Methods for Genetic Association

Chair: Angelo Canty (McMaster University)

ANTONIO CIAMPI, McGill University
The Role of Statistical Learning Methods in Genetic Association Studies

Genome-wide association studies (GWAS) often use univariate tests to identify associations between SNPs and traits, although biology suggests that there must be interactions between genes. There exist powerful but rarely-used methods originating from the statistical learning literature that may assist in identifying multi-SNP relationships and gaining understanding of SNP-trait associations. We demonstrate the performance of such methods on sets of SNPs within 100kb of approximately 20 candidate genes associated with bone density measures.

BINOD NEUPANE, McMaster University
Multivariate Meta-Analysis in Genetic Association Studies

In meta-analysis, multivariate approaches might be preferable over univariate methods for estimating parameters of interest based on correlated phenotypes. In genetic association studies, missing data are less likely while associations are generally very small to modest and could be heterogeneous, so power to detect such associations as well as their unbiased and precise estimations are the concerns. We carried out simulations to investigate performance of multivariate meta-analytic methods. Our preliminary results suggest that multivariate techniques perform better in terms of power, coverage and bias in some scenarios. We will present detailed results of our comparative analysis for different realistic scenarios.

JULIA TALEBAN, Samuel Lunenfeld Research Institute
Bootstrap Bias-reduction in Genetic Association Analysis of Time-to-event Outcomes

While it is necessary to control the false-positive error rate in genome-wide studies, selection of genetic markers (SNPs) with small association p-values introduces optimistic bias into parameter estimates. Motivated by an investigation of complications in diabetes, we extend a bootstrap-resampling method for quantitative/binary traits to analyse time to nephropathy under a proportional hazards model. We evaluate relative and absolute bias via simulations. For SNPs with low to moderate power, bootstrap estimates are closer to the truth than uncorrected estimates, but the method tends to over-correct when power is high. Among false positives, however, bias-reduction shrinks estimates appropriately toward the null.

ANDRIY DERKACH, University of Toronto
Combining Linear and Quadratic Tests for Rare Variants Provides a Robust Test Across Genetic Models

Rare variants play an important role in complex human diseases and traits. Although many association tests have been proposed for rare variants, there is much confusion about the practical choice of a good test. Recent evaluations categorize methods into two classes: linear statistics sensitive to specific directional alternatives or the omnibus quadratic statistics. However, neither type of tests consistently outperforms the other. To achieve robustness, we consider the minimal p-value approach and the Fisher's method of combining p-values from linear and quadratic tests. Analytical and extensive simulation studies show that both methods are robust across genetic models with varying parameters.

ZHIJIAN CHEN, Samuel Lunenfeld Research Institute
Sequential Two-Phase Stratified Designs for Regional Sequencing Following Genome-Wide Association Study: A Bayesian Approach

In focused follow-up studies, investigators may choose to sequence an entire genomic region of interest at the base pair level to identify potential causal variants. We consider Bayesian sequential two-phase designs, in which a subset of phase 1 subjects

are selected from genotype categories of a genetic marker and sequenced in a target region in phase 2. At each sampling point, we apply Bayesian model averaging to account for genetic model uncertainty with an adaptive sample size allocation strategy to improve estimation efficiency of the genetic association parameter. We present numerical results for quantitative and binary traits.

ZEYNEP BASKURT, University of Toronto

A Composite Likelihood Approach Using the Evidential Paradigm to Analyze Genetic Association in Pedigrees

Royall (1997) proposed the evidential paradigm, an alternative to Frequentist and Bayesian paradigms for interpreting data as evidence. The evidential paradigm uses the likelihood ratio (LR) for two simple hypotheses as an objective measure of the strength of statistical evidence. In genetic studies, LRs are commonly used to measure evidence. However, evidential association of pedigrees is not straightforward due to complex family structures. We propose to use composite likelihoods to construct LRs for evidential analysis of families. We show how to make these LRs robust from model misspecification; that they have good operational characteristics; and are consistent with competing methods.