博學
MPH
公共卫生硕士系列
总主编 姜庆五 俞顺章

M

分类数据的统计分析及SAS编程

主编 刘勤 金丕焕

复旦大學出版社

分类数据的统计分析及SAS编程
主编 刘勤 金丕焕
复旦大學出版社

MPH

公共卫生硕士系列

总主编 姜庆五 俞顺章

分类数据的统计分析及SAS编程

主编 刘勤 金丕焕

复旦大学出版社

图书在版编目(CIP)数据

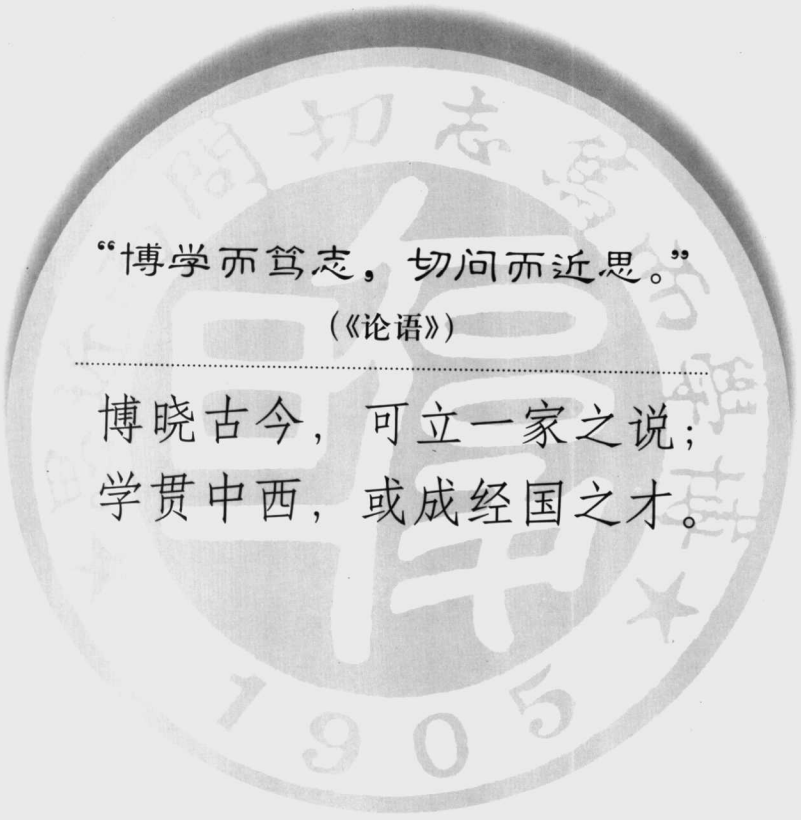
分类数据的统计分析及SAS编程/刘勤,金丕焕主编.—上海:复旦大学出版社,2002.9
(博学·MPH系列教材)
ISBN 7-309-03240-3

Ⅰ.分… Ⅱ.①刘…②金… Ⅲ.统计分析-应用软件,SAS-程序设计-教材 Ⅳ.C812

中国版本图书馆CIP数据核字(2002)第038656号

出版发行	复旦大学出版社
	上海市国权路579号 200433
	86-21-65118853(发行部) 86-21-65642892(编辑部)
	fupnet@fudanpress.com http://www.fudanpress.com
经销	新华书店上海发行所
印刷	同济大学印刷厂
开本	787×960 1/16
印张	16.5 插页 2
字数	245千
版次	2002年9月第一版 2002年9月第一次印刷
印数	1—3 000
定价	26.00元

如有印装质量问题,请向复旦大学出版社发行部调换。



“博学而笃志，切问而近思。”

（《论语》）

博晓古今，可立一家之说；

学贯中西，或成经国之才。

作者简介

刘勤，女，博士。毕业于山西医科大学公共卫生学院，并于1995年取得卫生统计学硕士学位。1998年在上海医科大学公共卫生学院取得博士学位。主要从事COX回归和临床试验的研究。目前在美国麻省大学医学院工作。

金丕焕，男，教授、博士生导师。1955年毕业于中国医科大学。1956年起在上海医科大学从事卫生统计学和计算机医学应用的教学和研究工作。著有卫生统计学、计算机医学应用及临床试验方面的多部著作和这些方面的论文。已培养多名硕士生和博士生。

内容提要

随着医学领域科学研究的飞速发展，以往研究生教材中的统计分析方法已不能满足实际工作的需要，尤其是当今医学科研中常常遇到反应变量是分类数据的资料。对于这类资料，以往国内统计书籍中介绍的方法远远不能达到充分利用分类资料提供的信息解决实际问题的要求。为了使医学工作者能充分利用科研及工作中千辛万苦收集到的宝贵资料，同时提高医学研究的统计分析水平，本书的编者们总结了近些年来用于分析分类数据的主要统计理论和方法，在本书中由浅入深地作了较详细的介绍。

本书首先介绍反应变量的分类尺度的种类，然后系统介绍了分析各种反应变量的统计分析方法，每种统计方法都附加有实例，同时列出了进行统计运算的SAS程序，并对程序及运算结果进行了详细说明和解释。全书内容具有新颖、实用、全面等特点，可使医学工作者在掌握了此统计学方法后，提高医学研究的统计分析水平。

本书为复旦版MPH(公共卫生硕士)系列教材之一，既可作为MPH课程的基本教材，又可作为医学科研人员和医学工作者的参考用书。

序　言

公共卫生硕士(MPH)是根据2002年国务院学位委员会批准设置的一个新的专业学位。MPH将成为公共卫生人才的重要职业教育形式。

MPH学位教育的目的是培养高层次卫生管理与疾病预防应用型人才。复旦大学培养的MPH的学生应该具备广博的专业知识、创新性的科学思维;勇于开拓、善于实践;能胜任卫生行政部门与医疗机构、疾病控制与卫生监督部门的高层次卫生管理与疾病预防的重要工作。在MPH学位教育过程中,我们将注重拓宽学生的知识面,注重现代科学技术的掌握,重点培养学生分析问题和解决问题的能力。

复旦大学公共卫生学院已经开展了5年的公共卫生应用型硕士研究生的教育,今年又被确定为我国MPH学位的首批试点单位。根据培养应用型研究生的经验,MPH学位教育过程中,我们将注重理论与实践,课堂教学和课题研究相结合。

我们设计的这套教材包括MPH学位的必修课,也有根据学生各自的基础和知识结构确立的选修课。其中不少教材已经在应用型研究生中应用,并收到良好效果。此系列教材包括:①MPH学位的"卫生事业(保健)管理"(health care management)专业方向课程,其中有卫生事业管理、卫生政策分析、医疗保险学、医院绩效管理、医学技术评估等课程。②"疾病控制"(disease control)专业方向课程,其中有流行病学基础、流行病学方法、卫生统计学、统计软件介绍、计算机在流行病学中的应用、重大传染性疾病防治案例等。③"环境医学与卫生监督"(environmental health and supervision)专业方向的课程,其中有环境卫生学、职业生命科学、营养与食品卫生、卫生法学与卫生监督学、毒理

学基础、卫生检验基础等。④“妇儿保健与健康促进”(women and child health and health promotion)专业方向课程,其中有妇幼卫生学、儿少卫生学、健康促进研究理论与实践、家庭与社区卫生服务等。

MPH 学位在我国尚处于试点时期,此套教材是我们开展 MPH 学位教育的探索,不当之处,请读者提出批评。我们将与全国的公共卫生教育者一起,为开拓与完善我国 MPH 学位教材建设作出贡献。

姜庆五 俞顺章

2002 年 9 月

前　言

根据当前研究生教学的需要，由复旦大学卫生统计教研室有关教师合作编写了《分类数据的统计分析及 SAS 编程》一书。本书主要面向医学工作者，具有理论结合实际、深入浅出和实用的特点。

随着医学领域科学研究的飞速发展，以往研究生教材中的统计分析方法已不能满足实际工作的需要，尤其是当今医学科研中常常遇到反应变量是分类数据的资料。对于这类资料，以往国内统计书籍中介绍的方法远远不能达到充分利用分类资料提供的信息解决实际问题的要求。为了使医学工作者能充分利用科研及工作中千辛万苦收集到的宝贵资料，同时提高医学研究的统计分析水平，本书的编者们总结了近些年来用于分析分类数据的主要统计理论和方法，在本书中由浅入深地作了较详细的介绍。

由于在分析分类数据时，其统计分析方法是根据反应变量的分类尺度来确定的，所以本书首先介绍了反应变量的分类尺度的种类，然后系统介绍了分析各种反应变量的统计分析方法。为了便于实际应用，每种统计方法都附加有实例，同时列出了进行统计运算的 SAS 程序，并对程序及运算结果进行了详细说明和解释。

本书是本人在原上海医科大学攻读博士学位及任教期间，在导师金丕焕教授的鼓励和指导下组织教研室的有关教师编写的。全书于 1999 年初完稿，年底定稿。本书在正式出版前曾作为原上海医科大学研究生试用教材。经过两年的教学试用，此次付梓前我们对部分章节进行了进一步的修订。本书的编写者已署名于章末。此外，上海日新医药发展公司统计部主管田晓燕硕士对本书原稿的部分内容提出了重要的修改意见并进行了详尽的计算，为本书的质量

提高作出了贡献。邓伟和罗剑锋老师参加了书稿的编辑工作，陈春巍及石桂芳老师参加了文字处理工作。

由于我们的知识和经验有限，本书难免存在各种问题，恳切希望读者提出意见和建议。

刘 勤

2002 年 2 月 20 日

于美国麻省大学医学院

目　录

1 概　述

1.1 分类数据

统计分析中经常遇到分类数据。所谓分类数据在这本书中指的是反应变量(应变量)为分类变量而不管说明变量(自变量)是分类变量或连续变量。

分类数据常常以列联表的形式表示。本书将讨论判定列联表或成套列联表中变量联系(association)的假设检验;也讨论描述说明变量与反应结果间联系的各种模型。

1.2 尺　度

分类数据中反应变量尺度的种类是选择正确分析策略的关键。选择适合于该数据尺度的分析方法可以得到良好的结果。如果不考虑尺度的种类,则可能导致用错方法而得出错误的结论。识别数据的尺度对正确应用分类数据的分析方法极为重要。

分类反应变量有以下几种尺度:

- 二分类尺度
- 顺序尺度
- 离散计数
- 名义尺度
- 分组生存时间

1.2.1 二分类尺度

二分类尺度是两种可能的结果，如治愈、未愈；生存、死亡；男性、女性等。

1.2.2 顺序尺度

有时结果不止两种可能性，而这些可能的结果存在顺序关系。如检验结果的－、＋、＋＋及＋＋＋；临床试验结果的无效、好转、显效和控制；水质的硬度为低、中、高等。

1.2.3 离散计数

与上述尺度不同的是尺度本身不是阴性、阳性或等级，而是离散计数本身。例如统计学生在一年中感冒的次数为 0,1,2,3 次等。

1.2.4 名义尺度

结果大于两种类别，而类别之间并没有顺序关系，例如职业、民族等类别。

1.2.5 分组生存时间

在生存时间数据中，如癌症病人手术后生存时间等，本来是连续数据，但有时可以把生存时间分组。这时，反应变量为在一定时间间隔内死亡的病人数。这就是分组生存时间数据。

1.3 分析策略

分类数据分析策略可以分成假设检验和建立模型。

假设检验法是建立一个关于联系(association)的假设。通常研究用随机化的方法进行。例如把病人随机分成两组，检验组别与疗效之间(列联表的行与列之间)是否有关。这种联系的无效假设为变量间没有联系，而备择假设一般有 3 种情况：①有一般联系(general associa-

tion)；②行平均分有差别(row mean scores differ)；③非零相关(non-zero correlation)。在以后讨论中我们将对各种不同的联系进行说明。

用建立模型的方法可求得各参数值，说明各因素的作用。通常用最大似然估计或加权最小二乘法估计。

本书前面部分介绍假设检验，后面部分介绍建立模型方法。

金丕焕