

普.通.高.等.学.校
计算机教育“十二五”规划教材

SPSS 统计分析 实用教程

(第2版)

SPSS GUIDE TO DATA ANALYSIS
(2nd edition)

谢蕾蕾 宋志刚 何旭洪 ◆ 主编



普.通.高.等.学.校
计算机教育“十二五”规划教材

SPSS 统计分析 实用教程

(第2版)

SPSS GUIDE TO DATA ANALYSIS
(2nd edition)

谢蕾蕾 宋志刚 何旭洪 ◆ 主编

人 民 邮 电 出 版 社
北 京

图书在版编目 (C I P) 数据

SPSS统计分析实用教程 / 谢蕾蕾, 宋志刚, 何旭洪
主编. -- 2版. -- 北京: 人民邮电出版社, 2013. 2
普通高等学校计算机教育“十二五”规划教材
ISBN 978-7-115-30221-2

I. ①S… II. ①谢… ②宋… ③何… III. ①统计分
析—软件包—高等学校—教材 IV. ①C819

中国版本图书馆CIP数据核字(2012)第308623号

内 容 提 要

SPSS 是目前应用最广泛的统计软件之一, 其在很多领域深受用户的好评。本书以 SPSS 16.0 为基础, 详细介绍了多种常用统计分析方法的原理和实现技术。全书分 11 章, 主要内容包括 SPSS 简介、变量、数据文件、系统参数、统计描述、统计图制作、均值比较和 T 检验、方差分析、相关分析、回归分析、聚类分析和判别分析、因子分析, 以及非参数检验等的原理, 并结合实际问题详细介绍使用 SPSS 解决这些问题的步骤和结果讨论。

本书统计理论严谨, 文字浅显易懂, 并配有大量的图表进行解释, 可作为计算机、金融、教育、统计等专业 SPSS 相关课程的教材, 同时也可供具备一定统计分析基础知识的读者和各行行业中非统计专业而需要使用统计方法的读者学习使用。



普通高等学校计算机教育“十二五”规划教材

SPSS 统计分析实用教程 (第 2 版)

- ◆ 主 编 谢蕾蕾 宋志刚 何旭洪
责任编辑 武恩玉
执行编辑 王 伟
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京昌平百善印刷厂印刷
- ◆ 开本: 787×1092 1/16
印张: 18.75 2013 年 2 月第 2 版
字数: 492 千字 2013 年 2 月北京第 1 次印刷

ISBN 978-7-115-30221-2

定价: 39.80 元

读者服务热线: (010)67170985 印装质量热线: (010)67129223
反盗版热线: (010)67171154

前 言

SPSS (Statistical Product and Service Solution, 统计产品与服务解决方案) 是全球专业统计分析软件的领导者之一, 其在电信、医疗、银行、证券、保险、制造、商业、科研、教育等行业都得到了广泛的应用。只要具有一定 Windows 操作技能, 并对统计分析原理稍有了解, 都可以通过学习本书, 满足利用 SPSS 进行一定科研工作的需求。继 SPSS 13.0 之后, SPSS 公司推出了 SPSS 16.0 版本。该版本在原版本的基础上不断改善, 页面更加人性化, 分析方法更加充实。本书即是以日常统计中经常用到的统计分析功能为目标, 结合具体实例来介绍 SPSS 的应用。

本书第 1 版自出版以来, 受到广大读者的喜爱。本书是在对第 1 版教材进行修订补充的基础上编写而成的。为了使本教材的系统性、实用性更强, 增加了统计图制作的内容, 对部分章节内容进行了调整, 修订了第 1 版的错误。相对于第 1 版教材, 第 2 版教材各章结构更加合理, 内容更加完整。

本书分为 11 章。第 1 章是对 SPSS 软件的简要介绍, 第 2 章是有关数据、文件的操作, 这两章是应用 SPSS 统计分析软件的基础; 第 3 章介绍了基本的描述统计分析的操作方法; 第 4 章介绍了常用的统计图形的制作方法; 第 5~11 章是对常用的统计分析功能的 SPSS 操作过程进行介绍。

该书各章节划分清晰, 首先对统计分析功能的原理进行介绍, 然后通过实例讲解 SPSS 操作过程, 最后是分析结果的解释。这样安排不仅使读者对统计分析功能的原理能有全面的理解, 并且方便读者掌握实际操作过程, 能够对分析结果进行判断和说明。

本书由谢蕾蕾(华北水利水电学院)、宋志刚、何旭洪主编, 其中谢蕾蕾编写第 3~11 章, 宋志刚编写第 1、2 章, 何旭洪负责统稿。同时, 全国多所高校的统计教师和统计专业人士参与了本书的编写, 在此, 编者表示衷心的感谢。

由于 SPSS 软件的统计分析功能非常强大, 本书只是对一些基本的统计分析方法进行介绍, 不可能涵盖所有的应用领域。

限于编者水平, 书中难免出现一些不足之处, 希望广大读者提出宝贵意见。

编者

2013 年 1 月

目 录

第 1 章 SPSS 简介1	
1.1 SPSS 的发展.....1	
1.2 SPSS 的版本.....2	
1.3 SPSS 的运行方式.....2	
1.4 SPSS 的启动、主界面和退出.....2	
1.4.1 启动 SPSS.....2	
1.4.2 SPSS 的数据编辑窗口.....3	
1.4.3 SPSS 结果输出窗口.....4	
1.4.4 退出 SPSS.....5	
1.5 SPSS 的帮助系统.....5	
1.5.1 Help 菜单中的 topics 命令.....5	
1.5.2 Help 菜单中的 tutorial 命令.....5	
1.5.3 各种对话框中的 Help 按钮.....5	
1.5.4 有关统计方法指导.....6	
小结.....6	
思考与练习.....6	
第 2 章 变量、数据文件、 系统参数7	
2.1 定义变量.....7	
2.1.1 变量的定义信息.....8	
2.1.2 变量定义信息的复制.....10	
2.2 数据的输入与保存.....10	
2.2.1 输入数据的一般方法.....10	
2.2.2 输入带有变量值标签的数据.....11	
2.2.3 SPSS 数据文件的保存.....11	
2.3 数据的编辑.....12	
2.3.1 单元值的修改.....12	
2.3.2 增加和删除一个个案.....13	
2.3.3 数据的排序.....13	
2.3.4 数据的行列互换.....13	
2.3.5 选取个案子集.....13	
2.3.6 数据分类汇总.....14	
2.3.7 缺失值的替代.....16	
2.3.8 数据次序确定.....17	
2.4 变量的操作.....18	
2.4.1 增加和删除一个变量.....18	
2.4.2 指定加权变量.....19	
2.4.3 根据已存在的变量建立新变量.....19	
2.4.4 产生计数变量.....20	
2.4.5 变量的重新赋值.....22	
2.4.6 变量的自动赋值.....23	
2.4.7 变量定义信息的查询.....24	
2.4.8 变量集的定义和使用.....25	
2.5 数据文件的合并和分组.....26	
2.5.1 数据文件的纵向合并.....26	
2.5.2 数据文件的横向合并.....27	
2.5.3 数据文件的分组.....28	
2.6 读入其他格式文件数据.....29	
2.6.1 读取固定格式的文本文件.....29	
2.6.2 读取自由格式的文本文件.....31	
2.6.3 读取 dBASE 软件文件.....33	
2.6.4 读取 Excel 软件文件.....34	
2.6.5 读取数据库文件.....34	
2.7 SPSS 运行环境设置.....38	
2.7.1 SPSS 状态栏的显示和隐藏.....38	
2.7.2 SPSS 网格线的显示和隐藏.....39	
2.7.3 SPSS 菜单的增加和删除.....39	
2.7.4 SPSS 字体的设置.....40	
小结.....41	
思考与练习.....41	
第 3 章 统计描述42	
3.1 基本描述统计分析.....42	
3.1.1 均值和均值标准误差.....42	
3.1.2 中位数.....42	
3.1.3 众数.....43	
3.1.4 全距.....43	
3.1.5 方差和标准差.....43	

3.1.6 四分位数、十分位数和百分位数	43	4.2.3 单个变量的多线图	96
3.1.7 峰度	44	4.3 饼图	97
3.1.8 偏度	44	4.3.1 定义和类型	97
3.1.9 SPSS 中实现过程	44	4.3.2 个案分组的饼图	97
3.2 频数	49	小结	98
3.2.1 统计学上的定义和计算公式	49	思考与练习	99
3.2.2 SPSS 中的实现过程	49	第 5 章 均值比较和 T 检验	100
3.2.3 结果和讨论	50	5.1 Means 过程	100
3.3 标准化 Z 分数及其线性转换	51	5.1.1 统计学上的定义和计算公式	100
3.3.1 统计学上的定义和计算公式	51	5.1.2 SPSS 中实现过程	100
3.3.2 SPSS 中实现过程	51	5.1.3 结果和讨论	102
3.3.3 结果和讨论	53	5.2 单样本 T 检验	102
3.4 探索分析	54	5.2.1 统计学上的定义和计算公式	102
3.4.1 统计学上的定义和计算公式	54	5.2.2 SPSS 中实现过程	103
3.4.2 SPSS 中的实现过程	54	5.2.3 结果和讨论	103
3.4.3 结果和讨论	58	5.3 两独立样本 T 检验	104
3.5 交叉列联表分析	62	5.3.1 统计学上的定义和计算公式	104
3.5.1 统计学上的定义和计算公式	62	5.3.2 SPSS 中实现过程	105
3.5.2 SPSS 中实现过程	63	5.3.3 结果和讨论	106
3.5.3 结果和讨论	66	5.4 两配对样本 T 检验	107
3.6 多选项分析	68	5.4.1 统计学上的定义和计算公式	107
3.6.1 统计学上的定义和计算公式	68	5.4.2 SPSS 中实现过程	107
3.6.2 SPSS 中的实现过程	69	5.4.3 结果和讨论	109
3.6.3 结果和讨论	74	小结	109
3.7 基本统计分析的报表制作	75	思考与练习	110
3.7.1 报表分类	75	第 6 章 方差分析	111
3.7.2 SPSS 中实现过程	75	6.1 方差分析的基本概念	111
3.7.3 结果和讨论	82	6.2 单因素方差分析	112
小结	83	6.2.1 统计学上的定义和计算公式	112
思考与练习	83	6.2.2 SPSS 中实现过程	113
第 4 章 统计图形	85	6.2.3 结果和讨论	115
4.1 条形图	85	6.3 多因素方差分析	117
4.1.1 定义和类型	85	6.3.1 统计学上的定义和计算公式	117
4.1.2 个案分组的简单条形图	86	6.3.2 SPSS 中实现过程	118
4.1.3 单个变量的简单条形图	91	6.3.3 结果和讨论	122
4.1.4 个案取值的简单条形图	92	6.4 协方差分析	125
4.2 线图	93	6.4.1 统计学上的定义和计算公式	125
4.2.1 定义和类型	93	6.4.2 SPSS 中实现过程	126
4.2.2 个案分组的单线图	94		

6.4.3 结果和讨论	127	8.5 曲线估计	178
小结	128	8.5.1 统计学上的定义和计算公式	178
思考与练习	128	8.5.2 SPSS 中实现过程	178
第 7 章 相关分析	130	8.5.3 结果和讨论	180
7.1 相关分析的基本概念	130	8.6 时间序列的曲线估计	184
7.2 二元定距变量的相关分析	131	8.6.1 统计学上的定义和计算公式	184
7.2.1 统计学上的定义和计算公式	131	8.6.2 SPSS 中实现过程	184
7.2.2 SPSS 中实现过程	132	8.6.3 结果和讨论	186
7.2.3 结果和讨论	134	8.7 含虚拟自变量的回归分析	188
7.2.4 绘制相关散点图	134	8.7.1 统计学上的定义和计算公式	188
7.3 二元定序变量的相关分析	136	8.7.2 SPSS 中实现过程	189
7.3.1 统计学上的定义和计算公式	136	8.7.3 结果和讨论	190
7.3.2 SPSS 中实现过程	136	8.8 逻辑回归分析	191
7.3.3 结果和讨论	137	8.8.1 统计学上的定义和计算公式	191
7.4 偏相关分析	138	8.8.2 SPSS 中实现过程	195
7.4.1 统计学上的定义和计算公式	138	8.8.3 结果和讨论	198
7.4.2 SPSS 中实现过程	139	小结	201
7.4.3 结果和讨论	140	思考与练习	202
7.5 距离相关分析	141	第 9 章 聚类分析与判别分析	204
7.5.1 统计学上的定义和计算公式	141	9.1 聚类分析与判别分析的基本概念	204
7.5.2 SPSS 中实现过程	142	9.2 层次聚类分析中的 Q 型聚类	205
7.5.3 结果和讨论	148	9.2.1 统计学上的定义和计算公式	205
小结	150	9.2.2 SPSS 中实现过程	208
思考与练习	150	9.2.3 结果和讨论	212
第 8 章 回归分析	152	9.3 层次聚类分析中的 R 型聚类	215
8.1 回归分析的基本概念	152	9.3.1 统计学上的定义和 计算公式	215
8.2 一元线性回归分析	153	9.3.2 SPSS 中实现过程	215
8.2.1 统计学上的定义和计算公式	153	9.3.3 结果和讨论	217
8.2.2 SPSS 中实现过程	157	9.4 快速聚类分析	218
8.2.3 结果和讨论	163	9.4.1 统计学上的定义和计算公式	218
8.3 多元线性回归分析	164	9.4.2 SPSS 中实现过程	219
8.3.1 统计学上的定义和计算公式	164	9.4.3 结果和讨论	222
8.3.2 SPSS 中实现过程	168	9.5 判别分析	225
8.3.3 结果和讨论	170	9.5.1 统计学上的定义和计算公式	225
8.4 非线性回归分析	173	9.5.2 SPSS 中实现过程	226
8.4.1 统计学上的定义和计算公式	173	9.5.3 结果和讨论	230
8.4.2 SPSS 中实现过程	174	小结	235
8.4.3 结果和讨论	177	思考与练习	236

第 10 章 因子分析	237	11.3 SPSS 单样本变量值随机性检验	262
10.1 因子分析的定义和数学模型	237	11.3.1 统计学上的定义和计算公式	262
10.1.1 统计学上的定义	237	11.3.2 SPSS 中实现过程	263
10.1.2 数学模型	238	11.3.3 结果和讨论	265
10.1.3 因子分析的 4 个基本步骤	239	11.4 SPSS 单样本 K-S 检验	265
10.1.4 确定待分析的原有若干变量 是否适合于因子分析	239	11.4.1 统计学上的定义和计算公式	265
10.1.5 构造因子变量	240	11.4.2 SPSS 中实现过程	266
10.1.6 因子变量的命名解释	242	11.4.3 结果和讨论	267
10.1.7 计算因子得分	242	11.5 两独立样本非参数检验	268
10.2 SPSS 中实现过程	242	11.5.1 统计学上的定义和计算公式	268
10.2.1 SPSS 中实现步骤	242	11.5.2 SPSS 中实现过程	270
10.2.2 SPSS 结果解释	247	11.5.3 结果和讨论	272
10.2.3 讨论	253	11.6 多独立样本非参数检验	274
小结	253	11.6.1 统计学上的定义和计算公式	274
思考与练习	254	11.6.2 SPSS 中实现过程	276
第 11 章 非参数检验	255	11.6.3 结果和讨论	277
11.1 总体分布的卡方检验	255	11.7 两配对样本非参数检验	278
11.1.1 统计学上的定义和计算公式	255	11.7.1 统计学上的定义和计算公式	278
11.1.2 SPSS 中实现过程	256	11.7.2 SPSS 中实现过程	280
11.1.3 结果和讨论	259	11.7.3 结果和讨论	283
11.2 二项分布检验	259	11.8 多配对样本非参数检验	284
11.2.1 统计学上的定义和计算公式	259	11.8.1 统计学上的定义和计算公式	284
11.2.2 SPSS 中实现过程	260	11.8.2 SPSS 中实现过程	286
11.2.3 结果和讨论	262	11.8.3 结果和讨论	289
		小结	291
		思考与练习	291

第 1 章

SPSS 简介

社会科学统计软件包 (Statistical Package for the Social Science, SPSS) 是世界著名的统计分析软件之一。经过近 40 年的发展, SPSS 在全球已拥有大量的用户。目前, SPSS 使用 Windows 的窗口方式展示各种管理和分析数据的方法, 可方便地用于特定的科研统计。本章概要介绍 SPSS 的发展背景、主要版本、运行方式、启动与退出以及帮助系统的使用。

1.1 SPSS 的发展

1968 年, 3 位美国斯坦福大学的学生开发了最早的 SPSS 系统, 并基于这一系统于 1975 年在芝加哥合作成立了 SPSS 公司。20 世纪 80 年代以前, SPSS 主要应用于企事业单位。1984 年, SPSS 总部推出了世界上第一个统计分析软件微机版本 SPSS/PC+, 开创了 SPSS 计算机系列产品的开发方向, 从而确立了该软件在个人用户市场第一的地位。迄今为止, SPSS 软件已有 40 余年历史, 拥有约 25 万用户, 它们分布于通信、医疗、银行、证券、保险、制造、商业、市场研究、科研教育等多个领域和行业, 是世界上应用最广泛的专业统计软件。

SPSS 的基本功能包括数据管理、统计分析、图表分析、输出管理等。其过程包括描述性统计、均值比较、一般线性模型、相关分析、回归分析、对数线性模型、聚类分析、数据简化、时间序列分析、因子分析等大类, 每类中又分为几个统计过程。如回归分析中包括线性回归分析、曲线估计、Logistic 回归等几个统计过程, 并且每个过程中又允许用户选择不同的方法及参数。SPSS 中还有专门的绘图系统, 可以根据数据绘制各种图形。

SPSS 名为社会学统计软件包, 这是为了强调其社会科学应用的一面 (因为社会科学研究中的许多现象都是随机的, 要使用统计学和概率论的定理来进行研究)。而实际上它在社会科学、自然科学的各个领域都能发挥巨大的作用。

1994~1998 年, SPSS 公司陆续并购了 SYSTAT 公司、BMDP 软件公司、Quantime 公司、ISL 公司等, 并将各公司的主打产品收纳 SPSS 旗下, 从而使 SPSS 公司由原来单一统计产品的开发与销售, 转向为企业、教育科研及政府机构提供全面信息统计决策支持服务, 成为走在了最新流行的“数据仓库”和“数据挖掘”领域前沿的一家综合统计软件公司。最近, 伴随着 SPSS 产品服务领域的扩大和服务深度的增加, SPSS 公司已决定将它的英文全称更改为 Statistical Product and Service Solutions, 意为“统计产品与服务解决方案”。

1.2 SPSS 的版本

到目前为止, SPSS 已成为适合 DOS、Windows、UNIX、Macintosh、OS/2 等多种操作系统使用的产品。国内常用的是其适用于 DOS 和 Windows 的版本。

SPSS for DOS 通常称为 SPSS/PC+, 现已较少使用。SPSS for Windows 界面友好, 功能强大, 用户越来越多。SPSS for Windows 目前最新的版本是 SPSS 18.0 等。SPSS 10.0 以上版本有两种结构: 一种结构是服务器 (Server) / 客户机 (Client) 结构, 由 SPSS Server 和 SPSS for Windows 两部分组成 (可以快速地对大规模数据进行处理); 另外一种结构是单机版本, 即 SPSS for Windows 标准版。

本书以运行于 Windows 9x/NT/2000/XP 上的 SPSS 16.0 for Windows 标准版为例, 并在本书后面的内容中简称为 SPSS。

1.3 SPSS 的运行方式

SPSS 主要有 3 种运行方式。

1. 批处理方式

把已编写好的程序 (语句程序) 存为一个文件, 在 SPSS 的 Production Facility 程序中打开运行。

2. 完全窗口菜单运行方式

这种方式通过选择窗口菜单和对话框完成各种操作。用户无须学会编程, 简单易用。本书中各个统计功能的实现都采用了这种方式。

3. 程序运行方式

这种方式是在命令 (Syntax) 窗口中直接运行编写好的程序, 或者在脚本 (Script) 窗口中运行脚本程序, 如图 1-1 所示。Syntax 窗口方式要求掌握 SPSS 的语法。

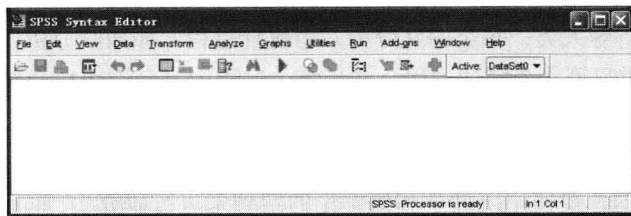


图 1-1 命令 (Syntax) 窗口

1.4 SPSS 的启动、主界面和退出

1.4.1 启动 SPSS

SPSS 安装完毕后, 系统会自动在 Windows 菜单中创建快捷方式。单击 Windows 的“开始”按钮, 在“程序”菜单项“SPSS for Windows”中选择“SPSS 16.0 for Windows”并单击, 即可启动 SPSS。

1.4.2 SPSS 的数据编辑窗口

SPSS 主界面主要有两个，一个是 SPSS 数据编辑窗口，另一个是 SPSS 输出窗口。下面先介绍数据编辑窗口。

SPSS 的数据编辑窗口与微软公司的 Excel 窗口有些相似，并且有一些功能也相同。不过 SPSS 的数据统计功能要比 Excel 强很多。

数据编辑窗口由标题栏、菜单栏、工具栏、编辑栏、变量名栏、内容区、窗口切换标签页和状态栏组成，如图 1-2 所示。

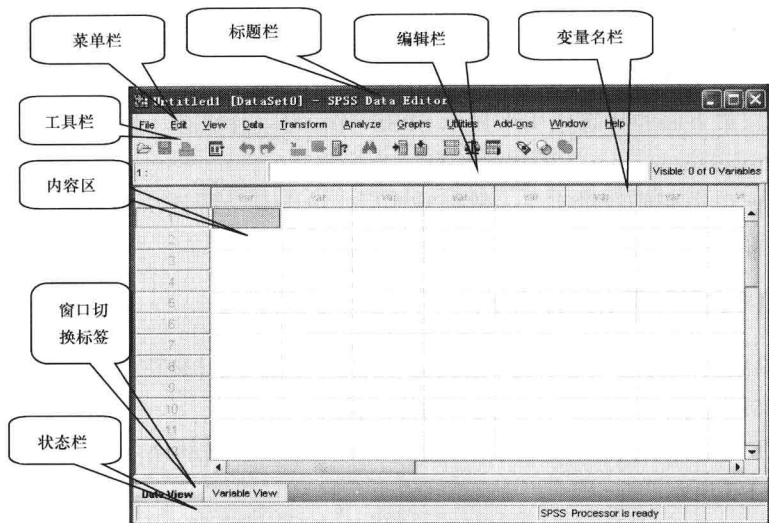


图 1-2 数据编辑窗口

标题栏中显示编辑的数据文件名。图 1-2 中所编辑的数据文件未命名，因而显示为“Untitled1”。

菜单栏中列出了 SPSS 的命令菜单，每个菜单对应一组相应的功能。其中“File”是对 SPSS 文件的操作；“Edit”是 SPSS 文件的编辑菜单；“View”是用户界面设置菜单；“Data”是数据文件的建立和编辑菜单；“Transform”是数据基本处理菜单；“Analyze”是统计分析菜单，主要统计功能都集中在该菜单中；“Graphs”是统计图形菜单；“Utilities”是相关应用和设置菜单；“Add-ons”是对各种模型的说明；“Windows”是 SPSS 各窗口切换菜单；“Help”是 SPSS 帮助菜单。

工具栏中列出了一些常用的快捷图标。

编辑栏中可以输入数据。

变量名栏中列出了该编辑文件中所含有的变量名。SPSS 自动命名变量名为 var0001, var0002 等。

内容区列出了所有个案在变量中的取值。SPSS 中每一行表示一个个案。内容区的最左边是行的标号，这和 Excel 类似。

该窗口下方有两个标签：“Data View”（数据视图）和“Variable View”（变量视图）。这两种视图提供了一种类似于电子表格的方法，用以产生和编辑 SPSS 数据文件中的变量和数据。“Data View”对应的表格用于查看、输入和修改数据，“Variable View”对应的表格用于输入和修改变量的定义。这样使用者就可以非常方便地进行变量类型的定义和数据的输入。在较早的 SPSS 版本中，变量的定义需要通过单击“Data”菜单的“Define Variable”子菜单，在弹出的“Define Variable”对话框中进行定义。

如果使用过电子表格,如 Microsoft Excel 等,那么数据编辑窗口中“Data View”所对应表格许多功能应该已经熟悉。但是它和一般的电子表格处理软件还有以下区别。

(1) 一个列对应一个变量,即每一列代表一个变量 (Variable) 或一个被观测量的特征。例如,问卷上的每一项就是一个变量。

(2) 行是观测,即每一行代表一个个体、一个观测、一个样品,在 SPSS 中称为事件 (Case)。例如,问卷上的每一个人就是一个观测。

(3) 单元包含值,即每个单元包括一个观测中的单个变量值。单元 (Cell) 是观测和变量的交叉。

(4) 数据文件是一张长方形的二维表。数据文件的范围是由观测和变量的数目决定的。可以在任一单元中输入数据。如果在定义好的数据文件边界以外输入数据,SPSS 将数据长方形延长到可包括那个单元和文件边界之间的任何行和列。

数据编辑窗口最下面的状态栏用来显示 SPSS 当前的运行状态。当 SPSS 等待用户操作时,会出现“SPSS Processor is ready”的提示信息。

在数据编辑窗口中完成变量定义、数据输入后,单击某个统计功能菜单,SPSS 会自动完成统计分析,并将弹出结果输出窗口,其中存放了数据统计的结果。

1.4.3 SPSS 结果输出窗口

SPSS 结果输出窗口名为 Viewer,它是显示和管理 SPSS 统计分析结果、报表及图形的窗口。读者可以将此窗口中的内容以结果文件.spo 的形式保存。

结果输出窗口中也有菜单栏和工具栏,它们下面是结果输出部分,然后是状态栏。标准的结果输出窗口如图 1-3 所示。

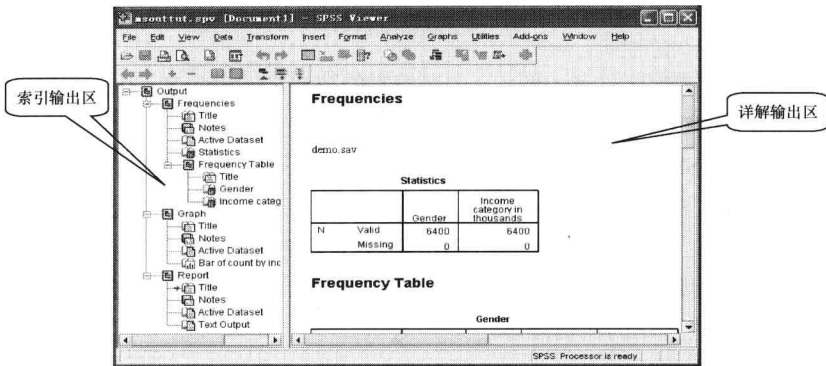


图 1-3 结果输出窗口

结果输出部分分成左右两个部分。左边部分是索引输出区,用于显示已有的分析结果标题和内容索引;右边部分是各个分析的具体结果,称为详解输出区。这和 Word 的文档结构视图十分类似。

索引输出区是详解输出区的一个视图,以简洁的方式反映出详解输出区中各个内容项,便于用户查找操作结果。它是一棵倒放的树,树根名为 Output,各内容以子树的形式向右下方排列,形成一棵完整的索引树。SPSS 默认将索引树和各个子树以展开的形式显示。在子树较多的情况下,为了方便查看和管理,可以将某些展开的子树以折叠的方式显示。只需要单击索引输出区中相应子树左边的 \square 符号,使其变为 \square 符号,就可以折叠展开的索引树,此时详解输出区该部分的内容就会隐藏起来。

可以对详解输出区中的表格进行编辑等操作。当某个表格处于选中状态时，相应表格的四周会出现黑色单线边框，通过鼠标单击来选中某个表格。如果要对选中表格进行编辑，可双击该表格，当表格四周出现黑色斜线边框时，即可对表格内数据进行修改。处于编辑状态的详解输出区菜单和正常时不同。

1.4.4 退出 SPSS

选择数据编辑窗口的“File”菜单中的“Exit”命令，或单击标题栏上的“关闭”按钮退出 SPSS。

1.5 SPSS 的帮助系统

SPSS 提供了丰富且详尽的在线帮助，主要有 4 种方式，下面分别介绍。

1.5.1 Help 菜单中的 topics 命令

在运行 SPSS 的任何时候，单击“Help”菜单中的“topics”命令，会弹出帮助主题窗口，如图 1-4 所示。在其中选择相关的命令，即可得到所需的各种帮助。

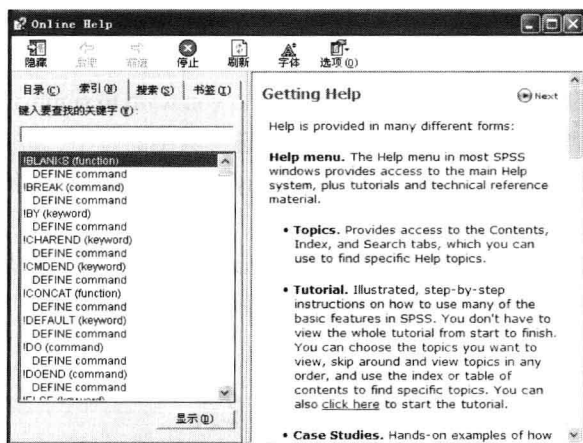


图 1-4 帮助主题窗口

该帮助主题窗口中有两个标签页。一个是“目录”标签页，在该页中可以按目录方式进行检索；另一个是“索引”标签页，在其中输入需要查询的关键字，下面的栏中则会显示出和该关键字相关的内容。

1.5.2 Help 菜单中的 tutorial 命令

选择“Help”菜单中的“tutorial”命令，则弹出浏览器形式的帮助文件窗口，如图 1-5 所示。在该窗口下面有 4 个按钮，分别对应“目录索引”、“帮助主页”、“前进”、“后退”功能。

1.5.3 各种对话框中的 Help 按钮

利用各种对话框中的“Help”按钮，可以直接获得 SPSS 相应命令的帮助，这是最简单也是最有效的获取帮助的方式。

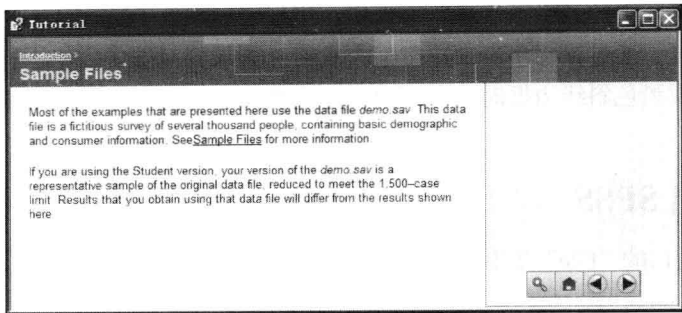


图 1-5 帮助窗口

在具体操作过程中，弹出某一个对话框时，一般总有“Help”按钮，单击该按钮，就可得到有关这一对话框选项内容的详细帮助。

1.5.4 有关统计方法指导

SPSS 对一些基本模块中的统计提供了 Statistics Coach 帮助，也就是“手把手”式的指导，步骤如下。

(1) 单击“Help”菜单中的“Statistics Coach”命令，就会弹出第 1 个对话框。该对话框用于选择要进行什么统计功能。

SPSS 一共提供了 7 个方面的统计指导。单击选择任何一个方面，进入下一步。

(2) 第 2 个对话框的标题为“What kind of data do you want to summarize?”，用于选择数据类型。SPSS 提供了 3 种类型的数据供用户选择。单击选择其中的一个，进入下一步。

(3) 第 3 个对话框的标题为“What kind of display do you want?”，用于选择结果显示的方式。单击选择其中的一个，进入下一步。

(4) 第 4 个对话框的标题为“What kind of summary do you want?”，用于选择要统计的命令，单击选择其中的一个，就会出现一个提示如何完成每步的帮助对话框。

小 结

SPSS 是在当前社会经济分析中应用最广泛的软件之一。继 SPSS 13.0 之后，SPSS 公司又推出了 SPSS 16.0，这是一个功能更加强大、设计更加人性化的版本。

SPSS 的主界面有两个，一个是 SPSS 数据编辑窗口，另一个是 SPSS 结果输出窗口。数据编辑窗口主要包括数据的编辑和统计分析；结果输出窗口主要的功能是显示和管理 SPSS 的各种统计分析的结果。

SPSS 的帮助系统可提供各种丰富详尽的在线帮助。

思考与练习

1-1 SPSS 的运行方式有几种？分别是什么？

1-2 SPSS 中“DataView”所对应的表格与一般的电子处理软件有什么区别？

第 2 章

变量、数据文件、系统参数

SPSS 对数据的处理是以变量为前提的，因此本章首先介绍定义变量、输入数据，再介绍保存数据、操作数据文件，最后介绍 SPSS 运行环境和设置系统参数。

2.1 定义变量

启动 SPSS 后，出现如图 2-1 所示数据编辑窗口。由于目前还没有输入数据，因此显示的是一个空文件。

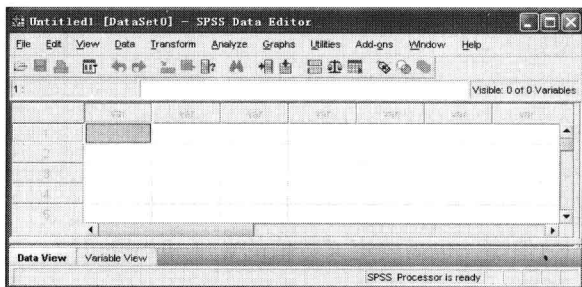


图 2-1 数据编辑窗口

输入数据前首先要定义变量。定义变量即要定义变量名、变量类型、变量长度（小数位数）、变量标签（或值标签）和变量的格式。

单击数据编辑窗口左下方的“Variable View”标签或双击列的题头（Var），进入如图 2-2 所示的变量定义视图窗口，在此窗口中即可定义变量。

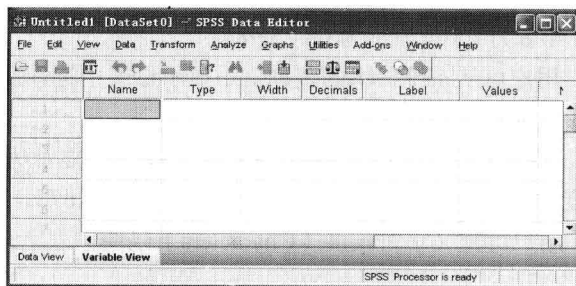


图 2-2 变量定义视图窗口

2.1.1 变量的定义信息

在图 2-2 所示的窗口中, 每一行表示一个变量的定义信息, 包括 Name、Type、Width、Decimal、Label、Values、Missing、Columns、Align、Measure 等。

1. 定义变量名

SPSS 默认的变量为 Var00001、Var00002 等, 用户也可以根据自己的需要来命名变量。SPSS 变量的命名和一般的编程语言一样, 有一定的命名规则, 具体内容如下。

- 变量名 (Name) 必须以字母、汉字或字符@开头, 其他字符可以是任何字母、数字或_、@、#、\$等符号。
- 变量最后一个字符不能是句号。
- 不能使用空白字符或其他特殊字符 (如 “!”、“?” 等)。
- 变量命名必须唯一, 不能有两个相同的变量名。
- 在 SPSS 中不区分大小写。例如, HXH、hxx 或 Hxx 对 SPSS 而言, 均为同一变量名称。
- SPSS 的保留字 (Reserved Keywords) 不能作为变量的名称, 如 ALL、AND、WITH、OR 等。

2. 定义变量类型

单击 Type 相应单元中的按钮, 弹出如图 2-3 所示的对话框。在对话框中选择合适的变量类型 (Type) 并单击 “OK” 按钮, 即可定义变量类型。

SPSS 的主要变量类型如下。

- **Numeric:** 数值型, 定义数值的宽度 (Width), 即整数部分+小数点+小数部分的位数, 默认为 8 位; 定义小数位数 (Decimal Places), 默认为 2 位。

- **Comma:** 加显逗号的数值型, 即整数部分每 3 位数加一逗号, 其余定义方式同数值型, 也需要定义数值的宽度和小数位数。

- **Dot:** 3 位加点数值型, 无论数值大小, 均以整数形式显示, 每 3 位加一小点 (但不是小数点); 可定义小数位置, 但都显示 0, 且小数点用逗号表示。如 1.2345 显示为 12.345,00 (实际是 12345E-4)。

- **Scientific notation:** 科学记数型。同时定义数值宽度 (Width) 和小数位数 (Decimal Places), 在数据编辑窗口中以指数形式显示。如定义数值宽度为 9, 小数位数为 2, 则 345.678 显示为 3.46E+02。

- **Date:** 日期型, 用户可从系统提供的多种日期显示形式中选择自己需要的形式。如选择 mm/dd/yy 形式, 则 1998 年 8 月 25 日显示为 08/25/98。

- **Dollar:** 货币型, 用户可从系统提供的多种货币显示形式中选择自己需要的, 并定义数值宽度和小数位数, 显示形式为数值前加 \$ 符号。

- **Custom currency:** 用户自定义型, 如果没有定义, 则默认显示为整数部分每 3 位加一逗号, 用户可定义数值宽度和小数位数。如 12345.678 显示为 12,345.678。

- **String:** 字符型, 用户可定义字符长度 (Characters) 以便输入字符。

3. 变量长度

设置变量的长度 (Width), 当变量为日期型时无效。

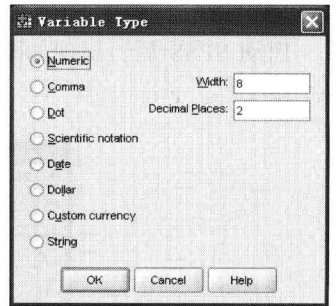


图 2-3 定义变量类型对话框

4. 变量小数点位数

设置变量的小数点位数 (Decimal)，当变量为日期型时无效。

5. 变量标签

变量标签 (Label) 是对变量名的进一步描述。变量只能由不超过 8 个字符组成，而 8 个字符经常不足以表示变量的含义。而变量标签可长达 120 个字符，变量标签可显示大小写，需要时可用变量标签对变量名的含义加以解释。

6. 变量值标签

值标签 (Values) 是对变量的每一个可能取值的进一步描述。当变量是定性或定序变量时，这是非常有用的。例如，在统计中经常用 1 代表“male”，2 代表“female”等。其具体的设置为：单击 Values 相应单元，弹出如图 2-4 所示的对话框；在第 1 个 Value 文本框内输入 1，在第 2 个 Value 文本框内输入“male”；单击“Add”按钮，完成添加该变量值标签。

7. 缺失值的定义方式

SPSS 有两类缺失值：系统缺失值和用户缺失值。在数据长方形中，任何空的数字单元都被认为是系统缺失值，用点号“.”表示。由于特殊原因造成的信息缺失值，称为用户缺失值。例如，在统计过程中可能需要区别一些被调查者不愿意回答的题目，然后将它们标为用户缺失值，统计过程可识别这种标识，带有缺失值的观测被特别处理。

单击 Missing 相应单元中的按钮，在弹出的如图 2-5 所示的对话框中可改变缺失值的定义方式 (Missing)，在 SPSS 中有两种定义缺失值的方式。

- 可以定义 3 个单独的缺失值。
- 可以定义 1 个缺失值范围和 1 个单独的缺失值。

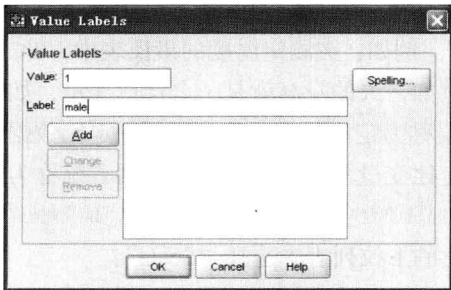


图 2-4 修改值标签

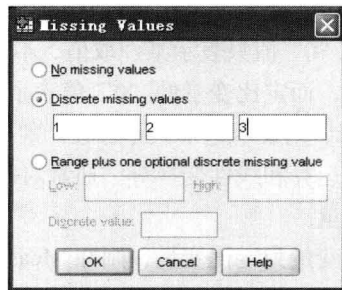


图 2-5 改变缺失值的定义方式

8. 变量的显示宽度

输入变量的显示宽度 (Columns)，默认为 8。

9. 变量显示的对齐方式

选择变量值显示时的对齐方式 (Align): Left (左对齐)、Right (右对齐)、Center (居中对齐)。默认是右对齐。

10. 变量的测量精度

变量按测量精度 (Measure) 可以分为定性变量、定序变量、定距变量和定比变量几种。

(1) 定性变量。定性变量又称为名义 (Nominal) 变量。这是一种测量精度最低、最粗略的基于“质”因素的变量，它的取值只代表观测对象的不同类别，例如，“姓名”变量、“性别”变量等都是定性变量。定性变量的取值称为定性数据或名义数据。定性数据的特点是用不多的名称来加以表达，并由被研究变量每一组出现的次数及其总计数所组成，这种数据是枚举型的，即由计