Deccan Education Society's
FERGUSSON COLLEGE (AUTONOMOUS), PUNE

Syllabus
for

T. Y. B. Sc. (Statistics)

[Pattern 2019]
(B.Sc. Semester-V)

From Academic Year
2021-22

Deccan Education Society's
FERGUSSON COLLEGE (AUTONOMOUS), PUNE
Department of Statistics
T.Y. B. Sc. Course Structure
Semester –V and VI

Semester	Paper No.	Course Code	Title	Credits	CE max marks	ESE max marks	Total max marks
V	**DSE-1A**	STS3501	Distribution Theory	2	50	50	100
V	**DSE-1B**	STS3502	Theory of Estimation	2	50	50	100
V	**DSE-2A**	STS3503	Introduction to Regression Analysis	2	50	50	100
V	**DSE -2B**	STS3504	Design of Experiments	2	50	50	100
V	**DSE-3A**	STS3505	Actuarial Statistics	2	50	50	100
V	**DSE-3B**	STS3506	Operations Research	2	50	50	100
V	**DSE-1**	STS 3507	Statistics Practical – I	2	50	50	100
V	**DSE -2**	STS3508	Statistics Practical – II	2	50	50	100
V	**DSE-3**	STS3509	Statistics Practical – III	2	50	50	100
V	**SEC-1**	STS3511	Data Analysis using C-programming	2	50	50	100
V	**SEC-2**	STS3512	Data Analysis using R software	2	50	50	100

Semester	Paper No.	Course Code	Title	Cre dits	CE max marks	ESE max marks	Total max marks
VI	**DSE-4A**	STS3601	Introduction to Stochastic Processes	2	50	50	100
VI	**DSE-4B**	STS3602	Testing of Hypotheses	2	50	50	100
VI	**DSE-5A**	STS3603	Reliability and Survival Analysis	2	50	50	100
VI	**DSE-5B**	STS3604	Applied Multivariate Analysis	2	50	50	100
VI	**DSE-6A**	STS3605	Time Series Analysis	2	50	50	100
VI	**DSE -6B**	STS3606	Biostatistics	2	50	50	100

VI	**DSE-4**	STS 3607	Statistics Practical – I	2	50	50	100
VI	**DSE-5**	STS3608	Statistics Practical – II	2	50	50	100
VI	**DSE-6**	STS3609	Statistics Practical – III	2	50	50	100
VI	**SEC-3**	STS3611	Elements of Statistical Computing and Data mining	2	50	50	100
VI	**SEC-4**	STS3612	Statistical Computing using R software	2	50	50	100

Title of the Course and Course Code	**Distribution Theory** **STS3501**	**Number of Credits : 2**
Course Outcomes (COs) **On completion of the course, the students will be able to:**		
CO1	Define various continuous probability distributions and outline the properties of probability density functions, cumulative distribution functions.	
CO2	Compute moment generating function, raw moments, central moments of different continuous probability distributions.	
CO3	Demonstrate the significance of the distributions and identify the real life situations for probability distributions.	
CO4	Analyse the relationship between different continuous distributions using the nature of the distributions.	
CO5	Determine and develop problem-solving techniques needed to accurately calculate probabilities.	
CO6	Relate the probability distributions to real life situations.	

Unit. No.	**Title of Unit and Contents**	**No. of Lectures**
I	**Gamma Distribution** $f(x) = \frac{\alpha^{\lambda}}{\Gamma\lambda} e^{-\alpha x} x^{\lambda-1}$, $x \geq 0$, $\lambda>0$, $\alpha>0$ $= 0$ otherwise Notation : $X \sim G(\alpha , \lambda)$ α : scale parameter, λ : shape parameter, Nature of probability curve for various values of shape parameter , m.g.f., c.g.f., moments, cumulants, β_1 , β_2 , γ_1, γ_2 , mode, additive property, Erlang distribution as a special case of gamma distribution, Distribution of sum of n iid exponential variables with same scale parameter.	**6**

II	**Weibull Distribution** $f(x)=\frac{\beta}{\alpha}\left(\frac{x-\gamma}{\alpha}\right)^{(\beta-1)}\exp\left\{-\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right\}\quad ,x\geq\gamma,\quad \alpha,\beta>0$ $=0$ otherwise. Notation: X~W(γ, α, β). Probability curve, location parameter, shape parameter, scale parameter. Derivation of distribution function, quartiles, mean and variance, coefficient of variation, relationship with gamma and exponential distribution. Application of this model for real life data.	**6**
III	**Cauchy Distribution** $f(x)=\frac{\lambda}{\pi}\frac{1}{\lambda^2+(x-\mu)^2}\quad -\infty<x<\infty,-\infty<\mu<\infty,\lambda>0.$ Notation: X~C(μ,λ) , Nature of the probability curve, comparison with tails of normal distribution, Derivation of distribution function, quartiles. Non – existence of moments, Statement of distribution of aX+b. Derivation of distribution of i) $\frac{1}{X}$ ii) X^2 ,where X ~ C (0,1), Problems based on these results, Statement of additive property for two Independent Cauchy variates, statement of distribution of the sample mean, comment on limiting distribution of $\overline{X}$., Statement of relationship with uniform, student's t and normal distributions. Application of this model for real life data.	**6**
IV	**Laplace (Double Exponential) Distribution** $f(x)=\frac{\lambda}{2}\exp(-\lambda\mid x-\mu\mid),\quad -\infty<x<\infty,-\infty<\mu<\infty,\lambda>0.$ Notation: X ~ L (μ , λ) Nature of probability curve, Derivation of distribution function, quartiles, MGF, CGF, Moments and cumulants, skewness and kurtosis, Derivation of Laplace distribution as the distribution of the, difference of two i.i.d. exponential random variables with mean $\frac{1}{\lambda}$, Application of this models for real life data.	**6**
V	**Lognormal Distribution** $f(x)=\frac{1}{(x-a)\sigma\sqrt{2\pi}}\exp\{\frac{-1}{2\sigma^2}[\log_e(x-a)-\mu]^2, x>a,$	**6**

	$, -\infty < \mu < \infty, \sigma > 0.$ $= 0$ otherwise. Notation : X~ LN (a, μ , σ^2), Derivation of relation with N (μ, σ^2) distribution, Nature of the probability curve, Derivation of moments (r^{th} moment of X), mean, variance, Karl Pearson's and Bowley's coefficient of skewness. Coefficient of kurtosis, derivation of quartiles and mode, Distribution of $\prod_{i=1}^{n} X_i$ when Xi's are independent lognormal random variables. Application of this model for real life data	
VI	**Order statistics** Order Statistics for a random sample of size from a continuous distribution, definition, derivation of distribution function and density function of the i^{th} order statistics $X_{(i)}$, particular cases fir I = 1 and i = n. Joint distribution of $(X)_{(i)}$, $(X)_{(i)}$ for a random sample from uniform and exponential distribution. Definition of p^{th} sample quantile X[(np]+1). Distribution of sample median for a random sample from uniform distribution**.**	**6**

References:

1. Hogg, R.V. and Craig A.T(1978). Introduction to Mathematical Stastistics, IV^{th} 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. III^{rd} Edition An Introduction to probability Theory and Mathematical Statistics Wiley Eastern Ltd .New Delhi.
6. Casella G. and Berger Robert L. (2002) Statistical Inference. 2^{nd} Edition, Duxbury Advanced series.
7. Dasgupta A. (2010) Fundamentals of Probability: A first course, , Springer, New York.

Title of the Course and Course Code	**Theory of Estimation STS3502**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Describe various terms for point estimation, interval estimations to understand problem of statistical inference. List and study the properties of point estimators.	
CO2	Explain the method to obtain estimators using maximum likelihood, method of moments, method of scoring and Fisher Information function.	
CO3	Demonstrate different situations with random sample from the standard distributions to obtain appropriate estimators. Execute the method to construct confidence intervals.	
CO4	Compare different estimators with random sample from the standard distributions with unknown parameters and examine the suitability of estimators.	
CO5	Evaluate efficiency of estimators and justify the importance of Fisher information function.	
CO6	Collect various situations to discuss about importance of an estimator of unknown parameter.	

Unit. No.	**Title of Unit and Contents**	**No. of Lectures**
I	**Point Estimation** Notion of a parameter, parameter space, sample space as a set of all possible values of $(X_1, X_2, . . , X_n)$, general problem of estimating an unknown parameter by point and interval estimation, Notion of a parameter, parameter space, sample space as a set of all possible values of $(X_1, X_2, . . , X_n)$, general problem of estimating an unknown parameter by point and interval estimation, Mean Square Error (MSE) of an estimator.	**3**
II	**Methods of Estimation** Method of moments: Derivation of moment estimators for standard distributions, Definition of likelihood as a function of unknown parameter, for a random sample from i) discrete distributions ii) continuous, distributions, distinction, between likelihood function and p.d.f./ p.m.f, Method of maximum likelihood: Derivation of maximum likelihood estimator (M.L.E.) for parameters of only standard distributions (case of two unknown parameters only for normal distribution). Use of iterative procedure to derive M.L.E. of	**3**

	location parameter μ of Cauchy distribution. Invariance property of M.L.E, M.L.E. of θ in uniform distribution over i) (0, θ) ii) (- θ, θ), iii) (mθ, nθ)(m<n) M.L.E. of θ in f(x ; θ)= exp {-(x- θ)}, x > θ, M.L.E. of location parameter in Laplace distribution, Illustrations of situations where M.L.E. and moment estimators are distinct and their comparison using mean square error.	
III	**Properties of estimators.** **Unbiasedness** Definition of an unbiased estimator, Proofs of the following results regarding unbiased estimators: a) Two distinct unbiased estimators of θ give rise to infinitely many estimators. b) If T is an unbiased estimator of θ, then ø (T) is unbiased estimator of ø (θ) provided ø (.) is a linear function. **Variance of the estimator** Notion of the Best Linear Unbiased Estimator and uniformly minimum variance unbiased estimator (UMVUE), uniqueness of UMVUE whenever it exists. Illustrations, **Sufficiency** Concept and definition of sufficiency, statement of the Fisher-Neyman factorization theorem with proof for discrete probability distribution. Pitmann – Koopman form and sufficient statistic; Exponential family of probability distributions and sufficient statistic. Proofs of the following properties of sufficient statistic: i) If T is sufficient for θ, then ø (T) is also sufficient for θ provided ø is a one to one and onto function. ii) If T is sufficient for θ then T is also sufficient for ø (θ). iii). M.L.E. is a function of sufficient statistic, **Efficiency** Fisher information function**:** Amount of information contained in statistic T=T (X_1, X_2, …, X_n), Statement	**22**

	regarding information in sample and in a sufficient statistic T, Cramer- Rao Inequality**:** Statement and proof of Cramer - Rao inequality, Cramer – Rao Lower Bound (CRLB), definition of minimum variance bound unbiased estimator (MVBUE) of ø(θ). Proofs of following results: a) If MVBUE exists for θ then MVBUE exists for ø (θ) where ø is a linear function. b) If T is MVBUE for θ then T is sufficient for θ, Comparison of variance with CRLB, relative efficiency of T_1 w.r.t. T_2 for (i) unbiased (ii) biased estimators, Efficiency of unbiased estimator T w.r.t. CRLB. **Asymptotic Behaviour of an Estimator** Consistency: Definition, proof of the following theorems:a. An estimator is consistent if its bias and variance both tend to zero as the sample size tends to infinity.b. If T is consistent estimator of θ and ø (.) is a continuous function, then ø (T) is a consistent estimator of ø (θ)	
IV	**Interval Estimation** Notion of interval estimation, definition of confidence interval (C.I), length of C.I., Confidence bounds, confidence coefficient. Definition of pivotal quantity and its use in obtaining confidence intervals. Interval estimation for the following cases: i) Mean (μ) of normal distribution (σ^2known and σ^2 unknown). ii) Variance (σ^2) of normal distribution (μ known and μ unknown) iii) Median, quartiles using order statistics.	**5**

References:

1. Dudewicz, E.J. and Mishra, S.N. (1988). Modern Mathematical Statistics, John Wiley, New York.
2. Hoel, P.G. Port, S. and Stone, C. (1972). Introduction to Statistical Theory, Houghton Mifflin Company (International) Dolphin Edition.
3. Hogg, R.V., McKean and Craig, A.T. (2019). Introduction to Mathematical Statistics (Eighth edition), Pearson.
4. Mood, A.M., Graybill, F. and Bose, D.C. (1974). Introduction to the theory of Statistics (third edition) International Student Edition, McGraw Hill.

5. Rohatagi , V.K. (2015). An introduction to Probability Theory and Mathematical Statistics, Wiley Eastern Ltd., New Delhi.
6. Kale B.K. and Murlidharan K. (2015) Introduction to Parametric Inference, Narosa Publication House, New Delhi. Kale B.K. and Murlidharan K. (2016) Introduction to Parametric Inference, Narosa Publication House, New Delhi.

Title of the Course and Course Code	**Introduction to Regression Analysis STS3503**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Recall the concept of fitting of simple regression models.	
CO2	Articulate the concept of Multicollinearity and ridge regression.	
CO3	Compare residual diagnostics and apply corrective measures.	
CO4	Analyse the multiple linear regression and logistic regression models.	
CO5	Justify weighted least squares method.	
CO6	Determine tests of hypothesis of model parameters, AIC and BIC criteria.	

Unit. No.	**TiTitle of Unit and Contents**	**No. of Lectures**
I	**Simple linear regression model** Review of simple linear regression model: $Y = \beta_0 + \beta_1 X + \varepsilon$ where ε is a continuous random variable with $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2$ Estimation of β_0 and β_1, by the method of least squares, Properties of estimators of β_0, and β_1, Estimation of σ^2, Assumption of normality of ε, Tests of hypothesis of β_1, Interval estimation in simple linear regression model, Coefficient of determination, Residual analysis: Standardized residuals, Studentized residuals, residual plots, Detection and treatment of outliers	**10**
II	**Multiple linear regression model** Review of multiple linear regression model $Y = \beta_0 + \beta_1 X_1 + . . , + \beta_p X_p +$, ε, where ε is a continuous random variable with $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2$, Estimation of regression parameters $\beta_0, \beta_1, . . .$ and β_p by the method of least squares, obtaining normal equations, solutions of normal equations, Estimation of σ^2, Assumption of normality of ε .Tests of hypothesis of regression parameters, Interval estimation in multiple linear regression model, Variable selection and model building, Residual diagnostics and corrective measures such as transformation of response variable, weighted least squares method,	**16**

	Introduction of multicolinearity, computation of VIF and brief introduction to ridge regression	
III	**Logistic regression model** Binary response variable, Logit transform, estimation of parameters, interpretation of parameters, Tests of hypotheses of model parameters, model deviance, LR test, AIC and BIC criteria for model selection, Multiple logistic regression**.**	**10**

References:

1. Draper, N. R. and Smith, H. (1998). Applied Regression Analysis Third Edition, John Wiley.
2. Hosmer, D. W. and Lemeshow, S. (2013). Applied Logistic Regression, Wiley.Third Edition
3. Montgomery, D. C., Peck, E. A. and Vining, G. G. (2012). Introduction to Linear Regression Analysis Wiley. Fifth Edition
4. Neter, J., W., Kutner, M. H., Nachtsheim, C.J. and Wasserman, W. (2005): Applied Linear Statistical Models, fifth edition, Irwin USA.
5. Chatterjee S. and Hadi A.S. (2012): Regression Analysis by Example, 5th Edition,Wiley.
6. Kleinbaum G. and Klein M. (2011): Logistic Regression, IIIrd Edition A Self learning text, Springer

Title of the Course and Course Code	**Design of Experiments** **STS3504**	**Number of Credits : 2**
Course Outcomes (COs) **On completion of the course, the students will be able to:**		
CO1	Identify relationships between cause and effect, planning and designing the experiments.	
CO2	Outline interactions among causative factors through factorial designs.	
CO3	Apply different experimental designs to real life situations.	
CO4	Analyse collected information through the experiments planned according to different designs using ANOVA and ANCOVA techniques.	
CO5	Validate the design employed in real life situations using residual analysis.	
CO6	Design a lay out of different statistical designs.	

Unit. No.	Title of Unit and Contents	No. of Lectures
I	**Design of Experiments** Analysis of variance (ANOVA): concept and technique, Basic terms of design of experiments: Experimental unit, treatments, layout of an experiment. Basic principles of design of experiments: replication, randomization and local control, choice of size and shape of a plot for uniformity trials, the empirical formula for the variance per unit area of plots. Completely Randomized Design (CRD): Application of the principles of design of experiment in CRD, Layout, Model: $X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$ i= 1,2, …….t. j = 1,2,……., n_i, assumptions and interpretations. Breakup of total sum of squares into components. Estimation of parameters, expected values of mean sums of squares, components of variance, preparation of (ANOVA) table, testing equality of treatment effects, hypothesis to be tested $H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$. Statement of Cochran's theorem. F test for testing Ho with justification (independence of chi-square is to be assumed), test for equality of two specified treatment effects using critical difference (C.D). **Randomized Block Design (RBD):** Application of the principles of design of experiments in RBD, Layout, Model: $X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$ i= 1,2, ……. t. j = 1, 2… b, assumptions and interpretations, Breakup of total sum of squares into components. Estimation of parameters, expected values of mean sums of squares, components of variance, preparation of analysis of variance table, Hypotheses to be tested, $H_{01}: \alpha_1 = \alpha_2 = \alpha_3 = \ldots = \alpha_t = 0$ $H_{02} : \beta_1 = \beta_2 = \beta_3 = \ldots = \beta_b = 0$, F test for testing H_{01} and H_{02} with justification (independence of chi- squares is to be assumed), test for equality of two specified treatment effects using critical difference (CD).	**18**

	Latin Square Design (LSD): Application of the principles of design of experiments in LSD, layout Model : $X_{ij(k)} = \mu + \alpha_i + \beta_j + Y_k + \varepsilon_{ij(k)}$, i = 1.2 … m, j = 1, 2 …, m, k = 1, 2 …., m. Assumptions and interpretations. Break up of total sum of squares into components. Estimation of parameters, expected values of mean sum of squares, components of variance, preparation of ANOVA table, hypothesis to be tested $H_{01}: \alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$, $H_{02} : \beta_1 = \beta_2 \cdots = \beta_m = 0$, $H_{03} : \gamma_1 = \gamma_2 = \cdots = \gamma_m = 0$, and their interpretations, Justification of F test for H_{01}, H_{02} and H_{03} (independence of , chi- square is to be assumed). Preparation of ANOVA table and F test for H_{01}, H_{02} and H_{03} testing for equality of two specified treatment effects, comparison of treatment effects using critical difference, linear treatment contrast and testing its significance, Linear treatment contrasts, orthogonal contrasts. Scheffe's method for comparing contrasts, Tukey's procedure for comparing pairs of treatment means(applicable to C.R.D., R.B.D. and L.S.D.), Identification of real life situations where the above designs are used and their analysis	
II	**Efficiency of Design** Concept and definition of efficiency of a design, Efficiency of RBD over CRD, Efficiency of LSD over (i) CRD (ii) RBD.	**4**
III	**Analysis of Covariance (ANOCOVA) with One Concomitant Variable** Model for covariance in CRD. Estimation of parameters (derivations are not expected), Preparation of analysis of variance – covariance table, test for β=0, test for equality of treatment effects (computational technique only)	**4**
IV	**Factorial Experiments** 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, 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	**10**

	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	

References:

1. Cochran W.G. and Cox, C.M. (1992) 2nd Edition Experimental Design, John Wiley and Sons, Inc., New York.
2. Dass, M.N. and Giri,N.C. (2020) Design and Analysis of Experiments, IIIrd Edition Wiley Eastern Ltd., New Delhi.
3. Federer W.T. (1977) Experimental Design: Oxford and IBH Publishing Co., New Delhi.
4. Goon, A.M., Gupta,M.K. and Dasgupta, B. (2016). Fundamentals of Statistics, Vol.II, The world Press Pvt. Ltd. Kolkata.
5. Johnson, R.A., Miller, I. and Freund, J(2011). Probability and Statistics for engineers, Prentice Hall, India.
6. Montgomery, D.C. (2019). Design and Analysis of Experiments, 10th John Wiley and sons Inc., New Delhi.
7. Snedecor, G.W. and Cochran, W.G. (1994). Statistical Methods, 8th edition, Affiliated East – West Press, New Delhi

Title of the Course and Course Code	**Actuarial Statistics STS3505**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Recall the concepts of financial mathematics and probability theory.	
CO2	Explain terms used in insurance business and survival analysis.	
CO3	Calculate actuarial present values and amount of premium for insurance policy.	
CO4	Classify risks into pure and speculative risk.	
CO5	Compare statistical distributions of life length random variable on the basis of survival curves and force of mortality curves.	
CO6	Construct life tables for different age groups of people.	

Unit. No.	Title of Unit and Contents	No. of Lectures
I	**Insurance Business** Insurance companies as business organizations, Role of insurance business in Economy, Concept of risk, types of risk, characteristics of insurable risk, Working of insurance business, introduction of terms such as premium, policy, policyholder and benefit, Role of Statistics in insurance, Insurance business in India, Measurement of adverse financial impact, expected value principle, Concept of utility function, Feasibility of insurance business, Illustrative examples.	7
II	**Survival Distribution and Life Tables** Time- until death random variable, its distribution function and survival function in actuarial notation, Force of mortality, Interrelations among distribution function, survival function, force of mortality and probability density function, Curate future life random variable, its probability mass function and survival function in actuarial notation, Construction of life table using random survivorship approach.	11
III	**Models for Life Insurance and Annuities** Theory of compound interest, effective rate of interest, discount factor, Insurance payable at the end of the year of death, present value random variable, actuarial present value, Derivation of actuarial present value for n-year term life insurance, whole life insurance and endowment insurance, Annuities – certain, annuity due, annuity immediate, Discrete life annuities: n-year temporary life annuity due and a whole life annuity due, present value random variables of the payment, and their actuarial present values.	12
IV	**Benefit Premiums** Concept of a loss random variable, Equivalence principle, Computation of fully discrete premium for n-year term life insurance, whole life insurance and endowment insurance, Variance of loss random variable	6

References:

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. (2010). 3rd revised edition Actuarial Statistics, Universities Press, Hyderabad, India.

Title of the Course and Course Code	**Operations Research STS3506**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Recall the concept of linear programming.	
CO2	Represent given situation into LPP and formulate the objective function , constraints and the network diagram.	
CO3	Apply the techniques of solving LPP to obtain optimal solution.	
CO4	Classify the solutions and interpret them according to the situations.	
CO5	Evaluate the CPM and PERT networks and apply project crashing techniques.	
CO6	Devise cost benefit analysis for different projects.	

Unit. No.	**Title of Unit and Contents**	**No. of Lectures**
I	**Linear Programming** Statement of the linear Programming Problem (LPP) with minimization or maximization of objective function. Formulation of problem as L.P. Problem. Definition of (i) A slack variable, (ii) A surplus Variable, L.P. Problem in (i) Canonical form (ii) standard form. Definition of (i) a solution (ii)basic and non-basic variables (iii) a feasible solution (iv) a basic feasible solution, (v) a degenerate and non–degenerate solution (vi) an optimal solution, Solution of L.P.P by (i) Graphical method (ii) Simplex Method: Obtaining Initial Basic Feasible Solution (IBFS), criteria for deciding whether obtained solution is optimal, criteria for unbounded solution, no solution, more than one solutions, introduction of artificial variable, Big-M method, Duality Theory: Writing dual of a primal problem, solution of a L.P.P. by using its dual problem.	**16**

II	**Transportation Problem** Transportation problem (T.P.), statement of T.P., balanced and unbalanced T.P. Minimization and maximization problem, obtaining basic feasible solution of T.P. by (i) Least cost method (ii) Vogel’s approximation method (VAM), u-v (MODI) method of obtaining Optimal solution of T.P., uniqueness and non-uniqueness of optimal solutions, degenerate solution. Removing degeneracy and obtaining optimal solution, Assignment Proble : Statement of an assignment problem , Minimization and maximization problem , balanced and unbalanced problem ,relation with transportation problem , optimal solution using Hungarian method , maximization case **.**	**10**
III	**Critical Path Method (CPM) and Project Evaluation and Review Techniques (PERT)** Definition of (i) Event, (ii) Node, (iii)Activity, (iv)Critical Activity, (v)Project Duration, CPM: Construction of network, Definitions, **i.** earliest start time. **ii.** earliest finish time. **iii**. latest start time.iv. latest finish time for an activity, Critical Path, Types of float, total floats, free float, independent float and their significance. Determination of critical path, 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, Cost Benefit Analysis, Definition of normal time, crash time, normal cost, crash cost, cost slope, direct cost, indirect cost, project cost, Determination of project duration and its associated cost when (i) Normal time are considered. (ii) Crash time are considered, Determination of optimal network.	**10**

References:

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.

4. Saceini, M., Yaspan,A. and Friedman, L.(2013).Operation Research methods and problems, Willey International Edition.

5. Sharma, J.K. (1989). Mathematical Models in Operation Research, Tata McGraw Hill Publishing Company Ltd., New Delhi.

6. Shrinath. L.S (1975). Linear Programming, Affiliated East- West Pvt. Ltd, New Delhi.

7. Taha, H.A. (2017). Operation research: An Introduction, 10th edition, Prentice Hall of India, New Delhi.

Title of the Course and Course Code	**Statistics Practical – I** **STS 3507**	**Number of Credits : 2**
Course Outcomes (COs) **On completion of the course, the students will be able to:**		
CO1	Recall the theory of continuous distributions.	
CO2	Carry out model sampling from Cauchy and Laplace distributions. Also carry out fitting of lognormal distribution.	
CO3	Apply the theory of confidence interval based on order statistics. Apply theory of continuous distributions to compute probability of various events.	
CO4	Analyze the regression models.	
CO5	Compare different distributions for real life situations.	
CO6	Develop method to obtain MLE of location parameter of Cauchy distribution.	

Sr. No.	Title of the Experiment
1.	Model sampling from Cauchy and Laplace distributions.
2.	Fitting of lognormal distribution.
3.	MLE of location parameter μ of Cauchy distribution.
4.	Applications of gamma distribution
5.	Applications of Weibull distribution
6.	Applications of lognormal distribution
7.	Applications of Laplace distribution.
8.	Constructions of confidence interval of μ and σ^2 of Lognormal distribution.

9.	Construction of confidence interval for population median and quartiles, based on order statistics.
10.	Simple linear regression analysis.
11.	Fitting of Multiple linear regression model.
12.	Testing hypotheses and interval estimation of regression parameters in multiple linear regression model.

Title of the Course and Course Code	**Statistics Practical – II STS3508**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Recall the principles of design of Experiments and planning of experiments.	
CO2	Illustrate the layout of designs employed.	
CO3	Apply parametric and non-parametric methods of analysis of variance.	
CO4	Analyse various experimental designs.	
CO5	Interpret the output of analysis of various statistical designs.	
CO6	Hypothesize the objectives in the given practical situation to employ a specific experimental design.	

Sr. No.	Title of the experiment
1.	Planning a lay out and analysis of CRD
2.	Analysis of R.B.D. ,
3.	Analysis of LSD.
4.	Multiple comparison (pairwise comparison) of treatment means using critical difference (C.D.), Scheffe's procedure in case of CRD, Tukey's method of pairwise comparison in case of CRD(for equal replications only).
5.	Multiple comparison (parirwise comparison) of treatment means using critical difference (C.D.), Tukey and Scheffe's procedure in case of RBD
6.	Multiple comparison (parirwise comparison) of treatment means using critical difference (C.D.), Tukey and Scheffe's procedure in case of LSD.
7.	Efficiency of RBD w.r.t. CRD, efficiency of LSD w.r.t. i) CRD ii) RBD.
8.	Analysis of covariance in CRD, testing $\beta = 0$.
9.	Analysis of 2^2 and 2^3 factorial experiments in CRD.
10.	Analysis of 2^2 and 2^3 factorial experiments in RBD.

11.	Analysis of 2^3 factorial experiments in RBD (Total confounding).
12.	Analysis of 2^3 factorial experiments in RBD (partial confounding).

Title of the Course and Course Code	**Statistics Practical –III STS3509**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Recall the concept of Linear Programming. Recall the concepts of financial mathematics and probability theory.	
CO2	Articulate the linear programming problem, Transportation problem, Assignment problem. Explain the concept of survival function and its properties.	
CO3	Apply the techniques to find optimal solution to all types of LPP. Calculate actuarial present values and amount of premium for insurance policy	
CO4	Compare statistical distributions of life length random variable .	
CO5	Determine and interpret the optimal solutions of all types of LPP	
CO6	Construct life tables for various age groups of people.	

Sr. No	Title of the Experiment
1.	Expected value principle and computation of minimum acceptable premium using utility function
2.	Survival function and force of mortality
3.	Distributions of future life time and curtate future life time random variable
4.	Construction of life tables
5.	Computation of APV for different insurance products and Discrete Annuities.
6.	To formulate real life situation into LPP, To obtain optimal solution of LPP by Simplex method.
7.	To obtain optimal solution of Primal LPP from dual LPP.
8.	Transportation Problem-I
9.	Transportation Problem-II
10.	To obtain optimal solution of assignment problem.

11.	Critical path Analysis
12.	Cost Benefit Analysis

Title of the Course and Course Code	**Data Analysis using C-programming STS3511**	**Number of Credits : 2**

Course Outcomes (COs) On completion of the course, the students will be able to:	
CO1	Introduce the concept of algorithm and programming and define differentiate terminology of C programming.
CO2	Articulate the different situations where programming is helpful for analysing the statistical data.
CO3	Apply the method to write simple C programs and use them to analyse the statistical data.
CO4	Examine the de-bugging of the programs as and when required to run the program successfully.
CO5	Consider and justify the use of parametric or non-parametric tests.
CO6	Write simple programs in C for statistical data analysis.

Unit. No.	**Title of Unit and Contents**	**No. of Lectures**
I	Algorithms and flowcharts, Data types, Operators, Data input/output, numeric and character data, printf ( ), scanf (), getchar ( ), putchar ( ), gets ( ), puts ( ).	**4**
II	Control Structure if, if….else, while, do....while, for, switch, goto, break, continue, nested loops	**3**
III	Concept, declaration, definition, initialization of array, problem using arrays, passing to function, arrays and string operations, string functions like strcpy(), strcat(), strlen(), strcmp(), strrev().	**4**
IV	Declaration, definition, recursion, user defined functions, library function, calling a function by reference and by value, local and global variables.	**3**

V	To carry out arithmetic calculations. To check whether given number is odd or even.To check whether given number m is divisible by n or not. To find maximum of 2 numbers or 3 numbers. To find area of triangle and circle. To find roots of quadratic equation. To check whether integer is prime or not. To find sum of digits of a number. To solve simultaneous linear equations. (two equations in two variables). To evaluate simple and compound interest. To prepare multiplication table. Programs using string function. To test palindrome string using string function. To sort a string using string function. To search string using string function. To combine given two strings using string function. To evaluate exp(x), sin (x), log (x) etc. using Taylor series expansion. To solve transcendental equations using Newton Raphson method. Program in C to prepare a discrete frequency distribution with given class interval from raw data To find mean, variance, coefficient of variation of n numbers To obtain correlation coefficient for given bivariate data and fit line of regression. To arrange the observations in ascending order of magnitude and find median.	**22**

Title of the Course and Course Code	**Data Analysis using R software** **STS 3512**	**Number of Credits : 2**
Course Outcomes (COs) **On completion of the course, the students will be able to:**		
CO1	Recall the principles of designs of experiments and prepare layout of designs employed. Recall the theory of continuous probability distributions.	

CO2	Estimate confidence intervals for quantiles. Demonstrate the computation of probabilities of events.
CO3	Apply parametric and non-parametric methods of analysis of variance. Apply the pivotal quantity technique to obtain confidence interval. Compare residual diagnostics using R.
CO4	Analyze various experimental designs, regression models and continuous probability distributions using R.
CO5	Interpret the output of analysis of various statistical designs, the output of analysis of regression models and MLE using R.
CO6	Hypothesize the objectives in the given practical situation to imply a specific experimental design and distributions. Test the hypotheses related to parameters in the regression models using R.

Unit. No.	**Title Of Unit and Contents**	**No. of Lectures**
I	Draw probability curves of Weibull distribution, Cauchy distribution, Laplace distribution, Log normal distribution and bi-variate normal distribution using R – software and study the characteristics of the distributions, Evaluation of probability of events related to Weibull distribution, Cauchy distribution, Laplace distribution, Log normal distribution using R – software, To obtain confidence interval of median using order statistics using R – software, MLE of location parameter μ of Cauchy distribution using R software, To study asymptotic behaviour of an estimator using R software.	**12**
II	Comparison of treatment means using box plot technique. Residual analysis, Testing normality of residuals graphically. Validating homoscedasticity, Scheffe's method for comparing contrasts, Tukey's procedure for comparing pairs of treatment means (applicable to C.R.D., R.B.D. and L.S.D.) using R software, Yate's procedure of computation of sum of squares of main and interaction effects in Factorial Experiments, using R programing, Analysis of non- normal data using. i) Square root transformation for counts. ii) $Sin^{-1}(.)$ transformation for proportions .iii) Kruskal Wallis test.	**12**
III	Fitting of Simple linear regression model using R, Test of hypothesis of β_1, interval estimation in simple linear regression model, interpretation from the value of R^2 using R, fitting of a multiple linear	**12**

	regression model using R, Test of hypothesis of regression parameters, interval estimation in multiple linear regression model, interpretation from the value of R^2 using R, Residual diagnostics: to draw residual plots and interpret it using R, Fitting and testing significance of logistic regression model with single regressor, multiple logistic regression model using R, Computation of odds ratio and writing interpretation. Test of hypothesis of model parameters. Test based on likelihood ratio, Wald's test using R.	

T.Y. B. Sc. SEM VI		
Title of the Course and Course Code	**Introduction to Stochastic Processes STS3601**	**Number of Credits : 2**
Course Outcomes (COs) **On completion of the course, the students will be able to:**		
CO1	Describe the concepts of stochastic processes in discrete time, especially concerning Markov chains and their classifications.	
CO2	Give examples of different types of stochastic processes and classify the states space of Markov chain.	
CO3	Demonstrate the computations of higher order probabilities of Markov chain using Kolmogorov's equations.	
CO4	Differentiate the stochastic processes based on classification of state space and explain the concept of stationary distribution of irreducible and ergodic Markov chain.	
CO5	Justify the stochastic simulation techniques.	
CO6	Design and develop the theory of stochastic process and its applications.	

Unit. No.	**Title of Unit and Contents**	**No. of Lectures**
I	Definition of a Stochastic process, state space, parameter space, types of stochastic processes, Markov chains (MC) $\{ X_n, n \geq 0\}$, finite MC, time homogeneous M.C. one step transition probabilities, and transition probability matrix (t.p.m.),stochastic matrix , partial sum of independent and identically distributed random variables as Markov chain.	**8**
II	Initial distribution, joint distribution function of $\{X_0, X_1, . . . , X_n\}$	**10**

	Chapman Kolmogorov equation, n-step transition probability matrix , , illustrations such as random walk, Gambler's ruin problem, partial sum of independent and identically distributed random variables as Markov chain, illustrations such as random walk, Gambler's ruin problem, Ehrenfest chain.	
III	Classification of states: Communicating states, first return probability, probability of ever return Classification of states, as persistent and transient states. Decomposition of state space, closed set of states, irreducible set of states, irreducible MC, periodicity of M.C. aperiodic M.C. ergodic M. C.	**12**
IV	Stationary distribution for a irreducible ergodic finite long run behavior of a M.C.	**6**

References:

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.

Title of the Course and Course Code	**Testing of Hypotheses** **STS3602**	**Number of Credits : 2**
Course Outcomes (COs) **On completion of the course, the students will be able to:**		
CO1	Describe the terms involved in the problem of testing of hypothesis to develop MP and UMP tests.	
CO2	Compute Type I error and Type II error to understand the concept of MP and UMP tests.	
CO3	Demonstrate MP test using NP Lemma and construction of LRT and SPRT	
CO4	Explain the situations when UMP test exists	
CO5	Justify the use of parametric or non-parametric tests.	
CO6	Develop Likelihood Ratio Test and illustrate that MP test is special case of LRT.	

Unit. No.	Title of Unit and Contents	No. of Lectures
I	**Most Powerful tests** Definition of most powerful (M.P.) level α test of simple null hypothesis against simple alternative. Statement of Neyman - Pearson (N-P) lemma for constructing the most powerful level α test of simple null hypothesis against simple alternative hypothesis. Illustrations, Power function of a test, power curve, definition of uniformly most powerful (UMP) level α test for one sided alternative. Illustrations	**10**
II	**Likelihood ratio tests** Notion of likelihood ratio test (LRT), $\lambda(x) = \sup_{\theta \epsilon \Theta 0} £(\theta \mid x) \;/\; \sup_{\theta \epsilon \Theta 0 \cup \Theta 1} £(\theta \mid x)$ 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). 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	**9**

	$\log_e \lambda(x)$ for testing $H_o : \theta \in Ho$ again $H_1 : \theta \in H_1$.	
III	**Sequential Tests** Wald's Sequential probability test procedure of strength (α, β) for simple null hypothesis of 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, Exponential. SPRT test statistic as a function of sufficient statistics. Graphical representation of SPRT.	**6**
IV	**Non-parametric Tests** Concept of non- parametric tests. Distinction between parametric and nonparametric tests. Concept of distribution free statistic. One tailed and two tailed test procedure of Run test, one sample and two sample problems, Empirical distribution function $S_n(x)$, Properties of $S_n(x)$ as estimator of F (.). Kolmogorov – Smirnov test for completely specified univariate distribution (one Sample problem only) for two sided alternative hypotheses. Comparison with chi-square test.	**11**

References:

1. Daniel, W.W. (2000): Applied Nonparametric Statistics, 2nd edn. Duxbury Press Boston.
2. Dudeweitz, E.J. and Mishra, S.N. (1988): Modern Mathematical Statistics, John Wiley and Sons, Inc.

3. Gibbons J.D.(2010): Non parametric Statistical Inference, 5th edn. CRC Press Book.
4. Hogg, R.V. and Craig, R.G. (1989): Introduction to Mathematical Statistics, fourth edition, Collier Macmillan
International Edition, Macmillan Publishing Co. Inc., New York.
5. Kale, B.K. and Muralidharan, K. (2015): Parametric Inference: An Introduction. Narosa, New Delhi.
6. Mood, A.M., Graybill, F. and Bose, D. C .(1974) : Introduction to the theory of Statistics, third edition International Student, Edition, McGraw Hill.
7. Rohatgi, V.K. (2008) : An introduction to Probability Theory Statistics, Wiley Eastern Ltd., New Delhi.

Title of the Course and Course Code	**Reliability and Survival Analysis STS3603**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Recall various probability distributions and their properties.	
CO2	Explain structural properties of coherent system.	
CO3	Compute reliability of coherent system.	
CO4	Explain the non-parametric estimation of survival function.	
CO5	Evaluate the structure function using modular and pivotal decomposition.	
CO6	Construct the reliability block diagrams and structure functions of coherent system using minimal path and cut sets.	

Unit. No.	Title of Unit and Contents	No. of Lectures
I	**Structural Properties of coherent system** Binary system of independent components, order of the system, different types of systems, concept of the structure function, structure function of series system, parallel system, k- out of- n system, (essentially parallel and series system), reliability block diagram, guidelines for construction of reliability block diagram, Coherent structure function (maximum 4 components), relevant, component, increasing structure function, pivotal decomposition of, structure function, dual of a structure function (proof of dual of series, system of	11

	order n is parallel system of order n, dual of the parallel, system of order n is a series system of order n, dual of k-out -of –n system is (n-k+1)-out of –n system). , path sets, cut sets, minimal path and cut set, representation of coherent system in terms of minimal path sets and cut sets, dual coherent structure function, relative importance of components, module of the coherent system, modular decomposition of coherent system.	
II	**Reliability of coherent system** Reliability of system of independent components, Basic properties of system reliability (such as reliability function is increasing function, system and component redundancy etc.), computation of reliability of coherent system by using minimal path and cut set representation, upper and lower bound on system reliability by using exact system reliability, relative importance of a component	**7**
III	**Ageing Properties** Survival function, probability density function, hazard function, cumulative hazard rate, mean residual life function, equilibrium residual life function, interrelation between all these function, no ageing, proof of following properties of no ageing 1. Cauchy functional equation. 2. Constant failure rate. 3. Constant mean residual life. 4. Exponential life distribution	**10**
IV	**Censoring and Nonparametric estimation of survival function** Concept of censoring, order censoring, time censoring, right random censoring, left random censoring, undersigned censoring, Nonparametric estimation of survival function, confidence band on survival function, actuarial estimator of survival function, Greenwood's formula, Kaplan Meier estimator of survival function in the presence of censored observations.	**8**

References:

1. Barlow, R. E. and Proschan F. (1975) : Statistical theory of Reliability and Life testing: Probability Models Holt, Rinehart and Winston Inc.
2. Barlow, R. E. and Proschan F. (1996) : Mathematical Theory of Reliability. John Wiley.
3. Cox, D.R. and Oakes, D. (1984) : Analysis of Survival Data, Chapman and Hall
4. Deshpande, J.V. and Purohit S.G. (2005) : Life Time Data: Statistical Models and Methods, Word Scientific.
5. Tobias, P.A. and Trindane, D. C. (1995) : Applied Reliability. Second edition. CRC Press
6. J. Gertsbakh (2005) : Reliabity theory with applications to prevective maintenance, springer
7. Jerald F. Lawless (2002): Statistical methods and methods for lifetime data 2nd edition. Wiley
8. William Q. *Meeker* and Luis A. Escobar (1998): Statistical methods for reliability data, John Wiley & Sons.

Title of the Course and Course Code	**Applied Multivariate Analysis STS3604**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Memorize the properties of univariate normal distribution. Define bivariate and multivariate distributions.	
CO2	Demonstrate the significance of the distributions and identify the real-life situations for bivariate and multivariate distributions.	
CO3	Relate multivariate normal distribution with Chi-square distribution. Sketch probability density curves for bivariate distribution.	
CO4	Determine and develop problem-solving techniques needed to accurately calculate probabilities.	
CO5	Apply discriminant analysis technique for solving appropriate real-world events.	
CO6	Construct clusters for data generated in day today life using cluster analysis.	

Unit. No.	Title of Unit and Contents	No. of Lectures
I	**Bivariate Normal Distribution.** $$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\{\frac{-1}{2(1-\rho^2)}[(\frac{x-\mu_1}{\sigma_1})^2+(\frac{y-\mu_2}{\sigma_2})^2-2\rho(\frac{x-\mu_1}{\sigma_1})(\frac{y-\mu_2}{\sigma_2})\}$$ $-\infty < x, y < \infty, -\infty < \mu_1, \mu_2 < \infty, \quad \sigma_1, \sigma_2 > 0, -1 < \rho < 1.$ $= 0$ *otherwise* . Notation: $(X, Y) \sim BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Nature of surface of p. d. f., marginal and conditional distributions, identification of parameters, regression of Y on X, independence and un-correlatedness, Derivation of MGF and moments. Statement of distribution of aX+ bY+ c , $\frac{X}{Y}$. Applications of this model for real life data.	**12**
II	**Multivariate Normal Distribution** Introduction to Multivariate Normal Distribution, Non-singular Multivariate Normal Distribution, Applications of Multivariate Normal Distribution.	**4**
III	**Discriminant Analysis** Separation and classification for two populations, Fisher's Approach to classification with two populations, Numerical examples.	**6**
	Cluster Analysis Hierarchical Clustering Methods: Single Linkage Method, Complete Linkage Method and Average Linkage Method, Non-Hierarchical Clustering Method: K- means method, Hierarchical Clustering Methods: Numerical examples, Non-Hierarchical Clustering Method: Numerical examples.	**14**

References:

1. Anderson, T. W. (2003), Introduction to Multivariate Analysis, 3rd edition, John Wiley
2. Morrison, D.F. (1990), Multivariate Statistical Methods, McGraw Hill Co.
3. Rao, C. R. (2002). Linear Statistical Inference and its Applications, Paperback edition Wiley Eastern

4. Timm, N. H. (2002), Applied Multivariate Analysis, Springer, New York
5. Johnson R.A. & Wichern, D.W. (2007). Applied Multivariate, Statistical Analysis, Sixth Edition, Prentice Hall Inc.

Title of the Course and Course Code	**Time Series Analysis STS3605**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Discuss the concept of time series and its components .	
CO2	Explain basic models of time series and different methods of estimation of trend and seasonal variation.	
CO3	Demonstrate exponential smoothing and autoregressive model fitting technique of time series analysis.	
CO4	Analyze the real life time series and carry out residual analysis.	
CO5	Determine an appropriate model to forecast future observations of the time series.	
CO6	Prepare an appropriate time series model for the given data.	

Unit. No.	Title of Unit and Contents	No. of Lectures
I	**Features of Time Series Data** Introduction to Time Series, components of Time Series, Trend, seasonality, changing variability, Test for randomness of a series against trend and seasonality: Run test, Chi-sq test	**5**
II	**Estimation of Trend and Seasonal Indices** Moving average and exponential smoothing, Forecasting based on smoothing, Double exponential smoothing, Halt Winter method with seasonality, Choosing parameters for smoothing and forecasting, Estimating mean square error of forecasting, Prediction, Intervals based on normality assumption, Estimation of seasonal Indices using ratio to trend, ratio to moving average and link relative method	**15**
III	**Time Series Analysis Through Regression Approach** Regression models for trend and seasonality, De- trending and de-seasonalizing a series, Analysis of Irregular components to examine whether the effect of trend , seasonality is removed	**8**

IV	**Introduction to Box Jenkins Techniques** Study of weak stationarity through plots involving differencing and seasonal differencing concepts, consequences of over differencing when its variance increases, Transformation of data: Transformation and differencing, AR(I) and MA (I) models, Autocorrelatation and Autocovariance, forecasts based on AR(I) and MA (I) models, MSE of forecasts. [under the assumption that parameters are estimated	**8**

References:

1. Montgomery, D.C. and Johnson L.A. (1976): Forecasting and Time Series Analysis, McGraw Hill.
2. Farmum, N.R. and Stantorr, L.W. (1989): Quantitative Forecasting Methods, PWSKent Publishing Company, Boston.
3. Christopher Chatfield (1975): The Analysis of Time Series, 6th edition, CRC Press.
4. Mukhopadhyay, P (2011): Applied Statistics, 2nd edition revised reprint, Books and Allied (P) Ltd.

Title of the Course and Course Code	**Biostatistics STS3606**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Recall concepts of health, illness, disease and socially defined concept of sickness and define the phases of clinical trials.	
CO2	Discuss the principles of epidemiology.	
CO3	Demonstrate the types of study designs used in clinical trials and use of randomization and blinding.	
CO4	Analyze 2 x 2 cross over design data.	
CO5	Evaluate proportional odds ratio.	
CO6	Write the objectives of clinical trials.	

Unit. No.	**Title of Unit and Contents**	**No. of Lectures**
I	**Epidemiology** Introduction to Epidemiology, Odds, odds ratio, relative risk, Estimation of odds ratio (OR), Confidence interval for OR. Relation	**12**

	with parameter in a logit model, Symmetry in square contingency tables, collapsing tables and Simpson's paradox.	
II	**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, Phases of clinical trial, purpose, duration, cost, drug regulatory bodies, ICH, statistical analysis plan, clinical study report	**8**
III	**Design of clinical trials** Design of clinical trials: parallel vs. cross-over designs, cross-sectional vs. longitudinal designs, objectives and end points of clinical trials, design of phase I trials, design of single – stage and multi-stage phase II trials. Design and monitoring of phase III trials with sequential stopping, design of bio-equivalence trials. Inference for 2x2 crossover design: Classical methods of interval testing for bioequivalence, nonparametric methods, 2 treatments, 2 periods cross over design	**12**
IV	**Bioequivalence and bio-availability** Bioequivalence and bio-availability, non-inferiority trial, non-parametric methods, Practice based medical research, evidence based medicine	**4**

References:

1. A.p .Gore and S. A, Paranjape ,(2000) Course on mathematical and statistical Ecology, Kluwer, publishing Holland.
2. M.B. Kulkarni, V.R. Prayag, (2004) "Introduction to Statistical Ecology, SIPF Academy, Nasik 41.
3. Agresti A. (1996) Categorical Data Analysis. Wiley, New York.
4. J.N.S. Matthews (2006) Introduction to Randomized Controlled Clinical Trials, Chapman and Hall.
5. Stephen Sann (2000) Statistical Issues in drug Development, John Wiley.
6. Steven Diantadosi (2000) Clinical Trials – A methodological Perspective, John Wiley.
7. 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.

Title of the Course and Course Code	**Statistics Practical - I** **STS 3607**	**Number of Credits : 2**
Course Outcomes (COs) **On completion of the course, the students will be able to:**		
CO1	Introduce the concept parametric and non-parametric tests. Further also introduce concept of sequential tests. Introduce concept of Markov chain. Define the various terms involved in Markov chain. Evaluate transition probability matrix of Markov chain. Recall the method of evaluation of reliability of any system.	
CO2	Demonstrate the various situations to compute probability of type-1 error, power of the test. Also demonstrate the MP tests and UMP tests.	
CO3	Apply various parametric and non-parametric tests to real life data. Also apply MP and UMP tests. Compute reliability of coherent system.	
CO4	Examine the suitability of tests. Verify the underlying conditions for applying the tests.	
CO5	Consider and justify the use parametric or non-parametric tests. Also justify ergodicity of Markov chain. Determine minimal path and cut sets for coherent systems.	
CO6	Develop the method to obtain the realization of Markov chain.	

Sr. No.	Title of the experiment
1	Testing of hypotheses (Probability of type I and II errors, power of a test etc).
2	Construction of most powerful (MP) test.
3	Construction of uniformly most powerful (UMP) test, plotting of power function of a test.
4	Determination of minimal i) cut sets ii) path sets for given coherent system.
5	Finding reliability h (p) of coherent system with i.i.d. components each with reliability p. Graph of h (p) against p. S shaped ness property.
6	Non-parametric tests: Run test, median test. Kolmogorov - Smirnov test.
7	SPRT for Bernoulli, Binomial, Poisson. (graphical representation also).
8	SPRT for normal, exponential distribution. (graphical representation also).
9	Realization of Stochastics processes
10	Stochastic Process-I

11	Stochastic Process-II
12	Stochastic Process-III

Title of the Course and Course Code	**Statistics Practical – II** **STS 3608**	**Number of Credits : 2**
Course Outcomes (COs) **On completion of the course, the students will be able to:**		
CO1	Discuss the techniques of smoothing, time series component estimation.	
CO2	Compare various time series models fitted through the residual analysis, Identify the real-life situations for bivariate and multivariate distributions.	
CO3	Use one sample test on mean vector in multivariate normal distribution.	
CO4	Analyse different time series data.	
CO5	Determine an appropriate model to forecast future observations of the time series.	
CO6	Build an appropriate model of time series for the real life data.	

Sr. No.	Title of Experiment
1.	Sketching probability curves for bivariate normal distribution for different parameters
2.	Applications of bivariate normal distribution
3.	Applications of multivariate normal distribution
4.	Inference of parameters of multivariate normal distribution-I
5.	Inference of parameters of multivariate normal distribution-II
6.	Analysis of Time Series _I (Estimation of trend in Time series by regression method, moving averages,)
7.	Analysis of Time Series _II (Single, Double and Triple exponential smoothing technique)
8.	Analysis of Time Series _III (Estimation of seasonal indices by ratio to trend method and method of link relatives)
9.	Analysis of Time Series _IV (Test of randomness, sample autocorrelations, L Jung Box test)
10.	Analysis of Time Series _V (Fitting of AR (1) model)
11	Bio Statistics practical-I
12	Bio Statistics practical-II

Title of the Course and Course Code	**Statistics Practical – III STS 3609**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Recall the techniques of data presentation like tabulation, bargraph , pie chart , box plot .	
CO2	Discuss different characteristics of the data	
CO3	Apply different methods of collection of data	
CO4	Analyze the data using statistical tools and R-software	
CO5	Determine the hypotheses and validate using appropriate statistical tests	
CO6	Prepare the project report and make the presentation.	

The students are supposed to carry out project work during this semester. This project work should be done in a group of not more than 6 students.

The project will be evaluated on the following parameters

E - 50 marks

SE -50 marks at the end of the semester

Concurrent Evaluation (CE)

(i) Initial presentation of the problem -10 marks

(ii) Data collection/ Methods to be used / layout of project ---10 marks

(iii) Submission of the project & final presentation in scheduled time----20 marks

(i) Attendance & active participation for project work ----10 marks

End Semester Examination (ESE)

Following are the details for evaluation of the project:

Details about submission marks		
1.	Technical problem and motivation	2
2.	Abstract and key words	2
3.	Conversion of the technical problem into statistical language	4
4.	Representation of raw data / questionnaire	2

5.	Exploratory data analysis	5
6.	Statistical analysis	5
7.	Conclusions in relevant language , scope / limitations	5
Total		25

Note: Viva would be conducted individually. Viva will be taken for 15 minutes for each candidate.

- The division of 25 marks for viva is as follows:

1.	Understanding of technical problem.	10
2.	Understanding of Statistical Techniques used to solve the problem	10
3.	Overall impression	5
Total		25

Title of the Course and Course Code	**Elements of Statistical Computing and Data mining STS 3611**	**Number of Credits : 2**
Course Outcomes (COs) On completion of the course, the students will be able to:		
CO1	Recall the concept of random numbers and interrelation among probability distributions	
CO2	Compare the procedures of random number generation and describe the concept of classification.	
CO3	Demonstrate different random samples generation procedures from probability distributions and apply binary classifier in real life situations.	
CO4	Compare the random number generation procedures on the basis of entropy.	
CO5	Estimate sensitivity, specificity and accuracy of a classifier	
CO6	Create confusion matrix and ROC curve	

Unit. No.	Title of Unit and Contents	No. of Lectures
I	**Random number generators** Concept of randomness, revision of random sample, generating randomness, practical difficulties in generating randomness, Mid square algorithm and its limitations Requisites of good random number generator, Test for randomness, Run test, poker's test , Concept of entropy as measure of randomness, Significance of Kendall's τ / spearman's rank correlation empirical test viz., Kolmogorov Smirnov test, Specific random number generators : Linear Congruential Generator(L.C.G.), Mixed L.C.G., Multiplicative generator.	8
II	**Generating random variates** Methods to generate random variates. (a) Inverse transformation method. (b) Method of composition (c) Method of convolution (d) Acceptance Rejection method, Generating discrete random variates. (Using recurrence relation and inverse transformation method) Generating samples from, Bernoulli distribution, Binomial distribution (Also using composition method.), Poisson distribution, Geometric distribution (also using relation with exponential distribution.), Generating continuous random variates. Generating samples from, Uniform distribution using inverse transformation, Exponential distribution, Normal distribution using, (i) Box muller method. (ii) Atkinson pearce method. (iii) Acceptance rejection method. (i)using Laplace distribution as majorising function. (ii)using exponential distribution as majorising function. (iii)using Cauchy as majorising function	8
III	**Applications of random number generation and random variate generations.** Monte carlo integration, Computation of definite integral, Error in computation. (confidence interval for the value of integral)**,** Computation of some special integrals, empirical MGF, empirical Laplace transform, Estimating moments	8

	Introduction to data mining Concept of database, Extraction of data from database with respect to research objective, Concept of receiving operating characteristics (ROC) of algorithm / method, Concept of specificity, sensitivity, accuracy, Mathew's correlation coefficient, applications of ROC to association of some simple statistical procedures.	**12**

References:

1. Averill M. Law (2000): Simulation Modelling and Analysis, M.C. Graw Hill
2. Debasis Kundu, Ayanendranath Basu (2004): Statistical Computing, Narosa Publishing House
3. Do Le *Minh* (2001): Applied Probability Models, Duxbury.

Title of the Course and Course Code	**Statistical Computing using R software** **STS 3612**	**Number of Credits : 2**
Course Outcomes (COs) **On completion of the course, the students will be able to:**		
CO1	Recall the concept of testing of hypotheses, stochastic process and survival function	
CO2	Compare different time series models observing residual plots and error measures using R.	
CO3	Apply autoregressive modelling and exponential smoothing procedures for time series data using R. Apply Markov chain to study the behavior of the process.	
CO4	Explain the non-parametric estimation of survival function using R.	
CO5	Judge models of time series using different error measures and identify the different states of Markov chain.	
CO6	Create an appropriate model for real life time series and realization of stochastic process.	

Unit. No.	**Title of Unit and Contents**	**No. of Lectures**
I	**Analysis of time series using R** Analysing time series using time series plots, Smoothing of time series and forecasting using r, Testing randomness and stationarity in time series, transformations, differencing, testing serial autocorrelations, fitting of autoregressive models to the time series, Residual Analysis, comparison of different time series models, validity of assumptions using normal probability plots	**12**

II	**Ageing and Nonparametric estimation of survival function** Positive and negative ageing: IFR, DFR ,IFRA, DFRA, bathtub failure rate, Classification of following parametric families of life distribution according to aging using R**:** Weibull, Gamma, lognormal, linear failure rate, Makeham, Pareto, Lehman, Actuarial estimator of survival function using R, Greenwood's formula, Kaplan Meier estimator of survival function in the presence of censored observations using R.	**12**
III	**Testing of Hypotheses & Stochastic Process** Testing of hypotheses (Probability of type I and II errors, power of a test etc.), Realization of stochastic process, Computations of n step transition probability matrix and discuss long term behaviour, Some illustrations of n step transition probability matrix to obtain periodicity of state	**12**