2.2 2^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

- 什么是因子设计?
- 什么是对照?
- 讲过哪些统计思想?
- 什么是可辨识?

引言

- 什么是因子设计?
- 什么是对照?
- 讲过哪些统计思想?
- 什么是可辨识?

引言

- 什么是因子设计?
- 什么是对照?
- 讲过哪些统计思想?
- 什么是可辨识?

引言

- 什么是因子设计?
- 什么是对照?
- 讲过哪些统计思想?
- 什么是可辨识?

引言

- 包含 k 个二水平因子的试验, 全部处理有 2^k 个;
- 主要应用于:
  - 定性考察因子对响应的影响;
  - 筛选大量因子中有实质影响的因子的初级研究阶段.
- 优势: 试验次数可以控制在较少的范围内.
- 缺陷: 对于连续变化的定量因子, 不能归纳非线性关系.

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 $L_4(2^3)$

(1) 固定效应模型

(2) 主效应与交互效应

(3) 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 $L_{2^k}(2^{2^k-1})$

2.2.4 2^k 因子试验的部分实施

(1) 2^2 因子设计的固定效应模型

- 两个二水平因子 A 和 B, 以 0、1 表示两个水平;
- 用 (1)、a、b、ab 表示 4 个处理 $(0,0)$, $(1,0)$, $(0,1)$, $(1,1)$, 当代表因子的字母出现时, 该因子水平取 1, 否则取 0;
- 设每个处理重复 m 次, 以 y_{ijk} 表示处理 (i,j) 处第 k 次重复试验的响应值, 则 2^2 试验的固定效应模型为:

$$\begin{cases} y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}, \\ \varepsilon_{ijk} \overset{\text{i.i.d.}}{\sim} N(0,\sigma^2), \quad i,j = 0,1, \quad k = 1,\cdots,m, \\ \tau_0 + \tau_1 = 0, \quad \beta_0 + \beta_1 = 0, \\ (\tau\beta)_{00} + (\tau\beta)_{01} = (\tau\beta)_{10} + (\tau\beta)_{11} = 0, \\ (\tau\beta)_{00} + (\tau\beta)_{10} = (\tau\beta)_{01} + (\tau\beta)_{11} = 0. \end{cases}$$

(1) 2^2 因子设计的固定效应模型

- 两个二水平因子 A 和 B, 以 0、1 表示两个水平;
- 用 (1)、a、b、ab 表示 4 个处理 $(0,0)$, $(1,0)$, $(0,1)$, $(1,1)$, 当代表因子的字母出现时, 该因子水平取 1, 否则取 0;
- 设每个处理重复 m 次, 以 y_{ijk} 表示处理 (i,j) 处第 k 次重复试验的响应值, 则 2^2 试验的固定效应模型为:

$$
\begin{cases}
y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}, \\
\varepsilon_{ijk} \overset{\text{i.i.d.}}{\sim} N(0,\sigma^2), \quad i,j = 0,1, \quad k = 1,\cdots,m, \\
\tau_0 + \tau_1 = 0, \quad \beta_0 + \beta_1 = 0, \\
(\tau\beta)_{00} + (\tau\beta)_{01} = (\tau\beta)_{10} + (\tau\beta)_{11} = 0, \\
(\tau\beta)_{00} + (\tau\beta)_{10} = (\tau\beta)_{01} + (\tau\beta)_{11} = 0.
\end{cases}
$$

(2) 主效应与交互效应

Definition

- 称 $\tau_1 - \tau_0$ 为因子 A 的主效应, 记作 τ, 其估计记作 A, 称 $\beta_1 - \beta_0$ 为因子 B 的主效应, 记作 β, 其估计记作 B;
- 称 $\frac{1}{2}[(\tau\beta)_{11} - (\tau\beta)_{01} - (\tau\beta)_{10} + (\tau\beta)_{00}]$ 为因子 A 与因子 B 的交互效应, 记作 $\tau\beta$, 其估计记作 AB.

(2) 主效应与交互效应

- 以符号 (1)、a、b、ab 表示各处理上 m 个响应值的总和, 则各效应的估计为

$$\begin{cases} A = [ab + a - b - (1)]/(2m); \\ B = [ab - a + b - (1)]/(2m); \\ AB = [ab - a - b + (1)]/(2m). \end{cases}$$

- 它们是互相正交的三个对照, 它们的平方和分别为

$$\begin{cases} SS_A = \dfrac{1}{4m}[ab + a - b - (1)]^2; \\ SS_B = \dfrac{1}{4m}[ab - a + b - (1)]^2; \\ SS_{AB} = \dfrac{1}{4m}[ab - a - b + (1)]^2. \end{cases}$$

(3) 正交表 $L_4(2^3)$

- 将计算诸效应估计量的对比的系数符号列成表:

	A	B	AB
(1)	$-$	$-$	$+$
b	$-$	$+$	$-$
a	$+$	$-$	$-$
ab	$+$	$+$	$+$

- 第一列为二分列, 称第二列为四分列. 此外:

  (1) 每列 “$+$” 号与 “$-$” 号的出现的次数相等;

  (2) 任何两列组成四组不同的符号对, 其出现的次数相等.

(3) 正交表 $L_4(2^3)$

Definition

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的重复次数相等.

性质:

- 任意一列中不同符号出现的次数相等;
- 从一张正交表中挑选出部分列组成的子表依然是正交表.

(3) 正交表 $L_4(2^3)$

- $L_4(2^3)$：L 表示正交表，2 代表正交表中不同水平数，4 代表表的行数，3 代表表的列数.

	A	B	AB
(1)	$-$	$-$	$+$
b	$-$	$+$	$-$
a	$+$	$-$	$-$
ab	$+$	$+$	$+$

- 任意两列对应符号相乘得出另一列，且任意一列均可由其余两列对应符号相乘得到，称这一性质为 $L_4(2^3)$ 的任何两列的交互作用列是另一列.

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 $L_8(2^7)$

(1) 固定效应模型

(2) 主效应与交互效应

(3) 正交表 $L_8(2^7)$

(4) 正交表的等价

2.2.3 2^k 设计与正交表 $L_{2^k}(2^{2^k-1})$

2.2.4 2^k 因子试验的部分实施

(1) 2^3 设计的固定效应模型

- 设 3 个因子为 A、B、C, 以 “1” 和 “0” 分别表示因子的两个水平.
- 三个主效应 A、B、C, 三个二阶交互效应 AB、BC、AC, 以及一个三阶交互效应 ABC.
- 全部处理组合共 8 个:

$$(0,0,0),(0,0,1),(0,1,0),(0,1,1),$$

$$(1,0,0),(1,0,1),(1,1,0),(1,1,1).$$

(1) 2^3 设计的固定效应模型

$$
\begin{cases}
y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \varepsilon_{ijkl};\\
\varepsilon_{ijkl} \overset{\text{i.i.d.}}{\sim} N(0,\sigma^2), \quad i=0,1, \quad j=0,1, \quad k=0,1, \quad l=1,\cdots,m;\\
\tau_0+\tau_1=\beta_0+\beta_1=\gamma_0+\gamma_1=0;\\
(\tau\beta)_{00}+(\tau\beta)_{10}=(\tau\beta)_{01}+(\tau\beta)_{11}=(\tau\beta)_{00}+(\tau\beta)_{01}=(\tau\beta)_{10}+(\tau\beta)_{11}=0;\\
(\tau\gamma)_{00}+(\tau\gamma)_{10}=(\tau\gamma)_{01}+(\tau\gamma)_{11}=(\tau\gamma)_{00}+(\tau\gamma)_{01}=(\tau\gamma)_{10}+(\tau\gamma)_{11}=0;\\
(\beta\gamma)_{00}+(\beta\gamma)_{10}=(\beta\gamma)_{01}+(\beta\gamma)_{11}=(\beta\gamma)_{00}+(\beta\gamma)_{01}=(\beta\gamma)_{10}+(\beta\gamma)_{11}=0;\\
(\tau\beta\gamma)_{0jk}+(\tau\beta\gamma)_{1jk}=0, \quad j=0,1, \quad k=0,1;\\
(\tau\beta\gamma)_{i0k}+(\tau\beta\gamma)_{i1k}=0, \quad i=0,1, \quad k=0,1;\\
(\tau\beta\gamma)_{ij0}+(\tau\beta\gamma)_{ij1}=0, \quad i=0,1, \quad j=0,1.
\end{cases}
$$

(1) 2^3 设计的固定效应模型

$$
\begin{cases}
y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \varepsilon_{ijkl}; \\
\varepsilon_{ijkl} \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2), \quad i = 0, 1, \quad j = 0, 1, \quad k = 0, 1, \quad l = 1, \cdots, m; \\
\tau_0 + \tau_1 = \beta_0 + \beta_1 = \gamma_0 + \gamma_1 = 0; \\
(\tau\beta)_{00} + (\tau\beta)_{10} = (\tau\beta)_{01} + (\tau\beta)_{11} = (\tau\beta)_{00} + (\tau\beta)_{01} = (\tau\beta)_{10} + (\tau\beta)_{11} = 0; \\
(\tau\gamma)_{00} + (\tau\gamma)_{10} = (\tau\gamma)_{01} + (\tau\gamma)_{11} = (\tau\gamma)_{00} + (\tau\gamma)_{01} = (\tau\gamma)_{10} + (\tau\gamma)_{11} = 0; \\
(\beta\gamma)_{00} + (\beta\gamma)_{10} = (\beta\gamma)_{01} + (\beta\gamma)_{11} = (\beta\gamma)_{00} + (\beta\gamma)_{01} = (\beta\gamma)_{10} + (\beta\gamma)_{11} = 0; \\
(\tau\beta\gamma)_{0jk} + (\tau\beta\gamma)_{1jk} = 0, \quad j = 0, 1, \quad k = 0, 1; \\
(\tau\beta\gamma)_{i0k} + (\tau\beta\gamma)_{i1k} = 0, \quad i = 0, 1, \quad k = 0, 1; \\
(\tau\beta\gamma)_{ij0} + (\tau\beta\gamma)_{ij1} = 0, \quad i = 0, 1, \quad j = 0, 1.
\end{cases}
$$

(2) 主效应与交互效应

- 称 $\tau_1 - \tau_0$ 为因子 A 的效应, 记作 τ, 其估计量记作 A;
- 称 $\beta_1 - \beta_0$ 为因子 B 的效应, 记作 β, 其估计量记作 B;
- 称 $\gamma_1 - \gamma_0$ 为因子 C 的效应, 记作 γ, 其估计量记作 C;

(2) 主效应与交互效应

- 称 $\frac{1}{2}[(\tau\beta)_{11} - (\tau\beta)_{10} - (\tau\beta)_{01} + (\tau\beta)_{00}]$ 为因子 A 与 B 的交互作用，记作 $\tau\beta$，其估计量记作 AB;
- 称 $\frac{1}{2}[(\tau\gamma)_{11} - (\tau\gamma)_{10} - (\tau\gamma)_{01} + (\tau\gamma)_{00}]$ 为因子 A 与 C 的交互作用，记作 $\tau\gamma$，其估计量记作 AC;
- 称 $\frac{1}{2}[(\beta\gamma)_{11} - (\beta\gamma)_{10} - (\beta\gamma)_{01} + (\beta\gamma)_{00}]$ 为因子 B 与 C 的交互作用，记作 $\beta\gamma$，其估计量记作 BC;

(2) 主效应与交互效应

- 称

$$\frac{1}{2}\left\{\frac{1}{2}\left[(\tau\beta\gamma)_{111}-(\tau\beta\gamma)_{011}-(\tau\beta\gamma)_{101}+(\tau\beta\gamma)_{001}\right]\right.$$
$$\left.-\frac{1}{2}\left[(\tau\beta\gamma)_{110}-(\tau\beta\gamma)_{010}-(\tau\beta\gamma)_{100}+(\tau\beta\gamma)_{000}\right]\right\}$$

为 3 个因子 A、B、C 的交互作用，记作 $\tau\beta\gamma$，它的估计量记作 ABC.

- 各试验点上 m 个观察值的总和 $y_{000\cdot}$、$y_{100\cdot}$、$y_{010\cdot}$、$y_{110\cdot}$、$y_{001\cdot}$、$y_{101\cdot}$、$y_{011\cdot}$、$y_{111\cdot}$ 分别用 (1)、a、b、ab、c、ac、bc、abc 表示，
- 则诸效应的无偏估计为

$$\left\{\begin{aligned}
A &= \frac{1}{4m}\left[-(1)-c-b-bc+a+ac+ab+abc\right];\\
B &= \frac{1}{4m}\left[-(1)-c+b+bc-a-ac+ab+abc\right];\\
C &= \frac{1}{4m}\left[-(1)+c-b+bc-a+ac-ab+abc\right];\\
AB &= \frac{1}{4m}\left[+(1)+c-b-bc-a-ac+ab+abc\right];\\
AC &= \frac{1}{4m}\left[+(1)-c+b-bc-a+ac-ab+abc\right];\\
BC &= \frac{1}{4m}\left[+(1)-c-b+bc+a-ac-ab+abc\right];\\
ABC &= \frac{1}{4m}\left[-(1)+c+b-bc+a-ac-ab+abc\right].
\end{aligned}\right.$$

- 它们是七个互相正交的对比!

(3) 正交表 $L_8(2^7)$

- 诸效应的对比系数的符号列表:

处理	A	B	AB	C	AC	BC	ABC
(1)	−	−	+	−	+	+	−
c	−	−	+	+	−	−	+
b	−	+	−	−	+	−	+
bc	−	+	−	+	−	+	−
a	+	−	−	−	−	+	+
ac	+	−	−	+	+	−	−
ab	+	+	+	−	−	−	−
abc	+	+	+	+	+	+	+

- 该表是一个正交表, 以 $L_8(2^7)$ 表示它;
- A 列为二分列, B 列为四分列, C 列为八分列, 这三列即可用来安排试验, 又可用来分析试验结果;

(3) 正交表 $L_8(2^7)$

- 正交性：任何两列符号乘积之和为 0;
- 任意两列相乘，得出表中的一列，如 $A\times B=AB$、$AB\times C=ABC$ 等;
- 可由其它两列运算得到的列称为那两列的交互作用列，$L_8(2^7)$ 中的交互作用关系:

1	2	3	4	5	6	7	列　号
	3	2	5	4	7	6	1
		1	6	7	4	5	2
			7	6	5	4	3
				1	2	3	4
					3	2	5
						1	6

(4) 正交表的等价

从试验设计的角度来看:

- 试验的次序可以自由选择: 正交表的任意两行可以互相置换;
- 因子可以自由地安排在正交表的列上: 正交表的任意两列可以互相置换;
- 因子的水平可以自由安排: 正交表每一列的水平可以互相置换;
- 正交表不唯一!

(4) 正交表的等价

Definition

称两张正交表**等价**, 如果对其中一张表进行适当的行置换和列置换可以得到另一张表; 称两张正交表**同构**, 如果对其中一张表进行适当的行置换、列置换和水平置换可以得到另一张表.

- 以水平记号“1”和“0”代替代替符号“+”和“−”，得到

No.	A	B	AB	C	AC	BC	ABC
1	0	0	1	0	1	1	0
2	0	0	1	1	0	0	1
3	0	1	0	0	1	0	1
4	0	1	0	1	0	1	0
5	1	0	0	0	0	1	1
6	1	0	0	1	1	0	0
7	1	1	1	0	0	0	0
8	1	1	1	1	1	1	1

- 把符号 “+” 和 “−” 互换, 然后分别以 “1” 和 “0” 替换 “+” 和 “−”, 得到

No.	A	B	AB	C	AC	BC	ABC
1	0	0	0	0	0	0	0
2	0	0	0	1	1	1	1
3	0	1	1	0	0	1	1
4	0	1	1	1	1	0	0
5	1	0	1	0	1	0	1
6	1	0	1	1	0	1	0
7	1	1	0	0	1	1	0
8	1	1	0	1	0	0	1

- AB 列由 A 列和 B 列按照模 2 加法生成, 表示交互效应列.

(5) $L_8(2^7)$ 的构造方法

- 利用诸效应的系数构造;
- 利用列名运算:

Step 1 构造二分列 A, 它的前四个元素为 0, 后四个元素为 1;

Step 2 构造四分列 B, 并利用对应元素的模 2 加法运算构造 B 与它前面的列 A 的交互作用列 AB;

Step 3 构造八分列 C, 并利用对应元素的模 2 加法运算依此构造 C 与它前面的 A 列、B 列和 AB 列的交互作用列 AC 列、BC 列和 ABC 列.

Example

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作用.

Example (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

- 为什么因子 C 不能安排在第三列?

试验号		$A(1)$	$B(2)$	$AB(3)$	$C(4)$	$AC(5)$	$BC(6)$	$ABC(7)$	棉结粒数
1	(1)	0	0	0	0	0	0	0	0.30
2	c	0	0	0	1	1	1	1	0.35
3	b	0	1	1	0	0	1	1	0.20
4	bc	0	1	1	1	1	0	0	0.30
5	a	1	0	1	0	1	0	1	0.15
6	ac	1	0	1	1	0	1	0	0.50
7	ab	1	1	0	0	1	1	0	0.15
8	abc	1	1	0	1	0	0	1	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	

- m_0 行表示该列中水平为 0 的行对应的试验结果的平均值, m_1 行表示该列中水平为 1 的行对应的试验结果的平均值.
- $R := \max\{m_i\} - \min\{m_i\}$ 为极差, 可用来衡量 3 个因子和它们交互作用的主次关系, 极差越大表明相应的效应越大.

两种方式计算各列的偏差平方和,

- 采用计算对比偏差平方和的公式:

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

- 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			

- 结果显示只有 C, B 和 AC 显著.

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 $L_{2^k}(2^{2^k-1})$

2.2.4 2^k 因子试验的部分实施

(1) 处理记号

- 从后往前依次引入因子, 每引入一个新的因子, 就依次和前面的因子组合.
- 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) 效应:

- 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, $\cdots$, 1 个 k 因子交互效应, 共 2^k-1 个效应.
- 诸效应自由度均为 1, 共 2^k-1.
- 若每个处理重复 m 次, 则总自由度是 2^km-1, 误差的自由度是 $2^k(m-1)$.

(1) 处理记号

- 从后往前依次引入因子, 每引入一个新的因子, 就依次和前面的因子组合.
- 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) 效应:

- 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, $\cdots$, 1 个 k 因子交互效应, 共 2^k-1 个效应.
- 诸效应自由度均为 1, 共 2^k-1.
- 若每个处理重复 m 次, 则总自由度是 $2^k m-1$, 误差的自由度是 $2^k(m-1)$.

(2) 效应

- 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.

- 2^3 设计中二因子交互效应 AC 的估计为

$$\begin{aligned}AC&=\frac{1}{4m}(a-1)(b+1)(c-1)\\&=\frac{1}{4m}[abc-ab+ac-a-bc+b-c+(1)].\end{aligned}$$

(2) 效应

- 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.

- 2^3 设计中二因子交互效应 AC 的估计为

$$\begin{aligned} AC &= \frac{1}{4m}(a-1)(b+1)(c-1) \\ &= \frac{1}{4m}\left[abc - ab + ac - a - bc + b - c + (1)\right]. \end{aligned}$$

(2) 效应

- 效应的平方和为:

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

(3) 正交表 $L_{2^k}(2^{2^k-1})$

- 用于设计和分析 2^k 因子试验
- 两种构造方式: 利用效应估计的符号和列名运算

(2) 效应

- 效应的平方和为:

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

(3) 正交表 $L_{2^k}(2^{2^k-1})$

- 用于设计和分析 2^k 因子试验
- 两种构造方式: 利用效应估计的符号和列名运算

- 利用列名运算构造正交表 $L_{2^k}(2^{2^k-1})$

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作用列;

⋮

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作用列.

- 试利用列名运算构造正交表 $L_{16}(2^{15})$.

- 根据对照的系数符号构造:
  - 列的次序, 按字母次序, 首先引入 A, 其后每引入一个字母, 就依次引入它与前面各列的交互作用列;
  - 行的次序, 按字母反序, 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.
- 根据正交表等价的定义, 行列次序不重要!
- 试利用对照的系数符号构造正交表 $L_{16}(2^{15})$.

- 利用正交表 $L_{16}(2^{15})$ 的二分列、四分列、八分列等主效应列安排试验
- 诸效应的估计和诸效应的平方和可由相应行来计算. 可使用公式:

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

T 表示所有试验数据之和, T_{i0} 表示第 i 列中水平为 0 对应的数据之和, T_{i1} 表示第 i 列中水平为 1 对应的数据之和.

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 $L_{2^k}(2^{2^k-1})$

2.2.4 2^k 因子试验的部分实施

- 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.

- 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:

- 如果以“$-$”和“$+$”表示两个水平, 且以对应水平的乘法运算表示列与列之间运算时, $\mathbf{I}$ 表示全部由“$+$”组成的列;
- 如果以“0”和“1”表示两个水平, 且以对应水平的模 2 加法运算表示列与列之间的运算时, $\mathbf{I}$ 表示全部由“0”组成的列.

- 由 $ABC = \mathbf{I}$ 可得到别名关系

$$BC = A, \quad AC = B, \quad AB = C.$$

- $ABC = \mathbf{I}$ 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).
- 部分实施是由其定义关系确定的. 讨论部分实施时, 都应指出它的定义关系.

- 当 k 的数值较大时, 通常采用 2^k 设计的部分实施.
- 采用正交表 $L_{2^{k-1}}(2^{2^{k-1}-1})$ 是 2^k 设计的 $1/2$ 实施, 采用正交表 $L_{2^{k-2}}(2^{2^{k-2}-1})$ 是 2^k 设计的 $1/4$ 实施, $\cdots$.
- 2^k 设计部分实施的方案不唯一, 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.

Example

$L_8(2^7)$ 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:

	A	B	AB	C	AC	BC	ABC	定义关系
4	A	B		C			D	$\mathbf{I}=ABCD$
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?

- 2^{4-1} 设计的设计矩阵为

$$\boldsymbol{D}_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

- 根据模 2 加法运算，观察得到其定义关系为

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

- 两边分别乘以 A, B, C, D, AB, AC, AD, 便该设计的得到一切别名关系: $A = BCD,\ B = ACD,\ D = ABC,\ AB = CD,\ AC = BD,\ BC = AD.$

- 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

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

- 由此也可以得到这个 2^{5-2} 的一切别名关系, 作为练习.

- 试验次数压缩得越多，定义关系就越长，别名关系也就越多.
- 一般地，2^{k-p} 设计不唯一，在不考虑实际问题背景的情况下，如何比较两个不同的 2^{k-p} 设计呢？
- 分辨度是一个常用的衡量部分实施方案性能的指标.

- 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).

Definition

设集合 G 为非空集, 如果集合 G 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, 如果 $a * b = b * a$. 则称 G 为可交换群, 或阿贝尔群.

- D_5 的生成字和单位元共同组成的集合

$$\{\mathbf{I}, ABE, CDE, ABCD\}$$

构成一个可交换群, 称这个群为 D_5 的定义关系子群;

- D_6 的定义关系子群

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

- 称一个设计的所有生成字的最小字长为这个设计的分辨度, 用大写罗马数字来表示.
- D_4 的是分辨度 IV 设计, D_5 是分辨度 III 设计, D_6 也是分辨度 III 设计.

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作用混杂;

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作用混杂;

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.

- 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.

总结

- 2^2 设计与正交表 $L_4(2^3)$;
- 2^3 设计与正交表 $L_8(2^7)$;
- 2^k 设计与正交表 $L_{2^k}(2^{2^k-1})$;
- 主效应, 交互效应;
- 定义关系, 混杂, 效应别名, 字长, 分辨度.

习题

1. 写出 2^2 因子设计和 2^3 因子设计的方差分析表.
2. 构造正交表 $L_{16}(2^{15})$.
3. 写出正交表 $L_{32}(2^{31})$ 的 $ABCD$ 列.
4. 给出一种 2^{5-2} 部分实施方案, 被给出它的定义关系和所有别名关系.

注意: 各种教材上符号有所不同!

请提问