

刘江涛 刘立佳 编著

SPSS

数据统计 与分析应用教程



◆ 从基础角度出发

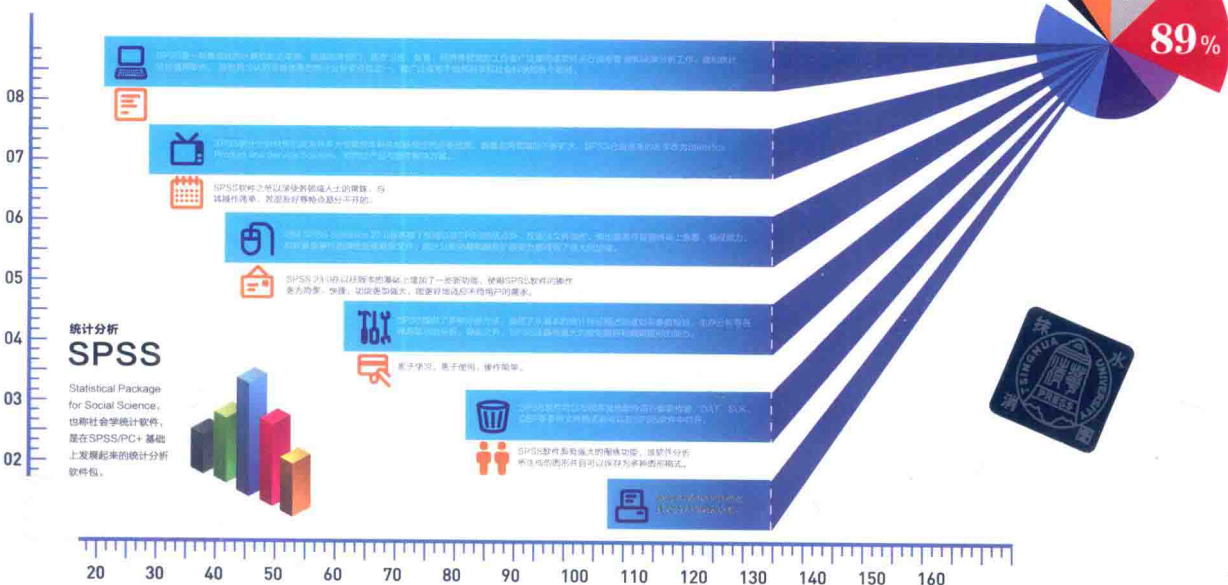
通过**34**个课堂练习全面讲解SPSS基本知识以及常见的统计分析功能

◆ 从实用角度出发

通过4个专业统计分析案例详细介绍SPSS在相关行业领域的具体应用

◆ 从读者角度出发

配备资源包括**50**个统计分析的数据文件以及**130**分钟的语音视频教学文件



SPSS 数据统计与分析应用教程：基础篇

刘江涛 刘立佳 编著

清华大学出版社

北京

内 容 简 介

本书从实用角度出发,以实例的形式详细介绍了 SPSS 常用的统计分析方法在相关行业领域的具体应用。书中首先从 SPSS 软件开始介绍,详细讲解了 SPSS 软件的基本知识和统计分析前的准备,描述统计分析,均值比较和 T 检验,SPSS 的方差分析,非参数检验,SPSS 的相关分析,SPSS 的回归分析,SPSS 的多元统计分析,SPSS 在时间序列预测中的应用,SPSS 在问卷缺失值、信度处理与多重响应分析中的应用,统计图形,酸奶饮料新产品口味测试研究案例,某汽车企业汽车年销量的预测,中国消费者信心指数影响因素分析,偏态分布的激素水平影响因素分析等内容。

配备资源包括书中实例用到的数据文件,以及讲解实例统计分析过程的语音视频教学文件,可帮助读者快速掌握书中介绍的内容。

本书内容编排科学、实例丰富、可操作性强,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,是相关专业本科生、研究生、专业统计分析人士及管理人员和决策者等理想的参考书,也可作为大中专院校和培训机构相关专业的教材。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

SPSS 数据统计与分析应用教程:基础篇/刘江涛,刘立佳编著. —北京:清华大学出版社,2017
ISBN 978-7-302-45092-4

I. ①S… II. ①刘… ②刘… III. ①统计分析—软件包—教材 IV. ①C819

中国版本图书馆 CIP 数据核字(2016)第 225933 号

责任编辑:陈冬梅

装帧设计:杨玉兰

责任校对:吴春华

责任印制:何 芊

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62791865

印 刷 者:北京鑫丰华彩印有限公司

装 订 者:三河市漂源装订厂

经 销:全国新华书店

开 本:190mm×260mm 印 张:24 字 数:581 千字

版 次:2017 年 1 月第 1 版 印 次:2017 年 1 月第 1 次印刷

印 数:1~3000

定 价:55.00 元

产品编号:066735-01

前 言

SPSS 是一种集成化的计算机处理和统计分析通用软件，是世界公认的统计分析软件包之一，非常适合进行数据分析。在经济学、医学、教育学、管理学、心理学、广告学、统计学及商业、工业、林业、农业等各个领域都有广泛的应用。

本书从实用角度出发，以实例的形式详细介绍了 SPSS 常用的统计分析方法在相关行业领域的具体应用。书中通过大量的实例，将作者的经验和心得与软件的操作技巧紧密结合。在内容安排上也尽量做到科学合理，把篇幅放在重点内容的讲解上，让读者花最少的时间掌握最关键的技术，可大幅度提高学习效率。

本书内容

本书首先从 SPSS 软件的基本知识和基本操作开始介绍，详细讲解了 SPSS 软件的特点、环境要求及各种常用设置。然后介绍 SPSS 统计分析前的准备工作，包括数据文件的建立、数据文件的属性、数据文件的整理等内容，为后面介绍具体应用打好基础。从第 3 章开始依次介绍了描述统计分析，均值比较和 T 检验，SPSS 的方差分析，非参数检验，SPSS 的相关分析，SPSS 的回归分析，SPSS 的多元统计分析，SPSS 在时间序列预测中的应用，SPSS 在问卷缺失值、信度处理与多重响应分析中的应用，统计图形等内容。最后 4 章，通过酸奶饮料新产品口味测试研究案例、某汽车企业汽车年销量的预测、中国消费者信心指数影响因素分析、偏态分布的激素水平影响因素分析等综合案例，详细深入地讲解 SPSS 在不同行业中的具体应用。

本书特点

本书以“实用、够用、精用”为原则，以熟练掌握 SPSS 应用为目的，以“学中练、练中学”的形式来组织内容，其主要特点如下。

编排科学、易学易用：书中由易到难安排内容，适合初学者循序渐进地学习。书中对每种方法都是先介绍概念、原理、操作方法，然后给出实例讲解具体应用，这样安排内容符合学习规律，方便读者掌握所学知识。

内容全面、讲解细致：书中对所有常见的统计分析方法都图文并茂地进行了深入详细的介绍，读者只要按照书中介绍的步骤进行操作，就可以轻松掌握所学内容。

案例丰富、技术实用：书中对每种分析方法都安排了实例讲解其在行业中的具体应用，在讲解过程中穿插了作者的经验和心得，可帮助读者举一反三，深入掌握所学内容。

视频教学，学习高效：配备书中实例的语音视频教学文件，不仅可帮助读者解决学习中遇到的问题，还可引导读者快速掌握所学内容，并进行技术拓展。

资源配备

本书配备的具体内容如下。

“数据文件”目录：提供了书中实例所用到的数据文件。

“视频文件”目录：书中讲解实例统计分析的语音视频教学文件。

本书配备资源请到清华大学出版社官方网站下载，网址 <http://www.tup.com.cn>。如在下载过程中遇到问题，可联系我们，联系电话：010-62792097-201，邮箱：2864186118@qq.com。

读者对象

本书内容编排科学，实例丰富，可操作性强，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，是相关专业本科生、研究生、专业统计分析人士及管理人员和决策者等理想的参考书，也可作为大中专院校和培训机构相关专业的教材。

本书由哈尔滨理工大学的刘江涛和黑龙江职业学院的刘立佳编著，其中刘江涛编写了第 8~16 章，刘立佳编写了第 1~7 章。其他参加编写的人员还有于香芝、杨旺功、江俊浩、王劲、田万勇、赵一飞、韩成斌、周艳山、田君、张博、吴艳臣、徐昱、王永忠、李明玉、武可元、于秀青等。由于作者水平有限，书中难免存在疏漏和错误之处，敬请读者批评指正。

编者

目 录

第 1 章 SPSS 23.0 概述	1	2.2.2 类型	21
1.1 SPSS 23.0 的特点与新功能	1	2.2.3 宽度	22
1.1.1 SPSS 23.0 的特点	1	2.2.4 小数位	22
1.1.2 SPSS 23.0 的新增功能	2	2.2.5 标签	22
1.2 SPSS 23.0 的环境要求	3	2.2.6 值	22
1.2.1 SPSS 23.0 对硬件的要求	3	2.2.7 缺失值	23
1.2.2 SPSS 23.0 对软件的要求	3	2.2.8 列	23
1.3 SPSS 23.0 的安装、卸载、启动和退出	4	2.2.9 对齐	23
1.3.1 SPSS 23.0 的安装与卸载	4	2.2.10 测量	23
1.3.2 SPSS 23.0 的启动与退出	7	2.2.11 角色	23
1.4 SPSS 23.0 相关设置	7	2.2.12 课堂练习：员工满意度调查表的数据属性设计	23
1.4.1 常规功能设置	7	2.3 SPSS 数据文件的整理	24
1.4.2 查看器功能设置	8	2.3.1 观测量排序：各地区粮食产量	24
1.4.3 设置有关数据的参数	9	2.3.2 数据的转置：5 个地区经济指标	26
1.4.4 设置自定义数值型变量的格式	10	2.4 思考与练习	27
1.4.5 设置输出的参数	11	第 3 章 描述统计分析	29
1.4.6 设置图表的参数	12	3.1 频数分析	29
1.4.7 设置输出表格的参数	13	3.1.1 频数分析概述	29
1.4.8 设置文件位置的参数	14	3.1.2 频数分析的 SPSS 操作详解	30
1.4.9 设置脚本的参数	15	3.1.3 课堂练习：分析学生身高分布特征	34
1.4.10 设置多重归因窗口的参数	15	3.2 描述统计分析	37
1.4.11 设置语法编辑器窗口的参数	16	3.2.1 描述统计分析概述	37
第 2 章 SPSS 统计分析前的准备	17	3.2.2 描述统计分析的 SPSS 操作详解	39
2.1 SPSS 数据文件的建立	17	3.2.3 课堂练习：大学新生的心理健康状况	40
2.1.1 新建数据文件	17	3.3 探索性分析	42
2.1.2 直接打开已有数据文件	18	3.3.1 探索性分析概述	42
2.1.3 使用数据库导入数据	18		
2.1.4 从文本向导导入数据	18		
2.1.5 课堂练习：股票指数的导入	19		
2.2 SPSS 数据文件的属性	20		
2.2.1 变量名：Name 栏	20		

3.3.2 探索性分析的 SPSS 操作详解	42	第 5 章 方差分析	74
3.3.3 课堂练习：大学新生的心理 健康状况	44	5.1 概述	74
3.4 列联表分析	49	5.1.1 方差分析的概念	74
3.4.1 列联表分析概述	49	5.1.2 方差分析的基本思想	74
3.4.2 列联表分析的 SPSS 操作详解	50	5.1.3 方差分析的基本假设	75
3.4.3 课堂练习：大学生出生年代与 感恩心理	51	5.2 单因素方差分析	75
3.5 比率分析	53	5.2.1 单因素方差分析的基本原理	75
3.5.1 比率分析概述	54	5.2.2 单因素方差分析的 SPSS 操作详解	76
3.5.2 比率分析的 SPSS 操作详解	54	5.2.3 课堂练习：化肥种类对粮食 产量的影响	79
3.5.3 课堂练习：年度销售净利润率 差异	54	5.3 多因素方差分析	82
3.6 思考与练习	56	5.3.1 多因素方差分析的基本原理	83
第 4 章 均值比较和 T 检验	58	5.3.2 多因素方差分析的 SPSS 操作详解	83
4.1 均值过程	58	5.3.3 课堂练习：动物饲料对小鼠 体重增加的影响	87
4.1.1 均值过程简介	58	5.4 协方差分析	91
4.1.2 均值过程的 SPSS 操作详解	58	5.4.1 协方差分析的基本原理	91
4.1.3 课堂练习：比较考试成绩	59	5.4.2 协方差分析的 SPSS 操作详解	91
4.2 单样本 T 检验	61	5.4.3 课堂练习：降压药治疗效果 分析	92
4.2.1 单样本 T 检验的基本原理	61	5.5 重复测量方差分析	94
4.2.2 单样本 T 检验的 SPSS 操作 详解	62	5.5.1 重复测量方差分析的原理	94
4.2.3 课堂练习：溶剂浓度测量方法 准确度分析	63	5.5.2 重复测量方差分析的 SPSS 操作详解	95
4.3 独立样本 T 检验	64	5.5.3 课堂练习：锻炼减肥效果 分析	96
4.3.1 独立样本 T 检验的基本原理	65	5.6 思考与练习	99
4.3.2 独立样本 T 检验的 SPSS 操作 步骤	65	第 6 章 非参数检验	101
4.3.3 课堂练习：教学质量评价	66	6.1 卡方检验	101
4.4 两配对样本 T 检验	68	6.1.1 卡方检验的原理	101
4.4.1 配对样本 T 检验的基本原理	68	6.1.2 卡方检验的 SPSS 操作 详解	102
4.4.2 两配对样本 T 检验的 SPSS 操作详解	69	6.1.3 课堂练习：小白鼠试验 数据验证	104
4.4.3 课堂练习：贫血儿童血红蛋白 平均水平	70	6.2 二项检验	107
4.5 思考与练习	71		



6.2.1	二项检验的原理	107	7.2.3	课堂练习：平均温度与日照 时数的相关性	130
6.2.2	二项检验的 SPSS 操作详解	107	7.3	偏相关分析	132
6.2.3	课堂练习：研究糖尿病患病率 高低	108	7.3.1	偏相关分析的基本原理	132
6.3	游程检验	110	7.3.2	偏相关分析的 SPSS 操作 详解	133
6.3.1	游程检验的原理	111	7.3.3	课堂练习：用偏相分析研究 学生学习成绩相关关系	134
6.3.2	游程检验的 SPSS 操作详解	111	7.4	距离分析	136
6.3.3	课堂练习：运动员状态稳定性 判断	112	7.4.1	距离分析的基本原理	136
6.4	单样本 K-S 检验	113	7.4.2	距离分析的 SPSS 操作详解	137
6.4.1	单样本 K-S 检验的原理	114	7.4.3	课堂练习：3 个地区月平均 气温的相似程度	139
6.4.2	单样本 K-S 检验的 SPSS 操作 详解	114	7.5	思考与练习	141
6.4.3	课堂练习：考试成绩是否服 从正态分布	115	第 8 章	回归分析	143
6.5	两独立样本的非参数检验	117	8.1	一元线性回归分析	143
6.5.1	两独立样本非参数检验的 原理	117	8.1.1	简单线性回归的基本原理	143
6.5.2	两独立样本非参数检验的 SPSS 操作详解	118	8.1.2	一元线性回归的 SPSS 操作 详解	144
6.5.3	课堂练习：比较病人检查结果差别	119	8.1.3	课堂练习：失业率与通货 膨胀率关系	147
6.6	多独立样本的非参数检验	121	8.2	多元线性回归分析	149
6.6.1	多独立样本非参数检验的 原理	121	8.2.1	多元线性回归的基本原理	149
6.6.2	多独立样本非参数检验的 SPSS 操作详解	121	8.2.2	多元线性回归的 SPSS 操作 详解	149
6.6.3	课堂练习：农药杀虫效果 比对	122	8.2.3	课堂练习：美国是否存在 规模经济	149
6.7	思考与练习	124	8.3	曲线回归	152
第 7 章	相关分析	126	8.3.1	曲线回归的基本原理	152
7.1	相关分析概述	126	8.3.2	曲线回归的 SPSS 操作详解	153
7.1.1	相关的基本概念	126	8.3.3	课堂练习：广告支出对 销售量产生的影响	154
7.1.2	相关分析	127	8.4	非线性回归分析	156
7.2	简单相关分析	127	8.4.1	非线性回归分析的基本原理	157
7.2.1	简单相关分析的基本原理	127	8.4.2	非线性回归分析的 SPSS 操作详解	157
7.2.2	简单相关分析的 SPSS 操作 详解	129	8.4.3	课堂练习：培训效果分析	159

8.5 思考与练习	161	10.2.5 课堂练习：某城市月度平均 气温的季节性分解分析	202
第 9 章 SPSS 的多元统计分析	163	10.3 时间序列的随机性分析	204
9.1 因子分析	163	10.3.1 随机性分析的原理	204
9.1.1 因子分析的基本原理	163	10.3.2 ARIMA 模型的 SPSS 操作详解	205
9.1.2 因子分析的 SPSS 操作详解	165	10.3.3 课堂练习：工业生产值和国库券 利率与基金利率差额 ARIMA 模型分析	206
9.1.3 课堂练习：国民经济主要指标 统计	168	10.4 思考与练习	209
9.2 聚类分析	173	第 11 章 问卷缺失值、信度处理与 多重响应分析	211
9.2.1 聚类分析的基本原理	173	11.1 调查问卷缺失值处理方法	211
9.2.2 快速聚类法的 SPSS 操作 详解	174	11.1.1 缺失值的类型与处理方法	211
9.2.3 课堂练习：全国环境污染 程度分析	176	11.1.2 替换缺失值的 SPSS 操作 详解	212
9.3 判别分析	179	11.1.3 缺失值分析的 SPSS 操作 详解	213
9.3.1 判别分析的基本原理	179	11.1.4 实例图文分析：电信公司客户 数据缺失值的分析	214
9.3.2 判别分析的 SPSS 操作详解	182	11.2 调查问卷的信度分析	220
9.3.3 课堂练习：判别分析过去和 潜在客户的财务以及人口统计 信息	185	11.2.1 信度分析概述	220
9.4 思考与练习	188	11.2.2 信度分析的 SPSS 操作详解	221
第 10 章 时间序列预测	190	11.2.3 实例图文分析：调查问卷的 信度	222
10.1 时间序列的预处理	190	11.3 调查问卷的多重响应分析	224
10.1.1 预处理的基本原理	190	11.3.1 多重响应分析概述	225
10.1.2 时间序列预处理的 操作详解	191	11.3.2 多重响应分析的 SPSS 操作 详解	225
10.1.3 课堂练习：某国国库券利率与 基金利率差额数据的 预处理	192	11.3.3 实例图文分析：消费者使用的 手机品牌调查	226
10.2 时间序列的确定性分析	194	11.4 思考与练习	229
10.2.1 确定性分析的基本原理	195	第 12 章 统计图形	231
10.2.2 指数平滑法的 SPSS 操作 详解	195	12.1 SPSS 图形的基本功能	231
10.2.3 课堂练习：进出口贸易总额的 指数平滑建模	199	12.1.1 图形生成器	231
10.2.4 季节分解的 SPSS 操作 详解	202	12.1.2 用传统模式创建图形	231
		12.2 条形图	232



12.2.1	条形图的类型和 SPSS 操作 详解	232	12.10.2	直方图实例	277
12.2.2	简单条形图 SPSS 操作详解	233	12.11	思考与练习	278
12.2.3	集群条形图	243	第 13 章 酸奶饮料新产品口味测试 研究案例		
12.2.4	堆积条形图实例	248	13.1	案例背景	279
12.3	折线图	252	13.1.1	研究项目概况	279
12.3.1	折线图的类型和 SPSS 操作	253	13.1.2	分析思路与商业理解	279
12.3.2	简单线图实例	253	13.2	数据理解	280
12.3.3	多线图实例	255	13.2.1	数据与设计框架一致性 检查	280
12.3.4	垂直线图实例	256	13.2.2	平均值的列表描述	281
12.4	面积图	257	13.2.3	平均值的图形描述	282
12.4.1	面积图的类型和 SPSS 操作	257	13.3	不同品牌的评分差异分析	283
12.4.2	简单箱图实例	259	13.3.1	单因素方差分析模型简介	283
12.4.3	堆积面积图实例	260	13.3.2	品牌作用的总体检验	284
12.5	饼图	262	13.3.3	组间两两比较	285
12.5.1	饼图的类型和 SPSS 操作	262	13.3.4	方差齐性检验	287
12.5.2	饼图实例	263	13.4	两因素方差分析模型	288
12.6	高低图	264	13.4.1	两因素方差分析模型简介	288
12.6.1	高低图的类型和 SPSS 操作	264	13.4.2	拟合包括交互项的饱和 模型	289
12.6.2	简单高低图实例	265	13.4.3	拟合只包含主效应的模型	290
12.6.3	分类高低收盘图实例	266	13.4.4	组间两两比较	291
12.7	箱图	267	13.4.5	随机因素分析	293
12.7.1	箱图的类型和 SPSS 操作	268	13.5	分析结论与讨论	294
12.7.2	简单箱图实例	268	13.5.1	分析结论	294
12.7.3	复合箱图实例	269	13.5.2	Benchmark: 用还是不用	295
12.8	误差条形图	270	第 14 章 某汽车企业汽车年销量的 预测		
12.8.1	误差条形图的类型和 SPSS 操作	271	14.1	案例背景	296
12.8.2	简单误差条形图实例	271	14.1.1	研究项目概况	296
12.8.3	集群误差条形图实例	272	14.1.2	分析思路和商业理解	296
12.9	散点图	273	14.2	数据理解	297
12.9.1	散点图的作图步骤	273	14.3	变量变换后的线性回归	298
12.9.2	简单分布实例	274	14.3.1	线性回归模型简介	298
12.9.3	重叠分布实例	275	14.3.2	变量变换后拟合线性 回归模型	300
12.9.4	3D 分布	275			
12.10	直方图	276			
12.10.1	直方图的类型和 SPSS 操作	276			



- 14.3.3 模型拟合效果的判断 302
- 14.3.4 存储预测值和区间估计值 304
- 14.4 曲线拟合 305
 - 14.4.1 用曲线估计过程同时拟合
多个曲线模型 305
 - 14.4.2 模型拟合效果的判断 307
 - 14.4.3 模型的预测 309
- 14.5 利用非线性回归进行拟合 309
 - 14.5.1 模型简介 310
 - 14.5.2 构建分段回归模型 310
 - 14.5.3 不同模型效果的比较 313
- 14.6 分析结论 313
- 第 15 章 中国消费者信心指数
影响因素分析 315**
 - 15.1 案例背景 315
 - 15.1.1 项目背景 315
 - 15.1.2 项目问卷 315
 - 15.1.3 分析思路和商业理解 317
 - 15.2 数据理解 318
 - 15.2.1 考察时间、地域对信心指数的
影响 318
 - 15.2.2 考察性别、职业、婚姻状况等
对信心指数的影响 320
 - 15.2.3 考察年龄对信心指数的
影响 322
 - 15.3 标准框架下的建模分析 323
 - 15.3.1 建立总模型 323
 - 15.3.2 两两比较的结果 325
 - 15.4 多元方差分析模型的结果 326
 - 15.4.1 模型简介 327

- 15.4.2 拟合多元方差分析模型 327
- 15.5 最优尺度回归 334
 - 15.5.1 方法简介 334
 - 15.5.2 利用最优尺度回归进行
分析 335
- 15.6 多水平模型框架下的建模分析 338
 - 15.6.1 模型简介 338
 - 15.6.2 针对时间拟合多水平模型 340
- 15.7 项目总结与讨论 344
 - 15.7.1 分析结论 344
 - 15.7.2 运用复杂模型建模的情况 345
- 第 16 章 偏态分布的激素水平
影响因素分析 346**
 - 16.1 案例背景 346
 - 16.1.1 研究项目概况 346
 - 16.1.2 分析思路与商业理解 346
 - 16.2 数据理解 347
 - 16.2.1 单变量描述 347
 - 16.2.2 变量关联探索 349
 - 16.3 对因变量变换后的建模分析 353
 - 16.3.1 常见的变量变换方法 353
 - 16.3.2 本案例的具体操作 354
 - 16.4 秩变换分析 356
 - 16.5 利用 Cox 模型进行分析 357
 - 16.5.1 Cox 回归模型的基本原理 357
 - 16.5.2 本案例的具体操作 358
 - 16.6 项目总结与讨论 360
 - 16.6.1 分析结论 360
 - 16.6.2 如何正确选择分析模型 360
- 思考与练习答案 362**

第 1 章 SPSS 23.0 概述

SPSS(Statistical Package for Social Science, 社会学统计软件)是在 SPSS/PC+基础上发展起来的统计分析软件包。SPSS 是一种集成化的计算机处理和统计分析通用软件,是世界公认最优秀的统计分析软件包之一,被广泛应用于自然科学和社会科学的各个领域。近年来,我国政府部门、医疗卫生、体育、经济等领域的工作者广泛使用该软件进行信息管理和决策分析工作。同时,掌握 SPSS 统计分析软件的应用已成为许多大专院校本科生和研究生的必备技能。随着应用领域的不断扩大,SPSS 已由原来的名字改为 Statistics Product and Service Solution,即统计产品与服务解决方案。

SPSS 统计软件自 1968 年开发至今,已经经历了多次改版,并于 20 世纪 90 年代以后推出了以交互式对话为主要特征的第 7 版。第 7 版以后的版本称为 SPSS for Windows 版,最新的版本是 2015 年推出的第 23 版。

IBM SPSS Statistics 23.0 版本除了保持以往 SPSS 的优点外,在语法文件加密、输出图表在智能终端上查看、编程能力、应对紧急事件的弹性处理数据文件、统计分析功能和编程扩展能力方面都得到了很大加强。另外,中文字符的兼容性问题得到了彻底解决。本书主要以 IBM SPSS 23.0(以下简称为 SPSS 23.0)版本为例,讲解 SPSS 统计软件的功能及使用方法。

1.1 SPSS 23.0 的特点与新功能

SPSS 软件之所以深受各领域人士的青睐,与其操作简单、界面友好等特点是分不开的。而 SPSS 23.0 在以往版本的基础上增加了一些新功能,使得 SPSS 软件的操作更为简便、快捷,功能更加强大,能更好地适应不同用户的需求。

1.1.1 SPSS 23.0 的特点

SPSS 软件风靡世界,与 SAS、SYSTAT 一起成为世界上公认的三大数据分析软件,而 SPSS 却为各个领域的广大科研工作者及其他用户所钟爱,其原因在于它具有以下特性。

(1) 多种实用分析方法。SPSS 提供了多种分析方法,涵盖了从基本的统计特征描述到诸如非参数检验、生存分析等各种高层次的分析。此外,SPSS 还具有强大的绘制图形和编辑图形的能力。

(2) 易于学习,易于使用,操作简单。对 SPSS 23.0 而言,除了数据输入工作要使用键盘外,其他大部分操作均可以使用菜单、对话框来完成。同时 SPSS 还保留了命令行方式的优点,采用菜单式操作与“语法”程序运行的完美结合,使熟悉 SPSS 语言的用户可以直接在语句窗口中输入 SPSS 命令,提交系统执行。还可以通过单击对话框中的【粘贴】按钮,自动生成“语言”程序代码,提交系统运行就可实现指定功能,并可以文件形式保存,从而减少了用户的工作量。这样用户不必记忆大量的命令,使操作更简单,也使 SPSS 软件变得更加易学易用。

(3) 兼容多种数据文件格式,具有强大的图表功能。SPSS 软件可以与很多其他软件进行

数据传输, DAT、SLK、DBF 等多种文件格式都可以在 SPSS 软件中打开。SPSS 软件还具有强大的图表功能, 该软件分析所生成的图形文件可以保存为多种格式。

(4) 可以根据用户的需要选择所需的模块。用户可以根据机器的配置情况, 自由选择模块来安装。

(5) 内置 SaxBasic 语言。SPSS 软件内置了 SaxBasic 语言, 该语言与“语法”命令混合编程, 可以提高效率, 便于高级用户使用。

1.1.2 SPSS 23.0 的新增功能

2015 年推出的第 23 版中, 在保留以往版本优良特性的基础上又增加了一些新的功能模块, 使得功能更加强大, 操作上更突出个性化, 更好地适应了不同用户的数据分析需求。新增功能具体介绍如下。

1. 地理空间关联规则

通过使用地理空间关联规则, 可以根据空间属性和非空间属性在数据中查找模式。例如, 可以通过位置属性和人口统计信息属性识别罪案数据中的模式。根据这些模式, 可以构建规则, 以预测有可能发生特定类型罪案的地点。

此过程可以在“基本统计信息”选项中获得。

2. 空间时间预测

空间时间预测使用包含位置数据、预测输入字段(预测变量)、时间字段和目标字段的数据。每个位置在数据中都有许多行, 这些行表示每个预测变量在每个位置、每个时间间隔的值。

此过程可以在“基本统计信息”选项中获得。

3. 时间因果模型

时间因果建模尝试发现时间序列数据中的关键因果关系。在时间因果建模中, 指定一组目标序列以及这些目标的候选输入集。这样, 过程将为每个目标构建一个自回归时间序列模型, 并且仅包括那些与目标具有因果关系的输入。此方法不同于传统时间序列建模, 在传统时间序列建模中, 必须为目标序列显式指定预测变量。由于时间因果建模通常涉及为多个相关的时间序列构建模型, 因此结果称为模型系统。

在“预测”选项中提供了时间因果建模。

4. 批量装入数据库中

将数据导出至数据库时, 批量装入会将数据成批提交到数据库, 而不是一次提交一条记录。此操作可以使数据处理速度更快, 对于大型数据文件尤其如此。

5. 可编程性增强功能

现在, 用户可以从任何外部 R 进程运行使用 R Integration Package for IBM® SPSS® Statistics 中函数的 R 程序, 如 R IDE 或 R 解释器; 还可以从 R 运行 SPSS Statistics 命令语法。通过 Python 或 R 实现的扩展命令现在支持在变量列表中使用 TO 和 ALL 关键字。

IBM SPSS Statistics - Essentials for R 和 IBM SPSS Statistics - Essentials for Python 现在包含



更多扩展命令以及关联的定制对话框。另外，可以通过在语法编辑器中按 F1 键来访问随 Essentials for R 和 Essentials for Python 一起安装的所有扩展命令的帮助。

1.2 SPSS 23.0 的环境要求

SPSS 23.0 采用较为成熟的技术，对运行环境要求不高，对用户硬件配置要求低，用户无须额外升级现有的计算机软硬件，也无须再购买配套的数据库软件等，简单安装便可投入使用。

1.2.1 SPSS 23.0 对硬件的要求

SPSS 23.0 对计算机硬件的要求并不高，但由于 SPSS 的主要用途是面向大型数据库，其运算涉及的数据量比较多。因此，一般需要有较高的内存，而且，如果用户还要进行多因素分析等大运算量分析，计算机至少要有 512MB 的内存。

SPSS 23.0 对计算机硬件的基本要求如下。

- (1) Pentium 以上系列的处理器。
- (2) 至少 512MB 的内存。对于巨量数据的管理和复杂的统计分析，1GB 的内存能够保证较为理想的数据运行速度。
- (3) 至少 1GB 的硬盘剩余空间。如果想要将 SPSS 23.0 的各模块(包括 SPSS Smart 查看器、SPSS Data Access Pack 等相关附件)全部安装，至少需要约 800MB 的硬盘剩余空间，只安装 SPSS 模块需要约 473MB 的硬盘剩余空间。
- (4) CD-ROM 光盘驱动器，这是用光盘安装 SPSS 的基本要求。从网上安装 SPSS 软件则不需要光驱。
- (5) S-VAD 显示器和 Windows 2000/XP/Vista/7 兼容的图形适配卡。
- (6) 支持 TCP/IP 网络协议的网络适配卡，用于访问 SPSS 公司的服务器，以获得服务和软件的升级。

1.2.2 SPSS 23.0 对软件的要求

SPSS 23.0 可以在中文操作系统下运行。SPSS 23.0 在含有中文的数据文件和图表等兼容性方面有了很大的改善，不会再出现在早期版本中经常遇到的乱码等问题。

建议安装以下软件。

- (1) Windows XP/7 操作系统。如果需要支持 SPSS 软件的中文输入和输出，应安装中文操作系统。
- (2) Internet Explorer 6 或以上版本的浏览器。
- (3) Adobe Reader。为阅读 PDF 格式的帮助文件和 SPSS 分析软件的相关文档，应安装 PDF 阅读器。安装光盘中已提供该软件，用户可根据自身需要选择安装。
- (4) SPSS Data Access Pack。此软件提供不同的数据类型和不同数据库共享的解决方案，如果需要 Access、Brieve、DB2、dBase、Excel 等常用数据文件，可选择安装此软件，安装光盘中已提供此软件。

1.3 SPSS 23.0 的安装、卸载、启动和退出

SPSS 23.0 的安装、卸载、启动与退出和一般的 Windows 应用软件基本一样，非常简便，本节将分别进行详细介绍。

1.3.1 SPSS 23.0 的安装与卸载

1. SPSS 23.0 的安装

SPSS 23.0 如同其他 Windows 应用软件的安装一样，非常容易。下面简要介绍如何安装 SPSS 23.0。

step 01 启动计算机，在 Windows 桌面上双击安装包开始安装。

step 02 进入【正在准备安装】界面，系统正在配置 Windows Installer，以向导方式完成软件安装，如图 1.1 所示。

step 03 出现 IBM 公司的 SPSS 软件简介，单击【下一步】按钮，如图 1.2 所示。

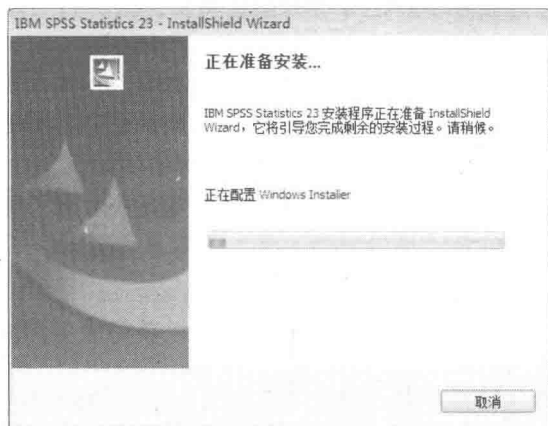


图 1.1

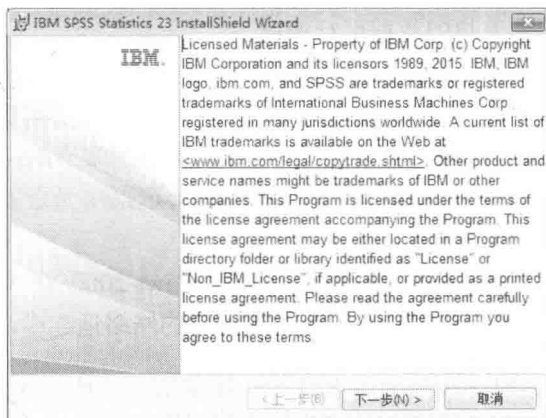


图 1.2

step 04 进入欢迎向导界面，选中【单个用户许可证】单选按钮并单击【下一步】按钮，如图 1.3 所示。

step 05 在【软件许可协议】界面中，选中【我接受许可协议中的全部条款】单选按钮，然后单击【下一步】按钮，如图 1.4 所示。

step 06 在【客户信息】界面中输入用户姓名及单位，然后单击【下一步】按钮，如图 1.5 所示。

step 07 在【帮助语言】界面中选择【简体中文】选项，然后单击【下一步】按钮，如图 1.6 所示。

step 08 在【辅助技术】界面中选择是否安装 JAWS for Windows 屏幕阅读软件，然后单击【下一步】按钮，如图 1.7 所示。

step 09 在进入的界面中选择是否安装 IBM SPSS Statistics-Essentials for Python，然后单击



【下一步】按钮，如图 1.8 所示。

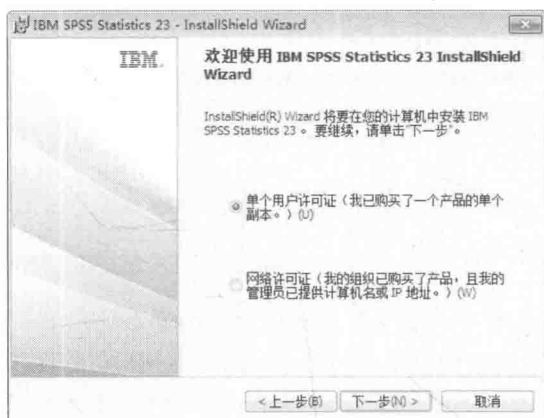


图 1.3

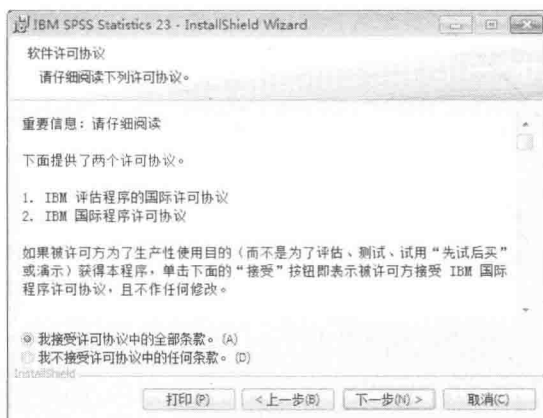


图 1.4



图 1.5

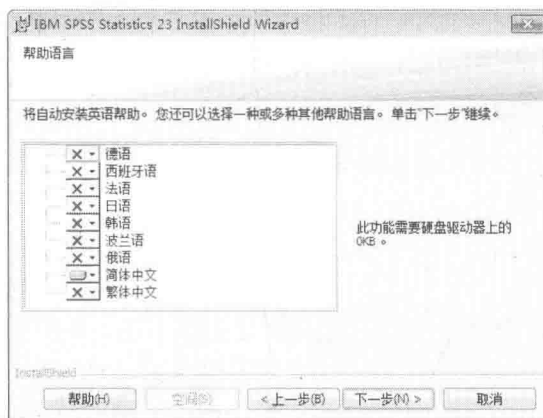


图 1.6

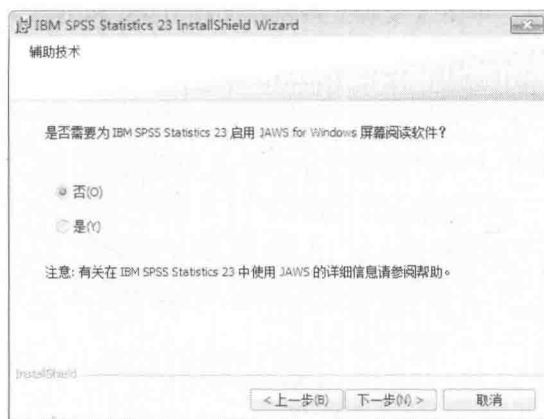


图 1.7



图 1.8

step 10 在【软件许可协议】界面中，选中【我接受许可协议中的全部条款】单选按钮，然后单击【下一步】按钮，如图 1.9 所示。

step 11 在【目的地文件夹】界面中更改安装路径，如图 1.10~图 1.12 所示。

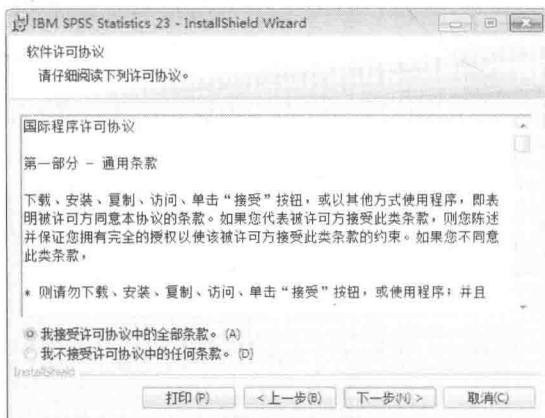


图 1.9



图 1.10



图 1.11



图 1.12

step 12 在【已做好安装程序的准备】界面中，单击【安装】按钮，如图 1.13 和图 1.14 所示。



图 1.13

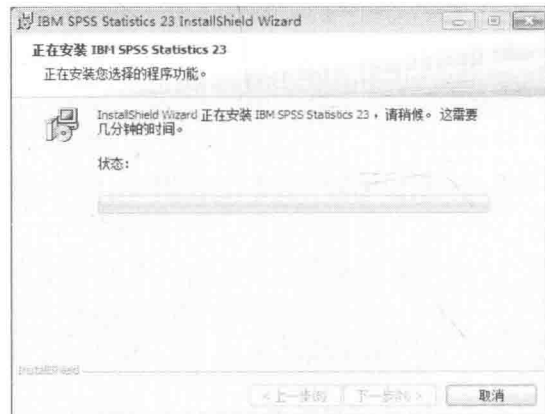


图 1.14