

结构方程模型应用中模型选择的原理和方法

柳恒超¹, 许 燕¹, 王 力²

(1. 北京师范大学 心理学院, 北京 100875 2. 华南师范大学 应用心理研究中心 广州 510631)

摘 要 随着结构方程模型在心理学研究中的应用, 针对研究问题建构理论模型成为研究者们所重视的问题。该文探讨了心理学研究中模型选择的一般原则和基本方法, 希望对广大研究者有所帮助。

关键词 模型选择 嵌套模型 非嵌套模型

中图分类号 B841.2

文献标识码 A

文章编号 1003-5184(2007)01-0075-04

1 引言

近年来, 随着统计研究方法的发展, 结构方程模型(Structural Equation Model)在心理学研究中的应用越来越广泛。研究者们对建构关于某一研究问题的理论模型也越来越感兴趣。建构模型的主要目的是把理论假设和观测数据连接起来, 用观测数据来检验研究假设的正确与否。但是即使观测数据能够很好地拟和模型, 也并不意味着就找到了唯一的真理, 因为可能存在着许多个与研究数据都具有良好拟合程度的模型, 这就涉及到了研究中的模型选择问题。模型选择是一个相对比较复杂的过程, 既依赖于相关统计工具的使用, 但是也不能够忽视模型建构的理论基础。现在很多研究者都对科学研究中的模型选择问题进行了深入的探讨, 文章旨在根据相关研究的成果和自身的研究经验, 对心理学研究中的模型选择的原理和方法进行介绍, 希望能够对广大从事心理学研究的研究者们有所帮助。

在科学研究中, 研究者感兴趣的中心问题是所得到的模型是否是最好的模型, 即是否是一个最接近真实情况的模型。首先, 有关研究者提出了进行模型推断和选择的三个基本原则^[1]: 1) 简单性和俭省化原则: 对于一个模型来讲, 如果待估计的参数太少, 这个模型可能是有偏差的, 也就是模型的过度拟和问题; 而拥有太多待估计参数的模型, 其精确性可能很差, 也会产生不能拟和这样的问题。所以从某种意义上, 所谓“模型选择问题”就是不能拟和和过度拟和的平衡问题。2) 多个假设模型原则: 首先针对研究问题, 根据已有的研究资料提出多个假设模

型, 然后收集数据、进行分析, 通过模型与实际数据的拟合情况和模型比较的结果, 确定最终的结果模型。3) 统计检验力原则: 根据统计分析结果所提供的数量化信息指标来判断假设模型成立的证据是否充足。由于模型的建构和选择是建立在统计推断的基础之上, 因而根据相应的统计检验指标的好坏来进行判断仍然是当前模型建构和选择的主要方法。

结构方程模型(SEM)是当今统计方法的一大发展, 它把测量方程和结构方程完美地结合了起来, 成为多变量数据分析的重要手段之一。现在, 其在社会科学和心理学研究中的应用日趋重要和普遍^[2]。研究者业已开发出很多用于结构方程模型分析的统计软件, 如 LISREL, EQS, AMOS 等, 其中, 专用统计软件为 LISREL。下面就结合 LISREL 介绍几种在心理学研究中常用的模型选择的原理和方法。

2 嵌套模型(Nested Model)情况下模型选择的方法

在心理学研究中, 通常会在已有研究的基础上, 针对某一研究问题, 提出很多相关的模型假设。如果这些模型之间具有嵌套关系, 就可以使用嵌套模型的模型选择方法。

2.1 嵌套模型的概念

Keith F. Widaman 和 Jane S. Thompson(2003)对嵌套模型的定义如下^[3]: 如果模型嵌套于模型需要满足以下两个条件: 第一, M_k 需要估计的参数应少于 M_l , 因而其自由度更高; 第二, M_k 不应该包括在 M_l 中没有出现过的参数。只要满足以上两个条件, 就可以说 M_k 嵌套于 M_l 。

2.2 嵌套模型的连续性

嵌套模型具有连续性,这种连续性可以表示为: $M_0 \dots M_k \dots M_l \dots M_s$,其中 M_s 为饱和模型(Saturated Model), M_0 为零模型(Null Model)。所谓饱和模型是指模型中待估计的参数的数量等于观测变量协方差矩阵中数据点的数量。饱和模型能够精确地再生出观测变量的协方差矩阵,完美地拟和数据。饱和模型的卡方值和自由度都为零。零模型的设定有很多方法,但是所有零模型都有共同的特点,即零模型中的待估计的参数的数量要少于能够从数据中得出的数量。从零模型至饱和模型之间的模型就是研究模型。

2.3 嵌套模型的特点

1) 嵌套模型之间可以做统计检验上的比较。互相嵌套的两个模型之间的卡方值和自由度可以直接相减。

2) 嵌套模型对参数的估计,都用标准误来建立了置信区间,其每个参数都有独立的标准误和置信区间。

3) 可以对嵌套模型进行似然比检验(Likelihood Ratio Test)。

正是因为嵌套模型具有以上几个特点,因此,当一系列相关模型具有嵌套关系时,根据简单性和俭省化原则就可以在拟和指数都很好的嵌套模型中选择出最好的模型。例如,假设模型 M_k 嵌套于模型 M_l , M_k 与 M_l 之间对数据的拟和差异就可以通过卡方检验来进行评定。模型 M_k 与 M_l 之间的卡方值的差为 $\Delta\chi^2$,自由度之间的差为 Δdf 。如果 $\Delta\chi^2$ 相对于 Δdf 在某一置信水平上不显著,那么根据简单性和俭省化原则就可以选择模型 M_k 作为被选模型,因为模型 M_k 比 M_l 具有较少的参数。如果 $\Delta\chi^2$ 相对于 Δdf 在某一置信水平上达到显著水平,那么就要拒绝模型 M_k 而把模型 M_l 作为被选模型,因为模型 M_l 能更好拟和数据。当然在这个过程中也要考虑模型在理论上的解释意义。

2.4 Anderson 和 Gerbing 的嵌套模型的选择方法

Anderson 和 Gerbing(1988)在前人研究的基础上提出了一个设计嵌套模型和模型选择的方法^[4]。首先,设定五个嵌套模型,分别为饱和模型 M_s 、独立零模型 M_n 、理论模型 M_t ,它代表了研究者感兴趣的理论假设,限制模型 M_c 和非限制模型 M_u ,这两个模型分别是研究者所假设的其次感兴趣的理论模型。这

五个模型为嵌套模型,他们之间的嵌套关系如下: $M_n < M_c < M_t < M_u < M_s$ 。接下来,研究者首先检验是否存在拟和指数可以接受的模型,可用伪卡方检验(pseudo chi-square test)方法^[5](Bentler & Bonett, 1980)进行这一步的检验。具体做法是以模型 M_s 的卡方值(最小的卡方值)和独立模型 M_n 的自由度(最大的自由度)进行卡方检验。如果这一伪卡方检验显著,那么在这一系列的嵌套模型中,绝对找不到一个与数据拟和良好的模型,需要研究者重新修正除 M_s 和 M_n 以外的模型。如果伪卡方检验不显著,就可进行序列卡方差异检验(Sequential Chi-Square Difference Tests)。具体来讲就是研究者要对模型 M_c , M_t 和 M_u 进行估计,获得每个模型的似然比卡方统计值。由于嵌套模型的卡方统计值具有上面所说的特点,就可以进行下面的模型检验,进行模型选择。如果要接受模型 M_t , M_t 与 M_s 以及 M_t 与 M_u 相比应该不显著,而 M_c 与 M_t 相比则应该显著。如果要放弃模型 M_t ,接受模型 M_c ,那么模型 M_c 与 M_t 以及 M_c 与 M_s 相比都应当不显著。如果要放弃模型 M_t ,接受模型 M_u ,那么模型 M_c 与 M_t 以及 M_t 与 M_u 相比都应当显著,同时还有一个附加条件,即如果 M_t 与 M_s 相比显著,那么 M_u 与 M_s 与相比就应当不显著。

3 非嵌套情况下的模型选择问题

假如研究者认为几个模型都有道理,而且他们之间又不具有嵌套关系,可以利用以下的方法进行模型的比较和选择。

3.1 交叉效度分析

优良模型的一个很重要的标准是其概括力,即使用不同的样本都能很好地复制出同一个模型,交叉效度分析就是从这个角度来分析在一系列相关模型中谁是最好的模型。交叉效度分析是指对在一个样本数据中所探索出的模型再用另外一个样本来进行验证。如果原来样本容量很大,也可将原来样本分为两半,其中一半进行探索性模型修正,另一半留作验证模型分析。Cudeck & Browne(1983)不仅介绍了一种检验交叉效度的方法,还建议采用交叉效度拟合指数(Cross-Validation Fit Indices)来检验模型的交叉效度,以推论模型在其他样本和总体中的合理性^[6]。交叉拟和指数的公式为:

$$CVI = F(S_b, \sum(\hat{\theta}_{k|a}))$$

公式中的 S_b 是样本 b 的协方差矩阵, $\sum(\hat{\theta}_{k|a})$ 是样本 a 的模型所衍生的协方差矩阵。交叉效度拟和指数(CVI)是一个随机变量,因两次使用随机样本,随机误差较为复杂。为此, Browne & Cudeck(1989) 利用单个样本导出了 CVI 的数学期望的估计即 $ECVI^{[7]}$, 公式为:

$$ECVI = F_k(\theta) + \frac{2t}{N - 1}$$

公式中的 $F_k(\theta)$ 是模型 M_k 对样本的拟和函数的极小值, t 是 M_k 的自由参数的数目, N 是样本量。在 LISREL 统计软件中能够直接给出研究模型的 $ECVI$, 为了便于根据研究模型的 $ECVI$ 的大小来判断模型对观测变量间的关系的拟和度, LISREL 统计软件还同时输出饱和模型和独立模型的 $ECVI$ 的值。在嵌套模型序列中, 饱和模型约束最少, 拟和最好, 所以 $ECVI$ 的值最小。独立模型约束最多, 拟和最差, 所以 $ECVI$ 的值最大。因此, 在一般情况下, 介于二者之间的模型的 $ECVI$ 值愈接近饱和模型的 $ECVI$ 值, 拟和愈好, 愈接近独立模型的 $ECVI$ 值, 表示拟和愈差。

3.2 阿凯克信息准则(Akaike 's information criterion , AIC)

阿凯克信息准则是由 Akaike 根据统计信息理论提出的一个模型拟和指标, 它的计算方法^[8]是

$$AIC = \chi^2 + 2t$$

公式中的 χ^2 为假设模型 M_k 的检验统计量, t 是模型

中自由估计的参数的数目。AIC 是表示根据某一样本估计所得的模型参数在使用另一样本进行估计时的有效程度的指数, 从这种意义上讲, 它也是一种交叉效度指数。AIC 的值越小, 表示模型的拟和度越好, 而且模型很简约。 LISREL 程序也会直接给出饱和模型和独立模型的 AIC, M_k 的 AIC 愈接近饱和模型的 AIC, 说明模型拟和越好。

但是, 单个的 AIC 指数包含一些来源不清的恒定因素, 并且也受样本容量的影响。为此, Kenneth P. Burnham 和 David R. Anderson(2002) 提出了模型比较和模型选择中的指数^[11]: $\triangle AIC_i = AIC_i - AIC_{min}$, 公式中的 AIC_{min} 是一系列相关模型的不同 AIC 值中最小的一个值。通过这种转换可以看出, 最好模型的 $\triangle AIC$ 值为 0, 其他模型的 $\triangle AIC$ 值都为正数。在一系列相关的候选模型中, $\triangle AIC$ 可以提供模型之间比较的充分的证据。 $\triangle AIC$ 的解释规则如下: $\triangle AIC_i \leq 2$ 时, 此模型受到最大的支持; $4 \leq \triangle AIC_i \leq 7$ 时, 此模型的支持力度减少; $\triangle AIC > 10$ 时, 此模型不再受到支持。

下面通过一个例子, 来看一下 $\triangle AIC$ 在模型选择中的运用。在一个关于 5 ~ 7 儿童人格结构的研究^[9]中, 研究者为了验证 5 ~ 7 岁儿童的人格也是大五结构, 分别选取儿童样本和成人样本进行施测, 并把成人样本作为验证儿童人格模型的外部效标。作者除了对自己提出的儿童人格的大五直交模型和大五斜交模型进行检验外, 还对其他研究者所提出的儿童人格结构模型进行了检验, 具体结果见下表:

表 1 对儿童和成人的人格自我报告的竞争模型比较的 CFA 结果汇总

检验模型	大学生样本					6.7 岁儿童的联合数据				
	χ^2	df	$\triangle \chi^2$	AIC	$\triangle AIC$	χ^2	df	$\triangle \chi^2$	AIC	$\triangle AIC$
1. 单因素模型: 整体的自我评价	2757.9	779	759***	46729	775	1753.4	799	198***	12039	118
2. 两因素: 学术和社会情绪	2423.7	779	461***	46431	477	1670.4	799	115***	12057	196
3. 两因素: 社会化和中介	2751.0	779	788***	46722	476	1770.7	799	215***	12056	195
4. 大五直交模型	2274.5	779	312***	46426	292	1606.8	799	51***	11992	131
5. 大五斜交模型	1962.8	769	—	45954	0	1555.5	769	—	11861	0

(注: 资料来源于 Jeffrey R. Measelle, Oliver P. John, Jennifer C. Ablow, Philip A. Cowan and Carolyn P. Cowan, 2005)

从表 1 中的结果可以看出, 根据 Kenneth P. Burnham 和 David R. Anderson(2002) 的检验的原则, 唯一可以接受的模型是大五斜交模型, 其他模型都不受支持。

4 结束语

在心理学研究中, 模型的推断和选择是一个相当复杂的问题。在具体的实践中, 研究者们应当特别注意如下两个原则: 建构任何一个假设模型时都

需要有相应的理论基础 ,竖立多个模型假设的观念 ,因为能够拟和任何一个数据的模型会有很多 ,一个能够很好拟和数据的模型并不一定是要找的最好的模型。总的来说 ,只有在充分的理论研究的基础上 ,竖立多个模型假设的观念 ,借助相应的统计手段 ,才能找到一个最接近真实情况的最好的模型。

参考文献

1 Burnham Kenneth P ,David R. Anderson. Model selection and multimodel inference : A practical information - theoretical approach. 2th ed. New York : Springer - Verlag , 2002.

2 Lin W Y , Hau K T. Structural equation modeling : model equivalency and respecification. Educational Journal ,1995 23 : 147 - 162.

3 Keith F. Widaman ,Jane S. Thompson. On specifying the null model for incremental fit indices in structural equation modeling. Psychological Methods 2003 8(1) :16 - 37.

4 James C. Anderson ,David W G. Structural equation modeling in practice : A review and recommended two - step approach. Psychological Bulletin ,1988 ,103(3) 411 - 423.

5 Bentler P M ,Bonnett D G. Significant tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin , 1980 88(3) 588 - 606.

6 Cudeck R ,Browne M W. Cross - validation of covariance structures. Multivariate Behavioral Research ,1983 ,18 :147 - 167.

7 Cudeck R ,Henly S. Model selection in covariance structures analysis and the “ problem ” of sample size : A clarification. Psychological Bulletin ,1991 ,109 512 - 519.

8 王权 ,李金波. 实证性因素分析. 浙江大学出版社 2002.

9 Jeffrey R. Measelle ,Oliver P. John , Jennifer C. Ablow , et al. . Can children provide coherent ,stable , and valid self - reports on the big five dimensions ? A longitudinal study from ages 5 to 7. Journal of Personality and Social Psychology 2005 89(1) :90 - 106.

The Principal and Methods of Model Selection in Application of SEM

Liu Hengchao¹ ,Xu Yan¹ ,Wang Li²

(1. Department of Psychology ,Beijing Normal University ,Beijing 100875 ;

2. Center for Studies of Psychological Application , South China Normal University , Guangzhou 510631)

Abstract :With the application of structural equation model in psychological research , more and more researchers emphasize the development of theoretical model in researches. This paper discussed that the general principal and basic methods of model selection in psychological researches , it can be helpful for researchers.

Key words :model selection ; nested model ; non - nested model