Statistical Foundations for Data Science

Probability spaces, conditional probability, independence; Random variables, distribution functions, probability mass and density functions, functions of random variables, standard univariate discrete and continuous distributions; Mathematical expectations, moments, moment generating functions, inequalities; Random vectors, joint, marginal and conditional distributions, conditional expectations, independence, covariance, correlation, standard multivariate distributions, functions of random vectors; Law of large numbers, sampling distribution and central limit theorem, Gaussian/Normal Distribution its pdf and cdf, Symmetric distribution, Skewness and Kurtosis, Standard normal variate (Z) and standardization, kernel density estimation, Q-Q plot, Chebyshev's inequality, Discrete and Continuous Uniform distribution, Bernoulli and Binomial distribution, Log Normal distribution, Power log normal distribution, Box cox transform. Sampling distributions; Point estimation - estimators, minimum variance unbiased estimation, maximum likelihood estimation, method of moments, consistency; Interval estimation; Testing of hypotheses - tests and critical regions, Chi-squared test, F-test, G-test, KS Test.

Textbooks:

• Douglas C. Montgomery and George C. Runger, Applied Statistics and Probability for Engineers, John Wiley and Sons Inc., 2005

• Amir D Azcel, Jayavel Sounderpandian, Palanisamy Saravanan and Rohit Joshi, Complete Business Statistics, 7th edition McGrawHill education 2012.

• Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers and Keying Ye, Probability and Statistics for Engineers and Scientists, 8th Edition, Pearson Education Asia, 2007.

• Ross S.M., Introduction to Probability and Statistics for Engineers and Scientists, 3rd edition, Elsevier Academic Press.