区组大小不等的主效应设计

陈雪平¹, 林金官^{2*}, 黄性芳², 汪红霞²

1. 江苏理工学院数理学院, 常州 213001;

2. 南京审计大学统计科学与大数据研究院, 南京 211815

E-mail: chenxueping@jstu.edu.cn, jglin@nau.edu.cn, xfhuang@seu.edu.cn, hxwang@nau.edu.cn

收稿日期: 2016-04-11; 接受日期: 2016-07-04; 网络出版日期: 2017-04-24; * 通信作者

国家自然科学基金 (批准号: 11571073, 11601195 和 11471172)、教育部高等学校博士点基金 (批准号: 20120092110021)、江苏省自然科学基金 (批准号: BK20141326 和 BK20160289)、江苏高校“青蓝工程”和江苏省高校自然科学基金 (批准号: 16KJB110005) 资助项目

摘要 区组设计被广泛应用在人类的各种探索试验中. 以往的研究中, 区组大小通常被假定是相等的, 但时常与实际情形相悖. 本文讨论了区组大小不等时的正交主效应设计, 该类设计的处理主效应可以通过区组因子达到正交; 同时讨论了相关的设计性质. 最后, 本文给出了若干构造方法, 包括如何添加新的区组, 以及如何叠加不同区组大小的设计. 结果表明相应设计具有较为灵活的区组数、区组大小和因子数.

关键词 关联矩阵 正交 等比例频数条件 加权相遇数

MSC (2010) 主题分类 62K10, 05B05

1 引言

很多试验中, 我们利用区组技术来进一步控制误差, 从而提高数据分析的精确性和可靠性. 常记区组大小为 k_i ($i = 1, 2, \dots, b$), b 为区组个数, 从文献 [1–4] 中可见, 区组大小常被假定是相等的, 且尽可能小. 但是, 正如文献 [5] 所指出的, 区组的大小完全依赖于实际情况, 而且在试验前应该已经确定, 即区组大小不必要相等. 例如, 一胎动物的数量可能是不等的, 一批树木之间的间隙也可能是不等的 (参见文献 [5]).

如果区组设计只包含一个处理因子, 对于相应的数据分析和方法, 无论是正交情形还是非正交情形都较成熟. 非正交情形的区组设计也经常得到应用, 如平衡不完全区组设计 (balanced incomplete blocked designs) 和部分平衡区组设计 (partially balanced block designs). 更多细节可以参见文献 [1, 2]. 对于这类情形, 文献 [5] 首次研究了区组大小不等的情形, 后期又有学者提出了若干数据分析方法, 参见文献 [6, 7]. 另外, 构造带有不等区组大小的等重复平衡区组设计的相关讨论可参见文献 [8].

当区组设计包含两个或者更多的处理因子时, 分辨度为 III 的设计在其中起到了重要的作用, 这些设计均满足文献 [9] 提出的等比例频数条件. 对于这类设计, 基本要求区组因子与处理因子是正交的, 于是, 区组大小和设计的试验次数需满足一定的条件, 导致试验次数往往较大, 特别是在处理因子

英文引用格式: Chen X P, Lin J G, Huang X F, et al. Main-effect plans with blocks of natural size (in Chinese). Sci Sin Math, 2017, 47: 765–778, doi: 10.1360/SCM-2016-0240

的水平数不全相等时. 为了克服这些缺点, 文献 [3] 提出了一类给定区组大小的主效应设计, 其中任意处理因子之间是相互正交的, 但处理因子与区组因子之间允许不正交. 由于该类设计建立在正交设计的基础上, 故试验次数仍然较大. 为了减少试验次数, 并保持数据分析的正交性, 文献 [10] 得到了一类新的区组大小为 2 的正交主效应设计, 其满足文献 [3] 中设计的性质要求, 但只需更少的区组. 通过推广文献 [10] 中的想法, 文献 [11, 12] 提出了一类处理因子通过区组因子达到正交的饱和主效应设计 (POTB). 相应设计的试验次数比正交主效应设计要少, 且处理因子虽与区组因子之间是不正交的, 但可以通过区组因子使得处理因子之间达到正交. 进一步地, 文献 [11] 证明了对于 POTB 设计, 每个因子的任意主效应对照是可估的, 且任意两个因子的主效应对照的最佳线性无偏估计量 (BLUE) 之间是不相关的.

然而, 前述文献 [10, 11] 中的方法不能处理区组大小不等时的情形. 事实上, 讨论区组大小不等时的多因子主效应设计一直是一个难点. 本文将提出若干该类设计的数据分析和构造方法.

全文结构如下. 第 2 节给出多因子主效应设计的基本概念和符号约定, 同时给出几个有关加权相遇数的引理, 其是处理区组大小不等的主效应设计的基础. 第 3 节给出了处理因子通过区组因子达到正交的一个充分条件, 讨论了相应设计的若干性质, 推广了文献 [11] 的结果. 第 4 节给出了若干区组大小不等的 POTB 设计的构造方法. 这些方法均较容易操作, 且所得设计的区组个数和因子数较为灵活. 第 5 节是本文的总结.

2 基本概念

2.1 符号和背景

首先给出一些符号和说明. 根据文献 [11], 记一个试验次数为 n 的设计 D 含有一个区组因子 F_0 和 m 个处理因子 $F_1, F_2, \dots, F_m$, 水平数分别为 $b, p_1, p_2, \dots, p_m$, 且该设计含有 b 个区组, 大小分别为 $k_1, k_2, \dots, k_b$. 称设计 D 是饱和的, 如果试验次数满足

$$\sum_{i=1}^m (p_i - 1) + (b - 1) = n - 1,$$

其中 $n = \sum_{l=1}^b k_l$ 表示 D 的试验次数.

考虑一个可加的、固定的主效应模型 (可参见文献 [1, 6, 11]), 且带有独立的同分布误差, 误差方差记为 σ^2 . 另记 $\mathbf{1}_n$ 表示元素全为 1 的列向量. $n \times p_s$ 阶矩阵 $\mathbf{X}_s$ 表示因子 F_s 的关联矩阵, 如果第 u 次试验时, 因子 F_s 的水平为 i , 则关联矩阵的第 (u, i) 个元素为 1, 否则为 0, s 的取值为 $0, 1, \dots, m$. 于是, 试验的响应向量 $\mathbf{Y} = (Y_1, \dots, Y_n)^T$ 可以表示为

$$\mathbf{Y} = \mathbf{1}_n \mu + \mathbf{X}_0 \boldsymbol{\alpha} + \mathbf{X}_1 \boldsymbol{\beta}_1 + \dots + \mathbf{X}_m \boldsymbol{\beta}_m + \boldsymbol{\epsilon}, \quad (2.1)$$

其中 μ 表示为总均值, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_b)^T$ 表示区组效应, $\boldsymbol{\beta}_s = (\beta_{s,1}, \dots, \beta_{s,p_s})^T$ 表示处理因子 F_s 的固定效应, $s = 1, 2, \dots, m$, 试验误差为 $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)^T$. 全文中上标 T 表示向量或者矩阵的转置.

对于两个因子的情形, 记 $p_s \times p_t$ 阶矩阵 $\mathbf{M}_{s,t}$ 表示处理因子 F_s 相对于处理因子 F_t 的关联矩阵, 其中 $s, t = 0, 1, \dots, m$. 类似地, 关联矩阵的第 (i, j) 个元素表示因子 F_s 在 i 水平, 同时因子 F_t 在 j 水平时的试验次数. 于是, 矩阵 $\mathbf{M}_{s,0}$ 表示处理因子 F_s 相对于区组因子 F_0 的关联矩阵. 进一步, 记 $\mathbf{r}_s$ 表示因子 F_s 的重复向量, 即由不同水平的重复次数组成. 于是, 根据关联矩阵的定义可知, $\mathbf{r}_i = \mathbf{X}_i^T \mathbf{1}_n$.

对于关联矩阵 $M_{s,t}$ 与 X_s 和 X_t 之间的关系, 有

$$M_{s,t} = X_s^T X_t, \quad s, t = 0, 1, \dots, m.$$

对于处理因子 F_s 和 F_t , 关联矩阵 $M_{s,t}$ 与因子水平的重复向量 r_s 和 r_t 之间的关系确定了它们之间的正交性情况. 在一个主效应设计中, 称因子 F_s 和 F_t 为正交的, 如果它们满足文献 [9] 提出的等比例频数条件, 即对于任意的水平 $i = 0, 1, \dots, p_s - 1$ 和 $j = 0, 1, \dots, p_t - 1$, 因子 F_s 在 i 水平、因子 F_t 在 j 水平时的试验次数是 i 水平的重复次数与 j 水平的重复次数乘积的倍数, 即

$$M_{s,t} = \frac{1}{n} r_s r_t^T.$$

于是称一个主效应设计 D 是正交的, 如果任意两个因子是正交的. 相反地, 如果两个处理因子 F_s 和 F_t 是不正交的, 则等式 $M_{s,t} = \frac{1}{n} r_s r_t^T$ 不再成立. 记两者的差为 $E_{s,t}$, 则有

$$E_{s,t} = M_{s,t} - \frac{1}{n} r_s r_t^T,$$

它将应用在后面的证明中.

例 1 表 1 给出了一个饱和的主效应设计, 其含有 4 个三水平的处理因子, 3 个二水平的处理因子, 同时, 该设计含有 4 个区组, 区组大小分别为 $k_1 = k_2 = k_3 = 4$ 和 $k_4 = 3$. 例如, 对于其中的二水平因子 F_4 和 F_5 , 如果只考虑前三个区组, 则关联矩阵为

$$M_{4,5} = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}, \quad \frac{1}{n} r_4 r_5^T = \frac{1}{12} \begin{bmatrix} 6 \\ 6 \end{bmatrix} \begin{bmatrix} 6 & 6 \end{bmatrix}.$$

表 1 区组大小分别为 $k_1 = k_2 = k_3 = 4$ 和 $k_4 = 3$ 的主效应设计

	F_0	F_1	F_2	F_3	F_4	F_5	F_6	F_7
区组 1	1	0	0	0	0	0	0	0
	1	0	1	1	1	1	0	0
	1	1	0	1	1	0	1	0
	1	1	1	0	0	1	1	0
区组 2	2	1	1	1	0	0	0	1
	2	1	2	2	1	1	0	1
	2	2	1	2	1	0	1	1
	2	2	2	1	0	1	1	1
区组 3	3	2	2	2	0	0	0	2
	3	2	0	0	1	1	0	2
	3	0	2	0	1	0	1	2
	3	0	0	2	0	1	1	2
区组 4	4	1	2	0	1	0	1	0
	4	1	2	0	1	0	1	1
	4	1	2	0	1	0	1	2

于是, 容易验证 $\mathbf{M}_{4,5} = \frac{1}{n} \mathbf{r}_4 \mathbf{r}_5^T$, 即根据等比例频数条件, 因子 F_4 和 F_5 是正交的. 但是, 对于三水平因子 F_1 和 F_2 , 当只考虑前三个区组时, 有

$$\mathbf{M}_{1,2} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad \frac{1}{n} \mathbf{r}_1 \mathbf{r}_2^T = \frac{1}{12} \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} [4 \ 4 \ 4].$$

显然, $\mathbf{M}_{1,2} \neq \frac{1}{n} \mathbf{r}_1 \mathbf{r}_2^T$. 于是, 根据等比例频数条件因子 F_1 和 F_2 不是正交的. 然而, 第 3 节的分析表明, 它们可以通过区组因子达到正交^[11], 即所有主效应对照是可估的, 且主效应对照的 BLUE 之间是不相关的. 另外, 当考虑所有区组时, 任意两个因子之间仍可以通过区组因子达到正交.

2.2 加权相遇数

对于区组大小不等情形, 当区组设计只含有一个处理因子时, 文献 [5, 8, 13] 提出的加权相遇数在理论推导中起到了关键的作用. 这里将上述概念推广至多个处理因子情形. 记 $\mathbf{K}$ 表示对角线元素为 $(k_1, k_2, \dots, k_b)$ 的对角矩阵, 即

$$\mathbf{K} = \begin{bmatrix} k_1 & & & \\ & k_2 & & \\ & & \ddots & \\ & & & k_b \end{bmatrix},$$

另外, $r_{s,l}$ 表示在第 l 个区组内的因子 F_s 的水平重复向量.

定义 1 如果在第 l 个区组内, 因子 F_s 的 i 水平出现 a 次, 因子 F_t 的 j 水平出现 b 次, 则称有 ab 次相遇. 如果以区组大小 k_l 为权数, 则称 ab/k_l 为加权相遇数. 容易验证

$$\mathbf{M}_{s,0} \mathbf{K}^{-1} \mathbf{M}_{0,t} = \sum_{l=1}^b \frac{r_{s,l} r_{t,l}^T}{k_l}$$

为因子 F_s 和 F_t 在 b 个区组的加权相遇数之和组成的矩阵, 其中 $\mathbf{K}^{-1}$ 表示对角矩阵 $\mathbf{K}$ 的逆.

下面给出两个有关加权相遇数性质的引理.

引理 1 令 F_s 和 F_t 为主效应设计的两个因子, 设计含有 b 个大小为 $k_1, k_2, \dots, k_b$ 的区组, 则

(i) 当 $s, t = 1, 2, \dots, m$ 时, 有

$$\mathbf{E}_{s,t} - \mathbf{E}_{s,0} \mathbf{K}^{-1} \mathbf{E}_{0,t} = \mathbf{M}_{s,t} - \mathbf{M}_{s,0} \mathbf{K}^{-1} \mathbf{M}_{0,t};$$

(ii) 当 $t = 0$ 且 $s = 1, 2, \dots, m$, 有

$$\mathbf{E}_{s,0} - \mathbf{E}_{s,0} \mathbf{K}^{-1} \mathbf{E}_{0,0} = \mathbf{0}.$$

证明 (i) 根据 $\mathbf{E}_{s,t}$ 的定义, 有

$$\begin{aligned} \mathbf{E}_{s,t} - \mathbf{E}_{s,0} \mathbf{K}^{-1} \mathbf{E}_{0,t} &= \left(\mathbf{M}_{s,t} - \frac{1}{n} \mathbf{r}_s \mathbf{r}_t^T \right) - \left(\mathbf{M}_{s,0} - \frac{1}{n} \mathbf{r}_s \mathbf{r}_0^T \right) \mathbf{K}^{-1} \left(\mathbf{M}_{0,t} - \frac{1}{n} \mathbf{r}_0 \mathbf{r}_t^T \right) \\ &= (\mathbf{M}_{s,t} - \mathbf{M}_{s,0} \mathbf{K}^{-1} \mathbf{M}_{0,t}) - \frac{1}{n} \mathbf{r}_s \mathbf{r}_t^T + \frac{1}{n} \mathbf{M}_{s,0} \mathbf{K}^{-1} \mathbf{r}_0 \mathbf{r}_t^T \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{n} \mathbf{r}_s \mathbf{r}_0^T \mathbf{K}^{-1} \mathbf{M}_{0,t} - \frac{1}{n^2} \mathbf{r}_s \mathbf{r}_0^T \mathbf{K}^{-1} \mathbf{r}_0 \mathbf{r}_t^T \\
& = \mathbf{M}_{s,t} - \mathbf{M}_{s,0} \mathbf{K}^{-1} \mathbf{M}_{0,t} - \frac{1}{n} \mathbf{r}_s \mathbf{r}_t^T + \frac{1}{n} \mathbf{r}_s \mathbf{r}_t^T - \frac{1}{n} \mathbf{r}_s \mathbf{r}_t^T + \frac{1}{n} \mathbf{r}_s \mathbf{r}_t^T \\
& = \mathbf{M}_{s,t} - \mathbf{M}_{s,0} \mathbf{K}^{-1} \mathbf{M}_{0,t}.
\end{aligned}$$

(ii) 由于 $\mathbf{E}_{s,0} = \mathbf{M}_{s,0} - \frac{1}{n} \mathbf{r}_s \mathbf{r}_0^T$, 故有

$$\begin{aligned}
\mathbf{E}_{s,0} \mathbf{K}^{-1} \mathbf{E}_{0,0} & = \left(\mathbf{M}_{s,0} - \frac{1}{n} \mathbf{r}_s \mathbf{r}_0^T \right) \mathbf{K}^{-1} \left(\mathbf{M}_{0,0} - \frac{1}{n} \mathbf{r}_0 \mathbf{r}_0^T \right) \\
& = \mathbf{M}_{s,0} - \frac{1}{n} \mathbf{r}_s \mathbf{r}_0^T - \frac{1}{n} \mathbf{M}_{s,0} \mathbf{K}^{-1} \mathbf{r}_0 \mathbf{r}_0^T + \frac{1}{n^2} \mathbf{r}_s \mathbf{r}_0^T \mathbf{K}^{-1} \mathbf{r}_0 \mathbf{r}_0^T \\
& = \mathbf{E}_{s,0} - \frac{1}{n} \mathbf{r}_s \mathbf{r}_0^T + \frac{1}{n} \mathbf{r}_s \mathbf{r}_0^T \\
& = \mathbf{E}_{s,0}.
\end{aligned}$$

证毕. □

例 2 (例 1 续) 对于例 1 中的主效应设计, 当考虑四个区组时, 有

$$\begin{aligned}
\mathbf{M}_{1,2} & = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 1 & 2 \end{bmatrix}, \quad \frac{1}{n} \mathbf{r}_1 \mathbf{r}_2^T = \frac{1}{15} \begin{bmatrix} 4 \\ 7 \\ 4 \end{bmatrix} \begin{bmatrix} 4 & 4 & 7 \end{bmatrix}, \\
\mathbf{M}_{1,0} & = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 2 & 2 & 0 & 3 \\ 0 & 2 & 2 & 0 \end{bmatrix}, \quad \frac{1}{n} \mathbf{r}_1 \mathbf{r}_0^T = \frac{1}{15} \begin{bmatrix} 4 \\ 7 \\ 4 \end{bmatrix} \begin{bmatrix} 4 & 4 & 4 & 3 \end{bmatrix}.
\end{aligned}$$

于是,

$$\mathbf{E}_{1,2} = \frac{1}{15} \begin{bmatrix} 14 & -1 & -13 \\ -13 & 2 & 11 \\ -1 & -1 & 2 \end{bmatrix}, \quad \mathbf{E}_{1,0} = \frac{1}{15} \begin{bmatrix} 14 & -16 & 14 & -12 \\ 2 & 2 & -28 & 24 \\ -16 & 14 & 14 & -12 \end{bmatrix}.$$

通过计算表明,

$$\mathbf{E}_{1,2} - \mathbf{E}_{1,0} \mathbf{K}^{-1} \mathbf{E}_{0,2} = \mathbf{M}_{1,2} - \mathbf{M}_{1,0} \mathbf{K}^{-1} \mathbf{M}_{0,2},$$

其中 $\mathbf{K} = \text{diag}(4, 4, 4, 3)$. 另一方面, 有 $\mathbf{E}_{1,0} - \mathbf{E}_{1,0} \mathbf{K}^{-1} \mathbf{E}_{0,0} = \mathbf{0}$. 当 $s, t = 0, 1, \dots, m$ 时, 类似的结论同样成立.

记 $\mathbf{U}_s$ 表示因子 F_s 的原始和 (raw totals) 向量, G 表示总和, 于是有

$$G = \mathbf{1}_n^T \mathbf{Y}, \quad \mathbf{U}_s = \mathbf{X}_s^T \mathbf{Y}, \quad s = 0, 1, \dots, m.$$

另外, 记 $\mathbf{U}_s^{\text{ad}}$ 为因子 F_s 的矫正和向量, 即 $\mathbf{U}_s^{\text{ad}} = \mathbf{U}_s - \mathbf{M}_{s,0} \mathbf{K}^{-1} \mathbf{U}_0$, $s = 1, 2, \dots, m$.

引理 2 关于区组大小不等时的设计的矫正和向量, 有

$$\mathbf{U}_s^{\text{ad}} = \tilde{\mathbf{U}}_s - \mathbf{E}_{s,0} \mathbf{K}^{-1} \tilde{\mathbf{U}}_0,$$

其中 $\tilde{U}_s = U_s - \frac{G}{n} r_s$.

证明 根据 $\tilde{U}_s$ 和 $E_{s,0}$ 的定义, 有

$$\begin{aligned} \tilde{U}_s - E_{s,0} K^{-1} \tilde{U}_0 &= \tilde{U}_s - \frac{G}{n} r_s - \left(M_{s,0} - \frac{1}{n} r_s r_0^T \right) K^{-1} \left(U_0 - \frac{G}{n} r_0 \right) \\ &= \tilde{U}_s - M_{s,0} K^{-1} U_0 - \frac{G}{n} r_s + \frac{G}{n} M_{s,0} K^{-1} r_0 \\ &\quad + \frac{1}{n} r_s r_0^T K^{-1} U_0 - \frac{G}{n^2} r_s r_0^T K^{-1} r_0 \\ &= U^{\text{ad}} - \frac{G}{n} r_s + \frac{G}{n} r_s + \frac{G}{n} r_s - \frac{G}{n} r_s \\ &= U^{\text{ad}}. \end{aligned}$$

证毕. □

3 广义 POTB 设计

文献 [11] 研究了区组大小相等时的 POTB 设计. 其中, 因子 F_s 和 F_t 称为通过区组因子达到正交, 如果满足

$$M_{s,0} M_{0,t} = k M_{s,t}.$$

于是, 如果任意两个因子都能通过区组因子达到正交, 则称设计 D 是一个 POTB 设计. 在该类设计中, 任意因子的主效应对照是可估的, 且任意两个因子的主效应对照的 BLUE 之间是不相关的. 这里将其推广至区组大小不等情形, 并称相应的设计为广义 POTB 设计. 本节先给出一个主效应设计是广义 POTB 设计的充分条件. 它依赖于关联矩阵与加权相遇数之间的关系. 更多的关于广义 POTB 设计的性质见第 3.2 小节.

3.1 广义 POTB 设计的充分条件

定理 1 假定设计 D 是一个含有 b 个区组大小为 $k_1, k_2, \dots, k_b$ 的主效应设计, 此处, 区组大小可不等. 矩阵 $M_{s,t}$ 和 K 的定义与前述一致. 如果设计 D 中任意两个因子的关联矩阵满足

$$M_{s,t} = M_{s,0} K^{-1} M_{0,t},$$

则设计 D 是广义 POTB 设计.

证明 为方便起见, 我们仅给出模型 (2.1) 下含有两个处理因子 F_1 和 F_2 时的证明过程, 含有 m 个处理因子时的证明方法类似, 但符号会更复杂. 此时的正规方程为 $\mathbf{X}^T \mathbf{X} \hat{\beta} = \mathbf{X}^T \mathbf{Y}$. 记

$$\mathbf{X} = (\mathbf{1}_n \ \mathbf{X}_0 \ \mathbf{X}_1 \ \mathbf{X}_2), \quad \beta^T = (\mu \ \alpha^T \ \beta_1^T \ \beta_2^T),$$

则有

$$\begin{pmatrix} n & r_0^T & r_1^T & r_2^T \\ r_0 & M_{0,0} & M_{0,1} & M_{0,2} \\ r_1 & M_{1,0} & M_{1,1} & M_{1,2} \\ r_2 & M_{2,0} & M_{2,1} & M_{2,2} \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{\alpha} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} G \\ U_0 \\ U_1 \\ U_2 \end{pmatrix}. \quad (3.1)$$

从上式方程两边消去 $\hat{\mu}$, 可以得到下述简化正规方程:

$$\begin{pmatrix} \mathbf{E}_{0,0} & \mathbf{E}_{0,1} & \mathbf{E}_{0,2} \\ \mathbf{E}_{1,0} & \mathbf{E}_{1,1} & \mathbf{E}_{1,2} \\ \mathbf{E}_{2,0} & \mathbf{E}_{2,1} & \mathbf{E}_{2,2} \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} \tilde{U}_0 \\ \tilde{U}_1 \\ \tilde{U}_2 \end{pmatrix}. \quad (3.2)$$

根据引理 1, 为了消去 $\hat{\alpha}$, 可以在第一个方程的两边左乘矩阵 $\mathbf{E}_{1,0}\mathbf{K}^{-1}$ 并减去第二个方程, 同时在第一个方程的两边左乘矩阵 $\mathbf{E}_{2,0}\mathbf{K}^{-1}$ 并减去第三个方程. 于是有

$$\begin{pmatrix} \mathbf{E}_{1,1} - \mathbf{E}_{1,0}\mathbf{K}^{-1}\mathbf{E}_{0,1} & \mathbf{E}_{1,2} - \mathbf{E}_{1,0}\mathbf{K}^{-1}\mathbf{E}_{0,2} \\ \mathbf{E}_{2,1} - \mathbf{E}_{2,0}\mathbf{K}^{-1}\mathbf{E}_{0,1} & \mathbf{E}_{2,2} - \mathbf{E}_{2,0}\mathbf{K}^{-1}\mathbf{E}_{0,2} \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} \tilde{U}_1 - \mathbf{E}_{1,0}\mathbf{K}^{-1}\tilde{U}_0 \\ \tilde{U}_2 - \mathbf{E}_{2,0}\mathbf{K}^{-1}\tilde{U}_0 \end{pmatrix}. \quad (3.3)$$

记 $\mathbf{C}_{s,t} = \mathbf{M}_{s,t} - \mathbf{M}_{s,0}\mathbf{K}^{-1}\mathbf{M}_{0,t}$. 由引理 2, 可以进一步得到如下简化的正规方程:

$$\begin{pmatrix} \mathbf{C}_{1,1} & \mathbf{C}_{1,2} \\ \mathbf{C}_{2,1} & \mathbf{C}_{2,2} \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{U}_1^{\text{ad}} \\ \mathbf{U}_2^{\text{ad}} \end{pmatrix}. \quad (3.4)$$

于是当 $\mathbf{M}_{s,t} = \mathbf{M}_{s,0}\mathbf{K}^{-1}\mathbf{M}_{0,t}$ 时, 处理因子可通过区组因子达到正交, 定理得证. $\square$

注 1 定理 1 的结果推广了文献 [11] 关于区组大小相等时的结果. 由定理 1 可知, 当区组大小相等时, 即 $k_1 = k_2 = \cdots = k_b = k$ 时, 有

$$\mathbf{M}_{s,0}\mathbf{M}_{0,t} = k\mathbf{M}_{s,t}.$$

注 2 对于表 1 中的设计 $\mathbf{D}$, 容易验证等式 $\mathbf{M}_{s,t} = \mathbf{M}_{s,0}\mathbf{K}^{-1}\mathbf{M}_{0,t}$ 对于任意的处理因子 F_s 和 F_t 均成立. 于是, 设计 $\mathbf{D}$ 是一个广义 POTB 设计.

3.2 广义 POTB 设计的若干性质

引理 3 考虑一个含有 b 个区组的广义 POTB 设计, 其区组大小分别为 $k_1, \dots, k_b$, 且含有 m 个处理因子 $F_1, \dots, F_m$. 于是, 因子 F_s 的固定效应的最小二乘估计可从下述方程中得到:

$$\mathbf{C}_{s,s}\hat{\beta}_s = \mathbf{U}_s^{\text{ad}}.$$

注 3 从定理 1 立得引理 3, 且其同时推广了文献 [11] 的结果.

定理 2 考虑一个含有 b 个区组的广义 POTB 设计, 其区组大小分别为 $k_1, \dots, k_b$, 且含有 m 个处理因子 $F_1, \dots, F_m$. 假定试验结果是独立的, 且带有方差为 σ^2 的试验误差, 于是下述结论成立:

- (i) 每个因子的主效应对照是可估的;
- (ii) 因子 F_s 的主效应对照 $\mathbf{l}^T\beta_s$ 的 BLUE 为 $\mathbf{l}^T\mathbf{C}_{s,s}^{-}\mathbf{U}_s^{\text{ad}}$, 其中 $\mathbf{C}^{-}$ 表示 $\mathbf{C}$ 的广义逆;
- (iii) 上述 BLUE 的方差为 $\mathbf{l}^T\mathbf{C}_{s,s}^{-}\mathbf{l}\sigma^2$;
- (iv) 对于任意的因子 F_s 和 F_t , 相应主效应对照的 BLUE 之间是不相关的.

证明 我们仅给出最后一个结论的证明过程, 其他三个可用文献 [11] 中的方法类似得到. 记

$$\mathbf{W}_s = \mathbf{X}_s^T\mathbf{H}, \quad \mathbf{W}_t = \mathbf{X}_t^T\mathbf{H}, \quad (3.5)$$

其中

$$H = I - X_0 K^{-1} X_0^T.$$

于是矩阵 $C_{s,t}$ 和 U_s^{ad} 可以表示为

$$C_{s,t} = W_s W_t^T, \quad U_s^{\text{ad}} = W_s Y. \quad (3.6)$$

因此, 对于主效应对照 $l_1^T \beta_s$ 和 $l_2^T \beta_t$ 的 BLUE, 有

$$l_1^T \hat{\beta}_1 = l_1^T (W_s W_s^T)^{-1} W_s Y, \quad l_2^T \hat{\beta}_2 = l_2^T (W_t W_t^T)^{-1} W_t Y. \quad (3.7)$$

当 $C_{s,t} = 0$ 时, 有 $W_s W_t^T = 0$, 于是, 其协方差矩阵 $\text{Cov}(l_1^T \hat{\beta}_s, l_2^T \hat{\beta}_t) = 0$. 证毕. $\square$

定义下列平方和:

$$\begin{aligned} SS^{\text{tot}} &= Y^T Y - \frac{G^2}{n}, \\ SS_0 &= U_0^T K^{-1} U_0 - \frac{G^2}{n}, \\ SS_t^{\text{adj}} &= (U_t^{\text{ad}})^T \hat{\beta}_t. \end{aligned}$$

引理 4 考虑一个含有 b 个区组的广义 POTB 设计, 其区组大小分别为 $k_1, \dots, k_b$, 且含有 2 个处理因子 F_1 和 F_2 . 于是上述定义的平方和满足下列等式:

$$U_1^T \hat{\beta}_1 + U_2^T \hat{\beta}_2 + U_0^T \hat{\alpha} = SS_1^{\text{adj}} + SS_2^{\text{adj}} + SS_0. \quad (3.8)$$

证明 由于正规方程可以用 (3.1) 中的四个方程来表示, 于是进一步详细给出其中一个的结果,

$$\hat{\mu} r_0 + M_{0,1} \hat{\beta}_1 + M_{0,2} \hat{\beta}_2 + M_{0,0} \hat{\alpha} = U_0. \quad (3.9)$$

根据因子 F_1 和 F_2 的原始和向量、矫正和向量的关系, 有

$$U_i = U_i^{\text{ad}} + M_{i,0} K^{-1} U_0, \quad i = 1, 2.$$

将上式转置, 然后乘以相应的参数估计, 并运用方程 (3.9) 进行等量代换, 可得 $U_1^T \hat{\beta}_1$ 和 $U_2^T \hat{\beta}_2$ 分别可以表示为

$$U_1^T \hat{\beta}_1 = (U_1^{\text{ad}})^T \hat{\beta}_1 + U_0^T K^{-1} (U_0 - \hat{\mu} r_0 - M_{0,2} \hat{\beta}_2 - M_{0,0} \hat{\alpha})$$

和

$$U_2^T \hat{\beta}_2 = (U_2^{\text{ad}})^T \hat{\beta}_2 + U_0^T K^{-1} (U_0 - \hat{\mu} r_0 - M_{0,1} \hat{\beta}_1 - M_{0,0} \hat{\alpha}).$$

将上述两个方程相加即得结论. $\square$

根据定理 2 和引理 4 的结论, 我们可以对文献 [11] 中关于区组大小相等时 POTB 设计方差分析的结论进行推广, 有如下定理.

定理 3 对于含有两个处理因子的广义 POTB 设计, 其试验结果可进行如下的方差分析, 见表 2.

注 4 上述定理的证明技巧与文献 [11] 中关于区组大小相等时的情形基本一致, 故这里省略. 另外, 引理 4 和定理 3 仅给出了两因子的结果, 多因子情形的结果可类似得到.

表 2 两因子广义 POTB 设计的方差分析表

来源	自由度	平方和
F_1 (矫正)	$p_1 - 1$	$(\mathbf{U}_1^{\text{ad}})^{\text{T}} \mathbf{C}_{1,1}^{-} \mathbf{U}_1^{\text{ad}}$
F_2 (矫正)	$p_2 - 1$	$(\mathbf{U}_2^{\text{ad}})^{\text{T}} \mathbf{C}_{2,2}^{-} \mathbf{U}_2^{\text{ad}}$
F_0	$b - 1$	$\mathbf{U}_0^{\text{T}} \mathbf{K}^{-1} \mathbf{U}_0 - G^2/n$
误差	差	差
总和	$n - 1$	$\mathbf{Y}^{\text{T}} \mathbf{Y} - G^2/n$

4 广义 POTB 设计的构造

本节记一个 n 行 b 个区组 (区组大小分别为 $k_1, k_2, \dots, k_b$) m 个处理因子 $p_1, p_2, \dots, p_m$ 的广义 POTB 设计为 $\rho_n(b; p_1 p_2 \cdots p_m)$. 因此, 当设计含有 m_i 个 p_i 水平处理因子时, 可记为 $\rho_n(b; p_1^{m_1} \cdots p_s^{m_s})$. 另一方面, 一个广义 POTB 设计 $\rho_n(b; p_1 p_2 \cdots p_m)$ 也可以用一个 $n \times (m+1)$ 阶矩阵

$$\mathbf{D} = [\mathbf{F}_0 \ \mathbf{F}_1 \ \mathbf{F}_2 \ \cdots \ \mathbf{F}_m]$$

表示, 其中第 1 个列向量 $\mathbf{F}_0$ 表示 b 水平区组因子, 后续列向量 $\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_m$ 表示处理因子 $F_1, F_2, \dots, F_m$. 在不至混淆的情形下, 我们采用记号 $\mathbf{F}_i$ 同时表示因子 F_i 和它对应的设计列. 因此, 如果 $\mathbf{F}_0$ 的第 i 行为 l 水平, 则表示该试验在第 l 个区组里, 其中 $1 \leq i \leq n, 1 \leq l \leq b$.

下面给出几个广义 POTB 设计的循环构造方法.

构造 1 (增加单元区组的构造法)

第 1 步 给定参数组合 $b_1, p_1, \dots, p_{m_1}$, 确定广义 POTB 设计 $\mathbf{D}_1 = [\mathbf{F}_0^1 \ \mathbf{F}_1^1 \ \mathbf{F}_2^1 \ \cdots \ \mathbf{F}_{m_1}^1]$, 区组大小分别为 $k_1, k_2, \dots, k_{b_1}$.

第 2 步 增加一个新的区组, 试验组合为 $\{p'_1 \cdots p'_m\}$, 其中 p'_i 为一个整数满足 $p'_i = i + b + 1 \pmod{p_i}$, $i = 1, 2, \dots, m$, 其中 “mod y ” 表示对 y 取余.

第 3 步 叠加上述区组, 即有

$$\mathbf{D}_2 = [\mathbf{F}_0^2 \ \mathbf{F}_1^2 \ \cdots \ \mathbf{F}_m^2] = \begin{bmatrix} \mathbf{F}_0^1 & \mathbf{F}_1^1 & \cdots & \mathbf{F}_m^1 \\ b+1 & p'_1 & \cdots & p'_m \end{bmatrix}. \quad (4.1)$$

定理 4 (4.1) 中的设计 $\mathbf{D}_2$ 是一个 $\rho_{n+1}(b+1; p_1 \cdots p_m)$, 其含有 $b+1$ 个区组, 分别为 $k_1, \dots, k_{b+1}$, 且 $k_{b+1} = 1$.

证明 根据关联矩阵 $\mathbf{M}_{s,t}$ 的定义, 有

$$\mathbf{M}_{s,t} = \sum_{l=1}^{b+1} \mathbf{M}_{s,t,l}, \quad \mathbf{M}_{s,0} \mathbf{K}^{-1} \mathbf{M}_{0,t} = \sum_{l=1}^{b+1} \frac{\mathbf{r}_{s,l} \mathbf{r}_{t,l}^{\text{T}}}{k_l}. \quad (4.2)$$

于是对于设计 $\mathbf{D}_2$ 的任意两个处理因子 F_s 和 F_t , 可得

$$\begin{aligned} \mathbf{M}_{s,t} &= \sum_{l=1}^{b+1} \mathbf{M}_{s,t,l} = \sum_{l=1}^b \mathbf{M}_{s,t,l} + \mathbf{M}_{s,t,(b+1)} \\ &= \sum_{l=1}^b \frac{\mathbf{r}_{s,l} \mathbf{r}_{t,l}^{\text{T}}}{k_l} + \frac{\mathbf{r}_{s,(b+1)} \mathbf{r}_{t,(b+1)}^{\text{T}}}{k_{b+1}} \end{aligned}$$

$$= \sum_{l=1}^{b+1} \frac{\mathbf{r}_{s,l} \mathbf{r}_{t,l}^T}{k_l},$$

其中 $k_{b+1} = 1$. 上述第 3 个等号右端第 1 项

$$\sum_{l=1}^b M_{s,t,l} = \sum_{l=1}^b \frac{\mathbf{r}_{s,l} \mathbf{r}_{t,l}^T}{k_l}$$

成立是基于 $\mathbf{D}_1$ 是一个广义 POTB 设计的事实. 根据 $\mathbf{r}_{s,(b+1)}, \mathbf{r}_{t,(b+1)}$ 和 $\mathbf{M}_{s,t,(b+1)}$ 的定义知, 第 2 项

$$\mathbf{M}_{s,t,(b+1)} = \frac{\mathbf{r}_{s,(b+1)} \mathbf{r}_{t,(b+1)}^T}{k_{b+1}}$$

成立. 事实上, 在第 l 个区组内, 如果有因子 F_s 为 i 且因子 F_t 为 j 水平的处理组合, 则 $\mathbf{M}_{s,t,(b+1)}$ 的第 (i, j) 个元素是 1, 否则取 0. 根据等式 (4.2) 和定理 1, 结论得证. $\square$

例 3 应用构造方法 1 可容易构造一个广义 POTB 设计 $\rho_5(2; 2^3)$ 如下. 记 $\mathbf{D}_1 = [\mathbf{1}_4 \ \mathbf{F}_1 \ \mathbf{F}_2 \ \mathbf{F}_3]$ 为一个 4 阶的 Hadamard 矩阵, 于是根据定理 4, 设计

$$\mathbf{D}_2 = \begin{bmatrix} \mathbf{1}_4 & \mathbf{F}_1 & \mathbf{F}_2 & \mathbf{F}_3 \\ 2 & 1 & 0 & 1 \end{bmatrix}$$

是一个 $\rho_5(2; 2^3)$. 另外, 易见 $\mathbf{D}_2$ 是饱和的.

注 5 在构造 1 中, 如果设计 $\mathbf{D}_1$ 是饱和的, 则设计 $\mathbf{D}_2$ 也是饱和的.

注 6 由于构造 1 中第 $b+1$ 个区组的处理水平组合是可变的, 因此设计 $\mathbf{D}_2$ 是不唯一的. 如果仅要求正交性, 则总计存在 $\prod_{i=1}^m p_i$ 种不同设计. 因此, 可以结合其他设计的优良性准则进一步优化广义 POTB 设计, 如文献 [14–21] 中的均匀性或者空间填充性等. 最近, 文献 [22] 提出了一种结合均匀性准则的最优或近似最优 POTB 设计. 显然, 该方法也可以适用于区组大小不等时的广义 POTB 设计. 但是, 随着因子数 m 的增加, 搜索算法就变得困难. 结合其他优良性准则, 构造相应的广义 POTB 设计是一个有意义的问题.

注 7 构造方法 1 可以被循环使用. 由于设计 $\mathbf{D}_2$ 是一个广义 POTB 设计, 于是一个新的广义 POTB 设计可以通过增加一个新的区组而得到. 令设计 $\mathbf{D}_2$ 是例 3 中得到的含有 2 个区组的广义 POTB 设计, 其中 $k_1 = 4, k_2 = 1$, 于是设计

$$\mathbf{D}_3 = \begin{bmatrix} \mathbf{1}_4 & a_1 & a_2 & a_3 \\ 2 & 1 & 0 & 1 \\ 3 & 0 & 1 & 0 \end{bmatrix}$$

也是一个广义 POTB 设计, 其含有三个区组, 大小分别为 $k_1 = 4, k_2 = 1, k_3 = 1$.

构造 2 (叠加不同区组大小的广义 POTB 设计)

第 1 步 对于给定的参数组合 $b_1, p_1, \dots, p_{m_1}$ 和 $b_2, p_{m_1+1}, \dots, p_{m_1+m_2}$, 确定两个广义 POTB 设计

$$\mathbf{D}_1 = [\mathbf{F}_0^1 \ \mathbf{F}_1^1 \ \mathbf{F}_2^1 \ \cdots \ \mathbf{F}_{m_1}^1] \quad \text{和} \quad \mathbf{D}_2 = [\mathbf{F}_0^2 \ \mathbf{F}_1^2 \ \mathbf{F}_2^2 \ \cdots \ \mathbf{F}_{m_2}^2],$$

其区组大小分别为 $k_1, \dots, k_{b_1}$ 和 $k_{b_1+1}, \dots, k_{b_1+b_2}$.

第 2 步 将设计 D_1 扩张到

$$D_1^T = \begin{bmatrix} F_0^1 & F_1^1 & \cdots & F_{m_1}^1 \\ F_0^2 \oplus b_1 & g_1 & \cdots & g_{m_1} \end{bmatrix},$$

其中 $g_i = F_0^2 \oplus i \pmod{p_i}$, $i = 1, \dots, m_1$, 这里, 符号 $F_0^2 \oplus b_1$ 表示向量 F_0^2 的每个元素均加上 b_1 .

第 3 步 删去设计 D_2 的第 1 列, 然后对设计 D_2 的剩余列进行类似地扩张. 令

$$D_2^T = \begin{bmatrix} g_{m_1+1} & \cdots & g_{m_1+m_2} \\ F_1^2 & \cdots & F_{m_2}^2 \end{bmatrix},$$

其中 $g_i = F_0^1 \oplus i \pmod{p_i}$, $i = m_1 + 1, \dots, m_1 + m_2$.

第 4 步 叠加上述两个设计, 可得 $D_3 = [D_1^T \ D_2^T]$.

定理 5 构造方法 2 中的设计 D_3 是一个 $\rho_{n_1+n_2}(b_1+b_2; p_1 \cdots p_{m_1} p_{m_1+1} \cdots p_{m_1+m_2})$, 其区组大小分别为 $k_1, k_2, \dots, k_{b_1+b_2}$.

证明 我们首先考虑当 $s, t = 1, \dots, m_1$ 时的处理因子 F_s 和 F_t . 将所有 $b_1 + b_2$ 个区组分成两部分: 第 1 部分包含前面 b_1 个区组, 第 2 部分包含剩余的 b_2 个区组, 有

$$\begin{aligned} M_{s,t} &= \sum_{l=1}^{b_1+b_2} M_{s,t,l} = \sum_{l=1}^{b_1} M_{s,t,l} + \sum_{l=b_1+1}^{b_1+b_2} M_{s,t,(l)} \\ &= \sum_{l=1}^{b_1} \frac{r_{s,l} r_{t,l}^T}{k_l} + \sum_{l=b_1+1}^{b_1+b_2} \frac{r_{s,l} r_{t,l}^T}{k_l} \\ &= \sum_{l=1}^{b_1} M_{s,t,l} + \sum_{l=b_1+1}^{b_1+b_2} M_{s,t,l} \\ &= M_{s,0} K^{-1} M_{0,t}, \end{aligned}$$

其中

$$K = \text{diag}(k_1, \dots, k_{b_1}, k_{b_1+1}, \dots, k_{b_1+b_2}).$$

由于 $D_1 = [F_0^1 \ F_1^1 \ F_2^1 \ \cdots \ F_{m_1}^1]$ 是一个广义 POTB 设计, 所以有

$$\sum_{l=1}^{b_1} M_{s,t,l} = \sum_{l=1}^{b_1} \frac{r_{s,l} r_{t,l}^T}{k_l}.$$

当 $l = b_1 + 1, \dots, b_1 + b_2$ 时, $M_{s,t,l}$ 或者 $\frac{r_{s,l} r_{t,l}^T}{k_l}$ 的第 (x, y) 个元素为

$$\begin{cases} k_l, & \text{如果 } g_s = x \cdot \mathbf{1}_{n_2}, \quad g_t = y \cdot \mathbf{1}_{n_2}, \\ 0, & \text{其他,} \end{cases}$$

其中 $0 \leq x \leq p_s - 1$, $0 \leq y \leq p_t - 1$. 当 $s, t = m_1 + 1, \dots, m_1 + m_2$ 时, 可得类似结论. 当 $s = 1, \dots, m_1$, $t = m_1 + 1, \dots, m_1 + m_2$ 时, 容易验证对于每个区组, 有 $M_{s,t,l} = \frac{r_{s,l} r_{t,l}^T}{k_l}$. 因此有

$$\sum_{l=1}^{b_1+b_2} M_{s,t,l} = \sum_{l=1}^{b_1+b_2} \frac{r_{s,l} r_{t,l}^T}{k_l}.$$

结论得证. $\square$

例 4 记设计 D_1 和 D_2 分别表示文献 [11] 中的两个 POTB 设计 $\rho_{12}(3; 2^3 3^2)$ 和 $\rho_{12}(6; 4^2)$, 其中第 1 个设计的区组大小为 4, 第 2 个设计的区组大小为 2. 于是, 根据构造 2 可得一个新的广义 POTB 设计 $\rho_{24}(9; 2^3 3^2 4^2)$, 其含有 9 个区组, 大小分别为 $k_1 = k_2 = k_3 = 4$ 和 $k_4 = \cdots = k_9 = 2$.

注 8 如果设计 D_1 和 D_2 是饱和的, 则在构造 2 中设计 D_3 也是饱和的.

注 9 在构造 2 中, 令 $D_2 = (0, 1, \dots, p_{m+1} - 1)^T$, 且 $k_{b+1} = p_{m+1}$. 于是可以得到一个新的广义 POTB 设计 $\rho_{n+k}(b+1; p_1 \cdots p_m p_{m+1})$, 其含有 $b+1$ 个区组, 大小分别为 $k_1, \dots, k_{b+1}$, 如表 1 中的广义 POTB 设计 $\rho_{15}(4; 3^4 2^3)$.

注 10 类似于构造 1, 可以在构造 2 中进一步采用搜索算法来获得具有其他优良性质的广义 POTB 设计, 如均匀性或者空间填充性.

利用构造 1 和 2 可得广义 POTB 设计, 我们将部分试验次数 $n \leq 20$ 的设计列在表 3 中. 表 3 中依次给出了试验次数 n , 区组数 b , 区组大小 $k_1, k_2, \dots, k_b$ 和因子的水平数 $p_1^{m_1} \cdots p_s^{m_s}$. 最后一列中的记号 “B” 用来表明此处相应的设计可由文献 [11] 构造得到, “A1” 和 “A2” 表示相应的设计可以分别由前述构造 1 和 2 得到. 注意到, 构造 1 和 2 可得到很多广义 POTB 设计, 为了节省篇幅, 此处仅给出一小部分设计.

表 3 部分试验次数 $n \leq 20$ 的广义 POTB 设计

n	b	$k_1, k_2, \dots, k_b$	$p_1^{m_1} \cdots p_s^{m_s}$	方法
5	2	4, 1	2^3	A1
6	2	4, 2	2^4	A2
7	2	4, 3	$2^3 \cdot 3^1$	A2
8	2	4, 4	$2^3 \cdot 4^1$	A2
9	2	4, 5	$2^3 \cdot 5^1$	A2
10	3	4, 3, 3	$2^3 \cdot 3^2$	A2
10	2	8, 2	2^8	A2
11	3	4, 3, 4	$2^3 \cdot 3^1 \cdot 4^1$	A2
12	2	9, 3	3^5	A2
12	3	4, 4, 4	$3^3 \cdot 2^3$	B
13	4	4, 4, 4, 1	$3^3 \cdot 2^3$	A1
12	6	2, 2, 2, 2, 2, 2	4^2	B
13	7	2, 2, 2, 2, 2, 2, 1	4^2	A1
24	9	4, 4, 4, 2, 2, 2, 2, 2, 2	$3^3 \cdot 2^3 \cdot 4^2$	A2
15	4	4, 4, 4, 3	$3^4 \cdot 2^3$	A2
16	5	4, 4, 4, 3, 1	$3^4 \cdot 2^3$	A1, A2
18	5	4, 4, 4, 3, 3	$3^5 \cdot 2^3$	A2
19	6	4, 4, 4, 3, 3, 1	$3^5 \cdot 2^3$	A1, A2
14	7	2, 2, 2, 2, 2, 2, 2	$4^2 \cdot 2^1$	A2
16	4	4, 4, 4, 4	$4^1 \cdot 3^3 \cdot 2^3$	A2
16	4	4, 4, 4, 4	$3^3 \cdot 2^6$	A2
20	5	4, 4, 4, 4, 4	$4^1 \cdot 3^3 \cdot 2^6$	A2

5 结论

本文主要讨论了区组大小不等时的主效应设计. 通过区组大小不等情形下的加权相遇数, 我们给出了主效应设计成为一个广义 POTB 设计的充分条件. 其推广了文献 [11] 中关于区组大小相等时的情形, 其中处理因子可以通过区组因子达到正交. 另外, 本文也给出了该类设计的若干性质和方差分析. 最后, 本文给出了两个构造方法, 均比较容易实现, 且得到了一批广义 POTB 设计.

另外, 从构造过程可见, 对于给定的参数 n 、 b 、 p_s 和 k_l , 目前提出的方法所得广义 POTB 设计是不唯一的. 于是通过结合其他优良性 (参见文献 [15, 22, 23]) 来构造广义 POTB 设计是一个很有意义的问题. 正如一位审稿专家建议, 设计的 A 、 D 和 E 最优是后续可以研究的问题. 这里仅对 A 最优作部分讨论, 即考虑效应 β 估计量的方差. 假定设计 D 是一个可行设计, $\hat{\Theta}_t$ 是模型 (2.1) 中参数 $\Theta_t = \mathbf{X}_t\beta_t$ 的最佳线性无偏估计, 其中 $t = 1, \dots, m$, 于是可以证明

$$\text{tr}(\text{Var}(\hat{\Theta}_t)) \geq \sigma^2(p_t - 1).$$

其中关键问题在于如何确定广义 POTB 设计, 使得协方差矩阵 $\text{Var}(\hat{\Theta}_t)$ 的迹达到其下界. 这将成为我们后续研究的问题.

参考文献

- 1 Dey A. Incomplete Block Designs. Singapore: World Scientific, 2010
- 2 Hinkelmann K, Kempthorne O. Design and Analysis of Experiments, vol. 2. New York: Wiley, 2005
- 3 Mukerjee R, Dey A, Chatterjee K. Optimal main effect plans with non-orthogonal blocking. *Biometrika*, 2002, 89: 225–229
- 4 Liu M Q, Fang K T. Some results on resolvable incomplete block designs. *Sci China Ser A*, 2005, 48: 503–512
- 5 Pearce S C. Experimenting with blocks of natural size. *Biometrics*, 1964, 20: 699–706
- 6 Patterson H D, Thompson R. Recovery of inter-block information when block sizes are unequal. *Biometrika*, 1971, 58: 545–554
- 7 Kulshreshtha A C, Dey A, Saha G M. Balanced designs with unequal replications and unequal sizes. *Ann Math Statist*, 1972, 43: 1342–1345
- 8 Gupta S C, Jones B. Equireplicate balanced block designs with unequal block sizes. *Biometrika*, 1983, 70: 433–440
- 9 Addelman S. Orthogonal main effect plans for asymmetrical factorial experiments. *Technometrics*, 1962, 4: 21–46
- 10 Bose M, Bagchi S. Optimal main effect plans in blocks of small size. *Statist Probab Lett*, 2007, 77: 142–147
- 11 Bagchi S. Main-effect plans orthogonal through the block factor. *Technometrics*, 2010, 52: 243–249
- 12 Chen X P, Lin J G, Yang J F, et al. Construction of main-effect plans orthogonal through the block factor. *Statist Probab Lett*, 2015, 106: 58–64
- 13 Pearce S C. Concurrence and quasi-replication: An alternative approach to precision in designed experiments. *Biom J*, 1976, 18: 106–116
- 14 Hickernell F J. Goodness-of-fit statistics, discrepancies and robust designs. *Statist Probab Lett*, 1999, 44: 73–78
- 15 Hickernell F J, Liu M Q. Uniform designs limit aliasing. *Biometrika*, 2002, 89: 893–904
- 16 Jiang B C, Ai M Y. Construction of uniform designs without replications. *J Complexity*, 2014, 30: 98–110
- 17 Zhou Y D, Fang K T, Ning J H. Constructing uniform designs: A heuristic integer programming method. *J Complexity*, 2012, 28: 224–237
- 18 Tang Y, Xu H, Lin D K J. Uniform fractional factorial designs. *Ann Statist*, 2012, 40: 891–907
- 19 Xu G, Zhang J, Tang Y. Level permutation method for constructing uniform designs under the wrap-around L_2 -discrepancy. *J Complexity*, 2014, 30: 46–53
- 20 Sun F S, Yin Y H, Liu M Q. Construction of nested space-filling designs using difference matrices. *J Statist Plann Inference*, 2013, 143: 160–166
- 21 Sun F S, Liu M Q, Qian P Z G. On the construction of nested space-filling designs. *Ann Statist*, 2014, 42: 1394–1425
- 22 Chen X P, Lin J G, Huang X F. Construction of main effects plans orthogonal through the block factor based on level permutation. *J Korean Statist Soc*, 2015, 44: 538–545

Main-effect plans with blocks of natural size

CHEN XuePing, LIN JinGuan, HUANG XingFang & WANG HongXia

Abstract Block designs have been widely used in almost all areas of human investigations. It is commonly assumed that all blocks have the same size which may not, however, always be available because of practical restrictions. In this paper, the problem of main-effect plans with blocks of unequal sizes is addressed. The sufficient condition for each pair of factors to be orthogonal through the block factor is indicated. Useful properties on ANOVA (analysis of variance) of such designs are discussed. Construction procedures via augmenting new blocks and combining main-effect plans with unequal block sizes are also provided. The resulting designs have flexible sizes of blocks and factors.

Keywords incidence matrix, orthogonal, proportional frequency condition, weighted concurrence

MSC(2010) 62K10, 05B05

doi: 10.1360/SCM-2016-0240