



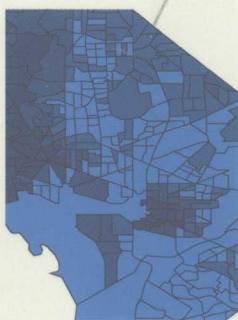
财经易文

ArcView GIS[®] 与ArcGIS[®] 地理信息统计分析

Statistical Analysis Of
Geographic Information
with ArcView GIS[®]
and ArcGIS[®]

[美] 戴维·W·S·黄 (David W.S. Wong) 著
杰·李 (Jay Lee)

张学良 译



中国财政经济出版社

ArcView GIS

与ArcGIS®

地理信息统计分析

Statistical Analysis Of Geographic Information
with ArcView GIS® and ArcGIS®

[美] 戴维·W·S·黄 (David W.S. Wong)

杰·李 (Jay Lee)

著

张学良 译

中国财政经济出版社

图书在版编目(CIP)数据

ArcView GIS® 与 ArcGIS® 地理信息统计分析/(美)黄,(美)李著;张学良译.

—北京:中国财政经济出版社,2008.2

书名原文:Statistical Analysis of Geographic Information with ArcView GIS® and ArcGIS®

ISBN 978-7-5095-0415-4

I. A… II. ①黄… ②李… ③张… III. 地理信息系统—应用软件, ArcView、ArcGIS IV. P208

中国版本图书馆 CIP 数据核字(2007)第 199994 号

著作权合同登记号:图字 01-2007-3227 号

David W. S. Wong, Jay Lee

Statistical Analysis of Geographic Information with ArcView GIS® and ArcGIS®

ISBN 978-0-471-46899-8

Copyright © 2005 by John Wiley & Sons, Inc. All Rights Reserved.

Authorized translation of the edition published by John Wiley & Sons. No part of this book may be reproduced in any form without the written permission of John Wiley & Sons, Inc.

Translation rights arranged through John Wiley & Sons, Inc.

本书中文简体字翻译版由美国 John Wiley & Sons 公司授权中国财政经济出版社在中华人民共和国境内独家出版发行。

未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

中国财政经济出版社 出版

URL: <http://www.cfeph.cn>

E-mail: webmaster@ewinbook.com

(版权所有 翻印必究)

社址:北京海淀区阜成路甲 28 号 邮政编码:100036

发行电话:010-88191017

三河市世纪兴源印刷有限公司印刷 各地新华书店经销

787×1092 毫米 16 开 26.25 印张 460 千字

2008 年 2 月第 1 版 2008 年 2 月北京第 1 次印刷

定价:58.00 元

ISBN 978-7-5095-0415-4/F·0345

(图书出现印装问题,本社负责调换)

使用指南

这是一本介绍基本的空间统计和空间分析知识以及如何运用 ArcView GIS 和 ArcGIS 软件进行相关分析的著作。在使用本书时请注意本书的一些特点：

1. 附带易于使用的空间分析和空间统计方面的工具和程序，并且与 GIS 软件包进行了整合。
2. 可让读者验证所讨论的空间分析和空间统计概念。
3. 提供了工具和数据，读者可以运用本书所介绍的统计分析方法进行实验。

读者可以到下列地址下载本书的软件包：<http://www.ewinbook.com/gis.rar>。下载问题请与我们联系，软件使用问题请与软件开发商联系。

电话：010-88191069

电子邮件：win@ewinbook.com



财经易文 学习的伙伴
www.ewinbook.com

谨以此书献给我们的妻子和女儿：

瑞吉娜·刘和麦丽迪·黄

林·李和瑞贝卡·李

没有他们的支持和在家庭生活上做出的巨大牺牲，

本书将无法完成。

译者序

这是一本介绍基本的空间统计和空间分析知识以及如何运用 ArcView GIS 和 ArcGIS 软件进行相关分析的著作。

长期以来,只考虑时间维度而不考虑空间维度一直是现代主流经济学的一个突出的特点。即便翻开当前最流行的经济学经典教材的最新版本,也很少能看到有关“空间”两字的只言片语。现代主流经济学思想体系与理论模型总是默认“空间”是“均质”和无差别的,因而无需加以考虑,这种回避或高度简化的处理方式,并不是因“疏忽”或“大意”造成的偶然事件,而更可能是现有研究体系中存在的“系统性偏差”。正如现代区域科学之父沃尔特·艾萨德所指出的,经济学分析是“在一个没有空间维度的空中楼阁中”进行的。

经济学研究对“空间”的忽略在很大程度上与现代经济学研究的模型化发展趋势紧密相关。现代经济学越来越成为一门模型的学问,其研究方法越来越模型化、工具化,“只有可以模型化的思想才会得到垂青”。“空间”长期未能真正被纳入主流经济学的研究框架,从方法论的角度讲,这一方面是由于空间经济学的先驱没有用模型化的方法来表达他们的思想;另一方面是因为,经济活动在空间中的聚集本质上是收益递增规律的表现,而收益递增比规模收益不变或收益递减的主流经济学思想更难模型化,因此随着经济学变得越来越严谨,对空间经济的研究就长期被推到学术的外围,具有许多科学思想和重要观点的空间经济学理论因不能满足现代经济学的模型化要求而被排除在主流经济学研究之外。

然而不可否认,与时间维度一样,对人类经济活动的分析也离不开空间维度。在现实生活中你会发现,经济活动不仅表现出时间维度上的联系,在空间维度上也表现出某种程度的相关,空间聚集、城市经济对国家经济与区域经济的发展越来越重要;在理论上,我们不仅要回答“生产什么”(What)、“怎样生产”(How)、“为谁生产”(Who)等问题,还要回答“在哪里生产”(Where)的问题,而这就涉及“空间”研究领域。因此,在现代经济学中真正引入“空间”因素,甚至将现代经济学从一门单纯的时间维度



的经济学发展成为包括时间与空间两个维度的“时空”经济学就非常有必要了。

值得庆幸的是,随着数理分析技术和计算机处理技术的不断进步,经济学家们开始尝试把“空间维度”纳入到现代经济学的分析框架中。近年来,空间经济学、新经济地理学与空间统计、空间计量经济学得到了很大的发展,与传统的区域经济学、城市经济学及经济地理学一样,它们将“空间”作为共同的研究范畴,保罗·克鲁格曼甚至认为“空间是经济学研究最后的前沿”,其本人也因为在“新经济地理”、“新贸易理论”等方面的突出贡献被普遍认为是可能获得诺贝尔经济学奖的热门人物之一。

近年来,之所以空间经济理论得到了快速的发展,空间经济学逐渐得到主流经济学的认同,部分原因可能是空间经济学在模型化与实证分析方面取得了很大的进步。在模型分析方面,以保罗·克鲁格曼为代表的新经济地理学者,在 Dixit-Stiglitz 垄断竞争模型的基础上,采用收益递增—不完全竞争的建模技巧,建立了以中心—外围模式为核心的区域模型、以城市层次体系演化为中心的城市模型以及以产业聚集和国际贸易为代表的国际模型,使得空间研究第一次从真正意义上被纳入主流经济学的研究范畴。1999 年藤田、克鲁格曼和维纳布尔斯合著的《空间经济学:城市、区域与国际贸易》是空间模型研究的集大成之作,以此为基础,经济学界兴起了空间经济学的研究浪潮,掀起了经济学研究中继产业组织理论、新贸易理论、新增长理论后的第四次革命——空间经济理论革命。

在实证分析方面,空间统计与经济计量学的发展,在拓宽了传统的计量经济学理论方法的同时,也为空间经济学提供了很好的实证分析工具。经典的计量经济学模型总是假定 Gauss-Markov、解释变量固定等条件,但是在区域经济分析中,空间依赖的存在打破了大多数经典统计和计量经济学分析中样本相互独立的基本假设,因此直接将经典统计和计量经济学的方法应用于与地理位置相关的数据时,通常会忽视数据的空间依赖性,进而引起各种问题。因此,在处理空间数据时,要适当引入空间统计和空间计量经济分析方法。当然空间统计与空间计量经济学也不是抛弃所有的经典统计与经典计量经济学技术,而是对这些技术加以修改,使它们适用于空间数据分析。另外,空间信息系统(GIS)的发展,Matlab、Stata、SAS 等软件中相关空间分析程序的开发,特别是 ARCGIS、Geoda 等空间分析专业软件的广泛运用,也为空间经济学的实证研究提供了条件。

可以说,空间分析理论与方法正方兴未艾。目前国内引进与出版的空间

经济学方面的专著与教材已非常丰富,空间经济学的相关理论也为国内相关学者所熟知,然而国内鲜有专门介绍空间统计与空间计量的书籍,结合具体案例进行分析的教材更是少见。虽然国外介绍空间统计与空间计量知识的书籍较多,但大多较为高深晦涩。而本书是一本十分通俗易懂的教科书,正如本书原作者在英文版序中所说,本书是目前唯一一本具有以下特点的著作:(1) 附带易于使用的空间分析和空间统计方面的工具和程序,并且与 GIS 软件包进行了整合;(2) 可让读者验证所讨论的空间分析和空间统计概念;(3) 提供了工具和数据,因此读者可以运用本书所介绍的统计分析方法进行实验,从而为自己日后在科研和实践活动中进行统计分析积累经验。这就使得本书中文版的出版更有意义。

当然,将本书读者对象完全定位在经济学领域有失偏颇。实际上,空间分析方法也为地理学者所熟知,地理信息系统相关专业人士对本书所介绍的 GIS 软件会比非专业人士更加熟悉。而就我理解,以“空间”作为研究范畴,地理学者与经济学者可以找到一些共同的研究主题,甚至可以进行一些交叉研究。我认为,本书体系完善,内容通俗易懂且操作性强,非常适合作为区域经济学、城市经济学、统计学、经济地理学与地理信息系统等相关专业的教科书,学生可以在一个学期的课程里学到相关的基础知识,还可以动手进行一些独立的研究。本书对相关研究人员与实际工作者也会很有帮助。

翻译本书是一项艰苦的工作,由于本书内容专业性强,涉及面广,我几乎放弃了手头上的其他工作全身心投入到翻译中来。我要特别感谢我的夫人王舒宁在翻译中给我的大量帮助。好友马君杰博士与钱陈博士也给了我很大的帮助,马博士甚至还参与了本书第 6 章的翻译工作;上海财经大学的孙海鸣教授、丛树海教授、赵晓雷教授与 ESRI 中国(北京)有限公司华东区总经理张乾乾先生也一直关心本书的出版。本书的出版得到了中国财政经济出版社财经易文出版中心的大力支持,总经理刘静女士和编辑刘占彬先生给予了热情的支持,刘占彬先生作为本书的责任编辑为本书的出版做了大量工作,在本书中文版付梓之际,在此一一表示感谢。另外,我还要感谢上海市重点学科建设项目(项目编号:B802)与上海市首届晨光计划(项目编号:2007CG47)对本书出版的资助。

最后,我还要感谢原书作者所提供的帮助。翻译过程中我与原书作者黄教授与李教授有过十余次的沟通,并将对原书的勘误一一请他们确定,两位教授总是耐心细致地给予回答,这也提高了本书中文版的质量。我也要感谢他们对本书的翻译工作所给予的积极评价。

本书介绍的空间统计分析方法应用面广,作为一名经济学者,本人在此



主要从经济学的视角来思考空间分析方法的重要性及其应用。实际上,本书介绍的空间统计与空间分析方法和工具对地理学、社会学等多种学科中所有涉及空间维度的研究均适用,不同领域的读者可以根据自己的专业背景与研究内容对本书介绍的方法和工具进行不同的应用。我与原书作者都希望本书中文版的出版能够在中国进一步推动相关领域的研究,也希望更多有志于空间研究的人士加入到这一行列中来。

张学良

2008年初于沪上

前言

尽管空间分析和空间统计已经存在了半个多世纪,但是 GIS——不论其中的“S”是代表“系统”(system)还是代表“科学”(science)——的历史却相对较短。从逻辑上讲, GIS 近些年的发展和成熟应该与空间分析和空间统计的进步密切相关。然而, 尽管空间分析被认为是 GIS 的一项强大功能 (Goodchild, 1987; Fotheringham and Rogerson, 1994), 但是相对于空间分析和空间统计的应用来说, GIS 的开发尚不能提供足够的支持。虽然已经有了一些空间统计软件包 (比如吕克·安塞林的 SpaceStat), 但是在 2000 年以前, 投入商业销售的 GIS 软件包中没有任何一种能够支持除缓冲、叠加和简单的距离计算等基本 GIS 程序之外的空间分析和空间统计程序。

我们从 1999 年开始努力将空间分析及空间统计功能与 GIS 整合起来, 并最终于 2001 年出版了《ArcView GIS 统计分析》(Statistical Analysis with ArcView GIS, John Wiley & Sons, Inc.) 一书, 据我们所知, 从严格意义上讲, 我们的这项工作是这方面最早的尝试。在该项目中, 我们使用 ArcView 3. X 中的 Avenue 脚本开发用于空间统计分析的程序和功能。这些脚本在单个项目文件 (. apr) 中以嵌入功能的形式提供。从那时起, 我们就收到了读者和用户的大量正面反馈和热切鼓励, 其中很多人并非来自地理或 GIS 学科领域。正是他们的反馈, 让我们知道我们在做一项正确的工作。

我们的第一本书和配套的 GIS 工具看起来已经实现了其教学目标, 即鼓励大家在更广泛的领域更好地运用空间统计。它们成功地向地理学者和非地理学者介绍了空间统计与空间分析方面的知识。但是, 读者告诉我们, 如果这本书编写成一本独立的教科书, 使得学生和科研人员都能够使用它, 那么它的影响将会更大, 贡献也会更多。在读者的鞭策和鼓励下, 我们编写了本书, 添加了一些基本统计知识。我们认为, 本书现在已经成为一本较为完备的教科书, 可以用来在定量分析方法教学中向不同层次的学生介绍基本统计概念和空间统计知识。从本书目前的形式来看, 我们认为它并不仅仅是一本教科书, 还可以用来向那些工作中较多涉及空间维度的专业人士、实务工作者和政策制定者介绍空间分析和空间统计知识。



自我们的第一本书推出以来,商家已经在一定程度上认识到空间分析和空间统计技术在 GIS 中的重要性。环境系统研究所(ESRI,加利福尼亚州雷德兰市)在开发最新的旗舰 GIS 软件包 ArcGIS 9 时就吸纳了我们在第一本书中讨论过的许多技术。这些技术目前在 ArcGIS 9. X 的空间统计向导中均可使用。用户可以利用这些工具计算基本的空间中心指标以及局部和全局空间自相关指标,此外还可利用它们进行热/冷点分析。尽管 ArcGIS 9. X 能够执行的空间统计只涵盖了本书所讨论的一部分,但我们很高兴看到渗透的过程已经开始。我们衷心希望 ArcGIS 及其他 GIS 软件包能够更多地引入空间统计工具,并能更好地执行空间统计。

为了更好地服务于读者,现在该是推出一本经过拓展的综合性新书的时候了。我们很高兴能够提供本书,并能得到 John Wiley & Sons 出版公司的支持。在此我们想指出,我们开发统计工具的途径与各商家略有不同。在本书中,我们仍使用 ArcView 3. X 及相关 Avenue 脚本作为我们的开发平台。确实,较新的 ArcGIS 在许多方面功能可能更丰富,尤其是在地图设计和渲染方面,但它似乎相对较慢,并且对硬件的要求也更高,这就使它在运行和使用分析工具进行空间统计时成本也更高(在内存和计算时间上)。此外,正如我们测试过的, ArcGIS 中的 Visual Basic for Applications 执行起来也比 ArcView 3. X 中的 Avenue 脚本慢得多。由于有些空间分析和空间统计程序涉及的计算量很大,因此带 Avenue 脚本的 ArcView 3. X 使用起来优势较为明显。

相对于第一本书来说,本书有以下改进:

- 第一本书是通过 ArcView 项目文件(.apr)介绍各种工具的,而本书则将各工具纳入不同的 ArcView 扩展模块(.avx)中。我们针对每一章的内容分别开发了专门的扩展模块。用户可以将多种扩展模块加载到同一个项目文件中来使用各套工具。
- 尽管这些工具是用 Avenue 脚本在 ArcView 3. X 中开发的,但本书附录概述了在 ArcGIS 9. X 中可以兼容的功能。扩展模块中的工具在更大范围上为一般的统计分析和空间分析—统计技术提供了支持。读者在执行 ArcGIS 9. X 中与本书提到的程序相对应的一些程序时可能要查看该附录。
- 本书各章都增加了 ArcView 范例,以说明如何运用空间统计量,以及如何使用配套的 ArcView 扩展模块。
- 传统统计学教科书通常会给出标准正态分布(z 值)、 t 分布、 F 分布、卡方分布(chi-square)等不同统计分布的概率表,而本书并未给出这些概率表,因为各 ArcView 扩展模块随时可以计算这些概率。

- 所有这些程序 (ArcView 3. X 扩展模块) 和数据集在配套的光盘中均能找到。光盘中的 README. TXT 文件对如何安装扩展模块、相关文件和数据集进行了详细说明。

据我们所知, 本书是唯一一本具有以下特点的著作: (1) 附带空间分析和空间统计方面的工具与程序, 这些工具和程序易于使用, 并且与极易获得的 GIS 软件包进行了整合; (2) 可让用户验证所讨论的空间分析和空间统计概念; (3) 提供了工具和数据, 因此用户可以运用书中所介绍的统计分析方法进行实验, 从而为自己日后在科研和实践活动中进行统计分析积累经验。

至于我们的下一个项目, 我们目前正在针对本书所讨论的所有空间统计工具开发新的工具包, 同样是以向导的形式。要了解相关更新和最新发展, 请直接询问我们或访问我们的网站 (<http://mason.gmu.edu/~dwong2> 或 <http://gis.geog.kent.edu/AVStat>)。

我们无意将为本书开发的统计工具作为商业软件包, 我们的主要目的是提供教学工具, 以分享我们的经验。尽管我们能够对 GIS 工具进行编程, 但我们并不是专业软件程序员, 我们的程序在性能上未必达到了最优, 但是可以确定的是它们能够正确运行, 并且相对高效。我们在程序中尽可能地加入了错误捕获功能。当然, 我们不可能开发出能够捕获所有由用户或数据带来的潜在错误的程序, 因此, 我们希望读者和用户能够为我们指出任何可能存在的问题、错误或改进空间。我们一直渴望与读者交流, 甚至合作开展研究。

部分程序在能够处理的记录个数上存在限制, 可能最多只能处理 256 个记录。这在一定程度上是由 dbf 格式的“列”限制造成的。对于部分这样的工具, 我们已经开发了替代性工具来绕开这些限制。如果读者需要这样的拓展工具, 可与我们联系。

戴维·黄和杰·李

2005 年 10 月

参考文献

- Fotheringham, A. S. and P. A. Rogerson (eds.). 1994. *Spatial Analysis and GIS*. Bristol, PA: Taylor and Francis.
- Goodchild, M. F. 1987. A spatial analytical perspective on geographical information systems. *International Journal of Geographical Information Systems* 7 (1): 3-19.

致 谢

我们要感谢乔治梅森大学 (George Mason University) 计算科学学院 (School of Computational Sciences) 地球系统和地理信息科学专业的研究生唐颖奇 (Yingqi Tang), 他为好几个扩展模块中多种统计分布的概率计算开发了 DLL 程序。我们还要对我们各自所在的单位——肯特州立大学 (Kent State University) 地理系和乔治梅森大学计算科学学院地球系统和地理信息科学专业——表示感谢, 感谢它们的支持, 感谢它们为本书的写作和配套工具的开发提供必要的实验室设备, 同时感谢它们让我们这两位主任从行政工作中抽出时间投入本书的写作。我们还要感谢 John Wiley & Sons 出版公司的吉姆·哈珀 (Jim Harper), 他对这第二个项目极为耐心, 并且给予了很多协助和支持。

最后, 我们要感谢父母支持我们在美国接受研究生教育, 并使我们最终能够自由地追求我们的兴趣和梦想。

目 录

前言	1
致谢	1
第1章 引言	1
1.1 为什么要进行统计和抽样 / 1	
1.2 空间数据有何特别之处? / 5	
1.2.1 可变区域单元问题(MAUP) / 6	
1.2.2 空间自相关 / 8	
1.3 空间数据与进行空间分析和空间统计的必要性 / 9	
1.4 空间分析与空间统计基础 / 11	
1.4.1 计量尺度 / 12	
1.4.2 数学符号 / 20	
1.4.3 尺度、范围和投影 / 23	
1.5 ArcView 说明——数据模型和范例 / 27	
1.5.1 ArcView GIS 中使用的数据模型 / 28	
1.5.2 随机抽样——一般方式 / 28	
1.5.3 随机和系统点抽样——一个扩展功能 / 33	
练习 / 39	
参考文献 / 40	

第1部分 经典统计

第2章 分布描述:一个变量(单变量)	45
2.1 集中趋势指标 / 45	
2.1.1 众数 / 46	



2.1.2	中位数	/ 48
2.1.3	均值	/ 50
2.1.4	分组平均数或加权平均数	/ 52
2.2	离散趋势指标	/ 54
2.2.1	极差、最小值、最大值和百分位数	/ 55
2.2.2	平均差	/ 56
2.2.3	方差和标准差	/ 57
2.2.4	加权方差和加权标准差	/ 60
2.2.5	变异系数	/ 61
2.3	ArcView 范例	/ 62
2.4	高阶矩统计量	/ 65
2.4.1	偏度和峰度	/ 65
2.5	ArcView 范例	/ 70
2.5.1	其他统计量	/ 76
2.6	应用范例	/ 80
2.7	小结	/ 86
	练习	/ 86
	参考文献	/ 87

第3章 关系描述:两个变量(双变量) 88

3.1	相关性分析	/ 88
3.2	相关性:定类尺度	/ 92
3.2.1	定类尺度和二元变量:phi 系数	/ 92
3.2.2	定类尺度和多元变量:卡方统计量	/ 94
3.3	相关性:定序尺度	/ 97
3.4	相关性:定距/定比尺度	/ 99
3.5	趋势分析	/ 102
3.5.1	简单线性回归模型	/ 102
3.5.2	判定系数	/ 105
3.5.3	实证范例	/ 106
3.6	ArcView 说明	/ 108
3.7	应用范例	/ 114
	练习	/ 117

参考文献 / 118

第4章 假设检验 119

4.1 与概率相关的几个概念 / 120

4.2 概率函数 / 122

4.2.1 二项分布 / 123

4.2.2 泊松分布 / 125

4.2.3 正态分布 / 127

4.2.4 ArcView 说明 / 131

4.3 中心极限定理与置信区间 / 132

4.4 假设检验 / 134

4.4.1 ArcView 说明 / 136

4.4.2 错误类型 / 136

4.5 参数检验统计量 / 137

4.5.1 方差差异 / 137

4.5.2 ArcView 说明 / 139

4.6 均值差异 / 143

4.6.1 小样本 / 143

4.6.2 大样本 / 146

4.6.3 ArcView 说明 / 147

4.7 均值与定值的差异 / 152

4.7.1 ArcView 说明 / 152

4.8 皮尔逊相关系数的显著性 / 153

4.8.1 ArcView 说明 / 153

4.9 回归系数的显著性 / 156

4.9.1 ArcView 说明 / 157

4.10 非参数检验统计量 / 160

4.10.1 卡方统计量 / 160

4.10.2 ArcView 说明 / 160

4.10.3 斯皮尔曼秩相关系数 / 163

4.10.4 ArcView 说明 / 163

4.10.5 柯尔莫洛夫—斯米尔诺夫检验 / 163

4.10.6 ArcView 说明 / 165



4.11 小结 / 165

练习 / 166

参考文献 / 168

第2部分 空间统计

第5章 点模式描述 171

5.1 点要素的性质 / 171

5.2 点分布的集中趋势 / 173

5.2.1 平均中心 / 175

5.2.2 加权平均中心 / 178

5.2.3 中位数中心 / 180

5.3 点分布的离散趋势与方位 / 181

5.3.1 标准距离 / 183

5.3.2 标准差椭圆 / 187

5.4 ArcView 说明 / 191

5.5 应用范例 / 197

练习 / 198

参考文献 / 200

第6章 点模式分析 201

6.1 尺度与范围 / 202

6.2 样方分析 / 204

6.2.1 基本概念 / 204

6.2.2 运用 K-S 检验比较观测分布与期望分布 / 209

6.2.3 运用方差—均值比率比较观测模式与期望模式 / 213

6.3 多阶邻点分析 / 218

6.3.1 最近邻点统计量 / 219

6.3.2 运用最近邻点统计量检验分布模式 / 222

6.3.3 高阶邻点统计量 / 225

6.3.4 最近邻点统计量的边界调整 / 227

6.4 K 函数 / 232

6.5 点的空间自相关 / 238