SAVITRIBAI PHULE PUNE UNIVERSITY

(*Formerly University of Pune*)

Three Year B. Sc. Degree Course in Statistics
(Faculty of Science and Technology)

T. Y. B. Sc. STATISTICS

CHOICE BASED CREDIT SYSTEM SYLLABUS FOR THIRD YEAR
(To be implemented from Academic Year 2021-22)

Course Structure
T. Y. B. Sc. Statistics

Semester V	Semester VI	Credits	Marks	
			CIA	ESE
ST 351: Distribution Theory – I	ST 361: Distribution Theory – II	2	15	35
ST 352: Theory of Estimation	ST 362: Testing of Hypothesis	2	15	35
ST 353: Design and Analysis of Experiments	ST 363: Sampling Theory	2	15	35
ST 354: Statistical Process and Product Control	ST 364: Introduction to Survival Analysis	2	15	35
ST 355: Operations Research – I	ST 365 (A): Actuarial Statistics OR ST 365 (B): Operations Research – II	2	15	35
ST 356: Regression Analysis	ST 366 (A): Stochastic Processes OR ST 366 (B): Reliability Theory and Applications OR ST 366 (C): Medical Statistics and Clinical Trials	2	15	35
ST 357: Practical Paper – I	ST 367: Practical Paper – IV	2	15	35
ST 358: Practical Paper – II	ST 368: Practical Paper – V	2	15	35
ST 359: Practical Paper – III	ST 369: Project	2	15	35
Skill Enhancement Courses (SEC)				
ST 3510: SEC 1: Turbo C **(Practical Course)**	ST 3610: SEC 3: Introduction to Python **(Practical Course)**	2	15	35
ST 3511: SEC 2: Statistical Computing using R-software **(Practical Course)**	ST 3611: SEC 4: Data Analytics **(Practical Course)**	2	15	35

Equivalence for old courses (from 2014-15) with new courses (w.e.f. from 2021-22) in Statistics:

Semester V	
Papers in Old Syllabus (2014–15 to 2019–20)	Equivalent papers in New Syllabus (2019–20 onwards)
ST331: Distribution Theory	No equivalent paper *
ST 332: Theory of estimation	ST 352: Theory of Estimation
ST 333: Sampling Methods	ST 363: Sampling Theory
ST 334: Design of Experiments	ST 353: Design and Analysis of Experiments
ST 335: C Programming (Turbo C)	ST 3510: SEC 1: Turbo C
ST 336: Introduction to Regression Analysis	ST 356: Regression Analysis

Semester VI	
Papers in Old Syllabus (2014–15 to 2019–20)	Equivalent papers in New Syllabus (2019–20 onwards)
ST 341: Actuarial Statistics	ST 365 (A): Actuarial Statistics
ST 342: Testing of Hypotheses	ST 362: Testing of Hypothesis
ST 343: Statistical Quality Control	ST 354: Statistical Process and Product Control
ST 344: Operations Research	ST 355: Operations Research-I
ST 345 (A): Reliability and Survival Analysis	No equivalent paper *
ST 345 (B): Introduction to Stochastic processes	ST 366 (A): Stochastic Processes
ST 346: Statistical Computing Using R Software	ST 3511: SEC 2:Statistical Computing using R software
ST 347 : Practical Paper – I	No equivalent paper *
ST 348 : Practical Paper – II	No equivalent paper *
ST 349 : Practical Paper – III	No equivalent paper *

*The Examination of the papers having 'No equivalent papers' will be conducted as per University provisions prescribed in such cases.

GENERAL INFORMATION

1. A student of the three year B.Sc. degree course will not be allowed to offer Statistics and Statistical Techniques simultaneously in any of the three years of the course.
2. Students offering Statistics at the first year of the three year B.Sc. course may be allowed to offer Statistical Techniques as one of their subjects in the second year of the three year B.Sc. course, in place of Statistics.
3. Students offering Statistical Techniques at the first year of the three year B.Sc. course may be allowed to offer Statistics as one of their subjects in the second year of the three year B.Sc. course in place of Statistical Techniques, provided they satisfy other requirements regarding subject combinations, if any.
4. Each theory courseis of 2 credits. There will be 03 lectures per week each of 50 minutes. Each practical course is of 1.5 credits. The duration of each practical course is 04 hour 20 minutes per week.
5. **Students must complete all the practicals (experiments) in each of the practical papers including skill enhancement courses(SEC) to the satisfaction of the teachers concerned.**
6. Students must produce the laboratory journal along with the completion certificate duly signed by the Head of the Departmentat the time of practical examination.
7. The practical examination will be conducted semester wise. Internal and external examiner will examine the candidates answer books jointly after examination is over at the centre.
8. In **semester VI**, a student shall opt for any one of the following papers:
   i) ST 365(A): Actuarial Statistics **OR** ST 365(B): Operations Research – II
   ii) ST 366 (A): Stochastic Processes **OR** ST 366(B): Reliability Theory and Applications **OR** ST 366 (C): Medical Statistics and Clinical Trials.
   All other papers of semester VI are compulsory.
9. The theory question paper for each theory paper shall cover all the topics in the pertaining syllabus with proportional weightage to the number of hours of instruction prescribed.
10. Use of statistical tables is allowed for theory as well as practical examination.

11. To perform and complete the practical, it is necessary to have computing facility at the college. So,college will provide sufficient number of computers, UPS, and printer in the laboratory.
12. **Study Tour**: In order to acquaint the students with applications of statistical methods in various fields such as industries, agricultural sectors, government institutes, etc. at least one Study Tour for T.Y. B.Sc. Statistics students should be arranged and study report should be attached in the journal.
13. **Instructions for Examination**:
    a) **Medium of Instructions**: English
    b) **Pattern of examination**: Semester wise.
    c) **Standard of passing**: 20 Marks out of 50 marks for each paper. There is separate passing for internal and end semester university examination. (Student should obtain minimum 14 marks out of 35 in the external university examinationand minimum 06 marks out of 15 in the internal assessment).
    d) **Verification / Revaluation**: Allowed for Theory papers only.
    e) **Pattern of university question papers for ESE (Total 35 marks):**

For ST−351 to ST−356 and ST−361 to ST−366 (A), (B), (C).

Q.1	***(A) Multiple choice questions: 03 questions each of 01 marks.*** ***(B) True or false:02 questions each of 01 marks.***	***[03 Marks]*** ***[02 Marks]***
Q.2.	***Attempt any 02 out of 03 questions each of 05 marks.***	***[10 Marks]***
Q.3.	***Attempt any 02 out of 03 questions each of 05 marks.***	***[10 Marks]***
Q.4.	***Attempt any 01 out of 02 questions each of 10 marks.***	***[10 Marks]***

14. **Skill Enhancement Courses (SEC)**
    a) Skill Enhancement Courses (SEC) are considered as the practical courses.
    b) For SEC concerned teacher should conduct atleast 10 practicals (experiments) and it shall cover all the topics in the respective course with weightageproportional to the number of contact hours prescribed.

c) Students must complete all the practicals (experiments) in each of the SEC to the satisfaction of concernedteachers. Also, students must produce the journal along with the completion certificate duly signed by the Head of the Departmentat the time of the examination of SEC.

15. **Structure of evaluation of practical courses: ST 357, ST 358, ST 359, ST 367, ST 368 and skill enhancement courses (SEC) as:**

A) Continuous internal assessment (CIA)	
Section	Marks
I) Journal	10
II) Viva-Voce	05
Total of A	**15**

B) End of semester examination (ESE)			
Section	Nature	Marks	Time
I	To solve 2 questions out of 4, each of 14 marks	28	2 hour 30 minutes (for calculator users) 2 hour (for computer users)
II	Viva-Voce	7	10 minutes
Total of B		**35**	**2 hour 40 minutes (for calculator users)** **2 hour 10 minutes (for computer users)**

Grand Total of A and B : 15 + 35 = 50 marks

16. **Guidelines for conducting University examination (as per circular number 61 of 2005) for skill enhancement courses (SEC),**

- **ST 3510: Turbo C (Practical course),**
- **ST 3511: Statistical Computing using R-software (Practical course),**
- **ST 3610: Introduction to Python (Practical course) and**
- **ST 3611: Data Analytics (Practical course)**

a. The examination will be conducted in Statistics laboratory on computers.

b. Provision of at least 15 computers with necessary Turbo C, R-software and Pythonsoftwares installed should be made available by the centre.

c. Duration of examination is TWO hours.

d. Examination will be conducted at the time of theory examination of T.Y.B.Sc. Semester V and VI.

e. The examination will be conducted batch wise. A batch size will preferably be 12 candidates.

f. There will be no verification and revaluation of the aforesaidSECs.

g. If possible, the candidate will attach the printout of his/her above referred answer book.

h. The duration of this examination will be extended by 10 minutes for getting printouts.

i. Printouts of charts, graphs are not required. However they should be shown to the examiner on computer itself.

j. In case of partial power failure proportionate additional time may be given at that centre for the concerned batch. In case of total power failure candidates are required to write the answers in the answer book as in the case of regular theory examination.

k. The slips (question paper) made available for skill enhancement courses (SEC) shall be allotted to the candidates at random so that the total of marks in the slips will be exactly 28.

l. A candidate will solve the problems in the slip given to him/ her on computer and the output of work obtained by him/her will be evaluated by the examiner.

m. Answer book for this examination will be the answer book which is used at the time of theory examination.

n. One internal examiner and one external examiner will be appointed by the University for this Examination.

o. Internal and External examiners will examine the answer books jointly at the centre immediately after examination is over. **They will prepare mark list and enter the marks on University marks entry link after examination is over.**

p. For the conduct of this examination following staff will be appointed at the centre for each batch: **One peon, one technical assistant (computer maintenance person), one junior supervisor, one laboratory supervisor, one laboratory assistant.**

q. Remuneration to both examiners will be paid separately at the rate which is prescribed by the University for the Examiner of Practical examination of science subjects after the examination work is over.

Semester V

ST 351: DISTRIBUTION THEORY – I

1. Beta Distributions **(8L)**

1.1 Beta distributions of first kind : p.d.f.

$$f(x) = \frac{1}{B(m,n)} x^{m-1}(1-x)^{n-1}; \quad 0 \le x \le 1, \quad m,n > 0$$

= 0; elsewhere

Notation :$X \sim \beta_1(m,n)$, Nature of probability curve, symmetry, mean, variance, properties, r[th] raw moment, harmonic mean, median for $\beta_1(m,m)$.

1.2 Relation with U(0,1). The probability distributions of $\frac{1}{X}, X+Y, X-Y, XY, \frac{X}{Y}$where X and Y are i.i.d. $\beta_1(1,1)$.

1.3 Beta distributions of second kind

p.d.f.

$$f(x) = \frac{1}{B(m,n)} \frac{x^{m+1}}{(1+x)^{m+n}}; \quad x \ge 0, \quad m,n > 0$$

= 0; elsewhere

Notation :$X \sim \beta_2(m,n)$, Nature of probability curve, symmetry, mean, variance, properties, r[th] raw moment, harmonic mean, median for $\beta_2(m,m)$.

1.4 Interrelation between $\beta_1(m,n)$ and $\beta_2(m,n)$.

1.5 Distribution of $\frac{X}{Y}$, $\frac{X}{X+Y}$ etc. when X and Y are independent gamma variates.

1.6 Relation between distribution functions of $\beta_1(m,n)$ and binomial distribution.

1.7 Real life situations and applications.

2. Order Statistics **(10L)**

2.1 Order statistics for a random sample of size n from a continuous distribution, definition, derivation of distribution function and density function of the i-th order statistic $X_{(i)}$, particular cases for $i = 1$ and $i = n$. Distribution of $X_{(i)}$froma random sample following uniform and exponential distributions.

2.2 Derivation of joint p.d.f. of $(X_{(i)}, X_{(j)})$, probability distribution of sample range $X_{(n)} - X_{(1)}$.

2.3 Distribution of sample median

2.4 $Corr(X_{(i)}, X_{(j)})$when$X_1, X_2, \ldots . X_n$are i.i.d. uniform r.v.s ,distribution of $X_{(n)} - X_{(1)}$ and sample median. Comment on unbiased estimator of θ for $U(0,\theta)$ and exponential(θ) based on order statistics.

2.5 Joint distribution of $X_{(1)}, X_{(2)}, \ldots . X_{(n)}$

3. Cauchy distribution **(6L)**

3.1 p.d.f.

$$f(x) = \frac{\lambda}{\pi}\frac{1}{\lambda^2 + (x-\mu)^2};\ -\infty < X < \infty;\ -\infty < \mu < \infty;\ \lambda > 0$$

=0; elsewhere

3.2 Nature of the probability curve.

3.3 Distribution function, quartiles, non – existence of moments, distribution of i) $1/X$ ii) X^2 where $X \sim C(0,1)$ iii) $aX + b$

3.4 Additive property for two independent Cauchy variates (statement only), statement of distribution of the sample mean, comment on limiting distribution of X .

3.5 Relationship with uniform, Student's-t and normal distributions.

3.6 Applications of $C(\mu, \lambda)$

4. Chebychev's Inequality. **(4L)**

4.1 Chebychev's theorem: If g (x) is a non – negative function of r.v. X such that $E[g(X)] < \infty$ then, $P[g(X) \geq k] \leq \{E[g(X)]/k\}$ where k is positive real number.

4.2 Chebychev's inequality for discrete and continuous distributions (with proof) in the forms $P[|X-\mu| \geq k] \leq \left\{\sigma^2/k^2\right\}$, where $k > 1$, $P[|X-\mu| \geq k\sigma] \leq \left\{1/k^2\right\}$, $k > 1$. Where $\mu = E(X)$ and $\sigma^2 = Var(X) < \infty$.

4.3 Applications of Chebychev's inequality in control charts, statistical inference.

5. Central Limit Theorem and Weak Law of Large Numbers **(8L)**

5.1 Sequence of r.v.s. , convergence of sequence of r.v. in a) probability b) distribution, with simple illustrations.

5.2 De Moivre - Laplace theorem (Binomial distribution tends to Normal distribution for large n), Statement and proof of the central limit theorem for i.i.d. r.v.s. (proof based on MGF), simulation-based demonstrations.

5.3 Weak law of large numbers (WLLN), Simulation-based demonstrations

5.4 Applications of CLT and WLLN.

Books Recommended :

1. H. Cramer (1992). Mathematical Methods of Statistics, Princeton University Press

2. Mood, A.M. Graybill, F.Boes, D. (2017). Introduction to Theory of Statistics, 3rdEdn.,Mc-Graw Hill Series.

3. B.W. Lindgren (1993). Statistical Theory, 4thEdn., CRC Press, London.

4. Hogg, R.V. McKean, J. and Craig A.T. (2012). Introduction to Mathematical Statistics, 7thEdn., Pearson,

5. Sanjay Arora and BansiLal (1989). New Mathematical Statistics: A problem-oriented First Course, SatyaPrakashan New Delhi.

6. S.C. Gupta and V. K. Kapoor (2020). Fundamentals of Mathematical Statistics, 12thEdn.,Sultan Chand and Sons, 88, Daryaganj, New Delhi, 2.

7. V.K. Rohatgi and Saleh A. K. Md. E. (2015). An Introduction to Probability Theory and

Statistics, 3rdEdn. Wiley, New York
8. Feller W. (1968). An Introduction of Probability Theory and Its Applications, Vol. I, 3rdEdn., Wiley.
9.Sheldon Ross (2018). A first course in probability, 10thEdn.,Pearson Education.
10. Buyan, K. C. (2010). Probability theory and Statistical inference ,1stEdn., New Central Book Agency.

ST 352: THEORY OF ESTIMATION

1. Point Estimation and Interval Estimation (4L)

1.1 Notion of a parameter, parameter space.

1.2 General problem of estimating an unknown parameter by point and interval estimation (using Pivotal quantity approach).

1.3 Point Estimation: Definition of an estimator, distinction between estimator and estimate, illustrative examples. Mean Square Error (MSE) of an estimator.

2. Methods of Estimation (8L)

2.1 Likelihood Function, Definition of likelihood as a function of unknown parameter for a random sample comes from i) discrete ii) continuous distribution.

Distinction between the likelihood function and p.d.f. or p.m.f.

2.2 Method of maximum likelihood: Derivation of maximum likelihood estimator (M.L.E.) for parameters of only standard distributions (case of two unknown parameters only for normal distribution). Invariance property of M.L.E. M.L.E. of θ in uniform distribution over i) (0, θ) ii) (- θ, θ) iii) (mθ, nθ) (m < n), M.L.E. of θ in f(x ; θ)= Exp {-(x- θ)}, x >θ, M.L.E. of location parameter in Laplacedistribution.

2.3 Method of moments: Derivation of moment estimator for standard distributions. Illustrations of situations where M.L.E. and moment estimators are distinct and their comparison using MSE.

3. Properties of Estimator (20L)

3.1 Unbiasedness: Definition of an unbiased estimator, biased estimator, positive and negative bias, illustrations and examples (these should include unbiased and biased estimators for the same parameters). Proofs of the following results regarding unbiased estimators: (i) Two distinct unbiased estimators of (θ) give rise to infinitely manyestimators. (ii) If T is an unbiased estimator of θ, then ϕ(T) is unbiased estimator of ϕ(θ) provided ϕ(.) is a linearfunction.

Variance oftheestimator: Notion of the Best Linear Unbiased Estimator and uniformly minimum variance unbiased estimator (UMVUE), uniqueness of UMVUE whenever it exists.

3.2 Sufficiency: Concept and definition of sufficiency, statement of the Fisher-Neyman factorization theorem with proof for discrete probability distribution. Pitmann – Koopman form and sufficient statistic; Exponential family of probability distributions and sufficient statistic. Proofs of the following properties of sufficient statistics: (i) If T is sufficient for θ, then ϕ(T) is also sufficient for θ provided ø is a oneto one and ontofunction. (ii) If T is sufficient for

θ then T is also sufficient for $\phi(\theta)$. (iii) M.L.E. is a function of sufficientstatistic.

3.3 Fisher information function: Amount of information contained in statistic $T = T(X_1, X_2, \ldots, X_n)$. Statement regarding information in sample and in a sufficient statistic T.

3.4 Cramer- Rao Inequality: Statement and proof of Cramer - Rao inequality, Cramer – Rao Lower Bound (CRLB), definition of minimum variance bound unbiased estimator (MVBUE) of ø (θ). Proofs of following results: (i) If MVBUE exists for θ then MVBUE exists for $\phi(\theta)$ where $\phi(.)$ is alinear function.(ii) If T is MVBUE for θ then T is sufficient forθ.

3.5 Efficiency: Comparison of variance with CRLB, relative efficiency of T_1 w.r.t. T_2 for (i) unbiased (ii) biased estimators. Efficiency of unbiased estimator T w.r.t. CRLB.

4. Asymptotic Behavior of an Estimator (4L)

4.1Consistency: Definition.

4.2 Proof of the following theorems: (i) An estimator is consistent if its bias and variance both tend to zero asthe sample size tends toinfinity. (ii) If T is consistent estimator of θ and $\phi(.)$ is a continuous function,then $\phi(T)$ is a consistent estimator of $\phi(\theta)$.

Books Recommended:

1. Dudewecz, E.J. and Mishra, S.N. (1988). Modern Mathematical Statistics, John Wiley and Sons,Inc.
2. Hoel, P.G. Port, S. and Stone, C.(1972). Introduction to Statistical Theory, Houghton Mifflin Company (International) Dolphin Edition.
3. Hogg, R.V. McKean, J. and Craig A.T. (2012). Introduction to Mathematical Statistics, 7thEdn., Pearson,
4. Kendall, M. and Stuart, A. (1943). The advanced Theory of Statistics, Vol 1, Charles and Company Ltd., London
5. B.W. Lindgren (1993). Statistical Theory, 4thEdn., CRC Press, London.
6. Mood, A.M. Graybill, F.Boes, D. (2017). Introduction to Theory of Statistics, 3rdEdn.,Mc-Graw Hill Series.
7. V.K. Rohatgi and Saleh A. K. Md. E. (2015). An Introduction to Probability Theory and Statistics, 3rdEdn. Wiley, New York
8. Ramchandran, K.M. and Tsokos C. P. (2009). Mathematical Statistics with Applications, Academic Press.
9. Buyan, K. C. (2010) (Probability theory and Statistical inference , 1stEdn., New Central Book Agency.

ST 353: DESIGN AND ANALYSIS OF EXPERIMENTS

1. Design of Experiments (22L)

1.1 Analysis of variance (ANOVA): only concept and technique.

1.2 Basic terms: Experimental unit, treatment, layout of an experiment.

1.3 Basic principles: Replication, randomization and local control. Choice of size and shape of a plot for uniformity trials, the empirical formula for the variance per unit area of plots, definitions of Linear Treatment contrasts, orthogonal treatment contrasts

1.4 Completely Randomized Design (CRD): Application of the principles of design of experiment in CRD, Layout, Model: $X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$; i= 1,2, ...t. j = 1,2,....., n_i, Assumptions and interpretations. Testing normality of residuals (ε_{ij}) graphically (normal probability plot). Breakup of total sum of squares into components. Estimation of parameters, expected values of mean sums of squares, components of variance, preparation of ANOVA table, testing equality of treatment effects, Hypothesis to be tested $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_t = 0$. Comparison of treatment means using box plot techniques. Statement of Cochran's theorem. F test for testing H_0 with justification (independence of chi- square is to be assumed), test for equality of two specified treatment effects using critical difference (C.D). Equivalence between ANOVA and regression model.

1.5 Randomized Block Design (RBD): Application of the principles of design of experiments in RBD, layout. Model: $X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$; i= 1,2, ...t. j = 1,2,..... b. Assumptions and interpretations. Breakup of total sum of squares into components. Estimation of parameters, expected values of mean sums of squares, components of variance, preparation of analysis of variance table, Hypotheses to be tested

$H_{01}: \alpha_1 = \alpha_2 = \alpha_3 = ... = \alpha_t = 0$, $H_{02}: \beta_1 = \beta_2 = ... = \beta_b = 0$.

F-test for testing H_{01} and H_{02} with justification (independence of chi-squares is to be assumed), test for equality of two specified treatment effects using CD.

1.6 Latin Square Design (LSD): Application of the principles of design of experiments in LSD, layout, Model :$X_{ij(k)} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ij(k)}$; i,j,k=1,2,⋯,m, Assumptions and interpretations. Breakup of total sum of squares into components. Estimation of parameters, expected values of mean sums of squares, components of variance, preparation of analysis of variance table, hypotheses to be tested.

$H_{01} : \alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$, $H_{02} : \beta_1 = \beta_2 = \cdots = \beta_m = 0$, $H_{03} : \gamma_1 = \gamma_2 = \cdots = \gamma_m = 0$ and their interpretation. Justification of F test for H_{01}, H_{02} and H_{03} Independence of chi-square is to be assumed). Preparation of ANOVA table and F test for H_{01}, H_{02} and H_{03} testing for equality of two specified treatment effects, comparison of treatment effects using C.D.

1.7 Identification of real life situations where the above designs are used.

1.8 Analysis of non- normal data using. i) Square root transformation for counts.

ii) $Sin^{-1}(.)$ transformation for proportions.

2. Efficiency of Design **(3L)**

2.1 Concept and definition of efficiency of a design.

2.2 Efficiency of RBD over CRD.

2.3 Efficiency of LSD over (i) CRD (ii) RBD.

3. Factorial Experiments **(11L)**

3.1 General description of m^nfactorial experiment, $2^2$and 2^3factorial experiments arranged in RBD.

3.2 Definitions of main effects and interaction effects in 2^2 and 2^3factorial experiments.

3.3 Yate's procedure, preparation of ANOVA table, test for main effects and

interaction effects.
3.4 General idea of confounding in factorial experiments.
3.5 Construction of layouts in total confounding and partial confounding in 2^2 and 2^3 factorial experiments.
3.6 Total confounding (confounding only one interaction) ANOVA table, testing main effects and interaction effects.
3.7 Partial confounding (confounding only one interaction per replicate); ANOVA table, testing main effects and interaction effects.

Books Recommended

1. Cochran W.G. and Cox, C.M. (1992) Experimental Design, John Wiley and Sons, Inc., New York.
2. Dass, M.N. and Giri,N.C. (1987) Design and Analysis of Experiments, II Edition Wiley Eastern Ltd., New Delhi
3. Federer W.T. (1967) Experimental Design : Oxford and IBH Publishing Co., New Delhi
4. Goon, A.M., Gupta,M.K. and Dasgupta, B. (1998). Fundamentals of Statistics, Vol.II, The world Press Pvt. Ltd. Kolkatta
5. Gupta S.C. and Kapoor V.K.(2006). Fundamentals of Applied Statistics, S.Chand Sons, New Delhi
6. Johnson, R.A., Miller, I. and Freund, J.(2010). Probability and Statistics for Engineers, Prentice Hall, India.
7. Kempthorne, O. (1994). Design of Experiments: Introduction to Experimental Design, Wiley , New York
8. Montgomery, D.C. (2012). Design and Analysis of Experiments, John Wiley and sons Inc., New Delhi.
9. Snedecor, G.W. and Cochran, W.G. (1994). Statistical Methods, 8Th edition, Affiliated East – West Press, New Delhi
10. Wu, C.F.J. and Hamda, M. (2009). Experiments, Planning, Analysis and Parameter Design Optimization, John Wiley &Sons,Inc.,Hoboken,New Jersey.

ST-354: STATISTICAL PROCESS AND PRODUCT CONTROL

1. Introduction to Statistical Quality Control (SQC) (5L)

1.1 Meaning of quality, total quality management (TQM), six-sigma, ISI, ISO,
1.2 Meaning and purpose of Statistical Quality Control (SQC), Seven Process Control
1.2 Tools of Statistical Quality Control (SQC) (i) Check Sheet, (ii) Cause and effect diagram(CED), (iii) Pareto Diagram, (iv) Histogram, (v) Control chart, (vi)Scatter Diagram,(vii) Design of Experiments (DOE). (Only introduction of 7 PC tools is expected).

2. Control Charts (4L)

2.1 Chance causes and assignable causes of variation, statistical basis of control charts, exact probability limits, k-sigma limits, justification for the use of 3- sigma limits for normal distribution and using Chebychev's inequality for non-normal distributions.

2.2 Criteria for detecting lack of control situations:

(i) At least one point outside the control limits

(ii) A run of seven or more points above or below central line.

(iii) Presence of a non random pattern eg. cycle or linear trends etc.

2.3 Control chart technique as hypotheses testing problem.

2.4 Construction of control charts for (i) standards given, (ii) standards not given.

3 Control charts for variables and attributes (12 L)

3.1 Introduction of R chart and $\bar{X}$ chart

3.2 a) Construction of R chart when the process standard deviation is specified: control limits, drawing of control chart, plotting of sample ranges, drawing conclusion-determination of state of control process, corrective action if the process is out of statistical control. Construction of $\bar{X}$chart when the process average is specified: control limits, drawing of control chart, plotting of sample means. Drawing conclusion - determination of state of control of process, corrective action if the process is out of statistical control.

b) Construction of R chart when the process standard deviation (σ) is not given:
control limits, drawing of control chart, plotting sample range values, revision of control limits if necessary, estimate of σ for future use.

3.3 Construction of $\bar{X}$chart when the process average (μ) is not given :

control limits based $\hat{\mu} = \bar{X}$, $\hat{\sigma} = {\bar{R}}/{d_2}$ drawing of control chart, plotting sample means, revision of control limits of $\bar{X}$ chart, if necessary.

Note: To find revised control limits of any control chart delete the sample points above UCL and points below LCL (assuming a search for assignable causes at those points), in case of R and$\bar{X}$charts, first of all, revisions of control limits of R is to be completed and then by using the observations for which R chart shows the process is under control, the control limits for $\bar{X}$ chart should be determined. Revision of control limits of $\bar{X}$ chart be continued without revising the value of R or Estimate of $\mu$and σ for further use. Determination of state of control of the process. Probability of catching a shift. Identification of real life situations where this technique can be used. Limitations of $\bar{X}$, R charts.

3.4 p- chart: a) Construction and working of p-chart when subgroup sizes are same and value of the process fraction defective P is specified: control limits, drawing of control chart, plotting of sample fraction defectives. Determination of state of control of the process.

b) p-chart when subgroups sizes are different and value of the process fraction defective P is not specified with i) separate control limits ii) average sample size iii) standardized control chart. Drawing of control chart, plotting sample fraction defectives, determination of state of control of the process. Interpretation of high and low spots. Identification of real life situations. Probability of catching a shift.

3.5 C - Chart: a) Construction of c-chart when standard is given; control limits justification of 3 sigma limits, drawing of control chart, plotting number of defects per unit.
(b) Construction of c chart when standard is not given; control limits, explanation for the use of 3-sigma limits, drawing of control chart. Plotting number of defects per unit. Determination of state of control, interpretation of high and low spots in above cases. Identification of real life situations.

4. Process Capability Analysis (5L)

4.1 Specification limits, natural tolerance limits and their comparisons, decisions based on these comparisons, estimate of percent defectives.
4.2 Capability ratio and capability indices (Cp), capability performance indices(Cpk) with respect to machine and process, interpretation, relationship between (i) Cp and Cpk (ii) defective parts per million and Cp.

5. Sampling Plans (10 L)

5.1 Acceptance Sampling for Attributes: Introduction, Concept of sampling inspection plan, Comparison between 100% inspection and sampling inspection. Procedures of acceptance sampling with rectification, Single sampling plan and double sampling plan. Probabilities of acceptance and rejection
5.2 Explanation of the terms: Producer's risk and Consumer's risk, Operating characteristic (OC) curve, Acceptable Quality Level (AQL), Lot Tolerance Fraction Defective (LTFD) and Lot Tolerance Percent Defective (LTPD), Average Outgoing Quality (AOQ) and Average Outgoing Quality Limit (AOQL), AOQ curve, Average Sample Number (ASN), Average Total Inspection (ATI)
5.3 Single Sampling Plan: Computation of probability of acceptance using Poisson approximation, Derivation of AOQ and ATI. Graphical determination of AOQL, Determination of a single sampling plan by: a) lot quality approach b) average quality approach.
5.4 Double Sampling Plan: Evaluation of probability of acceptance using Poisson distribution, Structure of OC Curve, Derivation of AOQ, ASN and ATI (with complete inspection of second sample), Graphical determination of AOQL, Comparison of single sampling plan and double sample plan.

Books Recommended

1. Besterfield, D.H. and Michna , C.B. et al. (2009). Total Quality Management, 3rd edition, Pearson Education, Delhi.34
2. Dodge, H.F. and Roming, H.G. Sampling Inspection tables, John Wiley and Sons, Inc. New York
3. Duncan A.J. (1974). Quality Control and Industrial Statistics, fourth edition D.B. Taraporewala Sons and Co. Pvt. Ltd., Mumbai.
4. Grant, E. L. and Leavenworth (1980). Statistical Quality Control, fifth edition, Mc-Graw Hill, New Delhi.
5. Johnson, N.L. and Kotz, S. (1993). Capability Studies, Chapman and Hall Publishers.
6. Kamji and Asher (1996). 100 Methods of TQM, Sage Publishers, Delhi
7. Montgomery, D. C. (2008). Statistical Quality Control, 6thEdn., John Wiley, New York.
8. SP20 : Handbook of SQC, Bureau of Indian Standards.

ST 355: OPERATION RESEARCH - I

1 Introduction (2L)

1.1 History (Sec. 1.2)

1.2 Definition – (Sec. 1.3 – i. Daellenbach& George, ii. Theierauf&Klekamp, H. A. Taha)

1.3 Features of Operations Research Approach – Interdisciplinary Approach, Scientific Approach, Holistic Approach, Objective-Oriented Approach (Sec. 1.4).

1.4 Operations Research Approach to problem solving – Judgment, Research, Action (Sec. 1.5).

1.5 Classification of models based on Static models, Dynamic models, Deterministic models, Analytical model, Simulation model (Sec. 1.6.3 & 1.6.4,1.6.5).

1.6 Methods for solving models – Analytical method, Numerical method, Monte Carlo Method (Sec. 1.8).

1.7 Advantages, Opportunities and disadvantages of Operations research (Sec. 1.10, Sec. 1.11).

1.8 Applications of Operation Research – Finance and Accounting, Marketing, Purchase, Procurement and Exploration, Production Management (Sec. 1.13).

2 Linear Programming (15L)

2.1 Statement of the Linear Programming Problem (LPP), Formulation of problem as L.P. problem. L.P. Problem in (i) Canonical form (ii) standard form. Definition of (i) A slack variable (ii) A surplus Variable (iii) a solution (iv) a feasible solution (v) a basic feasible solution, (vi) a degenerate and non –degenerate solution vii) an optimal solution (viii) basic and non-basic variables .

2.2 Solution of L.P.P by: i) Graphical Method: solution space, unique and non-unique solutions, obtaining an optimal solution. Ii) Simplex Method: a) Initial basic feasible solution (IBFS) is readily available: obtaining an IBFS, criteria for deciding whether obtained solution is optimal criteria for unbounded solution, no solution, and more than one solution. B) IBFS not readily available: introduction of artificial variable, Big-M method, modified objective function, modifications and application of simplex method to L.P.P. with artificial variables.

2.3 Duality Theory: i) Writing dual of a primal problem, ii) Solution of a L.P.P. by using its dual problem.

3 Transportation and Assignment Problem (11L)

3.1 Transportation problem (T.P.), statement of T.P., balanced and unbalanced T.P.

3.2 Methods of obtaining basic feasible solution of T.P: i) North-West corner rule ii) Method of matrix minima (least cost method), iii) Vogel's approximation method (VAM).

3.3 U-V method of obtaining Optimal solution of T.P., degenerate solution, uniqueness and non- uniqueness of optimal solutions.

3.4 Assignment problems: i) Statement of an assignment, balanced and unbalanced problem, ii)relation with T.P. iii) optimal solution of an assignment problem using Hungarian method.

3.5 Examples and numerical problems.

4 CPM and PERT **(8L)**

4.1 Introduction (Sec. 13.1)

4.2 Basic differences between PERT and CPM. (Sec. 13.2).

4.3 Definition of (i) Events with Merge and Burst Events,(ii) Node,(iii) Activities – Predecessor, Successor, Dummy,(iv)Critical Activity,(v)Project Duration. (Sec. 13.4).

4.4 CPM: Construction of network, Definitions of (i) earliest start time (ii) earliest finish time (iii) latest start time (iii) latest finish time for an activity. Types of float - total floats, free float, independent float and their significance. Determination of critical path (Sec. 13.5).

4.5 PERT: Construction of network; (i) pessimistic time estimate, (ii) optimistic time estimate (iii) most likely time estimates, Determination of critical path, determination of mean and standard deviation of project duration, computations of probability of completing the project in a specified duration. (Sec. 13.6).

Note: All sections are from the book "Operations Research-Theory and Application" by J. K. Sharma

Books Recommended

1. Gass, S. I. (2011). Linear Programming Method and Applications, 5thEdn. Dover Publications
2. Taha, H. A. (2017). Operation Research – An Introduction, 10th Edition, Pearson,
3. Saceini, M., Yaspan, A. and Friedman, L. (2013). Operation Research: Method and Problems, Literary Licensing
4. Shrinath L.S. (1982). Linear Programming: Principles & Applications, , Affiliated East-West Press Pvt. Ltd. New Delhi.
5. Phillips, D.T., Ravindra, A., Solberg, J. (1987). Operations Research Principles and Practice, 2ndEdn.,John Wiley and Sons Inc.
6. Sharma J. K. (2012). Mathematical Models in Operations Research,5thEdn., MACIN
7. Kapoor, V. K. (2012). Operation Research: Concepts, Problems and Solutions, Sultan Chand and Sons, New Delhi.
8. Gupta, P. K. and Hira, D.S. (1976). Operations Research, S. Chand and company Ltd., New Delhi.
9. Shrinath, L.S. (2001). PERT-CPM Principles and Applications, Affiliated East-West Press Pvt. Ltd., New Delhi.

ST 356: REGRESSION ANALYSIS

1. Simple Linear Regression Model **(11L)**

1.1 Review of simple linear regression model : $Y = \beta_0 + \beta_1 X + \epsilon$, where ϵ is continuous random variable with $E(\epsilon) = 0$, $V(\epsilon) = \sigma^2$. Assumptions of simple linear regression. Estimation of $\beta_0$and β_1, by the method of least squares.

1.2 Properties of estimators of β_0, and β_1. (Sec. 2. 2. 2)

1.3 Estimation of σ^2. (Sec. 2. 2. 3)

1.4 Tests of hypothesis of β_1. (Sec. 2. 3)

1.5 Interval estimation in simple linear regression model. (Sec. 2. 4)

1.6 Coefficient of determination.(Sec. 2. 6)
1.7 Residual analysis (Sec. 4.1, 4.2.1) : Standardized residuals, Studentized residuals, (Sec. 4. 2. 2), Residual plots (Sec. 4. 2. 3)
1.8 Detection and treatment of outliers. (Sec. 4. 4)
1.9 Interpretation of three plots produced by lm command in R.

2. Multiple Linear Regression Model (15L)

2.1 Review of multiple linear regression model $Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \epsilon$, where ϵ is a continuous random variable with $E(\epsilon) = 0$, $V(\epsilon) = \sigma^2$. Assumptions of multiple linear regression. Estimation of regression parameters β_0, β_1, … and β_p by the method of least squares, obtaining normal equations, solutions of normal equations. (Sec. 3. 2. 1)
2.2 Estimation of σ^2(Sec. 3. 2. 4)
2.3 Tests of hypothesis of regression parameters. (Sec. 3. 3. 1, 3. 3. 2)
2.4 Interval estimation in regression model. (Sec.3. 4. 1)
2.5 Residual diagnostics and corrective measures such as transformation of response variable, (Sec. 5. 1, 5. 2, 5. 3)

3. Logistic Regression Model (10 L)

3.1 Binary response variable, Logit transformation, estimation of parameters, interpretation of parameters (Sec. 13. 2. 1, 13. 2. 2, 13. 2. 3 or chapter 1 except sec. 1.5 of HL)
3.2 Tests of hypotheses of model parameters, model deviance, Test based on likelihood ratios(LR).(Sec. 13. 2. 4, or chapter 1 except sec. 1.5 of HL)
3.3 AIC and BIC criteria for model selection.
3.4 Interpretation of output produced by glm command in R.
3.5 Introduction to Multiple logistic regression (Sec.2.1, 2.2 and 2.3 of HL)

Note:
(i) All sections are from the book “Introduction to Linear Regression Analysis” by Montgomery, D. C., Peck, E.A. and Vining, G. G.
(ii) HL stands for the book “Applied Logistic Regression “ byHosmer, D.W. and Lemeshow, S.

Books Recommended

1. Draper, N. R. and Smith, H. (1998)Applied Regression analysis (John Wiley) Third Edition.
2. Hosmer, D.W. and Lameshow, S.(1989). Applied Logistic Regression (Wiley).
3. Montgomery, D. C., Peck, E. A. and Vining, G.G. (2003). Introduction to Linear Regression Analysis(Wiley).
4. Neter, J., W., Kutner, M. H. ;Nachtsheim, C. J. and Wasserman, W. (1996). Applied Linear Statistical Models, fourth edition, Irwin USA.
5. Chatterjee S. and Hadi A.S.(2012) : Regression Analysis by Example, 5 thEdition,Wiley.
6. Kleinbaum G. and Klein M. (2011) : Logistic Regression, IIIrd Edition A Self learning text, Springer.

ST 357: PRACTICAL PAPER-I

Sr. No.	Title of the Experiment	No. of Experiments
1.	Chebychev's inequality and Order Statistics.	1
2.	Central Limit Theorem and Weak law of large numbers. Verification using simulation approach.	1
3.	Model sampling from Cauchy distribution.	1
4.	Construction of confidence interval for population mean (μ) for normal distribution when σ^2 known and σ^2 unknown.	1
5.	Construction of confidence interval for Population variance (σ^2) for normal distribution when μ known and μ unknown.	1
6.	Construction of confidence interval for population median and quartiles based on order statistics.	1
7	M.L.E and moment estimator of truncated Binomial and truncated Poisson distributions (truncated at zero).	1
8.	Simple regression analysis and diagnostics. (Estimation of Parameter, Test for significance of parameters Regression model, Interval Estimation, Interpretation of parameters. Computation of R-square, Residual Analysis by graphical method).	1
9.	Multiple regression analysis and diagnostics. (Estimation of Parameter, Test for significance of parameters Regression model, Interval Estimation, Interpretation of parameters, Computation of R-square, Adjusted R-square, Residual Analysis by graphical method, Variable selection methods).	1
10.	Logistic regression. (Estimation of Parameter, Test for significance of parameters Regression model, Interval Estimation, Interpretation of parameter/(s)).	1
11.	Simulations from nonstandard distributions (for given pmf/pdf), chi-square, t-distributions and F-distributions.	1
12.	Monte Carlo techniques to find approximate numerical value of gamma and beta function with real arguments.	1
Total no. of practicals		**12**

Note: Regression analysis practicals are performed using MS-Excel or R-software. (Using built in commands/direct commands available in MS-Excel and R-software).

ST 358: PRACTICAL PAPER – II

Sr. No.	Title of the experiment	No. of experiments
1	Analysis of CRD (equal and unequal replications), pair wise comparison of treatments using critical difference (C.D) method.	1
2	Analysis of RBD pair wise comparison of treatments using i) C.D ii) Tukey and Scheff's procedure (explain its theory in brief). Efficiency of RBD w.r.t. CRD	2
3	Analysis of LSD, pair wise comparison of treatments using C.D. Efficiency of LSD w.r.t. i) CRD ii) RBD.	2
4	Analysis of variance (ANOVA) for CRD, RBD and LSD using Kruskal Wallis test (explain its theory in brief)	1
5	Analysis of covariance in CRD, testing $\beta = 0$,	1
6	Analysis of covariance in RBD , testing $\beta = 0$	1
7	Analysis of $2^2$and 2^3 factorial experiments in RBD.	1
8	Analysis of 2^3 factorial experiments in RBD (partial confounding).	2
9	Analysis of 2^3 factorial experiments in RBD (total confounding).	1
Total no. of practicals		**12**

ST 359: PRACTICAL PAPER - III

Sr. No.	Title of the Experiment	No. of Experiments
1	Construction of R and $\bar{X}$chart, $\overline{\mathrm{X}} - \mathrm{MR}$ control chart, Probability of Catching shift.	1
2	p-chart for(I) fixed sample size (ii) variable sample size based on individual control limits, probability of detecting shift.	1
3	Single sampling plan for attributes (OC Curve, Producer's and Consumer's risks, AOQ, AOQL, ATI)	1
4	Determination of single sampling plan by: a) lot quality approach b) average quality approach	1
5	Double sampling plan for attributes (OC curve, AOQ, AOQL, ATI, ASN using Poisson distribution).	1
6	Linear programming problem I (Simplex method)	1

7	Linear programming problem II (Dual Simplex method)	1
8	Transportation problem	1
9	Assignment problem	1
10	CPM and PERT	2
11	Cost Benefit Analysis: determination of project duration and its associated cost when (i) Normal times are considered (ii) Crash times are considered. Determination of optimal network.	1
Total no. of practicals		**12**

Note: Practicals Based on Operations Research are to be verified using TORA

ST 3510: TURBO C(PRACTICAL COURSE)

1. C- Fundamentals **(6 L)**

1.1 Introduction

1.1 Algorithms and flowcharts.

1.2 Introduction to procedural language, middle level language, higher level language, general language structure, character set, keywords, identifiers.

1.3 Data types: Numeric and character data types, Numeric and character constants, string constants, symbolic constants.

1.4 Operators: Numeric, logical, arithmetic, unary, relational, equality, decrement, increment, conditional assignments, precedence of operator expressions and their evaluation.

1.5 Data input/output, numeric and character data, printf ( ), scanf ( ), getchar ( ), putchar ( ), gets ( ), puts ( ).

1.6 Formatted output

Programs on this unit

i. Converting °C temperature to °F
ii. To carry out arithmetic calculations (addition, subtraction, multiplication, division).
iii. To find area and perimeter of triangle, rectangle and square.
iv. To find area and circumference of circle.

2. Control Structures **(8L)**

2.1 Introduction

2.2 If, if else, while, do....while, for, switch, goto, break, continue,nested loops, programs using control structures.

Programs on this unit

i. To find roots of quadratic equation and print the outputs according to Discriminant using switch and do while operations.
ii. To check whether given number is odd or even.
iii. To check whether given number m is divisible by n or not.
iv. To find maximum of 2 numbers or 3 numbers.
v. To check whether integer is prime or not.

3. Arrays **(11L)**

3.1 Introduction & Concept.

3.2 declaration, definition, initialization of array, problem using arrays, passing to function, arrays

3.3 string operations, string functions like strcpy(), strcat(), strlen(), strcmp(), strrev().

Programs on this unit

i. To find arithmetic mean, geometric mean, harmonic mean, median, variance and coefficient of variation of frequency distribution.
ii. To find correlation coefficient and least square regression line of Y on X for a given bivariate data.
iii. To arrange the given data in increasing/decreasing order of magnitude.
v. To obtain median of given n observations.
vi. To obtain addition of two matrices, multiplication of two matrices.
vii. To test palindrome string using string function.
viii. To sort a string using string function.
ix. To search string using string function.
x. To combine given two strings using string function.
xi. To copy the string using string copy function.
xii. To find the length of the given string using string length function (strlen()).
xiii. To reverse the given string using string reverse function (strrev()).
xiv. To concatenates destination and source string (strcat()).
xv. To compare two strings using string compare function (strcmp()).

4. Functions **(7L)**

4.1 Introduction

4.2 Declaration, definition, recursion, user defined functions, library function, calling a function by reference and by value, local and global variables.

Programs on this unit

i. To find factorial of integer number (both recursive and non-recursive).
ii. To find the value of $X_n$where n is integer(both recursive and non-recursive).
iii. To prepare a 2X2 contingency table for chi square test and to find the value of test statistic and to check whether two attributes are independent.
iv. To fit a Binomial distribution to given data.

5. Pointers **(4L)**

5.1 Introduction

5.2 Basic concept and relation to one dimensional array.

Programs on this unit:

i. To print the address of the stored variable.
ii. To print addition/ subtraction of two numbers using pointers.

Books Recommended

1. Gottfried, B.S. (1996). Programming with C (Schaum Outline series), McGraw Hill co., London

2. Kanitkar, Y. (2008). Let us C, BFB publishers, New Delhi.
3. Karnighan, B. W. and Ritchi, M.(1988). The C programming language, Second edition,Prentice Hall.
4. Rajaraman V. (2007). Computer programming in C, Prentice Hall of India.,
5. Balgurusami, E. (2004), Programming in ANSI C, Tata McGraw Hill Education India

ST 3511: STATISTICAL COMPUTING USING R SOFTWARE(PRACTICAL COURSE)

1. Fundamentals of R **(3L)**

1.1 Revision of commands and functions studied in S.Y.B.Sc. Practical Course. Creating a vector using scan function, creating a data frame using edit command, Importing data from MS- Excel file.

1.2 Using read.table command, saving the R-output in a file using MS-Excel, concept of R-script file, Graphics using R: (a) High level plotting functions (b) Low level plotting functions

(c) Interactive graphic functions. The following statistical methods using "R"

2. Diagrams **(3L)**

2.1 Simple bar diagram, Subdivided bar diagram, Multiple bar diagram,

2.2 Pie diagram, stem and leaf diagram.

3. Graphs **(3L)**

3.1 Boxplot for one and more than one variables, rod or spike plot,

3.2 histogram for raw data with **prob =TRUE option** and for both equal and unequal class intervals,

3.3 frequency polygon, ogive curves, empirical distribution function.

3.4 Saving the diagram and graphs in MS-Word file.

4. Measures of Central Tendency, dispersion, skewness and kurtosis **(5L)**

4.1 Computations of following measures for all types of data.

a) Central tendency mean, mode, median, quartiles, deciles, percentiles, geometric mean and harmonic mean.

b) Dispersion: variance, standard deviation, coefficient of variation, mean deviation.

c) Skewness: Bowley's coefficient and Karl Pearson's coefficient of skewness.

d) Moments: Computations of raw and central moments, measures of skewness and kurtosis based on it.

5. Probability distributions **(5L)** 5.1

Simulation from distributions, computations of probabilities, cumulative probabilities, quantiles and drawing random sample using d, p, q, r functions and graphs of pmf/ pdf by varying parameters for following distributions:

5.2 Binomial, Poisson, Hypergeometric , normal, exponential, gamma, Cauchy, lognormal, Weibull, uniform, Laplace.

5.3 Fitting of Poisson and normal distribution, testing normality of data by Shapiro- Wilks test.

6 . Testing of hypothesis (1L)

6.1 Chi -square test for independence of attributes.

7. ANOVA (3L)

7.1 One way and two way classification, Bartlett's test for homoscedasticity, Kruskal Wallis test.

8. Non – parametric Tests (3L)

8.1 Wilcoxon's signed rank test , Mann Whitney Test, Kolmogorov Smirnov test

9. Programming in R (10L)

9.1 Statements: if and if… else, for loop, cat and print commands.

Writing following Programs in R:

1. Formation of strata clusters using special criterion and hence drawing a sample by stratified sampling and cluster sampling method.
2. Testing normality of number of samples.
3. Verifying the assumptions in testing H_0: $\mu = \mu_0$ and then applying appropriate test.
4. Verifying the assumptions in testing H_0: $\mu_1 = \mu_2$ and then applying appropriate test.
5. Verifying the assumptions in testing H_0: $\mu_1 = \mu_2$ in paired data and then applying appropriate test.
6. Verifying the assumptions in testing H_0: $\sigma_1^2 = \sigma_2^2$ in paired data and then applying appropriate test.
7. Verifying the assumptions in one way ANOVA and then applying appropriate test.
8. Performing number of chi – square tests.

Books recommended:

1. Crawley, M.J. (2006). Statistics – An introduction using R. John Wiley London.
2. Purohit, S.G., Deshmukh, S.R. and Gore, S.D., (2015). Statistics using R. Alpha Science International.
3. Verzani, J., (2018). Using R for introductory statistics. CRC press.
4. Schumacker, R.E., (2014). Learning statistics using R. Sage Publications.

Semester VI

ST : 361 DISTRIBUTION THEORY – II

1. Weibull Distribution (5L)

1..1p.d.f.

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}}; x \geq 0, \alpha, \beta > 0$$

= 0 , otherwise

Notation :$X \sim W(\alpha, \beta)$.

1.2 Probability curve, location parameter, shape parameter, scale parameter, Distribution function, quartiles, mean and variance, coefficient of variation, relationship with gamma and exponential distribution, Hazard rate, IFR and DFR property.

1.3 Real life situations and applications.

2. Laplace (Double Exponential) Distribution (5L)

2.1 p.d.f

$$f(x) = \frac{\lambda}{2} e^{-\lambda|X-\mu|}; -\infty < X, \mu < \infty, \qquad \lambda > 0$$

= 0 , otherwise

Notation :$X \sim L\ (\mu,\ \lambda)$.

2.2 Nature of the probability curve.

2.3 Distribution function, quartiles, comment on MLE of μ, λ.

2.4 MGF, CGF, moments and cumulants, $\beta_1, \beta_2, \gamma_1, \gamma_2$

2.5 Laplace distribution as the distribution of the difference of two i.i.d. exponential variates with mean $1/\lambda$. Applications and real life situations.

3. Lognormal Distribution (6L)

3.1 p.d.f

$$f(x) = \frac{1}{(x-a)\sigma\sqrt{2\pi}} exp\left\{\frac{-1}{2\sigma^2}[\log(x-a) - \mu]^2\right\};\quad X > a, -\infty < \mu < \infty,\ \ \sigma > 0$$

= 0 , otherwise

Notation :$X \sim LN\ (a, \mu, \sigma^2)$.

3.2 Nature of the probability curve.

3.3 Moments (r- th moment of $X - a$), first four moments , β_1 and $\gamma_1$coefficients, quartiles, mode.

3.4 Relation with $N\ (\mu, \sigma^2)$ distribution

3.5 Distribution of $\prod X_i$ Xi's are independent lognormal variates.

3.6 Applications and real life situations

4. Truncated Normal Distribution (4L)

4.1 Normal distribution $N\ (\mu, \sigma^2)$ truncated i) to the left below a ii) to the right above b iii) to the left below a, and to the right above b, (a < b) its p.d.f. and derivation of mean and statement (without derivation) of variance.

4.2 Real life situations and applications.

5. Parato Distribution (4L)

5.1 p.d.f.

$$f(x) = \frac{\lambda}{x^{\lambda+1}}; \quad X \geq 1, \qquad \lambda > 0$$

= 0 , otherwise

5.2 Nature of pdf curve ,

5.3 Mean, variance and moments. Existence of moments for different values of λ. Symmetry, CDF

5.4 Applications in the field of economics.

6. Bivariate Normal Distribution (12L)

6.1p.d.f. of a bivariate normal distribution

$$f(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left\{\frac{x-\mu_1}{\sigma_1}\right\}^2 + \left\{\frac{y-\mu_2}{\sigma_2}\right\}^2 - 2\rho\left\{\frac{x-\mu_1}{\sigma_1}\right\}\left\{\frac{y-\mu_2}{\sigma_2}\right\}\right]\right\};$$

$$-\infty < X, Y, \mu_1, \mu_2 < \infty, ; \ \sigma_1, \sigma_2 > 0; \ -1 < \rho < 1$$

= 0 , otherwise

Notation :$(X, Y) \sim BN\ (\mu_1, \mu_2, \sigma_1, \sigma_2\rho\ ), \underline{X} \sim N_p(\underline{\mu}\ , \Sigma)$, use of matrix algebra is recommended.

6.2 Nature of surface of p.d.f. ,marginal and conditional distributions, identification of parameters, regression of Y on X and of X on Y, independence and uncorrelatedness, MGF and moments. Distribution of $aX\ + bY\ + c, \quad X/Y$.

6.3 Applications and real life situations

Books Recommended

1. H. Cramer (1992). Mathematical Methods of Statistics, Princeton University Press

2. Mood, A.M. Graybill, F.Boes, D. (2017). Introduction to Theory of Statistics, 3rdEdn.,Mc-Graw Hill Series.

3. B.W. Lindgren (1993). Statistical Theory, 4thEdn., CRC Press, London.

4. Hogg, R.V. McKean, J. and Craig A.T. (2012). Introduction to Mathematical Statistics, 7thEdn., Pearson,

5. Sanjay Arora and BansiLal (1989). New Mathematical Statistics: A problem-oriented First Course, SatyaPrakashan New Delhi.

6. S.C. Gupta and V. K. Kapoor (2020). Fundamentals of Mathematical Statistics, 12thEdn.,Sultan Chand and Sons, 88, Daryaganj, New Delhi, 2.

7. V.K. Rohatgi and Saleh A. K. Md. E. (2015). An Introduction to Probability Theory and Statistics, 3rdEdn. Wiley, New York

9. Feller W. (1968). An Introduction of Probability Theory and Its Applications, Vol. I, 3rdEdn., Wiley.

10.Sheldon Ross (2018). A first course in probability, 10thEdn.,Pearson Education.

11. Buyan, K. C. (2010). Probability theory and Statistical inference, 1stEdn., New Central Book Agency.

ST 362: TESTING OF HYPOTHESIS

1. Parametric Tests (15 L)

1.1 Statistical hypothesis, problem of testing of hypotheses. Definition and illustrations of (1) simple hypothesis, (2) composite hypothesis, (3) test of hypothesis, (4) critical region, (5) type I and type II errors. Probabilities of type I error and type II error. Problem of controlling the probabilities of errors of two kinds. (b) Definition and illustrations of (i) level of significance, (ii) observed level of significance (p-value), (iii) size of a test, (iv) power of a test.

1.2 Definition of most powerful (M.P.) level α test of simple null hypothesis against simple alternative. Statement of Neyman - Pearson (N-P) lemma for constructing the most powerful level α test of simple null hypothesis against simple alternative hypothesis. Illustrations.

1.3 Power function of a test, power curve, definition of uniformly most powerful (UMP) level α test for one sided alternative. Illustrations.

2. Likelihood ratio tests (9L)

2.1 Notion of likelihood ratio test (LRT), $\Lambda(x) = \frac{\underset{\theta \epsilon \text{Ⓗ}_0}{\text{Sup}} L(\theta|\underline{x})}{\underset{\theta \epsilon \text{Ⓗ}_0 \cup \text{Ⓗ}_1}{\text{Sup}} L(\theta|\underline{x})}$

2.2 Construction of LRT for H_0: $\theta = \theta_0$ against H_1: $\theta \neq \theta_0$ for the mean of normal distribution for i) Known σ^2 ii) unknown σ^2 (one sided and two sided alternatives).

2.3 LRT for variance of normal distribution for i) known μ ii) unknown μ (one sided and two sided alternatives hypotheses).

2.4 LRT for parameters of binomial and exponential distribution for two sided alternatives only.

2.5 LRT as a function of sufficient statistics, statement of asymptotic distribution of $-2\log_e\lambda(x)$.

3. Non-parametric Tests (12 L)

3.1 Concept of non-parametric tests. Distinguish between parametric and nonparametric Tests. Concept of distribution free statistic. One tailed and two tailed test procedure of i) Sign test, ii) Wilcoxon signed rank test iii) Mann Whitney U test, iv) Run test, one sample and two samples problems

3.2 Empirical distribution function $S_n(x)$. Properties of $S_n(x)$ as estimator of F(.). Kolmogorov-Smirnov test for completely specified univariate distribution (one Sample problem only) for two sided alternative hypotheses. Comparison with chi-square test.

Books Recommended

1. B.W. Lindgren (1993). Statistical Theory, 4thEdn., CRC Press, London.
2. Daniel, W.W. (2000) Applied Nonparametric Statistics, Duxbury Press Boston.
3. Dudewecz, E.J. and Mishra, S.N. (1988). Modern Mathematical Statistics, John Wiley and Sons,Inc.
4. Gibbons J.D. andChakraborti, S. (2010). Non parametric Statistical Inference, CRC Press, London
5. Hoel, P.G. Port, S. and Stone, C.(1972). Introduction to Statistical Theory, Houghton Mifflin Company (International) Dolphin Edition.

6. Hogg, R.V. McKean, J. and Craig A.T. (2012). Introduction to Mathematical Statistics, 7thEdn., Pearson,
7. Kendall, M. and Stuart, A. (1943). The advanced Theory of Statistics, Vol 1, Charles and Company Ltd., London
8. Mood, A.M. Graybill, F.Boes, D. (2017). Introduction to Theory of Statistics, 3rdEdn.,Mc-Graw Hill Series.
9. Ramchandran, K.M. and Tsokos C. P. (2009). Mathematical Statistics with Applications, Academic Press.
10. V.K. Rohatgi and Saleh A. K. Md. E. (2015). An Introduction to Probability Theory and Statistics, 3rdEdn. Wiley, New York
11. Buyan, K. C. (2010). Probability theory and Statistical inference,1stEdn., New Central Book Agency.

ST 363: SAMPLING THEORY

1. Role of Sample Surveys in Research Methodology (3L)

1.1 Objectives of a sample survey.

1.2 Designing a questionnaire, characteristics of a good questionnaire (Questions with codes & scores are to be discussed). Reliability and validity testing by using

i) Test – Retest method, ii) Internal Consistency: a) Kuder Rechardson Coefficient (KR-20)

b) Cronbach's Coefficient Alpha (α)

1.3 Planning, execution and analysis of a sample survey, practical problems at each of these stages.

1.4 Sampling and non-sampling errors with illustrations.

1.5 Study of some surveys illustrating the above ideas, rounds conducted by National Sample Surveys organization.

2. Simple Random Sampling (8L)

2.1 Concept of distinguishable elementary units, sampling units, sampling frame, random sample, Requisites of a good sample, Simple random sampling from finite population of size (N), i) with replacement (SRSWR) ii) without replacement (SRSWOR),

2.2 Definitions of population mean, population variance and population total parameters, inclusion probabilities.

2.3 Sample mean ($\bar{y}$) as an estimator of population mean ($\bar{y}$), derivation of expectation and standard error of$\bar{y}$, N$\bar{y}$ as an estimator of population total, derivation of expectation and standard error of N$\bar{y}$, Estimator of above standard errors, both in case of SRSWR and SRSWOR, confidence interval for population mean, population total, standard error.

2.4 Sampling for proportion as an application of a simple random sampling with X_i as zero or one. Sample proportion (p) as an estimator of population proportion (P) of units possessing a certain attribute, derivation of expectation and standard error of (p), Np as an estimator of total number of units in the population possessing a certain attribute, derivation of

expectation and standard error of Np. Estimator of above standard error both in case of SRSWR and SRSWOR.

3. Stratified Random Sampling (12L)

3.1 Stratification, basis of stratification, real life situation where stratification can be used.

3.2 Stratified random sampling as a sample drawn from individual strata using SRSWOR in each stratum.

3.3 (a)$\overline{y}_{st} = \frac{\sum N_i \bar{y}_i}{N}$as an estimator of population mean ($\bar{Y}$), derivation of expectation and standard error of $\overline{y}_{st}$, (b) N$\overline{y}_{st}$as an estimator of population total, derivation of expectation and standard error of N$\overline{y}_{st}$, (c) Estimator of above standard errors.

3.4 Problem of allocation, Proportional allocation, Neyman's allocation, derivation of the expressions for the standard errors of the above estimators when these allocations are used.

3.5 Gain in precision due to stratification, comparison amongst SRSWOR, stratification with proportional allocation and stratification with Neyman'sallocation.

3.6 Cost and variance analysis in stratified random sampling, minimization of variance for fixed cost, minimization of cost for fixed variance, optimum allocation, Neyman's allocation as a particular case of optimum allocation in cost and variance analysis.

4. Systematic Sampling (Population size divisible by sample size) (6L)

4.1 Real life situations where systematic sampling is appropriate. Techniques of drawing a sample using systematic sampling

4.2 Estimation of the population mean and population total, standard error of these estimators.

4.3 Comparison of systematic sampling with SRSWOR.

4.4 Comparison of systematic sampling with SRSWOR and stratified sampling in the presence of linear trend.

5. Determination of Sample Size (4L)

5.1 Determination of the sample size for estimating population mean, population total and population proportion for the given:

i) Margin of error and confidence coefficient.
ii) Coefficient of variation of the estimator and confidence coefficient.

6. **Ratio and Regression Methods of Estimation for SRSWOR (3L)**

6.1 Rationale behind using auxiliary variates in estimation.

6.2 Situations where (i) ratio method is appropriate, (ii) regression method is appropriate.

6.3 Ratio and regression estimators of the population mean and population total.

6.4 Comments regarding bias, statement of standard errors of ratio and regression estimators relative efficiency of these estimators, with respect to SRSWOR

(Derivations are not expected).

Books Recommended

1. Cochran, W.G.(2007):Sampling Techniques , Third Edition, Wiley India Pvt. Ltd., New Delhi.
2. Murthy, M. N. (1977): Sampling Theory and Methods, Statistical Publishing Society, Kolkata.

3. Singh, D. and Chaudhary, F. S. (1986): Theory and Analysis of Sample Survey Designs, Wiley Eastern Ltd., New Delhi.
4. Sukhatme, P.V., Sukhatme, B. V.(1984): Sampling theory of Surveys with Applications, Indian Society of Agricultural Statistics, New Delhi.
5. Sampath S. (2005): Sampling Theory and Methods, Second edition, Narosa, New Delhi.
6. Mukhopadhyay P (2008): Theory and methods of survey sampling. Prentice-Hall of India, New Delhi.
7. Chagbao Wu and Mary E. Thompson(2020) :Sampling Theory and Practice, Springer Nature Switzerland.
8. RaghunathArnab (2017): Survey Sampling Theory and Applications, Academic Press, Elsevier.

ST 364: INTRODUCTION TO SURVIVAL ANALYSIS

1. Ageing (18L)

1.1 Survival function, probability density function, hazard function, cumulative hazard rate, mean residual life function, equilibrium residual life function , interrelation between all these function, Total time on test transform (TTT).

1.2 No ageing, proof of the properties of no ageing:

i) Cauchy functional equation ii) Constant failure rate, iii) Constant mean residual life,

iv) Exponential life distribution, v) Exponential Equilibrium distribution, vi) Identity function as the TTT transforms

1.3 Positive and negative ageing: IFR, DFR IFRA, DFRA, NBU, NWU, NBUE, NWUE, DMRL, IMRL, bathtub failure rate. Proof of the Implications: IFR → IFRA, IFR → DMRL, IFRA → NBU, NBU → NBUE, DMRL → NBUE, DFR → DFRA, DFR → IMRL, DFRA → NWU, NWU→ NWUE, IMRL → NWUE.

1.4 Examples.

2. Some Parametric Families of life distributions (10L)

2.1 Classification of following parametric families of life distribution according to ageing: Weibull, Gamma, lognormal, linear failure rate, Makeham, Pareto, Lehman, distribution of specific parallel and series system.

2.2 Some Properties of Exponential Distribution:

i) If $T_1, T_2, \ldots\ldots, T_n$ are independent with$T_i \rightarrow$ EXP (λ_i) for $i = 1,2, \ldots\ldots, n$ and $T = \min\{T_1, T_2, \ldots\ldots, T_n\}$ then T→ Exp $(\sum_{i=1}^{n} \lambda_i)$

ii) If $T_1, T_2, \ldots\ldots, T_n$ are independent and identically distributed exponential random variable with parameter λ, then $2\lambda \sum_{i=1}^{n} T_i \rightarrow {x^2}_{2n}$

iii) If T is a continuous non-negative random variable with cumulative hazard rate R(t) then R(t) is exponential random variable with parameter 1.

iv) If T follows the Weibull distribution with parameter λ and ϒ then T^{Υ} has exponential distribution with parameter λ.

v) Concept of spacing and normalized spacing, distribution of spacing and normalized spacing

3. Censoring (8L)

3.1 Examples of life time situations, straggled and aligned entries in the survival time data, concept of censoring, time/type I censoring, order/type II censoring, right random censoring, left random censoring, undersigned censoring. Estimation of parameter of exponential distribution using Type I and Type II censoring.

3.2 Nonparametric estimation of survival function, confidence band on survival function, actuarial estimator of survival function, Greenwood's formula, Kaplan-Meier estimator of survival function in the presence of censored observations.

Books Recommended

1. Barlow, R. E. and Proschan F. (1975). Statistical theory of Reliability and Life testing: Probability Models Holt, Rinehart and Winston Inc.
2. Barlow, R. E. and Proschan F. (1996). Mathematical Theory of Reliability. John Wiley.
3. Cox, D.R. and Oakes, D. (1984).Analysis of Survival Data,Chapman and Hall.
4. Deshpande, J.V. and Purohit S.G. (2005). Life Time Data: Statistical Models and Methods, Word Scientific.
5. Tobias, P.A. and Trindane, D. C. (1995). Applied Reliability. Second edition. CRC Press.

ST 365 (A): ACTUARIAL STATISTICS

1. Introduction and Feasibility of Insurance Business. (5L)

1.1 Insurance companies as business organizations.

1.2 Role of insurance business in Economy. Concept of risk and their types

1.3 Introduction of terms: premium, policy, policyholder.

1.4 Role of Statistics in insurance business.

1.5 Expected value principle. Concept of utility function, Feasibility of insurance business

2. Survival Distribution and Life Tables (9L)

2.1 Time- until death random variable, its distribution function and survival function in actuarial notation.

2.2 Force of mortality. Curtate future life random variable, its probability mass function and survival function in actuarial notation.

2.3 Deferred probability.

2.4 Construction of life table using random survivorship approach.

3. Models for Life Insurance (9L)

3.1 Introduction of simple and compound interest rate policy.
3.2 Different types of Interest rates. Insurance payable at the end of the year of death, present value random variable, actuarial present value.
3.3 Derivation of actuarial present value for n-year term life insurance, whole life insurance and n-year endowment insurance.

4. Annuities (8L)

4.1 Annuities – certain, annuity due, annuity immediate.
4.2 Discrete life annuities: n-year temporary life annuity due and a whole life annuity due, present value random variables of the payment, and their actuarial present values.

5. Benefit Premiums (5L)

5.1 Concept of a loss random variable. Equivalence principle, Computation of fully discrete premium for n-year term life insurance, whole life insurance and endowment insurance

Books Recommended

1. Bowers N.L. Jr., H.S.Gerber, J.C. Hickan, D.A.Jones, C.J.Nesbitt, (1997). Actuarial Mathematics, Society of Actuaries, U.S.
2. Deshmukh, S. R. (2009). Actuarial Statistics, Universities Press, Hyderabad.

ST 365 (B): OPERATION RESEARCH – II

1. Replacement Theory (5L)

1.1 Introduction (Sec. 17.1)
1.2 Types of failures – Gradual failure and sudden failure (Sec. 17.2)
1.3 Replacement of item whose efficiency deteriorates with time when (a) value of money remains constant during the period (for time as a discrete variable and continuous variable), (b) value of money changes with constant rate during the period (Sec. 17.3 – Model I, Model II).

2. Inventory Theory (9L)

2.1 Introduction, meaning, functional role and reasons for carrying inventory (Sec. 14.1, 14.2, 14.3, 14.4)
2.2 Inventory Control Methods – (i) Always Better Control (ABC) analysis,(ii) Vital, Essential and Desirable (VED) Analysis,(iii) Fast moving, Nonmoving, Slow moving, Dead (FNSD) Analysis (Sec. 14.13).
2.3 Description of generalized inventory model (Sec 14.6).
2.4 Types of inventory models:
   i) The economic lot size model with uniform demand, instantaneous replenishment rate and no shortage(with derivation),
   ii) The Economic lot size model with uniform rate of demand, finite replenishment rate and no shortage(with derivation),

iii) The economic lot size model with uniform demand, instantaneous replenishment with shortage(without derivation),
iv) The economic lot size model with uniform rate of demand, finite replenishment rate with shortage (without derivation) (Sec. 14.7).

3. Decision Theory (9L)

3.1 Introduction, Types of Decision making Environment (Sec. 11.1,11.3)
3.2 Decision making under uncertainty - (i) Laplace criterion, (ii) Minimax / Maximax criterion, (iii) Savage minimax regret criterion, (iv) Hurwitz criterion (Sec. 11.4).
3.3 Decision making under risk: (i) Expected Monetary Value, (ii) Expected Opportunity Loss, (iii) Expected Value of Perfect Information (Sec. 11.5).
3.4 Decision Trees (Sec. 11.7)

4. Game Theory (8L)

4.1 Introduction, Definitions – (i) Number of players, (ii) sum of gains or losses, (iii) Strategy and its type (Sec. 12.1).
4.2 Two Person Zero-Sum Game (Sec. 12.2)
4.3 Pure & Mixed Strategies (Se. 12.3,12.4)
4.4 Principles of Dominance (Sec. 12.5)
4.5 Solution Methods for Games without Saddle Point – Algebraic, Arithmetic, Matrix and Graphical Method.

5. Sequencing (5L)

5.1 Introduction, notations, terminology, assumptions (Sec. 20.1, 20.2)
5.2 Processing n jobs through two machines (Sec. 20.3)
5.3 Processing n jobs through three machines, (reducible to two machines and n-jobs). (Sec. 20.4)
5.4 Examples and problems

Note: All sections are from the book "Operations Research-Theory and Application" by J. K. Sharma

Books Recommended

1. Gass,E.: Linear programming method and applications, Narosa Publishing House, New Delhi.
2. Taha, R.A.: Operation research, An Introduction, fifth edition, Prentice Hall of India, New Delhi.
3. Saceini, Yaspan, Friedman: Operation Research methods and problems, Willey International Edition.
4. Shrinath.L.S: Linear Programming, Affiliated East-West Pvt. Ltd, New Delhi.
5. Phillips,D.T, Ravindra , A, Solberg, I.: Operation Research principles and practice, John Willey and sons Inc.
6. Sharma, J.K.: Mathematical Models in Operation Research, Tata McGraw Hill Publishing Company Ltd., New Delhi.

7. Kapoor, V.K.: Operations Research, Sultan Chand and Sons. New Delhi.
8. Gupta, P.K.andHira, D.S.: Operation Research, S.Chand and company Ltd., New Delhi.

ST 366 (A): STOCHASTIC PROCESSES

1. Introduction (15L)

1.1 Definition of a Stochastic process, state space, index set, classification of stochastic processes with examples,

1.2 Markov Property, Markov chains (MC) {Xn, n ≥ 0}, finite MC, time homogeneous M.C.

1.3 One step transition probabilities, and transition probability matrix (t.p.m.), stochastic matrix, graphical representation Markov chain, n-step transition probability matrix ,

1.4 Chapman Kolmogorov equation, initial distribution,

1.5 joint distribution function of $\{X_0, X_1,\ldots,Xn\}$, partial sum of independent and identically distributed random variables as Markov chain.

1.6 Illustrations such as random walk, Gambler's ruin problem, Ehrenfest chain.

2. Classification of states (12L)

2.1 Communicating states (accessible, transitive, symmetry, reflexive states) first return probability, probability of ever return

2.2 Classification of states, as persistent (recurrent) and transient states,

2.3 Closed set of states, reducible state, reducible MC, irreducible states, irreducible MC, absorbing state, periodicity, aperiodic, Mean recurrence, non-null recurrent, ergodic state and ergodic M.C.

2.4 illustrate with random walk MC, Gambler's ruin MC, and other examples problem find is it ergodic MC or not.

3. Stationary distribution (5L)

3.1 Elementary property of stationary distribution, steady state distribution, for an irreducible ergodic finite

3.2 Long run behavior of a M.C**.** with illustration

4. Poisson Process (4L)

4.1 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 inter-arrival time ,mean, variance and covariance functions .

4.2 Definition of compound Poisson process mean and variance functions and its applications.

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.and Stone, C.J. (1972). Introduction to stochastic processes, Wiley Eastern.
4. Medhi J. (1982). Stochastic processes, Wiley Eastern
5. Ross, S. (2000). Introduction to probability models, 7thedn, Academic Press
6. Ross, S. (1996) Stochastic processes, John Wiley.
7. Srinivasan, S.K. and Mehta, K.M. (1981). Stochastic Processes, Tata Mc-Graw Hill.
8. Taylor, H N and Karlin, S. (1984). An introduction to stochastic modeling Academic Press.

ST 366 (B): RELIABILITY THEORY AND APPLICATIONS

1. Structural properties of coherent system (15L)

1.1 Binary system of independent components, order of the system, different types of systems, concept of the structure function, structure function of series system, parallel system,
k-out-of-n system, essentially series and parallel system, reliability block diagram, guidelines for construction of reliability block diagram.

1.2 Coherent structure function (maximum 4 components), relevant component, increasing structure function, pivotal decomposition of structure function, dual of a structure function (proof of dual of series system of order n is parallel system of order n,dual of the parallel system of order n is a series system of order n, dual of k-out -of –n system is (n-k+1)-out of –n system).

1.3 Path sets, cut sets, minimal path and cut set, representation of coherent system in terms of minimal path sets and cut sets, dual coherent structure function, relative/structural importance of components, module of the coherent system, modular decomposition of coherent system.

2. Reliability of coherent system (8L)

2.1 Reliability (h(p)) of system of independent components, Basic properties of system reliability (such as reliability function is increasing function, system and component redundancy etc.).

2.2 Computation of reliability of coherent system by using minimal path and cut set representation, upper and lower bound on system reliability by using exact system reliability, relative importance of a component.

2.3 Reliability of Binary Components and Systems, Examples to show that systems with higher reliabilities can be constructed using components of low values of reliabilities, i.e. unreliable components.

3. Ageing Properties: (7L)

3.1 Definitions: Hazard rates, hazard function, survival function, Concept of distributions with increasing and decreasing failure rate (IFR/DFR), Average Increasing (Decreasing) Failure Rate (IFRA/ DFRA)

3.2 Relationship betweena) Survival function and hazard function ii) Density function and hazard rate

3.3 Derivations of the following results :i) Hazard rate of a series system of components having independent life times is summation of component hazard rates. Ii) Lifetime of series system of independent components with independent IFR life times is IFR.

3.4 Illustrations: Exponential, gamma, Weibull, Pareto distributions.

4. Introduction to reliability data and related Inference (6L)

4.1 Reliability concepts and Reliability Data: Introduction (Sec. 1.1), Examples of reliability data (Sec. 1.2), General models for reliability data (Sec. 1.3), Repairable systems and Nonrepairable units (Sec. 1.4).

4.2 Parametric Likelihood: Likelihood for complete samples forexponential distribution (Sec. 7.2), confidence intervals for θ(Sec. 7.3) and for functions of θ(Sec. 7.4), Examples in the context of reliability estimation.

4.3Construction ofconfidence intervals for exponential distribution (Sec. 9.3).

Note: All sections of Chapter 4 are from the book "Statistical Methods for Reliability Data" by Meeker William and EscobarLuis (1998).

Books Recommended

1. Meeker William and Escobar Luis (1998). Statistical Methods for Reliability Data, AWiley-Interscience Publication, John Wiley & Sons, Inc.
2. Barlow R. E. and Proschan, Frank (1981). Statistical Theory of Reliability and Life Testing, Holt Rinebart and Winston Inc., New York.
3. Sinha, S. K. (1987). Reliability and Life testing, Second Edition, Wiley
4. Trivedi, R.S. (2001). Probability and Statistics with Reliability, Queuing and Computer Science Applications, Prentice - Hall of India Pvt. Ltd., New Delhi.
5. Besterfield, D.H. and Michna, C.B. et al. (2009). Total Quality Management, 3rd edition,Pearson Education, Delhi.34
6. Dodge, H.F. and Roming, H.G. (1959). Sampling Inspection tables, John Wiley and Sons, Inc.New York
7. Duncan A.J. (1974). Quality Control and Industrial Statistics, fourth edition D.B. Taraporewala Sons and Co. Pvt. Ltd., Mumbai.
8. Grant, E. L. and Leavenworth (1980). Statistical Quality Control, fifth edition, Mc-GrawHill,New Delhi.
9. Montgomery, D. C. (2008). Statistical Quality Control, 6thEdn., John Wiley, New York.

ST 366 (C): MEDICAL STATISTICS AND CLINICAL TRIALS

1. Population study **(6 L)**

1.1 India's population and census.

1.2 Population growth and models for population growth.

1.3 Birth and death rates.

2. Epidemiology **(10 L)**

2.1 Introduction to Epidemiology

2.2 Odds ratio: Properties, inference for odds ratios and Log odds ratios. Relationship between odds ratio and relative risk.

2.3 Estimation of odds ratio, Confidence interval for OR. Relation

2.4 Symmetry in square contingency tables, collapsing tables and Simpson's paradox.

3. Clinical trials **(10 L)**

3.1 General information on history of drug discovery including Louis Pasteur (rabies and small pox), Ronald Ross and malaria, Alexander Fleming and penicillin, Jonas Salk and polio, cholera, asthma, diabetes, blood pressure, heart attack, arthritis.

3.2 Phases of clinical trial, purpose, duration, cost, drug regulatory bodies, ICH, statistical analysis plan, clinical study report.

4. Design of clinical trials **(10 L)**

4.1 Phases of clinical trial, purpose, duration, cost, drug regulatory bodies, ICH, statistical analysis plan, clinical study report.

4.2 Parallel designs, case control studies, longitudinal studies, safety studies

4.3 Treatments, 2 periods cross over design.

4.4 Bioequivalence and bio-availability, non-inferiority trial

4.5 Practice based medical research, evidence based medicine

Books Recommended

1. Gore A. P. and ParanjapeS. A. (2000). Course on Mathematical and Statistical Ecology, Kluwer Publishing Holland,

2. Kulkarni M.B. andPrayag V. R.(2004). Introduction to Statistical Ecology,SIPF Academy, Nasik.

3.Agresti, A., (1996). An introduction to categorical data analysis. John Wiley & Sons.

4. Matthews J.N.S. (2006). Introduction to Randomized Controlled clinical Trials, Chapman and Hall

5. Stephen S.(2000). Statistical Issues in drug Development, John Wiley.

6. Steven D. (2000). Clinical Trials – A methodological perspective, John Wiley.

7. FriedmonL.M., ForbesC.D., DeMets D.L. (2015). Fundamentals of Clinical Trials, Spinner.
8. Selvin, S., (2004). Statistical analysis of epidemiologic data. Oxford University Press.
9. Shoukri, M.M. and Cihon, C., (1998). Statistical methods for health sciences. CRC press.
10. A.P. Gore, S.A.Paranjpe and M.B. Kulkarni. Lecture Notes on Medical Statistics.

ST 367: PRACTICAL PAPER - IV

Sr. No.	Title of the experiment	No. of experiment
1	Model sampling from Laplace distribution.	1
2	Applications of bivariate normal and Pareto distribution.	1
3	Fitting of log-normal distribution.	1
4	SPRT for Binomial, Poisson, Normal and Exponential distributions (graphical representation also)	2
5	Testing of hypotheses (Probability of type I and type II errors, power of a test etc).	1
6	Most Powerful (M.P.) test	1
7	Uniformly Most Powerful (U.M.P.) test. Plotting of power function of a test.	1
8	Non- parametric tests: Sign test and Wilcoxon's signed rank test	1
9	Mann-Whitney U test.	1
10	Run test and median test.	1
11	Kolmogorov- Smirnov test.	1
Total no. of Practicals		**12**

ST 368: PRACTICAL PAPER - V

Sr. No.	Title of the experiment	No. of experiment
1	Determination of sample size for variables and attributes.	1
2	Simple random sampling (estimation of population mean, population total with standard errors), i) with replacement, ii) without replacement. Confidence interval for population mean and population total.	1
3	Simple Random Sampling for Proportions (estimation of population proportion, population total with standard errors), confidence interval for population proportion and population total.	1

4	Stratified Random Sampling: Proportional and Neyman allocation, comparison with SRSWOR.	1
5	Cost and Variance Analysis in Stratified Random sampling.	2
6	Ratio and Regression methods of estimation. Comparison with SRSWOR.	1
7	M.L.E. of parameters of Weibull distribution using Newton-Raphson method.	1
8	M.L.E. of parameters of Gamma distribution using digamma function.	1
9	Estimation of parameter of exponential distribution for type I and typeII censored data.	1
10	Actuarial estimator of survival function and its standard error(in the presence of censored observations).	1
11	Kaplan-Meier estimator of survival function and its standard error(in the presence of censored observations). Plot of estimator against t.	1
Total no. of practicals		**12**

ST 369: PROJECT (EQUIVALENT TO 12 PRACTICALS)

1. For project maximum 5 students are allowed in a group.
2. Copy of the project report to be made available to the examiner on the first day of practical examination.
3. The following points should be included in the Project Report/ Dissertation:
   a. Title of the project, name(s) of the student(s), name of the Department and College. Acknowledgement, Data Sources, Description of the computing system/software(s), Programming language(s) used, etc. (if applicable)
   b. Motivation for selecting the topic, abstract of the project, key-words of the project.
   c. Text of the project. Broadly this should cover description of the selected problem using terminology in the field of application, conversion of the problem in statistical language, literature survey, description of collected data, small illustrative data set, methodology for the analysis, interpretation of the results, validation of the results, conclusions in statistical as well as user's language, limitation of proposed solutions, directions for future work, references used, etc.

4. The division of 50 marks for the project evaluation is,

A) Internal evaluation:	**15 Marks (Based on project report and viva)**
B) End semester evaluation	
i) Project report:	**25 Marks**
ii) Viva based on power point presentation:	**10 Marks**
Total marks (A+B):	**50 Marks**

i) Aspects of Assessment and marks assigned (The following are the guidelines, some modification can be done as and when required by the examiners)
Scheme of the marking of project report for end semester examination:

Sr.No.	Aspects in Project Report	Marks
i)	Dissertation in the proper format which includes (Title, abstract, Key words, methodology, conclusion, references, limitations and source of data etc.)	7
ii)	Conversion of the real life problem in statistical language	3
iii)	Appropriateness of tools (preferably learned in syllabus) used for analysis, testing the assumptions needed for analysis, methodology, program coding (if any) and numerical computations.	10
iv)	Conclusions in relevant language	5
Total		**25**

ii)**Viva of each project group consisting maximum 5 students is conducted using power point presentation only (Time duration: 10 Minutes).**

Division of marks for viva as:

Presentation (through slides)	**Marks**
Understating the Project problem	4
Understanding of statistical techniques used to solve the problem	4
Overall presentation	2
Total	**10**

ST 3610: INTRODUCTION TO PYTHON (PRACTICAL COURSE)

1. Introduction and Data types (11 L)

1.1 Introduction, Installation and Working with Python, Understanding Python variables, Python basic Operators, Understanding python blocks

1.2 Input and Output- print() statement, Numeric data types: int, float, complex, string data type and string operations, Defining list and list slicing, Sequences of Data, Dictionaries, Groupings of Data Indexed by Name, Special String Substitution Using Dictionaries, Arrays, Working with Sets

2. Control structures, Functions and File operations (15 L)

2.1 **Control structures**: Conditional blocks using if, else and elif for loop in python, Use of while loops in python, Loop manipulation using pass, continue, break.

2.2 **Python Functions, Modules and Packages**: Organizing python codes using functions, Importing own module as well as external modules, Understanding Packages, Powerful Lamda function in python, Programming using functions, modules and external packages.

2.3**File Operation**: Reading config files in python, Writing log files in python, Understanding read functions, read(), readline() and readlines(), Understanding write functions, write() and writelines(), Manipulating file pointer using seek, Programming using file operations

3. Manipulating Data (10 L)

3.1 **NumPy** : Introduction , NumPy Array , Array Indexing , Numpy Operations.

3.2 **Pandas** : Introduction , Series, DataFrame , Indexing and Slicing of DataFrame, handling missing Data , GroupBy , Merge Join and concatenate

3.3 **Data visualization using Matplotlib**: Simple line plot, simple scatter plot, Simple bar diagram,Subdivided bar diagram, Multiple bar diagram,Piediagram, Rod or Spike Plot, Boxplot, visualizing error, density and contour plot, histogram, binning and density, multiple subplot.

Books Recommended

1. Wes McKinney (2012). Python for Data Analysis, O'REILLY Publications,
2. Dr. R. Nageswara Rao (2017). Core Python Programming, Wiley Publication.
3. Lutz, M. (2013). Learning Python: Powerful Object-Oriented Programming. O'Reilly Media, Inc.
4. Jake Vanderplas (2017), Python Data Science Handbook, O'REILLY Publications, 2nd edition.
5. Mueller, J.P.&Massaron, L.(2015). Python for data science for dummies. Wiley & Sons.

ST 3611: DATA ANALYTICS(PRACTICAL COURSE)

1. **Introduction to Data Mining** **(8L)**
   1.1 Data preparation for knowledge discovery: Data understanding and data cleaning tools, Data transformation, Data Discretization, Data Visualization.
   1.2 Data Mining Process: CRISP and SEEMA, Supervised and unsupervised learning techniques.

2. **Classification** **(8L)**
   2.1 Problem of classification,
   2.2 classification techniques: k-nearest neighbour, decision tree, Naïve Bayesian, classification based on logistic regression

3. **Model Evaluation, Selection and Classification Accuracy** **(10L)**
   3.1 Model evaluation and selection: Metrics for Evaluating Classifier Performance,
   3.2 Concept of training data, testing data and validation of model, Cross-Validation, Bootstrap, Model Selection Using Statistical Tests of Significance,
   3.3 Techniques to Improve Classification
   3.4 Accuracy: Introduction to Ensemble Methods, Bagging,
   3.5 Boosting and AdaBoost, Improving Classification Accuracy of ClassIm balanced Data.
   3.6 Examples.

4. **ANN, SVM and Clustering** **(10L)**
   4.1 Introduction to Artificial Neutral Network (ANN) and Support Vector Machine (SVM),
   4.2 Clustering: k-means, k-mediods,
   4.3 Market Basket Analysis: Association rules and prediction, Apriori Algorithm, data attributes, applications to electronic commerce.
   4.4 Examples.

Books Recommended

1. Berson and Smith S.J. (1997): Data warehousing, Data Mining, and OLAP, McGraw-Hill.
2. Breiman J.H Friedman, R.A. Olshen and stone C.J. (1984): Classification and Regression Trees, Wadsworth and Brooks / Cole.
3. Han, J. and Kamber, M. and Pei, J. (2012): Data Mining: Concepts and Techniques. MorganGaufmann.3rd Edition.
4. Mitchell T.M. (1997): Machine Learning, McGraw-Hill.

5. Ripley B.D. (1996): Pattern Recognition and Neural Networks. Cambridge University Press.
6. Vapnik V.N. (2013). The nature of Statistical learning theory, Springer.
7. Cristianini N. and Shawe-Taylor J. (2000). An Introduction to support vector machines and other kernel-based learning methods, Cambridge University Press.
8. Mehrika, K., Mohan, C., and Ranka (1997) Elements of Artificial neural networks. Penram international.
9. Hastie T, Tibshirani R, Friedman J, (2009): The elements of statistical Learning, Springer.
10. Chattamvelli, R. (2015). Data mining methods. Alpha Science International.

— — — ****** — — —