

· 基础理论与方法 ·

结构方程模型方法在流行病学研究中的应用

杨廷忠 阮哈建 李甫中

【摘要】 目的 论述结构方程模型(SEM)方法在流行病学研究中的应用。方法 简述 SEM 的主要构成、统计假设和目前常用的软件及这一方法如何在流行病学研究中的应用和对应用中的有关问题的处理。结果 相对于传统的流行病学方法,SEM 是一种综合思维方法,不仅分析因素和疾病之间的关系,也分析因素和因素之间的关系;同样是一种验证性的方法,对于有些复杂问题的流行病学研究,特别是以理论为依据的研究颇为重要;SEM 分析能够得到潜在变量的有关参数,并对表述潜在变量的显变量的测量误差做出估计。结论 SEM 能够应用于流行病学的研究,且具有较传统流行病学分析方法无法比拟的优势。

【关键词】 流行病学; 结构方程模型; 潜在变量

The application of structural equation model approach in epidemiological research YANG Ting-zhong*, RUAN Ha-jian, LI Fu-zhong. *Department of Social Medicine, School of Medicine, Zhejiang University, Hangzhou 310031, China

【Abstract】 Objective To discuss application of structural equation model (SEM) approach in epidemiological research. **Methods** A brief overview on major components of SEM, statistical assumptions underlying the use of SEM, and current software available to users and how SEM can be used were discussed through a practical epidemiological research project. **Results** Advantages of SEM comparing with conventional epidemiological approach were shown. SEM, having the nature of comprehensive thinking and analytic approach, not only exploring the association between factors and diseases but also among factors. It also served a confirmatory, rather than exploratory approach on data modeling, as well as having the capability of correcting estimates by separating measurement error from the equations, to provide modeling the latent variables. **Conclusion** SEM approach could be used in epidemiological research as having some advantages comparing with conventional epidemiological approaches.

【Key words】 Epidemiology; Structural equation model; Latent variable

结构方程模型(structural equation model, SEM)最初应用于社会科学领域的研究。由于这种方法具有综合性和灵活性的特点,随后很快就在心理学等多个学科研究中得以广泛地传播和应用^[1-3]。当今的流行病学研究所面临的问题越来越复杂。对人类健康构成极大威胁的很多疾病,如心血管病、癌症和艾滋病的病因根本不是传统意义上人们所认为的那样简单和“纯朴”;目前流行病学研究已拓展到疾病以外更加广泛的领域,如突发公共卫生事件、心理压力及行为问题等;无论什么问题,想当然地提出预防和控制措施都无济于事。疾病或健康现象是多个因素共同作用于机体所产生的结局,但传统的流行病学因素分析的方法只是孤立地关注诸因素与疾病之间的关联,缺乏从总体结构上的把握,SEM 正好提

供了这样一种整体思维的方法。SEM 已在众多领域得到很好地应用,本文对这一问题的兴趣并不在于其有多新颖,主要在于这种思路对流行病学研究所具有的借鉴价值。

基本原理

1. SEM 的构成:SEM 由测量模型和潜在变量模型两部分构成^[1]。

(1)测量模型:该模型是用以描述一系列测量变量与其所负荷的潜在变量之间的关系。分别用方程表示

$$y(p \times 1) = \Lambda_y(P \times m) * \eta(m \times 1) + \epsilon(p \times 1)$$

$$x(q \times 1) = \Lambda_x(q \times n) * \xi(n \times 1) + \delta(q \times 1)$$

式中 y 和 x 分别为内生测量变量和外生测量变量所构成的向量($p \times 1$)和($q \times 1$)。 Λ_y 和 Λ_x 分别为 y 对 η 和 x 对 ξ 的因子负荷矩阵($P \times m$)和($q \times$

作者单位:310031 杭州,浙江大学医学院社会医学教研室(杨廷忠、阮哈建); Oregon Research Institute, U.S.(李甫中)

n)。 η 为内生潜在变量所构成的向量 ($m \times 1$)， ξ 为外生潜在变量所构成的向量 ($n \times 1$)。 ϵ 和 δ 分别为 y 和 x 的测量误差所构成的向量 ($p \times 1$) 和 ($q \times 1$)， 它们被假定与 ξ 和 ζ 不相关， 且它们之间也互不相关。 在流行病学研究中有时会遇到一些抽象的概念， 这些概念因为不能直接进行测量而作为潜在变量来看待， 它通过若干个其他测量题目来间接表达。 每个测量题目得到的指标， 而并非是对概念的一个完全的表达， 具有一定的测量误差。 测量误差表示在测量模型中每一个测量项目对变量表达的准确性， 其取决于两方面， 其一是理论假设的正确程度， 其二是实际操作的准确程度。 如果理论假设基本正确而实际测量问题很大， 同样会使测量项目偏离对变量的表达， 致使测量误差很大。 测量模型是在已有的假设下进行的， 用于对已提出的假设的判断和评价。 一个测量模型可以包含若干个潜在结构。

(2) 潜变量模型： 该模型表示外生潜变量和内生潜变量之间的因果关系， 这种关系可以是直接的也可以是间接的。 其形式为

$$\eta(m \times 1) = B(m \times m) * \eta(m \times 1) + \Gamma(m \times n) * \xi(n \times 1) + \zeta(m \times 1)$$

式中 B 为内生潜变量对内生潜变量 $\{\eta\}$ 效应系数的 $m \times m$ 矩阵， Γ 为外生潜变量 (ξ) 对内生潜变量 $\{\eta\}$ 效应系数的 $m \times n$ 矩阵， 最后一项为残差项构成的向量 ($m \times 1$)， 它被假定接近于 0 及与 ξ 不相关。

2. SEM 的建立与检验： 所有的 SEM 分析必须包括①依据有关理论和知识建立模型(一般以图形来勾画出有关变量之间的关系， 模型中的每个成分需经操作化处理)； ②确认， 也就是实际数据必须满足模型估计的需要， 对于每个未知参数要能够获得惟一的解； ③通过使用 SEM 程序可以完成参数估计； ④根据一系列的模型拟合指标， 对于实际操作模型和理论假设模型之间的吻合程度进行评价； 预先的理论假设模型常出现统计拟合不满意的情况， 这需要对原来的设想进行修改和完善； ⑤模型外推， 系指所建立的模型能否适合于其他资料。

SEM 是一个参数统计方法， 在应用中首先要涉及到样本数量问题， 它取决于模型中所包含的因素的数量、 资料分布和模型的复杂性。 一般而言， 少于 100 个样本的模型可能会产生不可靠的结果， 而大于 1000 个样本的模型则很易于成立。 一个普遍使用的估计样本量的规律是模型中样本数与参数量之比， 即 1 个参数 10 个样本^[4]。

SEM 统计假设为多变量正态分布， 这就意味着测量变量是连续的且呈正态分布。 但有些实际的资料往往不能满足这个条件， 常有些研究者则对此不屑一顾。 有研究表明， 非正态资料在合理假设情况下， 参数估计可靠但标准误估计会发生问题。 而很偏态的资料， 无论是用最大似然法还是最小二乘法进行估计， 都会产生不正确的统计检验和偏性的标准误^[5]。 现已有用于处理非正态问题的 SEM 程序， 如渐进式自由分布方法 (asymptotical distribution-free methods) 可以解决此类问题。 但此方法需要大的样本来获得稳定的权重矩阵。 Sattrra-Bentler 法克服了这一缺陷^[6]。

参数评价使用 t 检验法， 群内比较使用标准参数， 群间比较使用非标准参数^[7]。 结构方程模型运算可以在 SPSS 统计软件的 LISREL 程序上进行或在 SAS 软件 CALIS 模块的 EQS 程序上进行。 EQS 编程首先指明 CALIS 模块， 再在 Ram 语句中列出各变量(标号)矩阵， 结果变量在前， 原因变量在后； 第一层次运算负荷值和路径参数， 第二层次运算误差参数^[8]。

3. 模型评价： 模型的拟合优度检验最易于使人想到 χ^2 检验方法， 似然比 χ^2 检验无显著性表明观察和预测的共变异矩阵差别不明显， 方程拟合良好， 但值得注意的是在大样本条件下即使其他指标拟合情况很好， 很小的差别会使 χ^2 值变得有显著性， 故对该指标的要求不必很严格。

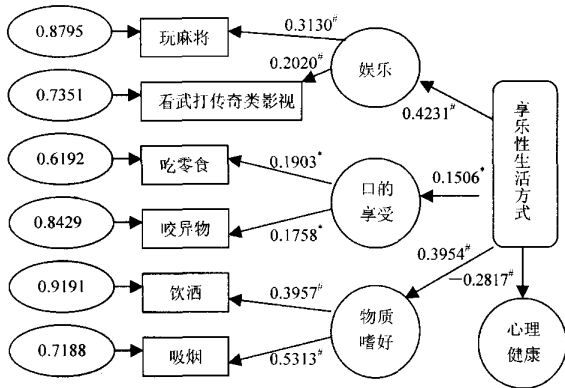
评价的其他指标还有① GFI (goodness of fit index) 表示由模型能够解释的变异和共变异， 其值在 0 和 1 之间， 在 0.9 及以上时表示模型拟合可以接受， 0.95 以上表示模型拟合良好； ② AGFI 为考虑了自由度后的 GFI， 其值通常 > 0.8 时表明拟合可以接受， 0.85 以上表示方程拟合良好； ③ RMR (root mean square residual) 为均方根残差， 表示平均拟合度的残差测量， 其值 < 0.1 表明拟合可以接受， < 0.06 表示方程拟合良好。

目前有几种 SEM 商业软件可用于资料分析， 包括 Amos、 EQS、 LISREL 和 Mplus [Amos (<http://www.smallwaters.com/index.html>)、 EQS (<http://www.mvsoft.com>)、 LISREL (<http://www.ssicentral.com>)]。 Mplus (<http://www.statmodel.com>)。 这些程序是免费学习版可以下载， 其程序一般限于很少的例数和变量， 可以进行一些简单的分析。 商业统计软件有 CALIS (有单独的也有包含

在 SAS 等大型软件包中的)、LISREL(有单独的也有包含在 SPSS 等大型软件包中的)、Ramona 和 SEPATH 等。这些程序具有共同的基本方面,但每种又各具特色。如 EQS 能够提供多种统计检验, Mplus 能够提供混合模拟。

实例分析

本例是 1426 名男性高中生吸烟和饮酒等行为与心理健康的研究资料,其变量的测量见文献[9]。结构模型包括测量模型和潜变量模型。前者被假定 6 个行为测定指标代表 5 个潜在因子,每一行为测量变量被假设只能负荷于其中的一个潜在因子上,由此构成享乐性行为结构;在这里娱乐、口的享受和物质嗜好都是一些抽象的概念,它通过一些具体的测量指标来表达。后者检验以上 3 个潜在因子能否由单一的二级潜在因子所解释,是否支持享乐性行为是享乐性生活方式的构成部分,及享乐性生活方式到心理健康之间路径。模型拟合结果 GFI: 0.9812; AGFI: 0.9714; $\chi^2 = 12.564$, $P(0.1213) > 0.05$, 模型拟合良好(图 1)。结果提示吸烟和饮酒是享乐性行为的构成部分,是享乐性生活方式的表现,而享乐型生活方式不利于心理健康。这种模型也可以推广到疾病的研究,将出口接在冠心病上就形成了冠心病行为致病结构模型。



图中圆圈代表潜在因子,方框代表行为测定变量,椭圆内数字为测量误差,参数值为标准化值,其显著水平以非标准化值用 t 检验决定, * $P < 0.05$, # $P < 0.01$

图1 行为与心理健康的结构模型

讨论

当代流行病学研究面临着很多复杂问题的挑战,应当重视以理论为依据的研究。SEM 是一种验证性的思维方法和分析技术,可以构建一个多维网状病因分析结构,明确各个变量在这个结构中的位点和各个位点之间的资源流向,在此基础上进行结构拟合。在 SEM 结构中,既有单个因素的连接点又有多个因素的交汇点;在这个结构中,疾病可以不是终点。就病因因素而言,它们不但具有直接作用,且还可具有间接作用;不但具有主效应作用,且还可具有反馈作用;特别是可能还有支配若干因素的潜在因子存在。显然,这比危险因素分析的方法获得的信息要更加充分和具体,对于现象的认识也更加深入。

在很多疾病和健康现象的流行病学研究中往往要涉及到一些社会、心理和行为变量,而后者大多是一些抽象的概念,往往变异性大,难以测量,其本身无法被直接观察,需要用其他的多重测量来表示。目前流行病学研究对于这一问题处理办法通常有两种,一是使用一些现成的测量工具,但现有的测量工具十分有限,难以适应对于各种疾病的研究,更不能表述研究者活跃的思维,这使其只局限于分析若干个因素与某些疾病之间的关系;二是直接用某一条测定项目表达某一概念,但这种做法的误差非常大。SEM 为抽象变量的表达和测量准确性的检验提供了方法。它将潜在变量指标的测量分为真正的变异和误差变异两种,获得对测量误差的估计。不但考虑如何应用若干个测量项目准确地表达某一变量的属性,而且应将这部分内容作为整个结构的有机部分进行整体思考和判断。传统的流行病学分析方法无法做到这一点。

SEM 属于多维立体网状结构,结构中所涉及的要素间不但呈上下、左右空间关系,而且呈前后的时间关系,故这一方法应是纵向观察研究。当然有些横向资料在逻辑上具有纵向资料的性质,如父母吸烟与子女吸烟的关系,故没有必要刻板地看待这一问题。通过横断面资料依据有关变量逻辑上的因果关系检验建立的模型,也许会为病因研究提供某些线索和为进一步的纵向研究打下基础。

SEM 是一种因果关系模型。在这个模型中,有些变量只影响其他变量,而不受另一些变量的影响,常见的是一些人口学和社会变量;而有些变量则处于以上两种情况,常见的是一些心理、行为变量;过去我们常将疾病作为只受其他变量影响的变量来处理,实际在很多情况下疾病对其他变量有反向作用,明确这一点对于认识疾病发生的机制大有裨益。

SEM 是一种因果关系模型。在这个模型中,有些变量只影响其他变量,而不受另一些变量的影响,常见的是一些人口学和社会变量;而有些变量则处于以上两种情况,常见的是一些心理、行为变量;过去我们常将疾病作为只受其他变量影响的变量来处理,实际在很多情况下疾病对其他变量有反向作用,明确这一点对于认识疾病发生的机制大有裨益。

SEM 最初的设计主要针对连续变量的分析,特别是结构中的出口变量应为连续变量。但在流行病学病因研究中常常为分类变量。现已发展出适合于分类资料的结构分析模型,其基本思路是对分类变量 Y 进行标准化处理^[10]。同时可以采取变通的办法,如果出口变量为分类变量,可以在此变量上连接某个或某些连续变量,使其变为中间变量。也可采取分步分析策略,但这样做的结果缺陷在于可能使整个模型的拟合优度有所下降。所要声明的是 SEM 分析能够用若干个指标对模型的拟合优度做出估计,这种估计是表示模型对数据总拟合度测量的统计学意义,而不能用来判断模型的质量和模型的理论及实际意义。尽管模型能对测定误差做出估计,但关键还是在实际测量中的控制。应强调的是应用 SEM 得出的结论只能用来排除不能用来肯定。

在技术上相对于传统的方程到方程的估计方法,SEM 的独特之处在于将因子分析和回归分析整合起来,可以同时进行假设结构之内和假设结构之间的分析。同时,SEM 可以进行多个样本的分析,显示不同样本结构关系参数的差异。随着研究的不断深入,SEM 已经被发展为较复杂的一些模型。对于纵向资料分析的变化率和变化的个体差别,已发展出测量误差的交互作用分析和多级资料结构设计等方法^[11,12]。这些都值得进一步关注。

值得强调的是,SEM 的分析应建立在符合实际的理论构思和严格的变量测量之上。如果无任何理论依据和实际工作的基础就直接建构模型,这种模型除了提供统计学相关的意义外,别无任何帮助。同样没有准确的数据为基础,就不会产生较理想的结构模型。对于病因潜在变量的统计分析首先看若

干病因因素之间有无关联,然后再分析这种联系能否被一个共同因子所解释及这一个因子是否导致疾病的发生。因此这一过程是一种验证的方法,而不是探索的手段。故这一方法的使用应该在其他流行病学病因研究方法的基础之上进行。

参 考 文 献

- 1 Bollen KA. Structural equations with latent variables Toronto: John Wiley & Sons, 1989. 318-403.
- 2 Cliff N. Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, 1983, 18: 115-126.
- 3 Wall MM, Li R. Comparison of multiple regression to two latent variable techniques for estimation and prediction. *Stat Med*, 2003, 22: 3671-3685.
- 4 McDonald RP, Ho M-HR. Principles and practice in reporting structural equation analyses. *Psychological Methods*, 2002, 7: 64-82.
- 5 West SG, Finch JF, Curran PJ. Structural equation models with non-normal variable. In: Hoyle GH, ed. *Structural equation modeling: Concepts issues and application*. Thousand Oaks: Sage, 1995. 56-75.
- 6 Chou CP, Bentler P. Estimates and tests in structural equation modeling. In: Hoyle GH, ed. *Structural equation modeling: Concepts issues and application*. Thousand Oaks: Sage, 1995. 167-244.
- 7 Satorra A, Bentler PM. Corrections to test statistic and standard errors in covariance structure analysis. In: Von Eye A, Clogg CC, eds. *Analysis of latent variable in developmental research*. Newbury Park: Sage, 1994. 399-419.
- 8 SAS Institute, Inc. SAS/STAT users guide. Cary, NC: Author, 1998. 370-432.
- 9 杨廷志, 施卫星. 男高中生吸烟、饮酒和享乐性行为关系的结构分析. *中国行为医学科学*, 2002, 11: 72-81.
- 10 Muthen BO. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 1984, 49: 115-132.
- 11 Willet JB, Sayer AG. Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, 1994, 116: 363-381.
- 12 Fuzhong L. Approaches to testing interaction effects using structural equation modeling methodology. *Multivariate Behavioral Research*, 1998, 33: 1-39.

(收稿日期: 2004-04-26)

(本文编辑: 张林东)

· 读者 · 作者 · 编者 ·

本刊对统计学符号及统计学方法的要求

按 GB 3358-82《统计学名词及符号》的有关规定书写,常用如下: (1) 样本的算术平均数用英文小写 \bar{x} (中位数仍用 M); (2) 标准差用英文小写 s ; (3) 标准误用英文小写 s_e ; (4) t 检验用英文小写 t ; (5) F 检验用英文大写 F ; (6) 卡方检验用希文小写 χ^2 ; (7) 相关系数用英文小写 r ; (8) 自由度用希文小写 v ; (9) 概率用英文大写 P (P 值前应给出具体检验值, 如 t 值、 χ^2 值、 q 值等)。以上符号均用斜体。关于资料的统计学分析: 对于定量资料, 应根据实验或调查设计类型和资料的条件选用合适的统计学分析方法, 不能盲目套用 t 检验和单因素方差分析; 对于定性资料, 应根据实验或调查设计类型、列联表中定性变量的性质和分析目的选用合适的统计学分析方法, 不能盲目套用 χ^2 检验; 对于回归分析, 应结合专业知识和散布图选用合适的回归类型, 不能盲目套用简单直线回归分析。

本刊编辑部