Instructor Name: Abhra Sarkar

Course Name: Statistical Methods for Categorical Data - Logistic regression and Beyond

Course Description: The course will cover traditional general-purpose models for categorical data, including logistic and probit binomial and multinomial regression models, Markov chains etc. which will then provide a background for more advanced methods suited to address modern era complex and high dimensional challenges.

The first two days will develop the background on general methods for (high-dimensional) linear and generalized linear (mixed) models. Topics to be covered include ordinary least squares, maximum likelihood estimation, ridge regression, LASSO, elastic net etc. The final two days will focus on adapting and applying these techniques to categorical data. In each case, relevant R packages will also be discussed in detail.

Course Outline:

Please complete an outline of topics you plan to cover for each day of the institute.

Day 1: Linear models, MLE, OLS, Ridge regression

Day 2: LASSO, Elastic net, REML estimation, mixed models

Day 3: Logit and probit (mixed) models, ROC curves, multinomial logit models, logit mixed models

Day 4: Modeling categorical sequences using Markov chains, etc.