

# The Unknown Unequal Variance Model of Bivariate Meta-Analysis Based on the Difference Mean

Pengfei Guo<sup>1,2,3</sup>, Gang Li<sup>4\*</sup>, Xiangru Sun<sup>5</sup>

<sup>1</sup>College of Computational Science, Zhongkai University of Agriculture and Engineering, Guangzhou Guangdong

<sup>2</sup>Intelligent Agriculture Engineering Research Center of Guangdong Higher Education Institute, Zhongkai University of Agriculture and Engineering, Guangzhou Guangdong

<sup>3</sup>Guangdong Province Key Laboratory of Waterfowl Healthy Breeding, Zhongkai University of Agriculture and Engineering, Guangzhou Guangdong

<sup>4</sup>Guangzhou E-Government Center, Guangzhou Guangdong

<sup>5</sup>Institute of Reproductive Medicine, Panyu Hexian Memorial Hospital of Guangzhou, Guangzhou Guangdong

Email: 450799063@qq.com, \*24715926@qq.com

Received: Jan. 15<sup>th</sup>, 2019; accepted: Feb. 5<sup>th</sup>, 2019; published: Feb. 12<sup>th</sup>, 2019

---

## Abstract

This paper introduces the unknown unequal variance model of bivariate meta-analysis based on the difference mean; we obtain the maximum likelihood estimator of the difference mean effect and unknown variances. Then we find the combining overall difference mean effect  $D_{overall}$  by the fixed-effect model of bivariate Meta-analysis based on the difference mean. Moreover, we obtain the corresponding covariance matrix and the  $(1-\alpha)100\%$  confidence interval for the overall difference mean effect  $D_{overall}$ .

## Keywords

Meta-Analysis, The Unknown Unequal Variance Model, Maximum Likelihood Estimate

---

# 基于均差估计二变量Meta-分析的未知不相等方差模型

郭鹏飞<sup>1,2,3</sup>, 李刚<sup>4\*</sup>, 孙向茹<sup>5</sup>

<sup>1</sup>仲恺农业工程学院, 计算科学学院, 广东 广州

\*通讯作者。

<sup>2</sup>仲恺农业工程学院, 广东省高校智慧农业工程研究中心, 广东 广州

<sup>3</sup>仲恺农业工程学院, 广东省水禽健康养殖重点实验室, 广东 广州

<sup>4</sup>广州市电子政务服务中心, 广东 广州

<sup>5</sup>番禺何贤纪念医院, 生殖医学科, 广东 广州

Email: 450799063@qq.com, \*24715926@qq.com

收稿日期: 2019年1月15日; 录用日期: 2019年2月5日; 发布日期: 2019年2月12日

## 摘要

本文介绍基于均差估值的二变量Meta-分析的未知不相等方差模型, 通过极大似然估计法给出效应量及方差估计, 进而得到各研究均值效应量协方差矩阵的估计量。然后, 通过基于均差效应量二变量Meta-分析的固定效应模型给出了合并均差统计量  $D_{overall}$  的具体形式和权重。进而得到了  $D_{overall}$  的协方差矩阵和两个变量的  $(1-\alpha)100\%$  置信区间。

## 关键词

Meta-分析, 未知不相等方差模型, 极大似然估计

Copyright © 2019 by author(s) and Hans Publishers Inc.

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

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



Open Access

## 1. 引言

Meta-分析回答了一个非常广为接受的问题和设计。即什么是研究的目标? 在一个广泛的整体中研究的目标是否能够得到确认? 什么是研究结果的可操作定义: 试验样本组还是研究对象的总体? 什么样类型的设计在搜索过程中是需要包括的? 只有检验研究假设的随机试验可以作为纳入文献? 非试验的研究是否可以作为纳入文献? 上述这些问题的答案影响了一般综述的方法, 统计推断的模式和研究结果的解释。

Meta-分析是一种对同一问题的不同研究进行定量合并的一种方法。针对具体问题的 Meta-分析及其统计方法研究已持续了一个世纪之久。最早是在 1904 年 Karl Pearson 针对伤寒疫苗效果进行了合并研究 [1]; 但是在过去的四十年中, 人们逐渐意识到医学实验和临床操作需要基于整体的相关性和可靠证据, Meta-分析的影响得到极大的扩展 [2]。

尽管 Meta-分析在很多领域都有较好的应用, 但针对 Meta-分析本身的统计方法研究与其应用研究相比较少。从应用的意义上讲, Meta-分析是以估计量的标准误差为权重的点估计加权平均。通过选择 Meta-分析两个合理假设(该假设对于每一个研究作为独立先验信息存在)中的一个来合并公共定量信息估计, 从而选择合并共同定量信息的估计模型: 固定效应模型和随机效应模型。固定效应模型的 Meta-分析依赖于假设各个研究具有相等效应量, 具体是以估计量方差的逆作为权重给出的加权平均效应量 [3]。

在具体应用中, 建立统计模型解决上述挑战中的问题是非常重要的事情。例如: 在临床医学中, 目前对于流行病学的 Meta-分析研究主要集中在单变量 Meta-分析。由于流行病(例如: 非霍其因淋巴瘤, 乙

/丙肝等)本身的致病原因很复杂、致病原因之间具有相关性,且前期检查也是多指标检验,因此单变量 Meta-分析给出的效应量并不能完全地描述流行病的致病机理。由此,希望展开对流行病学的多变量 Meta-分析研究。多变量 Meta-分析有很多优点:第一,可以在一个模型框架下得到所有效应量的估计量;第二,可以通过所有效应量的估计量的协方差矩阵描述多效应量之间的关系。第三,我们可以获得具有更好统计特性的参数估计量;第四,可以获得区别于单变量 Meta-分析的潜在机理原因;第五,多变量分析的方法可以在一定程度上降低纳入文献的偏倚性。

由于多变量 Meta-分析的诸多优点,针对多变量 Meta-分析的理论研究也在不断深入。1988年, Raudenbush SW 等人为了研究教练在 SAT 中的作用,通过广义最小二乘法建立了多效应量合并的模型[4]; 1993年, van Houwelingen HC 等人在文献[5]中首次给出了基于比值比的二变量 Meta-分析模型,讨论了该模型的异质性检验和敏感性分析; 2002年, van Houwelingen HC 等人通过似然估计方法建立了在广义多元混合线性模型的框架下的多元 Meta-回归模型,并将此模型扩展到了非正态分布情形[6]。2008年, Riley RD 等人在文献[7]中通过极大似然估计法给出了一些特殊相关系数情形下的二变量 Meta-分析的协方差矩阵估计; 2008年, Ritz J 等人在协方差矩阵已知情形下,通过极大似然估计和估计方程给出了多元效应量回归参数,并将此模型应用到肺癌发病率的临床诊断中[8]。2010年, Paul M 等人通过可积嵌套拉普拉斯近似贝叶斯方法给出了多变量 Meta-分析的合并效应量估计,这种方法得到的方差估计偏移量更小且稳定[9]。

本文通过极大似然估计法给出效应量及方差估计且得到各研究均值效应量协方差矩阵的估计量;然后,通过基于均差效应量二变量 Meta-分析的固定效应模型给出了合并均差统计量  $D_{\text{overall}}$  的具体形式和权重。建立了基于均差估值的二变量 Meta-分析的未知不相等方差模型并给出了  $D_{\text{overall}}$  的协方差矩阵和两个变量的  $(1-\alpha)100\%$  置信区间。

## 2. 问题描述

在协方差矩阵的基于均差估值的二变量 Meta-分析的未知不相等方差模型中,假设个体量  $X_{ij}^C, X_{ij}^T, X_{i2j}^C$  和  $X_{i2j}^T$  是独立的并且是正态分布,其均值分别为  $\mu_1^C, \mu_1^T, \mu_2^C$  和  $\mu_2^T$ 。假设两个变量研究的方差分别为  $\sigma_C^2, \sigma_T^2, \tau_C^2, \tau_T^2$ 。因此  $\bar{X}_{i1}^C, \bar{X}_{i1}^T, \bar{X}_{i2}^C$  和  $\bar{X}_{i2}^T$  是独立的且服从正态分布,其均值分别为  $\mu_1^C, \mu_1^T, \mu_2^C$  和  $\mu_2^T$ , 方差分别为  $\frac{\sigma_C^2}{n_{i1}^C}, \frac{\sigma_T^2}{n_{i1}^T}, \frac{\tau_C^2}{n_{i2}^C}, \frac{\tau_T^2}{n_{i2}^T}$ 。假设两个变量之间的关系是相互独立的。

由上述假设,可以得到均差效应量  $D_i$  的分布为:

$$D_i = \begin{pmatrix} D_{i1} \\ D_{i2} \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_C^2 w_{i1}^C + \sigma_T^2 w_{i1}^T & 0 \\ 0 & \tau_C^2 w_{i2}^C + \tau_T^2 w_{i2}^T \end{pmatrix} \right)$$

其中  $\mu_1 = \mu_1^T - \mu_1^C, \mu_2 = \mu_2^T - \mu_2^C$  为均值,  $w_{i1}^T = \frac{1}{n_{i1}^T}, w_{i1}^C = \frac{1}{n_{i1}^C}, w_{i2}^T = \frac{1}{n_{i2}^T}, w_{i2}^C = \frac{1}{n_{i2}^C}$  是非随机的。一般地,  $\mu_1, \mu_2$  为未知参数。接下来通过极大似然估计法来估计  $\mu_1^C, \mu_1^T, \mu_2^C, \mu_2^T$ 。假设操作组和对照组是相互独立的,记  $S_i = S_i^C + S_i^T$ , 其中  $S_i^C = \begin{pmatrix} \sigma_C^2 w_{i1}^C & 0 \\ 0 & \tau_C^2 w_{i2}^C \end{pmatrix}, S_i^T = \begin{pmatrix} \sigma_T^2 w_{i1}^T & 0 \\ 0 & \tau_T^2 w_{i2}^T \end{pmatrix}$ 。

## 3. 均值效应量的极大似然估计

根据之前的假设效应量  $D_i$  服从正态分布且每个研究的操作组和对照组是独立的,因此关于  $\mu^C, \mu^T$ ,

$\sigma_c^2$ ,  $\sigma_T^2$ ,  $\tau_c^2$ ,  $\tau_T^2$  的似然函数为:

$$\begin{aligned} L(\boldsymbol{\mu}, \mathbf{D}_i, \mathbf{S}_i) &= L(\boldsymbol{\mu}^C, \bar{\mathbf{X}}_i^C, \mathbf{S}_i^C) L(\boldsymbol{\mu}^T, \bar{\mathbf{X}}_i^T, \mathbf{S}_i^T) \\ &= \prod_{i=1}^k \frac{1}{2\pi \det(\mathbf{S}_i^C)^{\frac{1}{2}}} \exp \left\{ -\frac{(\bar{\mathbf{X}}_i^C - \boldsymbol{\mu}^C)^T (\mathbf{S}_i^C)^{-1} (\bar{\mathbf{X}}_i^C - \boldsymbol{\mu}^C)}{2} \right\} \\ &\quad \cdot \prod_{i=1}^k \frac{1}{2\pi \det(\mathbf{S}_i^T)^{\frac{1}{2}}} \exp \left\{ -\frac{(\bar{\mathbf{X}}_i^T - \boldsymbol{\mu}^T)^T (\mathbf{S}_i^T)^{-1} (\bar{\mathbf{X}}_i^T - \boldsymbol{\mu}^T)}{2} \right\} \end{aligned}$$

其中  $\boldsymbol{\mu}^C = \begin{pmatrix} \mu_1^C \\ \mu_2^C \end{pmatrix}$ ,  $\boldsymbol{\mu}^T = \begin{pmatrix} \mu_1^T \\ \mu_2^T \end{pmatrix}$ ,  $\mathbf{S}_i^C = \begin{pmatrix} \sigma_c^2 w_{i1}^C & 0 \\ 0 & \tau_c^2 w_{i2}^C \end{pmatrix}$ ,  $\mathbf{S}_i^T = \begin{pmatrix} \sigma_T^2 w_{i1}^T & 0 \\ 0 & \tau_T^2 w_{i2}^T \end{pmatrix}$ 。由操作组和对照组的独立性, 可以分别计算其极大似然估计。操作组和对照组所对应的对数似然函数分别为:

$$\begin{aligned} L(\boldsymbol{\mu}^C, \bar{\mathbf{X}}_i^C, \mathbf{S}_i^C) &= -k \ln(2\pi) - \frac{1}{2} \sum_{i=1}^k \ln[\det(\mathbf{S}_i^C)] - \frac{1}{2} \sum_{i=1}^k \frac{(\bar{\mathbf{X}}_i^C - \boldsymbol{\mu}^C)^T (\mathbf{S}_i^C)^{-1} (\bar{\mathbf{X}}_i^C - \boldsymbol{\mu}^C)}{2}, \\ L(\boldsymbol{\mu}^T, \bar{\mathbf{X}}_i^T, \mathbf{S}_i^T) &= -k \ln(2\pi) - \frac{1}{2} \sum_{i=1}^k \ln[\det(\mathbf{S}_i^T)] - \frac{1}{2} \sum_{i=1}^k \frac{(\bar{\mathbf{X}}_i^T - \boldsymbol{\mu}^T)^T (\mathbf{S}_i^T)^{-1} (\bar{\mathbf{X}}_i^T - \boldsymbol{\mu}^T)}{2}. \end{aligned}$$

因为操作组和对照组所对应的对数似然函数与基于均差估计二变量 Meta-分析的未知相等方差模型中的似然函数相同, 因此通过类似的计算可以给出参数  $\boldsymbol{\mu}^C$ ,  $\boldsymbol{\mu}^T$ ,  $\sigma_c^2$ ,  $\sigma_T^2$ ,  $\tau_c^2$ ,  $\tau_T^2$  的极大似然估计量及其部分性质:

**定理 3.1:** 设效应量  $\mathbf{D}_i$  服从  $N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_c^2 w_{i1}^C + \sigma_T^2 w_{i1}^T & 0 \\ 0 & \tau_c^2 w_{i2}^C + \tau_T^2 w_{i2}^T \end{pmatrix}\right)$  的正态分布, 则

1)  $\boldsymbol{\mu}$  的极大似然估计量的分量为:

$$\hat{\mu}_1 = \hat{\mu}_1^T - \hat{\mu}_1^C = \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^T}{w_{i1}^T}}{\sum_{i=1}^k \frac{1}{w_{i1}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^C}{w_{i1}^C}}{\sum_{i=1}^k \frac{1}{w_{i1}^C}}, \quad \hat{\mu}_2 = \hat{\mu}_2^T - \hat{\mu}_2^C = \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^T}{w_{i2}^T}}{\sum_{i=1}^k \frac{1}{w_{i2}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^C}{w_{i2}^C}}{\sum_{i=1}^k \frac{1}{w_{i2}^C}}.$$

2)  $\sigma^2$ ,  $\tau^2$  的极大似然估计量分别为:

$$\begin{aligned} \hat{\sigma}_{CML}^2 &= \frac{1}{k} \sum_{i=1}^k \frac{(\bar{X}_{i1}^C - \hat{\mu}_1^C)^2}{w_{i1}^C}, \quad \hat{\tau}_{CML}^2 = \frac{1}{k} \sum_{i=1}^k \frac{(\bar{X}_{i2}^C - \hat{\mu}_2^C)^2}{w_{i2}^C}; \\ \hat{\sigma}_{TML}^2 &= \frac{1}{k} \sum_{i=1}^k \frac{(\bar{X}_{i1}^T - \hat{\mu}_1^T)^2}{w_{i1}^T}, \quad \hat{\tau}_{TML}^2 = \frac{1}{k} \sum_{i=1}^k \frac{(\bar{X}_{i2}^T - \hat{\mu}_2^T)^2}{w_{i2}^T}. \end{aligned}$$

由上述知, 效应量满足正态分布合并效应量均值估计量分量的具体形式, 可以得到合并效应量均值估计量的如下性质:

**性质 3.1:** 设效应量  $\mathbf{D}_i$  服从  $N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_c^2 w_{i1}^C + \sigma_T^2 w_{i1}^T & 0 \\ 0 & \tau_c^2 w_{i2}^C + \tau_T^2 w_{i2}^T \end{pmatrix}\right)$  的正态分布, 那么  $\boldsymbol{\mu}$  的极大似然估计量的分量为:

$$\hat{\mu}_1 = \hat{\mu}_1^T - \hat{\mu}_1^C = \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^T}{w_{i1}^T}}{\sum_{i=1}^k \frac{1}{w_{i1}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^C}{w_{i1}^C}}{\sum_{i=1}^k \frac{1}{w_{i1}^C}}, \quad \hat{\mu}_2 = \hat{\mu}_2^T - \hat{\mu}_2^C = \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^T}{w_{i2}^T}}{\sum_{i=1}^k \frac{1}{w_{i2}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^C}{w_{i2}^C}}{\sum_{i=1}^k \frac{1}{w_{i2}^C}}.$$

则  $\hat{\mu}_{ML}$  对于  $\mu$  是无偏的。

**证明:** 要证明估计量的无偏性, 需证等式  $E(\hat{\mu}_{ML}) = \mu$  成立。根据我们给出的效应量  $D_i$  服从正态分布的合并效应量均值估计量分量的具体形式, 我们需要证明  $E(\hat{\mu}_1^T - \hat{\mu}_1^C) = \mu_1^T - \mu_1^C$ ,  $E(\hat{\mu}_2^T - \hat{\mu}_2^C) = \mu_2^T - \mu_2^C$ 。因为  $E(\bar{X}_{i1}^C) = \mu_1^C$ ,  $E(\bar{X}_{i1}^T) = \mu_1^T$ ,  $E(\bar{X}_{i2}^C) = \mu_2^C$ ,  $E(\bar{X}_{i2}^T) = \mu_2^T$ , 所以

$$\begin{aligned} E(\hat{\mu}_1^T - \hat{\mu}_1^C) &= E\left\{ \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^T}{w_{i1}^T}}{\sum_{i=1}^k \frac{1}{w_{i1}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^C}{w_{i1}^C}}{\sum_{i=1}^k \frac{1}{w_{i1}^C}} \right\} = E\left\{ \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^T}{w_{i1}^T}}{\sum_{i=1}^k \frac{1}{w_{i1}^T}} \right\} - E\left\{ \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^C}{w_{i1}^C}}{\sum_{i=1}^k \frac{1}{w_{i1}^C}} \right\} \\ &= \frac{\sum_{i=1}^k \frac{E\{\bar{X}_{i1}^T\}}{w_{i1}^T}}{\sum_{i=1}^k \frac{1}{w_{i1}^T}} - \frac{\sum_{i=1}^k \frac{E\{\bar{X}_{i1}^C\}}{w_{i1}^C}}{\sum_{i=1}^k \frac{1}{w_{i1}^C}} = \frac{\sum_{i=1}^k \frac{\mu_1^T}{w_{i1}^T}}{\sum_{i=1}^k \frac{1}{w_{i1}^T}} - \frac{\sum_{i=1}^k \frac{\mu_1^C}{w_{i1}^C}}{\sum_{i=1}^k \frac{1}{w_{i1}^C}} \\ &= \mu_1^T - \mu_1^C \\ E(\hat{\mu}_2^T - \hat{\mu}_2^C) &= E\left\{ \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^T}{w_{i2}^T}}{\sum_{i=1}^k \frac{1}{w_{i2}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^C}{w_{i2}^C}}{\sum_{i=1}^k \frac{1}{w_{i2}^C}} \right\} = E\left\{ \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^T}{w_{i2}^T}}{\sum_{i=1}^k \frac{1}{w_{i2}^T}} \right\} - E\left\{ \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^C}{w_{i2}^C}}{\sum_{i=1}^k \frac{1}{w_{i2}^C}} \right\} \\ &= \frac{\sum_{i=1}^k \frac{E\{\bar{X}_{i2}^T\}}{w_{i2}^T}}{\sum_{i=1}^k \frac{1}{w_{i2}^T}} - \frac{\sum_{i=1}^k \frac{E\{\bar{X}_{i2}^C\}}{w_{i2}^C}}{\sum_{i=1}^k \frac{1}{w_{i2}^C}} = \frac{\sum_{i=1}^k \frac{\mu_2^T}{w_{i2}^T}}{\sum_{i=1}^k \frac{1}{w_{i2}^T}} - \frac{\sum_{i=1}^k \frac{\mu_2^C}{w_{i2}^C}}{\sum_{i=1}^k \frac{1}{w_{i2}^C}} \\ &= \mu_2^T - \mu_2^C \end{aligned}$$

综上所述,  $\hat{\mu}_{ML}$  对于  $\mu$  是无偏的。

#### 4. 均值效应量极大似然估计量的协方差矩阵

假设所有的研究都提供所有的均值效应。由多变量统计学的大数定理, 合并效应量均值估计量可以近似为一个多元正态分布, 其对应的协方差矩阵可由下面的定理给出:

**定理 4.1:** 设效应量  $D_i$  服从  $N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_C^2 w_{i1}^C + \sigma_T^2 w_{i1}^T & 0 \\ 0 & \tau_C^2 w_{i2}^C + \tau_T^2 w_{i2}^T \end{pmatrix}\right)$  的正态分布, 那么  $\mu$  的合并均值

效应量极大似然估计量为:

$$\hat{\mu}_1 = \hat{\mu}_1^T - \hat{\mu}_1^C = \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^T}{w_{i1}^T}}{\sum_{i=1}^k \frac{1}{w_{i1}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^C}{w_{i1}^C}}{\sum_{i=1}^k \frac{1}{w_{i1}^C}}, \quad \hat{\mu}_2 = \hat{\mu}_2^T - \hat{\mu}_2^C = \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^T}{w_{i2}^T}}{\sum_{i=1}^k \frac{1}{w_{i2}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^C}{w_{i2}^C}}{\sum_{i=1}^k \frac{1}{w_{i2}^C}}.$$

则  $\hat{\mu}_{ML}$  所对应的协方差矩阵  $Var(\hat{\mu}_{ML})$  为:

$$Var(\hat{\boldsymbol{\mu}}_{ML}) = \begin{pmatrix} \frac{\sigma_C^2}{\sum_{i=1}^k \frac{1}{w_{i1}^C}} + \frac{\sigma_T^2}{\sum_{i=1}^k \frac{1}{w_{i1}^T}} & 0 \\ 0 & \frac{\tau_C^2}{\sum_{i=1}^k \frac{1}{w_{i2}^C}} + \frac{\tau_T^2}{\sum_{i=1}^k \frac{1}{w_{i2}^T}} \end{pmatrix}$$

通过合并均值效应量极大似然估计量  $\hat{\boldsymbol{\mu}}_{ML}$  所对应的协方差矩阵  $Var(\hat{\boldsymbol{\mu}}_{ML})$  可以给出  $\hat{\boldsymbol{\mu}}_{ML}$  的如下性质:

**性质 4.1:** 设效应量  $D_i$  服从  $N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_C^2 w_{i1}^C + \sigma_T^2 w_{i1}^T & 0 \\ 0 & \tau_C^2 w_{i2}^C + \tau_T^2 w_{i2}^T \end{pmatrix}\right)$  的正态分布,  $\boldsymbol{\mu}$  的合并均值效应量极大似然估计量为:

$$\hat{\mu}_1 = \hat{\mu}_1^T - \hat{\mu}_1^C = \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^T}{w_{i1}^T}}{\sum_{i=1}^k \frac{1}{w_{i1}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^C}{w_{i1}^C}}{\sum_{i=1}^k \frac{1}{w_{i1}^C}}, \quad \hat{\mu}_2 = \hat{\mu}_2^T - \hat{\mu}_2^C = \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^T}{w_{i2}^T}}{\sum_{i=1}^k \frac{1}{w_{i2}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^C}{w_{i2}^C}}{\sum_{i=1}^k \frac{1}{w_{i2}^C}}.$$

则  $\hat{\boldsymbol{\mu}}_{ML}$  对于  $\boldsymbol{\mu}$  是一致的。

## 5. 整体均值合并统计量的估计

在上述研究中, 效应量用的是均差。因此可以给出均值效应量  $D_i$  的协方差估计为:

$$\hat{\boldsymbol{S}}_{iML} = \begin{pmatrix} \hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T & 0 \\ 0 & \hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T \end{pmatrix}.$$

在 Meta-分析中, 合并统计量是由具体研究表现统计量的加权平均给出的。根据均差估计二变量 Meta-分析的固定效应模型可知, 通过  $\hat{\boldsymbol{S}}_{iML}$  可以给出每一个研究的权重及均差合并统计量。具体如下:

1) 对于均值合并统计量, 每个研究均差的权重为:

$$\boldsymbol{v}_i = \left( \sum_{i=1}^k \hat{\boldsymbol{S}}_{iML}^{-1} \right)^{-1} \hat{\boldsymbol{S}}_{iML}^{-1} = \begin{pmatrix} \frac{1}{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T} & 0 \\ \frac{1}{\sum_{i=1}^k \frac{1}{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T}} & 0 \\ 0 & \frac{1}{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T} \\ 0 & \frac{1}{\sum_{i=1}^k \frac{1}{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T}} \end{pmatrix}$$

注意到权重  $\boldsymbol{v}_i$  只与个体研究的样本量有关并且是非随机的。

2) 整体均值合并统计量的估计量为:

$$\boldsymbol{D}_{overall} = \sum_{i=1}^k \boldsymbol{D}_i \boldsymbol{v}_i = \begin{pmatrix} \frac{\bar{X}_{i1}^T - \bar{X}_{i1}^C}{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T} \\ \frac{1}{\sum_{i=1}^k \frac{1}{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T}} \\ \frac{\bar{X}_{i2}^T - \bar{X}_{i2}^C}{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T} \\ \frac{1}{\sum_{i=1}^k \frac{1}{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T}} \end{pmatrix}$$

通过  $\mathbf{D}_{\text{overall}}$ ，可以得到  $\mathbf{D}_{\text{overall}}$  的协方差阵的估计量为：

$$\begin{aligned} \text{Var}(\mathbf{D}_{\text{overall}}) &= \begin{pmatrix} \frac{\sum_{i=1}^k \text{Var}(\bar{X}_{i1}^T - \bar{X}_{i1}^C)}{\left(\sum_{i=1}^k \frac{1}{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T}\right)^2} & 0 \\ 0 & \frac{\sum_{i=1}^k \text{Var}(\bar{X}_{i2}^T - \bar{X}_{i2}^C)}{\left(\sum_{i=1}^k \frac{1}{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T}\right)^2} \\ \frac{\sum_{i=1}^k \frac{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T}{\left(\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T\right)^2}}{\left(\sum_{i=1}^k \frac{1}{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T}\right)^2} & 0 \\ 0 & \frac{\sum_{i=1}^k \frac{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T}{\left(\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T\right)^2}}{\left(\sum_{i=1}^k \frac{1}{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T}\right)^2} \\ \frac{\sum_{i=1}^k \frac{1}{\left(\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T\right)}}{\left(\sum_{i=1}^k \frac{1}{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T}\right)^2} & 0 \\ 0 & \frac{\sum_{i=1}^k \frac{1}{\left(\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T\right)}}{\left(\sum_{i=1}^k \frac{1}{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T}\right)^2} \\ \frac{\sum_{i=1}^k \frac{1}{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T}}{\sum_{i=1}^k \frac{1}{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T}} & 0 \\ 0 & \frac{\sum_{i=1}^k \frac{1}{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T}}{\sum_{i=1}^k \frac{1}{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\sum_{i=1}^k \frac{1}{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T}}{\sum_{i=1}^k \frac{1}{\hat{\sigma}_{CML}^2 w_{i1}^C + \hat{\sigma}_{TML}^2 w_{i1}^T}} & 0 \\ 0 & \frac{\sum_{i=1}^k \frac{1}{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T}}{\sum_{i=1}^k \frac{1}{\hat{\tau}_{CML}^2 w_{i2}^C + \hat{\tau}_{TML}^2 w_{i2}^T}} \end{pmatrix} \end{aligned}$$

通过  $\mathbf{D}_{\text{overall}}$  所对应的协方差矩阵  $\text{Var}(\mathbf{D}_{\text{overall}})$  可以给出单变量和联合变量的置信区域，其具体形式为：

**性质 5.1:** 设效应量  $\mathbf{D}_i$  服从  $N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_C^2 w_{i1}^C + \sigma_T^2 w_{i1}^T & 0 \\ 0 & \tau_C^2 w_{i2}^C + \tau_T^2 w_{i2}^T \end{pmatrix}\right)$  的正态分布， $\boldsymbol{\mu}$  的合并均值效应量极大似然估计量为：

$$\hat{\mu}_1 = \hat{\mu}_1^T - \hat{\mu}_1^C = \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^T}{w_{i1}^T}}{\sum_{i=1}^k \frac{1}{w_{i1}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i1}^C}{w_{i1}^C}}{\sum_{i=1}^k \frac{1}{w_{i1}^C}}, \quad \hat{\mu}_2 = \hat{\mu}_2^T - \hat{\mu}_2^C = \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^T}{w_{i2}^T}}{\sum_{i=1}^k \frac{1}{w_{i2}^T}} - \frac{\sum_{i=1}^k \frac{\bar{X}_{i2}^C}{w_{i2}^C}}{\sum_{i=1}^k \frac{1}{w_{i2}^C}}.$$



则合并均值效应量  $\hat{\mu}_{ML}$  所对应各个分量的  $(1-\alpha)100\%$  置信区间为:

$$\left( \mu_1 - Z_{\frac{\alpha}{2}} \sqrt{\text{Var}(\mathbf{D}_{\text{overall}})_{(1,1)}}, \mu_1 + Z_{\frac{\alpha}{2}} \sqrt{\text{Var}(\mathbf{D}_{\text{overall}})_{(1,1)}} \right) \text{ 和 } \left( \mu_2 - Z_{\frac{\alpha}{2}} \sqrt{\text{Var}(\mathbf{D}_{\text{overall}})_{(2,2)}}, \mu_2 + Z_{\frac{\alpha}{2}} \sqrt{\text{Var}(\mathbf{D}_{\text{overall}})_{(2,2)}} \right),$$

其中  $Z_{\frac{\alpha}{2}}$  表示正态分布的  $\frac{\alpha}{2}$  分位数,  $\text{Var}(\mathbf{D}_{\text{overall}})_{(i,j)}$  表示矩阵  $\text{Var}(\mathbf{D}_{\text{overall}})$  第  $i$  行第  $j$  列的分量。

## 6. 结束语

本文通过极大似然估计法给出效应量及方差估计,进而得到各研究均值效应量协方差矩阵的估计量。然后,由基于均差效应量二变量 Meta-分析的固定效应模型给出了合并均差统计量  $\mathbf{D}_{\text{overall}}$  的具体形式和权重,构建了基于均差估值的二变量 Meta-分析的未知不相等方差模型,进而得到了  $\mathbf{D}_{\text{overall}}$  的协方差矩阵和两个变量的  $(1-\alpha)100\%$  置信区间。本文的结果对流行病学的研究有着重要的统计学意义。

## 基金项目

本项目由如下基金支持:国家自然科学基金(61871475, 61471133, 61571444, 61473331),广东省科技计划(2017B010126001, 2017A070712019, 2016A040402043, 2015A070709015, 2015A020209171, 2016B010125004, 2014B040404070, 2015A040405014, 2016A070712020),广东省教育厅科技计划(2017GCZX001, 2016GCZX001, 2017KTSCX094, 2017KTSCX095, 2017KQNCX098),广州市科技计划(201707010221)。

## 参考文献

- [1] Pearson, K. (1904) Report on Certain Enteric Fever Inoculation Statistics. *British Medical Journal*, **3**, 1243-1246.
- [2] Lee, W.L., Bausell, R.B. and Berman, B.M. (2001) The Growth of Health-Related Meta-Analyses Published from 1980 to 2000. *Evaluation and the Health Professions*, **24**, 327-335. <https://doi.org/10.1177/01632780122034948>
- [3] Stijnen, T. (1999) Tutorial in Biostatistics. Meta-Analysis: Formulating, Evaluating, Combining, and Reporting by S-L. Normand. *Statistics in Medicine*, **18**, 321-359.
- [4] Raudenbush, S.W., Becker, B.J. and Kalaian, H. (1988) Modeling Multivariate Effect Sizes. *Psychological Bulletin*, **103**, 111-120. <https://doi.org/10.1037/0033-2909.103.1.111>
- [5] van Houwelingen, H.C., Zwinderman, K.H. and Stijnen, T. (1993) A Bivariate Approach to Meta-Analysis. *Statistics in Medicine*, **12**, 2273-2284. <https://doi.org/10.1002/sim.4780122405>
- [6] van Houwelingen, H.C., Arends, L.R. and Stijnen, T. (2002) Advanced Methods in Meta-Analysis: Multivariate Approach and Metaregression. *Statistics in Medicine*, **21**, 589-624. <https://doi.org/10.1002/sim.1040>
- [7] Riley, R.D., Abrams, K.R., Lambert, P.C., Sutton, A.J. and Thompson, J.R. (2007) Bivariate Random Effects Meta-Analysis and the Estimation of Between-Study Correlation. *BMC Medical Research Methodology*, **7**, 3. <https://doi.org/10.1186/1471-2288-7-3>
- [8] Ritz, J., Demidenko, E. and Spiegelman, D. (2008) Multivariate Meta-Analysis for Data Consortia, Individual Patient Meta-Analysis, and Pooling Projects. *Journal of Statistical Planning and Inference*, **138**, 1919-1933. <https://doi.org/10.1016/j.jspi.2007.07.004>
- [9] Paul, M., Riebler, A., Bachmann, L.M., Rue, H. and Held, L. (2010) Bayesian Bivariate Meta-Analysis of Diagnostic Test Studies Using Integrated Nested Laplace Approximations. *Statistics in Medicine*, **29**, 1325-1339. <https://doi.org/10.1002/sim.3858>



**知网检索的两种方式：**

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择：[ISSN]，输入期刊 ISSN：2163-1476，即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：[orf@hanspub.org](mailto:orf@hanspub.org)