

马春庭 主编

掌握和精通 SPSS 10

掌握和精通 SPSS10

马春庭 主编

高 萍 武海军 高智勇 副主编

机 械 工 业 出 版 社

SPSS (Statistics Package for Social Science) for Windows 是世界上著名的统计分析软件, 适用于自然科学、社会科学各领域的统计分析。全书共分 8 章, 主要介绍 SPSS 的基本概念、基本操作与统计分析方法, 其中包括系统运行管理方式、数据文件的建立与编辑、数据操作、描述性统计分析、相关分析、方差分析、回归分析、均值检验、聚类与判别分析、因子分析、非参数检验、统计图形的生成与编辑等内容。

本书是一本实用性很强的普及性读物。全书内容新颖、文字流畅、通俗易懂, 适用于 SPSS 的初学者及想利用 SPSS 进行统计分析的科研工作者使用。

图书在版编目 (CIP) 数据

掌握和精通 SPSS10/马春庭主编. —北京: 机械工业出版社, 2001.1
ISBN 7-111-08700-3

I. 掌… II. 马… III. 统计分析-软件包, SPSS10 IV. C819

中国版本图书馆 CIP 数据核字(2000)第 88118 号

机械工业出版社(北京市百万庄大街 22 号 邮政编码 100037)

责任编辑: 曲彩云

封面设计: 姚毅 责任印制: 路琳

中国建筑工业出版社密云印刷厂印刷·新华书店北京发行所发行

2001 年 2 月第 1 版第 1 次印刷

787mm×1092mm1/16·21.75 印张·537 千字

0001—4000 册

定价: 35.00 元

凡购本书, 如有缺页、倒页、脱页、由本社发行部调换

总 序

计算机的飞速发展推动了科学技术的进步和在人们日常生活的应用。而“打开科学大门的钥匙”——数学在科学技术进步方面的作用是决定性的。力图将人们从纷繁复杂的数学计算、推导中解放出来的数学软件也可谓“百花齐放、百家争鸣”。国外比较著名的数学软件主要包括：

(1) 符号计算、代数运算等数学软件

Albert、Bergman、CALI、Calculus、Cartan、CHEVIE、CLA、Classi、CoCoA、DELiA、Derive、DIMSYM、FeynArts、FORM、Galois、GROEBNER、GRTensorII、Ideals、KASH、LiE、Symmetries、LIE、liesymm、Macsyma、Macaulay、Maple、MAS、Mathcad、Mathematica、MuPad、Numbers、ODE、ODESolve、PARI、PARAMAX、Reduce、SIMATH、Singular、SPDE、STENSOR、SYM_DE、SymbMath、Theorist、VAXIMA、XPLORE 等等

(2) 矩阵运算软件

Caesy、CLAM、Ctrl-C、Mat/C、Matcalc、Matlab、MaTX、MEDAL、Gnu Octave、O-MATRIX、RlaB、XMATH、

(3) 科学数据可视化软件

AVS、Data Explorer、Explorer、Gnuplot、Interface Description、Language、PV-WAVE 等

(4) 数值、统计计算程序库

Netlib Numerical、Subroutines、Repository、Scilab、Statlib Statistical、Subroutine、Repository、IDL、Origin 6.0、Pro Fit 等

(5) 分子模拟计算软件

Amber、Charm、Discover、Gaussian 95 等

(6) 统计套装系统

Genstat、Glim、SAS、SPSS 等

(7) 地址信息系统

GRASS 等

其中的一些软件逐渐被引入我国，特别是在高等院校、研究所。这些软件的普及也在改变着我国科学计算、公式推导、处理数据、图形制作、动画设计以及文本编辑等的手段。对国外的几种主流科学计算软件，如美国 MathWorks 公司的 MATLAB、加拿大 Waterloo Maple Inc. 的 Maple、美国 MathSoft 公司的 Mathcad、美国 Wolfram 公司 Mathematica 以及 SPSS Inc. 的 SPSS 等我们组织了一批学者、硕士和博士编写了这一套图书。

《MATLAB 最新使用手册》针对当前 MATLAB 使用方面的教材比较多，而且 MATLAB 在我国今年呈比较普及的特点，是为满足读者需要一本放在计算机旁的函数使用手册而编写的，书中将最新版本的 MATLAB 的所有函数的语法结构、用法进行了详细的描述，并对绝大多数的函数举例说明。本书是经常使用 MATLAB 的用户必备的参考手册。

《掌握和精通 Maple》Maple 是世界上最著名的符号和代数运算的数学软件之一，包括 MATLAB、Mathcad 在内的许多科学计算软件均将 Maple 的“核”作为自己符号和代数

运算的插件。本书将引导读者初步掌握 Maple 语言的使用，读者也一定会被 Maple 强大的数学软件的功能所折服。

《掌握和精通 Mathcad2000》美国 MathSoft 公司也是世界上著名的科学计算软件厂商之一，它不但生产出 Mathcad 这一精品，而且推出了统计软件 S-PLUS。Mathcad 除了提供了常用的数学计算、数据统计分析的函数外，最大的特点是针对数学表达式的特殊性，提供了简单、方便的输入手段。其中的 Mathconnex 项目功能可将不同来源的数据集成在同一环境中。本书将会帮助读者尽快掌握这一数学工具。

《掌握和精通 Mathematica》Mathematica 也是最早进入我国的数学工具软件之一，相对其它数学软件，它的优势是符号运算和强大的图形处理功能。本书详细阐述 Mathematica 的用法和特点。

《掌握和精通 SPSS10》SPSS 是世界上著名的统计分析软件，适用于自然科学、社会科学各领域的统计分析，可完成对大量数据内在联系的提取。本书将 SPSS10 最常用的功能和使用方法介绍给大家。

这一套图书立足于软件的最新版本，图文并茂。古人曰“工欲善其事，必先利其器”，我们希望这套图书能够对读者朋友熟悉和掌握这集中科学计算软件提供帮助。

编 者

2001 年 1 月于北京

序

SPSS (Statistics Package for Social Science) for Windows 是世界上著名的统计分析软件,它是在SPSS/PC (for DOS)基础上发展起来的,适用于自然科学、社会科学各领域进行统计分析。

在 SPSS 中普遍使用 Windows 的窗口方式管理程序运行的全过程,通过对话框来实现各种命令参数的指定。只要掌握基本的 Windows 操作方法,粗通统计分析原理,就可以应用该软件得到具有专业水准的统计分析结果。

SPSS 具有强大的数据库互接功能,在其它数据库应用软件中建立的数据文件可以方便地导入到 SPSS 中,包括 dBASE、Excel、Foxpro、VF、MS Access、Text 等数据文件,利用 SPSS 进行统计分析。

SPSS 具有强大的图表功能,使用该软件不但可以得到数字式分析结果,还可以得到各种统计报表及形象直观的统计图形。可以实现统计结果与其它应用程序的共享。

SPSS 名为“社会科学软件包”,这是为了强调社会科学应用的一面,实际上它已经在社会科学、自然科学各个领域发挥着巨大作用。该软件可应用于经济学、生物学、矿物学、心理学、医疗卫生、体育、农业、商业、金融等各个部门。

本书共分 8 章。其中前 6 章为基础部分,主要介绍 SPSS 的基本概念和基本操作,通过前 6 章的学习读者可以基本掌握 SPSS 中的基本操作方法和步骤,为下面的统计分析作好准备;第 7 章介绍各种统计分析方法,包括描述性统计分析、相关分析、方差分析、回归分析、均值检验、聚类与判别分析、因子分析、非参数检验等。第 8 章介绍统计图形的生成与编辑。

本书由马春庭任主编,高萍、武海军、高智勇任副主编,马春庭负责全书的组织编写工作。参加本书编写工作的还有房立清、高秀峰、王希武、王蕾、谭业双、徐立新、吴开腾、赵红、王宁、方光、王立、张成卫等同志,对此表示诚挚的谢意。

由于篇幅的限制,目前国内较少使用的部分统计分析功能没有收入书中。鉴于本书的宗旨是介绍一种流行软件,而不是介绍统计分析原理和算法,因此对所涉及到的统计理论只作概念性介绍。若希望对统计分析原理有进一步的了解,请参考有关专业书籍。

编 者

2001 年 1 月 北京

目 录

总序

序

第 1 章 SPSS 简介	1
1.1 SPSS for Windows9. x(NT) 的特点	1
1.2 SPSS 对软硬件环境的要求.....	1
1.3 软件安装方法.....	2
第 2 章 认识 SPSS 系统	8
2.1 首次运行 SPSS.....	8
2.2 数据录入.....	10
2.2.1 定义变量的数据类型.....	10
2.2.2 录入数据.....	11
2.3 简单统计分析.....	12
2.4 统计分析的基本步骤.....	15
第 3 章 创建数据文件	16
3.1 常量与变量.....	16
3.1.1 常量.....	16
3.1.2 变量.....	16
3.1.3 变量属性的复制与粘贴.....	24
3.2 数据编辑与数据文件.....	25
3.2.1 数据编辑器.....	25
3.2.2 数据文件的结构.....	26
3.3 数据录入.....	27
3.3.1 数据录入方法.....	27
3.3.2 输入带有值标签的数据.....	29
3.4 数据文件的编辑.....	29
3.4.1 移动记录指针到指定观测量.....	30
3.4.2 定位到单元格.....	31
3.4.3 插入或删除一个变量.....	33
3.4.4 插入或删除一个观测量.....	35
3.4.5 数据剪切、复制与粘贴.....	35
3.5 由已存在的变量建立新变量.....	36
3.5.1 SPSS 的基本运算符.....	36
3.5.2 SPSS 中的基本函数.....	37
3.5.3 由已存在的变量建立新变量.....	44
3.6 数据文件的保存.....	48
3.6.1 保存为 SPSS 格式的数据文件.....	48
3.6.2 保存为其它格式的数据文件.....	49

第 4 章 从其它数据源读入数据文件	51
4.1 读入电子表格文件	51
4.2 读入数据库文件	56
4.3 读入文本型数据文件	65
第 5 章 SPSS10 的基本操作	72
5.1 系统运行管理方式	72
5.1.1 完全窗口菜单运行管理方式	72
5.1.2 程序运行管理方式	72
5.1.3 混合运行管理方式	72
5.1.4 程序自控运行方式	72
5.2 认识窗口	73
5.2.1 数据编辑窗口	73
5.2.2 输出窗口	73
5.2.3 语句窗口	74
5.2.4 脚本编辑窗口	75
5.2.5 统计图表编辑窗口	77
5.2.6 交互式图表编辑窗口	78
5.2.7 草稿输出窗口	79
5.3 输出窗口中的基本操作	80
5.3.1 查看输出对象	80
5.3.2 编辑输出对象	85
5.3.3 将输出对象应用于其它应用程序	87
5.3.4 保存输出	89
5.4 数据透视表的基本操作	91
5.4.1 改变数据表的外观	91
5.4.2 数据透视表的操作	93
5.5 脚本的建立及运行	100
5.5.1 运行脚本	100
5.5.2 创建脚本	104
5.5.3 自动脚本	107
5.6 程序自控运行方式	109
5.7 系统选项设置	116
5.7.1 General 选项卡	118
5.7.2 Viewer 选项卡	119
5.7.3 Draft Viewer 选项卡	120
5.7.4 Output Labels 选项卡	122
5.7.5 Charts 选项卡	122
5.7.6 Interactive 选项卡	124
5.7.7 Pivot Tables 选项卡	126

5.7.8	Data 选项卡	127
5.7.9	Currency 选项卡	128
5.7.10	Scripts 选项卡	130
第 6 章	数据文件的操作及变换	131
6.1	定义时间系列日期型变量	132
6.2	对观测量排序	134
6.3	数据文件转置	135
6.4	数据文件的合并	137
6.4.1	增加观测量	137
6.4.2	增加变量	140
6.5	分类汇总	144
6.5.1	分类汇总的概念	144
6.5.2	分类汇总的方法	145
6.6	拆分数据文件	148
6.7	选择观测量	151
6.8	对观测量作加权处理	156
6.9	设定随机数种子	158
6.10	对特定变量值计数	159
6.11	对变量值重新编码	162
6.11.1	编码成同类型的变量	163
6.11.2	编码成不同类型的变量	166
6.12	将变量值分类化	168
6.13	给观测量排秩	170
6.14	自动编码	173
6.15	时间系列的变换	175
6.16	替换缺失值	181
第 7 章	统计分析	187
7.1	基本概念和常用统计量	187
7.1.1	数理统计的基本概念	187
7.1.2	常用统计量	188
7.1.3	百分位的选择	188
7.2	统计分析过程一览	189
7.2.1	Reports 包含的分析功能	189
7.2.2	Descriptive Statistics 包含的分析功能	190
7.2.3	Compare Means 包含的分析功能	191
7.2.4	General Linear Model 包含的分析功能	192
7.2.5	Correlate 包含的分析功能	192
7.2.6	Regression 包含的分析功能	193
7.2.7	Loglinear 包含的分析功能	195

7.2.8	Classify 包含的分析功能	196
7.2.9	Data Reduction 包含的分析功能	196
7.2.10	Scale 包含的分析功能	197
7.2.11	Nonparametric Tests 包含的分析功能	197
7.2.12	Survival 包含的分析功能	198
7.2.13	Multiple Response 包含的分析功能	199
7.3	基本统计分析	199
7.3.1	一维频数分布表 Frequencies	200
7.3.2	求描述统计量的过程 Descriptives	203
7.3.3	考察数据的过程 Explore	205
7.3.4	多维频数分布表 Crosstabs	207
7.3.5	OLAP Cubes	211
7.3.6	对观测量列表	213
7.3.7	行形式输出报告	214
7.3.8	列形式输出报告	218
7.4	相关分析	219
7.4.1	相关分析的概念和相关系数	219
7.4.2	二元变量的相关分析	222
7.4.3	偏相关分析	224
7.4.4	距离分析	226
7.5	均值的比较与检验	229
7.5.1	Means 过程	230
7.5.2	独立样本 T 检验	232
7.5.3	配对样本 T 检验	234
7.6	方差分析	236
7.6.1	单因素方差分析	236
7.6.2	一个因变量的多因素方差分析	240
7.6.3	多元方差分析	244
7.6.4	重复试验的方差分析	247
7.6.5	方差成份分析	249
7.7	回归分析	251
7.7.1	线性回归	251
7.7.2	曲线估计	255
7.7.3	逻辑分析	257
7.7.4	概率分析	258
7.7.5	权重估计	260
7.7.6	最小二乘法	261
7.7.7	非线性回归	261
7.8	聚类与判别	264

7.8.1	快速样本聚类过程	265
7.8.2	分层聚类	268
7.8.3	判别分析	272
7.9	因子分析	277
7.10	非参数检验	283
7.10.1	卡方检验	284
7.10.2	二项分布检验	285
7.10.3	游程检验	286
7.10.4	一个样本的柯尔莫哥洛夫-斯米诺夫检验	287
7.10.5	两个独立样本检验	288
7.10.6	多个独立样本检验	289
7.10.7	两个相关样本检验	290
7.10.8	多个相关样本检验	291
第8章	统计图形的生成与编辑	292
8.1	交互式图形的生成与编辑	293
8.1.1	交互式图形的生成	294
8.1.2	交互式图形的编辑	296
8.2	统计图形的生成	302
8.2.1	条形图	302
8.2.2	线图	306
8.2.3	面积图	308
8.2.4	饼图	309
8.2.5	高低图	310
8.2.6	帕累托图	311
8.2.7	控制图	313
8.2.8	箱图	314
8.2.9	误差条图	316
8.2.10	散点图	317
8.2.11	直方图	318
8.2.12	P-P 正态概率图	319
8.2.13	Q-Q 正态概率图	321
8.2.14	系列图	323
8.2.15	时间系列图	325
8.3	统计图形的编辑	327

第 1 章 SPSS 简介

SPSS(Statistics Package for Social Science)是一个适用于自然科学、社会科学各领域的统计分析软件包，是世界上流行的统计分析软件。近年来，我国的医疗卫生、体育、经济各领域的科研工作者开始使用该软件进行研究工作，其应用前景逐步被人们所认识。

1.1 SPSS for Windows9.x(NT)的特点

1.SPSS for Windows 是在 SPSS/PC (for DOS) 的基础上发展起来的，目前版本号为 SPSS10.0。它具有其它 Windows 软件的共同特点，便于广大 Windows 用户使用。

2.在 SPSS 中，除了数据输入工作需要使用键盘来完成外，其它如命令语句的书写、语句参数的选取等绝大部分工作都可通过操作“对话框”自动完成。因此，用户无需花过多时间和精力记忆大量的命令、过程、选择项等。而对于熟悉 SPSS 语言的老用户，也可以直接书写出程序语句，然后提交系统运行。

3.具有第四代语言的特点，只要通过菜单的选择、对话框的操作告诉系统要做什么，无需告诉怎样做。用户只需粗通统计分析原理，即可得到统计分析结果。

4.分析方法丰富，提供了从简单描述统计分析到多因素的统计分析方法。可以满足绝大部分统计分析工作的需要。

5.具有完善的数据转换接口。其它软件生成的数据文件，可方便地转换成 SPSS 数据文件。这些数据文件包括 dBASE 数据文件、Excel 电子表格、FoxPro 数据库文件、MS Access 数据文件、用文本编辑软件生成的 ASCII 码数据文件等。

1.2 SPSS 对软硬件环境的要求

- 操作系统为 Windows 95, Windows 98, Windows NT 4.0, 或 Windows 2000。SPSS 目前没有汉化版本，若需要在程序中由用户输入的部分使用汉字，则应安装中文版的操作系统。
- Pentium 系列处理器（运行速度 90MHz 及以上）。
- 16兆内存。推荐 64兆及以上。
- 典型安装需要 80兆以上的硬盘空间。
- 安装时需配备有 CD-ROM 驱动器。
- 显示器要求有 800×600 像素及以上分辨率。
- 若需与 SPSS 服务器实现连接，需配备基于 TCP/IP 网络传输协议的网络适配器。

1.3 软件安装方法

将安装盘置入CD-ROM驱动器, 自动播放功能将启动“运行安装程序”菜单, 选择“Install

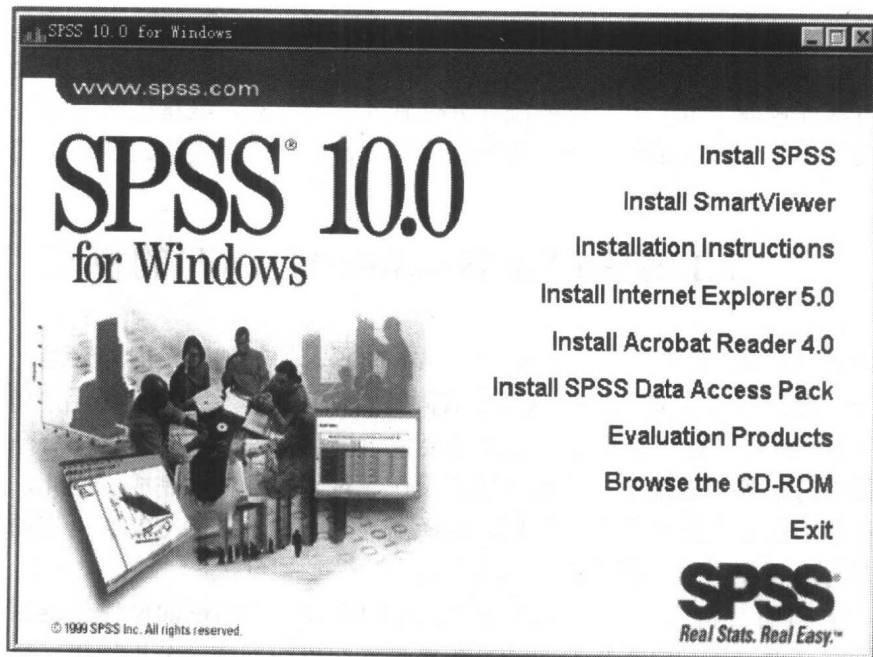


图1.1 SPSS10.0 安装启动画面

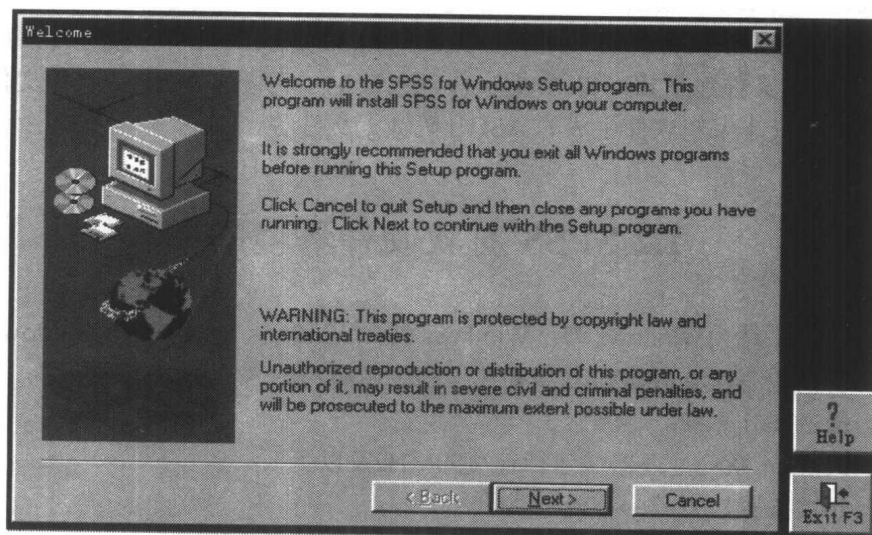


图1.2 SPSS 安装欢迎界面

SPSS”，如图1.1所示。

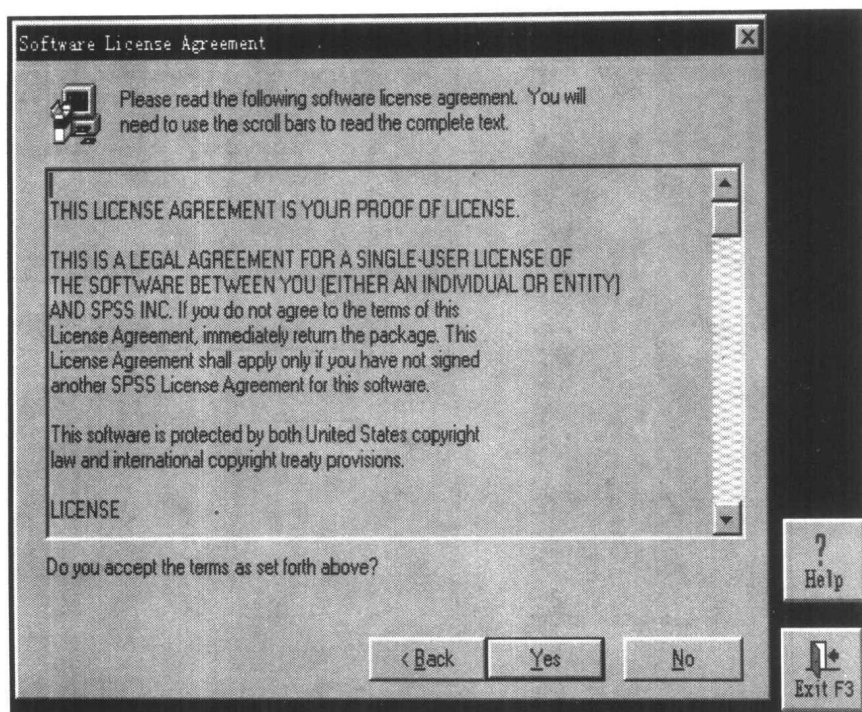


图1.3 SPSS10 安装许可协议

您也可以按下列步骤运行安装程序：

1. 将安装盘置入CD-ROM驱动器；
 2. 单击“开始”，再单击“运行”；
 3. 在“运行”对话框中键入“d:\setup”（这里假定D盘为CD-ROM驱动器）。
- 接着是安装的欢迎界面（如图1.2所示）和安装许可协议界面（如图1.3所示）。

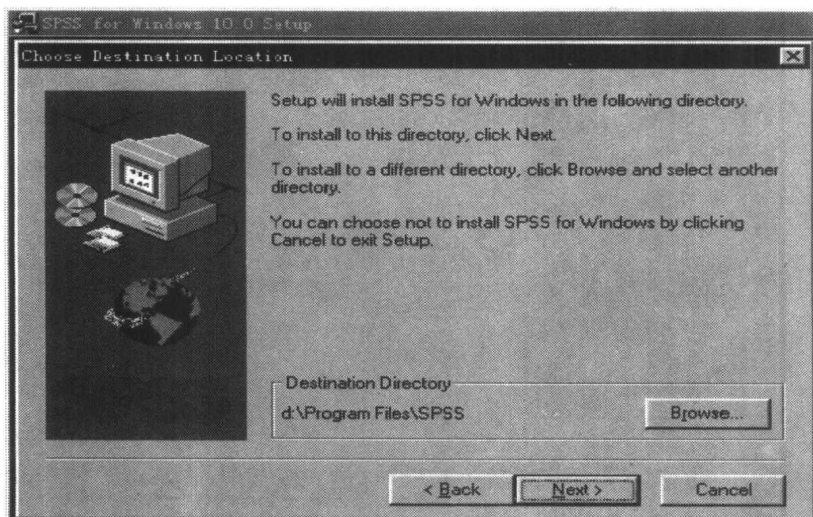


图1.4 SPSS10 安装路径设置界面

通过图1.4所示的路径设置可以选定SPSS10安装的目录。

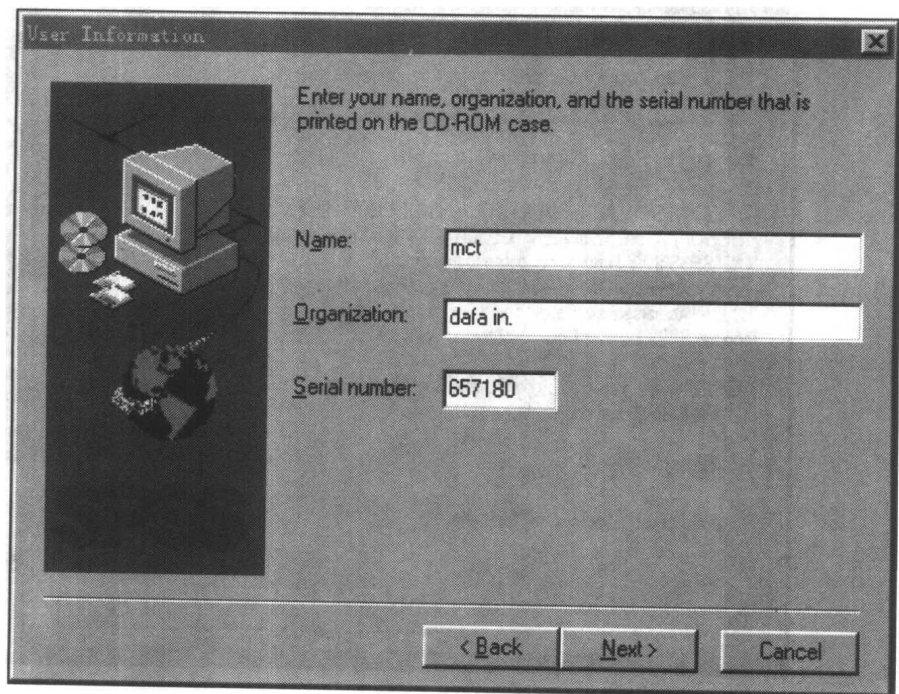


图1.5 用户信息输入界面

选定安装路径后进入用户信息输入界面，如图1.5所示，在本界面中输入用户名称、公司或厂商的名称以及序列号。

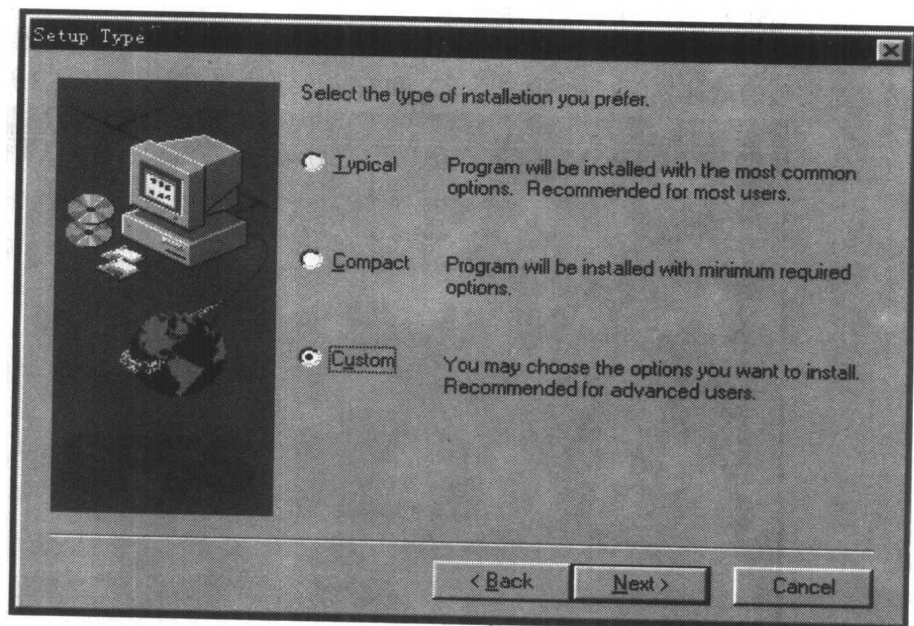


图1.6 安装类型设置界面

针对用户对SPSS熟悉程度的不同，系统提供了三种安装类型（典型安装、最小化安装、用户自定义安装），如图1.6所示：

典型安装 除了运行SPSS所需的基本文件外，还包括Sample data files（提供了建立数据文件的示例）、Help Files（用户帮助文件）、TableLooks（操作与保存数据表格式的工具），Statistics Coach（包含对部分基本类型的统计分析的指导范例）。

最小安装 仅包含运行SPSS所需的最少文件。

用户自定义安装 用户可选择安装的部件包含典型安装的全部组件，此外还可以选择安装下列组件：

- **Basic Scripting**。安装Sax Basic，以及软件中实现OLE自动化必须的文件。
- **Production Facility**。提供用自动化方式运行程序的能力。无需用户参与，程序自动运行直到执行完最后一个命令语句。如果您经常面对大量重复且耗时的分析任务，此时自动运行模式显得尤其方便。
- **Syntax Guides**。安装盘中包含的语法参考指导（PDF文件格式）。阅读PDF格式文件之前需要安装Adobe Acrobat Reader阅读器（随安装盘附带，参见图1.1所示）。

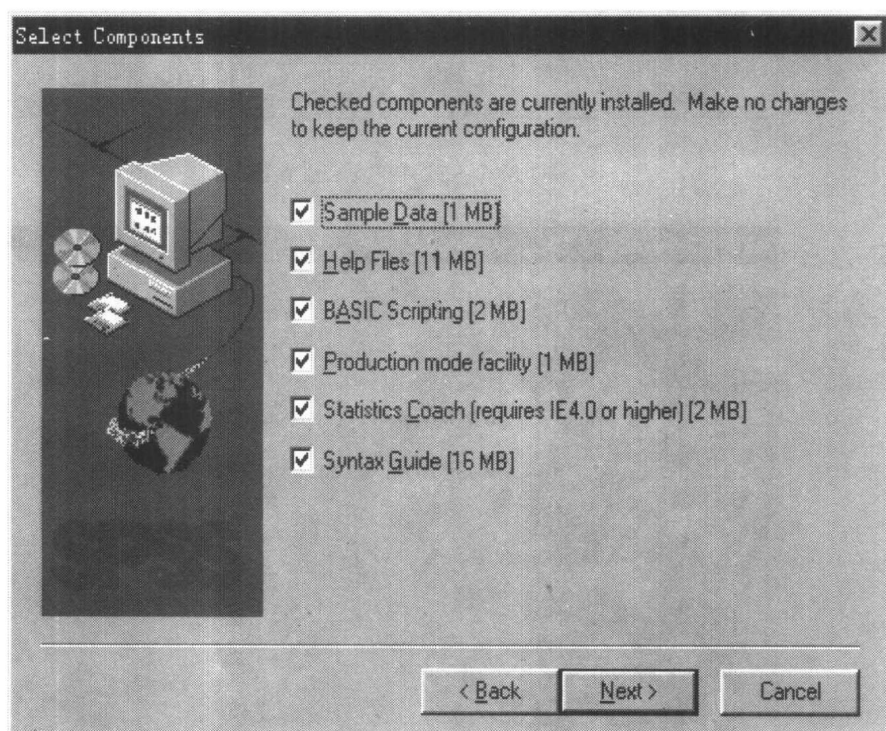


图1.7 SPSS10 自定义安装选项窗口

如果选定了用户自定义安装的方式，接着出现的界面如图1.7所示，用户根据熟悉程度和计算机的配置、实际需要进行选择。

SPSS10可以安装在单个个人计算机上，也可安装在联网的计算机上以实现资源共享，在安装过程中用户可自由选择，如图1.8所示。

接着出现的是SPSS10的选项窗口,通过本窗口可以在系统中加入两种模型:回归模型(Regression Models)和高级模型(Advanced Models),如图1.9所示。

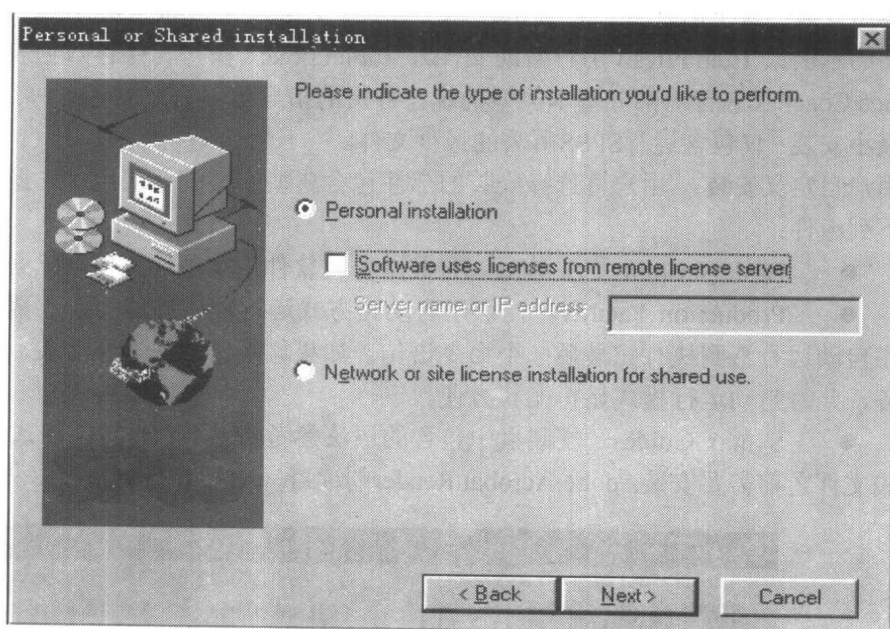


图1.8 个人或共享安装设置窗口

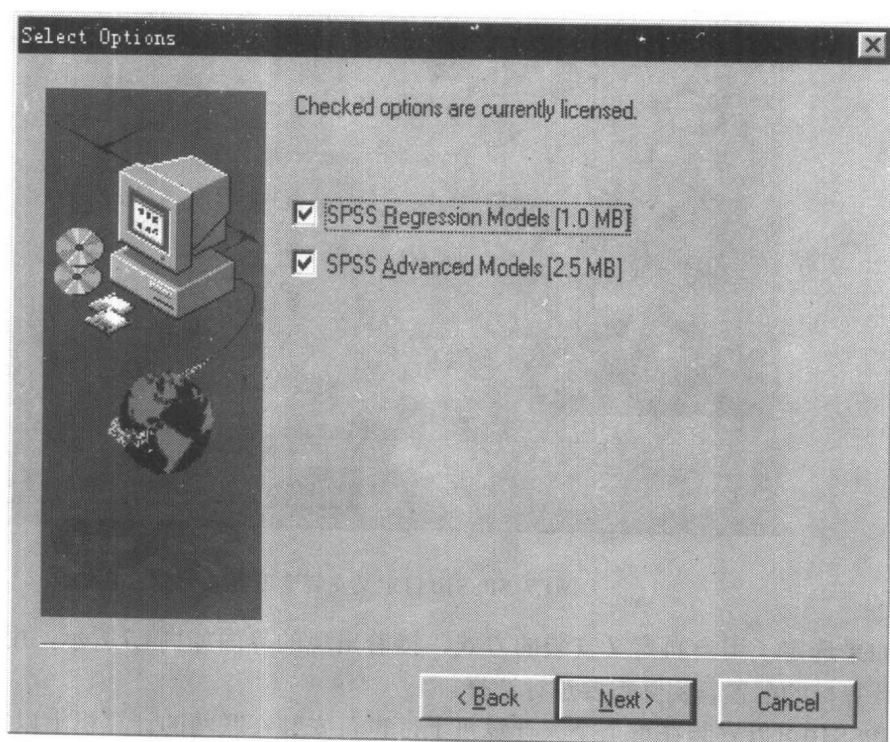


图1.9 SPSS10 选项窗口