

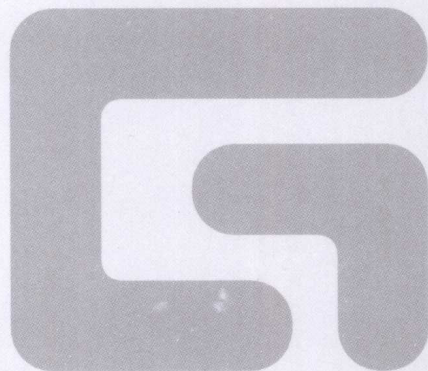
万卷方法

社会科学研究方法经典译丛



CRC Press  
Taylor & Francis Group

GENERALIZED LATENT  
VARIABLE MODELING :  
MULTILEVEL, LONGITUDINAL,  
AND STRUCTURAL  
EQUATION MODELS



# 广义潜变量模型

——多层次、纵贯性以及结构方程模型

Anders Skrondal 著  
Sophia Rabe-Hesketh

陈华珊 叶鹏飞 李洋 译



重庆大学出版社

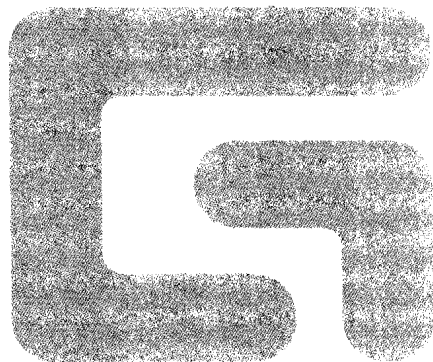
<http://www.cqup.com.cn>

万卷方法

社会科学研究方法经典译丛

主编 沈崇麟 夏传玲

GENERALIZED LATENT  
VARIABLE MODELING:  
MULTILEVEL, LONGITUDINAL,  
AND STRUCTURAL  
EQUATION MODELS



# 广义潜变量模型

——多层次、纵贯性以及结构方程模型

Anders Skrondal  
Sophia Rabe-Hesketh

著

陈华珊 叶鹏飞 李洋

译

重庆大学出版社

*Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*, by Anders Skrondal, Sophia Rabe-Hesketh, published by Chapman & Hall/CRC, Copyright © 2004 by Chapman & Hall/CRC

All Rights Reserved. Authorized translation from English language edition published by CRC Press, part of Taylor & Francis Group LLC.

版贸核渝字(2006)第 57 号

**图书在版编目(CIP)数据**

广义潜变量模型:多层次、纵贯性以及结构方程模型/(美)斯科隆多(Skrondal, A.), (美)赫斯基思(Hesketh, S. R.)著;陈华珊,叶鹏飞,李洋译. —重庆:重庆大学出版社,2011. 1

(万卷方法)

书名原文:Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models

ISBN 978-7-5624-5393-2

I. ①广… II. ①斯…②赫…③陈…④叶…⑤李…  
III. ①统计模型—研究 IV. ①C81

中国版本图书馆 CIP 数据核字(2010)第 089436 号

## 广义潜变量模型

——多层次、纵贯性以及结构方程模型

(美)安德斯·斯科隆多(Anders Skrondal)

(美)索菲亚·拉贝-赫斯基思(Sophia Rabe-Hesketh)

著

陈华珊 叶鹏飞 李洋译

策划编辑 雷少波 林佳木

责任编辑:李定群 高鸿宽 版式设计:林佳木

责任校对:贾梅 责任印制:赵晟

\*

重庆大学出版社出版发行

出版人:邓晓益

社址:重庆市沙坪坝正街 174 号重庆大学(A区)内

邮编:400030

电话:(023) 65102378 65105781

传真:(023) 65103686 65105565

网址:<http://www.cqup.com.cn>

邮箱:fxk@cqup.com.cn (营销中心)

全国新华书店经销

自贡新华印刷厂印刷

\*

开本:787×1092 1/16 印张:24.25 字数:619千 插页:16开2页

2011年1月第1版 2011年1月第1次印刷

印数:1—4 000

ISBN 978-7-5624-5393-2 定价:69.00元

---

本书如有印刷、装订等质量问题,本社负责调换

版权所有,请勿擅自翻印和用本书

制作各类出版物及配套用书,违者必究

## 作译者简介

---

安德斯·斯科隆多(Anders Skrondal)是英国伦敦经济学院统计学系教授。

索菲亚·拉贝-赫斯基思(Sophia Rabe-Hesketh)是美国加州伯克利大学教育学院教授及生物统计学系教授。

陈华珊:中国社会科学院社会学所助理研究员,从事社会学方法论及组织社会学研究。

叶鹏飞:中国劳动关系学院工会学系讲师,清华大学博士毕业,主要研究领域是劳工和城市社会学。

李洋:北京市社会科学院社会学研究所助理研究员,清华大学博士毕业。

陈华珊翻译第5章至第11章以及第13章和第14章。叶鹏飞翻译了第1章、第2章和第3章。李洋翻译了第4章和第12章。

# 总序

## 社会研究方法的现状及其发展趋势

---

近年来,社会调查技术和社会研究方法都有很大的发展。在调查技术方面,自 20 世纪 70 年代以来,社会变迁多次横断面的跟踪调查研究,几乎成为所有国家和地区了解社会结构转变和社会发展状况的基础性调查。这种调查不仅对社会学的研究有很大促进作用,而且对整个社会科学的研究都产生了重大影响,并且这些调查结果有的已作为政府有关部门决策的重要依据。国际上比较著名的此类调查有:美国芝加哥大学全国民意调查中心(National Opinion Research Center,简称 NORC)的“社会综合调查(General Social Survey,简称 GSS)”,英国埃塞克斯大学调查中心进行的“全国家庭生活和社会变迁调查”,法国经济和社会调查所进行的“全国经济社会调查”,日本社会学学会组织进行的“全国社会分层与社会流动调查(简称 SSM)”。中国台湾“中央”研究院社会学研究所,也每两年进行一次“台湾社会变迁基本调查”。美国的“社会基础调查”,现在已成为年度性的调查项目,它是美国国家基金会目前资助的最大的社会科学研究项目。以上这些调查,除美国的调查外,一般均因经费原因采用纵向的间隔性重复调查法,即每隔一段时间,进行一次全国规模的抽样调查。每次调查除保留社会研究所需的基本项目外,都有不同的主题。在间隔若干时间后,再重复同一主题的调查,这样的研究设计,使社会变迁研究在可以涉及更为广泛的研究领域的同时,具有更好的积累性和可比性。多年来,这些基础性调查获得的资料,滋养着大批的社会科学研究者,有时一项调查就有很多名博士生用来写博士论文,以此取得的研究成就,其可靠性受到社会科学界的广泛认同。例如 1997 年出版,以台湾地区社会变迁基本调查数据为基础的研究报告集《90 年代的台湾社会,社会变迁基本调查研究系列二》收集论文 16 篇,内容涉及社会生活的各个方面,在台湾地区引起了极大的反响。

国内社会科学界在这方面也有了长足的发展。笔者所在的中国社会科学院社会学研究所的社会调查和方法研究室,组织或参与了多项与社会变迁有关的大规模抽样调查,取得了一定的研究成果,并积累了大量有关社会变迁的宝贵数据资料,其中主要有:

1. 城乡家庭变迁系列调查:该课题是由中国社会科学院社会学研究所牵头,联合北京大学和地方社科院的研究人员展开的一项类似多次横断面的城乡家庭变迁调查。这一调查始于 1981 年的“中国五城市婚姻家庭调查”,而后有 1988 年的“中国农村家庭调

查”、1991年的“中国七城市家庭调查”、1998年的“中国城乡家庭变迁调查”。

2. 有关中国城乡社会变迁的系列调查:这一调查始于1991年的第二批国情调查,然后有1992年的“中国城乡居民生活调查”、1993年的“第三批国情调查”、1995年的“第四批国情调查”和1997年的“中国沿海发达地区社会变迁调查”。上述调查虽然还不是严格意义上的多次横断面的纵贯研究,但研究者已在研究设计中尽量考虑到纵贯研究的基本原则,如调查队伍的稳定、指标的可比性和样本空间的延续性等。

3. 中国城乡社会变迁调查:这一调查始于2000年,为中国社会科学院重大课题。目前已经完成第一期第一次调查和第二次调查,今后将把这一调查发展为连续的、定期进行的社会变迁调查。

在纵向调查技术取得长足进步的同时,20世纪末至今,电话调查也有很大发展。电话调查涉及的范围几乎与个别(面对面)访谈同样全面。电话调查中使用的一系列方法,是在20世纪70年代后期和面对面调查一起发展起来的。在20世纪80年代中期,电话调查开始变得很普遍,并且成为许多场合中各种调查方法的首选。正如某些学者所言,一种在公共和私营部门被人们用来帮助提高决策效率的收集信息的有效方法为人们所普遍认同时,这一现象本身就具有方法论上的意义。不仅如此,电话调查还有很大的实践意义,因为它为研究者提供了更多的控制调查质量的机会。这一机会包括抽样、被调查人的选择、问卷题项的提问、计算机辅助电话访谈(CATI)和数据录入。正因为如此,今天在各种社会调查中,如果没有发现其他重要的足以放弃使用电话调查的原因,电话调查由于其独特的对调查质量进行全面监控的优点,常常成为各种调查方式的首选。由笔者翻译,重庆大学出版社出版的《电话调查方法:抽样、选择和督导》一书,也于2005年面世。

无论是纵向调查抑或电话调查,实际上都是收集研究资料的方法,而应用社会科学的发展,不仅在于调查技术,即收集资料技术的发展,还在于研究方法和分析技术的发展。近年来,无论是定性研究方法,还是定量研究方法都有了长足的发展。

首先,计算机技术的发展可谓突飞猛进,它对当今社会生活的各个方面产生了巨大的影响,在悄悄地改变着社会科学的研究风格和研究方式的同时,也大大提升了社会科学学者的研究能力。这种影响表现在研究过程的各个阶段,从理论建构(概念映射)、问卷设计(专业的问卷设计软件)、调查实施(计算机辅助访谈、计算机辅助电话访问系统、网络在线调查系统)、数据录入(光学标记识别软件)到数据分析(包括文本、声音、图像资料的处理),甚至延伸到写作发表阶段。这样的过程发生在如社会学、经济学、政治学、心理学、教育学中,促进了学科之间的相互借鉴和交叉融合,至少在研究方法上呈现出这种趋势。随着计算机计算能力的大幅度提高,20世纪80年代后期,统计学领域内发生了一场“革命”,主要表现在对定类和定序变量的建模能力的大幅度提高上,以及与分布无关的统计分析模型的发展之上,特别是基于“Resampling”(包括Bootstrap、Jackknife、Monte Carlo模拟等)的建模技术。同时,计算能力的提高还带动了基于神经网络、动态模拟、人工智能、生态进化等新兴的分析和预测模型的发展。这些进展都为定量社会科学研究提供了更多的可供选择的工具。

亚德瑞安·E. 拉夫特里(Adrian E. Raftery)依据社会学家所处理的数据类型,将定量社会学在美国的发展划分为三个时代:第一代起始于20世纪40年代,交互表是其主要处理对象,研究重点是关联度和对数线性模型;第二代起始于20世纪60年代,主要处理单层次的调查数据,Lisrel类型的因果模型和事件史分析是其研究重点;第三代起始于20世纪80年代后期,开始处理诸如文本、空间、社会网络等非传统的数据类型,目前尚没有

形成成熟的形态。拉夫特里的综述,虽然更强调定量社会学研究对统计学的贡献,但也大致勾勒出定量社会学在国外的发​​展脉络。

从分析模型的角度来看,定量分析在以下几个方向有了突破性发展:

1. 缺失值处理:由于社会生活的复杂性,社会调查数据常常出现缺失值,传统的处理方式是忽略这些缺失值,或者用均值替代。但现在则倾向于用多重插值法(multiple imputation)或者其他基于模型的方法进行处理。这些技术的发展,不仅会增强我们对数据的处理能力,而且将改变我们设计问卷的方式。基于这些技术,我们在不增加受访者负担的前提下,大大增加了调查问卷的内容:每个受访者只回答问卷的一部分,然后通过对缺失值的处理,获得他们对未回答部分的估值。

2. 非线性关系:线性假定是经典定量分析的一个常见假定,但在实际研究当中,线性假定只能被看作是对社会现实的一个逼近和简化。面对具体的研究数据,如果没有理论上的明确指引(不幸的是,我们常常没有中程理论的指引),我们是无法在线性模型和非线性模型之间作出取舍的。但 MARS 模型的出现,让我们可以从经验数据当中获得最为拟合的变量之间的函数关系,而不必预先作出线性假定。这样,理论思考 and 数据分析就可以实现一个互动的循环过程,定量分析就不单单是对理论和假设的简单证伪过程,而是理论思维一个重要组成部分。

3. 测量层次:20 世纪六七十年代的统计模型,大多要求数据的测量层次在定距以上,如因素分析,但社会学的调查数据却大多为定类或定序数据。对应分析、Loglinear、Logit、Logistic Regression、潜类分析、Ordinal Regression、Normal Ogive Regression 等统计模型的出现,大大提高了定量社会学处理定类和定序数据的能力。

4. 测量模型:基于文化、社会、心理和认知等方面的考虑,在社会学界仍有人对问卷调查在中国的效度提出质疑。抛弃“本土化”的文化执著,我们更应当关注的是问卷调查的项目反应理论(item response theory),即受访者回答问卷题器时的过程模型。这方面的进展主要表现在两个方面:一是分解测量量表的成分,如 Rasch model、IRT 分析、Mokken 分析等;二是将测量模型与因果模型或其他分析模型结合在一起,明确把测量误差引入到分析当中,充分评估它们对分析结果的影响,如结构方程模型。

5. 潜变量模型:与测量模型相关联的另外一个发展方向是潜变量模型,例如,潜变量分层分析(latent class analysis)、潜变量结构分析(latent structure analysis)、潜变量赋值分析(latent budget analysis)等。“潜变量”这一概念表明,我们可以通过测量“显变量”来测量无法直接观察的理论概念,如权力、声望、地位等。这样,理论和现实之间,通过“潜变量”到“显变量”的映射(测量过程),就有了连接的桥梁。

6. 分析单元的层序性:在定量分析当中,我们常常强调要避免出现“生态谬误”,即分析单元的层次和结论或推论的层次不一致。与其相关的方法论争论是“宏观和微观”的问题。随着多层次模型的出现,我们可以同时考察多个层次上的问题,我们可以把个人放在其家庭背景中,再把家庭放在社区的背景下,考察个人层次的变量对社区变量的效应,或者社区层次的变量对个体行为的具体影响。在定量分析模型当中,“宏观和微观”的连接获得了建模技术上的支持。在这个领域当中,还有一个方向也值得关注:分析宏观层次的数据,对微观层次进行推论。

7. 社会网络模型:区分“关系数据”和“属性数据”,是把分析重点从个体/群体等社会单元转移到这些社会单元之间关系的第一步,社会网络模型是目前发展较快的一个定量分析领域,其理论根基是结构主义。社会网络分析目前仍然具有较浓厚的“形态学”特征(基于图论的缘故),但却为我们理解社会关系在社会空间上的形态奠定了基础,通过计



计算机模拟和研究社会网络的历期数据,研究社会结构的“发生学”性质模型也处在萌芽状态当中。

8. 系统动力学:如果说社会网络模型是在社会空间上拓展定量社会学的研究手段,那么社会过程在时间上和物理空间上的属性,则是事件史模型、事件数模型、历期分析、Cox 回归、时间序列分析、Cohort 分析、状态空间模型等模型的研究对象。在这个领域,计量经济学为定量社会学研究提供了许多有益的范例。

9. 预测模型:上述模型仍然是在分析主义的范式下。有些社会学的应用研究,更强调模型的预测精度,而不是模型的认知价值,例如,社会趋势的预测。由于计算能力的提高,神经网络、基因算法、人工智能、模式识别等数据挖掘技术有了长足发展,已经出现了许多拟合经验数据的预测模型,比较成功的应用出现在计量经济学领域(如对股市的预测)。

10. 计算机模拟:对于社会学应用研究而言,研究的对象具有历史性、规模大、变迁的过程不仅漫长且表现某种渐进性的特点,且因社会隔离/社会伦理原因无法接近或有实验禁忌等,无法直接进行观察和研究,这时计算机模拟就成为一个可供选择的替代方案。计算机模拟主要有两个类型:一是基于计算机网络的模拟:每台微机作为一个代理,整个网络作为“社会”实时演化,如法国的 Swarm 计划;二是基于概念模型的系统,在计算机时间上,按照既定规则运行,较有名的研究是罗马俱乐部的《增长的极限》,常见的软件有 Simul, Arena 等。自然科学家对此方向似乎比社会学家更有兴趣。

定性研究方法一直是社会学研究领域比较传统的研究方法,在社会学研究的古典时期,它甚至是社会学家手中唯一的研究方法。但随着定量研究方法在社会学研究中的广泛应用,定性研究方法就似乎越来越不受人们的重视。但需要澄清的事实是,在定量分析模型取得飞速发展的同时,在过去的二十多年里,定性研究方法也有了长足的进步。主要表现在以下六个方面:

1. 研究素材日益扩大:除了传统的参与观察、深度访谈、专题小组访谈之外,会话、交谈、电视、广播、文档、日记、叙事、自传 (autobiography) 等社会过程中自然产生的素材,甚至社会学理论本身 (理论的形式化),也开始进入定性分析的视野当中。所有这些资料,不仅可以以文本的格式存储,而且,新型的多媒体介质,如图像、声音和视频,作为原始的分析素材,也日益成为定性分析的新宠。

2. 分析方法更加多样:定性方法的种类在最近的二十多年中,更是有了一个质的飞跃。在比较传统的、源自语言学的方法,如内容分析、话语分析、修辞分析、语意分析、符号学、论据分析等方法之外,社会学家也创造出自己独特的定性分析方法,如施特劳斯 (Strauss) 等人的扎根理论、海斯 (Heise) 的事件结构分析、拉津 (Ragin) 的定性对比分析、Abbott 和 Hrycak 采用最优匹配技术的序列分析、亚贝儿 (Abell) 的形式叙事分析 (formal narrative analysis)、鲍尔 (Bauer) 等人的语库建设、Attride-Stirling 等人的主题网络分析和神经网络技术应用的定性分析领域。所有这些方法的一个共同特征是,把定性研究向更加系统、更加精确、更加严格、更加形式化的方向推进。

3. 认识论基础更加多元化:现象学、释义学和本土方法论 (ethnomethodology) 的认识论,一直是定性分析的大本营,但近年来,实证主义也开始逐渐为定性分析所接纳,解释和阐释之间,由激烈的对立关系,逐渐演变为相互融洽的关系。

4. 研究过程更加客观规范:定性分析的一个主要问题在于阐释过程中不可避免的主观性。为了尽可能消除“解释者偏见”和主观选择性,定性分析开始遵循严格的程序模板或程序规则,并尝试引入定量分析中的“信度”“效度”“代表性”等概念,通过编码和对比,再加上传统的定



性分析标准,如可解释性、透明性和一致性,使得定性研究的过程更加规范、阐释的结果更加客观,研究的结论更加可信。

5. 研究过程更加有效率:这主要应归功于大量计算机辅助定性数据分析(CAQDA)软件的涌现。从20世纪80年代以来,定性分析过程的数字化和计算机化,已经是一个不可逆转的大趋势。这种发展趋势与定性研究者的理论取向无关,不管他们的理论立场是实证主义、符号互动论,还是本土方法论,大多数定性研究者都在自己的研究当中,开始采用计算机来辅助定性资料的分析过程。据不完全统计,目前已经有二十多种定性分析的软件,分别隶属于德国、英国、法国、美国等国家。其中,有一些软件是国外研究机构的科研成果,可以免费使用,但比较成熟的定性辅助系统大多是商业软件。这些定性分析的辅助系统,不仅使得研究者从处理大量文字材料的繁复劳动中解放出来,而且能够让研究者共享他们各自分析的细节,从而改变定性研究的流程和研究集体之间的合作方式。同时,由于采用数据库结构,定性资料的管理也更加方便,这就为组织大型定性研究项目(包括多个研究地点、多个研究对象、历时的定性研究)提供了新的可能性。越来越多的定性研究人员开始走出他们的摇椅,坐到计算机屏幕前、湮没在访谈资料和故纸堆中的定性社会学家的形象已经一去不复返了。

6. 定性研究和定量研究的结合更加紧密:在定量分析方法的教材中,定性研究常常被看做是定量研究的前期准备工作,但定性研究者却持完全相反的观点,他们一般认为定性方法是自成一体的,可以完成从形成概念到检验假设的全部研究过程。在实际的应用研究中,定性方法和定量方法常常是交织在一起的,例如,克劳(Currall)等人在研究组织环境重要的群体过程时,通过内容分析把5年的参与观察资料量化,然后用统计分析来检验理论假定。格雷(Gray)和邓斯坦(Densten)在研究企业的控制能力时,利用潜变量模型把定性方法和定量方法有机结合在一起。雅各布斯(Jacobs)等人在研究比利时的家庭形态对配偶的家庭劳动分工影响时,首先用定量方法对纵向调查数据进行分析,从定量分析的结果中,又延伸出对核心概念的定性研究。这三个研究分别代表了定量和定性方法相互融合的三个方向:①克劳等人的研究代表着定性方法的实践者试图将定性数据尽可能量化的取向,近年来涌现出的处理调查数据中开放题器的编码问题的工具软件(如 Words at, Smarttext 等,注意:它们都是由著名的统计软件公司出品的处理定性资料的软件),处理定性资料的传统内容分析软件(如 Nvivo、MaxQDA、Kwalitan 等)也开始提供将定性资料转换到常用统计软件的数据接口,这些工具上的革新将加快这种趋势的发展。②格雷和邓斯坦的工作代表了“方法论多元论”的取向,即在应用研究过程中,通过核心概念的测量模型,把定性研究和定量研究结合在一起。③雅各布斯等人的工作则代表了一部分定量研究者对过度形式化的定量方法的不满,并试图通过定性方法加以弥补。在定量研究领域,对“模型设定”问题的关注,是定量方法重新试图返回定性研究这种取向的另外一种表现。

与社会调查技术和社会研究方法突飞猛进的现实相比,我国学术界在这些方面的论著的出版似乎显得有些迟缓。虽然已经翻译了美国的一小部分经典定量分析教材,如布莱洛克(Blalock)和巴比(Babie)的教材,也有自己编写的一些教材,如袁方等人的《社会研究原理和方法》、卢淑华的《社会统计学》等,此外,偏重软件操作的还有郭志刚的《社会统计分析方法——spss 软件应用》、郭志刚的《logistic 回归模型——方法与应用》、阮桂海的《spss for windows 高级应用教程》等。在《社会学研究》等专业杂志上,也常常有一些定量分析的应用研究,可是专门的方法和模型研究却没有,也没有专门的方法研究期刊。仅就定量研究方法的介绍而言,也存在一些缺陷,主要表现在:

1. 原理和操作脱节。
2. 过分依赖某些商业软件,不全面。
3. 与中国的实证研究相脱节。
4. 不能反映当前方法研究的最新进展。

与定量研究方法相比,由于各种原因,定性研究方法的引进和介绍都比较少。在福特基金会资助的方法高级研讨班上,曾讨论过一些定性研究方法。在定性方法研究方面也有少数专著,如袁方和王汉生 1997 年出版的教程,陈向明 2000 年出版的专著。但总体说来,我们对定性研究方法还停留在初步介绍的阶段,主要的介绍也局限在定性研究的研究设计和资料收集的阶段上,对定性分析方法的介绍,则没有能够反映出当代定性方法的最新进展。特别是在定性分析工具(定性分析软件)的引进和研究上,基本上还是一个空白。虽然不乏一些出色的定性研究报告,但从方法研究上讲,我们才刚刚起步。当然,我们同时还应该注意到,在历史学领域,我国对定性资料的鉴别、考据和分析,积累了大量的经验和知识,这也应当是定性方法研究的知识来源之一,应努力发扬光大。

令人欣慰的是,社会研究方法的引进和出版方面相对滞后的状况终于有所改观。重庆大学出版社的编辑,以独到的学术眼光,逆当前出版界唯利是图的不良选题风气,投入了大量的人力、物力,组织出版“万卷方法”。自 2004 年至今,已引进社会科学研究方法方面的专著十余种,在我国社会科学界已经引起了一定的反响。然而,更为可贵的是,重庆大学出版社并未以已经取得的成绩而自满,而是再接再厉,在原有“万卷方法”的基础上,进一步组织出版“万卷方法—社会科学研究方法经典译丛”。按我们的设想,“译丛”应该是一个开放的体系,旨在跟踪社会科学研究方法发展的前沿,引进和介绍这一方面的经典著作和最新成果。

“译丛”第一批有《抽样调查设计导论》《社会科学研究设计原理》《社会科学研究测量原理》《社会科学研究分析技术》《问卷设计手册》《回归分析法》《数据再分析法》《抽样调查设计导论》《社会网络分析法》《广义潜变量模型》《定性变量数据分析》和《复杂调查设计和分析方法》(书名也许有变化)等十余种,几乎囊括了研究设计、测量和分析方法的所有领域,涵盖从基础的回归分析到最前沿的潜变量分析和多水平模型等各种分析方法。无论是社会科学各专业的本科生、研究生,还是社会科学研究的学者都将从中有所收获。

“译丛”由中国社会科学院社会学所社会调查与方法研究室的多位研究人员担纲,主译者都是在社会研究方法各个领域中具有相当造诣的教师和研究人员。“译丛”的译者不仅仅把翻译看做是一个“翻译”,而且也把它看做是一次再学习和再创新。

我们期待“译丛”的出版能对社会研究方法的研究、应用和教学有所推动。

沈崇麟 夏传玲

2010 年 12 月于中国社科院社会学所社会调查与方法研究室

# 前 言

本书的主要目标是在尽可能广的范围上统一及扩展潜变量建模。本书所涵盖的模型包括多层次模型、纵贯性模型以及结构方程模型,另外还包括与它们相关的模型,如广义线性混合模型、随机系数模型、题器响应模型、因子模型、面板模型、重复测量模型、潜类模型及脆弱模型。

潜变量模型用于绝大多数经验研究学科当中,尽管常常被其他的名称所代替。本着本丛书的题目——“跨学科统计学丛书”,我们试图综合不同学科的视角,并在统计学、生物统计学、计量心理学以及计量经济学这几个学科的语言中进行转换(尽管我们并未声称掌握了所有这些学科的语言)。

我们认为模型的进展被这种“地方”行话所妨碍,并产生了隔阂。例如,计量经济学家和生物统计学家之间很少浏览对方的学术期刊。更加令人惊讶的还有学科内的部落化,这一点从《计量心理学》期刊上题器响应模型和因子模型之间缺乏交叉引用可以反映出来(尽管在同一个期刊上)。缺乏沟通的其中一个有害后果就是无法知晓对方学科所取得的有用进展,直到它们被“翻译”过来并发表在“适当的”文献上。例如,关于退出的模型(在社会科学中称为损耗)在过去的几十年中已经在计量经济学中取得了突出的进展,然而直到最近,它们才在统计学文献中被“发现”。

本书由两个部分所组成:方法论与应用研究。在第1章,我们讨论潜变量的概念、用途以及解释。在第2章,我们将不同学科所使用的针对不同响应类型的模型汇总在一起。在第3章对经典潜变量模型进行了回顾之后,我们在第4章将第2章所考察过的所有响应类型的模型进行统一及扩展。已有的以及最新的关于模型辨识、模型估计、潜变量预测以及模型诊断的方法将详尽地在第5章到第8章进行阐释。

在应用研究部分的第9章到第14章,我们使用第1部分所发展的方法来分析来自生物学、医学、心理学、教育学、社会学、政治学、经济学、市场营销以及其他领域的问题。所有的应用研究都使用真实数据,但出于教学原因,我们的分析通常是简化了的。在所有这些应用中,我们都使用了与安德鲁·皮科尔斯合作开发的 Stata 程序包 `gllamm`。

我们希望本书的两个部分中丰富的交叉索引可以让读者既容易在应用章节找到方法论的例子(忽略前半部分),也容易在方法论章节中找到应用研究的统计学背景(忽略后半部分)。

前面的3章多内容是作为对贯穿本书的建模视角的相对温和的介绍。后续的方法论章节具有更强的技术性,部分原因也在于本框架的通用性。然而,我们也同样考虑更简单的特殊例子,使得标记不那么复杂,想法更加明晰,结果更加符合直觉。对于潜变量模型的解释及应用更加感兴趣的读者可以忽略第4章到第8章的大多数内容,他们可关注于应用章节部分。

本书是过去4年中我们在合作开发“广义线性潜变量及混合模型(GLLAMM)”及其软件的一其中个成果。我们感谢我们的合作者安德鲁·皮科尔斯所做的贡献。安德斯感谢他的“老板”——皮埃·马格纳斯,索菲亚感谢她的前“老板”布莱恩·艾弗瑞特,是他们给予了支持并且接受本书的工作优先于其他项目。布莱恩·艾弗瑞特、莱昂纳多·格里利、卡拉·拉姆皮奇尼以及另外两个匿名评审员阅读了本书稿,并给予我们许多有用的建议。另外,我们还要感谢艾瑞特·艾特金、比尔·布朗尼、斯蒂芬·詹金斯、安德鲁·皮科尔斯、斯文·奥韦·萨缪尔森以及杰罗恩·沃芒所给予的建设性评论。

大卫·克雷顿、皮埃·卡拉格·安德森、安东尼·希思、安德鲁·皮科尔斯以及本特·特拉恩为本书的应用研究部分提供了数据。我们感谢帕特里克·赫加蒂、BUGS项目组、穆森夫妇、应用计量经济学杂志以及皇家统计学会将数据发布到互联网上以供访问。在海琳·罗肖夫的帮助下,我们还使用了来自英国数据中心以及挪威社会科学数据服务中心的数据。感谢查普曼-霍尔出版社的杰斯敏·奈姆和克斯蒂·斯特劳德,是他们的频繁而温和的提示保证本书取得稳固的进展。我们还感谢LATEX的开发者提供了这个不可估价的免费工具帮助我们准备书稿。

在每一个章节中,我们都共同写作,为本书做出几乎相同的贡献。尽管本书的写作是一项非常辛苦的工作,但我们在此过程中也享受了很多乐趣!

gllamm 软件、文档等可从<http://www.gllamm.org> 下载。本书应用研究中的部分数据和脚本可从<http://www.gllamm.org/books> 下载。

安德斯·斯科隆多  
索菲亚·拉贝-赫斯基思

第 1 部分 方法论 .....	1
1 潜变量无所不在 .....	3
1.1 介绍 .....	3
1.2 带有测量误差的“真实”变量 .....	4
1.3 假设构念 .....	5
1.4 未观测到的异质性 (Unobserved heterogeneity) .....	9
1.5 缺失值与反事实 .....	12
1.6 潜响应 .....	13
1.7 产生灵活分布 .....	14
1.8 组合不同来源的个体单元信息 .....	15
1.9 总结 .....	15
2 对不同响应过程建模 .....	16
2.1 介绍 .....	16
2.2 广义线性模型 .....	17
2.3 广义线性模型的扩展 .....	21
2.4 潜响应方程 (Latent response formulation) .....	25
2.5 存续或生存的建模 .....	30
2.6 总结与进一步阅读 .....	36
3 经典潜变量模型 .....	37
3.1 介绍 .....	37
3.2 多层次回归模型 .....	37
3.3 因子模型和题器响应模型 .....	48
3.4 潜类模型 .....	56
3.5 带有潜变量的结构方程模型 .....	57
3.6 纵贯性模型 .....	60
3.7 总结与进一步阅读 .....	70

4	一般模型框架	71
4.1	介绍	71
4.2	响应模型	72
4.3	潜变量的结构模型	81
4.4	干扰项分布	84
4.5	参数约束和基本参数	87
4.6	潜变量和线性估计量的简化式	89
4.7	潜变量的矩结构	90
4.8	观测响应和潜响应的边缘矩结构	91
4.9	简化式分布和似然	94
4.10	简化式参数	96
4.11	总结与进一步阅读	98
5	辨识与等价	99
5.1	介绍	99
5.2	辨识	100
5.3	等价	114
5.4	总结与进一步阅读	117
6	估计	119
6.1	引言	119
6.2	最大似然:封闭形式的边缘似然性	120
6.3	最大似然:近似边缘似然	121
6.4	似然最大化	129
6.5	非参数最大似然估计	136
6.6	受约束/残差最大似然(REML)	137
6.7	有限信息方法	138
6.8	最大伪似然估计	142
6.9	广义估计方程(GEE)	146
6.10	固定效应方法	148
6.11	贝叶斯方法	152
6.12	总结	158
7	潜变量赋值	162
7.1	介绍	162
7.2	后验分布	163
7.3	经验贝叶斯(EB)	165
7.4	经验贝叶斯众数(Empirical Bayes modal, EBM)	172
7.5	最大似然法	174
7.6	在“线性案例”中几种记分法的关系	176
7.7	专门的记分方法	177
7.8	潜变量计分和分类的使用	178

7.9	总结与进一步阅读	181
8	模型设定与推论	183
8.1	介绍	183
8.2	统计建模	183
8.3	推论(基于似然)	187
8.4	模型选择:相对拟合标准	191
8.5	模型充分性:全局绝对拟合标准	195
8.6	模型诊断:局部绝对拟合标准	198
8.7	总结与进一步阅读	204
<b>第2部分</b>	<b>应用研究</b>	<b>205</b>
9	二项响应	207
9.1	介绍	207
9.2	儿童呼吸道感染研究:随机截距模型	207
9.3	心肌梗塞诊断:潜类模型	210
9.4	数学推理:题器响应模型	212
9.5	尼古丁咀嚼胶与戒烟:元分析	218
9.6	妻子的就业转换:马尔可夫模型及未观测到的异质性	225
9.7	白靴兔计数:捕获-再捕获模型及异质性	228
9.8	对堕胎的态度:多层次题器响应模型	231
9.9	总结与进一步阅读	234
10	定序响应	235
10.1	介绍	235
10.2	性教育群组随机试验:潜增长曲线模型	235
10.3	政治效能:因子维度及题器偏误	240
10.4	生活满意度:定序有衡正态概率因子模型(Ordinal scaled probit factor model)	248
10.5	总结与进一步阅读	256
11	事件数	257
11.1	介绍	257
11.2	儿童牙病防治:过度离散模型	257
11.3	癫痫症治疗:随机系数模型	262
11.4	苏格兰唇癌:疾病制图	266
11.5	总结与进一步阅读	274
12	存续与生存响应	276
12.1	介绍	276
12.2	多重事件群集存续数据的建模	276
12.3	吸烟的肇端:离散时间脆弱模型	279
12.4	锻炼和心绞痛:成比例风险随机效应和因子模型	282
12.5	总结与进一步阅读	289



13	比较响应:多分类响应及排序数据 .....	290
13.1	介绍 .....	290
13.2	异质性及“无关选项独立性” .....	290
13.3	模型结构 .....	291
13.4	不列颠普选:离散选择及排序的多层次模型 .....	293
13.5	后物质主义:排序的潜类模型 .....	298
13.6	咖啡机的消费偏好:结合选择模型(conjoint choice model) .....	303
13.7	总结与进一步阅读 .....	307
14	多过程响应与混合响应 .....	309
14.1	介绍 .....	309
14.2	饮食与心脏病:协变量测量误差模型 .....	309
14.3	生殖器疱疹与子宫颈癌:一项病例对照研究的潜类协变量测量误差 模型 .....	313
14.4	职业培训与压力:服从者平均因果效应模型 .....	316
14.5	医师建议及饮酒:内生试验模型 .....	320
14.6	肝硬化治疗:生存与指标联合模型 .....	324
14.7	总结与进一步阅读 .....	328
	<b>术语对照表</b> .....	329
	<b>人名对照表</b> .....	335
	<b>参考文献</b> .....	340

# 第 1 部分 方法论