University of Pune

Two Year M. A. / M. Sc. Degree Program in Statistics

Revised Syllabi of Second Year M. A. / M. Sc. in Statistics

(Choice Based Credit System)

(To be implemented in the affiliated colleges of University of Pune)

(With effect from Academic Year 2014-2015)

Submitted by

Prof. S. R. Deshmukh
Chairperson,
Board of Studies in Statistics

1) **Title of the Program:** M. A. / M. Sc. in Statistics

2) **Preamble to the syllabus:** M. A. / M. Sc. Statistics program is of minimum 100 credits spread over four semesters. This program is offered at the colleges affiliated to the University of Pune. The program emphasizes both theory and applications of statistics and is structured to provide knowledge and skills in depth necessary for the employability of students in industry, other organizations, as well as in academics. Accordingly, the program has important features such as individual/ group projects, elective courses and courses on standard software packages such as MATLAB, MINITAB, SYSTAT, SPSS, R. Syllabus of the first two semesters covers core courses. The second year syllabus contains core, elective and open courses. It is possible for the students to study basic courses from other disciplines such as economics, life sciences, computer science, and mathematics in place of electives.

3) **Introduction:** (a) M. A. / M. Sc. Statistics program will be conducted under credit system in four semesters. There will be approximately 25 credits in each semester for a total of 100 credits. Each course is given credit values between 1 and 4 depending on the expected study load of the student. One credit is taken to be equivalent to 15 clock hours of study load.

(b) The program consists of core courses which may be hard-core (compulsory) or soft-core (Elective).

(c) Some courses are termed Open Courses (O). The open courses are those offered by other departments but relevant to M. A. /M. Sc. Statistics program.

(d) In addition, there are Lab courses and Project courses.

(e) For every course, there will be Continuous Internal Assessment (CIA) conducted by department or college and End of the Semester Examination (ESE) will be conducted by the University at the end of semester.

4) **Eligibility:** As per latest university rules.

5) **Examination:**

A) (i) **Pattern of examination**: There would be Continuous Internal Assessment (CIA) or in semester assessment and an End of Semester Examination (ESE) for each course.CIA includes written examinations, assignments, small projects, viva-voce examinations, presentations seminars, quizzes etc.

(ii) **Pattern of the question paper at End of Semester Examination (ESE)**:

Duration of ESE will be 45 minutes per credit having 20 marks per credit.

B) **Standard of passing**:

(i) For passing the course, student has to score at least 40% marks in ESE and CIA combined. Moreover student has to score minimum 30% marks in CIA as well as in ESE separately.

(ii)The grade for the course is declared on the basis of combined marks in CIA and ESE with a weightage of 50% marks in CIA and 50% marks in ESE.

(iii) If student fails in the course, he has to reappear for ESE. However his performance in CIA will be carried forward.

C) ATKT rules:

(i) A student can take admission to the third semester, if she/he completes 50% credits of the total credits expected to be completed within first two semesters.

(ii) If a student fails in a course, he/ she can reappear for the CIA by registering for the said courses during 5^{th} , 6^{th} , 7^{th} or 8^{th} semester. Otherwise the existing performance at CIA will be carried forward.

D) Award of class: As per the University rules.

E) External students: Not applicable

F) Setting of question paper: As per the University rules

G) Verification or revaluation and photo copy of answer book: As per the University rules.

H) Grade Improvement: For grade improvement minimum 30 credit courses should be taken by the student from parent department. Grade Improvement Program will be implemented at the end of the academic year. A student can opt for the grade improvement program only after declaration of result of final semester examination (i.e. at the end of the next academic year after passing the M.A./ M. Sc. Examination) within two years of completion of M.A./ M. Sc. and only once.

6) Structure of the program: Following is a structure of the two year M.Sc./M.A. (Statistics) program.

T: Theory P: Practical O: Open C: Compulsory E: Elective

Semester I

Course Code	T/ P	O/C/ E	Title	No. of credits	ESE Duration	Marks Assigned
ST - 11	**T**	**C**	Mathematical Analysis	**4**	3 Hours	80
ST - 12	**T**	**C**	Integral Calculus and statistical computing	**2**	1Hour 30 minutes	40
ST - 13	**T**	**C**	Linear Algebra	**4**	3 Hours	80
ST - 14	**T**	**C**	Probability Distribution I	**2**	1Hour 30 minutes	40
ST - 15	**T**	**C**	Probability Distribution II	**3**	2Hours 15 minutes	60
ST - 16	**T**	**C**	Sampling Theory	**4**	3 Hours	80
ST - 17	**P**	**C**	Practicals -I	**4**	3 Hours	80
ST - 18	**P**	**C**	Practicals -II	**2**	1Hour 30 minutes	40
			Total	**25**		500
Semester II						
ST - 21	**T**	**C**	Probability Theory	**3**	2Hours 15 minutes	60
ST - 22	**T**	**C**	Limit theorems and Convergences	**3**	2Hours 15 minutes	60
ST - 23	**T**	**C**	Regression Analysis	**4**	3 Hours	80
ST - 24	**T**	**C**	Parametric Inference (Estimation)	**4**	3 Hours	80
***ST - 25**	**T**	**C**	Testing of hypothesis	**1**	2Hours 15 minutes	20
***ST - 26**	**T**	**C**	Exploratory Multivariate Analysis *** ESE for ST 25 and ST 26 will be held together.**	**2**		40
ST - 27	**T**	**C**	Inference in Multivariate Analysis	**3**	2Hours 15 minutes	60
ST - 28	**P**	**C**	Practicals -III	**4**	3 Hours	80
			Total	**24**		480

Semester III

Course Code	T/ P	O/C/ E	Title	No. of credits	ESE Duration	Marks Assigned
ST - 31	**T**	**C**	Markov Chains	**3**	2Hours 15 minutes	60
ST - 32	**T**	**C**	Design and analysis of Experiments	**4**	3 Hours	80
ST - 33	**T**	**C**	Asymptotic Inference	**4**	3 Hours	80
ST - 34	**T**	**C**	Statistical Process Control	**3**	2Hours 15 minutes	60
ST - 35	**T**	**C**	Practicals IV	**4**	3 Hours	80
2 Elective courses				**4 credits each**	3 Hours	80
List of elective courses: (Any two courses are to be chosen for semester III)						
ST (E)36	**T**	**E/O**	Data Mining	**4**	3 Hours	80
ST (E)37	**T**	**E**	Bayesian Inference	**4**	3 Hours	80
ST (E)38	**T**	**E/O**	Optimization Techniques	**4**	3 Hours	80
ST (E)39	**T**	**E/O**	Discrete Event System Simulation	**4**	3 Hours	80
ST (E)40	**T**	**E**	Statistical Methods in Microarray Data Analysis	**4**	3 Hours	80
			Total	**26**		520
Semester IV						
ST - 41	**T**	**C**	Stochastic Processes	**3**	2Hours 15 minutes	60
ST - 42	**T**	**C**	Time Series Analysis	**4**	3 Hours	80
ST - 43	**T**	**C**	Actuarial Statistics	**4**	3 Hours	80
ST - 44	**T**	**C**	Survival Analysis	**4**	3 Hours	80
ST - 45	**T**	**C**	Practicals V	**4**	3 Hours	80
ST - 46			Project *	**2**	1 Hours 30 minutes	40
Elective course				**4**	3 Hours	80
List of elective courses: (Any one course is to be chosen for semester IV)						
ST (E) 47	**T**	**E**	Non parametric Statistical	**4**	3 Hours	80

			Inference			
ST (E) 48	**T**	**E/O**	Statistical Analysis of Clinical Trials	**4**	3 Hours	80
ST (E) 49	**T**	**E/O**	Statistics Education and Research Methodology	**4**	3 Hours	80
			Total	**25**		480

Note: Board of studies in statistics has discretion to update list of courses and syllabi.

7) Notes for implementation of the program

a. Number of lectures per credit is 15 , (13 Theory + 2 for Tutorials, seminars, tests etc.)

b. Each credit of practical course will be allotted 1hour 30 minutes laboratory work per week.

c. Nature of CIA of practical course :

i) Journal day to day completion 25 % marks.

ii) Viva – voce day to day experiments. 25% marks.

iii) Internal tests 50% marks.

d. Nature of ESE of practical course : ESE will be of 45 minutes duration per credit having 20 marks per credit of which 10% marks will be reserved for viva –voce at the time of examination

e. There should not be more than 10 students in a batch for M.Sc. practical course.

f. At least two interactive sessions per course per semester be conducted by concerned teachers.

g. In order to acquaint the students with applications of statistical methods in various fields such as industries, agricultural sectors, government institutes etc. study tour can be arranged.

8) Project: Project can be done in a group not more than 3 students.

Internal Evaluation: Guide(Supervisor) will work as a internal examiner.
Project internal a assessment is carry 40 marks.
Assessment will be done on followings aspects.

i) Timely Submission in the proper proforma which includes (Title, abstract, Key words, Methodology, conclusion, references, limitations and source of data etc.) **(20 marks)**

ii) Contents and understanding of the topic of dissertation **(20 marks)**

ESE evaluation : External examiner will evaluate project along with guide for 40 marks. Assessment will be done on followings aspects.

i) Presentation	i) Viva (Contents and Methodology ,	(**20 marks)**
	ii) Understanding of the subject matter of project).	**(20 marks)**

9) **Detailed Syllabus:** Detailed syllabus along with the list of recommended books for semester III and IV of M.Sc. / M. A. (Statistics) is given below:

ST- 11 (TC):Mathematical Analysis (from2013-14) 4 credits

Unit 1:

Set of real numbers, supremum and infimum of sets of real numbers, real field, existence theorem of ordered field R (with proof). countable and uncountable sets, countability of rational and unaccountability of the real numbers, Metric Spaces, Interior point, Exterior point, Boundary point, limit point of a set, open set, closed set, and compact sets. Bolzano-Weierstrass and Heine-Borel Theorem (with proof). Applications of these theorems.

[15L]

Unit 2:

Sequences of real numbers, convergence and divergence of sequences and subsequences,Cauchy sequence, Limit inferior, limit superior of the sequences, some special sequences.

[15L]

Unit 3:

Series of real numbers, series of non negative terms, convergence of series, tests of convergence of series (root test, ratio test with proof) absolute convergence, Power series and applications to power series, addition and multiplication of the series, radius of convergence of Binomial, exponential, geometric and log series. Differentiation and integration of power series.

[15 L]

Unit 4:

The derivative of a real function, mean value theorem, the continuity of derivatives, L'Hospital's Rule, Taylor's Theorem. Cauchy- Schwartz inequality, fundamental theorem of a calculus, integration by parts, differentiation under the sign of integral.

[15 L]

Books Recommended:

1. Apostol T. M. (1975) Mathematical Analysis: A Modern Approach to Advanced Calculus. (Addison - Wesley)
2. Bartle R. G. (1976), Elements of Real Analysis (Wiley)
3. Bartle G.R. & Sherbert D. R. (2000): Introduction to Real Analysis- John Wiley & Son Inc.
4. Goldberg R. R. (1964) : Methods of Real Analysis- Blaisdell Publishing company, New york, U.S.A.
5. Malik S. C. & Arora S. (1991), Mathematical Analysis (Wiley Eastern Limited 2nd edition)
6. Royden (1988), Principles of Real Analysis (Macmillian)
7. Rudin W.(1985), Principles of Mathematical Analysis (McGraw – Hill)
8. Trench : ( ) Real Analysis

ST- 12(TC): Integral Calculus and Statistical computing (from2013-14) 2 credits

Unit 1: Riemann and Riemann- Stieltjes Integral, applications in Statistics,

[10 L]

Unit2:Improper integrals of first and second kind for one variable, conditions for convergenceof beta and Gamma functions , relation between beta and gamma functions, properties of beta and gamma functions, duplication formula. Implict function Theorem / Inverse function theorem and their simple applications.

[5 L]

Unit 3 : (i)Newton–Raphson method for two or more simultaneous transcendental equations,

(ii) Newton's bivariate interpolation formula,

(iii) Unconstrained optimization : Grid search method, Gradient search : Seepest descent method ,Newton's , method.

(iv) Simpson's , Trapezoidal rule for bivariate integrals;

(v) Simulation : Linear congruential generator ; Monte Carlo method to evaluate single and multiple integrals.

(vi) Jack – Knife estimators.

(vii) Boot-Strap method

[15 L]

Books Recommended:

1. Apostol T. M. (1975) Mathematical Analysis: A Modern Approach to Advanced Calculus. (Addison - Wesley)
2. Bartle R. G. (1976), Elements of Real Analysis (Wiley)
3. Bartle G.R. & Sherbert D. R. (2000): Introduction to Real Analysis- John Wiley & Son Inc
4. Goldberg R. R. (1964) : Methods of Real Analysis- Blaisdell Publishing company, New york, U.S.A.
5. Krishnamurthy and Sen: Numerical Algorithms(East West press Pvt. Ltd.)
6. Malik S. C. & Arora S. (1991), Mathematical Analysis (Wiley Eastern Limited2nd edition)
7. Royden (1988), Principles of Real Analysis (Macmillian)

Rudin W.(1985), Principles of Mathematical Analysis (McGraw – Hill)

8. S.S. Sastry (4th edition ,2009)Prentice Hall : Introductory methods of Numerical Analysis.
10. V. Rajaraman (1993): Computer Oriented numerical methods. Prentice Hall

ST- 13 (TC)Linear Algebra (From 2013-14) 4 credits

Unit 1: Vector space, subspace, linear dependence and independence, basis of vector space, dimension of a vector space, orthogonal and orthonormal vectors, orthonormal basis, Gram- Schmidt orthogonalization Matrix algebra, special types of matrices, orthogonal matrix, idempotent matrix partitioned matrices, elementary operations, rank of a matrix, inverse of a matrix

[15L]

Unit2: Characteristic roots of a matrix, right and left characteristic vectors,

Properties of characteristic roots and vectors, algebraic and geometric multiplicities, spectral decomposition, n^{th} power of a matrix, Cayley-Hamilton theorem.

[15L]

Unit 3: g-inverse, Moore-Penrose g-inverse, solution of a system of homogeneous and non- homogeneous linear equations. Gauss seidal and Gauss Jacobbi iterative methods.

[15L]

Unit 4: Quadratic forms, definition, reduction and classification, simultaneous reduction of two quadratic forms, maxima and minima of ratio of quadratic forms.

[15L]

Books:

1. Graybill (1961): An introduction to linear Statistical models Volume I, Mc Graw Hill
2. Hadley G. (1987) :Linear Algebra, Narosa
3. Searle S.A.(1982): Matrix Algebra Useful for Statistics, Wiley
4. Rao C.R. (1973) :Linear Statistical Inference and its Applications, Wiley Eastern
5. Rao, C.R. and Bhimashankaram, P. (1992) :Linear Algebra,Tata Mc Graw Hill

Additional Reference : http// aix1.uottawa.ca/~ jkhoury/app.htm

ST 14 : (TC) Probability Distributions I (From 2013-14) 2 crdits

Unit 1: Brief review of a random variable, c.d.f and its characteristic properties with proof for univariate and bivariate probability distributions, quantiles, discrete, continuous distributions, p.m.f., p.d.f., symmetric distributions, mixtures of probability distributions, transformation of random variables, m.g.f, moments, p.g.f , compound distribution.

[15L]

Unit 2: Random vectors, joint probability distributions, joint m.g.f., mixed moments, variance covariance matrix, independence, sums of independent r.v.s , convolutions, conditional expectation and variance, regression function and best linear regression function, multiple and partial correlation coefficient

[15L]

1. BLS Prakash Rao ; First course in probability
2. Casella and Burger(2002) Statistical Inference (Duxbury advanced series II edition)
3. Hogg RV and Craig T T (1978) Introduction to mathematical Statistics 4th edn. (Collier McMillan)
4. Johnson N.L. & Kotz S.(1996) Distributions in statistics Vol.I .VolI and Vol III John Wiley and sons Inc.)
5. Johnson N.L., Kotz S., Balkrishnan, N. Multivariate Distributions (John Wiley and sons)
6. Rohatagi V.K. & Saleh A.K.(2001) Introduction to probability theory and mathematical statistics. (John Wiley and sons)

ST 15 (TC): Probability Distributions II(From 2013-14) 3 crdits

Unit 1: Bivariate Poisson, bivariate exponential (all 4 types), bivariate normal distribution, Dirichilet distribution and their properties

[15L]

Unit 2 : a)Exponential family of distributions, location and scale families, non-regular families

[7L]

b) Review of order statistics, joint distribution of all order of statistics, probability integral transformation,distribution of rank factors, distributions of sign statistic, Kolmogorov-Smirnov statistic and Wilcoxon sign rank statistic

[8L]

Unit 3: Sampling distributions of quadratic forms and linear forms for random samples from normal distribution, Fisher Cochran theorem, Non-central chi-square, t, F distribution

[15L]

Books :

1. BLS Prakash Rao ; First course in probability

2. Casella and Burger(2002) Statistical Inference (Duxbury advanced series II edition)
3. Hogg RV and Craig T T (1978) Introduction to mathematical Statistics 4th edn. (Collier McMillan)
4. Johnson N.L. & Kotz S.(1996) Distributions in statistics Vol.I .Vol I and Vol III (John Wiley and sons Inc.)

5. Johnson N.L., Kotz S., Balkrishnan, N. Multivariate Distributions (John Wiley and sons)
6. Rohatagi V.K. & Saleh A.K.(2001) Introduction to probability theory and mathematical statistics. (John Wiley and sons)

ST-16: (TC)Sampling Theory (From 2013-14) 4 crdits

Unit 1: Basic finite population techniques SRSWR, SRSWOR Inclusion probabilities, related results on estimation of population total, Determination of sample size in various aspects. Probability Proportional to Size With Replacement (PPSWR) methods, cumulative total method and Lahiri's method for estimation problem, estimation of finite population mean and total.

[15L]

Unit 2: Horwitz – Thompson estimator, its variance and properties, midzuno scheme of sampling. Stratified sampling, comparison of allocation problem of allocation in stratified sampling, construction of strata, deep stratification.

[15L]

Unit 3: Use of supplementary information for estimation, ratio and regression estimators using separate strata and combined strata, unbiased and almost unbiased ratio type estimators of population mean post stratification, variance of estimator of population mean under it.

Systematic sampling, sample mean and its variance, circular systematic sampling, two dimensional systematic sampling , comparison of systematic sampling with random sampling and stratified sampling with random sampling and stratified sampling.

[15L]

Unit 4 : Cluster sampling with cluster sampling with clusters of equal sizes and unequal sizes , estimation of population mean and its standard error, two stage sampling with equal first stage units, expected value and the variance of sample mean. Sampling and non – sampling errors, Response errors, mathematical model for Response errors, Hansen Horwitz technique, Randomized Response Technique (RRT). Warner's randomized response technique. **[15L]**

Books:

1. Des Raj & Chandhok P.(1998), Sample survey theory (Narosa)
2. Murthy M.N.(1977) Sampling theory and methods (Statistical Publishing Society)
3. Sukhatme P.V. Sukhatme B.V. and C. Ashok Sampling theory of survey and applications (Indian society for Agricultural statistics)
4. W. G. Cochran ,(1977) Sampling techniques (John Wiley and sons)

ST-17(PC) Practical s - I (From 2013-14) 4 credits (6 hours a week)

1. Introduction to Statistical Software-I
2. Introduction to Statistical Software-II
3. Matrices
   Contents: Properties of Matrices, Inverse of matrix and non – singular matrix, inverse by partitioning..
4. G-Inverse
5. MPG-Inverse
6. Eigen Value, eigen Vectors, Spectral Decomposition, Power of matrix.

7. Solution of System of Linear Equations using Gauss elimination , Gauss Jorden, Gauss Seidal and Gauss Jacobbi methods .
8. Classification and Reduction of Quadratic forms.
9. Plotting of density function and distribution functions.
10. Model sampling from Gamma , Chi-square ,Weibull ,lognormal probability distribution.
11. Model sampling from mixture of distribution
12. Model sampling from bivariate probability distribution
13. Computation of probability of events related to bivariate probability distribution Computation of probability of non-central χ^2, t, F-distributions
14. PPS sampling
15. Stratified sampling(using Ratio and Regression) , Ratio and Regression estimates
16. Circular Systematic Sampling
17. Cluster Sampling with equal cluster size
18. Cluster Sampling with unequal cluster size
19. Two stage sampling

ST-18(PC) Practical s II (From 2013-14) 2 credits (3 hours a week)

1. Computations of Summary Statistics using R.
2. Model sampling from density function and distribution function using R.
3. Simultaneous Transcendental equations N- R method.
4. Grid search , steepest descent and Newton's Method of optimization.
5. Bivariate interpolation.
6. Computations of double integral
7. Numerical integration using simulations.
8. Computation of integral by Riemann and Riemenn – Stiltjes sums.
9. Boots Trap method
10. Jack knife method

ST 21 : (TC) Probability Theory (From 2013-14) 3 credits

Unit 1: Review of algebra of sets, sequence of sets , limsup, liminf and limit of a sequence of sets, field, sigma field, minimal sigma field, Borel fields, measurable space , monotone classes. Probability measure on a measurable space, probability space, properties of probability measure: continuity, mixture of probability measures, Lebesgue and Lebesgue-Steltjes measures.

[**15L]**

Unit 2: Measurable function, Real and Vector valued random variables, simple r.v., r.v. as a limit of sequence of simple r.v.s, discrete and continuous type r.v., distribution function, decomposition of a distribution function

[15L]

Unit 3: Integration of a measurable function with respect to a probability measure, expectation of a r.v., properties of expectation, characteristic function and properties, Parseval relation, uniqueness theorem, inequalities of moments

[15L]

Books :

(1) Bhat B.R.(1985) Modern Probability theory (Wiley Eastern)
(2) Breiman : Probability Theory
(3) Billingsley P. (1986) Probability and Measure (Wiley)
(4) Feller W. (1969) Introduction to probability and its applications Vol I and Vol.II (Wiley Eastern)
(5) K. L. Chung ( ):Probability Theory.

ST 22 (TC) : Limit theorems and Convergence s : 3 Credits

Unit 1: Convergence of a sequence of r.v.s,

a) convergence in probability
b) convergence in distribution
c) convergence in r^{th} mean ,
d) almost sure convergence, their inter-relations
Slutkey's Theorem

(15L)

Unit 2: Independence of two and n (>2) events, sequence of independent events, π and λ systems, Dynkin's theorem (Introduction) independence of r.v.s ,Kolmogorov, zero-one law, Borel Cantelli lemma

(15L)

Unit 3: Law of Large Numbers : Weak Law of Large Numbers (WLLN), Khintchin's WLLN, Strong Law of Large Numbers (SLLN) (Statement only), Central Limit Theorem (CLT)

Levy continuity theorem, CLT for i.i.d. r.v.s, Liaponove's form, Lindeberg Feller form and their applications

(15L)

Books :

(1) Bhat B.R.(1985) Modern Probability theory (Wiley Eastern)
(2) Breiman : Probability Theory
(3) Billingsley P. (1986) Probability and Measure (Wiley)
(4) Feller W. (1969) Introduction to probability and its applications Vol I and Vol.II (Wiley Eastern)
(5) K. L. Chung: Probability theory.

ST-23 : (TC) Regression Analysis (From 2013-14) 4credits

Unit 1 Simple linear regression, assumptions, least square (LS) estimators of parameters, standard error . of estimators, testing of hypothesis for coefficient of regression, s.e. of prediction, testing of hypotheses about parallelism (Slopes) ,equality of intercepts, congruence, extrapolation, optimal choice of independent variables,diagnostic checks and correction: graphical technique, tests for normality , uncorrelated ness, homo scadasticity, lack of fit. modifications like polynomial regression, transformations on of dependent or independent variables , weighted LS, inverse regression.

[15L]

Unit 2 Multiple regression: Standard Gauss-Markov setup, least square estimation, error and estimation spaces, variance and covariance of LS estimators, properties of LS estimators, estimation of error variance, case with correlated observation, LS estimation with restriction on parameters, simultaneous estimation of linear parametric functions, testing of hypothesis for one and more than one linear parametric functions, confidence intervals and regions.Mallows Cp, forward, backward selection method.

[15L]

Unit 3: a) Multicollinearity : consequences, detection and remedies, autocorrelation consequences, Durbin Watson test, estimation of parameters in autocorrelation.

[4L]

b) Multiple correlation, adjusted multiple correlation coefficient, null distribution of simple correlation and multiple correlation coefficient, partial correlation coefficient and its relation with multiple correlation coefficient, test for significance of simple ,multiple and partial correlation coefficients, variable selection procedures.
Residual and residual diagnostics, transformation of variables: Box-Cox power Transformation, generalized weighted least sequence.

[11L]

Unit 4: a) Non-linear regression: Non-linear least squares transformation to a linear model, statistical inference in non-linear regression

[5L]

b) Logistic regression: Logit transform, ML estimation, tests of hypothesis, Wald test, LR test, score test, test for overall regression, introduction to link functions such as binomial, inverse binomial, inverse Gaussian and Gamma.

[7L]

c) Generalized linear model :

[3L]

Books:

1. Draper, N. R. and Smith H. (1998) Applied regression analysis 3rd edition (John Wiley)
2. Hosmer, D. W. and Lemeshow, S. (1989) Applied logistic regression (John Wiley)
3. McCullagh, P. and Nelder, J. A.(1989) Generalized linear models (Chapman and Hall)
4. Montogomery D.C. et. al.(2003) Introduction to linear regression analysis (Wiley Eastern)
5. Neter, J.; Wasserman, W. and Kutner, M.H.(1985) Applied linear statistical models
6. Ratkowsky, D. A.(1983) Nonlinear regression modeling (Marcel Dekker)

ST 24: (TC) Parametric Inference Estimation (From 2013-14) 4credits

Unit 1: Sufficiency :- Factorization theorem, joint sufficiency , likelihood equivalence , minimal sufficiency, construction of minimal sufficient statistics, Special classes of distribution, admitting minimal sufficient statistics.

[15L]

Unit 2: Fisher information & information matrix Completeness, bounded completeness, complete sufficient statistics, special classes of distribution admitting complete sufficient statistics.

[15L]

Unit 3:- Complete minimal sufficient statistics, ancillary statistics, Basu's theorem & its application unbiased estimator, UMVUE, n & s condition for existence of UMVUE (with proof), Rao- Blackwell theorem, Lehman- scheffe theorem, & there uses, C-R inequality, Regularity conditions , MVBUE , Chapman robin's bound, Bhattacharya bound (withproof).

[15L]

Unit 4:- Confidence interval, relation with testing of hypothesis , SELCI, UMACI. introduction to Bayesian estimation, prior & posterior distribution, loss function, principle of minimum expected posterior loss, quadratic & other common loss functions, conjugate family of prior distribution & its examples.

[15L]

Books:

1. Casella G. & Beregar R.L. (2002) Statistical Inference, 2nd Edition (Duxbury Advanced Series)
2. Dudewitz E.J. & Mishra S.N.(1988) Modern Mathematical Statistics (John Wiley)
3. Kale B.K. (1999) A First course on Parametric Inference (Narosa)
4. Lehman E.L (1988) Theory of point estimation (John Wiley)
5. Lehman E.L(1986) Testing of Statistical hypotheses (John Wiley)
6. Rohatagi V.K. (1976) Introduction to theory of probability & mathematical statistics (John Wiley & sons)

ST 25 : (TC) Testing of Hypothesis (From 2013-14) 1 credits

Unit 1:- Test function, NP lemma (with proof) for test function, UMP test for one-sided alternative for one parameter, exponential class of densities & extension to the distributions having MLR property. UMPU test.

[15L]

Books:

1. Casella G. & Beregar R.L. (2002) Statistical Inference, 2nd Edition (Duxbury Advanced Series)
2. Dudewitz E.J. & Mishra S.N.(1988) Modern Mathematical Statistics (John Wiley)
3. Kale B.K. (1999) A First course on Parametric Inference (Narosa)
4. Lehman E.L (1988) Theory of point estimation (John Wiley)
5. Lehman E.L(1986) Testing of Statistical hypotheses (John Wiley)
6. Rohatagi V.K. (1976) Introduction to theory of probability & mathematical statistics (John Wiley & sons)

ST-26 : (TC) Exploratory Multivariate Data Analysis (From 2013-14) 2credits

Unit 1 a) Exploratory Multivaritate Data Analysis: Sample mean vector, dispersion matrix, Correlation matrix.
Linear Transformation and its Mean, Variance. Covariances, Correlation between linear transformations. Graphical representation.

b) Cluster analysis

(15L)

Unit 2. Principal component analysis, Factor analysis, Canonical correlation, with applications.

(15L)

Books:

1. Anderson T.W.(1984) Introduction to multivariate analysis (John Wiley)

2. C.R.Rao (1985 Linear Statistical inference and its applications (Wiley Eastern Ltd)

3. Hardle, W. K. & Simar, L. (2012) , Applied Multivariate Statistical analysis (Springer, New York)
4. Johnson R.A. and Wichern D.W.(1988)Applied multivariate statistical analysis (Prentice hall Inc.)
5. Johnson R.A. & Wichern, D.W. (1988). Applied Multivariate Statistical analysis (Prentice Hall Inc.)

6. Kshirsagar A.M. (1983) Multivariate Analysis(Marcel Dekker.)

7. K.C. Bhuyan (2005) Multivariate Analysis and its application, New Central book agency,

LTD. Kolkatta

8. Morrison, D.F.(1990). Multivariate Statistical Methods (McGraw Hill Co.) (3rd ed.)

ST-27 : <u>(TC) Inference in Multivariate analysis (from 2013-14) (3 credits)</u>

Unit 1. Multivariate normal distribution, singular and non -singular normal distributions, m.g.f., Characteristic function, moments, distribution of a linear form and a quadratic form of normal variables, Marginal and conditional distribution. Tests for multivariate normality, Test of significance for multiple and partial correlation coefficients.

(15L)

Unit 2. M.L.E.S. of parameters of multivariate normal distribution and their sampling distribution. Wishart matrix, Wishart distribution and its properties. Tests of hypothesis about mean vector of a multivariate normal population. Hotelling T^2 statistic and its distribution, its applications. confidence region for mean vector of multivariate normal Distribution.

(15L)

Unit 3. Likelihood ratio test. Test for equality of dispersion matrices , Discriminant analysis, Mahalanobis D^2 Statistic, test for significance of the coefficients in discriminant function. Misclassification error.

(15L)

Books:

1. Anderson T.W.(1984) Introduction to multivariate analysis (John Wiley)

2. C.R.Rao (1985 Linear Statistical inference and its applications (Wiley Eastern Ltd)

3. Hardle, W. K. & Simar, L. (2012) , Applied Multivariate Statistical analysis (Springer,New york)

4. Johnson R.A. and Wichern D.W.(1988)Applied multivariate statistical analysis (Prentice hall Inc.)
5. Kshirsagar A.M. (1983) Multivariate Analysis(Marcel Dekker.)

6. K.C. Bhuyan (2005) Multivariate Analysis and its application, New Central book agency,
LTD. Kolkatta
7. Morrison, D.F.(1990). Multivariate Statistical Methods (McGraw Hill Co.) (3rd ed.)

ST 28: PRACTICAL PAPER III (Departmental Course) 4 credits

1 Simple regression and regression diagnostic I
2 Multiple regression
3 Lack of fit of the regression model
4 Multiple regression (selection of variable)
5 Nonlinear regression model
6 Multicollinearity and orthogonal polynomial regression.
7 Logistic regression I
8 Generalized Linear Model and Poisson regression.
9 Application of Central Limit Theorem and weak law of large number
10 Exploratory multivariate data analysis.
11 Multivariate analysis (contour plot).
12 Principal component analysis .
13 Factor Analysis.
14 Cluster Analysis.
15 Canonical correlation.
16 Model sampling from multivariate normal distribution. And computation of M.L.E.'s of parameters.
17 Likelihood ratio test.
18 Application of Hotelling T^2 statistics.
19 Discriminant analysis

ST 31(TC) Markov Chains (from 2014-15) (3 credits)

Unit 1:

Stochastic processes, Markov property, Markov chains (MC), finite MC, transition probabilities, initial distribution, illustrations such as random walk, Ehrenfest chain, gambler's ruin chain, queuing chain, birth death chain, branching chain, Chapman Kolmogorov equation, n-step transition probabilities, transition probability matrix (t.p.m.) hitting times, probability of ever return, transient and recurrent states, decomposition of state space, closed set of states, irreducible set of states, irreducible MC, absorption probabilities, martingales, classification of states of birth and death chains, branching chain, queuing chain, random walk, gambler's ruin chain etc.

(15L)

Unit 2:

Elementary properties of stationary distributions, illustrations such as birth and death chains, Ehrenfest chain, particles in box, average number of visits to recurrent state, non null and positive recurrent states, period of state, existence of uniqueness of stationary distributions, reducible chains, illustrations such as queuing chain finite chains, convergence to the stationary distribution. Steady state distribution, ergodic Markov chain. Ergodic theorem .

(15L)

Unit3:

(a)Gambler's ruin problem:

Random walk, random walk with absorbing and reflecting and elastic barrier, classification of states, probability of absorption in persistent class starting from transient state, application to gambler's ruin problem, probability of ruin cases (i)adversary is infinitely rich (ii) stakes are doubled or halved , expected gain, expected duration of the game.

(b) Branching Chain:

BGW branching process, offspring distribution, mean and variance, generating function for probability of ultimate extinction, nth generation size and related recurrence relations.

(15L)

Books Recommended:

1. Adke, S.R., Manjunath, S.M. (1984) An introduction to finite Markov processes (Wiley Eastern)
2. Bhat, B.R. (2000) stochastic models: Analysis and applications (New Age International)
3. Hoel , P.G.,Port, S.C. ,Stone, C.J. (1972) : Introduction to stochastic processes
4. MedhiJ. (1982) Stochastic processes (Wiley Eastern)
5. Ross, S. (2000) Introduction to probability models, 7th edn (Academic Press)
6. Ross, S. (1996) Stochastic processes (John Wiley)
7. Srinivas and Mehta (1976) Stochastic Processes (Tata Mc-Graw Hill)
8. Taylor, H N and Karlin, S. (1984) An introduction to stochastic modeling(Academic Press)

ST 32 (TC) Design of Experiments and Analysis of Experiments(from 2014-15)
(4 Credits)

Unit 1:

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

(15L)

Unit 2:

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

(15L)

Unit 3:

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

systems, multiple responses, concept of rotatable designs.

(15L)

Unit 4

Taguchi methods: concept of loss function, S/N ratio, orthogonal arrays, triangular tables, linear graphs, inner and outer arrays, nested and split plot designs

(15L)

Books Recommended:

1. Dean, A. and Voss, D. (1999). *Design and Analysis of Experiments*, Springer.
2. George E. P. Box, Draper N.R. (1987). *Empirical Model-Building and Response Surfaces*, Wiley.
3. Hicks, C.R., Kenneth V. and Turner, Jr. (1999). *Fundamental Concepts in the Design of Experiments*, Oxford University Press.
4. John P.W.M. (1971). *Linear Models*, Wiley.
5. Kshirsagar A.M. (1983). *Linear Models*, Marcel Dekker
6. Montgomery, D.C. (2001). *Design and Analysis of Experiments*, Wiley.
7. Ogawa J. (1974). *Statistical Theory of the Analysis of Experimental Design*, Marcel Dekker.

ST 33 :(TC) Asymptotic Inference (from 2014-15) (4 Credits)

Unit 1

Consistency : real and vector parameters, choosing between consistent estimators, implications.

consistency and asymptotic normality : real and vector parameters, continuous transformations, invariance of consistency under continuous transformation (delta method), differentiable transformations, generation of CAN estimators using central limit theorem. **(15 L)**

Unit 2

Maximum likelihood estimation, inconsistent MLEs, Cramer-Huzurbazar theorem, extension to vector–valued parameters, special cases such as exponential class of densities and multinomial distribution. **(15 L)**

Unit 3

Asymptotic theory of tests: tests based on MLEs, likelihood ratio tests, asymptotic distribution of log likelihood ratio, Bartlett Correction, Wald Test, Score Test, locally most powerful tests; comparison of tests: asymptotic relative efficiency of a test, Pitman asymptotic relative efficiency (ARE)

(15 L)

Unit 4

Asymptotic confidence intervals: construction and examples, applications to categorical data analysis , Three dimensional contingency tables.

(15 L)

Books Recommended:

1. Fergusson, T.S. (1996), A course in Large Sample Theory, Chapman and Hall.
2. Gupta Anirban Das (2008), Asymptotic Theory of Statistics and Probability, Springer, New York. Chapters 1, 4, 7, 13, 16, 21, 22, 27.
3. Kale B.K. (1999) A first Course in Parametric Inference, Narosa, New Delhi.
4. Lehmann E. L. and Casella G. (1999) Theory of Point Estimation, Springer, New York,
5. Rao C. R.(1995) Linear Statistical Inference and its Applications, Wiley, New York .

ST 34 : (TC)Statistical Process control (SPC) (from 2014-15) (3 credits)

Unit 1

(a)TQM Total quality Management : meaning and dimensions of quality, Quality improvement, Quality Philosophy, Introduction to TQM, six sigma, DMAIC, and other extension of TQM, quality systems, The ISO 9000 and other Quality systems.

(b) Control Chart : Revision of theory of control charts, Concepts of stable industrial processes, Systematic variation, random variation, variation within and between subgroups, estimation of process parameters,.

Equivalence between control chart and testing of hypothesis problem. Choice of control limits Operating characteristic (O C curve) of control chart. Probability of false alarm, probability of catching shift in parameter. Concept of Run length ,probability distribution of run length ,average run length (ARL). Comparison of control chart using ARL,OC curve, criteria of detecting lack of controls (sensitizing rules),patterns on control charts with justification and its effect on Probability of false alarm. $\overline{X}$ -S chart with subgroup size (i)fixed

(ii)variable, probability limits , S^2 chart .Applications of control charts situations other than manufacturing

(c) CUSUM Chart :,

Chart statistic (C_i^+, C_i^-)and chart parameters(k,h), construction and working of tabular CUSUM chart for mean and variance ,Statement of hypotheses. Estimation of shift in mean of process ,Fast initial response or headstart feature, Siegmund's approximation for ARL and determination of chart parameters. CUSUM chart for subgroup size n>1, comparison between Shewhart chart and CUSUM chart V mask procedure.

(15L)

Unit 2

(a) EWMA chart.

Chart statistic its expectation and variance. Choice of chart parameters (λ,L). Construction and working of EWMA chart for mean and variance.EWMA chart for subgroup size n>1,Comparison of Shewhart control charts with CUSUM charts. Simulation of ARL(δ).

(b)Processcapability: Different Process capability and performance indicesC_p,C_{pk}. C_{pm}.Propertis and relation between capability indices.Connection between proportion of defectives(DPPM) and C_p.Interval estimation of mean given C_{pm} ≥1.Estimation and confidence intervals of estimators of C_p..andC_{pk} Testing of hypothesis about C_p .

(c)Other control charts

(i) Synthetic control chart: Confirming run length (CRL) chart for attributes, Synthetic control chart , computations of chart parameters for given ARL(0),Steady state model ,Computations of ARL(δ), ATS(δ)., Comparison of with Shewhart control chart and CUSUM charts.

(ii) Non-parametric control chart.: Concept ,construction of non parametric chart using sign test.

(iii) Control charts for auto correlated observations: Need ,constructions of control chart for residuals after fitting first order auto correlated model.

(15L)

Unit 3: (a) Attribute control charts.: Revision of control charts for attributes, OC curve for P chart and C chart. Determination sample size for P chart by various criteria (i) probability of catching at least 0.5 (ii) to get LCL >0 (iii) To have at least some defectives in sample with given confidence coefficient.

(iv) minimizing ATS (δ) chart and O C Curve, U chart , Demerit control chart for number of defects. Nelsons control chart for low defect counts.

(b)General ideas of economic designing of control charts. Duncan's model for the economic control chart.

(b) Hoteling T^2Chart:Testing multivariate normality, Hotelling T^2multivariate control chart for mean vector when (i) dispersion matrix is (i) known (ii) unknown ARL(0), ARL(δ).Control chart for dispersion matrix when mean vectoris (i) known (ii) unknown.T^2control chart when subgroup size n=1

(c)Acceptance Sampling Plan

(i) Acceptance Sampling Plan for attributes : Equivalence between sampling plans and testing of hypothesis problem. Double and multiple and sequential sampling plans for attributes. Curtailed inspection sampling plan, Operating characteristic functions. AOQLATI, ASN. Continuous Sampling Plans, chain sampling plan. Description of MIL STD and Dodge Roming sampling plans

(ii) Acceptance sampling plan for variable**:** parameters of plan.(i) Critical distance method,(ii) critical proportion method. (15L)

Books Recommended :

1. Bourke P.D. (1991)Detecting shifts in fraction non – confirming using run length chart with 100% inspection. J ournal of Quality Technology 23 (3) 225-230
2. Besterfield, D.H..Besterfield – Michana, c, Besterfield,G.H. Besterfield-Sace, M(2001) Total Quality Management ; Pearson Education(Singapore) Pte. Ltd. India 2nd Edition.
3. Logotheris, N. (1992) Managing Total Quality; Prentic Hall of India.
4. Montgomery ,D.C. (1985) Introduction to Statistical Quality Control (Wiley)
5. Oakland J.S. (1989) Total Quality Management: Butterworth – Heinemann.
6. Raid W. Amin **a** Marion R. Reynolds Jr. b; Bakir Saad c :Nonparametric quality control charts based on the sign statistic :Communications in Statistics - Theory and Methods Vol.34,2005.
7. Wu, Yeu and Spedding (2001) Asymptotic control charts for detecting fraction non confirming increases JQT 33 (1)104-111

ST 35: (PC) Practical paper IV(from 2014-15) (3credits)

1. One way classification. Multiple comparison tests.
2. Two way classification with equal number of observations per cell (with interaction) and unequal number of observations per cell (without interaction).
3. BIBD. (Intra block analysis)
4. Analysis of covariance in one way and two way model.

5. 2^k Factorial Experiments, analysis of single replicate of 2^k factorial experiments Total and partial confounding in 2^k factorial experiments.
6. 2^k factional experiments.. 3^2 factorial experiments
7. Random effect model with one factor, estimation of variance.
8. Fitting first and second order response surface model, central composite design contour, surface plots, canonical analysis of stationery points.
9. Taguchi methods: S/N ratio, orthogonal arrays, triangular tables, linear graphs, inner and outer arrays.

10. Realization of Markov chain when t.p.m. is given and computation of transition probabilities and Stationary distribution of Markov chain.

11. Verification of consistency and asymptotic normality of the estimators
12. Comparing methods of estimation, MSE and sample size considerations

13. Power functions and comparison of tests & confidence intervals (LR, Wald, Rao),
14. Analysis of three dimensional contingency tables.
15. CUSUM, EWMA charts.
16. Synthetic and Hotelling T^2 chart for mean vector.

ST36(TE/ TO): Data Mining (from 2014-15)(4 credits)

Unit1:

Supervised Learning**:** Linear methods for classification, linear discriminant analysis (LDA), logistic regression, Bayes classifier, nearest neighbor classifier.

packages in R for these methods.

(15L)

Unit2:

Neural network (NN), support vector machine (SVM) packages in R for these methods.

(15L)

Unit3 :

Regression and classification trees (CART). Assessment and model selection: Bias-variance trade off, training error rate, AIC, BIC, CIC, DIC (information criterion), cross-validation. Ada boosting.

(15L)

Unit4:

Unsupervised learning: Clustering procedures- k-means, hierarchical, self-organizing map, EM algorithm, feature selection: principal component analysis,

association rules, software packages for these methods.

(15L)

Books Recommended

1. Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). Classification and Regression Trees. (Wadsworth and Brooks/Cole).
2. Daniel T.Larose, (2006). Data Mining Methods and Models. Wile-Interscience.
3. Galit Shmueli, Nitin Patel, Peter Bruce, (2010). Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner , Wiley
4. Hastie T., Tibshirani R. and Friedman J. H., (2003). The Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer.
5. Mitchell Tom, (1997). Machine Learning. McGraw-Hill.
6. Ripley, B.D. (1996)..Pattern Recognition and Neural Networks. (Cambridg University Press).

ST37(TE):Bayesian Inference (from 2014-15) (4 credits)

Unit 1:

Basics on minimaxity: subjective and frequents probability, Bayesian inference, Bayesian estimation , prior distributions, posterior distribution, loss function, principle of minimum expected posterior loss, quadratic and other common loss functions, Advantages of being a Bayesian HPD confidence intervals, testing, credible intervals, prediction of a future observation. (15L)

Unit 2:

Bayesian analysis with subjective prior, robustness and sensitivity, classes of priors, conjugate class, neighborhood class , density ratio class different methods of objective priors : Jeffrey's prior, probability matching prior, conjugate priors and mixtures, posterior robustness: measures and techniques. (15L)

Unit 3:

Model selection and hypothesis testing based on objective probabilities and Bayes' factors, large sample methods: limit of posterior distribution, consistency of posterior distribution, asymptotic normality of posterior distribution. (15L)

Unit 4:

Bayesian Computations : analytic approximation, E- M Algorithm, Monte Carlo sampling, Markov Chain Monte Carlo Methods, Metropolis – Hastings Algorithm, Gibbs sampling, examples, convergence issues (15L)

Reference Books:

1. Albert Jim (2009) Bayesian Computation with R, second edition, Springer, New York
2. Bolstad W. M. (2007) Introduction to Bayesian Statistics 2nd Ed. Wiley, New York
3. Christensen R. Johnson, W. Branscum A. and Hanson T.E. (2011) Bayesian Ideas and data analysis : A introduction for scientist and Statisticians, Chapman and Hall, London
4. Congdon P. (2006) Bayesian Statistical Modeling, Wiley , New York.
5. Ghosh, J.K. Delampady M. and T. Samantha (2006). An Introduction to Bayesian Analysis : Theory and Methods, Springer, New York.
6. Lee P.M. (2004) Bayesian Statistics : An Introduction, Hodder Arnold, New York.
7. Rao C.R. Day D. (2006) Bayesian Thinking, Modeling and Computation, Handbook of Statistics, Vol.25.

ST 38(TE/ TO): Optimization Techniques (from 2014-15) (4 credits)

Unit 1

Linear Programming: Review of simplex algorithm and simplex method, artificial variable technique methods. degeneracy, duality in linear programming, duality theorems, dual simplex method with justification.

Integer linear programming problem: pure and mixed integer programming problem, Gomory's all integer programming method. Fractional cut method- all integer and mixed integer linear programming problem, branch and bound method, dynamic programming, sensitivity. Bellman's optimality principle.

(15L)

Unit 2

Transportation and Assignment Problems: Balance and degeneracy in transportationproblem, transshipment problem, duality theory of testing optimality of solution in transportation problem and transshipment problem, Hungarian method of assignment, maximization, prohibitions and other variations of assignment problems, duality theory of assignment problems.

(15L)

Unit 3

Nonlinear programming: Karush-Kuhn-Tucker conditions, Quadratic programming, Wolfes, Beales and Fletchers algorithms for solving quadratic programming problems. Convex problems, mixed integer models.

(15L)

Unit 4

Networking models: Network flows, maximal flow in the network. Transportation problems, transshipment problems and assignment problems as networking problems. Network scheduling by PERT/CPM Techniques. Resource Analysis in network scheduling. **(15L)**

Books Recommended

1. Bertsekas, D. (1999). *Nonlinear Programming*, 2nd Edn. Athena Scientific.
2. Chong, E. K. P. and Zak, S. (2004). *An Introduction to Optimization*, Wiley.
3. Fletcher, R. (2000). *Practical Methods of Optimization*, Wiley
4. Hadley, G. (1987). *Linear Programming*. Addison-Wesley.
5. Kambo, N.S. (1991). *Mathematical Programming Techniques*. Affiliated East-West press.
6. Panneerselvam, R. (2012). *Operations Research*, 2nd Edn. Prentice Hall of India.
7. Taha, H.A. (1992). *Operations Research*, 5th ed. Macmillan.

ST 39(TE/TO) :Discrete-Event System Simulation (from 2014-15) (4 credits)

Unit1

Basic concepts of systems and examples of systems; Modeling of systems; Types of systems and models; Meaning of simulation; Advantages and disadvantages of simulation; concepts in discrete-event simulation.

(15 L)

Unit 2

Review of standard distributions, stationary and non-stationary poisson processes; Queueing systems- characteristics, long-run measures of performance of queueing systems, steady-state behaviour of infinite-population Markovian models, Steady-state behaviour of finite-population models, Networks of queues.

(15 L)

Unit 3

Definition and properties of random numbers, Generation of pseudo-random numbers, Techniques for generating random numbers, Tests for properties of random numbers, Random variate generation using Inverse-transform, Acceptance-rejection and other methods with applications.

(15 L)

Unit 4

Input-modelling: data collection, identifying the distribution with data, parameter estimation, Goodness-of-fit tests, fitting a non-stationary Poisson process, Selecting input-models without data, Multivariate and time-series input models.

Verification and validation of simulation models, Output-analysis for a single-model, comparison and analysis of alternative system designs.

Applications of simulation: modeling for queueing and inventory systems, exposure to current applications and research on simulation.

(15 L)

Book Recommended:

Banks J., Carson J.S., Nelson B.L., Nicol D.M., and Shahabudeen P. (2007) Discrete-Event System Simulation, Fourth Edition, Pearson PHI.

ST 40 :(TE) Statistical Methods in Microarray Data Analysis (from 2014-15)

(4 credits)

Unit 1

Background of Microarrays and Normalization techniques:Introduction to Biology relevant to microarray experiment. Microarray experimental set up and quantification of information available from microarray experiments. data cleaning, transformation of data.
Between arrays & within array normalization, in particular quantile and LOWESS normalization, stage wise normalization.
Concordance coefficient and its role in normalization.

(15L)

Unit 2

Statistical Inference procedures in comparative experiments:
Inference procedures for single channel microarray data. Application of two sample t – test. Tests for validating assumptions of two sample t-test. Application of Welch test and Wilcoxon rank sum test.
Inference procedures for two channel microarray data. Application of paired t –test. Tests for validating assumptions of paired t-test. Application of Wilcoxon signed rank test.
Inference procedures for comparing more than two types of mRNA samples in single channel or two channel microarray experiments. Application of one way ANOVA F test, one way ANOVA Welch F test, Kruskal-Wallis test, pairwise t-test, pairwise Welch test and pairwise Wilcoxon rank sum test. Strip charts and its role to decide the profile of differentially expressed genes.

(15L)

Unit 3

Multiple hypotheses testing problem and Principal component analysis:
Multiple hypotheses testing problem. Adjustments for multiple hypotheses testing, adjusted p-values.
False discovery rate and its application to microarray data analysis.
Principal component analysis for microarray data, scree plot, plot of scores to display sets of differentially expressed genes. Singular value decomposition of a rectangular matrix and the concept of ballot. Its application to microarray data analysis.

(15L)

Unit 4

Cluster analysis and Logistic regression:
Hierarchical cluster analysis of microarray data to identify groups of genes and outlying genes
K - means cluster analysis of microarray data to identify groups of genes
Application of logistic regression for microarray data. Concept of AIC and BIC and its role to identify marker genes.

(15L)

Note: R software will be heavily used in applications of all the statistical methods to microarray data to identify differentially expressed genes in two or more biological samples.

Books Recommended:

1. Amartunga D. and Cabrera J. (2004). *Exploration and Analysis of DNA Microarray and Protein Array Data*. Wiley.
2. Deshmukh S.R. (2007). *Microarray Data: Statistical Analysis Using R*, Narosa.

3. Draghici, S. (2003). *Data Analysis Tools for DNA Microarrays*, Chapman and Hall/CRC.
4. Dov, S. (2003). *Microarray Bioinformatics*, Cambridge University Press,
5. McLachlan, G.J.; Do, K.A. and Ambroise, C. (2004). *Analyzing Microarray Gene Expression Data*, Wiley.
6. Simon, R.M. ; Korn, E.L. ; McShane, L.M. ; Radmacher, M.D. ; Wright, G.W. and Zhao, y. (2003). *Design and Analysis of DNA Microarray Investigations*. Springer.
7. Speed, T. (2003). *Statistical Analysis of Gene Expression Microarray Data*, Chapman and Hall/CRC.

ST 41(TC): Stochastic Processes(from 2014-15) (3 credits)

Unit 1:

Kolmogorov convergence Theorem, Markov property in continuous time stochastic processes.

Poisson process: Postulates and properties of Poisson process, probability distribution of N(t) the number of occurrences of the event in (0,t], Poisson process and probability distribution of interarrival time, generalizations of Poisson process: (i) pure birth process: Yule Furry process (ii) birth immigration process. (15L)

Unit 2:

(a) Birth and death process: Birth and death process, particular cases: (i) immigration-emigration process, (ii) linear growth process, (iii) linear growth with immigration, (iv) immigration death process, (v) pure death process.

(b) Continuous time Markov chains(CTMC): Chapman Kolmogorov equations, limiting distributions, ergodicity of homogeneous Markov process.

(c) Markov processes with continuous state space: Introduction to Brownian motion and its properties, Wiener process. (15L)

Unit 3:

(a) Renewal process: renewal process in continuous time, renewal **f**unction and renewal density, renewal equation, stopping time: wald's equation, elementary renewal theorem and its applications: (i) Age and block replacement policies, (ii) Replacement on failure and block replacement, renewal theorems (Blackwell's and Smith's): (i) Blackwell's theorems, (ii)

Smith's theorem or Key Renewal theorem, Poisson process as renewal process, alternating or two stage renewal process.

(b) Theory of Queues: Queuing processes, steady state distribution, Little's formula, Birth and death processes in queuing theory, Markovian model, M/M/s model, M/M/s/s: Erlang loss model, GI/M/1 model.

(15L)

Books Recommended:

1. Adke, S.R. and Manjunath, S.M. (1984) An introduction to finite Markov processes (Wiley Eastern)
2. Bhat, B.R. (2000) Stochastic models: Analysis and applications (New Age International)
3. Hoel,Port ,Stone (1972) : Elementary stochastic processes
4. Medhi, J. (1982) Stochastic processes (Wiley Eastern)
5. Ross, S. (2000) Introduction to probability models, 7th edn (Academic Press)
6. Ross, S. (1996) Stochastic processes (John Wiley)
7. Srinivas and Mehta (1976) Stochastic Processes (Tata Mc-Graw Hill)
8. Taylor, H. N. and Karlin, S. (1984) An introduction to stochastic modeling(Academic Press)

ST42: (TC) Time Series Analysis(from 2014-15) (4 credits)

Unit 1:

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

(15L)

Unit 2:

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

(15L)

Unit 3:

Forecasting in time series models, Durbin-Levinson algorithm, innovation algorithm (without proof). Estimation of mean, auto covariance and autocorrelation functions, Yule-Walker estimation, Estimation of ARIMA models parameters, maximum likelihood method, large sample theory (without proofs).

Choice of AR and MA periods, FPE, AIC, BIC, residual analysis and diagnostic checking. Unit-root non stationarity, unit-root tests.

(15 L)

Unit 4:

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

(15 L)

Books Recommended:

1. Brockwell, P.J. and Davis, R. A. *Introduction to Time Series Analysis*, Springer
2. Chatfield, C. (2001). *Time Series Forecasting*, Chapmann & hall, London
3. Fuller, W. A. (1996). *Introduction to Statistical Time Series*, 2nd Ed. John Wiley,
4. Hamilton N. Y. (1994). *Time Series Analysis*. Princeton University press. Princeton
5. Kendall, Sir Maurice and Ord, J. K. (1990). Time Series (Third Edition), Edward Arnold.
6. Lutkepohl, H. and Kratzing, M. (Ed.) (2004). *Applied Time Series Econometrics*, Cambridge University Press, Cambridge
7. Shumway, R. H.and Stoffer D. S. (2010) Time Series Analysis & Its Applications, Springer, New York.
8. Tsay, R. S. (2010). *Analysis of Financial Time*

ST43(TC): Actuarial Statistics(from 2014-15) (4 credits)

Unit 1

Future life time random variable, its distribution function and density function, concept of force of mortality, curtate future life time random variable its probability mass function, deferred probabilities, all these functions in terms of international actuarial notation.

Analytical laws of mortality such as Gompertz' law and Makeham's law, single decrement life table, select and ultimate life table.

(15L)

Unit 2

Concept of compound interest rate, discount factor, present value of the money, nominal rate of interest, force of interest. Assurance contracts with level and varying benefits, such as whole life insurance, term insurance endowment insurance. Means and variances of the present value random variables of the payments under these contracts under the assumption of constant force of interest, when the benefit payments are made at the end of year of eath(discrete set up) or when it is paid at the epoch of death(continuous set up). Actuarial present value of the benefit. Net single premiums

(15L)

Unit 3

Annuity contracts, annuity certain, discrete annuity, m-thly annuity, continuous annuity, deferred annuity, present values and accumulated values of these annuities.
Continuous life annuity, discrete life annuity, such as whole life annuity, temporary life annuity, n-year certain and life annuity, life annuities with mthly payments.
Present value random variables for these annuity payments, their means and variances. Actuarial present value of the annuity.

(15L)

Unit 4

Loss at issue random variable, various principles to decide net premiums for insurance products and annuity schemes defined in unit II and III, fully continuous premiums and fully discrete premiums, True m-thly payment premiums. Extended equivalence principle to decide gross premiums. Concept of reserve, prospective & retrospective approach. Fully continuous reserve.Fully discrete reserve.

(15L)

Books Recommended:

1. Bowers, JR. N.L., Gerber, H.U., Hickman, J.C., Jones, D.A. and Nesbitt, C.J. (1997). *Actuarial Mathematics*, 2nd Edn., The Society of Actuaries.
2. Deshmukh S.R. (2009). *Actuarial Statistics: An Introduction Using R*, Universities Press.
3. Harriett, E.J. and Dani, L. L.(1999). *Principles of Insurance: Life,Health, and Annuities*, 2nd Edn., Life Office Management Association.
4. Neill, Alistair (1977). *Life Contingencies*, The Institute of Actuaries.
5. Palande, P. S., Shah, R. S. and Lunawat, M. L. (2003). *Insurance in India - Changing Policies and Emerging Opportunities*, Response Books.

ST 44(TC) :Survival Analysis (from 2014-15) (4 credits)

Unit1:

Meaning of censoring, concepts of time, order and random censoring(left and right), survival function, density function, hazard function (rate), cumulative hazard rate, mean residual life function, percentile residual life function, Equilibrium distribution function. Exponential distribution & its no ageing properties: Lack of memory property, constant failure rate, Cauchy-function equation, constant mean residual life function ,TTT transform, identity function as a TTT transform

b) Ageing classes - IFR, IFRA, NBU, NBUE, DMRL, HNBUE and their duals, and inter relationship among these classes. Bathtub Failure rate, IFRA closure property, bound on reliability function of an IFRA distribution.

(15L)

Unit 2:

Life distributions - Exponential Gamma, Weibull, Lognormal, Pareto, linear Failure rate, Makeham family, Lehman families (proportional hazard rate family), spacing, normalized spacing and results of an exponential distribution based on normalized spacing.

Parametric inference for complete data:

a) Exponential distribution: Point estimation of parameter of exponential distribution and Fisher information, exact and asymptotic Confidence Intervals for λ, obtaining minimal sufficient and consistent estimator of λ, Graphical method for checking exponetiality of data.

b) Weibull: Obtaining MLE of scale and shape parameter of Weibull distribution and sample information matrix.

c) Gamma: Obtaining MLE of scale and shape parameter of Gamma distribution and sample information matrix.

d) Lognormal: Obtaining MLE of of parameter μ and σ, Confidence Interval for μ and σ

(15L)

Unit3:

Parametric inference for censored data:

1) Type I censoring: Exponential distribution

2) Type II censoring: Exponential, gamma, Lognormal

3) Random censoring: Exponential, Lehman family, Weibull distribution,

Non-Parametric estimation of survival Function

a) For complete data: Non parametric estimator of distribution function and survival function, distribution of imperical survival function, confidence band for survival function (by Using Kolmogorov - Smirnov statistics)

b) For censored data: Actuarial estimator of survival Function, Estimator of variance of actuarial estimator (Greenwoods formula), product limit estimator and its variance, redistribution to right algorithm.

(15L)

Unit 4:

Test for Exponetiality: Estimable function of degree r, Kernel, symmetric Kernel, U- statistic, variance of U- Statistic, one sample U-Statistic theorem, Hollander and Proschan Test, Test for exponetiality against positive ageing based n sample spacing, Analytical test for exponentiality against NBUE, Deshpande's Test, Two sample U-statistic theorem, Wilcoxon and Mann –Whitney test, Gehan's test, Mantel- Haenzel test,

Log rank test, Semi-parametric regression for failure rate - Cox's proportional hazards model with one and several covariates.
(15L)

Books Recommended:

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

2. Deshpande ,J.V, Purohit, S. G.,(2005),Life Time Data :Statistical Models and Methods

3. Elandt - Johnson, R.E., Johnson N.L. (1980) Survival models and Data Analysis, John Wiley and Sons
4. Gross A.J. and Clark, V. A. (1975) Survival Distributions: Reliability Applications in the Biomedical Sciences, John Wiley and Sons.

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

ST 45(PC) Practicals Paper V (from 2014-15) (4 credits)

Practical No.	Title
1	Smoothing the series using various Filters: Other filters, data transfer Box- Cox transformation, differencing, checking stationery and normality after transformation.
2	ACF/PACF, Analysis of series and residuals, residual analysis.
3	Order selection in time series: use of ACF/PACF and ATC, BIC, fitting of AR, MA models (conditional least squares or maximum likelihood).
4	Fitting of ARMA, AIMA and SARIMA models (conditional least squares or maximum likelihood).
5	Forecasting using fitted linear models (recursively), Holt -Winters forecasts construction of forecast intervals.
6	Fitting heteroscedastic models: checking for heteroscedecity from residuals, Arch GARCH modeling.
7	Realization of queuing model M/M/1, M/M/S and estimation of queuing parameters.
8	Realization of G1/M1 model.
9	Realization of Poisson process.
10	Realization of birth and death process.
11	Parametric analysis of complete data.
12	Parametric analysis of censored data.
13	Computation of Actuarial estimator of survival function and PL – estimator.
14 and 15	Practical based on test for exponentiality (2 practicals).
16	Construction of life table.
17	Computation of Net single premiums when benefit is payable (i)at the moment of death, (ii) at the end of year of death.
18	Calculation of present values and accumulated value for different types of annuity.
19	Calculation of premiums when benefit is payable (i) at the moment of death (ii) at the end of year of death.
20	Calculate of prospective reserve and Retrospective reserve for the following policies: (i) Fully continuous (ii) Fully discrete.

ST 46(PC) : Project(from 2014-15) (2 credits) See note no. 8

This part of the course consists of one of the following two components –

(i) Summary of research articles

Or

(ii) Data Analysis

1. **Summary of Research Articles**

Students are expected to read some (not less than 2) articles (exact number of articles will be decided by the supervisor) on a specified topic or theme, summarize and write a comprehensive report and present the summary of the articles.

2. **Data Analysis**

Students are expected to choose her/his own project, wherein they are expected to analyze data pertaining to certain theme using a variety of statistical tools that they have studied so far.

Note: 1. Students have to prepare project report in two copies of which one is to be submitted for assessment.

2. Data analysis project can be done in a group of maximum 3 students, however summary of research articles has to be done individually.

ST 47(TE) : Non parametric Statistical Inference(from 2014-15) (4 credits)

Unit 1:

Review of Order statistics, empirical distribution function, Glivenko canteeli theorem, Kolmogorov- Smirnov one sample tests, comparison of Chi-square and Kolmogorov- Smirnov tests.

[15 L]

Unit 2:

One sample U statistic, Kernel and symmetric kernel, two sample U statistic, asymptotic distribution of U statistic, UMVUE property of U statistic, asymptotic distribution of linear function of order statistics.

[15L]

Unit 3:

Rank tests, locally most powerful rank tests, sign test and Willcoxon signed rank test, Kolmogorov- Smirnov two sample test, median test, Mann Whitney U test, normal score test.

[15 L]

Unit 4 :

ARE of various tests based on linear rank statistic, Friedman's two way ANOVA test, density estimation.

[15 L]

Books Recommended:

1. Casella, G. and Berger, R.L. (1990, 2002) Statistical Inference. Wadsworth publishing Co., Belmont, CA
2. Conover W. J. (1999) Practical Non-parametric Statistical Inference, Wiley India Edition
3. Gibbons J. D. (1971) Non-parametric Statistical Inference, McGraw Hill Kogakusha, Ltd. New Delhi
4. Lehmann, E.L. (1983) Theory of Point Estimation. Wiley, New York.
5. Lehmann, E.L. (1986) Testing Statistical Hypotheses. 2-nd edition Wiley, New York.
6. Silvey, S.D. (1991) Statistical Inference. Chapman & Hall, London

ST(TE/TO) 48 : Statistical Analysis of Clinical Trials (from 2014-15) (4 credits)

Unit 1

Introduction to clinical trials: 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. Data management: data definitions, case report forms, database design, data collection systems for good clinical practice. Bioavailability, pharmacokinetics and pharmacodynamics, two-compartment model.

(15L)

Unit 2

Design of clinical trials: parallel vs. cross-over designs, cross-sectional vs. longitudinal 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 bio-equivalence trials. Inference for 2x2 crossover design: Classical methods of interval hypothesis testing for bioequivalence, Bayesian methods, nonparametric methods.

(15L)

Unit 3

Power and sample size determination, multiplicative (or log-transformed) model, ML method of estimation, assessment of inter and intra subject variabilities, detection of outlying subjects.

Optimal crossover designs: Balaam's design, two-sequence dual design, optimal four period designs, assessment of bioequivalence for more than two drugs, Williams design.

(15L)

Unit 4

Designs based on clinical endpoints: Weighted least squares method, log-linear models, generalized estimating equations. Drug interaction study, dose proportionality study, steady state analysis. Interim analysis and group sequential tests, alpha spending functions. Analysis of categorical data.

(15L)

Books Recommended:

1. Chow S.C. and Liu J.P.(2009). Design and Analysis of Bioavailability and bioequivalence. 3rd Edn. CRC Press.
2. Chow S.C. and Liu J.P. (2004). *Design and Analysis of Clinical Trials*. 2nd Edn. Marcel Dekkar
3. Fleiss J. L.(1989). *The Design and Analysis of Clinical Experiments*. Wiley.
4. Friedman L. M.Furburg C. Demets D. L.(1998). *Fundamentals of Clinical Trials*, Springer.
5. Jennison .C. and Turnbull B. W. (1999). *Group Sequential Methods with Applications to Clinical Trails*, CRC Press.
6. Marubeni .E. and Valsecchi M. G. (1994). *Analyzing Survival Data from Clinical Trials and Observational Studies*, Wiley.

ST(TE/TO) 49:Statistics Education and Research Methodology (from 2014-15)

(4 credits)

Section I: Statistics Education

Unit 1:

Characteristics of 21st century professional – problem solver, innovator, effective communicator, collaborator, self-directed learner, information and media literate, globally aware, civically engaged, critical thinker; nature of Statistics discipline; definition and goals of Statistics education; characteristics of a good teacher, Bloom's Taxonomy of learning levels – knowledge level, comprehension, application, analysis, evaluation and synthesis; teaching statistical concepts at various levels of learning. Key verbs to test learning outcomes at various levels of learning; using computers and information technology to teach statistics – using simulations to teach statistical concepts.

(15 L)

Unit 2:

Introduction to design science and applications to designing controlled experiments, observational studies, survey instruments, preparing for statistical consultancy services-i to Informal inference; statistical thinking and communication; Monitoring the effective use of statistical methods and statistical software packages, topics from current research in statistics education.

(15 L)

Section II: Research Methodology

Unit1:

Academicians view and practitioners view of Statistics as a discipline; meaning of research; basic skills required for researching in statistics; Critical and creative reading; critical thinking and creative thinking; types of research, a Model and its components of research; understanding the various components.

(15 L)

Unit 2:

General research patterns and research patterns specific to research in Statistics. Detailed study of various researching strategies and some illustrations, designing of instruments for research and their validation, written and oral presentation of research outcomes, physical and logical structuring of research writings, illustrations of research problems and outcomes from literature.

(15 L)