

Unifying Stratified and Weighted FDR Methods with Applications to Large-Scale Genetic Studies

by

Lei Sun

University of Toronto

155 College Street

Toronto, ON M5T 3M7, Canada

`sun@utstat.toronto.edu`

Abstract

A central issue in high-dimensional genetic studies is how to assess statistical significance taking into account the inherent large-scale multiple hypothesis testing. To improve power, a number of studies have investigated the benefits of utilizing available prior information, however, the relative merits of different methods remain unknown. We focus on the stratified FDR (Sun et al., 2006) and weighted FDR (Genovese et al., 2006; Roeder et al., 2006) control methods. The two approaches model the prior distinctively. Weighted FDR converts the available prior information to test-specific weighting factor and adjusts the p-values accordingly. In contrast, stratified FDR divides tests into several disjoint strata based on the prior information and applies FDR control separately in each stratum. We first unify the two approaches in one framework and we show the trade-off between power and robustness by theoretical, simulation, and application studies. Robustness is desirable to safeguard against potential uninformative or even misleading prior information. We demonstrate the practical relevance by applying the two methods to three genome-wide association studies on diabetes and diabetes-related complications using previous genome-wide linkage results as the available prior information.

This is joint work with Yun Joo Yoo, Shelley Bull, Andrew Paterson and Daryl Waggott.