

全国高等医药教材建设研究会 卫生部规划教材
全国高等学校教材
供基础、临床、预防、口腔医学类专业用

医学统计学

第 4 版

主 编 马斌荣

 人民卫生出版社

全国高等学校教材
供基础、临床、预防、口腔医学类专业用

医学统计学

第 4 版

主 编 马斌荣

编 者(以姓氏汉语拼音字头为序)

陈 智(中国医科大学)

方 亚(华中科技大学同济医学院)

郭德成(首都医科大学)

李 康(哈尔滨医科大学)

刘启贵(大连医科大学)

马斌荣(首都医科大学)

施 榕(上海第二医科大学)

田考聪(重庆医科大学)

人 民 卫 生 出 版 社

图书在版编目(CIP)数据

医学统计学/马斌荣主编. —4版.—北京:
人民卫生出版社, 2004.6
ISBN 7-117-06225-8

I. 医… II. 马… 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(2004)第 049214 号

医学统计学
第 4 版

主 编: 马斌荣

出版发行: 人民卫生出版社(中继线 67616688)

地 址: (100078)北京市丰台区方庄芳群园 3 区 3 号楼

网 址: <http://www.pmph.com>

E - mail: pmph@pmph.com

印 刷: 北京智力达印刷有限公司(万通)

经 销: 新华书店

开 本: 850 × 1168 1/16 印张: 18

字 数: 416 千字

版 次: 1990 年 6 月第 1 版 2004 年 7 月第 4 版第 19 次印刷

标准书号: ISBN 7-117-06225-8/R·6226

定 价: 22.00 元

著作权所有, 请勿擅自用本书制作各类出版物, 违者必究

(凡属质量问题请与本社发行部联系退换)

全国高等学校五年制临床医学专业

第六轮规划教材修订说明

为适应我国高等医学教育改革的需要,经全国高等医药教材建设研究会和卫生部临床医学专业教材评审委员会审议,决定从2002年9月开始进行五年制临床医学专业规划教材第六轮的修订。第六轮的修订工作要以《中国医学教育改革发展纲要》和《关于“十五”期间普通高等教育教材建设与改革的意见》为指导,及时反映新世纪教学内容和课程改革的成果,在选择教材内容和编写体系时,应注意素质教育和创新能力与实践能力的培养,为学生知识、能力、素质协调发展创造条件。第六轮的修订要继承和发扬第五轮教材编写的优点,在坚持“三基”、“五性”、“三特定”的同时,提倡创新,可同时编写配套教材(含光盘);增加英文的词汇量;加强人文科学的内容;并强调增强学生的法律意识等,力争编出精品教材。

随着教材品种的不断增多和完善,第六轮教材将不再与七年制共用;并为适应各院校的具体情况,不再划分必修教材和选修教材,由各院校自行选择使用。

全套教材共50种,于2004年秋季全部出齐,其中24种同时为教育部确定的普通高等教育“十五”国家级规划教材。另根据学科发展的需要,本轮教材将原《耳鼻咽喉科学》更名为《耳鼻咽喉-头颈外科学》;将原《计算机应用基础》更名为《医学计算机应用基础》。

第六轮教材目录

- | | | | |
|------------------|---------|-----------------|---------|
| 1. 《医用高等数学》第4版 | 主编 张选群 | 14. 《病理学》第6版 | 主编 李玉林 |
| △2. 《医学物理学》第6版 | 主编 胡新珉 | | 副主编 唐建武 |
| 3. 《基础化学》第6版 | 主编 魏祖期 | △15. 《病理生理学》第6版 | 主编 金惠铭 |
| 4. 《有机化学》第6版 | 主编 吕以仙 | | 王建枝 |
| | 副主编 陆阳 | 16. 《药理学》第6版 | 主编 杨宝峰 |
| 5. 《医学生物学》第6版 | 主编 傅松滨 | | 副主编 苏定冯 |
| △6. 《系统解剖学》第6版 | 主编 柏树令 | 17. 《医学心理学》第4版 | 主编 姜乾金 |
| △7. 《局部解剖学》第6版 | 主编 彭裕文 | 18. 《法医学》第4版 | 主编 王保捷 |
| △8. 《组织学与胚胎学》第6版 | 主编 邹仲之 | △19. 《诊断学》第6版 | 主编 陈文彬 |
| △9. 《生物化学》第6版 | 主编 周爱儒 | | 潘祥林 |
| | 副主编 查锡良 | | 副主编 康熙雄 |
| △10. 《生理学》第6版 | 主编 姚泰 | △20. 《医学影像学》第5版 | 主编 吴恩惠 |
| | 副主编 吴博威 | | 副主编 冯敦生 |
| 11. 《医学微生物学》第6版 | 主编 周正任 | △21. 《内科学》第6版 | 主编 叶任高 |
| | 副主编 李凡 | | 陆再英 |
| 12. 《人体寄生虫学》第6版 | 主编 李雍龙 | | 副主编 谢毅 |
| 13. 《医学免疫学》第4版 | 主编 陈慰峰 | | 王辰 |
| | 副主编 金伯泉 | | |

- | | | |
|---------------------|----------------|--------------------|
| △22. 《外科学》第6版 | 主编 吴在德
吴肇汉 | 副主编 刘移民
傅 华 |
| | 副主编 郑 树
安 洪 | 副主编 段广才
主 编 李家邦 |
| △23. 《妇产科学》第6版 | 主 编 乐 杰 | 副主编 高鹏翔 |
| | 副主编 谢 幸 | 主 编 邹赛德 |
| 24. 《儿科学》第6版 | 主 编 杨锡强
易著文 | 副主编 杨长兴
主 编 裴海泓 |
| | 副主编 沈晓明
常立文 | 主 编 宋今丹 |
| △25. 《神经病学》第5版 | 主 编 王维治 | 主 编 药立波 |
| | 副主编 罗祖明 | 副主编 冯作化
周春燕 |
| △26. 《精神病学》第5版 | 主 编 郝 伟 | 主 编 左 伋 |
| △27. 《传染病学》第6版 | 主 编 彭文伟 | 主 编 徐叔云 |
| | 副主编 李兰娟
乔光彦 | 副主编 魏 伟
主 编 马斌荣 |
| △28. 《眼科学》第6版 | 主 编 惠延年 | 主 编 丘祥兴 |
| 29. 《耳鼻咽喉-头颈外科学》第6版 | 主 编 田勇泉 | 副主编 王明旭
主 编 王家良 |
| | 副主编 孙爱华 | 主 编 南登崑 |
| △30. 《口腔科学》第6版 | 主 编 张志愿 | 主 编 郭继军 |
| △31. 《皮肤性病学》第6版 | 主 编 张学军 | 主 编 赵同刚 |
| 32. 《核医学》第6版 | 主 编 李少林 | 副主编 达庆东
汪建荣 |
| | 副主编 张永学 | 主 编 文历阳 |
| △33. 《流行病学》第6版 | 主 编 王建华 | 主 编 杨乘辉 |
| 34. 《卫生学》第6版 | 主 编 仲来福 | |
| | | △35. 《预防医学》第4版 |
| | | 36. 《中医学》第6版 |
| | | 37. 《医学计算机应用基础》第3版 |
| | | 38. 《体育》第3版 |
| | | 39. 《医学细胞生物学》第3版 |
| | | 40. 《医学分子生物学》第2版 |
| | | 41. 《医学遗传学》第4版 |
| | | △42. 《临床药理学》第3版 |
| | | 43. 《医学统计学》第4版 |
| | | △44. 《医学伦理学》第2版 |
| | | △45. 《临床流行病学》第2版 |
| | | 46. 《康复医学》第3版 |
| | | 47. 《医学文献检索》第2版 |
| | | 48. 《卫生法》第2版 |
| | | 49. 《医学导论》第2版 |
| | | △50. 《全科医学概论》第2版 |

注：画△者为普通高等教育“十五”国家级规划教材

全国高等学校临床医学专业 第五届教材评审委员会

名誉主任委员 裘法祖

主任委员 陈灏珠 副主任委员 龚非力

委员（以姓氏笔画为序）

于修平 王卫平 王鸿利 文继舫 朱明德 刘国良
李焕章 杨世杰 张肇达 沈 悌 吴一龙 郑树森
原 林 曾因明 廖秦平 樊小力

秘书 孙利军

第4版前言

廿一世纪是信息技术和生命科学快速发展的时代,医学信息的正确收集、整理和分析对基础医学、临床医学、流行病学、医院信息管理和卫生经济的发展将起重要的促进作用。同时一门新兴的遵循证据评价及患者预后的循证医学(Evidence-based Medicine, EBM)愈来愈受到广大医学工作者的重视。

随着计算机科学的迅猛发展,以及医学统计软件包的日益完善和广泛应用,原来需很长时间才能计算出结果的内容,使用统计软件包瞬间就能准确地计算出来结果,大大促进了医学统计学的发展。但是,统计软件包的使用能给那些熟悉统计基本原理和统计方法的用户带来极大的方便,而那些统计概念不清,不能正确选择统计方法的用户将会感到困惑和不知所措。

根据上述理由,医学统计学(第4版)具有如下特点:

1. 尽可能深入浅出地对医学统计学的基本概念、基本理论、基本知识和基本技术,尤其对各种统计方法的使用条件、使用方法及统计结果的理解做了正确解释,而不拘泥于大量的繁杂的计算过程。对一些很有价值的值得引起注意的概念,比如回归方程中的决定系数、卡方检验中的行 \times 列分割及其检验水准的校正、方差分析中的析因分析及医学统计中的研究设计等,都做了较详细的讲解。各学校可以根据各自教学条件和学生水平进行因材施教,实现学生个性化学习。

2. 根据中华人民共和国教育部高教司[1999]46号文件,“关于转发《高等医药院校计算机基础教学基本要求(本科)》的通知”精神,要求医学生在学习医学统计学和文献检索课时,应该有必要的上机操作。因此,本次教材增加了第十五章医学统计学中的上机实习应用。它以SPSS for Windows Ver11.5为例,示范性地讲解了如何建立数据文件、进行相应的统计分析及其结果的正确解释。这样,既能提高学习效率,也有利于学生熟悉计算机技术的应用。

3. 增加了第十四章“循证医学”,简要讲解了循证医学的基本概念、基本特点、基本要素及临床使用方法和步骤。

4. 在每一章的末尾都增加了密切相关的英语片断。

本教材的参考教学学时为32学时至54学时,建议上机4~8学时。

本书附有英中名词索引和中英名词索引,汇集了全书中的主要专业术语,方便读者查阅。

限于我们的学识和能力,本书还可能有许多不足之处,我们愿意虚心听取广大读者的批评和建议,以便改进。

马斌荣

2004年3月28日

目 录

第一章 医学统计中的基本概念	1
第一节 绪论	1
第二节 医学统计工作的内容和资料类型	2
第三节 医学统计工作中的基本概念	3
第二章 平均水平(集中趋势)的统计描述	9
第一节 频数分布	9
第二节 平均数	11
第三章 离散趋势的统计描述	18
第一节 衡量变异程度(或离散程度)的指标	18
第二节 正态分布及应用	21
第三节 医学参考值范围	25
第四章 抽样误差与假设检验	29
第一节 均数的抽样误差与标准误差	29
第二节 总体均数的估计	30
第三节 假设检验的意义和步骤	32
第五章 t 检验	35
第一节 t 分布	35
第二节 单个样本 t 检验	36
第三节 配对样本 t 检验	36
第四节 两独立样本 t 检验	38
第五节 t 检验中的注意事项	44
第六节 假设检验中的两类错误	46
第六章 方差分析	49
第一节 完全随机设计的方差分析	49
第二节 随机区组设计的方差分析	53
第三节 多个样本均数的两两比较	57
第四节 2×2 析因设计的方差分析	58

第五节	方差齐性检验	63
第七章	相对数及其应用	67
第一节	相对数的概念及计算	67
第二节	相对数使用应注意的问题	69
第三节	率的标准化	70
第四节	医学中常用的相对数指标	72
第五节	率的抽样误差与区间估计	73
第八章	χ^2 检验	77
第一节	行 \times 列表 χ^2 检验	77
第二节	四格表 χ^2 检验	85
第三节	配对资料 χ^2 检验	92
第九章	非参数检验方法	97
第一节	配对资料的符号秩和检验	97
第二节	两样本比较的秩和检验	100
第三节	多个样本比较的秩和检验	103
第十章	线性相关与回归	109
第一节	线性相关	109
第二节	线性回归	113
第三节	线性相关和回归的区别与联系	118
第四节	等级相关	119
第十一章	多元线性回归与多元逐步回归	123
第一节	多元线性回归法	123
第二节	多元逐步回归	131
第三节	多元线性回归分析的注意事项	133
第十二章	统计表与统计图	137
第一节	统计表	137
第二节	统计图	140
第十三章	医学统计中的研究设计	152
第一节	实验研究设计	152
第二节	调查研究设计	157

第三节	诊断试验研究设计	160
第十四章	循证医学	163
第一节	概述	163
第二节	循证医学的证据分级	168
第三节	循证医学的基本过程	168
第四节	循证医学在临床医学中的应用实例	178
第十五章	医学统计学中的上机实习	182
第一节	统计软件概述	182
第二节	统计软件包的评价和选择	183
第三节	数据文件的建立	184
第四节	t 检验	196
第五节	方差分析	202
第六节	χ^2 检验	213
第七节	非参数检验	229
第八节	相关与回归	234
附表		245
英中索引表		264
中英索引表		270
参考文献		276

第一章 医学统计中的基本概念

第一节 绪 论

统计分析是科研工作中的重要组成部分。科研的总体设计、资料采集、资料整理、资料分析直到最后得出结论都与它有密切关系。掌握了它就可以使用较少的人力、物力和时间获得比较可靠的结果。需要运用统计分析的场合而不去运用它可能造成不应有的缺陷或得出错误的结论。

北京某医院某大夫使用“乌贝散”(乌贼骨9克,白及9克,川贝9克,甘草9克,共研细末过120目筛,冲服)治疗胃溃疡病出血107例,有效101例,有效率为94.4%。那么如果别的医院,其他大夫使用“乌贝散”来治疗胃溃疡病出血,其有效率也一定是94.4%吗?显然,不一定恰为94.4%,那么是多少呢?比它高还是比它低?这就是一个求置信区间的问题。

北京某医院有位老大夫,用“冠心灵”治疗冠心病,其对照组用西药,观察结果如表1-1所示:

表 1-1 冠心灵与单纯西药疗效对比

	显 效	有 效	无 效	合 计
单纯西药	9	25	6	40
冠心灵	19	18	5	42

请问冠心灵是否比单纯西药有效?这就涉及到统计学中的统计学检验(test of significance),俗称要算 P 值。

北京15所医院联合对心肌梗死抢救治疗后的患者进行了随访调查,其随访观察的部分资料如表1-2所示:

表 1-2 心肌梗死抢救治疗后的随访研究

	随 访 数	死 于 本 病	中 断 随 访
第 1 年	1342	130	41
第 2 年	1171	48	53
第 3 年	1070	42	73
第 4 年	955	56	48
第 5 年	851	35	

根据这些资料,可提出如下几个问题:

出院后第1年死亡率与以后几年的死亡率是否有显著性差别？

2年生存率是多少？3年生存率是多少？

其中有些患者中断随访了，中断随访的患者如何进行统计分析？怎样的随访才算有效？

这些问题的解决，将有赖于一系列统计方法。

某杂志中一篇文章报道，用某种中草药治疗玫瑰糠疹，有效率为78%，平均疗程为3周左右。问此药治疗玫瑰糠疹的确有效吗？由于玫瑰糠疹具有自然治愈的情况，一般不服药，多喝水，到3周左右有些患者也会自愈。所以，此药的78%疗效有待探讨。这就要求设立具有可比性的对照组。

用某药治某病，治疗1例有效，能说100%有效吗？治疗2例都有效，能说明100%有效吗？治10例，其中9例有效，能说90%有效吗？显然，观察的例数将是个重要因素，例数太少不足以说明问题，观察例数太多，将费时、加重科研经济负担，那么观察多少例能说明问题？这是一个样本可靠性问题。

第二节 医学统计工作的内容和资料类型

一、医学统计工作的内容

实验设计(experiment design)、收集资料(collecting data)、整理资料(sorting data)和分析资料(analyzing data)是医学统计工作的主要内容。

(一) 实验设计

就是根据研究的目的，制定总的研究方案。包括研究对象的纳入标准和排除标准，样本量和样本获取方法，实验组与对照组的分组原则，确定观察指标及精度，实验过程中的质量控制，拟使用的统计方法等等。

(二) 收集资料

就是根据研究的目的、实验设计的要求，收集准确、完整、充满信息的原始资料(raw data)。

医学统计资料主要有实验数据和现场调查资料、医疗卫生工作记录、报表和报告卡等。实验数据是指在试验过程中获得的数据；现场调查资料主要来源于大规模的流行病学调查获取的资料；医疗卫生工作纪录有门诊病历卡、住院病历卡、化验报告等；报表有卫生工作基本情况年报表、传染病年(日、月)报表、疫情旬(月、季、年)报表等；报告卡有传染病发病报告卡、出生报告卡、死亡报告卡等等。

在这些资料的收集过程中，必须进行质量控制(quality control)，包括它的统一性、确切性、可重复性。对这些原始数据的精度(precision)和偏性(bias)应有明确的控制范围。

(三) 整理资料

就是把收集到的原始资料，有目的地进行科学加工，使资料系统化、条理化，以便进行统计分析。一般来说，要注意如下几点：

资料的逻辑检查，检查报表(或报告卡)的纵向、横向的合计和总的合计是否吻合。

从专业的角度对资料的合理性进行检查，比如退休时的年龄不应出现小于20岁，出生婴儿的体重不应出现大于10000g，男性患者的调查表中不应出现妇科疾病等。

从专业的角度对资料的一致性进行检查,比如诊断和疗效的评定标准是否统一,又比如胃镜下萎缩性胃炎的严重程度的评定标准是否一致等。

有时,为了进行统计分析需要对原始数据进行加工,将其转化为频数分布表数据,比如,把年(月)龄进行分组,计算各组出现的频数。频数分布表(frequency distribution table)不仅可以表示观察值的分布情况,也有利于计算各种统计指标。

(四) 分析资料

就是把经过统计整理的资料,做一系列统计描述和统计推断,阐明事物的规律性。应该注意,不同的资料使用的统计描述和统计推断的方法也会不一样,本书在后面几章将做详细介绍。

学习统计方法应着重于:

- 理解医学统计方法的基本原理和基本概念。
- 掌握收集、整理与分析资料的基本知识与技能。
- 重视原始资料的完整性、可靠性及处理数据时的实事求是的科学态度。

二、资料的类型

医学统计资料一般可分为计量资料(measurement data)和计数资料(enumeration data)两大类,不同的统计资料应采用不同的统计分析方法。

计量资料是对每个观察对象的观察指标用定量方法测定其数值大小所得的资料,一般用度量衡单位表示,如身高(cm)、体重(kg)、浓度(mmol)、脉搏(次/分)、血磷(mmol/L)、血红蛋白(g/L)等。

计数资料是先将观察对象的观察指标按性质或类别进行分组,然后计数各组该观察指标的数目所得的资料。比如,用“生脉散”治疗脑血管病,其疗效可分为治愈、显效、有效、无效,那么其对应的人数将是计数资料。又如,调查某人群的血型分布,按照A、B、AB、O四型分组,计数所得该人群的各血型组的人数也是计数资料。

等级分组资料:在医学实践中,有些资料具有计数资料的特性,同时又兼有半定量的性质,被称为按等级分组资料。如临床化验中,将化验结果按-、+、++、+++等级分组,计数得到每组病人数,就是等级分组资料。

根据分析的需要,计量资料、计数资料和等级分组资料可以互相转化。例如每个人的血红蛋白原属计量资料;若按血红蛋白正常与异常分为两组,得出各组的人数,是计数资料;若按血红蛋白含量的多少分为五个等级:<60g/L(重度贫血)、60~90g/L(中度贫血)、90~125g/L(轻度贫血)、125~160g/L(血红蛋白正常)、>160g/L(血红蛋白增高),计算各等级人数,就是等级分组资料。

第三节 医学统计工作中的基本概念

一、变异(variance)

医学研究的对象是有机的生命体,其功能是十分复杂的。不同的个体在相同的条件下,对

外界环境因素可以发生不同的反应。例如，同种族、同年龄、同性别的健康人，在相同的条件下测其脉搏、呼吸、体温等生理指标可以有很大差异。在临床治疗中，用同样的药物治疗病情相同的病人，疗效也不尽相同。即使在实验室里，动物与动物之间也有明显的差异。这种现象称为个体差异或称为变异。变异是由众多的、偶然的、次要的因素造成的。

由于医学统计研究的对象是有变异的事物，因此，用观察 1~2 例的结果来推论出一般规律是不恰当的。例如，我们不能用某一健康成人的红细胞数作为一般健康成人的红细胞数；也不能因为用青木香治疗 1 例高血压患者有效，而断定青木香治疗高血压的有效率为 100%，因为观察例数太少，受个别的偶然的因素影响太大。科学研究的工作，主要任务之一就是要从表现为偶然性的大量数据中，分析出其中必然性的规律。而概率论和数理统计就是要通过这种偶然性去寻找隐藏在其内部的必然规律性。

二、总体(population)和样本(sample)

总体是同质的个体所构成的全体。医学研究的对象，一般都是数量巨大的群体。譬如，我们要研究华北地区成年人的血液的红细胞数的正常值，那么，华北地区全部健康成年人的血液都是我们的研究对象，这称为总体。但是，我们不可能对华北地区所有健康成年人都进行血液检查，而只能按照一定的科学方法从中抽取一部分人来进行血液检查。然后，根据检验结果来推论华北地区成年人的血液的红细胞数的正常值。

这种从总体中抽取部分个体的过程称为抽样(sampling)，所抽得的部分成为样本，在一个样本里含有的个体数可以不同，样本包含的个体数目称为样本含量(sample size)。如何正确地从样本观察到的情况来推测总体，这就是医学统计学要解决的问题之一。

三、抽样(sampling)

从总体中抽取样本，一定要遵循科学原则。一般来说，一个样本应具有代表性(representation)、随机性(randomization)和可靠性(reliability)，二个样本之间应具有可比性(comparable)。

(一) 代表性

就是要求样本中的每一个个体必须符合总体的规定。这就要求对总体有一个明确的规定。这种规定是根据研究目的而具体确定的。例如，要调查华北地区成年男子的红细胞数的正常值，我们可以规定研究对象为体温正常、无急性病、无血液消耗性疾病、肝在肋缘下 0.5cm 以内、血红蛋白在 125g/L 以上、血小板在 $100 \times 10^9/L$ 以上的华北地区 18 岁以上的男子。必须做肝功能、X 线胸透及某些血象检查等，那么所抽取的样本中的每个个体都必须做上述检查，并符合上述规定。

(二) 随机性

就是要保证总体中的每个个体有相同的几率被抽作样本。重要的是要避免主客观的“偏性”。必须指出，随机化抽样绝不等于随意抽样。例如，将 40 只小鼠分为两组。如果闭上眼睛，随意抓取 20 只作为第一组，留下的作为第二组，表面上看来是随机的，实际上是不随机的。因为体壮、活泼性强的小鼠难于抓到，故大部分留在第二组，两组的活泼性是不同的。

为了保证抽样的随机性，可用抽签法、机械抽样法、分层抽样法及利用随机数字表抽样等方法。

1. 抽签法 例如,要把 40 只动物分成两组,首先把每个动物编上号码(1~40),然后做 40 个签(1~40),混匀后,再通过抽签随机把它分成两组。

2. 机械抽样法 又称为等间隔抽样,即先将总的观察对象按某种顺序编号,再从这些编号中采用等间隔抽样。譬如,在某工厂 1000 名健康女工中抽取 100 人进行血液检查,可先将工人编号,然后从 0 到 9 共十个数字中,用抽签法任意抽取一个数字,比如抽到一个“8”,那么工人编号末尾逢“8”的即为检查对象,即 8、18、28、38、……、998 为检查对象。如果想在 1000 名工人抽取 200 人做检查,那么可在 0 到 9 十个数字中随机抽取 2 个数字,则工人编号末尾的数字与这两个数字相同者即为受检对象。

机械抽样的优点是易于理解,简便易行,容易得到一个按比例分配的样本。

3. 分层抽样法 即先按某种性质将总体分为若干组别、类别或区域(在统计上称为“层”)。再从每一层内按比例进行随机抽样,组成样本。比如,调查北京市胃癌发病情况,可先根据北京城区和农村人口比例,确定城区和农村选取的人数。然后,在城区和农村中用机械抽样或抽签法进一步抽样。当然,在城区和农村的抽样中,还应考虑到各年龄段的人口构成,按不同比例再分别随机抽样,即分层随机抽样(stratified random sampling)。分层随机抽样可根据实际科研题目要求分成多层,但一定要注意各层的比例和按比例分配抽取样本。

4. 使用随机数字表进行随机抽样 在许多统计书上都能找到随机数字表,现在的计算机也能产生一系列随机数字。如何用随机数字表进行随机分组,读者可参阅有关书籍,本处从略。

综上所述,随机化抽样绝不是随便抽样,应根据科研要求,做好设计,按设计抽样。

(三) 可靠性

即实验的结果要具有可重复性,即由科研课题的样本得出的结果所推测总体的结论有较大的可信度。由于个体之间存在差异,只有观察一定数量的个体方能体现出其客观规律性。如果根据少数几例就下结论,这种结论可靠性差,体现不了规律,可能被后人所否定,这种先例是曾经发生过的。诚然,每个样本的含量越多,可靠性就会越大,但是例数增加,人力、物力都会发生困难,所以应以“足够”为准。究竟需要多少例数,它与所观察的指标的变异程度有关。可在其他统计书中的“样本例数估计”中找到有关内容,本书从略。

(四) 可比性

如果进行两个或多个样本之间的比较,那么要求每二个样本之间应具有可比性(comparable)。

可比性是指处理组(临床设计中称为治疗组)与对照组之间,除处理因素不同外,其他可能影响实验结果的因素要求基本齐同,也称为齐同对比原则。

譬如,有人研究两种不同中草药对慢性支气管炎的疗效,用两组病人做比较。第一组病人选取农民,治疗时间在 12~2 月间,第二组病人选自工厂工人,治疗时间在 3~5 月间。结果认为第二组疗效显著高于第一组。这个结论显然不可靠。因为不仅农民与工人不同,更关键的是第二组病人的治疗是在天气回暖时,症状可自行缓解。这就难以区别疗效到底是由于药物的作用还是天气回暖的关系,这种对照组无可比性。

不同医院的住院病人差别较大,相互做比较时要注意其可比性。大医院或有特色的医院,接受治疗的病人中可能重病人、疑难杂症者比例较高,小医院或基层医院可能一般的轻症病人

比例较大，所以两家医院的住院病人病死率不具有可比性，不能直接进行比较，要进行标准化后才能进行比较。

在临床医学中，设置对照组是非常必要的，但也是比较困难的。对那些具有“自然痊愈”特性的疾病(譬如玫瑰糠疹、口腔溃疡)都应设立对照组。

空白对照在基础医学研究的动物实验中可以采用，但是，在临床医学研究中一般不采用。

对照组使用“安慰剂”只在少数病种中使用，使用时一定要遵循病人知情原则。往往用于与心理或精神因素有关的疾病，譬如疼痛、晕船、情绪不安等病人。这些安慰剂是外形和真药一样而实质上没有药理作用的物质，它通过医务人员的语言、态度、威信及病人的心理和精神因素而起作用。有的疾病使用“安慰剂”疗效可达35%，有的甚至更高。所以，这类疾病使用“安慰剂”是可行的，也是必要的。

在临床研究中，往往要判断某种新药或某种新疗法的疗效是否较已知的药物或疗法为好，所以对照组通常是采用已知的当前治疗方法中最好的一种。

四、配对设计(matching design)与随机区组设计(random block design)

在做试验或搞科研的开始，我们就应考虑把试验对象按“完全随机设计(completely random design)”、“配对设计(matching design)”或“配伍组设计”。设计不同不仅抽样方法不同，在处理其实验结果时使用的统计方法也不同。

(一) 完全随机设计

它是将受试对象随机分配到各个处理组或对照组中，或分别从不同总体中随机抽样进行研究。它可以是两样本比较，也可以是多样本比较；各样本含量可以相等，也可以不等，但也不宜差别太大。

(二) 配对设计

它是将受试对象按一定条件配成对子，称为配对样本(paired sample)，将它们随机分到两个组中，分别给予不同的处理。配对条件是指有可能影响实验效应的非主要因素。动物实验中，常将种属、性别相同，年(月)龄、体重相近的两个动物配成对子。在人群试验中，常将性别相同，年龄、生活劳动条件相近的两个人配成对子。这样就可增加各处理组之间的齐同对比性。

(三) 随机区组设计

随机区组设计(random block design)也称为配伍组设计，是配对设计的扩展。配对设计的每一“对子”分别随机分到两个处理组中，而配伍组设计中的每个“配伍组”，包含多个受试对象，要将它们分别随机分到各处理组。

五、误差(error)

统计上所说的误差，泛指观测值(observed value)与真实值之差，以及样本指标与总体指标之差，主要有以下三类：

(一) 系统误差

在收集资料过程中，由于仪器初始状态未调整到零、标准试剂未经校正、医生掌握疗效标准偏高或偏低等原因，可造成观察结果倾向性的偏大或偏小，这叫系统误差(system error)。系统误差影响原始资料的准确性，必须克服。如果已发生，要尽力查明其原因，并予以校正。

(二) 随机测量误差

在收集原始资料过程中,即使仪器初始状态及标准试剂已经校正,但是,由于各种偶然因素的影响也会造成同一对象多次测定的结果不完全一致。譬如,实验操作人员操作技术不稳定,不同实验操作人员之间的操作差异,电压不稳及环境温度差异等因素造成测量结果的误差。这种误差往往没有固定的倾向,有时高有时低,被称为随机测量误差(random measurement error)。对于这种误差应采取措施,尽最大可能来控制,至少应控制在一定的允许范围内。一般可以用技术培训、指定固定实验操作人员、加强责任感教育及购置一定精度的稳压器、恒温装置等措施,从而达到控制的目的。

(三) 抽样误差

即使在消除了系统误差,并把随机测量误差控制在允许范围内,样本均数(或其他统计量)与总体均数(或其他统计量)之间仍可能有差异。这种误差产生的原因是:①个体之间存在变异;②抽样时只能抽取总体中的一部分作为样本。由此样本的数据构成的统计指标(如均数)就会与总体的该统计指标有误差,这种差异是抽样引起的,故这种误差叫做抽样误差(sampling error),要用统计方法进行正确分析。

一般说来,样本含量越大,则抽样误差越小,样本的观察指标越与总体的该指标接近,即越能说明总体的规律。反之,样本含量越小,则抽样误差相应地越大。因此,我们不能仅仅凭观察指标的大小进行简单判断,而应该使用概率与数理统计方法来辨别哪些实验研究的结果是有实际意义的,哪些可能是由抽样误差所造成的,从而得出正确的结论。常用的统计方法是“统计学检验”,也称为“显著性检验”。

六、概率(probability)

概率是描写某一事件发生的可能性大小的一个量度。用 A 表示某一事件, P 表示该事件可能发生的概率,可记为 $P(A)$ 。

概率有古典概率与统计概率之分,医学上常用的是统计概率,即对某一随机现象进行大量观察后得到的一个统计百分数 f/N ,此处 N 为观察总数, f 为发生数或频数。譬如,某病病死率,乳腺癌术后五年生存率等等。

在一定条件下,肯定发生的事件称为必然事件,概率为 1;肯定不发生的事件称为不可能事件,概率为 0;可能发生也可能不发生的事件称为随机事件或偶然事件,其概率介于 0 与 1 之间。在统计学上,习惯将 $P \leq 0.05$ 或 $P \leq 0.01$ 的事件称为小概率事件,表示该事件发生的可能性很小。一般来说,如果小概率事件在现实中出现了,就要追究其原因。在医学科研中,常把 $P \leq 0.05$ 作为事物差别有统计意义, $P \leq 0.01$ 作为事物差别有高度统计意义的界限。

思考练习题

1. 常见的三类误差是什么?应采取什么措施和方法加以控制?
2. 统计分析要求每一个样本应该具有哪三性?
3. 什么是两个样本之间的可比性?

(马斌荣)

The modern study of the life sciences includes experimentation, data gathering, and interpretation.

The design and evaluation of experiments, formally known as the scientific method, is utilized in all scientific fields and often implied rather than explicitly outlined in many investigations. The components of the scientific method include observation, formulation of potential question or problem, construction of hypothesis, followed by a prediction, and the design of experiment to test the prediction.

Generally, an observation can be classified as either quantitative or qualitative. Quantitative observations are based on some sorts of measurement, e. g., length, weight, temperature, and pH. Qualitative observations are based on categories reflecting a quality question. This usually takes the form of a "way" question and implies a cause and effect relationship.

A hypothesis is a tentative explanation for the observations you made. A good hypothesis suggests a cause and effect relationship and is testable.

The entire purpose and design of an experiment is to accomplish one goal, that is, to test the hypothesis. An experiment tests the hypothesis testing the correctness or incorrectness of the predictions that came from it. Theoretically, an experiment should alter or test only the factor suggested by the prediction, while all other factors remain constant.