Khandesh College Education Society's

Moolji Jaitha College, Jalgaon

An "Autonomous College" Affiliated to

KBC North Maharashtra University, Jalgaon

KNOWLEDGE IS POWER

ESTD. 1945

SYLLABUS

Statistics

T.Y.B. Sc.

(Semester V & VI)

Under Choice Based Credit System (CBCS)

[w. e. f. Academic Year: 2021-22]

Program Specific Outcomes:

On completion of the B.Sc. (Statistics) students are able to:

- Serve as a statistician with sound theoretical, practical and computational skills.
- Work as researcher for formulation and solution of mathematical, scientific, societal and industrial problems.
- Understand the role of statistics in science, society and for National Development.
- Apply some discrete and continuous probability distributions which are highly useful in modelling real life.
- Investigate the relationship between a variable of interest (the response) and a set of related predictor variables and formulate and fit the appropriate regression model to the given dataset.
- Serve as Administrators/Investigators in the private as well as government sectors and worked as Analyst in Manufacturing (SQC Unit), Pharmaceutical industries.
- Work in service industries as consultant & analyst such as Banking and Insurance, Forest, Telecom, Transports, Hotel etc.
- Recognize and make appropriate use of different statistical softwares such as R, Python, MINITAB, SPSS, MS-EXCEL etc.
- Work as Programmer to write script using Python and R language and develop algorithms.
- Convert large amount of complex data into summarized form using graphical representation tools that helps in visual interpretation.
- Serve as Data Scientist/ Data Analyst in various companies of IT sector, Health Sector, Agricultural sector etc.

Learning Objectives:

- To prepare graduates who are not only statistically sound but also capable of using their appropriate statistical skills in interdisciplinary areas.
- To familiarize students with computational techniques and software used in the statistical arena.
- To provide a solid ground in the best practices of collating and disseminating information.
- To prepare students for undertaking further study.
- To teach students to construct practical statistical models for several processes in the real-world.

Exam Pattern:

- Each theory and practical course will be of 50 marks comprising of 10 marks internal and 40 marks external examination.

External Theory Examination (40 marks):

- External examination will be of two hours duration for each theory course. There shall be 4 questions each carrying equal marks (10 marks each) while the tentative pattern of question papers shall be as follows:

- Q1 (A), Q2 (A), Q3 (A) and Q4 (A), each will be of 6 marks (attempt any 2 out of 3 sub-questions).
- Q1 (B), Q2 (B), Q3 (B) and Q4 (B), each will be of 4 marks (attempt any 1 out of 2 sub-questions).

External Practical Examination (40 marks):

- Practical examination shall be conducted by the respective department at the end of the semester. Practical examination will be of minimum 3 hours duration and shall be conducted as per schedule. There shall be 05 marks for journal, 5 marks for *viva-voce*. Certified journal is compulsory to appear for practical examination.

Internal Theory/ Practical Examination (10 marks):

- Internal theory assessment of the student by respective teacher will be comprehensive and continuous, based on written test/ assignment. The written test may comprise of both objective and subjective type questions.
- Internal practical examination will be conducted by department as per schedule given. For internal practical examination student must produce the laboratory journal of practicals completed along with the completion certificate signed by the concerned teacher and the Head of the department.

Structure of T.Y.B.Sc. (Statistics) Curriculum Semester V

	Core I	ST-351	Distribution Theory-I	3	3	45	10	40
	Core II	ST-352	Statistical Inference-I	3	3	45	10	40
	Core III	ST-353	Design of Experiments	3	3	45	10	40
	Core IV	ST-354	Sampling Theory	3	3	45	10	40
	Core V	ST-355	Introduction to Regression Analysis	3	3	45	10	40
	Core VI	ST-356 (A) or ST-356 (B)	Statistical Quality Control or Introduction to Stochastic Processes	3	3	45	10	40
	Skill Based	ST-350	Introduction to Python	2	2	30	10	40
	Core (Practical)	ST-357	Statistics Practical-I	2	4 /Batch	60	10	40
		ST-358	Statistics Practical-II	2	4 /Batch	60	10	40
		ST-359	Statistics Practical-III	2	4 /Batch	60	10	40

Structure of T.Y.B.Sc. (Statistics) Curriculum Semester VI

	Core I	ST-361	Distribution Theory-II	3	3	45	10	40
	Core II	ST-362	Statistical Inference-II	3	3	45	10	40
	Core III	ST-363	C Programming	3	3	45	10	40
	Core IV	ST-364	Elements of Clinical Trials	3	3	45	10	40
	Core V	ST-365	Optimization Techniques	3	3	45	10	40
	Core VI	ST-366 (A) or ST-366 (B)	Official Statistics and Applied Statistics or Actuarial Statistics	3	3	45	10	40
	Skill Based	ST-360	Data Analysis using SPSS	2	2	30	10	40
	Core (Practical)	ST-367	Statistics Practical-IV	2	4 /Batch	60	10	40
		ST-368	Statistics Practical-V	2	4 /Batch	60	10	40
		ST-369	Statistics Practical-VI (Project)	2	4 /Batch	60	10	40

DSC: Discipline Specific Core Courses/Core Practical; SEC: Skill Enhancement Course; Int : Internal examination; Ext : External examination

T.Y. B.Sc. (Statistics): Semester V
Discipline Specific Core (DSC) Course
ST-351: Distribution Theory-I

Total Hours: 45 **Credits: 3**

Course objectives:

- To continue study of standard discrete and continuous probability distributions and their applications.
- To introduce truncated distributions, Chebychev's inequality, WLLN, CLT and order statistics.
- To illustrate examples on CLT and WLLN.

Course outcomes:

Students will be able to

- Use Chebychev's inequality and WLLN to solve statistical problem.
- Compute various events probability using Central Limit Theorem.
- Apply hyper geometric and negative binomial distribution in real life situations.
- Apply truncated distribution in real life situations.
- Obtain distributions of order statistics.

Unit-I: Probability Distributions **(8 h)**

- Random variable, its probability function and cumulative distribution function (cdf).
- Joint probability functions of bivariate r.v.s.
- Marginal and conditional distributions. Independence of r.v.s.
- Expectation of a r.v. and its properties. Moments, measures of location and dispersion of a r.v.
- Probability generating function (pgf) and moment generating function (mgf), cgf, fmgf, characteristic function of a r.v., their properties and uses. Indicator r.v. and Degenerate r.v.
- Discrete uniform, binomial, Poisson, geometric, exponential and normal distributions.
- Reproductive (additive) property of standard distributions.
  NOTE: DERIVATIONS ARE NOT EXPECTED.

Unit-II: Chebychev's Inequality and Weak Law of Large Numbers **(5 h)**

- Chebychev's theorem: If $g(X)$ is a non-negative function of a r.v. X, $E\{g(X)\} < \infty$ and if $k > 0$ then $P\{g(X) \geq k\} \leq E\{g(X)\}/k^2$.
- Chebychev's inequality for discrete & continuous distribution in the forms
  $P\{|X - \mu| \geq k\sigma\} \leq 1/k^2$ and
  $P\{|X - \mu| < k\sigma\} \geq 1 - 1/k^2$, where $\mu = E(X)$ and $\sigma^2 = V(X)$.
- Concept of convergence in probability.
- Statement and proof of WLLN based on Chebychev's theorem.
- Examples and problems.

Unit-III: Central Limit Theorem **(4 h)**

- Statement and proof of the central limit theorem for i.i.d.r.v.s. based on mgf.
- Examples and problems.

Unit-IV: Hyper geometric distribution **(4 h)**

- Probability mass function

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}},\ x = 0,1,2,\ldots,n;\ n \le M$$

Notation X~H(N, M, n)

- Applications of hyper geometric distributions
- Binomial approximation to hyper geometric distribution
- Conditional distribution of X given (X+Y), where X and Y are independent binomial random variables with parameters $(n_1, p)$and (n_2, p) respectively.
- Raw moments, factorial moments, mean and variance
- Examples and problems

Unit-V: Negative binomial distribution **(8 h)**

- Probability mass function

$$P(X = x) = \binom{x+k-1}{x} p^k q^x; \qquad x = 0,1,2,\ldots \qquad 0 < p < 1,\ q = 1-p.$$

Notation X~ NB(k, p), $k \ge 1$.

- Probability Generating function (pgf), mgf, cgf, fmgf, first four moments and cumulants, factorial moments, recurrence relation for probabilities.
- Additive property.
- NB distribution as a waiting time distribution.
- NB(k, p) as the distribution of sum of k i.i.d. geometric r.v.s. with common parameter p.
- NB distribution obtained from Poisson distribution with gamma distributed parameter.
- Poisson approximation to NB distribution.
- Examples and problems.

Unit-VI: Truncated Distributions **(8 h)**

- Truncated distribution as conditional distribution, truncation to the right, left and on both sides.
- Binomial distribution B(n p) left truncated at X = 0 (value zero not observable), its p.m.f, mean, variance.
- Poisson distribution P(λ) left truncated at X = 0 (value zero not observable), its p.m.f, mean, variance.
- Normal distribution N(μ, σ^2) truncated
  (i) to the left of X = a
  (ii) to the right of X = b

(iii) to the left of $X = a$ and to the right of $X = b$, its p.d.f and mean.

- Examples and problems.

Unit-VII: Order Statistics **(8 h)**

- Order statistics for a random sample from a continuous distribution.
- Distribution of the i^{th} order statistics $X_{(i)}$ (distribution function and probability density function).
- Joint distribution of $(X_{(i)}, X_{(j)})$.
- Distribution of the smallest order statistics $X_{(1)}$, distribution of largest order statistics $X_{(n)}$
- Distribution of the sample median, distribution of the sample range $X_{(n)} - X_{(1)}$. Distribution of $X_{(1)}$ and $X_{(n)}$ for uniform and exponential distributions.
- Examples and problems.

References:

- Gupta S.C. and Kapoor V. K. (2017). Fundamentals of Mathematical Statistics. S. Chand and Sons. New Delhi.
- Rohatgi V. K. (1976). An Introduction to Probability theory and Mathematical Statistics. John Wiley and Sons, New York.
- Hogg. R. V., M. McKean J. W. and Craig. A. J. (2019). Introduction to Mathematical Statistics. Pearson Education, Inc.
- Weatherburn C. E. (1968). A first course in Mathematical Statistics. Cambridge University Press.
- Kulkarni M.B. and Ghatpande S.B. (2007). Introduction to Discrete Probability and Probability Distributions. SIPF Academy.
- Mood A. M. and Graybill F. A and Boes D. C. (2001). Introduction to the Theory of Statistics, third edition. Mc Graw Hill Education.
- Dudewicz E.J. and Mishra S.N. (1988). Modern Mathematical Statistics, (Wiley Sons).
- Biswas S. and Sriwastav G. L. (2011). Mathematical Statistics; Narosa Pub.

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester V
Discipline Specific Core (DSC) Course
ST-352: Statistical Inference-I

Total Hours: 45 **Credits: 3**

Course objectives:

- To acquaint the students with point estimation.
- To introduce properties of estimators such as unbiasedness, relative efficiency, sufficiency, consistency.
- To explain Cramer-Rao inequality and methods of estimation.

Course outcomes:

Students will be able to

- Understand problem of estimation of parameters.
- Distinguish between estimator and estimate.
- Test whether estimator is unbiased or not, attains the Cramer-Rao lower bound or not.
- Find efficiency of estimator relative to another estimator.
- Test whether statistic is sufficient or not for unknown parameter.
- Find estimator of unknown parameter using maximum likelihood estimation and method of moments.

Unit-I: Point Estimation **(12 h)**

- Concept of random sample from a distribution, Notion of a Parameter, Parameter space, general problem of estimation. Types of estimation: Point estimation and interval estimation.
- Point estimation: Definition of estimator, distinction between estimator and estimate, illustrative examples.
- Unbiasedness: Definition of unbiased estimator, biased estimator, positive and negative biases. Illustrative examples (These should include unbiased and biased estimators for the same parameters)
- Proofs of the results regarding unbiased estimator:
  (a) Two distinct unbiased estimators of $f(\theta)$ give rise to infinitely many unbiased estimators of $f(\theta)$.
  (b) If T is an unbiased estimator of θ, then f (T) is an unbiased estimator $f(\theta)$, provided $f(T)$ is linear function of T.
- Discussion of the following results:-
  (a) If T is an unbiased estimator of θ, then $f(T)$ need not be an unbiased estimator of $f(\theta)$, illustrative examples.
  (b) Sample standard deviation is a biased estimator of population standard deviation.
- Examples and Problems.

Unit-II: Relative Efficiency **(3 h)**

- Relative efficiency of unbiased estimator T_1 with respect to another unbiased estimator T_2, use of mean square error to define relative efficiency of biased estimators.
- Notion of uniformly minimum variance unbiased estimator (UMVUE), uniqueness of UMVUE whenever it exists.
- Examples and Problems.

Unit-III: Sufficiency **(6 h)**

- Concept and definition of sufficiency
- Statement of Neyman's factorization theorem (proof for discrete case only).
- Proofs of the following properties of sufficient statistics:
  a) If T is sufficient for θ, f (T) is also sufficient for $f(\theta)$ provided f is one to one and onto function.
  b) If T is sufficient for θ then T also sufficient for f (θ).
- Definition of likelihood as a function of the parameter for a random sample from (i) discrete, (ii) continuous distribution. Definition of Fisher's information function. Amount of information regarding parameter contained in a statistic T and a sufficient statistic T.
- Examples and problems.

Unit-IV: Asymptotic Behavior of an Estimator **(4 h)**

- Consistency: Definition of consistent estimator, proof of the following theorems:
  (a) Biased 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 then f(T) is also consistent estimator of $f(\theta)$ provided f is continuous function of T. (Invariance property of consistent estimator).
- Examples and problems.

Unit-V: Cramer-Rao Inequality **(6 h)**

- Statement and proof of Cramer-Rao inequality.
- Definition of minimum variance bound unbiased estimator (MVBUE) of f (θ)
  Proofs of the following results:
  i) If MVBUE exists for θ, then MVBUE exists for $\phi(\theta)$, provided ϕ is linear function.
  ii) If T is MVBUE for θ then T is sufficient for θ.
- Examples and problems.

Unit-VI: Methods of Estimation **(14 h)**

- Method of maximum likelihood, derivation of maximum likelihood estimators (m.l.e.) for parameters of only standard distributions: binomial, normal. Invariance property of m.l.e., relation between m.l.e. and sufficient statistics.
  (a) m.l.e. of uniform distribution over (i) (a, b), (ii) $(-\theta, \theta)$.
  (b) m.l.e. of θ in $f(x,\theta) = \exp[-(x-\theta)]\ x \geq \theta$.

- Method of moments: Derivation of moment estimators for standard distributions: binomial, Poisson, normal, exponential and uniform, illustration of situations where m.l.e. and moment estimators are distinct and their comparison using mean square error.
- Examples and problems.

References:

- Gupta S.C. and Kapoor V. K. (2017). Fundamentals of Mathematical Statistics. S. Chand and Sons, New Delhi.
- Rohatgi V. K. (1976). An Introduction to Probability theory and Mathematical Statistics. John Wiley and Sons, New York.
- Hogg. R. V., M. McKean J. W. and Craig. A. J. (2019). Introduction to Mathematical Statistics. Pearson Education, Inc.
- Mood A. M. and Graybill F. A and Boes D. C. (2001). Introduction to the Theory of Statistics, third edition, Mc Graw Hill Education.
- Kale B. K. and Muraridharan. (2015). Parametric Inference: An Introduction, Alpha Science Intl Ltd.
- Dudewicz E.J. and Mishra S.N. (1988). Modern Mathematical Statistics, (Wiley Sons).
- Biswas S. and Sriwastav G. L. (2011). Mathematical Statistics; Narosa Pub.

Methods of Teaching:

- **Interactive lecture method, Problem solving method and ICT enabled teaching method.**

T.Y. B.Sc. (Statistics): Semester V
Discipline Specific Core (DSC) Course
ST-353: Design of Experiments

Total Hours: 45 **Credits: 3**

Course objectives:

- To introduce concept of design of experiments.
- To make student aware about standard designs of experiments such as CRD, RBD and LSD.
- To introduce efficiency of design, missing plot technique, BIBD and factorial experiments.

Course outcomes:

Students will be able to

1. Plan the experiment, obtain relevant information from it.
2. Understand basic principles of Design of Experiments.
3. Study Standard designs: CRD, RBD, LSD etc.
4. Identify real life situations where the above designs are useful.
5. Compare the various designs using the concept of efficiency.

Unit-I: Introduction to Design of Experiments **(4 h)**

- Concept of Design of Experiment (DOE), Introduction to basic terms of Design of Experiments, Experimental unit, treatments, layout of an experiment, factor, level, run of experiment, control experiment, test experiment.
- Basic principles of Design of Experiments, Randomization, Replication and Local control.
- Uniformity trials.
- Choice of size and shape of a plot.
- The empirical formula for the variance per unit area of plots.
- Examples and problems

Unit-II: Standard Designs of Experiments **(15 h)**

- Completely Randomized Design (CRD).
  Definition and model, Preparation of Analysis of Variance (ANOVA) table, testing of equality of treatment effects, testing equality of two specified treatment means, critical differences. Merits and demerits of CRD.
- Randomized Block Design (RBD).
  Definition and model, Preparation of ANOVA table, testing of equality of treatment effects and block effects, testing for equality of two specific treatment means, critical differences. Merits and demerits of RBD.
- Latin Square Design (LSD) : definition, model:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk} \quad \begin{matrix} i = 1,2,\ldots,m; \\ j = 1,2,\ldots,m; \\ k = 1,2,\ldots,m \end{matrix} \qquad (i,j,k) \in S$$

Assumptions and interpretation, Estimation of parameters, Expected value of Mean sum of squares, components of variance.

Hypothesis for the model:

$H_{01}: \alpha_1 = \alpha_2 = \ldots = \alpha_m$

$H_{02}: \beta_1 = \beta_2 = \ldots = \beta_m$

$H_{03}: \gamma_1 = \gamma_2 = \ldots = \gamma_m$

and its interpretation. Justification of use of F-test for H_{01}, H_{02} and H_{03}, (independence of Chi-squares 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 treatments effects, use of critical difference, testing for equality of two row effects, two column effects and treatment effects. Merits and demerits of LSD.

- 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.)
- Identification of real life situations where the above designs are useful.
- Applications of principles of Design of Experiments in CRD, RBD and LSD.
- Simple algebraic and numerical problems.

Unit-III: Efficiency of a Design **(4 h)**

- Concept and definition of efficiency of a design.
- Comparison of efficiencies between CRD and RBD.
- Comparison of efficiencies between LSD and RBD, LSD and CRD.
- Simple numerical problems.

Unit-IV: Missing Plot Technique **(6 h)**

- Situations where missing plot technique is applicable.
- Estimation of missing plots by minimizing error sum of squares in RBD and LSD with one or two observations are missing.
- Derivation of exact treatments sum of squares, preparing analysis of variance table and writing report.
- Iterative procedure in case of missing observations.
- t-test for comparing any two treatment effects.

Unit-V: Balanced Incomplete Block Design (B.I.B.D.) **(9 h)**

- Definition and simple relations between parameters:
  (i) $bk = rv$
  (ii) $\lambda(v-1) = r(k-1)$
  (iii) $b \geq v$
  (iv) $b \geq v + r - k$
  (v) $r \geq k$
- Model, estimation of parameters (derivation are not expected)
- Analysis of variance table (intra-block analysis only) for testing significance of treatment effects and block effects.
- Tests for comparing two treatment effects.

- Symmetric BIBD: definition, theorems on symmetric BIBD

Unit-VI: Factorial Experiments **(7 h)**

- General description of factorial experiments, 2^2, 2^3 factorial experiments arranged in RBD.
- Definitions of main effects and interactions in 2^2, 2^3 factorial experiments.
- Preparation of ANOVA table by Yates procedure, test for main effects and interactions, estimation of main effects and interaction effects.
- General idea of confounding in factorial experiments, total confounding, analysis of variance table, testing main effects and interactions (confounding only one interaction).
- Partial confounding (confounding only one interaction per replicate), ANOVA table, testing of main effects and interactions.
- Construction of layouts in total confounding and partial confounding for 2^2, 2^3 factorial experiments.

References:

- Federer W.T. (1963). Experimental Designs, Oxford & IDH Publishing Co., New Delhi.
- Cochren W.G. & Cox G,M. (1992). Experimental Designs, Second Edition, John Wiley & Sons Inc., New Delhi.
- Montgomery D.C. (2001). Design & Analysis of Experiments, John Wiley & Sons Inc., New Delhi.
- Das M. N. and Giri N.C. (1986). Design & Analysis of Experiments, Second edition, Wiley Eastern Ltd., New Delhi.
- Snedecor G.W. and Cochran W.G. (1989). Statistical Methods, 8th edition, Affiliated East West Press, New Delhi.
- Goon A. M., Gupta M. K. and Dasgupta B. (1986). Fundamentals of Statistics, Vol-II, The World Press Pvt. Ltd., Calcutta.
- Gupta S.C. and Kapoor V.K. (2007). Fundamentals of Applied Statistics, S. Chand and Sons, New Delhi.
- Parimal Mukhopadhyay (2005). Applied Statistics, Books and Allied (P) Ltd, Kolkata.

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester V
Discipline Specific Core (DSC) Course
ST-354: Sampling Theory

Total Hours: 45 **Credits: 3**

Course objectives:

- To introduce in detail sampling methods such as SRSWR, SRSWOR, stratified sampling, systematic sampling.
- To introduce the problem of estimation of the population mean and population total.
- To discuss ratio and regression methods of estimation.
- To make student aware about applications of sampling methods in real life situations.

Course outcomes:

Students will be able to

- Obtain simple random sample with replacement and without replacement.
- Identify unbiased estimator of population mean and population mean square.
- Derive variance of unbiased estimator in case of SRSWR and SRSWOR.
- Determine sample size to conduct sample survey.
- Compare various sampling methods.
- Construct stratified random sample and systematic random sample in real life situations where these sampling are appropriate.
- Apply ratio method of estimation and regression method of estimation

Unit-I: Sample Survey **(6 h)**

- Concept of distinguishable elementary units, sampling units, sampling frame.
- Objective of a sample survey.
- Designing questionnaire, characteristics of good questionnaire.
- Planning, execution and analysis of a sample survey. Practical problems in planning, execution and analysis of a sample survey.
- Sampling and non-sampling errors with illustrations.
- Study of some surveys illustrating the above ideas.

Unit-II: Simple Random Sampling **(12 h)**

- Simple random sampling with and without replacement: Definition, inclusion probabilities. Definitions of population mean, population total and population variance.
- Proof of the following results:
  1) Sample mean as an unbiased estimator of population mean.
  2) $N\bar{y}$ is an unbiased estimator of population total.
  3) Sample mean square is an unbiased estimator of population mean square for SRSWOR
  4) $Var(\bar{y}) = \frac{N-n}{Nn} S^2$ and $SE(\bar{y})$ in case of SRSWOR.

5) $\text{Var}(\bar{y}) = \frac{N-1}{Nn} S^2$ and $SE(\bar{y})$ in case of SRSWR

- Simple random sampling without replacement for proportions.
- Proof of the following results:
  1) Sample proportion is an unbiased estimator of Population proportion
  2) Np is an unbiased estimator of NP.
  3) $\text{Var}(p) = \frac{(N-n)}{N-1} \frac{P(1-P)}{n}$ and SE(p)
- Examples and problems.

Unit-III: Determination of Sample Size (In case of SRS) **(3 h)**

- Determination of sample size for estimating population mean and population proportion when
  1) Margin of error and confidence coefficient is given.
  2) Coefficient of variation and confidence coefficient is given.
- Examples and problems.

Unit-IV: Stratified Random Sampling **(12 h)**

- Introduction.
- Real life situations.
- Stratified random sampling as a sample drawn from individual stratum using SRSWOR in each stratum.
- Construction of strata
- Proof of the following results.
  a) $\bar{y}_{st}$ as an unbiased estimator of population mean $\bar{Y}$.
  b) $N\bar{y}_{st}$ as an unbiased estimator of population total.
- Standard error of $\bar{y}_{st}$ & $N\bar{y}_{st}$ and their estimation.
- 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.
- Gain in precision due to stratification, comparison amongst SRSWOR, stratification with proportional allocation & stratification with Neyman's allocation.
- Cost & 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.
- Examples and problems.

Unit-V: Systematic Sampling **(6 h)**

- Real life situation where systematic sampling is appropriate. Technique of drawing a sample using systematic sampling.
- Estimation of population mean and population total, standard errors of these estimators.
- Distinguishing between stratification and systematic sampling, between SRSWOR and systematic sampling through real life situations.

- Examples and problems.

Unit-VI: Ratio and Regression Methods of Estimation for SRSWOR **(6 h)**

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

References:

- Cochran W.G. (2007). Sampling Techniques, Third Edition, Wiley
- Sukhatme P.V. (1953). Sampling Theory of Surveys with Applications, Indian Society of Agricultural Statistics, New Delhi.
- Murty M.N. (1977). Sampling Methods, ISI, Kolkata.
- Daroga, Singh & Chaudhary F.S. (2013). Theory & Analysis of Sample Survey Designs, New Age International.
- Mukhopadhay P. (2002). Theory and Method of Sample Survey, (Chapman and Hall)
- Gupta S.C. and Kapoor V.K. (2007). Fundamentals of Applied Statistics, S. Chand and Sons, New Delhi.

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester V
Discipline Specific Core (DSC) Course
ST-355: Introduction to Regression Analysis

Total Hours: 45 **Credits: 3**

Course objectives:

- To discuss in detail simple linear regression model.
- To introduce in detail multiple linear regression model and logistic regression model.
- To make student aware about applications of these models in real life situations.

Course outcomes:

Students will be able to

- Investigate the relationship between a variable of interest (the response) and the set of related predictor variables.
- Formulate and fit the appropriate regression model to the given dataset.
- Statistical data analysis using regression in various real-life situations.
- Test the significance of regression parameters.
- Understand the concept of binary response variable, Logit transform, estimation of parameter, interpretation of parameters.
- Compare AIC and BIC criteria for model selection in regression analysis.

Unit-I: Simple Linear Regression Model **(14 h)**

- 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.
- Properties of estimators of β_0 and β_1
- Estimation of σ^2
- Assumption of normality of ε. Tests of hypothesis of β_1
- Interval estimation in simple linear regression model
- Coefficient of determination
- Residual analysis : Standardized residuals, Studentized residuals, residual plots
- Detection and treatment of outliers
- Interpretation of four plots produced by lm command in R

Unit-II: Multiple linear Regression Model **(16 h)**

- 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.
- Estimation of σ^2
- Assumption of normality of ε. Tests of hypothesis of regression parameters

- Interval estimation in simple linear regression model
- Variable selection and model building
- Residual diagnostics and corrective measures such as transformation of response variable, weighted least squares method
- Polynomial regression models

Unit-III: Logistic Regression Model **(15 h)**

- Binary response variable, Logit transform, estimation of parameters, interpretation of parameters.
- Tests of hypotheses of model parameters, model deviance, LR test.
- AIC and BIC criteria for model selection
- Interpretation of output produced by glm command in R
- Multiple logistic regression

References:

- Draper, N. R. and Smith, H. (1998). Applied Regression Analysis, John Wiley, 3rd Ed.
- Hosmer, D. W. and Lemeshow, S. (1989). Applied Logistic Regression, Wiley.
- Montgomery, D. C., Peck, E. A. and Vining, G. G. (2003). Introduction to Linear
- Regression Analysis, Wiley.
- Neter, J., W., Kutner, M. H.; Nachtsheim, C.J. and Wasserman, W.(1996). Applied
- Linear Statistical Models, fourth edition, Irwin USA.

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester V
Discipline Specific Core (DSC) Course
ST-356 (A): Statistical Quality Control

Credits: 3

Total Hours: 45

Course objectives:

- To make student aware about Indian standards and International standards.
- To introduce statistical process control.
- To make student aware about capability study and sampling plans.

Course outcomes:

Students will be able to

- Construct, read and interpret control charts for variables and attributes.
- Judge process capability.
- Design and use sampling inspection plan.
- Apply tools such as histogram, scatter diagram, cause and effect diagram etc.
- Understand role of statistical methods in ISO.
- Estimate percentage of defective product in a production process.
- Compute Producer's risk, Consumer's risk, Average Outgoing Quality Limits etc.

Unit-I: Indian Standards and International Standards **(4 h)**

- Introduction to IS series and ISO 9000: 2015 Series with reference to process Control and statistical techniques (History, Organization Structure and different Clauses), role of statistical methods.

Unit-II: Statistical Process Control (SPC) **(20 h)**

- Introduction
- Seven Process Control (PC) Tools of SPC
  (i) Check sheet (ii) Cause and Effect Diagram (CFD) (iii) Pareto Diagram (iv) Histogram (v) Control chart. (vi) Scatter diagram (vii) Designs of Experiment (DOE).
- Control Charts: 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. A point outside the control limits.
  ii. Non random pattern of variation of the following type.
  (a) Seven or more points above or below central line.
  (b) Presence of cycle or linear trends.
  Note: Mathematical justification for (ii) is not expected.
  Use of control charts when (i) standards are given (ii) standards are not given.
- Control charts for continuous variables
  Decisions preparatory to control charts:
  (i) Choice of the variable

(ii) Basis of subgroups.
(iii)Size of subgroups.
(iv) Frequency of subgroups (Periodicity)

- R Chart and $\overline{X}$ chart: Purpose of R and $\overline{X}$ chart. 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 $\overline{X}$ chart when the process standard deviation (σ) is not given. Control limits based on σ, drawing of control chart. Plotting sample means, revision of control limits of $\overline{X}$ chart, if necessary.
- Control charts for Attributes:
  Decision preparatory to control charts:
  (i) size of subgroups:
  (ii) frequency of subgroups (Periodicity).
- p-chart (for fixed and variable sample size), np-chart, c-chart and u-chart (for fixed sample size)
- X chart, MR chart.
- CUSUM chart (tabular method).
- Examples and problems.

Unit-III: Capability Studies **(8 h)**

- Specification Limits, natural tolerance limits and their comparisons.
- Decisions based on these comparisons, estimate of percent defective.
- Catching the shift on average, evaluation of probability of catching shift of the first sample or on the subsequent samples after the shift (when process standard deviation is fixed).
- Shift in the process fraction defective, Evaluation of probability (using normal approximation only) of catching the shift on the first sample or on the subsequent samples after the shift.
- Process Capability Indices C_p, C_{pk} .
- Examples and problems.

Unit-IV: Acceptance Sampling for Attributes **(13 h)**

- Introduction. Concept of sampling inspection plan, comparison between 100% inspection and sampling inspection. Rectification of single and double sampling plans.
- Explanation of the terms: Producer's risk, Consumer's risk, Acceptance Quality Level (AQL), LTFD, Average Outgoing Quality (AOQ), AOQL, Average Sample Number (ASN), Average Total inspection (ATI), Operating characteristic (OC) curve , AOQ curve, ATI curve.
- Single Sampling Plan:
  Evaluation of probability of acceptance using
  (i) Hypergeometric (ii) Binomial (iii) Poisson and (iv) Normal distributions.
  Derivation of AOQ and ATI. Graphical determination of AOQL, Determination of a single sampling plans by lot quality and average quality approaches (numerical problems

are not expected). Description of Dodge and Romig tables (numerical problems are not expected)

- Double Sampling Plan:
  Evaluation of probability acceptance using Poisson approximation, derivation of ASN and ATI (With complete inspection of second sample). Derivation of the approximate formula of AOQ. Description of Dodge Romig Tables.
- Comparisons of single sampling plan and double sampling plan.
- Example and problems.

References:

- Duncan A.J. (1986). Quality Control & Industrial Statistics, fifth edition, Irwin.
- Grant E.L. & Richard Leavenworth. (2005). Statistical Quality Control, McGraw Hill Education India.
- Montgomery D. C. (2009). Statistical Quality Control, John Wiley & Sons Inc.
- Western Electric. (1982). Statistical Quality Control Handbook, Second Edition.
- ISO 9001:2015 Standards, 2015.

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester V
Discipline Specific Core (DSC) Course
ST-356 (B): Introduction to Stochastic Processes

Total Hours: 45 **Credits: 3**

Course objectives:

- To make student aware about introductory stochastic processes.
- To introduce applications of stochastic processes.

Course outcomes:

Students will be able to

- Understand the definition of stochastic process.
- Classify stochastic process and give examples of each type process.
- Describe Markov chain and its transition matrix.
- Identify and apply appropriate stochastic technique for a given applied problem.

Unit-I: Notion of Stochastic Processes **(10 h)**

- Probability Distributions: Generating functions, Bivariate probability generating function.
- Stochastic Process: Introduction, Stationary Process. Examples of Stochastic Processes.

Unit-II: Discrete Time Markov Chains **(20 h)**

- Markov Chains: Definition of Markov Chain, transition probability matrix, order of Markov chain, Markov chain as graphs, higher transition probabilities. Generalization of independent Bernoulli trials, classification of states and chains, stability of Markov system, graph theoretic approach.

Unit-III: Continuous Time Markov Chains **(15 h)**

- Counting Process: Definition, examples and properties
- Poisson Process: Postulates of Poisson process, properties of Poisson process, inter-arrival time, pure birth process, Yule Furry process, birth and death process, pure death process.

References:

- Medhi, J. (2009). Stochastic Processes, New Age International Publishers.
- Basu, A.K. (2005). Introduction to Stochastic Processes, Narosa Publishing.
- Bhat, B.R.(2000). Stochastic Models: Analysis and Applications, New Age
- International Publishers.
- Feller, William (1968): Introduction to probability Theory and Its Applications, Vol I,
- 3rd Edition, Wiley International.
- Ross S. M. (2012). Stochastic Processes, Wiley India Pub.

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester V
Skill Enhancement Course (SEC)
ST-350: Introduction to Python

Total Hours: 30 **Credits: 2**

Course objectives:

- To make student aware about Python programming.
- To illustrate inbuilt and user defined functions in Python.
- To explain statistical problem solving using Python.

Course outcomes:

Students will be able to

- Create and run Python programs using required tools.
- Understand and explain the results of given Python programs.
- Use inbuilt data structures and modules in Python.
- Solve statistical problems using Python.

Unit-I: Introduction to Python **(3 h)**

- Python Shell
- Variables
- Arithmetic Operators
- Assignment Operators
- Comparison Operators
- Logical Operators
- Identity Operators
- Membership Operators
- Expressions

Unit-II: Data Types in Python **(4 h)**

- Integer (int)
- Float (float)
- Boolean (bool)
- String (str)
- Tuple (tuple)
- List (list)
- Set (set)
- Dictionary (dict)

Unit-III: Programming in Python **(5 h)**

- if conditional
- if...else conditional
- if...elif...else conditional
- while loop
- for loop

- User defined functions def

Unit-IV: Inbuilt Functions in Python **(6 h)**

- Iteration: range(), len(), enumerate()
- Find attributes: dir()
- Value: abs(), round()
- Check truthiness: all(), any()
- Evaluate: eval()
- User input: input(),
- Operate on collection: min(), max(), sum()
- Open files: open()

Unit-V: Scientific Libraries in Python **(6 h)**

- importing libraries,
- Data manipulation: pandas
- Numerical Analysis: numpy
- Scientific methods: scipy
- Plotting graphs: matplotlib

Unit-VI: Statistical Problem solving using Python **(6 h)**

- Measures of central Tendency
- Measures of dispersion
- Correlation and Regression

References:

- Al Sweigart (2015). Automate the Boring Stuff with Python, 2nd Edition, No Starch Press
- Eric Matthes (2019). Python Crash Course, No Starch Press
- Thomas Haslwanter (2016). An Introduction to Statistics with Python, Springer
- Jake VanderPlas (2016). Python Data Science Handbook, O'Reilly Media, Inc

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester V
Discipline Specific Core (DSC) Course
ST-357: Statistics Practical-I

Total Hours: 60 **Credits: 2**

Course objectives:

- To develop skill of the numerical calculations.
- To develop abilities to adopt proper statistical methods of analysis for a given data.
- To introduce procedure of identification of suitable distribution for given dada.

Course outcomes:

Students will be able to

- Fit negative binomial and truncated distributions.
- Model sample from negative binomial distribution.
- Estimate parameters of standard distributions.
- Apply hyper geometric, negative binomial and truncated distributions.

List of Practicals:

Sr. No.	Topic Particular	Hours
1	Fitting of Negative binomial Distribution	4
2	Fitting of Truncated binomial Distribution (Parameter should be estimate by using Newton Raphson Method)	8
3	Fitting of Truncated Poisson Distribution(Parameter should be estimate by using Newton Raphson Method)	8
4	Model sampling from Negative binomial Distribution	4
5	Estimation of parameters of standard probability distributions by the method of Maximum Likelihood Estimation.	8
6	Estimation of parameters of standard probability distributions by method of moments	8
7	Applications of hypergeometric and negative binomial Distribution	8
8	Applications of Truncated Poisson and Truncated binomial Distribution	8
9	Applications of Truncated Normal Distribution	4

General instructions:

- All Practicals of this paper are to be carried out by using R/MS-Excel/MINITAB software. Student must complete all the practicals to the satisfaction of concerned teacher.
- Students must be encouraged to collect live data from real life situations for practicals.

References:

- Gupta S.C. and Kapoor V. K. (2017). Fundamentals of Mathematical Statistics. S. Chand and Sons, New Delhi.
- Rohatgi V. K. (1976). An Introduction to Probability theory and Mathematical Statistics. John Wiley and Sons, New York.
- Hogg. R. V., M. McKean J. W. and Craig. A. J. (2019). Introduction to Mathematical Statistics. Pearson Education, Inc.
- Weatherburn C. E. (1968). A first course in Mathematical Statistics. Cambridge University Press.
- Kulkarni M.B. and Ghatpande S.B. (2007). Introduction to Discrete Probability and Probability Distributions. SIPF Academy.
- Mood A. M. and Graybill F. A and Boes D. C. (2001). Introduction to the Theory of Statistics, third edition. Mc Graw Hill Education.
- Dudewicz E.J. and Mishra S.N. (1988). Modern Mathematical Statistics, (Wiley Sons).
- Biswas S. and Sriwastav G. L. (2011). Mathematical Statistics; Narosa Pub.
- Purohit S.G., Gore S.D. and Deshmukh S.R. (2008). Statistics Using R. Narosa Pub.

Methods of Teaching:

- Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester V
Discipline Specific Core (DSC) Course
ST-358: Statistics Practical-II

Total Hours: 60 **Credits: 2**

Course objectives:

- To develop skill of the numerical calculations.
- To develop abilities to adopt proper DOE.
- To introduce analysis of designs using R software.

Course outcomes:

Students will be able to

- Analyze standard designs CRD, RBD and LSD using R.
- Estimate missing observations and then analyze design.
- Compare designs and analyze BIBD and factorial experiments.
- Analyze factorial experiment with total and partial confounding.

List of Practicals:

Sr. No.	Topic Particular	Hours
1	Analysis of CRD.	4
2	Analysis of RBD.	4
3	Analysis of LSD.	8
4	Efficiency of Designs	4
5	Missing Plot Technique in RBD	8
6	Missing Plot Technique in LSD.	8
7	Analysis of BIBD.	8
8	Analysis of 2^3 factorial experiment arranged in RBD.	8
9	Analysis of 2^3 factorial experiment with total & partial confounding.	8

General instructions:

- All Practicals of this paper are to be carried out by using R software.
- Student must complete all the practicals to the satisfaction of concerned teacher.

References:

- Federer W.T. (1963). Experimental Designs, Oxford & IDH Publishing Co., New Delhi.
- Cochren W.G. & Cox G,M. (1992). Experimental Designs, Second Edition, John Wiley & Sons Inc., New Delhi.
- Montgomery D.C. (2001). Design & Analysis of Experiments, John Wiley & Sons Inc., New Delhi.
- Das M. N. and Giri N.C. (1986). Design & Analysis of Experiments, Second edition, Wiley Eastern Ltd., New Delhi.

- Snedecor G.W. and Cochran W.G. (1989). Statistical Methods, 8th edition, Affiliated East West Press, New Delhi.
- Goon A. M., Gupta M. K. and Dasgupta B. (1986). Fundamentals of Statistics, Vol-II, The World Press Pvt. Ltd., Calcutta.
- Gupta S.C. and Kapoor V.K. (2007). Fundamentals of Applied Statistics, S. Chand and Sons, New Delhi.
- Parimal Mukhopadhyay (2005). Applied Statistics, Books and Allied (P) Ltd, Kolkata.
- Purohit S.G., Gore S.D. and Deshmukh S.R. (2008). Statistics Using R. Narosa Pub.

Methods of Teaching:

- Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester V
Discipline Specific Core (DSC) Course
ST-359: Statistics Practical-III

Total Hours: 60 **Credits: 2**

Course objectives:

- To develop skill of the numerical calculations.
- To develop abilities to adopt proper sampling techniques.
- To make student aware about real life problem solving regarding SQC, Regression and Stochastic Processes.

Course outcomes:

Students will be able to

- Obtain random samples using SRSWR, SRSWOR, Stratified and Systematic sampling.
- Estimate population parameters.
- Construct and interpret regression equations.
  Or Calculate transition probability matrix
- Construct control charts and OC curves, AOQ curves etc.
  Or Identify types of classes of Markov Chains etc.

List of Practicals:

Sr. No.	Topic Particular	Hours
1	Simple Random Sampling.	8
2	Stratified Random Sampling.	8
3	Systematic sampling..	4
4	Ratio and Regression methods of estimation	4
5	Simple regression analysis and dignostics by graphical method	4
6	Multiple regression analysis and dignostics by graphical method	8
7	Logistic regression	4
8	X & MR charts, CUSUM charts.	8
9	Single Sampling Plans (with OC, AOQ, AOQL, ATI, ASN curves)	4
10	Double Sampling Plans (with OC, AOQ, AOQL, ATI, ASN curves)	8
11	Calculation of transition probability matrix	8
12	Identification of characteristics of reducible and irreducible chains.	4
13	Identification of types of classes of Markov chain	4
14	Identification of ergodic transition probability matrix	4

(Note: Consider practical No. 8, 9 and 10 for ST-356(A) and 11, 12, 13 and 14 for ST-356(B))

General instructions:

- All Practicals of this paper are to be carried out by using R/Python/MINITAB software.
  Student must complete all the practicals to the satisfaction of concerned teacher.
- Students must be encouraged to collect live data from real life situations for practicals.

References:

- Mukhopadhay P. (2002) Theory and Method of Sample Survey, (Chapman and Hall)
- Gupta S.C. and Kapoor V K. (2007). Fundamentals of Applied Statistics, S. Chand and Sons, New Delhi.
- Montgomery D. C. (2009). Statistical Quality Control, John Wiley & Sons Inc.
- Medhi, J. (2009). Stochastic Processes, New Age International Publishers.
- Ross S. M. (2012) Stochastic Processes, Wiley India Pvt.
- Purohit S.G , Gore S D. and Deshmukh S.R. (2008). Statistics Using R. Narosa Pub.
- Thomas Haslwanter (2016) An Introduction to Statistics with Python, Springer

Methods of Teaching:

- Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester VI
Discipline Specific Core (DSC) Course
ST-361: Distribution Theory-II

Total Hours: 45 **Credits: 3**

Course objectives:

- To make student aware about uniform, lognormal, Weibull, Laplace and Cauchy distributions.
- To introduce multinomial distribution and bivariate normal distribution.
- To discuss derivations of distributions of functions of random variables.
- To acquaint students with the applications of probability distributions.

Course outcomes:

Students will be able to

- Use continuous uniform distribution in real life situations.
- Derive probability distributions of functions of uniform random variables.
- Understand relation between normal distribution and lognormal distribution.
- Apply lognormal and Weibull distribution in real life situations.
- Identify situations where Cauchy and Laplace distribution is applicable.
- Use multinomial and bivariate normal distribution in real life situations.

Unit-I: Rectangular (Uniform) distribution **(8 h)**

- P.d.f.

$$f(x) = \frac{1}{b-a} \qquad -\infty < a < x < b < \infty$$

$$= 0 \qquad \text{Otherwise}$$

- Distribution function, mean, variance, mgf, r^{th} raw moment.
- Standard form: U(0,1).
- U(0,1) as the distribution of F(X), where X is a continuous type r.v. with d.f. F(.), application to model sampling, Use of U(0,1) to generate integer valued random numbers.
- Distributions of X+Y, X-Y, XY, X/Y for X and Y are independent U(0,1) random variables.
- Real life situations.
- Examples and problems.

Unit-II: Log-normal distribution **(5 h)**

- P.d.f.

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

$$= 0 \qquad \text{otherwise.}$$

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

- Nature of the probability curve.
- Moments (r^{th} moment about x=a), first four moments, β_1 and γ_1 coefficients, quartiles.
- Relation with $N(\mu, \sigma^2)$ distribution.
- Examples and problems.

Unit-III: Weibull distribution **(5 h)**

- P.d.f.:

$$f(x) = \frac{\beta}{\alpha}\left(\frac{x-\gamma}{\alpha}\right)^{\beta-1} \exp\left\{-\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right\};\ \gamma \le x < \infty,\ -\infty < \gamma < \infty,\ \alpha, \beta > 0$$

$$= 0 \qquad otherwise.$$

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

- Distribution function, quartiles.
- r^{th} Moment about $x = \gamma$, mean and variance.
- Relation with exponential distribution.
- Examples and problems.

Unit-IV: Laplace Distribution (Double Exponential Distribution) **(6 h)**

- P.d.f :

$$f(x) = \frac{\lambda}{2}\exp\left[-\lambda|x-\mu|\right] \quad -\infty < x < \infty,\ -\infty < \mu < \infty,\ \lambda > 0$$

$$= 0 \qquad otherwise.$$

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

- Nature of probability curve.
- Distribution function, quartiles.
- mgf, cgf, moments and cumulants, $\beta_1, \beta_2, \gamma_1, \gamma_2$.
- Laplace distribution as the distribution of the difference of two i.i.d exponential variates with mean θ.
- Examples and problems.

Unit-V: Cauchy Distribution **(6 h)**

- P.d.f.:

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

$$= 0 \qquad otherwise.$$

Notation : $X \sim C(\mu, \lambda)$

- Nature of probability curve.
- Distribution function, quartiles, non-existence of moments.
- Additive property for two independent Cauchy variates (Statement only),
- Statement of distribution of the sample mean.
- Relationship with uniform and Student's 't' distribution.
- Examples and problems.

Unit-VI: Multinomial Distribution **(8 h)**

- Joint p.m.f.

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!\, p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}}{x_1!\, x_2! \ldots x_k!}; \; x_i = 0, 1, 2, \ldots n; \; i = 1, 2, \ldots k$$

$$x_1 + x_2 + \ldots + x_k = n; \; p_1 + p_2 + \ldots + p_k = 1; \; 0 < p_i < 1$$

$$= 0, \quad \text{otherwise}$$

Notation $(X_1, X_2, \ldots, X_k)' \sim MD(n, p_1, p_2, \ldots, p_k)$

- Joint mgf of $X_1, X_2 \ldots X_k$
- Use of joint mgf to obtain means, variances, covariances, total correlation coefficients, multiple and partial correlation coefficients for k = 3, univariate marginal distributions.
- Variance covariance matrix, Rank of Variance-Covariance matrix and its interpretation.
- Real life situations.
- Examples and problems.

Unit-VII: Bivariate Normal Distribution (BND) **(7 h)**

- P.d.f:

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

$$-\infty < x, y, \mu_1, \mu_2 < \infty \quad \sigma_1, \sigma_2 > 0 \; , -1 < \rho < 1$$

$= 0$ *otherwise.*

Notation : $(X,Y) \sim BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

- Marginal and conditional distributions, identification of parameters, regression of Y on X and X on Y, independence and uncorrelatedness, mgf and moments
- Cauchy distribution as the distribution of Z = X/Y where $(X, Y) \sim BN(0, 0, \sigma_1^2, \sigma_2^2, \rho)$
- Example and problems.

References:

- Gupta S.C. and Kapoor V. K. (2017). Fundamentals of Mathematical Statistics. S. Chand and Sons, New Delhi.
- Rohatgi V. K. (1976). An Introduction to Probability theory and Mathematical Statistics. John Wiley and Sons, New York.
- Hogg. R. V., M. McKean J. W. and Craig. A. J. (2019). Introduction to Mathematical Statistics. Pearson Education, Inc.
- Weatherburn C. E. (1968). A first course in Mathematical Statistics. Cambridge University Press.
- Mood A. M. and Graybill F. A and Boes D. C. (2001). Introduction to the Theory of Statistics, third edition. Mc Graw Hill Education.
- Dudewicz E.J. and Mishra S.N. (1988). Modern Mathematical Statistics, (Wiley Sons).
- Biswas S. and Sriwastav G. L. (2011). Mathematical Statistics; Narosa Pub.

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester VI
Discipline Specific Core (DSC) Course
ST-362: Statistical Inference-II

Total Hours: 45 **Credits: 3**

Course objectives:

- To make student aware about Interval estimation.
- To introduce basic concepts of testing of hypothesis.
- To acquaint students with the applications of parametric and nonparametric tests.
- To introduce sequential tests.

Course outcomes:

Students will be able to

1. Construct simple statistical hypothesis and composite statistical hypothesis.
2. Determine probability of the error of first kind and second kind.
3. Test simple hypothesis against the alternative hypothesis.
4. Write critical regions and identify the one that has maximum power among all critical regions.
5. Apply nonparametric and sequential tests where these tests are applicable.
6. Find confidence interval for unknown parameter.

Unit-I: Parametric Tests **(15 h)**

- Statistical hypothesis, problem of testing of hypothesis, Definition and Illustrations of (i) simple hypothesis, (ii) composite hypothesis, (iii) Two types of errors in testing of hypothesis (iv) sizes of two types of errors. Problem of controlling the sizes of two types of errors.
  Definition and illustrations of (i) level of significance (ii) observed level of significance (p value) (iii) power function of the test (iv) size of test (v) power of test.
- Definition of Most Powerful (MP) and Uniformly Most Powerful (UMP) tests of size α. Neyman-Pearson's (NP) lemma with proof for the construction of MP test, construction of UMP test for one sided alternative.
- Examples and problems.

Unit-II: Interval Estimation **(8 h)**

- Notion of interval estimation, definition of confidence interval, confidence bounds.
- Relation between confidence interval and testing of hypothesis, definition of pivotal quantity and its use in obtaining confidence interval and bounds.
- Interval estimation for the following cases:
  - Mean (μ) of normal distribution (when σ known and σ unknown)
  - Variance (σ^2) of normal distribution (when μ known and μ unknown)
  - Difference of two means $\mu_1 - \mu_2$ (a) for a sample from bivariate normal population (b) for samples from two independent normal populations.
  - Mean of exponential distribution.

- Population proportion (P) and difference of two population proportions $(P_1 - P_2)$ in case of two independent large samples.
- Population quantiles using order statistics.

- Examples and problems.

Unit-III: Non Parametric Tests **(14 h)**

- Meaning of ordinal and nominal data.
  Meaning of non-parametric problems, Distinction between parametric and Non-parametric methods. Concept of distribution free statistic, advantages and disadvantages of non parametric methods.
  Procedure of:
  - sign test for one and two samples problem
  - Wilcoxon signed rank T-test for paired observations.
  - Mann-Whitney U-test for two independent samples.
  - Run test for randomness of given observations.
  - Kolmogorov-Smirnov test for completely specified univariate distribution (only one sample problem)
- Examples and problems.

Unit-IV: Sequential Tests **(8 h)**

- Sequential test procedure for simple null hypothesis against simple alternative hypothesis and its comparison with fixed sample size test procedure. Definition of Wald's sequential probability ratio test (SPRT) of strength (α, β).
  Illustration for standard distributions: Bernoulli, Poisson, normal and exponential distribution, graphical and tabular procedures for carrying out the test.
- Examples and problems.

References:

- Gupta S.C. and Kapoor V. K. (2017). Fundamentals of Mathematical Statistics. S. Chand and Sons, New Delhi.
- Rohatgi V. K. (1976). An Introduction to Probability theory and Mathematical Statistics. John Wiley and Sons, New York.
- Hogg. R. V., M. McKean J. W. and Craig. A. J. (2019). Introduction to Mathematical Statistics. Pearson Education, Inc.
- Mood A. M. and Graybill F. A and Boes D. C. (2001). Introduction to the Theory of Statistics, third edition. Mc Graw Hill Education.
- Kale B. K. and Muraridharan. (2015). Parametric Inference: An Introduction, Alpha Science Intl Ltd.
- Dudewicz E.J. and Mishra S.N. (1988). Modern Mathematical Statistics, (Wiley Sons).
- Biswas S. and Sriwastav G. L. (2011). Mathematical Statistics; Narosa Pub.

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester VI
Discipline Specific Core (DSC) Course
ST-363: C Programming

Total Hours: 45 **Credits: 3**

Course objectives:

- To make student aware about C programming.
- To illustrate inbuilt and user defined functions in C.
- To explain statistical problem solving using C.

Course outcomes:

Students will be able to

- Create algorithm and flowchart to solve simple problems using C.
- Understand procedural language, middle level language, higher level language, general language structure.
- Design and implement programs that use loops.
- Solve statistical computational problems using C program.

Unit-I: Introduction to C **(10 h)**

- Algorithms and flowcharts.
- Introduction to procedural language, middle level language, higher level language, general language structure, character set, keywords, identifiers.
- Data types: Numeric and character data types, Numeric and character constants, string constants, symbolic constants.
- Operators: Numeric, logical, arithmetic, unary, relational, equality, decrement, increment, conditional assignments, precedence of operator expressions and their evaluation.
- Data input/output, numeric and character data, printf ( ), scanf ( ), getchar ( ), putchar ( ), gets ( ), puts ( ).
- Formatted output

Unit-II: Control Structures **(6 h)**

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

Unit-III: Arrays **(12 h)**

- 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.
  - To find mean, median, variance and coefficient of variation of frequency distribution.
  - To find correlation coefficient and least square regression line of Y on X for a given bivariate data.
  - To arrange the given data in increasing/decreasing order of magnitude.

- To obtain median of given n observations.
- To obtain addition of two matrices, multiplication of two matrices.

Unit-IV: Functions **(7 h)**

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

Unit-V: Pointers **(10 h)**

- Basic concept and relation to one dimensional array.
- Following programs must be covered under above sections.

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 find sum of digits of a number.
11. To solve simultaneous linear equations.(two equations in two variables)
12. To evaluate simple and compound interest
13. To convert decimal number to equivalent binary number.
14. To generate Fibonacci series like 0, 1, 1,2,3,5...
15. To sort a string using string function.
16. To combine given two strings using string function.

References:

- Gottfried, B.S. (1996). Programming with C (Schaum Outline series), McGraw Hill co.,
- London
- Kanitkar, Y (2008). Let us C, BFB publishers, New Delhi.
- Karnighan, B. W. and Ritchi, M. (1988). The C programming language, Second edition,
- Prentice Hall.
- Rajaraman V. (2007). Computer programming in C, Prentice Hall of India.,

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester VI
Discipline Specific Core (DSC) Course
ST-364: Elements of Clinical Trials

Total Hours: 45 **Credits: 3**

Course objectives:

- To introduce terminology used in Clinical trials and several common designs used for clinical trials.
- To explain techniques of clinical trials and analysis and interpretation of the experiments.
- To discuss treatment comparison.

Course outcomes:

Students will be able to

- Reduce the bias and variability involved during conduction of clinical trials.
- Estimate the true therapeutic effect of the drug.
- Analyse the outcomes of experiments using statistical methods.

Unit-I: Introduction to clinical trials **(8 h)**

- Definition, need and ethics of clinical trials,
- Drug Development Process,
- FDA and its guidelines,
- Protocol, objectives and end points of clinical trials,
- Possible bias and random errors in clinical studies,
- Conduct of clinical trials,
- Overview of phase I-IV trials.

Unit-II: Types of Clinical trials **(8 h)**

- Single site vs Multi site studies,
- Placebo/Active Control,
- Dose response and historical controls,
- Superiority and non-inferiority trials,
- Combination and Equivalence trials
- Vaccine Trials

Unit-III: Designs for Clinical Trials **(9 h)**

- Washout and Run-in period,
- Cluster Randomized design,
- Parallel and Cross-Over Designs,
- Factorial Designs.
- Balance Incomplete Block Design
- Titration Designs

Unit-IV: Randomization and Blinding **(10 h)**

- Needs and Benefits of Randomization and Blinding
- Fixed Allocation Randomization: Simple Randomization, Permuted Blocked Randomization, Stratified Randomization.
- Adaptive Randomization Procedures: Covariate, Treatment and Response Adaptive Randomization.
- Blinding: No, single, double and triple blinding.

Unit-V: Analysis Methods **(10 h)**

- Sample size determination and power calculations
- Treatment comparison:
  Parametric: Z, t, Proportions (One Sample-Two samples) tests, ANOVA
  Non-parametric tests: Run, Sign, Wilcoxon signed rank, Mann-Whitney U-test, Kolmogorov-Smirnov test
- Categorical data analysis: Chi-square

References:

- Chow, Shein-Chung, and Jen-pei Liu. (2008). Design and analysis of clinical trials: concepts and methodologies, John Wiley & Sons.
- Friedman, Lawrence M., et al. (2015) Fundamentals of clinical trials, springer
- Chen, Ding-Geng Din, Karl E. Peace, and Pinggao Zhang. (2017). Clinical trial data analysis using R and SAS, CRC Press.

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester VI
Discipline Specific Core (DSC) Course
ST-365: Optimization Techniques

Total Hours: 45 **Credits: 3**

Course objectives:

- To introduce linear programming problems.
- To explain LPP problems solving using graphical and simplex method.
- To introduce transportation and assignment problems with its solving methods.
- To acquaint students with PERT, networking analysis and Simulations

Course outcomes:

Students will be able to

- Understand the basics of Linear Programming Problems.
- Solve LPP by appropriate method such as graphical, simplex etc.
- Write dual of primal problem.
- Solve LPP using its dual.
- Solve assignment and transportation problem.
- Construct network and analyse using CPM and PERT.
- Apply Mont Carlo simulation method.

Unit-I: Linear Programming Problems **(12 h)**

- Statement of the linear programming problems. Simple examples and formulation of problems.
- Definitions of i) A Slack variable ii) Surplus variable iii) Unrestricted variable iv) Decision variable.
- Definition of i) a solution ii) feasible solution iii) a basic feasible solution (b.f.s. degenerate and non-degenerate solution) iv) Optimal solution v) basic and non basic variables vi) objective function vii) non- negativity conditions.
- Solutions of L.P.P. by i) graphical method: Solution space unique and non-unique solutions. Obtaining an optimum solution ii) Simplex method: initial b.f.s. is readily available, obtaining the initial basic feasible solution. Criterion for deciding whether obtained solution is optimal, method of improving a solution.
- Initial b.f.s. is not readily available, introduction to artificial variable. Big M. method (or penalty method) modified objective function. Modification and applications of simplex method L.P.P. with artificial variable.
- Examples and problems.

Unit-II: Theory of Duality **(4 h)**

- Writing a dual of primal problem.
- Solution of L.P.P. by using its dual
- Conversion of primal to dual and dual to primal
- Examples and problems.

Unit-III: Transportation Problem **(9 h)**

- Definition of i) a feasible solution, ii) a basic feasible solution and iii) optimal solution.
- Statement of transportation problem, balanced and unbalanced transportation problem.
- Methods of obtaining initial basic feasible solution:
  1) North west corner method.
  2) Method of matrix minima (least cost method)
  3) Vogel's Approximation Method (VAM).
- Optimal solution of transportation problem using *uv*-method (MODI), uniqueness and non uniqueness of optimal solution. Degeneracy and method of resolving degeneracy.
- Variants in transportation problem: No allocation in a particular cell, maximization problem.

Unit-IV: Assignment Problem **(4 h)**

- Assignment problem: Statement of assignment problem, relation to transportation problem and solution of assignment problem using Hungarian method.
- Special cases in the assignment problem: Unbalanced assignment problem, maximization problem, restrictions on assignments and alternate optimal solution.
- Examples and problems.

Unit-V: C.P.M. And Networking Analysis **(12 h)**

- Definition i) Event or node ii) Activity iii) critical activity iv) Project function v) Predecessor and successor activity vi) Predecessor and successor event vii) properties of network viii) numbering by Fulkerson's rule.
- Critical path method, constructions of a network
- Definition i) Earliest start time ii) Earliest finish time iii) latest start time iv) Latest finish time v) Critical path
- Float, Total float, Independent float & Free float, their significance.
- PERT : Definition of PERT, i) Pessimistic time ii) Optimistic time iii) Most likely time iv) Forward Pass Calculation v) Backward Pass calculation vi) Slack vii) Critical Path viii) Probability of meeting scheduled date.
- Calculation of expected time, S.D. of project duration.
- Distinguish between PERT and C.P.M.
- Examples and problems.

Unit-VI: Simulation **(4 h)**

- Introduction to simulation, merits, demerits, limitations.
- Pseudo random number generates: Linear congruential , mid square method.
- Model sample from normal distribution (using Box- Muller transformation), uniform distribution, exponential distribution.
- 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.

References:

- Taha H.A. (2007). Operations Research an Introduction, 8th Ed, Pearson Prentice Hall.
- Gupta P.K. , Hira D.S. and Kamboj A. (2012). Introduction to Operations Research, S. Chand.
- Shrinath L.S. PERT and CPM Principles & Applications, 3rd Edition, Affiliated East West Press Pvt. Ltd.
- Kapoor V.K. (2011). Operations Research, S. Chand & Sons, New Delhi.
- Sharma S.D. (2002). Operations Research, Kedarnath Ramnath & Co., Meerut.
- Kanti Swarup, Gupta P. K. And Man Mohan. (2020). Operations Research, Sultan Chand and Sons.

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester VI
Discipline Specific Core (DSC) Course
ST-366 (A): Official Statistics and Applied Statistics

Total Hours: 45 **Credits: 3**

Course objectives:

- To describe construction of life tables.
- To explain Indian Official Statistics.
- To introduce statistics in education and psychology.
- To acquaint students with elements of annuities and demand and supply analysis.

Course outcomes:

Students will be able to

- Understand about the agencies responsible for collection of data on official statistics and their important publications.
- Understand and construct life tables.
- Study statistics in psychology and education.
- Understand the Demand and Supply analysis.

Unit-I: Indian Official Statistics **(9 h)**

- Present official statistical system in India. Methods of collection of official statistics, their reliability and limitations.
- Role of Ministry of Statistics and Program Implementation (MoSPI), National Statistical Office (NSO), Central Statistical Office (CSO), National Sample Survey Office (NSSO), and National Statistical Commission.
- Government of India's Principal publications containing data on the topics such as population, industry and finance.

Unit-II: Life Tables **(8 h)**

- Introduction and meaning.
- Construction. Functions and their interrelations Complete life table
- Expectation of life.
- Numerical examples and problems.

Unit-III: Statistics in Psychology and Education **(10 h)**

- Introduction
- Scaling individual test items
- Scaling of scores on a test
  - Z scores and Z scaling
  - Standard scores, normalised scores, T scores, percentile scores
- Reliability of test scores
  - Methods of determining test reliability

    i)Test retest method

ii)Parallel forms method
iii)Split half method
iv)Kuder-Richardson method

- Validity of test scores
- Comparison of test scores
- Advantages and disadvantages of test scores
- I.Q. and E.Q.
- Examples and problems.

Unit-IV: Annuities **(8 h)**

- Accumulation and amount functions
- The effective rate of interest and discount
- Nominal rate of interest and discount
- Annuities: annuity certain, annuity due, annuity immediate, perpetuity
- Examples and problems

Unit-V: Demand and Supply Analysis **(10 h)**

- Demand: meaning, statement of law, assumptions, exceptions and determinants of demand, individual and market demand..
- Supply: meaning, statement of law, assumptions, exceptions and determinants of supply, individual and market supply
- Elasticity of demand: definition : i) price elasticity of demand ii) income elasticity of demand iii) cross elasticity of demand
- Method of measuring elasticity of demand: i) percentage method, ii) point method iii) total outlay method iv) ARC Method
- Demand forecasting: meaning need and methods of forecasting
- Examples and problems

References:

- Gupta S.C. and Kapoor V.K. (2007). Fundamentals of Applied Statistics, S. Chand and Sons, New Delhi.
- Goon, Gupta, Dasgupta. (1986). Fundamentals of Statistics, Vol-II, The World Press Pvt. Ltd., Calcutta.
- Parimal Mukhopadhyay. (2005). Applied Statistics, Books and Allied (P) Ltd, Kolkata.
- Guide to current Indian Official Statistics, Central Statistical Office, GOI, New Delhi. (Refer website-http://mospi.nic.in/)

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester VI
Discipline Specific Core (DSC) Course
ST-366 (B): Actuarial Statistics

Credits: 3

Total Hours: 45

Course objectives:

- To make student aware about Indian Insurance business.
- To describe the concept of risk and risk measurement models.
- To explain elementary concepts of financial mathematics.
- To acquaint students with statistical theory used in life insurance contracts.

Course outcomes:

Students will be able to

1. Identify basic risk available in the problem and formulate loss random variable.
2. Summarize different terms of life tables and their applications.
3. Apply tools of financial mathematics.
4. Simulate data related with actuarial statistics.

Unit-I: Insurance Business **(3 h)**

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

Unit-II: Feasibility of Insurance Business **(3 h)**

- Measurement of adverse financial impact, expected value principle.
- Concept of utility function
- Feasibility of insurance business.
- Illustrative examples.

Unit-III: Survival Distribution and Life Tables **(12 h)**

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

Unit-IV: Models for Life Insurance **(12 h)**

- Theory of compound interest, effective rate of interest, discount factor.

- Insurance payable at the end of the year of death, present value random variable, actuarial present value.
- Derivation of actuarial present value for n-year term life insurance, whole life insurance and endowment insurance.

Unit-V: Annuities **(8 h)**

- Annuities – certain, annuity due, annuity immediate.
- 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.

Unit-VI: Benefit Premiums **(7 h)**

- 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.
- Variance of loss random variable

References:

- Bowers N.L. Jr., H. S. Gerber, J. C. Hickan, D. A. Jones, C. J. Nesbitt (1997). Actuarial Mathematics, Society of Actuaries, U.S.
- Deshmukh, S. R. (2009). Actuarial Statistics, Universities Press, Hyderabad, India.
- Deshmukh, S.R. (2005). Actuarial Statistics: In Introduction Using R. Narosa Publishing House, New Delhi.
- Promislow, S. D. (2014). Fundamentals of Actuarial Mathematics. 3rd Edition. Wiley.

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester VI
Skill Enhancement Course (SEC)
ST-360: Data Analysis using SPSS

Total Hours: 30 **Credits: 2**

Course objectives:

- To make student aware about introduction to SPSS.
- To explain computations of descriptive statistics using SPSS.
- To describe statistical testing to solve real life problems using SPSS.

Course outcomes:

Students will be able to

1. Apply SPSS software to compute descriptive statistics.
2. Understand statistical hypothesis testing procedure using SPSS software.
3. Write SPSS commands to analyze data.

Unit-I: An Overview of SPSS **(6 h)**

- Introduction to SPSS
- Opening a Data File in SPSS
- SPSS Data Editor
  - Variable View
  - Data View
  - Entering Data into the Data Editor
  - Saving the Data File
  - Statistical Analysis
- SPSS: general description, functions, menus, commands

Unit-II: Managing Data and Results **(4 h)**

- Editing and Manipulating Data
- Missing Values
- Editing SPSS output
- Copying SPSS output
- Printing from SPSS
- Closing SPSS
- Importing Data

Unit-III: Descriptive Analysis with SPSS **(4 h)**

- Descriptive Statistics:
  - Frequencies
  - Measure of central tendency
  - Measure of Dispersion
  - Skewness
  - Kurtosis
- Charts and Graphs:

Bar charts
Pie chart

- Scatter Plots
- Line Graphs and Histogram
- Examples and problems

Unit-IV: Correlation and Regression with SPSS **(4 h)**

- Correlation simple and multiple
- Rank correlation
- Simple linear regression
- Multiple regression
- Examples and problems

Unit-V: Statistical Tests with SPSS **(6 h)**

- Parametric Tests
- Students t-test
  One sample t-test
  Two sample t-Test
  Paired t-test
- ANOVA
  One - Way ANOVA
  Two - Way ANOVA
- Chi-Square Test
  Chi - square test for goodness of fit
  Chi - square test for independence of attributes
- Examples and problems

Unit-VI: SPSS with command syntax **(6 h)**

- An overview of syntax and its grammar
- Working with variables, constants and strings
- SPSS basic operators, arithmetic expressions and logical expressions
- Syntax window and performing computations with syntax
- Controlling flow and executing conditionals.
- Computing random variables from binomial, Poisson, uniform and normal distributions
- Examples and problems

References:

- A. Rajathi and P.Chandran (2006). SPSS for you, MJP Publishers
- K. Pandya, Smruti Bulsari, Sanjay Sinha (2011). SPSS in Simple Steps, Dreamtech Press
- Robert H. Carver and Jane Gradwohl Nash (2011). Doing Data Analysis with SPSS, Cengage Learning.
- Andy Field (2009). Discovering Statistics Using SPSS, Third edition, SAGE

Methods of Teaching:

- Interactive lecture method, Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester VI
Discipline Specific Core (DSC) Course
ST-367: Statistics Practical-IV

Total Hours: 60 **Credits: 2**

Course objectives:

- To develop skill of the numerical calculations.
- To develop abilities to apply proper statistical tests for testing of hypothesis.
- To introduce confidence interval estimation practical problem solving using software and model sampling from bivariate distribution.

Course outcomes:

Students will be able to

- Fit lognormal distribution.
- Obtain model samples from Cauchy and Bivariate normal distributions.
- Solve parametric and nonparametric testing problems
- Construct confidence interval and SPRT.

List of Practicals:

Sr. No.	Topic Particular	Hours
1	Fitting of Log-normal Distribution.	4
2	Model Sampling from Cauchy Distribution	4
3	Model Sampling from bivariate normal distribution.	4
4	Testing of hypothesis – I (Prob. Type I & II errors, MP test, Power of test)	8
5	Testing of hypothesis – II (UMP test for simple Vs. composite for Binomial, Poisson, Normal & Exponential Distributions)	8
6	Confidence Interval estimation	8
7	Non – parametric test – I (Sign test, Wilcoxon's Signed Rank test, Run test.)	8
8	Non – parametric test – II (Mann-Whitney test, Kolmogorov-Smirnov test.)	8
9	SPRT – I (Binomial & Poisson Distributions.)	4
10	SPRT – II (Normal & Exponential Distributions.)	4

General instructions:

- All Practicals of this paper are to be carried out by using R/SPSS/MINITAB software. Student must complete all the practicals to the satisfaction of concerned teacher.
- Students must be encouraged to collect live data from real life situations for practicals.

References:

- Gupta S.C. and Kapoor V. K. (2017). Fundamentals of Mathematical Statistics. S. Chand and Sons, New Delhi.
- Hogg. R. V., M. McKean J. W. and Craig. A. J. (2019). Introduction to Mathematical Statistics. Pearson Education, Inc.

- Mood A. M. and Graybill F. A and Boes D. C. (2001). Introduction to the Theory of Statistics, third edition. Mc Graw Hill Education.
- Kale B. K. and Muraridharan. (2015). Parametric Inference: An Introduction, Alpha Science Intl Ltd.
- Dudewicz E.J. and Mishra S.N. (1988). Modern Mathematical Statistics, (Wiley Sons).
- Biswas S. and Srivastav G. L. (2011). Mathematical Statistics; Narosa Pub.
- Siegel S. and Castellan N. J. (1988). Non Parametric Statistics for the Behavioral Sciences. McGraw Hill Singapore.
- Gibbons J.D. and Chakraborti S. (2020). Non Parametric Statistical Inference, Chapman and Hall/CRC.
- Sprent P. and Smeeton N. C. (2001). Applied Non Parametric Statistical Methods, Chapman & Hall/CRC.
- Purohit S.G., Gore S.D. and Deshmukh S.R. (2008). Statistics Using R. Narosa Pub.
- Andy Field (2009). Discovering Statistics Using SPSS, Third edition, SAGE

Methods of Teaching:

- Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester VI
Discipline Specific Core (DSC) Course
ST-368: Statistics Practical-V

Total Hours: 60 **Credits: 2**

Course objectives:

- To develop skill of the numerical calculations.
- To develop abilities to adopt proper statistical methods of analysis for a given data.
- To introduce practical problem solving using C.

Course outcomes:

Students will be able to

- Analyse the outcomes of experiments using statistical methods.
- Compute descriptive statistics using C.
- Fit binomial distribution and regression lines using C.

List of Practicals:

Sr. No.	Topic Particular	Hours
1	Randomization Methods.	4
2	Statistical Analysis for Parallel Designs	4
3	Z-test, t-test, two sample t-test, paired t-test	8
4	Sample size calculation and power of the test	4
5	Nonparametric Tests	8
6	Analysis of Categorical Data	8
7	Construction of frequency distribution with given class interval from raw data.	8
8	Computation of mean, variance, standard deviation and quartiles for given n observations and frequency distribution.	4
9	Fitting of a Binomial distribution to given data.	4
10	Computation of correlation coefficient for a given bivariate data.	4
11	Fitting a line of regression of Y on X for a given bivariate data.	4

General instructions:

- Practicals from number 1 to 6 of this paper are to be carried out by using R/SPSS/MINITAB software and practicals from number 7 to 11 are to be carried out using C Programming.
- Student must complete all the practicals to the satisfaction of concerned teacher.
- Students must be encouraged to collect live data from real life situations for practicals.

References:

- Chow, Shein-Chung, and Jen-pei Liu. (2008). Design and analysis of clinical trials: concepts and methodologies, John Wiley & Sons.
- Friedman, Lawrence M., et al. (2015) Fundamentals of clinical trials, springer

- Chen, Ding-Geng Din, Karl E. Peace, and Pinggao Zhang (2017). Clinical trial data analysis using R and SAS, CRC Press.
- Kanitkar, Y (2008). Let us C, BFB publishers, New Delhi
- Purohit S.G., Gore S.D. and Deshmukh S.R. (2008). Statistics Using R. Narosa Pub.
- Andy Field (2009). Discovering Statistics Using SPSS, Third edition, SAGE

Methods of Teaching:

- Problem solving method and ICT enabled teaching method.

T.Y. B.Sc. (Statistics): Semester VI
Discipline Specific Core (DSC) Course
ST-369: Statistics Practical-VI (Project)

Total Hours: 60 **Credits: 2**

Course objectives:

- To develop skill of the formulating the real life statistical problem and collecting the data.
- To develop abilities to apply proper statistical methods to solve formulated problem.
- To develop skill of presentation and report writing.

Course outcomes:

Students will be able to

1. Collect real life data sets from various fields.
2. Gain proficiency for real life data analysis.
3. Work in team and communicate data analysis results to non-statisticians.

Project duration: November to February.

Project Guide:

Teacher from the Department of Statistics and/or personnel from organization where student is going to visit for field work or training. Each project group will be guided by concerned teacher (guide)

Project Topic:

Students in consultation with the guide will decide Project Topic/Area. Project work may be carried out in a group of students (maximum five) depending upon the depth of fieldwork/problem involved. Project work should be based on real life data related to social/ industrial/medical/ banking etc fields. Students may also use primary, secondary or simulated data sets for their project work.

Presentation:

Internal assessment of this paper will be based on project work presentation (power point presentation) by student or group of students before students and teachers of the department including guide.

Project Report:

Project work data analysis is to be carried out by using R/SPSS/PYTHON/MINITAB software. Students are supposed to write the project report on project work and submit a copy of project report to department before term end practical examinations.

References:

- Chow, Shein-Chung, and Jen-pei Liu. (2008). Design and analysis of clinical trials: concepts and methodologies. John Wiley & Sons.

- Chen, Ding-Geng Din, Karl E. Peace, and Pinggao Zhang (2017). Clinical trial data analysis using R and SAS, CRC Press.
- Purohit S.G., Gore S.D. and Deshmukh S.R. (2008). Statistics Using R. Narosa Pub.
- Andy Field (2009). Discovering Statistics Using SPSS, Third edition, SAGE.
- Parimal Mukhopadhyay. (2005). Applied Statistics, Books and Allied (P) Ltd, Kolkata.
- Sprent P. and Smeeton N. C. (2001). Applied Non Parametric Statistical Methods, Chapman & Hall/CRC.
- Montgomery D. C. (2009). Statistical Quality Control, John Wiley & Sons Inc.
- Mukhopadhay P. (2002). Theory and Method of Sample Survey, (Chapman and Hall)
- Montgomery, D. C., Peck, E. A. and Vining, G. G. (2003). Introduction to Linear Regression Analysis, Wiley.
- Montgomery D.C. (2001). Design & Analysis of Experiments, John Wiley & Sons Inc., New Delhi.

Methods of Teaching:

- Problem solving method and ICT enabled teaching method.

Skills acquired and Job prospects for the BSc Statistics students:

B.Sc. Statistics programme emphasizes both theory and applications of statistics and is structured to provide knowledge and skills in depth necessary for the employability of students in industry, other organizations, as well as in academics.

After successful completion of three years degree course in Statistics, student will be well versed with relevant generic skills and global competencies such as

1. **Problem-solving skills** that are required to solve different types of Statistics related problems with well-defined solutions, and tackle open-ended problems that belong to the disciplinary-area boundaries;
2. **Investigative skills**, including skills of independent investigation of Statistics related issues and problems;
3. **Communication skills** involving the ability to listen carefully, to read texts and research papers analytically and to present complex information in a concise manner to different groups/audiences of technical or popular nature;
4. **Analytical skills** involving paying attention to detail and ability to construct logical arguments using correct technical language related to Statistics and ability to translate them with popular language when needed;
5. **ICT skills**;
6. **Personal skills** such as the ability to work both independently and in a group.

Job Opportunities:

- The student who has thoroughly studied this syllabus of T.Y.B.Sc.(Statistics) can join for higher education at PG level towards M.Sc.(Statistics)
- Students with B.Sc.(Statistics) degree are expected to served as Statisticians/ Administrators / Investigators in the private as well as government sections.
- Students with B.Sc.(Statistics) degree under this syllabus will find better opportunities of Statistician/Analyst in Manufacturing (SQC unit), Pharmaceutical Industries, Service Industries such as Banking and Insurance, Railway, Forest, Telecom, Transports, Hotel etc services.

Head of the
Statistics Department
Moolji Jaitha College, Jalgaon