COLLOQUIUM

Department of Statistics and Probability
Michigan State University

Xihong Lin
Harvard School of Public Health

Statistical Methods for Rare Variant Association Testing for Sequencing Data

Tuesday, April 19, 2011
A405 Wells Hall
10:20 a.m. - 11:10 a.m.
Refreshments: 10:00 a.m.

Abstract

Sequencing studies are increasingly being conducted to identify rare variants associated with complex traits. The limited power of classical single marker association analysis for rare variants poses a central challenge in such studies. We propose the sequence kernel association test (SKAT), a supervised, flexible, computationally efficient regression method to test for association between genetic variants (common and rare) in a region and a continuous or dichotomous trait, while easily adjusting for covariates. As a score-based variance component test, SKAT can quickly calculate p-values analytically by fitting the null model containing only the covariates, and so can easily be applied to genome-wide data. Using SKAT to analyze a genome-wide sequencing study of 1000 individuals, by segmenting the whole genome into 30kb regions, requires only 7 hours on a laptop. Through analysis of simulated data across a wide range of practical scenarios and triglyceride data from the Dallas Heart Study, we show that SKAT can substantially outperform several alternative rare-variant association tests. We also provide analytic power and sample size calculations to help design candidate gene, whole exome, and whole genome sequence association studies.

To request an interpreter or other accomodations for people with disabilities, please call the Department of Statistics and Probability at 517-355-9589.