2.1.4 多重比较与对照

王正明　　易泰河

系统工程学院　　军事建模与仿真系

2019 年 11 月 22 日

2.1.4 多重比较与对照

(1) 固定效应模型

(2) 多重比较

(3) 对照

- 单因子固定效应模型

$$\begin{cases} y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad \varepsilon_{ij} \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2), \\ i = 1, \cdots, a, \quad j = 1, \cdots, n_i, \\ \sum_{i=1}^{a} n_i \tau_i = 0. \end{cases}$$

- A_i 是原因 (cause), τ_i 是结果 (effect).

- 双因子固定效应模型

$$
\begin{cases}
y_{ijl} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijl}, \quad \varepsilon_{ijl} \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2), \\
i = 1, \cdots, a, \quad j = 1, \cdots, b, \quad l = 1, \cdots, m, \\
\sum_{i=1}^{a} \tau_i = 0, \quad \sum_{j=1}^{b} \beta_j = 0, \\
\sum_{i=1}^{a} (\tau\beta)_{ij} = 0, \quad j = 1, \cdots, b, \\
\sum_{j=1}^{b} (\tau\beta)_{ij} = 0, \quad i = 1, \cdots, a.
\end{cases}
$$

- 三因子固定效应模型?

- 将多因子试验看作单因子试验，一共 N 个处理，

处理	观察值			
$\boldsymbol{x}_1$	y_{11}	y_{12}	$\cdots$	y_{1m_1}
$\boldsymbol{x}_2$	y_{21}	y_{22}	$\cdots$	y_{2m_2}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\boldsymbol{x}_N$	y_{N1}	y_{N2}	$\cdots$	y_{Nm_N}

- 固定效应模型:

$$y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad j = 1, 2, \cdots, m_i.$$

不改变自由度，不改变独立参数的个数!

- 参数化: 把测不同处理处的响应值转化为测诸效应参数.

2.1.4 多重比较与对照

(1) 固定效应模型

(2) 多重比较

(3) 对照

Example

设制造某新型手枪共有 A, B, C, D 四种不同工艺. 为研究四种工艺之间的差异, 命 a, b, c, d, e 五个战士打靶, 每人提供 400 发子弹. 命中频率数据如下, 四种工艺是否有差异?

	A	B	C	D
a	0.60	0.59	0.71	0.72
b	0.80	0.81	0.88	0.86
c	0.68	0.64	0.80	0.79
d	0.68	0.70	0.81	0.82
e	0.59	0.60	0.73	0.72
和	3.36	3.34	3.93	3.91
平均	0.672	0.668	0.786	0.782

- 多重比较: $H_0: \mu_i = \mu_j, \quad 1 \le i < j \le N$.
- C_N^2 个两样本比较: $H_0^{ij}: \mu_i = \mu_j, \quad H_1^{ij}: \mu_i \neq \mu_j$.
- 注意到

$$\bar{y}_{i\cdot} - \bar{y}_{j\cdot} \sim N\left(\mu_i - \mu_j, \frac{\sigma^2}{m_i} + \frac{\sigma^2}{m_j}\right), \quad \frac{MS_E}{n-N} \sim \chi^2(n-N),$$

当原假设 $H_0^{ij}: \mu_i = \mu_j$ 成立时

$$t_{ij} = \frac{\bar{y}_{j\cdot} - \bar{y}_{i\cdot}}{\sqrt{MS_E(1/m_i + 1/m_j)}} \sim t(n-N).$$

此即两样本 t 检验.

- 多重比较: $H_0: \mu_i = \mu_j, \quad 1 \le i < j \le N$.
- C_N^2 个两样本比较: $H_0^{ij}: \mu_i = \mu_j, \quad H_1^{ij}: \mu_i \neq \mu_j$.
- 注意到

$$\bar{y}_{i\cdot} - \bar{y}_{j\cdot} \sim N\left(\mu_i - \mu_j, \frac{\sigma^2}{m_i} + \frac{\sigma^2}{m_j}\right), \quad \frac{MS_E}{n-N} \sim \chi^2(n-N),$$

当原假设 $H_0^{ij}: \mu_i = \mu_j$ 成立时

$$t_{ij} = \frac{\bar{y}_{j\cdot} - \bar{y}_{i\cdot}}{\sqrt{MS_E\,(1/m_i + 1/m_j)}} \sim t(n-N).$$

此即两样本 t 检验.

多重比较犯错概率积累:

- 若单个检验犯第 I 类错误的概率为 α, 则 n 个检验犯第 I 类错误的概率为 $1-(1-\alpha)^n$;
- 当 $\alpha=0.05$、$N=5$ 时, $C_N^2=10$,
$$1-(1-0.05)^{10}\approx 0.40;$$
- 当 $\alpha=0.05$、$N=10$ 时, $C_N^2=45$,
$$1-(1-0.05)^{45}\approx 0.90.$$

多重比较犯错概率积累:

- 若单个检验犯第 I 类错误的概率为 α, 则 n 个检验犯第 I 类错误的概率为 $1-(1-\alpha)^n$;
- 当 $\alpha=0.05$、$N=5$ 时, $C_N^2=10$,

$$1-(1-0.05)^{10}\approx 0.40;$$

- 当 $\alpha=0.05$、$N=10$ 时, $C_N^2=45$,

$$1-(1-0.05)^{45}\approx 0.90.$$

- 多重检验的Bonferroni 法: 降低单个检验的水平使整体达到 α 的检验水平,

$$|t_{ij}| > t_{1-\alpha/(2C_N^2)}(n-N)$$

记由上式确定的拒绝域为 A_{ij}, 则 $P(A_{ij}|H_0^{ij}) = \alpha/C_N^2$,

$$P\left(\bigcup_{i<j} A_{ij}|H_0\right) < \sum_{i<j} P(A_{ij}|H_0^{ij}) = \sum_{i<j} \frac{\alpha}{C_N^2} = \alpha.$$

- 当 N 较大时, 多重检验属于大规模统计推断 (large-scale inference) 问题, 是近年来统计学领域研究的热点问题之一 (T. Tony Cai & Wenguang Sun, 2019).

- 多重检验的Bonferroni 法: 降低单个检验的水平使整体达到 α 的检验水平,

$$|t_{ij}| > t_{1-\alpha/(2C_N^2)}(n-N)$$

记由上式确定的拒绝域为 A_{ij}, 则 $P(A_{ij}|H_0^{ij}) = \alpha/C_N^2$,

$$P\left(\bigcup_{i<j} A_{ij}|H_0\right) < \sum_{i<j} P(A_{ij}|H_0^{ij}) = \sum_{i<j} \frac{\alpha}{C_N^2} = \alpha.$$

- 当 N 较大时, 多重检验属于大规模统计推断 (large-scale inference) 问题, 是近年来统计学领域研究的热点问题之一 (T. Tony Cai & Wenguang Sun, 2019).

小测试

- 两样本 t 检验?

$$t_{ij} = \frac{\bar{y}_{j\cdot} - \bar{y}_{i\cdot}}{\sqrt{MS_E(1/m_i + 1/m_j)}} \sim t(n-N).$$

- $|t_{ij}| > t_{1-\alpha/2}(n-N)$ 时拒绝 H_0^{ij}.

小测试

- 两样本 t 检验?

$$t_{ij} = \frac{\bar{y}_{j\cdot} - \bar{y}_{i\cdot}}{\sqrt{MS_E\left(1/m_i + 1/m_j\right)}} \sim t(n-N).$$

- $|t_{ij}| > t_{1-\alpha/2}(n-N)$ 时拒绝 H_0^{ij}.

2.1.4 多重比较与对照

(1) 固定效应模型

(2) 多重比较

(3) 对照

- $H_0^{ij}: \mu_i = \mu_j \quad \Longleftrightarrow \quad H_0^{ij}: \tau_i - \tau_j = 0.$
- 考虑线性假设

$$H_0^c: \sum_{i=1}^{N} c_i \mu_i = 0, \quad \text{or} \quad H_0^c: \sum_{i=1}^{N} c_i \tau_i = 0.$$

Definition

设 $c_1, c_2, \cdots, c_N$ 为满足 $c_1 + \cdots + c_N = 0$ 的 N 个不全为零的常数，称线性组合 $c_1\mu_1 + c_2\mu_2 + \cdots + c_N\mu_N$ 为一个对比或对照(contrast).

- $H_0^{ij}: \mu_i = \mu_j \quad \Longleftrightarrow \quad H_0^{ij}: \tau_i - \tau_j = 0.$
- 考虑线性假设

$$H_0^c: \sum_{i=1}^{N} c_i\mu_i = 0, \quad \text{or} \quad H_0^c: \sum_{i=1}^{N} c_i\tau_i = 0.$$

Definition

设 $c_1, c_2, \cdots, c_N$ 为满足 $c_1 + \cdots + c_N = 0$ 的 N 个不全为零的常数，称线性组合 $c_1\mu_1 + c_2\mu_2 + \cdots + c_N\mu_N$ 为一个对比或对照(contrast).

- $c = \mu_i - \mu_j$, $c = 2\mu_i - \mu_j - \mu_k$ 都是对照;
- 由于 $\bar{y}_{i\cdot} \sim N(\mu_i, \sigma^2/m_i)$, 故

$$\sum_{i=1}^{N} c_i \bar{y}_{i\cdot} \sim N\left(\sum_{i=1}^{N} c_i\mu_i, \sum_{i=1}^{N} \frac{c_i^2}{m_i}\sigma^2\right),$$

也称 $\sum\limits_{i=1}^{N} c_i\mu_i$ 的无偏估计 $\sum\limits_{i=1}^{N} c_i\bar{y}_{i\cdot}$ 为对照;

- 定义对照平方和为

$$SS_c = \frac{\left(\sum_{i=1}^{N} c_i\bar{y}_{i\cdot}\right)^2}{\sum_{i=1}^{N} \frac{c_i^2}{m_i}}.$$

- $c = \mu_i - \mu_j$, $c = 2\mu_i - \mu_j - \mu_k$ 都是对照;
- 由于 $\bar{y}_{i\cdot} \sim N(\mu_i, \sigma^2/m_i)$, 故

$$\sum_{i=1}^{N} c_i \bar{y}_{i\cdot} \sim N\left(\sum_{i=1}^{N} c_i \mu_i, \sum_{i=1}^{N} \frac{c_i^2}{m_i} \sigma^2\right),$$

也称 $\sum\limits_{i=1}^{N} c_i \mu_i$ 的无偏估计 $\sum\limits_{i=1}^{N} c_i \bar{y}_{i\cdot}$ 为对照;

- 定义对照平方和为

$$SS_c = \frac{\left(\sum_{i=1}^{N} c_i \bar{y}_{i\cdot}\right)^2}{\sum_{i=1}^{N} \frac{c_i^2}{m_i}}.$$

检验: $H_0:\sum\limits_{i=1}^{N} c_i\mu_i = 0,\quad H_1:\sum\limits_{i=1}^{N} c_i\mu_i \neq 0.$

Theorem

当原假设 $H_0:\sum\limits_{i=1}^{N} c_i\mu_i = 0$ 成立时,

(1) $\sum\limits_{i=1}^{N} c_i\bar{y}_{i\cdot} \sim N\left(0, \sum\limits_{i=1}^{N}\frac{c_i^2}{m_i}\sigma^2\right),\ \dfrac{\sum\limits_{i=1}^{N} c_i\bar{y}_{i\cdot}}{\sqrt{MS_E\sum\limits_{i=1}^{N}\frac{c_i^2}{m_i}}} \sim t(f_E);$

(2) $\frac{SS_c}{\sigma^2} \sim \chi^2(1),\ \frac{SS_c}{MS_E} \sim F(1, f_E).$

- 一切对照构成 $N-1$ 维线性空间;
- 考虑等重复情形, 称两个对照互相正交, 如果它们对应的向量互相正交. 正交对照互相独立.
- 平方和分解式:

$$SS_{\text{Total}} = SS_{\text{Treatments}} + SS_{\text{Error}};$$

- 双因子平方和分解式:

$$SS_{\text{Treatments}} = SS_A + SS_B + SS_{A\times B}.$$

- 一切对照构成 $N-1$ 维线性空间;
- 考虑等重复情形, 称两个对照互相正交, 如果它们对应的向量互相正交. 正交对照互相独立.
- 平方和分解式:

$$SS_{\text{Total}} = SS_{\text{Treatments}} + SS_{\text{Error}};$$

- 双因子平方和分解式:

$$SS_{\text{Treatments}} = SS_A + SS_B + SS_{A\times B}.$$

- 处理平方和的自由度为 $N-1$, 是否可分解为 $N-1$ 个互相正交的对照的平方和呢?

Theorem

如果每个处理重复试验次数相等, 则处理的平方和可分解为任意 $N-1$ 个互相正交的对照的平方和.

- 处理平方和的自由度为 $N-1$, 是否可分解为 $N-1$ 个互相正交的对照的平方和呢?

Theorem

如果每个处理重复试验次数相等, 则处理的平方和可分解为任意 $N-1$ 个互相正交的对照的平方和.

小测试

- 对照平方和公式?
- 处理平方和的分解?

Example

设制造某新型手枪共有 A, B, C, D 四种不同工艺. 为研究四种工艺之间的差异, 命 a, b, c, d, e 五个战士打靶, 每人提供 400 发子弹. 命中频率数据如下, 四种工艺是否有差异?

	A	B	C	D
a	0.60	0.59	0.71	0.72
b	0.80	0.81	0.88	0.86
c	0.68	0.64	0.80	0.79
d	0.68	0.70	0.81	0.82
e	0.59	0.60	0.73	0.72
和	3.36	3.34	3.93	3.91
平均	0.672	0.668	0.786	0.782

Example (cont.)

为检验四组工艺之间是否存在差异，构造三个互相正交的对照

$$c_1 = \tau_1 - \tau_2, \quad c_2 = \tau_3 - \tau_4, \quad c_3 = \tau_1 + \tau_2 - \tau_3 - \tau_4.$$

根据对照平方和的定义:

$$SS_{c_1} = \frac{(3.35 - 3.34)^2}{5 \times (1+1)} = 0.00001,$$

$$SS_{c_2} = \frac{(3.93 - 3.91)^2}{5 \times (1+1)} = 0.00004,$$

$$SS_{c_3} = \frac{(3.35 + 3.34 - 3.93 - 3.91)^2}{5 \times (1+1+1+1)} = 0.06613.$$

$SS_{c_1} + SS_{c_2} + SS_{c_3} = 0.06618$ 恰为因子 A 的平方和.

Example (cont.)

为检验四组工艺之间是否存在差异, 构造三个互相正交的对照

$$c_1 = \tau_1 - \tau_2, \quad c_2 = \tau_3 - \tau_4, \quad c_3 = \tau_1 + \tau_2 - \tau_3 - \tau_4.$$

根据对照平方和的定义:

$$SS_{c_1} = \frac{(3.35 - 3.34)^2}{5 \times (1+1)} = 0.00001,$$

$$SS_{c_2} = \frac{(3.93 - 3.91)^2}{5 \times (1+1)} = 0.00004,$$

$$SS_{c_3} = \frac{(3.35 + 3.34 - 3.93 - 3.91)^2}{5 \times (1+1+1+1)} = 0.06613.$$

$SS_{c_1} + SS_{c_2} + SS_{c_3} = 0.06618$ 恰为因子 A 的平方和.

Example (cont.)

利用 $MS_E = 0.00593$, 诸对照的 F 统计量为

$$F_{c_1} = \frac{SS_{c_1}}{MS_E} = 0.00169 < F_{0.95}(1, 16) = 4.494,$$

$$F_{c_2} = \frac{SS_{c_2}}{MS_E} = 0.00675 < F_{0.95}(1, 16) = 4.494,$$

$$F_{c_3} = \frac{SS_{c_2}}{MS_E} = 11.152 > F_{0.95}(1, 16) = 4.494,$$

总结

- 概念:
  (1) 效应与因果 (cause-effects)
  (2) 独立参数与自由度 (degree of freedom)
  (3) 对照
- 思想:
  (1) 犯错概率累积
  (2) 参数化 (parameterize)
- 方法:
  (1) 两样本 t 检验
  (2) 对照的 t 检验与 F 检验

习题

(1) 证明本节中的两个定理.