University of Pune

M.Sc.Part- II (Statistics)

Semester – III (With Effect from 2012-13)

1) ST - 31 Elementary Stochastic Processes
2) ST - 32 Design of Experiments
3) ST - 33 Statistical Inference II
4) ST - 34 (A) Data Mining (Departmental Course)
   Or
   (B) Analysis of Clinical Trials (Departmental Course)
   Or
   (C) Advanced Probability
5) ST - 35 Practicals

Semester – IV (With Effect from 2012-13)

1) ST - 41 Optimization Techniques
2) ST – 42 Statistical Process Control (SPC) and Survival Analysis
3) ST - 43 Time Series Analysis
4) ST- 44 (A) Actuarial Statistics (Departmental Course)
   Or
   (B) Stochastic Models in Finance (Departmental Course)
5) ST- 45 Practicals (50% weightage) and Project (50% weightage)

M.Sc.Part- II
SEMESTER III

ST 31 : Elementary Stochastic Process

(i) Markov chain with stationary transition probabilities, Chapman Kolmogorov equation, n-step transition probabilities, classification of states (persistent, transient, non null, null persistent, periodic, aperiodic), ergodic Markov chain, stationary distribution of Markov chain, existence and uniqueness of stationary distribution, interpretation of stationary probability.

(ii) Random walk, random walk with absorbing and reflecting barrier, classification of states, probability of absorption in persistent class starting from transient state, application to gambler's ruin problem: probability of ruin, expected gain, expected duration of a game.

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

(iv) Markov pure jump process, continuous time Markov process, Kolmogorov forward and backward equations, Poisson process, pure birth process, birth and death process, properties and results associated with these processes.

(v) Introduction to renewal process, key renewal and elementary theorem (without proof), renewal equations and its solution(without proof), Poisson process as a renewal process, applications of renewal process in reliability theory.

(vi) Elementary queuing models: M/M/1, M/M/K, M/G/K queuing systems, relationship between queuing models and birth and death process, applications of these to real life problems.

(vii) Introduction to Brownian motion, Weiner process and its properties.

Books Recomended:

(1) Medhi, J. (1982) Stochastic processes (Wiley Eastern)

(2) Taylor, H N and Karlin, S. (1984) An introduction to stochastic modeling(Academic Press)

(3) Ross, S. (2000) Introduction to probability models, 7th edn (Academic Press)

(4) Srinivas and Mehta (1976) Stochastic Processes (Tata Mc-Graw Hill)

(5) Ross, S. ( ) Stochastic processes (John Wiley)

(6) Adke, S.R. and Manjunath, S.M. (1984) An introduction to finite Markov processes (Wiley Eastern)

(7) Bhat, B.R. (2000) Stochastic models: Analysis and applications (New Age International)

(8) Athreya and Lahiri ( ) Probability theory

ST 32 :Design of Experiments

1. General Linear model : Definition, assumptions, (5L)
   concept of estimability. Least sq. estimation, BLUE,
   error space, Gauss Markov theorem, Variance and correlation
   of BLUE's.

2. ANOVA : One way and two way classification with (5L)
   more than one but equal number of observations per cell
   (with or without interaction) Estimation and related
   Test of hypotheses Tuckey's test of additivity. Connectedness
   balancedness, orthogonality.

3. 2^k full factorial experiment : diagrammatic representation (7L)
   of main effect, 1^{st} and 2^{nd} order interactions, model,
   analysis of single more than one replicates using ANOVA.

4. Total confounding of 2^k design in 2^p blocks, $p \geq 2$, partial (10L)
   confounding in 2^p blocks, $p = 2,3$.
   Fractional factorial experiments. Resolution III, IV and V
   of a design aberration of a design.
5. 3^k designs : Contrasts for linear and quadratic effects, statistical (6L)
   analysis of 3^k design , confounding and factorial experiments in
   3^k design.
6. RSM : Linear and quadratic model, stationary pt. CCD. (6L)
7. Taguchi method : Concept of loss function S?N ratio (7L)
   Linear graphs, inner and outer arrays, ANNOVA
8. Random effects model for one way ANNOVA. (4L)

Books Recomended:

1. Kshirsagar A.M. : Linear Models (Marcel Dekker)
2. Montgomery D.C. : Design and Analysis of Experiments (John Wiley)
3. Morrison D.F. : Multivariate Statistical Methods(Mc Graw Hill)
4. Rao C.R.: Linear Statistical Inference and its Application (Wiley Eastern)
5. Johnson and Wichern : Applied Multivariate Statistical Analysis (Prentic Hall Inc)
6. John P.W.M. (1971) Linear Models) (John Wiley Ltd.)
7. Dean A. and Voss, D. (1999). Design and Analysis of Experiments (Springer)
8. Hicks , C.R. and Turner, K.V. (1999) Fundamental Concepts in the Design of Experiments. (Oxford University Press)

ST 33: Statistical Inference II

1. Consistent estimation of real and vector valued parameters. Invariance of consistent estimator under continuous transformation, consistency of estimators by methods of moments and method of percentiles, mean squared error criterion, Asymptotic relative efficiency, error
2. probabilities and their rates of convergence. (16 L)

3. Consistent Asymptotic Normal (CAN) estimator, invariance of CAN estimator under differentiable transformation, CAN property of estimators obtained by moments and percentiles, CAN estimators obtained by moments and MLE method in one parameter exponential family, extension to multiparameter exponential family, examples of consistent but not asymptotically normal estimators from pitman family Cramer – Huzurbarar theorem. (16L)

4. likelihood ratio test, asymptotic distribution of LRT statistic, Wald's test, Rao's score test, Pearson's X^2 test for variances, large sample tests and confidence intervals based on CAN estimators, variance stabilizing transformation and large sample tests, consistency of large sample tests, asymptotic power of large sample tests. (16L)

5. Empirical distribution function, Glivenko – Cantelli theorem. Review of Kolmogorov goodness of fit test, sign test.
   Wilcoxon's signed rank test, two sample Kolmogorov Smirnov test, Kruskal – Wallis K- sample test. (12L)

References

(1) Kale, B.K (1999) A first course on parametric inference, Narosa Publishing House.
(2) Rohatgi V.K and Saleh A.K. Md. E (2001):Introduction to Probability Theory and mathematical Statistics, John – Wiley and Sons.
(3) Lehmann, E.L. (1986): Testing Statistical Hypotheses.
(4) Dudewicz E.J and Mishra S.N (1988) Modern Mathematical statistics, John Wiley and Sons.
(5) Ferguson T.S (1996): A course on large sample Theory, Chapman and Hall.
(6) Gibbons J.D. (1985): Non parametric statistical inference 2nd Ed. Marcal Dekker Inc.

ST 34 (A) : Data Mining (Departmental Course)

1.Review of classification methods from multivariate analysis, classification and decision trees.(10L)
2.Clustering methods from both statistical and data mining viewpoints, vector quantization. (10L)
3.Unsupervised learning from univariate and multivariate data, Dimension reduction and feature selection. (10L)
4.Supervised learning from moderate to high dimensional input spaces, artificial neural networks and extensions of regression models, regression trees. (10L)

5.Introduction to databases, including simple relational databases, data warehouses and introduction to online analytical data processing (10L)

Association rules and prediction, data attributes, applications to electronic commerce.

Books Recommended

1. Berson, A. and Smith, S.J. (1997). Data Warehousing, Data Mining, and OLAP. (McGraw-Hill.)
2. Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). Classification and Regression Trees. (Wadsworth and Brooks/Cole).
3.Han,J. and Kamber.M. (2000). Data Mining; Concepts and Techniques. (Morgan Kaufmann.)
4. Mitchell, T.M. (1997). Machine Learning. (McGraw-Hill.)
5. Ripley, B.D. (1996). Pattern Recognition and Neural Networks. (Cambridge University Press).

ST 34 (B) : Analysis of Clinical Trials (Departmental Course)

Introduction to clinical trials : The need and either of clinical trails, bias and random error in clinical studies, conduct of clinical trials, overview of Phase I – IV trails, multi- centre trails. (13L)

Data management : data definitions, case report forms, database design, data collection system for good clinical practice.
(13L)

Design of clinical trials : parallel vs cross- over designs, cross- sectional designs cross- sectional vs longitudinal designs, review of factorial designs, objectives and end points of clinical trails, design of Phase I trails, Design of single –stage and multi- stage Phase II trails, design and monitoring of Phase – II trails with sequential stopping, design of bio- equivalence trails. (12L)

Reporting and analysis : analysis of categorical outcomes from phase I – III trails, analysis of survival data from clinical trials. (12L)

Surrogate endpoints: selection and design of trails with surrogate endpoint data. Meta – analysis of clinical trails.

Books Recommended :

1. S. Piantadosi (1997) Clinical Trails: A methodological Perspective. Wiley and sons.
2. C. Jennison and B.W. Turnbull (1999). Group sequential methods with application to Clinical Trails CRC Press.
3. L.M. Friedman, C. Furburg, D.L Demets (1998) Fundamentals of Clinical Trials, Springer Verlag.
4. J.L. Fleiss(1989). The design and Analysis of Clinical Experiments. Wiley and sons.
5. E. Marubeni and M.G. Valsecchi (1994) Analyzing Survival Data Clinical Trials and Observational Studies, Wiley and sons.

ST 34 (C) Advanced Probability (Departmental Course)

Ring, σ-ring, Measure space, Caretheodory Extension theorem, Lebesgue measure, integral of a measurable function with respect to a measure, its properties. Hahn- Jordan decomposition, Lebesgue decomposition, Randon – Nikodym derivative, Product measure, Fubini's theorem. Convergence in measure, almost everywhere convergence, Kolmogorov Inequality. Kolmogorov three series criterion strong law of large numbers.

Conditional Probability and conditional expectations, their simple properties. Martingales (discrete parameter).

Books Recommended :

1. Billingsley, P. (1986) Probability and Measure (John Wiley and Sons)
2. Ash, R.B.(1972) Real Analysis and Probability (Academic Press)
3. Kingman and Taylor(1968) Introduction to Measure and Probability (Cambridge Univ. Press)

Additional Books for Reference :

1. Halmos P.R. (1962)

ST 35: Practicals

Titles of Experiments

1. Model sampling from MVN distribution, MLE's of parameters of MVN distribution
2. Multivariate Analysis (Wishart matrix contour plots)
3. Hoteling's T square test ,Likelihood Ratio Test
4. MANOVA and Discriminant analysis.
5. Principal Componant Analysis (PCA), Canonical correlation.
6. Cluster Analysis
7. One way classification, multiple comparison test.
8. Two way classification with (i) equal (ii) unequal number of observations per cell (model with interaction), connectedness.
9. LSD and BIBD
10. Analysis of covariance. (One way and Two way Analysis)
11. 2^k factorial Experiments, analysis of single replicate of 2^k.
12. Total and partial confounding2^k experiments.
13. Fractional experiments, Plackett- Burman Design.
14. 3^k factorial Experiments confounding and factorial experiments 3^k designs.
15. Random effect and mixed models (restricted and unrestricted).
16. Nested designs and mixed split plot designs.
17. Fitting first and second response surface model. Central Composite Design, Contour, surface plots, Spherical CCD, Box- Behnken design, face centered CCD.
18. Small composite designs, blocking in RSM, optimal designs, simplex lattice designs, simplex centroid designs.
19. Taguchi methods: S/N ratio orthogonal arrays, triangular tables, linear graphs, inner and outer arrays.
20. Non parametric Test.

M.Sc. Part- II
SEMESTER IV

ST 41 : Optimization Techniques

1.Linear Programming : Convex sets, Supporting and Separating Hyper- planes, Standard Linear Programming, Problem, basic feasible solution, simplex algorithm and simplex method, graphical solution, two phase method. Duality in linear programming, duality theorems, dual simplex method with justification, sensitivity. (10L)

2.Transportation and assignment algorithms, Hungarian method of assignment, transshipment problems, duality theory of testing optimality of solution in transportation problem and transshipment problems, transportation problem and transshipment problems as network problems. Balance and degeneracy in transportation problem. (10L)

3.Maximization, prohibitions and other variations of assignment problems. Integer linear Programming, Problem, branch and bound method, network. (10L)

4.Nonlinear Programming : Kuhn- Tucker conditions, Quadratic programming, Wolf's, Beale's and Fletcher's algorithms for solving quadratic programming problems. (10L)

5.Markovian and Non- Markovian queuing models, cost profit models of (M/M/l) and (M/M/s) queueing systems. Simulation, event type simulation, simulation of queuing systems.
Dynamic programming. (10L)

Books Recommended :

1. Kambo N.S. (1991) Mathematical Programming Techniques(Affiliated as west press Pvt. Ltd.)
2. Hadly, G. (1987) Linear Programming.
3. Taha, H.A. (1992) Operations Research 5th ed.(Macmillan)
4. Panneerselvam, R. Operations Research (Prentic hall of India)
5. Medhi J. (1984) Stochastic Processes 2nd edition (New International Pvt. Ltd.

ST : 42 Statistical Process control (SPC) and Survival Analysis.

Section – I Statistical Process control (SPC)

1.Total quality Management : Concept of quality, Quality improvement, Quality Philosophy, Introduction to TQM, six sigma and other Extension of TQM, quality systems, The ISO 9000 and other Quality systems. (4L)

2.Statistical Process Control : Concepts of stable industrial processes, Systematic variation, random variation, revision of theory of control charts,CUSUM, EWMA charts. Comparison of Shewhart control charts with CUSUM charts. General ideas on economic designing of control charts. Duncan's model for the economic control chart. Process capability and performance indices C_p C_{pk}. Estimation and confidence intervals of estimators of C_p. Connection between proportion of defectives and C_p. De merit control chart for number of defects. (8L)

3.Confirming run length (CRL) chart for attributes, Synthetic control chart Steady state model ARL ATS, Hoteling T square multivariate control chart. (5L)

4. Acceptance Sampling Plan for attributes : Double and multiple sampling plans for attributes. Curtailed sampling plan, Operating characteristic functions. (3L)

5.Acceptance sampling plan for variable **:** Designing variable acceptance sampling plans. AQL based sampling plans. Continuous Sampling Plans, chain sampling and sequential sampling plan. (5L)

Section- II Survival Analysis

6.Concepts of Time, Order and random Censoring.
Life distribution – Exponential gamma, Weibull, Lognormal, Pareto, Linear Failure rate. Likelihood based parametric inference Point estimation, Confidence Interval. (9L)

7.Failure rate, mean residual life and their elementary properties. Aging classes- IFR, IFRA, NBU, NBUE, HNBUE and their duals, Bathtub Failure rate.
Estimation of survival function – Actuarial Estimator, Kaplan- Meier Estimator, Empirical survival function (8L)

8.Test of expopnentiality against non-parametric classes- within NBU classes and NBUE classes Total time on test, Deshpande test. Two sample problem- Gehan Test, Log rank test. (8L)

Books Recommended :

1. Montgomery, D.C. (1985) Introduction to Statistical Quality Control (Wiley)
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. Oakland J.S. (1989) Total Quality Management: Butterworth – Heinemann.
5. 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
6. Wu, Yeu and Spedding (2001) Asymptotic control charts for detecting fraction non confirming increases JQT 33 (1)104-111
7. Cox, D.R. and Oaks,D.)1984) Analysis of Survival data, Chapman and Hall, New York.
8. Gross, A.J// and Clark, V.A. (1975) Survival Distiributions : Reliability Applications in the Biomedical Sciences, John Wiley and Sons.
9. Elandt- Johnson R.E. Johnson N.L. (1980) Survival models and Data Analysis , John Wiley and sons
10. Miller, R.G. (19981) Survival Analysis (Wiley)
11. Zacks, S. Reliabilty.
12. J.V. Deshpande and Sudha Purohhit : Life time data : Statistical models and methods, word scientific Publisher
13. Chap T.L.E. Applied Survival Analysis : John Wiley and Sons
14. J.V. Deshpande, A. P. Gore (1995) A. Shan bhouge : Statistical Analysis of Non- normal data : New Age international Ltd.

ST 43: Time series analysis

Time series as discrete parameter stochastic process

1. Features of time series Data :Trend, Seasonality, auto covariance and auto correlation function and their propertier two exploratory time – series analysis (10L)

2. Test for randomness of a series against trend and seasonality M.A and exponential smoothing, Hot winters smoothing, forecasting based on smoothing, adaptive smoothing. (10L)

3. Study of stationary process – (a)moving average (M.A), (b)auto regressive (AR), (c)ARMA and (d)ARIMA model box – Jenkins model (12L)
   Discussion (without proof) of estimation of means, auto covariance and auto correlation function under large sample theory.
   Choice of AR and MA periods.
   Estimation of ARIMA model parameters.
   Forecasting, residual analysis and diagnostic checking.

4. Introduction to spectral analysis of weekly stationary process.
   Periodogram and correlogram analysis.
   (8L)

5. Non – stationary and seasonal time models :Unit – root non – stationarits, unit – root test, integrated ARMA (ARIMA) models seasonal ARIMA (SARIMA) models.
   Conditional heteroschedastic models: volatility models, ARCH and GARCH, properties, examples, estimation and forecasting.
   Multivariate time series model, VAR models, Vector ARMA model, co integration models. (20L)

Books recommended:

1. Anderson T.W :The statistical analysis of time series Wiley.
2. Box G.E.P and Jenkins G.M :Time series analysis – forecasting and control, holdan – day, san Francisco.
3. Montgomery DC and Johnson LA :Forecasting and time series analysis, Mcgrew Hill.
4. Kendall and Ord :Time series (3rd Edition), Edward Arnold.
5. Brockwell P.J and Daris R.A :Time series : Theory and methods springer – verlag.

ST 44 (A) Actuarial Statistics (Departmental Course)

1.Utility theory, insurance and utility theory, models for individual claims and their sums, survival function, curtate future lifetime, force of mortality. (6L)

2.Life and its relation with survival function, examples, assumptions for fractional ages, some analytical laws of mortality, select and ultimate tables. (6L)

3.Principles of compound interest : Nominal and effective rates of interest and discount, compound interest, accumulation factor, continuous compounding. (6L)

4.Life insurance : Insurance payable at the moment's of death- level benefit insurance, endowment insurance, differed insurance, and varying benefit insurance. (6L)

5.Life annuities : Continuous life annuities, discrete life annuities, life annuities with monthly payments.

(6L)

6.Net premiums : Continues and discrete premiums, true monthly payment premiums.
Some practical considerations : Premiums that include expenses types of expenses, per policy expenses.

(5L)

7.Net premium reserves : Continuous and discrete net premium reserve, reserves on semi continuous basis, reserves based on true monthly premiums.

(5L)

8.Multiple life functions , joint life and last survivor states, insurance and annuity benefits through multiple life functions evaluation for special mortality laws.

(5L)

9.Distribution of aggregate claims, compound Poisson distribution and its applications. Distribution of aggregate claims, compound Poisson distribution and its application. (5L)

Books Recommended :

1. N.L. Bowers, H. U. Gerber, J.C. Hickman, D.A. Jones and C.J. Nesbitt, (1986), Actuarial Mathematics, Society of Actuaries, Itasca, IIIinios , U.S.A. Second Edition (1997)
   Section I -Chapters : 1,2,3,8,9 and 11
   Section II – Chapters : 4,5,6,7,13 and 14
2. Deshmukh S.R. (2009) An introduction to Actuarial Satistics Using R University Press.
3. Spurgeon E.T. (1972), Life Contingencies Cambridge University Press.
4. Neill, A.(1977) Life Contingencies, Heinemann.

ST 44 (B) Stochastic Models in Finance (Departmental Course)

1.Derivatives : forward and future contracts. Markets, prices, arbitrage and hedging. Complete market, market risk and credit risks in the use of derivatives. (10L)

2.Options Markets, properties of stock option prices, American and European options.
Binomial model : One step and two- step models, Binomial trees. Risk neutral valuations. (10L)

3.Behavior of stock prices : Conditional expectation, Martingale, Brownian Motion and Geometric Brownian motion, Markov property, Ito integral, Ito/diffusion and Mean- reverting processes, Processes, Ito Lemma. (10L)

4.Black Scholes model : Distribution of returns, volatility, risk neutral pricing equivalent martingale measure, Black- Scholes- Merton different equation. Estimating volatility. (10L)

5.Options on stock indices, currencies and futures. Some exotic equity and foreign exchange derivatives, Black model. Models of the term structure of interest retes : one factor diffusion model, Vasicek, Cox- Ingersoll- Ross and Hull white models. (10L)

Books Recommended:

1. John Hull ,(2008) Options, futuresand other derivatives, (International 7th Editional) Prentic Hall.
2. M. Baxter and A. Rennie, (1996) Financial Calculas, Cambridge University/ Press
3. N. Bingham and R. Keisel, (1998) Risk- Neutral Valuation, Springer.

ST 45 : Practicals (50marks) and Project (50marks)

A: Practicals

Title of Practicals

1. Smoothing the series using various Filters : MA filter, Exponential Smoothing, Other filters, Data Transfer : Box- Cox transformation, Differencing : Checking stationary 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 AIC, BIC, Fitting of AR, MA, ARMA, AIMA and SARIMA models : (conditional least squares or max. likelihood)
4. Forecasting :Using fitted linear (recursively), Holt- Winters forecasts, Construction of forecast intervals, Fitting heteroscedastic models : Checking for hetroscdeicity from residuals, Arch, GARCH modeling.
5. Non linear programming.
6. Computation of design parameters of synthetic control charts for means control limits for $\overline{X}$ chart when σ is known and capability indices.
7. Implementation of CUSUM by using both the methods.
8. Implementation of EWMA chart.
9. Obtaining design parabola of SSP, DSP and curtailed DSP.
10. Survival Analysis I
11. Survival Analysis II.

B: Project

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 articles (number 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 pairs of student however summary research article have to done individually

---

Examination: 1. Practical Examination of 50 marks(10marks for internal evaluation and 40 marks for final examination by University)with duration 1.5 hours will be conducted as usual.

2.Project evaluation (50 marks) (10marks for internal evaluation and 40 marks for final examination by University) will be done separately by conducting presentation and viva