Stanford University
Department of Statistics

4:15pm, Tuesday, July 28, 2015
Sequoia Hall Room 200
Cookies served at 3:45pm, 1st Floor Lounge.

Speaker:  Bhaswar B. Bhattacharya, Stanford University

Title:  Power of Graph-Based Two-Sample Tests

Abstract

Testing equality of two multivariate distributions is a classical problem for which many non-parametric tests have been proposed over the years. Most of the popular tests are based either on similarity graphs constructed using inter-point distances between the observations (multivariate generalizations of the Wald-Wolfowitz’s runs test) or on multivariate data-depth (generalizations of the Mann-Whitney rank test). These tests are known to be asymptotically normal under the null and consistent against all fixed alternatives.

In this talk, a general framework of graph-based tests will be introduced that includes all these tests. The asymptotic efficiency of a general graph-based test can be derived using Le Cam’s theory of local asymptotic normality, which provides a theoretical basis for comparing the performance of these tests. As a consequence, it will be shown that popular tests based on similarity graphs such as the Friedman-Rafsky test (1979), the test based on the K-nearest neighbor graph (1984), the minimum matching test of Rosenbaum (2005), among others have zero asymptotic (Pitman) efficiency against $O(N^{-\frac{1}{2}})$ alternatives. On the other hand, the tests based on multivariate depth functions (the Liu-Singh rank sum statistic (1993)), which include the Tukey depth (1975) and the projection depth (2003), have non-zero asymptotic efficiencies; though they might be computationally expensive when the dimension is large.

Finally, the limiting normal distribution of tests based on K-nearest neighbor graphs and other stabilizing random geometric graphs will be derived in the Poissonized setting. This shows that two-sample tests based on these graphs are powerful against $O(N^{-\frac{3}{4}})$ alternatives, that is, they have second-order efficiencies.