



卫生部“十一五”规划教材

全国高等医药教材建设研究会规划教材

全国高等学校教材

供基础、临床、口腔医学类专业用

# 医学统计学

第 5 版

主 编 马斌荣

副主编 李 康



人民卫生出版社

## 全国高等学校教材

### 供基础、临床、口腔医学类专业用

1. 医用高等数学 / 第5版
2. 医学物理学 / 第7版
3. 基础化学 / 第7版
4. 有机化学 / 第7版
5. 医学生物学 / 第7版
6. 系统解剖学 / 第7版
7. 局部解剖学 / 第7版
8. 组织学与胚胎学 / 第7版
9. 生物化学 / 第7版
10. 生理学 / 第7版
11. 医学微生物学 / 第7版
12. 人体寄生虫学 / 第7版
13. 医学免疫学 / 第5版
14. 病理学 / 第7版
15. 病理生理学 / 第7版
16. 药理学 / 第7版
17. 医学心理学 / 第5版
18. 法医学 / 第5版
19. 诊断学 / 第7版
20. 医学影像学 / 第6版
21. 内科学 / 第7版
22. 外科学 / 第7版
23. 妇产科学 / 第7版
24. 儿科学 / 第7版
25. 神经病学 / 第6版
26. 精神病学 / 第6版
27. 传染病学 / 第7版
28. 眼科学 / 第7版
29. 耳鼻咽喉-头颈外科学 / 第7版
30. 口腔科学 / 第7版
31. 皮肤性病学 / 第7版
32. 核医学 / 第7版
33. 流行病学 / 第7版
34. 卫生学 / 第7版
35. 预防医学 / 第5版
36. 中医学 / 第7版
37. 计算机应用基础 / 第4版
38. 体育 / 第4版
39. 医学细胞生物学 / 第4版
40. 医学分子生物学 / 第3版
41. 医学遗传学 / 第5版
42. 临床药理学 / 第4版
43. **医学统计学 / 第5版**
44. 医学伦理学 / 第3版
45. 临床流行病学 / 第3版
46. 康复医学 / 第4版
47. 医学文献检索 / 第3版
48. 卫生法 / 第3版
49. 医学导论 / 第3版
50. 全科医学概论 / 第3版
51. 麻醉学 / 第2版
52. 急诊医学

策划编辑… 窦天舒 杨 晋  
责任编辑… 樊京娜 杨 晋  
封面设计… 郭 森  
版式设计… 郭 森 陈 旻



ISBN 978-7-117-10090-8



9 787117 100908 >

定价(含光盘): 30.00 元



卫生部“十一五”规划教材  
全国高等医药教材建设研究会规划教材

全国高等学校教材  
供基础、临床、口腔医学类专业用

# 医学统计学

第 5 版

主 编 马斌荣

副主编 李 康

编 者 (以姓氏汉语拼音字头为序)

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)

沈其君 (宁波大学医学院)

王乐三 (中南大学湘雅医学院)

闫 岩 (首都医科大学)

尹 平 (华中科技大学同济医学院)

张丕德 (广东药学院)

钟晓妮 (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. —5 版. —北京:  
人民卫生出版社, 2008. 6  
ISBN 978-7-117-10090-8

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 (2008) 第 047123 号

本书本印次封底贴有防伪标, 请意识别。

**医学统计学**  
第 5 版

**主 编:** 马斌荣

**出版发行:** 人民卫生出版社 (中继线 010-67616688)

**地 址:** 北京市丰台区方庄芳群园 3 区 3 号楼

**邮 编:** 100078

**网 址:** <http://www.pmph.com>

**E - mail:** [pmph@pmph.com](mailto:pmph@pmph.com)

**购书热线:** 010-67605754 010-65264830

**印 刷:** 北京市卫顺印刷厂

**经 销:** 新华书店

**开 本:** 787×1092 1/16 **印张:** 17.5

**字 数:** 470 千字

**版 次:** 1990 年 6 月第 1 版 2008 年 6 月第 5 版第 28 次印刷

**标准书号:** ISBN 978-7-117-10090-8/R·10091

**定价 (含光盘):** 30.00 元

**版权所有, 侵权必究, 打击盗版举报电话: 010-87613394**

(凡属印装质量问题请与本社销售部联系退换)



# 全国高等学校五年制临床医学专业 第七轮 规划教材修订说明

全国高等学校五年制临床医学专业卫生部规划教材从第一轮编写出版至今已有30年的历史。几十年来,在卫生部的领导和支持下,以裘法祖院士为代表的一大批有丰富临床和教学经验、有高度责任感的老教授和医学教育家参与了本套教材的创建和每一轮的修订工作,使我国的五年制临床医学教材不断丰富、完善与更新,形成了一套课程门类齐全、学科系统优化、内容衔接合理的规划教材。本套教材为推动我国医学教育事业的改革和发展做出了历史性巨大贡献。正如老一辈医学教育家亲切地称这套教材是中国医学教育的“干细胞”教材,由她衍生出了八年制和研究生两套规划教材。今天,全国一大批在临床教学、科研、医疗第一线的中青年教授、学者继承和发扬了老一辈的优良传统,积极参与了本套第七轮教材的修订和建设工作,并借鉴国内外医学教育教 育的经验和成果,不断完善和提升编写的水平和质量,已逐渐将每一部教材打造成了精品,使第七轮教材更加成熟、完善和新颖。

## 第七轮教材的修订从2006年5月开始,其修订和编写特点如下:

●在全国广泛、深入调研基础上,总结和汲取了前六轮教材的编写经验和成果,尤其是对一些不足之处进行了大量的修改和完善,并在充分体现科学性、权威性的基础上,更考虑其全国范围的代表性和适用性。

●依然坚持教材编写“三基、五性、三特定”的原则。

●内容的深度和广度严格控制在五年制教学要求的范畴,精练文字压缩字数,以更适应广大五年制院校的要求,减轻学生的负担。

●在尽可能不增加学生负担的前提下,提高印刷装帧质量,根据学科需要,部分教材改为双色印刷、彩色印刷,以提升教材的质量和可读性。

●适应教学改革的需求,实现教材的系列化、立体化建设,本轮大部分教材配有《学习指导与习题集》、《实验指导》、《教师用书》以及配套光盘等,且与教材同期出版。

第七轮教材共52种,新增1种,即《急诊医学》。全套教材均为卫生部“十一五”规划教材,绝大部分为普通高等教育“十一五”国家级规划教材,分两批于2008年出版发行。

# 第七轮 教材目录

1. 医用高等数学 / 第5版 主编 张选群
2. 医学物理学 / 第7版 主编 胡新珉
3. 基础化学 / 第7版 主编 魏祖期
4. 有机化学 / 第7版 主编 吕以仙
5. 医学生物学 / 第7版 主编 傅松滨
6. 系统解剖学 / 第7版 主编 柏树令
7. 局部解剖学 / 第7版 主编 彭裕文
8. 组织学与胚胎学 / 第7版 主编 邹仲之 李继承
9. 生物化学 / 第7版 主编 查锡良
10. 生理学 / 第7版 主编 朱大年
11. 医学微生物学 / 第7版 主编 李凡 刘晶星
12. 人体寄生虫学 / 第7版 主编 李雍龙
13. 医学免疫学 / 第5版 主编 金伯泉
14. 病理学 / 第7版 主编 李玉林
15. 病理生理学 / 第7版 主编 金惠铭 王建枝
16. 药理学 / 第7版 主编 杨宝峰
17. 医学心理学 / 第5版 主编 姚树桥 孙学礼
18. 法医学 / 第5版 主编 王保捷
19. 诊断学 / 第7版 主编 陈文彬 潘祥林
20. 医学影像学 / 第6版 主编 吴恩惠 冯敢生
21. 内科学 / 第7版 主编 陆再英 钟南山
22. 外科学 / 第7版 主编 吴在德 吴肇汉
23. 妇产科学 / 第7版 主编 乐杰
24. 儿科学 / 第7版 主编 沈晓明 王卫平
25. 神经病学 / 第6版 主编 贾建平
26. 精神病学 / 第6版 主编 郝伟
27. 传染病学 / 第7版 主编 杨绍基 任红
28. 眼科学 / 第7版 主编 赵堪兴 杨培增
29. 耳鼻咽喉-头颈外科学 / 第7版 主编 田勇泉
30. 口腔科学 / 第7版 主编 张志愿
31. 皮肤性病学 / 第7版 主编 张学军
32. 核医学 / 第7版 主编 李少林 王荣福
33. 流行病学 / 第7版 主编 王建华
34. 卫生学 / 第7版 主编 仲来福
35. 预防医学 / 第5版 主编 傅华
36. 中医学 / 第7版 主编 李家邦
37. 计算机应用基础 / 第4版 主编 邹赛德
38. 体育 / 第4版 主编 裴海泓
39. 医学细胞生物学 / 第4版 主编 陈誉华
40. 医学分子生物学 / 第3版 主编 药立波
41. 医学遗传学 / 第5版 主编 左伋
42. 临床药理学 / 第4版 主编 李俊
43. 医学统计学 / 第5版 主编 马斌荣
44. 医学伦理学 / 第3版 主编 丘祥兴 孙福川
45. 临床流行病学 / 第3版 主编 王家良 王滨有
46. 康复医学 / 第4版 主编 南登崑
47. 医学文献检索 / 第3版 主编 郭继军
48. 卫生法 / 第3版 主编 赵同刚
49. 医学导论 / 第3版 主编 文历阳
50. 全科医学概论 / 第3版 主编 杨秉辉
51. 麻醉学 / 第2版 主编 曾因明
52. 急诊医学 主编 沈洪

## 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 樊小力

秘书 孙利军

21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到重要的促进作用。医学统计学作为医学生建立统计分析思维的入门课程是非常必要的,有利于医学生对临床工作和科学研究进行科学的整理和发表,有利于对多种渠道获取的信息进行分析和批判,了解其有用性和局限性。

随着计算机科学技术的迅猛发展,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,现在使用统计软件包瞬间就能准确地计算出来,大大促进了医学统计学的发展。统计软件包的使用为熟悉统计基本原理和统计方法的用户将带来极大的方便,而对于那些统计概念不清,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5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,尤其把各种统计方法的使用条件,使用方法及统计结果的理解作正确解释,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实验设计与诊断试验的评价等,都进行了较详细的讲解。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教高〔2007〕1号文件“积极探索专业评估制度改革,重点推进工程技术、医学等领域的专业认证试点工作,逐步建立适应职业制度需要的专业认证体系”、纽约中华医学基金会的《全球医学教育最低基本要求》、世界卫生组织西太平洋地区本科医学教育质量保障指南等文件等精神,医学生要具有掌握和使用信息技术的能力。因此,本次教材第十四章以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,既能用统计软件进行数据处理,也有利于学生熟悉、掌握计算机技术。

3. 在每一章的练习题中,都有“单项选择题”,熟悉该题型,有助于学生参加医师资格考试。

4. 在每一章的末尾都增加了与章节内容相关的英语片断,有助于学生了解相应部分内容的英文表达。

5. 本教材配有光盘,其内容为各章的“PPT”文件和第十四章以SPSS for Windows Ver11.5为例的上机操作的视频教材两部分。“PPT”文件由各章的作者编写,主编仅根据PPT文件的编写规范,从文本的角度进行了整合,对各章的素材、超媒体、超链接未作变动,保留了各章的风格。便于使用本教材的老师根据自己的兴趣、教学经验、艺术修养,参考本文档的素材、超媒体、超链接的思路,以某种风格制作自己的PPT文件,有





创造性地进行教学。以 SPSS for Windows Ver11.5 为例的上机操作的视频教材将有助于学生自学和教师教学。

本光盘仅供使用本教材的师生使用。未经授权，不得以光盘的内容从事商业及其他盈利性活动。

本教材的参考教学学时为 32~54 学时，建议上机 6~10 学时。

本书附有中英名词索引和英中名词索引，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

2008年2月28日

# 目 录

<b>第一章</b>	<b>医学统计中的基本概念</b> .....	1
第一节	绪论 / 1	
第二节	医学统计工作的内容及资料类型 / 2	
第三节	医学统计中的基本概念 / 3	
	练习题 / 6	
<b>第二章</b>	<b>集中趋势的统计描述</b> .....	8
第一节	频数分布 / 8	
第二节	平均数 / 10	
	练习题 / 14	
<b>第三章</b>	<b>变异程度的统计描述</b> .....	16
第一节	衡量变异程度的指标 / 16	
第二节	正态分布及应用 / 19	
第三节	医学参考值范围 / 22	
	练习题 / 24	
<b>第四章</b>	<b>抽样误差与假设检验</b> .....	27
第一节	均数的抽样误差与标准误 / 27	
第二节	总体均数的估计 / 28	
第三节	假设检验的意义和步骤 / 30	
	练习题 / 32	
<b>第五章</b>	<b>t 检验</b> .....	35
第一节	单样本 $t$ 检验 / 35	
第二节	配对样本均数 $t$ 检验 / 35	
第三节	两独立样本均数 $t$ 检验 / 37	
第四节	方差不齐时两样本均数的 $t'$ 检验 / 39	
第五节	$u$ 检验 / 41	
第六节	$t$ 检验中的注意事项 / 41	
第七节	假设检验中两类错误 / 42	
	练习题 / 43	
<b>第六章</b>	<b>方差分析</b> .....	46
第一节	完全随机设计的方差分析 / 46	
第二节	随机区组设计的方差分析 / 48	
第三节	多个样本均数的两两比较 / 50	
第四节	$2 \times 2$ 析因设计的方差分析 / 52	



第五节 方差齐性检验 / 54

练习题 / 55

**第七章 相对数及其应用** ..... 58

第一节 相对数的概念及计算 / 58

第二节 相对数使用应注意的问题 / 60

第三节 率的标准化 / 60

第四节 医学中常用的相对数指标 / 62

第五节 率的抽样误差与区间估计 / 63

练习题 / 64

**第八章  $\chi^2$  检验** ..... 67

第一节 四格表资料的  $\chi^2$  检验 / 67

第二节 配对四格表资料的  $\chi^2$  检验 / 71

第三节 行 $\times$ 列表资料的  $\chi^2$  检验 / 72

练习题 / 76

**第九章 非参数检验** ..... 79

第一节 配对资料的符号秩和检验 / 79

第二节 两样本比较的秩和检验 / 81

第三节 多个样本比较的秩和检验 / 84

练习题 / 86

**第十章 线性相关与回归** ..... 90

第一节 线性相关 / 90

第二节 线性回归 / 93

第三节 线性相关和回归的区别与联系 / 97

第四节 等级相关 / 98

练习题 / 99

**第十一章 多元线性回归与多元逐步回归** ..... 102

第一节 多元线性回归 / 102

第二节 多元逐步回归 / 107

第三节 多元线性回归的注意事项 / 109

练习题 / 109

**第十二章 统计表与统计图** ..... 113

第一节 统计表 / 113

第二节 统计图 / 116

练习题 / 121





<b>第十三章 医学实验设计与诊断试验的评价</b> .....	124
第一节 医学实验设计的三个要素 / 124	
第二节 医学实验研究设计的几个原则 / 125	
第三节 诊断试验的评价 / 127	
第四节 诊断试验的一致性检验 / 130	
练习题 / 131	
<b>第十四章 SPSS for Windows 实习</b> .....	133
第一节 SPSS for Windows 概述 / 133	
第二节 SPSS for Windows 的启动和退出 / 133	
第三节 数据文件的建立 / 135	
第四节 $t$ 检验 / 142	
第五节 方差分析 / 150	
第六节 $\chi^2$ 检验 / 163	
第七节 非参数检验 / 179	
第八节 一元线性相关与回归 / 185	
第九节 多元线性回归 / 189	
<b>附表</b> .....	194
<b>中英名词索引</b> .....	213
<b>英中名词索引</b> .....	218
<b>练习题答案</b> .....	223
<b>参考文献</b> .....	267

# 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、资料分析直到最后作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。在需要运用统计学原理的场合，如果不能合理地运用则可能造成不应有的缺陷或得出错误的结论。

北京某医院某医生使用“乌贝散”（乌贼骨 3 钱，白芨 3 钱，川贝 3 钱，甘草 3 钱，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，其他医生使用“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？这是一个求置信区间的问题。

北京某医院有位医生，用“冠心病”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 冠心病与单纯西药疗效比较

试验用药	例数	显效	有效	无效
单纯西药	40	9	25	6
冠心病	42	19	18	5

请问冠心病是否比单纯西药有效？这就涉及到统计学中的显著性检验（test of significance），简称计算  $P$  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，其随访调查的部分资料如表 1-2 所示。

表 1-2 心肌梗死抢救治疗后的随访研究

随访时间	随访数	死于本病	中断随访
第一年	1 342	130	41
第二年	1 171	48	53
第三年	1 070	42	73
第四年	955	56	48
第五年	851	35	0

根据这些资料，可提出如下几个问题：

- 出院后第一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，有赖于一系列统计学方法。

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%，平均疗程为三周左右，问此药治疗玫瑰糠疹的确有效吗？由于玫瑰糠疹具有自然痊愈性，一般不服药，多喝水，到三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。

再譬如，用某药治某病，治一例有效，能说 100% 有效吗？治疗二例都有效，能说明 100% 有效吗？治疗十例，其中九例有效，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及资料类型

### 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。

1. 实验设计 根据研究的目的，制定总的研究方案。具体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。

2. 收集资料 根据研究的目的，实验设计的要求，收集准确的完整的含有丰富信息的原始资料 (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；医疗卫生工作记录有门诊病历卡、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。这些资料的收集过程，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。

3. 整理资料 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：

(1) 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，并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 10 000 克，男性患者的调查表中不应出现妇科疾病等。

(2) 一致性检查：从专业的角度，如诊断和疗效的评定标准是否统一，又如胃镜下萎缩性胃炎的严重程度的评定标准是否一致等。

(3) 原始数据的加工：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。

4. 分析资料 对经过统计整理的资料进行一系列统计描述和统计推断，阐明事物的规律性。特别要注意，各种资料使用的统计学方法会有所不同，本书在后面几章将作详细介绍。

学习统计学方法应着重于：

- 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。

1. 计量资料 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 (mg/l)、脉搏 (次/分)、血磷 (mg%)、血红蛋白 (g%) 等。

2. 计数资料 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。

3. 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, 将化验结果按一、+、++、+++ 等级分组, 计数得到的每组病人人数, 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, 可将其若按血红蛋白正常与异常分为两组, 得出各组的人数, 转化为计数资料; 若按血红蛋白含量的多少分为五个等级:  $<6(g\%)$  (重度贫血)、 $6\sim9(g\%)$  (中度贫血)、 $9\sim12.5(g\%)$  (轻度贫血)、 $12.5\sim16(g\%)$  (血红蛋白正常)、 $>16(g\%)$  (血红蛋白增高), 计算各等级人数, 就是等级分组资料。

### 第三节 医学统计中的基本概念

#### 一、变异

医学研究的对象是有机的生命体, 其功能十分复杂, 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 (variance)。变异是由众多的、偶然的、次要的因素造成的。

由于医学统计学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来代替一般健康成人的红细胞数; 也不能因为用青木香治疗一例高血压患者有效, 而断定青木香治疗高血压有效率为 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、总体和样本

总体 (population) 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, 我们不可能把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), 所抽得的部分称为样本 (sample), 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)。如何正确地由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

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)”，两个样本之间应具有可比性 (comparability)。

#### (一) 代表性

要求样本能够充分反映总体的特征。为此应根据研究目的对总体有一个明确的规定。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 12.5 克以上、血小板在 10 万以上的华北地区 18 岁以上的男子，必须做肝功、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

即需要保证总体中的每个个体都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。表面上看是随机的，实际却是不随机的。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，可用机械抽样法、分层抽样法及利用随机数字表抽样等方法。

1. 机械抽样 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 1 000 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 1 000 名工人抽取 200 人作检查，那么可在 0 到 9 十个数字中随机抽取两个数字，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。

2. 分层抽样 (stratified sampling) 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。

3. 随机数字表法 目前多中心的临床试验中，比较多的是根据中心数、病例数、分组数，由生物统计学工作者，设置种子数，用通用统计软件生成随机分组表。在许多统计书上也找到随机数字表，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, 找到“样本例数估计”的有关内容, 本书从略。

#### (四) 可比性

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。如果进行两个或多个样本之间的比较, 则要求各样本之间应具有可比性。

譬如, 有人研究两种不同中草药对慢性支气管炎的疗效, 用两组病人作比较, 第一组病人选取农民, 治疗时间在12~2月间, 第二组病人选取工厂工人, 治疗时间在3~5月间, 其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, 慢性支气管炎的症状可自行缓解。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, 相互作用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, 设置对照组非常必要, 但有时也会比较困难。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。

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, 对照组通常是采用已知的当前治疗方法中最好的一种。

## 四、完全随机设计与随机区组设计

在试验或科研的开始阶段, 我们就应考虑把试验对象按不同的方法进行分组和设计, 常用的有完全随机设计(completely random design)、配对设计(paired design)和随机区组设计(randomized block design)等。设计不同不仅分组方法不同, 而且处理数据所使用的统计学方法也不同。

1. 完全随机设计 将受试对象随机分配到各处理组或对照组中, 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。

2. 配对设计 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, 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。

3. 随机区组设计 也称为配伍组设计, 它是配对设计的扩展。配对设计的每一“对子”中的受试对象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中。



## 五、误差

统计上所说的误差 (error), 泛指观测值 (observed value) 与真实值之差, 以及样本统计量与总体参数之差, 主要有以下三类。

1. 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, 称作系统误差 (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, 予以校正。

2. 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, 实验操作员操作技术不稳定, 不同实验操作员之间的操作差异, 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, 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。

3. 抽样误差 由抽样不同引起的样本均数 (或其他统计量) 与总体均数 (或其他参数) 之间的差异称作抽样误差 (sampling error)。即使消除了系统误差, 并把随机测量误差控制在允许范围内, 这种误差仍不可避免。抽样误差产生的原因是: 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。抽样误差可以用统计方法进行分析。

一般说来, 样本含量越大, 则抽样误差越小, 样本统计量与总体的参数越接近, 即越能说明总体的规律。反之, 样本含量越小, 则抽样误差相应的越大。因此, 我们不能仅仅凭观察指标的大小进行简单判断, 而应该使用统计学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。

## 六、概率

概率 (probability) 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,  $f$  为某事件发生的频数。如高血压的患病率、乳腺癌术后五年的生存率等。

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。

## 练习题

### 一、单项选择题

1. 医学统计学研究的对象是

- A. 医学中的小概率事件
- B. 各种类型的数据
- C. 动物和人的本质
- D. 疾病的预防与治疗



- E. 有变异的医学事件
2. 用样本推论总体，具有代表性的样本指的是
- A. 总体中最容易获得的部分个体  
B. 在总体中随意抽取任意个体  
C. 挑选总体中的有代表性的部分个体  
D. 用配对方法抽取的部分个体  
E. 依照随机原则抽取总体中的部分个体
3. 下列观测结果属于等级资料的是
- A. 收缩压测量值  
B. 脉搏数  
C. 住院天数  
D. 病情程度  
E. 四种血型
4. 随机误差指的是
- A. 测量不准引起的误差  
B. 由操作失误引起的误差  
C. 选择样本不当引起的误差  
D. 选择总体不当引起的误差  
E. 由偶然因素引起的误差
5. 收集资料不可避免的误差是
- A. 随机误差  
B. 系统误差  
C. 过失误差  
D. 记录误差  
E. 仪器故障误差

## 二、简答题

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. 抽样中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)

### Introduction

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.

## 第二章 集中趋势的统计描述

本章学习计量资料的统计描述方法。这一部分的内容主要包括统计图表和统计指标，利用统计图表可以清楚地揭示资料的分布类型和特征，利用统计指标则可以概括性地描述一组资料的集中趋势或平均水平。学习这些方法的目的在于能够有效地组织、整理和表达统计资料的信息。

### 第一节 频数分布

通过实验或临床观察等各种方式得到的原始资料，如果是计量资料并且观察的例数较多，可以对数据进行分组，然后制作频数表或绘制直方图，用以显示数据的分布规律。

#### 一、频数表

频数表 (frequency table) 是指如下一种格式的统计表，即同时列出观察指标的可能取值区间及其在各区间内出现的频数。具体作法是，先根据观察个体的数量大小进行分组，然后计算每组中观察值出现的次数。由于这种资料的表达方式较完整地体现了观察值的分布规律，所以也称作频数分布表。

用手工整理资料编制频数表时，通常先编制划记表，即先将选定的组列出，每一组段的起点称下限，终点称上限（上限一般不列出），然后在原始数据中逐一观察，观察到的数据应当归入哪一组，就在划记表的相应位置上划记，划够五道成一个“正”字。将全部数据划记完毕后计算各组中“正”字的笔画数目，即可得到各组的频数。

现结合实例说明频数表的编制方法和应注意的问题。

**例 2.1** 某地用随机抽样方法检查了 140 名正常成年男子的红细胞数，检测结果如表 2-1 所示：

表 2-1 某地 140 名正常成年男性红细胞数 ( $\times 10^{12}/L$ )

4.76	5.26	5.61	5.95	4.46	4.57	4.31	5.18	4.92	4.27	4.77	4.88
5.00	4.73	4.47	5.34	4.70	4.81	4.93	5.04	4.40	5.27	4.63	5.50
5.24	4.97	4.71	4.44	4.94	5.05	4.78	4.52	4.63	5.51	5.24	4.98
4.33	4.83	4.56	5.44	4.79	4.91	4.26	4.38	4.87	4.99	5.60	4.46
4.95	5.07	4.80	5.30	4.65	4.77	4.50	5.37	5.49	5.22	4.58	5.07
4.81	4.54	3.82	4.01	4.89	4.62	5.12	4.85	4.59	5.08	4.82	4.93
5.05	4.40	4.14	5.01	4.37	5.24	4.60	4.71	4.82	4.94	5.05	4.79
4.52	4.64	4.37	4.87	4.60	4.72	4.83	5.33	4.68	4.80	4.15	4.65
4.76	4.88	4.61	3.97	4.08	4.58	4.31	4.05	4.16	5.04	5.15	4.50
4.62	4.73	4.47	4.58	4.70	4.81	4.55	4.28	4.78	4.51	4.63	4.36
4.48	4.59	5.09	5.20	5.32	5.05	4.41	4.52	4.64	4.75	4.49	4.22
4.71	5.21	4.94	4.68	5.17	4.91	5.02	4.76				

1. 确定组数 对数据进行分组时首先要考虑的是应取的组数，分组过少会导致过多的信息损失，资料的表达过于粗略；分组过多则会使资料过于分散，分布的规律性不能明显地表示出来。组数通常选择在 8~15 之间，若资料在 100 例以上，一般取 10 个组左右；若例数较少，组数可相应的减少。总之，以能显示数据的分布规律为宜。



2. 确定组距 分组时必须事先规定组距, 组距的宽度按相邻两组的下限之差计算, 一般应取相同的组距。将全距除以组数可以得到组距的近似值。全距即数据中最大值与最小值之差。如果用  $R$  表示全距,  $k$  表示组数, 则参考组距为  $R/k$ 。组距的选择应符合专业习惯, 得到参考组距后再结合实际情况做适当地调整。如表 2-1 数据中的最大值为 5.95, 最小值为 3.82, 全距为 2.13, 现拟分 10 个组, 参考组距为

$$i = \frac{R}{k} = \frac{2.13}{10} \approx 0.21$$

因为通常比较习惯含有 0.20 或 0.25 个单位倍数的区间宽度, 因此可以在 0.20 和 0.25 两种组距中选择。若用 0.20 作为组距, 则总共有 11 个区间; 若用 0.25 作为组距, 则有 9 个区间。权衡后选择组距为 0.20 比较合适。

有些资料因为数据中有特大或特小的数值也可以采用不等组距, 如某些食物中毒的潜伏期, 大部分人在短时间内出现症状, 只有极少部分人经过较长的时间后才有反应, 这种情况可以将后面的一些组段作适当的合并, 并将最后一个组段以“大于……”表示, 以避免出现一些组段的频数为零的情况。

3. 确定组段 一个频数分布必须包括整个资料范围的全部数据, 即一个数据必须能够归属于某一组, 同时只能归属于一个组, 不能兼属。为此, 实际组段在每组中只包含下限而不包含上限, 如第一组为“3.80~”、第二组为“4.00~”, 凡小于 4.00 的测量值均应分入第一组, 大于或等于 4.00 且小于 4.20 的测量值归在第二组。

4. 用手工编制划记表, 可得到例 2.1 的频数表, 如表 2-2 所示。

表 2-2 某地 140 名正常成年男子红细胞数的频数表

红细胞数( $\times 10^{12}/L$ )	划记	组中值	频数	频率(%)
(1)	(2)	(3)	(4)	(5)
3.80~	丁	3.90	2	1.4
4.00~	正丁	4.10	6	4.3
4.20~	正正丁	4.30	11	7.9
4.40~	正正正正正	4.50	25	17.9
4.60~	正正正正正正丁	4.70	32	22.9
4.80~	正正正正正正丁	4.90	27	19.3
5.00~	正正正丁	5.10	17	12.1
5.20~	正正正	5.30	13	9.3
5.40~	正	5.50	4	2.9
5.60~	丁	5.70	2	1.4
5.80~6.00	—	5.90	1	0.7

## 二、直方图

将表 2-1 的资料编成频数表后, 可以看出数据的分布情况, 若绘成直方图 (histogram) 则更直观。直方图是以垂直条段代表频数分布的一种图形, 条段的高度代表各组的频数, 由纵轴标度; 各组的组限由横轴标度, 条段的宽度表示组距。如将表 2-2 资料绘制直方图, 如图 2.1 所示。

从图 2-1 明显看出该地区成年男性红细胞数的分布特点: 数据多集中在  $4.80 \times 10^{12}/L$  附近, 两侧对称下降, 最小值不低于  $3.80 \times 10^{12}/L$ , 最大值不超过  $6.00 \times 10^{12}/L$ 。绘制直方图的频数表资料一般为等距分组, 对于不等距资料应先将不等距的各组频数折算为等距频数, 然后再作图。



### 三、频数分布表的用途

1. 作为陈述资料的形式，可以代替原始资料，便于进一步分析。

2. 便于观察数据的分布类型 医学研究中常见的资料分布类型可分为对称分布和偏态分布两大类。在对称分布资料中有一种非常重要的分布类型—正态分布 (normal distribution)，其特征是

中间组段的频数最多，两侧的频数分布对称，并按一定的规律下降，表 2-2 的频数分布即近似呈正态分布 (参见图 2-1)。如果频数分布的高峰向左偏移，长尾向右侧延伸称为正偏态分布，相反则称为负偏态分布。在统计分析时常需要根据资料的分布形式选择相应的统计分析方法。

3. 便于发现资料中某些远离群体的特大或特小的可疑值 如在频数表中连续出现 0 的频数后，又出现了一些频数就值得怀疑，应进行检查和核对，必要时可通过统计检验决定取舍。

4. 当样本含量比较大时，可用各组段的频率作为概率的估计值 如表 2-2 第 (4) 栏的频数除以总例数即为第 (5) 栏的频率，由此可推测成年男子的红细胞计数值出现在各组段的概率分别为 0.014, 0.043, 0.079, ..., 0.007。

总之，通过频数分布表和直方图，可以大致看出观察值的形态和特征。如果需要进一步用数字概括、明确地描述频数分布的特征则应使用统计指标描述的方法。

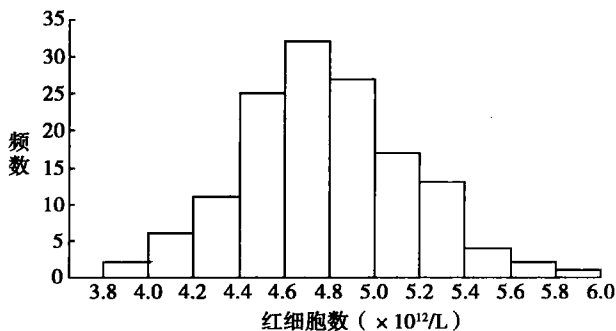


图 2-1 140 名正常成年男子红细胞计数的直方图

## 第二节 平均数

平均数 (average) 是描述一组观察值集中位置或平均水平的统计指标，它常作为一组数据的代表值用于分析和进行组间的比较。平均数有多种，常用的有算术均数、几何均数和中位数等。

### 一、算术均数

算术均数 (mean) 简称为均数，用于说明一组观察值的平均水平或集中趋势，是描述计量资料的一种最常用的方法。均数计算有直接法和加权法。

#### (一) 直接法

将所有的观察值  $X_1, X_2, \dots, X_n$  直接相加再除以观察例数，写成公式为

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{\sum X}{n} \quad (2-1)$$

式中， $\bar{X}$  表示样本均数； $\sum$  是希腊字母 (读作 sigma)，表示求和的符号； $n$  为样本观察例数。如对例 2.1 的数据用上面公式计算，可算得 140 名正常成年男子红细胞数的均值为

$$\bar{X} = \frac{4.76 + 5.26 + 5.61 + \dots + 5.02 + 4.76}{140} = 4.77 (\times 10^{12} / L)$$

#### (二) 加权法

加权法是根据频数表计算均数的一种方法。当观察例数较多时，资料通常要分组编成频数表，对已经编成频数表的资料不能再公式 (2-1) 计算，这种情况下我们可以把各





组的组中值视为各组观察值的代表值, 分别乘以各组的频数得到各组观察值之和, 然后将它们相加得到观察值的总和再除以总例数。用公式表示如下:

$$\bar{X} = \frac{f_1 x_1 + f_2 x_2 + \cdots + f_k x_k}{n} = \frac{\sum f x}{n} \quad (2-2)$$

式中  $k$  表示频数表的组段数;  $f_1, f_2, \cdots, f_k$  及  $x_1, x_2, \cdots, x_k$  表示 1 至  $k$  组的频数及组中值, 见表 2-2 第 (3) 栏和第 (4) 栏; 组中值 = (本组段下限 + 本组段上限) / 2, 如 “4.00~” 组段的组中值为  $(4.00 + 4.20) / 2 = 4.10$ , 余类推。

将表 2-2 的数据代入公式 (2-2), 有

$$\bar{X} = \frac{2 \times 3.90 + 6 \times 4.10 + 11 \times 4.30 + \cdots + 2 \times 5.70 + 1 \times 5.90}{140} = 4.78 (\times 10^{12} / L)$$

由此可见, 在样本例数较多的情况下, 加权法与直接法算得的结果相差不大。

### (三) 均数的应用

均数的意义很容易理解, 而且结果也比较稳定, 因而应用极为广泛。但是它主要适用于对称分布或偏度不大的资料, 尤其适合正态分布资料。由于在计算均数时用到了每一个观察值, 在偏态较大的情况下, 算出的均值容易受到频数分布尾端极大或极小值的影响, 不能真正地反映分布的集中位置, 这时应考虑改用其他方法。

## 二、几何均数

医学研究中有一类比较特殊的资料, 如抗体滴度、细菌计数、血清凝集效价、某些物质浓度等, 其数据特点是观察值间按倍数关系变化, 对此可以计算几何均数 (geometric mean) 以描述其平均水平。几何均数用  $G$  表示, 计算公式为

$$G = \sqrt[n]{X_1 X_2 \cdots X_n} \quad (2-3)$$

即将  $n$  个观察值连乘后开  $n$  次方。为了计算方便, 常改用对数的形式计算, 即

$$G = \lg^{-1} \left( \frac{\lg X_1 + \lg X_2 + \cdots + \lg X_n}{n} \right) = \lg^{-1} \left( \frac{\sum \lg X}{n} \right) \quad (2-4)$$

可以看出, 几何均数相当于各观察值对数的均值再取反对数。对于频数表资料, 若用  $f_1, f_2, \cdots, f_k$  及  $x_1, x_2, \cdots, x_k$  表示 1 至  $k$  组的频数及各组中位取值, 可用下式计算:

$$G = \lg^{-1} \left( \frac{f_1 \lg x_1 + f_2 \lg x_2 + \cdots + f_k \lg x_k}{n} \right) = \lg^{-1} \left( \frac{\sum f \lg x}{n} \right) \quad (2-5)$$

**例 2.2** 测得 10 个人的血清滴度的倒数分别为 2, 2, 4, 4, 8, 8, 8, 8, 32, 32, 求平均滴度。

如果计算均数, 其值为  $\bar{X} = 10.8$ , 现计算几何均数

$$G = \lg^{-1} \left( \frac{\lg 2 + \lg 2 + \lg 4 + \lg 4 + \lg 8 + \lg 8 + \lg 8 + \lg 8 + \lg 32 + \lg 32}{10} \right) \approx 7$$

显然在这里均数不能代表其平均水平, 选择几何均数则比较合适。故 10 份血清滴度的平均水平为 1:7。

**例 2.3** 某医师使用胎盘浸液钩端螺旋体菌苗对 326 名农民接种两月后测得血清 IgG 抗体滴度如表 2-3, 试计算平均抗体滴度。

表 2-3 胎盘浸液钩端螺旋体菌苗接种两月后血清 IgG 抗体滴度

IgG 滴度倒数	例数	IgG 滴度倒数	例数
20	16	320	54
40	57	640	25
80	76	1 280	23
160	75		



$$G = \lg^{-1} \left( \frac{16 \lg 20 + 57 \lg 40 + 76 \lg 80 + 75 \lg 160 + 54 \lg 320 + 25 \lg 640 + 23 \lg 1280}{326} \right) \approx 139$$

即胎盘浸液钩端螺旋体菌苗接种 2 个月后血清 IgG 抗体的平均滴度为 1 : 139。

几何均数在医学研究领域多用于血清学和微生物学中。有些明显呈偏态分布的资料经过对数变换后呈对称分布，也可以采用几何均数描述其平均水平，但要注意观察值中不能有 0 或负数，否则在作对数变换之前需要加一个常数。同一组观察值的几何均数总是小于它的算术均数。

### 三、中位数和百分位数

#### (一) 中位数

将一组观察值按从小到大顺序排列  $X_1 \leq X_2 \leq \dots \leq X_n$ ，居中心位置的数值即为中位数 (median)，记为  $M$ 。当观察例数  $n$  为奇数时，中位数是按顺序排列在第  $(n+1)/2$  项的观察值；当观察例数为偶数时，则中位数是按顺序排列在第  $n/2$  项和第  $(n/2)+1$  项之间，即该两项的平均值。在全部观察值中有半数的值比  $M$  大，另有半数的值比  $M$  小。例如，现测得 5 个人的极低密度脂蛋白 (VLDL) 中的载脂 B 蛋白的含量 (mmol/L) 分别为

0.0095, 0.0322, 0.0617, 0.0970, 0.1085

中位数等于 0.0617mmol/L；若仅测量了四个值：

0.0095, 0.0322, 0.0617, 0.0970

则中位数为  $(0.0322+0.0617)/2=0.04695$ (mmol/L)。

另外也可以根据频数表资料计算中位数。下面通过实例说明它的计算方法。

**例 2.4** 对某地 630 名 50 岁~60 岁的正常女性检查了血清甘油三酯含量 (mmol/L)，资料如表 2-4，试计算其中位数。

表 2-4 某地 630 名正常女性血清甘油三酯含量 (mmol/L) 的频数表

甘油三酯 (1)	频数 (2)	累积频数 (3)	累积频率(%) (4)
0.10~	27	27	4.3
0.40~	169	196	31.1
0.70~ $\checkmark$	167	363	57.6 $M$
1.00~	94	457	72.5
1.30~	81	538	85.4
1.60~	42	580	92.1
1.90~	28	608	96.5
2.20~	14	622	98.7
2.50~	4	626	99.4
2.80~	3	629	99.8
3.10~	1	630	100.0
合计	630	—	—

计算频数分布中位数的第一步是确定它所在区间的位置。资料总例数是 630，630 的 50% 是 315，当观察值按顺序排列时，第 315 个观察值就是中位数  $M$ 。从累积频数分布可以看出，第 315 个观察值处于 [0.70, 1.00] 区间内，第 (4) 栏的累积频率则更清楚地显示了它所在的区间。假定该区间内的观察值呈均匀分布，因为在此区间之前有 196 例，在区间内共有 167 例，所以可以推想，中位数位于该区间内离下限的距离为该区间宽度的



$(315-196)/167$  处。为了得出中位数的值,可将  $(315-196)/167$  乘以区间宽度 0.30,再将此乘积项加到所属区间的下限上去(图 2-2)。于是有

$$M=0.70+\frac{630\times 0.5-196}{167}\times 0.30=0.914(\text{mmol/L})$$

一般,可用以下公式计算中位数:

$$M=L+\left(\frac{0.5n-f_L}{f_M}\right)i_M \quad (2-6)$$

式中  $L$ 、 $i_M$ 、 $f_M$  分别为  $M$  所在组段的下限、组距和频数,  $f_L$  为  $M$  所在组段之前各组段的累积频数。

### (二) 百分位数

中位数可以用来描述一组观察值的中心位置。但有时我们还需要了解数据分布的其他位置,如资料分布的左侧占全部 25% 的位置,这时可以通过计算百分位数(per-

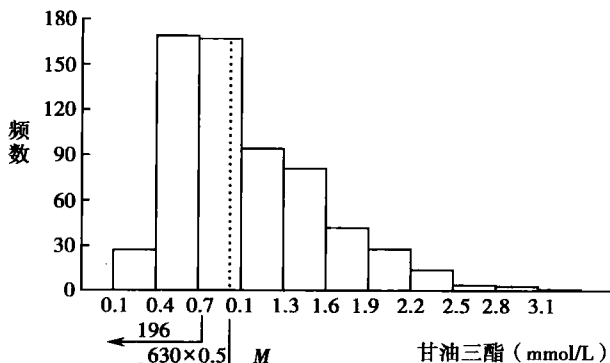


图 2-2 中位数计算方法示意图

centile) 确定。百分位数用符号  $P_x$  表示,  $x$  即百分位,所谓百分位数  $P_x$  是指在一组数据中找到这样一个值,全部观察值的  $x\%$  小于  $P_x$ ,其余  $(100-x)\%$  大于  $P_x$ 。如  $P_{25}$  表示资料在  $P_{25}$  位置左侧的累积频数占总数的 25%,右侧占 75%。 $P_{50}$  实际就是中位数  $M$ 。百分位数的手工计算可以根据频数表来做,计算原理与中位数完全相同,对式(2-6)只需将中位数换成任意百分位数,即

$$P_x=L+\left(\frac{n\cdot x\%-f_L}{f_x}\right)i_x \quad (2-7)$$

式中  $L$ 、 $i_x$ 、 $f_x$  分别为  $P_x$  所在组段的下限、组距和频数,  $f_L$  为  $P_x$  所在组段之前各组段的累积频数。

**例 2.5** 计算例 2.4 的百分位数  $P_{25}$ 、 $P_{75}$ 、 $P_{90}$ 。

利用公式(2-7),可分别求得

$$P_{25}=0.40+\frac{630\times 0.25-27}{169}\times 0.30=0.632(\text{mmol/L})$$

$$P_{75}=1.30+\frac{630\times 0.75-457}{81}\times 0.30=1.357(\text{mmol/L})$$

$$P_{90}=1.60+\frac{630\times 0.90-538}{42}\times 0.30=1.807(\text{mmol/L})$$

### (三) 百分位数的应用

1. 中位数与均数、几何均数的作用相同,都能用来反映一组数据的集中位置或平均水平,由于中位数的确定仅取决于它在数据序列中的位置,而不是由全部观察值综合计算出来的,因此不受少数特大或特小数值的影响,在这一点上它优于均数。一般说来,在频数分布呈明显偏态、或频数分布的两端无确定数值时,使用中位数描述集中位置或平均水平较为合理。当变量呈对称分布时,理论上中位数和均数相同。但对于样本资料,由于计算均数时利用了所有的观察值,所以较中位数稳定。另外,中位数还有一个很大的缺点,即不便于作统计运算,如根据两组资料的不同中位数无法算出合并的中位数。因此在统计分析中,中位数的用途远不如均数广泛。

2. 百分位数可用来描述资料的观察值序列在某百分位置的水平,中位数是其中的一个特例。多个百分位数结合使用常可以用来说明某一特定的问题,如用  $P_{25}$  及  $P_{75}$  描述资料



的离散程度,用  $P_{2.5}$  及  $P_{97.5}$  规定医学 95% 的参考值范围,在研究青少年生长发育时用百分位数  $P_5$ 、 $P_{25}$ 、 $P_{75}$ 、 $P_{95}$  划分等级等。百分位数可用于任何频数分布的资料,但靠近两端的百分位数仅在样本例数比较大时才比较稳定(如  $n > 100$ )。

## 练习 题

### 一、单项选择题

- 某医学资料数据大的一端没有确定数值,描述其集中趋势适用的统计指标是
  - 中位数
  - 几何均数
  - 均数
  - $P_{95}$  百分位数
  - 频数分布
- 算术均数与中位数相比,其特点是
  - 不易受极端值的影响
  - 能充分利用数据的信息
  - 抽样误差较大
  - 更适用于偏态分布资料
  - 更适用于分布不明确资料
- 一组原始数据呈正偏态分布,其数据的特点是
  - 数值离散度较小
  - 数值离散度较大
  - 数值分布偏向较大一侧
  - 数值分布偏向较小一侧
  - 数值分布不均匀
- 将一组计量资料整理成频数表的主要目的是
  - 化为计数资料
  - 便于计算
  - 形象描述数据的特点
  - 为了能够更精确地检验
  - 提供数据和描述数据的分布特征
- 6 人接种流感疫苗一个月后测定抗体滴度为 1:20、1:40、1:80、1:80、1:160、1:320,求平均滴度应选用的指标是
  - 均数
  - 几何均数
  - 中位数
  - 百分位数
  - 倒数的均数

### 二、计算与分析

1. 现测得 10 名乳腺癌患者化疗后血液尿素氮的含量 (mmol/L) 分别为 3.43, 2.96, 4.43, 3.03, 4.53, 5.25, 5.64, 3.82, 4.28, 5.25, 试计算其均数和中位数。

2. 某地 100 例 30-40 岁健康男子血清总胆固醇值 (mg/dl) 测定结果如下:

202	165	199	234	200	213	155	168	189	170	188	168	184	147	219	174	130	183	178	174
228	156	171	199	185	195	230	232	191	210	195	165	178	172	124	150	211	177	184	149
159	149	160	142	210	142	185	146	223	176	241	164	197	174	172	189	174	173	205	224
221	184	177	161	192	181	175	178	172	136	222	113	161	131	170	138	248	153	165	182
234	161	169	221	147	209	207	164	147	210	182	183	206	209	201	149	174	253	252	156

(1) 编制频数分布表并画出直方图;

(2) 根据频数表计算均值和中位数,并说明用哪一个指标比较合适;

(3) 计算百分位数  $P_5$ 、 $P_{25}$ 、 $P_{75}$  和  $P_{95}$ 。

3. 测得 10 名肝癌病人与 16 名正常人的血清乙型肝炎表面抗原 (HBsAg) 滴度如下表,试分别计算它们的平均滴度。



肝癌病人与正常人的血清乙肝表面抗原 (HBsAg) 滴度

滴度倒数	正常人数	肝癌病人数
8	7	1
16	5	2
32	1	3
64	3	2
128	0	1
256	0	1

(李 康)

### Measures of Central Tendency: Mean and Median.

One of the most important numerical measures is a measure of center—a measure along the horizontal axis that locates the center of the distribution. There are several commonly used measures to describe the location or center of a population or sample. The most widely utilized measure of central tendency is the arithmetic mean or average.

The arithmetic average of a set of measurement is very common and useful measure. To distinguish between the mean for the sample and the mean for the population, we will use the symbol  $\bar{X}$  for a sample mean and  $\mu$  for a population mean.

One measure of central tendency that is not as sensitive to the value of each measurement is the median. The median can be used as a summary measure for ordinal observations as well as for discrete and continuous data. The *median* is defined as the 50th percentile of a set of measurements; if a list of observations is ranked from smallest to largest, half the values are greater than or equal to the median, where the other half are less than or equal to it. Therefore, if a set of data contains a total of  $n$  observations where  $n$  is odd, the median is the middle value, or the  $[(n+1)/2]$ th largest measurement; if  $n$  is even, the median is usually taken to be the average of the two middlemost values, the  $(n/2)$ th and  $[(n/2)+1]$ th observations.

Although both the mean and the median are good measures of the center of the population, the median is less sensitive to the extreme values or outliers.

## 第三章 变异程度的统计描述

上一章所讲的几种反映集中趋势或平均水平的指标，是描述一组观察值重要特征的最基本的统计量。一般情形下，除要了解观察值的平均水平外，往往还需要同时了解这些观察值之间的变异程度或偏离集中位置的程度。为了说明这一问题，看下面的例子。

例 3.1 对甲乙两名高血压患者连续观察 5 天，测得的收缩压分别如下：

甲患者 (mmHg) 162 145 178 142 186 ( $\bar{X}_甲 = 162.6$ )

乙患者 (mmHg) 164 160 163 159 166 ( $\bar{X}_乙 = 162.4$ )

从列出的数据可以看出，两人收缩压的均数几乎没有差别，但甲患者的血压波动比较大，而乙患者相对比较稳定。因此，描述一组观察值，除需要表示其平均水平外，还要说明它的离散或变异的情况。

### 第一节 衡量变异程度的指标

衡量变异程度大小的指标有多种，但大体可以分为两类：一类是按间距计算，有极差和四分位数间距；另一类则按平均差距计算，有离均差平方和、方差、标准差和变异系数等。

#### 一、极差和四分位数间距

##### (一) 极差

极差 (range) 也称作全距，即观察值中最大值和最小值之差，用符号  $R$  表示，是变异指标中最简单的一种。极差大说明变异程度大，反之说明变异程度小，如上例甲乙两患者收缩压的极差分别为

$$R_甲 = 186 - 142 = 44(\text{mmHg})$$

$$R_乙 = 166 - 159 = 7(\text{mmHg})$$

可见甲患者收缩压的波动大，乙患者波动小。

极差的方法虽然简单，但在某些场合颇有实用价值，如用于说明传染病、食物中毒等的最短、最长潜伏期等。用极差说明数据分布的离散程度，简单明了、容易使用，但由于计算时仅用到了最大值和最小值，而没有利用观察值中的全部信息，随着观察例数的增多，抽到较大或较小数值的可能性越来越大，极差也会随之而变大，尤其当资料呈明显偏态时会显得更加不稳定，所以通常仅用于粗略地说明变量的变动范围。

##### (二) 四分位数间距

极差的不稳定主要受观察值大小两端极端数值的影响，如果将两端的数据去掉一定的比例，所得到的结果就会比较稳定。一种方法是把所有的观察值排序后，分成四个数目相等的段落，每个段落的观察值数目各占总例数的 25%，去掉两端的 25%，取中间 50% 观察值的数据范围即为四分位数间距 (quartile)。四分位数间距用符号  $Q$  表示，它可以通过计算百分位数  $P_{75}$  和  $P_{25}$  之差得到，即  $Q = P_{75} - P_{25}$ 。例如由上一章例 2.4 算出，50~60 岁正常女性血清甘油三酯含量的百分位数  $P_{25}$  和  $P_{75}$  的位置分别为 63.2mg/dl 和 135.7mg/dl， $Q = 135.7 - 63.2 = 72.5\text{mg/dl}$ ，说明有 50% 女性的甘油三脂在 63.2mg/dl 和 135.7mg/dl 之间，其四分位数间距为 72.5mg/dl。四分位数间距越大，说明数据的变异越大；反之，



四分位数间距越小说明变异越小。

用四分位数间距作为说明个体差异的指标，与极差相比不易受极端值的影响，但仍未用到每一个具体的观察值，在统计分析中应用得不够普遍。

## 二、离均差平方和、方差、标准差和变异系数

### (一) 平均偏差

为了利用每一个观察值的信息，一个自然的设想是计算各观察值偏离平均数的平均差距，为了避免正负抵消，可将每个观察值与均数之差的绝对值相加，然后取平均，称作平均偏差 (mean difference)，它可以表示为

$$\text{平均偏差} = \frac{\sum |X - \bar{X}|}{n} \quad (3-1)$$

如对于例 3.1 有

甲患者：

$$\text{平均偏差} = \frac{|162-162.6| + |145-162.6| + \dots + |186-162.6|}{5} = 15.52(\text{mmHg})$$

乙患者：

$$\text{平均偏差} = \frac{|164-162.4| + |160-162.4| + \dots + |166-162.4|}{5} = 2.32(\text{mmHg})$$

说明甲患者血压的波动较乙患者大。

平均偏差是一个很直观的变异量度，但由于用了绝对值，在数学上不便于继续处理，使它在应用上受到很大的限制，实际中很少使用。

### (二) 离均差平方和

为了克服平均偏差使用绝对值不便于进一步运算的缺点，可以不通过取绝对值，而是通过取平方来避免正负抵消，即使用离均差平方和 (sum of square, 简称为 SS)，其计算公式为

$$SS = \sum (X - \bar{X})^2 \quad (3-2)$$

它描述了每个观察值相对于集中位置  $\bar{X}$  的分散程度。通过计算可化为下式：

$$SS = \sum (X - \bar{X})^2 = \sum X^2 - \frac{(\sum X)^2}{n} \quad (3-3)$$

### (三) 方差

将离均差平方和再取平均，其结果称作方差 (variance)。需要注意的是，对于样本资料，在对离均差平方和取平均时分母用  $n-1$  代替  $n$ ，其计算公式为

$$S^2 = \frac{\sum (X - \bar{X})^2}{n-1} \quad (3-4)$$

式中  $S^2$  为样本方差，其值越大说明数据的变异越大。分母  $n-1$  称为自由度 (degree of freedom, 简记为  $df$ )，因为在所有的  $n$  个离均差平方项中，只有  $n-1$  个是独立的，即由于样本均数  $\bar{X}$  的限制，在所有的离均差平方项中只要有  $n-1$  个已知，最后一个离均差平方项便自动确定。

### (四) 标准差

方差的单位是原度量单位的平方，如果原始数据的单位为毫米汞柱，则方差的单位就是毫米汞柱的平方。在统计分析中为了方便，通常将方差取平方根，还原成与原始观察值单位相同的变异量度，其计算公式为

$$S = \sqrt{\frac{\sum (X - \bar{X})^2}{n-1}} \quad (3-5)$$



S 称为标准差 (standard deviation), 有时也记为 SD。显然, 一组观察值的标准差越大说明其变异程度越大。将式中的离均差平方和展开, 标准差也可写为

$$S = \sqrt{\frac{\sum X^2 - (\sum X)^2/n}{n-1}} \quad (3-6)$$

例如对于例 3.1 有

甲患者:  $\sum X = 162 + 145 + 178 + 142 + 186 = 813$   
 $\sum X^2 = 162^2 + 145^2 + 178^2 + 142^2 + 186^2 = 133\ 713$

$$S = \sqrt{\frac{133\ 713 - 813^2/5}{5-1}} = 19.49(\text{mmHg})$$

乙患者:  $\sum X = 164 + 160 + 163 + 159 + 166 = 812$   
 $\sum X^2 = 164^2 + 160^2 + 163^2 + 159^2 + 166^2 = 131\ 902$

$$S = \sqrt{\frac{131\ 902 - 812^2/5}{5-1}} = 2.88(\text{mmHg})$$

说明甲患者血压波动比乙患者血压的波动大。

如果是频数表资料, 可用如下公式计算:

$$S = \sqrt{\frac{\sum fx^2 - (\sum fx)^2/n}{n-1}} \quad (3-7)$$

式中  $x$  和  $f$  分别为各组段的组中值及出现的频数。

**例 3.2** 根据第二章表 2-2 的频数表资料, 计算成年男子红细胞数的标准差。计算表如下:

表 3-1 140 名正常成年男子红细胞计数 ( $\times 10^{12}/L$ ) 的标准差计算表

红细胞计数 (1)	组中值( $x$ ) (2)	频数( $f$ ) (3)	$fx$ (4)	$fx^2$ (5)
3.80~	3.90	2	7.80	30.42
4.00~	4.10	6	24.60	100.86
4.20~	4.30	11	47.30	203.39
4.40~	4.50	25	112.50	506.25
4.60~	4.70	32	150.40	706.88
4.80~	4.90	27	132.30	648.27
5.00~	5.10	17	86.70	442.17
5.20~	5.30	13	68.90	365.17
5.40~	5.50	4	22.00	121.00
5.60~	5.70	2	11.40	64.98
5.80~5.95	5.90	1	5.90	34.81
合计		140	669.80	3 224.20

由公式 (3-7) 得

$$S = \sqrt{\frac{3\ 224.20 - 669.80^2/140}{140-1}} = 0.38(\times 10^{12}/L)$$

即所求标准差为  $0.38 \times 10^{12}/L$ 。

在所有的变异指标中, 标准差或方差是其他变异指标所不能比拟的。首先标准差能够直接用于代数运算, 如根据来自同一总体的几个样本的标准差, 可以直接求得合并样本的标准差, 而不必根据合并样本重新计算。还有很重要的一点, 标准差与正态分布有明确的关系, 它与均数结合能够完整地概括一个正态分布。今后我们会逐渐看到怎样利用这些特





点进行各种运算。

### (五) 变异系数

标准差使用的度量衡单位与原始数据相同, 在两组数据的均数相差不大, 单位也相同时, 从标准差的大小就可以直接比较两个样本的变异程度。但是, 有时我们需要对均数相差较大或单位不同的几组观察值的变异程度进行比较, 这时使用标准差就不适宜了 (极差或四分位数间距同样不合适)。在这种情况下可以使用变异系数 (coefficient of variation, 简记为 CV), 其计算公式为

$$CV = \frac{S}{\bar{X}} \times 100\% \quad (3-8)$$

**例 3.3** 测得某地成年人舒张压的均数为 77.5mmHg, 标准差为 10.7mmHg; 收缩压的均数为 122.9mmHg, 标准差为 17.1mmHg。试比较舒张压和收缩压的变异程度。

舒张压和收缩压是两个不同的指标, 尽管单位相同, 但均数相差较大, 如直接比较两个标准差, 会得出收缩压变异较大的结论。现计算两者的变异系数

$$\text{舒张压} \quad CV = \frac{10.7}{77.5} \times 100\% = 13.8\%$$

$$\text{收缩压} \quad CV = \frac{17.1}{122.9} \times 100\% = 13.9\%$$

可见两种指标的变异度几乎没有什么差别。

## 第二节 正态分布及应用

### 一、正态分布

前面提到, 标准差与均数结合能够完整地描述一个正态分布。那么什么是正态分布呢? 我们可以用直方图加以说明。以上一章图 2.1 为例, 现将表示红细胞数的横坐标用变量  $X$  表示, 第  $i$  组的组距和人数分别用  $\Delta X_i$  和  $f_i$  表示,  $n$  为总观察例数, 则在  $[X, X + \Delta X_i)$  区间内每单位红细胞数的频率为

$$f(X) = (f_i/n) / \Delta X_i$$

$f(X)$  称作密度函数。将图 2.1 表示人数的纵坐标换成  $f(X)$  后可以得到图 3.1 (a)。虽然两个图的纵坐标含义各异, 但图的形状却完全相同。在图 3.1 (a) 中, 任一矩形的面积都有其特殊意义, 即矩形的面积恰好等于红细胞数在这一区间内出现的频率 ( $f(X)\Delta X_i = f_i/n$ ), 面积的总和为 1。在例数较大的情况下, 可以近似地看作红细胞数在这一区间出现的概率。现设想将观察人数增多, 组段不断细分, 图中的直条将逐渐变窄, 其分布图就会越来越表现出中间高, 两侧逐渐降低并完全对称的特点, 直方图的顶端将逐渐接近于一条均匀连续的曲线 (图 3-1 中 (b) 和 (c))。这条曲线所描述的分布, 便近似于我们通常所说的正态概率分布, 简称正态分布。

正态分布的密度函数, 即正态曲线的方程由下式给出:

$$f(X) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} \quad -\infty < X < +\infty \quad (3-9)$$

其中  $\pi$  和  $e$  是两个常数, 分别为圆周率和自然对数的底, 其近似值分别为 3.14159 和 2.71828;  $\mu$  和  $\sigma$  是正态分布的两个参数, 恰好是前面学过的最有实用意义的均数  $\bar{X}$  和标准差  $S$ 。需要注意的是,  $\mu$  与  $\bar{X}$ 、 $\sigma$  与  $S$  的意义并不完全相同,  $\mu$  和  $\sigma$  表示总体参数, 而  $\bar{X}$  和  $S$  表示由样本资料算得的统计量。如果  $\mu$  已知, 在计算  $\sigma$  时所用的自由度为  $n$ , 而不是  $n-1$ , 即

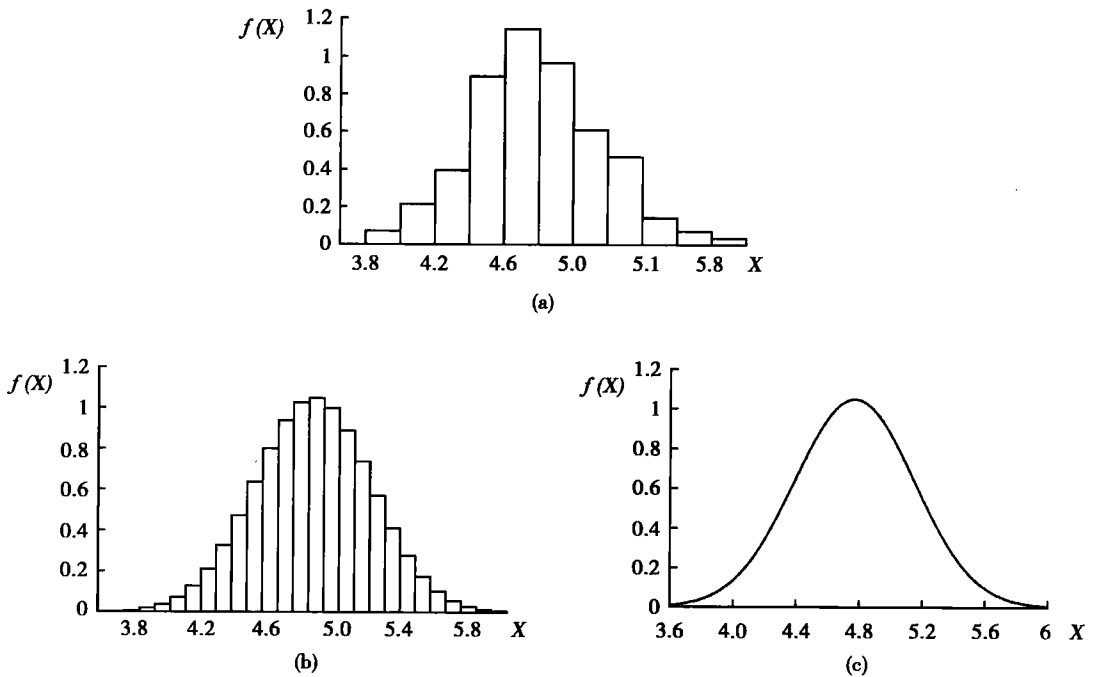


图 3-1 某地成年男子红细胞数的分布逐渐接近正态分布示意图

$$\sigma = \sqrt{\frac{\sum (X - \mu)^2}{n}} \tag{3-10}$$

实际应用中  $\mu$  和  $\sigma$  通常是未知的，但可以将  $\bar{X}$  和  $S$  作为总体参数的估计量来使用。正态分布有以下几个主要特征：

1. 正态分布以均值  $\mu$  为中心，左右对称  $X$  取值范围理论上没有边界 ( $-\infty < X < +\infty$ )， $X$  离  $\mu$  越远，函数  $f(X)$  值越接近于 0，但不会等于 0。
2. 正态分布中，曲线下面积集中在以均值  $\mu$  为中心的部分，越远离中心，曲线越接近  $X$  轴，曲线下面积越小，超过一定范围以外的面积（概率）可以忽略。

3. 正态曲线下的面积分布有一定的规律 即所有的正态分布曲线，在  $\mu$  左右的相同倍数的标准差范围内面积相同；一些特殊情况如在  $\mu \pm \sigma$  范围内的面积约为 68.3%，在  $\mu \pm 1.96\sigma$  范围内约为 95%；在  $\mu \pm 2.58\sigma$  范围内约为 99%，如图 3-2 所示。

4. 正态分布完全由参数  $\mu$  和  $\sigma$  决定  $\mu$  是位置（即平均水平）参数，决定分布曲线在横轴的偏移位置。在  $\sigma$  一定时， $\mu$  增大，曲线沿横轴向右移动；反之  $\mu$  减小，曲线沿横轴向左移动如图 3-3 所示。 $\sigma$  是变异参数，决定分布曲线的形态。 $\sigma$  越大，曲线的形状越“矮胖”，表示数据分布越分散； $\sigma$  越小，曲线的形状越“瘦高”，表示数据分布越集中，如图 3-4 所示。

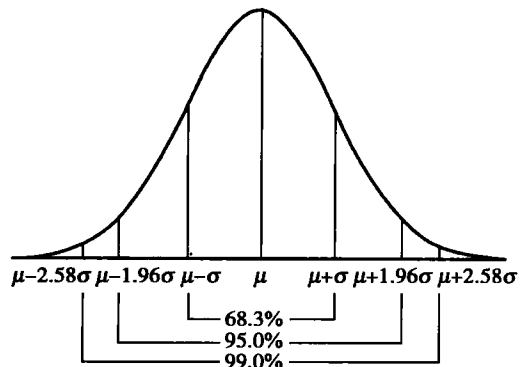


图 3-2 正态分布曲线下的面积

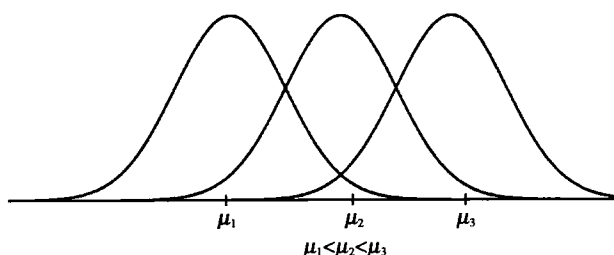


图 3-3 三种不同均值的正态分布

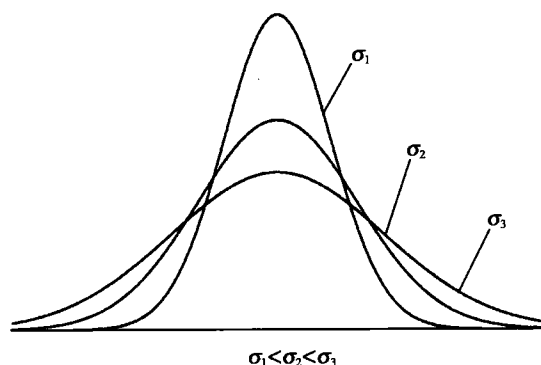


图 3-4 三种不同标准差的正态分布

## 二、标准正态分布

标准正态分布 (standard normal distribution) 是均数为 0、标准差为 1 的正态分布。在式 (3-9) 中令  $\mu=0$  和  $\sigma=1$ , 并用函数  $\varphi(u)$  代替函数  $f(X)$  以区别于一般的正态分布, 就可以得到标准正态分布曲线的函数:

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}, -\infty < u < +\infty \quad (3-11)$$

标准正态分布的曲线如图 3-5 所示。

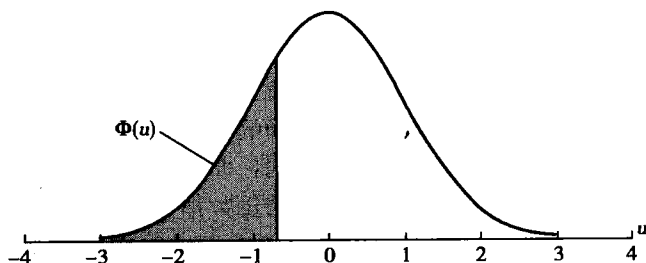


图 3-5 标准正态分布

标准正态分布曲线下, 左侧任一区间的面积可以通过对下式积分求得, 即

$$\Phi(u) = \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \quad (3-12)$$

为了应用方便, 积分结果被制成了一个表 (参见附表 1), 通过查表可以得到  $u$  值左侧的面积。由于分布曲线以 0 为中心, 两边完全对称, 表中只列出了  $u$  值的负数部分。



标准正态分布在实际中应用极为广泛。对任何参数  $\mu$  和  $\sigma$  的正态分布，都可以通过一个简单的变量变换转化成标准正态分布，即

$$u = \frac{X - \mu}{\sigma} \quad (3-13)$$

一旦完成这种变换，就可以利用标准正态分布表求出与原始变量  $X$  有关的概率值。例如成年男子的红细胞数近似服从正态分布，假设均值为  $4.78 \times 10^{12}/L$ ，标准差为  $0.38 \times 10^{12}/L$ ，现想知道红细胞数在  $4 \times 10^{12}/L$  以下所占的比例时，可以先将  $X = 4(\times 10^{12}/L)$  变换为相应的  $u$  值，即

$$u = \frac{4 - 4.78}{0.38} = -2.05$$

于是，问题转化成了求标准正态分布  $u$  值小于  $-2.05$  的概率。现查附表 1 得  $\Phi(-2.05) = 0.0202$ ，表明成年男子的红细胞数低于  $4 \times 10^{12}/L$  的人约占总体的 2%。若要求出在  $4.0 \times 10^{12}/L \sim 5.5 \times 10^{12}/L$  范围内所占的比例，可以通过下面简单计算获得：

$$\begin{aligned} P(4.0 \leq X \leq 5.5) &= \Phi\left(\frac{5.5 - 4.78}{0.38}\right) - \Phi\left(\frac{4 - 4.78}{0.38}\right) \\ &= [1 - \Phi(-1.89)] - \Phi(-2.05) \\ &= (1 - 0.0294) - 0.0202 \\ &= 0.9504 \end{aligned}$$

### 三、正态分布的应用

正态分布在医学领域中应用很广。首先，有不少医学现象服从或近似服从正态分布，如同性别、同年龄儿童的身高和体重，同性别健康成人的红细胞数、血红蛋白含量、脉搏数等。在这类情形下，利用正态分布可以很容易地确定其数值出现在任意指定范围内的概率，尤其是医学参考值范围的估计。实验中的测量误差一般也是服从正态分布的，利用这一点，可以准确地进行误差分析和质量控制。对一些偏态分布的资料，有时也能够通过对数等变换后转换成正态分布，然后按正态分布规律处理。还有很重要一点，正态分布是今后将要学习的各种统计推断方法的理论基础。

## 第三节 医学参考值范围

### 一、医学参考值范围的概念

医学参考值范围 (reference value range) 传统上称作正常值范围 (normal range)，指正常人的解剖、生理、生化、免疫及组织代谢产物的含量等各种数据的波动范围。由于存在生物个体变异，每个正常人的测量值会有所不同，即使是同一个人也会因机体的内外环境变化而改变。因此同属正常人也不能以某一个测量数据作为标准，而必须确定一个波动范围，如一般以  $4.0 \times 10^9/L \sim 10.0 \times 10^9/L$  作为成人白细胞数总数的参考值范围。关于“正常值范围”这一说法，在其意义、推理和观念上都比较模糊，现已很少使用，而改用“参考值范围”。其确切含义为从选择的参照总体上获得的所有检查结果，用统计方法建立百分位数界限时所得到的区间称为参考值范围。习惯上是包含 95% 的参照总体的范围。

使用“参考值范围”的目的主要有两个方面：一是基于临床实践，着眼于个体，作为划分正常人与异常人的界线；二是基于预防医学实践，着眼于人群，如制订不同性别、年龄的儿童某项发育指标的等级标准，用来评价儿童的发育水平等。需要注意的是，所谓正



常不是指机体任何器官、组织的形态和机能都正常的人，而是指符合特定健康水平的人，绝对健康是不存在的，在每个人身上都可能存在着某种程度的病理状态，在使用或制定临床参考值范围时，“正常人”的健康水平应有明确的界定。

## 二、医学参考值范围的制定方法

### (一) 选择足够数量的正常人作为参照样本

选择参照样本必须要考虑可能影响所要制定参考值范围指标的各种疾病及干扰因素，将这些人排除在外。例如在制定血清谷一丙转氨酶活性正常值时，选取正常人的条件为肝、肾、心、脑、肌肉等无器质性疾患，近期无特殊用药史等。参照样本含量的确定没有统一的规则，它与总体分布有关，如果接近正态分布，变异度又不是很大，需要的样本含量就可以少一些；反之，如果明显偏态或数据比较分散，需要的样本含量就相应地多一些。一般认为至少应在 120 例以上，如果例数过少，确定的参考值范围往往不够准确。

有时对于同一测量指标，需要考虑性别、年龄、民族、地理位置等因素并加以区分，对不同人群分组制定参考值范围，例如红细胞数及血红蛋白含量，男女各异，高原居民与平原不同，在这种情况下，应保证每组有足够的例数。

### (二) 对选定的参照样本进行准确地测定

为保证原始数据可靠，要严格控制检测误差，包括分析仪器的灵敏度、试剂的纯度、操作技术及标准的掌握等，同时必须对测量条件作出统一的规定和说明，如临床化验参考值范围的制定，应对收集样本时的环境和生理条件（温度、季节、体育活动强度、饮食、妊娠等），收集、转运和储藏样品的方法及时间有明确的规定。

### (三) 决定取单侧范围还是双侧范围值

有些指标如白细胞数过高或过低均属异常，故其参考值范围需要分别确定下限和上限，称作双侧。有些指标如 24 小时尿糖含量仅在过高、肺活量仅在过低时为异常，只需确定其上限或下限，称作单侧。

### (四) 选择适当的百分范围

参考值范围是指绝大多数正常人（参照总体）的测量值所在的范围。所谓的“绝大多数”究竟是多少，取决于资料的性质和研究目的所规定的百分界限。最常用的百分界限是 95%，也可以选 80%、90%、99% 等。

图 3-6 是正常人和病人的数据分布有重叠的示意图。以制定单侧上限值为例，从图 3-6 可以看出：提高上限值，假阳性可以减少，但假阴性必然增加；反之，若降低上限值，假阴性可以减少，但假阳性必然增加。

实际中最好结合正常人和病人的数据分布特点，权衡假阳性和假阴性的比例，选择一个适当的百分范围。若主要目的在于减少假阳性（如确诊病人），参考值范围的百分数范围要取大一些（如 95% 或 99%）；反之，若主要目的在于减少假阴性（如初筛病人），百分数范围可以小一些（如 90% 或 80%）。

### (五) 估计参考值范围的界限

参考值范围估计的方法有多种，其中最基本的有百分位数法和正态分布法。表 3-2 和表 3-3 分别给出了两种不同方法的 95% 和 99% 百分范围的参考值界限。

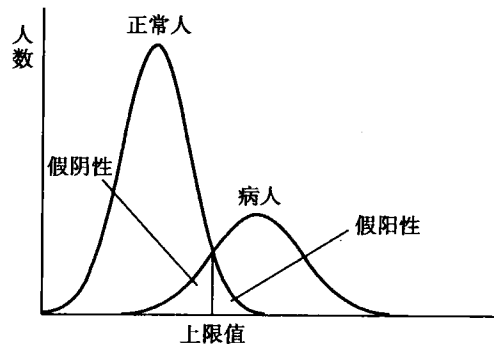


图 3-6 正常人和病人数据分布重叠



表 3-2 参考值范围所对应的百分位数

百分范围(%)	单 侧		双侧(对称)	
	下限	上限	下限	上限
95	$P_5$	$P_{95}$	$P_{2.5}$	$P_{97.5}$
99	$P_1$	$P_{99}$	$P_{0.5}$	$P_{99.5}$

表 3-3 参考值范围所对应的正态分布区间

百分范围(%)	单 侧		双侧(对称)	
	下限	上限	下限	上限
95	$\bar{X}-1.65S$	$\bar{X}+1.65S$	$\bar{X}-1.96S$	$\bar{X}+1.96S$
99	$\bar{X}-2.33S$	$\bar{X}+2.33S$	$\bar{X}-2.58S$	$\bar{X}+2.58S$

**例 3.4** 若已算得某地正常成年男子红细胞数的均数为  $4.78 \times 10^{12}/L$ ，标准差为  $0.38 \times 10^{12}/L$ ，试估计该地成年男子红细胞数的 95% 参考值范围。

正常成年男子的红细胞数近似服从正态分布，故可按正态分布法处理。又因红细胞数过多或过少均属异常，所以应计算双侧参考值范围。

下限： $\bar{X}-1.96S=4.78-1.96 \times 0.38=4.04(\times 10^{12}/L)$

上限： $\bar{X}+1.96S=4.78+1.96 \times 0.38=5.52(\times 10^{12}/L)$

即该地区成年男子红细胞数的 95% 参考值范围为  $4.04 \times 10^{12}/L \sim 5.52 \times 10^{12}/L$ 。

**例 3.5** 见第二章表 2-4 资料。为该地区 50~60 岁女性高血脂诊断与治疗提供参考依据，试估计其血清甘油三酯含量的 95% 单侧参考值范围。

由表 2-4 可见，甘油三酯呈正偏态分布，可考虑用百分位数方法处理。

$$P_{95}=1.90+\frac{(630 \times 0.95-580)}{28} \times 0.30=2.098(\text{mmol}/L)$$

即其 95% 单侧参考值范围为小于 2.098mmol/L。

两种方法比较，百分位数法适合于任何分布类型的资料，故在实际中最为常用。但由于参考值范围所涉及的常常是波动较大的两端数据，使用百分位数法必须有较大的样本含量，否则结果不稳定。正态分布法的应用条件是资料必须近似服从正态分布，优点是结果较稳定，缺点是医学上很多资料不呈正态分布，适用范围较窄。医学中有一部分偏态资料，经对数转换后能够近似地转换为正态分布，对这类资料也可以用正态分布法处理，具体作法请自己思考。关于资料是否符合正态分布的统计检验问题可参考其他有关书籍。

## 练 习 题

### 一、单项选择题

- 变异系数主要用于
  - 比较不同计量指标的变异程度
  - 衡量正态分布的变异程度
  - 衡量测量的准确度
  - 衡量偏态分布的变异程度
  - 衡量样本抽样误差的大小
- 对于近似正态分布的资料，描述其变异程度应选用的指标是
  - 变异系数
  - 离均差平方和
  - 极差
  - 四分位数间距
  - 标准差
- 某项指标 95% 医学参考值范围表示的是
  - 检测指标在此范围，判断“异常”正确的概率大于或等于 95%



- B. 检测指标在此范围, 判断“正常”正确的概率大于或等于 95%
  - C. 在“异常”总体中有 95%的人在此范围之外
  - D. 在“正常”总体中有 95%的人在此范围
  - E. 检测指标若超出此范围, 则有 95%的把握说明诊断对象为“异常”
4. 应用百分位数法估计参考值范围的条件是
- A. 数据服从正态分布
  - B. 数据服从偏态分布
  - C. 有大样本数据
  - D. 数据服从对称分布
  - E. 数据变异不能太大
5. 已知动脉硬化患者载脂蛋白 B 的含量 (mg/dl) 呈明显偏态分布, 描述其个体差异的统计指标应使用
- A. 全距
  - B. 标准差
  - C. 变异系数
  - D. 方差
  - E. 四分位数间距

**二、计算与分析**

1. 下表为 10 例垂体催乳素微腺瘤的病人手术前后的血催乳素浓度, 试说明用何种指标比较手术前后数据的变异情况较为合适。

手术前后患者血催乳素浓度 (ng/ml)

例号	血催乳素浓度		例号	血催乳素浓度	
	术前	术后		术前	术后
1	276	41	6	266	43
2	880	110	7	500	25
3	1 600	280	8	1 760	300
4	324	61	9	500	215
5	398	105	10	220	92

2. 某地 144 例 30~45 岁正常成年男子的血清总胆固醇测量值近似服从均数为 4.95mmol/L, 标准差为 0.85mmol/L 的正态分布。①试估计该地 30~45 岁成年男子血清总胆固醇的 95%参考值范围; ②血清总胆固醇大于 5.72mmol/L 的正常成年男子约占其总体的百分之多少?

3. 某地 200 例正常成人血铅含量的频数分布如下表。

- (1) 简述该资料的分布特征。
- (2) 若资料近似呈对数正态分布, 试分别用百分位数法和正态分布法估计该地正常成人血铅值的 95%参考值范围。

某地 200 例正常成人血铅含量 ( $\mu\text{mol/L}$ ) 的频数分布

血铅含量	频数	累积频数	血铅含量	频数	累积频数
0~	7	7	1.44~	14	188
0.24~	49	56	1.68~	4	192
0.48~	45	101	1.92~	4	196
0.72~	32	133	2.16~	1	197
0.96~	28	161	2.40~	2	199
1.20~	13	174	2.64~	1	200

(李 康)



### Measures of Dispersion and Variability: Sum of Square, Mean of Square and Standard Deviation

To develop a measure that uses all the data to form an index of dispersion consider the following. Suppose we express each observation as a distance from the mean  $X_i = (X_i - \bar{X})$ . These differences are called deviates and will be sometimes positive ( $X_i$  is above the mean) and sometimes negative ( $X_i$  is below the mean). The average of deviates always sum to zero. The widely used measure of dispersion called the sample variance utilizes the squares of the deviates. The quantity  $SS = \sum (X_i - \bar{X})^2$  is the sum of these squared deviates and referred to as the sum of squares. The corrected sum of squares is utilized in the formula for the sample variance.

$$S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$$

and is referred to as the mean of squares. A more "natural" calculation is the standard deviation, which is positive square root of  $S^2$ .

$$S = \sqrt{\frac{\sum X_i^2 - (\sum X_i)^2/n}{n-1}}$$



## 第四章 抽样误差与假设检验

本章开始学习统计推断 (statistical inference) 方法, 内容涉及参数估计和假设检验的基本原理。参数估计是指通过样本数据对总体的特征进行推测, 给出的结果通常是一个区间; 假设检验则是从另一角度去分析数据, 重点放在总体参数大小的比较上。

### 第一节 均数的抽样误差与标准误

在医学研究中, 绝大多数情况是由样本信息推断总体特征。由于个体存在差异, 因此通过样本推论总体时会存在一定的误差, 如样本均数  $\bar{X}$  往往不等于总体均数  $\mu$ , 这种由抽样造成的样本均数与总体均数的差异称为抽样误差 (sampling error)。例如, 假设健康成年男子的红细胞服从均数为  $\mu=4.75 \times 10^{12}/L$ , 标准差  $\sigma=0.38 \times 10^{12}/L$  的正态分布, 现随机抽取 140 人, 计算红细胞的样本均数为  $\bar{X}=4.77 \times 10^{12}/L$ , 两值之差  $0.02 \times 10^{12}/L$  即为抽样误差。可以设想, 若再随机抽取 140 名成年男子进行测量, 其平均红细胞数会是另一不同的结果。对于抽样研究, 抽样误差是不可避免的。

如何能精确地估计成年男子红细胞数的总体均数? 因为存在抽样误差, 基于 140 例研究对象的样本均数不可能与总体均数完全一样, 问题是我们能否知道估计值的精确程度, 对此可以利用抽样误差的分布规律对其进行分析。理论上可以证明: 若从正态总体  $N(\mu, \sigma^2)$  中, 反复多次随机抽取样本含量固定为  $n$  的样本, 那么这些样本均数  $\bar{X}$  也服从正态分布, 即  $\bar{X}$  的总体均数仍为  $\mu$ , 样本均数的标准差为  $\sigma/\sqrt{n}$ 。中心极限定理则告诉我们: 在样本含量  $n$  很大的情况下 (如  $n \geq 50$ ), 无论原始测量变量服从什么分布,  $\bar{X}$  的抽样分布都近似服从正态分布  $N(\mu, \sigma^2/n)$ 。

统计学中为了区别个体观察值之间变异的标准差与反映样本均数之间变异的标准差, 将后者称为标准误 (standard error)。显然, 标准误小于原始测量值的标准差, 标准误越小说明估计越精确, 因此可以用标准误表示抽样误差的大小。标准误用符号  $\sigma_{\bar{x}}$  表示, 计算公式为

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad (4-1)$$

由式 (4-1) 可见, 在样本含量一定的情况下, 标准误与标准差成正比, 说明当总体中各观测值变异较小时, 抽到的样本均数  $\bar{X}$  与总体均数  $\mu$  可能相差较小, 用  $\bar{X}$  估计  $\mu$  的可靠程度高; 当总体中各观测值变异较大时, 抽到的样本均数  $\bar{X}$  与总体均数  $\mu$  可能相差很大, 用  $\bar{X}$  估计  $\mu$  的可靠程度也相对较低。标准误与样本含量的平方根  $\sqrt{n}$  成反比, 说明在同一总体中随机抽样, 样本含量  $n$  越大, 标准误越小。标准误反映了样本均数间的离散程度, 也反映了样本均数与总体均数的差异。

实际工作中, 总体标准差  $\sigma$  往往未知, 因而通常用样本标准差  $S$  代替  $\sigma$ , 求得样本均数标准误的估计值  $S_{\bar{x}}$ , 计算公式为

$$S_{\bar{x}} = \frac{S}{\sqrt{n}} \quad (4-2)$$

**例 4.1** 在某地随机抽查成年男子 140 人, 得红细胞均数  $4.77 \times 10^{12}/L$ , 标准差  $0.38 \times 10^{12}/L$ , 试计算其标准误。



按公式 (4-2) 计算得:

$$S_{\bar{X}} = \frac{S}{\sqrt{n}} = \frac{0.38}{\sqrt{140}} = 0.032 (\times 10^{12}/L)$$

标准误是抽样分布的重要特征之一, 它可以用于参数的区间估计和对不同组之间的参数进行比较。

## 第二节 总体均数的估计

### 一、可信区间的概念

参数估计 (parameter estimation) 指由样本统计量估计总体参数, 是统计推断的重要内容之一。常用的估计方式有两种: 点估计 (point estimation) 和区间估计 (interval estimation)。

点估计是使用单一的数值直接作为总体参数的估计值, 如用  $\bar{X}$  估计相应的  $\mu$ 。该法表达简单, 但未考虑抽样误差的影响, 无法评价参数估计的准确程度。

区间估计是指按预先给定的概率, 计算出一个区间, 使它包含未知的总体均数。事先给定的概率  $1-\alpha$  称为可信度 (通常取 0.95 或 0.99), 计算得到的区间称为可信区间 (confidence interval, CI)。可信区间通常由两个数值界定的可信限 (confidence limit) 构成, 其中数值较小的一方称为下限, 数值较大的一方称为上限。总体均数估计的 95% 可信区间表示该区间包括总体均数  $\mu$  的概率为 95%, 即若作 100 次抽样算得 100 个可信区间, 则平均有 95 个可信区间包括  $\mu$  (估计正确), 只有 5 个可信区间不包括  $\mu$  (估计错误)。例如, 估计某地成年男子的红细胞平均水平, 其 95% 的可信区间为  $4.71 \times 10^{12}/L \sim 4.83 \times 10^{12}/L$ , 说明总体红细胞均数没有被包含在这一区间的概率只有 5%, 或有 95% 的把握认为总体红细胞均数在此区间内。

可信区间估计的效果, 一是由可信度  $1-\alpha$  来反映, 即计算出的区间包括总体均数  $\mu$  的概率大小, 其值越接近 1 越好; 二是由区间的宽度来反映, 区间越窄说明估计越精确。在样本含量一定的情况下, 二者是矛盾的, 若仅考虑提高可信度, 则使估计的区间变宽, 从而降低了使用可信区间的价值。一般情况下, 95% 可信区间更为常用; 在可信度确定的情况下, 增加样本量可以缩小区间宽度, 从后面区间估计的计算公式可以更清楚地看到这一点。

### 二、总体均数可信区间的计算

总体均数  $\mu$  可信区间的计算公式可以利用  $\bar{X}$  的抽样分布获得。实际中, 总体均数可信区间的计算方法, 根据总体标准差  $\sigma$  是否已知, 以及样本量  $n$  的大小而有所不同。

#### (一) $\sigma$ 已知

如果变量  $X$  服从均数为  $\mu$ 、标准差为  $\sigma$  的正态分布, 则

$$u = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad (4-3)$$

服从标准正态分布。按照标准正态分布规律, 95% 的  $u$  值在  $-1.96$  和  $1.96$  之间, 即

$$P\left(-1.96 \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq 1.96\right) = 0.95 \quad (4-4)$$

$$P\left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}\right) = 0.95$$



从而得到 95% 可信区间:

$$(\bar{X} - 1.96\sigma_{\bar{X}}, \bar{X} + 1.96\sigma_{\bar{X}}) \quad (4-5)$$

更一般的情况:

$$(\bar{X} - u_{\alpha/2}\sigma_{\bar{X}}, \bar{X} + u_{\alpha/2}\sigma_{\bar{X}}) \quad (4-6)$$

其中  $u_{\alpha/2}$  为标准正态分布的双侧界值。若取  $1-\alpha=0.95$ , 则为总体均数的 95% 可信区间, 或取  $1-\alpha=0.99$ , 则为总体均数 99% 可信区间。需要注意的是,  $\mu$  不是一个随机变量, 而是包含在可信区间内的一个参数值。

## (二) $\sigma$ 未知

事实上, 总体标准差  $\sigma$  通常是未知的, 这时我们可以用其估计量  $S$  代替  $\sigma$ , 但在这种情况下,  $(\bar{X} - \mu)/(S/\sqrt{n})$  已不再服从标准正态分布, 而是服从著名的  $t$  分布。

1.  $t$  分布  $t$  分布 ( $t$ -distribution) 由英国统计学家 W. S. Gosset 于 1908 年以 “Student” 为笔名提出, 他证明了在正态分布总体中进行抽样,  $(\bar{X} - \mu)/(S/\sqrt{n})$  服从自由度为  $\nu=n-1$  的  $t$  分布。 $t$  分布是以 0 为中心的对称分布, 参见图 4-1。

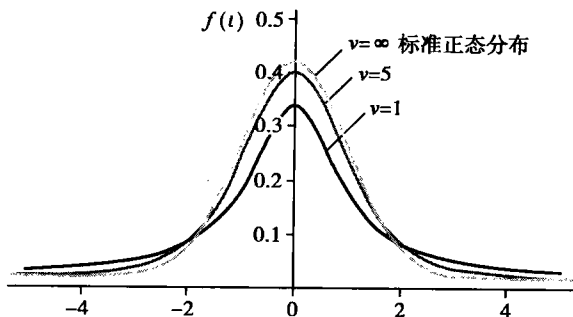


图 4-1 不同自由度的  $t$  分布图

由图 4-1 可见,  $t$  分布曲线的形态变化与自由度  $\nu=n-1$  有关。随着自由度  $\nu$  的增大,  $t$  分布曲线越来越接近于标准正态分布曲线; 当  $\nu \rightarrow \infty$

时,  $t$  分布的极限分布就是标准正态分布。因此,  $t$  分布曲线下面积 95% 的界值不是一个常量, 它随自由度大小不同而变化。为了应用方便, 可根据附表 2 查找相应的  $t$  界值。 $t$  界值表中, 横标目为自由度  $\nu$ , 纵标目为概率, 一侧尾部面积为单侧概率, 两侧尾部面积之和为双侧概率; 表中的值为相应  $t$  的界值。如当  $\nu=24$ 、双侧概率  $\alpha=0.05$  时, 由表中查得  $t_{0.05/2, 24}=2.064$ , 此处 2.064 即为两侧尾部概率各为 0.025 的  $t$  界限值。由于  $t$  分布的对称性, 表中只列出了正值。

2. 可信区间的计算 计算可信区间的原理与  $\sigma$  已知情况完全相同, 仅仅是两侧概率的界值有些差别。按  $t$  分布规律,  $100\% \times (1-\alpha)$  的  $t$  值在  $-t_{\alpha/2, \nu}$  和  $t_{\alpha/2, \nu}$  之间, 即

$$P(-t_{\alpha/2, \nu} < \frac{\bar{X} - \mu}{S/\sqrt{n}} < t_{\alpha/2, \nu}) = 1 - \alpha \quad (4-7)$$

将上式写成可信区间:

$$(\bar{X} - t_{\alpha/2, \nu} S_{\bar{X}}, \bar{X} + t_{\alpha/2, \nu} S_{\bar{X}}) \quad (4-8)$$

$t_{\alpha/2, \nu}$  界值可以由附表 2 查得。需要注意: 在小样本情况下, 应用这一公式的条件是原始变量  $X$  服从正态分布。

在大样本情况下 (如  $n > 50$ ), 无论变量  $X$  是否服从正态分布, 按照中心极限定理  $\bar{X}$  都服从正态分布, 同时  $t$  分布逼近标准正态分布, 可信区间可以用下式近似计算:

$$(\bar{X} - u_{\alpha/2} \cdot S_{\bar{X}}, \bar{X} + u_{\alpha/2} \cdot S_{\bar{X}}) \quad (4-9)$$

**例 4.2** 某医生测得 25 名动脉粥样硬化患者血浆纤维蛋白原含量的均数为 3.32g/L, 标准差为 0.57g/L, 试计算该种病人血浆纤维蛋白原含量总体均数的 95% 可信区间。

本例  $n=25$   $\bar{X}=3.32$ ,  $S=0.57$ ,  $\nu=n-1=25-1=24$ ,  $\alpha=0.05$ , 查  $t$  值表  $t_{0.05/2(24)}=2.064$ , 按公式 (4-8) 计算得:

$$\text{下限: } \bar{X} - t_{\alpha/2, \nu} S_{\bar{X}} = 3.32 - 2.064 \times 0.57 / \sqrt{25} = 3.08$$



上限:  $\bar{X} + t_{\alpha/2, \nu} S_{\bar{X}} = 3.32 + 2.064 \times 0.57 / \sqrt{25} = 3.56$

该种病人血浆纤维蛋白原含量总体均数的 95% 可信区间为 3.09g/L~3.56g/L。

**例 4.3** 试计算例 4.1 中该地成年男子红细胞总体均数的 95% 可信区间。

本例属于大样本, 可采用正态近似的方法计算可信区间 ( $\alpha=0.05$ )。因为  $\bar{X}=4.77$ ,  $S=0.38$ ,  $n=140$ , 则 95% 可信区间为

下限:  $\bar{X} - u_{\alpha/2} S_{\bar{X}} = 4.77 - 1.96 \times 0.38 / \sqrt{140} = 4.71$

上限:  $\bar{X} + u_{\alpha/2} S_{\bar{X}} = 4.77 + 1.96 \times 0.38 / \sqrt{140} = 4.83$

估计该地成年男子红细胞总体均数的 95% 可信区间为  $4.71 \times 10^{12}/L \sim 4.83 \times 10^{12}/L$ 。

3. 模拟实验 为了进一步理解可信区间的含义, 现假定  $\mu=4.75$ ,  $\sigma=0.39$ , 用计算机模拟健康成年男子红细胞数作抽样实验, 重复产生样本含量  $n=140$  的 100 个随机样本, 分别计算其 95% 可信区间。从图 4-2 可以看到, 绝大多数可信区间包含总体参数  $\mu=4.75$ , 有 6 个可信区间没有包含总体参数 (用星号标记)。

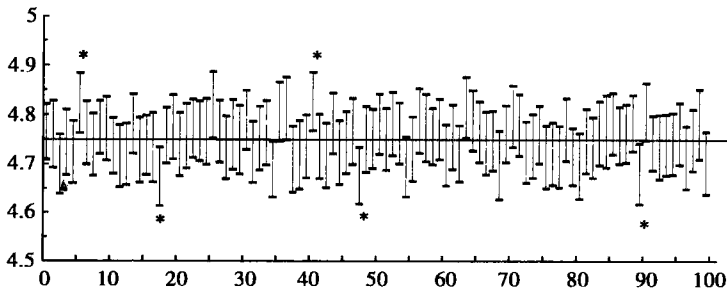


图 4-2 模拟抽样成年男子红细胞数的 95% 可信区间示意图

( $\mu=4.75$ ,  $\sigma=0.39$ ,  $n=140$ )

### (三) 单侧可信区间

前面涉及的都是双侧可信区间。但有些情况下, 我们所关心的仅仅是单侧的可信限。例如, 对于轻度原发性高血压患者进行治疗, 一般病人接受治疗后可以使舒张压平均降低 10mmHg, 现提出一种新的治疗方法, 我们仅对此疗法是否优于标准法感兴趣, 为此有 100 名患者接受了这种新疗法, 得到的数据显示平均降压 12.7mmHg, 血压治疗前后变化的标准差为 5.6mmHg, 能否说新疗法优于标准疗法 ( $\alpha=0.05$ )? 为此可以计算单侧可信区间, 其中只需关心可信区间的下限, 即如果下限值高于 10mmHg, 我们就有理由下结论说新疗法优于标准疗法。

单侧可信区间与双侧可信区间的计算公式基本相同, 只需将公式 (4-8) 和 (4-9) 中的抽样分布的双侧界值换成单侧界值, 同时只取下限或上限。本例:

95% 可信区间下限:  $\bar{X} - t_{0.05, 99} S_{\bar{X}} = 12.7 - 1.66 \times 5.6 / \sqrt{100} = 11.8$

$\mu \geq 11.8$ , 即有 95% 的把握推断新药平均降压至少为 11.8mmHg, 优于标准疗法。

## 第三节 假设检验的意义和步骤

### 一、假设检验的原理与思想

假设检验 (hypothesis test) 亦称显著性检验 (significant test), 是统计推断的另一重要内容, 其目的是比较总体参数之间有无差别。实际中多数情况是用样本数据去推断总



体, 由于存在生物个体变异和随机测量误差, 不能简单地根据样本统计量数值的大小直接比较总体参数。如在比较甲乙两种治疗高脂血症药物疗效的试验中, 甲乙两组各为 100 名患者, 甲药使血清甘油三酯平均下降 1.36mmol/L、乙药使血清甘油三酯平均下降 1.12mmol/L, 并不能说明甲药优于乙药, 因为如果再重新做一次试验其结果可能相反。假设检验的基本思想是首先对所需要比较的总体提出一个无差别的假设, 然后通过样本数据去推断是否拒绝这一假设。其意义可由下面的例子说明。

**例 4.4** 使用黑加仑油软胶囊治疗高脂血症, 30 名患者治疗前后血清甘油三酯检测结果的差值为  $1.38 \pm 0.76$  (mmol/L), 问治疗后血清甘油三酯是否有所改善?

对上面问题可以作如下考虑: 如果治疗前后血清甘油三酯没有改善, 则前后变化差值的总体均数应为  $\mu_d = 0$ , 但现在计算出的样本差值均数  $\bar{d} = 1.38$ , 其差别可能由两种原因引起, 即药物的作用和生物变异性波动。现用假设  $H_0: \mu_d = 0$  表示治疗前后没有差别, 需要判断的是这个假设是否符合实际检测结果。我们知道, 即使治疗前后血清甘油三酯没有改善, 由于生物的变异性,  $\bar{d}$  不会恰好等于 0; 但当  $\bar{d} - 0$  较大时, 血清甘油三酯的改变就难以完全用生物变异来解释, 从而认为药物具有作用。问题是  $\bar{d} - 0$  究竟要大到多少才能够下这样的结论?

由前一节知道, 如果治疗前后血清甘油三酯检测结果的差值服从正态分布, 在  $H_0: \mu_d = 0$  成立的条件下,  $t = (\bar{d} - 0) / (S_d / \sqrt{n})$  服从  $t$  分布 ( $S_d$  为差值的标准差)。因此, 可以利用现有样本计算  $t$  统计量, 如果计算出的  $t$  值超出了给定的界限 (说明在  $H_0$  为真时获得现有样本的概率很小), 则倾向于拒绝假设  $H_0$ , 认为治疗前后有差别 (图 4-3)。

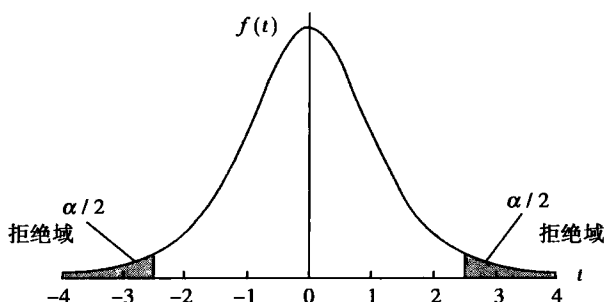


图 4-3 利用  $t$  分布进行假设检验原理示意图

假设检验的实质是判断观察到的“差别”是由抽样误差引起还是总体上的不同, 目的是评价两种不同处理引起效应不同的证据有多强, 这种证据的强度用概率  $P$  来度量和表示。除  $t$  分布外, 针对不同的资料还有其他各种检验统计量及分布, 如  $F$  分布、 $\chi^2$  分布等, 应用这些分布对不同类型的数据进行假设检验的步骤相同, 其差别仅仅是需要计算的检验统计量不同。

## 二、假设检验的基本步骤

1. 建立假设和确定检验水准 假设检验中, 包括无效假设 (null hypothesis) 和备择假设 (alternative hypothesis) 两种假设。

无效假设符号为  $H_0$ , 指需要检验的假设, 如治疗前后血清甘油三酯没有差别, 即

$$H_0: \mu_d = 0$$

这一假设通常与我们要验证的结论相反, 是计算检验统计量和  $P$  值的依据。

备择假设符号为  $H_1$ , 是在  $H_0$  成立证据不足的情况下而被接受的假设, 如拒绝治疗前后血清甘油三酯相同的假设, 可表示为

$$H_1: \mu_d \neq 0$$

备择假设有双侧和单侧两种情况。双侧检验指无论是正方向还是负方向的误差, 若显著地超出检验水准则拒绝  $H_0$ ,  $H_1: \mu_d \neq 0$  即为双侧检验。单侧检验指仅在正方向或负方



向误差超出规定的水准时则拒绝  $H_0$ ，如治疗后血清甘油三酯下降的假设可表示为

$$H_1: \mu_d > 0 \text{ (或 } H_1: \mu_d < 0\text{)}$$

双侧检验和单侧检验应如何选择，需根据研究目的和专业知识的而定。例如，比较两种降血脂药物的疗效，因无法判断两种药物的优劣，应选用双侧检验；如果是检验一种药物的降血脂作用，因为药物一般总是能降低血脂而不是升高血脂，这时可以采用单侧检验。由于双侧检验将拒绝域的概率等分在  $t$  分布两侧的尾部，因此单侧检验的  $t$  界值（绝对值）总是小于双侧检验所用的界值。对同一样本，双侧检验得出有显著性差别的结论，单侧检验也一定是显著的，因而在实际中使用较多的是双侧检验。本书中在未作说明的情况下，选用的均是双侧检验。

建立检验假设的同时，还必须给出检验水准。检验水准亦称显著性水准（significant level），用  $\alpha$  表示，是预先规定的拒绝域的概率值，实际中一般取  $\alpha=0.05$  或  $\alpha=0.01$ 。显然， $\alpha$  值越大越容易得出有差别的结论。

2. 选择检验方法和计算检验统计量 根据资料类型、研究设计方案和统计推断的目的，选择适当的检验方法和计算公式。许多假设检验方法是以检验统计量来命名的，如  $t$  检验、 $u$  检验、 $F$  检验和  $\chi^2$  检验等，这些都是后面章节要学习的重要内容。

3. 确定  $P$  值和作出统计推断结论 查表得到检验用的临界值，然后将算得的统计量与拒绝域的临界值作比较，确定  $P$  值。如对双侧  $t$  检验  $|t| \geq t_{\alpha/2, \nu}$ ，则  $P \leq \alpha$ ，按  $\alpha$  检验水准拒绝  $H_0$ ，接受  $H_1$ ；若  $P > \alpha$ ，则不能拒绝  $H_0$ 。 $P$  值是假设检验下结论的主要依据，其含义是指在原假设成立的条件下，观察到的样本差别是由于机遇所致的概率。因此， $P$  值越小越有理由拒绝无效假设，认为总体之间有差别的统计学证据越充分。需要注意：不拒绝  $H_0$  不等于支持  $H_0$  成立，仅表示现有样本信息不足以拒绝  $H_0$ 。

传统上，通常将  $P > 0.05$  称为“不显著”， $0.01 < P \leq 0.05$  称为“显著”， $P \leq 0.01$  称为“非常显著”。目前有学者主张应尽量使用有无统计学意义进行表述。由于统计软件的应用和普及，最好列出精确的  $P$  值，以便对检验结果的可靠性作出准确地判断。

## 练 习 题

### 一、单项选择题

- 样本均数的标准误越小说明
  - 观察个体的变异越小
  - 观察个体的变异越大
  - 抽样误差越大
  - 由样本均数估计总体均数的可靠性越小
  - 由样本均数估计总体均数的可靠性越大
- 抽样误差产生的原因是
  - 样本不是随机抽取
  - 测量不准确
  - 资料不是正态分布
  - 个体差异
  - 统计指标选择不当
- 对于正偏态分布的总体，当样本含量足够大时，样本均数的分布近似为
  - 正偏态分布
  - 负偏态分布
  - 正态分布
  - $t$  分布
  - 标准正态分布
- 假设检验的目的是
  - 检验参数估计的准确度
  - 检验样本统计量是否不同



C. 检验样本统计量与总体参数是否不同

D. 检验总体参数是否不同

E. 检验样本的  $P$  值是否为小概率

5. 根据样本资料算得健康成人白细胞计数的 95% 可信区间为  $7.2 \times 10^9/L \sim 9.1 \times 10^9/L$ , 其含义是

A. 估计总体中有 95% 的观察值在此范围内

B. 总体均数在该区间的概率为 95%

C. 样本中有 95% 的观察值在此范围内

D. 该区间包含样本均数的可能性为 95%

E. 该区间包含总体均数的可能性为 95%

## 二、计算与分析

1. 为了解某地区小学生血红蛋白含量的平均水平, 现随机抽取该地小学生 450 人, 算得其血红蛋白平均数为  $101.4 \text{g/L}$ , 标准差为  $1.5 \text{g/L}$ , 试计算该地小学生血红蛋白平均数的 95% 可信区间。

2. 研究高胆固醇是否有家庭聚集性, 已知正常儿童的总胆固醇平均水平是  $175 \text{mg/dl}$ , 现测得 100 名曾患心脏病且胆固醇高的子代儿童的胆固醇平均水平为  $207.5 \text{mg/dl}$ , 标准差为  $30 \text{mg/dl}$ 。试回答下列问题:

① 如何衡量这 100 名儿童总胆固醇样本平均数的抽样误差?

② 估计 100 名儿童的胆固醇平均水平的 95% 可信区间。

③ 根据可信区间判断高胆固醇是否有家庭聚集性, 并说明理由。

(李 康)

## Sampling Error and Hypothesis Testing

Perhaps the most important sampling distribution in applied statistics is the distribution of the sample mean. To construct a sampling distribution for the sample mean, We do the following:

1. From a finite population of size  $N$ , randomly draw all possible samples of size  $n$ .
2. For each sample compute  $\bar{X}$ .
3. List all the distinct observed values of  $\bar{X}$  and the frequency of each of these distinct values.

Construction of a sampling distribution for the sample mean is a very arduous task for samples of any appreciable size.

The variance of the sampling distribution of mean is not equal to the population variance. However, the variance of the sampling distribution of the mean is exactly the population variance divided by the sample size. In general, for a sample of size  $n$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

The  $\sigma_{\bar{X}}$  is called the standard error of the mean.

Confidence Interval: A  $(1 - \alpha)$  100% confidence interval for a unknown population mean  $\sigma$  and small sample size

$$\bar{X} \pm t_{\alpha/2, v} \frac{S}{\sqrt{n}}$$



where  $t_{\alpha/2, v}$  is the  $t$ -value under the *Student's t* distribution with  $n-1$  degree of freedom, and  $n$  symbolizes sample size,  $S$  symbolizes the standard deviation of the sample.

A statistical test of hypothesis consists of five parts: the null hypothesis, denoted by  $H_0$ ; the alternative hypothesis, denoted by  $H_1$ ; the test statistic and its  $P$ -value; the rejection region; the conclusion. When you specify these five parts, you define a particular test.



## 第五章 $t$ 检验

本章介绍计量资料两组均数比较的最常用的假设检验方法,即  $t$  检验 ( $t$  test) 和  $u$  检验 ( $u$  test)。根据研究设计和资料的性质有单个样本的  $t$  检验,配对样本  $t$  检验,两个独立样本  $t$  检验以及在方差不齐时的  $t'$  检验。实际应用时应把握各种检验方法的用途、适用条件和注意事项。

理论上, $t$  检验的应用条件要求样本来自正态分布总体,两样本均数比较时,还要求两总体方差相同,即具有方差齐性。但在实际工作中,与上述条件略有偏离,只要其分布为单峰且近似对称分布,也可应用。 $t$  检验的原理依据的是上一章介绍的  $t$  分布。当样本含量较大时,可用  $u$  检验。

### 第一节 单样本 $t$ 检验

单样本  $t$  检验又称单样本均数  $t$  检验 (one sample  $t$  test),适用于样本均数  $\bar{X}$  与已知总体均数  $\mu_0$  的比较,其比较目的是检验样本均数  $\bar{X}$  所代表的总体均数  $\mu$  是否与已知总体均数  $\mu_0$  有差别。已知总体均数  $\mu_0$ ,一般为标准值、理论值或经大量观察得到的较稳定的指标值。

单样本  $t$  检验用于总体标准差  $\sigma$  未知的资料,其统计量  $t$  值按公式 (5-1) 计算:

$$t = \frac{\bar{X} - \mu_0}{S_{\bar{X}}} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}, \nu = n - 1 \quad (5-1)$$

其中  $S$  为样本标准差,  $n$  为样本含量。

**例 5.1** 以往通过大规模调查已知某地新生儿出生体重为 3.30kg。从该地难产儿中随机抽取 35 名新生儿作为研究样本,平均出生体重为 3.42kg,标准差为 0.40kg,问该地难产儿出生体重是否与一般新生儿体重不同?

本例已知总体均数  $\mu_0 = 3.30\text{kg}$ ,但总体标准差  $\sigma$  未知,  $n = 35$ ,为小样本,  $\bar{X} = 3.42\text{kg}$ ,  $S = 0.40\text{kg}$ ,故选用单样本  $t$  检验。

(1) 建立检验假设,确定检验水准

$H_0: \mu = \mu_0$ , 该地难产儿与一般新生儿平均出生体重相同

$H_1: \mu \neq \mu_0$ , 该地难产儿与一般新生儿平均出生体重不同

$\alpha = 0.05$

(2) 计算检验统计量

在  $\mu = \mu_0$  成立的前提条件下,计算统计量为

$$t = \frac{\bar{X} - \mu_0}{S_{\bar{X}}} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{3.42 - 3.30}{0.40/\sqrt{35}} = 1.77$$

(3) 确定  $P$  值,作出推断结论

本例自由度  $\nu = n - 1 = 35 - 1 = 34$ ,查附表 2,得  $t_{0.05/2,34} = 2.032$ 。因为  $t < t_{0.05/2,34}$ ,故  $P > 0.05$ ,表明差异无统计学意义,按  $\alpha = 0.05$  水准不拒绝  $H_0$ ,即根据现有样本信息,尚不能认为该地难产儿与一般新生儿平均出生体重不同。

### 第二节 配对样本均数 $t$ 检验

配对样本均数  $t$  检验简称配对  $t$  检验 (paired  $t$  test),又称非独立两样本均数  $t$  检验。



适用于配对设计计量资料均数的比较，其比较目的是检验两相关样本均数所代表的未知总体均数是否有差别。配对设计 (paired design) 是将受试对象按某些重要特征相近的原则配成对子，每对中的两个个体随机地给予两种处理。应用配对设计可以减少实验的误差和控制非处理因素，提高统计处理的效率。配对设计处理分配方式主要有三种情况：①两个同质受试对象分别接受两种处理，如把同窝、同性别和体重相近的动物配成一对，或把同性别、年龄相近及病情相同的病人配成一对；②同一受试对象或同一标本的两个部分，随机分配接受两种不同处理，如例 5.2 资料；③自身对比 (self-contrast)，即将同一受试对象处理 (实验或治疗) 前后的结果进行比较，如对高血压患者治疗前后某一生理指标进行比较。前两种情况属于随机对照的配对设计，应采用配对  $t$  检验判断两种处理的效果有无差别；后一种自身对比不能随机分配处理，其前后的差别可能是处理因素的作用，也可能是其他因素的影响，因此在实验性研究中应设平行对照。

配对设计的资料具有对子内数据一一对应的特征，研究者应关心对子的效应差值而不是各自的效应值。因此进行配对  $t$  检验时，首选应计算各对数据间的差值  $d$ ，将  $d$  作为变量计算均数。配对样本  $t$  检验的基本原理是假设两种处理的效应相同，理论上差值  $d$  的总体均数  $\mu_d$  为 0，现有样本差值均数不等于 0 的  $\bar{d}$  可以来自  $\mu_d=0$  的总体，也可以来自  $\mu_d \neq 0$  的总体。因此可将该检验理解为差值样本均数  $\bar{d}$  与已知总体均数  $\mu_d (\mu_d=0)$  比较的单样本  $t$  检验，其检验统计量为

$$t = \frac{\bar{d} - \mu_d}{S_{\bar{d}}} = \frac{\bar{d} - 0}{S_{\bar{d}}} = \frac{\bar{d}}{S_d / \sqrt{n}} \quad (5-2)$$

式中  $d$  为每对数据的差值， $\bar{d}$  为差值样本的均数， $S_d$  为差值样本的标准差， $S_{\bar{d}}$  为差值样本均数的标准差，即差值样本的标准误， $n$  为配对样本的对子数。

**例 5.2** 有 12 名接种卡介苗的儿童，八周后用两批不同的结核菌素，一批是标准结核菌素，一批是新制结核菌素，分别注射在儿童的前臂，两种结核菌素的皮肤浸润反应平均直径 (mm) 如表 5-1 所示，问两种结核菌素的反应性有无差别。

表 5-1 12 名儿童分别用两种结核菌素的皮肤浸润反应结果 (mm)

编号	标准品	新制品	差值 $d$	$d^2$
1	12.0	10.0	2.0	4.00
2	14.5	10.0	4.5	20.25
3	15.5	12.5	3.0	9.00
4	12.0	13.0	-1.0	1.00
5	13.0	10.0	3.0	9.00
6	12.0	5.5	6.5	42.25
7	10.5	8.5	2.0	4.00
8	7.5	6.5	1.0	1.00
9	9.0	5.5	3.5	12.25
10	15.0	8.0	7.0	49.20
11	13.0	6.5	6.5	42.25
12	10.5	9.5	1.0	1.00
合计			39( $\sum d$ )	195( $\sum d^2$ )

(1) 建立检验假设，确定检验水准

$H_0: \mu_d=0$ ，两种结核菌素的皮肤浸润反应总体平均直径差异为零

$H_1: \mu_d \neq 0$ ，两种结核菌素的皮肤浸润反应总体平均直径差异不为零

$\alpha=0.05$



(2) 计算检验统计量

先计算差值  $d$  及  $d^2$ , 如表 5-1 第四、五列所示:

本例  $\sum d = 39$ ,  $\sum d^2 = 195$

计算差值均数:

$$\bar{d} = \sum d / n = 39 / 12 = 3.25$$

计算差值的标准差:

$$S_d = \sqrt{\frac{\sum d^2 - (\sum d)^2}{n-1}} = \sqrt{\frac{195 - \frac{(39)^2}{12}}{12-1}} = 2.4909$$

计算差值的标准误:

$$S_{\bar{d}} = \frac{S_d}{\sqrt{n}} = \frac{2.4909}{3.464} = 0.7191$$

按公式 (5-2) 计算, 得

$$t = \frac{\bar{d}}{S_{\bar{d}}} = \frac{3.25}{0.7191} = 4.5195$$

(3) 确定  $P$  值, 作出推断结论

自由度  $\nu = n - 1 = 12 - 1 = 11$ , 查附表 2, 得:

$$t_{0.05/2, 11} = 2.201, \quad t_{0.01/2, 11} = 3.106$$

本例  $t > t_{0.01/2, 11}$ ,  $P < 0.01$ , 差别有统计学意义, 拒绝  $H_0$ , 接受  $H_1$ , 可认为两种方法皮肤浸润反应结果的差别有统计学意义。

### 第三节 两独立样本均数 $t$ 检验

两独立样本  $t$  检验 (two independent samples  $t$ -test), 又称成组  $t$  检验, 它适用于完全随机设计的两样本均数的比较, 其目的是检验两样本所来自总体的均数是否相等。完全随机设计是将受试对象随机地分配到两组中, 每组对象分别接受不同的处理, 分析比较两组的处理效应。

两独立样本  $t$  检验要求两样本所代表的总体服从正态分布  $N(\mu_1, \sigma_1^2)$  和  $N(\mu_2, \sigma_2^2)$ , 且两总体方差  $\sigma_1^2 = \sigma_2^2$ , 即方差齐性 (homogeneity of variance)。若两总体方差不齐, 可采用  $t'$  检验, 或进行变量变换及用秩和检验方法处理 (见第九章)。

两独立样本  $t$  检验的检验假设是两总体均数相等, 即  $H_0: \mu_1 = \mu_2$ , 也可表述为  $\mu_1 - \mu_2 = 0$ , 这里可将两样本均数的差值  $\bar{X}_1 - \bar{X}_2$  看成一个统计量,  $S_{\bar{X}_1 - \bar{X}_2}$  就是差值的标准误, 则在  $H_0$  条件下两独立样本均数  $t$  检验可视为样本  $\bar{X}_1 - \bar{X}_2$  与已知总体均数  $\mu_1 - \mu_2 = 0$  的单样本  $t$  检验, 统计量计算公式为

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{S_{\bar{X}_1 - \bar{X}_2}} = \frac{|\bar{X}_1 - \bar{X}_2|}{S_{\bar{X}_1 - \bar{X}_2}}, \quad \nu = n_1 + n_2 - 2 \quad (5-3)$$

其中:

$$S_{\bar{X}_1 - \bar{X}_2} = \sqrt{S_c^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)} \quad (5-4)$$

$$S_c^2 = \frac{\sum X_1^2 - \frac{(\sum X_1)^2}{n_1} + \sum X_2^2 - \frac{(\sum X_2)^2}{n_2}}{n_1 + n_2 - 2} \quad (5-5)$$

$S_c^2$  称为合并方差 (pooled variance), 公式 (5-5) 可用于已知两样本观察值原始资料时计算, 当两样本标准差  $S_1$  和  $S_2$  已知时, 合并方差  $S_c^2$  为



$$S_c^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \quad (5-6)$$

**例 5.3** 25 例糖尿病患者随机分成两组，甲组单纯用药物治疗，乙组采用药物治疗合并饮食疗法，二个月后测空腹血糖 (mmol/L) 如表 5-2 所示，问两种疗法治疗后患者血糖值是否相同？

表 5-2 25 名糖尿病患者两种疗法治疗后二个月血糖值 (mmol/L)

编号	甲组血糖值( $X_1$ )	编号	乙组血糖值( $X_2$ )
1	8.4	1	5.4
2	10.5	2	6.4
3	12.0	3	6.4
4	12.0	4	7.5
5	13.9	5	7.6
6	15.3	6	8.1
7	16.7	7	11.6
8	18.0	8	12.0
9	18.7	9	13.4
10	20.7	10	13.5
11	21.1	11	14.8
12	15.2	12	15.6
		13	18.7

(1) 建立检验假设，确定检验水准

$H_0: \mu_1 = \mu_2$ ，两种疗法治疗后患者血糖值的总体均数相同

$H_1: \mu_1 \neq \mu_2$ ，两种疗法治疗后患者血糖值的总体均数不同

$\alpha = 0.05$

(2) 计算检验统计量

由原始数据计算得：

$n_1 = 12$ ， $\sum X_1 = 182.5$ ， $\sum X_1^2 = 2\ 953.43$ ， $\bar{X}_1 = 15.21$

$n_2 = 13$ ， $\sum X_2 = 141.0$ ， $\sum X_2^2 = 1\ 743.16$ ， $\bar{X}_2 = 10.85$

代入公式 (5-5) 和 (5-4)，得：

$$S_c^2 = \frac{2\ 953.43 - \frac{(182.5)^2}{12} + 1\ 743.16 - \frac{(141.0)^2}{13}}{12 + 13 - 2} = 17.03$$

$$S_{\bar{X}_1 - \bar{X}_2} = \sqrt{17.03 \left( \frac{1}{12} + \frac{1}{13} \right)} = 1.652$$

按公式 (5-3)，算得

$$t = \frac{15.21 - 10.85}{1.652} = 2.639$$

(3) 确定  $P$  值，作出推断结论

两独立样本  $t$  检验自由度  $\nu = n_1 + n_2 - 2 = 12 + 13 - 2 = 23$ ；查  $t$  界值表， $t_{0.05/2, 23} = 2.069$ ， $t_{0.01/2, 23} = 2.807$ 。由于  $t_{0.01/2, 23} > t > t_{0.05/2, 23}$ ， $0.01 < P < 0.05$ ，按  $\alpha = 0.05$  的水准拒绝  $H_0$ ，接受  $H_1$ ，差异有统计学意义，故可认为两种疗法治疗糖尿病患者二个月后测得的空腹血糖值的均数不同。

有些资料（如抗体滴度的资料），宜用几何均数表示其平均水平。由于这些资料不服从正态分布（常服从对数正态分布），两样本所代表的总体方差也可能不等。当进行几何



均数的假设检验时, 应先进行变量变换, 即将这些观察值  $X$  用  $\lg X$  来代替 (即作对数变换),  $\lg X$  往往近似服从正态分布, 相应的两总体方差也可能相等。因为数据变换并未改变两组数据间的相对关系, 故可用前述的三种形式的  $t$  检验对  $\lg X$  进行分析。

## 第四节 方差不齐时两样本均数的 $t'$ 检验

当两总体方差不等 (方差不齐) 时, 两独立样本均数的比较, 可采用  $t'$  检验, 亦称近似  $t$  检验 (separate variance estimation  $t$ -test)。

### 一、方差齐性检验

由两样本方差推断两总体方差是否相同的检验方法可用  $F$  检验。 $F$  检验要求资料服从正态分布。检验统计量  $F$  值按公式 (5-7) 计算:

$$F = \frac{S_1^2 \text{ (较大)}}{S_2^2 \text{ (较小)}}, \quad \nu_1 = n_1 - 1, \quad \nu_2 = n_2 - 1 \quad (5-7)$$

式中  $S_1^2$  为较大的样本方差,  $S_2^2$  为较小的样本方差。检验统计量  $F$  值为两个样本方差之比, 若样本方差的不同仅为抽样误差的影响,  $F$  值一般不会偏离 1 太远。求得  $F$  值后, 查附表 3 (方差齐性检验用的  $F$  界值表) 得  $P$  值。一般取  $\alpha = 0.05$  水准作判断, 若  $F \geq F_{0.05(\nu_1, \nu_2)}$ ,  $P \leq 0.05$ , 拒绝  $H_0$ , 接受  $H_1$ , 可认为两总体方差不齐。若  $F < F_{0.05(\nu_1, \nu_2)}$ ,  $P > 0.05$ , 则认为两总体方差具有齐性。方差齐性检验步骤与前述假设检验的基本步骤相同。

**例 5.4** 两组小白鼠分别饲以高蛋白和低蛋白饲料, 4 周后记录小白鼠体重增加量 (g) 如表 5-3 所示, 问两组动物体重增加量的均数是否相等?

表 5-3 两种饲料喂养小白鼠 4 周后体重增重 (g) 情况

编号	高蛋白组体重增加量( $X_1$ )	编号	低蛋白组体重增加量( $X_2$ )
1	50	1	36
2	47	2	38
3	42	3	37
4	43	4	38
5	39	5	36
6	51	6	39
7	43	7	37
8	48	8	35
9	51	9	33
10	42	10	37
11	50	11	39
12	43	12	34
		13	36

(1) 建立检验假设, 确定检验水准

$H_0: \sigma_1^2 = \sigma_2^2$ , 即高蛋白与低蛋白饲料喂养后小白鼠体重增加量的总体方差相同

$H_1: \sigma_1^2 \neq \sigma_2^2$ , 即高蛋白与低蛋白饲料喂养后小白鼠体重增加量的总体方差不同

$\alpha = 0.05$

(2) 计算检验统计量

对表 5-3 的数据计算可得  $S_1^2 = 17.659$ ,  $S_2^2 = 3.269$ , 按公式 (5-7) 有

$$F = \frac{S_1^2}{S_2^2} = \frac{17.659}{3.269} = 5.402$$



### (3) 确定 $P$ 值, 作出统计推论

自由度  $\nu_1 = n_1 - 1 = 12 - 1 = 11$ ,  $\nu_2 = n_2 - 1 = 13 - 1 = 12$ , 查方差齐性检验用的  $F$  界值表(附表 3),  $F_{0.05(11,12)} \approx 3.34$ ,  $F > F_{0.05(11,12)}$ , 故  $P < 0.05$ , 差别有统计学意义, 按  $\alpha = 0.05$  水准, 拒绝  $H_0$ , 接受  $H_1$ 。故认为两组体重增加量的总体方差不齐, 不可直接用两独立样本均数  $t$  检验, 而应采用下面的  $t'$  检验。

## 二、 $t'$ 检验

$t'$  检验有三种方法, 包括 Satterthwaite 法近似  $t$  检验、Welch 法近似  $t$  检验和 Cochran & Cox 法近似  $t$  检验。Cochran & Cox 法是对临界值校正, Satterthwaite 法和 Welch 法是对自由度进行校正。这里主要介绍 Satterthwaite 法和 Cochran & Cox 法。 $t'$  统计量的计算公式为:

$$t' = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad (5-8)$$

Cochran & Cox 法校正临界值  $t'_{\alpha/2}$  的公式为

$$t'_{\alpha/2} = \frac{S_{\bar{X}_1}^2 \times t_{\alpha(\nu_1)} + S_{\bar{X}_2}^2 \times t_{\alpha(\nu_2)}}{S_{\bar{X}_1}^2 + S_{\bar{X}_2}^2}, \quad \nu = n_1 + n_2 - 2 \quad (5-9)$$

式中  $\nu_1 = n_1 - 1$ ,  $\nu_2 = n_2 - 1$ , 根据校正的临界值, 作出推断结论。

Satterthwaite 法  $t'$  检验的自由度校正公式为

$$\nu = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}} \quad (5-10)$$

根据自由度查  $t$  界值表, 作出推断结论。

目前, Satterthwaite 法是统计软件中普遍使用的方法。

**例 5.5** 对例 5.4 资料进行  $t'$  检验, 比较两组小白鼠增重均数是否不同。

(1) 建立检验假设, 确定检验水准

$H_0: \mu_1 = \mu_2$ , 即两种饲料小白鼠增重总体均数相同

$H_1: \mu_1 \neq \mu_2$ , 即两种饲料小白鼠增重总体均数不相同

$\alpha = 0.05$

(2) 计算检验统计量

两总体方差不同, 应选用  $t'$  检验。即

$$t' = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{45.750 - 36.538}{\sqrt{\frac{17.659}{12} + \frac{3.269}{13}}} = 3.103$$

(3) 确定  $P$  值, 作出推断结论

按 Satterthwaite 法计算校正自由度, 得:

$$\nu = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}} = \frac{(17.659^2/12 + 3.269^2/13)^2}{\frac{(17.659^2/12)^2}{12 - 1} + \frac{(3.269^2/13)^2}{13 - 1}} = 11.696 \approx 12$$

查  $t$  界值表, 得  $t_{0.05/2(12)} = 2.179$ ,  $t' > t_{0.05/2(12)}$ ,  $P < 0.05$

按 Cochran & Cox 法计算校正界值, 先查  $t$  界值表, 得  $t_{0.05/2,11} = 2.201$ ,  $t_{0.05,12} = 2.179$ , 再按公式 (5-9), 算得:

$$t'_{0.05/2} = \frac{S_{\bar{X}_1}^2 \cdot t_{\alpha(\nu_1)} + S_{\bar{X}_2}^2 \cdot t_{\alpha(\nu_2)}}{S_{\bar{X}_1}^2 + S_{\bar{X}_2}^2} = \frac{1.472 \times 2.201 + 0.251 \times 2.179}{1.472 + 0.251} = 2.198$$

$t' > t'_{0.05/2}$ , 得  $P < 0.05$ 。



两种检验方法所获得的界值虽略有差异,但结论是一致的。按  $\alpha=0.05$  水准,拒绝  $H_0$ , 接受  $H_1$ , 差异有统计学意义。可认为两种饲料饲养后小白鼠增重的均数不同,高蛋白组高于低蛋白组。

应该注意:在 Cochran & Cox 法中,当  $n_1=n_2$  时,  $\nu_1=\nu_2=\nu$ ,  $t'=t$ ,  $t'_{\alpha/2}=t_{\alpha/2(\nu)}$ ,  $\nu=n-1$ , 此时即使方差不齐,也可用  $t$  检验。

## 第五节 $u$ 检验

由第四章可知,根据数理统计的中心极限定理,不论变量  $X$  的分布是否服从正态分布,当随机抽样的样本例数足够大时,样本均数  $\bar{X}$  服从正态分布  $N(\mu, \sigma_{\bar{X}}^2)$ , 其中  $\mu$  为原始的总体均数,  $\sigma$  为总体标准差,  $\sigma_{\bar{X}}=\sigma/\sqrt{n}$  为样本均数标准误。因此,可以计算  $u$  统计量:

$$u = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad (5-11)$$

如果变量  $X$  的分布服从正态分布,则不论样本例数  $n$  大小,样本均数  $\bar{X}$  均服从正态分布  $N(\mu, \sigma_{\bar{X}}^2)$ ,  $u$  服从标准正态分布。因此,当总体标准差  $\sigma$  已知,或样本量较大(如  $n>50$ ) 时,样本均数与总体均数比较、配对设计样本均数比较和两独立样本均数比较的假设检验,可以计算检验统计量  $u$  值。由于标准正态分布  $u$  的界值双侧时,有固定的  $u$  界值 ( $u_{0.05/2}=1.96$ ,  $u_{0.01/2}=2.58$ ), 因此用计算出的  $u$  值与这些界值比较,很容易确定  $P$  值和作出推断结论。在成组设计的两样本均数比较的统计量  $u$  值计算中,两样本均数差的标准误  $S_{\bar{X}_1 - \bar{X}_2} = \sqrt{S_1^2/n_1 + S_2^2/n_2}$ , 统计量  $u$  值的计算公式为

$$u = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \quad (5-12)$$

**例 5.6** 研究正常人与高血压患者胆固醇含量 ( $\text{mg}\%$ ) 的资料如下,试比较两组血清胆固醇含量有无差别。

正常人组  $n_1=506$ ,  $\bar{X}_1=180.6$ ,  $S_1=34.2$

高血压组  $n_2=142$ ,  $\bar{X}_2=223.6$ ,  $S_2=45.8$

(1) 建立检验假设,确定检验水平

$H_0: \mu_1 = \mu_2$ , 即正常人与高血压患者血清胆固醇值总体均数相同

$H_1: \mu_1 \neq \mu_2$ , 即正常人与高血压患者血清胆固醇值总体均数不同

$\alpha=0.05$ , 双侧

(2) 计算统计量  $u$  值

将已知数据代入公式 (5-12), 得

$$u = \frac{|180.6 - 223.6|}{\sqrt{34.2^2/506 + 45.8^2/142}} = 10.40$$

(3) 确定  $P$  值,作出推断结论

本例  $u=10.40 > 2.58$ , 故  $P < 0.01$ , 按  $\alpha=0.05$  水准拒绝  $H_0$ , 接受  $H_1$ , 可以认为正常人与高血压患者的血清胆固醇含量有差别,高血压患者高于正常人。

## 第六节 $t$ 检验中的注意事项

1. 假设检验结论正确的前提 作假设检验用的样本资料,必须能代表相应的总体,同时各对比组具有良好的组间均衡性,才能得出有意义的统计结论和有价值的专业结论。



这要求有严密的实验设计和抽样设计，如样本是从同质总体中抽取的一个随机样本，试验单位在干预前随机分组，有足够的样本量等。

2. 检验方法的选用及其适用条件 应根据分析目的、研究设计、资料类型、样本量大小等选用适当的检验方法。 $t$  检验是以正态分布为基础的，资料的正常性可用正态性检验方法检验，或直观地通过频数表以及相应的直方图予以判断。若资料为非正态分布，可采用数据变换的方法，尝试将资料变换成正态分布资料后进行分析；如果数据变换后仍为非正态分布，则可选用非参数检验。

配对设计的计量资料应采用配对  $t$  检验，而完全随机设计的两样本均数的比较，作两独立样本  $t$  检验时，并要求两组资料方差齐。若方差不齐，应选用近似的  $t'$  检验，或采用数据变换的方法使方差具有齐性，也可以采用第九章的非参数检验。若资料为大样本 ( $n > 50$ ) 或已知总体方差，则可选用  $u$  检验。如果有两个以上样本均数要作比较，则应使用第六章方差分析。

3. 双侧检验与单侧检验的选择 需根据研究目的和专业知予以选择。单侧检验和双侧检验中的  $t$  值计算过程相同，只是  $t$  界值不同，对同一资料作单侧检验更容易获得显著的结果。单双侧检验的选择，应在统计分析工作开始之前就决定，若缺乏这方面的依据，一般应选用双侧检验。

4. 假设检验的结论不能绝对化 假设检验统计结论的正确性是以概率作保证的，作统计结论时不能绝对化。在报告结论时，最好列出概率  $P$  的确切数值或给出  $P$  值的范围，如写成  $0.02 < P < 0.05$ ，同时应注明采用的是单侧检验还是双侧检验，以便读者与同类研究进行比较。当  $P$  接近临界值时，下结论应慎重。

5. 正确理解  $P$  值的统计意义  $P$  是指在无差别假设  $H_0$  的总体中进行随机抽样，所观察到的等于或大于现有统计量值的概率。其推断的基础是小概率事件的原理，即概率很小的事件在一次抽样研究中几乎是不可能发生的，如发生则拒绝  $H_0$ 。因此， $P \leq 0.05$  只能说明统计学意义的“显著”，并不表示其差异具有实际意义。从本章  $t$  检验的计算公式可以看出，假设检验的结论与样本大小有关，当样本量足够大时，标准误趋于零，无论两样本均数相差多少，都能得到足以拒绝  $H_0$  的  $t$  值和  $P$  值。

6. 假设检验和可信区间的区别 假设检验用以推断总体均数间是否相同，而可信区间则用于估计总体均数所在的范围。

## 第七节 假设检验中两类错误

假设检验无论是拒绝  $H_0$  还是不拒绝  $H_0$ ，都可能犯错误，见表 5-4。

表 5-4 假设检验的两类错误

真实情况	假设检验结论	
	拒绝 $H_0$	不拒绝 $H_0$
$H_0$ 成立	I 型错误 ( $\alpha$ )	推断正确 ( $1-\alpha$ )
$H_1$ 成立	推断正确 ( $1-\beta$ )	II 型错误 ( $\beta$ )

当  $H_0$  为真时，假设检验结论拒绝  $H_0$  接受  $H_1$ ，这类错误称为第一类错误或 I 型错误 (type I error)，亦称假阳性错误。前面所讲的检验水准  $\alpha$ ，就是预先规定的允许犯 I 型错误概率的最大值。如  $\alpha = 0.05$ ，表示在  $H_0$  为真的条件下重复 100 次试验，理论上会有 5 次拒绝  $H_0$ 。

当真实情况为  $H_0$  不成立时，假设检验结论不拒绝  $H_0$ ，这类错误称为第二类错误或 II





型错误 (type II error), 亦称假阴性错误。其概率大小用  $\beta$  表示。 $\beta$  只取单侧, 其值的大小一般未知, 必须在知道两总体的标准差、均数的实际差值和样本含量时, 才能算出。

为了更好地理解两类错误的意义, 以样本均数与总体均数比较的单侧  $u$  检验来说明, 见图 5-1。设  $H_0: \mu = \mu_0$ ,  $H_1: \mu > \mu_0$ ; 若  $H_0$  实际上是成立的, 但由于抽样误差的存在, 偶然得到较大的  $\bar{X}_1$  值, 使得  $u \geq u_\alpha$ , 按  $\alpha = 0.05$  的检验水准拒绝  $H_0$ , 接受  $H_1$ , 结论为  $\mu > \mu_0$ , 此时犯 I 型错误最大可能概率值为  $\alpha$ 。相反, 若  $\mu$  确实大于  $\mu_0$ , 即  $H_0$  不成立,  $H_1$  成立, 由于抽样的偶然性, 得到较小的  $\bar{X}_2$  值, 使得  $u < u_\alpha$ , 检验结论不拒绝  $H_0$ , 此时犯 II 型错误, 其概率为  $\beta$ 。 $1 - \beta$  称为检验效能 (power of a test), 也称把握度, 表示当两总体确实有差别时, 按规定的检验水准能发现其差别的能力。

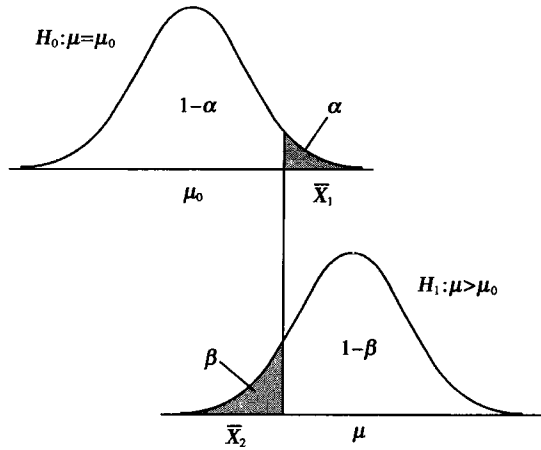


图 5-1 I 型错误与 II 型错误示意图  
(以单侧  $u$  检验为例)

由图 5-1 中可见,  $\alpha$  愈小,  $\beta$  愈大; 相反,  $\alpha$  愈大,  $\beta$  愈小。若要同时减小 I 型错误  $\alpha$  和 II 型错误  $\beta$ , 唯一的方法就是增加样本含量  $n$ 。

## 练 习 题

### 一、单项选择题

- 两样本均数比较, 检验结果  $P > 0.05$  说明
  - 两总体均数的差别较小
  - 两总体均数的差别较大
  - 支持两总体无差别的结论
  - 不支持两总体有差别的结论
  - 可以确认两总体无差别
- 由两样本均数的差别推断两总体均数的差别, 其差别有统计学意义是指
  - 两样本均数的差别具有实际意义
  - 两总体均数的差别具有实际意义
  - 两样本和两总体均数的差别都具有实际意义
  - 有理由认为两样本均数有差别
  - 有理由认为两总体均数有差别
- 两样本均数比较, 差别具有统计学意义时,  $P$  值越小说明
  - 两样本均数差别越大
  - 两总体均数差别越大
  - 越有理由认为两样本均数不同
  - 越有理由认为两总体均数不同
  - 越有理由认为两样本均数相同
- 减少假设检验的 II 类误差, 应该使用的方法是
  - 减少 I 类错误
  - 减少测量的系统误差
  - 减少测量的随机误差
  - 提高检验界值
  - 增加样本含量
- 两样本均数比较的  $t$  检验和  $u$  检验的主要差别是
  - $t$  检验只能用于小样本资料
  - $u$  检验要求大样本资料



C. *t* 检验要求数据方差相同

D. *t* 检验的检验效能更高

E. *u* 检验能用于两大样本均数比较

二、计算与分析

1. 已知正常成年男子血红蛋白均值为 140g/L, 今随机调查某厂成年男子 60 人, 测其血红蛋白均值为 125g/L, 标准差 15g/L。问该厂成年男子血红蛋白均值与一般成年男子是否不同?

2. 某研究者为比较耳垂血和手指血的白细胞数, 调查 12 名成年人, 同时采取耳垂血和手指血见下表, 试比较两者的白细胞数有无不同。

成人耳垂血和手指血白细胞数 (10g/L)

编号	耳垂血	手指血	编号	耳垂血	手指血
1	9.7	6.7	7	4.7	4.6
2	6.2	5.4	8	5.8	4.2
3	7.0	5.7	9	7.8	7.5
4	5.3	5.0	10	8.6	7.0
5	8.1	7.5	11	6.1	5.3
6	9.9	8.3	12	9.9	10.3

3. 分别测得 15 名健康人和 13 名 III 度肺气肿病人痰中  $\alpha_1$  抗胰蛋白酶含量 (g/L) 如下表, 问健康人与 III 度肺气肿病人  $\alpha_1$  抗胰蛋白酶含量是否不同?

健康人与 III 度肺气肿患者  $\alpha_1$  抗胰蛋白酶含量 (g/L)

健康人	III 度肺气肿患者	健康人	III 度肺气肿患者
2.7	3.6	1.9	4.8
2.2	3.4	1.3	5.6
4.1	3.7	1.5	4.1
4.3	5.4	1.7	3.3
2.6	3.6	1.3	4.3
1.9	6.8	1.3	
1.7	4.7	1.9	
0.6	2.9		

4. 某地对 241 例正常成年男性面部上颌间隙进行了测定, 得其结果如下表, 问不同身高正常男性其上颌间隙是否不同?

某地 241 名正常男性上颌间隙 (cm)

身高(cm)	例数	均数	标准差
161~	116	0.2189	0.2351
172~	125	0.2280	0.2561

5. 将钩端螺旋体病人的血清分别用标准株和水生株作凝溶试验, 测得稀释倍数如下表, 问两组的平均效价有无差别?

钩端螺旋体病患者凝溶试验的稀释倍数

标准株	100	200	400	400	400	400	800	1 600	1 600	1 600	3 200	3 200	3 200
水生株	100	100	100	200	200	200	200	400	400	800	1 600		

6. 为比较男、女大学生的血清谷胱甘肽过氧化物酶 (GSH-Px) 的活力是否相同, 某



医生从某大学 18~22 岁大学生中随机抽查男生 48 名, 女生 46 名, 测定其血清谷胱甘肽过氧化酶含量 (活力单位), 男、女性的均数分别为 96.53 和 93.73, 男、女性标准差分别为 7.66 和 14.97。问男女性的 GSH-Px 是否相同?

(沈其君 施 榕)

### *t* Test

W. S. Gosset in 1908 derived a complicated formula for the density function of  $t = (\bar{X} - \mu) / (S / \sqrt{n})$  for random samples of sized  $n$  from a normal population, and he published his results under the pen name 'Student'. Ever since, the statistic method has been known as 'Student's  $t$ '. It has the following characteristics:

- It is mound-shape and symmetric about  $t=0$ , just like  $z$ .
- It is more variable than  $z$ , with 'heavy tails'; that is, the  $t$  curve does not approach the horizontal axis as quickly as  $z$  does. This is because the  $t$  statistic involves two random quantities,  $\bar{X}$  and  $S$ , whereas the  $z$  statistic only involves the sample mean  $\bar{X}$ .
- The shape of the  $t$  distribution depends on the sample size  $n$ . As  $n$  increases, the variability of  $t$  decreases because the estimates  $s$  of  $\sigma$  is based on more and more information. Eventually, when  $n$  is infinitely large, the  $t$  and  $z$  distributions are identical.

#### Type I versus Type II Errors in Hypothesis Testing:

Because the predictions in  $H_0$  and  $H_1$  are written so that they are mutually exclusive and all inclusive. We have a situation where one is true and the other is automatically false. When  $H_0$  is true then  $H_1$  is false, if  $H_0$  is rejected we have made an error, called a Type I error. When  $H_0$  is false then  $H_1$  is true, if  $H_0$  is accepted we have made an error, called a Type II error.

- $P$  (Type I error) = the probability of rejecting a true null hypothesis.
- $P$  (Type II error) = the probability of accepting a false null-hypothesis.

## 第六章 方差分析

上一章介绍了两个样本均数比较的  $t$  检验, 而对于独立地来自于正态分布总体且总体方差相等的  $k$  ( $k \geq 3$ ) 个样本均数的比较,  $t$  检验不再适用, 方差分析 (analysis of variance, ANOVA) 则是解决上述问题的重要分析方法。方差分析还可用于多因素多水平组间效应分析、回归分析等。本章将主要介绍单因素多水平以及两因素多水平组间效应分析的方差分析。

方差分析由 R. A. Fisher (1923) 首先提出, 故又称为  $F$  检验。其基本思想是将全部观测值的总变异按影响因素分解为相应的若干部分变异, 在此基础上, 计算假设检验的统计量  $F$  值, 实现对总体均数是否有差别的推断。

### 第一节 完全随机设计的方差分析

完全随机设计 (completely random design) 又称单因素方差分析 (one-way analysis of variance, one-way ANOVA), 是一种将实验对象随机分配到不同处理组的单因素设计方法。此类设计只考察一个处理因素, 通过对该因素不同水平组均值的比较, 推断该处理因素不同水平组的均值之间的差异有无统计意义。完全随机设计的数据结构一般形式如表 6-1 所示, 其中  $k$  为处理因素的水平数,  $X_{ij}$  为处理因素第  $i$  水平的第  $j$  个观测值,  $n_i$  ( $i=1, 2, \dots, k$ ) 为处理因素第  $i$  水平组的观测例数,  $n$  为总例数,  $\bar{X}_i$  为处理因素第  $i$  水平组的均数,  $\bar{X}$  为总均数,  $S_i^2$  为处理因素第  $i$  水平组的方差,  $S^2$  为全部观测值的方差。

表 6-1 完全随机设计方差分析的数据结构

	处理因素				合计
	水平 1	水平 2	...	水平 $k$	
	$X_{11}$	$X_{21}$	...	$X_{k1}$	
	$X_{12}$	$X_{22}$	...	$X_{k2}$	
	...	...	...	...	
	$X_{1n}$	$X_{2n}$	...	$X_{kn}$	
$n_i$	$n_1$	$n_2$	...	$n_k$	$n$
$\bar{X}_i$	$\bar{X}_1$	$\bar{X}_2$	...	$\bar{X}_k$	$\bar{X}$
$S_i^2$	$S_1^2$	$S_2^2$	...	$S_k^2$	$S^2$

表中,  $n$  个观测值彼此不同, 可以用方差来反映该变异程度。方差的分子部分为  $n$  个观测值的离均差平方和, 被称为总变异, 记为  $SS_T$ 。对此可以作如下分解:

$$\begin{aligned} SS_T &= \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} [(X_{ij} - \bar{X}_i) + (\bar{X}_i - \bar{X})]^2 \\ &= \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 + 2 \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X}) \end{aligned}$$

其中

$$2 \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X}) = 0$$

故



$$SS_T = \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

其中,  $\sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2$  称为组间变异, 是组内均值  $\bar{X}_i$  与总均值  $\bar{X}$  之差的平方和, 记为

$SS_{TR}$ , 反映了处理因素各个水平组间的差异, 同时也包含了随机误差;  $\sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$  称为组内变异, 是组内各个观测值  $X_{ij}$  与本组内均值  $\bar{X}_i$  之差的平方和, 记为  $SS_E$ , 它反映了各组内样本的随机波动。

由上可见, 总变异  $SS_T$ 、组间变异  $SS_{TR}$  和组内变异  $SS_E$  之间满足下式:

$$SS_T = SS_{TR} + SS_E \quad (6-1)$$

其中, 总变异自由度  $\nu_T = n - 1$ , 组间变异自由度  $\nu_{TR} = k - 1$ , 组内变异自由度  $\nu_E = (n_1 - 1) + (n_2 - 1) + \dots + (n_k - 1) = n - k$ 。对于自由度, 同样有

$$\nu_T = \nu_{TR} + \nu_E \quad (6-2)$$

各变异除以相应自由度得到相应平均变异, 即方差, 又叫均方。组间均方:

$$MS_{TR} = \frac{SS_{TR}}{\nu_{TR}} = \frac{SS_{TR}}{k-1} \quad (6-3)$$

组内(误差)均方:

$$MS_E = \frac{SS_E}{\nu_E} = \frac{SS_E}{n-k} \quad (6-4)$$

检验各处理组均值之间有无差异这一假设可以代之以比较  $MS_{TR}$  和  $MS_E$  来实现。 $MS_{TR}$  与  $MS_E$  之比即构成了方差分析的统计量:

$$F = \frac{MS_{TR}}{MS_E} \quad (6-5)$$

可以证明, 当无效假设  $H_0$  成立时, 这样定义的统计量  $F$  服从自由度为  $(k-1, n-k)$  的  $F$  分布。 $F$  值接近 1, 可认为均值的差异只源于随机波动, 而非处理因素作用;  $F$  值大于 1 并且  $F > F_{\alpha(k-1, n-k)}$ ,  $P < \alpha$ , 则按  $\alpha$  水准拒绝  $H_0$ , 表明有随机波动之外的处理因素造成了均值的差异。

以上计算过程可用完全随机设计方差分析表进行概括(表 6-2)。

表 6-2 完全随机设计的方差分析表

变异来源	平方和 SS	自由度 $\nu$	均方 MS	F 值
总变异	$SS_T = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2$ $= (n-1)S^2$	$\nu_T = n-1$		
处理组间	$SS_{TR} = \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2$	$\nu_{TR} = k-1$	$MS_{TR} = SS_{TR}/\nu_{TR}$	$F = \frac{MS_{TR}}{MS_E}$
组内(误差)	$SS_E = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$ $= \sum_{i=1}^k (n_i - 1)S_i^2$	$\nu_E = n-k$	$MS_E = SS_E/\nu_E$	

**例 6.1** 某研究者拟探讨枸杞多糖(LBP)对酒精性脂肪肝(AFL)大鼠 GSH (mg/gprot) 的影响, 将周龄 7 周的 Wistar 大鼠 36 只随机分为甲、乙、丙三组, 其中甲(正常对照组) 12 只, 其余 24 只用乙醇灌胃 10 周造成大鼠慢性 AFL 模型后, 再随机分为 2 组, 乙(LBP 治疗组) 12 只, 丙(戒酒组) 12 只, 8 周后测量三组 GSH 值, 结果见表 6-3。试问三种处理方式大鼠的 GSH 值是否相同?



表 6-3 三组大鼠 GSH 值 (mg/gprot)

	甲	乙	丙	合计
	79.81	87.58	60.29	
	80.60	70.73	62.63	
	65.21	68.18	38.25	
	89.09	93.88	49.03	
	93.32	69.45	64.98	
	88.60	81.34	65.85	
	60.70	87.58	35.45	
	87.04	64.45	56.96	
	91.34	68.57	37.48	
	85.54	79.05	54.53	
	104.28	80.36	46.56	
	72.29	56.40	55.23	
$n_i$	12	12	12	36
$\bar{X}_i$	83.15	75.63	52.27	70.35
$S_i$	12.30	11.07	10.85	17.35

方差分析具体步骤:

(1) 提出检验假设, 确定检验水准

$H_0: \mu_1 = \mu_2 = \mu_3$ , 即三个组 GSH 值的总体均数相同

$H_1: \mu_1, \mu_2, \mu_3$  不全相同, 即三个组 GSH 值的总体均数不全相同

$\alpha = 0.05$

(2) 计算检验统计量  $F$  值

按表 6-2 中计算式计算检验统计量, 将其列于表 6-4 的方差分析表中。

表 6-4 完全随机设计的方差分析表

变异来源	平方和 SS	自由度 $\nu$	均方 MS	$F$ 值
总变异	10 530.6003	35		
处理组间	6 223.8752	2	3 111.9376	23.85
组内(误差)	4 306.7251	33	130.5068	

(3) 确定  $P$  值, 做出推断结论

分子自由度  $\nu_{TR} = 2$ , 分母自由度  $\nu_E = 33$ , 查  $F$  界值表 (方差分析用), 因  $F$  界值表中无  $\nu_E = 33$ , 取  $\nu_E = 32$ ,  $F_{0.05(2,32)} = 3.30$ 。由于  $F > F_{0.05(2,32)}$ , 从而  $P < 0.05$ , 按照  $\alpha = 0.05$  的显著性水准拒绝  $H_0$ , 可以认为三种处理方式大鼠的 GSH 值不全相同。

## 第二节 随机区组设计的方差分析

随机区组设计 (randomized block design) 又称为配伍组设计, 其做法是先将受试对象按条件相同或相近组成  $m$  个区组 (或称配伍组), 每个区组中有  $k$  个受试对象, 再将其随机地分到  $k$  个处理组中。随机区组设计在  $m$  个区组和  $k$  个处理水平组构成  $mk$  个格子, 每个格子仅一个数据  $X_{ij}$  ( $i = 1, 2, 3, \dots, k; j = 1, 2, 3, \dots, m$ ), 其方差分析属无重复数据的两因素方差分析 (two-way ANOVA), 数据结构如表 6-5 所示。



表 6-5 随机区组设计方差分析的数据结构

区组(B)	处理因素(A)			
	水平 1	水平 2	...	水平 k
区组 1	$X_{11}$	$X_{21}$	...	$X_{k1}$
区组 2	$X_{12}$	$X_{22}$	...	$X_{k2}$
...	...	...	...	...
区组 m	$X_{1m}$	$X_{2m}$	...	$X_{km}$

与完全随机设计方差分析情况类似,总变异  $SS_T$  可分解为:处理 A 因素的变异  $SS_A$ 、区组 B 因素的变异  $SS_B$  以及随机误差  $SS_E$ 。即

$$SS_T = SS_A + SS_B + SS_E \quad (6-6)$$

相应的自由度:

$$\nu_T = \nu_A + \nu_B + \nu_E \quad (6-7)$$

有关计算式归纳于表 6-6 的方差分析表中。

表 6-6 随机区组设计的方差分析表

变异来源	平方和 SS	自由度 $\nu$	均方 MS	F 值
总变异	$SS_T = \sum_{i=1}^k \sum_{j=1}^m (X_{ij} - \bar{X})^2$ $= S^2(n-1)$	$\nu_T = n-1$		
处理间	$SS_A = \sum_{i=1}^k m(\bar{X}_i - \bar{X})^2$	$\nu_A = k-1$	$MS_A = SS_A/\nu_A$	$F_A = \frac{MS_A}{MS_E}$
区组间	$SS_B = \sum_{j=1}^m k(\bar{X}_j - \bar{X})^2$	$\nu_B = m-1$	$MS_B = SS_B/\nu_B$	$F_B = \frac{MS_B}{MS_E}$
误差	$SS_E = SS_T - SS_A - SS_B$	$\nu_E = (k-1)(m-1)$	$MS_E = SS_E/\nu_E$	

其中,  $\bar{X}_i$  为 A 因素第  $i$  水平组的均数,  $\bar{X}_j$  为 B 因素第  $j$  区组的均数;  $n$  为总观察例数,  $\bar{X}$  为全部观测值总均数,  $S^2$  为全部观测值的方差。

可以看出,与完全随机设计相比,随机区组设计方差分析将总变异分解为三部分,即除处理组间变异之外,还将区组因素导致的变异也分离出来,从而减少了随机误差,提高了实验效率。

**例 6.2** 为探讨 Rgl 对镉诱导大鼠睾丸损伤的保护作用,研究者按照窝别把大鼠分成 10 个区组,然后将同一区组内的 3 只大鼠随机地分配到三个实验组,分别给予不同处理,一定时间后测量大鼠的睾丸 MT 含量 ( $\mu\text{g/g}$ ),数据如表 6-7 所示。试比较三种不同处理对大鼠 MT 含量有无差别?

表 6-7 三组大鼠 MT 含量值 ( $\mu\text{g/g}$ )

窝别	对照组	氯化镉组	Rgl+氯化镉组	$\bar{X}_j$
1	40.6	78.3	116.3	78.4
2	44.8	86.0	124.6	85.1
3	36.7	72.1	149.0	85.9
4	49.9	95.4	128.8	91.4
5	59.8	99.2	134.1	97.7
6	54.5	95.9	133.0	94.5
7	38.4	76.4	115.6	76.8



续表

窝别	对照组	氯化镉组	Rgl+氯化镉组	$\bar{X}_j$
8	41.6	79.9	117.0	79.5
9	46.8	86.5	128.4	87.2
10	44.7	85.3	124.3	84.8
$\bar{X}_i$	45.8	85.5	127.1	86.1( $\bar{X}$ )

方差分析步骤如下:

(1) 建立检验假设, 确定检验水准

$H_{0(A)}: \mu_1 = \mu_2 = \mu_3$ , 即三组大鼠 MT 含量的总体均值相同

$H_{1(A)}: \mu_1、\mu_2、\mu_3$  不全相等, 即三组大鼠 MT 含量的总体均值不全相同

$H_{0(B)}: \tau_1 = \tau_2 = \dots = \tau_{10}$ , 即不同窝别大鼠 MT 含量的总体均值相同

$H_{1(B)}: \tau_1, \tau_2, \dots, \tau_{10}$  不全相等, 不同窝别大鼠 MT 含量的总体均值不全相同

$\alpha = 0.05$

(2) 计算检验统计量  $F$  值

$$S^2 = 34.85$$

$$SS_T = S^2(n-1) = 34.85 \times (30-1) = 35\ 226.4630$$

$$SS_A = \sum_{i=1}^k m(\bar{X}_i - \bar{X})^2 = 10 \times (45.8 - 86.1)^2 + 10 \times (85.5 - 86.1)^2 + 10 \times (127.1 - 86.1)^2 \\ = 33\ 054.5$$

$$SS_B = \sum_{j=1}^m k(\bar{X}_j - \bar{X})^2 = 3 \times (78.4 - 86.1)^2 + 3 \times (85.1 - 86.1)^2 + \dots + 3 \times (84.8 - 86.1)^2 \\ = 1\ 279.47$$

$$SS_E = SS_T - SS_A - SS_B = 35\ 226.463 - 33\ 054.50 - 1\ 279.47 = 892.4930$$

按表 6-6 中计算式计算检验统计量, 将其列于表 6-8 的方差分析表中。

表 6-8 随机区组设计的方差分析表

变异来源	平方和 SS	自由度 $\nu$	均方 MS	F 值
总变异	35 226.4630	29		
处理组间	33 078.7980	2	16 539.3990	341.92
区组间	1 276.9630	9	141.8848	2.93
误差	870.7020	18	48.3723	

(3) 确定  $P$  值, 做出推断结论

对于 A 因素, 分子自由度  $\nu_A = 2$ , 分母自由度  $\nu_E = 18$ , 查  $F$  界值表 (方差分析用),  $F_{0.05(2,18)} = 3.55$ 。由于  $F = 341.92$ ,  $F > F_{0.05(2,18)}$ , 故  $P < 0.05$ , 差别有统计学意义, 按照  $\alpha = 0.05$  的显著性水准, 拒绝  $H_{0(A)}$ , 可认为三组大鼠 MT 含量的总体均值不全相同, 不同处理对大鼠 MT 含量有影响。

对于因素 B, 分子自由度  $\nu_B = 9$ , 分母自由度  $\nu_E = 18$ , 查  $F$  界值表 (方差分析用),  $F_{0.05(9,18)} = 2.46$ 。由于  $F = 2.93$ ,  $F > F_{0.05(9,18)}$ , 故  $P < 0.05$ , 按照  $\alpha = 0.05$  的显著性水准, 拒绝  $H_{0(B)}$ , 可看出不同窝别的大鼠 MT 含量的总体均值不同。

### 第三节 多个样本均数的两两比较

完成方差分析后, 我们要么不拒绝  $H_0$ , 推断  $\mu_1 = \mu_2 = \dots = \mu_k$ , 要么拒绝  $H_0$  而接受  $H_1$ , 推断  $k$  个均数不全相同。若不拒绝  $H_0$ , 则整个分析便可结束。然而, 若接受  $H_1$ , 说





明某些组间总体均数不同, 那么就需要进一步作多重比较 (multiple comparison), 以明确哪些组间有差异。

两两比较能否采用  $t$  检验呢? 如果采用  $t$  检验对  $k$  个均数进行两两比较, 共需比较  $\binom{k}{2} = \frac{k!}{2!(k-2)!}$  次, 如  $k=5$ , 需进行 10 次比较, 假设每次比较所确定的检验水准  $\alpha=0.05$ , 则每次检验拒绝  $H_0$  不犯 I 类错误的概率为  $1-0.05=0.95$ , 那么 10 次检验均不犯 I 类错误的概率为  $0.95^{10}=0.5987$ , 而累积犯 I 类错误的概率为  $1-0.5987=0.4013$ 。显然, 采用前述的  $t$  检验进行两两比较是不合理的, 这增大了 I 类错误的概率。

两两比较的方法很多, SNK (Student-Newman-Keuls) 法是较常用的方法之一, 其检验统计量为  $q$ , 故又称为  $q$ -检验 Newman-Keuls test。其计算公式为:

$$q = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{MS_E}{2} \left( \frac{1}{n_A} + \frac{1}{n_B} \right)}} \quad (6-8)$$

式中  $\bar{X}_A$ 、 $\bar{X}_B$  分别为任意两个对比组的样本均数; 分母为两组均值差值的样本标准误, 其中  $n_A$ 、 $n_B$  为相应的两个对比组的样本例数,  $MS_E$  为误差均方。下面用实例说明 SNK 法两两比较的具体方法。

**例 6.3** 对例 6.1 中三组大鼠 GSH 值的均数作两两比较。

1. 提出检验假设, 确定检验水准

$H_0: \mu_A = \mu_B$ , 即对比组 GSH 值的总体均数相等

$H_1: \mu_A \neq \mu_B$ , 即对比组 GSH 值的总体均数不相等

$\alpha=0.05$

2. 计算检验统计量  $q$  值

首先将三个样本均数由大到小排序, 并编组:

组别	甲	乙	丙
均数	83.15	75.63	52.27
组次	1	2	3

三个样本均数共需做  $\binom{3}{2} = \frac{3!}{2!(3-2)!} = 3$  次两两比较。

第 1 组与第 3 组比较:

$$MS_E = 130.5068, \bar{X}_1 = 83.15, \bar{X}_3 = 52.27, n_1 = 12, n_3 = 12$$

$$q_{13} = \frac{\bar{X}_1 - \bar{X}_3}{\sqrt{\frac{MS_E}{2} \left( \frac{1}{n_1} + \frac{1}{n_3} \right)}} = \frac{83.15 - 52.27}{\sqrt{\frac{130.5068}{2} \left( \frac{1}{12} + \frac{1}{12} \right)}} = 9.36$$

其余类推, 可以得到第 1 组与第 2 组、第 2 组与第 3 组比较的  $q$  值, 将所有计算结果列于表 6-9。

表 6-9 例 6.3 的 SNK 法两两比较计算表

对比组 A 与 B	$\bar{X}_A - \bar{X}_B$	$q$ 值	组数 $a$	$q_{0.05}$ 界值	$P$ 值
1 与 3	30.88	9.36	3	3.49	<0.05
1 与 2	7.52	2.28	2	2.89	>0.05
2 与 3	23.36	7.08	2	2.89	<0.05

3. 确定  $P$  值, 做出推断结论

以误差 (组内) 自由度  $\nu_E$  和  $a$  查  $q$  界值表 (附表 5), 其中  $a$  为对比组内所包含的组



数, 如第 1 组与第 3 组比较,  $\alpha=3$ ; 第 1 组与第 2 组比较,  $\alpha=2$ ; 第 2 组与第 3 组比较时,  $\alpha=2$ 。本例  $\nu_E=33$ , 因  $q$  界值表中无  $\nu=33$ , 取  $\nu=30$ ,  $q_{0.05(30,2)}=2.89$ ,  $q_{0.05(30,3)}=3.49$ 。

由表 6-9 可知, 甲组与丙组、乙组与丙组之间相比较时, 拒绝  $H_0$ , 大鼠 GSH 值均数的差别具有统计学意义, 而甲组与乙组之间差别不显著。

## 第四节 2×2 析因设计的方差分析

析因设计 (factorial design) 是一种多因素多水平交叉组合的实验设计方法。在医学研究中, 如果涉及两个或多个处理因素, 而研究者希望了解各处理因素的效应以及因素间的交互作用时, 则可以采用析因设计方法。这里只介绍具有 2 个处理因素、每个因素有 2 个水平即所谓的 2×2 析因设计的方差分析。

设因素 A 的 2 个水平分别用  $A_1$ 、 $A_2$  表示, 因素 B 的 2 个水平分别用  $B_1$ 、 $B_2$  表示, 共有四种不同的组合。现在每种组合下各进行了  $r$  次观测,  $X_{ijk}$  表示在  $A_i B_j$  条件下的第  $k$  次观测数据, 其数据结构见表 6-10。

表 6-10 2×2 析因设计的数据格式

A 因素	B 因素	
	$B_1$	$B_2$
$A_1$	$X_{111}, X_{112}, \dots, X_{11r}$	$X_{121}, X_{122}, \dots, X_{12r}$
$A_2$	$X_{211}, X_{212}, \dots, X_{21r}$	$X_{221}, X_{222}, \dots, X_{22r}$

2×2 析因设计方差分析的目的在于考察 A、B 两因素主效应及两因素间的交互作用。什么是“主效应”与“交互作用”, 以及如何进行方差分析, 通过下面实例加以说明。

**例 6.4** 在一项动物实验研究中, 40 只 5 周龄雌性 BALB/c 小鼠被随机分为 4 组, 接受不同的处理, 即  $A_1 B_1$ 、 $A_1 B_2$ 、 $A_2 B_1$ 、 $A_2 B_2$ , 观测小鼠 BALF 中 IL-4 (pg/ml) 的水平, 观测结果见表 6-11。

表 6-11 小鼠 BALF 中 IL-4 值 (pg/ml)

A 因素	B 因素							
	$B_1$				$B_2$			
$A_1$	37.43	35.04	32.02	33.30	37.94	39.21	39.58	39.77
	36.83	36.88	30.58	38.19	33.63	32.62	43.50	27.41
	36.32	36.60			39.30	46.43		
	30.57	36.27	37.76	30.63	45.11	54.15	51.61	52.45
$A_2$	35.95	33.95	29.69	36.06	46.77	47.66	54.99	50.65
	32.44	34.69			42.71	48.60		

为便于说明, 将表 6-11 数据的均数整理成表 6-12。

表 6-12 例 6.4 资料 IL-4 值均数整理表

A 因素	B 因素		均数	$B_2 - B_1$
	$B_1$	$B_2$		
$A_1$	35.32	37.94	36.63	2.62
$A_2$	33.80	49.47	41.64	15.67
均数	34.56	43.71	39.14	9.15
$A_2 - A_1$	-1.52	11.53	5.01	



1. 单独效应 (simple effect) 指其他因素水平固定在一个水平时, 某一因素不同水平之间均数的差别。如当 A 因素固定在  $A_1$  水平时, B 因素的单独效为 2.62; 当 A 因素固定在  $A_2$  水平时, B 因素的单独效为 15.67。同理, B 因素固定在  $B_1$  水平时, A 因素的单独效应为 -1.52; B 因素固定在  $B_2$  水平时, A 因素的单独效应为 11.53。

2. 主效应 (main effect) 指某一因素各个水平间的平均差别。如 A 因素的主效应记为  $X_A$ , 则  $X_A = 41.64 - 36.63 = 5.01$ ; 同理, B 因素的主效应  $X_B = 43.71 - 34.56 = 9.15$ 。

3. 交互效应 (interaction effect) 指如果一个因素的单独效应随另一因素的水平变化而变化, 而且其变化幅度不能用随机误差解释时, 则称这两个因素间存在交互效应。如对于  $B_1$  水平, A 因素的单独效应为 -1.52; 而对于  $B_2$  水平, A 因素的单独效应为 11.53, 这说明 A 因素与 B 因素的效应互不独立, A 因素单独效应依赖于 B 因素水平的选择, 即 A 因素与 B 因素之间存在交互效应, 记为  $X_{AB}$ 。

4. 方差分析 按照方差分析基本思想, 析因设计中, 观测值总变异  $SS_T$  可分解为以下四个部分:

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E \quad (6-9)$$

相应的自由度:

$$\nu_T = \nu_A + \nu_B + \nu_{AB} + \nu_E \quad (6-10)$$

有关计算式归纳于表 6-13 的方差分析表中。其中,  $\bar{X}_i$  为 A 因素  $i$  水平均数,  $\bar{X}_j$  为 B 因素  $j$  水平均数,  $\bar{X}_{ij}$ 、 $S_{ij}^2$  分别为 A 因素  $i$  水平、B 因素  $j$  水平组合下的均数和方差。

表 6-13  $2 \times 2$  析因设计方差分析表

变异来源	平方和 SS	自由度 $\nu$	均方 MS	F 值
总变异	$SS_T = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^r (X_{ijk} - \bar{X})^2$ $= (2 \times 2 \times r - 1) S^2$	$\nu_T = 2 \times 2 \times r - 1$		
A 因素	$SS_A = \sum_{i=1}^2 2r (\bar{X}_i - \bar{X})^2$	$\nu_A = 2 - 1$	$MS_A = SS_A / \nu_A$	$F_A = \frac{MS_A}{MS_E}$
B 因素	$SS_B = \sum_{j=1}^2 2r (\bar{X}_j - \bar{X})^2$	$\nu_B = 2 - 1$	$MS_B = SS_B / \nu_B$	$F_B = \frac{MS_B}{MS_E}$
AB 因素	$SS_{AB} = \sum_{i=1}^2 \sum_{j=1}^2 r (\bar{X}_{ij} - \bar{X})^2 -$ $SS_A - SS_B$	$\nu_{AB} = (2-1)(2-1)$	$MS_{AB} = SS_{AB} / \nu_{AB}$	$F_{AB} = \frac{MS_{AB}}{MS_E}$
误差	$SS_E = SS_T - SS_A - SS_B - SS_{AB}$	$\nu_E = 2 \times 2 \times (r-1)$	$MS_E = SS_E / \nu_E$	

本例  $2 \times 2$  析因设计的方差分析步骤如下:

(1) 提出检验假设, 确定检验水准

$H_{0(A)}: \mu_1 = \mu_2$ , 即 A 因素两个不同水平的 IL-4 测量结果相同

$H_{1(A)}: \mu_1 \neq \mu_2$ , 即 A 因素两个不同水平的 IL-4 测量结果不同

$H_{0(B)}: \tau_1 = \tau_2$ , 即 B 因素两个不同水平的 IL-4 测量结果相同

$H_{1(B)}: \tau_1 \neq \tau_2$ , 即 B 因素两个不同水平的 IL-4 测量结果不同

$H_{0(AB)}: A$  与  $B$  无交互效应

$H_{1(AB)}: A$  与  $B$  存在交互效应

$\alpha = 0.05$

(2) 计算检验统计量  $F$  值

按表 6-13 中的公式计算检验统计量。为简便, 通常使用统计软件 (如 SPSS) 进行计



算, 现将计算结果列于表 6-14 的方差分析表中。

表 6-14  $2 \times 2$  析因设计方差分析表

变异来源	平方和 SS	自由度 $\nu$	均方 MS	F 值	P 值
总变异	2 055. 1779	39			
A 因素	250. 6504	1	250. 6504	16. 63	<0. 05
B 因素	836. 2188	1	836. 2188	55. 48	<0. 05
AB 因素	425. 6910	1	425. 6910	28. 24	<0. 05
误差	542. 6177	36	15. 0727		

### (3) 确定 P 值, 作出推断结论

查 F 界值表, 对于 A 因素  $F_{0.05(1,36)} = 4.11$ ,  $F_A > F_{0.05(1,36)}$ ,  $P < 0.05$ , 按照  $\alpha = 0.05$  的检验水准拒绝  $H_{0(A)}$ , 认为  $A_1$  组与  $A_2$  组 IL-4 值总体均数不相等。同理, 对于 B 因素  $F_B > F_{0.05(1,36)}$ ,  $P < 0.05$ , 拒绝  $H_{0(B)}$ ,  $B_1$  组与  $B_2$  组 IL-4 值总体均数不相等; 对于 AB 交互作用,  $F_{AB} > F_{0.05(1,36)}$ ,  $P < 0.05$ , 拒绝  $H_{0(AB)}$ , A 与 B 两处理因素的交互作用显著。

## 第五节 方差齐性检验

方差分析有其应用条件, 理论上要求各样本相互独立, 服从正态分布且方差齐同 (homogeneity of variance)。相对而言, 方差是否齐同对检验的准确性影响更大些。对于多个方差的齐性检验, 广为应用的是 Bartlett 检验法和 Levene 检验法。Bartlett 检验法主要适用于正态分布资料的方差齐性检验问题, 如资料不服从正态分布, 则可采用 Levene 检验法。以下仅介绍 Bartlett 检验法, 对于 Levene 检验法, 有兴趣的读者可参考有关书籍。

对于 Bartlett 检验法, 其检验统计量为:

$$\chi^2 = \frac{Q_1}{Q_2}, \quad \nu = k - 1 \quad (6-11)$$

其中:

$$Q_1 = \sum_{i=1}^k (n_i - 1) \ln(S_i^2 / S^2) \quad (6-12)$$

$$Q_2 = 1 + \frac{1}{3(k-1)} \left( \sum_{i=1}^k \frac{1}{n_i - 1} - \frac{1}{n - k} \right) \quad (6-13)$$

式中,  $S_i^2$  为第  $i$  组的方差,  $S^2$  为合并方差 (对完全随机设计资料有  $S^2 = MS_E$ ),  $k$  为比较组数,  $n_i$  为第  $i$  组样本例数,  $n$  为总例数。

当样本来自独立正态总体时, 在  $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$  为真情况下, 该检验统计量服从  $\nu = k - 1$  的  $\chi^2$  分布。样本方差  $S_i^2$  差异越大,  $Q_1$  越大,  $\chi^2$  值亦越大。如果  $\chi^2 \geq \chi_{\alpha, \nu}^2$ , 则  $P \leq \alpha$ , 拒绝  $H_0$ , 可认为方差不齐; 反之, 若  $\chi^2 < \chi_{\alpha, \nu}^2$ , 则  $P > \alpha$ , 不拒绝  $H_0$ , 尚不能认为方差不齐。

**例 6.5** 对例 6.1 中三组资料作方差齐性检验。

(1) 提出检验假设, 确定检验水准

$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$ , 即三个总体方差相等

$H_1: \sigma_1^2, \sigma_2^2, \sigma_3^2$  不全相等

$\alpha = 0.05$

(2) 计算检验统计量  $\chi^2$  值



本例:

$$k=3, \quad n_1=12, \quad n_2=12, \quad n_3=12, \quad n=36$$

$$S_1=12.30, \quad S_2=11.07, \quad S_3=10.85, \quad MS_E=130.5068$$

$$\begin{aligned} Q_1 &= \sum_{i=1}^k (n_i-1) \ln(S_i^2/S_i^2) \\ &= (12-1) \ln \frac{130.5068}{12.30^2} + (12-1) \ln \frac{130.5068}{11.07^2} + (12-1) \ln \frac{130.5068}{10.85^2} \\ &= 0.2010 \end{aligned}$$

$$\begin{aligned} Q_2 &= 1 + \frac{1}{3(k-1)} \left( \sum_{i=1}^k \frac{1}{n_i-1} - \frac{1}{n-k} \right) \\ &= 1 + \frac{1}{3(3-1)} \left( \frac{1}{12-1} + \frac{1}{12-1} + \frac{1}{12-1} - \frac{1}{36-3} \right) \\ &= 1.0404 \end{aligned}$$

$$\chi^2 = \frac{Q_1}{Q_2} = \frac{0.2010}{1.0404} = 0.19, \quad \nu = 3-1=2$$

(3) 确定  $P$  值, 做出推断结论

自由度  $\nu=2$ , 查  $\chi^2$  界值表,  $\chi_{0.05,2}^2=5.99$ 。由于  $\chi^2=0.19$ ,  $\chi^2 < \chi_{0.05,2}^2$ , 故  $P > 0.05$ , 按照  $\alpha=0.05$  的显著性水准, 不拒绝  $H_0$ , 尚不能认为三个总体方差不齐。

## 练习 题

### 一、单项选择题

- 方差分析的基本思想和要点是
  - 组间均方大于组内均方
  - 组内均方大于组间均方
  - 不同来源的方差必须相等
  - 两方差之比服从  $F$  分布
  - 总变异及其自由度可按不同来源分解
- 方差分析的应用条件之一是方差齐性, 它是指
  - 各比较组相应的样本方差相等
  - 各比较组相应的总体方差相等
  - 组内方差=组间方差
  - 总方差=各组方差之和
  - 总方差=组内方差+组间方差
- 完全随机设计方差分析中的组间均方反映的是
  - 随机测量误差大小
  - 某因素效应大小
  - 处理因素效应与随机误差综合结果
  - 全部数据的离散度
  - 各组方差的平均水平
- 对于两组资料的比较, 方差分析与  $t$  检验的关系是
  - $t$  检验结果更准确
  - 方差分析结果更准确
  - $t$  检验对数据的要求更为严格
  - 近似等价
  - 完全等价
- 多组均数比较的方差分析, 如果  $P < 0.05$ , 则应该进一步做的是
  - 两均数的  $t$  检验
  - 区组方差分析
  - 方差齐性检验
  - $q$  检验
  - 确定单独效应

### 二、计算与分析

- 在评价某药物耐受性及安全性的 I 期临床试验中, 对符合纳入标准的 40 名健康自



愿者随机分为 4 组，每组 10 名，各组注射剂量分别为 0.5U、1U、2U、3U，观察 48 小时后部分凝血活酶时间 (S)。试比较任意两两剂量间的部分凝血活酶时间有无差别？

各剂量组 48 小时部分凝血活酶时间 (s)

0.5 U	1 U	2 U	3 U
36.8	40.0	32.9	33.0
34.4	35.5	37.9	30.7
34.3	36.7	30.5	35.3
35.7	39.3	31.1	32.3
33.2	40.1	34.7	37.4
31.1	36.8	37.6	39.1
34.3	33.4	40.2	33.5
29.8	38.3	38.1	36.6
35.4	38.4	32.4	32.0
31.2	39.8	35.6	33.8

2. 为探讨小剂量地塞米松对急性肺损伤动物模型肺脏的保护作用，将 36 只二级 SD 大鼠按性别、体重配成 12 个配伍组，每一配伍组的 3 只大鼠被随机分配到对照组、损伤组与激素组，实验 24 小时后测量支气管肺泡灌洗液总蛋白水平 (g/L)，结果如下表。问 3 组大鼠的总蛋白水平是否相同？

3 组大鼠总蛋白水平 (g/L)

配伍组	对照组	损伤组	激素组
1	0.36	1.48	0.30
2	0.28	1.42	0.32
3	0.26	1.33	0.29
4	0.25	1.48	0.16
5	0.36	1.26	0.35
6	0.31	1.53	0.43
7	0.33	1.40	0.31
8	0.28	1.30	0.13
9	0.35	1.58	0.33
10	0.41	1.24	0.32
11	0.49	1.47	0.26
12	0.27	1.32	0.26

3. 为研究喹啉酸对大鼠急性和亚急性炎症的影响，将 40 只体重为  $200 \pm 20(g)$  的雄性 Wistar 大鼠建立急性和亚急性炎症动物模型，然后随机分为 4 组，每组 10 只，给予不同的处理，观察其 WBC 值。4 种处理分别为：腹腔注射生理盐水后 3 小时处死、腹腔注射生理盐水后 6 小时处死、腹腔注射喹啉酸 (0.35mg/g) 后 3 小时处死，腹腔注射喹啉酸 (0.35mg/g) 后 6 小时处死。实验结果如下表。问喹啉酸与给药距处死的时间间隔 (简称时间) 对 WBC 值是否有影响？

不同药物与不同时间大鼠 WBC 值 ( $10^3$ )

时 间	药 物			
	生理盐水		喹啉酸	
3h	21.3	18.8	15.8	11.0
	21.9	13.5	8.7	12.8
	11.1	22.6	9.4	12.5
	16.3	17.1	5.3	9.3
	17.9	14.6	8.3	11.0
6h	19.0	23.0	13.9	19.0
	25.2	22.8	15.8	15.3
	22.9	17.8	18.3	19.2
	19.8	24.6	13.0	18.2
	22.7	25.3	14.0	17.3

(钟晓妮)

### Analysis of Variance

The responses that are generated in an experiment situation always exhibits a certain amount of variability. In an analysis of variance, we divide the total variation in the response measurements into portions that may be attributed to various factors of interest to the experimenter. If the experiment has been properly designed, these portions can be used to answer questions about the effects of the various factors on the response of interest. Among all the type of designs in the analysis of variance, one way that classification completely randomized design is popularly used. The model assumptions:

- The  $k$  samples represent independent random samples drawn from  $k$  specific populations with means  $\mu_1, \mu_2, \dots, \mu_k$ , where  $\mu_1, \mu_2, \dots, \mu_k$ , are unknown constants.
- Each of the  $k$  populations is normal.
- Each of the  $k$  population has the same variance  $\sigma^2$ .

These assumptions look very much like those used for the unpaired  $t$  test in chapter 5. If the null hypothesis is written  $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ , why not test this hypothesis a series of  $t$  tests?

Considering a three samples experiment like example 6.1, We would compare  $\bar{X}_1$  to  $\bar{X}_2$ ,  $\bar{X}_1$  to  $\bar{X}_3$ ,  $\bar{X}_2$  to  $\bar{X}_3$  in three separate  $t$  tests. To see why, note that with an alpha level at 0.05, we will correctly accept a true  $H_0$ , 95% of the time. For a set of three independent comparisons ( $H_0$ 's) all of which are true, the probability of correctly accepting all of them is  $(0.95)^3 = 0.86$ , so the probability of a Type I error (rejecting at least one true  $H_0$ ) becomes  $1 - 0.86 = 0.14$ , which is unacceptably high.

## 第七章 相对数及其应用

在临床研究的统计资料中,除了前述的计量资料外,还有阴性和阳性、有效和无效、治愈和未治愈、死亡与未死亡及各种疾病分类等类型的资料。这些资料的整理是先将研究对象按其性质或特征分类,再分别计数每一类的例数。对于这些资料常用率、构成比、相对比等指标来进行统计描述。

### 第一节 相对数的概念及计算

相对数 (relative number) 是两个有关的绝对数之比,也可以是两个统计指标之比。计算相对数的意义主要是把基数化作相等,便于相互比较。例如,某病用 A 法治疗 100 人,有效 80 人, B 法治疗 150 人,有效 100 人,若仅比较两组有效的绝对人数是不能得出正确结论的,而通过计算  $\frac{80}{100} \times 100\% = 80.0\%$  与  $\frac{100}{150} \times 100\% = 66.7\%$ ,可以很容易看出该样本 A 法优于 B 法。相对数常用的指标有率、构成比和相对比。

#### 一、率

率 (rate) 表示在一定范围内某现象的发生数与可能发生的总数之比,说明某现象出现的强度或频度,通常以百分率 (%),千分率 (‰),万分率 (/万),十万分率 (/10 万) 等表示。计算公式为:

$$\text{率} = \frac{\text{某事物或现象发生的实际数}}{\text{某事物或现象发生的所有可能数}} \times \text{比例基数} \quad (7-1)$$

公式中的“比例基数”通常依据习惯而定,治愈率、感染率用百分率,出生率、死亡率用千分率,某些疾病的死亡率用十万分率。总体率用  $\pi$  来表示,样本率用  $P$  来表示。

#### 二、构成比

构成比 (constituent ratio) 表示某事物内部各组成部分在整体中所占的比重,常以百分数表示,计算公式如下:

$$\text{构成比} = \frac{\text{某一组成部分的观察单位数}}{\text{同一事物各组成部分的观察单位总数}} \times 100\% \quad (7-2)$$

例 7.1 某医院某月各科室住院病人数及死亡人数如表 7-1 所示,其中第 (4) 列为构成比,是由第 (3) 列数据计算而得。第 (5) 列为率,是由第 (2) 与第 (3) 列数据计算而得。

表 7-1 某医院某月各科室住院病人数及死亡人数

科室(1)	病人数(2)	因病死亡人数(3)	死亡构成(%)(4)	病死率(‰)(5)
内 科	350	25	30.86	71.43
外 科	650	30	37.04	46.15
肿瘤科	120	20	24.69	166.67
妇产科	300	5	6.17	16.67
皮肤科	56	0	0.00	0.00
眼 科	45	0	0.00	0.00
小儿科	100	1	1.24	10.00
合计	1 621	81	100.00	49.97





由表 7-1 中的第 (4) 与第 (5) 列可看出, 构成比和率虽然同是相对数, 但却是两种不同的概念, 应用的场合不同, 应特别注意。构成比各构成部分的相对数之和应为 100%; 某一构成部分的增减会影响其他部分构成比相应的减少或增加。而某一部分率的变化并不影响其他部分率的变化, 且其平均率不能简单地将各率相加后平均求得。

### 三、相对比

相对比 (relative ratio) 是 A、B 两个有关联指标之比, 用以描述两者的对比水平, 其计算公式为:

$$\text{相对比} = \frac{\text{A 指标}}{\text{B 指标}} \quad (7-3)$$

1. 两类别例数之比 如我国 1982 年人口普查的男性人口数为 519 433 369, 女性人口数为 488 741 919 人, 则男女比例数为:

$$\frac{519\,433\,369}{488\,741\,919} = 1.063$$

即男女比例为 1.063 : 1。

2. 相对危险度 (relative risk, 简称为 RR) 是流行病学中常用的指标, 表示在两种不同条件下某疾病发生的概率之比。其计算公式为:

$$RR = \frac{P_1}{P_0} \quad (7-4)$$

**例 7.2** 某地某年龄组男性吸烟和非吸烟的冠心病死亡资料如表 7-2 所示, 试计算其相对危险度。

表 7-2 某地某年龄组男性吸烟和非吸烟的冠心病死亡资料

分组	死亡数	观察人年数	死亡率(1/10 万人年)
吸烟组	104	43 248	240.5
非吸烟组	12	10 673	112.4

$$RR = \frac{240.5}{112.4} = 2.140$$

说明该地男性吸烟者的冠心病死亡率是不吸烟者的冠心病死亡率的 2.140 倍。

3. 比数比 (odds ratio, 简记为 OR) 也称优势比, 常用于流行病学中病例-对照研究资料。其计算公式为:

$$OR = \frac{P_1 / (1 - P_1)}{P_0 / (1 - P_0)} \quad (7-5)$$

**例 7.3** 服用反应停与新生儿肢体缺陷关系的病例对照研究的资料如表 7-3 所示。试计算服用反应停与不服用反应停引起新生儿肢体缺陷的优势比。

表 7-3 服用反应停与肢体缺陷关系病例对照研究

服用反应停	畸形儿组	对照组	合计
使用	34(a)	2(b)	36
未使用	16(c)	88(d)	104
合计	50	90	140

由公式 (7-5) 可以得出:

$$OR = \frac{\text{疾病组的暴露比数}}{\text{对照组的暴露比数}} = \frac{a/c}{b/d} = \frac{ad}{bc} \quad (7-6)$$

本例:



$$OR = \frac{34 \times 88}{2 \times 16} = 93.5$$

## 第二节 相对数使用应注意的问题

1. 不要把构成比与率相混淆 构成比只能说明某事物内部各组成部分的比重和分布,不能说明该事物某一部分发生的强度与频率。如表 7-1, 从第(4)列的构成可看出外科的死亡占全院死亡的比例最大, 但从第(5)列各科的病死率来看, 肿瘤科的病死率最高, 它反映了各科的死亡危险性的大小。

2. 使用相对数时分母不宜过小 分母过小时相对数不稳定。在观察例数较小时, 最好直接用绝对数表示, 如“5 例中 4 例有效”, 而不要写有效率为 80%, 以免引起误解。

3. 注意资料的可比性 用以比较的资料应是同质的, 除了要比较的处理因素外, 其他条件应基本相同。对于不同时期、地区、条件下的资料比较时应注意是否具有可比性。

4. 要考虑存在抽样误差 比较两个样本率或构成比时, 由于存在抽样误差, 不能单凭数字表面相差的大小作结论, 对总体进行推断时应作统计学检验。

## 第三节 率的标准化

### 一、标准化的意义

率的标准化 (standardization) 是为了在比较两个不同人群的患病率、发病率、死亡率等资料时, 消除其内部构成 (如年龄、性别、工龄、病程长短等) 的影响。如年龄会影响死亡率, 年龄越大, 死亡率可能越大; 病情越重, 越难治愈, 治愈率就越低。

例 7.4 某病两种疗法治愈率比较的资料如表 7-4 所示:

表 7-4 某病两种疗法治愈率的比较

病 型	西医疗法组			中西医结合疗法组		
	病例数	治愈数	治愈率(%)	病例数	治愈数	治愈率(%)
普通型	60	36	60	20	13	65
重 型	20	8	40	60	27	45
暴发型	20	4	20	20	5	25
合 计	100	48	48	100	45	45

从表 7-4 可以看出内部构成对总治愈率的影响。就任一种病型来看, 西医疗法均不如中西医结合疗法, 但从总治愈率看, 中西医结合疗法反而低于西医疗法, 这种偏差源于内部构成的不同。西医疗法组的普通型多, 中西医结合疗法组的重型多。西医疗法的总治愈人数因病情较轻而增多, 结果是不合理的。为了正确比较两种疗法的疗效, 统计上常用标准化的方法来消除内部构成的影响。其方法有直接法、间接法和反推法三种, 本节介绍常用的直接法。

### 二、标准化直接法的计算步骤

#### (一) 标准化直接法的计算方法

$$P' = \frac{N_1 P_1 + N_2 P_2 + \dots + N_k P_k}{N} = \frac{\sum N_i P_i}{N} \quad (7-7)$$



式中  $P'$  为标准率,  $N_1, N_2, \dots, N_k$  为某一影响因素 (如病型、年龄等) 标准构成的每层例数,  $P_1, P_2, \dots, P_k$  为原始数据中各层的率,  $N$  为标准构成的总例数。上式也可写成:

$$P' = C_1 P_1 + C_2 P_2 + \dots + C_k P_k = \sum C_i P_i \quad (7-8)$$

式中  $C_i = N_i/N$  为标准构成的构成比。

## (二) 标准构成的选取

1. 选取有代表性的、较稳定的、数量较大的人群构成为标准, 如全国范围或全省范围的数据作为标准构成。国际间比较时需要采用世界通用的标准。

2. 选择相互比较的各组例数合计作为标准构成。

3. 从比较的各组中任选其一作为标准构成。

对表 7-4 的计算采用第二种方法, 取比较各组的各层例数的合计作为标准构成, 即将两组同病型的人数之和作为标准构成, 根据两组的原各层的治愈率, 计算两组各层的预期治愈数, 最后得到两组标准化后的总治愈率, 其计算如表 7-5 所示:

表 7-5 消除构成影响后两种疗法治愈率的比较

病型	标准构成人数	西医疗法组		中西医结合疗法组	
		原治愈率(%)	预期治愈数	原治愈率(%)	预期治愈数
普通型	80	60	48	65	52
重型	80	40	32	45	36
暴发型	40	20	8	25	10
合计	200	48	88	45	98

西医疗法组标准化后的总治愈率为:

$$P' = \frac{88}{200} \times 100\% = 44\%$$

中西医结合疗法组标准化后的总治愈率为:

$$P' = \frac{98}{200} \times 100\% = 49\%$$

可见, 中西医结合疗法组标准化后的总治愈率高于西医疗法组标准化后的总治愈率。

**例 7.5** 表 7-6 的数据给出了甲、乙两县的食管癌死亡率, 试计算标准化死亡率并进行比较。

表 7-6 甲、乙两县各年龄组的食管癌死亡率 (1/10 万)

年龄组(岁)	甲 县			乙 县		
	人口数	死亡数	死亡率	人口数	死亡数	死亡率
0~	378 977	2	0.5	282 762	1	0.4
30~	63 436	11	17.3	39 443	4	10.1
40~	54 910	55	100.2	40 488	29	71.6
50~	41 970	151	359.8	33 309	99	297.2
60~	25 060	163	650.4	23 167	122	526.6
70~	10 780	70	649.4	14 548	98	673.6
合计	575 133	452	78.6	433 717	353	81.4

由表 7-6 的数据可看出, 各年龄段甲县的死亡率大多数都高于乙县, 但总的死亡率却是乙县高于甲县, 其原因是两县人口构成的不同, 因此要采用标准率进行比较。表 7-7 给出了某地的标准人口及构成比。



表 7-7 某地的人口各年龄组人口数及构成

年龄组(岁)	人口数	构成比
0~	3 970 027	0.6452
30~	570 014	0.0926
40~	548 990	0.0892
50~	496 011	0.0806
60~	354 421	0.0576
70~	213 529	0.0347
合计	6 152 992	1.0000

表 7-8 给出了以某地的标准人口为标准进行标准化计算的结果。

表 7-8 以某地的人口为标准的标准化计算表

年龄组(岁)	标准人口数	甲 县		乙 县	
		原死亡率(3)	预期死亡数 (4)=(2)×(3)	原死亡率	预期死亡数 (6)=(2)×(5)
0~	3 970 027	0.5	20	0.4	16
30~	570 014	17.3	99	10.1	58
40~	548 990	100.2	550	71.6	393
50~	496 011	359.8	1 785	297.2	1 474
60~	354 421	650.4	2 305	526.6	1 866
70~	213 529	649.4	1 387	673.6	1 438
合计	6 152 992		6 146		5 245

甲县标化后的总死亡率为:

$$P'_{\text{甲}} = \frac{6\ 146}{6\ 152\ 992} \times 100\ 000/10\ \text{万} = 99.9\ (1/10\ \text{万})$$

乙县标化后的总死亡率为:

$$P'_{\text{乙}} = \frac{5\ 245}{6\ 152\ 992} \times 100\ 000/10\ \text{万} = 85.2\ (1/10\ \text{万})$$

从标准化死亡率看,甲县的食管癌死亡率高于乙县。

## 第四节 医学中常用的相对数指标

### 一、死亡统计指标

1. 死亡率 (death rate) 表示某地某年每 1 000 人中的死亡人数, 计算公式为:

$$\text{死亡率} = \frac{\text{某年死亡人口总数}}{\text{同年年平均人口数}} \times 1\ 000\%$$

2. 年龄别死亡率 (age-specific death rate) 表示某地某年龄组每 1 000 人口中的死亡数, 计算公式为:

$$\text{年龄别死亡率} = \frac{\text{某年某地某年龄组死亡人数}}{\text{同年龄别平均人口数}} \times 1\ 000\%$$

3. 死因别死亡率 (cause specific death rate) 表示某年每 10 万人中由于患有某种疾病死亡的人数, 计算公式为:

$$\text{某病死亡率} = \frac{\text{某年某地某病死亡人数}}{\text{同年该地平均人口数}} \times 100\ 000/10\ \text{万}$$



4. 死因构成 (proportion of dying of a specific cause) 也称相对死亡比, 是某种死因引起的死亡人数占总死亡数的百分比, 计算公式为:

$$\text{某种死因的构成比} = \frac{\text{因某种死因死亡的人数}}{\text{总死亡人数}} \times 100\%$$

## 二、疾病统计的指标

1. 发病率 (incidence rate) 表示在某一时期内特定人群中患某病新病例的频率, 计算公式为:

$$\text{某病发病率} = \frac{\text{某时期某病新病例数}}{\text{同期内平均人口数}} \times \text{比例基数}$$

2. 患病率 (prevalence rate) 也称现患率, 表示某一时刻某人群中患某病的频率, 计算公式为:

$$\text{某病患病率} = \frac{\text{某地某时点某病患者例数}}{\text{该地同期内平均人口数}} \times \text{比例基数}$$

以上比例基数可为 100%、1 000‰、10 000/万、100 000/10 万, 应根据临床医学或流行病学的专业要求决定。

3. 病死率 (cause fatality) 表示某期间内, 某病患者中因某病死亡的频率, 计算公式为:

$$\text{某病病死率} = \frac{\text{某期间因某病死亡人数}}{\text{同期某病的患病人数}} \times 100\%$$

4. 治愈率 (cure rate) 表示接受治疗的病人中治愈的频率, 计算公式为:

$$\text{治愈率} = \frac{\text{治愈病人数}}{\text{接受治疗病人数}} \times 100\%$$

## 第五节 率的抽样误差与区间估计

### 一、率的标准误

如同前面章节学过的样本均数与总体均数存在着抽样误差, 样本率与总体率同样存在着抽样误差。误差的大小可以用率的标准误  $\sigma_p$  来描述, 其计算公式为:

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} \quad (7-9)$$

式中  $\pi$  为总体率,  $n$  为样本例数。

由于实际中总体率  $\pi$  往往未知, 我们常用样本率  $P$  来代替总体率  $\pi$ , 则上述公式可以改写为:

$$S_p = \sqrt{\frac{P(1-P)}{n}} \quad (7-10)$$

**例 7.6** 为了解某药的疗效, 对 100 名患者治疗的结果进行调查, 结果为 80 人有效, 有效率为 80%。则样本率的抽样误差为:

$$S_p = \sqrt{\frac{P(1-P)}{n}} = \sqrt{\frac{0.80 \times (1-0.80)}{100}} = 0.04 = 4\%$$

### 二、率的区间估计

我们可以参照估计总体均数可信区间的概念与方法, 对总体率进行区间估计。具体有两种估计方法。



### (一) 正态近似法

当  $n$  足够大, 且  $nP$  和  $n(1-P)$  均大于 5 时,  $P$  的分布接近正态分布, 可用公式 (7-11) 来求总体率的可信区间:

$$(P - u_{\alpha/2} S_P, P + u_{\alpha/2} S_P) \quad (7-11)$$

式中  $P$  为样本率,  $S_P$  为样本率的标准误,  $u_{\alpha/2}$  是在可信度为  $1-\alpha$  时, 标准正态分布双侧概率为  $\alpha$  的界值, 如  $u_{0.05/2} = 1.96$ 。

如例 7.6, 该药物有效率的 95% 可信区间是:

$$\text{下限: } P - u_{\alpha/2} S_P = 0.80 - 1.96 \times 0.04 = 0.7216$$

$$\text{上限: } P + u_{\alpha/2} S_P = 0.80 + 1.96 \times 0.04 = 0.8784$$

即所求药物有效率的 95% 可信区间是: (72.16%, 87.84%)。

### (二) 查表法

当样本含量较小时, 比如  $n \leq 50$ , 需查附表 6 百分率的可信区间表, 得到总体率的可信区间。

例 7.7 某眼科医生在某校抽查 6 年级学生 30 名, 其中患近视眼的学生 12 名, 求该校 6 年级学生患近视眼率的 95% 可信区间。

查附表 6.2, 在  $n=30$ ,  $X=12$  的纵横交叉处可得到 95% 可信区间值为 23%~59%。

注意: 附表 6 中的  $X$  值只列出了  $X \leq n/2$  部分, 当  $X > n/2$  时, 应以  $n-X$  值查表, 然后用 100 减去查得的数值, 即为所求的区间。如对于上例, 若欲求视力正常率的 95% 可信区间, 若用 30 与  $30-12=18$  来查表, 则查表得不到结果。此时, 可用 100 减去上述所得区间得正常率的 95% 可信区间为: 41%~77%。

## 练 习 题

### 一、单项选择题

- 如果一种新的治疗方法能够使不能治愈的疾病得到缓解并延长生命, 则应发生的情况是
  - 该病患病率增加
  - 该病患病率减少
  - 该病的发病率增加
  - 该病的发病率减少
  - 该疾病的死因构成比增加
- 计算乙肝疫苗接种后血清学检查的阳转率, 分母为
  - 乙肝易感人数
  - 平均人口数
  - 乙肝疫苗接种人数
  - 乙肝患者人数
  - 乙肝疫苗接种后的阳转人数
- 计算标准化死亡率的目的是
  - 减少死亡率估计的偏倚
  - 减少死亡率估计的抽样误差
  - 便于进行不同地区死亡率的比较
  - 消除各地区内部构成不同的影响
  - 便于进行不同时间死亡率的比较
- 影响总体率估计的抽样误差大小的因素是
  - 总体率估计的容许误差
  - 样本率估计的容许误差
  - 检验水准和样本含量
  - 检验的把握度和样本含量
  - 总体率和样本含量
- 研究某种新药的降压效果, 对 100 人进行试验, 其显效率的 95% 可信区间为 0.862~0.926, 表示



- A. 样本显效率在 0.862~0.926 之间的概率是 95%
- B. 有 95% 的把握说总体显效率在此范围内波动
- C. 有 95% 的患者显效率在此范围
- D. 样本率估计的抽样误差有 95% 的可能在此范围
- E. 该区间包括总体显效率的可能性为 95%

## 二、计算与分析

1. 某工厂在“职工健康状况报告中”写到：“在 946 名工人中，患慢性病的有 274 人，其中女性 219 人，占 80%，男性 55 人，占 20%。所以女性易患慢性病”，你认为是否正确？为什么？

2. 在“锑剂短程疗法治疗血吸虫病病例的临床分析”一文中，根据下表资料认为“其中 10~岁组死亡率最高，其次为 20~岁组”，问这种说法是否正确？

锑剂治疗血吸虫不同性别死亡者年龄分布

年龄组	男	女	合计
0~	3	3	6
10~	11	7	18
20~	4	6	10
30~	5	3	8
40~	1	2	3
50~	5	1	6
合计	29	22	51

3. 某研究根据以下资料说明沙眼 20 岁患病率最高，年龄大的反而患病率下降，你同意吗？说明理由。

某研究资料沙眼病人的年龄分布

年龄组	沙眼人数	构成比(%)
0~	47	4.6
10~	198	19.3
20~	330	32.1
30~	198	19.3
40~	128	12.4
50~	80	7.8
60~	38	3.7
70~	8	0.8
合计	1 027	100.0

4. 今有两个煤矿的工人尘肺患病率(%)如下表，试比较两个煤矿的工人尘肺总的患病率。

两个煤矿的工人尘肺患病率情况

工龄(年)	甲 矿			乙 矿		
	检查人数	尘肺人数	患病率(%)	检查人数	尘肺人数	患病率(%)
<6	14 026	120	0.86	992	2	0.20
6~	4 285	168	3.92	1 905	8	0.42
10~	2 542	316	12.43	1 014	117	11.54
合计	20 853	604	2.90	3 911	127	3.25

5. 抽样调查了某校 10 岁儿童 200 名的牙齿，患龋 130 人，试求该校儿童患龋率的



95%的区间估计。

(刘启贵)

### Relative Number and Its Application

In this chapter, two basic measures to assess the frequency of health events are introduced. These measures, which play key roles in medicine, epidemiology, and public health, are risk and prevalence.

Risk is a measure of the occurrence of new cases of the disease of interest in population. More precisely, risk is the proportion of unaffected individuals who, on average, will contract the disease of interest over a specified period of time. Risk is estimated by observing a particular population for a defined period of time—the risk period. The estimated risk ( $R$ ) is a proportion; the numerator is the number of newly affected persons ( $A$ ) and denominator is the size ( $N$ ) of the population under observation:

$$R = \frac{\text{Newcases}}{\text{Persons at risk}} = \frac{A}{N}$$

Prevalence indicates the number of existing cases of the disease of interest in a population. Specifically, the point prevalence ( $P$ ) is the proportion of a population that has the disease of interest at a particular time, eg, on a given day. This value is estimated by dividing the number of existing affected individuals ( $C$ ) by the number of persons in the population ( $N$ ):

$$P = \frac{\text{number of existing affected individuals}}{\text{number of persons in the population}} = \frac{C}{N}$$



## 第八章 $\chi^2$ 检验

$\chi^2$  检验 (chi square test) 也称卡方检验, 是英国统计学家 Pearson 提出的一种用途广泛的假设检验方法。该检验以  $\chi^2$  分布 (chi square distribution) 为理论依据, 可以推断两个 (或多个) 总体率以及构成比之间有无差别。

### 第一节 四格表资料的 $\chi^2$ 检验

#### 一、四格表资料的 $\chi^2$ 检验

**例 8.1** 为了解某中药治疗原发性高血压的疗效, 将 70 名高血压患者随机分为两组, 试验组用该药加辅助治疗, 对照组用安慰剂加辅助治疗, 观察结果见表 8-1, 问该药治疗原发性高血压是否有效?

表 8-1 两种疗法治疗原发性高血压的疗效

组别	有效	无效	合计	有效率(%)
对照组	20(25.8) <i>a</i>	24(18.2) <i>b</i>	44( <i>a+b</i> )	45.45
试验组	21(15.2) <i>c</i>	5(10.8) <i>d</i>	26( <i>c+d</i> )	80.77
合计	41( <i>a+c</i> )	29( <i>b+d</i> )	70( <i>n</i> )	58.57

本例为两样本率比较的资料, 表 8-1 内有 4 个数  $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$  是该表的基本数据, 其余数据可以由这 4 个基本数据推算出来, 称为四格表 (fourfold table) 资料。四格表资料可以用  $\chi^2$  检验推断两个总体率 (或构成比) 之间有无差别。 $\chi^2$  检验的检验统计量为  $\chi^2$ , 其基本公式为:

$$\chi^2 = \sum \frac{(A-T)^2}{T} \quad (8-1)$$

自由度为:

$$\nu = (\text{行数} - 1)(\text{列数} - 1) \quad (8-2)$$

公式 (8-1) 亦称 Pearson  $\chi^2$ 。式中 *A* 为实际频数 (actual frequency), 如上例中的 4 个基本数据  $\begin{vmatrix} 20 & 24 \\ 21 & 5 \end{vmatrix}$ ; *T* 为理论频数 (theoretical frequency)。

理论频数 *T* 是根据检验假设  $H_0: \pi_1 = \pi_2$  确定的。如例 8.1, 无效假设是试验组与对照组治疗原发性高血压的总体有效率相等, 于是可以将两组合计有效率 58.57% 作为各组的有效率。按照这一假设, 理论上对照组 44 例原发性高血压病患者中有效者应为  $(44 \times 41)/70 = 25.8$ , 无效者为  $(44 \times 29)/70 = 18.2$ ; 同理, 试验组的 26 例原发性高血压病患者中有效者应为  $(26 \times 41)/70 = 15.2$ , 无效者为  $(26 \times 29)/70 = 10.8$ 。由此可得出理论频数 *T* 的计算公式为:

$$T_{RC} = \frac{n_{R.} \cdot n_{.C}}{n} \quad (8-3)$$

式中  $T_{RC}$  为第 *R* 行 (row) 第 *C* 列 (column) 的理论频数,  $n_{R.}$  为相应行的合计,  $n_{.C}$  为相应列的合计, *n* 为总例数。

由公式 (8-1) 可以看出,  $\chi^2$  值反映了实际频数与理论频数的吻合程度。若检验假设



$H_0$  成立, 实际频数与理论频数的差值会较小, 则  $\chi^2$  值也会较小; 反之, 若检验假设  $H_0$  不成立, 实际频数与理论频数就会相差较大, 则  $\chi^2$  值也会较大。由公式 (8-1) 还可以看出,  $\chi^2$  值的大小还取决于  $\frac{(A-T)^2}{T}$  的个数多少 (自由度  $\nu$  的大小)。由于各  $\frac{(A-T)^2}{T}$  皆是正值, 故自由度  $\nu$  愈大,  $\chi^2$  值也会愈大; 所以只有考虑了自由度  $\nu$  的影响,  $\chi^2$  值才能正确地反映实际频数  $A$  和理论频数  $T$  的吻合程度。

由公式 (8-2) 可见,  $\chi^2$  检验的自由度  $\nu$  取决于可以自由取值的格子数目, 而不是样本含量  $n$ 。四格表资料只有两行两列,  $\nu=1$ , 即在周边合计数固定的情况下, 4 个基本数据当中只有一个可以自由取值, 因此, 对于四格表资料, 只要根据公式 (8-3) 计算出一个理论值  $T_{1c}$  后, 其他 3 个理论值可用周边合计数减去相应的理论值  $T$  得出。如例 8-1 中,  $T_{11}=(44 \times 41)/70=25.8$ ,  $T_{12}=44-25.8=18.2$ ,  $T_{21}=41-25.8=15.2$ ,  $T_{22}=26-15.2=10.8$  (或  $T_{22}=29-18.2=10.8$ )。

$\chi^2$  检验时, 要根据自由度  $\nu$  查  $\chi^2$  界值表。当自由度  $\nu$  确定后,  $\chi^2$  分布曲线下右侧尾部的面积为  $\alpha$  时, 横轴上相应的  $\chi^2$  值记作  $\chi_{\alpha, \nu}^2$ , 即  $\chi^2$  分布的分位数。 $\chi^2$  值与  $P$  值的对应关系见附表 7 的  $\chi^2$  分布界值表。由附表 7 可知, 在自由度  $\nu$  确定后,  $\chi^2$  值愈大,  $P$  值愈小; 反之,  $\chi^2$  值愈小,  $P$  值愈大。若检验水准为  $\alpha$ , 当  $\chi^2 \geq \chi_{\alpha, \nu}^2$  时,  $P \leq \alpha$ , 拒绝  $H_0$ , 接受  $H_1$ ; 当  $\chi^2 < \chi_{\alpha, \nu}^2$  时,  $P > \alpha$ , 不拒绝  $H_0$ 。

现以例 8.1 为例说明  $\chi^2$  检验的步骤。

### 1. 建立检验假设并确定检验水准

$H_0: \pi_1 = \pi_2$ , 即试验组与对照组的总体有效率相等

$H_1: \pi_1 \neq \pi_2$ , 即试验组与对照组的总体有效率不等

$\alpha = 0.05$

### 2. 计算检验统计量

按公式 (8-3) 计算  $T_{11}$ , 然后用减法计算  $T_{12}$ 、 $T_{21}$  和  $T_{22}$ :

$$T_{11} = (44 \times 41)/70 = 25.8, \quad T_{12} = 44 - 25.8 = 18.2$$

$$T_{21} = 41 - 25.8 = 15.2, \quad T_{22} = 26 - 15.2 = 10.8$$

按公式 (8-1) 计算  $\chi^2$  值

$$\chi^2 = \frac{(20-25.8)^2}{25.8} + \frac{(24-18.2)^2}{18.2} + \frac{(21-15.2)^2}{15.2} + \frac{(5-10.8)^2}{10.8} = 8.40$$

再按公式 (8-2) 计算  $\nu$ ,  $\nu = (2-1)(2-1) = 1$ 。

### 3. 确定 $P$ 值, 作出推断结论

以  $\nu=1$  查附表 7 的  $\chi^2$  界值表, 得  $P < 0.005$ 。按  $\alpha = 0.05$  水准, 拒绝  $H_0$ , 接受  $H_1$ , 可以认为两组治疗原发性高血压的总体有效率不等, 即可认为该中药治疗原发性高血压有效。

两样本率比较时, 当总例数  $n \geq 40$  且所有格子的  $T \geq 5$  时, 可用  $\chi^2$  检验的基本公式 (8-1)。实际应用时, 为省去计算理论频数的步骤, 简化计算公式, 常用四格表资料  $\chi^2$  检验的专用公式 (8-4) 计算检验统计量  $\chi^2$  值, 即

$$\chi^2 = \frac{(ad-bc)^2 n}{(a+b)(c+d)(a+c)(b+d)} \quad (8-4)$$

式中  $a, b, c, d$  为四格表的实际频数;  $(a+b)$ ,  $(c+d)$ ,  $(a+c)$ ,  $(b+d)$  是周边合计数;  $n$  为总例数,  $n = a + b + c + d$ 。公式 (8-4) 中相应符号见表 8-1。

仍以例 8.1 资料为例, 用公式 (8-4) 计算  $\chi^2$  值, 有

$$\chi^2 = \frac{(20 \times 5 - 24 \times 21)^2 \times 70}{44 \times 26 \times 41 \times 29} = 8.40$$



结果与用公式 (8-1) 计算的相同。

## 二、四格表资料 $\chi^2$ 检验的校正公式

$\chi^2$  界值表的依据是  $\chi^2$  分布, 其分布是连续型分布, 而计数资料中的实际频数  $A$  为分类资料, 是不连续的。因此, 用公式 (8-1) 计算的  $\chi^2$  值查界值表所得的概率  $P$  偏小, 特别是对自由度  $\nu=1$  的四格表资料的影响。为此, 美国统计学家 F. Yates (1934 年) 提出了用  $|A-T|-0.5$  计算  $\chi^2$  的连续性校正法 (correction for continuity), 其校正公式为:

$$\chi_c^2 = \sum \frac{(|A-T|-0.5)^2}{T} \quad (8-5)$$

$$\chi_c^2 = \frac{(|ad-bc|-n/2)^2 n}{(a+b)(c+d)(a+c)(b+d)} \quad (8-6)$$

公式 (8-5) 和公式 (8-6) 分别是公式 (8-1) 和公式 (8-4) 的校正。在实际工作中, 对于四格表资料, 通常规定为:

(1) 当  $n \geq 40$  且所有的  $T \geq 5$  时, 用  $\chi^2$  检验的基本公式 (8-1) 或四格表资料  $\chi^2$  检验的专用公式 (8-4)。

(2) 当  $n \geq 40$  但有  $1 \leq T < 5$  时, 用四格表资料  $\chi^2$  检验的校正公式 (8-5) 或 (8-6); 或改用四格表资料的 Fisher 确切概率法。

(3) 当  $n < 40$ , 或  $T < 1$  时, 用四格表资料的 Fisher 确切概率法。

**例 8.2** 某医学院抽样调查大学四年级和五年级学生近视眼患病情况, 四年级学生的近视率为 7.14%, 五年级学生的近视率为 35.71%, 调查结果见表 8-2。问该大学四年级与五年级学生的近视眼患病率是否不同?

表 8-2 两个年级大学生的近视眼患病率比较

年级	近视	非近视	合计	近视率(%)
四年级	2(4.7)	26(23.3)	28	7.14
五年级	5(2.3)	9(11.7)	14	35.71
合计	7	35	42	16.67

注: 括号内数字为理论频数

(1) 建立检验假设并确定检验水准

$H_0: \pi_1 = \pi_2$ , 即四年级与五年级学生的近视眼患病率相同

$H_1: \pi_1 \neq \pi_2$ , 即四年级与五年级学生的近视眼患病率不相同

$\alpha = 0.05$

(2) 计算检验统计量

按照公式 (8-3) 计算各观察值的理论频数, 记于表 8-2 的括号中。本例  $n=42$ , 但有 2 个格子的理论频数分别 2.3 和 4.7 均小于 5, 需用四格表资料  $\chi^2$  检验的校正公式 (8-5) 或公式 (8-6)。本例用公式 (8-6) 计算校正  $\chi^2$  值:

$$\chi_c^2 = \frac{(|2 \times 9 - 26 \times 5| - 42/2)^2 \times 42}{(2+5)(26+9)(2+26)(5+9)} = 3.62$$

$$\nu = 1$$

(3) 确定  $P$  值, 作出推断结论

以  $\nu=1$  查附表 7 的  $\chi^2$  界值表得  $P > 0.05$ 。按  $\alpha = 0.05$  水准, 不拒绝  $H_0$ , 还不能认为四年级与五年级学生近视眼患病率不等。

本资料若不校正时,  $\chi^2 = 5.49$ ,  $P < 0.05$ , 结论与之相反。



### 三、四格表的确切概率法

当四格表资料中出现  $n < 40$  或  $T < 1$ , 需改用四格表资料的 Fisher 确切概率 (Fisher's exact test) 法。该法是一种直接计算概率的假设检验方法, 其理论依据是超几何分布 (hypergeometric distribution)。四格表的确切概率法不属于  $\chi^2$  检验的范畴, 但常作为四格表资料假设检验的补充。

确切概率计算法的基本思想是: 在四格表边缘合计固定不变的条件下, 利用公式 (8-7) 直接计算表内四个格子数据的各种组合的概率  $P_i$ , 然后计算单侧或双侧累计概率  $P$ , 并与检验水准  $\alpha$  比较, 作出是否拒绝  $H_0$  的结论。

$$P_i = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{a!b!c!d!n!} \quad (8-7)$$

式中  $a, b, c, d, n$  等符号的意义同表 8-1,  $\sum P_i = 1; !$  为阶乘符号,  $0! = 1$ 。

下面用实例说明其检验原理和检验方法。

**例 8.3** 将 17 名腰椎间盘突出症患者随机分到两组, 分别用两种方法治疗, 结果见表 8-3, 问两种疗法的疗效是否不同?

表 8-3 两种疗法对腰椎间盘突出症的疗效

疗 法	治 愈	未 治 愈	合 计	治 愈 率 (%)
新疗法	7	2	9	77.78
保守疗法	2	6	8	25.00
合 计	9	8	17	52.94

本例  $n = 17 < 40$ , 不满足  $\chi^2$  检验的应用条件, 宜用 Fisher 确切概率法。其假设检验步骤如下:

1. 建立检验假设并确定检验水准

$H_0: \pi_1 = \pi_2$ , 即两种疗法对腰椎间盘突出症的疗效相同

$H_1: \pi_1 \neq \pi_2$ , 即两种疗法对腰椎间盘突出症的疗效不同

$\alpha = 0.05$

2. 计算概率

在四格表周边合计数不变的条件下, 表 8-3 内 4 个实际频数变动的组合数共有“周边合计中最小数+1”个即  $8+1=9$  个, 根据公式 (8-7) 计算各种组合的四格表概率, 结果见表 8-4。例如实际观察到的四格表资料的概率为:

$$P^* = \frac{9!8!8!9!}{7!2!2!6!17!} = 0.041464$$

表 8-4 各种组合的四格表计算的概率

四格表序号	治 愈	未 治 愈	$a-T_a$	$P$
1	1	8	-3.76	0.000370
	8	0		
2	2	7	-2.76	0.011847
	7	1		
3	3	6	-1.76	0.096750
	6	2		
4	4	5	-0.76	0.290251
	5	3		



续表

四格表序号	治愈	未治愈	$a - T_a$	$P$
5	5	4	0.24	0.362814
	4	4		
6	6	3	1.24	0.193501
	3	5		
7*	7	2	2.24*	0.041464*
	2	6		
8	8	1	3.24	0.002962
	1	7		
9	9	0	4.24	0.000041
	0	8		

\* : 为实际四格表

3. 确定累计概率  $P$  值, 作出推断结论

双侧检验: 在没有足够临床经验认为新疗法治愈率可能高于保守疗法的情况下, 应做双侧检验假设。在四格表周边合计数不变的条件下,  $a$  值的理论频数为  $T_{11} = T_a = (9 \times 9) / 17 = 4.76$ ; 在实际观察频数  $a = 7$  时,  $|a - T_a| = |7 - 4.76| = 2.24$ 。观察上述 9 个  $2 \times 2$  表,  $a$  值越大,  $c$  值越小,  $|a - T_a|$  值越大;  $a$  值越小,  $c$  值越大,  $|a - T_a|$  值越大。欲拒绝  $H_0$ ,  $P$  值的计算应包括  $|a - T_a| \geq 2.24$  的四格表的概率之和。双侧累计概率  $P$  值为:

$$\begin{aligned} P &= P(1) + P(2) + P(7) + P(8) + P(9) \\ &= 0.000370 + 0.011847 + 0.041464 + 0.002962 + 0.000041 \\ &= 0.057 \end{aligned}$$

$P = 0.057 > 0.05$ , 在  $\alpha = 0.05$  检验水准下, 不拒绝  $H_0$ , 还不可以认为两种疗法不同。

单侧检验: 若本例有充足的医学知识认为新疗法不会比保守疗法差, 只需作单侧检验, 计算包括  $a - T_a \geq 2.24$  的四格表的概率之和。单侧累计概率  $P$  值为:

$$\begin{aligned} P &= P(7) + P(8) + P(9) = 0.041464 + 0.002962 + 0.000041 \\ &= 0.044 \end{aligned}$$

$P < 0.05$ , 在  $\alpha = 0.05$  检验水准下, 拒绝  $H_0$ , 接受  $H_1$ 。可以认为两种疗法不同, 新疗法的治愈率高于保守疗法。

同一个资料, 由于选择检验的单、双侧不同, 得出了不同的结论。如前所述, 在资料分析时, 用单侧检验还是用双侧检验, 应根据研究目的在实验设计时确定, 不应该在实验结束后进行统计分析时为达到主观愿望而临时作出选择。

## 第二节 配对四格表资料的 $\chi^2$ 检验

**例 8.4** 现有 198 份痰标本, 每份标本分别用 A、B 两种培养基培养结核菌, 结果见表 8-5。问 A、B 两种培养基的阳性培养率是否不等?

表 8-5 A、B 两种培养基的培养结果

A 培养基	B 培养基		合计
	+	-	
+	48(a)	24(b)	72
-	20(c)	106(d)	126
合计	68	130	198



本例为配对设计的计数资料。计数资料的配对设计常用于两种检验方法、培养方法、诊断方法的比较，其特点是对样本中各观察单位分别用两种方法处理，然后观察两种处理方法的某两分类变量的计数结果。观察结果有四种情况，可整理成表 8-5 的形式：①A、B 两种检测方法皆为阳性数 ( $a$ )；②A、B 两种检测方法皆为阴性数 ( $d$ )；③A 法为阳性、B 法为阴性数 ( $b$ )；④A 法为阴性、B 法为阳性数 ( $c$ )。其中， $a$ 、 $d$  为两法观察结果一致的两种情况， $b$ 、 $c$  为两法观察结果不一致的两种情况。当两种处理方法无差别时，对总体有  $B=C$ ，即两总体率相等  $\pi_1=\pi_2$ 。由于在抽样研究中，抽样误差是不可避免的，样本中的  $b$  和  $c$  往往不等 ( $b\neq c$ ，即两样本率不等： $p_1\neq p_2$ )。为此，需进行假设检验 (McNemar test)，其检验统计量为

$$\chi^2 = \frac{(b-c)^2}{b+c}, \nu=1 \quad (8-8)$$

$$\chi^2 = \frac{(|b-c|-1)^2}{b+c}, \nu=1 \quad (8-9)$$

公式 (8-8) 用于  $(b+c)\geq 40$  时，公式 (8-9) 用于  $(b+c) < 40$  时。值得注意的是，该法一般用于样本含量不太大的资料。因本法仅考虑了两法结果不一致的两种情况 ( $b$ 、 $c$ )，而未考虑样本含量  $n$  和两法结果一致的两种情况 ( $a$ 、 $d$ )。所以，当  $n$  很大且  $a$  与  $d$  的数值很大 (即两法的一致率较高)， $b$  与  $c$  的数值相对较小时，即便是检验结果有统计学意义，其实际意义往往也不大。

本例的检验步骤如下：

(1) 建立检验假设并确定检验水准

$H_0: B=C$ ，即两种培养基的阳性培养率相等

$H_1: B\neq C$ ，即两种培养基的阳性培养率不相等

$\alpha=0.05$

(2) 计算检验统计量

本例  $b+c>40$ ，用公式 (8-8) 计算得

$$\chi^2 = \frac{(24-20)^2}{24+20} = 0.36, \nu=1$$

(3) 确定  $P$  值，作出推断结论

查  $\chi^2$  界值表得  $P>0.05$ 。按  $\alpha=0.05$  水准，不拒绝  $H_0$ 。尚不能认为两种培养基的阳性培养率不同。

应该注意，配对设计的四格表资料只能用配对  $\chi^2$  检验，而不能随意转化为两组独立样本的  $\chi^2$  检验。

### 第三节 行×列表资料的 $\chi^2$ 检验

前面介绍了两个样本率比较的  $\chi^2$  检验方法，其基本数据有 2 行 2 列，称  $2\times 2$  表或四格表资料。本节介绍的行×列表资料的  $\chi^2$  检验，用于多个样本率的比较、两个或多个构成比的比较。其基本数据有以下三种情况：①多个样本率比较时，有  $R$  行 2 列，称为  $R\times 2$  表；②两个样本的构成比比较时，有 2 行  $C$  列，称  $2\times C$  表；③多个样本的构成比比较，有  $R$  行  $C$  列，称为  $R\times C$  表。以上三种情况可统称为行×列表资料。

行×列表资料的  $\chi^2$  检验仍用 Pearson  $\chi^2$  公式，即公式 (8-1) 计算检验统计量  $\chi^2$  值。因该式需先计算理论频数  $T_{RC}$ ，计算较繁琐，可将计算理论频数的公式 (8-3) 代入公式 (8-1)，化简后得行×列表资料  $\chi^2$  检验的专用公式为：



$$\chi^2 = n \left( \sum \frac{A^2}{n_{rnc}} - 1 \right), \nu = (\text{行数} - 1)(\text{列数} - 1) \quad (8-10)$$

式中各符号的意义同前。

## 一、行×列表的 $\chi^2$ 检验

**例 8.5** 某医院用 3 种方案治疗急性无黄疸型病毒肝炎 254 例，观察结果见表 8-6，问 3 种疗法的有效率是否不同。

表 8-6 3 种方案治疗肝炎的疗效

组别	有效	无效	合计	有效率(%)
西药组	51	49	100	51.00
中药组	35	45	80	43.75
中西药结合组	59	15	74	79.73
合计	145	109	254	57.09

本例为 3 个样本率的比较，是  $3 \times 2$  表资料。

(1) 建立检验假设并确定检验水准

$H_0$ : 3 种治疗方案的有效率相等

$H_1$ : 3 种治疗方案的有效率不全相等

$\alpha = 0.05$

(2) 计算检验统计量

按公式 (8-10) 计算  $\chi^2$  值:

$$\begin{aligned} \chi^2 &= 254 \times \left( \frac{51^2}{100 \times 145} + \frac{49^2}{100 \times 109} + \frac{35^2}{80 \times 145} + \frac{45^2}{80 \times 109} + \frac{59^2}{74 \times 145} + \frac{15^2}{74 \times 109} - 1 \right) \\ &= 254 \times (0.1794 + 0.2203 + 0.1056 + 0.2322 + 0.3244 + 0.0279 - 1) \\ &= 22.81 \end{aligned}$$

$$\nu = (3-1)(2-1) = 2$$

(3) 确定  $P$  值，作出推断结论

查  $\chi^2$  界值表得  $P < 0.05$ ，在  $\alpha = 0.05$  的检验水准下，拒绝  $H_0$ ，接受  $H_1$ ，可以认为三种疗法的有效率有差别。

**例 8.6** 某研究人员收集了亚洲、欧洲和北美洲人的 A、B、AB、O 血型资料，结果见表 8-7，问不同地区人群 ABO 血型分类构成比是否不同。

表 8-7 三个不同地区血型样本的频数分布

地区	A	B	AB	O	合计
亚洲	321	369	95	295	1 080
欧洲	258	43	22	194	517
北美洲	408	106	37	444	995
合计	987	518	154	933	2 592

本例为 3 个样本构成比的比较，是  $3 \times 4$  表资料。

(1) 建立检验假设并确定检验水准

$H_0$ : 不同地区人群血型分布总体构成比相同

$H_1$ : 不同地区人群血型分布总体构成比不全相同

$\alpha = 0.05$

(2) 计算检验统计量

按公式 (8-10) 计算  $\chi^2$  值:



$$\chi^2 = 2592 \left( \frac{321^2}{987 \times 1080} + \frac{369^2}{518 \times 1080} + \dots + \frac{444^2}{933 \times 995} - 1 \right) = 297.38$$

$$\nu = (3-1)(4-1) = 6$$

(3) 确定 P 值，作出推断结论

查  $\chi^2$  界值表得  $P < 0.05$ ，在  $\alpha = 0.05$  检验水准下，拒绝  $H_0$ ，认为三个不同地区的人群血型分布总体构成比有差别。

当多个样本率比较的  $R \times 2$  表资料  $\chi^2$  检验，推断结论为拒绝  $H_0$ ，接受  $H_1$  时，只能认为各总体率之间总的来说有差别，但不能说明任两个总体率之间有差别。多个样本率间的两两比较若直接用四格表资料的  $\chi^2$  检验进行多重比较，将会加大犯 I 类错误的概率。因此，样本率间的多重比较不能直接用四格表资料的  $\chi^2$  检验。下面介绍多个样本率间多重比较的  $\chi^2$  检验。

## 二、多个样本率间多重比较

多个样本率间多重比较有  $\chi^2$  分割法和 bonferroni 方法，应用这些方法的目的是为保证检验假设中 I 型错误  $\alpha$  的概率不变。这里仅介绍最简单的 bonferroni 方法，其基本思想是根据重复检验的次数重新规定检验水准  $\alpha'$ 。该方法是一种比较保守的方法，比较的组数不宜过多。根据分析目的不同、 $k$  个样本率两两比较的次数不同，重新规定的检验水准的估计方法亦不同。通常有两种情况：

1. 多个实验组间的两两比较 分析目的为  $k$  个实验组间，任两个率均进行比较，检验水准  $\alpha'$  可用下式估计：

$$\alpha' = \frac{\alpha}{\binom{k}{2}} \tag{8-11}$$

式中  $\binom{k}{2} = \frac{k(k-1)}{2}$ ， $k$  为参加检验的组数。

2. 实验组与同一个对照组的比较 分析目的为各实验组与同一个对照组比较，而各实验组间不需要比较。检验水准  $\alpha'$  可用下式估计：

$$\alpha' = \frac{\alpha}{k-1} \tag{8-12}$$

式中  $k$  为参加检验的组数。

为了方便查表和应用，现将多个样本率比较时常用的  $\chi^2$  值与对应的概率  $P$  值整理于表 8-8 中。

表 8-8  $\nu=1$  时的  $\chi^2$  界值表 (供多个样本率间的多重比较用)

$\chi^2$	$P$	$\chi^2$	$P$	$\chi^2$	$P$
5.02	0.02500	6.96	0.00833	7.88	0.00500
5.73	0.01666	7.24	0.00714	8.05	0.00455
6.24	0.01250	7.48	0.00625	8.21	0.00417

**例 8.7** 对例 8.5 中表 8-6 的资料进行两两比较，以推断是否任两种疗法治疗急性无黄疸型病毒肝炎的有效率均有差别？

本例研究目的为 3 个实验组间的两两比较，检验水准  $\alpha'$  用公式 (8-11) 估计得：

$$\alpha' = \frac{0.05}{3(3-1)/2} = \frac{0.05}{3} = 0.0167$$

(1) 建立检验假设并确定检验水准





$H_0: \pi_A = \pi_B$ , 即任两对比组的总体有效率相等

$H_1: \pi_A \neq \pi_B$ , 即任两对比组的总体有效率不等

$\alpha = 0.05$

(2) 计算检验统计量

本例研究目的为 3 个实验组间的两两比较, 两两比较的  $2 \times 2$  表及用公式 (8-4) 分别计算任两对比组的检验统计量  $\chi^2$  值结果见表 8-9。

(3) 确定  $P$  值, 作出推断结论

由表 8-8, 任两对比组的  $P$  值结果见表 8-9。按  $\alpha' = 0.0167$  水准, 西药组与中药组的疗效比较不拒绝  $H_0$ , 尚不能认为西药与中药治疗肝炎的有效率有差异; 中药组与中西药结合组的疗效比较拒绝  $H_0$ , 接受  $H_1$ , 可以认为中药与中西药结合治疗肝炎的有效率有差异, 中西药结合的疗法好于单纯用中药的疗法; 西药组与中西药结合组的疗效比较拒绝  $H_0$ , 接受  $H_1$ , 可以认为西药与中西药结合治疗肝炎的有效率有差异, 中西药结合的疗法好于单纯用西药的疗法。

表 8-9(a) 西药组与中药组的疗效比较

对比组	有效	无效	合计	$\chi^2$	$P$
西药组	51	49	100	0.94	>0.0167
中药组	35	45	80		
合计	86	94	180		

表 8-9(b) 中药组与中西药结合组的疗效比较

对比组	有效	无效	合计	$\chi^2$	$P$
中药组	35	45	80	20.93	<0.0167
中西药结合组	59	15	74		
合计	94	60	154		

表 8-9(c) 西药组与中西药结合组的疗效比较

对比组	有效	无效	合计	$\chi^2$	$P$
西药组	51	49	100	15.10	<0.0167
中西药结合组	59	15	74		
合计	110	64	174		

**例 8.8** 以例 8.5 中表 8-6 资料中的中药治疗组为对照组, 西药治疗组与中西药结合为试验组, 试分析两试验组与对照组的总体有效率有无差别?

本例研究目的为各试验组与同一对照组的比较, 检验水准  $\alpha'$  用公式 (8-12) 估计得:

$$\alpha' = \frac{\alpha}{3-1} = 0.025$$

(1) 建立检验假设并确定检验水准

$H_0: \pi_T = \pi_C$ , 即各试验组与对照组的总体有效率相等

$H_1: \pi_T \neq \pi_C$ , 即各试验组与对照组的总体有效率不等

$\alpha = 0.05$

(2) 计算检验统计量

本例为各试验组与同一对照组的比较, 由表 8-9 可知: 西药组与中药组的疗效比较  $\chi^2 = 0.94$ , 中西药结合组与中药组的疗效比较  $\chi^2 = 20.93$ 。

(3) 确定  $P$  值, 作出推断结论



查表 8-8, 各试验组与同一对照组比较的  $P$  值结果见表 8-9。按  $\alpha' = \frac{\alpha}{3-1} = 0.025$  水准, 西药组与中药组的疗效比较不拒绝  $H_0$ , 尚不能认为西药与中药治疗肝炎的有效率有差异; 中药组与中西药结合组的疗效比较拒绝  $H_0$ , 接受  $H_1$ , 可以认为中药与中西药结合治疗肝炎的有效率有差异, 中西药结合的疗法好于单纯用中药的疗法。

### 三、行×列表 $\chi^2$ 检验时的注意事项

1. 一般认为, 行×列表资料中各格的理论频数不应小于 1, 并且  $1 \leq T < 5$  的格子数不宜超过格子总数的  $1/5$ 。若出现上述情况, 可通过以下方法解决: ①最好是增加样本含量, 使理论频数增大; ②根据专业知识, 考虑能否删去理论频数太小的行或列, 能否将理论频数太小的行或列与性质相近的邻行或邻列合并; ③改用双向无序  $R \times C$  表资料的 Fisher 确切概率法。

2. 多个样本率比较, 若所得统计推断为拒绝  $H_0$ , 接受  $H_1$  时, 只能认为各总体率之间总的来说有差别, 但不能说明任两个总体率之间均有差别。要进一步推断哪两两总体率之间有差别, 需进一步做多个样本率的多重比较。

3. 在实际应用中, 对于  $R \times C$  表资料要根据其分类类型和研究目的选用恰当的检验方法。行×列表资料的  $\chi^2$  检验与分类变量的顺序无关, 对于有序的  $R \times C$  表资料不宜用  $\chi^2$  检验。

## 练 习 题

### 一、单项选择题

- 利用  $\chi^2$  检验公式不适合解决的实际问题是
  - 比较两种药物的有效率
  - 检验某种疾病与基因多态性的关系
  - 两组有序试验结果的药物疗效
  - 药物三种不同剂量显效率有无差别
  - 两组病情“轻、中、重”的构成比例
- 欲比较两组阳性反应率, 在样本量非常小的情况下 (如  $n_1 < 10, n_2 < 10$ ), 应采用
  - 四格表  $\chi^2$  检验
  - 校正四格表  $\chi^2$  检验
  - Fisher 确切概率法
  - 配对  $\chi^2$  检验
  - 校正配对  $\chi^2$  检验
- 进行四组样本率比较的  $\chi^2$  检验, 如  $\chi^2 > \chi_{0.01,3}^2$ , 可认为
  - 四组样本率均不相同
  - 四组总体率均不相同
  - 四组样本率相差较大
  - 至少有两组样本率不相同
  - 至少有两组总体率不相同
- 从甲、乙两文中, 查到同类研究的两个率比较的  $\chi^2$  检验, 甲文  $\chi^2 > \chi_{0.01,1}^2$ , 乙文  $\chi^2 > \chi_{0.05,1}^2$ , 可认为
  - 两文结果有矛盾
  - 两文结果完全相同
  - 甲文结果更为可信
  - 乙文结果更为可信
  - 甲文说明总体的差异较大
- 两组有效率比较检验功效的相关因素是
  - 检验水准和样本率
  - 总体率差别和样本含量
  - 样本含量和样本率
  - 总体率差别和理论频数
  - 容许误差和检验水准

### 二、计算与分析

- 某神经内科医师观察 291 例脑梗塞病人, 其中 102 例病人用西医疗法, 其他 189 例



病人采用西医疗法加中医疗法，观察一年后，单纯用西医疗法组的病人死亡 13 例，采用中西医疗法组的病人死亡 9 例，请分析两组病人的死亡率差异是否有统计学意义？

2. 某医院研究中药治疗急性心肌梗死的疗效，临床观察结果见下表。问接受两种不同疗法的患者病死率是否不同？

两种药治疗急性心肌梗死的疗效

组 别	存活	死亡	合计	病死率(%)
中药组	65	3	68	4.41
非中药组	12	2	14	14.29
合计	77	5	82	6.10

3. 某医师观察三种降血脂药 A, B, C 的临床疗效，观察 3 个月后，按照患者的血脂下降程度分为有效与无效，结果如下表，问三种药物的降血脂效果是否不同？

三种药物降血脂的疗效

药物	有效	无效	合计
A	120	25	145
B	60	27	87
C	40	22	62

4. 某医师按照白血病患者发病情况，将 308 例患者分为两组，并按 ABO 血型分类记数，试问两组患者血型总体构成有无差别？

308 例急、慢性白血病患者血型分布

组别	A	B	O	AB	合计
急性组	60	47	61	21	189
慢性组	42	30	34	13	119
合计	102	77	95	34	308

5. 为研究某补钙制剂的临床效果，观察 56 例儿童，其中一组给予这种新药，另一组给予钙片，观察结果如表，问两种药物预防儿童的佝偻病患率是否不同？

两组儿童的佝偻病患情况

组别	病例数	非病例数	合计	患病率(%)
新药组	8	32	40	20.0
钙片组	6	10	16	37.5
合计	14	42	56	25.0

6. 某医院 147 例大肠埃希菌标本分别在 A, B 两种培养基上培养，然后进行检验，资料见下表，试分析两种培养基的检验结果是否有显著性差别？

A、B 两种培养基上培养大肠埃希菌标本结果

A 培养基	B 培养基		合 计
	+	-	
+	59	36	95
-	15	37	52
合计	74	73	147

(王乐三)



### The Chi-Square Test

The chi-square test that we consider is a test of differences among distributions. Its purpose is to compare the observed frequencies of a discrete, ordinal, or categorical data set with those of theoretically expected distribution.

The assumptions for the  $\chi^2$  test for goodness of fit are as follows:

- An independent random sample is drawn from the population.
- The population can be divided into a set of  $k$  mutually exclusive categories.
- The expected frequencies for each category must be specified.

The hypothesis test takes only one form:

$H_0$ : The observed frequency distribution is the same as the hypothesized frequency distribution.

$H_1$ : The observed and hypothesized frequency distribution are different.

Let  $A_i$  denote the observed frequency of the  $i$  th category. Let  $T_i$  denote the expected frequency for the  $i$  th category. The test statistic is based on the difference between the observed and expected frequencies,  $A_i - T_i$ . It is defined as

$$\chi^2 = \sum_{i=1}^k \frac{(A_i - T_i)^2}{T_i}$$

## 第九章 非参数检验

前面讨论的方法都是基于总体分布为正态分布的前提下对参数进行的检验,即参数检验方法(parametric test)。但是实际上,有些资料并不符合上述条件,甚至常常分布是未知的。本章所讨论的非参数检验方法(nonparametric test),可以不考虑总体的参数和总体的分布类型,对总体的分布或分布位置进行检验。这种统计方法,有时也称为任意分布检验(distribution-free test)。非参数检验通常适用于下述资料:

- (1) 总体分布为偏态或分布未知的计量资料;
- (2) 等级资料;
- (3) 个别数据偏大或数据的某一端无确定的数值,如“ $<0.01\text{mg}$ ”、“ $>150\text{mg}$ ”等,只有一个下限或上限,而没有具体数值;
- (4) 各组离散程度相差悬殊,即各总体方差不齐。

非参数检验方法的优点是适应性强,但由于损失了部分信息,检验效率降低了。在资料服从正态分布的前提下,当 $H_0$ 不真时,非参数检验方法不如参数检验方法能灵敏地拒绝 $H_0$ ,换句话说犯第二类错误的可能性大于参数检验法。因此,对于适合参数检验的资料,最好还是用参数检验。

### 第一节 配对资料的符号秩和检验

配对资料的符号秩和检验(Wilcoxon signed rank test)是配对设计的非参数检验。

#### 一、检验步骤和基本思想

##### (一) 检验步骤

1. 求出各对数据的差值

2. 建立假设检验

$H_0$ : 差值的总体中位数为零

$H_1$ : 差值的总体中位数不为零

确定检验水准  $\alpha$

3. 编秩次并求秩和

依差值绝对值,从小到大编秩,并按差值的正负,标上正负号。编秩时,在正负号不同的差数中,若有绝对值相等的观测值,则取其平均秩次。对差值为0的对子,舍去不计,相应地总的对子数也要减去其对子数,记为 $n$ 。分别求正负秩次之和 $T_+$ 与 $T_-$ ,并以绝对值较小者作为统计量 $T$ 值,如公式(9-1)所示:

$$T = \min(T_+, T_-) \quad (9-1)$$

正、负秩和相加应等于总秩和,即 $T_+ + T_- = n(n+1)/2$ ,通过对其计算可以判断 $T_+$ 和 $T_-$ 的计算是否有误。

4. 查表确定 $P$ 值范围

当 $n \leq 25$ 时,可查附表8的 $T$ 界值表, $T$ 愈小则 $P$ 愈小。当 $T$ 恰为附表中的界值时, $P$ 值一般都小于表中对应的概率值。

应该注意,当 $n > 25$ 时,无法查附表8,可接近似正态分布用 $u$ 检验,其公式为:



$$u = \frac{|T - n(n+1)/4| - 0.5}{\sqrt{n(n+1)(2n+1)/24}} \quad (9-2)$$

当相同秩次较多时, 采用校正公式:

$$u = \frac{|T - n(n+1)/4| - 0.5}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\sum(t_i^2 - t_i)}{48}}} \quad (9-3)$$

其中  $t_i$  为相同秩次的个数。得出  $u$  值后, 按  $u$  界值确定  $P$  的范围。

式中 0.5 是连续校正数。因为  $T$  值是不连续的, 而  $u$  分布是连续的, 这种校正影响甚微, 也可以省去。

## (二) 基本思想

如果  $H_0$  成立, 即差值的总体中位数为 0, 则理论上样本的正负秩和应相等, 即  $T$  值应为总秩和  $\frac{n(n+1)}{2}$  的一半, 即  $T = \frac{n(n+1)}{4}$ 。由于存在抽样误差,  $T$  应接近  $\frac{n(n+1)}{4}$ ,  $T$  愈小,  $T$  与  $\frac{n(n+1)}{4}$  的差距越大, 相应的  $P$  值就愈小。当  $P \leq \alpha$  时, 拒绝  $H_0$ 。

## 二、配对资料的符号秩和检验

**例 9.1** 临床某医生研究白癜风病人的白介素 IL-6 指标在白斑部位与正常部位有无差异, 调查的资料如表 9-1 所示。

表 9-1 白癜风病人的不同部位白介素 IL6 指标 (pg/ml)

病人号(1)	白斑部位(2)	正常部位(3)	$d = \text{正常} - \text{白斑}$ (4)	秩次(5)
1	40.03	88.57	48.54	6
2	97.13	80.00	-17.13	-3
3	80.32	123.72	43.40	4
4	25.32	39.03	13.71	2
5	19.61	24.37	4.76	1
6	14.50	92.75	78.25	8
7	49.63	121.57	71.94	7
8	44.56	89.76	45.20	5
合计			$T_+ = 33$	$T_- = 3$

( $T_+ + T_- = 33 + 3 = 36$ , 总秩和  $\frac{n(n+1)}{2} = \frac{8(8+1)}{2} = 36$ , 计算准确无误)

### (1) 建立假设检验

$H_0$ : 差值的总体中位数为零

$H_1$ : 差值的总体中位数不为零

$\alpha = 0.05$

### (2) 计算统计量

首先计算每个对子的差值  $d$ , 见表 9-1 第 (4) 列, 根据 8 个  $d$  的绝对值, 由小到大编秩, 并冠以原  $d$  的正负号, 见表 9-1 第 (5) 列。然后分别相加正负秩次, 得到秩和  $T_+ = 33$ ,  $T_- = 3$ , 取统计量:

$$T = \min(T_+, T_-) = 3$$

### (3) 查表及结论

现  $n = 8$ , 查  $T$  界值表  $T_{0.05(8)} = 3 \sim 33$ ,  $T = 3$  落在界点上,  $P < 0.05$ , 按  $\alpha = 0.05$  水准拒绝  $H_0$ , 可认为白斑部位与正常部位的白介素 IL-6 有差异。因为负秩和小 ( $T_- = 3$ ) 可看出白斑部位白介素指标小于正常部位。



**例 9.2** 对 28 名有轻度牙周疾病的成年人, 指导他们实行良好的口腔卫生习惯, 6 个月 after, 牙周情况好转程度依高到低给予分数 +3, +2, +1; 牙周情况变差程度依次给予分数 -1, -2, -3; 没有变化给予 0 分。数据如表 9-2 所示, 试对此项指导的结果进行评价。

表 9-2 实行良好口腔卫生习惯 6 个月后牙周情况的变化程度

变化对应的分数	人数	变化对应的分数	人数
+3	4	-1	4
+2	5	-2	2
+1	6	-3	2
0	5		

(1) 建立假设检验

$H_0$ : 差值的总体中位数为零

$H_1$ : 差值的总体中位数不为零

$\alpha=0.05$

(2) 计算统计量

设变化对应分数为  $d$ , 编秩及计算正负秩和如表 9-3 所示。

表 9-3 正负秩和计算表

$d$ (1)	频 数			秩次范围 (5)	平均秩次 (6)	负秩和 (7)=(2)×(6)	正秩和 (8)=(3)×(6)
	- (2)	+ (3)	总 (4)				
1	4	6	10	1-10	5.5	22	33
2	2	5	7	11-17	14.0	28	70
3	2	4	6	18-23	20.5	41	82
合计	8	15	23			$T_- = 91$	$T_+ = 185$

(3) 查表及结论

查  $T$  界值,  $T_{0.05(23)} = 73 \sim 203$ 。因为  $T = \min(T_+, T_-) = T_- = 91 > 73$ ,  $P > 0.05$  按  $\alpha = 0.05$  水准, 不拒绝  $H_0$ , 即对 28 名有轻度牙周疾病的成年人, 实行良好的口腔卫生习惯 6 个月后, 牙周情况尚无显著好转。

本例若用  $u$  检验, 由于上述资料相同秩次较多, 用校正式 (9-3)。

$$\begin{aligned}
 u &= \frac{|T - n(n+1)/4| - 0.5}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\sum(t^3 - t)}{48}}} \\
 &= \frac{|91 - 23(23+1)/4| - 0.5}{\sqrt{\frac{23(23+1)(2 \times 23+1)}{24} - \frac{(10^3 - 10) + (7^3 - 7) + (6^3 - 6)}{48}}} \\
 &= 1.44
 \end{aligned}$$

查  $u$  值表  $u_{0.05/2} = 1.96$ ,  $u < u_{0.05/2}$ ,  $P > 0.05$  按  $\alpha = 0.05$  水准, 不拒绝  $H_0$ , 结论同前。

## 第二节 两样本比较的秩和检验

在  $t$  检验一章中, 曾介绍了两独立样本均数的比较, 但是样本资料要求总体服从正态分布。当此条件不满足时, 可采用本节介绍的 Wilcoxon Mann-Whitney test 方法, 其目的是比较两独立样本分别代表的总体分布有无差异。



## 一、检验步骤与基本思想

### (一) 检验步骤

1. 建立假设检验及确定显著水准  $\alpha$

$H_0$ : 两总体分布相同;

$H_1$ : 两总体分布不同;

$\alpha=0.05$ 。

2. 编秩号

两样本观察值从小到大混合编秩, 属不同组的相同观察值取原秩次的平均秩次。

3. 求秩和

设  $n_1$  与  $n_2$  分别为两样本的含量, 规定  $n_1 < n_2$ , 两组合计例数  $N = n_1 + n_2$ 。分别计算两样本含量为  $n_1$  和  $n_2$  组对应的秩和  $T_1$  和  $T_2$  (两组的秩和合计等于总秩和, 即  $T_1 + T_2 = N(N+1)/2$ , 可用于核对), 取样本含量小的  $n_1$  的秩和  $T_1$  为统计量  $T$  值。

4. 确定  $P$  值

当  $n_1 < 10$ ,  $n_2 - n_1 \leq 10$  时, 查附表 9 的  $T$  界值表。 $T$  值在表中范围外 (包括端点时),  $P$  值小于表中对应的概率值,  $T$  值在表中范围内,  $P$  值大于表中对应的概率值。

当  $n_1$  与  $n_2$  超出  $T$  界值表的范围时, 可接近似正态用  $u$  检验:

$$u = \frac{\left| T - \frac{1}{2} n_1 (N+1) \right| - 0.5}{\sqrt{n_1 \cdot n_2 (N+1) / 12}} \quad (9-4)$$

当相同秩次较多时 (如等级资料), 采用校正公式:

$$u = \frac{\left| T - \frac{1}{2} n_1 (N+1) \right| - 0.5}{\sqrt{\frac{n_1 n_2}{12 N (N-1)} (N^3 - N - \sum (t_i^3 - t_i))}} \quad (9-5)$$

其中  $t_i$  为相同秩次的个数。

### (二) 基本思想

如果  $H_0$  成立, 则两样本来自分布相同的总体, 两样本的平均秩次  $T_1/n_1$  与  $T_2/n_2$  应相等或很接近, 且都和总体的平均秩次  $(N+1)/2$  相差很小。含量为  $n_1$  样本的秩和  $T_1$ , 应在  $n_1(N+1)/2$  ( $T$  值表范围中心为  $n_1(N+1)/2$ ) 的左右变化, 当  $T$  值偏离此值太远,  $H_0$  发生的可能性就很小; 若偏离出给定  $\alpha$  值所确定的范围, 即  $P < \alpha$  时, 拒绝  $H_0$ 。

## 二、两样本比较的秩和检验

**例 9.3** 对无淋巴细胞转移与有淋巴细胞转移的胃癌患者进行随访, 观察其生存时间如表 9-4 所示, 问两组患者的生存时间是否不同?

表 9-4 两组患者生存时间 (月)

无淋巴细胞转移		有淋巴细胞转移	
时间	秩次	时间	秩次
12	4.5	5	1
25	10	8	2
27	11	12	4.5
29	12.5	12	4.5
38	17	12	4.5
42	19	17	7





续表

无淋巴细胞转移		有淋巴细胞转移	
时间	秩次	时间	秩次
46	20	21	8
46	21	24	9
56	23	29	12.5
60	24	30	14
		34	15
		36	16
		40	18
		48	22
$n_1=10$	$T_1=162$	$n_2=14$	$T_2=138$

( $T_1+T_2=162+138=300$  与  $N(N+1)/2=300$  相等, 核对计算无误)

### (1) 建立假设检验

$H_0$ : 两患者的生存时间总体分布相同;

$H_1$ : 两组患者的生存时间长短不同;

$\alpha=0.05$ 。

### (2) 计算统计量

将两样本 24 个数据由小到大统一编秩, 见表 9-4 的“秩次”一列, 由于两组生存时间有相同的观察值 12 与 29, 原秩次分别应为 3、4、5、6 与 12、13, 所以取平均秩次后每一个分别为 4.5 与 12.5。同一组内的相同值也可以按照排列的顺序取秩次, 而不取平均秩次, 因为它不影响该组总的秩和。本例两组例数各为 10、14, 取较小者为  $n_1$ , 因此  $n_1=10$ ,  $T_1=162$ , 故统计量  $T=T_1=162$ 。

### (3) 查表与结论

$n_1-n_2=4$ , 按  $\alpha=0.05$ , 查  $T$  值表 (附表 9) 得范围 91~159。因为  $T=162>159$ , 超出范围, 故  $P<0.05$ , 拒绝检验假设  $H_0$ , 即两组患者的生存时间不同。由于编秩次采用由小到大, 因此, 平均秩次大的其实际水平高, 平均秩次水平小的其实际水平低。无转移组的平均秩次为  $162/10=16.2$ , 而有转移组平均秩次为  $138/14=9.86$  小于无转移组, 所以无转移组的平均时间长于有转移组。

**例 9.4** 44 例健康人与 24 例慢性气管炎病人痰液嗜酸性粒细胞数的测量值如表 9-5 的 (2)、(3) 栏所示, 问健康人与慢性气管炎病人痰液嗜酸性粒细胞数有无显著差别?

表 9-5 两组人痰嗜酸性粒细胞的秩和计算

嗜酸性粒细胞	健康人	病人	秩范围	平均秩次	病人组的秩和
(1)	(2)	(3)	(4)	(5)	(6)=(3)×(5)
-	5(11.4)	11(45.8)	1—16	8.5	93.5
+	18(40.9)	10(41.7)	17—44	30.5	305.0
++	16(36.3)	3(12.5)	45—63	54.0	162.0
+++	5(11.4)	0(0.0)	64—68	66.0	0.0
合计	44	24			$T_1=560.5$

对于表 9-5 频数表资料, 可以采用下面计算方法进行检验。

### (1) 建立假设检验

$H_0$ : 健康人与病人痰液嗜酸性粒细胞数总体分布相同



$H_1$ : 健康人与病人痰液嗜酸性粒细胞数总体分布不同

$\alpha=0.05$

(2) 计算统计量

表 9-5 中第 (4) 列系按第 (2) 与 (3) 列数据统一编的秩号, 第 (5) 列为各等级的平均秩号, 第 (6) 列是较小样本的秩和  $T_1=560.5$ 。依公式 (9-5) 计算:

$$u = \frac{\left| T_1 - \frac{1}{2} n_1 (N+1) \right| - 0.5}{\sqrt{\frac{n_1 n_2}{12N(N-1)} (N^3 - N - \sum (t_i^3 - t_i))}}$$

$$= \frac{\left| 560.5 - \frac{1}{2} \times 24 \times (68+1) \right| - 0.5}{\sqrt{\frac{24 \times 44}{12 \times 68 \times (68-1)} - [68^3 - 68 - (16^3 - 16 + 28^3 - 28 + 19^3 - 19 + 5^3 - 5)]}}$$

$$= 3.62$$

(3) 确定  $P$  值, 作出结论

由于  $u_{0.05/2} = 1.96$ ,  $u > u_{0.05/2}$ ,  $P < 0.05$ , 有统计学意义, 所以拒绝  $H_0$ , 接受  $H_1$ , 可认为两总体分布不同。

### 第三节 多个样本比较的秩和检验

#### 一、检验的基本步骤

上节讨论了两样本比较的秩和检验, 如果比较的样本多于两个, 则可以用本节的 Kruskal-Wallis  $H$  检验方法。

设有  $k$  个样本, 每个样本含量为  $n_i (i=1, \dots, k)$ , 总例数  $N = \sum_{i=1}^k n_i$ 。检验的具体步骤如下:

1. 建立假设检验

$H_0$ : 各抽样总体分布相同;

$H_1$ : 各抽样总体的分布不同或不全相同;

$\alpha=0.05$ 。

2. 计算统计量

(1) 编秩次, 将各组数据统一从小到大编秩次, 对相等的数值, 如果分属不同组时应取平均秩次。

(2) 求秩和, 分别计算各组的秩和  $T_i$ , 可用关系式  $\sum T_i = N(N+1)/2$  检验  $T_i$  的计算是否正确。

(3) 计算  $H$  值:

$$H = \frac{12}{N(N+1)} \sum \frac{T_i^2}{n_i} - 3(N+1) \quad (9-6)$$

当相同秩次较多时 (如等级资料), 采用校正的  $H_c$  值, 即

$$H_c = \frac{H}{1 - \frac{\sum (t_i^3 - t_i)}{N^3 - N}} \quad (9-7)$$

式中  $t_i$  为相同秩次的个数。  $H$  或  $H_c$  近似服从自由度  $\nu = k - 1$  的  $\chi^2$  分布。按  $\chi^2$  的界值表确定  $P$  的范围。



## 二、实例计算

**例 9.5** 研究白血病时,测定四组鼠脾 DNA 的含量,结果列于表 9-6 的 (1)、(3)、(5)、(7) 列 (已分别由小到大排列),试分析各组 DNA 含量有无差别?

(1) 建立假设检验

$H_0$ : 四组鼠脾 DNA 含量的总体分布相同;

$H_1$ : 四组鼠脾 DNA 含量不全相同;

$\alpha=0.05$ 。

(2) 计算统计量

四个样本总例数  $N=8+7+9+8=32$ 。将四个样本 32 个观察值统一由小到大编秩,见表 9-6 第 (2)、(4)、(6)、(8) 列。其中有两个测量值为 10.3mg,有三个测量值为 12.3mg,均取各自的平均秩次。

表 9-6 各组鼠脾 DNA 含量 (mg) 的秩和计算

正常脾		自发性白血病		移植白血病甲		移植白血病乙	
含量(1)	秩次(2)	含量(3)	秩次(4)	含量(5)	秩次(6)	含量(7)	秩次(8)
12.3	18	10.8	8	9.3	1	9.5	2
13.2	22	11.6	13	10.3	3.5	10.3	3.5
13.7	26	12.3	18	11.1	11	10.5	5
15.2	28	12.7	21	11.7	14	10.5	6
15.8	29	13.5	23	11.7	15	10.5	7
16.9	30	13.5	24	12.0	16	10.9	9
17.3	31	14.8	27	12.3	18	11.0	10
17.4	32			12.4	20	11.5	12
				13.6	25		
$n_1=8$	$T_1=216$	$n_2=7$	$T_2=134$	$n_3=9$	$T_3=123.5$	$n_4=8$	$T_4=54.5$

( $\sum T_i = 216 + 134 + 123.5 + 54.5 = 528$  与  $\frac{N(N+1)}{2} = \frac{32(32+1)}{2} = 528$  相等,计算无误)

由式 (9-6) 得:

$$\begin{aligned}
 H &= \frac{12}{N(N+1)} \sum \frac{T_i^2}{n_i} - 3(N+1) \\
 &= \frac{12}{32(32+1)} \left( \frac{216^2}{8} + \frac{134^2}{7} + \frac{123.5^2}{9} + \frac{54.5^2}{8} \right) - 3(32+1) = 19.94
 \end{aligned}$$

(3) 查表及结论

现  $k=4$ ,  $\nu=4-1=3$ , 查  $\chi^2$  界值表,  $\chi_{0.05,3}^2 = 7.81$ ,  $\chi^2 > \chi_{0.05,3}^2$ ;  $P < 0.05$  按  $\alpha=0.05$  水准,拒绝  $H_0$ , 接受  $H_1$ , 故可认为四组 DNA 含量有显著的差别。

**例 9.6** 表 9-7 第 (2)~(4) 列为霍乱菌苗不同途径免疫 21 天后血清抗体滴度水平测定结果,问各组间的血清抗体滴度水平之间差异是否有统计学意义?

(1) 建立假设检验

$H_0$ : 三组血清抗体滴度的总体分布相同;

$H_1$ : 三组血清抗体滴度水平不全相同;

$\alpha=0.05$ 。

(2) 计算统计量

本例做多组秩和检验,表 9-7 中第 (7) 列是各滴度的平均秩次。以各组的例数乘平均秩次,即得第 (8)~(10) 列滴度的各组秩和。



表 9-7 三种不同途径免疫 21 天后血清抗体滴度的分布与秩和计算

抗体滴度 (1)	气 雾			合计 (5)	秩次范围 (6)	平均秩次 (7)	秩 和		
	80 亿 (2)	100 亿 (3)	皮下注射 (4)				80 亿 (8)	100 亿 (9)	皮下 (10)
1:10	2	4	2	8	1~8	4.5	9	18	9
1:20	15	7	1	23	9~31	20	300	140	20
1:40	10	12	13	35	32~66	49	490	588	637
1:80	5	7	9	21	67~87	77	385	539	693
1:160	1	2	5	8	88~95	91.5	91.5	183	475.5
1:320	—	—	1	1	96	96	—	—	96
合计	33	32	31	96	—	—	1 275.5	1 468	1 912.5

依公式 (9-6) 得:

$$H = \frac{12}{N(N+1)} \sum \frac{T_i^2}{n_i} - 3(N+1)$$

$$= \frac{12}{96(96+1)} \left( \frac{1\ 275.5^2}{33} + \frac{1\ 468^2}{32} + \frac{1\ 912^2}{31} \right) - 3(96+1) = 11.36$$

由于各组相同例数较多, 故用公式 (9-7) 校正:

$$H_c = \frac{11.36}{1 - \frac{8^3 - 8 + 23^3 - 23 + 35^3 - 35 + 21^3 - 21 + 8^3 - 8}{96^3 - 96}} = 12.27$$

(3) 查表及结论

查  $\chi^2$  界值表,  $\nu = 3 - 1 = 2$ ,  $\chi_{0.01(2)}^2 = 9.21$ , 因为  $H_c > \chi_{0.01(2)}^2$ ,  $P < 0.01$ , 三组血清抗体滴度水平的差别有统计学意义。

## 练 习 题

### 一、单项选择题

- 对医学计量资料成组比较, 相对参数检验来说, 非参数秩和检验的优点是
  - 适用范围广
  - 检验效能高
  - 检验结果更准确
  - 充分利用资料信息
  - 不易出现假阴性错误
- 对于计量资料的比较, 在满足参数法条件下用非参数方法分析, 可能产生的结果是
  - 增加 I 类错误
  - 增加 II 类错误
  - 减少 I 类错误
  - 减少 II 类错误
  - 两类错误都增加
- 两样本比较的秩和检验, 如果样本含量一定, 两组秩和的差别越大说明
  - 两总体的差别越大
  - 两总体的差别越小
  - 两样本的差别可能越大
  - 越有理由说明两总体有差别
  - 越有理由说明两总体无差别
- 多个计量资料的比较, 当分布类型不清时, 应选择的统计方法是
  - 方差分析
  - Wilcoxon T 检验
  - Kruskal-Wallis H 检验
  - u 检验
  - $\chi^2$  检验
- 在一项临床试验研究中, 疗效分为“痊愈、显效、有效、无效”四个等级, 现欲



比较试验组与对照组治疗效果有无差别,宜采用的统计方法是

- A. Wilcoxon 秩和检验  
 B.  $2 \times 4$  列联表  $\chi^2$  检验  
 C. 四格表  $\chi^2$  检验  
 D. Fisher 确切概率法  
 E. 计算标准化率

## 二、计算与分析

1. 某医院测定 10 名受试者针刺膻中穴前后痛阈的数据,见下表,试分析针刺膻中穴前后痛阈值的差异有无统计学意义?

10 名受试者针刺膻中穴前后痛阈资料

编号	针刺前	针刺后	编号	针刺前	针刺后
1	600	610	6	1 125	1 425
2	600	700	7	1 400	1 350
3	685	575	8	750	825
4	1 050	600	9	1 000	800
5	900	600	10	1 500	1 400

2. 8 名健康男子服用肠溶醋酸棉酚片前后的精液中精子浓度检查结果如下表(服用时间 3 月),问服用肠溶醋酸棉酚片前后精液中精子浓度有无下降?

8 名健康男子服用肠溶醋酸棉酚片前后的精液中精子浓度(万/ml)

编号	服药前	服药后	编号	服药前	服药后
1	6 000	660	5	6 000	6 300
2	22 000	5 600	6	6 500	1 200
3	5 900	3 700	7	26 000	1 800
4	4 400	5 000	8	5 800	2 200

3. 雌鼠两组分别给以高蛋白和低蛋白的饲料,实验时间自生后 28 天至 84 天止,计 8 周。观察各鼠所增体重,结果如下表,问两种饲料对雌鼠体重增加有无显著影响?

两种饲料雌鼠体重增加量(g)

高蛋白组	低蛋白组	高蛋白组	低蛋白组
83	65	123	101
97	70	124	107
104	70	129	122
107	78	134	
113	85	146	
119	94	161	

4. 测得铅作业与非铅作业工人的血铅值( $\mu\text{g}/100\text{g}$ ),问铅作业工人的血铅值是否高于非铅作业工人?

非铅作业组	铅作业组	非铅作业组	铅作业组
5	17	12	43
5	18	13	44
6	20	15	
7	25	18	
9	34	21	

5. 用  $V_{K3}$  眼药水对近视眼患者作治疗,对照组用生理盐水作安慰剂,对两组的疗效进



行观察，结果如下表，试分析  $V_{K3}$  眼药水对近视眼患者的治疗是否有疗效？

$V_{K3}$  眼药水治疗近视眼患者的疗效观察

疗 效	$V_{K3}$ 眼药水组	生理盐水组
退步	8	20
不变	93	60
进步	11	10
恢复	4	1
合计	116	91

6. 对正常、单纯性肥胖及皮质醇增多症三组人的血浆皮质醇含量进行测定，其结果见下表，问三组人的血浆皮质醇含量的差异有无统计学意义？

三组人的血浆皮质醇测定值 (nmol/L)

正常人	单纯性肥胖人	皮质醇增多症	正常人	单纯性肥胖人	皮质醇增多症
0.4	0.6	9.8	3.1	4.1	14.8
1.9	1.2	10.2	3.7	5.0	15.6
2.2	2.0	10.6	3.9	5.9	15.6
2.5	2.4	13.0	4.6	7.4	21.6
2.8	3.1	14.0	7.0	13.6	24.0

7. 在针刺麻醉下，对肺癌、肺化脓症及肺结核三组患者进行肺部手术，效果分四级，结果见下表，试比较针刺麻醉对三组病人的效果有无差异？

三组患者肺部手术的针麻效果

针麻效果	肺癌	肺化脓症	肺结核
I	10	24	48
II	17	41	65
III	19	33	36
IV	4	7	8
合计	50	105	157

(刘启贵)

### Nonparametric Method

Some experiments generate responses that can be ordered or the ranked, but the actual value of the response cannot be measured numerically except with an arbitrary scale that you might create. It may be that you are able to tell only whether one observation is larger than another, Perhaps you can rank a whole set of observation without actually knowing the exact numerical values of the measurements. When the data do not appear to satisfy the parametric assumptions that have been mentioned in the previous chapters, an alternative method of analysis can be used-nonparametric statistical methods. Non-parametric methods generally specify the hypothesis in terms of the population distribution rather than parameters such as the means and the standard deviations. Parametric assumptions are replaced by more general assumptions about the population distribution, and the ranks of the observations are often used in place of the actual measurement.



Here we just give you the example of the one-sample Wilcoxon signed-rank test about the assumption and hypotheses. The purpose is to test the null hypothesis that a particular population has a hypothesized median  $M_0$ .

Assumptions:

- The continuous random variable  $X$  is symmetric about a median  $M$ .
- $X_1, X_2, \dots, X_n$  denotes a random sample of size  $n$  from the distribution of  $X$ .
- $M_0$  denotes an hypothesized median for  $X$ . Also note that when a distribution is symmetric then the median and mean are equal. So “mean” can be substituted for “median” throughout this discussion.

Hypotheses:

- *Two-tailed.*  $H_0 : M = M_0$  versus  $H_1 : M \neq M_0$
- *Left-tailed.*  $H_0 : M \geq M_0$  versus  $H_1 : M < M_0$
- *Right-tailed.*  $H_0 : M \leq M_0$  versus  $H_1 : M > M_0$

# 第十章 线性相关与回归

在前面章节学习的统计方法中,所涉及的是单变量资料的统计分析方法,如描述某一变量的统计特征或是对该变量在各组间的差别作出统计推断。但在实际医学研究中,常常需要对两个变量间的关系进行分析,例如人的年龄变化与血压、糖尿病人的血糖与胰岛素水平、孕妇的雌三醇水平与出生儿体重的关系等。对此可以采用回归与相关的统计分析方法。

## 第一节 线性相关

### 一、线性相关的基本概念

为直观地判断两个变量之间的关系,可以在直角坐标系中把每对  $(X_i, Y_i)$  值所代表的点绘出来,形成散点图 (scatter diagram)。例如 11 名男青年身高与前臂长资料绘制的散点图如图 10-1 所示。

若一个变量  $X$  由小到大 (或由大到小), 另一变量  $Y$  亦相应地由小到大 (或由大到小), 则两个变量的散点图呈直线趋势, 我们称这种现象为共变, 也就是两个变量之间有“相关关系”。男青年身高与前臂长散点呈直线趋势, 即男青年身材高, 前臂亦长, 说明身高与前臂长之间存在线性相关关系, 我们把这种关系称为直线相关 (linear correlation) 或简单相关 (simple correlation)。

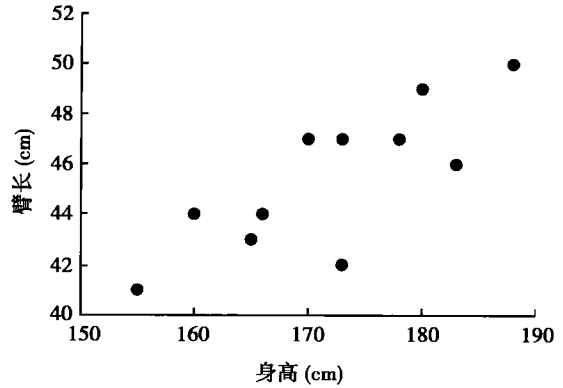


图 10-1 11 名男青年身高与前臂长资料绘制的散点图

直线相关用于双变量正态资料 (bivariate normal distribution)。直线相关的性质可由散点图直观地说明。散点图中点的分布即线性相关的性质和相关之间的密切程度, 可分为以下几种情况:

1. 正相关 (positive correlation) 如图 10-2 (1), 各点的分布呈现椭圆形,  $Y$  随  $X$  的增加而增加,  $X$  随  $Y$  的增加而增加。各点的排列越集中, 相关也就越密切, 当各点的分布如图 10-2 (2) 呈现直线型时,  $Y$  与  $X$  就完全正相关 (perfect positive correlation) 了。

2. 负相关 (negative correlation) 如图 10-2 (3), 各点的分布也呈现椭圆形,  $Y$  随  $X$  的增加而减少,  $X$  随  $Y$  的增加而减少。各点的排列越集中, 相关也就越密切, 当各点的分布如图 10-2 (4) 呈现直线型时, 则  $Y$  与  $X$  就完全负相关 (perfect negative correlation) 了。

在医学研究中, 完全正相关或完全负相关的情况几乎不存在。

3. 无相关 (zero correlation) 如图 10-2 中的 (5), 无论  $X$  增加还是减少,  $Y$  不受其影响, 反之,  $X$  也不受  $Y$  的影响。图 10-2 中的 (6)、(7) 两种情形, 虽然各点分布接近直线, 但直线与  $X$  轴或  $Y$  轴平行, 使  $X, Y$  互不受对方的影响, 所以无相关。

4. 图 10-2 (8) 中各点分布可能表示  $X$  与  $Y$  存在某种曲线相关, 但与线性相关已完全不同, 称为非线性相关。



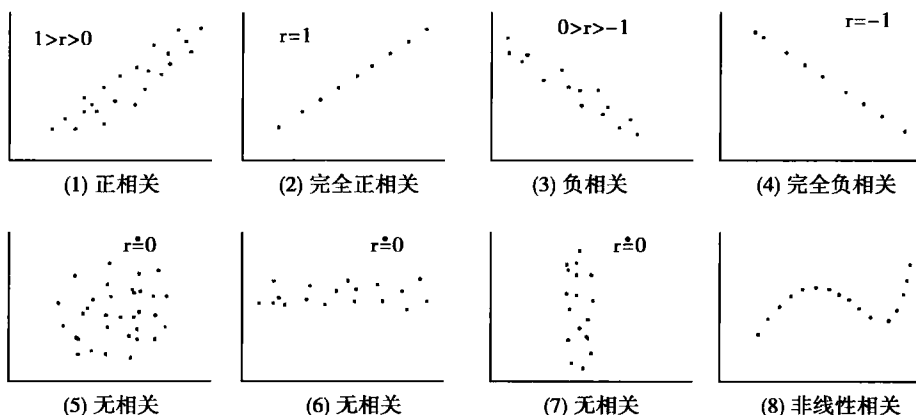


图 10-2 不同相关系数的散点图

用来分析这种呈现直线型相关关系的统计方法称为线性相关，它要求两个变量都是正态分布资料。线性相关可用统计量相关系数来定量地描述。

## 二、线性相关系数

### (一) 相关系数的意义

在分析两个变量  $X$  与  $Y$  之间关系时，常常要了解  $X$  与  $Y$  之间有无相关关系，相关是否密切，是呈正相关还是负相关。相关系数 (coefficient of correlation) 就是说明具有直线关系的两个变量间相关密切程度和相关方向的统计量。Pearson 相关系数的计算公式为：

$$r = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2 \sum(Y - \bar{Y})^2}} = \frac{l_{XY}}{\sqrt{l_{XX}l_{YY}}} \quad (10-1)$$

式中分子  $l_{XY}$  表示  $X$  与  $Y$  的离均差积和，分母  $l_{XX}$  表示  $X$  的离均差平方和， $l_{YY}$  表示  $Y$  的离均差平方和。

相关系数  $r$  没有测量单位，其数值为  $-1 \leq r \leq 1$ 。 $r$  值为正，表示正相关； $r$  值为负，表示负相关； $r$  值为 0，则称零相关即无直线关系。当  $r$  绝对值为 1 时，称完全相关。注意：生物界影响因素众多， $r$  的绝对值为 1 的机会极少，因而很少有完全相关，经常见到的是  $r$  值介于  $-1$  与  $1$  之间。在样本例数相等的情况下，计算出的相关系数的绝对值愈接近 1，相关愈密切；相关系数愈接近 0 时，相关愈不密切。

### (二) 相关系数的计算方法

公式 (10-1) 实际计算时，可简化为：

$$l_{XY} = \sum(X - \bar{X})(Y - \bar{Y}) = \sum XY - \frac{(\sum X)(\sum Y)}{n} \quad (10-2)$$

$$l_{XX} = \sum(X - \bar{X})^2 = \sum X^2 - \frac{(\sum X)^2}{n}$$

$$l_{YY} = \sum(Y - \bar{Y})^2 = \sum Y^2 - \frac{(\sum Y)^2}{n}$$

**例 10.1** 从男青年总体中随机抽取 11 名男青年组成样本，分别测量每个男青年的身高和前臂长，测量结果如表 10-1 所示，试计算身高与前臂长之间的相关系数。



表 10-1 11 名男青年身高与前臂长的测量结果 (cm)

编号	身高 X	前臂长 Y	XY	X <sup>2</sup>	Y <sup>2</sup>
1	170	47	7 990	28 900	2 209
2	173	42	7 266	29 929	1 764
3	160	44	7 040	25 600	1 936
4	155	41	6 355	24 025	1 681
5	173	47	8 131	29 929	2 209
6	188	50	9 400	35 344	2 500
7	178	47	8 366	31 684	2 209
8	183	46	8 418	33 489	2 116
9	180	49	8 820	32 400	2 401
10	165	43	7 095	27 225	1 849
11	166	44	3 174	28 561	2 116
合计	1 891	500	86 185	326 081	22 810

解:  $n=11$ ,  $\sum X=1\ 891$ ,  $\sum X^2=89\ 599$ ,  $\sum Y=500$ ,  $\sum Y^2=22\ 810$ ,  $\sum XY=86\ 185$ 。代入公式 (10-2), 得:

$$l_{XX} = \sum X^2 - \frac{(\sum X)^2}{n} = 326\ 081 - \frac{1\ 891^2}{11} = 1\ 000.909$$

$$l_{YY} = \sum Y^2 - \frac{(\sum Y)^2}{n} = 22\ 810 - \frac{500^2}{11} = 82.727$$

$$l_{XY} = \sum XY - \frac{(\sum X)(\sum Y)}{n} = 86\ 185 - \frac{1\ 891 \times 500}{11} = 230.455$$

按公式 (10-1) 计算相关系数:

$$r = \frac{230.455}{\sqrt{1\ 000.909 \times 82.727}} = 0.8009$$

结果  $r$  为正值, 表示前臂长与身高呈现正相关。由于  $r$  是样本统计量, 对于男青年总体的身高与前臂长之间的相关关系, 尚需进行进一步的假设检验。

### 三、线性相关系数的假设检验

例 10.1 计算的相关系数  $r=0.8009$  是根据样本资料计算出来的, 是样本相关系数。与前面讲的其他统计量一样, 根据样本资料计算出来的相关系数同样存在抽样误差。即使在一个  $X$  与  $Y$  无关总体中作随机抽样, 由于抽样误差的影响, 所得的样本相关系数也不会恰好等于零。因此要判断两个变量  $X$  与  $Y$  是否真的存在相关关系, 需根据  $r$  作总体相关系数  $\rho$  是否为零的假设检验。对此可以用  $t$  检验法, 其公式为:

$$t = \frac{|r-0|}{\sqrt{\frac{1-r^2}{n-2}}}, \text{ 自由度 } \nu = n-2 \quad (10-3)$$

式中分母为相关系数  $r$  的标准误。求得  $t$  值后查附表 2 的双侧  $t$  界值表得到  $P$  值, 然后根据  $P$  值的大小决定拒绝  $\rho$  为零的假设  $H_0$  还是不拒绝  $H_0$ 。

也可以根据计算出的  $r$  值按自由度  $\nu=n-2$  直接查附表 11 的  $r$  界值表得到  $P$  值, 若  $r > r_{\alpha/2, \nu}$  则可以认为两变量之间存在相关关系。

**例 10.2** 对例 10.1 所得的  $r$  值检验男青年身高与前臂长之间是否存在相关关系。

(1) 建立检验假设

$H_0: \rho=0$ , 即身高与前臂长之间不存在相关关系;

$H_1: \rho \neq 0$ , 即身高与前臂长之间存在相关关系;



$\alpha=0.05$ 。

(2) 计算统计量

$$t = \frac{|0.8009 - 0|}{\sqrt{\frac{1 - (0.8009)^2}{11 - 2}}} = 4.013, \nu = 11 - 2 = 9$$

(3) 确定  $P$  值, 作出结论

查  $t$  界值表, 得  $t_{0.005/2, 9} = 3.690$ ,  $t > t_{0.005/2, 9}$ ,  $P < 0.005$ , 故按  $\alpha = 0.05$  水准拒绝  $H_0$  接受  $H_1$ , 可以认为男青年身高与前臂长之间存在正相关关系。或查  $r$  界值表  $r_{0.005/2, 9} = 0.776$ ,  $r > r_{0.005/2, 9}$ ,  $P < 0.005$ , 结论同前。

#### 四、线性相关分析的注意事项

1. 分析两个变量之间有无相关关系可首先绘制散点图, 散点图呈现出直线趋势时, 再计算相关系数和作假设检验。

2. 相关系数的计算只适用于两个变量都服从正态分布的情形, 如果资料不服从正态分布, 应先通过变量变换, 使之正态化, 再根据变换值计算相关系数。

3. 依据公式计算出的相关系数仅是样本相关系数, 它是总体相关系数的一个估计值, 与总体相关系数之间存在着抽样误差, 要判断两个事物之间有无相关及相关的密切程度, 必须作假设检验。当检验拒绝了无效假设时, 才可以认为两个事物之间存在着相关关系, 然后再根据计算出的相关系数大小判断相关关系的密切程度。

4. 相关分析是用相关系数来描述两个变量间相互关系的密切程度和方向, 而两个事物之间的关系既可能是依存因果关系, 也可能仅是相互伴随的数量关系。因此, 决不可因为两事物间的相关系数有统计学意义, 就认为两者之间存在着因果关系, 要证明两事物间确实存在因果关系, 必须凭借专业知识加以阐明。

## 第二节 线性回归

### 一、线性回归的基本概念

在相关分析中, 通常不分  $X$  与  $Y$  何者为自变量和因变量。现在假设两个变量  $X$ 、 $Y$  中, 当一个变量  $X$  改变时, 另一个变量  $Y$  也相应地改变, 此时称  $X$  为自变量 (independent variable),  $Y$  为因变量 (dependent variable)。当这样的两个变量之间存在着直线关系时, 不仅可以用相关系数  $r$  表示变量  $Y$  与  $X$  线性关系的密切程度, 也可以用一个直线方程来表示  $Y$  与  $X$  的线性关系。在医学和生物学现象中, 许多变量间存在近似  $Y$  与  $X$  的线性关系, 这种关系有某种不确定性。例如, 儿童的年龄与体重, 一般来说儿童年龄越大其体重亦越大, 但很难说 8 岁男童体重一定是多少, 同样是 8 岁男童, 体重却有高有低。针对这种情况, 只能根据大量实测数据, 寻找出其规律性, 用一个直线方程来描述两个变量间依存变化的数量关系, 这样得出的直线方程叫做线性回归方程 (linear regression equation)。线性回归方程的形式为:

$$\hat{Y} = a + bX \quad (10-4)$$

其中:

$$b = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sum(X - \bar{X})^2} = \frac{l_{XY}}{l_{XX}} \quad (10-5)$$

$$a = \bar{Y} - b\bar{X} \quad (10-6)$$



$\hat{Y}$  是给定  $X$  时  $Y$  的估计值,  $b$  称为回归系数 (regression coefficient)。

回归系数  $b$  和常数项  $a$  是方程中两个待定的参数, 计算这两个参数的数学原理是最小二乘法 (method of least square), 该方法的原则是保证各实测点到回归直线的纵向距离的平方和最小, 并使计算出的回归直线最能代表实测数据所反映出的直线趋势。上述计算公式就是依据最小二乘法推导出的结果。

## 二、线性回归方程的计算

**例 10.3** 有人研究了温度对蛙的心率的影响, 得到了表 10-2 中所示的资料, 试进行回归分析。

表 10-2 温度 (°C) 与蛙的心率 (次/分)

对象	温度(X)	心率(Y)	XY	X <sup>2</sup>	Y <sup>2</sup>
1	2	5	10	4	25
2	4	11	44	16	121
3	6	11	66	36	121
4	8	14	112	64	196
5	10	22	220	100	484
6	12	23	276	144	529
7	14	32	448	196	1 024
8	16	29	464	256	841
9	18	32	576	324	1 024
10	20	34	680	400	1 156
11	22	33	726	484	1 089
合计	132	246	3 622	2 024	6 610

(1) 根据表 10-2 的数据绘制散点图 (图 10-3)。从绘制的散点图中可以看出温度与蛙的心率之间存在着明显的直线趋势, 所以可以进一步考虑建立二者之间的线性回归方程。

(2) 计算回归系数与常数项

在本例中:

$$\sum X = 132, \quad \sum X^2 = 2\,024, \quad \bar{X} = 12$$

$$\sum Y = 246, \quad \sum Y^2 = 6\,610, \quad \bar{Y} = 22.363$$

$$\sum XY = 3\,622$$

代入公式 (10-5) 和 (10-6) 得:

$$b = \frac{l_{XY}}{l_{XX}} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}} = \frac{3\,622 - \frac{(132)(246)}{11}}{2\,024 - \frac{132^2}{11}} = \frac{670}{440} = 1.523$$

$$a = \bar{Y} - b\bar{X} = 22.363 - 1.523 \times 12 = 4.087$$

则, 回归方程为:

$$\hat{Y} = 4.087 + 1.523X$$

(3) 作回归直线

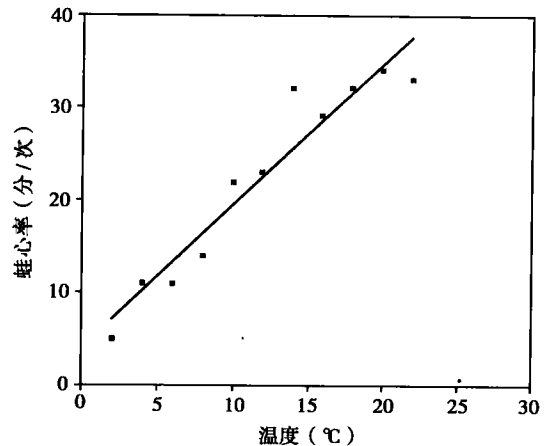


图 10-3 温度与蛙的心率之间关系散点图



按求得的回归方程, 在  $X$  实测值的范围内 (本例为 2~22) 任取两个相距较远的点  $A(X_1, \hat{Y}_1)$ 、 $B(X_2, \hat{Y}_2)$ , 连接  $A$ 、 $B$  两点即得到回归直线。本例可取  $X_1=3$ , 计算出  $\hat{Y}_1=8.65$ ;  $X_2=21$ , 计算出  $\hat{Y}_2=36.06$ ; 过  $(3, 8.65)$  和  $(21, 36.06)$  两点的连线即为所求的回归直线 (regression line), 参见图 10-3。

由图 10-3 可见, 温度升高, 蛙的心率也趋向加快, 蛙的心率  $Y$  依温度  $X$  的变化而变化, 散点呈直线趋势, 但并不完全在一条直线上。其中温度  $X$  是自变量, 是能够精确测量和严密控制的量; 蛙的心率  $Y$  是因变量, 是随机变化的。因为所有点并不完全在一条直线上, 说明除了受自变量的影响外, 还有其他因素对  $Y$  起作用。

线性回归方程可以应用于以下三个方面: ①分析两个变量之间是否存在线性依存关系; ②利用回归方程对因变量  $Y$  进行估计, 也就是把自变量  $X$  代入回归方程对因变量  $Y$  进行估计, 必要时可以作区间估计; ③利用回归方程进行统计控制, 实际就是利用回归方程进行逆运算, 通过控制自变量  $X$  取值来限定因变量  $Y$  在一定范围内波动。

### 三、线性回归方程的假设检验

由样本资料建立回归方程的目的是对两变量的回归关系进行推断, 也就是对总体回归方程  $\hat{Y}=\alpha+\beta X$  作假设检验。应该注意, 总体回归系数  $\beta$  是总体回归方程有无意义的关键, 如果  $\beta=0$ , 那么,  $\hat{Y}=\alpha$  是个常数, 无论  $X$  如何变化, 都不会影响  $\hat{Y}$ , 回归方程也就无实际意义。由样本资料计算的回归系数  $b$  和其他统计量一样, 它与总体回归系数间也存在抽样误差, 即如果总体回归系数为 0, 由样本资料计算的回归系数  $b$  也可能不为 0。所以对线性回归方程要进行假设检验, 就是要检验  $b$  是否为  $\beta=0$  的总体中的一个随机样本。该假设检验通常用方差分析或者  $t$  检验, 两者的检验效果等价。

#### (一) 方差分析

$Y$  值的变异可用式  $\sum(Y-\bar{Y})^2$  来反映, 而每个  $Y-\bar{Y}$  都可以分解成下式:

$$Y-\bar{Y}=(Y-\hat{Y})+(\hat{Y}-\bar{Y})$$

将此式两边平方然后展开, 得:

$$\begin{aligned}\sum(Y-\bar{Y})^2 &= \sum[(Y-\hat{Y})+(\hat{Y}-\bar{Y})]^2 \\ &= \sum(Y-\hat{Y})^2 + \sum(\hat{Y}-\bar{Y})^2 + 2\sum(Y-\hat{Y})(\hat{Y}-\bar{Y})\end{aligned}$$

其中

$$\begin{aligned}\sum(Y-\hat{Y})(\hat{Y}-\bar{Y}) &= \sum[Y-\bar{Y}-b(X-\bar{X})][\bar{Y}+b(X-\bar{X})-\bar{Y}] \\ &= \sum b(X-\bar{X})(Y-\bar{Y}) - \sum b^2(X-\bar{X})^2 \\ &= \sum b^2(X-\bar{X})^2 - \sum b^2(X-\bar{X})^2 = 0\end{aligned}$$

则

$$\sum(Y-\bar{Y})^2 = \sum(\hat{Y}-\bar{Y})^2 + \sum(Y-\hat{Y})^2 \quad (10-7)$$

此式中,  $\sum(Y-\bar{Y})^2$  为  $Y$  的离均差平方和, 又称总平方和, 表示因变量  $Y$  总的变异, 可用  $SS_{\text{总}}$  表示;  $\sum(\hat{Y}-\bar{Y})^2$  称回归平方和, 表示在  $Y$  总的变异中, 可以用  $X$  与  $Y$  的线性关系引起  $Y$  变异来解释的部分, 可用  $SS_{\text{回归}}$  表示;  $\sum(Y-\hat{Y})^2$  称剩余平方和或残差平方和, 说明除  $X$  对  $Y$  的线性影响之外的其他随机因素对  $Y$  的影响, 可用  $SS_{\text{剩余}}$  表示。

回归系数检验的基本思想是: 如果  $X$  与  $Y$  之间无线性回归关系, 则  $SS_{\text{回归}}$  与  $SS_{\text{剩余}}$  都是其他随机因素对  $Y$  的影响, 由此描写变异的  $MS_{\text{回归}}$  与  $MS_{\text{剩余}}$  应近似相等, 总体回归系数  $\beta=0$ , 反之,  $\beta \neq 0$ 。于是, 可用  $F$  检验对  $X$  与  $Y$  之间有无回归关系进行检验。

回归系数的假设检验可用下面简化公式计算:

$$SS_{\text{总}} = \sum(Y-\bar{Y})^2 = \sum Y^2 - \frac{(\sum Y)^2}{n} \quad (10-8)$$



$$SS_{\text{回归}} = \sum (\hat{Y} - \bar{Y})^2 = \sum (\bar{Y} + b(X - \bar{X}) - \bar{Y})^2 = b^2 \sum (X - \bar{X})^2$$

$$= b \frac{l_{XY}}{l_{XX}} l_{XX} = bl_{XY} = \frac{l_{XY}^2}{l_{XX}} \quad (10-9)$$

$$SS_{\text{剩余}} = SS_{\text{总}} - SS_{\text{回归}} \quad (10-10)$$

这三个平方和的自由度依次为:

$$\nu_{\text{总}} = n - 1, \quad \nu_{\text{回归}} = 1, \quad \nu_{\text{剩余}} = n - 2$$

可得:

$$MS_{\text{回归}} = \frac{SS_{\text{回归}}}{\nu_{\text{回归}}} \quad (10-11)$$

$$MS_{\text{剩余}} = \frac{SS_{\text{剩余}}}{\nu_{\text{剩余}}} \quad (10-12)$$

$$F = \frac{MS_{\text{回归}}}{MS_{\text{剩余}}} \quad (10-13)$$

下面对例 10.3 数据建立的回归方程进行假设检验。

### (1) 建立假设检验

在例 10.3 中

$$H_0: \beta = 0;$$

$$H_1: \beta \neq 0;$$

$$\alpha = 0.05.$$

### (2) 计算统计量

$$SS_{\text{总}} = \sum Y^2 - \frac{(\sum Y)^2}{n} = 6\,610 - \frac{246^2}{11} = 1\,108.54$$

$$SS_{\text{回归}} = \frac{l_{XY}^2}{l_{XX}} = \frac{670^2}{440} = 1\,020.23$$

$$SS_{\text{剩余}} = SS_{\text{总}} - SS_{\text{回归}} = 88.31$$

$$F = \frac{MS_{\text{回归}}}{MS_{\text{剩余}}} = \frac{SS_{\text{回归}}/\nu_{\text{回归}}}{SS_{\text{剩余}}/\nu_{\text{剩余}}} = \frac{1020.23/1}{88.31/9} = 103.97$$

### (3) 确定 P 值, 得出统计结论

查 F 界值表,  $\nu_{\text{回归}} = 1, \nu_{\text{剩余}} = 9, F_{0.01(1,9)} = 10.56, F > F_{0.01(1,9)}, P < 0.01$ , 拒绝  $H_0$ , 可以认为温度与蛙的心率之间存在线性回归关系。

表 10-3 方差分析表

变异来源	SS	$\nu$	MS	F	P
总变异	1 108.54	10			
回 归	1 020.23	1	1 020.23	103.97	<0.01
剩 余	88.31	9	9.81		

## (二) t 检验

回归系数 t 检验的基本思想与样本均数与总体均数比较的 t 检验类似。t 值可按下式计算:

$$t = \frac{|b - 0|}{s_b} \quad (10-14)$$

$$\nu = n - 2 \quad (10-15)$$

$$s_b = \frac{s_{Y \cdot X}}{\sqrt{l_{XX}}} \quad (10-16)$$

$$s_{Y \cdot X} = \sqrt{\frac{SS_{\text{剩余}}}{\nu_{\text{剩余}}}} = \sqrt{MS_{\text{剩余}}} \quad (10-17)$$

其中,  $s_b$  为样本回归系数的标准误, 反映样本回归系数与总体回归系数之间的抽样误



差。 $s_{y \cdot x}$ 为剩余标准差,表示因变量 $Y$ 值对于回归直线的离散程度。

对例 10.3 数据建立的回归方程进行  $t$  检验过程如下:

(1) 建立假设检验

在例 10.2 中,

$$H_0: \beta=0;$$

$$H_1: \beta \neq 0$$

$$\alpha=0.05$$

(2) 计算统计量

$$s_{y \cdot x} = \sqrt{\frac{88.31}{9}} = 3.13, \quad s_b = \frac{3.13}{\sqrt{440}} = 0.149$$

$$t = \frac{|1.523 - 0|}{0.149} = 10.22, \quad \nu = 11 - 2 = 9$$

(3) 确定  $P$  值, 作结论

根据  $\nu=9$ , 查  $t$  界值表,  $t_{0.01/2(9)} = 3.250$ ,  $P < 0.01$ , 拒绝  $H_0$ , 结论与  $F$  检验相同。

细心的读者可以发现统计量  $F$  与  $t$  之间存在着关系  $\sqrt{F} = t$ 。本例  $\sqrt{103.97} \approx 10.22$ 。

#### 四、线性回归分析的注意事项

1. 只有将两个内在有联系的变量放在一起进行回归分析才是有意义的。
2. 作回归分析时, 如果两个有内在联系的变量之间存在的是一种依存因果的关系, 那么应该以原因为  $X$ , 以结果为  $Y$ 。如果变量之间因果关系难以确定, 则应以易于测定、较为稳定或变异较小者为  $X$ 。
3. 在回归分析中, 因变量是一个服从正态分布的随机变量, 自变量既可以是随机变量 (II 型回归模型, 两个变量应该都服从正态分布), 也可以是给定的量 (I 型回归模型, 每个  $X$  取值相对应的  $Y$  服从正态分布)。如果  $Y$  不服从正态分布, 在进行回归分析前, 应先进行变量的变换以使变量符合要求。
4. 使用回归方程计算估计值时, 不要轻易把估计的范围扩大到建立方程时的自变量的取值范围之外。如例 10.3 中,  $X$  的取值范围为  $2 \sim 22$ , 计算估计值时  $X$  的取值最好在  $2 \sim 22$  之间。

### 第三节 线性相关和回归的区别与联系

相关与回归分析是研究两变量之间关系的两种不同的方法, 二者之间既有区别、又有联系。

#### 一、线性相关与回归的区别

1. 相关系数的计算只适用于两个变量都服从正态分布的情形, 而在回归分析中, 因变量是随机变量, 自变量既可以是随机变量 (II 型回归模型, 两个变量都应该服从正态分布), 也可以是给定的量 (I 型回归模型, 每个  $X$  取值相对应的变量  $Y$  服从正态分布)。

2. 线性相关表示两个变量之间的相互关系是双向的, 回归则反映两个变量之间的依存关系, 是单向的。

#### 二、线性相关与回归的联系

1. 如果对同一资料进行相关与回归分析, 则得到的相关系数  $r$  与回归方程中的  $b$  正负



号相同。 $r$  为正, 说明  $X$  与  $Y$  相关关系的方向是一致的,  $X$  增大 (或减小),  $Y$  也增大 (或减小);  $b$  为正, 说明  $X$  与  $Y$  之间由回归方程所确定的函数关系是递增的, 即  $X$  增大 (或减小),  $Y$  也增大 (或减小)。

2. 在相关分析中, 求出  $r$  后要要进行假设检验, 同样, 在回归分析中, 对  $b$  也要进行假设检验。实际上, 通过数学推导, 对同一样本可以得出  $r$  与  $b$  互相转化的公式, 同一样本的这两种假设检验也是等价的。因此, 由于  $r$  的假设检验可以直接查表, 较为简单, 所以也可以用其代替对  $b$  的假设检验。

3. 相关与回归可以互相解释。 $r$  的平方称为确定系数 (coefficient of determination), 可表示为:

$$R^2 = r^2 = \frac{l_{XY}^2}{l_{XX}l_{YY}} = \frac{l_{XY}^2/l_{XX}}{l_{YY}} = \frac{SS_{\text{回归}}}{SS_{\text{总}}}$$

此时说明了当  $SS_{\text{总}}$  不变的情况下, 回归平方和的大小决定了相关系数的大小,  $R^2$  反映出回归平方和在总平方和中所占的比重。 $R^2$  越接近 1, 回归效果越好。应用确定系数, 也可以从回归的角度对相关程度作进一步地了解, 例如,  $r=0.5$ ,  $R^2=0.25$ , 说明一个变量的变异仅有 25% 由另一变量所引起, 避免了对相关系数表示的相关程度做出不合适的解释。

## 第四节 等级相关

如果观测值是等级资料, 则可以用等级相关来表达两事物之间的关系。等级相关是分析  $X$ 、 $Y$  两变量等级间是否相关的一种非参数方法, 常用的等级相关方法是 Spearman 等级相关。此方法是将数据资料  $X$ 、 $Y$  按数值大小分别编成等级次序, 以等级次序作为新的变量代入公式, 计算等级相关系数 (rank correlation coefficient)  $r_s$ , 来说明两变量  $X$ 、 $Y$  间线性相关关系的密切程度和方向。与线性相关系数  $r$  一样, 等级相关系数  $r_s$  的数值亦在  $-1$  与  $1$  之间, 数值为正表示正相关, 数值为负表示负相关。

### 一、等级相关系数的计算

Spearman 等级相关系数  $r_s$  可由公式 (10-18) 计算, 即:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \quad (10-18)$$

式中,  $n$  表示样本含量;  $d$  表示  $X$ 、 $Y$  的秩次之差。

**例 10.4** 某医生做一种研究, 欲了解人群中氟骨症患病率 (%) 与饮用水中氟含量 (mg/L) 之间的关系。随机观察 8 个地区氟骨症患病率与饮用水中氟含量, 数据如表 10-4 中 (2)、(4) 两栏。试计算等级相关系数  $r_s$ 。

表 10-4 不同地区饮水中氟含量 (mg/L) 与氟骨症患病率 (%)

地区编号(1)	饮水中氟含量		氟骨症患病率		$d$ (6)=(3)-(5)	$d^2$ (7)=(6) <sup>2</sup>
	X(2)	秩次(3)	Y(4)	秩次(5)		
1	0.48	1	22.37	2	-1	1
2	0.64	2	23.31	3	-1	1
3	1.00	3	25.32	4	-1	1
4	1.47	4	22.29	1	3	9
5	1.60	5	35.00	5.5	-0.5	0.25
6	2.86	6	35.00	5.5	0.5	0.25
7	3.21	7	46.07	7	0	0
8	4.71	8	48.31	8	0	0





(1) 将  $X$ 、 $Y$  分别从小到大编秩, 见表 10-4 中的 (3)、(5) 两栏。若遇到相同观察值时, 则取平均秩次。如  $Y_5 = Y_6 = 35.00$ , 则  $Y_5$ 、 $Y_6$  分别标平均秩次  $(5+6)/2 = 5.5$ 。

(2) 计算差数  $d$ , 见 (6) 栏, 注意  $\sum d = 0$ 。

(3) 计算  $d^2$ , 见 (7) 栏,  $\sum d^2 = 12.5$ 。

(4) 代入公式 (10-18) 计算  $r_s$ 。

$$r_s = 1 - \frac{6 \times 12.5}{8(8^2 - 1)} = 0.85$$

注意: 此公式适用于资料中相同观察值例数不多的情形。若资料中相同观察值的例数较多时, 计算的结果偏差较大, 此时可由公式 (10-19) 计算校正的  $r_s$  值  $r'_s$ , 即

$$r'_s = \frac{(n^3 - n)/6 - (T_X + T_Y) - \sum d^2}{\sqrt{(n^3 - n)/6 - 2T_X} \sqrt{(n^3 - n)/6 - 2T_Y}} \quad (10-19)$$

式中  $T_X = \sum_{i=1}^k (t_i^3 - t_i)/12$ ,  $t_i$  为第  $i$  组相同秩次的个数,  $T_Y$  的计算方法相同;  $k$  为有相同秩次的组数。当  $T_X = T_Y = 0$  时, 公式 (10-18) 等于公式 (10-19)。

## 二、 $r_s$ 的假设检验

$r_s$  是由样本资料计算出的相关系数, 存在抽样误差问题, 故要推断总体中两变量间有无线性相关关系, 须进行假设检验。

(1) 建立假设检验

$$H_0: \rho_s = 0$$

$$H_1: \rho_s \neq 0$$

$$\alpha = 0.05$$

(2) 计算统计量

$$r_s = 1 - \frac{6 \times 12.5}{8(8^2 - 1)} = 0.85$$

(3) 确定  $P$  值, 作出结论。当  $n \leq 50$  时, 查附表 12 中的等级相关系数  $r_s$  界值表。如本例  $r_{s, 0.05/2, 8} = 0.738$ ,  $r_s > r_{s, 0.05/2, 8}$ ,  $P < 0.05$ , 在  $\alpha = 0.05$  水平上拒绝  $H_0$ , 接受  $H_1$ , 可以认为饮用水中氟含量与氟骨症患病率之间存在正相关关系。

注意: 当  $n > 50$  时, 可查附表 11 ( $r$  界值表), 取自由度  $\nu = n - 2$ , 查  $r_{\alpha/2, n-2}$ , 若  $r_s > r_{\alpha/2, n-2}$ ,  $P < \alpha$ , 则拒绝  $H_0$ , 若  $r_s < r_{\alpha/2, n-2}$ ,  $P > \alpha$ , 则接受  $H_1$ 。

## 练 习 题

### 一、单项选择题

1. 使用最小二乘法确定直线回归方程的原则是

- A. 各观测点距回归直线的纵向距离之和最小
- B. 各观测点距回归直线的横向距离之和最小
- C. 各观测点距回归直线的垂直距离平方和最小
- D. 各观测点距回归直线的纵向距离平方和最小
- E. 各观测点距回归直线的横向距离平方和最小

2. 两数值变量相关关系越强, 表示

- A. 相关系数越大
- B. 相关系数的绝对值越大
- C. 回归系数越大
- D. 回归系数的绝对值越大
- E. 相关系数检验统计量的  $t$  值越大



3. 回归分析的决定系数  $R^2$  越接近于 1, 说明
- A. 相关系数越大  
B. 回归方程的显著程度越高  
C. 应变量的变异越大  
D. 应变量的变异越小  
E. 自变量对应变量的影响越大
4. 两组资料作回归分析, 直线回归系数  $b$  较大的一组, 表示
- A. 两变量关系密切的可能性较大  
B. 检验显著的可能性较大  
C. 决定系数  $R^2$  较大  
D. 决定系数  $R^2$  可能大也可能小  
E. 数量依存关系更密切
5. 1~7 岁儿童可以用年龄 (岁) 估计体重 (市斤), 回归方程为  $\hat{Y}=14+4X$ , 若将体重换成国际单位 kg, 则此方程
- A. 常数项改变  
B. 回归系数改变  
C. 常数项和回归系数都改变  
D. 常数项和回归系数都不改变  
E. 决定系数改变

## 二、计算与分析

1. 12 名 20 岁女青年的身高与体重资料如下表, 试问女青年身高与体重之间有无相关关系?

12 名 20 岁女青年的身高与体重资料

编号	1	2	3	4	5	6	7	8	9	10	11	12
身高(cm)	164	156	172	172	177	180	166	162	172	167	158	152
体重(kg)	55	56	60	68	66	65	56	55	60	55	46	51

2. 某医师研究某种代乳粉价值时, 用大白鼠做实验, 得大白鼠进食量和体重增加量的资料如下, 试问大白鼠的进食量与体重的增加量之间有无关系? 能否用大白鼠的进食量来估计其体重的增加量?

大白鼠进食量和体重增加量的资料

动物编号	1	2	3	4	5	6	7	8	9	10	11
进食量(g)	820	780	720	867	690	787	934	679	639	820	780
增重量(g)	165	158	130	180	134	167	186	145	120	150	135

3. 测得 347 名 13 岁健康男童的身高和体重, 身高均数为 146.4cm, 标准差为 8.61cm, 体重均数为 37.04kg, 标准差为 6.67kg. 身高和体重的相关系数  $r=0.74$ , 试计算由身高推算体重的回归系数及由体重推算身高的回归系数。

4. 某防治所作病因研究, 对一些地区水质的平均碘含量 ( $\mu\text{g/L}$ ) 与地方性甲状腺肿患病率进行了调查, 结果如下表, 试问甲状腺肿患病率与水质中碘的含量有无相关关系?

局部地区水质的平均碘含量 ( $\mu\text{g/L}$ ) 与地方性甲状腺肿患病率

地区编号	1	2	3	4	5	6	7	8	9	10	11	12	13	14
患病率(%)	40.5	37.7	39.0	20.0	22.5	37.4	31.5	15.6	21.0	6.3	7.1	9.0	4.0	5.4
碘含量( $\mu\text{g/L}$ )	1.0	2.0	2.5	3.5	3.5	4.0	4.4	4.5	4.6	7.7	8.0	8.0	8.3	8.5

(闫岩)



## Linear Regression and Correlation

In this chapter we present analyses to determine the strength of relationship between two variables. In the case of linear regression, we will examine the amount of variability in one variable ( $Y$ , the dependent variable) that is explained by changes in another variable ( $X$ , the independent variable). Specifically, we will look for straight-line or linear changes in  $Y$  as  $X$  changes. Regression analysis is usually done in situations where we have controlled the  $X$  variable and measured it essentially without error. The goal of this analysis is to describe the functional relationship between two variables as a straight line, where  $X$  is the independent variable and  $Y$  is the dependent variable.

Correlation analysis is used when both variables are experimental and measured with error. It is more preliminary than regression analysis and generally measures the relationship between two variables of interest. We will present the Pearson product-moment correlation and one nonparametric analogy, the Spearman correlation coefficients, as methods of analysis in these situations.

# 第十一章 多元线性回归与多元逐步回归

在第十章中已介绍了用直线回归描述一个应变量与一个自变量之间的线性依存关系，但由于事物之间的联系往往是多方面的，特别在医学研究中，某个应变量往往受多个因素（自变量）的影响，如儿童身高不仅受年龄影响，而且受性别影响；肺活量的大小除与年龄、性别有关外，还受身高、体重以及胸围等因素的影响。如果这些因素与应变量之间的关系是线性的，则可以应用多元线性回归（multiple linear regression）来研究。多元线性回归是研究一个应变量与多个自变量之间线性依存关系的统计方法，可以对自变量的作用进行评价，也可以用作预测和判别。

## 第一节 多元线性回归

### 一、多元线性回归方程

设与应变量  $Y$  有关的自变量有  $k$  个，记为  $X_1, X_2, X_3, \dots, X_k$ 。观察到  $n$  例样本数据如表 11-1 所示。

表 11-1 多元回归分析数据格式

例号 $i$	$X_1$	$X_2$	...	$X_k$	$Y$
1	$X_{11}$	$X_{12}$	...	$X_{1k}$	$Y_1$
2	$X_{21}$	$X_{22}$	...	$X_{2k}$	$Y_2$
3	$X_{31}$	$X_{32}$	...	$X_{3k}$	$Y_3$
⋮	⋮	⋮	...	⋮	⋮
$n$	$X_{n1}$	$X_{n2}$	...	$X_{nk}$	$Y_n$

若应变量  $Y$  与  $X_1, X_2, X_3, \dots, X_k$  构成线性依存关系，则其多元线性回归方程可写为：

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_k X_k \quad (11-1)$$

其中， $\hat{Y}$  为  $X = (X_1, X_2, \dots, X_k)$  时，应变量  $Y$  的总体平均值的估计值； $b_0, b_1, b_2, \dots, b_k$  为待定系数。式中  $b_0$  相当于一元直线回归方程中的常数项  $a$ ，表示当所有自变量为 0 时应变量  $Y$  的总体平均值的估计值； $b_j (j=1, 2, \dots, k)$  为自变量  $X_j$  的偏回归系数（partial regression coefficient）， $b_j$  表示除  $X_j$  以外的其他自变量固定不变的情况下， $X_j$  每改变一个测量单位时所引起的应变量  $Y$  的平均改变量。

根据观察到的  $n$  例数据，代入公式 (11-1)，可得第  $i$  例的应变量  $Y_i$  的估计值  $\hat{Y}_i$ ，并得到估计残差：

$$SS_{\text{残差}} = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (11-2)$$

根据最小二乘法（method of least square）原理，求出使残差平方和  $SS_{\text{残差}}$  最小的  $b_1, b_2, \dots, b_k$ 。常数项  $b_0$  可用下式求出：

$$b_0 = \bar{Y} - (b_1 \bar{X}_1 + b_2 \bar{X}_2 + \dots + b_k \bar{X}_k) \quad (11-3)$$

由于多元线性回归方程的最小二乘法计算量较大，故往往借助如 SPSS 等统计软件完成。本章将主要通过实例介绍多元线性回归方程的基本分析步骤，并利用 SPSS 统计软件



完成计算。

**例 11.1** 20 名糖尿病人的血糖、胰岛素及生长素的测定值列于下表中，试建立血糖对于胰岛素及生长素的二元线性回归方程。

表 11-2 糖尿病人的血糖、胰岛素及生长素的测定值

病例号 <i>i</i>	血糖(mmol/L) <i>Y</i>	胰岛素(mu/L) <i>X</i> <sub>1</sub>	生长素(μg/L) <i>X</i> <sub>2</sub>
1	12.21	15.2	9.51
2	14.54	16.7	11.43
3	12.27	11.9	7.53
4	12.04	14.0	12.17
5	7.88	19.8	2.33
6	11.10	16.2	13.52
7	10.43	17.0	10.07
8	13.32	10.3	18.89
9	19.59	5.9	13.14
10	9.05	18.7	9.63
11	6.44	25.1	5.10
12	9.49	16.4	4.53
13	10.16	22.0	2.16
14	8.38	23.1	4.26
15	8.49	23.2	3.42
16	7.71	25.0	7.34
17	11.38	16.8	12.75
18	10.82	11.2	10.88
19	12.49	13.7	11.06
20	9.21	24.4	9.16

根据上述研究问题，采用多元线性回归方法，其回归方程为：

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2$$

根据表 11-2 资料，通过拟合该方程，采用最小二乘法估计出常数项  $b_0$  和偏回归系数  $b_1$ 、 $b_2$ ，使其残差平方和  $SS_{\text{残差}}$  达到最小。对本例由 SPSS 统计软件可得到如下表 11-3 的主要结果。

表 11-3 偏回归系数估计结果

变量	偏回归系数	标准误	<i>t</i>	<i>P</i>
常数项	17.011	2.472	6.880	0.000
<i>X</i> <sub>1</sub>	-0.406	0.094	-4.313	0.000
<i>X</i> <sub>2</sub>	0.098	0.116	0.843	0.411

由此得到回归方程为：

$$\hat{Y} = 17.011 - 0.406X_1 + 0.098X_2$$

## 二、多元线性回归方程的假设检验

由样本计算出的偏回归系数  $b_j$  是总体偏回归系数  $\beta_j$  的估计值 ( $j=1, 2, \dots, k$ )。如果总体偏回归系数等于 0，则由于存在抽样误差，仍可使样本估计的偏回归系数不等于 0。



与直线回归的情形相同,需对所建立的多元回归方程进行假设检验,以判断它是否具有统计学意义。多元线性回归方程的假设检验可以简单地分为模型检验和单个偏回归系数的检验。

### (一) 模型检验

可用方差分析方法来检验应变量  $Y$  与  $k$  个自变量之间是否存在线性回归关系。在多自变量情况下,应变量总离均差平方和  $SS_{\text{总}}$  可分解成回归平方和  $SS_{\text{回归}}$  和残差平方和  $SS_{\text{残差}}$  两个部分:

$$SS_{\text{总}} = SS_{\text{回归}} + SS_{\text{残差}} \quad (11-4)$$

其中,总离均差平方和为:

$$SS_{\text{总}} = \sum Y^2 - \frac{(\sum Y)^2}{n} \quad (11-5)$$

回归平方和为:

$$SS_{\text{回归}} = \sum (\hat{Y}_i - \bar{Y})^2 = \sum b_i L_{iY} \quad (11-6)$$

残差平方和为:

$$SS_{\text{残差}} = \sum (Y_i - \hat{Y}_i)^2 = SS_{\text{总}} - SS_{\text{回归}} \quad (11-7)$$

各离均差平方和所对应的自由度为  $\nu_{\text{总}} = n - 1$ ,  $\nu_{\text{回}} = k$ ,  $\nu_{\text{剩}} = n - k - 1$ 。从而得到  $F$  统计量计算公式为:

$$F = \frac{SS_{\text{回归}}/k}{SS_{\text{残差}}/(n-k-1)} = \frac{MS_{\text{回归}}}{MS_{\text{残差}}} \quad (11-8)$$

无效假设  $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$ , 即  $k$  个自变量对应变量  $Y$  无线性回归关系条件下,统计量  $F$  服从  $F_{\alpha(k, n-k-1)}$  分布,如果  $F \geq F_{\alpha(k, n-k-1)}$  则在  $\alpha$  水平上拒绝  $H_0$ , 认为  $k$  个自变量  $X$  与应变量  $Y$  之间存在线性回归关系,即回归方程有统计学意义。否则,不拒绝  $H_0$ , 即认为所有  $k$  个自变量与应变量之间无线性回归关系,即回归方程无统计学意义。

对于例 11.1, 模型检验如下:

$$H_0: \beta_1 = \beta_2 = 0$$

$$H_1: \beta_1 \neq 0 \text{ 和 } \beta_2 \neq 0 \text{ 至少有一个成立}$$

$$\alpha = 0.05$$

由 SPSS 统计软件可得到表 11-4 的模型检验结果。

表 11-4 回归方程的方差分析表

变异来源	SS	$\nu$	MS	F	P
回归	116.626	2	58.313	21.539	0.000
残差	46.025	17	2.707		
总变异	162.651	19			

由表 11-4 可知,  $SS_{\text{总}} = 162.65$ ,  $SS_{\text{回归}} = 116.63$ ,  $SS_{\text{残差}} = 46.03$ ,  $\nu_{\text{回}} = 2$ ,  $\nu_{\text{剩}} = 17$ ,  $F = 21.54$ ,  $P < 0.05$ 。从而拒绝  $H_0$ , 可以认为  $\beta_1$  和  $\beta_2$  不全为 0, 即所求回归方程有统计学意义。

### (二) 偏回归系数检验

模型的假设检验成立, 只能说明总的情况, 但对于回归方程中每个自变量来讲, 并不能保证它们都对应变量的线性关系具有统计学意义。因此, 需要对每个自变量(相应的偏回归系数)进行假设检验, 以免把作用不明显的自变量引入回归方程中。所用的检验方法为  $F$  检验或  $t$  检验法, 两种检验结果相同。

#### 1. $F$ 检验



在其他自变量存在于回归方程的条件下考究某一自变量  $X_j$  对应变量  $Y$  的回归效应。其无效假设为  $H_0: \beta_j = 0$ , 备择假设为  $H_1: \beta_j \neq 0$ 。具体步骤如下:

(1) 建立含  $k$  个自变量的回归方程, 并计算回归平方和  $SS_{\text{回归}}$  及残差平方和  $SS_{\text{残差}}$ ;

(2) 将拟检验的自变量  $X_j$  ( $j=1, 2, \dots, k$ ) 剔除, 重新建立一个含  $k-1$  个自变量 ( $X_1, X_2, \dots, X_{j-1}, X_{j+1}, \dots, X_k$ ) 的回归方程, 并计算不包含  $X_j$  的回归平方和  $SS_{\text{回归}(-j)}$ , 则  $U = SS_{\text{回归}} - SS_{\text{回归}(-j)}$  是在其他自变量已在回归方程中的条件下,  $X_j$  单独引起的回归平方和的改变量, 称其为  $X_j$  的偏回归平方和, 其自由度为 1。

(3) 用  $F$  统计量检验  $X_j$  的回归效应, 计算公式为:

$$F_j = \frac{U/1}{SS_{\text{残差}}/(n-k-1)} \quad (11-9)$$

(4)  $F_j$  服从  $F_{\alpha(1, n-k-1)}$  分布。根据  $\nu_1 = 1, \nu_2 = n-k-1$ , 查  $F$  值表, 如果  $F_j > F_{\alpha(1, n-k-1)}$ , 则在  $\alpha$  水平上拒绝  $H_0$ , 认为自变量  $X_j$  与应变量  $Y$  之间的线性回归关系具有统计学意义; 否则不拒绝  $H_0$ , 即尚不能认为自变量  $X_j$  与应变量  $Y$  之间存在线性回归关系。

如对于例 11.1, 将胰岛素  $X_1$  和生长素  $X_2$  全部纳入回归方程中, 可得到:

$$SS_{\text{回归}} = 116.6265, SS_{\text{残差}} = 46.0249$$

把  $X_1$  从回归方程中取出, 而单独建立  $X_2$  与  $Y$  的回归方程为:

$$\hat{Y} = 7.012 + 0.429X_2$$

此时, 有  $SS_{\text{回归}(-1)} = 66.2746$ 。从而  $U = 116.6265 - 66.2746 = 50.3519$ 。

若把  $X_2$  从回归方程中取出, 而单独建立  $X_1$  与  $Y$  的回归方程为:

$$\hat{Y} = 18.7961 - 0.4585X_1$$

此时, 有  $SS_{\text{回归}(-2)} = 114.7032$ 。从而  $U = 116.6265 - 114.7032 = 1.9233$ 。

由公式 (11-9) 有:

$$F_1 = \frac{50.3519/1}{46.0249/(20-2-1)} = 18.5982$$

$$F_2 = \frac{1.9233/1}{46.0249/(20-2-1)} = 0.7104$$

按  $\alpha = 0.05$  查  $F$  界值表, 得  $F_{0.05(1, 17)} = 3.59$ 。因此, 在  $\alpha = 0.05$  水平上, 可以认为胰岛素对血糖的线性回归关系有统计学意义, 而生长素对血糖的线性回归关系无统计学意义。所以应剔除  $X_2$ , 只建立  $X_1$  与  $Y$  的线性回归方程。

## 2. $t$ 检验

无效假设为  $H_0: \beta_j = 0$ , 备择假设为  $H_1: \beta_j \neq 0$ 。 $t$  检验方法比  $F$  检验要简便些, 检验统计量为:

$$t_{b_j} = \frac{b_j}{S_{b_j}} \quad (11-10)$$

其中,  $S_{b_j}$  为偏回归系数  $b_j$  的标准误。

由表 11-3 中的回归系数的估计结果, 按公式 (11-10) 可得:  $t_{b_1} = -4.31, P = 0.0005$ ;  $t_{b_2} = 0.84, P = 0.4110$ 。在  $\alpha = 0.05$  水平下, 认为血糖与胰岛素的线性回归关系有统计学意义, 而与生长素的线性回归关系无统计学意义。结论与  $F$  检验一致。

## 三、多个自变量的线性回归实例

**例 11.2** 为了研究影响糖尿病患者糖化血红蛋白 (HbA1c) 的主要危险因素, 某研究者收集了一部分糖尿病患者的糖化血红蛋白 ( $Y, \%$ )、年龄 ( $X_1$ , 岁)、体重指数 ( $X_2$ ,  $\text{kg}/\text{m}^2$ )、总胆固醇 ( $X_3$ ,  $\text{mmol}/\text{L}$ )、收缩压 ( $X_4$ ,  $\text{mmHg}$ ) 和舒张压 ( $X_5$ ,  $\text{mmHg}$ ) 等数据资料。现从中随机抽取了 20 例, 数据见表 11-5。试作多元线性回归分析。



表 11-5 20 例糖尿病患者的数据资料

编号	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	Y	编号	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	Y
1	49	32.19	6.0	148	86	7.6	11	53	23.43	7.1	161	86	7.5
2	67	24.77	2.7	151	98	7.4	12	46	30.56	2.9	146	79	7.3
3	64	25.24	7.0	151	80	7.4	13	59	25.19	6.0	158	80	7.3
4	66	24.26	4.8	157	87	7.2	14	76	27.26	5.4	124	85	6.9
5	68	30.28	3.5	136	83	7.3	15	63	23.93	6.7	133	89	7.5
6	48	26.18	7.6	137	87	7.6	16	74	24.94	7.9	166	82	7.9
7	66	26.36	5.9	157	91	7.5	17	52	22.82	5.3	149	71	7.3
8	47	32.07	5.7	157	89	7.7	18	64	24.34	2.5	126	93	6.8
9	64	28.44	6.1	154	82	7.3	19	54	25.44	2.6	151	83	6.9
10	75	30.65	6.9	137	86	7.7	20	78	28.98	7.2	147	74	7.5

对于表 11-5 的数据资料用 SPSS 统计软件计算的主要结果见表 11-6 和表 11-7。

表 11-6 回归方程的方差分析表

变异来源	SS	$\nu$	MS	F	P
回归	1.07906	5	0.21581	7.32	0.0015
残差	0.41294	14	0.02950		
总变异	1.49200	19			

表 11-7 偏回归系数估计结果

自变量	偏回归系数	标准误	t	P
常数项	3.87598	1.01115	3.83	0.0018
X <sub>1</sub>	-0.00153	0.00409	-0.37	0.7146
X <sub>2</sub>	0.03192	0.01348	2.37	0.0328
X <sub>3</sub>	0.10834	0.02451	4.42	0.0006
X <sub>4</sub>	0.0085	0.00368	2.31	0.0366
X <sub>5</sub>	0.01058	0.00663	1.60	0.1328

由表 11-6 可见,  $F=7.32$ ,  $P<0.01$ , 此回归方程有统计学意义。由表 11-7 可见, 自变量 X<sub>2</sub>、X<sub>3</sub>、X<sub>4</sub> 按  $\alpha=0.05$  水平有统计学意义, 但 X<sub>1</sub> 和 X<sub>5</sub> 无统计学意义。

这里应注意的是, 因为偏回归系数之间存在着相关性, 当从原方程剔除一个变量时, 其他变量, 特别是与它有密切关系的一些变量的偏回归系数都可能会受到影响。

有时需要比较各自变量对应变量的相对贡献大小, 但由于偏回归系数受变量度量衡的影响, 各自变量的度量衡又不同, 这样, 比较各偏回归系数的绝对值大小不易得出正确的结论。为此, 可以对偏回归系数进行标准化处理, 消除度量衡的影响, 即用标准化偏回归系数 (standardized partial regression coefficient) 反映各自变量对应变量的影响程度。标准化偏回归系数  $b'_j$  的计算公式为:

$$b'_j = b_j \left( \frac{S_j}{S_Y} \right) \quad (11-11)$$

式中  $S_j$  及  $S_Y$  分别为自变量 X<sub>j</sub> 及应变量 Y 的标准差,  $b_j$  为 X<sub>j</sub> 的偏回归系数。

#### 四、复相关系数和决定系数

1. 复相关系数 (multiple correlation coefficient) 又称多元相关系数或全相关系数,





表示回归方程中的全部自变量  $X$  与应变量  $Y$  的相关密切程度。和简单相关系数不同的是，复相关系数取值总为正值，在 0 与 1 之间，简记为  $R$ 。

$$R = \sqrt{\frac{SS_{\text{回归}}}{SS_{\text{总}}}} = \sqrt{1 - \frac{SS_{\text{残差}}}{SS_{\text{总}}}} \quad (11-12)$$

2. 决定系数 (coefficient of determination) 复相关系数的平方称决定系数，记为  $R^2$ ，用以反映线性回归方程能在多大程度上解释应变量  $Y$  的变异性。其定义为：

$$R^2 = \frac{SS_{\text{回归}}}{SS_{\text{总}}} = 1 - \frac{SS_{\text{残差}}}{SS_{\text{总}}} \quad (11-13)$$

$R^2$  是联系多元回归与相关的纽带，反映回归方程与数据的拟合程度，回归方程的拟合程度越好，残差平方和就越小，决定系数  $R^2$  越接近 1；反之， $R^2$  越接近 0。例 11.2 中  $R^2 = 0.7232$ ，说明由年龄、体重指数、总胆固醇、收缩压和舒张压可解释该样本糖化血红蛋白变异的 72.32%。此时， $R = 0.85$ 。

## 第二节 多元逐步回归

多元线性回归分析中当涉及的自变量较多时，这些自变量可能并不是全部都对应变量有显著影响，同时有些自变量之间也可能相关的。通常情况下，我们希望将有统计学意义的自变量引入回归方程，以使方程更为简单，容易解释；更重要的是把不显著的自变量排除后可以使残差的均方减小，有利于揭示其他自变量的作用。为此可以采用下面三种自变量筛选方法：

1. 向后法 (backward selection) 先建立一个包含所有自变量的回归方程，然后每次剔除一个偏回归平方和最小且无统计学意义的自变量，直到不能剔除时为止。这种方法在自变量不是很多的情况下效果较佳。

2. 向前法 (forward selection) 方程由一个自变量开始，每次引入一个偏回归平方和最大，且具有统计学意义的自变量，由少到多，直到无具有统计意义的变量可以引入为止。这种方法的主要问题是，先进入方程的变量有可能受到后进入方程变量的影响变得不显著。

3. 逐步法 (stepwise selection) 在前述两种方法的基础上，进行双向筛选的一种方法。即向前引入每一个新自变量之后都要重新对先前已选入的自变量进行检查，以评价其有无继续保留在方程中的价值。为此引入和剔除交替进行，直到无具有统计学意义的新变量可以引入也无自变量可以剔除时为止。

筛选自变量通常采用偏回归平方和的  $F$  检验。例如，方程中已引入了  $(m-1)$  个自变量，在此基础上考虑再引入变量  $X_j$ 。现记引入  $X_j$  后方程（即含  $m$  个自变量）的回归平方和为  $SS_{\text{回归}}$ ，残差为  $SS_{\text{残差}}$ ；之前含  $(m-1)$  个自变量（不包含  $X_j$ ）方程的回归平方和为  $SS_{\text{回归}(-j)}$ ，则  $X_j$  的偏回归平方和为

$$U = SS_{\text{回归}} - SS_{\text{回归}(-j)}$$

检验统计量为：

$$F_j = \frac{U/1}{SS_{\text{残差}}/(n-m-1)} \quad (11-14)$$

$F_j$  服从  $F_{\alpha(1, n-m-1)}$  分布。如果  $F_j = F_{\alpha(1, n-m-1)}$ ，则  $X_j$  选入方程；否则不入选。从方程中剔除无统计学作用的自变量，过程相反，但检验统计量一样。

有时，也采用校正决定系数 (adjusted determinant)  $R^2_{\text{校正}}$  作为判断标准。选择校正决定系数大者为“最优”方程。校正决定系数的计算公式为：



$$R^2 = 1 - \frac{MS_{\text{误差}}}{SS_{\text{总}} / (n - m - 1)} \quad (11-15)$$

式中  $m$  为方程中包含的自变量个数。

在进行逐步回归前，首先应确定检验每个自变量是否有统计学意义的  $F$  检验水平，以作为引入或剔除自变量的标准。 $F$  检验水平可以根据具体情况而定。一般而言，若使最终的回归方程中包含较多的自变量， $F$  水平不宜取得过高，可以适当放宽些，但也不能取得太低，否则就会失去筛选自变量的意义。一般可将  $F$  值定在  $\alpha$  为 0.05、0.10 或 0.20 水平上，对于回归方程的选入和剔除水准要求  $\alpha_{\text{选入}} \leq \alpha_{\text{剔除}}$ 。

虽然，逐步回归法最后所得的“最优”回归方程与所给的临界值的选择有关，不能保证所挑选的回归方程在任何准则下都是最优的。但从长期实践应用来看，一般情况下逐步回归分析法所挑选的回归方程至少是较理想的，所以目前被广泛使用。

对例 11.2 进行多元逐步回归，采用逐步法筛选自变量，选入水准为 0.10，剔除水准为 0.15，SPSS 软件计算过程及相应结果见表 11-8 至表 11-11。

表 11-8 多元逐步回归结果

模型	变量	偏回归系数	标准误	标准偏回归系数	$t$	$P$
第 1 步	常数	6.774	0.156	—	43.545	0.000
	$X_3$	0.110	0.027	0.693	4.079	0.001
第 2 步	常数	6.040	0.432	—	13.967	0.000
	$X_3$	0.111	0.025	0.697	4.350	0.000
	$X_2$	0.027	0.015	0.289	1.805	0.089
第 3 步	常数	4.799	0.667	—	7.193	0.000
	$X_3$	0.097	0.024	0.611	4.125	0.001
	$X_2$	0.031	0.014	0.330	2.287	0.036
	$X_4$	0.008	0.004	0.341	2.285	0.036

表 11-9 多元逐步回归方差分析表

模型	变异来源	SS	$\nu$	MS	$F$	$P$
第 1 步	回归	0.717	1	0.717	16.640	0.001
	误差	0.775	18	0.043		
	总变异	1.492	19			
第 2 步	回归	0.841	2	0.421	10.993	0.001
	误差	0.651	17	0.038		
	总变异	1.492	19			
第 3 步	回归	1.001	3	0.334	10.889	0.000
	误差	0.491	16	0.031		
	总变异	1.492	19			

表 11-10 多元逐步回归的复相关系数和决定系数表

模型	复相关系数	决定系数
第 1 步	0.693	0.480
第 2 步	0.751	0.564
第 3 步	0.819	0.671



表 11-11 未进入方程的自变量

模型	变量	标准化系数	t	P
第 1 步	$X_1$	-0.154	-0.889	0.386
	$X_2$	0.289	1.805	0.089
	$X_4$	0.298	1.803	0.089
	$X_5$	0.174	0.987	0.337
第 2 步	$X_1$	-0.128	-0.775	0.450
	$X_4$	0.341	2.285	0.036
	$X_5$	0.191	1.161	0.263
第 3 步	$X_1$	-0.043	-0.274	0.788
	$X_5$	0.232	1.626	0.125

最后得到的回归方程为：

$$\hat{Y} = 4.799 + 0.031X_2 + 0.097X_3 + 0.008X_4$$

根据上述结果，可以认为体重指数  $X_2$ 、总胆固醇  $X_3$  和收缩压  $X_4$  是影响糖化血红蛋白的主要因素。

### 第三节 多元线性回归的注意事项

1. 应变变量与自变量间具有线性依存关系是进行多元线性回归的前提条件，且这种线性依存关系必须是建立在专业依据的基础上。

2. 多元线性回归原则上要求应变变量是连续型变量，其预测值与实际观测值的差值（即残差）服从正态分布。另外，要求应变变量的观测值相互独立，对于传染性疾病等数据应谨慎处理。

3. 进行多元线性回归时，应注意样本量。一般应使样本量是自变量个数的 5 倍以上。

4. 进行多元线性回归时，决定系数  $R^2$  是一个重要的统计量。如果样本量已经足够大，但是决定系数  $R^2$  仍较小，应该考虑可能还有对应应变变量有实质影响的自变量未选入方程。

5. 在多元线性回归中，自变量通常是连续型变量，而对于无序分类变量必须数量化；等级变量则可根据实际情况选择直接引入回归方程或数量化后引入回归方程。数量化方法可以参考其他书籍。

6. 在逐步回归中，选择不同的检验水平，其回归方程的结果可能不一致，究竟应选择哪一个回归方程，必须结合医学的实际意义来确定。

7. 当自变量间存在较强的线性关系时，会使多元回归方程中的参数估计不准确，影响多元线性回归分析的结果。一般的处理方法是剔除造成共线性的自变量，再重新建立回归方程。

## 练 习 题

### 一、单项选择题

- 在疾病发生危险因素的研究中，采用多变量回归分析的主要目的是
  - 节省样本
  - 提高分析效率
  - 克服共线影响
  - 减少异常值的影响
  - 减少混杂的影响



2. 多元线性回归分析中, 反映回归平方和在应变量  $Y$  的总离均差平方和中所占比重的统计量是

- A. 简单相关系数                      B. 复相关系数                      C. 偏回归系数  
D. 回归均方                              E. 决定系数  $R^2$

3. 对同一资料作多变量线性回归分析, 若对两个具有不同个数自变量的回归方程进行比较, 应选用的指标是

- A. 决定系数                              B. 相关系数                              C. 偏回归平方和  
D. 校正决定系数                      E. 复相关系数

4. 多元线性回归分析, 对回归方程作方差分析, 检验统计量  $F$  值反映的是

- A. 所有自变量与应变量间是否存在线性回归关系  
B. 部分自变量与应变量间是否存在线性回归关系  
C. 自变量与应变量间存在的线性回归关系是否较强  
D. 自变量之间是否存在共线  
E. 回归方程的拟合优度

5. 在多元回归分析中, 若对某个自变量的值都乘以一个常数  $c$  ( $c \neq 0$ ), 则

- A. 偏回归系数不变、标准回归系数改变  
B. 偏回归系数改变、标准回归系数不变  
C. 偏回归系数与标准回归系数均不改变  
D. 偏回归系数与标准回归系数均改变  
E. 偏回归系数和决定系数均改变

## 二、计算与分析

1. 某种特殊营养缺乏状态下, 儿童年龄 (岁)、身高 (cm) 与体重 (kg) 测定结果见下表, ①试建立年龄、身高与体重的二元回归方程; ②对回归方程作检验; ③计算复相关系数与决定系数; ④计算年龄和身高的标准偏回归系数。

营养缺乏儿童年龄、身高、体重测定值

编号 $i$	1	2	3	4	5	6	7	8	9	10	11	12
身高 $X_1$	145	150	124	157	129	127	140	122	107	107	155	148
年龄 $X_2$	8	10	6	11	8	7	10	9	10	6	12	9
体重 $Y$	29	32	24	30	25	26	35	26	25	23	35	31

2. 有学者认为, 血清中低密度脂蛋白增高和高密度脂蛋白降低, 是引起动脉硬化的一个重要原因。现测量了 30 名动脉硬化疑似患者的载脂蛋白 AI、载脂蛋白 B、载脂蛋白 E、载脂蛋白 C、低密度脂蛋白中的胆固醇、高密度脂蛋白中的胆固醇含量, 资料如下表。

① 分别作  $Y_1$  和  $Y_2$  对  $X_1, X_2, X_3, X_4$  的多元线性回归分析。

② 作  $Y_2/Y_1$  对  $X_1, X_2, X_3, X_4$  的逐步回归分析, 并与前面的分析结果进行比较。

30 名动脉硬化疑似患者的观测资料

序号 $i$	载脂蛋白 AI	载脂蛋白 B	载脂蛋白 E	载脂蛋白 C	低密度脂蛋白	高密度脂蛋白
	(mg/dL)	(mg/dL)	(mg/dL)	(mg/dL)	(mg/dL)	(mg/dL)
	$X_1$	$X_2$	$X_3$	$X_4$	$Y_1$	$Y_2$
1	173	106	7	14.7	137	62
2	139	132	6.4	17.8	162	43
3	198	112	6.9	16.7	134	81



续表

序号 $i$	载脂蛋白 AI (mg/dL) $X_1$	载脂蛋白 B (mg/dL) $X_2$	载脂蛋白 E (mg/dL) $X_3$	载脂蛋白 C (mg/dL) $X_4$	低密度脂蛋白 (mg/dL) $Y_1$	高密度脂蛋白 (mg/dL) $Y_2$
4	118	138	7.1	15.7	188	39
5	139	94	8.6	13.6	138	51
6	175	160	12.1	20.3	215	65
7	131	154	11.2	21.5	171	40
8	158	141	9.7	29.6	148	42
9	158	137	7.4	18.2	197	56
10	132	151	7.5	17.2	113	37
11	162	110	6	15.9	145	70
12	144	113	10.1	42.8	81	41
13	162	137	7.2	20.7	185	56
14	169	129	8.5	16.7	157	58
15	129	138	6.3	10.1	197	47
16	166	148	11.5	33.4	156	49
17	185	118	6	17.5	156	69
18	155	121	6.1	20.4	154	57
19	175	111	4.1	27.2	144	74
20	136	110	9.4	26	90	39
21	153	133	8.5	16.9	215	65
22	110	149	9.5	24.7	184	40
23	160	86	5.3	10.8	118	57
24	112	123	8	16.6	127	34
25	147	110	8.5	18.4	137	54
26	204	122	6.1	21.0	126	72
27	131	102	6.6	13.4	130	51
28	170	127	8.4	24.7	135	62
29	173	123	8.7	19.0	188	85
30	132	131	13.8	29.2	122	38

(尹平)

### Multiple Regression Analysis

Regression analysis is the statistical methodology for predicting values of one response (dependent) variable from a collection of one or more predictor (independent) variable values. It can also be used for assessing the effects of the predictor variables on the responses.

The classical linear regression model is as follows:

Let  $X_1, X_2, X_3, \dots, X_k$  be  $k$  predictor variables thought to be related to a response variable  $Y$ .



The classical linear regression model states that  $Y$  is composed of a mean, which depends on a continuous manner on the  $X_i$ 's, and a random error  $\epsilon$ , which accounts for measurement error and the effects of other variables not explicitly considered in the model.

Specifically, the linear regression model with a single response takes the following form

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + e$$

The term 'linear' refers to the fact that the mean is a linear function of the unknown parameters  $\beta_1, \beta_2, \dots, \beta_k$ . The predictor variables may or may not enter the model as first-order terms.

## 第十二章 统计表与统计图

统计表 (statistical table) 和统计图 (statistical chart) 是统计描述的重要工具。在统计工作的整个过程, 从实验设计或调查设计开始, 直到最后分析总结, 为突出数据的说服力, 都要用统计表和统计图进行描述, 尤其在科研论文中, 表达统计结果及进行对比分析时, 应用更为广泛。目前统计表和统计图的制作主要使用统计软件, 因此本章对有关手工制作的方法不作介绍。

### 第一节 统计表

统计表是把统计资料 and 结果用表格的形式来表达, 其目的是简洁、清晰、直观, 方便对比和阅读。统计表有基本的制作要求, 并不是把数据放到表格里就形成统计表。同时也要注意, 不是所有的数据都需要制作统计表, 应该有选择性地对重点要表达的数据制作统计表。

#### 一、统计表的编制原则和结构

##### (一) 统计表的编制原则

1. 重点突出, 简单明了, 即一张表一般只表达一个中心内容和一个主题。若内容过多, 可分别制成若干张表。

2. 主谓分明, 层次清楚。统计表虽然是表格的形式, 但其内涵代表的是若干完整的文字语句, 因此, 主谓语的位置要准确。一般来说, 定语部分放在标题内, 主语放在表的左边作为横标目, 谓语放在右边作为纵标目, 横标目与纵标目交叉的格子放置数据, 从左向右读, 每一行便形成一个完整的句子。标目的安排及分组要层次清楚, 符合逻辑, 便于分析比较。

3. 数据表达规范、文字和线条尽量从简。

##### (二) 统计表的构成

从外形上看, 统计表可由标题、标目 (包括横标目、纵标目)、线条、数字和备注 5 部分构成, 如表 12-1。

表 12-1 某地 1988 年老年人口各年龄组死亡率和死亡百分比

年龄组(岁)	死亡率(‰)	死亡百分比(%)	年龄组(岁)	死亡率(‰)	死亡百分比(%)
60~	19.61	16.60	75~	76.05	20.94
65~	23.64	17.51	80~	115.45	25.51
70~	50.28	19.44			

1. 标题 它是统计表的总名称, 放在表的上方中间位置, 简明扼要地说明表的主要内容, 包括时间、地点和研究内容。表 12-1 的标题为“某地 1988 年老年人口各年龄组死亡率和死亡百分比”。若有多张表格, 标题前应加上标号, 以方便引用, 如“表 12-1”。如果表中所有数据指标的度量衡单位一致, 可以将其放在标题后面, 放于括号内。

2. 标目 横标目位于表的左侧, 说明各行数据的涵义, 如“60~”, “65~”等, 纵标目位于表头右侧, 说明各列数据的涵义, 如“死亡率”等。标目要文字简明, 有单位的标目要注明单位。总标目是对横标目和纵标目内容的概括, 横标目的总标目位于表的左上



角，如“年龄组（岁）”，纵标目的总标目在需要时才设置。

3. 线条 目前一般采用三横线表，表的顶线和底线把表的主要内容与标题分隔开，中间一条线把纵标目与数据分隔开，不宜使用竖线和斜线。如果某些标目或数据需要分层表示，可用短横线分隔。

4. 数字 用阿拉伯数字表示，位数对齐，小数位数一致。表内不留空格，无数字用“-”表示，缺失数字用“...”表示，最好以备注的形式说明。若数字是“0”，则填写“0”。

5. 备注 表中数据区一般不插入文字或其他说明，需要说明时可用“\*”号标出，将说明文字写在表格的下面。

## 二、统计表的种类

根据标目的层次复杂程度，统计表可以分成简单表和复合表。

1. 简单表 简单表的标目只有一个层次，主语按一个标志排列，一般用作横标目，而纵标目为统计指标名称，如表 12-1。每一行可用一个完整的句子表达，第一行可读为“某地 1988 年老年人口 60~年龄组死亡率为 19.61‰，死亡百分比为 16.60‰”。

2. 复合表 复合表的标目有两个以上层次，即主语按多个标志排列。复合表有两个或三个分组标志，一般把其中主要的或分项较多的一个作为横标目，而其余的则安排在纵标目与总标目上。如表 12-2。将老人的年龄和性别结合起来分组，可以分析不同年龄、性别的老年人的死亡率和死亡百分比。

表 12-2 某地 1988 年老年人口不同年龄、性别死亡率和死亡百分比

年龄组(岁)	男		女	
	死亡率(‰)	死亡百分比(%)	死亡率(‰)	死亡百分比(%)
60~	22.97	17.34	15.99	15.57
65~	29.46	18.65	17.89	15.93
70~	62.91	19.55	39.22	19.30
75~	177.05	20.28	59.48	21.86
80~	373.17	24.18	94.42	27.34

## 三、编制统计表的注意事项

统计表只要满足编制原则、结构合理，能够清晰地表达数据结果即可，实际应用中应注意以下问题：

1. 统计表不一定是唯一的 同一份数据经过标目重排或分解组合，可以根据需要构造不同形式的统计表，如表 12-2 也可以转换成表 12-3，便于将男女的死亡率、死亡百分比放在一起对比。也可以把复合表分解为多个简单表。

表 12-3 某地 1988 年老年人口不同年龄、性别死亡率和死亡百分比

年龄组(岁)	死亡率(‰)		死亡百分比(%)	
	男	女	男	女
60~	22.97	15.99	17.34	15.57
65~	29.46	17.89	18.65	15.93
70~	62.91	39.22	19.55	19.30
75~	177.05	59.48	20.28	21.86
80~	373.17	94.42	24.18	27.34





2. 避免内容混杂，表达不清，结构混乱 若标目层次多于四个以上，统计表就会变得繁冗，因此尽量少用。

**例 12.1** 某地 1974 年 111 例钩端螺旋体病患者发病季节、年龄和职业构成资料如表 12-4 所示，请按照编制统计表的基本要求，检查此表编制得是否合适，如不合适，加以修改。

表 12-4 流行病学有关的主要因素 (原表)

季 节							
8月上旬		8月中旬		8月下旬		9月上旬	
人数	%	人数	%	人数	%	人数	%
7	6.3	70	63.1	28	25.2	6	5.4
年 龄 (岁)							
15 以下		15~		45 及以上			
人数	%	人数	%	人数	%		
12	10.8	95	85.6	4	3.6		
职 业							
农 民		学 生		待业青年		其 他	
人数	%	人数	%	人数	%	人数	%
106	95.5	1	0.9	3	2.7	1	0.9

表 12-4 的缺点是：①标题不够明确，而且未注明时间、地点；②发病季节、年龄和职业三项指标无可比性，放在一起，内容繁杂；③标目设计不合理，“人数”、“%”多次重复，不便于比较分析；④表内文字和线条过多。因此可以将其分为三个表，即表 12-5、表 12-6 和表 12-7，具体修改如下：

表 12-5 某地 1974 年钩端螺旋体病患者发病季节分布 (修改表)

季节	人数	比率(%)
8月上旬	7	6.3
8月中旬	70	63.1
8月下旬	28	25.2
9月上旬	6	5.4
合计	111	100.0

表 12-6 某地 1974 年钩端螺旋体病患者发病年龄分布 (修改表)

年龄(岁)	人数	比率(%)
15 以下	12	10.8
15~	95	85.6
45 及以上	4	3.6
合计	111	100.0

表 12-7 某地 1974 年钩端螺旋体病患者发病职业分布 (修改表)

职业	人数	比率(%)
农民	106	95.5
学生	1	0.9
待业青年	3	2.7
其他	1	0.9
合计	111	100.0

**例 12.2** 某医院用麦芽根糖浆治疗急慢性肝炎 161 例，疗效资料如表 12-8，指出其缺点并加以改进。

表 12-8 麦芽根糖浆治疗急慢性肝炎疗效观察 (原表)

效果	有 效						无 效	
	小 计		近 期 痊 愈		好 转			
	例	%	例	%	例	%	例	%
总例数	108	67.1	70	43.5	38	23.6	53	32.9

此表的主要目的是表达用麦芽根糖浆治疗急慢性肝炎的疗效情况，但是其标题过于简单，主谓安排不合理，标目组合重复。可进行如下修改，见表 12-9。



表 12-9 某年某医院麦芽根糖浆治疗急性性肝炎的疗效观察 (修改表)

疗效	例数	百分比(%)
无效	53	32.9
好转	38	23.6
近期痊愈	70	43.5
合计	161	100.0

## 第二节 统计图

统计图是把数据资料以图示的形式表达,使数据对比更加形象、直观,一目了然。统计图利用点的位置、曲线的变化、直条的长短和面积的大小等各种几何图形来表达统计资料 and 指标,它将研究对象的特征、内部构成、相互关系、对比情况、频数分布等情况形象而生动地表达出来,更直观地反映出事物间的数量关系,更易于比较和理解。

### 一、统计图的制作原则和结构

#### (一) 统计图的制作原则

1. 必须根据资料的性质、分析目的选用适当的统计图,由于统计图不能精确地显示数据大小,所以经常需要与统计表一起使用。
2. 一个图通常只表达一个中心内容和一个主题,即一个统计指标。
3. 绘制图形应注意准确、美观,图线粗细应用适当,定点准确,不同事物用不同线条(实线、虚线、点线)或颜色表示,给人以清晰的印象。

#### (二) 统计图的结构

统计图通常由标题、图域、标目、图例和刻度 5 部分组成。

1. 标题 其作用是简明扼要地说明资料的内容、时间和地点,一般位于图的下方中央位置并编号,便于引用和说明。
2. 图域 即制图空间,除圆图外,一般用直角坐标系第一象限的位置表示图域,或者用长方形的框架表示。
3. 标目 分为纵标目和横标目,表示纵轴和横轴数字刻度的意义,一般有度量衡单位。
4. 图例 对图中不同颜色或图案代表的指标进行注释。图例通常放在横标目与标题之间,如果图域部分有较大空间,也可以放在图域中。
5. 刻度 即纵轴与横轴上的坐标。刻度数值按从小到大的顺序,纵轴由下向上,横轴由左向右。绘图时按照统计指标数值的大小,适当选择坐标原点和刻度的间隔。

#### (三) 统计图的种类

医学中常用的统计图有直条图、百分条图、圆图、普通线图、半对数线图、直方图、箱图和散点图等。应根据资料的性质和分析目的选择适当的图形。

## 二、常用统计图的绘制方法及基本要求

### (一) 直条图

直条图(bar chart)又称条图,用等宽直条的长短来表示相互独立的统计指标数值大小和它们之间的对比关系。直条图按直条是横放还是竖放分卧式和立式两种,按分组因素的多少分为单式和复式两种。主要适用于各自独立的统计指标。指标既可以是绝对数,也



可以是相对数。常用的条图有三种。

1. 单式条图 具有一个统计指标，一个分组因素。如根据表 12-10 绘成的图 12-1，描述按地区分组的发病率。

表 12-10 某省 1979 年四个地区脊髓灰质炎发病率 (1/10 万)

地区	发病率
甲	22.0
乙	4.5
丙	1.5
丁	1.0

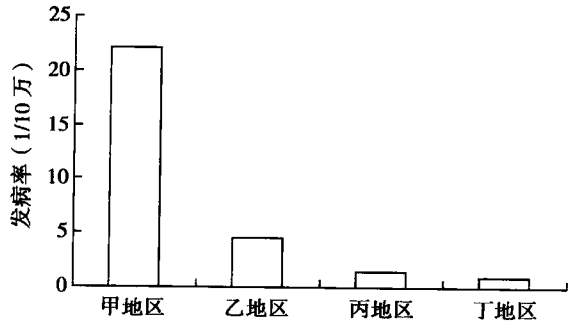


图 12-1 某省 1979 年四个地区脊髓灰质炎发病率

2. 复式条图 具有一个统计指标，两个分组因素。如根据表 12-11 绘成的图 12-2。

表 12-11 某地区 1952 年与 1972 年三种疾病的死亡率 (1/10 万)

死因	1952 年	1972 年
肺结核	163.2	27.4
心脏病	72.5	82.5
恶性肿瘤	52.5	183.5

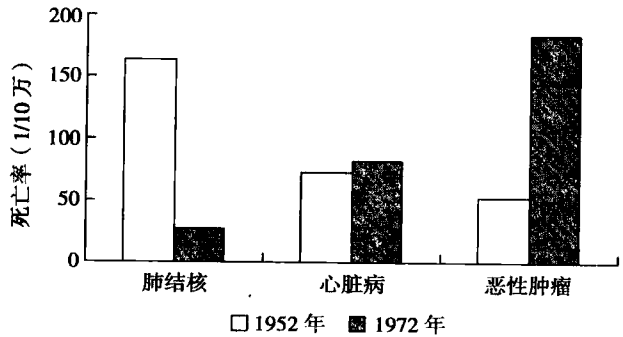


图 12-2 某地 1952 年与 1972 年三种疾病死亡率比较

绘制直条图时应注意：一般以横轴为基线，表示各个标志，纵轴表示各标志相应的值；纵轴尺度必须从“0”开始，而且要等距；直条的宽度要相等，直条之间的间隔要等距，通常与直条的宽度相等或略小。

## (二) 构成图

构成图常用于描述构成比的资料。常用的构成图有圆图和百分条图。

1. 圆图 (pie chart) 是一种构成图，把圆的总面积作为 100%，表示事物的全部，而圆内各扇形面积用来表示全体中各部分所占的比例。如根据表 12-12 单纯型老年性气管炎疗效绘成图 12-3。

表 12-12 复方猪胆囊治疗老年性气管炎近期疗效比较

近期疗效	单纯型		哮喘型	
	例数	构成比(%)	例数	构成比(%)
临床治愈	60	27.15	23	12.64
显效	98	44.34	82	45.06
有效	51	23.08	66	36.26
无效	12	5.43	11	6.04
合计	221	100.00	182	100.00



圆图绘制方法：以圆形的360°角为100%，将各组成部分的构成比(%)乘以3.6°，即得各组成部分应占的圆心角度数；再以某刻度(比如相当于时钟12点的位置)为起点，顺时针按圆心角度数大小或者自然顺序排列各个扇形。不同的扇形可以用不同的颜色或花纹区别，也可以简要注明文字和百分比。

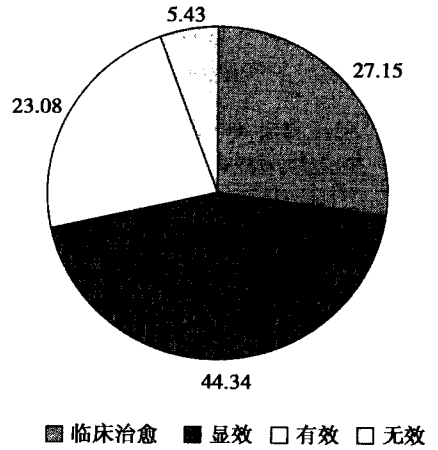


图 12-3 复方猪胆囊治疗单纯型老年性气管炎近期疗效比较

2. 百分比条图 百分比条图 (percentage chart) 也是一种构成图，用矩形条的长度表示100%，而用其中分割的各段表示各组成部分的百分比。如根据表 12-12 绘成的图 12-4。

百分条图绘制方法：绘一条等宽的水平直条作为100%，根据各部分所占百分比，按大小或资料的自然顺序把直条分成若干段，然后在直条的各分段上标出百分比。同一指标分层相互比较时，可以绘制多个直条。

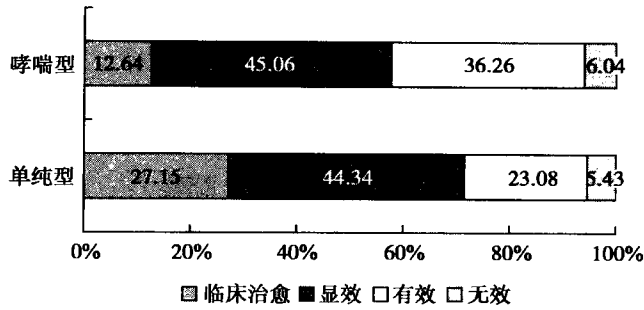


图 12-4 复方猪胆囊治疗老年性气管炎近期疗效比较

### (三) 线图

线图 (line chart) 是用线段的升降来表示指标的连续变化趋势，适用于描述一个变量随另一个变量变化的趋势。通常纵坐标是因变量或统计指标，横坐标是解释变量。普通线图的坐标都按算术尺度标记刻度。若纵坐标按对数尺度，横坐标按算术尺度，制作的线图为半对数线图 (semi-logarithmic linear chart)。

如根据表 12-13 绘成的图 12-5 为普通线图。

年份	男	女
1968	50.19	37.54
1969	42.97	25.00
1970	45.37	27.88
1971	44.42	25.10
1972	35.59	24.08
1973	38.31	24.10
1974	25.29	16.00

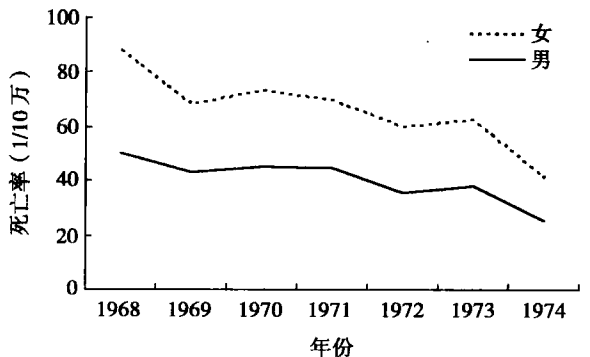


图 12-5 某地 1968~1974 年结核病死亡率



对于相对数指标，用普通线图有时难以准确对比表达变化趋势，可以使用半对数线图。如将表 12-14 的数据绘制成普通线图（图 12-6），呈现出白喉死亡率线条的坡度比伤寒、副伤寒死亡率线条变化快；如果绘制成半对数线图（图 12-7），则白喉死亡率与伤寒、副伤寒死亡率变化快慢速度相差不大。

表 12-14 某地 1949~1958 年白喉、伤寒、副伤寒的死亡率 (1/10 万)

年份	白喉		伤寒、副伤寒	
	死亡率	对数值	死亡率	对数值
1949	3.3	0.5185	1.1	0.0414
1950	5.9	0.7709	0.9	-0.0458
1951	8.4	0.9243	1.9	0.2788
1952	3.9	0.5911	1.0	0.0000
1953	2.5	0.3979	0.7	-0.1549
1954	1.5	0.1761	0.6	-0.2218
1955	3.3	0.5185	0.6	-0.2218
1956	1.1	0.0414	0.2	-0.6990
1957	1.0	0.0000	0.3	-0.5229
1958	0.6	-0.2218	0.05	-1.3010

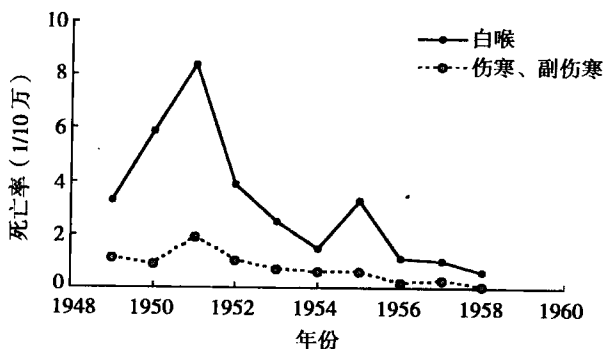


图 12-6 某地 1949~1958 年白喉、伤寒、副伤寒死亡率线图

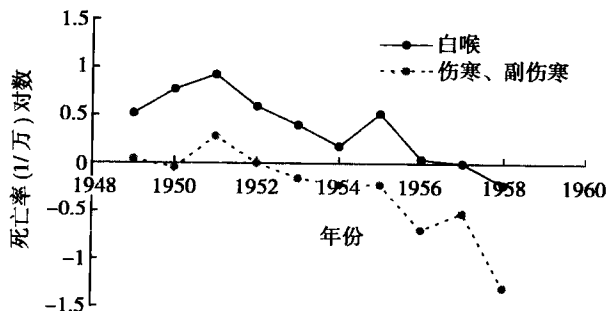


图 12-7 某地 1949~1958 年白喉、伤寒、副伤寒死亡率半对数线图

绘制线图时应注意：横轴代表分组标志，纵轴代表统计指标。横轴和纵轴的刻度都可以不从“0”开始；用短线依次将相邻各点连接即得线图，不应将折线描成光滑曲线；线



图中只有一条线，称为单式线图。若有两条及以上的线条，称为复式线图；在绘图时，需要注意纵横轴比例，比例不同，给人的印象也不同。

#### (四) 直方图

直方图 (histogram) 用直条矩形面积代表各组频数，各矩形面积总和代表各组频数的总和。它主要用于表示连续变量频数分布情况，如根据频数表 12-15 绘成频数分布图 12-8。

表 12-15 某市某年 150 名 3 岁女孩身高的频数分布

组段(cm)	频数	组段(cm)	频数
80~	1	94~	24
82~	3	96~	17
84~	8	98~	10
86~	10	100~	6
88~	19	102~	2
90~	23	104~106	1
92~	26		

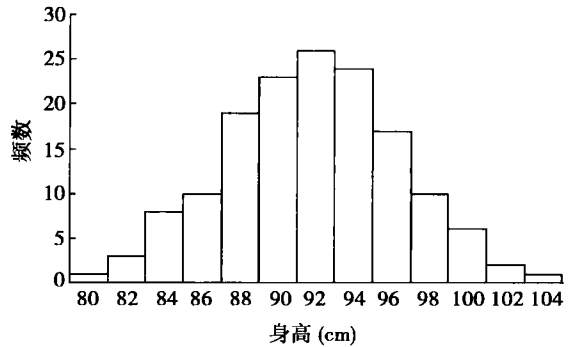


图 12-8 某市某年 150 名 3 岁女孩身高频数分布

绘制直方图时应注意：纵轴的刻度必须从“0”开始，而横轴的刻度按实际范围制定；各矩形的高度为频数或频率，宽度为组距。如果各组段的组距不同要调整各矩形的高，矩形高度 = 组段频数 / 组距。

#### (五) 箱图

箱图 (box plot) 用于比较两组或多组资料的集中趋势和离散趋势，箱图的中间横线表示中位数，箱子的长度表示四分位数间距，两端分别是  $P_{75}$  和  $P_{25}$ ，箱图最外面两端连线表示最大值和最小值，或  $P_{95}$  和  $P_{5}$ ，异常值另作标记。箱子越长表示数据离散程度越大。中间横线若在箱子中心位置，表示数据分布对称，中间横线偏离箱子正中心越远，表示数据分布越偏离中位数。根据表 12-16 绘成的图 12-9，可以看出，第 I 种饲料喂养得到的数据离散程度最大，第 III 种饲料喂养得到的数据基本对称且中位数最大。

表 12-16 10 窝大鼠分别用四种饲料喂养 60 天后体重增加量 (g)

窝别	饲料			
	I	II	III	IV
1	33	45	77	30
2	44	45	68	40
3	62	55	76	41
4	52	50	81	41
5	44	42	84	36
6	41	35	78	48
7	79	40	75	33
8	55	50	73	45
9	53	50	73	45
10	48	41	72	32

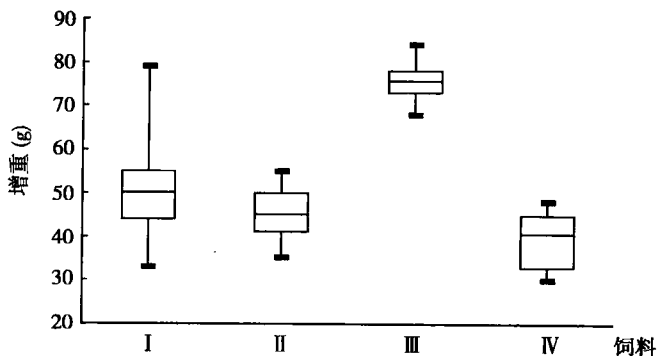


图 12-9 10 窝大鼠分别用四种饲料喂养 60 天后体重增加量箱图

### (六) 散点图

散点图 (scatter chart) 用点的密集程度和变化趋势来表示两指标之间的直线相关关系。如根据表 12-17 绘成图 12-10。

表 12-17 15 名 8 岁男孩身高与坐高

身高 X(cm)	坐高 Y(cm)
120.2	65.9
122.5	67.0
127.5	67.0
126.3	68.6
117.8	63.1
127.7	69.1
117.4	61.0
126.1	68.5
131.8	71.3
120.0	64.0
121.1	66.3
130.0	71.9
120.0	68.3
116.3	61.5
124.3	67.1

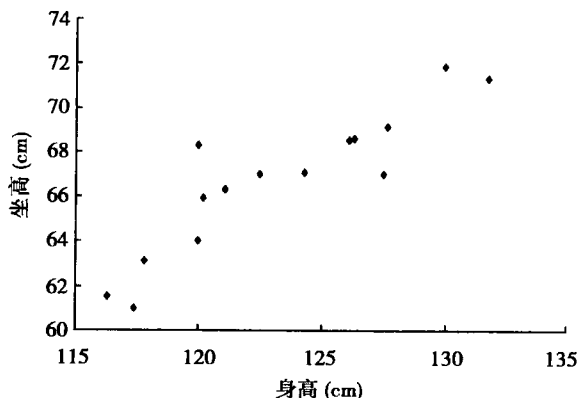


图 12-10 15 名 8 岁男孩身高与坐高散点图

绘制散点图时应注意：纵轴和横轴各代表一种事物，横轴代表自变量，纵轴代表因变量；纵轴和横轴的起点，不一定从“0”开始。

## 练 习 题

### 一、单项选择题

1. 统计表的主要作用是
  - A. 便于形象描述和表达结果
  - B. 客观表达实验的原始数据
  - C. 减少论文篇幅
  - D. 容易进行统计描述和推断
  - E. 代替冗长的文字叙述和便于分析对比
2. 描述某疾病患者年龄 (岁) 的分布, 应采用的统计图是
  - A. 线图
  - B. 条图
  - C. 百分条图
  - D. 直方图
  - E. 箱式图
3. 高血压临床试验分为试验组和对照组, 分析考虑治疗 0 周、2 周、4 周、6 周、8 周



血压的动态变化和改善情况，为了直观显示出两组血压平均变动情况，宜选用的统计图是

- A. 半对数图    B. 线图    C. 条图    D. 直方图    E. 百分条图

4. 研究三种不同麻醉剂在麻醉后的镇痛效果，采用计量评分法，分数呈偏态分布，比较终点时分数的平均水平及个体的变异程度，应使用的图形是

- A. 复式条图    B. 复式线图    C. 散点图    D. 直方图    E. 箱式图

5. 研究血清低密度脂蛋白 LDL 与载脂蛋白 B-100 的数量依存关系，应绘制的图形是

- A. 直方图    B. 箱式图    C. 线图    D. 散点图    E. 条图

## 二、改表和绘图

1. 某地调查脾肿大和疟疾临床分型的关系、程度与血片查疟原虫结果列表如下，此表有何缺点，请改进。

项目 脾肿程度	血膜阴性	血膜阳性				合 计		
		恶性疟		间日疟		例数	%	
		例数	%	例数	%			
脾肿者	174	28	12.6	20	9.04	222	48	21.6
脾 I	105	8	6.6	9	7.40	122	17	13.9
脾 II	51	14	20.0	5	7.10	70	19	27.1
脾 III	15	6	23.1	5	19.20	26	11	42.3
	3	0	0.0	1	25.00	4	1	25.0

2. 试根据下表资料绘制适当统计图形。

某地 1975 年 839 例正常人发汞值分布资料 ( $\mu\text{g/g}$ )

组段	0~	0.2~	0.4~	0.6~	0.8~	1.0~	1.2~	1.4~	1.6~2.2	合计
例数	133	193	190	111	83	34	43	16	36	839

3. 根据下表分别绘制普通线图和半对数线图，并说明两种统计图形的意义。

某地某年食管癌年龄别发病率 (1/10 万)

年龄(岁)	男	女	年龄(岁)	男	女
40~	4.4	2.1	65~	50.2	16.4
45~	7.2	3.3	70~	68.5	12.5
50~	7.3	4.5	75~	86.2	19.9
55~	6.9	5.5	80~	97.0	15.2
60~	19.3	6.7			

(张丕德 郭德成)

### Statistical Graphs and Tables

In this chapter, four basic chart types are introduced, which are bar chart, pie chart, Drop-line Chart and histogram.

**Bar Chart:** A simple bar chart presents a summary variable divided into discrete categories, making it easy to compare categories visually. A glance at a bar chart tells you the relative sizes of the quantities represented by the bars, which can be drawn vertically or horizontally.





**Pie Chart:** What proportion does each group contribute to the total? A pie chart demonstrates the contribution of parts to a whole. Each slice of the pie belongs to a group defined by a category of a variable. The size of a slice is proportional to the frequency of its group or to a function of a summary variable. For example, one pie chart shows the number of employees in each department, making it easy to compare relative department sizes.

**Drop-line Chart:** A dot or line chart usually presents a summary function plotted against a categorical variable. A point is plotted for each value of the category variable. A line chart is often used when the category axis is defined by an ordered series, although the points can represent any set of values. The data can be connected with lines and dot symbols can be displayed at the plotted points. Multiple-line charts can have drop lines connecting the categories. In a drop-line chart, the dots in each category are connected by line.

**Histograms:** An investigator wants to know how data are distributed across a range of value. A histogram shows the distribution of values in a quantitative variable by dividing the range into equally spaced intervals and plotting the count of cases in each interval as a bar. Instead of the count of cases, you can also plot the percentage.

## 第十三章 医学实验设计与诊断试验的评价

无论是临床研究或动物实验都需要根据研究的目的,制定总的研究方案,也称为研究设计。一个周密的研究设计,能用比较经济的人力、物力和时间,使研究因素的效应得到充分体现,获得准确和客观的结论,达到研究的目的。本章将对医学实验设计(medical experimental design)研究作一个较概括的介绍。另外,随着早期诊断、无创诊断的需求提高,大量新的诊断仪器设备及新的诊断方法的出现,对诊断试验方法的正确评价也日益迫切,本章也对诊断试验的评价方法作一介绍。

### 第一节 医学实验设计的三个要素

医学实验设计包括许多内容,但有三个基本要素:受试对象,处理因素和实验效应。

#### 一、受试对象

受试对象是根据研究目的而确定。根据实验的要求,受试对象可以是人,动物,细胞和血标本等。选择受试对象应注意以下几个问题:

1. 受试对象的纳入标准和排除标准 在医学研究中,首先根据研究目标确定研究对象,并且对研究对象的条件作严格的规定,即制定严格的纳入标准和排除标准,以保证他(它)们的同质性。例如,在研究某新药对慢性胃溃疡的治疗效果时,规定“胃溃疡十二个月以上未治愈者”为慢性胃溃疡病人,受试对象中就应当排除十二指肠溃疡以及胃溃疡发病等于和小于十二个月的患者。又例如,动物实验时,要求受试动物为同种属、同性别和体重相近,因为这些条件不一致,可能影响实验结果。

2. 选择对处理因素敏感性强的受试对象 如研究某新药对冠心病患者心律紊乱的疗效,则应选择心律紊乱经常发作的冠心病患者作为研究对象,而排除那些几天或几个月发作一次短暂心律失常的患者,因为他们在观察时期内对该药的治疗效果很可能没有反应。

3. 选择依从性好的患者作研究对象 即选择那些能够服从试验安排并坚持合作的患者。否则若不依从的数量较大,就会影响研究结果的准确性,导致研究结果出现偏倚。

4. 要注意医学伦理学问题 研究者应以患者的利益为最高准则,由受试者或其法定代理人在知情同意书上签字并注明日期,执行知情同意过程的研究者或其代表也需在知情同意书上签名并注明日期。在执行过程中,当科研与治疗发生冲突时,要服从医疗上的需要,这样做才符合医学伦理学要求。在动物实验中,也要注意善待动物。

#### 二、处理因素

处理因素是根据研究目的而施加于受试对象的干预措施,如研究某降压药的效果,降压药是处理因素,高血压患者为受试对象。有时,也将受试对象本身所具有的某些特征看作处理因素,如研究某工厂职业病的男女发病率的差异,则性别是处理因素。处理因素也可以是客观存在的因素,比如不同季节太湖中绿藻含量的差异,不同季节就是该实验的处理因素。

在设计中应将处理因素的实施方法规定得具体、细致。其目的是使处理因素在实验过



程中保持稳定性,有利于分析处理因素与实验结果之间的关系。例如:应明确给出受试对象的处理次数、每次的剂量,对化学试剂应规定试剂的生产单位、批号、纯度和配制规范,观察药物疗效应具体规定采集病人尿、血样本的时间及间隔等。

在实验过程中,除处理因素外也能使受试对象产生效应的因素属于“非处理因素”,每一项医学实验研究都可能受到非处理因素的影响。由于它可能干扰识别处理因素与效应间的关系,成为“混杂因素”,在进行实验之前,研究者非常有必要经过周密的思考,作出合理的实验设计来控制这些非处理因素,以排除非处理因素对处理因素的干扰。有时,即使设计时能够考虑到这些非处理因素,但实践中对这些非处理因素很难加以严格控制,其实验结果仍不免会受到影响,对此可以用统计学方法加以分析和校正,分离出非处理因素对实验结果的作用,发现真正的实验效应,作出客观的结论。

### 三、实验效应

实验效应是处理因素作用于受试对象的客观反映和结果,它通过观察指标来表达。因此,如何正确选择观察指标是关系研究成败的关键。

#### (一) 主观指标和客观指标

选用的指标要尽量客观,客观指标不易受主观因素影响。如电生理以及大多数临床化验数据(如血糖、甘油三酯、胆固醇等)都是客观指标。如果研究的指标是通过受试对象回答或描述症状得到,以及研究人员自行判断或医生通过体征检查所获得的结果大都是主观指标。主观指标易受观察对象和研究人员的心理状态、启发暗示和感官差异的影响,在实验设计中应尽量不作为主要指标。有时一些指标看似客观,实际上却受主观因素的影响,如眼底镜检查,可由检查人员掌握标准不同而异。在实验设计中,若一定要采取主观指标,那么就必须采取措施以减少或消除主观因素的影响。主观指标应尽可能满足临床诊断标准和疗效评价标准。

#### (二) 选择灵敏度和特异度高的指标

灵敏度是指某处理因素存在时所选指标能够表现出的效应程度,灵敏度不高就不能充分地反映出处理因素的作用,如治疗慢性脂肪肝选用谷丙转氨酶、谷草转氨酶作为疗效指标就不敏感,因为这些指标都是急性指标,并不是所有慢性脂肪肝患者这些指标都表现为异常,而B超或CT影像则随病情的变化而变化,可以较敏锐地反映出治疗的作用。特异度是指某处理因素不存在时所选指标不显示处理效应的程度。特异度高的指标更容易揭示处理因素的作用,不易受混杂因素的干扰,减少假阳性率。因此,所选指标最好同时具有较高的灵敏度与特异度。

#### (三) 观察指标的准确度和精密度

准确度是指研究结果与相应测定事物真实情况符合或接近的程度,主要受系统误差的影响。精密度是指相同条件下对同一对象的某指标进行重复测量时,观测值与其均值的接近程度,主要受随机因素的影响。精密度有两重含义:一是仪器设备本身具有的精度,另一种是指在反复测试过程中应控制的精度,它可以通过制定操作规范和技术培训进行控制。无论是准确度还是精密度,其水平高低都显示了研究工作质量的好坏,一般要将其控制在专业规定的容许范围内。

## 第二节 医学实验研究设计的几个原则

在医学实验研究中,由于存在各种非实验因素的干扰,使实验结果产生某种误差。为



了能够控制或避免各种误差，我们作研究设计时，必须遵循实验设计的基本统计原则，即对照原则、随机原则、重复原则和均衡原则。

## 一、对照原则

对照原则 (principle of control) 是指在实验中应设立对照组，其目的是通过与对照组效应对比鉴别出实验组的效应与大小。在对照组与实验组中，除处理因素不同外，应该使非处理因素基本齐同。医学研究中常用的对照组形式主要有以下几种，我们应该根据研究目的和研究条件等因素的不同，选择适当的对照形式。

1. 空白对照 (blank control) 不给予对照组的受试对象任何处理因素，也不给予任何干预，但其他实验条件和实验组相同。例如，进行可疑致癌物的研究，实验组给予干预，对照组不给予任何处理，比较两组的差异。空白对照主要用于无损伤、无刺激的实验性研究。

2. 自身对照 (self control) 同一受试对象既作对照者，又作试验者接受处理因素，这种对照形式叫自身对照。例如，研究某种药的降压效果，记录同一患者服用降压药前后的血压值，服用前的血压就可以作为对照。

对于服用药前后的自身对照，最好用临床住院患者，并且控制在同一个疗程之内，保证患者的依从性。

3. 标准对照 (standardized control) 对照组的干预采用现有标准方法或常规方法。如有些临床试验，出于伦理上的考虑，对照组可用传统疗法或已经使用的规范疗法，实验组使用新的疗法。有时，人们把公认的某种标准或正常值作为对照也称为标准对照。

4. 实验对照 (experimental control) 指与实验组操作相同，但不给予研究的处理因素，通常用于有损伤和刺激的动物实验。如实验组的动物注射某种药物，对照组注射无药理作用的生理盐水。

5. 安慰剂对照 (placebo control) 是一种假干预对照，主要用于临床试验。安慰剂是外形、气味、包装与受试药物一样的物质，但不含有有效药物成分，其目的是消除患者心理作用的影响。使用安慰剂要特别注意医学伦理学问题，应以患者的利益为最高准则，由受试者或其法定代理人在知情同意书上签字并注明日期，执行知情同意过程的研究者或其代表也需在知情同意书上签名并注明日期。

6. 相互对照 (mutual control) 指各实验组之间互为对照。例如，比较几种不同药物或同一种药物不同剂量对某种疾病的疗效，目的是比较其疗效差别时，可以使用相互对照。

## 二、随机化原则

随机化原则 (principle of random), 随机化是应用统计方法进行资料分析的基础。随机化就是在抽样时做到研究总体中任何一个个体都有同等的机会被抽取到各组样本中。其目的是要尽量使实验组与对照组之间各种非处理因素分布一致，提高组间均衡性，减少偏倚，将一些不可控制的因素的总效应归并到总的实验误差中。随机化 (randomization) 的方法很多，请参阅本书第一章。

## 三、重复的原则

重复原则 (principle of replication) 是指要在相同实验条件下重复进行多次观察。重复是消除非处理因素影响的又一重要方法，表现为样本量的大小和重复次数的多少。由于各种影响因素的存在，不同受试对象对同一处理因素的反应不同，表现为其效应指标的数值不同，只有在大量重复实验的条件下，实验的效应才能反应其真正的客观规律性；反



之, 样本量小, 重复性差, 若个别实验效应出现极端值, 指标就不够稳定, 会产生偏倚的结论。从统计学的角度讲, 观察次数越多, 从样本计算出的频数或平均数等统计量越接近总体参数。但重复原则并非要求无限大的样本, 样本量太大, 工作量也大, 增大人力物力的消耗, 难于控制实验条件, 工作不容易做得仔细, 影响研究结果的质量, 造成实验结果的不稳定。样本量究竟需要多大, 要根据样本含量的估计的公式进行计算。具体计算方法可参阅有关的专业书籍。

#### 四、均衡原则

均衡原则 (principle of balance) 是指实验中的各组之间除处理因素不同外, 要尽可能控制非处理因素, 使实验组与对照组在非处理因素方面基本一致, 具有齐同可比性。实验中的非实验因素较多, 一般应考虑:

1. 如做动物实验, 动物的种属、体重、月龄、性别应该保持一致, 如为临床试验, 除考虑年龄, 性别, 职业外, 还应考虑受试对象的病情, 病程及以前治疗情况。要尽可能在非处理因素方面保持一致。

2. 各实验组之间实验条件和实验环境应保持一致, 如同一实验, 各组的实验处理尽量由同一个人在同一时间完成, 或同一地点进行, 尽量减少误差。

3. 临床实验要尽量用盲法试验 (blind trial)。盲法试验有两种: 一种是单盲法, 即医生知道用什么药或疗法, 而病人不知道; 另一种是双盲法, 即主治医师及病人双方都不知道所用的是何种药物。这样做的目的是使疗效的评价保持客观性, 避免受心理影响。但实际操作上, 盲法试验存在一定的困难, 如病人是否愿意作为受试对象, 医生是否愿意配合。尤其有些药物有副作用, 一些患者不适合服用, 若不告诉主治医师, 容易出现意外。用盲法试验时, 要特别注意医学伦理学问题。研究者应以患者的利益为最高准则, 由受试者或其法定代理人在知情同意书上签字并注明日期, 执行知情同意过程的研究者或其代表也需在知情同意书上签名并注明日期。

### 第三节 诊断试验的评价

临床诊断包括各种实验室检查诊断、影像诊断和仪器诊断等, 各种方法的诊断价值如何, 必须通过诊断试验确定。好的诊断试验方法将对临床诊断的正确性和疾病的治疗效果起重要的作用。因此, 对于不同疾病不断发现新的诊断试验方法是医学进步的重要标志, 同时, 如何对新的诊断试验方法的正确性、可靠性进行评价也是一个重要问题。正确性是指新的诊断试验方法能否真实地反映健康或疾病的状况, 其评价指标主要有敏感度、特异度、预测值等。此外还有新的诊断试验方法与金标准诊断的一致性评价等。

#### 一、诊断试验评价的方法

##### (一) 诊断试验评价方法的金标准

诊断试验评价 (evaluation of diagnostic testing) 中应该有一个金标准 (gold standard), 临床上常用的金标准有:

- (1) 病理学检查 (组织活检或尸体解剖);
- (2) 外科学中的手术确诊;
- (3) 通过随访获得肯定的结论;
- (4) 影像诊断中使用的标准模具。

有时, 也使用同专业中大家公认的最好的诊断试验方法。



## (二) 新的诊断试验方法

- (1) 新的生化检测方法；
- (2) 新的 DNA 检测指标；
- (3) 新的影像检测方法；
- (4) 新的医学检测仪器的检测指标；
- (5) 根据多种指标提出的新的综合指标；
- (6) 根据多种指标提出的新的数学模型、判别准则。

## (三) 诊断试验评价的四格表

根据“金标准”诊断的结果，把受试对象分为实际患病组（阳性）和未患病组（阴性），用待评价的新诊断试验方法对这些受试对象进行诊断，得出“阳性”和“阴性”的结果。两种诊断方法的诊断结果有一致的，也有不一致的，于是可得四格表，如表 13-1 所示。

表 13-1 新的诊断试验方法与金标准比较的四格表

金标准诊断	新诊断试验方法	
	阳性	阴性
有某种疾病	真阳性( <i>a</i> )	假阴性( <i>b</i> )
没有某种疾病	假阳性( <i>c</i> )	真阴性( <i>d</i> )

## 二、诊断试验的评价指标

### (一) 基本指标

1. 敏感度 (sensitivity) 又被称为真阳性率，表示实际患病者且被待评价的诊断方法诊断为患者的概率，反映了待评价的诊断方法检出患者的能力，该值愈大愈好。用  $S_e$  表示敏感度的估计值，实际计算公式为：

$$S_e = \frac{a}{a+b} \quad (13-1)$$

2. 特异度 (specificity) 又称为真阴性率，表示实际未患病者且被待评价的诊断方法诊断为非患者的概率，反映了待评价的诊断方法检出非患者的能力，该值也愈大愈好。用  $S_p$  表示特异度的估计值，实际计算公式为：

$$S_p = \frac{d}{c+d} \quad (13-2)$$

3. 总符合率 (total consistent rate) 又称总的正确率 ( $\pi$ )，表示待评价诊断方法的诊断结果与金标准诊断结果符合的程度，它是真阳性与真阴性之和占总被检测人数的百分数，它反映正确诊断患者和非患者的能力，其计算公式为：

$$\hat{\pi} = \frac{a+d}{a+b+c+d} \quad (13-3)$$

4. 误诊率 (mistake diagnostic rate) 又被称为假阳性率 ( $\alpha$ )，表示实际未患病者中被待评价的诊断方法诊断为患者的概率，反映了待评价的诊断方法将非患者误诊为患者的判断错误的可能性，该值愈小愈好。其计算公式为：

$$\hat{\alpha} = 1 - S_p = \frac{c}{c+d} \quad (13-4)$$

5. 漏诊率 (omission diagnostic rate) 又被称为假阴性率 ( $\beta$ )，表示实际患病者中被待评价的诊断方法诊断为非患者的概率，反映了待评价的诊断方法将实际是患者诊断为非患者的判断错误的可能性，该值愈小愈好。其计算公式为：



$$\hat{\beta} = 1 - S_c = \frac{b}{a+b} \quad (13-5)$$

6. Youden 指数 (Youden index) 表示扣除了误诊率 (YI)、漏诊率之后的率, 该值愈大愈好, 其实际计算公式为:

$$YI = 1 - \hat{\alpha} - \hat{\beta} = S_c + S_p - 1 \quad (13-6)$$

7. 比数积 (odds product) 是把敏感度和特异度综合考虑的统计指标 (OP), 该值也愈大愈好, 它适用于大样本, 并且  $a$ 、 $b$ 、 $c$ 、 $d$  中没有为零的情况, 其计算公式为:

$$OP = \frac{S_c}{1-S_c} \cdot \frac{S_p}{1-S_p} = \frac{ad}{bc} \quad (13-7)$$

## (二) 预测值

预测值 (predictive value) 又被称为预告值, 表示待评价的诊断方法诊断结果与标准诊断方法诊断结果相符合的概率, 该值愈大愈好。

1. 阳性预测值 (positive predictive value) 表示待评价的诊断方法诊断为阳性的病例中实际真是患者的概率, 用  $PV_+$  表示其估计值, 其计算公式为:

$$PV_+ = \frac{a}{a+c} \quad (13-8)$$

2. 阴性预测值 (negative predictive value) 阴性预测值表示待评价的诊断方法诊断为阴性的病例中实际是非患者的概率。用  $PV_-$  表示其估计值, 其计算公式为

$$PV_- = \frac{d}{b+d} \quad (13-9)$$

应该注意: 敏感度、特异度、误诊率和漏诊率等指标, 只有将待评价的诊断试验方法与金标准比较时才有意义。另外, 在临床实践中,  $PV_+$  和  $PV_-$  的大小不仅与灵敏度和特异度有关, 还与诊断的目标人群中患有某种疾病的比例大小有关。

## 三、例题和统计分析

例 13.1 评价某种生化酶诊断前列腺癌的价值。经病理诊断为前列腺癌的 118 例, 诊断为非前列腺癌的 128 例; 用某种生化酶诊断试验方法诊断为前列腺癌 92 例, 非前列腺癌 154 例, 其结果如表 13-2 所示。

表 13-2 某种生化酶诊断前列腺癌与病理诊断结果的比较

病理诊断	某种生化酶诊断		合计
	有病(阳性)	无病(阴性)	
有病(阳性)	80	38	118
无病(阴性)	12	116	128
合计	92	154	246

本例:

$$Se = \frac{80}{118} = 0.678,$$

$$Sp = \frac{116}{128} = 0.906$$

$$\hat{\alpha} = \frac{12}{128} = 0.094,$$

$$\hat{\beta} = \frac{38}{118} = 0.322$$

$$\hat{\pi} = \frac{80+116}{80+38+12+116} = 0.797,$$

$$YI = 0.678 + 0.906 - 1 = 0.584$$

$$PV_+ = \frac{80}{92} = 0.870,$$

$$PV_- = \frac{116}{154} = 0.753$$

$$OP = \frac{80 \times 116}{12 \times 38} = 20.35$$



## 第四节 诊断试验的一致性检验

### 一、一致性检验的基本概念及计算公式

诊断试验的一致性检验经常用在下列两种情况：一种是评价新的诊断试验方法与金标准的一致性；另一种是评价两种诊断试验方法对同一个样本的化验结果的一致性、两个医务工作者对同一组病人的诊断结论的一致性、同一医务工作者对同一组病人前后进行两次观察作出的诊断的一致性。

1960年Cohen等提出用Kappa值作为评价判断的一致性程度的指标。实践证明，它是一个描述判断的一致性较为理想的指标，因此在临床试验中得到广泛的使用。

#### (一) Kappa 统计量的计算公式

Kappa 统计量的计算公式为：

$$Kappa = \frac{\sum P_A - P_e}{1 - P_e} \quad (13-10)$$

式中， $P_A$  为两次观察的一致性，可用下式表示：

$$P_A = \frac{\text{实际观察一致数}}{\text{总观察人数}} = \frac{\sum A}{N} \quad (13-11)$$

$P_e$  为期望一致率，即两次检查结果由于偶然机会所造成的一致率，也称期望率，其计算公式为：

$$P_e = \left( \frac{\sum R_i C_i}{N} \right) / N \quad (13-12)$$

$R_i$ 、 $C_i$  是该格点所对应的行的总例数和列的总例数， $N$  是总例数。

由上可见，Kappa 值实际为两个差值之比，分子为观察到的一致率与期望率之差，差值越大，说明观察到的一致率远比期望率高。分母为“1-期望率”，是为了减少期望率对一致性判断的影响。

可见 Kappa 值越大，表明一致性越好，一般来说，若  $Kappa \geq 0.75$ ，说明已经取得较好的一致性；若  $Kappa < 0.40$ ，说明一致程度不够理想。

#### (二) Kappa 统计量的标准误及假设检验

Kappa 标准误的计算公式为：

$$S_k = \frac{1}{(1 - P_e) \sqrt{N}} \sqrt{P_e + P_e^2 - \frac{\sum R_i C_i (R_i + C_i)}{N^3}} \quad (13-13)$$

Kappa 值是一个样本的统计量，因此可以对其是否具有统计学意义作假设检验，其方法如下：

1.  $H_0: \kappa = 0$ ，即总体 Kappa 值等于零  
 $H_1: \kappa \neq 0$ ，即总体 Kappa 值不等于零  
 确定检验水准  $\alpha$
2. 统计量  $u$  值

$$u = \frac{Kappa}{S_k} \quad (13-14)$$

$u$  近似服从标准正态分布。如果  $u > u_{\alpha/2}$ ，可以认为两种诊断试验方法有一定的一致性。

### 二、实例计算

例 13.2 根据例 13.1 的资料，评价某种生化酶诊断前列腺癌与病理诊断的一致性





(本例属于评价新的诊断试验方法与金标准的一致性)。

### 1. 计算 Kappa 值

$$P_A = \frac{80+116}{246} = 0.7967$$

$$P_e = \frac{1}{246} \left( \frac{118 \times 92}{246} + \frac{128 \times 154}{246} \right) = \frac{44.13+80.13}{246} = 0.5051$$

$$Kappa = \frac{P_A - P_e}{1 - P_e} = \frac{0.7967 - 0.5051}{1 - 0.5051} = \frac{0.2916}{0.4949} = 0.5893$$

### 2. Kappa 值的假设检验

$H_0$ : 总体 Kappa 值等于零

$H_1$ : 总体 Kappa 值不等于零

$\alpha = 0.05$

$$S_K = \frac{1}{(1-0.5051)\sqrt{246}} \sqrt{0.5051 + (0.5051)^2 - \frac{92 \times 118(92+118) + 128 \times 154(128+154)}{(246)^3}}$$

$$= \frac{1}{7.7622} \sqrt{0.5051 + 0.2551 - 0.5265}$$

$$= 0.0623$$

$$u = \frac{Kappa}{S_K} = \frac{0.5893}{0.0623} = 9.4591$$

$u > u_{0.01/2}$ ,  $P < 0.01$ , 按照  $\alpha = 0.05$  检验水准具有统计学意义。应该注意, 虽然本例检验的  $P$  值很小, 但 Kappa 值并不很大,  $P$  值的大小在一定程度上受样本例数的影响; 在样本例数较大的情况下, 即使 Kappa 值较小如  $Kappa < 0.40$ , 仍能得到较小的  $P$  值。

## 练 习 题

### 一、单项选择题

- 实验研究随机化分组的目的是
  - 减少抽样误差
  - 减少实验例数
  - 保证客观
  - 提高检验准确度
  - 保持各组的非处理因素均衡一致
- 关于实验指标的准确度和精密度, 正确的说法是
  - 精密度较准确度更重要
  - 准确度较精密度更重要
  - 精密度主要受随机误差的影响
  - 准确度主要受随机误差的影响
  - 精密度包含准确度
- 在临床试验设计选择对照时, 最可靠的对照形式是
  - 历史对照
  - 空白对照
  - 标准对照
  - 安慰对照
  - 自身对照
- 两名医生分别阅读同一组 CT 片诊断某种疾病, Kappa 值越大说明
  - 观察个体的变异越大
  - 观察个体的变异越小
  - 观察一致性越大
  - 机遇一致性越大
  - 实际一致性越大
- 下列叙述正确的有
  - 特异度高说明测量的稳定性好



- B. 灵敏度必须大于特异度才有实际意义
- C. 增大样本含量可以同时提高灵敏度和特异度
- D. 特异度高说明假阳性率低
- E. 阳性预测值高说明患病的概率大

## 二、计算与分析

将 100 名头外伤的急诊患者作为受检人群，用患者失去知觉的时间、头部表面外伤严重程度和神经检查结果作为有无颅骨骨折的新的综合诊断标准，与所有患者的影像诊断有无颅骨骨折进行比较，结果见下表，试计算其敏感度、特异度、总的符合率、阳性预测值、阴性预测值，并对该诊断试验标准进行评价。

100 名头外伤人群的疾病状态和诊断结果

影像诊断	新的诊断方法		合计
	阳性	阴性	
有骨折	48(a)	2(b)	50
无骨折	10(c)	40(d)	50
合计	58	42	100

(马斌荣)

### Experimental Design and Evaluation of Diagnostic Testing

The evaluation of diagnostic testing mainly involves the following indexes: sensitivity, specificity,  $PV_+$  and  $PV_-$ .

The sensitivity of a test is defined as the percentage of persons with the disease of interest who have positive test results. Sensitivity is calculated as follows:

$$\text{Sensitivity} = \frac{\text{True-positives}}{\text{True-positives} + \text{False-negatives}} \times 100\% = \frac{a}{a+b} \times 100\%$$

The specificity of a test is defined as the percentage of persons without the disease of interest who have negative test results. Specificity is calculated as follows:

$$\text{Specificity} = \frac{\text{True-negatives}}{\text{True-negatives} + \text{False-positives}} \times 100\% = \frac{d}{c+d} \times 100\%$$

The  $PV_+$  is defined as the percentage of persons with positive test results who actually have the disease of interest. The  $PV_+$  therefore, allows us to estimate how likely it is that the disease of interest is present if test is positive.

The  $PV_-$  is defined as the percentage of persons with negative test results who do not have the disease of interest.

# 第十四章 SPSS for Windows 实习

## 第一节 SPSS for Windows 概述

SPSS (statistics package for social science) 是社会科学用的统计软件包。它与 SAS 一样是世界上通用的统计软件包之一。事实上,它不仅适用于社会科学,同样可应用于经济学、心理学、医学等各个领域。早期推出的是 SPSS for DOS (DOS 版本)。上世纪九十年代推出了 SPSS for Windows (Windows 版本)。这样无须编写程序,完全可以在 Windows 下通过“菜单”、“对话框”使用鼠标来操作,为读者提供极大的方便,深受用户欢迎。本书是根据 SPSS for Windows Ver 11.5 版本编写的。

### SPSS for Windows 的特点

1. SPSS for Windows 可在 Windows 98/2000/XP 环境下使用。根据作者的经验,在英文 Windows 环境下运行,故障率较低。如果需要使用汉字,则可安装“中文之星”等软件。因此,建议选用 Windows98/2000/XP+中文之星。

2. 除了数据录入及部分命令程序等少数输入工作要用键盘键入外,大多数操作可通过鼠标单击或双击来完成。

3. 本软件与其他软件有转换接口。本软件与其他软件有数据转换接口,比如由 DBASE、FOXBASE、FOXPRO 生成的 (\*.dbf) 数据文件; Microsoft Excel 生成的 (\*.xls) 数据文件; 由 SPSS for DOS 生成的 (\*.sys) 数据文件; 由 SYSTAT 生成的 (\*.sys) 数据文件; 由 LOTUS 生成的 (\*.w\*) 数据文件等都可方便地读入。而本软件生成的数据文件也可以方便地转化为相应的其他数据文件。

4. SPSS for Windows 软件包为模块结构。SPSS for Windows 软件包分为 Base、Professional、Advance、Tables、Trends、Categories 和 LISREL 等功能模块,读者可根据需要购买其中相应的功能模块。

5. SPSS for Windows 使用的硬件环境。SPSS for Windows 安装时,根据所选的模块多少及版本的不同,所占的硬盘空间也不同,一般来说,需要 16MB~24MB。除此之外,为了运行此软件,一般还应有大于 10MB 的自由空间,否则,存取速度上会明显下降,甚至会出现出错信息。

在运行该软件时,至少有 4MB 以上的内存空间。

## 第二节 SPSS for Windows 的启动和退出

### 一、SPSS for Windows 的启动

首先进入 Windows98/2000/XP,然后找到 SPSS Ver11.5 的图标,如图 14-1 所示:

双击该图标中的“SPSS 11.5 for Windows”即可启动 SPSS for Windows;一旦启动,将在屏幕上显示 SPSS for Windows 主界面,如图 14-2 所示:

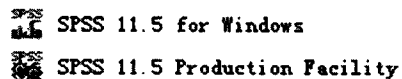


图 14-1 SPSS Ver11.5 的图标

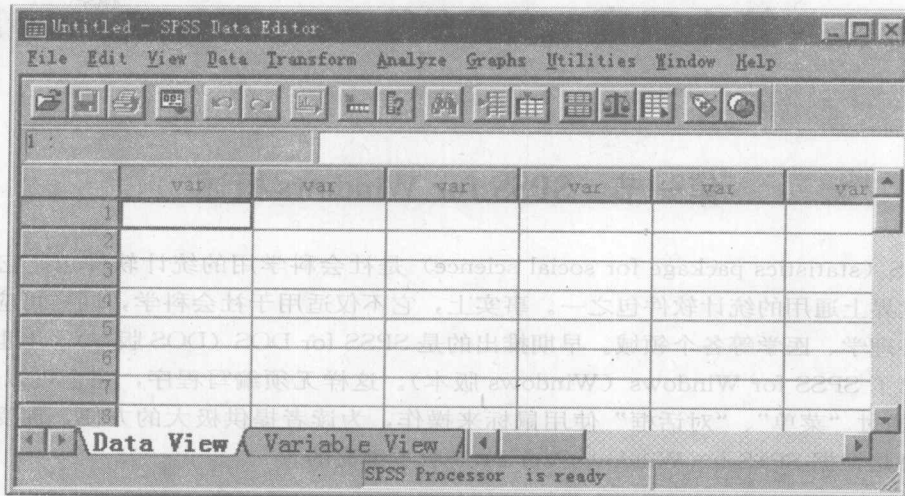


图 14-2 SPSS for Windows 主界面

SPSS for Windows 主界面上最上面一行是由 10 个菜单项目组成的主菜单。其内容为：

- (一) File                    文件操作
- (二) Edit                    文本编辑
- (三) View                    视图
- (四) Data                    数据文件的建立和编辑
- (五) Transform            数据转换
- (六) Analyze                统计分析
- (七) Graphs                统计图表的建立和编辑
- (八) Utilities                实用的一些工具
- (九) Window                窗口信息与控制
- (十) Help                    帮助

每个菜单都包括一系列功能，用鼠标单击可出现下拉式菜单，供读者进一步选择和操作。这些功能将在后面章节中陆续介绍。

## 二、SPSS for Windows 系统的退出

可按下列步骤退出 SPSS for Windows：

1. 首先要把新建的或修改过的文件的内容进行保存（存储）。具体的操作方法是：

(1) 如果是对一个新建的数据文件进行保存，单击主菜单中的“File”，在下拉菜单中，找到“Save As”并单击之，弹出如图 14-3 的对话框。

输入自选的数据文件名（其后缀应为 .sav）及相应的磁盘路径，单击“OK”键，即可存入数据文件。

(2) 如果是对一个旧的数据文件修改后进行保存，并且不再要保存原数据文件，只需在“File”的下拉菜单中找到“Save”并单击之，便可存入数据文件（注意，原数据文件将被覆盖）；如果要保留原数据文件，则选择“File”下拉菜单的“Save As”，并键入新的文件名，则修改过后数据将以新的文件名保存。

(3) 如果在运行过程中产生统计分析的结果，读者也想把它保存下来，以便今后查阅或打印，应把 Output 窗口列为当前窗口，然后使用上述同样方法进行存储。当然，其文

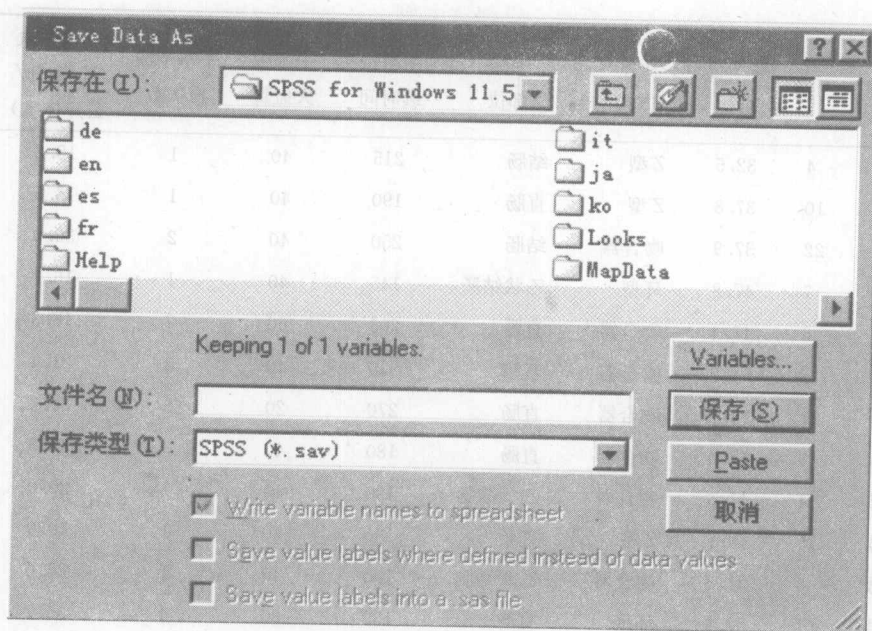


图 14-3 存储数据文件的对话框

件的后缀应为“.spo”。

(4) 如果运行中产生了统计图的话，那么把统计图列为当前窗口，再用类似方法存储。只是其后缀应为“.cht”。

2. 单击主菜单中的“File”，在下拉菜单中找到“Exit”，单击之，则可退出 SPSS for Windows。

如果读者忘了上述第一步的操作而直接操作第二步，屏幕上将出现一些对话框，询问读者是否要对修改过的文件内容进行保存，读者可以根据实际需要，进行人机对话。

### 第三节 数据文件的建立

数据文件的建立是指把科研工作过程中采集的各种信息 (Information)、数据 (Data) 以某种方式存入到计算机的磁盘中，建立可随时存取、修改、统计分析的数据文件的全过程。一般来说它包括数据编码、建立数据文件的框架结构、数据录入、选定数据文件名及保存数据等几个步骤。

**例 14.1** 北京儿童医院李龙教授进行了“围手术期输血对先天性巨结肠术后感染关系的探讨”的研究课题。共观察了 86 个病例，每个病例采集的数据为性别、年龄（月龄）、红细胞压积、手术方式、疾病部位、手术经历的时间（分钟）、手术中失血量（毫升）、手术中输血次数、手术中输血量（毫升/每公斤体重）及是否感染等。

我们取其前 20 个病例数据，其原始数据如表 14-1 所示：

表 14-1 围手术期输血对先天性巨结肠术后感染关系的探讨的数据

病例号	性别	月龄	红细胞压积	手术方式	疾病部位	手术持续时间	手术中失血量	手术中输血次数	手术输血量(毫升/公斤体重)	感染与否
1	男	11	56.4	环型	乙状结肠	200	40	1	10.0	无感染



续表

病例号	性别	月龄	红细胞压积	手术方式	疾病部位	手术持续时间	手术中失血量	手术中输血次数	手术输血量(毫升/公斤体重)	感染与否
2	男	4	32.5	Z型	结肠	215	40	1	15.2	无感染
3	男	10	37.8	Z型	直肠	190	40	1	13.5	无感染
4	男	22	37.9	吻合器	结肠	250	40	2	30.0	感染
5	女	7	47.8	环型	乙状结肠	145	40	1	16.7	无感染
6	女	6	47.4	吻合器	直肠	205	60	2	18.3	无感染
7	男	45	54.7	吻合器	直肠	210	40	3	21.4	无感染
8	男	1	98.3	吻合器	直肠	270	20	3	30.3	感染
9	男	1	47.0	吻合器	直肠	180	40	3	31.3	无感染
10	男	4	31.6	吻合器	乙状结肠	180	40	1	20.0	无感染
11	男	15	49.7	Z型	直肠	190	40	1	10.0	无感染
12	男	5	31.8	吻合器	乙状结肠	170	40	1	25.0	无感染
13	男	1	52.3	环型	直肠	135	30	4	35.7	无感染
14	男	9	46.6	Z型	直肠	245	40	1	12.5	感染
15	男	1	76.4	吻合器	乙状结肠	200	20	3	32.3	感染
16	男	144	48.1	吻合器	乙状结肠	325	40	1	7.4	感染
17	男	11	80.8	吻合器	乙状结肠	280	100	2	18.8	感染
18	男	2	56.1	吻合器	直肠	225	20	2	22.7	无感染
19	男	17	41.2	吻合器	直肠	225	40	1	16.7	感染
20	男	60	41.9	吻合器	结肠	270	40	1	10.5	无感染

## 一、数据编码

在建立数据文件之前首先要进行数据编码 (Data Code)。它包括如下三个层次:

1. 定义数据项的变量名 (Variable Name) 在例 14.1 中, 涉及病例号、性别、月龄、红细胞压积、手术方式、疾病部位、手术持续时间、手术中失血量、手术次数、手术中输血量及感染与否等 11 个数据项。我们应该对每一个数据项确立一个变量名。本例中的变量名为:

num	病例号
sex	性别
age	月龄
hct	红细胞压积
pt	手术方式
da	疾病部位
time	手术持续时间
lb	手术中失血量
bn	手术中输血次数
bc	手术中输血量
infect	感染与否



给数据项取变量名可以用汉语拼音或英文缩写。但是，为数据项取变量名时，应该遵循如下原则：

- (1) 变量名一般由不多于 8 个字符组成。
- (2) 首字符应该是英文字母，其后可为字母或数字及下划线“\_”和圆点“.”。“\_”可以放在最后，“.”不可以放在最后。
- (3) 变量名不能使用 SPSS 的保留字。SPSS 的保留字有：ALL、AND、OR、NOT、EQ、GE、GT、LE、LT、NE、TO、WITH 及一些常用的函数符号等。
- (4) 系统中不区分变量名中的大小写字母。例如 ABCD 与 abcd 被认为是同一变量。

2. 定义数据项变量的标签 (Variable Labels) 定义变量标签是对该变量名所表示的数据项内涵的进一步说明。由于变量名只由不超过 8 个字符组成，8 个字符不足以表示该变量的内涵。而变量标签不受字符位数的限制，可以用英文，在中文之星的中文平台上也可以使用中文。在统计分析过程的输出中，会在变量名对应的位置显示该变量的英文或中文的标签。这将有助于理解输出结果。

变量标签是一个可选择 (optional) 的属性，可以定义，也可以不定义。例 14.1 的变量标签可以定义为表 14-2 所示：

表 14-2 例 14.1 的变量标签

变量名	变量标签	变量名	变量标签
num	病例号	time	手术持续时间
sex	病人的性别	lb	手术中失血总量
age	病人月龄	bn	手术中输血次数
hct	红细胞压积	bc	手术中输血量
pt	手术方式	infect	感染与否
da	疾病部位		

3. 定义变量取值的变量值标签 (Value Labels) 每个变量名对应一个数据项，每个变量取不同的值 (简称变量值)，表示数据项中的不同信息。有时为了更好地理解统计分析过程中的输出结果，要进一步给变量的不同的取值赋以不同标签，以便对不同的信息作出正确解释。在输出结果的相应位置上就会出现该标签，使读者一目了然。

并不是所有变量值都要定义标签，一般来说只有离散变量才给变量值定义标签。例 14.1 的变量值标签如表 14-3 所示：

表 14-3 例 14.1 的变量值标签

变量名	变量值	变量标签	变量名	变量值	变量标签
sex	1	男	da	1	结肠
	2	女		2	乙状结肠
pt	1	环型		3	直肠
	2	Z 型	infect	0	无感染
	3	吻合器		1	感染

## 二、定义变量

定义变量包括定义变量名、变量类型、变量长度 (小数点位数)、变量标签、变量值标签、缺失值和变量显示格式 (宽度、对齐方式) 等。

定义变量的步骤如下：



### 进入数据窗口

1. 首先启动 SPSS for Windows, 进入 SPSS for Windows 的主界面。把光标移到 SPSS for Windows 的主界面下面的“Variable View”, 单击鼠标, 此时屏幕弹出含有“Variable View”的定义变量对话框, 见图 14-4:

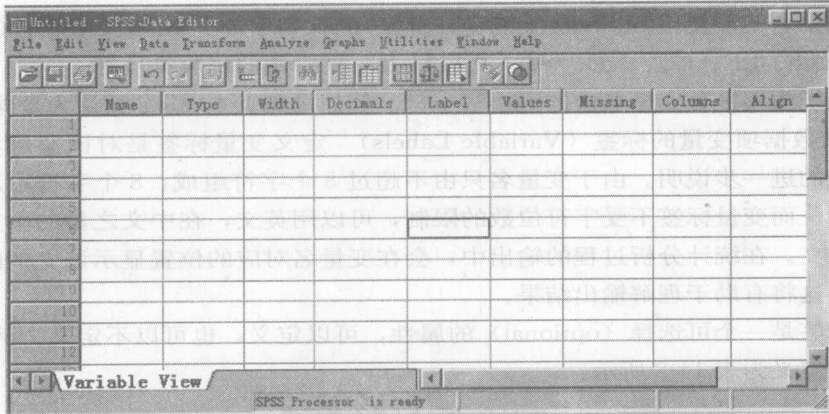


图 14-4 定义变量的对话框

此时, 屏幕出现与变量有关的字段如下:

Name 变量名

Type 变量类型

Width 变量宽度

Decimals 小数点

Label 变量标签

Values 变量值标签

Missing 缺失值

Columns 变量显示宽度

Align 变量对齐方式 (左、右或中心对齐)

2. 定义变量 以例 2.1 中的变量 num、sex、age、hct、pt、da、time、lb、bn、bc、infect 为例, 定义变量。

(1) 在图 14-4 中的 Name 下键入“num”

(2) 点击 Type, 在其下出现一个矩形框, 再点击该矩形框右侧的图标, 弹出定义变量类型的对话框, 如图 14-5 所示:

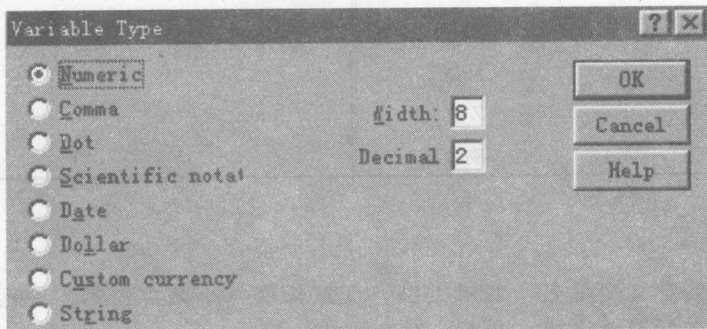


图 14-5 定义变量类型的对话框





在图 14-5 的左边出现八种可供选择的变量类型，如表 14-4 所示：

表 14-4 八种可供选择的变量类型

"Numeric"	数值型	"Date"	日期型
"Comma"	带逗点的数值型	"Dollar"	带有美元符号的数值型
"Dot"	逗号做小数的数值型	"Custom Currency"	用户自定义型
"Scientific Notation"	科学记数法	"String"	字符串型

读者可以用光标选择其中一个类型，并单击“OK”。

如果选了“Numeric”，则在图 14-4 的“Numeric”的右侧的“Width”、“Decimal”中，读者应调整相应的数值型的宽度（长度）及小数点位数。其默认值为宽度：8，小数位数：2。

应该注意，数值型的宽度应该大于该变量的数据的“整数位数+小数位数+1”。

如果选了“String”，则在图 14-4 的中部“Width”中让读者键入字符串的长度。

本例的“num”，可选变量类型为“Numeric”；其“Width”为 3，“Decimal”为 0；

(3) “label”中应录入相应的变量标签，对于“num”应录入“病人的病例号”；对于变量“num”没有变量值标签，故不用“Values”；

(4) 点击图 14-4 中的“Missing”，在其下出现一个矩形框，再点击该矩形框中右侧的图标，弹出定义变量“Missing（缺失值）”的对话框，如图 14-6 所示：

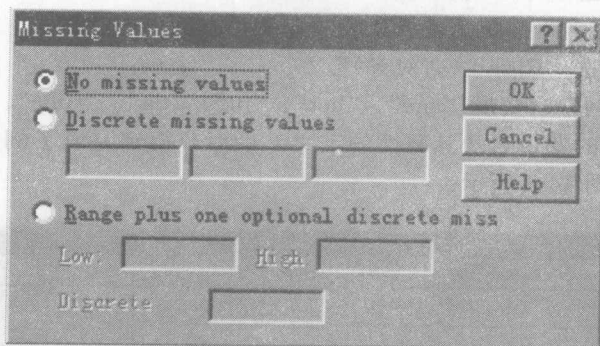


图 14-6 定义变量缺失值的对话框

在 SPSS 中缺失值有两类，即系统缺失值和用户缺失值。前者不需定义，系统自动生成；若无用户缺失值，可选图 14-6 中的“**No missing values**”；

若有用户缺失值，并且变量的观察值是离散值，可选图 14-6 中的“**Discrete missing values**”，并录入相应的缺失值；

如果变量的缺失值既有某个范围，也有离散值，可选图 14-6 中的“**Range plus one optional discrete missing value**”；

本例不需定义缺失值，故可不操作此部分内容；

(5) 点击图 14-6 的右侧“OK”表示确认，并回到图 14-4；

(6) 在图 14-4 “Columns”中，调整变量显示宽度的数值，本例可选“4”；

(7) 在图 14-4 “Align”中，可选“Right（右对齐）”；

至此，变量“num”已定义完毕。

再在图 14-4 的“name”下录入“sex”，同时可选变量类型为“Numeric”；其“Width”为 4，“Decimal”为 0；“label”中应录入相应的变量标签，对于“sex”录入“病人的性别”；对于变量“sex”应有变量值标签，故点击“Values”，在其下出现一个矩形



框，再点击该矩形框中右侧的图标，可弹出定义变量值标签“Value Labels (变量值标签)”的对话框，如图 14-7 所示：

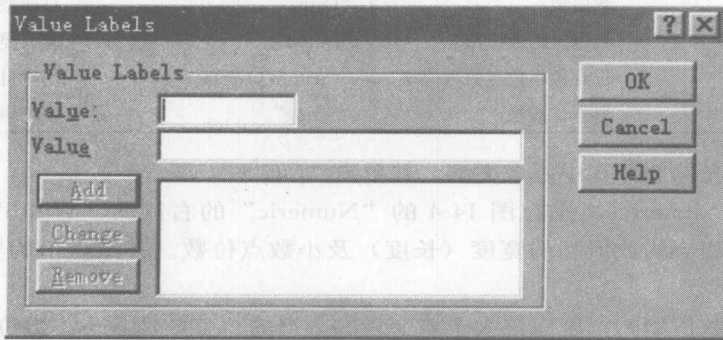


图 14-7 定义变量值标签的对话框

在图 14-7 中，让读者用键盘键入变量值标签：先键入变量值，再键入该变量值对应的标签，然后再点击“Add”；重复该过程，直到所有变量值及其标签都键完为止。本例，先键入变量值“1”，再键入该变量值的标签“男”，击“Add”按钮；键入变量值“2”，再键入该变量值的标签“女”，击“Add”按钮。

在图 14-7 的右侧击“OK”表示确认，并回到图 14-4。

在“Columns”中，调整变量显示宽度的数值，本例可选“4”；

在“Align”中可选“Right (右对齐)”。

读者应该模仿此例，对变量“age”、“hct”、“pt”、“da”、“time”、“lb”、“bn”、“bc”及“infect”作出相应的变量名及其属性的定义。最后，得到完成了例 14.1 的数据文件的变量的框架结构，如图 14-8 所示：

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
1 rum	Numeric	3	0	病人的病例号	None	None	5	Right
2 sex	Numeric	4	0	病人的性别	{1, 男}..	None	4	Right
3 age	Numeric	5	0	病人的月龄	None	None	6	Right
4 hct	Numeric	5	1	红血球压积	None	None	6	Right
5 pt	Numeric	4	0	手术方式	{1, 环状}	None	5	Right
6 da	Numeric	4	0	疾病部位	{1, 乙状结}	None	4	Right
7 time	Numeric	6	0	手术时间	None	None	6	Right
8 lb	Numeric	6	0	手术中失血量	None	None	5	Right
9 bn	Numeric	3	0	手术中输血次数	None	None	4	Right
10 bc	Numeric	6	1	手术中输血量	None	None	6	Right
11 infect	Numeric	3	0	感染与否	{0, 无感染}	None	4	Right
12								

图 14-8 例 14.1 的数据文件的变量的框架结构

### 3. 保存(存储)数据文件的变量的框架结构 具体的操作方法是：

单击主菜单中的“File”，在下拉菜单中，找到“Save As”并单击之，单击后弹出如图 14-3 的对话框。输入自选的数据文件名，本例可选“EG1401”（其后缀应为 .sav）及相应的磁盘路径后，点击“OK”键，即可存入数据文件。



### 三、数据录入

在图 14-8 的下部，点击“Data View”，可得到数据文件的二维表格，如图 14-9 所示：

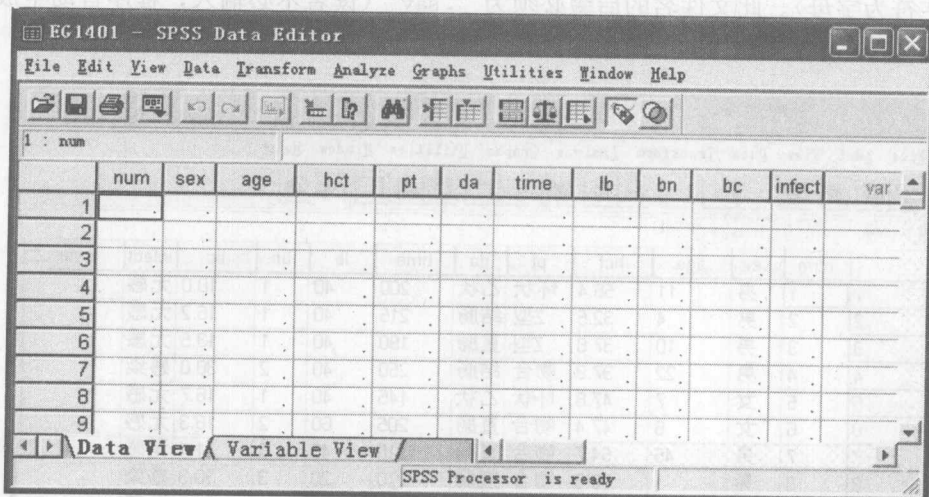


图 14-9 数据文件的二维表格

表格的顶部标有已定义的变量名“num”、“sex”、“age”、“hct”、“pt”、“da”、“time”、“lb”、“bn”、“bc”、“infect”；表格的左侧有观察量（case）的序号。一个变量名和一个观察量序号就对应了二维表格中的一个单元格，反之亦然。

#### （一）按变量输入数据

把光标移到变量名“num”列的顶部，即序号为 1 的单元格并单击之，使该单元格为当前操作的单元格，输入该变量的第一个值，回车；此时，序号为 1 的单元格接受该数值，并且当前操作单元格下移到序号为 2 的单元格，输入第二个值，回车；如此下去，一直到把该变量的数值输完为止。

再把光标移到变量名“sex”列的顶部，对应观察量序号为 1 的单元格，并单击之，使之成为当前操作单元格。键入 1，回车；1，回车；1，回车；1，回车；2，回车；2，回车；……；1，回车。把所有病人的性别全部输入完毕。

同样可对变量“hct”、“pt”、“da”、“time”、“lb”、“bn”、“bc”、“infect”的数据进行录入。

#### （二）按观察序号（病人）输入数据

按观察序号输入数据，就是把一个观察量序号（病人）的所有变量的数据都输入后，再输入下一个观察量序号（病人）的所有变量的数据……。从数据二维表格来看，它是横向输入数据。

在本例中，先把光标移第一个观察量序号（病人）的变量名为“num”的顶部，使之成为当前操作单元格，键入 1，按“Tab”键；键入 11，按“Tab”键；键入 56.4，按“Tab”键；……，键入 0，按“Tab”键。

再输入第二个观察量序号（病人）的全部数据；

如此继续，直到把最后一个观察序号（病人）的最后一个变量值输完为止。

### 四、数据文件的存储

1. 如果是要建立一个新的数据文件，则应在 SPSS for Windows 主界面中点击



“File”，并在其弹出菜单中，寻找“Save As”，并单击之，则出现保存数据文件对话框，如图 14-3 所示。

在图 14-3 的“File Name (文件名)”栏中，键入文件名（文件名应小于等于 8 位字符，首字符为字母）。此文件名的后缀必须为“.sav”（读者不必输入，程序自动生成）。

键入文件名后，按“OK”钮，即保存了例 14.1 的数据文件，如图 14-10 所示：

	num	sex	age	hct	pt	da	time	lb	bn	bc	infect	var
1	1	男	11	56.4	环状	乙状	200	40	1	10.0	无感	
2	2	男	4	32.5	Z型	结肠	215	40	1	15.2	无感	
3	3	男	10	37.8	Z型	直肠	190	40	1	13.5	无感	
4	4	男	22	37.9	吻合	结肠	250	40	2	30.0	感染	
5	5	女	7	47.8	环状	乙状	145	40	1	16.7	无感	
6	6	女	6	47.4	吻合	直肠	205	60	2	18.3	无感	
7	7	男	45	54.7	吻合	直肠	210	40	3	21.4	无感	
8	8	男	1	98.3	吻合	直肠	270	20	3	30.3	感染	
9	9	男	1	47.0	吻合	直肠	180	40	3	31.3	无感	
10	10	男	4	31.6	吻合	乙状	180	40	1	20.0	无感	
11	11	男	15	49.7	Z型	直肠	190	40	1	10.0	无感	
12	12	男	5	31.8	吻合	乙状	170	40	1	25.0	无感	
13	13	男	1	52.3	环状	直肠	135	30	4	35.7	无感	
14	14	男	9	46.6	Z型	直肠	245	40	1	12.5	感染	
15	15	男	1	76.4	吻合	乙状	200	20	3	32.3	感染	
16	16	男	144	48.1	吻合	乙状	325	40	1	7.4	感染	
17	17	男	11	80.8	吻合	乙状	280	100	2	18.8	感染	
18	18	男	2	56.1	吻合	直肠	225	20	2	22.7	无感	
19	19	男	17	41.2	吻合	直肠	225	40	1	16.7	感染	
20	20	男	60	41.9	吻合	结肠	270	40	1	10.5	无感	
21												

图 14-10 例 14.1 的数据文件

2. 如果是一个旧的数据文件，在其上修改或追加了一些数据，并且不再想保留原来的数据文件，只想保存经修改或追加后的数据文件。此时，只要在 SPSS for Windows 主界面的“File”的下拉菜单中寻找“Save”并用鼠标单击之，便可把该数据文件存入。

#### 第四节 t 检验

对于两组计量资料的均数作统计学检验可用  $t$  检验。它要求该两组资料都分别服从正态分布或近似正态分布，并且要求两组的方差具有齐同性。如果两组方差不齐，则要进行校正，常用的方法有校正界限值或者校正自由度，其本质都是校正界限值。

$t$  检验可分为配对  $t$  检验 (Paired-Samples T test) 和两组独立样本的  $t$  检验 (Independent-Samples T test)。现分别叙述于下：

##### 一、配对 $t$ 检验

例 14.2 有 12 名接种卡介苗的儿童，八周后用两批不同的结核菌素，一批是标准结





核菌素，一批是新制结核菌素，分别注射在儿童的前臂，两种结核菌素的皮肤浸润反应平均直径（mm）如表 14-5 所示，问两种结核菌素的反应性有无差别。

表 14-5 12 名儿童分别用两种结核菌素的皮肤浸润反应结果（mm）

编号	标准品	新制品	编号	标准品	新制品
1	12.0	10.0	7	10.5	8.5
2	14.5	10.0	8	7.5	6.5
3	15.5	12.5	9	9.0	5.5
4	12.0	13.0	10	15.0	8.0
5	13.0	10.0	11	13.0	6.5
6	12.0	5.5	12	10.5	9.5

### （一）建立数据文件

首先建立数据文件，此处有两个变量：

取变量名 standard 表示标准品

New 表示新制品

根据建立数据文件的方法，建立数据文件，取名为“EG1402.sav”，如图 14-11 所示：

	standard	new	var	var
1	12.0	10.0		
2	14.5	10.0		
3	15.5	12.5		
4	12.0	13.0		
5	13.0	10.0		
6	12.0	5.5		
7	10.5	8.5		
8	7.5	6.5		
9	9.0	5.5		
10	15.0	8.0		
11	13.0	6.5		
12	10.5	9.5		
13				
14				

图 14-11 例 14.2 数据文件

### （二）统计分析

#### 1. 作正态性检验

(1) 单击 SPSS for Windows 主画面的“Analyze”，展开下拉菜单；

(2) 在下拉菜单中，寻找“Nonparametric Tests”，弹出小菜单上寻找“1-Sample K-S”，单击之，弹出单样本正态性检验“One Sample Kolmogorov-Smirnov Test”对话框，如图 14-12 所示：

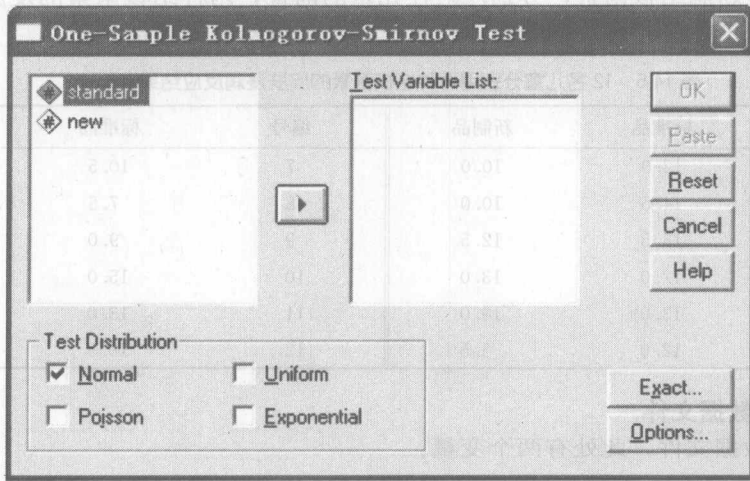


图 14-12 单样本正态性检验对话框

(3) 将图 14-12 左边矩形框的源变量“standard”、“new”调入右边的“Test Variable List”下的矩形框内；

(4) 在图 14-12 的下部的“Test Distribution”中，激活“Normal”（即选择正态分布的检验）；

(5) 单击“OK”钮，则得输出结果，如表 14-6 所示：

表 14-6 One-Sample Kolmogorov-Smirnov Test

		STANDARD	NEW
N		12	12
Normal Parameters(a, b)	Mean	12.042	8.792
	Std. Deviation	2.3975	2.5087
Most Extreme Differences	Absolute	.160	.153
	Positive	.095	.153
	Negative	-.160	-.111
Kolmogorov-Smirnov Z		.553	.529
Asymp. Sig. (2-tailed)		.919	.942

a Test distribution is Normal.

b Calculated from data.

(6) 从表 14-6 可见，经“Kolmogorov-Smirnov”正态性检验，变量“standard”、“new”的 P 值分别为 0.919 和 0.942，可认为近似正态分布；

## 2. 作配对 T 检验

(1) 单击 SPSS for Windows 主画面的“Analyze”，展开下拉菜单；

(2) 在下拉菜单中寻找“Compare Means”，弹出小菜单，在小菜单上寻找“Paired-Samples T Test”，单击之，弹出配对 t 检验“Paired-Samples T Test”对话框，如图 14-13 所示：

(3) 图 14-13 分为三部分，左上角为源变量，左下角为当前选择变量“Current Selections”，右边为配对后的变量对“Paired Variables”；

(4) 分别标记左上角源变量中的 standard 和 new，变量自动调入左下角，再点击向

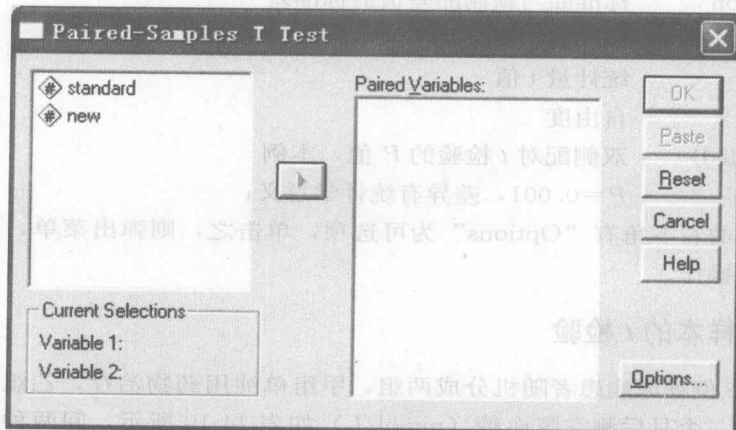


图 14-13 配对 T 检验对话框

箭头，已配对的变量调入右边的矩形框中，此时“OK”键被激活；

(5) 在图 14-13 中，点击“OK”键，得结果如表 14-7、表 14-8、表 14-9 所示：

表 14-7 Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	STANDARD	12.042	12	2.3975	.6921
	NEW	8.792	12	2.5087	.7242

表 14-7 中，Mean 均数  
 N 配对数  
 Std. Deviation 标准差  
 Std. Error of Mean 均数的标准误

表 14-8 Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	STANDARD & NEW	12	.485	.110

表 14-8 中，Correlation 表示标准品 (STANDARD) 与新制品 (NEW) 的相关系数，本例为 0.485，Sig 表示标准品 (STANDARD) 与新制品 (NEW) 的相关系数的双侧显著性检验，本例为： $P=0.110$ ，无相关关系。

表 14-9 Paired Samples Test

	Paired Differences	Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	STANDARD-NEW	3.250	2.4909	.7191	1.667	4.833	4.520	11	.001

表 14-9 中，Paired Differences 配对的差异  
 Mean 标准品与新制品差值的均数



Std. Deviation	标准品与新制品差值的标准差
Std. Error Mean	标准品与新制品差值的均数的标准误
t	统计量 $t$ 值
df	自由度
Sig. (2-tailed)	双侧配对 $t$ 检验的 $P$ 值, 本例 $P=0.001$ , 差异有统计学意义;

在图 14.13 的右下角有“Options”为可选项, 单击之, 则弹出菜单, 读者可以通过人机对话, 进行操作。

## 二、两组独立样本的 $t$ 检验

**例 14.3** 25 例糖尿病患者随机分成两组, 甲组单纯用药物治疗, 乙组采用药物治疗合并饮食疗法, 二个月后测空腹血糖 (mmol/L) 如表 14-10 所示, 问两种疗法治疗后患者血糖值是否相同?

表 14-10 25 名糖尿病患者两种疗法治疗后二个月血糖值 (mmol/L)

编号	甲组血糖值( $X_1$ )	编号	乙组血糖值( $X_2$ )
1	8.4	1	5.4
2	10.5	2	6.4
3	12.0	3	6.4
4	12.0	4	7.5
5	13.9	5	7.6
6	15.3	6	8.1
7	16.7	7	11.6
8	18.0	8	12.0
9	18.7	9	13.4
10	20.7	10	13.5
11	21.1	11	14.8
12	15.2	12	15.6
		13	18.7

### (一) 建立数据文件

两组独立样本  $t$  检验, 建立数据文件的方法与配对  $t$  检验有所不同。设立两个变量:

“Group” 其取值 1 表示甲组,

其取值 2 表示乙组,

“P” 表示血糖值。

根据本章第三节建立数据文件的方法, 建立数据文件, 取名为“EG1403.sav”, 其数据文件如图 14-14 所示:

### (二) 统计分析

#### 1. 正态性检验

(1) 点击 SPSS for Windows 主画面中的“Data”展开下拉菜单;

(2) 在下拉菜单中, 点击“Split File”得分割文件“Split File”对话框, 如图 14-15 所示:

(3) 激活“Compare groups”, 把变量“Group”调入“Groups Based on:”下, 点击“OK”, 回到 SPSS for Windows 的主画面;

(4) 单击 SPSS for Windows 主画面的“Analyze”, 展开下拉菜单;





EG1403 - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utili

1 : group 1

	group	value	var	var
1	1	8.4		
2	1	10.5		
3	1	12.0		
4	1	12.0		
5	1	13.9		
6	1	15.3		
7	1	16.7		
8	1	18.0		
9	1	18.7		
10	1	20.7		
11	1	21.1		
12	1	15.2		
13	2	5.4		
14	2	6.4		
15	2	6.4		
16	2	7.5		
17	2	7.6		

Data View Variable View

SPSS Processo

图 14-14 例 14.3 的数据文件

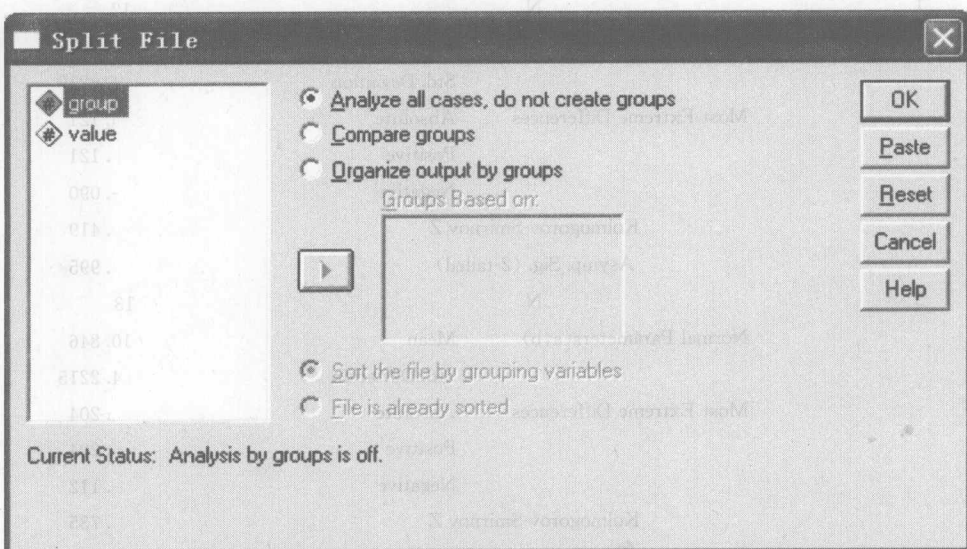


图 14-15 “Split File” 的对话框



(5) 在下拉菜单中, 寻找 “Nonparametric Tests”, 弹出小菜单上寻找 “1-Sample K-S”, 单击之, 弹出单样本正态性检验 “One Sample Kolmogorov-Smirnov Test” 对话框, 如图 14-16 所示;

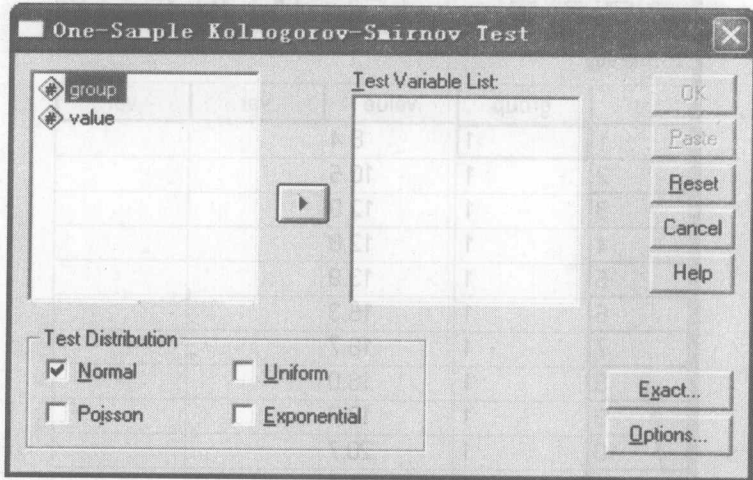


图 14-16 正态性检验对话框

(6) 将图 14-16 左边矩形框的源变量 “value” 调入右边的 “Test Variable List” 下的矩形框内;

(7) 在图 14-16 下部的 “Test Distribution” 中, 激活 “Normal” (即选择正态分布的检验);

(8) 单击 “OK” 钮, 则得输出结果, 如表 14-11 所示:

表 14-11 One-Sample Kolmogorov-Smirnov Test

GROUP		VALUE	
1	N	12	
	Normal Parameters(a, b)	Mean	15.208
		Std. Deviation	4.0216
	Most Extreme Differences	Absolute	.121
		Positive	.121
		Negative	-.090
	Kolmogorov-Smirnov Z		.419
Asymp. Sig. (2-tailed)		.995	
2	N	13	
	Normal Parameters(a, b)	Mean	10.846
		Std. Deviation	4.2215
	Most Extreme Differences	Absolute	.204
		Positive	.204
		Negative	-.112
	Kolmogorov-Smirnov Z		.735
Asymp. Sig. (2-tailed)		.652	

a Test distribution is Normal.

b Calculated from data.



从表 14-11 可见, Group 1 的正态性检验  $P = 0.995$ , Group 2 的正态性检验  $P = 0.652$ , 都可以认为近似正态分布;

(9) 再回到“Split File”对话框; 点击“Analyze all cases, do not create groups”, 点击“OK”键, 回到 SPSS for Windows 主画面。

## 2. 两组独立样本的 $t$ 检验

(1) 单击 SPSS for Windows 主画面的“Analyze”, 展开下拉菜单;

(2) 在下拉菜单中寻找“Compare Means”, 弹出小菜单, 在其上寻找“Independent-Samples T Test”, 单击之, 则弹出两独立样本的  $t$  检验“Independent-Samples T Test”对话框, 如图 14-17 所示:

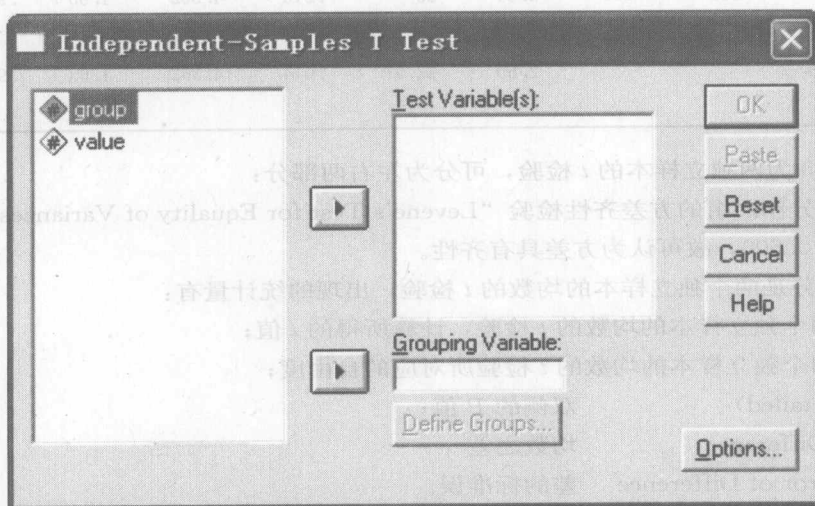


图 14-17 两独立样本的  $t$  检验的对话框

在图 14-17 的左边矩形框内出现源变量, 把 Group 调入中间下部的“Grouping Variable”下的对话框, 此时出现“Group (? ,?)”, 其下面有“Define Groups”, 单击“Define Groups”按钮, 弹出定义分组“Define Groups”的对话框, 如图 14-18 所示:

(3) 在图 14-18 中的 Group1 和 Group2 后, 分别键入“1”和“2”, 表示甲组单纯用药物治疗与乙组采用药物治疗合并饮食疗法作对比分析。击“Continue”按钮, 返回图 14-17;

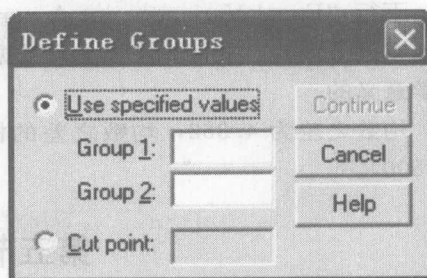


图 14-18 定义分组的对话框

(4) 把变量 value 调入“Test Variable (s):”下的对话框, 此时“OK”键被激活;

(5) 单击“OK”键, 得输出结果, 如表 14-12、表 14-13 所示:

表 14-12 Group Statistics

GROUP	N	Mean	Std. Deviation	Std. Error Mean
VALUE 1	12	15.208	4.0216	1.1609
2	13	10.846	4.2215	1.1708



表 14-12 显示两组的例数、均数、标准差和均数的标准误。

表 14-13 Independent Samples Test

	Levene's Test for Equality of Variances				t-test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
								Lower	Upper	
Equal variances assumed	.283	.600	2.64	23	.015	4.362	1.65	.9444	7.7800	
Equal variances not assumed			2.65	22.97	.014	4.362	1.65	.9511	7.7733	

表 14-13 为两独立样本的  $t$  检验，可分为左右两部分：

左面部分作两组的方差齐性检验 “Levene’s Test for Equality of Variances”，本例  $F=0.283$ ， $P=0.600$ ，故可认为方差具有齐性。

右面部分是两个独立样本的均数的  $t$  检验，出现的统计量有：

$t$  两个独立样本的均数的  $t$  检验，计算所得的  $t$  值；

$df$  两个独立样本的均数的  $t$  检验所对应的自由度；

Sig (2-tailed) 双侧的  $P$  值；

Mean Difference 均数之差

Std. Error of Difference 差的标准误

95% Confidence Interval of the Difference 差的 95% 的可信区间

表的下面有两行：

上行 “Equal Variances Assumed” 方差齐性被接受的

下行 “Equal Variances not Assumed” 方差齐性不被接受的

本例方差具有齐性，应读上行数据，即  $t=2.64$ ， $df=23$ ， $P=0.015$ ，该差异是有统计学意义的。

均数之差为 4.362，均数之差的标准误为 1.65，差的 95% 的可信区间为 (0.9444, 7.7800)。

## 第五节 方差分析

### 一、单因素方差分析

**例 14.4** 某研究者拟探讨枸杞多糖 (LBP) 对酒精性脂肪肝 (AFL) 大鼠 GSH (mg/gprot) 的影响，将周龄 7 周的 Wistar 大鼠 36 只随机分为甲、乙、丙三组，其中甲组 (正常对照组) 12 只，其余 24 只用乙醇灌胃 10 周造成大鼠慢性 AFL 模型后，再随机分为 2 组，乙组 (LBP 治疗组) 12 只，丙组 (戒酒组) 12 只，8 周后测量三组 GSH 值，结果见表 14-14，处理方式不同的大鼠的 GSH 值是否相同？

#### (一) 建立数据文件

本例共有 3 组，每组 12 只，共测量 36 个 GSH 值



表 14-14 三组大鼠 GSH 值

甲	乙	丙	甲	乙	丙
79.81	87.58	60.29	60.70	87.58	35.45
80.60	70.73	62.63	87.04	64.45	56.96
65.21	68.18	38.25	91.34	68.57	37.48
89.09	93.88	49.03	85.54	79.05	54.53
93.32	69.45	64.98	104.28	80.36	46.56
88.60	81.34	65.85	72.29	56.40	55.23

设变量“Group” 取值 1~3, 分别表示 A、B、C 组

变量“GSH” 各 GSH 测量值

建立数据文件, 取名为“EG1404.sav”, 如图 14-19 所示:

group	gsh	var	var
1	79.81		
2	80.60		
3	65.21		
4	89.09		
5	93.32		
6	88.60		
7	60.70		
8	87.04		
9	91.34		
10	85.54		
11	104.28		
12	72.29		
13	87.58		
14	70.73		
15	68.18		
16	93.88		
17	69.45		

图 14-19 例 14.4 的数据文件

## (二) 统计分析

1. 首先要对 3 组数据作正态性检验。读者可以模仿“两独立样本  $t$  检验中的正态性检验的方法”进行操作, 计算的结果是:

Group 1 的  $P=0.917$ , Group 2 的  $P=0.874$ , Group 3 的  $P=0.896$ , 都可以认为近似正态分布;

2. 单因素方差分析及方差齐性检验





(1) 单击 SPSS for Windows 主画面的“Analyze”，展开下拉菜单；在下拉菜单中寻找“Compare Means”，弹出小菜单，寻找“One-Way ANOVA”，单击之，弹出单因素方差分析“One-Way ANOVA”对话框，如图 14-20 所示：

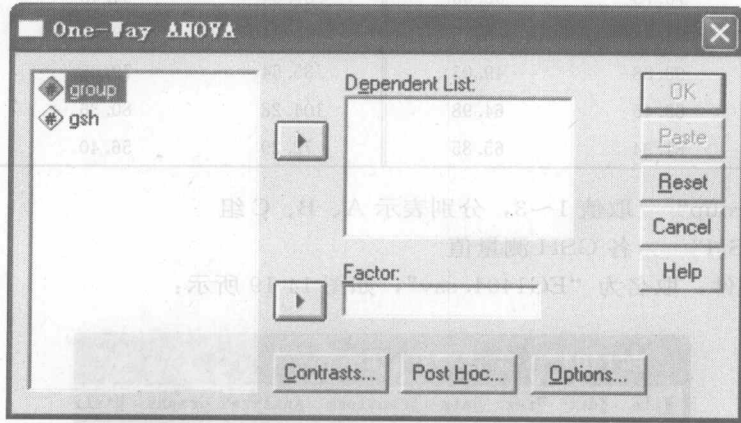


图 14-20 单因素方差分析对话框

(2) 在图 14-20 的左边出现源变量名，把“GSH”调入右边的“Dependent List”下的矩形框内；

(3) 在图 14-20 中，把变量“Group”调入右下部分的“Factor:”下的矩形框内；

(4) 在图 14-20 中，单击“Options”后，弹出单因素方差分析可选项“One-Way ANOVA: Options”对话框，如图 14-21 所示：

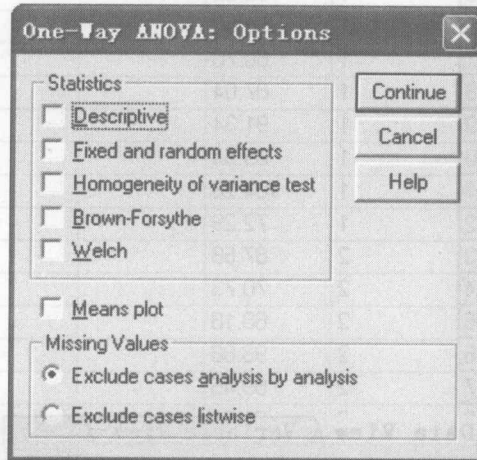


图 14-21 单因素方差分析的可选项对话框

在图 14-21 的上部是统计量：

Descriptive 是增加显示统计描述

Homogeneity of variance test 是方差齐性检验

在图 14. 21 中间有统计量：

Means plot 绘制各组的均数的连线图

在图 14. 21 的下部，是对缺失值的处理方法：

(5) 点击“Descriptive”，即要作基本统计描述，点击“Homogeneity of variance



test”，即要作方差齐性检验；

(6) 单击“OK”按钮，则得输出结果，如表 14-15、表 14-16、表 14-17 所示：

表 14-15 Descriptives

GSH

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
					1	12		
2	12	75.6308	11.06869	3.19526	68.5981	82.6635	56.40	93.88
3	12	52.2700	10.84936	3.13194	45.3766	59.1634	35.45	65.85
Total	36	70.3508	17.34573	2.89095	64.4819	76.2198	35.45	104.28

表 14-15 显示各组的基本统计描述：

表 14-16 Test of Homogeneity of Variances

GSH

Levene Statistic	df1	df2	Sig.
.015	2	33	.985

表 14-16 显示方差齐性的检验， $P=0.985$ ，可以认为方差具有齐性。

表 14-17 ANOVA

GSH

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6 223.875	2	3 111.938	23.845	.000
Within Groups	4 306.725	33	130.507		
Total	10 530.600	35			

表 14-17 显示单因素方差分析的结果：

Sum of Squares	离均差平方和
df	自由度
Mean Square	均方差
F	方差分析的统计量 $F$ 值
Sig.	方差分析的统计量 $P$ 值
Between Groups	组间变异
Within Groups	组内变异
Total	总的变异

本例  $F=23.845$ ， $P<0.001$ ，组间均数的差异有统计学意义。

A 组、B 组和 C 组，测得各个体的 GSH 不同，差异有统计学意义。但是，究竟哪些组的均数之间差异有统计学意义？应该进一步作两两比较。

## 二、方差分析中均数的两两比较

用方差分析对多组均数作显著性检验，如果差异有显著性意义，只说明总体来说各组均数之间有差异显著性，并不意味着任意两两之间的均数差异都有显著性意义。如果需要



进一步了解哪两个均数之间的差异有显著性意义，则需要进一步地作样本均数之间的两两比较，常用的统计方法是 Student Newman Keuls Test (简称为 NK 或 SNK) 和 Least Significant Difference Procedure (简称为 LSD) 法。

**例 14.5** 对于例 14.4，经单因素方差分析，得出结论总体来说各组均数之间差异有显著性意义，现在进一步作任意两组均数之间的两两比较。

其操作步骤是：

- (1) 读入数据文件“EG1404.sav”，如图 14-19 所示；
- (2) 单击 SPSS for Windows 主画面的“Analyze”，展开下拉菜单，在下拉菜单中寻找“Compare Means”，弹出小菜单，寻找“One-Way ANOVA”，单击之，弹出单因素方差分析“One-Way ANOVA”对话框，如图 14-20 所示；
- (3) 把图 14-20 的左边的源变量“GSH”调入右边的“Dependent List”下的矩形框；
- (4) 把“group”调入右边中部的“Factor:”下的矩形框；
- (5) 单击图 14-20 中的“Post Hoc...”键，弹出“One-Way ANOVA: Post Hoc Multiple Comparisons”对话框，如图 14-22 所示：

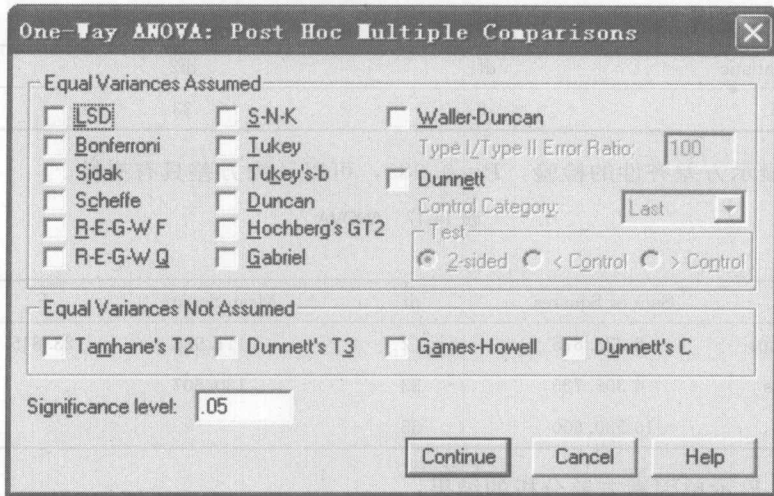


图 14-22 “One-Way ANOVA: Post Hoc Multiple Comparisons”对话框

- (6) 在图 14-21 的左下角的“Significance Level”中，读者可选“0.05”或“0.01”；
- (7) 点击图 14-22 中的“LSD”和“S-N-K”，再点击“Continue”，回到图 14-20；
- (8) 点击图 14-20 中的“OK”，除了再出现单因素方差分析结果表 14-17 外，可得两两比较的结果，如表 14-18、表 14-19 所示：

表 14-18 GSH

GROUP	N	Subset for alpha=.05	
		1	2
Student-Newman-Keuls(a)	3	12	52.2700
	2	12	75.6308
	1	12	83.1517
Sig.		1.000	.116

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size=12.000.





表 14-18 是选用“Post Hoc Tests For Group”中的“S-N-K”作两两比较的结果：

其中 Group1 与 Group2 其均数作为子集成对出现在表中，两者的  $P=0.116$ ，差异无显著性意义。

其余子集 Group1 与 Group 3、Group2 与 Group 3 的均数未作为子集成对出现在表中，表示两者之间差异有显著性意义；

表 14-19 Multiple Comparisons

Dependent Variable: GSH

(I)GROUP	(J)GROUP	Mean Difference(I-J)	Std. Error	Sig.	95% Confidence Interval		
					Lower Bound	Upper Bound	
LSD	1	2	7.5208	4.66381	.116	-1.9678	17.0094
		3	30.8817(*)	4.66381	.000	21.3931	40.3703
	2	1	-7.5208	4.66381	.116	-17.0094	1.9678
		3	23.3608(*)	4.66381	.000	13.8722	32.8494
	3	1	-30.8817(*)	4.66381	.000	-40.3703	-21.3931
		2	-23.3608(*)	4.66381	.000	-32.8494	-13.8722

\* The mean difference is significant at the .05 level.

表 14-19 是选用 LSD 法做的均数的多重两两比较的结果：

Group 1 与 Group 2 的  $P=0.116$  无显著性意义

Group 1 与 Group 3 的  $P<0.001$  有显著性意义

Group 2 与 Group 3 的  $P<0.001$  有显著性意义

表 14-19 也列出了两组均数之差、标准误、 $P$  值、95% 置信区间等统计参数。

### 三、随机区组设计的方差分析

**例 14.6** 为探讨 Rg1 对镉诱导大鼠睾丸损伤的保护作用，某研究者将同一窝别的 3 只大鼠随机地分到  $T_1$ 、 $T_2$ 、 $T_3$  三组，进行不同处理 ( $T_1$ : 对照组,  $T_2$ : 氯化镉组,  $T_3$ : Rg1+氯化镉组)，共观察了 10 个窝别大鼠的睾丸 MT 含量 ( $\mu\text{g/g}$ )，数据如表 14-20 所示，试问不同处理对大鼠 MT 含量有无影响？

表 14-20 三组大鼠 MT 含量值 ( $\mu\text{g/g}$ )

窝别	$T_1$	$T_2$	$T_3$
1	40.6	78.3	116.3
2	44.8	86.0	124.6
3	36.7	72.1	149.0
4	49.9	95.4	128.8
5	59.8	99.2	134.1
6	54.5	95.9	133.0
7	38.4	76.4	115.6
8	41.6	79.9	117.0
9	46.8	86.5	128.4
10	44.7	85.3	124.3



在例 14.6 中, 我们不仅考虑处理因素有  $T_1$ 、 $T_2$ 、 $T_3$  三个不同处理的差别, 还考虑窝别 (区组因素) 的不同。此时就构成双因素 (不同处理因素, 不同区组) 的方差分析, 也称为双因素方差分析 (Two Way ANOVA)。

### (一) 建立数据文件

取变量 group 表示  $T_1$ 、 $T_2$ 、 $T_3$  三个水平组,

block 表示 10 个不同窝别,

mt 表示 MT 含量值。

建立例 14.6 的数据文件 “EG1405.sav”, 如图 14-23 所示:

	block	group	mt	var
1	1	1	40.60	
2	2	1	44.80	
3	3	1	36.70	
4	4	1	49.90	
5	5	1	59.80	
6	6	1	54.50	
7	7	1	36.40	
8	8	1	41.60	
9	9	1	46.80	
10	10	1	44.70	
11	1	2	78.30	
12	2	2	86.00	
13	3	2	72.10	
14	4	2	95.40	
15	5	2	99.20	

图 14-23 例 14.6 的数据文件

### (二) 统计分析

(1) 单击 SPSS for Windows 主画面中的 “Analyze”, 展开下拉菜单;

(2) 用 “One Sample kolmogorov-Smirnov Test” 对  $T_1$ 、 $T_2$ 、 $T_3$  三个不同处理的数据, 作正态性检验, 得结果: Group 1 的  $P=0.972$ , Group 2 的  $P=0.955$ , Group 3 的  $P=0.983$ , 都可以认为近似正态分布;

(3) 在下拉菜单中寻找 “General Linear Model” 弹出小菜单, 单击 “Univariate...”, 则弹出方差分析 “Univariate” 对话框, 如图 14-24 所示:

(4) 把图 14-24 的左边的源变量 “mt” 调入右边的 “Dependent” 下的矩形框;

(5) 把 “group”、“block” 调入右边中部的 “Fixed Factors:” 下的矩形框;

(6) 单击 “Model” 键, 弹出 “Univariate: Model” 对话框, 如图 14-25 所示:

(7) 点击图 14-25 中的 “Custom”, 再点击中间的 “Build Terms” 下的矩形框的右侧的下箭头, 寻找 “Main Effects”, 再把 “group”、“block” 调入右边中部的 “Model” 下的矩形框, (本部分的操作是认为 group 与 block 是无交互作用的, 只考虑本因素的主效应), 再击 “Continue” 键, 回到图 14-24;

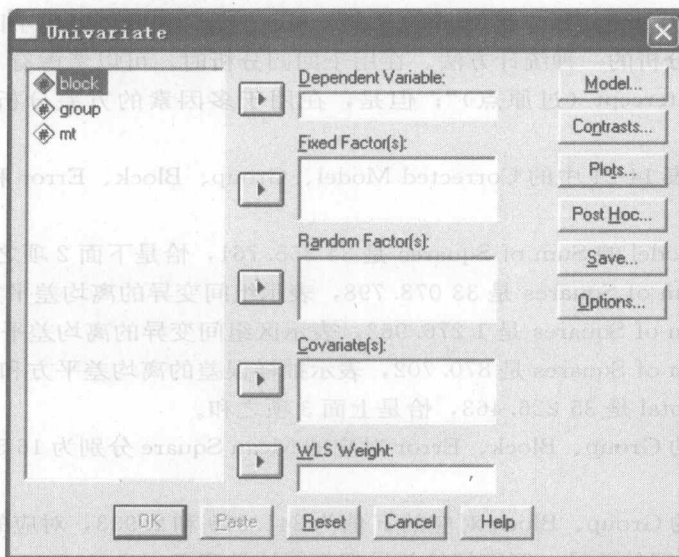


图 14-24 Univariate 的方差分析对话框

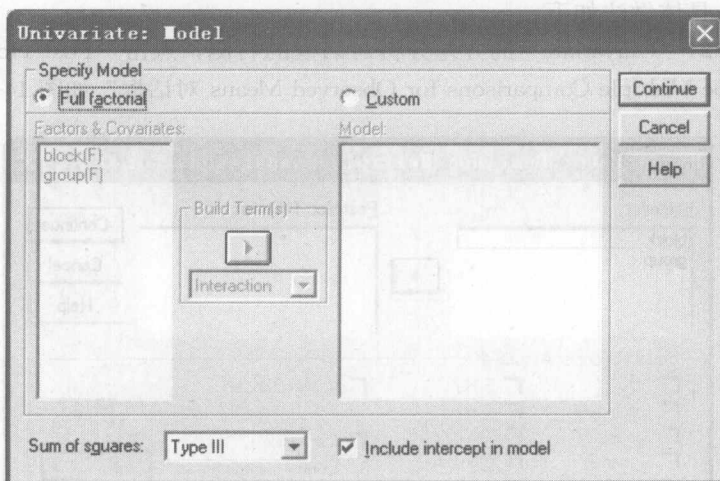


图 14-25 “Univariate: Model” 对话框

(8) 在图 14-24 中，击“OK”，可得输出结果如表 14-21 所示：

表 14-21 Tests of Between-Subjects Effects

Dependent Variable: MT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	34 355.761(a)	11	3 123.251	64.567	.000
Intercept	222 551.307	1	222 551.307	4 600.797	.000
BLOCK	1 276.963	9	141.885	2.933	.025
GROUP	33 078.798	2	16 539.399	341.919	.000
Error	870.702	18	48.372		
Total	257 777.770	30			
Corrected Total	35 226.463	29			

a R Squared=.975 (Adjusted R Squared=.960)



SPSS 中的“General Linear Model GLM Univariate”是既可以用于回归分析又可以用于多因素的方差分析的一种统计方法。在用于回归分析时，可以考虑有“Intercept（有截距）”或去掉“Intercept（过原点）”；但是，在用于多因素的方差分析时，就不要考虑“Intercept”项。

本例，只对表 14-21 中的 Corrected Model、Group、Block、Error 和 Corrected Total 进行解释：

Corrected Model 的 Sum of Squares 是 34 355.761，恰是下面 2 项之和：  
 Group 的 Sum of Squares 是 33 078.798，表示组间变异的离均差平方和；  
 Block 的 Sum of Squares 是 1 276.963，表示区组间变异的离均差平方和；  
 Error 的 Sum of Squares 是 870.702，表示抽样误差的离均差平方和；  
 Corrected Total 是 35 226.463，恰是上面 3 项之和。

表 14-21 中的 Group、Block、Error 对应的 Mean Square 分别为 16 539.399、141.885 和 48.372；

表 14-21 中的 Group、Block 对应的  $F$  值为 341.919 和 2.933，对应的  $P$  值为 0.000 和 0.025；可见，组间的差异、区组间的差异都有统计学意义。

(9) 既然组间的差异有统计学意义，我们在可以保留区组的条件下，进一步作组间的两两多重比较，具体做法如下：

1) 在图 14-24 “Univariate”的方差分析对话框的右侧，点击“Post Hoc”，弹出“Univariate: Post Hoc Multiple Comparisons for Observed Means 对话框”如图 14-26 所示：

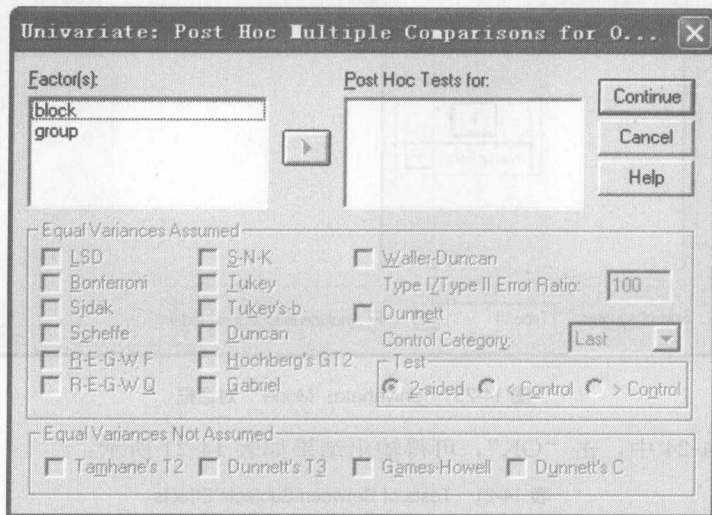


图 14-26 “Univariate: Post Hoc Multiple Comparisons...”对话框

2) 在图 14-26 中，把“factors”下的 group 调入“Post Hoc Tests for”下的矩形框中；

3) 点击图 14-26 中的“LSD”，再点击“Continue”，回到图 14-24；

4) 点击图 14-24 中的“OK”，输出的结果除了表 14-21 外，可得表 14-22：

表 14-22 是用 LSD 法作均数的多重两两比较结果：

Group 1 与 Group 2 的  $P < 0.001$  有显著性意义

Group 1 与 Group 3 的  $P < 0.001$  有显著性意义

Group 2 与 Group 3 的  $P < 0.001$  有显著性意义



表 14-22 Multiple Comparisons

Dependent Variable: MT

	(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	1	2	-39.7200(*)	3.11038	.000	-46.2547	-33.1853
		3	-81.3300(*)	3.11038	.000	-87.8647	-74.7953
	2	1	39.7200(*)	3.11038	.000	33.1853	46.2547
		3	-41.6100(*)	3.11038	.000	-48.1447	-35.0753
	3	1	81.3300(*)	3.11038	.000	74.7953	87.8647
		2	41.6100(*)	3.11038	.000	35.0753	48.1447

Based on observed means.

\* The mean difference is significant at the .05 level.

表 14-22 也列出了两组均数之差、标准误、P 值、95% 置信区间等统计参数。

#### 四、2×2 析因设计的方差分析

析因设计 (factorial design) 是将多个因素的各个水平进行排列组合, 在每一种可能的水平组合下进行试验, 以探讨各因素的效应以及各因素之间的交互效应, 而且通过比较各种组合效应, 找出最佳组合。它是一种全面的高效率的设计。本处只介绍 2 个处理因素、每个因素只有 2 个水平的析因设计方差分析, 即 2×2 析因设计方差分析。

**例 14.7** 某研究中, 40 只 5 周龄雌性 BALB/c 小鼠被随机分为 4 组, 接受不同的处理:  $A_1B_1$ 、 $A_1B_2$ 、 $A_2/B_1$ 、 $A_2B_2$ , 观测小鼠 BALF 中 IL-4 (pg/ml) 水平, 观测结果见表 14-23。

表 14-23 小鼠 BALF 中 IL-4 值 (pg/ml)

		B					
		$B_1$			$B_2$		
A	$A_1$	37.43	35.04	32.02	37.94	39.21	39.58
		33.30	36.83	36.88	39.77	33.63	32.62
		30.58	38.19	36.32	43.50	27.41	39.30
		36.60			46.43		
	$A_2$	30.57	36.27	37.76	45.11	54.15	51.61
		30.63	35.95	33.95	52.45	46.77	47.66
		29.69	36.06	32.44	54.99	50.65	42.71
		34.69			48.60		

##### (一) 建立数据文件

本例要建立三个变量:

B B=1 表示用  $B_1$  药 B=2 表示用  $B_2$  药

A A=1 表示用  $A_1$  药 A=2 表示用  $A_2$  药

IL-4 表示 IL-4 (pg/ml) 水平。

建立的数据文件模式如表 14-24 所示:



表 14-24 A、B 两药治疗血色素低的病人的血色素增加量

Case	A	B	II-4
1	1	1	37.43
2	1	1	35.04
3	1	1	32.02
4	1	1	33.30
5	1	1	36.83
6	1	1	36.88
...	...		
11	1	2	37.94
12	1	2	39.21
13	1	2	39.58
14	1	2	39.77
15	1	2	33.63
16	1	2	32.62
...	...		
21	2	1	30.57
22	2	1	36.27
23	2	1	37.76
24	2	1	30.63
25	2	1	35.95
26	2	1	33.95
...	...		
37	2	2	54.99
38	2	2	50.65
39	2	2	42.71
40	2	2	48.60

建立数据文件“EG1407.sav”，如图 14-27 所示：

## (二) 统计分析

(1) 单击 SPSS for Windows 主画面中的“Analyze”，展开下拉菜单；

(2) 在下拉菜单中寻找“General Linear Model”弹出小菜单，单击“Univariate...”，则弹出方差分析“Univariate”对话框，如图 14-28 所示：

(3) 把图 14-28 的左边的源变量“il”调入右边的“Dependent”下的矩形框；

(4) 把“a”、“b”调入右边中部的“Fixed Factors:”下的矩形框；

(5) 单击图 14-28 中的“Model”键，弹出“Univariate: Model”对话框，如图 14-29 所示：

(6) 点击图 14-29 中的“Custom”，再点击中间“Build Terms”下的矩形框的右侧的下箭头，寻找“Interaction”，把“a”和“b”作为交互项，调入右边中部的“Model”下



EC1407 - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window

34 :

	a	b	il	var	var
1	1	1	37.43		
2	1	1	35.04		
3	1	1	32.02		
4	1	1	33.30		
5	1	1	36.83		
6	1	1	36.88		
7	1	1	30.58		
8	1	1	38.19		
9	1	1	36.32		
10	1	1	36.60		
11	1	2	37.94		
12	1	2	39.21		
13	1	2	39.58		
14	1	2	39.77		
15	1	2	33.63		
16	1	2	32.62		
17	1	2	43.50		

Data View Variable View

SPSS Processor is ready

图 14-27 例 14.7 的数据文件

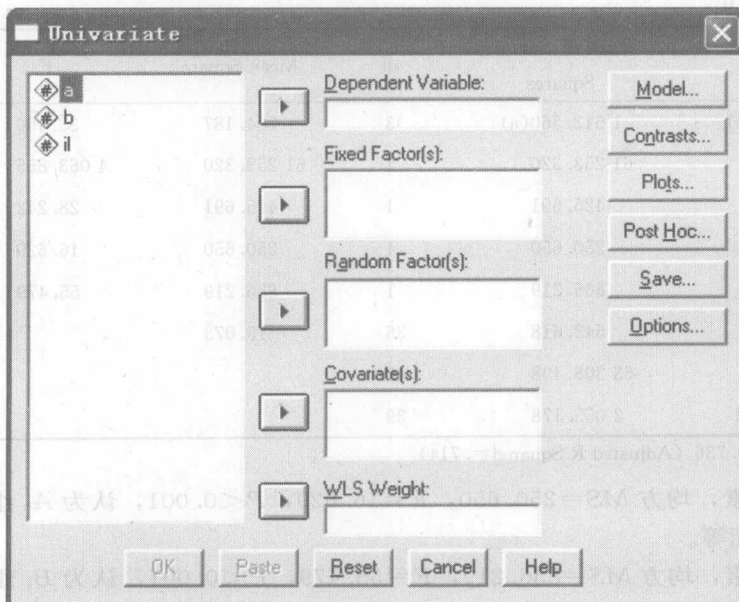


图 14-28 例 14.7 的 Univariate 的方差分析对话框

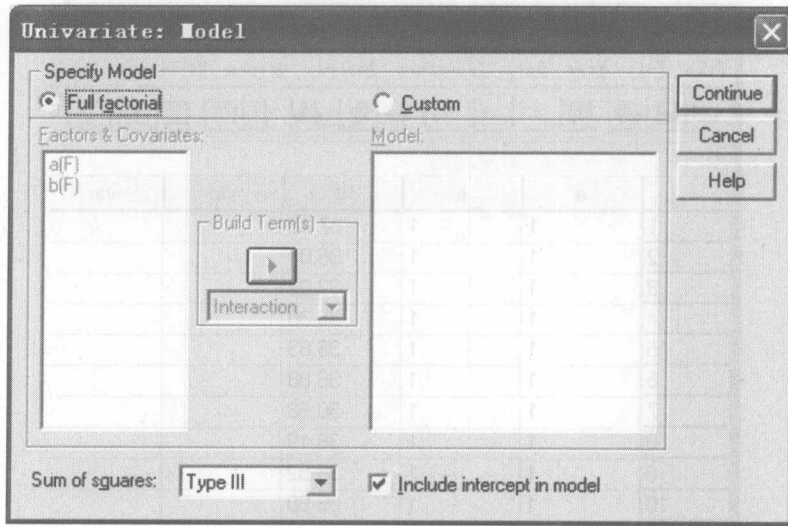


图 14-29 例 14.7 的 “Univariate: Model” 对话框

的矩形框；

(7) 再点击中间 “Build Terms” 下的矩形框的右侧的下箭头，寻找 “Main Effects”，再把 “a” 和 “b” 调入右边中部的 “Model” 下的矩形框；

((6) 与 (7) 的操作是认为 “a”、“b” 之间既有交互作用的，又考虑 “a”、“b” 因素的主效应)

(8) 再击 “Continue” 键，回到图 14-28；

(9) 在图 14-28 中，击 “OK”，可得输出结果如表 14-25 所示：

表 14-25 Tests of Between-Subjects Effects

Dependent Variable: IL

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1 512.560(a)	3	504.187	33.450	.000
Intercept	61 253.320	1	61 253.320	4 063.855	.000
A * B	425.691	1	425.691	28.242	.000
A	250.650	1	250.650	16.629	.000
B	836.219	1	836.219	55.479	.000
Error	542.618	36	15.073		
Total	63 308.498	40			
Corrected Total	2 055.178	39			

a R Squared=.736 (Adjusted R Squared=.714)

对于 A 因素，均方  $MS=250.650$ ， $F=16.629$ ， $P<0.001$ ，认为  $A_1$  组与  $A_2$  组 IL-4 值总体均数不相等。

对于 B 因素，均方  $MS=836.219$ ， $F=55.479$ ， $P<0.001$ ，认为  $B_1$  组与  $B_2$  组 IL-4 值总体均数不相等；

对于 AB 交互作用，均方  $MS=425.691$ ， $F=28.242$ ， $P<0.001$ ，认为 A 与 B 间存在交互效应。





## 第六节 $\chi^2$ 检验

### 一、四格表资料的 $\chi^2$ 检验

**例 14.8** 为了解某中药治疗原发性高血压的疗效，将 70 名高血压患者随机分为两组。试验组用该药加辅助治疗，对照组用安慰剂加辅助治疗，观察结果见表 14-26，问该药治疗原发性高血压是否有效？

表 14-26 两种疗法治疗原发性高血压的疗效

组别	有效	无效	合计
对照组	20	24	44
试验组	21	5	26
合计	41	29	70

#### (一) 建立数据文件

设三个变量：

t 表示处理方法

t=1 表示对照组

t=2 表示试验组

r 表示疗效

r=1 表示“有效”

r=2 表示“无效”

f 表示频数

建立数据文件“EG1408.sav”，如图 14-30 所示：

	t	r	f	var
1	对照组	有效	20	
2	对照组	无效	24	
3	试验组	有效	21	
4	试验组	无效	5	
5				
6				
7				
8				
9				

图 14-30 例 14.8 的数据文件

#### (二) 统计分析

(1) 在 SPSS for Windows 的主画面中，点击“Data”，展开下拉菜单，点击“Weight



cases”得“Weight cases”对话框，如图 14-31 所示：

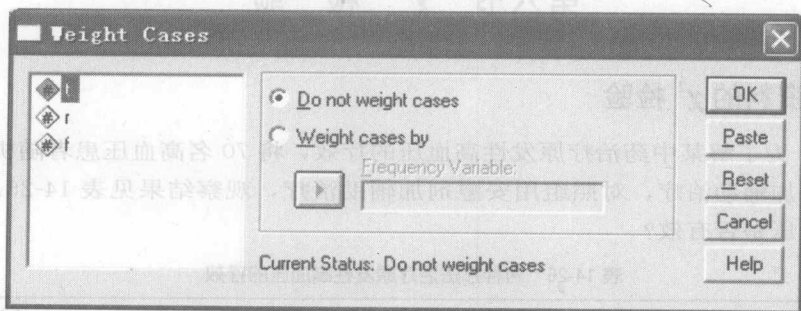


图 14-31 例 14.8 的“Weight cases”对话框

- (2) 激活“Weight cases by”，把变量“f”调入“Frequency Variable”下，点击“OK”，回到 SPSS for Windows 的主画面；
- (3) 点击 SPSS for Windows 主画面的“Analyze”，展开下拉菜单；
- (4) 在下拉菜单中，点击“Descriptive Statistics”弹出小菜单，选“Crosstabs”，点击之，则弹出“Crosstabs”对话框，如图 14-32 所示：

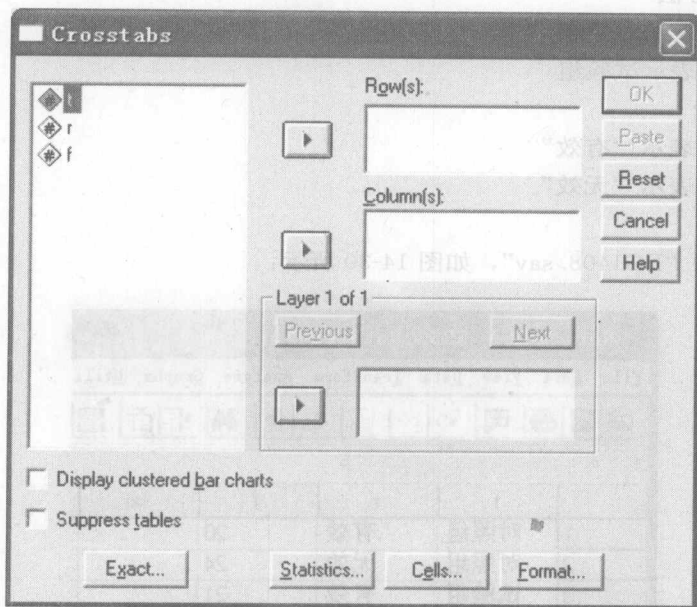


图 14-32 例 14.8 的“Crosstabs”对话框

- (5) 把图 14-32 左边源变量的“t”调入“Row(s)”下的矩形框；“r”调入“Column(s)”下的矩形框；
- (6) 点击图 14-32 下的“Statistics”钮，弹出“Crosstabs: Statistics”对话框，如图 14-33 所示；
- (7) 激活图 14-33 中的“Chi-Square”，点击“Continue”，回到图 14-32；
- (8) 点击图 14-32 下的“Cells”键，弹出“Crosstabs: Cell Display”对话框，如图 14-34 所示；
- (9) 在图 14-34 中，选“Expected”和“Row”，点击“Continue”，回到图 14-32；

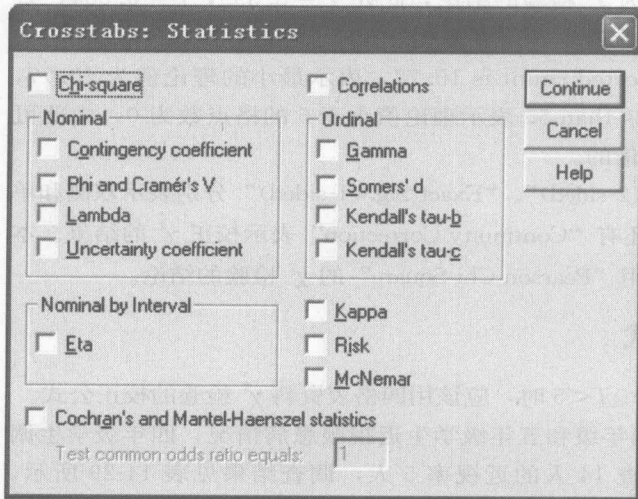


图 14-33 例 14.8 的“Crosstabs: Statistics”对话框

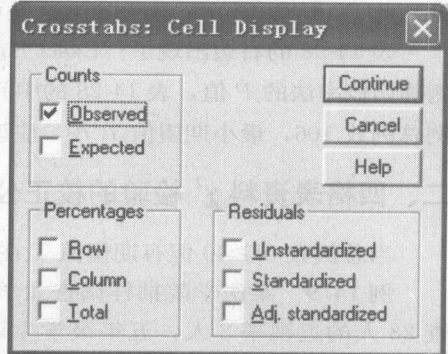


图 14-34 例 14.8 的“Crosstabs: Cell Display”对话框

(10) 点击图 14-32 中的“OK”键，得输出结果，如表 14-27、表 14-28 所示：

表 14-27 T \* R Crosstabulation

			R		Total
			有效	无效	
T	对照组	Count	20	24	44
		Expected Count	25.8	18.2	44.0
		% within T	45.5%	54.5%	100.0%
	试验组	Count	21	5	26
		Expected Count	15.2	10.8	26.0
		% within T	80.8%	19.2%	100.0%
Total	Count	41	29	70	
	Expected Count	41.0	29.0	70.0	
	% within T	58.6%	41.4%	100.0%	

表 14-27 表示对照组与试验组的疗效比较的行×列表，其中除了观察值以外，还有期望值和行百分数。

表 14-28 Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8.399(b)	1	.004		
Continuity Correction(a)	7.007	1	.008		
Likelihood Ratio	8.884	1	.003		
Fisher's Exact Test				.005	.003
Linear-by-Linear Association	8.279	1	.004		
N of Valid Cases	70				

a Computed only for a 2x2 table

b 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.77.



表 14-28 中部 Pearson Chi-Square 为  $\chi^2$  检验的结论, 此处  $\chi^2=8.399$ ,  $P=0.004$ , 差异有显著性意义, 即对照组与试验组的疗效差异有显著性。

表 14-28 的下部 The minimum expected count is 10.77, 表示最小的理论值为 10.77, 0 Cells (.0%) have expected count less than 5. 表示理论值小于 5 的格点数为 0, 由此可见, 此 Pearson Chi-Square 的结论是可信的。

表 14-28 的右边出现了 “Exact Sig. (2-sided)”、“Exact Sig. (1-sided)” 分别表示双侧和单侧确切概率法的 P 值, 表 14-28 的中间还有 “Continuity Correction” 表示校正  $\chi^2$  的结果。本例总例数 106, 最小期望值 7.70, 应该用 “Pearson Chi-Square” 的  $\chi^2$  检验的结论。

## 二、四格表资料 $\chi^2$ 检验的校正公式

当总例数  $N \geq 40$  但有期望值 T 在  $1 \leq T < 5$  时, 应该用四格表资料  $\chi^2$  检验的校正公式。

**例 14.9** 某医学院抽样调查大学四年级和五年级学生近视眼患病情况, 四年级学生调查 28 人的近视率 2 人, 五年级学生调查 14 人的近视率 5 人, 调查结果见表 14-29 所示。问该大学四年级与五年级学生的近视眼患病率是否不同?

表 14-29 四年级与五年级学生的近视眼患病率

年级	近视	非近视	合计
四年级	2	26	28
五年级	5	9	14
合计	7	35	42

### (一) 建立数据文件

设三个变量:

t 表示处理方法

t=1 表示四年级

t=2 表示五年级

r 表示疗效

r=1 表示 “近视”

r=2 表示 “非近视”

f 表示频数

建立数据文件 “EG1409.sav”, 如图 14-35 所示:

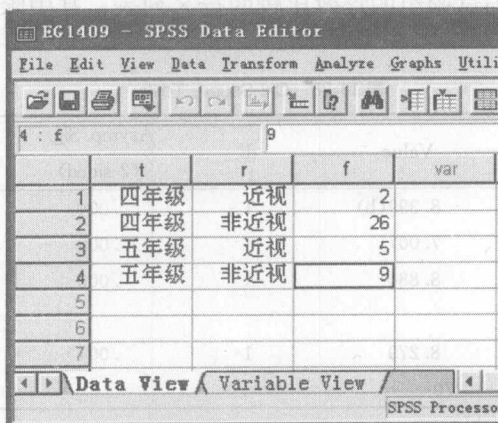


图 14-35 例 14.9 的数据文件



## (二) 统计分析

(1) 在 SPSS for Windows 的主画面中, 点击 “Data”, 展开下拉菜单, 点击 “Weight cases” 得 “Weight cases” 对话框, 如图 14-31 所示;

(2) 激活 “Weight cases by”, 把变量 “f” 调入 “Frequency Variable” 下, 点击 “OK”, 回到 SPSS for Windows 的主画面;

(3) 点击 SPSS for Windows 主画面的 “Analyze”, 展开下拉菜单;

(4) 在下拉菜单中, 点击 “Descriptive Statistics” 弹出小菜单, 选 “Crosstabs”, 点击之, 则弹出 “Crosstabs” 对话框, 如图 14-32 所示;

(5) 把图 14-32 左边源变量的 “t” 调入 “Row(s)” 下的矩形框; “r” 调入 “Column(s)” 下的矩形框;

(6) 点击图 14-32 下的 “Statistics” 钮, 弹出 “Crosstabs: Statistics” 对话框, 如图 14-33 所示;

(7) 激活图 14-33 中的 “Chi-Square”, 点击 “Continue”, 回到图 14-32;

(8) 点击图 14-32 下的 “Cells” 键, 弹出 “Crosstabs: Cell Display” 对话框, 如图 14-34 所示;

(9) 在图 14-34 中, 选 “Expected” 和 “Row”, 点击 “Continue”, 回到图 14-32;

(10) 点击图 14-32 中的 “OK” 键, 得输出结果, 如表 14-30、表 14-31 所示:

表 14-30 T \* R Crosstabulation

			R		Total
			近视	非近视	
T	四年级	Count	2	26	28
		Expected Count	4.7	23.3	28.0
		% within T	7.1%	92.9%	100.0%
	五年级	Count	5	9	14
		Expected Count	2.3	11.7	14.0
		% within T	35.7%	64.3%	100.0%
Total		Count	7	35	42
		Expected Count	7.0	35.0	42.0
		% within T	16.7%	83.3%	100.0%

表 14-30 表示四年级与五年级的近视情况的行×列表, 其中除了观察值以外, 还有期望值和行百分数。

表 14-31 Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5.486(b)	1	.019		
Continuity Correction(a)	3.621	1	.057		
Likelihood Ratio	5.188	1	.023		
Fisher's Exact Test				.031	.031



续表

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Linear-by-Linear Association	5.355	1	.021		
N of Valid Cases	42				

a. Computed only for a 2×2 table

b. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 2.33.

表 14-31 是四年级与五年级的近视情况比较的  $\chi^2$  检验的结果：

表 14-31 中部 Pearson Chi-Square 为  $\chi^2$  检验的结论，本例题总例数  $N=42$ ，表 14-31 的下部 The minimum expected count is 2.33，故应该用四格表资料  $\chi^2$  检验的校正公式。

应该读表 14-31 中的“Continuity Correction”， $\chi^2=3.621$ ， $P=0.057$ ，差异无显著性意义，即根据目前资料，尚未发现四年级与五年级的近视患病率有明显不同。

### 三、四格表的确切概率法

当四格表资料中出现  $N<40$  或  $T<1$ ，需改用四格表资料的 Fisher 确切概率（Fisher's Exact Test）法。该法是一种直接计算概率的假设检验方法。

**例 14.10** 将 17 名腰椎间盘突出症患者随机分成两组，分别用两种方法治疗，结果见表 14-32，问两种疗法的疗效是否不同？

表 14-32 17 名腰椎间盘突出症患者两种疗法的疗效表

疗法	治愈	未治愈	合计
新疗法	7	2	9
保守疗法	2	6	8
合计	9	8	17

本例  $n=17<40$ ，宜用四格表的确切概率法。

#### （一）建立数据文件：

设三个变量：

t 表示处理方法

t=1 表示“新疗法”

t=2 表示“保守疗法”

r 表示疗效

r=1 表示“治愈”

r=2 表示“未治愈”

f 表示频数

建立数据文件“EG1410.sav”，如图 14-36 所示：

#### （二）统计分析

(1) 在 SPSS for Windows 的主画面中，点击“Data”，展开下拉菜单，点击“Weight cases”得“Weight cases”对话框，如图 14-31 所示；

(2) 激活“Weight cases by”，把变量“f”调入“Frequency Variable”下，点击“OK”，回到 SPSS for Windows 的主画面；

(3) 点击 SPSS for Windows 主画面的“Analyze”，展开下拉菜单；在下拉菜单中，点击“Descriptive Statistics”弹出小菜单，选“Crosstabs”，点击之，则弹出“Crosstabs”对话

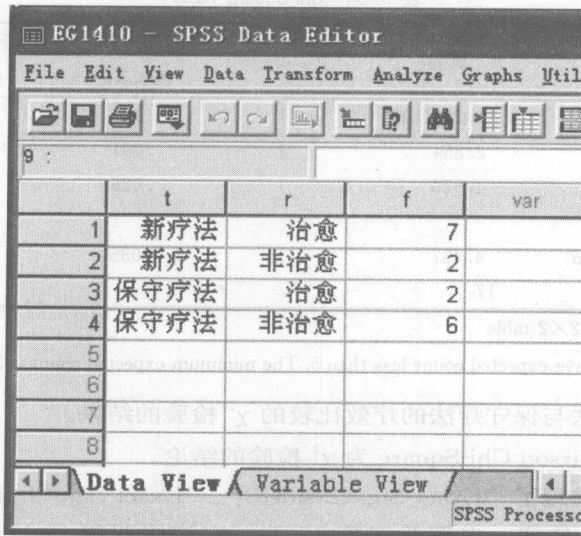


图 14-36 例 14.10 的数据文件

框，如图 14-32 所示；

(4) 把图 14-32 左边源变量的“t”调入“Row(s)”下的矩形框；“r”调入“Column(s)”下的矩形框；

(5) 点击图 14-32 下的“Statistics”钮，弹出“Crosstabs: Statistics”对话框，如图 14-32；

(6) 在图 14-33 激活“Chi-Square”，点击“Continue”，回到图 14-32；

(7) 点击图 14-32 下的“Cells”键，弹出“Crosstabs: Cell Display”对话框，如图 14-34 所示；

(8) 在图 14-34 中，选“Expected”和“Row”，点击“Continue”，回到图 14-32；

(9) 点击图 14-32 中的“OK”键，得输出结果，如表 14-33、表 14-34 所示：

表 14-33 T \* R Crosstabulation

		R		Total	
		治愈	非治愈		
T	新疗法	Count	7	2	9
		Expected Count	4.8	4.2	9.0
		% within T	77.8%	22.2%	100.0%
	保守疗法	Count	2	6	8
		Expected Count	4.2	3.8	8.0
		% within T	25.0%	75.0%	100.0%
Total		Count	9	8	17
		Expected Count	9.0	8.0	17.0
		% within T	52.9%	47.1%	100.0%

表 14-33 表示新疗法与保守疗法的疗效比较的行×列表，其中除了观察值以外，还有期望值和行百分数。





表 14-34 Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.735(b)	1	.030		
Continuity Correction(a)	2.854	1	.091		
Likelihood Ratio	4.976	1	.026		
Fisher's Exact Test				.057	.044
Linear-by-Linear Association	4.457	1	.035		
N of Valid Cases	17				

a Computed only for a 2×2 table

b 4 cells (100.0%) have expected count less than 5. The minimum expected count is 3.76.

表 14-34 是新疗法与保守疗法的疗效比较的  $\chi^2$  检验的结果。

表 14-34 中部 Pearson Chi-Square 为  $\chi^2$  检验的结论。

表 14-34 的右边出现了“Exact Sig. (2-sided)”、“Exact Sig. (1-sided)”分别表示双侧和单侧确切概率法的  $P$  值。本例总例数 17 例，应该用双侧 Fisher's Exact Test,  $P = 0.057$ ，差异无统计学意义，即新疗法与保守疗法的疗效尚未发现有显著性差异。

#### 四、配对四格表资料的 $\chi^2$ 检验

**例 14.11** 现有 198 份痰标本，每份标本分别用 A、B 两种培养基培养结核菌，结果见表 14-35。问 A、B 两种培养基的阳性培养率是否不等？

表 14-35 A、B 两种培养基的培养结果

A 培养基	B 培养基		合计
	+	-	
+	48(a)	24(b)	72
-	20(c)	106(d)	126
合计	68	130	198

##### (一) 建立数据文件：

设三个变量：

a 表示使用 A 培养基的生长情况

a=1 表示 A 培养基阳性

a=2 表示 A 培养基阴性

b 表示使用 B 培养基的生长情况

b=1 表示 B 培养基阳性

b=2 表示 B 培养基阴性

f 表示频数

建立数据文件“EG1411.sav”，如图 14-37 所示：

##### (二) 统计分析

(1) 在 SPSS for Windows 的主画面中，点击“Data”，展开下拉菜单，点击“Weight cases”得“Weight cases”对话框，如图 14-38 所示：

(2) 在图 14-38 中，激活“Weight cases by”，把变量“f”调入“Frequency Variable”下，点击“OK”，回到 SPSS for Windows 的主画面；

(3) 点击 SPSS for Windows 主画面的“Analyze”，展开下拉菜单；





	a	b	f	var
1	A阳性	B阳性	48	
2	A阳性	B阴性	24	
3	A阴性	B阳性	20	
4	A阴性	B阴性	106	
5				
6				
7				
8				

图 14-37 例 14.11 的数据文件

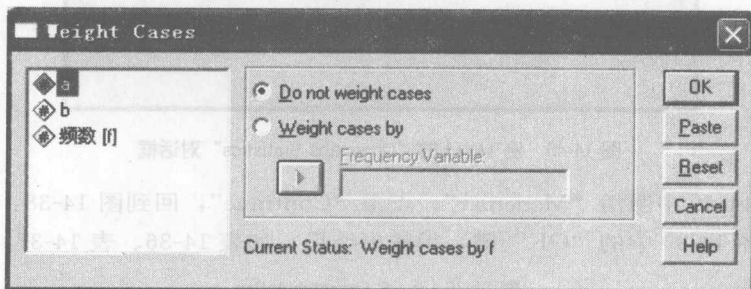


图 14-38 例 14.11 的“Weight cases”对话框

(4) 在下拉菜单中, 点击“Descriptive Statistics”弹出小菜单, 选“Crosstabs”, 点击之, 则弹出“Crosstabs”对话框, 如图 14-39 所示:

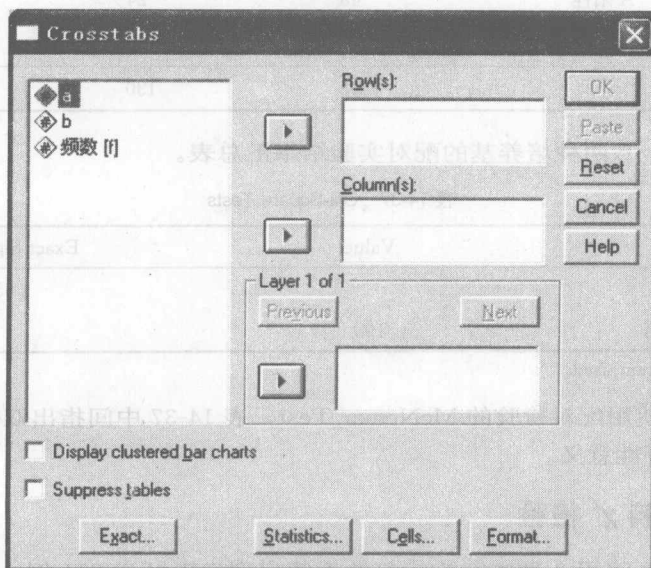


图 14-39 例 14.11 的“Crosstabs”对话框



(5) 把图 14-39 左边源变量的“a”调入“Row(s)”下的矩形框；“b”调入“Column(s)”下的矩形框；

(6) 点击图 14-39 下的“Statistics”钮，弹出“Crosstabs: Statistics”对话框，如图 14-40：

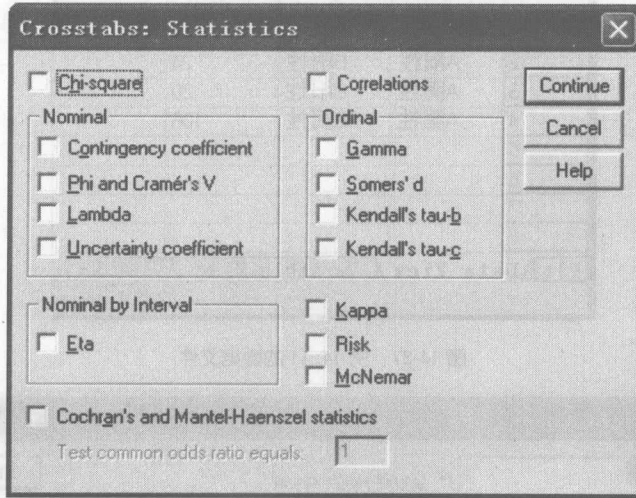


图 14-40 例 14.11 的“Crosstabs Statistics”对话框

(7) 在图 14-40 中激活“McNemar”，点击“Continue”，回到图 14-38：

(8) 点击图 14-39 中的“OK”键，得输出结果，如表 14-36、表 14-37：

表 14-36 A \* B Crosstabulation

Count	B		Total
	B 阳性	B 阴性	
A 阳性	48	24	72
A 阴性	20	106	126
Total	68	130	198

表 14-36 为 A、B 两种培养基的配对实验结果汇总表。

表 14-37 Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test.		.652(a)
N of Valid Cases	198	

a Binomial distribution used.

表 14-37 指出使用配对检验的 McNemar Test，表 14-37 中间指出双侧确切概率为  $P=0.652$ ，差异无显著性意义。

## 五、行×列表资料 $\chi^2$ 检验

例 14.12 某医院用 3 种治疗方案治疗急性无黄疸型病毒性肝炎 254 例，观察结果见表 14-38，问 3 种疗法的有效率是否不同？



表 14-38 不同方案治疗急性无黄疸型病毒性肝炎的疗效

组别	有效	无效	合计
西药组	51	49	100
中药组	35	45	80
中西药结合组	59	15	74
合计	145	109	254

### (一) 建立数据文件

取变量名及变量值标签如下：

plan	表示不同治疗方案
1	表示西药组
2	表示中药组
3	表示中西药结合组
effect	表示不同治疗效果
1	表示有效
2	表示无效
f	表示频数

建立数据文件“EG1412.sav”，如图 14-41 所示：

	plan	effect	f	var
1	西药	有效	51	
2	西药	无效	49	
3	中药	有效	35	
4	中药	无效	45	
5	中西药	有效	59	
6	中西药	无效	15	
7				
8				
9				

图 14-41 例 14.12 的数据文件

### (二) 统计分析

(1) 在 SPSS for Windows 的主画面中，点击“Data”，展开下拉菜单，点击“Weight cases”得“Weight cases”对话框，如图 14-42 所示：

(2) 击活“Weight cases by”，把变量“f”调入“Frequency Variable”下，点击“OK”，回到 SPSS for Windows 的主画面；

(3) 点击 SPSS for Windows 主画面的“Analyze”，展开下拉菜单，在下拉菜单中，点“Descriptive Statistics”弹出小菜单，选“Crosstabs”，点击之，则弹出“Crosstabs”对话框，如图 14-43 所示：

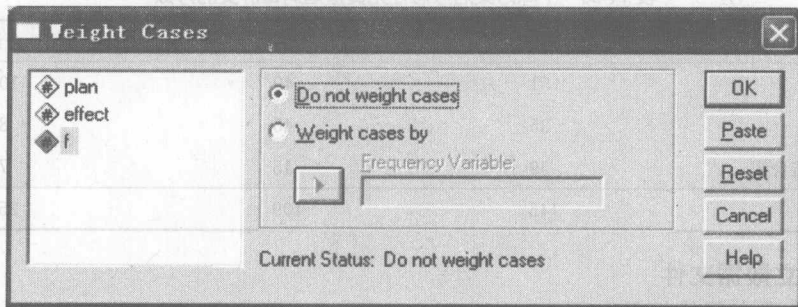


图 14-42 例 14.12 的“weight cases”对话框

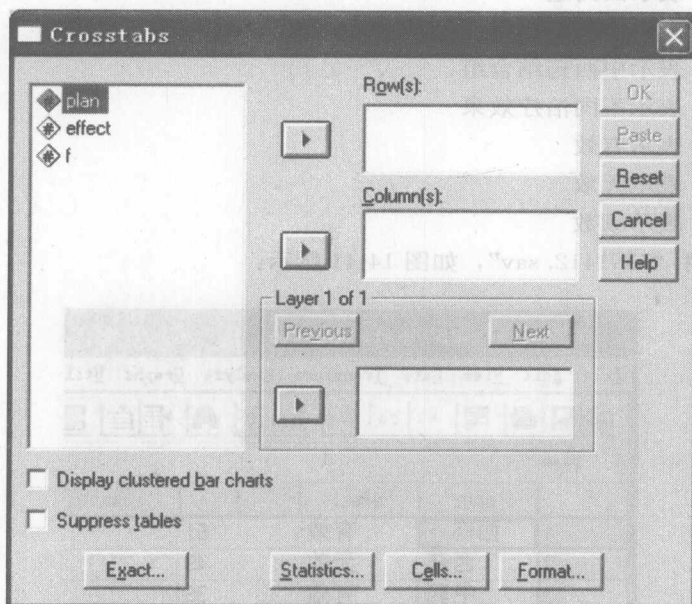


图 14-43 例 14.12 的“Crosstabs”对话框

(4) 把图 14-43 左边源变量的“plan”调入“Row(s)”下的矩形框；“effect”调入“Column(s)”下的矩形框；

(5) 点击图 14-43 下的“Statistics”钮，弹出“Crosstabs: Statistics”对话框，如图 14-44 所示；

(6) 在图 14-44 中，激活“Chi-Square”，点击“Continue”，回到图 14-43；

(7) 点击图 14-43 下的“Cells”键，弹出“Crosstabs: Cell Display”对话框，如图 14-45 所示；

(8) 在图 14-45 中，选“Row”即行百分数，回到图 14-43；

(9) 点击图 14-43 中的“OK”键，得输出结果，如表 14-39、表 14-40 所示；

表 14-39 是西药、中药和中西药疗效的行×列表，其中除了观察值以外，还有行百分数。

表 14-40 的下部 The minimum expected count is 31.76，表示最小的理论值为 31.76，0 Cells (.0%) have expected count less than 5. 表示理论值小于 5 的格点数为 0，由此可见，此 Pearson Chi-Square 的结论是可信的。



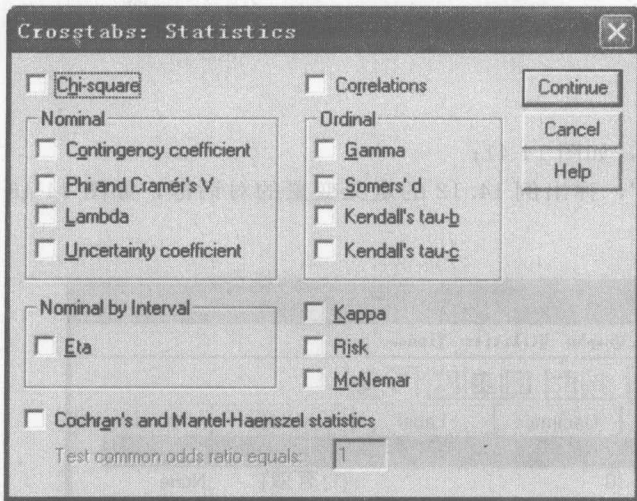


图 14-44 例 14.12 的“Crosstabs Statistics”对话框

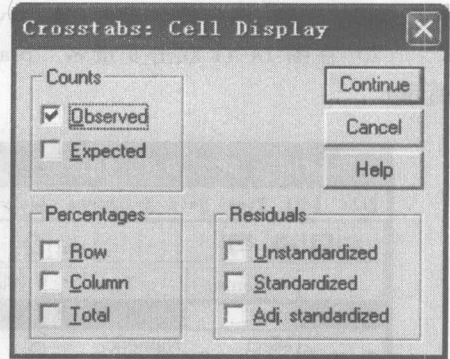


图 14-45 例 14.12 的“Crosstabs: Cell Display”对话框

表 14-39 PLAN \* EFFECT Crosstabulation

		EFFECT		Total	
		有效	无效		
PLAN	西药	Count	51	49	100
		% within PLAN	51.0%	49.0%	100.0%
	中药	Count	35	45	80
		% within PLAN	43.8%	56.3%	100.0%
	中西药	Count	59	15	74
		% within PLAN	79.7%	20.3%	100.0%
Total		Count	145	109	254
		% within PLAN	57.1%	42.9%	100.0%

表 14-40 Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	22.808(a)	2	.000
Likelihood Ratio	24.149	2	.000
Linear-by-Linear Association	12.382	1	.000
N of Valid Cases	254		

a 0 cells(.0%) have expected count less than 5.

The minimum expected count is 31.76.

表 14-40 中部 Pearson Chi-Square 为  $\chi^2$  检验的结论，此处  $P=0.000 < 0.001$ ，差异有显著性意义，即西药、中药和中西药疗效差异有显著性。

在进行  $\chi^2$  检验中，若  $P < 0.05$ ，我们拒绝无效假设  $H_0$ ，只能作出总的结论，即总的来说，西药、中药和中西药疗效差异有统计学意义。但不知道究竟哪两个组之间差异有统计学意义。因此，还要进行两两比较，即需要进行行  $\times$  列分割 (Subdividing  $R \times C$  table)，对每两组之间的差异作统计分析，然后作出结论。

重复多次的假设检验，将使第一类错误  $\alpha$  扩大，必须重新规定检验水准，详见本书第



八章，本例的检验水准  $\alpha$  应设为 0.017。

本例可以对变量“plan”设不同的“Missing value”，进行“行×列分割”计算。

1. 进行西药、中药组之间的比较

其操作步骤为：

(1) 打开数据文件“EG1412.sav”，如图 14-41；

(2) 在图 14-41 双击变量名“plan”，弹出例 14.12 的定义变量的对话框，如图 14-46 所示：

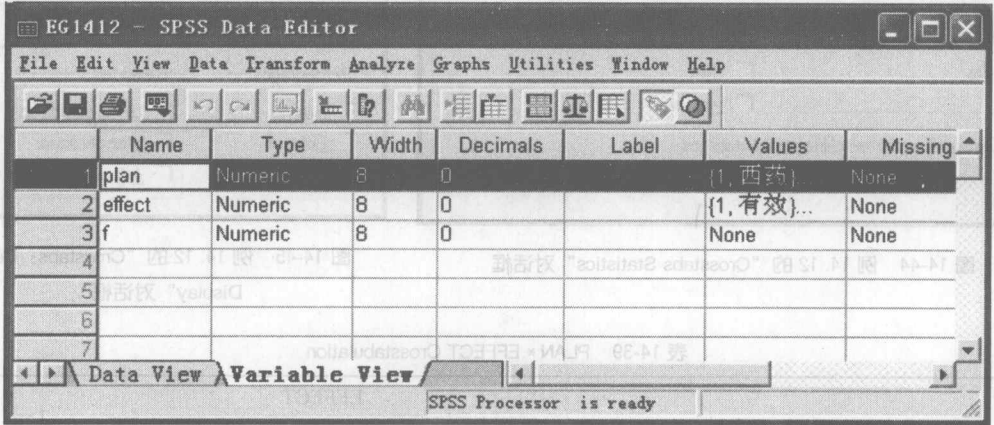


图 14-46 例 14.12 的定义变量的对话框

(3) 在图 14-46 的中，点击变量“plan”与“Missing”交叉的单元格的右侧的“...”，弹出“Missing Values”对话框，如图 14-47 所示：

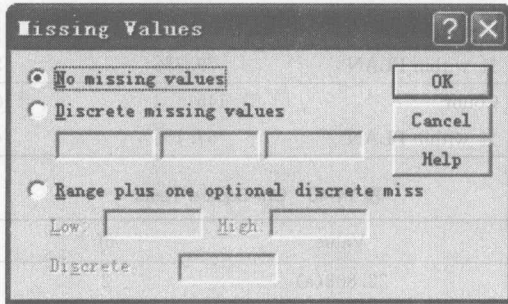


图 14-47 例 14.12 “Missing Values”对话框

(4) 在图 14-47 中，点击“Discrete Missing Values”，并在其下面的格子中输入“3”，点击“OK”键，回到图 14-46；

(5) 点击图 14-46 的左下角的“Data View”，回到图 14-41（即例 14.12 的数据文件）；

(6) 点击 SPSS for Windows 主画面的“Analyze”，展开下拉菜单；在下拉菜单中，点“Descriptive Statistics”弹出小菜单，选“Crosstabs”，点击之，则弹出“Crosstabs”对话框，如图 14-43 所示；

(7) 把图 14-43 左边源变量的“plan”调入“Row(s)”下的矩形框；“effect”调入“Column(s)”下的矩形框；

(8) 点击图 14-43 下的“Statistics”钮，弹出“Crosstabs: Statistics”对话框，如图



14-44; 激活“Chi-Square”, 点击“Continue”, 回到图 14-43;

(9) 点击图 14-43 下的“Cells”键, 弹出“Crosstabs: Cell Display”对话框, 如图 14-45 所示; 在图 14-45 中选“Row”, 点击“Continue”, 回到图 14-43;

(10) 点击图 14-43 中的“OK”键, 其结果如表 14-41、表 14-42 所示:

表 14-41 PLAN \* EFFECT Crosstabulation

			EFFECT		Total
			有效	无效	
PLAN	西药	Count	51	49	100
		% within PLAN	51.0%	49.0%	100.0%
	中药	Count	35	45	80
		% within PLAN	43.8%	56.3%	100.0%
Total	Count	86	94	180	
	% within PLAN	47.8%	52.2%	100.0%	

表 14. 41 表示西药和中药疗效的行×列表, 其中除了观察值以外, 还有行百分数。

表 14-42 Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.936(b)	1	.333		
Continuity Correction(a)	.668	1	.414		
Likelihood Ratio	.938	1	.333		
Fisher's Exact Test				.369	.207
Linear-by-Linear Association	.931	1	.335		
N of Valid Cases	180				

a Computed only for a 2×2 table

b 0 cells (.0%) have expected count less than 5.

The minimum expected count is 38.22.

表 14. 42 中部 Pearson Chi-Square 为  $\chi^2$  检验的结论, 此处  $P=0.333>0.017$ , 差异无统计学意义, 即西药组和中药组疗效无显著性差异。

2. 读者类似地在本例的图 14-46 中, 点击“Discrete Missing Values”, 并在其下面的格子中分别输入“1”或“2”, 可以得到中药组和中西药组、西药组和中西药组的两两比较的结果, 如表 14-43、表 14-44、表 14-45 和 14-46。

表 14-43 PLAN \* EFFECT Crosstabulation

			EFFECT		Total
			有效	无效	
PLAN	中药	Count	35	45	80
		% within PLAN	43.8%	56.3%	100.0%
	中西药	Count	59	15	74
		% within PLAN	79.7%	20.3%	100.0%
Total	Count	94	60	154	
	% within PLAN	61.0%	39.0%	100.0%	



表 14. 43 表示中药和中西药疗效的行×列表，其中除了观察值以外，还有行百分数。

表 14. 44 Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	20. 926(b)	1	. 000		
Continuity Correction(a)	19. 440	1	. 000		
Likelihood Ratio	21. 660	1	. 000		
Fisher's Exact Test				. 000	. 000
Linear-by-Linear Association	20. 790	1	. 000		
N of Valid Cases	154				

a Computed only for a 2×2 table

b 0 cells (. 0%) have expected count less than 5.

The minimum expected count is 28. 83.

此处 0 cells (. 0%) have expected count less than 5. The minimum expected count is 28. 83. 故可用 Pearson Chi-Square 的结果。

表 14. 44 中部 Pearson Chi-Square 为  $\chi^2$  检验的结论， $P < 0. 001 < 0. 017$ ，差异有统计学意义，即中药组和中西药组疗效差异有显著性。

表 14-45 PLAN \* EFFECT Crosstabulation

		EFFECT		Total	
		有效	无效		
PLAN	西药	Count	51	49	100
		% within PLAN	51. 0%	49. 0%	100. 0%
	中西药	Count	59	15	74
		% within PLAN	79. 7%	20. 3%	100. 0%
Total		Count	110	64	174
		% within PLAN	63. 2%	36. 8%	100. 0%

表 14. 45 表示西药和中西药疗效的行×列表，其中除了观察值以外，还有行百分数。

表 14-46 Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	15. 096(b)	1	. 000		
Continuity Correction(a)	13. 886	1	. 000		
Likelihood Ratio	15. 708	1	. 000		
Fisher's Exact Test				. 000	. 000
Linear-by-Linear Association	15. 010	1	. 000		
N of Valid Cases	174				

a Computed only for a 2×2 table

b 0 cells (. 0%) have expected count less than 5.

The minimum expected count is 27. 22.

此处 0 cells (. 0%) have expected count less than 5. The minimum expected count is 28. 83. 故可用 Pearson Chi-Square 的结果。

表 14. 46 中部 Pearson Chi-Square 为  $\chi^2$  检验的结论， $P < 0. 001 < 0. 017$ ，差异有统计





学意义，即西药组和中西药组疗效差异有显著性。

## 第七节 非参数检验

### 一、配对资料的符号秩和检验

配对资料的符号秩和检验 (Wilcoxon signed rank test) 是配对设计的非参数检验。

**例 14.13** 临床某医生研究白癜风病人的白介素 IL-6 指标在白斑部位与正常部位有无差异，调查的资料如表 14-47 所示：

表 14-47 白癜风病人的不同部位白介素 IL-6 指标

病人号	白斑部位	正常部位
1	40.03	88.57
2	97.13	80.00
3	80.32	123.72
4	25.32	39.03
5	19.61	24.37
6	14.50	92.75
7	49.63	121.57
8	44.56	89.76

#### (一) 建立数据文件

分别定义二个变量

Baiban 表示白斑部位

Zhengch 表示正常部位

录入数据，得数据文件“EG1413.sav”，如图 14-48 所示：

	baiban	zhengch	var
1	40.03	88.57	
2	97.13	80.00	
3	80.32	123.72	
4	25.32	39.03	
5	19.61	24.37	
6	14.50	92.75	
7	49.63	121.57	
8	44.56	89.76	
9			
10			

图 14-48 例 14.13 数据文件



## (二) 统计分析

(1) 单击 SPSS for Windows 主画面中“Analyze”，展开下拉菜单；在下拉菜单中，寻找“Nonparametric Tests”，自动弹出小菜单；

(2) 在小菜单中寻找“2 Related Samples...”，点击之，得两配对样本的秩和检验“Two-Related-Samples Tests”对话框，如图 14-49 所示：

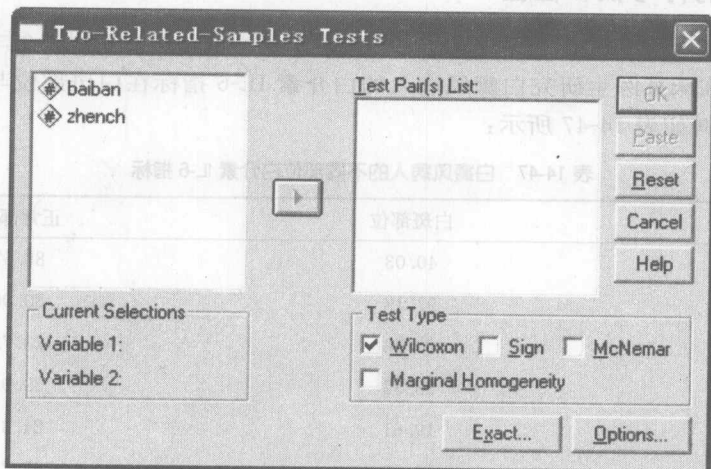


图 14-49 两配对样本的秩和检验

(3) 将图 14-49 左边的源变量“baiban”及“zhengch”按配对方式，调入右边“Test Pair (s) List”下的矩形框；

(4) 点击“OK”钮，得输出结果，如表 14-48、表 14-49 所示：

表 14-48 Ranks

		N	Mean Rank	Sum of Ranks
ZHENCH-BAIBAN	Negative Ranks	1(a)	3.00	3.00
	Positive Ranks	7(b)	4.71	33.00
	Ties	0(c)		
	Total	8		

a ZHENCH<BAIBAN

b ZHENCH>BAIBAN

c ZHENCH=BAIBAN

表 14-48 是两配对样本的秩和检验的统计描述，ZHENCH<BAIBAN 有 1 对；ZHENCH>BAIBAN 有 7 对，平均秩次分别为 3 和 4.71。

表 14-49 Test Statistics (b)

	ZHENCH-BAIBAN
Z	-2.100(a)
Asymp. Sig. (2-tailed)	.036

a Based on negative ranks.

b Wilcoxon Signed Ranks Test

表 14-49 是两配对样本的秩和检验的统计分析的结果，使用的方法是“Wilcoxon Signed Ranks Test”，z 为检验统计量，双侧检验的概率为  $P=0.036<0.05$ ，即差异有显著性意义。



## 二、两独立样本比较的秩和检验

**例 14.14** 对无淋巴细胞转移与有淋巴细胞转移的胃癌患者,观察其生存时间如表 14-50 所示,问两组患者的生存时间是否不同?

表 14-50 两组患者生存时间(月)

无淋巴细胞转移	有淋巴细胞转移	无淋巴细胞转移	有淋巴细胞转移
12	5	46	24
25	8	56	29
27	12	60	30
29	12		34
38	12		36
42	17		40
46	21		48

### (一) 建立数据文件

取二个变量:

Group 表示分组

1 表示“无淋巴细胞转移”

2 表示“有淋巴细胞转移”

time 表示患者的生存时间

建立数据文件“EG1414.sav”,如图 14-50 所示:

group	time	var	var
1	12		
2	25		
3	27		
4	29		
5	38		
6	42		
7	46		
8	46		
9	56		
10	60		
11	5		
12	8		
13	12		
14	12		
15	12		

图 14-50 例 14.14 的数据文件



## (二) 统计分析

(1) 单击 SPSS for Windows 主画面中“Analyze”，展开下拉菜单；在下拉菜单中，寻找“Nonparametric Tests”，自动弹出小菜单；

(2) 在小菜单中寻找“2 Independent Samples...”，点击之，弹出两独立样本的秩和检验“Two-Independent-Samples Tests”对话框，如图 14-51 所示：

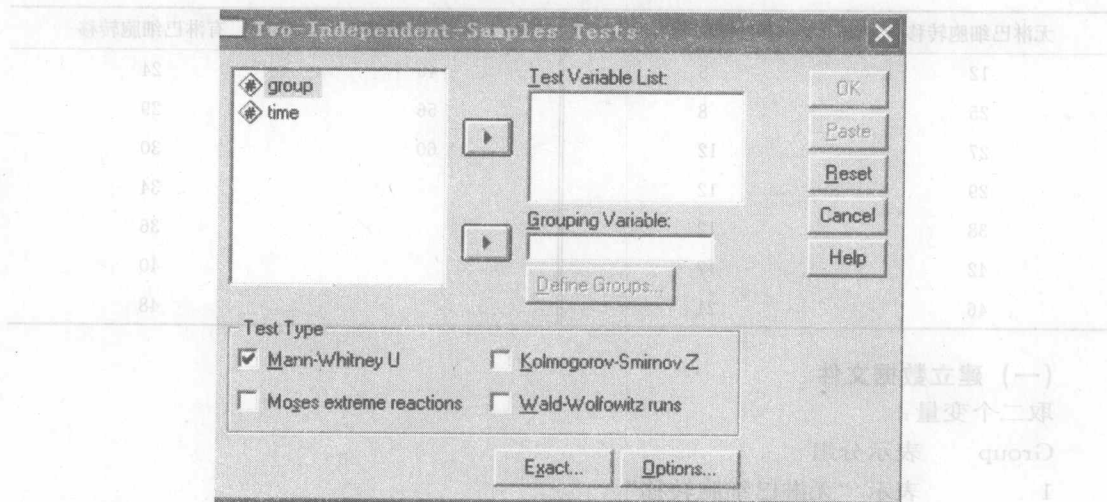


图 14-51 两独立样本的秩和检验对话框

(3) 将图 14-51 的左边的源变量中的“time”调入“Test Variable List”下的矩形框；

(4) 将图 14-51 的左边的源变量中的“Group”

调入“Grouping Variable:”下的矩形框；

(5) 并激活“Define Groups”，点击“Define Groups”键，得到“Two-Independent-Samples: Define Groups”对话框；如图 14-52 所示：

(6) 在图 14-52 中，给出组的范围“1”、“2”，点击“Continue”，回到图 14-51；

(7) 在图 14-51 中单击“OK”钮，得输出结果，如表 14-51、表 14-52 所示：

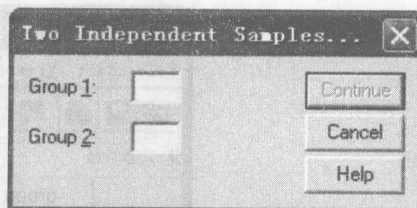


图 14-52 两独立样本（定义分组）的秩和检验对话框

表 14-51 Ranks

GROUP		N	Mean Rank	Sum of Ranks
TIME	1	10	16.20	162.00
	2	14	9.86	138.00
Total		24		

表 14-51 表示，无淋巴细胞转移组共 10 例，其平均秩次为 16.20，秩次之和为 162；有淋巴细胞转移组共 14 例，平均秩次为 9.86，秩次之和为 138。

表 14-52 的前三行指出，本检验使用的三个统计量“Mann-Whitney U”、“Wilcoxon Rank Sum W”和“Z”，他们的值分别为 33.000、138.000、-2.172；

表 14.52 的后二行双侧检验  $P=0.030$ ，双侧确切概率  $P=0.031$ ，两组的患者的生存时间差异有显著性，由于无淋巴细胞转移组的平均秩次为 16.20 大于有淋巴细胞转移组平均秩次为 9.86，故无淋巴细胞转移组的患者的生存时间明显大于有淋巴细胞转移组。



表 14-52 Test Statistics (b)

	TIME
Mann-Whitney U	33.000
Wilcoxon W	138.000
Z	-2.172
Asymp. Sig. (2-tailed)	.030
Exact Sig. [2 * (1-tailed Sig.)]	.031(a)

a Not corrected for ties.

b Grouping Variable: GROUP

### 三、多个样本比较的秩和检验

**例 14.15** 研究白血病时，测定四组鼠脾 DNA 的含量，结果如表 14-53 所示：试分析各组 DNA 含量有无差别？

表 14-53 各组鼠脾 DNA 含量 (mg) 的秩和计算

正常脾	患自发性白血病的脾	患移植白血病的脾(甲组)	患移植白血病的脾(乙组)
12.3	10.8	9.3	9.5
13.2	11.6	10.3	10.3
13.7	12.3	11.1	10.5
15.2	12.7	11.7	10.5
15.8	13.5	11.7	10.5
16.9	13.5	12.0	10.9
17.3	14.8	12.3	11.0
17.4		12.4	11.5
		13.6	

#### (一) 建立数据文件

取变量：

Group 表示分组

- 1 表示正常脾组
- 2 表示患自发性白血病的脾组
- 3 表示患移植白血病的脾 (甲组)
- 4 表示患移植白血病的脾 (乙组)

DNA 表示各组各白鼠的体重的增加量

建立数据文件为“EG1415.sav”，如图 14-53 所示：

#### (二) 统计分析

(1) 单击 SPSS for Windows 主画面中“Analyze”，展开下拉菜单；在下拉菜单中，寻找“Nonparametric Tests”，自动弹出小菜单；在小菜单中寻找“K Independent Samples”，单击之，得多个独立样本的秩和检验“Tests for Several Independent Samples”对话框，如图 14-54 所示：



EG1415 - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utili

1 : group 1

	group	dna	var	var
1	1	12.3		
2	1	13.2		
3	1	13.7		
4	1	15.2		
5	1	15.8		
6	1	16.9		
7	1	17.3		
8	1	17.4		
9	2	10.8		
10	2	11.6		
11	2	12.3		
12	2	12.7		
13	2	13.5		
14	2	13.5		
15	2	14.8		
16	3	9.3		
17	3	10.3		

Data View Variable View SPSS Processo

图 14-53 例 14.15 数据文件

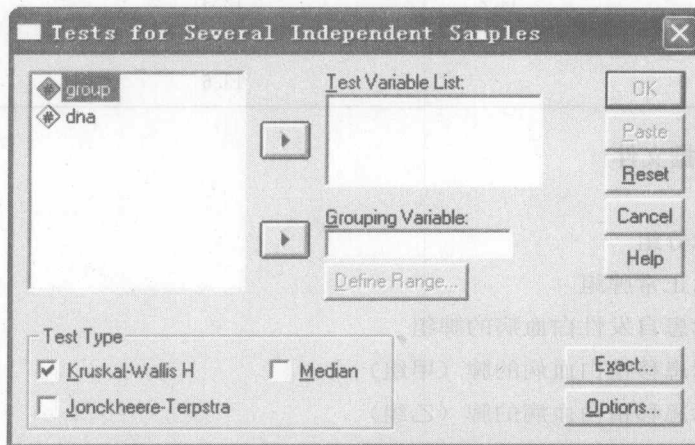


图 14-54 多个独立样本的秩和检验的对话框

(2) 将图 14-54 的左边的源变量中的“DNA”调入“Test Variable List”下的矩形框；

(3) 将图 14-54 的左边的源变量中的“Group”调入“Grouping Variable:”下的矩形框，并激活“Define Range”。点击“Define Range”键，得到“Several Independent Samples: Define Range”对话框，如图 14-55 所示：



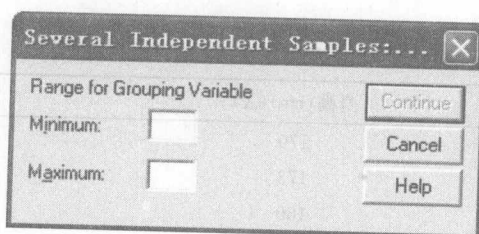


图 14-55 多个独立样本（定义分组）的秩和检验对话框

- (4) 在图 14-55 中，给出组的范围“1”和“4”，点击“Continue”，回到图 14-54；  
 (5) 在图 14-54 单击“OK”钮，得输出结果，如表 14-54、表 14-55 所示：

表 14-54 Ranks

GROUP	N	Mean Rank
DNA 1	8	27.00
2	7	19.14
3	9	13.72
4	8	6.81
Total	32	

表 14-54 表示正常脾组，共 8 例，平均秩次为 27；患自发性白血病的脾组共 7 例，平均秩次为 19.14；患移植白血病的脾（甲组）共 9 例，平均秩次为 13.72；患移植白血病的脾（乙组）共 8 例，平均秩次为 6.81。

表 14-55 Test Statistics (a, b)

	TIME
Chi-Square	19.939
df	3
Asymp. Sig.	.000

a Kruskal Wallis Test

b Grouping Variable: GROUP

表 14-55 表示用“Kruskal Wallis Test”（实质上用 chi-square 作近似计算） $P < 0.001$ ，差异有显著性意义。

经多个独立样本的秩和检验，若  $P < 0.05$ ，我们拒绝无效假设  $H_0$ ，只能作出总体上有显著性意义的总的结论，而不能对每两两之间有无显著性差异作出结论。若要进行两两比较，还需要进行分割，才能对每两两之间差异有无显著性作出结论。分割可分为多组间的两两比较和多个实验组与同一对照组比较两种。

因为重复多次的假设检验，将使第一类错误  $\alpha$  扩大，必须重新规定检验水准，不能再原来的检验水准  $\alpha = 0.05$  作为是否拒绝  $H_0$  的标准。具体校正的  $\alpha$  的算法可参照本书第八章的行  $\times$  列的  $\chi^2$  分割的算法。具体操作也类似，读者可模拟上机练习。

## 第八节 一元线性相关与回归

**例 14.16** 从男青年总体中随机抽取 11 名男青年组成样本，分别测量每个男青年的身高和前臂长，身高和前臂长均以 cm 为单位，测量结果如表 14-56 所示，试计算身高与前



臂长之间的相关系数。

表 14-56 11 名男青年身高与前臂长的测量结果

编号	身高(cm)(X)	前臂长(cm)(Y)
1	170	47
2	173	42
3	160	44
4	155	41
5	173	47
6	188	50
7	178	47
8	183	46
9	180	49
10	165	43
11	166	44

### (一) 建立数据文件

取变量

high 表示身高

length 表示前臂长

建立数据文件，取名为“EG1416.sav”，如图 14-56 所示：

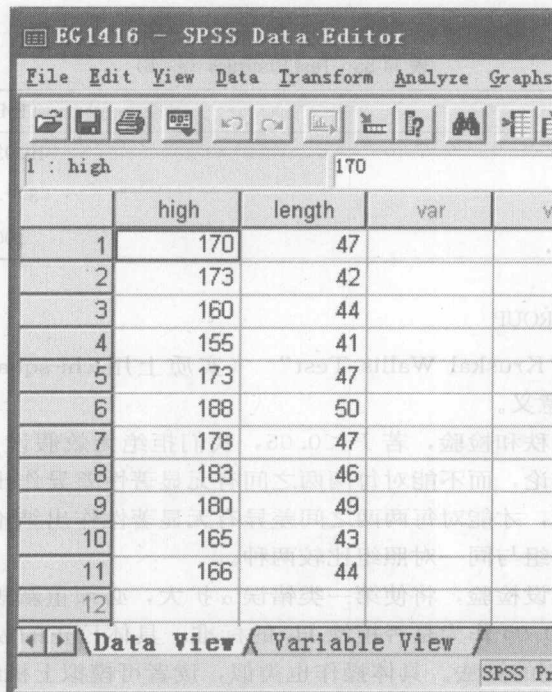


图 14-56 例 14.16 数据文件

### (二) 统计分析

(1) 首先用“One Sample Kolmogorov-Smirnov Test”方法，分别对“high”、“length”作正态性检验，其结果如表 14-57：





表 14-57 One-Sample Kolmogorov-Smirnov Test

		HIGH	LENGTH
Normal Parameters(a, b)			
	Mean	171.91	45.45
	Std. Deviation	10.005	2.876
Most Extreme Differences			
	Absolute	.093	.159
	Positive	.093	.148
	Negative	-.092	-.159
Kolmogorov-Smirnov Z		.308	.527
Asymp. Sig. (2-tailed)		1.000	.944

- a Test distribution is Normal.
- b Calculated from data.

表 14-57 中，“high”、“length”的  $P=1.000$ 、 $P=0.944$ ，都为近似正态分布；

(2) 单击 SPSS for Windows 主画面中的“Analyze”项，展开下拉菜单；

(3) 在下拉菜单中寻找“Correlate”，弹出小菜单，在其上寻找“Bivariate”，单击之，则弹出二元变量相关“Bivariate Correlations”对话框，如图 14-57 所示：

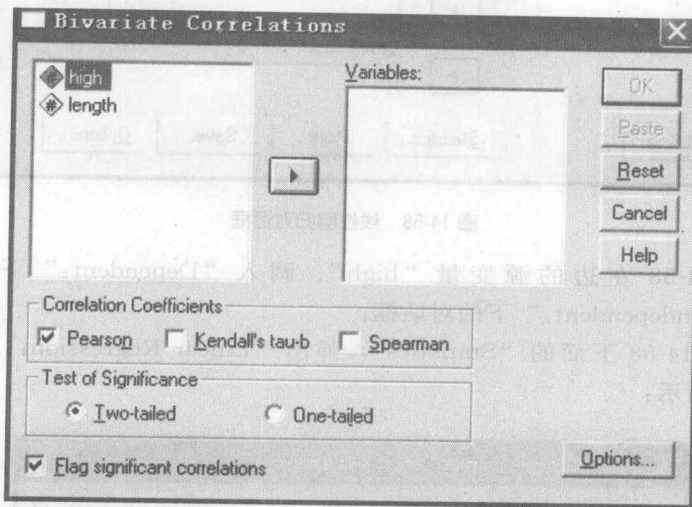


图 14-57 二元变量相关“Bivariate Correlations”对话框

(4) 将图 14-57 左边的源变量“high”、“length”调入“Variables”下的矩形框，点击“OK”，得相关分析结果如表 14-58 所示：

表 14-58 Correlations

		HIGH	LENGTH
HIGH	Pearson Correlation	1	.801(**)
	Sig. (2-tailed)	.	.003
	N	11	11
LENGTH	Pearson Correlation	.801(**)	1
	Sig. (2-tailed)	.003	.
	N	11	11

\*\* Correlation is significant at the 0.01 level (2-tailed).



表 14-58 显示相关系数  $r=0.801$ ，双侧 Pearson 检验  $P=0.003<0.01$ ，有统计学意义，可见，身高与前臂长高度相关。

(5) 单击 SPSS for Windows 主画面中的“Analyze”项，展开下拉菜单；

(6) 在下拉菜单中寻找“Regression”，弹出小菜单，在其上寻找“Linear”，单击之，则弹出线性回归“Linear Regression”对话框，如图 14-58 所示：

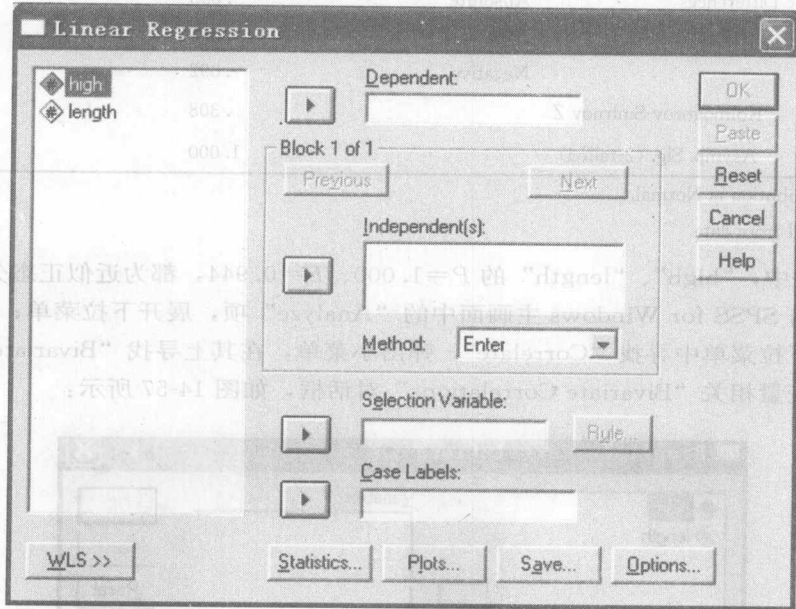


图 14-58 线性回归对话框

(7) 将图 14-58 左边的源变量“high”调入“Dependent:”下的矩形框；将“length”调入“Independent:”下的对话框；

(8) 点击图 14-58 下面的“Statistics”，弹出“Linear Regression: Statistics”对话框，如图 14-59 所示：

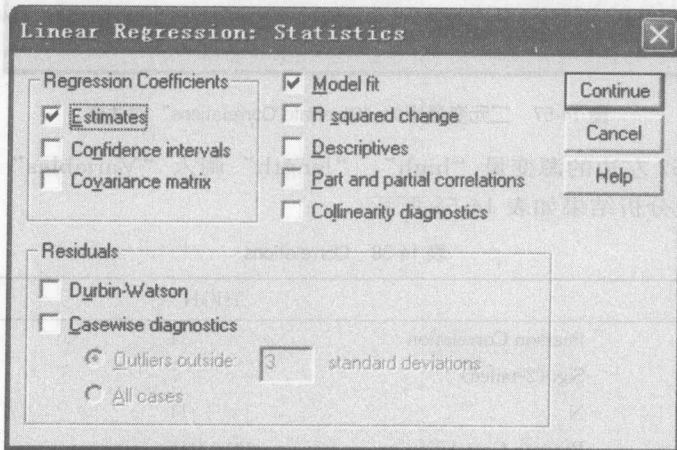


图 14-59 “Linear Regression: Statistics”对话框

(9) 在图 14-59 击活“Descriptive”，点击“Continue”键，回到图 14-58；



(10) 点击图 14-58 中的“OK”键，得输出结果，如表 14-59、表 14-60、表 14-61、表 14-62 所示：

表 14-59 Descriptive Statistics

	Mean	Std. Deviation	N
HIGH	171.91	10.005	11
LENGTH	45.45	2.876	11

表 14-59 是因变量“high”及自变量“Length”的统计描述。

表 14-60 ANOVA (b)

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	641.981	1	641.981	16.097	.003(a)
Residual	358.929	9	39.881		
Total	1 000.909	10			

a Predictors: (Constant), LENGTH

b Dependent Variable: HIGH

表 14-60 用方差分析对包含常数项的回归方程作显著性检验，Regression 回归  $SS_{\text{回归}} = 641.981$  其自由度为 1， $MS_{\text{回归}} = 641.981$  Residual 残差  $SS_{\text{残差}} = 358.929$  其自由度为 9， $MS_{\text{残差}} = 39.881$   $F = 16.097$ ， $P = 0.003 < 0.01$ ，可见经  $F$  显著性检验，差异有统计学意义，即此回归方程有意义（此回归方程可建立）。

表 14-61 Coefficients (a)

Model		Unstandardized Coefficients		Standardized	t	Sig.
		B	Std. Error	Coefficients		
1	(Constant)	45.286	31.617		1.432	.186
	LENGTH	2.786	.694	.801	4.012	.003

a Dependent Variable: HIGH

表 14-61 显示回归方程中的系数，常数项为 45.286， $P = 0.186$ ，自变量“length”的系数为 2.786， $P = 0.003 < 0.01$ ，其标准化系数为 0.801。故该回归方程为：

$$\hat{Y}(\text{high}) = 45.286 + 2.786X(\text{length})$$

表 14-62 Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.801(a)	.641	.602	6.315

a Predictors: (Constant), LENGTH

表 14-62 是有关模型的参数，R 表示相关系数为 0.801，R Square 即 R 平方值，也称为确定系数，其公式为：

$$R^2 = \frac{SS_{\text{回归}}}{SS_{\text{总}}}$$

此处  $R^2 = 0.641$  表示因变量的变异中有 64.1% 可由自变量的变化来解释，一般来说  $R^2$  值越大，回归方程越有价值，但是，是否真有意义还有待对回归方程作显著性检验。

## 第九节 多元线性回归

例 14.17 20 名糖尿病人的血糖、胰岛素及生长素的测定值列于下表中，试建立血糖



对于胰岛素及生长素的二元线性回归方程。

表 14-63 糖尿病人的血糖、胰岛素及生长素的测定值

病例号 $i$	血糖 $y(\text{mmol/L})$	胰岛素 ( $x_1, \mu\text{u/L}$ )	生长素 ( $x_2, \mu\text{g/L}$ )
1	12.21	15.2	9.51
2	14.54	16.7	11.43
3	12.27	11.9	7.53
4	12.04	14.0	12.17
5	7.88	19.8	2.33
6	11.10	16.2	13.52
7	10.43	17.0	10.07
8	13.32	10.3	18.89
9	19.59	5.9	13.14
10	9.05	18.7	9.63
11	6.44	25.1	5.10

(一) 建立数据文件

取变量 X1 表示胰岛素

X2 表示生长素

Y 表示血糖

建立数据文件“EG1417.sav”如图 14-60 所示：

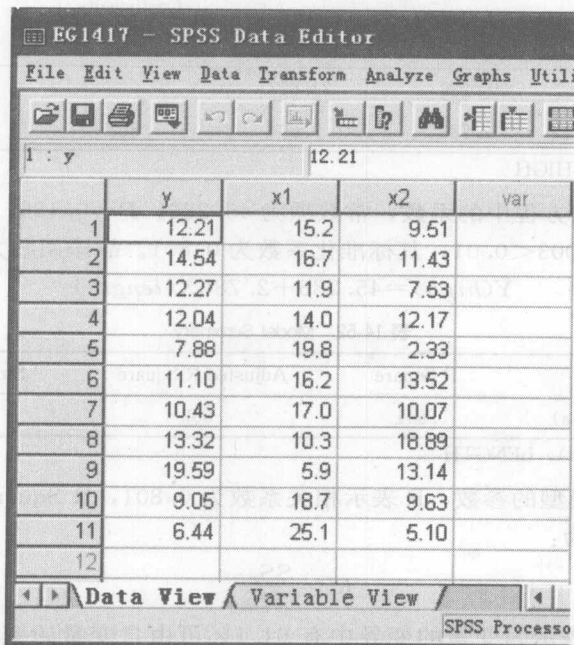


图 14-60 例 14.17 的数据文件

(二) 统计分析

(1) 首先用“One Sample Kolmogorov-Smirnov Test”方法，分别对“Y”、“X1”“X2”作正态性检验，其结果如表 14-64：



表 14-64 One-Sample Kolmogorov-Smirnov Test

		血糖	胰岛素	生长素
N		11	11	11
Normal Parameters(a, b)	Mean	11.7155	15.527	10.3018
	Std. Deviation	3.53406	5.0947	4.43120
Most Extreme Differences	Absolute	.165	.114	.156
	Positive	.165	.114	.143
	Negative	-.085	-.111	-.156
Kolmogorov-Smirnov Z		.547	.377	.519
Asymp. Sig. (2-tailed)		.926	.999	.951

a Test distribution is Normal.

b Calculated from data.

表 14-64 中，“血糖”、“胰岛素”、“生长素”的  $P=0.926$ 、 $P=0.999$ 、 $P=0.951$  都为近似正态分布；

(2) 点击 SPSS for Windows 的主画面的“Analyze”项，展开下拉菜单；在下拉菜单中寻找“Regression”，弹出小菜单，在其上寻找“Linear”，点击之，则弹出线性回归“Linear Regression”对话框，如图 14-61 所示：

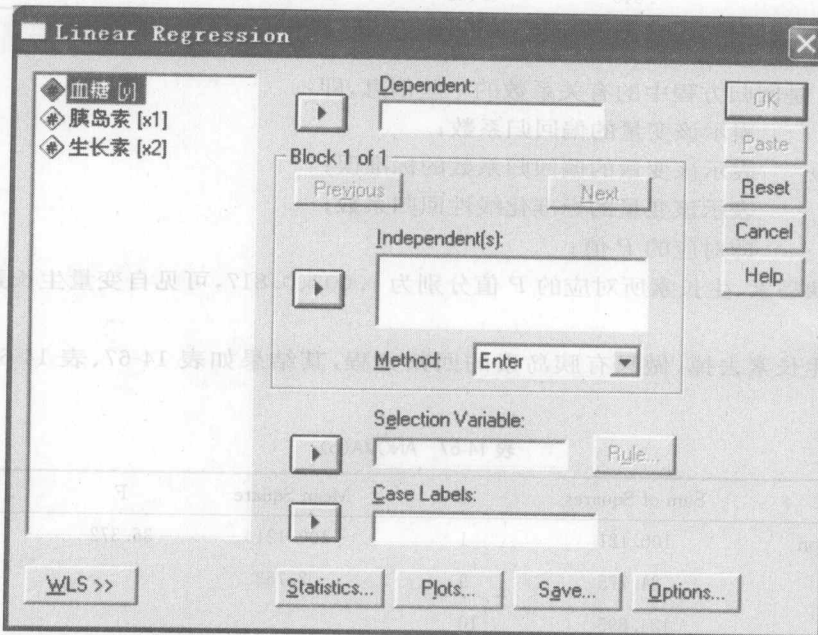


图 14-61 例 14.17 的“Linear Regression”对话框

(3) 将图 14-60 左边的源变量 X1、X2 调入“Independent:”下的对话框；将 Y 调入“Dependent:”下的矩形框；

(4) 单击“OK”按钮，得输出结果，如表 14-65、表 14-66

表 14-65 是用方差分析对回归方程作统计学检验；

左边把总的变异 (Total) 分解为二部分；



表 14-65 ANOVA (b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	100.297	2	50.149	16.310	.002(a)
	Residual	24.598	8	3.075		
	Total	124.895	10			

a Predictors: (Constant), 生长素, 胰岛素

b Dependent Variable: 血糖

Regression 回归 其自由度为 2  $SS_{\text{回归}} = 100.297$   $MS_{\text{回归}} = 50.149$

Residual 残差 其自由度为 8  $SS_{\text{残差}} = 24.568$   $MS_{\text{残差}} = 3.075$

$F = 16.310, P = 0.002 < 0.01$ , 可见经  $F$  检验, 回归方程 (含有常量的) 有统计学意义, 即此回归方程可建立。

表 14-66 Coefficients (a)

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	20.602	3.626		5.682	.000
胰岛素	-.599	.144	-.863	-4.161	.003
生长素	.040	.165	.050	.239	.817

a Dependent Variable: 血糖

表 14-66 是回归方程中的有关系数的一些信息, 即:

$B$  表示该变量的偏回归系数;

$Std\ error$  表示该变量的偏回归系数的标准误;

$Beta$  表示该变量的标准化线性回归系数;

$Sig$  即对应的  $P$  值;

自变量胰岛素、生长素所对应的  $P$  值分别为 0.003、0.817, 可见自变量生长素无统计学意义;

我们把生长素去掉, 做只有胰岛素的回归方程, 其结果如表 14-67、表 14-68、表 14-69 所示:

表 14-67 ANOVA(b)

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	100.121	1	100.121	36.372	.000(a)
Residual	24.775	9	2.753		
Total	124.895	10			

a Predictors: (Constant), 胰岛素

b Dependent Variable: 血糖

表 14-67 是用方差分析对回归方程作统计学检验;

左边把总的变异 (Total) 分解为二部分;

Regression 回归 其自由度为 1  $SS_{\text{回归}} = 100.121$   $MS_{\text{回归}} = 100.121$

Residual 残差 其自由度为 9  $SS_{\text{残差}} = 24.775$   $MS_{\text{残差}} = 2.753$

$F = 36.372, P < 0.001$ , 可见经  $F$  检验, 回归方程 (含有常量的) 有统计学意义, 即此回归



方程可建立。

表 14-68 Coefficients(a)

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	21.359	1.675		12.748	.000
胰岛素	-.621	.103	-.895	-6.031	.000

a. Dependent Variable: 血糖

表 14-68 是回归方程中的有关系数的一些信息；

自变量胰岛素对应的  $P < 0.001$ 、可见自变量胰岛素有统计学意义；可得回归方程为：

$$\hat{Y}(\text{血糖}) = 21.359 - 0.621X_1(\text{胰岛素})$$

表 14-69 Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.895(a)	.802	.780	1.65913

a. Predictors: (Constant), 胰岛素

表 14-69 中  $R$  Square 也称为确定系数。此处  $R^2 = 0.802$ ，表示因变量的变异中有 80.2% 可由自变量的变化来解释，一般来说  $R^2$  值越大，回归方程越有价值。

(马斌荣)

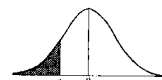
SPSS Ver. 11.5 is a comprehensive system for analyzing data. SPSS can make data file, also can take data from almost any type of file and use them to generate tabulated reports, charts and plots of distributions, descriptive statistics, and complex statistical analysis.

SPSS makes statistical analysis accessible for the casual user and convenient for the experienced user. The Data Editor offers a simple and efficient spreadsheet-like facility for entering data and browsing the working data file. Much of the output from the base system takes the form of flexible pilot tables that you can modify quickly and copy directly into other applications.

Beneath the menus and dialog box, SPSS uses a command language that can be used to create and run production jobs. The help system contains brief syntax diagrams in addition to full help on graphical user interface.

# 附表

附表1 标准正态分布曲线下左侧尾部面积,  $\Phi(u)$  值



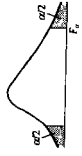
$u$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641





附表2  $t$  分布界值表 (双侧尾部面积)

自由 度 $\nu$	概率, $P$										
	单侧: 双侧:	0.25 0.50	0.20 0.40	0.10 0.20	0.05 0.10	0.025 0.05	0.01 0.02	0.005 0.01	0.0025 0.005	0.001 0.002	0.0005 0.001
1		1.000	1.376	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619
2		0.816	1.061	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3		0.765	0.978	1.638	2.353	3.182	4.541	5.841	7.543	10.215	12.924
4		0.741	0.941	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5		0.727	0.920	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6		0.718	0.906	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7		0.711	0.896	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8		0.706	0.889	1.397	1.860	2.306	2.896	2.355	3.833	4.501	5.041
9		0.703	0.883	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10		0.700	0.879	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11		0.697	0.876	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12		0.695	0.873	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13		0.694	0.870	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14		0.692	0.868	1.345	1.761	2.145	2.624	2.977	3.325	3.787	4.140
15		0.691	0.866	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16		0.690	0.865	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17		0.689	0.863	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18		0.688	0.862	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19		0.688	0.861	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20		0.687	0.860	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21		0.686	0.859	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22		0.686	0.858	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23		0.685	0.858	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.768
24		0.685	0.857	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25		0.684	0.856	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26		0.684	0.856	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27		0.684	0.855	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28		0.683	0.855	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29		0.683	0.854	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30		0.683	0.854	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
31		0.682	0.853	1.309	1.696	2.040	2.453	2.744	3.022	3.375	3.633
32		0.682	0.853	1.309	1.694	2.037	2.449	2.738	3.015	3.365	3.622
33		0.682	0.853	1.308	1.692	2.035	2.445	2.733	3.008	3.356	3.611
34		0.682	0.852	1.307	1.691	2.032	2.441	2.728	3.002	3.348	3.601
35		0.682	0.852	1.306	1.690	2.030	2.438	2.724	2.996	3.340	3.591
36		0.681	0.852	1.306	1.688	2.028	2.434	2.719	2.990	3.333	3.582
37		0.681	0.851	1.305	1.687	2.026	2.431	2.715	2.985	3.326	3.574
38		0.681	0.851	1.304	1.686	2.024	2.429	2.712	2.980	3.319	3.566
39		0.681	0.851	1.304	1.685	2.023	2.426	2.708	2.976	3.313	3.558
40		0.681	0.851	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
50		0.679	0.849	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496
60		0.679	0.848	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
70		0.678	0.847	1.294	1.667	1.994	2.381	2.648	2.899	3.211	3.435
80		0.678	0.846	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
90		0.677	0.846	1.291	1.662	1.987	2.368	2.632	2.878	3.183	3.402
100		0.677	0.845	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
200		0.676	0.843	1.286	1.653	1.972	2.345	2.601	2.839	3.131	3.340
500		0.675	0.842	1.283	1.648	1.965	2.334	2.586	2.820	3.137	3.310
1 000		0.675	0.842	1.282	1.646	1.962	2.330	2.581	2.813	3.098	3.300
$\infty$		0.6745	0.8416	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.0902	3.2905

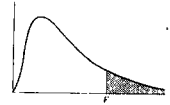


附表3 F分布界值表(方差齐性检验用, 双侧界值) P=0.05

分母的自由度 $v_2$	分子的自由度 $v_1$															
	1	2	3	4	5	6	7	8	9	10	12	15	20	30	60	$\infty$
1	647.79	799.50	864.16	899.58	937.11	937.11	948.22	956.66	963.29	968.63	976.71	984.87	993.10	1 001.41	1 009.80	1 018.26
2	38.51	39.00	39.17	39.25	39.33	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.48	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.08	13.99	13.90
4	12.22	10.05	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.46	8.36	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.23	6.12	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.07	4.96	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.36	4.25	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.89	3.78	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.56	3.45	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.31	3.20	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.12	3.00	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	2.96	2.85	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.84	2.72	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.73	2.61	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.64	2.52	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.57	2.45	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.50	2.38	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.44	2.32	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.39	2.27	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.35	2.22	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.31	2.18	2.04
22	5.79	4.38	3.75	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.27	2.14	2.00
23	5.75	4.35	3.72	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.44	2.36	2.24	2.11	1.97
24	5.72	4.32	3.69	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.21	2.08	1.94
25	5.69	4.29	3.67	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.18	2.05	1.91
26	5.66	4.27	3.65	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.16	2.03	1.88
27	5.63	4.24	3.63	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.13	2.00	1.85
28	5.61	4.22	3.61	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.11	1.98	1.83
29	5.59	4.20	3.59	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.09	1.96	1.81
30	5.57	4.18	3.46	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.07	1.94	1.79
40	5.42	4.05	3.34	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	1.94	1.80	1.64
60	5.29	3.93	3.31	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.82	1.67	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.69	1.53	1.31
$\infty$	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.57	1.39	1.00



附表 4.1 F 界值表 (方差分析用, 单侧界值) 上行:  $P=0.05$  下行:  $P=0.01$



分母的自由度 $\nu_2$	分子的自由度, $\nu_1$											
	1	2	3	4	5	6	7	8	9	10	11	12
1	161	200	216	225	230	234	237	239	241	242	243	224
	4 052	4 999	5 403	5 625	5 764	5 859	5 928	5 981	6 022	6 056	6 082	6 106
2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41
	98.49	99.00	99.17	99.25	99.30	99.33	99.34	99.36	99.38	99.40	99.41	99.42
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74
	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.31	27.05
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91
	21.20	18.00	16.59	15.98	15.52	15.21	14.98	14.80	14.66	14.54	14.45	14.37
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68
	16.26	17.27	12.06	11.39	10.97	10.67	10.45	10.27	10.15	10.05	9.96	9.89
6	5.99	5.15	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00
	13.74	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57
	12.25	9.55	8.45	7.85	7.46	7.19	7.00	6.84	6.71	6.62	6.54	6.47
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28
	11.26	8.65	7.59	7.01	6.63	6.37	6.19	6.03	5.91	5.82	5.74	5.67
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07
	10.56	8.02	6.99	6.42	6.06	5.80	5.62	5.47	5.35	5.26	5.18	5.11
10	4.69	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91
	10.04	7.56	6.55	5.99	5.64	5.39	5.21	5.06	4.95	4.85	4.78	4.71
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79
	9.65	7.20	6.22	5.67	5.32	5.07	4.88	4.74	4.63	4.54	4.46	4.40
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69
	9.33	6.93	5.95	5.41	5.06	4.82	4.65	4.50	4.39	4.30	4.22	4.16
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60
	9.07	6.70	5.74	5.20	4.85	4.62	4.44	4.30	4.19	4.10	4.02	3.96
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53
	8.86	6.51	5.56	5.03	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80
15	4.54	3.68	3.29	3.06	2.90	2.79	2.76	2.64	2.59	2.55	2.51	2.48
	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42
	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.61	3.55
17	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.41	2.38
	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.52	3.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34
	8.28	6.01	5.09	4.58	4.25	4.01	3.85	3.71	3.60	3.51	3.44	3.37
19	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38	2.34	2.31
	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.36	3.30
20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35	2.31	2.28
	8.10	5.85	4.94	4.43	4.10	3.87	3.71	3.56	3.45	3.37	3.30	3.23
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25
	8.02	5.78	4.87	4.37	4.04	3.81	3.65	3.51	3.40	3.31	3.24	3.17
22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30	2.26	2.23
	7.94	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12
23	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.32	2.28	2.24	2.20
	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.14	3.07
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26	2.22	2.18
	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.25	3.17	3.09	3.03
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24	2.20	2.16
	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.21	3.13	3.05	2.99

附表 4.2  $F$  界值表 (方差分析用, 单侧界值)

分母的自由度 $\nu_2$	分子的自由度, $\nu_1$											
	14	16	20	24	30	40	50	75	100	200	500	$\infty$
1	245	246	248	249	250	251	252	253	253	254	254	254
	6.142	6.169	6.208	6.234	6.258	6.286	6.302	6.323	6.334	6.352	6.361	6.366
2	19.42	19.43	19.44	19.45	19.46	19.47	19.47	19.48	19.49	19.49	19.50	19.50
	99.43	99.44	99.45	99.46	99.47	99.48	99.48	99.49	99.49	99.49	99.50	99.50
3	8.71	8.69	8.66	8.64	8.62	8.60	8.58	8.57	8.56	8.54	8.54	8.53
	26.92	26.83	26.69	26.60	26.50	26.41	26.35	26.27	26.23	26.18	26.14	26.12
4	5.87	5.84	5.80	5.77	5.74	5.71	5.70	5.68	5.66	5.65	5.64	5.63
	14.24	14.15	14.02	13.93	13.83	13.74	13.69	13.61	13.57	13.52	13.48	13.46
5	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.40	4.38	4.37	4.36
	9.77	9.68	9.55	9.47	9.38	9.29	9.24	9.17	9.13	9.07	9.04	9.02
6	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.72	3.71	3.69	3.68	3.67
	7.60	7.52	7.39	7.31	7.23	7.14	7.09	7.02	6.99	6.94	6.90	6.88
7	3.52	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.28	3.25	3.24	3.23
	6.35	6.27	6.15	6.07	5.98	5.90	5.85	5.78	5.75	5.70	5.67	5.65
8	3.23	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.94	2.93
	5.56	5.48	5.36	5.28	5.20	5.11	5.06	5.00	4.96	4.91	4.88	4.86
9	3.02	2.98	2.93	2.90	2.86	2.82	2.80	2.77	2.76	2.73	2.72	2.71
	5.00	4.92	4.80	4.73	4.64	4.56	4.51	4.45	4.41	4.36	4.33	4.31
10	2.86	2.82	2.77	2.74	2.70	2.67	2.64	2.61	2.59	2.56	2.55	2.54
	4.60	4.52	4.41	4.33	4.25	4.17	4.12	4.05	4.01	3.96	3.93	3.91
11	2.74	2.70	2.65	2.61	2.57	2.53	2.50	2.47	2.45	2.42	2.41	2.40
	4.29	4.21	4.10	4.02	3.94	3.86	3.80	3.74	3.70	3.66	3.62	3.60
12	2.64	2.60	2.54	2.50	2.46	2.42	2.40	2.36	2.35	2.32	2.31	2.30
	4.05	3.98	3.86	3.78	3.70	3.61	3.56	3.49	3.46	3.41	3.38	3.36
13	2.55	2.51	2.46	2.42	2.38	2.34	2.32	2.28	2.26	2.24	2.22	2.21
	3.85	3.78	3.67	3.59	3.51	3.42	3.37	3.30	3.27	3.21	3.18	3.16
14	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.16	2.14	2.13
	3.70	3.52	3.51	3.43	3.34	3.26	3.21	3.14	3.11	3.06	3.02	3.00
15	2.43	2.39	2.33	2.29	2.25	2.21	2.18	2.15	2.12	2.10	2.08	2.07
	3.56	3.48	3.36	3.29	3.20	3.12	3.07	3.00	2.97	2.92	2.89	2.87
16	2.37	2.33	2.28	2.24	2.20	2.16	2.13	2.09	2.07	2.04	2.02	2.01
	3.45	3.37	3.25	3.18	3.10	3.01	2.96	2.89	2.86	2.80	2.77	2.75
17	2.33	2.29	2.23	2.19	2.15	2.11	2.08	2.04	2.02	1.99	1.97	1.96
	3.35	3.27	3.16	3.08	3.00	2.92	2.86	2.79	2.76	2.70	2.67	2.65
18	2.29	2.25	2.19	2.15	2.11	2.07	2.04	2.00	1.98	1.95	1.93	1.92
	3.27	3.19	3.07	3.00	2.91	2.83	2.78	2.71	2.68	2.62	2.59	2.57
19	2.26	2.21	2.15	2.11	2.07	2.02	2.00	1.96	1.94	1.91	1.90	1.88
	3.19	3.12	3.00	2.92	2.84	2.76	2.70	2.63	2.60	2.54	2.51	2.49
20	2.23	2.18	2.12	2.08	2.04	1.99	1.96	1.92	1.90	1.87	1.85	1.84
	3.13	3.05	2.94	2.86	2.77	2.69	2.63	2.56	2.53	2.47	2.44	2.42
21	2.20	2.15	2.09	2.05	2.00	1.96	1.93	1.89	1.87	1.84	1.82	1.81
	3.07	2.99	2.88	2.80	2.72	2.63	2.58	2.51	2.47	2.42	2.38	2.36
22	2.18	2.13	2.07	2.03	1.98	1.93	1.91	1.87	1.84	1.81	1.80	1.78
	3.02	2.94	2.83	2.75	2.67	2.58	2.53	2.46	2.42	2.37	2.33	2.31
23	2.14	2.10	2.04	2.00	1.96	1.91	1.88	1.84	1.82	1.79	1.77	1.76
	2.97	2.89	2.78	2.70	2.62	2.53	2.48	2.41	2.37	2.32	2.28	2.26
24	2.13	2.09	2.02	1.98	1.94	1.89	1.86	1.82	1.80	1.76	1.74	1.73
	2.93	2.85	2.74	2.66	2.58	2.49	2.44	2.36	2.33	2.27	2.23	2.21
25	2.11	2.06	2.00	1.96	1.92	1.87	1.84	1.80	1.77	1.74	1.72	1.71
	2.89	2.81	2.70	2.62	2.54	2.45	2.40	2.32	2.29	2.23	2.19	2.17

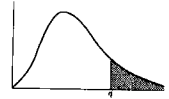
附表 4.3 F 界值表 (方差分析用, 单侧界值)

分母的自由度 $\nu_2$	分子的自由度, $\nu_1$											
	1	2	3	4	5	6	7	8	9	10	11	12
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15
	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.17	3.09	3.02	2.96
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25	2.20	2.16	2.13
	7.68	5.49	4.60	4.11	3.79	3.56	3.39	3.26	3.14	3.06	2.98	2.93
28	4.20	3.34	2.95	2.71	2.56	2.44	2.36	2.29	2.24	2.19	2.15	2.12
	7.64	5.45	4.57	4.07	3.76	3.53	3.36	3.23	3.11	3.03	2.95	2.90
29	4.18	3.33	2.93	2.70	2.54	2.43	2.35	2.28	2.22	2.18	2.14	2.10
	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.08	3.00	2.92	2.87
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21	2.16	2.12	2.09
	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.06	2.98	2.91	2.84
32	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19	2.14	2.10	2.07
	7.50	5.35	4.46	3.97	3.66	3.42	3.25	3.12	3.01	2.94	2.86	2.80
34	4.13	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17	2.12	2.08	2.05
	7.44	5.29	4.42	3.93	3.61	3.38	3.21	3.08	2.98	2.89	2.82	2.76
36	4.11	3.26	2.86	2.63	2.48	2.36	2.28	2.21	2.15	2.10	2.06	2.03
	7.39	5.25	4.38	3.89	3.58	3.35	3.18	3.04	2.94	2.86	2.78	2.72
38	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02
	7.35	5.21	4.31	3.86	3.54	3.32	3.15	3.02	2.91	2.82	2.75	2.69
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.07	2.04	2.00
	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.88	2.80	2.73	2.66
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.02	1.99
	7.27	5.15	4.29	3.80	3.49	3.26	3.10	2.96	2.86	2.77	2.70	2.64
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.96
	7.24	5.12	4.26	3.78	3.46	3.24	3.07	2.94	2.84	2.75	2.68	2.62
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.14	2.09	2.04	2.00	1.97
	7.21	5.10	4.24	3.76	3.44	3.22	3.05	2.92	2.82	2.73	2.66	2.60
48	4.04	3.19	2.80	2.56	2.41	2.30	2.21	2.14	2.08	2.03	1.99	1.96
	7.19	5.08	4.22	3.74	3.42	3.20	3.04	2.90	2.80	2.71	2.64	2.58
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.02	1.98	1.95
	7.17	5.06	4.20	3.72	3.41	3.18	3.02	2.88	2.78	2.70	2.62	2.56
60	4.00	3.15	2.76	2.52	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92
	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.01	1.97	1.93	1.89
	7.01	4.92	4.08	3.60	3.29	3.07	2.91	2.77	2.67	2.59	2.51	2.45
80	3.96	3.11	2.72	2.48	2.33	2.21	2.12	2.05	1.99	1.95	1.91	1.88
	6.96	4.88	4.04	3.56	3.25	3.04	2.87	2.74	2.64	2.55	2.48	2.41
100	3.94	3.09	2.70	2.46	2.30	2.19	2.10	2.03	1.97	1.92	1.88	1.85
	6.90	4.82	3.98	3.51	3.20	2.99	2.82	2.69	2.59	2.51	2.43	2.36
125	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.95	1.90	1.86	1.83
	6.84	4.78	3.94	3.47	3.17	2.95	2.79	2.65	2.56	2.47	2.40	2.33
150	3.91	3.06	2.67	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.85	1.82
	6.81	4.75	3.91	3.44	3.14	2.92	2.76	2.62	2.53	2.44	2.37	2.30
200	3.89	3.04	2.65	2.41	2.26	2.14	2.05	1.98	1.92	1.87	1.83	1.80
	6.76	4.71	3.88	3.41	3.11	2.90	2.73	2.60	2.50	2.41	2.34	2.28
400	3.86	3.02	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.78
	6.70	4.66	3.83	3.36	3.06	2.85	2.69	2.55	2.436	2.37	2.29	2.23
1 000	3.85	3.00	2.61	2.38	2.22	2.10	2.02	1.95	1.89	1.84	1.80	1.76
	6.66	4.62	3.80	3.34	3.04	2.82	2.66	2.53	2.43	2.34	2.26	2.20
$\infty$	3.84	2.99	2.60	2.37	2.21	2.09	2.01	1.94	1.88	1.83	1.79	1.75
	6.64	4.60	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.24	2.18



附表 4.4 F 界值表 (方差分析用, 单侧界值)

分母的自由度 $\nu_2$	分子的自由度, $\nu_1$											
	14	16	20	24	30	40	50	75	100	200	500	$\infty$
26	2.10	2.05	1.99	1.95	1.90	1.85	1.82	1.78	1.76	1.72	1.70	1.69
	2.86	2.77	2.66	2.58	2.50	2.41	2.36	2.28	2.25	2.19	2.15	2.13
27	2.08	2.03	1.97	1.93	1.88	1.84	1.80	1.76	1.74	1.71	1.68	1.67
	2.83	2.74	2.63	2.55	2.47	2.38	2.33	2.25	2.21	2.16	2.12	2.10
28	2.06	2.02	1.96	1.91	1.87	1.81	1.78	1.75	1.72	1.69	1.67	1.65
	2.80	2.71	2.60	2.52	2.44	2.35	2.30	2.22	2.18	2.13	2.09	2.06
29	2.05	2.00	1.94	1.90	1.85	1.80	1.77	1.73	1.71	1.68	1.65	1.64
	2.77	2.68	2.57	2.49	2.41	2.32	2.27	2.19	2.15	2.10	2.06	2.03
30	2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.69	1.66	1.64	1.62
	2.74	2.66	2.55	2.47	2.38	2.29	2.24	2.16	2.13	2.07	2.03	2.01
32	2.02	1.97	1.91	1.86	1.82	1.76	1.74	1.69	1.67	1.64	1.61	1.59
	2.70	2.62	2.51	2.42	2.34	2.25	2.20	2.12	2.08	2.02	1.98	1.96
34	2.00	1.95	1.89	1.84	1.80	1.74	1.71	1.67	1.64	1.61	1.59	1.57
	2.66	2.58	2.47	2.38	2.30	2.21	2.15	2.08	2.04	1.98	1.94	1.91
36	1.98	1.93	1.87	1.82	1.78	1.83	1.69	1.65	1.62	1.59	1.56	1.55
	2.62	2.54	2.43	2.35	2.26	2.17	2.12	2.04	2.00	1.94	1.90	1.87
38	1.96	1.92	1.85	1.80	1.76	1.71	1.67	1.63	1.60	1.57	1.54	1.53
	2.59	2.51	2.40	2.32	2.22	2.14	2.08	2.00	1.97	1.90	1.86	1.84
40	1.95	1.90	1.84	1.79	1.74	1.69	1.66	1.61	1.59	1.55	1.53	1.51
	2.56	2.49	2.37	2.29	2.20	2.11	2.05	1.97	1.94	1.88	1.84	1.81
42	1.94	1.89	1.82	1.78	1.73	1.68	1.64	1.60	1.57	1.54	1.51	1.49
	2.54	2.46	2.35	2.26	2.17	2.08	2.02	1.94	1.91	1.85	1.80	1.78
44	1.82	1.88	1.81	1.76	1.72	1.66	1.63	1.58	1.56	1.52	1.50	1.48
	2.52	2.44	2.32	2.24	2.15	2.06	2.00	1.92	1.88	1.82	1.78	1.75
46	1.91	1.87	1.80	1.75	1.71	1.65	1.62	1.57	1.54	1.51	1.48	1.46
	2.50	2.42	2.30	2.22	2.13	2.04	1.98	1.90	1.86	1.80	1.76	1.72
48	1.90	1.85	1.79	1.74	1.70	1.64	1.61	1.56	1.53	1.50	1.47	1.45
	2.48	2.40	2.28	2.20	2.11	2.02	1.96	1.88	1.84	1.78	1.73	1.70
50	1.90	1.85	1.78	1.74	1.69	1.63	1.60	1.55	1.52	1.48	1.46	1.44
	2.46	2.39	2.26	2.18	2.10	2.00	1.94	1.86	1.82	1.76	1.71	1.68
60	1.86	1.81	1.75	1.70	1.65	1.59	1.56	1.50	1.48	1.44	1.41	1.39
	2.40	2.32	2.20	2.12	2.03	1.93	1.87	1.79	1.74	1.68	1.63	1.60
70	1.84	1.79	1.72	1.67	1.62	1.56	1.53	1.47	1.45	1.40	1.37	1.35
	2.35	2.28	2.15	2.07	1.98	1.88	1.82	1.74	1.69	1.62	1.56	1.53
80	1.82	1.77	1.70	1.65	1.60	1.54	1.51	1.45	1.42	1.38	1.35	1.32
	2.32	2.24	2.11	2.03	1.94	1.84	1.78	1.70	1.65	1.57	1.52	1.49
100	1.79	1.75	1.68	1.63	1.57	1.51	1.48	1.42	1.39	1.34	1.30	1.28
	2.26	2.19	2.06	1.98	1.89	1.79	1.73	1.64	1.59	1.51	1.46	1.43
125	1.77	1.72	1.65	1.60	1.55	1.49	1.45	1.39	1.36	1.31	1.27	1.25
	2.23	2.15	2.03	1.94	1.85	1.75	1.68	1.59	1.54	1.46	1.40	1.37
150	1.76	1.71	1.64	1.59	1.54	1.47	1.44	1.37	1.34	1.29	1.25	1.22
	2.20	2.12	2.00	1.91	1.83	1.72	1.66	1.56	1.51	1.43	1.37	1.33
200	1.74	1.69	1.62	1.57	1.52	1.45	1.42	1.35	1.32	1.26	1.22	1.19
	2.17	2.09	1.97	1.88	1.79	1.69	1.62	1.53	1.48	1.39	1.33	1.28
400	1.72	1.67	1.60	1.54	1.49	1.42	1.38	1.32	1.28	1.22	1.16	1.13
	2.12	2.04	1.92	1.84	1.74	1.64	1.57	1.47	1.42	1.32	1.24	1.19
1 000	1.70	1.65	1.58	1.53	1.47	1.41	1.36	1.30	1.26	1.19	1.13	1.08
	2.09	2.01	1.89	1.81	1.71	1.61	1.54	1.44	1.38	1.28	1.19	1.11
$\infty$	1.69	1.64	1.57	1.52	1.46	1.40	1.35	1.28	1.24	1.17	1.11	1.00
	2.07	1.99	1.87	1.79	1.69	1.59	1.52	1.41	1.36	1.25	1.15	1.00

附表5  $q$  界值表 (Newman-Keuls 法用) 上行:  $P=0.05$  下行:  $P=0.01$ 

$\nu$	组数, $a$								
	2	3	4	5	6	7	8	9	10
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99
	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24
6	3.46	4.34	4.90	5.30	5.63	5.89	6.12	6.32	6.49
	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16
	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92
	4.75	5.64	6.20	6.62	6.96	7.24	7.47	7.68	7.86
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74
	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60
	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39
	4.32	5.05	5.50	5.84	6.10	6.12	6.51	6.67	6.81
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25
	4.21	4.89	5.32	5.63	5.88	6.08	5.26	6.41	6.54
16	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15
	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35
18	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07
	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20
20	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01
	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09
30	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82
	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73
	2.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65
	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45
120	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56
	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30
$\infty$	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47
	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16



附表 6.1 百分率的置信区间 上行: 95%置信区间 下行: 99%置信区间

n	x													
	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0~98													
	0~100													
2	0~84	1~99												
	0~93	0~100												
3	0~71	1~91	9~99											
	0~83	0~96	4~100											
4	0~60	1~81	7~93											
	0~73	0~89	3~97											
5	0~52	1~72	5~85	15~95										
	0~65	0~81	2~92	8~98										
6	0~46	0~64	4~78	12~88										
	0~59	0~75	2~86	7~93										
7	0~41	0~58	4~71	10~82	18~90									
	0~53	0~68	2~80	6~88	12~94									
8	0~37	0~53	3~65	9~76	16~84									
	0~48	0~63	1~74	5~83	10~90									
9	0~34	0~48	3~60	7~70	14~79	21~86								
	0~45	0~59	1~69	4~78	9~85	15~91								
10	0~31	0~45	3~56	7~65	12~74	19~81								
	0~41	0~54	1~65	4~74	8~81	13~87								
11	0~28	0~40	2~52	6~61	11~69	17~77	23~83							
	0~38	0~51	1~61	3~69	7~77	11~83	17~89							
12	0~26	0~38	2~48	5~57	10~65	15~72	21~79							
	0~36	0~48	1~57	3~66	6~73	10~79	15~85							
13	0~25	0~36	2~45	5~54	9~61	14~68	19~75	25~81						
	0~34	0~45	1~54	3~62	6~69	9~76	14~81	19~86						
14	0~23	0~34	2~43	5~51	8~58	13~65	18~71	23~77						
	0~32	0~42	1~51	3~59	5~66	9~72	13~78	17~83						
15	0~22	0~32	2~41	4~48	8~55	12~62	16~68	21~73	27~79					
	0~30	0~40	1~49	2~56	5~63	8~69	12~74	16~79	21~84					
16	0~21	0~30	2~38	4~46	7~52	11~59	15~65	20~70	25~75					
	0~28	0~38	1~46	2~53	5~60	8~66	11~71	15~76	19~81					
17	0~20	0~29	2~36	4~34	7~50	10~56	14~62	18~67	23~72	28~77				
	0~27	0~36	1~44	2~51	4~57	7~63	10~69	14~74	18~78	22~82				
18	0~19	0~27	1~35	3~41	6~48	10~54	13~59	17~64	22~69	26~74				
	0~26	0~35	1~42	2~49	4~55	7~61	10~66	13~71	17~75	21~79				
19	0~18	0~26	1~33	3~40	6~46	9~51	13~57	16~62	20~67	24~71	29~76			
	0~24	0~33	1~40	2~47	4~53	6~58	9~63	12~68	16~73	19~77	23~81			
20	0~17	0~25	1~32	3~38	6~44	9~49	12~54	15~59	19~64	23~69	27~73			
	0~23	0~32	1~39	2~45	4~51	6~56	9~61	11~66	15~70	18~74	22~78			
21	0~16	0~24	1~30	3~36	5~42	8~47	11~52	15~57	18~62	22~66	26~70	30~74		
	0~22	0~30	1~37	2~43	3~49	6~54	8~59	11~63	14~68	17~71	21~76	24~80		
22	0~15	0~23	1~29	3~35	5~40	8~45	11~50	14~55	17~59	21~64	24~68	28~72		
	0~21	0~29	1~36	2~42	3~47	5~52	8~57	10~61	13~66	16~70	20~73	23~77		
23	0~15	0~22	1~28	3~34	5~39	8~44	10~48	13~53	16~57	20~62	23~66	27~69	31~73	
	0~21	0~28	1~35	2~40	3~45	5~50	7~55	10~59	13~63	15~67	19~71	22~75	25~78	
24	0~14	0~21	1~27	3~32	5~37	7~42	10~47	13~51	16~55	19~59	22~63	26~67	29~71	
	0~20	0~27	0~33	2~39	3~44	5~49	7~53	9~57	12~61	15~65	18~69	21~73	24~76	
25	0~14	0~20	1~26	3~31	5~36	7~41	9~45	12~49	15~54	18~58	21~61	24~65	28~69	31~72
	0~19	0~16	0~32	1~37	3~42	5~47	7~51	9~56	11~60	14~63	17~67	20~71	23~74	26~77



附表 6.2 百分率的置信区间 上行: 95%置信区间 下行: 99%置信区间

n	x													
	0	1	2	3	4	5	6	7	8	9	10	11	12	13
26	0~13	0~20	1~25	2~30	4~35	7~39	9~44	12~48	14~52	17~56	20~60	23~63	27~67	30~70
	0~18	0~25	1~31	1~36	3~41	4~46	5~50	9~54	11~58	13~62	16~65	19~69	22~72	25~75
27	0~13	0~19	1~24	2~29	4~34	6~38	9~42	11~46	19~50	17~54	19~58	22~61	26~65	29~68
	0~18	0~25	0~30	1~35	3~40	4~44	6~48	8~52	10~56	13~60	15~63	18~67	21~70	24~73
28	0~12	0~18	1~24	2~28	4~33	6~37	8~41	11~45	13~49	16~52	19~56	22~59	25~63	28~66
	0~17	0~24	0~29	1~34	3~39	4~43	6~47	8~51	10~55	12~58	15~62	17~65	20~68	23~71
29	0~12	0~18	1~23	2~27	4~32	6~36	8~40	10~44	13~47	15~51	18~54	21~58	24~61	26~64
	0~17	0~23	0~28	1~33	2~37	4~42	6~46	8~49	10~53	12~57	14~60	17~63	19~66	22~70
30	0~12	0~17	1~22	2~27	4~31	6~35	8~39	10~42	12~46	15~49	17~53	20~56	23~59	26~43
	0~16	0~22	0~27	1~32	2~36	4~40	5~44	7~48	9~52	11~55	14~58	16~62	19~65	21~68
31	0~11	0~17	1~22	2~26	4~30	6~34	8~38	10~41	12~45	14~48	17~51	19~55	22~58	25~61
	0~16	0~22	0~27	1~31	2~35	4~39	5~43	7~47	9~50	11~54	13~57	16~60	18~63	20~66
32	0~11	0~16	1~21	2~25	4~29	5~33	7~36	9~40	12~43	14~47	16~50	19~53	21~56	24~59
	0~15	0~21	0~26	1~30	2~34	4~38	5~42	7~46	9~49	11~52	13~56	15~59	17~62	20~65
33	0~11	0~15	1~20	2~24	3~28	5~32	7~36	9~39	11~42	13~46	16~49	18~52	20~55	23~58
	0~15	0~20	0~25	130	2~34	3~37	5~41	7~44	8~48	10~51	12~54	14~57	17~60	19~63
34	0~10	0~15	1~19	2~23	3~28	5~31	7~35	9~38	11~41	13~44	15~48	17~51	20~54	22~56
	0~14	0~20	0~25	1~29	2~33	3~36	5~40	6~43	8~47	10~50	12~53	14~56	16~59	18~62
35	0~10	0~15	1~19	2~23	3~27	5~30	6~34	8~37	10~40	13~43	15~46	17~49	19~52	22~55
	0~14	0~20	0~24	1~28	2~32	3~35	5~39	6~42	8~45	10~49	12~52	14~55	16~57	18~60
36	0~10	0~15	1~18	2~22	3~26	5~29	6~33	8~36	10~39	12~42	14~45	16~48	19~51	21~54
	0~14	0~19	0~23	1~27	2~31	3~35	5~38	6~41	8~44	9~47	11~50	13~53	15~56	17~59
37	0~10	0~14	1~18	2~22	3~25	5~28	6~32	8~35	10~38	12~41	14~44	16~47	18~50	20~54
	0~13	0~18	0~23	1~27	2~30	3~34	4~37	6~40	7~43	9~46	11~49	13~52	15~55	17~58
38	0~10	0~14	1~18	2~21	3~25	5~28	6~32	8~34	10~37	11~40	13~43	15~46	18~49	20~51
	0~13	0~18	0~22	1~26	2~30	3~33	4~36	6~39	7~42	9~45	11~48	12~51	14~54	16~56
39	0~9	0~14	1~17	2~21	3~24	4~27	6~31	8~33	9~36	11~39	13~42	15~45	17~48	19~50
	0~13	0~18	0~21	1~25	2~29	3~32	4~35	6~38	7~41	9~44	10~47	12~49	14~53	16~55
40	0~9	0~13	1~17	2~21	3~24	4~27	6~30	8~33	9~35	11~38	13~41	15~44	14~47	19~49
	0~12	0~17	0~21	1~25	2~28	3~32	4~35	5~38	7~40	9~43	10~46	12~49	13~52	15~54
41	0~9	0~13	1~17	2~20	3~23	4~26	6~29	7~32	9~35	11~37	12~40	14~43	16~46	18~48
	0~12	0~17	0~21	1~24	2~28	3~31	4~34	5~37	7~40	8~42	10~45	11~48	13~50	15~53
42	0~9	0~13	1~16	2~20	3~23	4~26	6~28	7~31	9~34	10~37	12~39	14~42	16~45	18~47
	0~12	0~17	0~20	1~24	2~27	3~30	4~33	5~36	7~39	8~42	9~44	11~47	13~49	15~52
43	0~9	0~12	1~16	2~19	3~23	4~25	5~28	7~31	8~33	10~36	12~39	14~41	15~44	17~45
	0~12	0~16	0~20	1~23	2~26	3~30	4~33	5~35	6~38	8~41	9~43	11~46	13~49	14~51
44	0~9	0~12	1~15	2~19	3~22	4~25	5~28	7~30	8~33	10~35	11~38	13~40	15~43	17~45
	0~11	0~16	0~19	1~23	2~26	3~29	4~32	5~35	6~37	8~40	9~42	11~45	12~47	14~51
45	0~8	0~12	1~15	2~18	3~21	4~24	5~27	7~30	8~32	9~34	11~37	13~39	15~42	16~44
	0~11	0~15	0~19	1~22	2~25	3~28	4~31	5~34	6~37	8~39	9~42	10~44	12~47	14~49
46	0~8	0~12	1~15	2~18	3~21	4~24	5~26	7~29	8~31	9~34	11~36	13~39	14~41	16~43
	0~11	0~15	0~19	1~22	2~25	3~28	4~31	5~33	6~36	7~39	9~41	10~43	12~46	13~48
47	0~8	0~12	1~15	2~17	3~20	4~23	6~26	6~28	8~31	9~34	11~36	12~38	14~40	16~43
	0~11	0~15	0~18	1~21	2~24	2~27	3~30	5~33	6~35	7~38	9~40	10~42	11~45	13~47
48	0~8	0~11	1~14	2~17	3~20	4~22	5~25	6~28	8~30	9~33	11~35	12~37	14~40	15~42
	0~10	0~14	0~18	1~21	2~24	2~27	3~29	5~32	6~35	7~37	8~40	10~42	11~44	13~47
49	0~8	0~11	1~14	2~17	2~20	4~22	5~25	6~27	7~30	9~32	10~35	12~37	13~39	15~41
	0~10	0~14	0~17	1~20	1~24	2~26	3~29	4~32	6~34	7~36	8~39	9~41	11~44	12~46
50	0~7	0~11	1~14	2~17	2~19	3~22	5~24	6~26	7~29	9~31	10~34	11~36	13~38	15~41
	0~10	0~14	0~17	1~20	1~23	2~26	3~28	4~31	5~33	7~36	8~38	9~40	11~43	12~45



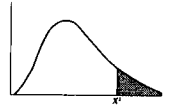
附表 6.3 百分率的置信区间 上行: 95%置信区间 下行: 99%置信区间

n	x											
	14	15	16	17	18	19	20	21	22	23	24	25
26												
27	32~71											
	27~76											
28	31~69											
	26~74											
29	30~68	33~71										
	25~72	28~75										
30	28~66	31~69										
	24~71	27~74										
31	27~64	30~67	33~70									
	23~69	26~72	28~75									
32	26~62	29~65	32~68									
	22~67	25~70	27~73									
33	26~61	28~64	31~67	34~69								
	21~66	24~69	26~71	29~74								
34	25~59	27~62	30~65	32~68								
	21~64	23~67	25~70	28~72								
35	24~58	26~61	29~63	31~66	34~69							
	20~63	22~66	24~68	27~71	29~73							
36	23~57	26~59	28~62	30~65	33~67							
	19~62	22~64	23~67	26~69	28~72							
37	23~55	25~58	27~61	30~63	32~66	34~68						
	19~60	21~63	23~65	25~68	28~70	30~73						
38	22~54	24~57	26~59	29~62	31~64	33~67						
	18~59	20~61	22~64	25~66	27~69	29~71						
39	21~53	23~55	26~58	28~60	30~63	32~65	35~68					
	18~58	20~60	22~63	24~65	26~68	28~70	30~72					
40	21~52	23~54	25~57	27~59	29~62	32~64	34~66					
	17~57	19~59	21~61	23~64	25~66	27~68	30~71					
41	20~51	22~53	24~56	26~58	29~60	31~63	33~65	35~67				
	17~55	19~58	21~60	23~63	25~65	27~67	29~69	31~71				
42	20~50	22~52	24~54	26~57	28~59	30~61	32~64	34~66				
	16~54	18~57	20~59	22~61	24~64	26~66	28~67	30~70				
43	19~49	21~51	23~53	25~56	27~58	29~60	31~62	33~65	36~67			
	16~53	18~56	19~58	21~60	23~62	25~65	27~66	29~69	31~71			
44	19~48	21~50	22~52	24~55	26~57	28~59	30~61	33~63	35~65			
	15~52	14~55	19~57	21~59	23~61	25~63	26~65	28~68	30~70			
45	18~47	20~49	22~51	24~54	26~56	28~58	30~60	32~62	34~64	36~66		
	15~51	17~54	19~56	20~58	22~60	24~62	26~64	28~66	30~68	32~70		
46	18~46	20~48	21~50	23~53	25~55	27~57	29~59	31~61	33~63	35~65		
	15~50	16~53	18~55	20~57	22~59	23~61	25~63	27~65	29~67	31~69		
47	18~45	19~47	21~49	23~52	25~54	26~56	28~58	30~60	32~62	34~64	36~66	
	14~19	16~52	18~54	19~56	21~58	23~60	25~62	26~64	28~66	30~68	32~70	
48	17~44	19~46	21~48	22~51	24~53	26~53	28~57	30~59	31~61	33~63	35~65	
	14~49	16~51	17~53	19~55	21~27	22~59	24~61	26~63	28~65	29~67	31~69	
49	17~43	18~45	20~47	22~50	24~52	25~54	27~56	29~58	31~60	33~62	34~64	36~66
	14~48	15~50	17~52	19~54	20~56	22~58	23~60	25~62	27~64	29~66	31~68	32~70
50	16~43	18~45	20~47	21~49	23~51	25~63	26~55	28~57	30~59	32~61	34~63	36~65
	14~47	15~49	17~51	18~53	20~55	21~57	23~59	25~61	26~63	28~65	30~67	32~68



附表 6.4 百分率的置信限 ( $50 \leq n \leq 100$ ) 上行: 95%置信限 下行: 99%置信限

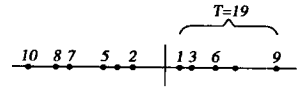
k	n						k	n					
	50	60	70	80	90	100		50	60	70	80	90	100
1	0~11	0~9	0~8	0~7	0~6	0~5	26	31~57	26~49	22~44	20~39	18~36	
	0~14	0~12	0~10	0~9	0~8	0~7		27~61	23~53	20~48	17~43	16~39	
2	0~14	1~11	0~10	1~9	0~8	0~7	27	32~58	27~51	24~45	21~40	19~37	
	0~17	0~14	0~13	0~11	0~10	0~9		29~62	24~55	21~49	18~44	16~40	
3	1~17	1~14	1~12	1~11	1~10	1~8	28	34~60	29~52	25~46	22~42	20~38	
	1~20	1~17	1~15	1~13	0~12	0~10		30~64	25~56	22~50	19~45	17~41	
4	2~19	2~16	2~14	2~13	1~11	1~10	29	35~62	30~54	26~48	23~43	20~39	
	1~23	1~20	1~17	1~15	1~14	1~12		32~65	27~57	23~51	20~46	18~42	
5	3~22	3~18	3~16	2~14	2~13	2~11	30	37~63	31~55	27~49	24~44	21~40	
	2~26	2~22	2~19	1~17	1~15	1~13		33~67	28~59	24~53	21~47	19~43	
6	5~24	4~20	3~18	3~16	3~14	2~12	31	33~57	28~50	25~45	22~41	20~36	
	3~29	3~24	2~21	2~19	2~17	2~14		29~60	25~45	22~49	20~44	18~39	
7	6~27	5~23	4~20	4~17	3~15	3~14	32	34~58	29~51	26~46	23~42	21~37	
	4~31	4~26	3~23	3~21	2~18	2~16		30~62	26~55	23~50	21~45	19~40	
8	7~29	6~25	5~21	5~19	4~17	4~15	33	35~59	31~53	27~47	24~43	22~38	
	6~33	4~29	4~25	3~22	3~20	3~17		32~63	27~56	24~51	21~46	19~41	
9	9~31	7~26	6~23	5~20	5~18	4~16	34	36~61	32~54	28~48	25~44	23~39	
	7~36	5~30	5~27	4~24	4~21	3~18		33~64	28~58	25~52	22~47	20~42	
10	10~34	8~29	7~25	6~22	6~20	5~18	35	38~62	33~55	29~50	26~45	24~39	
	8~38	7~32	6~28	5~25	4~22	4~19		34~66	30~59	26~53	23~48	21~43	
11	12~36	10~30	8~26	7~23	6~21	5~19	36	34~56	30~51	27~40	25~39	23~38	
	10~40	8~34	7~30	6~21	5~24	4~20		31~60	27~54	24~49	22~44	20~39	
12	13~38	11~32	9~28	8~25	7~22	6~20	37	35~58	31~52	28~47	26~46	24~45	
	11~43	9~36	7~32	6~28	6~25	5~21		32~61	28~55	25~50	23~48	21~46	
13	15~41	12~34	10~30	9~26	8~23	7~21	38	36~59	32~53	29~48	27~47	25~46	
	12~45	10~38	8~33	7~30	6~27	6~23		33~62	29~56	26~51	24~50	22~49	
14	16~43	13~36	11~31	10~27	9~25	8~22	39	37~60	33~54	29~49	27~52	25~51	
	14~47	11~40	9~35	8~31	7~28	6~24		34~64	30~57	27~52	25~51	23~50	
15	18~44	15~38	13~33	11~29	10~26	9~24	40	39~61	34~55	30~50	28~53	26~52	
	15~49	12~42	10~37	9~33	8~30	7~26		35~65	31~59	28~53	26~52	24~51	
16	20~46	16~40	14~34	12~30	11~27	9~25	41	35~56	31~51	28~46	26~54	24~53	
	17~51	14~44	11~38	10~34	9~31	8~27		32~60	29~54	27~53	25~52	23~51	
17	21~48	14~81	15~36	13~32	12~28	10~26	42	36~57	32~52	29~47	27~54	25~53	
	18~53	15~46	12~40	11~35	10~32	9~29		33~61	30~55	28~54	26~53	24~52	
18	23~50	19~43	16~37	14~33	12~30	11~27	43	37~59	33~53	30~48	28~55	26~54	
	20~55	16~47	14~41	12~37	10~33	9~30		34~62	31~56	29~54	27~53	25~52	
19	25~53	20~45	17~38	15~34	13~31	12~28	44	38~60	34~54	31~49	29~56	27~55	
	21~57	17~49	15~43	13~38	11~35	10~31		35~63	32~57	30~56	28~55	26~54	
20	27~55	22~47	18~40	16~36	14~32	13~29	45	39~61	35~55	32~50	30~57	28~56	
	23~59	19~51	16~44	14~39	12~36	11~32		36~64	33~58	31~57	29~56	27~55	
21	28~57	23~49	20~41	17~37	15~33	14~30	46	36~56	33~57	31~58	29~59	27~58	
	24~61	20~52	17~46	15~41	13~31	12~33		33~59	32~58	30~59	28~59	26~58	
22	30~59	25~50	21~43	18~39	16~35	14~31	47	37~57	34~57	32~60	30~61	28~60	
	26~63	22~54	18~47	16~42	14~38	12~34		34~60	33~60	31~61	29~61	27~60	
23	32~61	26~52	22~45	19~40	17~36	15~32	48	38~58	35~58	33~61	31~62	29~61	
	28~65	23~56	19~49	17~44	15~39	13~35		35~61	34~61	32~62	30~62	28~61	
24	34~63	28~53	23~46	20~41	18~37	16~33	49	39~59	36~60	34~62	32~63	30~62	
	29~67	24~58	21~50	18~45	16~41	14~36		36~62	35~61	33~63	31~64	29~63	
25	36~64	29~55	25~48	21~43	19~38	17~35	50	40~60	37~61	35~62	33~64	31~63	
	31~69	26~59	22~52	19~46	17~42	15~38		37~63	36~62	34~63	32~64	30~64	

附表7  $\chi^2$  分布界值表

自由度 $\nu$	概率, $P$ (右侧尾部面积)												
	0.995	0.990	0.975	0.950	0.900	0.750	0.500	0.250	0.100	0.050	0.025	0.010	0.005
1					0.02	0.10	0.45	1.32	2.71	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	0.21	0.58	1.39	2.77	4.11	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	0.58	1.21	2.37	4.11	6.25	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	1.06	1.92	3.36	5.39	7.78	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	1.61	2.67	4.35	6.63	9.24	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.09	21.95
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	11.04	14.34	18.25	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.49	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.60	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.72	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.83	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	16.34	20.34	24.93	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	17.24	21.34	26.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	18.14	22.34	27.14	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	19.04	23.34	28.24	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	19.94	24.34	29.34	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	20.84	25.34	30.43	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	21.75	26.34	31.53	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	22.66	27.34	32.62	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	23.57	28.34	33.71	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	24.48	29.34	34.80	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	33.66	39.34	45.62	51.81	55.70	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	42.94	49.33	56.33	63.17	67.50	70.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	52.29	59.33	66.98	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	61.70	69.33	77.58	85.53	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	64.28	71.14	79.33	88.13	96.58	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	73.29	80.62	89.33	98.64	107.56	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	82.36	90.13	99.33	109.14	118.50	124.34	129.56	135.81	140.17



附表8 T界值表(配对比较的符号秩和检验用)

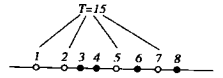


n	单侧:0.05	0.025	0.01	0.005
	双侧:0.10	0.05	0.02	0.010
5	1~15(0.0312)			
6	2~19(0.0469)	0~21(0.0156)		
7	3~25(0.0391)	3~26(0.0234)	0~28(0.0078)	
8	5~31(0.0391)	3~33(0.0195)	1~35(0.0078)	0~36(0.0039)
9	8~37(0.0488)	5~40(0.0195)	3~42(0.0098)	1~44(0.0039)
10	10~45(0.0420)	8~47(0.0244)	5~50(0.0098)	3~52(0.0049)
11	13~53(0.0415)	10~56(0.0210)	7~59(0.0093)	5~61(0.0019)
12	17~61(0.0461)	13~65(0.0212)	9~69(0.0081)	7~71(0.0046)
13	21~70(0.0471)	17~74(0.0239)	12~79(0.0085)	9~82(0.0040)
14	25~80(0.0453)	21~84(0.0247)	15~90(0.0083)	12~93(0.0043)
15	30~90(0.0473)	25~95(0.0240)	19~101(0.0090)	15~105(0.0042)
16	35~101(0.0467)	29~107(0.0222)	23~113(0.0091)	19~117(0.0046)
17	41~112(0.0492)	34~119(0.0224)	27~126(0.0087)	23~130(0.0047)
18	47~124(0.0494)	40~131(0.0241)	32~139(0.0091)	27~144(0.0045)
19	53~137(0.0478)	46~144(0.0247)	37~153(0.0090)	32~158(0.0047)
20	60~150(0.0487)	52~158(0.0242)	43~167(0.0096)	37~123(0.0047)
21	67~164(0.0479)	58~173(0.0230)	49~182(0.0097)	42~189(0.0045)
22	75~178(0.0492)	65~188(0.0231)	55~198(0.0095)	48~205(0.0046)
23	83~193(0.0490)	73~203(0.0242)	62~214(0.0098)	54~222(0.0046)
24	91~209(0.0475)	81~219(0.0245)	69~231(0.0097)	61~239(0.0048)
25	100~225(0.0479)	89~236(0.0241)	76~249(0.0094)	68~257(0.0048)

注:括号内为单侧确切概率



附表9 T界值表(两样本比较的秩和检验用)

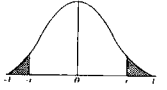


单侧 双侧 单侧 双侧  
 1行  $P=0.05$   $P=0.10$ ; 2行  $P=0.025$   $P=0.05$   
 3行  $P=0.01$   $P=0.02$ ; 4行  $P=0.005$   $P=0.01$

$n_1$ (较小 $n$ )	$n_2 - n_1$										
	0	1	2	3	4	5	6	7	8	9	10
2				3~13	3~15	3~17	4~18	4~20	4~22	4~24	5~25
							3~19	3~21	3~23	3~25	4~26
3	6~15	6~18	7~20	8~22	8~25	9~27	10~29	10~32	11~34	11~37	12~39
			6~21	7~23	7~26	8~28	8~31	9~33	9~36	10~38	10~41
					6~27	6~30	7~32	7~35	7~38	8~40	8~43
							6~33	6~36	6~39	7~41	7~44
4	11~25	12~28	13~31	14~34	15~37	16~40	17~43	18~46	19~49	20~52	21~55
	10~26	11~29	12~32	13~35	14~38	14~42	15~45	16~48	17~51	18~54	19~57
		10~30	11~33	11~37	12~40	13~43	13~47	14~50	15~53	15~57	16~60
			10~34	10~38	11~41	11~45	12~48	12~52	13~55	13~59	14~62
5	19~36	20~40	21~44	23~47	24~51	26~54	27~58	28~62	30~65	31~69	33~72
	17~38	18~42	20~45	21~49	22~53	23~57	24~61	26~64	27~68	28~72	29~76
	16~39	17~43	18~47	19~51	20~55	21~59	22~63	23~67	24~71	25~75	26~79
	15~40	16~44	16~49	17~53	18~57	19~61	20~65	21~69	22~73	22~78	23~82
6	28~50	29~55	31~59	33~63	35~67	37~71	38~76	40~80	42~84	44~88	46~92
	26~52	27~57	29~61	31~65	32~70	34~74	35~79	37~83	38~88	40~92	42~96
	24~54	25~59	27~63	28~68	29~73	30~78	32~82	33~87	34~92	36~96	37~101
	23~55	24~60	25~65	26~70	27~75	28~80	30~84	31~89	32~94	33~99	32~104
7	39~66	41~71	43~76	45~81	47~86	49~91	52~95	54~100	46~105	58~110	61~114
	36~69	38~74	40~79	42~84	44~89	46~94	48~99	50~104	52~109	54~114	56~119
	34~71	35~77	37~82	39~87	40~93	42~98	44~103	45~109	47~114	49~119	51~124
	32~73	34~78	35~84	37~89	38~95	40~100	41~106	43~111	44~117	45~122	47~128
8	51~85	54~90	56~96	59~101	62~106	64~112	67~117	69~123	72~128	75~133	77~139
	49~87	51~93	53~99	55~105	58~110	60~116	62~122	65~127	67~133	70~138	72~144
	45~91	47~97	49~103	51~109	53~115	56~120	58~126	60~132	62~138	64~144	66~150
	43~93	45~99	47~105	49~111	51~117	53~123	54~130	56~136	58~142	60~148	62~154
9	66~105	69~111	72~117	75~123	78~129	81~135	84~141	87~147	90~153	93~159	96~165
	62~109	65~115	68~121	71~127	73~134	76~140	79~146	82~152	84~159	87~165	90~171
	59~112	61~119	63~126	66~132	68~139	71~145	73~152	76~158	78~165	81~171	83~178
	56~115	58~122	61~128	63~135	65~142	67~149	69~156	72~162	74~169	76~176	78~183
10	82~128	86~134	89~141	92~148	96~154	99~161	103~167	106~174	110~180	113~187	117~193
	78~132	81~139	84~146	88~152	91~159	94~166	97~173	100~180	103~187	107~193	110~200
	74~136	77~143	79~151	82~158	85~165	88~172	91~179	93~187	96~194	99~201	102~208
	71~139	73~147	76~154	79~161	81~169	84~176	86~184	89~191	92~198	94~206	97~213

附表 10  $H$  界值表 (三样本比较的秩和检验用)

$n$	$n_1$	$n_2$	$n_3$	$P$	
				0.05	0.01
7	3	2	2	4.71	
	3	3	1	5.14	
8	3	3	2	5.36	
	4	2	2	5.33	
	4	3	1	5.21	
	5	2	1	5.00	
	3	3	3	5.60	7.20
9	4	3	2	5.44	6.44
	4	4	1	4.97	6.67
	5	2	2	5.16	6.53
	5	3	1	4.96	
	4	3	3	5.73	6.75
	4	4	2	5.49	7.04
10	5	3	2	5.25	6.82
	5	4	1	4.99	6.95
	4	4	3	5.60	7.14
	5	3	3	5.65	7.08
	5	4	2	5.27	7.12
11	5	5	1	5.13	7.31
	4	4	4	5.69	7.65
	5	4	3	5.63	7.44
	5	5	2	5.34	7.27
12	5	4	4	5.62	7.76
	5	5	3	5.71	7.54
	5	5	4	5.64	7.79
13	5	5	5	5.78	7.98

附表 11  $t$  界值表 (双侧尾部面积)

自由度 $\nu$	概率, $P$									
	单侧: 双侧:	0.25 0.50	0.10 0.20	0.05 0.10	0.025 0.05	0.01 0.02	0.005 0.01	0.0025 0.005	0.001 0.002	0.000 0.001
1		0.707	0.951	0.988	0.997	1.000	1.000	1.000	1.000	1.000
2		0.500	0.800	0.900	0.950	0.980	0.990	0.995	0.998	0.999
3		0.404	0.687	0.805	0.878	0.934	0.959	0.974	0.986	0.991
4		0.347	0.608	0.729	0.811	0.882	0.917	0.942	0.963	0.974
5		0.309	0.551	0.669	0.755	0.833	0.875	0.906	0.935	0.951
6		0.281	0.507	0.621	0.707	0.789	0.834	0.870	0.905	0.925
7		0.260	0.472	0.582	0.666	0.750	0.798	0.836	0.875	0.898
8		0.242	0.443	0.549	0.632	0.715	0.765	0.805	0.847	0.882
9		0.228	0.419	0.521	0.602	0.685	0.735	0.776	0.820	0.847
10		0.216	0.398	0.497	0.576	0.658	0.708	0.750	0.795	0.823
11		0.206	0.380	0.476	0.553	0.634	0.684	0.726	0.772	0.801
12		0.197	0.365	0.457	0.532	0.612	0.661	0.703	0.750	0.780
13		0.189	0.351	0.441	0.514	0.592	0.641	0.683	0.730	0.760
14		0.182	0.338	0.426	0.497	0.574	0.623	0.664	0.711	0.742
15		0.176	0.327	0.412	0.482	0.558	0.606	0.647	0.694	0.725
16		0.170	0.317	0.400	0.468	0.542	0.590	0.631	0.678	0.708
17		0.165	0.308	0.389	0.456	0.529	0.575	0.616	0.662	0.693
18		0.160	0.299	0.378	0.444	0.515	0.561	0.602	0.648	0.679
19		0.156	0.291	0.369	0.433	0.503	0.549	0.589	0.635	0.665
20		0.152	0.284	0.360	0.423	0.492	0.537	0.576	0.622	0.652
21		0.148	0.277	0.352	0.413	0.482	0.526	0.565	0.610	0.640
22		0.145	0.271	0.344	0.404	0.472	0.515	0.554	0.599	0.629
23		0.141	0.265	0.337	0.396	0.462	0.505	0.543	0.588	0.618
24		0.138	0.260	0.330	0.388	0.453	0.496	0.534	0.578	0.607
25		0.136	0.255	0.323	0.381	0.445	0.487	0.524	0.568	0.597
26		0.133	0.250	0.317	0.374	0.437	0.479	0.515	0.559	0.588
27		0.131	0.245	0.311	0.367	0.430	0.471	0.507	0.550	0.579
28		0.128	0.241	0.306	0.361	0.423	0.463	0.499	0.541	0.570
29		0.126	0.237	0.301	0.355	0.416	0.456	0.491	0.533	0.562
30		0.124	0.233	0.296	0.349	0.409	0.449	0.484	0.526	0.554
31		0.122	0.229	0.291	0.344	0.403	0.442	0.477	0.518	0.546
32		0.120	0.225	0.287	0.339	0.397	0.436	0.470	0.511	0.539
33		0.118	0.222	0.283	0.334	0.392	0.430	0.464	0.504	0.532
34		0.116	0.219	0.279	0.329	0.386	0.424	0.458	0.498	0.525
35		0.115	0.216	0.275	0.325	0.381	0.418	0.452	0.492	0.519
36		0.113	0.213	0.271	0.320	0.376	0.413	0.446	0.486	0.513
37		0.111	0.210	0.267	0.316	0.371	0.408	0.441	0.480	0.507
38		0.110	0.207	0.264	0.312	0.367	0.403	0.435	0.474	0.501
39		0.108	0.204	0.261	0.308	0.362	0.398	0.430	0.469	0.495
40		0.107	0.202	0.257	0.304	0.358	0.393	0.425	0.463	0.490
41		0.106	0.199	0.254	0.301	0.354	0.389	0.420	0.458	0.484
42		0.104	0.197	0.251	0.297	0.350	0.384	0.416	0.453	0.479
43		0.103	0.195	0.248	0.294	0.346	0.380	0.411	0.449	0.474
44		0.102	0.192	0.246	0.291	0.342	0.376	0.407	0.444	0.469
45		0.101	0.190	0.243	0.288	0.338	0.372	0.403	0.439	0.465
46		0.100	0.188	0.240	0.285	0.335	0.368	0.399	0.435	0.460
47		0.099	0.186	0.238	0.282	0.331	0.365	0.395	0.431	0.456
48		0.098	0.184	0.235	0.279	0.328	0.361	0.391	0.427	0.451
49		0.097	0.182	0.233	0.276	0.325	0.358	0.387	0.423	0.447
50		0.096	0.181	0.231	0.273	0.322	0.354	0.384	0.419	0.443



附表 12  $r_s$  界值表

$n$	概率, $P$									
	单侧:	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	双侧:	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
4		0.600	1.000	1.000						
5		0.500	0.800	0.900	1.000	1.000				
6		0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7		0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8		0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9		0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10		0.248	0.455	0.564	0.648	0.745	0.794	0.830	0.879	0.903
11		0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12		0.217	0.406	0.503	0.587	0.678	0.727	0.769	0.818	0.846
13		0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.791	0.824
14		0.200	0.367	0.464	0.538	0.626	0.679	0.723	0.771	0.802
15		0.189	0.354	0.446	0.521	0.604	0.650	0.700	0.750	0.779
16		0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.729	0.762
17		0.176	0.328	0.414	0.503	0.582	0.635	0.679	0.729	0.762
18		0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
19		0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
20		0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.662	0.696
21		0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22		0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23		0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24		0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25		0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26		0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27		0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28		0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.567	0.598
29		0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30		0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580
31		0.126	0.236	0.301	0.356	0.418	0.459	0.496	0.541	0.571
32		0.124	0.232	0.296	0.350	0.412	0.452	0.489	0.533	0.563
33		0.121	0.229	0.291	0.345	0.405	0.446	0.482	0.525	0.554
34		0.120	0.225	0.287	0.340	0.399	0.439	0.475	0.517	0.547
35		0.118	0.222	0.283	0.335	0.394	0.433	0.468	0.510	0.539
36		0.116	0.219	0.279	0.330	0.388	0.427	0.426	0.504	0.533
37		0.114	0.216	0.275	0.325	0.382	0.421	0.456	0.497	0.526
38		0.113	0.212	0.271	0.321	0.378	0.415	0.450	0.491	0.519
39		0.111	0.210	0.267	0.317	0.373	0.410	0.444	0.485	0.513
40		0.110	0.207	0.264	0.313	0.368	0.405	0.439	0.479	0.507
41		0.108	0.204	0.261	0.309	0.364	0.400	0.433	0.473	0.501
42		0.107	0.202	0.257	0.305	0.359	0.395	0.428	0.468	0.495
43		0.105	0.199	0.254	0.301	0.355	0.391	0.423	0.463	0.490
44		0.104	0.197	0.251	0.298	0.351	0.386	0.419	0.458	0.484
45		0.103	0.194	0.248	0.294	0.347	0.382	0.414	0.453	0.479
46		0.102	0.192	0.246	0.291	0.343	0.378	0.410	0.448	0.474
47		0.101	0.190	0.243	0.288	0.340	0.374	0.405	0.443	0.469
48		0.100	0.188	0.240	0.285	0.336	0.370	0.401	0.439	0.465
49		0.098	0.186	0.238	0.282	0.333	0.366	0.397	0.434	0.460
50		0.097	0.184	0.235	0.279	0.329	0.363	0.393	0.430	0.456



附表 13 随机数字表

88	56	53	27	59	33	35	72	67	47	77	34	55	45	70	08	18	27	38	90	16	95	86	70	75
09	72	95	84	29	49	41	31	06	70	42	38	06	45	18	64	84	73	31	65	52	53	37	97	15
12	96	88	17	31	65	19	69	02	83	60	75	86	90	68	24	64	19	35	51	56	61	87	39	12
85	94	57	24	16	92	09	84	38	76	22	00	27	69	85	29	81	94	78	10	21	94	47	90	12
38	64	43	59	93	98	77	87	68	07	91	51	67	62	44	40	98	05	93	78	23	32	65	41	18
53	44	09	42	72	00	41	86	79	79	68	47	22	00	20	35	55	31	51	51	00	83	63	22	55
40	76	66	26	84	57	99	99	90	37	36	63	32	08	58	37	40	13	68	97	87	64	81	07	83
02	17	79	18	05	12	59	52	57	02	22	07	90	47	03	28	14	11	30	79	20	69	22	40	98
95	17	82	06	53	31	51	10	96	46	92	06	88	07	77	56	11	50	81	69	40	23	72	51	39
35	76	22	42	92	96	11	83	44	80	34	68	35	48	77	33	42	40	90	60	73	96	53	97	86
26	29	13	56	41	85	47	04	66	08	34	72	57	59	13	82	43	80	46	15	38	26	61	70	04
77	80	20	75	82	72	82	32	99	90	63	95	73	76	63	89	73	44	99	05	48	67	26	43	18
46	40	66	44	52	91	36	74	43	53	30	82	13	54	00	78	45	63	98	35	55	03	36	67	68
37	56	08	18	09	77	53	84	46	47	31	91	18	95	58	24	16	74	11	53	44	10	13	85	57
61	65	61	68	66	37	27	47	39	19	84	83	70	07	48	53	21	40	06	71	95	06	79	88	54
93	43	69	64	07	34	18	04	52	35	56	27	09	24	86	61	85	53	83	45	19	90	70	99	00
21	96	60	12	99	11	20	99	45	18	48	13	93	55	34	18	37	79	49	90	65	97	38	20	46
95	20	47	97	97	27	37	83	28	71	00	06	41	41	74	45	89	09	39	84	51	67	11	52	49
97	86	21	78	73	10	65	81	92	59	58	76	17	14	97	04	76	62	16	17	17	95	70	45	80
62	92	06	34	13	59	71	74	17	32	27	55	10	24	19	23	71	82	13	74	63	52	52	01	41
04	31	17	21	56	33	73	99	19	87	26	72	39	27	67	53	77	57	68	93	60	61	97	22	61
61	06	98	03	91	87	14	77	43	96	43	00	65	98	50	45	60	33	01	07	98	99	46	50	47
85	93	85	86	88	72	87	08	62	40	16	06	10	89	20	23	21	34	74	97	76	38	03	29	63
21	74	32	47	45	73	96	07	94	52	09	65	90	77	47	25	76	16	19	33	53	05	70	53	30
15	69	53	82	88	79	96	23	53	10	65	39	07	16	29	45	33	02	43	70	02	87	40	41	45
02	89	08	04	49	20	21	14	68	86	87	63	93	95	17	11	29	01	95	80	35	14	97	35	33
87	18	15	89	79	85	43	01	72	73	08	61	74	51	69	89	74	39	82	15	94	51	33	41	67
98	83	71	94	22	59	97	50	99	52	08	52	85	08	40	87	80	61	65	31	91	51	80	32	44
10	08	58	21	66	72	68	49	29	31	89	85	84	46	06	59	73	19	85	23	65	09	29	75	63
47	90	56	10	08	88	02	84	27	83	42	29	72	23	19	66	56	45	65	79	20	71	53	20	25
22	85	61	68	90	49	64	92	85	44	16	40	12	89	88	50	14	49	81	06	01	82	77	45	12
67	80	43	79	33	12	83	11	41	16	25	58	19	68	70	77	02	54	00	52	53	43	37	15	26
27	62	50	96	72	79	44	61	40	15	14	53	40	65	39	27	31	58	50	28	11	39	03	34	25
33	78	80	87	15	38	30	06	38	21	14	47	47	07	26	54	96	87	53	32	40	36	40	69	76
13	13	92	66	99	47	24	49	57	74	32	25	43	62	17	10	97	11	69	84	99	63	22	32	98
10	27	53	96	23	71	50	54	36	23	54	31	04	82	98	04	14	12	15	09	26	78	25	47	47
28	41	50	61	88	64	85	27	20	18	83	36	36	05	56	39	71	65	09	62	94	76	62	11	89
34	21	42	57	02	59	19	18	97	48	80	30	03	30	98	05	24	67	70	07	84	97	50	87	46
61	81	77	23	23	82	82	11	54	08	53	28	70	58	96	44	07	39	55	43	42	34	43	39	28
61	15	18	13	54	16	86	20	26	88	90	74	80	55	09	14	53	90	51	17	52	01	63	01	59
91	76	21	64	64	44	91	13	32	97	75	31	62	66	54	84	80	32	75	77	56	08	25	70	29
00	97	79	08	06	37	30	28	59	85	53	56	68	53	40	01	74	39	59	73	30	19	99	85	48
36	46	18	34	94	75	20	80	27	77	78	91	69	16	00	08	43	18	73	68	67	69	61	34	25
88	98	99	60	50	65	95	79	42	94	93	62	40	89	96	43	56	47	71	66	46	76	29	67	02
04	37	59	87	21	05	02	03	24	17	47	97	81	56	51	92	34	86	01	82	55	51	33	12	91
63	62	06	34	41	94	21	78	55	09	72	76	45	16	94	29	95	81	83	83	79	88	01	97	30
78	47	23	53	90	34	41	92	45	71	09	23	70	07	12	12	38	92	79	43	14	85	11	47	23
87	68	62	15	43	53	14	36	59	25	54	47	33	70	15	59	24	48	40	35	50	03	42	99	36
47	60	92	10	77	88	59	53	11	52	66	25	69	07	64	48	68	64	71	06	61	65	70	22	12
56	88	87	59	41	65	28	04	67	53	95	79	88	37	31	50	41	06	94	76	81	83	17	16	33
02	57	45	86	67	73	43	07	34	48	44	26	87	93	29	77	09	61	67	84	06	69	44	77	75
31	54	14	13	17	48	62	11	90	60	68	12	93	64	28	46	24	79	16	76	14	60	25	51	01
28	50	16	43	36	28	97	85	58	99	67	22	52	76	23	24	70	36	54	54	59	28	61	71	96
63	29	62	66	50	02	63	45	52	38	67	63	47	54	75	83	24	78	43	20	92	63	13	47	48
45	65	58	26	51	76	96	59	38	72	86	57	45	71	46	44	67	76	14	55	44	88	01	62	12
39	65	36	63	70	77	45	85	50	51	74	13	39	35	22	30	53	36	02	95	49	34	88	73	61
73	71	98	16	04	29	18	94	51	23	76	51	94	84	86	79	93	96	38	63	08	58	25	58	94
72	20	56	20	11	72	65	71	08	86	79	57	95	13	91	97	48	72	66	48	09	71	17	24	89
75	17	26	99	76	89	37	20	70	01	77	31	61	95	46	26	97	05	73	51	53	33	18	72	87
37	48	60	82	29	81	30	15	39	14	48	38	75	93	29	06	87	37	78	48	45	56	00	84	47

## 中英名词索引

I 型错误	type I error	42
II 型错误	type II error	43
chi square 检验	chi square test	67
Fisher 确切概率	Fisher's exact test	69
q-检验	Newman-Keuls test	51
SPSS	statistics package for social science	133
t 分布	t-distribution	29
Youden 指数	Youden index	129
<b>A</b>		
安慰剂对照	placebo control	126
<b>B</b>		
百分比条图	percentage chart	118
百分位数	percentile	13
半对数线图	semi-logarithmic linear chart	118
备择假设	alternative hypothesis	31
比数积	odds product	129
变量的标签	Variable Labels	137
变量名	Variable Name	136
变量值标签	Value Labels	137
变异	variance	3
变异系数	coefficient of variation	19
标准差	standard deviation	18
标准对照	standardized control	126
标准化	standardization	60
标准化偏回归系数	standardized partial regression coefficient	106
标准误	standard error	27
标准正态分布	standard normal distribution	21
病死率	cause fatality	63
<b>C</b>		
参数估计	parameter estimation	28
抽样	sampling	3
抽样误差	sampling error	6, 27
重复原则	principle of replication	126
<b>D</b>		
代表性	representation	4



单独效应	simple effect	53
单样本 t 检验	one sample t test	35
单因素方差分析	one-way ANOVA	46, 152
等级相关系数	rank correlation coefficient	98
点估计	point estimation	28
对照原则	principle of control	126
多元线性回归	multiple linear regression	102
多重比较	multiple comparison	51

## F

发病率	incidence rate	63
方差	variance	17
方差分析	analysis of variance, ANOVA	46
方差齐性	homogeneity of variance	37
非参数检验方法	nonparametric test	79
分层抽样	stratified sampling	4
分层随机抽样	stratified random sampling	4
分析资料	analyzing data	2
复相关系数	multiple correlation coefficient	106

## G

概率	probability	6
构成比	constituent ratio	58

## H

合并方差	pooled variance	37
患病率	prevalence rate	63
回归系数	regression coefficient	94
回归直线	regression line	95

## J

极差	range	16
几何均数	geometric mean	11
计量资料	measurement data	2
计数资料	enumeration data	2
假设检验	hypothesis test	30
交互效用	interaction effect	53
金标准	gold standard	127
精度	precision	2
决定系数	coefficient of determination	107
均衡原则	principle of balance	127
均数	mean	10



## K

可比性	comparability	4
可靠性	reliability	4
可信区间	confidence interval	28
可信限	confidence limit	28
空白对照	blank control	126

## L

离均差平方和	sum of square	17
理论频数	theoretical frequency	67
连续性校正法	correction for continuity	69
漏诊率	omission diagnostic rate	128
率	rate	58

## M

盲法试验	blind trial	127
敏感度	sensitivity	128

## N

年龄别死亡率	age-specific death rate	62
--------	-------------------------	----

## P

配对资料的符号秩和检验	Wilcoxon signed rank test	79
配对设计	paired design	5
配对样本	paired sample	5
偏回归系数	partial regression coefficient	102
偏性	bias	2
频数表	frequency table	8
频数分布表	frequency distribution table	2
平均偏差	mean difference	17
平均数	average	10

## Q

区间估计	interval estimation	28
------	---------------------	----

## S

散点图	scatter diagram	90
实际频数	actual frequency	67
实验设计	experiment design	2
收集资料	collecting data	2
数据	Data	135
数据编码	Data Code	136



死亡率	death rate	62
死因别死亡率	cause specific death rate	62
四分位数间距	quartile	16
随机测量误差	random measurement error	6
随机化原则	principle of random	126
随机区组设计	randomized block design	5, 48

**T**

特异度	specificity	128
统计表	statistical table	113
统计图	statistical chart	113
统计推断	statistical inference	27

**W**

完全随机设计	completely random design	5
无效假设	null hypothesis	31
误差	error	6
误诊率	mistake diagnostic rate	128

**X**

析因设计	factorial design	52
系统误差	system error	6
显著性检验	significant test	30
显著性水准	significant level	32
线图	line chart	118
相对比	relative ratio	59
相对数	relative number	58
相对危险度	relative risk, RR	59
相关系数	coefficient of correlation	91
箱图	box plot	120

**Y**

阳性预测值	positive predictive value	129
样本	sample	3
样本含量	sample size	3
医学实验设计	medical experimental design	124
阴性预测值	negative predictive value	129
因变量	dependent variable	93
预测值	predictive value	129
原始资料	raw data	2
圆图	pie chart	117

**Z**

诊断试验评价	evaluation of diagnostic testing	127
--------	----------------------------------	-----



整理资料	sorting data	2
正态分布	normal distribution	10
直方图	histogram	9, 120
直条图	bar chart	116
治愈率	cure rate	63
中位数	median	12
主效应	main effect	53
自变量	independent variable	93
自身对照	self control	126
自由度	degree of freedom	17
总符合率	total consistent rate	128
总体	population	3
最小二乘法	method of least square	94, 102

---

# 英中名词索引

## A

actual frequency	实际频数	67
age-specific death rate	年龄别死亡率	62
alternative hypothesis	备择假设	31
analysis of variance, ANOVA	方差分析	46
analyzing data	分析资料	2
average	平均数	10

## B

bar chart	直条图	116
bias	偏性	2
blank control	空白对照	126
blind trial	盲法试验	127
box plot	箱图	120

## C

cause fatality	病死率	63
cause specific death rate	死因别死亡率	62
chi square test	chi square 检验	67
coefficient of correlation	相关系数	91
coefficient of determination	决定系数	107
coefficient of variation	变异系数	19
collecting data	收集资料	2
comparability	可比性	4
completely random design	完全随机设计	5
confidence interval	可信区间	28
confidence limit	可信限	28
constituent ratio	构成比	58
correction for continuity	连续性校正法	69
cure rate	治愈率	63

## D

Data	数据	135
Data Code	数据编码	136
death rate	死亡率	62
degree of freedom	自由度	17
dependent variable	因变量	93





## E

enumeration data	计数资料	2
error	误差	6
evaluation of diagnostic testing	诊断试验评价	127
experiment design	实验设计	2

## F

factorial design	析因设计	52
Fisher's exact test	Fisher 确切概率	70
frequency distribution table	频数分布表	2
frequency table	频数表	8

## G

geometric mean	几何均数	11
gold standard	金标准	127

## H

histogram	直方图	9, 120
homogeneity of variance	方差齐性	37
hypothesis test	假设检验	30

## I

incidence rate	发病率	63
independent variable	自变量	93
interaction effect	交互效用	53
interval estimation	区间估计	28

## L

line chart	线图	118
------------	----	-----

## M

main effect	主效应	53
mean	均数	10
mean difference	平均偏差	17
measurement data	计量资料	2
median	中位数	12
medical experimental design	医学实验设计	124
method of least square	最小二乘法	94, 102
mistake diagnostic rate	误诊率	128
multiple comparison	多重比较	51
multiple correlation coefficient	复相关系数	106
multiple linear regression	多元线性回归	102



negative predictive value	阴性预测值	129
<b>N</b>		
negative predictive value	阴性预测值	129
Newman-Keuls test	q-检验	51
nonparametric test	非参数检验方法	79
normal distribution	正态分布	10
null hypothesis	无效假设	31
<b>O</b>		
odds product	比数积	129
omission diagnostic rate	漏诊率	128
one sample t test	单样本 t 检验	35
one-way ANOVA	单因素方差分析	46, 152
<b>P</b>		
paired design	配对设计	5
paired sample	配对样本	5
parameter estimation	参数估计	28
partial regression coefficient	偏回归系数	102
percentage chart	百分比条图	118
percentile	百分位数	13
pie chart	圆图	117
placebo control	安慰剂对照	126
point estimation	点估计	28
pooled variance	合并方差	37
population	总体	3
positive predictive value	阳性预测值	129
precision	精度	2
predictive value	预测值	129
prevalence rate	患病率	63
principle of balance	均衡原则	127
principle of control	对照原则	126
principle of random	随机化原则	126
principle of replication	重复原则	126
probability	概率	6
<b>Q</b>		
quartile	四分位数间距	16
<b>R</b>		
random measurement error	随机测量误差	6
randomized block design	随机区组设计	5, 48



range	极差	16
rank correlation coefficient	等级相关系数	98
rate	率	58
raw data	原始资料	2
regression coefficient	回归系数	94
regression line	回归直线	95
relative number	相对数	58
relative ratio	相对比	59
relative risk, RR	相对危险度	59
reliability	可靠性	4
representation	代表性	4

### S

sample	样本	3
sample size	样本含量	3
sampling	抽样	3
sampling error	抽样误差	6, 27
scatter diagram	散点图	90
self control	自身对照	126
semi-logarithmic linear chart	半对数线图	118
sensitivity	敏感度	128
significant level	显著性水准	32
significant test	显著性检验	30
simple effect	单独效应	53
sorting data	整理资料	2
specificity	特异度	128
standard deviation	标准差	18
standard error	标准误	27
standard normal distribution	标准正态分布	21
standardization	标准化	60
standardized control	标准对照	126
standardized partial regression coefficient	标准化偏回归系数	106
statistical chart	统计图	113
statistical inference	统计推断	27
statistics package for social science	SPSS	133
statistical table	统计表	113
stratified random sampling	分层随机抽样	4
stratified sampling	分层抽样	4
sum of square	离均差平方和	17
system error	系统误差	6

### T

t-distribution	t 分布	29
----------------	------	----



theoretical frequency	理论频数	67
total consistent rate	总符合率	128
type I error	I 型错误	42
type II error	II 型错误	43
<b>V</b>		
Value Labels	变量值标签	137
Variable Name	变量名	136
Variable Labels	变量的标签	137
variance	方差	17
variance	变异	3
<b>W</b>		
Wilcoxon signed rank test	配对资料的符号秩和检验	79
<b>Y</b>		
Youden index	Youden 指数	129

## 第一章 医学统计中的基本概念

### 练习题

#### 一、单项选择题

1. 医学统计学研究的对象是  
A. 医学中的小概率事件  
B. 各种类型的数据  
C. 动物和人的本质  
D. 疾病的预防与治疗  
E. 有变异的医学事件
2. 用样本推论总体，具有代表性的样本指的是  
A. 总体中最容易获得的部分个体  
B. 在总体中随意抽取任意个体  
C. 挑选总体中的有代表性的部分个体  
D. 用配对方法抽取的部分个体  
E. 依照随机原则抽取总体中的部分个体
3. 下列观测结果属于等级资料的是  
A. 收缩压测量值  
B. 脉搏数  
C. 住院天数  
D. 病情程度  
E. 四种血型
4. 随机误差指的是  
A. 测量不准引起的误差  
B. 由操作失误引起的误差  
C. 选择样本不当引起的误差  
D. 选择总体不当引起的误差  
E. 由偶然因素引起的误差
5. 收集资料不可避免的误差是  
A. 随机误差  
B. 系统误差  
C. 过失误差  
D. 记录误差  
E. 仪器故障误差

答案：E E D E A

#### 二、简答题

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？

#### [参考答案]

常见的三类误差是：

(1) 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。要尽量查明其原因，必须克服。

(2) 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，实验操作员操作技术不稳定，不同实验操作员之间的操作差异，电压不稳及环境温度差异等因素造成测量结果的误差。对于这种误差应采取相应的措施加以控制，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。

(3) 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（或其他参数）之间仍可能有差异。这种差异是由抽样引起的，故这种误差叫做抽样误差，要用统计方法进行正确分析。

2. 抽样中要求每一个样本应该具有哪三性？

[参考答案]

从总体中抽取样本，其样本应具有“代表性”、“随机性”和“可靠性”。

(1) 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。

(2) 随机性：就是要保证总体中的每个个体均有相同的几率被抽作样本。

(3) 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。每个样本的含量越多，可靠性会越大，但是例数增加，人力、物力都会发生困难，所以应以“足够”为准。需要作“样本例数估计”。

3. 什么是两个样本之间的可比性？

[参考答案]

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。

(马斌荣)

## 第二章 集中趋势的统计描述

### 练习题

#### 一、单项选择题

- 某医学资料数据大的一端没有确定数值，描述其集中趋势适用的统计指标是
  - 中位数
  - 几何均数
  - 均数
  - $P_{95}$ 百分位数
  - 频数分布
- 算术均数与中位数相比，其特点是
  - 不易受极端值的影响
  - 能充分利用数据的信息
  - 抽样误差较大
  - 更适用于偏态分布资料
  - 更适用于分布不明确资料
- 一组原始数据呈正偏态分布，其数据的特点是
  - 数值离散度较小
  - 数值离散度较大
  - 数值分布偏向较大一侧
  - 数值分布偏向较小一侧
  - 数值分布不均匀
- 将一组计量资料整理成频数表的主要目的是
  - 化为计数资料
  - 便于计算
  - 形象描述数据的特点
  - 为了能够更精确地检验
  - 提供数据和描述数据的分布特征
- 6人接种流感疫苗一个月后测定抗体滴度为 1:20、1:40、1:80、1:80、1:160、1:320，求平均滴度应选用的指标是



- A. 均数
- B. 几何均数
- C. 中位数
- D. 百分位数
- E. 倒数的均数

答案: A B D E B

## 二、计算与分析

1. 现测得 10 名乳腺癌患者化疗后血液尿素氮的含量 (mmol/L) 分别为 3.43, 2.96, 4.43, 3.03, 4.53, 5.25, 5.64, 3.82, 4.28, 5.25, 试计算其均数和中位数。

[参考答案]

$$\bar{X} = \frac{3.43 + 2.96 + 4.43 + 3.03 + 4.53 + 5.25 + 5.64 + 3.82 + 4.28 + 5.25}{10} = 4.26(\text{mmol/L})$$

$$M = \frac{4.28 + 4.43}{2} = 4.355(\text{mmol/L})$$

2. 某地 100 例 30~40 岁健康男子血清总胆固醇值 (mg/dl) 测定结果如下:

202 165 199 234 200 213 155 168 189 170 188 168 184 147 219 174 130 183 178 174  
228 156 171 199 185 195 230 232 191 210 195 165 178 172 124 150 211 177 184 149  
159 149 160 142 210 142 185 146 223 176 241 164 197 174 172 189 174 173 205 224  
221 184 177 161 192 181 175 178 172 136 222 113 161 131 170 138 248 153 165 182  
234 161 169 221 147 209 207 164 147 210 182 183 206 209 201 149 174 253 252 156

- (1) 编制频数分布表并画出直方图;
- (2) 根据频数表计算均值和中位数, 并说明用哪一个指标比较合适;
- (3) 计算百分位数  $P_5$ 、 $P_{25}$ 、 $P_{75}$  和  $P_{95}$ 。

[参考答案]

- (1) 编制频数分布表:

某地 100 例 30~40 岁健康男子血清总胆固醇值的频数表

甘油三酯 (mg/dL)	频数	累积频数	累积频率
(1)	(2)	(3)	(4)
110~	2	2	2
125~	4	6	6
140~	11	17	17
155~	16	33	33
170~	27	60	60
185~	12	72	72
200~	13	85	85
215~	7	92	92
230~	5	97	97
245~	3	100	100
合计	100	—	—

画直方图:

- (2) 计算均数和中位数:

$$\bar{X} = \frac{(110 + 7.5) \times 2 + (125 + 7.5) \times 4 + \dots + (245 + 7.5) \times 3}{100} = 182.9(\text{mg/dl})$$

$$M = P_{50} = 170 + \frac{100 \times 0.5 - 33}{27} \times 15 = 179.4(\text{mg/dl})$$

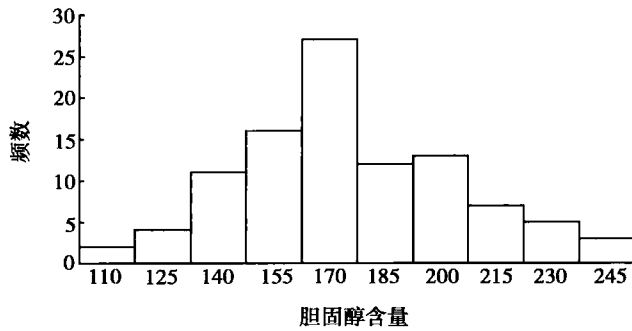


图 某地 100 例 30~40 岁健康男子血清总胆固醇值的频数分布

从上述直方图能够看出：此计量指标近似服从正态分布，选用均数较为合适。

(3) 计算百分位数：

$$P_5 = 125 + \frac{100 \times 0.05 - 2}{4} \times 15 = 136.25 (\text{mg/dl})$$

$$P_{25} = 155 + \frac{100 \times 0.25 - 17}{16} \times 15 = 162.5 (\text{mg/dl})$$

$$P_{75} = 200 + \frac{100 \times 0.75 - 72}{13} \times 15 = 203.5 (\text{mg/dl})$$

$$P_{95} = 230 + \frac{100 \times 0.95 - 92}{5} \times 15 = 239 (\text{mg/dl})$$

3. 测得 10 名肝癌病人与 16 名正常人的血清乙型肝炎表面抗原 (HBsAg) 滴度如下表，试分别计算它们的平均滴度。

肝癌病人与正常人的血清乙肝表面抗原 (HBsAg) 滴度

滴度倒数	正常人数	肝癌病人数
8	7	1
16	5	2
32	1	3
64	3	2
128	0	1
256	0	1

[ 参考答案 ]

肝癌病人与正常人的血清乙肝表面抗原 (HBsAg) 滴度测定结果

滴度倒数(X)	正常人数(f <sub>1</sub> )	肝癌病人数(f <sub>2</sub> )	lgx	f <sub>1</sub> lgx	f <sub>2</sub> lgx
8	7	1	0.90	6.30	0.90
16	5	2	1.20	6.00	2.40
32	1	3	1.50	1.50	4.50
64	3	2	1.81	5.43	3.62
128	0	1	2.11	0.00	2.11
256	0	1	2.41	0.00	2.41
合计	16	10	—	19.23	15.94

$$G_1 = \lg^{-1} \left( \frac{19.23}{16} \right) \approx 15.92 \quad G_2 = \lg^{-1} \left( \frac{15.94}{10} \right) \approx 39.26$$

正常人乙肝表面抗原 (HBsAg) 滴度为 1 : 15.92





肝癌病人乙肝表面抗原 (HBsAg) 滴度为 1 : 39.26

(李 康)

### 第三章 变异程度的统计描述

#### 练 习 题

##### 一、单项选择题

1. 变异系数主要用于
  - A. 比较不同计量指标的变异程度
  - B. 衡量正态分布的变异程度
  - C. 衡量测量的准确度
  - D. 衡量偏态分布的变异程度
  - E. 衡量样本抽样误差的大小
2. 对于近似正态分布的资料, 描述其变异程度应选用的指标是
  - A. 变异系数
  - B. 离均差平方和
  - C. 极差
  - D. 四分位数间距
  - E. 标准差
3. 某项指标 95% 医学参考值范围表示的是
  - A. 检测指标在此范围, 判断“异常”正确的概率大于或等于 95%
  - B. 检测指标在此范围, 判断“正常”正确的概率大于或等于 95%
  - C. 在“异常”总体中有 95% 的人在此范围之外
  - D. 在“正常”总体中有 95% 的人在此范围
  - E. 检测指标若超出此范围, 则有 95% 的把握说明诊断对象为“异常”
4. 应用百分位数法估计参考值范围的条件是
  - A. 数据服从正态分布
  - B. 数据服从偏态分布
  - C. 有大样本数据
  - D. 数据服从对称分布
  - E. 数据变异不能太大
5. 已知动脉硬化患者载脂蛋白 B 的含量 (mg/dl) 呈明显偏态分布, 描述其个体差异的统计指标应使用
  - A. 全距
  - B. 标准差
  - C. 变异系数
  - D. 方差
  - E. 四分位数间距

答案: A E D B E

##### 二、计算与分析

1. 下表为 10 例垂体催乳素微腺瘤的病人手术前后的血催乳素浓度, 试说明用何种指标比较手术前后数据的变异情况较为合适。

手术前后患者血催乳素浓度 (ng/ml)

例号	血催乳素浓度	
	术前	术后
1	276	41
2	880	110
3	1 600	280



续表

例号	血催乳素浓度	
	术前	术后
4	324	61
5	398	105
6	266	43
7	500	25
8	1 760	300
9	500	215
10	220	92

**[参考答案]**

血催乳素浓度术前均值 = 672.4ng/ml, 术后均值 = 127.2ng/ml。手术前后两组均值相差较大, 故选择变异系数作为比较手术前后数据变异情况比较合适。

术前:  $\bar{X} = 672.4, S = 564.65$

$$CV = \frac{564.65}{672.4} \times 100\% = 83.98\%$$

术后:  $\bar{X} = 127.2, S = 101.27$

$$CV = \frac{101.27}{127.2} \times 100\% = 79.61\%$$

可以看出: 以标准差作为比较两组变异情况的指标, 易夸大手术前血催乳素浓度的变异。

2. 某地 144 例 30~45 岁正常成年男子的血清总胆固醇测量值近似服从均数为 4.95mmol/L, 标准差为 0.85mmol/L 的正态分布。①试估计该地 30~45 岁成年男子血清总胆固醇的 95% 参考值范围; ②血清总胆固醇大于 5.72mmol/L 的正常成年男子约占其总体的百分之多少?

**[参考答案]**

① 正常成年男子的血清总胆固醇测量值近似服从正态分布, 故可按正态分布法处理。又因血清总胆固醇测量值过高或过低均属异常, 所以应计算双侧参考值范围。

下限:  $\bar{X} - 1.96S = 4.95 - 1.96 \times 0.85 = 3.28(\text{mol/L})$

上限:  $\bar{X} + 1.96S = 4.95 + 1.96 \times 0.85 = 6.62(\text{mmol/L})$

即该地区成年男子血清总胆固醇测量值的 95% 参考值范围为 3.28~6.62mmol/L。

② 该地正常成年男子的血清总胆固醇测量值近似服从均数为 4.95mmol/L, 标准差为 0.85mmol/L 的正态分布, 计算 5.72mmol/L 对应的标准正态分布  $u$  值:

$$u = \frac{5.72 - 4.95}{0.85} \approx 0.91$$

问题转化为求  $u$  值大于 0.91 的概率。由于标准正态分布具有对称性, 所以  $u$  值大于 0.91 的概率与  $u$  值小于 -0.91 的概率相同。查附表 1 得,  $\phi(-u) = 0.1814$ , 所以说血清总胆固醇大于 5.72mmol/L 的正常成年男子约占其总体的 18.14%。

3. 某地 200 例正常成人血铅含量的频数分布如下表。

(1) 简述该资料的分布特征。

(2) 若资料近似呈对数正态分布, 试分别用百分位数法和正态分布法估计该地正常成人血铅值的 95% 参考值范围。



某地 200 例正常成人血铅含量 ( $\mu\text{mol/L}$ ) 的频数分布

血铅含量	频数	累积频数
0.00~	7	7
0.24~	49	56
0.48~	45	101
0.72~	32	133
0.96~	28	161
1.20~	13	174
1.44~	14	188
1.68~	4	192
1.92~	4	196
2.16~	1	197
2.40~	2	199
2.64~	1	200

[参考答案]

(1) 从表可以看出, 血铅含量较低组段的频数明显高于较高组段, 分布不对称。同正态分布相比, 其分布高峰向血铅含量较低方向偏移, 长尾向血铅含量较高组段延伸, 数据为正偏态分布。

某地 200 例正常成人血铅含量 ( $\mu\text{mol/L}$ ) 的频数分布

血铅含量	组中值	频数	累积频数	累积频率
0.00~	0.12	7	7	3.5
0.24~	0.36	49	56	28.0
0.48~	0.60	45	101	50.5
0.72~	0.84	32	133	66.5
0.96~	1.08	28	161	80.5
1.20~	1.32	13	174	87.0
1.44~	1.56	14	188	94.0
1.68~	1.80	4	192	96.0
1.92~	2.04	4	196	98.0
2.16~	2.28	1	197	98.5
2.40~	2.52	2	199	99.5
2.64~	2.76	1	200	100

(2) 因为正常人血铅含量越低越好, 所以应计算单侧 95% 参考值范围。

百分位数法: 第 95% 百分位数位于 1.68~ 组段, 组距为 0.24, 频数为 4, 该组段以前的累积频数为 188, 故

$$P_{95} = 1.68 + \frac{(200 \times 0.95 - 188)}{4} \times 0.24 = 1.80 (\mu\text{mol/L})$$

即该地正常成人血铅值的 95% 参考值范围为小于 1.80  $\mu\text{mol/L}$ 。

正态分布法: 将组中值进行对数变换, 根据题中表格, 得到均值和标准差计算表。



某地 200 例正常成人血铅含量 ( $\mu\text{mol/L}$ ) 均值和标准差计算表

血铅含量	组中值	$\lg$ 组中值 $\lg(x)$	频数 ( $f$ )	$fx$	$fx^2$
0.00~	0.12	-0.92	7	-6.44	5.9248
0.24~	0.36	-0.44	49	-21.56	9.4864
0.48~	0.60	-0.22	45	-9.9	2.178
0.72~	0.84	-0.08	32	-2.56	0.2048
0.96~	1.08	0.03	28	0.84	0.0252
1.20~	1.32	0.12	13	1.56	0.1872
1.44~	1.56	0.19	14	2.66	0.5054
1.68~	1.80	0.26	4	1.04	0.2704
1.92~	2.04	0.31	4	1.24	0.3844
2.16~	2.28	0.36	1	0.36	0.1296
2.40~	2.52	0.40	2	0.80	0.3200
2.64~	2.76	0.44	1	0.44	0.1936
合计	—	—	200	-31.52	19.8098

计算均值和标准差:

$$\bar{X} = \frac{-31.52}{200} = -0.1576$$

$$S = \sqrt{\frac{19.8098 - (-31.52)^2/200}{200-1}} = 0.2731$$

单侧 95% 参考值范围:

$$\bar{X} + 1.65S = -0.1576 + 1.65 \times 0.2731 = 0.2930$$

$$\lg^{-1}(0.2930) = 1.96 (\mu\text{mol/L})$$

即该地正常成人血铅值的 95% 参考值范围为小于  $1.96 \mu\text{mol/L}$ , 与百分位数法相比两者相差不大。

(李 康)

## 第四章 抽样误差与假设检验

### 练习题

#### 一、单项选择题

1. 样本均数的标准误越小说明

- A. 观察个体的变异越小
- B. 观察个体的变异越大
- C. 抽样误差越大
- D. 由样本均数估计总体均数的可靠性越小
- E. 由样本均数估计总体均数的可靠性越大

2. 抽样误差产生的原因是

- A. 样本不是随机抽取
- B. 测量不准确
- C. 资料不是正态分布
- D. 个体差异
- E. 统计指标选择不当

3. 对于正偏态分布的总体, 当样本含量足够大时, 样本均数的分布近似为

- A. 正偏态分布
- B. 负偏态分布



C. 正态分布

D.  $t$  分布

E. 标准正态分布

4. 假设检验的目的是

A. 检验参数估计的准确度

B. 检验样本统计量是否不同

C. 检验样本统计量与总体参数是否不同

D. 检验总体参数是否不同

E. 检验样本的  $P$  值是否为小概率

5. 根据样本资料算得健康成人白细胞计数的 95% 可信区间为  $7.2 \times 10^9/L \sim 9.1 \times 10^9/L$ , 其含义是

A. 估计总体中有 95% 的观察值在此范围内

B. 总体均数在该区间的概率为 95%

C. 样本中有 95% 的观察值在此范围内

D. 该区间包含样本均数的可能性为 95%

E. 该区间包含总体均数的可能性为 95%

答案: E D C D E

## 二、计算与分析

1. 为了解某地区小学生血红蛋白含量的平均水平, 现随机抽取该地小学生 450 人, 算得其血红蛋白平均数为 101.4g/L, 标准差为 1.5g/L, 试计算该地小学生血红蛋白平均数的 95% 可信区间。

[参考答案]

样本含量为 450, 属于大样本, 可采用正态近似的方法计算可信区间。

$$\bar{X} = 101.4, S = 1.5, n = 450, S_{\bar{X}} = \frac{S}{\sqrt{n}} = \frac{1.5}{\sqrt{450}} = 0.07$$

95% 可信区间为

$$\text{下限: } \bar{X} - u_{\alpha/2} \cdot S_{\bar{X}} = 101.4 - 1.96 \times 0.07 = 101.26 (\text{g/L})$$

$$\text{上限: } \bar{X} + u_{\alpha/2} \cdot S_{\bar{X}} = 101.4 + 1.96 \times 0.07 = 101.54 (\text{g/L})$$

即该地成年男子红细胞总体均数的 95% 可信区间为 101.26g/L ~ 101.54g/L。

2. 研究高胆固醇是否有家庭聚集性, 已知正常儿童的总胆固醇平均水平是 175mg/dl, 现测得 100 名曾患心脏病且胆固醇高的子代儿童的胆固醇平均水平为 207.5mg/dl, 标准差为 30mg/dl。试回答下列问题:

① 如何衡量这 100 名儿童总胆固醇样本平均数的抽样误差?

② 估计 100 名儿童的胆固醇平均水平的 95% 可信区间;

③ 根据可信区间判断高胆固醇是否有家庭聚集性, 并说明理由。

[参考答案]

① 均数的标准误可以用来衡量样本均数的抽样误差大小, 即

$$S = 30 \text{mg/dl}, n = 100$$

$$S_{\bar{X}} = \frac{S}{\sqrt{n}} = \frac{30}{\sqrt{100}} = 3.0$$

② 样本含量为 100, 属于大样本, 可采用正态近似的方法计算可信区间。

$\bar{X} = 207.5, S = 30, n = 100, S_{\bar{X}} = 3$ , 则 95% 可信区间为

$$\text{下限: } \bar{X} - u_{\alpha/2} \cdot S_{\bar{X}} = 207.5 - 1.96 \times 3 = 201.62 (\text{mg/dl})$$

$$\text{上限: } \bar{X} + u_{\alpha/2} \cdot S_{\bar{X}} = 207.5 + 1.96 \times 3 = 213.38 (\text{mg/dl})$$

故该地 100 名儿童的胆固醇平均水平的 95% 可信区间为 201.62mg/dl ~ 213.38mg/dl。

③ 因为 100 名曾患心脏病且胆固醇高的子代儿童的胆固醇平均水平的 95% 可信区间



的下限高于正常儿童的总胆固醇平均水平 175mg/dl, 提示患心脏病且胆固醇高的父辈, 其子代胆固醇水平较高, 即高胆固醇具有一定的家庭聚集性。

(李 康)

## 第五章 t 检 验

### 练 习 题

#### 一、单项选择题

1. 两样本均数比较, 检验结果  $P > 0.05$  说明
  - A. 两总体均数的差别较小
  - B. 两总体均数的差别较大
  - C. 支持两总体无差别的结论
  - D. 不支持两总体有差别的结论
  - E. 可以确认两总体无差别
2. 由两样本均数的差别推断两总体均数的差别, 其差别具有统计学意义是指
  - A. 两样本均数的差别具有实际意义
  - B. 两总体均数的差别具有实际意义
  - C. 两样本和两总体均数的差别都具有实际意义
  - D. 有理由认为两样本均数有差别
  - E. 有理由认为两总体均数有差别
3. 两样本均数比较, 差别具有统计学意义时,  $P$  值越小说明
  - A. 两样本均数差别越大
  - B. 两总体均数差别越大
  - C. 越有理由认为两样本均数不同
  - D. 越有理由认为两总体均数不同
  - E. 越有理由认为两样本均数相同
4. 减少假设检验的 II 类误差, 应该使用的方法是
  - A. 减少 I 类错误
  - B. 减少测量的系统误差
  - C. 减少测量的随机误差
  - D. 提高检验界值
  - E. 增加样本含量
5. 两样本均数比较的  $t$  检验和  $u$  检验的主要差别是
  - A.  $t$  检验只能用于小样本资料
  - B.  $u$  检验要求大样本资料
  - C.  $t$  检验要求数据方差相同
  - D.  $t$  检验的检验效能更高
  - E.  $u$  检验能用于两大样本均数比较

答案: D E D E B

#### 二、计算与分析

1. 已知正常成年男子血红蛋白均值为 140g/L, 今随机调查某厂成年男子 60 人, 测其血红蛋白均值为 125g/L, 标准差 15g/L. 问该厂成年男子血红蛋白均值与一般成年男子是否不同?

[参考答案]

因样本含量  $n > 50$  ( $n = 60$ ), 故采用样本均数与总体均数比较的  $u$  检验。

(1) 建立检验假设, 确定检验水平

$H_0: \mu = \mu_0$ , 该厂成年男子血红蛋白均值与一般成年男子相同

$H_1: \mu \neq \mu_0$ , 该厂成年男子血红蛋白均值与一般成年男子不同

$\alpha = 0.05$



(2) 计算检验统计量

$$u = \frac{\bar{X} - \mu}{\sigma_{\bar{x}}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{140 - 125}{15/\sqrt{60}} = 7.75$$

(3) 确定  $P$  值, 作出推断结论

$7.75 > 1.96$ , 故  $P < 0.05$ , 按  $\alpha = 0.05$  水准, 拒绝  $H_0$ , 接受  $H_1$ , 可以认为该厂成年男子血红蛋白均值与一般成年男子不同, 该厂成年男子血红蛋白均值低于一般成年男子。

2. 某研究者为比较耳垂血和手指血的白细胞数, 调查 12 名成年人, 同时采取耳垂血和手指血见下表, 试比较两者的白细胞数有无不同。

成人耳垂血和手指血白细胞数 (10g/L)

编号	耳垂血	手指血
1	9.7	6.7
2	6.2	5.4
3	7.0	5.7
4	5.3	5.0
5	8.1	7.5
6	9.9	8.3
7	4.7	4.6
8	5.8	4.2
9	7.8	7.5
10	8.6	7.0
11	6.1	5.3
12	9.9	10.3

[参考答案]

本题为配对设计资料, 采用配对  $t$  检验进行分析

(1) 建立检验假设, 确定检验水平

$H_0: \mu_d = 0$ , 成人耳垂血和手指血白细胞数差异为零

$H_1: \mu_d \neq 0$ , 成人耳垂血和手指血白细胞数差异不为零

$\alpha = 0.05$

(2) 计算检验统计量

$$\sum d = 11.6, \sum d^2 = 20.36$$

$$\bar{d} = \sum d/n = 11.6/12 = 0.967$$

$$S_d = \sqrt{\frac{\sum d^2 - (\sum d)^2/n}{n-1}} = \sqrt{\frac{20.36 - \frac{(11.6)^2}{12}}{12-1}} = 0.912$$

$$t = \frac{\bar{d} - \mu_d}{S_{\bar{d}}} = \frac{\bar{d} - 0}{S_{\bar{d}}} = \frac{\bar{d}}{S_d/\sqrt{n}} = t = \frac{\bar{d}}{S_d/\sqrt{n}} = \frac{0.967}{0.912/\sqrt{12}} = 3.672$$

$t = 3.672 > t_{0.05/2, 11}$ ,  $P < 0.05$ , 拒绝  $H_0$ , 接受  $H_1$ , 差别有统计学意义, 可以认为两者的白细胞数不同。

3. 分别测得 15 名健康人和 13 名 III 度肺气肿病人痰中  $\alpha_1$  抗胰蛋白酶含量 (g/L) 如下表, 问健康人与 III 度肺气肿病人  $\alpha_1$  抗胰蛋白酶含量是否不同?



健康人与Ⅲ度肺气肿患者  $\alpha_1$  抗胰蛋白酶含量 (g/L)

健康人	Ⅲ度肺气肿患者	健康人	Ⅲ度肺气肿患者
2.7	3.6	1.9	4.8
2.2	3.4	1.3	5.6
4.1	3.7	1.5	4.1
4.3	5.4	1.7	3.3
2.6	3.6	1.3	4.3
1.9	6.8	1.3	
1.7	4.7	1.9	
0.6	2.9		

**[参考答案]**

由题意得,  $\bar{X}_1=2.067$ ,  $S_1=1.015$ ;  $\bar{X}_2=4.323$ ,  $S_2=1.107$

本题是两个小样本均数比较, 可用成组设计  $t$  检验, 首先检验两总体方差是否相等。

$H_0: \sigma_1^2 = \sigma_2^2$ , 即两总体方差相等

$H_1: \sigma_1^2 \neq \sigma_2^2$ , 即两总体方差不等

$\alpha=0.05$

$$F = \frac{S_2^2}{S_1^2} = \frac{1.107^2}{1.015^2} = 1.19$$

$F_{0.05(12,14)} = 2.53 > 1.19$ ,  $F < F_{0.05(12,14)}$ , 故  $P > 0.05$ , 按  $\alpha=0.05$  水准, 不拒绝  $H_0$ , 差别无统计学意义。故认为健康人与Ⅲ度肺气肿病人  $\alpha_1$  抗胰蛋白酶含量总体方差相等, 可直接用两独立样本均数比较的  $t$  检验。

(1) 建立检验假设, 确定检验水平

$H_0: \mu_1 = \mu_2$ , 健康人与Ⅲ度肺气肿病人  $\alpha_1$  抗胰蛋白酶含量相同

$H_1: \mu_1 \neq \mu_2$ , 健康人与Ⅲ度肺气肿病人  $\alpha_1$  抗胰蛋白酶含量不同

$\alpha=0.05$

(2) 计算检验统计量

$$S_c^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2} = 1.12$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{S_{\bar{X}_1 - \bar{X}_2}} = \frac{|\bar{X}_1 - \bar{X}_2|}{S_{\bar{X}_1 - \bar{X}_2}} = 5.63$$

(3) 确定  $P$  值, 作出推断结论

$t=5.63 > t_{0.001/2, 26}$ ,  $P < 0.001$ , 拒绝  $H_0$ , 接受  $H_1$ , 差别有统计学意义, 可认为健康人与Ⅲ度肺气肿病人  $\alpha_1$  抗胰蛋白酶含量不同。

4. 某地对 241 例正常成年男性面部上颌间隙进行了测定, 得其结果如下表, 问不同身高正常男性其上颌间隙是否不同?

某地 241 名正常男性上颌间隙 (cm)

身高(cm)	例数	均数	标准差
161~	116	0.2189	0.2351
172~	125	0.2280	0.2561

**[参考答案]**

本题属于大样本均数比较, 采用两独立样本均数比较的  $u$  检验。

由上表可知,





$$n_1 = 116, \bar{X}_1 = 0.2189, S_1 = 0.2351$$

$$n_2 = 125, \bar{X}_2 = 0.2280, S_2 = 0.2561$$

(1) 建立检验假设, 确定检验水平

$H_0: \mu_1 = \mu_2$ , 不同身高正常男性其上颌间隙均值相同

$H_1: \mu_1 \neq \mu_2$ , 不同身高正常男性其上颌间隙均值不同

$$\alpha = 0.05$$

(2) 计算检验统计量

$$u = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} = 0.91$$

(3) 确定  $P$  值, 作出推断结论

$u = 0.91 < 1.96$ , 故  $P > 0.05$ , 按  $\alpha = 0.05$  水准, 不拒绝  $H_0$ , 差别无统计学意义, 尚不能认为不同身高正常男性其上颌间隙不同。

5. 将钩端螺旋体病人的血清分别用标准株和水生株作凝溶试验, 测得稀释倍数如下表, 问两组的平均效价有无差别?

钩端螺旋体病患者凝溶试验的稀释倍数

标准株	100	200	400	400	400	400	800	1 600	1 600	1 600	3 200	3 200	3 200
水生株	100	100	100	200	200	200	200	400	400	800	1 600		

[参考答案]

本题采用两独立样本几何均数比较的  $t$  检验。

$t = 2.689 > t_{0.05/2, 22}$ ,  $P < 0.05$ , 拒绝  $H_0$ , 接受  $H_1$ , 差别有统计学意义, 可认为两组的平均效价有差别。

6. 为比较男、女大学生的血清谷胱甘肽过氧化物酶 (GSH-Px) 的活力是否相同, 某医生从某大学 18~22 岁大学生中随机抽查男生 48 名, 女生 46 名, 测定其血清谷胱甘肽过氧化物酶含量 (活力单位), 男、女性的均数分别为 96.53 和 93.73, 男、女性标准差分别为 7.66 和 14.97。问男女性的 GSH-Px 是否相同?

[参考答案]

由题意得  $n_1 = 48$ ,  $\bar{X}_1 = 96.53$ ,  $S_1 = 7.66$

$n_2 = 46$ ,  $\bar{X}_2 = 93.73$ ,  $S_2 = 14.97$

本题是两个小样本均数比较, 可用成组设计  $t$  检验或  $t'$  检验, 首先检验两总体方差是否相等。

$H_0: \mu_1^2 = \mu_2^2$ , 即两总体方差相等

$H_1: \mu_1^2 \neq \mu_2^2$ , 即两总体方差不等

$$\alpha = 0.05$$

$$F = \frac{S_2^2}{S_1^2} = \frac{7.66^2}{14.97^2} = 3.82$$

$F = 3.82 > F_{0.05(47, 45)}$ , 故  $P < 0.05$ , 差别有统计学意义, 按  $\alpha = 0.05$  水准, 拒绝  $H_0$ , 接受  $H_1$ , 故认为男、女大学生的血清谷胱甘肽过氧化物酶的活力总体方差不等, 不能直接用两独立样本均数比较的  $t$  检验, 而应用两独立样本均数比较的  $t'$  检验。

$$t' = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = 1.53, t'_{0.05/2} = 2.009, t' < t'_{0.05/2}, P > 0.05, \text{按 } \alpha = 0.05 \text{ 水准, 不}$$

拒绝  $H_0$ , 差别无统计学意义, 尚不能认为男性与女性的 GSH-Px 有差别。

(沈其君 施 榕)



## 第六章 方差分析

### 练习题

#### 一、单项选择题

- 方差分析的基本思想和要点是
  - 组间均方大于组内均方
  - 组内均方大于组间均方
  - 不同来源的方差必须相等
  - 两方差之比服从 F 分布
  - 总变异及其自由度可按不同来源分解
- 方差分析的应用条件之一是方差齐性, 它是指
  - 各比较组相应的样本方差相等
  - 各比较组相应的总体方差相等
  - 组内方差=组间方差
  - 总方差=各组方差之和
  - 总方差=组内方差+组间方差
- 完全随机设计方差分析中的组间均方反映的是
  - 随机测量误差大小
  - 某因素效应大小
  - 处理因素效应与随机误差综合结果
  - 全部数据的离散度
  - 各组方差的平均水平
- 对于两组资料的比较, 方差分析与  $t$  检验的关系是
  - $t$  检验结果更准确
  - 方差分析结果更准确
  - $t$  检验对数据的要求更为严格
  - 近似等价
  - 完全等价
- 多组均数比较的方差分析, 如果  $P < 0.05$ , 则应该进一步做的是
  - 两均数的  $t$  检验
  - 区组方差分析
  - 方差齐性检验
  - $q$  检验
  - 确定单独效应

答案: E B C E D

#### 二、计算与分析

1. 在评价某药物耐受性及安全性的 I 期临床试验中, 对符合纳入标准的 40 名健康自愿者随机分为 4 组, 每组 10 名, 各组注射剂量分别为 0.5U、1U、2U、3U, 观察 48 小时后部分凝血活酶时间 (s)。试比较任意两两剂量间的部分凝血活酶时间有无差别?

各剂量组 48 小时部分凝血活酶时间 (s)

0.5U	1U	2U	3U
36.8	40.0	32.9	33.0
34.4	35.5	37.9	30.7
34.3	36.7	30.5	35.3
35.7	39.3	31.1	32.3
33.2	40.1	34.7	37.4
31.1	36.8	37.6	39.1
34.3	33.4	40.2	33.5



续表

0.5U	1U	2U	3U
29.8	38.3	38.1	36.6
35.4	38.4	32.4	32.0
31.2	39.8	35.6	33.8

**[参考答案]**

如方差齐同，则采用完全随机设计的方差分析。

经 Bartlett 方差齐性检验， $\chi^2=1.8991$ ， $\nu=3$ 。由于  $\chi^2_{0.05,3}=7.81$ ， $\chi^2 < \chi^2_{0.05,3}$ ，故  $P > 0.05$ ，可认为四组 48 小时部分凝血活酶时间的总体方差齐同，于是采用完全随机设计的方差分析对四个剂量组部分凝血活酶时间进行比较。

(1) 提出检验假设，确定检验水准

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ ，即四个剂量组部分凝血活酶时间的总体均数相同

$H_1: \mu_1、\mu_2、\mu_3、\mu_4$  不全相同，即四个剂量组部分凝血活酶时间的总体均数不全相同  
 $\alpha=0.05$

(2) 计算检验统计量，列于方差分析表中

方差分析表

变异来源	平方和 SS	自由度 $\nu$	均方 MS	F 值
处理组间	101.0860	3	33.6953	4.80
组内(误差)	252.4780	36	7.0133	
总变异	353.5640	39		

(3) 确定 P 值，作出推断结论

分子自由度  $\nu_{TR}=3$ ，分母自由度  $\nu_E=36$ ，查 F 界表（方差分析用）， $F_{0.05(3,36)}=2.87$ 。由于  $F=4.80$ ， $F > F_{0.05(3,36)}$ ，故  $P < 0.05$ ，按照  $\alpha=0.05$  的显著性水准，拒绝  $H_0$ ，接受  $H_1$ ，差别有统计学意义，可认为四个剂量组部分凝血活酶时间的总体均数不全相同，进而需进行均数间多重比较。

本题采用 SNK 法进行多重比较。

(1) 提出检验假设，确定检验水准

$H_0: \mu_A = \mu_B$ ，即任意两组部分凝血活酶时间的总体均数相同

$H_1: \mu_A \neq \mu_B$ ，即任意两组部分凝血活酶时间的总体均数不相同  
 $\alpha=0.05$

(2) 计算检验统计量，用标记字母法标记

多重比较结果 ( $\alpha=0.05$ )

组别	均数	例数	SNK 标记
1U	37.830	10	A
2U	35.100	10	B
3U	34.370	10	B
0.5U	33.620	10	B

(3) 作出推断结论

1U 与 0.5U，1U 与 2U，1U 与 3U 间差别有统计学意义（标记字母不同），可认为



1U 与 0.5U, 1U 与 2U, 1U 与 3U 间部分凝血活酶时间的总体均数不同。

0.5U、2U、3U 组彼此间差别无统计学意义 (均含有字母 B), 可认为这三组部分凝血活酶时间的总体均数相同。

2. 为探讨小剂量地塞米松对急性肺损伤动物模型肺脏的保护作用, 将 36 只二级 SD 大鼠按性别、体重配成 12 个配伍组, 每一配伍组的 3 只大鼠被随机分配到对照组、损伤组与激素组, 实验 24 小时后测量支气管肺泡灌洗液总蛋白水平 (g/L), 结果如下表。问 3 组大鼠的总蛋白水平是否相同?

3 组大鼠总蛋白水平 (g/L)

配伍组	对照组	损伤组	激素组
1	0.36	1.48	0.30
2	0.28	1.42	0.32
3	0.26	1.33	0.29
4	0.25	1.48	0.16
5	0.36	1.26	0.35
6	0.31	1.53	0.43
7	0.33	1.40	0.31
8	0.28	1.30	0.13
9	0.35	1.58	0.33
10	0.41	1.24	0.32
11	0.49	1.47	0.26
12	0.27	1.32	0.26

[参考答案]

本题采用随机区组设计的方差分析。

(1) 提出检验假设, 确定检验水准

$H_{0(A)}: \mu_1 = \mu_2 = \mu_3$ , 即三组大鼠总蛋白水平的总体均值相同

$H_{1(A)}: \mu_1、\mu_2、\mu_3$  不全相同, 即三组大鼠总蛋白水平的总体均值不全相同

$H_{0(B)}: \tau_1 = \tau_2 = \dots = \tau_{12}$ , 即不同配伍组大鼠总蛋白水平的总体均值相同

$H_{1(B)}: \tau_1、\tau_2、\dots、\tau_{12}$  不全相同, 即不同配伍组大鼠总蛋白水平的总体均值不全相同

$\alpha = 0.05$

(2) 计算检验统计量, 列于方差分析表中

方差分析表

变异来源	平方和 SS	自由度 $\nu$	均方 MS	F 值
处理组间	9.5512	2	4.7756	719.80
区组间	0.1138	11	0.0103	1.56
误差	0.1460	22	0.0066	
总变异	9.8109	35		

(3) 确定 P 值, 作出推断结论。

对于处理因素, 分子自由度  $\nu_A = 2$ , 分母自由度  $\nu_E = 22$ , 查 F 界值表 (方差分析用),  $F_{0.05(2,22)} = 3.44$ 。由于  $F = 719.80, F > F_{0.05(2,22)}$ , 故  $P < 0.05$ , 按照  $\alpha = 0.05$  的显著性水



准, 拒绝  $H_{0(A)}$ , 差别有统计学意义, 可认为三组大鼠总蛋白水平的总体均值不全相同。

对于区组因素, 分子自由度  $\nu_B=11$ , 分母自由度  $\nu_E=22$ , 查  $F$  界值表 (方差分析用),  $F_{0.05(11,22)}=2.26$ 。由于  $F=1.56$ ,  $F < F_{0.05(11,22)}$ , 故  $P > 0.05$ , 照  $\alpha=0.05$  的显著性水准, 不拒绝  $H_{0(B)}$ , 差别无统计学意义, 尚不能认为区组因素对大鼠总蛋白水平有影响。

3. 为研究喹啉酸对大鼠急性和亚急性炎症的影响, 将 40 只体重为  $200 \pm 20(g)$  的雄性 Wistar 大鼠建立急性和亚急性炎症动物模型, 然后随机分为 4 组, 每组 10 只, 给予不同的处理, 观察其 WBC 值。4 种处理分别为: 腹腔注射生理盐水后 3 小时处死、腹腔注射生理盐水后 6 小时处死、腹腔注射喹啉酸 (0.35mg/g) 后 3 小时处死, 腹腔注射喹啉酸 (0.35mg/g) 后 6 小时处死。实验结果如下表。问喹啉酸与给药距处死的时间间隔 (简称时间) 对 WBC 值是否有影响?

不同药物与不同时间大鼠 WBC 值 ( $10^3$ )

时间	药 物			
	生理盐水		喹啉酸	
3h	21.3	18.8	15.8	11.0
	21.9	13.5	8.7	12.8
	11.1	22.6	9.4	12.5
	16.3	17.1	5.3	9.3
	17.9	14.6	8.3	11.0
6h	19.0	23.0	13.9	19.0
	25.2	22.8	15.8	15.3
	22.9	17.8	18.3	19.2
	19.8	24.6	13.0	18.2
	22.7	25.3	14.0	17.3

[参考答案]

本题采用  $2 \times 2$  析因设计方差分析。

1. 提出检验假设, 确定检验水准

$H_{0(A)}$ :  $\mu_1 = \mu_2$ , 即 A 因素两个水平组 WBC 值总体均数相等

$H_{1(A)}$ :  $\mu_1 \neq \mu_2$ , 即 A 因素两个水平组 WBC 值总体均数不相等

$H_{0(B)}$ :  $\tau_1 = \tau_2$ , 即 B 因素两个水平组 WBC 值总体均数相等

$H_{1(B)}$ :  $\tau_1 \neq \tau_2$ , 即 B 因素两个水平组 WBC 值总体均数不相等

$H_{0(AB)}$ : A 与 B 无交互效应

$H_{1(AB)}$ : A 与 B 存在交互效应

$\alpha=0.05$

2. 计算检验统计量, 列于方差分析表中。

方差分析表

变异来源	平方和 SS	自由度 $\nu$	均方 MS	F 值
A 因素	423.1502	1	423.1502	48.68
B 因素	291.0603	1	291.0603	33.48
A×B	3.5403	1	3.5403	0.41
误差	312.9470	36	8.6930	
总变异	1030.6978	39		



3. 确定  $P$  值, 作出推断结论。

对于 A 因素,  $\nu_A = 1, \nu_E = 36$ , 查  $F$  界值表 (方差分析用),  $F_{0.05(1,36)} = 4.11$ 。由于  $F_A = 48.68, F_A > F_{0.05(1,36)}$ , 故  $P < 0.05$ , 按照  $\alpha = 0.05$  的显著性水准, 拒绝  $H_{0(A)}$ , 接受  $H_{1(A)}$ , 认为 A 因素 (药物) 两个水平组 WBC 值总体均数不相等。

对于 B 因素,  $\nu_B = 1, \nu_E = 36$ , 查  $F$  界值表 (方差分析用),  $F_{0.05(1,36)} = 4.11$ 。由于  $F_B = 33.48, F_B > F_{0.05(1,36)}$ , 故  $P < 0.05$ , 按照  $\alpha = 0.05$  的显著性水准, 拒绝  $H_{0(B)}$ , 认为 B 因素 (时间) 两个水平组 WBC 值总体均数不相等。

对于 AB 交互作用,  $\nu_{AB} = 1, \nu_E = 36$ , 查  $F$  界值表 (方差分析用),  $F_{0.05(1,36)} = 4.11$ 。由于  $F_{AB} = 0.41, F_{AB} < F_{0.05(1,36)}$ , 故  $P > 0.05$ , 按照  $\alpha = 0.05$  的显著性水准, 不拒绝  $H_{0(AB)}$ , 认为 A (药物) 与 B (时间) 间无交互效应。

(钟晓妮)

## 第七章 相对数及其应用

### 练习题

#### 一、单项选择题

- 如果一种新的治疗方法能够使不能治愈的疾病得到缓解并延长生命, 则应发生的情况是:
  - 该病患病率增加
  - 该病患病率减少
  - 该病的发病率增加
  - 该病的发病率减少
  - 该疾病的死因构成比增加
- 计算乙肝疫苗接种后血清学检查的阳转率, 分母为:
  - 乙肝易感人数
  - 平均人口数
  - 乙肝疫苗接种人数
  - 乙肝患者人数
  - 乙肝疫苗接种后的阳转人数
- 计算标准化死亡率的目的是:
  - 减少死亡率估计的偏倚
  - 减少死亡率估计的抽样误差
  - 便于进行不同地区死亡率的比较
  - 消除各地区内部构成不同的影响
  - 便于进行不同时间死亡率的比较
- 影响总体率估计的抽样误差大小的因素是:
  - 总体率估计的容许误差
  - 样本率估计的容许误差
  - 检验水准和样本含量
  - 检验的把握度和样本含量
  - 总体率和样本含量
- 研究某种新药的降压效果, 对 100 人进行试验, 其显效率的 95% 可信区间为 0.862~0.926, 表示:
  - 样本显效率在 0.862~0.926 之间的概率是 95%
  - 有 95% 的把握说总体显效率在此范围内波动
  - 有 95% 的患者显效率在此范围
  - 样本率估计的抽样误差有 95% 的可能在此范围
  - 该区间包括总体显效率的可能性为 95%

答案: A C D E E



## 二、计算与分析

1. 某工厂在“职工健康状况报告中”写到：“在 946 名工人中，患慢性病的有 274 人，其中女性 219 人，占 80%，男性 55 人，占 20%。所以女性易患慢性病”，你认为是否正确？为什么？

[参考答案]

不正确，因为此百分比是构成比，不是率，要知道男女谁更易患病需知道 946 名工人中的男女比例，然后计算男女患病率。

2. 在“锑剂短程疗法治疗血吸虫病病例的临床分析”一文中，根据下表资料认为“其中 10~岁组死亡率最高，其次为 20~岁组”，问这种说法是否正确？

锑剂治疗血吸虫病不同性别死亡者年龄分布

年龄组	男	女	合计
0~	3	3	6
10~	11	7	18
20~	4	6	10
30~	5	3	8
40~	1	2	3
50~	5	1	6
合计	29	22	51

[参考答案]

不正确，此为构成比替代率来下结论，正确的计算是用各年龄段的死亡人数除各年龄段的调查人数得到死亡率。

3. 某研究根据以下资料说明沙眼 20 岁患病率最高，年龄大的反而患病率下降，你同意吗？说明理由。

某研究资料沙眼病人的年龄分布

年龄组	沙眼人数	构成比(%)
0~	47	4.6
10~	198	19.3
20~	330	32.1
30~	198	19.3
40~	128	12.4
50~	80	7.8
60~	38	3.7
70~	8	0.8
合计	1 027	100.0

[参考答案]

不正确，此为构成比替代率来下结论，正确的计算是用各年龄段的沙眼人数除各年龄段的调查人数得到患病率。

4. 今有两个煤矿的工人尘肺患病率(%)如下表，试比较两个煤矿的工人尘肺总的患病率。



两个煤矿的工人尘肺患病率情况 (%)

工龄(年)	甲 矿			乙 矿		
	检查人数	尘肺人数	患病率	检查人数	尘肺人数	患病率
<6	14 026	120	0.86	992	2	0.20
6~	4 285	168	3.92	1 905	8	0.42
10~	2 542	316	12.43	1 014	117	11.54
合计	20 853	604	2.90	3 911	127	3.25

[参考答案]

两个煤矿的工人尘肺标准化患病率 (%)

工龄(年)	标准构成	甲 矿		乙 矿	
		原患病率	预期患病人数	原患病率	预期患病人数
<6	15 018	0.86	129	0.20	30
6~	6 190	3.92	243	0.42	26
10~	3 556	12.43	442	11.54	410
合计	24 764	814		466	

$$\text{甲矿尘肺患病率} = \frac{814}{24\,764} \times 100\% = 3.29\%$$

$$\text{乙矿尘肺患病率} = \frac{466}{24\,764} \times 100\% = 1.88\%$$

甲矿尘肺患病率高于乙矿尘肺患病率。

5. 抽样调查了某校 10 岁儿童 200 名的牙齿, 患龋 130 人, 试求该校儿童患龋率的 95% 的区间估计。

[参考答案]

$$P = \frac{130}{200} \times 100\% = 65\%$$

$$S_p = \sqrt{\frac{P(1-P)}{n}} = \sqrt{\frac{0.65 \times 0.35}{200}} = 0.0337 = 3.37\%$$

$$(P - u_{0.05} S_p, P + u_{0.05} S_p) = (65\% - 1.96 \times 3.37\%, 65\% + 1.96 \times 3.37\%) \\ = (58.39\%, 71.61\%)$$

(刘启贵)

## 第八章 $\chi^2$ 检 验

### 练 习 题

#### 一、单项选择题

1. 利用  $\chi^2$  检验公式不适合解决的实际问题是

- A. 比较两种药物的有效率
- B. 检验某种疾病与基因多态性的关系
- C. 两组有序试验结果的药物疗效
- D. 药物三种不同剂量显效率有无差别
- E. 两组病情“轻、中、重”的构成比例





2. 欲比较两组阳性反应率，在样本量非常小的情况下（如  $n_1 < 10, n_2 < 10$ ），应采用
  - A. 四格表  $\chi^2$  检验
  - B. 校正四格表  $\chi^2$  检验
  - C. Fisher 确切概率法
  - D. 配对  $\chi^2$  检验
  - E. 校正配对  $\chi^2$  检验
3. 进行四组样本率比较的  $\chi^2$  检验，如  $\chi^2 > \chi_{0.01,3}^2$ ，可认为
  - A. 四组样本率均不相同
  - B. 四组总体率均不相同
  - C. 四组样本率相差较大
  - D. 至少有两组样本率不相同
  - E. 至少有两组总体率不相同
4. 从甲、乙两文中，查到同类研究的两个率比较的  $\chi^2$  检验，甲文  $\chi^2 > \chi_{0.01,1}^2$ ，乙文  $\chi^2 > \chi_{0.05,1}^2$ ，可认为
  - A. 两文结果有矛盾
  - B. 两文结果完全相同
  - C. 甲文结果更为可信
  - D. 乙文结果更为可信
  - E. 甲文说明总体的差异较大
5. 两组有效率比较检验功效的相关因素是
  - A. 检验水准和样本率
  - B. 总体率差别和样本含量
  - C. 样本含量和样本率
  - D. 总体率差别和理论频数
  - E. 容许误差和检验水准

答案：C C E C B

## 二、计算与分析

1. 某神经内科医师观察 291 例脑梗塞病人，其中 102 例病人用西医疗法，其他 189 例病人采用西医疗法加中医疗法，观察一年后，单纯用西医疗法组的病人死亡 13 例，采用中西医疗法组的病人死亡 9 例，请分析两组病人的死亡率差异是否有统计学意义？

[参考答案]

本题是两组频数分布的比较，资料可整理成下表的形式。

两组疗法病人的死亡率的比较

组别	死亡	存活	合计
西医疗法	13	89	102
西医疗法加中医疗法	9	180	189
合计	22	269	291

(1) 建立检验假设并确定检验水准

$H_0: \pi_1 = \pi_2$ ，即两组病人的死亡率相等

$H_1: \pi_1 \neq \pi_2$ ，即两组病人的死亡率不等

$\alpha = 0.05$

(2) 用四个表的专用公式，计算  $\chi^2$  检验统计量  $\chi^2$  值

$$\chi^2 = \frac{(ad-bc)^2 n}{(a+b)(c+d)(a+c)(b+d)} = \frac{(13 \times 180 - 89 \times 9)^2 \times 291}{22 \times 269 \times 102 \times 189} = 6.041$$

(3) 确定 P 值，作出推断结论

以  $\nu = 1$  查附表 7 的  $\chi^2$  界值表，得  $P < 0.05$ 。按  $\alpha = 0.05$  水准，拒绝  $H_0$ ，接受  $H_1$ ，可以认为两组病人的死亡率不等。

2. 某医院研究中药治疗急性心肌梗死的疗效，临床观察结果见下表。问接受两种不同疗法的患者病死率是否不同？



两种药治疗急性心肌梗死的疗效

组别	存活	死亡	合计	病死率(%)
中药组	65	3	68	4.41
非中药组	12	2	14	14.29
合计	77	5	82	6.10

[参考答案]

本题  $T_{22} = \frac{5 \times 14}{82} = 0.853 < 1$ , 宜用四格表的确切概率法

(1) 建立检验假设并确定检验水准

$H_0: \pi_1 = \pi_2$ , 即两种不同疗法的患者病死率相同

$H_1: \pi_1 \neq \pi_2$ , 即两种不同疗法的患者病死率不同

$\alpha = 0.05$

(2) 计算确切概率

$$P_i = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{a!b!c!d!n!} = 0.2001$$

(3) 作出推断结论

按  $\alpha = 0.05$  水准, 不拒绝  $H_0$ , 无统计学意义, 还不可以认为两种不同疗法的患者病死率不同。

3. 某医师观察三种降血脂药 A, B, C 的临床疗效, 观察 3 个月后, 按照患者的血脂下降程度分为有效与无效, 结果如下表, 问三种药物的降血脂效果是否不同?

三种药物降血脂的疗效

药物	有效	无效	合计
A	120	25	145
B	60	27	87
C	40	22	62

[参考答案]

本题为 3 个样本构成比的比较, 是  $3 \times 2$  表资料。

(1) 建立检验假设并确定检验水准

$H_0$ : 三种药物的降血脂有效的概率相同

$H_1$ : 三种药物的降血脂有效的概率不同

$\alpha = 0.05$

(2) 计算检验统计量

$$\chi^2 = n \left( \sum \frac{A^2}{n_{Rn_C}} - 1 \right)$$

$$= 294 \left( \frac{120^2}{220 \times 145} + \frac{25^2}{74 \times 145} + \frac{60^2}{220 \times 87} + \frac{27^2}{74 \times 87} + \frac{40^2}{220 \times 62} + \frac{22^2}{74 \times 62} - 1 \right) = 9.93$$

$$\nu = (3-1)(2-1) = 2$$

(3) 确定 P 值, 作出推断结论

查  $\chi^2$  界值表得  $P < 0.05$ , 在  $\alpha = 0.05$  检验水准下, 拒绝  $H_0$ , 接受  $H_1$ , 认为三种药物的降血脂有效率不同。

4. 某医师按照白血病患者的发病情况, 将 308 例患者分为两组, 并按 ABO 血型分类记数, 试问两组患者血型总体构成有无差别?



308 例急、慢性白血病患者血型分布

组别	A	B	O	AB	合计
急性组	60	47	61	21	189
慢性组	42	30	34	13	119
合计	102	77	95	34	308

[参考答案]

本例为 2 个样本构成比的比较, 是  $2 \times 4$  表资料。

(1) 建立检验假设并确定检验水准

$H_0$ : 两组患者血型总体构成比相同

$H_1$ : 两组患者血型总体构成比不全相同

$\alpha=0.05$

(2) 计算检验统计量

$$\begin{aligned} \chi^2 &= n \left( \sum \frac{A^2}{n_r n_c} - 1 \right) \\ &= 308 \left( \frac{60^2}{102 \times 189} + \frac{47^2}{77 \times 189} + \dots + \frac{13^2}{34 \times 119} - 1 \right) = 0.6081 \\ \nu &= (2-1)(4-1) = 3 \end{aligned}$$

(3) 确定  $P$  值, 作出推断结论

查  $\chi^2$  界值表得  $P > 0.05$ , 在  $\alpha=0.05$  检验水准下, 不拒绝  $H_0$ , 还不能认为两组患者血型总体构成比不全相同。

5. 为研究某补钙制剂的临床效果, 观察 56 例儿童, 其中一组给予这种新药, 另一组给予钙片, 观察结果如表, 问两种药物预防儿童的佝偻病患病率是否不同?

两组儿童的佝偻病患病情况

组别	病例数	非病例数	合计	患病率(%)
新药组	8	32	40	20.0
钙片组	6	10	16	37.5
合计	14	42	56	25.0

[参考答案]

本题是两组二分类频数分布的比较, 用四格表  $\chi^2$  检验。表中  $n=56 > 40$ , 且有一个格子的理论频数小于 5, 须采用四格表  $\chi^2$  检验的校正公式进行计算。

(1) 建立检验假设并确定检验水准

$H_0$ :  $\pi_1 = \pi_2$ , 即新药组与钙片组儿童佝偻病患病概率相同

$H_1$ :  $\pi_1 \neq \pi_2$ , 即新药组与钙片组儿童佝偻病患病概率不同

$\alpha=0.05$

(2) 用四格表  $\chi^2$  检验的校正公式, 计算检验统计量  $\chi^2$  值:

$$\begin{aligned} \chi^2_c &= \frac{(|ad-bc| - n/2)^2 n}{(a+b)(c+d)(a+c)(b+d)} = \frac{(|8 \times 10 - 32 \times 6| - \frac{56}{2})^2 \times 56}{40 \times 16 \times 14 \times 52} = 1.050 \\ \nu &= 1 \end{aligned}$$

(3) 确定  $P$  值, 作出推断结论

以  $\nu=1$  查附表 7 的  $\chi^2$  界值得  $\chi^2_{0.25(1)} = 1.32$ ,  $\chi^2 < \chi^2_{0.25(1)}$ ,  $P > 0.05$ 。按  $\alpha=0.05$  水



准，不拒绝  $H_0$ ，无统计学意义，还不能认为新药组与钙片组儿童佝偻病患病概率不同。

6. 某医院 147 例大肠杆菌标本分别在 A、B 两种培养基上培养，然后进行检验，资料见下表，试分析两种培养基的检验结果是否有显著性差别？

A、B 两种培养基上培养大肠杆菌标本结果

A 培养基	B 培养基		合计
	+	-	
+	59	36	95
-	15	37	52
合计	74	73	147

[参考答案]

本题是一个配对设计的二分类资料，采用配对四格表资料的  $\chi^2$  检验。

(1) 建立检验假设并确定检验水准

$H_0: B=C$ ，即两种培养基的阳性培养率相等

$H_1: B \neq C$ ，即两种培养基的阳性培养率不相等

$\alpha=0.05$

(2) 计算检验统计量

本例  $b+c=36+15=51 > 40$ ，用配对四格表  $\chi^2$  检验公式，计算检验统计量  $\chi^2$  值

$$\chi^2 = \frac{(b-c)^2}{b+c} = \frac{(36-15)^2}{36+15} = 8.65, \nu=1$$

(3) 确定 P 值，作出推断结论

查  $\chi^2$  界值表得  $P < 0.05$ 。按  $\alpha=0.05$  水准，拒绝  $H_0$ 。认为两种培养基的阳性培养率不同。

(王乐三)

## 第九章 非参数检验

### 练习题

#### 一、单项选择题

- 对医学计量资料成组比较，相对参数检验来说，非参数秩和检验的优点是
  - 适用范围广
  - 检验效能高
  - 检验结果更准确
  - 充分利用资料信息
  - 不易出现假阴性错误
- 对于计量资料的比较，在满足参数法条件下用非参数方法分析，可能产生的结果是
  - 增加 I 类错误
  - 增加 II 类错误
  - 减少 I 类错误
  - 减少 II 类错误
  - 两类错误都增加
- 两样本比较的秩和检验，如果样本含量一定，两组秩和的差别越大说明
  - 两总体的差别越大
  - 两总体的差别越小



- C. 两样本的差别可能越大
  - D. 越有理由说明两总体有差别
  - E. 越有理由说明两总体无差别
4. 多个计量资料的比较, 当分布类型不清时, 应选择的统计方法是
- A. 方差分析
  - B. Wilcoxon  $T$  检验
  - C. Kruskal-Wallis  $H$  检验
  - D.  $u$  检验
  - E.  $\chi^2$  检验
5. 在一项临床试验研究中, 疗效分为“痊愈、显效、有效、无效”四个等级, 现欲比较试验组与对照组治疗效果有无差别, 宜采用的统计方法是
- A. Wilcoxon 秩和检验
  - B.  $2 \times 4$  列联表  $\chi^2$  检验
  - C. 四格表  $\chi^2$  检验
  - D. Fisher 确切概率法
  - E. 计算标准化率

答案: A B D C A

## 二、计算与分析

1. 某医院测定 10 名受试者针刺膻中穴前后痛阈的数据, 见下表, 试分析针刺膻中穴前后痛阈值的差异有无统计学意义?

10 名受试者针刺膻中穴前后痛阈资料

编号	针刺前	针刺后
1	600	610
2	600	700
3	685	575
4	1 050	600
5	900	600
6	1 125	1 425
7	1 400	1 350
8	750	825
9	1 000	800
10	1 500	1 400

### [参考答案]

(1) 建立假设检验

$H_0$ : 差值总体中位数为零

$H_1$ : 差值总体中位数不为零

$\alpha=0.05$

(2) 计算统计量见下表

10 名受试者针刺膻中穴前后痛阈

编号	针刺前	针刺后	差值	秩次
1	600	610	10	1
2	600	700	100	4.5
3	685	575	-110	-6
4	1 050	600	-450	-10



续表

编号	针刺前	针刺后	差值	秩次
5	900	600	-300	-8.5
6	1 125	1 425	300	8.5
7	1 400	1 350	-50	-2
8	750	825	75	3
9	1 000	800	-200	-7
10	1 500	1 400	-100	-4.5
合计				$T_+ = 17$ $T_- = 38$

$T_+ + T_- = 17 + 38 = 55$ , 总秩和  $\frac{n(n+1)}{2} = \frac{10(10+1)}{2} = 55$ , 计算准确无误

$T = \min(T_+, T_-) = 17$ .

(3) 查表及结论

现  $n=10$ , 查  $T$  界值表  $T_{0.05(10)} = 8 \sim 47$ ,  $T=17$  落在此范围内, 所以  $P < 0.05$ , 按  $\alpha = 0.05$  水准, 不拒绝  $H_0$ , 针刺膻中穴前后痛阈值的差异无统计学意义。

2. 8 名健康男子服用肠溶醋酸棉酚片前后的精液中精子浓度检查结果如下表 (服用时间 3 月), 问服用肠溶醋酸棉酚片前后精液中精子浓度有无下降?

8 名健康男子服用肠溶醋酸棉酚片前后的精液中精子浓度 (万/ml)

编号	服药前	服药后
1	6 000	660
2	22 000	5 600
3	5 900	3 700
4	4 400	5 000
5	6 000	6 300
6	6 500	1 200
7	26 000	1 800
8	5 800	2 200

[参考答案]

(1) 建立假设检验

$H_0$ : 差值总体中位数为零

$H_1$ : 差值总体中位数不为零

$\alpha = 0.05$

(2) 计算统计量, 见下表

8 名健康男子服用肠溶醋酸棉酚片前后的精液中精子浓度 (万/ml)

编号	服药前	服药后	差值	秩次
1	6 000	660	-5 340	-6
2	22 000	5 600	-16 400	-7
3	5 900	3 700	-2 200	-3
4	4 400	5 000	600	2
5	6 000	6 300	300	1



续表

编号	服药前	服药后	差值	秩次
6	6 500	1 200	-5 300	-5
7	26 000	1 800	-24 200	-8
8	5 800	2 200	-3 600	-4
合计				$T_+ = 3 \quad T_- = 33$

$T_+ + T_- = 33 + 3 = 36$ , 总秩和  $\frac{n(n+1)}{2} = \frac{8 \times (8+1)}{2} = 36$ , 计算准确无误

$T = \min(T_+, T_-) = 3$ 。

(3) 查表及结论

现  $n=8$ , 查  $T$  界值表  $T_{0.05(8)} = 3 \sim 33$ ,  $T=3$  恰好落在界点上, 所以  $P \leq 0.05$ , 按  $\alpha = 0.05$  水准, 拒绝  $H_0$ , 接受  $H_1$ , 可认为健康男子服用肠溶醋酸棉酚片前后的精液中精子浓度有差异。

3. 雌鼠两组分别给以高蛋白和低蛋白的饲料, 实验时间自生后 28 天至 84 天止, 计 8 周。观察各鼠所增体重, 结果如下表, 问两种饲料对雌鼠体重增加有无显著影响?

两种饲料雌鼠体重增加量 (g)

高蛋白组	低蛋白组	高蛋白组	低蛋白组
83	65	123	101
97	70	124	107
104	70	129	122
107	78	134	
113	85	146	
119	94	161	

[参考答案]

(1) 建立假设检验

$H_0$ : 两总体分布相同

$H_1$ : 两总体分布不同

$\alpha = 0.05$

(2) 计算统计量

将两样本 21 个数据由小到大统一编秩, 见下表:

两种饲料雌鼠体重增加量 (g)

高蛋白组	秩次	低蛋白组	秩次
83	5	65	1
97	8	70	2.5
104	10	70	2.5
107	11.5	78	4
113	13	85	6
119	14	94	7
123	16	101	9
124	17	107	11.5



续表

高蛋白组	秩次	低蛋白组	秩次
129	18	122	15
134	19		
146	20		
161	21		
172.5( $T_1$ )		58.5( $T_2$ )	

$T=58.5$

(3) 查表与结论

$n_1=9, n_2=12, n_2-n_1=3$ , 按  $\alpha=0.05$ , 查  $T$  值表得范围  $71\sim 127$ , 因为  $T=58.5 < 71$ , 超出范围, 故  $P < 0.05$ , 拒绝  $H_0$ , 接受  $H_1$ , 接受即两种饲料对雌鼠体重增加有显著影响。

4. 测得铅作业与非铅作业工人的血铅值 ( $\mu\text{g}/100\text{g}$ ), 问铅作业工人的血铅值是否高于非铅作业工人?

铅作业与非铅作业工人的血铅值 ( $\mu\text{g}/100\text{g}$ )

非铅作业组	铅作业组	非铅作业组	铅作业组
5	17	12	43
5	18	13	44
6	20	15	
7	25	18	
9	34	21	

[参考答案]

(1) 建立假设检验

$H_0$ : 两总体分布相同

$H_1$ : 两总体分布不同

$\alpha=0.05$

(2) 计算统计量

将两样本 17 个数据由小到大统一编秩, 见下表:

铅作业与非铅作业工人的血铅值 ( $\mu\text{g}/100\text{g}$ )

非铅作业组	秩次	铅作业组	秩次
5	1.5	17	9
5	1.5	18	10.5
6	3	20	12
7	4	25	14
9	5	34	15
12	6	43	16
13	7	44	17
15	8		
18	10.5		
21	13		
59.5( $T_1$ )		93.5( $T_2$ )	

$T=93.5$





(3) 查表与结论

$n_1=7, n_2=10, n_2-n_1=3$ , 按  $\alpha=0.05$ , 查  $T$  值表得范围  $42\sim84$ , 因为  $T=93.5$ , 超出范围, 故  $P<0.05$ , 拒绝  $H_0$ , 接受  $H_1$ , 铅作业工人的血铅值高于非铅作业工人的血铅值。

5. 用  $V_{K3}$  眼药水对近视眼患者作治疗, 对照组用生理盐水作安慰剂, 对两组的疗效进行观察, 结果如下表, 试分析  $V_{K3}$  眼药水对近视眼患者的治疗是否有疗效?

$V_{K3}$  眼药水治疗近视眼患者的疗效观察

疗效	$V_{K3}$ 眼药水组	生理盐水组
退步	8	20
不变	93	60
进步	11	10
恢复	4	1
合计	116	91

[参考答案]

(1) 建立假设检验

$H_0$ : 两总体分布相同

$H_1$ : 两总体分布不同

$\alpha=0.05$

(2) 计算统计量

将两样本的资料统一由小到大统一编秩, 见下表:

$V_{K3}$  眼药水治疗近视眼患者的疗效观察

疗效	$V_{K3}$ 眼药水组	生理盐水组	合计	范围	平均秩次	盐水组秩和
退步	8	20	28	1~28	14.5	290
不变	93	60	153	29~181	105	6 300
进步	11	10	21	182~202	192	1 920
恢复	4	1	5	203~207	205	205
合计	116	91	207			8 715

$$\begin{aligned}
 u &= \frac{\left| T_1 - \frac{1}{2} n_1 (N+1) \right| - 0.5}{\sqrt{\frac{n_1 \times n_2}{12N(N-1)} [N^3 - N - \sum (t^3 - t)]}} \\
 &= \frac{\left| 8715 - \frac{1}{2} \times 91 (207+1) \right| - 0.5}{\sqrt{\frac{91 \times 116}{12 \times 207 (207-1)} [207^3 - 207 - (28^3 - 28 + 153^3 - 153 + 21^3 - 21 + 5^3 - 5)]}} \\
 &= 2.27
 \end{aligned}$$

由于  $u_{0.05/2} = 1.96$ ,  $u > u_{0.05/2}$ ,  $P < 0.05$ , 拒绝  $H_0$ , 接受  $H_1$ , 有统计学意义, 可认为两总体分布不同。即可认为  $V_{K3}$  眼药水对近视眼患者的治疗有疗效。

6. 对正常、单纯性肥胖及皮质醇增多症三组人的血浆皮质醇含量进行测定, 其结果见下表, 问三组人的血浆皮质醇含量的差异有无统计学意义?



三组人的血浆皮质醇测定值 (nmol/L)

正常人	单纯性肥胖人	皮质醇增多症
0.4	0.6	9.8
1.9	1.2	10.2
2.2	2.0	10.6
2.5	2.4	13.0
2.8	3.1	14.0
3.1	4.1	14.8
3.7	5.0	15.6
3.9	5.9	15.6
4.6	7.4	21.6
7.0	13.6	24.0

[参考答案]

(1) 建立假设检验

$H_0$ : 三组人的血浆皮质醇含量的总体分布相同

$H_1$ : 三组人的血浆皮质醇含量的总体分布不全相同

$\alpha=0.05$

(2) 计算统计量

将三样本 30 个观察值统一由小到大编秩, 见下表:

三组人的血浆皮质醇测定值 (nmol/L)

正常人	秩次	单纯性肥胖人	秩次	皮质醇增多症	秩次
0.4	1	0.6	2	9.8	20
1.9	4	1.2	3	10.2	21
2.2	6	2	5	10.6	22
2.5	8	2.4	7	13	23
2.8	9	3.1	10.5	14	25
3.1	10.5	4.1	14	14.8	26
3.7	12	5	16	15.6	27
3.9	13	5.9	17	15.6	28
4.6	15	7.4	19	21.6	29
7	18	13.6	24	24	30
	96.5		117.5		251

$$\begin{aligned}
 H &= \frac{12}{N(N+1)} \sum \frac{T_i^2}{n_i} - 3(N+1) \\
 &= \frac{12}{30(30+1)} \left( \frac{96.5^2}{10} + \frac{117.5^2}{10} + \frac{251^2}{10} \right) - 3(30+1) = 18.12
 \end{aligned}$$

(3) 查表及结论

现  $k=3$ ,  $\nu=k-1=3-1=2$  查  $\chi^2$  界值表  $\chi^2_{0.05,2}=5.99$ ,  $\chi^2 > \chi^2_{0.05,3}$ ;  $P < 0.05$  按  $\alpha=0.05$  水准, 拒绝  $H_0$ , 接受  $H_1$ , 故可认为三组人的血浆皮质醇含量的总体分布有差别。

7. 在针刺麻醉下, 对肺癌、肺化脓症及肺结核三组患者进行肺部手术, 效果分四级, 结果见下表, 试比较针刺麻醉对三组病人的效果有无差异?



三组患者肺部手术的针麻效果

针麻效果	肺癌	肺化脓症	肺结核
I	10	24	48
II	17	41	65
III	19	33	36
IV	4	7	8
合计	50	105	157

[参考答案]

(1) 建立假设检验

$H_0$ : 三组病人的总体效果相同

$H_1$ : 三组病人的总体效果不全相同

$\alpha=0.05$

(2) 计算统计量

将三个样本的资料统一由小到大编秩, 见下表:

三组患者肺部手术的针麻效果

针麻效果	例数			合计	范围	平均秩次	秩和		
	肺癌	肺化脓症	肺结核				肺癌	肺化脓症	肺结核
I	10	24	48	82	1—82	41.5	415	996	1 992
II	17	41	65	123	83—205	144	2 448	5 904	9 360
III	19	33	36	88	206—293	249.5	4 740.5	8 233.5	8 982
IV	4	7	8	19	294—312	303	1 212	2 121	2 424
合计	50	105	157	312			8 815.5	17 254.5	22 758

$$H = \frac{12}{N(N+1)} \sum \frac{T_i^2}{n_i} - 3(N+1)$$

$$= \frac{12}{312(312+1)} \left( \frac{815.5^2}{50} + \frac{17\ 254.5^2}{105} + \frac{22\ 758^2}{157} \right) - 3(312+1) = 5.77$$

$$H_c = \frac{5.77}{1 - \frac{82^3 - 82 + 123^3 - 123 + 88^3 - 88 + 19^3 - 19}{312^3 - 312}} = 6.43$$

(3) 查表及结论

现  $k=3$ ,  $\nu=k-1=3-1=2$  查  $\chi^2$  界值表  $\chi^2_{0.05,2} = 5.99$ ,  $\chi^2 > \chi^2_{0.05,3}$ ;  $P < 0.05$  按  $\alpha=0.05$  水准, 拒绝  $H_0$ , 接受  $H_1$ , 故可认为三组病人的总体效果不全相同。

(刘启贵)

## 第十章 线性相关与回归

### 练习题

#### 一、单项选择题

- 使用最小二乘法确定直线回归方程的原则是
  - 各观测点距回归直线的纵向距离之和最小
  - 各观测点距回归直线的横向距离之和最小



- C. 各观测点距回归直线的垂直距离平方和最小  
 D. 各观测点距回归直线的纵向距离平方和最小  
 E. 各观测点距回归直线的横向距离平方和最小
2. 两数值变量相关关系越强, 表示  
 A. 相关系数越大  
 B. 相关系数的绝对值越大  
 C. 回归系数越大  
 D. 回归系数的绝对值越大  
 E. 相关系数检验统计量的  $t$  值越大
3. 回归分析的决定系数  $R^2$  越接近于 1, 说明  
 A. 相关系数越大  
 B. 回归方程的显著程度越高  
 C. 应变量的变异越大  
 D. 应变量的变异越小  
 E. 自变量对应变量的影响越大
4. 两组资料作回归分析, 直线回归系数  $b$  较大的一组, 表示  
 A. 两变量关系密切的可能性较大  
 B. 检验显著的可能性较大  
 C. 决定系数  $R^2$  较大  
 D. 决定系数  $R^2$  可能大也可能小  
 E. 数量依存关系更密切
5. 1~7 岁儿童可以用年龄 (岁) 估计体重 (市斤), 回归方程为  $\hat{Y}=14+4X$ , 若将体重换成国际单位 kg, 则此方程  
 A. 常数项改变  
 B. 回归系数改变  
 C. 常数项和回归系数都改变  
 D. 常数项和回归系数都不改变  
 E. 决定系数改变

答案: D B E D C

## 二、计算与分析

1. 12 名 20 岁女青年的身高与体重资料如下表, 试问女青年身高与体重之间有无相关关系?

12 名 20 岁女青年的身高与体重资料

编号	1	2	3	4	5	6	7	8	9	10	11	12
身高(cm)	164	156	172	172	177	180	166	162	172	167	158	152
体重(kg)	55	56	60	68	66	65	56	55	60	55	46	51

[参考答案]

身高为  $X$ , 体重为  $Y$ 。

$$n=12, \sum X=1998, \sum X^2=333470, \sum Y=693,$$

$$\sum Y^2=40469, \sum XY=115885$$

代入公式 (10-2), 得:

$$l_{XX} = \sum X^2 - \frac{(\sum X)^2}{n} = 333470 - \frac{1998^2}{12} = 803$$

$$l_{YY} = \sum Y^2 - \frac{(\sum Y)^2}{n} = 40469 - \frac{693^2}{12} = 448.25$$

$$l_{XY} = \sum XY - \frac{(\sum X)(\sum Y)}{n} = 115885 - \frac{1998 \times 693}{12} = 500.5$$

由公式 (10-1) 计算相关系数

$$r = \frac{500.5}{\sqrt{803 \times 448.25}} = 0.834$$

下面采用  $t$  检验法对相关系数进行检验。



(1) 建立检验假设

$H_0: \rho=0$ , 即身高与体重之间不存在相关关系

$H_1: \rho \neq 0$ , 即身高与体重之间存在相关关系

$\alpha=0.05$

(2) 计算统计量

$$t = \frac{|0.834 - 0|}{\sqrt{\frac{1 - 0.834^2}{12 - 2}}} = 4.7799$$

$$\nu = 12 - 2 = 10$$

(3) 查界值表, 得统计结论

查  $t$  界值表, 得  $t_{0.005/2, 10} = 3.581$ ,  $t > t_{0.005/2, 10}$ ,  $P < 0.005$ , 按  $\alpha = 0.05$  水准, 拒绝  $H_0$ , 接受  $H_1$ , 可以认为女青年身高与体重之间存在正相关关系。

2. 某医师研究某种代乳粉价值时, 用大白鼠做实验, 得大白鼠进食量和体重增加量的资料如下, 试问大白鼠的进食量与体重的增加量之间有无关系? 能否用大白鼠的进食量来估计其体重的增加量?

大白鼠进食量和体重增加量的资料

动物编号	1	2	3	4	5	6	7	8	9	10	11
进食量(g)	820	780	720	867	690	787	934	679	639	820	780
增重量(g)	165	158	130	180	134	167	186	145	120	150	135

[参考答案]

进食量为  $X$ , 增重量为  $Y$ 。

$$n = 11, \sum X = 8516, \sum X^2 = 6668876, \sum Y = 1670,$$

$$\sum Y^2 = 258080, \sum XY = 1309248$$

由公式 (10-2) 计算可得:

$$l_{XX} = \sum X^2 - \frac{(\sum X)^2}{n} = 6668876 - \frac{8516^2}{11} = 75943.363$$

$$l_{YY} = \sum Y^2 - \frac{(\sum Y)^2}{n} = 258080 - \frac{1670^2}{11} = 4543.636$$

$$l_{XY} = \sum XY - \frac{(\sum X)(\sum Y)}{n} = 1309248 - \frac{8516 \times 1670}{11} = 16364.364$$

由公式 (10-1) 计算相关系数:  $\gamma = \frac{16364.364}{\sqrt{75943.363 \times 4543.636}} = 0.881$

下面用  $r$  检验法对相关系数进行检验:

由  $r = 0.881$ ,  $n = 11$ ,  $\nu = 11 - 2 = 9$

查  $r$  界值表, 得  $r_{0.005/2, 9} = 0.776$ , 因  $r > r_{0.005/2, 9}$ , 故  $P < 0.005$ , 按  $\alpha = 0.05$  水准, 拒绝  $H_0$ , 接受  $H_1$ , 可以认为大白鼠的进食量与体重增加量之间存在正相关关系。

我们根据上表测得数据绘制散点图, 见下图, 可以看出大白鼠的进食量与增重量有明显的直线趋势, 我们考虑建立二者之间的线性回归方程。

由上述计算

$$l_{XX} = 75943.636, l_{XY} = 16364.364, l_{YY} = 4543.636, \bar{X} = 774.182, \bar{Y} = 151.818$$

代入公式 (10-5) 和 (10-6) 得:

$$b = \frac{l_{XY}}{l_{XX}} = \frac{16364.364}{75943.636} = 0.215$$

$$a = \bar{Y} - b\bar{X} = 151.818 - 0.215 \times 774.182 = -15.003$$



则回归方程为： $\hat{Y} = -15.003 + 0.215X$

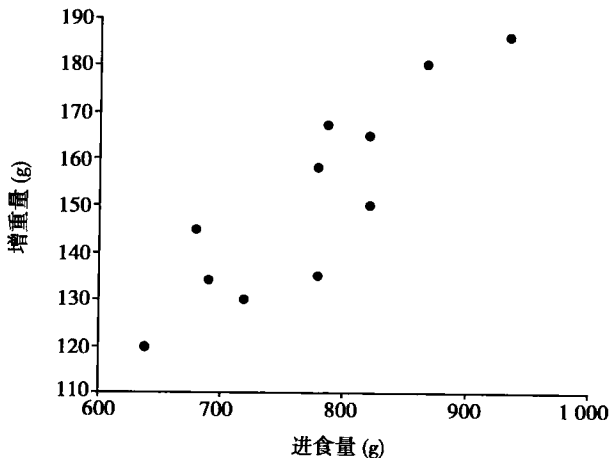


图 大白鼠的进食量与增重量之间关系散点图

最后我们采用方差分析法对回归方程作检验：

(1) 建立假设检验：

$$H_0: \beta=0; H_1: \beta \neq 0, \alpha=0.05,$$

(2) 由计算可得：

$$SS_{\text{总}} = \sum Y^2 - \frac{(\sum Y)^2}{n} = 4543.636, SS_{\text{回归}} = \frac{l_{XY}^2}{l_{XX}} = 3526.199$$

$$SS_{\text{剩余}} = SS_{\text{总}} - SS_{\text{回归}} = 1017.437$$

$$F = \frac{MS_{\text{回归}}}{MS_{\text{剩余}}} = \frac{SS_{\text{回归}}/\nu_{\text{回归}}}{SS_{\text{剩余}}/\nu_{\text{剩余}}} = 31.192$$

$$\nu_{\text{回归}} = 1, \nu_{\text{剩余}} = 11 - 2 = 9$$

(3) 查界值表，得统计结论

查  $F$  界值表， $F_{0.01(1,9)} = 10.56$ ， $F > F_{0.01(1,9)}$ ， $P < 0.01$ ，拒绝  $H_0$ ，接受  $H_1$ ，说明大白鼠的进食量与体重增加量之间存在线性回归关系，也就是说，可以用大白鼠的进食量来估计其体重的增加量。

3. 测得 347 名 13 岁健康男童的身高和体重，身高均数为 146.4cm，标准差为 8.61cm，体重均数为 37.04kg，标准差为 6.67kg。身高和体重的相关系数  $r=0.74$ ，试计算由身高推算体重的回归系数及由体重推算身高的回归系数。

[参考答案]

身高为  $X$ ，体重为  $Y$ 。

由题意， $\bar{X}=146.4$ ， $S_X=8.61$ ； $\bar{Y}=37.04$ ， $S_Y=6.67$ ， $r=0.74$

由身高推体重的回归系数为  $b_1$ ，则

$$b_1 = \frac{l_{XY}}{l_{XX}} = \frac{r \sqrt{l_{XX} \cdot l_{YY}}}{l_{XX}} = \frac{r S_Y}{S_X} = \frac{0.74 \times 6.67}{8.61} = 0.573$$

由体重推身高的回归系数为  $b_2$ ，则

$$b_2 = \frac{l_{XY}}{l_{YY}} = \frac{r \sqrt{l_{XX} \cdot l_{YY}}}{l_{YY}} = \frac{r S_X}{S_Y} = \frac{0.74 \times 8.61}{6.67} = 0.955$$

4. 某防治所作病因研究，对一些地区水质的平均碘含量 ( $\mu\text{g/L}$ ) 与地方性甲状腺肿患病率进行了调查，结果如下表，试问甲状腺肿患病率与水质中碘的含量有无相关关系？



局部地区水质的平均碘含量 ( $\mu\text{g/L}$ ) 与地方性甲状腺肿患病率

地区编号	1	2	3	4	5	6	7	8	9	10	11	12	13	14
患病率(%)	40.5	37.7	39.0	20.0	22.5	37.4	31.5	15.6	21.0	6.3	7.1	9.0	4.0	5.4
碘含量( $\mu\text{g/L}$ )	1.0	2.0	2.5	3.5	3.5	4.0	4.4	4.5	4.6	7.7	8.0	8.0	8.3	8.5

[参考答案]

甲状腺肿患病率为  $Y$ , 水质中碘含量为  $X$ 。

- ① 将  $X, Y$  分别从小到大编秩, 见下表 (3)、(5) 两栏;
- ② 计算差数  $d$ , 见 (6) 栏;
- ③ 计算  $d^2$ , 见 (7) 栏,  $\sum d^2 = 870$ ;
- ④ 代入公式 (10-18) 计算  $r_s$ ,

$$r_s = 1 - \frac{6 \times 870}{14(14^2 - 1)} = -0.912$$

下面对  $r_s$  进行检验

- (1) 建立假设检验

$$H_0: \rho_s = 0; H_1: \rho_s \neq 0, \alpha = 0.05$$

- (2) 计算统计量

$$r_s = -0.912$$

- (3) 结论

当  $n \leq 50$  时, 查附表 12 中的等级相关系数  $r_s$  界值表。

由于  $r_{0.05/2, 14} = 0.538$ ,  $|r_s| > r_{0.05/2, 14}$ ,  $P < 0.05$ , 拒绝  $H_0$ , 接受  $H_1$ , 可认为甲状腺肿患病率与水质中碘的含量存在负相关关系。

不同地区水质中碘含量 ( $\mu\text{g/L}$ ) 与甲状腺肿患病率 (%)

地区编号	水质中碘含量			甲状腺肿患病率		
	$X$	秩次	$Y$	秩次	$d$	$d^2$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	1.0	1	40.5	14	-13	169
2	2.0	2	37.7	12	-10	100
3	2.5	3	39.0	13	-10	100
4	3.5	4.5	20.0	7	-2.5	6.25
5	3.5	4.5	22.5	9	-4.5	20.25
6	4.0	6	37.4	11	-5	25
7	4.4	7	31.5	10	-3	9
8	4.5	8	15.6	6	2	4
9	4.6	9	21.0	8	1	1
10	7.7	10	6.3	3	7	49
11	8.0	11.5	7.1	4	7.5	56.25
12	8.0	11.5	9.0	5	6.5	42.25
13	8.3	13	4.0	1	12	144
14	8.5	14	5.4	2	12	144

(闫岩)



## 第十一章 多元线性回归与多元逐步回归

### 练习题

#### 一、单项选择题

- 在疾病发生危险因素的研究中,采用多变量回归分析的主要目的是
  - 节省样本
  - 提高分析效率
  - 克服共线影响
  - 减少异常值的影响
  - 减少混杂的影响
- 多元线性回归分析中,反映回归平方和在应变量  $Y$  的总离均差平方和中所占比重的统计量是
  - 简单相关系数
  - 复相关系数
  - 偏回归系数
  - 回归均方
  - 决定系数  $R^2$
- 对同一资料作多变量线性回归分析,若对两个具有不同个数自变量的回归方程进行比较,应选用的指标是
  - 决定系数
  - 相关系数
  - 偏回归平方和
  - 校正决定系数
  - 复相关系数
- 多元线性回归分析,对回归方程作方差分析,检验统计量  $F$  值反映的是
  - 所有自变量与应变量间是否存在线性回归关系
  - 部分自变量与应变量间是否存在线性回归关系
  - 自变量与应变量间存在的线性回归关系是否较强
  - 自变量之间是否存在共线
  - 回归方程的拟合优度
- 在多元回归分析中,若对某个自变量的值都乘以一个常数  $c$  ( $c \neq 0$ ),则
  - 偏回归系数不变、标准回归系数改变
  - 偏回归系数改变、标准回归系数不变
  - 偏回归系数与标准回归系数均不改变
  - 偏回归系数与标准回归系数均改变
  - 偏回归系数和决定系数均改变

答案: E E D A B

#### 二、计算与分析

- 某种特殊营养缺乏状态下,儿童年龄(岁)、身高(cm)与体重(kg)测定结果见下表,①试建立年龄、身高与体重的二元回归方程;②对回归方程作检验;③计算复相关系数与决定系数;④计算年龄和身高的标准偏回归系数。

营养缺乏儿童年龄、身高、体重测定值

编号 $i$	1	2	3	4	5	6	7	8	9	10	11	12
身高 $X_1$	145	150	124	157	129	127	140	122	107	107	155	148
年龄 $X_2$	8	10	6	11	8	7	10	9	10	6	12	9
体重 $Y$	29	32	24	30	25	26	35	26	25	23	35	31





[参考答案]

(1)

参数估计

模型	变量	偏回归系数	标准误	标准化偏回归系数	t	P
1	常数	2.114	5.048		0.419	0.685
	$X_1$	0.135	0.047	0.564	2.863	0.019
	$X_2$	0.923	0.434	0.419	2.126	0.062

回归方程为： $\hat{Y} = 2.114 + 0.135X_1 + 0.923X_2$

(2)

方差分析

变异来源	离均差平方和	自由度	均方差	F	P
回归	150.884	2	75.442	16.154	0.001
误差	42.033	9	4.670		
总变异	192.917	11			

从上表可见， $F = 16.154$ ， $P < 0.001$ ，此回归方程有统计学意义。

(3) 复相关系数  $R = 0.884$ ，决定系数  $R^2 = 0.782$ 。

(4) 年龄和身高的标准偏回归系数见参数估计表。

2. 有学者认为，血清中低密度脂蛋白增高和高密度脂蛋白降低，是引起动脉硬化的一个重要原因。现测量了 30 名动脉硬化疑似患者的载脂蛋白 AI、载脂蛋白 B、载脂蛋白 E、载脂蛋白 C、低密度脂蛋白中的胆固醇、高密度脂蛋白中的胆固醇含量，资料如下表。

① 分别作  $Y_1$  和  $Y_2$  对  $X_1, X_2, X_3, X_4$  的多元线性回归分析。

② 作  $Y_2/Y_1$  对  $X_1, X_2, X_3, X_4$  的逐步回归分析，并与前面的分析结果进行比较。

30 名动脉硬化疑似患者的观测资料

序号	载脂蛋白 AI (mg/dL)	载脂蛋白 B (mg/dL)	载脂蛋白 E (mg/dL)	载脂蛋白 C (mg/dL)	低密度脂蛋白 (mg/dL)	高密度脂蛋白 (mg/dL)
$i$	$X_1$	$X_2$	$X_3$	$X_4$	$Y_1$	$Y_2$
1	173	106	7	14.7	137	62
2	139	132	6.4	17.8	162	43
3	198	112	6.9	16.7	134	81
4	118	138	7.1	15.7	188	39
5	139	94	8.6	13.6	138	51
6	175	160	12.1	20.3	215	65
7	131	154	11.2	21.5	171	40
8	158	141	9.7	29.6	148	42
9	158	137	7.4	18.2	197	56
10	132	151	7.5	17.2	113	37
11	162	110	6	15.9	145	70
12	144	113	10.1	42.8	81	41
13	162	137	7.2	20.7	185	56



续表

序号	载脂蛋白 AI (mg/dL)	载脂蛋白 B (mg/dL)	载脂蛋白 E (mg/dL)	载脂蛋白 C (mg/dL)	低密度脂蛋白 (mg/dL)	高密度脂蛋白 (mg/dL)
$i$	$X_1$	$X_2$	$X_3$	$X_4$	$Y_1$	$Y_2$
14	169	129	8.5	16.7	157	58
15	129	138	6.3	10.1	197	47
16	166	148	11.5	33.4	156	49
17	185	118	6	17.5	156	69
18	155	121	6.1	20.4	154	57
19	175	111	4.1	27.2	144	74
20	136	110	9.4	26	90	39
21	153	133	8.5	16.9	215	65
22	110	149	9.5	24.7	184	40
23	160	86	5.3	10.8	118	57
24	112	123	8	16.6	127	34
25	147	110	8.5	18.4	137	54
26	204	122	6.1	21.0	126	72
27	131	102	6.6	13.4	130	51
28	170	127	8.4	24.7	135	62
29	173	123	8.7	19.0	188	85
30	132	131	13.8	29.2	122	38

[参考答案]

①  $Y_1$  和  $Y_2$  对  $X_1, X_2, X_3, X_4$  的多元线性回归分析。

$Y_1$  为应变量的方差分析表

变异来源	离均差平方和	自由度	均方差	$F$	$P$
回归	18 530.408	4	4 632.602	8.090	0.000
误差	14 316.258	25	572.650		
总变异	32 846.667	29			

$Y_1$  为应变量的参数估计

变量	偏回归系数	标准误	标准化偏回归系数	$t$	$P$
常数	-0.829	47.773	—	-0.017	0.986
$X_1$	0.233	0.197	0.165	1.181	0.249
$X_2$	1.325	0.282	0.714	4.699	0.000
$X_3$	-0.124	2.783	-0.008	-0.045	0.965
$X_4$	-2.385	0.765	-0.494	-3.119	0.005

复相关系数  $R=0.751$ , 决定系数  $R^2=0.564$ 。

按  $\alpha=0.05$  检验水准, 回归方程中  $X_2$  和  $X_4$  有统计学意义, 即低密度脂蛋白中的胆固醇与载脂蛋白 B 及载脂蛋白 C 有线性回归关系。与载脂蛋白 B 呈正相关, 而与载脂蛋白



呈负相关。

$Y_2$  为应变量的方差分析表

变异来源	离均差平方和	自由度	均方差	F	P
回归	4 392.581	4	1 098.145	22.487	0.000
误差	1 220.886	25	48.835		
总变异	5 613.467	29			

$Y_2$  为应变量的参数估计

变量	偏回归系数	标准误	标准化偏回归系数	t	P
常数	-2.132	13.951	—	-0.153	0.880
$X_1$	0.483	0.058	0.825	8.385	0.000
$X_2$	-0.053	0.082	-0.069	-0.640	0.528
$X_3$	-0.294	0.813	-0.046	-0.362	0.720
$X_4$	-0.415	0.223	-0.208	-1.858	0.075

此时，复相关系数  $R=0.885$ ，决定系数  $R^2=0.783$ 。

按  $\alpha=0.05$  检验水准，回归方程中只有  $X_1$  有统计学意义，即高密度脂蛋白中的胆固醇含量与载脂蛋白 AI 有线性回归关系，并呈正相关。

② 作  $Y_2/Y_1$  关于  $X_1, X_2, X_3, X_4$  的逐步回归，选入水准  $\alpha_{\text{选入}}=0.05$ ，剔除水准  $\alpha_{\text{剔除}}=0.10$ 。

$Y_2/Y_1$  为应变量的方差分析表

变异来源	离均差平方和	自由度	均方差	F	P
回归	0.283	3	0.094	46.846	0.000
误差	0.052	26	0.002		
总变异	0.336	29			

$Y_2/Y_1$  为应变量的参数估计

变量	偏回归系数	标准误	标准化偏回归系数	t	P
常数	0.355	0.088	—	4.018	0.000
$X_1$	0.003	0.000	0.583	7.357	0.000
$X_2$	-0.004	0.000	-0.612	-7.507	0.000
$X_4$	0.003	0.001	0.216	2.700	0.012

此时，复相关系数  $R=0.919$ ，决定系数  $R^2=0.844$ 。

按  $\alpha=0.05$  检验水准，回归方程中  $X_1, X_2$  和  $X_4$  有统计学意义，即高、低密度脂蛋白中的胆固醇含量的比值与载脂蛋白 B、载脂蛋白 C 和载脂蛋白 AI 有线性回归关系，并与载脂蛋白 C 及载脂蛋白 AI 呈正相关，而与载脂蛋白 B 呈负相关。

与前面的回归结果比较，用  $Y_2/Y_1$  作应变量得到的回归方程  $R^2=0.844$ ，比单独用  $Y_1$  或  $Y_2$  作应变量得到的回归方程 ( $Y_1: R^2=0.564; Y_2: R^2=0.783$ ) 要高，这提示：高、低密度脂蛋白中的胆固醇含量的比值，较低密度脂蛋白中的胆固醇含量或高密度脂蛋白中的胆固醇含量，对诊断动脉硬化可能更有价值。

(尹 平)



## 第十二章 统计表与统计图

### 练习题

#### 一、单项选择题

1. 统计表的主要作用是
  - A. 便于形象描述和表达结果
  - B. 客观表达实验的原始数据
  - C. 减少论文篇幅
  - D. 容易进行统计描述和推断
  - E. 代替冗长的文字叙述和便于分析对比
2. 描述某疾病患者年龄（岁）的分布，应采用的统计图是
  - A. 线图
  - B. 条图
  - C. 百分条图
  - D. 直方图
  - E. 箱式图
3. 高血压临床试验分为试验组和对照组，分析考虑治疗 0 周、2 周、4 周、6 周、8 周血压的动态变化和改善情况，为了直观显示出两组血压平均变动情况，宜选用的统计图是
  - A. 半对数图
  - B. 线图
  - C. 条图
  - D. 直方图
  - E. 百分条图
4. 研究三种不同麻醉剂在麻醉后的镇痛效果，采用计量评分法，分数呈偏态分布，比较终点时分数的平均水平及个体的变异程度，应使用的图形是
  - A. 复式条图
  - B. 复式线图
  - C. 散点图
  - D. 直方图
  - E. 箱式图
5. 研究血清低密度脂蛋白 LDL 与载脂蛋白 B-100 的数量依存关系，应绘制的图形是
  - A. 直方图
  - B. 箱式图
  - C. 线图
  - D. 散点图
  - E. 条图

答案：E D B E D

#### 二、改表和绘图

1. 某地调查脾肿大和疟疾临床分型的关系、程度与血片查疟原虫结果列表如下，此表有何缺点，请改进。

项目 脾 肿程度	血膜 阴性	血膜阳性				合计		
		恶性疟		间日疟		例数	%	
		例数	%	例数	%			
脾肿者	174	28	12.6	20	9.04	222	48	21.6
脾 I	105	8	6.6	9	7.40	122	17	13.9
脾 II	51	14	20.0	5	7.10	70	19	27.1
脾 III	15	6	23.1	5	19.20	26	11	42.3
	3	0	0.0	1	25.00	4	1	25.0



[参考答案]

本表的缺点有：(1) 无标题；(2) 横标目与纵标目分类不明确，标目设计不合理，“合计”不清晰，不便于比较分析；(3) 线条过多，比例数小数位不统一。具体修改如下，见下表，也可以把血膜阳性分类单独列表。

某地脾肿大程度和疟疾临床分型的关系

脾肿程度	血膜阴性		血膜阳性						合计	
	例数	%(*)	恶性疟		间日疟		合计		例数	%(△)
			例数	%(*)	例数	%(*)	例数	%(*)		
脾 I	105	86.06	8	6.56	9	7.38	17	13.94	122	54.96
脾 II	51	72.86	14	20.00	5	7.14	19	27.14	70	51.53
脾 III	15	57.69	6	23.08	5	19.23	11	42.31	26	11.71
其他	3	75.00	0	0.00	1	25.00	1	25.00	4	1.80
合计	174	78.38	28	12.61	20	9.01	48	21.60	222	100

注：(\*) 表示行百分比，(△) 表示列百分比

2. 试根据下表资料绘制适当统计图形。

某地 1975 年 839 例正常人发汞值分布资料 (μg/g)

组段	0~	0.2~	0.4~	0.6~	0.8~	1.0~	1.2~	1.4~	1.6~2.2	合计
例数	133	193	190	111	83	34	43	16	36	839

[参考答案]

本题需要应用直方图表示 839 例正常人发汞值分布情况，由于最后一组的组距与其他组不等，制图时转换成：矩形高度 = 组段频数 / 组距，即变成等组距。为保持原始数据的组距一致为 0.2，也可把最后一组频数转换为  $36 / (0.6 / 0.2) = 12$ ，频数图见图 1。

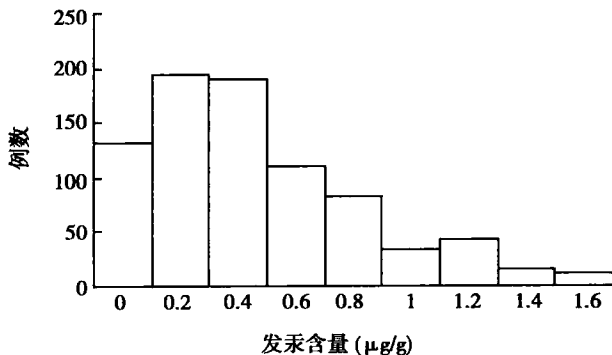


图 1 某地 1975 年 839 例正常人发汞值分布图

3. 根据下表分别绘制普通线图和半对数线图，并说明两种统计图型的意义。

某地某年食管癌年龄别发病率 (1/10 万)

年龄(岁)	男	女
40~	4.4	2.1
45~	7.2	3.3
50~	7.3	4.5



续表

年龄(岁)	男	女
55~	6.9	5.5
60~	19.3	6.7
65~	50.2	16.4
70~	68.5	12.5
75~	86.2	19.9
80~	97.0	15.2

[参考答案]

将表中数据绘制成普通线图(图2),可以看出,60岁之前,男女食管癌年龄别发病率随年龄增长的变化趋势差异较小,60岁之后,男性随年龄变化食管癌发病率比女性增长较快,差异明显扩大。将表中数据绘制成半对数线图(图3),不同性别随年龄变化食管癌年龄别发病率变化快慢程度相当,而且女性食管癌发病率趋势和转折点更清楚。应用半对数线图能够更恰当地表示相对指标的变化趋势。

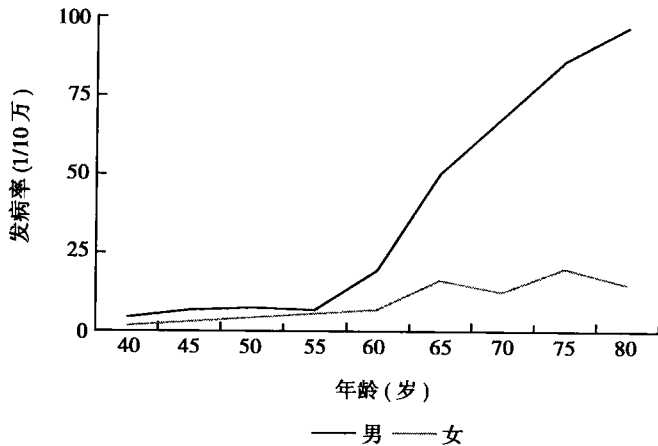


图2 某地某年不同性别食管癌年龄别发病率(1/10万)比较(普通线图)

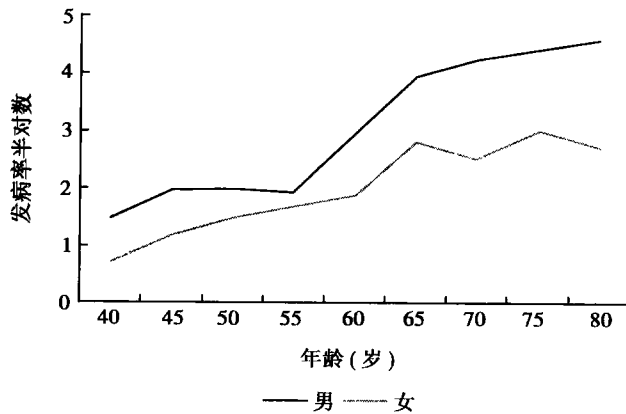


图3 某地某年不同性别食管癌年龄别发病率对数比较(半对数线图)

(张丕德 郭德)



## 第十三章 医学实验设计与诊断试验的评价

### 练习题

#### 一、单项选择题

- 实验研究随机化分组的目的是
  - 减少抽样误差
  - 减少实验例数
  - 保证客观
  - 提高检验准确度
  - 保持各组的非处理因素均衡一致
- 关于实验指标的准确度和精密度，正确的说法是
  - 精密度较准确度更重要
  - 准确度较精密度更重要
  - 精密度主要受随机误差的影响
  - 准确度主要受随机误差的影响
  - 精密度包含准确度
- 在临床试验设计选择对照时，最可靠的对照形式是
  - 历史对照
  - 空白对照
  - 标准对照
  - 安慰对照
  - 自身对照
- 两名医生分别阅读同一组 CT 片诊断某种疾病，Kappa 值越大说明
  - 观察个体的变异越大
  - 观察个体的变异越小
  - 观察一致性越大
  - 机遇一致性越大
  - 实际一致性越大
- 下列叙述正确的有
  - 特异度高说明测量的稳定性好
  - 灵敏度必须大于特异度才有实际意义
  - 增大样本含量可以同时提高灵敏度和特异度
  - 特异度高说明假阳性率低
  - 阳性预测值高说明患病的概率大

答案：E C D C C

#### 二、计算与分析

将 100 名头外伤的急诊患者作为受检人群，用患者失去知觉的时间、头部表面外伤严重程度和神经检查结果作为有无颅骨骨折的新的综合诊断标准，与所有患者的影像诊断有无颅骨骨折进行比较，结果见下表，试计算其敏感度、特异度、总的符合率、阳性预测值、阴性预测值，并对该诊断试验标准进行评价。

100 名头外伤人群的疾病状态和诊断结果

影像诊断	新的诊断方法		合计
	阳性	阴性	
有骨折	48(a)	2(b)	50
无骨折	10(c)	40(d)	50
合计	58	42	100



[参考答案]:

(1) 计算评价指标

$$\text{敏感度: } S_e = \frac{a}{a+b} = \frac{48}{50} = 0.96$$

$$\text{特异度: } S_p = \frac{d}{c+d} = \frac{40}{50} = 0.80$$

$$\text{总的符合率: } \hat{\pi} = \frac{a+d}{a+b+c+d} = \frac{48+40}{100} = 0.88$$

$$\text{阳性预测值: } PV_+ = \frac{a}{a+c} = \frac{48}{58} = 0.828$$

$$\text{阴性预测值: } PV_- = \frac{d}{d+c} = \frac{40}{42} = 0.952$$

(2) 对该诊断试验标准进行评价

$$P_A = \frac{48+40}{100} = 0.8800$$

$$P_e = \frac{1}{100} \left( \frac{58 \times 50}{100} + \frac{42 \times 50}{100} \right) = \frac{29+21}{100} = 0.50$$

$$Kappa = \frac{P_A - P_e}{1 - P_e} = \frac{0.88 - 0.50}{1 - 0.50} = \frac{0.38}{0.50} = 0.76$$

$$S_{\bar{k}} = \frac{1}{(1-0.50)\sqrt{100}} \sqrt{0.50 + (0.50)^2 - \frac{58 \times 50(58+50) + 42 \times 50(42+50)}{(100)^3}}$$

$$= \frac{1}{5} \sqrt{0.50 + 0.25 - 0.51} = 0.0987$$

$$U = \frac{Kappa}{S_{\bar{k}}} = \frac{0.76}{0.0987} = 7.700$$

$$U > U_{0.01}, P < 0.01$$

(马斌荣)



---

## 参考文献

1. 郭祖超主编. 医学统计学. 北京: 人民军医出版社, 2000
2. 蒋庆琅著. 方积乾等译. 实用统计分析方法. 北京: 北京医科大学中国协和医科大学联合出版社, 1998
3. 马斌荣编. 医学科研中的统计方法. 第3版. 北京: 北京科学技术出版社, 2005
4. 倪宗瓚主编. 卫生统计学, 第4版. 北京: 人民卫生出版社, 2000
5. 蒋知俭主编. 医学统计学, 北京: 人民卫生出版社, 1998
6. Xiao-Hua Zhou, Nancy A. Obuchowski, Donna K. McClish, 著. 宇传华译. 诊断医学统计学. 北京: 人民卫生出版社, 2005
7. 胡良平编. 现代统计学与SAS应用. 北京: 军事医学科学出版社, 2000
8. 方积乾主编. 医学统计学与电脑实验. 上海: 上海科学技术出版社, 1997
9. Ying Lu, Ji-Qian Fang Advanced Medical Statistics, World Scientific, 2003
10. 金丕焕主编. 医用统计方法. 上海: 上海医科大学出版社, 1999
11. Montgomery DC 著. 汪仁宫 陈荣昭译. 实验设计与分析. 北京: 中国统计出版社, 1998
12. 马斌荣主编. SPSS for Windows Ver11. 5 在医学统计中的应用. 第3版. 北京: 科学出版社, 2004
13. Raymond S. Greenberg, Medical Epidemiology, The McGraw-Hill Companies, Inc, 2001
14. Thomas Glover & Kevin Mitchell, An Introduction to Bio-statistics, The McGraw-Hill Companies, Inc, 2001