2.3 3^k 因子设计及其部分实施

王正明　易泰河

系统工程学院　军事建模与仿真系

2019 年 11 月 13 日

引言

- 2^k 设计与正交表 $L_{2^k}(2^{2^k-1})$.
- 混杂, 别名, 定义对比
- 如何把 2^k 因子设计及其部分实施的方法推广到 3^k 因子设计中来?

2.3 3^2 因子设计与正交表 $L_9(3^4)$

2.3.1 3^2 因子设计与正交表 $L_9(3^4)$

2.3.2 3^k 因子设计与正交表 $L_{3^k}(3^{\frac{3^k-1}{3-1}})$

2.3.3 3^k 因子试验的部分实施

(1) 3^2 设计的固定效应模型

- 设 3^2 设计的 2 个因子为 A 与 B, 3 个水平分别以 0、1 和 2 表示, 全部处理共 9 个: $\{(i,j):i,j=0,1,2\}$.

- 设每个处理重复 m 次, 则 3^2 试验的线性可加模型为

$$\begin{cases} y_{ijk}=\mu+\tau_i+\beta_j+(\tau\beta)_{ij}+\varepsilon_{ijk}, \\ \varepsilon_{ijk}\overset{\text{i.i.d.}}{\sim}N(0,\sigma^2), \quad i,j=0,1,2, \quad k=1,\cdots,m, \\ \sum_{i=0}^{3}\tau_i=0, \quad \sum_{j=0}^{3}\beta_j=0, \\ \sum_{i=0}^{2}(\tau\beta)_{ij}=0, \quad j=0,1,2, \quad \sum_{j=0}^{2}(\tau\beta)_{ij}=0, \quad i=0,1,2. \end{cases}$$

- 诸效应的估计?

(2) 利用正交表 $L_9(3^4)$ 安排 3^2 设计的试验

试验号	A	B	AB	A^2B
1	0	0	0	0
2	0	1	1	1
3	0	2	2	2
4	1	0	1	2
5	1	1	2	0
6	1	2	0	1
7	2	0	1	1
8	2	1	0	2
9	2	2	1	0

- A 为三分列, B 为九分列, 构成 3^2 设计的全面实施;
- 交互作用列 $AB: x_1 + x_2 \mod 3$, 交互作用列 $A^2B: 2x_1 + x_2 \mod 3$, 表示交互作用 $A \times B$ 的两个部分.

试验号	A	B	AB	A^2B
1	0	0	0	0
2	0	1	1	1
3	0	2	2	2
4	1	0	1	2
5	1	1	2	0
6	1	2	0	1
7	2	0	1	1
8	2	1	0	2
9	2	2	1	0

- 每列的自由度均为 2, 四列一共 8 个自由度, 恰为 9 个处理的自由度.
- 任意两列按照两种模 3 加法运算得到剩余两列, 称这种关系为任意两列的交互作用是其余两列.

- 如果表 $L_9(3^4)$ 的第四列按照 $x_1 + 2x_2 \mod 3$ 生成，得到新的正交表 $L_9(3^4)$. 这两张正交表有什么区别?

试验号	A	B	AB	AB^2
1	0	0	0	0
2	0	1	1	2
3	0	2	2	1
4	1	0	1	1
5	1	1	2	0
6	1	2	0	2
7	2	0	1	2
8	2	1	0	1
9	2	2	1	0

(3) 利用正交表 $L_9(3^4)$ 作 3^2 设计的方差分析

试验号	A	B	AB	A^2B	试验结果
1	0	0	0	0	$y_{00\cdot}$
2	0	1	1	1	$y_{01\cdot}$
3	0	2	2	2	$y_{02\cdot}$
4	1	0	1	2	$y_{10\cdot}$
5	1	1	2	0	$y_{11\cdot}$
6	1	2	0	1	$y_{12\cdot}$
7	2	0	2	1	$y_{20\cdot}$
8	2	1	0	2	$y_{21\cdot}$
9	2	2	1	0	$y_{22\cdot}$
自由度	2	2	2	2	
T_0	$y_{0\cdot\cdot}$	$y_{\cdot 0\cdot}$	$y_{00\cdot}+y_{12\cdot}+y_{21\cdot}$	$y_{00\cdot}+y_{11\cdot}+y_{22\cdot}$	
T_1	$y_{1\cdot\cdot}$	$y_{\cdot 1\cdot}$	$y_{01\cdot}+y_{10\cdot}+y_{22\cdot}$	$y_{01\cdot}+y_{12\cdot}+y_{20\cdot}$	$T=\sum_{i=0}^{2}\sum_{j=0}^{2} y_{ij\cdot}$
T_2	$y_{2\cdot\cdot}$	$y_{\cdot 2\cdot}$	$y_{02\cdot}+y_{11\cdot}+y_{20\cdot}$	$y_{02\cdot}+y_{10\cdot}+y_{21\cdot}$	

(3) 利用正交表 $L_9(3^4)$ 作 3^2 设计的方差分析

- 平方和计算公式

$$\begin{cases} SS_A = 3m\sum_{i=0}^{2}(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot})^2 = \dfrac{T_{A_0}^2 + T_{A_1}^2 + T_{A_2}^2}{3m} - \dfrac{T^2}{9m}, \\ SS_B = 3\sum_{j=0}^{2}(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2 = \dfrac{T_{B_0}^2 + T_{B_1}^2 + T_{B_2}^2}{3} - \dfrac{T^2}{9}, \\ SS_{AB} = \dfrac{T_{(AB)_0}^2 + T_{(AB)_1}^2 + T_{(AB)_2}^2}{3m} - \dfrac{T^2}{9m}, \\ SS_{A^2B} = \dfrac{T_{(A^2B)_0}^2 + T_{(A^2B)_1}^2 + T_{(A^2B)_2}^2}{3m} - \dfrac{T^2}{9m}, \\ SS_{A\times B} = SS_{AB} + SS_{A^2B}. \end{cases}$$

其中, $T_{(\cdot)_i}$ 分别表示对应列的 T_i.

- 每列的自由度都为 2, 交互作用 $A \times B$ 的自由度为 4, 被分解为两个自由度均为 2 的部分.
- 利用 F 统计量检验因子 A 的效应是否显著

$$F = \frac{SS_A/2}{SS_E/f_E},$$

当 A 的效应不显著时, 该统计量服从自由度为 $(2, f_E)$ 的 F 分布.

- 在 2^k 设计中, 以 "AB" 表示两因子交互效应; 这里以 "$A \times B$" 表示因子 A 和 B 的交互作用, AB 和 A^2B 分别表示它的两个互相正交的部分, 为什么?

2.3 3^2 因子设计与正交表 $L_9(3^4)$

(1) 3^k 设计的效应及其自由度

- 每个因子主效应的自由度为 2;
- C_k^2 个二因子交互效应, 自由度均为 $(3-1)^2=4$;
- C_k^3 个三因子交互效应, 自由度均为 $(3-1)^3=8$;
- 一般地, 有 $C_k^h(h \le k)$ 个 h 因子交互效应, 每个的自由度为 $(3-1)^h=2^h$.
- 如果每个处理均重复 m 次试验, 则有 $3^k m-1$ 个总自由度和

$$(3^k m-1)-\sum_{h=1}^{k} C_k^h 2^h = 3^k(m-1)$$

个误差自由度.

Example

以 3^3 因子试验为例,

- 每个因子的主效应的自由度为 ____,
- 二因子交互效应的自由度为 ____,
- 三因子交互效应的自由度为 ____.
- 如果每个处理组合均重复试验 m 次, 则总自由度为 ____, 误差自由度为 ____.

(3) 正交表 $L_{3^k}(3^{\frac{3^k-1}{3-1}})$ 的构造:

Step 1 构造三分列, 该列的前 3^{k-1} 行置水平 0, 中间 3^{k-1} 行置水平 1, 后 3^{k-1} 行置水平 2, 列名记作 A;

Step 2 构造九分列, 列名记作 B, 并按照两种模 3 加法运算构造交互作用列 AB 和 A^2B;

$\cdots$

Step k 构造 3^k 分列, 该列的各行按照 $0,1,2,0,1,2,\cdots$ 的次序交替排列, 列名记作 K, 并利用两种模 3 加法运算依此得到 3^k 分列与前面各列的交互作用列 AK, A^2K, BK, B^2K, $\cdots$.

其中, 三分列、九分列、2^k 分列统称为基本列.

Example

与 3^3 设计对应的正交表是 $L_{27}(3^{13})$, 它一共有 27 行 13 列, 第一列、第二列和第五列为 3^3 设计的全面实施, 其余 10 列为交互作用列. 其构造步骤如下:

Step 1 构造三分列

$A = [0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,2]^{\mathrm{T}}$;

Step 2 构造九分列

$B = [0,0,0,1,1,1,2,2,2,0,0,0,1,1,1,2,2,2,0,0,0,1,1,1,2,2,2]^{\mathrm{T}}$,

并利用两种模 3 加法运算得到第三列 AB 和第四列 A^2B;

Step 3 构造二十七分列

$C = [0,1,2,0,1,2,0,1,2,0,1,2,0,1,2,0,1,2,0,1,2,0,1,2,0,1,2]^{\mathrm{T}}$,

并利用两种模 3 加法运算依此得到第六列 AC、第七列 A^2C、第八列 BC、第九列 B^2C、第十列 ABC、第十一列 A^2B^2C、第十二列 A^2BC 和第十三列 AB^2C.

Example (Cont.)

AB 和 A^2B 两列可用于计算交互效应 $A \times B$ 的平方和, 即

$$SS_{A\times B} = SS_{AB} + SS_{A^2B};$$

类似地, AC 和 A^2C 两列可用于计算交互效应 $A \times C$ 的平方和, BC 和 B^2C 两列可用于计算交互效应 $B \times C$ 的平方和; ABC、A^2B^2C、A^2BC 和 AB^2C 四列可用于计算交互效应 $A \times B \times C$ 的平方和, 即

$$SS_{A\times B\times C} = SS_{ABC} + SS_{A^2B^2C} + SS_{A^2BC} + SS_{AB^2C}.$$

各列的平方和计算公式与 3^2 中一样. 以因子 A 为例, 其平方和计算公式为

$$SS_A = \frac{T_{A_0}^2 + T_{A_1}^2 + T_{A_2}^2}{9m} - \frac{T^2}{27m}.$$

- 称一个列名为标准化列名, 如果它最后一个字母的指数是 1, 否则称为非标准化列名.
- 称一组列名为完备的, 如果该组列名包含所有基本列的列名和任意两列的列名运算得到的列名.
- 称一组列名为标准化完备列名, 如果它们既是标准化的又是完备的.
- 按照前面介绍的标准化方法构造的正交表 $L_{2^k}(2^{2^k-1})$ 和 $L_{3^k}(3^{\frac{3^k-1}{3-1}})$ 的列名都是标准化完备列名.

Example

正交表 $L_9(3^4)$ 中, 列名 A^2B 是标准化列名, 而 AB^2 为非标准化列名.

(1) $L_8(2^7)$ 的列名组成的集合 $\{A, B, AB, C, AC, BC, ABC\}$ 是标准化完备的;

(2) 从正交表 $L_8(2^7)$ 中挑选出的列名集合 $\{A, BC, ABC\}$ 是不完备的, 它虽然对列名运算封闭, 但是没有包含基本列名 B 和 C;

(3) $L_9(3^4)$ 的列名组成的集合 $\{A, B, AB, A^2B\}$ 是标准化完备的.

- 在 3^k 因子设计中, 标准化列名与非标准化列名可通过列名的平方互相转化:

$$(A^2B)^2 = A^4B^2 = AB^2, \qquad (AB^2)^2 = A^2B^4 = A^2B.$$

- 列名的平方运算恰好对应水平置换: $0 \to 0$, $1 \to 2$, $2 \to 1$. 表明标准化列名与非标准化列名对应的正交表互相同构.
- 注意, 二水平正交表列名运算是在代数运算的基础上对幂指数进行模 2 运算, 而三水平正交表的列名运算是在代数运算的基础上对幂指数进行模 3 运算.

- k 越大，3^k 设计高阶交互作用越多，全面试验次数越多;
- 如果有高阶交互效应不显著的先验信息，那么可以采用部分实施;
- 如果没有高阶交互效应不显著的先验信息，部分实施带来效应混杂，通常假定:
  - 实际问题中，往往只有少部分两因子交互作用存在，而其它两因子交互作用和高阶因子交互作用都不存在，此即效应稀疏原则;
  - 尽量考虑低阶交互效应和放弃高阶交互效应，即效应有序原则.

Example

当已知 3^3 设计的因子间的交互作用都不存在时，可使用正交表 $L_9(3^4)$ 安排试验，有两种可供选择的表头设计方案:

列名	A	B	AB	A^2B	定义关系	分辨度
方案一	A	B	C		$\mathbf{I}=ABC^2$	III
方案二	A	B		C	$\mathbf{I}=A^2BC^2$	III

这里 $\mathbf{I}$ 表示全部由 0 组成的列. 表中还有一个空白列，在作方差分析时，这个列的平方和可以作为误差平方和. 表中两个方案都只包含 3^3 设计的 9 个试验点，都称为3^3 设计的一个 $1/3$ 实施，或3^3 设计的一个 3^{3-1} 实施. 它们的分辨度都为 III(字的平方也只能算一个字).

Example

当 3^4 设计的因子间的交互作用不存在时, 可以采用正交表 $L_9(3^4)$ 安排试验, 表头设计如下:

列名	A	B	AB	A^2B	定义关系	分辨度
因子	A	B	C	D	$\mathbf{I} = ABC^2 = A^2BD^2$	III

表中已没有空白列, $L_9(3^4)$ 的 8 个自由度全部用来估计因子主效应了. 如果要作方差分析, 需要作重复试验, 才能计算误差平方和. 这个设计只包含 3^4 设计的 9 个试验点, 称为3^4 设计的一个 $1/9$ 实施, 或3^4 设计的一个 3^{4-2} 实施.

总结

1. 3^2 设计与正交表 $L_9(3^4)$;
2. 正交表 $L_{3^k}(3^{\frac{3^k-1}{3-1}})$ 的构造方法;
3. 概念: 标准化列名, 完备列名组与完备正交表, 定义关系, 分辨度, 字长.

习题

1. 利用正交表 $L_{27}(3^{15})$ 可以把交互作用 $A\times B\times C$ 的平方和分解为
$$SS_{A\times B\times C}=SS_{ABC}+SS_{A^2B^2C}+SS_{A^2BC}+SS_{AB^2C},$$
试写出每一分量的表达式.
2. 试利用正交表 $L_{27}(3^{15})$ 分析讲义中例 2.5 的数据.
3. 能否根据正交表 $L_{2^k}(2^{2^k-1})$ 和 $L_{3^k}(3^{\frac{3^k-1}{3-1}})$ 的构造规律, 归纳出一般的完备正交表 $L_{q^k}(q^{\frac{q^k-1}{q-1}})$ 的构造方法.

有什么问题?