

文章编号: 1008-181X (1999) 04-0295-05

Meta 分析及 MetaWin 软件

彭少麟, 郑凤英

(中国科学院华南植物研究所, 广州 510650)

摘要: 综述 Meta 分析及 MetaWin 软件。MetaWin 是第一个专为生态学者设计的定量 Meta 分析软件, 它提供了两种假设模型、多种可选择的结合统计量, 不仅操作简单, 而且提出了一种新的非参数检验法——重取样检验。

关键词: Meta 分析; MetaWin 软件; 综述

中图分类号: Q141 **文献标识码:** A

生态学的发展已经由定性观察研究向定量实验研究迈进, 从本世纪初的竞争、捕食等较粗放的实验过度到了室内模拟的 CO₂ 倍增等较为精细的实验。

Scheiner (1993) 曾提出这样一个基本的问题: 生态学实验的目的何在? 通常, 实验的目的在于验证一个或一类特定的假说。但事实上, 一个实验的结果仅局限于某一特定时间、地点和物种, 显然单个实验无法对一个假说进行验证^[1]。Gause 的竞争排斥理论的提出仅用他本人的草履虫实验是无法证明的。应用生态学与理论生态学对各自实验结果的应用有所不同, 对于前者, 一定条件下的实验结果只要有利于其生态系统的管理便为有用, 但对于后者, 如果从独立实验中提炼不出综合性内容便无多大用处。那么一个独立实验结果究竟能对大自然揭示多少呢? 从一个实验结果中我们又能外推多少呢? 它是对以前研究结果的支持还是否定呢? 在什么样的条件下这些独立实验能够得出相同的或相互矛盾的结果呢? 如何回答这些问题, 生态学家们持有不同的见解。

虽然在生态学中绝对重复实验是不可能的, 但综合一类实验结果以便于找出一般性结论也同样重要, 也即综述是非常必要的。传统的描述性综述只对独立实验结果作一个简单罗列, 解释过程中往往添加了过多的综述者主观意见, 没有起到综合、提炼的作用。数表决法把综述引入量化的门槛, 却最终因为方法有失科学性而遭到遗弃。正在生态学家为不能找出一个较好的方法来解决上述问题困扰时, Meta 分析出现在他们的面前。

1 Meta 分析简介

Meta 分析是一种对同一主题下的多个独立实验 (研究) 进行综合的统计分析方法。它萌芽于本世纪初^[2]; 1976 年由美国教育学家定义为 Meta 分析, 并揭开了它在教育学、心理学及医学中的应用的新篇章^[3]。Meta 分析在这些学科的应用中取得了极大的成功, 发展出了多种分析方法。Mann 称其为医学方法学研究中的一次革命, 且羽翼渐丰^[4]。

基金项目: 国家自然科学基金重大项目 (39899370), 中国科学院重大项目 (KZ951-B1-110), 广东省自然科学基金项目 (980952)

作者简介: 彭少麟 (1956~), 男, 博士, 研究员, 博士生导师, 著名生态学家, 国家自然科学基金重大项目首席科学家, 全国“五一”劳动奖章获得者。

收稿日期: 1999-10-22

直到 90 年代,此方法才被生态学家发现,虽然目前它在生态学中的应用实例还很少,但已引起了生态学界的高度重视。Gurevitch (1993) 出版了第一部生态学中的 Meta 分析专著^[5], 并与人合作于 1997 年发行了 MetaWin 软件包。

在我国,彭少麟(1988)首次将此方法引入我国生态学界^[6], 并利用此方法进行生态学分析^[7]。

Meta 分析目前主要应用于对照实验的综合研究中, 目的为判断实验中的处理会对实验对象产生正或负效应; 效应是大还是小; 同一主题下不同独立实验的结果是否一致, 变异程度有多大等问题。

但 Meta 分析决不仅仅是一个数学分析过程, 它本身也是一项研究, 需要认真设计。主要步骤如下所述。

提出所要解决的问题并制定搜集、选择文献的标准。搜集文献, 这是一项非常繁重且关键的工作。为了能搜集到全面的文献, 通过各种途径来最大可能地收集已发表的和未发表文献(包括正式期刊中的论文、会议论文、摘要以及各种私人交换资料等)。

标定各研究的特点, 并对其进行分类。根据研究背景特点的不同将所有研究分为几个级别(class), 以作比较。

定量测度研究特点。为了避免分析时对质量不等的研究给予相同的结合标准, 导致分析结果的不准确, 分析家们提出了定性 Meta 分析, 即制定标准, 对研究特点进行打分评估; 综合研究结果并结合研究特点来分析结果。也有人称这一步为定量 Meta 分析, 以相对于定性 Meta 分析。

研究特征分析(敏感性分析), 分析研究的基本特征(研究对象、研究环境等的特征)和方法学特征对效应值之间的协变关系。

目前已有发展出多种定量 Meta 分析方法。但它们的基本思想是一致的, 那就是先提出假设, 构造一个结合统计量, 然后计算各研究的结合统计量, 并用其在定性 Meta 分析中所得分数去权重它的结合统计量; 计算各级别研究中的加权平均结合统计量(在平均过程中, 要根据其各结合统计量的方差进行权重); 做各级别研究间统计量的异质性检验。

定量 Meta 分析方法的不同主要在于结合统计量和统计假设的不同。

2 MetaWin 软件的特点

MetaWin 是一个主要为生态学工作者设计的定量 Meta 分析软件, 其主要特点如下所述。

2.1 提供了两种假设模型

这两种假设模型为固定效应模型和混合效应模型, 具体计算过程见文献[6]。两者的区别主要在于前者假设所综合的研究共享一个真实效应大小, 实际测量的效应大小不同是由于随机取样所导致, 而后者却假设研究间具有不同的真实效应大小, 即所测效应大小的不同是由两部分组成, 真实效应的不同, 随机取样造成误差。后者更切合实际, 区间估计较保守, 更受 Meta 分析家们欢迎。

2.2 提供多种可选择的结合统计量

在生态学领域内的 Meta 分析中最常用的结合统计量为 Hedges'd 效应值: $d = (X_e - X_c) / (S \cdot J)$ (其中, X_e 、 X_c 分别为实验组和对照组的测量平均值, S 为两组共同标准差, J 为小样本校正值), MetaWin 还提供了反应比(response ratio): $\ln(X_e / X_c)$ (X_e 、 X_c

的意义同上)这是从医学 Meta 中新引进的一种结合统计量;此外, MetaWin 还为对 Meta 分析较为熟悉的分析者提供了更多的选择机会,如相关系数 (correlation coefficient) 等。

2.3 提供了两种数据输入方式

对有经验的分析者可直接输入效应值、样本方差等所需数据,其格式称效应数据格式。这种数据输入法的好处在于分析者可根据所收集的文献的实际情况来自己构造结合统计量,也即 MetaWin 为分析者提供了较大的自由。在文献数据满足前两种结合统计量计算情况下,分析者可以输入原文献中的统计数据,如平均值、样本方差、样本大小来进行计算,比较方便,称原始数据格式。

2.4 提供了一项非参数检验——重取样检验

上述参数模型检验是在假设所有研究中的实验组和对照组观测值均遵循正态分布情况下进行的;许多 Meta 分析方法基于大样本近似原理,即当实验组和对照组样本大小不小于 10 时,效应值才趋于正态分布。但如果样本太小,实验组和对照组样本大小太悬殊或效应值太大时,大样本近似原理就变得不准确了^[8]。但事实上,许多生态学观察值却违背了上述情况^[9]。此外,只有当上述假设被满足时,用于检验研究间效应异质性的 Q 值才有近似的 X^2 分布^[7]。重取样检验法是取代传统参数和非参数检验的一种好方法。

重取样检验是一种计算机加强 (computer intensive) 非参数检验方法^[10]。MetaWin 中提供了随机化检验法 (randomization test) 和自助法或靴襻法 (bootstrap)。前者常被用来决定一个统计量的显著性水平,后者则用于给出统计量的置信区间。

MetaWin 中用自助法来计算所有研究总效应值和每一级别加权平均效应值的置信区间,对于样本含量为 i 的每一级别,我们均以放回式取样选取 i 个研究并计算其加权效应值,然后重复上述取样方法多次,按大小顺序将效应值排列起来,在两端取 2.5% 处的值做为 5% 至信区间的上下限,置信区间包括零在内的级别被认为没有显著不等于零。但当样本含量太小时,会出现区间估计过低,此时,可用偏差校正法^[9]。

MetaWin 中用随机检验来判断级别间效应大小的差异是否显著。首先用原始数算出 QB ,然后将 j 个级别里的所有研究混在一起,再随机将它们分成 j 个级别,级别含量仍与原来相同,算出 QB 值,重复此过程多次,得出一个 QB 值的分布, QB 的显著性水平为随机 QB 值大于等于实际 QB 值数占重复随机取样数的百分比。

3 MetaWin 软件的使用方法

3.1 MetaWin 软件构成

运行 MetaWin,只需一台装有 Windows95、Windows3.1 或 WindowsNT 的 IBM 兼容机,其中共包括 8 个文件。(1) MetaWin.exe: 在 Windows95 和 WindowsNT 下的可执行文件。(2) MetaWin.hlp: 在 Windows95 和 WindowsNT 下的帮助文件。(3) MetaWin.cnt: 在 Windows95 和 WindowsNT 下的帮助文件的内容。(4) MetaW16.exe: 在 Windows3.1 下的可执行文件。(5) Meta16.hlp: 在 Windows3.1 下的帮助文件。(6) Raw.dta: 以原文献统计数据输入数据的格式示范文件。(7) Effect.dta: 以效应大小输入数据的格式示范文件。(8) Gur-hed.dta: 作者的示范数据格式文件。

Windows95 和 Windows3.1 版本的不同之处主要在于研究特征类型量、每一特征类型中级别数、每一数据文件中所含研究量及非参数检验中的重复数的最大值的不同,Windows95 比 Windows3.1 范围更广。

3.2 MetaWin 软件使用方法

3.2.1 数据输入

打开 MetaWin 文件, 下拉 file 菜单, 点击 edit data file, 进入数据输入状态, 可直接在弹出的窗口中输入数据, 也可从 file 菜单中点击 load a file 上载已有的文件。原始数据输入格式如下:

```
sex tree state +/- Nc Ne Xc Xe Sc Se Label
m oak pa + 7 7 78.14 79.71 40.650 40.650 study1
m maple ny + 7 7 18.86 26.00 9.170 9.170 study2
f maple ny - 6 6 -1.80 -2.10 0.490 0.490 study3
```

其中, 第一行为标题行, 前 3 项为级别分类标准, 事实上, Windows95 版本可允许 10 个分类标准, Windows3.1 为 5 个; +/- 为方向符, 如果你所期待的效应值为正值时 (即实验中的处理会对实验对象产生正效应), 在按所期待趋势应该出现正效应值的研究项中加+, 负效应值的研究中加-, 它必须紧跟级别组, 否则程序运行时不能识别数据文件中共有几项划分级别标准; N_e 、 N_c 分别为实验组和对照组的样本含量; X_e 、 X_c 分别为实验组和对照组的测量平均值; S_e 、 S_c 分别为实验组和对照组的标准差; Label 为各研究的标记。标题行下面的每一行为一个研究的效应数据。效应数据输入格式为:

```
sex tree state +/- Nc Ne effect var Label
m oak pa + 7 7 78.036 0.286 study1
m maple ny + 7 7 0.565 0.347 study2
f maple ny - 6 6 1.533 0.517 study3
```

其中, effect 一列为效应值; var 为效应方差; 其它同原始数据格式。

3.2.2 数据分析

在打开 MetaWin 文件的同时, 会自动弹出一个 Meta-Analysis 窗口。在此窗口上部 Type of Input 一项中点击 raw 或 effect (确定数据为原始还是效应格式), 原来灰色的 Data file 就会加亮, 点击, 从弹出的‘打开’窗口中选定并打开要分析的数据文件。

在 Meta-Analysis 窗口中部选择固定效应或混合效应模型, 结合统计量, 也可增加重复检验; 窗口下部 group by 中可选择划分级别的标准, 并在 Refine Analysis 中可以在不改变数据文件的情况下去掉一些级别或研究来纯化分析。

所有这些选项选择好后, 即可点 start 键进行运算。运算结束后, 会自动弹出一个 Meta-Analysis output 窗口, 显示分析结果。

3.2.3 结果显示

在分析结果中, 可看到分析时间, 数据来源路径, 以及名为 Parametric methods 和 Meta-Analysis results for groups 的两个表。前者为所有研究的效应值表, 每一行代表一个研究, 包括其名称、所属级别、小样本校正 (J)、对照和实验两组的共有标准差 ($spool$)、效应值 (d)、95% 的置信区间 (95%CI)、各研究的方差 (V)、权重 (W)。

一般先假设所有研究享有共同的 d 值进行分析, 此时在第二个结果表中可看到所有研究的总平均效应值 (d_{++})、95% 的置信区间、同质性 (Q_{wi}), 自由度 (df)、 X^2 检验的 p 值。如果其级别内异质性经 X^2 检验显著, 则说明假设不正确, 此时按一定的标准将所有研究划分为几个级别, 再进行分析。此时的第二个结果表中会显示各级别内所有研究的加权平均效应值 (d_{i+})、95% 的置信区间、同质性 (Q_{wi}), 自由度 (df)、 X^2 检验

的 p 值以及级别间同质性 (Q_b)、级别内总同质性 (Q_w)、总的同质性 (Q_{total})。如果级别间同质性 (Q_b) 经检验后显著, 则说明级别间差异显著; 如果某一级别内同质性 (Q_{wi}) 经检验显著, 说明这一级别内各研究的效应值差异较大, 应该进一步划分此级别, 再分析, 直到 Q_{wi} 经检验不显著。

如我们在做捕食关系的 Meta 分析中发现, 捕食者导致被捕食者种群数量降低, $d_{++} = -0.3855$ (固定效应模型) $d_{++} = -0.4589$ (混合效应模型); 但不同标准划分的级别的效应大小有差异, 捕食效应随地带性而变化, 热带效应值最大, 亚热带、温带、寒带也有效应, 其中亚热带最小; 按所在生态系统划分级别时, 陆生生态系统级别为中效应, 淡水生态系统为小效应^[10]。

MetaWin 是一个操作简单且功能较全的 Meta 分析软件, 以 Windows 作支持, 用户通过界面与机器直接对话, 分析过程简单易学, 结果输出明了。遗憾的是 MetaWin 中没有考虑定性 Meta 分析所得出的研究质量评估值, 所以利用原始数据直接输入法不能对效应值进行研究质量权重。同时值得注意的是 MetaWin 只提供了定量 Meta 分析方法, 而 Meta 分析本是一项研究, 数量分析前需要认真设计, 分析后也需对结果进行研究特征分析。一个好的 Meta 分析不仅要选择好的定量分析方法, 而且分析前设计和分析后的特征分析都非常重要, 因为统计的目的是为解决问题提供科学依据。

参考文献:

- [1] Scheiner S M. Introduction: Theories, hypotheses and statistics. New York: Chapman and Hall, 1993
- [2] Kay D, Jesse A B. Meta-analysis: state of the science. *Epidemiological reviews*, 1992(14): 154~176
- [3] 赵宁, 俞顺章. Meta-analysis: 一种新的定量综合方法. *中国慢性病预防与控制*, 1993, 1(6): 277~281
- [4] Mann C. Meta-analysis in the breech. *Science*. 1990. 249: 476~480
- [5] Gurevitch J, Hedges V. Meta-analysis: combining the results of independent experiments. New York: Chapman & Hall. 1993
- [6] 彭少麟, 唐小焱. Meta 分析及其在生态学上应用. *生态学杂志*, 1998, 17(5): 74~79
- [7] 郑凤英, 彭少麟. 捕食关系的 Meta 分析. *生态学报*, 1999, 9(4): 448~452
- [8] Hedges L V, Olkin I. Statistical methods for meta-analysis. Orlando: FL Academic Press, 1985
- [9] Gurevitch J, Morrow L L, Wallace A, et al. A meta-analysis of field experiments on competition. *American Naturalist*. 1992(140): 539~572
- [10] Dean C, Gurevitch J, Michael S R. Resampling tests for meta-analysis of ecological data. *Ecology*. 1997, 78(5): 1 277~1 283

Introduction of MetaWin Software

PENG Shao-lin, ZHENG Feng-ying

(South China Institute of Botany, Chinese Academy of Sciences, Guangzhou 510650, China)

Abstract: MetaWin is the first software of quantitative meta-analysis for ecologists, which includes two hypothesis models and at least two cumulative statistics. It can be performed easily and proves a new non-parametric test, and resampling test.

Key words: meta-analysis; MetaWin; review