

SPSS统计分析

从入门到精通

● 杜强 贾丽艳 编著

源于实践 成就行录

SPSS

上海软件行业协会 秘书长 杨根兴

江苏省软件行业协会 副会长 徐雷

鼎力推荐

- ▶ 全面展示新版本SPSS 的功能应用。
- ▶ 包括各种数据分析应用的实例，从基本的描述统计、相关分析、回归分析等，到高级的生存分析、时序分析、决策树分析等。
- ▶ 6个完整的案例，帮助读者尽快融入实战。

 人民邮电出版社
POSTS & TELECOM PRESS

SPSS统计分析 从入门到精通

◎ 杜强 贾丽艳 编著

人民邮电出版社
北京

图书在版编目(CIP)数据

SPSS统计分析从入门到精通 / 杜强, 贾丽艳编著

— 北京: 人民邮电出版社, 2011.9

ISBN 978-7-115-25709-3

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

中国版本图书馆CIP数据核字(2011)第125561号

内 容 提 要

本书使用 IBM SPSS Statistics 19.0 进行讲解和操作, 致力于使读者全面了解 SPSS, 了解和学习如何使用 SPSS 进行数据融合、数据分析、结果展示等工作。本书介绍的是 SPSS 的窗口和对话框操作方式, 着重于介绍 SPSS 分析软件的实际应用。

全书 25 章, 第 1~3 章重点讲解了数据和文件的管理操作, 以及 SPSS 系统环境的设置。第 4~18 章主要介绍各种统计分析方法及对应的操作方式, 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、生存分析、时间序列分析、多重响应分析等几大类。第 19 章介绍各种统计图形的生成和编辑。第 20~25 章列举了用 SPSS 处理多种行业数据的案例, 包括上市公司财务数据分析、影响汇率的因素分析、多因素试验设计等多方面的应用。

本书适应自然科学和社会科学各领域、各专业研究人员的多层次需要, 可供相关专业本科生、研究生、专业统计分析人士以及管理人员和决策者等学习与参考。

SPSS 统计分析从入门到精通

◆ 编 著 杜 强 贾丽艳
责任编辑 张 涛

◆ 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>
大厂聚鑫印刷有限责任公司印刷

◆ 开本: 787×1092 1/16
印张: 33.5

字数: 883 千字

2011 年 9 月第 1 版

印数: 1-3 000 册

2011 年 9 月河北第 1 次印刷



ISBN 978-7-115-25709-3

定价: 69.00 元

读者服务热线: (010) 67132692 印装质量热线: (010) 67129223

反盗版热线: (010) 67171154

前 言

统计学描绘了一套包括观察事物现象、收集测量数据、处理所得数据的体系，是指导我们进行数据分析的优秀的方法论；而新的统计分析软件的出现，为我们把这套方法应用到实际生活中提供了极大的便利。把统计分析方法和计算机结合起来的较好范例就是 SPSS，它既适用于对数据分析感兴趣的初级用户，也适用于不同领域的统计分析专家。

在数据挖掘、机器学习、智能决策等领域欣欣向荣的同时，SPSS 软件也积极跟进、不断发展。2009 年 7 月，IBM 宣布收购 SPSS，优秀的统计分析软件 SPSS 结合 IBM 的资源优势，必将为用户提供更加实用的分析体验。随后推出的 IBM SPSS Statistics 18.0 和 19.0，其操作友好性、模型可靠性和与其他平台的兼容性等，都有了非常大的提升。

与 SPSS 15.0 相比，SPSS 19.0 增加了许多新的功能：增加了 Bootstrapping、直销（Direct Marketing）等模块，极大地方便了分析专家和市场研究人员应用相关功能；新模块 IBM SPSS Statistics Developer 使 SPSS 可以与 R 软件协作；增加了部分更新的分析方法（广义线性模型）和报告方式（轻量表）；对于从外部来源读取的数据以及在会话中创建的新变量，改进了确定其默认测量级别的方法；增强的语法编辑器，可将编辑器窗格拆分为上下两个，支持语法加亮和自动完成等功能；加入支持 salesforce.com 的数据库驱动程序；改进了软件架构和性能，如枢纽表的显示效果与交互性能、描述性分析过程的性能等。

本书是一本实用的参考手册，对于读者关心的各种模型参数含义及其设置方法，书中力求讲解详尽；从软件操作步骤到行业案例分析，都配有图形指示和样本数据实例分析；对于数理统计和分析方法的相关背景，书中做简单铺垫后，给出算法的基本思想或关键步骤，重点介绍软件操作方法和参数设置，以期达到读者可以通过本书理解和应用相关统计方法的目的；若读者再参考有关统计分析的专业书籍或不同背景知识的书籍，再加上不断的操作演练和实践，就能做更深入的算法改进、应用创新等研究性工作。

程序文件下载网址为：www.ptpress.com.cn。

致谢：

在此，感谢人民邮电出版社各位编辑同志和金羽工作室张铮老师的帮助和指导，感谢罗瑞芳博士、何力武博士、鲍青波、李广鹏、王杉和王命达等人的参与和给出的宝贵建议。

由于编写水平有限，统计理论和软件的发展非常迅速，书中难免会存在纰漏和不足之处，恳请广大读者提出批评与指正，我们真诚希望与大家相互交流，共同进步。读者可以发送邮件到 editor@book95.com，或登录我们的答疑互动网站 bbs.book95.com 找到本书的主题帖与作者交流。

编者

目 录

第1章 SPSS 19.0概述 ----- 1

- 1.1 SPSS 简介 ----- 1
- 1.2 SPSS 的安装、启动和退出 ----- 3
 - 1.2.1 SPSS 19.0 的安装 -----3
 - 1.2.2 SPSS 19.0 的启动 -----4
 - 1.2.3 SPSS 19.0 的退出 -----5
- 1.3 SPSS 19.0 的界面及设置 ----- 6
 - 1.3.1 常用界面 -----6
 - 1.3.2 常规选项参数-----8
 - 1.3.3 查看器选项参数----- 11
 - 1.3.4 文件位置选项参数----- 12
 - 1.3.5 输出选项参数----- 12
 - 1.3.6 图表选项参数----- 13
 - 1.3.7 多重归因选项参数----- 14
 - 1.3.8 枢纽表选项参数----- 15
 - 1.3.9 数据选项参数----- 17
 - 1.3.10 货币选项参数 ----- 18
 - 1.3.11 脚本选项参数 ----- 18
 - 1.3.12 语法编辑器选项参数 ----- 19

第2章 数据文件的建立 与操作 ----- 20

- 2.1 数据编辑器与数据文件-----20
 - 2.1.1 数据编辑器 ----- 20
 - 2.1.2 数据文件 ----- 23
- 2.2 常量、变量、操作符和表达式 ----23
 - 2.2.1 常量与变量 ----- 23
 - 2.2.2 操作符与表达式----- 28
 - 2.2.3 如何定义一个变量----- 28
 - 2.2.4 概率事件 ----- 32
- 2.3 输入数据 -----32
 - 2.3.1 输入数据的方法----- 32
 - 2.3.2 查看文件信息和变量信息 ----- 33
- 2.4 编辑数据文件 -----34
 - 2.4.1 在单元格中编辑数据 ----- 34

- 2.4.2 插入变量与删除变量 -----34
- 2.4.3 插入观测量与删除观测量 ----- 35
- 2.4.4 数据的剪切、复制和粘贴 ----- 35
- 2.4.5 撤销操作----- 36
- 2.5 对数据文件的操作----- 36
 - 2.5.1 数据文件的打开与保存 ----- 37
 - 2.5.2 数据库文件的转换 ----- 37
- 习题 2----- 42

第3章 数据文件的操作 ----- 43

- 3.1 数据文件的一般操作 ----- 43
 - 3.1.1 数据排序----- 43
 - 3.1.2 数据文件的拆分 ----- 44
 - 3.1.3 数据文件的合并 ----- 46
 - 3.1.4 数据文件的转置 ----- 49
 - 3.1.5 变量取值的求秩 ----- 50
 - 3.1.6 变量值的重新编码 ----- 52
 - 3.1.7 计算新变量----- 56
- 3.2 分类汇总 ----- 58
 - 3.2.1 数据描述----- 58
 - 3.2.2 分类汇总的参数设置 ----- 58
 - 3.2.3 分类汇总的结果 ----- 60
- 3.3 观测量的加权----- 60
- 3.4 数据文件的结构重组 ----- 62
 - 3.4.1 选择数据重组方式 ----- 63
 - 3.4.2 变量组到观测量组的重组 ----- 63
 - 3.4.3 观测量组到变量组的重组 ----- 66
 - 3.4.4 转置重组----- 69
- 习题 3----- 69

第4章 基本统计分析功能 ----- 70

- 4.1 OLAP 在线分析过程 ----- 70
 - 4.1.1 数据描述----- 70
 - 4.1.2 OLAP 过程的操作和设置 ----- 70
- 4.2 个案汇总分析----- 73
 - 4.2.1 个案汇总分析的参数设置 ----- 73

4.2.2	输出结果	75	6.1.2	非参数检验的优点	107
4.3	按行和列的汇总分析	75	6.1.3	非参数检验的缺点	108
4.3.1	按行汇总过程	75	6.2	卡方检验	108
4.3.2	按列汇总过程	79	6.2.1	原理与方法	108
4.4	频数分析	81	6.2.2	数据和问题描述	109
4.4.1	数据描述	81	6.2.3	卡方检验实例分析	109
4.4.2	对分类变量的频数分析	81	6.3	二项式检验	111
4.4.3	对连续变量的频数分析	83	6.3.1	原理与方法	111
4.5	描述性统计分析	85	6.3.2	数据和问题描述	111
4.5.1	数据描述	85	6.3.3	二项检验实例分析	111
4.5.2	描述分析过程	85	6.4	游程检验	112
4.6	探索分析过程	86	6.4.1	原理与方法	112
4.6.1	数据描述	86	6.4.2	数据和问题描述	113
4.6.2	探索分析实例	87	6.4.3	游程检验实例分析	113
4.7	列联表分析过程	90	6.5	Kolmogorov-Smirnov 单样本检验	114
4.7.1	数据描述	90	6.5.1	原理与方法	114
4.7.2	列联表分析的参数设置	90	6.5.2	数据和问题描述	115
4.7.3	列联表分析的输出结果	93	6.5.3	K-S 单样本检验实例分析	115
4.8	Bootstrap 简介与设置	95	6.6	两独立样本检验	116
4.8.1	Bootstrap 简介	95	6.6.1	原理与方法	116
4.8.2	Bootstrap 参数设置	95	6.6.2	数据和问题描述	117
习题 4		96	6.6.3	两独立样本检验实例分析	117
第5章 均值比较和T检验	97		6.7	k 个独立样本的检验	119
5.1	均值分析过程	98	6.7.1	原理与方法	119
5.1.1	原理与方法	98	6.7.2	数据和问题描述	120
5.1.2	SPSS 实例分析	98	6.7.3	k 个独立样本检验实例分析	120
5.2	单样本 T 检验	100	6.8	两个相关样本的检验	121
5.2.1	原理与方法	100	6.8.1	原理与方法	121
5.2.2	SPSS 实例分析	100	6.8.2	数据和问题描述	123
5.3	独立样本 T 检验	101	6.8.3	两个相关样本检验的 实例分析	123
5.3.1	原理与方法	101	6.9	k 个相关样本的检验	123
5.3.2	SPSS 实例分析	102	6.9.1	原理与方法	124
5.4	配对样本 T 检验	103	6.9.2	数据和问题描述	125
5.4.1	原理与方法	103	6.9.3	k 个相关样本检验的 实例分析	125
5.4.2	SPSS 实例分析	104	习题 6		126
习题 5		105			
第6章 非参数检验	106		第7章 多重响应分析	127	
6.1	非参数检验的简介	107	7.1	多重响应概述	127
6.1.1	非参数检验与参数检验	107	7.2	多重响应变量集的定义	128

定义多重响应变量集的实例 -----	128	8.5.2 问题描述和数据准备 -----	168
7.3 多重响应变量集的频率分析 -----	129	8.5.3 多元 Logistic 回归参数设置 -----	169
多重响应变量频数分析的实例 -----	129	8.5.4 案例的结果分析 -----	173
7.4 多重响应变量集的交叉表分析 -----	131	8.6 有序回归 -----	175
交叉表分析 -----	131	8.6.1 问题描述和数据准备 -----	175
多重响应变量交叉表分析的实例 -----	131	8.6.2 有序回归的参数设置 -----	175
7.5 用表过程研究多重响应变量集 -----	133	8.6.3 案例的结果分析 -----	178
7.5.1 多重响应变量集的定义 -----	133	8.7 概率单位回归分析 -----	180
7.5.2 建立包含多重响应变量集的表格 -----	134	8.7.1 概率单位回归分析简介 -----	180
习题 7 -----	135	8.7.2 问题描述和数据准备 -----	180
第8章 回归分析 ----- 137		8.7.3 概率单位回归的参数设置 -----	181
8.1 线性回归 -----	137	8.7.4 案例的结果分析 -----	182
8.1.1 一元线性回归的基本原理 -----	137	8.8 加权回归分析 -----	183
8.1.2 多元线性回归的基本原理 -----	139	8.8.1 加权回归分析简介 -----	184
8.1.3 模型假设的其他检验 -----	140	8.8.2 问题描述和数据准备 -----	184
8.1.4 问题描述和数据准备 -----	141	8.8.3 加权回归的参数设置 -----	185
8.1.5 线性回归分析和操作 -----	141	8.8.4 案例结果分析 -----	186
8.1.6 案例的结果分析 -----	146	8.9 二阶段最小二乘回归 -----	186
8.2 曲线回归 -----	148	8.9.1 二阶段最小二乘回归的基本原理 -----	186
8.2.1 曲线回归的基本原理 -----	149	8.9.2 问题描述和数据准备 -----	187
8.2.2 问题描述和数据准备 -----	149	8.9.3 二阶段最小二乘回归的参数设置 -----	188
8.2.3 曲线回归分析和操作 -----	150	8.9.4 案例的结果分析 -----	188
8.2.4 案例的结果分析 -----	151	8.10 最佳尺度回归 -----	189
8.3 非线性回归 -----	152	8.10.1 最佳尺度回归原理 -----	190
8.3.1 非线性回归简介 -----	153	8.10.2 问题描述和数据准备 -----	190
8.3.2 问题描述和数据准备 -----	154	8.10.3 最佳尺度回归的参数设置 -----	190
8.3.3 非线性回归的参数设置 -----	155	8.10.4 案例的结果分析 -----	194
8.3.4 案例的结果分析 -----	158	习题 8 -----	196
8.4 二元 Logistic 回归 -----	159	第9章 方差分析 ----- 198	
8.4.1 二元 Logistic 回归的数学原理 -----	159	9.1 方差分析简介 -----	198
8.4.2 问题描述和数据准备 -----	160	9.1.1 t 检验与方差分析的比较 -----	198
8.4.3 二元 Logistic 回归的参数设置 -----	161	9.1.2 方差分析的基本原理 -----	199
8.4.4 案例的结果分析 -----	164	9.2 单因素方差分析 -----	201
8.5 多元 Logistic 回归分析 -----	167	9.2.1 原理与方法 -----	201
8.5.1 多元 Logistic 回归的原理简介 -----	168	9.2.2 单因素方差分析实例 -----	201
		9.3 多因素方差分析过程 -----	205
		9.3.1 原理与方法 -----	206
		9.3.2 二因素方差分析实例 -----	208

9.3.3	协方差分析实例	215
9.3.4	交互效应中随机因素的分析	218
9.4	多元方差分析	221
9.4.1	原理与方法	221
9.4.2	多元方差分析实例	222
9.5	重复测量设计的方差分析	223
9.5.1	原理与方法	223
9.5.2	SPSS 实例分析	224
9.6	方差成分分析	227
9.6.1	原理简介	227
9.6.2	SPSS 实例分析	228
9.7	正交实验设计	230
9.7.1	正交实验设计简述	230
9.7.2	SPSS 实例分析	231
9.7.3	正交实验设计的方差分析	232
习题 9		233

第10章 相关分析----- 234

10.1	相关分析的基本概念	234
10.1.1	相关分析的特点和应用	234
10.1.2	相关系数的计算	235
10.1.3	SPSS 提供的相关分析功能	236
10.2	两变量相关分析	237
10.2.1	问题描述和数据准备	237
10.2.2	相关分析的参数设置	237
10.2.3	案例的结果分析	238
10.3	偏相关分析	239
10.3.1	偏相关分析的基本原理	239
10.3.2	偏相关分析实例	240
10.4	距离分析	241
10.4.1	距离分析的基本概念	241
10.4.2	距离分析的参数设置	242
10.4.3	距离分析实例	245
习题 10		246

第11章 因子分析----- 248

11.1	因子分析的原理简介	248
11.1.1	因子分析的基本思想	248
11.1.2	因子分析和主成分 分析的联系	249
11.1.3	因子分析的基本步骤	249

11.2	SPSS 因子分析的应用实例	250
11.2.1	数据描述	250
11.2.2	SPSS 因子分析过程的设置	251
11.2.3	结果分析	255
习题 11		258

第12章 分类分析----- 260

12.1	聚类分析的原理简介	260
12.1.1	聚类分析的基本概念	260
12.1.2	聚类分析的一般原理	261
12.2	快速样本聚类过程	263
12.2.1	快速聚类简介	263
12.2.2	问题描述和数据准备	263
12.2.3	SPSS 快速聚类的设置	264
12.2.4	案例的结果分析	266
12.3	系统聚类	267
12.3.1	系统聚类简介	267
12.3.2	问题描述和数据准备	268
12.3.3	SPSS 系统聚类的设置	268
12.3.4	案例的结果分析	271
12.3.5	对聚类结果的进一步分析	272
12.4	两步聚类分析	274
12.4.1	两步聚类简介	274
12.4.2	问题描述和数据准备	275
12.4.3	SPSS 两步聚类的设置	275
12.4.4	案例的结果分析	278
12.5	一般判别分析	280
12.5.1	判别分析的基本原理	280
12.5.2	问题描述和数据准备	282
12.5.3	判别分析的参数设置	282
12.5.4	案例的结果分析	284
12.6	逐步判别分析实例	287
12.6.1	问题描述和数据准备	288
12.6.2	逐步判别的参数设置	288
12.6.3	案例的结果分析	290
12.7	决策树分析	292
12.7.1	决策树分类的基本原理	292
12.7.2	决策树过程的参数设置	294
12.7.3	问题描述和数据准备	305
12.7.4	案例分析	305
习题 12		309

第13章 生存分析----- 311

- 13.1 生存分析简介----- 311
 - 13.1.1 生存分析的基本概念 -----311
 - 13.1.2 生存分析的数据特点 -----313
 - 13.1.3 生存分析的常用方法 -----313
 - 13.1.4 SPSS 中的生存分析过程 -----313
- 13.2 寿命表分析 ----- 314
 - 13.2.1 寿命表分析简介 -----314
 - 13.2.2 寿命表分析的基本步骤-----314
 - 13.2.3 寿命表实例分析 -----315
- 13.3 Kaplan-Meier 分析 ----- 318
 - 13.3.1 Kaplan-Meier 分析的步骤 -----318
 - 13.3.2 生存曲线的比较和检验-----319
 - 13.3.3 Kaplan-Meier 分析的实例 -----319
- 13.4 Cox 回归模型----- 322
 - 13.4.1 Cox 回归模型的原理简介 -----322
 - 13.4.2 Cox 回归实例分析-----324
- 习题 13----- 329

第14章 信度分析----- 331

- 14.1 信度分析 ----- 331
 - 14.1.1 信度分析的基本原理 -----331
 - 14.1.2 问题描述和数据准备 -----333
 - 14.1.3 信度分析的参数设置 -----333
 - 14.1.4 案例的结果分析 -----335
- 14.2 多维尺度分析----- 336
 - 14.2.1 多维尺度分析简介 -----336
 - 14.2.2 问题描述和数据准备 -----336
 - 14.2.3 ALSCAL 过程的参数设置 -----337
 - 14.2.4 案例的结果分析 -----339
- 习题 14----- 340

第15章 时间序列分析 ----- 342

- 15.1 SPSS 的时间序列分析概览 ---- 342
 - 15.1.1 创建模型的通用设置选项-----343
 - 15.1.2 应用模型的通用设置选项-----347
- 15.2 时间序列数据的预分析 ----- 348
 - 15.2.1 缺失值替换 -----348
 - 15.2.2 定义日期变量 -----349
 - 15.2.3 时间序列的平稳化 -----349
- 15.3 指数平滑模型----- 351

- 15.3.1 指数平滑的基本原理----- 351
- 15.3.2 指数平滑模型的参数设置 ----- 352
- 15.3.3 指数平滑模型实例分析----- 354

15.4 ARIMA 模型 -----357

- 15.4.1 ARIMA 模型的基本原理 ----- 357
- 15.4.2 ARIMA 模型的参数设置 ----- 359
- 15.4.3 ARIMA 模型实例分析 ----- 360

15.5 季节分解模型 -----362

- 15.5.1 季节分解法概述 ----- 362
- 15.5.2 季节分解模型实例分析----- 363

习题 15 -----366**第16章 对数线性模型 ----- 367**

- 16.1 对数线性模型概述 -----367
 - 16.1.1 简单列联表分析的不足----- 367
 - 16.1.2 对数线性模型的基本形式----- 368
- 16.2 常规对数线性模型过程-----368
 - 16.2.1 常规过程概述 ----- 368
 - 16.2.2 问题描述和数据准备----- 369
 - 16.2.3 常规过程的参数设置----- 369
 - 16.2.4 案例的结果分析 ----- 371
- 16.3 Logit 过程-----372
 - 16.3.1 Logit 过程概述 ----- 372
 - 16.3.2 问题描述和数据准备----- 372
 - 16.3.3 Logit 过程的参数设置 ----- 373
 - 16.3.4 案例的结果分析 ----- 373
- 16.4 模型选择过程 -----375
 - 16.4.1 模型选择过程概述----- 375
 - 16.4.2 问题描述和数据准备----- 376
 - 16.4.3 层次对数线性模型的操作过程 ----- 376
 - 16.4.4 案例的结果分析 ----- 377
- 习题 16 -----379

第17章 对应分析----- 380

- 17.1 对应分析的基本原理 -----380
 - 17.1.1 对应分析与因子分析----- 380
 - 17.1.2 SPSS 中的对应分析 ----- 381
 - 17.1.3 使用对应分析的注意事项 ----- 381
- 17.2 简单对应分析 -----382
 - 17.2.1 简单对应分析的数学原理----- 382

17.2.2	SPSS 简单对应分析实例	383	19.6.3	用对话框创建高低图	421
17.3	多元对应分析	387	19.7	帕累托图	425
17.3.1	多元对应分析基本概念 及其特点	387	19.7.1	数据和问题描述	425
17.3.2	多元对应分析的参数设置	388	19.7.2	用对话框创建帕累托图	425
17.3.3	实例的结果分析	392	19.8	控制图	427
习题 17		395	19.8.1	数据和问题描述	427
			19.8.2	用对话框创建控制图	427
第18章 缺失值分析	396		19.9	箱图	432
18.1	缺失值分析的概念	396	19.9.1	数据和问题描述	432
18.1.1	缺失值的表现方式	396	19.9.2	用图表构建程序作箱图	432
18.1.2	SPSS 中的缺失值处理方法	397	19.9.3	用对话框创建箱图	434
18.2	缺失值分析的参数设置	398	19.10	误差条图	434
18.3	缺失值分析的实例	401	19.10.1	数据和问题描述	434
习题 18		404	19.10.2	用对话框创建误差条图	435
第19章 统计图形	405		19.11	散点图	435
19.1	概述	405	19.11.1	数据和问题描述	436
19.1.1	数据和变量的准备	405	19.11.2	用图表构建程序作散点图	436
19.1.2	图表构建程序的基本操作	406	19.11.3	用对话框创建散点图	438
19.1.3	旧对话框作图	408	19.12	直方图	439
19.1.4	图形的编辑	408	19.12.1	数据和问题描述	440
19.2	条形图	408	19.12.2	用图表构建程序作直方图	440
19.2.1	数据和问题描述	409	19.13	P-P 概率图	441
19.2.2	用图表构建程序作条形图	409	19.13.1	数据和问题描述	441
19.2.3	用对话框创建条形图	412	19.13.2	用对话框创建 P-P 概率图	441
19.3	线形图	414	19.14	Q-Q 概率图	443
19.3.1	数据和问题描述	414	19.14.1	数据和问题描述	443
19.3.2	用图表构建程序作线形图	414	19.14.2	用对话框创建 Q-Q 概率图	443
19.3.3	用对话框创建线形图	415	19.15	时间序列图	444
19.4	面积图	416	19.15.1	普通序列图	445
19.4.1	数据和问题描述	416	19.15.2	自相关序列图	447
19.4.2	用图表构建程序作面积图	417	19.15.3	互相关序列图	449
19.4.3	用对话框创建面积图	418	19.16	双轴线图	450
19.5	饼图	419	19.16.1	数据和问题描述	450
19.5.1	数据和问题描述	419	19.16.2	用图表构建程序 作双轴线图	450
19.5.2	用图表构建程序作饼图	419	习题 19		452
19.5.3	用对话框创建饼图	420	第20章 上市公司财务危机 预警分析	453	
19.6	高低图	420	20.1	财务危机预警的应用简介	453
19.6.1	数据和问题描述	420			
19.6.2	用图表构建程序作高低图	421			

20.1.1	财务危机的定量定义方法	453
20.1.2	金融危机预警的模型选择	454
20.2	数据描述	454
20.2.1	数据说明	454
20.2.2	指标选择	455
20.2.3	补充说明	455
20.3	分析方法概述	456
20.3.1	判别分析	456
20.3.2	logistic 回归方法	456
20.4	SPSS 建模过程和结论分析	457
20.4.1	SPSS 数据筛选操作	457
20.4.2	SPSS 判别分析 建模与分析	460
20.4.3	logistic 回归建模与分析	463
20.5	进一步的分析与应用	466
20.5.1	分类结果的应用分析	466
20.5.2	建模方法的改进	467
20.6	建议和推广	467
20.6.1	时间序列研究	467
20.6.2	数据的有效预警期	467
20.6.3	指标的简化方法	467

第21章 影响汇率的 因素分析 ----- 469

21.1	汇率影响因素的简介	469
21.2	数据描述	470
21.3	分析方法概述	471
21.3.1	探索性分析	471
21.3.2	多元回归分析	472
21.4	SPSS 建模过程和结论分析	472
21.4.1	数据准备	472
21.4.2	探索性分析	473
21.4.3	多元回归分析	474
21.5	进一步的分析与应用	477
21.5.1	剔除存在共线性的 外汇储备变量	477
21.5.2	回归模型的进一步改进	478
21.5.3	两个回归模型的比较	479
21.6	建议和推广	479
21.6.1	时间序列研究	479

21.6.2	汇率影响因素的定性分析	480
--------	-------------	-----

第22章 因子分析在成绩综合 评价中的应用 ----- 481

22.1	学生成绩的综合评价简介	481
22.2	数据描述	482
22.3	分析方法概述	482
22.3.1	应用因子分析进行成绩综合 评价的步骤	482
22.3.2	应用因子分析法进行成绩 综合评价的注意事项	483
22.4	SPSS 建模过程和结论分析	484
22.4.1	数据准备	484
22.4.2	SPSS 因子分析建模与分析	486
22.5	进一步的分析与应用	489
22.6	建议和推广	490
22.6.1	高中生的成绩综合评价	490
22.6.2	对缺失数据的处理	490
22.6.3	多种方法结合的综合 评价模型	491

第23章 高等教育办学条件的 聚类分析 ----- 492

23.1	数据描述	492
23.1.1	关于基本办学条件指标 合格与否的判定	493
23.1.2	指标选取	494
23.1.3	数据格式	494
23.2	聚类分析法简述	494
23.3	SPSS 建模过程和结论分析	495
23.3.1	对专科院校进行聚类的 设置操作	495
23.3.2	对本科院校的分析	499
23.4	建议和推广	501

第24章 试卷信度的 检验与分析 ----- 502

24.1	试卷信度检验的背景简介	502
24.1.1	测验内容的自身方面	502
24.1.2	施测过程	502

24.1.3	被测试者的自身因素	503
24.2	数据描述	503
24.3	分析方法概述	503
24.3.1	试卷信度的基本计算公式	504
24.3.2	试卷信度的估计方法	504
24.4	SPSS 建模过程和结论分析	505
24.4.1	SPSS 信度分析的 参数设置	505
24.4.2	结果分析	505
24.5	建议和推广	507

第25章 多因素试验的 设计与分析 508

25.1	试验设计简介	508
25.1.1	试验设计的应用	509

25.1.2	试验设计问题的解决步骤	509
25.2	数据描述	510
25.3	分析方法概述	510
25.3.1	正交设计方法	510
25.3.2	综合评分方法	512
25.4	SPSS 建模过程和结论分析	512
25.4.1	数据标准化	512
25.4.2	性能指标权重的确定	514
25.4.3	利用权重求综合指标	515
25.4.4	对综合得分的进一步分析	515
25.5	建议和推广	516

习题提示 518

参考文献 524

第 1 章

SPSS 19.0概述

SPSS通过简单的菜单式操作，就可以方便地规范和融合搜集到的各种数据，并能实施从简单的描述性统计分析到复杂的时序分析等多种方法，对数据进行建模，返回有意义的分析结果，比如客户特征的分类、发展趋势的预测等。把这些结果应用于实际，可以帮助读者在发掘潜在客户、制定长远规划等工作上做出更加准确的判断。

本章首先介绍SPSS 19.0软件环境的设置内容和设置方式，让用户打造适合自己的SPSS 19.0工作环境。

1.1 SPSS 简介

IBM 于 2009 年收购 SPSS 后，连续推出了 IBM SPSS Statistics 18.0 和 IBM SPSS Statistics 19.0。业界领先的统计分析软件 SPSS，结合 IBM 的资源优势，必将为用户提供更加方便和强大的智能分析体验。

SPSS 19.0 加入了许多新的功能，包括：增加了 Bootstrapping、直销（Direct Marketing）等模块，极大地方便了分析专家和市场研究人士应用相关功能；新模块 IBM SPSS Statistics Developer 使 SPSS 可以与 R 软件协作；增加了部分更新的分析方法（广义线性模型）和报告方式（轻量表）；改进的默认测量级别，对于从外部来源读取的数据以及在会话中创建的新变量，改进了确定其默认测量级别的方法；增强的语法编辑器，可将编辑器窗格拆分为上下两个窗格，支持语法加亮和自动完成等功能；加入支持 salesforce.com 的数据库驱动程序；改进了软件架构和性能，例如枢纽表交互性能和显示效果、描述性分析过程的性能等。

1. SPSS 的特点

(1) 界面友好，操作简单。SPSS 的命令语句、子命令及各种选项绝大部分都包含在各种菜单和对话框中，因此用户无需花大量时间记忆繁杂的命令、过程和选项。在 SPSS 中，大多数操作可以通过菜单和对话框来完成，因此操作便捷，易于学习和使用。

(2) 适用性好, 因人而异。虽然大部分统计分析方法可以通过菜单和对话框来完成, 但是, 对于熟悉 SPSS 编程语法的用户, 也可以在语法窗口中直接编写程序语句, 从而灵活地完成各种复杂的统计分析任务。另外, 用对话框指定命令、子命令和选项之后, 通过单击界面上的“粘贴”按钮, 可把当前对话框设置对应的语法命令, 自动粘贴到语法编辑窗口中, 并允许保存为文件。

(3) 算法隐藏。具有第四代语言的特点, 只需通过菜单的选择以及对话框的操作告诉系统要做什么, 无需告诉系统怎样做。用户只需了解统计分析原理, 无需通晓统计分析的各种算法, 即可得到统计分析结果。

(4) 接口完善。具有完善的数据转换接口, 其他软件生成的数据文件(例如 Excel 文件、Access 文件、关系数据库生成的 DBF 文件、文本编辑软件生成的 ASCII 码数据文件、SAS 数据文件、salesforce.com 提供的数据等)均可方便地转换成可供 SPSS 分析的数据文件。

SPSS 支持 OLE 技术和 ActiveX 技术, 生成的交互表格和图形对象可以与同类软件进行自动嵌入与链接。现在, SPSS 19.0 还支持 Sax Basic 和 Python 语言。

(5) 功能强大。SPSS 的核心部分是统计功能, 可以完成数理统计分析任务, 提供了从简单的单变量分析到复杂的多变量分析的多种方法, 既包括常规的相关分析、回归分析、方差分析、卡方检验、t 检验和非参数检验, 也包括多元回归分析、聚类分析、判别分析、主成分分析和因子分析, 还包括时间序列分析、生存分析和可靠性分析等。

(6) 表格和图形化功能。SPSS 可以直接生成数十种风格的表格 (OLAP Cubes), 伴随其他分析过程又可生成一般表、多响应表和频数表等表格。利用专门的编辑窗口或结果查看窗口, 能编辑所生成的表格, 如表 1-1 所示。

表 1-1 SPSS 生成表格样例

OLAP Cubes						
时段: 6						
	名 称					
	四川长虹		招商银行		总 计	
	收 盘 价	交 易 量	收 盘 价	交 易 量	收 盘 价	交 易 量
合计	26.23	1 586 805	71	3 999 344	97.23	5 586 149
N	4	4	4	4	8	8
均值	6.557 5	396 701.3	17.75	999 836	12.153 8	698 268.6
标准差	0.085 61	13 123.4	0.5	556 136.7	5.991 83	493 885.5
总和的%	0.04	0.038	0.109	0.097	0.149	0.135
合计 N 的%	0.069	0.069	0.069	0.069	0.138	0.138

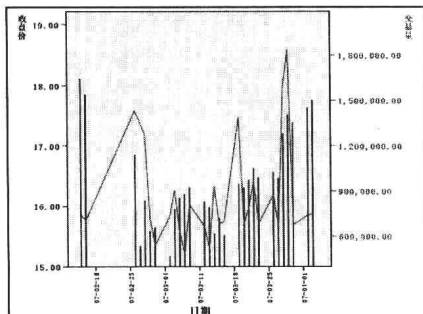


图 1-1 SPSS 图形样例

SPSS 拥有强大的图形功能, 能生成数十种统计图形, 包括条形图、线图、面积图、饼图、高低图、帕累托图、控制图、箱图、误差条图、散点图、直方图、P-P 概率图、Q-Q 概率图和时间序列图等。通过直观、漂亮的统计图形, 能更加形象地显示分析结果, 如图 1-1 所示。

SPSS 19.0 的模型浏览器, 能够交互式地改变部分图形的内容和结构, 使得以前输出的多个图表, 能在一个界面中以交互的方式进行分析, 再加上鼠标提示的功能, 可给用户带来更好的分析体验。

2. SPSS 19.0 的不同版本

SPSS 软件是分模块销售的，用户可根据自己的应用需要和用户数量，确定所需要使用的配置。下面所指 SPSS Statistics 客户端，即大家熟悉的 SPSS 统计分析软件。

(1) 教育版 SPSS Statistics 的客户端，有以下 4 种选择：SPSS 基本版，只有 Base 模块，适合做数据处理、交叉行列分析和描述统计；SPSS 标准版，含有 Base、Advanced、Regression 三个模块，能够完成常见的多元统计分析；SPSS 高级版，含有 Base、Advanced、Regression、Forecasting、Exact Tests 5 个模块，适合经济类和医学类等教学应用中的高级用户；SPSS 专业版，含有 SPSS 客户端的所有 16 个模块。

(2) 商业版 SPSS Statistics 的客户端，有以下三种：SPSS 标准版，含有 Base、Advanced、Regression、Custom Tables 4 个模块，能够进行数据探索及常见的多元统计分析；SPSS 专业版，在标准版基础上增加了 6 个模块，适合做预测分析和营销分析的高级用户；SPSS 白金版，含有 SPSS 客户端的所有 16 个模块，并进一步增加了三个独立产品 SamplePower、Visualization Designer 和 Amos。

(3) 对于企业级应用，SPSS 推荐采用企业级的三层体系架构，即 DB Server + SPSS Server + SPSS Clients。该方案具备有力的海量数据运算能力，充分利用了数据库系统的性能、多线程分析和多 CPU 分布式计算的优势及优化的变量挑选算法等。

1.2 SPSS 的安装、启动和退出

本节介绍 SPSS 安装和启动操作，从这里本书开始踏上 SPSS 应用的旅程。

1.2.1 SPSS 19.0 的安装

运行 SPSS 19.0 的安装程序后，首先见到的是图 1-2 所示的准备界面，稍等片刻即进入图 1-3 所示的安装向导界面，此后为标准的 Windows 安装过程，一直单击“下一步”按钮，直到完成安装。建议在安装的过程中不要运行其他程序。

安装的最后一步界面如图 1-4 所示，单击“确定”按钮后，进入授权和注册界面，要求用户选择授权类型，也可以单击“取消”按钮稍后注册。

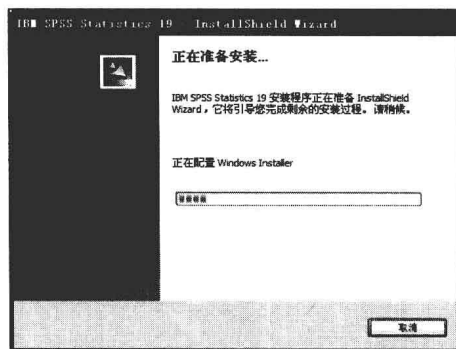


图 1-2 SPSS 安装准备界面

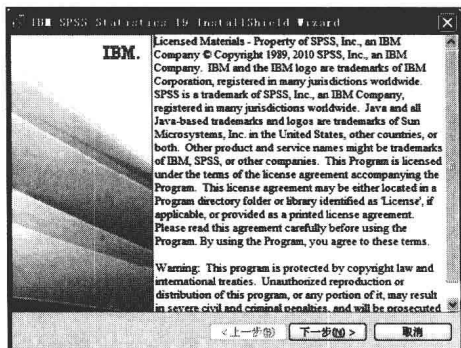


图 1-3 SPSS 安装向导界面

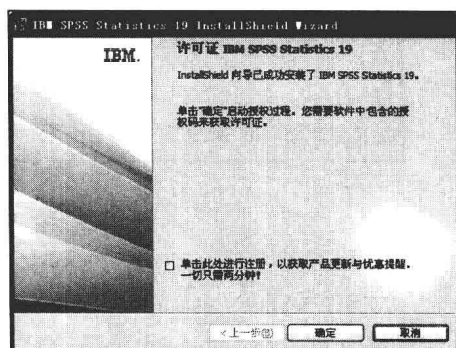


图 1-4 SPSS 安装确定界面

1.2.2 SPSS 19.0 的启动

本书主要介绍 SPSS 的窗口菜单运行方式，通过选择窗口、菜单与对话框来完成各种分析过程。

1. 启动

启动 SPSS 程序，可以双击图 1-5 所示的 SPSS 桌面图标，也可以在“开始”菜单中依次单击“程序→IBM SPSS Statistics→IBM SPSS Statistics 19”启动程序。启动后，首先出现图 1-6 所示的启动界面，显示该软件的版本信息。首次启动后，会先打开图 1-7 所示的选择界面，表明 SPSS 已正常启动，在此可以设置“以后不再显示此对话框”。



图 1-5 SPSS 的桌面图标

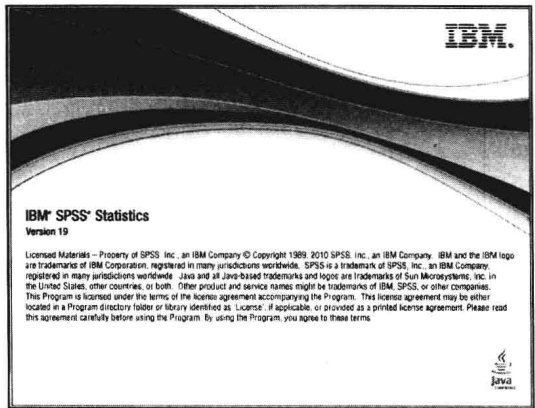


图 1-6 SPSS 的启动界面

对于初学者，一个好的开始是查看软件自身的帮助教程，SPSS 19.0 提供了较为完善和实用的帮助系统，并支持中文，不过需要单独安装 SPSS 官方提供的语言包“SPSS_Statistics_19_Language”。

在图 1-7 中，单击选中“运行教程”单选框，再单击“确定”按钮，就打开图 1-8 所示的教程界面（单击主菜单“帮助→教程”也可打开此教程），这是一个循序渐进的实践指南。另外，SPSS 19.0 的帮助系统是通过网页方式访问的。

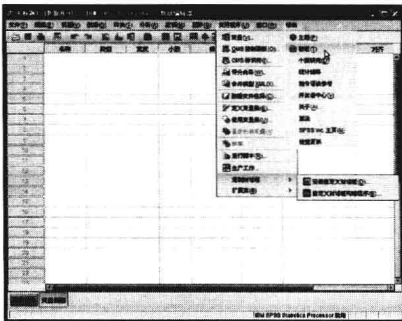


图 1-7 SPSS 启动选择界面

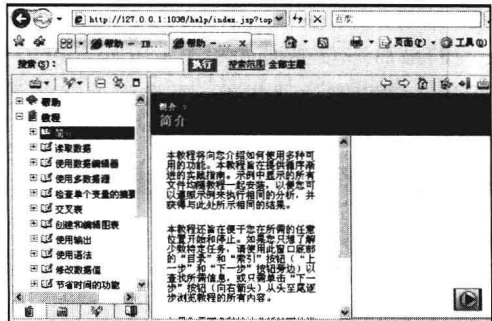


图 1-8 SPSS 帮助教程

2. SPSS 的注册

IBM SPSS Statistics 19.0 在 Windows 桌面“开始/程序”菜单中的项目如图 1-9 所示，单击“许可证授权向导”后，进入如图 1-10 所示的许可证状态界面，显示当前的软件许可信息（临时授权），单击“下一步”按钮，可更改 SPSS 19.0 软件的授权方式（许可证还是临时）。

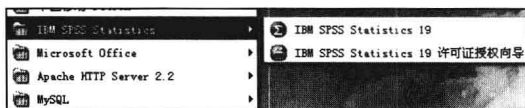


图 1-9 SPSS 注册

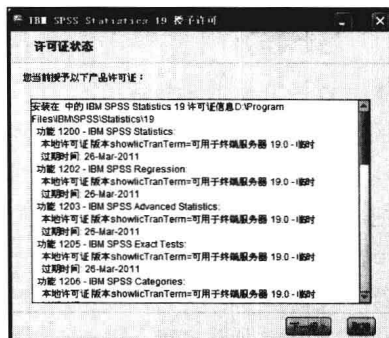


图 1-10 SPSS 的许可证状态界面

3. 其他三种运行方式

(1) 程序运行方式。直接在语法命令窗口或脚本窗口编辑和运行程序，这种方式要求用户掌握 SPSS 的语法命令（Syntax）语言或 Sax Basic、Python 编程语言。

(2) 混合运行方式。首先在“完全窗口”方式下的数据编辑器窗口中输入数据，或者利用“文件”菜单打开已经存在的数据文件；然后利用菜单和对话框操作，设置数据处理的参数；参数设置好后单击设置界面中的“粘贴”按钮，将选择的过程及参数转换成相应的语法命令，并自动粘贴至语法编辑窗口；然后，在语法编辑窗口中添加语句、参数，或者修改已有命令中的参数，单击窗口中的运行功能执行分析。

(3) SPSS 19.0 提供了用户自定义对话框的功能，在图 1-7 中单击菜单“实用程序→定制对话框→自定义对话框构建程序”，打开如图 1-11 所示的设计界面，在此用户可以定制设计能实现特定功能的对话框，包括运行各种 SPSS 功能和程序。

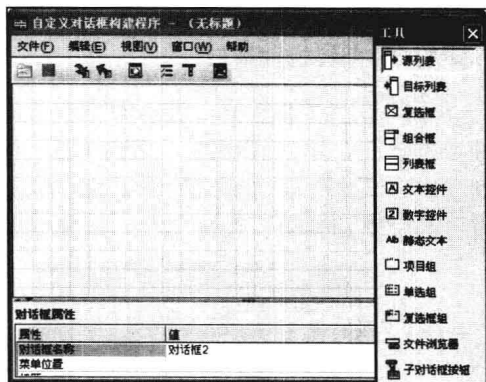


图 1-11 定制对话框的设计界面

1.2.3 SPSS 19.0 的退出

在图 1-11 所示的数据编辑器窗口，单击菜单“文件→退出”，或者单击窗口右上角的关闭按钮，都可以退出 SPSS。

如果进行了任何设置或分析操作，退出时首先弹出如图 1-12 所示的输出内容保存对话框，指定是否保存 SPSS 查看器窗口的输出项目。

如果对要关闭的文件进行了修改，退出时还会弹出如图 1-13 所示的文件保存对话框，保存后即可正常退出（注意：非正常退出可能引起数据丢失）。

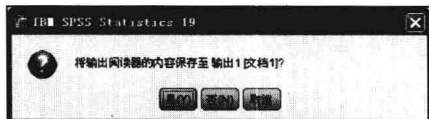


图 1-12 提示保存 SPSS 查看器的输出内容

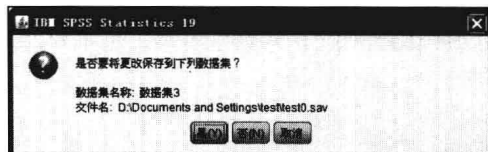


图 1-13 提示更改文件保存对话框