Deccan Education Society's

FERGUSSON COLLEGE, PUNE
(AUTONOMOUS)

SYLLABUS UNDER AUTONOMY

FIRST YEAR B.Sc.
SEMESTER - I

SYLLABUS FOR F.Y. B.Sc. STATISTICS

Academic Year 2016-2017

FERGUSSON COLLEGE (AUTONOMOUS), PUNE
DEPARTMENT OF STATISTICS

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

Semester	Paper Code	Title of Paper	No. of Credits
I	STS1101	Descriptive Statistics	2
	STS1102	Probability Theory	2
	STS1103	Statistics Practical - I	2
II	STS1201	Descriptive Statistics and Introduction to R - Software	2
	STS1202	Discrete Probability Distributions - I	2
	STS1203	Statistics Practical - II	2

PAPER CODE: STS1101
PAPER – I: DESCRIPTIVE STATISTICS
[Credit -2: No. of Lectures 36]

Objectives: The main objective of this course is to acquaint students with some basic concepts in Statistics. They will be introduced to some elementary statistical methods of analysis of data. At the end of this course students are expected to be able,

(i) to compute various measures of central tendency, dispersion, skewness and kurtosis.
(ii) to analyze data pertaining to attributes and to interpret the results.
(iii) to compute the correlation coefficient for bivariate data and interpret it.
(iv) to fit linear, quadratic and exponential curves to the bivariate data to investigate relation between two variables.
(v) to fit linear regression model to the bivariate data
(vi) to compute and interpret various index numbers.

	Title and Contents	**No. of Lectures**
Unit - I	**1. Introduction to Statistics** 1.1 Meaning of Statistics as a Science. 1.2 Importance of Statistics. 1.3 Scope of Statistics: In the field of Industry, Biological sciences, Medical sciences, Economics, Social Sciences, Management sciences, Agriculture, Insurance, Information technology, Education and Psychology. 1.4 Statistical organizations in India and their functions: CSO, ISI, NSS, IIPS (Devnar, Mumbai), Bureau of Economics and statistics.	**2**
Unit - II	**2. Population and Sample** 2.1 Types of characteristics: Attributes: Nominal scale, ordinal scale, Variables: Interval scale, ratio scale, discrete and continuous variables, difference between linear scale and circular scale 2.1 Types of data: (a) Primary data, Secondary data (b) Cross-sectional data, time series data, directional data. 2.2 Notion of a statistical population: Finite population, infinite population, homogeneous population and heterogeneous population. Notion of a sample and a random sample 2.3 Methods of sampling (Description only): Simple random sampling with and without replacement (SRSWR and SRSWOR), stratified random sampling, systematic sampling, cluster sampling and two-stage sampling.	**5**
Unit - III	**3. Summary Statistics.** 3.1 Classification: Raw data and its classification,	

	ungrouped frequency distribution, Sturges' rule, grouped frequency distribution, cumulative frequency distribution, inclusive and exclusive methods of classification, Open end classes, and relative frequency distribution. 3.2 **Measures of Central Tendency** Concept of central tendency of statistical data, Statistical averages, characteristics of a good statistical average. Arithmetic Mean (A.M.): Definition, effect of change of origin and scale, combined mean of a number of groups, merits and demerits, trimmed arithmetic mean. Mode and Median: Definition, merits and demerits. Empirical relation between mean, median and mode. Partition Values: Quartiles, Deciles and Percentiles, Box Plot. Geometric Mean (G.M.): Definition, merits and demerits. Harmonic Mean (H.M.): Definition , merits and demerits. Order relation between arithmetic mean, geometric mean, harmonic mean Weighted Mean: weighted A.M., G.M. and H.M. Situations where one kind of average is preferable to others. 3.3 **Measures of Dispersion** Concept of dispersion, characteristics of good measure of dispersion. Range, Semi-interquartile range (Quartile deviation): Definition, merits and demerits Mean deviation:Definition, merits and demerits, minimality property (without proof) Variance and standard deviation: Definition, merits and demerits, effect of change of origin and scale, combined variance for n groups(derivation for two groups). Mean squared deviation: Definition, minimality property of mean squared deviation , Measures of dispersion for comparison: Coefficient of range, coefficient of quartile deviation and	**12**

	coefficient of mean deviation, coefficient of variation (C.V.)	
Unit – IV	**4. Moments, Skewness and Kurtosis** 4.1 Moments: Raw moments (m'_r) for ungrouped and grouped data. Central moments (m_r) for ungrouped and grouped data, Effect of change of origin and scale. Relations between central moments and raw moments, upto 4-th order (without proof). 4.2 Concept of skewness of frequency distribution, positive skewness, negative skewness, symmetric frequency distribution. Bowley's coefficient of skewness : Bowley's coefficient of skewness lies between -1 to 1, interpretation using Box plot. Karl Pearson's coefficient of skewness. Measures of skewness based on moments (β_1, γ_1). 4.3 Concepts of kurtosis, Types of kurtosis, leptokurtic, mesokrtic and platykurtic frequency distributions. Measures of kurtosis based on moments (β_2, γ_2).	**9**
Unit –V	**5.Theory of Attributes** 5.1 Attributes: Concept of a Likert scale, classification, notion of manifold classification, dichotomy, class-frequency, order of a class, positive class- frequency, negative class frequency, ultimate class frequency, relationship among different class frequencies (up to three attributes), and dot operator to find the relation between frequencies, fundamental set of class frequencies. 5.2 Consistency of data upto 3 attributes. 5.3 Concepts of independence and association of two attributes. 5.4 Yule's coefficient of association (Q), $-1 \leq Q \leq 1$, interpretation.	**8**

References:

1. Agarwal, B. L. (2003). Programmed Statistics, Second Edition, New Age International Publishers, New Delhi.
2. Goon, A. M., Gupta, M. K. and Dasgupta, B. (1983). Fundamentals of Statistics, Vol. 1, Sixth Revised Edition,The World Press Pvt. Ltd., Calcutta.
3. Gupta, S. C. and Kapoor, V. K. (1983). Fundamentals of Mathematical Statistics, Eighth Edition, Sultan Chand and Sons Publishers, New Delhi.
4. Gupta, S. C. and Kapoor, V. K. (1997). Fundamentals of Applied Statistics, Third Edition, Sultan Chand and Sons Publishers, New Delhi.
5. Freund, J. E. (1977). Modern Elementary Statistics. Fourth Edition, Prentice Hall of India Private Limited, New Delhi.
6. Purohit, S. G., Gore S. D., Deshmukh S. R. (2008). Statistics Using R, Narosa

Publishing House, New Delhi.
7. Sarma, K. V. S. (2001). Statistics Made it Simple: Do it yourself on PC. Prentce Hall of India, New Delhi.
8. Snedecor G. W. and Cochran W. G.(1989). Statistical Methods, Eighth Ed. East-West Press.

PAPER CODE: STS1102
PAPER – II: PROBABILITY THEORY
[Credit -2: No. of Lectures 36]

Objectives

The main objective of this course is to introduce to the students the basic concepts of probability, axiomatic theory of probability, concept of random variable, probability distribution (univariate and bivariate) discrete random variables, expectation and moments of probability distribution. By the end of the course students are expected to be able

(i) to distinguish between random and non-random experiments.
(ii) to find the probabilities of events.
(iii) to obtain a probability distribution of random variable (one or two dimensional) in the given situation, and
(iv) to apply standard discrete probability distribution to different situations.

	Title and Contents	No. of Lectures
Unit -I	**1. Review of probability, conditional probability, independence** 1.1 Experiments/Models, Ideas of deterministic and non-deterministic models. Random Experiment, concept of statistical regularity. 1.2 Definitions of - (i) Sample space, (ii) Discrete sample space: finite and countably infinite, (iii) Event, (iv) Elementary event, (v) Complement of an event. (vi) Certain event (vii) Impossible event 1.3 Concept of occurrence of an event. 1.4 Algebra of events and its representation in set theory notation. Occurrence of following events. (i) at least one of the given events (ii) none of the given events, (iii) all of the given events, (iv) mutually exclusive events, (v) mutually exhaustive events, (vi) exactly one event out of the given events. 1.5 Classical definition of probability and its limitations. 1.6 Probability model, probability of an event, equiprobable and non-equiprobable sample space, 1.7 Axiomatic definition of probability. 1.8 Conditional probability of an event. 1.9 Independence of two events 1.10 Pairwise independence and mutual independence for three events	**6**

	1.11 Multiplication theorem $P(A \cap B) = P(A) \cdot P(B\|A)$. Generalization to $P(A \cap B \cap C)$. 1.12 Examples and Problems.	
Unit -II	**2. Bayes' Theorem** 2.1 Partition of the sample space 2.2 Proof of Bayes' theorem. Applications of Bayes' theorem in real life 2.3 Examples and Problems	**3**
Unit –III	**3.Univariate Probability Distributions** 3.1 Concept and definition of a discrete random variable. 3.2 Probability mass function (p.m.f.) and distribution function (d.f.), $F(\cdot)$ of discrete random variable, properties of distribution function. 3.3 Mode and median of a univariate discrete probability distribution. 3.4 Definition of expectation (Mean) of a random variable, expectation of a function of a random variable, moment generating function (m.g.f.) and cumulant generating function (c.g.f.) and their properties 3.5 Definitions of variance, standard deviation (s.d.) and Coefficient of variation (c.v.) of univariate probability distribution, effect of change of origin and scale on mean, variance and s.d. 3.6 Definition of raw, central moments and factorial raw moments upto order two of univariate probability distributions and their interrelations. 3.7 Coefficients of skewness and kurtosis based on moments. 3.8 Examples and Problems.	**12**
Unit –IV	**4. Discrete Probability Distributions** 4.1 Degenerate distribution (one point distribution), with pmf $P(X=c) = 1$, mean and variance. 4.2 Uniform discrete distribution on integers 1 to n: p.m.f., mean, variance 4.3 Bernoulli Distribution: p.m.f., mean, variance. 4.4 Binomial Distribution: p.m.f. $$P(X = x) = \begin{cases} \binom{n}{x} p^x q^{(n-x)}, & x = 0, 1, \ldots, n. \\ 0 & \text{otherwise.} \end{cases}$$	**15**

	Notation : $X \sim B(n, p)$. 4.5 Examples and Problems. Recurrence relation for successive probabilities, computation of probabilities of different events, mode of the distribution, mean, variance, m.g.f. and c.g.f. moments , skewness (comments when $p = 0.5$, $p > 0.5$, $p < 0.5$). Situations where this distribution is applicable. 4.6 Hypergeometric Distribution : p.m.f $P(X = x) = \begin{cases} \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} & \mathbf{x = 0, 1, \ldots \min\{ n , M \}} \\ \mathbf{0} & \mathbf{otherwise.} \end{cases}$ Notation : $X \sim H(N,M,n)$. Computation of probability, mean and variance, situations where this distribution is applicable, binomial approximation to hypergeometric distribution.	

References:

1. Agarwal B. L. (2003). Programmed Statistics, second edition, New Age International Publishers, New Delhi.
2. Gupta, S.C. and Kapoor, V. K. (1983). Fundamentals of Mathematical Statistics, Eighth Edition, Sultan Chand and Sons Publishers, New Delhi.
3. Hoel P. G. (1971). Introduction to Mathematical Statistics, John Wiley and Sons, New York.
4. Hogg, R. V. and Craig R. G. (1989). Introduction to Mathematical Statistics, Ed. 4. MacMillan Publishing Co., New York.
5. Mayer, P. (1972). Introductory Probability and Statistical Applications, Addison Wesley Publishing Co., London.
6. Mood, A. M. and Graybill, F. A. and Boes D.C. (1974). Introduction to the Theory of Statistics, Ed. 3, McGraw Hill Book Company.
7. Ross S. (2002). A First Course in Probability, Sixth Edition, Pearson Education, Inc. & Dorling Kindersley Publishing, Inc.

	PAPER CODE: STS1103 **PAPER – III: STATISTICS PRACTICAL - I** **[Credit -2: No. of Practicals 10]** **Pre-requisites**: Knowledge of the topics in theory papers I and II. **Objectives**: At the end of this course students are expected to be able (i) to use various graphical and diagrammatic techniques and interpret. (ii) to compute various measures of central tendency, dispersion, skewness and kurtosis, (iii) to compute correlation coefficient, regression coefficients, (iv) to fit binomial and Poisson distributions, (v) to analyse data pertaining to discrete and continuous variables and to interpret the results, (vi) to compute probabilities of bivariate distributions, (vii) to interpret summary statistics of computer output. (viii) to summarize and analyze the data using computer (ix) to draw random samples from Poisson and binomial distributions.
	Title of Experiment/ Practical
1	Diagrammatic representation of statistical data: simple and subdivided bar diagrams, multiple bar diagram, percentage bar diagram, pie diagram
2	Graphical representation of statistical data: histogram, frequency curve and ogive curves. Determination of mode and median graphically
3	Use of random number tables to draw SRSWOR, SRSWR, stratified sample and systematic sample
4	Computation of measures of central tendency and dispersion (ungrouped data). Use of an appropriate measure and interpretation of results and computation of partition values
5	Computation of measures of central tendency and dispersion (grouped data). Use of an appropriate measure and interpretation of results and computation of partition values
6	Measures of skewness and kurtosis, Box plot
7	Fitting of binomial distribution and computation of expected frequencies.
8	Application of binomial and hypergeometric distribution.
9	Graphical and diagrammatic representation of statistical data using MS - Excel
10	Computation of summary statistics using MS Excel

Note: Per practical number of lectures are 4.

Deccan Education Society's

FERGUSSON COLLEGE, PUNE
(AUTONOMOUS)

SYLLABUS UNDER AUTONOMY

FIRST YEAR B.Sc.
SEMESTER - II

SYLLABUS FOR F.Y. B.Sc. STATISTICS

Academic Year 2016-2017

PAPER CODE: STS1201
PAPER – I: DESCRIPTIVE STATISTICS AND INTRODUCTION TO R - SOFTWARE
[Credit -2: No. of Lectures 36]

	Title and Contents	No. of Lectures
Unit -I	**1. Correlation** 1.1 Bivariate data, Scatter diagram and interpretation. 1.2 Concept of correlation between two variables, positive correlation, negative correlation, no correlation. 1.3 Covariance between two variables (m_{11}) : Definition, computation, effect of change of origin and scale. 1.4 Karl Pearson's coefficient of correlation (r) : Definition, computation for ungrouped data and interpretation. Properties: (i) $-1 \leq r \leq 1$ (with proof), (ii) Effect of change of origin and scale (with proof). 1.5 Spearman's rank correlation coefficient: Definition, derivation of formula, computation and interpretation (without ties). In case of ties, compute Karl Pearson's correlation coefficient between ranks. (Spearman's rank correlation coefficient formula with correction for ties not expected.)	**10**
Unit -II	**2. Regression** 2.1 Meaning of regression, difference between correlation and regression, 2.2 Concept of error in regression, error modeled as a continuous random variable. Simple linear regression model: $Y= a + b X + \varepsilon$, where ε is a continuous random variable with $E(\varepsilon) =0$, $V(\varepsilon) = \sigma^2$. Estimation of a, b by the method of least squares. Interpretation of parameters. Statement of the estimator of σ^2. 2.3 Concept of residual, plot of residual against X, concept of coefficient of determination. 2.4 Fitting of second degree curve ($Y = a + b X + c X^2$), 2.5 Fitting of exponential curves of the type $Y = a b^X$ and $Y = aX^b$. In all these curves parameters are estimated by the method of least squares.	**14**
Unit –III	**3. Index Numbers** 3.1 Introduction.	

	3.2 Definition and Meaning. 3.3 Problems/considerations in the construction of index numbers. 3.4 Simple and weighted price index numbers based on price relatives. 3.5 Simple and weighted price index numbers based on aggregates. 3.6 Laspeyre's, Paasche's and Fisher's Index numbers 3.7 Consumer price index number: Considerations in its construction. 3.8 Methods of construction of consumer price index number - (i) family budget method (ii) aggregate expenditure method. 3.9 Shifting of base, splicing, deflating, purchasing power. 3.10 Description of the BSE sensitivity and similar index numbers.	**8**
Unit –IV	**4. Fundamentals of R-Software:** **4.1** Introduction to R, features of R, starting and ending R session, getting help in R, R commands and case sensitivity. 4.2 Vectors and vector arithmetic a) creation of vectors using functions c, seq, rep b) Arithmetic operations on vectors using operators +, - , * , / , ^ . c) Numerical functions: log10, log, sort, max, min, unique, range, length, var, prod, sum, summary, fivenum etc. d) accessing vectors 4.3 Data frames : creation using data.frame, subset and transform commands. 4.4 Resident data sets : Accession and summary p, q, d, r functions.	**4**

References:

1 Agarwal, B. L. (2003). Programmed Statistics, Second Edition, New Age International Publishers, New Delhi.

2 Goon, A. M., Gupta, M. K. and Dasgupta, B. (1983). Fundamentals of Statistics, Vol. 1, Sixth Revised Edition,The World Press Pvt. Ltd., Calcutta.

3 Gupta, S. C. and Kapoor, V. K. (1983). Fundamentals of Mathematical Statistics, Eighth Edition, Sultan Chand and Sons Publishers, New Delhi.

4 Gupta, S. C. and Kapoor, V. K. (1997). Fundamentals of Applied Statistics, Third Edition, Sultan Chand and Sons Publishers, New Delhi.

5 Freund, J. E. (1977). Modern Elementary Statistics. Fourth Edition, Prentice Hall of India Private Limited, New Delhi.

6 Purohit, S. G., Gore S. D., Deshmukh S. R. (2008). Statistics Using R, Narosa Publishing House, New Delhi.

7 Sarma, K. V. S. (2001). Statistics Made it Simple: Do it yourself on PC. Prentce Hall of India, New Delhi.
8 Snedecor G. W. and Cochran W. G.(1989). Statistical Methods, Eighth Ed. East- West Press.

PAPER CODE: STS1202

PAPER – II: DISCRETE PROBABILITY DISTRIBUTIONS - I

[Credit - 2: No. of Lectures 36]

	Title and Contents	**No. of Lectures**
Unit - I	**1. Discrete Probability Distributions** 1.1 **Poisson distribution**: Notation : $X \sim P(\lambda)$. $p(x) = \frac{e^{-\lambda}\lambda^x}{x!}$, $x = 0,1,2 \ldots\ldots$, $\lambda > 0$ $= 0$ otherwise Mean, variance, and m.g.f, c.g.f. moments , skewness , kurtosis Situations where this distribution is applicable. 1.2 **Geometric distribution**: Notation: $X \sim G(p)$, Geometric distribution on support (0, 1, 2…) with p.m.f. $p(x) = pq^x$, $x = 0, 1, 2\ldots$ $= 0$ otherwise Geometric distribution on support (1, 2,…….) with p.m.f. $p(x) = pq^{x-1}$, $x = 1, 2\ldots$ $0 < p < 1$, $q = 1 - p$. $= 0$ otherwise Mean, variance, m.g.f, moments , lack of memory property. Situations where this distribution is applicable. 1.3 **Negative Binomial Distribution**: p. m. f. : $P(X=x) = \binom{x+k-1}{x} p^k q^x$, $x = 0, 1, 2, 3 \ldots\ldots$, $0 \le p \le 1$, $q = 1-p$ $= 0$ otherwise Notation**:** $X \sim NB(k, p)$. Nature of p. m. f., negative binomial distribution as a waiting time distribution, Mean, variance, m.g.f, c.g.f, moments , skewness , and kurtosis. (recurrence relation between moments is not expected). Poisson approximation to negative binomial distribution. Real life situations.	**18**

	1.4 Examples and Problems.	
Unit -II	**2. Bivariate Discrete Probability Distribution** 2.1 Definition of two-dimensional discrete random variable, its joint p.m.f. and its distribution function and their properties, concept of identically distributed r.vs. 2.2 Computation of probabilities of events in bivariate probability distribution. 2.3 Concepts of marginal and conditional probability distributions. 2.4 Independence of two discrete random variables based on joint and marginal p.m.f.s 2.5 Definition of raw and central moments 2.6 Theorems on expectations of sum and product of two jointly distributed random variables. 2.7 Conditional expectation. 2.8 Definitions of conditional mean and conditional variance. 2.9 Definition of covariance, coefficient of correlation, independence and uncorrelation. 2.10 Variance of linear combination of variables i.e. Var(aX + bY). 2.11 Additive property of binomial and Poisson distributions Conditional probability distribution of X given (X+Y) for independent Poisson variables. negative binomial distribution as sum of k i.i.d.geometric random variables. Relation between geometric and negative binomial distribution. 2.12 Examples and Problems.	**18**

References:

1 Agarwal B. L. (2003). Programmed Statistics, second edition, New Age International Publishers, New Delhi.

2 Gupta, S.C. and Kapoor, V. K. (1983). Fundamentals of Mathematical Statistics,
Eighth Edition, Sultan Chand and Sons Publishers, New Delhi.

3 Hoel P. G. (1971). Introduction to Mathematical Statistics, John Wiley and Sons, New York.

4 Hogg, R. V. and Craig R. G. (1989). Introduction to Mathematical Statistics, Ed.
MacMillan Publishing Co., New York.

5 Mayer, P. (1972). Introductory Probability and Statistical Applications,

Addison
Wesley Publishing Co., London.

6 Mood, A. M. and Graybill, F. A. and Boes D.C. (1974). Introduction to the Theory
of Statistics, Ed. 3, McGraw Hill Book Company.

7 Ross S. (2002). A First Course in Probability, Sixth Edition, Pearson Education, Inc. & Dorling Kindersley Publishing, Inc.

	PAPER CODE: STS1203 **PAPER – III: STATISTICS PRACTICAL - II** **[Credit -2: No. of Practicals 10]** **Pre-requisites**: Knowledge of the topics in theory papers I and II. **Objectives**: At the end of this course students are expected to be able (i) to use various graphical and diagrammatic techniques and interpret. (ii) to compute various measures of central tendency, dispersion, skewness and kurtosis, (iii) to compute correlation coefficient, regression coefficients, (iv) to fit binomial and Poisson distributions, (v) to analyse data pertaining to discrete and continuous variables and to interpret the results, (vi) to compute probabilities of bivariate distributions, (vii) to interpret summary statistics of computer output. (viii) to summarize and analyze the data using computer (ix) to draw random samples from Poisson and binomial distributions.
	Title of Experiment / Practical
1	Scatter diagram, correlation coefficient (ungrouped data), fitting of line of regression, residual plot
2	Scatter diagram, correlation coefficient, fitting a line of regression, fitting of second degree curve, using MS Excel
3	Fitting of second degree curve, exponential curve of type $y = ab^x$, $y = ax^b$, comparison, finding the best fit using residual s.s. and coefficient of determination
4	Index numbers
5	Fitting of Poisson distribution and computation of expected frequencies.
6	Applications of Poisson and geometric distributions
7	Model sampling from poisson and binomial distributions.
8	Bivariate Probability Distributions
9	Use of R- Software and computation of probabilities using R- Software
10	Fitting of binomial and Poisson distribution and computation of expected frequencies using R- Software

Note: Per practical number of lectures are 4.