

“十二五”国家重点图书出版规划项目



应用统计工程前沿丛书

现代分层分位回归： 理论、方法与应用

田茂再 著



清华大学出版社

“十二五”国家重点图书出版规划项



应用统计工程前沿丛书

现代分层分位回归： 理论、方法与应用

田茂再 著



清华大学出版社
北 京

内 容 简 介

数据存在于特定的时间和空间中,其复杂的分层结构是一种普遍现象.充分借助于数据的这一特点,可以大大提高统计分析的有效性.本书致力于介绍复杂分层数据分析前沿知识,侧重于算法、仿真与实证研究.内容主要包括:分层线性模型、分层广义线性模型、分层非线性模型、分层半参数模型和分层分位回归模拟等.

本书可作为统计学及其相关领域的本科生、研究生的教材,亦可供教师和科技人员参考.

本书封面贴有清华大学出版社防伪标签,无标签者不得销售.

版权所有,侵权必究.侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

现代分层分位回归:理论、方法与应用/田茂再著.--北京:清华大学出版社,2015

(应用统计工程前沿丛书)

ISBN 978-7-302-39474-7

I. ①现… II. ①田… III. ①统计方法—研究 IV. ①C81

中国版本图书馆CIP数据核字(2015)第038030号

责任编辑:汪 操

封面设计:傅瑞学

责任校对:王淑云

责任印制:宋 林

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课 件 下 载: <http://www.tup.com.cn>, 010-62770175-4113

印 装 者:三河市金元印装有限公司

经 销:全国新华书店

开 本:170mm×230mm 印 张:23 字 数:435千字

版 次:2015年4月第1版 印 次:2015年4月第1次印刷

印 数:1~2000

定 价:59.00元

“应用统计工程前沿丛书”

编 委 会

顾问： 袁 卫 吴喜之 易丹辉 胡飞芳

主任： 赵彦云 金勇进

委员： 王晓军 张 波 孟生旺 许王莉 吕晓玲

蒋 妍 李静萍 王 星 肖宇谷

为中国的应用统计开拓奋进

(“应用统计工程前沿丛书”代序)

改革开放以来,我国统计事业取得了突飞猛进的发展.市场化、全球化和信息技术与网络经济的蓬勃发展,使统计在经济、社会、管理、医学、生物、农业、工程等领域中的应用迎来了又一春天.2011年2月,国务院学位委员会第28次会议通过了新的《学位授予和人才培养学科目录(2011)》,统计学上升为一级学科,这是国家对统计学科建设与发展的重大支持,它将全面推动统计学理论方法和应用研究的深入发展.天赐良机,我们应该为之而努力奋斗.



长期以来,我国统计学科主要在经济学门类、理学门类和医学门类下发展,未来进一步发展,一级统计学科将成为一面旗帜.世界各国实践充分表明,统计广泛应用在各个学科,在信息网络技术与计算机强大能力的推动下,统计学科发展特别是统计的应用正展示出一种前所未有的时代特征,它将为创造新的人类文明,提升人类发展能力做出新的重要的贡献.

新中国把中国从一盘散沙凝聚成高度集中的国家,推行计划经济发展模式.这一时期,统计为计划直接服务,为政府各级管理部门,为企事业单位的计划管理,为市场资源配置,为消费、投资的安排等提供全面系统的服务,因此表现出统计在经济社会管理上的重要作用.但是,由于权力至上的落后观念和体制机制呆滞,统计的科学性不被重视,统计数据搜集整理的简单化和主观操作酿成了很多不良的后果.改革开放之后,市场的作用强化了统计的社会影响和地位,但是,惯性的从上向下的主观思维方式仍然没有彻底的改观,因此,统计的科学应用仍然需要依靠内生发展的强大魅力不断深入和扩大.

近年来,全球化进一步加速了经济结构的转型与效率的提高.事实上,一国的稳步可持续发展离不开扎实的基础.在当今的信息化网络化时代,信息基础设施及其运用效率成为基础的基础,伴随而来的是统计在搜集数据、整理数据、数据分析上发挥的重要基础性作用.电子金融、电子政务、电子商务、网上购物、微博等一系列以网络信息技术为支撑的经济社会活动,创造了大数据的新时代,计算机科学、数据库技术、大数据统计分析成为新时代发展的耀眼之星,统计学理论方法在海量数据挖掘分析、高维分析和复杂系统模型分析,以及时空的统计图示图解分析等方面正显示出强劲发展的能量,应该讲现时期是统计应用最

好的发展机遇,它将大大提高人类发展的创造力、生产力,造福社会、造福人类。

二

在发展非凡的年代,谁能插上翅膀自由翱翔,谁能潜下海底自由鱼跃,统计学科需当仁不让,测度方位、穿透迷雾、指引方向、科学决策,助国家繁荣昌盛,立世界之林。这是当今中国人民大学统计学科建设的基本认知和理念。中国人民大学统计学科成立于1950年,已有60多年的发展历程,为共和国建设培养了大批优秀人才。他们广泛分布在政府部门以及银行、保险、证券、数据调查与咨询等商业企业,发挥了骨干作用。几代人大统计学人的辉煌历程和奉献,铸就了中国人民大学应用统计的特色,作为国家应用统计重点学科、教育部重点研究基地和国家统计局重点研究基地,在融入世界一流队列,开拓中国应用,培养高精尖应用统计人才,全方位支持国家建设和发展上,做出了重要的贡献。

今天,中国人民大学统计学科布局不仅深入经济社会发展领域和保险精算与金融风险管理领域,而且已经扩展到人文社会科学的许多领域,如法律、新闻、政治学、伦理学、教育学、心理学、文献计量等,展示出应用统计在量化人文社会科学研究中的重要作用。同时,我们也在生物、医学与公共健康领域开展了深入的统计交叉应用研究。建设扎实的概率论与数理统计基础,发展强大的应用统计是中国人民大学统计学院继往开来的基本目标。

三

为了系统总结和凝练中国人民大学在统计学各个领域的科研成果,引领和推动我国统计学学科建设,提高统计学在人文社会科学与自然科学各领域科学研究,以及在管理、决策支持等方面应用的科学化和普及水平,促进统计学及其交叉学科人才培养,我们组织编写了这套“应用统计工程前沿丛书”。丛书选题覆盖应用统计学的主要分支领域,如人文、社会、政治、经济、金融、管理、法律、教育、生物、卫生、网络、数据挖掘等,力求在科学性、应用性、创新性、前沿性和可读性上形成特色。

丛书针对各领域的实际问题,着重统计学方法、模型的创新、设计 and 应用。在应用领域的具体统计问题研究上,积极发展统计应用流程科学,强调应用背景描述清晰,基础问题明确,发挥对微观数据、大量数据归纳探索与挖掘的统计方法作用,发展标准化的统计思维方法,创建应用领域的重要统计模型,深入解决问题,推动应用领域适应信息社会的高速发展。我们首次提出应用统计工程一词。工程是将自然科学原理应用到工农业生产部门中去而形成的各学科的总称。

“工程”是科学的某种应用,通过这一应用,使自然界的物质和能源的特性能够通过各种结构、机器、产品、系统和过程,以最短的时间和精而少的人力做出高效、可靠且对人类有用的东西.我们强调应用统计的工程性,也就是强调统计的实际应用价值、科学流程与先进的统计应用技术.

丛书要反映统计学科多个前沿领域的科研进展,反映信息化和网络化背景下在诸多统计学应用领域产生的新的统计学问题及其方法和模型的发展,以及在人文社会科学各个领域的开创性应用研究.丛书选题覆盖了应用统计学的各主要分支学科和主要新兴应用领域,系统总结和凝练应用统计的专门技术方法,引领和推动我国大数据中的统计科学方法及其应用,提高网络信息统计处理与网络经济活动与经营活动的统计科学分析能力,提高统计学在企业经营管理、市场营销、科学决策,以及全面提升综合竞争力方面的作用,提高统计学在宏观经济产业政策、货币政策、收入分配政策等重大政策制定与效果分析,以及全面提升我国国际竞争力和国家软实力方面的作用.

本套丛书主要面向统计学及其交叉学科领域的科研人员、研究生和高年级本科生,以及在实际工作中需要应用统计学理论与方法的各领域专业人士.丛书在理论方法与应用领域深入结合研究上,强调增加关键点的细节内容,突出以统计知识为核心的应用领域的统计知识体系建设.各书在内容上力求拥有清晰的逻辑结构;对方法、概念和统计问题的描述增加相关概念知识和应用背景及交叉学科知识运用的铺垫;同时给出相关参考文献或推荐阅读书目,以帮助有兴趣的读者进一步深入学习.奉献给相关专业的读者能读懂并能够学以致用用的应用统计,这是本丛书追求的重要目标之一.

赵彦云 吕晓玲

2014 年 12 月

前 言

数据存在于特定的时间和空间中,离开数据的这一“时空性”,常常会造成传统的统计分析方法效果不佳,甚至失效.数据客观存在的时空性结构很普遍,其中常见的就是分层结构(hierarchical structure),它具有下述意义:我们用变量来描述个体,而个体嵌套在更大单元里,形成金字塔形状.

1. 分层分位回归(hierarchical quantile regression)

自20世纪末叶以来,针对数据分层结构的分层模型(hierarchical models)不论是在理论研究方面还是在应用方面都获得了长足的发展.该模型实质上就是条件独立分层模型.Hobert(2000)给出了有关分层模型目前计算方面的问题及前景展望.该类模型的最大优点就在于它能借助于组间与组内扰动力量.不过,现有的分层模型的理论本质上说就是给定预测变量 X 的值 x 后,响应变量 Y 的条件均值 $m(x)$ 的理论.这些理论没有也不可能给出响应变量的条件分位函数的全面刻画.所以,考虑给定协变量的条件下响应变量的全面刻画问题、估计子空间的稳健问题等,就显得非常必要了.

另一方面,Koenker和Bassett(1978)首先提出了分位回归模型的概念.分位回归是一种统计方法,它旨在对条件分位函数进行统计推断.正如基于残差平方和最小化的经典线性回归方法能估计条件均值函数一样,分位回归方法也为我们提供了一种估计条件分位函数的机制.有关分位回归的优点可以粗略地概括如下:(1)给定一组预测变量之后,它能全面刻画响应变量的整个条件分布;(2)分位回归模型有线性规划代理(LP),这使得估计简便;(3)就像LAD这一特例一样,分位回归的目标函数是加权的绝对偏差和,所以它能给出一个稳健的位置测度,因此被估计的系数向量对响应变量的离群点(outliers)不敏感;(4)当误差项服从非正态的时候,分位回归估计量要比最小二乘估计量更为有效,等等.

最近分位回归取得了长足发展.下面是几个典型的例子:(1)在参数分位回归模型方面,Portnoy和Koenker(1997)讨论了线性规划中内点问题的最新进展;(2)在非参数分位回归模型方面,Yu和Jones(1998)提出了“双核”(double-kernel)法;(3)在半非参数分位回归模型方面,Koenker等(1992)给出了一种解决基于罚似然估计(the penalized likelihood estimation)的算法;(4)目前,分位回归有几个热门话题:时间序列中的分位回归、分位回归的拟合优度以及贝叶斯分位回归,等等;(5)随着实际生活的需要,文献中出现了一些结合分层模型与分位回归方法优点的文章.Tian和Chen(2006)提出了分层线性分位回归模型.在那篇文章中,他们提出了一个基于高斯迭代并利用了分位回归和分层模型的

优点的新方法. 在理论方面, 考虑了渐近性质. 对于 n 收敛和渐近正态性, 我们有一些简单的条件. Tian, Tang 和 Chan (2010) 拓展了已经存在的分层线性模型和局部线性分位回归模型, 并提出了分层半参数分位回归模型. 该模型第一层为非参数. 在非参数的假设下, 非参数函数的偏导向量, 通常在经济学中叫做边际效应, 它将作为第二层的响应变量. 为了研究协方差效应对响应变量完整条件分布的影响, 该文章考虑分位回归系数. 不像普通的分层均值, 固定效应假定是常数, 而是允许固定效应的分位数是协变量的函数, 所以该方法对于很多统计应用者是很有吸引力的.

2. 复合分位回归 (composite quantile regression)

在最近几年里, 为了在多元线性回归中同时选择变量和估计系数, 大家提出了很多方法. 著名的方法有 Nonnegative Garrote (Breiman, 1995), Lasso (Tibshirani, 1996) 和 SCAD (Fan 和 Li, 2001). Fan 和 Li (2006) 综合性地描述了变量选择方面最新的优秀方法. 以上这些方法都是在方差有限的假设下进行的. 如果知道了误差项的分布, 那么最好的估计是已知真实的基本稀疏模型下的极大似然估计. 在线性回归问题中, 当我们遇到误差项分布 (似然函数) 未知的情形, 就得考虑实际最优过程. 我们注意到 LS-ORACLE 模型选择理论不需要误差项正态以及误差项分布具有有限方差的假设. 无论如何, 有限方差的假设对于基于最小二乘的最优模型选择的理论是很重要的. 另一方面, 模型的选择在于发现响应和预测之间关系的稀疏结构, 因此, 即便误差项方差无限, 模型选择也仍然是一个合法而又有趣的问题.

LS-ORACLE 中的局限性可见一斑. 人们希望找到一个可以克服 LS-ORACLE 问题的可选择最优 ORACLE. Zou 和 Yuan (2008) 介绍了一个新的回归模型, 叫做复合分位回归 (CQR), 从这种分位回归方法中可以找到满足前面所述性质的最优估计量. 他们定义了 CQR, 得到了 CQR-ORACLE 和 LS-ORACLE 相比的渐近最优性质和一个广泛使用的下界, 该下界说明相关效率大于 70%. Kai, Li 和 Zou (2009) 在复合分位回归的基础之上考虑了“复合分位回归平滑”方法. 该方法进一步改进了局部多项式回归. 在估计非参数回归函数和其导函数时局部的 CQR 估计具有很多优点, 例如: 对于非参数回归函数的局部线性 CQR 估计, 其渐近理论性质表明, 与传统的局部最小二乘估计相比, 这种新的估计在各种常见的非正态误差假设下其估计效率有显著的提高; 对于回归函数的导函数, 局部二次 CQR 估计的渐近理论结果显示该估计在误差非正态时能极大地提高与局部最小二乘估计相对应的各种估计效率, 同时, 在一些尤为不利的场合其损失的效率也最多只有 8.01%; 另外, 局部 p 次多项式 CQR 估计的一般渐近理论结果不需要假设误差具有有限方差, 因此, 局部 CQR 估计能在噪声方差为无穷大

时比局部多项式估计表现得更好. 局部线性 (多项式) 回归函数被证明是效率最优的线性光滑器 (Fan 和 Gijbels, 1996), 局部 CQR 估计的渐近理论结果与该结论并不矛盾, 因为它实际上是一种非线性光滑器.

3. 局部加权适应性分位回归 (locally weighted adaptive quantile Regression)

文献中一种非常流行的估计方法是基于局部常数拟合的思想. 局部平滑步骤在许多文章中已经研究过, 例如, 文献 (Tsybakov, 1986), (Truong, 1989), (Hall 和 Jones, 1990), (Chaudhuri, 1991), (Fan, 等, 1994), (Yu 和 Jones, 1998), (Tian 和 Chen, 2006) 以及 (Tian, 2009). 与常数窗宽相比, 变化的核估计量可以进行局部不同水平的平滑. 但是, 现有文章中的方法在计算均方误差 (MSE) 中的偏差部分时经常需要用到待估函数的高阶导数, 而这显然是比原估计更为困难的问题, 所以更加复杂. 再则, 现有的条件分位回归建模方法都是基于待估函数的平滑假设条件下的, 而这些假设在不连续点的邻域和陡峭的边界都是不能满足的, 所以经常会导致过度平滑问题和边界拟合失真问题, 这可参考文献 (Müller, 1992), (Wu 和 Chu, 1993), (Banerjee 和 Rosenfeld, 1993) 以及 (Speckman, 1994) 等. 对于均值回归 (与分位回归相比), 文献中已经提出了很多方法解决不连续性和陡峭边界问题. 例如, Polzehl 和 Spokoiny (2000, 2003) 针对附加偏误的局部多项式模型提出了一种自适应加权平滑过程. 局部加权适应性平滑方法是一种数据的适应性平滑的迭代技术, 它适合非连续的回归方程. 所用到的基本假设是这个回归函数可以由简单函数逼近, 如局部常数函数或局部线性函数. 另一方面, 局部边界问题在不同的均值回归方面已经有人研究过, 参见文献 (Korostelev 和 Tsybakov, 1993), (Scott, 1992), (Donoho, 1999), (Polzehl 和 Spokoiny, 2000) 及那里面相关的文献. 然而据我们了解, 类似的条件分位回归问题至今还没有人探讨过.

本书致力于介绍复杂分层数据分析前沿的知识, 侧重于算法、仿真与实证研究, 直接目的是给读者一些复杂分层数据的分位回归建模知识.

自 2004 年中国人民大学统计学院在全国首开“分层模型”课程以来, 本书作者一直担任本课程的主讲老师. 本书的大部分材料在课堂上讨论过. 本书在写作过程中, 自始至终有下面的硕士生、博士生参加过校正等工作: 李远、周朋朋、范洁瑜、张宁、戴成、钱政超、石恒泽、周健、安姝静、陈博钰、范博文、范燕、姜春波、马维华、苏宇楠、张圆圆、陈彦靓、郭洁、康雁飞、荣耀华、王伟、罗幼喜、储昭霖、封达道、李兆媛、司世景、夏文涛、熊巍、何静、胡亚南、黄雅丽、李茜、刘甦倩、吕爽、朱倩倩、田玉柱、梁晓琳、马春桃、马绰欣、孟令宾、王榛、杨亚琦、张亚丽、李二倩、罗静、史普欣、王晓荷、袁梦、吴延科、晏振等. 在此,

我对他们表示衷心的感谢!

本书得到下面基金的资助：国家自然科学基金 (No.11271368)、北京市社会科学基金重大项目 (No.15ZDA17)、教育部高等学校博士学科点专项科研基金 (No.20130004110007)、国家社会科学基金重点项目 (No.13AZD064)，以及兰州商学院“飞天学者特聘计划”。同时，感谢教育部人文社会科学重点研究基地中国人民大学应用统计研究中心和中国人民大学“985 工程”的大力支持。

由于本人水平有限，错误在所难免，甚望读者批评指正!

作者 田茂再

中国人民大学统计学院

2014 年 10 月

目 录

第一篇 分 层 模 拟

第 1 章 分层线性模型	3
1.1 概述	3
1.1.1 背景介绍	3
1.1.2 复杂数据界定	4
1.1.3 经典模型	5
1.1.4 主要参考文献	7
1.2 贝叶斯估计法	7
1.2.1 引言	7
1.2.2 例子	12
1.2.3 协方差结构未知时的估计	16
1.2.4 协方差结构未知的例子	17
1.2.5 多元回归方程间的可交换性	18
1.2.6 多元回归方程中的可交换性	21
1.2.7 主要参考文献	22
1.3 不完整数据的极大似然法	22
1.3.1 引言	22
1.3.2 EM 算法的定义	25
1.3.3 一般性质	29
1.3.4 例子	31
1.3.5 主要参考文献	44
1.4 EM 算法	45
1.4.1 介绍	45
1.4.2 协方差成分模型: 已知协方差情况下的理论	46
1.4.3 方差和协方差估计	49
1.4.4 例子	51
1.4.5 主要参考文献	62
1.5 迭代广义最小二乘法	62
1.5.1 引言	62
1.5.2 基本模型与符号	62
1.5.3 估计	65

1.5.4	随机系统	66
1.5.5	参数间的限制	67
1.5.6	未来的应用	68
1.5.7	纵向数据	68
1.5.8	测量误差	69
1.5.9	实例分析	70
1.5.10	主要参考文献	72
1.6	得分算法	72
1.6.1	引言	72
1.6.2	模型	73
1.6.3	对数似然函数	75
1.6.4	二水平嵌套	77
1.6.5	EM 算法	79
1.6.6	多于两水平嵌套	80
1.6.7	主要参考文献	83
1.7	Newton-Raphson 算法	83
1.7.1	引言	83
1.7.2	计算方法	84
1.7.3	Newton-Raphson 算法中对数似然的导数	85
1.7.4	用于 Newton-Raphson 算法的矩阵分解	89
1.7.5	通过 EM 算法估计 σ 与 D	92
1.7.6	例子	94
1.7.7	主要参考文献	96
第 2 章	分层广义线性模型	97
2.1	模型	97
2.1.1	介绍	97
2.1.2	分层广义线性模型	97
2.1.3	极大 h 似然估计的性质	105
2.1.4	估计过程	109
2.1.5	推广	113
2.1.6	分层广义线性模型分析举例	116
2.1.7	讨论	128
2.1.8	软件	129
2.1.9	主要参考文献	129
2.2	Gibbs 抽样方法	129

2.2.1	引言	129
2.2.2	随机效应广义线性模型	131
2.2.3	贝叶斯公式	132
2.2.4	Gibbs 抽样	132
2.2.5	条件分布	133
2.2.6	模拟与实例	136
2.2.7	主要参考文献	140
第 3 章	分层非线性模型	142
3.1	条件二阶广义估计方程	142
3.1.1	引言	142
3.1.2	模型	143
3.1.3	估计	144
3.1.4	条件方差 - 协方差的结构	146
3.1.5	惩罚尾似然和惩罚扩展最小二乘的关系	148
3.1.6	模拟	149
3.1.7	例子	155
3.1.8	主要参考文献	156
3.2	混合估计	157
3.2.1	引言	157
3.2.2	Lindstrom-Bates, Breslow-Clayton 和 Lee-Nelder 估计量 ...	158
3.2.3	混合估计	161
3.2.4	推广到分层广义线性模型	167
3.2.5	主要参考文献	170
第 4 章	分层半参数模型	171
4.1	分层半参数非线性模型	171
4.1.1	引言	171
4.1.2	半参数非线性混合效应模型	173
4.1.3	估计	176
4.1.4	计算	179
4.1.5	加拿大温度数据	182
4.1.6	模拟	184
4.1.7	主要参考文献	185
4.2	均值 - 协方差同时建模	186
4.2.1	背景	186

4.2.2 模型与估计方法	188
4.2.3 数值研究	191
4.2.4 主要参考文献	199

第二篇 分层分位回归模拟

第 5 章 分位回归引论	203
5.1 引言	203
5.1.1 分位数	203
5.1.2 分位回归	204
5.1.3 分位回归方法的演变	207
5.2 估计方法和算法	212
5.2.1 参数分位回归模型	212
5.2.2 Box-Cox 变换分位数模型	212
5.2.3 非参数分位回归模型	213
5.2.4 窗宽选择	216
5.2.5 半参数分位回归模型	217
5.2.6 两步法	217
5.3 分位回归应用领域	217
5.3.1 执行总裁 (CEO) 年报酬与公司股本的市场价值关系	218
5.3.2 分位数 Engel (恩格尔) 曲线	219
5.3.3 分位回归和婴儿体重的决定因素	220
5.3.4 医学中参考图表的应用	223
5.3.5 在生存分析方面的应用	223
5.3.6 风险值、分布尾部及分位数	224
5.3.7 经济	225
5.3.8 环境模型的应用	225
5.3.9 在检测异方差性上的应用	225
5.4 其他方面的进展	226
5.4.1 时间序列的分位回归	226
5.4.2 拟合优度	226
5.4.3 贝叶斯分位回归	228
5.5 软件 and 标准误差	228
5.6 主要参考文献	229

第 6 章 分层样条分位回归模拟	230
6.1 引言	230
6.2 条件分位函数的非参数估计	231
6.3 回归分位数模型的 Wald 检验	233
6.4 条件分位分层模型及其在家庭用电量需求上的应用	235
6.4.1 第一步: 家庭需求周期的时间序列模型	235
6.4.2 第二阶段: 需求周期的横截面模型	236
6.4.3 条件分位数分层模型	237
6.5 数据的描述	238
6.5.1 第一阶段结果	239
6.5.2 第二阶段结果	240
6.6 主要参考文献	245
第 7 章 分层线性分位回归模拟	247
7.1 引言	247
7.2 分层分位回归模型	248
7.3 EQ 算法	249
7.3.1 Q 步	249
7.3.2 E 步	249
7.3.3 迭代	250
7.3.4 初始值选取的基本方法	250
7.4 渐近性质	251
7.5 真实数据分析举例	252
7.5.1 数据描述	252
7.5.2 分位回归	253
7.5.3 两水平分层分位回归模型	254
7.5.4 部分结果	256
7.6 主要参考文献	258
第 8 章 分层半参数分位回归模拟	259
8.1 介绍	259
8.2 模型和估计	260
8.3 渐近结果	266
8.4 模拟分析	267
8.4.1 误差为多元柯西分布的层次线性模型	267
8.4.2 具有异方差的层次非参数分位回归模型	268

8.5	实际数据例子	269
8.6	主要参考文献	273
第 9 章	复合分层线性分位回归模拟	274
9.1	介绍	274
9.2	模型	275
9.3	估计	276
9.4	渐近性质	278
9.4.1	误差项为正态分布	278
9.4.2	误差项分布非正态	279
9.5	模拟	280
9.5.1	误差项正态	280
9.5.2	误差项为柯西分布	281
9.5.3	离群点	281
9.5.4	选择最优 K	283
9.6	实证部分	283
9.6.1	多水平模型中的数据分析	283
9.6.2	结果	285
9.7	主要参考文献	286
第 10 章	复合分层半参数分位回归模拟	287
10.1	介绍	287
10.2	模型	288
10.3	估计与算法	289
10.4	渐近性质	290
10.5	模拟研究	292
10.5.1	对于不同的误差项分布	292
10.5.2	对于 Y 存在异常值的情况	294
10.5.3	函数及其导数估计	294
10.6	实际数据分析	296
10.7	主要参考文献	299
参考文献		301
索引		347