Перестановочные критерии для множественной проверки гипотез

Грабовой Андрей

Московский физико-технический институт

Москва,
2019г

- *P. H. Westfall, J. F. Troendle* Multiple Testing with Minimal Assumptions. // Biometrical Journal, 2008.
- *M. R. Peritz, K. R. Gabriel* On closed testing procedure with special reference to ordered analysis of variance. // Biometrika, 1976.
- *F. Bretz, T. Hothorn, P. Westfall* Multiple Comparisons Using R. — Taylor and Francis Group, 2011. 200 p.

Таблица: Обозначение

	True H_0	True H_1	All
Reject	V	S	R
Non-Reject	U	T	$m - R$
All	m_0	$m - m_0$	m

FWER:

$$FWER = P\left(V \geq 1\right)$$

Хотим контролировать FWER:

$$FWER \leq \alpha$$

The permutation multiple test algorithm

Задано множество гипотез:

$$\mathcal{H} = \{H_{\mathcal{I}} | \mathcal{I} \subset \mathcal{S}\}, \quad \mathcal{S} = \{1, \cdots, m\}, \quad H_{\mathcal{I}} = \cap_{i \in \mathcal{I}} H_i$$

Max-t test:

$$t_{max} = \max_{i \in \mathcal{I}} \{t_i\}, \quad H_{\mathcal{I}} \text{ is rejected } \Leftrightarrow P\left(t_{max} \geq c | H_{\mathcal{I}}\right) \leq \alpha$$

Условие на контроль FWER:

$$H \in \mathcal{H} \text{ is rejected } \Leftrightarrow H_* \text{ is rejected } \forall H_* \subset H$$

The permutation multiple test algorithm

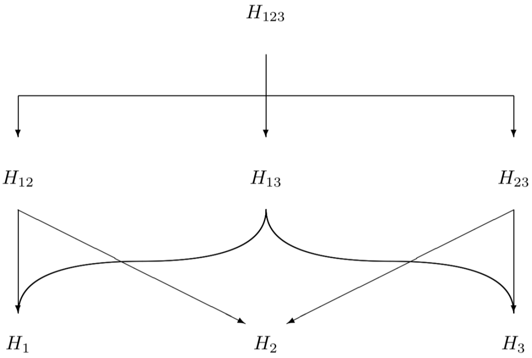

Рис.: Пример проверки множественной гипотезы из трех простых гипотез.

Получаем проверку $O(2^m)$ гипотез.

The permutation multiple test algorithm

Уменьшаем с $O(2^m)$ до $O(m)$.

Предположения:

- Используем t_{max} статистику
- следующие распределение равны

$$\max_{i\in\mathcal{I}}\{t_i\}|H_{\mathcal{I}}, \quad \max_{i\in\mathcal{I}}\{t_i\}|H_{\mathcal{S}}$$

Рассмотрим реализации $\{t_i\}_{i=1}^m$ для $\{H_i\}_{i=1}^m$, тогда:

$$p_{\mathcal{I}} = P\left(\max_{i\in\mathcal{I}} T_i \geq \max_{i\in\mathcal{I}} t_i|H_{\mathcal{I}}\right), \quad H_{\mathcal{I}} \text{ is rejected } \Leftrightarrow p_{\mathcal{I}} \leq \alpha$$

The permutation multiple test algorithm

1. By closure,

$$\text{reject } H_1 \text{ if } \max_{I:\, I \supseteq \{1\}} P\Big(\max_{i \in I} > T_i \geq \max_{i \in I} t_i \,|\, H_I\Big) \leq \alpha\,.$$

But if $I \supseteq \{1\}$, then $\max_{i \in I} t_i = t_1$, hence the rule is

$$\text{reject } H_1 \text{ if } \max_{I:\, I \supseteq \{1\}} P\Big(\max_{i \in I}\ T_i \geq t_1 \,|\, H_I\Big) \leq \alpha\,.$$

Using subset pivotality (B), the rule becomes

$$\text{reject } H_1 \text{ if } \max_{I:\, I \supseteq \{1\}} P\Big(\max_{i \in I}\ T_i \geq t_1 \,|\, H_{\{1,\dots,m\}}\Big) \leq \alpha\,.$$

Use of the “Max” statistic (A) implies

$$P\Big(\max_{i \in I} T_i \geq t_1 \,|\, H_{\{1,\dots,m\}}) \leq P\Big(\max_{i \in J}\ T_i \geq t_1 \,|\, H_{\{1,\dots,m\}}\Big) \quad \text{for} \quad I \subseteq J\,.$$

Hence, by subset pivotality and by use of the “Max” statistic, the rule by which we reject H_1 simplifies to this:

$$\text{reject } H_1 \text{ if } P\Big(\max_{i \in \{1,\dots,m\}}\ T_i \geq t_1 \,|\, H_{\{1,\dots,m\}}\Big) \leq \alpha\,.$$

The permutation multiple test algorithm

2. Again by closure and subset pivotality,

$$\text{reject } H_2 \text{ if } \max_{I:\, I \supseteq \{2\}} P\left(\max_{i \in I} T_i \geq \max_{i \in I} t_i \,|\, H_{\{1,\ldots,m\}}\right) \leq \alpha\,.$$

If $I \supseteq \{1\}$, then $\max_{i \in I} t_i = t_1$; else $\max_{i \in I} t_i = t_2$. Partitioning the set $\{I : I \supseteq \{2\}\}$ into two sets,

$$\boldsymbol{S}_1 = \{I : I \supseteq \{1,2\}\}\,, \qquad \boldsymbol{S}_2 = \{I : I \supseteq \{2\}, 1 \notin I\}\,,$$

we require

$$P\left(\max_{i \in I} T_i \geq t_1 \,|\, H_{\{1,\ldots,m\}}\right) \leq \alpha \quad \text{for all} \quad I \in \boldsymbol{S}_1$$

and

$$P\left(\max_{i \in I} T_i \geq t_2 \,|\, H_{\{1,\ldots,m\}}\right) \leq \alpha \quad \text{for all} \quad I \in \boldsymbol{S}_2\,.$$

Since we are using the "Max" statistic, these conditions are equivalent to the following rejection rule: reject H_2 if

$$P\left(\max_{i \in \{1,\ldots,m\}} T_i \geq t_1 \,|\, H_{\{1,\ldots,m\}}\right) \leq \alpha$$

and

$$P\left(\max_{i \in \{2,\ldots,m\}} T_i \geq t_2 \,|\, H_{\{1,\ldots,m\}}\right) \leq \alpha\,.$$

The permutation multiple test algorithm

$j.$: Continuing in this fashion, the rule is reject H_j if

$$P\left(\max_{i\in\{1,\dots,m\}} T_i \geq t_1 \mid H_{\{1,\dots,m\}}\right) \leq \alpha$$

and

$$P\left(\max_{i\in\{2,\dots,m\}} T_i \geq t_2 | H_{\{1,\dots,m\}}\right) \leq \alpha$$

and

$$\dots$$

and

$$P\left(\max_{i\in\{j,\dots,m\}} T_i \geq t_j | \, !H_{\{1,\dots,m\}}\right) \leq \alpha\,.$$

At step j the rule is equivalently stated in terms of p-values for the composite hypotheses as

$$\text{reject } H_j \text{ if } \max_{i\leq j} p_{\{i,\dots,m\}} \leq \alpha;$$

hence the rule reduces to

$$\text{reject } H_j \text{ if } \tilde{p}_j \leq \alpha,$$

where

$$\tilde{p}_j := \max_{i\leq j} p_{\{i,\dots,m\}}$$