Code No.: **MDS 553**
Course Title:**Statistical Methods for Data Science**   Full Mark: 75
Nature: **Theory and Practical** (Compulsory)   Credit: 3

Course Description:

The course explains different probability distributions and their applications, some non-parametric statistical tests and their applications, different aspects of the testing of hypothesis along with Neymann-Pearson's lemma, uniformly most powerful tests, likelihood ratio tests for testing means and variance in exponential families.

Course Objectives:

After completion of this course, students will be able to
- Understand the concept of multinomial probability distributions, their major characteristics and applications
- Be familiar with probability functions of extreme value distributions, their major characteristics and their applications
- Understand the concept of generalized power series distribution with special focus to Binomial, Poisson, Negative Binomial distributions, and examples
- Know meaning and importance of prior and posterior distributions, applications focusing on some particular distributions and examples
- Understand how the distributions are compounding, understand mixed type distributions and their applications
- Know the difference between parametric and non-parametric statistical tests
- Apply non-parametric statistical tests appropriately in real life data analysis
- Understand the different aspects of testing of hypothesis, likelihood ratio tests and their applications.

Course Contents:
Unit 1:Multinomial Distribution   [4 Hrs.]

Probability mass function, moment generating and characteristic function, moments, covariance and correlation, distribution fitting and examples.

Unit2: Extreme Value Distributions   [4 Hrs.]
Probability density, distribution functions, moments, properties and examples.

Unit3: Generalized Power Series Distribution   [6Hrs.]
Unified Probability mass function, it's special cases - Binomial, Poisson, Negative Binomial distributions and examples.

Unit4: Prior and Posterior Distributions   [6Hrs.]
Meaning and examples including cases where Binomial, Beta, Exponential, Gamma, Poisson, Negative Binomial distributions and examples.

Unit 5: Compound and Mixed Type Distribution   [6 Hrs.]
Compound Negative Exponential Distribution: Compounding of distributions, its moments.
Mixed Type Distribution: Mixed random variable, meaning and examples, computation of moments of mixed random variables, examples.

Unit6: Non-Parametric Tests   [11Hrs.]

An overview of parametric tests, need of non-parametric statistical tests, Wilcoxon-Mann-Whitney U test, Median test, Fisher exact test for 2×2 tables, median test, Wilcoxon Sign ranks test, McNemar test, Kruskal-Wallis one-Way Analysis of Variance, Kolmogorov-Smirnov one sample and two sample tests, Friedman two way analysis of variance, relevant examples.

Unit 7: Testing of Hypothesis   [11Hrs.]

General concept of simple and composite hypothesis, two types of errors, level of significance, power and size of a test.Most powerful test – Neymann Pearson's lemma and its application.Uniformly most powerful test- application to standard statistical distribution, unbiased test. Likelihood ratio test - Principle and properties, likelihood ratio test for testing means and variance in exponential families (without derivation), relevant examples.

Laboratory Works:

The applications of different probability distributions, testing of hypothesis using different statistical tests in real life data will be performed using appropriate software.

References:

1. Biswas, S. (1991). *Topics in Statistical Methodology.* India : Wiley Eastern
2. Chandra, T.K. and Chatterjee, D. (2003). *A First Course in Probability.* India: Narosa Publishing House.
3. Hoel, P.G., Port, S.C. and Stone, C.J. (1971). *Introduction to Probability Theory.* New Delhi India:Universal Book Stall.
4. Hogg, R.V. and Tanis, E.A. (2001). *Probability and Statistical Inference.* India: Pearson Education.
5. Kale, B.K. (1999). *A First Course on Parametric Inference.* Nindia: Narosa Publishing House.
6. Lehmann E.L. (1986). *Testing Statistical Hypotheses.* John Wiley and Sons.
7. Meyer, P.L. (1970). *Introductory Probability and Statistical Applications.* USA: Addison-Wesley.
8. Rohatgi, V.K. and Saleh, A.K.Md.E. (2005). *An Introduction to Probability and Statistics.* Singapore: John Wiley and Sons.
9. Shrestha, S. L. (2011). *Probability and Probability Distributions.* Kathmandu Nepal: S. Shrestha.
10. Zacks,S. (1971). *Theory of Statistical Inference.* John Wiley and Sons.

***