格致方法·定量研究系列　吴晓刚　主编

模糊集合理论在社会科学中的应用

[澳] 麦可·史密生（Michael Smithson）
[美] 杰·弗桂能（Jay Verkuilen）　著

林宗弘　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

35

格致出版社　上海人民出版社

格致方法·定量研究系列　吴晓刚　主编

模糊集合理论在社会科学中的应用

[澳] 麦可·史密生 (Michael Smithson)
[美] 杰·弗桂能 (Jay Verkuilen)　著
林宗弘　译

安徽大学图书馆
藏书
www.lib.ahu.edu.cn

SAGE Publications ,Inc.

格致出版社　上海人民出版社

图书在版编目(CIP)数据

模糊集合理论在社会科学中的应用/(澳)史密生(Smithson, M.),(美)弗桂能(Verkuilen, J.)著;林宗弘译.—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)

ISBN 978-7-5432-2197-0

Ⅰ.①模… Ⅱ.①史… ②弗… ③林… Ⅲ.①模糊集理论-研究 Ⅳ.①O159

中国版本图书馆CIP数据核字(2012)第271647号

责任编辑 罗 康

格致方法·定量研究系列

模糊集合理论在社会科学中的应用

[澳]麦可·史密生 [美]杰·弗桂能 著

林宗弘 译

出 版 世纪出版集团 www.ewen.cc

格致出版社 www.hibooks.cn

上海人民出版社

(200001 上海福建中路193号24层)

格致出版

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

字 数 76,000

版 次 2012年12月第1版

印 次 2012年12月第1次印刷

ISBN 978-7-5432-2197-0/C·94

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

此为试读，需要完整PDF请访问：www.ertongbook.com

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，大部分硕士生在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

对人类来说，集合概念的得来相当容易。人们可能发现，集合的逻辑最早来自亚里士多德，他在他经典的《逻辑学》里认为，人与牛都是“动物”，并且把动物分为有足的、双足的、有翅的与水生的几种不同类别。直到 1874 年，这种逻辑观念才得到精确的数学表述。乔治·肯特尔(George Canter)，这位双亲来自俄罗斯的丹麦学者，发表了一篇有数学公式与严谨的集合概念的论文，标志着集合理论作为一个数学分支的诞生。

集合理论，或本书中称为“经典集合论”的学说，长期以来支配着数学教学，现在仍然是高中数学教材里的重要组成部分。根据这个理论，成分函数，或以 m_x 来表示任何属于 X 集合的数据，只有两种数值，即 0 或 1，而且其函数图形可以表示为 $m_x: X \rightarrow \{0, 1\}$。然而，这个对现实的简化被罗飞·查迪(Lotfi Zadeh)于 1965 年所发表那篇关于模糊集合的革命性论文永远改写了。模糊集合的函数图形可以定义为 $m_x: X \rightarrow [0, 1]$，这允许数值出现在整个单位区间内。尽管模糊集合的逻辑基础可以追溯到柏拉图，但是直到查迪论文

的出现，才奠定了这类理论研究的基础，该理论随后被应用在计算机科学、工程科学与其他基础科学，包括社会科学之中。时至今日，模糊逻辑的运用已经涉及如何推动子弹列车、洗衣机与摄影机的运行等各方面。

模糊集合理论对社会科学有何贡献呢？在社会科学研究里，模糊性是稀松平常的事情，模糊集合理论为我们提供了一种适当的方式去系统性、建设性地处理模糊性。史密生与弗桂能的模糊集合理论是该领域的及时雨。本书提供了模糊集合理论的介绍，它超越了我们熟悉的清晰二分逻辑领域，是带领读者前往应许之地的一趟智识之旅。

廖福挺

此为试读，需要完整PDF请访问：www.ertongbook.com

目 录

序 1

第 1 章 导论 1

第 2 章 模糊集合教学的总纲 7

第 1 节 集合理论 9

第 2 节 为何要研究模糊集合? 12

第 3 节 归属函数 14

第 4 节 模糊集合理论的运算 17

第 5 节 模糊数据与模糊变量 25

第 6 节 模糊集合的图解 27

第 3 章 测量成员归属 29

第 1 节 导论 30

第 2 节 建构归属函数的方法 32

第 3 节 模糊集合所需的测量特质 40

第 4 节 归属函数的测量特质 43

第 5 节　成员归属定义的不确定性估计　47

第 4 章　模糊集合的内在结构与特质　55

第 1 节　集势:模糊集合的总量　58

第 2 节　模糊集合的概率分布　63

第 3 节　定义与测量模糊性　66

第 5 章　模糊集合之间的简单关系　75

第 1 节　交集、并集与包含　76

第 2 节　侦测与评估模糊包含　81

第 3 节　包含的量化与建模　86

第 4 节　量化与类似成员规模　94

第 6 章　多变量模糊集合的关系　101

第 1 节　组成集合指数　103

第 2 节　多集合关系:并发率、共变异数与共同发生率　107

第 3 节　多重与部分并集与包含　120

第 7 章　总结　123

参考文献　127

译名对照表　135

第1章

导　论

社会科学研究者长期以来观察到，尽管人类习惯于把他们的世界分成几个领域或类型，但他们却常使用一些边界模糊或是成分渐变的分类。社会科学界所使用的概念也常常如此。本书介绍模糊集合理论，它是罗飞·查迪对经典集合理论的延伸，为处理允许部分成员（有时译为元素）归属或归属程度不一的分类方式，提供了一个数学框架。

在最初允许样本部分地属于一个集合的直觉显现之后，模糊集合理论提出了并集与交集等集合概念的一般化理论。因此，这个理论把分类概念带到了向量的领域。如果经济贫困与心理忧郁的程度被认为是重要的，那么模糊集合理论认为，我们讨论穷人与忧郁人群交集的程度有多高之类的议题仍有意义。我们把模糊集合理论加入社会科学的工具箱有5大理由：

第一，该理论能够系统化地处理模糊性。

第二，社会科学的许多创见同时具备分类与向量的性质，即便是分类概念，也有重大的程度之分。

第三，该理论可以超越条件分析的工具与一般线性模型，用集合理论的一般化操作来分析多变量关系。

第四，它有理论上的精确性，大部分理论常常使用逻辑或集合导向的词汇来表达，但多数以连续变量为主的统计模型则不是。

第五，模糊集合理论以精确的方式结合了集合导向的思路与连续变量。

在过去的40年间，模糊集合理论被引介以来，已经累积了许多合适的应用经验，这种情况建议我们适时出版一本介绍模糊集合的书籍。举例来说，在心理学界，出现了以模糊集合为基础的认知理论（例如，Oden & Massaro，1978）或记忆理论（Massaro，Weldon & Kitzis，1991），模糊集合被用来解决测量问题，并且提供了新颖的分析工具（Hesketh，Pryor，Gleitzman & Hesketh，1988；Parasuraman，Masalonis & Hancock，2000；Smithson，1987；Wallsten，Budescu，Rappoport，Zwick & Forsyth，1986；Zwick，Budescu & Wallsten，1988；Smithson & Oden，1999）。此外，在社会学界与政治学界，拉津呼吁引进模糊集合来处理他所谓“多样性导向”的研究，并且强化理论与数据分析的关联性（Ragin，2000）。

因此，本书旨在引导社会科学研究者熟悉模糊集合与方法工具以便运用它。第2章介绍模糊集合理论的基本概念，包括成员归属等级、集合的理论运算、模糊数值与模糊变量。第3章着重在模糊集合与调查方法里赋予成员归属等级以建立归属函数。第4章探讨模糊集合的单变量性质——也就是集合的量与势——所拥有的概率分布及模糊性。第5章发展了集合之间（交集、并集与相容）的双变量关系。最后，第6章引进多集合关系与概念，包括组合集合指针、条件

元素函数以及多重与部分交集与兼容。纵贯全书，我们从不同的社会科学学科里找出范例，并且尽量建立起模糊集合取向与传统数据分析技巧之间的关联性。

与其他模糊集合理论的教科书不同，本书强调将模糊集合的概念与相当直接的统计技巧（尤其是许多已经用在标准化统计软件中的技巧）结合起来。我们相信，这种结合是必要的。当模糊集合理论被引入社会科学后，多数研究者常常仅仅运用模糊集合引入成员归属等级与重合分类的想法，某些明显是归属等级的运用也许混合了原型与类似原型的测量，少数运用模糊交集与并集、模糊逻辑或其他模糊推论方法。本书把模糊集合理论作为主要的核心议题，并且为那些希望驾驭模糊集合概念，以作出统计推论并检验他们模型的研究者，提供明确的指引。

有关模糊集合的某些观点并不包括在这本书里，不是因为我们认为那些方向不重要，而是因为讨论它们会使本书的篇幅加倍。反之，我们将集中讨论那些对广大社会科学圈内的、曾听说过模糊集合理论但是不了解它的读者们而言，最有趣也最容易理解的观点。然而，我们在此列举一些没包括在本书之内，但是仍然与社会科学家们有关的两个领域里的研究文献。模糊逻辑是模糊集合理论的直接延伸，并且包含在许多模糊推论与控制系统的运用中，范围从简单静态数据结构到复杂动态的系统分析。贝多西（Bárdossy）与杜克斯坦（Duckstein）讨论了区域规划所使用的数据系统。塞斯（Seitz）与他的同事们使用动态推论系统来模拟外交决策与组织行为（Seitz，1994；Seitz，Hulin & Hanisch，2001）。泰伯对部分使用模糊理性的计算模型提供了基本简介（Taber，

此为试读，需要完整PDF请访问：www.ertongbook.com