KLE Society's

Raja Lakhamagouda Science Institute (Autonomous), Belagavi

DEPARTMENT OF STATISTICS

I – SEMESTER

DSC - 21ST101: Descriptive Statistics

Total Teaching Hours: 56 **No. of Classes/Week: 04**

Course outcomes

After completing this course, the student will be able to:
CO1: Acquire knowledge of introductory statistics, its scope and importance in various areas such as Medical, Engineering, Agricultural and Social Sciences etc.
CO2: Get knowledge of various types of data, their organization and evaluation of summary measures such as measures of central tendency and dispersion etc.
CO3: Perceive the knowledge of correlation, regression analysis, regression diagnostics, partial and multiple correlations.
CO4: Learn different of types of data reflecting independence and association between two or more attributes.
CO5: Develop ability to critically assess a standard report having graphics, probability statements

Unit –1: Introduction to Statistics **13 Hours**
Statistics: Definition and scope. Concepts of statistical population and sample (SRS, Stratified, Systematic and Cluster sampling methods Definitions only). Data: quantitative and qualitative, cross sectional and time-series, discrete and continuous. Scales of measurement: nominal, ordinal, interval and ratio. Presentation of data: tabular and graphical. Frequency distributions, cumulative frequency distributions and the graphical representations. Stem and leaf displays.

Unit–2: Uni-Variate Data Analysis **18 Hours**
Measures of Central Tendency: Mean, weighted mean, trimmed mean, Median, Mode, Geometric and harmonic means, properties, merits and limitations, relation between these measures. Measures of Dispersion: Range, Quartile deviation, Mean deviation, Standard deviation and their relative measures. Gini's Coefficient, Lorenz Curve. Moments, Skewness and Kurtosis. Quantiles and Measures based on them. Box Plot. Outliers.

Unit– 3: Bi-Variate Data Analysis **15 Hours**
Bivariate Data, Scatter plot, Correlation, Karl Pearson's correlation coefficient, Rank correlation –Spearman's and Kendall's measures. Concept of errors, Principle of least squares, fitting of polynomial and exponential curves. Simple linear regression and its properties. Fitting of linear regression line and coefficient of determination.

Unit–4: Multi variate Data Analysis **10 Hours**
Analysis of Categorical Data: Contingency table, independence and association of attributes, measures of association - odds ratio, Pearson's and Yule's measure, Multivariate Frequencies, Multivariate Data Visualization, mean vector and dispersion matrix, Multiple linear regression, multiple and partial correlation coefficients. Residual error variance.

References:

1. Agresti, A. (2010): Analysis of Ordinal Categorical Data, 2nd Edition, Wiley.
2. Anderson T.W. and Jeremy D. Finn (1996). The New Statistical Analysis of Data, Springer
3. Freedman, D., Pisani, R. and Purves, R. (2014), Statistics, 4th Edition, W. W. Norton & Company.
4. Gupta, S. C. (2018), Fundamental of Statistics, Himalaya Publishing House,7th Edition.
5. Gupta S. C. and V. K. Kapoor (2020), Fundamental of Mathematical Statistics, Sultan Chand and Co. 12th Edition.
6. Hogg, R. V. McKean J. W. and Craig, A. T. (2012), Introduction to Mathematical Statistics, Pearson seventh Edition.
7. Joao Mendes Moreira, Andre CPLF De' Carvalho, Tomas Horvath (2018), General Introduction to Data Analytics, Wiley.
8. Johnson, R. A. and Bhattacharyya, G. K. (2006), Statistics: Principal sand methods. 5thEdition, John Wiley & Sons, New York.
9. Medhi, J. (2005), Statistical Methods, New Age International.
10. Ross, S. M. (2014), Introduction to Probability and Statistics for Engineers and Scientists, 5th Edition, Academic Press.
11. Tukey, J. W. (1977), Exploratory Data Analysis, Addison-Wesley Publishing Co.

Statistics Practical I

DSC - 21ST102: Statistics Practical I

Total Teaching Hours: 56 **No. of Classes/Week: 04**

Course Outcome

After completing this course, the student will be able to:

CO1: Acquire knowledge of introductory statistics, its scope and importance in various areas such as Medical, Engineering, Agricultural and Social Sciences etc.

CO2: Get knowledge of various types of data, their organization and evaluation of summary measures such as measures of central tendency and dispersion etc.

CO3: Perceive the knowledge of correlation, regression analysis, regression diagnostics, partial and multiple correlations.

CO4: Learn different of types of data reflecting independence and association between two or more attributes.

CO5: Develop ability to critically assess a standard report having graphics, probability statements.

1. Presentation of data by frequency tables, diagrams and graphs, stem and leaf, partition values.
2. Arithmetic Mean (AM), geometric mean, harmonic mean, weighted AM, trimmed mean, corrected mean.
3. Mode, median, partition values.
4. Absolute and relative measures of dispersion, Box plots.
5. Problems on moments, skewness and kurtosis.
6. Fitting of curves by least squares method.
7. Product moment correlation coefficient and rank correlation.
8. Regression of two variables.
9. Multivariate Descriptive statistics, mean Vector, dispersion matrix correlation matrix, Partial and Multiple correlation.
10. Problems on Association of attributes.

KLE Society's
Raja Lakhamagouda Science Institute (Autonomous), Belagavi

DEPARTMENT OF STATISTICS

II – SEMESTER

DSC - 21ST201: Probability and Distributions

Total Teaching Hours: 56 **No. of Classes/Week: 04**

Course Outcome

After completing this course, the student will be able to:

CO1: Conceptualize the probabilities of events including frequentist and axiomatic approach. Simultaneously, they will learn the notion of conditional probability including the concept of Bayes theorem.

CO2: Get knowledge related to concept of discrete and continuous random variables and their probability distribution including expectations and moments.

CO3: Learn knowledge of important discrete and continuous distribution such as Binomial, Poisson, Normal distributions.

CO4: Acquire knowledge on R-programming in the descriptive statistics and probability models.

Unit –1: Probability **15 Hours**

Random experiment, sample space and events, algebra of events. Definitions of Probability-Classical, statistical, subjective and axiomatic approaches – illustrations and applications, Addition rule, Conditional probability, independence of events and multiplication rule, Total Probability rule, Bayes theorem- applications.

Unit–2: Random Variables and Mathematical Expectation-(One Dimension) 14 Hours

Definitions of discrete and continuous random variables, Distribution function, probability mass and density functions – properties and illustrations, Expectation of a random variable and rules of expectation and related results, Moments and moment generating function – properties and uses.

Unit– 3: Standard Distributions **13 Hours**

Bernoulli, Binomial, Poisson, distributions– mean, variance, moments and m. g. f. recursive relations for probabilities and moments of Binomial and Poisson distributions, Normal distribution and its properties.

Unit–4: Data Analysis Using R **14 Hours**

Introduction to R: Installation, command line environment, overview of capabilities, brief mention of open-source philosophy. R as a calculator: The four basic arithmetic operations. Use of parentheses nesting up to arbitrary level. The power operation. Evaluation of simple expressions. Quotient and remainder operations for integers. Standard functions, e.g., sin, cos, exp, log. The different types of numbers in R: Division by zero leading to Inf or -Inf. NaN. NA. No need to go into details. Variables. Creating a vector using c(), seq() and colon operator. How functions map over vectors. Functions to summarize a vector: sum, mean, sd, median etc. Extracting a subset from the vector (by index, by property). R as a graphing calculator: Introduction to plotting. Plot(), lines(), abline(). No details about the graphics parameters except colour and line width. Barplot, Pie chart and Histogram. Box plot. Scatter plot and

simple linear regression using lm (y~x). Problems on discrete and continuous probability distributions.

References:

1. Dudewitz. E.J. and Mishra. S. N. (1998), Modern Mathematical Statistics. John Wiley.
2. Goon A.M., Gupta M.K., Das Gupta. B. (1991), Fundamentals of Statistics, Vol. I, World Press, Calcutta.
3. Gupta. S.C and V.K. Kapoor (2020), Fundamentals of Mathematical Statistics, Sultan Chand and Co, 12th Edition.
4. Hogg, R.V., Tanis, E.A. and Rao J.M. (2009), Probability and Statistical Inference, Seventh Edition, Pearson Education, New Delhi. 5. Mood, A.M., Graybill, F.A. and Boes, D.C. (2007), Introduction to the Theory of Statistics,3rd Edition. (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
6. Ross, S. (2002), A First Course in Probability, Prentice Hall. 7. Sudha G. Purohit, Sharad D. Gore, Shailaja R Deshmukh, (2009), Statistics Using R,Narosa Publishing House. 8. R for beginners by Emmanuel Paradis (freely available at https://cran.rproject.org/doc/contrib/Paradisrdebuts_en.pdf).

Statistics Practical II

DSC - 21ST202: Statistics Practical II

Total Teaching Hours: 56 **No. of Classes/Week: 04**

Course Outcomes:

After completing this course, the student will be able to:

CO1: Conceptualize the probabilities of events including frequentist and axiomatic approach. Simultaneously, they will learn the notion of conditional probability including the concept of Bayes theorem.

CO2: Get knowledge related to concept of discrete and continuous random variables and their probability distribution including expectations and moments.

CO3: Learn knowledge of important discrete and continuous distribution such as Binomial, Poisson, Normal distributions.

CO4: Acquire knowledge on R-programming in the descriptive statistics and probability models.

1. Two exercise on Descriptive statistics (Presentations, Summarizations, correlations, regression and Graphs using R).
2. Computing probability: using addition and multiplication theorems.
3. Conditional probability and Bayes' theorem.
4. Problems on pmf, expectation, variance, quantiles, skewness, kurtosis (Discrete case).
5. Problems on pdf, expectation, variance, quantiles, skewness, kurtosis (Continuous case).
6. Problems on discrete probability distributions (Binomial and Poisson).
7. Problems on Normal probability distributions.
8. Computation of moments and Moment generating functions (Discrete and Continuous Case).
9. Fitting of distributions Binomial, Poisson, Normal distributions.
10. Generation of random samples. (Binomial, Poisson, Normal)