

3.4 多重比较

3.4.1 例子

有4种新研制的镇痛药. 为检验新药相对于原有药物的镇痛效果, 选取情况相似的9个病人作试验. 对每个病人先后分别按随机排列的次序使用这4种新药和原有药物镇痛. 镇痛时间(min)见下表.

		病 人								
		1	2	3	4	5	6	7	8	9
原有药物		15.8	16.7	15.7	14.0	16.2	13.7	15.9	17.9	15.8
新 药	A	17.8	15.9	17.7	17.4	19.2	17.6	16.7	17.4	17.6
	B	19.1	20.0	18.0	19.3	20.0	19.1	19.0	20.4	19.4
	C	16.8	14.9	16.9	15.8	14.4	14.8	16.2	17.6	16.6
	D	21.4	20.4	20.1	21.3	19.4	20.2	21.1	21.2	20.3

问题：这4种新药的镇痛效果是否与原有药物等效？

统计问题的构建:

假设总体 $X = (X_1, \dots, X_5)' \stackrel{d}{\sim} N_5(\mu, \Sigma)$, 其中 X_1, \dots, X_5 分别表示原有药物以及新药A, B, C和D的镇痛时间. $\mu = (\mu_1, \dots, \mu_5)'$ 其中 μ_1, \dots, μ_5 分别表示原有药物以及新药A-D的平均镇痛时间.

检验问题:

$$\mathbf{H}_{i0} : \mu_{i+1} = \mu_1, \quad v.s. \quad \mathbf{H}_{i1} : \mu_{i+1} \neq \mu_1, \quad 1 \leq i \leq 4.$$

3.4.2 多重比较问题

总体 $X \stackrel{d}{\sim} N_p(\mu, \Sigma)$, $\mu_0 = (\mu_{10}, \dots, \mu_{p0})'$ 已知, 并假定 Σ 未知.

如果零假设“ $\mu = \mu_0$ ”被拒绝, 一个关心的问题是:

对哪些分量 i 有 $\mu_i \neq \mu_{i0}$, $1 \leq i \leq p$?

这意味着要进行下面 p 个检验问题: (多重比较问题)

$$\mathbf{H}_{i0} : \mu_i = \mu_{i0}, \quad v.s. \quad \mathbf{H}_{i1} : \mu_i \neq \mu_{i0}, \quad 1 \leq i \leq p. \quad (3.4.1)$$

3.4.3 错误率

记 $\Omega = \{1, 2, \dots, p\}$,

$$I = \{i : \mu_i = \mu_{i0}, 1 \leq i \leq p\},$$

I : 检验问题(3.4.1)中零假设为真的检验问题的集合.

犯错误: 对某个多重比较方法, 如果存在 $i \in I$, 但方法拒绝零假设 \mathbf{H}_{i0} , 则称该方法犯错误.

错误率: 对某个 I , 记

$$\Theta_I = \{\mu = (\mu_1, \dots, \mu_p)' : \mu_i = \mu_{i0}, i \in I; \mu_i \neq \mu_{i0}, i \notin I\}$$

为 I 所对应的参数空间.

那么对某个 I 而言, 多重比较方法犯错误的概率为

$$\sup_{\mu \in \Theta_I} [P \{ \bigcup_{i \in I} (\text{拒绝} \mathbf{H}_{i0}) \}],$$

因此, 对检验问题(3.4.1), 多重比较方法的错误率定义为

$$\sup_{I \subseteq \Omega} \left(\sup_{\mu \in \Theta_I} \left[P \left\{ \bigcup_{i \in I} (\text{拒绝} \mathbf{H}_{i0}) \right\} \right] \right).$$

例如:

$I = \{1\}$ 时, 犯错误的概率为在 $\mu_1 = \mu_{10}, \mu_2 \neq \mu_{20}, \dots, \mu_p \neq \mu_{p0}$ 下
概率 $P\{\text{拒绝}\mathbf{H}_{10}\}$ 的上确界.

$I = \{1, 2\}$ 时, 犯错误的概率为在 $\mu_1 = \mu_{10}, \mu_2 = \mu_{20}, \mu_3 \neq \mu_{30}, \dots$
 $\mu_p \neq \mu_{p0}$ 下概率

$P\{\text{拒绝}\mathbf{H}_{10}, \text{不拒绝}\mathbf{H}_{20}\} + P\{\text{拒绝}\mathbf{H}_{20}, \text{不拒绝}\mathbf{H}_{10}\} + P\{\text{拒绝}\mathbf{H}_{10} \text{ 与 } \mathbf{H}_{20}\}$
的上确界.

多重比较问题的主要目标是: 控制错误率, 构造错误率不大于显著性
水平 α 的检验方案.

3.4.4 联合置信区间法

对检验问题(3.4.1), 如果已构造出参数 μ_1, \dots, μ_p 的水平为 $1 - \alpha$ 的一个联合置信区间 $\{(\hat{\mu}_{iL}, \hat{\mu}_{iU}), 1 \leq i \leq p\}$, 满足

$$\begin{aligned} P(\hat{\mu}_{iL} < \mu_i < \hat{\mu}_{iU}, 1 \leq i \leq p) &= P\left(\bigcap_{i=1}^p \{\hat{\mu}_{iL} < \mu_i < \hat{\mu}_{iU}\}\right) \\ &\geq 1 - \alpha, \end{aligned} \quad (3.4.2)$$

对任意的 μ 成立.

则一个错误率不大于 α 的多重比较方案为:

对每个 $1 \leq i \leq p$,

若 $\mu_{i0} \leq \hat{\mu}_{iL}$ 或 $\mu_{i0} \geq \hat{\mu}_{iU}$, 则拒绝 H_{i0} , 认为 $\mu_i \neq \mu_{i0}$;

否则不拒绝 H_{i0} , 认为 $\mu_i = \mu_{i0}$.

证明：只要对任意 $I \subseteq \Omega = \{1, \dots, p\}$, 上面的检验方案犯错误的概率

$$\sup_{\mu \in \Theta_I} P\left\{ \bigcup_{i \in I} (\text{拒绝 } \mathbf{H}_{i0}) \right\} \leq \alpha, \quad (3.4.3)$$

则这个检验方案的错误率不大于 α . 事实上,

$$\begin{aligned} P\left(\bigcup_{i \in I} \{\text{拒绝 } \mathbf{H}_{i0}\}\right) &= P\left(\bigcup_{i \in I} [\{\mu_{i0} \leq \hat{\mu}_{iL}\} \cup \{\mu_{i0} \geq \hat{\mu}_{iU}\}]\right) \\ &= 1 - P\left(\bigcap_{i \in I} \{\hat{\mu}_{iL} < \mu_{i0} < \hat{\mu}_{iU}\}\right) \\ &\leq 1 - P\left(\left[\bigcap_{i \in I} \{\hat{\mu}_{iL} < \mu_{i0} < \hat{\mu}_{iU}\}\right] \cap \left[\bigcap_{i \notin I} \{\hat{\mu}_{iL} < \mu_i < \hat{\mu}_{iU}\}\right]\right). \end{aligned}$$

而由(3.4.2)知

$$P\left(\left[\bigcap_{i \in I} \{\hat{\mu}_{iL} < \mu_{i0} < \hat{\mu}_{iU}\}\right] \cap \left[\bigcap_{i \notin I} \{\hat{\mu}_{iL} < \mu_i < \hat{\mu}_{iU}\}\right]\right) \geq 1 - \alpha,$$

对任意的 μ 成立, 故(3.4.3)成立. #

3.4.5 Bonferroni不等式方法

设样本 $x_1, \dots, x_n \stackrel{i.i.d.}{\sim} N_p(\mu, \Sigma)$,

记样本均值 $\bar{x} = (\bar{x}_1, \dots, \bar{x}_p)'$, 样本协方差阵 $\mathbf{S} = (s_{ij})_{p \times p}$.

有

$$\frac{\sqrt{n}(\bar{x}_i - \mu_i)}{\sqrt{s_{ii}}} \stackrel{d}{\sim} t(n-1), \quad 1 \leq i \leq p.$$

因此, 可以取如下形式的 μ 的一个联合置信区间:

$$\bar{x}_i - c \frac{\sqrt{s_{ii}}}{\sqrt{n}} < \mu_i < \bar{x}_i + c \frac{\sqrt{s_{ii}}}{\sqrt{n}}, \quad 1 \leq i \leq p,$$

其中, 常数 c 待确定, 以满足

$$P \left(\bigcap_{i=1}^p \left\{ \bar{x}_i - c \frac{\sqrt{s_{ii}}}{\sqrt{n}} < \mu_i < \bar{x}_i + c \frac{\sqrt{s_{ii}}}{\sqrt{n}} \right\} \right) \geq 1 - \alpha. \quad (3.4.4)$$

Bonferroni不等式: 对事件 A_1, \dots, A_p , 有

$$P\left(\bigcup_{i=1}^p A_i\right) \leq \sum_{i=1}^p P(A_i).$$

那么

$$P\left(\bigcap_{i=1}^p A_i\right) = 1 - P\left(\bigcup_{i=1}^p A_i^c\right) \geq 1 - \sum_{i=1}^p P(A_i^c), \quad (3.4.5)$$

其中, A_i^c 表示事件 A_i 的补(逆)事件.

因此, 由(3.4.5)知

$$\begin{aligned} & P\left(\bigcap_{i=1}^p \left\{\bar{x}_i - c\frac{\sqrt{s_{ii}}}{\sqrt{n}} < \mu_i < \bar{x}_i + c\frac{\sqrt{s_{ii}}}{\sqrt{n}}\right\}\right) \\ & \geq 1 - \sum_{i=1}^p P\left(\left\{\mu_i \leq \bar{x}_i - c\frac{\sqrt{s_{ii}}}{\sqrt{n}}\right\} \cup \left\{\mu_i \geq \bar{x}_i + c\frac{\sqrt{s_{ii}}}{\sqrt{n}}\right\}\right) \quad (3.4.6) \\ & = 1 - \sum_{i=1}^p P\left(\frac{\sqrt{n}|\bar{x}_i - \mu_i|}{\sqrt{s_{ii}}} \geq c\right). \end{aligned}$$

因此, 为使(3.4.4)成立, 只需

$$P\left(\frac{\sqrt{n}|\bar{x}_i - \mu_i|}{\sqrt{s_{ii}}} \geq c\right) \leq \frac{\alpha}{p},$$

对所有 $1 \leq i \leq p$ 成立.

不难知道, 取 $c^* = t_{1-\alpha/(2p)}(n-1)$ 即可, 其中, $t_{1-\alpha/(2p)}(n-1)$ 是自由度为 $(n-1)$ 的 t 分布的 $1 - \alpha/(2p)$ 分位点.

则 μ 的一个水平为 $1 - \alpha$ 的联合置信区间为

$$\left(\bar{x}_i - c^* \sqrt{\frac{s_{ii}}{n}}, \bar{x}_i + c^* \sqrt{\frac{s_{ii}}{n}}\right), \quad 1 \leq i \leq p.$$

因此, 基于**Bonferroni**方法的多重比较的检验方案为:

对每个 $1 \leq i \leq p$,

若 $|t_i| \geq c^* = t_{1-\alpha/(2p)}(n-1)$, 则拒绝 \mathbf{H}_{i0} ;

否则, 不拒绝 \mathbf{H}_{i0} ,

其中,

$$t_i = \frac{\sqrt{n}(\bar{x}_i - \mu_{i0})}{\sqrt{s_{ii}}}, \quad 1 \leq i \leq p.$$

Bonferroni不等式方法偏“保守”.

3.4.6 Scheffe方法

考虑如下的多重比较问题:

$$\mathbf{H}_{a0} : a'\mu = a'\mu_0, \quad v.s. \quad \mathbf{H}_{a1} : a'\mu \neq a'\mu_0, \quad a \in R^p, a \neq 0. \quad (3.4.7)$$

对任意的 $a \in R^p, a \neq 0$, 有

$$\frac{\sqrt{n}(a'\bar{x} - a'\mu)}{\sqrt{a'\mathbf{S}a}} \stackrel{d}{\sim} t(n-1),$$

那么 $a'\mu$ 的一个置信区间可以取为

$$\frac{\sqrt{n}|a'\bar{x} - a'\mu|}{\sqrt{a'\mathbf{S}a}} < c,$$

其中, c 是参数待确定. 为使上面是一个联合置信区间, 要求满足

$$P \left(\bigcap_{\substack{a \in R^p \\ a \neq 0}} \left\{ \frac{\sqrt{n}|a'\bar{x} - a'\mu|}{\sqrt{a'\mathbf{S}a}} < c \right\} \right) \geq 1 - \alpha.$$

不难知道

$$P\left(\bigcap_{\substack{a \in R^p \\ a \neq 0}} \left\{ \frac{\sqrt{n}|a'\bar{x} - a'\mu|}{\sqrt{a'\mathbf{S}a}} < c \right\}\right) = P\left(\sup_{a \in R^p} \frac{\sqrt{n}|a'\bar{x} - a'\mu|}{\sqrt{a'\mathbf{S}a}} < c\right).$$

由二次型极值性质知

$$\sup_{a \in R^p} \frac{n(a'\bar{x} - a'\mu)^2}{a'\mathbf{S}a} = n(\bar{x} - \mu)'\mathbf{S}^{-1}(\bar{x} - \mu),$$

进而知

$$\begin{aligned} n(\bar{x} - \mu)'\mathbf{S}^{-1}(\bar{x} - \mu) &\stackrel{d}{\sim} T_p^2(n-1); \\ \frac{n(n-p)}{p(n-1)}(\bar{x} - \mu)'\mathbf{S}^{-1}(\bar{x} - \mu) &\stackrel{d}{\sim} F(p, n-p). \end{aligned}$$

因此, c 可以取为

$$c^{**} = \sqrt{\frac{p(n-1)}{n-p} F_{1-\alpha}(p, n-p)},$$

它满足

$$P \left(\sup_{a \in R^p} \frac{\sqrt{n}|a'\bar{x} - a'\mu|}{\sqrt{a'\mathbf{S}a}} < c^{**} \right) \geq 1 - \alpha.$$

可得 $\{a'\mu, a \in R^p\}$ 的一个联合置信区间为

$$a'\bar{x} - c^{**}n^{-1/2}\sqrt{a'\mathbf{S}a} < a'\mu < a'\bar{x} + c^{**}n^{-1/2}\sqrt{a'\mathbf{S}a}.$$

这时, 若取 $a = e_i$, e_i 为第 i 个元素是1, 其它元素均是0的向量, $1 \leq i \leq p$, 则可得 $\{\mu_1, \dots, \mu_p\}$ 的一个联合置信区间为

$$\left(\bar{x}_i - c^{**}\sqrt{\frac{s_{ii}}{n}}, \quad \bar{x}_i + c^{**}\sqrt{\frac{s_{ii}}{n}} \right), \quad 1 \leq i \leq p.$$

因此, 基于Scheffe方法的多重比较的检验方案为:

对每个 $1 \leq i \leq p$,

若 $|t_i| \geq c^{**} = \sqrt{\frac{p(n-1)}{n-p} F_{1-\alpha}(p, n-p)}$, 则拒绝 \mathbf{H}_{i0} ;

否则, 不拒绝 \mathbf{H}_{i0} ,

其中,

$$t_i = \frac{\sqrt{n}(\bar{x}_i - \mu_{i0})}{\sqrt{s_{ii}}}, \quad 1 \leq i \leq p.$$

Scheffe方法也偏“保守”.

3.4.7 Bonferroni方法与Scheffe方法的比较

注意到: Bonferroni方法和Scheffe方法的多重比较问题的检验方案不同之处仅在于它们的临界值不同.

Bonferroni的临界值为 $c^* = t_{1-\alpha/(2p)}(n-1)$,

Scheffe的临界值为 $c^{**} = \sqrt{\frac{p(n-1)}{n-p} F_{1-\alpha}(p, n-p)}$.

可以证明, 对常用的 p 和 α , 即当 $\alpha \leq 0.10$, $2 \leq p \leq 12$ 时, 对任意的 $n \geq (p+1)$, 都有 $c^* < c^{**}$.

因此, Bonferroni的联合置信区间长度比Scheffe的短.

Scheffe方法的多重比较问题的检验方案比Bonferroni的更保守.

对多重比较问题(3.4.1), 倾向于使用由Bonferroni方法给出的检验方案.

例子(3.4.1例续):

考虑原始问题的一个线性变换

$$\mathbf{Y} = \mathbf{C}\mathbf{X} = (Y_1, \dots, Y_4)' \stackrel{d}{\sim} N_4(\boldsymbol{\theta}, \mathbf{T}), \quad (3.4.8)$$

其中,

$$\mathbf{C} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\boldsymbol{\theta} = \mathbf{C}\boldsymbol{\mu} = (\theta_1, \dots, \theta_4)',$$

$$\theta_i = \mu_{i+1} - \mu_1, \quad 1 \leq i \leq 4,$$

$$\mathbf{T} = \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}'.$$

则检验问题转换为：

$$\mathbf{H}_{i0} : \theta_i = 0, \quad v.s. \quad \mathbf{H}_{i1} : \mu_i \neq 0, \quad 1 \leq i \leq 4.$$

变换后的数据表为：

	病 人								
	1	2	3	4	5	6	7	8	9
新药A-原有药物	2.0	-0.8	2.0	3.4	3.0	3.9	0.8	-0.5	1.8
新药B-原有药物	3.3	3.3	2.3	5.3	3.8	5.4	3.1	2.5	3.6
新药C-原有药物	1.0	-1.8	1.2	1.8	-1.8	1.1	0.3	-0.3	0.8
新药D-原有药物	5.6	3.7	4.4	7.3	3.2	6.5	5.2	3.3	4.5

计算出

$$\text{样本均值 } \bar{y} = \begin{pmatrix} 1.7333 \\ 3.6222 \\ 0.2556 \\ 4.8550 \end{pmatrix},$$

$$\text{样本协差阵 } \mathbf{S} = \begin{pmatrix} 2.6875 & 1.2804 & 1.0867 & 1.4592 \\ 1.2804 & 1.1869 & 0.4361 & 1.1299 \\ 1.0867 & 0.4361 & 1.7003 & 1.4690 \\ 1.4592 & 1.1299 & 1.4690 & 2.0228 \end{pmatrix},$$

检验统计量 $t_1 = 3.1718, t_2 = 9.9744, t_3 = 0.5881, t_4 = 10.2401$.

由于 $p = 4$, $n = 9$, 取显著性水平 $\alpha = 0.05$,
则采用Bonferroni不等式法的检验方案,
相应的临界值 $c^* = t_{1-\alpha/(2p)}(n-1) = t_{0.99375}(8) = 3.2060$.
有 $t_1 < c^*$, $t_2 > c^*$, $t_3 < c^*$, $t_4 > c^*$.

结论:

新药**A**和**C**的镇痛效果与原有药物都**没有显著差异**,
新药**B**和**D**的镇痛效果与原有药物都**有显著差异**.

如果对每种新药单独做比较的话, 即采用t-检验,
临界值为 $c^{***} = t_{1-\alpha/2}(n-1) = t_{0.975}(8) = 2.3060$,
有 $t_1 > c^{***}$, $t_2 > c^{***}$, $t_3 < c^{***}$, $t_4 > c^{***}$.

结论:

新药C的镇痛效果与原有药物没有显著差异,
新药A、B和D的镇痛效果与原有药物都有显著差异.

3.4.8 多元方差分析中的多重比较

假设有相互独立的 k 个总体 $\mathbf{X}_j \stackrel{d}{\sim} N_p(\mu_j, \Sigma)$, $1 \leq j \leq k$, $\Sigma > 0$.

如果零假设“ $\mu_1 = \cdots = \mu_k$ ”被拒绝, 一个关心的问题是:

对哪些总体对 (s, t) , $1 \leq s < t \leq k$, 使得 $\mu_s \neq \mu_t$?

记 $\mu_j = (\mu_{1j}, \cdots, \mu_{pj})'$, $1 \leq j \leq k$.

若“ $\mu_s = \mu_t$ ”被拒绝, 是否存在 $1 \leq i \leq p$, 使得 $\mu_{is} \neq \mu_{it}$?

这就是要进行下面的多重比较问题:

$$\mathbf{H}_{sti0} : \mu_{is} = \mu_{it}, \quad v.s. \quad \mathbf{H}_{sti1} : \mu_{is} \neq \mu_{it}, \quad 1 \leq s < t \leq k, \quad 1 \leq i \leq p. \quad (3.4.9)$$

Bonferroni不等式方法

设 x_{j1}, \dots, x_{jn_j} 是来自总体 \mathbf{X}_j 的样本, $1 \leq j \leq k, 1 \leq i \leq p$.

记 $n = \sum_{j=1}^k n_j, n \geq p + k$.

计算如下统计量:

第 j 个总体的样本均值: $\bar{x}_j = n_j^{-1} \sum_{i=1}^{n_j} x_{ji} = (\bar{x}_{1j}, \dots, \bar{x}_{pj})', 1 \leq j \leq k$;

第 j 个总体的样本离差阵: $\mathbf{V}_j = \sum_{i=1}^{n_j} (x_{ji} - \bar{x}_j)(x_{ji} - \bar{x}_j)', 1 \leq j \leq k$;

组内离差阵: $\mathbf{SSW} = \sum_{j=1}^k \mathbf{V}_j = (w_{ij})_{p \times p}$;

样本协差阵: $\mathbf{S} = \mathbf{SSW} / (n - k) = (s_{ij})_{p \times p}$.

利用Bonferroni不等式方法得到一个置信水平为 $1 - \alpha$ 的联合置信区间为

$$\mu_{is} - \mu_{it} : (\bar{x}_{is} - \bar{x}_{it}) \pm \sqrt{\frac{n_s + n_t}{n_s n_t}} \sqrt{s_{ii}} t_{1-\alpha/(k(k-1)p)}(n-k),$$

其中, $1 \leq s < t \leq k, 1 \leq i \leq p$.

因此, 多重比较问题(3.4.9)的错误率不大于 α 的一个检验方案为:

若 $|t_{sti}| \geq t_{1-\alpha/(k(k-1)p)}(n-k)$, 则拒绝 \mathbf{H}_{sti0} ;

否则, 不拒绝 \mathbf{H}_{sti0} ,

其中,

$$t_{sti} = \sqrt{\frac{n_s + n_t}{n_s n_t}} \cdot \frac{\bar{x}_{is} - \bar{x}_{it}}{\sqrt{s_{ii}}}, \quad 1 \leq s < t \leq k, 1 \leq i \leq p.$$

例子(3.3.4的例3续):

均值检验的结果是: 3个总体均值不完全相同.

多重比较检验: 3种生产方法在完成4种不同任务时的具体差异.

本例中, $p = 4$, $k = 3$, $n_1 = n_2 = n_3 = 10$, 总样本量 $n = 30$.

取显著性水平 $\alpha = 0.05$, 则本例对应的Bonferroni不等式方法的临界值为

$$C_r = t_{1-\alpha/(k(k-1)p)}(n-k) = t_{1-0.05/24}(27) = \mathbf{3.1301}.$$

经计算, 3个总体的样本均值分别为

$$\bar{x}_1 = \begin{pmatrix} 10.70 \\ 12.81 \\ 13.49 \\ 10.31 \end{pmatrix}, \quad \bar{x}_2 = \begin{pmatrix} 10.90 \\ 12.06 \\ 17.67 \\ 15.64 \end{pmatrix}, \quad \bar{x}_3 = \begin{pmatrix} 14.77 \\ 18.66 \\ 17.95 \\ 18.71 \end{pmatrix},$$

总的样本协方差阵为

$$S = \frac{\mathbf{SSW}}{n - k} = \begin{pmatrix} 23.08 & 20.70 & 24.17 & 22.12 \\ 20.70 & 22.06 & 24.94 & 21.80 \\ 24.17 & 24.94 & 31.85 & 26.22 \\ 22.12 & 21.80 & 26.22 & 26.10 \end{pmatrix}.$$

再计算第 s 个总体与第 t 个总体的第 i 个指标是否有差异的统计量 t_{sti} ($1 \leq s < t \leq 3$, $1 \leq i \leq 4$) 的值见下表

		$s = 1, t = 2$	$s = 1, t = 3$	$s = 2, t = 3$
i	1	-0.0931	-1.8945	-1.8014
	2	0.3571	-2.7856	-3.1427
	3	-1.6562	-1.7672	-0.1109
	4	-2.3330	-3.6768	-1.3438

结论:

- 1) 方法1和方法3在完成第4个任务时所花时间在显著差异;
- 2) 方法2和方法3在完成第2个任务时所花时间在显著差异.

3.4.9 其它多重比较方法

i) 基于显著性水平的多重比较方法)

假设有 m 个零假设 $\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_m$,

相应检验的 p 值为 P_1, P_2, \dots, P_m .

对这 m 个 p 值按由小到大排序, 得 $P_{(1)} \leq P_{(2)} \leq \dots \leq P_{(m)}$.

Sidak-Holm下降法

- 1) 若 $P_{(1)} > 1 - (1 - \alpha)^{1/m}$, 则接受所有零假设;
- 2) 若 $P_{(1)} \leq 1 - (1 - \alpha)^{1/m}$, 则拒绝 $\mathbf{H}_{(1)}$, 接下考虑 $\mathbf{H}_{(2)}$;
- 3) $\dots\dots$;
- 4) 直到找到第一个 j 满足: $P_{(j)} > 1 - (1 - \alpha)^{1/(m-j+1)}$,
则接受 $\mathbf{H}_{(j)}, \mathbf{H}_{(j+1)}, \dots, \mathbf{H}_{(m)}$, 拒绝 $\mathbf{H}_{(1)}, \dots, \mathbf{H}_{(j-1)}$.

ii) 基于错误发现率的多重比较方法

错误发现率-**FDR** (False Discovery Rate)

假设有 m 个零假设 $\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_m$, 其中有 m_0 个数正确的, R 表示被拒绝的零假设的个数.

零假设	接收	拒绝	合计
真	U	V	m_0
假	T	S	$m - m_0$
合计	W	R	m

则

$$\text{FDR} = \frac{V}{R}.$$

假设有 m 个零假设 $\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_m$,

相应检验的 p 值为 P_1, P_2, \dots, P_m .

对这 m 个 p 值按由小到大排序, 得 $P_{(1)} \leq P_{(2)} \leq \dots \leq P_{(m)}$.

(1) Benjamini-Hochberg (1995) 多重比较方法:

1) 对给定的FDR α , 找最大的 k , 使得 $P_{(k)} \leq \frac{k}{m}\alpha$, 即

$$k = \arg \max_{1 \leq i \leq m} \left\{ i : P_{(i)} \leq \frac{i}{m}\alpha \right\};$$

2) 拒绝零假设 $\mathbf{H}_{(1)}, \mathbf{H}_{(2)}, \dots, \mathbf{H}_{(k)}$.

则总检验方法的 $\text{FDR} \leq \alpha$.

可得修正的 p 值为

$$\tilde{P}_{(i)} = \min_{i \leq k \leq m} \left\{ \min \left(\frac{m}{k} P_{(k)}, 1 \right) \right\}, \quad 1 \leq i \leq m.$$

(2) Benjamini-Liu (1999) 多重比较方法:

1) 对给定的FDR α , 找最小的 k , 使得

$$k = \arg \min_{1 \leq i \leq m} \left\{ i : P_{(i)} \leq \min \left(\alpha, \frac{m}{(m+1-k)^2} \alpha \right) \right\};$$

2) 拒绝零假设 $\mathbf{H}_{(1)}, \mathbf{H}_{(2)}, \dots, \mathbf{H}_{(k)}$.

可得修正的 p 值为

$$\tilde{P}_{(i)} = \min_{i \leq k \leq m} \left\{ \min \left(\frac{(m+1-k)^2}{m} P_{(k)}, 1 \right) \right\}, \quad 1 \leq i \leq m.$$

(3) Benjamin-Yekutieli (2001) 多重比较方法:

1) 对给定的FDR α , 找最大的 k , 使得

$$k = \arg \max_{1 \leq i \leq m} \left\{ i : P_{(i)} \leq \frac{i}{m \sum_{j=1}^m \frac{1}{j}} \alpha \right\};$$

2) 拒绝零假设 $\mathbf{H}_{(1)}, \mathbf{H}_{(2)}, \dots, \mathbf{H}_{(k)}$.

可得修正的 p 值为

$$\tilde{P}_{(i)} = \max_{1 \leq k \leq i} \left\{ \min \left(\frac{m \sum_{j=1}^m \frac{1}{j}}{k} P_{(k)}, 1 \right) \right\}, \quad 1 \leq i \leq m.$$

例子:

假设有10个零假设. 在零假设下, 相应10个检验的 p 值排序为

k	1	2	3	4	5	6	7	8	9	10
$P_{(k)}$	0.0020	0.0045	0.0060	0.0080	0.0085	0.0090	0.0175	0.0250	0.1055	0.5350
$10P_{(k)}/k$	0.0200	0.0225	0.0200	0.0200	0.0170	0.0150	0.0250	0.0313	0.1172	0.5350

应用**Benjamini-Hochberg**方法, 若要总的FDR不超过 0.05,
则拒绝零假设 $H_{(1)}, H_{(2)}, \dots, H_{(8)}$.

例子(3.3.4的例3续):

共有12个零假设. 在零假设下, 相应的检验统计量均服从 $T(27)$, 即自由度为 27 的 T 分布.

检验统计量和相应的 p 值见下表:

		$s = 1, t = 2$	$s = 1, t = 3$	$s = 2, t = 3$
i	1	-0.0931(0.9265)	-1.8945(0.0689)	-1.8014(0.0680)
	2	0.3571(0.7238)	-2.7856(0.0097)	-3.1427(0.0040)
	3	-1.6562(0.1093)	-1.7672(0.0885)	-0.1109(0.9125)
	4	-2.3330(0.0273)	-3.6768(0.0010)	-1.3438(0.1902)

p 值由小到大排序后见下表:

k	1 H_{413}	2 H_{223}	3 H_{213}	4 H_{412}	5 H_{123}	6 H_{113}	7 H_{313}	8 H_{312}	9 H_{423}	10 H_{212}	11 H_{323}	12 H_{112}
$P_{(k)}$	0.0010	0.0040	0.0097	0.0273	0.0680	0.0689	0.0885	0.1093	0.1902	0.7238	0.9125	0.9265

基于Sidak-Holm方法的检验结果:

修正的 p 值见下表:

k	1 H_{413}	2 H_{223}	3 H_{213}	4 H_{412}	5 H_{123}	6 H_{113}	7 H_{313}	8 H_{312}	9 H_{423}	10 H_{212}	11 H_{323}	12 H_{112}
$P_{(k)}$	0.0010	0.0040	0.0097	0.0273	0.0680	0.0689	0.0885	0.1093	0.1902	0.7238	0.9125	0.9265
$\tilde{P}_{(k)}$	0.0119	0.0431	0.0929	0.2205	0.4307	0.3933	0.4265	0.4394	0.5700	0.9789	0.9923	0.9265

其中 $\tilde{P}_{(k)} = 1 - (1 - P_{(k)})^{m-k+1}$, $1 \leq k \leq m$.

若要多重检验的显著性水平不超过 0.05, 则拒绝零假设 H_{413} 和 H_{223} .

与Bonferroni检验结果一致.

基于Benjamini-Hochberg方法的检验结果:

修正的 p 值见下表:

k	1 H_{413}	2 H_{223}	3 H_{213}	4 H_{412}	5 H_{123}	6 H_{113}	7 H_{313}	8 H_{312}	9 H_{423}	10 H_{212}	11 H_{323}	12 H_{112}
$P_{(k)}$	0.0010	0.0040	0.0097	0.0273	0.0680	0.0689	0.0885	0.1093	0.1902	0.7238	0.9125	0.9265
$12P_{(k)}/k$	0.0120	0.0240	0.0388	0.0819	0.1632	0.1378	0.1517	0.1640	0.2536	0.8686	0.9955	0.9265

若要总的FDR不超过 0.05, 则拒绝零假设 $H_{413}, H_{223}, H_{213}$.

总结：对多元方差分析问题

- 1) 先检验各总体的均值是否有差异;
- 2) 总体均值有差异时, 再进行多重比较;
- 3) 多重比较不能代替均值的检验.