Core Courses in Statistics
for 3-years Multidisciplinary Program

STSMIN101T/STSCOR101T–Descriptive Statistics-I and Probability-I **[Credit3]** **[45 Lecture Hours]**
Unit1: Statistical Data **[12 Lecture Hours]**
Statistics: Definition and scope. Concepts of statistical population and sample. Data: quantitative and qualitative, cross-sectional and time-series, discrete and continuous. Scales of measurement: nominal, ordinal, interval and ratio. Collection of data, concept of questionnaire. Presentation of data: tabular and graphical. Frequency distributions, cumulative frequency distributions and their graphical representations. Stem and leaf displays.
Unit2: Univariate Data Analysis **[15 Lecture Hours]**
Measures of Central Tendency: Mean, Median, Mode. Measures of Dispersion: Range, Mean deviation, Standard deviation, Coefficient of variation, Gini's Coefficient, Lorenz Curve. Moments, skewness and kurtosis. Quantiles and measures based on them. Box Plot. Outliers and its detection using quantiles. Trimmed mean.
Unit3: Introduction to Probability **[18 Lecture Hours]**
Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability: classical, statistical and axiomatic. Probability space and different properties of probability function. Conditional Probability, laws of addition and multiplication, independent events, theorem of total probability, Bayes' theorem and its applications.
STSMIN101P/STSCOR101P: List of Practicals **[Credit2]** **[60 Lecture Hours]**

- Graphical representation of data.
- Problems based on construction of frequency distributions, cumulative frequency distributions and their graphical representations.
- 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.
- Problems related to quantiles and measures based on them.
- Problem of detection of outliers using quantiles, construction of boxplot.
- Numerical sums using classical definition of Probability.
- Numerical sums on conditional probability.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B.(2002):Fundamentals of Statistics, Vol. I & II, 8thEdition, World Press, Kolkata.
- Miller, Irwin and Miller, Marylees John E. Freunds(2006): MathematicalStatisticswithApplications,7thEdition, Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D. C.(2007): Introduction to the Theory of Statistics,3rdEdition, Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J. W. (1977): Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Freedman, D., Pisani, R. and Purves, R. (2014): Statistics,4thEdition, W.W. Norton & Company.
- Chung, K. L. (1983): Elementary Probability Theory with Stochastic Process, Springer/ Narosa.
- Feller, W. (1968): An Introduction to Probability Theory & its Applications, John Wiley.
- Goon, A.M., Gupta, M.K. & Dasgupta, B. (2003): An Outline of Statistical Theory Vol- I, World Press.
- Parzen, E. (1972): Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.(1937): Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.(1973): Exercises in Probability, Narosa.
- Ross, S. (2002): A First Course in Probability, Prentice Hall.
- Stirzaker, D. (2003): Elementary Probability, 2ndEdition, Cambridge University Press
- Rahman, N.A.(1983): Practical Exercises in Probability and Statistics, Griffin.
- Rohatgi, V. K. and Saleh, A.K. Md. E. (2009): An Introduction to Probability and Statistics. 2nd Edition, John Wiley and Sons.

STSMIN202T/STSCOR202T –Descriptive Statistics- II & Probability-II

[Credit3] **[45LectureHours]**

Unit1: Bivariate Data Analysis **[12 Lecture Hours]**

Bivariate data: Definition, scatter diagram, simple correlation, linear regression, principle of least squares, fitting o f polynomial and exponential curves, correlation ratio, correlation index, intra-class correlation. Rank correlation: Spearman's and Kendall's measures.

Unit2: Categorical Data Analysis **[12 Lecture Hours]**

Analysis of Categorical Data: Contingency table, independence & association of attributes. Ideas of complete and absolute association. Yule's measures of association and colligation, Cramer's measure of association, odds-ratio.

Unit3: Random Variables & Standard Discrete and Continuous Probability Distributions-I **[21 Lecture Hours]**

Random Variables: Definition of discrete and continuous random variables, cumulative distribution function (c.d.f.) and its properties (without proof), probability mass function (p.m.f.), and probability density function (p.d.f.). Expectation and Variance. Standard discrete probability distributions I: Discrete Uniform, Binomial, Poisson; Standard continuous probability distributions I: Rectangular, Exponential, Normal.

STSMIN202P/STSCOR202P: List of Practicals

[Credit2] **[60 Lecture Hours]**

- Correlation coefficient for a bivariate frequency distribution.
- Lines of regression, angle between lines and estimated values of variables.
- Fitting of polynomials, exponential curves.
- Spearman rank correlation with and with outlies.
- Computation of correlation ratio.
- Computation of intra class correlation coefficient.
- Fitting of binomial distribution for given n and p.
- Fitting of binomial distribution after computing mean and variance.
- Fitting of Poisson distribution for given value of lambda.
- Fitting of Poisson distribution after computing mean.
- Fitting of exponential distribution.
- Fitting of normal distribution.
- Application problem based on binomial distribution.
- Application problem based on Poisson distribution.
- Application problem based on negative binomial distribution.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta,B. (2002): Fundamentals of Statistics, Vol.I & II,8thEdition, World Press, Kolkata.

- Miller, Irwin and Miller, Marylees(2006):John E.Freunds Mathematical Statistics with Applications,7thEdition, Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes,D.C.(2007): Introduction to the Theory of Statistics, 3rd Edition,Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W.(1977): Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Freedman, D., Pisani, R. and Purves, R.(2014): Statistics, 4thEdition, W. W. Norton & Company.
- Agresti, A. (2010): Analysis of Ordinal Categorical Data,2nd Edition, Wiley.
- Chung, K.L. (1983): Elementary Probability Theory with Stochastic Process, Springer/ Narosa.
- Feller, W.(1968): An Introduction to Probability Theory & its Applications, John Wiley.
- Goon, A.M., Gupta, M.K. & Dasgupta, B. (2003): An Outline of Statistical Theory Vol- I, World Press.
- Parzen, E.(1972):Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V. (1937): Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T.(1973):Exercises in Probability, Narosa.
- Ross, S. (2002): A First Course in Probability, Prentice Hall.
- Stirzaker, D.(2003): Elementary Probability, 2ndEdition, Cambridge University Press
- Rahman, N.A.(1983):Practical Exercises in Probability and Statistics, Griffin.
- Rohatgi, V.K. and Saleh, A.K.Md.E.(2009): An Introduction to Probability and Statistics. 2nd Edition, John Wiley and Sons.

STSMIN303T/STSCOR303T –Survey Sampling
[Credit3] **[45 Lecture Hours]**

Unit1: Simple Random Sample **[15 Lecture Hours]**

Concept of population and sample, complete enumeration versus sampling, sampling and non- sampling errors. Types of sampling: non-probability and probability sampling, basic principles 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 and sample size determination.
Simple Random sampling using auxiliary information: Ratio and Regression methods of estimation.

Unit2: Stratified and Systematic Sampling **[18 Lecture Hours]**

Stratified random sampling, technique, estimates of population mean and total, variances of these estimates, proportional and optimum allocations and their comparison with SRS. Practical difficulties in allocation, estimation of gain in precision.
Systematic Sampling, Technique, estimates of population mean and total, variances of these estimates (N=n x k case). Comparison of systematic sampling

with SRS and stratified sampling in the presence of linear trend and corrections.

Unit3: Other Sampling Methods [12 Lecture Hours]

Cluster sampling (equal-size clusters only) estimation of population mean and its variance, Concept of sub-sampling. Two-stage sampling, Estimation of Population mean and variance of the estimate, comparison between two-stage, cluster and uni-stage sampling. Randomized response technique: Warner's method.

STSMIN303P/STSCOR303P: List of Practicals
[Credit2] [60 Lecture Hours]

- Select an 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
- Ratio and Regression estimation: Calculate the population mean or total of the population. Calculate mean squares. Compare the efficiencies of ratio and regression estimators relative to SRS.
- Stratified Sampling: allocation of sample to strata by proportional and Neyman's methods. Compare the efficiencies of above two methods relative to SRS.
- Estimation of gain in precision in stratified sampling.
- Comparison of systematic with stratified sampling and SRS in the presence of a linear trend.
- Cluster sampling: estimation of mean or total, variance of the estimate, estimate of intra-class correlation coefficient, efficiency as compared to SRS.
- Two stage sampling.

Reference Books

- Cochran, W.G. (1984): Sampling Techniques (3rd Ed.), Wiley Eastern
- Sukhatme, P.V., Sukhatme, B.V. Sukhatme, S. (1984). Sampling Theories of Survey with Application, IOWA State University Press and Indian Society of Agricultural Statistics
- Murthy, M.N. (1977): Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta
- Des Raj and Chandhok P. (1998): Sample Survey Theory, Narosa PublishingHouse
- Goon A.M., Gupta M.K. and Dasgupta B. (2008): Fundamentals of Statistics, Vol-2, World Press

STSCOR404T – Basics of Statistical Inference

[Credit 3] **[45 Lecture Hours]**

Unit1: Basic concepts of Estimation and Testing [15 Lecture Hours]

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

Unit2: Tests for categorical data and Nonparametric tests [15 Lecture Hours]

Categorical data: Tests of proportions, tests of association and goodness-of-fit using Chi-square test, Yates' correction.
Tests for the significance of correlation coefficient. Sign test for median, Sign test for symmetry, Wilcoxon two-sample test.

Unit3: Analysis of Variance and Design of Experiments [15 Lecture Hours]

Analysis of variance, one-way and two-way classification. Brief exposure of three basic principles of design of experiments, treatment, plot and block. Analysis of completely randomized design, randomized complete block design. Bioassay.

STSCOR404P: List of Practical

[Credit 2] **[60 Lecture Hours]**

- 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).
- Chi-square test of proportions.
- Chi-square tests of association.
- Chi-square test of goodness-of-fit.
- Test for correlation coefficient.
- Sign test for median.
- Sign test for symmetry.
- Wilcoxon two-sample test.
- 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.

Reference Books

- Daniel, Wayne W., Bio-statistics: A Foundation for Analysis in the Health Sciences. John Wiley (2005).
- Goon, A.M., Gupta M.K. & Das Gupta, Fundamentals of statistics, Vol.-I & II (2005).
- Das, M. N. &Giri, N. C.: Design and analysis of experiments. John Wiley.
- Dunn, O.J Basic Statistics: A primer for the Biomedical Sciences .(1964, 1977) byJohn Wiley.
- Bancroft, Holdon Introduction to Bio-Statistics (1962) P.B. Hoebar New York.
- Goldstein, A Biostatistics-An introductory text (1971). The Macmillion New York.

STSCOR505T –Applied Statistics

[Credit 3] **[45 Lecture Hours]**

Unit1: Time Series and Index Number **[20 Lecture Hours]**

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 semi-averages and method of least squares (linear, quadratic and modified exponential). Measurement of seasonal variations by method of ratio to trend.
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. Uses and limitations of index numbers.

Unit2: Statistical Quality Control **[10LectureHours]**

Statistical Quality Control: Importance of statistical methods in industrial research and practice. Determination of tolerance limits. Causes of variations in quality: chance and assignable. General theory of control charts, process & product control, Control charts for variables: X- bar and R-charts. Control charts for attributes: p and c-charts

Unit3: Analysis of Variance and Design of Experiment **[15 Lecture Hours]**

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

STSCOR505P: List of Practical
[Credit 2] **[60 Lecture Hours]**

- Estimators of population mean.
- Confidence interval for the parameters of a normal distribution (one sample and twosample problems).
- Tests of hypotheses for the parameters of a normal distribution (one sample and twosample problems).
- Chi-square test of proportions.
- Chi-square tests of association.
- Chi-square test of goodness-of-fit.
- Test for correlation coefficient.
- Sign test for median.
- Sign test for symmetry.
- Wilcoxon two-sample test.
- Run test for randomness.
- Analysis of Variance of a one way classified data
- Analysis of Variance of a two way classified data.

ReferenceBooks

- Daniel, Wayne W., Bio-statistics: A Foundation for Analysis in the Health Sciences.John Wiley (2005).
- Goon, A.M., Gupta M.K. & Das Gupta, Fundamentals of statistics, Vol.-I & II (2005).
- Das, M. N. &Giri, N. C.: Design and analysis of experiments. John Wiley.
- Dunn, O.J Basic Statistics: A primer for the Biomedical Sciences .(1964, 1977) byJohn Wiley.
- Bancroft, Holdon Introduction to Bio-Statistics (1962) P.B. Hoebar New York.
- Goldstein, A Biostatistics-An introductory text (1971). The Macmillion New York.

STSCOR606T –Statistics in Finance

[Credit3] **[45LectureHours]**

Unit1: Derivatives **[10 Lecture Hours]**

Introduction to derivatives: Forward contracts, spot price, forward price, future price. Options, zero-coupon bonds and discount bonds.

Unit2: Some Stochastic Processes **[15 Lecture Hours]**

Introduction to Random walk, Brownian Motion and Martingales.
Option pricing: One-step and Two-step Binomial Models, General Binomial Tree Model, Black-Scholes formula, Implied Volatility, Properties of the European Call Option, Put Options, Put-call parity for European options

Unit3: Portfolio Optimization **[20 Lecture Hours]**

Portfolio Optimization: Efficient frontier and Tangency portfolio - Efficient portfolio with N risky assets and one risk-free asset, Notion of VaR, Capital Asset Pricing Model, Capital Market Line (CML). Security Market Line (SML) - Security Characteristic Line (SCL), Testing for CAPM.

Problems based on Derivatives:

Finding European Call/Put Option using One-Step and Two-Step Binomial Tree Models. Finding European Call/Put Option using Black Scholes Formula.

Finding an Efficient Portfolio of N risky assets and one risk free asset.

STSCOR606P: List of Practical

[Credit2] **[60 Lecture Hours]**

- Practical related to Unit 1.
- Practical related to Unit 2.
- Practical related to Unit 3.

Reference Books

- Introduces Quantitative Finance – Paul Wilmott
- Options, Futures and other derivatives – John C Hull.
- An Elementary Introduction to Statistics in Finance – S M Ross
- Statistics in Finance – David Ruppert
- Statistics of Financial Markets – J Franker, C M Hafner Stochastic Processes- S.M.Ross

OR

STSCOR607T –Survival Analysis and Clinical Trials

[Credit3] **[45 Lecture Hours]**

Unit1: Introduction to Survival Analysis **[20 Lecture Hours]**

Survival Analysis: Functions of survival times, survival distributions and their applications, exponential, gamma, Weibull, Rayleigh, lognormal, death density function for a distribution having bath-tub-shaped hazard function. 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. Non-parametric methods: Actuarial and Kaplan-Meier methods for estimating survival function and variance of the Estimator.

Unit2: Risk Theory **[10 Lecture Hours]**

Competing Risk Theory: Indices for measurement of probability of death under competing risks and their inter-relations. Estimation of probabilities of death using maximum likelihood principle and modified minimum Chi-square methods. Theory of independent and dependent risks. Bivariate normal dependent risk model.

Unit3: Clinical Trial **[15 Lecture Hours]**

What is clinical trial? Different phases; Major steps of executing a controlled clinical trial; Type of control groups; Blinding; Bias; Ethics of randomization.

Determination of trial size; Randomized clinical trial

STSCOR607P: List of Practical

[Credit2] **[60LectureHours]**

(The entire practical are to be done preferably by using R/statistical packages.)

- 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 Nonparametric methods with Actuarial methods
- To estimate the survival function and variance of the estimator using

Nonparametric methods with Kaplan-Meier method

- To estimate Crude probability of death
- To estimate Net-type I probability of death
  To estimate Net-type II probability of death
- To estimate partially crude probability of death
- To simulate the random sequence of treatment assignments.

Reference Books

- Lee, E.T. and Wang, J.W.(2003): Statistical Methods for Survival data Analysis, 3rdEdition, JohnWiley and Sons.
- Biswas,S.(2007):AppliedStochasticProcesses:ABiostatisticalandPopulation OrientedApproach, Reprinted 2ndCentralEdition,NewCentralBook Agency.
- Chiang,C.L.(1968):Introduction to Stochastic Processes in Bio StatisticsJohnWileyandSons.
- Indrayan, A.(2008):Medical Biostatistics,2nd Edition Chapman and Hall/CRC.Rosenberger and Lachin: Randomized Clinical Trials: Theory and Practice
- Ding-Geng(Din)Chen and Karl E. Peace: Clinical Trial Data Analysis Using R

N.B.: Student may opt any one course between STSCOR606 and STSCOR607 in Semester 6.

STSGSE301M /STSGSE501M– C++ Programming
[Credit2] **[90 Lecture Hours]**

Unit1: Introduction to C++ **[70 Lecture Hours]**

Components, basic structure of programming. Notion of header file. Concept of character and variable; allocation of memory. Declaration and assignment of variables and array variables. Input and output operations in C. Use of conditional operations and loops; if…else, for, while, do…while etc. dim arrays. User defined functions.

Unit 2: Applications of C++ **[20 Lecture Hours]**

Applications in Simple mathematical operations, Sorting of an array and finding quantiles, Preparing a frequency table, Mean, median and mode of a grouped frequency Data, Variance and coefficient of variation of a grouped frequency data.

Reference Books

- Kernighan, B.W. and Ritchie, D. (1988): C Programming Language, 2ndEdition, Prentice
- Balagurusamy,E. (2011):ProgramminginANSIC,6thEdition Tata McGraw
- Gottfried, B.S. (1998): Schaums Outlines: Programming with C, 2ndEdition, Tata McGraw Hill.

STSGSE402M/STSGSE602M –Computation using Software
[Credit3] **[90 Lecture Hours]**

Unit1: Introduction to Excel **[40 Lecture Hours]**

Basic idea about software. Input and modification of data. Basic cell operations; operation of mathematical and inbuilt functions on cell. Display summary statistics for univariate and bivariate data. Regression and correlation computation. Use of Data Analysis' tool (only the applicable functions). Construction of data table and operations on it.

Graphical representation of data: Column diagram, pie diagram, line diagram, bar diagram, scatterplot, boxplot, stock diagram, surface plot and radar plot, histogram.

Unit2: Introduction to R **[50 Lecture Hours]**

Use of Rascal calculator. Operations within build mathematical functions. Input a vector, numeric and non-numeric vectors. Addition and deletion of data from a vector. Logical operations and use of different logical functions. Understanding the non-numeric outputs, like—NULL, NA and NaN. Array and Matrix with associated operations. Construction of new function in R. Use the help in R. Loading and installing packages in R.

ReferenceBooks

- Davies, T.M. (2016): The Book of R: A First Course in Programming and Statistics,1st Edition, NoStarch Press, USA.