FDR Control in Ordered Hypothesis Testing

Jinxi Liu

October 18, 2017

The Question

Suppose that we have a sequence of null hypotheses, $H_1, H_2, ..., H_m$, and that we want to reject some hypotheses while controlling the False Discovery Rate. Moreover, suppose that these hypotheses must be rejected in an ordered fashion: a test procedure must reject hypotheses $H_1, ..., H_k$ for some $k \in \{0, 1, ..., m\}$.

A rejection rule in this setting amounts to a procedure for choosing the stopping point k.

Solution 1

Classical methods for FDR control, such as the original Benjamini-Hochberg selection procedure, are ruled out by the requirement that the hypotheses be rejected in order.

However, we might be able to transform the sequence of p-values $p_1, ..., p_m$ into a monotone increasing sequence of statistics $0 \leq q1 \leq ... \leq q_m \leq 1$. We then can achieve ordered FDR control by applying the original Benjamini-Hochberg procedure on the monotone test statistics q_i.

Solution 1: Inspiration

Rényi representation theorem

$$\left(\frac{Y_1}{m}, \frac{Y_1}{m} + \frac{Y_2}{m-1}, \ldots, \sum_{i=1}^{m} \frac{Y_i}{m-i+1}\right) \stackrel{d}{=} E_{1,m}, E_{2,m}, \ldots, E_{m,m},$$

where the $E_{i,m}$ are exponential order statistics, meaning that the $E_{i,m}$ have the same distribution as a sorted list of independent standard exponential random variables.

Solution 1: Inspiration

In our context, let

$$
\begin{aligned}
Y_i &= -log(1 - p_i), \\
z_i &= \sum_{j=1}^{i} Y_j/(m - j + 1), \textit{and} \\
q_i &= 1 - e^{-Z_i}.
\end{aligned}
$$

Under the global null, the Y_i are distributed as independent exponential random variables. Thus, by Rényi representation, the Z_i are distributed as exponential order statistics, and so the q_i are distributed like uniform order statistics.

Solution 1: ForwardStop Procedure

Procedure 1 (ForwardStop).

Let $p_1, .., p_m \in [0, 1]$, *and let* $0 < \alpha < 1$. *We reject hypotheses* $1, ..., \hat{k}_F$, *where*

$$\hat{k}_F = max \left\{ k \in \{1, .., m\} : \frac{1}{k} \sum_{i=1}^{k} Y_i \leq \alpha \right\},$$

and

$$Y_i = -log(1 - p_i).$$

This procedure is called ForwardStop because it scans the p-values in a forward manner: If $\frac{1}{k} \sum_{i=1}^{k} Y_i \leq \alpha$, then we can reject the first k hypotheses regardless of the remaining p-values.

Solution 1: ForwardStop Procedure

THEOREM

Suppose that we have p-values $p_1, ..., p_m \in (0, 1)$, a subset $N \subseteq \{1, .., m\}$ and are null i.e., independently drawn from $U([0, 1])$. Then, the ForwardStop procedure controls FDR at level α, meaning that

$$E\left[\left|\left\{1, .., \hat{k}_F\right\} \cap N\right| \Big/ max\left\{\hat{k}_F, 1\right\}\right] \leq \alpha$$

Note: ForwardStop provides FDR control even when some null hypotheses are interspersed among the non-null ones.

Solution 1: StrongStop Procedure

In ForwardStop procedure, we created the ordered test statistics Z_i by summing transformed p-values starting from the first p-value. Under the global null, we could just as well obtain uniform order statistics q_i by summing from the back:

$$\tilde{Y}_i = -log(1 - p_i),$$
$$\tilde{z}_i = \sum_{j=i}^{m} Y_j/j, \text{ and}$$
$$\tilde{q}_i = 1 - e^{-\tilde{z}_i}.$$

Solution 1: StrongStop Procedure

Procedure 2 (StrongStop).

Let $p_1, .., p_m \in [0, 1]$, and let $0 < \alpha < 1$. We reject hypotheses $1, ..., \hat{k}_S$, where

$$\hat{k}_S = max \left\{ k \in \{1, .., m\} : \tilde{q}_k \leq \frac{\alpha k}{m} \right\},$$

and $\tilde{q}_k$ is as defined in the above page.

Unlike ForwardStop, this new procedure needs to look at the p-values corresponding to the last hypotheses before it can choose to make any rejections.

Solution 1: StrongStop Procedure

THEOREM

Suppose that we have p-values $p_1, ..., p_m \in (0, 1)$, the last $m - s$ of which are null (i.e., independently drawn from $U([0, 1])$). Then the above rule controls the FWER at level α, meaning that

$$P\left[\hat{k}_S > s\right] \leq \alpha.$$

FWER control is stronger than FDR control, and so we immediately conclude that StrongStop also controls the FDR. Note that the guarantees only hold when the non-null p-values all precede the null ones.

Solution 1: Discussion

A major advantage of both ForwardStop and StrongStop is that these procedures seek the largest k at which an inequality holds, even if the inequality may not hold for some index l with $l < k$. This property enables them to get past some isolated large p-values for the early hypotheses, thus resulting in a substantial increase in power.

The methods in this article require that the null p-values be independent.

Solution 2

Lynch et al. (2016) discussed a conventional fixed sequence method that stops testing once an acceptance occurs, and develop such a method controlling the FDR under both arbitrary and negative dependencies. They also extend the conventional fixed sequence method to one that allows more but a given number of acceptances.

Solution 2: Assumptions

They assume that the true null p-values, $\hat{P}_i$, for $i = 1, ..., m_0$, are marginally distributed as follows:

$$Pr\left(\hat{P}_i \leq p\right) \leq p \text{ for any } p \in (0, 1).$$

Solution 2: Assumptions

Positive regression dependence(PRDS)

The vector of p-values $\overrightarrow{P}$ is PRDS on the vector of null p-values $\overrightarrow{P_0} = (\hat{P}_1, ..., \hat{P}_{m_0})$ if for every increasing set D and for each i = 1,...,m_0, the conditional probability Pr $\left(\overrightarrow{P} \in D | \hat{P}_i = p\right)$ non-decreasing in p.

Solution 2: Assumptions

Negative Association

The vector of p-values $\overrightarrow{P}$ is negatively associated with null p-values if for each i = 1,...,m_0, the following inequality holds:

$$Pr\left(\hat{P}_i \leq p_{\mu_i}, P_j \leq p_j, j = 1, ...m, with\ j \neq \mu_i\right) \leq$$
$$Pr\left(\hat{P}_i \leq p_{\mu_i}\right) Pr\left(P_j \leq p_j, j = 1, ...m, with\ j \neq \mu_i\right)$$

for all fixed p_j's.

Several multivariate distributions posses the conventional negative association property, including multivariate normal with non-positive correlation, multinomial, dirichlet, and multivariate hypergeometric. Independence is a special case of negative dependence.

Solution 2: Conventional Fixed Sequence Procedures

Consider a conventional fixed sequence procedure with critical constants

$$\alpha_i^{(1)} = min\left(\frac{m\alpha}{m-i+1}, 1\right), i = 1, ..., m.$$

(i) This procedure strongly controls the FDR at level α **under arbitrary dependence** of the p-values.

(ii) One **cannot** increase even one of the critical constants $\alpha_i^{(1)}$, i $= 1,...,m$, while keeping the remaining fixed without losing control of the FDR. This is true even when $\overrightarrow{P}$ is assumed to be PRDS on $\overrightarrow{P_0}$.

Solution 2: Conventional Fixed Sequence Procedures

When the p-values are **negatively associated**, the critical constants of the conventional fixed sequence procedure can be further improved as in the following: The conventional fixed sequence method with critical constants

$$\alpha_i^{(2)} = \frac{i\alpha}{1 + (i-1)\alpha}, i = 1, ..., m.$$

strongly controls the FDR at level α when the p-values are negatively associated on the true null p-values.

Solution 2: Fixed Sequence Procedures that Allow More Acceptances

A conventional fixed sequence method might potentially lose power if an early null hypothesis fails to be rejected, with the remaining hypotheses having no chance of being tested. To remedy this, they generalize a conventional fixed sequence method to one that allows a certain number of acceptances.

Solution 2: Fixed Sequence Procedures that Allow More Acceptances Under Arbitrary Dependence

The fixed sequence method stopping on the k_{th} acceptance with critical constants

$$\alpha_i^{(3)} = \begin{cases} \frac{\alpha}{k}, & \text{if } i = 1, ..., k \\ \frac{(m-k+1)\alpha}{(m-i+1)k}, & \text{if } i = k+1, ..., m \end{cases}$$

strongly controls the FDR at level α **under arbitrary dependence** of p-values.

Solution 2: Fixed Sequence Procedures that Allow More Acceptances

This procedure reduces to the Conventional Fixed Sequence Procedures when k $= 1$. However, we cannot prove that this procedure is optimal in the sense that its critical constants cannot be further improved without losing control of the FDR under arbitrary dependence of the *p*-values.

Under certain distributional assumptions on the p-values, the critical constants of this procedure can potentially be improved. In particular, we have the following result:

Solution 2: Fixed Sequence Procedures that Allow More Acceptances Under Independence

Consider a fixed sequence method stopping on the k_{th} acceptance with critical constants

$$\alpha_i^{(4)} = \frac{(r_{i-1}+1)\alpha}{k+(i-k)\alpha}, i = 1, ...m,$$

where r_i is the number of rejections among the first i tested hypotheses, with $r_0 = 0$, for i = 1,...m. This procedure strongly controls the FDR at level α if the true null p-values are mutually independent and are independent of the false null p-values.

Solution 2: Data Driven Ordering

The applicability of the aforementioned fixed sequence methods depends on availability of natural ordering structure among the hypotheses. When the hypotheses cannot be pre-ordered, one can use pilot data available to establish a good ordering among the hypotheses in some cases. For example, in replicated studies, the hypotheses for the follow-up study can be ordered using the data from the primary study. However, when prior information is unavailable ordering information can usually be assessed from the data itself.

Solution 2: Data Driven Ordering

Assume that the variables of interest are $X_i, i = 1, ..., m$, with n independent observations $X_{i1}, ...X_{in}$ on each X_i. An ordering statistic, Y_i, and a test statistic, T_i, are determined for each i = 1,...,m. The Y_i's are used to order all of the hypotheses $H_1, ..., H_m$, Ti is used to test the hypothesis H_i; i = 1,...,m, and P_i is the corresponding p-value. In addition, Y_i is chosen such that it is independent of the T_i's under H_i and tends to be larger as the effect size increases.

Solution 2: Data Driven Ordering

Example: Two sample T-test. Consider testing $H_i : \mu_i^{(1)} = \mu_i^{(2)}$ against $H_i' : \mu_i^{(1)} \neq \mu_i^{(2)}$ simultaneously using $n = n_1 + n_2$ data vectors. Suppose $X_{ij}^{(l)}, j = 1, ..., n_l$, follows a $N(\mu_i^{(l)}, \sigma^2)$ distribution, for l = 1, 2. Then, the hypotheses can be tested using the two-sample t-test statistics T_i and are ordered through the values of the total sum of squares, which is
$Y_i = \sum_{l=1}^{2} \sum_{j=1}^{n_l} X_{ij}^{(l)}/n$, for i = 1,...m.

The rationale behind this is independence between the Y_i's and T_i under H_i, and the following result:
$E[Y_i] = (n-1)\sigma^2 + n_1 n_2 (\mu_i^{(1)} - \mu_i^{(2)})^2/n$.

Extension: Controlling FDR for Testing Hierarchically Ordered Hypotheses

In many problems involving the testing of multiple hypotheses, the hypotheses have an intrinsic, hierarchical structure such as a tree-like or graphical structure.

Notations and Definitions

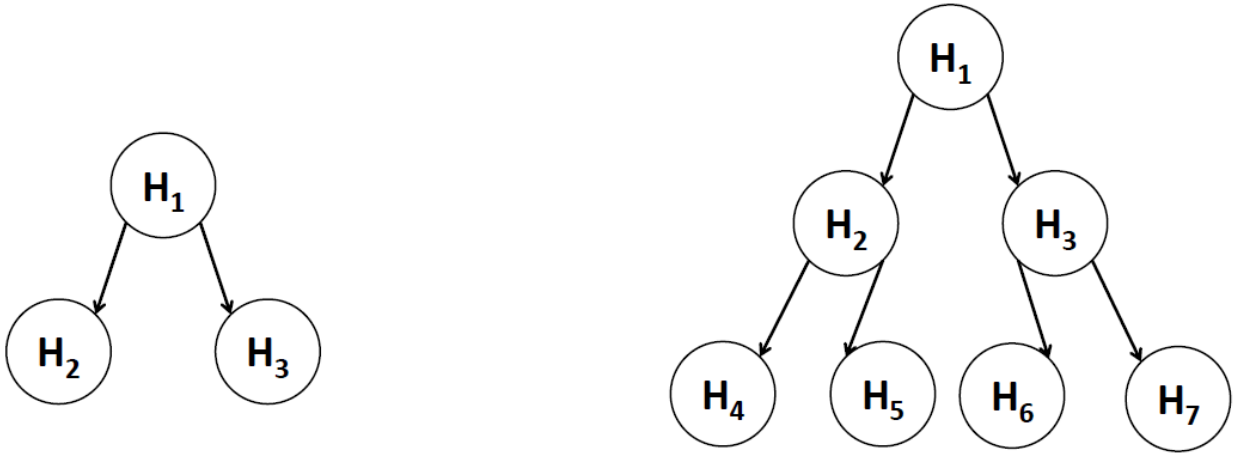

Figure 1: (a) An example of a hierarchical structure with 3 hypotheses for which H_2 and H_3 are only tested if H_1 is rejected. (b) An example of a hierarchical structure with 7 hypotheses for which H_2 and H_3 are only tested if H_1 is rejected, H_4 and H_5 are only tested if H_2 is rejected, and H_6 and H_7 are only tested if H_3 is rejected.

Notations and Definitions

1. Let M $= \{H_1, ..., H_m\}$ be the set of the m tested hypotheses.
2. Let T:{0,...,m} $\rightarrow$ {0,...,m} be a function that takes an index of a hypothesis and returns the index of the parent hypothesis with T(0) $= 0$.
3. D_i is the set of all ancestor hypotheses of H_i, which includes H_i.
4. d_i is the cardinality of D_i. The depth of H_i in the hierarchy if defined as d_i.
5. Let D be the maximum depth of the m hypotheses to be tested.
6. M_i is the set of all descendant hypotheses of H_i, which also includes H_i.
7. Let m_i be the cardinality of M_i.

Notations and Definitions

8. If $m_i = 1$, then H_i has no children and it is referred to as a leaf hypothesis.
9. Let l be the number of leaf hypotheses in the whole hierarchy and l_i be the number of leaf hypotheses in the subtree under H_i.
10. Hypotheses are grouped into D families by depth where family d contains all hypotheses with depth d. That is, $F_d = \{H_i \in M, : d_i = d\}$.

For example, in Figure 1 (a), $T(2) = T(3) = 1$ and H_2 and H_3 are leaf hypotheses. In Figure 1(b), $T(6) = T(7) = 3$; $D_6 = \{H_1, H_3, H_6\}$, $M_2 = \{H_2, H_4, H_5\}$, and $F_3 = \{H_4, H_5, H_6, H_7\}$.

Assumptions

Throughout this paper the authors make use of the following basic assumption regarding marginal p-values: for any p-value P_i with H_i being true,

$$Pr(P_i \leq p) \leq p \text{ for any } 0 \leq p \leq 1.$$

Assumptions

Positive Dependence Assumption *For any coordinatewise non-decreasing function of the p-values* ψ, $E(\psi(P_1, ..., P_m)|P_i \leq p)$ *is non-decreasing in* p *for each p-value* P_i *such that* H_i *is true.*

Positive Dependence Assumption is slightly more relaxed than the condition of positive regression dependence on a subset (PRDS).

Assumptions

Block Dependence Assumption *For each d =1,...,D, the p-values corresponding to the hypotheses in F_d are independent of the p-values corresponding to the hypotheses not in F_d.*

Block Dependence Assumption only characterizes the joint dependence of the p-values across families but does not describe the joint dependence within families.

Generalized Stepwise Procedure

Reference

1. G'Sell, Max Grazier, et al. "Sequential selection procedures and false discovery rate control." Journal of the royal statistical society: series B (statistical methodology) 78.2 (2016): 423-444.
2. Lynch, Gavin, et al. "The control of the false discovery rate in fixed sequence multiple testing." arXiv preprint arXiv:1611.03146 (2016).
3. Lynch, Gavin, and Wenge Guo. "On Procedures Controlling the FDR for Testing Hierarchically Ordered Hypotheses." arXiv preprint arXiv:1612.04467 (2016).