RUTGERS UNIVERSITY
DEPARTMENT OF STATISTICS AND BIOSTATISTICS
HILL CENTER #501, BUSCH CAMPUS, PISCATAWAY

www.stat.rutgers.edu

Seminar

Speaker: Shaw-Hua Lo, Department of Statistics, Columbia University

Title: Discovering Influential Variables: A Method of Partitions

Date: Wednesday October 1, 2008

Time: 3:20 PM

Place: 552 Hill Center

Abstract

We shall introduce a general computer intensive approach, based on a method proposed earlier by us for detecting which, of many potential explanatory variables, have an influence on a dependent variable Y. This approach is suited to detect influential variables, where causal effects depend on the confluence of values with other variables. It has the advantage of avoiding a difficult direct analysis involving possibly thousands of variables, by dealing with many randomly selected small subsets. The main objective is to discover the influential variables, rather than to measure their effects. Once they are detected, the problem of dealing with a much smaller group of influential variables should be vulnerable to appropriate analysis. In a sense, we are confining our attention to locating a few needles in a haystack. If time permits, we shall include a real application by using a variation of proposed methods to a case-control sporadic breast cancer study/data. Interactions of gene pairs associated with breast cancer are reported (PNAS, Aug. 2008).