

成人高等教育 (本科) 专业教材

医学统计学

YIXUE TONGJIXUE

李君荣 孙 峰 主编

江苏大学出版社

成人高等教育(本科)专业教材

医学统计学

主 编 李君荣 孙 峰
编 者 (以姓氏笔画为序)
王劲松 刘丽群 孙 峰
孙 蓉 李中兴 李君荣
唐艳林

江苏大学出版社

内容提要

本书按照医学科研的流程,以研究设计作为起点,对与医学研究最为密切的医学统计学内容进行了全面介绍。本书条理清楚,深入浅出,密切联系医学科研实践,具有三大特色:打破传统医学统计学章节编排顺序,按照医学科研人员分析资料的思维习惯安排章节顺序,便于将基本统计学方法应用于医学科研实践;增加医学科研中的基本统计方法选择等内容,特别是基于研究设计作为起点的统计流程图对于初学者具有重要的指导意义;增加附录 SPSS 软件使用精要,所有案例均来自本书中的例题,对于学生理解统计理论和统计软件,非常有帮助。

本书系医药卫生类成人高等教育的教材,也可供相关医学类专业的全日制本科学生使用,还可作为医卫人员的参考用书。

图书在版编目(CIP)数据

医学统计学/李君荣,孙峰主编. —镇江:江苏大学出版社,2009.7

ISBN 978-7-81130-097-0

I. 医… II. ①李…②孙… III. 医学统计 IV. R195.1

中国版本图书馆 CIP 数据核字(2009)第 111301 号

医学统计学

主 编/李君荣 孙 峰

责任编辑/许 龙

出版发行/江苏大学出版社

地 址/江苏省镇江市梦溪园巷 30 号(邮编:212003)

电 话/0511-84446464

排 版/镇江文苑制版印刷有限责任公司

印 刷/丹阳市兴华印刷厂

经 销/江苏省新华书店

开 本/787 mm×960 mm 1/16

印 张/13.5

字 数/285 千字

版 次/2009 年 7 月第 1 版 2009 年 7 月第 1 次印刷

书 号/ISBN 978-7-81130-097-0

定 价/25.00 元

本书如有印装错误请与本社发行部联系调换

前言

Foreword

随着我国卫生事业的不断发展,医学科学研究所面临的任务日益艰巨,各种医学信息的传播和交流日益广泛。医学科学研究以及医药卫生工作的实践都会涉及大量的信息资料,如何整理、分析、挖掘这些信息,以达到充分利用信息的目的,就要对事物的变化有更深入的认识。事物的变化有必然性和偶然性,且经常是偶然性(不确定性)掩盖了必然性,妨碍了人们对客观规律的认识。如何去伪存真、去粗取精,准确地从表面纷繁杂乱的不确定性现象中揭示其内在的规律(必然性),正是医学统计学的基本任务。学习和掌握医学统计学的原理和方法,正确地进行研究设计,搜集、整理和分析医药卫生信息,将有助于人们科学地总结和分析疾病发生、发展规律,为相关决策提供依据,促进医学教育和医学科研的发展。

为满足 21 世纪高等院校医药卫生专业人才的培养需要,培养高质量的医学专业人才,江苏大学继续教育学院组织长期从事成人教育的富有经验的有关专家学者,在充分考虑成人教育特点的基础上,总结编写了本书。本书内容简明实用,对概念和原理解释力求准确,对理论、方法、技术等介绍努力做到科学规范,对语言文字的叙述力戒繁琐。书中所有统计方法均从医学实例入手,由浅入深讲述其基本概念、基本原理、基本方法以及各种统计方法的用途和应用条件,便于学生实际操作时模仿和借鉴,是一本富有特色和创新意义的教材。

全书共分九章,包括绪论、实验设计基础、数值变量资料的统计描述、数值变量资料的统计推断、分类变量资料的统计描述、分类变量资料的统计推断、两

变量关系的统计描述与统计推断、医学论文中的统计图表、医学科研中的基本统计方法选择。每章附有小结和复习思考题,可供学生自测和实习之用。

本书系医药卫生类专业成人高等教育的教材,也可供相关医学类专业的全日制本科学生使用,还可作为医卫人员的参考用书。

参加本书编写的有:江苏大学的李君荣(第一、九章)、刘丽群(第三、八章)、唐艳林(第四章),扬州大学的孙峰(第五、六章)、孙蓉(第七章)和王劲松(第二章),镇江市第二人民医院的李中兴(第九章)。本书由李君荣拟定大纲并对全书进行统稿。附录由李君荣、唐艳林编写及整理。

本书在编写过程中,得到了相关单位的大力支持,特别是江苏大学继续教育学院领导的支持。全书的复算由江苏大学医学统计教研室的全体老师完成,特别是刘丽群、唐艳林两位老师和研究生李春楠、陈鑫铭付出了辛勤劳动。在此,向以上所有对本书给予关心和支持的各位领导和同仁以及被引用文献的有关作者致以诚挚的谢意。

限于编者水平,本书不当之处在所难免,恳请广大读者提出批评和建议,以便再版时修订。

编 者

2009年6月

目 录 Contents

第一章 绪 论	(1)
第一节 统计学发展简史	(1)
第二节 医学统计学在医学科研中的地位和研究内容	(3)
第三节 统计工作的基本步骤	(4)
第四节 常用的统计学基本概念	(6)
第二章 实验设计基础	(10)
第一节 实验设计的意义	(10)
第二节 实验研究的基本要素	(11)
第三节 实验设计的基本原则	(14)
第四节 两种简单的实验设计方案	(17)
第三章 数值变量资料的统计描述	(21)
第一节 数值变量资料的整理与描述	(21)
第二节 正态分布	(31)
第四章 数值变量资料的统计推断	(39)
第一节 均数的抽样误差	(39)
第二节 t 分布	(41)
第三节 统计推断基础	(43)
第四节 t 检验	(46)
第五节 方差分析	(51)
第六节 t 检验与方差分析的应用条件检验	(59)
第七节 非参数统计方法	(62)
第八节 假设检验应注意的问题	(71)

第五章 分类变量资料的统计描述	(75)
第一节 分类变量资料的整理与统计描述	(75)
第二节 率的标准化法	(80)
第三节 二项分布	(83)
第六章 分类变量资料的统计推断	(89)
第一节 率的抽样误差	(89)
第二节 总体率的估计	(90)
第三节 单样本设计的 u 检验	(91)
第四节 卡方检验	(92)
第七章 两变量关系的统计描述与统计推断	(107)
第一节 直线回归	(107)
第二节 直线相关	(116)
第三节 直线回归与相关分析时的注意事项	(119)
第四节 等级相关	(120)
第八章 医学论文中的统计图表	(125)
第一节 统计表	(125)
第二节 统计图	(128)
第九章 医学科研中的基本统计方法选择	(135)
第一节 基本统计方法选择	(135)
第二节 基本统计方法选择流程图	(139)
参考文献	(143)
附录 A 统计用表	(144)
附录 B 统计软件 SPSS 使用精要	(178)
附录 C 常用英汉医学统计学词汇	(205)

第一章 绪 论

人类社会已经进入了科技高速发展的时代,各种信息的传播和交流日益加快。医学科研和医药卫生工作的实践以及居民健康状况研究,都会涉及大量的数据信息资料,如何整理、分析、挖掘这些信息,达到充分利用信息的目的,就要对事物的变化有更深入的认识。事物的变化有必然性和偶然性,且经常是偶然性(不确定性)掩盖了必然性,妨碍了人们对客观规律的认识。如何去伪存真、去粗取精,正确地从事物纷繁杂乱的不确定性现象中揭示其内在的规律,正是统计学所面临的根本任务。

第一节 统计学发展简史

统计一词由来已久,它的含义在历史上不断演变。拉丁语“state”的意思是各种现象的状态和状况。在原始社会里,人类最初的一般计数活动就蕴藏着统计萌芽。直到 17 世纪中叶以后,随着统计实践的发展,客观上要求总结经验,使之上升为理论,并进一步指导实践。由此,欧洲开始出现一些统计理论著作,统计学(statistics)应运而生,并出现了一些不同的学派。

一、古典统计学时期

这是统计学的萌芽期,指 17 世纪中叶至 18 世纪中叶。当时统计学主要有两大学派。

(一) 政治算术学派

该学派产生于 17 世纪中叶的英国,代表人物是 W. Petty,代表作是《政治算术》。在这部著作中,Petty 首次用计量和比较的方法对英、法、荷三国的经济实力、军事实力和内在潜力进行了分析,主张一切论述都用数字和尺度来进行,并提出用图表形式概括数字资料的理论和方法。其理论与方法对后来的统计学形成与发展有深远的影响。

(二) 国势学派(亦称记述学派)

该学派产生于 18 世纪的德国,主要代表人物是 H. Coning 和 G. Achenwall。Achenwall 的一个重要贡献是,他首次把国势学定义为“statistik”,统计学的名称由此沿用而来。该学派后来逐渐发展为政府统计。

二、近代统计学时期

这是统计学的发展期,指 18 世纪末至 19 世纪末。在这段时期内,统计学发展并形成了许多学派,其中最主要的是数理统计学派和社会统计学派。

(一) 数理统计学派

它是在概率论已有相当发展的基础上,将概率论引入统计学而形成的。该学派的奠基人是比利时学者 A. Quetelet。他的老师 P. S. Laplace 主张用自然科学的方法研究社会现象,并正式将概率论引入统计学,使统计学进入了一个新的发展阶段。Quetelet 最早用大数定律论证了社会现象复杂变化的偶然性中存在着规律性,并提出了误差理论,用以解决统计学的准确性问题。此外,他还将统计学明确定义为一门既研究社会现象又研究自然现象的独立的方法论科学。他的思想在英美等国得以发展,至 19 世纪末,数理统计自成一派。

(二) 社会统计学派

该学派以德国为中心,由 K. G. A. Knies 首创,主要代表人物有 C. L. E. Engel 和 G. V. Mayr 等人。他们认为统计学是一门社会科学,是研究社会现象变动原因和规律性的实质性科学,这与数理统计学派的通用方法论相对立。社会统计学派研究社会总体而不是个别的社会现象,将政府统计与社会调查结合起来以形成自己的特点,在国际统计学界占有一定的地位,对日本等国的统计学界有一定的影响。

随着社会经济的不断发展,人们要求统计学提供更多的方法,社会科学本身也不断向定量化发展,日益重视方法论的研究,从而出现了统计学从实质性科学向方法论转化的趋势。

三、现代统计学时期

这是指 20 世纪初迄至今日的统计学发展时期。这一时期的科学技术迅猛发展,社会生产发生巨大变化。数理统计学同自然科学、工程技术紧密结合而被广泛应用,获得迅速发展,进入了鼎盛时期。首先, W. S. Gosset 提出了小样本的 t 分布理论,随后, R. A. Fisher 对

其加以充实,并由 J. Neyman 等人进一步发展,建立了统计假设理论。后来,A. Wald 创立了决策理论;S. S. Wilks 和 J. Wishart 等对样本分布理论加以充实和发展;R. A. Fisher 提出了实验设计的概念和三个基本原则(对照原则、随机化原则和重复原则),W. G. Cochran 等对其加以丰富和发展,进一步拓宽了统计学的范围。在这期间特别值得一提的是,数理统计学派吸取了生物学研究的有益成果,F. Galton, K. Pearson, W. S. Gosset, R. A. Fisher 等一批杰出的统计学家提出并发展了回归分析、相关分析、假设检验等方法。20 世纪 60 年代以后,数理统计学的发展呈现出三个明显的趋势:一是越来越广泛地应用数学方法;二是出现了一些新的分支(如抽样理论、非参数统计、多元分析、时间序列分析等)以及一些边缘学科(如经济计量学、医学统计学、天文统计学、工程统计学等);三是应用更加广泛,尤其伴随着计算机的发展,各种统计软件层出不穷(如 SAS、SPSS、Stata 等),使其所发挥的作用也愈加明显。

第二节 医学统计学在医学科研中的地位和研究内容

统计学原理和方法几乎应用到自然科学和社会科学的各个领域,产生了许多应用性分支,诸如社会经济统计学、工业统计学、生物统计学、教育统计学等。

医学领域的研究对象主要是人体以及与人体健康相关的各种因素,具有其特殊性,并受到社会、经济和心理等诸多因素的影响。这些影响具有不确定性,必须透过这些不确定性来探测其内部潜含的规律性,统计学便有了用武之地。医学统计学(medical statistics)就是运用概率论和数理统计原理、方法结合医药卫生工作的实际情况,阐述医学科研设计的基本原理,研究医学资料(信息)的搜集、整理和分析的方法学总称,它是认识医学现象数量特征的重要工具。

医学生学习统计方法,主要是因为:① 医学上许多现象(如血压、脉搏、SGPT 等生理生化指标测定)都是随机现象。随机现象广泛存在于生物医学的各个领域,对于这些“随机现象”,由于其不确定性,只有借助概率论原理,运用统计学方法,帮助我们透过偶然性来认清事物内部潜在的客观规律。② 现在可获得的书刊资料很多,如何识别错误信息,必须掌握一定的统计学知识。20 世纪 70 年代以后发展起来的 DME(Design, Measurement & Evaluation)就是应用统计学原理和分析方法,结合流行病学等相关学科,帮助临床医师阅读文献资料、评价医学文献、开展医学科研和总结工作经验,使医学统计方法的应用范围更加广泛。③ 通过本学科的学习,充分认识研究设计和原始资料的重要性,有助于培养学生如实反映研究中的现象这一每位科研人员必备的品质。医学科学研究大量运用到统计学,其作为一种工具帮助我们揭示医学表象后的内在规律,但其本身并不能创造任何规律,为获得规

律性的统计结果而伪造和篡改数字是极其不可取的。

医学统计学的研究内容非常广泛,几乎所有的统计原理和方法都能在医学领域得以直接或间接地应用,根据目前该领域研究的现状以及教学学时的安排,本书着重介绍的内容有:研究设计基础、统计分布理论、样本资料的统计描述、统计推断(包括参数估计及假设检验)、相关分析与回归分析等。书中的医学科研中的基本统计方法选择一章,便于学生学完本课程以后,能够较为快捷和正确地将所学的知识直接应用于未来的医学研究实践;另外,书中的统计软件包 SPSS 使用精要,为读者快速掌握并利用 SPSS 分析实际资料提供了有效的手段,同时也便于提高课堂教学的效率与效果,让学生不要过多地关注公式的推导和计算,而应充分认识到统计概念和统计思维的重要性,掌握各种统计方法的应用条件,从专业上合理地解释统计结果。

学习本课程的目的并非让每位读者成为统计学家,而是在于使医学生掌握统计学的思维方法,学会从不确定性和概率的角度去考虑医学问题。一要能够结合自己所学的医学专业知识,合理地进行研究设计,获得可靠、准确的资料,充分挖掘资料中隐含的信息,恰如其分地作出理性的概括,撰写出具有较高水平的医学科学论文;二要能运用自己所学的统计学知识,批判性地阅读、评价文献资料。这对于提高自身的科研素质具有非常积极意义。

第三节 统计工作的基本步骤

统计工作的步骤类似于科研工作的步骤,一般分为四步:研究设计(design)、搜集资料(collection of data)、整理资料(sorting data)和分析资料(analysis of data)。需要说明的是,上述四个步骤在实际工作中并不是完全分开的,前一步骤是后一步骤的基础,后者又对前一步骤提出要求,相辅相成,缺一不可。任何一个步骤有缺陷,都会影响统计分析的结果,乃至研究结论的可靠性。

一、研究设计

统计学所讲的设计,是指用尽可能少的实验次数来获得足够有效的资料,从而得出较为可靠的结论。一份良好的设计应该是专业设计和统计设计的有机结合,没有良好的统计设计,就得不到可靠的数据,也不便于对资料进行统计分析。统计工作首先要有一个全过程的设计,这是很重要的一步(详见第二章)。

二、搜集资料

多数学者认为用“搜集”比较合适,即有目的地搜集资料,而不是漫无目的地“收集”。资料搜集的好坏是统计工作(科研工作)成败的关键,如果搜集原始资料的计划不周密,原始记录不准确,就会造成整理资料、分析资料的困难,甚至得出错误的结论。资料的来源主要有以下几方面。

(一) 统计报表

统计报表有多种,包括医院工作报表、疫情报表、职业病报表、出生或死亡报表等。这类报表通常是由于工作需要,由国家或地方统一设计,要求有关卫生机构定期、逐级上报,其所含信息完全可以为科研所利用。

(二) 日常医疗卫生工作记录

在日常工作中,住院、门诊病历以及各种医学检查记录等,不仅是临床医疗工作的凭证,如果资料准确、完整,也是医学科学研究非常宝贵的资料来源。

以上是经常性资料。为了使这些资料可以供医学科研使用,必须注意资料的完整性和正确性,防止重复、遗漏或者项目填写不清等情况发生;一旦发现错误,应及时纠正,否则得到的资料就不可靠。因此,医疗卫生人员必须从思想上认识原始记录的重要性。

(三) 专题调查或实验性研究资料

有些资料不能通过上述途径得到,如肝癌的病因学研究、某种药物的疗效观察等,必须进行专题调查或实验才能取得所需的资料。这类资料由事先具有较为严密的设计而获得的,其价值相对较高。

三、整理资料

整理资料又称统计归纳,即根据研究者整理分析计划的要求对资料进行分组与汇总,使资料条理化、系统化的过程,其目的是便于分析。

在整理资料之前,首先应对原始数据(raw data)进行复核、检查,数据准确无误后方可进行整理、分析。复核、检查的手段有多种,可以通过手工核实、逻辑检查等方法发现一些常见的错误。值得一提的是,随着计算机应用的普及,数据管理大多采用数据库进行。在数据录入的过程中,录入错误不容忽视,它对统计分析的质量带来严重影响,甚至会歪曲结论。

除了加强责任心之外,目前较好的解决方法是输入两次,建立两个数据库,再编程比较这两个数据库,若发现不一致,即表示有录入错误。

四、分析资料

分析资料是对整理过的资料作进一步的统计计算、分析的过程。统计分析包括统计描述和统计推断,其结果可用适当的统计图表表示;同时,还必须结合专业知识对统计结果给出恰如其分的专业解释结论。

第四节 常用的统计学基本概念

一、同质与变异

同质是指性质相同或相近的事物。统计学常要求的同质观察对象,是指性质相同或相近的观察对象。例如,研究某种疾病时,一定要有统一的诊断标准;研究儿童身体发育情况时,同性别、同年龄的儿童称为同质观察单位(unit)。然而,即使是同质观察单位,就其某一特征(观察指标)而言,个体之间也是存在差异的,这种差异称为变异(variation)。例如,同年龄、同性别的儿童,其体重不一定相同;某药物治疗患有同一疾病的患者,有的患者能治愈,有的患者好转,有的患者则无效甚至恶化、死亡。

没有变异就没有统计学。由于生物医学领域变异的广泛存在,统计学在该领域的应用就显得非常重要和必要。

二、总体与样本

总体(population)是指根据研究目的所确定的同质观察对象的全体,它是由客观存在的具有相同或类似性质的观察对象组成。总体分为有限总体和无限总体。例如,调查某市小学一年级入学儿童的体重情况,则该市所有的一年级入学儿童就构成了本次研究的总体,作为有限总体,其同质的基础是生活在该市范围内的一年级入学儿童。样本(sample)是从总体中随机抽取的部分有代表性的观察单位,某指标的实测值即构成了样本。例如为了解该市一年级入学儿童的平均体重,可以对总体中的所有儿童调查,称之为普查。但由于时间、人力、物力和财力等方面的限制,普查的难度比较大,如其他方法也能达到同样的研究目

的,普查就没有必要。此时,可以在总体中随机抽取一部分有代表性的一年级入学儿童作为本次研究的调查对象,这一部分儿童就组成了本次研究的样本。

医学研究大都研究的是样本。了解这一概念非常重要,这对于掌握后述的许多统计方法,特别对抽样误差以及统计推断的必要性至关重要。

三、参数与统计量

描述总体特征的统计指标统称为参数(parameter),用希腊字母表示;而由样本的信息所计算的统计指标,称之为统计量(statistic)。例如,某市所有一年级入学儿童的平均体重就是一个参数,而根据样本所计算出的平均体重则是一个统计量。

四、确定性现象与随机现象

所谓确定性现象是在一定条件下,一定会发生或一定不会发生的现象。其表现结果为两种事件:肯定发生某种结果的事件称为必然事件;肯定不发生某种结果的事件称为不可能事件。所谓随机现象是在同样条件下可能会出现两种或多种结果,究竟会发生哪种结果,事先不能确定,其表现结果称为随机事件。随机事件必须具备两个特征,即不确定性和规律性(每次发生的可能性的确定)。

五、抽样研究与抽样误差

从总体中随机地抽出一部分个体(样本)进行研究,称为抽样研究(sampling study)。总体中的个体往往存在着变异,而样本只是总体中的一部分,因而,由样本所计算出的统计量常常与总体的参数之间存在着差异,这种由于随机抽样所造成的样本统计量与总体参数之间的差异,称之为抽样误差(sampling error)(详见第四章)。

六、频率与概率

在 n 次随机试验中,事件 A 发生 m 次,则以下比值

$$f = \frac{m}{n} = \frac{A \text{ 发生的试验次数}}{\text{试验的总次数}} \quad (1-1)$$

称为事件 A 在 n 次试验中出现的频率(frequency), m 称为频数(frequency number)。频率常用小数或百分数表示,显然有: $0 \leq f \leq 1$ 。在医学研究中,常用的发病率、病死率、阳性率

等都可看作是频率。

实践表明,在重复试验中,事件 A 出现的频率,随着试验次数的不断增加,愈来愈接近一个稳定的常数,这个常数就称为事件 A 出现的概率(probability),可以记为 $P(A)$ 。概率的这一定义称为概率的统计定义,它描述了事件 A 发生的可能性,其值介于 0 和 1 之间。概率愈近于 1,该事件发生的可能性就愈大;某事件发生的概率等于 0,称为不可能事件;某事件发生的概率等于 1,称为必然事件;介于两者之间,称之为随机事件。统计学研究关注的是随机事件。概率很小的随机事件称为小概率事件。

七、统计资料的类型与变量

统计资料是指观察单位的原始记录。关于统计资料的类型,从不同的角度和研究目的出发,其分类不完全相同。但从应用的角度而言,可以将原始资料分为:

(1) 定量资料(quantitative data): 对每个观察单位用定量方法测定某项指标量的大小,所得资料称为定量资料。它一般有度量衡单位,例如体重(kg)、身高(cm)、血红蛋白含量(g/L)等。

(2) 定性资料(qualitative data): 以定性的方式来表达观察单位的某项指标。例如性别、血型、职业等。

(3) 等级资料(ranked data): 介于上述两资料之间,以等级来表达观察单位的某项指标。等级资料类似于定性资料,不同的是属性分组有程度差异,等级资料所分各组按大小顺序排列。例如疗效(治愈、好转、无效、死亡)、大便隐血试验(一、±、+、++、+++、++++)等,各类结果之间有程度上的差别(等级关系),即所谓的“半定量资料”。

值得注意的是,由于研究目的不同,原始资料的记录方法也常常不同,上述类型的资料在使用时需要相互转换。随着计算机技术的发展,各种统计软件层出不穷,如 SAS、SPSS、Stata 等,这些软件无不使用变量来处理数据。所谓变量(variable)在统计学上是对随机变量(random variable)的简称,指的是研究人员对观察对象感兴趣的某一特征或指标。变量根据其取值可以分为:

(1) 数值变量(numerical variable)或称定量变量(quantitative variable): 即变量的取值是定量的,表现为数值大小。它相当于定量资料的概念。

(2) 分类变量(categorical variable)亦称定性变量(qualitative variable): 变量的取值是定性的,表现为互不相容的类别或属性。根据属性或者类别的多少又可分为二分类变量和多分类变量。

① 二分类变量: 将观察单位按两种属性分类,如性别作为一个变量,表现为互不相容的两类结果(男或女);再如生存和死亡,有效和无效,阳性和阴性等均为两互不相容的类别,

属于定性资料。

② 多分类变量：观察单位的属性分类超过两种为多分类变量，它又包括两种情况。一种为无序分类(unordered categories)，如血型变量可以取 A, B, O 以及 AB 型四个值，表现为四类结果之间彼此互斥，属于定性资料；另一种为有序分类(ordered categories)，如临床疗效的观察（治愈、显效、好转、无效等），尿蛋白试验（结果为一、±、+、++、+++、++++等），各类结果之间有程度上的差别（等级关系），属于等级资料。

【本章小结】

医学统计学是医学科研和医药卫生管理领域的一门重要基础课程，它的研究内容非常广泛，涉及研究设计、资料搜集、整理与分析等，探索从不确定性现象中揭示其内在的规律。医学统计学不仅理论性较强，而且其实践性也非常强。希望读者在今后的科研实践中，将所学的医学统计学知识批判性地运用于文献阅读与理解之中，同时正确地选用统计方法处理科研资料。本章学习的重点应放在掌握几个重要的统计学基本概念上，且能举一反三。

【复习思考题】

1. 何谓医学统计学？其研究的主要内容是什么？
2. 医学统计工作的基本步骤有哪些？
3. 统计资料的类型有哪些？之间的关系如何？

第二章 实验设计基础

良好的实验设计是实验研究中各因素得以合理安排,实验结果准确、研究结论可靠的基本保证,合理地选择与运用实验设计方法是医学科研的基本要求,也是统计分析的前提。本章就实验性研究的有关概念、实验设计的基本原则以及常用的两种实验设计方案进行介绍。

第一节 实验设计的意义

一、医学科学研究类型

根据研究者是否能够人为地设置处理因素,即是否主动对研究对象施加一定的干预措施,科学研究方法分为实验性研究(experiment study)和观察性研究(observational study)两类。

(一) 实验性研究

实验性研究亦简称实验,是指研究者能够人为控制某些条件,将干预措施(处理因素)施加给研究对象,观察其变化及结果。例如,为研究健康教育等综合性干预措施对心脑血管疾病的预防效果,将符合条件的社区居民随机分成实验组和对照组,实验组给予健康知识教育、饮食指导等干预措施,对照组不实施干预。一段时间以后,比较两组人群血压值、血清胆固醇等指标的大小,从而推断出综合性干预措施是否有效。实验性研究的优点在于能够较好地控制非处理因素(即混杂因素)的影响,使得各对比组之间均衡可比,其缺点是当样本含量比较小时,不能保证非处理因素对各组的影响相同。在实际工作中,习惯上将以人为对象的实验,称为试验(trial),如临床试验、现场试验、社区干预试验等。