Statistics and Application 统计学与应用, 2022, 11(3), 660-668
Published Online June 2022 in Hans. http://www.hanspub.org/journal/sa
https://doi.org/10.12677/sa.2022.113071

Hans汉斯

拉丁方设计的统计量分布证明过程教学设计

唐安民

云南大学，统计系，云南 昆明

收稿日期：2022年5月29日；录用日期：2022年6月20日；发布日期：2022年6月29日

摘 要

拉丁方设计是《试验设计》析因试验部分实施的重要方法，也包含了方差分析内容，以往的关于拉丁方设计检验统计量的证明较为繁杂，大多教材只给出其结论。本文基于自定义的水平均值索引矩阵及其性质，证明拉丁方设计的检验统计量的分布，通俗易懂，并且容易推广到其他的方差分析模型。

关键词

试验设计，方差分析，拉丁方设计，索引矩阵，统计量

Teaching Design of Proving of Statistics Distribution for Latin Design

Anmin Tang

Department of Statistics, Yunnan University, Kunming Yunnan

Received: May 29^{th}, 2022; accepted: Jun. 20^{th}, 2022; published: Jun. 29^{th}, 2022

Abstract

Latin square design is an important method for analysis of variance (ANOVA) in the

文章引用: 唐安民. 拉丁方设计的统计量分布证明过程教学设计[J]. 统计学与应用, 2022, 11(3): 660-668.
DOI: 10.12677/sa.2022.113071

course of experimental design. The previous proof of test statistics for Latin design is difficult for most students to understand, therefore many textbooks only give some conclusions about Latin design. In this paper, based on the defined horizontal mean index matrix and its properties, the proving process about test statistics for Latin design is easy to understand, and can be easily extended to other models for analysis of variance.

Keywords

Experience Design, Analysis of Variance, Latin Design, Index Matrix, Test Statistics

1. 引言

在试验设计中，当因子个数与诸因子水平数较多时，即使每个水平组合只做一次试验，总试验次数也可能多到实际上难以承受的地步。这就提出了一个问题：设计试验时如何能做到减少试验次数，只做全部水平组合中的一部分，又能达到试验目的（能比较诸因子的重要性，又能比较诸水平与诸水平组合的优劣），这就是全因子试验（析因试验）的部分实施，正交设计是不考虑交互作用时一种重要的部分实施方法，但以往关于其检验统计量的证明相对繁琐和难懂 [1]，所以大多比较经典的试验设计教程 [2,3]只给出其结论。本文基于自定义的水平均值索引矩阵及相关性质，可通俗易懂地证明拉丁方设计的检验统计量的分布。

一般地，在由p个拉丁字母（或p个不同符号）排成的p行p列的方阵中，若每行的p个元素不同，每列的p个元素也不同，即每个字母在每行每列出现一次，也只出现一次，则称这个方阵为一个p阶拉丁方。每个p阶拉丁方可以用来安排三因子试验其中两个因子分别当作行因子和列因子，另一个因子当作拉丁因子，其水平用拉丁字母表示，每个因子都是p个水平，总共包括p^2次试验，用拉丁方作出的试验设计称为拉丁方设计。在不考虑交互效应时，拉丁方设计可以做到三个因子的全因子试验的部分实施，从全因子试验所需要的p^3 次试验挑选其中p^2 次试验，确保了在任一因子的任一水平下，其他因子的每个水平重复相同次数，即任意两个因子是“正交”，从而能达到试验目的（能比较诸因子水平的重要性），并减少试验次数。正因为这些优良性质，拉丁方设计及它的推广（希腊拉丁方设计和超方设计）在实践中有着广泛的应用。然而，对于这些设计中检验统计量分布，大多教材只给出其结论，或给出的证明繁杂难以理解，本文基于定义的水平均值索引矩阵及其性质，证明这些检验统计量的分布，简单直观，并且容易推广到其他的方差分析模型。我们以下面

的拉丁方模型的检验统计量分布的证明来说明我们方法的有效性。

定义 1.1 *[p阶拉丁方设计的统计模型]*

$$\begin{cases} y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \varepsilon_{ijk}, \\ \quad i = 1, \dots, p, j = 1, \dots, p, k = 1, \dots, p, \\ \text{诸}\varepsilon_{ijk} \sim N(0, \sigma^2), \\ \text{约束条件:} \sum_{i=1}^{p} \alpha_i = 0, \sum_{j=1}^{p} \tau_j = 0, \sum_{k=1}^{p} \beta_k = 0. \end{cases}$$

其中y_{ijk}是行、列因子的水平组合为(i, k)、拉丁因子取第j个水平这一试验条件下的试验观察值, 由所有观察值构成的$n(n = p^2)$维向量记作$\boldsymbol{y}$, μ 是一般平均，$\alpha_i, \tau_j, \beta_k$ 分别是行因子第i个水平、拉丁因子第j个水平和列因子第k个水平的效应，ε_{ijk} 是随机误差，这是一个不考虑交互效应的可加线性统计模型. 需要说明，i、j与k这三个下标不是独立，给定i 与k，拉丁因子的水平水平编号j是惟一被确定. 易知：

$$\boldsymbol{y} \sim N_n(\boldsymbol{\mu}^*, \sigma^2 \boldsymbol{I}_n),$$

n维均值向量$\boldsymbol{\mu}^*$是集合$\{\mu + \alpha_i + \tau_j + \beta_k | i = 1, \dots, p, j = 1, \dots, p, k = 1, \dots, p\}$ 相应于$\boldsymbol{y}$的组合向量，$\boldsymbol{I}_n$是n阶单位矩阵。由模型的约束条件容易得：

$$\mathbf{1}_n^{\mathrm{T}} \boldsymbol{\mu}^* = n\mu. \tag{1.1}$$

为了检验这三个因子的诸水平之间有没有显著性差异，对应的是检验下述三个假设：

$$\begin{aligned} &\mathrm{H}_{01} : \alpha_i = 0, \text{对一切} i = 1, \dots, p; \\ &\mathrm{H}_{02} : \tau_j = 0, \text{对一切} j = 1, \dots, p; \\ &\mathrm{H}_{03} : \beta_k = 0, \text{对一切} k = 1, \dots, p. \end{aligned} \tag{1.2}$$

检验(1.2)式中三个假设通常构造以下三个检验统计量，并用到它们的分布如下：

$$\begin{aligned} F_{\text{行}} &= \frac{SS_{\text{行}}/(p-1)}{SS_{\mathrm{e}}/((p-1)(p-2))} \overset{\mathrm{H}_{01}}{\sim} F(p-1, (p-1)(p-2)), \\ F_{\text{拉丁}} &= \frac{SS_{\text{拉丁}}/(p-1)}{SS_{\mathrm{e}}/((p-1)(p-2))} \overset{\mathrm{H}_{02}}{\sim} F(p-1, (p-1)(p-2)), \\ F_{\text{列}} &= \frac{SS_{\text{列}}/(p-1)}{SS_{\mathrm{e}}/((p-1)(p-2))} \overset{\mathrm{H}_{03}}{\sim} F(p-1, (p-1)(p-2)), \end{aligned} \tag{1.3}$$

这里的行因子平方和$SS_{\text{行}} = \sum_{i=1}^{p} p(\overline{y}_{i..} - \overline{y}_{...})^2$, 拉丁因子平方和$SS_{\text{拉丁}} = \sum_{j=1}^{p} p(\overline{y}_{.j.} - \overline{y}_{...})^2$, 列因子平方和$SS_{\text{列}} = \sum_{k=1}^{p} p(\overline{y}_{..k} - \overline{y}_{...})^2$, 误差平方和$SS_{\mathrm{e}} = SS_{\mathrm{T}} - SS_{\text{行}} - SS_{\text{列}} - SS_{\text{拉丁}}$, 其中总平方和$SS_{\mathrm{T}} = \sum_i \sum_j \sum_k (y_{ijk} - \overline{y}_{...})^2$. 要证明式子(1.3)本个统计量的分布，要用到多元统计中常用的两个定理，这里不加证明地给出如下：

定理 1.1 (正态变量的二次型分布) 设n维随机向量$\boldsymbol{X} = (x_1, \dots, x_n)^{\mathrm{T}} \sim N_n(\boldsymbol{\nu}, \boldsymbol{I}_n)$, 这里$\boldsymbol{\nu}$是$n$维均

DOI: 10.12677/sa.2022.113071

值向量, 如果$\boldsymbol{A}$为n阶对称方阵,那么二次型

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}\sim\chi^2(r,\lambda)$$

的充分必要条件是

$$\boldsymbol{A}^2=\boldsymbol{A},\text{ 并且}r\text{是矩阵}\boldsymbol{A}\text{的秩, 非中心参数}\lambda=\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\nu}.$$

进一步地, 如果有$\lambda=0$, 那么就说$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$服从中心化$\chi^2$分布$\chi^2(r)$.

定理 1.2 (正态变量二次型的独立性) *设n维随机向量$\boldsymbol{X}=(x_1,\ldots,x_n)^{\mathrm{T}}\sim N_n(\boldsymbol{\nu},\boldsymbol{I}_n)$, 这里$\boldsymbol{\nu}$是$n$维均值向量, 如果$\boldsymbol{A}_1$和$\boldsymbol{A}_2$ 均为n 阶对称方阵,并且二次型$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{X}$和$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}_2\boldsymbol{X}$都服从$\chi^2$分布, 则$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{X}$和$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}_2\boldsymbol{X}$相互独立的充分必要条件是$\boldsymbol{A}_1\boldsymbol{A}_2=\boldsymbol{0}$.*

2. 水平索引向量和索引矩阵

本章构造水平索引向量和水平均值索引矩阵，并给出它们的性质和包含于式子(1.3)里的平方和的二次型表达式。行因子第$i(i=1,\ldots,p)$ 个水平索引向量$\boldsymbol{r}_i$可构造如下：

$$\boldsymbol{r}_i=(\underbrace{0,\ldots,0}_{p\text{个}0},\ldots,\underbrace{0,\ldots,0}_{p\text{个}0},\underbrace{1,\ldots,1}_{\text{第}i\text{个分块}p\text{个}1},\underbrace{0,\ldots,0}_{p\text{个}0},\ldots,\underbrace{0,\ldots,0}_{p\text{个}0})^{\mathrm{T}}=\boldsymbol{e}_i\otimes\boldsymbol{1}_p,$$

显然n维向量$\boldsymbol{r}_i$可以分成p个分块, 第i个分块是元素全为1的p维向量$\boldsymbol{1}_p$, 其余的分块全为p维0向量，且有行因子第i 个水平观察值的和$y_{i\cdot\cdot}=\frac{1}{p}\boldsymbol{r}_i^{\mathrm{T}}\boldsymbol{y}$, 也可构造列因子第$k(k=1,\ldots,p)$ 个水平索引向量$\boldsymbol{c}_j$如下：

$$\boldsymbol{c}_k=(\underbrace{0,\ldots,0,1,0,\ldots,0}_{\text{第}k\text{个元素}1},\ldots,\underbrace{0,\ldots,0,1,0,\ldots,0}_{\text{第}k\text{个元素}1})^{\mathrm{T}}=\boldsymbol{1}_p\otimes\boldsymbol{e}_k,$$

显然n维向量$\boldsymbol{c}_k$可以分成p个分块, 每个分块都是第k个元素为1其余元素均为0的p维向量, 且有列因子第k个水平观察值的和$y_{\cdot\cdot k}=\boldsymbol{c}_k^{\mathrm{T}}\boldsymbol{y}$. 由于拉丁方设计的不惟一性，拉丁因子水平索引向量也是不惟一的, 但由拉丁方设计的正交性可以得拉丁因子水平索引向量一般可以构造如下：

定义 2.1 *[水平索引向量和矩阵] 对于拉丁因子, 构造第$j(j=1,\ldots,p)$个水平均值的索引向量$\boldsymbol{l}_j=(l_{j1},\ldots,l_{jn})^{\mathrm{T}}$, 其中$l_{jm}=1$, 当观测值向量$\boldsymbol{y}$ 的第m个观察值是拉丁因子第j个水平的观测值，否则就是0. 进而拉丁因子的水平均值索引矩阵定义如下：$\boldsymbol{L}=[\boldsymbol{l}_1,\ldots,\boldsymbol{l}_p]^{\mathrm{T}}$, 类似地也可定义行因子水平均值索引矩阵$\boldsymbol{R}=[\boldsymbol{r}_1,\ldots,\boldsymbol{r}_p]^{\mathrm{T}}$ 和列因子的水平均值索引矩阵$\boldsymbol{C}=[\boldsymbol{c}_1,\ldots,\boldsymbol{c}_p]^{\mathrm{T}}$.*

易得拉丁因子的第j个水平观察值的和$y_{\cdot j\cdot}=\boldsymbol{l}_j^T\boldsymbol{y}$, 不难验证，上述定义行因子水平索引向量$\boldsymbol{r}_i$和列因子的水平索引向量$\boldsymbol{c}_k$尽管有具体的表达式, 但它们同样满足均值向量的一般定义(2.1). 由正交设计的定义, 水平均值的索引向量应满足以下的定理：

定理 2.1 (水平索引向量性质) *(1).$\boldsymbol{l}_j^{\mathrm{T}}\boldsymbol{1}_n=p$且$\boldsymbol{l}_j^{\mathrm{T}}\boldsymbol{l}_j=p$; (2).$\boldsymbol{l}_{j_1}^{\mathrm{T}}\boldsymbol{l}_{j_2}=0$, 当$j_1\neq j_2$; (3).$\boldsymbol{l}_j^{\mathrm{T}}\boldsymbol{\mu}^*=$*

$p\mu + p\tau_j$; (4).$\boldsymbol{r}_i$和$\boldsymbol{c}_k$也有类似的结论,并且$\boldsymbol{r}_i^{\mathrm{T}}\boldsymbol{l}_j = 1$, $\boldsymbol{r}_i^{\mathrm{T}}\boldsymbol{c}_k = 1$和$\boldsymbol{l}_j^{\mathrm{T}}\boldsymbol{c}_k = 1$.

拉丁因子第j个水平有p个观察值,所以$\boldsymbol{l}_j$有p个元素为1，其余元素为0，从而得结论(1)；任一观察值只属于拉丁因子的一个水平, 拉丁因子的两个不同的水平索引向量同一位置的元素不可能同时非0, 因而拉丁因子不同的两个水平均值索引向量的内积为0，即结论(2)；由拉丁方设计的正交性，来自于拉丁因子第j 个水平的p 个观察值一定是来自于行因子p 个不同水平，也来自于列因子p 个不同水平，所以$\boldsymbol{l}_j^{\mathrm{T}}\boldsymbol{\mu}^* = p\mu + p\tau_j + \sum_{i=1}^{p}\alpha_i + \sum_{k=1}^{p}\beta_k = \mu + \tau_j$, 从而证明结论(3); 行因子水平索引向量$\boldsymbol{r}_i$ 和列因子的水平索引向量$\boldsymbol{c}_k$也可类似由(2.1)定义, 所以也应满足结论(1)-(4), 由正交性可得这三类均值索引向量的任意两类向量的内积为$\frac{1}{n}$. 由定理(2.1), 经简单推导可得水平均值索引矩阵的性质如下：

定理 2.2 (水平均值索引矩阵性质) *(1).* $\boldsymbol{R}\boldsymbol{R}^{\mathrm{T}} = \boldsymbol{L}\boldsymbol{L}^{\mathrm{T}} = \boldsymbol{C}\boldsymbol{C}^{\mathrm{T}} = p\boldsymbol{I}_p$;
(2). $\boldsymbol{R}\boldsymbol{1}_n = \boldsymbol{L}\boldsymbol{1}_n = \boldsymbol{C}\boldsymbol{1}_n = p\boldsymbol{1}_p$, $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{1}_p = \boldsymbol{L}^{\mathrm{T}}\boldsymbol{1}_p = \boldsymbol{C}^{\mathrm{T}}\boldsymbol{1}_p = \boldsymbol{1}_n$;
(3).$\boldsymbol{L}\boldsymbol{\mu}^* = p\mu\boldsymbol{1}_p + p(\tau_1, \ldots, \tau_p)^{\mathrm{T}}$; *矩阵*$\boldsymbol{R}$*和*$\boldsymbol{C}$ *也有类似的结论;*
(4).$\boldsymbol{R}\boldsymbol{L}^{\mathrm{T}} = \boldsymbol{R}\boldsymbol{C}^{\mathrm{T}} = \boldsymbol{C}\boldsymbol{L}^{\mathrm{T}} = \boldsymbol{C}\boldsymbol{R}^{\mathrm{T}} = \boldsymbol{L}\boldsymbol{R}^{\mathrm{T}} = \boldsymbol{L}\boldsymbol{C}^{\mathrm{T}} = \boldsymbol{1}_p\boldsymbol{1}_p^{\mathrm{T}}$.

证 由定理(2.1), $\mathbb{A}\mathbb{A}^{\mathrm{T}} = [\boldsymbol{l}_1, \cdots, \boldsymbol{l}_p]^{\mathrm{T}}[\boldsymbol{l}_1, \cdots, \boldsymbol{l}_p]$ 的第(j_1, j_2) 分块$\boldsymbol{l}_{j_1}^{\mathrm{T}}\boldsymbol{l}_{j_2} = p$当$j_1 = j_2$, 否则为零, 所以有$[\boldsymbol{l}_1, \cdots, \boldsymbol{l}_p]^{\mathrm{T}}[\boldsymbol{l}_1, \cdots, \boldsymbol{l}_p] = p\boldsymbol{I}_p$, 从而$\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}} = p\boldsymbol{I}_p$, (1)中其它结论可类似证明; 给定$j = 1, \ldots, p$, $\boldsymbol{l}_j^{\mathrm{T}}\boldsymbol{1}_n = p$, 所以有$\boldsymbol{L}\boldsymbol{1}_n = [\boldsymbol{l}_1, \cdots, \boldsymbol{l}_p]^{\mathrm{T}}\boldsymbol{1}_n = p\boldsymbol{1}_p$, 任何一个观察值只可能是拉丁因子一个水平上的，拉丁因子所有水平均值向量同一位置上的元素只有一个为1，所以有$\boldsymbol{L}^{\mathrm{T}}\boldsymbol{1}_p = \sum_{j=1}^{p}\boldsymbol{l}_j = \boldsymbol{1}_n$, (2)中其它结论可类似证明; 由定理(2.1)易得(3), 并有$\boldsymbol{r}_i^{\mathrm{T}}\boldsymbol{l}_j = 1$这恰好是$\boldsymbol{R}\boldsymbol{L}^{\mathrm{T}}$的第$(i, j)$ 个元素，所以有结论(4). ▮

3. 检验统计量分布的证明

基于上节引入的水平索引向量和索引矩阵及其它们性质，可以用以下的二次型表示诸因子平方和。

定理 3.1 *式子(1.3)中的诸平方和可以用*$\boldsymbol{y}$*的二次型表示如下：*

$$SS_T = \boldsymbol{y}^{\mathrm{T}}\mathbb{P}\boldsymbol{y}, SS_{行} = \boldsymbol{y}^{\mathrm{T}}\mathbb{P}_1\boldsymbol{y}, SS_{拉丁} = \boldsymbol{y}^{\mathrm{T}}\mathbb{P}_2\boldsymbol{y}, SS_{列} = \boldsymbol{y}^{\mathrm{T}}\mathbb{P}_3\boldsymbol{y}, SS_e = \boldsymbol{y}^{\mathrm{T}}\mathbb{P}_e\boldsymbol{y},$$

其中$\mathbb{P} = \boldsymbol{I}_n - \frac{1}{n}\boldsymbol{1}_n\boldsymbol{1}_n^{\mathrm{T}}$, $\mathbb{P}_1 = \frac{1}{p}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} - \frac{1}{n}\boldsymbol{1}_n\boldsymbol{1}_n^{\mathrm{T}}$, $\mathbb{P}_2 = \frac{1}{p}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L} - \frac{1}{n}\boldsymbol{1}_n\boldsymbol{1}_n^{\mathrm{T}}$, $\mathbb{P}_3 = \frac{1}{p}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C} - \frac{1}{n}\boldsymbol{1}_n\boldsymbol{1}_n^{\mathrm{T}}$, $\mathbb{P}_e = \mathbb{P} - \mathbb{P}_1 - \mathbb{P}_2 - \mathbb{P}_3$.

证 用矩阵表示总和$y_{\ldots} = \boldsymbol{1}_n^{\mathrm{T}}\boldsymbol{y}$, 那么总平方和

$$SS_{\mathrm{T}} = \sum_i\sum_j\sum_k(y_{ijk} - \overline{y}_{\ldots})^2 = \sum_i\sum_j\sum_k y_{ijk}^2 - \frac{y_{\ldots}^2}{n} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} - \frac{1}{n}(\boldsymbol{1}_n^{\mathrm{T}}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{1}_n^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}\mathbb{P}\boldsymbol{y},$$

这里的$\mathbb{P} = \boldsymbol{I}_n - \frac{1}{n}\boldsymbol{1}_n\boldsymbol{1}_n^{\mathrm{T}}$.

因为诸因子平方和二次型表达式证明类似，为了节约篇幅, 只给出$SS_{拉丁}$的二次型表达式的证

DOI: 10.12677/sa.2022.113071

明如下：

$$SS_{拉丁} = p\sum_{j=1}^{p}(\overline{y}_{.j.}-\overline{y}_{...})^2 = \sum_{j=1}^{p}\frac{y_{.j.}^2}{p}-\frac{y_{...}^2}{n}$$

$$= \frac{1}{p}(y_{.1.},\ldots,y_{.p.})(y_{.1.},\ldots,y_{.p.})^{\mathrm{T}}-\frac{1}{n}y_{...}^{\mathrm{T}}y_{...},$$

用矩阵表示$(y_{.1.},\ldots,y_{.p.})^{\mathrm{T}}=\boldsymbol{L}\boldsymbol{y}$, 那么

$$SS_{拉丁} = \boldsymbol{y}^{\mathrm{T}}(\frac{1}{p}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}\mathbb{P}_2\boldsymbol{y},$$

这里$\mathbb{P}_2=\frac{1}{p}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}$. 由$SS_{\mathrm{e}}=SS_{\mathrm{T}}-SS_{行}-SS_{列}-SS_{拉丁}$, 易得$SS_{\mathrm{e}}=\boldsymbol{y}^{\mathrm{T}}\mathbb{P}_{\mathrm{e}}\boldsymbol{y}$, 这里$\mathbb{P}_{\mathrm{e}}=\mathbb{P}-\mathbb{P}_1-\mathbb{P}_2-\mathbb{P}_3$. ∎

定理 3.2 定理*(3.1)*中的矩阵$\mathbb{P}$, $\mathbb{P}_1$, $\mathbb{P}_2$, $\mathbb{P}_3$ 和$\mathbb{P}_{\mathrm{e}}$都是对称矩阵，并满足以下性质：

(1).$\mathbb{P}\mathbb{P}_1=\mathbb{P}_1\mathbb{P}=\mathbb{P}_1$, $\mathbb{P}\mathbb{P}_2=\mathbb{P}_2\mathbb{P}=\mathbb{P}_2$, $\mathbb{P}\mathbb{P}_3=\mathbb{P}_3\mathbb{P}=\mathbb{P}_3$, $\mathbb{P}\mathbb{P}_{\mathrm{e}}=\mathbb{P}_{\mathrm{e}}$;

(2). $\mathbb{P}$, $\mathbb{P}_1$, $\mathbb{P}_2$, $\mathbb{P}_3$和$\mathbb{P}_{\mathrm{e}}$都是幂等矩阵;

(3).$\mathbb{P}_1\mathbb{P}_2=\mathbb{P}_1\mathbb{P}_3=\mathbb{P}_2\mathbb{P}_3=\mathbf{0}$, $\mathbb{P}_1\mathbb{P}_{\mathrm{e}}=\mathbb{P}_2\mathbb{P}_{\mathrm{e}}=\mathbb{P}_3\mathbb{P}_{\mathrm{e}}=\mathbf{0}$, 并满足交换律;

(4). $\mathrm{tr}(\mathbb{P})=p^2-1$, $\mathrm{tr}(\mathbb{P}_1)=p-1$, $\mathrm{tr}(\mathbb{P}_2)=p-1$, $\mathrm{tr}(\mathbb{P}_3)=p-1$和$\mathrm{tr}(\mathbb{P}_{\mathrm{e}})=(p-1)(p-2)$.

证 由定理(3.1)易得矩阵$\mathbb{P}$, $\mathbb{P}_1$, $\mathbb{P}_2$, $\mathbb{P}_3$和$\mathbb{P}_{\mathrm{e}}$ 都是对称矩阵。

$$\mathbb{P}\mathbb{P}_2=\mathbb{P}_2-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}(\frac{1}{p}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})=\mathbb{P}_2-\frac{1}{pn}\mathbf{1}_n(\boldsymbol{L}\mathbf{1}_n)^T\boldsymbol{L}+\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}},$$

由定理(2.2)可得$\boldsymbol{L}\mathbf{1}_n=p\mathbf{1}_p$和$\mathbf{1}_p^{\mathrm{T}}\boldsymbol{L}=\mathbf{1}_n^{\mathrm{T}}$, 从而$\mathbb{P}\mathbb{P}_2=\mathbb{P}_2$, 同理可证$\mathbb{P}\mathbb{P}_1=\mathbb{P}_1$和$\mathbb{P}\mathbb{P}_3=\mathbb{P}_3$, 由$\mathbb{P}$, $\mathbb{P}_1$, $\mathbb{P}_2$ 和$\mathbb{P}_3$是对称矩阵，可得$\mathbb{P}_1\mathbb{P}=\mathbb{P}_1$, $\mathbb{P}_2\mathbb{P}=\mathbb{P}_2$, $\mathbb{P}_3\mathbb{P}=\mathbb{P}_3$; 接下来，证明$\mathbb{P}_2$是幂等矩阵，只需证明$\mathbb{P}^2=\mathbb{P}$和$\mathbb{P}_2^2=\mathbb{P}_2$如下：

$$\mathbb{P}^2=(\boldsymbol{I}_n-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})^2=\boldsymbol{I}_n-\frac{2}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}+\frac{1}{n^2}\mathbf{1}_n(\mathbf{1}_n^{\mathrm{T}}\mathbf{1}_n)\mathbf{1}_n^{\mathrm{T}}=\boldsymbol{I}_n-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}=\mathbb{P},$$

$$\mathbb{P}_2^2 = (\frac{1}{p}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})(\frac{1}{p}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})$$

$$= \frac{1}{p^2}\boldsymbol{L}^{\mathrm{T}}(\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}})\boldsymbol{L}-\frac{1}{pn}\mathbf{1}_n(\boldsymbol{L}\mathbf{1}_n)^{\mathrm{T}}\boldsymbol{L}-\frac{1}{pn}\boldsymbol{L}^{\mathrm{T}}(\boldsymbol{L}\mathbf{1}_n)\mathbf{1}_n^{\mathrm{T}}+\frac{1}{n^2}\mathbf{1}_n(\mathbf{1}_n^{\mathrm{T}}\mathbf{1}_n)\mathbf{1}_n^{\mathrm{T}},$$

由定理(2.2)易得：

$$\mathbb{P}_2^2 = \frac{1}{p^2}\boldsymbol{L}^{\mathrm{T}}p\boldsymbol{I}_p\boldsymbol{L}-\frac{1}{pn}\mathbf{1}_np\mathbf{1}_p^{\mathrm{T}}\boldsymbol{L}-\frac{1}{pn}\boldsymbol{L}^{\mathrm{T}}p\mathbf{1}_p\mathbf{1}_n^{\mathrm{T}}+\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}$$

$$= \frac{1}{p}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}=\mathbb{P}_2,$$

类似地，可证$\mathbb{P}_1$和$\mathbb{P}_3$也是幂等矩阵. 从而, $\mathbb{P}\mathbb{P}_{\mathrm{e}}=\mathbb{P}(\mathbb{P}-\mathbb{P}_1-\mathbb{P}_2-\mathbb{P}_3)=\mathbb{P}-\mathbb{P}_1-\mathbb{P}_2-\mathbb{P}_3=\mathbb{P}_{\mathrm{e}}$.

由定理(2.2)也容易得到:

$$\begin{aligned}\mathbb{P}_1\mathbb{P}_2 =&(\frac{1}{p}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})(\frac{1}{p}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})^{\mathrm{T}}\\ =&\frac{1}{p^2}\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{R}\boldsymbol{L}^{\mathrm{T}})\boldsymbol{L}-\frac{1}{pn}\mathbf{1}_n(\boldsymbol{L}\mathbf{1}_n)^{\mathrm{T}}\boldsymbol{L}-\frac{1}{pn}\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{R}\mathbf{1}_n)\mathbf{1}_n^{\mathrm{T}}+\frac{1}{n^2}\mathbf{1}_n(\mathbf{1}_n^{\mathrm{T}}\mathbf{1}_n)\mathbf{1}_n^{\mathrm{T}}\\ =&\frac{1}{p^2}\boldsymbol{R}^{\mathrm{T}}\mathbf{1}_p\mathbf{1}_p^{\mathrm{T}}\boldsymbol{L}-\frac{1}{pn}\mathbf{1}_n p\mathbf{1}_p^{\mathrm{T}}\boldsymbol{L}-\frac{1}{pn}\boldsymbol{R}^{\mathrm{T}}p\mathbf{1}_p\mathbf{1}_n^{\mathrm{T}}+\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}\\ =&\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}+\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}=\mathbf{0},\end{aligned}$$

也可类似证明$\mathbb{P}_1\mathbb{P}_3=\mathbf{0}$和$\mathbb{P}_2\mathbb{P}_3=\mathbf{0}$, 由结论(i)和(ii)可得：

$$\mathbb{P}_1\mathbb{P}_{\mathrm{e}}=\mathbb{P}_1(\mathbb{P}-\mathbb{P}_1-\mathbb{P}_2-\mathbb{P}_3)=\mathbb{P}_1-\mathbb{P}_1=\mathbf{0},$$

同理可证明$\mathbb{P}_2\mathbb{P}_{\mathrm{e}}=\mathbf{0}$和$\mathbb{P}_3\mathbb{P}_{\mathrm{e}}=\mathbf{0}$，由$\mathbb{P}$, $\mathbb{P}_1$, $\mathbb{P}_2$ 和$\mathbb{P}_3$是对称矩阵，容易证明(ii)中这些式子满足交换律。

$\mathbb{P}_{\mathrm{e}}^2=(\mathbb{P}-\mathbb{P}_1-\mathbb{P}_2-\mathbb{P}_3)^2=\mathbb{P}^2+\mathbb{P}_1^2+\mathbb{P}_2^2+\mathbb{P}_3^2-\mathbb{P}\mathbb{P}_1-\mathbb{P}\mathbb{P}_2-\mathbb{P}\mathbb{P}_3-\mathbb{P}_1\mathbb{P}-\mathbb{P}_2\mathbb{P}-\mathbb{P}_3\mathbb{P}+\mathbb{P}_1\mathbb{P}_2+\mathbb{P}_1\mathbb{P}_3+\mathbb{P}_2\mathbb{P}_1+\mathbb{P}_2\mathbb{P}_3+\mathbb{P}_3\mathbb{P}_1+\mathbb{P}_3\mathbb{P}_2$, 由结论(i)和(ii)易得$\mathbb{P}_{\mathrm{e}}^2=\mathbb{P}_{\mathrm{e}}$. 由迹函数的性质可得：

$$\mathrm{tr}(\mathbb{P})=\mathrm{tr}(\boldsymbol{I}_n-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})=\mathrm{tr}(\boldsymbol{I}_n)-\frac{1}{n}\mathrm{tr}(\mathbf{1}_n^{\mathrm{T}}\mathbf{1}_n)=n-1=p^2-1,$$

再结合定理(2.2)的结论有：

$$\mathrm{tr}(\mathbb{P}_2)=\mathrm{tr}(\frac{1}{p}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})=\frac{1}{p}\mathrm{tr}(\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}})-1=\frac{1}{p}\mathrm{tr}(p\boldsymbol{I}_p)-1=p-1,$$

同理$\mathrm{tr}(\mathbb{P}_1)=p-1$和$\mathrm{tr}(\mathbb{P}_3)=p-1$,

$$\mathrm{tr}(\mathbb{P}_{\mathrm{e}})=\mathrm{tr}(\mathbb{P}_{\mathrm{e}}-\mathbb{P}_1-\mathbb{P}_2-\mathbb{P}_3)=p^2-1-3(p-1)=(p-1)(p-2),$$

这个结论也说明，要运用拉丁方设计做方差分析，其阶数p至少是3. ▮

定理 3.3 *以$\boldsymbol{\mu}^*$和矩阵$\mathbb{P}$, $\mathbb{P}_1$, $\mathbb{P}_2$, $\mathbb{P}_3$和$\mathbb{P}_{\mathrm{e}}$ 构成的二次型满足以下性质：*
(i). $\boldsymbol{\mu}^{*\mathrm{T}}\mathbb{P}\boldsymbol{\mu}^*=\Sigma_{i=1}^p\alpha_i^2+\Sigma_{j=1}^p\tau_j^2+\Sigma_{k=1}^p\beta_k^2$;
(ii). $\boldsymbol{\mu}^{*\mathrm{T}}\mathbb{P}_1\boldsymbol{\mu}^*=\Sigma_{i=1}^p\alpha_i^2$, $\boldsymbol{\mu}^{*\mathrm{T}}\mathbb{P}_2\boldsymbol{\mu}^*=\Sigma_{j=1}^p\tau_j^2$, $\boldsymbol{\mu}^{*\mathrm{T}}\mathbb{P}_3\boldsymbol{\mu}^*=\Sigma_{j=1}^p\beta_k^2$;
(iii). $\boldsymbol{\mu}^{*\mathrm{T}}\mathbb{P}_{\mathrm{e}}\boldsymbol{\mu}^*=0$.

证 由式子(1.1), 得$\mathbb{P}\boldsymbol{\mu}^*=(\boldsymbol{I}_n-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})\boldsymbol{\mu}^*=\boldsymbol{\mu}^*-\mu\mathbf{1}_n$, 由$\mathbb{P}$ 是对称幂等阵，从而$\boldsymbol{\mu}^{*\mathrm{T}}\mathbb{P}\boldsymbol{\mu}^*=(\mathbb{P}\boldsymbol{\mu}^*)^{\mathrm{T}}\mathbb{P}\boldsymbol{\mu}^*=(\boldsymbol{\mu}^*-\mu\mathbf{1}_n)^{\mathrm{T}}(\boldsymbol{\mu}^*-\mu\mathbf{1}_n)=\Sigma_{i=1}^p\alpha_i^2+\Sigma_{j=1}^p\tau_j^2+\Sigma_{k=1}^p\beta_k^2$;

由定理(2.1)和定理(1.3)可得：$\mathbb{P}_2\boldsymbol{\mu}^*=(\frac{1}{p}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})\boldsymbol{\mu}^*=\mu\boldsymbol{L}^{\mathrm{T}}\mathbf{1}_p+\boldsymbol{L}^{\mathrm{T}}(\tau_1,\ldots,\tau_p)^{\mathrm{T}}-\mu\mathbf{1}_n=((\tau_1,\ldots,\tau_p)\boldsymbol{L})^{\mathrm{T}}$, 由$\mathbb{P}_2$是对称幂等阵，从而

$$\boldsymbol{\mu}^{*\mathrm{T}}\mathbb{P}_2\boldsymbol{\mu}^*=(\mathbb{P}_2\boldsymbol{\mu}^*)^{\mathrm{T}}\mathbb{P}_2\boldsymbol{\mu}^*=(\tau_1,\ldots,\tau_p)(\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}})(\tau_1,\ldots,\tau_p)=p\Sigma_{j=1}^p\tau_j^2;$$

由结论(i)和(ii)易得结论(iii). ∎

由定理(3.3), 在原假设H_{01}, H_{02}和H_{03}分别成立时，可得：

$$\boldsymbol{\mu}^{*\mathrm{T}}\mathbb{P}_1\boldsymbol{\mu}^* \overset{\mathrm{H}_{01}}{=} 0, \boldsymbol{\mu}^{*\mathrm{T}}\mathbb{P}_2\boldsymbol{\mu}^* \overset{\mathrm{H}_{02}}{=} 0, \boldsymbol{\mu}^{*\mathrm{T}}\mathbb{P}_3\boldsymbol{\mu}^* \overset{\mathrm{H}_{03}}{=} 0, \tag{3.1}$$

结合定理(3.1)-(3.2)和式子(3.1)，由定理(1.1)可得：

$$\begin{aligned}
&\frac{SS_{行}}{\sigma^2} \overset{\mathrm{H}_{01}}{\sim} \chi^2(p-1),\\
&\frac{SS_{拉丁}}{\sigma^2} \overset{\mathrm{H}_{02}}{\sim} \chi^2(p-1),\\
&\frac{SS_{列}}{\sigma^2} \overset{\mathrm{H}_{03}}{\sim} \chi^2(p-1),\\
&\frac{SS_{\mathrm{e}}}{\sigma^2}\sim\chi^2((p-1)(p-2)),
\end{aligned} \tag{3.2}$$

结合定理(3.2)和定理(1.2)，可得$SS_{行}$、$SS_{拉丁}$ 和$SS_{列}$ 均与SS_{e} 独立。运用以上结论，根据F分布的定义，可证明式子(1.3).

接下来，基于二次型的期望公式可以证明:

$$\begin{aligned}
E(SS_{\mathrm{T}}) =&(p^2-1)\sigma^2 + p\Sigma_{j=1}^{p}\tau_j^2,\\
E(SS_{行}) =&(p-1)\sigma^2 + p\Sigma_{i=1}^{p}\alpha_i^2,\\
E(SS_{拉丁}) =&(p-1)\sigma^2 + p\Sigma_{j=1}^{p}\tau_j^2,\\
E(SS_{列}) =&(p-1)\sigma^2 + p\Sigma_{k=1}^{p}\beta_k^2,\\
E(SS_{\mathrm{e}}) =&(p-1)(p-2)\sigma^2.
\end{aligned} \tag{3.3}$$

证 由迹函数的性质可得：$E(SS_{Lat}) = E(\boldsymbol{y}^{\mathrm{T}}\mathbb{P}_2\boldsymbol{y}) = E(\mathrm{tr}(\boldsymbol{y}^{\mathrm{T}}\mathbb{P}_2\boldsymbol{y})) = \mathrm{tr}(\mathbb{P}_2E(\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}))$, 其中由$\mathrm{Cov}(\boldsymbol{y}) = E(\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}) - E(\boldsymbol{y})E(\boldsymbol{y})^{\mathrm{T}}$, 可得$E(\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}) = \sigma^2\boldsymbol{I}_n + \boldsymbol{\mu}^*\boldsymbol{\mu}^{*\mathrm{T}}$, 从而由定理(3.2)~(3.3)得：

$$E(SS_{Lat}) = \sigma^2\mathrm{tr}(\mathbb{P}_2) + \mathrm{tr}(\mathbb{P}_2\boldsymbol{\mu}^*\boldsymbol{\mu}^{*\mathrm{T}}) = \sigma^2\mathrm{tr}(\mathbb{P}_2) + \boldsymbol{\mu}^{*\mathrm{T}}\mathbb{P}_2\boldsymbol{\mu}^* = (p-1)\sigma^2 + p\Sigma_{j=1}^{p}\tau_j^2,$$

同理可证明其它结论。 ∎

4. 结论

作为析因试验的部分实施，拉丁方设计在试验设计中占据着重要的位置，本文基于均值索引矩阵性质证明拉丁方设计检验统计量，简单易懂，有助于理解拉丁方设计的原理，并且容易推广到其他方差分析模型中相关统计量的证明，包括多因子方差分析固定效应模型、多因子方差分析

随机效应模型、多因子方差分析混合效应模型、正交设计方差分析模型 [4] 等。

基金项目

云南大学研究生创新人才研究培育项目————研究生课程教材建设质量提升计划(CZ22622202)，云南大学2020年教育教学改革研究项目《试验设计实验》非标准答案考核。

参考文献

[1] Montgomery, D.C. (1976) Design and Analysis of Experiments. Willey, New York.

[2] 王万中. 试验的设计与分析[M]. 北京: 高等教育出版社, 2004.

[3] 茆诗松, 周纪芗, 陈颖. 试验设计[M]. 第2版. 北京: 中国统计出版社, 2012.

[4] 方开泰, 马长兴. 正交与均匀试验设计[M]. 北京: 科学出版社, 2001.