

MULTIPLE TESTING OF ONE-SIDED HYPOTHESES UNDER UNKNOWN DEPENDENCE

Seonghun Cho¹, Youngraee Kim², Johan Lim², Hyungwon Choi³,
DoHwan Park⁴ and Woncheol Jang^{*2}

¹*Inha University*, ²*Seoul National University*,
³*National University of Singapore*
and ⁴*University of Maryland at Baltimore County*

Abstract: The one-sided hypotheses in a multiple testing problem make the empirical null distribution (or p -values) conservative. Furthermore, it introduces a significant loss of power if not appropriately considered. We propose a multiple testing procedure named discarding adaptively with bounding on principal factor approximation (DAB-PFA) to simultaneously test a number of one-sided hypotheses under the general dependency of test statistics. Specifically, we use the principal factor approximation (PFA) by Fan and Han (2017) to account for the dependence structure among test statistics and adaptively discard small or large p -values when estimating the realized false discovery proportion (FDP). We derive the convergence rate of the proposed estimator and numerically compare the false discovery rate (FDR) and the true positive rate (TPR) of our method to many existing procedures, including those from Benjamini and Hochberg (1995), Efron (2004), and Wang and Fan (2017). We demonstrate our method through simulation studies and analysis of protein phosphorylation levels for serous ovarian adenocarcinoma samples.

Key words and phrases: Conservative null, discarding adaptively with bounding (DAB), false discovery rate, multiple testing, one-sided hypothesis, principal factor approximation.

1. Introduction

1.1. Multiple testing

In the last two decades, multiple testing, in which many hypotheses are tested simultaneously, has been one of the few central topics of statistics. The multiple testing problem arises in various applications, including microarray analysis in genetics, functional magnetic resonance imaging (fMRI) studies of the brain, clinical trials with multiple endpoints, and tens of thousands of A/B tests performed by major internet companies. Early works on this topic introduced various type I errors, including the family-wise error rate (FWER), the generalized FWER, the false discovery rate (FDR), the positive FDR, and proposed procedures to control the aforementioned errors at the nominal level.

^{*}Corresponding author. E-mail: wjang@snu.ac.kr