VIDYASAGAR UNIVERSITY

Curriculum for 3-Year BSc (General)
in

Statistics
Under Choice Based Credit System (CBCS)
[w.e.f 2018-2019]

VIDYASAGAR UNIVERSITY
B Sc (General) in Statistics
[Choice Based Credit System]

Year	Semester	Course Type	Course Code	Course Title	Credit	L-T-P	Marks		
							CA	ESE	TOTAL
1	I			**SEMESTER-I**			CA	ESE	TOTAL
		Core-1 (DSC-1A)		Descriptive Statistics - **Lab**	6	4-0-4	15	60	75
		Core-2 (DSC-2A)		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		Core-3 (DSC-3A)		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		AECC-1 (Elective)		English/MIL	2	1-1-0	10	40	50
				Semester - I : Total	**20**				**275**
	II			**SEMESTER-II**					
		Core-4 (DSC-1B)		Elementary Probability Theory - **Lab**	6	4-0-4	15	60	75
		Core-5 (DSC-2B)		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		Core-6 (DSC-3B)		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		AECC-2 (Elective)		Environmental Studies	4		20	80	100
				Semester - 2 : Total	**22**				**325**

2

Year	Semester	Course Type	Course Code	Course Title	Credit	L-T-P	Marks		
2	**III**			**SEMESTER-III**			**CA**	**ESE**	**TOTAL**
		Core-7 (DSC-1C)		Introduction to Statistical Inference - **lab**	6	4-0-4	15	60	75
		Core-8 (DSC-2C)		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		Core-9 (DSC-3C)		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		SEC-1		TBD	2	1-1-0	10	40	50
				Semester - 3 : Total	**20**				**275**
	IV			**SEMESTER-IV**					
		Core-10 (DSC-1D)		Applications of Statistics - **Lab**	6	4-0-4	15	60	75
		Core-11 (DSC-2D)		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		Core-12 (DSC-3D)		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		SEC-2		TBD	2	1-1-0	10	40	50
				Semester - 4 : Total	**20**				**275**

Year	Semester	Course Type	Course Code	Course Title	Credit	L-T-P			
3	**V**			**SEMESTER-V**			**CA**	**ESE**	**TOTAL**
		DSE-1A		Discipline-1(Statistics)	6	4-0-4/ 5-1-0	15	60	75
		DSE-2A		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		DSE-3A		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		SEC-3		TBD	2	1-1-0	10	40	50
				Semester - 5 : Total	**20**				**275**
	VI			**SEMESTER-VI**					
		DSE-1B		Discipline-1(Statistics)	6	4-0-4/ 0-0-12	15	60	75
		DSE-2B		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		DSE-3B		Other Discipline/TBD	6	4-0-4/ 5-1-0	15	60	75
		SEC-4		TBD	2	1-1-0	10	40	50
				Semester - 6 : Total	**20**				**275**
				Total in all semester:	**122**				**1700**

CC = Core Course , **AECC** = Ability Enhancement Compulsory Course , **GE** = Generic Elective , **SEC** = Skill Enhancement Course , **DSE** = Discipline Specific Elective , **CA**= Continuous Assessment , **ESE**= End Semester Examination , **TBD**=To be decided , **CT** = Core Theory, **CP**=Core Practical , **L** = Lecture, **T** = Tutorial ,**P** = Practical , **MI**L = Modern Indian Language , **ENV**S = Environmental Studies

4

List of Core Course & Elective Courses

Core Courses (CC)

DSC-1A: **Descriptive Statistics**
DSC-1B: **Elementary Probability Theory**
DSC-1C: **Introduction to Statistical Inference**
DSC-1D: **Applications of Statistics**

Discipline Specific Electives (DSE)

DSE-1:   **Operations Research**
Or
DSE-1:   **Survival Analysis**
Or
DSE–1:   **Statistical Quality Control**

DSE- 2:   **Econometrics**
Or
DSE - 2:   **Financial Statistics**
Or
DSE - 2:   **Computation and Data Analysis**

Skill Enhancement Course (SEC)

SEC-1:   **Statistical Data Analysis using R**
SEC-2:   **Research Methodology**
SEC-3:   **Monte Carlo Method**
SEC-4:   **Data Base Management Systems**

Core Course

DSC-1A (CC-1): Descriptive Statistics   **Credits 06**

DSC1AT (CC-1): Descriptive Statistics   **Credits 04**

Course Contents:

Unit 1:
Introduction: Definition and scope of Statistics, concepts of statistical population and sample. Data: quantitative and qualitative, attributes, variables, scales of measurement - nominal, ordinal, interval and ratio. Frequency distribution, Presentation: tabular and graphic, including histogram and ogives.

Unit 2:
Measures of Central Tendency: mathematical and positional. Measures of Dispersion: range, quartile deviation, mean deviation, standard deviation, coefficient of variation, moments, skewness and kurtosis.

Unit 3:

Bivariate data: Definition, scatter diagram, simple, partial and multiple correlation (3 variables only), rank correlation (Spearman). Simple linear regression, principle of least squares and fitting of polynomials and exponential curves.

DSC1AP: Practical   **Credits 02**

1. Graphical representation of data
2. Problems based on measures of central tendency
3. Problems based on measures of dispersion
4. Problems based on combined mean and variance and coefficient of variation
5. Problems based on moments, skewness and kurtosis
6. Fitting of polynomials, exponential curves
7. Karl Pearson correlation coefficient
8. Partial and multiple correlations (3 variables only)
9. Spearman rank correlation with and without ties.
10. Correlation coefficient for a bivariate frequency distribution
11. Lines of regression, angle between lines and estimated values of variables.

Suggested Readings:

- Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I& II, 8th Edn. The World Press, Kolkata.
- Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.
- Mood, A.M. Graybill, F.A. AndBoes, D.C. (2007): Introduction to theTheory of Statistics, 3rd Edn. (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
- Goon A.M., Gupta M.K. and Dasgupta B. : Basic Statistics. The World Press, Kolkata.
- Chakraborty, Arnab (2016) : Probability and Statistics. Sarat Book House.

6

DSC-1B (CC-2): Elementary Probability Theory   **Credits 06**

DSC1BT (CC-2): Elementary Probability Theory   **Credits 04**

Course Contents:

Unit 1:
Probability: Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability – classical, statistical, and axiomatic. Conditional Probability, laws of addition and multiplication, independent events, theorem of total probability, Bayes' theorem and its applications.

Unit 2:

Random Variables: Discrete and continuous random variables, p.m.f., p.d.f., c.d.f. Illustrations of random variables and its properties. Expectation, variance, moments.

Unit 3:

Standard probability distributions: Binomial, Poisson, geometric, negative binomial, Uniform, normal, exponential. Weak law of large numbers and Lindeberg-Levy Central Limit Theorem (C.L.T).

DSC1BP: Elementary Probability Theory (Lab)   **Credits 02**

List of Practical:
1. Fitting of binomial distributions for n and p = q = ½
2. Fitting of binomial distributions for n and p given
3. Fitting of binomial distributions computing mean and variance
4. Fitting of Poisson distributions for given value of lambda
5. Fitting of Poisson distributions after computing mean
6. Application problems based on binomial distribution
7. Application problems based on Poisson distribution
8. Problems based on area property of normal distribution
9. To find the ordinate for a given area for normal distribution
10. Application based problems using normal distribution
11. Fitting of normal distribution when parameters are given
12. Fitting of normal distribution when parameters are not given

Suggested Readings:

➢ Hogg, R.V., Tanis, E.A. and Rao J.M. (2009): Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi.
➢ Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.
➢ Myer, P.L. (1970): Introductory Probability and Statistical Applications, Oxford & IBH Publishing, New Delhi

- Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I& II, 8th Edn. The World Press, Kolkata.
- Chakraborty, Arnab (2016) : Probability and Statistics. Sarat Book House.
- Ross, S. (2002): A First Course in Probability, Prentice Hall.

DSC-1C (CC-3): Introduction to Statistical Inference   **Credits:06**

DSC-1CT: Introduction to Statistical Inference   **Credits 04**

Course Contents:

Unit 1:
Population and Sample, Parameter and Statistic, Population distribution and Sampling distribution. Statistical Inference: Point Estimation, Interval Estimation and Testing of Statistical Hypothesis. Four useful distributions for statistical Inference; Normal, χ , t and F (Statement of the pdf's & shape of the curves)

Unit 2:
Estimation of population mean, confidence intervals for the parameters of a normal distribution (one sample and two sample problems). The basic idea of significance test. Null and alternative hypothesis. Type I& Type II errors, level of significance, concept of p-value. Tests of proportions. Tests of hypotheses for the parameters of a normal distribution (one sample and two sample problems). Sign test

Unit 3:
Analysis of variance, one-way and two-way classification. Brief exposure of three basic principles of design of experiments,Statistical concepts of "treatment", "plot" and "block". Analysis of completely randomized design, randomized complete block design.

DSC1CP: Introduction to Statistical Inference Lab   **Credits 02**

Practical:
1. Estimators of population mean.
2. Confidence interval for the parameters of a normal distribution (one sample and two sample problems).
3. Tests of hypotheses for the parameters of a normal distribution (one sample and two sample problems).
4. Test for proportions.
5. Sign test for median.
6. Analysis of Variance of a one way classified data
7. Analysis of Variance of a two way classified data.
8. Analysis of a CRD.
9. Analysis of an RBD.

Suggested Readings:
- Daniel, Wayne W., Bio-statistics (2005): A Foundation for Analysis in the Health Sciences. John Wiley .
- Goon, A.M., Gupta M.K. & Das Gupta (2005): Fundamentals of statistics, Vol.-I & II.
- Dass, M. N. &Giri, N. C.: Design and analysis of experiments. John Wiley.

> Dunn, O.J (1977): Basic Statistics: A primer for the Biomedical Sciences. John Wiley.

DSC-1D (CC-4): Applications of Statistics   Credits 06

DSC1DT (CC-4): Applications of Statistics   Credits 04

Course Contents:

Unit 1:
Concept of population and sample, complete enumeration versus sampling, sampling and non-sampling errors. Types of sampling: non-probability and probability sampling, basic principle of sample survey, simple random sampling with and without replacement, definition and procedure of selecting a sample, estimates of: population mean, total and proportion, variances of these estimates, estimates of their variances. Stratified random sampling: Technique, estimates of population mean and total, variances of these estimates, proportional and optimum allocations.

Unit 2:
Index numbers: Definition, Criteria for a good index number, different types of index numbers. Construction of index numbers of prices and quantities, consumer price index number & wholesale price index number. Uses and limitations of index numbers. Economic Time Series: Components of time series, Decomposition of time series-Additive and multiplicative model with their merits and demerits, Illustrations of time series. Measurement of trend by method of free-hand curve, method of least squares.

Unit 3:

Demographic Methods: Introduction, measurement of population, rates and ratios of vital events. Measurement of mortality: CDR, SDR (w.r.t. Age and sex), IMR, Standardized death rates. Life (mortality) tables: definition of its main functions and uses. Measurement of fertility and reproduction: CBR, GFR, and TFR. Measurement of population growth: GRR, NRR.

DSC1DP: Applications of Statistics Lab   Credits 02

Practical:
1. To select a SRS with and without replacement.
2. For a population of size 5, estimate population mean, population mean square and population variance. Enumerate all possible samples of size 2 by WR and WOR and establish all properties relative to SRS.
3. For SRSWOR, estimate mean, standard error, the sample size.
4. Stratified Sampling: allocation of sample to strata by proportional and Neyman's methods.
5. Construction of price and quantity index numbers by Laspeyre's formula, Paasche's formula, Marshall-Edgeworth's formula, Fisher's Formula. Comparison and interpretation.
6. Construction of Consumer and wholesale price index numbers, fixed base index number and consumer price index number with interpretation.

9

7. Measurement of trend: Fitting of linear, quadratic trend, exponential curve and plotting of trend values and comparing with given data graphically.
8. Computation of measures of mortality.
9. Completion of life table.
10. Computation of measures of fertility and population growth.

Suggested Readings:

➢ Mukhopadhyay, P. (1999): Applied Statistics, New Central Book Agency, Calcutta.
➢ Goon, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. II,9th Edition World Press, Kolkata.
➢ Gupta, S. C. and Kapoor, V.K. (2008): Fundamentals Of Applied Statistics, 4th Edition (Reprint), Sultan Chand &Sons

Discipline Specific Electives (DSE)

DSE-1: Operations Research   **Credits 06**

DSE1T: Operations Research   **Credits 04**

Course Contents:

Unit 1:
Introduction and Historical Background, Phases of Operations Research, model building, various types of O.R. problems. Linear Programming Problem, Requirements of LPP, Mathematical Formulation of LPP, Graphical Methods to Solve Linear Programming Problems.

Unit 2:
Simplex method for solving L.P.P. Charne's M-technique for solving L.P.P. involving artificial variables. Special cases of L.P.P. Concept of Duality in L.P.P: Dual simplex method.

Unit 3:
Introduction, Formulation of Transportation Problem (TP). Initial solution by North West corner rule, Least cost method and Vogel's approximation method (VAM), MODI's method to find the optimal solution, special cases of transportation problem. Assignment problem: Hungarian method to find optimal assignment, special cases of assignment problem.

DSE1P: Operations Research Lab   **Credits 02**

Practical (**Using TORA/WINQSB/LINDO/LINGO**):

1. Mathematical formulation of L.P.P and solving the problem using graphical method, Simplex technique and Charne's Big M method involving artificial variables.

2. Identifying Special cases by Graphical and Simplex method and interpretation
   a. Degenerate solution
   b. Unbounded solution
   c. Alternate solution
   d. Infeasible solution
3. Allocation problem using Transportation model.
4. Allocation problem using Assignment model.

Suggested Readings:
- Taha, H. A. (2007): Operations Research: An Introduction, 8 Hall of India.
- KantiSwarup, Gupta, P.K. and Manmohan (2007): Operations Research, 13th Edition, Sultan Chand and Sons.
- Hadley, G: (2002) : Linear Programming, Narosa Publications
- Hillier, F.A and Lieberman, G.J. (2010): Introduction to Operations Research-Concepts and cases, 9th Edition, Tata McGraw Hill

Or

DSE-1: Survival Analysis   **Credits 06**

DSE1T: Survival Analysis   **Credits 04**

Course Contents:

Unit 1:

Survival Analysis: Functions of survival times, survival distributions and their applications- exponential, gamma, Weibull, Rayleigh, lognormal distributions, and distribution having bath-tub shaped hazard function. Mean Residual Time.

Unit 2:
Censoring Schemes: Type I, Type II and progressive or random censoring with biological examples. Estimation of mean survival time and variance of the estimator for Type I and Type II censored data with numerical examples.

Unit 3:

Non-parametric methods: Actuarial and Kaplan-Meier methods for estimating survival function and variance of the Estimator.

DSE1P: Survival Analysis Lab   **Credits 02**

Practical:
1. To estimate survival function.
2. To determine death density function and hazard function.
3. To identify type of censoring and to estimate survival time for type I censored data.
4. To identify type of censoring and to estimate survival time for type II censored data.
5. To identify type of censoring and to estimate survival time for progressively type I censored data.

11

6. Estimation of mean survival time and variance of the estimator for type I censored data.
7. Estimation of mean survival time and variance of the estimator for type II censored data.
8. Estimation of mean survival time and variance of the estimator for progressively type I censored data.
9. To estimate the survival function and variance of the estimator using Non-parametric methods with Actuarial methods.
10. To estimate the survival function and variance of the estimator using Non-parametric methods with Kaplan-Meier method.

Suggested Readings:

➢ Lee, E.T. and Wang, J.W. (2003): Statistical Methods for Survival data Analysis, 3rd Edition, John Wiley and Sons.
➢ Kleinbaum, D.G. (1996): Survival Analysis, Springer.
➢ Chiang, C.L. (1968): Introduction to Stochastic Processes in Bio Statistics, John Wiley and Sons.
➢ Indrayan, A. (2008): Medical Biostatistics, 2nd Edition Chapman and Hall/CRC.

Or

DSE–1: Statistical Quality Control   **Credits 06**

DSE1T: Statistical Quality Control   **Credit 04**

Course Contents:

Unit 1:

Quality: Definition, dimensions of quality, Difference between product control and process control, Statistical Process Control - Seven tools of SPC, chance and assignable Causes of quality variation.

Unit 2:
Statistical Control Charts - Construction and Statistical basis of 3-σ Control charts, Rational Sub-grouping, Control charts for variables: X-bar & R-chart, X-bar & s-chart. Control charts for attributes: np chart, p-chart, c-chart and u-chart. Comparison between control charts for variables and control charts for attributes. Analysis of patterns on control chart, Estimation of process capability.

Unit 3:
Definitions related to product control, Acceptance sampling plan, Principle of acceptance sampling plans, Single sampling plan - their OC, AQL, LTPD, AOQ, AOQL, ASN, ATI functions with graphical interpretation, Double sampling plan - their OC, AQL, LTPD, AOQ, AOQL, ASN, ATI functions with graphical interpretation, use and interpretation of Dodge and Roming sampling inspection plan tables.

Unit 4:

Introduction to Six-Sigma: Overview of Six Sigma, Lean Manufacturing and Total Quality Management (TQM), Introduction to ISO quality standards: ISO 9001, ISO 14001, BIS.

DSE1P: Statistical Quality Control Lab   **Credits 02**

List of Practical
Construction and interpretation of statistical control charts
> X-bar & R-chart
> X-bar & s-chart
> np-chart
> p-chart
> c-chart
> u-chart

Single sample inspection plan: Construction and interpretation of OC, AQL, LTPD, ASN, ATI, AOQ, AOQL curves
Calculation of process capability

Suggested Readings:

- ➢ Montogomery, D. C. (2009): Introduction to Statistical Quality Control, 6th Edition, Wiley India Pvt. Ltd.
- ➢ Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. II, 8th Edn. The World Press, Kolkata.
- ➢ Mukhopadhyay, P (2011): Applied Statistics, 2nd edition revised reprint, Books and Allied (P) Ltd.
- ➢ Montogomery, D. C. and Runger, G.C. (2008): Applied Statistics and Probability for Engineers, 3rd Edition reprint, Wiley India Pvt. Ltd.
- ➢ Ehrlich,B.Harris(2002):Transactional Six Sigma and Lean Servicing, 2nd Edition St. Lucie Press.
- ➢ Hoyle, David (1995): ISO Quality Systems Handbook, Heinemann Publication. 2nd Edition, Butterworth

DSE- 2: Econometrics   **Credits 06**

DSE2T: Econometrics   **Credits 04**

Course Contents:

Unit 1:

Introduction: Objective behind building econometric models, nature of econometrics, model building, role of econometrics, structural and reduced forms. Estimation under linear restrictions. Dummy variables, Qualitative data.

Unit 2:

Multi-collinearity: Introduction and concepts, detection of multi-collinearity, consequences and solutions of multi-collinearity,.

13

Unit 3:

Autocorrelation: Concept, consequences of auto correlated disturbances, detection and solution of autocorrelation. Generalized least squares estimation.

Unit 4:

Heteroscedastic disturbances: Concepts and efficiency of Aitken estimator with OLS estimator under heteroscedasticity. Consequences of heteroscedasticity. Tests and solutions of hetero scedasticity. Errors in variables: Correlation between error and regressors. Instrumental variable method (Single-equation model with one explanatory variable)

DSE2P: Econometrics Lab   **Credits 02**

Practical:
1. Problems related to consequences of Multicollinearity.
2. Diagnostics of Multicollinearity.
3. Problems related to consequences of Autocorrelation (AR(I)).
4. Diagnostics of Autocorrelation.
5. Problems related to consequences Heteroscedasticity.
6. Diagnostics of Heteroscedasticity.
7. Estimation of problems of General linear model under Heteroscedastic distance terms.
8. Problems on Autoregressive models.
9. Problems on Instrumental variable.

Suggested Readings:
➢ Gujarati, D. and Sangeetha, S. (2007): Basic Econometrics, 4th Edition McGraw Hill Companies
➢ Johnston, J. (1972): Econometric Methods, 2nd Edition, McGraw Hill International.
➢ Koutsoyiannis, A. (2004): Theory of Econometrics, 2nd Edition, , Palgrave Macmillan Limited
➢ Maddala, G.S. and Lahiri, K. (2009): Introduction to Econometrics, 4th Edition, John Wiley & Sons.

<div align="center">Or</div>

DSE - 2: Financial Statistics   **Credits 06**

DSE2T: Financial Statistics   **Credits 04**

Course Contents:

Unit 1:
Probability review: Real valued random variables, expectation and variance, skewness and kurtosis, conditional probabilities and expectations. Discrete Stochastic Processes,

Binomial processes, General random walks, Geometric random walks, Binomial models with state dependent increments.

Unit 2:
Tools needed for Option Pricing: Wiener process, stochastic integration, and stochastic differential equations. Introduction to derivatives: Forward contracts, spot price, forward price, future price. Call and put options, zero-coupon bonds and discount bonds

Unit 3:
Pricing Derivatives: Arbitrage relations and perfect financial markets, pricing futures, put-call parity for European options, relationship between strike price and option price. Stochastic Models in Finance: Discrete time process- binomial model with period one.

Unit 4:
Stochastic Models in Finance: Continuous time process- geometric Brownian motion. Ito's lemma, Black-Scholes differential equation, Black-Scholes formula for European options, Hedging portfolios: Delta, Gamma and Theta hedging. Binomial Model for European options: Cox-Ross-Rubinstein approach to option pricing. Discrete dividends

DSE2P: Financial Statistics Lab   **Credits 02**

List of Practical
1. To verify "no arbitrage" principle
2. To verify relationship between spot price, forward price, future price
3. To price future contracts
4. To verify put-call parity for European options
5. To construct binomial trees and to evaluate options using these trees
6. To price options using black – Scholes formula
7. To hedge portfolios using delta and gamma hedging
8. To hedge portfolios theta hedging
9. Pricing of call options using binomial model
10. Computation of dividends on call options as a percentage of stock price.
11. Computation of dividends on call options as a fixed amount of money.
12. Pricing of put options using binomial model
13. Call-put parity for options following binomial models.
14. Effect of dividends on put options.

Suggested Readings:
➢ Franke, J., Hurdle, W.K. And Hafner, C.M. (2011): Statistics of Financial Markets: An Introduction, 3rd Edition,
➢ Stanley L. S. (2012): A Course on Statistics for Finance, Chapman and Hall/CRC.

DSE-2: Computation and Data Analysis   **Credits 06**

DSE2T: Computation and Data Analysis   **Credits 04**

Course Contents:

<u>**Computer Programming with C:**</u>

Input - output statements. Operator - relational and logical, Conditional operator. Library functions. Data type. Decision making and branching-if, if-else, Nesting of if statement, while, do-while. Arrays. Use of Functions. File handling,

Following problems to be studied:
(i)   Selection and Bubble sort, Computation of quantiles, Computation of Spearman's rank correlation coefficient (no tie case)
(ii)   Fitting of Binomial and Poisson distributions
(iii)   Interpolation by Lagrange's formula.
(iv)   Numerical integration (Trapezoidal and Simpson's1/3rule) with convergence;
(v)   Solution of numerical equations by Newton Raphson and iterative method (single variable);
(vi)   Addition, multiplication, transpose of matrices
(vii)   Correlation & Regression (2 and 3 variables)
(viii) Generation of random samples from Normal,Chi-square, t and F distributions, ANOVA.

DSE2P: Computation and Data Analysis Lab   Credits 02

The problem listed above to be solved by using C language for practical.

Skill Enhancement Course (SEC)

SEC-1 – Statistical Data Analysis using R   Credits 02

SEC1T: Statistical Data Analysis using R

Course Contents:

Unit 1:
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.

Unit 2:
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 summaries 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).

16

Unit 3:

Matrix operations in R: Creation. Basic operations. Extracting sub matrices. Loading data from a file: read. Table () and read.csv(). Mention of head = TRUE and head = FALSE. Data frames. Mention that these are like matrices, except that different columns may be of different types.

Unit 4:

Problems on discrete and continuous probability distributions.

Suggested Readings:

- Gardener, M (2012) Beginning R: The Statistical Programming Language, Wiley Publications.
- Braun W J, Murdoch D J (2007): A First Course in Statistical Programming with R. Cambridge University Press. New York
- A simple introduction to R by Arnab Chakraborty (freely   available at http://www.isical.ac.in/~arnabc/)
- R for beginners by Emmanuel Paradis (freely available at **https://cran.r-project.org/doc/contrib/Paradisrdebuts_en.pdf**)

SEC - 2: Research Methodology   Credits 02

SEC2T: Research Methodology

Course Contents:

Unit 1:

What is Research? Role of Research in important areas. Characteristics of Scientific Method. Process of research: Stating Hypothesis or Research question, Concepts & Constructs, Units of analysis & characteristics of interest, Independent and Dependent variables, Extraneous or Confounding variables. Measurements and scales of Measurements. Types of research: Qualitative & Quantitative Research, Longitudinal Research, Survey & Experimental Research

Unit 2:

Survey Methodology and Data Collection, sampling frames and coverage error, non-response.

Unit 3:

Review of various techniques for data analysis covered in core statistics papers, techniques of interpretation, precaution in interpretation.

Unit 4:

Develop a questionnaire, collect survey data pertaining to a research problem (such as gender discriminations in private v/s government sector, unemployment rates, removal of subsidy, impact on service class v/s unorganized sectors), questions and answers in surveys, Internal & External validity, , interpret the results and draw inferences. Formats and presentations of Reports – an overview.

Suggested Readings:

- Kothari, C.R. (2009): Research Methodology: Methods and Techniques, 2nd Revised Edition reprint, New Age International Publishers.
- Kumar, R (2011): Research Methodology: A Step - by - Step Guide for Beginners, SAGE publications.

SEC-3: Monte Carlo Method   Credits 02

SEC3T: Monte Carlo Method

Course Contents:

Unit 1:
Using the computer for random number generation. (treated as a black box). A brief look at some popular approaches (no mathematical justification needed). Simulating a coin toss, a die roll and a card shuffle.

Unit 2:
CDF inversion method. Simulation from standard distributions. Finding probabilities and moments using simulation

Unit 3:
Monte Carlo integration. Basic idea of importance sampling. MCMC not included.

Unit 4:
Generating from Binomial and Poisson distributions, and comparing the histograms to the PMFs. Generating from Uniform(0,1) distribution, and applying inverse CDF transforms. Simulating Gaussian distribution using Box-Muller method. Approximating the expectation of a given function of a random variable using simulation. Graphical demonstration of the Law of Large Numbers.Approximating the value of pi by simulating dart throwing.

Suggested Readings:

- Shonkwiler, Ronald W. and Mendivil, Franklin (2009):Explorations in Monte Carlo Methods (Undergraduate Texts in Mathematics)
- Carsey,Thomas M. and Harden, Jeffrey J. (2014):Monte Carlo Simulation and Resampling Methods for Social Science

SEC-4: Data Base Management Systems   Credits 02

SEC4T: Data Base Management Systems

Course Contents:

Unit 1:
Introduction: Overview of Database Management System, Introduction to Database Languages, advantages of DBMS over file processing systems.

Unit 2:
Relational Database Management System: The Relational Model, Introduction to SQL: Basic Data Types, Working with relations of RDBMS: Creating relations e.g. Bank, College Database (create table statement).

Unit 3:
Modifying relations (alter table statement), Integrity constraints over the relation like Primary Key , Foreign key, NOT NULL to the tables, advantages and disadvantages of relational Database System.

Unit 4:
Database Structure: Introduction, Levels of abstraction in DBMS, View of data, Role of Database users and administrators, Database Structure: DDL, DML, Data Manager (Database Control System).Types of Data Models Hierarchical databases, Network databases, Relational databases, Object oriented databases.

Suggested Readings:

- ➤ Gruber, M(1990): Understanding SQL, BPB publication.
- ➤ Silberschatz, A, Korth, H and Sudarshan, S(2011) "Database System and Concepts", 6th   Edition McGraw-Hill.
- ➤ Desai, B. (1991): Introduction to Database Management system, Galgotia Publications..