

双边试验设计下基于优势比的区间估计

李红琼^{1,2}, 江绍萍^{1,*}, 喻扬涛¹, 赵硕¹, 董竹月¹

(1. 云南民族大学 数学与计算机科学学院, 云南昆明 650500;

2. 楚雄师范学院 数学与统计学院, 云南楚雄 675000)

摘要: 针对双边试验设计下优势比的置信区间估计问题, 基于独立和非独立假设下的双边数据, 提出了四种基于Wald统计量、对数-Wald统计量的置信区间以及MOVER型置信区间和Bootstrap重抽样置信区间. 通过置信区间的经验覆盖概率、置信下限的均值、达到固定功效试验所需的样本量以及此样本下的经验覆盖概率的模拟研究, 比较各置信区间的统计性能. 结果表明: 基于样本方差的对数-Wald统计量的置信区间具有令人满意的结果, 它既能保证经验覆盖概率接近置信水平, 且在相同参数设置下达到固定功效时所需的样本量较少. 最后, 结合耳科临床数据对所提方法进行验证.

关键词: 双边试验设计; 优势比; 置信区间; 经验覆盖概率; 样本量

中图分类号: O212

文献标识码: A **文章编号:** 1000-4424(2026)01-0037-18

§1 引言

在生物医学和流行病学研究中, 经常遇到成对身体器官(如耳朵、眼睛、肾脏)或身体成对部分(如手臂和脚)患病, 进行治疗产生的数据, 即双边数据. 这样的双边数据具有组内相关性, 忽略它们之间的相关性进行统计推断, 会得到不合理的结果^[1-3]. 此外, 对于具有组内相关性的双边数据, 人们感兴趣的问题是两种不同的治疗药物(方法)对于疾病的治愈率是否有显著性差异. 判断不同治疗药物的显著性差异时, 等价性和非劣效性检验是两种重要的统计方法. 进行等价性或非劣效性检验, 常用的指标主要有: 比例差(Rate Difference, RD)、比例比(Relative Risk, RR)以及优势比(Odds Ratio, OR). 虽然RD与RR能够直观地反映两个治疗组之间的绝对或相对差异, 但对于RD而言在进行等价性或非劣效性检验时, 需要设定固定的边际值, 这种边际选择往往较为困难, 且可能因选择不当而影响检验结果的准确性和可靠性; 对于RR而言其对事件发生率的假设和样本量要求都较高, 当事件发生率较低时, 比例比的估计可能不准确. 与RD, RR相比, 虽然OR的解释相对复杂, 但其适用性广, 无需固定边际假设.

收稿日期: 2023-12-08 修回日期: 2025-03-20

*通讯作者, E-mail: jiang2005124@163.com

基金项目: 国家自然科学基金(62362071); 云南民族大学研究生科研基金(2022SKY085)

对于具有组内相关性的双边数据, Rosner^[1]在独立和非独立模型下研究了两种治疗的等价性评价问题; Tang等人^[4]在独立和非独立模型下基于比例差(治愈率之差)提出了八种检验统计量和对应的近似非条件检验过程; Tang等人^[5]对比例差的区间估计问题进行了研究; Zhuang等人^[6]在相关系数相等的模型下, 提出了基于比例比(治愈率之比)的五种置信区间; 邱世芳等人^[7]在独立和非独立模型下从区间估计角度, 给出了样本量的近似公式. 以上的研究都是在双边数据下对两种治疗基于比例差或比例比的等价性检验, 而对于双边数据下基于优势比的非劣效性检验却少有研究. 因此, 本文在双边试验设计下, 基于优势比从区间估计的角度研究两种治疗药物(方法)的非劣效性评价问题.

§2 数据结构和统计模型

假设有 N 个成对器官或身体成对部分都需要治疗的病人被随机分配到两个治疗组中: 新治疗药物(或方法, $i = 0$) 和标准治疗药物(或方法, $i = 1$), 且两种治疗药物间相互独立. 换言之, 对每一组都有三种可能的治疗结果, 一是全部治愈(成对器官都治愈), 二是部分治愈(成对器官中只有一只被治愈), 三是未治愈, 此试验数据可归结为如表1的数据.

表 1 双边试验设计下的观测频数及概率

成对器官的治愈只数	治疗药物 (方法)	
	$i = 0$	$i = 1$
$k = 0$	n_{00} (p_{00})	n_{01} (p_{01})
$k = 1$	n_{10} (p_{10})	n_{11} (p_{11})
$k = 2$	n_{20} ($p_{20} = 1 - p_{00} - p_{10}$)	n_{21} ($p_{21} = 1 - p_{01} - p_{11}$)
合计	n_{+0} (1.0)	n_{+1} (1.0)

表中 n_{ki} ($k = 0, 1, 2; i = 0, 1$)表示接受第 i 种治疗药物后有 k 只器官治愈的病人个数, p_{ki} 表示相应的概率, 且 $n_{+i} = n_{0i} + n_{1i} + n_{2i}$, $p_{0i} + p_{1i} + p_{2i} = 1$, $N = n_{+0} + n_{+1}$. 根据文献[1], 假设在治疗结束时, 接受第 i 种药物治疗的第 j 个病人有 k 只耳被治愈记为 $z_{ijk} = 1$ ($i = 0, 1; j = 1, 2, \dots, n_{+i}; k = 1, 2$), 否则 $z_{ijk} = 0$. 设 $p(z_{ijk} = 1) = \lambda_i$ 且 $p(z_{ijk} = 1 | z_{ij,3-k} = 1) = R\lambda_i$, 其中 R 是一个大于0的常数, 用来衡量同一个患者的两只成对器官相关性的指标, 如果 $R = 1$, 则两只成对器官完全独立; 若 $R\lambda_i = 1$, 则两只成对器官完全非独立. 基于以上的假设, 有

$$p_{0i} = 1 + R\lambda_i^2 - 2\lambda_i, \quad p_{1i} = 2\lambda_i(1 - R\lambda_i), \quad p_{2i} = R\lambda_i^2; \quad i = 0, 1. \quad (1)$$

为了检验新治疗药物是否不劣于标准治疗药物, 本文基于优势比进行非劣效性检验. 优势比定义为

$$\delta = \frac{1 - p_{00} - p_{10}/2}{p_{00} + p_{10}/2} / \frac{1 - p_{01} - p_{11}/2}{p_{01} + p_{11}/2} = \frac{\lambda_0(1 - \lambda_1)}{\lambda_1(1 - \lambda_0)}, \quad (2)$$

由(2)有

$$\lambda_0 = \frac{\delta\lambda_1}{1 + (\delta - 1)\lambda_1}. \quad (3)$$

建立非劣效假设检验问题

$$H_0: \delta \leq \delta_0 \quad \text{versus} \quad H_1: \delta > \delta_0, \quad (4)$$

其中 $\delta_0(0 < \delta_0 \leq 1)$ 是事先给定的非劣效性边界值, 如果 $\delta \leq \delta_0$, 则新治疗药物疗效劣于标准药物, 反之则不劣于标准药物.

基于观测数据 $(n_{00}, n_{01}, n_{10}, n_{11}, n_{20}, n_{21})$ 的对数似然函数可表示为 $\theta = (\delta, \lambda_1, R)^\top$ 的函数, 即有

$$\begin{aligned} \ln l(\theta) = & C + n_{00} \ln[(1 - \lambda_1)^2 + \delta^2 \lambda_1^2 (R - 1)] - 2(n_{00} + n_{10} + n_{20}) \ln[1 + (\delta - 1)\lambda_1] + \\ & n_{01} \ln(1 + R\lambda_1^2 - 2\lambda_1) + n_{10} \ln[1 - \lambda_1 + \delta\lambda_1(1 - R)] + n_{11} \ln(1 - R\lambda_1) + \\ & (n_{10} + 2n_{20}) \ln \delta + (n_{10} + n_{11} + 2n_{20} + 2n_{21}) \ln \lambda_1 + (n_{20} + n_{21}) \ln R, \end{aligned} \quad (5)$$

其中 C 是与 δ, λ_1, R 无关的常数, δ 为感兴趣参数, λ_1 与 R 为讨厌参数, 且 $0 \leq \lambda_1 \leq 1, 0 \leq \delta \leq \infty$, $\max\{0, [\delta^2 \lambda_1^2 - (\lambda_1 - 1)^2]/\delta^2 \lambda_1^2, (2\lambda_1 - 1)/\lambda_1^2\} \leq R \leq \min\{[1 + (\delta - 1)\lambda_1]/\delta\lambda_1, 1/\lambda_1\}$ (推导过程见附录). 本文主要在独立($R = 1$)与非独立($R \neq 1$)假设下基于优势比 δ 进行置信区间的构造, 提出有效的区间估计方法.

§3 参数估计及置信区间的构造

3.1 参数估计及方差表达式

设 $\hat{\lambda}_0$ 和 $\hat{\lambda}_1$ 为参数 λ_0 和 λ_1 的极大似然估计(MLE), 且 $\hat{\lambda}_0 = (n_{10} + 2n_{20})/2n_{+0}$ 和 $\hat{\lambda}_1 = (n_{11} + 2n_{21})/2n_{+1}$, 从而可以得到参数 δ 的极大似然估计为

$$\hat{\delta} = \frac{(2n_{01} + n_{11})(n_{10} + 2n_{20})}{(2n_{00} + n_{10})(n_{11} + 2n_{21})}.$$

通过(1)中的 $p_{2i} = R\lambda_i^2 (i = 0, 1)$, 可得 R 的极大似然估计为

$$\hat{R} = \frac{n_{20}/n_{+0} + n_{21}/n_{+1}}{\hat{\lambda}_0^2 + \hat{\lambda}_1^2} = \frac{4n_{+0}n_{+1}(n_{20}n_{+1} + n_{21}n_{+0})}{n_{+1}^2(n_{10} + 2n_{20})^2 + n_{+0}^2(n_{11} + 2n_{21})^2}.$$

此外, 可以求解 $H_0: \delta \leq \delta_0$ 条件下讨厌参数的限制性极大似然估计(REML). 在非独立假设下 $\tilde{\lambda}_1$ 与 \tilde{R} 为参数 λ_1 和 R 的限制性极大似然估计, $\tilde{\lambda}_1$ 和 \tilde{R} 通过迭代法求解方程组

$$\begin{cases} \left. \begin{aligned} \frac{\partial \ln l(\theta)}{\partial \lambda_1} \Big|_{\delta=\delta_0} &= \frac{2n_{00}[\lambda_1 - 1 + \delta^2 \lambda_1 (R - 1)]}{(1 - \lambda_1)^2 + \delta^2 \lambda_1^2 (R - 1)} - \frac{2(n_{00} + n_{10} + n_{20})(\delta - 1)}{1 + (\delta - 1)\lambda_1} + \frac{2n_{01}(R\lambda_1 - 1)}{1 + R\lambda_1^2 - 2\lambda_1} + \\ & \frac{n_{10}[\delta(1 - R) - 1]}{1 - \lambda_1 + \delta\lambda_1(1 - R)} - \frac{n_{11}R}{1 - R\lambda_1} + \frac{n_{10} + n_{11} + 2n_{20} + 2n_{21}}{\lambda_1} = 0; \\ \frac{\partial \ln l(\theta)}{\partial R} \Big|_{\delta=\delta_0} &= \frac{n_{00}\delta^2 \lambda_1^2}{(1 - \lambda_1)^2 + \delta^2 \lambda_1^2 (R - 1)} + \frac{n_{01}\lambda_1^2}{1 + R\lambda_1^2 - 2\lambda_1} - \frac{n_{10}\delta\lambda_1}{1 - \lambda_1 + \delta\lambda_1(1 - R)} - \\ & \frac{n_{11}\lambda_1}{1 - R\lambda_1} + \frac{n_{20} + n_{21}}{R} = 0 \end{aligned} \right\}$$

得到. 而在独立假设下 $\tilde{\lambda}_1$ 通过求解方程

$$\begin{aligned} \frac{\partial \ln l(\theta)}{\partial \lambda_1} \Big|_{\delta=\delta_0} &= \frac{2n_{00} + 2n_{01} + n_{10} + n_{11}}{\lambda_1 - 1} - \frac{2(n_{00} + n_{10} + n_{20})(\delta - 1)}{1 + (\delta - 1)\lambda_1} + \\ & \frac{n_{10} + n_{11} + 2n_{20} + 2n_{21}}{\lambda_1} = 0 \end{aligned}$$

得到.

以往研究中通常采用delta方法求解感兴趣参数的方差^[8-11],但delta方法是一种近似求解的方法,得到的结果带有一定的偏差.为了避免这种偏差,文中采用Fisher-score方法求解感兴趣参数的方差^[12-14](推导过程见附录)为

$$\text{Var}(\hat{\delta}) = \begin{cases} \frac{I_{22}I_{33} - I_{23}^2}{I_{11}I_{22}I_{33} + 2I_{12}I_{13}I_{23} - I_{11}I_{23}^2 - I_{13}^2I_{22} - I_{12}^2I_{33}} \triangleq \Sigma_1(\theta_1), & R \neq 1; \\ \frac{\delta[1 + (\delta - 1)\lambda_1]^2\{2n_{+0}\delta + 2n_{+1}[1 + (\delta - 1)\lambda_1]^2\}}{4n_{+0}n_{+1}\lambda_1\{(1 - \lambda_1)^3 + \delta\lambda_1[\lambda_1^2(2 - \delta) + \lambda_1(\delta - 4) + 2]\}} \triangleq \Sigma_1(\theta_2), & R = 1. \end{cases} \quad (6)$$

从而得到

$$\text{Var}(\log \hat{\delta}) = \begin{cases} \frac{I_{22}I_{33} - I_{23}^2}{\delta^2(I_{11}I_{22}I_{33} + 2I_{12}I_{13}I_{23} - I_{11}I_{23}^2 - I_{13}^2I_{22} - I_{12}^2I_{33})} \triangleq \Sigma_2(\theta_1), & R \neq 1; \\ \frac{[1 + (\delta - 1)\lambda_1]^2\{2n_{+0}\delta + 2n_{+1}[1 + (\delta - 1)\lambda_1]^2\}}{4n_{+0}n_{+1}\delta\lambda_1\{(1 - \lambda_1)^3 + \delta\lambda_1[\lambda_1^2(2 - \delta) + \lambda_1(\delta - 4) + 2]\}} \triangleq \Sigma_2(\theta_2), & R = 1. \end{cases} \quad (7)$$

其中 $\theta_1 = (\delta, \lambda_1, R)$, $\theta_2 = (\delta, \lambda_1)$.

3.2 置信区间的构造

3.2.1 基于样本方差的Wald统计量的置信区间

本文考虑基于样本方差的Wald统计量,具体形式为

$$T_{W1} = \frac{\hat{\delta} - \delta}{\sqrt{\hat{\text{Var}}(\hat{\delta})}} = \frac{\hat{\delta} - \delta}{\sqrt{\Sigma_1(\hat{\theta}_h)}}$$

其中 $h = 1, 2$. $h = 1$ 表示两只器官不独立,且 $\hat{\theta}_1 = (\hat{\delta}, \hat{\lambda}_1, \hat{R})$, $h = 2$ 表示两器官独立,且 $\hat{\theta}_2 = (\hat{\delta}, \hat{\lambda}_1)$.根据中心极限定理, T_{W1} 渐近服从标准正态分布.所以优势比 δ 基于样本方差的Wald统计量的置信水平为 $100(1 - \alpha)\%$ 的置信区间为

$$CI_{W1} = (\max(0, \hat{\delta} - Z_{1-\alpha}\sqrt{\Sigma_1(\hat{\theta}_h)}), \infty). \quad (8)$$

3.2.2 基于原假设下方差的Wald统计量的置信区间

基于原假设下方差的Wald统计量的具体形式为

$$T_{W2} = \frac{\hat{\delta} - \delta}{\sqrt{\hat{\text{Var}}(\hat{\delta}|H_0)}} = \frac{\hat{\delta} - \delta}{\sqrt{\Sigma_1(\tilde{\theta}_h)}}$$

根据中心极限定理, T_{W2} 渐近服从标准正态分布,所以优势比 δ 基于原假设下方差的Wald统计量的置信水平为 $100(1 - \alpha)\%$ 的置信区间为

$$CI_{W2} = (\max(0, \hat{\delta} - Z_{1-\alpha}\sqrt{\Sigma_1(\tilde{\theta}_h)}), \infty). \quad (9)$$

其中 $h = 1, 2$, $\tilde{\theta}_1 = (\delta_0, \tilde{\lambda}_1, \tilde{R})$, $\tilde{\theta}_2 = (\delta_0, \tilde{\lambda}_1)$.

3.2.3 基于样本方差的对数-Wald统计量的置信区间

基于样本方差的对数-Wald统计量的具体形式为

$$T_{W3} = \frac{\log \hat{\delta} - \log \delta}{\sqrt{\hat{\text{Var}}(\log \hat{\delta})}} = \frac{\log \hat{\delta} - \log \delta}{\sqrt{\Sigma_2(\hat{\theta}_h)}},$$

根据中心极限定理, T_{W3} 渐近服从标准正态分布, 所以优势比 δ 基于样本方差的对数-Wald统计量的置信水平为 $100(1 - \alpha)\%$ 的置信区间为

$$CI_{W3} = (\exp(\log \hat{\delta} - Z_{1-\alpha} \sqrt{\Sigma_2(\hat{\theta}_h)}), \infty). \quad (10)$$

3.2.4 基于原假设下方差的对数-Wald统计量的置信区间

基于原假设下方差的对数-Wald统计量的具体形式为

$$T_{W4} = \frac{\log \hat{\delta} - \log \delta}{\sqrt{\hat{\text{Var}}(\log \hat{\delta} | H_0)}} = \frac{\log \hat{\delta} - \log \delta}{\sqrt{\Sigma_2(\tilde{\theta}_h)}},$$

根据中心极限定理, T_{W4} 渐近服从标准正态分布, 所以优势比 δ 基于原假设下方差的对数-Wald统计量的置信水平为 $100(1 - \alpha)\%$ 的置信区间为

$$CI_{W4} = (\exp(\log \hat{\delta} - Z_{1-\alpha} \sqrt{\Sigma_2(\tilde{\theta}_h)}), \infty). \quad (11)$$

3.2.5 MOVER型置信区间

MOVER方法即方差重估法, 由Zou等人^[15-16]提出构造比值置信区间的方法, 其主要思想是利用比值中分子、分母的置信区间中将各自的方差“复原”出来构造比值的置信区间^[17].

根据 $\delta = \lambda_0(1 - \lambda_1)/[\lambda_1(1 - \lambda_0)]$, 令 $\delta = Y_1/Y_2$ 并将其转换为 $Y_1 - \delta Y_2 = 0$, \hat{Y}_1 与 \hat{Y}_2 分别表示 Y_1 与 Y_2 的极大似然估计, $\hat{\rho}$ 是 \hat{Y}_1 与 $\delta \hat{Y}_2$ 的相关系数 ρ 的估计值. 由中心极限定理有

$$\frac{(\hat{Y}_1 - \delta \hat{Y}_2) - (Y_1 - \delta Y_2)}{\sqrt{\text{Var}(\text{Var}Y_1) + \text{Var}(\delta \hat{Y}_2) - 2\hat{\rho}\sqrt{\text{Var}(\hat{Y}_1)\text{Var}(\delta \hat{Y}_2)}}} \sim N(0, 1),$$

则 $Y_1 - \delta Y_2$ 的置信下限为

$$l = (\hat{Y}_1 - \delta \hat{Y}_2) - Z_{1-\alpha} \sqrt{\text{Var}(\hat{Y}_1) + \text{Var}(\delta \hat{Y}_2) - 2\hat{\rho}\sqrt{\text{Var}(\hat{Y}_1)\text{Var}(\delta \hat{Y}_2)}}, \quad (12)$$

其中 $Z_{1-\alpha}$ 为标准正态分布的 $1 - \alpha$ 分位数. 将 Y_1 的置信下限令作 l_1 , Y_2 的置信上限令作 u_2 . 根据中心极限定理: $(\hat{Y}_1 - Y_1)/\sqrt{\text{Var}(\hat{Y}_1)} \sim N(0, 1)$, 从而有 $l_1 = \hat{Y}_1 - Z_{1-\alpha}\sqrt{\text{Var}(\hat{Y}_1)}$, 则可以得到 $\text{Var}(\hat{Y}_1) = (\hat{Y}_1 - l_1)^2/Z_{1-\alpha}^2$, 同理可得 $\text{Var}(\delta \hat{Y}_2) = \delta^2(u_2 - \hat{Y}_2)^2/Z_{1-\alpha}^2$, 将其带入(12)得

$$l = (\hat{Y}_1 - \delta \hat{Y}_2) - \sqrt{(\hat{Y}_1 - l_1)^2 + \delta^2(u_2 - \hat{Y}_2)^2 - 2\delta\hat{\rho}(\hat{Y}_1 - l_1)(u_2 - \hat{Y}_2)}. \quad (13)$$

由于有 $Y_1 - \delta Y_2 = 0$ 的限制, 进一步令上述得到的下限等于0, 反推出 δ 的置信区间下限, 则(13)可以通过加和的MOVER方法^[18]得到 δ 的置信下限为

$$\delta_L = \frac{\hat{Y}_1 \hat{Y}_2 - \hat{\rho}(\hat{Y}_1 - l_1)(u_2 - \hat{Y}_2) - \sqrt{[\hat{Y}_1 \hat{Y}_2 - \hat{\rho}(\hat{Y}_1 - l_1)(u_2 - \hat{Y}_2)]^2 - l_1 u_2 (2\hat{Y}_1 - l_1)(2\hat{Y}_2 - u_2)}}{u_2 (2\hat{Y}_2 - u_2)}.$$

其中

$$\hat{Y}_1 = \frac{(2n_{01} + n_{11})(n_{10} + 2n_{20})}{4n_{+0}n_{+1}}, \hat{Y}_2 = \frac{(2n_{00} + n_{10})(n_{11} + 2n_{21})}{4n_{+0}n_{+1}},$$

$$\text{Var}(\hat{Y}_1) = \frac{\lambda_0(1 + R\lambda_0 - 2\lambda_0)[\lambda_1(1 + R\lambda_1 - 2\lambda_1) + 2n_{+1}(1 - \lambda_1)^2] + 2n_{+0}\lambda_0^2\lambda_1(1 + R\lambda_1 - 2\lambda_1)}{4n_{+0}n_{+1}},$$

$$\text{Var}(\hat{Y}_2) = \frac{\lambda_0\lambda_1(1 + R\lambda_0 - 2\lambda_0)(1 + R\lambda_1 - 2\lambda_1 + 2n_{+1}\lambda_1) + 2n_{+0}(1 - \lambda_0)^2\lambda_1(1 + R\lambda_1 - 2\lambda_1)}{4n_{+0}n_{+1}},$$

$$\hat{\rho} = \frac{\lambda_0\lambda_1\{(1 + R\lambda_0 - 2\lambda_0)[1 + R\lambda_1 - 2\lambda_1 + 2n_{+1}(\lambda_1 - 1)] + 2n_{+0}(\lambda_0 - 1)(1 + R\lambda_1 - 2\lambda_1)\}}{4n_{+0}n_{+1}\sqrt{\text{Var}(\hat{Y}_1)\text{Var}(\hat{Y}_2)}}.$$

且 $\text{Var}(\hat{Y}_1)$, $\text{Var}(\hat{Y}_2)$, $\hat{\rho}$ 中的参数 λ_0 , λ_1 , R 都用其极大似然估计代替得到, 因而优势比 δ 的置信水平为 $100(1 - \alpha)\%$ 的置信区间为

$$CI_M = (\max(0, \delta_L), \infty). \quad (14)$$

3.2.6 Bootstrap重抽样置信区间

当样本量较小时, 考虑采用小样本下表现较好的Bootstrap重抽样置信区间, 具体步骤如下.

第一步: 给定观测数据 $n = (n_{00}, n_{01}, n_{10}, n_{11}, n_{20}, n_{21})$, 计算参数 δ , λ_1 与 R 的估计值 $\hat{\delta}$, $\hat{\lambda}_1$ 与 \hat{R} (在独立假设下 $R = 1$ 为一个常数);

第二步: 基于第一步计算出的 $\hat{\delta}$, $\hat{\lambda}_1$ 与 \hat{R} , 产生Bootstrap样本 $n^* = (n_{00}^*, n_{01}^*, n_{10}^*, n_{11}^*, n_{20}^*, n_{21}^*)$;

第三步: 对每一组Bootstrap样本 n^* 都可以计算参数 δ 的估计值 $\hat{\delta}^*$. 重复第二步 M_2 次, 得到 M_2 个Bootstrap样本下的估计 $\hat{\delta}_m^*$ ($m = 1, 2, \dots, M_2$), 并将这 M_2 个 $\hat{\delta}_m^*$ 按从小到大的顺序排列为 $\{\hat{\delta}_{(1)}^*, \hat{\delta}_{(2)}^*, \dots, \hat{\delta}_{(M_2)}^*\}$, 由此得到基于Bootstrap重抽样方法构造的置信区间.

根据Shao等人^[19]介绍的方法, 优势比 δ 的置信水平为 $100(1 - \alpha)\%$ 的置信区间为

$$CI_B = \left(\hat{\delta}_{([\alpha \cdot M_2])}^*, \infty \right). \quad (15)$$

其中 $[a]$ 表示对 a 进行取整.

3.3 样本量的计算

在检验问题 $H_0 : \delta \leq \delta_0$, $H_1 : \delta > \delta_0$ 下, 还关心给定显著性水平 $\alpha = 0.05$, 达到指定的 $1 - \beta = 95\%$ 功效, 需要多少的样本量? 以及此样本下的经验覆盖概率为多少? 为回答这个问题, 文章考虑计算样本量的迭代算法, 具体步骤如下.

第一步: 在 H_1 下, 给定 N ($N = n_{+0} + n_{+1}$), δ_0 , δ_1 , λ_1 , R 产生随机数据 $(n_{00}, n_{01}, n_{10}, n_{11}, n_{20}, n_{21})$;

第二步: 基于第一步产生的数据, 计算近似概率 $p(T_q > Z_{1-\alpha} | H_1)$, 记为 p^* , 其中 $q = W_1, W_2, W_3, W_4, M$;

第三步: 若第二步中计算出的近似概率 p^* 小于指定的功效 $1 - \beta$, 则增大 N 的值, 反之则减小 N 的值, 重复第一步和第二步, 不断调整 N 的值;

第四步: 重复第三步, 直到计算出各种方法的近似概率非常接近于指定的功效 $1 - \beta$, 即 $(|p^* - (1 - \beta)| < 0.005)$, 则可得到各检验统计量达到指定功效的近似样本量. 最后根据此样

本量在 H_0 下, 可求出对应的经验覆盖概率.

§4 随机模拟

为评价本文提出的四种Wald置信区间、MOVER型置信区间以及Bootstrap重抽样置信区间的统计性能, 通过Monte Carlo模拟计算各置信区间的经验覆盖概率(Empirical Coverage Probability, ECP)、置信下限(Lower Confidence Limit, LCL)的均值以及达到固定功效所需的样本量且在此样本下对应的经验覆盖概率. 制定检验的显著性水平 $\alpha = 0.05$, 且为了解各置信区间在不同参数设置下的表现, 考虑平衡设计下的样本量 $(n_{+0}, n_{+1}) = (20, 20), (50, 50)$, 以及非平衡设计下的样本量 $(n_{+0}, n_{+1}) = (20, 30), (30, 40)$. 计算经验覆盖概率及置信下限均值时, 非独立假设下参数设置为: $\delta_0 = 0.3, 0.5, 0.7, 1.0, \lambda_1 = 0.30, 0.45, R = 0.8, 1.0, 1.2$. 独立假设下当真实模型独立时, 参数设置为: $\delta_0 = 0.3, 0.5, 0.7, 1.0, \lambda_1 = 0.30, 0.45$; 当真实模型非独立时, 参数设置为: $\delta_0 = 0.3, 0.5, 0.7, 1.0, \lambda_1 = 0.30, 0.45, R = 1.2$. 在每个样本量设置下, 对每一组参数组合, 通过重复模拟 M_1 次分别计算各置信区间的经验覆盖概率, 计算公式为

$$ECP = \frac{1}{M_1} \sum_{m=1}^{M_1} I(\delta \in (\delta_l(n_m), \infty)),$$

其中 $(\delta_l(n_m), \infty)$ 是感兴趣参数 δ 模拟第 m 次的置信区间, n_m 表示第 m 次模拟的样本, $I(A)$ 为示性函数, 当 A 成立时取1, 不成立时取0. 令 $M_1 = 10000, M_2 = 5000$, 不同情形下的模拟结果见后文所列的表2-表4, 图1-图3(表2仅列出参数 $\delta_0 = 0.5, 0.7$ 的模拟结果).

参照Tang等人^[20]的标准, ECP的模拟结果越接近置信水平95%表明效果越好, 且原则上若 $94\% \leq ECP \leq 96\%$, 则表明覆盖效果优良, 若 $ECP < 94\%$, 则置信区间覆盖过于紧缩, 若 $ECP > 96\%$, 则置信区间覆盖过于保守. 通过表2-表4和图1-图3的模拟结果, 可得到如下的结论.

1) 在非独立假设下, 即便真实模型独立时, 对数-Wald置信区间(基于样本方差和原假设下方差)、MOVER型置信区间、Bootstrap重抽样置信区间, 即 $CI_{W3}, CI_{W4}, CI_M, CI_B$ 具有令人满意的结果, 它们在不同样本量以及不同参数的组合下, 经验覆盖概率都接近预先给定的置信水平. 从置信下限的均值来看, 它们的置信下限均值都相对较大, 相当于区间宽度更窄. 其中基于样本方差的对数-Wald置信区间是最优的, 其在保证覆盖概率接近95%的前提下, 经验覆盖概率取值离散程度最小, 区间宽度也最窄.

2) 在非独立假设下, Wald置信区间(基于样本方差和原假设下方差)的表现不是令人满意的. 基于样本方差的对数-Wald置信区间是保守的, 其不同样本量以及不同参数的组合下计算出的经验覆盖概率都大于96%; 基于原假设下方差的Wald置信区间是紧缩的, 其不同样本量以及不同参数的组合下计算出的经验覆盖概率都小于94%; 且从置信下限的均值来看, 它们置信区间下限的均值较小, 即区间宽度较宽.

3) 在独立假设下, 当真实模型独立时, $CI_{W3}, CI_{W4}, CI_M, CI_B$ 具有令人满意的结果, 它们在不同样本量的分配设计以及不同参数的组合下, 经验覆盖概率都接近预先给定的置信水平, 且它们的置信下限均值都相对较大; 而当真实模型非独立时, 各种方法得到的置信区间的经验覆盖概率都不准确, 虽然 $CI_{W3}, CI_{W4}, CI_M, CI_B$ 的经验覆盖概率大致在覆盖效果优良的范围, 但相比非独立假设与独立假设下真实模型独立时的结果来看, 其经验覆盖概率都有一定的偏差.

表 2 非独立假设下模拟10000次得到的经验覆盖概率及置信下限的均值

δ_0	λ_1	R	CI_{W1}	CI_{W2}	CI_{W3}	CI_{W4}	CI_M	CI_B	
平衡设计: $(n_{+0}, n_{+1}) = (20, 20)$									
0.5	0.30	0.8	99.95(0.084)	88.88(0.183)	95.01(0.239)	95.62(0.237)	95.94(0.199)	95.60(0.219)	
		1.0	99.95(0.072)	88.05(0.189)	94.63(0.238)	95.17(0.236)	95.59(0.196)	95.28(0.217)	
		1.2	99.98(0.061)	87.83(0.185)	94.62(0.231)	95.21(0.230)	95.81(0.188)	95.46(0.210)	
	0.45	0.8	99.35(0.156)	89.38(0.205)	94.07(0.267)	95.06(0.265)	95.05(0.242)	94.88(0.253)	
		1.0	99.70(0.134)	88.05(0.206)	93.45(0.263)	94.65(0.260)	94.83(0.236)	94.60(0.248)	
		1.2	99.77(0.114)	88.46(0.199)	94.04(0.253)	94.89(0.252)	95.25(0.222)	95.07(0.236)	
	0.7	0.30	0.8	99.89(0.151)	88.13(0.276)	94.84(0.351)	95.01(0.352)	95.23(0.319)	94.86(0.335)
			1.0	99.98(0.129)	87.86(0.274)	94.43(0.344)	94.61(0.345)	94.96(0.309)	94.53(0.327)
			1.2	99.97(0.114)	88.08(0.278)	94.56(0.339)	94.60(0.341)	95.05(0.302)	94.69(0.320)
		0.45	0.8	99.28(0.243)	88.92(0.306)	94.24(0.387)	94.59(0.385)	94.67(0.364)	94.31(0.374)
			1.0	99.71(0.202)	89.01(0.291)	94.45(0.371)	94.79(0.370)	94.87(0.346)	94.49(0.357)
			1.2	99.94(0.175)	89.02(0.284)	94.71(0.360)	94.82(0.361)	95.22(0.328)	94.93(0.340)
平衡设计: $(n_{+0}, n_{+1}) = (50, 50)$									
0.5	0.30	0.8	98.92(0.239)	91.01(0.253)	94.55(0.305)	95.08(0.304)	95.32(0.288)	94.90(0.295)	
		1.0	99.09(0.231)	91.30(0.246)	95.03(0.300)	95.58(0.299)	95.89(0.282)	95.35(0.290)	
		1.2	99.01(0.224)	90.28(0.243)	94.44(0.297)	94.99(0.296)	95.41(0.279)	94.93(0.287)	
	0.45	0.8	98.10(0.283)	91.44(0.290)	94.41(0.329)	94.93(0.328)	95.16(0.318)	94.85(0.322)	
		1.0	98.63(0.267)	91.14(0.276)	94.70(0.320)	95.05(0.319)	95.31(0.308)	95.01(0.313)	
		1.2	98.76(0.255)	90.60(0.266)	94.52(0.313)	94.94(0.312)	95.47(0.299)	95.21(0.304)	
	0.7	0.30	0.8	98.65(0.359)	91.01(0.378)	94.56(0.441)	94.86(0.441)	94.96(0.428)	94.77(0.433)
			1.0	98.89(0.346)	90.85(0.366)	95.15(0.433)	95.39(0.433)	95.58(0.419)	95.37(0.425)
			1.2	99.02(0.332)	90.44(0.353)	94.92(0.424)	95.07(0.424)	95.40(0.409)	95.27(0.415)
		0.45	0.8	98.04(0.410)	90.91(0.420)	94.31(0.469)	94.54(0.469)	94.70(0.459)	94.73(0.463)
			1.0	98.70(0.384)	91.30(0.396)	94.85(0.453)	95.06(0.453)	95.15(0.443)	95.18(0.447)
			1.2	98.66(0.370)	90.58(0.387)	94.32(0.447)	94.43(0.447)	94.84(0.432)	94.87(0.436)
非平衡设计: $(n_{+0}, n_{+1}) = (20, 30)$									
0.5	0.30	0.8	99.68(0.116)	90.14(0.184)	94.59(0.249)	96.30(0.245)	96.43(0.205)	95.57(0.224)	
		1.0	99.76(0.106)	89.48(0.184)	94.52(0.246)	95.84(0.242)	96.06(0.200)	95.25(0.221)	
		1.2	99.77(0.097)	88.66(0.188)	94.03(0.245)	95.47(0.241)	95.97(0.196)	95.13(0.218)	
	0.45	0.8	99.00(0.182)	90.29(0.214)	94.10(0.277)	95.28(0.275)	95.29(0.252)	94.79(0.263)	
		1.0	99.43(0.159)	90.02(0.200)	94.91(0.266)	95.80(0.263)	96.00(0.239)	95.56(0.251)	
		1.2	99.56(0.147)	89.06(0.202)	94.33(0.264)	95.30(0.261)	95.87(0.232)	95.30(0.245)	
	0.7	0.30	0.8	99.61(0.202)	89.63(0.283)	94.34(0.368)	95.26(0.365)	95.07(0.330)	94.78(0.346)
			1.0	99.63(0.183)	89.16(0.275)	94.36(0.360)	95.19(0.357)	95.23(0.319)	95.00(0.336)
			1.2	99.81(0.170)	88.28(0.282)	93.92(0.360)	94.87(0.357)	94.98(0.315)	94.68(0.334)
		0.45	0.8	98.93(0.280)	89.88(0.316)	94.29(0.400)	95.05(0.397)	94.91(0.377)	94.80(0.386)
			1.0	99.46(0.243)	90.14(0.293)	94.87(0.382)	95.36(0.380)	95.49(0.358)	95.29(0.367)
			1.2	99.67(0.223)	88.93(0.294)	94.02(0.377)	94.49(0.376)	95.06(0.345)	94.81(0.356)
非平衡设计: $(n_{+0}, n_{+1}) = (30, 40)$									
0.5	0.30	0.8	99.31(0.180)	90.39(0.213)	94.89(0.276)	95.82(0.274)	95.92(0.246)	95.70(0.259)	
		1.0	99.45(0.169)	90.67(0.204)	94.80(0.269)	95.71(0.268)	95.95(0.238)	95.70(0.252)	
		1.2	99.51(0.162)	89.76(0.204)	94.35(0.268)	95.55(0.266)	95.78(0.235)	95.60(0.249)	
	0.45	0.8	98.26(0.236)	90.06(0.251)	93.61(0.304)	94.63(0.302)	94.73(0.287)	94.32(0.294)	
		1.0	98.98(0.218)	90.52(0.237)	94.37(0.294)	95.17(0.292)	95.46(0.275)	94.99(0.283)	
		1.2	99.11(0.204)	90.01(0.227)	94.67(0.287)	95.36(0.286)	95.85(0.265)	95.47(0.274)	
	0.7	0.30	0.8	99.20(0.285)	90.27(0.320)	94.79(0.402)	95.41(0.401)	95.56(0.377)	95.38(0.387)
			1.0	99.28(0.272)	89.46(0.318)	94.49(0.399)	95.07(0.398)	95.23(0.372)	94.88(0.383)
			1.2	99.46(0.256)	90.08(0.305)	94.73(0.389)	95.27(0.388)	95.51(0.360)	95.19(0.372)
		0.45	0.8	98.43(0.349)	90.62(0.367)	93.98(0.435)	94.58(0.434)	94.61(0.419)	94.39(0.425)
			1.0	98.82(0.319)	90.46(0.344)	94.69(0.419)	95.11(0.418)	95.25(0.403)	95.03(0.409)
			1.2	99.22(0.300)	89.70(0.332)	94.48(0.410)	94.76(0.409)	95.34(0.388)	95.15(0.395)

表 3 独立假设下真实模型独立时模拟10000次得到的经验覆盖概率及置信下限的均值

δ_0	λ_1	CI_{W1}	CI_{W2}	CI_{W3}	CI_{W4}	CI_M	CI_B
平衡设计: $(n_{+0}, n_{+1}) = (20, 20)$							
0.3	0.30	100.0(0.018)	88.38(0.102)	94.85(0.129)	96.11(0.126)	97.04(0.084)	96.01(0.107)
	0.45	99.92(0.055)	89.19(0.107)	94.95(0.144)	95.82(0.142)	96.68(0.115)	95.83(0.130)
0.5	0.30	99.99(0.062)	88.82(0.181)	95.15(0.231)	95.59(0.231)	95.98(0.190)	95.71(0.212)
	0.45	99.94(0.121)	88.63(0.194)	94.37(0.253)	94.80(0.253)	95.47(0.228)	94.80(0.242)
0.7	0.30	100.0(0.116)	85.87(0.273)	94.57(0.340)	94.57(0.343)	94.57(0.308)	94.60(0.325)
	0.45	99.90(0.186)	87.04(0.280)	94.16(0.362)	94.16(0.363)	94.16(0.341)	94.26(0.354)
1.0	0.30	100.0(0.200)	88.55(0.388)	95.27(0.495)	94.66(0.501)	94.74(0.475)	93.83(0.487)
	0.45	99.95(0.278)	89.32(0.415)	95.27(0.527)	95.02(0.530)	95.02(0.514)	92.68(0.528)
平衡设计: $(n_{+0}, n_{+1}) = (50, 50)$							
0.3	0.30	99.14(0.117)	91.53(0.130)	95.02(0.168)	96.28(0.167)	96.67(0.147)	95.93(0.157)
	0.45	98.87(0.146)	91.85(0.152)	95.10(0.183)	95.81(0.182)	96.30(0.171)	95.88(0.176)
0.5	0.30	99.08(0.229)	91.14(0.245)	94.94(0.299)	95.66(0.299)	95.95(0.282)	95.48(0.289)
	0.45	98.75(0.265)	91.11(0.275)	94.95(0.318)	95.36(0.318)	95.62(0.308)	95.27(0.312)
0.7	0.30	99.18(0.340)	91.38(0.359)	95.52(0.428)	95.53(0.429)	95.53(0.415)	95.47(0.421)
	0.45	98.62(0.382)	90.39(0.397)	94.36(0.453)	94.43(0.453)	94.83(0.444)	95.01(0.448)
1.0	0.30	99.14(0.514)	90.63(0.546)	94.79(0.632)	94.52(0.634)	94.60(0.625)	93.99(0.629)
	0.45	98.55(0.552)	91.30(0.573)	94.95(0.651)	94.95(0.652)	94.95(0.647)	94.20(0.648)
非平衡设计: $(n_{+0}, n_{+1}) = (20, 30)$							
0.3	0.30	99.92(0.033)	90.04(0.097)	94.62(0.132)	96.94(0.129)	97.35(0.085)	96.23(0.107)
	0.45	99.82(0.073)	91.40(0.108)	95.43(0.149)	96.76(0.148)	96.81(0.121)	96.25(0.134)
0.5	0.30	99.95(0.099)	89.48(0.181)	94.63(0.242)	95.83(0.240)	96.33(0.197)	95.52(0.218)
	0.45	99.70(0.152)	90.54(0.196)	95.29(0.262)	95.80(0.261)	96.23(0.236)	95.76(0.247)
0.7	0.30	99.90(0.173)	89.86(0.270)	95.19(0.355)	95.44(0.354)	95.97(0.315)	95.39(0.333)
	0.45	99.72(0.232)	89.24(0.292)	95.34(0.378)	95.35(0.379)	95.53(0.357)	95.43(0.366)
1.0	0.30	99.95(0.286)	89.15(0.409)	95.03(0.526)	95.03(0.528)	95.31(0.497)	94.62(0.510)
	0.45	99.80(0.346)	89.65(0.434)	94.97(0.550)	94.45(0.552)	94.44(0.537)	94.32(0.543)
非平衡设计: $(n_{+0}, n_{+1}) = (30, 40)$							
0.3	0.30	99.69(0.074)	90.79(0.109)	94.93(0.150)	96.59(0.148)	96.92(0.115)	96.25(0.131)
	0.45	99.38(0.112)	90.29(0.127)	94.98(0.166)	96.16(0.165)	96.32(0.146)	95.65(0.155)
0.5	0.30	99.52(0.167)	90.62(0.205)	94.89(0.269)	95.83(0.268)	96.26(0.238)	95.88(0.252)
	0.45	99.29(0.212)	90.74(0.232)	95.21(0.290)	95.33(0.290)	95.78(0.272)	95.37(0.280)
0.7	0.30	99.51(0.266)	90.01(0.313)	94.65(0.395)	94.95(0.395)	95.21(0.369)	94.93(0.380)
	0.45	99.29(0.312)	90.45(0.340)	94.77(0.415)	94.82(0.415)	95.06(0.401)	94.98(0.407)
1.0	0.30	99.58(0.407)	90.11(0.462)	94.79(0.575)	94.79(0.576)	94.79(0.557)	94.50(0.565)
	0.45	99.47(0.459)	90.34(0.501)	94.93(0.603)	94.93(0.604)	94.93(0.594)	94.60(0.598)

表 4 独立假设下真实模型非独立时模拟10000次得到的经验覆盖概率及置信下限的均值

δ_0	λ_1	CI_{W1}	CI_{W2}	CI_{W3}	CI_{W4}	CI_M	CI_B
平衡设计: $(n_{+0}, n_{+1}) = (20, 20)$							
0.3	0.30	100.0(0.018)	87.69(0.106)	94.36(0.130)	95.90(0.127)	96.80(0.085)	95.66(0.108)
	0.45	99.93(0.055)	88.33(0.111)	94.28(0.145)	95.25(0.143)	96.21(0.117)	95.33(0.131)
0.5	0.30	99.99(0.061)	88.28(0.188)	94.74(0.233)	95.19(0.233)	95.70(0.192)	95.40(0.213)
	0.45	99.88(0.122)	87.31(0.204)	93.37(0.257)	93.89(0.256)	94.70(0.231)	93.81(0.245)
0.7	0.30	100.0(0.116)	84.93(0.281)	94.12(0.342)	94.12(0.345)	94.12(0.310)	94.20(0.327)
	0.45	99.86(0.187)	85.68(0.293)	93.05(0.366)	93.05(0.368)	93.05(0.346)	93.09(0.359)
1.0	0.30	100.0(0.199)	87.33(0.407)	94.63(0.500)	93.92(0.507)	93.98(0.481)	93.21(0.493)
	0.45	99.88(0.280)	87.35(0.442)	93.83(0.535)	93.61(0.540)	93.61(0.524)	91.32(0.537)
平衡设计: $(n_{+0}, n_{+1}) = (50, 50)$							
0.3	0.30	99.10(0.117)	91.08(0.130)	94.56(0.168)	95.85(0.167)	96.23(0.147)	95.58(0.157)
	0.45	98.61(0.146)	91.02(0.153)	94.35(0.183)	95.12(0.183)	95.58(0.171)	95.23(0.177)
0.5	0.30	98.97(0.229)	90.48(0.247)	94.29(0.299)	95.09(0.299)	95.47(0.282)	94.84(0.290)
	0.45	98.22(0.266)	89.93(0.278)	94.12(0.320)	94.55(0.320)	94.87(0.309)	94.46(0.314)
0.7	0.30	98.97(0.341)	90.32(0.362)	94.90(0.430)	94.90(0.430)	94.92(0.416)	94.87(0.422)
	0.45	98.17(0.383)	88.91(0.400)	93.39(0.455)	93.41(0.455)	93.69(0.446)	94.01(0.450)
1.0	0.30	98.83(0.515)	89.68(0.548)	94.44(0.633)	94.21(0.635)	94.27(0.626)	93.36(0.630)
	0.45	97.63(0.556)	89.72(0.581)	93.79(0.656)	93.79(0.657)	93.79(0.652)	92.66(0.653)
非平衡设计: $(n_{+0}, n_{+1}) = (20, 30)$							
0.3	0.30	99.92(0.033)	89.30(0.100)	94.26(0.133)	96.55(0.129)	97.06(0.086)	95.94(0.108)
	0.45	99.69(0.073)	90.50(0.112)	94.69(0.150)	96.12(0.148)	96.20(0.122)	95.63(0.135)
0.5	0.30	99.97(0.099)	88.92(0.185)	94.23(0.244)	95.38(0.241)	95.84(0.198)	95.14(0.219)
	0.45	99.59(0.152)	89.59(0.202)	94.31(0.263)	94.89(0.263)	95.51(0.238)	94.79(0.249)
0.7	0.30	99.86(0.173)	89.12(0.275)	94.64(0.356)	95.00(0.355)	95.56(0.316)	94.97(0.335)
	0.45	99.64(0.234)	87.95(0.303)	94.04(0.382)	94.04(0.382)	94.22(0.361)	94.11(0.371)
1.0	0.30	99.91(0.287)	88.02(0.421)	94.34(0.530)	94.33(0.532)	94.57(0.501)	93.80(0.514)
	0.45	99.63(0.349)	87.75(0.456)	93.91(0.558)	93.16(0.561)	93.16(0.546)	93.04(0.552)
非平衡设计: $(n_{+0}, n_{+1}) = (30, 40)$							
0.3	0.30	99.64(0.075)	90.37(0.111)	94.59(0.151)	96.16(0.148)	96.47(0.116)	95.82(0.132)
	0.45	99.06(0.112)	89.33(0.129)	93.94(0.167)	95.42(0.166)	95.65(0.147)	94.81(0.156)
0.5	0.30	99.39(0.168)	89.96(0.208)	94.35(0.270)	95.32(0.269)	95.80(0.239)	95.44(0.253)
	0.45	98.92(0.214)	89.29(0.237)	94.37(0.292)	94.46(0.292)	95.12(0.275)	94.60(0.283)
0.7	0.30	99.49(0.267)	89.14(0.318)	94.18(0.397)	94.56(0.396)	94.91(0.371)	94.56(0.382)
	0.45	99.07(0.314)	89.04(0.347)	93.88(0.419)	93.91(0.419)	94.11(0.404)	94.03(0.411)
1.0	0.30	99.38(0.409)	89.22(0.471)	94.06(0.579)	94.06(0.580)	94.06(0.561)	93.84(0.569)
	0.45	99.08(0.462)	88.45(0.513)	93.55(0.608)	93.55(0.610)	93.55(0.600)	93.28(0.604)

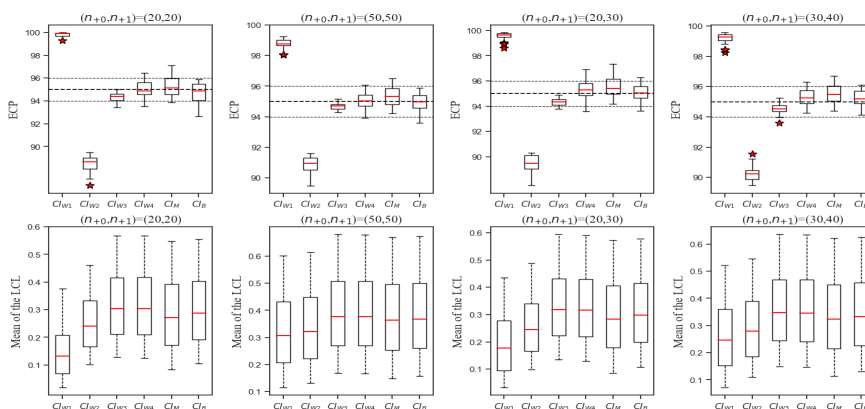


图 1 非独立假设下的经验覆盖概率及置信下限均值的箱线图

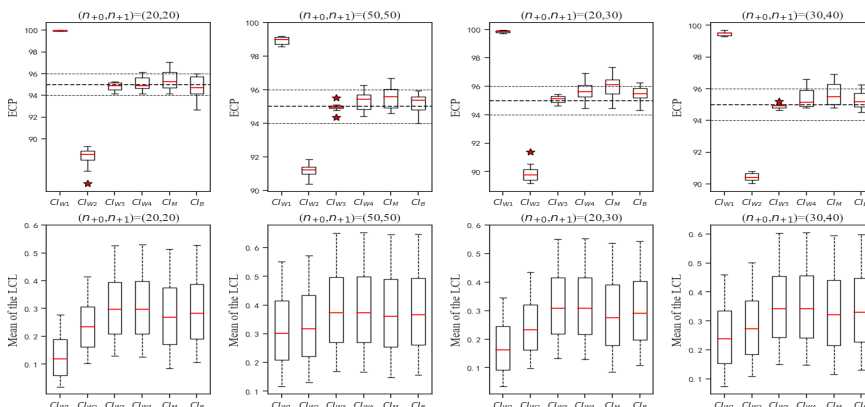


图 2 独立假设下真实模型独立时的经验覆盖概率及置信下限均值的箱线图

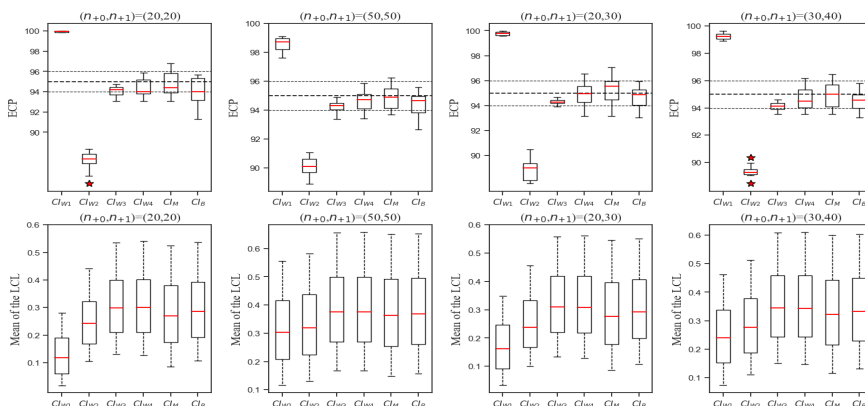


图 3 独立假设下真实模型非独立时的经验覆盖概率及置信下限均值的箱线图

根据上述结论可知, 在研究过程中, 若不清楚真实模型是否独立时, 推荐使用非独立假设下的方法.

由于独立假设下当真实模型非独立时效果不佳, 所以对于样本量的确定问题中, 本文仅从非独立假设下进行模拟研究. 计算达到固定功效所需的样本量以及此样本下的经验覆盖概率时, 参数设置为: $\delta_0 = 0.5, 0.7, \lambda_1 = 0.30, 0.45, R = 0.8, 1.0, 1.2$, 模拟结果见表5.

表 5 $\alpha = 0.05$ 下功效达到95%所需的样本量和其对应的经验覆盖概率

δ_0	λ_1	R	CI_{W1}	CI_{W2}	CI_{W3}	CI_{W4}	CI_M	
平衡设计: $n_{+0} : n_{+1} = 1 : 1$								
0.5	0.30	0.8	235(97.56)	176(91.98)	194(94.43)	200(95.37)	200(95.55)	
		1.0	257(97.44)	182(92.26)	212(95.17)	212(95.53)	215(95.16)	
		1.2	274(97.85)	200(91.95)	227(95.17)	231(95.39)	236(95.55)	
	0.45	0.8	178(97.76)	130(91.86)	144(94.56)	160(95.48)	151(95.25)	
		1.0	214(97.88)	160(92.51)	180(95.09)	182(95.24)	182(95.39)	
		1.2	242(97.51)	173(91.16)	202(94.72)	210(95.20)	212(95.98)	
	0.7	0.30	0.8	351(97.38)	271(91.93)	300(94.85)	303(95.11)	303(95.15)
			1.0	382(97.18)	300(92.50)	331(95.05)	340(94.81)	331(95.26)
			1.2	424(96.69)	321(91.77)	362(94.84)	365(95.25)	365(95.36)
0.45		0.8	270(97.02)	210(92.57)	233(95.22)	240(95.30)	241(94.97)	
		1.0	334(97.10)	255(92.66)	300(95.27)	291(95.31)	296(94.94)	
		1.2	382(96.97)	300(92.89)	331(94.31)	331(94.42)	351(95.23)	
非平衡设计: $n_{+0} : n_{+1} = 3 : 4$								
0.5	0.30	0.8	243(97.67)	180(91.95)	200(94.68)	210(95.57)	206(95.45)	
		1.0	270(97.80)	200(92.49)	220(94.72)	230(95.60)	222(95.78)	
		1.2	286(97.53)	210(91.85)	241(94.57)	250(95.61)	251(95.74)	
	0.45	0.8	185(97.36)	134(91.69)	151(94.30)	160(95.63)	157(95.44)	
		1.0	221(97.61)	160(92.52)	190(94.84)	186(95.26)	192(95.73)	
		1.2	251(97.72)	182(92.09)	210(94.74)	211(95.47)	220(95.65)	
	0.7	0.30	0.8	360(97.31)	280(93.06)	310(94.93)	320(95.18)	317(95.50)
			1.0	400(97.35)	310(92.74)	341(94.61)	350(95.33)	351(95.34)
			1.2	425(97.37)	331(92.54)	370(94.72)	382(95.45)	385(95.30)
0.45		0.8	275(96.99)	220(92.55)	241(94.50)	250(94.93)	246(94.96)	
		1.0	350(97.18)	270(92.66)	294(94.60)	301(94.86)	301(94.99)	
		1.2	391(96.66)	310(92.98)	331(94.11)	331(94.32)	353(95.32)	

从表5中可以得到下列结论.

- 1) 在不同参数组合下, 样本量的值随着 δ_0 和 R 的增加而增加, 随着 λ_1 的增加而减少.
- 2) 不同样本量的分配设计, 对样本量的变化无显著影响.

3) 基于样本方差的Wald统计量在达到固定功效时, 所需的样本量最多, 且在此样本量下置信区间的覆盖过于保守. 基于原假设下方差的Wald统计量在达到固定功效时, 所需的样本量最少, 但在此样本量下所对应置信区间的覆盖过于紧缩. 对数-Wald统计量(基于样本方差和原假设下方差)、MOVER型统计量的样本量表现是相对较好的, 它们不仅在达到固定功效时所需的样本量较少, 此时的样本量所对应的置信区间的覆盖概率也非常接近于置信水平.

§5 实例分析

以Mandel等人^[21]关于儿童中耳炎疾病的双盲临床试验为例, 此试验中研究者将接受治疗的214个儿童随机的分配到两种药物(Cefaclor和Amoxicillin)的治疗组中, 每个儿童都接受为期14天的治疗. 经过14天的治疗后, 有两只耳朵患病的儿童会出现三种治疗结果: 患者两只耳朵都被治愈、患者只有一只耳朵被治愈以及两只耳朵都未被治愈, 此临床试验数据如表6所示.

表 6 两只耳朵患病的儿童接受14天治疗后的数据

治愈的耳朵只数	治疗药物 (方法)	
	Cefaclor	Amoxicillin
0	14	15
1	9	3
2	21	13
合计	44	31

为了判断Cefaclor的疗效是否非劣效于Amoxicillin, 对 $H_0 : \delta \leq \delta_0$, $H_1 : \delta > \delta_0$ 进行检验, 其中取 $\delta_0 = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0$. 根据表6中的试验数据, 有 $n_{00} = 14, n_{01} = 15, n_{10} = 9, n_{11} = 3, n_{20} = 21$ 及 $n_{21} = 31$. 通过计算可得 $\hat{R} = 1.6165$. 为此认为此实验数据具有组内相关性. 因此在非独立假设下, 采用本文提出的方法计算出各置信区间的下限如表7所示(表7仅列出 $\delta_0 = 0.5, 0.6, 0.7, 0.8, 0.9, 1.0$ 的结果).

表 7 基于优势比 δ 的不同置信区间的置信下限

δ_0	CI_{W1}	CI_{W2}	CI_{W3}	CI_{W4}	CI_M	CI_B
0.5	0.939	1.348	1.050	1.009	0.769	0.775
0.6	0.939	1.334	1.050	1.062	0.769	0.771
0.7	0.939	1.320	1.050	1.099	0.769	0.767
0.8	0.939	1.297	1.050	1.117	0.769	0.767
0.9	0.939	1.261	1.050	1.114	0.769	0.787
1.0	0.939	1.212	1.050	1.098	0.769	0.771

由表7可以看出, 当 $\delta_0 \leq 0.7$ 时本文提出的六种方法的置信区间下限都大于 δ_0 , 则可以认为Cefaclor的药效非劣效于Amoxicillin. 而当非劣效性边界值 $\delta_0 > 0.7$ 时, 所提的方法不能全部

支持拒绝原假设, 这是因为 δ_0 越大, 对药物的要求越高, 从而越难成立.

§6 结论

本文在双边试验设计下基于优势比研究了两种治疗药物(方法)的区间估计问题, 分别给出了四种不同形式的Wald置信区间、MOVER型置信区间以及Bootstrap重抽样置信区间的构造方法, 并通过Monte Carlo模拟研究了各种置信区间的经验覆盖概率以及置信下限的均值. 结果表明: 在非独立假设下无论真实模型是否独立, 以及独立假设下真实模型独立时, 对数-Wald置信区间(基于样本方差和基于原假设下方差)、MOVER型置信区间以及Bootstrap重抽样置信区间, 即 CI_{W3} , CI_{W4} , CI_M , CI_B 的表现是令人满意, 且 CI_{W3} 是最优的, 它既能保证覆盖概率接近95%, 又能使经验覆盖概率取值离散程度较小, 其区间宽度也最窄. 在独立假设下当真实模型非独立时, 得到的结果带有一定的偏差. 此外, 本文还在非独立假设下达到固定功效进行了样本量的确定研究, 结果表明基于对数-Wald统计量(基于样本方差和原假设下方差)的置信区间、MOVER型置信区间的样本量相对较少. 综上所述, 无论是置信区间的经验覆盖概率、置信下限均值以及样本量的确定研究中, 基于样本方差的对数-Wald统计量的置信区间是最有效的, 可应用于实际问题.

附录

1 参数空间的推导

由 $0 \leq \lambda_1 \leq 1$ 和 $0 \leq \delta \leq \infty$ 可以推导出参数 R 的约束空间为

$$\begin{cases} 0 \leq 1 + \frac{R\delta^2\lambda_1^2}{[1+(\delta-1)\lambda_1]^2} - \frac{2\delta\lambda_1}{1+(\delta-1)\lambda_1} \leq 1 \\ 0 \leq \frac{2\delta\lambda_1}{1+(\delta-1)\lambda_1} \left[1 - \frac{R\delta\lambda_1}{1+(\delta-1)\lambda_1}\right] \leq 1 \\ 0 \leq \frac{R\delta^2\lambda_1^2}{[1+(\delta-1)\lambda_1]^2} \leq 1 \end{cases}, \quad \begin{cases} 0 \leq 1 + R\lambda_1^2 - 2\lambda_1 \leq 1 \\ 0 \leq 2\lambda_1(1 - R\lambda_1) \leq 1 \\ 0 \leq R\lambda_1^2 \leq 1 \end{cases}.$$

从而有

$$\begin{cases} \frac{\delta^2\lambda_1^2 - (\lambda_1 - 1)^2}{\delta^2\lambda_1^2} \leq R \leq \frac{2[1+(\delta-1)\lambda_1]}{\delta\lambda_1} \\ \frac{\delta^2\lambda_1^2 - (\lambda_1 - 1)^2}{2\delta^2\lambda_1^2} \leq R \leq \frac{1+(\delta-1)\lambda_1}{\delta\lambda_1} \\ 0 \leq R \leq \frac{[1+(\delta-1)\lambda_1]^2}{\delta^2\lambda_1^2} \end{cases}, \quad \begin{cases} \frac{2\lambda_1 - 1}{\lambda_1^2} \leq R \leq \frac{2}{\lambda_1} \\ \frac{2\lambda_1 - 1}{2\lambda_1^2} \leq R \leq \frac{1}{\lambda_1} \\ 0 \leq R \leq \frac{1}{\lambda_1^2} \end{cases}.$$

则 R 的参数空间为

$$\max \left\{ 0, \frac{\delta^2\lambda_1^2 - (\lambda_1 - 1)^2}{\delta^2\lambda_1^2}, \frac{2\lambda_1 - 1}{\lambda_1^2} \right\} \leq R \leq \min \left\{ \frac{1+(\delta-1)\lambda_1}{\delta\lambda_1}, \frac{1}{\lambda_1} \right\}.$$

2 感兴趣参数方差的推导

为建立Fisher信息阵, 首先找出参数 $\theta = (\delta, \lambda_1, R)^T$ 的得分函数 $S(\delta, \lambda_1, R)$, 即

$$S(\delta, \lambda_1, R) = \left(\frac{\partial \ln l(\theta)}{\partial \delta}, \frac{\partial \ln l(\theta)}{\partial \lambda_1}, \frac{\partial \ln l(\theta)}{\partial R} \right)^T =$$

$$\left(\begin{array}{l} \frac{2n_{00}\delta\lambda_1^2(R-1)}{(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)} - \frac{2(n_{00} + n_{10} + n_{20})\lambda_1}{1 + (\delta-1)\lambda_1} + \frac{n_{10}\lambda_1(1-R)}{1-\lambda_1 + \delta\lambda_1(1-R)} + \\ \frac{n_{10} + 2n_{20}}{\delta} \\ \frac{2n_{00}[\lambda_1 - 1 + \delta^2\lambda_1(R-1)]}{(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)} - \frac{2(n_{00} + n_{10} + n_{20})(\delta-1)}{1 + (\delta-1)\lambda_1} + \frac{2n_{01}(R\lambda_1 - 1)}{1 + R\lambda_1^2 - 2\lambda_1} + \\ \frac{n_{10}[\delta(1-R) - 1]}{1-\lambda_1 + \delta\lambda_1(1-R)} - \frac{n_{11}R}{1-R\lambda_1} + \frac{n_{10} + n_{11} + 2n_{20} + 2n_{21}}{\lambda_1} \\ \frac{n_{00}\delta^2\lambda_1^2}{(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)} + \frac{n_{01}\lambda_1^2}{1 + R\lambda_1^2 - 2\lambda_1} - \frac{n_{10}\delta\lambda_1}{1-\lambda_1 + \delta\lambda_1(1-R)} - \frac{n_{11}\lambda_1}{1-R\lambda_1} + \\ \frac{n_{20} + n_{21}}{R} \end{array} \right)$$

进一步有

$$\begin{aligned} \frac{\partial^2 \ln l(\theta)}{\partial \delta^2} &= \frac{2n_{00}\lambda_1^2(R-1)[(1-\lambda_1)^2 + \delta^2\lambda_1^2(1-R)]}{[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)]^2} + \frac{2(n_{00} + n_{10} + n_{20})\lambda_1^2}{[1 + (\delta-1)\lambda_1]^2} - \frac{n_{10} + 2n_{20}}{\delta^2} - \\ &\quad \frac{n_{10}\lambda_1^2(1-R)^2}{[1-\lambda_1 + \delta\lambda_1(1-R)]^2}, \\ \frac{\partial^2 \ln l(\theta)}{\partial \delta \partial \lambda_1} &= \frac{4n_{00}\delta\lambda_1(1-\lambda_1)(R-1)}{[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)]^2} - \frac{2(n_{00} + n_{10} + n_{20})}{[1 + (\delta-1)\lambda_1]^2} + \frac{n_{10}(1-R)}{[1-\lambda_1 + \delta\lambda_1(1-R)]^2}, \\ \frac{\partial^2 \ln l(\theta)}{\partial \delta \partial R} &= \frac{2n_{00}\delta\lambda_1^2(1-\lambda_1)^2}{[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)]^2} + \frac{n_{10}\lambda_1(\lambda_1-1)}{[1-\lambda_1 + \delta\lambda_1(1-R)]^2}, \\ \frac{\partial^2 \ln l(\theta)}{\partial \lambda_1^2} &= \frac{2n_{00}\{\delta^2(R-1)[1-2\lambda_1^2 + 2\lambda_1 - \delta^2\lambda_1^2(R-1)] - (1-\lambda_1)^2\}}{[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)]^2} - \frac{n_{11}R^2}{(1-R\lambda_1)^2} + \\ &\quad \frac{2(n_{00} + n_{10} + n_{20})(\delta-1)^2}{[1 + (\delta-1)\lambda_1]^2} + \frac{2n_{01}[R-1 - (R\lambda_1-1)^2]}{(1+R\lambda_1^2-2\lambda_1)^2} - \frac{n_{10}[\delta(1-R)-1]^2}{[1-\lambda_1 + \delta\lambda_1(1-R)]^2} - \\ &\quad \frac{n_{10} + n_{11} + 2n_{20} + 2n_{21}}{\lambda_1^2}, \\ \frac{\partial^2 \ln l(\theta)}{\partial R \partial \lambda_1} &= \frac{2n_{00}\delta^2\lambda_1(1-\lambda_1)}{[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)]^2} + \frac{2n_{01}\lambda_1(1-\lambda_1)}{(1+R\lambda_1^2-2\lambda_1)^2} - \frac{n_{10}\delta}{[1-\lambda_1 + \delta\lambda_1(1-R)]^2} - \\ &\quad \frac{n_{11}}{(1-R\lambda_1)^2}, \\ \frac{\partial^2 \ln l(\theta)}{\partial R^2} &= \frac{-n_{00}\delta^4\lambda_1^4}{[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)]^2} - \frac{n_{01}\lambda_1^4}{(1+R\lambda_1^2-2\lambda_1)^2} - \frac{n_{10}\delta^2\lambda_1^2}{[1-\lambda_1 + \delta\lambda_1(1-R)]^2} - \\ &\quad \frac{n_{11}\lambda_1^2}{(1-R\lambda_1)^2} - \frac{n_{20} + n_{21}}{R^2}. \end{aligned}$$

利用以上的二阶偏导数, 并将以下 n_{ki} ($k=0, 1, 2; i=0, 1$)的期望

$$\begin{aligned} E(n_{00}) &= \frac{n_{+0}[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)]}{[1 + (\delta-1)\lambda_1]^2}, E(n_{01}) = n_{+1}(1 + R\lambda_1^2 - 2\lambda_1), \\ E(n_{10}) &= \frac{2n_{+0}\delta\lambda_1[1-\lambda_1 + \delta\lambda_1(1-R)]}{[1 + (\delta-1)\lambda_1]^2}, E(n_{11}) = 2n_{+1}\lambda_1(1-R\lambda_1), \end{aligned}$$

$$E(n_{20}) = \frac{n_{+0}\delta^2\lambda_1^2 R}{[1 + (\delta - 1)\lambda_1]^2}, E(n_{21}) = n_{+1}R\lambda_1^2$$

代入相应的表达式, 可以得到参数 $\theta = (\delta, \lambda_1, R)^\top$ 的Fisher信息矩阵

$$I(\delta, \lambda_1, R) = \begin{pmatrix} E(-\frac{\partial^2 \ln l(\theta)}{\partial \delta^2}) & E(-\frac{\partial^2 \ln l(\theta)}{\partial \delta \partial \lambda_1}) & E(-\frac{\partial^2 \ln l(\theta)}{\partial \delta \partial R}) \\ E(-\frac{\partial^2 \ln l(\theta)}{\partial \lambda_1 \partial \delta}) & E(-\frac{\partial^2 \ln l(\theta)}{\partial \lambda_1^2}) & E(-\frac{\partial^2 \ln l(\theta)}{\partial \lambda_1 \partial R}) \\ E(-\frac{\partial^2 \ln l(\theta)}{\partial R \partial \delta}) & E(-\frac{\partial^2 \ln l(\theta)}{\partial R \partial \lambda_1}) & E(-\frac{\partial^2 \ln l(\theta)}{\partial R^2}) \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix}.$$

其中

$$I_{11} = \frac{-2n_{+0}\lambda_1^2(R-1)[(1-\lambda_1)^2 + \delta^2\lambda_1^2(1-R)]}{[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)][1 + (\delta-1)\lambda_1]^2} - \frac{2n_{+0}\lambda_1^2}{[1 + (\delta-1)\lambda_1]^2} + \frac{2n_{+0}\lambda_1}{\delta[1 + (\delta-1)\lambda_1]} + \frac{2n_{+0}\delta\lambda_1^3(1-R)^2}{[1 - \lambda_1 + \delta\lambda_1(1-R)][1 + (\delta-1)\lambda_1]^2},$$

$$I_{12} = \frac{-4n_{+0}\delta\lambda_1(1-\lambda_1)(R-1)}{[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)][1 + (\delta-1)\lambda_1]^2} - \frac{2n_{+0}\delta\lambda_1(1-R)}{[1 - \lambda_1 + \delta\lambda_1(1-R)][1 + (\delta-1)\lambda_1]^2} + \frac{2n_{+0}}{[1 + (\delta-1)\lambda_1]^2},$$

$$I_{13} = \frac{-2n_{+0}\delta^2\lambda_1^3(\lambda_1-1)(R-1)}{[1 - \lambda_1 + \delta\lambda_1(1-R)][(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)][1 + (\delta-1)\lambda_1]},$$

$$I_{22} = \frac{-2n_{+0}\{\delta^2(R-1)[1 - 2\lambda_1^2 + 2\lambda_1 - \delta^2\lambda_1^2(R-1)] - (1-\lambda_1)^2\}}{[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)][1 + (\delta-1)\lambda_1]^2} - \frac{2n_{+0}(\delta-1)^2}{[1 + (\delta-1)\lambda_1]^2} + \frac{2n_{+0}\delta\lambda_1[\delta(1-R) - 1]^2}{[1 - \lambda_1 + \delta\lambda_1(1-R)][1 + (\delta-1)\lambda_1]^2} - \frac{2n_{+1}[R-1 - (R\lambda_1-1)^2]}{1 + R\lambda_1^2 - 2\lambda_1} + \frac{2n_{+1}\lambda_1 R^2}{1 - R\lambda_1} + \frac{2n_{+0}\delta + 2n_{+1}[1 + (\delta-1)\lambda_1]}{\lambda_1[1 + (\delta-1)\lambda_1]},$$

$$I_{32} = \frac{-2n_{+0}\delta^2\lambda_1(1-\lambda_1)}{[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)][1 + (\delta-1)\lambda_1]^2} + \frac{2n_{+0}\delta^2\lambda_1}{[1 - \lambda_1 + \delta\lambda_1(1-R)][1 + (\delta-1)\lambda_1]^2} - \frac{2n_{+1}\lambda_1(1-\lambda_1)}{1 + R\lambda_1^2 - 2\lambda_1} + \frac{2n_{+1}\lambda_1}{1 - R\lambda_1},$$

$$I_{33} = \frac{n_{+0}\delta^4\lambda_1^4}{[(1-\lambda_1)^2 + \delta^2\lambda_1^2(R-1)][1 + (\delta-1)\lambda_1]^2} + \frac{2n_{+0}\delta^3\lambda_1^3}{[1 - \lambda_1 + \delta\lambda_1(1-R)][1 + (\delta-1)\lambda_1]^2} + \frac{n_{+1}\lambda_1^4}{1 + R\lambda_1^2 - 2\lambda_1} + \frac{2n_{+1}\lambda_1^3}{1 - R\lambda_1} + \frac{\lambda_1^2\{n_{+0}\delta^2 + n_{+1}[1 + (\delta-1)\lambda_1]^2\}}{R[1 + (\delta-1)\lambda_1]^2},$$

$$I_{21} = I_{12}, I_{23} = I_{32}, I_{31} = I_{13}.$$

通过求解Fisher信息阵的逆矩阵的第(1,1)个元素, 可以得到非独立假设下感兴趣参数 δ 的方差, 即有

$$\text{Var}(\hat{\delta}) = \left[I_{11} - \begin{pmatrix} I_{12} & I_{13} \end{pmatrix} \begin{pmatrix} I_{22} & I_{23} \\ I_{23} & I_{33} \end{pmatrix}^{-1} \begin{pmatrix} I_{12} \\ I_{13} \end{pmatrix} \right]^{-1} = \frac{I_{22}I_{33} - I_{23}^2}{I_{11}I_{22}I_{33} + 2I_{12}I_{13}I_{23} - I_{11}I_{23}^2 - I_{13}^2I_{22} - I_{12}^2I_{33}}.$$

在独立假设下, $R = 1$ 为一个常数, 则参数 $\theta = (\delta, \lambda_1)^\top$ 的 Fisher 信息矩阵为

$$I(\delta, \lambda_1) = \begin{pmatrix} E\left(-\frac{\partial^2 \ln l(\theta)}{\partial \delta^2}\right) & E\left(-\frac{\partial^2 \ln l(\theta)}{\partial \delta \partial \lambda_1}\right) \\ E\left(-\frac{\partial^2 \ln l(\theta)}{\partial \lambda_1 \partial \delta}\right) & E\left(-\frac{\partial^2 \ln l(\theta)}{\partial \lambda_1^2}\right) \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix},$$

其中

$$\begin{aligned} I_{11} &= E\left(-\frac{\partial^2 \ln l(\theta)}{\partial \delta^2}\right) = \frac{2n_{+0}\lambda_1(1-\lambda_1)}{\delta[1+(\delta-1)\lambda_1]^2}, \\ I_{12} &= E\left(-\frac{\partial^2 \ln l(\theta)}{\partial \delta \partial \lambda_1}\right) = \frac{2n_{+0}}{[1+(\delta-1)\lambda_1]^2} = I_{21}, \\ I_{22} &= E\left(-\frac{\partial^2 \ln l(\theta)}{\partial \lambda_1^2}\right) = \frac{2n_{+0}\delta + 2n_{+1}[1+(\delta-1)\lambda_1]^2}{\lambda_1(1-\lambda_1)[1+(\delta-1)\lambda_1]^2}. \end{aligned}$$

通过求解 Fisher 信息阵的逆矩阵的第(1,1)个元素, 可以得到独立假设下感兴趣参数 δ 的方差为

$$\text{Var}(\hat{\delta}) = (I_{11} - I_{12}I_{22}^{-1}I_{21})^{-1} = \frac{\delta[1+(\delta-1)\lambda_1]^2\{2n_{+0}\delta + 2n_{+1}[1+(\delta-1)\lambda_1]^2\}}{4n_{+0}n_{+1}\lambda_1\{(1-\lambda_1)^3 + \delta\lambda_1[\lambda_1^2(2-\delta) + \lambda_1(\delta-4) + 2]\}}.$$

参考文献:

- [1] Rosner B. Statistical methods in ophthalmology: An adjustment for the intraclass correlation between eyes[J]. Biometrics, 1982, 38(1): 105-114.
- [2] Rosner B, Milton R C. Significance testing for correlated binary outcome data[J]. Biometrics, 1988, 44(2): 505-512.
- [3] Donner A. Statistical methods in ophthalmology: an adjusted chi-square approach[J]. Biometrics, 1989, 45(2): 605-611.
- [4] Tang Niansheng, Tang Manlai, Qiu Shifang. Testing the equality of proportions for correlated otolaryngologic data[J]. Computational Statistics and Data Analysis, 2008, 52(7): 3719-3729.
- [5] Tang Niansheng, Qiu Shifang, Tang Manlai, et al. Asymptotic confidence interval construction for proportion difference in medical studies with bilateral data[J]. Statistical Methods in Medical Research, 2011, 20(3): 233-259.
- [6] Zhuang Tingting, Tian Guoliang, Ma Changxing. Confidence intervals for proportion ratios of stratified correlated bilateral data[J]. Journal of Biopharmaceutical Statistics, 2019, 29(1): 203-225.
- [7] 邱世芳, 郭黎萱. 双边试验设计下基于区间估计的样本量的确定[J]. 重庆理工大学学报(自然科学), 2018, 32(4): 219-229.
- [8] 孟令宾, 李二倩, 田茂再. 基于鞍点逼近的二项抽样下优势比的置信区间构造[J]. 数理统计与管理, 2017, 36(1): 85-102.
- [9] 钱政超, 张晨阳, 孟令宾, 等. 二项抽样下基于鞍点逼近方法的流行病相对风险置信区间构造[J]. 数学的实践与认识, 2014, 44(21): 204-217.
- [10] 王维贤, 田茂再. 鞍点逼近下相关差的置信区间构造[J]. 高校应用数学学报, 2019, 34(3): 264-272.
- [11] 古丽斯坦·库尔班尼牙孜, 孟丽君, 田茂再. 配对设计中条件优势比的置信区间构造[J]. 系统科学与数学, 2021, 41(3): 824-836.

- [12] 江绍萍, 李慧敏, 缪清. 逆抽样条件下配对试验设计中优比的统计推断[J]. 统计与决策, 2018, 34(17): 5-8.
- [13] 江绍萍. 逆抽样条件下含结构零的 2×2 列联表中相对差的估计[J]. 统计与决策, 2010(2): 14-16.
- [14] 古丽斯坦·库尔班尼牙孜, 田茂再. 独立逆抽样下优势比检验统计量的构造[J]. 统计与决策, 2022, 38(5): 5-10.
- [15] Zou Guangyong, Donner A, Qiu Shifang. MOVER-R for confidence intervals of ratios[J]. Wiley StatsRef: Statistics Reference Online, 2018, 1-11.
- [16] Donner A, Zou Guangyong. Interval estimation for a difference between intraclass kappa statistics[J]. Biometrics, 2002, 58(1): 209-215.
- [17] 杨梓滢, 余雯慧, 张玉, 等. 一种基于MOVER法中介效应置信区间的构造方法[J]. 中国卫生统计, 2021, 38(2): 204-207+214.
- [18] 段重阳. 配对设计率差及率比置信区间构建新方法[D]. 广州: 南方医科大学, 2017.
- [19] Shao Jun, Tu Dongsheng. The Jackknife and Bootstrap[M]. London: Chapman and Hall, 1993.
- [20] Tang Niansheng, Yu Bin. Simultaneous confidence interval for assessing non-inferiority with assay sensitivity in a three-arm trial with binary endpoints[J]. Pharmaceutical Statistics, 2020, 19(5): 518-531.
- [21] Mandel E M, Bluestone C D, Rockette H E, et al. Duration of effusion after antibiotic treatment for acute Otitis Media: comparison of Cefaclor and Amoxicillin[J]. Pediatric infectious disease, 1982, 1(5): 310-316.

Interval estimation based on odds ratio in bilateral experimental design

LI Hong-qiong^{1,2}, JIANG Shao-ping¹, YU Yang-tao¹, ZHAO Shuo¹, DONG Zhu-yue¹

(1. School of Mathematics and Computer Science, Yunnan Minzu University, Kunming 650500, China;

2. School of Mathematics and Statistics, Chuxiong Normal University, Chuxiong 675000, China)

Abstract: For the problem of confidence interval estimation of the odds ratio under a bilateral experimental design, four types of confidence intervals based on Wald statistic, logarithm-Wald statistic, MOVER type confidence interval and Bootstrap resampling confidence interval are proposed based on bilateral data under independent and non-independent hypotheses. The statistical performance of these intervals is compared through simulation studies of empirical coverage probability, mean of the lower confidence limit, the required sample size to achieve a fixed power in the experiment, and empirical coverage probability under this sample. The results show that the logarithm-Wald interval based on the sample variance yields satisfactory outcomes, It not only ensures that the empirical coverage probability is close to the confidence level but also requires a smaller sample size to achieve a fixed power under the same parameter setting. Finally, the proposed methods are validated using clinical data from the field of otology.

Keywords: bilateral experimental design; odds ratio; confidence interval; empirical coverage probability; sample size

MR Subject Classification: 62F03