



普通高等教育“十一五”国家级规划教材

全国高等学校医学规划教材

(供临床·基础·预防·护理·检验·口腔·药学等专业用)

医学统计学

第2版

主编 李晓松



高等教育出版社

Higher Education Press



普通高等教育“十一五”国家级规划教材

全国高等学校医学规划教材

医学统计学 第2版

(供临床·基础·预防·护理·检验·口腔·药学等专业用)

主 编 李晓松

副主编 郝元涛 张菊英 王 彤 仇小强

编 委 (按章节顺序排列)

李晓松	四川大学	杨土保	中南大学
张菊英	四川大学	李新华	贵阳医学院
贾 红	泸州医学院	薛 茜	新疆医科大学
郝元涛	中山大学	张福良	大连医科大学
刘美娜	哈尔滨医科大学	尹 平	华中科技大学
王 彤	山西医科大学	丁元林	广东医学院
薛付忠	山东大学	仇小强	广西医科大学
杨兴华	首都医科大学	毛宗福	武汉大学
张 强	四川大学	孙 高	中国医科大学
曾 庆	重庆医科大学	朱彩蓉	四川大学

学 术 秘 书

沈卓之 四川大学

曾芳芳 中山大学



高等教育出版社
Higher Education Press

内容简介

本教材着力突出对学生开展医学科学研究、进行研究设计、运用统计方法分析资料、运用统计软件实现数据管理和分析、正确解释和表达统计分析结果的综合能力培养。在文字上力求通俗易懂,在公式上进一步淡化推导及计算过程。整个教材力图达成定位明确、内容优化、重点突出、循序渐进、便于自学的编写特色与风格。

本教材共 25 章,内容涵盖了国家执业医师资格考试的全部知识点,新增了统计方法选择与结果解释、统计分析结果的正确表达、数据管理及 EpiData 软件简介、SPSS 软件简介 4 章内容。配有《医学统计学实习指导》。适合于临床医学、基础医学、预防医学、护理、检验、口腔及药学等专业各年制本科生使用,同时也可供研究生、临床医师、公共卫生专业人员及相关科研工作者参考使用。

此外,为方便教与学,编者还另外制作了与教材内容相关的光盘,内容包括 CAI 课件、例题和习题数据。

图书在版编目(CIP)数据

医学统计学/李晓松主编. —2 版. —北京:高等教育出版社,2008. 12

供临床、基础、预防、护理、检验、口腔、药学等专业用
ISBN 978 - 7 - 04 - 025525 - 6

I. 医… II. 李… 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(2008)第 188886 号

策划编辑 冯娟 席雁 责任编辑 孙葵葵 封面设计 张楠 责任绘图 黄建英
版式设计 张岚 责任校对 金辉 责任印制 朱学忠

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 人民教育出版社印刷厂

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>

开 本 787 × 1092 1/16
印 张 30
字 数 740 000

版 次 2003 年 12 月第 1 版
2008 年 12 月第 2 版
印 次 2008 年 12 月第 1 次印刷
定 价 35.70 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 25525 - 00

前 言

医学统计学是医学科学研究重要的方法学基础,在医学教育和医学生培养中具有举足轻重的地位。2003年8月四川大学组织编写了全国高等学校医学规划教材《医学统计学》,该教材密切联系医疗卫生实际,内容深度和广度适于医学各专业学生培养目标要求,具有适教适学特点,现已发行11万余册。本版教材基于传承与创新的编写原则,通过各位编者的通力协作,对上版教材进行了修订改版。

针对医学生在医学科研实践中研究设计功底不足、正确运用基本统计方法能力不够,以及对统计分析结果解释和表达不恰当等普遍存在的问题,本教材在注意课程内容完整性和系统性的基础上,强化统计设计在医学统计学中的地位,注重学生统计思维和统计方法运用能力的培养和训练。着力突出对学生正确开展医学科学研究、进行研究设计、运用统计方法分析资料、运用统计软件实现数据管理和分析、正确解释和表达统计分析结果的综合能力培养。针对医学生的思维特点和专业背景,尽量从实际的医学问题出发,讲解医学统计学的基本概念、原理、方法以及适用条件和注意事项。在文字上力求通俗易懂,增强可读性。在公式上进一步淡化推导及计算过程。整个教材力图达成定位明确、内容优化、重点突出、循序渐进、便于自学的编写特色与风格。

本教材共25章,内容包括研究设计和基本的统计分析方法,涵盖了国家执业医师资格考试的全部知识点。新增了统计方法选择与结果解释、统计分析结果的正确表达、数据管理及EpiData软件简介、SPSS软件简介4章内容,对于常用多变量统计方法,本教材仅作简要介绍。此外,本教材还提供了《医学统计学实习指导》,内容包括习题的SPSS软件实现、结果解释及其表达,思考题和案例辨析题的释疑或参考答案。

本教材适合于临床医学、基础医学、预防医学、护理、检验、口腔及药学等专业各年制本科生使用,由于各校开设的课程内容和学时不尽相同,因此可根据专业和学制特点选择相应章节讲授。同时本书也可供研究生、临床医师、公共卫生专业人员及相关科研工作者参考使用。

在本教材编写过程中,得到了高等教育出版社、四川大学教务处、华西公共卫生学院以及纽约中华医学基金会(CMB)的关心与支持,定稿会得到了广西医科大学和桂林医学院的热情接待和大力支持。卫生统计学前辈杨树勤教授和倪宗瓚教授也十分关心本教材的编写并提出若干建议。华西公共卫生学院卫生统计学教研室博士和硕士研究生全婷、陈俊、郑敏、陈可莉、曹明芹、张俊辉、郑杨、张子武、曹静、唐立、周艳、谭柯等对本教材的编辑、排版、例题计算结果复核等付出了艰辛的劳动。在此一并致以衷心的感谢。

本教材的编写虽经全体编委的努力工作和反复修改,但限于编者水平,难免存有疏漏之处或缺陷,欢迎同仁与读者批评指正。



2008年8月于成都

目 录

第一章 绪论	1	第四节 率的标准化	66
第一节 医学统计学的地位和作用	1	第六章 总体均数的估计	70
第二节 医学统计学基本内容和统计工作基本步骤	2	第一节 均数的抽样误差与标准误	70
第三节 统计学的若干基本概念	3	第二节 t 分布	75
第四节 统计思维的培养和本书内容的安排	5	第三节 总体均数的估计	77
第二章 调查研究设计	8	第七章 假设检验	82
第一节 调查研究的特点	8	第一节 基本思想	82
第二节 调查研究设计的基本内容	9	第二节 基本步骤	84
第三节 常用的抽样方法	14	第三节 I 型错误与 II 型错误	85
第四节 调查研究的质量控制	17	第四节 单侧检验与双侧检验	87
第三章 实验研究设计	20	第五节 假设检验需要注意的问题	88
第一节 实验设计的基本原则	20	第六节 假设检验与区间估计的联系	89
第二节 实验设计的基本内容	23	第八章 t 检验	91
第三节 常用的实验设计类型	27	第一节 样本均数与总体均数的比较	91
第四节 临床试验设计	33	第二节 配对设计均数的比较	92
第四章 定量资料的统计描述	39	第三节 两样本均数的比较	93
第一节 频数分布表和频数分布图	39	第四节 正态性检验与方差齐性检验	96
第二节 集中位置的描述	43	第五节 变量变换	98
第三节 离散程度的描述	47	第九章 方差分析	101
第四节 正态分布及其应用	51	第一节 方差分析的基本思想和应用条件	101
第五章 定性资料的统计描述	59	第二节 完全随机设计资料的方差分析	103
第一节 常用相对数及其应用	59		
第二节 应用相对数的注意事项	62		
第三节 动态数列及其应用	64		

II 目 录

第三节	随机区组设计资料的方差分析	104	第二节	秩相关	169
第四节	多个样本均数的两两比较	107	第三节	分类变量的关联性分析	170
第五节	交叉设计资料的方差分析	110	第十四章 直线回归分析	175	
第六节	析因设计资料的方差分析	112	第一节	直线回归方程的建立	175
第七节	重复测量资料的方差分析	116	第二节	直线回归的统计推断	177
第八节	多个样本的方差齐性检验	120	第三节	直线回归分析的应用	180
第十章 二项分布和 Poisson 分布及其应用	123		第四节	直线回归分析需注意的问题	182
第一节	二项分布	123	第五节	直线回归与直线相关分析的区别与联系	183
第二节	Poisson 分布	128	第十五章 生存分析	185	
第十一章 χ^2 检验	134		第一节	生存资料的特点	185
第一节	独立样本列联表资料的 χ^2 检验	134	第二节	生存分析的基本内容及几个基本概念	186
第二节	配对设计资料的 χ^2 检验	141	第三节	未分组资料的生存分析	189
第三节	拟合优度的 χ^2 检验	144	第四节	分组资料的生存分析	191
第四节	线性趋势 χ^2 检验	146	第五节	生存曲线的比较	194
第五节	四格表的 Fisher 确切概率法	147	第十六章 常用多变量统计方法简介	198	
第十二章 秩和检验	150		第一节	多重线性回归	198
第一节	Wilcoxon 符号秩和检验	150	第二节	Logistic 回归	205
第二节	成组设计两样本比较的秩和检验	153	第三节	Cox 比例风险回归	211
第三节	成组设计多个样本比较的秩和检验	156	第四节	其他常用多变量统计方法概述	216
第四节	随机区组设计的秩和检验	160	第十七章 诊断和筛检试验的评价	223	
第十三章 双变量关联性分析	164		第一节	诊断和筛检试验的基本概念和应用	223
第一节	直线相关	164	第二节	诊断和筛检试验的评价体系	225
			第三节	ROC 曲线及其应用	234
			第十八章 Meta 分析	239	
			第一节	Meta 分析的基本原理和步骤	239

第二节	Meta 分析的基本方法	244	第一节	统计方法的正确选择	302
第三节	Meta 分析的注意事项	252	第二节	统计结果的正确解释	306
第十九章	样本含量估计	255	第二十三章	统计分析结果的正确	
第一节	样本含量估计的意义及其		表达	309	
基本条件	255	第一节	常用统计表与统计图	309	
第二节	调查研究常用样本含量		第二节	医学论文统计报告的基本	
估计	256	要求	316		
第三节	实验研究常用样本含量		第三节	统计分析结果表达的常见	
估计	258	错误	319		
第四节	检验效能的估计	263	第二十四章	数据管理及 EpiData 软	
第二十章	测量手段的效度和信度		件简介	322	
评价	267	第一节	数据录入	322	
第一节	效度和信度的概念	267	第二节	数据核查与清理	339
第二节	效度的评价方法	270	第三节	数据预处理	343
第三节	信度的评价方法	271	第二十五章	SPSS 软件简介	352
第四节	量表的效度和信度评价		第一节	SPSS 窗口及菜单	352
方法	273	第二节	SPSS 的基本统计分析		
第二十一章	医学人口与疾病统计		功能	353	
常用指标	276	附录一	思考与练习	388	
第一节	医学人口统计常用指标	276	附录二	统计用表	437
第二节	疾病统计常用指标	287	附录三	英汉专业术语	457
第三节	寿命表及其应用	291	参考文献	469	
第二十二章	统计方法选择与结果				
解释	302				

第一节 医学统计学的地位和作用

当人们研发了一种治疗高血压病的新药,应该怎样评价该新药的疗效?最基本的方法就是比较。通常将患者以随机的方式分成两个组,一组服用该新药,另一组服用对照药物,观测并记录两种药物的疗效,最后统计分析该新药的有效性和安全性,这就是一个常见的临床试验。其中,统计学扮演什么角色?在这个临床试验中有诸多问题需要回答:需要多少名患者参加试验?如何随机地将患者分为两个组?哪些措施可以保证两组患者除了接受不同药物治疗外,其他影响疗效的因素在两组的分布是一致的?如果分布不一致,如何在诸多的影响因素中,分离出药物因素的效应?应采用什么样的指标来反映新药的有效性和安全性?怎样测量这些指标以保证数据的准确性和可靠性?如何控制临床试验的误差?如果两组疗效存在一定差别,怎样比较两种药物的疗效到底是否存在差别?换言之,我们需要了解这种差别是机会造成的,还是真实存在的?统计学可以回答上述问题。

我们再看另一种情形,假定为了解一个城市居民高血压病的患病现状,通常的做法是在这个城市调查一部分个体,利用这一部分个体的高血压病患病状况来反映整个城市的患病状况。那么,如何在这个城市选取这一部分个体?因为只有这部分个体能够很好地代表整个城市人群,用这种部分推论全体的做法才是准确的。此外,需要选取多少人进行调查?如何保证收集到的资料是准确和可靠的,又如何评价这种准确性和可靠性?几百人的血压(如收缩压)各不相同,看上去是一堆“杂乱无章”的数据,如何描述高血压病的患病状况,如何才能推论到整个城市人群?我们对于这种推论的正确性抱有多大的信心?统计学也可以回答上述问题。

每个人的血压都不一样,每个高血压病患者对同一种药物治疗的反应也存在着差别,这就是所谓的个体差异和不确定性。个体差异是自然界普遍存在的现象,个体结构和功能千差万别,机体反应受到各种自然和社会环境因素的影响和制约,对内外环境刺激的反应同样千差万别。在统计学中,我们将这种差异称为变异(variation)。由于变异的存在,同样条件下对同一个体的重复观测结果就具有不确定性。William Osler 爵士曾指出:“医学就是关于不确定性的科学和概率的艺术(Medicine is a science of uncertainty and an art of probability)”。生物医学中充满了个体变异和不确定性,其原因有些是已知的,有些是未知的,有的是可以人为控制的,有的则是无法控制的。事实上,客观事物在数量上所表现出来的现象既受到本质规律的制约,又受到诸多偶然因素的影响,这就妨碍了我们对事物规律性的认识。统计学正是处理数据中变异和不确定性的一门科学和艺术。它透过具有偶然性的现象来探测和揭示那些令人困惑的医学问题的规律性,对不确定性的数据作出科学推断,它是认识客观世界的重要工具和手段。因此,医学与统计

学的结合是必然的,它们的联姻就催生了医学统计学。更准确地说,医学统计学就是运用统计学的基本原理和方法来研究医学问题的一门学科,它包括了研究设计、数据收集、整理、分析以及分析结果的正确解释和表达。

医学统计学在整个医学科学研究中的地位和作用是举足轻重的,但这一点并非从一开始就被人们所认识。历史经验表明,它是在人类社会近百年的探索和实践逐步形成,甚至是在付出了若干生命代价后才逐步达成共识的。现实中,由于研究设计、数据收集、统计分析、结果解释以及结论报告任何一个环节中的缺陷或错误,已经断送了许多先进的甚至非常前沿的医学研究成果,而且即使是已经发表的研究成果也存在着大量的统计学缺陷或错误。现在,人们已逐步认识到了统计学在医学和公共卫生科学研究中的极端重要性,越来越多的临床医师、公共卫生专业人员以及实验室科学家主动寻求与统计学家合作,许多医学专业期刊邀请统计学家审稿,医学科研基金评审邀请统计学家参加,基金申请要求有统计学家合作等,这些都彰显了医学统计学蓬勃的生命力和广阔的应用空间。

第二节 医学统计学基本内容和统计工作基本步骤

在统计设计的基础上,对数据收集、整理、分析以及对分析结果的正确解释和表达是医学统计学的基本内容以及统计工作的基本步骤。值得强调指出的是,医学科研的统计设计是医学统计学的重要内容,也是统计工作的第一步和最关键的一步。一个常见和普遍的误解认为“统计”就是分析数据。我们经常会遇到这样的情形,医生或研究人员在研究结束后拿着数据咨询统计学专业人员或请其代为分析,但其研究的统计设计却存在着缺陷甚至错误。现代统计学的奠基人之一、著名统计学家 Fisher 曾精辟地指出,“做完实验后才找统计学家无异于请他作尸体解剖,他能做的全部事情就是告诉你这实验死于什么原因”。没有科学严谨的统计设计,数据的收集以及分析常常是没有价值的。对于不准确或不可靠的数据,试图寻求统计方法加以弥补亦是徒劳无益的,即使再高深的统计方法也一样于事无补,基于这些不准确或不可靠数据的统计分析所得出的结论常常是站不住脚的,甚至是误导的。因此,研究的统计设计和统计分析是统计学不可分割的两个重要组成部分。

研究的统计设计按照是否对研究对象施加干预措施常区分为观察性研究(observational study)和实验性研究(experimental study)两大类,前者又可区分为横断面研究(cross-sectional study)、病例对照研究(case-control study)和队列研究(cohort study);后者又可根据研究对象的不同区分为动物实验(animal experiment)、临床试验(clinical trial)和社区干预试验(community intervention trial)。对于不同研究设计所获得的数据,采用的统计分析方法常常是不同的,而且所得研究结论也存在差异。这在后面的相关章节将详细讲述。

数据的统计分析(statistical analysis)主要包括两个方面的内容:一是统计描述(statistical description),主要是运用一些统计指标诸如均数、标准差、率以及统计表和统计图等,对数据的数量特征及其分布规律进行客观地描述和表达,不涉及样本推断总体的问题;二是统计推断(statistical inference),即在一定的置信度或概率保证下,根据样本信息去推断总体特征。统计推断通常包括参数估计和假设检验两个内容:参数估计是指用样本指标推断总体相应的指标,例如根据部分城市人群的原发性高血压患病率去估计整个城市原发性高血压患病率;假设检验

是指由样本之间的差异推断总体之间是否可能存在差异,例如原发性高血压治疗药物在两组的疗效存在一定差别,假设检验回答这种差别是机会造成的,还是真实存在的。

第三节 统计学的若干基本概念

一、总体与样本

上面已屡屡提及总体与样本的概念。什么是总体与样本?人们通常需要了解或研究整个的一类个体,简单地说这个类就是总体(population)。但是,研究整个总体一般并不现实,能研究的通常只是它的一个部分,这个部分就是样本(sample)。人们再根据部分对整体进行推断,用统计学专业术语就是根据样本信息对总体特征作出统计推断。

更准确地讲,总体就是所有同质观察单位某种观察值(即变量值)的全体,例如调查某地2007年正常成年男子的红细胞数,则观察对象是该地2007年的正常成年男子,观察单位是每个人,观察值是每个人测得的红细胞数,该地2007年全部正常成年男子的红细胞数就构成一个总体。样本是总体中抽取部分观察单位的观测值的集合,如从该地2007年正常成年男子中,抽取300人,分别测得其红细胞数就组成样本。注意,个体间的同质性是构成总体的必备条件,也是进行研究的基本前提。

总体是根据研究目的所确定的,一般有无限总体和有限总体之分。前者指总体中的个体是无限的,如研究药物疗效,某病患者就是无限总体;后者指总体中的个体是有限的,它是指特定时间、空间中有限的研究个体。在医学研究中,多数情况总体是无限的,直接研究总体是不可能的。即使是有限总体,直接研究总体也是不现实的,它既无必要,有时也不可能。这里值得注意的是,总体中的个体在多数情形下是人,但也可以是其他个体(individual),如动物、家庭、学校、工厂、地区等,还可以是一个器官、一个细胞等。这里的个体是研究的基本单位,也是统计分析的基本单位。注意鉴别基本单位到底是什么,这在统计分析以及结果解释中都是十分重要的。

医学研究通常都想了解关于总体的某些数值特征,这些数值特征称为参数(parameter),如整个城市的原发性高血压患病率;根据样本算得的某些数值特征称为统计量(statistic),如根据几百人的调查数据所算得的样本人群原发性高血压患病率。后者是研究人员知道的,而前者是研究人员想知道的。显而易见,只有当样本代表了总体时,根据样本统计量所估计的总体参数才是准确的。

因此,选择样本的方法至关重要。机会是没有偏袒的,因而正确的方法就是采用客观的概率抽样方法选择样本。在实验设计中,将实验对象分配到处理组或对照组也应该采用随机分配的方法(详见第三章实验研究设计)。将具有某一类特征的个体排除在样本之外所表现出的系统倾向性称为选择偏性。当存在选择偏性时,抽取一个更大的样本无助于统计推断。相反,它只是在一个更大的规模上去重复错误。概率抽样方法的最重要标志就是总体中的每一个体都有同样的机会被选入样本,在理论上可计算出总体中任一个体被选入样本的机会大小。样本包含的观察单位数称为样本含量或样本大小(sample size)。

即便采用概率抽样方法抽取样本,但毕竟样本只是总体的一部分,这就存在着误差,统计学上将其称为抽样误差(sampling error)。那么,对于一次具体的概率抽样,抽样误差可能有多大?

它对样本含量的依赖程度怎样?为了将抽样误差置于控制之下,样本含量需取多少?这时用样本统计量去估计总体参数时,将会达到怎样的接近程度?这些都是统计学中十分重要的问题,大家将在后面的相关章节学习。

二、变量

变量(variable)是观测单位的某种特征或属性,变量的观测值就是所谓变量值,有时也称数据或资料(data)。更准确地讲,数据或资料是由具有若干变量值的观测单位所组成的。例如在调查中常规问及的问题:你年龄多大?是什么学历?结婚了吗?有工作吗?家里有多少人?对应的变量就是:年龄、学历、婚姻状况、就业情况、家庭人口数。有些问题的答案如年龄、家庭人口数是具体的数值,所对应的变量是定量的,称为定量变量。有些问题的答案如学历(文盲、小学、初中、高中、大学、研究生)、婚姻状况和就业状况是用语言来描述的,对应的变量是定性的,称为定性变量。

定量变量有连续和离散之分。年龄就是一个连续变量,因为不同人的年龄差异在理论上可以任意地小,如1年、1个月、1天、1小时等,它一般有度量衡单位。而家庭人口数就是一个离散变量,不同家庭的人口数可相差0、1、2等,在这些量之间不可能取其他量。当然,一个定量变量要么是连续的,要么是离散的。而对于定性变量,其取值是定性的,往往表现为互不相容的类别或属性。根据其取值特征,定性变量又可以区分为有序和无序分类的变量。所谓有序分类变量,是指其取值的各类别之间存在着程度上的差别,给人以“半定量”的感觉,因此也称为等级变量,如学历;无序分类变量又可区分为二项分类变量和多项分类变量,前者取值为相互对立的两类,如性别;后者取值为互不相容的多个类别,如血型。

上述对变量类型的区分在统计学中至关重要,因为它在很大程度上决定了统计分析方法的选择。当然,出于某些研究目的,各种类型变量间可以进行转换,如血压值为定量变量,可按照一定的临床标准,将其转换为高血压、正常血压和低血压。这种变量的转换通常具有方向性,一般从定量到半定量,再到定性,但须知这种转换后的数据,其信息量将减少。另一方面,为了对定性变量进行统计学处理,往往需要对其进行编码,例如性别的编码:男为1、女为0。这里,值得再次指出,变量类型的区分也与分析的基本单位有关。例如,患病与否的问题,若以人为基本测量和分析单位,它是二项分类变量,但若以乡为基本测量和分析单位,患病率则为定量变量。

三、误差

误差(error)泛指实测值与真实值之差,一般可区分为随机误差和非随机误差两大类。随机误差是一类不恒定的、随机变化的误差,往往使实测值无方向性地围绕着某一数值左右波动。抽样误差即为随机误差,即由于随机抽样造成的实测值与真实值之差。随机误差在随机抽样和观测中是不可避免的,但一般服从正态分布,可以通过统计学方法进行分析。

最常见的非随机误差即所谓系统误差,是指使实测值系统偏离真实值的、具有方向性的误差,因此也常称为偏性或偏倚(bias)。其产生的原因往往是可知的或可掌握的,例如仪器未校正、操作不规范等。因此,通过完善研究设计、规范操作流程、改进技术手段等方式,可以降低或消除系统误差。此外,还有一些非随机误差是在研究过程中由于研究者的偶然失误造成的,例如误读检验结果、记录失误等,即所谓过失误差。

四、概率

概率(probability)是统计推断中最为重要的概念之一。生活中,人们经常谈论诸如中奖的机会是多少或毕业后找到工作的机会是多少?数学家们为了给“机会”这个词一个确切和清晰的解释而奋斗了几个世纪。机会的含义可以用简单的抛硬币游戏加以说明,如获得反面的机会是50%,那么最终结果大约就有50%的次数反面向上。因此,当某一实验是在相同条件下独立地一次又一次重复进行时,某事件发生次数的百分数就是该事件的机会,这就是机会的频数理论。一个事件概率的大小,取决于该事件的可能发生机会数与该事件的可能发生及不发生的所有机会数之比,这个比值就是这一事件发生概率的度量。如一个事件有4次机会发生和6次机会不发生,4/10就表示了该事件发生的概率,并且可取作该概率的度量值。概率论和数学的其他组成部分一起组成了推理艺术的重要基础,统计学的重要组成部分统计推断就是建立在此基础上的。

医学研究的现象,大多是随机现象。根据某一研究目的,在一定条件下对某一随机现象(不确定现象)进行观测,其结果在事先是不确定的,将其称为随机事件(random event),在统计学中简称事件。如同一药物治疗同一疾病,治疗后究竟会产生什么结果是不确定的,可能治愈、好转,也可能无效、死亡。这里的每一种可能发生的结果都是一个随机事件。概率是度量随机事件发生可能性大小的数值。如治疗200例患者,120名患者治愈,治愈率为60%,这是一个频率(frequency)。在现实中,概率是难以获得的,在观测单位足够多时,可以将其频率作为概率的估计值。但在观测单位较少时,用频率估计概率是不可靠的。

概率通常用 P 表示,如果某事件不可能发生,其发生概率为0,另一个极端就是,如果某事件肯定要发生,那么其发生概率为1,概率取值界于0与1两个极端之间。当某事件发生的概率小于或等于0.05时,统计学习惯上称该事件为小概率事件(small probability event),其含义是该事件发生的可能性很小,进而认为它在一次抽样中不可能发生,这就是所谓小概率事件原理,它是进行统计推断的重要基础。

第四节 统计思维的培养和本书内容的安排

一、统计思维的培养

医学生学习统计学,并非要成为医学统计学的专业人员,其目的是培养统计思维,掌握基本的统计设计方法和收集准确可靠的数据,运用基本的统计分析方法来正确分析数据,掌握操作统计软件的技能,正确解释和表达分析结果。

学习统计学的关键是通过统计思维的培养,提高科学素养和科研能力。所谓统计思维就是指统计学独特的逻辑思维方法。由于存在个体差异,用样本推断总体就会出现误差,但这种误差是有规律的,它构成了统计推断的理论基石。理解了假设检验的推理逻辑,也就理解了统计结论的概率性。因此,学习统计学就需要牢固树立起个体变异和不确定性的观念、抽样误差的观念、假设检验结论的概率性观念等。前面反复提及的一个例子就是如果两种药物疗效存在差别,那么这是机会造成的,还是真的存在?统计学用假设检验的方法来回答这一问题(这在后面的相应

章节将进行详细讲解)。现在阅读一篇医学科研论文,不遇到假设检验和 P 值几乎是不可能的。因此,弄清楚假设检验和 P 值的真实含义是学习统计学和培养统计思维的核心问题之一,尚需结合后面的学习内容加以细心体会和领悟。

此外,值得指出的另一个重要问题是,医学统计学与数学和计算机运用联系密切,但作为一门应用学科,学习它不能脱离医学背景,必须紧密结合医学专业的实际问题。学习医学统计学的最终目的是运用统计思维和统计方法去分析和解决医学实际问题。医学生由于习惯于观察、记忆、判断和操作,逐步淡化了抽象思维和逻辑思维,但这并不意味着学习医学统计学一定要具备高深的数学知识,事实上,医学统计学的许多先驱并非数学家。对于统计公式,我们认为不必深究其数学推导,重要的是要了解其意义、用途和应用条件。随着计算机技术的迅速发展和普及,带来了统计计算的简化并推动了统计学的发展,但与此同时也出现了诸如多因素统计方法的误用甚至滥用的问题。须知,医学统计中的每个数据都有其特定的专业含义,而不是抽象的数字,但计算机并不能识别这是什么含义的数据,数据是否准确可靠,也不知道如何进行分组和分析等。如果不紧密结合医学专业背景,将会导致统计方法的误用或滥用,以及统计结果的错误解释。

二、本书内容的安排

简要概述本书的框架结构和章节安排,有助于学生从整体上把握本书的内容及其逻辑关系。

第一章绪论讲解了医学统计学的基本内容以及统计学的若干基本概念,它是后面章节的重要基础;第二、三章是医学科学研究的统计设计,包括调查和实验研究设计两个部分,主要内容包包括两种设计类型的区别与特点、基本内容和原则,在调查研究设计部分介绍了常用的抽样方法,在实验研究设计部分介绍了常用的实验设计类型,尤其较为详细地介绍了临床试验设计。样本含量估计属于统计设计部分内容,但由于涉及若干尚未学习的统计学概念,故放在第十九章介绍,该章包括了两种设计类型常用的样本含量估计方法。

第四、五章为数据的统计描述,主要涉及描述定量和定性资料的常见统计指标等;第六、七章为统计推断的基本内容,包括参数估计和假设检验两个部分,主要涉及其基本思想、概念及原理,如抽样误差的概念与标准误、假设检验的基本思想、I型与II型错误、总体均数估计的方法和假设检验的基本步骤等;第八至十二章为基本的和常用的单变量统计推断方法,针对不同类型的资料 and 不同类型的设计,分别介绍了常用的各种 t 检验、方差分析、 χ^2 检验和秩和检验等,主要涉及组间比较的假设检验;第十三、十四章介绍了双变量的相关和回归分析,亦包括统计推断的内容。

第十五章介绍了针对生存资料的一类统计分析方法(即生存分析),主要涉及生存资料的特点、生存分析的基本概念,以及生存分析中统计描述和推断的内容;第十六章对最常用的多重线性回归、logistic 回归和 Cox 比例风险回归等多变量统计方法进行了简介;第十七章主要介绍了诊断和筛检试验的评价指标体系及其应用;第十八章主要介绍了 Meta 分析的基本原理、基本方法及其注意事项;第二十章主要介绍了效度和信度的概念及其评价方法;第二十一章介绍了医学人口和疾病统计常用指标、寿命表及其应用。

第二十二、二十三章阐述了统计方法的正确选择、结果的正确解释和表达。增加这两章的目的是针对在统计方法选择、结果解释和表达中大量存在的问题和常见错误,帮助学生梳理正确选

择统计方法的基本思路 and 原则、阐释在结果理解和解释上容易发生的错误,介绍了医学论文统计报告的基本要求。

最后两章(即第二十四、二十五章)简介了常用的数据管理和数据分析软件。增加这两章的目的是使学生在学习了基本的统计分析方法后,能够了解或掌握基本的数据管理和分析技能。

小 结



1. 从医学科学研究中的个体差异和不确定性出发,介绍了什么是医学统计学、医学统计学在医学科研中的地位 and 作用、医学统计学的基本内容和统计工作的基本步骤。尤其强调了医学科研的统计设计是医学统计学的重要内容之一,统计设计和统计分析是统计学不可分割的两个重要组成部分。

2. 介绍了统计学的若干基本概念,包括总体与样本、参数和统计量、概率抽样和抽样误差、定量与定性变量、随机与非随机误差、概率与频率等,它们是学习医学统计学的重要基础。

3. 学习医学统计学的目的是培养独特的统计学逻辑思维方法,掌握统计设计方法和收集准确可靠的数据,运用统计分析方法正确分析数据,掌握操作统计软件的技能,正确解释和表达研究结果。关键是运用统计思维和统计方法去分析和解决医学科研问题,提高科学素养和科研能力。

(李晓松)

2

第二章 调查研究设计

第一节 调查研究的特点

医学研究可分为观察性研究(observational study)和实验性研究(experimental study)两大类。观察性研究是一种客观地观察、记录和描述事物或现象的认识活动。作为人类获取科学事实和认识客观世界的基本方法,它具有一个显著特征,即是对处于自然状态下的事物或现象进行观察。换言之,是观察者对被观察事物或现象不进行任何干预的情况下所作的观察。通过观察性研究通常只能得到有关事物或现象及其相关因素关系的线索。实验性研究通常是在观察性研究的基础上,在人为控制实验条件或对研究事物或现象施加一定干预措施的条件下,所作的进一步研究,第三章“实验研究设计”将作详细介绍。

这里讨论的调查(survey)研究属于一种观察性研究。在医疗卫生领域,观察性研究通常可分为横断面研究(cross-sectional study)、回顾性研究(retrospective study)和前瞻性研究(prospective study)。常见的回顾性研究方法有病例对照研究(case control study),常见的前瞻性研究方法有队列研究(cohort study),流行病学教材都详尽介绍了病例对照研究和队列研究。本章所讨论的调查研究是指横断面研究,或称横断面调查、现况研究或现况调查。

横断面研究是医学领域常见的研究方法,通过收集特定时间、空间和人群中疾病或卫生事件以及相关因素,描述疾病或卫生事件在该时间、空间和人群中的分布状况以及初步探讨与之相关的因素。例如通过调查发现肝癌高发区人群的乙型肝炎表面抗原阳性率比低发区高,这就提示乙型肝炎病毒感染与肝癌高发有一定联系。由于这种调查通常是在一个较短的时间段完成的,因此被称为横断面调查。它通常是在对研究事物或现象不太了解时进行的,可为进一步的相关因素研究打下基础和提供线索。调查研究按其研究目的一般又可分为描述性研究和分析性研究两种类型。描述性研究是对疾病或卫生事件在时间、空间和人群的分布和强度进行描述,如通过样本数据推断相应的总体参数(如描述某地居民冠心病患病率现状),它通常回答研究事物或现象“是什么”的问题;分析性研究侧重探讨和分析疾病或卫生事件的相关因素及其影响作用的大小,它通常回答研究事物或现象“为什么”的问题,即分析变量间相互关系(如分析患冠心病与血压等的关系)。但在实际工作中,描述性研究和分析性研究并无截然界线,常结合使用。因此,通过横断面研究,可以了解某一特定时间断面上特定人群中疾病或卫生事件的现状及其相关因素的分布情况。

横断面研究的主要特点是:①研究过程中没有人为施加的干预措施,研究事物或现象及其相关特征(包括研究因素和非研究因素)是客观存在的。②不能将研究因素随机地分配到研究对象中,如在冠心病患病危险因素的调查中,研究者不能随机地要求某些人吸烟,而要求另外一些人

不吸烟;也不能用随机化分组来平衡非研究因素对调查结果的影响。这是调查研究区别于实验研究的另一个重要特征。例如,比较两个社区某肿瘤患病率时,年龄和性别构成不同可能成为混杂因素,但不能在调查设计阶段通过随机分配的方法进行控制。故常在资料分析阶段利用标准化法、分层分析或多因素分析等统计方法对混杂因素进行调整或控制。③横断面研究反映的是某个时间断面上的情况,事物或现象的现状以及与之相联系的因素可能是同时存在的,两者的先后顺序不清楚,因此不能得出因果关系的结论。综上所述,横断面研究一般可为进一步的相关因素研究尤其是因果关系研究提供线索,下结论时应十分慎重。

第二节 调查研究设计的基本内容

调查研究设计是调查研究取得真实和可靠结果的重要保证,它包括资料收集、整理和分析各个环节以及整个调查过程的统计设想和科学安排。调查研究设计的要点就是根据研究目的,确定调查对象和观察单位,进而确定调查方法和样本含量;将观察指标转化为调查项目,进而设计成调查表或问卷;然后确定资料收集方法和资料整理、分析计划;最后制定调查组织计划以及调查质量控制措施等。现将调查设计的基本内容介绍如下。

一、明确调查目的和指标

首先应根据研究工作的需要,明确调查目的。虽然各种调查的具体目的不同,但从统计学角度来说,进行参数估计用以说明总体特征,或是研究事物或现象间的相关联系用以探索病因或相关因素,这些都需要通过具体指标来说明。调查者应该在设计中把调查目的转化成具体指标及调查项目。

例如,2003年我国开展了第三次国家卫生服务调查,拟通过对全国城乡及不同类型地区家庭居民健康询问调查,了解居民卫生服务需要及主要健康问题,医疗保障制度对居民卫生服务利用产生的影响,居民医疗卫生服务需求和利用水平及特点,卫生服务利用的经济负担,及对卫生服务的反应性等。具体调查指标包括:通过样本地区居民两周患病率、慢病患病率、伤残率、疾病严重程度及其丧失劳动能力程度等健康状况指标,了解我国不同类型地区居民和特殊人群卫生服务需要量以及存在的主要健康问题;通过居民就诊率、住院率等卫生服务利用指标,了解居民卫生服务需要向需求转化的程度、卫生服务需求与供给之间的关系;通过样本地区居民医疗保障制度和医疗保健费用等方面的指标,了解城乡和不同类型地区医疗保障制度改革进展、各种医疗保障制度覆盖范围、居民医疗保健负担能力和负担水平。

由此可见,调查目的是选定调查指标的依据,而调查指标又是调查目的的具体体现。因此,调查研究设计时,应将调查目的转化为具体的观察指标,通过指标来达到目的。

测量事物或现象的性质或数量的指标有主观指标与客观指标之分。显然,客观指标可靠性较高,应尽量采用。此外,指标的选择还应注意其灵敏度和特异性。指标选择应紧扣研究目的,做到少而精。如果在调查中纳入与调查目的无关的内容,既增加调查成本,也可能影响资料的准确性和可靠性。

二、确定调查对象和观察单位

根据调查目的确定调查对象,即明确调查总体的同质范围。按调查目的所确定的总体称为

目标总体(target population)。目标总体中组成调查对象的每个“个体”即为观察单位,可以是一个人、一个家庭或一个群体(如一家医院或一所学校等),以人为观察单位的调查较多见。例如,调查某乡 3~7 岁儿童生长发育状况,调查对象即该乡某年全体 3~7 岁儿童,观察单位即每个 3~7 岁儿童。

应根据不同的研究目的选择不同的调查对象。若调查目的是早发现、早诊断、早治疗,则可选择高危人群,如为早期发现糖尿病患者,就可以选择有糖尿病家族史、肥胖、缺乏锻炼等高危人群进行筛查;若调查目的是了解暴露因素与疾病的关联,则可选择暴露人群或职业人群等。例如,第三次国家卫生服务调查的目的是提供人群健康状况以及卫生服务状况等信息,调查的对象为全国抽样住户的实际人口(凡居住并生活在一起的常住人口)。

确定调查对象和观察单位后,还要确定需要调查多少观察单位,即样本含量(sample size)。调查的抽样设计中,样本含量的估计是一个十分重要的问题。抽样误差大小与样本含量直接相关,因此确定一个恰当的样本含量可将抽样误差控制在一定范围内。样本含量过少,所得指标不稳定,推断总体精度差,检验效能低;样本含量过多,不仅增加调查成本,而且可能增大各种非抽样误差,给调查的质量控制带来更多的困难。因此,样本含量估计的目的是在保证一定推断精度和检验效能的前提下,确定最少的观察单位数。调查设计中常涉及的是估计总体均数或总体率时所需要的样本含量,由于其具体计算涉及统计学中其他知识(后续章节将陆续介绍),因此,调查研究设计中的样本含量估计问题安排在本书的第十九章介绍。

三、确定调查方法

(一) 普查

普查(overall survey)就是全面调查(complete survey),即调查目标总体中全部观察对象,如我国的人口普查。理论上普查没有抽样误差,可以直接得到总体参数。但如普查规模太大,往往非抽样误差较大。普查可以了解总体在某一特定“时点”上的情况,如年中人口数、时点患病率等。但因普查成本较高,应用时应注意其成本效益和社会影响。一般在疾病普查时,应考虑以下条件:①疾病患病率的高低和病程的长短;②是否具有灵敏度和特异度较高的检查或诊断方法;③普查方法是否便于普查人员操作、易被群众接受;④是否具有实施以及治疗条件。此外,在调查时点患病率时,还应尽可能在短期内完成。

(二) 抽样调查

抽样调查(sampling survey)是一种非全面调查,即从总体中抽取一定数量的观察单位组成样本,对样本进行调查。抽样调查不仅节省调查成本,还有助于获得较为深入细致和准确的资料。因此,它在医疗卫生工作中应用最多。况且,许多医疗卫生问题只能作抽样调查,如大气或水污染的调查、食品卫生质量的检查等。此外,抽样调查还可用于评价普查质量,如人口普查中的 5% 抽样调查。

抽样调查有概率抽样(probability sampling)与非概率抽样(non-probability sampling)之分。概率抽样是指总体中观察单位被抽中的概率是已知的或可以计算的。概率抽样的样本对总体代表性较好,可以计算抽样误差,可以对总体进行统计推断。应用概率抽样的前提是目标总体和抽样框架(sampling frame)明确。所谓抽样框架是指由全部抽样单位组成的可用于抽样的清单。根据观察对象的特点,采用的概率抽样方法有所不同。调查研究中常用的几种概率抽样方法有