

偏正态数据的三臂非劣性检验

吴 祺^{1,2}, 梁 帆^{1,2*}, 吴刘仓^{1,2*}

(1. 昆明理工大学 理学院, 云南昆明 650500;

2. 昆明理工大学 应用统计学研究中心, 云南昆明 650500)

摘 要: 非劣性检验是一种重要的药效检验方法, 旨在证明某种新的药物在预先设定的非劣效界值下不劣于标准药物. 三臂非劣性检验包括安慰剂组、标准药物组和新药物组, 可同时检验标准药物相对于安慰剂的优效性以及新药物相较于标准药物的非劣性. 在现实生活中实验得到的数据往往不是严格的正态分布, 而是具有偏态特征, 因此需要研究用于偏正态三臂非劣性检验的新方法. 该文考虑用矩估计的方法来估计未知参数, 利用中心极限定理构造近似检验统计量, 并进一步基于Bootstrap方法, 利用Bootstrap样本估计未知参数, 构造Bootstrap检验统计量, 给出检验方法. 在不同的已知参数条件下通过Monte Carlo模拟得到检验犯第一类错误的概率、功效和样本量, 模拟结果表明该检验方法具有有效性和稳定性. 最后将该方法应用于一项HIV治疗的实例数据以判断其非劣性.

关键词: 三臂非劣性检验; 偏正态分布; Bootstrap抽样; Monte Carlo模拟

中图分类号: O212.1

文献标识码: A **文章编号:** 1000-4424(2026)01-0001-22

§1 引 言

随着生物医学技术的不断发展, 为满足人类治疗各种疾病的需求, 越来越多新研发的药物不断涌现. 尽管临床上针对某种疾病已有能够治疗的有效药物, 但这些药物可能存在成本高或副作用大等缺点. 因此科研人员有必要继续研发价格更低廉或副作用更小的新药物. 而在新药的研发中, 如何有效评价新药是否可以与已有药效的标准治疗药物相互替换使用, 是临床医学研究中非常关注的问题. 这就需要发展合适的统计方法对药物药效进行科学合理的生物评价. 非劣性评价是生物评价中备受关注的的一个重要方法.

非劣性评价主要是用来验证新药的药效在一定的非劣性界值下不比目前使用的标准药的药效差. 传统的非劣性评价中, 由于考虑伦理道德等因素, 仅对新药和标准药进行两臂非劣性试

收稿日期: 2024-01-02 修回日期: 2024-05-02

*通讯作者, email: 梁帆, fional1217@sina.com; 吴刘仓, wuliucang@163.com

基金项目: 云南省教育厅科学研究基金项目“大数据下三臂非劣性药效检验”(2024J0082); 云南省科技厅科技计划项目“大数据下序贯多任务随机试验非劣性检验”(202401AU070137); 国家自然科学基金(12261051)

验. 其中非劣性界值的选择则取决于统计推断和临床判断的结合^[1]. 然而, 由于两臂非劣性试验中不包含安慰剂, 因此难以证明标准药的有效性^[2], 即检验的灵敏度评价无法确定. 为了解决上述问题, 在所在地法律和伦理道德允许的前提下, 用安慰剂做对照的三臂非劣性试验是一个不错的选择. 三臂非劣性试验包含试验药(新药)组, 标准药(阳性)对照组和安慰剂对照组, 其不仅可以比较新药是否非劣(效)性于标准药, 也可以比较新药是否优于安慰剂. 这样, 既可以保证药物的有效性, 也可以达到检验内部和外部的确证. 目前, 众多学者关于三臂非劣性试验进行了深入研究, Pigeot等人^[3]采用保留标准药物效应法来确定三臂非劣性检验中的非劣性界值, 推导出Fieller置信区间, 并给出了样本量估计. Kieser等人^[4]提出三臂非劣性试验的疗效数据服从二项分布时的统计检验方法, 计算出犯第一类错误率, 并推导出了近似样本量公式. Hasler等人^[5]通过对Pigeot所提出的检验过程进行改进, 给出异方差情况下的检验过程. Kim等人^[6]提出了在两臂非劣性试验中采用Wilcoxon秩和检验的非参数检验过程, 以及在三臂非劣性试验中采用线性对比检验的非参数检验过程, 并通过Monte Carlo模拟研究, 对提出的非参数过程和参数过程进行了比较. Li等人^[7]利用封闭测试原则建立了三臂非劣性试验中多重比较的分层检验过程, 推导出了样本量最佳分配比例的公式. Tang等人^[8]在二元响应变量的三臂非劣性检验下提出了无条件准确检验和近似检验. Ghosh等人^[9]考虑了Poisson离散响应变量下结合灵敏度的非劣性条件检验. Tang等人^[10]构建了4种正态数据下三臂试验中同时评估非劣性和灵敏度的置信区间, 并给出了相应的算法和样本量计算公式. Tang等人^[11]基于正态数据利用保留标准药物效应法确定非劣性界值, 在半参数Bayesian框架下对灵敏度和非劣性进行同时检验, 对感兴趣参数的先验分布设定为DP先验. 且给出了对应的Gibbs抽样算法和后验分布形式.

然而这些文献大多考虑的是三臂非劣性试验所包含的试验药(新药)组, 标准药(阳性)对照组和安慰剂对照组的疗效数据都是正态分布的情况, 并没有考虑疗效数据服从偏正态分布的三臂非劣性检验. 但是在实际应用中, 数据更多的是呈现出单峰, 非对称的偏正态分布特征, 当数据呈现偏正态分布时, 意味着数据的分布不是完全对称的, 而是有一个长尾延伸到较高的值或较低的值, 如果假定数据服从正态分布进行推断, 可能导致已有的统计检验方法的估计不准确, 则得到的统计决策可能会出现偏差. 在非劣(效)性研究中, 研究者希望证明一种新治疗至少不比现有治疗差. 如果数据分布偏斜, 使用不恰当的统计方法可能会影响非劣性的结论, 适当的数据分析方法可以确保研究结果的可靠性. 因此为了帮助研究者更好地理解治疗效果, 为临床实践提供有价值的见解, 本文考虑疗效数据服从偏正态分布的情况下的三臂非劣性检验.

许多学者从不同方面对偏正态数据进行了研究. 比如, Azzalini^[12]最早提出了偏正态分布的概念, 构造出偏正态分布的概率密度函数, 并推导出相应的性质. 马婷等人^[13]通过极大似然迭代算法给出了联合模型参数的估计方法. Phontita等人^[14-15]提出了偏正态总体的尺度参数和偏度参数都已知的情况下, 位置参数的点估计方法, 并将该方法的思想应用于尺度和偏度参数未知的情况, 给出了矩估计和极大似然估计的方法. Wu等人^[16]研究了联合位置尺度和偏度的统计诊断问题. Morán-Vásquez等人^[17]推导出将对数偏态正态分布的分位数与其参数相关联的性质, 并通过将对数偏态分布研究了正偏态响应变量的分位数与一组解释变量之间的关系. 叶仁道等人^[18]研究了多个尺度参数和偏度参数未知情况下的偏正态总体, 基于Bootstrap方法研究了位置参数的齐次检验问题和共同位置参数的区间估计问题. 方冰妮^[19]研究了在尺度参数和偏度参数

未知的情况下的单个和两个偏正态总体, 运用中心参数化和Bootstrap方法, 得到位置参数的假设检验和区间估计. 本文考虑将文献[19]的方法应用到三臂非劣性试验中, 给出偏正态分布下位置参数、尺度参数和偏度参数的矩估计, 分别构造似检验统计量 and 的Bootstrap检验统计量, 给出了相应的算法并讨论了样本量的确定问题.

本文的主要工作如下: §2介绍了本文所研究的三臂非劣性检验的模型及其相关符号. §3首先给出疗效数据为偏正态数据时的参数矩估计, 然后构造相应的检验统计量, 进而通过Bootstrap方法给出判断非劣性的统计决策. §4通过模拟研究得到犯第一类错误的概率、检验功效以及所需样本量. §5将本文所提出的方法应用于一个HIV实例中检验非劣性. §6对本文进行总结和展望.

§2 模型与符号

三臂非劣性检验包含三个药物组, 分别是试验药(新药)组 E (experimental), 对照(标准药)组 R (reference)和安慰剂组 P (lacebo). 令

$$X_{kj} \sim SN(\mu_k, \sigma_k^2, \lambda_k), k = E, R, P, j = 1, \dots, n_k. \quad (1)$$

其中 $X_{k,j}$ 表示第 k 组第 j 个个体的疗效值, μ_k 表示第 k 组的疗效值, n_k 表示第 k 组的样本量, σ_k^2 表示第 k 组的方差, λ_k 表示第 k 组的偏度. 不失一般性, 假定 μ_k 的值越大表示第 k 组疗效效果越好. 通常两臂非劣性检验可表示为

$$H_0: \mu_E - \mu_R \leq \delta \quad \text{vs} \quad H_1: \mu_E - \mu_R > \delta. \quad (2)$$

其中 $\delta < 0$ 是根据保留标准药物效应法确定的非劣性界值. 在三臂非劣性检验中, 根据文献[20]采用阳性保留效应法. 根据CPMP(欧洲药品评价局专利药品委员会)的指导意见, 其中 f 可在区间 $[-\frac{1}{2}, -\frac{1}{3}]$ 内取值. 因此, 三臂非劣性检验可表示为

$$H_0: \mu_E - \mu_R \leq f(\mu_R - \mu_P) \quad \text{vs} \quad H_1: \mu_E - \mu_R > f(\mu_R - \mu_P). \quad (3)$$

令 $\theta = 1 + f$, 则

$$H_0: \mu_E - \theta\mu_R - (1 - \theta)\mu_P \leq 0 \quad \text{vs} \quad H_1: \mu_E - \theta\mu_R - (1 - \theta)\mu_P > 0. \quad (4)$$

其中 $\theta \in [0.5, 1]$ 表示标准药与安慰剂疗效之差的一部分. 显然拒绝原假设 H_0 就证明了新药对标准药的非劣性.

§3 非劣性检验的方法

根据文献[19], 首先给出偏正态分布的参数估计, 然后构造(4)的近似检验统计量和Bootstrap检验统计量.

3.1 偏正态总体的参数矩估计

若 $X \sim SN(\mu, \sigma^2, \lambda)$, 则

- 1) X 的矩生成函数为 $M_X(t) = 2 \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \Phi(t\sigma\delta)$.

- 2) X 的前三阶矩为 $E(X) = \mu + b\sigma\delta$, $S^2 = \text{Var}(X) = \sigma^2(1 - b^2\delta^2)$, $S^3 = \sigma^3\delta^3b(2b^2 - 1)$.

其中 $b = (2/\pi)^{1/2}$, $\delta = \lambda/(1 + \lambda^2)^{1/2}$. 通过对 $\mathbf{1}$ 中的 $M_X(t)$ 分别关于 t 求前三阶导, 易得 X 的前三阶原点矩, 从而进一步得到 $\mathbf{2}$ 中的均值, 方差和三阶中心矩. 则在三臂非劣性检验中, 设 X_{kj} , $k = E, R, P, j = 1, \dots, n_k$ 是根据模型(1)产生的试验数据, 则样本均值, 样本二阶中心距和样本三阶中心距可分别表示为

$$\bar{X}_k = \frac{1}{n_k} \sum_{j=1}^{n_k} X_{kj}, \quad S_{k2} = \frac{1}{n_k} \sum_{j=1}^{n_k} (X_{kj} - \bar{X}_k)^2, \quad S_{k3} = \frac{1}{n_k} \sum_{j=1}^{n_k} (X_{kj} - \bar{X}_k)^3. \quad (5)$$

定理3.1 设 $\delta_k = \lambda_k/(1 + \lambda_k^2)^{1/2}$, 且 $\delta_k \in (-1, 1)$, 若 $X_{kj} \sim SN(\mu_k, \sigma_k^2, \lambda_k)$, $k = E, R, P$, $j = 1, \dots, n_k$, 则参数 $(\mu_k, \sigma_k^2, \lambda_k)$ 的矩估计量可表示为

$$\hat{\mu}_k = \bar{X}_k - cS_{k3}^{1/3}, \quad \hat{\sigma}_k^2 = S_{k2} + c^2S_{k3}^{2/3}, \quad \hat{\lambda}_k = \hat{\delta}_k / (1 + \hat{\delta}_k^2)^{1/2}. \quad (6)$$

其中 $k = E, R, P$, $b = (2/\pi)^{1/2}$, $c = [2/(4 - \pi)]^{1/3}$, $\hat{\delta}_k = cS_{k3}^{1/3}/b(S_{k2} + c^2S_{k3}^{2/3})^{1/2}$, \bar{X}_k , S_{k2} , S_{k3} 分别表示 X_k 的样本均值, 样本二阶中心距和样本三阶中心距. 根据定理3.1可以得到偏正态分布中参数的矩估计. 以下的定理3.2给出偏正态分布下参数矩估计的无偏性.

定理3.2 若 $X_{kj} \sim SN(\mu_k, \sigma_k^2, \lambda_k)$, $k = E, R, P, j = 1, \dots, n_k$, 则由定理3.1得到参数的矩估计 $\hat{\mu}_k$, $\hat{\sigma}_k^2$ 和 $\hat{\delta}_k$ 为无偏估计.

以下是定理的证明.

$$\begin{aligned} E(\hat{\mu}_k) &= E(\bar{X}_k - cS_{k3}^{1/3}) = \mu_k + b\sigma_k\delta_k - c[\sigma_k^3\delta_k^3b(2b^2 - 1)]^{1/3} = \\ &= \mu_k + b\sigma_k\delta_k - \sigma_k\delta_k \left(\frac{2}{4 - \pi}\right)^{1/3} \left[\sqrt{\frac{2}{\pi}} \left(\frac{4}{\pi} - 1\right)\right]^{1/3} = \\ &= \mu_k + b\sigma_k\delta_k - \sigma_k\delta_k \left(\frac{2}{\pi}\right)^{\frac{3}{2 \times 3}} = \mu_k. \end{aligned} \quad (7)$$

$$\begin{aligned} E(\hat{\sigma}_k^2) &= E(S_{k2} + c^2S_{k3}^{2/3}) = \sigma_k^2(1 - b^2\delta_k^2) + \left(\frac{2}{4 - \pi}\right)^{2/3} S_{k3}^{2/3} = \\ &= \sigma_k^2 - b^2\sigma_k^2\delta_k^2 + \left\{\left(\frac{2}{4 - \pi}\right) [\sigma_k^3\delta_k^3b(2b^2 - 1)]\right\}^{2/3} = \\ &= \sigma_k^2 - b^2\sigma_k^2\delta_k^2 + \sigma_k^2\delta_k^2 \left(\frac{2}{4 - \pi} \sqrt{\frac{2}{\pi}} \frac{4 - \pi}{\pi}\right)^{2/3} = \\ &= \sigma_k^2 - b^2\sigma_k^2\delta_k^2 + \sigma_k^2\delta_k^2 \left(\frac{2}{\pi}\right)^{\frac{3 \times 2}{2 \times 3}} = \sigma_k^2 - b^2\sigma_k^2\delta_k^2 + \sigma_k^2\delta_k^2b^2 = \sigma_k^2. \end{aligned} \quad (8)$$

$$\begin{aligned} E(\hat{\delta}_k) &= E\left[cS_{k3}^{1/3}/b(S_{k2} + c^2S_{k3}^{2/3})^{1/2}\right] = \\ &= b\sigma_k\delta_k/b[\sigma_k^2(1 - b^2\delta_k^2) + b^2\sigma_k^2\delta_k^2]^{1/2} = b\sigma_k\delta_k/b\sigma_k = \delta_k. \end{aligned} \quad (9)$$

3.2 检验统计量

记线性组合 $Y = X_E - \theta X_R - (1 - \theta) X_P$, 由定理3.1和定理3.2可得

$$\begin{aligned} E(Y) &= E(X_E - \theta X_R - (1 - \theta) X_P) = \\ &= \mu_E + b\sigma_E\delta_E - \theta(\mu_R + b\sigma_R\delta_R) - (1 - \theta)(\mu_P + b\sigma_P\delta_P) = \\ &= \mu_E - \theta\mu_R - (1 - \theta)\mu_P + b(\sigma_E\delta_E - \theta\sigma_R\delta_R - (1 - \theta)\sigma_P\delta_P), \end{aligned} \quad (10)$$

$$\begin{aligned} \text{Var}(Y) &= \text{Var}(X_E - \theta X_R - (1 - \theta) X_P) = \\ &= \text{Var}(X_E) + \theta^2 \text{Var}(X_R) + (1 - \theta)^2 \text{Var}(X_P) = \\ &= \sigma_E^2(1 - b^2\delta_E^2) + \theta^2\sigma_R^2(1 - b^2\delta_R^2) + (1 - \theta)^2\sigma_P^2(1 - b^2\delta_P^2). \end{aligned} \quad (11)$$

当(4) H_0 成立时, 由中心极限定理可得

$$T = \frac{\bar{X}_E - \theta\bar{X}_R - (1 - \theta)\bar{X}_P - b(\sigma_E\delta_E - \theta\sigma_R\delta_R - (1 - \theta)\sigma_P\delta_P)}{\sqrt{\frac{\sigma_E^2(1 - b^2\delta_E^2)}{n_E} + \frac{\theta^2\sigma_R^2(1 - b^2\delta_R^2)}{n_R} + \frac{(1 - \theta)^2\sigma_P^2(1 - b^2\delta_P^2)}{n_P}}}. \quad (12)$$

其中 $k = E, R, P$, T 是(4)的近似检验统计量.

然而在实际问题中 σ_k 和 δ_k 往往是未知的, 因此可将其矩估计量 $\hat{\sigma}_k, \hat{\delta}_k$ 代入 T 的表达式, 则有如下检验统计量

$$T_M = \frac{\bar{X}_E - \theta\bar{X}_R - (1 - \theta)\bar{X}_P - b(\hat{\sigma}_E\hat{\delta}_E - \theta\hat{\sigma}_R\hat{\delta}_R - (1 - \theta)\hat{\sigma}_P\hat{\delta}_P)}{\sqrt{\frac{\hat{\sigma}_E^2(1 - b^2\hat{\delta}_E^2)}{n_E} + \frac{\theta^2\hat{\sigma}_R^2(1 - b^2\hat{\delta}_R^2)}{n_R} + \frac{(1 - \theta)^2\hat{\sigma}_P^2(1 - b^2\hat{\delta}_P^2)}{n_P}}}. \quad (13)$$

其中 $\bar{X}_k, k = E, R, P$ 是样本均值, T_M 的精确分布是未知的, 因而接下来考虑采用Bootstrap方法构造检验统计量进行检验.

在原假设 H_0 成立时, 令 $X_{BMk1}, X_{BMk2}, \dots, X_{BMkn_k}$ 是来自 $SN(\hat{\mu}_k, \hat{\sigma}_k^2, \hat{\lambda}_k)$ 的Bootstrap样本, 其中 $\hat{\mu}_k, \hat{\sigma}_k^2, \hat{\lambda}_k$ 是由定理3.1得到 $\mu_k, \sigma_k^2, \lambda_k$ 的矩估计量, 且 $X_{BMk1}, X_{BMk2}, \dots, X_{BMkn_k}$ 的样本均值, 样本二阶中心距和样本三阶中心距分别为 $\bar{X}_{BMk}, S_{BMk2}, S_{BMk3}, k = E, R, P$. 由定理3.1, $SN(\hat{\mu}_k, \hat{\sigma}_k^2, \hat{\lambda}_k)$ 的总体参数的矩估计值分别为 $\hat{\mu}_{BMk} = \bar{X}_{BMk} - cS_{BMk3}^{1/3}, \hat{\sigma}_{BMk}^2 = S_{BMk2} + c^2S_{BMk3}^{2/3}, \hat{\lambda}_{BMk} = \hat{\delta}_{BMk}/(1 + \hat{\delta}_{BMk}^2)^{1/2}, \delta_{BMk} = cS_{BMk3}^{1/3}/b(S_{BMk2} + c^2S_{BMk3}^{2/3})^{1/2}$. 接下来考虑利用中心极限定理构造检验统计量, 所以将三个总体的样本均值 \bar{X}_{BMk} 和参数的矩估计值 $(\hat{\mu}_{BMk}, \hat{\sigma}_{BMk}^2, \hat{\lambda}_{BMk}, \hat{\delta}_{BMk}), k = E, R, P$ 代入 T_M 得到Bootstrap检验统计量为

$$T_{BM} = \frac{\bar{X}_{BME} - \theta\bar{X}_{BMR} - (1 - \theta)\bar{X}_{BMP} - b(\hat{\sigma}_{BME}\hat{\delta}_{BME} - \theta\hat{\sigma}_{BMR}\hat{\delta}_{BMR} - (1 - \theta)\hat{\sigma}_{BMP}\hat{\delta}_{BMP})}{\sqrt{\frac{\hat{\sigma}_{BME}^2(1 - b^2\hat{\delta}_{BME}^2)}{n_E} + \frac{\theta^2\hat{\sigma}_{BMR}^2(1 - b^2\hat{\delta}_{BMR}^2)}{n_R} + \frac{(1 - \theta)^2\hat{\sigma}_{BMP}^2(1 - b^2\hat{\delta}_{BMP}^2)}{n_P}}}. \quad (14)$$

此时由 T_M, T_{BM} 可得

$$p_1 = P(T_{BM} \geq T_M). \quad (15)$$

若 $p_1 < \alpha$, 则在显著性水平 α 下拒绝原假设 H_0 , 即认为新药相较于标准药具有非劣性.

通过Monte Carlo模拟计算犯第一类错误的概率、检验功效和样本量.

在原假设成立时生成模拟数据, 通过Bootstrap方法得到检验结果, 通过多次模拟累计拒绝原假设的次数, 得到该检验过程犯第一类错误的概率, 步骤如下.

步骤1: 在原假设成立时生成模拟数据, 得到参数的矩估计值和检验统计量 T_M .

步骤2: 通过Bootstrap抽样数据得到 T_{BM} , 计算Bootstrap参数矩估计.

步骤3: 多次重复步骤2, 通过Bootstrap抽样得到检验的 p 值.

步骤4: 多次重复步骤1-步骤3, 通过多次模拟累计拒绝原假设的次数, 得到犯第一类错误的概率.

在备择假设成立时生成模拟数据, 可类似的通过Bootstrap方法得到检验结果, 通过多次模拟累计拒绝原假设即接受备择假设的次数, 得到检验功效; 同时在样本量不同的情况下计算出相应的功效, 由此可得达到确定的功效所需的样本量. 计算功效的步骤如下.

步骤1: 在备择假设成立时生成模拟数据, 得到参数的矩估计值和检验统计量 T_M .

步骤2: 通过Bootstrap抽样数据得到 T_{BM} , 计算Bootstrap参数矩估计.

步骤3: 多次重复步骤2, 通过Bootstrap抽样得到检验的 p 值.

步骤4: 多次重复步骤1-步骤3, 通过多次模拟累计接受备择假设的次数, 得到检验的功效.

§4 模拟研究

本节通过Monte Carlo模拟, 从数值上研究上述检验方法犯第一类错误的概率和功效以及相应样本量. 下面针对假设检验问题(4), 给出基于矩估计的Bootstrap检验方法犯第一类错误概率和功效的算法. 算法1给出了计算犯第一类错误概率的详细计算步骤, 其中内循环 n_1 和外循环 n_2 都为1000.

在 H_1 成立的情况下, (15)给出了在矩估计下正确接受备择假设的概率, 继而由算法1可以类似的计算出§3所描述的方法的检验功效, 若要求功效达到80%, 同样由此方法可得到所需的样本量.

下面将基于§2中提出的检验过程和Pigeot的正态三臂非劣(效)性检验方法进行比较. 在实验、对照和安慰剂组的治疗结果数据都服从偏正态分布且分布的位置、尺度和偏度参数皆未知的情况下, 同时分别使用以上算法和Pigeot所提出的检验方法得到相应的一类错误.

考虑 $\theta = 0.8$ 和 $\theta = 0.6$ 两种情况, 当 $\theta = 0.8$ 时, 取 $\mu_E = 7.8, \mu_R = 8, \mu_P = 7$ 和 $\mu_E = 2.8, \mu_R = 3, \mu_P = 2$ 以及当 $\theta = 0.6$ 时, 取 $\mu_E = 7.6, \mu_R = 8, \mu_P = 7$ 和 $\mu_E = 2.6, \mu_R = 3, \mu_P = 2$, 令尺度参数为 $\sigma_E^2 = 3.2, \sigma_R^2 = 3.5, \sigma_P^2 = 3$, 偏度参数分别为 $(\lambda_E, \lambda_R, \lambda_P) = (0.3, 0.4, 0.5), (-0.1, -0.2, -0.3), (0.2, 0.4, 0.1), (-0.6, -0.5, -0.4)$. 考虑试验、阳性对照和安慰剂三组样本量比例为 $n_E : n_R : n_P = 1 : 1 : 1, n_E : n_R : n_P = 2 : 2 : 1$ 和 $n_E : n_R : n_P = 3 : 2 : 1$ 的情况, 同时考虑在三个不同的显著性水平0.05, 0.075和0.1下进行检验.

出于方便排版的原因, 表格展示在后文.

表1-表12展示了本文所提出的方法和Pigeot所提出的正态数据检验方法犯第一类错误概率的比较, 其中 n 表示总样本量, psn 表示所提出的方法犯第一类错误的概率, pn 表示Pigeot所提出的正态数据下犯第一类错误的概率. 表1-表3分别为 $\theta = 0.6$ 时 $\mu_E = 2.6, \mu_R = 3, \mu_P = 2$ 的情况下不同显著性水平犯第一类错误的概率, 从表1-表3展示的数据中可以看到, 不同的样本量分配比例情况下, 在两种假设检验方法中, 前者针对偏态的方法随着样本量的增大能够较好的控制犯第一类错误的概率, 且在样本量较大时更接近预设的显著性水平0.05, 0.075和0.1, 而如果假定数据服从正态分布, 则犯第一类错误的概率明显偏离预设的显著性水平. 表4-表6是 $\theta = 0.8, \mu_E = 2.8, \mu_R = 3, \mu_P = 2$ 时检验犯第一类错误概率的比较, 可以看到此时提出的

偏正态检验方法同样比正态的方法展现出更好的效果. 同时, 也考虑了一组位置参数较大时的情况, 即 $\theta = 0.6, \mu_E = 7.6, \mu_R = 8, \mu_P = 7$ 和 $\theta = 0.8, \mu_E = 7.8, \mu_R = 8, \mu_P = 7$ 时本文中提出的方法与Pigeot所提出方法的一类错误概率的比较, 比较结果在表7-表12中展示. 从表7-表12可以看出在位置参数较大时该检验方法得到一类错误概率同样较正态方法更接近预设的显著性水平, 表明该检验方法较为有效和稳定.

接下来模拟了所提出的偏正态三臂非劣性检验方法的功效. 表13-表24给出了在上述有效控制犯第一类错误概率时的检验功效. 这里为使备择假设的条件成立分别取 $\mu_E = 6, \mu_R = 3, \mu_P = 2$ 和 $\mu_E = 12, \mu_R = 8, \mu_P = 7$. 可以从表13-表15看出 $\theta = 0.6$ 时在三种不同的样本量比例情况下, 随着总样本量的增加, 检验的功效都明显增加. 样本量分配比例为 $n_E : n_R : n_P = 2 : 2 : 1$ 时所得功效优于样本量分配比例为 $n_E : n_R : n_P = 1 : 1 : 1$, 而样本量分配比例为 $n_E : n_R : n_P = 3 : 2 : 1$ 时, 其功效优于 $n_E : n_R : n_P = 2 : 2 : 1$. 试验、对照和安慰剂三组样本不平衡分配时功效在样本量较小时就能达到80%, 而三组样本平衡分配时则需要较大的样本量. 比较不同显著性水平可知, 在相同的样本量总数及分配比例下, 显著性水平从0.05到0.075, 再到0.1检验的功效随之增加. 对应的比较表16-表18, $\theta = 0.8$ 时与 $\theta = 0.6$ 的功效表现的特征一致, 同时表现出比 $\theta = 0.6$ 时需要更大的样本量来达到相同的功效. 从表19-表24可以得到在位置参数设定较大时结论相同.

偏正态三臂非劣性检验方法在不同的参数设置及总样本量分配时达到功效80%时所需的样本量见表25和表26. $\theta = 0.8$ 与 $\theta = 0.6$ 两种情况下都在显著性水平 α 相同时, 总样本量分配比例为 $n_E : n_R : n_P = 1 : 1 : 1$ 的检验达到功效80%时所需的样本量明显多于其他两种比例, 而第三种分配比例所需的样本量最少. 而在同一总样本分配比例下显著性水平为0.1时达到功效80%所需总样本量最少, 显著性水平为0.05时达到功效80%所需总样本量最多.

算法 1: Monte Carlo模拟计算犯第一类错误的概率

输入: 实验组、对照组和安慰剂组的位置, 尺度、偏度参数及样本量 $\mu_i, \sigma_i^2, \lambda_i, n_i, i = E, R, P$, 循环的次數 n_1, n_2 , 预先指定的 θ

输出: 犯第一类错误的概率 P_1

- 1 初始化: $\mu_E - \theta\mu_R - (1 - \theta)\mu_P = 0, p = 0$;
- 2 **for** $j = 1, 2, \dots, n_2$ **do**
- 3 当原假设 H_0 成立时, 生成三组来自于偏正态总体 $SN(\mu_i, \sigma_i^2, \lambda_i)$ 的简单随机样本 $x_{i1}, x_{i2}, \dots, x_{in_i}, i = E, R, P$, 由(5)可得样本的前三阶矩: $\bar{x}_i \leftarrow \frac{1}{n} \sum_{k=1}^{n_i} x_{ik}$, $s_{i2} \leftarrow \frac{1}{n} \sum_{k=1}^{n_i} (x_{ik} - \bar{x}_i)^2$, $s_{i3} \leftarrow \frac{1}{n} \sum_{k=1}^{n_i} (x_{ik} - \bar{x}_i)^3$, 并由定理3.1得到相应的矩估计值: $\hat{\mu}_i \leftarrow \bar{x}_i - cs_{i3}^{1/3}$, $\hat{\sigma}_i^2 \leftarrow s_{i2} + c^2 s_{i3}^{2/3}$, $\hat{\delta}_i \leftarrow cs_{i3}^{1/3}/b \left(s_{i2} + c^2 s_{i3}^{2/3} \right)^{1/2}$, $\hat{\lambda}_i \leftarrow \hat{\delta}_i / \left(1 + \hat{\delta}_i^2 \right)^{1/2}$, 进一步由(13)可得检验统计量: $t_M \leftarrow \frac{\bar{X}_E - \theta\bar{X}_R - (1-\theta)\bar{X}_P - b(\hat{\sigma}_E\hat{\delta}_E - \theta\hat{\sigma}_R\hat{\delta}_R - (1-\theta)\hat{\sigma}_P\hat{\delta}_P)}{\sqrt{\frac{\hat{\sigma}_E^2(1-b^2\hat{\delta}_E^2)}{n_E} + \frac{\theta^2\hat{\sigma}_R^2(1-b^2\hat{\delta}_R^2)}{n_R} + \frac{(1-\theta)^2\hat{\sigma}_P^2(1-b^2\hat{\delta}_P^2)}{n_P}}}$
- 4 **for** $l = 1, 2, \dots, n_1$ **do**
- 5 生成Bootstrap样本 $x_{BMij} \sim SN(\hat{\mu}_i, \hat{\sigma}_i^2, \hat{\lambda}_i), i = E, R, P, j = 1, 2, \dots, n_i$, 由(5)计算Bootstrap样本前三阶矩 $(\bar{x}_{BMi}, s_{BMi2}, s_{BMi3})$, 由定理3.1利用Bootstrap样本 x_{BMij} , 计算矩估计值 $(\hat{\mu}_{BMi}, \hat{\sigma}_{BMi}^2, \hat{\delta}_{BMi})$, 进而由(14)得到Bootstrap检验统计量: $t_{BM} \leftarrow \frac{\bar{X}_{BME} - \theta\bar{X}_{BMR} - (1-\theta)\bar{X}_{BMP} - b(\hat{\sigma}_{BME}\hat{\delta}_{BME} - \theta\hat{\sigma}_{BMR}\hat{\delta}_{BMR} - (1-\theta)\hat{\sigma}_{BMP}\hat{\delta}_{BMP})}{\sqrt{\frac{\hat{\sigma}_{BME}^2(1-b^2\hat{\delta}_{BME}^2)}{n_{BME}} + \frac{\theta^2\hat{\sigma}_{BMR}^2(1-b^2\hat{\delta}_{BMR}^2)}{n_{BMR}} + \frac{(1-\theta)^2\hat{\sigma}_{BMP}^2(1-b^2\hat{\delta}_{BMP}^2)}{n_{BMP}}}}$
- 6 **if** $t_{BM} \geq t_M$ **then**
- 7 | $p \leftarrow p + 1$
- 8 **end**
- 9 **end**
- 10 $p_1 \leftarrow p/n_1$
- 11 **if** $p_1 \leq 0.05$ **then**
- 12 | $q_j \leftarrow 1$
- 13 **else**
- 14 | $q_j \leftarrow 0$
- 15 **end**
- 16 **end**
- 17 $P_1 \leftarrow \left(\frac{1}{n_2} \right) \sum_{j=1}^{n_2} q_j$

表1 $\theta = 0.6, \mu_E = 2.6, \mu_R = 3, \mu_P = 2$ 显著性水平为0.05时两种检验方法犯第一类错误概率的比较

	$\lambda_E, \lambda_R, \lambda_P$	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
		psn	pn	psn	pn	psn	pn	psn	pn
1: 1: 1	420	0.030	0.008	0.042	0.294	0.023	0.015	0.032	0.011
	540	0.033	0.007	0.051	0.333	0.032	0.013	0.046	0.006
	660	0.049	0.003	0.048	0.333	0.051	0.010	0.042	0.005
	780	0.048	0.000	0.048	0.402	0.051	0.002	0.033	0.001
	900	0.050	0.002	0.034	0.463	0.045	0.011	0.042	0.002
2: 2: 1	420	0.029	0.005	0.041	0.337	0.020	0.012	0.043	0.003
	540	0.048	0.003	0.048	0.309	0.052	0.011	0.055	0.008
	660	0.054	0.004	0.042	0.342	0.055	0.008	0.037	0.005
	780	0.050	0.001	0.032	0.425	0.040	0.008	0.031	0.004
	900	0.055	0.000	0.042	0.467	0.051	0.006	0.057	0.003
3: 2: 1	420	0.041	0.005	0.047	0.302	0.030	0.014	0.045	0.006
	540	0.047	0.003	0.035	0.332	0.051	0.012	0.039	0.006
	660	0.033	0.003	0.031	0.412	0.032	0.011	0.034	0.004
	780	0.051	0.003	0.052	0.477	0.041	0.005	0.049	0.002
	900	0.058	0.001	0.046	0.482	0.052	0.003	0.054	0.001

表2 $\theta = 0.6, \mu_E = 2.6, \mu_R = 3, \mu_P = 2$ 显著性水平为0.075时两种检验方法犯第一类错误概率的比较

	$\lambda_E, \lambda_R, \lambda_P$	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
		n	psn	pn	psn	pn	psn	pn	psn
1: 1: 1	420	0.053	0.008	0.068	0.361	0.035	0.018	0.048	0.017
	540	0.053	0.005	0.076	0.394	0.054	0.021	0.070	0.010
	660	0.080	0.007	0.075	0.466	0.062	0.014	0.065	0.011
	780	0.072	0.002	0.081	0.494	0.079	0.013	0.071	0.003
	900	0.078	0.003	0.065	0.560	0.077	0.011	0.072	0.002
2: 2: 1	420	0.045	0.008	0.064	0.361	0.047	0.018	0.066	0.017
	540	0.071	0.005	0.083	0.394	0.063	0.021	0.076	0.010
	660	0.079	0.007	0.062	0.466	0.079	0.014	0.052	0.011
	780	0.069	0.002	0.059	0.494	0.055	0.013	0.057	0.003
	900	0.079	0.003	0.076	0.560	0.081	0.011	0.081	0.002
3: 2: 1	420	0.062	0.007	0.067	0.336	0.054	0.018	0.058	0.011
	540	0.080	0.007	0.076	0.442	0.076	0.017	0.074	0.005
	660	0.067	0.004	0.051	0.468	0.055	0.016	0.053	0.008
	780	0.085	0.002	0.085	0.558	0.077	0.009	0.081	0.004
	900	0.086	0.001	0.085	0.521	0.081	0.009	0.079	0.007

表3 $\theta = 0.6, \mu_E = 2.6, \mu_R = 3, \mu_P = 2$ 显著性水平为0.1时两种检验方法犯第一类错误概率的比较

	$\lambda_E, \lambda_R, \lambda_P$	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
		n	psn	pn	psn	pn	psn	pn	psn
1: 1: 1	420	0.072	0.009	0.083	0.428	0.060	0.027	0.064	0.026
	540	0.069	0.013	0.098	0.464	0.071	0.015	0.077	0.014
	660	0.102	0.006	0.101	0.526	0.082	0.028	0.080	0.007
	780	0.091	0.005	0.098	0.588	0.096	0.011	0.080	0.006
	900	0.096	0.008	0.082	0.636	0.106	0.012	0.093	0.007
2: 2: 1	420	0.056	0.009	0.081	0.428	0.059	0.027	0.082	0.026
	540	0.100	0.013	0.102	0.464	0.068	0.015	0.108	0.014
	660	0.102	0.006	0.090	0.526	0.090	0.028	0.084	0.007
	780	0.097	0.005	0.092	0.588	0.070	0.011	0.088	0.006
	900	0.110	0.008	0.089	0.636	0.092	0.012	0.111	0.007
3: 2: 1	420	0.074	0.013	0.074	0.436	0.069	0.018	0.068	0.019
	540	0.091	0.006	0.091	0.505	0.095	0.019	0.087	0.016
	660	0.083	0.005	0.083	0.518	0.082	0.017	0.077	0.007
	780	0.097	0.004	0.097	0.596	0.090	0.024	0.117	0.015
	900	0.109	0.002	0.109	0.616	0.111	0.015	0.105	0.006

表4 $\theta = 0.8, \mu_E = 2.8, \mu_R = 3, \mu_P = 2$ 显著性水平为0.05时两种检验方法犯第一类错误概率的比较

	$\lambda_E, \lambda_R, \lambda_P$	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
		n	psn	pn	psn	pn	psn	pn	psn
1: 1: 1	420	0.029	0.007	0.034	0.220	0.020	0.006	0.039	0.014
	540	0.026	0.004	0.040	0.249	0.025	0.003	0.049	0.010
	660	0.050	0.004	0.054	0.295	0.048	0.001	0.060	0.013
	780	0.043	0.002	0.049	0.340	0.043	0.000	0.040	0.007
	900	0.047	0.002	0.037	0.388	0.053	0.000	0.040	0.008
2: 2: 1	420	0.035	0.005	0.054	0.225	0.031	0.001	0.033	0.009
	540	0.043	0.004	0.063	0.285	0.042	0.002	0.050	0.013
	660	0.056	0.000	0.051	0.334	0.051	0.000	0.045	0.006
	780	0.054	0.001	0.035	0.373	0.051	0.000	0.041	0.007
	900	0.057	0.001	0.059	0.426	0.049	0.000	0.049	0.005
3: 2: 1	420	0.037	0.006	0.054	0.238	0.035	0.004	0.048	0.014
	540	0.046	0.001	0.044	0.294	0.050	0.000	0.050	0.007
	660	0.032	0.002	0.037	0.352	0.033	0.001	0.041	0.007
	780	0.057	0.000	0.050	0.396	0.053	0.000	0.051	0.004
	900	0.059	0.002	0.041	0.438	0.064	0.002	0.046	0.006

表5 $\theta = 0.8, \mu_E = 2.8, \mu_R = 3, \mu_P = 2$ 显著性水平为0.075时两种检验方法犯第一类错误概率的比较

$\lambda_E, \lambda_R, \lambda_P$	n	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
		<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>
1: 1: 1	420	0.057	0.010	0.068	0.284	0.048	0.007	0.071	0.024
	540	0.053	0.008	0.079	0.329	0.040	0.004	0.078	0.018
	660	0.085	0.007	0.085	0.373	0.085	0.004	0.082	0.019
	780	0.071	0.005	0.072	0.423	0.075	0.002	0.060	0.014
	900	0.085	0.005	0.060	0.464	0.086	0.000	0.079	0.015
2: 2: 1	420	0.048	0.008	0.060	0.284	0.055	0.004	0.055	0.019
	540	0.064	0.010	0.079	0.358	0.068	0.004	0.068	0.017
	660	0.072	0.003	0.062	0.413	0.074	0.000	0.073	0.013
	780	0.081	0.002	0.047	0.465	0.082	0.001	0.057	0.012
	900	0.080	0.002	0.075	0.523	0.078	0.002	0.076	0.010
3: 2: 1	420	0.063	0.008	0.088	0.310	0.058	0.006	0.084	0.020
	540	0.077	0.004	0.088	0.367	0.072	0.002	0.074	0.015
	660	0.066	0.002	0.072	0.429	0.056	0.002	0.068	0.011
	780	0.089	0.001	0.077	0.485	0.086	0.000	0.082	0.010
	900	0.084	0.004	0.074	0.516	0.072	0.002	0.078	0.011

表6 $\theta = 0.8, \mu_E = 2.8, \mu_R = 3, \mu_P = 2$ 显著性水平为0.1时两种检验方法犯第一类错误概率的比较

$\lambda_E, \lambda_R, \lambda_P$	n	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
		<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>
1: 1: 1	420	0.067	0.013	0.098	0.341	0.062	0.009	0.080	0.028
	540	0.064	0.016	0.102	0.377	0.069	0.004	0.105	0.026
	660	0.099	0.009	0.103	0.427	0.109	0.007	0.106	0.028
	780	0.103	0.009	0.107	0.488	0.110	0.003	0.087	0.017
	900	0.094	0.008	0.082	0.524	0.108	0.004	0.096	0.016
2: 2: 1	420	0.068	0.012	0.086	0.345	0.069	0.008	0.095	0.029
	540	0.096	0.012	0.115	0.418	0.085	0.009	0.105	0.026
	660	0.107	0.006	0.094	0.486	0.103	0.002	0.097	0.020
	780	0.121	0.006	0.082	0.533	0.106	0.001	0.078	0.020
	900	0.106	0.002	0.113	0.576	0.111	0.002	0.105	0.016
3: 2: 1	420	0.075	0.015	0.105	0.376	0.073	0.008	0.103	0.028
	540	0.092	0.007	0.101	0.430	0.088	0.003	0.102	0.024
	660	0.085	0.006	0.080	0.496	0.078	0.002	0.098	0.018
	780	0.108	0.003	0.098	0.538	0.095	0.001	0.102	0.021
	900	0.114	0.005	0.087	0.578	0.107	0.003	0.105	0.018

表7 $\theta = 0.6, \mu_E = 7.6, \mu_R = 8, \mu_P = 7$ 显著性水平为0.05时两种检验方法犯第一类错误概率的比较

$\lambda_E, \lambda_R, \lambda_P$	n	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
		<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>
1: 1: 1	420	0.035	0.008	0.050	0.294	0.026	0.015	0.042	0.011
	540	0.034	0.007	0.059	0.333	0.040	0.013	0.051	0.006
	660	0.056	0.003	0.050	0.333	0.059	0.010	0.054	0.005
	780	0.056	0.000	0.061	0.402	0.059	0.002	0.050	0.001
	900	0.041	0.002	0.041	0.463	0.056	0.011	0.057	0.002
2: 2: 1	420	0.030	0.005	0.039	0.337	0.036	0.012	0.046	0.003
	540	0.051	0.003	0.066	0.309	0.044	0.011	0.058	0.008
	660	0.055	0.004	0.051	0.342	0.054	0.008	0.036	0.005
	780	0.050	0.001	0.041	0.425	0.044	0.008	0.041	0.004
	900	0.057	0.000	0.056	0.467	0.053	0.006	0.067	0.003
3: 2: 1	420	0.042	0.005	0.054	0.302	0.036	0.014	0.050	0.006
	540	0.044	0.003	0.044	0.332	0.056	0.012	0.043	0.006
	660	0.031	0.003	0.040	0.412	0.043	0.011	0.037	0.004
	780	0.056	0.003	0.062	0.477	0.054	0.005	0.052	0.002
	900	0.058	0.001	0.053	0.482	0.066	0.003	0.058	0.001

表8 $\theta = 0.6, \mu_E = 7.6, \mu_R = 8, \mu_P = 7$ 显著性水平为0.075时两种检验方法犯第一类错误概率的比较

$\lambda_E, \lambda_R, \lambda_P$	n	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
		<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>
1: 1: 1	420	0.053	0.008	0.070	0.361	0.035	0.018	0.054	0.017
	540	0.053	0.005	0.077	0.394	0.054	0.021	0.087	0.010
	660	0.080	0.007	0.079	0.466	0.062	0.014	0.076	0.011
	780	0.072	0.002	0.083	0.494	0.079	0.013	0.063	0.003
	900	0.078	0.003	0.066	0.560	0.077	0.011	0.075	0.002
2: 2: 1	420	0.053	0.008	0.076	0.361	0.051	0.018	0.067	0.017
	540	0.071	0.005	0.087	0.394	0.067	0.021	0.078	0.010
	660	0.075	0.007	0.071	0.466	0.081	0.014	0.063	0.011
	780	0.075	0.002	0.063	0.494	0.061	0.013	0.071	0.003
	900	0.082	0.003	0.089	0.560	0.090	0.011	0.087	0.002
3: 2: 1	420	0.064	0.007	0.073	0.336	0.058	0.018	0.068	0.011
	540	0.080	0.007	0.079	0.442	0.085	0.017	0.076	0.005
	660	0.067	0.004	0.053	0.468	0.072	0.016	0.061	0.008
	780	0.089	0.002	0.086	0.558	0.084	0.009	0.085	0.004
	900	0.085	0.001	0.087	0.521	0.088	0.009	0.085	0.007

表9 $\theta = 0.6, \mu_E = 7.6, \mu_R = 8, \mu_P = 7$ 显著性水平为0.1时两种检验方法犯第一类错误概率的比较

$\lambda_E, \lambda_R, \lambda_P$	n	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
		<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>
1: 1: 1	420	0.086	0.009	0.095	0.428	0.066	0.027	0.102	0.026
	540	0.075	0.013	0.099	0.464	0.070	0.015	0.109	0.014
	660	0.111	0.006	0.108	0.526	0.098	0.028	0.104	0.007
	780	0.107	0.005	0.103	0.588	0.098	0.011	0.091	0.006
	900	0.108	0.008	0.082	0.636	0.116	0.012	0.114	0.007
2: 2: 1	420	0.071	0.009	0.091	0.428	0.066	0.027	0.087	0.026
	540	0.091	0.013	0.110	0.464	0.083	0.015	0.110	0.014
	660	0.106	0.006	0.083	0.526	0.093	0.028	0.085	0.007
	780	0.110	0.005	0.088	0.588	0.084	0.011	0.088	0.006
	900	0.116	0.008	0.110	0.636	0.114	0.012	0.110	0.007
3: 2: 1	420	0.074	0.013	0.094	0.436	0.072	0.018	0.071	0.019
	540	0.091	0.006	0.112	0.505	0.096	0.019	0.087	0.016
	660	0.083	0.005	0.075	0.518	0.082	0.017	0.081	0.007
	780	0.097	0.004	0.106	0.596	0.092	0.024	0.115	0.015
	900	0.109	0.002	0.100	0.616	0.111	0.015	0.109	0.006

表10 $\theta = 0.8, \mu_E = 7.8, \mu_R = 8, \mu_P = 7$ 显著性水平为0.05时两种检验方法犯第一类错误概率的比较

$\lambda_E, \lambda_R, \lambda_P$	n	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
		<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>	<i>psn</i>	<i>pn</i>
1: 1: 1	420	0.029	0.008	0.032	0.234	0.020	0.005	0.021	0.020
	540	0.026	0.002	0.040	0.255	0.025	0.004	0.026	0.017
	660	0.050	0.006	0.054	0.302	0.048	0.002	0.050	0.003
	780	0.043	0.003	0.049	0.318	0.043	0.002	0.041	0.012
	900	0.047	0.002	0.037	0.359	0.053	0.002	0.057	0.006
2: 2: 1	420	0.032	0.007	0.045	0.276	0.031	0.004	0.041	0.010
	540	0.043	0.003	0.062	0.266	0.042	0.002	0.058	0.015
	660	0.060	0.004	0.049	0.289	0.051	0.003	0.063	0.009
	780	0.050	0.002	0.033	0.358	0.051	0.002	0.052	0.008
	900	0.049	0.001	0.053	0.396	0.049	0.002	0.060	0.007
3: 2: 1	420	0.041	0.008	0.063	0.244	0.038	0.001	0.051	0.011
	540	0.048	0.004	0.058	0.297	0.051	0.004	0.054	0.006
	660	0.031	0.005	0.043	0.320	0.033	0.000	0.044	0.009
	780	0.059	0.003	0.054	0.383	0.060	0.001	0.050	0.011
	900	0.053	0.001	0.047	0.399	0.055	0.002	0.055	0.004

表11 $\theta = 0.8, \mu_E = 7.8, \mu_R = 8, \mu_P = 7$ 显著性水平为0.075时两种检验方法犯第一类错误概率的比较

$\lambda_E, \lambda_R, \lambda_P$		0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
n		psn	pn	psn	pn	psn	pn	psn	pn
1: 1: 1	420	0.060	0.011	0.068	0.287	0.045	0.006	0.071	0.021
	540	0.050	0.006	0.079	0.344	0.043	0.005	0.080	0.028
	660	0.089	0.007	0.085	0.343	0.079	0.001	0.088	0.018
	780	0.072	0.006	0.072	0.424	0.079	0.003	0.068	0.010
	900	0.089	0.002	0.060	0.439	0.086	0.002	0.083	0.021
2: 2: 1	420	0.045	0.014	0.074	0.317	0.061	0.004	0.066	0.022
	540	0.068	0.006	0.087	0.340	0.071	0.007	0.080	0.023
	660	0.079	0.006	0.075	0.396	0.078	0.005	0.082	0.017
	780	0.073	0.002	0.048	0.432	0.086	0.002	0.061	0.013
	900	0.073	0.003	0.083	0.498	0.084	0.004	0.082	0.009
3: 2: 1	420	0.060	0.008	0.088	0.308	0.060	0.007	0.086	0.022
	540	0.077	0.008	0.074	0.382	0.076	0.006	0.068	0.020
	660	0.062	0.006	0.061	0.401	0.057	0.004	0.062	0.015
	780	0.089	0.002	0.084	0.469	0.090	0.002	0.085	0.008
	900	0.084	0.002	0.072	0.453	0.076	0.001	0.082	0.018

表12 $\theta = 0.8, \mu_E = 7.8, \mu_R = 8, \mu_P = 7$ 显著性水平为0.1时两种检验方法犯第一类错误概率的比较

$\lambda_E, \lambda_R, \lambda_P$		0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
n		psn	pn	psn	pn	psn	pn	psn	pn
1: 1: 1	420	0.079	0.025	0.100	0.372	0.081	0.014	0.096	0.040
	540	0.070	0.019	0.107	0.399	0.079	0.012	0.107	0.034
	660	0.110	0.013	0.112	0.410	0.110	0.005	0.100	0.030
	780	0.109	0.013	0.111	0.460	0.109	0.002	0.093	0.013
	900	0.116	0.008	0.081	0.509	0.114	0.009	0.093	0.017
2: 2: 1	420	0.065	0.012	0.080	0.373	0.067	0.007	0.098	0.042
	540	0.094	0.014	0.111	0.420	0.083	0.008	0.111	0.031
	660	0.105	0.010	0.090	0.453	0.100	0.008	0.104	0.024
	780	0.115	0.006	0.072	0.500	0.105	0.000	0.083	0.018
	900	0.100	0.009	0.110	0.542	0.108	0.001	0.109	0.015
3: 2: 1	420	0.081	0.014	0.115	0.390	0.084	0.007	0.100	0.031
	540	0.098	0.012	0.110	0.414	0.096	0.007	0.102	0.032
	660	0.086	0.012	0.083	0.483	0.082	0.007	0.098	0.028
	780	0.111	0.008	0.102	0.558	0.104	0.000	0.102	0.021
	900	0.109	0.004	0.091	0.551	0.114	0.003	0.102	0.019

表13 $\theta = 0.6, \mu_E = 6, \mu_R = 3, \mu_P = 2$ 显著性水平为0.05时偏正态检验方法的检验功效

n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4	
1: 1: 1	420	0.751	0.767	0.746	0.781
	540	0.775	0.794	0.799	0.810
	660	0.806	0.835	0.820	0.842
	780	0.822	0.858	0.836	0.859
	900	0.841	0.870	0.853	0.879
2: 2: 1	420	0.783	0.820	0.794	0.830
	540	0.818	0.841	0.824	0.857
	660	0.828	0.857	0.842	0.884
	780	0.866	0.889	0.862	0.897
	900	0.868	0.909	0.883	0.925
3: 2: 1	420	0.822	0.856	0.835	0.866
	540	0.845	0.873	0.847	0.886
	660	0.847	0.897	0.859	0.883
	780	0.881	0.915	0.880	0.924
	900	0.877	0.931	0.895	0.943

表14 $\theta = 0.6, \mu_E = 6, \mu_R = 3, \mu_P = 2$ 显著性水平为0.075时偏正态检验方法的检验功效

	n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4
1: 1: 1	420	0.787	0.825	0.784	0.824
	540	0.809	0.846	0.832	0.849
	660	0.843	0.874	0.855	0.877
	780	0.867	0.900	0.880	0.887
	900	0.875	0.909	0.884	0.910
2: 2: 1	420	0.825	0.858	0.830	0.857
	540	0.848	0.874	0.856	0.889
	660	0.862	0.894	0.876	0.908
	780	0.901	0.923	0.902	0.930
	900	0.913	0.944	0.913	0.945
3: 2: 1	420	0.866	0.891	0.869	0.890
	540	0.880	0.915	0.886	0.922
	660	0.883	0.926	0.905	0.921
	780	0.915	0.948	0.918	0.948
	900	0.916	0.955	0.927	0.958

表15 $\theta = 0.6, \mu_E = 6, \mu_R = 3, \mu_P = 2$ 显著性水平为0.1时偏正态检验方法的检验功效

	n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4
1: 1: 1	420	0.822	0.864	0.820	0.864
	540	0.854	0.882	0.870	0.883
	660	0.878	0.904	0.887	0.909
	780	0.896	0.927	0.905	0.921
	900	0.905	0.931	0.913	0.938
2: 2: 1	420	0.856	0.888	0.866	0.896
	540	0.878	0.906	0.884	0.920
	660	0.888	0.923	0.911	0.929
	780	0.928	0.951	0.930	0.955
	900	0.939	0.965	0.937	0.970
3: 2: 1	420	0.903	0.920	0.892	0.916
	540	0.913	0.941	0.915	0.940
	660	0.926	0.963	0.931	0.958
	780	0.941	0.965	0.941	0.968
	900	0.942	0.978	0.948	0.970

表16 $\theta = 0.8, \mu_E = 6, \mu_R = 3, \mu_P = 2$ 显著性水平为0.05时偏正态检验方法的检验功效

	n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4
1: 1: 1	420	0.705	0.715	0.702	0.750
	540	0.722	0.756	0.738	0.780
	660	0.748	0.772	0.763	0.803
	780	0.776	0.795	0.781	0.812
	900	0.779	0.805	0.796	0.831
2: 2: 1	420	0.740	0.767	0.743	0.786
	540	0.778	0.802	0.784	0.826
	660	0.790	0.815	0.809	0.839
	780	0.824	0.862	0.832	0.867
	900	0.838	0.880	0.850	0.899
3: 2: 1	420	0.772	0.812	0.786	0.832
	540	0.797	0.848	0.815	0.855
	660	0.804	0.852	0.812	0.859
	780	0.845	0.874	0.845	0.900
	900	0.843	0.890	0.850	0.920

表17 $\theta = 0.8, \mu_E = 6, \mu_R = 3, \mu_P = 2$ 显著性水平为0.075时偏正态检验方法的检验功效

	n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4
1: 1: 1	420	0.741	0.761	0.739	0.791
	540	0.761	0.800	0.796	0.815
	660	0.792	0.833	0.810	0.849
	780	0.817	0.854	0.828	0.844
	900	0.830	0.859	0.844	0.862
2: 2: 1	420	0.787	0.818	0.797	0.839
	540	0.814	0.842	0.817	0.864
	660	0.832	0.860	0.850	0.877
	780	0.868	0.889	0.876	0.901
	900	0.878	0.918	0.886	0.927
3: 2: 1	420	0.819	0.857	0.826	0.871
	540	0.841	0.889	0.857	0.892
	660	0.850	0.891	0.871	0.903
	780	0.878	0.914	0.884	0.930
	900	0.887	0.925	0.891	0.941

表18 $\theta = 0.8, \mu_E = 6, \mu_R = 3, \mu_P = 2$ 显著性水平为0.1时偏正态检验方法的检验功效

	n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4
1: 1: 1	420	0.785	0.800	0.771	0.816
	540	0.805	0.836	0.834	0.844
	660	0.839	0.877	0.849	0.882
	780	0.856	0.893	0.864	0.886
	900	0.864	0.905	0.884	0.907
2: 2: 1	420	0.833	0.854	0.842	0.875
	540	0.848	0.879	0.855	0.897
	660	0.868	0.900	0.877	0.918
	780	0.895	0.926	0.901	0.940
	900	0.910	0.947	0.910	0.948
3: 2: 1	420	0.858	0.886	0.862	0.895
	540	0.880	0.917	0.885	0.931
	660	0.889	0.930	0.902	0.933
	780	0.913	0.951	0.910	0.951
	900	0.915	0.952	0.930	0.956

表19 $\theta = 0.6, \mu_E = 12, \mu_R = 8, \mu_P = 7$ 显著性水平为0.05时偏正态检验方法的检验功效

	n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4
1: 1: 1	420	0.887	0.923	0.890	0.937
	540	0.933	0.962	0.946	0.962
	660	0.948	0.978	0.957	0.983
	780	0.967	0.991	0.977	0.991
	900	0.980	0.993	0.986	0.995
2: 2: 1	420	0.944	0.971	0.959	0.972
	540	0.964	0.987	0.961	0.989
	660	0.980	0.995	0.984	0.998
	780	0.991	0.999	0.996	0.998
	900	0.995	0.999	0.996	0.999
3: 2: 1	420	0.982	0.988	0.979	0.989
	540	0.981	0.995	0.988	0.998
	660	0.988	0.999	0.989	0.999
	780	0.996	1.000	0.997	1.000
	900	0.999	1.000	0.998	1.000

表20 $\theta = 0.6, \mu_E = 12, \mu_R = 8, \mu_P = 7$ 显著性水平为0.075时偏正态检验方法的检验功效

	n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4
1: 1: 1	420	0.926	0.963	0.931	0.965
	540	0.968	0.977	0.975	0.984
	660	0.974	0.994	0.980	0.993
	780	0.990	0.999	0.995	0.998
	900	0.992	0.996	0.995	0.996
2: 2: 1	420	0.976	0.983	0.978	0.986
	540	0.978	0.993	0.985	0.995
	660	0.991	1.000	0.991	1.000
	780	0.996	1.000	0.999	1.000
	900	0.999	1.000	0.999	1.000
3: 2: 1	420	0.989	0.994	0.989	0.994
	540	0.993	0.998	0.993	0.999
	660	0.997	1.000	0.998	1.000
	780	1.000	1.000	1.000	1.000
	900	1.000	1.000	1.000	1.000

表21 $\theta = 0.6, \mu_E = 12, \mu_R = 8, \mu_P = 7$ 显著性水平为0.1时偏正态检验方法的检验功效

	n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4
1: 1: 1	420	0.946	0.984	0.967	0.986
	540	0.985	0.992	0.984	0.995
	660	0.989	0.998	0.990	0.997
	780	0.997	0.999	0.999	1.000
	900	0.996	1.000	0.999	0.999
2: 2: 1	420	0.988	0.992	0.987	0.998
	540	0.986	0.997	0.991	0.998
	660	0.996	1.000	0.997	1.000
	780	0.999	1.000	0.999	1.000
	900	1.000	1.000	1.000	1.000
3: 2: 1	420	0.993	0.999	0.993	0.997
	540	0.997	0.999	0.995	0.999
	660	1.000	1.000	1.000	1.000
	780	1.000	1.000	1.000	1.000
	900	1.000	1.000	1.000	1.000

表22 $\theta = 0.8, \mu_E = 12, \mu_R = 8, \mu_P = 7$ 显著性水平为0.05时偏正态检验方法的检验功效

	n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4
1: 1: 1	420	0.847	0.865	0.842	0.887
	540	0.892	0.925	0.907	0.940
	660	0.915	0.948	0.927	0.962
	780	0.937	0.968	0.940	0.979
	900	0.944	0.976	0.946	0.985
2: 2: 1	420	0.924	0.948	0.933	0.960
	540	0.939	0.974	0.946	0.981
	660	0.958	0.982	0.966	0.992
	780	0.979	0.997	0.985	0.998
	900	0.990	0.995	0.990	0.997
3: 2: 1	420	0.952	0.972	0.959	0.976
	540	0.966	0.987	0.966	0.990
	660	0.976	0.993	0.976	0.996
	780	0.989	0.997	0.988	0.999
	900	0.995	0.998	0.995	0.999

表23 $\theta = 0.8, \mu_E = 12, \mu_R = 8, \mu_P = 7$ 显著性水平为0.075时偏正态检验方法的检验功效

	n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4
1: 1: 1	420	0.883	0.928	0.884	0.928
	540	0.943	0.957	0.948	0.969
	660	0.951	0.981	0.959	0.980
	780	0.963	0.986	0.973	0.992
	900	0.968	0.991	0.972	0.993
2: 2: 1	420	0.954	0.971	0.957	0.977
	540	0.963	0.986	0.971	0.988
	660	0.978	0.996	0.982	0.997
	780	0.989	0.999	0.992	0.999
	900	0.995	0.999	0.997	0.999
3: 2: 1	420	0.982	0.985	0.980	0.991
	540	0.981	0.994	0.983	0.997
	660	0.985	1.000	0.989	1.000
	780	0.997	1.000	0.996	1.000
	900	0.998	1.000	0.999	1.000

表24 $\theta = 0.8, \mu_E = 12, \mu_R = 8, \mu_P = 7$ 显著性水平为0.1时偏正态检验方法的检验功效

	n	0.3, 0.4, 0.5	-0.1, -0.2, -0.3	0.2, 0.4, 0.1	-0.6, -0.5, -0.4
1: 1: 1	420	0.919	0.955	0.922	0.960
	540	0.969	0.981	0.968	0.984
	660	0.976	0.991	0.977	0.991
	780	0.991	0.996	0.989	0.995
	900	0.985	0.996	0.991	0.996
2: 2: 1	420	0.974	0.984	0.982	0.992
	540	0.976	0.993	0.981	0.996
	660	0.990	0.998	0.992	1.000
	780	0.997	0.999	0.998	0.999
	900	0.998	1.000	0.998	1.000
3: 2: 1	420	0.991	0.993	0.989	0.997
	540	0.991	0.997	0.994	0.999
	660	0.994	1.000	0.998	1.000
	780	0.999	1.000	0.999	1.000
	900	0.999	1.000	0.999	1.000

表25 $\mu_E = 6, \mu_R = 3, \mu_P = 2$ 时偏正态检验方法的功效达到80%时所需的样本量

	α	θ	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
			n	power	n	power	n	power	n	power
1: 1: 1	0.05	0.6	660	0.806	570	0.830	570	0.811	540	0.810
		0.8	1020	0.806	900	0.805	960	0.810	660	0.803
	0.075	0.6	480	0.808	330	0.803	450	0.802	300	0.808
		0.8	780	0.817	540	0.800	660	0.810	450	0.813
	0.1	0.6	330	0.806	270	0.816	300	0.807	270	0.815
		0.8	540	0.805	420	0.800	450	0.800	360	0.814
2: 2: 1	0.05	0.6	480	0.809	390	0.802	450	0.811	300	0.802
		0.8	600	0.812	480	0.820	600	0.813	450	0.812
	0.075	0.6	360	0.808	240	0.803	330	0.811	240	0.811
		0.8	480	0.807	420	0.817	450	0.801	330	0.804
	0.1	0.6	240	0.806	240	0.840	240	0.820	180	0.802
		0.8	390	0.833	240	0.801	330	0.811	240	0.815
3: 2: 1	0.05	0.6	420	0.822	270	0.805	420	0.835	270	0.816
		0.8	660	0.804	420	0.812	540	0.815	360	0.808
	0.075	0.6	330	0.826	210	0.802	270	0.808	180	0.803
		0.8	390	0.801	330	0.808	390	0.807	240	0.801
	0.1	0.6	240	0.808	150	0.813	180	0.801	150	0.808
		0.8	330	0.810	210	0.811	300	0.810	180	0.802

表26 $\mu_E = 12, \mu_R = 8, \mu_P = 7$ 时偏正态检验方法的功效达到80%时所需的样本量

	α	θ	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4	
			n	power	n	power	n	power	n	power
1: 1: 1	0.05	0.6	210	0.808	150	0.822	180	0.805	150	0.820
		0.8	270	0.814	270	0.821	270	0.801	210	0.808
	0.075	0.6	120	0.804	120	0.823	120	0.809	120	0.828
		0.8	210	0.814	150	0.811	150	0.805	150	0.812
	0.1	0.6	120	0.835	90	0.817	120	0.84	90	0.818
		0.8	150	0.829	120	0.811	150	0.839	120	0.826
2: 2: 1	0.05	0.6	150	0.825	120	0.817	150	0.837	120	0.814
		0.8	180	0.815	150	0.820	180	0.816	150	0.830
	0.075	0.6	120	0.836	90	0.830	90	0.804	90	0.830
		0.8	150	0.828	120	0.832	120	0.801	120	0.831
	0.1	0.6	90	0.826	60	0.804	90	0.846	60	0.808
		0.8	120	0.839	90	0.821	120	0.841	90	0.837

表26(续) $\mu_E = 12, \mu_R = 8, \mu_P = 7$ 时偏正态检验方法的功效达到80%时所需的样本量

α	θ	0.3, 0.4, 0.5		-0.1, -0.2, -0.3		0.2, 0.4, 0.1		-0.6, -0.5, -0.4		
		n	power	n	power	n	power	n	power	
3: 2: 1	0.05	0.6	120	0.836	90	0.804	120	0.834	120	0.867
		0.8	180	0.832	120	0.817	120	0.803	120	0.825
	0.075	0.6	90	0.821	60	0.817	90	0.836	60	0.806
		0.8	120	0.834	90	0.806	120	0.838	90	0.811
	0.1	0.6	60	0.82	60	0.844	60	0.811	60	0.838
		0.8	90	0.819	60	0.8	90	0.83	60	0.81

§5 实例分析

本节将本文所提出的方法应用于HIV的非劣性试验, 以患者细胞表面的CD4抗原作为判断治疗效果的指标, 一般CD4计数越高免疫系统越健康, 这里比较三种药物组患者的logCD4. 数据是第7到第8周三组患者的测量结果, 数据集由总样本量 $n = 501$ 名患者组成, 每组患者数 $n_E = n_R = n_P = 167$, 以zA400d作为新药(试验组xExp), zA400dA400n作为标准药(对照组xRef), zA225z作为安慰剂(安慰剂组xPla). 数据描述如下: 三组数据的总体参数值未知, 从图1带有概率密度曲线的频率直方图中可以看出三组数据的偏态特征不太明显, 表27中给出三组数据的特征, 三组数据均值的大小相当相似, 三组数据的方差也比较相近, 由偏度公式得到的偏度数据显示三组数据都是负偏的. 为了检验数据的偏正态性, 参考[29-30]先分别采用Shapiro-Wilk检验、Kolmogorov-Smirnov检验和Cramer-von Mises检验, 实验组的检验 p 值依次为0.000391639, 0.000791679, 0.003596725, 可以得到在5%显著性水平时实验组不是正态分布, 进一步, 使用卡方拟合优度检验原假设为数据服从偏正态分布. 检验统计量的值 $\chi^2 = 6.9453 < \chi^2_{0.05,3} = 7.815$, 所以原假设在5%显著性水平下不被拒绝. 因此, 实验组的分布可以被认为是近似偏正态的. 类似的得到对照组的正态检验 p 值依次为0.005087814, 0.009443264, 0.01567969, 安慰剂的正态检验 p 值依次为0.000335353, 0.001029581, 0.004230334, 由此对照组和安慰剂都不是正态分布, 对照组的检验统计量的值 $\chi^2 = 2.0433 < \chi^2_{0.05,3} = 7.815$, 安慰剂的检验统计量的值 $\chi^2 = 2.8386 < \chi^2_{0.05,3} = 7.815$, 对照组和安慰剂都可被认为是近似偏正态的.

表27 HIV患者logCD4计数结果数据描述

	xExp	xRef	xPla
均值	3.14348089	3.193709022	2.934908505
方差	1.459695401	1.496688256	1.209267686
偏度	-0.513371565	-0.492869016	-0.636804521

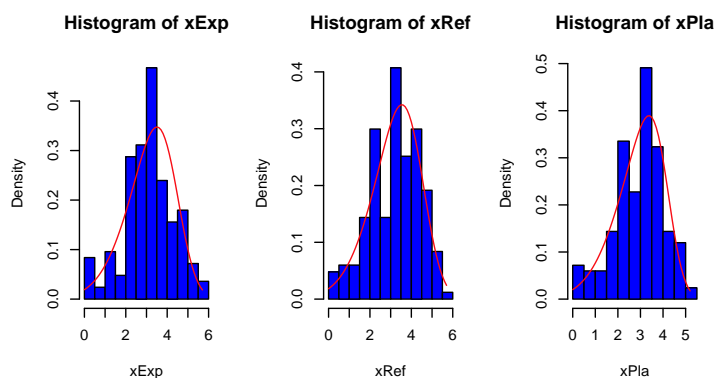


图1 带有概率密度曲线的频率直方图

在该实例数据下采用检验方法一循环1000次研究假设检验问题

$$H_0: \mu_E - \theta\mu_R - (1 - \theta)\mu_P \leq 0 \quad \text{vs} \quad H_1: \mu_E - \theta\mu_R - (1 - \theta)\mu_P > 0. \quad (16)$$

得到检验的 p 值为0.5333, 该值大于预先指定的显著性水平0.05, 故接受原假设, 认为新药zA400d相较于标准药zA400dA400n的疗效不满足非劣性.

§6 总结与展望

本文参考了文献[19], 运用参数矩估计和Bootstrap方法, 提出了偏正态数据下三臂非劣(效)性检验的方法. 通过矩估计得到三臂非劣(效)性试验的三组数据的位置、尺度和偏度的估计值, 并构造了总体位置参数的近似检验统计量, 进一步基于矩估计构造Bootstrap检验统计量. 在尺度参数和偏度参数未知的情况下, 模拟了两种不同方法的检验过程. Monte Carlo模拟结果表明, 基于矩估计的Bootstrap检验方法在多数参数设置下优于Pigeot提出的假设数据正态的检验方法, 并且验证了实验组、阳性对照组和安慰剂组在3: 2: 1的患者数量分配时所需的总人数最少. 将上述方法应用于HIV的非劣性试验实例分析, 以验证本文所提出方法的合理性和有效性. 综上所述对于三臂非劣性实验, 在数据服从偏正态分布时, 本文提出的Bootstrap检验方法比传统的正态检验方法更加有效.

此外, 本文旨在研究偏正态数据的三臂非劣性检验过程, 偏正态数据的三臂非劣性检验的置信区间并未考虑, 位置、尺度和偏度都未知时偏正态数据的三臂非劣性检验的置信区间有待研究.

参考文献:

- [1] Ghosh S, Ghosh S, Tiwari R C. Bayesian approach for assessing non-inferiority in a three-arm trial with pre-specified margin[J]. *Statistics in medicine*, 2016, 35(5): 695-708.
- [2] 许金芳. 包含安慰剂的三臂非劣效试验的设计及其两阶段样本量调整的研究[D]. 西安: 第四军医大学, 2009.
- [3] Pigeot I, Schäfer J, Röhmel J, et al. Assessing non-inferiority of a new treatment in a three-arm clinical trial including a placebo[J]. *Statistics in Medicine*, 2003, 22(6): 883-899.

-
- [4] Kieser M, Friede T. Planning and analysis of three-arm non-inferiority trials with binary endpoints[J]. *Statistics in medicine*, 2007, 26(2): 253-273.
- [5] Hasler M, Vonk R, Hothorn L A. Assessing non-inferiority of a new treatment in a three-arm trial in the presence of heteroscedasticity[J]. *Statistics in Medicine*, 2008, 27(4): 490-503.
- [6] Lee J S, Kim D J. Non-inferiority test in a two-arm trial and a three-arm trial including a placebo[J]. *The Korean Journal of Applied Statistics*, 2008, 21(6): 947-957.
- [7] Li Gang, Gao Shan. A group sequential type design for three-arm non-inferiority trials with binary endpoints[J]. *Biometrical Journal*, 2010, 52(4): 504-518.
- [8] Tang Niansheng, Yu Bin, Tang Manlai. Testing non-inferiority of a new treatment in three-arm clinical trials with binary endpoints[J]. *BMC Medical Research Methodology*, 2014, 14(1): 1-9.
- [9] Ghosh S, Paul E, Chowdhury S, et al. New approaches for testing non-inferiority for three-arm trials with Poisson distributed outcomes[J]. *Biostatistics*, 2022, 23(1): 136-156.
- [10] Tang Niansheng, Liang Fan. Confidence intervals for assessing non-inferiority with assay sensitivity in a three-arm trial with normally distributed endpoints[J]. *Mathematics*, 2022, 10(2): 167.
- [11] Tang Niansheng, Liang Fan, Jiang Depeng. Semiparametric Bayesian approach to assess non-inferiority with assay sensitivity in a three-arm trial with normally distributed endpoints[J]. *Computational Statistics*, 2023: 1-25.
- [12] Azzalini A. A class of distributions which includes the normal ones[J]. *Scandinavian Journal of Statistics*, 1985, 12(2): 171-178.
- [13] 马婷, 吴刘仓, 黄丽. 基于偏正态分布联合位置、尺度与偏度模型的极大似然估计[J]. *数理统计与管理*, 2013, 32(3): 433-439.
- [14] Thiuthad P, Pal N. Hypothesis testing on the location parameter of a skew-normal distribution (SND) with application[C]. *ITM Web of Conferences, EDP Sciences*, 2018, 20: 03003.
- [15] Thiuthad P, Pal N. Point estimation of the location parameter of a skew-normal distribution: some fixed sample and asymptotic results[J]. *Journal of Statistical Theory and Practice*, 2019, 13: 1-27.
- [16] Wu Liucang, Nie Xingfeng, Zheng Guifen. Statistical diagnostics for joint location, scale and skewness models with skew-normal data[J]. *Chinese Journal of Engineering Mathematics*, 2021, 38(2): 180-194.
- [17] Morán-Vásquez R A, Giraldo-Melo A D, Mazo-Lopera M A. Quantile estimation using the logskew-normal linear regression model with application to children's weight data[J]. *Mathematics*, 2023, 11(17): 3736.
- [18] 叶仁道, 王仲池, 罗堃, 等. 多个偏正态总体共同位置参数的Bootstrap置信区间[J]. *数学物理学报*, 2021, 41(1): 194-216.
- [19] 方冰妮. 偏正态总体位置参数的Bootstrap推断[D]. 杭州: 杭州电子科技大学, 2021.
- [20] 余彬. 二元响应变量的三臂非劣性试验[D]. 昆明: 云南大学, 2015.
- [21] Ghosh P, Nathoo F, Gönen M, et al. Assessing noninferiority in a three-arm trial using the Bayesian approach[J]. *Statistics in Medicine*, 2011, 30(15): 1795-1808.
- [22] Ye Rendao, Fang Bingni, Du Weixiao, et al. Bootstraptests for the location parameter under the skew-normal population with unknown scale parameter and skewness parameter[J]. *Mathematics*, 2022, 10(6): 921.
- [23] 刘堂勇. 假设检验的两类错误与Monte Carlo模拟[J]. *中国医院统计*, 2013(6): 416-418.
- [24] Pewsey A. Problems of inference for Azzalini's skewnormal distribution[J]. *Journal of Applied Statistics*, 2000, 27(7): 859-870.

- [25] Ghosh S, Tiwari R C, Ghosh S. Bayesian approach for assessing noninferiority in a three-arm trial with binary endpoint[J]. *Pharmaceutical Statistics*, 2018, 17(4): 342-357.
- [26] Koch A, Röhm J. Hypothesis testing in the “gold standard” design for proving the efficacy of an experimental treatment relative to placebo and a reference[J]. *Journal of Biopharmaceutical Statistics*, 2004, 14(2): 315-325.
- [27] Kwong K S, Cheung S H, Hayter A J, et al. Extension of three-arm non-inferiority studies to trials with multiple new treatments[J]. *Statistics in Medicine*, 2012, 31(24): 2833-2843.
- [28] Ghosh S, Paul E, Chowdhury S, et al. New approaches for testing non-inferiority for three-arm trials with Poisson distributed outcomes[J]. *Biostatistics*, 2022, 23(1): 136-156.
- [29] Ye Rendao, An Na, Luo Kun, et al. Bootstrap inference of the skew-normal two-way classification random effects model with interaction[J]. *Applied Mathematics-A Journal of Chinese Universities Series B*, 2022, 37(3): 435-452.
- [30] Ye Rendao, Xu Lijun, Luo Kun, et al. A parametric bootstrap approach for one-way classification model with skew-normal random effects[J]. *Applied Mathematics-A Journal of Chinese Universities Series B*, 2019, 34(4): 423-435.
- [31] 叶仁道, 姜玲. Panel数据模型的参数bootstrap推断[J]. *高校应用数学学报*, 2018, 33(4): 379-386.
- [32] 叶仁道, 安娜. 偏正态混合效应模型中方差分量函数的参数Bootstrap推断[J]. *应用数学*, 2023, 36(2): 353-363.
- [33] 叶仁道, 杜微晓. 偏正态非平衡面板单因素随机效应模型的Bootstrap推断[J]. *系统科学与数学*, 2023, 43(1): 244-270.

Three-arm non-inferiority test for skew-normal data

WU Qi^{1,2}, LIANG Fan^{1,2}, WU Liu-cang^{1,2}

(1. Faculty of Science, Kunming University of Science and Technology, Kunming 650500, China

2. Center for Applied Statistics, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: Noninferiority test is an important pharmacodynamic test, which aims to prove that a new drug is not inferior to a standard drug under a pre-set noninferiority margin. The three-arm non-inferiority test includes a placebo group, a standard drug group, and a new drug group, and can simultaneously test the superiority of the standard drug over the placebo and the non-inferiority of the new drug over the standard drug. In real life, the data obtained from experiments are often not strictly normal distribution, but have the characteristics of skewness, so it is necessary to study a new method for the non-inferiority test of skewness three-arm. The method of moment estimation is considered to estimate the unknown parameters, uses the central limit theorem to construct approximate test statistics, and further based on Bootstrap method, uses Bootstrap sample to estimate unknown parameters. Construct the Bootstrap test statistic and give the test method. The probability, efficiency and sample size of the first type of error were obtained by Monte Carlo simulation under different known parameters. The simulation results show that the method is effective and stable. Finally, the method was applied to the case data of an HIV treatment to determine its non-inferiority.

Keywords: three-arm non-inferiority trial; skew-normal distribution; Bootstrap sampling; Monte Carlo simulation

MR Subject Classification: 62F03; 62F10