Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous) (Affiliated to Savitribai Phule Pune University, Pune)

Choice Based Credit System (CBCS)

Framework for Syllabus

Bachelor of Science (B. Sc.) Statistics

Implemented from

Academic year 2022 -23

F.Y. B. Sc.(Statistics)

Semester - I

Course Type	Course Code	Course Title	Credits
DSCC A - 1	BSC-ST 101 T	Descriptive Statistics-I Using MS-Excel	02
DSCC A - 2	BSC-ST 102 T	Introduction to Probability	02
DSCC A -3 Practical	BSC-ST 103 P	Practical (Based on MS-Excel)-I	1.5

Semester - II

Course Type	Course Code	Course Title	Credits
DSCC A-4	BSC-ST 201 T	Descriptive Statistics-II Using R	02
DSCC A-5	BSC-ST 202 T	Discrete Probability Distributions	02
DSCC A-6 Practical	BSC-ST 203 P	Practical (Based on R) - II	1.5

S.Y. B. Sc. (Statistics)

Semester -III

Course Type	Course Code	Course Title	Credits
DSCC A-7	BSC-ST 301 T	Discrete Probability Distributions , Regression Analysis and Vital Statistics	02
DSCC A-8	BSC-ST 302 T	Continuous Probability Distributions	02
DSCC A-9 Practical	BSC-ST 303 P	Practical (Based on R) - III	02

Semester - IV

Course Type	Course Code	Course Title	Credits
DSCC A-10	BSC-ST 401 T	Sampling Distributions	02
DSCC A-11	BSC-ST 402 T	Large and Small Sample Statistical tests using R	02
DSCC A-12 Practical	BSC-ST 403 P	Practical (Based on R) - IV	02

T.Y. B. Sc. (Statistics)
Semester -V
Academic year 2022 -23

Course Type	Course Code	Course Title	Credits
DSCC A-1	BSC-ST 501 T	Probability Distributions -I	02
DSCC A-2	BSC-ST 502 T	Theory of Estimation	02
DSCC A-3	BSC-ST 503 T	Sampling Theory and Methods	02
DSCC A-4	BSC-ST 504 T	Statistical Process Control	02
DSCC A-5	BSC-ST 505 T	Regression Analysis	02
DSCC A-6	BSC-ST 506 T	Operations Research	02
DSCC A-7 Practical	BSC-ST 507 P	Practical-V (Based on BSC-ST-501 T and BSC-ST-505 T)	02
DSCC A-8 Practical	BSC-ST 508 P	Practical-VI (Based on BSC-ST-502 T and BSC-ST-503T)	02
DSCC A-9 Practical	BSC-ST 509 P	Practical-VII (Based on BSC-ST-504 and BSC-ST-506)	02
DSEC-A-10	BSC-ST 510 T	Python Programming	02
DSEC-A-11	BSC-ST 511 T	Actuarial Statistics	02

Course Type	Course Code	Course Title	Credits
DSCC A-12	BSC-ST 601 T	Probability Distributions -II	02
DSCC A-13	BSC-ST 602 T	Theory of Testing of Hypothesis	02
DSCC A-14	BSC-ST 603 T	Bio Statistics	02
DSCC A-15	BSC-ST 604 T	Reliability Analysis	02
DSCC A-16	BSC-ST 605 T	Design of Experiments Using MINITAB	02
DSCC A-17	BSC-ST 606 T	Stochastic Process and Time Series Analysis	02
DSCC A-18 Practical	BSC-ST 607 P	Practical-VIII (Based on BSC-ST-602 T and BSC-ST-604 T)	02
DSCC A-19 Practical	BSC-ST 608 P	Practical-IX (Based on BSC-ST-603 T and BSC-ST-605 T)	02
DSCC A-20 Practical	BSC-ST 609 P	Project	02
DSEC-A-21	BSC-ST 610 T	Data Visualization Using Tableau software	02
DSEC-A-22	BSC-ST 611 T	Basics of Data Science	02

Note: The End of Semester (EOS) Examination of following courses will be in online mode.

DSEC-A-10	BSC-ST 510 T	Python Programming	02
DSEC-A-21	BSC-ST 610 T	Data Visualization Using Tableau software	02

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

(Affiliated to Savitribai Phule Pune University, Pune)

Choice Based Credit System (CBCS)

Bachelor of Science (B. Sc.) Statistics

Syllabus of

T. Y. B. Sc. Statistics

Implemented from

Academic year 2022 -23

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Board of Studies in Statistics

Sr. No.	Name	Designation
1	Dr. A. A. Kulkarni	Chairman
2	Dr. A. J. Shivagaje	Member, Nominated by Academic Council
3	Dr. A. K. Khamborkar	Member, Nominated by Academic Council
4	Prof. S. V. Kawale	Member, Nominated by Hon. Vice Chancellor, SPPU, Pune
5	Dr. S. B. Pathare	Alumni Member, Nominated by Hon. Principal
6	Mr. Anirudha Deshmukh	Industry Expert Member, Nominated by Hon. Principal
7	Dr. V. P. Narkhede	Subject Expert, Special Nominated Member by Hon. Principal
8	Prof. S.A. Tarate	Member (Co-opted)
9	Dr. S. D. Jagtap	Member (Co-opted)
10	Dr. N.T. Shelke	Member (Co-opted)
11	Dr. B.P. Thakur	Member (Co-opted)

1. **Prologue/ Introduction of the programme: Atleast one page**

It is known that in economic activities are of three types, agriculture, industrial and service. In the same way the subject Statistics is a SERVICE SCIENCE having potential to address the problems in these three fields. In research application of Statistics is mandatory. In the present days, apart from traditional field of career, Data Science, Data Analytics, Data Mining, Data Visualization are the upcoming field of career for Statistics students. In these field student must have mathematical ability, statistical thinking, computer (Software and programming) knowledge and communication (Verbal and written). These points are taken into consideration to design the syllabus and examination pattern of Statistics. In addition to academics, the department takes care to arrange a series of lectures on interview skills, preparation of CV, improve communication skill and overall personality development. The students are given the task of event management so that they can practice the principles of management such as leadership, creativity, communication, time management, group activity, team work, etc. In general, through curricular, co-curricular and extra-curricular activities student in three years is developed as thought provoker, problem solver, technologically sound, with command on communication, strong self-confidence.

B. Sc. in Statistics program is of three years' duration, with semester pattern for all the three years. The important feature of the syllabus is that, all practical's form first year to third year will be conducted on computer using MS-EXCEL/ R Suit, Python programming and Tableau.

The feature is at the in all courses, end of every chapter, self-learning activities are listed. These self-learning activities will play important role in creating interest in the subject and also boost their confidence. Further group activities will give the chance to explore their creativity and ideas. In addition, the verbal and written communication will be improved. These self-learning activities are expected to motivate students to participate in various student related academic events organized by home college or by other colleges too.

The course on Tableau will give an opportunity to learn thousands of various data presentation types and to present the complex data by easy way. The practical examinations of all courses will be on computer.

In T.Y. B.Sc. examination of one theory course at each semester will be on computer. In short, maximum exposure is given to students to work on computer and evaluate them on computer.

The syllabus is framed with appropriate weightage of theory, applied and skill enhancement courses. After receiving B.Sc. degree, student is expected to have minimum knowledge of various courses and student will have ability to analyse the data with relevant interpretation of results.

2. Programme outcomes (PoS)

Students enrolled in the program complete a curriculum that exposes and trains students in a full range of essential skills and abilities. They will have the opportunity to master the following objectives.

a) Student will achieve the skill of understanding the data.
b) Student will be able to develop the data collection instrument.
c) Student will have skill to write a story using data visualization.
d) Student will understand the interdisciplinary approach to correlate the statistical concepts with concepts in other subjects.
e) Student will be made aware of history of Statistics and hence of its past, present and future role as part of our culture.
f) Students will demonstrate conceptual domain knowledge of the Statistics in an integrated manner.
g) Student will play the key role in management for effective functioning of MIS.

I. Programme Structure and Course Titles

Sr. No.	Class	Semester	Course Code	Course Title	Credits
1.	F.Y. B.Sc.	I	BSC-ST 101 T	Descriptive Statistics -I Using MS-Excel	02
2.	F.Y. B.Sc.	I	BSC-ST 102 T	Introduction to Probability	02
3.	F.Y. B.Sc.	I	BSC-ST103 P	Practical (Based on MS-Excel)-I	1.5
4.	F.Y. B.Sc.	II	BSC-ST 201 T	Descriptive Statistics-II Using R	02
5.	F.Y. B.Sc.	II	BSC-ST 202 T	Discrete Probability Distributions	02
6.	F.Y. B.Sc.	II	BSC-ST 203 P	Practical (Based on R) - II	1.5
7.	S.Y. B.Sc.	III	BSC-ST 301 T	Discrete Probability Distributions , Regression Analysis and Vital Statistics	02
8.	S.Y. B.Sc.	III	BSC-ST 302 T	Continuous Probability Distributions	02
9.	S.Y. B.Sc.	III	BSC-ST 303 P	Practical (Based on R) - III	02
10.	S.Y. B.Sc.	IV	BSC-ST 401 T	Sampling Distributions	02
11.	S.Y. B.Sc.	IV	BSC-ST 402 T	Large and Small Sample Statistical tests using R	02
12.	S.Y. B.Sc.	IV	BSC-ST 403 P	Practical (Based on R) - IV	02
13.	T.Y. B.Sc.	V	BSC-ST 501 T	Probability Distributions -I	02
14.	T.Y. B.Sc.	V	BSC-ST 502 T	Theory of Estimation	02
15.	T.Y. B.Sc.	V	BSC-ST 503 T	Sampling Theory and Methods	02
16.	T.Y. B.Sc.	V	BSC-ST 504 T	Statistical Process Control	02
17.	T.Y. B.Sc.	V	BSC-ST 505 T	Regression Analysis	02
18.	T.Y. B.Sc.	V	BSC-ST 506 T	Operations Research	02
19.	T.Y. B.Sc.	V	BSC-ST 507 P	Practical-V (Based on BSC-ST-501 T and BSC-ST-505 T)	02

20.	T.Y. B.Sc.	V	BSC-ST 508 P	Practical-VI (Based on BSC-ST-502 T and BSC-ST-503T)	02
21.	T.Y. B.Sc.	V	BSC-ST 509 P	Practical-VII (Based on BSC-ST-504 and BSC-ST-506)	02
22.	T.Y. B.Sc.	V	BSC-ST 510 T	Python Programming	02
23.	T.Y. B.Sc.	V	BSC-ST 511 T	Actuarial Statistics	02
24.	T.Y. B.Sc.	VI	BSC-ST 601 T	Probability Distributions -II	02
25.	T.Y. B.Sc.	VI	BSC-ST 602 T	Theory of Testing of Hypothesis	02
26.	T.Y. B.Sc.	VI	BSC-ST 603 T	Biostatistics	02
27.	T.Y. B.Sc.	VI	BSC-ST 604 T	Reliability Analysis	02
28.	T.Y. B.Sc.	VI	BSC-ST 605 T	Design of Experiments Using MINITAB	02
29.	T.Y. B.Sc.	VI	BSC-ST 606 T	Stochastic Process and Time Series Analysis	02
30.	T.Y. B.Sc.	VI	BSC-ST 607 P	Practical-VIII (Based on BSC-ST-602 T and BSC-ST-604 T)	02
31.	T.Y. B.Sc.	VI	BSC-ST 608 P	Practical-IX (Based on BSC-ST-603 T and BSC-ST-605 T)	02
32.	T.Y. B.Sc.	VI	BSC-ST 609 P	Project	02
33.	T.Y. B.Sc.	VI	BSC-ST 610 T	Data Visualization Using Tableau software	02
34.	T.Y. B.Sc.	VI	BSC-ST 611 T	Basics of Data Science	02

Ahmednagar Jilha Maratha Vidya Prasarak Samaj's

New Arts, Commerce and Science College, Ahmednagar (Autonomous)

Syllabus of T Y. B. Sc. Statistics under Faculty of Science

Semester – V	**Paper – I**
Course Code: BSC-ST 501 T	**Title of the Course: Probability Distributions –I**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. On completion of this course students will get the knowledge about distributions and their real life applications.
b. Students will learn the inter relationship among continuous distributions
c. Students will able to apply CLT to solve real life problems.
d. Students will get idea to simulate data from various continuous distributions.

Detailed Syllabus:

Unit - I		**Beta Distributions**	**09L**
	1.1.	Introduction of Gamma and Beta integral with their properties (Without proof) Beta distributions of first kind: p. d. f. $f(x) = \frac{1}{\beta(m,n)} x^{m-1}(1-x)^{n-1}$; $0 \leq x \leq 1, m, n > 0$ $= 0$; elsewhere Notation :$X \sim \beta_1(m, n)$, Nature of probability curve, skewness, mean, variance, properties, r^{th} raw moment, harmonic mean, median(for m=n) and mode for $\beta_1(m, n)$,.	
	1.2	Relation with U (0,1). The probability distributions of $1 - X$, Distribution function of $\beta_1(m, n)$, relationship between distribution function of beta variate first kind and Binomial distribution. Real life applications and Numerical examples.	
	1.3	Beta distributions of second kind p.d.f. $f(x) = \frac{1}{\beta(m,n)} \frac{x^{m+1}}{(1+x)^{m+n}}$; $x \geq 0, m, n > 0$ $= 0$; *elsewhere* Notation: $X \sim \beta_2(m, n)$, Nature of probability curve, skewness, mean, variance, r^{th} raw moment, harmonic mean, mode, median for $\beta_2(m, m)$. Distribution of $\frac{1}{X}$.	

		Interrelation between $\beta_1(m, n)$ and $\beta_2(m, n)$ variates. Proofs of the following: 1. If X follows G (α, λ_1) and Y follows G (α, λ_2), X and Y are independent then i. $X + Y$ and $\frac{X}{X+Y}$ are independent. ii. $X + Y$ follows G (α, $\lambda_1+\lambda_2$). iii. $\frac{X}{X+Y}$ follows $\beta_1(\lambda_1, \lambda_2)$. 2. If X follows G (α, λ_1) and Y follows G (α, λ_2), X and Y are independent then i. $X + Y$ and $\frac{X}{Y}$ are independent. ii. $X + Y$ follows G (α, $\lambda_1+\lambda_2$). iii. $\frac{X}{Y}$ follows $\beta_2(\lambda_1, \lambda_2)$. Real life applications and Numerical examples.	
Unit - II		**Weibull distribution**	**06L**
	2.1	History of the Weibull Distribution, p.d.f. of two parameter Weibull distribution $f(x) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}}$; $x \geq 0, \alpha, \beta > 0$ $= 0$, otherwise Notation: X~ W (Scale=α, Shape= β). Probability curve for different values of shape parameter, Exponential distribution as a particular case of Weibull distribution, mean and variance, coefficient of variation, Nature of the CV.	
	2.2	Distribution function, Survival Function, Nature of Hazard rate, quartiles, quartile deviation. Distribution of $\left(\frac{X}{\alpha}\right)^{\beta}$ Real life situations and applications of Weibull Distribution. Numerical examples.	
Unit-III		**Order Statistics.**	**06L**
	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 r[th] order statistic $X_{(r)}$, Distribution function and density function of $X_{(1)}$, Distribution function of $X_{(n)}$, Distribution function and density function of $X_{(r)}$, Distribution of $X_{(r)}$ from a random sample from uniform and exponential distributions.	
	3.2	Derivation of joint p.d.f. of $(X_{(r)}, X_{(s)})$, Probability distribution of sample range $X_{(n)} - X_{(1)}$. Probability Distribution of sample median. $Corr(X_{(r)}, X_{(s)})$, Numerical examples.	

Unit-IV		**Chebychev's Inequality, Central Limit Theorem and Weak Law of Large Numbers**	**09L**
	4.1	Chebychev's theorem: If g (x) is a non – negative function of r.v. X such that $E[g(X)] < \infty$ then, $P[g(X) \geq k] \leq \{\frac{E[g(X)]}{k}\}$ where k is positive real number.	
	4.2	Chebychev's inequality for discrete and continuous distributions (with proof) in the forms $P[\|X - \mu\| \geq k] \leq \frac{\sigma^2}{k^2}$, where $k > 1$, $P[\|X - \mu\| \geq k\sigma] \leq \frac{1}{k^2}, k > 1$. Where $\mu = E(X)$ and $\sigma^2 = Var(X) < \infty$. Applications of Chebychev's inequality in control charts, statistical inference. Numerical examples.	
	4.3	Sequence of r. v. s, convergence of sequence of r. v. in a) probability b) distribution, with simple illustrations. Statement and proof of the central limit theorem for i.i.d. r.v.s. (proof based on MGF), simulation-based demonstrations.	
	4.4	Weak law of large numbers (WLLN), Simulation-based demonstrations. Applications of CLT and WLLN. Numerical examples.	

Suggested Readings:

1. Hogg, R.V. and Craig A.T(1978). Introduction to Mathematical Stastistics, IVth Edition, Macmillan Publishing Company.Inc. New York.
2. Lindgren B.W.: (1976) Statistical Theory III rd Edition Collier Macmillan international Edition, Macmillan Publishing Co. Inc. New York.
3. Mood. A.M., Graybill , F.Bose ,D.C.: (1974) Introduction to theory of Statistics. III rd Edition Mc- Graw Hill Series.
4. Mukhopdhyay, P (1996). Mathematical Statistics, New Central Book Agency.
5. Rohatgi , V.K. IIIrd Edition An Introduction to probability Theory and Mathematical Statistics Wiley Eastern Ltd .New Delhi.
6. Casella G. and Berger Robert L. (2002) Statistical Inference. 2nd Edition, Duxbury Advanced series.
7. Dasgupta A. (2010) Fundamentals of Probability: A first course, , Springer, New York.

Semester – V	Paper – II
Course Code: BSC-ST 502 T	**Title of the Course: Theory of Estimation**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. Students will get in depth knowledge about different methods of estimation.
b. Students will understand the basic properties of estimator.
c. Students will get an idea about asymptotic behaviour of an estimator.

Detailed Syllabus:

Unit - I		**Preliminaries and Moment method of Estimation**	**06L**
	1.1.	Notion of a parameter, parameter space, General problem of estimating an unknown parameter by point and interval estimation (using Pivotal quantity approach), Point Estimation: Definition of an estimator, distinction between estimator and estimate, illustrative examples. Mean Square Error (MSE) of an estimator.	
	1.2	Method of moments: Procedure to obtain moment estimator, examples to find moment estimators for the parameters of standard distributions (Case of one or two parameters only), non-existence of moment estimator.	
Unit - II		**Methods of Estimation**	**09L**
	2.1	Likelihood Function, Definition of likelihood as a function of unknown parameter for a random sample comes from i) discrete ii) continuous distribution. Distinction between the likelihood function and Joint p.d.f. or p. m. f., Method of maximum likelihood (M.L.E) : Method of maximum likelihood estimation, examples of finding M.L.E of parameters of standard distributions (Case of two parameters only for Normal distribution), invariance property of M.L.E. and its application, Invariance property of M.L.E.	
	2.2	M.L.E. of the parameter of distribution M.L.E. of parameter(s) when support depend of parameter, parameters when θ in uniform distribution over i) $(0,\theta)$ ii) $(-\theta,\theta)$ iii) $(m\theta,n\theta)$ $(m<n)$, M.L.E. of exponential distribution with location parameter θ, M.L.E. of location parameter in Laplace distribution	
	2.3	M.L.E. using Newton-Raphson's Method for Parameter of 1) Truncated Binomial distribution truncated at 0 2) Truncated Binomial distribution truncated at 0	

		3) Cauchy distribution Illustrations of situations where M.L.E. and moment estimators are distinct and their comparison using MSE.	
Unit-III		**Properties of Estimator**	**09L**
	3.1	Unbiasedness: Definition of an unbiased estimator, biased estimator, positive and negative bias with its interpretation, illustrations and examples (these should include unbiased and biased estimators for the same parameters). Proofs of the following results: i. If there exist two unbiased estimator foe the parameter θ, then there exists infinitely many unbiased estimators for the parameter θ. ii. Invariance Property of Unbiasedness: If T is an unbiased estimator of θ, then $\boldsymbol{\phi}(\boldsymbol{T})$ is unbiased estimator of $\boldsymbol{\phi}(\boldsymbol{\theta})$ provided $\boldsymbol{\phi}(.)$ is a linear function. Notion of the Best Linear Unbiased and it's application.	
	3.2	Proof of the result: If $\mathrm{L}(\theta \mid \underline{x})$ be the likelihood function of sample of size n from the pdf $f(x, \theta)$ then i. $\mathrm{E}\left(\frac{\partial}{\partial\theta} log \mathrm{L}(\theta \mid \underline{x})\right) = 0$ ii. $\mathrm{E}\left(\frac{\partial^2}{\partial\theta^2} log \mathrm{L}(\theta \mid \underline{x})\right) = -\mathrm{E}\left(\frac{\partial}{\partial\theta} log \mathrm{L}(\theta \mid \underline{x})\right)^2$ Fisher information function: Amount of information contained in statistic $T = T(X_1, X_2, \ldots, X_n)$. Fisher Information contained in single observation $I_X(\theta)$, Relation between $I_X(\theta)$ *and* $I_{\underline{X}}(\theta)$. Cramer- Regularity conditions, Cramer – Rao Lower Bound (CRLB) with proof, definition of minimum variance bound unbiased estimator (MVBUE) of $\boldsymbol{\phi}(\boldsymbol{\theta})$. Proofs of following results: i. If MVBUE exists for θ then MVBUE exists for $\boldsymbol{\phi}$(θ) where $\boldsymbol{\phi}$(.) is a linear function. ii. If T is MVBUE for θ then T is sufficient for θ. Uniformly Minimum Variance Unbiased Estimator (UMVUE). Proof of the theorem: i. If UMVUE exist, then it is unique.	
	3.2	Sufficiency: Concept and definition of sufficiency, Examples, Need of Neyman Factorization theorem, Neyman factorization theorem with proof(for discrete case only). Pitmann–Koopman form and sufficient statistic; Exponential family of probability distributions and sufficient statistic. Proofs of the following properties: i. If T is sufficient for θ, then $\boldsymbol{\phi}(\boldsymbol{T})$ is also sufficient for θ provided ø is a one to one and onto function. ii. M.L.E. is a function of sufficient statistic.	

		iii. If $X1, X2, \ldots\ldots Xn$ be random sample from $f(x,\theta)$ which is a member of exponential family, then likelihood function is also a member of exponential family. iv. If $f(x,\theta)$ is a member of exponential family, then $\sum_{i=1}^{n} k(x_i)$ is sufficient statistic for θ.	
Unit-IV		**Efficiency and Asymptotic Behavior of an Estimator**	**06L**
	4.1	Efficiency: Concepts, definition, 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, Examples, Asymptotic Behavior of an Estimator.	
	4.2	Consistency as a property of sequence of estimators: Two Definitions. Consistency of an estimator T_n using Chebyshev's inequality, Examples. Proof of the following theorems: i. If T_n is unbiased for θ then it is consistent provided var(T_n) $\rightarrow 0$ as $n \rightarrow \infty$ ii. If T_n is biased for θ then it is consistent provided bias(T_n) and var(T_n) $\rightarrow 0$ as $n \rightarrow \infty$ iii. r^{th} raw moment of the sample is consistent estimator for r^{th} raw moment of population. iv. If T_n is consistent estimator of θ and $\phi(.)$ is a continuous function, then $\phi(T_n)$ is a consistent estimator of $\phi(\theta)$.	
	4.3	Interval Estimation: Concept of exact level of significance and observed level of significance, confidence coefficient, pivotal quantity, construction of confidence interval.	

Suggested Readings:

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

Semester – V	**Paper – III**
Course Code: BSC-ST 503 T	**Title of the Course: Sampling theory**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. Students will able to frame good questionnaire and test its reliability and validity.
b. Students will get an idea about different techniques of sample size determination.
c. Students will learn different techniques of selecting the random samples.
d. After completion of this course students will able to apply different sampling techniques to handle different real life problems.

Detailed Syllabus:

Unit - I		**Role of Sample Survey in Research Methodology**	**05L**
	1.1.	Objectives of Sample Survey, Methods to conduct Survey: In-person and telephone interviews, mailed and online questionnaires, Longitudinal survey method, cross sectional survey method, designing a questionnaire, characteristics of a good questionnaire, steps in implementing survey methods, scaling methods involve in survey, merits and demerits of sample survey, practical problems in planning and execution of sample survey.	
	1.2	Reliability and validity testing by using methods: Test- Retest reliability for the stability, split half test, Kuder Rechardson Coefficient (KR-20), Cronbach's Coefficient Alpha	
	1.3	Types of errors in sampling: Sampling and Non-sampling errors, reasons for the sampling and non-sampling error.	
Unit - II		**Simple Random Sampling**	**09L**
	2.1	Review of concepts of sampling, (Probability sampling and non-probability sampling), sampling units, sampled units, sampling frame, random sample, methods of drawing random sample, requisite of good sample, Simple random sampling and types of simple random sampling Results: In simple random sampling probability of selecting specified unit of the population at any given draw is equal to probability of its being selected at first draw. In simple random sampling probability that specified unit is included a sample is $\frac{n}{N}$, N= Population size, n= Sample size	
	2.2	Notations, definition of population total, population mean and population variance,	

		Simple random sampling without replacement (SRSWOR): probability of selection of particular sample using SRSWOR, unbiased estimator of population mean and population total, derivation of variance of estimator of population mean, variance of estimator of population total, derivation of expected value of sample mean sum of square, standard error of estimator of population mean and total, estimate of above standard error, confidence interval for the population mean and population total. Real life situations where SRSWOR is used, advantages and disadvantages of SRSWOR, Simple random sampling without replacement (SRSWR): probability of selection of particular sample using SRSWR, Unbiased estimator of population mean and population total, variance of estimator of population mean and population total, estimator of population mean sum of square, standard error of the estimator of population mean and total, estimators of standard error, confidence interval for the population mean and total. Real life situations where SRSWR is used, advantages and disadvantages of SRSWR, Comparison between SRSWOR and SRSWR method. Numerical examples.	
	2.3	Simple random sampling for attribute for both SRSWOR and SRSWR (Case of only two Classes): Sampling for proportion as an application of a simple random sampling with Xi as zero or one, sample proportion (p) as an estimator of population proportion (P) of units possessing a certain attribute, derivation of expectation and standard error of (p), Np as an estimator of total number of units in the population possessing a certain attribute, derivation of expectation and standard error of Np, estimation of above standard error. Numerical examples.	
Unit-III		**Stratified random Sampling**	**09L**
	3.1	Need of stratified random sampling, stratification, stratifying factor and stratum, stratified random sample as a sample drawn from individual strata using SRSWOR in each stratum. Estimation of population mean: $\bar{y}_{st} = \sum_{i=1}^{k} P_i \; \bar{y}_{n_i}$ as an estimator of population of mean, expectation of standard error of the $\bar{y}_{st}$, estimator of standard error of $\bar{y}_{st}$, confidence interval for the population mean.	
	3.2	Estimation of population total: N $\bar{y}_{st}$ as a estimator of population total, expected value and standard error of N $\bar{y}_{st}$, estimator of standard error of $N\,\bar{y}_{st}$, confidence interval for the population total. Problem of allocation, methods of allocations: Equal allocation, proportion allocation, Neyman's allocation, standard error of $\bar{y}_{st}$ in case of all the three-allocation method, comparison between SRSWOR, stratified random sampling with proportion and Nyman's allocation.	

	3.3	Optimum Allocation: Allocation using Following two principles: i. Determination of n_i's when cost (Linear Cost function) of the survey is fixed by minimizing variance of $\bar{y}_{st}$. ii. Determination of n_i's when variance of the estimator is fixed by minimizing cost (Linear cost function) of the survey. Optimum total sample size and variance of estimator and in case 1 and, optimum total sample size and total cost of survey in case 2, Gain in precision due to stratification.	
Unit-IV		**Systematic Sampling and Determination of Sample size**	**07L**
	4.1	Real life situations where systematic sampling is applicable, definition systematic sampling, only introduction of circular systematic sampling. Linear systematic sampling: Concept of linear systematic sampling, procedure of drawing a sample using systematic sample, estimation of population mean and population total in case of linear systematic sampling, standard error of the estimators.	
	4.2	Proof of the following results: i. Systematic sampling more efficient than SRSWOR if mean square among the unit of the same systematic sample is larger than population mean sum of square. ii. Systematic sampling is more efficient than SRSWOR if intraclass correlation coefficient is less than $-\frac{1}{N-1}$. iii. Systematic sampling more efficient than SRSWOR in presence of linear trend. Comparison of systematic sampling with the stratified random sampling.	
	4.3	Determination of the sample size for estimating population mean, population total and population proportion for the given: i) Margin of error and confidence coefficient. ii) Coefficient of variation of the estimator and confidence coefficient.	

Suggested Readings:

1) Cochran, W.G. (2007): Sampling Techniques, Third Edition, Wiley India Pvt. Ltd., New Delhi.
2) Murthy, M. N. (1977): Sampling Theory and Methods, Statistical Publishing Society, Kolkata
3) Singh, D. and Chaudhary, F. S. (1986): Theory and Analysis of Sample Survey Designs, Wiley Eastern Ltd., New Delhi.
4) Sukhatme, P.V., Sukhatme, B. V. (1984): Sampling theory of Surveys with Applications, Indian Society of Agricultural Statistics, New Delhi.
5) Sampath S. (2005): Sampling Theory and Methods, Second edition, Narosa, New Delhi.
6) Mukhopadhyay P (2008): Theory and methods of survey sampling. Prentice-Hall of India, New Delhi.
7) Chagbao Wu and Mary E. Thompson (2020): Sampling Theory and Practice, Springer Nature Switzerland.
8) Raghunath Arnab (2017): Survey Sampling Theory and Applications, Academic Press, Elsevier.

Semester – V	**Paper – IV**
Course Code: BSC-ST 504 T	**Title of the Course: Statistical Process Control**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. On completion of this course students will get thorough knowledge about online and offline techniques to improve quality of the process/ product.
b. Students will get sound knowledge about applications of Statistical Quality Control in the various sectors.
c. Students will able to apply appropriate tools to improve the quality of the product.

Detailed Syllabus:

Unit-I		**Introduction to Statistical Quality Control (SQC)**	**06L**
	1.1	Meaning and purpose of Statistical Quality Control (SQC), dimension of quality, application of SQC. Seven Process Control. Tools of Statistical Quality Control (SQC): (i) Check Sheet, (ii) Cause and effect diagram(CED), (iii) Pareto Diagram, (iv) Histogram, (v) Control chart, (vi)Scatter Diagram, (vii) Design of Experiments (DOE). (Only introduction of 7 PC tools is expected).	
	1.2	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. Relationship between testing of hypothesis and control chart. Construction of control charts for (i) standards given, (ii) standards not given.	
Unit-II	2.1	Control chart for variable (X-MR Chart, R chart and $\bar{X}$ chart) a) Construction and plotting of X-MR Chart. b) 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. c) 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.	**09L**

		d) 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. e) Construction of $\bar{X}$ chart when the process average (μ) is not given: drawing of control chart, plotting sample means, revision of control limits of $\bar{X}$ chart, if necessary and control limits for future purpose.	
	2.2	Control chart for attributes: p- chart: a) Construction and working of p-chart when subgroup sizes are same and value of the process fraction defective P is specified: control limits, drawing of control chart, plotting of sample fraction defectives. Determination of state of control of the process. b) p-chart when subgroups sizes are different and value of the process fraction defective P is not specified with i) separate control limits ii) average sample size iii) standardized control chart. Drawing of control chart, plotting sample fraction defectives, determination of state of control of the process. Interpretation of high and low spots. Identification of real life situations. Probability of catching a shift.	
	2.3	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.	
Unit III		Process Capability Analysis: a) Specification limits, natural tolerance limits and their comparisons, decisions based on these comparisons, estimate of percent defectives. b) 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.	**06L**
Unit IV	4.1	Acceptance Sampling for Attributes: Introduction, Concept of sampling inspection plan, Comparison between 100% inspection and sampling inspection. Procedures of acceptance sampling with rectification, Single sampling plan and double sampling plan. Probabilities of acceptance and rejection. Explanation of the terms: Producer's risk (Type I error) and Consumer's risk (Type II error), Operating characteristic (OC) curve, Acceptable Quality Level (AQL), Lot Tolerance Fraction Defective (LTFD) and Lot Tolerance Percent Defective (LTPD), Average Outgoing Quality (AOQ) and Average Outgoing Quality Limit (AOQL), Indifference quality limit, AOQ curve, Average Sample Number (ASN), Average Total Inspection (ATI), ATI curve.	**09L**

	4.2	Single Sampling Plan: Computation of probability of acceptance using Hyper geometric, binomial, Poisson and Normal approximation, Derivation and curve of OC, AOQ, ATI and ASN. Graphical determination of AOQL, Determination of a single sampling plan by: a) lot quality approach b) average quality approach. Double Sampling Plan: Evaluation of probability of acceptance using Hyper geometric, binomial, Poisson and Normal approximation, Structure of OC Curve, Derivation of AOQ, ASN and ATI (with complete inspection of second sample), Graphical determination of AOQL, Comparison of single sampling plan and double sample plan.	

Suggested Readings:

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

Semester – V	Paper – V
Course Code: BSC-ST 505 T	**Title of the Course: Regression Analysis**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. On completion of this course Students will able to handle any real life data easily and can infer conclusions by software.
b. Students will understand the different types of regression models and their applications in various field.
c. Students can learn role of residuals by using different methods.

Detailed Syllabus:

Unit-I		**Simple Linear Regression Model**	**09L**
	1.1	Regression and Model Building, Data Collection, Uses of Regression, Importance of regression in data research.	
	1.2	Simple linear regression: Linear regression as Conditional expectation of random variable Y given X, model: $Y = \beta_0 + \beta_1 X + \epsilon$, Assumptions of simple linear regression. Estimation of β_0 and β_1 by the method of least squares, Properties of estimators of β_0, and β_1, BLUE of β_0, and β_1, point estimator and interval estimator for σ^2. Hypothesis Testing on the Slope and Intercept, Interval Estimation of parameters in the model.	
	1.3	Prediction of New Observations, Coefficient of determination, No- intercept regression model, Estimation by Maximum Likelihood.	
Unit-II		**Multiple Linear Regression Model**	**09L**
	2.1	Introduction, Multiple linear regression model $Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \epsilon$, where ϵ is a continuous random variable with $E(\epsilon) = 0$, $V(\epsilon) = \sigma^2$. Assumptions of multiple linear regression, Least-Squares Estimation of the Regression Coefficients, obtaining normal equations, solutions of normal equations, Properties of the Least-Squares Estimators.	
	2.2	Estimation of σ^2, Maximum Likelihood Estimation, Tests of significance of regression parameters, Test on individual regression coefficients and subsets of coefficients, Interval estimation in regression model.	
	2.3	Residuals Analysis: Residuals, Scaling residuals (Standardized and Studentized), PRESS Residuals, Residual Plots, Partial regression and partial residual plots, Detection and treatment of outliers (introduction).	
Unit-III		**Variable Transformation and Model Building**	**06L**
	3.1	Variance- Stabilizing transformation, Various linear transformations, Weighted least square transformation.	

	3.2	Criteria for evaluating subset regression model: Coefficient of multiple determination, Adjusted R^2, Residual mean square, Mallow's Cp statistic, AIC, BIC.	
	3.3	Computational techniques for variable selection: All possible regressions, Stepwise regression methods: Forward selection, Backward Elimination, Stepwise regression (general comments and stopping rules).	
Unit-IV		**Logistic Regression Model**	**06L**
	4.1	Binary response variable, Logit transformation, estimation of parameters, interpretation of parameters, Confidence interval estimation, Tests of hypotheses of model parameters, model deviance, Test based on likelihood ratios (LR).	
	4.2	Diagnostic Checking in Logistic Regression, Other Models for Binary Response Data.	

Suggested Readings:

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

Semester – V	**Paper – VI**
Course Code: BSC-ST 506 T	**Title of the Course: Operations Research**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. After completion of this course students will formulate the real life problems in equations.
b. Student will understand the mathematical tool that are needed for solving optimization problems.
c. Students will be able to plan, schedule and control the project by using CPM and PERT techniques.

Detailed Syllabus:

Unit-I		**Introduction**	**09L**
	1.1	Definition, Application of Operations Research. Statement of the linear Programming Problem (LPP), Formulation of problem as LPP. Definition of convex region, and concave region (i) slack variable and (ii) Surplus Variable (iii) solution (iv) basic and non-basic variables (v) a feasible solution (vi) a basic feasible solution, (vii) a degenerate and non–degenerate solution and (viii) an optimal solution.	
	1.2	LPP in Canonical form and LPP in Standard form. Solution of L.P.P i) Graphical Method: solution space, unique and non-unique solutions, feasible solution, unbounded solution, infeasible solution, infinite solution, obtaining an optimal solution.	
	1.3	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 solution.	
Unit-II		**Big- M Techniques and Dual simplex method**	**06L**
	2.1	Introduction of artificial variable, Big-M method with examples.	
	2.2	Duality Theory: Writing dual of a primal problem, and its properties of (Solving problem in practical).	
Unit-III		**Transportation and Assignment Problem**	**06L**
	3.1	Transportation problem (T.P.), statement of T.P., Relation between LPP and TP balanced and unbalanced T.P. Methods of obtaining initial basic feasible solution of T. P.: i) North-West corner rule ii) Method of matrix minima (least cost method), iii) Vogel's approximation method (VAM).	

	3.2	U-V method and MODI method of obtaining Optimal solution of T.P., degenerate solution, uniqueness and non- uniqueness of optimal solutions.	
	3.3	Assignment problems: i) Statement of an assignment, relation between LPP and AP, Relation between TP and AP, balanced and unbalanced problem, ii) Relation with T.P. iii) optimal solution of an assignment problem using Hungarian method. Examples and numerical problems.	
Unit-IV		**CPM and PERT**	**09L**
	4.1	Introduction Definition of (i) Events with Merge and Burst Events, (ii) Node, (iii) Activities – Predecessor, Successor, Dummy, (iv)Critical Activity, (v)Project Duration.	
	4.2	CPM: Construction of network, Definitions of (i) earliest start time (ii) earliest finish time (iii) latest start time (iii) latest finish time for an activity. Types of float - total floats, free float, independent float and their significance. Determination of critical path.	
	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.	

Suggested Readings:

1. Gass, S.L. (2011). Linear programming methods and applications, Fifth Edition Dover Publications Inc.
2. Gupta, P.K. and Hira, D.S. Operation Research, 7th edition S. Chand and company Ltd., New Delhi. 3. Kapoor, V. K. (2006). Operations Research, S. Chand and Sons. New Delhi.
3. Saceini, M., Yaspan,A. and Friedman, L.(2013).Operation Research methods and problems, Willey International Edition.
4. Sharma, J.K. (1989). Mathematical Models in Operation Research, Tata McGraw Hill Publishing Company Ltd., New Delhi.
5. Shrinath. L.S (1975). Linear Programming, Affiliated East- West Pvt. Ltd, New Delhi.
6. Taha, H.A. (2017). Operation research: An Introduction, 10th edition, Prentice Hall of India, New Delhi.

Semester – V	**Paper – VII**
Course Code: BSC-ST 507 P	**Title of the Course: Practical V** **(Based on BSC-ST-501 T and BSC-ST-505 T)**
Credits: 02	**Total Lectures: 30 Hrs.**

Sr. No.	**Title of the Practical**	**No. of Practical's**
1	Model sampling from Weibull distribution.	1
2	Application of Beta & Weibull Distribution.	1
3	Order Statistics and Chebyshev's Inequality.	1
4	Verification of CLT.	1
5	X-MR chart	1
6	$\bar{X}$ and R chart when standers are given.	1
7	$\bar{X}$ and R chart when standers are not given.	1
8	P chart when standers are given and standers are not given.	1
9	C chart when standers are given and standers are not given.	1
10	Single sampling plan.	1
11	Determination of single sampling double sampling plan.	1

Semester – V	Paper – VIII
Course Code: BSC-ST 508 P	**Title of the Course: Practical VI** **(Based on BSC-ST-502 T and BSC-ST-503T)**
Credits: 02	**Total Lectures: 30 Hrs.**

Sr. No.	Title of the Practical	No. of Practical's
1	Simple random sampling without and with replacement (estimation of population mean, population total with standard errors. Confidence interval for population mean and population total	1
2	Simple Random Sampling for Proportions (estimation of population proportion, population total with standard errors), confidence interval for population proportion and population total.	1
3	Determination of sample size for variables and attributes.	1
4	Stratified Random Sampling: Equal, Proportional and Neyman allocation, comparison with SRSWOR.	1
5	Cost and Variance Analysis in Stratified Random sampling.	1
6	Systematic sampling. Estimation of population mean, population total with standard errors.	1
7	Method of Moments and Maximum Likelihood estimation.	1
8	Maximum Likelihood estimation (MLE) using iterative procedure.	1
9	Efficiency of (bias and unbiased) estimators.	1
10	Verification of consistent estimator.	1

Semester – V	Paper – IX
Course Code: BSC-ST 509 P	**Title of the Course: Practical VII** **(Based on BSC-ST-504 and BSC-ST-506)**
Credits: 02	**Total Lectures: 30 Hrs.**

Sr. No.	Title of the Practical	No. of Practical's
1	Simple Linear Regression: Estimation of parameters, test for significance of parameters, Interval estimation, Interpretation of parameters, Interpretation of regression model, Estimation of σ^2, alternate form of regression model.	1
2	Multiple linear regression: Estimation of parameters, test for significance of parameters, Interval estimation, computation of R^2, Adjusted R^2.	1
3	Model adequacy: Measures of model adequacy, Residual analysis by graphical method, interpretation of plots, Method of scaling, detection of outliers.	1
4	Variable transformation and model building: Techniques for variable selection(Forward selection, Backward elimination, Stepwise regression, best subset regression model)	1
5	Logistic regression: Estimation of parameters, test for significance of parameters, Interval estimation, Interpretation of parameters.	1
6	Linear programming problem by graphical method.	1
7	Linear programming problem I (Simplex method).	1
8	Linear programming problem II (Dual Simplex method).	1
9	Transportation problem.	1
10	Assignment problem.	1
11	Construction of CPM.	1
12	Construction of PERT.	1

Semester – V	**Paper – X**
Course Code: BSC-ST 510 T	**Title of the Course: Python Programming**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. Students will get sound knowledge of python programming with lot of applications in industrial field.
b. Students will build programs that use Python libraries.
c. Students will learn how to leverage the power of Python to solve tasks.
d. Practicing and learning different concepts in this course using Python coding will improve their logical thinking.

Detailed Syllabus:

Unit-I		**Introduction and Data types**	**09L**
	1.1	Introduction to Python: Installation and working with Python, Introduction of various IDEs, Python variables, Python basic Operators, Understanding python blocks.	
	1.2	Python Data Types, Declaring and using Numeric data types: int, float etc. Python Complex data types: Using string data type and string operations, Defining list and list slicing, Use of Tuple data type. String, List and Dictionary, Manipulations Building blocks of python programs, string manipulation methods, List manipulation. Dictionary manipulation, Programming using string, list and dictionary in-built functions. Working with Sets.	
	1.3	Python File Operations: Reading files, Writing files in python, Understanding read functions, read(), readline(), readlines(). Understanding write functions, write() and writelines(),Programming using file operations. Database Programming: Connecting to a database, Creating Tables, INSERT, UPDATE, DELETE and READ operations, Transaction Control, Disconnecting from a database, Exception Handling in Databases.	
Unit-II		**Control structures, Functions and Modules**	**06L**
	2.1	Control statements: Conditional blocks using if, else and elif, for loop in python, Use of while loops in python, Boolean Operators *and*, *or* and *not*, Loop manipulation using pass, continue, break. Iterables, Lists and Iterators. Augmented Assignments (‘+=’,’*=’).	

		Defining functions: def statement, Organizing Python codes using functions. Built-In range Function, Functions with Multiple Parameters. Random-Number Generation using random module.	
	2.2	Python Modules and Packages: Importing own module as well as external modules, Understanding Packages, Powerful Lamda function in python, Programming using functions, modules and external packages. Introduction to Formatted Strings.	
Unit-III		**Data Processing with NumPy and Pandas**	**06L**
	3.1	Installing and Importing NumPy, NumPy Arrays – indexing, slicing, reshaping, transpose, mathematical operations etc, Multi-dimensional NumPy Arrays, Computation on Arrays – broadcasting, comparisons, sorting, Fancy indexing etc , Structured Arrays. Different Functions on NumPy Array. Generating NumPy arrays programmatically (Generating random numbers, creating random Normal (Gaussian) distribution).	
	3.2	Introducing Pandas Objects – series, data frames, index. Operations on series and dataframes. Processing CSV and Excel files. Operations on Pandas Objects – indexing and selection, universal functions, missing data, hierarchical indexing, extraction using loc[] and iloc[]. Manipulation of data, Grouping data. Combining Dataset – concat and append, merge and join. Pivot table. Aggregation and grouping. Vectorized string operations.	
Unit-IV		**Data Visualization in Python**	**09L**
	4.1	Data Visualization using Matplotlib: Simple line plot, simple scatter plot, Simple bar diagram, Subdivided bar diagram, Multiple bar diagram, Histogram, Rod or Spike Plot, density and contour plot.	
	4.2	Formatting plots: Setting the transparency and size of axis labels, adding a shadow to the chart line, Adding a data table to the figure. Customizing Matplotlib with style. Drawing plots with colour markers, Setting ticks, labels, and grids, Adding legends and annotations.	
	4.3	Data Visualization using Seaborn library: Installing and loading Seaborn, scatter plots using seaborn, customizing seaborn plots, Adding titles and Labels, Creating Multiple Charts, Creating Categorical Plots.	

List of Practical:

Sr. No.	Practical Name	No. of Practical
1	Introduction and Installation of Anaconda Distribution	1
2	List and Tuple	1
3	Dictionary and Set	1
4	Introduction to Formatted Strings	1
5	Creating and Formatting tables using Python	1
6	Control Structure and Defining Function	1
7	Descriptive Statistics using NumPy Library	1
8	Introduction to Pandas	1
9	Simple bar diagram, Subdivided bar diagram, Multiple bar diagram using Matplotlib & Seaborn	1
10	Histogram, Rod or Spike Plot, density and contour plot using Matplotlib & Seaborn	1

Suggested Readings:

1. Data Visualization With Python For Beginners by Ai Publishing.
2. Python Data Visualization Cookbook, Second Edition by Igor Milovanović, Dimitry Foures, Giuseppe Vettigli.
3. Python for Programmers by Paul Deitel, Harvey Deitel.

Semester – V	Paper – XI
Course Code: BSC-ST 511 T	**Title of the Course: Actuarial Statistics**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. On completion of this course students will get basics concepts of Programming Language Turbo C.
b. Students will able to create programs in C related to statistical problems.
c. Practicing and learning of different concepts in this course using Turbo C will improve as well as develop their logical thinking.

Detailed Syllabus:

Unit-I		**Future Life Time Distribution**	**09L**
	1.1	Introduction: What is an Actuarial science, Basic terminology in Actuarial Science, Role of statistics in Insurance	
	1.2	Future Life Time Random Variable: definition , notation, density function, ,distribution function, survival function, Example.	
	1.3	Force of Mortality, Curtate Future Life time random variable: definition , notation, density function, ,distribution function, survival function, Example.	
Unit-II		**Life Table:**	**06L**
	2.1	Introduction, columns of life table , interpretation	
	2.2	Construction of life table under Gompert'z mortality law.	
	2.3	Construction of life table under Mekeham's mortality law.	
Unit-III		**Models for Life Insurance**	**09L**
	3.1	Introduction of simple and compound interest rate policy.	
	3.2	Different types of Interest rates. Insurance payable at the end of the year of death, present value random variable, actuarial present value	

	3.3	Derivation of actuarial present value for n-year term life insurance, whole life insurance and n-year endowment insurance	
IV		**Annuities**	**06L**
	4.1	Annuities – certain, annuity due, annuity immediate.	
	4.2	Discrete life annuities: n-year temporary life annuity due and a whole life annuity due, present value random variables of the payment, and their actuarial present values	

List of Practical:

Sr. No.	Practical Name	No. of Practical
1	Construction of life table using random survivorship approach.	1
2	Estimation of Survival Function.	1
3	Construction of life tables using Gompert'z mortality law.	1
4	Construction of life tables using Mekeham's mortality law.	1
5	Simple and Compound Interest Sum and Effective rate of interest.	1
6	Computation of net single premiums.	1
7	Calculation of present values and accumulated value for annuities certain, annuity due, annuity immediate.	1
8	Calculation of present values and accumulated value for discrete life annuities	1

Suggested Readings:

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

Semester – VI	Paper – I
Course Code: BSC-ST 601 T	**Title of the Course: Probability Distribution – II.**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. On completion of this course students will get the knowledge about different continuous distributions and their real life applications.
b. Students will learn the inter relationship among continuous distributions.
c. Students will get idea to simulate data from various continuous distributions.

Detailed Syllabus:

Unit-I		**Cauchy Distribution**	**09L**
	1.1	History of Cauchy Distribution. p.d.f. of Cauchy distribution $f(x) = \frac{\lambda}{\pi} \frac{1}{\lambda^2 + (x-\mu)^2} \quad -\infty < x < \infty; -\infty < \mu < \infty; \lambda > 0$ $= 0; \ elsewhere$ Notation: X~ C (Location=μ, Scale= λ). Nature of the probability curve. Symmetry of the density, Standard Cauchy Distribution, Distinction among t, Cauchy and Normal distribution using density curve, non – existence of moments, Distribution function, quartiles, Quartile deviation, IQR, Bowley's Coefficient of Skewness.	
	1.2	Mode, Distribution of the sample mean from Cauchy distribution and it's interpretation. Proofs of the following results: i. If X ~ t_1 then X ~C (0,1). ii. If X ~ U$\left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$, then Y = tan X has Standard Cauchy distribution. iii. If X ~C(0,1) then Y = X^2 follows $\beta_2\left(\frac{1}{2}, \frac{1}{2}\right)$. iv. If X ~C(0,1) then Y=$\frac{1}{X}$ follows C(0,1). v. If X and Y are Standard Normal Variates then $\frac{X}{Y}$ has Cauchy distribution. Applications of $C(\mu, \lambda)$.	
	1.3	History of Pareto Distribution. p.d.f. of Pareto Distribution. f(x) $= \frac{\lambda}{x^{\lambda+1}}$; $X \geq 1, \lambda > 0$ = 0 , otherwise	

		Nature of pdf curve, Mean, variance and moments. Existence of moments for different values of λ, CDF. Applications in the field of Economics and Survival analysis.	
Unit-II		**Laplace (Double Exponential) Distribution**	**06L**
	2.1	History of Laplace Distribution, p.d.f of Laplace Distribution $f(x) = \frac{\lambda}{2} e^{-\lambda\|x-\mu\|}$; $-\infty < x, \mu < \infty, \lambda > 0$ $= 0$, otherwise Notation :X~ L (Location=μ, Scale=λ). Reason for being called as Double Exponential, Nature of the probability curve. Mean and Variance, Distribution function, quartiles, IQR, Quartile deviation, Bowley's Coefficient of Skewness, MLE of μ, λ and comment on it.	
	2.2	MGF, CGF, moments and cumulants, Coefficients of skewness and kurtosis ($\beta1, \beta2, \gamma1, \gamma2$), mean deviation about mean. Proof of the following result: i. If X and Y have Standard Exponential variates then $X - Y$ follows L(0, 1). Applications and real life situations. Numerical examples.	
Unit-III		**Lognormal Distribution**	**06L**
	3.1	History of Lognormal Distribution, p.d.f of Lognormal Distribution $f(x) = \frac{1}{(x-a)\sigma\sqrt{2\pi}} exp \left\{ \frac{-1}{2\sigma^2} [\log(x - a) - \mu]^2 \right\}$; $x > a, \mu \in R, \sigma^2 > 0$ $= 0$, otherwise Notation :$X \sim LN(a, \mu, \sigma^2)$. Nature of the probability curve. $LN(0, \mu, \sigma^2)$, Distribution of $\log(X - a)$, distribution of $\log(X)$ when $X \sim LN(0, \mu, \sigma^2)$. Mean and Variance, r^{th} moment about a, central moments, $\beta1$ and $\gamma1$coefficients, quartiles, Quartile deviation, IQR, Mode.	
	3.2	Proofs of following results: i. If $X \sim N(\mu, \sigma^2)$ then $Y = e^x$ follows $LN(0, \mu, \sigma^2)$. ii. If $X \sim LN(0,0,1)$ then $Y = X^\alpha$ follows $LN(0,0,\alpha^2)$. iii. If $Xi \sim LN(0, \mu i, \sigma i^2), i = 1,2, \dots, n$ and Xi are independent then $\prod_{i=1}^{n} Xi$ follows $LN(0, \sum \mu_i, \sum {\sigma^2}_i)$. iv. If $X1, X2, \dots Xn$ are iid $LN(0, \mu, \sigma^2)$ then distribution of GM is $LN\left(0, \mu, \frac{\sigma^2}{n}\right)$. Applications and real life situations. Numerical examples.	

Unit-IV		**Bivariate Normal Distribution.**	09L
	4.1	Need of Truncation, Truncated Normal Distribution. Normal distribution N (μ, σ^2) truncated i) below a ii) above b iii)on both side(below a and above b), (a < b) its p.d.f. and derivation of mean, statement of variance(without proof). Real life situations and applications.	
	4.2	p.d.f. of a bivariate normal distribution $f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho 2}}\ exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)\right]\right\}$; $-\infty < x, y, \mu_1, \mu_2 < \infty$, ; $\sigma_1, \sigma_2 > 0$; $-1 < \rho < 1$ = 0 , otherwise Notation :$(X, Y) \sim BN\ (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Nature of surface of p.d.f, marginal and conditional distributions, identification of parameters.	
	4.3	Conditional expectation of Y given X=x as a regression line of Y on X, Conditional expectation of X given Y=y as a regression line of X on Y, independence and uncorrelatedness, joint MGF, moments using MGF. Distribution of $aX + bY + c, X/Y$. Applications and real life situations.	

Suggested Readings:

1. H. Cramer (1992). Mathematical Methods of Statistics, Princeton University Press.
2. Mood, A.M. Graybill, F.Boes, D. (2017). Introduction to Theory of Statistics, 3rd Edn., Mc-Graw Hill Series.
3. B.W. Lindgren (1993). Statistical Theory, 4th Edn., CRC Press, London.
4. Hogg, R.V. McKean, J. and Craig A.T. (2012). Introduction to Mathematical Statistics, 7th Edn., Pearson.
5. Sanjay Arora and BansiLal (1989). New Mathematical Statistics: A problem-oriented First Course, Satya Prakashan New Delhi.
6. S.C. Gupta and V. K. Kapoor (2020). Fundamentals of Mathematical Statistics, 12th Edn., Sultan Chand and Sons, 88, Daryaganj, New Delhi, 2.
7. V.K. Rohatgi and Saleh A. K. Md. E. (2015). An Introduction to Probability Theory and Statistics, 3rd Edn. Wiley, New York.
8. Feller W. (1968). An Introduction of Probability Theory and Its Applications, Vol. I, 3rd Edn., Wiley.
9. Sheldon Ross (2018). A first course in probability, 10th Edn., Pearson Education.
10. Bhuyan, K. C. (2010). Probability theory and Statistical inference, 1st Edn., New Central Book Agency.

Semester – VI	Paper – II
Course Code: BSC-ST 602 T	**Title of the Course: Testing of Hypothesis**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. On completion of this course students will get in depth knowledge of theoretical background of statistical tests.
b. Students will able to differentiate between parametric and non-parametric tests.
c. Students will understand the logic and framework of the inference of hypothesis testing.

Detailed Syllabus:

Unit - I		**Preliminaries**	**06L**
	1.1	Statistical hypothesis, problem of testing of hypotheses, Null and Alternative hypothesis. Definition and illustrations of (1) simple hypothesis, (2) composite hypothesis, (3) test of hypothesis, (4) critical region, (5) type I and type II errors.	
	1.2	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.	
Unit - II		**MP and UMP Test**	**09L**
	2.1	Definition of most powerful (M.P.) level α test of simple null hypothesis against simple alternative, Statement of Neyman - Pearson (NP) lemma for constructing the most powerful level α test of simple null hypothesis against simple alternative hypothesis. Illustrations.	
	2.2	Power function of a test, power curve, definition of uniformly most powerful (UMP) level α test for one sided alternative. Illustrations.	
Unit-III		**LRT and SPRT**	**09L**
	3.1	Notion of likelihood ratio test (LRT), $\lambda(\underline{x}) = \frac{\sup_{\theta \in \Theta_0} L(\theta\|x)}{\sup_{\theta \in \Theta} L(\theta\|x)}$. Construction of LRT for $H_0: \theta = \theta_0$ against $H_1: \theta \neq \theta_0$ for the mean of normal distribution for i) Known σ^2 ii) unknown σ^2 (one sided and two sided alternatives). LRT for variance of normal distribution for i) known μ ii) unknown μ (one sided and two sided alternatives hypotheses).	

	3.2	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 $-2\log_e \lambda(\underline{x})$	
	3.3	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.	
Unit-IV		**Non-parametric Tests**	**06L**
	4.1	Concept of non-parametric tests. Distinguish between parametric and nonparametric Tests. Concept of distribution free statistic. One tailed and two tailed test procedure of i. One sample and two samples Run test ii. Kolmogorov Smirnov test for completely specified univariate distribution (one Sample problem only) for two sided alternative hypotheses. Algebraic and graphical method, problem of ties. iii. Sign test (Normal approximation, problem of ties, sign test for paired data) iv. Wilcoxon signed rank test (Normal approximation, problem of ties, sign test for paired data) v. Median test (Normal approximation, problem of ties and Chi-square approximation). vi. Mann Whitney U test (Large sample approximation and problem of ties).	

Suggested Readings:

1. B.W. Lindgren (1993). Statistical Theory, 4th Edn., CRC Press, London.
2. Daniel, W.W. (2000) Applied Nonparametric Statistics, Duxbury Press Boston.
3. Dudewecz, E.J. and Mishra, S.N. (1988). Modern Mathematical Statistics, John Wiley and Sons,Inc.
4. Gibbons J.D. and Chakraborti, S. (2010). Non parametric Statistical Inference, CRC Press, London.
5. Hoel, P.G. Port, S. and Stone, C.(1972). Introduction to Statistical Theory, Houghton Mifflin Company (International) Dolphin Edition.
6. Hogg, R.V. McKean, J. and Craig A.T. (2012). Introduction to Mathematical Statistics, 7th Edn., Pearson.
7. Kendall, M. and Stuart, A. (1943). The advanced Theory of Statistics, Vol 1, Charles and Company Ltd., London.
8. Mood, A.M. Graybill, F.Boes, D. (2017). Introduction to Theory of Statistics, 3rd Edn.,Mc-Graw Hill Series.
9. Ramchandran, K.M. and Tsokos C. P. (2009). Mathematical Statistics with Applications, Academic Press.
10. V.K. Rohatgi and Saleh A. K. Md. E. (2015). An Introduction to Probability Theory and Statistics, 3rd Edn. Wiley, New York.
11. Bhuyan, K. C. (2010). Probability theory and Statistical inference,1st Edn., New Central Book Agency.

Semester – VI	**Paper – III**
Course Code: BSC-ST 603 T	**Title of the Course: Bio Statistics**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. Students will understand concepts of health, illness, disease and socially defined concept of sickness and define the phases of clinical trials.
b. The aim of the course is to provide contemporary education for applying statistical methods to epidemiology such as public health and related fields.
c. Students will able to solve healthcare problems by using statistical techniques.

Detailed Syllabus:

Unit - I		**Epidemiology**	**06L**
	1.1	Introduction to Epidemiology. Odds ratio: Properties, inference for odds ratios and Log odds ratios. Relationship between odds ratio and relative risk.	
	1.2	Estimation of odds ratio, Confidence interval for odds ratio. Relation with parameter in logit model, Symmetry in square contingency tables, collapsing tables and Simpson's paradox.	
Unit - II		**Introduction and Basic Concepts**	**09L**
	2.1	Clinical trials General information on history of drug discovery including Louis Pasteur (rabies and small pox), Ronald Ross and malaria, Alexander Fleming and penicillin, Jonas Salk and polio, cholera, asthma, diabetes, blood pressure, heart attack, arthritis.	
	2.2	Definition of Clinical Trial, Treatment, Placebo and Evaluation. Introduction to Randomization and Blinding, Case Report, Multicenter trial.	
	2.3	Uncertainty and Probability, Bias and Variability, Confounding and interaction.	
Unit- III		**Basic Design Considerations**	**09L**
	3.1	Goals of Clinical Trials, Target Population and patient selection, Design of clinical trials: parallel vs. cross-over designs, cross-sectional vs. longitudinal designs, objectives and end points of clinical trials.	
	3.2	Design of phase I trials, design of single – stage and multi-stage phase II trials.	
	3.3	Design and monitoring of phase III trials with sequential stopping.	

Unit-IV		**Bioequivalence**	**06L**
	4.1	Sample Size determination: Basic Concept (study objective, hypothesis, Type I and Type II error ,Power Analysis, Precision analysis,) , Sample size determination for one sample and two sample	
	4.2	Bioequivalence and bio-availability, non-inferiority trial, Practice based medical research, evidence based medicine	

Suggested Readings:

1. Shein-Chung Chow, Jen-Pei Liu : Design and Analysis of Clinical Trials.
2. A.p .Gore and S. A, Paranjape ,(2000) Course on mathematical and statistical Ecology, Kluwer, publishing Holland.
3. M.B. Kulkarni, V.R. Prayag, (2004) "Introduction to Statistical Ecology, SIPF Academy, Nasik 41.
4. Agresti A. (1996) Categorical Data Analysis. Wiley, New York.
5. J.N.S. Matthews (2006) Introduction to Randomized Controlled Clinical Trials, Chapman and Hall.
6. Stephen Sann (2000) Statistical Issues in drug Development, John Wiley.
7. Steven Diantadosi (2000) Clinical Trials – A methodological Perspective, John Wiley.
8. L.M. Friedmon, C.D. Forbes, D.L. Demats (2000) Fundamentals of Clinics Trials, Spinner. 8. Steve selvin (2004) Epidemiologic Analysis, Oxford Press.
9. M.M. Shoukni, C.A. Pavse (1999) Statistical Methods for Health Sciences, CPC Pree.
10. Steve Salvin, (1999) Statistical Analysis of Epidomiologic Data, Oxford.

Semester – VI	**Paper – IV**
Course Code: BSC-ST 604 T	**Title of the Course: Reliability Theory**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. Students will able to study different properties of non-repairable system.
b. After completion of this course, students can solve different industrial problems related to reliability.
c. Students will understand difference between component redundancy and system redundancy.
d. Students will able to analyse the system in terms of components and reliability importance.

Detailed Syllabus:

Unit - I		**Coherent System**	**08L**
	1.1.	Concept of reliability, definition of reliability, consequence of terms involves in definition of reliability, Binary system of independent components, order of the system, structure function of binary system of n components, structure function for different types of system such as series system, parallel system, k-out-n system (2-out of -3, 3-out of -4, 2-out of-4), essential parallel and series system, real life examples of all these types of systems.	
	1.2	Reliability block diagram, guidance to draw reliability block diagram, reliability block diagram for the series system, parallel system, k-out-n system (2-out of -3 3-out of -4, 2-out of-4), examples and problems.	
	1.3	Relevant and irrelevant components, unreasonable system, Coherent structure function and Coherent system, pivotal decomposition of structure function, bounds on coherent structure function, concept of redundancy, component and system redundancy, representation of above two types of redundancy using reliability block diagram. Proofs of the following theorem: 1) Let Φ be the coherent structure function of system of order n then a. $\Phi(\underline{X} \coprod \underline{Y}) \geq \Phi(\underline{X}) \coprod \Phi(\underline{Y})$. b. $\Phi(\underline{X}.\underline{Y}) \geq \Phi(\underline{X}).\Phi(\underline{Y})$.	
Unit - II		**Structural properties of Coherent system**	**09L**
	2.1	Dual of the coherent system, proofs of following results:	

		i. Dual of series system of n components is parallel system of n components. ii. Dual of the parallel system of n components is series system of n components. iii. Dual of k-out -of –n system is (n-k+1)-out of – n system.	
	2.2	Path vector and cut vector, path set and cut set, minimal path vector and minimal cut vector, minimal path set and minimal cut set, computation of minimal path and cut set for the systems up to 4 components, reliability block diagram of coherent system using minimal path sets, reliability block diagram of coherent system using minimal cut set. Proofs of the following results: i. Any coherent system can be represented as parallel arrangement of minimal series structure. ii. Any coherent system can be represented as series arrangement of minimal cut parallel structures.	
	2.3	Relative/structural importance of components with an illustration, module of the coherent system, modular decomposition of coherent system, examples of modular decomposition of the system.	
Unit-III		**Reliability of Coherent System**	**07 L**
	3.1	Reliability of components in binary system, reliability function of binary system of n components, reliability of system of n independent components, reliability function of series system, parallel system, general expression for reliability function of k-out of-n system, Basic properties of system reliability (such as reliability function is increasing function, system and component redundancy etc.	
	3.2	S-shape property of reliability function, 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 with illustration.	
	3.3	Event tree diagram, Fault tree diagram and success tress diagram, procedure to construct fault tree diagram, symbols used in fault tree diagram, fault and tree diagram for series system, parallel system, k-out of-n system, essentially parallel or series system.	
Unit-IV		**Ageing properties of coherent system**	**06L**
	4.1	Definition of survival function, hazard rate, cumulative hazard rate, residual survival function, interrelation among this function, computation of these function for the distributions such as	

		exponential, Weibull, Pareto, Hazard rate for the human being as an example of bathtub failure rate.	
	4.2	Derivations of the following results: i. Hazard rate of a series system of components having independent life times is summation of component hazard rates. ii. Lifetime of series system of independent components with independent IFR life times is IFR. IFR (DFR) class of life distribution, two definitions of IFR class life distributions and their equivalence, IFRA (DFRA) class of life distribution.	

Suggested Readings:

1. Meeker William and Escobar Luis (1998). Statistical Methods for Reliability Data, Wiley Interscience Publication, John Wiley & Sons.
2. Barlow R. E. and Proschan, Frank (1981). Statistical Theory of Reliability and Life Testing, Holt Rinebart and Winston Inc., New York.
3. Sinha, S. K. (1987). Reliability and Life testing, Second Edition, Wiley.
4. Trivedi, R.S. (2001). Probability and Statistics with Reliability, Queuing and Computer Science Applications, Prentice - Hall of India Pvt. Ltd., New Delhi.
5. Besterfield, D.H. and Michna, C.B. et al. (2009). Total Quality Management, 3rd edition, Pearson Education, Delhi.34.

Semester – VI	**Paper – V**
Course Code: BSC-ST 605 T	**Title of the Course: Design of Experiments Using MINITAB**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. On completion of this course students will get an idea about the planning, designing, conduction and analysis of experiment.
b. Students will able to apply different methods of DOE to solve different real life problems.
c. Students will able to use appropriate design for problem under consideration.

Detailed Syllabus:

Unit - I		**Basics of DOE and Completely Randomized Design**	**09L**
	1.1.	Analysis of variance (ANOVA): only concept and technique, Cochran's Theorem and its applications, fixed and random effect model. Concept of the design of the experiment and Important Phases of any Project (Plaining phase, design phase and Analysis phase) Preliminaries: Types of Experiment, Experimental unit, treatment, block, yield, replication, precision, experimental error, uniformity trials, layout of an experiment.	
	1.2	Principals of the Design: Replication, randomization and local control. Choice of size and shape of a plot for uniformity trials, the empirical formula for the variance per unit area of plots, definitions of Linear Treatment contrasts, orthogonal treatment contrasts.	
	1.3	Completely Randomized Design (CRD): Application of the principles of design of experiment in CRD, Layout of CRD, Model:: $x_{ij} = \mu + \alpha_i + €_{ij}$; $i = 1,2,\ldots\ldots.t,\ j = 1,2,\ldots\ldots\ldots.,n_i$, Assumptions in the model and interpretations, partition of total sum of squares into components. Estimation of parameters of the model, expected values of mean sums of squares, components of variance, preparation of ANOVA table, testing equality of treatment effects, Hypothesis to be tested	
		H0 : $\alpha_1 = \alpha_2 = \alpha_3 = ... = \alpha_t = 0$, F test for testing H0 with justification (independence of chi- square is to be assumed), test for equality of two specified treatment effects using critical difference (C.D).	

Unit - II		**Randomized Block Design and Latin Square Design**	**09L**
	2.1	Randomized Block Design (RBD): Application of the principles of design of experiments in RBD, layout. Model: $x_{ij} = \mu + \alpha_i + \beta_j + €_{ij}$, i= 1,2, ...t, j = 1,2,.......... b. Assumptions and interpretations. Breakup of total sum of squares into components. Estimation of parameters, expected values of mean sums of squares, components of variance, preparation of analysis of variance table, Hypotheses to be tested H_{01}: $\alpha_1 = \alpha_2 = \alpha_3 = ... = \alpha_t = 0$, H_{02}: $\beta_1 = \beta_2 = ... = \beta_b = 0$. F-test for testing H01 and H02 with justification (independence of chi-squares is to be assumed), test for equality of two specified treatment effects using CD	
	2.2	Latin Square Design (LSD): Application of the principles of design of experiments in LSD, layout, Model : $x_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + €_{ijk}$; i,j,k=1,2,⋯, m, Assumptions in the model and interpretations. Partition of total sum of squares into components. Estimation of parameters, expected values of mean sums of squares, components of variance, preparation of analysis of variance table, hypotheses to be tested. H_{01} : $\alpha_1 = \alpha_2 = \alpha_3 = ... = \alpha_m = 0$, H_{02} : $\beta_1 = \beta_2 = ... = \beta_b = 0$, H_{03} : $\gamma_1 = \gamma_2 = \cdots = \gamma_m = 0$. Justification of F test for H01, H02 and H03 Independence of chi-square is to be assumed). Preparation of ANOVA table, testing for equality of two specified treatment effects, comparison of treatment effects using C.D.	
Unit-III		**Kruskal Walis test and Efficiency of Designs**	**04L**
	3.1	Analysis of non- normal data using. i. Square root transformation for counts. ii. $\sin^{-1}(\,.\,)$ transformation for proportions, Kurskal Walis test.	
	3.2	Concept and definition of efficiency of a design, Efficiency of RBD over CRD, Efficiency of LSD over CRD, Efficiency of LSD over RBD when rows are used as blocks, LSD over RBD when columns are used as blocks.	
Unit-IV		**Total and Partial Confounding**	**08L**
	4.1	General description of m^n factorial experiment, 2^2 and 2^3 factorial experiments arranged in RBD, Definitions of main effects and interaction effects in 2^2 and 2^3 factorial experiments. Yate's procedure, preparation of ANOVA table, test for main effects and interaction effects.	
	4.2	General idea of confounding in factorial experiments, Construction of layouts in total confounding and partial confounding in 2^2 and 2^3	

		factorial experiments, Total confounding (confounding only one interaction) ANOVA table, testing main effects and interaction effects, Partial confounding (confounding only one interaction per replicate); ANOVA table, testing main effects and interaction effects.	

Suggested Readings:

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

Semester – VI	Paper – VI
Course Code: BSC-ST 606 T	**Title of the Course: Stochastic Process & Time Series Analysis.**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. Students will able to get basics concepts of stochastic Process.
b. Students will able to apply knowledge of stochastic in real life problems.
c. Students will able to fit appropriate time series model for data.
d. After completion of this course student will able to handle real time series data.

Detailed Syllabus:

Unit - I		**Introduction of Stochastic Process**	**06L**
	1.1.	Definition Random variable, Stochastic Process, state space, index set, classification of stochastic processes with examples, Types of Stochastic Process. Markov Property, Markov process, Markov chains (MC) $\{X_n, n \geq 0\}$, finite MC, time homogeneous M.C. & non homogeneous MC.	
	1.2	One step transition probabilities, and transition probability matrix (t. p. m.), stochastic matrix, Doubly stochastic matrix, graphical representation Markov chain, n-step transition probability matrix	
Unit-II		**Higher order TPM**	**09L**
	2.1	Chapman Kolmogorov equation, initial distribution. Joint distribution function of $\{ X_0, X_1, \ldots, X_n \}$, partial sum of independent and identically distributed random variables as Markov chain. Absorbing and reflecting barrier.	
	2.2	Construction of TPM: Random walk, Gambler's ruin problem, Ehrenfest chain, queueing chain, birth and death chain and real life examples graphical representation.	
Unit-III		**Introduction of Time Series**	**09L**
	3.1	Definition of time series, relation between time series and stochastic process, time series and regression. Components of time series: trend, seasonal variations, cyclical variations, irregular (error) fluctuations or noise. Exploratory data analysis: Time series plot to (i) check any trend and seasonality in the time series (ii) identify the nature of trend.	
	3.2	Methods of trend estimation and smoothing: (i) moving average, (ii) linear, parabolic, exponential, Pareto curve fitting by least squares	

		principle (iii) exponential smoothing. Choosing parameters for smoothing and forecasting. Forecasting based on exponential	
Unit-IV			**06L**
	4.1	Measurement of seasonal variations: i) simple average method, ii) ratio to moving average method, iii) ratio to trend where linear trend is calculated by method of least squares. Fitting of autoregressive model where AR(p) where p=1,2.	
	4.2	Case studies of real life Time Series: Price index series, share price index series, economic time series: temperature and rainfall time series, wind speed time series, pollution levels.	

Suggested Readings:

1. Bhat, B.R. (2000): Stochastic models: Analysis and applications, New Age International.
2. Hoel, P. G., Port, S.C. and Stone, C.J. (1972): Introduction to stochastic processes, Wiley Eastern.
3. Medhi J. (2009): Stochastic processes, New age International.
4. Ross, S. (2014): Introduction to probability models, 11th edn, Elsevier.
5. Ross, S. (1996): Stochastic processes, 2nd edn. John Wiley.
6. Taylor, H N and Karlin, S. (2010): An introduction to stochastic modelling 4th edn. Academic Press.
7. Vidyadhar Kulkarni (2016): Modelling and Analysis Stochastic Systems.3rd edn., CRC press.
8. Montgomery, D.C. and Johnson L.A. (1976): Forecasting and Time Series Analysis, McGraw Hill.
9. Farmum, N.R. and Stantorr, L.W. (1989): Quantitative Forecasting Methods, PWS Kent Publishing Company, Boston.
10. Christopher Chatfield (1975): The Analysis of Time Series, 6th edition, CRC Press.
11. Mukhopadhyay, P (2011): Applied Statistics, 2nd edition revised reprint, Books and Allied (P) Ltd.

Semester – VI	**Paper – VII**
Course Code: BSC-ST 607 P	**Title of the Course: Practical VIII** **(Based on BSC-ST-602 T and BSC-ST-604 T)**
Credits: 02	**Total Lectures: 30 Hrs.**

Sr. No.	**Title of the Practical**	**No. of Practical's**
1	Testing of hypotheses (Probability of type I and type II errors, power of the test etc.).	1
2	Construction of most powerful (MP) and uniformly most powerful (UMP) test, plotting of power function of a test.	2
3	SPRT for Bernoulli, Binomial, Poisson, Hypergeometric distributions (graphical representation also).	1
4	SPRT for normal, exponential distribution (graphical representation also).	1
5	Non- parametric tests : Sign test, Wilcoxon's signed rank test, Mann-Whitney U test.	2
6	Non- parametric tests : Run test, median test, Kolmogorov- Smirnov test.	1
7	Likelihood Ratio Test (LRT)	1
8	Determination of minimal i) cut sets ii) path sets for given coherent system.	1
9	Finding reliability h (p) of coherent system with i.i.d. components each with reliability p.	1
10	Structural and Reliability importance of the components.	1
11	Computation and plotting of hazard Rate for standard lifetime distribution (Exponential, Weibull, Gamma, Pareto). Classification of families of life distribution.	1

Semester – VI	Paper – VIII
Course Code: BSC-ST 608 P	**Title of the Course: Practical IX** **(Based on BSC-ST-603 T and BSC-ST-605 T)**
Credits: 02	**Total Lectures: 30 Hrs.**

Sr. No.	Title of the Practical	No. of Practical's
1	Analysis of CRD (equal and unequal replications) and pair wise comparison of treatment using C.D. method.	1
2	Analysis of RBD, pair wise comparison of treatments using i)C.D. method ii) Tuckey's test iii) Scheffe's test. Efficiency of R.B. D. w.r.t. C.R.D.	1
3	Analysis of LSD, pair wise comparison of treatments using C.D., Efficiency of LSD w.r.t. i) CRD ii) RBD	1
4	Non-Parametric one-way ANOVA (Kruskal Wallis test)	1
5	Analysis of $2^2$and 2^3 factorial experiments conducted in RBD.	1
6	Total confounding in 2^3 factorial experiments conducted in RBD.	1
7	Partial confounding in 2^3 factorial experiments conducted in RBD.	1
8	Construction of TPM, initial distribution given and not given and stationary probability.	1
9	Methods of trend estimation and smoothing.	1
10	Methods to measure of seasonal variations.	
11	Fitting of AR(1) and MA(1) model.	1

Semester – VI	**Paper – IX**
Course Code: BSC-ST 609 P	**Title of the Course: Project**
Credits: 02	

❖ **Guidelines for the Project:**

1) For project maximum 5 students are allowed in a group.
2) Copy of the project report to be made available to the examiner on the first day of practical examination.
3) The following points should be included in the Project Report/ Dissertation:
    a. Title of the project, name(s) of the student(s), name of the Department and College. Acknowledgement, Data Sources, Description of the computing system/software(s), Programming language(s) used, etc. (if applicable)
    b. Motivation for selecting the topic, abstract of the project, key-words of the project.
    c. Text of the project. Broadly this should cover description of the selected problem using terminology in the field of application, conversion of the problem in statistical language, literature survey, description of collected data, small illustrative data set, methodology for the analysis, interpretation of the results, validation of the results, conclusions in statistical as well as user's language, limitation of proposed solutions, directions for future work, references used, etc.
4) The division of 50 marks for the project evaluation is
    A) Internal Evaluation: Maximum marks 15
    B) End of semester Evaluation
        i Project Report : Maximum Marks 25
        ii Viva-Voce : Maximum Marks 10

Semester – V	**Paper – X**
Course Code: BSC-ST 610 T	**Title of the Course: Data Visualization Using Tableau software**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. On completion of this course students will get in depth knowledge about data visualization with Tableau.
b. Students will able to tackle the tricks used for data visualization with Tableau workbooks with attractive looks and ideas.
c. Learn tangible tactics for storytelling with data.

Detailed Syllabus:

Unit-I		**Launching and Introduction to Tableau**	**06 L**
	1.1	History of Tableau, advantages and disadvantages, Downloading and Installing Tableau Desktop, logging in to Tableau and its introduction with various features, connecting to Excel spreadsheet data source, opening the workbooks from file/from source/ or from saved data sources.	
	1.2	Renaming the field of data, adding the field in data, joining the two tables, finding the field. Tableau Interface: Importance of Data tab and Analytics tab, Viewing / Exploring different options or tabs in menu bar of Tableau, creating new worksheets or creating new dashboard or new story.	
	1.3	Data Preparation: Crosstab reports with wide tables, preparing your data for analysis, understanding the data. Data extracts and Live connection, Data types, Data Collection with IFTTT and Google Sheets.	
Unit-II		**Developing a Simple Visualization**	**09L**
	2.1	Creating Data Visualizations: Chart types, bar charts, legends, filters, line charts, step charts, jump lines, heat map, bullet charts, side by side bar, sorted bar charts, Pies and donuts, Pareto charts, custom polygons, Dragging and dropping the different features of own data into appropriate rows and columns. Dimensions used for slicing and dicing the data, Measures are numbers used for analysis.	

<table>
<tr><td></td><td>2.2</td><td>Saving the Tableau workbook and Tableau Packaged Workbook:
Two options for saving a workbook:
Tableau workbook – Without the data
Tableau packaged workbook – With embedded data.
a. It can be saved as. twb file (which is default type)
b. Data associated with the workbook is packaged and saved as .twbx file (File -Export Packaged Workbook and then save the file type as .twbx)
Opening the saved workbook, Sharing the workbooks with Tableau Reader, Saving the workbook in PDF format.</td><td></td></tr>
<tr><td></td><td>2.3</td><td>Formatting: Increasing / Decreasing the bar width, showing mark labels, changing colours of bars, changing font size, highlighting, axis formatting, formatting the tables, Top N function and its uses, inserting the lines.</td><td></td></tr>
<tr><td>Unit-III</td><td></td><td>Tableau for Presentations</td><td>09L</td></tr>
<tr><td></td><td></td><td>• Exporting an Image to Power Point
• Embedding Tableau in Power Point
• Exporting the data
• Displaying underlying data
• Exporting crosstab data
• Creating a dashboard- its importance
• Formatting the dashboard
• Adding, Removing and Renaming a dimension
• Copying or deleting a worksheet
• Changing the display from one chart type to another by (Show Me function)
• Swapping the crosstab, resizing
• Sorting (in detail too)
• Viewing details of underlying data
• Building a Hierarchy
• Aggregate measure by Analysis
• Use of Exclude and Keep function
• Use of filters with quick filters
Working with dashboard content</td><td></td></tr>
<tr><td>Unit-IV</td><td></td><td>Mapping</td><td>06L</td></tr>
<tr><td></td><td>4.1</td><td>Maps: Symbol maps, Filles maps, Density maps, map layers, maps with pie charts, mapbox maps, using the background map, Heatmaps, Dual axes, Using custom maps with a Web Map Service, Swapping maps.</td><td></td></tr>
</table>

	4.2	Clustering Analysis: Clustering in Tableau, Using custom maps with a Web Map Service. Python, R and MATLAB integration: Example: Local Regression with R	

List of Practical:

Sr. No.	Practical Name	No. of Practical
1	Introduction and Installation of Tableau	1
2	Data Processing using Tableau	1
3	Connect, Organize and Slice data using Tableau	1
4	Tableau to perform Exploratory Data Analysis	1
5	Quick table calculations to analyze the data in Tableau	1
6	Making dashboards and stories using Tableau	1
7	Designing visualizations for presentations	1
8	Map and customize data using Tableau	1
9	Clustering Analysis using Tableau	1
10	Python, R and MATLAB Integration	1

Suggested Readings:

1) Jumstart Tableau – A step by step guide to better data visualization By Arshad Khan- Apress.
2) Visual Analytics with Tableau- Alexander Loth, by Wiley.
3) Mastering Tableau 2021 by Marleen Meier, David Baldwin Third Edition- Expert Insight.
4) Tableau Your Data: Fast and Easy Visual Analysis with Tableau Software by Dan Murray- Wiley
5) Practical Tableau: 100 Tips, Tutorials, and Strategies from a Tableau Zen Master by

Semester – VI	**Paper – XI**
Course Code: BSC-ST 611 T	**Title of the Course: Basics of Data Science**
Credits: 02	**Total Lectures: 30 Hrs.**

Course Outcomes (COs):

a. On completion of this course students will get in depth knowledge of basics concepts of Data Science.
b. Students will get the skill sets needed to be a data scientist.
c. Students will get sound knowledge of career opportunities in IT Sector.
d. Students should be able to handle the different case studies using techniques of Data Science.
e. On completion of this course students will able to handle small data science projects smartly and effectively in team.

Detailed Syllabus:

Unit-I		**Introduction to Data Science**	**09L**
	1.1	Relation between Big Data and Data Science, Introduction and Importance of Data Science. Data science role and skill tracks. Current landscape of perspectives of Data Science. Role of Statistician in Data Science, Comparison between Statistician and Data Scientist, Difference between Data analysis and Data analytics.	
	1.2	Statistical Foundations of Data Science. Introduction to Data structures and Algorithms. Data preparation for knowledge discovery: Data understanding and data cleaning tools, Data transformation, Data Discretization, Data Visualization.	
Unit-II		**Exploratory Data Analysis (EDA) and Data Pipelining**	**06 L**
	2.1	Soft Skills for Data Scientists: Essential knowledge for a successful data scientist. Types of Data Science Projects. Common Mistakes in Data Science.	
	2.2	Basic tools (plots, graphs and summary statistics) of EDA. Problem of classification, classification techniques: k-nearest neighbour, decision tree, Random Forest, Naïve Bayesian, classification based on logistic regression.	
Unit-III		**Model Adequacy**	**06 L**
	3.1	Classification Model Performance: Confusion Matrix, Kappa Statistic. Learning and validation curves: Train Learning Curve, Validation Learning Curve, Accuracy vs loss, AOC and ROC Building new feature,	

		Dimensionality reduction, Feature selection, Detection and treatment of outlier.	
	3.2	Random variables, Independence, Sample space, Odds and risks, Calculate odds and risks with example, Expected values, Standard errors.	
Unit-IV		**Statistical Data Analysis**	**09 L**
	4.1	Descriptive Analytics: Measures of frequency(Displaying Frequency Distribution, Displaying Visualization Graph), Measures of dispersion, Measures of position (Two methods: Percentile ranks, Interquartile ranks).	
	4.2	Diagnostic Analytics (Functions and Applications of diagnostic analytics), Predictive Analytics (Process and applications of predictive analytics), Prescriptive Analytics.	
	4.3	Introduction to New Software used in Data Science.	

List of Practical:

Sr. No.	Practical Name	No. of Practical
1	Data Cleaning, Data Integration and Data Transformation	1
2	Classification by using K-Nearest Neighbor Algorithm	1
3	Classification by using Decision Tree Algorithm	1
4	Classification by using Random Forest Algorithm	1
5	Classification by using Naïve-Bayes Classifier Algorithm	1
6	Classification by using Logistic Regression Model	1
7	Finding the value of number of clusters k	1
8	Model Performance, Learning and Validation Curve	1
9	Descriptive Statistics in Data Science	1
10	Risk Ratio and Odds Ratio	1

Suggested Readings:

1. Introduction to Data Science by Hui Lin and Ming Li
2. Data Science Fundamentals and Practical Approaches, Understand Why Data Science is the Next by Dr. Gypsy Nandi, Dr. Rupam Kumar Sharma
3. A Textbook of Big Data by Varsha Gidde.
4. Mark Lutz’s, Learning Python, O’Really
5. Mark Lutz’s, Programming Python, O’Really
6. Jake VanderPlas, Python Data Science Handbook, O’ Reilly
7. https://docs.python.org/3/tutorial/
8. https://wiki.python.org
9. https://www.numpy.org/devdocs/user/quickstart.html
10. https://www.learnpython.org/en/Pandas_Basics