



刘先勇 袁长迎 段宝福 周方洁 编著

SPSS 10.0

统计分析软件与应用

工程师工具软件应用系列

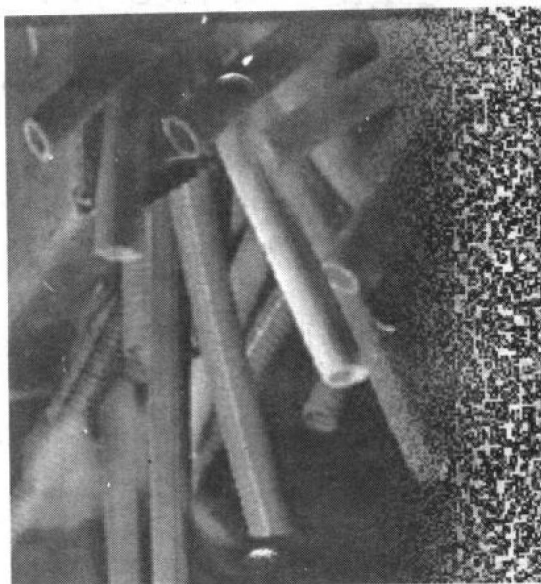
国防工业出版社

C819

7

国防工业出版社

www.ndip.com.cn



北方工业大学图书馆



00506344

刘先勇 袁长迎 段宝福 周方洁 编著

SPSS 10.0

统计分析软件与应用

工程师工具软件应用系列

内 容 简 介

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、社会科学各领域进行统计分析。全书共分 17 章,主要介绍 SPSS 的基本概念、基本操作与统计分析方法,其中包括数据文件的建立与编辑、数据操作、报表生成、描述性统计分析、均值比较与检验、方差分析、相关分析、回归分析、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 的初学者及各类统计分析人员阅读使用。

图书在版编目(CIP)数据

SPSS 10.0 统计分析软件与应用/刘先勇等编著.

北京:国防工业出版社,2002.1

(工程师工具软件应用系列)

ISBN 7-118-02751-0

I . S... II . 刘... III . 统计分析—软件包,SPSS 10.0 IV . C819

中国版本图书馆 CIP 数据核字(2001)第 086271 号

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号)

(邮政编码 100044)

北京奥隆印刷厂印刷

新华书店经售

*

开本 787×1092 1/16 印张 27 $\frac{1}{4}$ 633 千字

2002 年 1 月第 1 版 2002 年 1 月北京第 1 次印刷

印数:1—4000 册 定价:38.00 元

(本书如有印装错误,我社负责调换)

前 言

随着信息技术的普及和发展,人们对计算机应用的要求已不再局限于一般的文字处理,利用计算机进行数据统计分析、信息管理已经是许多企业和科研单位进行生产、管理及科学研究的基本手段。SPSS 统计分析软件包(Statistics Package for Social Science)对计算机在数据统计方面的应用提供了强有力的支持。

SPSS 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。鉴于 SPSS 在我国自然科学和社会科学方面的应用还刚刚起步,我们编写了本书,以帮助大家了解和应用 SPSS。本书重点讲述了 SPSS 的常用功能,通俗易懂并列举了大量的应用实例。既可供熟悉该软件的专业人士作为参考,也可作为初学者理想的实用教材。

SPSS 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,包括 dBASE、Excel、FoxPro、VF、MS Access、Text 等的数据库文件,都可以方便地导入到 SPSS 中进行统计分析。

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。

本书包括三大部分,共分 17 章。其中第一部分包括前 6 章,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 的基本操作方法和步骤,为下面的统计分析作好准备。这 6 章为基础部分,建议初学者通读;第二部分包括第 7 章至第 16 章,介绍各种统计分析方法,包括报表生成、描述性统计分析、均值比较与检验、方差分析、相关分析、回归分析、聚类与判别分析、因子分析、非参数检验等;第三部分为第 17 章,重点介绍了统计图形的生成与编辑方法。

本书由刘先勇、袁长迎、段宝福、周方洁合作编写。西南科技大学数学系徐明民教授审阅了全稿,并提出了具体的修改意见。参加本书编写工作的还有闫旭骞博士、朱景和博士等,在此一并表示诚挚的谢意。

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。

由于时间仓促,水平有限,书中错误和不妥之处在所难免,敬请广大读者批评、指正。我们的电子邮箱为:ligongboshi@sina.com、baofu@sina.com。

编 者

97 14

目 录

第 1 章 SPSS 简介	1
1.1 SPSS for Windows 的基本特点	1
1.2 SPSS for Windows 的运行环境	1
1.3 SPSS 10.0 for Windows 的安装	2
1.4 SPSS 10.0 for Windows 的启动和退出	6
1.4.1 SPSS 的启动	6
1.4.2 SPSS 的退出	7
1.5 SPSS 的帮助系统	7
1.5.1 联机帮助	7
1.5.2 操作指南	9
1.5.3 语法指南	9
1.5.4 统计分析指南	10
1.5.5 在线帮助	11
第 2 章 使用 SPSS 的一个简单实例	12
2.1 首次运行 SPSS	12
2.2 数据录入	14
2.2.1 定义变量的数据类型	14
2.2.2 录入数据	15
2.3 简单统计分析	15
2.4 统计分析的基本步骤	17
第 3 章 创建数据文件	18
3.1 常量与变量	18
3.1.1 常量	18
3.1.2 变量	18
3.1.3 变量属性的复制与粘贴	25
3.2 数据编辑与数据文件	26
3.2.1 数据编辑器	26
3.2.2 数据文件的结构	27
3.3 数据录入	28
3.3.1 数据录入方法	28
3.3.2 输入带有值标签的数据	29
3.4 数据文件的编辑	30

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(查找指定变量中的指定数值)	31
3.4.3	插入或删除一个变量	32
3.4.4	插入或删除一个观测量	34
3.4.5	数据剪切、复制与粘贴	34
3.5	由已存在的变量建立新变量	35
3.5.1	SPSS 的基本运算符	35
3.5.2	SPSS 中的基本函数	35
3.5.3	由已存在的变量建立新变量	36
3.6	数据文件的保存	39
3.6.1	保存为 SPSS 格式的数据文件	39
3.6.2	保存为其它格式的数据文件	39
第 4 章	从其它数据源读入数据文件	41
4.1	读入电子表格文件	41
4.2	读入数据库文件	45
4.3	读入文本型数据文件	52
第 5 章	SPSS 的基本操作	58
5.1	系统运行管理方式	58
5.1.1	完全窗口菜单运行管理方式	58
5.1.2	程序运行管理方式	58
5.1.3	混合运行管理方式	58
5.1.4	程序自控运行方式(SPSS Production Facility)	58
5.2	认识窗口	59
5.2.1	数据编辑窗口(Data Editor)	59
5.2.2	输出窗口(Output Viewer)	59
5.2.3	语句窗口(Syntax Editor)	60
5.2.4	脚本编辑窗口(Script)	61
5.2.5	统计图表编辑窗口(Chart Editor)	62
5.2.6	交互式图表编辑窗口(Interactive Chart)	63
5.2.7	草稿输出窗口(Draft)	64
5.3	输出窗口中的基本操作	65
5.3.1	查看输出对象	65
5.3.2	编辑输出对象	69
5.3.3	将输出对象应用于其它应用程序	72
5.3.4	保存输出	73
5.4	数据透视表的基本操作	75
5.4.1	改变数据表的外观	75
5.4.2	数据透视表的操作	77
5.5	创建及运行脚本	83

5.5.1 运行脚本	83
5.5.2 创建脚本	84
5.5.3 自动脚本(Autoscript)	88
5.6 程序自控运行方式(Production Facility)	90
5.7 系统选项设置	95
5.7.1 General 选项卡	95
5.7.2 Viewer 选项卡	97
5.7.3 Draft Viewer 选项卡	99
5.7.4 Output Labels 选项卡	100
5.7.5 Charts 选项卡	100
5.7.6 Interactive 选项卡	102
5.7.7 Pivot Tables 选项卡	104
5.7.8 Data 选项卡	105
5.7.9 Currency 选项卡	106
5.7.10 Scripts 选项卡	107
第 6 章 数据文件的操作及变换	109
6.1 定义时间系列日期型变量	110
6.2 对观测量排序	111
6.3 数据文件转置	112
6.4 数据文件的合并	114
6.4.1 增加观测量(Add Cases)	114
6.4.2 增加变量(Add Variables)	117
6.5 分类汇总	120
6.5.1 分类汇总的概念	120
6.5.2 分类汇总的方法	121
6.6 拆分数据文件	124
6.7 选择观测量	127
6.8 对观测量作加权处理	131
6.9 设定随机数种子	132
6.10 对特定变量值计数	133
6.11 对变量值重新编码	136
6.11.1 编码成同类型的变量	137
6.11.2 编码成不同类型的变量	140
6.12 将变量值离散化	142
6.13 给观测量排秩	143
6.14 自动编码	147
6.15 时间系列的变换	148
6.16 替换缺失值	154
第 7 章 统计分析功能一览	159

7.1 Reports 包含的分析功能	160
7.1.1 OLAP Cubes	160
7.1.2 Case Summaries	160
7.1.3 Report Summaries in Rows	160
7.1.4 Report Summaries in Columns	160
7.2 Descriptive Statistics 包含的分析功能	160
7.2.1 Frequencies	161
7.2.2 Descriptives	161
7.2.3 Explore	161
7.2.4 Crosstabs	161
7.3 Compare Means 包含的分析功能	162
7.3.1 Means	162
7.3.2 One – Sample T Test	162
7.3.3 Independent – Samples T Test	162
7.3.4 Paired – Samples T Test	163
7.3.5 One – Way ANOVA	163
7.4 General Linear Model 包含的分析功能	163
7.4.1 Univariate	163
7.4.2 Multivariate	163
7.4.3 Repeated Measures	164
7.4.4 Variance Components	164
7.5 Correlate 包含的分析功能	164
7.5.1 Bivariate	164
7.5.2 Partial	165
7.5.3 Distances	165
7.6 Regression 包含的分析功能	165
7.6.1 Linear	165
7.6.2 Curve Estimation	166
7.6.3 Binary Logistic	166
7.6.4 Multinomial Logistic	166
7.6.5 Ordinal	166
7.6.6 Probit	167
7.6.7 Nonlinear	167
7.6.8 Weight Estimation	167
7.6.9 2 – Stage Least Squares	167
7.7 Loglinear 包含的分析功能	167
7.7.1 General	168
7.7.2 Logit	168
7.7.3 Model Selection	168

7.8 Classify 包含的分析功能	168
7.8.1 K – Means Cluster	168
7.8.2 Hierarchical Cluster	169
7.8.3 Discriminant	169
7.9 Data Reduction 包含的分析功能	169
7.10 Scale 包含的分析功能	170
7.10.1 Reliability Analysis	170
7.10.2 Multidimensinal Scaling	170
7.11 Nonparametric Tests 包含的分析功能	170
7.11.1 Chi – Square	170
7.11.2 Binomial	171
7.11.3 Runs	171
7.11.4 1 – Sample K – S	171
7.11.5 2 Independent Samples	171
7.11.6 K Independent Samples	171
7.11.7 2 Related Samples	172
7.11.8 K Related Samples	172
7.12 Survival 包含的分析功能	172
7.12.1 Life Tables	172
7.12.2 Kaplan – Meier	172
7.12.3 Cox Regression	173
7.12.4 Cox w/ Time – Deep COV	173
7.13 Multiple Response 包含的分析功能	173
7.13.1 Define Set	173
7.13.2 Frequencies	174
7.13.3 Crosstabs	174
第 8 章 统计报表	175
8.1 分层摘要报告表	175
8.1.1 对话框介绍	175
8.1.2 应用举例	177
8.2 对观测量列表	178
8.2.1 对话框介绍	178
8.2.2 应用举例	180
8.3 行形式输出报告	181
8.4 列形式输出报告	186
第 9 章 描述性统计分析	189
9.1 一维频数分布分析	189
9.1.1 对话框介绍	189
9.1.2 应用举例	192

9.2 描述性统计分析	195
9.2.1 对话框介绍	195
9.2.2 应用举例	196
9.3 探索分析(Explore)	197
9.3.1 对话框介绍	197
9.3.2 应用举例	200
9.4 多维频数分布表	205
9.4.1 对话框介绍	206
9.4.2 应用举例	209
第 10 章 均值比较与检验	212
10.1 平均数检验	212
10.1.1 对话框介绍	213
10.1.2 应用举例	215
10.2 单一样本 T 检验	216
10.2.1 对话框介绍	216
10.2.2 应用举例	217
10.3 独立样本 T 检验	218
10.3.1 对话框介绍	218
10.3.2 应用举例	220
10.4 配对样本 T 检验	221
10.4.1 对话框介绍	222
10.4.2 应用举例	223
第 11 章 方差分析	225
11.1 单因素方差分析	225
11.1.1 对话框介绍	225
11.1.2 应用举例	228
11.2 一个因变量的多因素方差分析	232
11.2.1 对话框介绍	232
11.2.2 应用举例	238
11.3 多因变量方差分析	242
11.3.1 对话框介绍	243
11.3.2 应用举例	245
11.4 重复测量方差分析	247
11.4.1 对话框介绍	247
11.4.2 应用举例	250
11.5 方差成分分析	254
11.5.1 对话框介绍	254
11.5.2 应用举例	256
第 12 章 相关分析	258

12.1 相关分析的概念及相关系数·····	258
12.2 二元变量的相关分析(Bivariate) ·····	260
12.2.1 对话框介绍·····	261
12.2.2 应用举例·····	262
12.3 偏相关分析·····	264
12.3.1 对话框介绍·····	265
12.3.2 应用举例·····	266
12.4 距离分析·····	267
12.4.1 对话框介绍·····	268
12.4.2 应用举例·····	274
第 13 章 回归分析 ·····	276
13.1 线性回归·····	276
13.1.1 一元线性回归基本原理·····	277
13.1.2 多元线性回归基本原理·····	279
13.1.3 对话框介绍·····	280
13.1.4 应用举例·····	286
13.2 曲线估计·····	294
13.2.1 对话框介绍·····	294
13.2.2 应用举例·····	296
13.3 非线性回归·····	298
13.3.1 对话框介绍·····	298
13.3.2 应用举例·····	302
13.4 逻辑回归·····	305
13.4.1 对话框介绍·····	305
13.4.2 应用举例·····	309
13.5 概率分析·····	312
13.6 权重估计·····	313
13.7 最小二乘法·····	314
第 14 章 聚类与判别 ·····	315
14.1 快速样本聚类过程·····	315
14.1.1 基本概念·····	315
14.1.2 对话框介绍·····	316
14.1.3 应用实例·····	318
14.2 分层聚类·····	320
14.2.1 基本概念·····	320
14.2.2 对话框介绍·····	321
14.2.3 应用实例·····	326
14.3 判别分析·····	329
14.3.1 基本概念·····	329

14.3.2 对话框介绍·····	330
14.3.3 应用实例·····	334
第 15 章 因子分析 ·····	339
15.1 基本概念·····	339
15.2 对话框介绍·····	340
15.3 应用举例·····	346
第 16 章 非参数检验 ·····	351
16.1 卡方检验·····	351
16.1.1 对话框介绍·····	351
16.1.2 应用实例·····	353
16.2 二项分布检验·····	354
16.2.1 对话框介绍·····	354
16.2.2 应用实例·····	355
16.3 游程检验·····	356
16.3.1 对话框介绍·····	356
16.3.2 应用实例·····	356
16.4 单样本 K-S 检验 ·····	357
16.4.1 对话框介绍·····	358
16.4.2 应用实例·····	358
16.5 两个独立样本检验·····	359
16.5.1 对话框介绍·····	359
16.5.2 应用实例·····	361
16.6 多个独立样本检验·····	362
16.6.1 对话框介绍·····	362
16.6.2 应用实例·····	363
16.7 两个相关样本检验·····	364
16.7.1 对话框介绍·····	364
16.7.2 应用实例·····	365
16.8 多个相关样本检验·····	366
16.8.1 对话框介绍·····	366
16.8.2 应用实例·····	367
第 17 章 统计图形的生成与编辑 ·····	368
17.1 交互式图形的生成与编辑·····	369
17.1.1 交互式图形的生成·····	370
17.1.2 交互式图形的编辑·····	372
17.2 统计图形的生成·····	378
17.2.1 条形图·····	378
17.2.2 线图·····	382
17.2.3 面积图·····	383

17.2.4	饼图	384
17.2.5	高低图	384
17.2.6	帕累托图	386
17.2.7	控制图	387
17.2.8	箱图	388
17.2.9	误差条图	389
17.2.10	散点图	390
17.2.11	直方图	390
17.2.12	P - P 正态概率图	391
17.2.13	Q - Q 正态概率图	393
17.2.14	系列图	394
17.2.15	时间系列图	395
17.3	统计图形的编辑	398
附录 I	SPSS 各版本功能对照表	405
附录 II	SPSS 统计功能速查	407
附录 III	SPSS 中的函数	411
附录 IV	常用统计词汇英汉对照表	420

第 1 章 SPSS 简介

SPSS(Statistics Package for Social Science)是适用于自然科学、社会科学各领域的统计分析软件包,是目前非常流行的统计分析软件。SPSS 公司自 1972 年创立以来,不断推出新的 SPSS 版本,从最初的 SPSS/PC⁺ for DOS 到 SPSS 6.x、7.0、8.0、9.0、10.0 for Windows。随着版本的不断提高,其功能越来越强大,操作使用也越来越简单。目前的 SPSS 版本,不仅可以实现多种统计分析,还能通过清晰直观的表格及数十种生动形象的二维、三维图形将分析结果表达出来。

本章介绍 SPSS 的基本概况、软件运行环境、安装使用方法,以及如何在使用中获得帮助。

1.1 SPSS for Windows 的基本特点

(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 Windows 软件的共同特点,便于熟悉 Windows 操作方法的广大用户使用。

(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”来完成。只需要简单地点击几下鼠标,就可以实现所需的统计分析功能。用户无需花太多时间记忆大量的命令、过程、选择项等。同时 SPSS 还兼容了早期版本的使用方法,对于熟悉 SPSS 语言的老用户,可以使用内置的 Sax Basic 语言直接书写程序语句,提交系统运行。

(3) 具有第四代语言的特点,只需通过菜单的选择、对话框的操作告诉系统要做什么,无需告诉系统怎样做。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。

(6) 支持 OLE 和 ActiveX 技术,可以实现与其它应用程序的通信。

1.2 SPSS for Windows 的运行环境

(1) 操作系统为 Windows 95、Windows 98、Windows 2000 或 Windows NT 4.0,推荐使用 Windows 2000 操作系统。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或汉字平台。

(2) 586DX 或以上,IBM®兼容 PC。

(3) 最小 32MB 内存,32MB 虚拟内存。

(4) 所需硬盘空间:最小化安装需要 61MB 硬盘空间,典型安装需要 80MB 硬盘空间,自定义安装最大需要 95MB 硬盘空间。

(5) 安装时需配备有 CD-ROM 驱动器。

(6) SVGA 显示器。

(7) Windows 兼容鼠标。

1.3 SPSS 10.0 for Windows 的安装

1) 确认计算机上安装了 Windows 95、Windows 98、Windows 2000 或 Windows NT 4.0 操作系统。

2) 将安装盘置入 CD-ROM 驱动器,自动播放功能将打开“SPSS 10.0 for Windows”安装动画面,如图 1-1 所示。



图 1-1 SPSS 10.0 安装启动画面

3) 选择“Install SPSS”开始安装 SPSS 10.0。

4) 接下来进入 SPSS 安装欢迎界面,单击“Next”进入下一步。

5) 打开 SPSS 安装许可协议,如图 1-2 所示。请仔细阅读该协议中的条款,并单击“Next”进入下一步。

6) 在如图 1-3 所示的对话框中指定安装 SPSS 的路径。系统默认的路径是“C: \ Program Files \ SPSS”。如果要改变安装路径,单击“Browse”图标按钮,在打开的子对话框中重新输入新的安装路径,并单击“OK”返回图 1-3 的对话框。单击“Next”进入下一步。

7) 在接下来的对话框中输入有关用户信息,包括用户名、公司或单位名称、产品序列

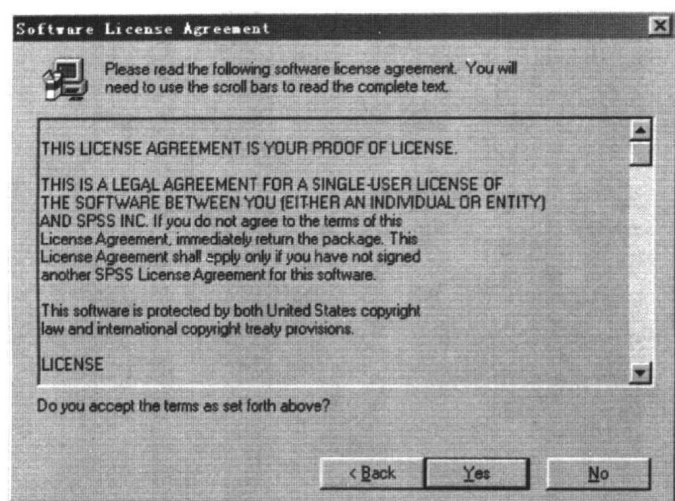


图 1-2 SPSS 安装许可协议

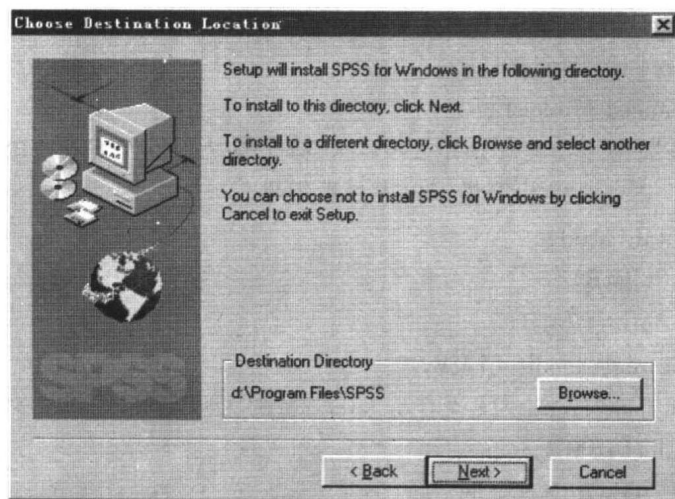


图 1-3 设置安装路径

号。单击“Next”进入下一步。

8) 弹出如图 1-4 所示的对话框,要求指定安装类型。系统提供了三种不同安装类型:最小化安装(Compact)、典型安装(Typical)、自定义安装(Custom)。根据实际需要选择一种安装类型。

(1) 最小化安装只安装基本统计模块。约需要 61MB 硬盘空间。

(2) 典型安装对大多数用户来说最常用的模块,需要 80MB 硬盘空间。建议初学者选择典型安装。典型安装包括以下基本组件:

- ① Sample Data
- ② Help Files
- ③ Basic Scripting
- ④ Production mode facility

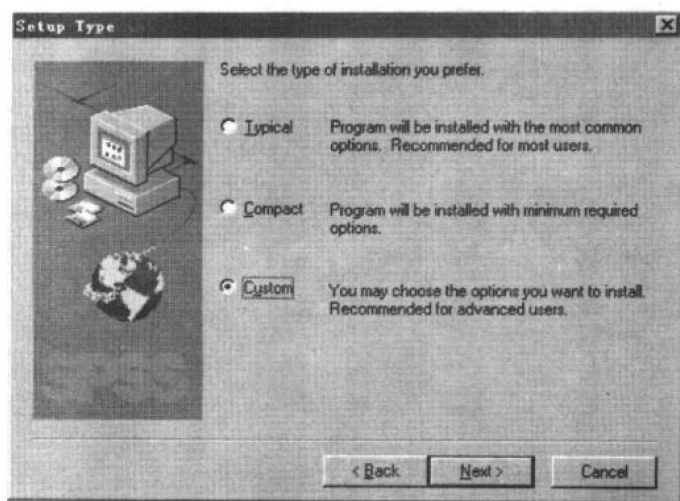


图 1-4 指定安装类型

⑤ Statistics Coach(require IE 4.0 or later)

⑥ SPSS Regression Models(可选)

⑦ SPSS Advanced Models(可选)

(3) 自定义安装由用户自行指定要安装的功能模块。已经熟悉 SPSS 软件使用的用户可以选择该项。自定义安装提供了以下可选组件：

① Sample Data(1MB)

② Help Files(11MB)

③ BASIC Scripting(2MB)

④ Production mode facility(1MB)

⑤ Statistics Coach(require IE 4.0 or later 2MB)

⑥ Syntax Guide(16MB)

全部安装需要 95MB 硬盘空间。

9) 在对话框中让用户选择安装的类型：个人用户安装(Personal installation)或共享安装(Network or site licence installation for shared use)。SPSS 既可以安装为单机使用，也可以安装在网络计算机上供多用户使用并实现资源共享。

10) 接下来的对话框让用户选择是否要安装高级统计模块(SPSS Advanced Models 2.5MB)和专业统计模块(SPSS Regression Models 1MB)。如果硬盘空间允许，建议读者选中这两个项目。否则在 SPSS 的菜单中无法调用与这两个模块有关的统计分析功能。

11) 在收集完安装设置的有关信息后，将弹出如图 1~5 所示的对话框，其中列出了用户所作的各种设置的具体内容，如安装路径、将要安装的组件、安装类型、安装所需的磁盘空间、用户名、产品序列号等。如果需要对前面的设置作更改，则单击“Back”按钮回到前面的对话框重新进行设置。单击“Next”图标按钮确认所作设置，则系统开始复制有关文件，将 SPSS 安装在硬盘中。

12) 根据机器的配置不同，安装过程大约需要 5~10 分钟时间。如图 1-6 所示为安