

| 全国高等学校医学规划教材 |

YIXUE TONGJI FENXI FANGFA

医学统计分析方法

主审 黄品贤 主编 朱继民



中国科学技术大学出版社

| 全国高等学校医学规划教材 |

YIXUE TONGJI FENXI FANGFA

医学统计分析方法

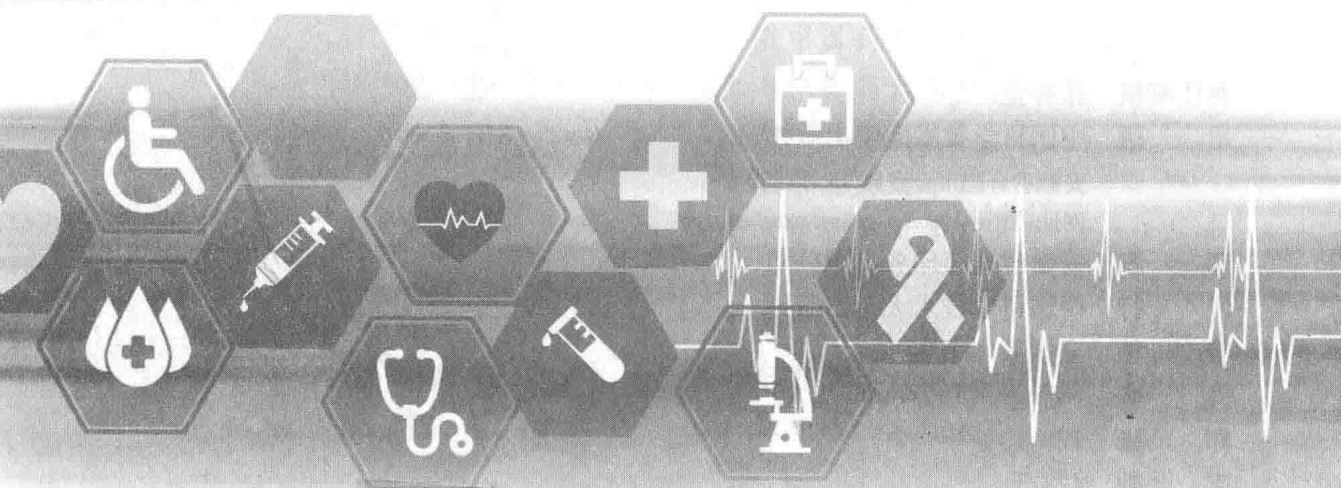
主 审 黄品贤

主 编 朱继民

副主编 闫国立 武

崔 宁 董

魏 沙



中国科学技术大学出版社

内 容 简 介

本书为全国高等学校医学规划教材之一,由6所高等中医药院校联合编写。全书15章,系统地介绍了医学统计分析的基本理论和基本方法,强调资料属性、设计类型和分布特征等在统计分析方法选用中的作用,注重统计分析结果的正确解释与表达,内容丰富,理实交融。

本书以医学科研资料的统计分析过程为主线进行章节安排,以体现统计分析先描述、再估计、最后检验的一般顺序,引导读者树立设计优先的理念,学会统计分析的一般过程与假设检验方法选用的一般思路,熟悉常用统计分析方法的SPSS软件操作与结果解读,非常适合高等医学院校选作相关专业本科教育教材,也可供相关专业研究生和科技工作者学习参考。

图书在版编目(CIP)数据

医学统计分析方法/朱继民主编. —合肥:中国科学技术大学出版社,2016.8
ISBN 978-7-312-04020-7

I. 医… II. 朱… III. 医学统计—分析方法—医学院校—教材 IV. R195.1

中国版本图书馆CIP数据核字(2016)第141646号

责任编辑 张善金

出 版 中国科学技术大学出版社

安徽省合肥市金寨路96号,230026

网址:<http://www.press.ustc.edu.cn>

印 刷 安徽省瑞隆印务有限公司

发 行 中国科学技术大学出版社

经 销 全国新华书店

开 本 787mm×1092mm 1/16

印 张 17

字 数 424千

版 次 2016年8月第1版

印 次 2016年8月第1次印刷

印 数 1—5000册

定 价 35.00元

《医学统计分析方法》

编 审 委 员 会

主 审 黄品贤 (上海中医药大学)

主 编 朱继民 (安徽中医药大学)

副主编(以姓氏笔画为序)

闫国立 (河南中医药大学)

武 松 (安徽中医药大学)

崔 宁 (山东中医药大学)

董 菊 (南京中医药大学)

魏 沙 (湖北中医药大学)

编 委(以姓氏笔画为序)

马晓梅 (河南中医药大学)

王东芳 (山东中医药大学)

王亚云 (安徽中医药大学)

王均琴 (南京中医药大学)

王鸣瑞 (安徽中医药大学)

王瑾瑾 (河南中医药大学)

冯 鑫 (安徽中医药大学)

戎 芬 (上海中医药大学)

孙春阳 (河南中医药大学)

杨 洁 (安徽中医药大学)

吴 娟 (南京中医药大学)

汪婷婷 (安徽中医药大学)

陈学芬 (上海中医药大学)

韩晓春 (山东中医药大学)

秘 书 李 静 李白坤 (安徽中医药大学)

前 言



众所周知,高等学校历来是我国科学研究的主战场之一。随着高等教育教学改革的深入开展,大学生参加科研实践活动已经成为培养高水平人才的必由之路。科研活动离不开统计学,科学的统计设计、恰当的统计分析及合理的专业解释才能够获得高质量的研究成果。

目前,有关医学统计学的教材和专著不少,但多以统计分析方法由浅入深为编写思路进行分类介绍。由于每种统计分析方法有特定的应用场合及适应条件,而实际的统计分析往往是依据研究设计与资料类型去寻求适宜的统计分析方法,这与现行教材和专著的编写思路恰恰是相反的。这可能是学生和科研工作者在寻求具体统计分析方法时经常存在迷茫和困惑的原因之一。本书以资料属性和设计类型为依据进行分类,并且一步到位地介绍某类资料的统计分析全过程,理顺处理医学研究数据资料实际问题时的统计学分析思路,消除选用统计分析方法时的迷茫与困惑,便于实施正确的统计分析。

在本书的编写过程中,我们注重对常用统计方法的基本思想和原理的通俗解释,强调资料属性、设计类型和分布特征等在统计分析方法选用中的作用,注重统计分析结果的正确解释与表达。在不打破医学统计学传统理论框架的前提下,以医学科研资料的统计分析过程为主线,按照基本概念、统计描述、参数估计和假设检验进行章节安排,以体现统计分析先描述、再估计、最后检验的一般顺序。在假设检验部分,以资料属性和设计类型为依据进行章节分类,较为全面地介绍医学科研资料的统计分析。例如,对于完全随机设计两样本数值变量资料的比较,可能用到的假设检验方法有两独立样本 t 检验、两独立样本校正 t 检验、Wilcoxon 秩和检验,这在以前的相关书籍中一般都要涉及两三个章节的内容,而在本书中,我们将其整合到一个章节内,根据应用条件的满足情况分别设置案例,给出分析思路和正确的分析方法。

本书共 15 章,除绪论外,分别介绍了实验研究设计、常用概率分布、数值变量资料的统计描述、分类变量资料的统计描述、统计表与统计图、参数估计、假设检验、完全随机设计数值变量资料的假设检验、配对与随机区组设计数值变量资料的假设检验、析因设计与重复测量设计数值变量资料的方差分析、无序分类变量资料的假设检验、有序分类变量资

料的假设检验、双变量相关与回归等。鉴于实际工作中统计分析方法错用误用现象较为多见的情况,本书还特意安排了“统计分析的常见错误”一章。全书参考学时(含上机实验教学)40~60个。

本书是集体智慧和力量的结晶,也是校际合作、资源共享的创新成果。上海中医药大学、河南中医药大学、湖北中医药大学、南京中医药大学、山东中医药大学和安徽中医药大学等兄弟院校具有丰富教学和科研经验的医学统计学专家们通力合作,精心策划,反复论证,悉心编写,字里行间倾注着他们大量的心血;上海中医药大学医学统计学专家黄品贤教授于百忙中挤出时间认真审阅了本书的全部内容,提出了许多宝贵的建议和修改意见,她为本书的出版同样付出了诸多的心血,尤其是对关键知识点和总体理论架构及学术质量的把关,为本书的科学性和实践指导性提供了保证;安徽中医药大学教务处、中西医结合临床学院领导对本书出版给予了热情的支持和指导,在此一并表示最真诚的感谢。此外,我还要特别感谢我的家人和所有关心本书出版的同事们,他们的支持和理解也为本书如期出版提供了很好的帮助。

本书由朱继民教授担任主编,闫国立、武松、崔宁、董菊、魏沙担任副主编。各章的写作分工如下:第1章朱继民,第2章董菊,第3章王亚云,第4章吴娟,第5章王均琴,第6章马晓梅,第7章陈学芬,第8章王东芳、韩晓春,第9章魏沙,第10章崔宁,第11章及附录A武松,第12章孙春阳,第13章闫国立,第14章王瑾瑾,第15章戎芬。全书由朱继民教授负责总纂和定稿,黄品贤教授担任主审。总校稿由李静、李白坤完成,正文校稿由王鸣瑞、杨洁完成,实验校稿由冯鑫、汪婷婷完成。

本教材的出版得到安徽省2014年高校优秀青年人才支持计划(2014-077)、安徽省精品资源共享课程“医学统计学”项目(2013gxxk055)、安徽中医药大学中医硕士专业学位研究生课程建设项目(2015YJG004)的支持和资助。

本书可作为高等医学院校本科教育的医学统计学教材,特别适合中医药院校本科生使用(案例多源自中医药方面的研究资料),也可作为参考书供医学高等学校研究生以及相关科研工作者使用。书中SPSS软件实现和结果判读方法方面的内容更是本书理实交融的点睛之笔,为读者在该领域的深入学习与研究提供了参考。

尽管我们作了很大的努力,但限于学识能力,书中难免存在瑕疵和不足,恳请同行专家和读者不吝赐教,以便我们在将来修订再版时使相关内容更臻完美。

朱继民

2016年5月26日

目 录



前言	(i)
第 1 章 绪论	(001)
1.1 统计学的定义与研究内容	(001)
1.2 统计学的基本概念	(002)
1.3 资料的类型	(004)
1.4 统计工作的基本步骤	(005)
第 2 章 实验研究设计	(008)
2.1 实验研究设计的基本要素与基本原则	(008)
2.2 完全随机设计	(012)
2.3 配对设计	(014)
2.4 随机区组设计	(016)
2.5 样本量估计	(018)
上机实验 常用实验设计的随机分组	(020)
第 3 章 常用概率分布	(024)
3.1 正态分布	(024)
3.2 二项分布和泊松分布	(029)
3.3 抽样分布	(033)
第 4 章 数值变量资料的统计描述	(037)
4.1 频数分布	(037)
4.2 集中趋势的统计描述指标	(040)
4.3 离散趋势的统计描述指标	(042)
上机实验 数值变量资料的统计描述	(044)

第5章 分类变量资料的统计描述	(046)
5.1 相对数	(046)
5.2 率的标准化	(048)
第6章 统计表与统计图	(052)
6.1 统计表	(052)
6.2 统计图	(054)
上机实验 常用统计图的制作	(059)
第7章 参数估计	(065)
7.1 抽样误差与标准误	(065)
7.2 总体均数的估计	(068)
7.3 总体率的估计	(071)
上机实验 参数估计	(071)
第8章 假设检验	(075)
8.1 假设检验的原理与步骤	(075)
8.2 假设检验的种类	(077)
8.3 两类错误与注意事项	(078)
8.4 正态性检验与数据变换	(080)
8.5 方差齐性检验	(083)
上机实验 正态性检验与方差齐性检验	(085)
第9章 完全随机设计数值变量资料的假设检验	(088)
9.1 单样本数值变量资料的假设检验	(088)
9.2 两样本数值变量资料的假设检验	(091)
9.3 多样本数值变量资料的假设检验	(096)
上机实验 完全随机设计数值变量资料的假设检验	(104)
第10章 配对与随机区组设计数值变量资料的假设检验	(111)
10.1 配对设计数值变量资料的假设检验	(111)
10.2 随机区组设计数值变量资料的假设检验	(114)
上机实验 配对与随机区组设计数值变量资料的假设检验	(120)
第11章 析因设计与重复测量设计数值变量资料的方差分析	(126)
11.1 析因设计数值变量资料的方差分析	(126)
11.2 重复测量设计数值变量资料的方差分析	(128)
上机实验 析因设计与重复测量设计数值变量资料的假设检验	(131)

第 12 章 无序分类变量资料的假设检验	(136)
12.1 完全随机设计无序分类变量资料的 χ^2 检验	(136)
12.2 配对设计无序分类变量资料的假设检验	(146)
上机实验 无序分类变量资料的假设检验	(152)
第 13 章 有序分类变量资料的假设检验	(159)
13.1 完全随机设计有序分类变量资料的秩和检验	(159)
13.2 配对设计有序分类变量资料的秩和检验	(165)
13.3 随机区组设计有序分类变量资料的秩和检验	(167)
上机实验 有序分类变量资料的假设检验	(170)
第 14 章 双变量相关与回归	(174)
14.1 直线相关	(174)
14.2 等级相关	(179)
14.3 直线回归	(183)
14.4 直线相关与直线回归的区别与联系	(194)
14.5 曲线拟合	(194)
上机实验 双变量相关与回归	(199)
第 15 章 统计分析的常见错误	(204)
15.1 统计描述中的常见错误	(204)
15.2 统计推断中的常见错误	(207)
附录 A SPSS 软件简介	(217)
附录 B 统计学报告准则简介	(224)
附录 C 医学统计学附表	(228)
附表 C.1 随机数字表	(228)
附表 C.2 标准正态分布函数 $\Phi(-z)$ 值表	(230)
附表 C.3 二项分布函数 $P(x \geq k)$ 值表	(231)
附表 C.4 泊松分布函数 $P(x \geq k)$ 值表	(233)
附表 C.5 t 界值表	(235)
附表 C.6 F 界值表(方差分析用)	(237)
附表 C.7 χ^2 界值表	(243)
附表 C.8 百分率的可信区间	(244)
附表 C.9 F 界值表(方差齐性检验用)	(247)

附表 C. 10	T 界值表(符号秩和检验用)	(248)
附表 C. 11	T 界值表(Wilcoxon 秩和检验用)	(249)
附表 C. 12	q 界值表	(250)
附表 C. 13(1)	q' 界值表	(251)
附表 C. 13(2)	Dunnett- t 界值表(双侧)	(252)
附表 C. 14	H 界值表(Kruskal-Wallis 秩和检验用)	(253)
附表 C. 15	M 界值表(Friedman M 秩和检验用)	(254)
附表 C. 16	r 界值表	(255)
附表 C. 17	r_s 界值表	(257)
附录 D	中英文术语对照表	(258)
参考文献		(262)

第1章 绪 论



学习目的:了解统计学的工作内容与步骤,熟悉统计学的基本概念与资料类型;学习统计思想,激发学习兴趣。

知识要点:统计学的定义、内容、基本概念,资料类型,统计工作的基本步骤。

医学统计分析方法是从事物同质性和变异性的数量表现(不确定性)出发,通过一定数量的观察、比较和分析,揭示医学问题的分布特征、相互关系及变化规律等(确定性),以指导医学理论与实践的科学方法。随着医学的发展,作为医学科研方法学的医学统计分析,日益受到广大医学工作者的青睐,并得到了广泛的应用。

1.1 统计学的定义与研究内容

1.1.1 定义

统计学(statistics)是关于资料的收集、整理、分析和表达的科学。它从随机现象的数据中提取信息与知识,并研究随机现象的数量规律。统计学是一门十分重要的方法学科,帮助人们去正确地认识世界和改造世界。

理论统计学(theoretical statistics)即数理统计学(mathematical statistics),是以概率论为基础,从理论的角度,对统计方法加以推导论证;其核心内容是统计推断方法,实质上是以归纳法研究随机现象的一般规律。英国三位著名统计学家皮尔逊(Karl Pearson)、戈赛特(William Sealy Gosset)和费希尔(Ronald Aylmer Fisher)对数理统计学的贡献卓著,被誉为现代数理统计学的“三剑客”。

应用统计学(applied statistics)是数理统计学的原理与方法在不同学科领域的具体应用。如生物统计学(biostatistics)、医学统计学(medical statistics)、中医药统计学等(statistics for chinese medicine)。应用统计学起源于德国的社会统计流派(大陆派),其中克尼斯(Karl Gustav Adolf Knies)、恩格尔(Christian Lonrenz Ernst Engel)和梅尔(Georg von Mayr)三位德国著名统计学家的贡献很大,被誉为社会统计学的“三剑客”。

医学统计分析(medical statistics analysis)属于应用统计学的范畴,它是应用数理统计的

原理与方法,研究医学领域资料的收集、整理、分析与表达的科学。是医学科研十分重要的工具和手段。

1.1.2 研究内容

统计学的研究内容主要包括统计设计和统计分析。

统计设计是统计工作的首要环节,是根据研究目的和研究对象的特点,确定统计设计类型、估算所需研究对象的数量(样本量)、确定研究对象的随机抽样方法与随机分组方法、明确统计指标和分析方法等的统计活动。

统计分析主要包括统计描述、统计推断和关系分析。在实际工作中,应依据研究目的、资料属性、设计类型等,选用恰当的统计分析方法进行资料处理,以保证统计分析结果的可信性和结论的科学性。统计分析方法没有优劣之分,只有用得是否恰当之说。

1.2 统计学的基本概念

学习统计学的基本概念是掌握统计分析方法的前提;正确理解统计学的基本概念,有助于领会统计学思想,从而培养统计思维;有助于开展周密的研究设计和实施合理的统计处理,也便于和统计学专业人士进行交流。

1.2.1 同质与变异

观察单位(observed unit)或个体(individual)是统计研究的基本单位。根据不同的研究目的,观察单位可以是一个人、一只动物或一个细胞,也可以是一个家庭、一所医院或一个地区。

1. 同质

同质(homogeneity)即性质相同,指观察单位间被研究指标的影响因素相同。由于医学研究观察指标的影响因素往往多而复杂,甚至未知;所以同质是相对的。理论上,同质要求影响研究指标的各种因素,在各观察单位上完全相同;但实际上,同质要求对研究指标影响较大,且可控的主要因素在各观察单位上表现相同或相似。例如,要研究某年某地区儿童的身高,那么年份、地区、年龄、性别、民族等是可控的主要影响因素,而遗传因素和膳食情况等则难以控制。

2. 变异

变异(variation)是指在同质基础上,不同观察单位间某研究指标值的差异。例如,某年某地区的儿童即使年龄、性别和民族都相同,其身高值却不尽相同。

1.2.2 总体与样本

1. 总体

总体(population)是根据研究目的所确定的同质观察单位某项变量值的集合。例如,研究某年某地 60 岁以上男性居民的血糖水平,则该地这一年所有 60 岁以上男性居民的血糖测量值就构成了研究总体,该地这一年每位 60 岁以上的男性居民就是观察单位。

根据研究总体的观察单位数是否明确,总体常分为有限总体和无限总体。有限总体中观察单位数是有限的或可知的,无限总体中的观察单位数则是无限的或不可知的。实际工作中,

对总体特征与性质的认识常常采用抽样研究的方法进行,而不是对总体中的观察对象进行逐个研究。

2. 样本

样本(sample)是从总体中随机抽取的具有代表性观察单位某项变量值的集合。样本所包含的观察单位数称为样本量。抽样研究(sampling study)是从总体中随机抽取样本,通过对样本的定量或定性测量结果来推断总体的特征与性质。为保证抽样研究结果的正确性,要求样本有良好的代表性。

所谓代表性是指样本某些指标的分布特征与总体相应指标的分布特征相一致;二者的一致性程度越高,样本的代表性就越好。为保证样本具有良好的代表性,应遵循随机抽样和样本量满足统计学要求这两个原则。

1.2.3 参数与统计量

1. 参数

反映总体的统计指标称为参数(parameter),一般用希腊字母表示。如 μ (总体均数)、 σ (总体标准差)、 π (总体率)等。就某总体而言,其参数一般是固定不变的,具有唯一性;参数通常是未知的。

2. 统计量

反映样本的统计指标称为统计量(statistic),一般用英文字母或拉丁字母表示。如 $\bar{x}$ (样本均数)、 s (样本标准差)、 p (样本率)等。与参数不同,统计量是随机的,通常会随着抽样方法、样本量、测量方法等的改变而不同;但统计量的分布是有规律的,是可计算出来的。

1.2.4 误差

误差(error)是指观测值与真实值之差以及样本统计量与总体参数之差。按照误差产生的原因和性质,常分为随机误差和非随机误差两类。

1. 随机误差

随机误差(random error)是指由于一些偶然因素或随机抽样造成的误差,包括随机测量误差(random measurement error)和随机抽样误差(random sampling error)。随机误差很难完全避免,其大小无方向性;但其分布有一定的规律,是可以估计和控制的。

2. 非随机误差

非随机误差(nonrandom error)是指由于一些非偶然因素造成的误差,包括系统误差(systematic error)和过失误差(gross error)。其中,系统误差具有方向性,研究中应尽量避免;过失误差通常是由于不够细心造成的错误判断和记录,研究中应杜绝此类误差。

1.2.5 频率与概率

1. 频率

若在试验条件相同的情况下,对某随机事件进行 n 次重复试验,那么出现某种结果的次数 k 与总试验次数 n 之比,即为频率(frequency)。当 n 足够大时,频率趋向于概率。

2. 概率

概率(probability)是反映随机事件发生可能性大小的度量,用 P 表示,其取值范围为: $0 \leq P \leq 1$,常用小数或百分数表示。 P 越接近 1,表示某事件发生的可能性越大; P 越接近 0,表示某事件发生的可能性越小。 $P=1$ 时为必然事件, $P=0$ 时为不可能事件,二者是随机事件的特例。

3. 小概率事件

在统计学中,习惯上将 $P \leq 0.05$ 或 $P \leq 0.01$ 的随机事件称为小概率事件(small probability event)。认为小概率事件在一次随机试验中不大可能发生,这就是小概率原理,它是统计推断的重要原理。基于小概率原理作出的统计推断能保证很高的正确性,但也存在推断错误的风险。

1.3 资料的类型

资料(data)又称数据,由变量及其变量值组成。统计资料一般分为数值变量资料和分类变量资料两大类。资料类型也称资料属性,是选取恰当统计分析方法的重要前提。

1.3.1 变量与变量值

1. 变量

变量(variable)是指观察单位的某项特征或属性,为研究者关心的各种指标。例如病人的年龄、病情和中医证型等,都可视为变量。

2. 变量值

变量值(value of variable)是指变量的测量值或观察值。例如某患者,男性,62岁、病情重、气血亏虚证等,就分别是该患者性别、年龄、病情和中医证型的变量值。

1.3.2 数值变量资料

1. 定义

数值变量资料(numerical data)又称定量资料(quantitative data)或计量资料(measurement data),是借助某种工具对观察单位的某项变量值的大小进行定量测量所得到的资料。例如,测量 125 名冠心病患者的甘油三酯水平所获得的资料即为数值变量资料。

2. 特征

数值变量资料通常具备三个特征:① 变量值有大小之分;② 一般有度量衡单位;③ 变量值的获取通常需要借助测量工具。

1.3.3 分类变量资料

分类变量资料(categorical data)的变量值是定性的,表现为互不相容的类别或属性。常分为无序分类变量资料(unordered categorical data)和有序分类变量资料(ordinal categorical data)两类。

1. 无序分类变量资料

无序分类变量资料又称定性资料(qualitative data)或计数资料(enumeration data),是先将事物按不同的属性归类,再清点每一类的数量多少所得到的资料。依据类别数的不同,又可分为:二分类资料,如性别(男、女)、治疗效果(有效和无效)等;多分类资料,如血型(A、B、O、AB)、职业(工、农、商、学、兵)等。无序分类变量资料的变量值无大小之分,也无度量衡单位。

2. 有序分类变量资料

有序分类变量资料又称等级资料(ordinal data)或半定量资料(semi-quantitative data),是将事物属性按等级顺序进行归类计数所得到的资料,各类别之间有程度的差别,但无度量衡单位。如用某中药复方治疗产后血瘀 110 例,按临床疗效分为治愈 60 例、显效 26 例、好转 15 例、无效 9 例,该资料就是等级资料。

1.3.4 资料类型转换

根据研究目的和统计分析的需要,三种资料类型之间可以相互转换。例如,一组血红蛋白量测量值(g/L)的数值变量资料,可以根据血红蛋白量参考值范围,划分为贫血、正常、增高的等级资料;也可以划分为血红蛋白量正常与异常的二分类定性资料。反过来,亦可将分类变量资料进行数量化,转换为数值变量资料。

就信息含量的丰富程度而言,数值变量资料最高,其次为有序分类变量资料,无序分类变量资料最低。所以在实际研究中,应尽量收集数值变量资料。

1.4 统计工作的基本步骤

统计工作的基本步骤包括研究设计、资料收集、资料整理、资料分析和结果报告与表达等。

1.4.1 研究设计

1. 研究设计

研究设计(research design)是对要开展的研究工作所做的全面设想,拟定一份合理的研究计划,以确保取得的研究结果和结论的科学性、实用性、创新性和先进性。研究设计包括专业设计和统计设计。

2. 专业设计

专业设计(specialized design)是根据研究目的运用专业理论知识来构思研究框架。专业设计内容包括:确定用什么方式、方法验证假说或回答有关专业问题;研究对象的纳入标准、排除标准和剔除标准;研究因素的施加方案;效应指标的选用等。

3. 统计设计

统计设计(statistical design)是根据研究目的确定研究总体、进行样本量估计、拟定随机抽样与随机分组方法以及统计分析方法等,旨在保证研究结果的经济性、可靠性、科学性和可重复性。

研究设计是任何一项研究工作的第一步,也是最关键的一环。研究者一定要树立“设计优先”的思想,一个严谨的研究设计,必然是专业设计与统计设计的完美结合。

1.4.2 资料收集

1. 资料收集

资料收集(data collection)即根据研究目的,按照设计要求,收集原始资料;其过程就是具体调查或实验实施的过程。资料的准确性、完整性和及时性,是得到正确统计结论的前提和基础。

2. 资料来源

医学统计资料主要来自日常医疗工作记录、统计报表、专题调查与实验研究资料、统计年鉴等。

1.4.3 资料整理

1. 资料整理

资料整理(data sorting)就是把收集到的原始资料,有目的、有计划地进行科学加工,使其系统化、条理化,以便于更好地统计分析,揭示所研究事物的本质及规律性。

2. 主要内容

- (1) 资料审核,查漏补缺和检错纠错。
- (2) 根据研究目的及资料的属性或数量特征,对资料进行分组以及变量变换等。

1.4.4 资料分析

1. 资料分析

资料分析(data analysis)就是根据研究目的、设计类型、资料属性与特征等,选取恰当的统计分析方法,对资料进行计算分析。

2. 主要内容

统计分析主要包括统计描述和统计推断两个方面。前者是用适当的统计指标与统计图表等,对资料的数量特征及其分布规律进行表达;后者是利用样本资料提供的信息,对未知总体的特征作出概率性的估计和推断,即统计推断包括参数估计和假设检验。

3. 分析过程

比较性研究资料的统计分析一般包括:

- (1) 统计描述,以展示资料的数量特征及分布规律。
- (2) 参数估计,即根据样本信息估计总体参数的大小及其可能所在的范围。
- (3) 假设检验(必要时),即基于样本资料提供的信息,运用恰当的假设检验方法,推测两个(或多个)总体的参数是否相同(或分布是否一致)。

4. 统计分析思路

(1) 明确研究目的:是比较两个(或多个)总体的参数(或分布)是否相同(或一致),还是探寻两个(或多个)变量间的相互关系。

(2) 识别设计类型:是完全随机设计,配对设计,还是随机区组设计等。

(3) 辨明资料属性:是数值变量资料,无序分类变量资料,还是有序分类变量资料。

(4) 拟定假设检验方法:包括参数检验和非参数检验等。

(5) 考察统计分析方法的应用条件,确定合适方法。

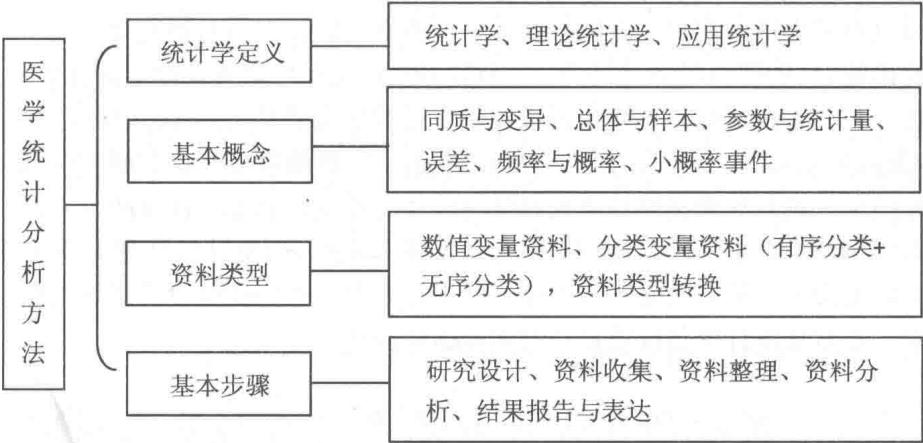
1.4.5 结果报告与表达

当今信息时代,统计设计、数据库的建立与管理、统计分析与结果表达等都由计算机及统计软件完成;因而正确理解和应用统计思想、统计原理和统计分析方法,建立科学的统计思维方式,对统计软件的输出结果进行规范表达和合理的专业解释显得越来越重要。

统计分析结果的表达,除了恰当的统计指标和统计图表外,一般还要介绍研究对象的随机抽样与分组方法、统计分析软件及版本、检验统计量的大小及其 P 值等,并结合学科专业知识对统计分析结果作出合理的专业解释。统计分析结果的规范报告和表达,不但能提高研究结果的认可度,而且有利于学术交流。

学习小结

医学统计分析方法是医学科研的重要工具,统计思维和方法已渗透到医学研究和管理决策的方方面面。学习本课程时,应建立设计优先的意识,掌握统计学的基本概念、基本原理和基本方法,遵循统计分析方法的应用条件,选择恰当的统计分析方法,规范表达统计分析结果,并对统计结果进行合理的专业解释。



思考题

- 1. 常见的医学资料类型有哪些？各有何特点？
- 2. 简述统计工作的基本步骤以及各步骤的主要作用。
- 3. 抽样研究时要求样本有良好的代表性,请谈谈你对样本代表性的理解？