

2.1.4 多重比较与对照

易泰河

系统工程学院 军事建模与仿真系

2019 年 12 月 27 日

引言

例 (手枪工艺)

设制造某新型手枪共有 A_1, A_2, A_3, A_4 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, 命中频率数据如下, 四种工艺是否有差异?

	A_1	A_2	A_3	A_4
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	3.35	3.34	3.93	3.91
平均	0.670	0.668	0.786	0.782

引言

例 (手枪工艺)

设制造某新型手枪共有 A_1, A_2, A_3, A_4 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, 命中频率数据如下, 四种工艺是否有差异?

	A_1	A_2	A_3	A_4
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	3.35	3.34	3.93	3.91
平均	0.670	0.668	0.786	0.782

模型假定

- 设某试验共 N 个处理, 第 i 个处理重复 m_i 次, 共 $n = m_1 + \cdots + m_N$ 次试验:

处理	观察值				平均
x_1	y_{11}	y_{12}	\cdots	y_{1m_1}	$\bar{y}_{1\cdot}$
x_2	y_{21}	y_{22}	\cdots	y_{2m_2}	$\bar{y}_{2\cdot}$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
x_N	y_{N1}	y_{N2}	\cdots	y_{Nm_N}	$\bar{y}_{N\cdot}$

- 固定效应模型:

$$\begin{cases} y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad \varepsilon_{ij} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2), \\ m_1\tau_1 + m_2\tau_2 + \cdots + m_N\tau_N = 0. \end{cases}$$

问题提出

- 多重比较: 比较任意两处理之间的差异,

$$H_0^{ij} : \tau_i = \tau_j, \quad H_1^{ij} : \tau_i \neq \tau_j; \quad \forall i < j$$

- 对照: 把处理分成两组, 判断两组之间是否存在差异, 如
 - 组 $\{A_1, A_2\}$ 与组 $\{A_3, A_4\}$ 之间是否存在差异?
 - 组 $\{A_1\}$ 与组 $\{A_3, A_4\}$ 之间是否存在差异?

教学目标

- 1 掌握多重比较的 t 检验法;
- 2 掌握对照的 t 检验与 F 检验;
- 3 理解正交对照的概念.

知识回顾

- 如果 $\xi \sim \mathcal{N}(0, 1)$, $\eta \sim \chi^2(n)$, 且互相独立, 则

$$\frac{\xi}{\sqrt{\eta/n}} \sim t(n);$$

- 如果 $\xi \sim \chi^2(m)$, $\eta \sim \chi^2(n)$, 且互相独立, 则

$$F := \frac{\xi/m}{\eta/n} \sim F(m, n);$$

- 如果 $y_1, y_2, \dots, y_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, 则样本均值 \bar{y} 与样本方差 S_n^2 互相独立, 且 $S_n^2 \sim \chi^2(n-1)$.

知识回顾

方差分析:

$$H_0 : (\tau_1, \tau_2, \dots, \tau_N) = \mathbf{0} \text{ vs } H_1 : (\tau_1, \tau_2, \dots, \tau_N) \neq \mathbf{0}.$$

- 总偏差平方和: $SS_T = \sum_{i=1}^N \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_{..})^2;$
- 因子 A 的平方和: $SS_A = \sum_{i=1}^N m_i (\bar{y}_{i.} - \bar{y}_{..})^2;$
- 误差平方和: $SS_E = \sum_{i=1}^N \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_{i.})^2;$
- $F = \frac{SS_A/(N-1)}{SS_E/(n-N)} = \frac{MS_A}{MS_E} \sim F(N-1, n-N).$

2.1.4 多重比较与对照

一 多重比较

- 两样本 t 检验
- 多重比较犯错概率的累积
- 多重比较的 Bonferroni 法

二 对照

- 考虑检验问题: $H_0^{ij} : \tau_i = \tau_j, \quad H_1^{ij} : \tau_i \neq \tau_j.$

- 注意到

$$\bar{y}_i - \bar{y}_j \sim \mathcal{N} \left(\tau_i - \tau_j, \frac{\sigma^2}{m_i} + \frac{\sigma^2}{m_j} \right), \quad \frac{MS_E}{\sigma^2} \sim \chi^2(n-N),$$

- 当原假设 $H_0^{ij} : \mu_i = \mu_j$ 成立时

$$t_{ij} = \frac{\bar{y}_j - \bar{y}_i}{\sqrt{MS_E(1/m_i + 1/m_j)}} \sim t(n-N),$$

两样本 t 检验

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

- 考虑检验问题: $H_0^{ij} : \tau_i = \tau_j, \quad H_1^{ij} : \tau_i \neq \tau_j.$
- 注意到

$$\bar{y}_i - \bar{y}_j \sim \mathcal{N} \left(\tau_i - \tau_j, \frac{\sigma^2}{m_i} + \frac{\sigma^2}{m_j} \right), \quad \frac{MS_E}{\sigma^2} \sim \chi^2(n-N),$$

- 当原假设 $H_0^{ij} : \mu_i = \mu_j$ 成立时

$$t_{ij} = \frac{\bar{y}_j - \bar{y}_i}{\sqrt{MS_E(1/m_i + 1/m_j)}} \sim t(n-N),$$

两样本 t 检验

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

- 考虑检验问题: $H_0^{ij} : \tau_i = \tau_j, \quad H_1^{ij} : \tau_i \neq \tau_j.$
- 注意到

$$\bar{y}_i - \bar{y}_j \sim \mathcal{N} \left(\tau_i - \tau_j, \frac{\sigma^2}{m_i} + \frac{\sigma^2}{m_j} \right), \quad \frac{MS_E}{\sigma^2} \sim \chi^2(n-N),$$

- 当原假设 $H_0^{ij} : \mu_i = \mu_j$ 成立时

$$t_{ij} = \frac{\bar{y}_j - \bar{y}_i}{\sqrt{MS_E(1/m_i + 1/m_j)}} \sim t(n - N),$$

两样本 t 检验

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

- 考虑检验问题: $H_0^{ij} : \tau_i = \tau_j, \quad H_1^{ij} : \tau_i \neq \tau_j.$
- 注意到

$$\bar{y}_i - \bar{y}_j \sim \mathcal{N} \left(\tau_i - \tau_j, \frac{\sigma^2}{m_i} + \frac{\sigma^2}{m_j} \right), \quad \frac{MS_E}{\sigma^2} \sim \chi^2(n-N),$$

- 当原假设 $H_0^{ij} : \mu_i = \mu_j$ 成立时

$$t_{ij} = \frac{\bar{y}_j - \bar{y}_i}{\sqrt{MS_E(1/m_i + 1/m_j)}} \sim t(n - N),$$

两样本 t 检验

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

例 (手枪工艺)

设制造某新型手枪共有 A_1, A_2, A_3, A_4 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, 命中频率数据如下:

	A_1	A_2	A_3	A_4
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	3.35	3.34	3.93	3.91
平均	0.670	0.668	0.786	0.782

工艺 A_1 和工艺 A_2 之间是否存在差异?

例

由于

$$\begin{aligned} |t_{12}| &= \left| \frac{\bar{y}_{2\cdot} - \bar{y}_{1\cdot}}{\sqrt{MS_E (1/m_1 + 1/m_2)}} \right| \\ &= \frac{|0.668 - 0.670|}{\sqrt{0.00593 \times 0.4}} \\ &= 0.04107 \leq t_{0.975}(16) = 2.11991. \end{aligned}$$

在检验水平 0.05 下, A_1 和 A_2 没有显著差异.

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, $C_N^2 = 45$,

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

多重比较的 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

$$P(A_{ij}|H_0^{ij}) = \alpha / C_N^2,$$

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

多重比较的 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

- 两样本检验的 t 统计量?

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

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

小测试

- 两样本检验的 t 统计量?

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

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

2.1.4 多重比较与对照

一 多重比较

二 对照

- 对照的定义
- 对照的 t 检验与 F 检验
- 正交对照与处理平方和的分解

注意到 $\tau_i = \tau_j \iff \mu_i - \mu_j = 0$, 考虑线性假设

$$H_0^c : \sum_{i=1}^N c_i \mu_i = 0.$$

定义 (对照)

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

例: $c = \mu_i - \mu_j$, $c = 2\mu_i - \mu_j - \mu_k$ 都是对照.

注意到 $\tau_i = \tau_j \iff \mu_i - \mu_j = 0$, 考虑线性假设

$$H_0^c : \sum_{i=1}^N c_i \mu_i = 0.$$

定义 (对照)

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

例: $c = \mu_i - \mu_j$, $c = 2\mu_i - \mu_j - \mu_k$ 都是对照.

注意到 $\tau_i = \tau_j \iff \mu_i - \mu_j = 0$, 考虑线性假设

$$H_0^c : \sum_{i=1}^N c_i \mu_i = 0.$$

定义 (对照)

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

例: $c = \mu_i - \mu_j$, $c = 2\mu_i - \mu_j - \mu_k$ 都是对照.

- 由于 $\bar{y}_i \sim \mathcal{N}(\mu_i, \sigma^2/m_i)$, 故

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

也称 $\sum_{i=1}^N c_i \mu_i$ 的无偏估计 $\sum_{i=1}^N c_i \bar{y}_i$ 为对照;

- 定义对照平方和为

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

- 由于 $\bar{y}_{i.} \sim \mathcal{N}(\mu_i, \sigma^2/m_i)$, 故

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

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

- 定义**对照平方和**为

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

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

定理 (对照的显著性检验)

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

$$\textcircled{1} \quad \sum_{i=1}^N c_i \bar{y}_i \sim \mathcal{N} \left(0, \sum_{i=1}^N \frac{c_i^2}{m_i} \sigma^2 \right), \quad \frac{\sum_{i=1}^N c_i \bar{y}_i}{\sqrt{MS_E \sum_{i=1}^N \frac{c_i^2}{m_i}}} \sim t(f_E);$$

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

思考: 对照的 t 检验与 F 检验之间有什么关系?

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

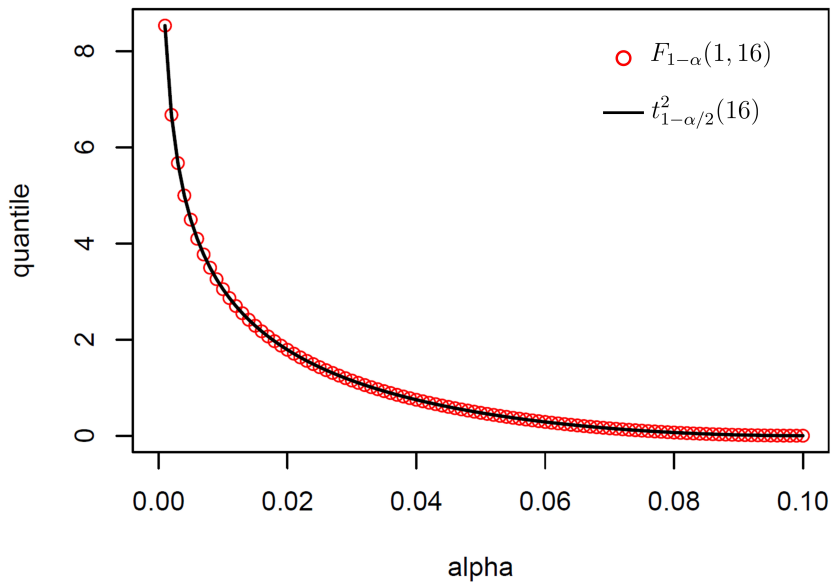
定理 (对照的显著性检验)

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

$$\textcircled{1} \quad \sum_{i=1}^N c_i \bar{y}_i \sim \mathcal{N} \left(0, \sum_{i=1}^N \frac{c_i^2}{m_i} \sigma^2 \right), \quad \frac{\sum_{i=1}^N c_i \bar{y}_i}{\sqrt{MS_E \sum_{i=1}^N \frac{c_i^2}{m_i}}} \sim t(f_E);$$

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

思考: 对照的 t 检验与 F 检验之间有什么关系?



小测试

- 上述定理对于一般的线性假设 $H_0 : \sum_{i=1}^N c_i \mu_i = 0$ 是否也适用?
- $\mu_1 - \mu_2$ 和 $\frac{1}{2}\mu_1 - \frac{1}{2}\mu_2$ 都是对照, 它们的平方和是否相等?

- 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:

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

- 将多因子试验视作单因子试验**不改变自由度!**

- 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:

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

- 将多因子试验视作单因子试验不改变自由度!

- 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:

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

- 将多因子试验视作单因子试验不改变自由度!

- 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?

定理 (处理平方和按正交对照分解)

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?

定理 (处理平方和按正交对照分解)

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.

例 (手枪工艺)

设制造某新型手枪共有 A_1, A_2, A_3, A_4 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, 命中频率数据如下, 四种工艺是否有差异?

	A_1	A_2	A_3	A_4
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	3.35	3.34	3.93	3.91
平均	0.670	0.668	0.786	0.782

例 (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.

例 (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.

例 (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

- ① 概念: 多重比较与对照;
- ② 思想: 犯错概率累积;
- ③ 方法: t 检验与 F 检验,
 - 多重比较的 t 检验, Bonferroni 法;
 - 对照的 t 检验和 F 检验;
 - 等重复情形, 处理平方和可分解为任意 $N - 1$ 个互相正交的对照的平方和.

习题

- (1) 手枪工艺案例中, 利用 t 法判断 A_2 与 A_3 之间差异的显著性;
- (2) 手枪工艺案例中, 判断对照 $2\mu_1 - \mu_2 - \mu_3$ 的显著性;
- (3) 证明处理平方和按正交对照分解的定理.