

List of MDC courses:

1. Introductory Data Analysis (Sem1) for all students. This is a new course which is given as an additional option.
2. Statistics for Practitioners (Sem 2) for other departments. This is an old course which is already in the syllabus.
3. Advanced concepts for data handling (Sem 2) for Statistics students. This is a new course which is given as an additional option.

Introductory Data Analysis (STAT105MDC01B) (The old paper Statistics for All will be STAT105MDC01A)

Objective of the course is not to include any derivation of the theory but to demonstrate the theory and conceptual clarifications by as many data-driven applications / examples as possible.

Unit 1:

What is Statistics? Statistics as a science and technology, Role of Probability in Statistics, Brief Overview of Data Analysis: Exploratory data analysis and Statistical Inference, Observational study and planning of experiments.

Unit 2:

Brief overview of classification of data on the basis of source and type, nature of variables, scales of measurements with examples, Brief idea of modern multivariate data with different applications, Overview of techniques of collection of primary data, Challenges in Data processing: scrutiny of data, outliers, variable transformation, handling missing data, variable selection. Presentation of data: Role of graphs, identifying misleading presentation with examples.

Unit 3:

Histogram as tool of summary, overview of different univariate summary measures (proofs of results not included), scatterplot as a tool of bivariate summary, idea of correlation and regression (proofs not included), idea of fitting mathematical curves, method of least squares (detail proof not included), correlation and causation, ecological correlation, concept of regression towards mean.

Unit 4:

Concept of population and sample, role of probability distributions in characterizing population: some standard probability models as examples, concept of frequency curve, population mean and variance in terms of expectation, idea of parameter and statistic, concept of sampling distributions and standard error, brief overview of inference with examples.

Reference

1. Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata.
2. Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd.
3. Johnson, R.A. and Wichern, D.W. Applied Multivariate Statistical Analysis, PHI.

Advanced concepts of data handling (STAT155MDC02B) (The old paper : Statistics for Bioscience will be STAT155MDC02B)

Objective of the course is not to include any derivation of the theory but to demonstrate the theory and conceptual clarifications by as many data-driven applications / examples as possible.

Unit 1:

Review of basic concepts of probability, application of standard discrete and continuous probability models, fitting of probability distributions, truncated distributions, review and applications of probability inequalities, bivariate normal (different concepts without proofs).

Unit 2:

Review of univariate and bivariate summary measures, review of simple linear regression, idea of non-linear regression, polynomial regression and correlation index, exponential curves, log-linear and power transformations. Idea of correlation ratio, rank correlation (Spearman's and Kendall's). Notion of independence & association in contingency tables; Concept of odds ratio, Pearson's chi-square. idea of logistic regression and regression for count data.

Unit 3:

Concept of population and sample, concept of bias and sampling error, idea of probability sampling with the help of examples, notion of parametric and non-parametric inference, idea of parameter and statistic, sampling distribution of statistic, standard error with the help of examples.

Unit 4:

Basic idea of estimation, testing of hypothesis and confidence estimation, idea of mean square error, notion of hypothesis, level of significance, p-value, critical value with the help of examples, some tests for mean and proportions, tests for several means, chi-square tests (derivation not included, only illustrations)

Reference

1. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I& II, 8th Edn. The World Press, Kolkata.
2. Das, N.G.: Statistical Methods, Vol I and II, Tata McGraw Hill Pub. Co. Ltd.
3. Gupta, S. C. and Kapoor, V.K. (2008): Fundamentals Of Mathematical Statistics, Sultan Chand & Sons
4. Hogg, R.V., Tanis, E.A. and Rao J.M. (2009): Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi.
5. Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia.
6. Myer, P.L. (1970): Introductory Probability and Statistical Applications, Oxford & IBH Publishing, New Delhi

Summer Internship (STAT341SI01)

A course requiring students to participate for a minimum of 30 hours in professional employment related activity or academic research works related to the discipline at any one of the following entities:

1. A private sector office, bank or industry
2. Any government aided University or Institute
3. Any Central/ State Government Statistical office or agency.
4. Any hospital or recognized healthcare center or allied research bodies.

However in exigencies, the Departmental Academic Committee (DAC) may allow any candidate to pursue his/her internship under bodies not included in the above list. In all cases, a candidate has to take prior approval/ NOC from the DAC through the Head of the Department. A candidate is normally expected to complete his/her internship during the semester break after the completion of Semester 4, but he/she may be allowed to complete the same before the start of Semester 7. Upon completion of the internship, the candidate has to submit a complete report of his/her work along with a completion certificate from the concerned organization to the DAC for evaluation and has to present his/her work before the committee. The decision of DAC regarding the approval of entity, topic of work and duration of leave (if any) for the internship shall be considered final.

STAT255MC04B: Applied Multivariate and Regression Models (The old paper Applied Statistics will be STAT255MC04A)

Unit 1

Introduction to multivariate data: mean vector, dispersion matrix, correlation matrix. Multiple linear regressions, multiple correlation (R) and partial correlation, Use of R^2 and Adjusted R^2 .

Unit 2

Principal component analysis and Factor analysis, Cluster analysis and Discriminant analysis (data oriented approach).

Unit 3

Concept of General linear model, use of dummy variables, concept of interaction, models including quantitative and qualitative covariates (examples only). Estimation under General Linear Model: Least Square Technique, Gauss-Markov Theorem (Statement Only), General linear Hypothesis (Statement only) and applications in different models, confidence interval and prediction interval.

Unit 4

Outliers and influential observations, Idea of problem of heteroscedasticity and autocorrelation, Multicollinearity, Model selection techniques: AIC, Mallow Cp (Derivations Excluded). Concept of training set and test set, Cross-validation. Forward and Backward selection. Idea of regression for binary and count data.

Reference

1. Gun, A.M., Gupta, M.K. and Dasgupta, B. (2008): Fundamentals of Statistics, Vol. I, 9th Edition World Press, Kolkata.
2. Das, N.G.: Statistical Methods, Vol I, Tata McGraw Hill Pub. Co. Ltd.
3. Johnson, R.A. and Wichern, D.W. Applied Multivariate Statistical Analysis, PHI.
4. Hardle W. and Simar, L. Applied Multivariate Statistical Analysis.
5. Kutner, M.H. et.al., Applied Linear Statistical Models.
6. Belsley D.A. et.al., Regression Diagnostics.
7. Draper N.R. and Smith, H. Applied Regression Analysis.