

□ 应用统计学丛书

An Introduction to
Categorical Data Analysis

属性数据分析引论

(第二版)

Alan Agresti 著

张淑梅 王 睿 曾 莉 译

辛 涛 校



高等教育出版社
HIGHER EDUCATION PRESS

属性数据分析广泛应用于社会科学、行为科学、生物医学、公共卫生、市场营销、教育和农业科学等许多领域。本书要求读者具有基础统计课程的知识，包括参数估计、假设检验、回归分析模型及方差分析模型。

全书分为十一章，分别介绍了与属性数据有关的统计学基础知识，列联表的相关知识，广义线性模型理论，logistic回归模型的理论、应用及其推广，对数线性模型，配对数据模型，聚簇关联响应的建模，广义线性混合模型以及属性数据分析方法的发展史。

本书配备了大量经典案例及适量习题，同时附有每种统计方法的SAS及SPSS的实现过程。

Copies of this book sold without a Wiley sticker on the cover are unauthorized and illegal.



学科类别：数学 / 统计
academic.hep.com.cn

ISBN 978-7-04-024751-0



9 787040 247510 >

定价 55.00 元

□ 应用统计学丛书

An Introduction to
Categorical Data Analysis

属性数据分析引论

(第二版)

Alan Agresti 著

张淑梅 王 睿 曾 莉 译

辛 涛 校



高等教育出版社
HIGHER EDUCATION PRESS

内容简介

属性数据分析广泛应用于社会科学、行为科学、生物医学、公共卫生、市场营销、教育和农业科学等诸多领域。本书的起点较低, 仅仅要求读者具有基础统计课程的知识, 包括参数估计、假设检验、回归分析模型及方差分析模型。全书分为十一章, 分别介绍了与属性数据有关的统计学基础知识, 列联表的相关知识, 广义线性模型理论, logistic 回归模型的理论、应用及其推广, 对数线性模型, 配对数据模型, 聚簇关联响应的建模, 广义线性混合模型以及属性数据分析方法的发展史。

本书内容详尽而不失精炼, 整个体系力求完整而不失紧凑, 语言深入浅出, 通俗易懂。配备了大量经典案例, 每章附有适量习题。这些习题编排精心, 由浅入深, 能够引导读者进一步凝练正文内容, 锻炼解决实际问题的能力。本书约有 100 个实际案例, 同时附有每种统计方法的 SAS 及 SPSS 的实现过程。

该书可以作为大学生及研究生学习“属性数据分析”入门课程的教材, 也可作为应用统计学家和在实际工作中涉及数据分析的专家学者的参考书。

图字: 01-2008-5529 号

Translation from the English language edition:
An Introduction to Categorical Data Analysis by Alan Agresti
Copyright ©2007 by John Wiley & Sons, Inc.
All rights reserved. This translation published under license

图书在版编目 (CIP) 数据

属性数据分析引论: 第 2 版 / (美)阿格雷斯特新
(Agresti, A.) 著; 张淑梅, 王睿, 曾莉译. — 北京: 高等教育出版社, 2008.12

书名原文: An Introduction to Categorical Data Analysis
ISBN 978-7-04-024751-0

I. 属… II. ①阿… ②张… ③王… ④曾… III. 社会调查-统计分析-统计程序 IV. C915-39

中国版本图书馆 CIP 数据核字 (2008) 第 154641 号

策划编辑	王丽萍	责任编辑	崔梅萍	封面设计	王凌波	责任绘图	黄建英
版式设计	马敬茹	责任校对	姜国萍	责任印制	韩 刚		

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100120
总 机 010-58581000

经 销 蓝色畅想图书发行有限公司
印 刷 廊坊市科通印业有限公司

开 本 787 × 1092 1/16
印 张 21.25
字 数 430 000

购书热线 010-58581118
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landraco.com>
<http://www.landraco.com.cn>
畅想教育 <http://www.widedu.com>

版 次 2008 年 12 月第 1 版
印 次 2008 年 12 月第 1 次印刷
定 价 55.00 元

译 序

属性变量不同于连续变量,属性变量的观测值仅表明结果所属的类别。如果我们研究的指标是一个属性变量,并希望用其它变量来说明或预测这个属性变量的取值时,不管用以说明的变量是属性的或连续的,使用的统计方法统称为属性数据分析^①。属性数据的分析方法和连续型数据的分析方法差异颇大。针对连续型数据的分析方法并不适合于分析属性数据。

专门针对属性数据的分析方法起源于 20 世纪初的英国。20 世纪早期,属性数据领域的研究主要集中于变量间的关联性。Pearson 于 1900 年引入了检验双向列联表统计独立性的卡方统计量。其后, Yule 提出了属性变量关联性的优势比度量。同为统计学家和遗传学家的 Fisher 于 1922 年引入了刻画卡方分布族特征的自由度,并于 1934 年引入了 Fisher 精确检验。针对属性数据的模型最早出现在 20 世纪 30 年代, probit 模型在处理毒理学中二分响应的应用中流行起来。1938 年 R. A. Fisher 和 Frank Yates 提出了 $\log[\pi/(1-\pi)]$ 变换。同是医师及统计学家的 Joseph Berkson 于 1944 年为该变换引入了术语“logit”。1951 年, Jerome Cornfield 通过该模型,在案例对照研究中使用优势比去近似相对风险, logistic 模型逐步流行起来。20 世纪 50 年代和 60 年代早期,出现了大量有关多向列联表关联性和交互结构的工作,这激起了 1965 至 1975 年间人们对对数线性模型的研究。20 世纪 70 年代早期,丹麦统计学家及数学家 Georg Rasch 在项目反应模型方面做出了杰出的贡献。近来,属性数据领域的研究焦点在于对聚簇关联响应数据拟合 logistic 回归模型。这些研究一方面对聚簇关联数据的边缘进行建模,另一方面利用广义线性混合模型进行建模。

近年来,伴随着计算机技术的迅猛发展,属性数据的分析方法被广泛应用于社会科学,心理科学,生物医学,教育,市场营销,金融信用评估等诸多领域。Alan Agresti 所著的“An Introduction to Categorical Data Analysis”得到了读者的一致好评。我们相信引进该书一定会对国内上述领域的数据分析师们有所帮助。鉴于此,我们产

^①引自:高慧璇,实用统计分析,北京大学出版社,2001。

生了翻译 Alan Agresti 这本 “An Introduction to Categorical Data Analysis” 的想法, 希望借此达到明理而兼顾应用的目的。

原著共十一章, 分别介绍了与属性数据有关的统计学基础知识, 列联表的相关知识, 广义线性模型理论, logistic 回归模型的理论、应用及其推广, 对数线性模型, 配对数据模型, 聚簇关联响应的建模, 广义线性混合模型以及属性数据分析方法的发展史。原著内容详尽而不失精炼, 整个体系力求完整而不失紧凑, 语言深入浅出, 通俗易懂。原著配备了大量经典案例, 每章附有适量习题。这些习题编排精心, 由浅入深, 能够引导读者进一步凝练正文内容, 锻炼解决实际问题的能力。

本书翻译过程中, 同时开设了两学期的研究生讨论班。在讨论班中, 我们除系统地研读原著之外, 还研读了 Alan Agresti 的另一本更富技术难度的 “Categorical Data Analysis” (2nd edition, Wiley, 2002) 以及一些相关的文献资料, 并且利用 SAS 软件系统重现了原著中的每个案例。感谢北京师范大学数学科学学院的研究生王奋际和孙佳楠, 他们对本书习题的翻译整理, SAS 程序的编译整理做出了颇多贡献。在此, 一并向参加讨论班的各位老师和研究生表示感谢, 感谢他们为译稿提出的各种宝贵意见。对王丽萍编辑为本书中译本的出版所付出的辛勤工作表示衷心的感谢。

原著涉及范围广泛, 而我们水平、经验有限, 加之时间仓促, 虽诚惶诚恐, 殚思极虑, 但疏漏之处在所难免。望各位专家, 同行和广大读者不吝赐教, 斧正。

译者

2008 年 5 月

于北京师范大学

第二版前言

近年来,针对属性数据的特殊统计方法的应用日益广泛,尤其是在生物医学和社会科学领域。这个现象一定程度上反映了过去几十年里属性数据分析方法的发展,同时也反映了科学家和应用统计学家方法论的日益精湛。如今,绝大多数科学家和应用统计学家已经意识到,将针对连续数据的分析方法应用于属性数据是没有必要而且通常是不合适的。

本书展示了针对属性数据最重要的分析方法,概述了诸如卡方检验等长期占主导地位的方法。除此之外,本书特别强调了建模技巧,尤其是 logistic 回归。

本书的表述基于一个较低的技术层面上,并不要求熟悉如微积分学和矩阵代数等高等数学内容。但是读者应该具备一些相关统计背景知识。这些知识包括非统计专业的本科生或研究生两学期的统计方法系列教材所含的内容,即估计、假设检验、回归模型。

本书是为学习属性数据分析入门课程的学生而设计。同时我也为应用统计学家和在实际工作中涉及数据分析的科学家而写。我希望这本书能对社会科学、行为科学和生物医学领域,以及公共卫生、市场营销、教育、生物、农业科学和工业质量控制等领域的分析师们处理属性响应数据有所帮助。

本书 1—8 章涵盖了最基础的属性数据分析方法。第 2 章将介绍诸如优势比、独立性检验、条件关联性以及边缘关联性等针对列联表的标准描述方法和推断方法。然而,我觉得可以通过在统计模型背景里观察统计方法以加强对它们的理解。因此,本书其余篇幅主要聚焦于属性响应数据的建模。第 3 章将介绍二分数据和计数数据的广义线性模型。第 4 章以及第 5 章将讨论二项(二分)数据的最重要的模型——logistic 回归模型。第 6 章将介绍名义和有序多项响应的 logistic 回归模型。第 7 章将讨论泊松(计数)数据的对数线性模型。第 8 章将展示针对配对数据的分析方法。

我认为 logistic 回归模型比对数线性模型更加重要,因为在实际应用中多数属性响应就是一个单独的二项或多项响应变量。因此我将主要精力放在了与此模型有关

的章节以及后面章节中讨论此模型推广的部分。与本书第一版相比,第二版更加强调了 logistic 回归而较少强调对数线性模型。

我喜欢通过使前述模型与通常的回归分析和 ANOVA 模型一致化来教授属性数据分析方法。第 3 章将通过广义线性模型来完成这个工作。一些教师可以选择轻描淡写地讲授这一章,主要将这一章用来介绍二项数据的 logistic 回归模型(3.1 及 3.2 节)。

第二版相对于第一版的主要变化是增加了两章,这两章主要是分析诸如具有个体重复测量的纵向研究中出现的聚簇关联属性数据。第 9 章和第 10 章把第 8 章的配对方法推广到聚簇数据中。第 9 章通过边缘模型完成这项工作,并强调了广义估计方程(GEE)方法,而第 10 章则基于随机效应更充分地对关联性进行建模。本书最后一章以历史的角度回顾了属性数据发展的历程(第 11 章),附录给出了本书中出现的几乎所有方法的 SAS 程序。

第 1—8 章的内容是属性数据分析入门课程的核心部分。若想为其它主题留出余力,可以略过 2.5、2.6、3.3、3.5、5.3—5.5、6.3、6.4、7.4、7.5 以及 8.3—8.6 节。教师可以选择 9—11 章的内容作为 1—8 章基础主题的补充。本书加星号的章节相对次要,想要迅速接触主题的读者可以略过这些章节。

相比于我的另一本书“*Categorical Data Analysis*”(2nd edition, Wiley, 2002),本书处于一个较低的技术水平上。我希望本书比那本书能吸引想要获得更多应用指导的读者。举例来说,本书并不试图推导似然方程、证明渐近分布、讨论前沿的研究工作,或罗列一个完整的参考目录。

本书所提到的大多数方法都需要大量的运算。对于大部分内容,我避开了复杂计算的细节,因为我觉得计算软件可以解决这个问题。针对属性数据分析的软件在多数大的商业软件包里很容易获得。我建议本书的读者尽可能地使用软件完成作业和检验书中的例子。附录里讨论了如何使用 SAS(特别是 PROC GENMOD)实现本书中所展示的几乎所有的分析方法。附录里的程序和许多本书里被分析的数据集可以从以下网页获得: <http://www.stat.ufl.edu/~aa/intro-cda/appendix.html>。网页 <http://www.stat.ufl.edu/~aa/cda/software.html> 包含了诸如 S-Plus 和 R、Stata 以及 SPSS 等其它软件的使用信息,并链接了由 Laura Thompson 提供的免费手册,这份出色的手册展示了 R 和 S-Plus 如何实现本书中几乎所有的例子以及它们的高级手册。网页中还罗列了本书第一版的一些错别字和谬误。

我要特别感谢 Brian Marx,过去十年里他为此书提出了诸多建议。他毫不吝惜他的时间,反馈他多次讲授此书所得的建议。他和 Bernhard Klingenberg 还非常慷慨的帮我审阅了本书第二版原稿并提出了许多有益的意见。我还要感谢那些对原稿提出意见以及对本书的案例和材料的覆盖范围给与建议的人们。他们包括: Anna Gottard(对 7.4 节提出建议)、Judy Breiner、Brian Caffo、Allen Hammer 和 Carlar Rampichini。也感谢对本书第一版给与帮助的人们,特别是 Patricia Altham、James

Booth、Jane Brockmann、Brent Coull、Al DeMaris、Joan Hilton、Peter Imrey、Harry Khamis、Svend Kreiner、Stephen Stigler 和 Larry Winner。最后，感谢那些为我的另一本书 (**Categorical Data Analysis**) 提供材料而引用于此的人们，特别是 Bernhard Klingenberg、Yongyi Min 和 Brian Caffo。非常感谢 Wiley 的 Steph Quigley，感谢他长期的关注。也感谢 Wiley 的全体工作人员一直以来的全力支持。

一如既往地，特别感谢给予我建议和鼓励的妻子 Jacki Levine。最后，著书真正令人愉快的副产品是有机会用其讲授短期课程并到各种机构去访学。这样，我有机会访问了约 30 个国家，认识了许多优秀的人。他们中的一些已经成为益友。我为他们奉上此书。

ALAN AGRESTI

英国，伦敦

2007 年 1 月

目 录

第 1 章 导 言	1
1.1 属性响应数据	1
1.1.1 响应变量和解释变量的区别	2
1.1.2 名义量表和有序量表的区别	2
1.1.3 本书的结构	3
1.2 属性数据的概率分布	3
1.2.1 二项分布	3
1.2.2 多项分布	5
1.3 比例的统计推断	5
1.3.1 似然函数和极大似然估计	5
1.3.2 二项比例的显著性检验	7
1.3.3 案例: 关于堕胎合法化的调查结果	7
1.3.4 二项比例的置信区间	8
1.4 关于离散数据的更多统计推断	9
1.4.1 Wald, 似然比和得分推断	9
1.4.2 二项参数的 Wald, 得分和似然比推断	11
1.4.3 小样本二项推断	11
1.4.4 小样本离散数据推断的保守性 *	12
1.4.5 基于中间 P -值的推断 *	13
1.4.6 小结	13
习题	14

第 2 章 列联表	18
2.1 列联表的概率结构	18
2.1.1 联合概率, 边缘概率以及条件概率	19
2.1.2 案例: 关于来世	19
2.1.3 诊断检验的敏感度和特异度	20
2.1.4 独立性	21
2.1.5 二项抽样和多项抽样	21
2.2 2×2 表比例的比较	22
2.2.1 比例差	22
2.2.2 案例: 阿司匹林与心脏病	22
2.2.3 相对风险	23
2.3 优势比	24
2.3.1 优势比的性质	24
2.3.2 案例: 阿司匹林和心脏病案例中的优势比	25
2.3.3 优势比和对数优势比的推断	25
2.3.4 优势比和相对风险的联系	27
2.3.5 案例对照研究中优势比的应用	27
2.3.6 观测研究的种类	28
2.4 独立性的卡方检验	29
2.4.1 皮尔逊统计量和卡方分布	29
2.4.2 似然比统计量	30
2.4.3 独立性检验	31
2.4.4 案例: 政党认同中的性别差异	31
2.4.5 列联表的单元残差	32
2.4.6 卡方统计量的分解	33
2.4.7 卡方检验的小结	34
2.5 有序数据的独立性检验	34
2.5.1 线性趋势与独立性	35
2.5.2 案例: 饮酒与婴儿畸形	35
2.5.3 有序检验的特殊功效	36
2.5.4 得分的选择	37
2.5.5 $I \times 2$ 表和 $2 \times J$ 表趋势的检验	38
2.5.6 名义变量 — 有序变量列联表	38
2.6 小样本的精确推断	38
2.6.1 2×2 表的费希尔精确检验	38
2.6.2 案例: 费希尔的品茶者试验	39

2.6.3	P -值和真实 P -值的保守性 (第 I 类错误)	40
2.6.4	优势比的小样本置信区间 *	41
2.7	三向列联表的关联性	41
2.7.1	部分表	42
2.7.2	条件关联与边缘关联: 死刑判决的案例	42
2.7.3	辛普森悖论	43
2.7.4	条件优势比和边缘优势比	44
2.7.5	条件独立和边缘独立	45
2.7.6	齐次关联性	46
	习题	46
第 3 章	广义线性模型	55
3.1	广义线性模型的构成部分	55
3.1.1	随机部分	56
3.1.2	系统部分	56
3.1.3	联系函数	56
3.1.4	正态 GLM	57
3.2	二分数据的广义线性模型	57
3.2.1	线性概率模型	58
3.2.2	案例: 打鼾与心脏病	58
3.2.3	logistic 回归模型	59
3.2.4	probit 回归模型	61
3.2.5	二分回归和累积分布函数 *	61
3.3	计数数据的广义线性模型	62
3.3.1	泊松回归	63
3.3.2	案例: 母鲨及其追随者	64
3.3.3	超散布性: 超出预期的变异性	67
3.3.4	负二项分布 *	68
3.3.5	比率数据的计数回归 *	69
3.3.6	案例: 英国的火车事故 *	69
3.4	统计推断和模型检验	70
3.4.1	关于模型参数的推断	70
3.4.2	案例: 再访打鼾与心脏病的案例	71
3.4.3	偏差	71
3.4.4	基于偏差的模型比较	72
3.4.5	比较观测和模型拟合的残差	72

3.5 广义线性模型的拟合	73
3.5.1 GLM 拟合的 Newton-Raphson 算法	73
3.5.2 依赖于似然函数的 Wald, 似然比以及得分推断	74
3.5.3 GLM 的优势	75
习题	75
第 4 章 logistic 回归	83
4.1 logistic 回归模型的解释	83
4.1.1 线性近似解释	84
4.1.2 母鲳: 观察并平滑二分结果	85
4.1.3 鲳: logistic 回归拟合的解释	86
4.1.4 优势比解释	88
4.1.5 回顾性研究中的 logistic 回归	88
4.1.6 X 服从正态分布意味着 Y 适合 logistic 回归	89
4.2 logistic 回归的推断	89
4.2.1 分组或未分组的二分数据	89
4.2.2 效应的置信区间	90
4.2.3 显著性检验	90
4.2.4 概率的置信区间	91
4.2.5 为什么使用模型估计概率?	91
4.2.6 概率的置信区间: 细节 *	91
4.2.7 模型参数估计的标准误 *	92
4.3 属性预测变量的 logistic 回归	93
4.3.1 用指示变量表示属性预测变量	93
4.3.2 案例: AZT 和 AIDS	94
4.3.3 因子的 ANOVA 型模型表示	95
4.3.4 $2 \times 2 \times K$ 列联表的 Cochran-Mantel-Haenszel 检验	96
4.3.5 优势比齐次性检验 *	97
4.4 多元 logistic 回归	97
4.4.1 案例: 以颜色和宽度作为预测变量的母鲳案例	97
4.4.2 通过模型对比确认某项是否必要	99
4.4.3 有序预测变量的定量化处理	100
4.4.4 容许交互效应	101
4.5 logistic 回归效应的概括	101
4.5.1 基于概率的解释	101
4.5.2 标准化解释	102
习题	103

第 5 章 logistic 回归模型的构建及应用	115
5.1 模型选择策略	115
5.1.1 运用多少个预测变量?	115
5.1.2 案例: 再访鲎的数据	116
5.1.3 逐步变量选择算法	117
5.1.4 案例: 鲎数据的向后剔除	118
5.1.5 AIC, 模型选择及“正确”模型	118
5.1.6 概括预测功效: 分类表 *	119
5.1.7 概括预测功效: ROC 曲线 *	120
5.1.8 概括预测功效: 相关 *	121
5.2 模型检验	121
5.2.1 模型比较的似然比检验	122
5.2.2 拟合优度与偏差	122
5.2.3 检验拟合: 分组数据, 未分组数据及连续预测变量	123
5.2.4 logit 模型的残差	124
5.2.5 案例: 佛罗里达大学的研究生入学	125
5.2.6 logistic 回归的影响诊断	126
5.2.7 案例: 心脏病与血压的关系	127
5.3 稀疏数据效应	128
5.3.1 无穷效应估计: 定量预测变量	128
5.3.2 无穷效应估计: 属性预测变量	129
5.3.3 案例: 带有稀疏数据的临床试验结果	130
5.3.4 小样本对 X^2 和 G^2 检验的影响	131
5.4 条件 logistic 回归与精确推断	132
5.4.1 条件极大似然推断	132
5.4.2 列联表的小样本检验	133
5.4.3 案例: 晋升能力	133
5.4.4 logistic 参数和优势比的小样本置信区间	134
5.4.5 小样本精确方法的局限性 *	134
5.5 logistic 回归的样本量与功效	135
5.5.1 比较两个比例所需的样本量	135
5.5.2 logistic 回归中的样本量 *	136
5.5.3 多重 logistic 回归中的样本量 *	137
习题	137

第 6 章 多类别 logit 模型	146
6.1 名义响应变量的 logit 模型	146
6.1.1 基线-类别 logit	146
6.1.2 案例: 钝吻鳄食物选择	147
6.1.3 估计响应概率	149
6.1.4 案例: 是否相信来世	150
6.1.5 离散选择模型	151
6.2 有序响应变量的累积 logit 模型	152
6.2.1 具有比例优势特性的累积 logit 模型	152
6.2.2 案例: 政治意识形态与隶属党派的关系	154
6.2.3 模型参数的推断	155
6.2.4 模型拟合的检验	156
6.2.5 案例: 对心理健康建模	157
6.2.6 比较累积概率的解释	158
6.2.7 潜变量诱导 *	159
6.2.8 响应类别选择的不变性	160
6.3 成对类别有序 logit	160
6.3.1 相邻类别 logit	161
6.3.2 案例: 再访政治意识形态	161
6.3.3 相继比 logit	161
6.3.4 案例: 发育毒性研究	162
6.3.5 聚簇数据中的超散布性	163
6.4 条件独立性检验	163
6.4.1 案例: 工作满意度和收入	163
6.4.2 推广的 Cochran-Mantel-Haenszel 检验 *	165
6.4.3 探测名义-有序条件关联	165
6.4.4 探测名义-名义条件关联	166
习题	166
第 7 章 列联表的对数线性模型	173
7.1 双向表和三向表的对数线性模型	173
7.1.1 双向表的独立性对数线性模型	174
7.1.2 独立性模型中的参数解释	174
7.1.3 双向表的饱和模型	175
7.1.4 三向表的对数线性模型	177
7.1.5 两因子参数描述条件关联	177
7.1.6 案例: 酒、香烟、大麻的使用	178

7.2	对数线性模型的推断	180
7.2.1	卡方拟合优度检验	180
7.2.2	对数线性单元残差	181
7.2.3	条件关联的检验	182
7.2.4	条件优势比的置信区间	182
7.2.5	高维对数线性模型	183
7.2.6	案例: 汽车事故与安全带	183
7.2.7	三因子交互作用	185
7.2.8	大样本和统计与实践显著性	186
7.3	对数线性模型与 logistic 模型的联系	186
7.3.1	利用 logistic 模型解释对数线性模型	187
7.3.2	案例: 再访汽车事故数据	187
7.3.3	对数线性模型和 logistic 模型间的对应	188
7.3.4	模型选择策略	189
7.4	独立图和衰退	189
7.4.1	独立图	190
7.4.2	三向表的衰退条件	190
7.4.3	衰退与 logistic 模型	191
7.4.4	多向表的衰退与独立图	192
7.4.5	案例: 学生成瘾物质使用的模型构建	192
7.4.6	图模型	194
7.5	对有序关联建模	194
7.5.1	线性—线性关联模型	196
7.5.2	案例: 性选择	197
7.5.3	有序的独立性检验	197
	习题	198
第 8 章	配对数据的模型	209
8.1	比较关联样本的比例	210
8.1.1	比较边缘比例的 McNemar 检验	210
8.1.2	比例的差的估计	211
8.2	配对的 logistic 回归	212
8.2.1	针对边缘比例的边缘模型	212
8.2.2	特定个体表与平均总体表	212
8.2.3	配对的条件 logistic 回归	213
8.2.4	匹配案例对照研究的 logistic 回归 *	214
8.2.5	McNemar 与 Cochran-Mantel-Haenszel 检验的联系 *	216