



格致方法·定量研究系列 吴晓刚 主编

多元回归中的交互作用

[美] 詹姆斯·杰卡德 (James Jaccard) 著
罗伯特·图里西 (Robert Turrisi) 著
蒋勤 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

格致方法·定量研究系列 吴晓刚 主编

多元回归中的交互作用

[美] 詹姆斯·杰卡德(James Jaccard) 著
罗伯特·图里西(Robert Turrisi)
蒋 勤 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

多元回归中的交互作用/(美)杰卡德
(Jaccard, J.), (美)图里西(Turrisi, R.)著;蒋勤
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 978-7-5432-2105-5

I. ①多… II. ①杰… ②图… ③蒋… III. ①多元回
归分析-研究 IV. ①0212.1

中国版本图书馆 CIP 数据核字(2012)第 122824 号

责任编辑 顾悦

格致方法·定量研究系列

多元回归中的交互作用

[美]詹姆斯·杰卡德 罗伯特·图里西 著
蒋勤 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
字 数 94,000
版 次 2012 年 7 月第 1 版
印 次 2012 年 7 月第 1 次印刷
ISBN 978-7-5432-2105-5/C·60
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

经典回归模型的表达式很直接,以最常见的三变量多元回归为例, Y 是因变量, X_1 和 X_2 是自变量,可把方程 1 写成如下形式:

$$Y = a + b_1 X_1 + b_2 X_2 + e \quad [1]$$

这里的 a 、 b_1 和 b_2 都是待估计的总体参数, e 是残差。假设满足必要的假定,那么普通最小二乘法(OLS)估计就是最佳的线性无偏估计。这个 OLS 方程是加叠性的,各项共同解释 Y ,没有自变量相乘的项。假设自变量相乘产生乘积项 $X_1 X_2$,那么就可改写上面的方程为方程 2:

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 (X_1 X_2) + e \quad [2]$$

现在,我们有了一个既包括加叠性变量,又包括非加叠变量的方程。而乘积项或乘法的项目($X_1 X_2$)也被称为“交互作用项”,它的系数 b_3 估计交互作用。问题是,它的含义是什么呢?

通常,社会科学专业一年级的研究生会觉得这个问题挺难回答。他们这么理解加叠模型(方程 1):当 X_2 保持不变时, X_1 变化一个单位,会导致 Y 变化 b_1 个单位。但如何理解 b_3 ,

即方程 2 中非加叠项的系数？当 X_1 对 Y 的影响取决于 X_2 的取值时，就存在交互作用。 b_3 估计的正是这种交互作用。请考虑这个例子： Y = 个人对竞选运动的捐献额(美元)， X_1 = 收入(美元)， X_2 = 教育程度(0 = 未上大学, 1 = 大学及以上)。如果研究者认为对受过大学教育的人而言，其收入对其捐献额影响更大(相比未受过大学教育者收入的作用)，那么研究者会倾向于使用包含交互作用的模型设定(方程 2)，而非方程 1。

很明显，交互作用假设——一个变量的影响取决于另一个变量的取值——应当需要检验。回顾已发表的研究，我们发现事实并非如此(很多研究不检验变量间的交互作用)，而我认为其原因在于交互作用这个想法较难捕捉，因此本书非常有意义。Jaccard 博士和 Turrissi 博士使理解和估计回归模型中的交互作用变得非常清楚。比如，他们强调，仅仅通过调节变量在不同组之间的作用差异来检验交互作用是不够的，如在上例中，仅仅分别观察上大学和没上大学两组人的收入在(分别的)回归方程中对政治捐献的影响是不够的。进而，他们提供了一个常常被忽略的见解：真正的度量来自数据本身，而不是测度。因此，要用回归检验交互作用，重要的是测度接近数据内在定距属性的程度。同时，本书也有许多新主题，比如聚类数据的交互作用以及随机系数模型。事实上，仔细阅读参考文献部分，我们能发现，1990 年之后出版的文献有 30 多种。

研究实践者必须利用工具去假定、理解和估计交互作用。仔细学习此书能使研究者在回归分析中达到上述要求。若要在其他或更高级的研究情境下研究交互作用，读者可参考 Jaccard 教授的其他著作。

Interaction Effects in Multiple Regression (2nd edition)

Copyright © 2003 by SAGE Publications, Inc.

All rights reserved. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

This simplified Chinese edition for the People's Republic of China is published by arrangement with SAGE Publications, Inc. © SAGE Publications, Inc. & TRUTH & WISDOM PRESS 2012.

本书版权归 SAGE Publications 所有。由 SAGE Publications 授权翻译出版。
上海市版权局著作权合同登记号：图字 09-2009-548

目 录

序	1
第 1 章 导论	1
第 1 节 交互的概念	4
第 2 节 简单效应与交互比较	7
第 3 节 多元回归分析的回顾	11
第 4 节 本书概览	22
第 2 章 双向交互作用	23
第 1 节 包含乘积项的回归模型	25
第 2 节 两个连续性解释变量	28
第 3 节 一个定性解释变量与一个连续解释变量	46
第 4 节 小结	58
第 3 章 三向交互作用	59
第 1 节 连续解释变量	61
第 2 节 定性和连续解释变量	68

第3节 小结	79
第4章 其他重要问题	81
第1节 连续变量之交互作用的双线性本质	83
第2节 计算关键自变量在调节变量取不同值时的系数	86
第3节 计算偏组成项	89
第4节 变换	91
第5节 多重交互作用	92
第6节 标准化与非标准化系数	95
第7节 度量性质	97
第8节 测量误差	101
第9节 稳健分析与假定的违反	104
第10节 被试者内和重复测量设计	106
第11节 同序和非同序交互	109
第12节 显著性的区间	113
第13节 混合交互	115
第14节 最佳实验设计和统计解释力	117
第15节 协变量	119
第16节 控制实验误差	120
第17节 整体检验和交互作用	122
第18节 常见的误用	123
第19节 聚类数据和随机系数模型的交互作用	124

第 20 节 连续还是离散解释变量	126
第 21 节 调节框架的回顾	127
注释	128
参考文献	130
译名对照表	134

第 **1** 章

导 论

许多社会科学理论框架都关注因果模型。这些模型试图设定一个或多个自变量对一个或多个因变量或结果变量的影响。简单地说,因果模型存在六种可能关系,如图 1.1 所示。直接因果是一个变量 X 为另一个变量 Y 的直接原因。在理论系统语境下, X 就是 Y 的直接决定因素。间接因果是 X 对 Y 有因果作用,但 X 对 Y 产生的作用只能通过另一个变量 Z 来实现;虚假因果就是 X 和 Y 相关,但这种相关性完全是因为存在一个共同原因 Z ,即 X 和 Y 不存在真正的因果关系;互为因果是指 X 对 Y 有因果影响, Y 对 X 也有因果影响;待定因果是指 X 对 Y 有因果影响, Y 对 X 也有因果影响;待定因

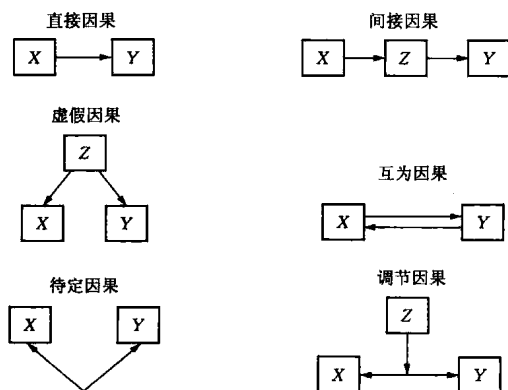


图 1.1 因果关系的类型