

SPSS 应用系列丛书 (1)

世界优秀统计工具

SPSS 11

统计分析教程

基础篇

张文彤 主编



北京希望电子出版社
Beijing Hope Electronic Press
www.bhp.com.cn

R195.1-39
Z256
JC
2002
C-1

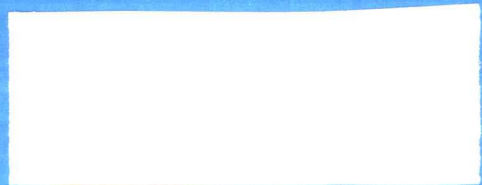
SPSS 应用系列丛书 (1)

世界优秀统计工具

SPSS 11 统计分析教程

基础篇

张文彤 主编



北京希望电子出版社
Beijing Hope Electronic Press
www.bhp.com.cn

内 容 简 介

SPSS 是世界最为优秀的统计工具之一, 深受各行业用户的青睐, SPSS 11.0 是其最新版本。

本书为《SPSS 11.0 统计分析教程》的基础篇, 由 3 部分 16 章及 3 个附录组成。主要内容包括: SPSS 数据分析实例, 数据编辑窗口用法及 Transform/Data 菜单详解, SPSS 编程操作入门, SPSS 结果窗口用法和 Help 菜单详解, 常用统计图形的制作和编辑, 交互式统计图, 报表, 描述性统计分析, 均数间的比较, 非参数统计分析方法, 卡方世界及相关分析等。

本书作者从统计专业用户的角度出发, 结合自身多年的 SPSS 使用经验, 在以风趣、明快的笔触介绍软件操作的同时, 注意将相应的统计学知识融入其中。书中既有深入浅出的软件功能介绍, 又有针对实际问题的解决办法, 更侧重于对统计新方法、新观点的讲解。

本书不仅是 SPSS 10~11 版的通用入门教材, 也是各行业从事数据统计与分析人员的重要指导书, 同时也可作为大专院校相关专业的参考教材。

本版 CD 为 SPSS11.0 相关材料 and 书中有关数据, 并赠送 SPSS11.0 试用版软件。

- 系列盘书名 : SPSS应用系列丛书(1)
盘 书 名 : 世界优秀统计工具 SPSS 11.0统计分析教程(基础篇)
文 本 著 者 : 张文彤 等
责 任 编 辑 : 郭淑珍
CD 制 作 者 : 希望多媒体开发中心
CD 测 试 者 : 希望多媒体测试部
出 版、发 行 者 : 北京希望电子出版社
地 址 : 北京市海淀区知春路63号卫星大厦三层 100080
网 址 : www.bhp.com.cn
E-mail: lwm@bhp.com.cn
电 话 : 010-62520290,62521724,62528991,62630301,62524940,62521921,82610344
(发行) 010-62613322-215(门市) 010-82675588-501,82675588-201(编辑部)
- 经 销 : 各地新华书店、软件连锁店
排 版 : 希望图书输出中心 全卫
CD 生 产 者 : 北京中新联光盘有限责任公司
文 本 印 刷 者 : 北京媛明印刷厂
开 本 / 规 格 : 787 毫米×1092 毫米 1/16 19.50 印张 450 千字
版 次 / 印 次 : 2002 年 6 月第 1 版 2002 年 6 月第 1 次印刷
印 数 : 0001-4000 册
本 版 号 : ISBN 7-900101-22-5
定 价 : 35.00 元(本版 CD)

说明: 凡我社产品如有残缺, 可执相关凭证与本社调换。

前 言

计算机永远是属于年轻人的，统计软件也是如此。作为一个极具活力和开拓精神的软件，SPSS (Statistical Product and Service Solutions) 自身的进步非常迅速，而它近几年来更是以令人惊讶的速度在国内得到了迅速普及，这无疑使用者对该软件本身的肯定。但是，在拥有这一优秀软件的同时，广大用户希望得到一本优秀 SPSS 参考书的呼声也越来越高。现在随着 11.0 版的正式推出，这一问题变量更加明显。近年来，国内也出现了数种非官方的 SPSS 教材，总体看来，各有优势，但不足之处也比较明显：

- ◇ SPSS 是一个非常权威而严肃的统计软件，可现在的许多应用型教材都存在着各种各样的常识性错误，统计理论也似是而非，用户完全无法从中体会到 SPSS 的强大功能，最后以为该软件只能作一些简单的分析，华而不实。这对 SPSS 的形象造成了极坏的影响。
- ◇ 现有几种统计专业人士编写的 SPSS 书籍在统计理论上是非常严肃的，但大多以编程为主线，或者仍然以老版本界面操作为主，内容从 6.0 一直到 10.0 都通用，完全没有体现出 SPSS 10~11 版许多出色的新功能。同时这些教材没有照顾到非统计专业人员的特点，写得过于专业、晦涩难懂，并不适合初学者入门。
- ◇ 据笔者不完全统计，所有的教材都只涉及到 SPSS 约 2/3 的常用功能，中、高级统计分析要么完全不涉及，要么走马观花的一笔带过，完全没有实用价值。特别是在市场研究领域被广泛应用的几个 SPSS 模块始终没有相应教材可用，这不能不说是一个遗憾。
- ◇ 没有真正易学易用的教材。理想的教材应当是深入浅出，同时幽默风趣，以激发读者自学的兴趣。非常可惜，完全满足这些特色的教材还没有出现。

针对以上问题，本书的写作目的是尽快提高广大使用者的水平，真正掌握最新的 SPSS 11.0 版的强大功能。按照读者的不同层次，本书分为基础篇和高级篇两册，共 31 章，内容以 SPSS 11 为准，包括了全部 10 个模块的所有主要功能。在写法上充分注意了通俗易懂：基础篇涵盖了常用的统计分析方法，入门部分强调文字轻松愉快，同时体现出 SPSS 操作上最具特色的功能和操作技巧。其中第 1 章专为初学者而准备，如果希望快速入门，读者可以在学习完第 1 章后直接跳到基础统计部分继续学习；基础统计部分充分考虑到了非统计专业人员的特点，将统计理论融入软件介绍之中，力求深入浅出；高级篇中的统计模块介绍则以统计学理论为准绳，立足于应用实例将统计方法、界面操作与结果解释结合讲述，使读者学后能真正掌握相应方法，而不是只明白了对话框的中文含义。为保证质量，其中比较重要的一般线性（方差分析）模型、多元线性回归模型、Logistic 模型、时间序列模型、生存分析均不惜篇幅详细介绍，并多由专人负责编写。

所谓术业有专攻，笔者主要从事的是医学统计和市场研究统计分析，因此书中的大部分数据实例都来自这两个领域，但本书更多的是作为 11.0 版的通用入门教材而编写，因此适用于各行业的初/中级用户。由于 SPSS 的功能极为强大，全部学习完毕需要相当长的时间，对于希望快速入门的朋友，这里针对不同专业给出参考阅读顺序如下：

- ◇ 临床科研工作者：基础篇 1、7、11~15 章，高级篇 1、4、5、13 章。但如果需要从事新药临床试验工作，则请务必阅读基础篇第 4 章，以补充编程知识。
- ◇ 市场研究工作者：基础篇 1、7、9~16 章，高级篇 1、4、8~11 章，其中高级篇 8~11 章涉及到了许多市场研究专用方法，是学习的重点。
- ◇ 社会学工作者：基础篇 1、7、11~15 章，高级篇 1、4、8~10 章。

但是，为了保证学习的系统性和连贯性，只要时间允许，笔者仍然强烈建议朋友们按章节顺序渐进学习，这样才能真正体会到 SPSS 软件的强大实力。

本书的雏形来自笔者在医学统计之星网站 (<http://www.MesStatStar.com>) 上连载的 SPSS 10 教程，但更多的心得体会则来自于长期从事 SPSS、SAS 教学的积累，以及数年来的 SPSS 使用经验。从 2000 年 4 月创作网上教程算起，全书的写作一共用了 20 个月的时间，虽然网络教程一面世就立刻受到了广大网友的热烈欢迎，但为了保证质量，真正向大家奉献一本精品，我们没有急于求成，将其匆匆结册出版，而是三易其稿，力求文字浅显易读，统计理论正确无误，并结合最新的 11.0 版的情况，对内容进行了长达三个月的详细修订。一言以蔽之，我们无愧于心。

本书同时也是作为复旦大学研究生用教材，编委基本上都是复旦大学卫生统计与社会医学教研室的年轻统计教师。其中田晓燕负责编写统计绘图部分，刘晓云负责分类资料的回归分析部分，罗剑峰负责时间序列部分，董伟负责生存分析部分，张文彤负责其余各章节和全书的统稿工作。

非常感谢苏立民总经理为我提供了这样一个机会，使我能够将自己的使用经验和大家一起分享。在成书过程中我还得到了卫生统计学、数理统计学和社会学界许多前辈、师长和朋友们的热心指导和帮助，在此一并致谢。

为节省篇幅，全书大部分的实例数据并未在书中列出，大家可以在书后所附光盘上找到它们，也可以到我的医学统计之星网站上下载。书后光盘中同时包括了 SPSS 公司主要产品的一月试用版，包括 SPSS 11 Base 一月试用版，从而大家就可以按照书中的叙述进行操作，让结果在屏幕上真实再现，相信这样对朋友们的学习更有帮助。

教材的编写有各种各样的风格，本书风格的最大特点就是幽默风趣，易于理解。如果将 SPSS 用户手册比作辞海的话，我们的编写目的就是提供一本大家学习 SPSS 时得心应手、图文并茂的新华字典，成为初学者快速成长为 SPSS 专家的桥梁。SPSS 是一款非常出色的统计软件，祝大家在本书的帮助下能够将它用得开心，玩得愉快。

限于作者水平，书中错缪之处难免，还请同行专家和广大读者不吝赐教。为便于交流，现列出各编委的电子信箱如下，欢迎大家就任何问题与我们联系。

张文彤：wtzhang@spss.com.cn

田晓燕：xytian@epscn.com

刘晓云：liuxiaoyun@hotmail.com

罗剑峰：jfluo@shmu.edu.cn

董伟：dwshmu@sina.com

张文彤

2002 年元旦于复旦公卫学院

目 录

第一部分 基本操作入门	
第 1 章 初探 SPSS——数据分析实例 2	
1.1 数据的输入和保存..... 3	
1.1.1 SPSS 的界面..... 3	
1.1.2 定义变量..... 3	
1.1.3 输入数据..... 4	
1.1.4 保存数据..... 5	
1.2 数据的预分析..... 6	
1.2.1 数据的简单描述..... 6	
1.2.2 绘制直方图..... 8	
1.3 按题目要求进行统计分析..... 9	
1.4 保存和导出分析结果..... 10	
1.4.1 保存结果文件..... 10	
1.4.2 导出分析结果..... 10	
第 2 章 数据编辑窗口用法详解 12	
2.1 数据编辑窗口操作入门..... 12	
2.1.1 SPSS 中的变量类型 与测量尺度..... 12	
2.1.2 定义新变量..... 13	
2.1.3 数据的录入与修改..... 15	
2.1.4 数据编辑窗口操作技巧集锦.. 16	
2.2 数据文件的管理——File 菜单详解.. 18	
2.2.1 新建数据文件..... 18	
2.2.2 打开其他格式的数据文件..... 19	
2.2.3 保存数据文件..... 23	
2.2.4 File 菜单中的其他条目..... 24	
2.3 Utilities 菜单详解..... 25	
第 3 章 数据文件的进一步整理 ——Transform/Data 菜单详解..... 28	
3.1 SPSS 对话框元素介绍..... 28	
3.1.1 一级对话框元素介绍..... 28	
3.1.2 二级对话框元素介绍..... 29	
3.1.3 对话框常用操作技巧..... 30	
3.2 Transform 菜单详解..... 30	
3.2.1 Compute 过程..... 30	
3.2.2 Count 过程..... 32	
3.2.3 Recode 过程..... 33	
3.2.4 Categorize Variables 过程..... 34	
3.2.5 Rank Cases 过程..... 34	
3.2.6 Transform 菜单中的其他命令.. 36	
3.3 Data 菜单详解..... 36	
3.3.1 Sort Cases 过程..... 36	
3.3.2 Transpose 过程..... 37	
3.3.3 Restructure 过程..... 38	
3.3.4 Merge Files 过程..... 39	
3.3.5 Aggregate 过程..... 42	
3.3.6 Split File 过程..... 42	
3.3.7 Select Cases 过程..... 43	
3.3.8 Weight Cases 过程..... 44	
3.3.9 Data 菜单中的其他命令..... 44	
3.4 Transform/Data 菜单命令 高级应用技巧集锦..... 44	
3.4.1 二项分布/Poisson 分布 累计概率的计算..... 44	
3.4.2 抽奖问题..... 45	
3.4.3 随机分配 / 抽样问题..... 46	
3.4.4 从原始数据生成简易频数表.. 46	
3.4.5 按字符型变量取值筛选记录.. 47	
第 4 章 SPSS 编程操作入门 48	
4.1 程序编辑窗口操作入门..... 48	
4.1.1 程序编辑窗口界面..... 48	
4.1.2 邂逅 Paste 按钮..... 49	
4.1.3 用程序加快我们的工作..... 49	
4.1.4 编程进行对话框无法 完成的工作..... 50	
4.2 结构化语句简介..... 51	
4.2.1 分支(条件)语句..... 51	
4.2.2 循环语句..... 52	
4.2.3 一个复杂程序示例..... 53	
4.3 SPSS Production Facility 操作入门... 53	
4.3.1 界面说明..... 54	

4.3.2 使用实例.....55

第5章 SPSS 结果窗口用法详解.....56

5.1 结果窗口元素介绍.....56

5.1.1 结果浏览窗口.....56

5.1.2 结果草稿浏览窗口.....58

5.1.3 如何美化输出结果
——SPSS 选项设置.....59

5.2 驾驭结果浏览窗口.....60

5.2.1 结果窗口的一般操作.....60

5.2.2 结果的导出.....61

5.2.3 如何在 Word 等软件中
使用输出结果.....62

5.3 文本输出结果的编辑.....64

5.4 表格编辑方法详解.....64

5.4.1 基本操作.....64

5.4.2 特色菜单内容详解.....65

第6章 自学 SPSS——Help 菜单详解.....69

6.1 初生牛犊.....70

6.1.1 自学模块 (Tutorial).....70

6.1.2 统计教练 (Statistics Coach).....71

6.2 笑傲江湖.....74

6.2.1 无处不在的界面帮助.....74

6.2.2 如何使用标准帮助文件.....75

6.3 凤舞九天.....77

6.3.1 语法指南 (Syntax Guide).....77

6.3.2 网络上的可用资源.....78

第二部分 统计图和统计报表

第7章 常用统计图形的制作和编辑.....82

7.1 SPSS 绘图功能简介.....83

7.2 统计图操作入门——条图.....84

7.2.1 引例和预定义对话框介绍.....84

7.2.2 条图的通用界面.....85

7.2.3 其他汇总模式的简单条图.....87

7.2.4 复式条图与分段条图.....89

7.3 统计图编辑方法详解.....90

7.3.1 编辑窗口介绍.....90

7.3.2 Gallery 菜单详解.....91

7.3.3 Chart 菜单.....92

7.3.4 Format 菜单.....94

7.3.5 其他菜单介绍.....96

7.4 线图和面积图.....97

7.4.1 线图.....97

7.4.2 线图特色编辑功能详解.....98

7.4.3 面积图制作及编辑详解.....99

7.5 散点图.....99

7.5.1 简单散点图.....100

7.5.2 散点图矩阵.....100

7.5.3 重叠散点图.....101

7.5.4 三维散点图.....102

7.5.5 散点图特色编辑功能详解.....102

7.6 Pareto 图和控制图.....106

7.6.1 Pareto 图.....106

7.6.2 控制图.....107

7.7 箱式图和误差图.....109

7.7.1 箱式图.....109

7.7.2 误差图.....110

7.8 P-P 图和 Q-Q 图.....111

7.8.1 P-P 概率图.....111

7.8.2 Q-Q 图.....113

7.9 其他统计图.....113

7.9.1 饼图.....113

7.9.2 直方图.....114

7.9.3 高一低图.....116

7.9.4 ROC 曲线.....117

7.10 统计地图简介.....119

第8章 交互式统计图.....122

8.1 概述.....122

8.2 交互式条图.....123

8.2.1 预备知识.....123

8.2.2 条图的界面.....124

8.2.3 跟我学——制作特色条图.....127

8.3 交互式点图、线图、条带图、
垂线图与面积图.....129

8.3.1 界面介绍.....130

8.3.2 实例操作.....131

8.4 交互式饼图.....133

8.4.1 简单饼图.....133

8.4.2 复式饼图..... 134

8.4.3 散点饼图..... 135

8.5 交互式箱式图和误差图..... 136

8.5.1 箱式图 136

8.5.2 误差图 137

8.6 交互式直方图..... 139

8.6.1 界面介绍..... 139

8.6.2 实例操作..... 140

8.7 交互式散点图..... 141

8.7.1 界面介绍..... 141

8.7.2 实例操作..... 143

8.8 交互图的编辑..... 145

8.8.1 概述 145

8.8.2 图形管理员
(Chart Manager)
——Plot Area 区..... 146

8.8.3 图形管理员
——Legend 区..... 152

8.8.4 图形管理员
——Elements 区..... 154

8.8.5 3-D 调色板..... 159

8.8.6 Utility 工具栏的其他选项..... 159

8.8.7 Cursor、Text
与 Style 工具栏..... 161

8.8.8 右键菜单中提供的
其他编辑功能..... 162

第 9 章 报表天地 (上)

——Reports 菜单详解..... 164

9.1 OLAP Cubes 过程..... 165

9.1.1 引例 165

9.1.2 界面说明..... 165

9.1.3 结果解释..... 167

9.1.4 对引例的进一步分析..... 168

9.2 Case Summaries 过程..... 168

9.2.1 引例 169

9.2.2 界面说明..... 169

9.2.3 结果解释..... 170

9.3 Report Summaries in Rows 过程..... 171

9.3.1 引例 171

9.3.2 界面说明..... 172

9.3.3 结果解释..... 176

9.4 Report Summaries in
Columns 过程..... 177

9.4.1 引例..... 177

9.4.2 界面说明..... 178

9.4.3 结果解释..... 180

第 10 章 报表天地 (下)

——Custom Tables 菜单详解... 182

10.1 Basic Tables 过程..... 182

10.1.1 引例..... 182

10.1.2 界面说明..... 183

10.1.3 结果解释..... 186

10.2 General Tables 过程..... 187

10.2.1 引例..... 187

10.2.2 界面说明..... 188

10.2.3 结果解释..... 190

10.3 Multiple Response Tables 过程..... 191

10.3.1 引例..... 192

10.3.2 界面说明..... 192

10.3.3 结果解释..... 193

10.4 Tables of Frequencies 过程..... 194

10.4.1 引例..... 194

10.4.2 界面说明..... 194

10.4.3 结果解释..... 196

第三部分 基础统计分析方法

第 11 章 描述性统计分析

——Descriptive Statistics
菜单详解..... 198

11.1 Frequencies 过程..... 199

11.1.1 引例..... 199

11.1.2 界面说明..... 200

11.1.3 结果解释..... 202

11.2 Descriptives 过程..... 203

11.2.1 界面说明..... 203

11.2.2 结果解释..... 204

11.3 Explore 过程..... 204

11.3.1 界面说明..... 204

11.3.2 结果解释206

11.3.3 对引例的进一步分析208

11.4 Ratio 过程209

11.4.1 引例与界面说明209

11.4.2 结果解释211

11.4.3 对引例的进一步分析211

11.5 综合分析实例212

11.5.1 频数表数据212

11.5.2 偏态分布数据的参考值范围212

11.5.3 标准正态变换213

11.5.4 探索性分析214

第 12 章 均数间的比较

——Compare Means 菜单详解... 218

12.1 Means 过程219

12.1.1 引例219

12.1.2 界面说明219

12.1.3 结果解释220

12.1.4 对引例的进一步分析222

12.2 One-Samples T Test 过程223

12.2.1 引例224

12.2.2 界面说明224

12.2.3 结果解释225

12.3 Independent-Samples T Test 过程225

12.3.1 引例225

12.3.2 界面说明226

12.3.3 结果解释227

12.4 Paired-Samples T Test 过程228

12.4.1 引例228

12.4.2 界面说明229

12.4.3 结果解释229

12.5 One-Way ANOVA 过程230

12.5.1 引例231

12.5.2 界面说明232

12.5.3 结果解释235

12.5.4 进一步分析的结果236

12.6 综合分析实例237

12.6.1 两样本 t 检验:
巧用 Cutpoint237

12.6.2 方差分析: 指定

均数的比较237

12.6.3 方差分析: 均数间曲线
趋势的判断239

第 13 章 非参数统计分析

——Nonparametric Tests

菜单详解242

13.1 分布类型检验方法243

13.1.1 Chi-Square 过程243

13.1.2 Binomial 过程245

13.1.3 Runs 过程247

13.1.4 1-Sample K-S 过程249

13.2 分布位置检验方法251

13.2.1 2 Independent Samples 与
K Independent Samples
过程251

13.2.2 2 Related Samples
与 K Related Samples 过程254

13.2.3 非参数检验中的一些问题256

13.3 秩变换分析方法257

13.3.1 原理简介257

13.3.2 应用实例257

第 14 章 卡方世界

——Crosstabs 过程详解259

14.1 分类资料数据录入格式简介260

14.2 Crosstabs 过程详解260

14.2.1 引例260

14.2.2 界面说明261

14.2.3 结果解释265

14.3 综合分析实例266

14.3.1 确切概率的计算266

14.3.2 配对卡方与一致性检验267

14.3.3 分层卡方检验269

第 15 章 相关分析

——Correlate 菜单详解271

15.1 Bivariate 过程272

15.1.1 引例272

15.1.2 界面说明272

15.1.3 结果解释273

15.2 Partial 过程274

15.2.1 引例	275	16.1.3 引例	282
15.2.2 界面说明	275	16.2 Define Sets 过程	282
15.2.3 结果解释	276	16.3 Frequencies 过程	283
15.3 Distances 过程	277	16.3.1 操作与界面说明	283
15.3.1 引例与界面说明	277	16.3.2 结果解释	284
15.3.2 结果解释	279	16.4 Crosstabs 过程	285
第 16 章 多选题分析		16.4.1 操作与界面说明	285
— Multiple Response		16.4.2 结果解释	286
菜单详解	280	附录 1 SPSS 11.0 新增功能介绍	288
16.1 多选题概述	280	附录 2 SPSS 11.0 常用函数一览表	291
16.1.1 基本概念	280	附录 3 SPSS 公司部分软件介绍	297
16.1.2 SPSS 中相应模块简介	281	参考文献	302

第一部分 基本操作入门

- 第 1 章 初探 SPSS——数据分析实例
- 第 2 章 数据编辑窗口用法详解
- 第 3 章 数据文件的进一步整理
——Transform/Data 菜单详解
- 第 4 章 SPSS 编程操作入门
- 第 5 章 SPSS 结果窗口用法详解
- 第 6 章 自学 SPSS——Help 菜单详解

第1章 初探 SPSS——数据分析实例

学习游泳有两种策略：第一种是先正规而系统的学习理论知识，包括流体力学、运动生理学、运动心理学等二十余门课程，等每一根汗毛会如何影响泳姿都搞清楚后才开始下水练习；第二种则是不管那么多，先下水试试，连呛带狗刨的折腾几次，然后再理论和实践并进。显然，我们都更喜欢采用后者。

——张文彤

欢迎加入 SPSS 使用者的行列，首先祝贺你选择了几大权威统计软件中界面最为友好，使用最为方便的 SPSS 来完成自己的工作。由于该软件极为易学易用（当然还至少要有不太高的英语水平），我们准备在章节安排上做一个新的尝试，即在第一章中不急于介绍它的界面，而是先为初学者准备了一个数据分析实例：当你将这个例题做完，SPSS 的基本使用方法也就已经被掌握了。从下一章开始，我们再详细介绍 SPSS 各个模块的精确用法。

我们的教材是以 SPSS 11.0 版为准编写的——什么？你还在用 7.0 版！还好，虽然 SPSS 从 10.0 版起在数据管理的界面操作上和以前版本有较大区别，但统计分析的基本操作是完全一样的，学习的时候应当不会有太大的障碍。好了，说了这么多废话，等急了吧，就让我们开始吧！



和 10.0 版相比，SPSS 11.0 在操作界面上基本没有变化，只是增加了一些功能和统计方法，因此本书的绝大多数内容对 10.0 版用户（包括 10.01~10.15 版）是完全通用的，对于不同之处，书中都会在相应处专门注明。

例 1.1 某克山病区测得 11 例克山病患者与 13 名健康人的血磷值(mmol/L)如下，问该地急性克山病患者与健康人的血磷值是否不同？

患者：0.84 1.05 1.20 1.20 1.39 1.53 1.67 1.80 1.87 2.07 2.11

健康人：0.54 0.64 0.64 0.75 0.76 0.81 1.16 1.20 1.34 1.35 1.48 1.56 1.87

让我们把要做的事情理理：首先要做的肯定是打开计算机（废话），然后进入 Win98 或 Win2000，在进入 SPSS 后，具体工作流程如下：

1. 将数据输入 SPSS，并及时存盘以防断电。
2. 根据设计初步判断可能用到的检验方法，然后进行必要的预分析（分布图、均数和标准差的描述等），以确定数据是否满足方法要求，最终选择一种合适的检验方法。
3. 按上一步的结论进行统计分析。
4. 保存和导出分析结果。

下面就按这几步依次讲解。

1.1 数据的输入和保存

1.1.1 SPSS 的界面

当打开 SPSS 后, 展现在我们面前的界面如图 1.1 所示。

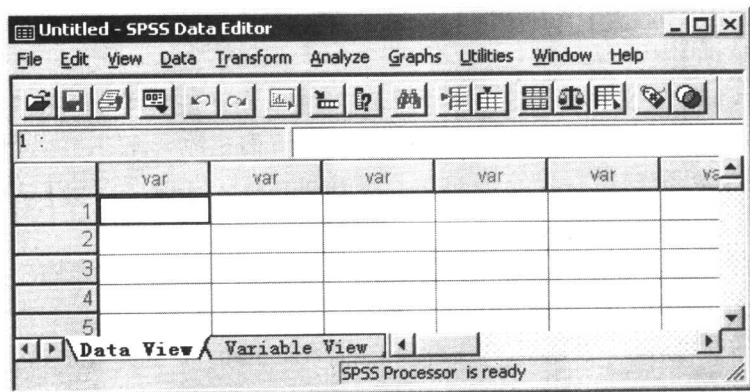



图 1.1 SPSS 的数据编辑窗口

请注意窗口顶部显示为“SPSS Data Editor”，表明现在所看到的是 SPSS 的数据编辑窗口。这是一个典型的 Windows 软件界面，有菜单栏、工具栏。特别的，工具栏下方的是数据栏，数据栏下方则是数据编辑窗口的主界面。该界面和 EXCEL 极为相似，由若干行和列组成，每行对应了一条记录，每列则对应了一个变量。由于现在我们没有输入任何数据，所以行、列的标号都是灰色的。请注意第一行第一列的单元格边框为深色，表明该数据单元格为当前单元格。


 初次进入 SPSS 系统时会出现一个导航对话框，请单击右下方的 Cancel 按钮，即可进入上面的主界面。

1.1.2 定义变量

【统计软件中数据的录入格式】

在统计软件中数据的录入格式和我们平时记录数据用的格式不太相同，大致上原则如下：

1. 不同观察对象的数据不能在同一条记录中出现，即同一观察对象的数据应当独占一行。
2. 每一个测量指标/影响因素只能占据一列的位置，即同一个指标的测量数值都应当录入到同一个变量中去。
3. 最终的数据集应当能够包含原始数据的所有信息。

 以上只是基本原则，第一条一般不会被违反，但有时候所用的分析方法会对数据格式有特别的要求，此时可能表面上会违反第二条规律。这种情况在配对数据和重复测量数据中最多见。在这些情况下，对同一个观察对象某个观

察指标的不同次测量被看成是不同的指标，因此被录入成了不同的变量。

比如本例，健康人和克山病人的血磷值虽然在不同的组中，但它们都是血磷值，因此在录入时就应当作为一个变量录入。也就是说，我们需要建立两个变量，一个变量代表血磷值，习惯上取名为 X，另一个变量代表观察对象是健康人还是克山病人，习惯上取名为 GROUP。

【在变量视图中定义变量】

现在让我们来定义变量：注意到了图 1.1 中窗口左下角的两个标签了吗？最左侧的“Data View”现在以白色显示，表明我们当前的界面是数据视图，不过我们要定义变量的话就应当切换到紧挨在它右侧的“Variable View”视图里去。单击“Variable View”标签，看到了吗，程序切换到变量视图窗口，如图 1.2 所示。

变量视图的操作界面和 FoxPro 等数据库非常相似，熟悉数据库操作的朋友应当对它感到十分亲切，这个视图的每一行代表了对一个变量的定义，每一列则代表定义该变量时用到的某种属性，如大家现在看到的名称(Name)、变量类型(Type)等，现在，请在第一、二行的 Name 列中分别输入变量名 X 和 Group，大家同时可以看到 SPSS 会在变量类型等列自动填入默认值，就用这些默认值挺好的，我们无需再多事。因此，变量定义到此就结束了。



有没有搞错?! 折腾了半天就起个名字! 其他的就都不管了? 是这样的，在绝大多数情况下，SPSS 给出的默认数据类型和数据精度完全可以满足需要，只是不太好看而已。至于标签等比较花哨的选项，现在我们才刚刚入门，一切从简。以后会详细介绍各种设置的用法。

让我们切换回数据视图，此时数据编辑窗口如图 1.3 所示。现在，第一、第二列的名称均为深色显示，并且就是我们刚才定义的 X 和 Group，表明这两列已经被定义为变量，其余各列的名称仍为灰色的“var”，表示尚未使用。同样地，各行的标号也为灰色，表明现在还未输入过数据，即该数据集内没有记录。

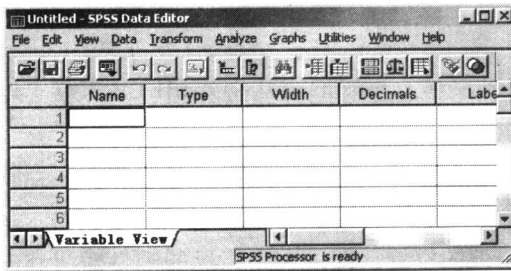


图 1.2 SPSS 的变量视图窗口

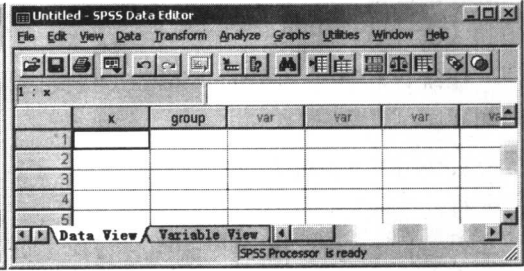


图 1.3 变量定义完毕的数据视图窗口

1.1.3 输入数据

我们先来输入变量 X 的值，请确认一行一列单元格为当前单元格，弃鼠标而用键盘，输入第一个数据 0.84，此时界面显示如图 1.4 所示。

请注意：在回车之前，输入的数据在数据单元格内左对齐显示，表示该单元格为第一次录入数据，同时数据栏内同步显示出你输入的数值。现在回车，界面如图 1.5 所示。首先，当前单元格下移，变成了二行一列单元格，而一行二列单元格的内容则被替换成了 0.84；其次，第一行的标号变黑，表明该行已输入了数据；第三，一行二列单元格因为没有输入过数据，显示为“.”，这代表该数据为缺失值。用类似的输入方式，我们将患者的血磷值输入完毕，并将相应的变量 GROUP 均取值为 1，此时数据编辑窗口如图 1.6 所示。

	x	group	var
1	0.84		
2			
3			

图 1.4 输入数据时的状态

	x	group	var
1	.84		
2			
3			

图 1.5 回车后的状态

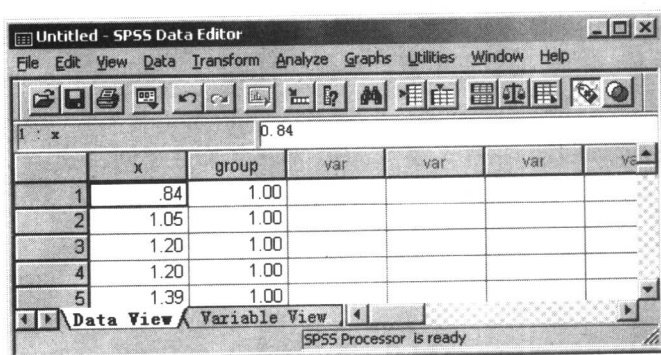


图 1.6 数据录入完毕的窗口

从第 12 行开始输入健康人的数据，并将相应的 GROUP 变量取值为 2。最终该数据集应该有 24 条记录。

1.1.4 保存数据

选择菜单 File→Save，由于该数据从来没有被保存过，所以弹出 Save Data As 对话框如图 1.7 所示。注意在 11.0 版中该对话框在中下部比以前多出了一个 Variables 钮，可以用于选择是保存数据集中的某些变量，这里我们暂时不管他，就使用默认的“Keeping 2 of 2 variables”。最下方有一个保存类型列表框，单击后可以看到 SPSS 中数据可以保存的各种数据类型，有 DBF、FoxPro、Excel、Access 等，这里我们仍然将其存为 SPSS 自己的数据格式 (*.sav 文件)。在文件名框内键入 xuelin 并回车，可以看到系统回到数据编辑窗口，其左上角已经由 Untitled 变为了现在的 xuelin，表明当前编辑文件的名称为 xuelin.sav。

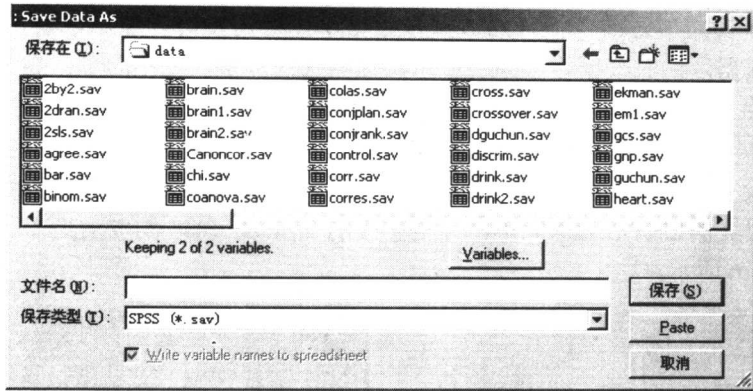


图 1.7 Save Data As 对话框

1.2 数据的预分析

现在来把我们掌握的情况理清清楚：该数据是定量资料，设计是两样本均数的比较，针对这种目的可用的检验方法有 t 检验、u 检验以及秩和检验。由于样本量太小，u 检验不能考虑；秩和检验的效能偏低，是最后的武器；现在最合适的就是无所不在的 t 检验了。它相应的假设如下：

$$H_0: \text{两总体均数相同, } \mu_1 = \mu_2 \quad H_1: \text{两总体均数不同, } \mu_1 \neq \mu_2$$

但是两样本 t 检验对数据是有要求的，在小样本时它要求分布不太偏，方差也得齐。该数据是否符合它的要求呢？还是做个描述看看吧！

1.2.1 数据的简单描述

首先我们需要知道数据的基本情况，如均数、标准差等。选择菜单项 Analyze → Descriptive Statistics → Descriptives，系统会弹出描述对话框如图 1.8 所示。

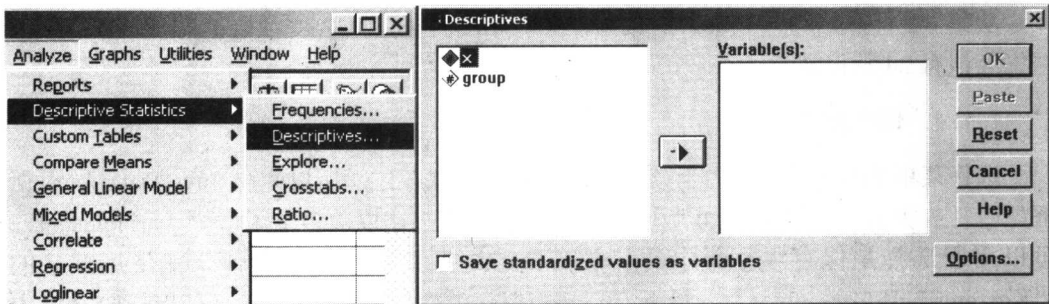



图 1.8 菜单操作示意及弹出的 Descriptives 主对话框

该对话框可分为左右两大部分，左侧为所有可用的候选变量列表，右侧为选入变量列表。我们只需要描述 X，用鼠标选中 X，单击中间的 ，变量 X 的标签就会移入右侧，注意这时 OK 按钮变黑，表明已经可以进行分析了，单击它，系统会弹出一个新的界面如图 1.9 所示。

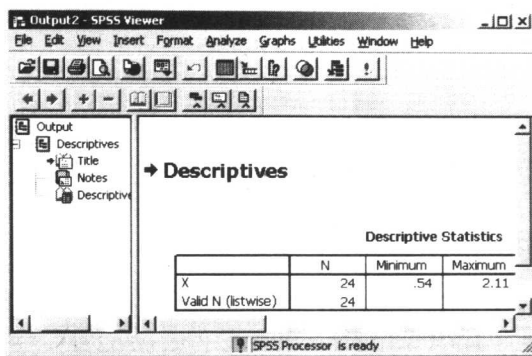


图 1.9 SPSS 的结果浏览窗口

2 如果按 SPSS 标准的叫法，这里应该是调用了 Descriptives 过程，为了避免太生硬，我们暂时称为调用对话框，等大家熟悉了 SPSS 以后，在统计分析各章中可能两种称呼会混用。

该窗口上方的名称为 SPSS Viewer，即（结果）浏览窗口，整个的结构和资源管理器类似，左侧为导航栏，右侧为具体的输出结果。结果表格给出了样本数、最小值、最大值、均数和标准差这几个常用的统计量。从中可以看到，24 个数据总的均数为 1.2846 mmol/L，标准差为 0.46866 mmol/L（SPSS 10 中显示为 0.4687，精度少一位）。

我们以上的做法对吗？当然有问题！光看总的描述是不够的，还应当看看分组的描述情况，不然怎么知道方差相差的大概有多远呢？这里要用到文件拆分功能，请切换回数据编辑窗口，选择菜单 **Data**→**Split File**，系统弹出文件拆分对话框如图 1.10 所示。

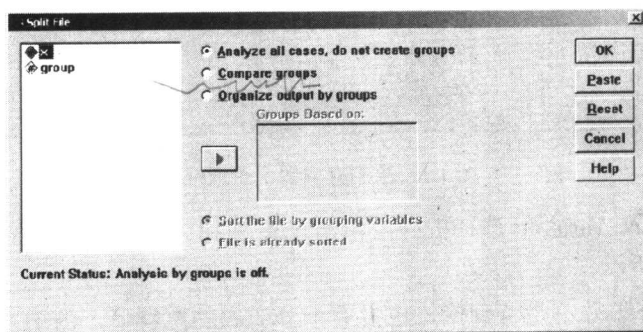


图 1.10 Split File 过程主对话框

选择 **Compare groups** 单选框，将变量 **GROUP** 选入右侧的 **Group Based on** 框，单击 **OK** 按钮，此时界面似乎没有任何改变，但请注意数据窗口状态栏的右侧，可以看到有黑色的 **Split File On** 出现，表明数据文件正处于拆分状态。现在再做一次刚才的数据描述，就会发现输出结果发生了改变如下：