

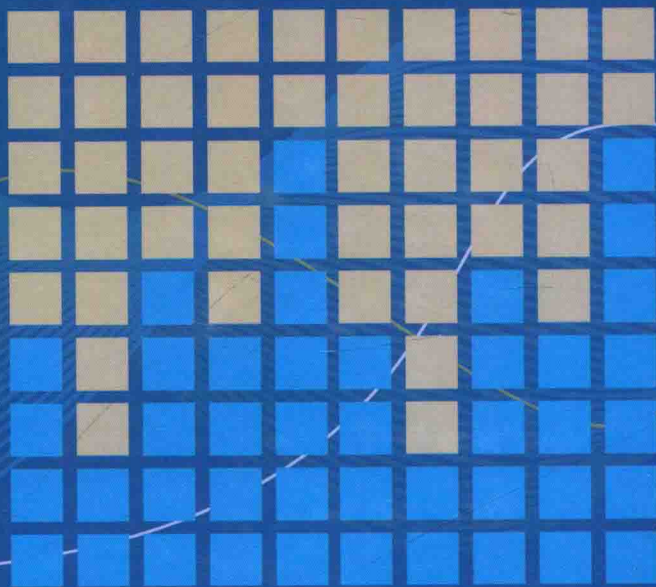
中国科学院教材建设专家委员会规划教材

全国高等医学院校规划教材

高级医学统计学

Advanced Medical Statistics

主 编 万崇华 罗家洪



科学出版社

中国科学院教材建设专家委员会规划教材
全国高等医学院校规划教材

高级医学统计学

Advanced Medical Statistics

主 编 万崇华 罗家洪
主 审 孟 群 丁元林
副主编 王心旺 施学忠 王乐三 王 玫
参加编著人员(以姓氏笔画为序)

万崇华(广东医学院)	武淑琴(山西医科大学)
马 菲(美国 Rochester 大学)	罗家洪(昆明医科大学)
王心旺(广州医科大学)	周 旋(济宁医学院)
王乐三(中南大学)	郎建英(广州中医药大学)
王立芹(河北医科大学)	孟 琼(昆明医科大学)
王 玫(滨州医学院)	赵 倩(广州医科大学)
王济川(美国 George Washington 大学)	胡利人(广东医学院)
冯常勇(美国 Rochester 大学)	柳 青(中山大学)
闫宇翔(首都医科大学)	郜艳晖(广东药学院)
安胜利(南方医科大学)	施学忠(郑州大学)
孙红卫(滨州医学院)	夏英林(美国 Rochester 大学)
杨永利(郑州大学)	郭秀花(首都医科大学)
李运明(成都军区总医院)	屠心铭(美国 Rochester 大学)
李晓梅(昆明医科大学)	韩 煜(美国 Rochester 大学)
李晓翠(昆明医科大学)	谢海义(美国 Dartmouth 医学院)
何利平(昆明医科大学)	潘海燕(广东医学院)
张岩波(山西医科大学)	

学术秘书 孟 琼 于 磊

科 学 出 版 社

北 京

· 版权所有 侵权必究 ·

举报电话:010-64030229;010-64034315;13501151303(打假办)

内 容 简 介

本书由中、美两国中青年统计学专家共同撰写,主要介绍高级统计学知识,包含常用的多元统计方法和统计模型以及一些实用的专题统计。全书共22章,包括多变量方差分析、多重线性回归分析进阶、Logistic回归分析进阶、主成份分析与因子分析、聚类分析、判别分析、典型相关分析、对应分析、多维尺度分析、结构方程模型、多水平模型、线性混合效应模型、对数线性模型、广义线性模型、纵向(重复测量)资料分析、生存分析、Meta分析、多元分析的样本含量估计、量表测评常用统计方法、医学中的特殊实验设计及其分析、现场(市场)调查常用统计方法等内容。各章均按目的要求、内容概要、实例分析、思考练习、延伸阅读等安排。对一些易混淆的问题多以表格的方式进行对比和归纳,以图形的方式突出直观性;实例分析同时使用了SPSS和SAS两个软件包,给出分析思路、操作、结果、解释并适当引申推广。

本书可作为医学类研究生(硕士、博士)及统计或对统计要求高的本科专业(如卫生统计、生物统计、生物信息专业)参考书,也可以作为“多元统计分析”、“高级医学统计方法”等课程的教材或辅导用书。

图书在版编目(CIP)数据

高级医学统计学 / 万崇华, 罗家洪主编. —北京: 科学出版社, 2014. 4
中国科学院教材建设专家委员会规划教材·全国高等医学院校规划教材
ISBN 978-7-03-039754-6

I. 高… II. ①万… ②罗… III. 医学统计-医学院校-教材 IV. R195-1

中国版本图书馆CIP数据核字(2014)第025966号

责任编辑:李国红 朱 华 / 责任校对:李 影 桂伟利
责任印制:肖 兴 / 封面设计:范璧合

版权所有, 违者必究。未经本社许可, 数字图书馆不得使用

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

http://www.sciencep.com

双青印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2014年4月第 一 版 开本:787×1092 1/16

2014年4月第一次印刷 印张:32 1/2

字数:765 000

定价:108.00元

(如有印装质量问题, 我社负责调换)

主编副主编介绍

万崇华,1964年生,医学博士,医学/管理学双硕士,博士/博士后导师,1999年破格教授。曾任昆明医学院公共卫生学院副院长、云南省中青年学术技术带头人和云南省高等学校教学科研带头人、云南省政协委员和民进云南省委员会常务委员。现任广东医学院生命质量与应用心理研究中心主任、人文与管理学院院长。



国际生命质量研究会 (ISOQOL) 委员及亚洲华人分会副会长,中国信息学会医院统计专业委员会副主任委员、医学统计教育专业委员会常委、中华医学会临床流行病学学会委员、广东省卫生经济学会卫生经济政策专委会及卫生资源配置与绩效评价专委会副主任委员、广东省医学会行为与心身医学分会委员。国家自然科学基金评审专家。《中国卫生统计》《中国医院统计》等多个杂志编委。

主要从事流行病与卫生统计学、社会医学与卫生管理两个学科的教学科研工作,研究方向是生命质量与现代心理测评、卫生改革与医疗保险、生物信息挖掘与疾病诊疗。主持国家自然科学基金课题4项,国家973和科技支撑计划子课题各1项,省部级课题多项。以第一作者或通讯作者发表论文近200篇,其中英文SCI刊物20多篇。主持“癌症患者生命质量测量表体系”和“慢性病患者生命质量测量表体系”的研制,主著《生命质量测定与评价方法》《癌症患者生命质量测定与应用》《健康测量》和《卫生资源配置与区域卫生规划的理论与实践》等多部专著,主编(副主编)《卫生统计学学习辅导》《医学统计学学习辅导》《当代大学生心理健康教育》《中国医学统计百科全书——单变量推断统计分册》等多部教材或教学参考书。主持完成卫生统计学计算机辅助教学系统HSCAIS98的研制、考卷分析与教学评估系统EQATES的研制。获国家级教学成果二等奖1项,省教学成果一等奖和二等奖各1项,省科技进步奖三等奖2项。获国家版权证书(相当于专利)6项。

先后到澳大利亚和美国留学并到法国、意大利等二十多国访问考察。



罗家洪,男,佤族,1958年10月生,教授,硕士研究生导师,昆明医科大学公共卫生学院流行病学与卫生统计学系主任,骨干教师,云南省民进昆医支部副主任。中国卫生信息学会医学统计教育专业委员会常委,云南省预防医学会流行病学分会副主任委员,云南新型农村合作医疗专家组成员,中国卫生质量管理杂志编委,昆明医科大学学报统计审读员,卫生软科学特邀审稿专家,昆明市西山区政府西山区第一届应急管理专家组公共卫生事件类专家。

借鉴国外先进经验创新性主编了一系列案例式教材,为全国医学统计学和流行病学案例式教学提供了教学模式和教材。主持云南省卫生统计学精品课程建设、昆明医科大学卫生统计学精品课程建设。主编副主编主审教材专著17部,参编参译教材专著8部。近几年主持参与各种科研项目16项。获得云南省科学技术进步三等奖3个,获得中国卫生经济学会三等奖一个。获得云南省教学成果奖二等奖2个,云南省优秀教材1

部,自编优秀教材1部;获昆明伍达观奖教金二等奖和三等奖各一个,昆明医科大学教学成果一等奖2个、二等奖1个,优秀教材二等奖2个,优秀教材2部。总计撰写公开发表科研论文200多篇(其中SCI收录8篇,第一作者80多篇,通讯作者30多篇)。2005~2013年公开发表科研论文132篇(其中SCI 8篇;非SCI论文:第一作者37多篇;通讯作者34篇;第二及以后作者53篇)。

研究方向:新型农村合作医疗、生命质量、艾滋病、公共卫生服务。



王心旺,男,博士,教授,1958年3月出生。2001级中山大学公共卫生学院流行病与卫生统计学专业博士研究生。现为广州医科大学流行病与卫生统计学教授、硕士生导师、广州医科大学公共卫生学院应用统计学专业学术带头人,中国卫生信息学会医学统计专业委员会常委,《中华医学研究杂志》《中华医学教育杂志》《中华医学教育探索》和《西北医学教育》杂志编委。现任广州医科大学继续教育学院副院长、广东省全科医学教育培训中心副主任、广州市慢性病防控与管理学会副会长、广州市医学会全科医学专委会主任委员。

主要研究方向为疾病负担测量与人群健康效应研究、医疗保险精算方案研究,先后主持多项省市医学科研课题,现今主持广东省自然科学基金项目“COPD经济负担测量及对人群健康总体效应的量化研究”和广东省科技厅科技计划研究项目“广州市新农合医疗基金风险管理研究”。

近年来在 *Medical Care*《中华医学杂志(英文版)》《中国卫生统计》《中国自然医学》《中国全科医学》和《中华医学教育杂志》等医学杂志发表论文50多篇,作为副主编参与编写出版《卫生事业管理》和《医学科研设计》教学用书,参编《卫生统计学(第7版)》《医疗保险学》《中国卫生统计方法研究进展》等教材与专著。

施学忠,博士,教授,郑州大学卫生统计学教研室主任,郑州大学学报(医学版)编辑部主任, *Asia Pacific Journal of Clinical Nutrition* 编辑部主任,中国高校科技期刊研究会理事、医学期刊专业委员会副主任委员,中国卫生信息学会医学统计教育专业委员会委员。近15年来主持国家“十五”“十一五”和“十二五”科技重大专项及河南省自然科学基金与科技攻关项目等共16项。获河南省科技进步二等奖2项,编写著作15部,发表论文120余篇。研究方向:卫生服务统计方法与应用。



王乐三,男,1963年10月生于湖南省邵阳市。博士,副教授,硕士生导师。现任中南大学公共卫生学院流行病与卫生统计系副主任。国家精品课程《医学(卫生)统计学》和国家精品资源共享课《医学(卫生)统计学》主讲教师,美国华盛顿大学公共卫生学院生物统计系访问学者。从事教学工作至今近30年,主要研究领域为流行病与卫生统计学方法及其医学应用,综合评价方法及其医学应用。先后参与多项国家级研究项目,获省部级教学科研成果奖3项,参加国家规划教材编写10余部。主要社会兼职:湖南省卫生统计专业委员会副主任委员,中国卫生信息学会统计理论与方法专业委员会委员,中国卫生信息学会卫生统计学教育专业委员会委员。



王玖,男,1971年8月生,山东高青县人,博士,滨州医学院卫生统计学教研室主任,滨州医学院学科带头人与学术骨干。中国卫生信息学会卫生统计教育委员会理事、统计理论与方法专业委员会理事,中国现场统计研究会统计综合评价研究分会理事,山东省应用统计学会常务理事。主讲统计学本科专业《卫生统计学》《统计软件》等,研究生及其他专业《医学统计学》等。

研究方向为临床数据分析、卫生管理统计,主持完成省、厅局级课题5项,主持在研省部级课题3项、厅级1项,发表论文20余篇;获第六届滨州市青年科技奖,山东软科学优秀成果奖多项,烟台市、滨州市社会科学优秀成果奖多项,第四届调研山东大学生社会调查省优秀指导教师。



其他作者简介(按姓氏笔画排列)

马菲,女,于2009年获上海财经大学统计学学士学位,现在美国罗切斯特大学攻读统计学博士学位。毕业论文主攻高维混合模型的参数估计,以及其在流式细胞仪中的应用。她的研究领域广泛,涉及半参数估计,生存分析,纵向数据分析,聚类分析以及各种统计方法在临床医学的应用。

王立芹,女,汉族,1966年8月出生,河北省吴桥县人,博士,河北医科大学公共卫生学院流行病学与卫生统计学教授,硕士研究生导师。1985年就读天津医科大学卫生系,1990年获学士学位,2004年获河北医科大学流行病学与卫生统计学硕士学位,2010年河北医科大学流行病学与卫生统计学博士毕业。长期从事医学统计学与卫生统计学的教学与科研工作。承担参与多项国家级省级课题。近几年发表的论文40余篇,参编多部统计学教材。

王济川,男,1947年出生,1982年四川大学经济系毕业,1986年获美国康奈尔大学社会学硕士学位,1990年获该校社会学博士学位,1989年9月至1990年8月于美国密西根大学人口中心作博士后研究。1991年9月任职美国俄亥俄州莱特州立大学医学院社区卫生系。2000年7月至2008年5月任该系教授。2008年6月至今任乔治·华盛顿大学医学院流行病学及生物统计系教授。主要研究领域为社会科学定量分析方法、多层统计分析模型、结构方程模型及公共卫生和疾病预防研究。

冯常勇,男,1991年获中国科学技术大学运筹学学士学位,2002年获美国罗切斯特大学(University of Rochester)数理统计学博士学位。现为罗切斯特大学医学中心(University of Rochester Medical Center)生物统计与计算生物学系(Department of Biostatistics and Computational Biology)副教授。主要研究方向为生存分析(Survival Analysis)和统计方法在流行病学与临床实验中的应用。现已发表论文94篇(统计理论文章28篇,医学文章66篇)。指导统计学博士生(在读)一名。

孙红卫,女,1980年7月生,山东莱州人,2006年7月毕业于四川大学,获得概率论与数理统计专业硕士学位,2006年于滨州医学院卫生统计学教研室工作至今。担任《多元统计方法》《概率论》《数理统计》《医学统计学》的教学工作。主要研究领域为“综合评价理论及应用”“多元统计方法原理及应用”,主持山东省统计科研重点课题3项,校级课题1项。参与多项省部级、厅局级课题多项。发表论文10余篇,其中核心期刊5篇,英文文章2篇。

安胜利,男,副教授。第一军医大学卫生统计学硕士,南方医科大学卫生统计学博士。

2001 年至今在南方医科大学公共卫生与热带医学学院生物统计学系任教。副主编教材 3 部,参编教材 6 部,发表科研论文 40 余篇。

杨永利,女,博士,郑州大学公共卫生学院卫生统计学教研室副教授,硕士生导师。河南省中西医结合循证医学专业委员会副主任委员,河南省应用统计学学会理事,郑州市预防医学会理事。主持或参与国家“十二五”科技重大专项、国家“十一五”科技重大专项以及国家自然科学基金等项目。获国家发明专利 5 项,厅级科研成果 10 项,编写著作 6 部,发表论文 50 篇。研究方向:卫生服务统计方法与应用。

李运明,男,1982 年生,江苏徐州人,第四军医大学流行病与卫生统计学专业生物统计方向硕士、卫生信息管理方向博士,成都军区总医院博士后工作站临床医学专业(在站)博士后,医务部助理员,主管技师。主要研究方向:创伤性颅脑外伤流行病学调查、健康风险评估模型、重复测量资料统计分析方法、混合效应模型理论及应用、微阵列数据分析、医院信息管理。承担中国博士后基金项目 1 项、四川省卫生厅科研项目 1 项,参与多项国家自然科学基金项目研究;参编专著 3 部;发表中文学术论文 90 余篇其中第一作者中文学术论文 20 篇,SCI 论文 28 篇(含第一及共同第一作者 7 篇);2012 年被成都军区总医院遴选为首批研究型人才,荣立三等功一次。

李晓梅,女,1984 年毕业于华西医科大学卫生专业,2000 年获泰国 Mahidol 大学健康社会学硕士学位。现昆明医科大学公共卫生学院流行病学与卫生统计学系教授,硕士生导师。

2000 年以来,主持及参与国家自然科学基金项目 3 项、国家软科学项目 1 项、云南省自然科学基金 3 项、其他项目 10 余项,发表论文 50 余篇,参编书籍 5 部。参加人民卫生出版社规划教材《卫生统计学》第 6、第 7 版编撰,参编科学出版社《医学统计学》(案例版)第 1、第 2 版及云南民族出版社《卫生统计学学习辅导》和《医学统计学学习辅导》等多部教材及教辅书籍。

闫宇翔,女,1973 年 9 月生,首都医科大学公共卫生学院流行病与卫生统计学系副教授。先后获得“北京市中青年骨干教师培养计划”与“北京市属高等学校青年拔尖人才培养计划”资助。目前主持国家自然科学基金和北京市自然科学基金课题各 1 项,发表科研论文 20 余篇,参编教材 6 部。

何利平,女,汉族,1976 年 11 月生,博士,副教授,籍贯云南宣威。从本科到博士专业都是概率论与数理统计,1997 年获云南大学经济学学士学位;2000 年获云南大学理学硕士学位;2007 年获云南大学概率论与数理统计专业理学博士学位。2000 年至今在昆明医科大学公共卫生学院流行病与卫生统计学系任教,学生层次包括留学生、博士、硕士和本科生,教学课程包括多元统计分析 with SPSS 软件包、医学统计学、概率论与数理统计等。主要研究方向:卫生服务公平性研究,艾滋病研究。发表学术论文 43 篇,其中第 1 作者 10 篇。参与 3 部教材和 1 部医学统计百科全书的分册的编写工作,参与 1 部专著的译编工作。

张岩波,男,流行病与卫生统计学博士,教授,博士生导师。研究方向为生物信息与健康评价方法研究。先后主持承担国家自然科学基金 4 项,山西省科技创新平台项目 1 项,山西省自然科学基金 2 项,国际合作项目 1 项,参加其他多项省部项目。主编《潜变量分析》(高等教育出版社,2009)与《医学与人文》(当代中国出版社,2004);副主编《生物医学研究中的统计方法》(高等教育出版社,2007),并参加了多部教材的编写;发表论文 50 余篇。

郎建英,女,流行病与卫生统计学硕士,针灸推拿学博士。广州中医药大学基础医学院预防医学教研室副教授。从事医学统计学和流行病学教学和科研工作 31 年。主持或参与科研课题 21 项,发表或参与发表教学或科研论文 33 篇。

孟琼,女,医学硕士,副教授,1978 年出生,现就职于昆明医科大学公共卫生学院。

主要研究方向为生命质量研究、概化理论的应用、区域卫生规划。近五年来主持昆明医科大学教研教改课题 2 项,云南省教育厅项目 1 项,参加国家自然科学基金资助项目 1 项,国家软科学基金资助项目 1 项,云南省自然科学基金资助项目 1 项,云南省卫生厅资助项目 1 项。共参加撰写专著 4 部,共发表文章 52 篇,其中第一作者 8 篇(1 篇为 SCI 收录的共享作者),参与 5 本教材编写(其中 1 本为副主编)。获昆明医科大学青年教师讲课比赛三等奖 2 次,云南省多媒体软件大赛一等奖 1 项(排名第四),获云南省科技进步奖三等奖 1 项(排名第七),获国家版权证书 1 个(排名第六)。

柳青,男,肿瘤流行病学教授、博士研究生导师,1982 年毕业于广东医药学院预防医学专业,获医学学士学位。1985 年毕业于中山医科大学卫生统计学专业,获硕士学位。1996 年于中山医科大学在职攻读博士学位毕业,获卫生统计专业医学博士。1990 年至 1991 年在法国里昂国际癌症研究中心分析流行病学部门进修一年。自工作以来,一直从事肿瘤流行病学及有关统计分析方法的研究。承担过多项新药临床试验的统计咨询和统计分析报告工作。参加编写了教育部和卫生部的预防专业本科生“卫生统计学”、8 年制医学生“医学统计学”、医学硕士研究生“医学统计学”和“医学统计学和电脑实验”等推荐教材和规划教材的编写,参与和主持了美国 NIH 的 R01 鼻咽癌病因研究项目、国家“八五”“九五”“十五”和“十一五”鼻咽癌病因筛查方案和效果评价的重大攻关项目,广东省 SARS 防治重点项目和广州市卫生局社区肿瘤防治重大项目,在国内外专业杂志发表论文多篇,曾获得广东省教育成果一等奖,广东省科技进步二等奖、三等奖。

胡利人,副教授、硕士生导师,广东医学院流行病与卫生统计学教研室副主任。先后主持完成中国成人教育协会“十二五”成人教育科研规划课题、广东省成人教育协会“十二五”成人教育科研规划课题重点课题 1 项,湛江市科技攻关项目 1 项。目前主持广东省教育科研“十二五”规划课题 1 项,湛江市科技攻关项目 2 项,主要参与国家自然科学基金项目 1 项(第 3)、中华医学会研究项目 1 项(第 2)。第一作者或通讯作者发表专业学术论文 50 多篇。全国高等院校医学实验教学规划教材《卫生统计学实习指导》副主编(2010),全国高等院校医学实验教学规划教材《流行病学学习指导》副主编(2011),全国高等医药院校规划教材案例版《卫生统计学》编委(2008)。

赵倩,女,博士,毕业于中山大学流行病与卫生统计学专业,现为广州医科大学公共卫生学院统计系教师。2007 年获国家留学基金委奖学金资助,赴美国加州大学旧金山分校(UCSF)进行博士联合培养,主持国家自然科学基金和广东省自然科学基金各一项,目前的研究方向为脑肿瘤影像分析的统计学方法。

郜艳晖,女,1973 年 9 月出生,山西长治人,广东药学院公共卫生学院流行病与卫生统计学系教授,硕士研究生导师。先后获得山西医科大学卫生统计学硕士学位和复旦大学流行病与卫生统计学博士学位,曾在加拿大渥太华大学社会医学与流行病学系进行博士后研究。研究方向为遗传流行病学统计方法,主持国家自然科学基金、广东省自然科学基金等项目,发表论文 30 余篇,作为副主编及编委参与编写教材 10 余部。

夏英林,于1999年获中山大学哲学博士学位,2006年获美国德保罗大学统计与应用数学硕士学位,2011年获罗切斯特大学医学统计学硕士学位。现为美国罗切斯特大学生物统计系研究助理教授,科研经验丰富,尤其擅长于多级建模、结构方程建模、潜在成长曲线模型、增长混合模型、中值分析、零膨胀计数建模、因子分析、聚类分析、主成份分析、心理理论、保真度研究、方程反应理论、非劣效性试验。他在各类学术期刊发表统计方法和应用研究论文60余篇。研究领域涉及产后抑郁症、抑郁症、家庭暴力、酗酒、青少年暴力、艾滋病毒的风险预防干预、睡眠障碍、疼痛、癌症、剂量反应关系,炎症和感染。

郭秀花,首都医科大学公共卫生学院副院长,流行病与卫生统计学系主任、教授、博士生导师,首都医科大学循证医学联合教研室副主任。目前主持国家科技部重大传染病“十二五”专项课题2项、国家科技部“十二五”支撑项目课题1项、“973”分课题2项、国家自然科学基金2项、北京市自然科学基金课题2项、北京市科委重大项目1项;已发表科研论文100多篇,其中SCI论文近20篇;主编教材9部,其中1部为北京市精品教材;获批国家专利3项;获批省部级科技进步三等奖7项;荣获北京市学术创新团队带头人;兼任中国现场统计生物统计专业委员会副理事长、中国卫生信息学会卫生统计理论与方法专业委员会常务委员;被聘为国际杂志 *Global Health Promotion* 编委、《中国预防医学杂志》等刊物编委,并为 *Plos One*、*International Journal for Quality in Health Care*、*International Journal of Infectious Diseases*、《北京大学学报(医学版)》等近10种杂志审稿。

屠心铭,于1982年获得复旦大学数学学士学位,后赴美攻读博士学位,于1989年获杜克大学统计学与应用数学博士学位。现为罗切斯特大学生物统计系教授、副系主任,心理健康统计部主任和生物统计系咨询中心主任。有20多年的生物统计学和心理学的研究经验。除参与众多书籍章节撰写和在各类学术期刊发表170多篇统计方法等科学论文外,撰写专著2本,主编书籍2部,其中 *Applied Categorical and Count Data Analysis*、*Modern Applied U-statistics* 深受同行好评。研究兴趣广泛,其中在纵向数据分析、结构方程模型、因果关系分析、自由分布模型、潜在成长混合模型、方程反应模型以及它们在一系列的观察性研究和随机对照中的应用颇有建树。

韩煜,毕业于南京大学数学系,现为美国罗切斯特大学(University of Rochester)生物统计与计算生物系(Department of Biostatistics and Computational Biology)在读统计学博士研究生。研究领域包括生存分析和统计临床试验分析,现已发表论文4篇。

谢海义,1982年兰州大学经济系毕业,1986年获美国乔治·华盛顿大学统计学硕士学位,1994年获美国犹他大学社会学博士学位(主修定量方法)。曾供职中国国家统计局。1994年起供职美国达特茅斯学院(Dartmouth College)医学院,现为该院家庭及社区卫生系副教授。主要研究领域为纵向统计分析方法及其在医学、公共卫生学及卫生经济学中的应用。

潘海燕,女,山东青岛人,1982年2月出生。广东医学院流行病与卫生统计学教研室副教授,博士。从事“慢性病患者生存质量量表、生存质量影响因素及流行病统计方法学”的研究工作。近五年来,参与编写教材2部,主持获得5项市厅级科研项目和4项校级科研项目的资助。发表学术论文20余篇,其中第一作者发表12篇,参与发表SCI论文5篇。2010年和2011年有两篇论文分别获东莞市预防医学会优秀论文一等奖。因教学、科研工作认真扎实,2012年被评为“广东省千百十工程”校级学术培养人。

前 言

统计学是研究数据的收集、整理、分析、解释和表达的原理和方法的一门综合学科,目前已经形成许多分支学科,如生物统计学、社会统计学、医学统计学等。《管子》中曾有过这样的阐述“不明数欲举大事,如船之无橹而欲行于大海也”。意思是说在不清楚相关数据的情况下,想做大事,无疑像没有橹的船航行于汪洋大海之中一样。可见,古代先贤早已知道了统计的重要性。无论是“结绳记事”,还是土地丈量、国势描述等无不体现着统计的存在。而如今什么都能“数字化”的时代更是如此,统计“无处不在,无时不有”。我们每天看新闻、读报纸会看到大量的统计数据,如全球到2020年将会有100亿人口、道琼斯指数又涨了30点、我国人均GDP比去年同期增长了7.7%,等等。

随着计算机网络的发展,人类在各个领域所涉猎的深度和广度都已今非昔比,需要处理的问题千头万绪,筹划决策、科学管理等理论和方法备受瞩目。这些理论和方法的基础就在于正确地分析各种信息和数据,为充分合理地利用各种资源提供科学依据。作为卫生信息处理与分析的理论和技艺的医学统计学也就成为不可或缺的重要工具。其作用与地位与日俱增,先后被列为医科院校本专科生和硕士甚至博士研究生的必修课程。对于医学类研究生或对统计要求高的一些本科专业的学生而言,在学习了一些基本统计学知识和原理后,必须有一些高级统计学知识来帮助进行科学研究。为此,不少院校在研究生中纷纷开设了“多元统计分析”“高级医学统计学”等课程。为便于广大科技工作者进一步学习统计学,我们集多年的教学经验编著了这本《高级医学统计学》。可作为医学类研究生(硕士/博士)及统计或对统计要求高的本科专业(如卫生统计/生物统计/生物信息专业)参考书,也可以作为“多元统计分析”“高级医学统计方法”“现场调查研究分析方法”等课程的教材或辅导用书。

本书力图体现以下特色:

1. 兼顾专著与教材性质。本书具有一定的深度和广度,知识性强,体现了认知与学术水平等专著的性质;同时也具备教材的特点,突出“深入浅出、循序渐进”,将案例式教学及参与式教学融入了本书中,是编著者多年来教学经验的总结及教学改革成果体现。

2. 内容丰富,实用性强。本书主要介绍高级统计学知识,包含了常用的多元统计方法和统计模型,还介绍了一些实用的专题统计。各章均按目的要求、内容概要、实例分析、思考练习、延伸阅读等安排。目的要求指明了学习的方向和重点;内容概要则简要地概括了应掌握的基本知识,尤其对一些易混淆的问题多以表格的方式进行对比和归纳,以图形的方式突出直观性;实例分析给出分析思路、操作、结果、解释并适当引申推广,便于学生举一反三的学习;思考练习部分,给出少量思考题和一定数量的应用分析题,供学生上机练习用,通过练习帮助学生掌握基本概念和方法。延伸阅读给出了大量被引用或相关的文献,供学生进一步查阅和学习。

3. 核心与特异模块结合。本书的编撰思路采用的是核心与特异模块结合的方式,核心模块包括了常用的多元统计方法和常用统计模型如结构方程模型、多水平模型、广义线性模型等。特异模块主要是一些专题统计,各高校结合自身科研采用的一些特色高级统计方法,如量表测评中的统计方法。

4. 操作以菜单点菜式为主,并兼顾常用的统计软件包。本书另一特色就是对实例分析

同时使用了SPSS和SAS两个软件包,每个方法均以其中一个软件的菜单操作为主,同时简单介绍另一个软件的不同之处。不能通过菜单实现的方法,也给出了相应的程序(如SAS程序),读者可以根据自己的喜好及自己已经具备的基础来进行选择。

本书编撰者是中、美两国从事医学统计工作多年的中青年学者,全部是统计学博士或高级职称。本书既是其教学经验的总结,也是其相关科研的反映。尽管如此,限于水平,不足之处敬请同行专家读者不吝赐教!

在本书编撰和出版过程中,我国著名的统计学专家方积乾教授、金水高教授、孙振球教授等给予了诸多的帮助指导。广东医学院江文富书记、郑学宝院长、杨云滨副院长,昆明医科大学姜润生校长、李燕副校长,东莞市石龙博爱医院黄新萍以及本书的责任编辑给予了大力的支持和帮助。卫生部信息统计中心主任孟群教授和广东医学院副院长丁元林教授在百忙中审阅了全书并提出了许多宝贵的修改意见。在此一并致谢!

万崇华 于广东医学院

罗家洪 于昆明医科大学

2013年12月

目 录

前言

第 1 章 绪论 Introduction	(1)
1.1 多元统计学概况 Overview of Multivariate Statistics	(1)
1.2 多元统计方法分类与选择 Classification and Selection of Multivariate Analysis	(3)
1.3 多元数据描述 Description of Multivariate Data	(7)
1.4 专题统计概述 Outlines of Specific Statistical Methods	(11)
思考练习 Exercises	(14)
延伸阅读 Further Readings	(14)
第 2 章 多变量方差分析 Multivariate Analysis of Variance	(16)
2.1 多变量方差分析概况 Overview of Multivariate Analysis of Variance	(16)
2.2 单样本多变量方差分析 One Sample Multivariate Analysis of Variance	(18)
2.3 单因素多变量方差分析 One-way Multivariate Analysis of Variance	(18)
2.4 多因素多变量方差分析 Multi-factor Multivariate Analysis of Variance	(19)
2.5 含协变量的多变量方差分析 Multivariate Analysis of Variance with Covariates	(20)
2.6 实例分析 Examples Analysis	(21)
思考练习 Exercises	(39)
延伸阅读 Further Readings	(42)
第 3 章 多重线性回归分析进阶 Advanced Multiple Linear Regression Analysis	(44)
3.1 多重线性回归回顾 Review of Multiple Linear Regression	(44)
3.2 加权最小二乘法——方差不齐的处理 Weighted Least Squares-Treatment for Heterogeneity of Variance	(45)
3.3 岭回归——共线性的处理 Ridge Regression-Treatment for Collinearity	(47)
3.4 最优尺度回归——分类变量的数值化 Optimal Scaling Regression-Quantifying Category Variables	(50)
3.5 两阶段最小二乘回归——因果模型构建 Two-stage Least Squares Regression- Construction of Causal Model	(51)
3.6 实例分析 Examples Analysis	(52)
思考练习 Exercises	(65)
延伸阅读 Further Readings	(66)
第 4 章 Logistic 回归分析进阶 Advanced Logistic Regression	(68)
4.1 Logistic 回归回顾 Logistic Regression Review	(68)
4.2 Logistic 回归诊断 Logistic Regression Diagnosis	(71)
4.3 无序多分类 Logistic 回归 Multinomial Logistic Regression	(74)

4.4	有序多分类 Logistic 回归 Ordinal Logistic Regression	(75)
4.5	条件 Logistic 回归 Conditional Logistic Regression	(76)
4.6	实例分析 Examples Analysis	(79)
	思考练习 Exercises	(88)
	延伸阅读 Further Readings	(90)
第 5 章	主成份分析与因子分析 Principal Component Analysis and Factor Analysis	
		(91)
5.1	主成份分析 Principal Component Analysis	(91)
5.2	因子分析 Factor Analysis	(96)
5.3	主成份分析与因子分析的关系 The Relationship between Principal Component Analysis and Factor Analysis	(101)
5.4	实例分析 Examples Analysis	(103)
	思考练习 Exercises	(114)
	延伸阅读 Further Readings	(116)
第 6 章	聚类分析 Cluster Analysis	(119)
6.1	聚类分析概况 Overview of Cluster Analysis	(119)
6.2	系统聚类法 Hierarchical Clustering Method	(122)
6.3	快速聚类法 Faster Clustering	(123)
6.4	其他聚类方法 Other Clustering Methods	(124)
6.5	聚类方法的选择 Sections of Clustering Methods	(125)
6.6	聚类分析注意事项 Some Notes of Cluster Analysis	(126)
6.7	实例分析 Examples Analysis	(127)
	思考练习 Exercises	(134)
	延伸阅读 Further Readings	(136)
第 7 章	判别分析 Discriminant Analysis	(138)
7.1	判别分析概况 Overview of Discriminant Analysis	(138)
7.2	距离判别分析 Distance Discriminant Analysis	(140)
7.3	Fisher 判别分析 Fisher Discriminant Analysis	(141)
7.4	Bayes 判别分析 Bayes Discriminant Analysis	(142)
7.5	分类资料判别分析(Bayes 公式法) Discriminant Analysis for Qualitative Data (Bayes Formula Method)	(143)
7.6	逐步判别分析 Stepwise Discriminant Analysis	(144)
7.7	聚类分析与判别分析的关系 The Relationship between Cluster Analysis and Discriminant Analysis	(145)
7.8	实例分析 Examples Analysis	(145)
	思考练习 Exercises	(164)
	延伸阅读 Further Readings	(168)
第 8 章	典型相关分析 Canonical Correlation Analysis	(170)
8.1	典型相关分析概况 Overview of Canonical Correlation Analysis	(170)
8.2	典型相关分析的统计思想 Statistical ideology of Canonical Correlation Analysis	

.....	(170)
8.3 典型相关分析的基本理论与模型假定 The Basics and Model Assumption of Canonical Correlation Analysis	(171)
8.4 典型相关分析的基本步骤 Basic Steps in Canonical Correlation Analysis	(172)
8.5 典型变量的性质及其意义解释 The Nature and Meaning of Canonical Variables	(173)
8.6 实例分析 Examples Analysis	(174)
思考练习 Exercises	(184)
延伸阅读 Further Readings	(186)
第9章 对应分析 Correspondence Analysis	(187)
9.1 概念与计算 The Concept and Calculation	(187)
9.2 对应分析中的假设检验问题 Hypothesis Testing Problems in Correspondence Analysis	(189)
9.3 多重对应分析 Multiple Correspondence Analysis	(189)
9.4 对应分析用于定量变量的情况 Correspondence Analysis for Quantitative Variables	(190)
9.5 需要注意的问题 Some Notes on Correspondence Analysis	(190)
9.6 实例分析 Examples Analysis	(190)
思考练习 Exercises	(196)
延伸阅读 Further Readings	(198)
第10章 多维尺度分析 Multi-Dimensional Scaling	(199)
10.1 多维尺度分析概述 Overview of Multidimensional Scaling	(199)
10.2 多维尺度分析原理 The Principle of Multidimensional Scaling	(201)
10.3 多维尺度分析步骤 Steps in Multidimensional Scaling	(205)
10.4 实例分析 Examples Analysis	(207)
思考练习 Exercises	(216)
延伸阅读 Further Readings	(217)
第11章 结构方程模型 Structural Equation Modeling	(218)
11.1 结构方程模型概况 Overview of Structural Equation Modeling	(218)
11.2 结构方程模型思想 The Basic Ideas of Structural Equation Modeling	(221)
11.3 结构方程模型分析步骤 Steps of the Structural Equation Modeling	(223)
11.4 均值结构模型 The Mean Structure Model	(227)
11.5 实例分析 Examples Analysis	(230)
思考练习 Exercises	(232)
延伸阅读 Further Readings	(233)
第12章 多水平模型 Multilevel Models	(235)
12.1 多水平模型的优点 Advantages of Multievel Models	(235)
12.2 多水平模型的基本公式表述、参数估计、模型的评估及假设检验 Basic Multiple Level Model Formulation, Parameter Estimation, Model Fit Evaluation and Hypothesis Testing	(238)

12.3	多水平模型的构建步骤	Steps in Building Multilevel Model	(243)
12.4	实例分析	Examples Analysis	(246)
	思考练习	Exercises	(262)
	延伸阅读	Further Readings	(262)
第13章	线性混合效应模型	Linear Mixed Effects Model	(265)
13.1	线性混合效应模型的数据类型及应用条件	Data Types and Conditions of Linear Mixed Effects Model	(265)
13.2	线性混合效应模型结构	The Structure of Linear Mixed Effects Model	(267)
13.3	参数估计和假设检验	Parameter Estimation and Hypothesis Testing	(268)
13.4	实例分析	Examples Analysis	(270)
	思考练习	Exercises	(278)
	延伸阅读	Further Readings	(279)
第14章	对数线性模型	Log-linear Model	(281)
14.1	对数线性模型概况	Overview of Log-linear Model	(281)
14.2	一般对数线性模型	General Log-linear Model	(282)
14.3	Logit 对数线性模型	Logit Log-linear Model	(284)
14.4	实例分析	Examples Analysis	(286)
	思考练习	Exercises	(300)
	延伸阅读	Further Readings	(301)
第15章	广义线性模型	Generalized Linear Models	(302)
15.1	广义线性模型概况	Overview of Generalized Linear Models	(302)
15.2	二分类数据的广义线性模型	Generalized Linear Models for Binary Data	(305)
15.3	多分类数据的广义线性模型	Generalized Linear Models for Polytomous Data	(306)
15.4	Poisson 回归模型	Poisson Regression Models	(308)
15.5	实例分析	Examples Analysis	(309)
	思考练习	Exercises	(317)
	延伸阅读	Further Readings	(318)
第16章	纵向(重复测量)资料分析	Longitudinal(repeated measure) Data Analysis	(319)
16.1	纵向(重复测量)资料概况	Overview of Longitudinal(repeated measure) Data Analysis	(319)
16.2	纵向(重复测量)资料的方差分析	Longitudinal(repeated measure) Data Analysis of Variance	(323)
16.3	广义估计方程模型	Generalized Estimating Equations	(327)
16.4	潜变量增长曲线模型	Latent Growth Curve Model	(329)
16.5	时间序列分析简介	Introduction to Time Series Analysis	(331)
16.6	实例分析	Example Analysis	(334)
	思考练习	Exercises	(368)
	延伸阅读	Further Readings	(369)

第 17 章 生存分析 Survival Analysis	(372)
17.1 生存分析概况 Survival Analysis Overview	(372)
17.2 生存分析的一些基本概念 Basic Concepts of Survival Analysis	(373)
17.3 尼尔森-阿兰 (Nelson-Aalen) 累积危险率估计 Nelson-Aalen Estimator of Cumulative Hazard Function	(375)
17.4 生存函数的 Kaplan-Meier 估计 Kaplan-Meier Estimation of Survival Function	(378)
17.5 Log-rank 检验 Log-rank test	(379)
17.6 Cox 比例危险率模型 Cox Proportional Hazards Model	(380)
17.7 实例分析 Example Analysis	(383)
思考练习 Exercises	(388)
延伸阅读 Further Readings	(390)
第 18 章 Meta 分析 Meta Analysis	(391)
18.1 Meta 分析概况 Overview of Meta Analysis	(391)
18.2 定量资料的 Meta 分析 Meta Analysis for Quantitative Data	(398)
18.3 定性资料的 Meta 分析 Meta Analysis for Qualitative Data	(400)
18.4 Meta 分析的偏倚 The Bias of Meta Analysis	(403)
18.5 实例分析与 RevMan 软件 Examples Analysis and RevMan Software	(405)
思考练习 Exercises	(410)
延伸阅读 Further Readings	(410)
第 19 章 多元分析的样本含量估计 Sample Size Estimation of Multivariate Analysis	(412)
19.1 样本量估计的主要参数和其他影响因素 Main Parameters and Other Influence Factors of Sample Size Estimation	(412)
19.2 基本的样本量估计 Basic Sample Size Estimation	(413)
19.3 方差膨胀因子的基本校正 The Corrections on Variance Inflation Factor ...	(415)
19.4 多因素回归分析的方差膨胀因子校正 The Variance Inflation Factor in Multiple Regression Analysis	(418)
19.5 实例分析 Examples Analysis	(421)
思考练习 Exercises	(423)
延伸阅读 Further Readings	(423)
第 20 章 量表测评常用统计方法 Statistical Methods Used in Measurements and Assessments of Scales	(425)
20.1 量表研制概况 Overview of the Scale Development	(425)
20.2 信度分析常用方法 Methods Commonly Used in Reliability Analysis	(427)
20.3 效度分析常用方法 Methods Commonly Used in Validity Analysis	(429)
20.4 反应度分析常用方法 Methods Commonly Used in Responsiveness Analysis	(430)
20.5 量表资料的统计分析 Statistical Analysis for Scale Data	(431)
20.6 实例分析 Examples Analysis	(432)

思考练习 Exercises	(434)
延伸阅读 Further Readings	(435)
第 21 章 医学中的特殊实验设计及其分析 Special Experiment Designs and Analysis in Medicine	(436)
21.1 不完全区组设计 Incomplete Block Design	(436)
21.2 嵌套设计 Nested Design	(441)
21.3 序贯设计 Sequential Design	(443)
21.4 响应曲面设计 Response Surface Design	(452)
21.5 实例分析 Example analysis	(456)
思考练习 Exercises	(459)
延伸阅读 Further Readings	(460)
第 22 章 现场(市场)调查常用统计方法 The Methods Commonly Used in Field (Market) Survey	(461)
22.1 特殊的调查设计 Special Survey Designs	(461)
22.2 现场(市场)调查中的轮廓分析 The Profile Analysis in Field (Market) Survey	(473)
22.3 现场(市场)调查中的多维列联表分析 Multidimensional Contingency Table Analysis in Field (Market) Survey	(477)
思考练习 Exercises	(479)
延伸阅读 Further Readings	(480)
附录 1 基础统计学方法概要	(482)
1.1 基础统计学方法概要	(482)
1.2 常用基础统计方法选择	(485)
附录 2 SAS 统计软件包简介	(488)
2.1 概述	(488)
2.2 SAS 的安装与启动	(489)
2.3 SAS 数据管理	(490)
2.4 常用统计分析	(491)
附录 3 SPSS 统计软件包简介	(494)
3.1 SPSS 概述	(494)
3.2 SPSS 的安装与启动	(494)
3.3 SPSS 的菜单	(495)
3.4 SPSS 的数据输入与保存(数据准备)	(501)
3.5 t 检验	(502)
3.6 其他检验	(503)