Journal of Statistical Planning and Inference 199 (2019) 219–235

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Group sequential BH and its adaptive versions controlling the FDR

Sanat K. Sarkar [a,*], Aiying Chen [b], Li He [c], Wenge Guo [d]

[a] Temple University, United States
[b] Sanofi Pasteur, United States
[c] Merck Research Laboratories, United States
[d] New Jersey Institute of Technology, United States

ARTICLE INFO

Article history:
Received 17 November 2017
Received in revised form 30 May 2018
Accepted 4 July 2018
Available online 19 July 2018

Keywords:
False discovery rate
Multiple testing
Group sequential design
BH procedure
Adaptive procedure
Positive dependence

ABSTRACT

This paper considers the problem of simultaneous testing of multiple hypotheses in a multi-stage group sequential setting subject to control over the false discovery rate (FDR). A multi-stage group sequential form of the BH procedure is developed, and a proof of its FDR control for *p*-values satisfying a positive dependence condition both between and within stages is given. This group sequential BH is adapted to the proportion of true nulls in two different ways, resulting in the proposal of two adaptive group sequential BH. While one of these adaptive procedures is theoretically shown to control its FDR when the *p*-values are positively dependent between but independent within stages, the other one's FDR control is assessed through simulations. Comparative performance studies of the proposed procedures in terms of FDR control, power, and proportion of sample saved carried out through extensive simulations provide evidence of superior performance of the proposed adaptive procedures.

© 2018 Published by Elsevier B.V.

1. Introduction

In many modern scientific investigations, such as those in gene and protein expression studies where thousands of genes are tested for possible association with some disease condition, and in pharmacogenetics research where genetic contributions are studied in evaluating safety and efficacy of drugs, questions are investigated through large-scale, long-term follow up studies. For economic benefits and safety reasons, data are often accrued sequentially in these studies allowing interim analyses to be performed for making early decisions. Statistical analyses of data in these studies often involve simultaneous testing of a large number of hypotheses, making multiple testing in a sequential framework involving multiple stages a frequently arising statistical problem. This brings newer challenge for developing large-scale multiple testing method that is applicable to a sequential setting with multiple stages and controls an appropriate error rate such as the false discovery rate (FDR), which is the expected proportion of false rejections out of the total number of rejections.

The notion of FDR has been introduced by Benjamini and Hochberg (1995), along with a powerful and easy-to-use method, known as the BH method, that controls the FDR when multiple hypotheses are simultaneously tested in a non-sequential or single-stage setting. Given a set of m null hypotheses $H_1, \ldots, H_m$ to be simultaneously tested using their respective p-values $P_1, \ldots, P_m$, the level α BH method is a step-up method with the critical constants $\lambda_i = i\alpha/m$, $i = 1, \ldots, m$; that is, it rejects H_i for all i such that $P_i \le P_{(R)}$, where $R = \max\{i : P_{(i)} \le \lambda_i\}$, provided the maximum exists, otherwise, it rejects none, with $P_{(1)} \le \cdots \le P_{(m)}$ being the ordered values of the P_i's. Benjamini and Hochberg (1995)

* Corresponding author.
E-mail addresses: sanat@temple.edu (S.K. Sarkar), aiying.chen@sanofipasteur.com (A. Chen), li.he@merck.com (L. He), wenge.guo@njit.edu (W. Guo).

https://doi.org/10.1016/j.jspi.2018.07.001
0378-3758/© 2018 Published by Elsevier B.V.

showed that the FDR of their method is less than or equal to $m_0\alpha/m$, where m_0 is the number of true null hypotheses, and hence the FDR is controlled at α, when the p-values are independent. Later on, Benjamini and Yekutieli (2001), Finner and Roters (2001), Sarkar (2002), and Storey et al. (2004) proved that the BH method's FDR under independence of the p-values is actually exactly equal to $m_0\alpha/m$. Benjamini and Yekutieli (2001) and Sarkar (2002) further showed that the FDR of the BH method is less than or equal to $m_0\alpha/m$ under a form of positive dependence condition that is shared by p-values in many multiple testing situations; see also Finner et al. (2009) and Sarkar (2008). The BH method, as a single-stage multiple testing method, has gained much popularity because of its applicability to a wide variety of scientific investigations and other desirable theoretical properties. Benjamini and Hochberg (2000) suggested the idea of adapting it to the data in an attempt to tighten its FDR control by estimating m_0 from the data and incorporating the estimate into it. They, of course, did not offer any theoretical proof of its ultimate FDR control. Several such adaptive versions of the BH method utilizing a wide variety of estimators for π_0, with theoretical proofs of their FDR control being offered only under independence, have been put forward in the literature. See, for example, Storey (2002), Storey et al. (2004), Benjamini et al. (2006), Gavrilov et al. (2009), Blanchard and Roquain (2009), and Sarkar (2008).

Papers dealing with FDR control in multi-stage statistical experiments involving simultaneous testing of multiple hypotheses do exist in the literature prior to this work.

Benjamini and Yekutieli (2005) introduced a two-stage procedure in the context of quantitative trait locus (QTL) mapping analysis, where the BH procedure is applied at each stage using the data available only at that stage and the hypotheses rejected at the first stage are further tested at the second stage. With m_0 being the number of true nulls and α_1 and α_2 being chosen differently for the first and second stages respectively, it controls the FDR at level $m_0\alpha_1\alpha_2/m$ under the same positive dependence condition for which the BH procedure is known to control the FDR.

Zehetmayer and coauthors in a sequence of papers considered the problem of controlling the FDR in the context of gene association or gene expression studies under a sequential setting when the test statistics are normally and independently distributed. Zehetmayer et al. (2005) considered a two-stage design where the total number of observations is fixed with certain fraction of these observations allocated to the first stage and the remaining observations distributed among the hypotheses whose first stage p-values are less than or equal to a prefixed futility boundary. They defined a sequential p-value and derived an estimation based approach to controlling the FDR following Storey (2002). Specifically, an estimate of the FDR is derived using the sequential p-values and a rejection threshold is chosen so that this estimate is less than or equal to a nominal level α. The hypotheses whose sequential p-values are less than or equal to this rejection threshold are then rejected. This approach was later extended by Zehetmayer et al. (2008) to control the FDR under multi-stage designs with fixed stagewise sample sizes as well as under multi-stage designs where the overall number of observations is fixed and at each stage a pre-specified fraction of observations is evenly distributed among the selected hypotheses according to some futility boundaries. Zehetmayer and Posch (2012) further assessed the following selection methods used at the first stage in a two-stage design — (i) the hypotheses whose first stage p-values are less than or equal to some prefixed futility boundary are selected; (ii) a prefixed number of most significant hypotheses are selected; and (iii) the hypotheses rejected by the BH procedure in Benjamini and Yekutieli (2005) at some fixed level α_1 are selected — with simulation studies showing that the FDR is controlled in all these scenarios. Zehetmayer et al. (2015) further proposed a sample size reassessment procedure controlling the FDR under a two-stage design. Based on the data available at the first stage, they derived an asymptotic expression of a selected power measure and determined a sample size for the second stage so that the power of the FDR controlling procedure is at some specified value. Their simulation results showed that their sample size reassessment procedure controls the FDR despite the data dependent choice of sample size.

Victor and Hommel (2007) also considered extending the BH procedure. Focusing on a two-stage adaptive design, they derived a method using a generalized definition of sequential p-value allowing for both early rejection and early acceptance at the interim analysis. They noted, however, that unlike in single testing where a futility boundary can often be determined based on sample size considerations, the choice of futility boundaries in multiple testing is often challenging since the joint distribution of the underlying test statistics is rarely known.

By utilizing sequential p-values, Malek et al. (2017) proposed a sequential conversion method to transform a fixed sample multiple testing procedure controlling some type I error rate, such as the FDR, to a sequential multiple testing procedure that still controls the same error rate. Specifically, their method applies the fixed sample multiple testing procedure, for example, the BH method, to the sequential p-values at each stage and allows an early rejection once sufficient evidence has accumulated against the null hypothesis.

Sarkar et al. (2013) extended the BH method and its adaptive version from single- to a two-stage adaptive design setting. More specifically, they considered screening the null hypotheses sequentially at the first stage as being rejected or accepted subject to certain boundaries on the FDR across all hypotheses and testing the remaining null hypotheses at the second stage having combined their p-values from the two stages using some combination function. These methods were theoretically proved to control the FDR under the assumption that the pairs of first and second-stage p-values across all hypotheses are independent and those which correspond to the null hypotheses are identically distributed as a pair (p_1, p_2) satisfying the p-clud property of Brannath et al. (2002). Bartroff and Song (2013) further considered the problem of developing a multi-stage FDR controlling procedure and developed such a procedure by appropriately adjusting the BH critical values at each stage, and assumed independence of the p-values across the hypotheses to prove its FDR control. However, in most studies involving group sequential design, the p-values are rarely independent across hypotheses, just as in the case of fixed sample design, and the underlying dependence structure can often be characterized by assuming a positive dependence condition.

Download English Version:

https://daneshyari.com/en/article/11020307

Download Persian Version:

https://daneshyari.com/article/11020307

Daneshyari.com