



THE UNIVERSITY OF CHICAGO

Department of Statistics

DISSERTATION PRESENTATION AND DEFENSE

DUO JIANG

Department of Statistics
The University of Chicago

Statistical Methods for Genetic Association Analysis in Samples with Related Individuals and Population Structure

TUESDAY, July 1, 2014, at 2:30 PM
117 Eckhart Hall, 5734 S. University Avenue

ABSTRACT

In this dissertation, we develop statistical methods for analyzing data from genetics association studies in the presence of sample structure. The recent development of high-throughput sequencing technologies calls for powerful statistical tests to detect rare genetic variants associated with complex human traits. Recognizing the limitations of existing rare-variant association tests for family data, we propose MONSTER, a robust rare-variant association test, which generalizes the SKAT-O method for independent samples. MONSTER uses a mixed effects model that accounts for covariates and family correlation, and simultaneously tests for joint association with a group of genetic variants. To obtain a powerful test, MONSTER adaptively adjusts to the unknown configuration of effects of genetic variants. MONSTER also offers an analytical way of assessing p -values, which is desirable because permutation is not straightforward to conduct in dependent samples. In simulation studies, we demonstrate that MONSTER effectively accounts for family structure, is computationally efficient and compares very favorably, in terms of power, to previously proposed tests that allow related individuals. We further illustrate the proposed approach using a candidate gene study for high-density lipoprotein cholesterol in the Framingham Heart Study data. We also consider association testing for a binary trait in samples with population structure. Many recently proposed methods to account for population structure are based on the linear mixed model approach, which is primarily designed for quantitative traits. We develop a method that assumes a quasi-likelihood framework for correlated binary observations, where population structure is accounted for using a covariance matrix estimated from genome-wide data. The testing method assesses significance through a retrospective approach by modeling the genotypes as random. Compared with previous methods for population structure, our approach exploits the dichotomous nature of the trait, and features the ability to adjust for covariates and efficient computation. Simulation studies demonstrate that our method properly controls for population structure including stratification and admixture, and outperforms the linear mixed model approach in a wide range of settings.

For information about building access for persons with disabilities, please contact Matt Johnston at 773.702-0541 or send an email to mhj@galton.uchicago.edu. If you wish to subscribe to our email list, please visit the following web site: <https://lists.uchicago.edu/web/arc/statseminars>.