



University of Calcutta

B.Sc.STATISTICS(General)

2018

Credit Distribution across Courses

		Credits	
Course Type	Total Papers	Theory + Practical	Theory*
Core Courses	12	12*4 =48 12*2 =24	12*5 =60 12*1=12
Discipline Specific Electives	6	6*4=24 6*2=12	6*5=30 6*1=6
Ability Enhancement Language Courses	2	2*2=4	2*2=4
Skill Enhancement Courses	4	4*2=8	4*2=8
Totals	24	120	120

*Tutorials of 1 Credit will be conducted in case there is no practical component

Core Course: 4 courses each from **three** subjects, total 12 (1 course from each subject under each semester)

Discipline Specific Electives: 2 courses each from **three** subjects, total 6 (1 course from each subject under each semester)

Skill Enhancement Course: 2 courses each from **two** subjects, total 4

Choices for Discipline Specific Electives

Discipline Specific Elective			
Operations Research (Gr-A) Sem-5	Survival Analysis (Gr-B) Sem-6		
Econometrics (Gr-A) Sem-5	Project work (Gr-B) Sem-6		

Choices for Skill Enhancement Courses

Skill Enhancement Course Semesters 3-6	
Statistical Data Analysis Using R (Gr-A1) Sem-3	Research Methodology (Gr-A2) Sem-5
Data Base Management System (Gr B1) Sem4	Monte Carlo Methods (Gr-B2) Sem-6

Subject Code: Subject(STS)-Hons/Gen(A/G)-Course(CC/DSE/SEC/GE/AECC)-Semester(1-6)-Paper No-Paper Comp

B. Sc 3 years general course Syllabus (Statistics)

Core Course (and Generic Elective for Honours Course)

Descriptive Statistics

CC-1 Sem-1 Descriptive Statistics (STS-G-CC-1-1-TH) & (STS-A-GE-1-1-TH)	
	4 Credits
Unit 1	20 L
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	15 L
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	25 L
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.	
Reference Books	
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.	

Descriptive Statistics

CC-1 Sem-1 Descriptive Statistics Lab (STS-G-CC-1-1-P) & (STS-A-GE-1-1-P)

2 Credits

List of Practical

Graphical representation of data

Problems based on measures of central tendency

Problems based on measures of dispersion

Problems based on combined mean and variance and coefficient of variation

Problems based on moments, skewness and kurtosis

Fitting of polynomials, exponential curves

Karl Pearson correlation coefficient

Partial and multiple correlations (3 variables only)

Spearman rank correlation with and without ties.

Correlation coefficient for a bivariate frequency distribution

Lines of regression, angle between lines and estimated values of variables.

Elementary Probability Theory

CC-2 Sem-2 Elementary Probability Theory (STS-G-CC-2-2-TH) & (STS-A-GE-2-2-TH)	
	4 Credits
Unit 1	20 L
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	15 L
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	25 L
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).	
Reference Books	
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.	

Elementary Probability Theory Lab

CC-2 Sem-2 Elementary Probability Theory Lab (STS-G-CC-2-2-P) & (STS-A-GE-2-2-P)

2 Credits

List of Practical

- Fitting of binomial distributions for n and $p = q = \frac{1}{2}$
- Fitting of binomial distributions for n and p given
- Fitting of binomial distributions computing mean and variance
- Fitting of Poisson distributions for given value of λ
- Fitting of Poisson distributions after computing mean
- Application problems based on binomial distribution
- Application problems based on Poisson distribution
- Problems based on area property of normal distribution
- To find the ordinate for a given area for normal distribution
- Application based problems using normal distribution
- Fitting of normal distribution when parameters are given
- Fitting of normal distribution when parameters are not given

Introduction to Statistical Inference

CC-3 Sem-3 Introduction to Statistical Inference (STS-G-CC-3-3-TH) & (STS-A-GE-3-3-TH)	
	4 Credits
Unit 1	15 L
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^2 , t and F (Statement of the pdf's & shape of the curves)	
Unit 2	25 L
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 (Single Sample)	
Unit 3	20 L
Analysis of variance, one-way and two-way classification (one & multiple observation(s) per cell). 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.	
Reference Books	
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.	

Introduction to Statistical Inference Lab

CC-3 Sem-3 Introduction to Statistical Inference Lab (STS-G-CC-3-3-P) & (STS-A-GE-3-3-P)

2 Credits

List of Practical

Estimators of population mean.

Confidence interval for the parameters of a normal distribution (one sample and two sample problems).

Tests of hypotheses for the parameters of a normal distribution (one sample and two sample problems).

Test for proportions.

Sign test for median.

Analysis of Variance of a one way classified data

Analysis of Variance of a two way classified data.

Analysis of a CRD.

Analysis of an RBD.

Applications of Statistics

CC-4 Sem-4 Applications of Statistics (STS-G-CC-4-4-TH) & (STS-A-GE-4-4-TH)	
	4 Credits
Unit 1	20L
<p>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, estimates of their variances, proportional and optimum allocations.</p>	
Unit 2	20L
<p>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. Tests for index numbers</p> <p>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. Moving average method.</p>	
Unit 3	20L
<p>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.</p> <p>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.</p>	
Reference Books	
<p>Mukhopadhyay, P. (1999): Applied Statistics, New Central Book Agency, Calcutta.</p> <p>Goon, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. II, 9th Edition World Press, Kolkata.</p> <p>Gupta, S. C. and Kapoor, V.K. (2008): Fundamentals Of Applied Statistics, 4th Edition (Reprint), Sultan Chand & Sons</p>	

Applications of Statistics Lab

CC-4 Sem-4 Applications of Statistics Lab (STS-G-CC-4-4-P) & (STS-A-GE-4-4-P)

2 Credits

List of Practical

To select a SRS with and without replacement.

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.

For SRSWOR, estimate mean, standard error, the sample size.

Stratified Sampling: allocation of sample to strata by proportional and Neyman's methods.

Construction of price and quantity index numbers by Laspeyre's formula, Paasche's formula, Marshall-Edgeworth's formula, Fisher's Formula. Comparison and interpretation.

Construction of Consumer and wholesale price index numbers, fixed base index number and consumer price index number with interpretation.

Measurement of trend: Fitting of linear, quadratic trend, exponential curve and plotting of trend values and comparing with given data graphically.

Computation of measures of mortality.

Completion of life table.

Computation of measures of fertility and population growth.

Discipline Specific Electives

Operations Research

DSE-A Sem-5 Operations Research (STS-G-DSE-A-5-1-TH)	
	4 Credits
Unit 1	20L
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	15L
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	25L
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.	
Reference Books	
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	

DES-A Sem-5 Operations Research Lab (STS-G-DSE-A-5-1-P)

2 Credits

List of Practical (Using TORA/WINQSB/LINDO/LINGO)

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.

Survival Analysis

DSE-B Sem-6 Survival Analysis (STS-G-DSE-B-6-2-TH)	
	4 Credits
Unit 1	20L
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	20L
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	20L
Non-parametric methods: Actuarial and Kaplan-Meier methods for estimating survival function and variance of the Estimator.	
Reference Books	
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.	

Survival Analysis Lab

DSE-B Sem-6 Survival Analysis Lab (STS-G-DSE-B-6-2-P)

2 Credits

List of Practical

To estimate survival function.

To determine death density function and hazard function.

To identify type of censoring and to estimate survival time for type I censored data.

To identify type of censoring and to estimate survival time for type II censored data.

To identify type of censoring and to estimate survival time for progressively type I censored data.

Estimation of mean survival time and variance of the estimator for type I censored data.

Estimation of mean survival time and variance of the estimator for type II censored data.

Estimation of mean survival time and variance of the estimator for progressively type I censored data.

To estimate the survival function and variance of the estimator using Non-parametric methods with Actuarial methods.

To estimate the survival function and variance of the estimator using Non-parametric methods with Kaplan-Meier method.

Econometrics

DSE-A Sem 5 Econometrics (STS-G-DSE-A-5-1-TH)

4 Credits

Unit 1

15L

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

15L

Multicollinearity: Introduction and concepts, detection of multicollinearity, consequences and solutions of multicollinearity,.

Unit 3

15L

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

Heteroscedastic disturbances: Concepts and efficiency of Aitken estimator with OLS estimator under heteroscedasticity. Consequences of heteroscedasticity. Tests and solutions of heteroscedasticity.

Errors in variables: Correlation between error and regressors. Instrumental variable method (Single-equation model with one explanatory variable)

Reference Books

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.

Econometrics Lab

DSE-A Sem-5 Econometrics Lab (STS-G-DSE-A-5-1-P)

2 Credits

List of Practical

Problems related to consequences of Multicollinearity.

Diagnostics of Multicollinearity.

Problems related to consequences of Autocorrelation (AR(I)).

Diagnostics of Autocorrelation.

Problems related to consequences Heteroscedasticity.

Diagnostics of Heteroscedasticity.

Estimation of problems of General linear model under Heteroscedastic distance terms.

Problems on Autoregressive models.

Problems on Instrumental variable.

Project Work

DSE-B Sem-6 Project Work (STS-G-DSE-B-6-2-P)

6 Credits

Analysing Social Change in Historical Perspective

Objective: The aim of the course is to initiate students to write and present a statistical report, under the supervision of a faculty, on some area of human interest. The project work will provide hands on training to the students to deal with data emanating from some real life situation and propel them to dwell on some theory or relate it to some theoretical concepts.

Skill Enhancement Courses

Statistical Data Analysis using R-I

SEC-A1 sem-3 Statistical Data Analysis Using R (STS-G-SEC-A-3-1-TH)	
	2 Credit
Unit 1	10 L
<p>Introduction to R: Installation, commandline environment, overview of capabilities, brief mention of open source philosophy.</p> <p>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.</p> <p>Standard functions, e.g., sin, cos, exp, log.</p>	
Unit 2	10 L
<p>The different types of numbers in R: Division by zero leading to Infor -Inf. NaN. NA. No need to go into details.</p> <p>Variables. Creating a vector using c(), seq() and colon operator. How functions map over vectors.</p> <p>Functions to summarise a vector: sum, mean, sd, medianetc. Extracting a subset from the vector (by index, by property).</p> <p>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).</p>	
Unit 3	10 L
<p>Matrix operations in R: Creation. Basic operations. Extracting submatrices.</p> <p>Loading data from a file: read.table() and read.csv(). Mention of head=TRUE and head=FALSE. Dataframes. Mention that these are like matrices, except that different columns may be of different types.</p>	
Reference Books	

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/Paradis-rdebuts_en.pdf)

Data Base Management Systems

SEC-B1 Sem-4 Data Base Management Systems (STS-G-SEC-B-4-2-TH)	
	2 Credits
Unit 1	5 L
Introduction: Overview of Database Management System, Introduction to Database Languages, advantages of DBMS over file processing systems.	
Unit 2	5 L
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	5 L
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	5 L
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.	
Reference Books	
<ul style="list-style-type: none"> ▶ 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. 	

Research Methodology_

SEC-A2 Sem-5 Research Methodology (STS-G-SEC-A-5-3-TH)	
	2 Credit
Unit 1	16 L
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	14 L
Survey Methodology and Data Collection, sampling frames and coverage error, non-response.	
Reference Books	
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-B2 Sem-6 Monte Carlo Method (STS-G-SEC-B-6-4-TH)**2 Credit****Unit 1****8L**

Using the computer for random number generation. (treated as a black box)
A brief look at some popular approaches (nomathematical justification needed).
Simulating a coin toss, a die roll and a card shuffle.

Unit 2**8 L**

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

Unit 3**6L**

Monte Carlo integration. Basic idea of importance sampling.

Unit 4**8 L**

Generating from Binomial, 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.

Reference Books

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.