

21 世纪统计学系列教材

统计分析 与SPSS的应用

(第二版)

薛薇 编著

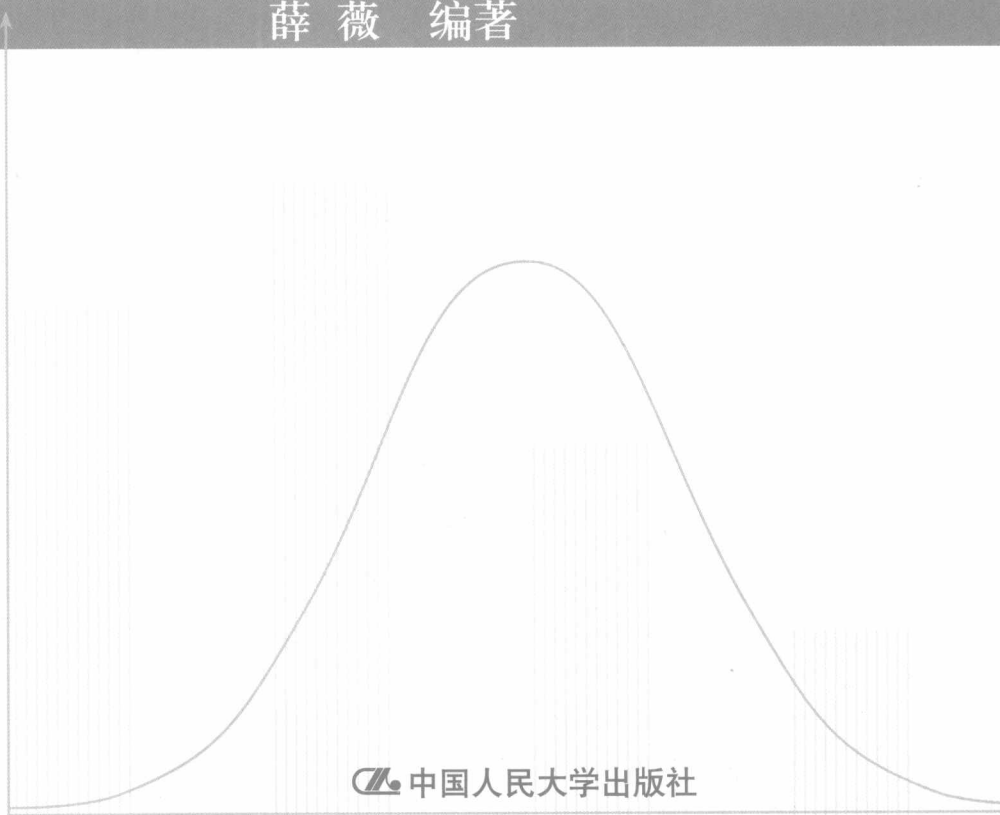
 中国人民大学出版社

21 世纪统计学系列教材

统计分析 与SPSS的应用

(第二版)

薛薇 编著



中国人民大学出版社

图书在版编目 (CIP) 数据

统计分析方法与 SPSS 的应用/薛薇编著. —2 版

北京: 中国人民大学出版社, 2007

(21 世纪统计学系列教材)

ISBN 978-7-300-08675-0

I. 统…

II. 薛…

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 164403 号

21 世纪统计学系列教材

统计分析方法与 SPSS 的应用 (第二版)

薛 薇 编著

出版发行 中国人民大学出版社

社 址 北京中关村大街 31 号

邮政编码 100080

电 话 010-62511242 (总编室)

010-62511398 (质管部)

010-82501766 (邮购部)

010-62514148 (门市部)

010-62515195 (发行公司)

010-62515275 (盗版举报)

网 址 <http://www.crup.com.cn>

<http://www.ttrnet.com>(人大教研网)

经 销 新华书店

印 刷 北京民族印刷厂

版 次 2001 年 11 月第 1 版

规 格 170 mm×228 mm 16 开本

2008 年 1 月第 2 版

印 张 24.5 插页 1

印 次 2008 年 1 月第 1 次印刷

字 数 446 000

定 价 35.00 元

版权所有 侵权必究 印装差错 负责调换

《21世纪统计学系列教材》编委会

编委会主任 易丹辉

编委会委员 (按姓氏笔画排序)

尹德光 冯士雍 张尧庭

陈希孺 吴喜之 赵彦云

柯惠新 袁 卫 倪加勋

顾 岚 袁寿庄 耿 直

总 序

改革开放以来，高等统计教育有了很大的发展。随着课程设置的不断调整，有不少教材出版，同时也翻译引进了一些国外优秀教材。作为培养我国统计专门人才的摇篮，中国人民大学统计学系自 1952 年创建以来，走过了风风雨雨，一直坚持着理论与应用相结合的办学方向，培养能够理论联系实际、解决实际问题的高层次人才。随着新知识经济和网络时代的到来，我们在教学科研的实践中，深切地感受到，无论是自然科学领域、社会科学领域的研究，还是国家宏观管理和企业生产经营管理，甚至在人们的日常生活中，信息需求量日益增多，信息处理技术更加复杂，作为信息技术支柱的统计方法，越来越广泛地应用于各个领域。

面对新的形势，我们一直在思索，课程设置、教材选择、教学方式等怎样才能使学生适应社会经济发展的客观需要。在反复酝酿、不断尝试的基础上，我们决定与统计学界的同仁，共同编写、出版一套面向 21 世纪的统计学系列教材。

这套系列教材聘请了中国科学院院士、中国科技大学陈希孺教授，上海财经大学数量经济研究院张尧庭教授，中国科学院数学与系统科学研究所冯士雍研究员等作为编委。他们长期任中国人民大学的兼职教授，一直关心、支持着统计学的学科建设和应用统计的发展。中国人民大学应用统计科学研究中心 2000 年已成为国家级研究基地，这些专家是首批专职或兼职研究人员。这一开放性研究基地的运作，将有利于提升我国应用统计科学研究的水平，也必将进一步促进高等统计教育的发展。

这套教材是我们奉献给新世纪的,希望它能够为促进应用统计教育水平的提高增添一份力量。这套教材力求体现以下特点:

第一,在教材选择上,主要面向经济类统计学专业。选材既包括统计教材也包括风险管理与精算方面的教材。尽管名为统计学系列教材,但并不求大、求全,而是力求精选。对于目前已有的内容较为成熟、适合教学需要、公认的较好的教材,并未列入本次出版计划。

第二,每部教材的内容和写作,注意广泛吸收国内外优秀教材的成果。教材力求简明易懂、内容系统和实用,注重对统计方法思想的阐述,并结合大量实际数据和实例说明统计方法的特点及应用条件。

第三,强调与计算机的结合。为着力提高学生运用统计方法分析解决问题的能力,教材所涉及的统计计算,要求运用目前已有的统计软件。根据教材内容,选择使用 SAS, SPSS, TSP, STATISTICA, EViews, MINITAB, Excel 等。

感谢中国人民大学出版社的同志们,他们怀着发展我国应用统计科学的热情和提高统计教育水平的愿望,经过反复论证,使这套教材得以出版。感谢参与教材编写的同行专家、统计学系的教师。愿大家的辛勤劳动能够结出丰硕的果实。我们期待着与统计学界的同仁,共同创造应用统计辉煌的明天。

易丹辉

第二版前言

《统计分析与 SPSS 应用》出版以来，受到了广大读者的青睐。对统计分析方法深入浅出的讲解风格，以实际数据为纽带说明 SPSS 的操作，以应用案例为背景阐述统计分析的思路，是本书的主要特点。

随着 SPSS 软件版本的不断升级，原书对软件操作部分的讲解内容已显得陈旧。在听取读者各方面意见的基础上，我们对原书进行了修订。修订后的《统计分析与 SPSS 应用》在延续原有风格的基础上，对以下方面进行了改进和调整：

第一，统计分析方法的讲解更加详尽。本书对统计分析方法的基本原理和基本思想进行了更加详尽的讲解，目的是希望读者能够领会统计分析方法的精髓和适用场合，以避免方法的滥用和误用。

第二，SPSS 操作的讲解与应用案例结合得更加紧密。本书更加注重 SPSS 操作步骤和应用的直观性，书中给出的 SPSS 窗口和对应用案例的分析操作是完全吻合的，目的是希望读者能够在理解应用案例问题的基础上，方便地对照所给的操作窗口，亲自利用 SPSS 实现相应的操作和分析。

第三，案例的选用更具实用性。本书选择的应用案例更加丰富，涉及面也有了进一步拓展，目的是希望启发读者操作案例，在各自不同的行业应用 SPSS 解决现实问题。

第四，提供了书中涉及案例的所有数据的电子文件，读者可到 www.rdjg.com.cn 上下载，从而方便地对照书中相关内容自行操作，得到相应的计算结果。

第五, 每章后面都增加了练习题, 部分练习题中涉及的数据, 读者可到 www.rdjg.com.cn 上下载。通过练习, 读者可进一步加深对相关内容的理解, 并熟练相关操作。

第六, 目前 SPSS 软件的最新版本为 15 版。但考虑到诸多因素, 本书以目前软件的可获得性最强, 可安装性最便利的 12 版为主要讲解对象。在操作方面, SPSS 最新版本与我们所选用的版本之间并无大的差异。

随着数据分析应用需求的不断加强, 掌握像 SPSS 这样流行的统计分析软件是十分必要的。真诚地希望读者能够通过阅读本书, 逐步领会统计分析方法的精髓, 掌握 SPSS 软件的操作, 并能够利用 SPSS 解决实际统计分析问题。

本书不妥和错误之处, 敬请各位读者指正。

薛薇

目 录

第 1 章 SPSS 统计分析软件概述	1
1.1 SPSS 使用基础	3
1.2 SPSS 的基本运行方式	9
1.3 利用 SPSS 进行数据分析的基本步骤	15
1.4 利用统计教练快速入门 SPSS	18
练习题	19
第 2 章 SPSS 数据文件的建立和管理	20
2.1 SPSS 数据文件	20
2.2 SPSS 数据的结构和定义方法	24
2.3 SPSS 数据的录入与编辑	33
2.4 SPSS 数据的保存	37
2.5 读取其他格式的数据文件	40
2.6 SPSS 数据文件合并	47
练习题	55
第 3 章 SPSS 数据的预处理	57
3.1 数据的排序	58
3.2 变量计算	61
3.3 数据选取	70
3.4 计数	74

3.5	分类汇总	76
3.6	数据分组	79
3.7	数据预处理的其他功能	83
	练习题	89
第 4 章	SPSS 基本统计分析	91
4.1	频数分析	92
4.2	计算基本描述统计量	101
4.3	交叉分组下的频数分析	107
4.4	多选项分析	119
4.5	比率分析	129
	练习题	133
第 5 章	SPSS 的参数检验	136
5.1	参数检验概述	136
5.2	单样本 t 检验	140
5.3	两独立样本 t 检验	145
5.4	两配对样本 t 检验	152
	练习题	156
第 6 章	SPSS 的方差分析	158
6.1	方差分析概述	158
6.2	单因素方差分析	160
6.3	多因素方差分析	179
6.4	协方差分析	194
	练习题	203
第 7 章	SPSS 的非参数检验	207
7.1	单样本的非参数检验	208
7.2	两独立样本的非参数检验	219
7.3	多独立样本的非参数检验	229
7.4	两配对样本的非参数检验	238
7.5	多配对样本的非参数检验	246
	练习题	255
第 8 章	SPSS 的相关分析	257
8.1	相关分析	257
8.2	绘制散点图	258

8.3 计算相关系数	266
8.4 偏相关分析	272
练习题	275
第 9 章 SPSS 的线性回归分析	277
9.1 回归分析概述	277
9.2 线性回归分析和线性回归模型	280
9.3 回归方程的统计检验	283
9.4 多元回归分析中的其他问题	295
9.5 线性回归分析的基本操作	298
9.6 线性回归分析的应用举例	304
9.7 曲线估计	312
练习题	319
第 10 章 SPSS 的聚类分析	321
10.1 聚类分析的一般问题	321
10.2 层次聚类	329
10.3 K-Means 聚类	343
练习题	350
第 11 章 SPSS 的因子分析	352
11.1 因子分析概述	352
11.2 因子分析的基本内容	356
11.3 因子分析的基本操作及案例	365
练习题	377



第1章

SPSS 统计分析软件概述

学习目标

1. 明确 SPSS 软件是一种专业的统计分析软件，了解 SPSS 的主要应用领域。
2. 熟练掌握 SPSS 进入和退出等基本操作，了解 SPSS 的基本窗口和菜单安排。
3. 掌握 SPSS 的三种使用方式以及它们的特点和应用场合。
4. 掌握 SPSS 联机帮助。
5. 掌握利用 SPSS 进行资料分析的基本步骤。

SPSS 是 Statistical Package for the Social Science 的英文缩写，意思是社会科学统计软件包。

SPSS 也是 Statistical Product and Service Solutions 的英文缩写，意思是统计产品与服务解决方案。

社会科学统计软件包 (SPSS) 是世界著名的统计分析软件之一。20 世纪 60 年代末，美国斯坦福大学的三位研究生研制开发了最早的统计分析软件 SPSS，并于 1975 年在芝加哥成立了专门研发和经营 SPSS 软件的 SPSS 公司。此时的 SPSS 软件主要运行在中小型计算机上，统称为 SPSSx 版，面向企事业单位用

户。20 世纪 80 年代初微型计算机出现, SPSS 公司以其敏锐的市场洞察力和雄厚的技术实力, 于 1984 年推出了运行在 DOS 操作系统上的 SPSS 微机版第一版, 随后又相继推出了第二版、第三版等, 统称为 SPSS/PC+ 版, 并确立了个人用户市场第一的地位。90 年代, 随着微机 Windows 图形操作系统的出现和盛行, SPSS 公司又研制出了以 Windows 为运行平台的 SPSS 第五版、第六版。90 年代中后期, 为适应用户在 Windows 95 操作系统环境下工作的习惯, 并迎合 Internet 的广泛使用, SPSS 第七版至第十五版又相继诞生, 统称为 SPSS for Windows 版。

1994—1998 年间, SPSS 公司陆续购并了 SYSTAT, BMDP, Quantime, ISL, ShowCase 等公司, 并将其各自的主打产品收纳于 SPSS 麾下, 从而使 SPSS 由原来单一的统计分析软件向为服务于企业、教育科研、政府机构等统计决策服务的综合性产品发展。为此 SPSS 公司已将原英文名称更改为 Statistical Product and Service Solutions, 即统计产品与服务解决方案。目前, SPSS 在全球约有几十万家产品用户, 分布于通讯、医疗、银行、证券、保险、制造、商业、市场研究、科研教育等多个行业和领域, 已成为世界上最流行、应用最广泛的专业统计分析软件。

从 SPSS/PC+ 版本到 SPSS for Windows 版本, SPSS 在用户操作和分析结果的展现方面做了非常大的改进。SPSS/PC+ 版本是通过命令行方式 (即用户输入 SPSS 命令和参数) 完成数据的管理和统计分析工作的, 统计分析文字结果和图形结果均以文本字符方式展现。SPSS for Windows 的第五版、第六版, 在保留以前版本的手工输入命令、参数操作方式的同时, 还为用户提供了直观的图形化菜单界面。用户的数据管理和统计分析工作可以非常方便地通过鼠标点选菜单或按钮并配合简单的对话框输入来实现, 从而免去了记忆命令和参数的负担, 也不需要任何计算机编程。同时, 对于图形分析结果, 能够以图形点阵的方式显示, 并允许用户对其进行编辑。SPSS for Windows 第七版以后的版本, 进一步完善了用户的操作界面, 增加了用户熟悉的与 Windows 95 风格一致的工具栏按钮。在分析结果的输出处理方面, 改变了原来版本的输出方式, 以崭新的树形结构方式管理分析结果, 以文字、表格、图形混排的形式展现分析结果, 以强大、灵活的编辑功能随心所欲地编辑分析结果, 以内容丰富的“统计教练”联机帮助方式帮助用户理解分析结果。同时, 还实现了对互联网的全面适应, 并支持数据的动态收集、分析和 HTML 格式报告等。

“易学，易用，易普及”已成为 SPSS 软件最大的竞争优势之一，也是广大数据分析人员对其偏爱有加的主要原因。而大量成熟的统计分析方法、完善的数据定义操作管理、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。

1.1 SPSS 使用基础

快速掌握 SPSS for Windows，需要首先了解 SPSS 安装的基本步骤，熟悉 SPSS 的基本操作环境，掌握启动和退出 SPSS 的方法。

1.1.1 SPSS 软件的安装和启动

作为 Windows 操作系统的应用软件产品，SPSS for Windows 安装的基本步骤与其他常用软件是基本相同的。基本步骤如下：

- (1) 启动计算机。将 SPSS 软件安装光盘放入光盘驱动器。
- (2) 运行资源管理器，鼠标双击光盘驱动器图标。
- (3) 在资源管理器目录窗口中找到 SPSS 的起始安装文件 setup 并执行。于是会看到 SPSS 安装的初始窗口，系统将自动进行安装前的准备工作。
- (4) 按照安装程序的提示，用户根据自己的需要填写和选择必要的参数。

例如：

- 接受软件使用协议。
- 指定将 SPSS 软件安装到计算机的哪个目录下。
- 选择安装类型。SPSS 有典型安装 (typical)、压缩安装 (compact) 和用户自定义安装 (custom) 三种安装类型。通常选择典型安装。
- 选择安装组件。SPSS 具有组合式软件的特征，在安装时用户可以根据自己的分析，选择部分组件 (模块) 安装。通常可接受安装程序的默认选择。
- 选择将软件安装在网络服务器上还是本地计算机上。通常安装在本地计算机上。
- 输入软件的合法序列号。在购买 SPSS 软件时厂商会提供序列号。

安装完毕后，用户应注意查看是否有安装成功的提示信息出现，进而判断是

否已经将 SPSS 成功地安装在计算机上了。安装成功后就可以启动运行 SPSS for Windows 软件了。

SPSS 的启动方法同一般常用软件的启动执行方法完全相同, 只需按以下顺序操作即可:

【开始】→【程序】→【SPSS for Windows】

1.1.2 SPSS 的基本窗口

了解 SPSS 的基本窗口是学习使用 SPSS 的入门点。应了解 SPSS 有哪些基本操作窗口, 各个窗口的功能和特点是什么, 各窗口之间的关系怎样。掌握了这些, 用户就能够很快理清软件使用的总体框架和脉络, 进而快速进入 SPSS 的核心。

SPSS 软件运行时多个窗口, 各窗口有各自的作用。但要快速入门, 只需要熟悉两个基本窗口即可, 它们是数据编辑窗口和结果输出窗口。

一、SPSS 数据编辑窗口

当 SPSS 启动后, 屏幕会显示如图 1—1 所示的窗口, 这就是数据编辑窗口。

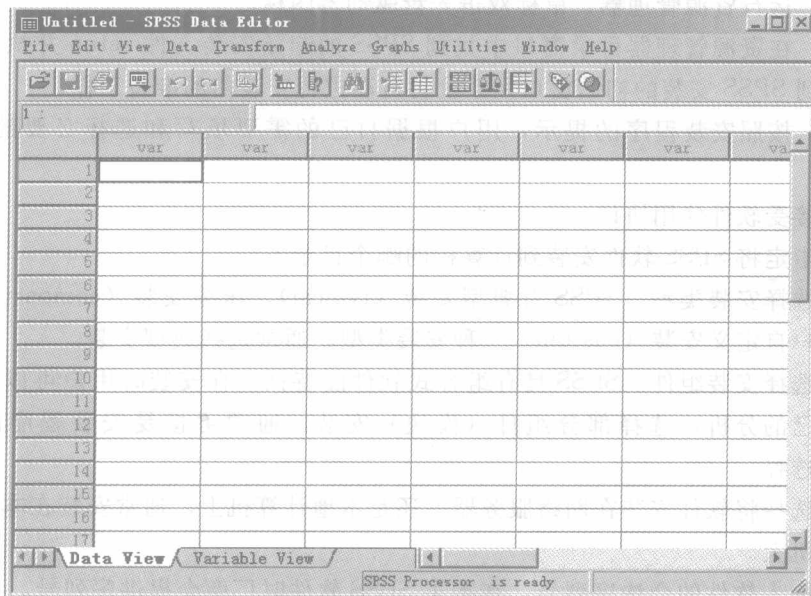


图 1—1 SPSS 数据编辑窗口

SPSS 数据编辑窗口（窗口标题为 SPSS Data Editor）是 SPSS 的主程序窗口，该窗口的主要功能是定义 SPSS 数据的结构、录入编辑和管理待分析的数据。

SPSS 数据编辑窗口是 SPSS 的主程序窗口。它在软件启动时自动打开，并一直保持打开状态直到退出 SPSS。在 SPSS 运行时只能打开一个数据编辑窗口，关闭数据编辑窗口意味着退出 SPSS。但在 SPSS 15 版中，允许同时打开多个数据编辑窗口。

数据编辑窗口的主要功能是：定义 SPSS 数据的结构、录入编辑和管理待分析的数据。SPSS 的所有统计分析功能都是针对该窗口中的数据。这些数据通常以 SPSS 数据文件的形式保存在计算机磁盘上，其文件扩展名为 .sav。



sav 文件格式是 SPSS 独有的，一般无法通过其他如 Word, Excel 等软件打开。

数据编辑窗口由窗口主菜单、工具栏、数据编辑区、系统状态显示区组成。

1. 窗口主菜单

窗口主菜单将 SPSS 常用的数据编辑、加工和分析的功能列了出来。用户可以通过点菜单完成相应的操作。菜单项对应的功能见表 1—1。

表 1—1 主窗口菜单及功能

菜单名	功能	解释
File	文件操作	对 SPSS 相关文件进行基本管理（如新建、打开、保存、打印等）
Edit	数据编辑	对数据编辑窗口中的数据进行基本编辑（如撤销/恢复、剪切、复制、粘贴），并实现数据查找、软件参数设置等功能
View	窗口外观 状态管理	对 SPSS 窗口外观等进行设置（如状态栏、表格线、变量值标签等是否显示、字体设置等）
Data	数据的操作 和管理	对数据编辑窗口中的数据进行加工整理（如数据的排序、转置、抽样、分类汇总、加权等）
Transform	数据基本 处理	对数据编辑窗口中的数据进行基本处理（如生成新变量、计数、分组等）
Analyze	统计分析	对数据编辑窗口中的数据进行统计分析和建模（如基本统计分析、均值比较、相关分析、回归分析、非参数检验等）
Graphs	制作统计 图形	对数据编辑窗口中的数据生成各种统计图形（如条形图、直方图、饼图、线图、散点图等）

续前表

菜单名	功能	解释
Utilities	实用程序	SPSS 其他辅助管理 (如显示变量信息、定义变量集、菜单编辑器等)
Windows	窗口管理	对 SPSS 的多个窗口进行管理 (如窗口切换、最小化窗口等)
Help	帮助	实现 SPSS 的联机帮助 (如语句检索、统计教练等)

在 SPSS 15 版菜单和子菜单的安排又略进行了调整。



初学 SPSS 时不必掌握菜单中的所有功能, 只需通过浏览菜单了解 SPSS 的大致功能即可。

2. 工具栏

同其他常用软件一样, SPSS 也将一些常用的功能以图形按钮的形式组织在工具栏中。图形按钮的功能都能在窗口主菜单中找到。用户可以直接点击工具栏上的某个按钮完成其相应的功能, 使操作更加快捷和方便。当鼠标停留在工具栏按钮上时, 计算机会自动提示相应按钮的功能说明。

3. 数据编辑区

数据编辑区是显示和管理 SPSS 数据结构和数据内容的区域。在数据编辑区中有两个视图, 分别是数据视图和变量视图, 依次用来以电子表格的形式录入和编辑管理 SPSS 的数据, 定义和修改 SPSS 数据的结构。每条数据都有一个顺序编号显示在编辑区的最左列。数据编辑区中的表格线可以通过 View 菜单下的 Grid Lines 选项设置成显示或不显示两种状态。

4. 系统状态显示区

系统状态显示区用来显示系统的当前运行状态。当系统等待用户操作时, 会出现“SPSS processor is ready”的提示信息, 该信息可以作为检查 SPSS 是否成功安装的手段。通过 View 主菜单下的 Status Bar 选项可以将系统状态区设置成显示或不显示两种状态。

二、SPSS 结果输出窗口

SPSS 结果输出窗口 (窗口标题为 SPSS Viewer) 是 SPSS 的另一个主要窗口, 该窗口的主要功能是显示管理 SPSS 统计分析结果、报表及图形。

SPSS 结果输出窗口如图 1—2 所示。输出窗口是显示管理 SPSS 统计分析结