



档案数据处理： 生活经历研究

格伦·H.埃尔德 (Glen H. Elder, Jr.)

[美] 伊莉莎·K.帕瓦尔科 (Eliza K. Pavalko) 著

伊丽莎白·C.柯里普 (Elizabeth C. Clipp)

洪岩璧 曾迪洋 译 武玲蔚 校



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

格致方法·定量研究系列 吴晓刚 主编

档案数据处理： 生活经历研究

格伦·H.埃尔德(Glen H. Elder, Jr.)

[美] 伊莉莎·K.帕瓦尔科(Eliza K. Pavalko) 著

伊丽莎白·C.柯里普(Elizabeth C. Clipp)

洪岩璧 曾迪洋 译 武玲蔚 校

SAGE Publications, Inc.

格致出版社 上海人民出版社

图书在版编目(CIP)数据

档案数据处理:生活经历研究/(美)格伦·H.埃尔德(Glen H. Elder, Jr.), (美)伊莉莎·K.帕瓦尔科(Eliza K. Pavalko), (美)伊丽莎白·C.柯里普(Elizabeth C. Clipp)著;洪岩璧,曾迪洋译. —上海:格致出版社:上海人民出版社,2017.3

(格致方法·定量研究系列)

ISBN 978-7-5432-2724-8

I. ①档… II. ①格… ②伊… ③伊… ④洪… ⑤曾… III. ①档案资料-数据处理 IV. ①G272

中国版本图书馆 CIP 数据核字(2017)第 029976 号

责任编辑 贺俊逸

格致方法·定量研究系列

档案数据处理:生活经历研究

格伦·H.埃尔德

[美]伊莉莎·K.帕瓦尔科 著

伊丽莎白·C.柯里普

洪岩璧 曾迪洋 译 武玲蔚 校

出版 世纪出版股份有限公司 格致出版社
世纪出版集团 上海人民出版社
(200001 上海福建中路 193 号 www.ewen.co)



编辑部热线 021-63914988
市场部热线 021-63914081
www.hibooks.cn

发行 上海世纪出版股份有限公司发行中心

印刷 上海商务联西印刷有限公司
开本 920×1168 1/32
印张 5.5
字数 91,000
版次 2017 年 3 月第 1 版
印次 2017 年 3 月第 1 次印刷

ISBN 978-7-5432-2724-8/C·170

定价:32.00 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,翻译成中文,起初集结成八册,于 2011 年出版。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热烈欢迎。为了给广大读者提供更多的方便和选择,该丛书经过修订和校正,于 2012 年以单行本的形式再次出版发行,共 37 本。我们衷心感谢广大读者的支持和建议。

随着与 SAGE 出版社合作的进一步深化,我们又从丛书中精选了三十多个品种,译成中文,以飨读者。丛书新增品种涵盖了更多的定量研究方法。我们希望本丛书单行本的继续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

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,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”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,这两本书拖了多年还没有完成。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,与我的教学理念是相通的。当格致出版社向我提出从这套丛书中精选一批翻译,以飨中文读者时,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香港、台湾地区的二十几位

研究生参与了这项工程,他们当时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;哈佛大学社会学系博士研究生郭茂灿和周韵。

参与这项工作的许多译者目前都已经毕业,大多成为中国内地以及香港、台湾等地区高校和研究机构定量社会科学方法教学和研究的骨干。不少译者反映,翻译工作本身也是他们学习相关定量方法的有效途径。鉴于此,当格致出版社和 SAGE 出版社决定在“格致方法·定量研究系列”丛书中推出另外一批新品种时,香港科技大学社会科学部的研究生仍然是主要力量。特别值得一提的是,香港科技大学应用社会经济研究中心与上海大学社会学院自 2012 年夏季开始,在上海(夏季)和广州南沙(冬季)联合举办《应用社会科学研究方法研修班》,至今已经成功举办三届。研修课程设计体现“化整为零、循序渐进、中文教学、学以致用”的方针,吸引了一大批有志于从事定量社会科学研究博士生和青年学者。他们中的不少人也参与了翻译和校对的工作。他们在

繁忙的学习和研究之余,历经近两年的时间,完成了三十多本新书的翻译任务,使得“格致方法·定量研究系列”丛书更加丰富和完善。他们是:东南大学社会学系副教授洪岩璧,香港科技大学社会科学部博士研究生贺光烨、李忠路、王佳、王彦蓉、许多多,硕士研究生范新光、缪佳、武玲蔚、臧晓露、曾东林,原硕士研究生李兰,密歇根大学社会学系博士研究生王骁,纽约大学社会学系博士研究生温芳琪,牛津大学社会学系研究生周穆之,上海大学社会学院博士研究生陈伟等。

陈伟、范新光、贺光烨、洪岩璧、李忠路、缪佳、王佳、武玲蔚、许多多、曾东林、周穆之,以及香港科技大学社会科学部硕士研究生陈佳莹,上海大学社会学院硕士研究生梁海祥还协助主编做了大量的审校工作。格致出版社编辑高璇不遗余力地推动本丛书的继续出版,并且在这个过程中表现出极大的耐心和高度的专业精神。对他们付出的劳动,我在此致以诚挚的谢意。当然,每本书因本身内容和译者的行文风格有所差异,校对未免挂一漏万,术语的标准译法方面还有很大的改进空间。我们欢迎广大读者提出建设性的批评和建议,以便再版时修订。

我们希望本丛书的持续出版,能为进一步提升国内社会科学定量教学和研究水平作出一点贡献。

吴晓刚

于香港九龙清水湾

序

在 20 世纪 60 年代,一批主要的纵贯研究(longitudinal studies)开始开展调查,例如密歇根收入动态追踪调查(Michigan Panel Study of Income Dynamics)、天才工程(Project Talent)和全国纵贯调查(National Longitudinal Surveys)。以第一项调查为例,密歇根追踪调查包括了 6 000 多户家庭。这项研究得出了一个出乎意料的发现:仅对极少数人来说,贫困才是一种持续状况(即持续一年以上)。这一发现打破了流行的说法,而只有纵贯研究才可能有这样的发现。自 1968 年开始,该研究致力于对同样的家庭进行回访。但是,即使像密歇根追踪调查这样弥足珍贵的调查,它也不是为研究生命历程(life course)或研究重大历史事件对生命历程的影响而设计的。因此,如果研究者想要用这些数据来进行生活经历研究,那么就不得不对这些数据进行重构,如果做不到这一点,就只能去找其他档案数

据了。

这就是埃尔德教授和他的同事在“对旧数据提出新问题”时所做的事情。他们对在其他纵贯研究中为别的目的而搜集的原始资料进行考察,把这些资料进行重构,产生适合于分析个人生活的观测数据。在阐述这一方法时,他们借用了知名的路易斯·特曼(Lewis Terman)研究。该调查自1921年开始,对一个由天才青少年所构成的样本(857名男孩和671名女孩)进行了10年的追踪调查。从那时开始,后来又搜集了12轮访谈数据,最近一次是在1991年。除访谈之外,特曼研究中的档案数据还提供了大量的补充信息,包括新闻故事、配偶和父母的访谈以及信件。本书作者描述了他们为重构这些档案所做的每一步:评估现存的资料、决定如何编码、修改研究问题、评估编码主题、编写编码本、对个案进行编码、评价信度和效度、进行数据分析。在建构新的健康编码(第3章)、职业模式(第4章)和测量第二次世界大战的影响(第5章)中,他们对上述过程进行具体说明。

集中关注特曼研究很有用,因为它能显示本方法的长处。当然,这一方法也能同样适用于其他纵贯数据。如果读者脑海中想不起来这样的数据,他/她或许可以从作者所提及的几个档案数据来源中获得一些帮助。而且,正如作者在附录中所提到的,现在大约有200个纵贯数据集可

以被使用,它们被存放在亨利·A.莫瑞(Henry A. Murray)研究中心的国家资料库里。因此本书所展现的丰富的分析方法可以用于诸多其他研究主题。

迈克尔·S.刘易斯-贝克

致 谢

一些著作把它们的出版归因于其作者一生中与他人
的交往和专业经历。本书就是这样一本著作。它展现了
我们如何使用档案数据来调查变迁社会中关于生命历程
的问题。从非正式的理解到本书所描述的逐步递进方法
(stepwise method),这一研究路径已经发展了 30 多年。美
国国立精神卫生研究院(The National Institute of Mental
Health)自 20 世纪 60 年代末开始到现在,就一直慷慨解囊
资助这一类的研究。

最初的研究步骤是由第一作者在加州大学伯克利分校
的人类发展研究所(The Institute of Human Development)
工作期间,进行一个关于纵贯数据档案项目的过程发展而
来的。我们要感激支持这些研究的历任所长,也要感激进
行“奥克兰成长研究”(Oakland Growth Study)的哈罗德·
E.琼斯(Harold E. Jones)和赫伯特·斯托尔兹(Herbert

Stolz),以及玛丽·琼斯(Mary Jones)和约翰·A.克劳森(John A.Clausen)的英明领导;还有进行伯克利辅导研究(Berkeley Guidance Study)中的琼·麦克法兰(Jean MacFarlane)和玛乔丽·杭齐克(Marjorie Honzik)。这些研究所搜集的丰富数据为后来者提供了肥沃的土壤,使他们可以用不同的方法对这些旧数据提出新问题。人类发展研究所为这些努力提供了极大的支持和鼓励,我们非常感谢这些年来一贯支持这类研究的所长,他们是约翰·克劳森,保罗·马森(Paul Mussen),M.布鲁斯特·史密斯(M. Brewster Smith),盖伊·斯旺森(Guy Swanson)和乔·坎波斯(Joe Campos)。

1983年的寒冬,第一作者在康奈尔大学发现路易斯·特曼纵贯调查可用于研究战争给男性生活带来的影响。我们对这一可能性进行了深入挖掘,并于20世纪80年代中期开启了一项关于第二次世界大战服役经历对成人发展及老龄化影响的研究。罗伯特·希尔斯(Robert Sears)是当时特曼研究的主任,他投入了很多资源用于数据编码,并对我们所提出的研究表示了特别的兴趣和支持。但他(甚至我们自己)可能不知道以后我们会有多少次要回去利用数据档案进行工作。如果没有希尔斯、档案学家埃莉诺·沃克(Eleanor Walker)和现在的主任阿尔·哈斯托夫(Al Hastorf)一如既往的支持,我们不可能完成了新的研究目的而重构数据这一艰难的任务。美国国立精神

卫生研究院的一项基金和老兵管理部的一项能力回顾基金(Merit Review Grant)使我们能够开展这一冒险性的研究。对于所有这些支持,我们都深怀感激。

我们之所以能够成功对特曼调查中关于男性的数据进行重构,实现我们的目的,主要应该归功于程序员和项目经理凯瑟琳·科若斯(Catherine Cross)卓越的领导,她为我们带来了高水平的专业编码技术。一起工作时,在辛苦地准备编码本和培训编码员的工作中,她起了关键性作用。在对战时后方动员进行编码过程中,安德鲁·沃克曼(Andrew Walkman)对凯瑟琳进行了协助,安德鲁是北卡大学教堂山分校历史系的一名高年级研究生。我们非常幸运地拥有一些富有思想的、工作精确的编码员,他们是Karen Catoe, Marty Williams Deane, Laura Kline 和 Michael Riddle。乔伊斯·泰伯(Joyce Tabor)技术娴熟地负责对位于卡莱罗那人口中心的编码组进行指导。我们非常感谢他们,以及林恩·艾戈(Lynn Igoe)的编辑工作。他们一起使这项研究进行得井井有条,并且非常成功,比我们所希望的做得好多了。作为感谢的一部分,我们把本书献给在本研究的所有重要的编码工作中发挥了作用的每一个人。

莫瑞研究中心的主任安妮·科尔比(Anne Colby)提供了机会,让我们在工作坊上报告了我们的方法。我们感谢科尔比博士和她出色的团队所提供的这些教学机会。

我们希望本书有助于加强莫瑞研究中心推动在个人生活研究中更有效的使用纵贯数据的努力。很多同事通过他们仔细的阅读和反馈,帮助我们不断完善本书,他们是科尔比博士、汤姆·库克(Tom Cook)和珍妮·布鲁克斯-冈恩(Jeanne Brooks-Gunn)。我们再次感谢所有人,尤其是罗伯特·默顿(Robert Merton)无与伦比的编辑建议。当老师、使用者和学生运用这些方法来回答他们的研究问题时,对本书的最终检验就将到来。我们希望收到您的意见和建议。

格伦·H.埃尔德

伊莉莎·K.帕瓦尔科

伊丽莎白·C.柯里普

目 录

序	1
---	---

致谢	1
----	---

第 1 章 导论	1
----------	---

第 2 章 处理档案数据	11
--------------	----

第 1 节 生命历程革命	15
--------------	----

第 2 节 向旧数据提出新问题	21
-----------------	----

第 3 节 特曼研究	26
------------	----

第 4 节 从旧数据到新问题:大萧条时期的孩子	30
-------------------------	----

第 5 节 从生命记录到编码:伯克利指南档案	37
------------------------	----

第 6 节 契合研究问题与数据	41
-----------------	----

第 7 节 结论	44
----------	----

第 3 章 重构档案	47
------------	----

第 1 节 第一步	52
-----------	----

第 2 节 创造新编码	66
-------------	----

第 3 节 小结	85
----------	----

第4章	从退休到晚年工作经历	87
第1节	第一步	89
第2节	创造新编码	98
第5章	测量历史影响	111
第1节	联结历史影响与个人生活	115
第2节	为研究问题开发的新编码	125
第3节	结论	139
附录	雷德克里夫学院的亨利·莫瑞研究中心	141
参考文献		143
译名对照表		149