Savitribai Phule Pune University

(Formerly Known as University of Pune)

Revised Syllabi

of

T.Y. B.Sc. Statistics (Principal)

(Semester System)

To be effective from

Academic Year 2015-2016

T.Y. B.Sc. Statistics (Principal)

The code numbers and titles of theory and Practical papers are as follows:

Semester III

ST 331: Distribution Theory

ST 332: Theory of estimation

ST 333: Sampling Methods

ST 334: Design of Experiments

ST 335: C Programming (Turbo C)

ST 336: Introduction to Regression Analysis

Semester IV

ST 341: Actuarial Statistics

ST 342: Testing of Hypotheses

ST 343: Statistical Quality Control

ST 344: Operations Research

ST 345 (A): Reliability and Survival Analysis

OR

ST 345 (B): Introduction to Stochastic processes

ST 346: Statistical Computing Using R software

Practical Papers

ST 347: Practical Paper I

ST 348: Practical Paper II

ST 349: Practical Paper III

Equivalences for the Old Courses (2010-11 to 2014-15) with New Courses (2015-16 onwards) in Statistics

T. Y. B.Sc. Statistics

Papers in Old Course (2010-11 to 2014-15)	**Equivalent papers in New Course (2015 -16 onwards)**
ST-331: Distribution Theory – I	No equivalent paper *
ST-332: Theory of estimation	ST-332: Theory of estimation
ST-333: Statistical Process Control(On line methods)	No equivalent paper *
ST-334: Design of Experiments	ST-334: Design of Experiments
ST-335: C Programming (Turbo C)	ST-335: C Programming (Turbo C)
ST-336 A) : Operations Management	No equivalent paper *
ST-336 B) : Actuarial Statistics	ST-341 : Actuarial Statistics
ST-336 C): Time Series Analysis	No equivalent paper *
ST-341 : Distribution Theory – II	No equivalent paper *
ST-342: Testing of Hypotheses	ST-342: Testing of Hypotheses
ST-343 : Statistical Process Control (off -line methods)	No equivalent paper *
ST-344 : Sampling Methods	ST-333 : Sampling Methods
ST-345 : Operation Research	No equivalent paper *
ST-346 A) : Medical Statistics	No equivalent paper *
ST-346 B) Statistical Ecology	No equivalent paper *
ST-346 C) Statistical Computing Using R – software	No equivalent paper *
Practicals	
ST-347 : Paper I	No equivalent paper *
ST-348 : Paper II	No equivalent paper *
ST-349 : 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.**

Notes

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. Students must complete all the practicals in each of the practical papers to the satisfaction of the teachers concerned.

5. Students must produce at the time of the practical examination, the laboratory journal along with the completion certificate signed by the Head of the Department.

6. The practical Examination will be conducted annually. Internal and external examination will examine the answer books jointly candidates after examination is over at the centre.

7. Of the 100 marks for each practical paper, 10 marks shall be reserved for journal and 10 marks for viva. These marks are to be given by concerned teacher in the college. A practical examination paper shall actually carry 80 marks.

8. In semester IV, a student shall opt for any one of the following papers.

ST- 345 (A): Reliability and survival Analysis

ST- 345 (B): Introduction to Stochastic Processes

All other papers of semester III and IV are compulsory.

9. Nature of practical Paper I and II

A) Continuous internal evaluation	**Marks**	**Duration**
i) Journal	10	
ii) Viva-voce	10	
Total of A	20	
B) Annual Practical examination:		
To solve 2 equations out of 4, each of 35 marks	70	3 hours on calculator 2 hours 30 minutes on computer
Viva-voce	10	10 minutes
Total of B	80	
Grand Total of A & B	100	3 hours & 10 minutes

Nature of Practical paper III

A) Continuous Internal Evaluation	**Marks**	**Duration**
i) Journal for C, R	5	
ii) Viva for C, R	5	
iii) Project Report (Dissertation)	5	
iv) Project Viva	5	
Total of A	20	
B) Annual Practical Exami		
Section I C Programming on computer, using slips prepared by University. No vivavoca conducted for this part.	20	1 hour 40 minutes for section I and II
Section II R programming on computer, using slips prepared by University. No viva – voce to be conducted for this part.	20	
Section III Project: Evaluation of Project Report (Dissertation) Viva of individual student	25 15	1 hours 30 min.
Total B	80	
Total of A & B	100	3 hours 10 minutes

10. Total duration of practical examination will be 3 hours and 10 minutes

11. Batch size should be of maximum 12 students.

12. For Practical papers I & II total time duration using (i) computer is 2 ½ hours (ii) calculator is 3 hours and additional 10 minutes for viva voce in each case.

13. For practical paper III along with internal examiner, expert be appointed to find the solution of Section I (C programming part) and Section II (R)

14. To perform and complete the practicals, it is necessary to have computing facility. So there should be sufficient number of computers, UPS, printer and scientific non programmable calculators in the laboratory.

15. 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 must be arranged.

16. For project under practical paper III (ST 349) maximum 5 students are allowed in a group.

17. Copy of the project be made available to the examiner on the first day of practical examination.

18. Use of scientific calculators and statistical tables is allowed for **theory** as well as **practical** examination.

19. Internal and External examiner will examine the answer books jointly at the centre immediately after examination is over.

20. **Guidelines for conducting University examination of Paper ST 346 Statistical computing using R software at T.Y. B.Sc. Semester IV (as per circular number 61 of 2005)**

a. The examination will be conducted in Statistics laboratory on computers.
b. Provision of at least 15 computers with necessary R software installed should be made available by the centre. Battery backup in case of power failure is essential.
c. Duration of examination is TWO hours.
d. Examination will be conducted at the time of theory examination of T.Y. B.Sc. Semester IV. The day, date and timing of this examination will be same as that of ST-346 course.
e. The examination will be conducted batch wise. A batch will consist of at most 12 candidates.
f. The batches examined simultaneously will have same question paper. However there will be separate question paper for each batch in case more (than one) batches are required to be formed.
g. A candidate will solve the question paper given to him/ her on computer and the output of work done by him/her will be evaluated by the examiner.
h. Answer book for this examination will be the answer book which is used at the time of theory examination in other science subject examination.
i. There will be no verification & revaluation of this paper.
j. If possible, the candidate will attach the printout of his/her above referred answer book.

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

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

m. 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.

n. One internal examiner and one external examiner will be appointed for this examination.

o. Examiner for this examination will be from the board of paper setters and examiners of TY B.Sc. / TY B.A. theory examination. The concerned chairman of board of examination will communicate the names of examiners to the University office.

p. Paper setting of this course will be done along with T Y B.Sc. theory examination. Manuscripts of the question papers for each batch of ST 346 will be packed separately and submitted to the University office.

q. University office will prepare 15 copies of that set and send them to the concerned centres.

r. Internal and External examiner will examine the answer books jointly at the centre immediately after examination is over. They will prepare mark list and send it to the University office / CAP Director. They will send the assessed answer books to the University.

s. For the conduct of this examination the following staff will be appointed at the centre for each batch. Two peons, one junior supervisor, one laboratory supervisor, one laboratory assistant and one machine mechanic.

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

Coursewise Syllabii

Semester III

ST 331: DISTRIBUTION THEORY

1. **Beta distribution** (9L)

1.1 **Beta distribution of first kind**

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

$= 0$, elsewhere

Notation: $X \sim \beta_1(m,n)$

Nature of probability curve,

Derivation of mean, variance, r^{th} raw moment, harmonic mean, mode.

Symmetry of the distribution .

1.2 Relation with U (0, 1).

Probability distributions of

$\frac{1}{X}$, X+Y, X-Y, XY, $\frac{X}{Y}$, where X and Y are iid β_1 (1,1),

1.3 **Beta distribution of second kind**

p.d.f $f(x) = \frac{1}{B(m,n)} \cdot \frac{x^{m-1}}{(1+x)^{m+n}}$, $x \ge 0$, $m, n > 0$

$= 0$, elsewhere

Notation: $X \sim \beta_2(m,n)$

Nature of probability curve ,

Derivation of mean, variance, r^{th} raw moment, harmonic mean, mode .

1.4 Derivation of interrelation between $\beta_1(m, n)$ and $\beta_2(m, n)$.

1.5 Derivation of distribution of $\frac{X}{Y}$, $\frac{X}{X+Y}$, when X and Y are independent gamma variates,

1.6 Statement of relation between distribution function of $\beta_1(m, n)$ and binomial distribution

2. **Weibull Distribution** (5L)

2.1 p.d.f .

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

$= 0$, elsewhere

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

2.2 Probability curve, location parameter, shape parameter, scale parameter. Derivation of distribution function, quartiles, mean and variance, coefficient of variation, relationship with gamma and exponential distribution.

3. **Order statistics** (5 L)

3.1 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 statistics $X_{(i)}$, particular cases for i = 1 and i = n. Distribution of $X_{(i)}$ for a random sample from uniform and exponential distribution. Definition of p-th sample quantile $X_{([np]+1)}$. Distribution of sample median for a random sample from uniform distribution.

4. **Cauchy Distribution** (6L)

4.1 p.d.f. $f(x) = \frac{\lambda}{\pi} \frac{1}{\lambda^2+(x-\mu)^2}$; -∞<x<∞, -∞< μ <∞, λ >0.
Notation : X~ C (μ , λ)

4.2 Nature of the probability curve, comparison with tails of normal distribution.

4.3 Derivation of distribution function, quartiles.
Non – existence of moments, Statement of distribution of ɑX + b.
Derivation of distribution of i) $\frac{1}{X}$ ii) X^2 where X~ C (0,1), Problems based on these results.

4.4 Statement of additive property for two Independent Cauchy variates, statement of distribution of the sample mean, comment on limiting distribution of $\bar{X}$.

4.5 Statement of relationship with uniform, student's t and normal distributions.

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

5.1 p.d.f. $f(x) = \frac{\lambda}{2}\exp(-\lambda|x-\mu|)$; -∞<x<∞, -∞< μ <∞, λ >0,
= 0 ; elsewhere
Notation: X~ L (μ , λ)

5.2 Nature of the probability curve .

5.3 Derivation of distribution function, quartiles.

5.4 MGF, CGF, Moments and cumulants, skewness and kurtosis

5.5 Derivation of Laplace distribution as the distribution of the difference of two i.i.d. exponential random variables with mean $\frac{1}{\lambda}$.

6. **Lognormal Distribution** (9 L)

6.1 p.d.f $f(x) = \frac{1}{(x-a)\sigma\sqrt{2\pi}} \exp\left\{\frac{-1}{2\sigma}[\log_e(x-a)-\mu]^2\right\}$, a<x, -∞< μ <∞, σ >0,
= 0 ; elsewhere
Notation: X~ LN (a ,μ , σ^2)

6.2 Derivation of relation with N (μ, σ²) distribution

6.3 Nature of the probability curve.

6.4 Derivation of moments (r-th moment of X-a), mean, variance, quartile, mode, Karl Pearson's and Bowley's coefficient of skewness and kurtosis, derivation of quartiles and mode.

6.5 Distribution of (Π Xi), when Xi's independent lognormal random variables .

7. Bivariate Normal Distribution. (9L)

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

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

-∞< x, y <∞,
-∞<μ_1 , μ_2< ∞ ,
σ_1, σ_2 >0,-1<ρ<+1

Notation: (X, Y) ~ BN(μ_1,μ_2,σ_1^2, σ_2^2,ρ),

7.2 Nature of surface of p. d. f., marginal and conditional distributions , identification of parameters , regression of Y on X, independence and

uncorrelatedness, Derivation of MGF and moments . Statement of distribution of aX+ bY+ c , X/Y.

Books Recommended

1. Arora Sanjay and Bansi Lal (1989). Mathematical Statistics (1st Edition), Satya Prakashan 16/17698 ,New Delhi .
2. Cramer H.: (1962) Mathematical Method of Stastistics , Asia Publishing House ,Mumbai
3. Gupta S.C and.Kapoor V.K: (2006). Fundamental Mathematical Statistics, Sultan Chand and Sons , 88,Daryaganj , New Delhi.
4. Hogg, R.V.and Craig A.T.(1970). Introduction Mathematical Stastistics (III rd Edition), Macmillan Publishing Company .Inc. New York
5. Lindgren B.W.: (1976) Statistical Theory (III rd Edition) Collier Macmillan international Edition , Macmillan Publishing Co.Inc. New York.
6. Mood. A.M. , Graybill , F.Bose ,D.C.: (1974) Introduction to theory of Statistics. (III rd Edition) Mc- Graw Hill Series.
7. Mukhopdhyay, P (1996). Mathematical Statistics, New Central Book Agency.
8. Rohatgi , V.K. (1975) An Introduction to probability Theory and Mathematical Statistics ,Wiley Eastern Ltd .New Delhi

.

ST 332: THEORY OF ESTIMATION

1. Point Estimation ***(4L)***

1.1 Notion of a parameter, parameter space, sample space as a set of all possible values of $(X_1, X_2, \ldots, X_n)$,general problem of estimating an unknown parameter by point and interval estimation.

1.2 Point Estimation: Definition of an estimator, distinction between estimator and estimate, illustrative examples.

1.3 Mean Square Error (MSE) of an estimator.

2. Methods of Estimation **(10L)**

2.1Method of moments: Derivation of moment estimators for standard distributions.Illustrations of situations where M.L.E. and moment estimators are distinct and their comparison using mean square error.

2.2 Definition of likelihood as a function of unknown parameter, for a random sample from i) discrete distribution ii) continuous distribution, distinction between likelihood function and p.d.f./ p.m.f.

2.3 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). Use of iterative procedure to derive

M.L.E.of location parameter μ of Cauchy distribution. Invariance property of M.L.E.

2.4 a) M.L.E. of θ in uniform distribution over i) (0, θ) ii) (- θ, θ) iii) (mθ, nθ)(m<n)
b) M.L.E. of θ in f(x ; θ)= Exp {-(x- θ)}, x > θ.
c)) M.L.E. of location parameter in Laplace distribution.

3. Criteria of Estimation

3.1 Unbiasedness (4L)

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:

a) Two distinct unbiased estimators of (θ) give rise to infinitely many estimators.

b) If T is an unbiased estimator of θ , then ø (T) is unbiased estimator of ø (θ) provided ø (.) is a linear function.

3.2 Variance of the estimator (4L)

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

3.3 Sufficiency (7L)

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 ø (T) is also sufficient for θ provided ø is a one to one and onto function.

ii) If T is sufficient for θ then T is also sufficient for ø (θ).

iii) M.L.E. is a function of sufficient statistic.

3.4 Efficiency (7L)

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.

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:

a). If MVBUE exists for θ then MVBUE exists for ø (θ) where ø (.) is a linear function.

b). If T is MVBUE for θ then T is sufficient for θ.

Comparison of variance with CRLB, relative efficiency of T1 w.r.t. T2 for (i) unbiased (ii) biased estimators. Efficiency of unbiased estimator T w.r.t. CRLB.

3.9 Asymptotic Behaviour of an Estimator (6L)
Chebychev's inequality for discrete and continuous distributions.

Consistency: Definition, proof of the following theorems:
a. An estimator is consistent if its bias and variance both tend to zero as the sample size tends to infinity.
b. If T is consistent estimator of θ and φ (.) is a continuous function, then Φ (T) is a consistent estimator of φ (θ).

4. Interval Estimation (6L)
Notion of interval estimation, definition of confidence interval (C.I), length of C.I., Confidence bounds, confidence coefficient. Definition of pivotal quantity and its use in obtaining confidence intervals.
Interval estimation for the following cases:
i) Mean (μ) of normal distribution (σ^2known and σ^2 unknown).
ii) Variance (σ^2) of normal distribution (μ known and μ unknown).
iii) Median, quartiles using order statistics.

Books Recommended

1 Dudeweitz, 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. and Craig, A.T. (1978). Introduction to Mathematical Statistics (fourthedition), Collier Macmillan International Edition, Macmillan Publishing Co. Inc.,New York
4. Kendall, M. and Stuart, A. (1943). The advanced Theory of Statistics, Vol 1, Charles and Company Ltd., London
5. Lindgren, B.W.(1976) Statistical Theory (third edition) Collier Macmillan International Edition, Macmillan Publishing Co., Inc. New York
6. Mood, A.M., Graybill, F. and Bose, D.C. (1974). Introduction to the theory of Statistics (third edition) International Student Edition, McGraw Hill.
7. Rohatagi , V.K. (1975). An introduction to Probability Theory and Mathematical Statistics, Wiley Eastern Ltd., New Delhi.
8. Ramchandran, K.M. and Tsokos C. P. (2009). Mathematical Statistics with Applications, Academic Press.

.

ST 333: SAMPLING METHODS

1. Sampling (8 L)

1.1 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) definitions, population mean and population total as parameters, inclusion probabilities.

1.2 (a) Sample mean $\bar{y}$ as an estimator of population mean,

derivation of expectation and standard error of $\bar{y}$, confidence interval for population mean, population total standard error.
(b) N $\bar{y}$ as an estimator of population total, derivation of expectation and standard error of N $\bar{y}$

(c) Estimator of above standard errors, both in case of SRSWR and SRSWOR.

1.3 Sampling for proportion as an application of a simple random sampling with X_i as zero or one.
(a) sample proportion as an estimator of population proportion of units possessing a certain attribute, derivation of expectation and standard error of (p).
(b) N_p as an estimator of total number of units in the population possessing a certain attribute, derivation of expectation and standard error of N_p
(c) Estimator of above standard error both in case of SRSWR and SRSWOR.

2 Determination of Sample Size (8 L)

Determination of the sample size for the given:
i) Margin of error and confidence coefficient.
ii) Coefficient of variation of the estimator and confidence coefficient.

.

3. Stratified Random Sample Size (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) $\overline{\overline{y_{st}}} = \frac{\sum N_i \overline{y_i}}{N}$ as an estimator of population mean ($\bar{Y}$),

Derivation of expectation and standard error of $\overline{\overline{y_{st}}}$.

(b) N $\overline{\overline{y_{st}}}$ as an estimator of population total, derivation of expectation and standard error of N $\overline{\overline{y_{st}}}$.

(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's allocation.
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.Ratio and Regression Methods of Estimation for SRSWOR (6L)

4.1 Rationale behind using auxiliary variates in estimation.
4.2 Situations where (i) ratio method is appropriate, (ii) regression method is appropriate.
4.3 Ratio and regression estimators of the population mean and population total.
4.4 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).

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

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

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

5.3 Comparison of systematic sampling with SRSWOR.

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

6. Role of Sample Surveys in Research Methodology (8L)

6.1 Objectives of a sample survey.

6.2 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

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

6.4 Sampling and non-sampling errors with illustrations.

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

Books Recommended

1. Cochran, W.G.(1977) Sampling Techniques, third EditionWiley Eastern Ltd., New Delhi.

2. Malhotra N. (2008). Marketing Research and Applied Orientation (third edition), Prentice Hall of India.
3. Mukhopadhyay P (2008). Sampling theory and methods of survey sampling. Prentice-Hall of India, New Delhi.

4. Murthy, M. N. (1967). Sampling methods, Indian Statistical Institute, Kolkata.
5. Singh, D. and Chaudhary, F. S. (1986). Theory and Analysis of Sample Survey Designs, Wiley Eastern Ltd., New Delhi.
6. Sukhatme, P.V., Sukhatme, B. V.(1984). Sampling theory of Surveys with Applications, Indian Society of Agricultural Statistics, New Delhi.

ST 334: DESIGN OF EXPERIMENTS

1. Design of Experiments (24 L)

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

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

1.3 Basic principles of design of experiments : 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.

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 graphically. 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 $H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_t = 0$. 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).

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

$H_{01} : \alpha_1 = \alpha_2 = \alpha_3 = \ldots = \alpha_t = 0$

$H_{02} : \beta_1 = \beta_2 = \beta_3 = \ldots = \beta_b = 0$

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 critical difference(CD).

1.6 Latin Square Design (LSD): Application of the principles of design of experiments in LSD, layout,

Model : $X_{ij(k)} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ij(k)}$ $i = 1,2, \cdots m$, $j = 1,2\cdots,m$, $k=1,2\cdots,m$.

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 critical difference, linear treatment contrast and testing its significance.

1.7 Linear treatment contrasts, orthogonal contrasts. Scheffe's method for comparing contrasts, Tuckey's procedure for comparing pairs of treatment means(applicable to C.R.D., R.B.D. and L.S.D.)

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

1.9 Analysis of non- normal data using.

i) Square root transformation for counts.

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

iii) Kruskal Wallis test.

2. Efficiency of Design (5L)

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. Analysis of Covariance (ANOCOVA) with One Concomitant Variable (7L)

3.1 Situations where analysis of covariance is applicable.

3.2 Model for covariance in CRD, RBD. Estimation of parameters (derivations are not expected)

3.3 Preparation of analysis of variance – covariance table, test for $\beta=0$, test for equality of treatment effects (computational technique only).

4. Factorial Experiments (12L)

4.1 General description of m^n factorial experiment, 2^2 and 2^3 factorial experiments arranged in RBD.

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

4.3 Yate's procedure, preparation of ANOVA table, test for main effects and interaction effects.

4.4 General idea of confounding in factorial experiments.

4.5 Construction of layouts in total confounding and partial confounding in 2^2 and 2^3 factorial experiments.

4.6 Total confounding (confounding only one interaction) ANOVA table, testing main effects and interaction effects.

4.7 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. (1968) Experimental Design, John Wiley and Sons, Inc., New York.
2. Dass, M.N. and Giri,N.C. (1986) 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. (1952). Design of Experiments, Wiley Eastern Ltd., New Delhi.
8. Montgomery, D.C. (2001). 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-335: C PROGRAMMING (Turbo C)

1. Introduction (11 L)

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

2. Control Structures (15L)

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

3. Arrays (11L)

3.1 Concept, declaration, definition, initialization of array, problem using arrays, passing to function, arrays and string operations, string functions like strcpy(), strcat(), strlen(), strcmp(), strrev().
List of programs using arrays.

3.2 To find mean, median, variance and coefficient of variation of frequency distribution.
3.3 To find correlation coefficient and least square regression line of Y on X for a given bivariate data.
3.4 To arrange the given data in increasing/decreasing order of magnitude.
3.5 To obtain median of given n observations.
3.6 To obtain addition of two matrices, multiplication of two matrices.

4. Functions (7L)

Declaration, definition, recursion, user defined functions, library function, calling a function by reference and by value, local and global variables.
List of writing functions:
1. To find factorial of integer number (both recursive and non-recursive)
2. To find the value of $X^n$where n is integer.(both recursive and non-recursive)
3. To find GCD of two integer numbers(both recursive and non-recursive)
4. To find maximum/minimum of n numbers.(non-recursive)

5. Pointers (4L)

Basic concept and relation to one dimensional array.

List of Simple Programs (short programs)

1. Converting °C temperature to °F.
2. To carry out arithmetic calculations.
3. To check whether given number is odd or even.
4. To check whether given number m is divisible by n or not.
5. To find maximum of 2 numbers or 3 numbers.
6. To find area of triangle and circle.
7. To find roots of quadratic equation.
8. To check whether integer is prime or not.
9. To find mean, Geometric mean and Harmonic Mean of n numbers.
10. To prepare multiplication table.
11. To find sum of digits of a number.
12. To solve simultaneous linear equations.(two equations in two variables)
13. To evaluate simple and compound interest
14. To solve transcendental equations using Newton-Raphson method.
15. To evaluate exp(x), sin (x), log (x) etc. using Taylor‟s series expansion.
16. To convert decimal number to equivalent binary number.
17. To generate Fibonacci series like 0, 1, 1,2,3,5…
18. To test palindrome string using string function.
19. To sort a string using string function.
20. To search string using string function.
21. To combine given two strings using string function.

List of programs (long programs)

22. Program in C to prepare a frequency distribution with given class interval from raw data.
23. Program in C to find mean, variance,standard deviation and quartiles for given n observations and frequency distribution.
24. Program in C to fit a Binomial distribution to given data.

25. Program in C 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.

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.,

ST 336: INTRODUCTION to REGRESSION ANALYSIS

1. Simple linear regression model **(14L)**

(i) Review of simple linear regression model: $Y = \beta_0 + \beta_1 X + \varepsilon$, where ε is a continuous random variable with $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2$. Estimation of β_0 and β_1, by the method of least squares.

(ii) Properties of estimators of β_0,and β_1 (Sec. 2.2.2)

(iii) Estimation of σ^2 (Sec. 2.2.3)

(iv) Assumption of normality of ε .Tests of hypothesis of β_1 (Sec. 2.3)

(v) Interval estimation in simple linear regression model (Sec 2.4)

(vi) Coefficient of determination (Sec 2.6)

(vii) Residual analysis (Sec 4.1, 4.2.1): Standardized residuals, Studentized residuals, sec. 4.2.2), residual plots (sec. 4.2.3)

(viii) Detection and treatment of outliers (Sec 4.4)

(ix) Interpretation of four plots produced by lm command in R

2. Multiple linear regression model **(20L)**

(i) Review of multiple linear regression model $Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \varepsilon$, where ε is a continuous random variable with $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2$. Estimation of regression parameters $\beta_0, \beta_1, \ldots$ and β_p by the method of least squares, obtaining normal equations, solutions of normal equations. (Sec. 3.2.1)

(ii) Estimation of σ^2 (Sec. 3.2.4)

(iii) Assumption of normality of ε .Tests of hypothesis of regression parameters (Sec. 3.3.1, 3.3.2)

(iv) Interval estimation in simple linear regression model (Sec 3.4.1)

(v) Variable selection and model building (9.1, 9.2)

(vi) Residual diagnostics and corrective measures such as transformation of response variable, (Sec. 5.1, 5.2, 5.3) weighted least squares method (Sec. 5.5, (except 5.5.1 and 5.5.2), 5.5.3.

(vii) Polynomial regression models (Sec. 7.1, 7.2.1)

3. Logistic regression model **(14L)**

(i) Binary response variable, Logit transform, estimation of parameters, interpretation of parameters.(Sec. 13.2.1, 13.2.2, 13.2.3 or chapter 1except sec.1.5 of HL)

(ii) Tests of hypotheses of model parameters, model deviance, LR test.(Sec. 13.2.4 or chapter 1except sec.1.5 of HL)

(iii) AIC and BIC criteria for model selection

(iv) Interpretation of output produced by glm command in R

(v) 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 Vinig, G.C.

1. (ii) HL stands for the book "Applied Logistic Regression" by Hosmer, 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 Lemeshow, 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.

Semester IV

ST-341: ACTUARIAL STATISTICS

1) Insurance Business (3L)

1.1 Insurance companies as business organizations.
1.2 Role of insurance business in Economy.
1.3 Concept of risk, types of risk, characteristics of insurable risk.
1.4 Working of insurance business, introduction of terms such as premium, policy, policyholder and benefit.
1.5 Role of Statistics in insurance.
1.6 Insurance business in India.

2) Feasibility of Insurance Business (4L)

2.1 Measurement of adverse financial impact, expected value principle.
2.2 Concept of utility function
2.3 Feasibility of insurance business.
2.4 Illustrative examples.

3) Survival Distribution and Life Tables (12L)

3.1 Time- until death random variable, its d.f. and survival function in actuarial notation.
3.2 Force of mortality.
3.3 Interrelations among d.f., survival function, force of mortality and p.d.f.
3.4 Curtate future life random variable, its p.m.f. and survival function in actuarial notation.
3.5 Construction of life table using random survivorship approach.

4) Models for Life Insurance (11 L)

4.1 Theory of compound interest, effective rate of interest, discount factor.
4.2 Insurance payable at the end of the year of death, present value random variable, actuarial present value.
4.3 Derivation of actuarial present value for n-year term life insurance, whole life insurance and endowment insurance.

5) Annuities (10L)

5.1 Annuities – certain, annuity due, annuity immediate.
5.2 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.

6) Benefit Premiums (8L)

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

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, India.

ST 342: TESTING OF HYPOTHESES

1. Parametric Tests (15 L)

1.1 (a) 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)
Notion of likelihood ratio test (LRT), $\lambda(x) = \text{Sup } L(\theta_0|x) / \text{Sup } L(\theta|x)$ Construction of LRT for $H_0 : \theta = \theta_0$ against $H1: \theta \neq \theta_0$ for the mean of normal distribution for i) known σ^2 ii) unknown σ^2 (one sided and two sided alternatives). LRT for variance of normal distribution for i) known μ ii) unknown μ (one sided and two sided alternatives hypotheses). LRT for parameters of binomial and exponential distribution for two sided alternatives only. LRT as a function of sufficient statistics , statement of asymptotic distribution of $-2 \log_e \lambda(x)$.

3. Sequential Tests (9 L)
Sequential test procedure for simple null hypothesis against simple alternative hypothesis and its comparison with fixed sample size N-P test procedure. Definition of Wald's SPRT of strength (α, β). Illustration for standard distributions like Bernoulli, Poisson, Normal and Exponential. SPRT as a function of sufficient statistics. Graphical representation of SPRT.

4. Non-parametric Tests (15 L)
4.1 Concept of non- parametric tests. Distinction between a parametric and a 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, (iii) Run test, one sample and two samples problems
4.2 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. Agarwal, B.L. (2003). Programmed Statistics, second edition, New Age International Publications , Delhi
2. Arora, S. and Bansi Lal. (1989) : New Mathematical Statistics, first edition, Satya Prakashan, New Delhi.
3. Daniel, W.W. (2000) Applied Nonparametric Statistics, Duxbury Press Boston.
4. Dudeweitz, E.J. and Mishra, S.N. (1988). Modern Mathematical Statistics, John Wiley and Sons, Inc
5. Gibbons J.D.(1971). Non parametric Statistical Inference, McGraw Hill Book Company, New York.
6. Hoel, P.G.,Port, S. and Stone, C. (1971). Introduction to Statistical Theory, Houghton Mifflin Company (International) Dolphin

7. Hogg, R.V. and Craig, R.G. (1989). Introduction to Mathematical Statistics (fourth edition,.Collier Macmillan International Edition, Macmillan Publishing Co. Inc., New York.

8. Kale, B.K. and Muralidharan, K. (2015). Parametric Inference: An Introduction. Narosa, New Delhi
9. Kendall, M. and Stuart, A. (1943) The advanced Theory of Statistics, Vol 1, Charles and Company Ltd., London.
10. Lindgren, B.W.(1976). Statistical Theory (third edition), Collier Macmillan International Edition,Macmillan publishing Co., Inc. New York.
11. Kunte, S ., Purohit, S.G. and Wanjale, S.K. : Lecture Notes On Nonparametric Tests
12. Mood, A.M., Graybill, F. and Bose, D. C .(1974). Introduction to the theory of Statistics (third edition) International Student Edition, McGraw Hill.
13. . Rohatgi, V.K. (1976). An introduction to Probability Theory and Mathematical Statistics, Wiley Eastern Ltd., New Delhi.
14. Siegel, S. (1956). Nonparametric methods for the behavioral sciences, International Student Edition, McGraw Hill,New York.

ST 343: STATISTICAL QUALITY CONTROL

1. Introduction (4 L)

Meaning and purpose of Statistical Quality Control (SPC), on line process control methods
(control charts) and offline process control methods (Sampling plans).

Seven Process Control (PC) Tools of SPC

(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 (24L)

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.
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.
Control chart technique as hypotheses testing problem.
Construction of control charts for (i) standards given, (ii) standards not given.

2.2 Control charts for variables

(I) R chart and $\bar{X}$ chart
Purpose of R and $\bar{X}$ chart, normal probability plot for checking normality assumption, 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.

(II) 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. Construction of $\bar{X}$ chart when the process average μis not given : control limits based on $\hat{\mu} = \bar{\bar{X}}$, $\hat{\sigma} = \frac{\bar{R}}{d_2}$, drawing of control chart, plotting sample means, revision of control limits of $\bar{X}$ chart, if necessary. Probability of catching a shift.

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 $\bar{R}$ or $\hat{\sigma}$.

Estimate of μ and σ for further use. Determination of state of control of the process. Identification of real life situations where this technique can be used. Limitations of $\bar{X}$, R charts.

2.3 **Control charts for Attributes**

(I) 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 separate control limits, 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.

(II) **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.

3. Capability Studies (6 L)

3.1 Specification limits, natural tolerance limits and their comparisons, decisions based on these comparisons, estimate of percent defectives.

3.2 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.*

4. Offline Methods (Lot Control)
Acceptance Sampling for Attributes
(14L)

4.1 **Introduction:** Concept of sampling inspection plan, comparison between 100% inspection and sampling inspection. Procedure of acceptance sampling with rectification.

4.2 **Single Sampling Plan**: Working of SSP, Evaluation of probability of acceptance using Poisson distribution. Producer's risk. Consumer's risk, Acceptable Quality Level (AQL). Lot Tolerance Fraction Defective (LTFD), LTPD, Average Outgoing Quality (AOQ), Average Outgoing Quality Limit (AOQL), Average Sample Number (ASN), Average Total Inspection (ATI), Operating characteristic (OC) curve, AOQ curve.

4.3 **Double Sampling Plan**: Working of DSP, Evaluation of probability of acceptance using Poisson distribution. Producer's risk. Consumer's risk, O.C.curve, Average Outgoing Quality (AOQ), AOQ curve, Average Outgoing Quality Limit (AOQL), Average Sample Number (ASN), Average Total Inspection (ATI) (with complete inspection of second sample).

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. (1983). Statistical Quality Control, John Wiley and Sons, Inc., New York.
8.SP20 : Handbook of SQC, Bureau of Indian Standards

ST 344: OPERATIONS RESEARCH

1) Linear Programming (16 L)

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

1.2 Solution of L.P.P by Simplex Method:
Obtaining Initial Basic Feasible Solution (IBFS) , criteria for deciding whether obtained solution is optimal ,criteria for unbounded solution , no solution , more than one solutions , introduction of artificial variable, Big-M method.

1.3 Duality Theory: Writing dual of a primal problem, solution of a L.P.P. by using its dual problem.

1.4 Examples and problems.

2) Transportation Problem (12 L)

2.1 Transportation problem (T.P.), statement of T.P., balanced and unbalanced T.P. Minimization and maximization problem.
2.2 Obtaining basic feasible solution of T.P. by (i) Least cost method (ii) Vogel's approximation method (VAM).
2.3 u-v (MODI) method of obtaining Optimal solution of T.P., uniqueness and non- uniqueness of optimal solutions, degenerate solution
2.4 Assignment Problem : Statement of an assignment problem , Minimization and maximization problem , balanced and unbalanced problem ,relation with transportation problem , optimal solution using Hungarian method , maximization case
2.6 Examples and problems

3. Simulation

(9 L)

3.1 Introduction to simulation, merits, demerits, limitations.
3.2 Pseudo random number generates: Linear congruential , mid square method.
3.3 Model sample from normal distribution (using Box- Muller transformation), uniform distribution, exponential distribution.
3.4 Monte Carlo method of simulation: Statistical applications of simulation in numerical integration such as computation of probabilities of events related to gamma, beta and bivariate normal distribution.

4. Critical Path Method (CPM) and Project Evaluation and Review Techniques (PERT) (11 L)

4.1 Definition of (i) Event,(ii) Node,(iii)Activity,(iv)Critical Activity,(v)Project Duration.

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

4.3 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.

Books Recommended

1. Gass, S.L. (1997). Linear programming methods and applications, Narosa Publishing House, New Delhi.
2. Gupta, P.K. and Hira, D.S.(2008). Operation Research, 3rd edition S. Chand and company Ltd., New Delhi.
3. Kapoor, V. K.(2006). Operations Research, S. Chand and Sons. New Delhi.
4. Phillips, D.T and Solberg, R.A.(1976). Operation Research principles and practice, John Willey and sons Inc.
5. Saceini, M., Yaspan,A.. and Friedman, L.(2013). Operation Research methods and problems, Willey International Edition.
6. Sharma, J.K. (1989). Mathematical Models in Operation Research, Tata McGraw Hill Publishing Company Ltd., New Delhi.
7. Shrinath.L.S (1975). Linear Programming, Affiliated East-West Pvt. Ltd, New Delhi.
8. Taha, H.A. (2007). Operation research: An Introduction, eighth edition, Prentice Hall of India, New Delhi.

ST 345 (A): RELIABILITY and SURVIVAL ANALYSIS

1. Structural Properties of coherent system (15 L)

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 parallel and series system, reliability block diagram, guidelines for construction of reliability block diagram.

1.2 Coherent structure function (maximum 4 components), relevent 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). , 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 importance of components, module of the coherent system, modular decomposition of coherent system.

2. Reliability of coherent system (8 L)

Reliability of system of independent components, Basic properties of system reliability (such as reliability function is increasing function, system and

component redundancy etc.), 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 .

3. Ageing Properties (19 L)

3.1 Survival function , probability density function, hazard function, cumulative hazard rate, mean residual life function, equilibrium residual life function , interrelation between all these function, no ageing, proof of following properties of no ageing

1) Cauchy functional equation
2) Constant failure rate
3) Constant mean residual life
4) Exponential life distribution

3.2 Positive and negative ageing: IFR, DFR IFRA, DFRA, NBU, NWU, NBUE, NWUE, bathtub failure rate.

Classification of following parametric families of life distribution according to aging**:** Weibull, Gamma, lognormal, linear failure rate, Makeham, Pareto, Lehman.

4. Censoring and Nonparametric estimation of survival function (6 L)

Concept of censoring, order censoring, time censoring, right random censoring, left random censoring, undersigned censoring, Nonparametric estimation of survival function, confidence band on survival function, actuarial estimator of survival function, green wood'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 345 (B): INTRODUCTION to STOCHASTIC PROCESSES

1. Definition of a Stochastic process, state space ,parameter space, types of stochastic processes , Markov chains (MC) { X_n, n ≥ 0}, finite MC, time homogeneous M.C. one step transition probabilities, and transition probability matrix (t.p.m.),stochastic matrix, Chapman Kolmogorov equation, n-step transition probability matrix , initial distribution, joint distribution function of {X_0, X_1, . . . , X_n} , partial sum of independent and identically distributed random variables as Markov chain, illustrations such as random walk, Gambler's ruin problem, Ehrenfest chain. (18L)

2. Classification of states: Communicating states, first return probability, probability of ever return Classification of states, as persistent and transient states . Decomposition of state space, closed set of states, irreducible set of states, irreducible MC, periodicity of M.C. aperiodic M.C. ergodic M. C. (12L)

3. Stationary distribution for an irreducible ergodic finite Long run behaviour of a M.C**.** (6L)

4. **Poisson process**: Postulates and properties of Poisson process, probability distribution of N(t), the number of occurrences of the event in (0,t], Poisson process and probability distribution of inter-arrival time ,mean, variance and covariance functions . Definition of compound Poisson process mean and variance functions and its applications. (12L)

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, 7th edn, 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 modelling Academic Press.

ST–346: STATISTICAL COMPUTING USING R SOFTWARE

Note: Students are expected to write commands in script file wherever applicable

1. Fundamentals of R (5L)

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

Simple bar diagram,Subdivided bar diagram,multiple bar diagram,Piediagram,Stem and leaf diagram

3. Graphs (5 L)

Boxplot for one and more than one variables, rod or spike plot, histogram for raw data with prob=T option and for both equal and unequal class intervals, frequency polygon,ogivecurves,empirical distribution function
Saving the diagram and graph in MS-Word file.

4. Measures of central tendency, dispersion, skewness and kurtosis. (6 L)

Computations of following measures for all types of data
(a)central tendency mean,mode,median,quartiles,deciles,percentiles , g.m. and h.m

(b)Dispersion:variance,standard deviation, coefficient of variation,mean deviation
(c)Skewness:Bowley"scoefficient and KarlPearson"scoefficient of skewness
(d) Moments: Computations of raw and central moments,measure of skewness and kurtosis based on it.

5. Probability distributions: (6 L)

SImulation from distributions, computations of probabilities,cumulative probabilities, quantiles and drawing random sample using d,p,q,r functions for following distributions.
Binomial,Poisson,Hypergeometric,normal,exponential,gamma,Cauchy,lognormal, Weibull,uniform,laplace ,Graphs of pmf/pdf by varying parameters for above distributions.
Fitting of Poisson and normal distribution, testing normality of data by Shapiro-Wilks test.

6 Testing of hypothesis (1 L)

Chi-square test for independence of attributes

7. ANOVA (5 L)

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

8. Non parametric tests (4L)

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

9. Programming in R: (12L)

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

Writing programs in R.

1. Formation of strata clusters using specified 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: $\sigma_1^2 = \sigma_2^2$ and then applying appropriate test.
7. Verifying the assumptions in one way ANOVA and then applying appropriate test.
8. Testing consistency
9. Testing normality of number of samples
10. 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.; Gore, S.D. and Deshmukh, S.R. (2015). Statistics using R, second edition. Narosa Publishing House, New Delhi.

3. Shahababa , B. (2011). Biostatistics with R, Springer, New York

4. Verzani, J. (2005). Using R for Introductory Statistics, Chapman and Hall /CRC Press, New York

ST 347: PRACTICAL PAPER I

Sr No.	Title of the experiment	No. of experiments
1.	Model sampling from Cauchy and Laplace distributions	(1)
2.	Fitting of lognormal distribution.	(1)
3.	Construction of confidence interval for population median and quartiles, based on order statistics.	(1)
4.	Testing of hypotheses (Probability of type I and type II errors, power of a test etc).	(1)
5.	Construction of uniformally most powerful (UMP) test, plotting of power function of a test.	(2)
7.	Non- parametric tests : Sign test, Wilcoxon's signed rank test , Mann-Whitney U test.	(2)
8.	Non- parametric tests : Run test, median test,	(1)
9.	Kolmogorov- Smirnov test.	(1)
10.	SPRT for Bernoulli, Binomial, Poisson, Hypergeometric distributions. (graphical representation also)	(1)
11.	SPRT for normal, exponential distribution (graphical representation also)	(1)
12.	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)
13.	Simple random sampling for proportions . (estimation of population proportion, population total with standard errors), confidence interval for population proportion and population total.	(1)
14.	Stratified random sampling : Proportional and Neyman allocation, comparison with SRSWOR.	(1)
15.	Stratified random sampling : cost and variance analysis.	(1)
16.	Ratio and Regression methods of estimation. Comparison with SRSWOR.	(2)
17.	M.L.E and moment estimator of truncated Binomial and truncated Poisson distributions (truncated at zero).	(1)
18.	Problems based on accumulated value, present value and effective rate of discount.	(1)
19.	Construction of life table.	(1)
20.	Computation of benefit premium for n-year term insurance, whole life insurance and endowment insurance.	(1)
	Total number of practicals.	(23)

ST 348: RACTICAL PAPER II

Sr. No. Title of the Experiment No.of experiments

1. Analysis of CRD (equal and unequal replications) (1)
pairwise comparison of treatments, using critical difference (C.D). Check normality using normal probability plot.

2 Analysis of R.B. D. pairwise comparison of treatments using (2)
i) C.D ii) Tukey and Scheff's procedure. Efficiency of RBD w.r.t. CRD
3. Analysis of L.S.D., pairwise comparison of treatments using C.D .(1)
and box plot, efficiency of LSD w.r.t. i) CRD ii) RBD.
4. Kruskal Wallis test (1)
5. Analysis of covariance in CRD, testing B = 0, (1)
6. Analysis of covariance in RBD , testing B= 0(1)
7. Analysis of 2^2 and 2^3 factorial experiments in RBD. (2)
8. Analysis of 2^3 factorial experiments in RBD (partial confounding) (1)
9. Analysis of 2^3 factorial experiments in RBD. (total confounding) (1)
10. R , $\bar{X}$ chart, probability of detecting shift, .for $\bar{X}$ chart, computations of C_p, $C_{\bar{pk}}$. (2)
11. p-chart for (i)fixed sample size (ii) variable sample size based on individual control limits ,probability of detecting shift. (1)
.
12. Single sampling plan for attributes (OC curve, AOQ,(1)
AOQL, ATI using Poisson distribution).
13. Determination of single sampling plan for attributes by (1)
i) lot quality approach ii) average quality approach
14. Double sampling plan for attributes(1)
(OC curve, AOQ, AOQL, ATI, ASN using Poisson distribution).
15. Linear programming problem I (Simplex method) (1)
16. Linear programming problem II (Simplex method) (1)
17. Transportation problem and assignment problem (2)
18. CPM
19 PERT
20. Simulation: (2)
i) Simulation from standard probability distributions: a) normal distribution using Box -Muller transformation, b) exponential distribution.
ii) Simulation of approximate integrals to find numerical value of gamma function, beta function with real arguments etc

Note: Answers of experiment No. 15 to 17 are to be verified using TORA package.

ST-349: PRACTICAL PAPER III

Section : I : C –Programming (List of programs using C)

1.A) Converting °C temperature to °F.
B) To check whether given number is odd or even.
C) To find maximum of 2 numbers or 3 numbers. D) To find area of triangle and circle.

2.A) To find roots of quadratic equation.
B) To solve transcendental equations using Newton-Raphson method.

3.A) To find mean, variance and coefficient of variation of n observations.
B) Arrange the observations in ascending order of magnitude and find median of n observations.

4. To find mean, median, variance and coefficient of variation of frequency distribution when fi and xi are given.

5. To find correlation coefficient for a given bivariate data.

6. To fit a line of regression of Y on X for a given bivariate data.

Section II: R programming (List **of Programs based on R)**

7. Simple regression analysis and dignostics by graphical method

8. Multiple regression analysis and dignostics by graphical method

9. Logistic regression

10. Computing mean, median, mode, quartile , decile , percentile for n observations

11. (a) Computation of probabilities of type
P [a < X <b], P[X >= a],P[X < b] etc for Standard discrete and continuous distributions
(b) Model sampling from standard discrete and continuous distributions
(Use of functions starting with d , p, q, r)

12. Non-parametric tests

Section III
Project: Equivalent to 12practicals
Data Analysis, report in dissertation form