BANGALORE UNIVERSITY
DEPARTMENT OF STATISTICS

MSc COURSE
in
STATISTICS

REGULATIONS AND SYLLABUS

SEMESTER SCHEME
(CBCS)
Year 2017

JNANA BHARATHI
BENGALURU 560 056

DEPARTMENT OF STATISTICS
BANGALORE UNIVERSITY

Extract of the Proceedings of the Meeting of the Board of Studies in Statistics held on 04 May 2017

REVISED SYLLABUS FOR M.Sc. (Statistics) and Statistics as an optional subject for B.Sc. for Approval by the Academic Council

Members of the Board of Studies in Statistics (PG and UG) met on 04 May 2017 at 10:15 AM and transacted the items as listed in the Agenda, vide Meeting Notice dated (copy enclosed).

Members present:

1. Prof. M R Srinivasan, University of Madras,
2. Prof. B Ismail, University of Mangalore,
3. Sri. R Prakash, Vijaya College, Bangalore
4. Smt. Mamta Ramesh, MES College, Bangalore
5. Prof. P Rajalakshmi, Bangalore University
6. Prof. G Nanjundan, Bangalore University
7. Prof. Parameshwar V. Pandit, Bangalore University
8. Dr. V Srinivas, Bangalore University
9. Prof. J V Janhavi (Chairperson)

Smt. K. S. Radhamani, Karnatak University's Karnatak College, expressed her inability to attend the meeting.

Resolutions of the Meeting:

1) Revision of the Syllabus for M.Sc. in Statistics and Statistics as one of the optional subjects for B.Sc.

The Chairperson presented the draft syllabus for M.Sc.(Statistics) and Statistics as one of the optional for B.Sc., prepared by the local members of the Board.

The members, after a thorough discussion on the proposed syllabi, suggested certain changes and approved the syllabi. The syllabi approved by the Board is appended (APPENDIX 1 and APPENDIX 2).

Mamta Ramesh 4/5/17

R-Prakash 4/5/17

G. Nan

[signature] 4/5/17

[signature] 4.5.17

[signature]

[signature]

P. Ra

Prof. J V Janhavi (Chairperson)

CHAIRPERSON
Department of Statistics
Bangalore University
BANGALORE - 560 056

REGULATIONS

M Sc COURSE (CBCS)
in
STATISTICS
(Semester Scheme)

1. **Name of the course : MSc in STATISTICS**
2. **Duration of the course : Four Semesters**
3. **Objectives:** The course is designed to cover modern Statistical theory and methods providing a good foundation for research in Statistics and for a professional career in Statistical Applications as well. The course provides opportunity to develop skills for data analysis using modern computing facilities. The course is structured such that a student can specialize in at least three specific areas of Statistical applications. The individual project work provides field experience and hands-on training.

3. **Eligibility:** A candidate who has passed B.Sc. Degree Examination with Statistics and Mathematics as optional subjects, securing a minimum of 50% in the Statistics and a minimum of 40% in the aggregate of this University or any other University equivalent thereto is eligible. Further, a candidate who has passed B.Sc. degree examination of any other university with Statistics as major/main subject and Mathematics as minor subject and a minimum of 50% marks in Statistics and 40% in aggregate is also eligible.

Scheme of Examination

Paper Code	Paper Title	Inst. Hrs/ Week	Duration of Exam (in hrs)	Int. Ass. Marks	Exam. Marks	Credits
I Semester						
STA 101	Sampling Theory	4	3	30	70	4
STA 102	Probability Theory - I	4	3	30	70	4
STA 103	Distribution Theory – I	4	3	30	70	4
STA 104	Quality Assurance and Reliability	4	3	30	70	4
STA 105	Real Analysis	4	3	30	70	4
STA 106	Matrix Algebra (Soft Core)	3	3	30	70	2
STA 107	Practical – I on R programming	4	3	15	35	2
STA 108	Practical – II (based on STA 101, STA 103, and STA 104)	4	3	15	35	2
II Semester						
STA 201	Statistical Inference -I	4	3	30	70	4
STA 202	Probability Theory –II	4	3	30	70	4
STA 203	Distribution Theory -II	4	3	30	70	4
STA 204	Linear Models and Regression Analysis	4	3	30	70	4
STA 205	Multivariate Analysis	4	3	30	70	4
STA 206	Statistics for National Development and Demography (Soft Core)	3	3	30	70	2
STA 207	Practical III(based on STA 201, STA 203, and STA 205)	4	3	15	35	2
STA 208	Practical IV (based on STA 204)	4	3	15	35	2

III Semester						
Paper Code	**Paper Title**	**Inst. Hrs/ Week**	**Duration of Exam (in hrs)**	**Int. Ass. Marks**	**Exam. Marks**	**Credits**
STA 301	Statistical Inference - II	4	3	30	70	4
STA 302	Data mining and Machine Learning	4	3	30	70	4
STA 303	Stochastic Processes	4	3	30	70	4
STA 304	Elective – I	4	3	30	70	4
STA 305	Open Elective	4	3	30	70	4
STA 306	Practical V (based on STA 301 and STA 303)	4	3	15	35	2
STA 307	Practical VI (based on STA 302 and Elective - I)	4	3	15	35	2
IV Semester						
STA 401	Statistical Inference -III	4	3	30	70	4
STA 402	Design and Analysis of Experiments	4	3	30	70	4
STA 403	Elective – II	4	3	30	70	4
STA 404	Elective – III	4	3	30	70	4
STA 405	Practical - VII (based on STA 401 and Electives - II and III	4	3	15	35	2
STA 406	Practical -VIII (based on STA 402)	4	3	15	35	2
STA 407	Project Work	8	Evaluation of reports	30	70	4
Total credits for all the four semesters						100

FIRST SEMESTER

STA 101: SAMPLING THEORY

(52 hours : 4 credits)

Unit 1

Standard random sampling methods: Simple random sampling, stratified random sampling, and systematic sampling. Determination of sample size. Determination of number of strata, construction of strata, post stratification. Design of field surveys. Questionnaire designing. Interviewing. Preparation of reports

10 hrs

Unit 2

Unequal probability sampling: Probability proportional to size (PPS) with replacement: Cumulative total method and Lahiri's scheme. Hansen-Hurwitz estimator. PPS without replacement: Horvitz-Thompson estimator, Midzuno-Sen estimator, Des Raj estimator for a general sample size and Murthy's estimator for a sample of size 2.

14 hrs

Unit 3

Cluster sampling with equal and unequal cluster sizes. Two-stage sampling with equal number of second stage units. SRS at both the stages. Two-stage sampling with unequal number of second stage units.

10 hrs

Unit 4

Ratio and regression estimators based on SRSWOR sampling: Bias, mean squared error, and variance estimation. Double sampling for ratio and regression estimation.

8 hrs

Unit 5

Issues in small area estimation - direct and synthetic estimators. Errors in surveys. Non-sampling errors. Randomized responses for variables: Warner and Greenberg models. Modeling observational errors, estimation of variance components, applications to longitudinal studies, repetitive surveys.

10 hrs

References

1. Chaudhuri, A. and Mukherjee, R. (1988). *Randomized Response: Theory and Techniques*, Marcel Dekker, New York.
2. Cochran, W. G. (1977). *Sampling Techniques*, 3/e; John Wiley, New York.
3. Des Raj and Chandok, P. (1998). *Sample Survey Theory*, Narosa Publishing House, New Delhi.
4. Mukhupadhyay, P. (2009). *Theory and Methods of Survey Sampling*, 2/e, Prentice Hall, New Delhi.
5. Murthy, M. N. (1977). *Sampling Theory and Methods*, Statistical Publishing Society, Calcutta.

6. Sampath, S. (2006). *Sampling Theory and Methods*, 2/e, Narosa Publishing House, New Delhi.

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

8. Sukhatme et al. (1984). *Sampling Theory of Surveys with Applications,* Iowa State University Press, USA.

STA 102: PROBABILITY THEORY – I

(52 hours : 4 credits)

Unit 1

Classes of sets: Field, σ -field, minimal σ-field, Borel field on R and R^n. Sequences of sets and their limits. Measure and its properties - σ-finite mesure. Counting measure. Lebesgue and Lebesgue- Stieltjes measures. Probability measure and its properties. **10 hrs**

Unit 2

Random variables, algebra of random variables, sequences of random variables, and vector random variables. Probability measure induced by a random variable. Distribution function of a random variable (d.f.), its properties. Decomposition of a d.f. into discrete and continuous parts. **12 hrs**

Unit 3

Expectation of a random variable. Properties. Monotone convergence theorem. Statement of dominated convergence theorem. Markov, Chebycheff, Jensen, Minkowski, and Holder inequalities. **12 hrs**

Unit 4

Characteristic function. Properties. Inversion theorem and its applications. Uniqueness theorem.

8 hrs

Unit 5

Product measure space. Fubini’s theorem. Fatou’s lemma, Statement of Radon-Nikodym theorem and its applications. **10 hrs**

References

1. Ash, R.B. and Doleans-Dade, C.A. (2000). *Probability and Measure Theory,* Academic Press, New York.

2. Bhat, B.R. (1999). *Modern Probability Theory,* 2/e, New Age International, New Delhi.

3. Billingsley, P. (1995). *Probability and Measure*, 3/e, John Wiley, New York.

4. Burril, C. (1972). *Measure, Integration, and Probability,* McGrawHill International, New York.

5. Chung, K.L.(2001). *A Course in Probability, 3/e,* Academic Press, New York.

6. Clarke, L.E. (1975). Random Variables, Longman Mathematical Texts, London.
7. Khosnevisan, D. (2013). *Probability*, American Mathematical Society, Indian Edition, Universities Press, Hyderabad.
8. Rao, C.R. (1973). *Linear Statistical Inference and Its Applications*, John Wiley, New York.

STA 103: DISTRIBUTION THEORY – I

(52 hours : 4 credits)

Unit 1
Distribution function of a random variable. Moment generating function. Probability generating function. **6 hrs**

Unit 2
Standard discrete and continuous distributions: Discrete uniform, binomial, Poisson, geometric, negative binomial, hyper geometric, uniform, normal. lognormal, exponential, gamma, beta, Weibull, Laplace, Pareto, Chisquare, t, and F distributions. Non-central chi-square, t, and F distributions. **12 hrs**

Unit 3
Truncated, power series, modified power series, compound, and mixture distributions. Distributions of functions of random variables: Transformation and moment generating function techniques. **12 hrs**

Unit 4
Distribution function of a random vector. Joint, marginal, and conditional distributions. Conditional expectation and conditional variance. Distributions of functions of several random variables - change of variables technique. Convolution of two random variables. **12 hrs**

Unit 5
Bivariate Normal Distribution: Marginal and conditional distributions, moment generating function. Bivariate exponential distribution: Gumbel and Marshall-Olkin. **10 hrs**

References

1. Dudewicz, E. J. and Mishra, S. N. (1980). *Modern Mathematical Statistics,* John Wiley, New York.
2. Hogg, R.V. and Tanis, E.A. (2005). *Probability and Statistical Inference,* 6/e, Macmillan, New York.
3. Johnson, N. L. and Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions,* Vol. 1 and Vol. 2, Wiley, New York.

4. Mukhopadhyay, P. (2015). *Mathematical Statistics*, Books and Allied (P) Ltd., Kolkata.
5. Rohatgi, V.K. and Saleh, A.K.Md. E. (2002). *An introduction to Probability and Statistics*, John Wiley, New York.
6. Rao, C.R. (1973). *Linear Statistical Inference and Its Applications,* Wiley Eastern, New Delhi.

STA 104: QUALITY ASSURANCE AND RELIABILITY ANALYSIS

(52 hours : 4 credits)

Unit 1

Concept of quality. Quality function and quality characteristics. Statistical modeling. Quality assurance - its evolution and modern trends. Statistically controlled processes. Chance and assignable causes. Principles of a process control chart and associated decision rules. Shewhart control charts for monitoring process level and process dispersion. Rational subgroups. Pre-control and analysis of patterns on a control chart. The seven QC tools. Design quality and conformance quality. Quality costs. Quality and productivity. Design and implementation of SPC and six-sigma programmes. BIS and ISO certification.

10 hrs

Unit 2

Analysis of OC, ARL, and other measures. Techniques for improving sensitivity of a chart. Natural tolerances. Process capability and its measures. CUSUM and EWMA charts. Process control with autocorrelated observations. Modifications of Shewhart control chart. Multivariate control charts.

10 hrs

Unit 3

Single and double lot-by-lot acceptance sampling plans. Performance analysis of sampling plans: OC, ASN, AOQ, and AOQL. Dodge-Romig and MIL-STD systems.

6 hrs

Unit 4

Reliability and hazard rate functions of a single component. Classes of lifetime distributions. Concept of ageing, positive and negative ageing, IFR, IFRA, NBU, NBUE, DMRL classes of distributions and their dual classes. Interrelations among the classes of life time distributions. Closures of these classes under formation of coherent systems, convolutions, and mixtures. Series and parallel systems. k-out- of -n system. Structure function and block diagrams of these systems. Minimal path and minimal cut sets.

12 hrs

Unit 5

Reliability of systems of independent components. Bounds of reliability functions of these systems. System life as a function of component lives. Expected system lifetime. Systems with repairs. Maintenance policies: Age replacement and block replacement policies and their characteristics. Reliability modeling: Introduction to shock models, stress-strength models, and proportional hazard models. **14 hrs**

References

1. Alwan, L.C. (2000). *Statistical Process Analysis,* McGraw Hill, New York.
2. Barlow, R.E. and Proschan, F. (1981). *Statistical Theory of Reliability and Life Testing*, 2/e, To Begin With, Silver Spring, MD, USA.
3. Burr, ????
4. Grant, E. L. and Leavenworth, R. S. (1996). *Statistical Quality Control*. 7th edition, McGrawHill, New York.
5. Mittage, H.J. and Rinne, H. (1993). *Statistical Methods of Quality Assurance,* Chapman and Hall, London, UK.
6. Montgomery, D.C. (2012). *Introduction to Statistical Quality Control*, 7/e, John Wiley, New York.
7. Ross, S.M. (2010). *Introduction to Probability Models,* 10/e, Academic Press, New York.
8. Smith, G.M. (1991). *Statistical Process Control and Quality Improvement,* 3/e, Prentice Hall, New York.
9. Wetherill, G.B. and Brown, D.W. (1991). *Statistical Process Control: Theory and Practice,* Chapman and Hall, London, UK.

STA 105: REAL ANALYSIS

(52 hours : 4 credits)

Unit 1

Interior points and limit points of subsets of R. Open and closed subsets of R. Bolzanao-Weierstrass theorem.

6 hrs

Unit 2

Riemann-Stieltjes (R-S) integral of a bounded real valued function. Necessary and sufficient condition for R-S integrability. Properties of R-S integrals. Integration by parts. Change of variables in R-S integrals. Mean value theorems for R-S integrals.

15 hrs

Unit 3

Improper Riemann and Riemann -Stieltjes integrals. Convergence and absolute convergence of improper integrals. Abel's and Dirichlet's theorems on the convergence of the integrals of the product of two integrands. Beta and Gamma integrals and their properties. Legendre's duplication formula. Stirling's approximation for n! Integrals involving parameters. Improper integrals involving parameters and their uniform convergence. Differentiation under the integral sign.

12 hrs

Unit 4

Maxima and minima of functions of several variables. Lagrangian multipliers. Double integrals. Leibnitz rule.

9 hrs

Unit 5

Sequences and series of functions. Pointwise and uniform convergence. Weierstrass test. Consequences of uniform convergence of sequences and series of functions Term by term integration and differentiation **10 hrs**

References

1. Apostol, T.M. (1986). *Mathematical Analysis*, 2/e, Narosa Publishing House, New Delhi.
2. Bartle, R.G. (1975). *The Elements of Real Analysis*, 2/e, John Wiley.
3. Bilodeau, G.G., Thie, P.R., and Keough, G.E. (2010). *An Introduction to Analysis*, 2/e, Jones and Bartlett (Indian Edition), New Delhi.
4. Goldberg, R.R. (1970). *Methods of Real Analysis,* Oxford and IBH Publishing Company, New Delhi.
5. Malik, S.C. and Arora, S. (1998). *Mathematical Analysis,* New Age, New Delhi.
6. Rudin, W. (2013). Principles of Mathematical Analysis, 3/e, Indian Print, Tata McGrawhill, New Delhi.

STA 106: MATRIX ALGEBRA (SOFT CORE)

(40 hours : 2 credits)

Unit 1

Vector spaces, subspaces, linear dependence and independence, basis and dimension of a vector space, inner product and orthogonallity of vectors, orthonormal basis, orthogonal matrix and its properties. column and row null spaces of a matrix, nullity of a matrix. Linear transformations, non-singular and orthogonal linear transformations and their properties. Trace of a matrix: definition, properties, idempotent matrix and its properties, Partitioned matrices: computation of determinant and inverse.

10 hrs

Unit 2

Characteristic roots, characteristic vectors. Cayley-Hamilton theorem, algebraic and geometric multiplicities of a characteristic root and their relationship, similar matrices, symmetric matrices and properties of their characteristic roots and characteristic vectors, spectral decomposition of a real symmetric matrix.

8 hrs

Unit 3

Matrix decomposition: Spectral decomposition, LU decomposition, SVD, Gram-Schmidt orthogonalization process.

6hrs

Unit 4

Quadratic forms: Congruent transformations, congruence of symmetric matrices. Canonical reduction and orthogonal reduction of real quadratic forms. Sylvester's law of inertia, nature of a quadratic form, necessary and sufficient condition for a real quadratic form to be positive definite, Gram matrix and properties, simultaneous reduction of a pair of quadratic forms. Derivative of a function with respect to a vector and a matrix.

10 hrs

Unit 5

Generalised inverse of a matrix: g-inverse and Moore – Penrose inverse: properties and computations. Systems of linear equations: consistency, existence of solutions, number of solutions, and solving the system of equations.

6 hrs

References

1. Graybill, F. A. (2002). *Matrices and Applications in Statistics*, 2/e, Wadsworth Publishing Co., Belmont, California, USA.
2. Rao, A. R. and Bhimasankaram, P. (2000). *Linear Algebra*, Hindustan Book Agency, New Delhi.
3. Searle, S. R. (2017). *Matrix Algebra Useful for Statistics*, 2/e, Wiley, New York.
4. Narayan, S. (2010). *A Text Book of Matrices*, S. Chand and Company Ltd., New Delhi.
5. Harville (1997). *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag, New York.

STA 107: Practical Paper

(4 hours/ week)

(2 credits)

Introduction to R. Data and file handling in R. Descriptive statistics. Exploratory Data Analysis (EDA). Bivariate and multivariate data: numeric and character variable. Various plots in R. Simulation in R. Fitting regression models. Graphical interface in R. Programming in R- editing functions – Functions- conditional evaluation – looping and object oriented programming in R.

List of Assignments

1. R Basics and Objects. Vectors, Matrices, data.frame, and lists. R functions and packages, datasets, and Import/Export functions.
2. Descriptive statistics for univariate (numeric) data: Measures of central tendency, EDA, and dispersion. Visualization techniques: bar-plot, histogram, box plot, and stem- and-leaf diagram.
3. Bivariate and multivariate data: Plots, correlation coefficient, and regression lines.
4. Discrete and continuous distributions – pmf and pdf plots, Sampling distributions.
5. Fitting of distributions to given data – raw programs and R functions [fitdistr for example.]

6. Simulation from discrete, continuous, and mixture distributions.
7. Illustration of limit theorems.
8. Using R functions for standard parametric statistical tests.
9. Selection of SRSWR and SRSWOR.
10. Optimization: maxima and minima of functions of one and two variables. Use of *optimize* and *optim* functions.

References

1. Dalgaard, P. (2008). *Introductory Statistics with R*, Springer, New York.
2. Horgan, J. M. (2009). *Probability with R*, John Wiley, New York.
3. Prabhanjan, N.T., Suresh, R., and Manjunath, B.G. (2016). *A Course in Statistics with R*, Wiley, New York.
4. Purohit, S. Gore, Desmukh, S. (2007). *Statistics Using R*, Narosa, New Delhi.
5. Teetor, P. (2011). R Cook book, Oreilly Media Inc., Cambridge, USA.
6. Venebles, W. N. and Smith, D. M. and the R Core Development Team (2013). *An Introduction to R*, Available in Installed R Software.
7. Verzani, J. (2005). *Using R for Introductory Statistics*, Chapman and Hall, London.

STA 108: Practical Paper

(Assignments based on STA 101 and STA 104)

(4 hours/ Week)

(2 credits)

List of Assignments

1. Fitting discrete distributions and normal distribution
2. Ratio and regression methods of estimation
3. Cluster sampling
4. PPSWR
5. PPSWOR
6. Two stage sampling and double sampling methods
7. OC and ARL curves of $\overline{X}$ and R control charts
8. CUSUM control charts
9. EWMA control charts
10. Multivariate control charts
11. Single and double attribute sampling plans.

SECOND SEMESTER

STA 201: STATISTICAL INFERENCE – I

(52 hours : 4 credits)

Unit 1

Families of distributions: Location, scale, location-scale, exponential, Pitman, regular, and Cramer families**.**

4 hrs

Unit 2

Sufficiency and completeness: Sufficiency, factorization theorem, minimal sufficiency, likelihood equivalence, completeness, bounded completeness. Reparametrization, statement of the theorem on complete sufficient statistic in k-parameter exponential family.

12 hrs

Unit 3

Unbiased estimation, combining unbiased estimators. Mean square error. Fisher information function. Cramer-Rao inequality, Rao-Blackwell, and Lehmann-Scheffe theorems. Uniformly minimum variance unbiased (UMVU) estimation. Fisher information matrix, simultaneous estimation of parameters of multinomial and normal distributions. Ancillary statistics, Basu's theorem, and its application in UMVU estimation.

14 hrs

Unit 4

Methods of estimation: Maximum likelihood (ML) and moment estimation. Properties of ML estimators. Computation of ML estimates using Newton-Raphson, method of scoring, and EM-algorithm.

8 hrs

Unit 5

Testing of hypotheses: Randomized and non-randomized tests. Power and size of a test. Most powerful (MP) test. Neyman–Pearson lemma. Uniformly most powerful (UMP) tests, Monotone likelihood ratio (MLR) property and construction of UMP tests for the families of distributions possessing MLR property and applications. Non- existence of UMP tests for two-sided alternatives.

14 hrs

References

1. Casella, G. and Berger, R.L. (2002). *Statistical Inference, 2/e,* Duxbury Press, Belmont, California, USA.
2. Dudewicz, E.J. and Mishra, S.N. (1980). *Modern Mathematical Statistics*, John Wiley, New York.
3. Kale, B.K. and Muralidharan, K. (2015). *Parametric Inference: An Introduction,* Narosa, New Delhi.
4. Lehmann, E.L. and Casella, G. (1998). *Theory of Point Estimation*, Springer, New York.

5. Lehmann, E.L. and Romano, J.P. (2005). *Testing Statistical Hypotheses,* 2/e, John Wiley, New York.
6. Rohatgi, V.K.and Saleh, A.K.Md.E. (2002): *An Introduction to Probability and Statistics*, John Wiley, New York.
7. Zacks, S. (1981). *Parametric Statistical Inference*, John Wiley, New York.

STA 202: PROBABILITY THEORY – II

(52 hours : 4 credits)

Unit 1

Independence of events and of random variables. Borel-Cantelli lemma. Convergence of sequences of random variables in distribution and in probability. Almost sure convergence and convergence in the r^{th} mean Relationship between these modes of convergence. Slutsky's theorem. **14 hrs**

Unit 2

Levy's continuity thorem. Khintchin and Tchebycheff weak laws of large numbers (WLLN). Markov's condition for WLLN. Generalized WLL.N Kolmogorov's inequality. Kolmogorov strong law of large numbers (SLLN). Kolmogorov's condition for the SLLN. **12 hrs**

Unit 3

Levy, Lindeberg-Levy, and Liapunov central limit theorems. Statement of Lindeberg-Feller central limit theorem. Applications of these theorems. Delta method and its applications.

12 hrs

Unit 4

Conditional probability – properties, Conditional Expectation. Smoothing and other properties. **4 hrs**

Unit 5

Martingales. Sub and super martingales. Doob decomposition. Martingale convergence thorem. Stopping times. Doob's optional sampling theorem. Wald's fundamental identity.

10 hrs

References

1. Ash, R.B. and Doleans-Dade, C.A. (2000). Probability and Measure Theory, Academic Press, New York.
2. Bhat, B.R. (1999). Modern Probability Theory, 2/e, New Age International Publishers, New Delhi.
3. Bhat, B.R. (2000). *Stochastic Models: Analysis and Applications,* New Age International Publishers, New Delhi.
4. Billingsley, P. (1995). Probability and Measure, 3/e, John Wiley, New York.

5. Burril, C.W. (1972). Measure, Integration, and Probability, McGrawHill International, New York.

6. Chung, K.L.(2001). A Course in Probability, *3/e*, Academic Press, New York.

7. Laha, R.G. and Rohatgi, V.K. (1979). *Probability Theory*, John Wiley, New York.

8. Khosnevisan, D. (2013). *Probability*, American Mathematical Society, Indian Edition, Universities Press, Hyderabad.

9. Rao, C.R. (1973). *Linear Statistical Inference and Its Applications,* John Wiley, New York.

10. Walsh, J.B. (2010). *Knowing the Odds: An Introduction to Probability*, American Mathematical Society, Providence, Rhode Island, USA.

STA 203: DISTRIBUTION THEORY – II

(52 hours : 4 credits)

Unit 1

Multinomial and multivariate normal distributions- marginal and conditional distributions. Independence of sub vectors. Distribution of a linear function of normal random variables.

8 hrs

Unit 2

Characteristic function of the multivariate normal distribution and its applications. Distribution of quadratic forms. Cochran's theorem. Independence of quadratic forms, independence of linear and quadratic forms. **10 hrs**

Unit 3

Sampling from Multivariate Normal Distribution – sample mean vector, sample covariance matrix. Distribution of the sample mean vector, independence of sample mean vector and sample covariance matrix. Assessing multivariate normality. Q-Q and chi-square plots. multiple and Partial Correlation coefficients and their distributions(statements only). Distribution of Hotelling's T^2 and Mahalanobis' D^2 statistics. Wishart Distribution and Distribution of sample generalized variance. **16 hrs**

Unit 4

Order statistics. Distribution of functions of order statistics. Distributions of range and sample median. Extreme value distribution. Record values.

8 hrs

Unit 5

Simulation: Design of simulation models, Technique for generating random deviates- inverse transformation method, rejection technique, Box- Muller technique for generating normal deviates. Generation of observations from multivariate normal distributions. Variance reduction technique. Statistical analysis of simulation output.

10 hrs

References

1. Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis,* 3/e, John Wiley, New York.

2. Arnold, B. C., Balakrishnan, N., and Nagaraja, H.N. (1992). *A First Course in Order Statistics,* John Wiley, New York.

3. Dudewicz, E. J. and Mishra, S. N. (1980). *Modern Mathematical Statistics,* John Wiley, New York.

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

5. Mukhopadhyay, P. (2015). *Mathematical Statistics*, Books and Allied (P) Ltd., Kolkata.

6. Rohatgi, V.K. and Saleh, A.K.Md.E. (2002). *An introduction to Probability and Statistics,* John Wiley, New York.

7. Kshirsagar,A.M. (1972). *Multivariate Analysis*, Marcel-Dekker, New York.

8. Hogg, R.V. and Tanis, E.A.(1983). *Probability and Statistical Inference*, 3/e, Macmillan, New York.

STA 204: LINEAR MODELS AND REGRESSION ANALYSIS

(52 hours : 4 credits)

Unit 1

Linear estimation: Gauss-Markov model, least squares estimation, BLUE, Gauss-Markov theorem. Distributional properties of least squares estimators, confidence intervals, general linear hypotheses, testable hypotheses, and their likelihood ratio test procedure. Estimation under linear restrictions involving estimable functions. **14 hrs**

Unit 2

Multiple linear regression, estimation, and porperties. Prediction of new observations and prediction interval. Hidden extrapolation. Use of dummy variables. Generalized linear models: link function, binary logistic regression, and Poisson regression. **12 hrs**

Unit 3

Measures of model adequacy, coefficient of determination R^2, lack of fit test, residuals, scaling residuals, and residual analysis: residual plots as tests for departure from assumptions of homoscedasticity, normality (Q-Q plot), non-linearity, and detection of outliers. Detecting influential observations. Transformations: Box-Cox transformation and transforming the predictors. Subset selection of regressors: Mallow's C_p statistic, all possible, stepwise, forward and backward regressions.

12 hrs

Unit 4

Heteroscedasticity and autocorrelation: sources, consequences, detection, and remedial procedures. **6 hrs**

Unit 5

Multicollinearity: sources, consequences, detection, and remedial procedures. Ridge regression and generalised least squares. Validation of regression models. Analysis of model coefficients and predicted values. Collecting fresh data. Data splitting. **8 hrs**

References

1. Cook, R. D. and Weisberg, S. (1982). *Residual and Influence in Regression,* Chapman and Hall, London,UK.
2. Draper, N. R. and Smith, H. (1998). *Applied Regression Analysis, 3/e,* John Wiley, NewYork.
3. Gunst, R. F. and Mason, R. L. (1980). *Regression Analysis and Its Applications - A Data Oriented Approach,* Marcel and Dekker, New York.
4. Montgomery, D. C. and Peck, E. A., and Vining, G. G. (2012). *Introduction to Linear Regression*, 5/e, John Wiley, New York.
5. Ryan, T. P. (1997). *Modern Regression Methods,* John Wiley, New York.
6. Searle, S. R. (1971). *Linear Models,* John Wiley, New York.
7. Weisberg, S. (1985). *Applied Linear Regression*, John Wiley, New York.

STA 205: MULTIVARIATE ANALYSIS

(52 hours : 4 credits)

Unit 1

MLE's of the parameters of multivariate normal distribution and their sampling distributions. Likelihood ratio tests: Tests of hypotheses about the mean vectors and covariance matrices for multivariate normal populations. Independence of sub vectors and sphericity test. **12 hrs**

Unit 2

Classification and discriminant procedures: Bayes, minimax, and Fisher's criteria for discrimination between two multivariate normal populations. Sample discriminant function. Tests associated with discriminant functions. Probabilities of misclassification and their estimation. Discrimination for several multivariate normal populations. Canonical discriminant function. **10 hrs**

Unit 3

Multivariate regression model: Estimation of parameters, tests of linear hypotheses about regression coefficients. Multivariate analysis of variance (MANOVA) of one and two- way classified data. **8 hrs**

Unit 4

Principal components, sample principal components asymptotic properties. Canonical variables and canonical correlations: definition, estimation, computations. Test for significance of canonical correlations. Factor analysis: Orthogonal factor model, factor loadings, estimation of factor loadings, factor scores. Applications. **12 hrs**

Unit 5

Cluster Analysis: distances and similarity measures, hierarchical clustering methods, K – means method. Multidimensional scaling: nature of the problem, classical solution.

10 hrs

References

1. Anderson, T. W. (2004). *An Introduction to Multivariate Statistical analysis,* 3/e, John Wiley, New York.
2. Giri, N. C. (1977). *Multivariate Statistical Inference,* Academic Press, New York.
3. Johnson, R. A. and Wichern, D.W. (2003). *An Introduction to Multivariate Statistical Analysis*, 5/e, Pearson Education.
4. Kshirsagar, A. M. (1972). *Multivariate Analysis,* Marcel Dekker, New York.
5. Morrison, D. F. (2005). *Multivariate Statistical Methods,* 4/e, McGrawhill, New York.
6. Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory,* John Wiley, New York.
7. Rao, C. R. (1973). *Linear Statistical Inference and Its Applications,* 2/e, John Wiley, New York.
8. Seber, G. A. F (1984). *Multivariate Observations*, John Wiley, New York.
9. Sharma, S. (1996). *Applied Multivariate Techniques*, John Wiley, New York.
10. Srivastava, M. S. (1979). *An Introduction to Multivariate Statistics,* North Holland.
11. Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979). *Multivariate Analysis,* Academic Press, New York.

STA 206: STATISTICS FOR NATIONAL DEVELOPMENT AND DEMOGRAPHY (SOFT CORE)

(40 hours : 2 credits)

Unit 1

Economic development: Classical growth models of Adam Smith, Malthus, David Ricardo. Growth in per capita income and distributive justice. Indices of development, human development index. **10 hrs**

Unit 2

Estimation of national income: product approach, income approach, and expenditure approach. GNP and GDP of India. **4 hrs**

Unit 3

Measuring inequality in incomes, Gini coefficient, and Theil's measure. Poverty measurement: measures of incidence and intensity, combined measures, Kakwani and Sen indices.

6 hrs

Unit 4

Measures of fertility. Distribution of time to first birth. Inter-live birth intervals and number of births (for both homogeneous and non-homogeneous groups). Estimation of parity progression ratios. Measures of mortality and cohort measures. Construction of abridged life tables. Distribution of life table functions.

12 hrs

Unit 5

Population growth models- exponential, logistic, Gompertz models. Population projection using Leslie matrix.

8 hrs

References

1. Biswas, S. (1988). *Stochastic Processes in Demography and Applications,* Wiley Easter, New Delhi.
2. Chiang, C. L. (1968). *Introduction to Stochastic Processes in Biostatistics,* John Wiley, New York.
3. Gupta, K. R. (2010). *Economic Growth Models,* Atlantic Publishers and Distributors, New Delhi.
4. Keyfitz, N. (1977). *Applied Mathematical Demography,* Springer Verlag.
5. Ramkumar, R. (1986). *Technical Demography, Wiley Eastern, New Delhi.*
6. Sen, A. (1977). *Poverty and Inequality,* Stanford University Press, USA.

STA 207: Practical Paper

(Assignments Based on STA 201 and STA 205)

(4 hours / Week)

(2 credits)

1. Generating random samples given sufficient statistics.
2. UMVU estimates.
3. MP tests.
4. UMP tests.
5 Maximum likelihood and moment estimates.
6. Maximum likelihood estimate: Method of scoring.
7. Applications of Hotelling's T^2.
8. Tests concerning covariance matrices.
9. Multivariate regression analysis and MANOVA.
10. Discriminant analysis and classification.
11. Principal component analysis and canonical correlations.
12. Factor analysis.

STA 208: Practical Paper

(Assignments Based on STA 203 and STA 204)

(4 hours / Week)

(2 credits)

1. Computation of mean vector, covariance matrix, partial and multiple correlations. from a multivariate data.
2. Test for multivariate normal distribution.
3. Simulation of observations from univariate and multivariate normal distributions.
4. Fitting multiple linear regression models.
5. Stepwise regression analysis.
6. Multicollinearity diagnostics.
7. Residual analysis.
8. Tests for autocorrelation.
9. Fitting a ridge regression model.
10. Fitting logistic and Poisson regression models.

THIRD SEMESTER

STA 301: STATISTICAL INFERENCE - II

(52 hours : 4 credits)

Unit 1

Consistency: Definition and criteria for consistency. Weak and strong consistencies. Marginal and jointly consistent estimators. Methods of obtaining consistent estimators, Invariance property. Comparison of consistent estimators. Consistent asymptotically normal (CAN) property: Definition and methods of obtaining CAN estimators. Example of consistent but not asymptotic normal in Pitman family. Invariance property. CAN property of MLE in Cramer's family. Best asymptotically normal (BAN) estimator. Asymptotic relative efficiency (ARE). **16 hrs**

Unit 2

Robust estimation: The influence curve and empirical influence curve. M-estimation: Median, Trimmed and winsorized mean. Influence curve for M-estimators. Limiting distribution of M-estimators. Resampling methods: Quenouille's Jackknife estimation, parametric and nonparametric bootstrap methods. **7 hrs**

Unit 3

Likelihood ratio (LR) test: Standard applications including one and two sample problems for the normal and exponential distributions. Role of sufficiency in LR tests, Relationship between LR test and most powerful test. Asymptotic distribution of the LR test statistic. Wald and Rao's score tests. Consistency of LR test. **10 hrs**

Unit 4

Similar tests and unbiased tests. Statements of theorems for the construction of UMPU tests for exponential families of distributions and their application to samples from normal distribution leading to t, χ^2 and F tests – one and two-sample problems. Student's t- test for correlation coefficient. **12 hrs**

Unit 5

Interval estimation: Concept of confidence level and confidence co-efficient. Duality between acceptance region of a test and a confidence interval. Pivotal quantity method. Shortest length confidence intervals. UMA and UMAU confidence intervals. Large Sample confidence regions based on CAN estimators. **7 hrs**

References

1. Casella, G. and Berger, R. L. (1990). *Statistical Inference*, 2/e, Duxbury Press, Belmont, California, USA.
2. Dudewicz, E. J. and Mishra, S. N. (1980). *Modern Mathematical Statistics*, John Wiley, New York.

3. Kale, B. K. and Muralidharan, K. (2015). *Parametric Inference: An Introduction*, Narosa, New Delhi.
4. Lehmann, E. L. and Cassella, G. (1998). *Theory of Point Estimation*, 2/e, Springer-Verlag, New York.
5. Rohatgi, V. K. and Saleh, A. K. Md. E. (2002). *An Introduction to Probability and Statistics*, 2/e, John Wiley, New York.
6. Zacks, S. (1981). *Parametric Statistical Inference*, John Wiley, New York.

STA 302: DATA MINING AND MACHINE LEARNING

(52 hours : 4 credits)

Unit 1

Essential Machine Learning: Empirical Probabilities, apriori algorithm, naïve Bayes, graphical models, simulation, SMOTE, interest measures, ROC tests, k-fold cross-validation, regularization and bias-variance tradeoff, and EM algorithm. **10 hrs**

Unit 2

Neural Networks: Neural network (NN) architecture, activation functions, single layer perceptron model, multi-layer NN, back-propagation algorithm, and convolutional NN. **8 hrs**

Unit 3

CART: Constructing a tree, regression, classification, and survival tree, pruning trees, trees for imbalanced data, bagging, random forests, and boosting techniques. **10 hrs|**

Unit 4

Support Vector Machine: Kernel based regression, SVM and discriminant analysis, Gaussian kernel based SVM, and SVM regression. **7 hrs**

Unit 5

High Dimensional Data and Unsupervised Learning: Least absolute squares selection operator (LASSO), least angle regression (LARS), smoothly clipped absolute deviation, and high-dimensional inequalities. Unsupervised learning: k-means, k-mediods, hierarchical clustering, principal component analysis (PCA), and high-dimensional PCA. **17 hrs**

References

1. Alpaydin, E. (2014). *Introduction to Machine Learning*, 3e., MIT Press, USA.
2. Bishop, C. (2007). *Pattern Recognition and Machine Learning*, Springer, New York.
3. Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and Regression Trees*. CRC Press, London.

4. Bühlmann, P., and Van De Geer, S. (2011). *Statistics for High-dimensional Data: Methods, Theory, and Applications*, Springer Science and Business Media, USA.

5. Efron, B, and Hastie, T. (2016). *Computer Age Statistical Inference*, Cambridge University Press, London.

6. Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning*, 2/e, Springer, New York, USA.

7. Haykin, S. (2009). *Neural Networks and Learning Machines*, Pearson, New Delhi.

8. Mitchel, T. M. (2013). *Machine Learning*, McGraw Hill, New York, USA.

9. Tattar, P. N. (2013). *R Statistical Application Development by Example Beginner's Guide*. Packt Publishing Ltd.

STA 303: STOCHASTIC PROCESSES

(52 hours : 4 credits)

Unit 1

Stochastic processes and their classification according to state space and time domain. Countable state space Markov chains. Chapman-Kolmogorov equations, n-step transition probabilities and first passage probabilities. Computing n-step transition probability matrix by spectral representation.

12 hrs

Unit 2

Classification of states and chains. Invariant distributions. Estimation of transition probabilities of a Markov chain. Absorption probabilities. Absorption and recurrence times. Random walk. Applications from social, biological and physical sciences.

10 hrs

Unit 3

Continuous time Markov processes. Kolmogorov-Feller differential equations. Poisson process. Pure birth process, Yule-Furry process. Birth and death process. Applications to queuing theory.

12 hrs

Unit 4

Renewal process, renewal function. Renewal equation. Elementary renewal theorem and its applications.

8 hrs

Unit 5

Galton - Watson branching processes. Properties of generating functions. Offspring mean and classification of Galton – Watson Processes. Probability of extinction.

10 hrs

References

1. Adke, S. R. and Manjunath, S. M. (1984). *An Introduction to Finite Markov Processes*, Wiley Eastern. New Delhi.
2. Issacson,D. L, and Madsen, R. W. (1976). *Markov Chains and Applications*, John Wiley, New York.
3. Karlin, S and Taylor, H. M. (1975). *A First Course in Stochastic Processes*, Academic Press, New York.
4. Bhat, B. R. (2000). *Stochastic Models,* New Age International, New Delhi.
5. Medhi, J. (2017). *Stochastic Processes*, 4/e. New Age International, New Delhi.
6. Ross, S.M. (1996). *Stochastic Processes*, 2/e, John Wiley, New York.
7. Hoel, P.G., Port, S.C., and Stone, C.J. (1991). *Introduction to Stochastic Processes*, Universal Book Stall, New Delhi.
8. Jones, P.W. and Smith, P. (2001). *Stochastic Processes: An Introduction*, Arnold Press, New York.

STA 306: Practical Paper

(Assignments based on STA 301 and STA 303)

(4 hours / Week)

(2 credits)

1. Computing n-step transition probabilities and stationary probabilities
2. Generating a sample path of a Markov chain and estimating the transition probabilities
3. Classification of the states of a Markov chain and computation of absorption probabilities and mean recurrence times
4. Resampling methods
5. UMPU tests
6. Likelihood ratio tests
7. Wald test and Rao's score test.
8. Confidence intervals- pivotal quantity method.
9. UMA and UMAU confidence intervals
10. Large sample confidence intervals based on CAN estimates

STA 307: Practical Paper

(Assignments based on STA 302 and Elective – I)

(4 hours / Week)

(2 credits)

1. Naïve Bayes Classification
2. EM algorithm
3. Neural networks
4. Classification and regression trees
5. Support vector machine
6. LASSO and LARS regression
7. ROC analysis.

8 – 10 Assignments based on Elective - 1

Open Elective: STATISTICAL METHODS

(52 hours : 4 credits)

Unit 1

Statistics: meaning and role as a decision making science. Data types and scales of measurement. Descriptive Statistics: measures of central tendency, positional averages, measures of dispersion, skewness, and kurtosis. Presentation: tables, diagrammatic and graphical methods. Exploratory data analysis using descriptive measures and graphical tools.

10 hrs

Unit 2

Probability theory: random experiment, sample space, simple events, types of events, probability of an event, rules of probability, conditional probability, Bayes theorem. Random variables - discrete and continuous types. Probability distributions: Bernoulli, binomial, Poisson and normal distributions and their applications.

10 hrs

Unit 3

Sampling methods: population and sample, parameter and statistic, concept of a random sample, simple random sampling, stratified sampling, systematic sampling, sample size determination.

8 hrs

Unit 4

Testing of hypothesis: null hypothesis, alternate hypothesis, test statistic, level of significance, p-value. Testing hypothesis about population mean, tests for proportions. Confidence intervals. Contingency tables, chi-square test for independence of attributes.

12 hrs

Unit 5

Bivariate data, correlation, scatter plot, correlation coefficient and its properties, testing for correlation coefficient, rank correlation coefficient. Regression: linear relationship, linear regression model, simple linear regression, fitting the regression model, coefficient of determination, standard error of the estimate of the regression coefficient.

12 hrs

References

1. Campbell, R.C. (1974). *Statistics for Biologists*, Cambridge University Press, London.
2. Chatfield, C. (1981). *Statistics for Technology*, Chapman and Hall, London.
3. Frank, H. and Athoen, S.C. (1997). *Statistics: Concepts and Applications*, Cambridge University Press, London.
4. Medhi, J. (1992). *Statistical Methods: An Introductory Text*, Wiley Eastern, New Delhi.
5. Lind, D.A., Marchal, W.C., and Wathen, S.A. (2012). *Basic Statistics for Business and Economics,* McGrawHill, London.

FOURTH SEMESTER

STA 401: STATISTICAL INFERENCE – III

(52 hours : 4 credits)

Unit 1
Decision Theory and Bayesian Analysis: Elements of a decision problem. Estimation and testing as decision problems. Bayes paradigm. Prior and posterior distributions. Conjugate and non-informative priors. Construction of Bayes estimators relative to squared error, weighted squared error, absolute error, Stein and LINEX loss functions. Minimax estimation. **12 hrs**

Unit 2
Parametric empirical Bayes estimation: Application in binomial and normal models. Hierarchical Bayes estimation, Bayesian credible intervals. Predictive distributions. Admissibility and inadmissibility, inadmissibility of sample mean vector from a normal distribution. Application to statistical computing through MCMC. James -Stein estimator. Bayes and minimax tests. **10 hrs**

Unit 3
Introduction to sequential procedures. Stopping time. Wald equation. Sequential probability ratio test: termination property, approximations to stopping bounds and construction of SPRT for standard distributions. Statement of Wald fundamental identity. Operating characteristic and average sample number functions and their plotting.
10 hrs

Unit 4
Nonparametric Tests: Definition of U-statistic and properties. Hoeffding's one-sample U – statistic theorem. Tests for randomness. Standard one and two sample nonparametric tests for location. Mood's test for two sample scale problem, Kolmogorov – Smirnov tests. Analysis of variance by ranks. **12 hrs**

Unit 5
Discrete data analysis: Binary variables, inference for proportions. Relative risk and odds Ratio. Categorical variables. Two-way contingency tables: Probability structure, comparing proportions, inference for odds and log-odds ratios, chi -square tests of independence for ordinal data, Fisher's exact test. Small sample inference. Three- way contingency tables- partial and conditional association, Simpson's paradox, conditional and marginal odds ratios. **8 hrs**

References

1. Berger, J. O. (1980). *Decision Theory and Bayesian Analysis*, John Wiley, New York.
2. Carlin, B. P. and Louis, T. A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*, 2/e, Chapman and Hall, London.
3. Ghosh, B. K. (1970). *Sequential Tests of Statistical Hypotheses*, Addison Wesley, New York.

4. Gibbons, J.D. (1985). *Nonparametric Statistical Inference,* Marcel Dekker, New York.
5. Kale, B. K. and Muralidharan, K. (2015). *Parametric Inference: An Introduction,* Narosa, New Delhi.
6. Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation,* 2/e, Springer-Verlag, New York.
7. Randles, R. H. and Wolfe, D. A. (1987). *Introduction to the Theory of non-parametric Statistics,* John Wiley and Sons, New York.
8. Rohatgi, V. K. and Saleh, A. K.Md.E. (2002). *An Introduction to Probability and Statistics,* John Wiley, New York.
9. Sinha, S. K. (1988). *Bayesian Estimation,* New Age International, New Delhi.
10. Wald, A. (1949). *Sequential Analysis,* John Wiley, New York.

STA 402: DESIGN AND ANALYSIS OF EXPERIMENTS

(52 hours : 4 credits)

Unit 1
Introduction to design of experiments. Fixed, random and mixed effects models. General block designs: C matrix and its properties, concepts of connectedness, orthogonality, variance balance. Intra block analysis of general block design: Estimability, best point estimates, interval estimates of estimable linear parametric functions and testing of linear hypotheses. **10 hrs**

Unit 2
BIBD and: Parameter relationship, properties estimation and Intra block Analysis. Youden Square design: Intra block Analysis. Multiple comparison tests: Scheffe, Tukey, Duncan and Dunnett's procedures. Model adequacy checking. Missing plot techniques for RBD and LSD. **12 hrs**

Unit 3
Two-way classification with interaction: Analysis, Tukey's test for non-additivity, test for interaction in multiple but equal number of observations per cell. Two –way random effects model: variance estimation and tests of hypotheses. **8 hrs**

Unit 4
Factorial experiments: concepts, symmetric factorial experiments. Analysis of 2^n and 3^n factorial experiments in randomized blocks. Complete and partial confounding, Layout and analysis of confounded 2^n and 3^n factorials. Fractional replication for 2^n factorials.
16 hrs

Unit 5
Analysis of covariance for CRD and RBD. Split-plot experiments in RBD.
6 hrs

References

1. Dey, A. (1986). *Theory of Block Designs,* Wiley Eastern, New Delhi.
2. Dean, A. and Voss, D. (1999). *Design and Analysis of Experiments*, Springer, New York.
3. Das, M. N. and Giri, N. (1979). *Design and Analysis of Experiments*, Wiley Eastern, New Delhi.
4. John, P. W. M. (1971). *Statistical Design and Analysis of Experiments*, Macmillan, London.
5. Mukhopadhyay, P. (1998). *Applied Statistics,* Allied Publishers, New Delhi.
6. Joshi, D. D. (1987). *Linear Estimation and Design of Experiments*, Wiley Eastern, New Delhi.
7. Montgomery, D.C. (2004). *Design and Analysis of Experiments*, 5/e, Wiley, New York.
8. Pearce, S. C. (1984). *Design of Experiments,* Wiley, New York.
9. Chakrabarti, M.C. (1962). *Mathematics of Design and Analysis of Experiments*, Asia Publishing House, New Delhi.
10. Kempthorne, O. (1952). *Design and Analysis of Experiments*, Wiley Eastern, New Delhi.
11. Cochran, W.G. and Cox, G. M. (1957). *Experimental Designs,* 2/e, John Wiley, New York.

STA 405: Practical Paper

(Assignments based on STA 401, Elective – II and Elective – III)

(4 hours/ Week)

(2 credits)

1. Construction of SPRTs
2. OC and ASN curves for SPRTs
3. Non-parametric Tests
4. Bayes and Minimax Rules

5 – 7 Assignments from Elective II

8 – 10 Assignments from Elective III

STA 406: Practical Paper

(Assignments based on STA 402)

(4 hours / Week)

(2 credits)

1. Two – way ANOVA
2. Missing Plot Techniques for RBD and LSD
3. Analysis of Covariance for CRD
4. Analysis of Covariance for RBD
5. Analysis of BIBD and Youden square Design
6. Analysis of 2^3 and 2^4 factorial experiments
7. Analysis of 3^2 and 3^3 factorial experiments
8. Confounding in 2^3 factorial experiments
9. Confounding in 3^2 and 3^3 factorial experiments
10. Analysis of split – plot experiments.

ELECTIVE PAPERS

S1: SURVIVAL ANALYSIS

(52 hours : 4 credits)

Unit 1

Basic Elements and Parametric Interference: Parametric models for study of life time: Exponential, Raleigh, Weibull, extreme value, gamma, Pareto, logistic, normal and log – normal. Survival function, hazard rate, cumulative hazard function, and mean residual life. Longitudinal studies.

12 hrs

Unit 2

Censoring mechanisms: type I, type II and random censoring. Parametric estimation with complete and censored samples. Large sample tests under censored data. The E – M algorithm.

8 hrs

Unit 3

Nonparametric Inference: Life table, Actuarial Estimator, Kaplan – Meier (product – limit) estimation. Self consistency. Statement of asymptotic properties of K – M estimator. Nelson-Aalen Estimator. Treatment of ties (Peto's method). Weighted log – rank estimator. Two – sample methods. Gehan and Mantel – Haenszel test and Tarone-Ware tests.

14 hrs

Unit 4

Semi-parametric Inference: Cox proportional hazards model. The partial likelihood justification. Statement of asymptotic properties of the estimator. Estimation of the baseline hazard function. Residuals and model checking., Graphical methods: Hazard plots and Survival plots. **10 hrs**

Unit 5

Regression for grouped data: logistic and proportional hazards approaches. Accelerated models. Competing risk model and estimation of cumulative hazard function.

8 hrs

References

1. Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*, 2/e, Wiley, New York.
2. Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data,* Wiley, New York.
3. Nelson, W. (1982). *Applied Life Data Analysis*, Wiley, New York.
4. Miller, J. (1980). *Survival Analysis*, Wiley, New York.
5. Klein, J. P. (2003). Survival Analysis, Springer Verlag, New York.
6. Kleinbaum, D. G. (1997). *Survival Analysis,* Springer Verlag, New York.

S2: OPERATIONS RESEARCH

(52 hours : 4 credits)

Unit 1

Queueing systems: General description and characteristics of a queueing system. M/M/1 and M/M/c queueing systems and their waiting time distributions. M/M/1/N and $M/M/\infty$ queues. Transient solution of $M/M/\infty$ queueing system. Non-Markovian queues. Imbedded Markov chain analysis of M/G/1 and GI/M/1 queues. Pollackzek – Khintchine formula. **14 hrs**

Unit 2

Inventory models: Basic characteristics of inventory systems (models). ABC analysis Deterministic inventory systems EOQ Models with quantity discounts, price breaks, and storage limitations. Multiperiod dynamic inventory models. Continuous review stochastic inventory systems. The (s, S) policy. Multiperiod stochastic inventory systems.

12 hrs

Unit 3

Integer programming: Pure and mixed Integer programming problems. Cutting plane methods – Gomory's algorithms. Branch and bound technique. Zero-one programming.

10 hrs

Unit 4

Nonlinear programming: Formulation of nonlinear programs. Unconstrained and constrained optimization problems. The Lagrangian method. Karush-Kuhn-Tucker optimality conditions. Quadratic programming. Wolfe's modified simplex method. Sequential unconstrained minimization technique (SUMT). **8 hrs**

Unit 5

Dynamic programming: Multistage dynamic programming. Bellman's principle of optimality. General formulation. Forward and backward recursion, computational methods, and applications of dynamic programming. Dynamic lot size determination. **8 hrs**

References

1. Gross, D and Harris, C. M. (1986). *Fundamentals of Queueing Theory*, 2/e, John Wiley.
2. Taha, H. A. (2002). *Operations Research*, 7/e; Macmillan.
3. Medhi. J. (1991). *Stochastic models in queueing theory*, Academic Press.
4. Bazaara, M. S. and Shetty, C. M. (1979). *Nonlinear Programming: Theory and Algorithms*, John Wiley, New York.
5. Hillier, F. S. and Liebermann, G. J. (1986). *Introduction to Operations Research,* Holden Day, New York.
6. Kambo, N. S. (1991). *Mathematical Programming Techniques*, Affiliated East-West Press, New Delhi.
7. Murthy, K. G. (1995). *Operations Research: Deterministic Optimization Models*, Prentice Hall, New Delhi.
8. Swarup, K. et. al. (1985). *Operations Research*, Sultan Chand and Co., New Delhi.
9. Wayne, L. W. (1996). *Introduction to Mathematical Programming*, 2/e, Duxbury Press, New York.

S3: TIME SERIES ANALYSIS

(52 hours : 4 credits)

Unit 1

Exploratory Time Series Analysis, test for randomness, Tests for trend and seasonality. Estimation of trend by moving average, estimation of seasonal effect for additive and multiplicative models, deseasonalising and detrending an observed time series.

6 hrs

Unit 2

Time-series(t.s) as discrete parameter stochastic process, definition of strict and weak stationarity of a t.s., Gaussian t.s., ergodicity, autocovariance and autocorrelation functions(ACF) and their properties, partial autocorrelation function(PACF). Genaral linear processes(G.L.P), autocovariance generating function, stationarity and invertibility conditions of a G.L.P; autoregressive processes(AR(p)), stationarity condition, ACF, PACF, Yule-Walker equations, Moving average (MA(q)) processes, Invertibility condition, ACF, PACF, duality between AR(p) and MA(q) processes; ARMA(p,q) processes, stationarity, invbertibility, ACF, PACF, particular cases of these processes.

18 hrs

Unit 3

Linear Non-stationary time Series models: ARIMA(p,d,q) processes, general form, three explicit forms, IMA(0,1,1) process, seasonal ARIMA processes. Forecasting : minimum mean square error forecast, BLUP, three basic forms for the forecast, forecast error and its properties, examples; forecasting through exponential and Holt-Winter smoothing.

12 hrs

Unit 4

Estimation: sample ACF. Sample PACF, fitting AR(p), MA(q), ARMA(p,q) models; model identification: determination of p, d, q: method of differencing, unit root test, using sample ACF, sample PACF, Bartlett, Quenouille and Anderson bounds; diagnostics: residual analysis, Box-Pierce postmanteau statistic, Ljung-Box test; AIC and BIC criteria.

10 hrs

Unit 5

Time series models of heterocsedasticity: Some common features of financial t.s., ARCH and GARCH models, test for ARCH effect, maximum likelihood estimation.

6 hrs

References

1. Anderson, T. W. (1971). *The Statistical Analysis of Time Series*, Wiley, New York.
2. Box, G. E. P, Jenkins, G. M, Reinsel, G. C. and Ljung, G. M. (2015). *Time Series Analysis - Forecasting and Control*, 5/e, Wiley.
3. Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*, 2/e, Indian Print, Springer, New Delhi.
4. Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*, 2/e, Springer, New York.
5. Chatfield, C. (1996). *The Analysis of Time Series: Theory and Practice*, 5/e, Chapman and Hall, London.
6. Chatfield, C. (2003). *Analysis of Time Series: An Introduction*, CRC Press, New Delhi.
7. Nachane, D. M. (2006). *Econometrics: Theoretical Foundations and Empirical Perspectives*, Oxford University Press, London.
8. Cryer, J. D. and Chan, K. S. (2008). *Time Series Analysis with Application in R*, 2/e, Springer, New York.
9. Kendall, M. G. and Ord, J. K. (1990). *Time Series*, 3/e, Edward Arnold, New York.
10. Montgemory, D. C. and Johnson, L. A. (1977). *Forecasting and Time Series Analysis*, McGrawHill, New York.

S4: STATISTICS OF FINANCE AND INSURANCE

(52 hours : 4 credits)

Unit 1

Basic concepts of financial markets - stocks, shares and assets. Put and call options, exercise price and exercise time. Arbitrage and hedging. Interest rates. Returns and log-returns. Random walk model, lognormal model. Call options. Law of one price. Simple binomial model and multiperiod binomial model. Option pricing. Arbitrage theorem.

12 hrs

Unit 2

Brownian motion and geometric Brownian motion models. Applications in stock market analysis. The Black-Scholes formula and its properties. The Greeks. Delta hedging arbitrage strategy. Volatality and estimating the volatility parameter. **10 hrs**

Unit 3

Portfolio theory- trading off expected return and risk, one risky asset and one risk free asset. Two risky assets, estimated expected return. Optimal mix of portfolio. Value at Risk (VAR). Estimation of VAR. **8 hrs**

Unit 4

Basic concepts of insurance portfolio. Individual claims and aggregate claim. Aggregate claim model for single period. Security loading and relative security loading. Adjustment coefficient. Discrete and continuous time surplus models. Probability of ruin. Compound Poisson process in continuous time surplus model. Relationship between ruin probability and adjustment coefficient. Bounds for ruin probability. Lundberg inequality and its inpretation in exponential claims. **14 hrs**

Unit 5

Fisrt surplus below the initial level and its its probability density function. Maximal aggregate loss. MGFs of the maximal aggregate loss. Claim Distributions, modeling of individual and aggregate claims. Fitting distributions to claim data, approximation to the distribution of total claim using the central limit theorem and Monte Carlo techniques. Light and heavy tailed distributions as models of claim amounts. Millls ratio. **8 hrs**

References

1. Baxtor, M. and Rennie, A. (2004). *Financial Calculus,* Cambridege University Press, UK.
2. Bouchard, J. P. and Potters, M. (2000). *Theory of Financial Risks*, Cambridge University Press, UK.
3. Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A. and Nesbnitt, C. J. (1997). *Actuarial Mathematics*, 2/e, Society of Actuaries, London.
4. Dickson, (2005). *Insurance Risk and Ruin,* Cambridge University Press, Cambridge, UK.
5. Etheridge, A. (2002). *A Course in Financial Calculus*, Cambridge University Press, UK.
6. Hull, J. C. (1993). *Options, Futures, and Other Derivative Securities*, 2/e, Prentice Hall of India, New Delhi.
7. Mikosch. T. (2004). *Non-life Insurance Mathematics*, Springer, New York.
8. Ross, S. M. (2004). *Introduction to Mathematical Finance,* Cambridge University Press, UK.

S5: NONPARAMETRIC STATISTICAL INFERENCE

(52 hours : 4 credits)

Unit 1

Introduction: Equal in distribution technique, Distribution-free over a class, Counting and Ranking statistics, Kolmogorov-Smirnov statistics for one and sample problems. **6 hrs**

Unit 2

U-statistics: Definition of one sample U-statistics, Projection principle and one sample U-statistics theorem, Two and k-sample U-statistics, Two sample U-statistics theorem, Asymptotic distribution of k-sample U-statistics, Joint distribution of k-sample U-statistics.
Linear rank statistics: Definition, examples of linear rank statistics, mean and variance of linear rank statistics, Distributional properties of linear rank statistics, Asymptotic normality of linear rank statistics – related results and Hajek's theorem. **18 hrs**

Unit 3

Asymptotic relative efficiency: Definition of Pitman asymptotic relative efficiency (ARE), Methods of evaluating Pitman ARE - Noether theorem, Examples of Pitman ARE for translation alternatives. **8 hrs**

Unit 4

Locally most powerful rank(LMPR) tests : Locally most powerful rank(LMPR) tests for one and two sample problems, asymptotic properties of LMPR statistic, Wilcoxon test, normal scores test, superiority of normal scores test over t-test in the sense of ARE, LMPR test for bivariate independence. **10 hrs**

Unit 5

Nonparametric Regression: Nature and scope of nonparametric regression Basic idea of smoothing, Smoothing histograms and nonparametric probability density function. Random design and fixed design model, Bin smoothers, running mean and running line smoothes. Univariate kernel density estimation. Local linear regression estimate.

10 hrs

References

1. Gibbons, J.D. and Chakraborti, S. (2004). *Nonparametric Statistical Inference,* 4/e, Marcel Dekker, New York.
2. Hajek, J., Sidak, Z. and Sen, P.K. (1999).*Theory of Rank Tests,* 2/e, Academic Press.
3. Hardle (1990).Applied Nonparametric Regression. Cambridge University Press, New York.
4. Hetmansperger, T.P. (1984). *Statistical Inference Based on Ranks,* John Wiley, New York.
5. Randles, R.H. and Wolfe, D.A. (1979). *Introduction to the Theory of Nonparametric Statistics,* John Wiley, New York.
6. Takezawa, K.(2005). *Introduction to Nonparametric Regression*, John Wiley, New York.

S6: ADVANCED SAMPLING THEORY

(52 hours : 4 credits)

Unit 1

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 techniques.

10 hrs

Unit 2

Introduction to the unified theory of finite population sampling and super population models.

6 hrs

Unit 3

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

11 hrs

Unit 4

Issues in small area estimation – synthetic and generalized regression estimators. Errors in surveys. Non-sampling errors and biased responses, randomized responses for variables, modeling observational errors, estimation of variance components, application to longitudinal studies, repetitive surveys.

13 hrs

Unit 5

Variance estimation, method of random groups, balanced half samples (IPNSS), Jack knife method. Design of field surveys. Questionnaire designing. Interviewing. Preparation of reports.

12 hrs

References

1. Chaudhuri, A. and Vos, J.W.E. (1988). *Unified Theory and Strategies of Survey Sampling*, North-Holland, Amsterdam.
2. Chaudhuri, A. and Mukerjee, R. (1988). *Randomized Response: Theory and Techniques,* Marcel Dekker, New York.
3. Cochran, W.G. (1984). *Sampling Techniques*, 3/e; John Wiley, New York.
4. Des Raj and Chandok, P. (1998). *Sampling Theory,* Narosa, New Delhi.
5. Hedayat, A.S. and Sinha, B.K. (1991). *Design and Inference in Finite Population Sampling,* John Wiley, New York.

6. Mukhopadhyay, P. (1996). *Inferential Problems in Survey Sampling*, New Age International, New Delhi.

7. Mukhopadhyay, P. (1998). *Small Area Estimation in Survey Sampling*, Narosa, New Delhi.

8. Murthy, M.N. (1977). *Sampling Theory and Methods*, Statistical Publishing House, Calcutta.

9. Wolter, K.M. (1985). *Introduction to Variance Estimation*, Springer Verlag, New York.

* * * * *