

SPSS 统计分析 实用教程

谭荣波 梅晓仁 编著

科学出版社

北京

内 容 简 介

SPSS 是世界上最优秀、用户使用最广泛的统计软件之一。本书以 SPSS13.0 版本为基础,结合实例对 SPSS 进行了深入浅出的分析,突出统计分析方法、实例分析和 SPSS 操作的有机结合,详细阐述 SPSS 中常用统计方法的操作步骤和输出结果的分析讨论与结论的陈述;同时又简明扼要、通俗易懂地介绍统计方法的基本原理和功能,把统计学原理的阐释与 SPSS 的操作应用有机结合起来,使读者知其然知其所以然。

本书第 1 章介绍 SPSS13.0 的特点、功能,运行环境与安装,帮助系统等;第 2 章讲述 SPSS 的数据编码、录入与整理;第 3~12 章介绍 SPSS 的各种统计方法,包括统计报表、描述统计、均值比较、方差分析、非参数检验、相关分析、回归分析、聚类分析与判别分析、因子分析、信度分析等。

本书可作为高等院校教育学、心理学、体育学、经济学等专业的本科生和研究生的学习教材和指导书,也可作为从事数据统计有关工作的教师和科研工作者的参考工具和使用手册。

图书在版编目(CIP)数据

SPSS 统计分析实用教程/谭荣波,梅晓仁编著. —北京:科学出版社,2007

ISBN 978-7-03-018932-5

I. S… II. ①谭…②梅… III. 统计分析—软件包, SPSS—教材
IV. C819

中国版本图书馆 CIP 数据核字(2007)第 065599 号

责任编辑:张丽娜 赵丽艳 / 责任制作:魏 谨

责任印制:赵德静 / 封面设计:李 祥

北京东方科龙图文有限公司 制作

<http://www.okbook.com.cn>

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

中国科学院印刷厂印刷

科学出版社发行 各地新华书店经销

*

2007 年 5 月第 一 版 开本: B5(720×1000)

2007 年 5 月第一次印刷 印张: 14 1/2

印数: 1—3 000 字数: 286 000

定 价: 28.00 元

(如有印装质量问题,我社负责调换(新欣))

序

——从新技术到新教材和新课程

在特定的历史时代,技术的变革将置换理论的内容,新技术的推广将带来课程理念的突变与飞跃。所以,一个国家、一个民族总是把技术的因素当作社会生产力的重要要素;一个科研工作者、一个担负着教书育人职责的人民教师,总是对新技术的出现特别是对新技术介入教学和科研充满了渴望和激情。谭荣波、梅晓仁就是这样的科研人员,就是这样的优秀的新型教师。

谭荣波作为一个教育学科班出身的科研人员、梅晓仁作为一个计算机科班出身的商学院教师,他们不仅努力学习新技术,而且还把这些新技术运用于各自的研究领域。谭荣波早在两年前就在中国社会科学出版社出版了《基于网络的研究性学习》。在这本书里,他把网络技术、计算机技术给学习带来的深刻变化作了准确的描写和预测。新时代的新的学习论就是在新技术的强劲介入和巨力推动下诞生的。谭荣波继续他的教育研究,梅晓仁继续他的信息管理学研究,他们都没有忘记培育和发挥自己的“新技术翅膀”。只有“新技术翅膀”奋力搏动,他们才能在各自的研究领域里飞得更高、飞得更远。

这本《SPSS 统计分析实用教程》就是他们采用“新技术翅膀”奋力搏击的精彩结晶。

SPSS 是目前世界上比较优秀的统计分析软件之一,它的全称为“Statistical Package for the Social Science”,随着该新统计技术介入各个学科领域以及它的服务深度的增加,现在英文名称则为“Statistical Product and Service Solutions”。在 20 世纪末(1998 年),该新技术和新软件开始登陆中国大陆的高校,成为了国内教育科研人员使用频率较高的统计工具,特别是它对于心理学、教育统计学、生物学、经济管理、医疗卫生、农林科学以及金融、体育等学科专业有着意义巨大的实用价值。时至今日,随着网络技术和计算机技术的快速发展,计算信息技能已经成为大学本科生、研究生以及从事各学科研究工作的教师必不可少的基本技能。对于那些常常要处理大量繁复数据的各学科本科生、研究生以及教学科研人员,掌握 SPSS 的新技术,并熟练地在学习和科研中运用它,则显得更为迫切和重要。

谭荣波、梅晓仁在这本《SPSS 统计分析实用教程》里,并不限于一般性地介绍 SPSS,可以想象,如果他们只限于介绍,那就成了具体化的产品中文说明书了。他们结合各自专业的学习、研究的需求,一方面侧重于传授、讲解 SPSS 操作技能的实用性,另一方面又强调了统计学专业知识的应用,把统计学原理的阐释与

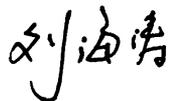
SPSS 的操作应用有机地结合起来。这就是新技术用于学科专业产生新的课程理论和方法的典型案例,这也构成了这本教材不同于其他同类教材的如下新的亮点:

第一,这本教材立足于数据处理技能的教学和训练。它根据数据处理的方法与模型,分门别类地阐释每种数据处理的统计学原理方法。在此基础上,精选了一些案例,具体说明如何组织数据文件,选用哪些处理过程(命令),进行哪些必要的设定,如何解释统计结果的意义。而对 SPSS 的其他许多专业功能,按“用多少,学多少”的原则进行大胆处理,达到了化繁为简、易教易学的效果。

第二,这本教材突出了实践教学环节和过程。作者不但精选了数量较多的教学案例,而且还对教学案例进行了必要的精确的分析,编制了丰富的习题库,为初学者处理数据、撰写学术论文提供了有效的参考。这也是其他软件介绍类的书籍所不具备的。因此,它的教学目标定位很准——这门课程和这本教材不一定要培养 SPSS 的专家,但是它能培养初学者和研究人员成为进行专业数据处理的行家。

谭荣波、梅晓仁从学习新技术开始,编写了新教材,开发了新课程,总结出新的课程理论和方法,这一过程正是一个优秀教师和优秀科研人员的成长过程。

湛江师范学院副院长



2007年1月8日

前 言

新世纪以来,随着以网络为主的信息技术迅速发展,信息化的浪潮席卷世界各国的教育。信息化已成为世界范围内教育现代化建设的标志,以信息化带动教育现代化已成为教育跨越式发展的必由之路。对学生的信息获取、储存、提取和应用等信息素养与创新能力的培养已成为当前信息技术教育的核心内容。本书是根据教育部制定的《高等学校文科类专业大学计算机教学基本要求(2003版)》中关于“在开设必修的计算机公共基础课程后,应当增设更加结合所在专业需要的后续课程”的要求组织编写的教材,它是结合教育学、心理学、经济学、管理学等专业的学习需要,对数据统计分析方法和手段进行学习的后续课程教材。

SPSS是世界著名的统计分析软件之一。在国际学术界有一条不成文的规定:凡是用SPSS软件完成的计算和统计分析,可以不必说明算法。由此可见其影响之大和信誉之高。鉴于目前国内出版的有关SPSS的书籍多从介绍该软件功能使用的角度来阐述,让初学者感觉“学会了SPSS,但不知道怎么用?”为此,本书结合编者多年从事SPSS教学、教育统计学、计算机教学等相关学科的教学经验和多次利用SPSS完成科研项目中大量数据的统计分析的科研工作经验,从用户角度出发,将统计分析方法、案例分析和SPSS操作有机地结合起来,力争使初学者或研究工作者成为进行专业数据处理的行家但不一定是SPSS的专家。

本书共12章,具体分为SPSS概述、数据的编码录入与整理、统计报表、描述性统计、两总体均值比较、方差分析、非参数检验、相关分析、回归分析、聚类分析和判别分析、因子分析、信度分析等内容。每章的主要内容安排为:①统计学原理:阐释本章节中每一种统计方法的统计学原理及其具体要求条件等;②案例与数据文件:给出具体问题的典型实际案例及数据储存格式;③统计操作过程:具体说明在SPSS上如何进行操作,解决实际案例中提出的问题;④结果分析:对SPSS给出的各种结果进行解释、分析,回答典型案例中提出的问题。

本书由谭荣波拟订提纲,执笔第2章、第5~7章、第10~12章,并负责全书的统稿;梅晓仁博士执笔第1章、第3~4章、第8~9章。另外,书中案例数据、习题分析和SPSS相关资源可登陆<http://202.192.128.41/zsb/spssweb>下载和浏览。

本书的完稿、出版得到了多方人士的关怀、指导和帮助。我院副院长刘海涛教授百忙中详细地审阅了书稿,提出了许多宝贵的意见,并亲自为本书作序。没有他的鞭策与支持,本书稿就不会如此顺利地完成。本书的出版还得到了广东省教育科学“十五”规划重点项目的基金资助和湛江师范学院博士基金的资助。对

此,一一致谢。

本书从写作结构、体例和案例选择上借鉴了多种相关书籍,吸取了一些有益的观点,引用了一些宝贵的资料,力图做到易学、简明、实用。在此向诸位作者表示深切的谢意。由于水平有限,加上编著时间仓促,书中难免有错漏之处,敬请不吝指正。

编 者

2007 年 1 月

电子信箱:zjtanrb@163.com

meixr0506@163.com

目 录

第 1 章 SPSS 13.0 for Windows 概述

1.1	SPSS 的发展与特点	1
1.2	SPSS 的功能介绍	2
1.2.1	数据的输入与编辑	2
1.2.2	数据的一般性管理	2
1.2.3	各种统计分析	2
1.2.4	图形与输出报告	3
1.2.5	与其他程序交换数据	3
1.2.6	二次开发功能	4
1.3	SPSS 的运行环境与安装	4
1.3.1	SPSS 13.0 for Windows 运行的硬件环境	4
1.3.2	SPSS 13.0 for Windows 运行的软件环境	4
1.3.3	SPSS 13.0 for Windows 的安装	5
1.4	SPSS 的帮助系统	7
1.4.1	联机帮助	7
1.4.2	在线帮助	8
1.4.3	图解帮助	8
1.5	利用 SPSS 进行数据分析的基本步骤	9
1.5.1	数据分析的一般步骤	9
1.5.2	利用 SPSS 进行数据分析的一般步骤	10

第 2 章 数据的编码、录入与整理

2.1	数据的编码	12
2.1.1	问卷与编码	12
2.1.2	编码应注意事项	13
2.2	数据的录入	14
2.2.1	操作术语	14
2.2.2	定义变量	14
2.2.3	数据的录入	17

2.2.4 数据的导入	18
2.3 数据的整理	19
2.3.1 数据分值的转换	19
2.3.2 量表的统分	21
2.3.3 数据的限选	22
2.4 思考题	24

第3章 统计报告

3.1 在线分析处理报告	26
3.1.1 在线分析处理报告的对话框	27
3.1.2 案例分析	30
3.2 个案摘要报告	32
3.2.1 个案摘要报告的对话框	32
3.2.2 案例分析	33
3.3 行形式摘要报告	35
3.3.1 行形式摘要报告的对话框	35
3.3.2 案例分析	39
3.4 列形式摘要报告	41
3.4.1 列形式摘要报告的操作	41
3.4.2 案例分析	43
3.5 思考题	44

第4章 描述性统计分析

4.1 频数分析	45
4.1.1 统计学原理	46
4.1.2 SPSS 操作及案例分析	46
4.2 基本描述统计量	51
4.2.1 统计学原理	51
4.2.2 SPSS 操作及案例分析	53
4.3 探索性分析	55
4.3.1 统计学原理	55
4.3.2 SPSS 操作及案例分析	59
4.4 交叉列联表分析	63
4.4.1 统计学原理	63
4.4.2 SPSS 操作及案例分析	65
4.5 比率分析	71

4.5.1 统计学原理	71
4.5.2 SPSS 操作及案例分析	72
4.6 多选项分析	75
4.6.1 SPSS 操作	75
4.6.2 案例分析	78
4.7 思考题	79

第 5 章 两总体均值比较

5.1 单样本 T 检验	82
5.1.1 统计学原理	82
5.1.2 SPSS 操作及案例分析	82
5.2 独立样本 T 检验	84
5.2.1 统计学原理	84
5.2.2 SPSS 操作及案例分析	85
5.3 配对样本 T 检验	87
5.3.1 统计学原理	87
5.3.2 SPSS 操作及案例分析	88
5.4 思考题	90

第 6 章 方差分析

6.1 单因素方差分析	92
6.1.1 统计学原理	92
6.1.2 SPSS 操作及案例分析	93
6.2 单因变量多因素方差分析	96
6.2.1 统计学原理	96
6.2.2 SPSS 操作及案例分析	97
6.3 协方差分析	100
6.3.1 统计学原理	100
6.3.2 SPSS 操作及案例分析	101
6.4 重复测量方差分析	102
6.4.1 统计学原理	102
6.4.2 SPSS 操作及案例分析	102
6.5 思考题	106

第 7 章 非参数检验

7.1 卡方检验	109
----------------	-----

7.1.1 统计学原理	109
7.1.2 SPSS 操作及案例分析	109
7.2 二项分布检验	111
7.2.1 统计学原理	111
7.2.2 SPSS 操作及案例分析	112
7.3 游程检验	113
7.3.1 统计学原理	113
7.3.2 SPSS 操作及案例分析	114
7.4 单个样本 K-S 检验	115
7.4.1 统计学原理	115
7.4.2 SPSS 操作及案例分析	116
7.5 两独立样本的非参数检验	117
7.5.1 统计学原理	117
7.5.2 SPSS 操作及案例分析	118
7.6 多独立样本的非参数检验	121
7.6.1 统计学原理	121
7.6.2 SPSS 操作及案例分析	122
7.7 两相关样本的非参数检验	124
7.7.1 统计学原理	124
7.7.2 SPSS 操作及案例分析	124
7.8 多相关样本的非参数检验	126
7.8.1 统计学原理	126
7.8.2 SPSS 操作及案例分析	127
7.9 思考题	128

第 8 章 相关分析

8.1 二元变量分析	130
8.1.1 统计学原理	130
8.1.2 SPSS 操作及案例分析	134
8.2 偏相关分析	140
8.2.1 统计学原理	140
8.2.2 SPSS 操作及案例分析	140
8.3 距离相关分析	142
8.3.1 统计学原理	142
8.3.2 SPSS 操作及案例分析	142
8.4 思考题	146

第 9 章 回归分析

9.1 线性回归分析	149
9.1.1 统计学原理	149
9.1.2 SPSS 操作及案例分析	150
9.2 曲线估计	165
9.2.1 统计学原理	165
9.2.2 SPSS 操作及案例分析	166
9.3 非线性回归分析	171
9.3.1 统计学原理	171
9.3.2 SPSS 操作及案例分析	172
9.4 思考题	179

第 10 章 聚类分析和判别分析

10.1 快速聚类	182
10.1.1 统计学原理	182
10.1.2 SPSS 操作及案例分析	182
10.2 分层聚类	185
10.2.1 统计学原理	185
10.2.2 SPSS 操作及案例分析	185
10.3 判别分析	190
10.3.1 统计学原理	190
10.3.2 SPSS 操作及案例分析	191
10.4 思考题	194

第 11 章 因子分析

11.1 统计学原理	195
11.1.1 因子分析的含义	195
11.1.2 数学模型	196
11.1.3 因子分析的基本步骤	196
11.2 SPSS 操作及案例分析	198
11.2.1 因子分析的对话框介绍	198
11.2.2 案例分析	200
11.3 思考题	208

第 12 章 信度分析

12.1 统计学原理	209
------------------	-----

12.2 SPSS 操作及案例分析	211
12.2.1 信度分析的对话框介绍	211
12.2.2 案例分析	212
12.3 思考题	217

参考文献	218
------------	------------

第 1 章 SPSS 13.0 for Windows 概述

SPSS 是 Statistical Package for the Social Science 的缩写,即“社会科学统计软件包”,是世界著名的统计分析软件之一。它广泛应用于经济学、生物学、心理学、教育学、医疗卫生、体育、农业、林业、商业、金融等各个领域。在国际学术界有一条不成文的规定,凡是用 SPSS 软件完成的计算和统计分析,可以不必说明算法,由此可见其影响之大和信誉之高。

1.1 SPSS 的发展与特点

SPSS 是世界上最早的统计分析软件之一,由美国斯坦福大学三位研究生于 20 世纪 60 年代末研制开发,并于 1975 年在芝加哥成立了 SPSS 公司。1984 年 SPSS 首先推出了世界上第一个统计分析软件微机版本 SPSS/PC+(版本为 SPSS/PC+x.x),它迅速占领了微机市场,扩大了自己的用户量,开创了 SPSS 微机系列产品的开发方向。20 世纪 80 年代末,Microsoft 公司发布 Windows 操作系统后,SPSS 迅速向 Windows 移植。1993 年 6 月,SPSS 公司正式推出 SPSS 6.0 for Windows 版本。随着 SPSS 版本的升级、功能的不断完善,许多有影响的报刊杂志纷纷就 SPSS 的自动统计绘图、数据的深入分析、使用方便、功能齐全等方面给予了高度的评价和称赞。

1994 至 1998 年间,SPSS 公司陆续并购了 SYSTAT、BMDP 和 ISL 等公司,由原来单一统计产品开发转向企业、教育科研及政府机构提供全面信息统计决策支持服务,发展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。伴随 SPSS 产品服务领域的扩大和服务深度的增加,SPSS 公司于 2000 年正式将其名称改为 Statistical Product and Service solutions,意为“统计产品与服务解决方案”。

SPSS 的突出优点是操作界面友好,输出结果美观漂亮,是非专业统计人员的首选统计软件。SPSS 是世界上最早采用图形菜单驱动界面的统计软件,它几乎将所有的功能都以统一、规范的界面展现出来。SPSS 采用 Windows 窗口方式展示各种管理和分析数据方法的功能,使用对话框展示各种功能选择项,只要掌握一定的 Windows 操作技能,略通统计分析原理,就可以使用该软件为特定的科研工作服务。在众多用户对国际常用统计软件 SAS、BMDP、GLIM、GENSTAT、EPILOG 和 MiniTab 的总体印象中,SPSS 诸项功能均获得最高分。

SPSS 采用类似 EXCEL 表格的方式输入与管理数据,数据接口较为通用,能方便地从其他数据库中读入数据。对于熟悉老版本编程运行方式的用户,还特别设计了语法生成窗口,用户只需在菜单中选好各个选项,然后单击“粘贴”按钮即可自动生成标准的 SPSS 程序,极大地方便了中、高级用户,也有利于重复或批量进行数据统计分析。

但是 SPSS 软件并不是完美的,它自身还存在着一些不足,如输出结果虽然漂亮,但不能为 Word 等常用文字处理软件直接打开,只能采用拷贝、粘贴的方式加以交互;又如针对树结构模型的 Answer Tree,针对神经网络技术的 Neural Connection、专门用于数据挖掘的 Clementine 等都不是直接纳入 SPSS,等等。为此 SPSS 公司也采取了相应的举措。

SPSS 作为统计分析工具,理论严谨、内容丰富,数据管理、统计分析、趋势研究、制表绘图、文字处理等功能,几乎无所不包。本书以 SPSS 最新的单机标准版 SPSS 13.0 for Windows 为蓝本,介绍 SPSS 的具体使用方法。

1.2 SPSS 的功能介绍

1.2.1 数据的输入与编辑

在数据输入方面,SPSS 不仅可以直接读取很多其他系统的数据,比如 Microsoft Office Excel、dBase、Lotus、SAS 等系统的数据,而且高版本的 SPSS 增加了很多支持 ODBC 数据库的驱动程序,包括 Oracle、SQL Server、DB2 UDB、Microsoft Access、Sybase 等,以便获得不同系统的数据。

SPSS 系统提供数据编辑器 SPSS Data Editor,该编辑器可以经过编码产生数据表格。SPSS 10.0 for Windows 以后的版本对数据属性的设置界面进行了较大的改动,对于大量变量的数据文件,数据属性的设置更为方便。

1.2.2 数据的一般性管理

SPSS 数据管理涵盖了数据库管理的大部分功能和数据统计分析的特殊功能,使数据统计分析非常简易、方便。

SPSS 数据管理包括变量及其观测值管理、个案管理、文件及数据管理,可实现例如数据的转置、样本排序、抽样、加权、变量的计算赋值、计数赋值、条件赋值及文件的拆分、合并等功能。

1.2.3 各种统计分析

SPSS 的统计分析功能包括基本统计分析和高级统计分析。SPSS 13.0 for Windows 的基本统计(SPSS Base for Windows)分析模块提供了多种常用统计运算方法,例如统计描述、方差分析、相关分析等,详见表 1.1。这些常用统计表方法广泛应用于教育、心理及其相关学科专业中。

SPSS 13.0 for Windows 的高级统计分析模块(SPSS Advanced Models)提供

了多种更为深入地考察数据内部关系的统计运算功能,例如可靠性分析、非线性回归、概率分析回归等,详见表 1.2。

表 1.1 SPSS 基本统计分析运算

运 算	功 能	运 算	功 能
Frequencies	频度分析	Correlation	相关分析
Descriptives	统计描述	Partial CORR	偏相关分析
Examine	数据考察分析	Distances	距离分析
T-Test	T 检验	Regression	回归分析
NonPAR Test	非参数检验	Curve fit	曲线分析
Multiple Response	多重响应分析	Cluster	聚类分析
One-way ANOVA	一元方差分析	Quick Cluster	快速聚类分析
UNIANOVA	多元方差分析	Factor	因子分析
ANOVA	协方差分析	Discriminant	判别分析
MANOVA	复方差分析		

表 1.2 SPSS 高级统计分析运算

运 算	功 能
ALSCAL	多维标度
Reliability	可靠性分析
NLR	非线性回归
PROBIT	概率分析回归
WLS	加权最小二乘法
2SLS	两分段最小二乘法
MIXED	拟合多种混合线性模型
Time Program 和 COXREG	以时间为协变量的 Cox 回归分析
LOGLINEAR	对数线性回归分析
GENLOG	最大似然法计算一般对数线性模型参数的回归分析
LOGISTIC	二分变量或连续变量的准独立性模型 Logistic 回归过程
HILOG Linear	多维列联表的有序对数模型。分层对数线性模型的回归分析
LOGIT	指定分类变量作为因变量的 Logit 模型
Survival Table	生存分析中的生命表分组比较
PLUM	预示顺序结果
KAPLAN-MEIER	用 Kaplan-Meier 估计方法估计事件的时间长度
GLM: Multivariate	通用线性模型的复方差分析
GLM: Repeated Measures	重复测量的多变量方差分析

1.2.4 图形与输出报告

SPSS 提供了完善的输出表格和高分辨率图形,包括十种分类图形、六种质量控制图、六种直方图散点图、八种诊断和探测图等近百种图形功能。

1.2.5 与其他程序交换数据

SPSS 13.0 for Windows 软件包提供了与其他软件包共享数据的功能,从而

极大地扩展了 SPSS 的用途,SPSS 13.0 for Windows 能够打开和保存的数据文件格式有:

- * .sys:SPSS / PC +(DOS 版的 SPSS);
- * .por:SPSS 可移植文件;
- * .xls:Microsoft 公司的 Excel 格式数据文件;
- * .slk:Multiplan 公司的 symbolic 格式数据文件;
- * .dbf:数据库 dBase、FoxBase 等格式数据文件;
- * .w * :Lotus 1-2-3 格式数据文件;
- * .dat:自由格式的纯文本数据文件;
- * .txt:纯文本文件;
- * .syd:Systat 软件包的数据文件;
- * .sys:Systat 软件包的系统文件;
- * .sas7dbat:SAS 软件包的长文件名的数据文件;
- * .sd7:SAS 软件包的短文件名的数据文件;
- * .sd2:SAS v6 for Windows 格式数据文件;
- * .ssd01:SAS v6 for Unix 格式数据文件。

1.2.6 二次开发功能

SPSS 13.0 for Windows 具有比较完备的二次开发功能。不但有源于 DOS 版本的 Syntax 命令集,还有内置的 VBA 客户语言——SaxBasic 语言。在 SPSS 的脚本编辑器中混合使用 SaxBasic 和 Syntax,可以很方便地实现脚本的编辑,从而提高统计分析的效率和功能。

除了能编写内部脚本,SPSS 13.0 for Windows 还支持 DLL 技术,调用 Windows 核心库或外部库中的 API 函数。它还支持 DDE、OLE 和 ActiveX 等技术,可以与其他同样支持该技术的程序进行接口。

1.3 SPSS 的运行环境与安装

1.3.1 SPSS 13.0 for Windows 运行的硬件环境

SPSS 13.0 for Windows 运行的硬件环境要求:

- ① Pentium II 以上的计算机;
- ② 有效内存不少于 128MB;
- ③ 硬盘空间在 300MB 以上;
- ④ 显示器可以选 SVGA 800×600 像素或更高。

1.3.2 SPSS 13.0 for Windows 运行的软件环境

SPSS 13.0 for Windows 可以在 Windows9x/2000/NT/XP 环境下运行。

1.3.3 SPSS 13.0 for Windows 的安装

① SPSS 13.0 for Windows 软件通常存放在一张光盘上,可以通过 Windows 的“资源管理器”找到软件的可执行文件 setup.exe 所在的文件夹,执行 setup.exe,出现图 1.1 所示的窗口。

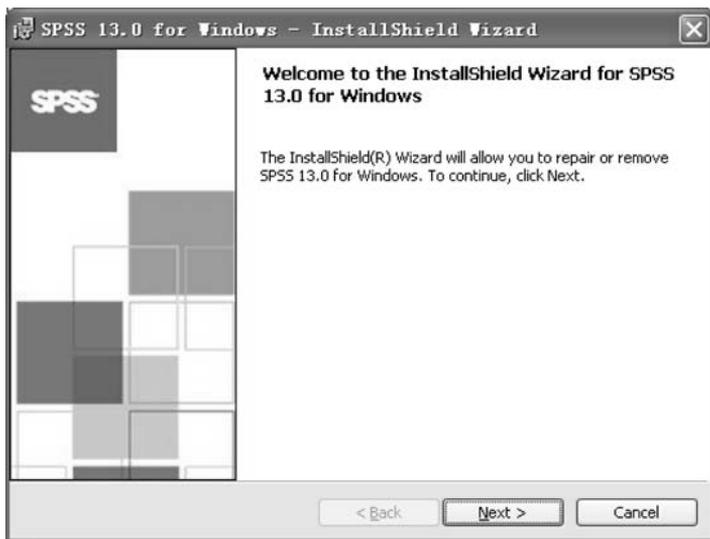


图 1.1 SPSS13.0 for Windows 安装窗口

② 点击“Next”按钮,出现“License Agreement”窗口,如图 1.2 所示。选择“I accept the terms in the license agreement”,点击“Next”按钮。

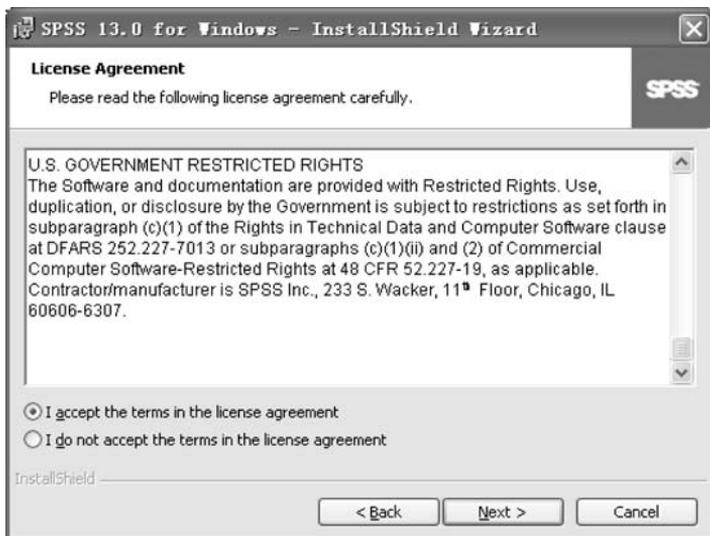


图 1.2 “License Agreement”窗口

③ 在“Destination Folder”窗口中,指定 SPSS 13.0 for Windows 为软件的安装位置。系统默认安装在 Program Files 文件夹中。用户可以自行指定安装的位置,并指定文件夹的名称,如图 1.3 所示。

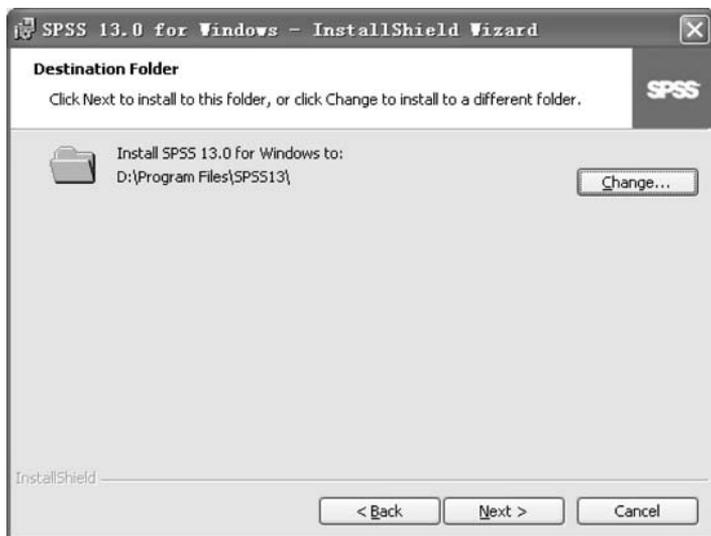


图 1.3 指定安装文件夹位置

④ 点击“Next”按钮,出现“Installing SPSS 13.0 for Windows”窗口,如图 1.4 所示,系统开始安装。



图 1.4 “Installing SPSS 13.0 for Windows”窗口

系统安装完成后,出现的信息如图 1.5 所示。

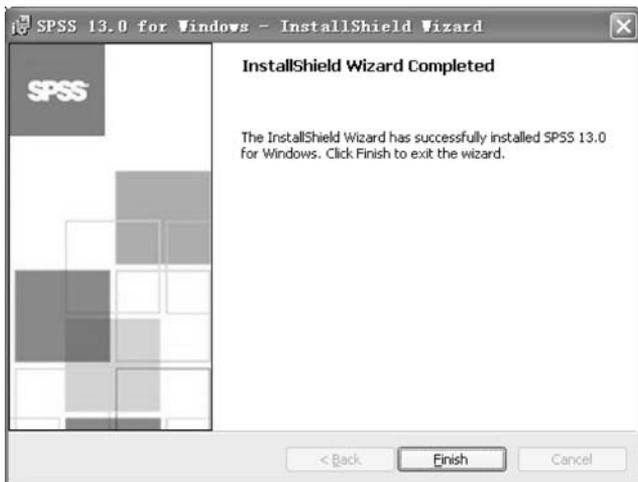


图 1.5 SPSS 13.0 for Windows 安装完毕信息

1.4 SPSS 的帮助系统

SPSS 13.0 for Windows 提供了功能强大的帮助系统,包括联机帮助、在线帮助、图解帮助等多种帮助手段。利用这些功能,使用者可以方便地实现查询,更快地掌握 SPSS 和提高工作效率。

1.4.1 联机帮助

在主菜单中单击“Help”选项,然后在打开的下拉式菜单中单击“Topics”选项,打开联机帮助对话框,如图 1.6 所示。在该对话框中根据需要逐级打开目录,可以获得帮助信息。例如,在对话框中单击“Overview”目录,然后单击“*What's New in SPSS 13.0*”选项,打开图 1.6 所示的帮助信息。

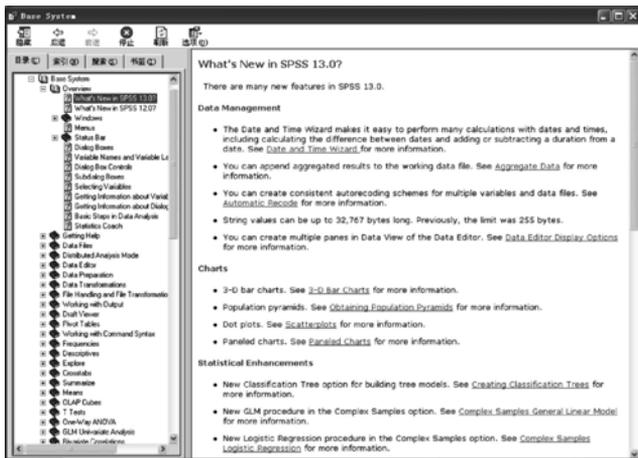


图 1.6 联机帮助

1.4.2 在线帮助

在 SPSS 中为实现一定功能准备数据时,需要打开一系列的对话框,这些对话框中一般都有一个“Help”按钮,该按钮提供在线帮助。

假设现有某学期某班学生 SPSS 考试成绩,现在想把百分制转换成五分制,打开考试成绩数据文件后,调用 Transform→Recode→into Different Variables 命令打开 Recode into Different Variables 对话框,如图 1.7 所示。如果用户不知道如何使用,除了可以使用联机帮助外,最快捷的方法就是直接在对话框中单击“Help”按钮,打开图 1.8 所示的帮助信息。利用该信息及相关信息,就可以了解条件赋值生成新变量的概念、方法、对话框中各控件的意义等功能。



图 1.7 “Recode into Different Variables”对话框

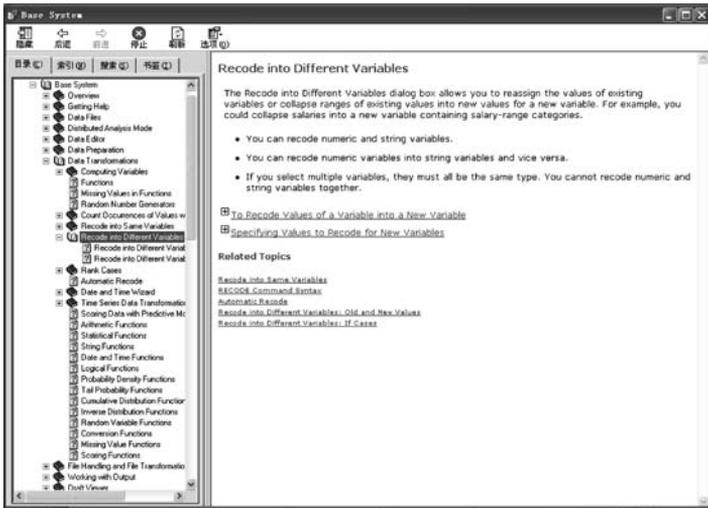


图 1.8 帮助信息

1.4.3 图解帮助

对于 SPSS 中的基本操作方法和与数据文件有关的一些操作,该软件提供了

比较具体的图解帮助功能。利用该功能,用户可以迅速地了解 SPSS 的基本使用方法。获得该帮助的方法是在“Help”主菜单中单击“Tutorial”选项,打开图解帮助窗口,如图 1.9 所示。单击某一项目,可以获得对应主题的图解帮助信息。

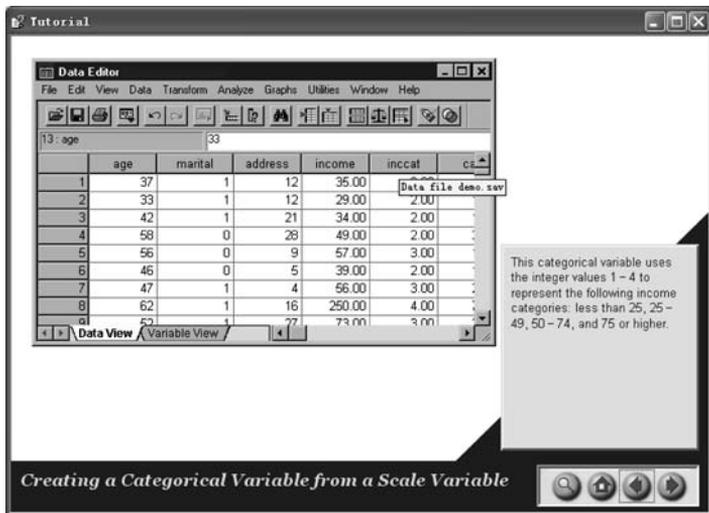


图 1.9 图解帮助

1.5 利用 SPSS 进行数据分析的基本步骤

学习和应用 SPSS 软件的过程并不是单纯地学习和应用一种计算机软件的过程。由于 SPSS 是一种专业性较强的统计软件,因此,学习和应用它时必须了解和掌握必要的统计学专业知识和数据分析的一般步骤和原则。这样才能避免滥用和误用方法,不至于因引用偏差甚至错误的数据分析结论而做出错误的决策。

1.5.1 数据分析的一般步骤

数据分析一般经过收集数据、加工和整理数据、分析数据等三个主要阶段,统计学对此有非常完整和严谨的论述。在数据分析的实践中,用统计学的理论指导应用是必不可少的,也是极为重要的。

(1) 明确数据分析目标

明确数据分析目标是数据分析的出发点。明确数据分析目标就是要明确本次数据分析要研究的主要问题和预期的分析目标等。只有这样,才能正确地制定数据收集方法和方案,例如,应该收集哪些数据,应采用怎样的方法收集等,进而为数据分析做好准备。

(2) 正确收集数据

正确收集数据是指应从分析目标出发,排除干扰因素,正确收集服务于既定分析目标的数据。正确的数据对于实现数据分析目标起到关键性的作用。在收

集数据的过程中,经常会获得一些与分析目标无关或者对分析目标起相反作用的干扰数据,排除这些数据是数据收集的重要环节。数据分析并不仅仅是对数据进行数学建模,收集的数据是否真正迎合数据分析的目标,其中是否包含其他因素的影响及其影响程度,应如何剔除这些影响等问题都是数据分析过程中必须注意的重要问题。

(3) 数据的加工整理

在明确数据分析目标的基础上收集到的数据,往往还需进行必要的加工整理后才能真正用于分析建模。通过数据的加工整理,人们能够大致掌握数据的总体分布特征,这是今后进一步深入分析和建模的基础。数据的加工整理通常包括数据的缺失值处理、数据的分组、基本描述统计量的计算、基本统计图形的绘制和数据标准化处理等。

(4) 明确统计方法的含义和适用范围

数据加工整理完成后,一般就可以做进一步的数据分析了。分析时应切忌滥用和误用统计分析方法。滥用和误用统计分析方法主要是由于对方法能解决哪类问题、方法适用的前提、方法对数据的要求不清楚等原因造成的。另外,统计软件的不断普及和应用中的不求甚解也会加重这种现象。因此,在数据分析中应避免盲目的“拿来主义”,否则,得到的分析结论可能会偏差较大,甚至错误。

(5) 读懂分析结果,正确解释分析结果

数据分析的直接结果是统计指标和统计参数。正确理解这些指标和参数的统计含义是一切分析结论的基础。正确理解统计指标和统计参数的含义不仅能够帮助人们有效避免毫无根据地随意引用统计数字的错误,同时也是证实分析结论正确性和可信性的依据。而这一切都取决于人们能否正确地把握统计分析方法的核心思想。

另外,将统计指标和统计参数与实际问题相结合也是非常重要的。客观地说,统计方法仅仅是一种有用的数量分析工具,它绝不是万能的。统计方法是否能够正确地解决各学科的具体问题,不仅取决于正确地选择统计方法,还取决于人们是否具有深厚的实际应用背景。只有将各学科的专业知识与统计指标和统计参数相结合,才能得出令人满意的分析结论。

1.5.2 利用 SPSS 进行数据分析的一般步骤

利用 SPSS 进行数据分析也应遵循数据分析的步骤,但涉及的方面会相对较少。主要工作分以下四个步骤。

(1) SPSS 数据的准备阶段

按照 SPSS 的要求,利用 SPSS 提供的功能准备 SPSS 数据文件。其中包括在数据编辑窗口中定义 SPSS 数据的结构、录入和修改 SPSS 数据等。

(2) SPSS 数据的加工整理阶段

对数据编辑窗口中的数据进行必要的预处理。

(3) SPSS 数据的分析阶段

选择正确的统计分析方法,对数据编辑窗口中的数据进行分析建模。由于 SPSS 能够自动完成数据建模中的数学计算并给出计算结果,使分析人员无需记忆数学公式,这无疑给统计分析和 SPSS 的广泛应用铺平了道路。

(4) SPSS 分析结果的解释

读懂 SPSS 输出窗口中的分析结果,明确其统计含义,并结合应用背景知识做出切合实际的合理解释。

利用 SPSS 进行数据分析的基本步骤如图 1.10 所示。

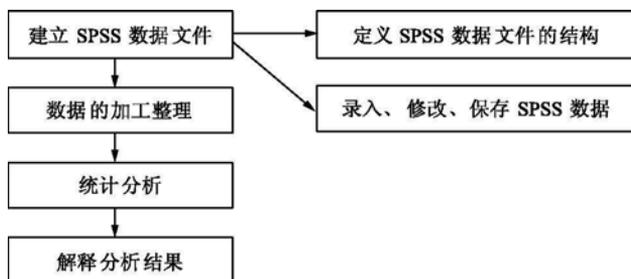


图 1.10 利用 SPSS 进行数据分析的基本步骤

第 2 章 数据的编码、录入与整理

数据的编码、录入与整理是研究者利用 SPSS 进行统计分析的必要前提。实际上这一工作就是把通过问卷调查获得的资料转变为 SPSS 能够识别、统计的数据文件,为各种统计分析做好准备。

2.1 数据的编码

2.1.1 问卷与编码

表 2.1 是一个假设的问卷量表,在实际使用中并没有操作意义,这里只解释如何进行问卷编码,其编码方案见表 2.2。

表 2.1 问卷量表

1. 您的性别:(1)男 (2)女
2. 您的文化程度:(1)没上过学 (2)小学 (3)初中 (4)高中 (5)大专以上
3. 您想选择下列哪些择偶条件(任选 3 项):(1)相貌 (2)文化水准 (3)气质风度 (4)志同道合 (5)人品 (6)家庭条件 (7)个人收入 (8)其他
4. 您购买房屋时,会考虑哪些因素(任选):(1)离工作地点的远近 (2)小孩所就读的学校 (3)居家附近的环境
5. 您对心理学感兴趣:(1)不感兴趣 (2)感兴趣 (3)非常感兴趣
6. 您对学英语感到厌恶:(1)不厌恶 (2)厌恶 (3)非常厌恶

表 2.2 问卷编码方案

变量名	编 码	样 例	
第 1 题	1 男 0 女	男	1
第 2 题	1 没上过学 2 小学 3 初中 4 高中 5 大专以上	小学	2
第 3 题-1	1 相貌 2 文化水准 3 气质风度 4 志同道合 5 人品 6 家庭条件 7 个人收入 8 其他	人品	2
第 3 题-2	1 相貌 2 文化水准 3 气质风度 4 志同道合 5 人品 6 家庭条件 7 个人收入 8 其他	相貌	4
第 3 题-3	1 相貌 2 文化水准 3 气质风度 4 志同道合 5 人品 6 家庭条件 7 个人收入 8 其他	其他	5
第 4 题-工作	1 选 0 不选	选	1
第 4 题-学校	1 选 0 不选	不选	0
第 4 题-环境	1 选 0 不选	选	1
第 5 题	A B C(A-1, B-2, C-3)	A	1
第 6 题	A B C(A-3, B-2, C-1)	A	3

2.1.2 编码应注意事项

1. 数值型数据的编码

社会科学领域尤其是教育心理领域的调查问卷多数属于数值型数据的量表。一般地,数值型数据的编码方法较为简单,即根据调查问卷的评分标准赋予分值就可以了。^① 态度量表通常采用三点计分、四点计分和五点计分等方式进行评分,如选项 A、B、C 计分为 1、2、3,选项 A、B、C、D 计分为 1、2、3、4。表 2.1 第 5、6 题就是属于数值型题目,其中第 5 题是正向数值型题目(即被选项的程度越高,分值越大),其计分方法为 1、2、3;第 6 题是反向数值型题目(即被选项的程度越高,分值越小),其计分方法为 3、2、1。当然,量表中每道题的评分标准各不相同。

2. 非数值型数据的编码

对于非数值型数据的编码,首先要确定一个编码规则,然后根据该规则对相应变量赋予分值。一般地,对于双值型变量的编码,多采用“1,0”或“1,2”赋值,例如表 2.1 第 1 题用“1”表示“男”,用“女”表示“0”。对于多值型变量的编码,可采用“1,2,3,…”对变量进行赋值,例如表 2.1 第 2 题“文化程度”就有“没上过学”、“小学”、“初中”、“高中”和“大专以上”5 个选项,可以分别用“1,2,3,4,5”来表示这 5 个选项。

一般情况下,对非数值型数据的编码,主要是起到分组的作用,不能对它们进行各种算术运算。

3. 多项选择题(多选题)的编码

多项选择题在社会科学领域中有广泛的应用。所谓多项选择题即题目的答案不止一个,答案的选项可以是多选项,多选项又可分为限定多选项和任意多选项两种,例如表 2.1 第 3 题是一个限定多选项(限选 3 项)题,第 4 题是一个任意多选项题。因此,多选项的编码方法可分为两种:

(1) 限定多选项分类法

对于限定多选项的编码,不能直接对题目进行编码,因为此类问题的选择,一次最多可以选择限定多选项的最大数,例如表 2.1 第 3 题是一个限定 3 选项题目,一次可以选择 1 个选项,也可以选择 2 个选项,最多可以选择 3 个选项。因此这类问题的编码方法是首先估算这类题目可被选择的最多选项,并给每个选项建立一个变量,然后给每个新变量用一个代码(如 1,2,3,…)表示。表 2.1 第 3 题的编码见表 2.2,如某被试同时选择“1 相貌”和“2 文化水准”,则变量“第 3 题-1”=1,“第 3 题-2”=2;又如某被试同时选择“2 文化水准”、“4 志同道合”和“5 人品”,则变量“第 3 题-1”=2,“第 3 题-2”=4,“第 3 题-3”=5,依此类推。

^① 为了确保数据的准确性和提高数据录入的效率,在数据录入过程中,此类数据多不作二次编码。可以将被试的原始选项直接录入(如录入选项 A、B、C),原始选项赋予分值由 SPSS 系统进行转换,具体操作见本章 2.3 节的内容。

(2) 任意多选项二分法

任意多选项就意味着每个被试一次可以任选 1 个选项、2 个选项、3 个选项等等,因此这类问题的编码方法为把每一个被选项作为一个变量来定义,每个变量只有“0”或“1”两种选择,“0”为不选,“1”为选中。例如表 2.1 第 4 题是一个任选多项题,被选项有 3 个,为此定义 3 个新变量(“第 4 题—工作”、“第 4 题—学校”、“第 4 题—环境”),每个变量只有“0”或“1”两种选择,具体编码见表 2.2 第 4 题的编码。

4. 缺失值的处理

缺失值分为用户缺失值(User Missing Value)和系统缺失值(System Missing Value)。用户缺失值指在问卷调查中,把被试不回答的一些选项当作缺失值来处理。用户缺失值的编码一般用研究者自己能够识别的数字来表示,如“0”、“9”、“99”等。系统缺失值主要指计算机默认的缺失方式,如果在输入数据时空缺了某些数据或输入了非法的字符,计算机就把其界定为缺失值,这时的数据标记为“.”。缺失值往往会给统计分析带来一些麻烦和误差,尤其在时间序列分析中更是如此。一般地,对缺失值的处理可采用如下方法:第一,替代法。即采用统计命令 Transform→Replace Missing Values 进行替代,或在相关统计功能中利用其【Options】等参数进行替代。第二,剔除法。即剔除有缺失值的题目,或剔除有缺失值的整份问卷。

2.2 数据的录入

2.2.1 操作术语

为了能够顺利地使用 SPSS 处理数据,我们必须掌握如下概念。

① 个案(Cases)。在数据处理中,一个研究对象就是一个个案,也相当于一个记录,在数据表格中表现为“一行”。每一个个案记录的是一个研究对象各个属性的具体数值,如某学生的所有信息(姓名、性别、年龄、各科学业成绩等等)。

② 样本(Sample)。样本是指具有共同属性的所有研究对象,如某学校一年级学生的所有信息。样本含多个个案,在数据表格中表现为“n 行”。

③ 变量(Variable)。问卷中的每个问题,在 SPSS 中称之为“变量”,相当于 dBASE 数据库中的“字段(Field)”,在数据表格里表现为“一列”。

④ 变量值(Value)。问卷中的答案就是变量值,也称为观测值。在 SPSS 系统里,单元格中的数值就是变量值。

2.2.2 定义变量

启动 SPSS 后便进入数据编辑窗口(Data View),见图 2.1。

由于当前没有输入数据,显示的是一个空文件,因此输入数据前首先要定义变量。定义变量即要定义变量名(Name)、变量类型(Type)、宽度(Width)、小数位

数(Decimals)、变量标签(Label)、数值标签(Values)、缺失值(Missing Values)、单元格列长度(Columns)、单元格字符排列方向(Align)和数据量度(Measure)。

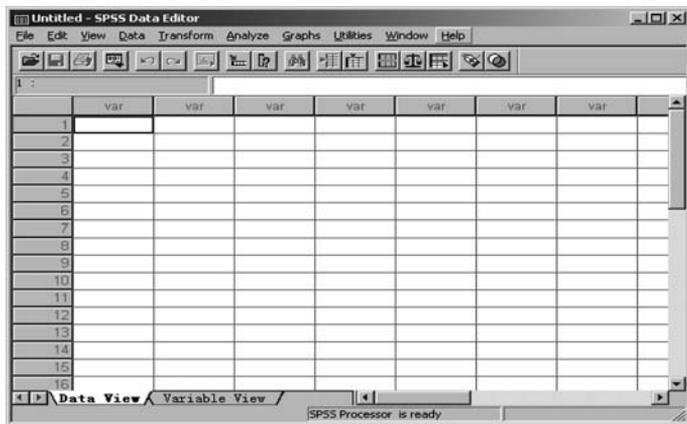


图 2.1 数据编辑窗口

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 所示的变量定义窗口,在此窗口即可定义变量。



图 2.2 变量定义窗口

1. 定义变量名

在 Name 下的单元格中输入变量名,如图 2.2 中的 id、name、sex、cj。变量名必须是以字母为首,其长度不能超过 8 个字母数的字符串。SPSS 的保留关键词不能作为变量名。如: AND、OR、NOT、WITH、TO、BY、GT、GE、LT、LE、EQ、NE、ALL 等。在 SPSS 系统里,也可以用中文作变量名,不过尽量不要使用它,这里涉及一个兼容性的问题。

2. 定义变量类型、宽度及小数位数

在 Type 下选择变量类型,单击该单元格,出现图标,再单击这个图标中的按钮,打开如图 2.3 所示的 Variable Type(变量类型)对话框,从中选择变量类型。系统默认值为数值型。变量类型有数值型(Numeric)、加显逗号的数值型(Com-

ma)、3 位加点点数值型(Dot)、科学计数型(Scientific notation)、日期型(Date)、货币型(Dollar)、用户自定义型(Custom currency)、字符型(String)。

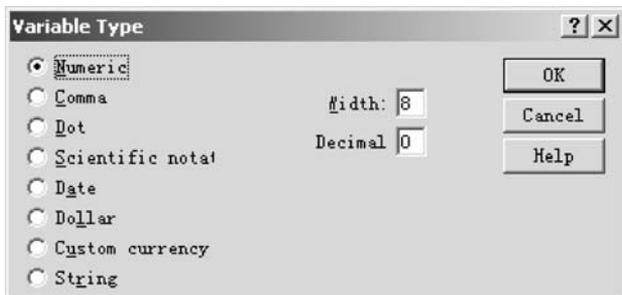


图 2.3 变量类型对话框

在 Width、Decimal 下选择宽度和小数位数,系统默认宽度为 8,小数位数为 2。单击 Width、Decimal 下面的某一单元格,将出现 , 点击箭头可以调大或调小变量的宽度和小数位数的数值,或者在数字框里直接输入数字即可。

3. 定义变量标签

在 Label 下的单元格中,输入标签,如图 2.2“编号”。SPSS 允许定义长达 255 字节的变量标签。

4. 定义数值标签

数值标签是对变量的每一可能取值进一步描述,当变量是定性或定序变量时,这是非常有用的。系统默认为 None(无),当需要定义数值标签时,单击含有“None”的单元格,便进入图 2.4 的“数值标签”窗口。

给性别变量 sex 的取值定为“1”表示“男”,“2”表示“女”,则在第一个“Value”框中输入“1”,再在第二个“Value”框中输入“男”,单击 Add 按钮,即可定义“1 = ‘男’”。依此类推,最后点击 OK 即可。

5. 定义缺失值

SPSS 有系统缺失值和用户缺失值两类缺失值,系统默认为 None(无)。当需要定义缺失值时,单击 Missing 下的含有“None”单元格,便进入图 2.5 的“缺失值”窗口。缺失值有以下三种选项。



图 2.4 数值标签窗口

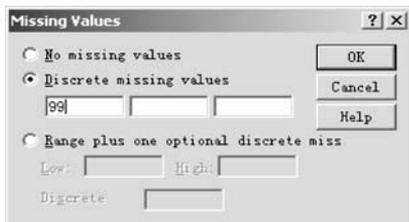


图 2.5 缺失值窗口

No missing values:没有缺失值。

Discrete missing values:定义 1~3 个单一数为缺失值。

Range plus one optional discrete missing values:定义指定范围为缺失值,同时指定另外一个不在这一范围的单一数为缺失值。

至于其他如单元格列长度(Columns)、单元格字符排列方向(Align)和数据量度(Measure)等均是不常用,一般使用系统默认值就可以了,以便减少工作量。

6. 变量定义的信息复制

如果有多个变量的类型相同,可以先定义一个变量,然后把该变量的定义信息复制给新变量,具体操作为:先定义好一个变量,在该变量的行号上单击右键,在弹出的快捷菜单中选择“Copy”命令,然后选择其他同类型所在行,单击鼠标右键,在弹出的快捷菜单中选择“Paste”,这样就自动复制了多个变量的定义信息。

2.2.3 数据的录入

定义了所有变量后,单击“Data View”标签,便可在数据编辑窗口中输入数据。数据编辑窗口中黑框所在的单元格为当前的数据单元格,表示用户正在对该数据单元格录入数据或正在修改该单元格中的数据,因此,在录入数据时,用户应首先将黑框移到想要输入数据的单元格上。

数据录入时可以逐行录入,即录入完一个数据后,按 Tab 键,黑框自动移到本行的下一个变量列上。也可以逐列录入,即按照变量录入数据,录入完一个数据后,按回车键,黑框自动移到本列的下一行上。

在数据录入过程中,我们应该把握一个原则即录入的数据应是最原始状态的数量化数据,而不是经过各种计算后产生的数据。

在数据录入过程中,为了避免数据意外丢失,应及时保存文件。可通过命令菜单 File→Save 或使用快捷命令 Ctrl+S 来保存文件。第一次保存文件时,将会出现图 2.6 所示的对话框。

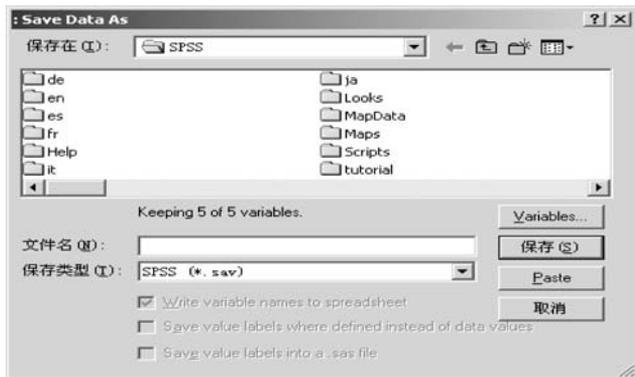


图 2.6 数据文件保存窗口

其中,“保存在(I)”右边的框中显示的是数据文件要保存的文件夹地址。“文件名(N)”框显示的是数据文件存盘的文件名,“保存类型(T)”框显示数据文件保存的类型,系统默认为“.sav”,一般不用改变。所有设定完成后,点击“保存”按钮,数据文件将保存在指定的文件夹里。然后,系统自动回到 SPSS 数据编辑窗口。

2.2.4 数据的导入

除了直接录入数据以外,SPSS 还能够导入各种格式的数据,因为 SPSS 提供强大的与其他软件包共享数据文件的功能,较好地扩展了 SPSS 的用途。

能够与 SPSS 共享数据文件的软件包有:

- * .sys:SPSS/PC+,DOS 版的 SPSS;
 - * .por:SPSS 可移植文件;
 - * .xls:Microsoft 公司的 Excel 格式的数据文件;
 - * .slk:Multiplan 公司的 symbolic 格式的数据文件;
 - * .dbf:数据库 dBase、Foxbase 等格式的数据文件;
 - * .w * :Lotus 公司的 1-2-3 格式的数据文件;
 - * .dat:Tab-Delimited 自由格式的纯文本数据文件;
 - * .syd:Systat 软件包的数据文件;
 - * .sys:Systat 软件包的系统文件;
 - * .sas7dbat:SAS 软件包的长文件名的数据文件;
- 等等。

导入其他软件包数据文件的操作如下:

① 执行菜单命令“File”→“Open”→“Data”,出现图 2.7 所示对话框。

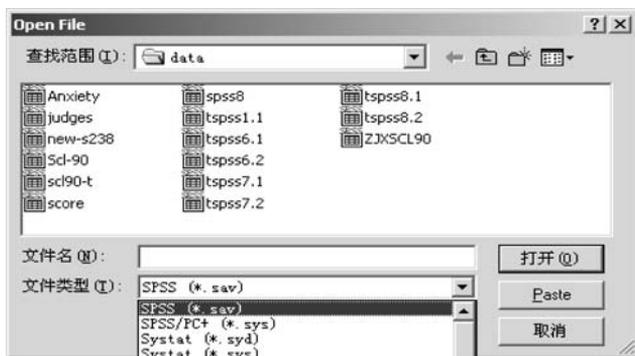


图 2.7 打开数据文件对话框

② 在“文件类型(T)”框中提供了 SPSS 共享的软件包数据文件,只要打开该框中下拉列表框,如图 2.7 所示,选择对应的数据文件格式,然后点击“打开(O)”按钮即可。

这里,提供一种实用的数据导入方法,即通过 Windows 操作系统提供的复