Deccan Education Society's

FERGUSSON COLLEGE, PUNE
(AUTONOMOUS)

SYLLABUS UNDER AUTONOMY

FIRST YEAR B.Sc.(Computer Science)
SEMESTER –I

SYLLABUS FOR F.Y.B.Sc.(Computer Science)
STATISTICS

Academic Year 2016-2017

FERGUSSON COLLEGE (AUTONOMOUS), PUNE
DEPARTMENT OF STATISTICS

Course Structure for F.Y. B.Sc. Statistics for Computer Science

F.Y. B.Sc. Statistics for Computer Science			
Sr. No.	**Subject of Title**	**Semester**	**Subject Code**
1	Descriptive Statistics	1	STC1101
2	Probability Theory and Discrete Probability Distributions		STC1102
3	Statistics Practical Semester I		STC1103
4	Regression Analysis and Time Series	2	STC1201
5	Continuous Probability Distributions and Inference		STC1202
6	Statistics Practical Semester II		STC1203

Learning Objectives:

Statistics deals with any decision making activity in which there is certain degree of uncertainty and helps in taking decisions in an objective & rational way. Statistics is required and extensively used in a vast spectrum of computer based applications. This course covers basic concepts and terminology in Statistics and covers basic tools and methods required for data analysis.

Data Mining and Warehousing, Theoretical Computer Science, Reliability of a Computer Programme or software, Digital Image Processing, Embedded Systems are few applications where Statistics can be extensively used.

PAPER CODE:STC-1101 **PAPER –I: TITLE: Descriptive Statistics** **[Credit -2: No. of Lectures 36]**		
	Title and Contents	**No. of Lectures**
Unit -I	**Data condensation and Graphical methods** 1.1 Raw data, attributes and variables, discrete and continuous variables. 1.2 Presentation of data using frequency distribution and cumulative frequency distribution. (Construction of frequency distribution is not expected) 1.3 Graphical Presentation of frequency distribution –histogram, stem and leaf chart, less than and more than type ogive curves. 1.4Numerical problems related to real life situations.	**7**
Unit -II	**Measures of central tendency and dispersion** 2.1 Measures of Central tendency: Mean, Mode, Median. Examples where each one of these is most appropriate. 2.2 Partition values: Quartiles, Box-Plot. 2.3Variance, Standard Deviation, Coefficient of Variation. (Section 2.1 to 2.3 should be covered	**10**

	for raw data, ungrouped frequency distribution and exclusive type grouped frequency distribution)	
Unit –III	**Moments** 3.1 Raw and Central moments: definition, computations for ungrouped and grouped data (only up to first four moments). 3.2 Relation between raw and central moments upto fourth order. 3.3 Numerical problems related to real life situations.	**3**
Unit –IV	**Measures of Skewness and Kurtosis** 4.1 Concept of symmetric frequency distribution, skewness, positive and negative skewness. 4.2 Measures of skewness-Pearson's measure, Bowley's measure, $\beta 1, \gamma 1$. 4.3 Kurtosis of a frequency distribution, measure of kurtosis($\beta 2, \gamma 2$) based upon moments, type of kurtosis: leptokurtic, platykurtic and mesokurtic. 4.4 Numerical problems related to real life situations.	**6**
Unit –V	**Correlation and Linear Regression** 5.1 Bivariate data, Scatter diagram. 5.2 Correlation, Positive Correlation, Negative correlation, Zero Correlation 5.3 Karl Pearson's coefficient of correlation (r), limits of r ($-1 \leq r \leq 1$), interpretation of r, Coefficient of determination (r^2) 5.4 Karl Pearson's coefficient of correlation between ranks 5.5 Meaning of regression, difference between correlation and regression. 5.6 Fitting of line $Y = a+bX$ 5.7 Concept of residual plot and mean residual sum of squares. 5.8 Numerical Problems.	**10**

References:

1. Fundamentals of Applied Statistics(3rd Edition), Gupta and Kapoor, S.Chand and Sons, New Delhi, 1987.

2.An Introductory Statistics, Kennedy and Gentle.
3. Statistical Methods, G.W. Snedecor, W.G. Cochran, John Wiley & sons, 1989.
4. Introduction to Linear Regression Analysis**,** Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, Wiley

PAPER CODE:STC-1102

PAPER –II: TITLE: Probability theory and Discrete probability distributions

[Credit -2: No. of Lectures 36]

	Title and Contents	No. of Lectures
Unit -I	**Detailed Review / Revision of Theory of Probability** 1.1 Counting Principles, Permutation, and Combination. 1.2 Deterministic and non-determination models. 1.3 Random Experiment, Sample Spaces (finite and countably infinite) 1.4 Events: types of events, Operations on events. 1.5 Probability - classical definition, probability models, axioms of probability, probability of an event. 1.6 Theorems of probability (with proof) i) $0 \leq P(A) \leq 1$ ii) $P(A) + P(A') = 1$ iii) $P(A) \leq P(B)$ when $A \subset B$ iv) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ 1.7 Numerical problems related to real life situations.	**5**
Unit -II	**Advanced Theory of Probability** 2.1Concepts and definitions of conditional probability, multiplication theorem $P(A \cap B)=P(A).P(B\|A)$ 2.2 Bayes' theorem (without proof) 2.3 Concept of Posterior probability, problems on posterior probability. 2.4 Definition of sensitivity of a procedure, specificity of a procedure. Application of Bayes' theorem to design a procedure for false positive and false negative.	**10**

	2.5 Concept and definition of independence of two events. 2.6 Numerical problems related to real life situations.	
Unit –III	**Discrete Random variable** 3.1 Definition of random variable and discrete random variable. 3.2 Definition of probability mass function, distribution function and its properties. 3.3 Definition of expectation and variance, theorem on expectation. 3.4 Determination of median and mode using p.m.f. 3.5 Numerical problems related to real life situations.	**8**
Unit –IV	**Standard Discrete Probability Distributions** 4.1Discrete Uniform Distribution: definition, mean, variance. 4.2 Bernoulli Distribution: definition, mean, variance, additive property. 4.3 Binomial Distribution: definition, mean, variance, additive property. 4.4 Geometric Distribution (p.m.f $p(x) = pq^x$, $x = 0,1,2.....$): definition, mean,variance. 4.5 Poisson Distribution: definition, mean, variance, mode, additive property, limiting case of B(n, p) 4.6 Illustration of real life situations. 4.7 Numerical problems related to real life situations.	**13**

References:

1.Statistical Methods, G.W. Snedecor, W.G. Cochran, John Wiley & sons, 1989.

2.Fundamentals of Applied Statistics(3^{rd} Edition), Gupta and Kapoor, S.Chand and Sons, New Delhi, 1987.

3. Modern Elementary Statistics, Freund J.E., Pearson Publication, 2005.

4. Probability, Statistics, Design of Experiments and Queuing theory with applications Computer Science, Trivedi K.S. ,Prentice Hall of

India, New Delhi,2001.
5. A First course in Probability 6th Edition , Ross, Pearson Publication, 2006.
6. Introduction to Discrete Probability and Probability Distributions, Kulkarni M.B., Ghatpande S.B.,SIPF Academy, 2007.

	PAPER CODE:STC-1103 **PAPER –III: Statistics Practical Semester I** **[Credit -2: No. of Practicals 10]**
	Title of Experiment/ Practical
1	Measures of central tendency
2	Measure of dispersion
3	Measures of skewness and kurtosis
4	Advanced theory of probability
5	Fitting of binomial and Poisson distribution
6	Correlation and Linear Regression
7	Discrete Probability Distributions
8	Introduction to R and graphical methods using R
9	Descriptive statistics using R
10	Fitting of discrete probability distributions using R

Deccan Education Society's

FERGUSSON COLLEGE, PUNE
(AUTONOMOUS)

SYLLABUS UNDER AUTONOMY

FIRST YEAR B.Sc.(Computer Science)
SEMESTER –II

SYLLABUS FOR F.Y.B.Sc. (Computer Science)
Statistics
Academic Year 2016-2017

PAPER CODE: STC-1201

PAPER –I: TITLE: Regression Analysis and Time series

[Credit -2: No. of Lectures 36]

	Title and Contents	No. of Lectures
Unit -I	**Multiple and Partial Correlation and Regression (for trivariate data)** 3.1 Yule's notation and concept of multiple regression. 3.2 Fitting of multiple regression plane. 3.3 Partial regression coefficient, interpretation. 3.4 Multiple correlation coefficient, concept, definition, computation and interpretation. 3.5 Partial correlation coefficient, concept, definition, computation and interpretation.	**8**
Unit -II	**Non Linear Regression** 2.1Second degree curve 2.2 Growth curve models of the type i) $Y = ae^{bX}$ ii) $Y = ab^{X}$ iii) $Y = aX^{b}$ 2.3 Logistic model $Y = k / (1+e^{a+bx})$ 2.4 Numerical problems related to real life situations.	**4**
Unit –III	**Time Series** 3.1 Meaning and Utility. 3.2 Components of Time Series. 3.3 Additive and Multiplicative models. 3.4 Methods of estimating trend: moving average method, least squares method and exponential smoothing method. 3.5 Elimination of trend using additive and multiplicative models. 3.6Measurement and estimation of seasonal variations using link relative method and ratio to trend method 3.7 Simple time series models: AR (1), AR (2). 3.8 Numerical problems related to real life situations.	**10**

Unit –IV	**Non parametric tests** 4.1 Run test 4.2 Sign test. 4.3 Kolmogrov - Smirnov test 4.4 Mann – Whitney test 4.5 Numerical problems related to real life situations.	**6**
Unit –V	**Simulation** 5.1 Introduction to Simulation, merits and demerits. 5.2 Monte Carlo Simulation 5.3 Pseudo-random number generator ,requisites of a good random number generator, Testing these requirements by using various test of hypothesis using Run test, goodness of fit test, Sign test etc. 5.4 Model Sampling from uniform and exponential distribution. 5.5 Model sampling from Normal distribution using Box-Muller transformation. 5.6 Numerical problems related to real life situations.	**8**

References:

1.Statistical Methods, G.W. Snedecor, W.G. Cochran, John Wiley & sons, 1989.

2. Statistical Methods, J. Medhi, New Age International, 1992.

3. Time Series Methods, Brockell and Devis, Springer, 2006.

4. Time Series Analysis,4th Edition, Box and Jenkin, Wiley, 2008.

5. Simulation and Modelling Analysis,Kelton and Law, Tata McGraw Hill, 2007.

6.System Simulation with Digital Computer, Narsingh Dev, Prentice Hall, 2003.

PAPER CODE:STC -1202

PAPER–II:TITLE: Continuous Probability Distributions and Inference

[Credit -2: No. of Lectures 36]

	Title and Contents	No. of Lectures
Unit -I	**Continuous Random Variable** 1.1 Definition of continuous random variable (r.v.), 1.2 Probability density function (p.d.f.), 1.3 Cumulative distribution function (c.d.f.), its properties. 1.4 Calculation of mean, mode, median, variance, standard deviation for continuous r. v. 1.5 Numerical problems related to real life situations.	**6**
Unit -II	**Standard Continuous Probability Distributions** 2.1 Uniform Distribution: statement of p.d.f., mean, variance, nature of probability curve. 2.2 Exponential Distribution: statement of p.d.f. of the form, $f(x) = (1/\theta)\, e(-x/\theta)$, mean, variance, nature of probability curve, lack of memory property. 2.3 Normal Distribution: statement of p.d.f., identification of parameters, nature of probability density curve, standard normal distribution, symmetry, distribution of aX+b, aX+bY+c where X and Y are independent normal variables, computations of probabilities using normal probability table, normal approximation to binomial and Poisson distribution , central limit theorem (statement only), normal probability plot. 2.4 Pareto Distribution: p.d.f., mean, variance, applications. 2.5 Numerical problems related to real life situations.	**12**
Unit –III	**Testing of hypothesis** 3.1Definitions: population, statistic, SRSWR, SRSWOR, random sample from a probability	**4**

	distribution, parameter, statistic, standard error of estimator. 3.2 Concept of null hypothesis and alternative hypothesis, critical region, level of significance, type I and type II error, one sided and two sided tests, p-value.	
Unit –IV	**Large Sample Tests** 4.1 Ho: $\mu = \mu o$ Vs H1: $\mu \neq \mu o$, $\mu < \mu o$, $\mu > \mu o$ (One sided and two sided tests) 4.2 Ho: $\mu 1 = \mu 2$ Vs H1: $\mu 1 \neq \mu 2$, $\mu 1 < \mu 2$, $\mu 1 > \mu 2$(One sided and two sided tests) 4.3 Ho: P = Po Vs H1: P≠Po, P < Po, P > Po (One sided and two sided tests) 4.4 Ho: P1 = P2 Vs H1: P1≠P2, P1< P2, P1> P2 (One sided and two sided tests) 4.5 Numerical problems related to real life situations.	**5**
Unit –V	**Tests based on t, Chi-square and F-distribution** 5.1 Ho: $\mu = \mu o$ Vs H1: $\mu \neq \mu o$, $\mu < \mu o$, $\mu > \mu o$ (One sided and two sided tests) 5.2 Ho: $\mu 1 = \mu 2$ Vs H1: $\mu 1 \neq \mu 2$, $\mu 1 < \mu 2$, $\mu 1 > \mu 2$ (One sided and two sided tests) 5.3 Paired t-test. 5.4 Chi square test for goodness of fit 5.5 Test for independence of attributes (m X n contingency table) 5.6 Test for significance of variation for a population. (One sided and two sided tests) 5.7 Test for equality of population variances(One sided and two sided tests) 5.8 Numerical problems related to real life situations.	**9**

References:

1.A First course in Probability 6th Edition , Ross, Pearson Publication, 2006.

2.Modern Elementary Statistics, Freund J.E., Pearson Publication, 2005.

3.Probability, Statistics, Design of Experiments and Queuing theory with applications Computer Science, Trivedi K.S. ,Prentice Hall of India, New Delhi, 2001.

4Common Statistical Tests, Kulkarni M.B., Ghatpande S.B., Gore S.D., Satyajeet Prakashan,Pune, 1999.

	PAPER CODE: STC-1203 **PAPER –III: Statistics Practical Semester II** **[Credit -2: No. of Practicals 10]**
	Title of Experiment/ Practical
1	Multiple Regression, Multiple and Partial Correlation Coefficient
2	Non linear regression
3	Time Series
4	Fitting of Normal distribution
5	Large sample tests of hypothesis
6	Non parametric tests
7	Tests based on t , chi-square and F distribution
8	Computations of probabilities of Normal and Exponential distributions using R
9	Tests of hypothesis using R
10	Simulation using R