.Pielou (1977) : An Introduction to Mathmatical Ecology, John Wiley.

ELECTIVE 3 : INDUSTRIAL STATISTICS AND OPERATIONS MANAGEMENT

Part-1 : Operations Research: Linear programming problems(LPP): Canonical and standard forms, and dual of an LPP, graphical method to solve two variable LPP, solving LPP manually using Simplex procedure in presence of slack or/and surplus or/and artificial variables, solving LPP using computer package Applications of LPP: representation of

transportation and assignment problems as LPP Solution of a transportation problem using initial basic feasible solution by North West corner, matrix minimization and Vogels method, Solution of assignment problem using Hungarian method.

CPM and PERT: determination of critical path and calculation of probabilities of completing a project with in specified period. (30L)

Part-2: Inventory planning: Concept of planned inventory policies. Deterministic Models-policy when inventory levels are reviewed continuously and demands occur uniformly with and without shortage costs, Economic order quantity, policy for production planning when inventory levels are reviewed periodically, solving special cases using computer package.

Stochastic models- a single period model with no set up cost having zero or non-zero initial stock {(s, S) policy }(4), solving special cases of these stochastic models when demand distribution is specified using computer package. (20L)

Part 3: Forecasting: Concept of forecasting and its applications in manufacturing and non manufacturing industrial situations, different methods of forecasting including average, last value, weighted average (exponential smoothing), forecasting in presence of linear trends using least square methods, forecasting in presence of seasonal effects, solving special cases using computer package. (15L)

Part 4: Reliability: definitions and relationships between survival function, hazard function, hazard rate of a non negative random variable, parametric distributions - Weibull, gamma, lognormal and exponential as life time distributions, concept of aging, IFR, IFRA classes of distributions and their dual, coherent system as binary function: minimal cut and path sets(vectors), representation of structure function of series, parallel and k out of n : G systems of independent components. Using minimal cut and path structure functions, dual of a coherent structure, derivation of reliabilities of above structures. (25L)

(90L)

REFERENCES

H. A. Taha: Operations Research, Macmillan Publishing Co. Inc (Edition 6, 1999).

F. S. Hillier and G. J. Libermann: Introduction to Operations Research, McGraw Hill, Inc.(Edition 6, 1995).

L. J. Bain and M. Englhardt: Statistical analysis of reliability and life testing models, Marcel Dekker (1991).

S. Zacks: Introduction to reliability analysis Probability models and statistical methods, Springer Verlag (1992).

R. E. Barlow and F. Proschan: Statistical theory of reliability and life testing: Probability models, Holt, Rinehart and Winston,(1975).

ELECTIVE 4 : PROGRAMMING WITH C

History and features of C language. Components of C language, Structure of a C Program. Data type : Basic data types, Enumerated data types, Derived data types. Variable Declaration : Local, Global, Parametric Variables, Assignment of variables. Numeric, Character, real and string constants. Arithmetic, Relation and logical operators. Assignment operators. Increment and Decrement operators, conditional operators, Bitwise operators. Type modifiers and expressions, Writing and interpreting expressions, using expressions in statements. Basic input/output.

Control Construct - I

Control Statements, conditional statements, if...else, Nesting of if....else, elseif ladder, switch statements, Loops in C : for, while, do... while loops.

Control Constructs - II

break, Continue, exit( ), goto and label declarations. (15L)

One dimensional, two dimensional and multidimensional arrays.

Functions, classification of functions, functions definition and declaration, assessing a function, return statement.

Storage classes : Automatic variables, External variables, static variables, register variables, Scope and lifetime of declarations.

Parameter Passing in functions, recursion in Functions.

Pointers; Pointer Notation : and * operators.

Pointer declaration and Initialization, Accessing a variable through pointer, pointer expressions, pointer arithmetic, pointer comarison, pointer increment/decrement and scale factor.

Pointer and Array : Pointers and one dimensional arrays, dynamic memory allocation functions malloc and calloc, pointers and multidimensional arrays, arrays of pointers.

Pointer and functions : Pointers to pointers, pointers and functions, pointers to functions, function returning pointers, functions with variable number of arguments. (15L)

Structure : Definition and declaration, structure (initialization, comparision of structure variables, array of structures; array within structures, structures within structures. Passing structures to functions; structure pointers. Unions - Definition and declaration, Accessing a union member, union of structure, initialization of a union variable, uses of union.

Introduction to linked list, linear linked list, insertion of a node in list, removal of a node from list.

Files in C : Defining and opening a file, closing a file, input-output operation on file. Creating a file, reading a file.

Preprocessors

Introduction to preprocessors, Macro Substitution, Simple Macro Substitution, Macro with arguments, Nesting of Macros, Undefining a Macro, File inclusion, Conditional Compilation Directives.

Standard header files, Library functions. String functions, Mathematical functions, Date and Time functions, variables argument list function, utility functions, character class test functions. (20L)

(50 Theory + 40 Laboratory hours besides the practical paper associated with this theory paper)

REFERENCES

Balagurusamy, E. : Programming in ANSI C. Tata McGraw Hill.

Kanetkar, Y.P. : Working with C. BPB Publication.

Gottfried, Byron S. : Theory and problems of programming with C. TMH

Schildt, Herbert : C : The complete reference III Ed.. TMH.

Schildt, Herbert : C Made easy. Mc Graw Hill.

ELECTIVE 5. APPLIED DECISION-MAKING

The decision problem—goals and objectives, conflict between goals, possible solutions, constraints, feasible solutions, objective function, costs and benefits, notions and criteria for optimality. Steps in decision-making: determining objective(s),identifying alternative feasible solutions, determining (expected) costs and benefits associated with a feasible solution, developing a measure of effectiveness, finding the optimal solution. Sensitivity analysis and post-optimality problems. Controlling a solution. (10L)

The structure of decisions, development of the pay-off measure, Bernoullian utility, expected value, pay-off without a natural measure. The standard gamble. Strategies and states of nature. (5L)

The analysis of decisions, the pay-off matrix, decisions under certainty, uncertainty, risk and competition; optimality criteria of pessimism, optimism and regret. A decision among decision criteria. Laplace criterion. (8L)

Sequential decisions, decision trees, informal analysis of decision trees, cutting decision trees. (6L)

Decision-making using expected money value and utility, expected Profit with perfect information, conditional profits, value of perfect information, value of sample information, expected net gain due to sampling. (8L)

Decisions under competition, two-person zero-sum games, games in normal and extensive forms, pure and mixed strategies, saddle points, dominance principle, uniqueness of value, solving 2x2 games, solving 2xm games graphically, co-operative, competitive and vindictive solutions. Elements of non-zero sum games, equilibrium points. (15L)

Decision problems in production—allocation of resources and linear programming formulations, the transportation problem, the assignment problem, economic order quantity model for inventory, inventory problems with random demand, critical path methods for production planning-PERT and CPM, project crashing. (20L)

Decision problems in marketing—brand-loyalty model, brand-share model, pricing problem, competitive bidding, allocation of advertising funds. (10L)

Decision problems in finance—investment decision trees, risk analysis, portfolio selection, dividend policy. (8L)

(90L)

REFERENCES

Raiffa, H. and Schlaifer, R. (1968). Applied Statistical Decision Theory. MIT Press.

Luce, R. D. and Raiffa, H. (1957). Games and Decisions. Wiley.

Levin, R. I. And Kirkpatrick, C. A. (1965). Quantitative Approaches to Management. Mc Graw Hill, Int'l Student Edition.

Johnson, R. D. and Siskin, B. R.(1976). Quantitative Techniques for Business Decisions. Prentice-Hal, India.

Kemeny, J. G., Schleifer Jr.,A., Snell, J. L. and Thompson, G. L. (1968). Finite Mathematics with Business Applications. Prentice-Hall, India.

Miller,D.W. and Starr,M.K.(1973) Executive Decisions and Operations Research, Prentice Hall of India.

Guidelines for the Project Work : B.Sc. Statistics (Honours):

1. A project work shall be normally offered in the third year of the honours program.
2. A project work shall be spread over the complete academic year.

3. A project work shall be assessed for a maximum of 150 marks. The assessment will be based on the project report (100 marks), presentation (25 marks) and viva-voce (25 marks).

4. A project may be undertaken by a group of students and the maximum number of students in a team shall not exceed five. However, the project report shall be submitted by each member of the team separately.

A project work shall be supervised by a faculty member assigned by the Head of the Department.

There shall be an external examiner and an internal examiner (preferably the supervisor) for the evaluation of the project work.

A project work should encourage a student to be able to interact with the end user.

A project work should be chosen such that there is enough scope to apply and demonstrate the statistical techniques learnt in the honours program.

A project report shall clearly state the problem addressed, the methodology adopted, the assumptions and hypotheses formulated, any previous references to the study undertaken, statistical analyses performed and the broad conclusion drawn.

M.A./M.Sc. Statistics Syllabus

Course No.	Title	Contact hours per week
Semester I		
101	Real Analysis	3 lectures, 1 tutorial
102	Linear Algebra	3 lectures, 1 tutorial, 1 practical
103	Distribution Theory	3 lectures, 1 tutorial
104	Sample Surveys and Statistics for National Development	3 lectures, 2 Practicals
105	Statistical Computing	2 lectures, 3 practicals
Semester II		
201	Probability	3 lectures, 1 tutorial
202	Stochastic Processes	3 lectures, 1 tutorial
203	Inference I	3 lectures, 1 tutorial, 1 practical
204	Linear Models and Regression Analysis	3 lectures, 1-2 practicals
205	Multivariate Analysis	3 lectures, 2 practicals
Semester III		
301	Inference II	3 lectures, 1 tutorial, 1 practical
302	Design and Analysis of of Experiments	3 lectures, 2 practicals
303	Course from the selected module	
304	Course from the selected module	
305	Project	5 hours per week
Semester IV		
401	Course from the selected module	
402	-do-	
403	-do-	
404	-do-	
405	Project	5 Hours per week

SEMESTER I

101 : REAL ANALYSIS

Recap of elements of set theory; Introduction to real numbers, Introduction to n-dimensional Euclidian space; open and closed intervals (rectangles), compact sets, Bolzano - Weirstrass theorem, Heine – Borel theorem. (15L)

Sequences and series; their convergence. (7L)

Real valued functions, continuous functions; uniform continuity, sequences of functions, uniform convergence; Power series and radius of convergence. (8L)

Differentiation, maxima - minima of functions; functions of several variables, constrained maxima - minima of functions. Multiple integrals and their evaluation by repeated integration. change of variables in multiple integration. Uniform convergence in improper integrals, differentiation under the sign of integral - Leibnitz rule. (15L)

(45L + 15T)

REFERENCES

Apostol, T. M. (1985). Mathematical Analysis, Narosa, Indian Ed.

Courant, R. and John, F. (1965). Introduction to Calculus and Analysis, Wiley.

Miller, K. S. (1957). Advanced Real Calculus, Harper, New York.

Rudin, Walter (1976). Principles of Mathematical Analysis, McGraw Hill.

102 : LINEAR ALGEBRA

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. (6L)

Vector spaces with an inner product, Gram-Schmidt orthogonalization process, orthonormal basis and orthogonal projection of a vector. (4L)

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. (6L)

Hermite canonical form, generalized inverse, Moore-Penrose generalized inverse, Idempotent matrices, Solutions of matrix equations. (5L)

Real quadratic forms, reduction and classification of quadratic forms, index and signature, triangular reduction of a positive definite matrix. (6L)

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. (12L)

Singular values and singular value decomposition, Jordan decomposition, extrema of quadratic forms, vector and matrix differentiation. (6L)

(45L, 10P)

REFERENCES

Graybill, F. A. (1983). Matrices with applications in statistics, 2nd Ed. Wadsworth.

Rao, C. R. (1973). Linear statistical inference and its applications, 2nd ed. John Wiley and Sons, Inc.

Searle, S. R. (1982). Matrix Algebra useful for Statistics. John Wiley and Sons. Inc.

ADDITIONAL REFERENCES

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

Biswas, S. (1984). Topics in Algebra of Matrices, Academic Publications.

Hadley, G. (1987). Linear Algebra, Narosa Publishing House.

Halmos, P. R. (1958). Finite-dimensional Vector Spaces, 2nd ed. D. Van Nostrand Company, Inc.

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

Rao, A. R. and Bhimasankaram, P. (1992). Linear Algebra, Tata McGraw Hill Publishing Company Ltd.

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

103 : DISTRIBUTION THEORY

Brief review of basic distribution theory. Joint, marginal and conditional p.m. fs. and p.d.fs. Standard discrete and continuous distributions. Bivariate normal, Bivariate exponential, multivariate normal and multinomial distributions. Functions of random variables and their distributions using Jacobian of transformation and other tools. (12L)

Compound, truncated and mixture distributions. Conditional expectation, correlation, multiple and partial correlation. Linear and multiple regression. (8L)

Markov, Holder, Jensen and Liapunov inequalities. (4L)

Approximating distributions. Delta method and its applications. Approximating distributions of sample moments. Transformation of statisitcs. (4L)

Sampling distributions. Non-central chi-square, t- and F- distributions and their properties. Distributions of quadratic forms under normality and related distribution theory. (10L)

Order statistics- their distributions and properties. Joint and marginal distributions of order statistics. Extreme values and their asymptotic distributions (Statement only) with applications. (7L)

(45L + 10P)

REFERENCES

Dudewicz, E.J. and Mishra, S.N. (1988) : Modern Mathematical Statistics, Wiley, Int'l Students' Edition.

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

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

ADDITIONAL REFERENCES

Pitman, J. (1993) : Probability, Narosa Publishing House.

Johnson, S. and Kotz, (1972) : Distributions in Statistics, Vol. I, II and III, Houghton and Miffin.

Cramer H. (1946), Mathematical Methods of Statistics, Princeton.

104 : SAMPLE SURVEYS AND STATISTICS FOR NATIONAL DEVELOPMENT

A. Sample Surveys : Review of basic finite population sampling techniques [srs wr/wor, stratified, systematic] and related results on estimation of population mean/total. Allocation problem in stratified sampling. (5L)

Unequal probability sampling: pps wr/wor methods [including Lahiri's scheme] and related estimators of a finite population mean [Hansen-Hurwitz and Desraj estimators for a general sample size and Murthy's estimator for a sample of size 2]. (8L)

Ratio and regression estimators based on srswor method of sampling. Two-stage sampling with equal number of second stage units. Double sampling. Cluster sampling. (8L)

Randomized response technique [Warner's model : related and unrelated questionnaire methods] (4L)

(25L + 15P)

REFERENCES

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

Cochran, W. G. : Sampling Techniques [3rd Edition, 1977). Wiley

Des Raj and Chandak (1998) : Sampling Theory. Narosa

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

Sukhatme et al (1984). Sampling Theory of Surveys with Applications. Iowa State University Press & IARS.

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

B. Statistics for National Development.

Economic development : Growth in per capita income and distributive justice. Indices of development, Human Development Index. (4L)

Estimation of national income - product approach, income approach and expenditure approach. (4L)

Population growth in developing and developed countries. Population projection using Leslie marix. Labour force projection. (4L)

Measuring inequality in incomes, Gini coefficient, Theil's measure. (2L)

Poverty measurement - different issues, measures of incidence and intensity, combined measures e.q., indices due to Kakwani, Sen etc. (6L)

(20L + 2T + 3P)

REFERENCES

CSO (1980): National Accounts Statistics - Sources and Health.

Keyfitz, N. (1977) : Applied Mathematical Demography, Springer Verlag.

UNESCO : Principles for Vital Statistics Systems. Series M -12.

Sen, A. (1997) : Poverty and Inequality.

105 : STATISTICAL COMPUTING

Programming in a high level language such as C (preferred) or FORTRAN. 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 by 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. (10L)

Numerical analysis and statistical applications. The purpose of this unit is to apply programming skills in methods and algorithms useful in probability, statistics, and data analysis. Topics should include numerical integration, root extraction, random number generation, Monte Carlo integration, and matrix computations. Should there be previous exposure to numerical analysis, more advanced techniques such as permutation tests and simulation of Poisson processes can be presented. (10L)

A statistical package such as MINITAB, SAS, or SPSS. The purpose of this unit is to use a statistical package to carry out statistical procedures already known to students. No 'new' statistical methods should be presented but interesting data can be analyzed using known methods on the package. Topics should include graphics, descriptive statistics, representation of multivariate data, simple hypothesis tests, analysis of variance, and linear regression. (10L)

(30L+45P)

REFERENCES

B.W. Kernighan and D.M. Ritchie (1988). The C Programming Language, Second Edition. Prentice Hall.

W.H. Press, S.A. Teukolsky, W.T. Vellering, and B.P. Flannery (1993). Numerical Recipes in C, Second Edition. Cambridge University Press.

B. Ryan and B.L. Joiner (2001). MINITAB Handbook, Fourth Edition. Duxbury.

R.A. Thisted (1988). Elements of Statistical Computing. Chapman and Hall.

Note on suggested books: The choice of textbooks and references depends on the programming language and statistical package used. The above assumes C and INITAB.

SEMESTER II

201 : PROBABILITY

Classes of sets, fields, sigma-fields, minimal sigma-field, Borel sigma-field in R_k, sequence of sets, limsup and liminf of a sequence of sets. Measure, Probability measure, properties of a measure, Caratheodory extension theorem (statement only), Lebesgue and Lebesgue-Steljes measures on R_k. (12L)

Measurable functions, Random variables, sequence of random variables, almost sure convergence, convergence in probability (and in measure). Integration of a measurable function with respect to a measure, Monotone convergence theorem, Fatou's lemma, Dominated convergence theorem. (12L)

Borel- Cantelli Lemma, Independence, Weak law and strong law of large numbers for iid sequences, Definition and examples of Markov dependence, Exchangeable sequences, m-dependent sequences, stationary sequences. (9L)

Convergence in distribution, characteristic function, uniqueness theorem, Levy's continuity theorem (statement only), CLT for a sequence of independent random variables under Lindeberg's condition, CLT for iid random variables. (12L)

(45L + 15T)

REFERENCES

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

Billingsley, P. (1986). Probability and Measure. Wiley.

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

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

202 : STOCHASTIC PROCESSES

Introduction to stochastic processes (sp's).; classification of sp's according to state space and time domain. Countable state Markov chains (MC's), Chapman-Kolmogorov equations; calculation of n-step transition probability and its limit. Stationary distribution, classification of states; transient MC; random walk and gambler's ruin problem; Applications from social, biological and physical sciences. (12L)

Discrete state space continuous time MC: Kolmogorov- Feller differential equations; Poisson process, birth and death process; Applications to queues and storage problems. Wiener process as a limit of random walk; first -passage time and other problems. (8L)

Renewal theory: Elementary renewal theorem and applications. Statement and uses of key renewal theorem; study of residual life time process. Stationary process: weakly stationary and strongly stationary processes; Moving average and auto regressive processes. Branching process: Galton-Watson branching process, probability of ultimate extinction, distribution of population size. Martingale in discrete time, inequality, convergence and smoothing properties. Statistical inference in MC and Markov processes. (15L)

(45L + 15T)

REFERENCES

Adke, S. R. and Manjunath, S. M. (1984): An Introduction to Finite Markov Processes, Wiley Eastern.

Bhat,B. R. (2000): Stochastic Models: Analysis and Applications, New Age International, India.

Cinlar, E. (1975): Introduction to Stochastic Processes, Prentice Hall.

Feller, W. (1968): Introduction to Probability and its Applications, Vol.1, Wiley Eastern.

Harris, T.E. (1963): The Theory of Branching Processes, Springer- Verlag.

Hoel, P.G.., Port, S.C. and Stone, C.J. (1972): Introduction to Stochastic Processes, Houghton Miffin & Co.

Jagers, P. (1974): Branching Processes with Biological Applications, Wiley.

Karlin, S. and Taylor, H. M. (1975): A First Course in Stochastic Processes, Vol.1, Academic Press.

Medhi, J. (1982): Stochastic Processes, Wiley Eastern.

Parzen, E. (1962): Stochastic Processes, Holden- Day.

203 : INFERENCE - I

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 Function, Examples from standard discrete and continuous models (such as Bernoulli, Poisson, Negative Binomial, Normal, exponential, Gamma, Pareto etc.) Plotting Likelihood Functions for these models upto two parameters. (6L)

Information in data about the parameters as variation in Likelihood Function, concept of no information, Sufficiency, Neyman Factorizability Criterion, likelihood Equivalence, Minimal Sufficient Statistic, Exponential families and Pitman Families, Invariance property

of sufficiency under one-one transformation of sample space and parameter space. Fisher Information for one and several parameters models. (7L)

Methods of estimation : maximum likelihood method, methods of moments and percentiles, 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. (11L)

Tests of Hypotheses, Concepts of critical regions, test functions, two kinds of errors, size function, power function, level, MP and UMP test in class of size \al tests, Wald's SPRT with prescribed errors of two types, 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. Definition of U-statistics and its properties as an estimator of its expectation, Introduction to Standard one sample and two sample non-parametric tests for location. Non-parametric confidence intervals for percentiles. (17L)

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. (4L)

(45L + 15T + 10P)

REFERENCES

Kale, B. K. (1999) A first Course on Parametric Inference, Narosa Publishing House.

Rohatgi V. (1988). An Introduction to Probability and Mathematical Statistics. Wiley Eastern Ltd. New Delhi (Student Edition)

ADDITIONAL REFERENCES

Lehmann E. L. (1986) - (Latest) Theory of Point Estimation (Student Edition)

Lehmann, E. L. (1986). Testing Statistical hypotheses (Student Edition)

Rao, C. R. (1973) : Linear Statistical Inference.

Dudewicz, E. J. and Mishra, S. N. (1988). Modern Mathematical Statistics. Wiley Series in Prob. Math. Stat., John Wiley and Sons, New York (International Student Edition)

Ferguson T. S. (1967). Mathematical Statistics. Academic Press.

Zacks, S. (1971). Theory of Statistical Inference, John Wiley and Sons, New York.

204 : LINEAR MODELS AND REGRESSION ANALYSIS

Gauss-Markov set-up, Normal equations and Least squares estimates, Error and estimation spaces, variances and covariances of least squares estimates, estimation of error variance, estimation with correlated observations, least squares estimates with restriction on parameters, simultaneous estimates of linear parametric functions. (12L)

Tests of hypotheses for one and more than one linear parametric functions, confidence intervals and regions, Analysis of Variance, Power of F-test, Multiple comparison tests due to Tukey and Scheffe, simultaneous confidence intervals. (9L)

Introduction to One-way random effects linear models and estimation of variance components. (4L)

Simple linear regression, multiple regression, fit of polynomials and use of orthogonal polynomials. (4L)

Residuals and their plots as tests for departure from assumptions such as fitness of the model, normality, homogeneity of variances and detection of outliers. Remedies. (6L)

Introduction to non-linear models. (4L)

Multicollinearity, Ridge regression and principal component regression, subset selection of explanatory variables, Mallow's Cp statistic. (6L)

(45L + 20P)

REFERENCES

Cook, R. D. and Weisberg, S. (1982). Residual and Influence in Regression. Chapman and Hall.

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

Gunst, R. F. and Mason, R. L. (1980). Regression Analysis and Its Applications - A Data Oriented Aproach. Marcel and Dekker.

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

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

205 : MULTIVARIATE ANALYSIS

Random sampling from a multivariate normal distribution. Maximum likelihood estimators of parameters. Distribution of sample mean vector. (3L)

Wishart matrix - its distribution and properties. Distribution of sample generalized variance. Null and non-null distribution of simple correlation coefficient. Null distribution of

partial and multiple correlation coefficient. Distribution of sample regression coefficients. Application in testing and interval estimation. (10L)

Distribution of sample intra - class correlation – coefficient in a random sample from a symmetric multivariate normal distribution. Application in testing and interval estimation. (2L)

Null distribution of Hotelling's T^2 statistic. Application in tests on mean vector for one and more multivariate normal populations and also on equality of the components of a mean vector in a multivariate normal population. (4L)

Multivariate linear regression model - estimation of parameters, tests of linear hypotheses about regression coefficients. Likelihood ratio test criterion. Multivariate Analysis of Variance [MANOVA] of one-and two-way classified data. (10L)

Classification and discrimination procedures for discrimination between two multivariate normal populations – sample discriminant function, tests ssociated with discriminant functions, probabilities of misclassification and their estimation, classification into more than two multivariate normal populations. (10L)

Principal components, Dimension reduction, Canonical variables and canonical correlation - definition, use, estimation and computation. (6L)

(45L + 30P)

REFERENCES

Anderson, T. W. (1983) : An Introduction to multivariate statistical analysis. 2nd Ed. Wiley.

Giri, N. C. (1977) : Multivariate Statistical inference. Academic Press.

Kshirsagar, A. M. (1972) : Multivariate Analysis. Marcel Dekker.

Morrison, D. F. (1976) : Multivariate statistical methods. 2nd. Ed. McGraw Hill.

Muirhead, R. J. (1982) : Aspects of multivariate statistical theory, J. Wiley.

Rao, C. R. (1973) : Linear statistical inference and its applications. 2nd Ed. Wiley.

Seber, G. A. F. (1984) : Multivariate observations. Wiley.

Sharma, S. (1996) : Applied multivariate techniques. Wiley.

Srivastava, M. S. and Khatri, C. G. (1979). : An introduction to multivariate statistics. North Holland.

Johnson, R. and Wychern (1992) : Applied multivariate Statistical analysis, Prentice - Hall, 3rd Ed.

SEMESTER III

301 : INFERENCE - II

Review of convergence in probability and convergence in distribution, Cramer and Slutsky's Theorems. (4L)

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. (8L)

Consistent Asymptotic Normal (CAN) estimator, Invariance of CAN estimator under differentiable transformation, CAN property of estimators obtained by moments and percentiles, CAN estimators obtained by moment and MLE method in one parameter exponential family, Extension to multiparameter exponential family, Examples of consistent but not asymptotically normal estimators from Pitman family. Method of maximum likelihood, CAN estimators for one-parameter Cramer family, Cramer - Huzurbazar theorem, Solution of likelihood equations, Method of scoring, Newton - Raphson and other iterative procedures, Fisher Lower Bound to asymptotic variance, extension to multiparameter case (without proof). Multinomial distribution with cell probabilities depending on a parameter. (12L)

MLE in Pitman Family and Double Exponential distribution, MLE in censored and truncated distributions. (6L)

Likelihood Ratio Test (LRT), Asymptotic distribution of LRT statistic, Wald Test, Rao's score test, Pearson \chi^2 test for Goodness of fit, Bartlett's Test for homogeneity of variances. Large Sample Tests and confidence intervals based on CAN estimators, Variance stabilizing transformation and large sample tests. Consistency of Large Sample Tests, Asymptotic power of large sample tests. (15L)

(45L + 15T + 10P)

REFERENCES

Kale, B. K. (1999) A first Course on Parametric Inference, Narosa Publishing House.

Rohatgi V. (1988). An Introduction to Probability and Mathematical Statistics. Wiley Eastern Ltd. New Delhi (Student Edition)

ADDITIONAL REFERENCES

Lehmann, E. L. (1986). Testing Statistical hypotheses (Student Edition)

Rao, C. R. (1973) : Linear Statistical Inference.

Dudewicz, E. J. and Mishra, S. N. (1988). Modern Mathematical Statistics. Wiley Series in Prob. Math. Stat., John Wiley and Sons, New York (International Student Edition)

Ferguson, T. S. (1996). A course on Large Sample Theory. Chapman and Hall, London.

302: DESIGN AND ANALYSIS OF EXPERIMENTS

Introduction to designed experiments; General block design and its information matrix (C), criteria for connectedness, balance and orthogonality; Intrablock analysis (estimability, best point estimates/interval estimates of estimable linear parametric functions and testing of linear hypotheses); BIBD- recovery of interblock information; Youden design - intrablock analysis. Analysis of covariance in a general Gauss-Markov model, applications to standard designs. (17L)

Fixed, mixed and random effects models; Variance components estimation - study of various methods; Tests for variance components; Missing plot technique - general theory and applications. (8L)

General factorial experiments, factorial effects; best estimates and testing the significance of factorial effects; study of 2 and 3 factorial experiments in randomized blocks; Complete and partial confounding. Fractional replication for symmetric factorials. Split plot and split block experiments. (10L)

Application areas: Response surface experiments; first order designs and orthogonal designs; clinical trials, longitudinal data, treatment- control designs; Model validation and use of transformation; Tukey's test for additivity. (10L)

(45L + 30P)

REFERENCES

Aloke Dey (1986): Theory of Block Designs, Wiley Eastern.

Angela Dean and Daniel Voss (1999): Design and Analysis of Experiments, Springer.

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

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

John, P.W.M.(1971): Statistical Design and Analysis of Experiments, Macmillan

Joshi,D.D.(1987): Linear Estimation and Design of Experiments, Wiley Eastern

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

Myers, R.H..(1971): Response Surface Methodology, Allyn & Bacon

Pearce,S.C.(1984): Design of Experiments, Wiley, New York

Rao,C.R..and Kleffe, J.(1988): Estimation of Variance Components and applications, North Holland.

Searle, S. R., Casella, G. and McCulloch, C. E.. (1992): Variance Components, Wiley.

MODULE 1 : PROBABILITY

M.1.1: ADVANCED DISTRIBUTION THEORY

Infinitely divisible distributions; basic properties, canonical representation of the characteristic function due to Levy - Khintchine (statement only). Limit theorems for sums of independent random variables. Characterization of lass L. Limit laws for sums of i.i.d random variables; stable laws and canonical representation of their characteristic functions (statement only). Domains of attraction and of partial attraction. (15L)

Order statistics; distributions of median and range. Discrete order statistics and their joint probability mass function. Limit distribution of k-th order statistics. Extreme value laws and their properties; asymptotic joint distribution of extreme order statistics, asymptotic distribution of central order statistic. (5L)

Max-infinite divisible and max -stable distributions. Multivariate extreme value distributions; dependence functions. (5L)

Record values and limit laws for record values. (5L)

Asymptotic distributions of U-statistics, linear functions of order statistics etc. (5L)

(45L+ 15T)

REFERENCES

Arnold, B.C., Balakrishnan, N and Nagaraja, H. N. (1992): A First Course in Order Statistics, Wiley.

Feller, W. (1971): An Introduction to Probability Theory and its Applications, Vol. II, Wiley.

Galambos, J. (1978): The Asymptotic Theory of Extreme Order Statistics, Wiley.

Gnedenko, B. V. and Kolmogorov, A. N. (1954): Limit Distributions for Sums of Independent Random Variables, Addison-Wesley Publishing Company.

Laha, R. G. and Rohatgi, V. K. (1979): Probability Theory, Wiley.

Resnick, S. I. (1987): Extreme Values, Regular Variation and Point Processes, Springer - Verlag.

M.1.2 : FUNCTIONAL ANALYSIS

Metric Spaces, topological notions such as limit points, convergence, closed and open sets, continuity, compactness in metric spaces, isometries and homeomorphisms, complete metric spaces, separable spaces, nowhere-dense sets, the metric spaces R(infinity) and c[0,1]. (12L)

Normed linear spaces, metric defined by a norm, equivalent norms, Banach spaces and examples (c[a,b], lp, Lp(m) spaces). (3L)

Inner product spaces, Hilbert spaces, Bessel's inequality, complete orthonormal sets, Parseval's identity, L2 space. The projection theorem for Hilbert spaces. (12L)

Bounded linear transformations, Banach Steinhaus theorem (principle of uniform boundedness), Open mapping theorem, the Hahn Banach theorem and applications. (14L)

The dual of a Banach space, (dual of c[a,b], lp, Lp spaces). Riesz representation theorem. (4L)

(45L + 15T)

REFERENCES

Functional Analysis (1966) George Bachman and Lawrence Narici, Academic Press.

Real Analysis (1988) H. L. Royden, MacMillan, New York.

Real and complex analysis (1966) Walter Rudin, McGraw Hill

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

Principles of Functional Analysis (1971) Martin Schechter, Academic

Functional Analysis (1973) Walter Rudin, McGraw Hill.

M.1.3: MARTINGALES

Radon - Nikodym theorem (statement only), conditional expectation, optimality property of conditional expectation in L2. (8L)

Martingales in discrete time, Doob's maximal inequality, up crossings inequality. Martingale convergence theorem, reverse martingale convergence theorem, stopping time and interplay with martingales. (20L)

Various applications of martingale convergence theorem and reverse martingale convergence theorem (including Levy's theorem on convergence of conditional expectation), proof of strong law of large numbers via martingales. Choquet-Deny theorem, Kolmogorov and Hewitt - Savage zero - one laws. (10L)

Introduction to martingales in continuous time, path properties and examples: Brownian motion and Poisson process. (7L)

(45L + 15T)

REFERENCES

Billingsley, P. (1979): Probability and Measure, 3/e, Wiley, New York.

Brieman, L.(1968): Probability, Addison-Wesley.

M.1.4: PROBABILITY MEASURES ON METRIC SPACES

Probability measures on metric spaces: Measures and integrals, tightness, (measure) determining class. Weak convergence of measures: properties of weak onvergence and Pontmanteau theorem. Convergence determining class. (10L)

Some special cases: Convergence of probability measures and convergence determining class in the Euclidean space, the circle, the space R infinity, the product spaces etc. (5L)

Random element of a metric space and its distribution, convergence in distribution of a sequence of random elements - various equivalent criteria for onvergence; convergence in probability; weak convergence and mappings. (10L)

Relative compactness and Prohorov's theorem. Weak convergence and tightness in C[0,1]; Wiener measure. Weak convergence and tightness in D[0,1]; Donsker's theorem. (20L)

(45L + 15T)

REFERENCES

Billingsley, P. (1968): Convergence of Probability Measures, Wiley.

Parthasarathy, K. R. (1967): Probability Measures on Metric Spaces, Academic Press.

M.1.5 : INFERENCE IN STOCHASTIC PROCESSES

Inference in Markov chains, estimation of transition probabilities, testing for order of a Markov chain, estimation of functions of transition probabilities, Parametric models and their goodness of fit. (6L)

Markov sequences, estimation of paramaters based on likelihood and conditional least squares, auto-regressive series. (6L)

Statement of martingale strong law of large numbers and CLT for martingales, CAN property of the mle from a general sequence of dependent random variables, Fisher information. Applications to Markov chains and sequences. (6L)

Likelihood of Poisson and other Pure Jump Markov processes from first principles, CAN property of mles , testing for a Poisson process, non-homogeneous processes, Analysis of parametric Pure Jump processes, Birth-Death-Immigration processes, testing goodness of fit of such models (8L)

Diffusion processes and their likelihood, properties of estimators (without proof). (6L)

Branching processes, Inconsistency of mle/moment estimators, Properties of estimators on the non-extinction path, Asymptotic distribution theory (6L)

Elements of semi-parametric and non-parametric analysis, Theory andapplications of optimal estimating functions, estimation of transition and stationary density, intensity function of a counting process. (7L)

(45L + 15T)

REFERENCES

Billingsley, P. (1962) Statistical Inference for Markov chains , Chicago University Press, Chicago.

Basawa, I.V. and Prakasa Rao, B.L.S.(1980) Statistical Inference for Stochastic Processes, Academic Press, London.

Adke, S.R. and Manjunath, S.M.(1984) An introduction to Finite Markov Processes, Wiley Eastern, New Delhi.

Guttorp, P. (1991) Statistical Inference for Branching Processes, Wiley.

Guttorp, P.(1995) Stochastic Modelling for Scientific Data, Springer.

Bhat, B.R.(2000) Stochastic models : Analysis and Applications, New Age International Publishers, New Delhi.

Prakasa Rao, B.L.S. and Bhat, B.R.(1996) Stochastic Processes and Statistical Inference, New Age International Publishers, New Delhi.

M.1.6 : APPLIED STOCHASTIC MODELS

Renewal theory : renewal function and its properties, elementary and key renewal theorems, asymptotic normality of the number of renewals, cost/rewards associated with renewals, regenerative processes. Applications (12L)

A review of Markov chains analysis, computing stationary distribution of a large(finite) state-space, applications. (5L)

Poisson process and Pure Jump Markov processes : a review of main results. Birth and Death processes, Applications to Queuing theory, inventory analysis, communication models, population biology and finance, ecology and epidemilogy (depending upon the interests of students and the teacher). Computational issues. (14L)

Brownian motion, applications to finance. (8L)

Semi-Markov processes and Markov decision processes : an introduction and applications to various areas. (6L)

(45L + 15P)

REFERENCES

Tijms, H.C. (1986) Stochastic Modelling and Analysis, WIley.

Medhi, J. (1982) Stochastic Processes, Wiley Eastern.

Ross, S.M. (1983) Stochastic Processes, Wiley.

Bhat, B.R. (2000) Stochastic Models : Analysis and Applications, New Age International Publications.

Feller, W. (1971) An introduction to Probability theory and its applications, Vol II.

Ross, S.M. (1970) Applied Probability models with optimization applications. Holden-Day, San Franscisco.

Wolff, R.W. (1989) Stochastic Modelling and the Theroy of Queues, Prentice Hall.

M.1.7: ADVANCED STOCHASTIC PROCESSES

Point processes: Definition and examples. Product densities and characteristic functional. Stationary point processes- Khintchine's existence theorem, intensity of point processes. Poisson process, Renewal process, doubly stochastic process. (15L)

Stationary processes: wide sense and strict sense stationarity, covariance functions-properties. Bochner's theorem, Herglotz theorem, Karunen's representation theorem. Khintchine's and Cramer's representation of wide sense stationary process. Spectral decomposition for weakly stationary process. Ergodic theorems. (15L)

Markov processes- continuous time and continuous state space, time homogeneous Markov processes, Kolmogorov's equations. Diffusion processes with Wiener process and Ornstein-Uhlenbeck process as particular cases. First passage time and reflecting barriers for Wiener process. Relation between Wiener and Uhlenbec processes. (15L)

(45L + 15T)

REFERENCES

Cox, D. R. and Miller, H. D. (1965): The theory of Stochastic Processes, Mathuen & Co, London.

Cramer, H. and Leadbetter, M. R. (1967): Stationary and Related Stochastic Processes, Wiley.

Daley, D. J. and Vere- Jones (1988): An Introduction to the Theory of Point Processes. Springer- Verlag.

Karlin, S. and Taylor, H. M. (1981): A Second Course in Stochastic Processes Academic Press.

Ross, S. M. (1983): Stochastic Processes, Wiley.

MODULE 2 : ADVANCED STATISTICAL ANALYSIS

M.2.1 : ADVANCED DISTRIBUTION THEORY

Same as M.1.1.

M.2.2: STATISTICAL DECISION THEORY

Decision problem and 2-person game, utility theory, loss functions, expected loss, decision rules (non-randomized and randomized), decision principles (conditional Bayes, frequentist), inference problems as decision problems, optimal decision rules. (6L)

Concepts of admissibility and completeness, Bayes rules, admissibility of Bayes rules. (4L)

Supporting and separating hyperplane theorems, minimax theorem of for finite parameter space, minimax estimators of Normal and Poisson means, admissibility of minimax rules. (8L)

Invariant decision rules - location parameter problems, invariance and minimaxity, admissibility of invariant rules, complete class theorem, complete and essentially complete classes in simple estimation and testing situations, estimation of a distribution function. (12L)

Multivariate normal distribution, exponential family of distributions, sufficient statistics, essentially complete classes of rules based on sufficient statistics, complete sufficient statistics. (10L)

Sequential decision rules, Bayes and minimax sequential decision rules, invariant sequential decision problems, sequential tests of a simple hypothesis against a simple alternative, SPRT and stopping rule principle. (5L)

(45L + 15T)

REFERENCES

Berger, J. O. (1985). Statistical Decision Theory and Bayesian Analysis. 2nd Ed. Springer.

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

M.2.3 : BAYESIAN INFERENCE

Subjective interpretation of probability in terms of fair odds. Evaluation of (i) subjective probability of an event using a subjectively unbiased coin (ii) subjective prior distribution of a parameter. Bayes theorem and computation of the posterior distribution. (5L)

Natural Conjugate family of priors for a model. Hyper parameters of a prior from conjugate family. Conjugate families for (i) exponential family models, (ii) models admitting sufficient statistics of fixed dimension. Enlarging the natural conjugate family by (i) enlarging hyper parameter space (ii) mixtures from conjugate family, choosing an appropriate member of conjugate prior family. Non informative, improper and invariant priors. Jeffrey's invariant prior. (8L)

Bayesian point estimation : as a prediction problem from posterior distribution. Bayes estimators for (i) absolute error loss (ii) squared error loss (iii) 0 - 1 loss. Generalization to convex loss functions. Evaluation of the estimate in terms of the posterior risk. (6L)

Bayesian interval estimation : Credible intervals. Highest posterior density regions. Interpretation of the confidence coefficient of an interval and its comparison with the interpretation of the confidence coefficient for a classical confidence interval. (5L)

Bayesian testing of Hypothesis : Specification of the appropriate form of the prior distribution for a Bayesian testing of hypothesis problem. Prior odds, Posterior odds, Bayes factor for various types of testing hypothesis problems depending upon whether the null hypothesis and the alternative hypothesis are simple or composite. Specification of the Bayes tests in the above cases. Discussion of Lindley's paradox for testing a point hypothesis for normal mean against the two sided alternative hypothesis. (7L)

Bayesian prediction problem. (2L)

Large sample approximations for the posterior distribution. (2L)

Bayesian calculations for non conjugate priors : (i) Importance sampling, (ii) Obtaining a large sample of parameter values from the posterior distribution using Acceptance - Rejection methods, Markov Chain Monte Carlo methods and other computer simulation methods. (10L)

(45L + 15T)

REFERENCES

Berger, J. O. Statistical Decision Theory and Bayesian Analysis, Springer Verlag.

Robert C. P. and Casella, G. Monte Carlo Statistical Methods, Springer Verlag.

Leonard T. and Hsu, J. S. J. Bayesian Methods. Cambridge University Press.

ADDITIONAL REFERENCES

DeGroot M. H. Optimal Statistical Decisions. McGraw Hill.

Bernando J. M. and Smith, A. F. M. Bayesian Theory, John Wiley and Sons.

Robert, C. P. The Bayesian Choice : A decision Theoretic Motivation, Springer.

Gemerman, D. Markov Chain Monte Carlo : Stochastic Simulation for Bayesian Inference, Chapman Hall.

Box, G. P. and Tiao, G. C. Bayesian Inference in Statistical Analysis, Addison - Wesley.

M.2.4 : NONPARAMETRIC AND SEMIPARAMETRIC METHODS

Empirical distribution function, Glivenko Cantelli Theorem, Kolmogorov Goodness of fit test. (6L)

One sample U-statistics, Kernel and symmetric kernel, Two sample U-statistics, Asymptotic distribution of U-statistics. UMVUE property of U-statistics,Asymptotic distribution of linear function of order statistics. (12L)

Rank tests, Locally most powerful rank tests, Linear rank statistics and their distributional properties under null hypothesis, Pitman's asymptotic relative efficiency. (9L)

One sample location problem, sign test and signed rank test, two sample Kolmogorov Smirnov tests. Two sample location and scale problems. Wilcoxon-Mann-Whitney test, normal score test, ARE of various tests based on linear rank statistics. Kruskal-Wallis K sample test. (10L)

Cox's Proportional Hazards Model, rank test (partial likelihood) for regression coefficients. Concepts of jackknifing method of Quenouille for reducing bias, Bootstrap methods, Confidence intervals. (8L)

(45L + 15P)

REFERENCES

Davison, A.C. and Hinkley, D.V. (1997) : Bootstrap methods and their application, Cambridge University Press.

Gibbons, J.D. (1985) : Nonparametric statistical inference, 2^{nd ed., Marcel Dekker, Inc.

Randles, R.H. and Wolfe, D.A. (1979) : Introduction to the theory of nonparametric statistics, John Wiley & Sons, Inc.

ADDITIONAL REFERENCES

Fraser, D.A.S. (1957) : Nonparametric methods in statistics, John wiley & sons, Inc.

Hajek, J. and Sidak, Z. (1967) : Theory of rank tests, Academic Press.

Puri, M.L. and Sen, P.K. (1971) : Nonparametric methods in multivariate analysis, John Wiley & Sons, Inc.

Cox, D.R. and Oakes, D. (1983) : Survival analysis, Chapman and Hall.

M.2.5 : ADVANCED MULTIVARIATE ANALYSIS

James-Stein estimator of the mean vector and improved estimation of dispersion matrix of a multivariate normal distribution. (3L)

Distribution of characteristic roots and vectors of Wishart matrices. Distribution of sample canonical correlations, Principal components and their sample variances. Inferences on canonical correlations and principal components. (10L)

Likelihood ratio, union-intersection and other test criteria for testing (1) hypotheses in multivariate linear and generalized linear models (growth curve models), (2) independence of sets of variables, (3) equality of covariance matrices, (4) identity of several normal populations, (5) sphericity, (6) equality of a covariance matrix to a given matrix, (7) equality of a mean vector and a covariance matrix to a given vector and a given matrix. Exact and asymptotic distributions of the test criteria. Associated confidence regions and simultaneous confidence intervals. Unbiasedness and monotonicity of the power function of the tests. (18L)

Invariant multivariate tests. (4L)

Factor analysis, Linear factor models, Estimation of factor loadings, Factor rotation, Estimation of factor scores, Testing goodness of fit. Cluster analysis. (10L)

(45L + 20P)

REFERENCES

Anderson T.W. (1984) An introduction to multivariate statistical analysis, 2nd Ed., J. Wiley.

Eaton M.L. (1983) Multivariate statistics-a vector space approach, J. Wiley.

Giri N.C. (1977) Multivariate statistical inference, Academic Press.

Kshirsagar A.M. (1972) Multivariate analysis, Marcel Dekker.

Morrison D.F. (1976). Multivariate statistical methods, McGraw Hill.

Muirhead, R. J. (1982) Aspects of multivariate statistical theory, J. Wiley.

Rao C.R. (1973) Linear statistical inference and its applications, J. Wiley.

Roy S.N. (1957) Some aspects of multivariate analysis, J. Wiley.

Srivastava M.S. and Khatri C.G. (1979)An introduction to multivariate statistics, North Holland.

M.2.6 : ADVANCED SAMPLING THEORY

Two-stage sampling with unequal number of second stage units. Issues in stratified sampling : allocation problems involving several study variables, stratum boundary determination problems, Double Sampling. (7L)

Introduction to he unified theory of finite population sampling. (3L)

Horvitz - Thompson Estimator (HTE) of a finite population total/mean, expressions for V(HTE) and its unbiased estimator. Issues in non-negative variance estimation. IPPS schemes of sampling due to Midzuno-Sen, Brewer, Durbin and JNK Rao (sample size 2 only), Rao-Hartley-Cochran sampling scheme for sample size n with random grouping. (12L)

Issues in small area estimation - synthetic and generalized regression estimators. (6L)

Non-sampling errors and biased responses, randomized responses for variables, errors in surveys, modeling observational errors, estimation of variance components, application to longitudinal studies (repetitive surveys). (8L)

Variance estimation, method of random groups, balanced half samples (IPNSS), Jack-knife method. (5L)

Introduction to superpopulation models. (4L)

(45L + 30P)

REFERENCES

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

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

Cochran, W. G. (1984). Sampling Techniques. 3rd Ed. Wiley.

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

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

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

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

Murthy, M. N. (1977) Sampling theory and methods. Stat. Publ. House, Calcutta.

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

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

M.2.7 : TIME SERIES ANALYSIS

Same as M.6.1.

M.2.8 : APPLIED REGRESSION ANALYSIS

Residuals and their analysis, Influential observations, Power transformations for dependent and independent variables. (6L)

Robust and L-1 regression, Estimation of prediction error by cross-validation and boot-strap. (8L)

Non-linear regression models, Different methods of estimation (Least squares, Maximum Likelihood), Asymptotic properties of estimators (8L)

Generalized linear models, Analysis of binary and grouped data by using logistic models, Log-linear models. (8L)

Random and mixed effect models, Max imum likelihood, MINQUE and restricted maximum likelihood estimators of variance components, Best linear unbiased predictors (BLUP), Growth curves. (15L)

(45L + 20P)

REFERENCES

Bates, D.M. and Watts,D.G. (1988). Nonlinear Regression Analysis and its Application, Wiley, New York.

Cook,R.D. and Weisberg,S. (1982). Residuals and Inference in Regression, Chapman and Hall, London.

Draper,N.R. and Smith,H. (1998). Applied Regression Analysis, 3rd Ed.,Wiley,New York

Efron,B. and Tibsirani,J.R. 1993). An Introduction to the Bootstrap, Chapman and Hall,New York

Kshirsagar,A.M. 1995). Growth Curves, Marcel and Dekker, New York.

McCullagh,P. and Nelder,J.A.(1989). Generalized Linear Models, 2nd Ed.,Chapman and Hall, London.

Searle,S.R. (1987). Linear Models for Unbalanced Data, Wiley, New York.

Seber,G.A. and Wild,G.J. (1989). Nonlinear Regression, Wiley, New York.

M.2.9 : ADVANCED DESIGN & ANALYSIS OF EXPERIMENTS

Finite group and finite field, finite geometry - projective and Euclidean, construction of complete set of mols, lattice designs and their analyses, construction of BIBDs using mols, finite geometry and difference method of lose, inter and intra-block analyses of a BIBD. (10L)

Two -associate PBIB designs - association scheme and intra-block analysis, group divisible designs, dual and linked block designs, resolvable and affine-resolvable designs, general row-column designs - connectedness and intra-block analysis. (10L)

Fractional factorial designs, orthogonal and balanced arrays and their connections with confounded and fractional factorials (6L)

Response surface designs - orthogonality, rotatability and blocking, construction and analysis, method of steepest ascent (6L)

Experiments with mixtures - models, analysis and designs (4L)

Optimum designs - various optimality criteria and their interpretations, regression designs - exact and approximate designs, optimal linear and quadratic regression designs over [-1,1], Equivalence Theorem [statement and simple applications only], optimality of BIBD, optimal chemical balance weighing designs, optimality of 2 factorial designs (7L)

Repeated measurements designs [first order residual effects] - analysis (2L)

(45L + 20P)

REFERENCES

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

Bose, R. C. and Shimamoto, T. (1952, 1973). Classification and analysis of PBIB Designs with two associate classes. Jour. Amer. Stat. Assoc. Vol. 47, pp 151-184.

Chakrabarty, M. C. (1962). Mathematics of Design of Experiments. Asia Publ. House.

Cornell, M. (1963). Mixture Experiments. Wiley.

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

Khuri, A. and Cornell. M. (1991). Response Surface Methodology. Marcel Dekker.

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

Raghavarao, D. (1971). Construction and Combinatorial Problems in Design of Experiments. Wiley.

Rao, C.R. and Kleffe, J. (1988). Estimation of variance Components and Applications. North-Holland, Amsterdam.

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

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

MODULE 3 : INDUSTRIAL STATISTICS AND OPERATIONS RESEARCH

M.3.1 : STATISTICAL PROCESS AND QUALITY CONTROL

Basic concept of process monitoring and control, process capability and process optimization. (5L)

General theory and review of control charts for attribute and variable data; O.C. and A.R.L. of control charts; control by gauging; Moving average and exponentially weighted moving average charts; Cu-sum charts using V-masks and decision intervals; Economic design of X-bar chart. (10L)

Acceptance sampling plans for attribute inspection; single, double and sequential sampling plans and their properties; Plans for inspection by variables for one-sided and two-sided specifications; Mil Std and IS plans; Continuous sampling plans of Dodge type and Wald-Wolfiwitz type and their properties. Bayesian sampling plans. (10L)

Capability indices Cp, Cpk and Cpm ; estimation, confidence intervals and tests of hypotheses relating to capability indices for Normally distributed characteristics. (5L)

Use of Design of Experiments in SPC; factorial experiments, fractional factorial designs, construction of such designs and analysis of data. (10L)

Multivariate quality control; use of control ellipsoid and of utility functions. (5L)

(45L+20P)

REFERENCES

Montgomery, D.C. (1985) Introduction to Statistical Quality Control; Wiley

Montgomery, D.C. (1985) Design and Analysis of Experiments; Wiley

Ott, E.R. (1975) Process Quality Control; McGraw Hill

Phadke, M.S. (1989) Quality Engineering through Robust Design; Prentice Hall

Wetherill, G.B. (1977) Sampling Inspection and Quality Control; Halsted Press

Wetherill, G.B. and Brown, D.W. ( ) Statistical Process Control, Theory and Practice; Chapman and Hall

M.3.2 : PLANNING AND ANALYSIS OF INDUSTRIAL EXPERIMENTS

Analysis of single replicate of 2k Full Factorial Experiment, total and artial confounding in 2^k Full Factorial Experiment, Resolution III, IV, and V fractions of 2k experiments.

Criteria in selecting factorial designs: Criteria based on the Spectrum of the information matrix-A and D- optimality, criteria based on alias matrix.

Construction of layouts of orthogonal array experiments and associated linear graphs to study some of the main effects and first order interactions of 2^k designs which need not be resolution 3 designs, (designs known as Taguchi designs) with special cases of L_8, and L_{16. (20L)

3^k Full factorial designs, . Total and partial confounding in 3k Factorial experiments. Construction of Orthogonal array experiments involving three level factors, with special cases of L_9 , and L_{18. (10L)

Roll of Center Composite Designs (CCD) as alternative to 3k designs, rotatability of CCD, Linear and quadratic Response surfaces, contour plots. (8L)

Roll of non normality, Box- Cox transformation, Generalized linear models(GLIM), For exponential family of distributions, (7L)

(45L + 30P)

REFERENCES

D. C. Montgomery: Design and Analysis of Experiments, J. Wiley and Sons(Asia) 5th edition (2001).

R. H. Myers & D.C. Montgomery: Response Surface Methodology, J. Wiley and Sons.

J. Fox: Quality through Design, McGraw-Hill Book Company, 1993.

J. A. Nelder and P. McCullagh: Generalized Linear Models, 2nd edition.

B. L. Raktoe, A. Hedayat and W. T. Federer, Factorial Designs, J. Wiley and Sons, 1981.

M.3.3 : RELIABILITY

Reliability concepts and measures; components and systems; coherent systems; reliability of coherent systems; cuts and paths; modular decomposition; bounds on system reliability; structural and reliability importance of components. (8L)

Life distributions; reliability function; hazard rate; common life distributions-exponential, Weibull, gamma etc. Estimation of parameters and tests in these models. (10L)

Notions of ageing; IFR, IFRA, NBU, DMRL, and NBUE Classes and their duals; loss of memory property of the exponential distribution; closures or these classes under formation of coherent systems, convolutions and mixtures. (10L)

Univariate shock models and life distributions arising out of them; bivariate shock models; common bivariate exponential distributions and their properties. (5L)

Reliability estimation based on failure times in variously censored life tests and in tests with replacement of failed items; stress-strength reliability and its estimation. (5L)

Maintenance and replacement policies; availability of repairable systems; modeling of a repairable system by a non-homogeneous Poisson process. (4L)

Reliability growth models; probability plotting techniques; Hollander-Proschan and Deshpande tests for exponentiality; tests for HPP vs. NHPP with repairable systems. (5L)

Basic ideas of accelerated life testing. (2L)

(45L+16P)

REFERENCES

Barlow R.E. and Proschan F.(1985) Statistical Theory of Reliability and Life Testing;Holt,Rinehart and Winston.

Lawless J.F. (1982) Statistical Models and Methods of Life Time Data; John Wiley.

Bain L.J. and Engelhardt (1991) Statistical Analysis of Reliability and Life Testing Models; Marcel Dekker.

Nelson, W (1982) Applied Life Data analysis; John Wiley.

Zacks S. Reliability Theory, Springer.

M.3.4 : STATISTICAL METHODS FOR TOTAL QUALITY MANAGEMENT

Quality Systems: ISO 9000 standards, QS 9000 standards. Concept of six-sigma and the Define-Measure-Analyse-Improve-Control Approach. Precision and accuracy in measurement systems. Estimation of Measurement Uncertainty. (8L)

Total Quality Management, Process Analysis and Optimization. (8L)

Quality at Design stage, Quality Function Deployment, Failure Mode and Effect Analysis, Conjoint Analysis. System, parameter and tolerance designs. Planning and analysis of fractional factorial experiments. Basic ideas of response surface methodology and contour plots. (15L)

Quality in manufacturing, control charts for attribute and variable characteristics, process adjustments based on control chart evidences. Process capability and performance indices. Evolutionary operations. (10L)

Measuring customer satisfaction, American Customer Satisfaction Index Model. (4L)

(45L+18P)

REFERENCES

Logothetis, N. (1992). Managing Total Quality; Prentice Hall of India.

Oakland J.S. (1989). Total Quality Management; Butterworth-Heinemann.

Mittag H.J. and Rinne H. (1993) Statistical Methods of Quality Assurance.

Montgomery D.C.(1985); Statistical Process Control; John Wiley.

Montgomery D.C.(1999) Design and Analysis of Experiments; John Wiley

M.3.5 : OPERATIONS RESEARCH I

Definition and scope of Operational research; phases in Operations Research; models and their solutions; decision-making under uncertainty and risk, use of different criteria; sensitivity analysis. (9L)

Review of LP problems; duality theorem; transportation and assignment problems; Non-linear programming-Kuhn Tucker conditions, Wolfe's and Beale's algorithms for solving quadratic programming problems. Bellman's principle of optimality, general formulation, computational methods and application of dynamic programming. (10L)

Analytical structure of inventory problems; EOQ formula of Harris, its sensitivity analysis and extensions allowing quantity discounts and shortages. Multi-item inventory subject to constraints. Models with random demand, the static risk model. P and Q-systems with constant and random lead times. (9L)

Queueing models-specifications and effectiveness measures. Steady-state solutions of M/M/1 and M/M/c models with associated distributions of queue-length and waiting time. M/G/1 queue and Pollazcek Khinchine result. Steady-state solutions of M/Ek/1 and Ek/M/1 queues. Machine interference problem. (9L)

Sequencing and scheduling problems. 2 machine n-job and 3-machine n-job problems with identical machine sequence for all jobs; 2-job n-machine problem with different routings. Branch and bound method for solving travelling salesman problem. (8L)

(45L+28P)

REFERENCES

Taha H.A. (1982) Operational Research: An Introduction; Macmillan.

Hillier F.S. and Leiberman G.J. (1962) Introduction to Operations Research; Holden Day.

Kanti Swarup, Gupta,P.K. and Singh,M.M.. (1985) Operations Research; Sultan Chand & Sons.

Philips D.T.,Ravindran A. and Solberg J.( ) Operations Research, Principles and Practice.

Churchman C.W., Ackoff R.L. and Arnoff E.L. (1957) Introduction to Operations Research; John Wiley.

M.3.6 :OPERATIONS RESEARCH II

Multi-stage decision processes and Dynamic Programming. Integer programming-branch and bound algorithm and cutting plane algorithm. Multi-criterion and goal programming. Stochastic programming; quantile rules, two-stage programming, use of fractional programming. (12L)

Decision-making in the face of competition, two-person games, pure and mixed strategies, existence of solution and uniqueness of value in zero-sum games, finding solutions in 2x2, 2xm and mxn games. Non-zero sum games, co-operative and competitive games, equilibrium solutions and their existence in bi-matrix games. Nash equilibrium solution. (11L)

s-S policy for inventory and its derivation in the case of exponential demand; multi-echelon inventory models; models with variable supply and models for perishable items; estimation of EOQ in some simple cases. (6L)

Transient solution of M/M/1 queue; bulk queues (bulk arrival and bulk service); finite queues; queues in tandem; GI/G/1 queue and its solution; simulation of queues. Flows in networks; max-flow min-cut theorem. (8L)

Replacement problems; block and age replacement policies; dynamic programming approach for maintenance problems; replacement of items with long life. Project management; PERT and CPM; probability of project completion, PERT-crashing. (8L)

(45L+24P)

REFERENCES

Hadley G. (1964) Non-linear and Dynamic programming; Addison Wesley

Murthy K.G. (1976) Linear and Combinatorial Programming; John Wiley

Kleinrock L. (1975) Queueing Systems, vol. 1, Theory; John Wiley

Saaty T.L. (1961) Elements of Queueing Theory with Applications; McGraw Hill

Hadley G. and Whitin T.M. (1963) Analysis of Inventory Systems; Prentice Hall

Starr M.K. and Miller D.W. (1962) Inventory Control-Theory and Practice; Prentice Hall

Mckinsey J.C.C. (1952) Introduction to the Theory of Games; McGraw Hill

Wagner H.M. (1973) Principles of O.R. with Applications to Managerial Decisions; Prentice Hall

Gross, D. and Harris,C.M. (1974) Fundamentals of Queueing Theory; John Wiley

MODULE 4 : BIOSTATISTICS

M.4.1 : BIOASSAY

Types of biological assays; Direct assays; Ratio estimators, asymptotic distributions; Fieller's theorem (5L)

Regression approaches to estimating dose-response relationships Logit and probit approaches when dose-response curve for standard preparation is unknown; Quantal responses; Methods of estimation of parameters; Estimation of extreme quantiles; Dose allocation schemes; Polychotomous quantal response (11L)

Estimation of points on the quantal response function (5L)

Sequential procedures (7L)

Estimation of safe doses (6L)

Bayesian approach to bioassay (5L)

(45L + 18P)

REFERENCES

Z. Govindarajulu (2000). Statistical Techniques in Bioassay, S. Kargar.

D. J. Finney (1971). Statistical Method in Bioassay, Griffin.

D. J. Finney (1971). Probit Analysis (3rd Ed.), Griffin.

G. B. Weatherile (1966). Sequential Methodsin Statistics, Methuen.

M.4.2 : STATISTICAL GENETICS

Basic biological concepts in genetics (relevant to this course). (4L)

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. (19L)

Non-random mating, inbreeding, phenotypic assortative mating. (8L)

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. (9L)

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 likelihood algorithms. (5L)

(45L + 10P)

REFERENCES

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.

ADDITIONAL REFERENCES

R. C. Elandt - Johnson (975). Probability Models and Statistical Methods in Genetics. John Wiley.

M.4.3 - SURVIVAL ANALYSIS

Concepts of time, Order and random Censoring, likelihood in these cases. Life distributions-Exponential Gamma, Weibull, Lognormal, Pareto, Linear Failure rate. Parametric inference (Point estimation, Confidence Intervals, Scores, LR, MLE tests (Rao-Willks-Wald)) for these distributions. (9L)

Life tables, failure rate, mean residual life and their elementary properties. Ageing classes - and their properties, Bathtub Failure rate. (4L)

Estimation of survival function - Acturial Estimator, Kaplan -Meier Estimator, Estimation under the assumption of IFR/DFR. Tests of exponentiality against non-parametric classes-Total time on test, Deshpande test. (10L)

Two sample problem-Gehan test, Log rank test. Mantel-Haenszel test, Tarone - Ware tests. (6L)

Semi-parametric regression for failure rate - Cox's proportional hazards model with one and several convariates. Rank test for the regression coefficients. (6L)

Competing risks model, parametric and non-parametric inference for this model. (6L)

Multiple decrement life table. (4L)

(45L + 15P)

REFERENCES

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 Distribution : Reliability applications in the Biomedical Sciences, John Wiley and Sons.

Elandt - Johnson, R.E. Johnson N.L. : Survival Models and Data Analysis, John Wiley and Sons.

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

Kalbfleisch J.D. and Prentice R.L. (1980), The Statistical Analysis of Failure Time Data, John Wiley.

M.4.4 : STATISTICAL ECOLOGY

Introduction to Ecology and evolution. (2L)

Population dynamics : single species - Exponential, logistic and Gompertz models, Leslie matrix model for age and stage structured population, Survivorship curves - Constant, monotone and bath tub shaped hazard rates. (10L)

Two species : Lotka - Volterra equations, isoclines, competition and coexistence, predator-pray oscillations. (8L)

Abundance estimation : Capture - recapture, Nearest neighbour, line transect sampling, indirect methods. (10L)

Ecological Diversity : Species abundance curve, Indices of diversity (Simpson's index, Shannon - Wiener index), Diversity as average rarity. (4L)

Harvesting renewable biological resources - Maximum sustainable yield, tragedy of the commons. (4L)

Game theory in ecology - Evolutionarily stable strategy, its properties, simple games such as Hawk - Dove game. (4L)

Foraging theory : Optimal foraging, diet choice, mean variance trade-off. (3L)

PRACTICALS

Fitting exponential and logistic growth models to data by linearisation method and other methods.

Verification of properties of the discrete logistic model for different values of r the exponential growth rate.

Fitting monotone and bath-tub shaped hazard functions to cohort data.

Simulation study of various abundance estimation methods by the following route : (a) choosing convenient values for parameters (b) generating data sets from such selected configurations (c) obtaining estimates of population density and their sampling distributions.

Writing programs for computing various indices of biological diversity.

(45L + 15P)

REFERENCES

Anil Gore and S. A. Paranjpe (2000). A course on mathematical and Statistical Ecology (Kluwer)

Clark, C. W. (1976). Mathematical Bioeconomics : Optimal Management of Renewable Resources (Wiley).

Maynard Smith, J. (1982). Evolution and the Theory of Games (Cambridge University Press).

Pielou, E. C. (1977). An Introduction to Mathematical Ecology (Wiley).

Seber, G. A. F. (1982). Estimation of Animal Abundance and Related Parameters (Charles Griffin).

Stephens, D. W. and Krebs, J. R. (1986). Foraging Theory (Princeton University Press).

M.4.5 : QUANTITATIVE EPIDEMIOLOGY

Introduction to modern epidemiology, principles of epidemiologic investigation, surveillance and disease monitoring in populations, epidemiologic resources on the Web. (3L)

Epidemiologic measures : organizing and presenting epidemiologic data, measures of disease frequency, review of vital statistics and demography, measures of effect and association, causation and causal inference. (5L)

Design and analysis of epidemiologic studies : types of studies, case-control studies, cohort studies, cross-over designs, regression models for the estimation of relative risk, meta-analysis, quantitative methods in screening. (25L)

Special topics : epidemiology of infectious and chronic diseases, epidemiology of cancer and cancer prevention, environmental epidemiology, molecular and genetic epidemiology. (12L)

(45L + 15P)

REFERENCES

K. J. Rothman and S. Geenland (ed.) (1998). Modern Epidemiology, Lippincott-Raven.

S. Selvin (1996). Statistical Analysis of Epidemiologic Data, Oxford University Press.

D. McNeil (1996). Epidemiological Research Methods. Wiley and Sons.

J. F. Jekel, J. G. Elmore, D.L. Katz (1996). Epidemiology, Biostatistics and Preventive Medicine. WB Saunders Co.

M.4.6 : CLINICAL TRIALS

Introduction to clinical trials : the need and ethics of clinical trials, bias and random error in clinical studies, conduct of clinical trials, overview of Phase I-IV trials, multi-center trials. (3L)

Data management : data definitions, case report forms, database design, data collection systems for good clinical practice. (2L)

Design of clinical trials : parallel vs. cross-over designs, cross-sectional vs. longitudinal designs, review of factorial designs, objectives and endpoints of clinical trials, design of Phase I trials, design of single-stage and multi-stage Phase II trials, design and monitoring of Phase III trials with sequential stopping, design of bioequivalence trials. (20L)

Reporting and analysis : analysis of categorical outcomes from Phase I - III trials, analysis of survival data from clinical trials. (15L)

Surrogate endpoints : selection and design of trials with surrogate endpoints, analysis of surrogate endpoint data. (2L)

Meta-analysis of clinical trials. (3L)

(45L + 18P)

REFERENCES

S. Piantadosi (1997). Clinical Trials : A Methodologic Perspective. Wiley and Sons.

C. Jennison and B. W. Turnbull (1999). Group Sequential Methods with Applications to Clinical Trials, CRC Press.

L. M. Friedman, C. Furburg, D. L. Demets (1998). Fundamentals of Clinical Trials, Springer Verlag.

J. L. Fleiss (1989). The Design and Analysis of Clinical Experiments. Wiley and Sons.

E. Marubeni and M. G. Valsecchi (1994). Analyzing Survival Data from Clinical Trials and Observational Studies, Wiley and Sons.

M.4.7 : DEMOGRAPHY

Coverage and content errors in demographic data, use of balancing equations and Chandrasekharan - Deming formula to check completeness of registration data. Adjustment of age data - use of Whipple, Myer and UN indices. Population composition, dependency ratio. (7L)

Measures of fertility; stochastic models for reproduction, distributions of time to first birth, inter-live birth intervals and of number of births (for both homogeneous and non-homogeneous groups of women), estimation of parameters; estimation of parity progression ratios from open birth interval data. (12L)

Measures of Mortality; construction of abridged life tables. Distributions of life table functions and their estimation.

Stable and quasi-stable populations, intrinsic growth rate. Models for population growth and their fitting to population data. Stochastic models for population growth. (8L)

Stochastic models for migration and for social and occupational mobility based on Markov chains. Estimation of measures of mobility. (6L)

Methods for population projection. Use of Leslie mateix. (6L)

(45L+ 10P)

REFERENCES

Bartholomew, D. J. (1982). Stochastic Models for Social Processes, John Wiley.

Benjamin, B. (1969). Demographic Analysis, George, Allen and Unwin.

Chiang, C. L. (1968). Introduction to Stochastic Processes in Biostatistics; John Wiley.

Cox, P. R. (1970). Demography, Cambridge University Press.

Keyfitz, N. (1977). Applied Mathematical Demography; Springer Verlag.

Spiegelman, M. (1969). Introduction to Demographic Analysis; Harvard University Press.

Wolfenden, H. H. (1954). Population Statistics and Their Compilation; American Actuarial Society.

M.4.8 :GENERALIZED LINEAR MODELS

Logisitic and Poisson regression: logit model for dichotomous data with single and multiple explanatory variables, ML estimation, large sample tests about parameters, goodness of fit, analysis of deviance, variable selection, extension to polytomous data, introduction to Poisson regression. (15L)

Log line ar models for two and three dimensional contingency tables: interpretation of parameters, comparison with ANOVA and regression, ML estimation of parameters, likelihood ratio tests for various hypotheses including independence, marginal and conditional independence, partial association, models with quantitative levels. (15L)

Nonparametric regression and generalized linear models: interpolating and smoothing splines for simple regression, use of cross-validation, applications to logistic and Poisson regression introduction to additive models and generalized additive models. (15L)

(45L + 18P)

REFERENCES

D.W. Hosmer and S. Lemeshow (2000). Applied Logistic Regression, Second edition. Wiley, New York.

A. Agesti (1990). Categorical Data Analysis. Wiley, New York.

R. Christensen (1997). Log-linear Models and Logistic Regression, Second edition. Springer, New York.

P. McCullagh and J.A. Nelder (1999). Generalized Linear Models, Second edition. Chapman and Hall, New York.

P.J. Green and B.W. Silverman (1994). Nonparametric Regression and Generalized Linear Models. Chapman and Hall, New York.

T.J. Hastie and R.J. Tibshirani (1999). Generalized ed Additive Models. Second edition. Chapman and Hall, New York.

M.4.9 : APPLIED REGRESSION ANALYSIS

Same as M.2.8.

MODULE 5 : COMPUTATIONAL STATISTICS

M.5.1 COMPUTER-INTENSIVE STATISTICAL METHODS - I

Exploratory data analysis: transforming data, graphical methods of clustering, outliers. (5L)

Linear regression: influential observations and diagnostics, robust methods, collinearity, variable selection. (11L)

Generalized linear models: exponential families and ML estimation, analysis of deviance and variable selection, logistic regression. (10L)

Nonlinear regression: estimation, hypothesis testing, goodness of fit. (6L)

EM algorithm: applications to missing and incomplete data problems, mixture models. (6L)

Smoothing with kernels: density estimation, simple nonparametric regression. (7L)

(45L + 20P)

Note on practicals: Each practical session should correspond to two teaching hours. Practical work should be done on statistical packages or using high level languages as taught in the core course on Statistical Computing.

REFERENCES

R. Gnanadesikan (1997); Methods for Statistical Data Analysis of Multivariate Observations, Second edition. Wiley.

D.A. Belsley, E. Kuh, and R.E. Welsch (1980); Regression Diagnostics. Wiley.

P. McCullagh and J.A. Nelder (1999); Generalized Linear Models, Third edition. Chapman and Hall.

G.E.F. Seber and C.J. Wild (1989); Nonlinear Regression. Wiley.

G.J. McLachlan and T. Krishnan (1997); The EM Algorithms and Extensions. Wiley.

J.S. Simonoff (1996); Smoothing Methods in Statistics. Springer.

M.5.2 COMPUTER - INTENSIVE STATISTICAL METHODS-II

Stochastic simulation : generating random variables, simulating multivariate distributions, simulating stochastic processes such as simple queues. (10L)

Variance reduction: importance sampling for integration, control variates and antithetic variables. (10L)

Markov Chain Monte Carlo methods: Gibbs sampling for multivariate simulation, simulated annealing for optimization. (7L)

Simulation based testing: simulating test statistics and power functions, permutation tests. (6L)

Bootstrap methods: resampling paradigms, bias and standard errors, confidence intervals, bootstrapping in regression. (12L)

Jackknife and cross-validation: jackknife in sample surveys, cross-validation for tuning parameters. (5L)

(45L + 20P)

Note on practicals: Each practical session should correspond to two teaching hours. Practical work should be done on statistical packages or using high level languages as taught in the core course on Statistical Computing.

REFERENCES

G.S. Fishman (1996); Monte Carlo: Concepts, Algorithms, and Applications. Springer.

R.Y. Rubinstein (1981); Simulation and the Monte Carlo Method. Wiley.

M.A. Tanner (1996); Tools for Statistical Inference, Third edition. Springer.

B. Efron and R.J. Tibshirani (1993); An Introduction to the Bootstrap. Chapman and Hall.

J. Shao and D.Tu (1995); The Jackknife and the Bootstrap. Springer Verlag.

M.5.3 COMPUTER PROGRAMMING

This course is intended to introduce object-oriented computer programming. It assumes prior exposure to programming in languages such as C and/or Fortran. The language of choice here is C++ . Should there be previous exposure to C++, Java can be substituted. The course can be followed by the Computer Science course.

Introduction to object-oriented programming concepts and design. (5L)

Programming in C++: data types and operations, functions and parameters, classes, constructors, input/output, control statements such as if-else, switch, for, while and do-while, pointers and references, dynamic allocation, processing of linked lists, arrays and character strings, libraries. (20L)

Introduction to program analysis: simple testing and debugging (5L)

Introduction to Web programming: simple examples in Java and the concept of bytecode. (5L)

(35L + 30P)

Note on practicals/tutorials: As this is a programming course, hand-on practical sessions are important and should be held in conjunction with lectures.

REFERENCES

R. Decker and S. Hirshfield (1998). The Object Concept: An Introduction to Computer Programming using C++. PWS Publishing.

S.B. Lippmann and J. Lajoie (1998). C++ Primer. Third edition. Addison-Wesley.

P. Naughton (1996). The Java Handbook. Tata McGraw-Hill.

W.J. Savitch (2001). Problem Solving with C++: The Object of Programming. Third Edition. Addison-Wesley Longman.

M.5.4 : KNOWLEDGE DISCOVERY AND DATA MINING

Review of classification methods from multivariate analysis; classification and decision trees. (7L)

Clustering methods from both statistical and data mining viewpoints; vector quantization. (7L)

Unsupervised learning from univariate and multivariate data; dimension reduction and feature selection. (7L)

Supervised learning from moderate to high dimensional input spaces; artificial neural networks and extensions of regression models, regression trees. (10L)

Introduction to databases, including simple relational databases; data warehouses and introduction to online analytical data processing. (8L)

Association rules and prediction; data attributes, applications to electronic commerce. (6L)

(45L + 16P)

Note on practicals/tutorials: Each practical session should correspond to two teaching hours. Practical work should be done on statistical packages or using high level languages as taught in the core course on Statistical Computing. Note on prerequisites: This course assumes successful completion of Computer-intensive Statistical Methods I. It can be presented in conjunction with Computer-intensive Statistical Methods II.

REFERENCES

A. Berson and S.J. Smith (1997). Data Warehousing, Data Mining, and OLAP. McGraw-Hill.

L. Breiman, J.H. Friedman, R.A. Olshen, and C.J. Stone (1984). Classification and Regression Trees. Wadsworth and Brooks/Cole.

J. Han and M. Kamber (2000). Data Mining; Concepts and Techniques. Morgan Gaufmann.

T.M. Mitchell (1997). Machine Learning. McGraw-Hill.

B.D. Ripley (1996). Pattern Recognition and Neural Networks. Cambridge University Press.

M.5.5 : STATISTICAL PATTERN RECOGNITION

Linear classifiers: linear discriminant function (LDF) for minimum squared error, LDF for binary outputs, perception learning algorithm. (12L)

Nearest neighbour decision rules: description, convergence, finite sample considerations, use of branch and bound methods. (12L)

Probability of errors: two classes, normal distributions, equal covariance matrix assumptions, Chernoff bounds and Bhattacharya distance, estimation of probability of error. (8L)

Feature selection and extraction: interclass distance measures, discriminant analysis, probabilistic distance measures, principal components. (13L)

(45L + 16P)

Note on practicals/tutorials: Each practical session should correspond to two teaching hours. Practical work should be done on statistical packages or using high level languages as taught in the core course on Statistical Computing.

Note on prerequisites: This course assumes successful completion of Computer-intensive Statistical Methods I. It can be presented in conjunction with Computer-intensive Statistical Methods II.

REFERENCES

R.O. Duda and P.E. Hart (1973). Pattern Recognition and Scene Analysis. Wiley.

K Fukunaga (1990). Introduction to Statistical Pattern Recognition, Second Edition. Academic Press.

G.J. McLachlan (1992). Discriminant Analysis and Statistical Pattern Recognition. Wiley.

B.D. Ripley (1996). Pattern Recognition and Neural Networks. Cambridge University Press.

MODULE 6 : ECONOMIC AND FINANCIAL STATISTICS

M.6.1 : TIME SERIES ANALYSIS

Time-series as discrete parameter stochastic process. Auto covariance and autocorrelation functions and their properties.

Exploratory Time Series Analysis, Tests for trend and Seasonality. Exponential and Moving Average Smoothing. Holt and Winters smoothing. Forecasting based on smoothing, Adaptive smoothing. (9L)

Detailed study of the stationary processes: (1) moving average (MA), (2) Auto regressive (AR), (3) ARMA and (4) AR integrated MA (ARIMA) models. Box-Jenkins models. Discussion (without proof) of estimation of mean, auto covariance and autocorrelation functions under large sample theory. Choice of AR and MA periods. Estimation of ARIMA model parameters. Forecasting. Residual analysis and diagnostic checking. Use of computer packages like SPSS. (24L)

Spectral analysis of weakly stationary process. Periodogram and correlogram analyses. Computations based on Fourier transform. Spectral decomposition of weakly AR process and representation as a one-sided MA process - necessary and sufficient conditions. Implication in prediction problems. (12L)

Use of software package is desirable :

1. Exploratory Analysis
2. Smoothing
3. Numerical exercises on MA and AR models - Forecasting
4. Numerical exercises on ARMA and ARIMA models - Forecasting
5. Numerical exercises on Box-Jenkins models.
6. Residual analysis and diagnostic checking
7. Computations based on Fourier transform
8. Periodogram analysis and interpretation
9. Correlogram analysis and interpretation

(45L + 18P)

REFERENCES

Box, G. E. P. and Jenkins, G. M. (1976). Time Series Analysis - Forecasting and Control, Holden-day, San Francisco.

Anderson, T. W. (1971). The Statistical Analysis of Time Series, Wiley, N. Y.

Montgemory, D. C. and Johnson, L. A. (1977). Forecasting and Time Series Analysis, McGraw Hill.

Kendall, Sir Maurice and Ord, J. K. (1990). Time Series (Third Edition), Edward Arnold.

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

ADDITIONAL REFERENCES

Fuller, W. A. (1976). Introduction to Statistical Time Series, John Wiley, N. Y.

Granger, C.W. J. and Newbold (1984). Forecasting Econometric Time Series, Third Edition, Academic Press.

Priestley, M. B. (1981). Spectral Analysis & Time Series, Griffin, London.

Kendall, M. G. and Stuart A. (1966). The Advanced Theory of Statistics, Volume 3, Charles Griffin, London.

Bloomfield, P. (1976). Fourier Analysis of Time Series – An Introduction, Wiley.

Granger, C. W. J. and Hatanka, M. (1964). Spectral Analysis of Economic Time Series, Princeton Univ. Press, N. J.

Koopmans, L. H. (1974). The Spectral Analysis of Time Series, Academic Press.

Nelson, C. R. (1973). Applied Time Series for Managerial Forecasting, Holden-Day.

Findley, D. F. (Ed.) (1981). Applied Time Series Analysis II, Academic Press.

M.6.2 : ECONOMETRICS

Nature of econometrics. The general linear model (GLM) and its extensions. Ordinary least squares (OLS) estimation and prediction. Use of dummy variables and seasonal adjustment. Generalized least squares (GLS) estimation and prediction. Heteroscedastic disturbances. Pure and mixed estimation. Grouping of observations and of equations. (10L)

Auto correlation, its consequences and tests. Theil BLUS procedure. Estimation and prediction. Multicollinearity problem, its implications and tools for handling the problem. Ridge regression. (10L)

Linear regression with stochastic regressors. Instrumental variable estimation. Errors in variables. Autoregressive linear regression. Distributed lag models. Use of principal components, canonical correlations and discriminant analyses in econometrics. (10L)

Simultaneous linear equations model. Examples. Identification problem. Restrictions on structural parameters - rank and order conditions. Restrictions on variances and covariances. (5L)

Estimation in simultaneous equations model. Recursive systems. 2 SLS Estimators. Limited information estimators, k-class estimators. 3 SLS estimation. Full information maximum likelihood method. Prediction and simultaneous confidence intervals. Monte Carlo studies and simulation. (10L)

(45L + 24P)

LAB WORK (Assignments of 2hrs each)

Use of Software packages is desirable.

1. OLS estimation and prediction in GLM.
2. Use of dummy variables (dummy variable trap) and seasonal adjustment.
3. GLS estimation and prediction.
4. Tests for heteroscedasticity ; pure and mixed estimation.
5. Tests for autocorrelation. BLUS procedure.
6. Ridge regression.
7. Industrumental variable estimation.
8. Estimation with lagged dependent variables.
9. Identification problems - checking rank and order conditions.
10. Estimation in recursive systems.
11. Two SLS estimation.
12. Simulation studies to compare OLS, 2SLS, LISE and FIML methods.

REFERENCES

Apte PG (1990); Text book of Econometrics. Tata McGraw Hill.

Cramer, J.S. (1971) : Empirical Econometrics, North Holland.

Gujarathi, D (1979) : Basic Econometrics, McGraw Hill.

Intrulligator, MD (1980) : Econometric models - Techniques and applications, Prentice Hall of India.

Johnston, J. (1984) : Econometric methods, Third edition, McGraw Hill.

Klein, L.R. (1962) : An introduction to Econometrics, Prentice Hall of India.

Koutsoyiannis, A (1979) : Theory of Econometrics, Macmillan Press.

Malinvaud, E (1966) : Statistical methods of Econometrics, North Holland.

Srivastava, V.K. and Giles D.A.E (1987) : Seemingly unrelated regression equations models, Maicel Dekker.

Theil, H. (1982) : Introduction to the theory and practice of Econometrics, John Wiley.

Walters, A (1970) : An introduction to Econometrics, McMillan & Co.

Wetherill, G.B. (1986) : Regression analysis with applications, Chapman Hall.

M.6.3 : GAME THEORY WITH APPLICATIONS IN ECONOMICS

Games as decision problems, classification of games. (2L)

Two-person Zero-sum games: normal and extensive forms, maximum criterion, Dominance, mixed strategies, minimax theorem, solutions of 2x2 and 2xm games And of mxn games (using simplex algorithm), equilibrium pairs, uniqueness of Value, games with perfect information. (8L)

Two-person non-zero-sum games : equilibrium pairs and maximin-maximin pairs Finding equilibrium pairs (Swastika method), solution concepts of such games, Proof of Nash's theorem, co-operative games, bargaining or negotiation set, Maximin bargaining solution, threat bargaining solution. (8L)

N-person games : non-cooperative games, imputations, stable sets, Shapley value Other solution concepts (4L)

Multi-stage games: stochastic games, recursive games, discounted stochastic Games, existence of value, bounds on value iteration, supergames (6L)

Game theory and public economics : allocation of pure public goods, externalities Coalition structures and local public goods, the second best tax Game, Shapley value of public goods games. (6L)

Differential games: existence of value and of saddle-point, solving the Hamilton-Jacobi equation, applications in exhaustible and renewable common property Resources. (6L)

Market games and oligopoly: 1-1 games, N-N games and M-N games, Cournot Equilibrium. (4L)

Bidding and auctions with known valuations and unknown valuations, auction of Two or more objects, horse market auctions (6L)

(50L+13P)

REFERENCES

Bacharach, M. (1977). Economics and the Theory of Games, Mcmillan Press.

Owen, G. (1968). Game Theory, Saunders.

Shubik, M. (1982). game Theory in the Social Sciences, M.I.T. Press.

Thomas, L. C. (1984). Games, Theory and Applications, John Wiley.

M.6.4 INVESTMENTS UNDER UNCERTAINTY

Main theme : Risk - Return Trade off. (3L)

Money market, Fixed income, equity, stocks and bonds, Treasury notes, Market indexes, Rates of interest, compound interest, inflation, Risk in a portfolio context, law of one price and arbitrage. (9L)

Risk and risk aversion, mean variance analysis, allocation between risky and risk free portfolios. (6L)

Diversification and portofolio risk, Markovitz portfolio selection, optimal portfolios. (6L)

Capital assets, pricing model, passive strategy, risk premium, Index models and diversification, CAPM and index model. (6L)

Options markets, american and european options, call and put options, open strategies, option like instruments, option valuation. Binomial option pricing, Black-Scholes option valuation, uses of Black-Scholes formula. Futures markets, Mechanics and strategies, Futures prices, expected spot prices. (15L)

(45L + 15T)

REFERENCES

Bodie, Z., Kane, A., and Marcus, A.J.(1996), Investments 4th Edition, Irwin.

Chapters : 1, 2, 4, 5, 6, 7, 8, 9, 10, 20, 21, 22.

ADDITIONAL REFERENCES

Arrow, K. J. (1971), Essays in the Theory of Risk Bearing, North Holland.

Hull John C. (1993), Options, Futures and Other Derivative Securities. 2nd Ed. Prentice Hall.

M.6.5 : OPTIMIZATION IN ECONOMICS

Production Theory : Production function, diminishing returns, so product curves and production frontiers, Cobb – Douglas production functions. Elasticity of substitutes. Homogeneous and homothetic production functions and their properties. (6L)

Long run and short run average costs. Linear programming for production theory. Indifference curves and their construction and properties, graphic and alternative solutions of the programming problems. (8L)

Comparative statics, endogeneous variables, second order conditions, multivariable models, Slutsky theorem in n variables. Revealed preferences and index numbers. Expenditure theorem, compensated demand function. Cost, revenue and profit functions. (6L)

Theory of the firm : Objectives of the firm, profit maximization, multiple products and inputs, sales maximization, choice of input and output combinations. Market structures, profit maximizing competitive firm, pure monopoly and monopolistic competition, Bilateral and oligopolistic monopoly. reaction curves, oligopolistic pricing. (10L)

Equilibrium and Welfare . Interdependence in the economy, equations of general equilibrium, Walras law, general equilibrium analysis, optimal cash balances, resource allocation, consumer's and producer's surplus. Pareto optimality, optimal use of resources in producing given outputs, optimal price system, breakeven constraints, welfare theory, welfare judgements. (15L)

(45L + 15T)

REFERENCES

W. J. Baumol (1996), Economic Theory and Operations Analysis, 4th Ed. Prentice Hall of India, Chapters 11, 12, 13, 14, 15, 16, 20, 21.

C. Birchenhall and P. Grout (1987). Mathematics for Modern Economics, Heritage.

A. K. Dixit, ( ) Optimization in Economic Theory, Oxford University Press.

A. K. Sen (1970). Collective choice and social welfare, Holden Day.

M.6.6 : ACTUARIAL STATISTICS

Section I - Probability Models and Life Tables

Utility theory, insurance and utility theory, models for individual claims and their sums, survival function, curtate future lifetime, force of mortality. (3L)

Life table and its relation with survival function, examples, assumptions for fractional ages, some analytical laws of mortality, select and ultimate tables. (6L)

Multiple life functions, joint life and last survivor status, insurance and annuity benefits through multiple life functions evaluation for special mortality laws. (6L)

Multiple decrement models, deterministic and random survivorship groups, associated single decrement tables, central rates of multiple decrement, net single premiums and their numerical evaluations. (3L)

Distribution of aggregate claims, compound Poisson distribution and its applications. Distribution of aggregate claims, compound Poisson distribution and its applications. (3L)

Section II - Insurance and Annuities

Principles of compound interest : Nominal and effective rates of interest and discount, force of interest and discount, compound interest, accumulation factor, continuous compounding.

Life insurance : Insurance payable at the moment of death and at the end of the year of death-level benefit insurance, endowment insurance, diferred insurance and varying benefit insurance, recursions, commutation functions. (6L)

Life annuities : Single payment, continuous life annuities, discrete life annuities, life annuities with monthly payments, commutation functions, varying annuities, recursions, complete annuities-immediate and apportionable annuities-due. (6L)

Net premiums : Continuous and discrete premiums, true monthly payment premiums, apportionable premiums, commutation functions, accumulation type benefits.

Payment premiums, apportionable premiums, commutation functions, accumulation type benefits. (6L)

Net premium reserves : Continuous and discrete net premium reserve, reserves on a semicontinuous basis, reserves based on true monthly premiums, reserves on an apportionable or discounted continuous basis, reserves at fractional durations, allocations of loss to policy years, recursive formulas and differential equations for reserves, commutation functions.

Some practical considerations : Premiums that include expenses-general expenses types of expenses, per policy expenses.

Claim amount distributions, approximating the individual model, stop-loss insurance. (6L)

(45L + 15T)

REFERENCES

N.L. Bowers, H.U. Gerber, J.C. Hickman, D.A. Jones and C.J. Nesbitt, (1986), 'Actuarial Mathematics', Society of Actuarias, Ithaca, Illiois, U.S.A. Second Edition (1997)

Section I - Chapters : 1, 2, 3, 8, 9, 11.

Section II - Chapters : 4, 5, 6, 7, 13, 14.

ADDITIONAL REFERENCES

Spurgeon E.T. (1972), Life Contingencies, Cambridge University Press.

Neill, A. (1977). Life Contingencies, Heineman.

M.6.7 : OFFICIAL STATISTICS

Introduction to Indian and International statistical systems. Role, function and activities of Central and State statistical organizations. Organization of large scale sample surveys. Role of National Sample Survey Organization. General and special data dissemination systems. (12L)

Population growth in developed and developing countries, evaluation of performance of family welfare programmes, projections of labour force and manpower. Scope and content of population census of India. (12L)

System of collection of Agricultural Statistics. Crop forecasting and estimation, productivity, fragmentation of holdings, support prices, buffer stocks, impact of irrigation projects. (11L)

Statistics related to industries, foreign trade, balance of payment, cost of living, inflation, educational and other social statistics. (10L)

(45L)

REFERENCES

Basic Statistics Relating to the Indian Economy (CSO) 1990.

Guide to Official Statistics (CSO) 1999.

Statistical System in India (CSO) 1995.

Principles and accommodation of National Population Censuses, UNESCO.

Panse, V. G., Estimation of Crop Yields (FAO).

Family Welfare Yearbook. Annual Publication of D/o Family Welfare.

Monthly Statistics of Foreign Trade in India, DGCIS, Calcutta and other Govt. Publications.

UGC MODEL CURRICULUM
STATISTICS

UNIVERSITY GRANTS COMMISSION
New Delhi

CONTENTS