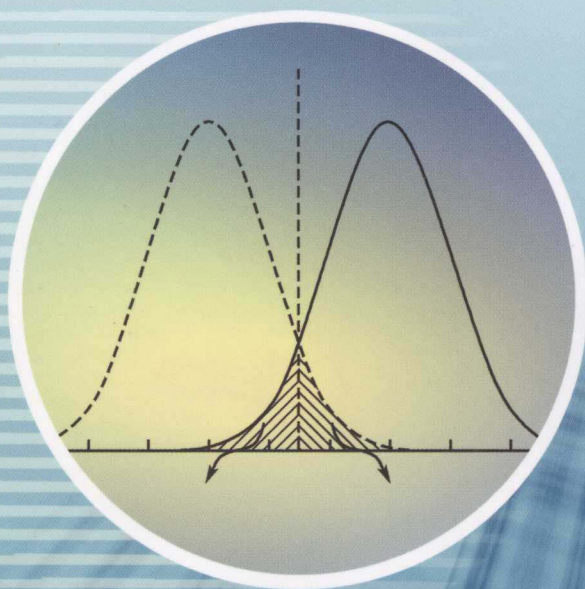




中国科学院教材建设专家委员会规划教材
全国高等医药院校规划教材

医学统计学

◎钟晓妮 主编



科学出版社

014308037

R195.1-43

10

中国科学院教材建设专家委员会规划教材
全国高等医药院校规划教材

医学统计学

主 编 钟晓妮

副主编 张菊英 方 亚

主 审 周燕荣

编 委 (按姓氏汉语拼音排序)

陈卫中 四川大学华西公共卫生学院

邓 丹 重庆医科大学公共卫生与管理学院

方 亚 厦门大学公共卫生学院

高菲菲 辽宁医学院公共卫生管理学院

高晓凤 川北医学院预防医学系

宋桂荣 大连医科大学公共卫生学院

王润华 重庆医科大学公共卫生与管理学院

王学梅 内蒙古医学院公共卫生学院

叶孟良 重庆医科大学公共卫生与管理学院

张 燕 重庆医科大学公共卫生与管理学院

张菊英 四川大学华西公共卫生学院

张星光 内蒙古医学院公共卫生学院

钟晓妮 重庆医科大学公共卫生与管理学院

科学出版社

北 京



北航

C1694350

R195.1-43

10

内 容 简 介

从实际应用角度出发,按章节顺序本书包含绪论、医学研究统计设计、估计性分析、组间统计量差异比较性分析、两变量间的关联性与趋势性分析、诊断试验评价、临床疗效的评价、统计表与统计图以及常见错误案例分析九个章节内容。第一章至第五章、第八章、第九章介绍医学统计学基本知识,第六章、第七章主要为医学统计学知识在临床上的应用。标注*章节可根据具体情况作为选学内容。

本书适合临床医学专业及其他医学专业学生、医学科研人员等使用。

图书在版编目(CIP)数据

医学统计学 / 钟晓妮主编. —北京:科学出版社,2013
中国科学院教材建设专家委员会规划教材·全国高等医药院校规划教材
ISBN 978-7-03-038909-1

I. 医… II. 钟… 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(2013)第 247367 号

责任编辑:邹梦娜 / 责任校对:赵桂芬

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>

北京市文林印务有限公司 印刷

科学出版社发行 各地新华书店经销

*

2014 年 1 月第 一 版 开本:787×1092 1/16

2014 年 1 月第一次印刷 印张:11 3/4

字数:272 000

定价:29.80 元

(如有印装质量问题,我社负责调换)

前 言

统计学是一门认识方法论的科学,它是使用一套特有的统计思维逻辑和科学的分析方法来认识不确定现象的数量规律性的方法论科学。而医学统计学则是医学科学研究方法的应用统计学,可以帮助人们占有信息,达到去伪存真、去粗取精,正确认识医学领域相关规律性的目的,对培养医学生科研能力具有特别重要的作用。因此,医学统计学已成为医学各专业各层次学生的必修课,是医学学生必须掌握的工具性学科。

本书为临床医学专业类本科生的医学统计学教学用书。作为教材,不仅注重知识的系统性,还要考虑实用性。因而,本书坚持理论联系实际,以解决实际问题为目的,以临床实例为切入点,以避免误用为警示。全书共九个章节,构成临床医学学生应掌握的医学统计方法基本知识体系。在内容安排上,力求结合医学临床工作实际,突出临床医学特点,较详细地介绍临床诊断与疗效评价方法。此外通过常见错误案例分析,提高学生灵活运用与正确应用医学统计学知识的能力。

参加本书编写的有王润华教授、钟晓妮教授、张菊英教授、方亚教授、邓丹副教授、王学梅副教授、宋桂荣副教授、高晓凤副教授、张燕讲师、叶孟良讲师、张星光讲师、陈卫中老师以及高菲菲老师,张燕担任秘书工作。

本书编写过程中,得到重庆医科大学公共卫生与管理学院、四川大学华西公共卫生学院、厦门大学公共卫生学院、内蒙古医学院公共卫生学院、大连医科大学公共卫生学院、川北医学院预防医学系以及辽宁医学院公共卫生管理学院的大力支持,特此表示衷心感谢。

限于编者的水平以及时间所限,难免存在不足之处,恳请读者与使用本书的老师与同学们提出宝贵意见。

钟晓妮

2013年5月于重庆

目 录

前言

第一章 绪论	(1)
第一节 医学统计学的内容、学习方法和学习态度	(1)
第二节 医学统计学的几个重要基本概念	(2)
一、总体与样本	(2)
二、观察性研究与实(试)验性研究	(3)
三、同质与变异	(3)
四、变量与资料	(4)
五、参数与统计量	(5)
六、概率与频率	(5)
七、随机误差与系统误差	(6)
第三节 医学统计工作的基本步骤	(7)
一、医学研究的设计与规划	(7)
二、预调查和预实验	(7)
三、搜集资料	(8)
四、清理资料	(8)
五、分析资料	(8)
第二章 医学研究统计设计	(9)
第一节 实验性研究统计设计	(9)
一、实验性研究的基本要素	(9)
二、实验性研究的基本原则	(11)
三、系统误差与控制方法	(15)
第二节 临床上常用的观察性研究统计设计	(16)
一、观察性研究统计设计	(17)
二、几种常见观察性研究设计	(17)
第三章 估计性分析	(19)
第一节 常见连续型随机变量的分布及特征点估计	(19)
一、频数分布表的制作	(19)
二、正态分布	(21)
三、正态分布定量变量资料特征点估计	(23)
四、对数正态分布定量变量资料特征点估计	(26)
五、其他偏态分布定量变量资料特征点估计	(27)
第二节 常见离散型随机变量的分布及特征点估计	(29)
一、二项分布	(29)

二、Poisson 分布	(30)
三、定性变量资料特征点估计	(31)
第三节 抽样研究及其常见统计量的抽样分布	(35)
一、抽样误差	(35)
二、 t 分布	(36)
三、 χ^2 分布*	(38)
四、 F 分布*	(39)
第四节 总体参数的区间估计	(40)
一、总体均数的区间估计	(40)
二、总体概率的区间估计	(42)
第五节 医学参考值范围	(42)
第六节 二项分布与 Poisson 分布的应用*	(44)
一、二项分布的应用	(44)
二、Poisson 分布的应用*	(45)
第四章 组间统计量差异比较性分析	(49)
第一节 假设检验的基本思想与步骤	(49)
一、假设检验的基本思想	(49)
二、假设检验的基本步骤	(50)
第二节 t 检验	(51)
一、单样本资料的 t 检验	(51)
二、配对样本资料的 t 检验	(52)
三、两独立样本资料的 t 检验	(53)
四、方差不齐时两样本资料的 t' 检验	(54)
五、假设检验的两类错误	(56)
六、假设检验的注意事项	(56)
第三节 方差分析	(58)
一、方差分析的基本思想及应用的基本条件	(58)
二、完全随机设计资料的方差分析	(60)
三、随机区组设计资料的方差分析	(61)
四、重复测量设计资料的方差分析*	(63)
五、多个样本均数的多重比较	(64)
第四节 χ^2 检验	(66)
一、四格表资料的 χ^2 检验	(66)
二、配对四格表资料的 χ^2 检验	(70)
三、行 $\times$ 列表资料的 χ^2 检验	(70)
第五节 秩和检验	(72)
一、Wilcoxon 符号秩检验	(73)
二、两独立样本比较的 Wilcoxon 秩和检验	(77)
三、多个独立样本比较的秩和检验	(80)
四、多个相关样本比较的秩和检验	(82)

五、多个样本两两比较的秩和检验	(84)
第五章 两变量间的关联性与趋势性分析	(91)
第一节 概述	(91)
第二节 直线相关与直线回归分析	(93)
一、直线相关分析	(93)
二、直线回归分析	(96)
第三节 资料不满足直线趋势模型条件时的趋势性分析*	(101)
一、秩相关	(101)
二、级序相关分析法	(103)
三、趋势性分析	(107)
第四节 两变量一致性分析*	(108)
一、定量测定的一致性分析	(108)
二、等级测定的一致性分析(Kappa 分析)	(108)
第五节 两变量的关联性分析*	(110)
一、结果变量为正态定距(定量)变量的关联性分析	(110)
二、结果变量为分类变量的关联性分析	(110)
三、结果变量为等级变量的关联性分析	(111)
四、两变量关联性分析的注意点	(111)
第六章 诊断试验评价	(113)
第一节 概述	(113)
一、诊断试验与筛检试验	(113)
二、诊断试验结果与疾病状态的关系	(114)
三、诊断试验评价的几种情况	(114)
第二节 诊断试验的评价设计	(115)
一、评价的设计程序	(115)
二、诊断试验评价内容和指标	(117)
第三节 ROC 曲线的制作与分析	(121)
一、ROC 曲线的基本形式和作用	(122)
二、ROC 曲线下面积计算及其假设检验	(122)
第七章 临床疗效的评价	(124)
第一节 临床疗效研究概述	(124)
一、临床疗效研究的特征	(124)
二、临床疗效研究常见设计类型	(125)
第二节 疗效评价指标与内容	(126)
一、临床疗效评价指标	(127)
二、临床疗效研究统计分析的内容	(128)
第三节 优效、等效和非劣效性试验及其假设检验	(129)
一、优效、等效和非劣效性试验的含义	(129)
二、优效、等效和非劣效试验中的假设检验	(130)

第四节 临床疗效研究设计与评价中应注意的问题	(131)
一、研究对象的特征和数量	(131)
二、分析数据集	(132)
三、统计学方法的正确选择	(132)
四、假设检验中统计学结果的正确解释	(133)
第八章 统计表与统计图	(134)
第一节 统计表	(134)
一、统计表的结构	(134)
二、统计表的种类	(135)
三、编制统计表的注意事项	(135)
第二节 统计图	(135)
一、绘制统计图的基本要求	(136)
二、常用统计图的实用条件与绘制	(136)
第九章 常见错误案例分析	(142)
第一节 概论	(142)
第二节 常见错误案例分析	(142)
附录一 统计用表	(154)
附录二 汉英名词对照表	(176)

第一章 绪 论

医学统计学(medical statistics)是一门医学科学研究方法学。从19世纪概率论与数理统计学开始应用于医学研究起,医学学科的形成和发展经历了200余年的历程。承续长期的时间和积淀,如今已形成了一门具有完整内容架构,理论与实践密切结合和交融的知识系统。它蕴涵丰富的科学思想和方法对于指导医学科技工作者选择医学研究课题,进行医学科研设计,确定质量控制方法,学习如何搜集、清理、识别和分析资料,充分发掘资料信息,在逻辑顺畅和有充足可信依据基础上引出客观结论,以及辨偏、减误、克混杂等方面都起着极其重要的作用。

医学统计学与流行病学、计算机应用、外语和文献检索等共同铸造了医学生必备的工具学科知识链。这对于培养医学生和医务工作者终生学习和汲取文献经验,提高业务素质起着导向和支撑作用。因此,医学统计学是医学各专业各层次学生的必修课。

第一节 医学统计学的内容、学习方法和学习态度

医学统计学的内容结构、章节编撰、理法深浅、叙述繁简主要取决于学习对象的层次和实际需求。作为临床医学本科类医学统计学教材,可以根据具体情况纳入以下内容。

统计学基础理论:包括随机变量及常见随机变量的分布,统计量和统计量的抽样分布,数学期望与方差,参数的点估计和区间估计,临床参考值范围的制定,假设检验的基本理论和适应不同性质资料、适应不同条件、适应不同设计方案的假设检验方法,两变量关联性分析与趋势性分析,观测结果的一致性分析,医学现象或疾病的聚集性分析,以及两种或多种认知方式的协调性分析,统计图表的制作等。

医学科研思想方法学包括医学科研设计的基本理论和方法分类,应详细、具体、缜密考虑和斟酌的基本要素、应严格遵循的基本原则,适合不同研究方式和研究思路的设计方案及对应统计分析方法,常用的医学评价方法等。

关于多元统计分析方法,可以根据具体情况机动安排,本书未作专题介绍。

学习方法上,理论与实践密切结合,学以致用、循环往复;承前启后、温故纳新;分析样本信息与推断总体特征和引出结论密切结合,有的放矢、不偏离研究目标;选用统计分析方法与注意适用条件密切结合,深入体会科学真理的相对性、适用性;一般资料分析与发掘信息相结合,锻炼深入思维、灵活思维、启迪研究潜能。如上这些都是十分重要的经验总结。特别值得提倡的是根据学习者个人特点,创造和实施适合自身的学习方法更可获得事半功倍的效果。一般而言,认真听课、做重点和难点笔记、课外复习、阶段总结、深入思考、多提问题、及时求解、结合实际等是学好医学统计学的必由之路。医学统计学的实践性很强,必须学会将学到的医学统计思想和方法应用于解决实际问题,多加锻炼,才能游刃有余;持之以恒,才能做到“识之准确、析之有理、处之合规、论之有据”,才能避免“一听就懂、一过就忘、一用就错”的窘境。

总之医学统计学是培养医学生和医学科技工作者科学思维方法和思维逻辑的重要知

识结构。它是指导我们围绕医学科学研究目的和要解决的实际问题,进行研究设计,指导我们搜集、清理、识别和分析资料;指导我们依据自身的研究条件、研究技术、研究过程、研究因素、指标和数据以及质控水平引出结论,并能为同条件下的新的实践所验证,具有可重复性和再现性的理论与方法。其内涵与人类认识世界(表象与水平、关联与协变、机制与规律)和改造世界(规划与设计、干预与评价、预测与控制)的框架相对应,且协调一致。因此认真对待,刻苦学习,充分使用,在实践中不断提高水平是每一名学习者应具有的正确认识和态度。

第二节 医学统计学的几个重要基本概念

这里介绍的是在医学统计学具有共性的,经常应用的名词概念。在其他章节的内容中不再做重复的涵义解释。

一、总体与样本

总体(population)是依据研究目的所指向的全部同质研究对象的某项或某些专业指标(变量)测量值的集合。样本(sample)是从总体中抽取的有代表性的部分同质研究对象某项或某些对应研究指标的测量值子集合。抽样研究的目的在于了解或推断目的总体特征。

要做到顺利了解和推断总体特征,首先要求样本对总体有代表性。什么样的样本对研究的目标总体具有好的代表性呢?这就要求样本至少具有以下三个特点。

(1) 是随机抽样获得的样本。随机抽取样本是依据总体规模和抽样框架(sampling frame)显示的总体多种结构特征,采取相应方式使每一观察单位都有同样的机会被抽取进入样本,这样获得的样本成为随机样本(random sample)。抽样方式常用单纯随机抽样(simple random sampling)、系统随机抽样(systematic sampling)、分层随机抽样(stratified random sampling)、整群随机抽样(cluster random sampling)以及多种方法结合应用的抽样方法,如分层整群抽样和多阶段抽样。随机抽样的目的是为了减少样本对总体的偏性(bias),增强代表性。

(2) 样本中具有足够的研究对象数。医学研究面对的是具有变异性的研究对象和指标,且不同指标测量值变异度大小不一致。这使得医学研究必须面临样本含量(sample size)的抉择。一项研究中,研究对象数过少,样本指标数据水平容易偏离总体水平,其变异度很不稳定;样本研究对象数过多时,又会带来成本投入的巨大和质量控制的困难。因此医学科学研究最好是做到“研究对象数的恰当和足够”。其科学含义是指能真实回答研究目的提出的问题,且具有节约的特点。科学样本含量是在研究者提出一定条件集合(估计或鉴别参数的误差或差别、指标变异度,获得真阳性结果的把握度,以及单、双侧分析和I类错误概率)时,经计算得到的,因此是有条件的,可以设计估计的。对样本含量规模的理解没有一定的严格标准。一般在方法学研究的重复测定中, $n > 30$,即为大样本,而在独立个体的抽样研究中,如指标变异度小, $n > 50$ 可算较大样本;若变异度较大,则至少 $n > 100$ 。此外,应注意样本含量越小,越不容易发现统计量(表示样本特征的量,statistics)与参数(表示总体特征的量,parameter)或统计量之间的统计意义的差别。因此经科学计算的样本含量在研究者不期望获得具有统计学意义差别时,是足够的样本含量;在研究期望获得具有统计学意义时,则是在规定条件下的最小样本含量。

(3) 样本的结构特征,特别是那些影响研究结果的结构特征应与总体的结构特征相一致。虽然随机抽样关注了总体与样本结构特征问题,只是从宏观方面,如人群分布、地区分

布、年龄分布等,更细微的问题也需给予重点关注。进行动物实验研究时,随机分组要注意品系、性别、体重、健康状态等;研究对象为某病患者群体时,抽样要依据规定的纳入标准和排除标准,分组要注意病情、病期、病型、并发症、年龄及已治疗情况,有时也有性别、遗传及其他社会学特征的均衡;对于社区人群,主要指性别、年龄、经济收入、文化程度、职业状况、民族、疾病健康状态、所处环境条件和医保状况等。

原则上讲,非随机抽样获得的样本是不能用来对目的总体特征进行推断的。但有时如样本含量大,且样本结构特征在主要方面能够反映总体特征时,推断时应十分谨慎,且必须认真分析样本可能存在的偏性。

临床试验和观察性研究中切忌以虚假“随机”用词来掩饰非随机和张冠李戴。

二、观察性研究与实(试)验性研究

对医学科学研究进行分类的方式很多,一般可以分为两大类,即观察性研究(observation study)和实验性研究(experimental study)或试验性研究(trial study)。研究对象为动物,一般称为实验研究;研究对象为社会人群(包括患者人群),一般称试验研究。试验研究又包括临床试验研究和社区干预试验研究。观察性研究与实验(试验)研究的根本区别在于:①观察性研究中,研究者不能将随机抽取的研究对象进行主动随机分组,只能按其固有的自然和社会学特征分组。而实验(试验)研究中,研究者可以主动地将随机抽取的研究对象进行随机分组,使它(他)们有同等机会被分配到观察因素水平组和对照组。②观察性研究中,研究者只能客观地、被动地记录被研究对象的各种状况和社会、态度、行为特征,以及受到的各种因素的影响等,不能主动施加给被研究对象某种干预措施;而实验(试验)研究中,研究者可以对研究对象施加某种干预措施,以便在一定条件下测试这种干预措施表现在不同因素水平下的客观效应指标值。

由于以上两个根本性区别,导致两种研究类型在其他方面也会呈现出不同的倾向性特征。如观察性研究,研究对象所处自然环境和社会环境影响因素较多,且研究者无法严格控制,故研究所需样本含量较大,分析方法必须包括多因素研究分析方法,也必须重视偏性分析,归因结论较难确定。在实验(试验)性研究中,由于研究者可以严格控制对比组间均衡性,明确干预处理因素方式、时间、剂量和水平等,故其他混杂(confounder)因素影响较少,研究精确度相对较高,必然呈现所需样本含量较少,分析方法多采用单因素分析法,在均衡组间对比条件下,归因结论较易确定。

观察性研究和实验(试验)性研究还可以进一步细分为许多指导实际工作的研究方式类别。

三、同质与变异

同质(homogeneity)是指符合研究目的规定的总体特征的研究对象的条件组合。这种条件组合同时也是随机抽样中样本对象的纳入标准和排除标准。不同的研究目的对总体和样本同质性的规定是不一样的,各具针对性。这在研究不同疾病患者群的治疗试验中是明确而突出的。研究对象的“同质性”不满足,必然给研究过程和结果带来观察因素以外的混杂因素干扰,可能导致研究结果的偏性。因此,在临床试验研究和结果报告中十分强调对比组间的均衡性分析。这是为了避免将抽样中的某种“不同质”进一步带到对比组中去,

造成对比组条件的不同,给分析结果时的“归因”造成混淆。

变异(variation)是指同质研究对象某效应指标测量值的波动性。这里的变异不是绝对的含义,而是相对的含义。医学研究中,一般称“效应指标”为结果变量,称处理或观察因素为原因变量。研究对象是原因变量的承受者,同时也是结果变量的反映者,是双变量的载体。一项严肃、严谨的医学科学研究必定要认真对待原因变量、研究载体和结果变量,故这三者称为医学科学研究设计中的三个基本要素。

四、变量与资料

变量(variable)在数理统计学中称为“随机变量”(random variable),是指在随机试验中,可被测定的,具有一定概率分布的量。它的取值是随机事件(random event)的数量化。随机变量按其取值形式的不同,一般可分为离散型随机变量(discrete random variable)和连续型随机变量(continuous random variable)。二者均可分别用概率分布函数表达其概率分布性质和特征。随机变量具有两个基本特征,一是取值的随机性;二是取值的统计规律性。

离散型随机变量是指该随机变量的有限可能的取值和取值概率是一一可列的。如下面的分布列表达:

$$\begin{array}{ccccccc} X & X_1 & X_2 & X_3 & \cdots & X_n \\ P(X=X_j) & P_1 & P_2 & P_3 & \cdots & P_n \end{array} \quad (1.1)$$

离散型随机变量的分布函数可记作

$$F(X) = P(X \leq X_j) = \sum_{X \leq X_j} P(X = X_j), \quad \sum_{X_j=1}^n P(X = X_j) = 1 \quad (1.2)$$

常见的离散型随机变量分布有二项分布、泊松分布、超几何分布、几何分布等。

连续型随机变量是指存在某非负函数(概率密度函数, probability density function) $f(X)$, 使随机变量 X 取值为任意区间 (a, b) 的概率为

$$P(a < X < b) = \int_a^b f(X) d(X), \quad \int_{-\infty}^{\infty} f(X) d(X) = 1 \quad (1.3)$$

常见的连续型随机变量分布有均匀分布、正态分布、对数正态分布、指数分布等。

对随机变量分布特征的表达和描述一般采用它们的特征参数,常用的是数学期望 $E(X)$ 和方差 $V(X)$ 。 $E(X)$ 是分布的均值(或平均水平), $V(X)$ 是分布的离散度(或取值变异程度)。

离散型随机变量在满足 $\sum_{j=1}^{\infty} |X_j| P_j < \infty$ 条件下,其数学期望定义为 $E(X) = \sum X P(X)$, 方差定义为 $V(X) = E[X - E(X)]^2$ 。

连续型随机变量在满足 $\int_{-\infty}^{\infty} |X| f(X) d(X) < \infty$ 条件下,其数学期望定义为 $E(X) = \int_{-\infty}^{\infty} X f(X) d(X)$, 方差仍定义为 $V(X) = E[X - E(X)]^2$ 。

医学统计学侧重于数理统计原理和方法在医学中的实际应用,基于应用中涉及的实际问题广泛,故对变量的分类和称谓更趋于复杂化。医学统计学中可以按三种情况对随机变量命名。第一,按其所起的作用命名,可有原因变量、结果变量、混杂变量、协变量;第二,按其性质命名,可有定量变量(数值变量)、定性变量(分类变量)、等级变量;第三,按其直接可测性,可有表象变量、潜变量。此外,依据变量取值排列是否有序或是否与时间顺序相一

致,可分为无序变量和有序变量,有序变量即前述等级变量。

随机变量取值表现在一定研究条件下的对同一样品同时多次检测或按时间序列检测获得的资料称为重复测量资料(repeated measurement data),在医学检测方法学研究和临床上经常涉及该类资料。随机变量表现在研究条件下对群体独立样品各1次的检测结果取值称为独立测量资料(independent measurement data),它是医学抽样研究常见的结果。这两种情况的资料在统计分析方法上是有区别的,要特别给予注意。识别资料的性质在医学研究资料分析中是很重要的,因为它决定着资料分析的方向。

从性质上区分,医学统计资料可区分为定量变量资料或称计量资料(measurement data),定性变量资料或称计数资料(count data)。在定性变量资料中,又可细分为对立的二值分类变量资料、互不相容的多值分类变量资料、有序等级分类变量资料。对资料性质不熟悉的医务工作者或学生常常受到三种误导而曲解资料的性质。第一种是受到数学概念的约束,即认为凡连续型随机变量取值集合都是定量变量资料,而离散型随机变量取值必为定性变量资料;第二种是受到医学研究对象某指标测定量纲(理解为度、量、衡单位)的影响,认为有明确量纲的,一定是定量变量资料,而表面上无直观量纲的,一定是定性变量资料(其实“量纲”的含义可以很广泛,它可以表现为某计数值/某观察单位);第三种是受到医学指标通俗名称的影响。如听到或看到诸如“××率”或“××比”,即认为一定是定性变量资料,殊不知这种情况体现在群体就会出现一群 N 个“率”或“比”,它仍为定量变量资料。因此识别资料的性质主要依据是否客观存在“明确的数或值/每一个观察单位”,或先验的“类别属性/每一个观察单位”。前者在群体即构成定量变量资料,后者在群体经归类点数即构成定性变量资料。如果属于后者情况且类别属性为有序的等级分类,则群体归类点数即取得等级资料。在特殊情况下,必要时可以按需要将定量变量资料转化为等级资料或二分类定性变量资料。

五、参数与统计量

医学统计学中的参数(parameter)是指描述总体或总体分布的特征数,常用希腊字母表示,如前面提到的总体分布的数学期望与方差。统计量(statistics)是指描述样本或样本分布的特征数,常以拉丁字母表示。

在理论模拟实验研究中,参数可以人为给定或使其改变,故常称为参变量。在实际医学研究中,目的总体往往无穷大或具可变性,无法获得其确切真值,这就必然提出如何通过随机样本统计量估计相应参数的问题。统计量是样本特征数,它不等于参数,存在着抽样误差(sampling error)。要利用统计量估计对应参数,必须首先对统计量本身有一些特性要求,那就是统计量应满足无偏性、一致性和有效性。它们各自的含义如下。

若令 θ 为参数, $\hat{\theta}$ 为统计量,定义 $E(\hat{\theta}) = \theta$ 为无偏性, $\lim_{n \rightarrow \infty} P\{|\hat{\theta} - \theta| < \varepsilon\} = 1$ 为一致性(ε 为任意无穷小量), $S_{\hat{\theta}_1}^2 < S_{\hat{\theta}_2}^2$, $\hat{\theta}_1$ 较 $\hat{\theta}_2$ 更有效($S_{\hat{\theta}_i}^2$ 表示样本方差)。

在以上条件下,统计量、统计量的抽样误差、统计量围绕对应参数的波动或分布规律是估计参数的三个必要条件。

六、概率与频率

随机变量的取值称为随机事件。对于离散型随机变量,随机事件出现的可能大小的度量称为概率(probability)。对于连续型随机变量,随机事件只具有一定的概率密度(probability

density), 只有随机变量在一定的区间取值时,才具有分布概率。

历史上人们对概率的认识和理解主要有三个方面。一是古典概型的认识,它起源于17世纪对赌博输赢的估计。它假定样本空间中的样本点只有有限个,而且每个样本点出现的概率完全相同。最典型的实例是投掷质地均匀的骰子。它仅有6种点数可能,且每种点数出现的概率完全相同。二是从频率(frequency)来认识概率。如果存在完全相同条件下,可进行多次重复试验的情况。若重复试验 n 次,实验者感兴趣的有利事件 A 发生 m 次,则记 m/n 为事件 A 发生的频率,记为 $f(A)$ 。虽然每重复 n 次实验中,事件 A 发生的频率有所波动,但当 n 逐渐趋向于无穷大时,频率 $f(A)$ 会逐渐趋向于一个稳定值,这个稳定值就是概率。三是主观经验的认识和估计,特别是长期的专业观察会使人们对那种难以用实验的结果来代表的事件发生的可能性给出某种有经验理由的估计,具有一定的准确性和可信性。随着这些认识的深入和发展,孕育和促进了统计学中两大学派的产生、成长与成熟,这就是频率学派和贝叶斯(bayes)学派。

上述三种对概率的认识虽然基本条件不同,思维方式不同,但共同点都认定概率的取值范围是 $[0,1]$ 。一般称概率等于零的随机事件为“不可能事件”,概率等于1的事件为“必然事件”。除以上两种极端情况外,一般随机事件 A 的发生概率是 $0 < P(A) < 1$ 。统计学中,人们基于大量的实践和经验,共同认定 $P \leq 0.05$ 的随机事件为“小概率事件”,而且“小概率事件”在一次随机试验中是不太可能发生的,这种不太可能发生的小概率事件被统计学者们用来表达一种通过样本推断总体的思维方法。如果我们把样本统计量也看成随机事件,基于样本统计量围绕参数形成抽样分布的特征和规律,小概率事件总是分布在远离参数的抽样分布的两侧尾部。可以人为地规定出现小概率事件的样本统计量临界点,比临界点更远离参数的全部样本随机事件就形成了拒绝总体范围,可称“拒绝域”。出现达到临界点或比临界点更偏的情况虽有发生,但发生可能性很小。此时如有发生,则据此拒绝总体犯“异质”错误的可能性很小,因此有理由拒绝“同质性”。这就是统计学中用以判断样本与总体差别意义的小概率推断法。

七、随机误差与系统误差

医学研究的对象绝大多数情况是有生命的个体与群体,以及来自这些生命体的生物材料,只有少数情况面对无生命的材料。由于生命对象都是处在一定自然环境条件下或(和)处在一定社会环境条件下,受到的各种影响因素作用不尽一致。即使是一同质群体处在相同自然与社会环境条件下,受到基本一致的各种因素的综合影响,由于他们个体存在的耐受性和反应差异性,也会表现出生物的、社会的、心理的不同应激,这必然导致同质样本群体某种检测指标(变量)取值的差异性,这是造成随机样本统计量差异性的根本原因。它表明在抽样研究中,生物的、社会的、心理的各种指标样本测量值集合都存在着抽样误差,这是对由单一独立个体组成的群体而言。另一种情况属于方法学研究范畴,即在相同条件下,对同一样品或标准品进行多次重复测定,这时对象是唯一的,条件是严格的,方法是相同的,是否多个重复测定的结果就一定完全一致呢?并不一定。这是由于影响检测结果的因素会有微小随机波动,只要检测精度足够,就能显示检测的误差。误差客观存在,无法回避,必须正视。了解、认识、掌握与应用误差及其分布规律是医学各专业人员 and 医学生的必要能力。

误差在方法学研究中称为测量误差(measurement error),是指在相同条件下(检测者——man、检测仪器设备——machine、检测用材料——material、检测方法——method、环境条件——environment,简称4M1E),对同一样品或标准品进行多次重复测定,测定值的波

动性和对真值的偏离。它包括两个方面内涵:测定值的随机波动性称为随机测量误差或精密度(precision),也称为随机不确定度;测定值对真值的偏离称为系统误差或准确度(accuracy),也称为系统不确定度。

误差在独立个体群的抽样研究中,随机误差表现为抽样误差(其中也隐含有随机测量误差),它反映样本统计量对参数的随机偏离,形成统计量的抽样分布;系统误差在独立个体群的抽样研究中,习称为偏性,它反映样本统计量对参数的具有倾向性的较大系统偏离。

随机误差在概念上应包括随机测量误差和抽样误差。这种误差在统计学上常应用正态分布(normal distribution)理论来进行分析。系统误差主要是识别其是否存在,寻找产生的原因和及时给予纠正和校正。表现为方法学研究中的系统误差形式常见有恒定系统误差、线性系统误差、非线性系统误差和周期性系统误差;表现为独立个体群抽样研究中的系统误差形式常见有选择性偏倚、测量性偏倚和混杂性偏倚。

实际工作中,有时会由于过失原因产生偶然误差。包括由于工作人员引起的错误操作、错读、错记、错录,甚至缺乏科学道德的故意修改等。这种情况无规律可循,一旦发现,无论测定值有无偏离或偏离大小,必须予以舍弃。

第三节 医学统计工作的基本步骤

医学统计工作的基本步骤主要是指对正式调查研究(survey study)和实验(试验)研究中产生的有关三个基本要素(研究因素、研究对象、效应指标)数据和现场及参考资料信息实施搜集、清理、识别、分析与推断,引出结论并与参考文献相佐证的全过程程序。但从医学研究的角度考虑,事前尚应包括研究的设计与规划、研究条件的准备与优化,预调查与预实验;事后也需要有统计学知识的检查与评价。

一、医学研究的设计与规划

首先应明确研究目的,在这一总前提下,才能有的放矢地进行全面设计与规划。

医学研究要涉及选题,确定框架研究方法(调查与实验)与细化研究方法(时间类别、规模类别、概率类别、方式类别等),准备研究工具与研究条件,确定研究总体、搜集抽样框架、决定抽样方式,拟定技术路线和关键问题解决办法,计划落实研究的三个基本要素,严格规定应遵循的基本原则,拟定各环节质量控制方法与标准,确定资料的搜集方法、资料清理核对方法及整理分组方法,按研究目的匹配规范的资料分析方案和分析方法、分析软件,预定的考察偏性及分析混杂因素的方法,弥补和挽救数据缺失的措施。甚至包括将来的报告与论文的写作框架及参考文献的利用等。

拟定设计与规划要全面考虑科学性、创新性、可行性和有效性,要把握理论方法、资料和推理依据性与准确性。设计中尚应注意为使研究有条不紊地进行,对培训、组织、宣传、后勤等给予充分重视。

科研标书实际上是一份精炼的科研设计计划书,为了加强训练,建议医学本科生最后一年应该争取写一份医学科研标书。

二、预调查和预实验

这一过程的主要目的是探索研究的可行性,完善设计和修正缺陷,并初步摸索研究中

主要内容指标数据的水平和范围,修订样本量、修订问卷,为正式研究奠定扎实的基础,并提出和完备解决问题与困难的措施与思路。虽然是预调查和预实验,执行时应等同于正式调查和正式实验(试验),只不过观测量较小而已。

三、搜集资料

搜集资料(collection of data)要求准确、及时、完整、规范。多数情况下,是边研究、边测试、边收集资料。因此要求记录规范、及时检查、保证质量。对有怀疑和缺失的数字,要及时查错补漏。记录要字迹工整,切忌潦草,要及时编目建档,妥善保存。

四、清理资料

清理资料包括清查资料(cleaning data)和整理资料(sorting data),是按设计要求系统、全面地将分散的原始数据进行清查与整理,使其系统化、条理化,以便为进一步分析资料打下良好基础。

这一步骤是在建立数据库过程中完成的,一般要求两名录人员分别独立完成数据录入,在数据库建立后还应对数据进行核对和逻辑核查。Excel、SPSS、SAS 等软件均可对变量建立二维结构数据库。一般“行”编号表示观察对象或观察单位,“列”表示记录指标或变量。“列”方向的变量可以是表达对应观察对象的各种变量取值或赋值。

整理资料主要是根据研究目的、资料性质和行编号量及取值特征考虑分组。按研究对象的属性特征归组称为质量分组或类别分组,如性别、地区、民族等;按研究对象对应该变量的数量取值大小来归组的称为数量分组或数量频数分组。在整理资料基础上,还可以进行基础性预分析,如类别频数,绘直条图;数量组别频数,绘直方图;计算正态偏度与峰度系数,简单分析等。

五、分析资料

分析资料(analysis of data)时,特别要注意结合研究目的识别资料性质、数据分布、样本大小和非对比参数齐同性。基本方法包括:估计性分析、聚集性分析、组间参数(或非参数)估计值差异的对比性分析、两变量或多变量间关联性分析、两变量间或多变量间的协变趋势性分析,两变量或多变量的一致性或协调性分析。其他尚有如多元统计等其他方法。分析资料原则上依据研究目的按设计方案进行。如果资料信息丰富,且有必要性,尚可考虑进一步挖掘资料信息,这时可以超越设计方案。

在分析资料基础上,要强调依靠自身研究的数据依据和分析结果依据,逻辑顺畅地导出推论和结论。切忌将来成文的报告中只单纯广泛涉及文献,而对自身的研究束之高阁或轻描淡写。

总之以上各个步骤密不可分,环环相扣,前制约后,是一个完整系统。任何一步出现质量问题,都将对最后结论产生影响,甚至导致研究失败。至于谈到形成论文和研究报告,除上述外尚有许多工作要做。请读者深入学习本书,并经过自身实践,定会产生较深刻的体会和感悟。

第二章 医学研究统计设计

医学统计学是现代医学研究中不可缺少的工具,从第一章绪论中我们知道,医学统计工作的基本步骤,即研究设计、收集资料、整理资料和分析资料,而研究设计是最为关键的一步。医学研究设计是指制定周密的医学研究计划,包括专业设计和统计学设计两个部分。专业设计是从专业角度考虑实验的科学安排,它包括选题、建立假设、确定研究对象和技术方法等;统计学设计是按照统计学原则,运用统计学原理和技术,对研究资料的收集、整理和分析进行科学设计,以适宜的人力、物力和时间获得可靠的结论。本章将主要介绍医学研究中的统计设计。

第一节 实验性研究统计设计

医学研究的类型很多,不同的分类标准有不同的研究类型,相应地有不同的研究设计类型。根据对研究对象是否主动进行干预,医学研究一般分为观察性研究(observational study)与实验性研究(experimental study)。

实验的目的是要阐明所要研究的处理因素作用于受试对象后所产生的实验效应。因此,在实验研究的统计设计中,需要考虑处理因素、受试对象、实验效应这三个基本要素,需要遵循随机、对照、重复三个基本原则。对于试验研究,往往还要求遵循盲法的原则。

一、实验性研究的基本要素

实验性研究的三个基本要素分别为处理因素、受试对象和实验效应,如用两种药物治疗高血压患者,比较两组患者舒张压的变化情况,所用药物为处理因素,高血压患者为受试对象,舒张压变化值为实验效应。

(一) 处理因素

1. 处理因素和处理水平 处理因素(treatment factor)是指研究者根据研究目的施加于受试对象,且在实验中需要观察并阐明其效应的因素,包括物理因素、化学因素和生物因素。可采用单一的处理因素,如药物处理;也可以采用多种因素同时处理,如在非药物治疗的基础上加用药物处理。

为了了解处理因素对实验效应的影响特点,常需要将处理因素分为不同的水平,如药物治疗时,可给予不同的剂量或不同的时间。

处理因素数量与处理水平数量形成多种组合方式,如单因素多水平、多因素多水平。

2. 确定处理因素的注意事项

(1) 抓住实验中的主要因素:任何一个实验都包含着多个因素,在设计时需抓住对实验结果影响较大的主要因素,而且除了要考虑各因素单独对受试对象的效应,还要考虑各因素对受试对象的交互作用。

(2) 明确处理因素和非处理因素:处理因素是实验中所要观察和阐明的因素;而参与