



The University of Chicago
Department of Statistics

Seminars for Fourth Year Ph.D. Students

ZUOHENG WANG
Department of Statistics
The University of Chicago

**An Incomplete-data Quasi-likelihood Framework with
Application to Genetic Association Studies on Related Individuals**

THURSDAY, November 29, 2007 at 3:30 PM
110 Eckhart Hall, 5734 S. University Avenue

ABSTRACT

Haplotype-based association analysis has been widely used in case-control studies, with haplotypes potentially providing more information on untyped variants as well as on interactions among tightly-linked typed variants. We consider the problems of testing association between a binary trait and one or more haplotypes, as well as testing for association of a trait with an untyped SNP, based on data from multiple tightly-linked genetic markers, using case-control samples containing arbitrary combinations of related and unrelated individuals with relationships specified by known pedigrees. Such samples commonly arise when families sampled for a linkage study are included in an association study. Current routine genotyping methods typically do not provide haplotype information; only the unphased genotype is directly observable. Statistically, this is a missing-data problem. Sampling multiple members from a pedigree provides additional haplotype information, but it also creates dependence among the genotypes of related individuals. This presents a dependent-data problem. We propose an incomplete-data quasi-likelihood (IQL) framework to accommodate both missing and dependent data. By forming a quasi-likelihood score function based on the conditional expectation of the marginal natural minimal sufficient statistic, we draw inference without specifying the full joint distribution. The resulting IQL score function retains the optimality properties of the quasi-likelihood approach. The consistency and asymptotic normality of the IQL estimator are established. Simulation studies are conducted to evaluate the power and type I error of our proposed method for genetic association testing with haplotypes on related individuals. And the application of the IQL score method to test for association with untyped SNPs using multiple typed SNPs in genome-wide association studies will be discussed.