

宝典丛书

300万

SPSS

(第3版)

宝典

本书基于最新的IBM SPSS Statistics 19版本, 涉及很多新特性, 具有技术前瞻性

不深入理论讲解, 集中介绍如何使用SPSS解决各种实际问题

结合具体的应用案例介绍技术, 有很强的项目背景, 读者可以从中学会统计的思想



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

杨 丹 编著

宝典丛书

SPSS 宝典

(第3版)

杨 丹 编著

电子工业出版社

Publishing House of Electronics Industry

北京 · BEIJING

内 容 简 介

本书基于最新版本 IBM SPSS Statistics 19 进行编写, 结合统计教学的特点, 以循序渐进的方式介绍了世界著名统计分析软件 SPSS 的多种模块和编程功能, 并对数据处理和统计分析的输出结果进行了详尽的解释。

本书共分 23 章, 主要介绍了 IBM SPSS Statistics 19 的基础知识、统计数据的创建和管理、SPSS 统计分析功能、SPSS 的图形绘制功能以及 SPSS 编程功能。其中包括 SPSS 的窗口及其设置、统计数据的创建与编辑、SPSS 数据的管理、数据转换与 SPSS 函数、SPSS 基本统计分析、多重反应分析、SPSS 的自定义表格、均值的比较与检验、方差分析、非参数检验、相关分析、回归分析、对数线性模型、聚类分析、判别分析、因子分析、对应分析、信度分析、统计图形的创建和编辑、SPSS 的命令语句程序设计、利用 SPSS 语句读取数据文件、宏等内容。

本书内容全面, 论述翔实, 深入浅出。全书以 SPSS 统计功能为主线, 涵盖数据管理和 SPSS 高级编程等内容, 既可供高等院校相关专业本科生、研究生, 以及从事统计分析和决策的各领域相关专业的读者学习参考, 亦可用作 SPSS 统计分析培训和自学教材。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有, 侵权必究。

图书在版编目 (CIP) 数据

SPSS 宝典 / 杨丹编著. —3 版. —北京: 电子工业出版社, 2013.10

(宝典丛书)

ISBN 978-7-121-21297-0

I. ①S… II. ①杨… III. ①统计分析—软件包 IV. ①C819

中国版本图书馆 CIP 数据核字 (2013) 第 195018 号

责任编辑: 徐津平

印 刷: 三河市双峰印刷装订有限公司

装 订: 三河市双峰印刷装订有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱

邮编: 100036

开 本: 787×1092 1/16

印张: 33

字数: 920 千字

印 次: 2013 年 10 月第 1 次印刷

印 数: 3000 册 定 价: 79.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zltts@phei.com.cn, 盗版侵权举报请发邮件到 dbqq@phei.com.cn。

服务热线: (010) 88258888。

前 言

在当今信息化时代背景下，无论是个人、政府或是企业，都需要在海量的信息中获取有价值的信息，并据此做出科学的评估和决策。因此，对信息的采集、处理和分析，并给出专业人士可接受的评估和预测报告等工作就变得十分重要。SPSS 正是为此功能设计了一整套集数据处理、评估和预测的解决方案。

SPSS (Statistical Product and Service Solutions, 统计产品与服务解决方案) 是当今世界上应用最为广泛的统计分析软件之一。尤其是近几年的发展更为迅速，并以每年一个新版本的速度进行更新，特别是 IBM SPSS Statistics 19 版本，在操作界面、数据管理、统计分析、图形处理以及可扩展编程能力和中文操作的兼容等方面都有很大的改进。

近几年，SPSS 的产品和服务已成功进入国内的多个政府部门和大型企业，并取得了卓有成效的业绩。国内科研院所的很多专业已经将会利用 SPSS 进行数据处理和统计决策分析作为培养目标的必备技能之一。

为了帮助学习 SPSS 的人员提高统计分析的理论水平和软件操作应用能力，笔者精心编著了本书。本书依照统计分析的学习规律，首先介绍了 SPSS 的基本概念和基本操作，在读者具有一定数据管理和初步分析能力的基础上，再对 SPSS 的统计分析模块进行分门别类的讲解，尽量做到统计分析理论和操作方法的完美结合；同时，针对读者的不同需求，本书利用数据管理和统计分析示例，介绍了 SPSS 的高级编程功能，以帮助高级用户对 SPSS 的统计分析功能进行二次开发。

与其他书籍相比，本书的特色之处

1. **版本新**：本书讲解的是最新的 IBM SPSS Statistics 19 版本，其中提到了该版本的很多新特性，具有技术前瞻性。
2. **内容实用**：很多书籍讲解了很多统计方面的理论基础，笔者认为，这是没有必要的。我们默认读者已经掌握了统计学的相关知识，但是不会使用 SPSS 软件。在这种情况下，本书集中篇幅讲解 SPSS 的各种使用方法，更有针对性，力争做到用更少的页数讲解更多的 SPSS 使用方法。
3. **实例有针对性**：介绍某个技术时，都是结合具体的应用实例来讲解的，有很强的项目背景，读者可以从中学会统计的思想。
4. **参数介绍详细**：本书的目标是成为一本最为全面的 SPSS 书籍，因此在介绍各个参数的时候，力求全面、详细，让读者可以作为参考手册使用。

本书包括的内容

本书基于最新的 IBM SPSS Statistics 19 版本编写，共分 23 章，包括 SPSS 入门和利用 SPSS 进行统计分析两大部分。

第1章首先带领读者了解 SPSS 的发展历程,了解 IBM SPSS Statistics 19 版本的新增功能,熟悉 SPSS 的帮助界面,并指导读者安装 SPSS 软件。

第2章即切入正题,介绍了 SPSS 的窗口、运行模式、各种对话框的特点和功能,使读者对 SPSS 有一个初步、直观的了解。

第3章开始了 SPSS 数据管理的第一步,创建和编辑数据文件。包括 SPSS 数据文件的结构、数据的编辑和读取其他格式数据文件的方法。

第4章介绍了数据管理的核心内容,包括数据的排序、拆分、合并以及分类汇总等操作。

第5章是数据转换,介绍了 SPSS 的内部函数和数据转换的方法,与此前的第2章到第4章构成了 SPSS 的数据管理部分,这是进行统计分析之前必经的一个步骤,也是 SPSS 软件的有机组成部分。

第6章开始介绍 SPSS 的统计分析功能。这一章给出了 SPSS Descriptive Statistics 模块的基本功能,包括统计分析中最常用的统计量和探索分析。

第7章介绍了 SPSS 的多重反应分析,包括多重反应分析的概念和 SPSS 操作。对于从事与问卷调查相关工作的读者而言,本章值得一读。

第8章介绍了 SPSS 自定义表格的方法。SPSS 的优势不仅在于强大的统计分析功能,制作条理清晰的表格和高质量的图形也是 SPSS 一个不可忽视的亮点。

第9章到第19章是全书的核心内容,即 SPSS 统计分析功能介绍,包括均值的比较与检验、方差分析、非参数检验、相关分析、回归分析、对数线性模型、聚类分析、判别分析、因子分析、对应分析、信度分析等常用的专业统计分析功能。不仅给出了统计方法的原理介绍和 SPSS 操作方法,同时还提供了大量精心选择的分析示例,涵盖金融、经济、医药卫生、社会科学、工程等领域。

第20章是统计图形的创建和编辑,介绍了绘制 SPSS 各种图形的方法和操作示例。

第21章介绍了 SPSS 的命令语句程序设计方法,包括 SPSS 命令的定义规则、SPSS 语句的控制流程结构和 SPSS 命令语句的语法规则,并结合 SPSS 的统计分析功能,给出统计分析的命令语句程序。

第22章介绍了 SPSS 命令语句在数据管理方面的实例分析,主要包括利用 SPSS 命令语句读取和创建各种格式的数据文件。

第23章简要介绍了 SPSS 的宏工具。最后3章内容是 SPSS 的程序编制功能,这也是 SPSS 另一独具特色的功能,与前面介绍的数据管理功能、统计分析功能和图形绘制功能共同构成了 SPSS 的四大功能,这也是 SPSS 不同于其他统计分析软件的特别之处。

适合阅读本书的读者

本书内容和实例适用于金融、经济、生物、医疗卫生保健和社会人文等多个学科的需求,既可供高等院校相关专业本科生、研究生以及从事统计分析的专业人士学习参考,亦可用作 SPSS 统计分析培训和自学教材。

本书的相关资源请到博文视点公司网站(www.broadview.com.cn)的“资源下载”栏目下载。

编著者

2013年6月

目 录

第 1 部分 IBM SPSS 19 入门..... 1	
第 1 章 IBM SPSS Statistics 概述.....2	
1.1 SPSS 的发展简史..... 2	
1.2 IBM SPSS Statistics 19 的特点..... 3	
1.3 IBM SPSS Statistics 19 对环境的 要求..... 3	
1.3.1 对硬件环境的要求..... 3	
1.3.2 对软件环境的要求..... 4	
1.4 安装和卸载 IBM SPSS Statistics 19.... 4	
1.4.1 安装 IBM SPSS Statistics 19... 4	
1.4.2 卸载 IBM SPSS Statistics 19... 5	
1.5 IBM SPSS Statistics 19 的启动与 退出..... 6	
1.5.1 启动 IBM SPSS Statistics 19.... 6	
1.5.2 退出 IBM SPSS Statistics 19.... 6	
1.6 IBM SPSS Statistics 19 新增功能..... 7	
1.7 IBM SPSS Statistics 19 的帮助系统..... 8	
1.7.1 Topics (主题帮助)..... 8	
1.7.2 Tutorial (自学指导)..... 9	
1.7.3 Cases Studies (案例研究)..... 9	
1.7.4 Statistics Coach (统计辅导)..... 10	
1.7.5 Command Syntax Reference (命令语句参考)..... 10	
1.7.6 Developer Central (研发中心)..... 11	
1.7.7 Algorithms (统计算法) ... 11	
1.8 小结..... 12	
第 2 章 IBM SPSS 19 的窗口及其设置..... 13	
2.1 IBM SPSS Statistics 19 的三种 运行方式..... 13	
2.1.1 完全窗口菜单运行方式..... 13	
2.1.2 程序运行方式..... 13	
2.1.3 混合运行方式..... 13	
2.2 IBM SPSS Statistics 19 常用对 话框..... 14	
2.2.1 SPSS 文件操作对话框..... 14	
2.2.2 SPSS 统计分析主对话框..... 16	
2.2.3 对话框的控制与选择..... 19	
2.3 IBM SPSS Statistics 19 的窗口及 其设置..... 21	
2.3.1 Data Editor (数据编辑 窗口)..... 22	
2.3.2 Output (结果输出窗口)..... 22	
2.3.3 SPSS Syntax Editor (语 句窗口)..... 23	
2.3.4 Script (脚本编辑窗口)..... 24	
2.4 小结..... 24	
第 3 章 SPSS 统计数据的创建与编辑..... 25	
3.1 SPSS 数据的属性及其定义方法..... 25	
3.1.1 变量名 (Name)..... 25	
3.1.2 变量类型 (Type)..... 25	
3.1.3 变量标签 (Label)..... 31	
3.1.4 Missing (变量缺失数据) ... 32	
3.1.5 Measure (度量尺度)..... 33	
3.2 SPSS 数据文件的构成..... 33	
3.2.1 SPSS 数据文件的结构..... 33	
3.2.2 SPSS 数据文件中的个案..... 34	
3.2.3 创建 SPSS 数据文件..... 34	
3.2.4 应用实例: 定义 SPSS 数 据结构..... 37	
3.2.5 实例: 大学教师的问卷调 查表..... 37	
3.2.6 调查表数据变量名及其属 性的设计..... 37	



3.3	SPSS 数据的录入	39	5.2.2	统计函数 (Statistical Functions)	79
3.3.1	认识数据录入窗口	39	5.2.3	字符串函数 (String Functions)	80
3.3.2	在数据编辑窗口中录入数据	40	5.2.4	日期和时间函数 (Date and Time Functions)	80
3.4	SPSS 数据的编辑	41	5.2.5	累计分布函数 (Cumulative Distribution Functions)	81
3.4.1	插入和删除变量	41	5.2.6	逆分布函数 (Inverse Distribution Functions)	82
3.4.2	插入和删除个案	42	5.2.7	随机变量函数 (Random Variable Functions)	83
3.4.3	数据的移动、复制和删除 ..	43	5.2.8	概率密度函数 (Probability Density Function)	84
3.4.4	SPSS 数据的定位	44	5.2.9	尾部概率函数 (Tail Probability Function)	85
3.5	数据文件的操作	45	5.2.10	缺失值函数 (Missing Value Functions)	86
3.5.1	创建和读取 SPSS 数据文件	45	5.3	变量的计算	86
3.5.2	保存 SPSS 数据文件	52	5.3.1	变量计算的基本操作	86
3.6	小结	54	5.3.2	实例: 高校提前录取名单的确定	89
第 4 章	SPSS 数据的管理	55	5.4	数据的选取	92
4.1	数据的排序	55	5.4.1	数据选取的基本操作	92
4.1.1	SPSS 数据排序的基本操作 ..	55	5.4.2	数据选取的应用实例	95
4.1.2	数据排序的一个例子	56	5.5	个案计数	97
4.2	数据的转置	57	5.5.1	个案计数的基本操作	97
4.3	重复个案的识别	59	5.5.2	实例: 央视主要节目的收视率调查	99
4.4	数据文件的拆分	62	5.6	数据的重新编码	100
4.5	数据文件的合并	65	5.6.1	数据的重新编码	100
4.5.1	个案合并	65	5.6.2	数据的自动重新编码	104
4.5.2	变量合并	67	5.6.3	重新编码的应用实例	105
4.6	数据的分类汇总	70	5.7	SPSS 变量集	108
4.6.1	分类汇总的概念	70	5.7.1	SPSS 变量集的分类	108
4.6.2	分类汇总的操作步骤	70	5.7.2	定义用户变量集	108
4.7	数据的加权	74	5.7.3	使用用户变量集	109
4.8	小结	76	5.8	小结	110
第 5 章	数据转换与 SPSS 函数	77			
5.1	SPSS 的表达式	77			
5.1.1	SPSS 算术表达式	77			
5.1.2	SPSS 条件表达式	77			
5.1.3	SPSS 逻辑表达式	78			
5.2	SPSS 函数	78			
5.2.1	算术函数 (Arithmetic Functions)	79			



第 2 部分 利用 SPSS 进行统计分析	111
第 6 章 SPSS 的基本统计分析	112
6.1 基本统计量的定义和计算	113
6.1.1 描述集中趋势的统计量	113
6.1.2 描述离散趋势的统计量	113
6.1.3 描述分布形态的统计量	114
6.2 频数分析过程	115
6.2.1 频数分析过程的 SPSS 操作	115
6.2.2 实例：成年男子血压数据的频数分析	118
6.2.3 频数分析的结果	119
6.3 描述性分析过程	120
6.3.1 描述性分析的 SPSS 操作	120
6.3.2 实例：学生体检数据的描述性分析	122
6.4 探索性分析过程	124
6.4.1 探索性分析问题的提出	124
6.4.2 探索分析的 SPSS 操作	124
6.4.3 Statistic 选项	125
6.4.4 Plots 选项	126
6.4.5 Options 选项	127
6.4.6 实例：新型合金耐热性的探索分析	128
6.4.7 探索分析的结果	128
6.5 小结	134
第 7 章 多重反应分析	135
7.1 多重反应分析的基本方法	135
7.1.1 二分法 (Multiple Dichotomies Method)	135
7.1.2 分类法 (Multiple Category Method)	136
7.2 定义多重反应分析数据集	136
7.2.1 实例：消费者购物场所的调查	136
7.2.2 定义多重数据集的 SPSS 操作	136
7.3 多重反应频数分析	138
7.3.1 多重反应频数分析的 SPSS 操作	138
7.3.2 多重反应频数分析的结果	139
7.4 多重反应列联表分析	139
7.4.1 SPSS 基本操作	139
7.4.2 多重反应列联表分析的结果	141
7.5 小结	142
第 8 章 SPSS 的自定义表格	143
8.1 认识自定义表格对话框	143
8.1.1 实例：个人家庭情况调查数据	143
8.1.2 自定义表格的 SPSS 操作	144
8.2 表格的结构和 SPSS 实现	149
8.2.1 堆栈和嵌套的基本定义	149
8.2.2 堆栈和嵌套的 SPSS 实现	149
8.3 自定义表格的选项及其独立性检验	151
8.3.1 实例：对婚姻幸福感认同的分析和列表	151
8.3.2 SPSS 基本操作	151
8.3.3 Titles 选项	151
8.3.4 Test Statistics 选项	152
8.3.5 Options 选项	153
8.3.6 自定义表格的结果	155
8.4 多重反应变量集的自定义表格	156
8.4.1 一般多重变量集自定义表格	156
8.4.2 SPSS 基本操作	156
8.4.3 多重反应变量集表格的结果	158
8.4.4 进一步的分析	159
8.4.5 具有嵌套结构的多重变量集自定义表格	160
8.4.6 多重变量集的独立性检验	161
8.4.7 多重变量集的列均值比较和检验	163
8.5 小结	164



第9章 均值的比较与检验	165	10.3.5 实例：来自4个不同行业的投诉数据	186
9.1 参数检验问题的提出	165	10.3.6 单因素方差分析的结果解释	187
9.2 参数检验的基本步骤	166	10.4 多因素方差分析	192
9.3 单样本 T 检验	167	10.4.1 多因素方差分析的 SPSS 操作	192
9.3.1 单样本 T 检验的基本方法	167	10.4.2 Model 选项	193
9.3.2 单样本 T 检验的 SPSS 操作	168	10.4.3 Contrasts 选项	194
9.3.3 实例：圆盘制动闸直径的 T 检验	169	10.4.4 Plots 选项	195
9.3.4 单样本 T 检验的结果	170	10.4.5 Post Hoc Multiple Comparisons 选项	195
9.4 独立样本 T 检验	171	10.4.6 Save 选项	196
9.4.1 独立样本 T 检验的基本方法	171	10.4.7 Options 选项	197
9.4.2 独立样本 T 检验的 SPSS 操作	172	10.4.8 实例：机器和工人两因素的方差分析	198
9.4.3 实例：两组乒乓球得分数据的 T 检验	174	10.4.9 两因素方差分析的结果解释	199
9.4.4 独立样本 T 检验的结果	174	10.5 协方差分析	201
9.5 配对样本 T 检验	175	10.5.1 协方差分析的基本思想	201
9.5.1 配对样本 T 检验问题的基本方法	175	10.5.2 实例：政府部门对培训效果的分析	201
9.5.2 配对样本 T 检验的 SPSS 操作	176	10.5.3 非饱和模型的 SPSS 操作	202
9.5.3 实例：新药疗效的测试结果检验	177	10.5.4 结果分析	203
9.5.4 配对样本 T 检验的结果	178	10.5.5 进一步分析：实例中的饱和模型	204
9.6 小结	180	10.5.6 饱和模型的结果解释	205
第10章 方差分析	181	10.6 小结	207
10.1 方差分析的基本原理	181	第11章 非参数检验	208
10.2 方差分析的概念和假设	182	11.1 卡方检验	209
10.3 单因素方差分析	183	11.1.1 卡方检验的基本原理	209
10.3.1 单因素方差分析的 SPSS 操作	183	11.1.2 卡方检验的 SPSS 操作	209
10.3.2 Contrasts 选项	183	11.1.3 实例：网站流量的泊松分布检验	213
10.3.3 Post Hoc Multiple Comparisons 选项	184	11.2 二项分布检验	215
10.3.4 Options 选项	186	11.2.1 二项分布检验的基本原理	215
		11.2.2 二项分布检验的 SPSS 操作	216



11.2.3 实例：抛硬币实验的二项分布检验.....	216	12.3.6 进一步分析的结果.....	239
11.3 游程检验.....	218	12.4 偏相关分析.....	240
11.3.1 游程检验的基本原理.....	218	12.4.1 偏相关分析的基本原理.....	240
11.3.2 游程检验的 SPSS 操作.....	218	12.4.2 偏相关分析的 SPSS 操作步骤.....	240
11.3.3 实例：体育达标成绩的检验.....	219	12.4.3 Options 选项.....	241
11.4 单样本 K-S 检验.....	221	12.4.4 实例：对一批体检数据的偏相关分析.....	242
11.4.1 单样本 K-S 检验的基本原理.....	221	12.4.5 偏相关分析的结果.....	243
11.4.2 单样本 K-S 检验的 SPSS 操作.....	221	12.4.6 改变控制变量后的结果.....	244
11.4.3 实例：车祸事故数的泊松分布检验.....	222	12.5 距离分析.....	245
11.4.4 实例的进一步分析：考虑性别因素.....	223	12.5.1 距离分析的基本概念.....	245
11.5 两独立样本的非参数检验.....	225	12.5.2 距离分析的 SPSS 操作.....	245
11.5.1 两独立样本检验的 SPSS 操作.....	225	12.5.3 Similarity Measure 选项和相似性测度.....	246
11.5.2 两独立样本 Mann-Whitney U 检验.....	227	12.5.4 Dissimilarity Measure 选项和不相似性测度.....	248
11.5.3 两独立样本的 K-S 检验.....	229	12.5.5 实例：体检数据的变量距离分析.....	250
11.6 小结.....	231	12.5.6 变量距离分析的结果.....	250
第 12 章 相关分析.....	232	12.5.7 实例：对飞机叶片的个案距离分析.....	251
12.1 相关分析的基本方法.....	232	12.5.8 个案距离分析的结果.....	252
12.2 相关系数的计算.....	233	12.6 小结.....	252
12.2.1 Pearson 简单相关系数.....	233	第 13 章 回归分析.....	253
12.2.2 Spearman 等级相关系数.....	233	13.1 线性回归分析.....	254
12.2.3 Kendall τ 相关系数.....	234	13.1.1 线性回归分析的方法概述.....	254
12.3 双变量相关分析.....	234	13.1.2 线性回归分析的数学模型.....	254
12.3.1 双变量相关分析的 SPSS 操作.....	234	13.1.3 线性回归方程的显著性检验.....	255
12.3.2 Options 选项.....	235	13.1.4 线性回归方程的残差分析.....	256
12.3.3 实例：汽车价格和汽车燃油率的分析.....	236	13.1.5 线性回归分析的 SPSS 操作.....	257
12.3.4 双变量相关分析的结果.....	237	13.1.6 实例：体检数据中的体重和肺活量的分析.....	263
12.3.5 进一步分析：定序型变量的分析.....	238	13.1.7 一元线性回归分析的结果.....	264
		13.1.8 多元线性回归：小学生语言测试得分分析.....	269



13.1.9	多元线性回归的结果	270
13.1.10	实例: 加权最小二乘回归...	273
13.1.11	加权最小二乘回归分析的结果.....	275
13.2	曲线参数估计	277
13.2.1	曲线参数估计的基本原理...	277
13.2.2	曲线参数估计的基本步骤...	277
13.2.3	实例: 金属强度测试的回归分析.....	280
13.2.4	曲线参数估计的结果	281
13.3	小结	285
第 14 章	对数线性模型	286
14.1	对数线性模型的基本概念	286
14.1.1	实例: 育龄妇女生育调查	286
14.1.2	对数线性模型的对数频数表.....	287
14.2	对数线性模型的建立	288
14.2.1	对数线性模型的效应项	288
14.2.2	效应项的计算方法	289
14.3	对数线性模型的检验	289
14.3.1	对数线性模型的检验项目	289
14.3.2	对数线性模型的参数估计	290
14.3.3	单项效应的参数估计和检验.....	290
14.3.4	拟合优度检验	291
14.3.5	交互效应检验	291
14.3.6	分层效应检验	292
14.3.7	饱和模型的偏关联检验	293
14.4	对数线性模型的 Model Selection 过程.....	293
14.4.1	层次对数线性模型的基本方法.....	293
14.4.2	层次对数线性模型的 SPSS 操作.....	294
14.4.3	Model Building 选项和模型的选择.....	294
14.4.4	Model 选项	295
14.4.5	Options 选项	296
14.4.6	实例: 育龄妇女生育调查	297
14.4.7	实例分析的结果	298
14.5	对数线性模型的 General 过程	304
14.5.1	General 过程的基本思想	304
14.5.2	General 过程的 SPSS 步骤	304
14.5.3	Save 选项	305
14.5.4	Model 选项	305
14.5.5	Options 选项	306
14.5.6	实例分析	307
14.5.7	General 对数线性模型的结果与分析	308
14.6	对数线性模型的 Logit 过程	311
14.6.1	Logit 过程的基本思想	311
14.6.2	Logit 过程的 SPSS 操作	311
14.6.3	实例分析	312
14.6.4	Logit 模型的结果与分析	312
14.7	小结	315
第 15 章	聚类分析	316
15.1	聚类分析的基本方法	316
15.1.1	实例: 不同学科的能力测试调查	316
15.1.2	距离描述	317
15.1.3	聚类类型	317
15.1.4	聚类方法	318
15.2	不相似测度的度量方法	318
15.2.1	定距型变量的不相似测度	318
15.2.2	计数变量的不相似测度	320
15.2.3	二值变量的不相似测度	320
15.3	分层聚类	322
15.3.1	分层聚类的基本方法	322
15.3.2	分层聚类的 SPSS 操作	323
15.3.3	Method 选项	324
15.3.4	Save 选项	327
15.3.5	Statistics 选项	328



15.3.6	Plots 选项	330	16.2.5	Save 选项	365
15.3.7	个案聚类: 16 个地区的 农民支出情况分析	332	16.3	实例: 对一组体检数据的 判别分析	366
15.3.8	个案聚类的结果分析	334	16.3.1	实例数据	366
15.3.9	确定分类数的讨论	337	16.3.2	操作步骤	366
15.3.10	变量聚类: SPSS 软件模 块使用的调查	338	16.4	判别分析的结果	367
15.3.11	变量聚类的结果分析	339	16.5	小结	371
15.4	逐步聚类分析	341	第 17 章	因子分析	372
15.4.1	逐步聚类分析的基本方法	341	17.1	因子分析的基本概念	372
15.4.2	逐步聚类分析的 SPSS 操作	342	17.2	因子分析的数学模型	373
15.4.3	Iterate 选项	343	17.3	因子分析的基本方法	373
15.4.4	Save 选项	344	17.3.1	因子提取	373
15.4.5	Options 选项	344	17.3.2	因子旋转	373
15.4.6	个案逐步聚类分析的结果	345	17.3.3	计算因子得分	373
15.4.7	逐步聚类: 变量聚类模 式的实例	348	17.4	因子分析的 SPSS 操作	374
15.4.8	变量逐步聚类的结果 分析	349	17.4.1	基本操作	374
15.5	二阶段聚类分析	351	17.4.2	Descriptives 选项	375
15.5.1	二阶段聚类分析的基本 原理	351	17.4.3	Extraction 选项	376
15.5.2	二阶段聚类分析的 SPSS 操作	352	17.4.4	Rotation 选项	377
15.5.3	Options 选项	353	17.4.5	Scores 选项	378
15.5.4	Output 选项	355	17.4.6	Options 选项	379
15.5.5	个案聚类: 不同汽车车 型的聚类分析	355	17.4.7	因子分析的共同度	380
15.5.6	二阶段个案聚类的结果	356	17.4.8	因子分析中的总方差 解释	380
15.6	小结	359	17.4.9	因子分析的碎石图和 解释	381
第 16 章	判别分析	360	17.4.10	旋转后的因子载荷矩阵	382
16.1	判别分析过程的基本原理	360	17.4.11	因子得分	382
16.2	判别分析的 SPSS 操作	360	17.5	实例: 奥运项目的因子分析	383
16.2.1	基本操作	360	17.5.1	操作步骤	383
16.2.2	Method 选项	362	17.5.2	结果分析	384
16.2.3	Statistics 选项	363	17.6	小结	388
16.2.4	Classify 选项	364	第 18 章	对应分析	389
			18.1	对应分析的基本方法	389
			18.2	对应分析的 SPSS 操作	390
			18.2.1	实例: 不同职业人员的 吸烟程度调查	390
			18.2.2	对应分析的基本操作	391



18.2.3	Model 选项	392	19.6.2	Guttman 模型	426
18.2.4	Statistic 选项	393	19.6.3	平行模型 (Parallel Model)	427
18.2.5	Plots 选项	394	19.7	小结	428
18.2.6	对应分析的结果	395	第 20 章	统计图形的创建和编辑	429
18.3	实例: 美国哲学博士学位毕业 人数的对应分析	399	20.1	SPSS 图形的基本功能介绍	429
18.3.1	实例数据	399	20.1.1	图形生成器 (Chart Builder)	429
18.3.2	操作步骤	399	20.1.2	交互模式创建图形 (Interactive)	430
18.3.3	实例的对应分析结果	400	20.1.3	传统模式创建图形 (Legacy Chart)	430
18.4	多元对应分析	404	20.2	条形图	431
18.4.1	多元对应分析的基本 方法	404	20.2.1	条形图的类型和基本 操作	431
18.4.2	多元对应分析的 SPSS 操作	405	20.2.2	简单条形图	433
18.4.3	Discretization 选项	407	20.2.3	复合条形图	442
18.4.4	Missing Values 选项	408	20.2.4	堆栈条形图	447
18.4.5	Options 选项	408	20.3	三维条形图 (3-D Bar)	452
18.4.6	Output 选项	410	20.3.1	三维条形图的类型	452
18.4.7	Save 选项	411	20.3.2	个案分组模式	454
18.5	实例: 雇员和汽车的多元对 应分析	412	20.4	小结	455
18.5.1	实例数据	412	第 21 章	SPSS 的命令语句程序设计	456
18.5.2	SPSS 操作	412	21.1	Syntax 程序语言概述	456
18.5.3	多元对应分析的结果	413	21.1.1	Syntax 命令语句的特性	456
18.6	小结	417	21.1.2	统计分析方式的选择	456
第 19 章	信度分析	418	21.2	Syntax 命令语句及其定义规则	457
19.1	信度分析的概念	418	21.2.1	SPSS 命令的类型	457
19.2	信度分析的基本方法	419	21.2.2	SPSS 命令的定义规则 及其构成	458
19.3	信度分析的 SPSS 操作	420	21.2.3	SPSS 命令的显示顺序	459
19.3.1	基本操作	420	21.2.4	SPSS 命令语句的状态	460
19.3.2	Statistics 选项	421	21.2.5	SPSS 的子命令	461
19.4	实例: 节目是否继续开办的 调查分析	423	21.2.6	命令执行的两种方式	462
19.4.1	实例中的调查数据	423	21.3	创建和运行 Syntax	462
19.4.2	操作步骤	423	21.3.1	认识 Syntax 窗口	462
19.5	信度分析的结果	423	21.3.2	定制 SPSS Syntax 编程 环境	463
19.6	进一步的分析	425			
19.6.1	拆半信度系数模型	425			



21.3.3	创建和运行 Syntax 文件 ..	464	22.4.1	读取混合数据文件	492
21.3.4	SPSS 程序的类型	465	22.4.2	读取分组数据文件	493
21.4	SPSS Syntax 语法	466	22.4.3	读取嵌套数据文件	496
21.4.1	流程结构语句	466	22.5	小结	498
21.4.2	过程语句	471	第 23 章	宏	499
21.4.3	文件操作语句	474	23.1	SPSS 宏概述	499
21.5	Syntax 语句错误与程序调试	478	23.1.1	SPSS 宏工具概述	499
21.5.1	识别 Syntax 命令的语 法错误	479	23.1.2	SPSS 宏的定义	499
21.5.2	错误信息	479	23.2	SPSS 宏的参数	500
21.5.3	理解错误信息的含义	480	23.2.1	SPSS 宏参数的定义及 分类	500
21.5.4	修改 Syntax 语法错误	480	23.2.2	SPSS 宏的关键字参数	501
21.5.5	Syntax 语法错误的实例 分析	480	23.2.3	SPSS 宏的位置参数	502
21.6	小结	482	23.2.4	参数的标识符	502
第 22 章	利用 SPSS 语句读取数据文件 ...	483	23.2.5	参数的默认设置	505
22.1	读取数据库中的数据	483	23.3	SPSS 宏的流程控制结构	505
22.1.1	安装数据库驱动	483	23.3.1	IF 结构	505
22.1.2	读取数据库表中的数据 ...	483	23.3.2	循环结构	506
22.1.3	读取数据库的多个表	485	23.4	宏的其他操作	507
22.2	读取 Excel 文件数据	486	23.4.1	宏变量的数学运算	507
22.2.1	读取 Excel 工作表数据 ...	486	23.4.2	宏变量中直接读取分配	508
22.2.2	读取多个工作表数据	487	23.4.3	字符串操作函数	508
22.3	读取文本文件数据	488	23.5	SPSS 宏的调试	509
22.3.1	读取定界文本数据文件 ...	488	23.5.1	输出窗口中显示扩展宏 语句	510
22.3.2	读取逗号分隔值 (CSV) 数据文件	489	23.5.2	结果窗口中显示宏参数	510
22.3.3	读取固定宽度文本数据 文件	491	23.6	小结	510
22.4	读取复杂文本数据	492	参考文献		511

Part

第 1 部分 IBM SPSS 19 入门

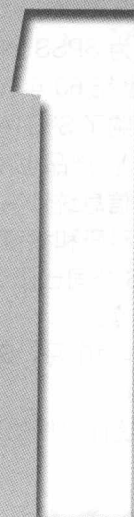
第 1 章 IBM SPSS Statistics 19 概述

第 2 章 IBM SPSS 19 的窗口及其设置

第 3 章 SPSS 统计数据的创建与编辑

第 4 章 SPSS 数据的管理

第 5 章 数据转换与 SPSS 函数



第 1 章 IBM SPSS Statistics 概述

本章包括

- ◆ SPSS 的发展简史
- ◆ IBM SPSS Statistics 19 对环境的要求
- ◆ IBM SPSS Statistics 19 的启动与退出
- ◆ IBM SPSS Statistics 19 的帮助系统
- ◆ IBM SPSS Statistics 19 的特点
- ◆ 安装和卸载 SPSS
- ◆ IBM SPSS Statistics 19 的新增功能

SPSS (Statistical Package for the Social Science) 是在 SPSS/PC+基础上发展起来的,是集成了数据管理、统计分析、图形报表和编程扩展于一身的功能强大的应用统计分析软件包。该软件适用于自然科学和社会科学的各个领域,是世界上最流行的统计软件之一。近年来,我国的政府部门、医疗卫生、体育、经济等领域的工作人员已广泛使用该软件进行信息管理和决策分析等工作。同时,SPSS 统计分析软件已经成为许多院校本科生和研究生的必备技能之一。

1.1 SPSS 的发展简史

SPSS 也称社会科学统计软件包,是一种集成化的计算机数据处理和统计分析通用软件。该软件是世界上公认的最优秀的统计分析软件包之一,现已广泛应用于经济学、生物学、心理学、医疗卫生、体育、农业、商业和金融等多个领域。

1968 年美国斯坦福大学的三位研究生开发了最早的 SPSS 统计软件,当时主要面向中小型的计算机用户和企业级用户,统称为 SPSSx 版,并于 1975 年在美国的芝加哥成立了 SPSS 总部。后来随着微型计算机和 DOS 操作系统的出现,SPSS 公司又于 1984 推出了 DOS 操作系统上的 SPSS 微机版,即 SPSS/PC+版,一共推出了三个版本。20 世纪 90 年代,随着 Windows 操作系统的出现和普及,相继出现了面向 Windows 操作系统的 SPSS 第 5 版、第 6 版。20 世纪 90 年代中后期,又推出以交互式对话为主要特征的第 7 版,目前最新的版本是第 19 版,是一种用于分析数据的综合系统,并且可以选择包括简体中文在内的多种语言,更广泛地适应不同用户的需求。第 7 版以后的版本称为 SPSS for Windows 版。

从 20 世纪 60 年代末到今天,SPSS 的发展已经历了半个世纪。1994 年至 1998 年间,SPSS 公司陆续并购了 SYSTAT、BMDP 等公司,成为世界上公认的三大数据分析软件之一 (SAS、SPSS 和 SYSTAT)。产品也由原来的单一统计分析软件开发为主,转变为向企业、教育科研及政府机构提供全面的信息统计决策支持服务,并广泛分布于通信、医疗、银行、证券、保险、制造、商业、市场研究、科研和教育等多个领域和行业,在全球财富 1000 强企业中,有 95%的企业是 SPSS 的客户。SPSS 公司已于 2000 年正式更名为 Statistical Product and Service Solutions,即统计产品与服务解决方案。

2009 年 10 月,SPSS 公司被 IBM 并购,因此在推出 SPSS 19 时将软件命名为 IBM SPSS Statistics 19。

SPSS 统计软件的易学易用以及专业化的统计分析结果是 SPSS 软件强大的竞争优势之一,也

是广大数据分析人员对其偏爱的主要原因。大量成熟的统计分析方法、完善的数据定义界面、开放的数据接口,以及灵活的统计表格和统计图形,更是 SPSS 长盛不衰的重要法宝。

1.2 IBM SPSS Statistics 19 的特点

从 SPSS/PC+ 版本到现在的 IBM SPSS Statistics 19 版本,SPSS 在用户操作方面最大的改进就是由原来的命令行方式,转变为以对话框输入方式为主,进行数据管理和统计分析,免去了记忆命令和参数的负担;同时,还保留了原来的命令行方式的优点。另外,分析结果的输出方式也由原来的以字符形式为主,转变为以文字、表格和图形混排的形式输出。IBM SPSS Statistics 19 的特点如下。

(1) 操作简单

IBM SPSS Statistics 19 的工具栏提供了方便用户进行各种不同操作的按钮,用户可以根据不同的需要增加或减少操作按钮。同时,IBM SPSS Statistics 19 还具有记忆功能,能够记忆用户最近打开的数个文件及当前执行的统计分析与绘图功能。

(2) 统计功能强大

用户可根据不同的需要,选择不同的统计模块,完成问题的基本统计分析和专业统计分析任务。同时,随着统计学理论的发展以及用户的需要,SPSS 软件在版本更新过程中还增添了最新的统计学研究成果。

(3) 丰富的图表功能

用户可根据需要,选择不同风格、不同功能的表和图来整理数据,使报告、论文更简洁、更美观。SPSS 提供了多种不同的图形和表格形式供用户选择,以满足用户的不同需求。

(4) 兼容多种不同的数据格式

SPSS 软件可以与很多其他软件文件格式进行数据传输。该软件可以打开扩展名为 .dat、.xls、.slk、.dbf 和 .wk3 等多种常用的数据文件。该软件生成的图形也可保存为多种图形格式。

(5) 完善的帮助系统

SPSS 根据不同层次的用户提供不同的帮助,即使是初次接触 IBM SPSS Statistics 19 的用户,也可以轻松完成基本的统计分析任务。用户在使用过程中,还可利用 Help 菜单中的“Developer Central”(研发中心)直接在线访问软件最新信息或寻求相关帮助。

1.3 IBM SPSS Statistics 19 对环境的要求

1.3.1 对硬件环境的要求

IBM SPSS Statistics 19 对计算机硬件的要求不高,但对于大数据量的统计分析问题,则需要较大的内存才能得到满意的运行速度。

对硬件的基本要求如下。

- ◆ 至少 128MB 内存。但对于海量数据的管理和复杂的统计分析(例如多因素分析等),则需要至少 256MB 及以上的内存方可获得满意的运行速度。