

正规部分因析设计别名成分数型的一些性质 *

李 智 李智明*

(新疆大学数学与系统科学学院, 乌鲁木齐, 830046)

摘要: 一般最小低阶混杂准则和最小混杂准则是挑选 s ($s \geq 2$) 水平最优正规部分因析设计的两个重要准则, 其分类模式分别为别名成分数型与字长型. 本文主要研究 s 水平正规设计的别名成分数型的一些性质, 得到了 s 水平下字长型中的元素可以表示为别名成分数型的函数形式, 揭示了别名成分数型与字长型之间的关系. 另一方面, 通过字长型也可以计算某些别名成分数型. 进一步, 得到了二水平设计的一些别名成分数型之间的表达公式.

关键词: 正规设计; 定义关系; 字长型; 最小混杂准则; 别名成分数型; 一般最小低级混杂准则

中图分类号: O212.6

英文引用格式: LI Z, LI Z M. Some properties of aliased component-number pattern for regular fractional factorial designs [J]. Chinese J Appl Probab Statist, 2021, 37(6): 585–597. (in Chinese)

§1. 引 言

试验设计是统计学科的重要分支之一, 它在农业、医药、化工以及生物等领域具有广泛的应用. 自 Fisher 创立试验设计以来, 为了选择最优设计, 相继产生各种最优准则. 例如: 最大分辨度准则、最小混杂准则、以及最大估计容量准则等. Box 和 Hunter^[1] 提出最大分辨度 (maximum resolution, MR) 准则, 并认为分辨度最大的设计为最优的. 然而, 当设计的分辨度相等时, MR 准则将无法判断设计的优劣. 假定低阶因子效应比高阶因子效应更重要; 同阶因子效应同等重要, 被称为效应排序原则. 基于效应排序原则, Fries 和 Hunter^[2] 提出最小混杂 (minimum aberration, MA) 准则, 它是在 MR 准则的基础上发展而来, 通过定义关系中的字长来选择最优设计. 在 MA 准则中, A_i ($i = 1, 2, \dots, n$) 表示为在 s^{n-m} 设计 D 中字长为 i 的定义字的个数. 按照字长的下标 i 从小到大进行排列, 定义 $W(D) = (A_1, A_2, \dots, A_n)$ 为设计 D 的字长型 (word-length pattern, WLP). 顺序最小化向量 $W(D)$ 的设计为最小混杂设计, 即 MA 设计. 分辨度 R 定义为字长型 $W(D)$ 中第一个非零 A_i 的下标 i . 因此, MR 和 MA 准则均以字长型为基础. 为了能够估计尽可能多的包含所有主效应和部分二因子交互效应的模型, Sun^[3] 提出了最大估计容量 (maximum

*国家自然科学基金项目 (批准号: 12061070、11661076) 和新疆维吾尔自治区自然科学基金项目 (批准号: 2021D01E13、2018Q011) 资助.

*通讯作者, E-mail: zml@xju.edu.cn.

本文 2020 年 5 月 7 日收到, 2020 年 7 月 27 日收到修改稿.

estimation capacity, MEC) 准则. MEC 准则似乎与字长型没有直接的联系, 然而 Hu 和 Zhang^[4] 指出如果 MEC 设计存在且仅关注主效应和二因子交互效应的条件下, MEC 设计就等价于 MA 设计. 在这种情况下可以认为 MEC 准则也是建立在字长型上. 因此, 对于上述准则, 为了获得最优设计研究其字长型至关重要. 更多的内容可参阅文献 [5, 6].

为了揭示上述准则的内在本质, Zhang 等^[7] 针对二水平设计在效应排序原则的基础上提出别名效应数型, 得到以上准则的分类模式均是别名效应数型的某种函数. 他们根据别名效应数型, 提出一般最小低阶混杂准则来选择最优设计. Zhang 和 Cheng^[8] 通过二水平别名效应数型得到了字长型中部分元素 A_i ($i = 3, 4, 5$) 的计算公式, Hu 和 Zhang^[4] 也验证了该结论. 对分辨度为 IV 的二水平设计, Chen 和 Liu^[9] 得到了 A_i ($i = 4, 5, 6$) 与别名效应数之间的表达形式. 这些结论揭示了 MA 准则中的字长型 A_i ($i = 3, 4, 5, 6$) 本质上是别名效应数型中某些元素的平均, 这说明 MA 设计具平均意义上的最小低阶混杂性质. 然而, 对 A_i ($i > 6$) 是否也是别名效应数型某些元素的平均仍需要进一步研究.

Zhang 和 Mukerjee^[10] 把该准则推广到 s ($s \geq 2$) 水平情形, 其中 s 为素数或素数幂. 对二水平设计, 因子与成分之间是一一对应的关系, 但三水平以上的因子交互效应对应于多个正交成分, 因此需要通过因子成分效应的别名关系来揭示 s 水平设计的一般混杂信息. 假设低阶成分效应比高阶成分效应更重要; 同阶成分效应同等重要, 称之为成分效应排序原则. 基于成分效应排序原则, Li 等^[11] 提出了三水平的别名成分数型及 GMC 准则, 得到了顺序最小化 A_i 等价于顺序最大化别名成分数型的某些元素. 进一步, Li 等^[12] 研究了 s 水平的别名成分数型与上述准则之间的关系, 并证明了 A_3 是主效应与二因子交互成分之间混杂的平均, 在分辨度 $R \geq IV$ 的条件下, A_4 是二因子交互成分混杂的平均.

然而, 对 s 水平正规设计, 其别名成分数型的性质、以及它与字长型 A_i ($i \geq 5$) 之间的关系研究成果较少. 基于此目的, 本文主要研究别名成分数型与字长型、别名成分数之间的性质, 得到了更一般的结果, 完善了 s 水平 GMC 理论. 在本文中, 考虑分辨度 $R \geq III$ 的设计. 其内容结构如下: 第 2 节主要回顾了有关 GMC 准则的一些基本概念、记号及部分结果. 第 3 节研究了 s 水平设计的别名成分数型与字长型之间的关系. 通过别名成分数可以计算字长型, 并且根据已知字长型可以计算某些别名成分数. 在第 4 节中, 得到了二水平正规设计的别名成分数之间的计算公式. 第 5 节提供了一个小结.

§2. 基本概念及记号

在试验设计中, 将影响输出结果的变量(因素)称为因子, 用 F_1, F_2, \dots, F_n 表示. 因子的取值称为水平, 通常情况下用数字 $0, 1, \dots, s$ 来表示. 若所有因子的水平都是相同的, 则称该设计为对称的, 反之称为非对称的. 将包含所有因子水平组合的设计称为完全设计. 对于一个 s^n 对称设计, 它表示包含 n 个因子的 s 水平完全设计. 当 n 较小时, 完全设计可以给出所有因子的主效应和交互效应, 较好地分析试验数据. 但当 n 较大时, 将所有组合

都列出来, 所消耗的时间、成本往往是巨大的, 因此产生了部分因析设计. 令 $q = n - m$ 和 $N = s^q, s^{n-m}$ 设计是完全设计 s^n 的 s^{-m} 部分, 它有 n 个因子 F_i ($i = 1, 2, \dots, n$)、 q 个独立列、 m 个独立的定义字和 N 个水平组合. m 个独立的定义字生成的子群称为定义对照群. 在定义对照子群中, 除了单位元外定义字包含字母的个数, 称为该定义字的字长. $F_{i_1} \times F_{i_2}$ 称为 F_{i_1} 和 F_{i_2} 之间的两因子交互效应, 以及 $F_{j_1} \times F_{j_2} \times \dots \times F_{j_i}$ 被称为 i 阶因子交互效应, 其中 $i_1, i_2, j_i \in \{1, 2, \dots, n\}$. 任何一个 i 阶因子交互效应可以分解成 $(s-1)^{i-1}$ 个正交的 i 阶成分效应 (简称成分). 如, 三水平设计中两因子交互效应 $F_1 \times F_2$ 分解成两个 $F_1 F_2$ 和 $F_1 F_2^2$ 正交成分.

基于效应排序原则, Zhang 等^[7] 提出了二水平正规设计的别名效应数型来揭示其因子之间的混杂信息. 令 $\#_i C_j^{(k)}$ 表示与 k 个 j 阶效应别名的 i 阶效应的个数, 记

$$\#_i C_j = (\#_i C_j^{(0)}, \#_i C_j^{(1)}, \dots, \#_i C_j^{(K_j)}), \quad i, j = 0, 1, \dots, n,$$

其中 $K_j = \binom{n}{j}$. 根据如下规则将 $\#_i C_j$ 进行排列: (i) $\max(i, j) < \max(s, t)$; (ii) 当 $i < s$ 时, $\max(i, j) = \max(s, t)$; (iii) 当 $i = s$ 和 $j < t$ 时, $\max(i, j) = \max(s, t)$, 若满足 (i)–(iii) 中任意一个, 则将 $\#_i C_j$ 置于 $\#_s C_t$ 之前. 称

$$\#C = (\#_1 C_1, \#_0 C_2, \#_1 C_2, \#_2 C_1, \#_2 C_2, \dots) \quad (1)$$

为别名效应数型 (aliased effect-number pattern, AENP). 基于别名效应数型, 称顺序最大化 (1) 的二水平设计具有一般最小低阶混杂 (general minimum lower-confounding, GMC).

Zhang 和 Mukerjee^[10], Li 等^[12] 推广 GMC 准则到 s ($s \geq 2$) 水平情形, 在成分效应排序原则基础上引入别名成分数型, 仍沿用记号 $\#_i C_j^{(k)}$, 但其含义表示与 k 个 j 阶成分效应别名的 i 阶成分效应的个数. Zhang 和 Mukerjee^[10] 将新定义下的序列 (1) 简化为

$$\#C = (\#_1 C_2, \#_2 C_2, \#_1 C_3, \#_2 C_3, \#_3 C_2, \#_3 C_3, \dots). \quad (2)$$

称它为别名成分数型 (aliased component-number pattern, ACNP). 特别的, 二水平的 AENP 也称为 ACNP. 在 s 水平下, 顺序最大化序列 (2) 的设计被称为 GMC 设计. 对于向量 $\#_i C_j$ 而言, 定义 0^t 表示向量 $\#_i C_j$ 中连续 t 个零. 如果向量 $\#_i C_j$ 中有连续零元素的尾部, 可将其忽略不计.

例 1 考虑一个 3^{5-2} 设计 $D = \{1, 2, 3, 12^2, 12^2 3^2\}$ 且定义关系为 $4 = 12^2, 5 = 12^2 3^2$, 称 $1, 2, 3, 4, 5$ 为主效应, $24, 14^2, 45^2$ 为两因子交互成分. 计算得其定义对照子群为 $G = \{I, 12^2 4^2, 12^2 3^2 5^2, 12^2 345, 34^2 5\}$, 或者记为

$$I = 12^2 4^2 = 12^2 3^2 5^2 = 12^2 345 = 34^2 5.$$

由此可知, 字长型 $W(D) = (0, 0, 2, 1, 1)$ 且分辨度 $R = III$. 此外, 所有的主效应和两因子交互成分之间的混杂关系如下:

$$1 = 24, \quad 2 = 14^2, \quad 3 = 45^2, \quad 4 = 12^2 = 35, \quad 5 = 34^2,$$

$$13^2 = 25, \quad 23 = 15^2, \quad 12 = 14 = 24^2, \quad 34 = 35^2 = 45.$$

下面计算设计 D 的 ACNP 前两个元素: 在向量 $\#C_2$ 中, $\#C_2^{(0)} = 0$, $\#C_2^{(1)} = 4$, $\#C_2^{(2)} = 1$, 以及对 $k \geq 3$ 时 $\#C_2^{(k)} = 0$, 因此 $\#C_2 = (0, 4, 1)$. 对于向量 $\#C_2$, $\#C_2^{(0)} = 8$, $\#C_2^{(1)} = 6$, $\#C_2^{(2)} = 6$, 以及对 $k \geq 3$ 时 $\#C_2^{(k)} = 0$, 因此, $\#C_2 = (8, 6, 6)$.

从例 1 可知, 字长型与分辨度通过定义对照子群得到, 而别名成分数型中的元素既可以通过定义对照子群, 也可以通过因子成分之间的混杂得到. 然而, 当因子的个数 n 及 m 增大时, 对应的定义关系子群就变得更加繁杂, 尤其对 s ($s \geq 3$) 水平下的设计. 因此, 研究字长型、别名成分数型、以及他们之间的换算关系显得尤为重要, 即在别名成分数型已知的条件下, 可计算其字长型; 同样的, 由字长型可以得到别名成分数.

§3. 别名成分数型与字长型的关系

定理 2 对一个分辨度 $R \geq III$ 的 s^{n-m} 正规设计, 它的字长型 $W = (A_1, A_2, \dots, A_n)$ 与别名成分数型 $\#C_i$ 和 $\#C_1$ 满足下式

$$\begin{aligned} A_{i+1} &= \frac{1}{i+1} \sum_{k=1}^{K_i} k \#C_i^{(k)} - \left(\frac{i}{i+1} \right) (s-2) A_i - \left(\frac{n+2}{i+1} - 1 \right) (s-1) A_{i-1}, \\ &= \frac{1}{i+1} \#C_1^{(1)} - \left(\frac{i}{i+1} \right) (s-2) A_i - \left(\frac{n+2}{i+1} - 1 \right) (s-1) A_{i-1}, \quad i \geq 2, \end{aligned}$$

其中 $K_i = \binom{n}{i}$. 特别的, $A_3 = 3^{-1} \sum_{k=1}^{K_3} k \#C_2^{(k)} = 3^{-1} \#C_1^{(1)}$.

证明: 当 $i \geq 3$ 时, 由文献 [10] 可知,

$$\sum_{k \geq 1} k \#C_i^{(k)} = (n-i+1)(s-1)A_{i-1} + i(s-2)A_i + (i+1)A_{i+1}.$$

将上式进行变形, 可以得到关于 A_{i+1} 的值等于 $\sum_{k \geq 1} k \#C_i^{(k)}$, A_i , A_{i-1} 的加权之和. 又因为分辨度 $R \geq III$, 故主效应之间不混杂, 即 $\sum_{k \geq 1} k \#C_i^{(k)} = \#C_1^{(1)}$. 特别的, 当 $R \geq III$ 时, 有 $A_1 = A_2 = 0$. \square

下面将定理 2 应用于二水平及三水平正规设计.

推论 3 对于分辨度 $R \geq III$ 的 2^{n-m} 正规设计, 则

$$\begin{aligned} A_i &= \frac{1}{i} \sum_{k=1}^{K_{i-1}} k \#C_{i-1}^{(k)} - \left(\frac{n+2}{i} - 1 \right) A_{i-2} \\ &= \frac{1}{i} \#C_1^{(1)} - \left(\frac{n-i+2}{i} \right) A_{i-2}, \quad i \geq 3. \end{aligned}$$

对于二水平正规设计, 推论3揭示了字长型中元素 A_i 与 A_{i-2} 之间的关系. 此外, 当 i 取不同值时, 可以由 A_{i-2} 得到二水平下的字长 A_i 与 $\#C_{i-1}$ (或 $\#C_1$) 之间的表达公式, 如:

$$A_3 = \frac{1}{3} \sum_{k=1}^{K_2} k \# C_2^{(k)}, \quad A_4 = \frac{1}{4} \sum_{k=1}^{K_2} k \# C_3^{(k)}, \quad A_5 = \frac{1}{5} \sum_{k=1}^{K_4} k \# C_4^{(k)} - \frac{n-3}{15} \sum_{k=1}^{K_2} k \# C_2^{(k)}.$$

例4 考虑分辨度 $R = IV$ 的两个 2^{9-4} 设计 d_1 和 d_2 :

$$d_1 : I = 1236 = 1247 = 1258 = 13459,$$

$$d_2 : I = 1236 = 1247 = 1348 = 23459.$$

因为设计 d_1, d_2 的分辨度 $R = IV$, 所以 $\sum_{k=1}^{K_1} k \# C_1^{(k)}(d_i) = \sum_{k=1}^{K_2} k \# C_2^{(k)}(d_i) = 0$ ($i = 1, 2$).

表1 设计 d_1, d_2 的 $\sum_{k=1}^{K_i} k \# C_i^{(k)}$ ($i = 3, 4, \dots, 9$)

设计	$\sum_{k=1}^{K_3} k \# C_3^{(k)}$	$\sum_{k=1}^{K_4} k \# C_4^{(k)}$	$\sum_{k=1}^{K_5} k \# C_5^{(k)}$	$\sum_{k=1}^{K_6} k \# C_6^{(k)}$	$\sum_{k=1}^{K_7} k \# C_7^{(k)}$	$\sum_{k=1}^{K_8} k \# C_8^{(k)}$	$\sum_{k=1}^{K_9} k \# C_9^{(k)}$
d_1	24	40	30	32	8	0	1
d_2	28	35	35	28	0	9	0

由表1知, $\sum_{k=1}^{K_3} k \# C_3^{(k)}(d_1) = 24$, $\sum_{k=1}^{K_3} k \# C_3^{(k)}(d_2) = 28$, 根据推论3, 可以得出 $A_4(d_1) = 6$, $A_4(d_2) = 7$. 同理, $A_5(d_1) = 8$, $A_6(d_1) = 0$, $A_7(d_1) = 0$, $A_8(d_1) = 1$, $A_9(d_1) = 0$ 以及 $A_5(d_2) = 7$, $A_6(d_2) = 0$, $A_7(d_2) = 0$, $A_8(d_2) = 0$, $A_9(d_2) = 1$. 所以设计 d_1, d_2 的字长型分别为 $W(d_1) = (0, 0, 0, 6, 8, 0, 0, 1, 0)$ 和 $W(d_2) = (0, 0, 0, 7, 7, 0, 0, 0, 1)$. 类似的, 也可通过推论3的第二种表达来计算设计的字长型.

根据推论3, 可以得到 A_i 与 $\#C_j$ 之间更一般的表达公式.

推论5 对于 2^{n-m} 正规设计, 则

$$A_i = \begin{cases} \sum_{j=1}^{i/2} (-1)^j \frac{(2j-2)!!(n-2j)!!}{i!!(n-i)!!} \sum_{k=1}^{K_{2j-1}} k \# C_{2j-1}^{(k)}, & \text{若 } i = 4t; \\ \sum_{j=1}^{(i-1)/2} (-1)^j \frac{(2j-1)!!(n-2j-1)!!}{i!!(n-i)!!} \sum_{k=1}^{K_{2j}} k \# C_{2j}^{(k)}, & \text{若 } i = 4t+1; \\ \sum_{j=1}^{i/2} (-1)^{j+1} \frac{(2j-2)!!(n-2j)!!}{i!!(n-i)!!} \sum_{k=1}^{K_{2j-1}} k \# C_{2j-1}^{(k)}, & \text{若 } i = 4t+2; \\ \sum_{j=1}^{(i-1)/2} (-1)^{j+1} \frac{(2j-1)!!(n-2j-1)!!}{i!!(n-i)!!} \sum_{k=1}^{K_{2j}} k \# C_{2j}^{(k)}, & \text{若 } i = 4t+3. \end{cases}$$

规定 $0!! = 1$, $(2t)!! = 2t \cdot (2t-2) \cdots 4 \cdot 2$, $(2t+1)!! = (2t+1) \cdot (2t-1) \cdots 3 \cdot 1$.

证明: 由推论 3 可知, 在二水平设计下的字长 A_i 可由 $\sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)}$ 和 A_{i-2} 表示. 同样地, A_{i-2} 也可以由 $\sum_{k=1}^{K_{i-3}} k_1^{\#} C_{i-3}^{(k)}$ 和 A_{i-4} 表示. 重复上述过程, 字长 A_i 可由 $\sum_{k=1}^{K_j} k_1^{\#} C_j^{(k)}$, $j = i-1, i-3, \dots, 2$ 或 1 表示. 另一方面, 根据推论 3, $\sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)}$ 的系数为 $1/i$, $\sum_{k=1}^{K_{i-3}} k_1^{\#} C_{i-3}^{(k)}$ 的系数为 $-(n-i+2)/[i(i-2)]$, $\sum_{k=1}^{K_{i-5}} k_1^{\#} C_{i-5}^{(k)}$ 的系数为 $(n-i+2)(n-i)/[i(i-2+4)(i-4)]$, 以此类推. 因此, 我们不仅要考虑 i 的奇偶性, 同时也要考虑 $i/2$ 的奇偶性, 故将 i 分为以下四类:

(i) 当 $i = 4t$ 时, 则

$$\begin{aligned} A_i &= \frac{1}{i} \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} - \frac{n-i+2}{i(i-2)} \sum_{k=1}^{K_{i-3}} k_1^{\#} C_{i-3}^{(k)} + \dots - \frac{(n-i+2) \cdots (n-2)n}{i \cdot (i-2) \cdots 4 \cdot 2} \sum_{k=1}^{K_1} k_1^{\#} C_1^{(k)}, \\ &= \sum_{j=1}^{i/2} (-1)^j \frac{(2j-2)!!(n-2j)!!}{i!!(n-i)!!} \sum_{k=1}^{K_{2j-1}} k_1^{\#} C_{2j-1}^{(k)}. \end{aligned}$$

(ii) 当 $i = 4t+1$ 时, 则

$$\begin{aligned} A_i &= \frac{1}{i} \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} - \frac{n-i+2}{i(i-2)} \sum_{k=1}^{K_{i-3}} k_1^{\#} C_{i-3}^{(k)} + \dots - \frac{(n-i+2) \cdots (n-5)(n-3)}{i \cdot (i-2) \cdots 3 \cdot 1} \sum_{k=1}^{K_2} k_1^{\#} C_2^{(k)}, \\ &= \sum_{j=1}^{(i-1)/2} (-1)^j \frac{(2j-1)!!(n-2j-1)!!}{i!!(n-i)!!} \sum_{k=1}^{K_{2j}} k_1^{\#} C_{2j}^{(k)}. \end{aligned}$$

(iii) 当 $i = 4t+2$ 时, 则

$$\begin{aligned} A_i &= \frac{1}{i} \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} - \frac{n-i+2}{i(i-2)} \sum_{k=1}^{K_{i-3}} k_1^{\#} C_{i-3}^{(k)} + \dots + \frac{(n-i+2) \cdots (n-2)n}{i \times (i-2) \cdots 4 \cdot 2} \sum_{k=1}^{K_1} k_1^{\#} C_1^{(k)}, \\ &= \sum_{j=1}^{i/2} (-1)^{j+1} \frac{(2j-2)!!(n-2j)!!}{i!!(n-i)!!} \sum_{k=1}^{K_{2j-1}} k_1^{\#} C_{2j-1}^{(k)}. \end{aligned}$$

(iv) 当 $i = 4t+3$ 时, 则

$$\begin{aligned} A_i &= \frac{1}{i} \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} - \frac{n-i+2}{i(i-2)} \sum_{k=1}^{K_{i-3}} k_1^{\#} C_{i-3}^{(k)} + \dots + \frac{(n-i+2) \cdots (n-5)(n-3)}{i \cdot (i-2) \cdots 3 \cdot 1} \sum_{k=1}^{K_2} k_1^{\#} C_2^{(k)}, \\ &= \sum_{j=1}^{(i-1)/2} (-1)^{j+1} \frac{(2j-1)!!(n-2j-1)!!}{i!!(n-i)!!} \sum_{k=1}^{K_{2j}} k_1^{\#} C_{2j}^{(k)}. \end{aligned}$$

故得证. \square

推论 5 揭示了二水平下字长型与别名效应数型之间的关系, 是推论 3 更一般的推广. 在别名效应数型 ${}_1^{\#} C_j$ 已知的条件下, 可以使用它来计算该设计的字长型.

推论 6 对于分辨度 $R \geq III$ 的 3^{n-m} 正规设计, 则

$$\begin{aligned} A_i &= \frac{1}{i} \sum_{k=1}^{K_{i-1}} k \# C_{i-1}^{(k)} - 2 \left(\frac{n+2-i}{i} \right) A_{i-2} - \left(\frac{i-1}{i} \right) A_{i-1} \\ &= \frac{1}{i} \# C_1^{(1)} - 2 \left(\frac{n+2-i}{i} \right) A_{i-2} - \left(\frac{i-1}{i} \right) A_{i-1}. \end{aligned}$$

对于三水平正规设计, 推论 6 揭示了字长型中 A_i 与 A_{i-1}, A_{i-2} 之间的关系. 类似于推论 3, 通过简单计算, 可以得到三水平下字长型 A_i 与别名成分数型之间的关系, 如:

$$A_3 = \frac{1}{3} \# C_1^{(1)}, \quad A_4 = \frac{1}{4} \# C_1^{(1)} - \frac{1}{4} \# C_1^{(1)}, \quad A_5 = \frac{1}{5} \# C_1^{(1)} - \frac{1}{5} \# C_1^{(1)} - \frac{2n-3}{15} \# C_1^{(1)}.$$

定理 7 对分辨度 $R \geq III$ 的 s^{n-m} 正规设计, 则

$$\begin{aligned} (a) \quad \# C_0^{(k)} &= \begin{cases} N_i - A_i, & k = 0, \\ A_i, & k = 1, \\ 0, & k \geq 2; \end{cases} \\ (b) \quad \# C_1^{(k)} &= \begin{cases} N_i - (n-i+1)(s-1)A_{i-1} - [i(s-2)-1]A_i - (i+1)A_{i+1}, & k = 0, \\ (n-i+1)(s-1)A_{i-1} + i(s-2)A_i + (i+1)A_{i+1}, & k = 1, \\ 0, & k \geq 2, \end{cases} \end{aligned}$$

其中 $N_i = \binom{n}{i}(s-1)^{i-1}$. 规定 $A_0 = 0$.

证明: (a) 因为 $\# C_0^{(k)}$ 表示 i 阶成分效应与 k 个 I 别名的个数. 由定义可知, 当 $k=1$ 时, 它表示定义关系子群中定义字长的个数, 因此 $\# C_0^{(1)} = A_i$. 当 $k=0$ 时, 表示 i 阶成分效应不与 I 别名的个数, 因为 i 阶成分效应的总个数为 $\binom{n}{i}(s-1)^{i-1}$, 因此 $\# C_0^{(0)} = \binom{n}{i}(s-1)^{i-1} - A_i$. 此外, 当 $k \geq 2$ 时, $\# C_0^{(k)} = 0$.

(b) 因为分辨度 $R \geq III$, 故主效应之间不混杂, 因此当 $k \geq 2$ 时, $\# C_1^{(k)} = 0$. 当 $k=1$ 时, 由定理 2 可知, $\# C_1^{(1)} = \sum_{k=1}^{K_i} k \# C_i^{(k)}$, 而 $\sum_{k=1}^{K_i} k \# C_i^{(k)} = (n-i+1)(s-1)A_{i-1} + i(s-2)A_i + (i+1)A_{i+1}$, 故 $\# C_1^{(1)} = (n-i+1)(s-1)A_{i-1} + i(s-2)A_i + (i+1)A_{i+1}$. 因为 i 阶成分效应的总个数为 $\binom{n}{i}(s-1)^{i-1}$, 故 $\# C_1^{(0)} = \binom{n}{i}(s-1)^{i-1} - (n-i+1)(s-1)A_{i-1} - [i(s-2)-1]A_i - (i+1)A_{i+1}$. \square

定理 7 得到了别名成分数 $\# C_0^{(k)}$ 与 A_i , 以及 $\# C_1^{(k)}$ 与 A_j ($j = i-1, i, i+1$) 之间的关系. 也就是说, 通过一个设计的字长型 W 可计算出 ACNP 中的元素 $\# C_0$ 和 $\# C_1$. 下面的结论可以由定理 7 得到.

推论 8 对任意 s^{n-m} 正规设计, 则

$$(a) \text{ 若分辨度 } R \geq III, \text{ 则 } \# C_1 = \left(\binom{n}{2}(s-1) - 3A_3, 3A_3 \right);$$

(b) 若分辨度 $R = IV$, 则 $\#_2 C_1 = \left(\binom{n}{2}(s-1), 0\right)$ 且 $\#_3 C_1 = \left(\binom{n}{3}(s-1)^2 - 4A_3, 4A_3\right)$.

证明: (a) 当分辨度 $R \geq III$, 则 $A_1 = A_2 = 0$. 由定理 7 可知, 当 $i = 2$ 时, 有

$$\begin{aligned}\#_2 C_1^{(0)} &= \binom{n}{2}(s-1) - (n-1)(s-1)A_1 - [2(s-2)-1]A_2 - 3A_3, \\ \#_2 C_1^{(1)} &= (n-1)(s-1)A_1 + 2(s-2)A_2 + 3A_3.\end{aligned}$$

所以 $\#_2 C_1 = \left(\binom{n}{2}(s-1) - 3A_3, 3A_3\right)$.

(b) 当分辨度 $R = IV$ 时, $A_3 = 0$. 由定理 7 可得, $\#_2 C_1 = \left(\binom{n}{2}(s-1), 0\right)$. 同理, $\#_3 C_1 = \left(\binom{n}{3}(s-1)^2 - 4A_4, 4A_4\right)$. \square

§4. 二水平下别名效应数型之间的关系

针对二水平设计, Zhang 和 Cheng^[8] 通过别名效应数型得到了 A_3, A_4 以及 A_5 的计算公式, 其中 A_3 的表达式与本文中一致, 而 A_4 及 A_5 的表达如下:

$$A_4 = \frac{1}{6} \sum_{k=1}^{K_2} k \#_2 C_2^{(k)}, \quad A_5 = \frac{1}{10} \sum_{k=1}^{K_3} k \#_2 C_3^{(k)} - \frac{n-3}{10} \sum_{k=1}^{K_2} k \#_1 C_2^{(k)}. \quad (3)$$

根据推论 5 中字长 A_i 与 $\sum_{k=1}^{K_j} k \#_1 C_j^{(k)}$ 之间的关系, 可得

$$A_4 = \frac{1}{4} \sum_{k=1}^{K_2} k \#_1 C_3^{(k)}, \quad A_5 = \frac{1}{5} \sum_{k=1}^{K_4} k \#_1 C_4^{(k)} - \frac{(n-3)}{5!!} \sum_{k=1}^{K_2} k \#_1 C_2^{(k)}. \quad (4)$$

将 (3) 和 (4) 进行对比, 发现字长 A_4 和 A_5 的表达不同. 但在同一个设计中, A_i 是唯一确定的. 因此, 将两种关于字长的表达式进行联立, 从而得到

$$\begin{aligned}\frac{1}{4} \sum_{k=1}^{K_2} k \#_1 C_3^{(k)} &= \frac{1}{6} \sum_{k=1}^{K_2} k \#_2 C_2^{(k)}, \\ \frac{1}{5} \sum_{k=1}^{K_4} k \#_1 C_4^{(k)} - \frac{(n-3)}{5!!} \sum_{k=1}^{K_2} k \#_1 C_2^{(k)} &= \frac{1}{10} \sum_{k=1}^{K_3} k \#_2 C_3^{(k)} - \frac{n-3}{10} \sum_{k=1}^{K_2} k \#_1 C_2^{(k)}.\end{aligned}$$

对于二水平设计, Zhang 和 Park^[13] 定义了 ${}_l C_m$ 为 l 阶与 m 阶因子交互效应的别名系数, 得到了 ${}_l C_m$ 与字长型之间的关系

$${}_l C_m = \sum_{k=0}^l \binom{n-(m-l+2k)}{l-k} \binom{m-l+2k}{k} A_{m-l+2k}. \quad (5)$$

这里规定 $\binom{n}{0} = 0$, 如果 $x < y$ 或者 $x < 0$ 时, 则 $\binom{x}{y} = 0$. Zhang 等^[7] 得到别名关系数与别名效应数型的关系为

$${}_i C_j = \sum_{k=1}^{K_j} k \#_i C_j^{(k)}. \quad (6)$$

从上面的式子中, 我们发现 $\sum_{k=1}^{K_i} k_2^{\#} C_i^{(k)}$ 可以由 $\sum_{k=1}^{K_j} k_1^{\#} C_j^{(k)} (j = 1, 2, \dots, i+1)$ 表示. 下面的结论揭示了这一规律.

定理 9 对于 2^{n-m} 正规设计, 若 $i \geq 2$ 且 t 为正整数, 则

(i) 当 $i = 4t$ 时, 有

$$\begin{aligned} \sum_{k=1}^{K_i} k_2^{\#} C_i^{(k)} &= \frac{n}{2i} (n-i+2) \sum_{j=1}^{(i-2)/2} (-1)^{j+1} \frac{(2j-2)!!(n-2j)!!}{(i-2)!!(n-i+2)!!} \sum_{k=1}^{K_{2j-1}} k_1^{\#} C_{2j-1}^{(k)} \\ &\quad + \frac{1}{2i} (i-1)(n-i) \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} + \frac{i+1}{2} \sum_{k=1}^{K_{i+1}} k_1^{\#} C_{i+1}^{(k)}. \end{aligned}$$

(ii) 当 $i = 4t+1$ 时, 有

$$\begin{aligned} \sum_{k=1}^{K_i} k_2^{\#} C_i^{(k)} &= \frac{n}{2i} (n-i+2) \sum_{j=1}^{(i-3)/2} (-1)^{j+1} \frac{(2j-1)!!(n-2j-1)!!}{(i-2)!!(n-i+2)!!} \sum_{k=1}^{K_{2j}} k_1^{\#} C_{2j}^{(k)} \\ &\quad + \frac{1}{2i} (i-1)(n-i) \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} + \frac{i+1}{2} \sum_{k=1}^{K_{i+1}} k_1^{\#} C_{i+1}^{(k)}. \end{aligned}$$

(iii) 当 $i = 4t+2$ 时, 有

$$\begin{aligned} \sum_{k=1}^{K_i} k_2^{\#} C_i^{(k)} &= \frac{n}{2i} (n-i+2) \sum_{j=1}^{(i-2)/2} (-1)^j \frac{(2j-2)!!(n-2j)!!}{(i-2)!!(n-i+2)!!} \sum_{k=1}^{K_{2j-1}} k_1^{\#} C_{2j-1}^{(k)} \\ &\quad + \frac{1}{2i} (i-1)(n-i) \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} + \frac{i+1}{2} \sum_{k=1}^{K_{i+1}} k_1^{\#} C_{i+1}^{(k)}. \end{aligned}$$

(iv) 当 $i = 4t+3$ 时, 有

$$\begin{aligned} \sum_{k=1}^{K_i} k_2^{\#} C_i^{(k)} &= \frac{n}{2i} (n-i+2) \sum_{j=1}^{(i-3)/2} (-1)^j \frac{(2j-1)!!(n-2j-1)!!}{(i-2)!!(n-i+2)!!} \sum_{k=1}^{K_{2j}} k_1^{\#} C_{2j}^{(k)} \\ &\quad + \frac{1}{2i} (i-1)(n-i) \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} + \frac{i+1}{2} \sum_{k=1}^{K_{i+1}} k_1^{\#} C_{i+1}^{(k)}. \end{aligned}$$

证明: 当 $i < j$ 时, 根据 (5) 和 (6), 可得

$$\sum_{k=1}^{K_j} k_1^{\#} C_j^{(k)} = \sum_{l=0}^i \binom{n-(j-i+2l)}{i-l} \binom{j-i+2l}{l} A_{j-i+2l}.$$

当 $i = 1$ 时, 有

$$\sum_{k=1}^{K_j} k_1^{\#} C_j^{(k)} = (n-j+1)A_{j-1} + (j+1)A_{j+1}. \quad (7)$$

当 $i = 2$ 时, 有

$$2 \sum_{k=1}^{K_j} k_2^{\#} C_j^{(k)} = (n-j+2)(n-j+1)A_{j-2} + 2j(n-j)A_j + (j+2)(j+1)A_{j+2}. \quad (8)$$

在 (7) 中用 $j+1$ 替换 j , 得

$$\sum_{k=1}^{K_{j+1}} k_1^{\#} C_{j+1}^{(k)} = (n-j)A_j + (j+2)A_{j+2}. \quad (9)$$

所以,

$$(j+2)A_{j+2} = \sum_{k=1}^{K_{j+1}} k_1^{\#} C_{j+1}^{(k)} - (n-j)A_j.$$

将其代入 (8) 式中, 整理得

$$2 \sum_{k=1}^{K_j} k_2^{\#} C_j^{(k)} = (n-j+2)(n-j+1)A_{j-2} + (j-1)(n-j)A_j + (j+1) \sum_{k=1}^{K_{j+1}} k_1^{\#} C_{j+1}^{(k)}.$$

根据推论 3 知

$$A_j = \frac{1}{j} \sum_{k=1}^{K_{j-1}} k_1^{\#} C_{j-1}^{(k)} - \left(\frac{n-j+2}{j} \right) A_{j-2},$$

将 A_j 代入上式得

$$\sum_{k=1}^{K_i} k_2^{\#} C_i^{(k)} = \frac{n}{2i} (n-i+2) A_{i-2} + \frac{1}{2i} (i-1)(n-i) \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} + \frac{i+1}{2} \sum_{k=1}^{K_{i+1}} k_1^{\#} C_{i+1}^{(k)}.$$

由推论 5, 我们将 i 分为四类:

(i) 当 $i = 4t$ 时, $i-2$ 为 $4t+2$ 类, 则

$$\begin{aligned} \sum_{k=1}^{K_i} k_2^{\#} C_i^{(k)} &= \frac{n}{2i} (n-i+2) \sum_{j=1}^{(i-2)/2} (-1)^{j+1} \frac{(2j-2)!!(n-2j)!!}{(i-2)!!(n-i+2)!!} \sum_{k=1}^{K_{2j-1}} k_1^{\#} C_{2j-1}^{(k)} \\ &\quad + \frac{1}{2i} (i-1)(n-i) \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} + \frac{i+1}{2} \sum_{k=1}^{K_{i+1}} k_1^{\#} C_{i+1}^{(k)}. \end{aligned}$$

(ii) 当 $i = 4t+1$ 时, $i-2$ 为 $4t+3$ 类, 则

$$\begin{aligned} \sum_{k=1}^{K_i} k_2^{\#} C_i^{(k)} &= \frac{n}{2i} (n-i+2) \sum_{j=1}^{(i-3)/2} (-1)^{j+1} \frac{(2j-1)!!(n-2j-1)!!}{(i-2)!!(n-i+2)!!} \sum_{k=1}^{K_{2j}} k_1^{\#} C_{2j}^{(k)} \\ &\quad + \frac{1}{2i} (i-1)(n-i) \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} + \frac{i+1}{2} \sum_{k=1}^{K_{i+1}} k_1^{\#} C_{i+1}^{(k)}. \end{aligned}$$

(iii) 当 $i = 4t+2$ 时, $i-2$ 为 $4t$ 类, 则

$$\sum_{k=1}^{K_i} k_2^{\#} C_i^{(k)} = \frac{n}{2i} (n-i+2) \sum_{j=1}^{(i-2)/2} (-1)^j \frac{(2j-2)!!(n-2j)!!}{(i-2)!!(n-i+2)!!} \sum_{k=1}^{K_{2j-1}} k_1^{\#} C_{2j-1}^{(k)}$$

$$+ \frac{1}{2i}(i-1)(n-i) \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} + \frac{i+1}{2} \sum_{k=1}^{K_{i+1}} k_1^{\#} C_{i+1}^{(k)}.$$

(iv) 当 $i = 4t+3$ 时, $i-2$ 为 $4t+1$ 类, 则

$$\begin{aligned} \sum_{k=1}^{K_i} k_2^{\#} C_i^{(k)} &= \frac{n}{2i}(n-i+2) \sum_{j=1}^{(i-3)/2} (-1)^j \frac{(2j-1)!!(n-2j-1)!!}{(i-2)!!(n-i+2)!!} \sum_{k=1}^{K_{2j}} k_1^{\#} C_{2j}^{(k)} \\ &\quad + \frac{1}{2i}(i-1)(n-i) \sum_{k=1}^{K_{i-1}} k_1^{\#} C_{i-1}^{(k)} + \frac{i+1}{2} \sum_{k=1}^{K_{i+1}} k_1^{\#} C_{i+1}^{(k)}. \end{aligned}$$

上述条件都需满足 $i > 2$, 当 $i = 2$ 时,

$$\sum_{k=1}^{K_2} k_2^{\#} C_2^{(k)} = \frac{3}{2} \sum_{k=1}^{K_3} k_1^{\#} C_3^{(k)}$$

仍满足, 故 $i \geq 2$ 时, 上式也成立. \square

针对二水平设计, 定理 9 得到了 ${}_2^{\#} C_i$ 与 ${}_1^{\#} C_j$ 之间的关系. 下面的推论为定理 9 提供了一些特例.

推论 10 对分辨度 $R \geq III$ 的 2^{n-m} 正规设计, 则

$$\begin{aligned} \sum_{k=1}^{K_2} k_2^{\#} C_2^{(k)} &= \frac{3}{2} \sum_{k=1}^{K_3} k_1^{\#} C_3^{(k)}, \\ \sum_{k=1}^{K_3} k_2^{\#} C_3^{(k)} &= \frac{1}{3}(n-3) \sum_{k=1}^{K_2} k_1^{\#} C_2^{(k)} + 2 \sum_{k=1}^{K_4} k_1^{\#} C_4^{(k)}, \\ \sum_{k=1}^{K_3} k_3^{\#} C_3^{(k)} &= \frac{2}{3}(n-4) \sum_{k=1}^{K_3} k_1^{\#} C_3^{(k)} + \frac{10}{3} \sum_{k=1}^{K_5} k_1^{\#} C_5^{(k)}. \end{aligned}$$

证明: 由定理 9, 当 $i = 2, 3$ 时, 计算可得二水平下 $\sum_{k=1}^{K_2} k_2^{\#} C_2^{(k)}$, $\sum_{k=1}^{K_3} k_2^{\#} C_3^{(k)}$ 的表达式.

Chen 和 Liu^[9] 给出了 A_6 的表达式

$$A_6 = \frac{1}{20} \sum_{k=1}^{K_3} k_3^{\#} C_3^{(k)} - \frac{n-4}{20} \sum_{k=1}^{K_2} k_2^{\#} C_2^{(k)}. \quad (10)$$

由推论 5 可得

$$A_6 = \frac{1}{6} \sum_{k=1}^{K_5} k_1^{\#} C_5^{(k)} - \frac{(n-4)}{6 \times 4} \sum_{k=1}^{K_3} k_1^{\#} C_3^{(k)}. \quad (11)$$

将 (10), (11) 联立, 知

$$\sum_{k=1}^{K_3} k_3^{\#} C_3^{(k)} = \frac{10}{3} \sum_{k=1}^{K_5} k_1^{\#} C_5^{(k)} - \frac{5(n-4)}{6} \sum_{k=1}^{K_3} k_1^{\#} C_3^{(k)} + (n-4) \sum_{k=1}^{K_2} k_2^{\#} C_2^{(k)}.$$

又因为

$$\sum_{k=1}^{K_2} k_2^{\#} C_2^{(k)} = \frac{3}{2} \sum_{k=1}^{K_3} k_1^{\#} C_3^{(k)},$$

代入即得

$$\sum_{k=1}^{K_3} k \# C_3^{(k)} = \frac{2}{3}(n-4) \sum_{k=1}^{K_3} k \# C_3^{(k)} + \frac{10}{3} \sum_{k=1}^{K_5} k \# C_5^{(k)}.$$

得证. \square

§5. 小结

本文主要研究 s ($s \geq 2$) 水平正规设计的别名成分数型与字长型、字长型各元素、以及部分别名成分数型之间的关系. 这些结论推广了文献 [8, 9, 12] 中别名成分数型与字长型之间的结论, 并且得到了一些新的结论, 进一步完善了 GMC 理论. 如, 定理 2 通过别名成分数型中的 $\#C_i$ 得到了字长型 A_{i+1} 与 A_i 和 A_{i-1} ($i \geq 2$) 之间的迭代关系, 为计算字长型提供依据. 由定理 2 得到了若干推论, 如二、三水平正规设计中别名成分数型与字长型之间的联系 (推论 3 和 6). 进一步, 得到二水平下字长 A_i 与 $\#C_j$ 之间的一般表达 (推论 5). 另一方面, 对 s 水平设计, 定理 7 揭示了通过字长型可以计算部分别名成分数型, 如由 A_3 得到 $\#C_1$ 和 $\#C_2$ 的计算公式 (推论 8). 在研究中, 我们发现别名成分数型之间也存在着某种联系, 定理 9 和推论 10 给出了二水平正规设计中某些别名成分数型之间的数学表达. 然而, 我们尚未得到三水平以上设计的别名成分数型之间的关系, 这将在以后的工作中继续研究.

参 考 文 献

- [1] BOX G E P, HUNTER J S. The 2^{k-p} fractional factorial designs part I, part II [J]. *Technometrics*, 1961, **3**(3): 311–351 and **3**(4): 449–458.
- [2] FRIES A, HUNTER W G. Minimum aberration 2^{k-p} designs [J]. *Technometrics*, 1980, **22**(4): 601–608.
- [3] SUN D X. Estimation capacity and related topics in experimental designs [D]. Waterloo, Canada: University of Waterloo, 1993.
- [4] HU J W, ZHANG R C. Some results on two-level regular designs with general minimum lower-order confounding [J]. *J Statist Plann Inference*, 2011, **141**(5): 1774–1782.
- [5] MUKERJEE R, WU C F J. *A Modern Theory of Factorial Designs* [M]. New York: Springer, 2006.
- [6] 赵胜利. 正规部分因子设计的最优化理论与构造方法 [J]. 应用概率统计, 2015, **31**(3): 320–336.
- [7] ZHANG R C, LI P, ZHAO S L, et al. A general minimum lower-order confounding criterion for two-level regular designs [J]. *Statist Sinica*, 2008, **18**(4): 1689–1705.
- [8] ZHANG R C, CHENG Y. General minimum lower order confounding designs: an overview and a construction theory [J]. *J Statist Plann Inference*, 2010, **140**(7): 1719–1730.
- [9] CHEN J, LIU M Q. Some theory for constructing general minimum lower order confounding designs [J]. *Statist Sinica*, 2011, **21**(4): 1541–1555.
- [10] ZHANG R C, MUKERJEE R. Characterization of general minimum lower order confounding via complementary sets [J]. *Statist Sinica*, 2009, **19**(1): 363–375.

- [11] LI Z M, ZHANG T F, ZHANG R C. Three-level regular designs with general minimum lower-order confounding [J]. *Canad J Statist*, 2013, **41**(1): 192–210.
- [12] LI Z M, ZHAO S L, ZHANG R C. On general minimum lower order confounding criterion for s -level regular designs [J]. *Statist Probab Lett*, 2015, **99**: 202–209.
- [13] ZHANG R C, PARK D K. Optimal blocking of two-level fractional factorial designs [J]. *J Statist Plann Inference*, 2000, **91**(1): 107–121.

Some Properties of Aliased Component-Number Pattern for Regular Fractional Factorial Designs

LI Zhi LI Zhiming

(College of Mathematics and System Science, Xinjiang University, Urumqi, 830046, China)

Abstract: General minimum lower-order confounding and minimum aberration are two important criteria to select s ($s \geq 2$)-level optimal regular fractional factorial designs. Their classification are based on the aliased component-number and word-length patterns, respectively. The paper mainly studies some properties of the aliased component-number pattern for s -level regular designs. We obtain that the elements of word-length pattern are expressed as some functions of aliased component-numbers under s -level case. It reveals the relationship between the aliased component-number and word-length patterns. On the other hand, we can calculate some aliased component-numbers by word-length pattern. Further, the formulas of some aliased component-numbers are provided for two-level designs.

Keywords: regular design; defining relation; word-length pattern; minimum aberration criterion; aliased component-number pattern; general minimum lower-order confounding criterion

2020 Mathematics Subject Classification: 62K15; 62K05