

CURRICULUM

SEMESTER I

DSC 511 Statistical Foundations for Data Science [2-0-0-2]

Objectives

- To learn basic and some advanced concepts in probability and statistics.
- To learn the concepts of statistics and random process in solving problems arising in data science.

Outcomes

- Students will be able to model uncertain phenomena using probability models and calculate the uncertainty in systems where such phenomena are a part of the system.
- Students will be able to implement statistical analysis techniques for solving practical problems.

Syllabus

Probability: Sample space, events and axioms; conditional probability; Bayes theorem; Random variables; Standard discrete and continuous probability distributions; Expectations and moments; Covariance and correlation; Linear Regression; Central limit theorem.

Statistics: Sampling distributions of the sample mean and the sample variance for a normal population; Point and interval estimation; Sampling distributions (Chi-square, t,F,Z), Hypothesis testing ; One tailed and two-tailed tests; Analysis of variance, ANOVA, One way and two way classifications.

Random Processes: Definition and classification of random processes, Poisson process, Gaussian white noise. Statistical analysis using R.

Learning Resources

1. S. Ross, Introduction to Probability and Statistics for and Engineers and Scientists, Third Edition, Elsevier, 2004.
2. G. R. Grimmett and D. R. Stirzaker, Probability and Random Processes, Oxford University Press, 2001
3. R.V. Hogg, J.W. McKean & A. Craig, Introduction to Mathematical Statistics, 6th Edition.
4. Montgomery, D. C. and G. C. Runger, Applied Statistics and Probability for Engineers. 5th Edition. John Wiley & Sons, Inc., NY, USA. 2009
5. Robert H. Shumway and David S. Stoffer, Time Series Analysis and Its Applications with R Examples, Third edition, Springer Texts in Statistics, 2006.
6. Athanasios Papoulis, Probability Random Variables and Stochastic Processes, 4th edition, McGraw-Hill, 2002.