



Statistics

21世纪统计学系列教材

Statistical Analysis and SPSS Application

统计分析 与SPSS的应用 (第四版)

薛 薇 编著



中国人民大学出版社



Statistics 21世纪统计学系列教材

Statistical Analysis and
SPSS Application

统计分析
与SPSS的应用 (第四版)

薛 薇 编著

中国人民大学出版社
· 北京 ·

图书在版编目 (CIP) 数据

统计分析与 SPSS 的应用/薛薇编著. —4 版. —北京: 中国人民大学出版社, 2014. 8
21 世纪统计学系列教材
ISBN 978-7-300-19863-7

I . ①统… II . ①薛… III . ①统计分析-软件包-高等学校-教材 IV . ①C819

中国版本图书馆 CIP 数据核字 (2014) 第 185265 号

21 世纪统计学系列教材
统计分析与 SPSS 的应用 (第四版)
薛 薇 编著
Tongji Fenxi yu SPSS de Yingyong

出版发行	中国人民大学出版社	邮政编码	100080
社 址	北京中关村大街 31 号	010 - 62511770 (质管部)	
电 话	010 - 62511242 (总编室) 010 - 82501766 (邮购部) 010 - 62515195 (发行公司)	010 - 62514148 (门市部) 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	版 次	2001 年 11 月第 1 版 2014 年 8 月第 4 版
印 刷	涿州市星河印刷有限公司	印 次	2014 年 8 月第 1 次印刷
规 格	185 mm×260 mm 16 开本	定 价	39.00 元
印 张	20 插页 1		
字 数	465 000		

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的研究水平，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，EVViews，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

前 言

《统计分析与 SPSS 的应用》自 2001 年第一版出版以来，至今已 13 年有余。它是编者在长期从事统计教学、科研以及企事业统计分析项目实践的基础上形成并不断完善的心血之作。本书自出版以来深受统计、经济、管理、生物医学、心理学等众多领域从事数据工作的教师、学者、学生以及实际工作者的广泛关注和好评，被许多高校选作统计和经济管理类相关专业统计分析软件课程的指定教材，也成为国内统计分析软件书籍的标杆之作。

对统计分析方法深入浅出的讲解，以实际数据为纽带说明 SPSS 的操作，以应用案例为背景阐述数据分析的思路，是本书一直坚持的风格。其特点在于：

第一，体系完整，内容规范，结构安排面向实践。本教材吸收了国外统计软件和统计学教材的共同特点，确保了教材的体系完整性及内容规范性。同时，面向实践，以统计分析的实际过程为主线，从软件基础和数据管理入手，按照数据分析需求由浅入深、分析方法从易到难的实战思路组织全书内容。全书涵盖了统计分析和软件的各个重要模块，囊括了描述统计、参数检验、方差分析、非参数检验、一般线性回归、聚类分析、因子分析等基本统计分析方法、经典多元统计方法和软件操作，能够提纲挈领地有效帮助读者全面快速掌握数据分析和软件应用。

第二，讲解的通俗性和严谨性相得益彰，使读者不仅能够知其然，而且能够知其所以然。本教材立足于统计方法的正确应用，注重以通俗易懂的语言论述统计方法的核心原理，注重对软件参数选项和分析结果统计含义的解释，注重对软件参数选项设置目的和应用必要性的讲解。同时，在提升可读性，弱化数学证明枯燥性的同时，兼顾严谨性，强化关键数学公式的论证，以加深读者对分析原理的理解。

第三，以应用为目标，精心选取典型的实际统计分析案例和数据，循序渐进地引导读者选择恰当的统计分析方法，一步步地指导读者利用统计软件进行数据的组织、整理、描述和建模分析等。

第三版出版后，编写者根据广大教师、学生和行业读者提出的反馈意见，在第四版中重点进行了如下修订：第一，迎合国内读者的语言习惯，软件采用了中文版本。第二，调整和更新了部分案例和数据，不仅提升了教材的时效性和现实性，而且确保案例在具有普遍参照性的同时更具方法应用的典型性。此外，增加了更为详尽的结果解释，使教材的可读性更强。第三，适当增加和调整了各章的练习题。第四，增加了带虚拟解释变量的回归分析。第五，强化了术语和用词的严谨性，使整本教材更加精炼和通顺。



随着数据分析应用需求的不断加强，掌握 SPSS 这样的权威统计分析软件是十分必要的。真诚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 解决实际数据分析问题。

本书不妥和错误之处，敬请各位读者不吝指正。

目 录

第 1 章 SPSS 统计分析软件概述	1
1. 1 SPSS 使用基础	2
1. 2 SPSS 的基本运行方式	7
1. 3 利用 SPSS 进行数据分析的基本步骤	11
练习题	13
第 2 章 SPSS 数据文件的建立和管理	14
2. 1 SPSS 数据文件	14
2. 2 SPSS 数据的结构和定义方法	18
2. 3 SPSS 数据的录入与编辑	25
2. 4 SPSS 数据的保存	28
2. 5 读取其他格式的数据文件	30
2. 6 SPSS 数据文件合并	36
练习题	42
第 3 章 SPSS 数据的预处理	44
3. 1 数据的排序	45
3. 2 查找重复个案	47
3. 3 变量计算	48
3. 4 数据选取	56
3. 5 计 数	59
3. 6 分类汇总	61
3. 7 数据分组	64
3. 8 数据预处理的其他功能	66
练习题	69
第 4 章 SPSS 基本统计分析	71
4. 1 频数分析	72
4. 2 计算基本描述统计量	78



4.3 交叉分组下的频数分析	85
4.4 多选项分析	94
4.5 比率分析	102
练习题	105
第 5 章 SPSS 的参数检验	107
5.1 参数检验概述	107
5.2 单样本 <i>t</i> 检验	110
5.3 两独立样本 <i>t</i> 检验	114
5.4 两配对样本 <i>t</i> 检验	120
练习题	123
第 6 章 SPSS 的方差分析	126
6.1 方差分析概述	126
6.2 单因素方差分析	128
6.3 多因素方差分析	144
6.4 协方差分析	156
练习题	164
第 7 章 SPSS 的非参数检验	168
7.1 单样本的非参数检验	169
7.2 两独立样本的非参数检验	179
7.3 多独立样本的非参数检验	186
7.4 两配对样本的非参数检验	193
7.5 多配对样本的非参数检验	200
练习题	207
第 8 章 SPSS 的相关分析	209
8.1 相关分析	209
8.2 绘制散点图	210
8.3 计算相关系数	217
8.4 偏相关分析	222
练习题	224
第 9 章 SPSS 的线性回归分析	226
9.1 回归分析概述	226
9.2 线性回归分析和线性回归模型	228
9.3 回归方程的统计检验	231
9.4 多元回归分析中的其他问题	241
9.5 线性回归分析的基本操作	243
9.6 线性回归分析的应用举例	248
9.7 曲线估计	254
练习题	260



第 10 章 SPSS 的聚类分析	262
10.1 聚类分析的一般问题	262
10.2 层次聚类	269
10.3 K-Means 聚类	281
练习题	286
第 11 章 SPSS 的因子分析	288
11.1 因子分析概述	288
11.2 因子分析的基本内容	291
11.3 因子分析的基本操作及案例	299
练习题	309

C 第1章

Chapter 1 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. 掌握利用 SPSS 进行数据分析的基本步骤。

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1版，随后又相继推出了第2版、第3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5版、第6版。90年代中后期，为适应用户在Windows95操作系统环境下工作的习惯，并迎合Internet的广泛使用，SPSS第7版至第17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之麾下，从而使 SPSS 由原来单一的统计分析软件向为企业、教育科研、政府等机构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企业用户，分布于通信、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。

2009 年 7 月，IBM 公司斥资 12 亿美元收购了 SPSS 软件公司。SPSS 18 和 19 版更新命名为 PASW (Predictive Analytics SoftWare) Statistics。20 版至目前的 22 版命名为 IBM SPSS Statistics。

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 5 版、第 6 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 7 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“统计辅导”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。18 版以后增加了多语言版，支持中文窗口显示和中文结果输出，更便于中国用户使用。

“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 需要首先熟悉 SPSS 的基本操作环境，掌握启动和退出 SPSS 的方法。

1.1.1 SPSS 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 软件运行时有多个窗口，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，它们是数据编辑器窗口和结果查看器窗口。

一、SPSS 数据编辑器窗口

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

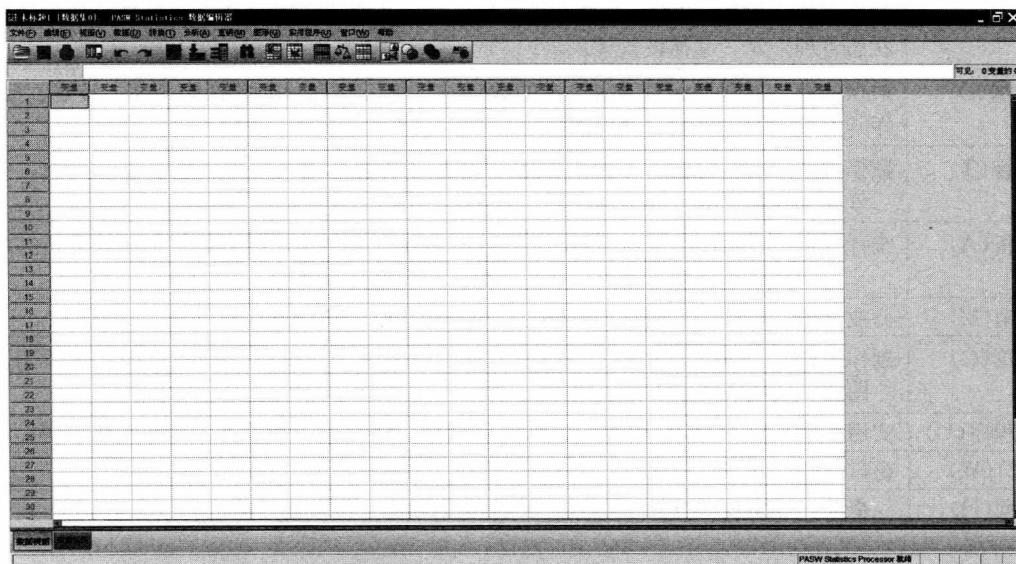


图 1—1 SPSS 数据编辑器窗口

SPSS 数据编辑器窗口是 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。其中，【数据视图】选项卡用于显示 SPSS 数据的内容；【变量视图】选项卡用于显示 SPSS 数据的结构。

SPSS 运行时可同时打开多个数据编辑器窗口。各数据编辑器窗口分别显示不同的数据集合（简称数据集）。按打开的先后顺序，各数据集依次自动命名为：数据集 0，数据集 1，数据集 2，等等。其中只有一个数据集为当前数据集，称为活动数据集。用户只能对某一时刻活动数据集中的数据进行分析。此外，数据集通常以 SPSS 数据文件的特有格式保存在磁盘上，其文件扩展名为 .sav。

.sav 文件格式是 SPSS 独有的，一般无法通过 Word、Excel 等其他软件打开。

数据编辑器窗口由窗口主菜单、工具栏、数据编辑区、系统状态显示区组成。

1. 窗口主菜单

窗口主菜单列出了 SPSS 常用的数据编辑、加工和分析的功能。用户可以通过点击菜单完成相应的操作。菜单项对应的功能如表 1—1 所示。



表 1—1

主窗口菜单及其功能

菜单名	功能	解释
文件(F)	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）
编辑(E)	数据编辑	对数据编辑器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
视图(V)	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）
数据(D)	数据的操作和管理	对数据编辑器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）
转换(T)	数据基本 处理	对数据编辑器窗口中的数据进行基本处理（如生成新变量、计数、分组等）
分析(A)	统计分析	对数据编辑器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）
直销(M)	商业营销	提供了商业营销分析的专用方法（如 RFM 分析等）
图形(G)	制作统计 图形	对数据编辑器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）
应用程序(U)	应用程序	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）
窗口(W)	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）
帮助(H)	帮助	实现 SPSS 的联机帮助（如语句检索、统计辅导等）

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 将一些常用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功能，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，以及定义和修改 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【视图(V)】菜单下的【网格线】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“PASW Statistics processor 就绪”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【视图(V)】菜单下的【状态栏(S)】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 结果查看器窗口

SPSS 结果查看器窗口是 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 结果查看器窗口如图 1—2 所示。查看器窗口是显示管理 SPSS 统计分析结果、报表及图形的窗口。

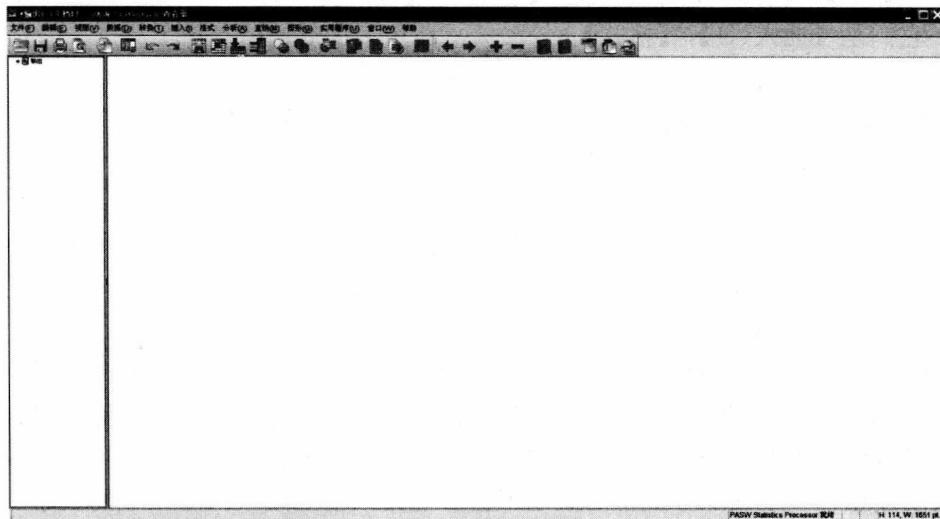


图 1—2 SPSS 结果查看器窗口

SPSS 统计分析的所有输出结果都显示在该窗口中。输出结果通常以 SPSS 输出文件的形式保存在计算机磁盘上，其文件扩展名为 . spv。

spv 文件格式是 SPSS 独有的，一般无法通过 Word、Excel 等其他软件打开。

结果查看器窗口有如下特点：

1. 创建或打开窗口的时机

查看器窗口通常在以下两种时机打开：第一，第一次打开 SPSS 数据文件时，由 SPSS 自动创建并打开；第二，在 SPSS 运行过程中由用户手工创建或打开，依次选择的菜单为：

【文件(F)】→【新建(N)】/【打开(O)】→【输出(O)】

2. 允许同时创建或打开多个查看器窗口

查看器窗口可以同时创建或打开多个，且可以利用主菜单中的【窗口(W)】菜单实现各个窗口间的相互切换。多个查看器窗口中只有一个为目标查看器窗口或称当前查看器窗口，意味着以后的统计分析结果将输出到该窗口中。设置多个查看器窗口的目的是允许用户将同一批数据的不同分析结果指定输出到不同的窗口中，这样就可以将它们以不同的文件名分别存放到磁盘上。

结果查看器窗口由窗口主菜单、工具栏、分析结果显示区组成。

1. 窗口主菜单

查看器窗口主菜单中的菜单选项可以大致分成三大类。第一类：菜单名（包括子菜单）、功能、操作对象与数据编辑器窗口完全相同，它们是【分析(A)】,【图形(G)】,【实用程序(U)】,【窗口(W)】,【帮助】。第二类：一部分菜单名（包括子菜单）、功能、操作对象与数据编辑器窗口相同，另一部分子菜单是附加的且针对查看器窗口，它们是【文件(F)】,【编辑(E)】,【视图(V)】。其中如【编辑(E)】菜单中的【大纲(I)】,【视图(V)】菜单中的【显示】,【隐藏(H)】等都是专门针对查看器窗口的。第三类：查看器窗口独有的菜单，它们是【插入(I)】,【格式】。

查看器窗口主菜单如此设置的目的是方便用户操作。但初学者应特别注意，虽然 SPSS 中有些窗口的主菜单名是相同的，但它们下面的子菜单却因操作窗口不同而存在较大的差异。因此，在使用时用户应注意区分，以免发生操作上的混乱。

2. 工具栏

查看器窗口的工具栏除了保留数据编辑器窗口中的某些功能按钮，还增添了一些自己特有的功能按钮，它们在查看器窗口主菜单中也有各自的对应项。

3. 分析结果显示区

分析结果显示区是专门显示统计分析结果的区域。该区域分成左右两个部分。左边区域以类似于 Windows 资源管理器的树形形式显示已有分析结果的目录，称为目录区。右边区域是各分析结果的详细报告，称为内容区。目录区和内容区都有各自独立的纵向和横向屏幕滚动按钮，一个区域中屏幕的滚动不会影响另一个区域。两个区域的大小可以随意调整，只需将鼠标移动到两区域中间的分隔栏上，然后左右拖动鼠标即可。目录区和内容区中均有一个红色的箭头，它们所指示的内容一一对应。用户可以分别对两个区域中的内容进行增、删、改等编辑管理操作。

总之，SPSS 的数据编辑器窗口专门负责输入和管理待分析数据，结果查看器窗口负责接收和管理统计分析的结果。数据的输入和结果的输出是在不同窗口中进行的，这点与 Excel 等其他有统计分析功能的软件有较大差异。

1.1.2 SPSS 软件的退出

退出 SPSS 的方法与退出一般常用软件的方法基本相同，只需要依次选择 SPSS 中的菜单即可，即

【文件(F)】→【退出(X)】

在退出 SPSS 之前，一般会提示用户以下两个问题：

- 是否将数据编辑器窗口中的数据存到磁盘上，文件扩展名为.sav。
- 是否将结果查看器窗口中的分析结果存到磁盘上，文件扩展名为.spv。

这时，用户应根据实际情况，指定将 SPSS 数据文件或结果文件存放到哪个磁盘上，并输入用户名。

1.2 SPSS的基本运行方式

SPSS为用户提供了三种基本运行方式，它们是完全窗口菜单方式、程序运行方式、混合运行方式。这三种运行方式分别适合于不同的用户和不同的统计分析要求。

1.2.1 完全窗口菜单方式

完全窗口菜单方式，是指在使用SPSS的过程中，所有的分析操作都通过菜单、按钮、输入对话框等方式来完成。

完全窗口菜单方式是一种最常见和普遍的使用方式，其最大的优点是简洁和直观。用户不需要了解任何计算机编程的概念，只要熟悉Windows的基本操作（如复选框、单选框、下拉框、对话框）并懂得相应的统计知识，就可以非常方便地完成统计分析工作。图1—3所示的就是SPSS中符合Windows使用习惯的几种基本操作方式。

另外，SPSS还有一种较常见但又很有特点的操作方式，如图1—4所示。



图1—3 SPSS常见操作方式

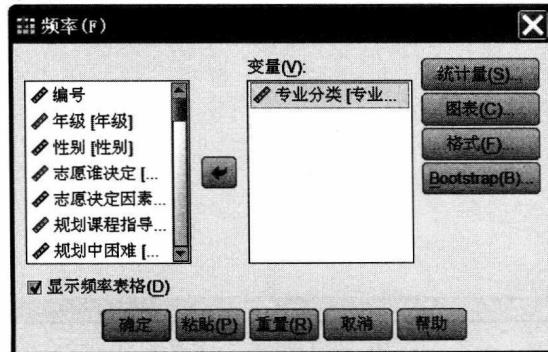


图1—4 SPSS变量选择操作方式