

格致方法·定量研究系列 吴晓刚 主编




潜变量增长曲线模型

[美] 克里斯托弗·普里彻 (Kristopher Preacher)
阿斯荣·威克曼 (Asron Wichman) 著
罗伯特·麦卡勒姆 (Robert MacCallum)
南希·布里格斯 (Nancy Briggs)

姜念涛 译

- ★ 革新研究理念
- ★ 丰富研究工具
- ★ 最权威、最前沿的定量研究方法指南

格致出版社  上海人民出版社

28

格致方法·定量研究系列 吴晓刚 主编

潜变量增长曲线模型

克里斯托弗·普里彻(Kristopher Preacher)
[美] 阿斯荣·威克曼(Asron Wichman) 著
罗伯特·麦卡勒姆(Robert MacCallum)
南希·布里格斯(Nancy Briggs)
姜念涛 译

SAGE Publications, Inc.

格致出版社 上海人民出版社

图书在版编目(CIP)数据

潜变量增长曲线模型/(美)普里彻
(Preacher, K.)等著;姜念涛译. —上海:格致出版社;上海人民出版社,2012
(格致方法·定量研究系列)
书名原文:Latent Growth Curve Modeling
ISBN 978-7-5432-2157-4

I. ①潜… II. ①普… ②姜… III. ①统计模型-研究 IV. ①C81

中国版本图书馆 CIP 数据核字(2012)第 214227 号

责任编辑 顾悦

格致方法·定量研究系列

潜变量增长曲线模型

[美] 克里斯托弗·普里彻 阿斯荣·威克曼 著
罗伯特·麦卡勒姆 南希·布里格斯
姜念涛译

出版 世纪出版集团 **格致出版社**
www.ewen.cc www.hibooks.cn
上海人民出版社

(200001 上海福建中路193号24层)



编辑部热线 021-63914988

市场部热线 021-63914081

发行 世纪出版集团发行中心
印刷 浙江临安曙光印务有限公司
开本 920×1168 毫米 1/32
印张 5.5
字数 107,000
版次 2012年10月第1版
印次 2012年10月第1次印刷
ISBN 978-7-5432-2157-4/C·84
定价 15.00元

出版说明

由香港科技大学社会科学部吴晓刚教授主编的“格致方法·定量研究系列”丛书,精选了世界著名的 SAGE 出版社定量社会科学研究丛书中的 35 种,翻译成中文,集结成八册,于 2011 年出版。这八册书分别是:《线性回归分析基础》、《高级回归分析》、《广义线性模型》、《纵贯数据分析》、《因果关系模型》、《社会科学中的数理基础及应用》、《数据分析方法五种》和《列表数据分析》。这套丛书自出版以来,受到广大读者特别是年轻一代社会科学工作者的欢迎,他们针对丛书的内容和翻译都提出了很多中肯的建议。我们对此表示衷心的感谢。

基于读者的热烈反馈,同时也为了向广大读者提供更多的方便和选择,我们将该丛书以单行本的形式再次出版发行。在此过程中,主编和译者对已出版的书做了必要的修订和校正,还新增加了两个品种。此外,曾东林、许多多、范新光、李忠路协助主编参加了校订。今后我们将继续与 SAGE 出版社合作,陆续推出新的品种。我们希望本丛书单行本的出版能为推动国内社会科学定量研究的教学和研究作出一点贡献。

总序

往事如烟，光阴如梭。转眼间，出国已然十年有余。1996年赴美留学，最初选择的主攻方向是比较历史社会学，研究的兴趣是中国的制度变迁问题。以我以前在国内所受的学术训练，基本是看不上定量研究的。一方面，我们倾向于研究大问题，不喜欢纠缠于细枝末节。国内一位老师的话给我的印象很深，大致是说：如果你看到一堵墙就要倒了，还用得着纠缠于那堵墙的倾斜角度究竟是几度吗？所以，很多研究都是大而化之，只要说得通即可。另一方面，国内（十年前）的统计教学，总的来说与社会研究中的实际问题是相脱节的。结果是，很多原先对定量研究感兴趣的学生在学完统计之后，依旧无从下手，逐渐失去了对定量研究的兴趣。

我所就读的美国加州大学洛杉矶分校社会学系，在定量研究方面有着系统的博士训练课程。不论研究兴趣是定量还是定性的，所有的研究生第一年的头两个学期必须修两门中级统计课，最后一个学期的系列课程则是简单介绍线性回归以外的其他统计方法，是选修课。希望进一步学习定量研

究方法的可以在第二年修读另外一个三学期的系列课程,其中头两门课叫“调查数据分析”,第三门叫“研究设计”。除此以外,还有如“定类数据分析”、“人口学方法与技术”、“事件史分析”、“多层线性模型”等专门课程供学生选修。该学校的统计系、心理系、教育系、经济系也有一批蜚声国际的学者,提供不同的、更加专业化的课程供学生选修。2001年完成博士学业之后,我又受安德鲁·梅隆基金会资助,在世界定量社会科学研究的重镇密歇根大学从事两年的博士后研究,其间旁听谢宇教授为博士生讲授的统计课程,并参与该校社会研究院(Institute for Social Research)定量社会研究方法项目的一些讨论会,受益良多。

2003年,我赴港工作,在香港科技大学社会科学部,教授研究生的两门核心定量方法课程。香港科技大学社会科学部自创建以来,非常重视社会科学研究方法论的训练。我开设的第一门课“社会科学里的统计学”(Statistics for Social Science)为所有研究型硕士生和博士生的必修课,而第二门课“社会科学中的定量分析”为博士生的必修课(事实上,大部分硕士生修完第一门课后都会继续选修第二门课)。我在讲授这两门课的时候,根据社会科学研究生的数理基础比较薄弱的特点,尽量避免复杂的数学公式推导,而用具体的例子,结合语言和图形,帮助学生理解统计的基本概念和模型。课程的重点放在如何应用定量分析模型研究社会实际问题上,即社会研究者主要为定量统计方法的“消费者”而非“生产者”。作为“消费者”,学完这些课程后,我们一方面能够读懂、欣赏和评价别人在同行评议的刊物上发表的定量研究的文章;另一方面,也能在自己的研究中运用这些成熟的

方法论技术。

上述两门课的内容,尽管在线性回归模型的内容上有少量重复,但各有侧重。“社会科学里的统计学”(Statistics for Social Science)从介绍最基本的社会研究方法论和统计学原理开始,到多元线性回归模型结束,内容涵盖了描述性统计的基本方法、统计推论的原理、假设检验、列联表分析、方差和协方差分析、简单线性回归模型、多元线性回归模型,以及线性回归模型的假设和模型诊断。“社会科学中的定量分析”则介绍在经典线性回归模型的假设不成立的情况下的一些模型和方法,将重点放在因变量为定类数据的分析模型上,包括两分类的 logistic 回归模型、多分类 logistic 回归模型、定序 logistic 回归模型、条件 logistic 回归模型、多维列联表的对数线性和对数乘积模型、有关删节数据的模型、纵贯数据的分析模型,包括追踪研究和事件史的分析方法。这些模型在社会科学研究中有着更加广泛的应用。

修读过这些课程的香港科技大学的研究生,一直鼓励和支持我将两门课的讲稿结集出版,并帮助我将原来的英文课程讲稿译成了中文。但是,由于种种原因,这两本书拖了四年多还没有完成。世界著名的出版社 SAGE 的“定量社会科学研究”丛书闻名遐迩,每本书都写得通俗易懂。中山大学马骏教授向格致出版社何元龙社长推荐了这套书,当格致出版社向我提出从这套丛书中精选一批翻译,以给中文读者时,我非常支持这个想法,因为这从某种程度上弥补了我的教科书未能出版的遗憾。

翻译是一件吃力不讨好的事。不但要有对中英文两种

语言的精准把握能力,还要有对实质内容有较深的理解能力,而这套丛书涵盖的又恰恰是社会科学中技术性非常强的内容,只有语言能力是远远不能胜任的。在短短的一年时间里,我们组织了来自中国内地及港台地区的二十几位研究生参与了这项工程,他们目前大部分是香港科技大学的硕士和博士研究生,受过严格的社会科学统计方法的训练,也有来自美国等地对定量研究感兴趣的博士研究生。他们是:

香港科技大学社会科学部博士研究生蒋勤、李骏、盛智明、叶华、张卓妮、郑冰岛,硕士研究生贺光烨、李兰、林毓玲、肖东亮、辛济云、於嘉、余珊珊,应用社会经济研究中心研究员李俊秀;香港大学教育学院博士研究生洪岩璧;北京大学社会学系博士研究生李丁、赵亮员;中国人民大学人口学系讲师巫锡炜;中国台湾“中央”研究院社会学所助理研究员林宗弘;南京师范大学心理学系副教授陈陈;美国北卡罗来纳大学教堂山分校社会学系博士候选人姜念涛;美国加州大学洛杉矶分校社会学系博士研究生宋曦。

关于每一位译者的学术背景,书中相关部分都有简单的介绍。尽管每本书因本身内容和译者的行文风格有所差异,校对也未免挂一漏万,术语的标准译法方面还有很大的改进空间,但所有的参与者都做了最大的努力,在繁忙的学习和研究之余,在不到一年的时间内,完成了三十五本书、超过百万字的翻译任务。李骏、叶华、张卓妮、贺光烨、宋曦、於嘉、郑冰岛和林宗弘除了承担自己的翻译任务之外,还在初稿校对方面付出了大量的劳动。香港科技大学霍英东南沙研究院的工作人员曾东林,协助我通读了全稿,在此

我也致以诚挚的谢意。有些作者,如香港科技大学黄善国教授、美国约翰·霍普金斯大学郝令昕教授,也参与了审校工作。

我们希望本丛书的出版,能为建设国内社会科学定量研究的扎实学风作出一点贡献。

吴晓刚

于香港九龙清水湾

序

回顾过去 20 多年的社会科学研究,我们提出了一个始终如一的发现:婚姻会增加幸福感。Stack 与 Eshleman 在一篇讨论这一问题的文献中总结道:“已婚者相对于未婚者在幸福感方面的一个明确的指标——总体幸福感上的优势似乎是真实的。”(1998:527)通过研究 17 个国家的数据,他们得出“对男性与女性,婚姻同等地增加幸福感”的结论(1998:527)。存在争论的是两个问题——婚姻是否能增加幸福感和两性之间的增加是否相同(或者如某些研究者认为的,男性比女性增加得多)。另外一个我们关心的问题是,无论两性不同还是相同,这一增加的幸福感是否在婚姻的历程中有所变化。通过分析一个全国范围的 17 年、5 期的样本, VanLangingham、Johnson 和 Amato(2001)发现,没有证据支持文献所建议的生命过程中婚姻幸福感的 U 形模式(幸福感的减少及在晚年的回升)。

对于上面提到的研究问题,有两种方法论的途径:使用横截面数据或者纵向数据比较结婚的人和没有结婚的人。很多关于这个问题的研究使用了横向数据。然而,要对“婚姻增加幸福感”这个简单的论述作出合适的评价,使用纵向

数据是必要的。一些研究者,比如 VanLaningham(2001)等人,确实使用了纵向数据(他们使用了固定效果模型)。我们真正想研究的是以下情形:在时点 0 上,有一个随机未婚人群的样本,这些人被追踪;在时点 1 的时候,有些人结了婚;在时点 2 的时候,又有一些人结了婚。但是有一些人在整个追踪的时期内都没有结婚,即使当他们完成自己生命历程的时候,也还是有人没有结婚。横向数据只是截取一个时点,然后比较那些结过婚的人和没有结过婚的人,而忽略了人与人之间不可观测的差异性。纵向分析模型(例如,固定效果模型、随机影响模型和混合影响模型)恰当地处理了不可观测的差异性问题。更合适的方法是潜变量增长曲线模型,这种方法不仅能够处理不可观测的差异性问题,也能够对男性和女性分别就他们生命历程中婚后的幸福感进行潜在趋势研究。

Preacher、Wichman、MacCallum 和 Briggs 的潜变量增长曲线模型为社会科学量化应用的研究者提供了一个学习一种新方法的研究,像上面提到的实体问题的独一无二的机会,并可以超越经验性的研究。他们显示了潜变量增长曲线模型作为结构方程模型是如何表达和估计的,并说明了它和多层次增长模型的关系。在本书中,作者显示了各种各样的模型,例如固定/随机截距和/或者固定/随机坡度、有条件增长曲线模型、平行过程模型、多组分析。他们也讨论了模型的各种延伸,例如分段增长、结构潜在增长、类型和有序变量结果等。他们用了大量说明性的图形来表述和讨论,这是使用结构方程模型的一个优势。

廖福挺

目 录

序	1
第 1 章 简介	1
第 1 节 本书概述	4
第 2 节 潜变量增长曲线模型:简短的历史与概述	6
第 3 节 模型设定和参数解释	9
第 4 节 时间的测量尺度	15
第 5 节 异步测量	19
第 6 节 前提假设	21
第 7 节 参数估计与缺失数据	23
第 8 节 模型评估与选择	27
第 9 节 统计效力	30
第 2 章 将潜变量增长曲线模型应用于经验数据	33
第 1 节 数据	34
第 2 节 软件	37
第 3 节 模型拟合策略的概要	38
第 4 节 模型 0:空模型	40
第 5 节 模型 1:随机截距模型	42
第 6 节 模型 2:固定截距、固定斜率	44

第 7 节	模型 3:随机截距、固定斜率	47
第 8 节	模型 4:随机截距、随机斜率	49
第 9 节	模型 5:多群体分析	53
第 10 节	模型 6:条件增长曲线模型	56
第 11 节	模型 7:并行过程模型	60
第 12 节	模型 8:队列序列设计	64
第 13 节	模型 9:时变协变量	70
第 14 节	模型 10:多项式增长曲线	75
第 15 节	模型 11:未设定轨迹	78
第 16 节	总结	81
第 3 章	模型的特别拓展	83
第 1 节	增长混合模型	85
第 2 节	分阶段增长	88
第 3 节	多指标潜变量的变化模型	92
第 4 节	结构潜变量曲线	95
第 5 节	自回归潜变量曲线模型	99
第 6 节	分类变量和次序变量模型	101
第 7 节	在变化中对因果关系建模	103
第 8 节	总结	106
第 4 章	潜变量增长模型和多层模型的关系	109
第 1 节	MLM 在重复测量数据中的应用	112
第 2 节	模型设置	114
第 3 节	参数估计	116
第 4 节	模型评估	117

第 5 节	MLM 和 LGM 重合的方面	118
第 6 节	MLM 和 LGM 的不同之处	121
第 7 节	软件	124
第 5 章	结 论	125
附录		129
注释		134
参考文献		137
译名对照表		154

第 **1** 章

简 介

社会和行为科学的一个根本观察就是人随时间而变化,但并不是以同样的方式或同样的速度。譬如,词汇能力在小学阶段会稳步地提高,但并不是所有的学生都会以同样的速度提高。在很多情况下,变化的速度和方向都存在个体差异,而这些变化中的个体差异往往具有科学上或实践上的意义。时间上的变化可以用秒(如在心脏反应研究中)或用十几年来测量(如在生命发展研究中)。例如,夫妻争执会在短短的几分钟内引起显著的内分泌变化,但这些变化对于丈夫和妻子而言,可能会十分不同(Kiecolt-Glaser et al., 1997)。通过分析青少年跟踪调查数据,Flora 和 Chassin(2005)检验了父母酗酒对青少年使用毒品增长的影响。早期研究变化的方法有很大的局限:(1)它们只关注群体层面或个体层面的增长;(2)它们只着眼于两个时间点上的测量,导致数据过于贫乏而不能检验一些最基本的关于变化的有意思的假设。科学家常常需要更广泛的数据和更高级的统计方法来区分和理解变化的形态和方向(轨迹),并确定变化的原因和后果。

纵向设计能够得到关于心理现象的趋势和个体变化差异的有价值信息。这类数据的丰富性随着数据收集次数的增加而增加。Willett(1989)以及 Willett 与 Sayer(1994)讨论

了相对于二期数据,多期纵向数据的几个优势:(1)研究结果的质量会得到提高;(2)根据心理学理论可以建立正确的增长函数;(3)可以检验系统中个体间成长差异的假设;(4)可以将成长的特性与背景特点联系起来;(5)成长测度的精确性与准确性随着数据期数的增加而快速提高。简言之,纵向数据不仅提高了假设检验的统计效力,而且使研究者可以检验他们用横截面或二期数据无法检验的假设。

虽然利用纵向数据的这些优势使许多方法发展起来,但本书所关注的是潜变量增长曲线模型(LGM)。潜变量增长曲线模型代表了一类广泛的统计方法,这些方法允许更好的假设表述,提供更好的统计效力,并且比起别的方法,这些方法使统计模型与理论更好地衔接在一起。潜变量增长曲线模型不仅可以直接检测个人在不同时点的变化,而且能够检查不同个体之间的差异。这个模型的好处不仅在于它能够建构变化模型,还在于它允许研究者探寻变化的前因后果。

潜变量增长曲线模型框架可以让科学家清楚地表达和检验如下的问题:(1)随着时间的推移,均值变化趋势的形态是什么样的?(2)最初的水平是否可以预测变化速度?(3)两组或多组分组是否在变化轨迹上有区别?(4)是否可以通过均值变化趋势的速度或弯曲度预测重要的结果?(5)哪些变量与随着时间的推移而产生的变化有系统的关联?(6)根据观测数据,关于变化轨迹的理论假设是否站得住脚?(7)变化轨迹的形态是否存在显著的个体间差异?(8)一个变量的变化是否与另一个变量的变化相关?

上述问题当然不是详尽无遗的。接下来,我们将对本书的其余部分进行概述。