常见的异质性估计方法及代码示例

陈奕昕 3227042017

2025-05-20

参数 被定义为效应量正式值大小的方差。换句话说，如果我们有一个无限大的研究样本，当样本量足够大时，每个研究的估计值都可以近似看做真实值），并计算出真实效应大小的方差,那么这个方差就是真实效应的方差, 这个方差就是 。

由于我们无法观察到真实效应量，因此无法直接计算出这一方差。但是我们可以根据观测值对其进行估计，估计值记为为了得出这个估计值，我们从差值 开始，表示真实效应的离散度。我们除以一个量，这样做的效果是把测量结果放回到原来的度量中，并使其成为一个平均值、而不是偏差平方的总和。

### 一、常见估计方法

|  |  |  |
| --- | --- | --- |
| 方法 | 原理 | 适用场景 |
| **DerSimonian-Laird (DL)** | 基于矩估计法，通过Q统计量与自由度之差计算 | 传统方法，计算速度快，适用于小样本 |
| **REML** | 限制性最大似然估计，避免固定效应参数估计的偏差 | 推荐默认方法，平衡精度与计算效率 |
| **Maximum Likelihood (ML)** | 标准最大似然估计，但可能低估 | 需大样本支持，理论分析为主 |
| **Paule-Mandel (PM)** | 基于加权残差平方和最小化，无需分布假设 | 鲁棒性强，适用于复杂异质性结构 |
| **Sidik-Jonkman (SJ)** | 基于样本方差的加权平均，假设异质性方差与研究内方差成比例 | 异质性较高时表现稳定 |
| **Hedges (HE)** | 改进的矩估计法，修正小样本偏差 | 小样本Meta分析 |

### 二、R语言实现示例

#### **1. 数据准备**

假设研究数据包含效应量（TE）及其方差（seTE）：

# 示例数据（效应量TE和标准误seTE）  
data <- read.csv("作业.csv")  
# 计算合并标准差和Cohen's d  
data$S\_pooled <- sqrt(  
 ((data$Total\_T - 1) \* data$SD\_T^2 + (data$Total\_C - 1) \* data$SD\_C^2) /   
 (data$Total\_T + data$Total\_C - 2))  
data$d <- (data$Mean\_T - data$Mean\_C) / data$S\_pooled  
   
# 计算方差  
data$var\_d <- (data$Total\_T + data$Total\_C) / (data$Total\_T \* data$Total\_C) +   
 data$d^2 / (2 \* (data$Total\_T + data$Total\_C))

#### **2. 使用meta包估计**

# 安装并加载meta包  
#install.packages("meta")  
library(meta)

## Loading required package: metadat

## Loading 'meta' package (version 8.0-2).  
## Type 'help(meta)' for a brief overview.

# 使用不同方法估计τ²  
TE <- data$d  
seTE <- sqrt(data$var\_d)  
meta\_DL <- metagen(TE = TE, seTE = seTE, data = data, method.tau = "DL")  
meta\_REML <- metagen(TE = TE, seTE = seTE, data = data, method.tau = "REML")  
meta\_ML <- metagen(TE = TE, seTE = seTE, data = data, method.tau = "ML")  
meta\_PM <- metagen(TE = TE, seTE = seTE, data = data, method.tau = "PM")  
meta\_SJ <- metagen(TE = TE, seTE = seTE, data = data, method.tau = "SJ")  
meta\_HE <- metagen(TE = TE, seTE = seTE, data = data, method.tau = "HE")  
  
# 查看结果（以REML为例）  
summary(meta\_REML)

## 95%-CI %W(common) %W(random)  
## 1 0.0951 [-0.2629; 0.4532] 11.5 15.6  
## 2 0.2790 [-0.0665; 0.6245] 12.3 16.1  
## 3 0.3701 [-0.0719; 0.8121] 7.5 12.4  
## 4 0.6732 [ 0.4894; 0.8569] 43.6 24.3  
## 5 0.4656 [ 0.0574; 0.8737] 8.8 13.6  
## 6 0.1860 [-0.1153; 0.4873] 16.2 18.1  
##   
## Number of studies: k = 6  
##   
## 95%-CI z p-value  
## Common effect model 0.4380 [0.3167; 0.5594] 7.08 < 0.0001  
## Random effects model 0.3658 [0.1636; 0.5680] 3.55 0.0004  
##   
## Quantifying heterogeneity (with 95%-CIs):  
## tau^2 = 0.0351 [0.0007; 0.2385]; tau = 0.1873 [0.0258; 0.4884]  
## I^2 = 62.8% [9.6%; 84.7%]; H = 1.64 [1.05; 2.55]  
##   
## Test of heterogeneity:  
## Q d.f. p-value  
## 13.43 5 0.0197  
##   
## Details of meta-analysis methods:  
## - Inverse variance method  
## - Restricted maximum-likelihood estimator for tau^2  
## - Q-Profile method for confidence interval of tau^2 and tau  
## - Calculation of I^2 based on Q

#### **3. 输出解读**

* ：异质性方差估计值（tau^2 = 0.0351 ）。
* ：异质性标准差（如tau = 0.1873）。
* **I²**：异质性比例（如I^2 = 62.8%）。
* **Q检验**：异质性检验的p值（如0.0197）。

### 三、不同方法对比示例

# 提取各方法的τ²估计值  
tau2\_values <- c(  
 DL = meta\_DL$tau2,  
 REML = meta\_REML$tau2,  
 ML = meta\_ML$tau2,  
 PM = meta\_PM$tau2,  
 SJ = meta\_SJ$tau2,  
 HE = meta\_HE$tau2  
)  
  
# 打印结果  
print(tau2\_values)

## DL REML ML PM SJ HE   
## 0.04352709 0.03508292 0.02744647 0.02490810 0.02836434 0.01114134

### 四、方法选择建议

1. **默认推荐**：使用**REML**（平衡精度与计算效率）。
2. **小样本**：优先选择**DL**或**HE**（计算快，避免过拟合）。
3. **高异质性**：尝试**PM**或**SJ**（对复杂异质性更鲁棒）。

### 五、注意事项

1. **置信区间**：使用meta包的metagen函数时，可通过hakn = TRUE启用Hartung-Knapp调整，提高小样本置信区间准确性。
2. **结果报告**：始终报告、I²和Q检验结果，如：  
   “采用REML法估计异质性方差（），I²=65%，提示高度异质性（Q=12.5, p=0.03）。”

通过上述代码，可灵活选择异质性估计方法，为Meta分析提供统计学支持。