STAT201B Lecture 13 Multiple Hypothesis Testing

Logic

很多时候我们需要检验多个 hypotheses：

1. pairwise multiple hypothesis testing (逐个检验多个 hypotheses 是否成立)
   e.g. 判断病人与正常人之间, 具体哪些基因发生了差异表达
2. global testing / simultaneous multiple hypothesis testing (检验多个 hypotheses 是否同时成立)
   e.g. 判断病人与正常人之间, 是否有基因发生了差异表达

假设 significance level 为 α, 共有 m 个 (独立的) sub-hypotheses
倘若对每个 sub-hypothesis 都进行 size α test, 那么整体来说, 犯 type-I error 的概率是 $1-(1-\alpha)^m \to 1$ as $m \to \infty$

因此我们需要做出改进, 主要有以下思路:

1. 控制 FWER (Family-Wise Error Rate), 即 m 次检验中出现任何 type-I error 的概率, 为 α (保守)
2. 控制 FDR (False Detection Rate), 即允许出现 type-I error, type-I error 占显著结果的比例, 为 α (宽松)
3. 利用 sub-hypotheses 的结果 (比如 p 值) 构造新的 global testing, 确保新的 testing 的 type-I error 为 α (仅适用于 global testing)

Logic

关于 Multiple hypothesis testing 的详细论述, 见:

- STA4100 Lecture 14, 内容包括:
  - Global testing problem
  - Bonferroni approach
  - Sidak approach
  - Fisher's combination test
  - Simes approach
  - Dunnet's approach
- STA4100 Lecture 15, 内容包括:
  - Bonferroni global test 的 optimality
  - Asymptotic power
  - "Needle in a haystack" problem
- STA4100 Lecture 16, 内容包括:
  - Optimality of detection threshold
  - χ^2 test
  - Bonferroni 和 χ^2 test 的比较
  - Bonferroni 和其他 global tests 的比较
  - Tukey's second-level significance testing
- STA4100 Lecture 17, 内容包括:
  - FWER 和 FDR 的定义和性质
  - Benjamini-Hochberg procedure
  - Hochberg procedure
  - BH procedure 的 empirical process viepoint
  - 对 BH procedure 的优化

- STA4100 Lecture 18, 内容包括:
  - BH procedure under dependency
  - The PRDS property

1 Multiple Hypothesis Testing

1.1 Multiple hypothesis testing 问题

Multiple hypothesis testing 问题的特点包括:

- 同时检验 $m > 1$ 个 hypotheses
- 希望不仅仅控制 the error rate for each test

Logic

第一种思路是控制 FWER, 即控制 "m 次检验中就犯哪怕 1 次 false rejection" 的概率

一种常见的方法是 Bonferroni method, 通过将每个 tests 的 level 调整为 α/m 来达到

这种方法非常 intuitive, 由于我们之前已经推出在 m 次检验中犯至少一次 false rejection 的概率为 $1-(1-\alpha)^m$, 我们直接控制它小于 α, 有:

$$
\begin{aligned}
& 1-(1-\alpha_0)^m < \alpha \\
\iff \quad & 1-\alpha_0 > (1-\alpha)^{\frac{1}{m}} \\
\iff \quad & \alpha_0 < 1-(1-\alpha)^{\frac{1}{m}} \overset{\text{Taylor expansion}}{\approx} 1-\left(1-\frac{\alpha}{m}\right) = \frac{\alpha}{m}
\end{aligned}
$$

1.2 Bonferroni method

令 significance level 为 α,
则 Bonferroni method reject the i^{th} null hypothesis, 若

$$p_i < \frac{\alpha}{m}$$

Remark (补充)

- Bonferroni method 对 $\{H_{0i}\}$ 没有任何要求, 永远可以使用它来控制 FWER
- Bonferroni method 非常严格, 很可能导致 power 较小
- 即使证明中的放缩非常宽松, 但对于 global test 且 H_{0i} independent 的情况, Bonferroni method 的 type-I error 会接近 α

Proof: Bonferroni method 的 validity

$$
\begin{aligned}
\mathbb{P}(\text{at least one Type I error}) &= \mathbb{P}\left(\bigcup_{i=1}^{m} \text{Type I error in the } i^{th} \text{ test}\right) \\
&\le \sum_{i=1}^{m} \mathbb{P}(\text{Type I error in the } i^{th} \text{ test}) \\
&= \sum_{i=1}^{m} \alpha/m \\
&= \alpha
\end{aligned}
$$

1.3 Definition: False discovery rate (FDR)

False discovery rate (FDR) 被定义为:

$$FDR = \mathbb{E}\left[\frac{\text{Number of false rejections}}{\text{Total number of rejections}}\right]$$

Logic

为了解决 Bonferroni method 过于 conservative 的问题, 第二种思路是控制 $FDR \leq \alpha$

一种常见的方法是 Benjamini-Hochberg procedure

1.4 Benjamini-Hochberg procedure

Benjamini-Hochberg procedure 的步骤如下:

1. 对于每个 test, 计算其 p-value P_i, 令 $P_{(1)} < \ldots, < P_{(m)}$ 表示 ordered p-values
2. 若 p-values 为 independent, 则选择 threshold 为 $R = \max\left(i : P_{(i)} < \frac{i\alpha}{m}\right)$
3. Reject 所有 $\text{p-value} \leq P_{(R)}$ 的 null hypothesis

由此能确保 $FDR \leq \alpha$