

从基础到实践丛书

# SPSS 统计分析 从基础到实践

罗应婷 杨钰娟 编著  
易为科技 审校



配套光盘提供了本书所有输入数据文件和典型结果文件。



电子工业出版社  
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY  
<http://www.phei.com.cn>

从基础到实践丛书

# SPSS 统计分析 从基础到实践

罗应婷 杨钰娟 编著  
易为科技 审校

电子工业出版社

Publishing House of Electronics Industry  
北京·BEIJING

## 内 容 简 介

本书作为 SPSS 的实用参考手册，是基于 SPSS 14.0 版本进行编写的，共 14 章。书中既详细介绍了 SPSS 各菜单的使用方法，又给出了其相应统计方法的基本原理和适用条件。同时，对于复杂的统计方法，都通过引例讲解说明。这样有利于实际工作者学习和真正熟练运用 SPSS 的强大统计功能。

同时，本书最后给出了 SPSS 在各个应用领域的使用实例，其中所用到的统计方法和思想也可以作为实际工作者在处理具体问题时的一个参考。应用实例涵盖管理决策、生物技术、工程分析、金融系统等领域，所选择的例子不仅具有典型性，而且具有很强的工程参考价值。

本书图文并茂，层次清晰明了，案例丰富多彩，为读者提供愉快的阅读享受。

本书特别适合希望提升数据统计分析能力的管理者，以及从事统计分析、市场分析、社会学、医药统计分析、金融专业的人员。既可以作为利用 SPSS 软件进行数据分析的实际工作者的一本参考手册，也可以作为各大院校学生学习 SPSS 软件的教材。

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 统计分析从基础到实践 / 罗应婷，杨钰娟编著. —北京：电子工业出版社，2007.6  
(从基础到实践丛书)

ISBN 978-7-121-04263-8

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

中国版本图书馆 CIP 数据核字 (2007) 第 056304 号

责任编辑：胡铭娅

印 刷：北京智力达印刷有限公司

装 订：北京中新伟业印刷有限公司

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：24.5 字数：555 千字

印 次：2007 年 6 月第 1 次印刷

印 数：5000 册 定价：45.00 元（含光盘 1 张）

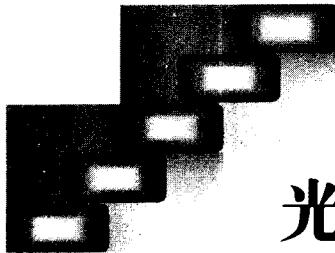
凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，  
联系及邮购电话：(010) 88254888。

质量投诉请发邮件至 [zlts@phei.com.cn](mailto:zlts@phei.com.cn)，盗版侵权举报请发邮件至 [dbqq@phei.com.cn](mailto:dbqq@phei.com.cn)。

服务热线：(010) 88258888。

# 丛 书 特 色

- 坚持“基础为本源，实践出真知”的写作理念，即首先掌握基本理论和基础技能，然后在实践中锻炼提高。
- 丛书内容“全、精、实用”，只要一本书，初学者就能入门，并完成实际工作。
- 循序渐进地介绍基本知识，通过穿插的小实例，以深入浅出之法加深读者的理解和应用能力，同时强调重点、技巧和注意点。
- 带领读者亲自完成多个项目开发。项目为实践中常用的、典型的應用问题。书中所有实例均调试通过。
- 在配书光盘（或网上下载）中，提供所有练习、实例及实战部分的代码和素材，方便读者学习和使用。
- 贴心顾问技术支持 E-mail: jsj@phei.com.cn，及时解答读者在阅读本书中的问题。



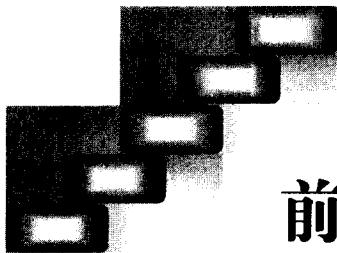
## 光盘使用说明

---

配套光盘包含图书中所有输入数据文件和典型的结果文件，相关内容按照图书章节的顺序归类，以方便读者查找。

- 输入文件的格式包括.sav 和.TXT 两种通用格式；
- 随书的结果输出文件的格式为.spo；
- 相关的输入输出文件都可用 SPSS 软件打开查看；
- 适合的 SPSS 软件版本包括 SPSS10.0 ~ SPSS15.0 的各个软件版本。

该数据文件是为了提高学习效率，请读者按配套图书的对应章节来学习。数据的版权属相关出处，仅供学习研究使用。



## 前 言

作为全球应用最广泛的统计软件，SPSS 已有 30 余年的历史。它广泛应用于农业、工业、商业、医学、社会学、市场分析、股市行情、旅游业等多个行业和领域。同时，由于 SPSS 的操作简捷、界面美观，受到越来越多非统计专业人士的青睐。

目前国内也有很多优秀的 SPSS 参考书。但是它们通常要求读者具有比较高的统计学基础。对于一些没有系统学习过统计学，但是又急切需要应用 SPSS 来处理实际问题的用户而言，这些参考书的起点就显得有些高了。因此，本书的编写目的是为各行业中需要应用 SPSS 软件处理实际问题的非统计专业人士提供一本容易理解、易学、易用的 SPSS 基础教材。

根据作者的经验，非统计专业的用户在使用 SPSS 时通常会遇到如下问题：

- 对 SPSS 中软件提供的各种统计方法的原理和适用条件不清楚，有时胡乱套用方法，出现问题与方法张冠李戴的情况。
- 面对 SPSS 对话框中复杂的选项设置而不知所措。
- 对 SPSS 最后输出的统计图表不知应该如何解读。
- 面对有大量数据的复杂问题时不知道如何应用 SPSS 软件来层层剖析数据，解决问题。

## 本书特色

在本书的编写过程中，一直致力于帮助软件的初学者解决以上问题，我们做了以下尝试和努力。

- **结构清晰，易学易用。**本书采用双线索形式安排文章结构。章节设置既符合“输入数据——整理数据——基础统计分析——高级统计分析”这一传统统计分析过程，又严格按照 SPSS 窗口各菜单顺序行文，便于读者尽快熟悉 SPSS 的操作界面。
- **目录明了，便于查询。**由于大部分实际工作者通常只使用 SPSS 的部分功能，因此，本书的目录编排采取了统计名称与功能键名相结合的方法，便于读者快速查询所需方法的操作过程。这样大大克服了一些书籍直接以功能键为目录而使读者无法快速查阅具体所需功能的缺点。
- **详细介绍了 SPSS 的常用功能所对应的对话框。**对其中选项的具体意义、适用情况都有介绍。这样，即使是英语水平不高的读者，也可以清晰地了解各个对话框中选

项的意义。

- 对于各种统计方法，在具体介绍其在 SPSS 的界面操作之前，都对它们的原理和适用条件做了详细的介绍。这样，使那些即使没有系统学习过统计学的读者也能够恰当地选择正确的统计方法。
- 对于每种统计方法都给出了具体的例子及其在 SPSS 中的实现，并对最后的输出结果做出了详细解释。初学者可以通过这些例子尽快掌握如何应用 SPSS 处理实际问题。
- 在本书的最后，给出了多个 SPSS 在行业中的应用实例。这些实例既有编者做过的实际问题，又有其他统计工作者的成功案例。希望读者通过这些案例的学习，能够感受到 SPSS 在解决实际问题中的应用。同时也能学习到其他统计工作者在处理实际问题的思想，并对自己处理实际问题有所启迪。

## 主要内容

本书以 SPSS 14.0 为基础，但也适用于其他版本的 SPSS 软件使用者。全书共 14 章，分为 SPSS 概述（第 1 章）、数据文件的建立与整理（第 2~5 章）、统计分析（第 6~13 章）、应用实例（第 14 章）四大部分。具体内容包括：SPSS 14.0 概述、SPSS 帮助系统、SPSS 数据文件的建立与编辑、数据整理、统计图、SPSS 报表、描述性统计分析、均值比较与 t 检验、方差分析、相关分析、回归分析、聚类分析、判别分析、因子分析、对应分析、非参数检验、SPSS 的应用实例等。

本书既详细介绍了 SPSS 各菜单常用过程的具体使用方法，又给出了其相应统计方法基本原理和适用条件介绍。同时，对于复杂的统计方法，都通过引例讲解说明。有利于实际工作者学习和真正熟练运用 SPSS 的强大统计功能。

同时，本书最后给出了 SPSS 在各个应用领域的使用实例，其中所用到的统计方法和思想也可以作为实际工作者在处理具体问题时的一个参考。应用实例涵盖：管理决策、生物技术、工程分析、金融系统的应用，所选择的例子具有典型性，而且具有很强的工程参考价值。

本书图文并茂，层次清晰明了，案例丰富多彩，提供读者愉快的阅读享受。

## 本书的读者对象

本书特别适合提升数据统计分析能力的管理者，以及从事统计分析、市场分析、社会学、医药统计分析、金融等专业的人员。本书既可以作为利用 SPSS 软件进行数据分析的实际工作者的一本参考手册，也可以作为各院校学生学习 SPSS 软件的教材。

## 致谢与分工

本书由罗应婷、杨钰娟编著，成都易为科技有限责任公司参与前期的策划和全书的审校工作。北京博文视点资讯公司朱沫红女士参与全程的策划工作。参与具体编辑、排版、

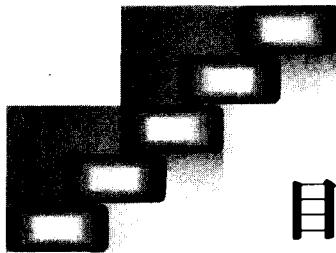
审校工作的还有张赛桥、周忻、王斌、万雷、王晓艳、田原、刘军华、姚新军等。

所谓厚积薄发，在知识的掌握过程中，我们更加强调循序渐进的学习方法。因此，如果要全面掌握 SPSS 的强大功能，读者需要静下心来，认真阅读本书的各个章节。但是对于应用工作者，如果需要尽快掌握 SPSS 来处理实际问题的话，那么由于本书编写过程中一直注意各章节间的独立性，只阅读相关的统计方法部分也可以学习用 SPSS 来处理实际问题的方法。

由于编者水平有限，书中谬误之处在所难免，希望广大专家读者及时指正批评。

作者

2007 年 4 月



# 目 录

## 第 1 篇 SPSS 概述

<b>第 1 章 SPSS14.0 for Windows 基础</b>	<b>2</b>
1.1 SPSS 简介	2
1.1.1 SPSS 的产生与发展	2
1.1.2 SPSS14.0 的新特性	3
1.1.3 SPSS 与其他常用统计软件比较	3
1.1.4 SPSS 的主要应用领域简介	4
1.2 SPSS14.0 的安装、启动与退出	4
1.3 SPSS14.0 窗口简介	6
1.3.1 数据编辑窗口 SPSS Data Editor	7
1.3.2 结果浏览窗口 SPSS Viewer	9
1.3.3 程序编辑窗口 SPSS Syntax Editor	12
1.3.4 结果草稿浏览窗口 SPSS Draft Viewer	12
1.3.5 VBs 宏程序编辑窗口 Script	13
1.4 SPSS14.0 的帮助系统	14
1.4.1 对话框上的 Help 按钮	14
1.4.2 主题词获得帮助—Topics 过程	15
1.4.3 新手入门—Tutorial 过程	16
1.4.4 实例学习—Case Studies 过程	16
1.4.5 统计教练—Statistics Coach 过程	17

1.4.6 语法指南—Command Syntax Reference 过程	17
1.4.7 算法介绍—Algorithms 过程	18
1.4.8 访问 SPSS 官方主页	18

## 第 2 篇 数据文件的建立与整理

<b>第 2 章 SPSS 数据文件的 建立与编辑</b>	<b>20</b>
-----------------------------------	-----------

2.1 数据的输入与编辑	20
2.1.1 SPSS 的变量类型	20
2.1.2 定义新变量	21
2.1.3 数据的录入与编辑	24
2.2 数据文件的创建—File 菜单详解	24
2.2.1 新建 SPSS 数据文件	24
2.2.2 从其他数据文件导入数据	25
2.2.3 File 菜单的其他命令	28
2.3 数据文件的编辑与管理	
—Edit/Utilities 菜单详解	28
2.3.1 Edit 菜单详解	29
2.3.2 Utilities 菜单详解	29

<b>第 3 章 SPSS 数据文件的整理</b>	<b>32</b>
---------------------------	-----------

3.1 数据文件引论	32
------------	----

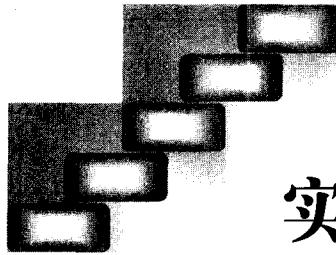
3.1.1 数据文件的整理在实际 问题中的重要性 .....	32	4.1.1 Graph 菜单简介 .....	64
3.1.2 一个数据文件整理的案例 .....	32	4.1.2 常用统计图简介 .....	67
<b>3.2 数据文件的整理——Data</b>		<b>4.2 常见统计图</b> .....	68
菜单详解 .....	33	4.2.1 条形图 (Bar Charts) .....	68
3.2.1 观测量排序——Sort Case 过程 .....	33	4.2.2 线图 (Line Charts) .....	75
3.2.2 数据文件转置——Transpose 过程 .....	34	4.2.3 面积图 (Area Charts) .....	77
3.2.3 数据格式重排——Restructure 过程 .....	35	4.2.4 饼图 (Pie Charts) .....	77
3.2.4 数据文件合并——Merge File 过程 .....	38	4.2.5 高低图 (High-Low Charts) .....	78
3.2.5 数据分类汇总——Aggregate 过程 .....	42	4.2.6 帕累托图 (Pareto Charts) .....	79
3.2.6 数据文件的拆分——Split File 过程 .....	45	4.2.7 质量控制图 (Control Charts) .....	81
3.2.7 选择观测量——Select Cases 过程 .....	46	4.2.8 箱图 (Boxplot) 与误差条图 ( Error Bar ) .....	83
3.2.8 观测量加权——Weight Cases 过程 .....	48	4.2.9 金字塔图 (Population Pyramid) .....	84
3.2.9 Data 菜单其他过程简介 .....	49	4.2.10 散点图 (Scatter/Dot) .....	85
<b>3.3 变量的变换和计算——</b>		4.2.11 直方图 (Histogram) .....	86
Transform 菜单详解 .....	49	4.2.12 P-P 图&Q-Q 图 .....	87
3.3.1 变量计算——Compute 过程 .....	49	4.2.13 普通序列图 (Sequence Charts) .....	89
3.3.2 变量值计数——Count 过程 .....	52	4.2.14 ROC 曲线 .....	91
3.3.3 变量重新赋值——Recode 过程 .....	54	4.2.15 时间序列图 ( Time Series Charts ) .....	93
3.3.4 变量值排序——Rank Cases 过程 .....	57	<b>4.3 SPSS 图形编辑</b> .....	95
3.3.5 变量组段划分—— Visual Bander 过程 .....	59	4.3.1 图形编辑概述 .....	95
3.3.6 自动重新赋值——Automatic Recode 过程 .....	61	4.3.2 图形基本设定——Edit 菜单 .....	96
3.3.7 Transform 菜单其他过程简介 .....	63	4.3.3 图形高级设定——Options 菜单 &Elements 菜单 .....	97
<b>第 4 章 数据的统计图表示 .....</b>	<b>64</b>	4.3.4 图形转换——Transform 菜单 .....	98
4.1 统计图概述 .....	64	<b>4.4 交互式统计图</b> .....	99
5.1 简单记录报表——Reports 子菜单 .....	105	4.4.1 交互式统计图概述 .....	99
5.1.1 在线分析处理——OLAP 过程 .....	105	4.4.2 交互式条图的界面 .....	100
		4.4.3 交互式条图实例 .....	102
		4.4.4 交互式条图的编辑 .....	102
		<b>第 5 章 SPSS 报表功能 .....</b>	<b>105</b>

5.1.2 观测量概述——Case Summaries 过程 ..... 109	6.6.2 引例及结果解释 ..... 152
5.1.3 生成商务报表——Report Summaries in Rows/Columns 过程 ..... 112	<b>第 7 章 均值比较与 t 检验 ..... 154</b>
5.2 高级报表——Tables 子菜单 ..... 120	7.1 t 检验简介 ..... 154
5.2.1 定义复选变量集—— Multiple Response Sets 过程 ..... 120	7.1.1 t 检验的概念及一般步骤 ..... 154
5.2.2 定制报表——Custom Tables 过程 ..... 122	7.1.2 t 检验的类型 ..... 154
5.2.3 Tables 子菜单其他过程简介 ..... 127	7.2 均值描述——Means 过程 ..... 155
<b>第 3 篇 统计分析</b>	7.2.1 Means 过程的操作界面 ..... 155
<b>第 6 章 描述性统计分析 ..... 130</b>	7.2.2 引例及结果解释 ..... 156
6.1 描述性统计量 ..... 130	7.2.3 分组变量的层次说明 ..... 158
6.1.1 描述性统计量 ..... 130	7.3 单样本 t 检验——One-Sample T Test 过程 ..... 159
6.1.2 Descriptive Statistics 子菜单概述 ..... 131	7.3.1 单样本 t 检验的一般步骤 ..... 159
6.2 频数分布表分析—— Frequencies 过程 ..... 132	7.3.2 One-Sample T Test 过程 的操作界面 ..... 159
6.2.1 Frequencies 过程的操作界面 ..... 132	7.3.3 引例及结果解释 ..... 160
6.2.2 引例 ..... 134	7.4 独立两样本 t 检验—— Independent-Sample T Test 过程 ..... 161
6.3 最基础的统计量分析—— Descriptive 过程 ..... 135	7.4.1 独立两样本 t 检验的一般步骤 ..... 161
6.3.1 Descriptive 过程的操作界面 ..... 136	7.4.2 Independent-Sample T Test 过程的操作界面 ..... 162
6.3.2 引例及结果解释 ..... 137	7.4.3 引例及结果解释 ..... 163
6.4 探索性分析——Explore 过程 ..... 137	7.5 配对样本 t 检验—— Paired-Sample T Test 过程 ..... 165
6.4.1 Explore 过程的操作界面 ..... 138	7.5.1 配对样本 t 检验一般步骤 ..... 165
6.4.2 引例及结果解释 ..... 139	7.5.2 Paired-Sample T Test 过程 的操作界面 ..... 166
6.5 列联表分析——Crosstabs 过程 ..... 145	7.5.3 引例及结果解释 ..... 167
6.5.1 Crosstabs 过程的操作界面 ..... 145	<b>第 8 章 方差分析 ..... 169</b>
6.5.2 引例 ..... 148	8.1 方差分析简介 ..... 169
6.5.3 结果解释 ..... 149	8.1.1 方差分析的提出 ..... 169
6.6 相对比描述——Ratio 过程 ..... 150	8.1.2 方差分析的基本概念 ..... 169
6.6.1 Ratio 过程的操作界面 ..... 150	8.1.3 方差分析的类型 ..... 170

8.2 单因素方差分析——One-Way	10.1.1 回归分析的概念.....	212
ANOVA 过程.....	10.1.2 回归分析的应用.....	213
8.2.1 单因素方差分析简介 .....	10.1.3 回归分析的类型.....	213
8.2.2 One-Way ANOVA 过程 的操作界面 .....	10.1.4 回归分析的一般步骤.....	214
8.2.3 引例及结果解释 .....	10.2 线性回归——Linear 过程.....	215
8.3 多因素方差分析——Univariate 过程 (1) .....	10.2.1 线性回归简介.....	215
8.3.1 多因素方差分析简介 .....	10.2.2 Linear 过程的操作界面 .....	217
8.3.2 Univariate 过程的操作界面.....	10.2.3 一元线性回归的例子.....	222
8.3.3 引例及结果解释 .....	10.2.4 多元线性回归的例子.....	225
8.4 协方差分析——Univariate 过程 (2) .....	10.2.5 小结 .....	229
8.4.1 协方差分析简介 .....	10.3 曲线拟合——Curve Estimation 过程 .....	230
8.4.2 引例及结果解释 .....	10.3.1 曲线拟合简介.....	230
8.4.3 小结 .....	10.3.2 Curve Estimation 过程的 操作界面.....	230
<b>第 9 章 相关分析.....</b>	10.3.3 引例及结果解释.....	232
9.1 相关分析简介.....	10.4 二分类变量 Logistic 回归—— Binary Logistic 过程.....	235
9.1.1 相关分析的概念 .....	10.4.1 Logistic 回归简介.....	235
9.1.2 Correlate 子菜单概述.....	10.4.2 Binary Logistic 过程的 操作界面.....	236
9.2 两变量相关分析——Bivariate 过程.....	10.4.3 引例及结果解释.....	239
9.2.1 两变量相关分析简介 .....	10.4.4 小结 .....	243
9.2.2 Bivariate 过程的操作界面.....	10.5 非线性回归——Nonlinear 过程 .....	244
9.2.3 引例及结果解释 .....	10.5.1 非线性回归简介.....	244
9.3 偏相关分析——Partial 过程 .....	10.5.2 Nonlinear 过程的操作界面 .....	244
9.3.1 偏相关分析简介 .....	10.5.3 引例及结果解释.....	248
9.3.2 Partial 过程的操作界面 .....	10.5.4 小结 .....	251
9.3.3 引例及结果解释 .....	<b>第 11 章 聚类分析与判别分析.....</b>	252
9.4 距离分析——Distances 过程 .....	11.1 聚类分析与判别分析相关 原理简介.....	252
9.4.1 距离分析简介 .....	11.1.1 聚类分析.....	252
9.4.2 Distances 过程的操作界面 .....	11.1.2 判别分析.....	252
9.4.3 引例及结果解释 .....	11.2 K-均值聚类分析——K-means Cluster 过程.....	253
<b>第 10 章 回归分析.....</b>		
10.1 回归分析简介 .....		

11.2.1 K-均值聚类法基本原理 .....	253	12.3 最优尺度分析——Optimal Scaling 过程初步认识 .....	306
11.2.2 K-means Cluster 过程 界面操作介绍 .....	253		
11.2.3 引例及结果解释 .....	256		
<b>11.3 系统聚类法——Hierarchical Cluster 过程 .....</b>	<b>258</b>	<b>第 13 章 非参数检验 .....</b>	<b>309</b>
11.3.1 系统聚类法基本原理 .....	258	13.1 非参数检验相关原理简介 .....	309
11.3.2 Hierarchical Cluster 过程 界面操作介绍 .....	258	13.1.1 非参数检验的概念 .....	309
11.3.3 引例及结果解释 .....	261	13.1.2 非参数检验的优缺点 .....	310
<b>11.4 两步聚类法——TwoStep Cluster 过程 .....</b>	<b>267</b>	13.1.3 非参数检验的类型 .....	310
11.4.1 两步聚类法基本原理 .....	267	13.2 分布类型的检验 .....	311
11.4.2 TwoStep Cluster 过程 界面操作介绍 .....	268	13.2.1 卡方检验——Chi-Square 过程 .....	311
11.4.3 引例及结果解释 .....	270	13.2.2 二项分布检验——Binomial 过程 .....	319
<b>11.5 判别分析——Discriminant 过程 .....</b>	<b>276</b>	13.2.3 游程检验——Runs 过程 .....	321
11.5.1 判别分析基本原理 .....	276	13.2.4 单个样本的 K-S 检验 —1-Sample K-S 过程 .....	324
11.5.2 Discriminant 过程界面 操作介绍 .....	277	13.3 分布位置检验 .....	326
11.5.3 引例及结果解释 .....	280	13.3.1 两个独立样本分布位置检验 —2 Independent Samples 过程 .....	326
<b>第 12 章 因子分析与对应分析 .....</b>	<b>285</b>	13.3.2 多个独立样本分布位置检验 —K Independent Samples 过程 .....	329
12.1 因子分析——Factor Analysis 过程 .....	285	13.3.3 两个相关样本分布位置检验 —2 Relate Samples 过程 .....	332
12.1.1 因子分析基本原理 .....	285	13.3.4 多个相关样本分布位置检验 —K Relate Samples 过程 .....	335
12.1.2 Factor Analysis 过程 界面操作介绍 .....	287		
12.1.3 引例及结果解释 .....	291		
<b>12.2 简单对应分析——Correspondence Analysis 过程 .....</b>	<b>301</b>	<b>第 4 篇 应用实例</b>	
12.2.1 简单对应分析基本原理 .....	301		
12.2.2 Correspondence Analysis 过程界面操作介绍 .....	301	<b>第 14 章 SPSS 在各领域的应用实例 .....</b>	<b>340</b>
12.2.3 引例及结果分析 .....	304		
		14.1 SPSS 在房地产决策中的应用 .....	340
		14.1.1 问题描述 .....	341

14.1.2 问题建模 .....	341	14.3.1 问题描述 .....	354
14.1.3 模型的验证 .....	348	14.3.2 问题建模 .....	355
14.2 SPSS 在生物模型中的应用 .....	348	14.3.3 模型的检验 .....	360
14.2.1 问题描述 .....	349	14.4 SPSS 在证券分析中的应用 .....	361
14.2.2 问题建模 .....	349	14.4.1 问题描述 .....	361
14.2.3 模型的讨论 .....	353	14.4.2 问题建模 .....	362
14.3 SPSS 在工程问题中的应用 .....	354	14.4.3 模型的讨论 .....	370



# 实例目录

第3章 SPSS数据文件的整理 ..... 32

例3.1	数据文件的整理	32
例3.2	升序和降序排列	34
例3.3	变量的转置	35
例3.4	将原文件的一组观测量变成两组相关的观测量	36
例3.5	将数据文件以增加观测量的方式添加到其他的文件	38
例3.6	增加变量	41
例3.7	合并后的文件只保留原文件和外部文件中关键变量值相等的观测量	42
例3.8	对变量进行分组，并且计算其算术平均数，计算其最大观测值，分类汇总	45
例3.9	为分组变量，对文件进行拆分，对文件进行排序，比较统计分析结果	46
例3.10	产生两列完全相同的随机数，它们的分布服从标准正态分布 $N(0, 1)$	51
例3.11	对目标变量的特定值计数	53
例3.12	对观测量进行重新赋值	56
例3.13	将变量依据参数进行组段划分	61
例3.14	将字符型变量自动重新赋值为数值型变量	63

第4章 数据的统计图表示 ..... 64

例4.1	由图表生成器绘制变量“Model Year”与变量“Mean Horsepower”关系的条图	66
例4.2	按班级绘制学生语文平均成绩的条图	68
例4.3	学生成绩的分组条图	72
例4.4	绘制各班语文平均成绩的简单线图和各班三门成绩平均分的复式线图	74
例4.5	学生成绩的线图	75
例4.6	商店营业额的饼图	77

例 4.7 股票 10 日价格高低图 .....	78
例 4.8 商店营业额的帕累托图 .....	80
例 4.9 零件质量控制的均数控制图和极差控制图 .....	81
例 4.10 人口收入的金字塔图 .....	84
例 4.11 房屋土地成本、房屋增值 成本、房屋售价之间的矩阵散点图 .....	85
例 4.12 依据突发性流体的实际观测数据，绘制流量关于龙头时间的变化趋势 .....	90
例 4.13 仪器观测准确度的比较，绘制 ROC 曲线比较两个传感器的优劣 .....	91
例 4.14 公司销售额的时间序列图，判断其是否存在周期性变动 .....	93
例 4.15 分段条图的图形编辑 .....	95
例 4.16 绘制交互式条图，并比较它与普通统计图的区别 .....	102

## 第 5 章 SPSS 报表功能 ..... 105

例 5.1 分析概述变量的观测量数、均值、标准差、占全部观测量数的百分比以及极差 .....	108
例 5.2 计算观测量数目、均值和极值，并且不排除缺失值 .....	111
例 5.3 对各个分组及所有观测量统计它们的均值、最小值、最大值和观测量数目 .....	115
例 5.4 统计不同工种工资情况，包括工资的最小值、最大值、工资>\$30 000 的人数占每个组人数的百分比，以及工资的均值，并总结每个分组中工资最大值与最小值的差值 .....	119
例 5.5 对休闲数据的调研分析，定义复选变量集 .....	120
例 5.6 对数据文件分类，按照不同描述变量汇总，生成报表 .....	126

## 第 6 章 描述性统计分析 ..... 130

例 6.1 编制学生身高的简易频数表 .....	134
例 6.2 学生身高的统计描述 .....	137
例 6.3 学生身高的探索性分析 .....	139
例 6.4 吸烟习惯与患病率的关系 .....	148
例 6.5 各地区城乡居民消费水平比较 .....	152

## 第 7 章 均值比较与 t 检验 ..... 154

例 7.1 利用均值过程比较身高是否受地区和性别的影响 .....	156
例 7.2 比较样本均值和总体均值的 t 检验 .....	160
例 7.3 药品的独立两样本 t 检验 .....	163
例 7.4 通过配对样本 t 检验查看病人与正常人的血液中胆碱酯酶活性是否有明显差别 .....	167

## 第8章 方差分析..... 169

- 例 8.1 为了寻求高产油菜品种的方差分析..... 169
- 例 8.2 为考察合成纤维中收缩率与总拉伸倍数对纤维弹性影响性..... 186
- 例 8.3 研究杨树一年生长量与施用氮肥和钾肥的关系..... 189

## 第9章 相关分析..... 195

- 例 9.1 抽样的文盲率和各省人均 GDP 的两变量相关分析..... 199
- 例 9.2 分析温度与河水流量之间的相关关系..... 204
- 例 9.3 分析各城市日照数是否近似..... 210

## 第10章 回归分析..... 212

- 例 10.1 通过线形回归拟合气压与沸点的关系..... 222
- 例 10.2 解释主管性格与雇员对其整体满意度之间的关系..... 225
- 例 10.3 拟合各阵泥石流泥面宽与泥深之间的关系..... 232
- 例 10.4 诊断发现运营不良的金融商业机构..... 239
- 例 10.5 棉花单株在不同时期的成铃数 (Y) 与初花后天数 (X) 存在非线性的关系 .. 248

## 第11章 聚类分析与判别分析..... 252

- 例 11.1 主要城市空气质量指标的 K 均值聚类分析..... 256
- 例 11.2 各地区交通事故情况系统聚类分析..... 261
- 例 11.3 主要城市日照时数的变量聚类过程..... 264
- 例 11.4 University of Florida graduate salarye 的两步聚类法..... 271
- 例 11.5 1995 年人类发展报告部分数据判别样品分析..... 280

## 第12章 因子分析与对应分析..... 285

- 例 12.1 各地区城市市政设施的变量的因子分析..... 291
- 例 12.2 主要城市日照时数的因子旋转结果..... 296
- 例 12.3 利用简单因子分析来确定不同年龄段选民的倾向..... 304

## 第13章 非参数检验..... 309

- 例 13.1 调查某美发店上半年各月顾客数量, 检验各月顾客数是否服从均匀分布..... 313
- 例 13.2 分析蚜虫在农作物根部分布是否为泊松分布..... 315
- 例 13.3 学生体重的分布是否为正态分布..... 317