

挖掘数据宝藏
助力决策支持

SPSS

统计分析 从入门到精通

○ 杜强 贾丽艳 编著

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 (C I P) 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, 2009. 3
ISBN 978-7-115-19651-4

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

中国版本图书馆CIP数据核字 (2009) 第006153号

内 容 提 要

本书基于 SPSS 个人版本 SPSS 15.0 for Windows 编写, 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 SPSS 分析软件的实际应用。

全书 25 章, 分 4 个部分。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其对应 SPSS 过程的操作方式, 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。

本书适合自然科学和社会科学各领域、各专业的研究人员多层次的需要, 也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 刘 浩
 - ◆ 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 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
三河市海波印务有限公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 37
字数: 931 千字 2009 年 3 月第 1 版
印数: 1-4 000 册 2009 年 3 月河北第 1 次印刷

ISBN 978-7-115-19651-4/TP

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

SPSS 是面向行业应用人员的统计分析软件，它的统计方法成熟，具有实用性、易用性以及与其他软件（如文字处理软件等）交互良好的特点。SPSS 操作简便深受普通用户所喜爱。同时，它包括高效的数据处理编程环境、自动化程序开发设计、扩展功能开发和二次开发等许多实用并且具有挑战性的功能，也适用于专业人士。

本书基于 SPSS 15.0 个人版本编写，目的是让读者与 SPSS 全面接触，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，本书采用 SPSS 的窗口和对话框操作方式进行介绍，着重突出 SPSS 分析软件的实用性。

SPSS 15.0 版本有许多突出的新特点和新功能，例如：使用次序回归（ordinal regression）算法对两种以上变量的次序输出进行预测；新增了广义线性模型（GZLMs）和广义估计方程（GEEs），可用于处理类型更广泛的统计模型，GZLMs、GEEs 所代表的一体化框架，包括了传统的因变量正态分布线性模型、Logistic 模型、二分变量的 Probit 模型、计数数据的 loglinear 模型，以及各种其他非标准的回归模型；从 SPSS 14 开始使用的高度可视化的图表构造交互界面——图形构建器，在 SPSS 15 中得到了进一步的加强，不仅增加了许多新的图形样式（例如双轴线图形等），还可以输出带标签的 PDF 格式的结果；关于更多的功能，将在书中不同章节详细介绍。

写作上，本书图文搭配，而且图形上给出了指导性的图释，使读者能更方便地定位相关细节，再结合文字叙述部分理解具体含义，快速掌握操作的具体方式和步骤，以节省读者的宝贵时间。

本书全面、系统地介绍了 SPSS 15.0 各个方面的操作，全书共 25 章，分如下 4 个部分：

- 第 1~3 章介绍 SPSS 主要界面和一些基本概念，重点讲解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其对应 SPSS 过程的操作方式，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。每一类又包含多个 SPSS 统计分析过程，比如回归分析中又分线性回归、曲线回归、Logistic 回归、Probit 回归、Ordinal 回归等多个分析过程。在第 12 章，还详细介绍了分类和预测算法中的决策树分析。
- 第 19 章介绍各种统计图形的生成和编辑，SPSS 15.0 的绘图系统能根据用户需要更方便地绘制各种图形。
- 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、多因素试验设计等应用。从问题的背景知识开始，到数据的搜集、整理和融合，再到模型的选择、参数的设置，最后到结果的分析，都给出了基本的分析原理和操作步骤。这些案例把统计思想、SPSS 操作、实际应用很好地联系起来，希望从系统性、完整性和实用性方面为读者提供有价值的参考。

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；而且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，书中作简单铺垫后，给出算法的基本思想和关键步骤，重点介绍在软件操作方法和参数设置上，以期达到读者通过本书可以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。

本书由多人合作完成，作者具有统计学专业背景和从事数据分析的工作经验，其中：贾丽艳负责第 5、7、8、9 章，鲍青波负责第 3、4 章，杜强负责第 1、2、6、10~19 章，李永刚、尹婧、董桂梅、卢文涛、孙志伟、邓博参与整理了最后 6 章的案例；参与编写的还有：马娜、孙洁、刘新梅、平海燕、张纬、魏铁军、马晓东、周永涛、于艳丽等；杜强、贾丽艳负责全书的内容校审和修改。在此，感谢两位授业恩师阮吉寿教授、宋向东教授对我的教诲；感谢所有参与编写人员的辛勤工作。

由于水平有限，书中纰漏和不足会有许多，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学习，共同进步（电子邮箱为：book_better@sina.com）。

作者
2009 年 1 月

第 1 章 SPSS 15.0 概述	1	2.4.1 在单元格中编辑数据	38
1.1 SPSS 简介	1	2.4.2 插入变量与删除变量	38
1.2 SPSS 的安装、启动和退出	3	2.4.3 插入观测量与删除观测量	39
1.2.1 SPSS 15.0 的安装	3	2.4.4 数据的剪切、复制和粘贴	39
1.2.2 SPSS 的启动	3	2.4.5 撤销操作	41
1.2.3 SPSS 15.0 的退出	5	2.5 对数据文件的操作	41
1.3 SPSS 15.0 的界面及设置	5	2.5.1 数据文件的打开与保存	41
1.3.1 常用界面	6	2.5.2 数据库文件的转换	41
1.3.2 通用 General 功能参数	9	第 3 章 数据文件的操作	50
1.3.3 Viewer 视图窗口参数	11	3.1 数据文件的一般操作	50
1.3.4 Draft Viewer (草稿窗口) 参数	12	3.1.1 数据排序	50
1.3.5 Output Labels 输出标签参数	14	3.1.2 数据文件的分组	51
1.3.6 Charts 图形参数	14	3.1.3 数据文件的合并	52
1.3.7 Interactive 交互图形窗口参数	16	3.1.4 数据文件的转置	55
1.3.8 Pivot Table 枢纽表参数	17	3.1.5 变量取值的求秩	56
1.3.9 Data 数据参数	19	3.1.6 变量值的重新编码	58
1.3.10 Currency 数值型变量格式参数	20	3.1.7 计算新变量	62
1.3.11 Scripts 脚本编辑窗口	21	3.2 分类汇总	64
第 2 章 数据文件的建立与操作	22	3.2.1 数据描述	65
2.1 数据编辑器与数据文件	22	3.2.2 分类汇总的参数设置	65
2.1.1 数据编辑器	22	3.2.3 分类汇总的结果	67
2.1.2 数据文件	25	3.3 观测量的加权	68
2.2 常量、变量、操作符和表达式	25	3.4 数据文件的结构重组	69
2.2.1 常量与变量	26	3.4.1 选择数据重组方式	70
2.2.2 操作符与表达式	30	3.4.2 变量组到观测量组的重组	71
2.2.3 如何定义一个变量	31	3.4.3 观测量组到变量组的重组	75
2.2.4 概率事件	35	3.4.4 转置重组	78
2.3 输入数据	35	第 4 章 基本统计分析功能	79
2.3.1 输入数据的方法	36	4.1 OLAP 在线分析过程	79
2.3.2 查看文件信息和变量信息	36	4.1.1 数据描述	79
2.4 编辑数据文件	38	4.1.2 OLAP 过程的操作和设置	79

4.2 观测的摘要报告分析	83	6.2.2 数据和问题描述	121
4.2.1 观测摘要分析的参数设置	83	6.2.3 卡方检验实例分析	121
4.2.2 输出结果	85	6.3 二项检验	123
4.3 行和列的摘要报告分析	85	6.3.1 原理与方法	123
4.3.1 行形式摘要报告	86	6.3.2 数据和问题描述	123
4.3.2 列形式摘要报告	90	6.3.3 二项检验实例分析	123
4.4 频数分析	93	6.4 游程检验	124
4.4.1 数据描述	93	6.4.1 原理与方法	125
4.4.2 对分类变量的频数分析	93	6.4.2 数据和问题描述	125
4.4.3 对连续变量的频数分析	95	6.4.3 游程检验实例分析	126
4.5 描述性统计分析	97	6.5 Kolmogorov-Smirnov 单样本检验	126
4.5.1 数据描述	97	6.5.1 原理与方法	127
4.5.2 Descriptives 分析	97	6.5.2 数据和问题描述	127
4.6 探索分析过程	98	6.5.3 K-S 单样本检验实例分析	127
4.6.1 数据描述	99	6.6 两独立样本检验	128
4.6.2 Explore 实例分析	99	6.6.1 原理与方法	129
4.7 列联表分析过程	102	6.6.2 数据和问题描述	130
4.7.1 数据描述	103	6.6.3 两独立样本检验实例分析	130
4.7.2 列联表分析的参数设置	103	6.7 k 个独立样本的检验	131
4.7.3 列联表分析的输出结果	106	6.7.1 原理与方法	132
第 5 章 均值比较和 T 检验	108	6.7.2 数据和问题描述	132
5.1 Means 过程	109	6.7.3 k 个独立样本检验实例分析	133
5.1.1 原理与方法	109	6.8 两个相关样本的检验	134
5.1.2 SPSS 实例分析	109	6.8.1 原理与方法	134
5.2 单样本 T 检验	111	6.8.2 数据和问题描述	135
5.2.1 原理与方法	111	6.8.3 两个相关样本检验的实例分析	135
5.2.2 SPSS 实例分析	112	6.9 k 个相关样本的检验	136
5.3 两独立样本 T 检验	113	6.9.1 原理与方法	137
5.3.1 原理与方法	113	6.9.2 数据和问题描述	138
5.3.2 SPSS 实例分析	114	6.9.3 k 个相关样本检验的实例分析	138
5.4 配对样本 T 检验	115	第 7 章 多重响应分析	140
5.4.1 原理与方法	115	7.1 多重响应概述	140
5.4.2 SPSS 实例分析	116	7.2 多重响应变量集的定义	140
第 6 章 非参数检验	118	7.2.1 定义多重响应变量集的实例	141
6.1 非参数检验的简介	118	7.3 多重响应变量集的频数分析	142
6.1.1 非参数检验与参数检验	118	7.3.1 多重响应变量频数分析的实例	142
6.1.2 非参数检验的优点	119	7.4 多重响应变量集的交叉表分析	144
6.1.3 非参数检验的缺点	119	7.4.1 多重响应变量交叉表分析的实例	144
6.2 卡方检验	119	7.5 使用 Tables 过程研究多重响应变量集	146
6.2.1 原理与方法	120		

7.5.1 多重响应变量集的定义	146	8.8.1 加权回归分析简介	202
7.5.2 用 Tables 过程建立包含多重 响应变量集的表格	147	8.8.2 问题描述和数据准备	203
第 8 章 回归分析	150	8.8.3 加权回归的参数设置	203
8.1 线性回归	150	8.8.4 案例的结果分析	204
8.1.1 一元线性回归的基本原理	150	8.9 二阶段最小二乘回归	205
8.1.2 多元线性回归的基本原理	152	8.9.1 二阶段最小二乘回归的 基本原理	206
8.1.3 模型假设的其他检验	153	8.9.2 问题描述和数据准备	206
8.1.4 问题描述和数据准备	154	8.9.3 二阶段最小二乘回归的 参数设置	207
8.1.5 线性回归分析的设置和操作	154	8.9.4 案例的结果分析	208
8.1.6 案例的结果分析	159	8.10 最优尺度回归	209
8.2 曲线回归	162	8.10.1 最优尺度回归原理	209
8.2.1 曲线回归的基本原理	162	8.10.2 问题描述和数据准备	209
8.2.2 问题描述和数据准备	163	8.10.3 最优尺度回归的参数设置	210
8.2.3 曲线回归分析的设置和操作	163	8.10.4 案例的结果分析	214
8.2.4 案例的结果分析	165	第 9 章 方差分析	217
8.3 非线性回归	166	9.1 方差分析简介	217
8.3.1 非线性回归简介	167	9.1.1 t 检验与方差分析的比较	217
8.3.2 问题描述和数据准备	168	9.1.2 方差分析的基本原理	218
8.3.3 非线性回归的参数设置	169	9.2 单因素方差分析	220
8.3.4 案例的结果分析	173	9.2.1 原理与方法	220
8.4 二元 Logistic 回归	173	9.2.2 单因素方差分析实例	220
8.4.1 二元 Logistic 回归的数学原理	174	9.3 多因素方差分析过程	225
8.4.2 问题描述和数据准备	175	9.3.1 原理与方法	225
8.4.3 二元 Logistic 回归的参数设置	176	9.3.2 二因素方差分析实例	228
8.4.4 案例的结果分析	180	9.3.3 协方差分析实例	236
8.5 多元 Logistic 回归分析	184	9.3.4 交互效应中随机因素的分析	238
8.5.1 多元 Logistic 回归的原理简介	184	9.4 多元方差分析	242
8.5.2 问题描述和数据准备	184	9.4.1 原理与方法	242
8.5.3 多元 Logistic 回归参数设置	185	9.4.2 多元方差分析实例	243
8.5.4 案例的结果分析	189	9.5 重复测量设计的方差分析	244
8.6 Ordinal 回归	191	9.5.1 原理与方法	244
8.6.1 问题描述和数据准备	192	9.5.2 SPSS 实例分析	245
8.6.2 Ordinal 回归的参数设置	192	9.6 方差成分分析	250
8.6.3 案例的结果分析	196	9.6.1 原理简介	250
8.7 概率单位回归分析	197	9.6.2 SPSS 实例分析	250
8.7.1 概率单位回归分析简介	198	9.7 正交实验设计	253
8.7.2 问题描述和数据准备	198	9.7.1 正交实验设计简述	253
8.7.3 概率单位回归的参数设置	199	9.7.2 SPSS 实例分析	254
8.7.4 案例的结果分析	200	9.7.3 正交实验设计的方差分析	256
8.8 加权回归分析	202		

第 10 章 相关分析	257	12.3.4 案例的结果分析	295
10.1 相关分析的基本概念	257	12.3.5 对聚类结果的进一步分析	296
10.1.1 相关分析的特点和应用	257	12.4 两阶段聚类分析	298
10.1.2 相关系数的计算	258	12.4.1 两阶段聚类简介	298
10.1.3 SPSS 提供的相关分析功能	259	12.4.2 问题描述和数据准备	299
10.2 两变量相关分析	260	12.4.3 SPSS 两阶段聚类的设置	299
10.2.1 问题描述和数据准备	260	12.4.4 案例的结果分析	304
10.2.2 相关分析的参数设置	260	12.5 一般判别分析	307
10.2.3 案例的结果分析	261	12.5.1 判别分析的基本原理	307
10.3 偏相关分析	262	12.5.2 问题描述和数据准备	308
10.3.1 偏相关分析的基本原理	262	12.5.3 判别分析的参数设置	309
10.3.2 偏相关分析实例	263	12.5.4 案例的结果分析	312
10.4 距离分析	264	12.6 逐步判别分析实例	315
10.4.1 距离分析的基本概念	265	12.6.1 问题描述和数据准备	315
10.4.2 距离分析的参数设置	265	12.6.2 逐步判别的参数设置	316
10.4.3 距离分析实例	269	12.6.3 案例的结果分析	318
第 11 章 因子分析	271	12.7 决策树分析	321
11.1 因子分析的原理简介	271	12.7.1 决策树分类的基本原理	321
11.1.1 因子分析的基本思想	271	12.7.2 决策树过程的参数设置	323
11.1.2 因子分析和主成分 分析的联系	272	12.7.3 问题描述和数据准备	338
11.1.3 因子分析的基本步骤	272	12.7.4 案例分析	338
11.2 SPSS 因子分析的应用实例	273	第 13 章 生存分析	344
11.2.1 数据描述	273	13.1 生存分析简介	344
11.2.2 SPSS 因子分析过程的设置	274	13.1.1 生存分析的基本概念	344
11.2.3 结果分析	278	13.1.2 生存分析的数据特点	346
第 12 章 分类分析	283	13.1.3 生存分析的常用方法	346
12.1 聚类分析的原理简介	283	13.1.4 SPSS 中的生存分析过程	346
12.1.1 聚类分析的基本概念	283	13.2 生命表分析	347
12.1.2 聚类分析的一般原理	284	13.2.1 生命表分析简介	347
12.2 快速样本聚类过程	286	13.2.2 生命表分析的基本步骤	347
12.2.1 快速聚类简介	286	13.2.3 生命表实例分析	348
12.2.2 问题描述和数据准备	286	13.3 Kaplan-Meier 分析	352
12.2.3 SPSS 快速聚类的设置	287	13.3.1 Kaplan-Meier 分析的步骤	352
12.2.4 案例的结果分析	289	13.3.2 生存曲线的比较和检验	352
12.3 分层聚类	291	13.3.3 Kaplan-Meier 分析的实例	353
12.3.1 分层聚类简介	291	13.4 Cox 回归模型	357
12.3.2 问题描述和数据准备	291	13.4.1 Cox 回归模型的原理简介	357
12.3.3 SPSS 分层聚类的设置	292	13.4.2 Cox 回归实例分析	358
		第 14 章 信度分析	366
		14.1 信度分析	366

14.1.1 信度分析的基本原理	366	16.3.1 Logit 过程概述	412
14.1.2 问题描述和数据准备	368	16.3.2 问题描述和数据准备	412
14.1.3 信度分析的参数设置	368	16.3.3 Logit 过程的参数设置	412
14.1.4 案例的结果分析	370	16.3.4 案例的结果分析	413
14.2 多维尺度分析	371	16.4 Model Selection 过程	415
14.2.1 多维尺度分析简介	371	16.4.1 Model Selection 过程概述	415
14.2.2 问题描述和数据准备	371	16.4.2 问题描述和数据准备	415
14.2.3 ALSICAL 过程的参数设置	371	16.4.3 层次对数线性模型的操作 过程	416
14.2.4 案例的结果分析	374	16.4.4 案例的结果分析	417
第 15 章 时间序列分析	377	第 17 章 对应分析	419
15.1 SPSS15 的时间序列分析概览	377	17.1 对应分析的基本原理	419
15.1.1 Create Models 的通用 设置选项	378	17.1.1 对应分析与因子分析	419
15.1.2 Apply Models 的通用 设置选项	384	17.1.2 SPSS 中的对应分析	420
15.2 时间序列数据的预分析	384	17.1.3 使用对应分析的注意事项	420
15.2.1 缺失值替换	385	17.2 简单对应分析	421
15.2.2 定义时间变量	385	17.2.1 简单对应分析的数学原理	421
15.2.3 时间序列的平稳化	386	17.2.2 SPSS 简单对应分析实例	422
15.3 指数平滑模型	388	17.3 多元对应分析	427
15.3.1 指数平滑的基本原理	389	17.3.1 多元对应分析基本概念及 其特点	428
15.3.2 指数平滑模型的参数设置	389	17.3.2 多元对应分析的参数设置	428
15.3.3 指数平滑模型实例分析	391	17.3.3 实例的结果分析	435
15.4 ARIMA 模型	395	第 18 章 缺失值分析	438
15.4.1 ARIMA 模型的基本原理	395	18.1 缺失值分析的概念	438
15.4.2 ARIMA 模型的参数设置	397	18.1.1 缺失值的表现方式	438
15.4.3 ARIMA 模型实例分析	398	18.1.2 SPSS 中的缺失值处理方法	439
15.5 季节分解模型	401	18.2 缺失值分析的参数设置	440
15.5.1 季节分解法概述	401	18.3 缺失值分析的实例	444
15.5.2 季节分解模型实例分析	402	第 19 章 统计图形	449
第 16 章 对数线性模型	406	19.1 概述	449
16.1 对数线性模型概述	406	19.1.1 数据和变量的准备	449
16.1.1 简单列联表分析的不足	406	19.1.2 图形构建器的基本操作	451
16.1.2 对数线性模型的基本形式	406	19.1.3 交互式作图和对话框作图	452
16.2 General 过程	407	19.1.4 图形的编辑	453
16.2.1 General 过程概述	407	19.2 条形图	453
16.2.2 问题描述和数据准备	408	19.2.1 数据和问题描述	453
16.2.3 General 过程的参数设置	408	19.2.2 用图形构建器作条形图	453
16.2.4 案例的结果分析	410	19.2.3 交互式条形图	457
16.3 Logit 过程	411		

19.2.4 用对话框创建条形图	459	19.12.1 数据和问题描述	499
19.3 线形图	460	19.12.2 用图形构建器作直方图	499
19.3.1 数据和问题描述	461	19.13 P-P 概率图	500
19.3.2 用图形构建器作线形图	461	19.13.1 数据和问题描述	500
19.3.3 交互式线形图	462	19.13.2 用对话框创建帕 P-P 概率图	501
19.3.4 用对话框创建线形图	464	19.14 Q-Q 概率图	502
19.4 面积图	465	19.14.1 数据和问题描述	503
19.4.1 数据和问题描述	465	19.14.2 用对话框创建 Q-Q 概率图	503
19.4.2 用图形构建器作面积图	465	19.15 时间序列图	504
19.4.3 交互式面积图	467	19.15.1 普通序列图	504
19.4.4 用对话框创建面积图	467	19.15.2 自相关序列图	507
19.5 饼图	468	19.15.3 互相关序列图	509
19.5.1 数据和问题描述	468	19.16 双轴线图	511
19.5.2 用图形构建器作饼图	468	19.16.1 数据和问题描述	511
19.5.3 交互式饼图	469	19.16.2 用图形构建器作双轴线图	511
19.5.4 用对话框创建饼图	470	第 20 章 上市公司财务危机预警分析	513
19.6 高低图	470	20.1 财务危机预警的应用简介	513
19.6.1 数据和问题描述	471	20.1.1 财务危机的定量定义方法	513
19.6.2 用图形构建器作高低图	471	20.1.2 财务危机预警的模型选择	514
19.6.3 交互式高低图	472	20.2 数据描述	514
19.6.4 用对话框创建高低图	473	20.2.1 数据说明	514
19.7 帕累托图	477	20.2.2 指标选择	515
19.7.1 数据和问题描述	478	20.2.3 补充说明	515
19.7.2 用对话框创建帕累托图	478	20.3 分析方法概述	516
19.8 控制图	479	20.3.1 判别分析	516
19.8.1 数据和问题描述	480	20.3.2 logistic 回归方法	516
19.8.2 用对话框创建控制图	480	20.4 SPSS 建模过程和结论分析	517
19.9 箱图	486	20.4.1 SPSS 数据筛选操作	517
19.9.1 数据和问题描述	486	20.4.2 SPSS 判别分析建模与分析	521
19.9.2 用图形构建器作箱图	486	20.4.3 logistic 回归建模与分析	525
19.9.3 交互式箱图	488	20.5 进一步的分析与应用	528
19.9.4 用对话框创建箱图	489	20.5.1 分类结果的应用分析	528
19.10 误差条图	490	20.5.2 建模方法的改进	529
19.10.1 数据和问题描述	490	20.6 建议和推广	529
19.10.2 交互式误差条图	490	20.6.1 时间序列研究	529
19.10.3 用对话框创建误差条图	491	20.6.2 数据的有效预警期	529
19.11 散点图	492	20.6.3 指标的简化方法	529
19.11.1 数据和问题描述	492	第 21 章 影响汇率的因素分析	531
19.11.2 用图形构建器作高低图	493	21.1 汇率影响因素的简介	531
19.11.3 交互式散点图	495	21.2 数据描述	532
19.11.4 用对话框创建散点图	498		
19.12 直方图	498		

21.3 分析方法概述	533	23.1.2 指标选取	557
21.3.1 探索性分析	533	23.1.3 数据格式	557
21.3.2 多元回归分析	534	23.2 聚类分析法简述	557
21.4 SPSS 建模过程和结论分析	534	23.3 SPSS 建模过程和结论分析	558
21.4.1 数据准备	534	23.3.1 对本科院校进行聚类的 设置操作	558
21.4.2 探索性分析	535	23.3.2 对本科院校的分析	561
21.4.3 多元回归分析	536	23.4 建议和推广	563
21.5 进一步的分析和应用	539	第 24 章 试卷信度的检验与分析	565
21.5.1 剔除存在共线性的外汇 储备变量	540	24.1 试卷信度检验的背景简介	565
21.5.2 回归模型的进一步改进	540	24.1.1 测验内容的自身方面	565
21.5.3 两个回归模型比较	541	24.1.2 施测过程	565
21.6 建议和推广	542	24.1.3 被测试者的自身因素	566
21.6.1 时间序列研究	542	24.2 数据描述	566
21.6.2 汇率影响因素的定性分析	542	24.3 分析方法概述	566
第 22 章 因子分析在成绩综合评价 中的应用	543	24.3.1 试卷信度的基本计算公式	566
22.1 学生成绩的综合评价简介	543	24.3.2 试卷信度的估计方法	567
22.2 数据描述	544	24.4 SPSS 建模过程和结论分析	568
22.3 分析方法概述	544	24.4.1 SPSS 信度分析的参数设置	568
22.3.1 应用因子分析进行成绩 综合评价的步骤	544	24.4.2 结果分析	569
22.3.2 应用因子分析法进行成绩 综合评价的注意事项	545	24.5 建议和推广	570
22.4 SPSS 建模过程和结论分析	546	第 25 章 多因素试验的设计与分析	571
22.4.1 数据准备	546	25.1 试验设计简介	571
22.4.2 SPSS 因子分析建模与分析	548	25.1.1 试验设计的应用	571
22.5 进一步的分析和应用	553	25.1.2 试验设计问题的解决步骤	572
22.6 建议和推广	553	25.2 数据描述	573
22.6.1 高中生的成绩综合评价	553	25.3 分析方法概述	573
22.6.2 对缺失数据的处理	554	25.3.1 正交设计方法	573
22.6.3 多种方法结合的综合 评价模型	554	25.3.2 综合评分方法	575
第 23 章 高等教育办学条件的聚类分析	555	25.4 SPSS 建模过程和结论分析	575
23.1 数据描述	555	25.4.1 数据标准化	576
23.1.1 关于基本办学条件指标 合格与否的判定	555	25.4.2 性能指标权重的确定	577
		25.4.3 利用权重求综合指标	578
		25.4.4 对综合得分的进一步分析	578
		25.5 建议和推广	579

第 1 章 SPSS 15.0 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原始数据,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。

本章详细介绍 SPSS 15.0 软件环境的设置内容和设置方式,打造适合自己的 SPSS 15.0 工作环境。

1.1 SPSS 简介

业界领先的统计分析软件提供商 SPSS 公司,推出的旗舰统计分析软件 SPSS 15.0 除了继承原有产品的特点之外,还增加了许多显著的新特点。这些增强特性包括:SPSS 15.0 提供了更强大的数据管理功能,能帮助用户更加方便地使用其他应用程序和数据库;新式的图表能够让用户将复杂的信息清晰地表现出来,SPSS 15.0 进一步增强了图形构建器的功能——高度可视化的图表创建界面;SPSS 15.0 还包括了次序回归 (Ordinal Regression) 分析算法,对两种以上变量的次序输出进行预测,例如预测客户忠诚度及其与客户满意度的相关性。另外,在统计模型与算法、可编程性等方面,SPSS 15.0 都进行了功能更完善、使用更便利等方面的改进。

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,因此用户无须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 语言的用户,也可以在语句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 Paste 按钮,可把当前对话框设置对应的语句,自动粘贴到命令语句窗口中,并允许保存为文件。因此 SPSS 能既适用于新用户也适用于老用户。

(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,无须告诉系统怎样做。用户只需了解统计分析原理,无须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,生成的表格和交互图对象可以与同类软件进行自动嵌入与链接。它还有内置的 VBA 客户语言(Sax Basic),能进行编程。

(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

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

SPSS 拥有强大的图形功能,能生成数十种基本图和交互图。基本图包括条形图、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。交互图比基本图漂亮,有二维和三维形式,包括条形交互图、点形交互图、线形交互图、带形交互图、饼形交互图、箱形交互图、误差条形交互图、直方交互图和散点交互图等。同表格一样,图形生成以后,也可以进行编辑。SPSS 的输出图形可以保存为多种格式。

通过直观、漂亮的统计图形,能形象地显示分析结果,如图 1-1 所示。

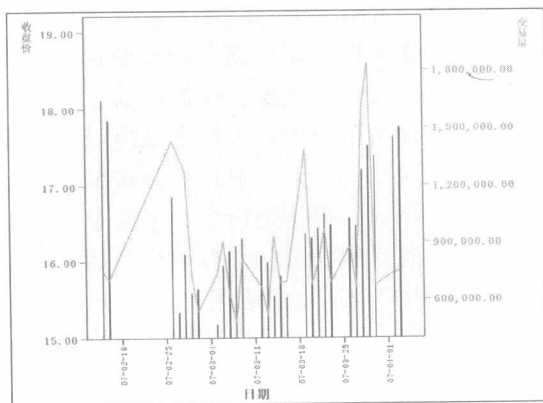


图 1-1 SPSS 生成图形

2. SPSS 15.0 的不同版本

对普通用户而言,可能会用到两个版本:个人版 SPSS 15.0 和服务器版 SPSS Server 15.0。

(1) 使用个人版 SPSS,能方便的实现数据过滤、数据筛选、统计分析和特定结果输出等功能,不仅帮助用户挖掘到隐含在大量数据背后的知识,而且能节省用户宝贵的时间。现在的个人电脑一般都能支持 SPSS 15.0 的运行,可适用操作系统为 Microsoft Windows XP 或 Windows 2000。

(2) 把 SPSS Server 15.0 安装在服务器上,把个人版 SPSS 15.0 安装在客户端上,通过分布模式连接服务器,这样用户在客户端就能直接存取服务器上的数据,并能在服务器上运行分析过程。SPSS Server 15.0 配有专属的结果发布系统,能在更广的范围以更方便的方式分享分析结果。

1.2 SPSS 的安装、启动和退出

本节介绍 SPSS 安装和启动操作,从这里我们开始踏上 SPSS 应用旅程。

1.2.1 SPSS 15.0 的安装

装入 SPSS 15.0 的安装盘,自动运行后弹出如图 1-2 所示的选择界面,单击 Install SPSS (若安装其他组件可单击相应项目),弹出如图 1-3 所示的界面,提示检查当前的系统环境。若光盘内容未自动运行,可右键单击光驱盘符,在弹出的菜单中选择“自动播放”,或者直接双击光驱盘符运行。

待图 1-3 所示的界面自动完成后,弹出如图 1-4 所示的界面,一直单击 Next 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。

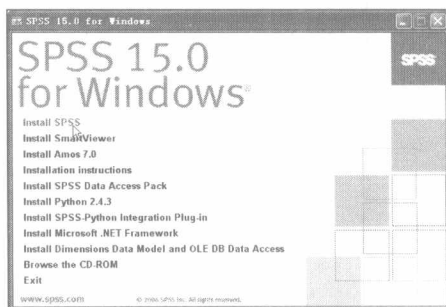


图 1-2 SPSS 安装 1

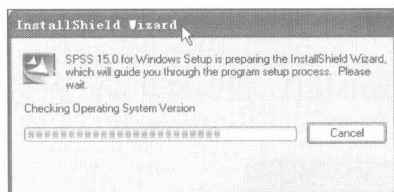


图 1-3 SPSS 安装 2

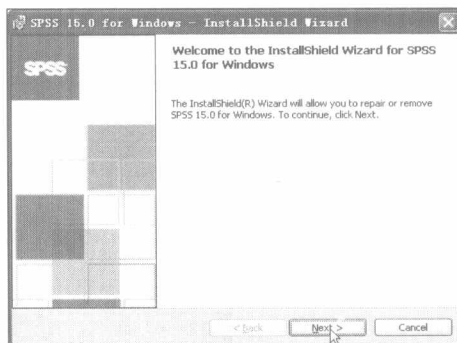


图 1-4 SPSS 安装 3

1.2.2 SPSS 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,可以双击如图 1-5 所示的桌面图标,也可以在“开始”菜单中依次单

击“程序→SPSS for Windows→SPSS15.0 for Windows”。启动后，会出现图 1-6 所示的启动界面，显示本软件的版本和注册信息，如果没有出现界面底部的注册信息，需要按后面介绍的方法进行注册，或者使用试用版本。



图 1-5 SPSS 的桌面图标

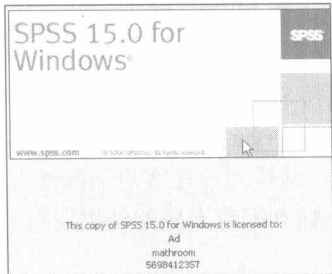


图 1-6 SPSS 的启动界面

之后会出现如图 1-7 所示的选择界面，表明 SPSS 已正常启动。单击选中“Run the tutorial”单选框，再单击“OK”按钮，会打开如图 1-8 所示的图形教程界面，它是一个学习 SPSS 操作的捷径。

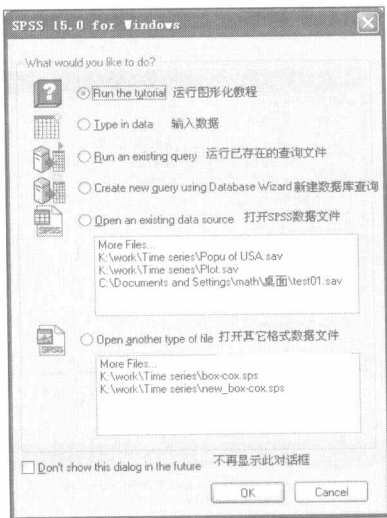


图 1-7 SPSS 启动选项

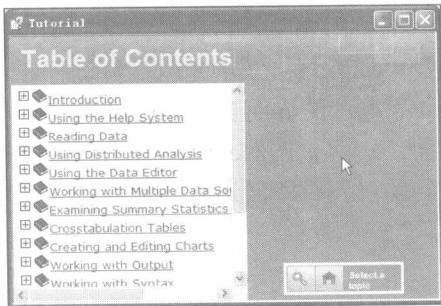


图 1-8 图形教程

2. SPSS 的注册方法

在桌面的“开始”菜单中单击如图 1-9 所示的 Wizard 项目，进入如图 1-10 所示的 SPSS 注册界面，单击“Start”后开始注册。

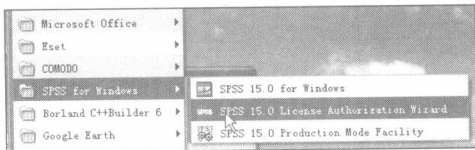


图 1-9 SPSS 注册 1

3. 其他 3 种运行方式

(1) 批处理方式。在桌面的“开始”菜单中依次单击“程序→SPSS for Windows→SPSS 15.0

Production Mode Facility”，打开如图 1-11 所示的操作界面，它用来运行指定的 Syntax 程序文件。

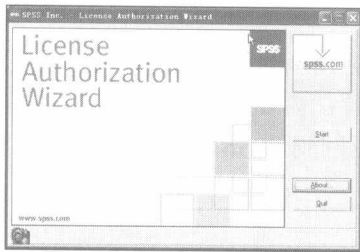


图 1-10 SPSS 注册 2

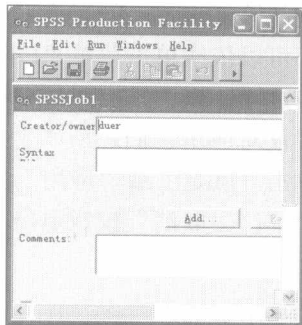


图 1-11 SPSS Production Facility 界面

(2) 程序运行方式。直接在 Syntax 语句窗口或 Script 脚本窗口编辑和运行程序，这种方式要求掌握 SPSS 的 Syntax 语言或 Sax Basic 脚本语言。

(3) 混合运行方式。首先在“完全窗口”方式下的数据编辑窗口中输入数据，或者利用“File”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 Paste 按钮，将选择的过程及参数转换成相应的命令语句，置于 Syntax 语句窗口；最后，在语句窗口中添加语句、参数，或者修改已有命令中的参数，单击窗口中的 Run 按钮执行分析。

1.2.3 SPSS 15.0 的退出

单击菜单“File→Exit”，或者单击 Data Editor 窗口右上角的关闭按钮，都可以退出 SPSS。

第一次退出时会弹出如图 1-12 所示的提示框，提示用户这是退出操作，勾选复选框后，再单击 Yes 按钮，下次退出时就不会再显示这个提示框了。

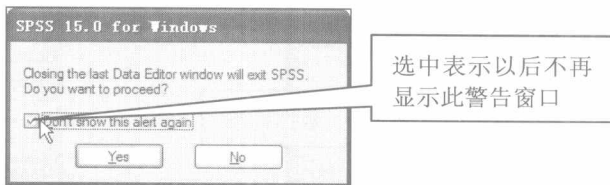


图 1-12 SPSS 退出提示框

如果对文件进行了修改，或者输出了新的结果，退出时还会弹出如图 1-13 所示的文件保存对话框，对相关内容保存后即可正常退出（注意：非正常退出可能引起数据丢失）。

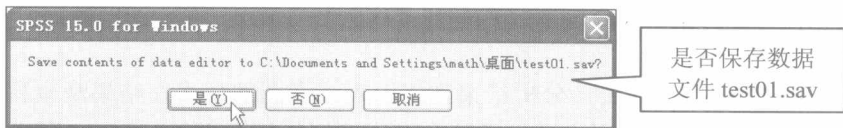


图 1-13 SPSS 提示文件保存对话框

1.3 SPSS 15.0 的界面及设置

本节首先介绍 SPSS 常用的几个界面及其功能，对大多数用户来讲，SPSS 是“窗口+对话框”方式的应用工具，熟悉和了解这些界面，对提高使用 SPSS 进行统计分析的工作效率