

21世纪统计学系列教材

统计分析与应用 SPSS的应用

薛薇 编著

中国人民大学

社

361

CS11-43
XSP

21 世纪统计学系列教材

统计分析与 SPSS 的应用

薛 薇 编著

中国人民大学出版社

图书在版编目 (CIP) 数据

统计分析与 SPSS 的应用/薛薇编著.
北京: 中国人民大学出版社, 2001
21 世纪统计学系列教材

ISBN 7-300-03907-3/F·1177

I. 统…

II. 薛…

III. ①统计分析-高等学校-教材

②统计分析-软件包, SPSS-高等学校-教材

IV. C81

中国版本图书馆 CIP 数据核字 (2001) 第 067321 号

21 世纪统计学系列教材
统计分析与 SPSS 的应用
薛薇 编著

出版发行: 中国人民大学出版社

(北京中关村大街 31 号 邮编 100080)

邮购部: 62515351 门市部: 62514148

总编室: 62511242 出版部: 62511239

E-mail: rendafx@public3. bta. net. cn

经 销: 新华书店

印 刷: 三河市新世纪印刷厂

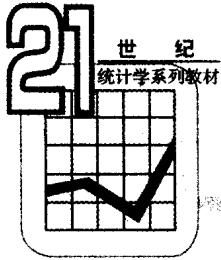
开本: 787×980 毫米 1/16 印张: 17.25

2001 年 11 月第 1 版 2001 年 11 月第 1 次印刷

字数: 315 000

定价: 20.00 元

(图书出现印装问题, 本社负责调换)



前 言

随着市场经济的发展和竞争的日益激烈，如何运用科学的分析方法，对收集到的数据作出准确、及时的分析并制定正确的决策，已经成为企业极为关注的问题。统计学作为收集、加工和分析数据的方法科学，在企业决策中无疑起着重要作用。同时，随着现代计算机技术的不断发展和成熟，统计分析方法的实际运用和实施已经极为普遍，其中统计分析软件扮演着不可缺少的关键性角色。

本书力求从统计分析方法的具体实现角度出发，根据当前流行的计算机统计分析软件 SPSS 自身的特点，以实际数据的分析为纽带，将对统计分析方法基本思想的理解与计算机软件的实际操作有机结合，使读者在理解常用统计分析方法思想的同时，能够将统计分析软件灵活运用并实施到实际统计分析工作中去，从而为实际统计分析工作提供一种切实可行的现实思路和方案。

本书在统计分析方法的讲解上，着重分析方法基本思想的讨论，力求通俗易懂、深入浅出，并通过实例，帮助读者理解方法和运用方法。这样，你会惊喜地发现统计就在你我身边。在内容安排上，包括基本统计分析方法，如：描述统计、频数分析、交叉列联表分析、多选项分析、参数检验、方差分析、非参数检验；同时也包括多元统计分析方法，如：多元回归分析、聚类分析、因子分析，从而使读者能够由浅入深地掌握统计分析方法的思想精华。

本书在统计分析软件 SPSS 的讲解上，以统计分析方法为主线，较为系统地讨论了软件的使用方法，并注意对实例分析结果加以解释说明。这样，读者依照教材就能够逐步独立地进行实际的统计分析和工作。

本书选用 SPSS for Windows 8.0 版本为讲解对象。在以后的新版本中，对主菜单的安排、文件的操作、图形的处理等方面作了一些改进。例如：在主菜单安排上，8.0 版本中的 Statistics 菜单改为 Analyze，Summarize 被分成了 Reports 和 Descriptive Statistics 两个子菜单项。在文件处理上，提供了 TextWizard 过程，使用户更为方便地打开文本数据文件。

本书讲解时注意兼顾统计分析方法和操作实现两个方面，力求克服有些 SPSS 软件使用手册中，仅讲解操作步骤，不讲解方法和输出结果的弱点。在这本教材的编写过程中，编审委员会的易丹辉教授、顾岚教授、吴喜之教授以及张云副教授、陈虹老师等都给予了具体的指导和帮助，并对初稿多次提出了许多极为宝贵的意见和建议，使编者受益匪浅，这里一并对各位老师表示真挚的感谢。

本书不妥和错误之处，敬请各位读者批评指正。

薛 薇

于北京，中国人民大学

2001 年 3 月



目 录

第 1 章 统计分析软件 SPSS for Windows 概述	(1)
§ 1.1 软件的发展及特点	(1)
§ 1.2 软件基础知识	(2)
第 2 章 SPSS 数据文件的建立和管理	(9)
§ 2.1 数据文件的特点与结构	(9)
§ 2.2 定义数据文件的结构	(15)
§ 2.3 数据的录入与保存	(18)
§ 2.4 数据的编辑	(20)
§ 2.5 与其他计算机软件数据共享	(23)
第 3 章 SPSS 数据文件的基本加工和处理	(28)
§ 3.1 数据文件的整理	(28)
§ 3.2 数据加工	(38)
§ 3.3 数据分组	(50)
§ 3.4 数据文件的其他处理功能	(55)

第 4 章 SPSS 输出窗口的操作和管理	(59)
§ 4.1 索引输出区管理	(60)
§ 4.2 详解输出区管理	(63)
§ 4.3 输出窗口中分析结果的存取	(69)
第 5 章 SPSS 基本统计分析	(71)
§ 5.1 案例说明	(71)
§ 5.2 计算基本描述统计量	(73)
§ 5.3 变量的频数统计	(79)
§ 5.4 交叉分组下的频数分析	(84)
§ 5.5 多选项分析	(92)
§ 5.6 基本统计分析的报表制作	(105)
第 6 章 SPSS 的参数检验	(113)
§ 6.1 统计推断的基本方法	(113)
§ 6.2 单样本 t 检验	(115)
§ 6.3 两独立样本 t 检验	(117)
§ 6.4 两配对样本 t 检验	(122)
第 7 章 SPSS 的非参数检验	(127)
§ 7.1 单样本非参数检验	(127)
§ 7.2 两独立样本非参数检验	(138)
§ 7.3 多独立样本非参数检验	(146)
§ 7.4 两配对样本非参数检验	(152)
§ 7.5 多配对样本非参数检验	(158)
第 8 章 SPSS 的方差分析	(166)
§ 8.1 方差分析概述	(166)
§ 8.2 单因素方差分析	(167)
§ 8.3 单因素方差分析案例分析	(172)
§ 8.4 均值比较	(175)
§ 8.5 多因素方差分析和协方差分析	(178)
§ 8.6 多因素方差分析案例分析	(183)

第 9 章 SPSS 的相关分析和线性回归分析	(185)
§ 9.1 相关分析	(185)
§ 9.2 相关分析案例分析	(191)
§ 9.3 偏相关分析	(192)
§ 9.4 线性回归分析	(195)
§ 9.5 线性回归分析案例分析	(209)
§ 9.6 曲线估计	(217)
§ 9.7 权重估计	(221)
第 10 章 SPSS 的聚类分析	(224)
§ 10.1 聚类分析概述	(224)
§ 10.2 层次聚类分析	(225)
§ 10.3 快速聚类分析	(236)
§ 10.4 案例分析	(242)
第 11 章 SPSS 的因子分析	(250)
§ 11.1 什么是因子分析	(250)
§ 11.2 因子分析的数学模型和相关概念	(251)
§ 11.3 因子分析的基本步骤	(253)
§ 11.4 案例分析	(259)
参考文献	(268)



第 1 章

统计分析软件 SPSS for Windows 概述

§ 1.1 软件的发展及特点

SPSS 是 Statistical Package for Social Science 的英文缩写，即：社会科学统计分析软件包。它是当今世界上公认和流行的综合统计分析软件包。在我国，SPSS for Windows 以其强大的统计分析功能、方便的用户操作界面、灵活的表格式分析报告及其精美的图形展现，受到了社会各界统计分析人员的喜爱。目前，SPSS 软件使用已经成为许多大专院校统计学专业和财经类、管理类专业本科学生的必修课程。

SPSS 起源于 20 世纪 60 年代的美国斯坦福大学。进入 70 年代，专门研制和经营 SPSS 软件的 SPSS 公司成立。此时的 SPSS 软件是在中小型计算机上运行，其版本统称为 SPSSx。80 年代初，随着微型计算机的出现，SPSS 公司以其敏锐的目光，迅速研制成功了运行在微型计算机 DOS 操作系统上的 SPSS 的第一版、第二版、第三版等，从而使 SPSS 得到了更为广泛的使用，并占领了计算机统计分析软件的大部分市场份额。此时的 SPSS 版本统称为 SPSS/PC+。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 SPSS for

Windows 的第五版、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以及对国际互联网 Internet 的广泛使用,SPSS for Windows 的第七版、第八版、第九版、第十版又相继诞生。目前,SPSS for Windows 已经成为世界各地流行使用的统计分析软件,并和权威的 SAS 等统计分析软件共享市场。

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有了非常大的改进。SPSS/PC+ 版本是通过用户输入 SPSS 命令程序和参数的方式来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 的第五版和第六版, 在保留以前版本的人工输入命令、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击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, 以崭新的树型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。本书以 SPSS for Windows 8.0 版本为讲解内容。

§ 1.2 软件基础知识

一、SPSS for Windows 的安装和启动

SPSS for Windows 作为 Windows 95/98 下的软件产品, 具有和其他软件基本相同的安装步骤。基本步骤概括如下:

- (1) 启动 Windows, 将存有 SPSS for Windows 的光盘放入光盘驱动器。
- (2) 运行资源管理器, 读取光盘文件。
- (3) 双击执行安装程序文件 setup。
- (4) 根据用户的需要选择所需要安装的程序项。

正确完成安装步骤以后, 就可以启动运行 SPSS for Windows 软件了。

二、SPSS for Windows 中的两个基本窗口

1. SPSS 的数据编辑窗口

正常启动 SPSS for Windows 后屏幕显示如图 1—1 所示。

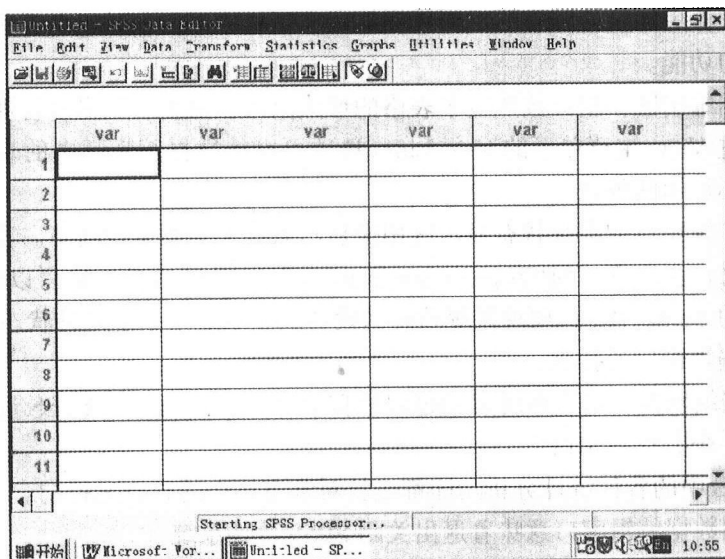


图 1—1 SPSS 数据编辑窗口

SPSS 数据编辑窗口的窗口标题为 data editor，它是 SPSS for Windows 的主程序窗口。数据编辑窗口是对分析对象——SPSS 数据文件进行录入、修改、管理等基本操作的窗口。SPSS 数据文件均以 .sav 作为文件扩展名存储在磁盘中。

数据编辑窗口包括几个部分：窗口主菜单、工具栏、数据编辑区、系统状态显示区。

数据编辑窗口的窗口主菜单基本功能是：

File: SPSS 文件的操作菜单

Edit: SPSS 文件的编辑菜单

Data: SPSS 数据文件的建立和编辑菜单

View: SPSS 用户界面基本状态设置等菜单

Transform: SPSS 数据的基本处理菜单

Statistics: SPSS 的统计分析菜单

Graphs: SPSS 的统计图形菜单

Utilities: SPSS 的相关应用和设置菜单

Windows: SPSS 中各窗口的切换菜单

Help: SPSS 的帮助菜单

以图形按钮组成的工具栏将常用的 SPSS 功能列示出来，这些功能都可以在

相应的功能菜单中找到。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，使得操作更加快捷和方便。当鼠标停留在工具栏按钮上时，系统会自动提示相应按钮的功能。

在数据编辑区，用户按照电子表格的形式录入、修改、编辑和管理待分析的数据。通过 View 主菜单下的 Grid Lines 选项可以将数据编辑区中的表格线设置成显示或不显示两种状态。

数据编辑窗口的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。也可以通过 View 主菜单下的 Status Bar 选项将系统状态区设置成显示或不显示两种状态。

数据编辑窗口的特点：

- 数据编辑窗口是在系统启动过程中自动打开的窗口，在 SPSS 运行过程中只能打开一个数据编辑窗口。
- SPSS 中的各种统计分析功能都是针对该窗口中的数据进行的。
- 关闭数据编辑窗口意味着退出关闭 SPSS 软件系统。

2. SPSS 的输出窗口

SPSS 输出窗口的窗口名为 Viewer，它是显示和管理 SPSS 统计分析结果、报表及图形的窗口，如图 1—2 所示。出现在该输出窗口中的内容均以 .spo 作为文件扩展名存储在磁盘上。

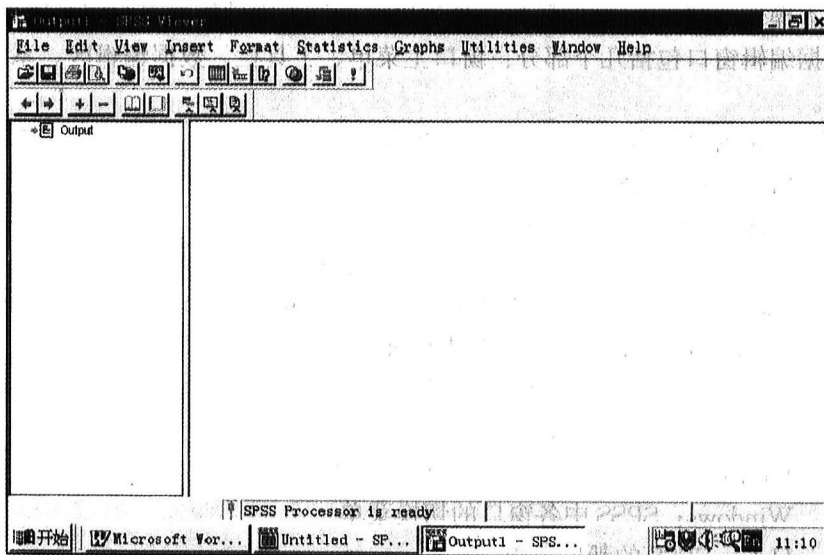


图 1—2 SPSS 输出窗口

输出窗口包括几个部分：窗口主菜单、工具栏、分析结果显示区、状态显示区。

输出窗口的窗口主菜单的一部分菜单，如：Statistics、Graphs、Utilities、Windows、Help 的功能与数据编辑窗口中的相同，是针对数据编辑窗口中的分析数据的；还有一部分功能菜单，如：View、Insert、Format 是输出窗口的专用菜单；其余的功能菜单，如：File、Edit 则是数据编辑窗口和输出窗口共用的。菜单这样设置的目的是为了更方便用户的操作。

需要说明的是，虽然 SPSS 中有些窗口的主菜单名有重名的情况，但是它们下面的子菜单功能却是随不同的当前窗口而不同的，其服务的对象也是不同的。因此，在使用时，用户应注意哪个窗口是当前窗口。

输出窗口的工具栏除了保留了数据编辑窗口中的某些图形按钮外，还增添了自己特有的功能按钮。

分析结果显示区是显示统计分析结果的地方。区域分成左右两个部分。左边以类似于 Windows 资源管理器的树形结构形式显示已有的分析结果的索引；右边显示的是各个分析结果的详细内容。左右两边的内容以两个红色的右箭头为标记——对应起来。用户可以对该区域中的内容进行增、删、改等编辑管理操作。

输出窗口的状态显示区与数据编辑窗口的状态区的区别是多了一个红色的“!”标记。该标记表示当前的输出窗口为主输出窗口。主输出窗口就是下一次分析结果将输出到的那个窗口。

输出窗口的特点：

- 输出窗口是在用户进行第一次统计分析时被系统自动打开的。
- 在 SPSS 运行过程中用户还可以手工创建若干个新输出窗口。菜单选项为：File -> New -> Output。手工创建若干个新输出窗口后，用户可以将同一批分析数据的不同分析结果指定输出到不同的输出窗口中，这样就可以将不同输出窗口中的内容以不同的文件名分别存放到磁盘上。

● 不同的输出窗口之间可以相互自由切换，可以利用主菜单中的 Windows 功能菜单实现。它们中只有一个主输出窗口。主输出窗口的状态栏上有一个红色的“!”标记，而其他输出窗口的状态栏上则没有这个标记。如果用户希望将以下的统计分析结果输出到某个输出窗口中，需要按动工具栏上的“!”图形按钮来指定某输出窗口为主输出窗口。

另外，SPSS 还有一个窗口名为 draft Viewer 的输出窗口。它可以通过 File -> New -> draft view 的菜单选项来创建。该窗口功能与上面提到的输出窗口的功能完全相同。但是，出现在 draft viewer 输出窗口中的内容是以 .rtf 作为文件扩展名存放到磁盘上的。.rtf 文件是按丰富文本格式 (rich text format)

存储的，它可以被 word 等其他软件直接读取和编辑。

三、SPSS for Windows 的退出

SPSS 退出可以通过 File -> Exit 的菜单选项实现，也可以直接按数据编辑窗口右上角的 按钮。在退出 SPSS 之前，一般会向用户提出以下两个问题：

- 是否将 data editor 窗口中的数据存到磁盘上，数据文件扩展名为 .sav；
- 是否将 viewer 窗口中的数据存到磁盘上，结果文件扩展名为 .spo 或 .rtf。

这时，用户应指定将 SPSS 数据文件或结果文件存放到哪个磁盘上，并输入相应文件名的用户名部分。

四、SPSS for Windows 的运行方式

SPSS 为用户提供了三种基本运行方式，分别为：完全窗口菜单方式、程序运行方式、混合运行方式。

1. 完全窗口菜单方式

完全窗口菜单方式是在 SPSS 运行中，所有分析操作过程都是用户通过点菜单或按钮和输入对话框的方式进行的。如果用户不清楚某些统计选项的含义，可以将鼠标移动到相应的选项上，然后单击鼠标右键。于是，在屏幕上就会出现相应的解释说明，从而有助于用户理解统计选项的意义。

完全窗口菜单方式是一种经常使用的运行方式，用户不需要任何计算机编程的概念，只要熟悉 Windows 的基本操作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。这种工作方式很适用于一般的统计分析人员和 SPSS 的初学者。在以后的章节中，我们主要讲解这种方式。

2. 程序运行方式

程序运行方式是统计分析人员首先根据自己的分析步骤，手工编写相应的 SPSS 命令程序；然后将编写好的程序一次性提交计算机运行。SPSS 软件系统会自动按照程序命令语句的前后顺序自动逐句执行 SPSS 的命令。程序运行方式很适用于熟练的 SPSS 程序员进行特殊分析的需要。在进行大规模的统计分析工作时，程序运行方式能够按照程序自动进行分析，不需要分析过程中的人工干预。

程序运行方式的实现步骤有以下两大步骤：

第一步，打开窗口名为 Syntax Editor 的语句窗口并编写 SPSS 程序。

语句窗口是专门用来供 SPSS 程序员编写 SPSS 程序的窗口。该窗口中的 SPSS 命令程序是按照 SPSS 软件的语法规则编写的，它所分析处理的数据是数

据编辑窗口中的数据。需要说明的是，SPSS for Windows 版本中的 SPSS 命令语法与 DOS 下的 SPSS 各版本的语法是完全兼容的。因此，熟悉 DOS 版本下的 SPSS 用户可以沿袭以前版本的语法来编写 Windows 版本下的 SPSS 程序。语句窗口中的 SPSS 命令程序均以 .sps 为文件扩展名存放在磁盘上。

语句窗口有自己的主菜单项、工具栏和系统状态栏，其功能与数据编辑窗口和输出窗口类似。

打开语句窗口的方式可以有两种：

(1) 手工打开

菜单选项：File -> New -> Syntax。可以创建一个新的语句窗口。

(2) SPSS 启动时自动打开

菜单选项：Edit -> Option。如图 1—3 所示。这样，在 SPSS 启动时会自动启动语句窗口。

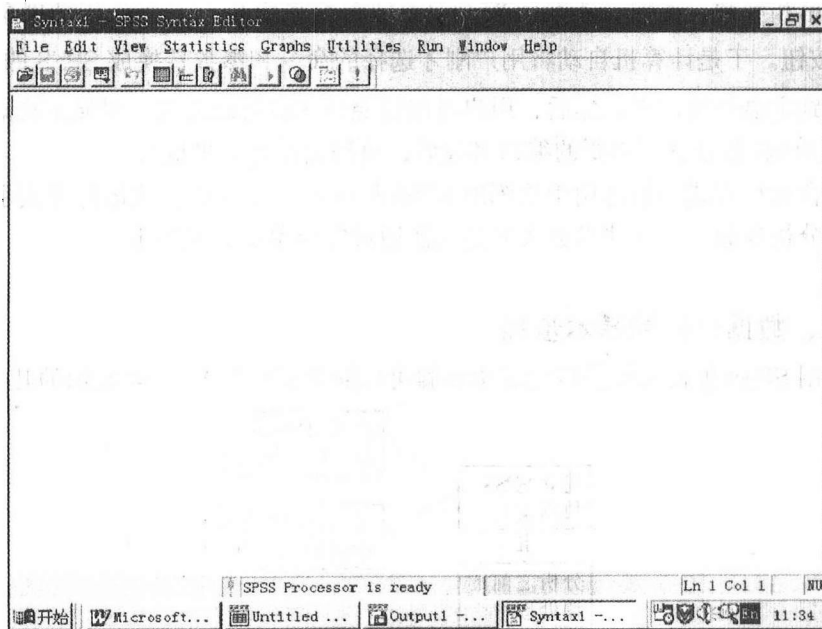


图 1—3 SPSS 语句窗口

语句窗口的特点是：

- 在 SPSS 运行过程中用户可以手工创建若干个新语句窗口。这样就允许用户将对同一批数据的不同分析命令程序分别写在不同的语句窗口中，从而可以将它们分别以不同的文件保存在磁盘上。

● 用户可以在多个语句窗口之间相互切换，同时编辑多个 SPSS 程序。其中在语句窗口的系统状态栏上有红色的“!”的为主语句窗口。

第二步，执行语句窗口中的 SPSS 程序。

选择语句窗口主菜单中的 Run 项。SPSS 提供了四种程序执行的方式：

(1) 执行全部 All：执行当前语句窗口中的所有 SPSS 命令程序。

(2) 执行部分 Selection：仅执行语句窗口中当前已选中的 SPSS 命令程序。命令选中可以通过鼠标在相应 SPSS 命令行上拖动来实现。

(3) 执行当前 Current：仅执行语句窗口中当前光标所在行的 SPSS 命令。

(4) 执行到最后 To end：仅执行语句窗口中当前光标所在行之后的所有 SPSS 命令程序。

3. 混合运行方式

混合运行方式是前两种运行方式的综合。首先按照菜单运行方式选择统计分析的菜单和参数。此时，不马上按 **OK** 按钮提交系统执行，而是按动屏幕上的 **Paste** 按钮。于是计算机自动将用户刚才选择的命令和参数转换成 SPSS 的命令，并粘贴到主语句窗口中；然后，用户可以按照程序运行的方式，对显示在语句窗口中的 SPSS 命令进行必要的编辑修改后，再提交给计算机执行。

混合运行方式一般适用于熟练的 SPSS 程序员。可以通过该运行方式手工输入某些分析参数，而这些参数大多是无法通过完全菜单方式给定。

五、数据分析的基本步骤

利用 SPSS 进行数据分析的基本步骤可以简单概括为图 1—4 所示的几步。

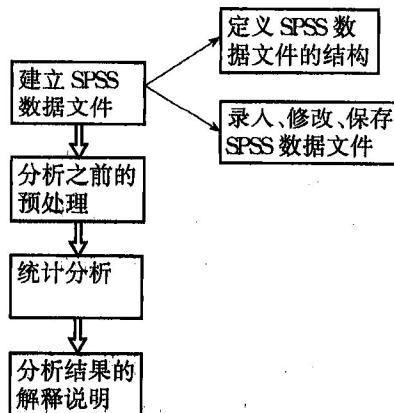
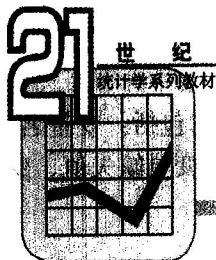


图 1—4



第 2 章

SPSS 数据文件的建立和管理

§ 2.1 数据文件的特点与结构

一、数据文件的特点

利用 SPSS 进行数据分析的第一步工作就是要建立 SPSS 的数据文件。应首先了解什么是 SPSS 的数据文件。

SPSS 数据文件是一种有结构的数据文件，它包括文件结构和数据两部分。

SPSS 数据文件中的一列数据称为一个变量。每个变量都应有一个名称，即：变量名，变量名是存取 SPSS 每个变量的惟一标识。SPSS 数据文件的结构是对每个变量及相关属性的描述。SPSS 数据文件中的一行数据称为一条个案 (Case)。所有个案组成了 SPSS 数据文件的数据部分。

SPSS 数据文件与一般的文本不同点在于：一般的文本文件仅有纯数据部分，而没有关于结构描述的部分。因此，SPSS 数据文件不能像一般的文本文件那样可以直接被大多数编辑软件读取，而只能在 SPSS 软件中打开。

建立 SPSS 数据文件应通过两大步骤实现：

第一步：定义 SPSS 数据文件的结构；