


T 高等院校研究生用书
EXTBOOKS FOR
GRADUATES

现代统计 分析方法与应用

何晓群 编著

(第二版)


 中国人民大学出版社

T 高等院校研究生用书
EXTBOOKS FOR
GRADUATES

现代统计 分析方法与应用

(第二版)

何晓群 编著

 中国人民大学出版社

图书在版编目 (CIP) 数据

现代统计分析方法与应用/何晓群编著. 2版
北京: 中国人民大学出版社, 2007
高等院校研究生用书
ISBN 978-7-300-08259-2

- I. 现…
- II. 何…
- III. 统计分析-研究生-教材
- IV. C813

中国版本图书馆 CIP 数据核字 (2007) 第 091224 号

高等院校研究生用书
现代统计分析方法与应用 (第二版)
何晓群 编著

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)		010-62511398 (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	北京民族印刷厂		
规 格	170 mm×228 mm 16 开本	版 次	1998 年 5 月第 1 版 2007 年 8 月第 2 版
印 张	28 插页 1	印 次	2007 年 8 月第 1 次印刷
印 数	512 000	定 价	33.00 元

版权所有 侵权必究 印装差错 负责调换

内容简介

统计方法与技术是现代社会科学、经济研究运用的基本方法。本书旨在提高研究生量化分析能力，在众多统计方法中，选择最实用的现代统计分析方法，在不失理论严密性的前提下，力求理清问题背景、方法思想、具体步骤、分析技巧。为重点突出方法的思想 and 应用，尽可能结合中国社会、经济、管理方面的实际问题，辅以案例研究，力求为非统计专业研究生进行量化分析起到一定示范作用。

本书既可作为非统计专业，如人文社会科学、财经管理类研究生学习现代统计分析方法之用，也可作为经济统计专业学生学习用书，同时也适合从事社会、经济、管理等研究和实际工作的从业人员进行量化研究的参考书。



作者简介

何晓群 中国人民大学统计学院教授，中国人民大学六西格玛质量管理研究中心主任，中国现场统计研究会副秘书长，国家税务总局特邀监察员。主要研究领域：应用数理统计，多元统计分析，六西格玛管理。多次参加国际学术会议，主持和参与多项国家和省部级及企业横向课题研究，发表论文80余篇。

策划编辑 陈永凤
责任编辑 邢伯春
封面设计 李尘工作室
版式设计 赵星华



第二版前言

自 1998 年本书出版以来,承蒙数万读者的厚爱,许多高校都采用本书作为研究生教材。有许多教师和学生给予我热情的鼓励和中肯的批评,这都是读者对我及本书的无私关心和奉献。在此我谨表衷心感谢。

几年来,一种责任心总在督促我对本书作出修改,重新出版,以表达对读者的谢意。在中国人民大学出版社陈永凤女士的精心策划下,我的几个研究生杨岚、孟鹏辉、郝燕梅、殷晓然为本书例题的运算做了大量工作,在此表示感谢。

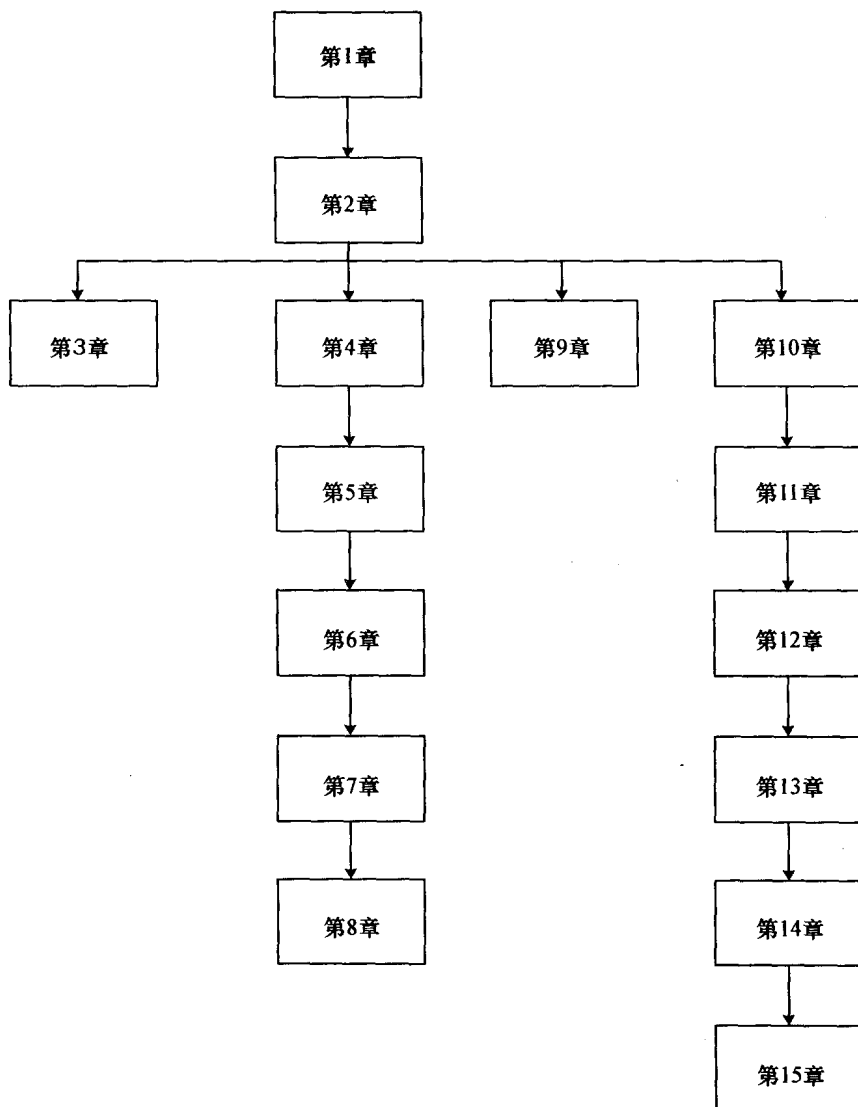
本次再版,我对本书做了如下修改:

1. 采用最新数据更换了大部分例题;
2. 大部分例题采用 SPSS (社会科学专用统计软件) 进行分析;
3. 增加了定性数据建模分析和多变量图表示法;
4. 对原书中一些错误之处进行了修正,并且斟酌了原书中的一些文字。

本书总的参考教学课时为 72 学时。根据笔者的经验,学生掌握这些基本方法和技能并不困难。如有计算机配合,教学效果将会更好。选用本书的教师在讲授时可有一定的灵活性,根据专业的不同有选择地讲授该书内容。一般 2 学分的课程可只讲授前 8 章的内容。从第 9 章开始是多元统计分析的内容,可看成高级统计分析方法。下面是本书一个建议讲授或阅



读顺序图:



路线 1: 对欲掌握基本统计分析方法者 36 学时可选讲第 1 章、第 2 章、第 4 章、第 5 章、第 6 章、第 7 章、第 8 章。

路线 2: 对欲掌握较为高级统计分析方法者 36 学时可选讲第 1 章、第 2



章、第 10 章、第 11 章、第 12 章、第 13 章、第 14 章、第 15 章。

第 3 章和第 9 章可根据需要任意选择在两条路线中讲授。

中国人民大学应用统计科学研究中心
中国人民大学六西格玛质量管理研究中心

何晓群

2007 年 5 月 1 日于北京九鼎山庄长安仁居



• 第一版前言 •

随着社会主义市场经济体制在中国的快速推进,许多非常复杂和重要的社会、经济问题急切地需要我们去研究。如金融风暴、通货膨胀、税务稽查、保险精算、环境保护、持续发展、国际竞争、公众意向,传统的定性研究理论和方法已无法满足这些热点问题的研究需要。面对 21 世纪,深刻的社会变革、迅猛的经济发展,使我国的人文社会科学、财经管理类研究生面临严峻的挑战和难得的机遇。时代呼唤我们精通定量分析的研究方法,掌握定性定量有机结合的研究技能。《现代统计分析方法与应用》一书正是适应这一需要,为非统计专业的人文社会科学、财经管理类硕士研究生学习现代统计分析方法而编著的。

统计方法与技术是现代社会科学、经济研究运用的基本方法。自 1969 年设立诺贝尔经济学奖以来,已有 40 多位学者获奖。这些获奖者大都精通现代统计方法,对统计方法的运用极为娴熟,在社会经济研究中取得了举世瞩目的成就。学习和运用统计方法已成为时代对我们的要求。

作者假定学生已具有线性代数、概率论与数理统计的基础知识,本着提高研究生量化分析能力的宗旨,在众多统计方法中,仅选择部分最实用的现代统计分析方法。在不失理论严密性的前提下,力求将问题的背景、方法的思想、具体的步骤、分析的技巧讲清楚。为重点突出方法思想和应用,每种方法尽可能结合中国社会、经济、管理方面的实际问题,以案例研究为导



向, 为非统计专业研究生进行量化分析起一定示范作用。

1996年, 中国人民大学率先在非统计专业的人文社会科学、财经管理类研究生中开设“统计方法与技术”必修课, 作者有幸给中国人民大学96级、97级硕士研究生和MBA主讲此课。在教学实践中, 学生们给了我许多启发和鼓励, 因为他们结合自己的专业, 对统计方法的学习产生了浓厚的兴趣, 看到了统计方法的用武之地, 清楚哪些方法最有用; 他们在学习的过程中也渴望拥有一本合适的教材。

本书的大部分内容都给非统计专业研究生讲授过, 根据笔者的经验, 如有计算机配合, 学生掌握这些基本方法和技能并不困难。选用本书的教师可有一定的灵活性, 根据不同专业有选择地讲授该书内容。本书参考教学课时为54学时。

本书也可作为经济统计专业现代统计分析方法课的教材。此书还可作为从事社会、经济、管理等研究和实际工作的同志进行量化研究的参考书。

本书在写作过程中, 始终得到中国人民大学研究生教材工作委员会和林岗、袁卫两位副校长的关注。中国人民大学研究生院高培勇教授、任兵处长, 统计系赵彦云、易丹辉、顾岚教授, 经济科学出版社莫霓舫、高续增同志给予作者许多鼓励和支持。香港浸会大学数学系首席教授方开泰先生对本书的写作给予悉心指点。我系兼职教授、中国科学院院士陈希孺教授, 上海财经大学张尧庭教授对本书的具体内容提出过不少中肯建议。台湾辅仁大学统计系主任谢邦昌教授与笔者的多次交谈, 也给我许多有益启示。我系博士孟生旺、刘勤、宋光辉、李序颖、刘长标等为本书的部分案例做过一些验算。合肥工业大学戴俭华教授, 西安统计学院王国梁教授, 华北电力大学张永林教授也曾给作者许多热情鼓励。在此, 我谨向对本书出版给予支持的师长和朋友表示衷心的感谢。

由于本人学识有限, 书中谬误之处在所难免, 恳请读者批评指正。

何晓群

1998年5月1日于中国人民大学静园

出教材学术精品 育人文社科英才
中国人民大学出版社读者信息反馈表

尊敬的读者：

感谢您购买和使用中国人民大学出版社的图书，我们希望通过这张小小的反馈卡来获得您更多的建议和意见，以改进我们的工作，加强我们双方的沟通和联系。我们期待着能为更多的读者提供更多的好书。

请您填妥下表后，寄回或传真回复我们，对您的支持我们不胜感激！

1. 您是从何种途径得知本书的：

书店 网上 报刊杂志 朋友推荐

2. 您为什么决定购买本书：

工作需要 学习参考 对本书主题感兴趣

随便翻翻

3. 您对本书内容的评价是：

很好 好 一般 差 很差

4. 您在阅读本书的过程中有没有发现明显的专业及编校错误，如果有，它们是：_____

5. 您对哪些专业的图书信息比较感兴趣：_____

6. 如果方便，请提供您的个人信息，以便于我们和您联系（您的个人资料我们将严格保密）：

您供职的单位：

您教授的课程（教师填写）：

您的通信地址：

您的电子邮箱：

请联系我们：

电话：82501868-633，541 62514162

传真：62514775

E-mail: rdcbsjg@crup.com.cn

通讯地址：北京市海淀区中关村大街甲59号文化大厦15层 100872

中国人民大学出版社工商管理出版事业部



目 录

第 1 章 概论	1
§ 1.1 为统计学正名	1
§ 1.2 市场呼唤真统计	3
§ 1.3 统计分析方法及其应用概述	5
思考与练习	13
第 2 章 统计学基础回顾	14
§ 2.1 统计数据的整理与描述	14
§ 2.2 几种重要的概率分布	18
§ 2.3 多元分布的基本概念	24
§ 2.4 多元正态分布	29
§ 2.5 参数估计	35
§ 2.6 假设检验	38
思考与练习	42
第 3 章 定性数据的 χ^2 检验	43
§ 3.1 多项分布与 χ^2 检验	43
§ 3.2 列联表分析	48
§ 3.3 一致性检验	57



§ 3.4	拟合优度检验	60
	思考与练习	66
第 4 章	一元线性回归	68
§ 4.1	一元线性回归模型	68
§ 4.2	回归参数 β_0, β_1 的估计	73
§ 4.3	最小二乘估计的性质	77
§ 4.4	回归方程的显著性检验	80
§ 4.5	残差分析	89
§ 4.6	预测和控制	93
§ 4.7	建模总结和应注意的问题	98
	思考与练习	104
第 5 章	多元线性回归	107
§ 5.1	多元线性回归模型	107
§ 5.2	回归参数的估计	110
§ 5.3	参数估计量的性质	118
§ 5.4	回归方程的显著性检验	121
§ 5.5	中心化和标准化	128
§ 5.6	相关阵与偏相关系数	130
§ 5.7	多元建模总结与评注	138
	思考与练习	145
第 6 章	违背基本假设的回归分析	147
§ 6.1	关于异方差性问题	147
§ 6.2	关于自相关性问题	160
§ 6.3	关于多重共线性问题	173
§ 6.4	异常值与强影响值	186
	思考与练习	192
第 7 章	自变量选择与逐步回归	194
§ 7.1	自变量选择对估计和预测的影响	195



§ 7.2	所有子集回归	198
§ 7.3	逐步回归	206
§ 7.4	实例与评注	217
	思考与练习	224
第 8 章	定性数据的建模分析	226
§ 8.1	对数线性模型基本理论和方法	227
§ 8.2	对数线性模型分析的上机实践	229
§ 8.3	Logistic 回归基本理论和方法	234
§ 8.4	Logistic 回归的方法及步骤	243
	思考与练习	245
第 9 章	多变量的图表示法	246
§ 9.1	散点图矩阵	246
§ 9.2	脸谱图	248
§ 9.3	雷达图与星图	252
§ 9.4	星座图	256
	思考与练习	258
第 10 章	聚类分析	259
§ 10.1	聚类分析的基本思想	259
§ 10.2	相似性度量	263
§ 10.3	类和类的特征	267
§ 10.4	聚类方法	270
§ 10.5	实际例子	281
	思考与练习	290
第 11 章	判别分析	292
§ 11.1	判别分析的基本思想	292
§ 11.2	距离判别	293
§ 11.3	Bayes 判别	296
§ 11.4	Fisher 判别	297



§ 11.5	逐步判别	299
§ 11.6	判别分析应用的几个例子	300
	思考与练习	317
第 12 章	主成分分析	319
§ 12.1	主成分分析的基本思想	319
§ 12.2	主成分分析的几何意义	321
§ 12.3	总体主成分及其性质	325
§ 12.4	样本主成分的导出	330
§ 12.5	主成分分析步骤及框图	333
§ 12.6	主成分分析的应用	335
	思考与练习	349
第 13 章	因子分析	350
§ 13.1	因子分析的基本思想	350
§ 13.2	因子载荷的求解	355
§ 13.3	因子分析的步骤与逻辑框图	362
§ 13.4	因子分析的上机实现	362
	思考与练习	377
第 14 章	对应分析	378
§ 14.1	对应分析的基本理论	378
§ 14.2	对应分析的步骤及逻辑框图	386
§ 14.3	对应分析的上机实现	388
	思考与练习	394
第 15 章	典型相关分析	395
§ 15.1	典型相关分析的基本理论	395
§ 15.2	典型相关分析的步骤及逻辑框图	403
§ 15.3	典型相关分析的上机实现	408
	思考与练习	412



参考文献	413
附录	415
表 1 泊松分布表	415
表 2 标准正态分布表	417
表 3 正态分布分位数表	420
表 4 卡方(χ^2)分布表	423
表 5 相关系数检验表	424
表 6 t 分布的临界点	425
表 7 F 分布表	426
表 8 D. W 检验上下界表	432



• 第 1 章 •

概 论

为使读者对现代统计学及其应用有较为系统的认识，本章将对统计学及其学科性质，统计学在现代社会、经济和管理等方面的应用给予扼要论述，并对统计分析方法与技术应用的步骤流程进行描述。

§ 1.1 为统计学正名

人文社会科学、财经管理类研究生在本科阶段大都学过“社会经济统计学原理”或“统计学”课程。“统计”一词对中国的普通老百姓来说也并不陌生。不少统计学者更认为政府统计才是统计。1988年出版的《中国大百科全书》经济卷Ⅱ，把统计学定义为一门社会科学。《大英百科全书》认为统计学是一门收集数据、分析数据，并根据数据进行推断的艺术和科学。然而，究竟什么是统计学，统计学的学科性质是什么，我们将要学习的统计学理论和方法究竟是社会科学，还是自然科学，同样需要我们去认识和思考。

1990年以前中国的高等统计教育，尤其是财经类的统计专业，基本上前苏联1954年统计工作会议决议的那种模式，即把统计学定义为一门具有阶级性、党性的社会科学。这种统计理论体系被中国统计界的一些学者称为“马克思主义的无产阶级统计学”。令人遗憾的是至今有不少统计界学者仍持这种观点。统计学是否具有阶级性我们姑且不论，但前苏联统计模式给中国统计界提供的示范作用以及对统计学在中国的发展所产生的历史性不良影响是不容置疑的。