

关于基因多效性检验中互斥假设检验的总结与探索

范宇轩, 李兰玉*

青岛大学数学与统计学院, 山东 青岛

收稿日期: 2023年3月19日; 录用日期: 2023年4月15日; 发布日期: 2023年4月21日

摘要

基因多效性是指单个基因能够影响多个性状的表达水平。能够有效检测基因多效性的存在并了解其产生的原因具有重要的统计学以及生物学意义。复杂原假设条件下的互斥检验方法十分适用于该类问题且具有较好的说服力。不失一般性的情况下, 本文主要针对两个性状的汇总统计量进行研究。我们将新构建的方法与IUT (intersection-union test)、PLACO (pleiotropic analysis under composite null hypothesis)、DACT (divide-aggregate composite-null test)、MAIUP (mixture-adjusted intersection-union test for pleiotropy test)几种方法进行了对比。模拟结果显示, 新方法与MAIUP方法相比在能够在某些特定的情形下能够更好的控制一类错误同时保证检验功效。最后将上述方法应用于冠状动脉疾病以及扁桃体切除的数据来进行实证分析。

关键词

互斥检验, 基因多效性, 汇总统计量

Summary and Exploration of Exclusive Hypothesis Test of Pleiotropy

Yuxuan Fan, Lanyu Li*

School of Mathematics and Statistics, Qingdao University, Qingdao Shandong

Received: Mar. 19th, 2023; accepted: Apr. 15th, 2023; published: Apr. 21st, 2023

Abstract

Pleiotropy means single gene can influence multiple traits. Detection and understanding of pleio-
*通讯作者。

trophy are statistically and biologically important. Mutually exclusive test under composite null hypothesis is a suitable and convincing approach. Without loss of generality, we mainly focus on analyzing summary statistics of two traits in this study. We compared our methods with IUT, PLACO, DACT, and MAIUP through simulations. Simulations show that the type I error can be controlled better by proposed method compared with MAIUP under specific situations with maintaining the statistical power simultaneously.

Keywords

Exclusive Hypothesis Test, Pleiotropy, Summary Statistics

Copyright © 2023 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

在遗传学发展的开始阶段, 人们对于基因的了解主要借助于对性状的研究, 而在其中也产生出单个基因影响单个性状的最初概念。随着时间的推移, 人们逐渐认识到基因不仅能直接控制某一性状的表现, 而且通过某一性状的变化, 间接能够影响到许多性状。比如, 翻毛鸡的翻毛基因, 不仅影响鸡羽的反卷, 而且还影响其体温下降, 心跳增快, 心、脾扩大, 生育力降低等, 这说明基因通过生理生化过程而影响一系列性状的表现[1]。

探索基因多效性对于统计学和生物学都具有重要意义。从统计学角度来看, 基因多效性的引入提高了关联分析的检验功效[2], 使得更加准确地发现的相关表型之间的遗传关系[3], 揭示无法通过分析单一表型来识别的遗传结构, 提高表型预测的准确性[4]。此外对于生物学而言, 有可能提高对全基因组学等基因医学的理解, 或者是定位可影响多种性状或疾病的药物遗传靶点, 从而推动高效药物的研发。

因此, 大量关于研究基因多效性的统计学方法应运而生。从单个遗传变异位点与多个性状的单变量边际关联的比较, 到所有性状对一个遗传变异位点同时进行回归的多变量分析, 再到一个遗传变异位点对所有性状进行反向回归分析的方法, 各类方法仍然在持续发展。其中针对单变量分析经常使用 meta 分析[5], 此方法虽然简单可行但忽略了性状之间的相关性; 而针对多变量的分析中常用‘所有性状同时对单个变异的回归’、‘小等位基因剂量对所有性状的回归’以及‘性状与小等位基因剂量矩阵的典型相关’等方法[6]。

尽管对于多效性研究投入了大量的关注, 但上述方法所构造的原假设并不能正确的反映基因多效性的含义。这是由于上述方法中的原假设是: 没有一个性状与一个变异位点相关, 但是仅有一个性状与变异有关时也可以拒绝这一原假设, 此时并不符合多效性的定义, 因此需要事后分析来解释是否存在多效性。然而随着基因变异数量的增加, 工作量也会不断加大。本文所研究的互斥检验的统计方法, 能够很好的提高检验多效性方法的可解释性和效率, 它将原假设拆分为一系列的子原假设, 并且这些子原假设互斥。即当且仅当子原假设其一成立时, 原假设才成立。

Schaid 等人[7]基于两个性状情形下的交并检验思想出发, 提出了一个基于线性模型和多元正态分布的顺序检验框架来分析多效性, 并且与以往的假设检验原假设的构造不同, 此时将原假设更改为: 至多

有一个性状与基因变异相关联。此后基于 Schaid 等人提出的序列检验框架汪毅等人[8]给予其赋予了互斥检验的名称。互斥检验改变了之前原有的原假设的构成, 从根本上做出了修正, 此时拒绝原假设则能够表明存在多效性, 因此给检验生物基因多效性提供了强有力的研究方向。

在 Schaid 等人提出的方法中, 当检验没有个性状与基因变异相关这一子假设时, 存在第一类错误十分保守的问题。之后, 汪毅等人参考 Schaid 等人的方法对P值的构造实施了改进。他们基于似然比以及 BIC 等相关知识给予原假设中拆分成的每个子假设不同的权重, 由此得到原假设的加权P值, 结果显示此时得到的P值保守程度减少, 能够基本上达到一类错误率的合理范围。

但上述方法仍然存在局限性, 其方法构建主要使用个体水平数据(即已知基因型、性状等的相关数据), 容易涉及样本人群隐私信息, 所以该类方法无法得到广泛应用。

汇总统计量的出现有效解决了这一问题, 其主要包含单个变异位点与单个性状之间关联性效应大小、回归估计误差以及检验P值等总括信息。在这一信息来源的支持下, 又促成了诸多基于汇总统计量的基因多效性检验方法。在其中较为突出的有 Debashree Ray 等人[9]提出的 PLACO (pleiotropic analysis under composite null hypothesis)、Zhonghua Liu 等人[10]提出的 DACT (divide-aggregate composite-null test) 以及 Ting Wang 等人[11]提出的 MAIUP (mixture-adjusted intersection-union test for pleiotropy test) 法。本文主要针对这几种方法进行对比分析, 并做出了一定的改进。

2. 符号定义

在不失一般性的前提下, 本文主要研究各类方法针对于两个性状的情况。

假设收集的两个性状的相关数据来自样本数为 n_1 , n_2 的个体, 设 $n_1 = n_2 = n$, 假设个体 n_1 与 n_2 之间独立, 即研究之间没有重叠样本。 $Y_k (k=1,2)$ 代表第 k 个性状的值, X 表示基因型。其中两种性状的值表示为: $Y_1 = (y_{11}, y_{12}, \dots, y_{1n})$, $Y_2 = (y_{21}, y_{22}, \dots, y_{2n})$, $y_{ij} (i=1,2; j=1,2, \dots, n)$ 表示来自第 j 个个体的第 i 性状的值; 位点的基因型值表示为: $X = (x_1, x_2, \dots, x_n)$, $x_i (i=1,2, \dots, n)$ 表示基因值。若所考虑的遗传变异是双等位单核苷酸多态性(SNP), 个体的基因型可以取值为 0、1 或 2。为简单起见, 我们假设不存在协变量, 通过考虑性状残差(通过对性状上的协变量进行回归得到)而不是原始的性状值, 可以很容易地放宽这一假设条件。尽管残差结果数据不是标准化的数据, 但 Broadaway KA [12] 等人的研究表明它不会影响基因关联测试的有效性。

对于给定基因变量, 即有标准单变量回归模型表示为:

$$y_1 = X\beta_1 + \varepsilon_1, \quad y_2 = X\beta_2 + \varepsilon_2, \quad (2.1)$$

其中, $\beta_k (k=1,2)$ 是第 k 个性状的遗传效应, $\varepsilon_1 \sim N(0, I_n \sigma_1^2)$ 、 $\varepsilon_2 \sim N(0, I_n \sigma_2^2)$, I_n 表示为 n 维单位矩阵。

Wald 检验统计量定义为: $Z_k = \frac{\widehat{\beta}_k}{\widehat{\sigma}_k}$ 。其中, $\widehat{\sigma}_k = se(\widehat{\beta}_k)$, $\widehat{\beta}_k$ 是关于 β_k 的极大似然估计, $\widehat{\sigma}_k$ 是其标准差的估计。

基于上述模型检验多效性的假设检验设定为:

$$H_0: \text{至多只有一个 } \beta_k \text{ 不为零} \quad \text{vs} \quad H_1: \beta_1 \neq 0 \text{ 且 } \beta_2 \neq 0 \quad (2.2)$$

由互斥检验的定义, 可以将原假设 H_0 分为三个互斥的子假设 H_{00} 、 H_{01} 、 H_{10} 组成。即:

$$H_0 = \begin{cases} H_{00}: \beta_1 = \beta_2 = 0 & \text{①} \\ H_{01}: \beta_1 = 0, \beta_2 \neq 0 & \text{②} \\ H_{10}: \beta_1 \neq 0, \beta_2 = 0 & \text{③} \end{cases} \quad \text{vs} \quad H_1: \beta_1 \neq 0 \text{ 且 } \beta_2 \neq 0, \quad (2.3)$$

即有 $H_0 = H_{00} \cup H_{01} \cup H_{10}$, 而 H_1 代表我们最感兴趣的多效性的情况。

3. P 值的构建

3.1. IUT 法

上文公式(2.3)可以变换为 $H_0 = H_0^{(1)} \cup H_0^{(2)}$ 。其中:

$$H_0^{(k)} : \beta_k = 0 \quad \text{vs} \quad H_1^{(k)} : \beta_k \neq 0 \quad (3.1)$$

针对每个 $H_0^{(k)}$ 检验都会有一个对应的 p_k 值, 因此关于原假设 H_0 的 p 值由交并检验知识可知为:

$$p = \max\{p_1, p_2\}. \quad (3.2)$$

此时 p_1, p_2 值计算依赖于 Wald 检验统计量 $Z_k (k=1,2)$ 在 $H_0^{(k)}$ 下服从标准正态分布计算而得。因此 IUT 方法也称为最大 p 值法或联合显著性检验[13]。

然而此方法在任意有限样本量的情况下一类错误会出现保守现象, 尤其是当 H_{00} 成立时候, 原因如下:

假设两性状独立, 则有 p_1 和 p_2 是独立的, 所以有

$$\Pr(p < \alpha) = \Pr(p_1 < \alpha)\Pr(p_2 < \alpha). \quad (3.3)$$

在有限样本量下, H_{01} 情形有 $\Pr(p_1 < \alpha | \beta_1 = 0) = \alpha$, 并且 $\Pr(p_2 < \alpha | p_2 \neq 0) < 1$, 所以有 $\Pr(p < \alpha) < \alpha$; 同理在 H_{10} 情形下亦是如此。

然而在 H_{00} 会更保守的原因是:

由于 $\Pr(p_1 < \alpha) = \Pr(p_2 < \alpha) = \alpha$, 即有 $\Pr(p < \alpha) = \alpha^2$, 对任意的 $\alpha \in (0,1)$ 。因此在此种情形下一类错误会更加保守。

3.2. PLACO 法

为了能够改进 IUT 方法下一类错误的保守性, 同时考虑多效性检验的复合假设的特性, Debashree Ray [9]提出了 PLACO, 此方法的不同之处在于它没有依赖于 p 值, 而依赖于 Z 检验统计量。通过观察假设(3.1)式中的原假设发现, 原假设可以改写为:

$$H_0 : \beta_1\beta_2 = 0 \quad \text{vs} \quad H_1 : \beta_1\beta_2 \neq 0 \quad (3.4)$$

因此需要考虑 $Z_1 * Z_2$ 的分布, 可以借助 Craig [14]提到的关于两变量 a, b 独立且均服从于正态分布, 则两变量分别标准化后得到的乘积 X 的密度函数为: $f(x) = \frac{\kappa_0(x)}{\pi}$, 其中 $\kappa_0(x)$ 为第二类的阶数为零阶的修正贝塞尔函数。

换句话说, 假设有如下情形: H_{00} 成立时, Z_1 和 Z_2 独立且服从于 $N(0,1)$; H_{01} 成立时, Z_1 和 Z_2 独立且分别服从于 $N(0,1)$ 和 $N(0,1+\tau_2^2)$; H_{10} 成立时, Z_1 和 Z_2 独立且分别服从于 $N(0,1+\tau_1^2)$ 和 $N(0,1)$ 。假设在全局假设下 H_{00} 成立的概率为 π_{00} , H_{01} 以 π_{01} 的概率成立, H_{10} 以概率 π_{10} 成立, 则 Z_2 服从标准正态分布 $N(0,1)$, 而 Z_1 在给定均值 μ_1 服从 $N(0, \tau_1^2)$ 前提下, 服从 $N(\mu_1, 1)$ 的分布。在此假设下 p 值计算为:

$$\begin{aligned} p_{z_1 z_2} &= 2 * P_{H_0} (Z_1 Z_2 > |z_1 z_2|) \\ &= 2 * \{P(H_{00})P_{H_{00}} (Z_1 Z_2 > |z_1 z_2|) + P(H_{01})P_{H_{01}} (Z_1 Z_2 > |z_1 z_2|) + P(H_{10})P_{H_{10}} (Z_1 Z_2 > |z_1 z_2|)\} \\ &= \pi_{00} F(z_1 z_2) + \pi_{01} F\left(\frac{z_1 z_2}{\sqrt{1+\tau_2^2}}\right) + \pi_{10} F\left(\frac{z_1 z_2}{\sqrt{1+\tau_1^2}}\right) \end{aligned} \quad (3.5)$$

其中 z_1 和 z_2 为观测值, 且 $F(u) = 2 \int_{|u|}^{\infty} f(x) dx$ 是标准正态乘积分布的双边尾部概率。

因此, 该方法主要需要知道均值 μ_1 和 μ_2 的数值, 以及了解三个子原假设可能成立的概率。PLACO 方法通过设置额外的模型假设(即假设两性状独立)来克服上述问题带来的困难, 最终根据 Huang YT [15] 中提到的方法中在不涉及估计未知参数的同时可以将 p 值近似的估计为:

$$\hat{P}_{z_1 z_2} \approx F\left(z_1 \frac{z_2}{\sqrt{\text{var}(z_2)}}\right) + F\left(\frac{z_1}{\sqrt{\text{var}(z_1)}}\right) - F(z_1 z_2). \quad (3.6)$$

3.3. DACT 法

与 PLACO 不同, DACT 直接估计三个子原假设的比例, 最终将单个子原假设获得的 P 值的加权总和作为结果。

假设取自 m 个样本, 则需要进行 m 个假设检验来进行检验 $H_0: \beta_1 \beta_2 = 0$ vs $H_1: \beta_1 \beta_2 \neq 0$ 。在回归模型中针对于第一个性状所涉及到的参数 β_1 检验有 m 个假设检验: $H_{0j}^{\beta_1}: \beta_1 = 0$, 同样针对第二个性状所涉及到的参数 β_2 检验亦有 m 个假设检验: $H_{0j}^{\beta_2}: \beta_2 = 0$, 这里 $(1 \leq j \leq m)$ 。同时将 $H_j^{\beta_1}$ (对于 $H_j^{\beta_2}$ 亦是如此) 定义为伯努利随机变量, 即: 如果假设 $H_{0j}^{\beta_1}$ 为真, 则 $H_j^{\beta_1} = 0$; 如果假设 $H_{0j}^{\beta_1}$ 为假, 则 $H_j^{\beta_1} = 1$, $(1 \leq j \leq m)$ 。正如前文独立的假设条件, 可以得到 $H_j^{\beta_1}$ 和 $H_j^{\beta_2}$ 是独立的。对于每一个基因, 通过检验统计量可以获得检验 β_1 的 p 值 p_{β_1} 以及检验 β_2 的 p 值 p_{β_2} , $(1 \leq j \leq m)$ 。

由 Efron [16] 假设 $P(H_j^{\beta_1} = 0) = \pi_0^{\beta_1}$, $P(H_j^{\beta_2} = 0) = \pi_0^{\beta_2}$, 其中, 参数 $\pi_0^{\beta_1}$ 是第一个性状不受遗传基因影响在全基因组中的比例, $\pi_0^{\beta_2}$ 是第二个性状不受遗传基因影响在全基因组中的比例。即有: (2.8) 式情况①有 $\Pr(H_j^{\beta_1} = 0, H_j^{\beta_2} = 0) = \pi_0^{\beta_1} \pi_0^{\beta_2}$; 情况②有 $\Pr(H_j^{\beta_1} = 0, H_j^{\beta_2} = 1) = \pi_0^{\beta_1} (1 - \pi_0^{\beta_2})$; 情况③有 $\Pr(H_j^{\beta_1} = 1, H_j^{\beta_2} = 0) = (1 - \pi_0^{\beta_1}) \pi_0^{\beta_2}$; 备择假设有 $\Pr(H_j^{\beta_1} = 1, H_j^{\beta_2} = 1) = (1 - \pi_0^{\beta_1}) (1 - \pi_0^{\beta_2})$ 。

由上述公式成立可知, 多效性的假设检验(2.8)式中的三个子假设的相对比例则分别为: $\omega_1 = \frac{\pi_0^{\beta_1} \pi_0^{\beta_2}}{c}$, $\omega_2 = \frac{\pi_0^{\beta_1} (1 - \pi_0^{\beta_2})}{c}$, $\omega_3 = \frac{(1 - \pi_0^{\beta_1}) \pi_0^{\beta_2}}{c}$, $c = \pi_0^{\beta_1} \pi_0^{\beta_2} + \pi_0^{\beta_1} (1 - \pi_0^{\beta_2}) + (1 - \pi_0^{\beta_1}) \pi_0^{\beta_2}$ 。

此时只需要需要估计 $\pi_0^{\beta_1}, \pi_0^{\beta_2}$ 即可。通过 Jin 和 Cai [17] 提出使用经验特征函数和傅里叶分析来估计真原假设比例的估计方法, 并且是基于检验 $\beta_1 = 0, \beta_2 = 0$ 的 Z 统计量来进行估计。假设有 m 个 Z 检验统计量, 即: $Z_j \sim N(\mu_j, \sigma_j^2), (1 \leq j \leq m)$ 。这里我们设置为 $\mu_j = \mu_0 = 0, \sigma_j^2 = \sigma_0^2 = 1$ 。给定 $r \in \left(0, \frac{1}{2}\right)$, 关于真原假设 π_0 的比例可以一致地估计为:

$$\hat{\pi}_0 = \sup_{\{0 \leq t \leq \sqrt{2r \log(m)}\}} \left\{ \int_{-1}^1 (1 - |\zeta|) \left(\text{Re} \left(\psi_m(t\zeta, Z_1, \dots, Z_m) \exp(-i\mu_0 t\zeta + \sigma_0^2 t^2 \zeta^2 / 2) \right) \right) d\zeta \right\} \quad (3.7)$$

此处 $\text{Re}(a)$ 表示复数 a 的实部。其中 ψ_m 为经验特征函数, 为 $\psi_m(t) = \frac{1}{m} \sum_{j=1}^m \exp(itZ_j)$ 。因此通过此方法可以估计 $\pi_0^{\beta_1}$ 和 $\pi_0^{\beta_2}$ 以及相应的权重 $\omega_1, \omega_2, \omega_3$ 。

在有了权重后, 还需要考虑如何定义三个子原假设中进行多效性的检验。则有在检验 $\beta_1 \neq 0$ 且 $\beta_2 = 0$ 这种情形时只需要通过 p_{β_2} 来检验 $\beta_2 = 0$ 即可, 这是因为 $\beta_1 \neq 0$; 同样在检验 $\beta_1 = 0$ 且 $\beta_2 \neq 0$ 时, 由于 $\beta_2 \neq 0$, 所以只需要通过 p_{β_1} 来检验 $\beta_1 = 0$ 即可; 对于 $\beta_1 = \beta_2 = 0$ 这种情况, 需要检验 β_1 和 β_2 是否都非零。在给定显著性水平 α 下, 当 $\text{Max}P = \max(p_{\beta_1}, p_{\beta_2}) < \alpha$ 时, 就会拒绝此种情况。给出的子假设 $\beta_1 = \beta_2 = 0$ 情况下中, 由于 $\text{Max}P$ 服从 $\beta(2, 1)$ 分布, 其平方服从均匀分布, 因此在此情况下的 p 值取为

$(MaxP)^2$ 。

即有 p 值计算为:

$$p = \begin{cases} p_{1j} = (MaxP_j)^2; & \text{if } \beta_1 = \beta_2 = 0 \\ p_{2j} = p_{\beta_1}; & \text{if } \beta_1 = 0, \beta_2 \neq 0 \\ p_{3j} = p_{\beta_2}; & \text{if } \beta_1 \neq 0, \beta_2 = 0 \end{cases} \quad (3.8)$$

因此原假设的 p 值计算为:

$$p = \widehat{\omega}_1 p_{1j} + \widehat{\omega}_2 p_{2j} + \widehat{\omega}_3 p_{3j}. \quad (3.9)$$

3.4. MAIUP 法

由于基于 IUT 的方法时, 尤其是在检验多效性中的两个基因与某一性状都无关这一检验的时候, 所获得的一类错误往往过于保守。为了修正 IUT 在探索多效性时一类错误的保守性, Ting Wang [11]提出了 MAIUP 的方法。MAIUP 基于传统的交并检验, 以两组独立的 p 值作为输入值, 并基于最初在高维中介分析[18] (mediation analysis)框架下而提出的一个新想法。MAIUP 的主要修正在于它通过拟合三组子原假设的分布, 将多效性检验的复合假设检验的形式原假设考虑在内, 从而最终生成有良好校准性的 p 值。简言之, MAIUP 通过遵循 IUT 的原则来检测基因多效性, 并且用一对表型的每个遗传位点的两个 p 值的最大值(用 p_{\max})表示作为显著性的度量。具体来说, 从 IUT 方法中可以获得的一组 $p_{\max,j} (j=1,2,\dots,S)$; (S 表示样本个数)。此时, 我们不依赖于其从 0 到 1 的常规的均匀分布, 而是通过借用整个基因组中所有基因的有用信息来估计每个子原假设的比例, 并通过拟合三个子原假设混合的分布来构建 $p_{\max,j}$ 的新的分布。即:

$$\begin{aligned} & \Pr(p_{\max,j} \leq t | H_{0j}) \\ &= \Pr(p_{\max,j} \leq t | H_{01,j})\Pr(H_{01,j}) + \Pr(p_{\max,j} \leq t | H_{10,j})\Pr(H_{10,j}) \\ & \quad + \Pr(p_{\max,j} \leq t | H_{00,j})\Pr(H_{00,j}) \\ &= \pi_{01} \Pr(p_{1j} \leq t | H_{01,j})\Pr(p_{2j} \leq t | H_{01,j}) + \pi_{10} \Pr(p_{1j} \leq t | H_{10,j})\Pr(p_{2j} \leq t | H_{10,j}) \\ & \quad + \pi_{00} \Pr(p_{1j} \leq t | H_{00,j})\Pr(p_{2j} \leq t | H_{00,j}) \\ &= \pi_{01}t \Pr(p_{2j} \leq t | H_{01,j}) + \pi_{10}t \Pr(p_{2j} \leq t | H_{10,j}) + \pi_{00}t^2 \end{aligned} \quad (3.10)$$

上式成立前提是 p_{1j} 和 p_{2j} 两者独立, 其中, π_{00} 、 π_{01} 和 π_{10} 是相对应的三个子原假设所对应的比例, t 是评价多效性显著性给定的一个阈值。同时, $\Pr(p_{1j} \leq t | H_{10,j})$ 是在假设 $H_{10,j}$ 下拒绝 $\beta_{1j} = 0$ 的功效函数, 同理 $\Pr(p_{2j} \leq t | H_{01,j})$ 是在假设 $H_{01,j}$ 下拒绝 $\beta_{2j} = 0$ 的功效函数。当在样本量比较大时, 两者趋于 1, 因此最终结果变为 $\pi_{01}t + \pi_{10}t + \pi_{00}t^2$ 。

在计算上式时还需要估计参数 $\pi_{01}, \pi_{10}, \pi_{00}$ 。依据 James Y. Dai [18]和 John D. Storey [19]提出的估计方法, 可以进行上述参数的估计。以 $\pi_{0+}(\lambda_1)$ 表示为 $\beta_1 = 0$ 时的假设下的原假设成立的比例, $\pi_{+0}(\lambda_2)$ 表示为 $\beta_2 = 0$ 时的假设下的原假设成立的比例。因此 $\pi_{0+}(\lambda_1)$ 和 $\pi_{+0}(\lambda_2)$ 可以保守的估计为:

$$\widehat{\pi}_{0+}(\lambda_1) = \frac{1}{S(1-\lambda_1)} \#\{p_{1j} > \lambda_1\}, \quad \widehat{\pi}_{+0}(\lambda_2) = \frac{1}{S(1-\lambda_2)} \#\{p_{2j} > \lambda_2\}. \quad (3.11)$$

其中 S 表示为在分析中样本的所有数目, λ_1, λ_2 表示两个调谐参数, 用于确定观察到的 p 值是否来自原假

设情况。与此同时 π_{00} 可以估计为:

$$\widehat{\pi}_{00}(\lambda_1, \lambda_2) = \frac{1}{S(1-\lambda_1)(1-\lambda_2)} \#\{p_{1j} > \lambda_1, p_{2j} > \lambda_2\}. \quad (3.12)$$

所以 π_{01}, π_{10} 计算分别为:

$$\widehat{\pi}_{01}(\lambda_1) = \widehat{\pi}_{0+}(\lambda_1) - \widehat{\pi}_{00}(\lambda_1, \lambda_2); \quad \widehat{\pi}_{10}(\lambda_1) = \widehat{\pi}_{+0}(\lambda_1) - \widehat{\pi}_{00}(\lambda_1, \lambda_2). \quad (3.13)$$

借助 James Y. Dai [18] 试用的结果, 此处采用 $\lambda_1 = \lambda_2 = 0.5$ 。将上述估计值带入式(3.10)中, 即可获得最终的结果。

3.5. 新构造的 P 值

DACT 方法中在估计三个子原假设的比例时借用了独立的严格前提条件, 而在 MAIUP 方法中估计三个子原假设的比例时条件相对较为宽泛, 所以在两种方法估计出来的比例进行比较后, 发现 MAIUP 方法做得到的比例更为接近, 因此考虑利用 MAIUP 方法中估计三个子原假设的比例作为 DACT 方法中子原假设的比例。即有 $\pi_{00}, \pi_{01}, \pi_{10}$ 估计为:

$$\begin{aligned} \widehat{\pi}_{00}(\lambda_1, \lambda_2) &= \frac{1}{S(1-\lambda_1)(1-\lambda_2)} \#\{p_{1j} > \lambda_1, p_{2j} > \lambda_2\}; \\ \widehat{\pi}_{01}(\lambda_1) &= \widehat{\pi}_{0+}(\lambda_1) - \widehat{\pi}_{00}(\lambda_1, \lambda_2); \\ \widehat{\pi}_{10}(\lambda_1) &= \widehat{\pi}_{+0}(\lambda_1) - \widehat{\pi}_{00}(\lambda_1, \lambda_2) \end{aligned}$$

其中, $\widehat{\pi}_{0+}(\lambda_1) = \frac{1}{S(1-\lambda_1)} \#\{p_{1j} > \lambda_1\}$, $\widehat{\pi}_{+0}(\lambda_2) = \frac{1}{S(1-\lambda_2)} \#\{p_{2j} > \lambda_2\}$, 并且 S 表示为在分析中样本的所有数目, λ_1 和 λ_2 表示两个调谐参数, 用于确定观察到的 p 值是否来自原假设情况。借助 James Y. Dai [18] 比较试用的结果, 此处采用 $\lambda_1 = \lambda_2 = 0.5$ 。即各个子原假设权重为 $\omega_1 = \frac{\widehat{\pi}_{00}}{c}$, $\omega_2 = \frac{\widehat{\pi}_{01}}{c}$, $\omega_3 = \frac{\widehat{\pi}_{10}}{c}$, $c = \widehat{\pi}_{00} + \widehat{\pi}_{01} + \widehat{\pi}_{10}$ 。

对于提出的新方法里各个子假设情形下每种情况 p 值取值同 DACT 方法。即在检验 $\beta_1 \neq 0, \beta_2 = 0$ 这种情形时只需要通过 p_{β_2} 来检验 $\beta_2 = 0$, 是因为 $\beta_1 \neq 0$; 同样在检验 $\beta_1 = 0, \beta_2 \neq 0$ 时, 由于 $\beta_2 \neq 0$, 所以只需要通过 p 值 p_{β_1} 来检验 $\beta_1 = 0$ 即可; 而对于 $\beta_1 = 0, \beta_2 = 0$ 这种情况, 需要检验 β_1 和 β_2 是否都非零, 在此种情况下取 $(MaxP)^2$ 。即有 p 值计算为:

$$p = \begin{cases} p_{1j} = (MaxP_j)^2; \beta_1 = \beta_2 = 0 \\ p_{2j} = p_{\beta_1}; \beta_1 = 0, \beta_2 \neq 0 \\ p_{3j} = p_{\beta_2}; \beta_1 \neq 0, \beta_2 = 0 \end{cases} \quad (3.14)$$

因此原假设的 p 值的最终计算为:

$$p = \omega_1 p_{1j} + \omega_2 p_{2j} + \omega_3 p_{3j}. \quad (3.15)$$

4. 模拟检验

4.1. 模拟设置

为了评估这些多效性检验的方法的一类错误以及功效的性能, 本文进行广泛的模拟研究。由于所有

这些方法都是使用汇总统计量类型的数据集进行的, 因此我们直接从给定的二元正态分布在各种子原假设情景下生成 Z 统计量。具体地说, 在 H_{00} 情形下以 π_{00} 的概率生成 Z 统计量为 $BVN((0,0),\Sigma)$, 在 H_{01} 情形下以 π_{01} 的概率生成 Z 统计量为 $BVN((0,\mu_{01}),\Sigma)$, 在 H_{10} 情形下以 π_{10} 的概率生成 Z 统计量为 $BVN((\mu_{10},0),\Sigma)$, 或者在 H_1 情形下以 π_{11} 的概率生成 Z 统计量为 $BVN((\mu_1,\mu_2),\Sigma)$ 。同时设置了样本量的数量 S 为 100、1000 或 2000, $\mu_{10} = \mu_{01} = \mu_1 = \mu_2$ 且取值为 4、10 或 20, 并考虑了多种概率设置, 主要以设置 $\pi_{11} = 0$ 用于检验控制一类错误的情况, 以及设置 $\pi_{11} \neq 0$ 来检验功效水平。与此同时 μ_{10} (或 μ_{01} 、 μ_1 和 μ_2) 的大小量化了基因与性状之间关联的强度大小, 值越大表示关联的信号越强。同时基于正态分布的知识近似的将 Z 统计量转换为对应的 p 值。在主要的模拟设置中, 将 Σ 设置为二维的单位矩阵; 同时进行了另一项模拟研究, 以评估在分析的性状中存在相关关系的情形, 研究去相关性后和不去相关性时的这些方法的一类错误和功效的表现情况, 来检验相关性将如何影响这些基于原假设的复合多效性测试方法的性能。此时将 Σ 的非对角线元素指定为非零值, 以生成相关检验统计的数据。

对于最终一类错误的控制和功效的评估, 我们在每个设置中重复模拟 1000 次, 最终以这些重复结果得到的平均值作为最终结果。

4.2. 比例估计比较

我们首先通过设置在不同基因个数、不同均值大小等情况下来比较 DACT 和 MAIUP 两种方法对于子原假设比例的估计。具体结果见表 1~10, 表头下标用于区分不同方法。

从表格可以发现, 对于 DACT 方法比例估计的精确度要逊于 MAIUP 方法获得的比例估计, 即估计比例与真实比例之间偏差更大, 同时可以看出 MAIUP 方法在均值大于等于 6 时估计能够趋于稳定, 同时随着基因个数的增加, 估计的精确度会更加精准。

Table 1. Estimation of ratio of sub-null hypotheses (100 genes; $\pi_{00} = 0.8, \pi_{01} = 0.1, \pi_{10} = 0.1, \pi_1 = 0$)

表 1. 子原假设的比例估计(基因个数为 100; $\pi_{00} = 0.8, \pi_{01} = 0.1, \pi_{10} = 0.1, \pi_1 = 0$)

μ_{01} 或 μ_{10}	π_{00_DACT}	π_{01_DACT}	π_{10_DACT}	π_{00_MAIUP}	π_{01_MAIUP}	π_{10_MAIUP}
2	0.005	0.063	0.062	0.934	0.030	0.036
4	0.012	0.096	0.096	0.862	0.064	0.070
6	0.013	0.100	0.101	0.860	0.064	0.072
8	0.013	0.101	0.102	0.860	0.064	0.072
10	0.014	0.102	0.103	0.860	0.064	0.072
12	0.014	0.103	0.103	0.860	0.064	0.072

Table 2. Estimation of ratio of sub-null hypotheses (100 genes; $\pi_{00} = 0.04, \pi_{01} = 0.48, \pi_{10} = 0.48, \pi_1 = 0$)

表 2. 子原假设的比例估计(基因个数为 100; $\pi_{00} = 0.04, \pi_{01} = 0.48, \pi_{10} = 0.48, \pi_1 = 0$)

μ_{01} 或 μ_{10}	π_{00_DACT}	π_{01_DACT}	π_{10_DACT}	π_{00_MAIUP}	π_{01_MAIUP}	π_{10_MAIUP}
2	0.474	0.214	0.215	0.233	0.367	0.369
4	0.280	0.248	0.250	0.058	0.456	0.458
6	0.269	0.249	0.251	0.058	0.456	0.459
8	0.268	0.249	0.250	0.058	0.456	0.459
10	0.266	0.249	0.251	0.058	0.456	0.459
12	0.265	0.249	0.251	0.058	0.456	0.459

Table 3. Estimation of ratio of sub-null hypotheses (100 genes; $\pi_{00} = 1, \pi_{01} = 0, \pi_{10} = 0, \pi_{11} = 0$)**表 3.** 子原假设的比例估计(基因个数为 100; $\pi_{00} = 1, \pi_{01} = 0, \pi_{10} = 0, \pi_{11} = 0$)

μ_{01} 或 μ_{10}	π_{00_DACT}	π_{01_DACT}	π_{10_DACT}	π_{00_MAIUP}	π_{01_MAIUP}	π_{10_MAIUP}
2	0.000	0.019	0.020	0.995	0.002	0.003
4	0.000	0.019	0.020	0.995	0.002	0.003
6	0.000	0.019	0.020	0.995	0.002	0.003
8	0.000	0.019	0.020	0.995	0.002	0.003
10	0.000	0.019	0.020	0.995	0.002	0.003
12	0.000	0.019	0.020	0.995	0.002	0.003

Table 4. Estimation of ratio of sub-null hypotheses (100 genes; $\pi_{00} = 0, \pi_{01} = 0, \pi_{10} = 0, \pi_{11} = 1$)**表 4.** 子原假设的比例估计(基因个数为 100; $\pi_{00} = 0, \pi_{01} = 0, \pi_{10} = 0, \pi_{11} = 1$)

μ_{01} 或 μ_{10}	π_{00_DACT}	π_{01_DACT}	π_{10_DACT}	π_{11_DACT}	π_{00_MAIUP}	π_{01_MAIUP}	π_{10_MAIUP}	π_{11_MAIUP}
2	0.111	0.222	0.223	0.444	0.030	0.147	0.148	0.675
4	0.000	0.017	0.017	0.965	0.000	0.001	0.001	0.999
6	0.000	0.009	0.009	0.982	0.000	0.000	0.000	1.000
8	0.000	0.010	0.010	0.981	0.000	0.000	0.000	1.000
10	0.000	0.010	0.010	0.981	0.000	0.000	0.000	1.000
12	0.000	0.010	0.010	0.981	0.000	0.000	0.000	1.000

Table 5. Estimation of ratio of sub-null hypotheses (100 genes; $\pi_{00} = 0.1, \pi_{01} = 0.05, \pi_{10} = 0.05, \pi_{11} = 0.8$)**表 5.** 子原假设的比例估计(基因个数为 100; $\pi_{00} = 0.1, \pi_{01} = 0.05, \pi_{10} = 0.05, \pi_{11} = 0.8$)

μ_{01} 或 μ_{10}	π_{00_DACT}	π_{01_DACT}	π_{10_DACT}	π_{11_DACT}	π_{00_MAIUP}	π_{01_MAIUP}	π_{10_MAIUP}	π_{11_MAIUP}
2	0.307	0.249	0.245	0.198	0.137	0.156	0.157	0.549
4	0.028	0.140	0.140	0.691	0.090	0.057	0.057	0.796
6	0.133	0.133	0.025	0.709	0.090	0.056	0.057	0.797
8	0.025	0.133	0.133	0.710	0.090	0.056	0.057	0.797
10	0.025	0.133	0.132	0.710	0.090	0.056	0.057	0.797
12	0.024	0.132	0.132	0.712	0.090	0.056	0.057	0.797

Table 6. Estimation of ratio of sub-null hypotheses (1000 genes; $\pi_{00} = 0.8, \pi_{01} = 0.1, \pi_{10} = 0.1, \pi_{11} = 0$)**表 6.** 子原假设的比例估计(基因个数为 1000; $\pi_{00} = 0.8, \pi_{01} = 0.1, \pi_{10} = 0.1, \pi_{11} = 0$)

μ_{01} 或 μ_{10}	π_{00_DACT}	π_{01_DACT}	π_{10_DACT}	π_{00_MAIUP}	π_{01_MAIUP}	π_{10_MAIUP}
2	0.007	0.076	0.075	0.847	0.071	0.070
4	0.011	0.095	0.095	0.812	0.089	0.088
6	0.012	0.098	0.098	0.812	0.089	0.088
8	0.012	0.098	0.098	0.812	0.089	0.088
10	0.012	0.099	0.099	0.812	0.089	0.088
12	0.012	0.099	0.099	0.812	0.089	0.088

Table 7. Estimation of ratio of sub-null hypotheses (1000 genes; $\pi_{00} = 0.04, \pi_{01} = 0.48, \pi_{10} = 0.48, \pi_1 = 0$)

表 7. 子原假设的比例估计(基因个数为 1000; $\pi_{00} = 0.04, \pi_{01} = 0.48, \pi_{10} = 0.48, \pi_1 = 0$)

μ_{01} 或 μ_{10}	π_{00_DACT}	π_{01_DACT}	π_{10_DACT}	π_{00_MAIUP}	π_{01_MAIUP}	π_{10_MAIUP}
2	0.146	0.235	0.237	0.220	0.383	0.385
4	0.277	0.249	0.250	0.050	0.469	0.470
6	0.270	0.249	0.250	0.049	0.470	0.471
8	0.268	0.249	0.250	0.049	0.470	0.471
10	0.267	0.249	0.250	0.049	0.470	0.471
12	0.267	0.249	0.250	0.049	0.470	0.471

Table 8. Estimation of ratio of sub-null hypotheses (1000 genes; $\pi_{00} = 1, \pi_{01} = 0, \pi_{10} = 0, \pi_1 = 0$)

表 8. 子原假设的比例估计(基因个数为 1000; $\pi_{00} = 1, \pi_{01} = 0, \pi_{10} = 0, \pi_1 = 0$)

μ_{01} 或 μ_{10}	π_{00_DACT}	π_{01_DACT}	π_{10_DACT}	π_{00_MAIUP}	π_{01_MAIUP}	π_{10_MAIUP}
2	0.000	0.013	0.013	0.998	0.001	0.001
4	0.000	0.013	0.013	0.998	0.001	0.001
6	0.000	0.013	0.013	0.998	0.001	0.001
8	0.000	0.013	0.013	0.998	0.001	0.001
10	0.000	0.013	0.013	0.998	0.001	0.001
12	0.000	0.013	0.013	0.998	0.001	0.001

Table 9. Estimation of ratio of sub-null hypotheses (1000 genes; $\pi_{00} = 0, \pi_{01} = 0, \pi_{10} = 0, \pi_1 = 1$)

表 9. 子原假设的比例估计(基因个数为 1000; $\pi_{00} = 0, \pi_{01} = 0, \pi_{10} = 0, \pi_1 = 1$)

μ_{01} 或 μ_{10}	π_{00_DACT}	π_{01_DACT}	π_{10_DACT}	π_{1_DACT}	π_{00_MAIUP}	π_{01_MAIUP}	π_{10_MAIUP}	π_{1_MAIUP}
2	0.040	0.160	0.161	0.638	0.031	0.147	0.146	0.676
4	0.000	0.018	0.018	0.963	0.000	0.001	0.001	0.998
6	0.000	0.010	0.010	0.979	0.000	0.000	0.000	1.000
8	0.000	0.009	0.009	0.981	0.000	0.000	0.000	1.000
10	0.000	0.009	0.009	0.981	0.000	0.000	0.000	1.000
12	0.000	0.010	0.010	0.981	0.000	0.000	0.000	1.000

Table 10. Estimation of ratio of sub-null hypotheses (1000 genes; $\pi_{00} = 0.1, \pi_{01} = 0.05, \pi_{10} = 0.05, \pi_1 = 0.8$)

表 10. 子原假设的比例估计(基因个数为 1000; $\pi_{00} = 0.1, \pi_{01} = 0.05, \pi_{10} = 0.05, \pi_1 = 0.8$)

μ_{01} 或 μ_{10}	π_{00_DACT}	π_{01_DACT}	π_{10_DACT}	π_{1_DACT}	π_{00_MAIUP}	π_{01_MAIUP}	π_{10_MAIUP}	π_{1_MAIUP}
2	0.027	0.138	0.138	0.696	0.142	0.158	0.158	0.542
4	0.027	0.138	0.138	0.696	0.099	0.051	0.051	0.799
6	0.025	0.133	0.133	0.710	0.099	0.050	0.050	0.800
8	0.024	0.131	0.131	0.714	0.099	0.050	0.050	0.800
10	0.024	0.131	0.131	0.714	0.099	0.050	0.050	0.800
12	0.024	0.131	0.131	0.715	0.099	0.050	0.050	0.800

4.3. 一类错误

在假设检验中一类错误是一个重要的衡量指标, 在检验水平为 0.01、0.05 下, 我们设置在不同的均值、基因个数以及各个子原假设在不同比例下进行了相应一类错误的模拟得到如表 11、表 12 的结果。表中 New_method 表示新构造的 p 值方法。

Table 11. Comparison of type I error of different methods (significance level is 0.01)
表 11. 不同方法下一类错误的比较(显著性水平为 0.01)

子原假设比例	均值; 基因数	IUT	PLAO	DAT	MAIUP	New_method
(0.80,0.10,0.10,0.00)	4; 100	0.000	0.019	0.001	0.028	0.017
(0.04,0.48,0.48,0.00)		0.007	0.024	0.022	0.019	0.016
(0.34,0.33,0.33,0.00)		0.004	0.020	0.012	0.020	0.017
(1.00,0.00,0.00,0.00)		0.000	0.009	0.006	0.019	0.009
(0.80,0.10,0.10,0.00)	4; 1000	0.002	0.022	0.004	0.013	0.015
(0.04,0.48,0.48,0.00)		0.009	0.027	0.025	0.011	0.014
(0.34,0.33,0.33,0.00)		0.006	0.023	0.015	0.011	0.015
(1.00,0.00,0.00,0.00)		0.000	0.010	0.006	0.011	0.010
(0.80,0.10,0.10,0.00)	4; 2000	0.002	0.022	0.004	0.012	0.016
(0.04,0.48,0.48,0.00)		0.009	0.027	0.025	0.011	0.014
(0.34,0.33,0.33,0.00)		0.006	0.023	0.015	0.011	0.015
(1.00,0.00,0.00,0.00)		0.000	0.010	0.005	0.010	0.010
(0.80,0.10,0.10,0.00)	10; 100	0.000	0.034	0.001	0.027	0.018
(0.04,0.48,0.48,0.00)		0.007	0.028	0.024	0.018	0.018
(0.34,0.33,0.33,0.00)		0.004	0.025	0.013	0.020	0.019
(1.00,0.00,0.00,0.00)		0.000	0.009	0.006	0.014	0.009
(0.80,0.10,0.10,0.00)	10; 1000	0.002	0.036	0.004	0.013	0.013
(0.04,0.48,0.48,0.00)		0.009	0.031	0.027	0.011	0.014
(0.34,0.33,0.33,0.00)		0.006	0.029	0.016	0.011	0.020
(1.00,0.00,0.00,0.00)		0.000	0.009	0.006	0.011	0.010
(0.80,0.10,0.10,0.00)	10; 2000	0.002	0.036	0.004	0.012	0.012
(0.04,0.48,0.48,0.00)		0.009	0.032	0.027	0.011	0.013
(0.34,0.33,0.33,0.00)		0.006	0.029	0.016	0.011	0.020
(1.00,0.00,0.00,0.00)		0.000	0.010	0.005	0.010	0.010
(0.80,0.10,0.10,0.00)	20; 100	0.000	0.037	0.001	0.027	0.018
(0.04,0.48,0.48,0.00)		0.007	0.028	0.024	0.018	0.018
(0.34,0.33,0.33,0.00)		0.004	0.026	0.013	0.020	0.019
(1.00,0.00,0.00,0.00)		0.000	0.009	0.006	0.019	0.009

Continued

(0.80,0.10,0.10,0.00)		0.002	0.039	0.004	0.013	0.013
(0.04,0.48,0.48,0.00)	20; 1000	0.009	0.032	0.027	0.011	0.014
(0.34,0.33,0.33,0.00)		0.006	0.030	0.016	0.011	0.020
(1.00,0.00,0.00,0.00)		0.000	0.009	0.006	0.011	0.010
(0.80,0.10,0.10,0.00)		0.002	0.040	0.004	0.012	0.012
(0.04,0.48,0.48,0.00)	20; 2000	0.009	0.033	0.027	0.011	0.013
(0.34,0.33,0.33,0.00)		0.006	0.030	0.016	0.011	0.020
(1.00,0.00,0.00,0.00)		0.000	0.010	0.005	0.010	0.010
(0.80,0.10,0.10,0.00)		0.002	0.040	0.004	0.012	0.012

Table 12. Comparison of type I error of different methods (significance level is 0.05)

表 12. 不同方法下一类错误的比较(显著性水平为 0.05)

子原假设比例	均值; 基因数	IUT	PLAO	DAT	MAIUP	New_method
(0.80,0.10,0.10,0.00)	4; 100	0.007	0.060	0.022	0.074	0.071
(0.04,0.48,0.48,0.00)		0.044	0.121	0.120	0.061	0.082
(0.34,0.33,0.33,0.00)		0.029	0.094	0.075	0.066	0.086
(1.00,0.00,0.00,0.00)		0.002	0.047	0.036	0.058	0.049
(0.80,0.10,0.10,0.00)	4; 1000	0.011	0.060	0.025	0.056	0.066
(0.04,0.48,0.48,0.00)		0.047	0.122	0.122	0.052	0.075
(0.34,0.33,0.33,0.00)		0.033	0.095	0.077	0.053	0.071
(1.00,0.00,0.00,0.00)		0.002	0.050	0.036	0.051	0.050
(0.80,0.10,0.10,0.00)	10; 100	0.007	0.076	0.022	0.073	0.071
(0.04,0.48,0.48,0.00)		0.044	0.147	0.125	0.061	0.083
(0.34,0.33,0.33,0.00)		0.030	0.116	0.078	0.066	0.088
(1.00,0.00,0.00,0.00)		0.002	0.047	0.036	0.058	0.049
(0.80,0.10,0.10,0.00)	10; 1000	0.012	0.075	0.025	0.056	0.066
(0.04,0.48,0.48,0.00)		0.048	0.149	0.126	0.052	0.077
(0.34,0.33,0.33,0.00)		0.033	0.119	0.080	0.053	0.073
(1.00,0.00,0.00,0.00)		0.002	0.050	0.036	0.051	0.050
(0.80,0.10,0.10,0.00)	20; 100	0.007	0.080	0.022	0.073	0.071
(0.04,0.48,0.48,0.00)		0.044	0.152	0.125	0.061	0.087
(0.34,0.33,0.33,0.00)		0.030	0.121	0.078	0.066	0.088
(1.00,0.00,0.00,0.00)		0.002	0.047	0.036	0.058	0.049
(0.80,0.10,0.10,0.00)	20; 1000	0.012	0.079	0.025	0.056	0.066
(0.04,0.48,0.48,0.00)		0.048	0.155	0.126	0.052	0.078
(0.34,0.33,0.33,0.00)		0.033	0.123	0.080	0.053	0.793
(1.00,0.00,0.00,0.00)		0.002	0.050	0.036	0.051	0.050

我们发现在两个性状都与同一基因都无关这一情况成立比例较高时, 不管均值多大 IUT 方法都十分的保守, 同时 DACT 方法相对来说也比较保守; 但当两个性状都与同一基因都无关这一情况成立比例较低时, IUT 方法能够控制一类错误, 而 DACT 方法出现膨胀现象。而 PLACO 方法只有在三个子假设比例分别为 1、0、0 时能够控制住一类错误, 其余情况出现膨胀状况。新构建的 p 值、MAIUP 方法在三个子假设比例分别为 1、0、0 时虽然能够控制一类错误, 但 MAIUP 方法相对来说有点膨胀, 但新构建的 p 值的方法在此种情况下表现更好

为了更加直观地进行对比, 绘制了在均值为 10、20, 样本量为 1000、2000 情况下三个子假设比例分别为 1、0、0; 0.8、0.1、0.1; 0.04、0.48、0.48 这三种情形下在显著性水平为 0.01 的一类错误柱状图。根据图 1 可以更直观的理解上述的一类错误数据规律。

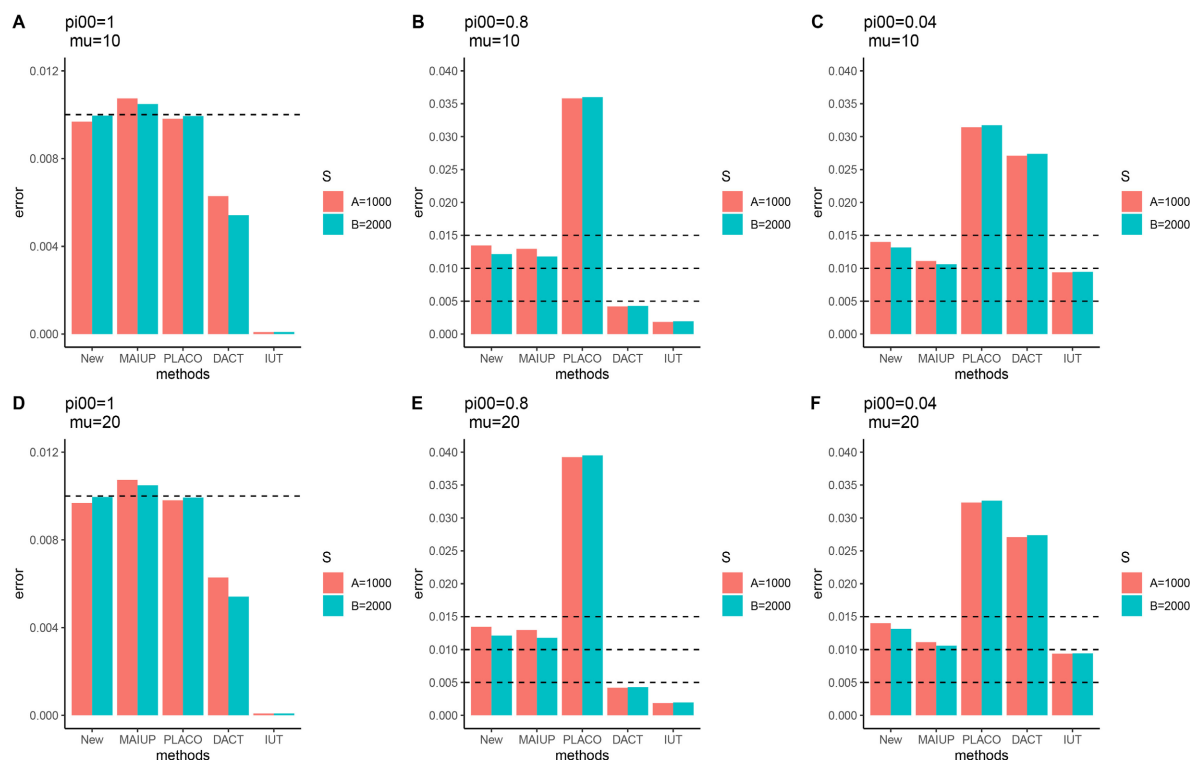


Figure 1. Comparison of type I error (significance level is 0.01)

图 1. 一类错误比较(显著性水平为 0.01)

上述结果是在两个性状独立时所得, 现在比较分析两性状存在相关关系时一类错误在去相关性(图 2 中 C、D)和不去相关性(图 2 中 A、B)大小比较如图 2 所见, 未去相关性时的一类错误随着相关性的加强而变大, 此时 PLACO、新构建的 p 值、MAIUP 方法尤为明显, 在去完相关性后, PLACO 方法在检验公式(2.3)中三个子假设比例分别为 0.8、0.1、0.1 情形下去完相关性后的一类错误有些膨胀(如图 2 中 C 所见)。而新构建的 p 值、MAIUP 方法能够较好地维持一类错误, 但新构建的 p 值的方法在 D 情形下更为稳定。

4.4. 功效

功效是假设检验中另一重要衡量目标, 在了解了相关一类错误后, 下面我们模拟了不同功效水平不同均值以及基因数下的各个方法的功效大小, 如表 13、表 14。

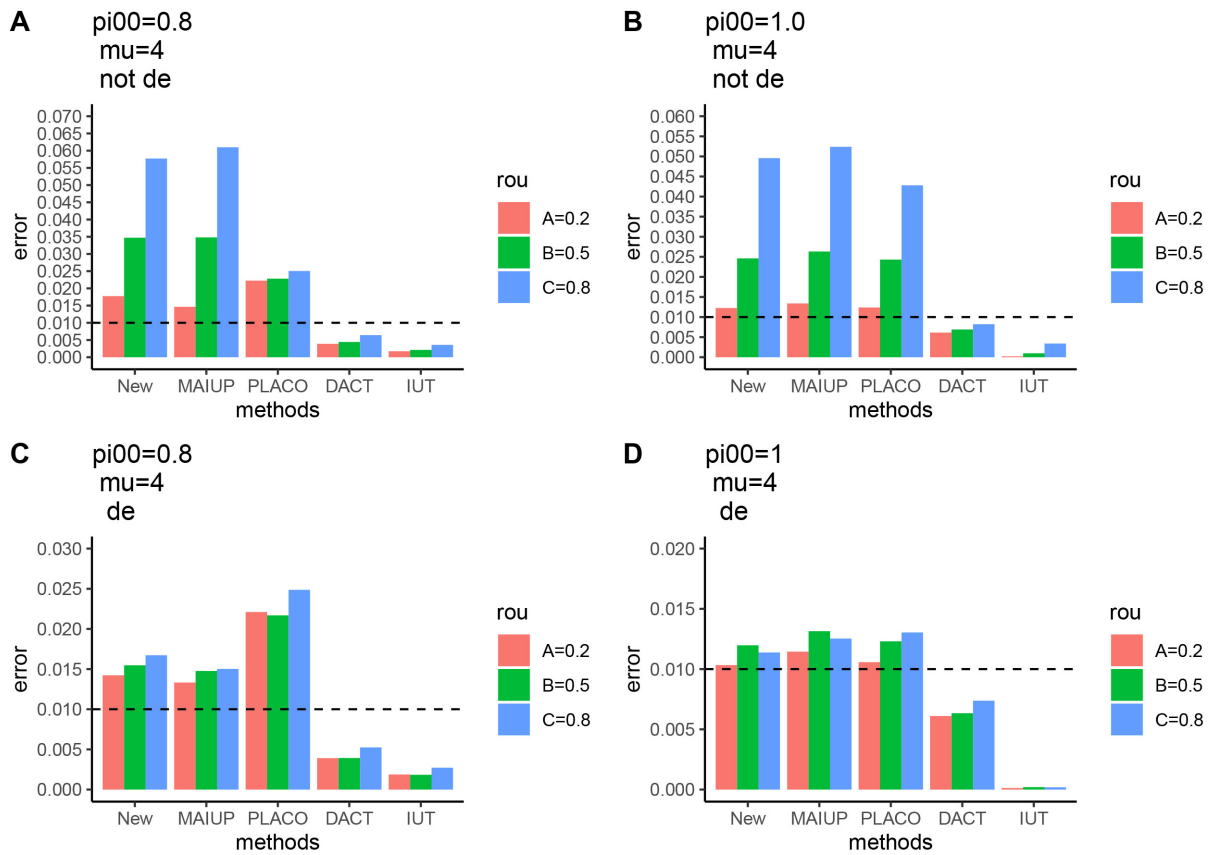


Figure 2. Comparison of type I error after decorrelation (significance level is 0.01)

图 2. 一类错误去相关性对比(显著性水平为 0.01)

Table 13. Comparison of power of different methods (significance level is 0.01)

表 13. 不同方法下一类功效的比较(显著性水平为 0.01)

子原假设比例	均值; 基因数	IUT	PLACO	DACT	MAIUP	New_method
(0.04,0.05,0.05,0.86)	10、100	0.860	0.878	0.860	0.874	0.876
(0.04,0.03,0.03,0.90)		0.900	0.900	0.900	0.909	0.910
(0.04,0.05,0.05,0.86)	10、1000	0.861	0.880	0.867	0.871	0.878
(0.04,0.03,0.03,0.90)		0.900	0.915	0.904	0.911	0.917
(0.04,0.05,0.05,0.86)	10、2000	0.860	0.880	0.867	0.870	0.878
(0.04,0.03,0.03,0.90)		0.900	0.915	0.904	0.911	0.917
(0.04,0.05,0.05,0.86)	20、100	0.860	0.880	0.860	0.864	0.873
(0.04,0.03,0.03,0.90)		0.900	0.900	0.900	0.929	0.920
(0.04,0.05,0.05,0.86)	20、1000	0.861	0.882	0.867	0.871	0.878
(0.04,0.03,0.03,0.90)		0.900	0.916	0.904	0.911	0.917
(0.04,0.05,0.05,0.86)	20、2000	0.861	0.882	0.867	0.870	0.878
(0.04,0.03,0.03,0.90)		0.900	0.916	0.904	0.911	0.917

Table 14. Comparison of power of different methods (significance level is 0.05)
表 14. 不同方法下一类功效的比较(显著性水平为 0.05)

子原假设比例	均值; 基因数	IUT	PLACO	DACT	MAIUP	New_method
(0.04,0.05,0.05,0.86)	10、100	0.860	0.902	0.878	0.908	0.906
(0.04,0.03,0.03,0.90)		0.900	0.932	0.900	0.905	0.905
(0.04,0.05,0.05,0.86)	10、1000	0.865	0.900	0.881	0.921	0.937
(0.04,0.03,0.03,0.90)		0.903	0.927	0.913	0.944	0.968
(0.04,0.05,0.05,0.86)	10、2000	0.865	0.900	0.881	0.944	0.937
(0.04,0.03,0.03,0.90)		0.903	0.927	0.914	0.965	0.968
(0.04,0.05,0.05,0.86)	20、100	0.860	0.904	0.878	0.908	0.906
(0.04,0.03,0.03,0.90)		0.900	0.934	0.900	0.915	0.935
(0.04,0.05,0.05,0.86)	20、1000	0.865	0.902	0.881	0.931	0.937
(0.04,0.03,0.03,0.90)		0.903	0.928	0.913	0.964	0.968
(0.04,0.05,0.05,0.86)	20、2000	0.865	0.902	0.881	0.934	0.937
(0.04,0.03,0.03,0.90)		0.903	0.928	0.914	0.965	0.968

通过表 13、表 14 发现在备择假设成立的概率越高时, 各个方法的功效水平会有所提升, 这符合常理。同时发现 MAIUP 以及 New_method 所得到的功效水平更高。

同时通过设置备择假设成立的概率在 0.7 时, 三个子假设成立的比例分别为 0.1、0.1、0.1 的情形, 比较在水平为 0.01 下的不同均值以及样本量间功效如图 3, 我们发现在均值较小的时候 IUT 以及 DACT 方法功效较差, 同时可见 New_method 以及 MAIUP 方法能够维持一个稳定功效水平。

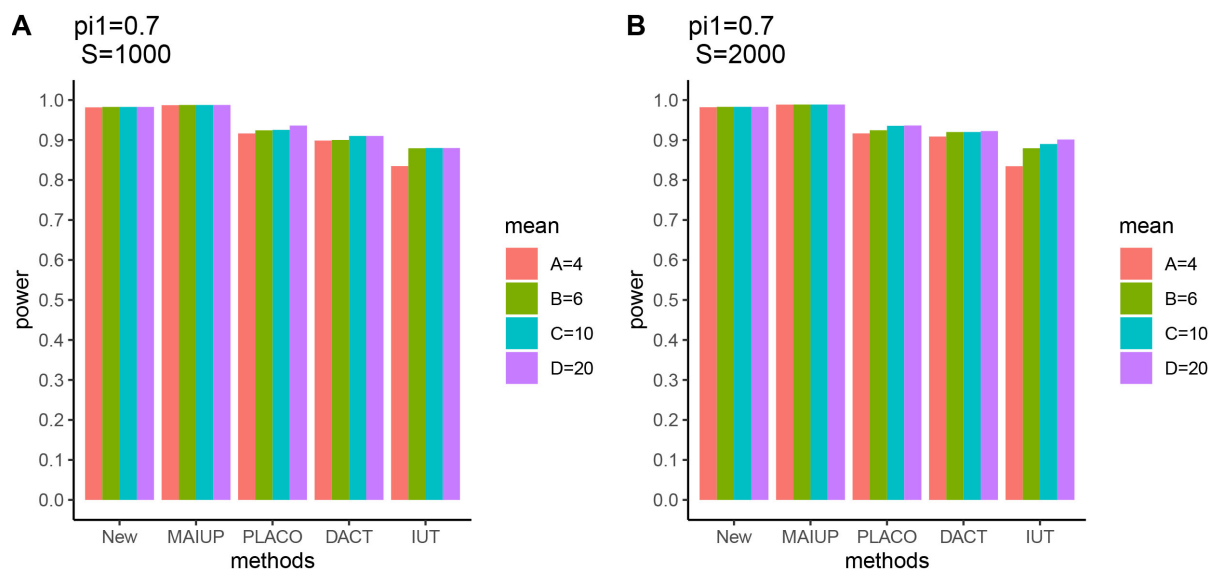


Figure 3. Comparison of power (significance level is 0.01)
图 3. 功效对比(显著性水平为 0.01)

同时将上述方法应用于去相关性的功效中, 在备择假设成立的比例为 0.7, 三个子假设成立的比例分别为 0.1、0.1、0.1, 均值为 4 的情形下进行对比(如图 4, A 表示未去相关性、B 表示去相关性)。两性状

在有相关性时去相关性前后功效进行了对比, IUT、DACT 方法会受相关性的影响, 去相关性前后的差距较大。而 PLACO、MAIUP 方法和新构建的 p 值的方法则受影响较小, 但后两者的功效较高, 效果较好。

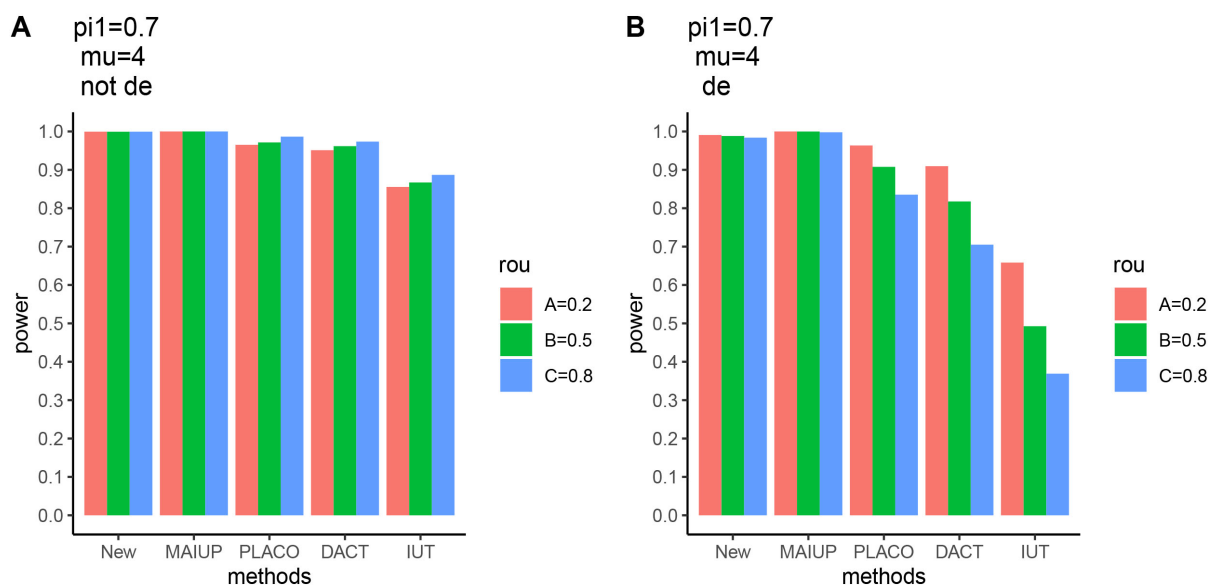


Figure 4. Comparison of power after decorrelation (significance level is 0.01)

图 4. 去相关性功效对比(显著性水平为 0.01)

4.5. 实际应用

多效性的显著例子在许多疾病和性状中广泛存在, 如自身免疫性疾病[20]和精神疾病[21]。除此之外多效性也存在于看似不相关的性状中。例如哮喘和精神疾病[22]、前列腺癌和血脂[23], 以及冠状动脉疾病和扁桃体切除术[24]。本文将上述五种方法应用于冠状动脉疾病以及扁桃体切除的实际数据中, 检验多

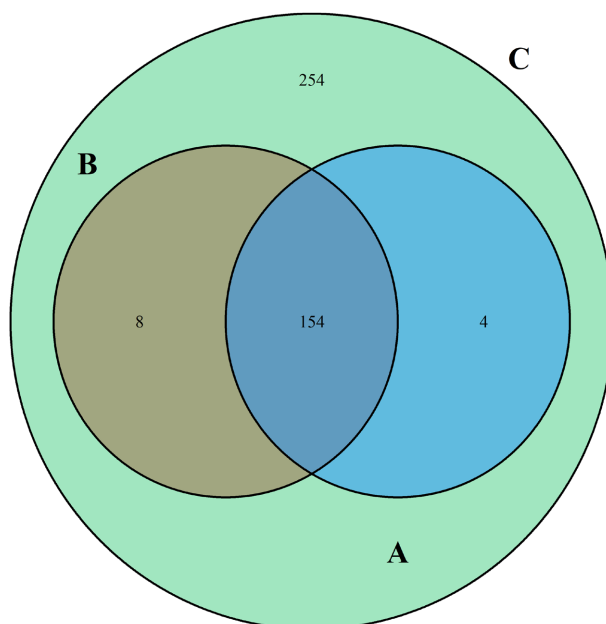


Figure 5. Venn diagram of detected genes of different methods

图 5. 不同方法检测出基因的韦恩图

效性基因的个数。从网站(<https://grasp.nhlbi.nih.gov/FullResults.aspx>)上查找相关的数据并进行下载并随机抽取两千个相关的基因进行多效性检验。由于 DACT 方法的不及新构建的 p 值的方法, 而 MAIUP 方法的结果和新构建的 p 值的方法结果相差无几, 因此使用 IUT、PLACO、新构建的 p 值的三种方法来检验多效性基因的个数, 分别检测到 158、162、420 个。并绘制韦恩图如图 5。

其中 A 代表 IUT 方法、B 代表 PLACO 方法、C 代表新构建的 p 值方法。可以看到新构建的 p 值的方法能够额外检测到更多多效性基因, 也许可以为研究提供一定参考价值。

5. 结论

本文在此次检验多效性的研究中, 以两性状为例, 从互斥检验角度出发, 介绍了 IUT、PLACO、DACT、MAIUP 方法, 同时在 DACT 方法的基础上借助 MAIUP 方法来修正 DACT 方法即新构造的 p 值。

通过模拟研究发现, MAIUP 方法无论是在维持一类错误还是保持功效水平等方面, 都表现更好。IUT 方法在检验两性状与一个基因无关时一类错误出现保守现象; PLACO 方法在基因与性状之间关联的强度越大时, 即关联的信号越强时会不能很好的控制一类错误, 一类错误出现膨胀现象, 但对于功效能够较好的维持功效水平。因此可以在数据关联的信号强度较小时采用此方法进行检验。DACT 方法在检验两性状与一个基因无关这一情形成立概率较大时一类错误会出现保守的现象, 而提出的新方法能够较好地控制一类错误, 同时较未修正的 DACT 方法能够提高检验的功效水平, 所以表现良好。值得注意, 上述情形是基于两性状之间是独立的情形, 当两性状之间存在关联时, 首先采用去相关性, 然后再试用上述方法, 此时新方法、MAIUP 方法无论是在控制一类错误还是检验功效水平方面都有较好的效果, 在实际应用中若出现两性状关联的情形时, 检验多效性可以优先考虑新构造的 p 值的方法和 MAIUP 方法。

参考文献

- [1] 徐大胜. 遗传与作物育种[M]. 成都: 四川大学出版社, 2011: 48.
- [2] Chung, D., Yang, C., Li, C., Gelernter, J. and Zhao, H. (2014) GPA: A Statistical Approach to Prioritizing GWAS Results by Integrating Pleiotropy and Annotation. *PLOS Genetics*, **10**, e1004787. <https://doi.org/10.1371/journal.pgen.1004787>
- [3] Hackinger, S. and Zeggini, E. (2017) Statistical Methods to Detect Pleiotropy in Human Complex Traits. *Open Biology*, **7**, Article ID: 170125. <https://doi.org/10.1098/rsob.170125>
- [4] Maier, R.M., Zhu, Z., Lee, S.H., et al. (2018) Improving Genetic Prediction by Leveraging Genetic Correlations among Human Diseases and Traits. *Nature Communications*, **9**, Article No. 989. <https://doi.org/10.1038/s41467-017-02769-6>
- [5] Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M. and Smoller, J.W. (2013) Pleiotropy in Complex Traits: Challenges and Strategies. *Nature Reviews Genetics*, **14**, 483-495. <https://doi.org/10.1038/nrg3461>
- [6] Ferreira, M.A. and Purcell, S.M. (2009) A Multivariate Test of Association. *Bioinformatics*, **25**, 132-133. <https://doi.org/10.1093/bioinformatics/btn563>
- [7] Schaid, D.J., Tong, X., Larrabee, B., et al. (2016) Statistical Methods for Testing Genetic Pleiotropy. *Genetics*, **204**, 483-497. <https://doi.org/10.1534/genetics.116.189308>
- [8] Wang, Y., Wu, P., Tong, X. and Sun, J. (2021) A Weighted Method for the Exclusive Hypothesis Test with Application to Typhoon Data. *Canadian Journal of Statistics*, **49**, 1258-1272. <https://doi.org/10.1002/cjs.11618>
- [9] Ray, D. and Chatterjee, N. (2020) A Powerful Method for Pleiotropic Analysis under Composite Null Hypothesis Identifies Novel Shared Loci between Type 2 Diabetes and Prostate Cancer. *PLOS Genetics*, **16**, e1009218. <https://doi.org/10.1371/journal.pgen.1009218>
- [10] Liu, Z., Shen, J., Barfield, R., et al. (2022) Large-Scale Hypothesis Testing for Causal Mediation Effects with Applications in Genome-wide Epigenetic Studies. *Journal of the American Statistical Association*, **117**, 67-81. <https://doi.org/10.1080/01621459.2021.1914634>
- [11] Wang, T., Lu, H. and Zeng, P. (2021) Identifying Pleiotropic Genes for Complex Phenotypes with Summary Statistics From a Perspective of Composite Null Hypothesis Testing. *Briefings in Bioinformatics*, **23**, Article ID: Bbab389. <https://doi.org/10.1093/bib/bbab389>

-
- [12] Broadaway, K.A., Cutler, D.J., Duncan, R., Moore, J.L., Ware, E.B., Jhun, M.A., *et al.* (2016) A Statistical Approach for Testing Cross-Phenotype Effects of Rare Variants. *The American Journal of Human Genetics*, **98**, 525-540. <https://doi.org/10.1016/j.ajhg.2016.01.017>
- [13] Barfield, R., Shen, J., Just, A.C., *et al.* (2017) Testing for the Indirect Effect under the Null for Genome-Wide Mediation Analyses. *Genetic Epidemiology*, **41**, 824-833. <https://doi.org/10.1002/gepi.22084>
- [14] Craig, C.C. (1936) On the Frequency Function of xy . *Annals of Mathematical Statistics*, **7**, 1-15. <https://doi.org/10.1214/aoms/1177732541>
- [15] Huang, Y-T. (2019) Genome-Wide Analyses of Sparse Mediation Effects under Composite Null Hypotheses. *Annals of Applied Statistics*, **13**, 60-84. <https://doi.org/10.1214/18-AOAS1181>
- [16] Efron, B., Tibshirani, R., Storey, J.D. and Tusher, V. (2001) Empirical Bayes Analysis of a Microarray Experiment. *Journal of the American Statistical Association*, **96**, 1151-1160. <https://doi.org/10.1198/016214501753382129>
- [17] Jin, J. and Cai, T.T. (2007) Estimating the Null and the Proportion of Nonnull Effects in Large-Scale Multiple Comparisons. *Journal of the American Statistical Association*, **102**, 495-506. <https://doi.org/10.1198/016214507000000167>
- [18] Dai, J.Y., Stanford, J.L. and LeBlanc, M. (2022) A Multiple-Testing Procedure for High-Dimensional Mediation Hypotheses. *Journal of the American Statistical Association*, **117**, 198-213. <https://doi.org/10.1080/01621459.2020.1765785>
- [19] Storey, J.D. and Tibshirani, R. (2003) Statistical Significance for Genomewide Studies. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 9440-9445. <https://doi.org/10.1073/pnas.1530509100>
- [20] Cotsapas, C., Voight, B.F., Rossin, E., *et al.* (2011) Pervasive Sharing of Genetic Effects in Autoimmune Disease. *PLOS Genetics*, **7**, e1002254. <https://doi.org/10.1371/journal.pgen.1002254>
- [21] Amare, A.T., Vaez, A., Hsu, Y.-H., *et al.* (2020) Bivariate Genome-Wide Association Analyses of the Broad Depression Phenotype Combined with Major Depressive Disorder, Bipolar Disorder or Schizophrenia Reveal Eight Novel Genetic Loci for Depression. *Molecular Psychiatry*, **25**, 1420-1429. <https://doi.org/10.1038/s41380-018-0336-6>
- [22] Zhu, Z., Zhu, X., Liu, C.-L., *et al.* (2019) Shared Genetics of Asthma and Mental Health Disorders: A Large-Scale Genome-Wide Cross-Trait Analysis. *European Respiratory Journal*, **54**, Article ID: 1901507. <https://doi.org/10.1183/13993003.01507-2019>
- [23] Andreassen, O.A., Zuber, V., Thompson, W.K., *et al.* (2014) Shared Common Variants in Prostate Cancer and Blood Lipids. *International Journal of Epidemiology*, **43**, 1205-1214. <https://doi.org/10.1093/ije/dyu090>
- [24] Pickrell, J., Berisa, T., Liu, J.Z., *et al.* (2016) Detection and Interpretation of Shared Genetic Influences on 42 Human Traits. *Nature Genetics*, **48**, 709-717. <https://doi.org/10.1038/ng.3570>