

Adapting Benjamini-Hochberg by Simes Inequality

Una modifica della procedura di Benjamini-Hochberg basata sulla diseguaglianza di Simes

Aldo Solari and Jelle J. Goeman

Abstract Benjamini and Hochberg introduced, in 1995, a simple but very successful procedure which bounds the False Discovery Rate at $\alpha\pi_0$, where π_0 is the unknown proportion of true null hypotheses [1]. Since that time, various Authors have tried to achieve adaptive control of the FDR by incorporating an estimate of π_0 into the Benjamini and Hochberg procedure.

We propose a two-stage procedure in which Simes inequality [2] is used to derive an upper bound for π_0 in the first stage, providing an adjusted level for use of the Benjamini and Hochberg procedure in the second stage. The advantage of this approach lies in the link between Simes inequality and Benjamini-Hochberg procedure, thereby making the first stage a very natural way to estimate π_0 .

Abstract Benjamini e Hochberg hanno introdotto nel 1995 una semplice procedura capace di controllare il False Discovery Rate al livello $\alpha\pi_0$, dove π_0 rappresenta la proporzione di ipotesi nulle vere [1]. Da allora, diversi Autori hanno proposto di modificare la procedura di Benjamini-Hochberg incorporando una stima di π_0 nella procedura. Proponiamo di modificare la procedura di Benjamini-Hochberg utilizzando la diseguaglianza di Simes [2] per stimare π_0 .

Key words: Adaptive procedure, False Discovery Rate, Simes Inequality

1 Introduction

The seminal paper in which Benjamini and Hochberg [1] introduced the concept of the False Discovery Rate (FDR) has changed thinking about multiple testing quite

Aldo Solari
University of Milano-Bicocca, Italy. e-mail: aldo.solari@unimib.it

Jelle J. Goeman
Radboud University Medical Centre, The Netherlands. e-mail: Jelle.Goeman@radboudumc.nl

radically, showing that Familywise Error Rate (FWER) control is not only way to do of multiple testing, and stimulating the field of multiple testing enormously [3].

Suppose we have a collection $H_1, \dots, H_m$ of hypotheses of interest, and an unknown number m_0 of these hypotheses is true, while the other $m_1 = m - m_0$ is false. We denote the proportion of true hypotheses $\pi_0 = m_0/m$. The goal of a multiple testing procedure is to choose a collection of hypotheses to reject. Let V denote the number of true null hypotheses that are erroneously rejected and let R be the total number of hypotheses that are rejected. Then, the FDR is defined by $E(V/R \vee 1)$, where $R \vee 1 = \max\{R, 1\}$.

Suppose we have p -values $p_1, \dots, p_m$ for each of the hypotheses $H_1, \dots, H_m$, and let $p_{(1)} \leq \dots \leq p_{(m)}$ be their ordered values. The Benjamini & Hochberg (BH) procedure is a linear step-up procedure that compares each ordered p -value $p_{(i)}$ with the critical value $i\alpha/m$, finds the largest j such that $p_{(j)}$ is smaller than its corresponding critical value, and rejects the j hypotheses with the j smallest p -values.

The BH procedure controls the FDR at a desired level $\alpha\pi_0$ when the p -values are independent or positively dependent. This suggests the possibility of an alternative, more powerful BH procedure that uses critical values $i\alpha/(\hat{\pi}_0 m)$ rather than $i\alpha/m$ if a good estimate $\hat{\pi}_0$ of π_0 would be available. Such a procedure might have an FDR closer to the chosen level α , and would be even more powerful than the original procedure if many hypotheses were false.

Such procedures are called *adaptive* procedures, and many have been proposed based on various estimates of π_0 . A problem with the adaptive approach, however, is that estimates of π_0 can have high variance, especially if p -values are strongly correlated. Naive plug-in procedures, in which this variance is not taken into account, will therefore generally not have FDR control, especially if $\pi_0 \approx 1$. More sophisticated methods are needed that do take the estimation error of π_0 into account. One such procedure, by Benjamini, Krieger and Yekutieli (BKY) [4], adjusts the α -level slightly from α to $\alpha^* = \alpha/(1 + \alpha)$ to adjust for the additional variance from estimation of π_0 . This procedure estimates π_0 by first performing an initial BH procedure at the slightly reduced level α^* , estimating π_0 by $\hat{\pi}_0 = (m - \hat{m}_1)/m$, where $\hat{m}_1$ is the number of rejections obtained in this first step. In the second and final step, a subsequent BH procedure is done at level $\alpha^*/\hat{\pi}_0$. Note that the BYK procedure is not guaranteed to give more rejections than the regular nonadaptive BH procedure, because $\alpha^*/\hat{\pi}_0$ may be smaller than α . FDR control for the BKY procedure has only yet been proven under independence, although simulations suggest FDR control under positive dependence as well.

2 A simple improvement based on Simes inequality

The BH method is closely related to Simes inequality [2], which says that

$$P\left(\bigcap_{i=1}^m \left\{p_{(i)} > \frac{i\alpha}{m}\right\}\right) \geq 1 - \alpha. \quad (1)$$

The close connection between the BH method and Simes inequality is immediately obvious from the fact that the BH and Simes critical values are identical. Weak control of FWER, a necessary condition for FDR control for the BH method follows immediately from Simes' inequality.

Consider the following adaptive BH procedure based on Simes inequality (BH-S in short):

1. If Simes inequality in (1) is not violated, i.e. $p_{(i)} > i\alpha/m$ for all $i = 1, \dots, m$, then accept all the hypotheses.
2. If Simes inequality in (1) is violated, i.e. $p_{(i)} \leq i\alpha/m$ for at least one i , then perform the BH procedure at $\alpha m/(m-1) > \alpha$.

In the first stage, the BH-S procedure verifies whether Simes inequality in (1) is violated or not. Violation of the Simes inequality provides evidence that at least one hypothesis is false, i.e. $m_0 < m$ and thus π_0 is at most $(m-1)/m$. This additional information is used in the second stage by performing the BH procedure at $\alpha m/(m-1)$. The following theorem shows that the BH-S procedure controls the FDR at α under the same assumption of the BH procedure.

Theorem 1. *Under the same assumption of the BH procedure, the BH-S procedure controls the FDR at level α .*

Proof. It has been shown that the BH procedure controls the FDR under a positive dependence assumption. This assumption implies that Simes inequality holds for the p-values corresponding to the true null hypotheses. Thus, if $m_0 = m$, the event $\bigcup_{i=1}^m \{p_{(i)} \leq \frac{i\alpha}{m}\}$ happens with probability at most α by Simes inequality. This guarantees weak control of the FWER at α , and thus FDR control at α .

On the other hand, if $m_0 < m$, it is easy to see that performing the BH procedure at $\alpha m/(m-1)$ guarantees FDR control at α because the FDR is bounded by $\alpha m/(m-1) \times \pi_0 = \alpha m_0/(m-1) \leq \alpha$.

Theorem 1 shows that the BH-S procedure represents a negligible but uniform improvement over the BH procedure. In the same spirit, we can also adapt the Benjamini & Yekutieli procedure by using Rütgers inequality [5].

3 The BH-GS procedure

A more powerful two-stage adaptive BH procedure can be obtained by using the estimator of π_0 proposed by Goeman and Solari [6]. This estimator is based on the Closed Testing procedure which uses Simes local tests [5,6]:

$$\hat{\pi}_0 = \max \left\{ i \in \{1, \dots, m\} : p_{(m-i+k)} > \frac{k\alpha}{i} \text{ for } k = 1, \dots, i \right\} / m. \quad (2)$$

If the maximum does not exist, then set $\hat{\pi}_0 = 0$. Goeman and Solari [6] showed that

$$P(\hat{\pi}_0 \leq \pi_0) \leq \alpha. \quad (3)$$

when the p-values corresponding to true null hypotheses satisfy Simes' inequality. Using $\hat{\pi}_0$ in (2) yields the following adaptive BH procedure (BH-GS in short):

1. Compute $\hat{\pi}_0$ in (2).
2. If $\hat{\pi}_0 = 0$ reject all hypotheses; otherwise perform the BH procedure at $\alpha/\hat{\pi}_0$.

It is easy to see that, under the same assumption of the BH procedure, the BH-GS guarantees FDR control at α when $m_0 = m$. However, if $m_0 < m$, the BH-GS procedure is not guaranteed to have FDR control at α . Indeed, it can be shown that the BH-GS procedure has FDR equal to $\alpha + \alpha^2/4$ under the Dirac-Uniform configuration with $m = 3$ and $m_0 = 2$. In this case, to control the FDR at α , we have to perform the BH-GS at the adjusted level of $\alpha^* = 2 \cdot (-1 + \sqrt{1 + \alpha})$.

For the specific multiple problem described in detail in Benjamini and Hochberg [1], the significance of the treatment effect on each of the 15 endpoints is given by the ordered $p_{(i)}$ s: 0.0001, 0.0004, 0.0019, 0.0095, 0.0201, 0.0278, 0.0298, 0.0344, 0.0459, 0.3240, 0.4262, 0.5719, 0.6528, 0.7590 and 1.000. Four hypotheses were rejected using the BH procedure at level $\alpha = 0.05$. The BH-GS procedure estimates π_0 equal to 0.8 at the first stage, and uses the BH procedure at level $\alpha^*/0.8 = 0.06173769$ at the second stage, resulting in the rejection of five hypotheses.

References

1. Benjamini, Y., Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*. **57**, 280–300 (1995)
2. Simes, R.J.: An improved Bonferroni procedure for multiple tests of significance. *Biometrika*. **73**, 751–754 (1986).
3. Goeman, J.J., Solari, A.: Tutorial in biostatistics: multiple testing in genomics. *Statistics in Medicine*. **33**, 1–33 (2014).
4. Benjamini, Y., Krieger, A., Yekutieli, D.: Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*. **93**, 491–507 (2006).
5. Hommel G.: A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*. **75**, 383–386 (1988).
6. Goeman, J.J., Solari, A.: Multiple testing for exploratory research. *Statistical Science*. **26**, 584–597 (2011).