

“十三五”国家重点出版物出版规划项目

经济科学译丛

空间计量经济学

空间数据的分位数回归

丹尼尔·P. 麦克米伦 (Daniel P. McMillen) 著

Quantile Regression for
Spatial Data

非
外
借

 中国人民大学出版社

“十三五”国家重点出版物出版规划项目

经济科学译丛

空间计量经济学

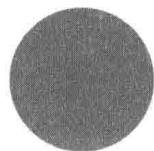
空间数据的分位数回归

丹尼尔·P. 麦克米伦 (Daniel P. McMillen) 著

闫吉丽 肖光恩 译

肖光恩 审校

Quantile Regression for
Spatial Data



中国人民大学出版社
· 北京 ·

图书在版编目 (CIP) 数据

空间计量经济学：空间数据的分位数回归/ () 丹尼尔·P. 麦克米伦著；闫吉丽，肖光恩译. —北京：中国人民大学出版社，2018.1

(经济科学译丛)

ISBN 978-7-300-23949-1

I. ①空… II. ①丹… ②闫… ③肖… III. ①区位经济学—计量经济学—研究
IV. ①F224.0

中国版本图书馆 CIP 数据核字 (2017) 第 013706 号

“十三五”国家重点出版物出版规划项目

经济科学译丛

空间计量经济学——空间数据的分位数回归

丹尼尔·P. 麦克米伦 著

闫吉丽 肖光恩 译

肖光恩 审校

Kongjian Jiliang Jingjixue

出版发行 中国人民大学出版社

社 址 北京中关村大街 31 号

邮政编码 100080

电 话 010-62511242 (总编室)

010-62511770 (质管部)

010-82501766 (邮购部)

010-62514148 (门市部)

010-62515195 (发行公司)

010-62515275 (盗版举报)

网 址 <http://www.crup.com.cn>

<http://www.ttrnet.com> (人大教研网)

经 销 新华书店

印 刷 北京溢漾印刷有限公司

规 格 185 mm×260 mm 16 开本

版 次 2018 年 1 月第 1 版

印 张 7.25 插页 1

印 次 2018 年 1 月第 1 次印刷

字 数 66 000

定 价 30.00 元

版权所有 侵权必究 印装差错 负责调换

《经济科学译丛》

编辑委员会

学术顾问

高鸿业 王传纶 胡代光 范家骧 朱绍文 吴易风

主 编

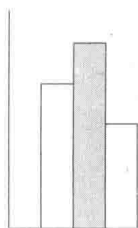
陈岱孙

副主编

梁 晶 海 闻

编 委 (按姓氏笔画排序)

王一江 王利民 王逸舟 贝多广 平新乔 白重恩
刘 伟 朱 玲 许成钢 张宇燕 张维迎 李 扬
李晓西 李稻葵 杨小凯 汪丁丁 易 纲 林毅夫
金 碚 姚开建 徐 宽 钱颖一 高培勇 梁小民
盛 洪 樊 纲



《经济科学译丛》总序

中国是一个文明古国，有着几千年的辉煌历史。近百年来，中国由盛而衰，一度成为世界上最贫穷、落后的国家之一。1949年中国共产党领导的革命，把中国从饥饿、贫困、被欺侮、被奴役的境地中解放出来。1978年以来的改革开放，使中国真正走上了通向繁荣富强的道路。

中国改革开放的目标是建立一个有效的社会主义市场经济体制，加速发展经济，提高人民生活水平。但是，要完成这一历史使命绝非易事，我们不仅需要从自己的实践中总结教训，也要从别人的实践中获取经验，还要用理论来指导我们的改革。市场经济虽然对我们这个共和国来说是全新的，但市场经济的运行在发达国家已有几百年的历史，市场经济的理论亦在不断发展完善，并形成了一个现代经济学理论体系。虽然许多经济学名著出自西方学者之手，研究的是西方国家的经济问题，但他们归纳出来的许多经济学理论反映的是人类社会的普遍行为，这些理论是全人类的共同财富。要想迅速稳定地改革和发

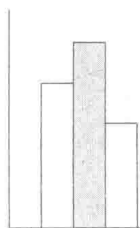
展我国的经济，我们必须学习和借鉴世界各国包括西方国家在内的先进经济学的理论与知识。

本着这一目的，我们组织翻译了这套经济学教科书系列。这套译丛的特点是：第一，全面系统。除了经济学、宏观经济学、微观经济学等基本原理之外，这套译丛还包括了产业组织理论、国际经济学、发展经济学、货币金融学、公共财政、劳动经济学、计量经济学等重要领域。第二，简明通俗。与经济学的经典名著不同，这套丛书都是国外大学通用的经济学教科书，大部分都已发行了几版或十几版。作者尽可能地用简明通俗的语言来阐述深奥的经济学原理，并附有案例与习题，对于初学者来说，更容易理解与掌握。

经济学是一门社会科学，许多基本原理的应用受各种不同的社会、政治或经济体制的影响，许多经济学理论是建立在一定的假设条件上的，假设条件不同，结论也就不一定成立。因此，正确理解掌握经济分析的方法而不是生搬硬套某些不同条件下产生的结论，才是我们学习当代经济学的正确方法。

本套译丛于 1995 年春由中国人民大学出版社发起筹备并成立了由许多经济学专家学者组织的编辑委员会。中国留美经济学会的许多学者参与了原著的推荐工作。中国人民大学出版社向所有原著的出版社购买了翻译版权。北京大学、中国人民大学、复旦大学以及中国社会科学院的许多专家教授参与了翻译工作。前任策划编辑梁晶女士为本套译丛的出版做出了重要贡献，在此表示衷心的感谢。在中国经济体制转轨的历史时期，我们把这套译丛献给读者，希望为中国经济的深入改革与发展作出贡献。

《经济科学译丛》编辑委员会



译者序言

向特定方法论方向扩展已经是当前空间计量经济学的发展的一个重要的新特征，而空间分位数回归正是空间计量经济学这种发展趋势的一种体现。我到美国路易斯安那州立大学奥尔斯商学院做访问学者后不久，我的指导老师空间统计与计量分析专家 R. Kelley Pace 教授就送给我一本《分位数回归：理论与应用》(*Quantile Regression: Theory and Application*) 让我学习，并指出这种分析方法的重要性及其应用与拓展的前景。尽管分位数回归已经是一种很成熟的分析方法，但是其应用的潜力仍未被充分地发掘，而其在空间领域的应用更是一种重要的发展。分位数回归的魅力可以让我们超越传统计量经济学分析中均值的“正宗”地位，并能欣赏到被解释变量以系列解释变量为条件的分布变化的奇妙；而空间分位数回归还可以让我们领略到空间自回归“溢出效应”的独特作用。

早在翻译《空间计量经济学：从横截面到空间面板》一书时，就已经有意翻译《空间计量经济学：空间数据的分位数回归》这本书。2015年6月到湖南大学经济与贸易学院参加全国高校国际贸易学科协作组首届青年论坛时认识了闫吉丽，向她介绍了这本书并达成了翻译的初步意向。正是受到 R. Kelley Pace 教授的影响，我才决定把这本书正式翻译出版。非常高兴的是中国人民大学出版社拥有卓识慧眼，很快就与外方出版社达成了知识产权的引进。同时，这本书也是我带领研究生团队学习空间计量经济学系列著作之后的第三本译作。本书的翻译也得到了武汉大学研究生院国际商务专业硕士精品课程《国际商务》、武汉大学经济与管理学院“计量经济学”平台课程等建设经费和“中央高校基本科研业务费专项资金”的资助，是武汉大学自主科研项目（人文社会科学）的研究成果。

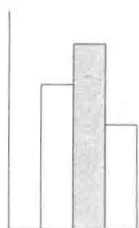
本书的翻译得到了美国路易斯安那州立大学奥尔斯商学院金融系 R. Kelley Pace 教授的鼓励，也得到了中国人民大学出版社王晗霞的大力支持，她工作高效，很快确定了翻译的相关事宜，使本书能及时出版。

由于水平有限，错误与疏漏之处在所难免。敬请读者批评指正，以便再版时修正。

肖光恩

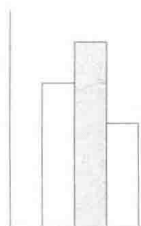
于美国路易斯安那州立大学奥尔斯商学院

2017年5月2日



致 谢

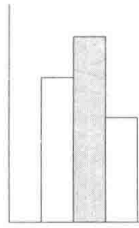
非常感谢 Roger Koenker, Paul Carrillo, Catia Nicodemo 和 Mark Partridge 提出的有帮助性的意见和建议。



摘要

分位数回归分析不同于大多数传统回归模型的地方就在于它更强调分布。标准回归方法一般是预测因变量的值如何随解释变量的变化而变化，而分位数回归则是预测因变量整个分布的变化。尽管分位数回归具有它的优势，但它在空间数据的分析中仍不常用。而本书的目的就是让研究人员更容易应用分位数回归方法处理空间数据集。本书的重点是对分位数回归分析的结果进行解释。书中运用了一系列基于模拟数据集和实际数据集拟合的例子来展示如何轻松地把看似复杂的分位数回归结果转化为结构良好的图表。与此同时，本书也详细考虑了参数空间模型和非参数空间模型。

关键词：分位数回归 空间计量经济学 非参数 局部加权回归

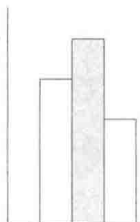


目 录

第 1 章 分位数回归：综述	1
1.1 对“旧城改造”的蒙特卡罗研究	6
1.2 分位数回归估计	9
1.3 销售价格的隐含分布	10
1.4 非线性分位数回归	14
1.5 本章总结	17
第 2 章 线性非参数分位数回归	19
2.1 基于模拟数据的线性分位数回归	21
2.2 模拟因变量的分布	27
2.3 解释变量离散的影响	31

2.4	非参数分位数回归	34
2.5	本章总结	41
第3章	累退性评估的分位数回归分析	43
3.1	评估率的蒙特卡罗分析	47
3.2	伊利诺伊州杜佩奇郡的评估率	50
3.3	本章总结	53
第4章	空间分位数自回归模型	55
4.1	含有内生解释变量的分位数回归	59
4.2	Hedonic 住房价格函数的应用	61
4.3	本章总结	70
第5章	条件参数分位数回归	73
5.1	空间数据的 CPAR 分位数回归	76
5.2	实证案例：华盛顿塔科马的房价	77
5.3	伊利诺伊州库克郡的评估率	83
5.4	本章总结	88
第6章	深入阅读指南	91
参考文献	96

第 1 章



分位数回归：综述



线性回归是大多数社会科学在进行实证研究时运用的标准工具。当因变量 y 和若干个解释变量 X 的关系可以表示为 $y = X\beta + u$ 时， y 对 X 的普通最小二乘（OLS）回归也许能提供参数 β 的无偏估计，以及给定 X 值下 y 的最好估计，即 $\hat{y} = X\hat{\beta}$ 。因此，当我们浏览任何社会科学杂志时都能够快速地发现回归分析作为实证分析工具的地位。

但是对线性回归模型的依赖性也被带到了空间数据的分析中。最常用的空间模型一般在解释变量中加入因变量的邻近观察值的加权平均值作为权重，如公式： $y = \rho WY + \beta X + \mu$ 。在这个模型中， W 就是表明观察值之间关系的“空间权重”矩阵。比如说一个城市的税率可能会受到其邻近城市税率的影响，如果这个城市有四个邻近城市并且每个邻近城市对这个城市的影响都相同，那么 WY 就是指这四个邻近城市税率的平均值。另外举一个例子，一栋房子的销售价格可能受其附近房屋销售价格的影响。在这两个例子中，回归分析的目的就是估计系数 ρ 和 β 的值，进而得到给定 X 值下 y 的预测值。

同时，回归分析也不太适合描述变量的分布。当回归的预测值是 $\hat{y} = X\hat{\beta}$ 时，预测值的分布仅仅是简单模拟变量 X 的分布。也就是说，一个解释变量的变化所暗含的效应就是导致预测值 \hat{y} 的平移，其平移的幅度由这个变量的估计系数决定。虽然在某些情况下平移是合理的，但是它的局限就在于研究者可能并不想提前施加这种约束。

对回归施加这种约束最好的一个例子就是“旧城改造”（gentrification）* 对房价的影响。对“旧城改造”的回归分析是为了寻求某地

* gentrification 指西方国家在城市化过程中，城市中心区更新（复兴）的一种新的社会空间现象，类似于我国大中城市内的“旧城改造”，也有人翻译成“城市绅士化”或“城市更新”。——译者注

区的周边街区一旦经过“旧城改造”后会导致该地区房屋价格上升多少。在确定了城市中哪些街区已经经过“旧城改造”后，研究者将“旧城改造”虚拟变量添加到一组住房特征的控制量中。“旧城改造”变量显著的正系数表明在其他条件相同的情况下，经过“旧城改造”的地区房价更高。这一发现可能会引起房价倡导者的关注，因为它意味着低收入居民可能会被迫搬出由于“旧城改造”而导致房屋价格上升的街区。

但是我们所得到的“旧城改造”街区的所有房屋价格上涨的结论可能是被大大简化后的结果。假设富人进入一个街区使得该地区价格位于前10%的房屋价格上升，而这些房价上升的房主直接迁移到别处。这样一来前10%的房屋价格上涨，而其余90%的房屋价格保持不变。然而，在对一组包含“旧城改造”指标在内的影响房屋价格的变量进行回归分析时，会得到经过“旧城改造”的街区房价预期较高的结果。这很容易导致研究人员得到错误的结论：“旧城改造”通过提高住房成本而损害了低收入家庭，然而事实上低价房屋的价格并没有改变。

而分位数回归更适合分析那些涉及因变量的分布变化的问题。粗略地讲——这个问题的重要性将会在本书的例子中得到证实——分位数回归可以分别估计各个自变量对因变量分布的不同点的单独影响。例如在“旧城改造”的例子中，对销售价格上涨的那部分房屋的分位数回归将揭示“旧城改造”对房屋出售价格具有显著影响，而对低价房屋的分位数回归将揭示“旧城改造”对房屋出售价格的影响很小或没有影响。

虽然分位数回归已很普遍，但是人们还没有意识到分位数回归估计的全部含义。研究人员通常应用各种分位数回归，10%、25%、50%、75%、90%分位数是常见的选择。尽管分位数回归可以对因变量分布中的不同点进行估计，但我们还是经常将其系数估计认为是类似于标准线性回归的一种估计。我们很少意识到分位数回归能够估计自变量值发生变化所导致的因变量整个分布的变化。而“旧城改造”变量的系数集就表明了当某地区的周边街区“旧城改造”后所导致的房屋价格整个分布的变化。根据结果做出图形后，我们可以看到销售价格分布中，高价部分向右移动而剩下的大部分则保持不变。

同时，利用分位数回归分析空间数据也可以避免出现一些特殊问题。例如，研究者们提出空间自回归（AR）模型的变体 $y = \rho WY + X\beta + \mu$ 用以进行回归分析。在这个过程中，他们仅把 WY 看作另外一个内生变量而已。然而空间自回归模型并不一定是空间建模的最好选择，特别是对由单个地区点而非更大地区或地块所构成的大数据集而言。当这种因变量的分布在空间中平稳变化时，反而可能更适宜应用非参数估计。

我写本书的目的是为了让研究人员更容易应用分位数回归对空间数据进行分析。书中的重点不是理论而是怎样对回归结果进行解释。一系列分位数回归代表着因变量服从什么分布？我在整本书中都强调，看似复杂的分位数回归结果可以轻松地用一组结构良好的图表来解释，就算其中运用了复杂的非参数估计量也是如此。