2.2 2^k 因子设计及其部分实施（二）

王正明　易泰河

系统工程学院　军事建模与仿真系

2019 年 11 月 27 日

知识回顾

定义

称由一些符号组成的矩阵为**正交表**(orthogonal table)，如果任意两列中同行符号构成的若干符号对的重复次数相等.

- 2^2 设计与正交表 $L_4(2^3)$;
- 2^3 设计与正交表 $L_8(2^7)$;
- 正交表的等价与同构;
- 利用正交表 $L_8(2^7)$ 设计试验和分析数据.

2.2 2^k 因子设计及其部分实施

2.2.1 2^2 设计与正交表 $L_4(2^3)$

2.2.2 2^3 设计与正交表 $L_8(2^7)$

2.2.3 2^k 设计与正交表 $L_{2^k}(2^{2^k-1})$

(1) 处理与效应

(2) 正交表 $L_{2^k}(2^{2^k-1})$ 的构造

(3) 正交表 $L_{2^k}(2^{2^k-1})$ 的应用

2.2.4 2^k 因子试验的部分实施

- 2^k 设计的处理是 k 维空间中顶点坐标用 0 或者 1 表示的立方体的 2^k 个顶点.
- 处理的记号: 从后往前依次引入试验因子, 每引入一个新的因子, 就依次和已引入的因子组合.
- 如 2^4 设计:

$$(1), d, c, cd, b, bd, bc, bcd,$$

$$a, ad, ac, acd, ab, abd, abc, abcd.$$

(1) 表示处理 $(0,0,0,0)$, d 表示处理 $(0,0,0,1)$, $\cdots$.

- 2^k 设计的处理是 k 维空间中顶点坐标用 0 或者 1 表示的立方体的 2^k 个顶点.
- 处理的记号: 从后往前依次引入试验因子, 每引入一个新的因子, 就依次和已引入的因子组合.
- 如 2^4 设计:

$$(1), d, c, cd, b, bd, bc, bcd,$$

$$a, ad, ac, acd, ab, abd, abc, abcd.$$

(1) 表示处理 $(0,0,0,0)$, d 表示处理 $(0,0,0,1)$, $\cdots$.

- 2^k 设计的处理是 k 维空间中顶点坐标用 0 或者 1 表示的立方体的 2^k 个顶点.
- 处理的记号: 从后往前依次引入试验因子, 每引入一个新的因子, 就依次和已引入的因子组合.
- 如 2^4 设计:

$$(1), d, c, cd, b, bd, bc, bcd,$$

$$a, ad, ac, acd, ab, abd, abc, abcd.$$

(1) 表示处理 $(0,0,0,0)$, d 表示处理 $(0,0,0,1)$, $\cdots$.

- k 个主效应, C_k^2 个二因子交互效应, C_k^3 个三因子交互效应, $\cdots$, 以及 1 个 k 因子交互效应;
- 一共$C_k^1 + C_k^2 + \cdots + C_k^k = 2^k - 1$个效应, 每个效应的自由度均为 1.
- 如果每个处理重复 m 次, 则总自由度为 $2^k m - 1$, 误差自由度为 $2^k m - 1 - 2^k + 1 = 2^k(m-1)$.

- k 个主效应, C_k^2 个二因子交互效应, C_k^3 个三因子交互效应, $\cdots$, 以及 1 个 k 因子交互效应;
- 一共 $C_k^1 + C_k^2 + \cdots + C_k^k = 2^k - 1$ 个效应, 每个效应的自由度均为 1.
- 如果每个处理重复 m 次, 则总自由度为 $2^k m - 1$, 误差自由度为 $2^k m - 1 - 2^k + 1 = 2^k(m-1)$.

- k 个主效应, C_k^2 个二因子交互效应, C_k^3 个三因子交互效应, $\cdots$, 以及 1 个 k 因子交互效应;
- 一共$C_k^1 + C_k^2 + \cdots + C_k^k = 2^k - 1$个效应, 每个效应的自由度均为 1.
- 如果每个处理重复 m 次, 则总自由度为 $2^k m - 1$, 误差自由度为 $2^k m - 1 - 2^k + 1 = 2^k(m-1)$.

- 效应的估计为:

$$AB\cdots K=\frac{1}{2^{k-1}m}(a\pm 1)(b\pm 1)\cdots(k\pm 1)$$

  - 如果左边有某个因子时, 右边相应括号内取 “$-$”, 否则取 “$+$”;
  - 右侧按代数方法展开后, 以 (1) 代替 1, 每一字母组合表示对应处理处 m 次试验观察值的总和.

- 效应的平方和为:

$$SS_{AB\cdots K}=\frac{1}{2^k m}\left[(a\pm 1)(b\pm 1)\cdots(k\pm 1)\right]^2.$$

- 效应的估计为:

$$AB\cdots K=\frac{1}{2^{k-1}m}(a\pm 1)(b\pm 1)\cdots(k\pm 1)$$

  - 如果左边有某个因子时, 右边相应括号内取 “$-$”, 否则取 “$+$”;
  - 右侧按代数方法展开后, 以 (1) 代替 1, 每一字母组合表示对应处理处 m 次试验观察值的总和.

- 效应的平方和为:

$$SS_{AB\cdots K}=\frac{1}{2^k m}\left[(a\pm 1)(b\pm 1)\cdots(k\pm 1)\right]^2.$$

小测试

1. 2^3 设计中, 交互效应 AC 的估计和平方和分别为

$$AC = \underline{\hspace{10cm}},$$

$$SS_{AC} = \underline{\hspace{10cm}},$$

2. 2^4 设计中, 交互效应 AC 的估计和平方和分别为

$$AC = \underline{\hspace{10cm}},$$

$$SS_{AC} = \underline{\hspace{10cm}}.$$

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

一 利用估计诸效应的对照的符号

Step 1 按字母正序安排列: 首先引入 A 列, 每引入一个新的列就依次引入它与前面各列的交互效应列, 如 $L_8(2^7)$ 的列为 A, B, AB, C, AC, BC, ABC;

Step 2 按字母反序安排行: 首先引入 (1), 每引入一个新的因子就依次和前面已引入的因子组合, 如 $L_8(2^7)$ 的行 $(1), c, b, bc, a, ac, ab, abc$;

Step 3 计算各列效应的估计的系数, 如 $L_8(2^7)$ 的 A 列

$$A = \frac{1}{4m}\left[-(1) - c - b - bc + a + ac + ab + abc\right]$$

故 A 列为 $[-,-,-,-,+,+,+,+]^{\mathrm{T}}$.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

— 利用估计诸效应的对照的符号

Step 1 按字母正序安排列：首先引入 A 列，每引入一个新的列就依次引入它与前面各列的交互效应列，如 $L_8(2^7)$ 的列为 A, B, AB, C, AC, BC, ABC;

Step 2 按字母反序安排行：首先引入 (1)，每引入一个新的因子就依次和前面已引入的因子组合，如 $L_8(2^7)$ 的行 $(1), c, b, bc, a, ac, ab, abc$;

Step 3 计算各列效应的估计的系数，如 $L_8(2^7)$ 的 A 列

$$A = \frac{1}{4m}\left[-(1) - c - b - bc + a + ac + ab + abc\right]$$

故 A 列为 $[-, -, -, -, +, +, +, +]^{\mathrm{T}}$.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

— 利用估计诸效应的对照的符号

Step 1 按字母正序安排列：首先引入 A 列，每引入一个新的列就依次引入它与前面各列的交互效应列，如 $L_8(2^7)$ 的列为 A, B, AB, C, AC, BC, ABC;

Step 2 按字母反序安排行：首先引入 (1)，每引入一个新的因子就依次和前面已引入的因子组合，如 $L_8(2^7)$ 的行 $(1), c, b, bc, a, ac, ab, abc$;

Step 3 计算各列效应的估计的系数，如 $L_8(2^7)$ 的 A 列

$$A = \frac{1}{4m}[-(1) - c - b - bc + a + ac + ab + abc]$$

故 A 列为 $[-,-,-,-,+,+,+,+]^{\mathrm{T}}$.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

一 利用估计诸效应的对照的符号

Step 1 按字母正序安排列: 首先引入 A 列, 每引入一个新的列就依次引入它与前面各列的交互效应列, 如 $L_8(2^7)$ 的列为 A, B, AB, C, AC, BC, ABC;

Step 2 按字母反序安排行: 首先引入 (1), 每引入一个新的因子就依次和前面已引入的因子组合, 如 $L_8(2^7)$ 的行 $(1), c, b, bc, a, ac, ab, abc$;

Step 3 计算各列效应的估计的系数, 如 $L_8(2^7)$ 的 A 列

$$A = \frac{1}{4m}[-(1) - c - b - bc + a + ac + ab + abc]$$

故 A 列为 $[-, -, -, -, +, +, +, +]^{\mathrm{T}}$.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

一 利用估计诸效应的对照的符号

二 利用基本列和交互效应列等列名运算

Step 1 构造二分列, 前 2^{k-1} 行置水平 0, 后 2^{k-1} 行置水平 1, 列名记为 A;

Step 2 构造四分列, 将二分列的两部分再次二分, 然后利用模 2 加法运算依次得到四分列与前面各列的交互效应列;

$\vdots$

Step k 构造 2^k 分列, 该列的各行两个水平交替安排, 并利用模 2 加法运算依次得到 2^k 分列与前面各列的交互效应列.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

一　利用估计诸效应的对照的符号

二　利用基本列和交互效应列等列名运算

Step 1　构造二分列, 前 2^{k-1} 行置水平 0, 后 2^{k-1} 行置水平 1, 列名记为 A;

Step 2　构造四分列, 将二分列的两部分再次二分, 然后利用模 2 加法运算依次得到四分列与前面各列的交互效应列;

$\vdots$

Step k　构造 2^k 分列, 该列的各行两个水平交替安排, 并利用模 2 加法运算依次得到 2^k 分列与前面各列的交互效应列.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

一 利用估计诸效应的对照的符号

二 利用基本列和交互效应列等列名运算

Step 1 构造二分列, 前 2^{k-1} 行置水平 0, 后 2^{k-1} 行置水平 1, 列名记为 A;

Step 2 构造四分列, 将二分列的两部分再次二分, 然后利用模 2 加法运算依次得到四分列与前面各列的交互效应列;

$\vdots$

Step k 构造 2^k 分列, 该列的各行两个水平交替安排, 并利用模 2 加法运算依次得到 2^k 分列与前面各列的交互效应列.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

一 利用估计诸效应的对照的符号

二 利用基本列和交互效应列等列名运算

Step 1 构造二分列, 前 2^{k-1} 行置水平 0, 后 2^{k-1} 行置水平 1, 列名记为 A;

Step 2 构造四分列, 将二分列的两部分再次二分, 然后利用模 2 加法运算依次得到四分列与前面各列的交互效应列;

$\vdots$

Step k 构造 2^k 分列, 该列的各行两个水平交替安排, 并利用模 2 加法运算依次得到 2^k 分列与前面各列的交互效应列.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

一 利用估计诸效应的对照的符号

二 利用基本列和交互效应列等列名运算

Step 1 构造二分列, 前 2^{k-1} 行置水平 0, 后 2^{k-1} 行置水平 1, 列名记为 A;

Step 2 构造四分列, 将二分列的两部分再次二分, 然后利用模 2 加法运算依次得到四分列与前面各列的交互效应列;

$\vdots$

Step k 构造 2^k 分列, 该列的各行两个水平交替安排, 并利用模 2 加法运算依次得到 2^k 分列与前面各列的交互效应列.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

- 表 $L_{2^k}(2^{2^k-1})$ 的列名运算规则：首先按照代数运算化简，然后将幂指数按照模 2 运算化简.
- 如：

$$ABC \times BC = AB^2C^2 = AB^0C^0 = A,$$

表明 ABC 列和 BC 列的交互效应列为 A 列.
- 利用列名运算可得到任意两列的交互效应列.
- $L_{2^k}(2^{2^k-1})$ 型正交表是完备的.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

- 表 $L_{2^k}(2^{2^k-1})$ 的列名运算规则: 首先按照代数运算化简, 然后将幂指数按照模 2 运算化简.
- 如:

$$ABC \times BC = AB^2C^2 = AB^0C^0 = A,$$

表明 ABC 列和 BC 列的交互效应列为 A 列.

- 利用列名运算可得到任意两列的交互效应列.
- $L_{2^k}(2^{2^k-1})$ 型正交表是完备的.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

- 表 $L_{2^k}(2^{2^k-1})$ 的列名运算规则: 首先按照代数运算化简, 然后将幂指数按照模 2 运算化简.
- 如:

$$ABC \times BC = AB^2C^2 = AB^0C^0 = A,$$

表明 ABC 列和 BC 列的交互效应列为 A 列.

- 利用列名运算可得到任意两列的交互效应列.
- $L_{2^k}(2^{2^k-1})$ 型正交表是完备的.

$L_{2^k}(2^{2^k-1})$ 型正交表的两种构造方式

- 表 $L_{2^k}(2^{2^k-1})$ 的列名运算规则: 首先按照代数运算化简, 然后将幂指数按照模 2 运算化简.
- 如:

$$ABC \times BC = AB^2C^2 = AB^0C^0 = A,$$

表明 ABC 列和 BC 列的交互效应列为 A 列.

- 利用列名运算可得到任意两列的交互效应列.
- $L_{2^k}(2^{2^k-1})$ 型正交表是完备的.

小测试

1. 试利用对照的系数符号构造正交表 $L_8(2^7)$.
2. 试利用列名运算构造正交表 $L_{16}(2^{15})$.

- 称二分列、$\cdots$、2^k 分列为主效应列或基本列, 2^k 完全因子设计中主效应列用于安排试验;
- 效应的估计: $(T_{i1}-T_{i0})/(2^{k-1}m)$, T_{i0} 表示第 i 列中 0 水平对应的数据之和, T_{i1} 表示第 i 列中 1 水平对应的数据之和;
- 诸效应的平方和可由相应行来计算:

$$SS_i=\frac{T_{i0}^2+T_{i1}^2}{2^{k-1}m}-\frac{T^2}{2^km},$$

$T=T_{i0}+T_{i1}$ 表示所有数据的和;

- 误差平方和: $SS_E=SS_T-SS_A-SS_B-SS_{AB}-\cdots$.

- 称二分列、$\cdots$、2^k 分列为主效应列或基本列, 2^k 完全因子设计中主效应列用于安排试验;
- 效应的估计: $(T_{i1} - T_{i0})/(2^{k-1}m)$, T_{i0} 表示第 i 列中 0 水平对应的数据之和, T_{i1} 表示第 i 列中 1 水平对应的数据之和;
- 诸效应的平方和可由相应行来计算:

$$SS_i = \frac{T_{i0}^2 + T_{i1}^2}{2^{k-1}m} - \frac{T^2}{2^k m},$$

$T = T_{i0} + T_{i1}$ 表示所有数据的和;

- 误差平方和: $SS_E = SS_T - SS_A - SS_B - SS_{AB} - \cdots$.

- 称二分列、$\cdots$、2^k 分列为主效应列或基本列, 2^k 完全因子设计中主效应列用于安排试验;
- 效应的估计: $(T_{i1} - T_{i0})/(2^{k-1}m)$, T_{i0} 表示第 i 列中 0 水平对应的数据之和, T_{i1} 表示第 i 列中 1 水平对应的数据之和;
- 诸效应的平方和可由相应行来计算:

$$SS_i = \frac{T_{i0}^2 + T_{i1}^2}{2^{k-1}m} - \frac{T^2}{2^k m},$$

$T = T_{i0} + T_{i1}$ 表示所有数据的和;

- 误差平方和: $SS_E = SS_T - SS_A - SS_B - SS_{AB} - \cdots$.

- 称二分列、$\cdots$、2^k 分列为主效应列或基本列, 2^k 完全因子设计中主效应列用于安排试验;
- 效应的估计: $(T_{i1}-T_{i0})/(2^{k-1}m)$, T_{i0} 表示第 i 列中 0 水平对应的数据之和, T_{i1} 表示第 i 列中 1 水平对应的数据之和;
- 诸效应的平方和可由相应行来计算:

$$SS_i=\frac{T_{i0}^2+T_{i1}^2}{2^{k-1}m}-\frac{T^2}{2^k m},$$

$T=T_{i0}+T_{i1}$ 表示所有数据的和;

- 误差平方和: $SS_E=SS_T-SS_A-SS_B-SS_{AB}-\cdots$.

例

在梳棉机上纺粘锦混纺纱, 为了提高质量, 选了 3 个因子, 每个因子 2 个水平,

- A 金属针布: $A_1 =$ 日本产, $A_2 =$ 青岛产;
- B 产量水平: $B_1 = 6$ 公斤, $B_2 = 10$ 公斤;
- C 锡林速度: $C_1 = 238$ 转/分, $C_2 = 320$ 转/分.

实践经验表明, 因子间可能有二因子交互效应.

例 (Cont.)

- 用正交表 $L_8(2^7)$ 来安排试验:

列号	1	2	3	4	5	6	7
因子	A	B	AB	C	AC	BC	ABC

试验号	$A(1)$	$B(2)$	$C(4)$
1	日本产	6	238
2	日本产	6	320
3	日本产	10	238
4	日本产	10	320
5	青岛产	6	238
6	青岛产	6	320
7	青岛产	10	238
8	青岛产	10	320

试验号	$A(1)$	$B(2)$	$AB(3)$	$C(4)$	$AC(5)$	$BC(6)$	$ABC(7)$	棉结粒数
1	0	0	0	0	0	0	0	$(1)=0.30$
2	0	0	0	1	1	1	1	$c=0.35$
3	0	1	1	0	0	1	1	$b=0.20$
4	0	1	1	1	1	0	0	$bc=0.30$
5	1	0	1	0	1	0	1	$a=0.15$
6	1	0	1	1	0	1	0	$ac=0.50$
7	1	1	0	0	1	1	0	$ab=0.15$
8	1	1	0	1	0	0	1	$abc=0.40$
T_0	1.15	1.30	1.20	0.80	1.40	1.15	1.25	$T=2.35$
T_1	1.20	1.05	1.15	1.55	0.95	1.20	1.10	
m_0	0.2875	0.3250	0.3000	0.2000	0.3500	0.2875	0.3125	
m_1	0.3000	0.2625	0.2875	0.3875	0.2375	0.3000	0.2750	
R	0.0125	0.0625	0.0125	0.1875	0.1125	0.0125	0.0375	

- $R := \max\{m_0, m_1\} - \min\{m_0, m_1\}$ 为极差, 可用来衡量诸效应的主次关系, 极差越大表明相应的效应越大.
- 2^k 因子设计中, 极差就是效应的估计的绝对值!

两种方式计算各列的偏差平方和,

- 采用计算对照偏差平方和的公式:

$$
\begin{aligned}
SS_A &= \frac{1}{8}\left[abc + ab + ac + a - bc - b - c - (1)\right]^2 \\
&= \frac{1}{8}\left[T_{A_1} - T_{A_0}\right]^2 = 0.0003125;
\end{aligned}
$$

- 正交表中常用的偏差平方和计算公式:

$$
\begin{aligned}
SS_A &= \frac{T_{A_0}^2 + T_{A_1}^2}{4} - \frac{T^2}{8} = \frac{1.15^2 + 1.20^2}{4} - \frac{2.35^2}{8} \\
&= 0.0003125.
\end{aligned}
$$

方差来源	平方和	自由度	均方	F 值	p 值
B	0.0078125	1	0.0078125	8.33	0.0447
C	0.0203125	1	0.0203125	75.00	0.0010
AC	0.0253125	1	0.0253125	27.00	0.0065
误差	0.0037500	4	0.0009375		
A	0.0003125	1			
AB	0.0003125	1			
BC	0.0003125	1			
ABC	0.0028125	1			
总和	0.1071825	7			

2.2 2^k 因子设计及其部分实施

引言

- 2^k 因子试验全面实施能够估计所有效应.
- 当 k 增加时, 全面实施所需的试验次数可能超出实验者所拥有的资源.
- 例如, $2^6 = 64$, 共 63 个自由度:
  - 仅有 6 个与主效应对应;
  - 仅有 15 个与二因子交互效应对应;
  - 其余 42 个自由度与三阶或更高阶交互效应对应.
- 如何采用 2^k 设计的一个试验点数较少的部分实施来估计感兴趣的效应?

引言

- 2^k 因子试验全面实施能够估计所有效应.
- 当 k 增加时, 全面实施所需的试验次数可能超出实验者所拥有的资源.
- 例如, $2^6 = 64$, 共 63 个自由度:
  - 仅有 6 个与主效应对应;
  - 仅有 15 个与二因子交互效应对应;
  - 其余 42 个自由度与三阶或更高阶交互效应对应.
- 如何采用 2^k 设计的一个试验点数较少的部分实施来估计感兴趣的效应?

2.2.4 2^k 因子试验的部分实施

(1) 混杂与定义关系

(2) 2^{k-p} 部分因子设计的构造

(3) 字长型与分辨度

例

2^3 设计包括 8 个处理, 如果只做 2^3 设计的一半

列　名	A	B	AB
因　子	A	B	C

称它为 2^3 设计的 $1/2$ 实施, 也称为 2^{3-1} 设计.

- 如果 AB 显著, 则无法区分 AB 和 C 的效应, 产生了混杂(confounding).
- 称 AB 与 C 互为别名(alias), 用 $C = AB$ 或 $AB = C$ 表示别名关系.

例

2^3 设计包括 8 个处理, 如果只做 2^3 设计的一半

列　名	A	B	AB
因　子	A	B	C

称它为 2^3 设计的 $1/2$ 实施, 也称为 2^{3-1} 设计.

- 如果 AB 显著, 则无法区分 AB 和 C 的效应, 产生了混杂(confounding).
- 称 AB 与 C 互为别名(alias), 用 $C = AB$ 或 $AB = C$ 表示别名关系.

- 别名关系 $C = AB$ 的两端同时乘 C 得到

$$ABC = C^2 = C^0 = \mathbf{I}.$$

$\mathbf{I}$ 称为单位元, 表示全部由 “$+$” 组成的列, 或全部由 “0” 组成的列.

- 由 $ABC = \mathbf{I}$ 可得到别名关系: $BC = A$, $AC = B$, $AB = C$.

- $ABC = \mathbf{I}$ 表达了该方案的全部别名关系, 称它为这个 2^{3-1} 设计的定义关系(defining relations).

- 别名关系 $C = AB$ 的两端同时乘 C 得到

$$ABC = C^2 = C^0 = \mathbf{I}.$$

$\mathbf{I}$ 称为单位元, 表示全部由 “$+$” 组成的列, 或全部由 “0” 组成的列.

- 由 $ABC = \mathbf{I}$ 可得到别名关系: $BC = A$, $AC = B$, $AB = C$.
- $ABC = \mathbf{I}$ 表达了该方案的全部别名关系, 称它为这个 2^{3-1} 设计的定义关系(defining relations).

- 别名关系 $C = AB$ 的两端同时乘 C 得到

$$ABC = C^2 = C^0 = \mathbf{I}.$$

$\mathbf{I}$ 称为单位元, 表示全部由 “$+$” 组成的列, 或全部由 “0” 组成的列.

- 由 $ABC = \mathbf{I}$ 可得到别名关系: $BC = A$, $AC = B$, $AB = C$.

- $ABC = \mathbf{I}$ 表达了该方案的全部别名关系, 称它为这个 2^{3-1} 设计的定义关系(defining relations).

- 定义关系为 $\mathbf{I} = ABC$ 的 2^{3-1} 设计是正交表 $L_8(2^7)$ 中 ABC 列为 $+$ 号对应的四行

处理	$\mathbf{I}$	A	B	AB	C	AC	BC	ABC
c	$+$	$-$	$-$	$+$	$+$	$-$	$-$	$+$
b	$+$	$-$	$+$	$-$	$-$	$+$	$-$	$+$
a	$+$	$+$	$-$	$-$	$-$	$-$	$+$	$+$
abc	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
(1)	$+$	$-$	$-$	$+$	$-$	$+$	$+$	$-$
bc	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$
ac	$+$	$+$	$-$	$-$	$+$	$+$	$-$	$-$
ab	$+$	$+$	$+$	$+$	$-$	$-$	$-$	$-$

- 下半部分构成一个定义关系为 $\mathbf{I} = -ABC$ 的 2^{3-1} 设计. 这两个设计构成互补关系, 即上一个 2^{3-1} 设计的 A 列估计出列的效应是 $A + BC$, 而下一个 2^{3-1} 设计的 A 列估计出来的效应则是 $A - BC$.

- 一般地,
  - 采用正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 是 2^k 因子试验的 $1/2$ 实施;
  - 采用正交表 $L_{2^{k-2}}(2^{2^{k-2}-1})$ 是 2^k 因子试验的 $1/4$ 实施;
  - $\cdots$
  - 采用正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 是 2^k 因子试验的 $1/2^p$ 实施.
- 部分实施是由其定义关系确定的. 讨论部分实施时, 应指出它的定义关系.
- 试验次数压缩得越多, 定义关系就越长, 效应混杂越多.

- 一般地,
  - 采用正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 是 2^k 因子试验的 $1/2$ 实施;
  - 采用正交表 $L_{2^{k-2}}(2^{2^{k-2}-1})$ 是 2^k 因子试验的 $1/4$ 实施;
  - $\cdots$
  - 采用正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 是 2^k 因子试验的 $1/2^p$ 实施.
- 部分实施是由其定义关系确定的. 讨论部分实施时, 应指出它的定义关系.
- 试验次数压缩得越多, 定义关系就越长, 效应混杂越多.

例

$L_8(2^7)$ 可用来安排 2^4 设计的 $1/2$ 实施、2^5 设计的 $1/4$ 实施、2^6 设计的 $1/8$ 实施和 2^7 设计的 $1/16$ 实施:

因子数量	A	B	AB	C	AC	BC	ABC
4	A	B		C			D
5	A	B	E	C			D
6	A	B	E	C	F		D
7	A	B	E	C	F	G	D

如何求定义关系?

- 可直接观察, 也通过观察设计矩阵.
- 如 2^{4-1} 的定义关系为 $\mathbf{I} = ABCD$, 两边分别乘以 A, B, C, D, AB, AC, AD, 便得到该设计的一切别名关系.

例

$L_8(2^7)$ 可用来安排 2^4 设计的 $1/2$ 实施、2^5 设计的 $1/4$ 实施、2^6 设计的 $1/8$ 实施和 2^7 设计的 $1/16$ 实施:

因子数量	A	B	AB	C	AC	BC	ABC
4	A	B		C			D
5	A	B	E	C			D
6	A	B	E	C	F		D
7	A	B	E	C	F	G	D

如何求定义关系?

- 可直接观察, 也通过观察设计矩阵.
- 如 2^{4-1} 的定义关系为 $\mathbf{I} = ABCD$, 两边分别乘以 A, B, C, D, AB, AC, AD, 便得到该设计的一切别名关系.

- 2^{5-2} 设计的设计矩阵为

$$\boldsymbol{D}_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

- 根据模 2 加法运算, 观察得到其定义关系为

$$\mathbf{I} = ABE = CDE = ABCD.$$

定义

部分因子设计中, 称列名记号 A、B 等为字母(letter), 称字母串 $ABCD$、AD 等为字(word), 一个字所含字母的个数称为这个字的字长(word length). 若一个字经列名运算化简后得到单位元 $\mathbf{I}$, 则称这个字为生成字(generator).

- 任意一个 2^{k-p} 设计的所有生成字和单位元 $\mathbf{I}$ 组成的集合在列名运算意义下构成一个群(group).

定义

设集合 G 为非空集, 如果 G 上定义了一个二元运算 “$*$”, 满足

1. 若 $a \in G$, $b \in G$, 则 $a * b \in G$;
2. 对任意 $a, b, c \in G$, $a * (b * c) = (a * b) * c$;
3. G 中有一单位元 e, 使得对任意 $b \in G$, $b * e = e * b = b$;
4. 对任意 $a \in G$, 存在逆元 a^{-1}, 使得 $a * a^{-1} = e$;

则称 G 为一个群 (group), 如果 $a * b = b * a$. 则称 G 为可交换群, 或阿贝尔群 (Abel group).

例

$L_8(2^7)$ 可用来安排 2^4 设计的 $1/2$ 实施、2^5 设计的 $1/4$ 实施、2^6 设计的 $1/8$ 实施和 2^7 设计的 $1/16$ 实施:

因子数量	A	B	AB	C	AC	BC	ABC
4	A	B		C			D
5	A	B	E	C			D
6	A	B	E	C	F		D
7	A	B	E	C	F	G	D

- $\boldsymbol{D}_5$ 的定义关系子群为 $\{\mathbf{I}, ABE, CDE, ABCD\}$;
- $\boldsymbol{D}_6$ 的定义关系子群

$$\{\mathbf{I}, ABE, CDE, ACF, BDF, ABCD, BCEF, ADEF\}.$$

2.2.4 2^k 因子试验的部分实施

(1) 混杂与定义关系

(2) 2^{k-p} 部分因子设计的构造

(3) 字长型与分辨度

1. 定义关系为 $\mathbf{I} = AB\cdots K$ 的 2^{k-1} 设计的构造:

   Step 1 以正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 的基本列安排前 $k-1$ 个试验因子;

   Step 2 以正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 中 $k-1$ 因子交互效应列作为第 k 个试验因子所在的列.

2. 一般地, 2^{k-p} 设计的构造方法为:

   Step 1 以正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 的基本列安排前 $k-p$ 个试验因子;

   Step 2 从正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 的交互效应列中, 选择 p 个合适的高阶交互效应列安排剩余的 p 个因子.

3. 2^{k-p} 设计设计有 p 个基本生成字.

1. 定义关系为 $\mathbf{I}=AB\cdots K$ 的 2^{k-1} 设计的构造:
   - Step 1 以正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 的基本列安排前 $k-1$ 个试验因子;
   - Step 2 以正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 中 $k-1$ 因子交互效应列作为第 k 个试验因子所在的列.
2. 一般地, 2^{k-p} 设计的构造方法为:
   - Step 1 以正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 的基本列安排前 $k-p$ 个试验因子;
   - Step 2 从正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 的交互效应列中, 选择 p 个合适的高阶交互效应列安排剩余的 p 个因子.
3. 2^{k-p} 设计设计有 p 个基本生成字.

1. 定义关系为 $\mathbf{I} = AB\cdots K$ 的 2^{k-1} 设计的构造:
   - Step 1 以正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 的基本列安排前 $k-1$ 个试验因子;
   - Step 2 以正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 中 $k-1$ 因子交互效应列作为第 k 个试验因子所在的列.
2. 一般地, 2^{k-p} 设计的构造方法为:
   - Step 1 以正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 的基本列安排前 $k-p$ 个试验因子;
   - Step 2 从正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 的交互效应列中, 选择 p 个合适的高阶交互效应列安排剩余的 p 个因子.
3. 2^{k-p} 设计设计有 p 个基本生成字.

1. 定义关系为 $\mathbf{I}=AB\cdots K$ 的 2^{k-1} 设计的构造:

   Step 1 以正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 的基本列安排前 $k-1$ 个试验因子;

   Step 2 以正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 中 $k-1$ 因子交互效应列作为第 k 个试验因子所在的列.

2. 一般地, 2^{k-p} 设计的构造方法为:

   Step 1 以正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 的基本列安排前 $k-p$ 个试验因子;

   Step 2 从正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 的交互效应列中, 选择 p 个合适的高阶交互效应列安排剩余的 p 个因子.

3. 2^{k-p} 设计设计有 p 个基本生成字.

1. 定义关系为 $\mathbf{I} = AB\cdots K$ 的 2^{k-1} 设计的构造:

   Step 1 以正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 的基本列安排前 $k-1$ 个试验因子;

   Step 2 以正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 中 $k-1$ 因子交互效应列作为第 k 个试验因子所在的列.

2. 一般地, 2^{k-p} 设计的构造方法为:

   Step 1 以正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 的基本列安排前 $k-p$ 个试验因子;

   Step 2 从正交表 $L_{2^{k-p}}(2^{2^{k-p}-1})$ 的交互效应列中, 选择 p 个合适的高阶交互效应列安排剩余的 p 个因子.

3. 2^{k-p} 设计设计有 p 个基本生成字.

2.2.4 2^k 因子试验的部分实施

(1) 混杂与定义关系

(2) 2^{k-p} 部分因子设计的构造

(3) 字长型与分辨度

- 2^{k-p} 设计不唯一，挑选合适方案的准则是利用最小的正交表估计出感兴趣的效应.
- 在不考虑先验信息的情况下，如何比较两个不同的 2^{k-p} 设计?
- 分辨度是一个常用的衡量部分实施方案性能的指标.

- 2^{k-p} 设计不唯一, 挑选合适方案的准则是利用最小的正交表估计出感兴趣的效应.
- 在不考虑先验信息的情况下, 如何比较两个不同的 2^{k-p} 设计?
- 分辨度是一个常用的衡量部分实施方案性能的指标.

- 2^{k-p} 设计不唯一, 挑选合适方案的准则是利用最小的正交表估计出感兴趣的效应.
- 在不考虑先验信息的情况下, 如何比较两个不同的 2^{k-p} 设计?
- 分辨度是一个常用的衡量部分实施方案性能的指标.

定义

用 $A_i(\boldsymbol{D})$ 表示设计 $\boldsymbol{D}$ 的生成字中字长为 i 的字的个数, 称

$$W(\boldsymbol{D}) = \{A_1(\boldsymbol{D}), A_2(\boldsymbol{D}), \cdots, A_k(\boldsymbol{D})\}$$

为设计 $\boldsymbol{D}$ 的字长型(word length pattern). 在一个设计的字长型中, 所有生成字的最小字长为这个设计的分辨度(resolution).

- 若当 $i < t$ 时 $A_i(\boldsymbol{D}) = 0$, 而 $A_t(\boldsymbol{D}) > 0$, 则 t 为设计 $\boldsymbol{D}$ 的分辨度.
- 如果大于 t 阶的效应不存在, 则分辨度为 $2t+1$ 的设计中任何不超过 t 阶的效应都是可估计的.

定义

用 $A_i(\boldsymbol{D})$ 表示设计 $\boldsymbol{D}$ 的生成字中字长为 i 的字的个数, 称

$$W(\boldsymbol{D}) = \{A_1(\boldsymbol{D}), A_2(\boldsymbol{D}), \cdots, A_k(\boldsymbol{D})\}$$

为设计 $\boldsymbol{D}$ 的字长型(word length pattern). 在一个设计的字长型中, 所有生成字的最小字长为这个设计的分辨度(resolution).

- 若当 $i < t$ 时 $A_i(\boldsymbol{D}) = 0$, 而 $A_t(\boldsymbol{D}) > 0$, 则 t 为设计 $\boldsymbol{D}$ 的分辨度.
- 如果大于 t 阶的效应不存在, 则分辨度为 $2t+1$ 的设计中任何不超过 t 阶的效应都是可估计的.

定义

用 $A_i(\boldsymbol{D})$ 表示设计 $\boldsymbol{D}$ 的生成字中字长为 i 的字的个数, 称

$$W(\boldsymbol{D}) = \{A_1(\boldsymbol{D}), A_2(\boldsymbol{D}), \cdots, A_k(\boldsymbol{D})\}$$

为设计 $\boldsymbol{D}$ 的字长型(word length pattern). 在一个设计的字长型中, 所有生成字的最小字长为这个设计的分辨度(resolution).

- 若当 $i < t$ 时 $A_i(\boldsymbol{D}) = 0$, 而 $A_t(\boldsymbol{D}) > 0$, 则 t 为设计 $\boldsymbol{D}$ 的分辨度.
- 如果大于 t 阶的效应不存在, 则分辨度为 $2t + 1$ 的设计中任何不超过 t 阶的效应都是可估计的.

用大写罗马数字表示分辨度:

分辨度 III: 如果二阶和二阶以上交互效应可以忽略, 则主效应之间没有混杂, 但至少有一个主效应与某个二阶交互效应混杂;

分辨度 IV: 如果三阶和三阶以上交互效应可以忽略, 则主效应之间、主效应和二阶交互效应之间没有混杂, 但至少有一个主效应与某个三阶交互效应混杂;

分辨度 V: 如果三阶和三阶以上交互效应可以忽略, 则主效应之间、主效应和二阶交互效应之间, 以及任意两对二阶交互效应之间没有混杂.

用大写罗马数字表示分辨度:

分辨度 III: 如果二阶和二阶以上交互效应可以忽略, 则主效应之间没有混杂, 但至少有一个主效应与某个二阶交互效应混杂;

分辨度 IV: 如果三阶和三阶以上交互效应可以忽略, 则主效应之间、主效应和二阶交互效应之间没有混杂, 但至少有一个主效应与某个三阶交互效应混杂;

分辨度 V: 如果三阶和三阶以上交互效应可以忽略, 则主效应之间、主效应和二阶交互效应之间, 以及任意两对二阶交互效应之间没有混杂.

用大写罗马数字表示分辨度:

分辨度 III: 如果二阶和二阶以上交互效应可以忽略, 则主效应之间没有混杂, 但至少有一个主效应与某个二阶交互效应混杂;

分辨度 IV: 如果三阶和三阶以上交互效应可以忽略, 则主效应之间、主效应和二阶交互效应之间没有混杂, 但至少有一个主效应与某个三阶交互效应混杂;

分辨度 V: 如果三阶和三阶以上交互效应可以忽略, 则主效应之间、主效应和二阶交互效应之间, 以及任意两对二阶交互效应之间没有混杂.

定义

- 称一个 2^{k-p} 设计 $\boldsymbol{D}$ 有最大分辨度(maximum resolution), 如果不存在分辨度更高的 2^{k-p} 设计.
- 设 $\boldsymbol{D}_1$ 和 $\boldsymbol{D}_2$ 是两个 2^{k-p} 设计, 如果存在整数 r, 使得
  1. 当 $1 \leq i < r$ 时, $A_i(\boldsymbol{D}_1) = A_i(\boldsymbol{D}_2)$,
  2. 且 $A_r(\boldsymbol{D}_1) < A_r(\boldsymbol{D}_2)$,

  则称 $\boldsymbol{D}_1$ 比 $\boldsymbol{D}_2$ 有较小的低阶混杂. 如果不存在比 $\boldsymbol{D}_1$ 有更小低阶混杂的 2^{k-p} 设计, 则称 $\boldsymbol{D}_1$ 为最小低阶混杂(minimum aberration) 设计.

小测试

例

$L_8(2^7)$ 可用来安排 2^4 设计的 $1/2$ 实施、2^5 设计的 $1/4$ 实施、2^6 设计的 $1/8$ 实施和 2^7 设计的 $1/16$ 实施:

因子数量	A	B	AB	C	AC	BC	ABC
4	A	B		C			D
5	A	B	E	C			D
6	A	B	E	C	F		D
7	A	B	E	C	F	G	D

- 这四个设计的定义关系子群、字长型、分辨度分别是什么?
- 能否构造出比他们具有更高分辨度或更小低阶混杂的设计?

总结

- 2^k 完全因子设计
  - 主效应、交互效应
  - 正交表 $L_{2^k}(2^{2^k-1})$ 的构造与应用
- 2^{k-p} 部分因子设计
  - 混杂与效应别名
  - 定义关系与定义关系子群
  - 2^{k-p} 设计的构造
  - 部分因子设计的比较：字长型、分辨度、最小低阶混杂
- 理解试验资源与获取信息之间的均衡

习题

1. 构造正交表 $L_{16}(2^{15})$.
2. 写出正交表 $L_{32}(2^{31})$ 的 $ABCD$ 列.
3. 给出一种 2^{5-2} 部分实施方案, 并给出它的定义关系、所有别名关系、字长型以及分辨度.
4. R 添加包 `FrF2` 的函数 `FrF2()` 可以用于生成 2^k 设计的部分实施方案, 函数 `aliases()` 可以给出该方案的全部别名关系. 尝试使用这两个函数.