第四章 模型估计与分析：结构方程模型

本章内容

第一节　结构方程模型概述
第二节　结构方程模型的组成
第三节　参数估计与识别问题
第四节　模型拟合评鉴
第五节　验证性因素分析
第六节　路径分析
第七节　结构方程模型：统合模型分析
第八节　中介与调节
第九节　结构方程模型的正确运用

第一节 结构方程模型概述

一、结构方程模型的特性

1. 外显变量与潜在变量

SEM的一个重要特性是能够对于抽象的构念（construct）进行估计与检验。社会与行为研究经常必须处理一些抽象的概念，针对抽象的概念，我们必须给予一个操作化的定义，以便通过该程序来得到具体的数据，用以反映不同个体在该概念上的强度。此时，我们具体测量的变量被称为**外显变量**（manifest variable）、观察变量（observed variable）或测量变量（measured variable）。这些可以直接获得数据的变量若是受到同一个潜在构念的影响，则会具有共同性，反映在变量之间的共变关系上。如果针对这些变量之间的共同性加以估计，所得到的能够反映该潜在构念的强度的数据被称为**潜在变量**（latent variable）。

在具体用来获得潜在变量的研究策略中，最典型的代表是因素分析（factor analysis：又可译为因子分析，factor表示潜在变量或构念）。因素分析利用一组测量同一个构念的观察变量来估计背后的潜在变量。传统因素分析针对的是观察变量的背后具有哪几个潜在变量，以及潜在变量与观察变量之间的关系为何。它无法事前预知，直到数据搜集完成之后，才进行变量间的共变关系分析，抽取出最适当的因素，确立一个最佳的因素结构模型，并为潜在变量命名。从这一程序进行的因素分析而得到的潜在因素是一种经验性的潜在变量，因而被称为探索性因素分析（exploratory factor analysis，EFA）。

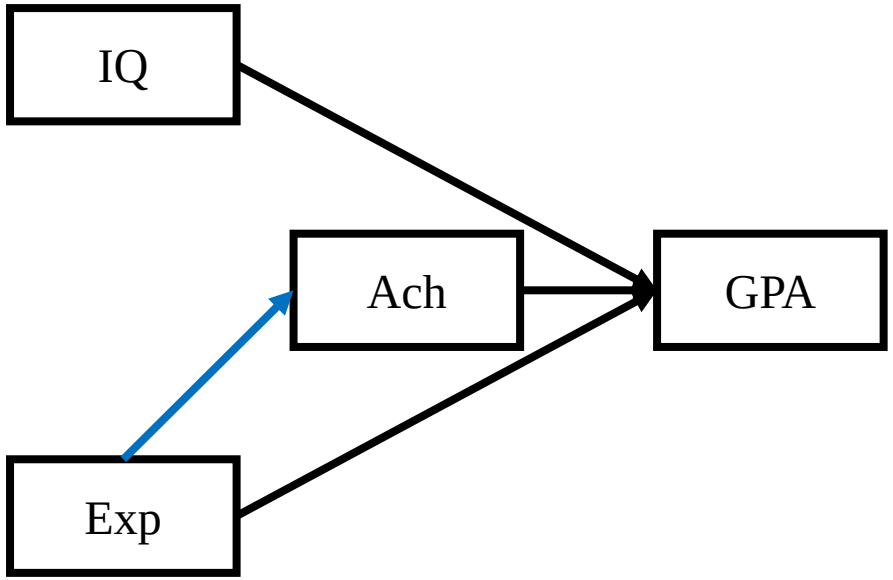
相对而言，在SEM中，潜在变量的概念与内涵基于理论的推导，且潜在变量与观察变量的关系是在资料搜集完成之前（事先）提出的假设性概念，然后通过实际搜集的数据，分析比对假设模型与观察到的数据之间的一致性或差异性，来决定研究者对于潜在变量所提出的假设性看法是否恰当，即模型拟合分析。以此种模型进行的因素分析称为验证性因素分析（confirmatory factor analysis，CFA），即一种先验性的、由事前的潜在变量定义的模型。

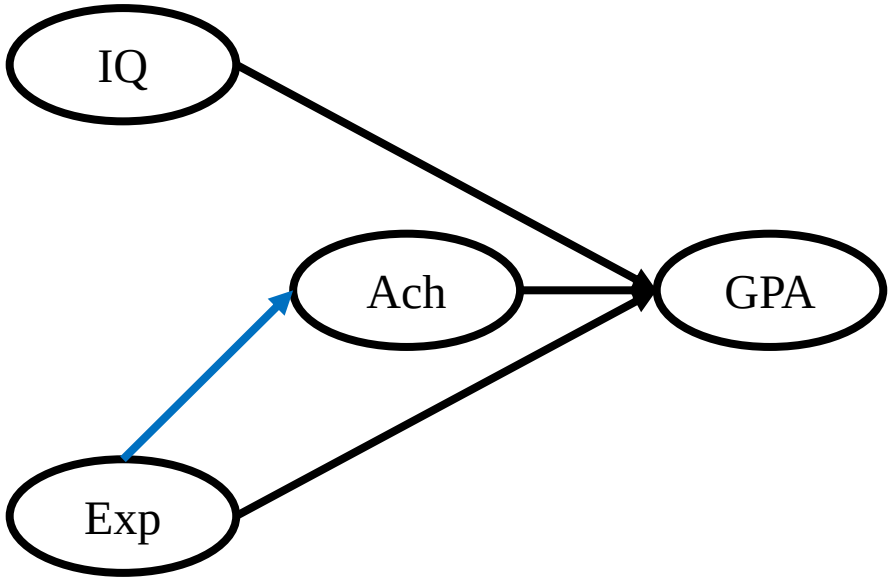
第一节 结构方程模型概述

一、结构方程模型的特性

2. 对变量关系的探究

社会及行为科学研究的变量关系通常并不是单纯地对一个变量的推论或两变量关系的讨论，而是涉及一组变量之间关系的讨论，这一组变量除了存在数学的、表面上的关系外，可能还存有潜在的因果性（causality）或阶层性（hierarchy）。例如，在一项对于学业表现的研究中，最常用的解释变量为智力。然而，研究者可能会考虑到除了智力（IQ）因素，学生先前的学习经验（Exp）也会影响学生的学习表现（GPA）。而先前学习经验的影响还可能基于成就动机（Ach）的中介作用，间接影响学业成就。从上述概念中，可以得到如下的研究方程：GPA=a×IQ+b×Exp+c×Ach+e1；Ach=d×Exp+e2。





第一节 结构方程模型概述

一、结构方程模型的特性

3. 模型比较分析

SEM的另一个特征是模组化分析的应用：SEM可以将一系列研究假设同时构造成一个有意义的假设模型（hypothetical model），然后经由统计程序对这一模型进行检验，在不同的模型之间则可进行竞争比较。

研究者往往会因为理论观点的不同而对同一组变量之间的假设关系产生不同的主张。因此，研究者可以基于不同的理论与假设前提，发展出不同的替代模型（alternative model），进行模型间的竞争比较。如此利用假设模型进行统计检验的优点是大大改善了传统路径分析对在多组回归等式间进行同时估计的限制，也提高了分析的应用广度。

Joreskog与Sorbom（1996）指出，SEM的模组化应用策略有三个层次：

① 单纯的验证（confirmatory）：针对单一的先验假设模型，评估其适切性，称为验证型研究；
② 模型的产生（model generation）：其程序是先设定一个起始模型，再与实际观察数据进行比较，从而进行必要的修正，反复进行估计程序以得到最佳拟合的模型，称为产生型研究；
③ 替代模型的竞争比较，以决定何者最能反映真实资料，称为竞争型研究（competitive modeling）。

第一节 结构方程模型概述

一、结构方程模型的特性

4. 结构方程模型的技术特性

①SEM具有理论先验性

SEM分析最重要的一个特性是它必须建立在一定的理论基础之上。也就是说，SEM是一个用以检验某一先期提出的理论模型（priori theoretical model）的适切性的统计技术。这也是SEM被视为一种验证性而非探索性统计方法的主要原因。在SEM的分析过程中，从变量内容的界定、变量关系的假设、参数的设定、模型的安排与修正，一直到应用分析软件来进行估计，其间的每一步都必须有清楚的理论概念或逻辑推理作为依据。从统计的原理来看，SEM也必须同时符合多项传统统计分析的基本假设（例如线性关系、正态性）以及SEM分析软件所特有的假设要件，否则所获得的统计数据无法采信。

以因素分析为例，SEM所使用的因素模型采取了相当严格的限制。对于潜在变量的内容与性质，研究者在测量之初就必须有非常明确的说明，或有具体的理论基础，并已先期决定了相对应的观察变量的组成模型。分析的进行即在考验这一先期提出的因素结构的适切性。除了在开发测量工具时可以利用这一程序来检验其结构的有效性，也用这一程序检验理论架构，因此称为验证性因素分析。

第一节 结构方程模型概述

一、结构方程模型的特性

4. 结构方程模型的技术特性

②SEM同时处理测量与分析问题

不论分析的内容为何，过去传统的统计方法多把变量视为“真实”“具体”“可观测”的测量数据，在分析过程中并不去处理测量过程所存在的问题，也就是说，“测量”与“统计”是两个独立分离的程序。传统上，如果变量所涉及的概念如同“智力”或“焦虑”等不易界定的心理概念，研究者为了获得可以分析的数据，会先行讨论测量的方法，并以信度与效度的概念程序先行进行评估。一旦通过评估的标准，即对所获得的测量数据进行分析。

相对于传统的做法，SEM是一套可以将“测量”与“分析”整合为一的计量研究技术。关键在于SEM以潜在变量的形式，利用对观察变量的模型化分析，对不可直接观察的构念或概念加以估计，不仅可以估计测量过程当中的误差，也可以评估测量的信度与效度（例如因素效度），甚至可以超越古典测量理论的一些基本假设，针对特定的测量现象（例如，误差的相关性）加以检测。另一方面，在探讨变量之间关系的时候，测量过程所产生的误差并没有被排除在外，而是同时包含在分析的过程当中，使得测量信度的概念可以整合到路径分析等统计推论的决策过程中。

第一节 结构方程模型概述

一、结构方程模型的特性

4. 结构方程模型的技术特性

③SEM以协方差的运用为核心，亦可处理平均数估计

SEM分析的核心概念是变量的协方差（covariance）。协方差是描述统计中的一种离散量数，利用方差（variance）的离均差和的数学原理，计算出两个连续变量配对分数的变异量，用以反映两个变量的共同变异或相互关联程度。协方差是一个非标准化的统计量数，受到两个变量所使用的量尺或单位的影响，数值可能介于-∞~+∞之间，如果将协方差除以两个变量的标准差，即可得出标准化协方差（即Pearson相关系数）。在SEM中，协方差具有两种功能：
第一种功能是描述性功能，利用变量之间的协方差矩阵可以观察出多个连续变量之间的关联情形；
第二种功能是验证性功能，用以反映理论模型所导出的协方差与实际观测得到的协方差的差异。

分析过程中最重要的数学程序即计算导出协方差矩阵（∑matrix）。如果研究者所设定的模型有问题，或是数据估计过程导致∑矩阵无法导出，整个SEM即无法完成。

除了协方差以外，SEM也可以处理变量的集中趋势的分析与比较，也就是平均数的检验。传统上，平均数的检验是以t检验或方差分析（ANOVA）进行的。由于SEM可以对于截距进行估计，使得SEM可以将平均数差异的比较纳入分析模型，同时若配合潜在变量的概念，SEM更可以估计潜在变量平均数，使应用范围更为广泛。

第一节 结构方程模型概述

一、结构方程模型的特性

4. 结构方程模型的技术特性

④SEM适用于大样本分析

由于SEM所处理的变量数目较多，变量之间的关系较为复杂，为了保证不违反统计假设，必须使用较大的样本量。同时，样本规模的大小也影响SEM分析的稳定性与各种指数的适用性。与其他统计技术一样，SEM分析所使用的样本规模当然是越大越好，但是所谓的最适合的规模会随着SEM模型的复杂度与分析的目的与种类而有相当大的变化。但是，一般来说，当样本量低于100时，几乎所有的SEM分析都是不稳定的。

Breckler (1990)曾针对人格与社会心理学领域的72个SEM实证研究进行分析，样本规模为40~8650，中位数为198。有1/4的研究的样本量小于500，约20%的研究样本的规模小于100。因此，一般而言，大于200的样本才可以称得上是一个中型的样本。若要追求稳定的SEM分析结果，低于200的样本量是不鼓励的。

如果样本量小是无法改变的既成事实，可以使用PLS-SEM。

第一节 结构方程模型概述

一、结构方程模型的特性

4. 结构方程模型的技术特性

⑤SEM包含了许多不同的统计技术

综观统计分析技术的内容，可以概略分为平均数检验的方差分析与探讨线性关系的回归分析两大范畴。事实上，这两者并无本质差异，前者可以被归为一般线性模型（general linear model）分析技术，后者则是以变量间的线性关系为分析的内容。随着计算机科技的发展以及分析软件功能的提升，两种统计模型已经可以互通，合而为一。

第一节 结构方程模型概述

一、结构方程模型的特性

4. 结构方程模型的技术特性

⑥SEM重视多重统计指标的运用

虽然SEM囊括多种不同的统计技术于一身，但是对于统计显著性的依赖性远不及一般统计分析，主要理由有三：

第一，SEM所处理的是整体模型的比较，因此所参考的指标不是单一的参数为主，而是整合性的系数，所以个别检验是否具有特定的统计显著性不是SEM分析的重点；

第二，SEM发展出多种不同的统计评估指标，使得使用者可以从不同的角度进行分析，避免过度倚赖单一指标；

第三，由于SEM涉及大样本分析，样本越大，SEM分析的核心概念——卡方统计量的显著性越会受到相当的扭曲，因此SEM的评估指数都特意避免碰触到卡方检验的显著性检验。也因为这个原因，在SEM分析当中，较少讨论与统计显著性决策有关的一类与二类错误议题，显示了SEM技术的优势在于整体层次（micro-level）而非个别或微观层次（macro-level）。

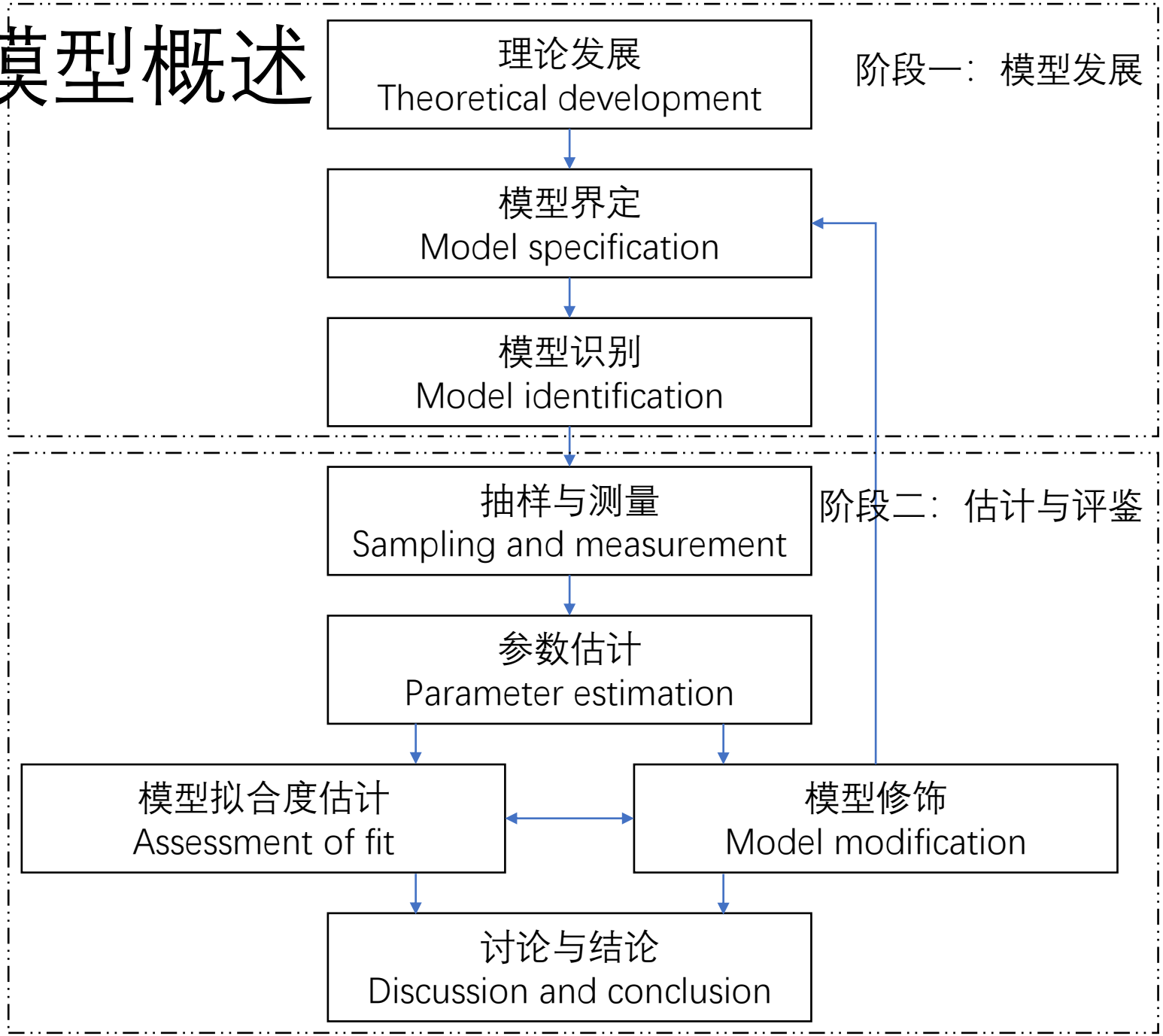
第一节 结构方程模型概述

二、结构方程模型的执行

SEM分析的基本程序可以分为模型发展与估计评鉴两个阶段：

前者发展SEM分析的原理基础并使SEM模型符合特定的技术要求，此时研究者的主要工作是概念推导与对SEM分析的技术原理的考虑；

后者则是产生SEM的计量数据来评估SEM模型的优劣好坏。有关SEM分析的执行流程如右图所示。



第一节 结构方程模型概述

二、结构方程模型的执行

1. 模型发展阶段

模型发展阶段的主要目的是建立一个适用SEM分析概念与技术需要的假设模型，牵涉理论发展、模型界定与模型识别等三个概念。在图中，这三个概念虽然是以连续的流程图来表示，但是三者间的关系只是说明概念发生的先后顺序，在实际操作上，这三个概念的运作则是一个相互作用的不断往复的过程。

SEM模型的建立必定以理论为基础。所谓的以理论为基础，并不是说SEM模型必须建立在某一个特定的理论之上，而是强调SEM模型的建立必须经过观念的厘清、文献的整理与推导或是研究假设的发展等理论性的辩证与演绎过程，最终提出一套有待检验的假设模型。前文已经指出了SEM的一个重要特性是理论的先验性，因此，SEM分析的第一个阶段的主要目的便是建构SEM的理论基础。另外两个概念——模型界定与模型识别——即是在理论性推演过程的基础上，将SEM模型的理论假设转换成适当的技术语言。

第一节 结构方程模型概述

二、结构方程模型的执行

1. 模型发展阶段

为配合理论推导的过程，研究者同时必须进行模型界定的重要工作。模型界定可以说是第一个阶段当中最为具体的步骤，目的是发展可供SEM进行检验与估计的变量关系与假设模型。一般来说，SEM模型的构成主要有两个部分：第一部分是理论或概念的基础，或是研究者个人的先备知识与经验；第二部分是SEM的技术语言与方法要求。当研究者面对所关心的研究问题时，除了基于自己的知识基础与研究兴趣来推演出值得探讨的研究命题之外，还需对于理论文献详加检阅，以建立严谨的科学假设或提出有待验证的理论模型。

如果研究者选择使用SEM来探讨所提出的假设与理论模型，则必须配合特定的SEM技术语言与各项操作要求，将研究者所提出的假设与理论模型转换成SEM模型。与此同时也需考虑SEM分析当中可能涉及的各种统计概念所使用的统计原理，并将其纳入SEM模型的设定之中。这个由理论发展到技术性模型建立的一整套程序就是“模型界定”的主要任务。而模型界定的具体产品是建立一个SEM路径图。该路径图就是模型识别步骤据以评估的依据，也是第二阶段进行估计分析的地图。

第一节 结构方程模型概述

二、结构方程模型的执行

1. 模型发展阶段

在模型界定的过程中有一个非常重要的技术问题，是必须让SEM模型具有可识别性，使SEM的各项数学估计程序可以顺利地进行。SEM设定的假设模型是基于研究者的研究需求而提出的，但是模型的分析必须利用实际搜集得到的数据，利用分析软件来进行估计工作。只有在模型符合统计分析与软件执行的要求（也就是能够被有效识别）的情况下，SEM分析才能顺利进行。此时，一个模型可以被有效进行识别的程度称为模型识别度（identifiability）。有关模型识别度的估算过程是“模型识别”的主要任务。

第一节 结构方程模型概述

二、结构方程模型的执行

2. 估计与评鉴阶段

一旦SEM模型发展完成，研究者必须搜集实际的测量资料来检验所提出的概念模型的适当性。这一阶段开始于样本的建立与测量工作的进行，所获得的观察资料经过处理后，即依照SEM分析工具的要求进行各项估计。

样本的性质对于SEM分析有重要影响。除了样本规模大小的影响，由于SEM涉及潜在变量的测量，因此SEM分析的结果也与样本结构及测量质量有密切的关系，也就是具有样本依赖性。Bollen (2002)指出，对于某些个体而言，潜在变量可能是有意义的概念，但是对另一些个体来说可能不然，或有不同的意义。因此在不同的样本间，SEM分析可能得到非常不同的结果。

由于SEM随着样本的特性而改变结果的性质，SEM的分析工具多提供了评鉴的指标，以反映样本规模与性质的影响。同时，SEM分析本身亦可以处理对测量误差的估计，使测量质量的影响可以被有效地控制。但是，研究者仍然必须谨慎选取研究样本，维护测量的质量，因为SEM分析的复杂性高，任何一个环节的瑕疵或失误都可能造成SEM分析结果产生变化。

第一节 结构方程模型概述

二、结构方程模型的执行

2. 估计与评鉴阶段

SEM的参数估计可以说完全是由计算机进行的，只有少数部分必须由人工计算完成（例如，测量模型的信度估计），但是如何让SEM分析顺利完成，仍有赖研究者对计算机软件正确无误地下达指令，以及对分析工具的各种选项的正确选择。不同的分析软件各有其优劣，因此操作方法各有不同。常用软件包括LISREL、Mplus、Amos、EQS、R、Python、Stata等。SEM的使用者必须对于SEM分析工具的一般性基本原理有深入的了解，也必须从操作演练中积累经验，熟知每一套软件的优劣与限制，了解每一个参数或警告讯息的意义与作用，如此才能顺利完成各项估计与评估程序。

值得注意的是，在估计与评鉴过程中，SEM分析工具通常会提供模型调整与修饰的计量信息，使用者可以根据这些指数或统计检验数据调整先前提出的假设模型，重新反复进行估计与模型评估，这一过程称为模型修饰。虽然这一做法违反了SEM分析理论先验性的精神，但是观察数据背后潜藏的各种信息也是科学研究相当珍贵的线索，从中既可能看出研究者在理论推导过程当中的疏忽或盲点，也可能引导研究者继续推导出更有意义的概念或假设，重新提出一套更趋合理的SEM模型。因此，模型的修正步骤一般也是SEM使用者相当重视的部分。

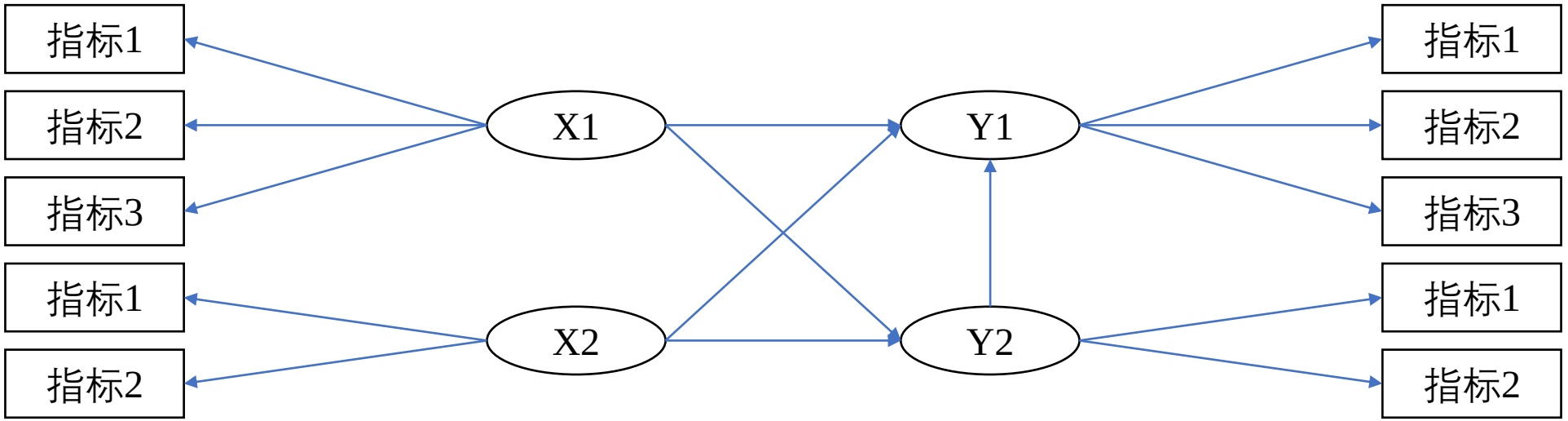
第一节 结构方程模型概述

三、结构方程模型的执行重点

1. 模型的描述与设定

典型的SEM分析由一系列代表特定研究假设的回归方程组成。一个SEM模型不仅牵涉不同类型的变量的处理（例如观察变量与潜在变量、内生变量与外源变量等），也涉及不同关系形态的设定（例如回归、相关与残差估计等）。因此，SEM分析首先重视的就是模型的正确设定与呈现，尤其重视概念路径图的运用，路径图不仅可以协助研究者呈现他们的研究架构，更可以促成研究者与其他读者之间的沟通。

在SEM的应用中，模型有两种呈现形式：概念模型（conceptual model）与统计模型（statistical model）。概念模型主要在说明一个SEM研究当中所探讨的概念间的关系，可以利用路径图呈现。



第一节 结构方程模型概述

三、结构方程模型的执行重点

1. 模型的描述与设定

统计模型则是指可以利用分析工具加以检测的操作性模型（operational model）。这里所谓的操作性模型，其意义类似操作性定义（operational definition），表示经过操作程序得到概念结果的整个程序。而此时的操作性除了指符合统计运算的原理外，同时也需符合统计分析工具的技术原理与限制。

统计模型与概念模型的差别是概念模型可以利用路径图来具体说明，但是统计模型通常是进行SEM分析过程中的一些思考概念与计算原则，而不是具体的实体模型。

通常，统计模型最具体的呈现就是语法指令的内容。例如，在一个LISREL语法当中会详细列举哪些参数是被估计参数、哪些参数是固定参数等讯息。然而统计模型的真正内涵是概念模型得以进行估计分析的整个统计过程。

第一节 结构方程模型概述

三、结构方程模型的执行重点

2. 资料的准备

用于SEM分析的数据可以是原始数据或矩阵数据。SEM分析的始祖LISREL最早使用的数据形态是矩阵。由于影响数据质量的原因很多，因此在进行SEM分析之前，也必须先行进行资料完整性与正确性的筛检，使得SEM分析可以在一个干净、完整、正确的资料基础上进行。

①矩阵数据的应用

Cudeck (1989)建议，SEM分析最好使用协方差矩阵，而非相关系数矩阵：协方差矩阵能够涵盖相关系数矩阵，最重要的是它能够推导出SEM分析所需的各种重要数据，例如变量的方差与协方差等。

②原始数据的使用

SEM分析也可以直接读取原始资料来进行分析。当研究者使用的变量包括了类别或顺序变量时，即无法产生协方差矩阵或相关矩阵，此时必须先行处理这些非等距数据，使其在格式上符合SEM分析的需求。这时原始数据就显得格外重要。

第一节 结构方程模型概述

三、结构方程模型的执行重点

2. 资料的准备

③变量的分布特性

SEM分析必须建立在一定的统计假设基础之上，当违反统计假设时，SEM分析的数据是值得怀疑的。因此，一般撰写研究论文时，应列举变量的分布特征与统计假设的检验结果，尤其是当研究者以矩阵数据作为输入数据时，由于缺乏各变量的原始数据以助判断变量的分布特性，因此更需要揭示各变量的频数分布的特征，证明变量的偏态与峰度处在合宜的水平，或是没有遗漏与偏离的状况。

Hoyle与Panter (1995)建议，在撰写研究报告时，应说明变量的正态性、多变量正态性以及峰度的数据，因为某些估计程序明显受到正态性不足的影响，例如最大似然法与一般化最小平方法程序。所以，完整清楚地列举检验结果是保证研究报告正确性的一个负责任的做法。

第一节 结构方程模型概述

三、结构方程模型的执行重点

3. 报表的整理与分析

SEM分析涉及一连串的参数估计、模型检验与模型修饰的技术处理程序。因此，使用者必须非常熟悉每一个步骤的原理与目的，才能理解报表的内容与分析说明的重点。

在阅读分析软件的报表时，应分别就两个层次的数据进行处理：第一层是过程性的资料，也就是在完成最终结论之前，我们必须详细检阅SEM分析的各项数据，观察这些数据的状态，必要时加以记录，以备撰写报告之用；第二层是最终解（final solution）的报告，也就是SEM分析的最后结论的各种参数数据，以及模型拟合度的最终数据。这两项处理的完成有下列重要原则。

①估计方法的选择

SEM分析可以用不同的估计方法进行参数估计，而不同的方法所得到的结果也有所不同。因此，在SEM的研究报告中应说明使用何种估计策略，并说明为配合该种策略，有无特殊的处理（例如关于样本规模的决定、变量经过正态校正等），使得读者可以清楚地了解SEM的各项参数是在何种基础上估计出来的。最常用的估计方法是最大似然法（maximum likelihood，ML）。

第一节 结构方程模型概述

三、结构方程模型的执行重点

3. 报表的整理与分析

②模型拟合指数

其功能是评估一个SEM模型是否与观测数据相拟合。在SEM的具体应用上，拟合有两种意义:

第一种是绝对拟合（absolute fit）：反映的是模型导出的协方差矩阵与实际观测的协方差矩阵之间的拟合情形，拟合度的数值大小表示模型导出数与实际观测数差异的多寡;

第二种是增量拟合（incremental fit）：是指某一个模型的拟合度与另一个替代模型的拟合度相比，增加或减少了多少拟合度。例如，一个模型假设潜在变量之间具有相关，替代模型则假设潜在变量之间没有相关（称为虚无或独立模型），计算出两个模型的拟合度差异量后，推知何者较能拟合观测资料。

这两种拟合度的概念适用不同的模型拟合指数。然而，所有的SEM分析都应先报告卡方统计量以及与卡方统计量的计算有关的讯息（自由度、样本量、显著性数据）。如果是经过校正的非正态性数据，在报告传统的卡方值之外，还应报告调整后的卡方值（Scaled X^2）。

除此之外，拟合指数[goodness-of-fit (GFI) index]可以说是每一个SEM研究都会报告的数据。因为GFI指数的性质类似于回归分析的R^2，数值越大表示实际观察的协方差矩阵能够被假设模型解释的比例越高，模型拟合度越佳。GFI指数可以说是反映绝对拟合的最佳指数。

第一节 结构方程模型概述

三、结构方程模型的执行重点

3. 报表的整理与分析

②模型拟合指数

增量拟合的评估可以利用NNFI、IFI等指数，这些指数的基础是模型间的卡方值差异值，也就是Hu与Bentler (1995)所称的第二类指数。如果研究者使用的是ML估计程序，NNFI[或称为TLI指数(Tucker & Lewis's index)]是较常用的指数；但是当样本量少时（例如低于150）则不建议使用，此时可以改用IFI指数。如果研究者采用的是GLS估计方法，则IFI指数的表现较理想。

若以非中央卡方为基础来比较模型增量拟合[Ho与Bentler (1995)所称的第三类型指数]，较佳的选择是CFI指数（又称为BFI指数或RNI指数）；RMSEA指数则是近年来逐渐被普遍采用的指数，因为RMSEA是第三类的非中央卡方指数当中不受样本分布影响的指数。Hu与Bentler (1999)主张CFI指数与RMSEA指数都需报告在论文中。当研究者想去估计统计检验力时，RMSEA指数是非常适合的。当研究者想要比较不同的模型，但是没有嵌套关系时，则可使用ECVI、AIC或CAIC指数。

在呈现这些数据时，如果分析的模型很多，可以利用表格来整理呈现，做到一目了然。在论文的文字叙述中，可以写为X^2 (128, N=284) =506.23, p<0.001, NNFI=0.89, CFI=0.91的形态。在呈现X^2数值时应一并报告自由度（e.g., 128）与样本量（e.g., 284）数据，然后再就数值的内容与意义加以说明。

第一节 结构方程模型概述

三、结构方程模型的执行重点

3. 报表的整理与分析

③参数的报告

当模型拟合指数显示某一个模型是适合的模型之后，研究者应着手整理从该模型估计的最终解当中得出的各参数数据。参数的报告应该尽可能充分翔实，使得读者可以清楚地看出每一个参数的特性与代表的意义，三种重要的讯息——参数的合理性、显著性检验和标准化解——都应完整地揭示。

首先，参数估计的合理性反映的是该参数是否符合数学或统计学理上的可能性，或是实证资料的可能性。一般而言，参数的方差是衡量参数估计数最重要的资料。当残差的方差出现了负值（称为Heywood现象）或是超过范围的协方差（当标准化的估计数大于1时）时，表示参数估计是有问题的。在报告SEM分析结果时，如果有方差的数值，都应在报告中予以揭示。

其次，各参数的显著性检验数据应翔实列举。除了指出检验值的大小与显著性以外（例如t检验值与p值），标准误也是重要的数据。从标准误的大小中可以看出参数估计是否存在潜在的问题。值得注意的是，在SEM模型当中，可能有某些参数被设定为固定值（例如，被用来作为潜在变量量尺化的因素载荷通常被设定为1.00），因此没有经过估计与显著性检验，在说明显著性结果时，亦应予以标注说明。

最后，标准化解的呈现通常是SEM分析最重要的一部分，因为标准化解反映了SEM模型中各参数估计的最后结果，而且是以标准化的形式出现的。

第一节 结构方程模型概述

三、结构方程模型的执行重点

3. 报表的整理与分析

④标准化解与完全标准化解

值得注意的是，LISREL分析会产生两种不同的标准化解：

一是以SS（standard solution）指令所获得的标准化解；

二是以SC（solution completely standardized）指令得到的完全标准化解。

标准化解的数学原理是针对潜在变量的方差估计数进行标准化，也就是将潜在变量的估计数除以潜在变量的估计标准差（以去除每一个潜在变量各有不同的变异情形的影响），然后计算出所有参数的数值。

完全标准化解则是除了将潜在变量的数据加以标准化之外，还将观察变量的估计数加以标准化。也就是将每一个观察变量的数据除以各变量的标准差（以去除每一个观察变量各有不同的变异情形的影响），然后计算出所有参数的标准化估计数。

第一节 结构方程模型概述

三、结构方程模型的执行重点

4. 替代模型的使用

①演绎取向的替代模型

最理想的SEM模型是基于理论观点提出的模型，因此替代模型的提出最好是能够运用演绎取向的策略。Joreskog (1993)指出，任何一个SEM模型都可能具有其他的替代模型。因此，SEM研究应善用替代模型的相互比较，来决定哪一个模型最能够反映实际观察数据，Joreskog (1993)称之为替代模型取向（alternative models approach）的SEM研究。但是，替代模型的提出是基于理论上或概念上的考虑，因此都是在分析进行之前即已经提出。

演绎取向的替代模型最重要的价值在于具有先验的理论基础。此外，也正由于模型的建立是在数据分析之前进行的，因此模型的设定可以经过详细的计算与安排，不受限于资料的计量特性。在结果的分析上，每一个替代模型的重要性和理论内涵都十分清晰明确；在操作层面上看，可以免除许多人为操纵的疑虑，减少犯错的可能。

第一节 结构方程模型概述

三、结构方程模型的执行重点

4. 替代模型的使用

②归纳取向的替代模型

如果说演绎取向的替代模型是理论概念的产物，那么归纳取向的替代模型就是现实世界的产物，虽然可能缺乏理论的正当性，但是归纳取向所提出的替代模型能够反映真实数据的特性，得到的结果最符合真实世界的描述。此类研究被Joreskog (1993)称为模型产生取向（model generating approach）的SEM研究。

在某些情况下，归纳取向的替代模型分析有其必要性。例如，当样本规模很小时，或是某些变量的测量质量较差时，参数估计的稳定性不佳，演绎取向所提出的替代模型之间的比较难以进行。此时，适当地利用参数的估计结果进行修正，可以让参数估计较为顺利。

其次，当研究的性质偏向探索性研究时，归纳取向的替代模型可以较演绎取向的替代模型提出更多有建设性的信息。尤其在应用研究领域，例如，教育研究、消费调查等，理论的引导性不如实际现象与数据具有启发性。此时，从研究数据反映的各项修饰建议所累积出来的假设模型更具有解释力。

第二节 结构方程模型的组成

一、结构方程模型的变量

1. 变量的特性

相对于常数（constant）而言，变量（variable）反映了变动与差异，指某一属性因时间、地点和人物的不同而不同，或是在质或量上的个别差异。基本上，变量可分成两大类：连续变量与类别变量。其中，连续变量指利用等距或比率尺度等有特定单位的工具测量得到的变量，变量中的每一个数值皆代表强度上的意义，又称为量化变量（quantitative variable ）。

相对而言，当变量数值所代表的意义为质性的概念时，则称这类变量为类别变量（categorical variable）或质性变量（qualitative variable）。类别变量的重点在于分门别类，而非反映强度，因此变量数值仅具有相等或不等的数学特性，或是大于与小于的顺序关系，这类变量无法像一般的变量那样以数学算式计算出有意义的统计量，仅能利用数值的频数（frequency）或百分比（percentage）来进行数据的分析与检验。相对而言，连续变量的数值反映了被观察现象或特质的程度大小，因此变量数值可通过加减乘除等数学运算法则获得各种统计量（例如，平均数、方差、标准分数等），进行进一步的统计分析。

由于SEM是一套用来分析变量间复杂的共变关系的统计方法，因此SEM的基本组成单位是连续变量，类别变量仅作为辅助或分组讨论的调节变量之用。换言之，在一个SEM模型中，凡是被视为变量的，就存在平均数与方差的统计量信息，变量之间的关系由协方差反映。

第二节 结构方程模型的组成

一、结构方程模型的变量

2. 测量变量与潜在变量

在SEM模型当中，变量有两种基本的形态：测量变量与潜在变量。研究者测量得到的测量变量资料是真正被SEM用来分析与计算的基本元素；潜在变量则是由测量变量推估出来的变量。在典型的SEM分析中，测量变量的变异受某一个或某几个潜在变量的影响，因此又被称为潜在变量的测量指标或外显变量。

在SEM的路径图中，测量变量以长方形来表示。当测量变量是一个无误差的测量之时，我们可以视此变量为一个真实有效的测量。通常，人口学变量属于这一类型，例如，性别、年龄等。另一类测量变量可能伴随一定的测量误差，或是反映某种抽象的概念内涵，以SEM的术语来说，这些测量变量受特定潜在变量影响，这使得测量变量分数呈现高低变化。

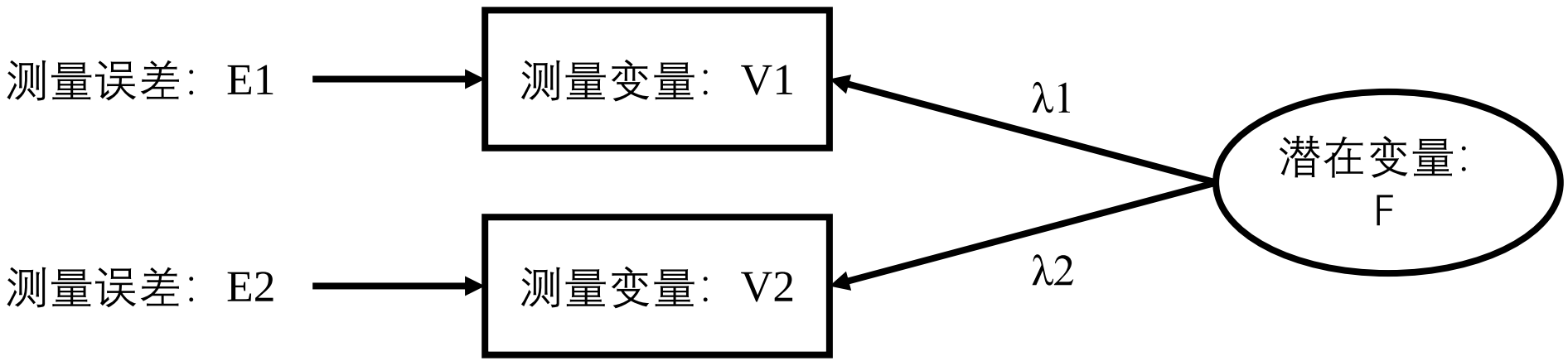
在SEM分析的路径图中，潜在变量以椭圆形的符号来表示。由于潜在变量无法由单一变量反映其抽象内容，必须通过测量变量来推估，因此一个潜在变量必须以两个以上的测量变量来估计，称为多元指标原则。不同变量之间的协方差反映了潜在变量的共同影响。在一个SEM模型中，测量变量一定存在，但潜在变量不可能单独存在，因为在研究过程当中，潜在变量并不是真实存在的变量，而是由测量变量估计出来的。

潜在变量的另一个特性是具有测量误差，测量变量的变异无法被共同的潜在变量充分解释的部分称为测量误差。测量误差可以被视为一个潜在变量，其平均数为零，变异量则可被估计，因此在路径图中有时会以椭圆形符号来表示残差（表示一个潜在变量），例如，Amos软件的绘图。

第二节 结构方程模型的组成

一、结构方程模型的变量

2. 测量变量与潜在变量



第二节 结构方程模型的组成

一、结构方程模型的变量

3. 内生变量与外源变量

沿用过去路径分析的术语，SEM模型中的变量可以分为内生变量与外源变量。所谓内生变量（endogenous variables）是指模型当中会受到任何一个其他变量影响的变量，也就是路径图中会受到任何一个其他变量以单箭头指涉的变量。外源变量（exogenous variables）则是模型当中不受任何其他变量影响但影响其他变量的变量，也就是路径图中会指向任何一个其他变量，但不被任何变量以单箭头指涉的变量。

在SEM模型中，如果在以内生及外源变量区分的基础上加上前述测量变量与潜在变量之分，那么SEM中的变量可以分为内生测量变量、外源测量变量、内生潜在变量与外源潜在变量四种类型。当一个潜在变量作为内生变量时，称为内生潜在变量（以希腊字母η表示，读作eta），它所影响的测量变量则称为内生测量变量（以y表示）；相对的，当一个潜在变量作为外源变量时，称为外源潜在变量（以希腊字母ξ表示，读作ksi），它所影响的测量变量则称为外源测量变量（以x表示）。

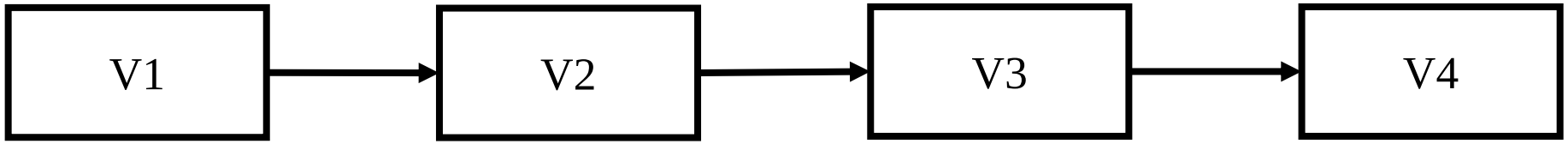
若以传统的自变量（independent variable）与因变量（dependent variable）的关系来看，外源变量因为不受其他变量影响，所以必为自变量，而内生变量多做因变量，但也可能作为影响其他变量的自变量。在回归分析中，因变量被视为被解释变量（或称为效标变量），是一个研究关心的焦点，因此SEM中的内生变量可以说是模型中的关键变量。

第二节 结构方程模型的组成

一、结构方程模型的变量

3. 内生变量与外源变量

当内生变量同时做因变量与自变量之时，如图中的V2与V3，表示该变量不仅被他人影响，还可能对其他变量产生作用，即成了一个中介变量（mediator）。由此可知，内生变量之间的关系与其本身的参数估计是SEM分析中相当重要的部分。



内生变量的一个重要性质是具有残差，因为内生变量的变异量不一定能够被模型当中的其他变量完全解释，而其他变量解释内生变量的不足之处即为残差。对于测量变量而言，其变异量无法被完全解释的残差部分称为测量残差或独特变异量（uniqueness）。如果内生测量变量存在于测量模型当中作为潜在因素的测量指标，残差可被视为测量误差。

对于潜在变量而言，内生潜在变量的变异量已经排除了测量误差的影响，因此无法被完全解释的残差部分不能被视为测量误差，而是其他变量无法解释的独特变异量。这部分变异量所反映的是模型无法有效解释内生潜在变量的部分，也就是过去回归分析的$1-R^2$，解释因素必定来自模型之外，因此被称为SEM模型的干扰项（disturbance）。由于它是无法被其他变量有效解释的变异量，因此也可以称之为解释残差或预测残差。

第二节 结构方程模型的组成

二、结构方程模型的参数

1. 参数的概念

在统计学当中，参数（parameter）是一个相当重要但很抽象的名词。参数的概念与推论统计有关，带有“未知”与“估计”这两个基本特质。

有时，研究者想要通过样本数据（已知）对于总体特性（未知）进行了解，例如，如果我们想要了解现在学生拥有零用钱的状况，我们可以从一群学生（样本）的每月零用钱的平均数去推估其他所有学生拥有零用钱的状况（总体）。此时，每月零用钱的平均数称为样本统计量，所有学生每月零用钱的平均数是我们想要估计得知的参数。值得注意的是，从统计的观点来看，参数所指的是一个未知而需要进行推估的量数，是一个计量的概念，而非总体本身。有时候，参数会被误当作总体的代名词，这是不正确的。

研究者所关心的是潜在变量的抽取或是对变量之间的关系的探讨，因此参数多半与潜在变量本身有关（例如，对总体的平均数或方差的估计），或是反映变量之间关系的参数（例如，对因素载荷、路径系数或协方差的估计）。研究者所拥有的只是样本观察数据的相关矩阵或协方差结构，再利用一些SEM软件，可以估计这些参数数值，并进行显著性的假设检验。

第二节 结构方程模型的组成

二、结构方程模型的参数

2. 自由参数、固定参数与限定参数

除了参数的性质之外，在SEM当中，参数就其是否有估计的需要而分成三种类型：自由参数（free parameter）、固定参数（fixed parameter）与限定参数（constrained parameter）。在SEM模型中，联结变量与变量之间关系的关键就是参数，参数的强弱大小必须通过统计程序加以估计，因此都是自由参数。当SEM模型中需要估计的自由参数越少时，称为越简效（既简单又有效率）的模型，自由度越大。

其次，SEM模型当中不被估计的参数将被设定为0，因此也被视为固定参数（固定为0）。因为某些理由被设定为常数（通常是1）而不被估计者，亦被称为固定参数。

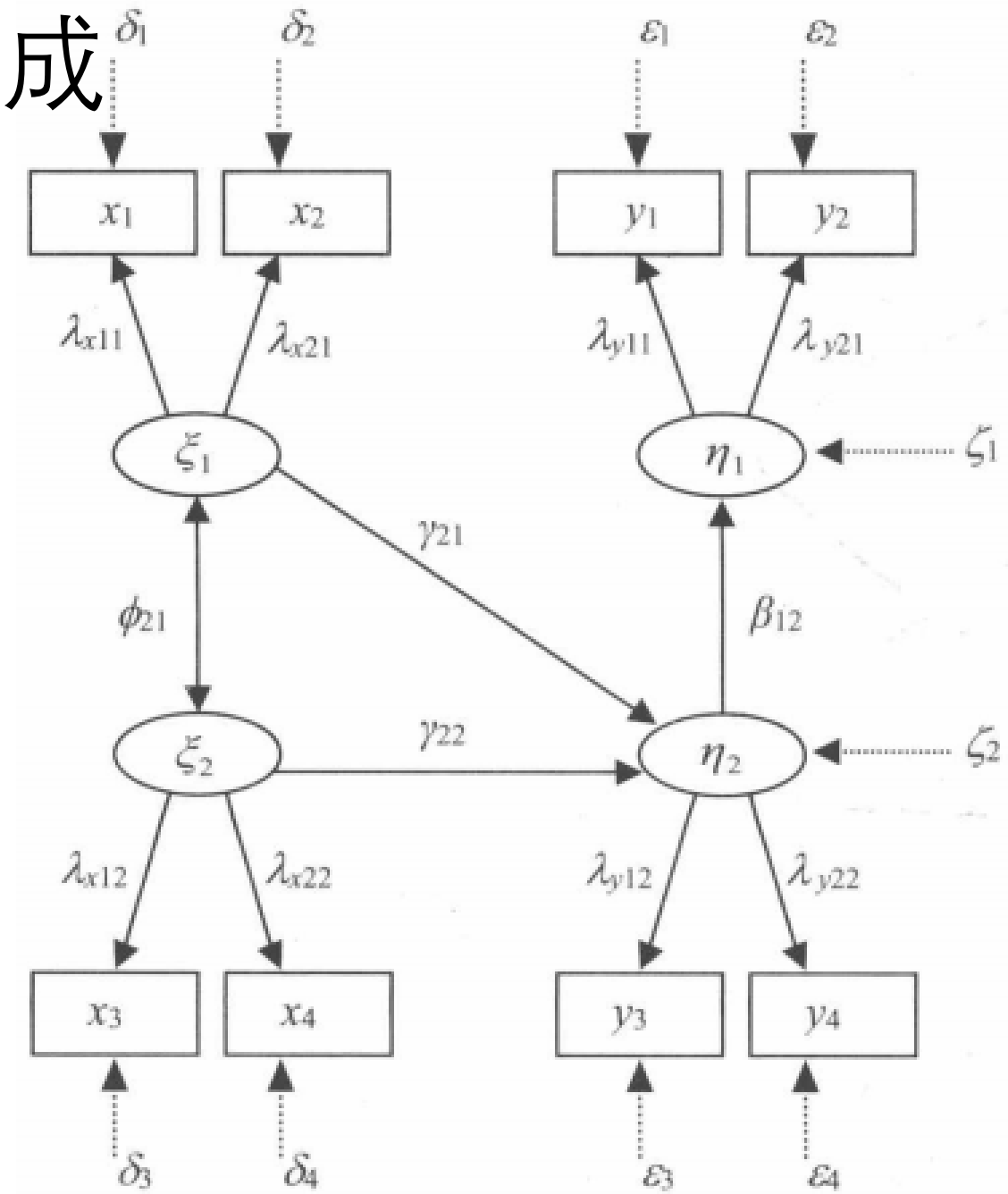
限定参数的使用多半与多样本间的比较有关，例如，某一个参数在甲样本与乙样本间被设定为等值（equivalent）的，此时的SEM对于这两个参数仅进行了一次估计，是为限定参数。限制参数通常是研究者基于特殊的需要而提出的，因此多与特定的理论或逻辑推导有关。从概念上来看，限定参数介于自由参数与固定参数之间，可被视为半自由参数。但是由于限定参数的数据仍然是由估计得出的，因此限定参数与自由参数被视为模型当中必须进行估计的参数。

第二节 结构方程模型的组成

二、结构方程模型的参数

3. 直接与非直接关系

在SEM模型中，参数所联结的变量关系有直接（方向性）关系（directional relationship）与非直接（非方向性）关系（non-directional relationship）两种类型。直接关系表示参数带有特定的影响方向，变量之间具有假设性的线性因果或预测关系，在路径图中以单向箭头（→）来表示。非直接关系则表示参数不带有特定方向，即变量之间虽然具有关系，但影响方向无法辨认，在SEM路径图中，以带有双箭头的线段或曲线表示。



第二节 结构方程模型的组成

二、结构方程模型的参数

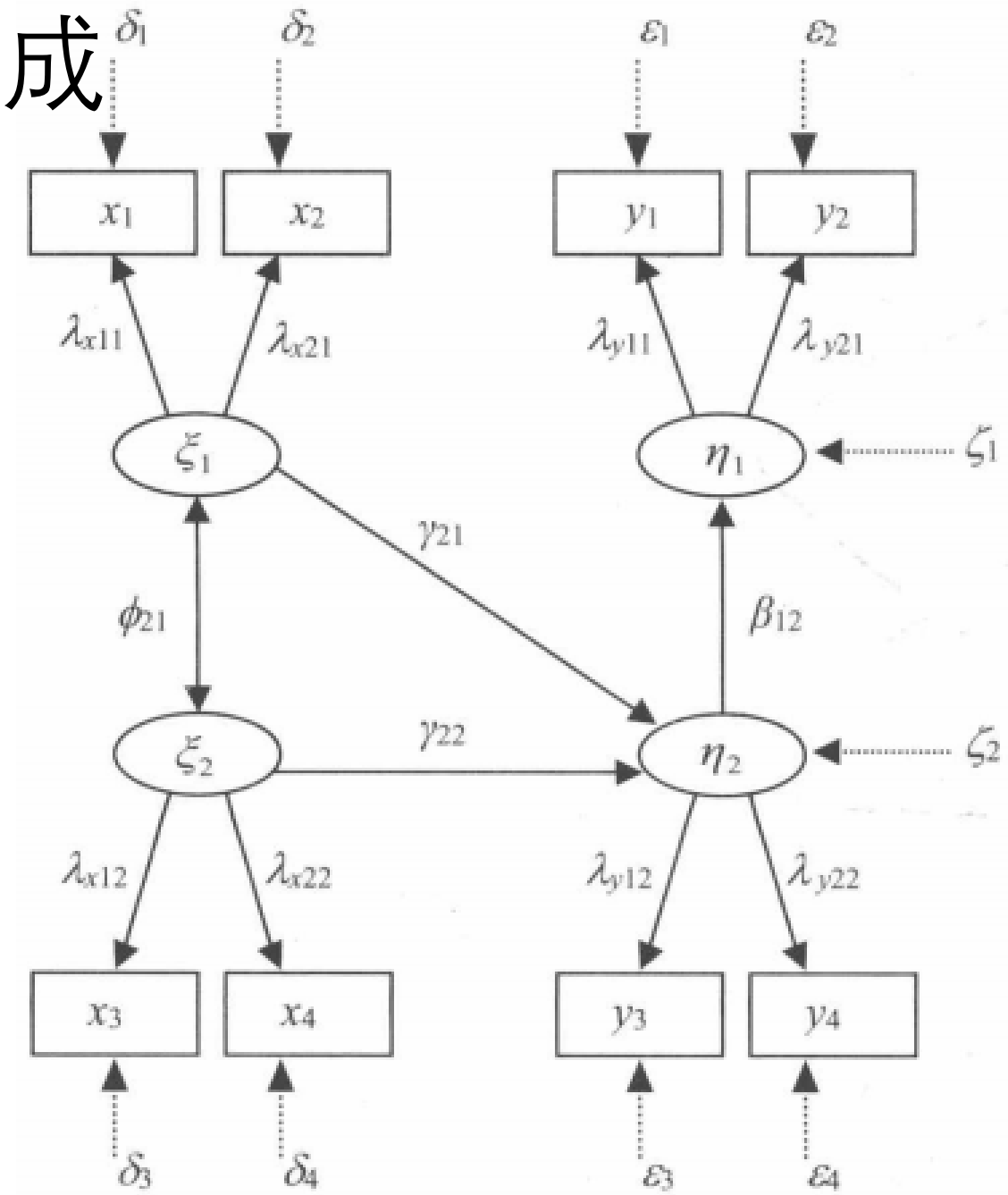
4. 模型参数与方程

一个完整的SEM模型包括:

- 测量模型（measurement model）：实际测量变量与潜在变量（特质）的的相互关系，构成模型的参数被称为测量模型参数;
- 结构模型（structure model）：说明潜在变量之间的关系，构成模型的参数被称为结构模型参数。

在右图中，SEM模型被分成左侧的外源变量关系与右侧的内生变量关系两部分，潜在变量与观察变量的关联强度由λ（lambda）参数表示，又称为因素载荷（factor loading），所构成的模型即为测量模型。外源与内生潜在变量之间的关系则以γ（gamma）参数反映，内生潜在变量之间的关系由β（beta）参数反映，由γ与β参数联结构成的模型即是结构模型。

模型中的δ（delta）与ε（epslon）分别表示外源观察变量与内生观察变量被潜在变量解释不完全的测量残差，ζ（zeta）则为内生潜在变量无法被完全解释的估计误差。残差的单箭头表示其方差可以被估计，而残差之间的共变关系则没有在图中描述出来，但它可能存在，可被视为自由估计参数来估计。



第二节 结构方程模型的组成

二、结构方程模型的参数

4. 模型参数与方程

上面所提及的各种参数与变量的关系可以利用一般线性方程加以描述。公式2.1与公式2.2是反映测量模型的一般方程，公式2.3是反映结构模型的一般方程，这三个一般方程即可构成一个一般结构方程模型（general structural equation model），这就是结构方程（structural equation）一词的含义：

$$y = \boldsymbol{\Lambda}_y \eta + \varepsilon \qquad (2.1)$$

$$x = \boldsymbol{\Lambda}_x \xi + \delta \qquad (2.2)$$

$$\eta = \boldsymbol{B}\eta + \boldsymbol{\Gamma}\xi + \zeta \qquad (2.3)$$

所谓一般方程，是指一般性的通式，而不反映特定变量与特定参数的关系。如果针对特定变量与参数的关系，则可利用明确的方程来表现。例如，图中的结构模型可以用下列两个方程来说明参数的关系：

$$\eta_1 = \beta_{12}\eta_2 + \zeta_1 \qquad (2.4)$$

$$\eta_2 = \gamma_{21}\xi_1 + \gamma_{22}\xi_2 + \zeta_2 \qquad (2.5)$$

第二节 结构方程模型的组成

二、结构方程模型的参数

4. 模型参数与方程

在SEM中，当单独使用测量模型，也就是只有测量模型而没有结构模型的回归关系假设时，即为验证性因素分析，因其检测的内容是测量题目的因素结构与测量误差。进一步地，单独看待结构模型，它其实就是一个传统的路径分析模型，可以用多元回归的概念来说明变量的因果或预测关系。事实上，如果一个SEM模型当中没有任何潜在变量的假设（即没有测量模型），只存在测量变量，并探讨这些测量变量的因果/预测关系时，就与传统的路径分析模型无异。

值得注意的是SEM模型与传统的回归分析不同。SEM不仅可以同时处理对多组回归方程的估计，更重要的是变量关系的处理更具有弹性。首先，在回归分析当中，变量仅分为自变量与因变量，同时这些变量都是无误差的测量变量。但是在SEM模型中，变量的关系除了具有测量关系之外，还可以利用潜在变量进行观察值的残差估计。因此，在SEM的模型中，残差的概念远比传统回归分析复杂。其次，在回归分析中，因变量被自变量解释后的残差被假设与自变量之间的关系是相互独立的；但是在SEM分析中，残差项是允许与变量之间有关联的，但前提是这些特殊的假设也应具有一定的理论逻辑基础。

第二节 结构方程模型的组成

二、结构方程模型的参数

4. 模型参数与方程

在SEM模型中，测量模型与结构模型的变量和参数可利用八种主要的矩阵概念来整合（见图2.4）。

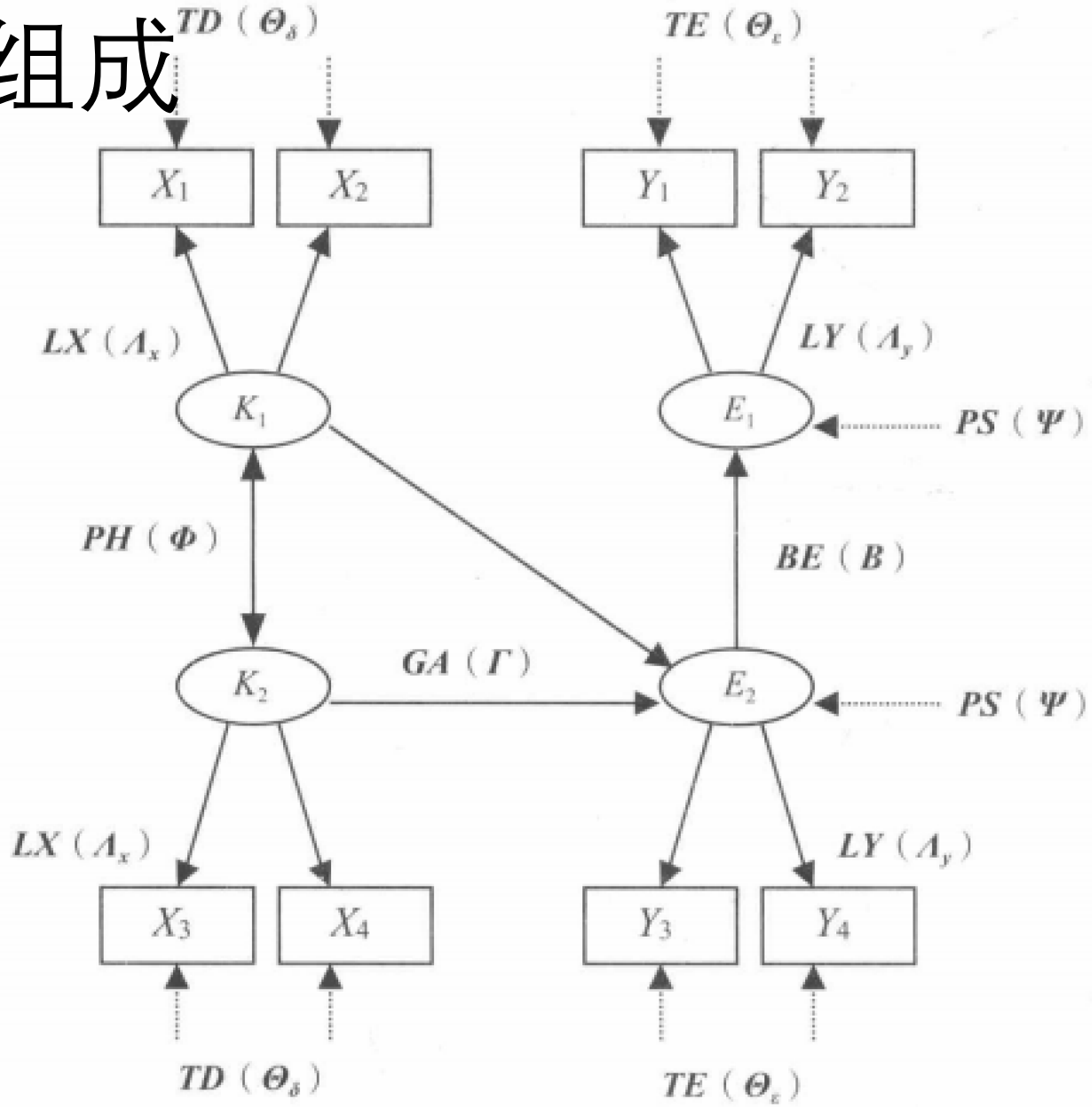


图 2.4　完整的 LISREL 模型矩阵的概念图示

第二节 结构方程模型的组成

二、结构方程模型的参数

4. 模型参数与方程

这八种矩阵的名称、代号、发音、意义与矩阵状态如表2.1所示。

表2.1　　SEM模型的八种矩阵概念列表

符号与发音		缩写	代表的含义	*mm*	*mf*	order
结构模型矩阵						
B	beta	**BE**	内生潜在变量被内生潜在变量解释的回归矩阵（E到E的回归系数）	ZE	FI	NE×NE
Γ	gamma	**GA**	内生潜在变量被外源潜在变量解释的回归矩阵（E到K的回归系数）	FU	FR	NE×NK
测量模型矩阵						
$\boldsymbol{\Lambda}_x$	lambda x	**LX**	外源观察变量被外源潜在变量解释的回归矩阵（K到X的因素载荷）	FU	FI	NY×NE
$\boldsymbol{\Lambda}_y$	lambda y	**LY**	内生观察变量被内生潜在变量解释的回归矩阵（E到Y的因素载荷）	FU	FI	NX×NK
Φ	phi	**PI**	外源潜在变量协方差矩阵（K到K的因素共变）	SY	FR	NK×NK
残差矩阵						
Ψ	psi	**PS**	内生潜在变量被外源潜在变量解释的误差项协方差矩阵（解释残差）	SY	FR	NE×NE
$\boldsymbol{\Theta}_\delta$	theta-delta	**TD**	外源观察变量被外源潜在变量解释的误差项协方差矩阵（X变量残差）	DI	FR	NX×NX
$\boldsymbol{\Theta}_\varepsilon$	theta-epsilon	**TE**	内生观察变量被内生潜在变量解释的误差项协方差矩阵（Y变量残差）	DI	FR	NY×NY

第二节 结构方程模型的组成

三、模型界定

1. 模型界定的概念

SEM分析的第一个具体步骤就是发展假设模型。通过本节所介绍的统计符号与观念所建立的结构方程可以协助研究者进行分析与决策。而模型界定的内容可以从两个方面来获得：第一方面是理论的基础或研究者个人的先备知识与经验，第二方面是SEM分析工具的技术语言与方法要求。首先，当研究者面对所关心的研究问题时，除了基于自己的知识基础与研究兴趣来推演值得探讨的研究命题之外，还需对理论文献详加检阅，以建立严谨的科学假设，或提出有待验证的理论模型。

进一步地，如果研究者决定使用SEM来检验他所提出的假设与理论模型，则必须配合SEM技术语言的规范与各项操作要求，将研究者所提出的假设与理论模型转换成SEM模型。在此同时，必须考虑SEM分析当中可能涉及的各种统计概念的基本原理与要求（例如，因素分析、路径分析、潜在变量的设定、平均数的估计等），将之一并纳入SEM模型界定之中。这个由理论发展到技术性模型建立的一整套程序被称为模型界定。

在模型界定的过程中，除了依据SEM原理来设定模型之外，另外一个非常重要的技术问题就是让SEM模型具有技术上的可识别性，使SEM的各项数学估计程序与统计决策过程顺利进行下去。SEM所设定的理论模型是基于研究者的研究需求提出的。但是，模型的分析与最终的统计决策的获得，必须利用实际搜集到的数据，在转换成为SEM分析的数据格式之后，利用分析软件来进行分析工作。只有假设模型符合统计分析与软件执行的要求，也就是在能够被有效识别的情况下，SEM分析才能顺利进行。此时，一个模型可以被分析工具有效识别的程度称为模型识别度，有关议题将在下一节介绍。

第二节 结构方程模型的组成

三、模型界定

2. 模型界定的简效原则

对共变结构关系的分析可以说是SEM最主要的核心概念，一个SEM模型往往涉及对数十个变量的共变关系的探讨，这些变量可以组合成无数种不同的假设模型。当SEM分析牵涉对如此复杂的变量关系的探究时，一个重要的基本原则是将这些共变关系以最符合理论意义且最简明扼要的方式加以界定，并使之最符合实际观测到的数据结构，这一原则称为简效原则（principle of parsimony）。

简效原则反映了SEM分析的一个主要限制，即研究者无法精确地说明一组变量的共变结构关系究竟以何种模型来表示会最好。因为同样一组变量的组合有无限种可能，而不同的关系模型可能代表了特定的理论意义。因此，如果研究者可以用一个比较单纯简单的模型来解释较多的实际观察数据的变化，那么以这个模型来说明变量间的真实关系时，比较不会得到错误的结论，避免犯下一类错误。换句话说，SEM分析对于理论模型的检验只能说明一个模型不至于有错，但是很难证明它是真正正确的模型。

第二节 结构方程模型的组成

三、模型界定

3. 等值模型问题

所谓等值模型（equivalent models）是指不同的SEM模型具有相同的模型拟合度，也就是模型拟合指数（例如χ^2与p value）统计量在不同的模型间具有相等的数值。等值模型的出现使得统计上“相等”的模型有不同的解释方法。如此一来，关于SEM的研究产生了相当大的困扰，不仅在概念层次上无法有效地厘清变量的关系与模型的优劣，在SEM的操作层次上也使得模型比较无法进行。

另一方面，即使等值模型的现象没有发生在某一个研究中，SEM的操作者也必须明了潜在的等值模型现象是可能存在的。换句话说，任何模型的提出，都可能发生模型界定方式不同，但是模型的拟合度相同的情形。研究者对于SEM分析结果的解释虽然是就他所提出的模型以及所获得的数据进行的，但是同样的一套数据也可以不同的方式来诠释其意义。有许多研究已经证实，对于特定的SEM模型，操作者可以在没有任何理论基础的情况下，利用SEM模型界定策略提出具有相同拟合度的等值模型。

第二节 结构方程模型的组成

三、模型界定

3. 等值模型问题

对于等值模型现象在SEM分析中的影响，目前最常被引用的解决方案称为前导理论（priori theory）策略，仍然是非常原始地从模型发展的初始，以特定理论基础或前导研究发现作为模型界定的依据，而其他各种等值模型被视为不合理但可能存在的模型。如果没有必要，一般皆把等值模型排除在研究讨论之外，以减少不必要的困扰。

另一个普遍的做法是提出几个具有特殊理论内涵的代表性对等模型，利用SEM估计获得各参数的估计数，通过对参数估计的比较，分析哪一个对等模型具有较高的解释力，这一策略可以被称为参数比较（parameter comparison）策略。由于对等模型的卡方值相等，因此所有的拟合指数都失去了比较的意义，仅可逐一检查参数估计结果来检查模型的优劣。

虽然等值模型在单一模型的评估上有困难，但是每一个等值模型都可以继续发展不同的嵌套模型以进行竞争比较。因此，Kline(1998)建议，可以利用代表性的等值模型对每个等值模型进行竞争比较，来检验哪一个等值模型具有最佳的解释性。然而，最终的决策仍在于何者最符合理论上的解释或最具有逻辑推导的正当性，再次显示了理论的考虑在SEM分析中的重要性。

第三节 参数估计与识别问题

一、模型识别问题

1. 参数数目与识别

识别性的问题可以模拟为解联立方程的过程，例如，对于一个二元一次方程：X+Y=10，在没有指定X或Y值或提供第二组方程的情况下，符合该方程条件的(X,Y)解有无数个。此时，就是所谓无法识别或识别不足（under-identified）的状况，也就是数学上无法得到一组特殊解的现象。

除非提出第二个方程：X+2Y=20，此时便能够求出一组且唯一一组特殊解(0,10)。以SEM的术语来说，此种提供足够的条件去求取联立方程未知变量的解的过程即是一种可以识别的状况，称为充分识别（just-identified）。

但是，若再提出第三个方程，例如：X+3Y=40，此时，有两个未知数，却有三个方程，也会造成无特定解的状况，但是可以利用估计的方式，求出符合这三个方程的最佳解，称为过度识别（over-identified）。例如，以(0,10)代人三式，分别得到10、20、30，此解在第三个方程产生10的差距（残差）；若改以(-1,11)代入三式，则可以得到10、21、32，此解在三个方程的残差分别为0、1、8，总和为9。也就是说，(-1,11)是一个较佳解（better solution）。

第三节 参数估计与识别问题

一、模型识别问题

2. 整体模型识别

决定模型识别性的具体步骤，首先是计算用以产生共变结构的观测值数目，称为测量数据数（numbers of data Points，DP）。测量数据数与样本测量变量协方差矩阵当中的协方差与方差数目有关，可利用下式来计算:

$$DP = \frac{(p+q)(p+q+1)}{2}$$

其中，p+q表示测量变量的个数，其中，p为外源测量变量的数目，q为内生测量变量的数目。假设有10个测量变量，总计可以产生10个方差与$C_2^{10} = 45$个协方差，合计为55个数据点，DP=55。至于对被估计的参数数目的计算，则牵涉研究者所提出的模型的各种设定，包括回归系数、协方差和方差三类参数，这些未知的参数必须能够顺利地估计。

第三节 参数估计与识别问题

一、模型识别问题

2. 整体模型识别

①t法则

Bollen (1989)利用DP数与参数估计数目的比较来判断模型的识别性，提出了一个必要非充分的识别条件计算法则——t法则（t-Rule），t值代表模型中的自由估计参数数目。SEM模型要能被识别，必须符合下列关系式：$t \leq \frac{(p+q)(p+q+1)}{2} = DP$。

t法则的判断原则如下：

- 当t<DP时，为过度识别，有过多的方程，但是只需要求少数几个因素解；
- 当t=DP时，为充分识别，例如用两个方程来求二元因素的解；
- 当t>DP时，为识别不足，以太少的条件求取过多的因素解，在SEM分析中，识别不足的情况将导致无法进行任何参数估计。

第三节 参数估计与识别问题

一、模型识别问题

2. 整体模型识别

②虚无B矩阵法则与递归法则

当一个SEM模型中没有任何内生变量影响其他的内生变量时，即没有结构关系的假设时，也就没有任何的结构参数（β系数）的估计，即B矩阵为0。此时，Φ、Γ与Ψ等各矩阵的未知参数数目为导出矩阵（∑矩阵）识别参数数目的函数，因此整个模型可以顺利识别。换句话说，当B矩阵为0时，模型自动识别，Bollen (1989)称之为虚无B矩阵法则（Null B Rule）。虚无B矩阵法则可以说是反映模型识别性的充分条件，当一个SEM模型符合虚无B矩阵法则时，模型的识别性就不成为问题，不必计算识别性数值。

另一个模型识别的充分条件是模型中的B矩阵呈现三角形状态（对称矩阵，对所有变量间的结构参数均加以估计），而Ψ（PST）矩阵呈现对角线状态（表示仅估计干扰项的方差，干扰项的相关不列入估计），此时为一个递归模型且为饱和模型。符合这一条件的模型会自动识别，称为递归法则（recursive rule）。

在递归模型（recursive model）中，变量的因果关系是单一方向的，预测残差项是彼此独立的独立残差模型；非递归模型（nonrecursive model）则允许干扰项具有相关的假设关系而列入估计，变量间亦可能具有回溯关系。当递归模型条件成立时，B、Φ、Γ与Ψ等各矩阵自动获得识别。因此，递归法则显示的也是识别性的充分条件。

第三节 参数估计与识别问题

一、模型识别问题

3. 测量模型的识别性

在SEM分析中，测量模型所决定的是整体模型当中的外显变量（测量变量）与潜在变量之间的关系。一个SEM可能包括多组测量模型，每一个测量模型的识别性都应符合识别性原则，否则可能影响整体模型评估的稳定性。

对于个别测量模型的识别，主要牵涉潜在变量量尺的设定问题。因为潜在变量是由SEM分析估计出来的，而非实际的观察变量，因此必须给定一个单位尺度。此时，可以将潜在变量的方差设定为1，也就是强制以标准化方差作为潜在变量的共同单位。另一种做法是在潜在变量所影响的各测量变量中指定一个变量的因素载荷（或回归系数）为1，也就是将测量变量的单位设定为潜在变量的参考标尺，并使潜在变量的方差得以自由估计。

当把潜在变量作为影响别人的外源变量（外源潜在变量）时，上述两种方式都可以用来设定测量模型的单位。但是当潜在变量作为被他人决定与影响的内生变量时（内生潜在变量），潜在变量本身成为估计的主要对象，此时应设定测量变量的因素载荷为1，使潜在变量的方差可以自由估计。

一个理想的测量模型应是潜在变量所影响的每一个测量变量的因素载荷越接近1，且需达到统计的显著性。同时每一组测量模型各自形成一个独立的丛集，彼此间没有假设性的因素载荷假设，也就是每一个测量变量仅受到单一潜在变量的影响。此时，测量变量被称为纯化指标（pure indicator），此要求是一个非常严格的模型界定策略。

第三节 参数估计与识别问题

一、模型识别问题

3. 测量模型的识别性

在严格的模型界定策略下，测量模型的识别性的计算主要涉及测量模型中设定的每一个潜在变量所影响的测量变量数目的多寡。对于只有一个潜在变量的测量模型，至少要有三个测量变量才能满足测量模型的识别性。同时，这三个测量变量的因素载荷都必须不等于0，且测量残差之间没有任何相关的假设。

如果测量模型当中有超过一个以上的潜在变量，每一个潜在变量只要有至少三个测量变量来估计，而每一个测量变量只用以估计单一一个潜在变量（或被单一一个潜在变量影响），残差之间没有共变假设，且潜在变量的方差被自由估计，此时测量模型就可以被有效识别。另外，如果潜在变量只以两个测量变量来估计，残差无相关，每一个测量变量只用以估计单一一个潜在变量，且没有任何一个潜在变量的共变或方差为0，则模型可以维持可识别的状况。

McDonald (1999)提出了一个较为温和的模型界定策略，称为独立丛集策略（independent cluster basis）。当潜在变量之间具有相关假设时，每一个潜在变量至少需要两个纯化指标；或是当潜在变量之间不具有相关的假设时，每一个潜在变量需要具有至少三个纯化指标。此概念是一个简单的概念，可以套用在验证性因素分析的模型识别性计算中。McDonald与Ho(2002)的分析指出，多数SEM研究均忽略了测量模型的识别性报告，使得读者无从判断测量模型的设定是否恰当，因此呼吁研究者应主动提供测量模型的识别性信息。

第三节 参数估计与识别问题

一、模型识别问题

4. 结构模型的识别性

对于整个SEM模型的识别，如果模型当中包含结构模型（也就是统合模型的情况），除了测量模型的识别问题之外，必须进一步评估结构模型的识别性，以确认结构模型当中的参数设定是恰当的。在计算上，结构模型识别性与测量模型识别性是两个独立的判断过程。结构模型的识别性判定主要牵涉结构模型的结构参数的设定，也就是潜在变量之间的参数数目，无关乎测量模型的参数设定。

结构模型的识别性决定于内生变量之间的关系的假设。如果内生变量之间没有任何预测关系假设（B矩阵为0），可直接应用前面提及的虚无B矩阵法则，此时的结构模型必然可以识别。如果内生变量之间具有特定的预测关系假设，此时需检验这些预测关系是否属于递归模型的路径模型。如果该结构模型是递归模型，也就是说没有回溯关系，且没有干扰残差的相关假设，那么此时的结构模型仍然可以识别，这就是递归法则的概念。

在非递归模型的路径模型中，潜在内生变量之间具有回溯关系（feedback loop），干扰残差项具有相关的假设，此时的结构模型因为涉及过多参数的估计而无法识别，研究者必须使用别的策略来使模型能够识别。除此之外，非递归模型的结构模型有两个附加的条件：第一，每一个方程至少要有潜在内生变量数目减一个变量不属于非递归模型路径；第二，用以计算标准误的信息矩阵（information matrix）必须可以被完全估计，并可以求出倒置信息矩阵（inverted information matrix）。如果从 SEM的分析软件中无法得到一个倒置信息矩阵，表示第二个原则无法满足，结构模型的识别性不足，在进行SEM分析时，将会出现警语，警告基于无法求出倒置信息矩阵的识别不足问题。

第三节 参数估计与识别问题

一、模型识别问题

4. 结构模型的识别性

McDonald (1997)提出了前置原则（precedence rule）与直交/正交法则（orthogonality rule）的概念，是两种较简易的判断原理。所谓前置原则是指在具有因果的先后次序的变量中（也就是具有因果关系的那一对变量），干扰项协方差均设定为零；而当所有的内生变量的干扰项之间都不具有相关的假设时，称为直交法则。在这两条法则下，结构模型的识别性是足够的。在一般的递归模型中，直交法则可以作为判断识别性的依据，但是在非递归模型中，直交法则不适用。

第三节 参数估计与识别问题

一、模型识别问题

5. 潜在变量的量尺化与识别性

潜在变量与一般测量变量最大的不同在于其“不可直接测量”的特性，因此潜在变量缺乏一个自然存在的尺度，必须以人为的手段设定尺度。这里所指的尺度可以被视为测量使用的量尺或单位。

在SEM当中，潜在变量的尺度会因为设定方法的不同而不同，具有非决定性（indetermination），往往会造成SEM参数估计的问题。对于SEM模型中的外源潜在变量的设定，最常使用的方法是将潜在外源变量的方差设定成一个常数（通常为1.00），也就是对模型当中所有的潜在变量进行标准化，使方差维持一致，如此可以对于其他参数进行比较。对于潜在内生变量，通常是将其中的一个测量变量与潜在变量的因素载荷设定为常数（通常为1.00）。

针对潜在变量进行量尺化（scaling）设定的具体做法有二：一是可以将潜在变量的方差设定为1；二是指定潜在变量中任何一个变量的因素载荷（或回归系数）为1。但是，当一个SEM模型混合了测量与结构模型时，潜在变量的量尺化便较为复杂。

在兼含测量与结构模型的统合模型中，对于外源潜在变量的设定可以是单纯地将外源潜在变量的方差设定为1来进行量尺化。但是，对于内生潜在变量，由于作为内生变量者的变异量被外源变量解释，并非被估计参数解释，因此对内生潜在变量的量尺化必须特别小心地处理，否则将造成外源变量无法顺利估计内生潜在变量方差的识别错误。

第三节 参数估计与识别问题

一、模型识别问题

5. 潜在变量的量尺化与识别性

内生变量的量尺化有三种常用策略：第一是Browne与Du Toit (1992)所提出的，将内生潜在变量的方差设定为1，进行设限的参数估计程序；第二是McDonald, Parker 与 Ishizuka (1993)的再参数化策略（reparameterization method），也是将内生潜在变量的方差设定为1，但是仅应用于递归模型的设定；第三是取内生潜在变量中一个因素载荷为固定参数，进行参数估计后，将潜在变量量尺固定后，再进行内生潜在变量的方差估计，称为再量尺法（rescaling method）。

前两种策略分别可由不同的分析软件来设定（例如，RAMONA、SEPATH、CALIS、COSAN），其优点是在完全标准化的参数估计下，标准误能维持不偏，但是结构参数的估计数值存在偏差。第三种再量尺策略应用于LISREL与EQS软件，优点是在完全标准化的参数估计下，可获得正确的参数估计值，但是某些标准误无法正确估计，使得显著性检验参考性降低。目前，大多数的SEM研究均是以LISREL或EQS所提供的再量尺法来进行内生潜在变量的量尺化设定的，这使得这些研究能够获得正确的标准化解，但是无法获得充分的标准误与显著性检验数据。

McDonald与Ho (2002)指出，虽然标准化估计解的报告是证明理论模型各参数意义的重要数据，因此强调精确的标准化解是一个正确的做法，但是这并不意味着在内生潜在变量中，参数估计的显著性检验可以被忽略。因此，搭配不同的检验程序，可以正确报告参数估计的显著性检验，这是未来研究可以努力的目标，但是至少在目前，LISREL与EQS等常用软件仍未能积极处理这一问题。

第三节 参数估计与识别问题

二、参数估计

1. 相关与共变

相关（correlation）是反映两个变量线性关系强度的统计概念。两个连续变量的关联情形除了可以用散点图的方式来表达，也可以利用数学模型来呈现其计量特性，即建立一个用以描述相关情形的量数——相关系数（coefficient of correlation）——来表示线性关系的强度。

线性关系的定义可以用一条最具代表性的直线来表示两个变量的关系，并以直线方程中的斜率来表现变动关系的计量关系。但是，斜率并不足以说明两个变量观察值的分布情形。若要以相关系数反映两个变量的配对观察值的分布，其运算必须考虑两个变量各自的集中与分散状况，以及配对分数的集中与分散状况，将所有观察值的分布情形纳入考虑，以协方差的概念进行处理。而SEM就是用来处理多个变量之间复杂共变结构的统计技术。协方差的公式如下：

$$Cov(X,Y)=\frac{\sum(X-\bar{X})(Y-\bar{Y})}{N-1}$$

协方差的正负号代表两变量是正向或负向关系。

第三节 参数估计与识别问题

二、参数估计

1. 相关与共变

值得注意的是，协方差的数值会因为两个变量的不同单位而没有一定的范围，因此协方差的大小可以反映两个变量共同变化时的原始量的变动，但两个协方差数值无法直接用于比较。若将协方差公式除以两个变量的标准差，协方差即成为相关系数，也就是去除单位的标准化关联系数。公式如下：

$$r = \frac{Cov(X,Y)}{s_x s_y}$$

相关系数作为一个标准化的协方差，其目的在于反映两个变量的关联强度：系数值越高（绝对值越接近1），表示线性关系越强；系数值越低（绝对值越接近0），表示线性关系越弱。SEM的分析过程主要探讨了两两变量的变化关系的结构特性。因此，适当的数据格式是输入协方差数据，协方差的优点是带有各变量的原始计量特征。如果是输入相关系数，那么因为标准化过程已经消除了变量的原始分布特征，所以必须另外输入平均数与标准差，在分析时才可获得测量变量的完整性质。

第三节 参数估计与识别问题

二、参数估计

2. SEM中的共变推导

SEM对于参数的估计主要与方差及共变结构的导出过程有关。Hays (1994)指出了四种与方差（协方差）计算有关的定理:

定理一：某一个变量与自己的共变即等于该变量的方差，即:

$$Cov(X,Y) = Var(X)$$

定理二：经过线性整合后的变量的协方差为:

$$Cov(aX + bY, cZ + dU) = acCov(X,Z) + adCov(X,U) + bcCov(Y,Z) + bdCov(Y,U)$$

定理三：经过线性整合后的变量的方差为:

$$Var(aX + bY) = Cov(aX + bY, aX + bY) = a^2Cov(X,X) + b^2Cov(Y,Y) + 2abCov(X,Y)$$

定理四：独立的两个变量的线性整合后的方差为:

$$Var(aX + bY) = a^2Cov(X,X) + b^2Cov(Y,Y)$$

第三节 参数估计与识别问题

二、参数估计

3. 方差与协方差导出矩阵

现以图3.1中的SEM分析为例。

对于两个观察变量V1与V2，其数学方程为：

$V_1 = \lambda_1 F_1 + E_1$，$V_2 = \lambda_2 F_1 + E_2$。

由定理二，计算两个观察变量的协方差为：

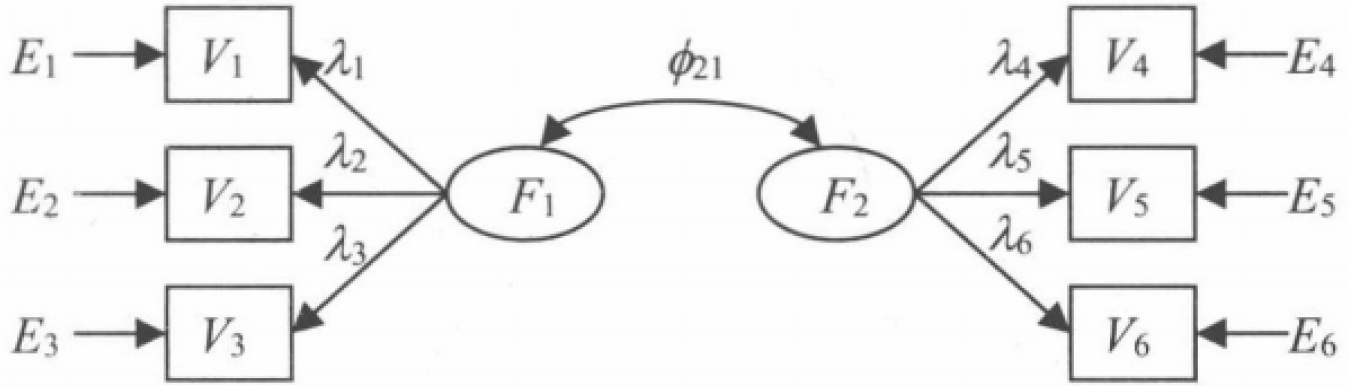


图 3.1 两个具有相关的潜在变量的 CFA 图示

$$Cov(V_1,V_2)=Cov(\lambda_1F_1+E_1,\lambda_2F_1+E_2)$$

$$=\lambda_1\lambda_2Cov(F_1,F_1)+\lambda_1Cov(F_1,E_2)+\lambda_2Cov(E_1,F_1)+Cov(E_1,E_2)$$

$$=\lambda_1\lambda_2Cov(F_1,F_1)=\lambda_1\lambda_2Var(F_1,F_1)=\lambda_1\lambda_2$$

要得到上式的结果，必须符合三个条件：第一，两个误差项的共变为0；第二，误差项与潜在变量的共变为0；第三，潜在变量F1的方差为1。也就是说，符合前述条件时，受到同一个潜在变量影响的两个观察变量的协方差等于潜在变量所属的两个观察变量的因素载荷乘积。

第三节 参数估计与识别问题

二、参数估计

3. 方差与协方差导出矩阵

进一步地，对于V1与V4这两个不同的潜在变量的观察变量，其协方差计算公式为：

$$Cov(V_1,V_4)=Cov(\lambda_1F_1+E_1,\lambda_4F_2+E_4)$$

$$=\lambda_1\lambda_4Cov(F_1,F_2)+\lambda_1Cov(F_1,E_4)+\lambda_4Cov(E_1,F_2)+Cov(E_1,E_4)$$

$$=\lambda_1\lambda_4Cov(F_1,F_2)=\lambda_1\lambda_4\phi_{21}$$

$$Var(V_1)=Cov(\lambda_1F_1+E_1,\lambda_1F_1+E_1)$$

$$=\lambda_1^2Cov(F_1,F_1)+\lambda_1Cov(F_1,E_1)+\lambda_1Cov(E_1,F_1)+Cov(E_1,E_1)$$

$$=\lambda_1^2Var(F_1)+Var(E_1)=\lambda_1^2+\theta_1$$

也就是说，观察变量的方差等于各观察变量的因素载荷的平方加上误差项的方差（θ1）。

第三节 参数估计与识别问题

二、参数估计

3. 方差与协方差导出矩阵

利用上述各式的概念，可以逐一计算出图3.1当中的六个观察变量的方差与配对协方差，进而产生一个由参数所导出的方差与协方差矩阵（∑），内容如下：

$$\boldsymbol{\Sigma}=\begin{pmatrix}
\lambda_1^2+\theta_1 & & & & & \\
\lambda_1\lambda_2 & \lambda_2^2+\theta_2 & & & & \\
\lambda_1\lambda_3 & \lambda_2\lambda_3 & \lambda_3^2+\theta_3 & & & \\
\lambda_1\lambda_4\phi_{21} & \lambda_2\lambda_4\phi_{21} & \lambda_3\lambda_4\phi_{21} & \lambda_4^2+\theta_4 & & \\
\lambda_1\lambda_5\phi_{21} & \lambda_2\lambda_5\phi_{21} & \lambda_3\lambda_5\phi_{21} & \lambda_4\lambda_5 & \lambda_5^2+\theta_5 & \\
\lambda_1\lambda_6\phi_{21} & \lambda_2\lambda_6\phi_{21} & \lambda_3\lambda_6\phi_{21} & \lambda_4\lambda_6 & \lambda_5\lambda_6 & \lambda_6^2+\theta_6
\end{pmatrix}$$

这一由各参数导出的∑矩阵，称为导出矩阵（reproduced matrix ∑）。也就是说，矩阵中的每一个量数，都是由SEM的假设模型与参数估计推导得出的，而非由实际观测到的数值。∑矩阵中的每一个元素都具有一个对应的实际观测值。也就是说，由样本测量得到的六个观察变量的方差与协方差也可以利用一个6×6的矩阵来表示，称之为S矩阵。矩阵中的每一个数值都是具体的实际观测值。将两个矩阵的数值相减，可以得到一个残差矩阵，用来评估每一个量数的拟合情形。所有残差整合之后，即可以用来评估理论模型与实际模型的拟合度。

第三节 参数估计与识别问题

三、参数估计策略

在SEM中，至少有下列方法可以估计各参数：工具性变量方法（instrumental variable method, IV）、两阶段最小平方法（two-stage least squares, TSLS）、未加权最小平方法（unweighted least squares, ULS）、最大概似（maximum likelihood, ML）、一般化最小平方法（generalized least squares, GLS）、椭圆分布理论法（elliptical distribution theory，EDT）以及渐近分布自由法（asymptotic distribution free，ADF）。这些方法的共同点是求取观察与估计共变结构之间差异的最小值，来导出各参数的最佳估计数。

在各种方法当中，最简单的方法是直接分别估计每一个方程，求得每一个方程的解，而不必参考模型当中其他方程的条件，称为有限讯息技术（limited-information techniques），IV法与TSLS法即属此类技术。相对而言，完全讯息技术（full-information techniques）则是同时考虑模型当中所有方程，以获得各参数的估计数。所谓完全讯息是指充分运用模型当中的所有计量信息，以获得最理想的估计数，一般常用的ULS、GLS、ML、ADF法等均属此类。

从统计原理来看，有限讯息技术是单变量（univariate）估计程序，也就是说，估计程序是各方程单独进行的；而完全讯息技术则是多变量（multivariate）估计程序，在估计每一个参数时，要同时考虑其他的方程。因此，有限讯息技术并没有迭代估计（iterative estimation）的问题，而完全讯息技术则必须进行迭代程序，以求取最理想的终解。一般而言，有些学者利用IV法或TSLS法来估计起始值（starting values），这些起始值可以用于完全讯息技术法的参数估计，并且节省估计的时间，提高SEM分析的效率。

第三节 参数估计与识别问题

三、参数估计策略

由于有限讯息技术并不需要进行迭代估计，因此具有速度快、识别错误问题较少、不受某些高阶统计假设限制等优点，因此统计的强韧性/稳健性（robustness）较高。相对而言，完全讯息技术则有较佳的统计效度，得到的结论较为合理。

经由迭代程序，完全讯息技术可以在一定的重复估计程序中找到最理想的解，这一程序称为有效的收敛（convergence）。所得到的估计结果应该是可以被理解的，称为可接受解（admissible solution）或是适当解（proper solution）。相对的，有时候研究者所提出的模型不恰当，模型数据与观察数据相差太大，或是起始值与最终解的差距过大，导致迭代程序无法在少数几次内完成。即使完成了估计，得到的估计解也无法理解、无法解释，称为不可接受解（non-admissible solution）或不当解（improper solution），例如，超过范围的标准化估计数、方差出现负值（Heywood现象），以及非正定问题。由此可知，一个具有理论与实务合理性的模型对于SEM的参数估计有决定性的影响。而起始值的估计虽然是改善终解适合性的一种策略，但是基本问题还是在待验证的假设模型之上。

第三节 参数估计与识别问题

三、参数估计策略

1. 加权最小平方策略（Weighted Least-Squares, WLS）

①定义

加权最小平方策略是一类通过拟合函数F(Q)最小化来估计参数的方法，其核心是使估计协方差矩阵（Σ）与观察协方差矩阵（S）的加权残差平方和达到最小。拟合函数的一般形式为：

$$F(Q) = [s - \sigma(\Theta)]' W^{-1} [s - \sigma(\Theta)]$$

其中W^{-1}是校正权数矩阵，不同估计方法的差异主要在于W^{-1}的选择。

②估计策略（WLS是一个统称）

- 无加权最小平方法（ULS）：权数矩阵W^{-1}为单位矩阵（即所有残差等权），拟合函数为 $F_{ULS} = \frac{1}{2} \operatorname{tr}\{[S - \Sigma(\Theta)]^2\}$ 。直接求残差平方和的最小值。
- 一般化最小平方法（GLS）：权数矩阵取观察协方差矩阵的逆矩阵S^{-1}，拟合函数为 $F_{GLS} = \frac{1}{2} \operatorname{tr}\{[S - \Sigma(\Theta)]S^{-1}\}^2$ 。通过加权校正异质性。

第三节 参数估计与识别问题

三、参数估计策略

1. 加权最小平方策略（Weighted Least-Squares, WLS）

③优点

- ULS：计算简单、快速，无需迭代。
- GLS：比ULS更有效，能处理残差异质性问题。
- WLS整体：弹性大，可适用于多种模型设定。

④缺点

- ULS：未考虑变量量尺差异，大样本时残差变异大，估计效果差；依赖正态假设。
- GLS：仍依赖正态假设；样本量较小时表现不如ML。
- WLS整体：需要较大样本（通常要求自由参数数目的10倍以上）；当存在遗漏值时需使用全列排除法，造成样本流失；对正态性敏感，违反正态时结果易扭曲。

⑤适用条件与场景

- ULS：适用于所有观察变量具有类似测量尺度（单位相同）的情况，且样本量不大、对精度要求不高时可用。
- GLS：适用于中等样本量（例如500以下），且变量近似正态分布的场景。
- WLS整体：当研究者希望使用加权最小平方框架，且样本量充足、数据质量较好时可采用。但在实际SEM应用中，ML法更为常用。

第三节 参数估计与识别问题

三、参数估计策略

2. 最大概似法（Maximum Likelihood, ML）

①定义

最大概似法基于概率原理，假设观察数据来自多变量正态总体，通过寻找使样本出现概率（似然）最大的参数值作为估计值。其拟合函数为：

$$F_{ML} = \log|\Sigma| - \log|S| + \mathrm{tr}(S\Sigma^{-1}) - \rho$$

其中ρ为测量变量数目。当模型完美拟合时，$F_{ML} = 0$。

②估计策略

ML法属于完全讯息技术，同时利用模型所有方程的信息，通过迭代程序（例如Newton-Raphson）不断调整参数，使拟合函数最小化，直至收敛。估计过程中以Σ^{-1}作为加权矩阵。

第三节 参数估计与识别问题

三、参数估计策略

2. 最大概似法（Maximum Likelihood, ML）

③优点

- 最常用的SEM估计方法，理论成熟，软件支持广泛。
- 具有渐近无偏性、一致性和有效性（大样本下）。
- 在小样本或变量峰度不太理想时，仍能获得较稳定的估计结果。
- 可配合Satorra-Bentler校正处理非正态数据。

④缺点

- 严格依赖多变量正态假设；违反正态时，卡方值易膨胀，标准误可能偏误。
- 大样本下卡方检验容易显著，需结合其他拟合指数判断。
- 迭代过程可能不收敛或得到非正定解（例如Heywood现象）。

⑤适用条件与场景

- 适用条件：样本量一般建议在200以上（中型样本）；变量近似正态分布；若轻微违反正态，可使用校正ML（例如scaled ML）。
- 场景：绝大多数SEM 研究（验证性因素分析、路径分析、统合模型等）的首选方法。

第三节 参数估计与识别问题

三、参数估计策略

3. 渐近分布自由法（Asymptotic Distribution Free, ADF）

①定义

ADF法由Bollen（1984）提出，是一种不依赖多变量正态假设的估计方法。它通过构造特殊的权数矩阵（包含四阶矩信息）来消除非正态性的影响，因此称为“分布自由”。其拟合函数为：

$$F_{\mathrm{ADF}} = [s - \sigma(\Theta)]' W^{-1} [s - \sigma(\Theta)]$$

其中权数矩阵的元素包含峰度与协方差的组合：$w_{ijkl} = \sigma_{ijkl} - \sigma_{ij}\sigma_{kl}$ 。

②估计策略

ADF法属于完全讯息技术，需要直接使用原始数据（不能仅用协方差矩阵），通过迭代估计参数。计算过程涉及四级动差（峰度），因此对计算机内存和计算时间要求较高。

第三节 参数估计与识别问题

三、参数估计策略

3. 渐近分布自由法（Asymptotic Distribution Free, ADF）

③优点

- 无需正态假设，适用于非正态分布的数据。
- 理论上可提供正确的标准误和卡方检验。

④缺点

- 样本量要求极大：通常需要2500人以上才能得到稳定结果。
- 计算繁复：耗时且占用大量内存，硬件要求高。
- 必须使用原始数据，无法直接输入协方差矩阵。
- 在小样本下表现不理想，甚至不如ML或GLS。

⑤适用条件与场景

- 适用条件：样本量极大（例如2500以上），数据明显违反正态假设（例如严重偏态、厚尾），且计算资源充足。
- 场景：当研究者无法通过数据转换或校正方法满足正态性，且拥有大样本时，可考虑ADF法。但在实际应用中，由于样本量限制，ADF法使用较少，更多采用校正ML法（例如scaled ML）来处理非正态数据。

第三节 参数估计与识别问题

四、参数估计的相关议题

1. 参数估计与样本量的关系

为了使SEM分析能够在正态化假设成立的条件下进行，维持一定规模的样本量是必要的。一般而言，使用最大概似法的参数估计，样本量需到达500人，正态假设的共变结构分析才能够维系；在500人以下时，GLS方法较佳。

Bentler与Yuan (1999)提出了一个统计量——Yuan-Bentler的T，稍微修正了ADF法的计算过程，发现样本规模为60~120时，仍有相当稳定的估计结果。公式如下：

$$T = \frac{[N - (p^* - q)]T_{ADF}}{[(N-1)(p^* - q)]}$$

其中，N为样本量，p为测量变量的数目，p*=[p(p+1)]/2为观察变量所提供的SEM分析的数据数，q为估计参数数目，T_{ADF}为基于ADF方法所估计出来的统计量。

当违反正态性假设时，以ML法与GLS法来估计参数，样本规模必须达2500以上，参数估计才能趋于稳定。样本越少，GLS法较ML法稍微好一点。ADF法在样本量在2500以下时表现得不理想。如果违反误差项独立假设，ML法与GLS法的估计效果在任何样本规模下均不理想，ADF法在样本量大于2500时表现理想。EDT法比ML、GLS与ADF法都更好，在中度到大型样本量的情况下，度量化ML法（scaled ML，是修正正态化不足后的最大概似法）表现得最好。

第三节 参数估计与识别问题

四、参数估计的相关议题

1. 参数估计与样本量的关系

总而言之，测量变量的正态性假设与独立性假设是影响SEM参数估计与假设检验的重要因素。ML法、度量化ML法和GLS法适用于中到大规模样本的SEM分析，度量化ML法尤其依赖计算机的校正。因此，如果考虑成本因素，ML法与GLS法是适合的参数估计方法，尤其是ML法，已经成为SEM分析参数估计的主要方法。如果配合ML的度量化校正，非正态的威胁可以加以修正，使得参数估计更为稳定。但是由于度量化ML法必须针对所有的观察变量进行调整，即使使用大型计算机工作站，估计过程也非常耗费时间。因此，使用者可以先利用ML法进行整体SEM模型检验。当进行参数估计时，则使用度量化ML法，将得到最佳的效果。在小样本下，可以使用Yuan-Bentler所发展的T统计量。

第三节 参数估计与识别问题

四、参数估计的相关议题

2. 模型参数估计的迭代

由于参数的数值大小是被估计出来的，因此，估计的程序必须能够不断尝试各种可能的数值，以获得最适合的结果（使估计矩阵与观察矩阵的差异最小化）。通过分析软件，计算机可以快速地反复进行估计。每一次的估计都可以稍微缩小观察矩阵与理论模型矩阵的差异，直到无法进一步有效改善模型的拟合度，计算机即自动停止参数估计，达成收敛，得出一组优化的参数值，称为最终解，这一过程就是所谓的迭代估计。有时，SEM的迭代估计在一定的次数内无法获得有效收敛以获得终解，即使得到任何参数估计值，也是不值得采信的。

一旦获得终解，SEM估计也可以获得每一个参数的估计标准误（standard error），用以反映每一个参数估计的可能的波动范围，也就是估计误差的大小。利用参数估计值除以标准误，可以计算出统计检验值（t值），再配合t分布，可以进行参数的显著性检验。一般而言，如果t值的绝对值大于1.96，则该参数即可达到0.05的显著水平。在样本量低于30之时，样本量越小，t值要越大才能超越显著水平的门槛。

第三节 参数估计与识别问题

四、参数估计的相关议题

3. 非正定（non-positive definition）问题

该问题是执行SEM参数估计过程中比较常见的技术问题，能够导致程序不正常地终止。

在SEM分析中，实际观察到的协方差矩阵与理论导出的协方差矩阵必须是非歧异性（non-singular）的，否则将使SEM的数学计算陷入无意义的困境。

SEM矩阵估计的非正定问题通常发生于下列状况：

① 矩阵中对角线的数值通常为观察变量的方差或自身相关系数（1.00），其数值必须为正值，否则违反数学原理。

② 对角线的数值是其他元素的基本条件，也就是说，方差是计算协方差的基础，在一定的方差的条件下，协方差有其存在的合理范围。因此，当矩阵中有超出一定合理范围的情况时，即违反数学原理，称为三角不均等条件（triangular inequality condition）。

出现方差为负值或协方差数值超过合理范围的现象，可能是数据输入错误造成的。SEM分析可以直接读取协方差矩阵，不必从原始数据逐个读取数据。但是在矩阵数据处理过程中，可能由于操作者的疏失产生错误，造成不合理的现象。

此外，当数据库当中存在遗漏值且呈现非系统性遗漏现象时，如果使用配对排除法或以其他方式填补遗漏值，都可能造成不合理数值的出现。

第三节 参数估计与识别问题

四、参数估计的相关议题

3. 非正定问题

第三，为了使SEM的参数估计顺利得到反矩阵，矩阵必须符合非歧义的要件。如果在倒置过程中分母项为零，将造成数学的无意义除式，致使分析中止。通常这一问题发生在变量之间高度相关的情况下，这就是变量间线性依赖或共线性（collinearity）的问题。当某一个变量是其他变量的线性整合时，反矩阵即可能出现非正定问题。

整体而言，造成SEM参数估计的非正定问题的原因不止上述三类，还可能有其他问题（例如，不恰当的假设模型、不恰当的起始值、模型的识别性问题等），或是多重原因造成的影响。

第四节 模型拟合评鉴

一、模型评鉴的基本概念

再次强调，SEM是一套用以检验特定假设模型的统计方法学。因此，SEM最主要的目的之一是检验研究者所提出的理论或概念架构是否具有实证的意义。整个SEM的分析程序都离不开研究者所提出的假设模型，因而研究者是否可以在提出研究问题的第一时间就通过理论推导与文献检阅过程来选择适当的研究变量，提出有意义的研究假设去说明变量的关系，进而发展出适切理想的假设模型，即成为SEM模型的计量检验程序是否可以顺利完成的一个基本条件。换句话说，SEM分析技术只是一套统计的方法与分析的策略，SEM本身并无法创造理论或知识，而是需要研究者以其智慧去整理前人发展出来的理论或知识，建构一套适当的概念模型，然后再以SEM技术协助研究者完成模型的分析与讨论。

1. 测量质量与模型评鉴

要能够顺利完成SEM的计量分析，除了假设模型的发展，还依赖严谨的抽样与测量程序，以获得稳定、有效的观察数据来评估SEM模型的适切性。

在计量领域中，测量与分析可以说是两套独立的程序，如果抽样有偏或测量的过程带有大量误差，不仅在检验过程当中会导致统计上的诸多问题，更可能导致错误的结论。换句话说，当观察变量的信效度不佳、测量质量低时，SEM的分析也不会有理想的结果，而且这一不理想的结果可能完全与假设模型的优劣无关，却让人们误将研究结果的问题指向是假设模型本身。因此，在进行SEM分析之前，研究者必须具体指出变量的操作性定义，谨慎小心地处理测量的信效度问题，以避免上述困境的发生。

第四节 模型拟合评鉴

一、模型评鉴的基本概念

2. 模型评鉴的假设检验

SEM模型评鉴的一个重要概念是SEM分析只能用来评估研究者所提出的假设模型是否适切。但是，究竟何者才是真正反映真实世界的变量之间关系的模型，并不能从模型评鉴程序中得出这一个结论。因为除了研究者所提出的模型，对于同样的一组观察变量，可能还有许多不同的模型组合，这些基于同样的观察数据的基础假设模型可能都有理想的拟合度，SEM分析无法区辨这些计量特征类似的理论模型哪个为真。SEM的使用者不但必须谨记统计方法学本身的限制，也必须避免陷入过度推论的陷阱中。

SEM分析策略与其他推论统计方法有一个明显差异：以支持虚无假设为模型拟合度存在的证据。一般来说，推论统计的各种统计检验多以推翻某一寻常假设（虚无假设）、获得研究者所提出的特定假设（对立假设）为真的结论为目标。当虚无假设不成立之时，研究者即可以“证明”所提出的假设有其存在的意义与价值。

但是，在SEM中，假设模型的适切性是以虚无假设的形式存在的，虚无假设代表SEM假设模型与实际观察所得数据相符合，对立假设则表示理论模型不能反映观察数据。很明显，基于这一假设检验逻辑获得的虚无假设成立的统计决策（也就是我们可以主张被检验的SEM模型是一个适当的模型时），只能说明该SEM假设模型与观察数据没有不同，却无法“证明”该SEM假设模型是否有别于任何其他的模型。这也就是为何SEM的研究者无不挖空心思去发展计量技术，以区辨、评估SEM模型的好坏优劣。

第四节 模型拟合评鉴

一、模型评鉴的基本概念

3. 参数估计与模型评鉴

从SEM的执行技术来看，假设模型的检验是在完成参数估计之后才进行的工作。因此，评估模型适切性的第一步应是SEM的模型与各项参数能够顺利被识别、收敛和估计。然后借由对模型参数的质量与适当性的检验，对于假设模型有一个初步的筛检。

具体来说，研究者可以逐一检查参数估计的结果，检查每一个参数的正负号、数值大小是否符合理论预期；或是检查测量误差的大小，分析这些残差项当中是否透露了某些变量的测量质量不佳的讯息。如果某些变量的测量误差过于严重，研究者应先行解决测量的问题，重新检讨参数的估计，而非进入模型评鉴的程序。

第四节 模型拟合评鉴

二、模型评鉴的方法

一旦SEM假设模型中的每一个参数被顺利估计，SEM即可以进行整体模型的评估。通过不同的统计程序或拟合指数的计算，研究者可以分析假设模型与实际观察数据的拟合情形。如果模型拟合度不理想，代表研究者所提出的假设模型可能存在某些问题，可能是模型的设定、参数的估计或是技术上的问题导致了假设模型无法与观察数据拟合。此时，研究者可以应用模型修饰的原则，调整假设模型的参数估计内容，重新加以估计，直到模型拟合度达到理想的水平。在此过程中也必须反复检验模型拟合指数，可见模型拟合度评估在SEM分析当中的重要性。

第四节 模型拟合评鉴

Log-likelihood Values

	Estimated Model	Saturated Model
Number of free parameters(t)	51	171
-2ln(L)	3168.281	2925.983
AIC (Akaike, 1974)*	3270.281	3267.983
BIC (Schwarz, 1978)*	3461.337	3908.584

*LISREL uses AIC= 2t - 2ln(L) and BIC = tln(N)- 2ln(L)

Goodness-of-Fit Statistics

Degrees of Freedom for (C1)-(C2)	120
Maximum Likelihood Ratio Chi-Square (C1)	242.298 (P = 0.0000)
Browne's (1984) ADF Chi-Square (C2_NT)	230.062 (P = 0.0000)
Estimated Non-centrality Parameter (NCP)	122.298
90 Percent Confidence Interval for NCP	(81.721 ; 170.655)
Minimum Fit Function Value	0.774
Population Discrepancy Function Value (F0)	0.391
90 Percent Confidence Interval for F0	(0.261 ; 0.545)
Root Mean Square Error of Approximation (RMSEA)	0.0571
90 Percent Confidence Interval for RMSEA	(0.0466 ; 0.0674)
P-Value for Test of Close Fit (RMSEA < 0.05)	0.128
Expected Cross-Validation Index (ECVI)	1.100
90 Percent Confidence Interval for ECVI	(0.970 ; 1.254)
ECVI for Saturated Model	1.093
ECVI for Independence Model	9.196
Chi-Square for Independence Model (153 df)	2842.325

Normed Fit Index (NFI)	0.915
Non-Normed Fit Index (NNFI)	0.942
Parsimony Normed Fit Index (PNFI)	0.717
Comparative Fit Index (CFI)	0.955
Incremental Fit Index (IFI)	0.955
Relative Fit Index (RFI)	0.891
Critical N (CN)	205.676
Root Mean Square Residual (RMR)	0.0536
Standardized RMR	0.0518
Goodness of Fit Index (GFI)	0.924
Adjusted Goodness of Fit Index (AGFI)	0.892
Parsimony Goodness of Fit Index (PGFI)	0.649

图4.1
验证性因素分析（范例一）的模型拟合度分析数据

第四节 模型拟合评鉴

二、模型评鉴的方法

1. 卡方检验

①原理

在SEM分析当中，最常用的模型评鉴方式是卡方检验（χ^2 test）。在SEM中，卡方值是由拟合函数转换而来的统计量，反映了SEM假设模型的导出矩阵与观察矩阵的差异程度。卡方值的导出式如下：

$$T = (N-1)F_{min}$$

其中，T代表模型拟合度的检验值，其性质与卡方值相同，因此可以视为卡方值；N为样本量，F_{min}表示以各种不同参数估计方法（例如ML、GLS、ADF法等）得到的拟合函数的最小函数估计值，该数值也反映了优化的参数统计量（假设模型与观察数据差异的最小值）。

在符合卡方分布的条件下，可以对于T值进行卡方检验来检验其显著性。此时，虚无假设为理论矩阵与观察矩阵没有差异[H0：$\Sigma = \Sigma(\theta)$]（模型拟合度良好）。当T值到达显著水平，代表虚无假设不成立，模型拟合度不佳；反之，当T值未达显著水平，代表虚无假设成立，模型拟合度良好。

第四节 模型拟合评鉴

二、模型评鉴的方法

1. 卡方检验

②正规卡方值（卡方自由度比）

在卡方检验的概念下，我们可以理解自由度越大的模型在卡方统计检验上越处于不利的地位。因此，如果对两个模型同时进行SEM分析，得到不显著的卡方值时，自由度越大的模型越有能力反映真实的数据，这就是SEM中常见的简效原理的统计原理。由此可见，如果利用卡方数据去检验模型拟合度，除了依据卡方值越小越好的统计显著性原理之外，还需考虑自由度大小的影响。

在SEM中，可以计算出一个卡方自由度比（$\frac{\chi^2}{df}$），称为正规卡方值（normed chi-square），来进行模型间拟合比较。正规卡方值越小，表示模型拟合度越高；反之则表示模型拟合度越差。一般而言，正规卡方值小于2或3时，模型具有理想的适配度，小于1则有过度适配的问题。

第四节 模型拟合评鉴

二、模型评鉴的方法

1. 卡方检验

③卡方检验的相关议题

以卡方检验来评估理论模型的适切性，应注意几个问题。第一，在SEM当中，被检验的假设模型若能够有效反映实际观察到的数据，其虚无假设应成立，而对立假设不成立。因此，利用卡方分布来进行假设模型的统计检验并无法推翻不良的模型以支持特定的模型，只能确认虚无假设是否成立。此程序无法检验假设模型的特定意义，只能说明假设模型并不是无法反映观察资料，只是不能证明假设模型是否真的反映了它所代表的理论概念。

进一步的，由于卡方分布受到自由度的影响，所以自由度越大，卡方值越大。也就是说，当自由度越大，所欲估计的参数数目越多，影响一个假设模型的因素就越多，造成假设模型拟合度不佳的可能性越大，此时的卡方值越不能反映理论模型是否能够反映观察数据的程度。

此外，卡方分布也与样本规模有关：样本越大，所累积的卡方值也就越大。但是大样本虽然提高了观察数据的稳定性，却也造成了卡方值扩大的效果。虚无假设被拒绝的概率与自由度及样本量具有正比函数关系。因此，当利用卡方分布来检验SEM模型时，会因为参数数目与样本量的技术特性，影响假设模型的拟合度检验。这也是一般SEM使用者舍卡方而用其他拟合指数的主要原因。

另一方面，基于中央极限定理，当样本够大时，从样本上获得的参数估计统计量应会越接近中央极限卡方分布，称为中央卡方分布（central chi-square distribution）。中央卡方分布代表假设模型参数是否能够拟合观测数据的卡方真分数。由SEM分析当中获得的T值除了可以通过卡方检验来检验其显著性，亦可以与中央卡方分布进行差异比较，来决定T值的统计意义。

第四节 模型拟合评鉴

二、模型评鉴的方法

2. 模型拟合指数

①GFI与AGFI指数

GFI指数即是拟合指数（goodness-of-fit index）的缩写，类似于回归分析当中的可解释变异量（R2），表示假设模型可以解释观察数据的方差与协方差的比例：

$$GFI = \frac{tr(\hat{\sigma}W\sigma)}{tr(s'Ws)}$$

其中，分子是理论假设模型的协方差所导出加权方差和，分母是样本实际观察所得到的协方差导出的加权方差和，W是加权矩阵。由于模型导出值会小于实际观测值，因此GFI是小于1的比值。GFI值越接近1，分子分母越接近，表示模型拟合度越高。相对地，GFI数值越小，分子分母差距越大，表示模型拟合度越低。

第四节 模型拟合评鉴

二、模型评鉴的方法

2. 模型拟合指数

①GFI与AGFI指数

AGFI（adjusted GFI）则类似于回归分析当中的调整后可解释变异量（adjusted R2），是在计算GFI系数时，将自由度纳入考虑之后所计算出来的模型拟合指数。参数越多，AGFI指数数值越大，越有利于得到理想的拟合度结论，其公式如下：

$$AGFI = 1 - \frac{1 - GFI}{1 - \frac{t}{DP}}$$

t为估计参数数目，DP为观察资料数。GFI与AGFI均具有标准化的特性，数值介于0~1，数值越大（越接近1）表示拟合越佳，越接近0表示拟合越差。一个能够拟合观察数据的SEM模型，其GFI与AGFI都会非常接近1.00，一般需要大于0.90才可被视为具有理想的拟合度。

到目前为止，并无这两个指数的统计概率分布，因此无法直接针对GFI与AGFI进行显著性检验。

第四节 模型拟合评鉴

二、模型评鉴的方法

2. 模型拟合指数

②PGFI指数

GFI指数一种变形是PGFI指数（Parsimony Goodness-of-fit Index），其计算式下：

$$PGFI = \left(1 - \frac{t}{DP}\right) GFI$$

从计算式可以看出，PGFI指数考虑到了模型当中估计参数的多寡，可以用来反映SEM假设模型的简效程度（degree of parsimony）。其判断原理与GFI相似，PGFI指数越接近1，显示模型越简单。但是，根据其计算式，可以发现除非估计参数的数目远小于观察资料数，否则PGFI指数会远小于GFI指数数值。

Mulaik等人（1989）指出，对于一个良好的模型来说，PGFI指数大约在0.5以上都是可能的。

第四节 模型拟合评鉴

二、模型评鉴的方法

2. 模型拟合指数

③NFI与NNFI指数

另外两种相当常用的拟合指数是正规拟合指数（normed fit index，NFI）与非正规拟合指数（non-normed fit index，NNFI），这两种指数是利用嵌套模型的比较原理所计算出来的一种相对性指数，反映了假设模型与一个观察变量间没有任何共变假设的独立模型的差异程度。

在SEM模型中，将观察变量之间设定为没有任何共变情况所得到的独立模型（independent model），可以说是利用同一组观察变量可能组合而成的无数个假设模型当中最基本的一种状况，在概念上可以作为所有其他模型的基准模型（baseline model）。独立模型表示了拟合状况最不理想的一种模型，反映了所有的观察变量之间没有任何关联，且自由度最大。故以独立模型导出的卡方值（χ^2_{indep}）是所有可能模型的卡方值的最大值。而其他所有的模型（称为比较模型，以χ^2_{test}表示）都是从虚无模型加以延伸的嵌套模型，卡方值会较独立模型的卡方值为低，可以与虚无模型相比较：

$$NFI = \frac{\chi^2_{indep} - \chi^2_{test}}{\chi^2_{indep}}$$

NFI指数的原理是计算假设模型的卡方值（χ^2_{test}）与虚无模型的卡方值（χ^2_{indep}）的差异量，可以视为某一个假设模型比起最糟糕模型的改善情形。

第四节 模型拟合评鉴

二、模型评鉴的方法

2. 模型拟合指数

③NFI与NNFI指数

根据研究发现，在小样本与大自由度的情况下，对于一个拟合度理想的假设模型，以NF1来检验拟合度会有低估的现象。所以，其他研究者提出了另一个NNFI指数，考虑了自由度的影响，类似于前述AGFI对GFI的调整，由此可以避免模型复杂度的影响，其计算式如下：

$$NNFI = \frac{\chi^2_{indep} - \frac{df_{indep}}{df_{test}}\chi^2_{test}}{\chi^2_{indep} - df_{indep}}$$

调整后的NNFI指数改善了NFI的问题，却使得NNFI有时会有超越0~1范围的数值出现，显示NNFI的波动性较大。同时，NNFI可能会比其他指数小，导致了在其他指数显示模型拟合的情况下，NNFI显示出模型拟合度不理想的矛盾结论。

第四节 模型拟合评鉴

二、模型评鉴的方法

2. 模型拟合指数

④IFI指数

Bollen(1989)提出了一个增量拟合指数（incremental fit index，IFI）来处理NNFI波动的问题以及样本大小对于NFI指数的影响，公式如下：

$$IFI = \frac{\chi^2_{indep} - \chi^2_{test}}{\chi^2_{indep} - df_{test}}$$

在一般正常情况下，IFI、NFI与NNFI值都会介于0~1，数值越大，表示拟合度越佳，同时系数值需大于0.90才可被视为具有理想的拟合度。

第四节 模型拟合评鉴

二、模型评鉴的方法

3. 替代指数

替代指数（alternative index）与模型拟合指数的主要不同在于替代指数不是以卡方统计量的假设检验来进行模型拟合度的评估的。替代性指数的基本想法是假设模型与实际观察矩阵的比较不是全有/全无的概念。也就是说，模型拟合度的检验并非针对假设模型导出矩阵与观察矩阵是否相同这一个虚无假设进行检验。因为观察数据本身是否能够反映真实变量的关系也是有待考验的问题，若以观察数据矩阵作为比较基准，并不一定能够反映假设模型的优劣。例如，观察数据本身的测量质量不良（信效度不佳）可能影响观察矩阵的内容。因此，替代性指数不再关心虚无假设是否成立，而是去直接估计假设模型与由抽样理论导出的卡方值的差异程度。

由于替代指数是直接估计被检验模型与理论分布的差异程度，因此可以在中央极限定理的基础上，以区间估计的概念来进行显著性检验。其做法是考虑抽样误差对于指数估计的影响，将所计算出来的指数转换成特定信心水平下（例如90%）真实指数出现的范围。其优点是可以反映抽样误差的影响。

第四节 模型拟合评鉴

二、模型评鉴的方法

3. 替代指数

①非中心性参数（NCP）

一种主要的替代指数是非中心性参数（non-centrality parameter, NCP）。Steiger (1990, p.177)称NCP为“模型不拟合的自然量数（natural measure of badness-of-fit of a covariance structure model），其原理类似离散量数的测量，即计算SEM模型估计得到的卡方统计量距离卡方分布期望值（中央点）的距离[discrepancy between $\sum$ and $\sum(\theta)$]。模型越不理想，距离卡方分布中央越远，以几个标准差（δ）来表示（即NCP数值）。越大的NCP值代表模型越不理想，当NCP为0时，代表模型具有完美的拟合度。在LISREL报表中，除了NCP值之外，也报告NCP在90%的信心水平下的估计范围，若这一区间涵盖了0，亦表示模型拟合度完美。

第四节 模型拟合评鉴

二、模型评鉴的方法

3. 替代指数

②RMSEA指数

另一个重要的替代性指数为平均概似平方误根系数（root mean square error of approximation, RMSEA），其计算式如下：

$$estimated\ RMSEA = \sqrt{\frac{\widehat{F_0}}{df_{test}}}$$

其中，$\widehat{F_0}$是被检验模型的卡方值减去自由度再除以样本量的函数值：

$$\widehat{F_0} = \frac{\chi^2_{test} - df_{test}}{N}$$

由上式可知，RMSEA系数不受样本量大小与模型复杂度的影响，当模型趋近完美拟合时，$\widehat{F_0}$接近0，RMSEA指数亦接近0。RMSEA指数越小，表示模型拟合度越佳。

第四节 模型拟合评鉴

二、模型评鉴的方法

3. 替代指数

②RMSEA指数

由计算原理可以看出，RMSEA与CFI及NFI的不同是：RMSEA指数在比较理论模型与完美拟合的饱和模型上的差距程度。数值越大代表模型越不理想，数值越小代表模型拟合度越理想。Hu与Bentler (1999)建议：当RMSEA指数低于0.06，模型可以视为一个好模型；大于0.10表示模型不理想。McDonald与Ho (2002)则建议以0.05为良好拟合的门槛，以0.08为可接受的拟合门槛。

近年来，RMSEA指数相当受到重视，由于其他模型拟合指数多受到样本大小与观察变量分布（例如平均数）的影响，而RMSEA可摆脱这项困扰，且具有NCP指数反映假设模型与中央卡方分布的离散性特质，因此RMSEA指数近年来普遍为大家接受。但是最近的研究指出，RMSEA指数在小样本时有高估的现象，使拟合模型被视为不理想的模型，因此在小样本时应谨慎使用RMSEA数值。

第四节 模型拟合评鉴

二、模型评鉴的方法

3. 替代指数

③CFI指数

CFI指数（comparative-fit index）反映了假设模型与无任何共变关系的独立模型差异程度的量数，也考虑到被检验模型与中央卡方分布的离散性。其计算原理基于非中央性改善比（ratio of improvement in noncentrality；假设模型距离中央卡方分布的移动情形），得出一个非中央性参数（noncentrality parameter，τ_i）：τ_i越大，代表拟合越不理想；当$\tau_i = 0$时，假设模型具有完美适切性。其概念式如下：

$$\tau_{indep.test} = \chi^2_{indep.test} - df_{indep.test}$$

$$\tau_{est.test} = \chi^2_{est.test} - df_{est.test}$$

$\tau_{est.test}$为理论假设模型非中央性参数估计数，$\tau_{indep.test}$为虚无模型相对于假设模型的非中央性参数。根据上式，得到CFI指数公式如下：

$$CFI = 1 - \frac{\tau_{est.test}}{\tau_{indep.test}}$$

由于独立模型（观察变量之间没有任何共变）是最不理想的模型，任何模型一定比独立模型的拟合度更佳。因此，CFI指数的数值也是越接近1越理想，表示能够有效改善非中央性的程度。其性质与NFI接近，一般是以0.95为通用的门槛。同时在小样本的SEM分析中，CFI指数用来评估模型拟合度十分稳定。

第四节 模型拟合评鉴

二、模型评鉴的方法

3. 替代指数

④ECVI与AIC指数

Browne与Cudeck (1993)提出了一个期望交叉效度指数（expected cross-validation index，ECVI），扩大了非中央性参数的应用。ECVI指数反映了在相同的总体之下，不同样本重复获得同一个假设模型的拟合度的期望值，是诊断模型的复核效化（cross-validation）的良好指数。ECVI值越小，表示模型拟合度的波动性越小，该假设模型越好；反之，ECVI越大，表示模型拟合度在不同样本上的波动性越大，该假设模型的拟合度越不理想。

在进行不同样本间的比较时，ECVI较RMSEA的优势是RMSEA指数的计算不考虑样本量与自由度的影响，因此无法应用于不同样本量的假设模型拟合度的比较。尤其当样本量很小但是自由度很大（模型很复杂时）时，ECVI指数能够反映假设模型的适切性。当必须进行不同竞争模型（competing models）的比较时，或是从多个模型当中挑选一个最佳模型时，ECVI指数与另外两个类似的指数——Akaike讯息指数（Akaike information criterion, AIC）以及Akaike一致讯息指数（Consistent Akaike information criterion, CAIC）——是非常理想的指数，因为ECVI指数与AIC指数可以同时考虑样本量与模型复杂度的非中央性参数。

一般的，ECVI、AIC与CAIC指数越小表示模型越简效。此时如果模型可以具有一定的拟合度，则越简效的模型越理想。换言之，这些指数可以作为模型选择的依据。可惜的是，这两种指数的数值范围不是0~1，因此仅适合进行模型间的比较，从AIC、CAIC或ECVI数值的绝对值大小，很难判定个别模型的拟合性。

第四节 模型拟合评鉴

二、模型评鉴的方法

3. 替代指数

⑤CN指数

LISREL报表提供了一个特别的模型拟合统计量——关键样本指数（Critical N，CN）。CN指数是由Hoelter (1983)提出的，用以说明样本规模的适切性。其原理是估计若要产生一个适当的模型拟合度（不显著的卡方统计量），所需要的样本量为多少。Hoelter (1983)认为，当CN指数大于200时，表示该模型可以适当地反映样本的数据。Byrne (1998)主张除了200是一个门槛之外，一个研究的样本量只有大于CN指数所估计出来的样本量，该SEM分析才具有合理性。

第四节 模型拟合评鉴

二、模型评鉴的方法

4. 残差分析指数

一般而言，SEM分析均提供了两种残差的数据：非标准化残差（un-standardized residuals）与标准化残差（standardized residuals）。非标准化残差就是假设模型与观察数据之间差距的原始量数，也就是参数估计无法反映实际观察数据的变异量（unexplained variance or covariance）。在LISREL中，残差原量尺估计数列于Fitted Residuals报表中。标准化残差将残差量转换为标准Z分数，因此残差值将大约落于±3.5的区间中。

使用非标准化残差来了解各估计数良好与否的优点是可以直接应用测量的原始量尺，来了解残差数值大小的具体意义。然而，这不利于相互的比较。当不同的观察变量的量尺或单位趋近一致时，对非标准化残差的直接比较无异议；但是当观察变量的性质差异较大或使用不同的量尺或单位时，对残差的数值就无法直接加以比较，而需先行转换成为标准化分数（例如，Z分数），也就是取标准化的残差。将所有的残差置于相同的量尺单位之上，才能显示出残差的比较意义。

对SEM的标准化残差分析，与其他统计分析（例如，多元回归分析）的做法类似。当标准化残差大于+3时，代表该估计变异量或协方差不足；当标准化残差小于-3时，代表该估计变异量或协方差对于两个观察变量的共变有过度解释的现象。两者都会造成模型不良拟合。

第四节 模型拟合评鉴

二、模型评鉴的方法

4. 残差分析指数

残差的大小反映了不良拟合的问题，除了从数值大小来判断，还可以通过绘图的方式来观察残差变动的趋势。例如，使用标准化残差分布Q图可以看出每一个估计残差的差距分布图。如果是一个好模型，标准化残差Q图内的观察点应呈现水平直线，当观察点分布在水平直线之外时，代表有不寻常的估计数存在。严重的残差除了可能来自该估计数不适当的理论假设，也可能来自非正态化的观察数据。因此，在诊断残差的大小之时，应配合观察数据的正态性检验，才能得知是不是假设模型界定有问题，或观察数据本质有问题。目前，各主要的SEM分析软件均提供残差均方根指数（root mean square residual, RMR）与标准化残差均方根指数（standardized root mean square residual, SRMR）来反映理论假设模型的整体残差：

$$RMR = \sqrt{2\sum_{i=1}^{q}\sum_{j=1}^{i}\frac{\left(s_{ij}-\hat{\sigma}_{ij}\right)^2}{q(q+1)}}$$

其中，$s_{ij}-\hat{\sigma}_{ij}$代表样本（观察）与估计（理论假设模型）的方差或协方差差异。RMR与SRMR越小，代表模型越能拟合观察值。由于RMR基于未标准化残差值计算得出，其数值没有标准化的特性，较难解释。因此，学者多采用标准化后的SRMR指数来评估模型的优劣。SRMR指数的数值介于0~1，当数值低于0.08时，表示模型拟合度佳。

第四节 模型拟合评鉴

二、模型评鉴的方法

5. 拟合指数的比较与运用

①模型评鉴指数的比较

目前最常见到的拟合度评估策略除了卡方值及卡方显著性、卡方自由度比这两种传统方式之外，还有CFI与RMSEA指数。有关各项指数运用的时机与判断标准，如表所示。

指标名称与性质	范围	判断值	适用情形
卡方检验			
χ^2 test 理论模型与观察模型的拟合程度	–	p>0.05	说明模型解释力
χ^2/df（Wheaton et al.） 考虑模型复杂度后的卡方值	–	1～3	不受模型复杂度影响
适合度指数			
GFI（Bentler，1983） 假设模型可以解释观察数据的比例	0～1	>0.90	说明模型解释力
AGFI（Bentler，1983） 考虑模型复杂度后的 *GFI*	0～1*	>0.90	不受模型复杂度影响
PGFI（Mulaik，1989） 考虑模型的简效性	0～1	>0.50	说明模型的简单程度
NFI（Bentler & Bonett，1980） 比较假设模型与独立模型的卡方差异	0～1	>0.90	说明模型较虚无模型的改善程度
NNFI（Bentler & Bonett，1980） 考虑模型复杂度后的 *NFI*	0～1*	>0.90	不受模型复杂度影响

第四节 模型拟合评鉴

二、模型评鉴的方法

5. 拟合指数的比较与运用

①模型评鉴指数的比较

目前最常见到的拟合度评估策略除了卡方值及卡方显著性、卡方自由度比这两种传统方式之外，还有CFI与RMSEA指数。有关各项指数运用的时机与判断标准，如表所示。

指标名称与性质	范围	判断值	适用情形
替代性指数			
NCP（Bentler，1988） 假设模型的卡方值距离中央卡方分布的离散程度	–	越接近0，越好	说明假设模型距离中央性卡方的程度
CFI（Bentler，1988） 假设模型与独立模型的非中央性差异	0～1	>0.95	说明模型较虚无模型的改善程度 特别适合小样本
RMSEA（Browne & Cudeck，1993） 比较理论模型与饱和模型的差距	0～1	<0.05	不受样本量与模型复杂度影响
AIC（Akaike，1987） 经过简效调整的模型拟合度的波动性	–	越小越好	适用于效度复核 非嵌套模型比较
CAIC（Akaike，1987） 经过简效调整的模型拟合度的波动性	–	越小越好	适用于效度复核 非嵌套模型比较
CN（Hoelter，1983） 产生不显著卡方值的样本规模	–	>200	反映样本规模的适切性
残差分析			
RMR 未标准化假设模型的整体残差	–	越小越好	了解残差特性
SRMR 标准化假设模型的整体残差	0～1	<0.08	了解残差特性

注：*指数数值可能会超过范围。

第四节 模型拟合评鉴

二、模型评鉴的方法

5. 拟合指数的比较与运用

②模型拟合指数的报告

Hu与Bentler (1999)主张将CFI与RMSEA都报告在论文中。尤其是RMSEA指数，当研究者想去估计统计检验力时特别适合。另外，当研究者想要比较不同的模型，但是没有嵌套关系时，则可使用ECVI、AIC或CAIC指数。

在报告模型拟合度的评鉴数据时，McDonald与Ho (2002)建议，必须将测量模型与结构模型的模型拟合度分开列举说明。Anderson与Gerbing (1988)曾经提出一个二阶段估计程序（two-stage procedure），他们将结构模型视为测量模型的嵌套模型，将一个完整的SEM模型的测量与结构部分的模型拟合度评鉴区分为两个独立的步骤来进行，如此即符合McDonald与Ho (2002)的建议。

如果不是使用两阶段估计程序，在ML与GLS这两种估计程序中，一个兼含测量模型与结构模型的SEM模型的卡方值可以直接切割成两个独立的卡方值，一个反映测量模型，另一个反映结构模型，然后分别利用相对应的卡方分布进行显著性检验。