## 19EMA110 : STATISTICAL MODELING

**L T P C**

**3 0 0 3**

## UNIT I 9 L

**Linear Statistical Models, Estimation and Sufficient Statistic:**

Simple linear regression & correlation, multiple regression & multiple correlation, Analysis of variance (one way, two way with as well as without interaction) , **Estimation**: Point estimation, criteria for good estimates (un-biasedness, consistency), Methods of estimation including maximum likelihood estimation, **Sufficient Statistic**: Concept & examples, complete sufficiency, their application in estimation.

## Learning Outcomes:

At the end of this unit, the student will be able to

* + use the equation of a linear function to model a data set (L3)
  + apply their knowledge of linear model to solve real world problems in comparing two measurement data (L3)
  + approximate the value of a population parameter on the basis of a sample statistic (L4)
  + learn a formal definition of sufficiency (L5)
  + learn how to apply the Factorization Theorem to identify a sufficient statistic (L5)

## UNIT II 9 L

**Test of hypothesis:** Concept & formulation, Type I and Type II errors, Neyman Pearson lemma, Procedures of testing (single proportion and mean, double proportions and Means for Large Samples, t-test, F-test and Chi-Square tests for Small samples).

## Learning Outcomes:

At the end of this unit, the student will be able to

* + understand how to develop Null and Alternative Hypotheses (L3)
  + understand Type I and Type II Errors (L3)
  + learn how to perform hypothesis testing for population proportion by the p-value approach(L4)
  + use confidence interval to draw conclusion about two-sided test(L5).

## UNIT III 8 L

**Non-parametric Inference:** Comparison with parametric inference, Use of order statistics. Sign test, Wilcoxon signed rank test, Mann-Whitney test, Run test, Kolmogorov-Smirnov test. Spearman’s and Kendall’s test. Tolerance region.

## Learning Outcomes:

At the end of this unit, the student will be able to

* + find the differences among parametric, nonparametric and semi-parametric inferences (L2)
  + learn nonparametric procedures for analyzing real data (L4)
  + perform and interpret the Mann Whitney U Test, Run test, Kolmogorov-Smirnov test, Spearman’s and Kendall’s test (L3)
  + identify the appropriate nonparametric hypothesis testing procedure based on type of outcome variable and number of samples (L4)

## UNIT IV 8 L

**Basics of Time Series Analysis & Forecasting:** Stationary, ARIMA Models: Identification, Estimation and Forecasting.

## Learning Outcomes:

At the end of this unit, the student will be able to

* + identify and interpret a non-seasonal ARIMA model(L4)
  + distinguish ARIMA terms from simultaneously exploring an ACF and PACF(L5)
  + test that all residual autocorrelations are zero(L4)
  + convert ARIMA models to infinite order Moving Average models(L4)
  + forecast with ARIMA models(L5)

## UNIT V 8 L

**R statistical programming language:** Introduction to R, Functions, Control flow and Loops, Working with Vectors and Matrices, Reading in Data, Writing Data, Working with Data, Manipulating Data, Simulation, Linear model, Data Frame, Graphics in R.

## Learning Outcomes:

* At the end of this unit, the student will be able to
* motivate for learning a programming language (L3)
* access online resources for R and import new function packages into the R workspace (L4)
* import, review, manipulate and summarize data-sets in R (L4)
* explore data-sets to create testable hypotheses and identify appropriate statistical tests using R (L5) create and edit visualizations with R (L4)

## Text Book(s):

* 1. Probability and Statistics for Engineers (4th Edition), I.R. Miller, J.E. Freund and R. Johnson.
  2. Fundamentals of Statistics (Vol. I & Vol. II), A. Goon, M. Gupta and B.Dasgupta.
  3. The Analysis of Time Series: An Introduction, Chris Chatfield.
  4. Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, Introduction to Linear Regression Analysis, 5/e, Wiley, 2015.
  5. Chris Chatfield, The Analysis of Time Series: Texts in Statistical Science, 6/e, Chapman & Hall/CRC, 2016.
  6. Garrett Grolemund, Hands-On Programming with R: write your own functions and simulations, Hadley Wickham, 2014.
  7. Jared P. Lander, R for EveryOne: Advanced Analytics and Graphics, 2/e, Addison Wesley Data and Analytics Series, 2017

## Reference Book(s):

1. Introduction to Linear Regression Analysis, D.C. Montgomery &E.Peck
2. Introduction to the Theory of Statistics, A.M. Mood, F.A. Graybill& D.C. Boes.
3. Applied Regression Analysis, N. Draper & H. Smith
4. Hands-on Programming with R,- Garrett Grolemund
5. R for Everyone: Advanced Analytics and Graphics, Jared P. Lander