

# 分层单双边配对数据优势比的同质性检验及样本量的确定

计逸宁

南京邮电大学理学院, 江苏 南京

收稿日期: 2024年4月23日; 录用日期: 2024年5月17日; 发布日期: 2024年5月24日

## 摘要

在临床试验中, 会对对身体器官或部位进行试验, 获得单边或双边的配对数据; 同时, 由于患者年龄、性别等混杂因素的影响, 相同的治疗方法可能会产生不同的治疗效果。因此, 本文基于优势比这一衡量治疗效果的指标对分层单双边混合配对数据进行研究。在Donner模型下提出对分层两组的单双边混合数据中各层优势比的同质性假设。利用迭代算法得到最大似然估计量; 构造似然比检验统计量( $T_{LR}$ )、Wald检验统计量( $T_W$ )和拉格朗日乘数检验统计量( $T_{SC}$ )进行同质性检验; 基于第一类错误率和功效比较三种检验统计量的效果。为临床试验设计和试验数据的检验方法选择提供建议, 以提高检验效果。本文还通过对比各检验方法, 预测在固定的显著性水平下达到目标功效所需的最小样本量, 既减少临床试验成本, 又使后续进行的试验有更稳健的第一类错误率和更好的功效。蒙特卡洛模拟结果显示: 在同质性假设下,  $T_W$ 在三个检验统计量中表现最为冒进;  $T_{LR}$ 相对冒进, 总体表现较为稳健; 而 $T_{SC}$ 在各种情况下均表现稳健, 且功效较高, 因此最为推荐使用。对于试验样本量的确定, 为了达到相同的功效,  $T_W$ 所需样本量较小; 随着层数的增加, 三个检验统计量所需样本量差异逐渐减小; 随着优势比增加, 越来越接近于1, 三个检验统计量所需样本量也会逐渐增加。

## 关键词

配对数据, 优势比, 最大似然估计, 同质性检验, 样本量

## Homogeneity Test and Sample Size Determination of Odds Ratios for Stratified Unilateral and Bilateral Paired Data

Yining Ji

College of Science, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu

Received: Apr. 23<sup>rd</sup>, 2024; accepted: May 17<sup>th</sup>, 2024; published: May 24<sup>th</sup>, 2024

文章引用: 计逸宁. 分层单双边配对数据优势比的同质性检验及样本量的确定[J]. 应用数学进展, 2024, 13(5): 2029-2044. DOI: 10.12677/aam.2024.135190

## Abstract

In clinical trials, pairs of body organs or parts are tested to obtain unilateral or bilateral paired data. At the same time, due to the influence of confounding factors such as patient age and gender, the same treatment method may produce different therapeutic effects. Therefore, this paper studied stratified unilateral and bilateral paired data based on the odds ratio which is an index to measure the treatment effect. In the Donner model, the homogeneity hypothesis of the odds ratios of each stratum in the stratified unilateral and bilateral paired data in two groups is proposed. The maximum likelihood estimator is obtained by iterative algorithm. The likelihood ratio test statistic ( $T_{LR}$ ), Wald test statistic ( $T_W$ ) and Lagrange multiplier test statistic ( $T_{SC}$ ) are constructed for homogeneity test. The effects of the three test statistics are compared based on the type I error rate and power, to provide suggestions for clinical trial design and test method selection of test data, so as to improve test results. This paper also predicts the minimum sample size required to achieve the target power at a fixed significance level by comparing these testing methods, which not only reduces the cost of clinical trials, but also enables the subsequent trials to have a more robust type I error rate and better power. Monte Carlo simulation results show that  $T_W$  is the most aggressive among the three test statistics under the homogeneity hypothesis.  $T_{LR}$  is relatively aggressive, the overall performance is relatively stable;  $T_{SC}$  is robust in all situations and has high power, so it is most recommended. For the determination of the sample size of the test, in order to achieve the same power, the sample size required by  $T_W$  is small. With the increase of the number of stratum, the difference of sample size required for the three test statistics gradually decreases. As the odds ratio increases and gets closer to 1, the sample size required for the three test statistics will gradually increase.

## Keywords

Paired Data, Odds Ratio, Maximum Likelihood Estimation, Homogeneity Test, Sample Size

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在临床试验中, 研究人员想要了解疾病的治疗情况, 可以从患者的特定组织或者器官收集数据, 并将患者根据治疗方案分组。对于眼睛、耳朵等成对的身体器官, 会带来单边或双边的配对数据, 在这种情况下, 只研究单边或双边的检验不太合适, 会使得从数据中获得分析结果不完整、不准确[1]。Rosner (1982) [2]指出成对身体器官的治疗结果通常是高度相关的, 它们之间的相关性是不可忽视的, 任何忽略这一依赖特征的统计过程都将导致不正确的推论。Donner (1989) [3]提出了一个配对数据的相关系数模型, Thompson (1993) [4]证明了 Donner 的模型可以使充分利用单个或成对器官的数据来提高检测的有效性。Ma 和 Wang (2021) [5]构造似然比检验、Wald 检验和 score 检验统计量, 检验了单双边混合配对数据比例的同质性, 并比较了这些方法下的经验性第一类错误率和功效。治疗效果不仅取决于治疗方案, 也可能与病人的性别、年龄、不同治疗环境等影响有关。Kahan 和 Morris (2013) [6]指出, 对这类配对数据的分析如果受到分层的影响可能会产生混杂因素, 不能够忽略这些因素。因此, 需要对数据进行分层, 以判断不同层间是否存在分层因素的影响。Sun 等(2022) [7]基于单边和双边的分层配对数据进行了风险差的同质性检验, 她进一步研究了使检验满足所需功效的样本量。样本量的确定也是试验设计的重要组成部分

分[8], 因此将其融入配对数据的研究。此外, 优势比是一个比较两组治疗效果的指标, 是一种首选的前瞻性联想测量方法, 主要用于回顾性或横断面抽样设计[9], 并可对此进行同质性检验。

本文基于分层单双边混合配对数据得到的优势比对同质性假设的检验方法进行研究, 可以为生物医学的临床试验数据在检验过程中检验方法的选择上提供建议; 对于样本量的研究, 可以为将来临床试验的方案设计提供建议, 以减少临床实验中人力、物力等成本的消耗。

## 2. 数据结构和最大似然估计

### 2.1. 数据结构

假设  $m_{ij}^{(q)} (l=0,1,2, i=1,2, q=1,2, j=1, \dots, J)$  是指第  $j$  层个体在接受组  $i$  类别的治疗后, 治疗部位的响应个数。相应的,  $p_{ij}^{(q)}$  分别表示在第  $j$  层中接受组  $i$  类别的治疗后, 没有响应, 单边响应, 双边响应的概率。

**Table 1.** Data structure of the  $j$ th stratum  
**表 1.** 第  $j$  层的单双边混合数据结构

响应个数( $l$ )	单边数据		双边数据		合计
	组 1 ( $i=1$ )	组 2 ( $i=2$ )	组 1 ( $i=1$ )	组 2 ( $i=2$ )	
0	$m_{01j}^{(1)} (p_{01j}^{(1)})$	$m_{02j}^{(1)} (p_{02j}^{(1)})$	$m_{01j}^{(2)} (p_{01j}^{(2)})$	$m_{02j}^{(2)} (p_{02j}^{(2)})$	$S_{0j}$
1	$m_{11j}^{(1)} (p_{11j}^{(1)})$	$m_{12j}^{(1)} (p_{12j}^{(1)})$	$m_{11j}^{(2)} (p_{11j}^{(2)})$	$m_{12j}^{(2)} (p_{12j}^{(2)})$	$S_{1j}$
2	-	-	$m_{21j}^{(2)} (p_{21j}^{(2)})$	$m_{22j}^{(2)} (p_{22j}^{(2)})$	$S_{2j}$
合计	$m_{+1j}^{(1)} (p_{+1j}^{(1)})$	$m_{+2j}^{(1)} (p_{+2j}^{(1)})$	$m_{+1j}^{(2)} (p_{+1j}^{(2)})$	$m_{+2j}^{(2)} (p_{+2j}^{(2)})$	$N_j$

当  $q=1$  时为单边数据,  $m_{ij}^{(1)} = (m_{01j}^{(1)}, m_{11j}^{(1)})$  服从二项分布  $B \sim (1, p_{ij}^{(1)})$ ; 当  $q=2$  时为双边数据,  $m_{ij}^{(2)} = (m_{0ij}^{(2)}, m_{1ij}^{(2)}, m_{2ij}^{(2)})$  服从三项分布  $m_{ij}^{(2)} \sim (m_{+ij}^{(2)}, p_{0ij}^{(2)}, p_{1ij}^{(2)}, p_{2ij}^{(2)})$ 。令  $Z_{hij}^{(1)}$  表示治疗单边部位的个体在第  $j$  层的治疗部位的响应的指示变量,  $Z_{hij}^{(1)} = 1$  表示存在响应,  $Z_{hij}^{(1)} = 0$  表示未响应。 $Z_{hijk}^{(2)} (k=1,2)$  表示治疗双边部位的个体在第  $j$  层第  $k$  个部位的响应的指示变量,  $Z_{hijk}^{(2)} = 1$  表示存在响应,  $Z_{hijk}^{(2)} = 0$  表示未响应。在 Donner 模型中, 假设  $P(Z_{hij}^{(1)} = 1) = P(Z_{hijk}^{(2)} = 1) = \pi_{ij} (0 \leq \pi_{ij} \leq 1)$ ,  $Corr(Z_{hijk}^{(2)}, Z_{hij(3-k)}^{(2)}) = \rho_j (0 \leq \rho_j \leq 1)$  [5]。

基于表 1 中的数据结构和 Donner 模型的假设, 得到  $p_{ij}^{(q)}$  的概率:

$$\begin{aligned}
 p_{01j}^{(1)} &= 1 - \pi_{ij}, \quad p_{11j}^{(1)} = \pi_{ij}; \\
 p_{01j}^{(2)} &= \rho_j (1 - \pi_{ij}) + (1 - \rho_j) (1 - \pi_{ij})^2, \\
 p_{11j}^{(2)} &= 2\pi_{ij} (1 - \rho_j) (1 - \pi_{ij}), \\
 p_{21j}^{(2)} &= \rho_j \pi_{ij} + (1 - \rho_j) \pi_{ij}^2; \quad i=1,2, \quad j=1, \dots, J.
 \end{aligned}$$

令  $\pi_i = (\pi_{i1}, \pi_{i2}, \dots, \pi_{iJ})^T$ ,  $\rho = (\rho_1, \rho_2, \dots, \rho_J)^T$ , 又  $m_{ij}^{(q)} = (m_{0ij}^{(q)}, m_{1ij}^{(q)}, m_{2ij}^{(q)}) (i=1,2)$  相互独立, 因此可以得到它的似然函数为:

$$L(\pi_1, \pi_2, \rho | m_{ij}^{(q)}) = f(m_{0ij}^{(1)}, m_{1ij}^{(1)}, m_{0ij}^{(2)}, m_{1ij}^{(2)}, m_{2ij}^{(2)} | \pi_i, \rho) = \prod_{i=1}^2 f(m_{ij}^{(1)} | \pi_i, \rho) \prod_{i=1}^2 f(m_{ij}^{(2)} | \pi_i, \rho),$$

则可以得到如下对数似然函数:

$$l(\pi_1, \pi_2, \rho) = \sum_{j=1}^J l_j(\pi_{1j}, \pi_{2j}, \rho_j) + \log C,$$

其中 
$$l_j(\pi_{1j}, \pi_{2j}, \rho_j) = \sum_{i=1}^2 \left\{ m_{0ij}^{(1)} \log(1 - \pi_{ij}) + m_{1ij}^{(1)} \log \pi_{ij} + m_{0ij}^{(2)} \log [\rho_j (1 - \rho_j) (1 - \pi_{ij})^2] \right. \\ \left. + m_{1ij}^{(2)} \log [2\pi_{ij} (1 - \rho_j) (1 - \pi_{ij})] + m_{2ij}^{(2)} \log [\rho_j \pi_{ij} + (1 - \rho_j) \pi_{ij}^2] \right\},$$

$$C = \prod_{j=1}^J \prod_{i=1}^2 \frac{(m_{+ij}^{(1)})!}{m_{0ij}^{(1)}! m_{1ij}^{(1)}!} \frac{(m_{+ij}^{(2)})!}{m_{0ij}^{(2)}! m_{1ij}^{(2)}! m_{2ij}^{(2)}!}$$

为一个与参数  $\pi_{ij}$  和  $\rho_j$  无关的常数。

选取优势比作为衡量两组间治疗效果的指标，在分层模型中，令  $\theta_j = \frac{\pi_{1j}/(1 - \pi_{1j})}{\pi_{2j}/(1 - \pi_{2j})}$  为第  $j$  层中两组之间的优势比，反映了各层的治疗方案在治疗效果上的差异，从而建立治疗方案与治疗效果间的联系。本文进行如下假设： $H_0: \theta_1 = \theta_2 = \dots = \theta_J = \theta \triangleq \theta_0$ ， $H_1: \theta \neq \theta_0$ ，其中  $\theta_0$  为一个大于零的常数。

### 2.2. 全局最大似然估计

令  $\hat{\pi}_{1j}$ ， $\hat{\pi}_{2j}$ ， $\hat{\rho}_j$  为  $\pi_{1j}$ ， $\pi_{2j}$ ， $\rho_j$  的估计值，对于  $\pi_{ij}$  和  $\rho_j$  ( $i=1, 2; j=1, \dots, J$ )，通过如下等式求得  $\pi_{1j}$ ， $\pi_{2j}$ ， $\rho_j$  的全局最大似然估计值：

$$\frac{\partial l_j}{\partial \pi_{ij}} = 0, \quad \frac{\partial l_j}{\partial \rho_j} = 0.$$

计算得到全局最大似然估计值  $\hat{\pi}_{ij}$ ， $\hat{\rho}_j$ 。由于上述式子不存在显式解，因此需要使用迭代算法对方程进行求解。使用 Fisher 得分法求解得到最大似然估计值  $\hat{\pi}_{ij}$ ， $\hat{\rho}_j$ 。

首先设置初值：

$$\pi_{1j}^{(0)} = \frac{m_{11j}^{(1)} + m_{11j}^{(2)} + 2m_{21j}^{(2)}}{m_{+1j}^{(1)} + 2m_{+1j}^{(2)}}, \quad \rho_j^{(0)} = \frac{4(m_{01j}^{(2)} + m_{02j}^{(2)})(m_{21j}^{(2)} + m_{22j}^{(2)}) - (m_{11j}^{(2)} + m_{12j}^{(2)})^2}{[m_{11j}^{(2)} + m_{12j}^{(2)} + 2(m_{21j}^{(2)} + m_{22j}^{(2)})][m_{11j}^{(2)} + m_{12j}^{(2)} + 2(m_{01j}^{(2)} + m_{02j}^{(2)})]},$$

参数的第  $(t+1)$  次迭代值可以通过 Fisher 得分法得到：

$$\begin{bmatrix} \pi_{1j}^{(t+1)} \\ \pi_{2j}^{(t+1)} \\ \rho_j^{(t+1)} \end{bmatrix} = \begin{bmatrix} \pi_{1j}^{(t)} \\ \pi_{2j}^{(t)} \\ \rho_j^{(t)} \end{bmatrix} + [I_{0j}]^{-1} (\pi_{1j}, \pi_{2j}, \rho_j) \times \begin{bmatrix} \frac{\partial l_j}{\partial \pi_{1j}} \\ \frac{\partial l_j}{\partial \pi_{2j}} \\ \frac{\partial l_j}{\partial \rho_j} \end{bmatrix},$$

其中  $I_{0j}$  为关于  $\pi_{1j}$ ， $\pi_{2j}$ ， $\rho_j$  的信息矩阵， $I_{0j} = \begin{bmatrix} I_{011j} & I_{012j} & I_{013j} \\ I_{021j} & I_{022j} & I_{023j} \\ I_{031j} & I_{032j} & I_{033j} \end{bmatrix}$ ， $I_{011j} = E\left(-\frac{\partial^2 l_j}{\partial \pi_{1j}^2}\right)$ ， $I_{012j} = E\left(-\frac{\partial^2 l_j}{\partial \pi_{1j} \partial \pi_{2j}}\right)$ ，

$$I_{013j} = E\left(-\frac{\partial^2 l_j}{\partial \pi_{1j} \partial \rho_j}\right), \quad I_{022j} = E\left(-\frac{\partial^2 l_j}{\partial \pi_{2j}^2}\right), \quad I_{023j} = E\left(-\frac{\partial^2 l_j}{\partial \pi_{2j} \partial \rho_j}\right), \quad I_{033j} = E\left(-\frac{\partial^2 l_j}{\partial \rho_j^2}\right).$$

### 2.3. 同质性假设下的最大似然估计

在研究中，对每层的优势比做一致性检验后，如果优势比的假设成立，那么这个优势比的值是多少

就需要关注。因此在等优势比假设成立的基础上，对优势比进行同质性检验。

在  $H_0: \theta_1 = \theta_2 = \dots = \theta_J = \theta \triangleq \theta_0$  假设下， $\pi_{2j} = \frac{\pi_{1j}}{\pi_{1j} - \pi_{1j}\theta_0 + \theta_0}$  ( $j=1, \dots, J$ )。此时的未知参数只有  $\pi_{ij}$  和  $\rho_j$ 。则相应的对数似然函数可以写作： $l^c(\pi_1, \rho, \theta_0) = \sum_{j=1}^J l_j(\pi_{1j}, \rho_j, \theta_0) + \log C$ 。

令  $\tilde{\pi}_{1j}$ ,  $\tilde{\rho}_j$  为  $\pi_{1j}$ ,  $\rho_j$  在  $H_0$  假设下的限制性最大似然估计值，即求下列式子的解：

$$\frac{\partial l_j^c}{\partial \pi_{1j}} = 0, \quad \frac{\partial l_j^c}{\partial \rho_j} = 0, \quad j=1, \dots, J。$$

使用 Fisher 得分法计算  $H_0$  假设下的最大似然估计值  $\tilde{\pi}_{1j}$ ,  $\tilde{\rho}_j$ 。

首先，对于第  $j$  层，设置  $\pi_{1j}$  和  $\rho_j$  的初值为：

$$\pi_{1j}^{(0)} = \frac{m_{11j}^{(1)} + m_{11j}^{(2)} + 2m_{21j}^{(2)}}{m_{+1j}^{(1)} + 2m_{+1j}^{(2)}}, \quad \rho_j^{(0)} = \frac{4(m_{01j}^{(2)} + m_{02j}^{(2)})(m_{21j}^{(2)} + m_{22j}^{(2)}) - (m_{11j}^{(2)} + m_{12j}^{(2)})^2}{[m_{11j}^{(2)} + m_{12j}^{(2)} + 2(m_{21j}^{(2)} + m_{22j}^{(2)})][m_{11j}^{(2)} + m_{12j}^{(2)} + 2(m_{01j}^{(2)} + m_{02j}^{(2)})]},$$

参数  $\pi_{1j}$  和  $\rho_j$  ( $j=1, \dots, J$ ) 的第  $(t+1)$  次迭代值可以通过 Fisher 得分法得到：

$$\begin{bmatrix} \pi_{1j}^{(t+1)} \\ \rho_j^{(t+1)} \end{bmatrix} = \begin{bmatrix} \pi_{1j}^{(t)} \\ \rho_j^{(t)} \end{bmatrix} + I_{1j}^{-1}(\pi_{1j}, \rho_j) \times \begin{bmatrix} \frac{\partial l_j^c}{\partial \pi_{1j}} \\ \frac{\partial l_j^c}{\partial \rho_j} \end{bmatrix},$$

其中  $I_{1j} = \begin{bmatrix} I_{111j} & I_{113j} \\ I_{113j} & I_{133j} \end{bmatrix}$ ,  $I_{111j} = E\left(-\frac{\partial^2 l_j^c}{\partial \pi_{1j}^2}\right)$ ,  $I_{113j} = E\left(-\frac{\partial^2 l_j^c}{\partial \pi_{1j} \partial \rho_j}\right)$ ,  $I_{133j} = E\left(-\frac{\partial^2 l_j^c}{\partial \rho_j^2}\right)$ 。

由于当样本量足够大时，即  $m \rightarrow +\infty$  时，似然比检验统计量、Wald 检验统计量、score 检验统计量等均渐进服从自由度为  $J-1$  的  $\chi^2$  分布，选择这三种方法对假设  $H_0$  进行检验，比较它们的第一类错误率和功效。

### 3. 同质性检验

#### 3.1. 检验统计量的构造

似然比检验分别计算在约束和无约束条件下的参数估计值，然后计算二者的对数似然函数是否足够接近[10]，所构造的似然比检验统计量的具体公式如下：

$$T_{LR} = 2 \ln \frac{\hat{L}^c}{\tilde{L}} = -2 \left[ l^c(\tilde{\pi}_1, \tilde{\rho}, \tilde{\theta}) - l(\hat{\pi}_1, \hat{\pi}_2, \hat{\rho}) \right] = -2 \sum_{j=1}^J \left[ l_j^c(\tilde{\pi}_{1j}, \tilde{\rho}_j, \tilde{\theta}_j) - l_j(\hat{\pi}_{1j}, \hat{\pi}_{2j}, \hat{\rho}_j) \right],$$

其中  $\hat{\pi}_{1j}$ ,  $\hat{\pi}_{2j}$ ,  $\hat{\rho}_j$  是全局最大似然估计， $\tilde{\pi}_{1j}$ ,  $\tilde{\rho}_j$ ,  $\tilde{\theta}_j$  是  $H_0$  下的最大似然估计。在给定显著性水平  $\alpha$  的情况下， $T_{LR} > \chi_{1-\alpha}^2(J-1)$  时拒绝原假设， $T_{LR} \leq \chi_{1-\alpha}^2(J-1)$  时接受原假设。

Wald 检验是检验无约束的估计值满足约束条件的程度。令  $\beta = (\theta, \pi_{11}, \rho_1, \dots, \pi_{1J}, \rho_J)_{2J+1}^T$ ，相应的  $H_0$  下的最大似然估计值为  $\hat{\beta} = (\hat{\theta}, \hat{\pi}_{11}, \hat{\rho}_1, \dots, \hat{\pi}_{1J}, \hat{\rho}_J)_{2J+1}^T$ ， $K = (1, 0, 0, \dots, 0)_{2J+1}$ ，那么  $K \times \hat{\beta}$  为  $H_0$  下的最大似然估计值，则构造在假设  $H_0: \theta_1 = \theta_2 = \dots = \theta_J = \theta \triangleq \theta_0$  下的 Wald 检验统计量为：

$$T_W = \frac{(\hat{\theta} - \theta_0)^2}{\text{Var}(\beta)} = \frac{(K\hat{\beta} - \theta_0)^2}{I_2^{-1}} = \frac{(\hat{\theta} - \theta_0)^2}{KI_2^{-1}K^T} = \frac{(\hat{\theta} - \theta_0)^2}{I_2^{-1}(1,1)},$$

$$I_2 = \begin{pmatrix} I_{244} & B \\ B^T & I_{NJ} \end{pmatrix}$$

为关于  $\beta$  的 Fisher 信息矩阵,  $I_2^{-1}(1,1)$  为  $I_2$  的逆矩阵对角线上的第一个元素, 其中

$$I_{244} = E \left( -\sum_{j=1}^J \frac{\partial^2 l_j^c}{\partial \theta^2} \right), \quad B = \left( E \left( -\frac{\partial^2 l_1^c}{\partial \theta \pi_{11}} \right), E \left( -\frac{\partial^2 l_1^c}{\partial \theta \rho_1} \right), \dots, E \left( -\frac{\partial^2 l_J^c}{\partial \theta \pi_{1J}} \right), E \left( -\frac{\partial^2 l_J^c}{\partial \theta \rho_J} \right) \right),$$

$$I_{NJ} = \text{diag}(I_{n1}, I_{n2}, \dots, I_{nJ}), \quad I_{nj} = \begin{pmatrix} \frac{\partial^2 l_j^c}{\partial \pi_{1j}^2} & \frac{\partial^2 l_j^c}{\partial \pi_{1j} \rho_j} \\ \frac{\partial^2 l_j^c}{\partial \pi_{1j} \rho_j} & \frac{\partial^2 l_j^c}{\partial \rho_j^2} \end{pmatrix}.$$

score 检验所感兴趣的参数为  $\theta$ , 因此令计分矩阵  $U = \left( \frac{\partial l^c}{\partial \theta}, 0, 0, \dots, 0, 0 \right)$ , score 检验统计量就转化为:

$$T_{sc}^c = \left( \sum_{j=1}^J \frac{\partial l_j^c}{\partial \theta} \right)^2 I_2^{-1}(1,1) |_{\theta=\bar{\theta}},$$

其中  $I_2^{-1}(1,1)$  为  $I_2$  的逆矩阵对角线上的第一个元素。

### 3.2. 仿真模拟

本节使用蒙特卡罗模拟来比较似然比统计量、Wald 检验统计量和 score 检验统计量的经验性的一类错误率和功效, 用于衡量三个统计量的性能。第一类错误率是原假设成立时拒绝原假设的概率。当第一类错误率介于 0.04 和 0.06 之间时, 检验被认为是稳健的; 大于 0.06 时, 检验被认为是冒进的; 小于 0.04 时, 被认为是保守的。功效是当原假设为假时接受它的概率 [11]。一个好的检验应该同时具有稳健的第一类错误率和尽可能高的功效。

设样本量  $m = (m_{ij}^{(1)}, m_{ij}^{(1)}) = (25, 25), (50, 50), (100, 100)$ ,  $J = 2, 4, 6$ 。首先, 计算不同参数配置下检验统计量的经验性第一类错误率, 具体参数配置如表 2 所示。

**Table 2.** Parameter configuration of Monte Carlo simulation

**表 2.** 蒙特卡罗模拟的参数配置

		层数		
		$J = 2$	$J = 4$	$J = 6$
$\pi_{1j}$	I	(0.3, 0.5)	(0.3, 0.5, 0.3, 0.5)	(0.3, 0.5, 0.3, 0.5, 0.3, 0.5)
	II	(0.4, 0.5)	(0.4, 0.5, 0.4, 0.5)	(0.4, 0.5, 0.4, 0.5, 0.4, 0.5)
	III	(0.4, 0.4)	(0.4, 0.4, 0.4, 0.4)	(0.4, 0.4, 0.4, 0.4, 0.4, 0.4)
$\pi_{2j}$	A	(0.4, 0.5)	(0.4, 0.5, 0.4, 0.5)	(0.4, 0.5, 0.4, 0.5, 0.4, 0.5)
	B	(0.5, 0.6)	(0.5, 0.6, 0.5, 0.6)	(0.5, 0.6, 0.5, 0.6, 0.5, 0.6)
	C	(0.5, 0.5)	(0.5, 0.5, 0.5, 0.5)	(0.5, 0.5, 0.5, 0.5, 0.5, 0.5)

模拟同质性检验的经验性第一类错误率时, 在每组的样本量相同的情况下进行构造。对于表 2 中的每种参数组合进行 10000 次模拟, 经验性第一类错误率即为在原假设为真的情况下, 拒绝原假设的次数除以 10000。表 3 和附表 1、附表 2 展示了层数  $J = 2$  时,  $\theta = 0.2, 0.4, 0.6$  的经验性第一类错误率大小。

从表中可以看出, 当层数相同时, 三种检验方法的结果都可以稳定在 0.04 至 0.06 的范围内, 其中  $T_w$

更为冒进，而  $T_{LR}$  则更为保守。

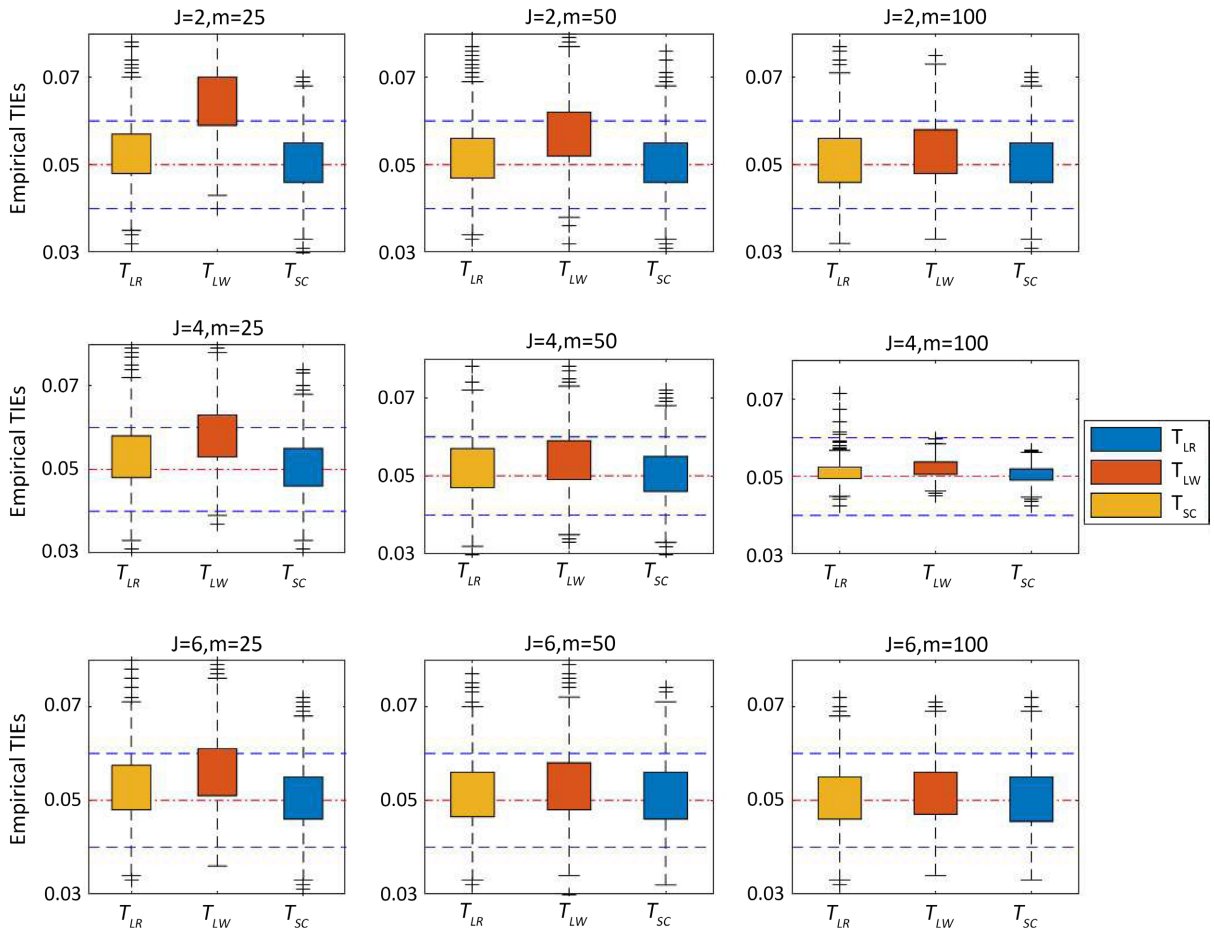
为了更直观反应三种不同方法在经验性第一类错误率上的差异以及它们的渐近性质，下面分别在不同层与不同相关系数的情形下随机生成 1000 组参数并做蒙特卡洛模拟，每一层分别绘制出三种方法在样本量为 25,50,100 情况下的经验性第一类错误率的箱线图。各层情况下的箱线图如图所示。

**Table 3.** The empirical TIEs (%) for  $J = 2$

**表 3.** 经验性第一类错误率 ( $J = 2$ )

$\theta$	$\pi_1$	$\rho$	$m = (25, 25)$			$m = (50, 50)$			$m = (100, 100)$		
			$T_{LR}^C$	$T_w^C$	$T_{SC}^C$	$T_{LR}^C$	$T_w^C$	$T_{SC}^C$	$T_{LR}^C$	$T_w^C$	$T_{SC}^C$
0.2	I	A	0.0490	0.0742	0.0471	0.0523	0.0613	0.0514	0.0530	0.0542	0.0534
		B	0.0536	0.0723	0.0514	0.0520	0.0611	0.0519	0.0487	0.0522	0.0479
		C	0.0535	0.0720	0.0507	0.0529	0.0622	0.0510	0.0481	0.0530	0.0479
	II	A	0.0515	0.0688	0.0496	0.0498	0.0600	0.0489	0.0516	0.0555	0.0513
		B	0.0514	0.0704	0.0505	0.0487	0.0588	0.0479	0.0476	0.0518	0.0465
		C	0.0554	0.0724	0.0537	0.0505	0.0594	0.0503	0.0502	0.0542	0.0504
	III	A	0.0582	0.0745	0.0562	0.0504	0.0583	0.0502	0.0456	0.0510	0.0462
		B	0.0564	0.0750	0.0551	0.0486	0.0575	0.0481	0.0497	0.0547	0.0489
		C	0.0527	0.0734	0.0507	0.0514	0.0611	0.0502	0.0485	0.0555	0.0484
0.4	I	A	0.0551	0.0719	0.0539	0.0538	0.0577	0.0538	0.0517	0.0543	0.0517
		B	0.0559	0.0702	0.0549	0.0546	0.0602	0.0542	0.0536	0.0562	0.0535
		C	0.0545	0.0674	0.0549	0.0487	0.0556	0.0484	0.0491	0.0544	0.0486
	II	A	0.0537	0.0672	0.0524	0.0478	0.0579	0.0477	0.0513	0.0546	0.0514
		B	0.0501	0.0651	0.0492	0.0515	0.0592	0.0508	0.0497	0.0505	0.0499
		C	0.0512	0.0666	0.0501	0.0501	0.0557	0.0499	0.0537	0.0550	0.0534
	III	A	0.0559	0.0727	0.0548	0.0533	0.0599	0.0531	0.0530	0.0583	0.0529
		B	0.0522	0.0650	0.0505	0.0507	0.0601	0.0505	0.0534	0.0544	0.0532
		C	0.0545	0.0686	0.0533	0.0496	0.0569	0.0494	0.0529	0.0586	0.0529
0.6	I	A	0.0482	0.0622	0.0471	0.0535	0.0586	0.0526	0.0465	0.0477	0.0469
		B	0.0519	0.0656	0.0511	0.0535	0.0600	0.0539	0.0519	0.0556	0.0516
		C	0.0483	0.0603	0.0477	0.0490	0.0594	0.0485	0.0472	0.0506	0.0469
	II	A	0.0534	0.0647	0.0518	0.0503	0.0565	0.0498	0.0550	0.0555	0.0546
		B	0.0582	0.0648	0.0573	0.0517	0.0596	0.0510	0.0451	0.0505	0.0450
		C	0.0534	0.0664	0.0526	0.0517	0.0574	0.0515	0.0498	0.0525	0.0499
	III	A	0.0519	0.0615	0.0512	0.0510	0.0570	0.0501	0.0480	0.0512	0.0478
		B	0.0485	0.0611	0.0481	0.0525	0.0550	0.0521	0.0523	0.0561	0.0520
		C	0.0552	0.0628	0.0539	0.0528	0.0515	0.0523	0.0496	0.0558	0.0493

图 1 可以看出，当  $J$  和  $m$  增加时，各检验统计量的经验性的一类错误率显著趋向于 0.05。其中， $T_{SC}$  在各种情况下均表现稳健； $T_w$  在层数和每组人数较少的情况下表现冒进，当层数和每组人数都较高时，才表现稳



**Figure 1.** The box plots of empirical TIEs  
**图 1.** 不同样本量下第一类错误率的箱线图

**Table 4.** The empirical powers (%) for  $J = 2$   
**表 4.** 经验性功效 ( $J = 2$ )

$\theta$	$\pi_1$	$\rho$	$m = (25, 25)$			$m = (50, 50)$			$m = (100, 100)$		
			$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$
0.2	I	A	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	0.9999	0.9999	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	0.9999	0.9999	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	II	A	0.9997	0.9999	0.9997	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	0.9997	0.9999	0.9997	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
III	A	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
	B	0.9999	1.0000	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
	C	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
0.4	I	A	0.9355	0.9716	0.9344	0.9980	0.9996	0.9980	1.0000	1.0000	1.0000
		B	0.9196	0.9655	0.9185	0.9976	0.9993	0.9976	1.0000	1.0000	1.0000



续表

0.4	I	C	0.9258	0.9671	0.9246	0.9986	0.9993	0.9985	1.0000	1.0000	1.0000
		A	0.9382	0.9733	0.9369	0.9986	0.9995	0.9986	1.0000	1.0000	1.0000
	II	B	0.9245	0.9653	0.9235	0.9980	0.9991	0.9980	1.0000	1.0000	1.0000
		C	0.9308	0.9671	0.9297	0.9975	0.9994	0.9975	1.0000	1.0000	1.0000
	III	A	0.9390	0.9727	0.9381	0.9986	0.9996	0.9986	1.0000	1.0000	1.0000
		B	0.9304	0.9698	0.9295	0.9987	0.9999	0.9987	1.0000	1.0000	1.0000
0.6	I	C	0.9328	0.9698	0.9322	0.9981	0.9991	0.9981	1.0000	1.0000	1.0000
		A	0.4930	0.6386	0.4899	0.7652	0.8498	0.7641	0.9729	0.9833	0.9728
		B	0.4693	0.6276	0.4671	0.7623	0.8439	0.7610	0.9630	0.9776	0.9629
	II	C	0.4793	0.6303	0.4754	0.7669	0.8446	0.7659	0.9695	0.9834	0.9694
		A	0.5025	0.6497	0.4989	0.7919	0.8653	0.7907	0.9737	0.9838	0.9735
		B	0.4819	0.6290	0.4792	0.7753	0.8514	0.7747	0.9689	0.9811	0.9688
	III	C	0.4933	0.6444	0.4900	0.7829	0.8599	0.7819	0.9716	0.9819	0.9716
		A	0.4947	0.6387	0.4907	0.7965	0.8683	0.7953	0.9742	0.9850	0.9740
		B	0.4832	0.6389	0.4817	0.7743	0.8547	0.7727	0.9730	0.9841	0.9728
		C	0.4880	0.6351	0.4861	0.7797	0.8555	0.7789	0.9719	0.9830	0.9718

健;  $T_{LR}$  与  $T_{SC}$  相比较为冒进, 总体表现较为稳健。因此, 基于第一类错误率的表现, 最推荐的检验统计量为  $T_{SC}$ 。

在参数配置相同的备择假设  $H_1$  下, 计算三个统计量的经验性功效。具体结果见表 4 和附表 3、附表 4。可以看出  $T_{LR}$ ,  $T_W$  和  $T_{SC}$  的功效随着  $J$  和  $m$  的增大而增大, 且在样本量较大的情况下表现更为良好。

### 3.3. 样本量的确定

由检验结果可知, 样本量对检验统计量的第一类错误率和功效有显著影响, 如果能提前预知在固定的显著性水平下达到期望功效所需的样本总量, 就能使后续进行的试验有更稳健的第一类错误率和更好的功效。于是, 提出了检验统计量  $T_{LR}$ ,  $T_W$ ,  $T_{SC}$  的渐近样本量计算方法。

在  $H_1$  假设下, 检验统计量  $T_{LR}$ ,  $T_{LW}$ ,  $T_{SC}$  达到特定功效所需的渐近样本量可由下式计算得出:

$$\Pr(T_{LR} \geq \chi_{1-\alpha}^2(J-1) | H_{11}) = 1 - \beta,$$

$$\Pr(T_W \geq \chi_{1-\alpha}^2(J-1) | H_{11}) = 1 - \beta,$$

$$\Pr(T_{SC} \geq \chi_{1-\alpha}^2(J-1) | H_{11}) = 1 - \beta.$$

由于上述式子不存在显式解, 因此可通过迭代的方法对渐进样本量大小进行求解:

样本总量  $N = \sum_{j=1}^J N_j$ ,  $N_j = m_{++j}^{(1)} + m_{++j}^{(2)}$ ,  $N_j = r_j N$ ,  $m_{++j}^{(1)} = t_j N$ ,  $m_{++j}^{(2)} = t_j N_j$ ,  $j = 1, 2, \dots, J$ 。在  $H_1$  假设条件下, 通过设置  $r_j, t_j, \pi_{ij}, \rho_j$  的参数值, 随机生成样本量  $N$ ;

基于(1)中生成的样本量  $N$ , 代入上式中求解得到渐近功效;

若步骤(2)中所得到的渐近功效小于所期望的功效  $1 - \beta$ , 则增加样本量  $N$ ; 若所得到的渐近功效大于  $1 - \beta$ , 则减小样本量  $N$ 。最终使得渐近功效接近  $1 - \beta$ 。

令  $r_j = \frac{1}{J}$ ,  $t_j = 0.5$ ,  $\pi_{1j} = (0.3, 0.5)$ ,  $\theta = (\theta_0, \theta_j)$ , 由于当  $\theta = 1$  时表示一层间的两组治疗效果相同, 因此选择  $\theta_0 = 1$ 。利用上述迭代方法, 求得  $\theta_j$  或  $\rho_j$  变化时, 达到显著性水平  $\alpha = 0.05$  或功效  $1 - \beta = 0.8$  所需的渐近样本量(具体结果如表 5 所示)。

**Table 5.** Sample size with 80% power  
**表 5.** 80% 功效下的渐进样本量

$\theta$	$\pi_1$	$\rho$	$J = 2$			$J = 4$			$J = 6$		
			$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$
0.2	I	A	48	32	48	64	48	64	72	72	72
		B	48	32	48	64	48	64	72	72	72
		C	48	32	48	64	48	64	72	72	72
	II	A	48	32	56	64	48	64	72	72	72
		B	48	32	48	64	48	64	72	72	72
		C	48	32	48	64	48	64	72	72	72
	III	A	48	32	48	64	32	64	72	72	72
		B	48	32	48	64	32	64	72	72	72
		C	48	32	48	64	32	64	72	72	72
0.4	I	A	128	88	128	128	96	128	144	144	144
		B	136	96	144	144	96	144	144	144	144
		C	136	88	136	128	96	128	144	144	144
	II	A	128	88	136	128	96	128	144	144	144
		B	136	88	136	128	96	128	144	144	144
		C	128	96	136	128	96	128	144	144	144
	III	A	128	80	128	128	96	128	144	144	144
		B	136	88	136	128	96	128	144	144	144
		C	136	88	136	128	96	128	144	144	144
0.6	I	A	424	344	416	416	320	416	432	432	432
		B	440	352	432	432	352	432	432	432	432
		C	432	344	424	432	352	416	432	432	432
	II	A	408	296	424	416	320	432	408	408	408
		B	416	344	432	432	336	432	432	432	432
		C	424	336	424	416	336	416	408	408	408
	III	A	408	328	424	416	320	416	432	432	432
		B	432	344	432	416	352	432	408	408	408
		C	416	336	408	416	320	416	408	408	408

#### 4. 实证研究

Mandel 等(1982) [12]进行了一项研究双盲随机对照试验, 该试验对 214 名患有中耳炎的儿童随机分配服用头孢克洛或阿莫西林进行治疗, 以比较这两种抗生素的效果。经过 14 天的治疗试验, 共有 203 例有效的治疗数据, 由表 6 所示。

基于表 6 中的数据进行实证分析来展示上面所提到的所有检验以及确定样本量的方法，进而得到一定的应用价值。假设  $H_0: \theta_1 = \theta_2 = \theta_3 \triangleq \theta$ ， $H_1: \theta \neq \theta_0$ ，来检验不同年龄下哪种抗生素的治疗效果较好。通过样本量估计，Fisher 得分法和牛顿迭代算法，我们可以得到样本量估计值  $\pi_{1j}^{(0)}, \pi_{2j}^{(0)}, \rho_j^{(0)}$  全局最大似然估计量  $\hat{\pi}_1, \hat{\pi}_2, \hat{\rho}_j$ ， $H_0$  假设下的最大似然估计量  $\tilde{\pi}_1, \tilde{\pi}_2, \rho_j$  结果汇总在表 7 中。

**Table 6.** The number of children after receiving treatment  
**表 6.** 各年龄层患病儿童接受治疗后改善情况

改善个数	年龄 < 2 岁				年龄 2~5 岁				年龄 ≥ 6 岁			
	头孢克洛		阿莫西林		头孢克洛		阿莫西林		头孢克洛		阿莫西林	
	单边	双边	单边	双边	单边	双边	单边	双边	单边	双边	单边	双边
0	10	11	9	8	22	3	7	6	7	1	8	0
1	2	2	3	2	14	1	24	6	11	0	11	1
2	-	2	-	8	-	5	-	10	-	6	-	3

**Table 7.** Maximum likelihood estimator of each parameter

**表 7.** 各参数的最大似然估计量

年龄	全局 MLEs			$H_0$ 假设下的 MLEs		
	$\hat{\pi}_1$	$\hat{\pi}_2$	$\hat{\rho}$	$\tilde{\pi}_{1j}$	$\tilde{\pi}_{2j}$	$\tilde{\rho}_j$
<2 岁	0.1928	0.5000	0.7085	0.3036	0.3036	0.7307
2-5 岁	0.4421	0.6830	0.5375	0.5783	0.5783	0.5334
≥6 岁	0.6911	0.6329	0.7534	0.6626	0.6626	0.7560

基于三类最大似然估计量构造检验统计量，三种检验统计量和对应 p 值如表 8 所示。

**Table 8.** Statistic values and p-value of homogeneity test

**表 8.** 同质性检验下的统计值和 p 值

检验统计量	$T_{LR}$	$T_W$	$T_{SC}$
统计值	6.4373	6.2064	6.3841
p 值	0.0112	0.0119	0.0115

结果表明，在原假设  $H_0$  下，p 值均小于 0.05，认为拒绝原假设，各层的优势比并不等于 1。优势比近似等于 0.2389，表示两种抗生素对于同一年龄段的患者的治疗效果不同，头孢克洛对儿童中耳炎有更好的治疗效果。然后根据提出的方法确定样本量。

**Table 9.** Estimated sample size under the desired power

**表 9.** 目标功效下的样本量

目标功效	$T_{LR}$	$T_W$	$T_{SC}$
80%	312	588	324
90%	620	712	708

再根据所提出的方法进行样本量的确定。考虑  $r_1 = r_2 = r_3 = \frac{1}{3}$ ， $t_1 = t_2 = 0.5$ ，参数估计值与上述相应

的最大似然估计值相同。在期望功效为 80% 和 90% 的情况下, 表 9 给出了显著性水平为 0.05 条件下的近似样本量。结果表明, 在同质性假设下,  $T_{LR}$  达到期望功效所需的样本量最少,  $T_w$  达到期望功效所需的样本量最大, 因此最推荐使用似然比检验。

根据以上结果, 表明头孢克洛对儿童中耳炎有更好的治疗效果, 因此更推荐使用头孢克洛治疗儿童中耳炎。而分层效应不明显, 说明年龄对儿童中耳炎的治疗影响不大。此外, 根据估计, 选取 312 名儿童急性中耳炎患者能够在研究不同治疗方案对治疗效果的影响时, 使用可能少的人力成本下达到期望的功效。

## 5. 总结

本文对于分层的单双边混合配对数据各层的优势比构造检验统计量进行同质性检验并估计渐近样本量的大小。首先, 构造基于 Donner 模型的同质性假设下参数的三种最大似然估计量。构造同质性假设下的最大似然估计量。利用 Fisher 得分算法迭代得到所需要的最大似然估计量。再通过估计量构造似然比检验统计量( $T_{LR}$ )、Wald 检验统计量( $T_w$ )和 score 检验统计量( $T_{SC}$ ), 选择第一类错误率和功效作为判断指标, 对三种统计量进行比较。蒙特卡洛模拟结果显示, 在同质性假设下,  $T_w$  在三个检验统计量中表现最为冒进;  $T_{LR}$  相对冒进, 总体表现较为稳健; 而  $T_{SC}$  在各种情况下均表现稳健, 且功效较高, 因此最推荐的检验统计量为  $T_{SC}$ 。对于临床试验中样本量的确定, 为了达到相同的功效,  $T_w$  所需样本量较小; 随着层数  $J$  的增加, 三个检验统计量所需样本量差异逐渐减小; 随着优势比的增加, 越来越接近于 1, 三个检验统计量所需样本量也会逐渐增加。

## 参考文献

- [1] Qiu, S.F. and Tao, J.R. (2022) Confidence Intervals for Assessing Equivalence of Two Treatments with Combined Unilateral and Bilateral Data. *Journal of Applied Statistics*, **49**, 3414-3435. <https://doi.org/10.1080/02664763.2021.1949440>
- [2] Ronser, B. (1982) Statistical Methods in Ophthalmology: An Adjustment for the Intraclass Correlation between Eyes. *Biometrics*, **38**, 105-114. <https://doi.org/10.2307/2530293>
- [3] Donner, A. (1989) Statistical Methods in Ophthalmology: An Adjusted Chi-Square Approach. *Biometrics*, **45**, 605-611. <https://doi.org/10.2307/2531501>
- [4] Thompson, J.R. (1993) The  $\chi^2$  Test for Data Collected on Eyes. *British Journal of Ophthalmology*, **77**, 115-117. <https://doi.org/10.1136/bjo.77.2.115>
- [5] Ma, C.X. and Wang, K.J. (2021) Testing the Homogeneity of Proportions for Combined Unilateral and Bilateral Data. *Journal of Biopharmaceutical Statistics*, **31**, 686-704. <https://doi.org/10.1080/10543406.2021.1946693>
- [6] Kahan, B.C. and Morris, T.P. (2013) Analysis of Multicentre Trials with Continuous Outcomes: When and How Should We Account for Centre Effects? *Statistics in Medicine*, **32**, 1136-1149. <https://doi.org/10.1002/sim.5667>
- [7] Sun, S.M., Li, Z.M. and Jiang, H.J. (2022) Homogeneity Test and Sample Size of Risk Difference for Stratified Unilateral and Bilateral Data. *Communications in Statistics-Simulation and Computation*, 1-24. <https://doi.org/10.1080/03610918.2022.2142240>
- [8] Nam, J.M. (1998) Power and Sample Size for Stratified Prospective Studies Using the Score Method for Testing Relative Risk. *Biometrics*, **54**, 331-336. <https://doi.org/10.2307/2534020>
- [9] Li, Z.M. and Ma, C.X. (2022) Statistical Inference for Odds Ratio of Two Proportions in Bilateral Correlated Data. *Axioms*, **11**, Article 502. <https://doi.org/10.3390/axioms11100502>
- [10] 靳云汇, 金赛男. 高级计量经济学[M]. 北京: 北京大学出版社, 2007.
- [11] Tang, M.L., Tang, N.S. and Rosner, B. (2006) Statistical Inference for Correlated Data in Ophthalmologic Studies. *Statistics in Medicine*, **25**, 2771-2783. <https://doi.org/10.1002/sim.2425>
- [12] Mandel, E.M., Bluestone, C.D., Rockette, H.E., Blatter, M.M., Reisinger, K.S., Wucher, F.P. and Harper, J. (1982) Duration of Effusion after Antibiotic Treatment for Acute Otitis Media: Comparison of Cefaclor and Amoxicillin. *Pediatric Infectious Disease Journal*, **1**, 310-316. <https://doi.org/10.1097/00006454-198209000-00006>

## 附录

Table A1. The empirical TIEs (%) for  $J = 4$ 附表 1. 经验性第一类错误率( $J = 4$ )

$\theta$	$\pi_1$	$\rho$	$m = (25, 25)$			$m = (50, 50)$			$m = (100, 100)$		
			$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$
0.2	I	A	0.0528	0.0628	0.0512	0.0521	0.0567	0.0507	0.0481	0.0524	0.0480
		B	0.0555	0.0663	0.0540	0.0495	0.0573	0.0481	0.0540	0.0544	0.0537
		C	0.0532	0.0636	0.0522	0.0545	0.0576	0.0539	0.0505	0.0526	0.0502
	II	A	0.0531	0.0631	0.0520	0.0495	0.0587	0.0485	0.0454	0.0492	0.0456
		B	0.0556	0.0653	0.0536	0.0559	0.0615	0.0545	0.0577	0.0593	0.0571
		C	0.0558	0.0636	0.0532	0.0526	0.0594	0.0522	0.0542	0.0583	0.0535
	III	A	0.0525	0.0637	0.0511	0.0495	0.0566	0.0486	0.0498	0.0537	0.0485
		B	0.0552	0.0624	0.0547	0.0511	0.0563	0.0500	0.0509	0.0548	0.0500
		C	0.0537	0.0683	0.0521	0.0534	0.0604	0.0521	0.0514	0.0558	0.0506
0.4	I	A	0.0502	0.0597	0.0495	0.0497	0.0531	0.0496	0.0519	0.0548	0.0517
		B	0.0521	0.0629	0.0510	0.0523	0.0557	0.0516	0.0507	0.0504	0.0506
		C	0.0520	0.0603	0.0504	0.0524	0.0562	0.0518	0.0512	0.0532	0.0508
	II	A	0.0523	0.0622	0.0511	0.0497	0.0546	0.0494	0.0486	0.0516	0.0485
		B	0.0520	0.0604	0.0502	0.0531	0.0596	0.0529	0.0509	0.0542	0.0510
		C	0.0531	0.0611	0.0524	0.0507	0.0538	0.0503	0.0516	0.0524	0.0513
	III	A	0.0551	0.0615	0.0534	0.0529	0.0590	0.0523	0.0474	0.0488	0.0479
		B	0.0507	0.0623	0.0502	0.0450	0.0529	0.0447	0.0504	0.0524	0.0502
		C	0.0528	0.0625	0.0510	0.0482	0.0533	0.0479	0.0517	0.0513	0.0516
0.6	I	A	0.0490	0.0559	0.0484	0.0487	0.0490	0.0487	0.0520	0.0536	0.0518
		B	0.0544	0.0579	0.0540	0.0466	0.0528	0.0460	0.0531	0.0544	0.0531
		C	0.0514	0.0605	0.0508	0.0498	0.0534	0.0492	0.0502	0.0510	0.0500
	II	A	0.0529	0.0579	0.0529	0.0527	0.0567	0.0525	0.0495	0.0504	0.0495
		B	0.0542	0.0592	0.0533	0.0539	0.0576	0.0532	0.0480	0.0509	0.0477
		C	0.0478	0.0552	0.0473	0.0513	0.0532	0.0512	0.0477	0.0501	0.0477
	III	A	0.0549	0.0636	0.0543	0.0487	0.0548	0.0485	0.0493	0.0515	0.0494
		B	0.0535	0.0596	0.0529	0.0468	0.0523	0.0461	0.0538	0.0551	0.0532
		C	0.0478	0.0558	0.0467	0.0498	0.0532	0.0493	0.0470	0.0501	0.0466

**Table A2.** The empirical TIEs (%) for J = 6  
**附表 2.** 经验性第一类错误率(J = 6)

$\theta$	$\pi_1$	$\rho$	$m = (25, 25)$			$m = (50, 50)$			$m = (100, 100)$		
			$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$
0.2	I	A	0.0500	0.0622	0.0488	0.0511	0.0570	0.0500	0.0475	0.0503	0.0477
		B	0.0506	0.0621	0.0492	0.0484	0.0541	0.0479	0.0515	0.0537	0.0521
		C	0.0540	0.0637	0.0523	0.0539	0.0572	0.0535	0.0530	0.0546	0.0517
	II	A	0.0519	0.0623	0.0511	0.0511	0.0587	0.0497	0.0448	0.0459	0.0449
		B	0.0527	0.0630	0.0503	0.0545	0.0583	0.0539	0.0514	0.0535	0.0509
		C	0.0516	0.0600	0.0507	0.0486	0.0575	0.0479	0.0507	0.0548	0.0503
	III	A	0.0568	0.0676	0.0555	0.0549	0.0597	0.0543	0.0499	0.0541	0.0489
		B	0.0526	0.0647	0.0515	0.0538	0.0578	0.0532	0.0495	0.0555	0.0497
		C	0.0510	0.0632	0.0494	0.0545	0.0586	0.0537	0.0483	0.0500	0.0474
0.4	I	A	0.0518	0.0597	0.0507	0.0513	0.0546	0.0510	0.0494	0.0522	0.0493
		B	0.0536	0.0611	0.0518	0.0506	0.0557	0.0502	0.0538	0.0573	0.0541
		C	0.0521	0.0594	0.0512	0.0497	0.0554	0.0496	0.0506	0.0508	0.0503
	II	A	0.0539	0.0580	0.0527	0.0508	0.0530	0.0506	0.0492	0.0504	0.0488
		B	0.0562	0.0633	0.0554	0.0513	0.0571	0.0508	0.0506	0.0516	0.0505
		C	0.0561	0.0601	0.0554	0.0497	0.0508	0.0499	0.0488	0.0519	0.0485
	III	A	0.0523	0.0586	0.0512	0.0535	0.0563	0.0527	0.0514	0.0527	0.0513
		B	0.0533	0.0566	0.0524	0.0475	0.0529	0.0471	0.0480	0.0507	0.0483
		C	0.0523	0.0587	0.0511	0.0518	0.0550	0.0512	0.0517	0.0534	0.0519
0.6	I	A	0.0549	0.0610	0.0538	0.0507	0.0530	0.0505	0.0510	0.0546	0.0512
		B	0.0491	0.0517	0.0483	0.0506	0.0520	0.0501	0.0515	0.0535	0.0514
		C	0.0546	0.0587	0.0541	0.0516	0.0520	0.0513	0.0495	0.0517	0.0491
	II	A	0.0532	0.0600	0.0520	0.0529	0.0558	0.0525	0.0508	0.0524	0.0506
		B	0.0513	0.0560	0.0509	0.0474	0.0510	0.0470	0.0489	0.0504	0.0484
		C	0.0508	0.0571	0.0505	0.0524	0.0544	0.0521	0.0519	0.0556	0.0520
	III	A	0.0508	0.0558	0.0499	0.0497	0.0540	0.0491	0.0503	0.0509	0.0500
		B	0.0542	0.0591	0.0537	0.0529	0.0554	0.0528	0.0515	0.0530	0.0515
		C	0.0510	0.0571	0.0503	0.0520	0.0552	0.0521	0.0476	0.0477	0.0474

**Table A3.** The empirical powers (%) for J = 4  
**附表 3.** 经验性功效(J = 4)

$\theta$	$\pi_1$	$\rho$	$m = (25, 25)$			$m = (50, 50)$			$m = (100, 100)$		
			$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$
0.2	I	A	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	II	A	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	III	A	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.4	I	A	0.9984	0.9991	0.9984	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	0.9971	0.9990	0.9971	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	0.9982	0.9994	0.9981	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	II	A	0.9981	0.9992	0.9981	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	0.9980	0.9994	0.9980	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	0.9980	0.9991	0.9980	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	III	A	0.9991	0.9999	0.9991	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	0.9982	0.9990	0.9981	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	0.9987	0.9996	0.9987	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.6	I	A	0.7786	0.8532	0.7764	0.9686	0.9814	0.9682	0.9997	0.9999	0.9997
		B	0.7702	0.8523	0.7689	0.9692	0.9808	0.9690	0.9997	0.9999	0.9997
		C	0.7748	0.8488	0.7733	0.9660	0.9789	0.9654	0.9995	0.9995	0.9995
	II	A	0.7855	0.8627	0.7837	0.9768	0.9852	0.9766	1.0000	1.0000	1.0000
		B	0.7725	0.8472	0.7709	0.9699	0.9834	0.9696	0.9995	0.9997	0.9995
		C	0.7801	0.8589	0.7783	0.9703	0.9834	0.9700	0.9998	0.9999	0.9998
	III	A	0.7981	0.8700	0.7963	0.9711	0.9839	0.9708	0.9998	0.9998	0.9998
		B	0.7673	0.8463	0.7656	0.9694	0.9813	0.9693	0.9999	1.0000	0.9999
		C	0.7882	0.8617	0.7863	0.9742	0.9856	0.9739	0.9999	1.0000	0.9999

**Table A4.** The empirical powers (%) for J = 6  
**附表 4.** 经验性功效(J = 6)

$\theta$	$\pi_1$	$\rho$	$m = (25, 25)$			$m = (50, 50)$			$m = (100, 100)$		
			$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$	$T_{LR}^C$	$T_W^C$	$T_{SC}^C$
0.2	I	A	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	II	A	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	III	A	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.4	I	A	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	II	A	0.9999	1.0000	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	0.9999	0.9999	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	III	A	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		B	0.9999	0.9999	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		C	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.6	I	A	0.9133	0.9461	0.9125	0.9960	0.9984	0.9960	1.0000	1.0000	1.0000
		B	0.9041	0.9371	0.9034	0.9957	0.9981	0.9956	1.0000	1.0000	1.0000
		C	0.9103	0.9442	0.9100	0.9966	0.9980	0.9963	1.0000	1.0000	1.0000
	II	A	0.9277	0.9564	0.9270	0.9971	0.9981	0.9971	1.0000	1.0000	1.0000
		B	0.9189	0.9493	0.9183	0.9967	0.9984	0.9967	1.0000	1.0000	1.0000
		C	0.9194	0.9503	0.9188	0.9977	0.9987	0.9977	0.9999	0.9999	0.9999
	III	A	0.9223	0.9539	0.9216	0.9970	0.9982	0.9970	1.0000	1.0000	1.0000
		B	0.9152	0.9432	0.9144	0.9970	0.9984	0.9970	1.0000	1.0000	1.0000
		C	0.9215	0.9525	0.9212	0.9979	0.9986	0.9978	1.0000	1.0000	1.0000