

研究生教学用书

教育部研究生工作办公室推荐

医学统计学与电脑实验 (附光盘)

*Medical Statistics and
Computerized Experiment*

With CD-ROM

(第二版)

方积乾 主编

上海科学技术出版社

研究生教学用书

教育部研究生工作办公室推荐

医学统计学与电脑实验 (附光盘)

Medical Statistics and
Computerized Experiment

With CD-ROM

(第二版)

方积乾 主编

上海科学技术出版社

图书在版编目 (CIP) 数据

医学统计学与电脑实验 / 方积乾主编. - 2 版. - 上海: 上海科学技术出版社, 2001, 7

ISBN 7-5323-5983-2

I. 医... II. 方... III. ①医学统计②计算机应用-医学统计 IV. R195.1

中国版本图书馆 CIP 数据核字(2001)第 032482 号

上海科学技术出版社出版发行

(上海瑞金二路 450 号 邮政编码 200020)

上海新华印刷厂印刷 新华书店上海发行所经销

1997 年 4 月第 1 版

2001 年 7 月第 2 版 2001 年 7 月第 2 次印刷

开本 787×960 1/16 印张 42.25 字数 761 000

印数 10 001-20 000 定价: 55.00 元(附光盘)

本书如有缺页、错装或坏损等严重质量问题,

请向本社出版科联系调换

内 容 提 要

本书共三篇,第一篇为统计学基本概念,第二篇为医学研究的设计与分析,第三篇为现代多元统计方法,共三十章。以医学统计学基本概念和常用的设计与分析为主干,辅以统计软件包 SAS 的操作和关于重要统计现象的电脑实验。

主 编 方积乾

副 主 编 (按姓氏笔画为序)

苏炳华 余松林 柳 青
骆福添 徐勇勇

编 者 (按姓氏笔画为序)

丁建生	兰州医学院
王 彤	山西医学院
王增珍	华中科技大学同济医学院
方积乾	中山医科大学
尹 平	华中科技大学同济医学院
朱淑明(兼秘书)	中山医科大学
孙 高	中国医科大学
刘 颜	中山医科大学
许宗利	中山医科大学
宇传华	第四军医大学
苏炳华	上海第二医科大学
李晓松	华西医科大学
何清波	上海第二医科大学
余松林	华中科技大学同济医学院
陈平雁	第一军医大学
陈 峰	南通医学院
林爱华	中山医科大学
赵耐青	复旦大学医学院
郝元涛	中山医科大学
柳 青	中山医科大学
骆福添	中山医科大学
姚 晨	军医进修学院
倪涛洋	中山医科大学
高桂明	中山医科大学

徐勇勇

凌莉(兼秘书)

康晓平

董时富

游志颖

熊国强

黎黎

第四军医大学

中山医科大学

北京大学医学部

华中科技大学同济医学院

中山医科大学

中南大学医学院

中山医科大学

前 言

(第一版)

我国医学院校和研究机构现行的研究生教学计划大多将医学统计学列为必修课,约 100 学时;有些还另设选修课,约 50 学时。可见医学统计学在医学科研中地位之重要。国内外虽不乏医学统计学佳作,但取之作为上述研究生课程的教材却有诸多不便。为此,许多同行和学生鼓励我们尽快编写一本与现行教学计划基本相吻合的教材。我们这个编写组于 1993 年成立。往日的友谊使我们很容易切磋交流,统一意向。我们将书名定为《医学统计学与电脑实验》,全书以医学统计学基本概念和常用的设计与分析为主干,辅以统计软件包 SAS 的操作和关于重要统计现象的电脑实验。我们商定共同努力营造如下特色:以中学初等数学为起点深入浅出地解释统计学概念与思维逻辑;覆盖经典的和现代的重要医学统计方法;将医学研究的设计与分析有机结合;精心设计每章的电脑实验使抽象理论变得具体生动;每章设思考、练习与实验帮助学有所用,用有所思。我们于 1995 年 6 月完成初稿,经互相审阅后一并交主编协调、润色,求得内容、形式和风格的统一。付印之前,主要章节又在作者们各自的学校试教,并征求同行意见。本书共 30 章,每 10 章为一篇。第一篇为统计学基本概念,第二篇为医学研究的设计与分析,第三篇为现代多元统计方法。连同理论课与实习课,每篇各需 50 学时。第一、二两篇合起来可作医药卫生领域研究生必修课医学统计学的教材;第三篇可作医药卫生领域研究生选修课的教材。实际上,第一篇也可作医疗系本科生医学统计学课程的教材,第一、二两篇也可作预防医学系本科生必修课《卫生统计学》的教学参考书。附录中有统计软件包 SAS 的入门简介。每章之后各设一节指导读者利用 SAS 进行电脑实验。电脑设备较多的院校可将现行的实习课扩展为电脑实习课,在教员指导下利用书中提供的 SAS 程序,学员可主动地观察随机现象的统计规律。电脑设备尚感不足的院校可由教员演示电脑实验过程或讲解实验所见。由于电脑实验一节相对独立,完全不具备示教条件的单位不妨跳过这一节,仍可进行其他各节的教学。每章之后的思考、练习与实验可供实习讨论课选用,部分也可作为课外作业和巩固复习内容。多数题目来自实践或文献,某些甚至没有最佳而只有较佳答案,颇有思考、讨论、实验和回味的余地。本书的附录二提供了若干医学实例和数据,读者可借以练习解决

实际问题的能力,附录三是汉英名词对照,汇集了全书散在的主要术语,方便读者查阅。

我们衷心地感谢本书的二位顾问第四军医大学郭祖超教授和中山医科大学胡孟璇教授的指导,感谢中山医科大学研究生处的支持,也感谢上海科学技术出版社唐仲华编辑的帮助。还要感谢作者们所在单位的同事和研究生们,在搜集资料、调试程序、打字、校对以及其他事务性工作方面他们默默地伸出援助之手,他们的名字不胜枚举,一一记在作者们的心中。

限于我们的学识和精力,这本书还有许多不足之处,我们将虚心吸取广大读者的批评与建议,争取在再版中给予弥补。

方积乾

1995年9月于广州

前 言

(第二版)

接中华人民共和国教育部教研办[2000]12号文件,得悉经专家评审《医学统计学与电脑实验》(上海科学技术出版社,1997)一书被列为教育部推荐的研究生教学用书。教育部研究生工作办公室建议第二版适当补充新内容和参编单位。于是,我们于2000年6月24~26日在广州中山医科大学召开了修编会议。参加会议的有20多名统计学专家,分别来自复旦大学医学院、上海第二医科大学、华西医科大学、第四军医大学、华中科技大学同济医学院以及中山医科大学等14所高等医学院校。经3日认真地讨论,确定了这次修编的原则是:尽量维持原书结构;适当增加新内容;强化概念,突出应用,淡化计算。经过4个月的共同努力,第二版脱稿了。

从全貌来看,第二版较之第一版在可读性方面有较大进步。第一,每章都是从医学问题的挑战开始,围绕实例讨论思路与对策、介绍统计学概念与方法,最终落实于问题的解决和结果的解释。第二,除正面介绍现代医学研究中常用的统计学概念、设计与方法外,本书还不时通过展示典型的不妥与错误和医学实例点评等从反面提醒读者,加深印象。第三,电脑实验更容易实施,除原有的SAS程序外,还用SPSS和EXCEL平行地编写了全部程序。

第二版的实用性也较前鲜明。第一,完整地新增了两章:连续型重复测量资料的分析(现第二十二章)和遗传资料统计分析方法(现第三十一章);前者是当前应用最广和出错最多的内容,后者是人类基因组学推动之下必定越来越广泛采用的内容。第二,大部分章节都适当增加了现代医学研究中十分重要的内容,如临床试验中的统计学原则(第十一章)、ROC曲线(第十八章)、通径分析(原第二十二章,现第二十一章)、量表的效度、信度与反应度(原第二十章,现第十三章)、确证性因子分析(第二十六章)、对应分析(第二十七章)、多分类logistic回归模型(第二十八章)和Poisson回归(第三十章)。第三,附录扩充成书面和光盘两部分,前者包括SAS简介、统计用表、参考答案、参考文献和汉英、英汉两种统计词汇对照与索引;后者包括电脑实验所需的SAS和SPSS可调用程序以及医学科研数据实例。

第二版仍然分三篇。第一篇统计学基本概念,包含十章;第二篇医学研究的

设计与分析,包含九章;第三篇现代统计方法,包含十二章。第一篇和第二篇可用于必修课,理论课和实习课各 50 学时左右;第三篇可用于选修课,理论课约 30 学时,实习课约 20 学时。若全书内容一章不漏地用于教学则需 170~180 学时,各校可根据自己的教学计划和学时数酌情取舍。有些章节标有星号,可在课程进行期间组织学生课外阅读或作读书报告用。

这次修编首先得力于许多兄弟院校工作在教学第一线的优秀中青年教师们,通过参与修编他们无私地贡献了各自的智慧和经验;中山医科大学医学统计学教研室的老师们虚心学习,认真总结,努力将本校和兄弟院校多年来使用第一版进行教学的成功经验融入新版之中;该教研室的技术员吴少敏、曾芳芳等和研究生王梅华、王奇玲、陈妍、刘国辉和郑迎东等全力支援,既为老师们尽心服务,又确实得到了锻炼。谨代表编委会对参与此次修编的全体人员一并致谢。

在短短 4 个月内,有这么多学校、这么多专家参与修编,千头万绪难免疏漏;何况,智者千虑必有一失,也许旧错误改掉了,新毛病又出现了。敬请读者和同行继续呵护这本书,严格要求,不吝指正,以便不断修编,一版胜过一版。

方积乾

2000 年 10 月于广州

目 录

第一篇 统计学基本概念

绪 论	3
第一章 描述性统计	8
第一节 变量与数据	8
第二节 频数表与直方图	10
第三节 样本平均水平的度量	14
第四节 样本变异性的度量	18
第五节 相对数与率的标准化的度量	20
第六节 电脑实验	24
思考、练习与实验	26
第二章 概率与分布	29
第一节 概率的意义及相关的一些概念	29
第二节 随机变量的分布特征	31
第三节 二项分布	34
第四节 Poisson 分布	38
第五节 正态分布	40
*第六节 概率的基本运算	46
第七节 电脑实验	50
思考、练习与实验	51
第三章 样本均数的抽样误差与置信区间	53
第一节 样本均数的分布	53
第二节 t 分布	57
第三节 正态分布总体均数的置信区间	60

第四节	两正态总体均数之差的置信区间	61
第五节	二项分布总体概率以及概率之差的置信区间	63
第六节	估计置信区间所需的样本量	65
第七节	电脑实验	66
	思考、练习与实验	68
第四章	连续型资料的假设检验	71
第一节	假设检验的独特逻辑	71
第二节	单组完全随机化设计资料均数的 t 检验	73
第三节	随机化配对设计资料均数的 t 检验	75
第四节	两组完全随机化设计资料均数的 t 检验	77
第五节	两组完全随机化设计资料方差齐性的 F 检验	80
第六节	二项分布和 Poisson 分布大样本资料参数的 Z 检验	82
第七节	实例点评	88
第八节	电脑实验	89
	思考、练习与实验	91
第五章	假设检验的功效与样本量	94
第一节	两类错误与功效	94
第二节	影响功效的四要素	96
第三节	功效与四要素的定量关系	97
第四节	常用统计检验的样本量估算	102
第五节	实例点评	104
第六节	电脑实验	105
	思考、练习与实验	106
第六章	离散型分类计数资料的 χ^2 检验	108
第一节	χ^2 分布和 Pearson 拟合优度检验	108
第二节	比较两独立样本比率的 χ^2 检验	109
第三节	2×2 交叉分类资料的 χ^2 检验	113
第四节	$R \times C$ 列联表资料的 χ^2 检验	116
第五节	检验频数分布的 χ^2 检验	121
* 第六节	四格表精确概率检验法	122
第七节	两个标准化率相等的假设检验	124

第八节 电脑实验	126
思考、练习与实验	129
第七章 基于秩次的非参数检验	132
第一节 配对样本的符号秩检验	132
第二节 两独立样本分布位置相同的假设检验	136
第三节 多个样本分布位置相同的假设检验	139
第四节 电脑实验	146
思考、练习与实验	149
第八章 线性相关	151
第一节 线性相关概念	151
第二节 相关系数	153
第三节 相关系数的统计推断	154
第四节 等级相关	156
第五节 线性相关分析的注意事项	158
第六节 电脑实验	159
思考、练习与讨论	160
第九章 线性回归	162
第一节 线性回归的统计描述	162
第二节 回归系数的统计推断	164
第三节 线性回归的应用及其前提条件	168
第四节 电脑实验	173
思考、练习与讨论	175
第十章 非线性回归	177
第一节 利用线性回归拟合曲线	177
第二节 非线性回归的最小二乘拟合	181
第三节 电脑实验	183
思考、练习与实验	185

第二篇 医学研究的设计与分析

第十一章 实验设计的统计学基本原则	189
第一节 实验设计中的基本概念	189
第二节 临床试验中的统计学要求	192
第三节 随机化技术	199
第四节 医学实例点评	201
第五节 电脑实验	203
思考、练习与实验	206
第十二章 方差分析	207
第一节 完全随机设计与试验结果分析	207
第二节 随机区组设计与试验结果的分析	214
第三节 拉丁方设计与试验结果的分析	216
第四节 析因设计与试验结果的分析	219
* 第五节 裂区设计与试验结果的分析	224
* 第六节 二阶段交叉设计与试验结果的分析	227
第七节 电脑实验	230
思考、练习与实验	234
第十三章 测量手段的效度和信度	238
第一节 测量手段的统计评价	238
第二节 描述信度的统计量	239
第三节 描述标准效度的统计量	244
第四节 量表的信度、效度和反应度	246
第五节 电脑实验	251
思考、练习与实验	254
* 第十四章 序贯试验设计与分析	256
第一节 基本概念	256
第二节 质反应序贯设计与分析	256
第三节 量反应序贯设计与分析	258

第四节 成组序贯试验·····	259
第五节 电脑实验·····	264
思考、练习与实验·····	265
第十五章 横断面研究的设计与分析·····	268
第一节 研究设计·····	268
第二节 估计总体参数的抽样方法与参数估计·····	269
第三节 样本含量的估计·····	274
第四节 现时寿命表·····	278
第五节 电脑实验·····	283
思考、练习与实验·····	284
第十六章 追踪研究的设计与分析·····	286
第一节 研究设计·····	286
第二节 发病率的计算·····	287
第三节 追踪资料的分析·····	291
第四节 电脑实验·····	300
思考、练习与实验·····	307
第十七章 病例-对照研究的设计与分析·····	309
第一节 病例-对照研究的设计·····	309
第二节 成组比较资料的分析·····	313
第三节 匹配比较资料的分析·····	319
第四节 电脑实验·····	323
思考、练习与实验·····	325
第十八章 诊断和筛查试验的研究设计与分析·····	327
第一节 试验设计与资料形式·····	327
第二节 常用诊断试验的评价指标·····	327
第三节 诊断或筛查试验的 ROC 分析·····	331
第四节 基于诊断或筛查试验的决策分析·····	337
第五节 电脑实验·····	340
思考、练习与实验·····	343

第十九章 医学文献综合研究与 meta 分析	345
第一节 基本概念	345
第二节 meta 分析常用统计方法	348
第三节 电脑实验	354
思考、练习与实验	357

第三篇 现代统计方法

第二十章 多元统计量与多元方差分析	361
第一节 常用多元统计量	361
第二节 两个均数向量的比较——Hotelling T^2 检验	365
* 第三节 多个均数向量的比较——多元方差分析	369
第四节 电脑实验	372
思考、练习与实验	375
第二十一章 多重回归与相关	377
第一节 多重线性回归模型及基本分析步骤	377
第二节 多重线性相关	381
第三节 回归分析中自变量的选择	384
第四节 多重线性回归应用中的若干技巧	388
* 第五节 通径分析	392
第六节 电脑实验	396
思考、练习与实验	398
第二十二章 连续型重复测量资料的分析	400
第一节 重复测量资料举例	400
第二节 分析中常见的不妥及其原因	401
第三节 综合指标的分析	403
第四节 重复测量资料的一般线性模型	404
第五节 电脑实验	408
思考、练习与实验	410

第二十三章 判别分析	412
第一节 判别分析的基本思想.....	412
第二节 Fisher 准则下的判别分析.....	413
第三节 Bayes 准则下的判别分析.....	415
第四节 逐步判别分析.....	418
第五节 回顾性考核和前瞻性考核.....	419
第六节 应用中的注意事项.....	421
第七节 电脑实验.....	422
思考、练习与实验.....	423
第二十四章 聚类分析	424
第一节 聚类分析的意义.....	424
第二节 系统聚类法.....	426
第三节 快速聚类法.....	429
第四节 对变量的聚类.....	430
第五节 电脑实验.....	431
思考、练习与实验.....	433
第二十五章 主成分分析	435
第一节 主成分的基本概念.....	435
第二节 主成分的计算与解释.....	438
第三节 主成分回归.....	441
第四节 电脑实验.....	443
思考、练习与实验.....	445
第二十六章 因子分析	449
第一节 因子模型.....	449
第二节 初始因子的提取.....	450
第三节 因子图和因子的旋转.....	453
第四节 因子得分与因子模型的应用.....	457
* 第五节 证实性因子分析.....	458
第六节 电脑实验.....	462
思考、练习与实验.....	463