



“十三五”全国统计规划教材

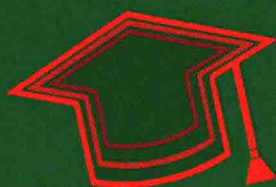
教育部学位管理与研究生教育司推荐研究生教学用书

(第三版)

医用多元统计 分析方法

陈峰 编著

陈启光 审阅



研究生用书



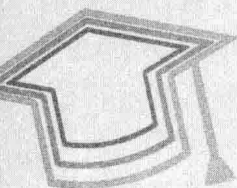
“十三五”全国统计规划教材

教育部学位管理与研究生教育司推荐研究生教学用书

(第三版)

医用多元统计 分析方法

陈 峰 编著
陈启光 审阅



研究生用书

图书在版编目(CIP)数据

医用多元统计分析方法 / 陈峰编著. — 3 版. — 北京 : 中国统计出版社, 2018.11

ISBN 978-7-5037-8704-1

I. ①医… II. ①陈… III. ①医学统计—分析方法—医学院校—教材 IV. ①R195.1

中国版本图书馆 CIP 数据核字(2018)第 222597 号

医用多元统计分析方法(第三版)

作 者/陈 峰

责任编辑/杨映霜

封面设计/张 冰

出版发行/中国统计出版社

通信地址/北京市丰台区西三环南路甲 6 号 邮政编码/100073

电 话/邮购(010)63376909 书店(010)68783171

网 址/<http://www.zgtjcs.com>

印 刷/河北鑫兆源印刷有限公司

经 销/新华书店

开 本/787×1092mm 1/16

字 数/600 千字

印 张/25.75

版 别/2018 年 12 月第 1 版

版 次/2018 年 12 月第 1 次印刷

定 价/72.00 元

版权所有。未经许可,本书的任何部分不得以任何方式在

世界任何地区以任何文字翻印、拷贝、仿制或转载。

如有印装差错,由本社发行部调换。

国家统计局

全国统计教材编审委员会第七届委员会

主任委员：宋跃征

副主任委员：叶植材 许亦频 赵彦云 邱 东 徐勇勇 肖红叶
耿 直

常务委员（按姓氏笔划排序）：

万东华	叶植材	许亦频	李金昌	杨映霜	肖红叶
邱 东	宋跃征	陈 峰	周 勇	赵彦云	耿 直
徐 辉	徐勇勇	郭建华	程维虎	曾五一	

学术委员（按姓氏笔划排序）：

方积乾	冯士雍	刘 扬	杨 灿	肖红叶	吴喜之
何书元	汪荣明	金勇进	郑京平	赵彦云	柯惠新
贺 铿	耿 直	徐一帆	徐勇勇	蒋 萍	曾五一

专业委员（按姓氏笔划排序）：

万崇华	马 骏	王汉生	王兆军	王志电	王 彤
王学钦	王振龙	王 震	尹建鑫	石玉峰	石 磊
史代敏	冯兴东	朱启贵	朱建平	朱 胜	向书坚
刘玉秀	刘立丰	刘立新	米子川	苏为华	杜金柱
李 元	李金昌	李 勇	李晓松	李 萍	李朝鲜
杨仲山	杨 军	杨汭华	杨贵军	吴海山	吴德胜
余华银	宋旭光	张 波	张宝学	陈 峰	林华珍
罗良清	周 勇	郑 明	房祥忠	郝元涛	胡太忠
洪永森	夏结来	徐国祥	郭建华	唐年胜	程维虎
傅德印	虞文武	薛付忠			

出版说明

全国统计教材编审委员会成立于1988年，是国家统计局领导下的全国统计教材建设工作的最高指导机构和咨询机构。自编审委员会成立以来，分别制定并实施了“七五”至“十三五”全国统计教材建设规划，共组织编写和出版了“七五”至“十二五”六轮“全国统计教材编审委员会规划教材”，这些规划教材被全国各院校师生广泛使用，对中国的统计和教育事业作出了积极贡献。自本轮规划教材起，“全国统计教材编审委员会规划教材”更名为“全国统计规划教材”，将以全新的面貌和更积极的精神，继续服务全国院校师生。

《国家教育事业发展规划“十三五”规划》指出，要实行产学研用协同育人，探索通识教育和专业教育相结合的人才培养方式，推动高校针对不同层次、不同类型人才培养的特点，改进专业培养方案，构建科学的课程体系和学习支持体系。强化课程研发、教材编写、教学成果推广，及时将最新科研成果、企业先进技术等转化为教学内容。加快培养能够解决一线实际问题、宽口径的高层次复合型人才。提高应用型、技术技能型和复合型人才培养比重。

《“十三五”时期统计改革发展规划纲要》指出，“十三五”时期，统计改革发展的总体目标是：形成依靠创新驱动、坚持依法治统、更加公开透明的统计工作格局，逐步实现统计调查的科学规范，统计管理的严谨高效，统计服务的普惠优质，统计运作效率、数据质量和服务水平明显提升，建立适应全面建成小康社会要求的现代统计调查体系，保障统计数据真实准确完整及时，为实现统计现代化打下坚实基础。

围绕新时代中国特色社会主义教育事业和统计事业新特点，全国统计教材编审委员会将组织编写和出版适应新时代特色、高质量、高水平的优秀统计规划教材，以培养出应用型、复合型、高素质、创新型的统计人才。

2015年9月，经李克强总理签批，国务院印发了《促进大数据发展行动纲要》系统部署大数据发展工作，我国各项工作进入大数据时代，拉开了统计教育和统计教材建设的大数据新时代。因此，在完成以往传统统计专业规划教材的编写和出版外，本轮规划教材要把编写大数据内容统计规划教材作为重点工作，以培养新一代适应大数据时代需要的统计人才。

为了适应新时代对统计人才的需要，组织编写出版高质量、高水平教材，本轮规划教材在组织编写和出版中，将坚持以下原则：

1. 坚持质量第一的原则。本轮规划教材将从内容编写、装帧设计、排版印刷等各环节把好质量关，组织编写和出版高质量的统计规划教材。

2. 坚持高水平原则。本轮规划教材将在作者选定、选题编写内容确定、编辑加工等环节上严格把关，确保规划教材在专业内容和写作水平等各方面，保证高水平高标准，坚决杜绝在低水平上重复编写。

3. 坚持创新的原则。无论是对以往规划教材进行修订改版，还是组织编写新编教材，本轮规划教材将把统计工作、统计科研、统计教学以及教学方法、方式的新内容融合在教材中，从规划教材的内容和传播方式上，实行创新。

4. 坚持多层次、多样性规划的原则。本轮规划教材将组织编写出版专科类、本科类、研究生和职业教育类等不同层次的统计教材，并可以考虑根据需要组织编写社会培训类教材；对于同一门课程，鼓励教师编写若干不同风格和适应不同专业培养对象的教材。

5. 坚持教材编写与教材研讨并重的原则。本轮规划教材将注重帮助院校师生学习和使用这些教材，使他们对教材中一些重要概念进一步理解，使教材内容的安排与学生的认知规律相符，发挥教材对统计教学的指导作用，进一步加强统计教材研讨工作，对教材进行分课程的研讨，以促进统计教材的向前发展。

6. 坚持创品牌、出精品、育经典的原则。本轮规划教材将继续修订改版已经出版的优秀规划教材，使它们成为精品，乃至经典，与此同时，将有意识地培养优秀的新作者和新内容规划教材，为以后培养新的精品教材打下基础，把“全国统计规划教材”打造成国内具有巨大影响力的统计教材品牌。

7. 坚持向国际优秀统计教材学习和看齐的原则。不论是修订改版教材还是新编教材，本轮规划教材将坚持与国际接轨，积极吸收国内外统计科学的新成果和统计教学改革的新成就，把这些优秀内容融进去。

8. 坚持积极利用新的教学方式和教学科技成果的原则。本轮规划教材将积极利用数据和互联网发展成果，适应院校教学方式、教学方法以及教材编写方式的重大变化，立体发展纸介质和利用数据、互联网传播方式的统计规划教材内容，适应新时代发展需要。

总之，全国统计教材编审委员会将不忘初心，牢记使命，积极组织各院校统计专家学者参与编写和评审本轮规划教材，虚心听取读者的积极建议，努力组织编写出版好本轮规划教材，使本轮规划教材能够在以往的基础上，百尺竿头，更进一步，为我国的统计和教育事业作出更大贡献。

国家统计局

全国统计教材编审委员会

第三版前言

本书 2000 年 12 月初版,2006 年修订。2005 年本书被教育部推荐为研究生教学用书,2015 年南京医科大学基于该教材开设的《医用多元统计分析方法》课程被江苏省认定为优秀研究生课程。该书被很多院校选择为医学研究生或生物统计学本科生的教材。在大家的鼓励下,2016 年再次修订出版。

本次再版,保持了前两版的风格。同时,增加了确证性因子分析和结构方程模型、典型相关分析和对应分析等内容。这些内容是目前医学科研中,尤其是高维数据分析中经常用到的方法。

修订后的第三版重新调整了章节内容,使得全书脉络更为清晰。第 1~2 章是多元分析基础,包括常用多变量统计量、多元正态分布,并增加了与多元正态分布有关的几种多元抽样分布,包括, Wishart 分布、Hotelling T^2 分布和 Wilks 分布;第 3~7 章是多重回归分析模型,包括多重线性回归、logistic 回归、Probit 回归、Poisson 回归、负二项回归、Cox 回归等,以及不同模型中逐步筛选变量的方法等,这些都属于线性或广义线性模型;第 8~12 章是降维分析技术,包括主成分分析,探索性和确证性因子分析,典型相关分析和对应分析,这些都属于对高维数据结构进行提炼、简化、降维后再分析的方法;第 13~14 章是分类技术,包括无师分类和有师分类,前者如系统聚类分析、条件聚类分析、动态聚类分析等,后者如距离判别分析、Fisher 判别分析和 Bayes 判别分析,以及逐步判别分析等。

本次再版更新了部分练习。读者如果需要实例和练习的原始数据以及相应的 SAS 代码,可以发邮件到 MVStat@163.com 索取。有关本书的建议和意见也可以通过该邮箱发给作者。

陈启光教授、荀鹏程博士、柏建岭博士和万雅同学仔细阅读了第三版部分书稿,并提出了很多中肯的建议;陈梦锴同学重新撰写了书中例题和习题的 SAS 代码;中国统计出版社的杨映霜编审为第三版的出版做了大量的组织协调工作。作者在此一并表示衷心感谢。

多元统计分析方法在医学研究中发挥了巨大的作用,已经成为医学研究中不可或缺的方法学之一。翻开医学学术期刊可以看到,几乎所有的医学研究,特别是基于人群的研究:临床试验、现场调查、队列研究、人群干预性试验等,无一不用到多元统计方法和统

计学思维。但是,读者需清晰地认识到各种统计分析方法的应用条件。一个在设计上存在严重缺陷,或设计严谨但实施不严肃、管理与质控缺位,甚至根本就没有设计或质控的研究,即使运用了统计学方法对资料进行了分析,也无法弥补设计中或实施中存在的缺陷。

陈 峰

2016 年 1 月于 南京

第二版前言

本书自 2000 年 12 月第一版问世以来,得到了读者的普遍关注,来信、来电不断,给了作者很多鼓励、建议和批评,有些学校将其列为多元统计分析的教材。去年岁末,本书被教育部推荐为研究生教学用书。甚为欣慰。借此重印之际,对原书作了修订。

此次修订对几个统计学名词作了统一:①将“multiple regression”译为多重回归,避免与“multivariate regression(多元回归)”相混淆,前者指一个因变量与多个自变量的回归,而后者是指多个因变量与多个自变量的回归;②将“odds”译为优势,“odds ratio”译为优势比,不再用“比数”和“比数比”的说法;③将“statistically significant”译为有统计学意义,不再用“显著性”一词,以免引起与实际资料中指标间的“显著差别”相混淆。

陈启光教授、于浩教授、荀鹏程博士、柏建岭博士仔细阅读了第二版书稿并提出了很多中肯的建议。中国统计出版社的杨映霜主任和吕军主任为本书的再版做了大量的组织协调工作。作者在此一并表示衷心感谢。

统计学是一门从数据中学习的科学,它植根于概率论和数学,并受计算科学的影响。面向应用是统计学自身发展的最直接的原动力,从应用中发现自身的不足,从而激发人们对新理论的探索和新方法的研究,似乎是统计学发展的永恒的轨迹。随着计算机技术和信息技术的迅猛发展,各门学科均得到了前所未有发展契机,统计学更是如此。而伴随着“一 omics”时代的到来,统计学在生物医学领域中的应用研究越来越广泛和深入,并不断与信息、计算科学交叉融合,在统计信息学、计算基因组学、遗传流行病学、进化和种群遗传学、计算神经学等方面有了突破性的进展;在统计理论研究方面上,Bayes 统计、计算密集型统计、高维数据统计等方面也得到了一定的发展,新的理论和方法不断涌现。

但是,统计学在其发展过程中已形成的理论和方法,不会因为新方法的涌现而被淘汰,尤其是较为经典者。例如,1713 年 J.Bernoulli 提出的二项分布,1761 年 T.Bayes 提出的 Bayes 定理,1800 年左右 Legendre 提出的最小二乘法,1809 年 C.F.Gauss 提出的正态分布,1900 年 K.Pearson 发展的 χ^2 检验,1921~1935 年期间 R.A.Fisher 提出的实验设计的基本原则、方差分析的基本思想、极大似然函数等,这些经典的理论和方法几乎每天都有成千上万的科学家们在使用。对经典理论和方法的深刻理解和领悟,必将有助于掌握现代统计思想和方法;相反,如不了解甚至误解经典理论和思想,则很难真正掌握和理解现代统计的精髓。因此,无论是统计理论工作者还是应用工作者,都应掌握并正确理

解经典统计的概念、理论和方法。这一点再怎么强调都不为过！本书现有内容是多元统计分析中最为基础的，且是医学研究生必须要掌握的。因此，这次修订仅限于修辞层面，而没有增添新的内容。

学习本身也是多元的，学习统计学更应该如此。很多读者反映统计学难学，有些人听了好几遍课仍不得要领。事实上，统计学是一门集理论性与实践性于一体的学科。要学好统计学，既要掌握其基本概念、基本原理和基本方法，又要不断练习、反复实践。模仿是学习的开端，因此，要找一本好书细细读，慢慢品，做做练习，谈谈体会。但不唯书。只看看书，听听课，无法超越书本，也难以领会统计学的精髓。实践出真知。统计学的理论和方法只有以“鲜活、多样”的实际问题为载体，方有用武之地；只有在实践中，所学的理论和方法才能得到最直接、最生动的验证；只有不断在实践中学习，才会加深对统计理论、概念、方法和分析策略等的领悟，方能超然书外，得心应手，运用自如。学而时习之，不亦说乎？与君共勉。

陈 峰

2006 年 12 月于 南京

第一版前言

第一次接触医用多元统计分析是在 1985 年史秉璋教授的讲习班上,他向我们展示了多元统计分析理论的博大精深,以及在医学各领域中的应用的精妙,他精湛的授课艺术和严谨治学的作风对我产生了极大的影响。从此,我迷上了多元统计分析。学习、讲授、应用,乐此不疲。

多元统计学起源于 20 世纪 20 年代,Wishart, Hotelling, Fisher, Roy 等是该领域的先驱。由于多元统计分析的计算量较大,开始时多局限于理论问题的研究。20 世纪 50 年代以后,随着计算机及统计分析软件的日益发展,多元统计方法越来越广泛地得到应用,并渗透到自然科学和社会科学的各个领域。与数学的其它理论一样,统计学的很多内容都是因实际需要而产生的,伴随着应用面的扩大和深入,多元统计分析的理论亦得到了突飞猛进的发展。实践证明,多元统计分析方法是一种有效的数据处理工具。

多元统计分析的书籍很多,有偏于理论方面的,也有偏于应用方面的。本书立足于应用,略涉及一些简单的理论问题,以使读者知其然,亦知其所以然。

多元统计分析方法的内容非常丰富,在选材时,笔者根据实际情况,选择了目前医学研究中最常用的一些方法,组成 10 个专题,每个专题一章。第 1 章是多元分析的基础,介绍多元分析中的基本统计量和多元正态分布及其应用。第 2 章介绍多元 T 检验和多元方差分析,详细讨论了成组设计、配对设计、区组设计、析因设计资料的多元方差分析。第 3 章介绍多元线性回归,详细讨论了衡量回归方程优劣的标准和变量筛选的技巧。第 4、5 章介绍主成分和因子分析,讨论了主成分的应用、因子分析的策略等。第 6 章介绍几种基于 logistic 分布的回归,讨论了条件的和非条件的 logistic 回归,多类结果和有序结果的 logistic 回归等,以及建模策略。第 7 章介绍广义线性模型的一般理论,包括参数估计、残差分析、拟合优度等,讨论了 logistic 回归与 probit 回归的区别, Poisson 回归与负二项回归的关系等。第 8 章介绍生存资料的分析,包括指数分布、Weibull 分布的拟合,指数回归、Weibull 回归和 Cox 回归模型的建立。第 9 章介绍了聚类分析,讨论了 8 种系统聚类及其联系,快速聚类,有序样品的聚类及条件系统聚类。第 10 章介绍了判别分析,讨论了距离判别, Bayes 判别, Fisher 判别等。其中 2~5、9、10 章介绍的方法属经典的多元分析方法,第 6、7、8 三章中介绍的方法是现代多元分析方法。为便于读者理解、掌握和正确应用,我们重点介绍各种方法所能解决的问题、应用条件及其局限性,

在每章的最后安排了“正确应用”一节,指出应用中常遇到的一些具体问题及解决办法。

附录 A 是专门为那些希望深入了解极大似然估计的读者而撰写的,这在练习编程时可能有帮助。书中的例题和练习,除个别为说明一些理论问题而杜撰的外,均为实际资料。这些资料大都短小精悍,除囿于篇幅,更重要的原因是便于说明问题。作者亦有意识地安排了个别大型资料的分析实例,目的是使读者从中领略到多元分析的一些策略和思路。为便于读者学习和练习,书中所有需计算的例题均给出了 SAS 程序,练习亦给出了参考答案。然而,要建立统计学思维,理解统计学方法,合理解释统计分析结果,不仅仅是读几本书,做几个练习,恐怕最重要的是在实践中反复应用,不断加深理解。

在应用多元分析时应注意:(1)必须思路清晰,知道自己要干什么;(2)在作多元分析前,必须先作描述性分析。只有在充分了解资料性质的基础上,才有可能正确选择方法,得出有价值的结论;(3)当所得结果不符逻辑,或有悖于专业知识时,既不要轻易接受,亦不要轻易放弃,必须弄清楚为什么。

限于作者水平,书中错缪之处在所难免,恳请同行专家和广大读者不吝赐教。

本书获江苏省跨世纪学术带头人专项基金和交通部学术专著出版基金联合资助。陈启光教授仔细审阅了原稿,字字斟酌,提出了许多中肯而富有指导性的意见,使许多错误在付印前得以更正。中国统计出版社的范仲实先生为本书的出版做了大量的组织工作。习题参考答案是荀鹏程、沈毅两位同志完成的。书中还大量引用了其它文献中的原始资料。作者在此一并表示衷心感谢。

最后,感谢我的导师陆守曾教授和杨树勤教授,正是在他们的指导和熏陶下,我才懂得统计学是 20 世纪最伟大的学科之一。

陈 峰

2000 年岁末 于南通

医学和生物现象变化万端,其因果关系更是错综复杂。人们所看到的某一结果的发生往往是众多因素综合作用的结果,通常并非某一因素的单一作用所致。例如,在大体相似的自然环境和社会条件下有些人得了某种疾病,而其余的人却不患该病,其致病因素往往不是单一的和唯一的,寻找这些致病因素的独立作用和联合作用是医学研究者的重要任务。

对于这些多因素共同作用的医学现象,要想探讨和澄清其中的必然规律,常用的单因素分析法(即一元分析)显然是无能为力的。因此,从上个世纪后半叶起,借助电子计算机这一有力计算工具的出现,国内外许多统计学先驱陆续把数理统计中的多元统计分析方法引入到医学研究中来,并在实践中获得了广泛的应用和取得了丰硕的成果。

实践中的成果和进一步的需要反过来又促进了医用多元统计分析方法不断向广度和深度发展,丰富和完善新的、功能优异的统计软件包的陆续涌现更使其如虎添翼。

进入 20 世纪末期,医学统计学界更是硕果累累,人才辈出,他们不仅有坚实的数学基础,更掌握电脑操作及其统计软件包使用方面的熟练技巧,再加上在医学应用方面的广泛实践经验,使得医用多元统计分析这一边缘性学科得到了迅猛的发展。陈峰教授就是其中的杰出代表之一。他在攻读硕士、博士期间,师从医学统计学界的著名学者,并继承了师辈们严谨的治学态度和开拓进取的好学精神,对本学科的发展前沿作了不懈的努力。近年来,随着其工作面的不断扩展,实践经验和理论修养也日益深化。本书就是在此背景下的一个宝贵产物。

该书较系统、全面地总结和反映了医用多元统计分析方法在目前阶段的发展现状及其应用成果和前景。作者在介绍统计方法时,从讲解实例入手,逐步深入,理论与实际并茂,突出应用背景,强调分析思路,并精心安排了大量习题,提供了参考答案和相应的 SAS 程序,有助于培养统计思维,指导研究设计,提高应用技巧。不同背景的读者均可从中获得系统的多元统计分析知识和实例借鉴。全书布局严谨,思路缜密;

文字流畅,可读性强。为本专业的教学、科研人员提供了一本有益的教材和参考专著,广大医务人员和医学院校师生亦可从中得到不少启示。

欣闻书稿即将付梓,特撰数言,权且为序。正值新世纪来临之际,相信该书的出版对于医学科学的发展将起到添砖加瓦的作用。

史果璋

2001 年 2 月于上海第二医科大学

目 录

1 绪论	(1)
1.1 多元分析常用统计量	(1)
1.1.1 均向量	(2)
1.1.2 方差—协方差矩阵	(2)
1.1.3 离均差平方和与离均差积和矩阵	(2)
1.1.4 相关系数矩阵	(3)
1.2 距离和相似系数	(3)
1.2.1 距离	(4)
1.2.2 相似系数	(6)
1.2.3 列联系数	(7)
1.3 多元分析在医学上的应用	(8)
1.3.1 比较	(8)
1.3.2 关系	(8)
1.3.3 综合	(8)
1.3.4 归类	(9)
2 多元正态分布	(10)
2.1 定义及性质	(10)
2.1.1 定义	(10)
2.1.2 性质	(12)
2.2 二元正态相关变量的参考值范围	(13)
2.3 其他多元分布	(15)
2.4 几个有用的抽样分布	(16)
2.4.1 Wishart 分布	(16)
2.4.2 Hotelling T^2 分布	(16)
2.4.3 Wilks 分布	(17)
3 均向量的统计推断	(18)
3.1 多元 T 检验	(18)
3.1.1 多元配对设计的均向量检验	(18)
3.1.2 多元成组设计两样本的均向量检验	(20)

3.2 多元方差分析	(21)
3.2.1 多元成组设计资料的分析	(22)
3.2.2 多元区组设计资料的分析	(24)
3.2.3 多元析因设计资料的分析	(26)
3.3 协方差阵的检验	(29)
3.3.1 $V=V_0$ 的检验	(29)
3.3.2 $V=\sigma^2 V_0$ 的检验	(30)
3.3.3 $V_1=V_2=\dots=V_k$ 的检验	(30)
3.4 多元方差分析的正确应用	(31)
4 多重线性回归	(33)
4.1 多重线性回归模型简介	(33)
4.2 回归系数的估计	(34)
4.2.1 矩阵计算法	(35)
4.2.2 消去变换法	(37)
4.3 方程的假设检验	(38)
4.3.1 y 方面变异的分解	(39)
4.3.2 回归方程的方差分析	(39)
4.4 决定系数与剩余标准差	(40)
4.5 偏回归系数的假设检验与区间估计	(41)
4.6 标准偏回归系数与自变量的贡献	(42)
4.6.1 标准偏回归系数	(42)
4.6.2 自变量作用的分解	(42)
4.6.3 复相关系数的分解	(43)
4.7 因变量的区间估计	(43)
4.7.1 y 的可信区间估计	(43)
4.7.2 y 的容许区间估计	(44)
4.8 指标的量化	(44)
4.9 衡量回归方程的标准	(45)
4.9.1 复相关系数	(45)
4.9.2 校正复相关系数	(45)
4.9.3 剩余标准差	(46)
4.9.4 赤池信息准则	(46)
4.9.5 C_p 统计量	(46)
4.10 逐步回归	(48)

4.11	回归系数反常的原因	(57)
4.12	岭回归	(59)
4.13	回归分析的正确应用	(61)
5	logistic 族回归	(63)
5.1	多元 logistic 回归	(63)
5.1.1	多元 logistic 回归模型	(63)
5.1.2	系数的解释	(64)
5.1.3	变量的假设检验	(67)
5.1.4	建模策略	(70)
5.1.5	四格表资料的 logistic 回归	(76)
5.2	配比设计的条件 logistic 回归	(78)
5.2.1	条件 logistic 回归模型	(78)
5.2.2	配对四格表资料的条件 logistic 回归	(82)
5.3	多类结果变量的 logistic 回归	(83)
5.3.1	多类结果变量的 logistic 回归模型	(83)
5.3.2	系数的解释与检验	(84)
5.3.3	建模策略	(87)
5.4	有序结果的累积优势 logistic 回归	(87)
5.4.1	累积优势 logistic 回归模型	(87)
5.4.2	累积优势模型的应用条件	(90)
5.5	有序结果的相邻等级 logistic 回归模型	(92)
5.6	logistic 族回归模型的正确应用	(94)
6	线性与广义线性模型	(97)
6.1	模型简介	(97)
6.1.1	线性模型	(97)
6.1.2	广义线性模型	(98)
6.1.3	指数分布族	(99)
6.1.4	联接函数	(100)
6.2	广义线性模型的建立	(101)
6.2.1	GLM 的参数估计	(101)
6.2.2	GLM 的假设检验	(103)
6.2.3	拟合优度	(104)
6.2.4	残差分析	(106)

6.3	logistic 回归与 Probit 回归	(107)
6.4	Poisson 回归	(109)
6.5	负二项回归	(113)
6.6	线性与广义线性模型的正确应用	(116)
7	生存分析模型	(117)
7.1	随访研究的特点	(117)
7.1.1	截尾数据	(117)
7.1.2	几个基本概念	(119)
7.1.3	随访资料的特点	(120)
7.2	生存分析的理论体系与常用指标	(121)
7.3	指数模型	(125)
7.3.1	指数分布模型	(125)
7.3.2	指数分布模型的参数估计	(126)
7.3.3	两个指数分布模型的比较	(126)
7.3.4	指数回归模型	(128)
7.4	Weibull 模型	(132)
7.4.1	Weibull 分布模型	(132)
7.4.2	Weibull 分布模型的参数估计	(133)
7.4.3	Weibull 回归模型	(135)
7.4.4	Weibull 回归与指数回归的比较	(135)
7.5	Cox 比例风险模型	(143)
7.6	生存分析的正确应用	(145)
8	主成分分析	(148)
8.1	主成分的定义	(148)
8.2	主成分的计算	(150)
8.3	主成分的性质	(152)
8.4	主成分的应用	(155)
8.4.1	主成分评价	(155)
8.4.2	主成分回归	(159)
8.5	有关的统计推断	(161)
8.5.1	特征根的可信区间估计	(161)
8.5.2	等相关性检验	(161)
8.5.3	主成分相等的检验	(162)