Mangalore University

Syllabus of

STATISTICS

(An optional subject for B.Sc. Degree Programmme (CBCS)

(With effect from 2019-20 onwards)

Mangalore University

Revised syllabus of STATISTICS, an optional subject for B.Sc. Degree Program

PREAMBLE

The last revision of the syllabus of Statistics, an optional subject for the three years B.Sc. degree program was undertaken in the year 2013-14. Statistics is vigorously gaining importance and recognition in today's society. For a long time, Statistics was identified solely with the displays of data and charts pertaining to the economic, demographic, and political situations prevailing in a country. However, gigantic advances during the twentieth century have enabled statistics to grow and assume its present importance as a discipline of data-based reasoning. Statistics, as a scientific discipline, stretch well beyond the confines of data display. It deals with collecting informative data, analysis interpreting these data, and drawing conclusions about a phenomenon under study. The scope of this subject naturally extends to all processes of acquiring knowledge that involve fact finding through collection and examination of data. Also, Statistics plays vital role in Big data& Analytics, which is now an emerging field in all spheres of human activity. There is a great demand for data analytics.

The University Grants Commission, New Delhi, has directed all Universities in the country to implement the Choice Based Credit System (CBCS, Semester Scheme) in both the under –graduate and post-graduate programmes. Mangalore University considered the feasibility of CBCS at several levels and through the meeting of several of its statutory bodies. After due consideration, Mangalore University decided to adopt the CBCS so that the acceptability of the programmes and degree offered by the University becomes comparable and readily acceptable.

Accordingly as per CBCS Semester scheme the board designed the curriculum of B.Sc. Statistics keeping in view that the curriculum of B.Sc. Statistics should be flexible enough to prepare graduates to either be functioning statisticians or to go on to post-graduation studies.

The following changes were made in the curriculum.

1) The Programme of study consists of Core Courses and Elective courses.

2) Core courses will consist of compulsory/Optional courses which are compulsorily studied by a student in all the six semesters. Elective courses are opted in I to IV Semesters.

3) Introduced a unit on Statistical computing using R software in each semester.

4) Introduced scaling techniques which are useful in the measurements of qualitative characteristics in social science research.

5) Introduced various multivariate data analysis methods (only description of methods).

6) Paper VIII of sixth semester deals with the application of statistical methods in fields such as Industrial quality control, Economics and Demography. This enables the students to understand and appreciate the applications of statistical methods in various fields.

ELIGIBILITY

1. Only those candidates who have passed Pre-University Course in Science or an equivalent course with Mathematics as one of the subject are eligible to take Statistics as one of the optional subjects in B.Sc. Course.

2. Any student taking Statistics as one of the optional subjects in the B.Sc. course shall take Mathematics as another optional subject.

Scheme of Instruction/Examination:

The board recommends that the existing scheme of instruction and examination in theory and practical may be continued. The theory question paper in all semesters shall cover all the topics in the pertaining syllabus with proportional weightage to the number of hours of instruction prescribed.

Theory question paper pattern

	No. of questions to be answered	Marks
Part A	Ten questions out of 12 questions	10 X 2 = 20
Part B	Five questions out of 8 questions	5 X 6 = 30
Part C	Three questions out of 5 questions Each question may have sub questions	3 X 10 = 30

Practical Examination:

1) Students will have to answer any four questions out of 8 questions. All questions carry equal marks.

2) Practical paper in each semester carries 50 marks and the split up of the practical marks are as given below.

a) Three hour examination with two examiners which includes at least one external examiner carries 30 marks.

b) Ten marks for class records.

c) Five marks each for the preparatory practical examination using calculator and for statistical computing using R software. In the semesters where R language is not discussed, ten marks to be allotted for the preparatory practical examination.

MANGALORE UNIVERSITY
CHOICE BASED CREDIT SYSTEM
STATISTICS
SCHEME

	Course Code	Particulars	Instruction hours/ week	Duration of the exam (hrs)	Max. Marks			Credits
					IA	Exam	Total	
I SEMESTER								
Group I Core Subject	BSCSTC 131	Descriptive Statistics	4	3	20	80	100	2
	BSCSTP 132	Statistics Practical - I	3	3	10	40	50	1
Group II Elective	BSCSTCE 133	Official Statistics	2	2	10	40	50	1
					Total number of Credits for Core Subject in I Semester: 04			
II SEMESTER								
Group I Core Subject	BSCSTC 181	Probability Theory	4	3	20	80	100	2
	BSCSTP-182	Statistics Practical - II	3	3	10	40	50	1
Group II Elective	BSCSTCE 183	Statistical Analysis of Experimental Data.	2	2	10	40	50	1
					Total number of Credits for Core Subject in II Semester: 04			
III SEMESTER								
Group I Core Subject	BSCSTC 231	Sampling Theory	4	3	20	80	100	2
	BSCSTP 232	Statistics Practical - III	3	3	10	40	50	1
Group II Elective	BSCSTCE 233	Introduction to Data Science and Big Data Analytics.	2	2	10	40	50	1
					Total number of Credits for Core Subject in III Semester: 04			
IV SEMESTER								
Group I Core Subject	BSCSTC 281	Statistical Inference – I	4	3	20	80	100	2
	BSCSTP-282	Statistics Practical - IV	3	3	10	40	50	1
Group II Elective	BSCSTOE 283	Basic Statistics	2	2	10	40	50	1
					Total number of Credits for Core Subject in IV Semester: 04			

	Course Code	Particulars	Instruction hours/ week	Duration of the exam (hrs)	Max. Marks			Credits
					IA	Exam	Total	
V SEMESTER								
Group I Core Subject	BSCSTC 331	Regression Analysis	3	3	20	80	100	2
	BSCSTC 332	Operations Research	3	3	20	80	100	2
	BSCSTP 333	Practical based on BSCSTC 331 & BSCSTC 332	4	3	20	80	100	2
Total number of Credits for Core Subject in V Semester: 06								
VI SEMESTER								
Group I Core Subject	BSCSTC 381	Statistical Inference II & Design and Analysis of experiments	3	3	20	80	100	2
	BSCSTC 382	Applied Statistics	3	3	20	80	100	2
	BSCSTP 383	Practical based on BSCSTC 381 & BSCSTC 382	4	3	20	80	100	2
Total number of Credits for Core Subject in VI Semester: 06								
Total number of Credits for Core Subject in I-VI Semesters: 28								

FIRST SEMESTER

BSCSTC 131: Descriptive Statistics (48 Lecture hours) (2 Credits)

UNIT I: Organisation and Presentation of Data (8L)

Meaning of Statistics as a Science, Importance of Statistics

Types of characteristics: Attributes - Nominal scale, ordinal scale, Variables - Interval scale, ratio scale, discrete and continuous variables,

Types of data: (a) Primary data, Secondary data. (b) Cross-sectional data, time series data,

Classification and Tabulation, Construction of frequency distribution, Graphical representation- Histogram, frequency polygon, frequency curve, ogives

Notion of a statistical population - Finite population, infinite population, homogeneous population and heterogeneous population

Notion of sample - random sample and non-random sample

Statistical organizations in India and their functions – MOSPI, CSO, NSSO, ISI, IIPS (Devnar, Mumbai), Bureau of Economics and statistics.

UNIT II – Univariate data analysis - Measures of central tendency (10L)

Concept of central tendency of statistical data: Statistical average, characteristics of a good statistical average Arithmetic Mean (A.M.), properties, trimmed arithmetic mean,

Median - Definition, formula for computation (with derivation) graphical method of determination of median, merits and demerits, Mode - Definition, formula for computation (with derivation) graphical method of determination of mode, merits and demerits

Empirical relation between mean, median and mode

Partition Values - Quartiles, Deciles and Percentiles, Percentile ranks,

Means of transformed data - 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 (proof for n = 2). Weighted Mean - Weighted A.M., G.M. and H.M. Situations where one kind of average is preferred to others. Examples

UNIT III - Univariate data analysis - Measures of Dispersion (10L)

Measures of Dispersion - Concept of dispersion, characteristics of good measure of dispersion. Range- Definition, merits and demerits. Semi-interquartile range (Quartile deviation), Mean deviation- Definition, merits and demerits, minimality property (without proof). Mean square deviation: Definition, minimal property of mean square deviation (with proof), Variance and standard deviation: Definition, merits and demerits, properties Measures

of dispersion for comparison: coefficient of range, coefficient of quartile deviation and coefficient of mean deviation, coefficient of variation (C.V.) Examples

UNIT IV – Moments, skewness and kurtosis (10L)

Moments- Raw moments for grouped and ungrouped data, Moments about an arbitrary constant for grouped and ungrouped data, Central moments for grouped and ungrouped data, Effect of change of origin and scale, Relations between central moments and raw moments

Skewness and Kurtosis - Concept of skewness of frequency distribution, positive skewness, negative skewness, symmetric frequency distribution, Interpretation with examples, Bowley's coefficient of skewness, Karl Pearson's coefficient of skewness. Measures of skewness based on moments. Kurtosis, leptokurtic, mesokurtic and platykurtic frequency distributions. Measures of kurtosis based on moments, Examples

UNIT V - Statistical computing (R software) (10L)

Introduction to R, R as a calculator, statistical software and a programming language, R preliminaries, getting help, data inputting methods(direct and importing from other spread sheet applications like Excel), data accessing, and indexing, packages, Graphics in R, built in functions, saving, storing and retrieving work. Descriptive statistics:, diagrammatic representation of univariate and bivariate data (box plots, stem and leaf diagrams, bar plots, pie diagram), measures of central tendency (mean, median and mode), partition values, measures of dispersion (range, standard deviation, mean deviation and inter quartile range), summaries of a numerical data, skewness and kurtosis,

References

1) Goon, A. M., Gupta, M. K., & Das Gupta, B. (1991). *Fundamentals of Statistics, Vol. 1.* World Press, Calcutta.

2) Gupta, S. C., & Kapoor, V. K. (2002). *Fundamental of Mathematical Statistics.* Sultan Chand & sons.

3) Medhi, J. (2006). *Statistical Methods: An Introductory Text.* New Age International(P) Limited, New Delhi.

4) Robert Kabacoff, I. (2015). *R in Action - Data Analysis and Graphics with R, second edition.* dreamtech Press.

5) Sudha Purohit, G., Sharad Gore, D., & Shailaja Deshmukh, R. (2008). *Statistics Using R.* Narosa Publishing House.

FIRST SEMESTER

BSCSTP 132: Statistics Practical – I -Descriptive Statistics (1 credit)

List of Practical Assignments

1) Graphical presentation of Statistical Data
2) Measures of central tendency-1
3) Measures of central tendency -2
4) Measures of dispersion-1
5) Measures of dispersion-2
6) Computation of central moments
7) Measures of skewness
8) Measures of kurtosis
9) Statistical Computing using R software - Descriptive statistics

SECOND SEMESTER

BSCSTC 181: Probability Theory (48 Lecture hours) (2 Credits)

UNIT I – Probability (14L)

Random Experiments, Sample space (Discreet and continuous), Elementary events and compound events, Algebra of events, Classical definition of probability and its limitations, relative frequency approach. Axioms of probability, Deduction of classical definition from Axiomatic definition.   Addition theorem, Conditional Probability and Independence, Multiplication theorem, Bayes' Theorem (with proof) and its applications

UNIT II - Random Variables (12L)

Random variables (discrete and continuous), properties, probability mass function, probability density function – their properties, Distribution function – Properties, Bivariate p.m.f.'s and p.d.f.'s , marginal and conditional probability distributions for two r.v.'s, Independence of random variables. Expectation of random variables – Rules of expectation, addition and multiplication theorems of expectation, variance and covariance. Mean and covariance of linear combination of random variables, moments, measures of location and dispersion – skewness and kurtosis of a random variable Expectation of functions of random variables, M.G.F. and its properties, cumulant generating function, cumulants.

UNIT III – Standard Probability distributions

Discrete distributions (9L)

Bernoulli, Binomial, Poisson, discreet uniform, Geometric, Negative Binomial, Hyper geometric distributions- definition through pmf. – Mean, Variance, MGF, CGF, Recurrence relation for moments, Properties, Interrelationships, Applications, Limiting Distributions

Unit IV - Continuous univariate distributions(8L)

Uniform, exponential, Gamma, Normal Distributions, Cauchy – definition through p.d.f – Distribution function – computation of moments, Median, Mode, MGF

UNIT V – Statistical Computing (5L)

Probability Distributions: R as a set of statistical tables- cumulative distribution, probability density function, quantile function, and simulate from the distribution, plotting probability curves for standard distributions**.**

References

1) Goon, A. M., Gupta, M. K., & Das Gupta, B. (1991). *Fundamentals of Statistics, Vol. 1.* World Press, Calcutta.

2) Gupta, S. C., & Kapoor, V. K. (2002). *Fundamental of Mathematical Statistics.* Sultan Chand & sons.

3) Hogg, R. V., & Craig, A. T. (1995). *Introduction to Mathematical Statistics.* 5/e, Prentice Hall, New Jersy, USA.

4) Levin, R. (2013). *Statistics for Management.* Prentice Hall India.

5) Medhi, J. (2006). *Statistical Methods: An Introductory Text.* New Age International(P) Limited, New delhi.

6) Robert Kabacoff, I. (2015). *R in Action - Data Analysis and Graphics with R, second edition.* dreamtech Press.

7) Rohatgi, V. K. (2002). *An Introduction to Probability theory and Mathematical Statistics.* Wiley Eastern Limited.

8) Ross, S. M. (2003). *Introduction to Probability Models.* 10e, Academic Press, UK.

9) Sudha Purohit, G., Sharad Gore, D., & Shailaja Deshmukh, R. (2008). *Statistics Using R.* Narosa Publishing House.

SECOND SEMESTER

BSCSTP 182: Statistics Practical – II – Probability Theory (1 credit)

List of Practical Assignments

1) Probability- Addition theorem, multiplication theorem and conditional probability

2) Probability- Bayes' theorem.

3) Random variables (Univariate), pmf, pdf and Distribution functions, mean and variance

4) Random variables (Bivariate), conditional and marginal distributions, covariance

5) Exercise on Binomial Distribution

6) Exercise on Poisson Distribution

7) Exercise on geometric, Negative Binomial, Hyper geometric distribution

8) Exercise on Normal Distributions

9) Statistical Computing using R software – Probability Distributions

THIRD SEMESTER

BSCSTC 231: Sampling Theory (48 Lecture hours) (2 Credits)

UNIT I – Statistical Investigation (6L)

Population and Sample.

Complete enumeration v/s sample surveys - merits and demerits. Need for sampling, random and non-random sampling, limitations of non-random sampling and judgment sampling, Principles of sample surveys. Principal steps in a sample survey, Errors in sampling, parameter and statistic, Unbiasedness, variance and precision of estimators, pilot survey, determination of sample size, Selection of sample using random numbers. Drawing samples from finite population with and without replacement. Sampling from frequency distribution and contingency tables.

UNIT II - Simple random sample (10L)

Simple random sampling with replacement (SRSWR) and simple random sampling without replacement (SRSWOR), Unbiased estimators of mean, variance and population total, Sampling variances, standard errors and their estimation, comparison of SRSWR with SRSWOR.

Sampling of attributes - Sampling for proportions, Estimation of population proportion and its Standard Error (SE).

UNIT III - Stratified random sampling (12L)

Need for stratification, Stratified sampling under SRSWR and SRSWOR, Unbiased estimators of mean and total, Variances of these estimators and their estimation. Allocation of sample size - proportional and optimum allocation (w.r.t. SRSWOR stratification only), Neyman's allocation, allocation with cost functions, Comparison of SRSWOR and stratified sampling. Gain in efficiency due to stratification.

Linear systematic sampling

Estimation of mean, Variance of the estimator of mean in terms of intra-class correlation, Comparison of SRSWOR and linear systematic sampling.

UNIT IV – Transformation of random variables and Sampling Distributions (12L)

Definition and derivation of students' t, Chi-square and F- distributions – their properties, mean and variance. Distribution of sample mean, sample variance under normality assumption, Distribution of $\frac{ns^2}{\sigma^2}$ using the independence of $\bar{x}$ and s^2 when sampling from normal population. Interrelationship between the distributions, Distribution of minimum and maximum.

UNIT V- Limit theorems (8L)

Markov's inequality, statement and proof of Chebychev's inequality, sequence of random variables, convergence in probability: basic results (without proof), Weak law of large numbers, central limit theorem for i. i. d. random variables and its application

References

1) Cochran, W. G., & Cos. (1977). *Sampling Techniques, 3rd Edition.* Wiley Eastern.

2) Goon, A. M., Gupta, M. K., & Das Gupta, B. (1991). *Fundamentals of Statistics, Vol. 1.* World Press, Calcutta.

3) Gupta, S. C., & Kapoor, V. K. (2002). *Fundamental of Mathematical Statistics.* Sultan Chand & sons.

4) Guptha, S. C., & Kapoor, V. K. (1984). *Fundamentals of Applied Statistics.* Sulthan Chand and Sons.

5) Hogg, R. V., & Craig, A. T. (1995). *Introduction to Mathematical Statistics.* 5/e, Prentice Hall, New Jersy, USA.

6) Rohatgi, V. K. (2002). *An Introduction to Probability theory and Mathematical Statistics.* Wiley Eastern Limited.

7) Sampath, S. (2006). *Sampling Theory and Methods.* 2/e, Narosa, New Delhi.

8). Mukhopadhyay P. (1998): Theory and Methods of Survey Sampling, Prentice-Hall of India

THIRD SEMESTER

BSCSTP 232: Statistics Practical – III – Sampling Theory (1 credit)

List of Practical Assignments

1) Drawing simple random samples using random numbers and from a frequency distribution and contingency table

2) Exercise on SRSWR

3) Exercise on SRSWOR and sampling of attributes

4) Exercise on Stratified random sampling.

5) Exercise on Allocation of samples in Stratified sampling.

6) Linear systematic sampling.

7) Exercise on Chebychev's inequality inequality.

8) Exercise on Central limit theorems.

FOURTH SEMESTER

BSCSTC 281: Statistical Inference – I (48 Lecture hours) (2 Credits)

Unit I - Point Estimation (12L)

Estimator and estimate, Unbiasedness, and consistency of estimators, Sufficient condition for consistency, Relative efficiency, Sufficient statistic. Factorization theorem and its applications, Fisher Information function, Cramer Rao Inequality (statement only), Minimum variance unbiased estimator, minimum variance bound estimator, Exercises

Maximum likelihood and moment methods of moments estimation-properties of these methods (without proof)- (discussion of examples to be restricted to the standard distributions studied during 2nd semester.)

UNIT II - Interval estimation (6L)

Confidence coefficient, confidence interval using Pivotal Quantity method with examples. Confidence interval for mean, difference between means, variance and ratio of variances under normality. Large sample confidence interval for mean, difference of means. Proportion and difference between proportions.

Unit III - Testing of Hypotheses (12L)

Statistical Hypotheses -Null and alternative, Simple and composite hypotheses. Critical region. Concepts of type I and type II errors, level of significance and p-value, power of test. Power function -power curve. Relationship between testing of hypothesis and interval estimation. Most powerful - test, and best critical region. Statement of Neyman and Pearson Lemma and its use.

UNIT IV - Likelihood ratio tests (LRT) (10L)

Derivation of tests for normal distribution only with testing for mean and variance - one sample and two sample tests - two sided and one sided alternatives, Paired t test. Test for significance of correlation coefficient, properties of LRT (without proof). Large Sample test and Chi-square tests.

UNIT – V Statistical Computing (8L)

Statistical Inference: classical tests: One- and two-sample tests, , t-test, F-test under normal assumption, chi-square test of independence and goodness of fit, interval estimation for mean, difference of mean and variance, tests for normality (shapiro-wilks test, and q-q plot),

References

1) Gupta, S. C., & Kapoor, V. K. (2002). *Fundamental of Mathematical Statistics.* Sultan Chand & sons.

2) Hogg, R. V., & Craig, A. T. (1995). *Introduction to Mathematical Statistics.* 5/e, Prentice Hall, New Jersy, USA.

3) John freund, E. (2001). *Mathematical statistics.* Prentice hall of India, New Delhi.

4) Levin, R. (2013). *Statistics for Management.* Prentice Hall India.

5) Mood, A. M., Graybill, A., & Boes, C. (2001). *Introduction to the theory of Statistics.* Tata McGraw-Hill Publishing company Limited.

6) Robert Kabacoff, I. (2015). *R in Action - Data Analysis and Graphics with R, second edition.* dreamtech Press.

7) Ross, S. M. (2003). *Introduction to Probability Models.* 10e, Academic Press, UK.

8) Sudha Purohit, G., Sharad Gore, D., & Shailaja Deshmukh, R. (2008). *Statistics Using R.* Narosa Publishing House.

FOURTH SEMESTER

BSCSTP 282: Statistics Practical – IV – Statistical Inference – I (1 credit)

List of Practical Assignments

1) Point Estimation -1
2) Point Estimation -2
3) Exercise on size and power
4) Large sample tests and confidence intervals-1
5) Large sample tests and confidence intervals-2
6) 't' test for testing mean, difference of means, paired t test and small sample confidence interval for mean and difference of means.
7) Chi square test for testing variance of a normal population, confidence interval for variance, testing goodness of fit, and testing the independence of attributes.
8) F test for testing the ratio of variances of normal population and confidence interval for the ratio of variances
9) Statistical Computing using R software – Statistical inference - I

FIFTH SEMESTER

BSCSTC 331: Regression Analysis (48 Lecture Hours) (2 Credits)

UNIT I - Scaling Techniques (07L)

Measurement statistical survey, Sources of error in measurement, Tests of measurement

Scaling – Meaning of scaling, important scaling techniques, Scale construction techniques

Guttman scale - Likert scale, selection of appropriate scale- limitations of scaling techniques

UNIT II - Bivariate Data analysis: Correlation (07L)

Bivariate data, bivariate frequency distribution , Concept of correlation between two variables, types of correlation, Scatter diagram, Covariance between two variables - Definition, computation, effect of change of origin and scale. Karl Pearson's coefficient of correlation, Spearman's rank correlation coefficient, Testing the significance of correlation

Unit III - Regression Analysis (09L)

Concept of errors in functional relationship, simple linear regression model. Assumptions, Estimation of model parameters, least squares method, Interpretation of Regression coefficient properties, residuals, Mean residual sum of squares, estimation of error variance, Residual plot and its interpretation, Explained and unexplained variation, coefficient of determination, testing the significance of regression coefficients. Fitting of second degree curve, exponential curve by the method of least squares after logarithmic transformation

UNIT IV - Multivariate data analysis (15L)

Multiple linear regression (3 variables only), Assumption, Estimation of Regression by least squares method. Estimation of regression coefficients. Partial, multiple correlation coefficients. Coefficient of Determination (R^2)

Multivariate Data Analysis - (Description of various multivariate methods to be given)

Logistic regression, Factor Analysis, Structural Equation Modelling, Cluster Analysis, Discriminant Analysis, conjoint analysis, Correspondence Analysis

UNIT V - Statistical Computing (10L)

Packages, GGplot2 package, likert package, correlation and regression analysis (bivariate and multivariate data), polynomial regression

References

1) Gupta, S. C., & Kapoor, V. K. (2002). *Fundamental of Mathematical Statistics.* Sultan Chand & sons.

2) Joseph Hair, F., Willium Black, C., Barry Babin, J., & Rolph Anderson, E. (2010). *Multivariate Data Analysis, Seventh edition.* Pearson Prentice Hall.

3) Kothari, C. R. (2004). *Research methodology.* New Age International Publishers.

4) Levin, R. (2013). *Statistics for Management.* Prentice Hall India.

5) Medhi, J. (2006). *Statistical Methods: An Introductory Text.* New Age International(P) Limited, New delhi.

6) Montgomery, D. C. (2007). *Introduction to Linear Regression analysis.* John Willey & sons.

7) Mukhopadhyay, P. (2000). *Mathematical Statistics.* Books & Allied Pvt. Ltd.

8) Robert Kabacoff, I. (2015). *R in Action - Data Analysis and Graphics with R, second edition.* dreamtech Press.

9) Sudha Purohit, G., Sharad Gore, D., & Shailaja Deshmukh, R. (2008). *Statistics Using R.* Narosa Publishing House.

FIFTH SEMESTER

BSCSTP 333: Statistics Practical – V – Regression Analysis (1 credit)

List of Practical Assignments

1) Exercise on Correlation
2) Spearman's rank correlation coefficient.
3) Simple linear regression
4) Multiple linear regression- 1
5) Multiple linear regression -2
6) Testing the significance of correlation coefficient and equality of correlation coefficients.
7) Testing the significance of regression coefficients. Coefficient of determination, Standard Error of Regression, ANOVA.
8) Fitting of quadratic curve and exponential curve by the method of least squares
9) Statistical Computing using R software – Regression analysis

FIFTH SEMESTER

BSCSTC 332: Operations Research (48 Lecture Hours) (2 Credits)

Unit I (14L)

Operations Research (OR): Origin, definition, phases of OR- types of models.

Linear Programming Problem (LPP): General model, formulation, graphical solution, and standard form of LPP. Simplex algorithm (without proof), Chame'sbig M method - indication of unique solution, multiple solution, unbounded solution, no solution and degeneracy, dual LPP and its properties.

Unit II(07L)

Transportation Problem: General description and Statement of T.P. , Balanced and unbalanced T.P. - initial solution by north west corner rule , matrix minima and Vogel's method, MODI method of optimization (without proof), degeneracy.

Unit III (06L)

Assignment Problem: General description and Statement of assignment problem , Hungarian method, maximization and minimization problems.

Unit IV (06L)

Game theory: Nature of games, Two person zero sum games -Pay off Matrix

- Pure and mixed strategies- Showing A's Problem as the Dual of the B's problem -Principles of Dominance -Algebraic solution of rectangular games (Zero sum) - solution of 2 x 2 games - Graphic solution of 2 x n and m x 2 games.

Unit V (15L)

Inventory Theory: Basic concepts, deterministic models with instantaneous / finite production with / without shortages (derivations for continuous case), Models with one and two price breaks, stochastic model -single period problem. Newspaper boy problem

References

1) Kalavathy, S. K.(2013). *Operations Research.* Vikas Publishing House Pvt. Ltd.

2) Kanthiswarup, Manmohan, & Guptha, P. K. (2002). *Operations Research.* Sulthan Chand & Co.

3) Sharma, S. D. (2009). *Operations Research.* Kedar Nath Ram Nath.

4) Taha, H. A. (2002). *Operations Research, An Introduction.* Mac Millan.

FIFTH SEMESTER

BSCSTP 333: Statistics Practical – VI – Operations Research (1 credit)

List of Practical Assignments

1) Formulation of linear programming problem and its solution for two variables
2) Simplex method of solving linear programming problems.
3) Charne's Big M method of solving linear programming problems.
4) Transportation problems.
5) Assignment problems.
6) Exercise on game theory.
7) Inventory problems(deterministic models)
8) Inventory problems (Stochastic model)

SIXTH SEMESTER

BSCSTC 381: Statistical Inference II & Design and Analysis of Experiments

(48 Lecture Hours) (2 Credits)

Unit I (12L)

Sequential   testing: Need for sequential test. Wald's sequential probability ratio test (SPRT) Strength of sequential tests. Wald's SPRT applied to Bernoulli, Poisson and Normal distributions. Expressions to constants to be given without proof.

Non-parametric tests: Advantages and limitations. Sign test for one sample problem and for pairs of observations. Two sample median test. Run test for randomness. Two sample run test. Null distribution of test statistic to be derived in each case, Large sample approximation to these tests.

Unit II(11L)

Analysis of variance: Meaning, basic assumptions, fixed effect model. Analysis of one way, two way and three way classified data with one observation per cell -mathematical model. Least square estimates, splitting of total sum of squares, expectation of sums of squares and mean sums of squares (under appropriate hypotheses) and ANOVA table**.**

Unit III (14L)

Designs of experiments - Meaning and terminology - experiment, 'treatment, experimental unit, experimental error and precision. Basic Principles of experimental design, Randomisation, Replication, Local Control

CRD, RBD and LSD: Layout model, splitting of Total variati in to che to different components. least square estimates of effects, ANOVA tables, Multiple comparisons: Tuckey's method, critical difference, advantages and limitations of each design

Unit IV (05L)

Missing plot technique: Estimation of one or two missing observations in RBD and LSD (least square estimates). ANOVA in case of missing observations.

Unit V (06L)

Factorial experiments: Meaning and advantages. 2^2 and 2^3 factorial experiments in RBD and LSD, main and interaction effects. Yates' method of computing factorial effect totals, ANOVA table and inferences. Contrasts and orthogonal contrasts

References

1) Cochran, W. G., & Cox, G. M. (1959). *Experimental Designs.* Wiley Eastern.
2) Federer, W. T. (1963). *Experimental Designs.* Oxfotd & IBH Publishing Co. Giri and Das
3) Goon, A. M., Gupta, M. K., & Das Gupta, B. (1983). *Fundamentals of Statistics, Vol. II.* World Press, Calcutta.
4) Guptha, S. C., & Kapoor, V. K. (1984). *Fundamentals of Applied Statistics.* Sulthan Chand and Sons.
5) Mukhopadhyaya, P. (1999). *Applied Statistics.* Books and Allied Pvt. Ltd.

SIXTH SEMESTER

BSCSTP 383: Statistics Practical – VII – Statistical Inference II & Design and Analysis of Experiments (1 credit)

List of Practical Assignments

1. Sequential probability ratio test.
2. Non-parametric inference
3. Exercise on Completely Randomised Design
4. Exercise on Randomised Block Design.
5. Exercise on Latin square Design.
6. Exercise on Missing observations
7. 2^2 Factorial experiment.
8. 2^3 Factorial experiment.

SIXTH SEMESTER

BSCSTC 382: Applied statistics (48 Lecture Hours) (2 Credits)

Unit I – Statistical Quality Control (11L)

Meaning of quality, quality characteristics - variables and attributes. Causes of variation - Assjgnable and Chance Causes, process control and product control. Control limits, specification limits, Natural tolerance limits, action limits, warning limits, Probability limits. General theory of control charts, Criteria for lack of Control, Selection of rational subgroups.

Unit II (11L)

Control chart for variables: $\bar{x}$ - R charts, $\bar{x}$ - s charts, control limits with and without standard values. Revised control charts. Construction and working of these charts,

Control chart for attributes: Need for attribute charts p, np, C and U charts, analysis and interpretation.

Unit III (11L)

Acceptance sampling Plans: Single sample plan by attributes, P.R, C.R, AQL, LTPD. Derivation of AOQ, O.C. ATI functions and their graphs, AOQL and indiffence Quality. Construction of SSP. Given AQL, LTPD, P.R.C.R. -Double sampling plan,

SSP by variablessingle specification, σ known and unknown plans. O.C. function Construction of these plans

Unit IV - **Index numbers, Demography (09L)**

Index Numbers: Introduction , Definition and Meaning, Points to be considered in construction of Index numbers, Simple and weighted price index numbers, Laspeyre's, Passche's and Fisher's Index numbers. Description of following index numbers - CPI, BSE, SENSEX

Demography:

Vital Statistics- Sources of vital statistics in India, functions of vital Statistics, rates and ratios, - Mortality rates- Crude, Age Specific and Standard Death rates - Fertility and reproduction rates, Crude birth rates general and specific fertility rates, gross and net reproductive rates, life – tables

Unit V - Time Series (06L)

Time Series: Time Series Meaning and Components of time series.

Additive and multiplicative models. Methods of estimating trend : moving average method, least squares method and exponential smoothing method Elimination of trend using additive

and multiplicative models Measurement and estimation of seasonal variations using Link Relative method and Ratio to trend method

References

1) Douglas Montgomery, C. (2001). *Introduction to Statistical Quality Control.* Wiely & Sons.

2) Goon, A. M., Gupta, M. K., & Das Gupta, B. (1991). *Fundamentals of Statistics, Vol. 1.* World Press, Calcutta.

3) Grant, E. L., & Richard, S. L. (2001). *Statistical Quality Control.* Tata McGraw Ltd.

4) Gupta, S. C., & Kapoor, V. K. (2002). *Fundamental of Mathematical Statistics.* Sultan Chand & sons.

5) Guptha, S. C., & Kapoor, V. K. (1984). *Fundamentals of Applied Statistics.* Sulthan Chand and Sons.

6) Levin, R. (2013). *Statistics for Management.* Prentice Hall India.

7) Medhi, J. (2006). *Statistical Methods: An Introductory Text.* New Age International(P) Limited, New Delhi.

8) Mitra, A. (2001). *Fundamentals of Quality Control and Improvement.* Pearson education, Asia.

9) Mukhopadhyaya, P. (1999). *Applied Statistics.* Books and Allied Pvt. Ltd.

SIXTH SEMESTER

BSCSTP 383: Statistics Practical – VIII – Applied Statistics (1 credit)

List of Practical Assignments

1) Exercise on $\bar{X}$ –R chart
2) Exercise on $\bar{X}$ –s chart
3) Control chart for attributes- p chart and np chart.
4) Control chart for attributes c chart and U chart.
5) Single sampling plan for attributes
6) Single sampling plan for variables.
7) Exercise on Index numbers
8) Exercise on vital statistics.
9) Exercise on Time Series. Graph of Time series data: identification of Components. Estimation of trend.

ELECTIVE COURSES

BSCSTCE 133: Official Statistics (24 L) (1 Credit)

Unit I (08L)

Introduction: Introduction to Indian and International Statistical system. Role, function and activities of Central and State Statistical organization. National Sample Survey Organization (NSSO) - functions. Ministry of programe implementation and Statistics.

Unit II (08L)

Population Statistics: Scope and content of Population census of India. Vital Statistics, methods of collecting population census, economic census.

Income Statistics: Uses, Methods of national income estimation, problems in the estimation of national income.

Unit III (08L)

Agricultural Statistics: System of collection of Agricultural Statistics. . Crop yield, Random sampling method. Production Statistics, Traditional Method, Crop estimation and forecasting.

Statistics related to industries, foreign trade, balance of payment, cost of living, inflation, educational and other social statistics.

References:

1. B. L. Agarwal (2003), Basic statistics, New age International Publishers
2. Guide to Official Statistics,Central Statistical Organisation, Department of Statistics, Ministry of Planning, Government of India, 1985 - India

3. Indian official statistical systems, M. R. Saluja, Statistical Pub. Society, 1972 - India

BSCSTCE 183: Statistical Analysis of Experimental Data (24L) (1 Credit)

Unit I - Preliminary description of error analysis (06L)

Introduction to measurements, measurement categories - Direct and derived quantities, errors in measurement, propagation of errors – sums and differences; products and quotients, independent errors in a sum, Nature of measurement errors-Systematic errors and random errors, Statistical analysis of random errors

Unit II - Normal Distribution (06L)

Histograms and distributions, The normal distribution, confidence interval, justification of mean as the best estimate, standard deviation of mean, The problem of combining separate measurements, weighted average and examples

Unit III - Regression Analysis (12L)

Bi-variate and Multivariate data, Correlation, Concept of Errors, Simple linear regression, interpretation of Regression coefficients, Principles of least squares, goodness of fit, multiple regression, non-linear regression, Examples using Excel

Monte Carlo techniques

Introduction, Random numbers, random number generation from probability distributions, specific distributions, Efficient Monte Carlo generation

References

1) John R Taylor (1996); An introduction to error analysis – The study of uncertainties in physical measurements, , University Science books, Sausalito, California

2) R Bevington, D Keith Robinson (2003): Data reduction and error analysis for the Physical sciences, Philip Third edition, Mc Graw Hill publications

3)Mechanical Measurement and Metrology, NPTEL online courses

   (https://nptel.ac.in/courses/112106138/8), course coordinated by IIT MADRAS

BSCSTCE 233: Introduction to Data Science and Big Data Analytics (24L) (1 Credit)

Unit I - Introduction to Big data and Data Analytics (08L)

Business Intelligence, Decision Support Systems, Data Warehousing; Definition of Big Data, Big data characteristics & considerations.

Data Analytics Lifecycle

Need of Data analytic lifecycle, Key roles for successful analytic projects, various phases of Data analytic lifecycle: Discovery, Data Preparation, Model Planning, Model Building, Communicating Results, Operationalization.

Unit II - Exploring Data in R and Machine Learning (10L)

Basic features of R, Exploring R GUI, Data Frames & Lists, Handling Data in R Workspace, Reading Data Sets & Exporting Data from R, Manipulating & Processing Data in R

Machine Learning: Supervised Learning

What is Machine Learning? Applications of Machine Learning; Supervised Learning: Structure of Regression Model, Linear Regression, Logistics Regression

Unit III - Classification & Unsupervised Learning (06L)

Classification: Classification Problem, Classification Models, Classification Trees, Bayesian Method; Association Rule: Structure of Association Rule, Apriori Algorithm, General Association; Clustering: Clustering Methods, Partition Methods, Hierarchical Methods.

.

References

- Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data by EMC Education Services (2015)
- Shmueli, G., Patel, N. R., & Bruce, P. C. (2010): Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XL Miner
- Y Lakshmi Prasad(2016) , Big Data Analytics made easy, Notion Press, Chennai
- Seema Acharya & Subhashini Chellappan (2015): Big Data & Analytics , Wiley Publications
- Carlo Vercellis (2009), Business Intelligence – Data Mining and Optimization for Decision Making - Wiley Publications

BSCSTOE 283: Basic Statistics

Unit I (08L)

Meaning and definition of Statistics. Collection of data, types of data – Primary data, Secondary data, Cross sectional data, longitudinal data. Measurement scales- nominal, interval and ratio. Types of data – qualitative and quantitative data - Classification of data

Elements of Probability Theory: Random experiments, Uncertainty, sample space, events, types of events, probability of an event, conditional probability, Bayes' Theorem.

Unit II - Analysis of Univariate data: (08L)

Measures of location- Mean, Median, Mode, Geometric mean and Harmonic mean and partition values, Measures of variation- Range, Quartile deviation, Mean deviation Standard deviation Coefficient of variation – Skewness and kurtosis, Stem and leaf diagram, Box plot – its interpretation.

Unit III (08L)

Analysis of Bivariate data: Correlation, types of correlation- Scatter diagram, Karl Pearson's method of computing correlation and its interpretation.

Association of Attributes: Class frequency, Inconsistency of data, consistency of data, coefficient of association and its interpretation.

Analysis of Univariate and bivariate data using R software.

References:

1. B. L. Agarwal (2003), Basic statistics, New age International Publishers
2. J Medhi (2006), Statistical Methods, New age International Publishers
3. S. C. Guptha (2016), Fundamentals of Statistics - Business statistics, Himalaya Publishing House
4. Levin R (2013), Statistics for Management, Prentice hall India