



全国普通高等医学院校五年制临床医学专业“十三五”规划教材
(供五年制临床医学专业用)

医学统计学

主 编 吴学森

副主编 苏 俐 李 栋 朱继民

中国医药科技出版社



全国普通高等医学院校五年制临床医学专业“十三五”规划教材
(供五年制临床医学专业用)

医学统计学

主 编 吴学森

副主编 苏 俐 李 栋 朱继民

编 者 (以姓氏笔画为序)

丁 宇 (新乡医学院)

刘一志 (泰山医学院)

朱继民 (安徽中医药大学)

苏 俐 (内蒙古医科大学)

张星光 (内蒙古医科大学)

贾贤杰 (蚌埠医学院)

韦 杰 (贵阳中医学院)

关红军 (牡丹江医学院)

李 栋 (泰山医学院)

吴学森 (蚌埠医学院)

姚 政 (云南中医学院)

曹文君 (长治医学院)

秘 书 贾贤杰 (蚌埠医学院)

中国医药科技出版社

内 容 提 要

本教材为全国普通高等医学院校五年制临床医学专业“十三五”规划教材之一。本书系根据全国普通高等医学院校五年制临床医学专业“十三五”规划教材编写总体原则、要求和医学统计学课程教学大纲的基本要求及课程特点编写而成,其内容主要包括绪论、统计描述、抽样分布与参数估计、数值变量资料的假设检验、方差分析、分类变量资料的假设检验、非参数检验、直线相关与回归、实验设计概述、临床试验设计概述等内容。并在各章设有“学习要求”“知识链接”“案例讨论”“本章小结”及“练习题”等模块。同时配有“爱慕课”在线学习平台(包括电子教材、教学大纲、教学指南、视频、课件、题库、图片等),从而使教材内容立体化、生动化,便教易学。本教材具有简洁、实用,便于学生自学特点。

本教材供全国普通高等医学院校基础、临床、预防、口腔医学类专业师生教学使用。

图书在版编目(CIP)数据

医学统计学/吴学森主编. —北京:中国医药科技出版社, 2016. 8

全国普通高等医学院校五年制临床医学专业“十三五”规划教材

ISBN 978 - 7 - 5067 - 8200 - 5

I. ①医… II. ①吴… III. ①医学统计 - 医学院校 - 教材 IV. ①R195.1

中国版本图书馆 CIP 数据核字(2016)第 170509 号

美术编辑 陈君杞

版式设计 张 璐

出版 中国医药科技出版社

地址 北京市海淀区文慧园北路甲 22 号

邮编 100082

电话 发行: 010 - 62227427 邮购: 010 - 62236938

网址 [www. cmstp. com](http://www.cmstp.com)

规格 889 × 1194mm¹/₁₆

印张 10

字数 238 千字

版次 2016 年 8 月第 1 版

印次 2016 年 8 月第 1 次印刷

印刷 三河市国英印务有限公司

经销 全国各地新华书店

书号 ISBN 978 - 7 - 5067 - 8200 - 5

定价 28.00 元

版权所有 盗版必究

举报电话: 010 - 62228771

本社图书如存在印装质量问题请与本社联系调换

全国普通高等医学院校五年制临床医学专业“十三五”规划教材

出版说明

为面向全国省属院校五年制临床医学专业教学实际编写出版一套切实满足培养应用型、复合型、技能型临床医学人才需求和“老师好教、学生好学及学后好用”的五年制临床医学专业教材，在教育部、国家卫生和计划生育委员会、国家食品药品监督管理总局的支持下，根据以“5+3”为主体的临床医学教育综合改革和国家医药卫生体制改革新精神，依据“强化医学生职业道德、医学人文素养教育”“提升临床胜任力”“培养学生临床思维能力和临床实践操作能力”等人才培养要求，在中国工程院副院长、第四军医大学原校长、中华医学会消化病学分会原主任委员樊代明院士等专家的悉心指导下，中国医药科技出版社组织全国近100所省属高等医学院校为主体的具有丰富教学经验和较高学术水平的550余位专家教授历时1年余的编撰，全国普通高等医学院校五年制临床医学专业“十三五”规划教材即将付梓出版。

本套教材包括五年制临床医学专业理论课程主干教材共计40门。将于2016年8月由中国医药科技出版社出版发行。主要供全国普通高等医学院校五年制临床医学专业教学使用，基础课程教材也可供基础医学、预防医学、口腔医学等专业教学使用。

本套教材定位清晰、特色鲜明，主要体现在以下方面：

1. 切合院校教学实际，突显教材针对性和适应性

在编写本套教材过程中，编者们始终坚持从全国省属医学院校五年制临床医学专业教学实际出发，并根据培养应用型临床医学人才的需求和基层医疗机构对医学生临床实践操作能力等要求，结合国家执业医师资格考试和住院医师规范化培训新要求，同时适当吸收行业发展的新知识、新技术、新方法，从而保证教材内容具有针对性、适应性和权威性。

2. 提升临床胜任能力，满足应用型人才需求

本套教材的内容和体系构建以强化医学生职业道德、医学人文素养教育和临床实践能力培养为核心，以提升临床胜任力为导向，体现“早临床、多临床、反复临床”，推进医学基础课程与临床课程相结合，转变重理论而轻临床实践、重医学而轻职业道德、人文素养的传统观念，注重培养学生临床思维能力和临床实践操作能力，满足培养应用型、复合型、技能型临床医学人才的要求。

3. 体现整合医学理念，强化医德与人文情感教育

本套教材基础课程与临床课程教材通过临床问题或者典型的案例来实现双向渗透与重组，

各临床课程教材之间考虑了各专科之间的联系和融通，逐步形成立体式模块课程知识体系。基础课程注重临床实践环节的设置，以体现医学特色，医学专业课程注重体现人文关怀，强化学生的人文情感和人际沟通能力的培养。

4. 创新教材编写模式，增强内容的可读性实用性

在遵循教材“三基、五性、三特定”的建设规律基础上，创新编写模式，引入“临床讨论”（或“案例讨论”）内容，同时设计“学习要求”“知识链接”“本章小结”及“练习题”或“思考题”模块，以增强教材内容的可读性和实用性，更好地培养学生学习的自觉性和主动性以及理论联系实践的能力、创新思维能力和综合分析能力。

5. 搭建在线学习平台，立体化资源促进数字教学

在编写出版整套纸质教材的同时，编者与出版社为师生均免费搭建了与每门纸质教材配套的“爱慕课”在线学习平台（含电子教材、教学课件、图片、微课、视频、动画及练习题等教学资源），使教学内容资源更加丰富和多样化、立体化，更好地满足在线教学信息发布、师生答疑互动及学生在线测试等教学需求，促进学生自主学习，为提高教育教学水平和质量，实现教学形成性评价等、提升教学管理手段和水平提供支撑。

编写出版本套高质量教材，得到了全国知名专家的精心指导和各有关院校领导与编者的大力支持，同时本套教材专门成立了评审委员会，十余位院士和专家教授对教材内容进行了认真审定并提出了宝贵意见，在此一并表示衷心感谢。出版发行本套教材，希望受到广大师生欢迎，并在教学中积极使用本套教材和提出宝贵意见，以便修订完善，共同打造精品教材，为促进我国五年制临床医学专业教育教学改革和人才培养作出积极贡献。

中国医药科技出版社

2016年7月

全国普通高等医学院校五年制临床医学专业“十三五”规划教材

教材建设指导委员会

主任委员 樊代明（中国工程院、第四军医大学）

副主任委员 （以姓氏笔画为序）

冯向先（长治医学院）

刘志跃（内蒙古医科大学）

杨 柱（贵阳中医学院）

吴开春（第四军医大学）

郑建中（长治医学院）

蔡映云（复旦大学附属中山医院）

委 员 （以姓氏笔画为序）

丰慧根（新乡医学院）

王旭霞（山东大学齐鲁医学部）

王金胜（长治医学院）

王桂琴（山西医科大学）

王雪梅（内蒙古医科大学）

王勤英（山西医科大学）

石秀梅（牡丹江医学院）

卢 海（首都医科大学附属北京同仁医院）

叶本兰（厦门大学医学院）

付升旗（新乡医学院）

邢 健（牡丹江医学院）

吕 丹（温州医科大学）

吕杰强（温州医科大学）

朱金富（新乡医学院）

任明姬（内蒙古医科大学）

刘学敏（长治医学院）

刘挨师（内蒙古医科大学）

孙思琴（泰山医学院）

孙钰玮（牡丹江医学院）

杨 征（四川大学华西口腔医院）

杨少华（桂林医学院）

李永芳（泰山医学院）

李建华（青海大学医学院）

吴学森（蚌埠医学院）

邱丽颖（江南大学无锡医学院）

何志巍（广东医科大学）

邹义洲（中南大学湘雅医学院）

张 闻（昆明医科大学）

张 燕（广西医科大学）

张丽芳（长治医学院）

张轩萍（山西医科大学）

张秀花（江南大学无锡医学院）

张荣波（安徽理工大学医学院）

张福良（大连医科大学）

林 昶（福建医科大学）

林友文（福建医科大学）

林贤浩（福建医科大学）

明海霞（甘肃中医药大学）

罗晓红 (成都中医药大学)	金子兵 (温州医科大学)
金美玲 (复旦大学附属中山医院)	郑 多 (深圳大学医学院)
赵小菲 (成都中医药大学)	赵幸福 (江南大学无锡医学院)
郝岗平 (泰山医学院)	柳雅玲 (泰山医学院)
段 斐 (河北大学医学院)	费 舟 (第四军医大学)
姚应水 (皖南医学院)	夏 寅 (首都医科大学附属北京天坛医院)
夏超明 (苏州大学医学部)	钱睿哲 (复旦大学基础医学院)
高凤敏 (牡丹江医学院)	郭子健 (江南大学无锡医学院)
郭艳芹 (牡丹江医学院)	郭晓玲 (承德医学院)
郭崇政 (长治医学院)	郭嘉泰 (长治医学院)
席 彪 (河北医科大学)	黄利华 (江南大学无锡医学院)
曹颖平 (福建医科大学)	彭鸿娟 (南方医科大学)
韩光亮 (新乡医学院)	游言文 (河南中医药大学)
强 华 (福建医科大学)	路孝琴 (首都医科大学)
窦晓兵 (浙江中医药大学)	

全国普通高等医学院校五年制临床医学专业“十三五”规划教材

教材评审委员会

主任委员 樊代明 (中国工程院副院长、院士,第四军医大学)

副主任委员 (以姓氏笔画为序)

刘昌孝 (中国工程院院士,天津药物研究院)

张志愿 (中国工程院院士,上海交通大学医学院)

张伯礼 (中国工程院院士,天津中医药大学、中国中医科学院)

顾晓松 (中国工程院院士,南通大学)

委员 (以姓氏笔画为序)

丁 炯 (南京医科大学)

王海波 (山东大学医学院)

吕广明 (南通大学医学院)

刘志跃 (内蒙古医科大学)

李亚明 (中国医科大学)

杨 柱 (贵阳中医学院)

吴开春 (第四军医大学)

张亚林 (中南大学湘雅二医院)

郑建中 (长治医学院)

洪 洋 (中国医科大学)

高志强 (北京协和医院)

葛金文 (湖南中医药大学)

蔡映云 (复旦大学附属中山医院)

全国普通高等医学院校五年制临床医学专业“十三五”规划教材

书 目

序号	教材名称	主编	ISBN
1	医用高等数学	吕 丹 张福良	978 - 7 - 5067 - 8193 - 0
2	医学统计学	吴学森	978 - 7 - 5067 - 8200 - 5
3	医用物理学	张 燕 郭嘉泰	978 - 7 - 5067 - 8195 - 4
4	有机化学	林友文 石秀梅	978 - 7 - 5067 - 8196 - 1
5	生物化学与分子生物学	郝岗平	978 - 7 - 5067 - 8194 - 7
6	系统解剖学	付升旗 游言文	978 - 7 - 5067 - 8198 - 5
7	局部解剖学	李建华 刘学敏	978 - 7 - 5067 - 8199 - 2
8	组织学与胚胎学	段 斐 任明姬	978 - 7 - 5067 - 8217 - 3
9	医学微生物学	王桂琴 强 华	978 - 7 - 5067 - 8219 - 7
10	医学免疫学	张荣波 邹义洲	978 - 7 - 5067 - 8221 - 0
11	医学生物学	张 闻 郑 多	978 - 7 - 5067 - 8197 - 8
12	医学细胞生物学	丰慧根 窦晓兵	978 - 7 - 5067 - 8201 - 2
13	人体寄生虫学	夏超明 彭鸿娟	978 - 7 - 5067 - 8220 - 3
14	生理学	叶本兰 明海霞	978 - 7 - 5067 - 8218 - 0
15	病理学	柳雅玲 王金胜	978 - 7 - 5067 - 8222 - 7
16	病理生理学	钱睿哲 何志巍	978 - 7 - 5067 - 8223 - 4
17	药理学	邱丽颖 张轩萍	978 - 7 - 5067 - 8224 - 1
18	临床医学导论	郑建中	978 - 7 - 5067 - 8215 - 9
19	诊断学	高凤敏 曹颖平	978 - 7 - 5067 - 8226 - 5
20	内科学	吴开春 金美玲	978 - 7 - 5067 - 8231 - 9
21	外科学	郭子健 费 舟	978 - 7 - 5067 - 8229 - 6
22	妇产科学	吕杰强 罗晓红	978 - 7 - 5067 - 8230 - 2
23	儿科学	孙钰玮 赵小菲	978 - 7 - 5067 - 8227 - 2
24	中医学	杨 柱	978 - 7 - 5067 - 8212 - 8
25	口腔科学	王旭霞 杨 征	978 - 7 - 5067 - 8205 - 0
26	耳鼻咽喉头颈外科学	夏 寅 林 昶	978 - 7 - 5067 - 8204 - 3
27	眼科学	卢 海 金子兵	978 - 7 - 5067 - 8203 - 6
28	神经病学	郭艳芹 郭晓玲	978 - 7 - 5067 - 8202 - 9
29	精神病学	赵幸福 张丽芳	978 - 7 - 5067 - 8207 - 4
30	传染病学	王勤英 黄利华	978 - 7 - 5067 - 8208 - 1
31	医学心理学	朱金富 林贤浩	978 - 7 - 5067 - 8225 - 8
32	医学影像学	邢 健 刘挨师	978 - 7 - 5067 - 8228 - 9
33	医学遗传学	李永芳	978 - 7 - 5067 - 8206 - 7
34	核医学	王雪梅	978 - 7 - 5067 - 8209 - 8
35	全科医学概论	路孝琴 席 彪	978 - 7 - 5067 - 8192 - 3
36	临床循证医学	韩光亮 郭崇政	978 - 7 - 5067 - 8213 - 5
37	流行病学	冯向先	978 - 7 - 5067 - 8210 - 4
38	预防医学	姚应水	978 - 7 - 5067 - 8211 - 1
39	康复医学	杨少华 张秀花	978 - 7 - 5067 - 8214 - 4
40	医学文献检索	孙思琴	978 - 7 - 5067 - 8216 - 6

注:40 门主干教材均配套有中国医药科技出版社“爱慕课”在线学习平台。

随着大数据时代的到来,临床实践每天都产生海量的数据;如何从这些海量的数据中发现规律,对临床医生们来说是一种挑战。以前,大部分高等医药学校临床专业把《医学统计学》中的部分内容归并在《预防医学》这门课中,这不利于培养学生分析处理数据的思维能力。此次,把《医学统计学》单独作为一门教材编写,不仅是顺应学科发展的要求,而且是提高学生分析数据的能力、培养具有创新理念的医药人才的重要保证。

本教材是全国普通高等医学院校五年制临床医学专业“十三五”规划教材之一,其通过案例引导着重体现预防医学与临床医学的“整合”,着力提升学生健康教育、预防保健、卫生管理等方面的数据处理能力,以“大预防观”去看待临床积累的数据,而不是单纯地从临床个案去理解健康。本教材的编者来自全国10所普通高等医药院校,他们长期在医学院校从事医学统计学的教学与科研工作,有着丰富的理论与实践经验,在内容的取舍上更能贴近现实,从而有利于教和学。

本教材在编写上遵循“够用、实用”的原则;且大部分章节以案例为先导,针对不同的数据类型采用不同的统计分析方法,方法的学习以解决实际问题为目的,不追求理论的系统性和完整性,着重培养学生分析处理数据的思维能力;因此,简洁、实用,针对性强,便于学生自学等是本教材的特点。

作者根据临床医学专业《医学统计学》教学大纲的基本要求和课程特点编写成本教材,内容上涵盖:第一章 绪论(吴学森编写)、第二章 统计描述(贾贤杰编写)、第三章 抽样分布与参数估计(曹文君编写)、第四章 数值变量资料的假设检验(姚政编写)、第五章 方差分析(苏俐、张星光编写)、第六章 分类变量资料的假设检验(李栋、刘一志编写)、第七章 非参数检验(关红军编写)、第八章 直线相关与回归(朱继民编写)、第九章 实验设计概述(韦杰编写)、第十章 临床试验设计概述(丁宇编写)等内容。本教材适合于临床医学专业、口腔医学专业、药学专业及其他相关专业本科生使用。

由于时间仓促,书中难免存在不足,还请同仁斧正!

CONTENTS

第一章 绪论	1
第一节 医学统计学的概述	1
第二节 医学统计工作的步骤	1
第三节 统计学中的几个基本概念	2
第二章 统计描述	5
第一节 数值变量资料的频数表	5
第二节 数值变量资料的描述性指标	6
一、描述数值变量集中趋势的统计指标	6
二、描述数值变量离散程度的统计指标	8
第三节 正态分布及其应用	10
一、正态分布的概念和特征	10
二、正态分布的应用	11
第四节 分类变量资料的统计描述	12
第五节 统计表和统计图	15
一、常用统计表	15
二、常用统计图	16
第三章 抽样分布与参数估计	22
第一节 抽样与抽样分布	22
一、常用的随机抽样方法	22
二、统计量及其抽样分布	23
三、抽样误差	23
第二节 t 分布和总体均数的估计	25
一、 t 分布的概念	25
二、 t 分布的图形与特征	26
三、总体均数的估计	27
第三节 二项分布和总体率的估计	30
一、二项分布	30
二、二项分布的特征	30
三、二项分布的应用	32
第四节 Poisson 分布和总体均数的估计	33
一、Poisson 分布	33
二、Poisson 分布的特征	33
三、Poisson 分布的应用	34

第四章 数值变量资料的假设检验	38
第一节 假设检验的基本思想	38
一、假设检验的原理	38
二、假设检验的基本步骤	39
第二节 t 检验与 u 检验	40
一、单样本 t 检验	40
二、配对 t 检验	40
三、两独立样本 t 检验	41
四、 u 检验	43
第三节 正态性检验	44
一、简单判断法	44
二、图示法	44
三、矩法	45
四、假设检验法	45
第四节 两个方差的齐性检验	46
一、 F 检验	46
二、Levene 检验	46
第五节 I 型错误和 II 型错误	47
一、I 型错误	47
二、II 型错误	47
第六节 假设检验时应注意的事项	48
第五章 方差分析	52
第一节 方差分析的基本思想及应用的基本条件	52
一、方差分析的基本思想	52
二、方差分析应用的基本条件	54
第二节 完全随机设计资料的方差分析	54
一、变异与自由度的分解	55
二、完全随机设计资料方差分析的基本步骤	55
第三节 随机区组设计资料的方差分析	56
一、变异与自由度的分解	56
二、随机区组设计资料方差分析的基本步骤	57
第四节 重复测量设计资料的方差分析	58
一、变异与自由度的分解	59
二、重复测量设计资料方差分析的基本步骤	60
第五节 多个样本均数的多重比较	61
一、LSD - t 检验	61
二、SNK - q 检验	61
三、Dunnett - t 检验	62
第六节 多个样本的方差齐性检验	63
一、Bartlett χ^2 检验	63
二、Levene 检验	63
第七节 变量变换	64
一、对数变换	64

二、平方根变换	64
三、平方根反正弦变换	64
第六章 分类变量资料的假设检验	68
第一节 χ^2 检验	68
一、 χ^2 检验的基本思想	68
二、四格表资料的 χ^2 检验	71
三、配对设计四格表资料的 χ^2 检验	74
四、行 $\times$ 列表资料的 χ^2 检验	75
第二节 服从二项分布资料的假设检验	81
一、样本率与总体率的比较	81
二、两样本率的比较——Z 检验	82
第三节 服从 Poisson 分布资料的假设检验	83
一、样本均数与总体均数的比较	83
二、两样本均数的比较	84
三、应用 Poisson 分布的注意事项	85
第七章 非参数检验	89
第一节 配对设计的符号秩和检验	89
一、检验步骤与基本思想	89
二、配对设计的符号秩和检验	90
第二节 成组设计两样本比较的秩和检验	92
一、检验步骤与基本思想	92
二、成组设计两样本比较的秩和检验	92
第三节 成组设计多样本比较的秩和检验	94
一、假设检验的基本步骤	94
二、实例计算	95
三、成组设计多个样本间的多重比较	96
第四节 随机区组设计资料的秩和检验	97
一、检验步骤与基本思想	97
二、随机区组设计资料的秩和检验	98
三、随机区组设计资料的多重比较	99
第八章 直线相关与回归	103
第一节 直线相关	103
一、概念	103
二、相关系数的计算与假设检验	104
三、相关分析的步骤	105
四、注意事项	106
第二节 等级相关	106
一、等级相关	106
二、等级相关系数	106
第三节 直线回归	108
一、概念	108
二、直线回归方程的建立与检验	109
三、决定系数	110

四、残差分析	110
五、直线回归方程的应用	111
第四节 直线相关与直线回归的区别与联系	113
一、区别	113
二、联系	113
第五节 曲线拟合	113
一、概念	113
二、常见曲线类型	114
三、一般步骤	114
第九章 实验设计概述	117
第一节 实验设计的特点及分类	117
一、实验设计的特点	117
二、实验设计的分类	117
第二节 实验设计的基本要素	118
一、研究对象	118
二、处理因素	119
三、实验效应	119
第三节 实验设计的基本原则	120
一、对照原则	120
二、随机化原则	121
三、重复的原则	123
第四节 样本含量的估算	123
一、样本含量估算的条件	123
二、常用的样本含量估算方法	124
第十章 临床试验设计概述	127
一、临床试验的特点	127
二、临床试验的基本步骤	128
附表	131
参考文献	146

第一章 绪 论

学习要求

1. 掌握 医学统计学的定义；统计工作的步骤；统计学中的几个基本概念；学习医学统计学应注意的问题。
2. 了解 医学统计学的内容。

本章主要是介绍学习《医学统计学》这本教材所要达到的目的，以及学好本门课程后所能解决的实际问题。通过介绍本课程的一些基本概念和原理，初步告诉大家学什么、怎样学。

第一节 医学统计学的概述

统计学（statistics）是研究数据的收集、整理和分析的一门科学，是帮助人们分析所占有的信息，达到去伪存真、去粗取精、正确认识世界的一种重要手段。

医学统计学（medical statistics）是应用数理统计学的原理与方法研究医学科研、居民健康状况以及卫生服务领域中数据的收集、整理和分析的一门应用科学。其目的是探索事物发展的内在规律性。

《Webster 国际大辞典》（第3版）对统计学的定义是“a science dealing with the collection, analysis, interpretation and presentation of numerical data”。由此看出：统计学是处理资料中变异的科学和艺术，是在收集、归类、分析和解释大量数据的过程中获取可靠结果的一门学科。这里强调了“过程”，但在实际工作中，许多人往往是忽略了设计、收集和归类（整理），到了分析数据时才想到统计学，此时难免发生“悔之晚矣”的憾事。作为医学统计学的应用者应充分认识到这一点。

第二节 医学统计工作的步骤

统计学对统计工作的全过程起指导作用，任何统计工作和统计研究的全过程都可分为以下四个步骤。

1. 设计 (design) 在进行统计工作和研究工作之前必须有一个周密的设计。设计是在广泛查阅文献、全面了解国内外现状、充分咨询相关专家意见的基础上，对要开展的研究工作所做的全面设想。其内容主要包括：明确研究目的和研究假说，做到目标明确、假说清晰；确定观察对象、观察单位、样本含量和抽样方法，拟定研究方案、预期分析指标、误差控制措施、研究进度及经费预算等。设计是整个研究工作中最关键的一环，也是指导以后工作的依据（详见第九章、十章）。

2. 收集资料 (collection of data) 按设计的要求，遵循统计学原理采取必要措施及时

取得准确、完整的原始资料。及时、准确、完整是收集资料的基本原则。医学工作中的统计资料主要来自以下三个方面：①统计报表，是由国家统一设计，有关医疗卫生机构定期逐级上报，提供居民健康状况和医疗卫生机构工作的主要数据，是制定卫生工作计划与措施、检查与总结工作的依据。如法定传染病报表、职业病报表、医院工作报表等。②经常性工作记录，如卫生监测记录、健康检查记录等。③专题调查或实验。

3. 整理资料 (sorting data) 收集来的资料在整理之前称为原始资料，原始资料通常是一堆杂乱无章的数据。整理资料的目的是通过科学的分组和归纳，使原始资料系统化、条理化，便于进一步计算统计指标和分析。其过程是：首先对原始资料进行准确性审查（逻辑审查与技术审查）和完整性审查；再拟定整理表，按照“同质者合并，非同质者分开”的原则对资料进行质量分组，并在同质基础上根据数值大小进行数量分组；最后汇总归纳。

4. 分析资料 (analysis of data) 其目的是依据数据资料的特点，计算有关指标，反映数据的综合特征，阐明事物的内在联系和规律。统计分析包括统计描述 (descriptive statistics) 和统计推断 (inferential statistics)。前者是用统计指标与统计图 (表) 等工具对样本资料的数量特征及其分布规律进行描述 (详见第二章)；后者是指如何抽样，以及如何用样本信息推断总体特征 (详见第三、四、五、六、七、八章)。进行资料分析时，需根据研究目的、设计类型和资料类型选择恰当的描述性指标和统计推断方法。

统计工作的四个步骤紧密相连、不可分割，任何一步的缺陷，都将影响最终分析的结果。

第三节 统计学中的几个基本概念

1. 同质 (homogeneity) 与变异 (variation) 严格地讲，同质是指被研究指标的影响因素完全相同。但在医学研究中，有些影响因素往往是难以控制的 (如遗传、营养等)，甚至是未知的。所以，在统计学中常把同质理解为对研究指标影响较大的、可以控制的主要因素尽可能相同。例如研究儿童的身高时，要求性别、年龄、民族、地区等影响身高较大的、易控制的因素要相同，而不易控制的遗传、营养等影响因素可以忽略。

同质基础上的个体差异称为变异。如同性别、同年龄、同民族、同地区的健康儿童的身高、体重不尽相同。事实上，客观世界充满了变异，生物医学领域更是如此。哪里有变异，哪里就需要统计学。若所研究的两类群体中所有个体一模一样，只需观察任一个体即可，无需进行统计研究。

2. 总体 (population) 与样本 (sample) 任何统计研究都必须首先确定观察单位 (observed unit)，亦称个体 (individual)。观察单位是统计研究中最基本的单位，可以是一个人、一个家庭、一个地区、一个样品、一个采样点等。

总体是根据研究目的确定的同质观察单位的全体，或者说，是同质的所有观察单位某种观察值 (变量值) 的集合。例如欲研究某省 2012 年 7 岁健康男孩的身高，那么，观察对象是该省 2012 年的 7 岁健康男孩，观察单位是每个 7 岁健康男孩，变量是身高，变量值 (观察值) 是身高测量值，则该省 2012 年全体 7 岁健康男孩的身高值构成一个总体。它的同质基础是同地区、同年份、同性别、同为健康的儿童。总体又分为有限总体 (finite population) 和无限总体 (infinite population)。有限总体是指在某特定的时间与空间范围内，同质研究对象的所有观察单位的某变量值的个数为有限个，如上例。无限总体是抽象的，无时间和空间的限制，观察单位数是无限的，如研究碘盐对缺碘性甲状腺病的防治效果，该总体的同质基础是缺碘性甲状腺病患者，同用碘盐防治；该总体应包括已使用和设想要使用碘盐防治的所有缺碘性甲状腺病患者的防治效果，没有时间和空间范围的限制，因而观察单位数无限，该总体为无限总体。

在实际工作中,所要研究的总体无论是有限的还是无限的,通常都是采用抽样研究。样本是按照随机化原则,从总体中抽取的有代表性的部分观察单位的变量值的集合。如从上例的有限总体(某省2012年7岁健康男孩)中,按照随机化原则抽取100名7岁健康男孩,他们的身高值即为样本。从总体中抽取样本的过程为抽样,抽样方法有多种,详见第三章。抽样研究的目的是用样本信息描述、推断总体特征。

统计学好比是总体与样本间的桥梁,能帮助人们设计与实施如何从总体中科学地抽取样本,使样本中的观察单位数(亦称样本含量, sample size)恰当,信息丰富,代表性好;能帮助人们挖掘样本中的信息,推断总体的规律性。

3. 资料 (data) 与变量 (variable) 及其分类 总体确定之后,研究者应对每个观察单位的某项特征进行测量或观察,特征称为变量。如“身高”“体重”“性别”“血型”“疗效”等。变量的测定值或观察值称为变量值 (value of variable) 或观察值 (observed value),亦称为资料。

在处理资料之前,首先要分清变量类型。

(1) 数值变量 (numerical variable) 其变量值是定量的,表现为数值大小,可经测量取得数值,多有度量衡单位。如身高 (cm)、体重 (kg)、血压 (mmHg, 或 kPa)、脉搏 (次/min) 和白细胞计数 ($\times 10^9/L$) 等。这种由数值变量的测量值构成的资料称为数值变量资料,亦称为定量资料 (quantitative data)。大多数的数值变量为连续型变量,如身高、体重、血压等;而有的数值变量的测定值只能是正整数,如脉搏、白细胞计数等,在医学统计学中把它们也视为连续型变量。

(2) 分类变量 (categorical variable) 其变量值是定性的,表现为互不相容的类别或属性。分类变量可分为无序变量和有序变量两类:

① 无序分类变量 (unordered categorical variable) 是指所分类别或属性之间无程度和顺序的差别。它又可分为: A. 二项分类,如性别(男、女),药物反应(阴性和阳性)等; B. 多项分类,如血型(O、A、B、AB),职业(工、农、商、学、兵)等。对于无序分类变量的分析,应先按类别分组,清点各组的观察单位数,编制分类变量的频数表,所得资料为无序分类资料,亦称计数资料。

② 有序分类变量 (ordinal categorical variable) 各类别之间有程度的差别。如尿糖化验结果按-、±、+、++、+++分类;疗效按治愈、显效、好转、无效分类。对于有序分类变量,应先按等级顺序分组,清点各组的观察单位个数,编制有序变量(各等级)的频数表,所得资料称为等级资料。

变量类型不是一成不变的,根据研究目的的需要,各类变量之间可以进行转化。例如血红蛋白(g/L)原属数值变量,若按血红蛋白正常与偏低分为两类时,可按二项分类资料分析;若按重度贫血、中度贫血、轻度贫血、正常、血红蛋白增高分为5个等级时,可按等级资料分析。有时亦可将分类资料数量化,如可将患者的恶性反应以0、1、2、3表示,则可按数值变量资料(定量资料)分析。

变量的类型不同,其分布规律亦不同,对它们采用的统计分析方法也不同。

4. 随机事件 (random event) 与概率 (probability) 医学研究的现象,大多数是随机现象,对随机现象进行实验或观察称为随机试验。随机试验的各种可能结果的集合称为随机事件,亦称偶然事件,简称事件。例如用相同治疗方案治疗一批某病的患者,治疗转归可能为治愈、好转、无效、死亡4种结果;对于一个刚入院的患者,治疗后究竟产生哪一种结果是不确定的,可能产生的每一种结果都是一个随机事件。

对于随机事件来说,在一次随机试验中,某个随机事件可能发生也可能不发生,但在一定数量的重复试验后,该随机事件的发生情况是有规律可循的。概率是描述随机事件发生可