

# 定量化文献综述方法与元分析

崔智敏, 宁泽逵

(西安财经学院, 西安 710100)

**摘 要:**文献综述与文献述评在科学研究中占据着越来越重要的位置,体现了学术科研继承性与严肃性,但是,很长一段时间内,传统的描述性文献综述方法由于方法论上存在的缺陷,使得描述性文献综述成果之间由于作者主观评判的差异而存在严重的偏差。近年来,国外出现了一种基于元分析理论与方法的系统性文献综述方法,为定量化文献综述提供了方法与工具,这一方面体现了统计学方法论在应用层面的拓展,另一方面也体现了科学研究中文献综述的规范化发展。

**关键词:**描述性文献综述;系统性文献综述;元分析

**中图分类号:**C32

**文献标识码:**A

**文章编号:**1002-6487(2010)19-0166-03

## 1 传统描述性文献综述及其缺陷

文献综述在科学研究中占据重要的基础性地位,是通往专题研究的桥梁(朱玲,2006)。这是因为任何一项研究都是建立在前人贡献的基础上,开展某项专题研究之前对该领域的前人所作贡献作一个回顾性评论,不仅是对前人贡献的尊重,也是通过比较阐述本研究创新之处的一种方法(张昕竹、陈志俊,2003)。在经验研究中,由于传统的文献综述研究尚无统一的规范,其实现形式较为灵活,研究者多是“按照一个自己认为合理的研究逻辑主线将该研究领域的研究成果串联在一起”(张昕竹、陈志俊,2003)。这种传统的文献综述方法,Borenstein et al.(2009)称之为描述性文献综述。

然而,在阅读文献的实践中,人们常会遇到“针对同样问题的不同研究所做的结论各异”的情况(卫林英、段兴民,2006)。这是因为,在科学研究中,由于学多实际困难,如研究对象数量及经费的限制、各种环境因素的影响以及研究本身的或然性等原因,往往难以得到真正的随机样本或大样本。针对同一问题常常同时或先后有许多类似的研究,尽管研究人员在研究中采用尽可能统一的标准,并进行质量控制,但是研究结果往往有很大差异,甚至完全相反。而且这些情况在人文社会科学研究尤为突出,使得按照传统描述性文献综述方法得到的文献综述结论往往很难消除读者的种种质疑,如研究者是否会在主观上为了确保研究逻辑的一致性,突出研究结论,而回避既有研究中一些重要冲突,“适当”取舍文献,进而丢掉科学研究的客观公正性呢。因此,传统描述性文献综述方法存在很多缺陷(夏凌翔,2005;Borenstein et al.,

2009),即:

(1)传统描述性文献综述法未使用任何系统方法来对所综述内容的原始数据进行收集、综合,从而使得文献选择标准主观随意,选择过程不透明。

(2)随着电子文献检索技术的发展,各类相关研究信息越来越容易获得,传统描述性文献综述法无法处理这种“信息爆炸”局面,也不能赋予各类相关研究成果以合理的权重,进而使得所得综述结论多是有偏。

(3)由于(1)和(2)的原因,使用传统描述性文献综述法所得最终研究结论不可避免带有主观随意性。

(4)使用传统描述性文献综述法进行文献综述往往只是罗列以往的研究结果,多只能述而不评,丢掉了文献综述的本来意义。

(5)传统描述性文献综述法以描述既有文献研究成果为主,无法实现定量的综合研究,也查不出造成同一问题不同研究结论各异情况的真正原因,无法消除对各种表象的争论,进而不能对后继研究者贡献清晰的研究思路。

## 2 系统性文献综述及其实现

鉴于传统描述性文献综述存在的缺陷,国外不同领域的许多研究者早在1980s中期开始探索跳出传统描述性文献综述的局限而采用现代系统性文献综述,并开发出一套相配套的元分析统计理论与技术。

### 2.1 系统性文献综述原理与元分析

系统性文献综述最早也被称之为综合研究,它是指采用一套事先确定且透明的文献取舍标准,就某特定研究主题选

基金项目:国家社科基金资助项目(10XJY033);陕西省科技厅软科学资助项目(2010KRM50);陕西省教育厅人文社科专项资助项目(111M015);西安财经学院科研基金资助项目(07JD01)

作者简介:崔智敏(1961-),男,陕西户县人,副教授,研究方向:现代企业管理。

宁泽逵(1979-),男,湖南邵东人,博士研究生,讲师,研究方向:发展经济学。

取大量相关或相近的研究成果,并用一套特定的统计分析技术从这些分散的研究成果中总结出该研究主题的主要结论(Cooper,1989;Borenstein et al.,2009)。系统性文献综述法确保了从文献取舍标准的确定、到文献的具体选择、到分散研究文献结论的提炼等所有过程的明晰,它实质上是一种透明的文献综述研究机制(Borenstein et al.,2009),其突出特点是在传统文献综述过程中融入了对文献数据进行统计综合分析(即元分析)。因此,在使用元分析进行系统性文献综述实践中,往往要求按照事前根据数学原理确定的标准对所选取的每项研究文献赋予特定的权重。然而,即便如此,使用系统性文献综述分析方法所得结论可能仍不能调和综述者与读者间差异,但是至少,元分析这种统计学分析方法提供了一套客观的、可重复的文献分析讨论框架。

其实,元分析中所使用的各项规则是各项原始文献研究中所使用规则的一个拓展。例如,在原始文献中研究者常要报告其研究主题的均值与方差,而做文献综述的研究者则对同一研究主题不同研究成果所汇报的均值与方差之间的真实差异感兴趣,如果方法论上不存在障碍,那么,他们将会使用方差分析或多元回归技术来揭示导致该主题研究结论差异的影响因素及其贡献程度。Borenstein et al.(2009)认为实践中使用元分析进行分析文献综述的原因可能很多,但一定要明确四个目的,即(1)何时开展文献的元分析,(2)应该选用哪种模型来分析搜集到数据资料,(3)应该使用何种敏感性分析,(4)如何解释所得研究结论。

## 2.2 元分析的统计学原理与技术

### (1)元分析及其统计学思路

据 Glass (1976)元分析是指以综合已有的发现为目的,对单个研究结果进行综合的统计分析方法。但是随着元分析方法的应用范围的扩大和自身的发展,元分析中既可包括了定量分析也可包括定性分析,它是一种综合多个同类研究的结果,对研究效应进行定量合并的分析研究过程和系统方法。因此,简单地讲,元分析就是应用特定的设计和统计学方法对以往的研究结果进行整体的和系统的定性与定量分析(夏凌翔,2005)。元分析的基本统计学思路是,以原始研究结果为单位,设计较严密,强调对有关研究进行全面的文献检索,有明确的文献纳入和排除的标准,系统地考虑了研究的对象、方法、测量指标等对分析结果的影响,对纳入文献进行了严格评价,并在此基础上对结果进行定量的合并。因此,以元分析为基础的系统文献综述相对于传统描述性文献综述,能最大限度地减少各种偏倚,确保结论的科学性、客观性和真实性。

### (2)元分析的统计学原理

效应值。在元分析中,需要对许多实验研究的结果进行定量综合,首先应计算出每一研究结果的效应值  $d$  (Effect size),它反映了两组变量间的关联程度,计算公式为  $d = (\bar{X}_1 - \bar{X}_2) / S_{within}$  (其中  $S_{within} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$ ),即实验组与控制组

的平均数之差再除以组内标准差所得的值。如果研究中没有提供均数和标准差,但提供了  $t$  值、 $r$  值或  $F$  值等显著性检验参数时,也可通过转换公式求出相应的  $d$  值 (Borenstein et al.,2009)。

Borenstein et al.(2009)指出通过计算效应值的方差、标准误、置信区间等统计量可以衡量效应值的精度。但是,研究中所取样本规模与研究设计会极大地影响到效应值的精度:样本规模越大,精度则越高;配组实验设计的精度要大于对比独立组实验设计,大于聚类组实验设计。

效应值的齐性检验。齐性检验又叫抽样样本效果大小的一致性分析,它是指所抽取的样本效果大小是否来自于共同的总体,因而可以看作是效应值之间的同质性检验。齐性检验的目的是指出不是所有的研究结果都能被综合进同一元分析中,如果研究结果不齐性,调查者应考虑是否是由随机抽样误差所致,如果不是随机抽样误差所致,则应该考虑将这些研究结果分成不同的子集合,使这些集合之间呈齐性关系,再对它们分别进行元分析(夏凌翔,2005)。齐性检验的

统计量为:  $Q = \sum_{i=1}^k W_i(Y_i - M)^2$ , 其中  $W$  是每项研究的权重(等于效应值方差的倒数),  $Y$  是效应值,  $M$  是效应值之和,  $k$  是研究数量(Borenstein et al.,2009)。

固定效应与随机效应模型。根据各样本效应值是否来自同一总体,元分析提供了两种统计模型假设。固定效应模型假设所分析的所有原始研究样本文献中,只有一个真实的效应值,所表现出来的差异只是统计抽样误差。随机效应模型则假设各原始研究样本文献中的真实效应值是来自于一个真实效应值分布集合的抽样,因此,真实效应值在研究样本文献间可以有所不同。如此,随机效应模型分析的重要任务是精确地估计这个真实效应值分布。理论分析表明各项研究成果所占权重在随机模型中更为均衡一点,相对于固定效应模型而言,大样本研究文献成果会赋予较低的权重,小样本研究文献成果会赋予较高的权重。而且,总效应值的标准误及其置信区间在随机效应模型下会更大更宽一些。Borenstein et al.(2009)认为,这两种统计模型目前尚无明显优劣之分,至于具体选择哪个模型则取决于样本效应值分布和相关资料误差,而实践中所出现的先用固定效应模型分析,如果异方差显著再用随机模型分析的策略通常是错误的。

### (3)元回归及其拓展

事实上早在元分析出现之际,有研究者就将该方法拓展到多变量元回归分析,其中,Stanley and Jarrell(1989)则并成功地该元分析方法应用到经济学研究当中,形成经济学文献分析的元回归分析。Stanley and Jarrell(1989)指出所谓元回归分析(MRA)是回归分析的回归分析,它将文献综述过程客观化,给研究者提供了一种用于分析、评估、评价各种可供选择的模型设定以及特定的搜索所产生的影响。

根据 Stanley and Jarrell(1989)和徐少君、金雪军、皮永华(2006),元回归分析要求因变量是一个来自于各原始研究综合的统计量,如回归参数;自变量则可能包括这些研究中

使用的方法、设计或数据的一些特征。在使用元回归分析方法时,首先要求收集所有的相关文献,包括发表和未发表的,从而降低由于任何非随机的文献选择所导致的潜在偏差,一个好的元回归分析需要提供有关诸如年份、文献入选标准等搜索细节,以便其他学者进行“复制”研究。其次,元回归分析需要选择一个代表统计量,并将其转化为统一的可比较的尺度。而且,在转化为一个统一的可比较的尺度的方法上,元回归分析一般不使用元分析中普遍采用的“效应值”,而使用其他的统计量,如标准正态分布和  $t$  统计量等。

例如,对于应用经济学文献中常见的计量模型:  $Y = X\beta + \varepsilon$ , 其中  $Y$  是  $(n \times 1)$  维的因变量,  $X$  是一组  $(n \times m)$  维的自变量,  $\beta$  是一组  $(m \times 1)$  维的回归系数,  $\varepsilon$  是随机扰动因子。利用元回归分析可以找出不同研究文献中原始回归模型  $(Y = X\beta + \varepsilon)$  间的真实差异,首先可以构建元回归分析的基本模型:  $b_j = \beta +$

$\sum_{k=1}^k \alpha_k Z_{jk} + e_j (j=1, 2, \dots, L)$ , 其中, 因变量  $b_j$  为“统一尺度值”, 是第  $j$  篇文献研究中的所关心的变量的系数估计值(一般是原始回归模型中的  $\beta$  值); 自变量(也称为“调节变量”)  $Z_{jk}$  代表原始实证文献中所涉及到的一些重要特征, 如各研究所使用的不同数据集、不同模型的哑变量等, 但由于不同的研究往往采用了的数据集、自变量和时间段, 因而在有限的研究文献数量的样本空间里, 不可能包括所有的这些特征, 需要有重点的选择。其系数  $\alpha_k$  可用来解释特定文献中某一特征与其他文献的偏离效应,  $e_j$  是元回归分析的随机扰动因子。在这个模型中  $\beta$  则是我们所关注的真正的参数估计值。但是, 就像其他的统计模型一样, 元回归分析也可能出现一些问题, 在一般情况下, 可以通过标准的统计检验进行自相关、异方差等的检验, 而对于元回归本身设定的一些问题, 则需要通过专门的统计检验加以鉴别或解决。例如, Borenstein et al. (2009) 认为元回归分析可以基于固定效应模型假设或随机效应模型假设, 但是在多数案例表明基于随机效应模型假设更具适应性。

### 2.3 元分析主要应用软件简介

Borenstein et al. (2009) 在其著作《Introduction to meta-analysis》附录部分重点推介了三种元分析应用软件, 即 Comprehensive Meta-Analysis (CMA) 2.0、RevMan 5.0 和 Stata macros with Stata 10.0。其中 CMA 2.0 是由 Borenstein et al. 开发的与《Introduction to meta-analysis》配套的收费软件, 其官网为 [www.Meta-Analysis.com](http://www.Meta-Analysis.com)。CMA 2.0 的操作界面友好, 使用者可以按照其格式要求提示录入原始数据, 操作方法与微软 Excel 2003 类似。同时 CMA 2.0 能够实现如下功能, 即计算每项研究的效应值与方差、执行元分析、执行敏感性分析、创建森林图、复杂数据结构分析、子群分析与元回归、出版者偏好检验等。

Stata 是一个功能强大而又小巧玲珑的统计分析软件, 随着元分析的快速发展, Stata 爱好者早在 1997 就开发出了基于 Stata 的元分析程序, 发展至今, 其功能已十分全面 (Sterne, 2009)。柏建岭等 (2007) 通过三个实例分析显示, 与 RevMan 5.0 相比, Stata 可更个性化、更高效专业地完成元分

析。但是, Stata 作为一种指令驱动程序, 其元分析的实现也是需要编辑 Stata 程序指令才能得以实现, 而且, Stata 元分析包所接受的数据录入格式相对 CMA 2.0 而言较为有限 (Borenstein et al., 2009)。

其实就关于元分析理论学习与软件操作而言, CMA 2.0 不但与《Introduction to meta-analysis》配套, 而且还提供了丰富的用户操作手册“Meta-Analysis Manual”, 加之 CMA 2.0 还提供 30 天的免费试用期, 因此, 对于初学者而言 CMA 2.0 是较为方便的。但是, 鉴于 Stata 软件分析结果在当前学术研究中的普遍影响力, 其分析结果更易为高影响力杂志所认同 (柏建岭等, 2007), 因此, 基于 Stata 软件的元分析更值得推广应用。

### 3 小结

据了解, 基于元分析的系统性文献综述方法, 国外已广泛地应用于医学、药学、心理学、教育学、犯罪学、生态学、体育学以及商学之中 (Borenstein et al., 2009)。但是, 国内直到 1990 年代才开始探索将该方法应用到医学、药学、心理学、体育学、教育学、生态学等学科, 而且, 在其它人文社会科学中的推广应用也尚未真正开始 (夏凌翔, 2005; 徐少君、金雪军、皮永华, 2006)。因此, 有必要号召统计学界与其他人文社会科学学界合作, 通过引进教材、软件与培训专门人才, 将这种方法引入国内人文社会科学界, 以规范较为混乱的文献综述研究现状。

#### 参考文献:

- [1] Borenstein, M., Hedges, L.V., Higgins, J.P.T. and Rothstein, H.R. Introduction to meta-analysis[M]. Chichester, UK: Wiley and Sons, Ltd. 2009,
- [2] Cooper, H.M. Integrating research: a guide for literature reviews (2nd ed)[M]. Newbury Park, CA: Sage Publications. 1989.
- [3] Glass, G.V. Primary, secondary and meta-analysis of research[J]. Education Research, 1976, 6(5)
- [4] Stanley, T.D. and Jarrell, S.B. Meta-regression analysis: a quantitative method of literature surveys [J]. Journal of Economic Surveys, 1989, 19(3).
- [5] Sterne, J.A.C. Meta-analysis in stata: an updated collection from the Stata Journal[M]. Texas, USA: A Stata Press Publication. 2009.
- [6] 柏建岭, 钟文昭, 郑明华, 陈峰. Stata 在 Meta 分析中的应用[J]. 循证医学, 2007, (6).
- [7] 夏凌翔. 元分析及其在社会科学研究中的应用[J]. 西北师大学报 (社会科学版), 2005, (5).
- [8] 徐少君, 金雪军, 皮永华. 一种新的经济学文献综述方法: MRA 研究述评[J]. 统计研究, 2006, (8).
- [9] 张昕竹, 陈志俊. 经济学论文的写作规范[J]. 数量经济技术经济研究, 2003, (8).
- [10] 朱玲. 文献研究的途径[J]. 经济研究, 2006, (2).
- [11] 卫林英, 段兴民. Meta 分析在科学研究中的应用与展望[J]. 生产力研究, 2006, (6).

(责任编辑/易永生)