



高等学校教材

# SPSS for Windows 高级应用教程

阮桂海 蔡建瓴 佟福玲 等编著

大专院校计算机统计分析教程

965  
P34  
898



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

URL: <http://www.phei.com.cn>

高等学校教材

# SPSS for Windows 高级应用教程

阮桂海 蔡建瓴 佟福玲 等编著

电子工业出版社

**Publishing House of Electronics Industry**

北京·BEIJING

## 内 容 简 介

本书注重最常用的高级统计和专业统计,编著了 Hierarchical Loglinear Models(分层对数线性分析模型)、Discriminant(判别分析)、Factor(因素分析)、Cluster(聚类分析)、Quick Cluster(大文件快速聚类分析)等的应用。

这些统计应用不但有命令、对话框的使用举例,而且极为详尽地启发读者巧用 SPSS 中的统计技术;对各种各样的输出结果进行准确无误的分析与解释;根据定量分析写出高质量的、决策性的论文。诸如此类的内容,尽管以前已编著了四册,但这一册中的内容则是前所未有的。此外,书中还编著了 Explore(数据错误的探测与图示)、Multiple Linear Regression(多元线性回归分析)、Multiple Response(多项选择的统计分析)以及 SPSS for Windows 快速入门知识等等。这是 SPSS V7.0 版本新增的命令或功能,根据这些知识引导,任凭读者有多大多难的课题研究,亦能迎刃而解。

此书可作为大专院校计算机统计分析的教程,也是社会科学、人文地理、心理学、医学、市场调查研究及统计学家必不可少的应用指南。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。  
版权所有,翻版必究。

丛 书 名: 高等学校教材  
书 名: SPSS for Windows 高级应用教程  
编 著 者: 阮桂海 蔡建瓴 佟福玲等  
策 划: 胡毓坚  
责任编辑: 张燕虹  
排版制作: 北京华旗资讯科技发展有限公司  
印 刷 者: 北京京安达明印刷厂  
出版发行: 电子工业出版社 URL: <http://www.phei.com.cn>  
北京市海淀区万寿路 173 信箱 邮编 100036  
经 销: 各地新华书店  
开 本: 787×1092 1/16 印张: 20.50 字数: 525 千字  
版 次: 1998 年 11 月第 1 版 1998 年 11 月第 1 次印刷  
书 号: ISBN 7-5053-4955-4  
G·390  
定 价: 26.00 元



凡购买电子工业出版社的图书,如有缺页、倒页、脱页者,请向购买书店调换。  
若购买书店售缺,请与本社发行部联系调换。电话 68279077

# 前 言

SPSS 公司创建于 1966 年,至今已发展到 SPSS V7. X for Windows 版本。它以命令简单易记、界面丰富清晰、技术新颖超群而深受用户青睐。目前,使用 SPSS 的用户遍及全世界。在我国,SPSS、SAS 等知名度已与 FORTRAN 语言、C 语言等可谓不相上下。

但是由于版权的缘故,有关 SPSS 的技术运用手册未能公开面市;北京大学人口所授权翻译过 SPSS V6.0 Base 技术教程,但印数甚少且尚未公开发售,也形成了“杯水车薪”的憾事。值此,我们北京大学社会学系计算机室在教学过程中,应急编著了 SPSS 6.0 高级应用教程(上)。这次又应急编著了姊妹篇——SPSS 高级应用教程,作为中册出版,故以本书封面的《SPSS for Windows 高级应用教程》命名。

按照 SPSS for Windows 的约定,凡是 DOS 环境支持的 SPSS /PC<sup>+</sup> V4.0 及其低版本,统称为 SPSS/PC<sup>+</sup>;凡是 Windows 环境支持的 SPSS V5.0 及以后版本,统称为 SPSS。

本书共有十二章,着眼于统计分析和结果剖析,把笔者十年的实践经验和苦苦探索向诚挚的读者毫无保留地奉献出来。

本书第一、二章叙述 SPSS for Windows 快速入门知识,是专为 SPSS 新读者而开设的重要一课。对于只有 DOS 下的 SPSS/PC<sup>+</sup> 知识,而无 Windows 下的 SPSS 新知识的读者,值得仔细阅读与上机实习。对于老读者,也应温故知新。

第三章分层对数线性分析是高级统计技术,适用于标称(定类)——标称(定类)数据的对数线性分析。

第四章判别分析,第五章因素分析、第六章聚类分析、第七章大样本快速聚类,是高级统计,是人人应该掌握的、最重要、最有实用价值的高层次统计技术。在以前出版的四册 SPSS/PC<sup>+</sup> 应用教程中曾想涉及但未深入下去,由于深为国内读者求知欲望所感动,这次对它们作了较为详尽的理论及方法研究与奉献。尽管某些调查数据还不尽人意,但作为抛砖引玉,不可谓不可。

第八章至第十二章也是我们新编著的数据分析与数据检验技巧,特别是多元线性回归一章,其作用不亚于高级专业统计过程——Factor 重要。

与众不同的是,本书的每一章后面都有习题压阵,旨在让读者学习了一章之后,对重要的技能知识学深、记牢、用活。这就体现了作为一本教科书的应有特色。

每一章习题的答案,也是极其重要的,我们已经逐一解答过了,准备在下一册附录中与读者见面。

总之,此书对于社会学、医学、心理学、商学、经济学领域的应用可望起到左右翅的腾飞作用。

如果您是 SPSS 的初学者或初级者,建议从第一、二章开始学用结合。如果您对统计知识欠缺,或对统计结果把握性欠佳,建议您花点本钱和精力,参加 SPSS 的有关培训班并熟读本书第三章至第十二章,其效果将突飞猛进!

参加本书写作的有阮桂海、蔡建瓴、佟福玲、苏红、周爽、贾克云、陈根洪、孙光斗、陈卓、周爱群、郑金山、徐金本、卢元禄、洪秀花、闫桂珍、曲庆云、蔡建琼、唐军、柳玉芝、王肖群、杨武栓

等同志。

在出版前,社会学系卢淑华、郭崇德等教授提出了宝贵的指导,蔡文英等同志付出了辛勤的劳动,值此,我们谨代表广大读者,向软件的作者——美国 SPSS INC 公司和促使我们执笔的同仁们致以深深的谢意!

由于时间匆促且经验不足,加上水平有限,书中不妥之处谨盼专家能人批评指正。

为使本书学用结合,每年4月25日和7月10日在北京大学社会学系计算机室举办全国性 SPSS 新技术培训,时间一周左右。电话(010)62756149 或 62751676。

北京大学社会学系计算机室

阮桂海 1998年4月

# 目 录

|  |      |
|--|------|
| <b>第一章 SPSS for Windows 快速入门知识</b> .....                         | (1)  |
| <b>第一节 用 Windows 环境下的 SPSS 做统计分析的三大步骤</b> .....                  | (1)  |
| 一、编辑 SPSS 的程序 .....  | (1)  |
| 二、在 Syntax 窗口 Paste(贴回)命令 .....                                  | (4)  |
| 三、从 Windows 分菜单快速进入 Syntax 窗口 .....                              | (5)  |
| 四、从 Windows 分菜单快速显示 Newdata 窗口中的数据 .....                         | (6)  |
| <b>第二节 点 Windows 菜单以显示和打印结果</b> .....                            | (6)  |
| <b>第三节 点 Windows 菜单项以快速处理 Newdata 窗口中的数据</b> .....               | (8)  |
| 一、移动、拷贝 Newdata 窗口中的单元数据 .....                                   | (9)  |
| 二、删除变量和个案 .....  | (9)  |
| 三、插入新个案和拷贝新个案 .....  | (10) |
| <b>第四节 运行一个 SPSS 过程</b> .....                                    | (11) |
| <b>第五节 打开一个 * .SAV 目标数据文件</b> .....                              | (14) |
| <b>第六节 如何读取一个电子表格数据</b> .....                                    | (16) |
| 一、SPSS 数据文件的特点与规则 .....  | (16) |
| 二、读取电子表格数据的规则 .....  | (16) |
| 三、读取 Excel 电子表格数据 .....  | (17) |
| <b>第七节 SPSS 如何读取 dBASE 数据和 Foxbase 数据</b> .....                  | (18) |
| 一、建立数据库 .....  | (18) |
| 二、读取 Foxbase 数据库数据,进入 SPSS 系统 .....                              | (19) |
| <b>第八节 SPSS 如何保存数据文件</b> .....                                   | (20) |
| 一、将 * .SAV 文件类型保存成数据库类型 * .DBF .....                             | (20) |
| 二、将 Newdata 窗口中的 * .SAV 类型转换为 ASCII 码类型 * .DAT .....             | (22) |
| <b>习题一</b> .....   | (23) |
| <b>第二章 在 SPSS for Windows 中调用 DOS 版 SPSS/PC+ 的程序文件</b> .....     | (25) |
| <b>第一节 Windows 下的 SPSS 如何调用 DOS 版 SPSS/PC+ 中的程序 * .INC</b> ..... | (25) |
| 一、在 Windows 下的 SPSS 中为何不宜重新输入原始数据 .....                          | (25) |
| 二、SPSS 直接调用 DOS 版的 SPSS/PC+ 中的程序 * .INC .....                    | (28) |
| 三、用 SPSS 的 New 命令调用 SPSS/PC+ (V4.0 以前)程序的步骤 .....                | (30) |
| <b>第二节 用 SPSS 的 Open 命令调用 SPSS/PC+ (V4.0 以前)的程序</b> .....        | (31) |
| 一、简单步骤介绍 .....   | (31) |
| 二、具体操作步骤 .....   | (31) |
| <b>习题二</b> .....   | (32) |
| <b>第三章 分层对数线性分析模型(Hierarchical Loglinear Models)的应用</b> .....    | (35) |
| <b>第一节 对数线性分析的应用</b> .....                                       | (35) |

|  |             |
|--|-------------|
| 第二节 对数线性关系的模型 .....                          | (35)        |
| 一、饱和模型(Saturated Model) .....                | (36)        |
| 二、饱和模型的观察值与期望值 .....                         | (37)        |
| 第三节 参数估计 .....                               | (38)        |
| 第四节 独立模型——非饱和模型 .....                        | (39)        |
| 一、卡方拟合度的检验 .....                             | (40)        |
| 二、残差 .....                                   | (41)        |
| 第五节 分层模型 .....                               | (42)        |
| 第六节 选择模型 .....                               | (43)        |
| 第七节 卡方统计量的分解 .....                           | (43)        |
| 一、K次项(及其以上)效应项为零的假设检验 .....                  | (43)        |
| 二、K次项效应为零的检验 .....                           | (46)        |
| 第八节 检验模型中个别效应 .....                          | (46)        |
| 第九节 产生各次项效应的程序 .....                         | (48)        |
| 第十节 利用自后淘汰法建立模型 .....                        | (48)        |
| 第十一节 在 Windwos 中通过对话框进行 Loglinear 分析 .....   | (50)        |
| 一、进入 Hiloglinear 主对话框的方法 .....               | (50)        |
| 二、模型选择(Model) .....                          | (51)        |
| 三、Option 选项 .....                            | (52)        |
| 四、对话框的扩充 .....                               | (53)        |
| 第十二节 在 Windows 对话框中运行 Hiloglinear 过程实例 ..... | (53)        |
| 一、进入 Hiloglinear 主对话框 .....                  | (53)        |
| 二、将对话框中已点选的命令、子命令项粘贴回来,供修改或保存 .....          | (55)        |
| 习题三 .....                                    | (56)        |
| <b>第四章 DISCRIMINANT 过程的应用 .....</b>          | <b>(59)</b> |
| 第一节 从逃课现象研究入手 .....                          | (59)        |
| 一、选择分析的个案 .....                              | (60)        |
| 二、组间均值差别 .....                               | (60)        |
| 三、Wilks 的 $\lambda$ (Lambda)值 .....          | (62)        |
| 四、相关系数的评价 .....                              | (62)        |
| 五、判别系数的估计 .....                              | (63)        |
| 六、判别分的计算 .....                               | (64)        |
| 七、Bayes 的分组规则 .....                          | (65)        |
| 八、分组输出 .....                                 | (66)        |
| 九、个案分组后的概括统计表 .....                          | (66)        |
| 十、判别分的直方图 .....                              | (68)        |
| 十一、计算错误分组的比例 .....                           | (68)        |
| 十二、不正确分组的期望比例 .....                          | (70)        |
| 十三、判别分析的其它统计量 .....                          | (70)        |

|  |      |
|--|------|
| 十四、判别函数系数的解释 .....                               | (72) |
| 十五、判别函数与变量二者的相关 .....                            | (72) |
| 十六、费歇尔分组函数系数 .....                               | (73) |
| 十七、与多元回归分析的关系 .....                              | (74) |
| 第二节 选择变量的方法 .....                                | (74) |
| 一、逐步选择变量的例子 .....                                | (74) |
| 二、变量选择的其它标准 .....                                | (78) |
| 第三节 三组判别分析 .....                                 | (80) |
| 一、函数数目 .....                                     | (80) |
| 二、判别函数的分组 .....                                  | (81) |
| 三、附加的统计量 .....                                   | (83) |
| 四、检验判别函数的显著性水平 .....                             | (84) |
| 五、变量对判别函数的贡献 .....                               | (86) |
| 第四节 当违背假设时 .....                                 | (86) |
| 第五节 用 Windows 下的 SPSS 作判别分析 .....                | (87) |
| 一、首先从 Spsswin 主菜单中选择 Discriminant 过程命令 .....     | (88) |
| 二、定义因变量的判别范围 .....                               | (89) |
| 三、选择一子集数据进行判别分析 .....                            | (90) |
| 四、判别分析中的统计量选择 .....                              | (90) |
| 五、判别分析的逐步选择变量法 .....                             | (91) |
| 六、分类(组)判别 .....                                  | (93) |
| 七、保存判别分析中的新变量 .....                              | (94) |
| 八、附加命令语句提高判别特征 .....                             | (94) |
| 第六节 用编程法做判别分析 .....                              | (95) |
| 一、DISCRIMINANT 命令表 .....                         | (95) |
| 二、下面子命令只用于 Windows 版的 SPSS 中 .....               | (95) |
| 三、“/OPPTION = /STATISTICS = ”仅用于 V4.0 前的版本 ..... | (96) |
| 习题四 .....  | (97) |

## 第五章 因素分析 .....

|                                   |       |
|-----------------------------------|-------|
| 第一节 因素分析的模型 .....                 | (99)  |
| 第二节 因素分析的步骤 .....                 | (100) |
| 第三节 因素分析所用的数据和程序 .....            | (101) |
| 一、因素分析的变量 .....                   | (101) |
| 二、因素分析的命令文件(程序) .....             | (102) |
| 第四节 抽取因素(Factor Extraction) ..... | (103) |
| 第五节 抽取前 5 个因素 .....               | (104) |
| 第六节 变量的共通性与惟一性 .....              | (105) |
| 第七节 用残差评估相关矩阵和因素模型 .....          | (107) |
| 第八节 用最大似然性法则抽取因素 .....            | (107) |



|                                 |              |
|---------------------------------|--------------|
| 第九节 抽取因素经验小结                    | (110)        |
| 第十节 转轴前后的对比                     | (111)        |
| 一、转轴之前的因素矩阵及图形                  | (111)        |
| 二、转轴后的因素矩阵与因素装载量图               | (112)        |
| 三、转轴前后的对比                       | (112)        |
| 第十一节 转轴的方式                      | (113)        |
| 一、转轴的方式                         | (113)        |
| 二、转轴前的情形与经过各种转轴后的结果             | (114)        |
| 第十二节 斜交转轴                       | (115)        |
| 一、斜交转轴的原理                       | (115)        |
| 二、斜交转轴后的因素图形阵和结构阵               | (115)        |
| 第十三节 直角转轴前后的结果比较                | (117)        |
| 一、直角转轴前后的因素装载量比较                | (117)        |
| 二、直角转轴前后的图形对比                   | (118)        |
| 第十四节 对转轴效果的验证                   | (119)        |
| 第十五节 因素分析                       | (119)        |
| 一、因素解释                          | (119)        |
| 二、因素模型                          | (120)        |
| 第十六节 因素分数                       | (120)        |
| 第十七节 在 Windows 下的 SPSS 中作因素分析   | (122)        |
| 一、进入 Factor 主对话框(主命令)的步骤        | (122)        |
| 二、Factor 分析的 <b>Option</b> 框    | (123)        |
| 三、因素分析时的描述性统计量                  | (124)        |
| 四、抽取因素个数                        | (124)        |
| 五、因素装载量的转轴(Rotation)            | (125)        |
| 六、因素得分                          | (126)        |
| 第十八节 在 Syntax 窗口编辑程序做 Factor 分析 | (127)        |
| 习题五                             | (128)        |
| <b>第六章 聚类分析的应用</b>              | <b>(131)</b> |
| 第一节 聚类分析与判别分析的区别                | (131)        |
| 第二节 聚类的对象                       | (131)        |
| 第三节 个案聚类                        | (131)        |
| 一、个案聚类前的变量选择                    | (131)        |
| 二、个案聚类前的变量标准化                   | (132)        |
| 第四节 聚类的形成法                      | (132)        |
| 一、凝聚法(也称连接阶层聚类)                 | (132)        |
| 二、凝聚法的标准                        | (133)        |
| 第五节 聚类分析的全过程                    | (133)        |
| 一、原始变量及其标准化值                    | (133)        |

|  |              |
|--|--------------|
| 二、欧几里德距离平方矩阵 .....                       | (134)        |
| 三、从冰柱图看聚类情形 .....                        | (135)        |
| 四、平均连接法的图表 .....                         | (136)        |
| 五、从树形图分析聚类成员 .....                       | (137)        |
| 第六节 图形显示的改正 .....                        | (138)        |
| 第七节 对变量的聚类 .....                         | (139)        |
| 第八节 采用编程法做聚类分析 .....                     | (140)        |
| 一、Cluster 命令表 .....                      | (140)        |
| 二、子命令注解 .....                            | (141)        |
| 第九节 利用 Windows 下的 SPSS 做分层聚类分析 .....     | (147)        |
| 一、进入 Hierarchical(下文简称“分层”聚类分析)子菜单 ..... | (147)        |
| 二、选择统计量 .....                            | (148)        |
| 三、画出聚类分析图 .....                          | (149)        |
| 四、分层聚类法的选择 .....                         | (149)        |
| 五、保存分层聚类时的聚类解新变量 .....                   | (153)        |
| 六、增加命令语句和丰富聚类特征 .....                    | (153)        |
| 习题六 .....                                | (153)        |
| <b>第七章 大文件的聚类分析 .....</b>                | <b>(155)</b> |
| 第一节 聚类方法 .....                           | (155)        |
| 一、聚心已知后的聚类 .....                         | (155)        |
| 二、聚心未知时的聚类 .....                         | (159)        |
| 第二节 采用 Windows 版本的 SPSS 做大文件聚类分析 .....   | (162)        |
| 一、进入 K - Means 聚类分析子菜单 .....             | (162)        |
| 二、决定聚类中心(聚心) .....                       | (163)        |
| 三、K - Means 聚类分析的迭代 .....                | (164)        |
| 四、K - Means 聚类分析中新变量的保存 .....            | (165)        |
| 五、K - Means 聚类中的统计量选择 .....              | (165)        |
| 六、本章例子详解 .....                           | (166)        |
| 第三节 用编程法进行大文件的聚类 .....                   | (169)        |
| 一、QUICK Cluster 的命令表 .....               | (169)        |
| 二、QUICK Cluster 过程命令的条件限制 .....          | (170)        |
| 三、工作原理 .....                             | (170)        |
| 四、大样本个案的聚类法 .....                        | (170)        |
| 五、CRITERIA 子命令的用法 .....                  | (171)        |
| 六、METHOD 子命令的用法 .....                    | (171)        |
| 七、INITIAL 子命令的用法 .....                   | (172)        |
| 八、FILE 子命令的用法 .....                      | (172)        |
| 九、PRINT 子命令的用法 .....                     | (172)        |
| 十、MISSING 子命令的用法 .....                   | (173)        |

|                                       |              |
|---------------------------------------|--------------|
| 十一、OUTFILE 子命令的用法 .....               | (173)        |
| 十二、SAVE 子命令的用法 .....                  | (173)        |
| 习题七 .....                             | (174)        |
| <b>第八章 利用 Explore 过程探测数据 .....</b>    | <b>(177)</b> |
| 第一节 Explore 过程可探测数据的哪些错误 .....        | (177)        |
| 第二节 做 Explore 探测的数据文件 .....           | (178)        |
| 第三节 数据探测之一:用直方图 .....                 | (179)        |
| 第四节 图形显示之二:茎—叶图(Stem - Leaf) .....    | (180)        |
| 第五节 框图(Box plot) .....                | (182)        |
| 第六节 统计分析前对假设的检验 .....                 | (184)        |
| 一、Levene 检验 .....                     | (184)        |
| 二、用幅度—数值图示法检验(Spread - Level) .....   | (185)        |
| 三、转换的形式 .....                         | (185)        |
| 第七节 正态性检验 .....                       | (186)        |
| 第八节 集中趋势分布的三种较佳平稳测度 .....             | (188)        |
| 一、较佳测度之一:中位数 .....                    | (188)        |
| 二、较佳测度之二:修正均值 .....                   | (188)        |
| 三、较佳测度之三:M 估计 .....                   | (189)        |
| 第九节 在 Windows 中进行 Explore 的数据探测 ..... | (190)        |
| 一、进入 Explore 对话框的步骤 .....             | (190)        |
| 二、Statistics 对话框 .....                | (191)        |
| 三、利用图形探查 .....                        | (192)        |
| 四、缺省值处理 .....                         | (193)        |
| 第十节 在 Syntax 窗口编程 .....               | (194)        |
| 一、EXAMINE 命令语句的格式 .....               | (194)        |
| 二、EXAMINE 过程命令的主要功能 .....             | (195)        |
| 三、EXAMINE 的命令顺序及基本语句 .....            | (195)        |
| 四、VARIABLES 子命令 .....                 | (195)        |
| 五、COMPARE 子命令 .....                   | (196)        |
| 习题八 .....                             | (196)        |
| <b>第九章 多元线性回归分析 .....</b>             | <b>(199)</b> |
| 第一节 什么是多元线性回归分析 .....                 | (199)        |
| 第二节 选择一条回归线 .....                     | (200)        |
| 第三节 异常点(Outliers) .....               | (201)        |
| 第四节 标准化的回归系数 .....                    | (201)        |
| 第五节 从样本到总体 .....                      | (201)        |
| 一、总体参数的估计 .....                       | (202)        |
| 二、假设的检验 .....                         | (203)        |

|  |       |
|--|-------|
| 三、置信区间 .....                             | (203) |
| 第六节 模型的拟合优度 .....                        | (204) |
| 一、 $R^2$ 系数 .....                        | (204) |
| 二、方差分析 .....                             | (204) |
| 三、 $R^2$ 的另一解释 .....                     | (206) |
| 第七节 预测值及其标准误差 .....                      | (206) |
| 一、预测均值的特性曲线 .....                        | (207) |
| 二、预测一个新值时误差的来源 .....                     | (208) |
| 三、读个案的预测值和标准误差图 .....                    | (210) |
| 第八节 寻找是否满足假设 .....                       | (211) |
| 一、残差 .....                               | (211) |
| 二、直线性(Linearity) .....                   | (211) |
| 三、等方差性 .....                             | (212) |
| 四、误差的独立性 .....                           | (213) |
| 五、残差的正态性 .....                           | (214) |
| 第九节 设置异常值(Outliers) .....                | (216) |
| 一、有多个预测变量时的异常值测量法: Mahalanobis 距离法 ..... | (217) |
| 二、探测影响点(删除残差与 Cook 距离法) .....            | (218) |
| 第十节 数据不符合线性回归的基本假设时的处理法 .....            | (221) |
| 一、变非线性为线性关系 .....                        | (221) |
| 二、如何变非线性为线性关系 .....                      | (221) |
| 三、偏态分布的改善 .....                          | (222) |
| 四、方差的稳定 .....                            | (222) |
| 五、数据的对数转换 .....                          | (222) |
| 六、假设的重要性 .....                           | (223) |
| 第十一节 多元线性回归模型 .....                      | (223) |
| 一、起始成绩的预测 .....                          | (224) |
| 二、相关系数阵 .....                            | (224) |
| 三、偏回归系数 .....                            | (225) |
| 第十二节 确定重要的变量 .....                       | (226) |
| 一、B 系数与 Beta 系数的不足之处 .....               | (226) |
| 二、部分相关与偏相关 .....                         | (227) |
| 第十三节 回归模型的讨论 .....                       | (228) |
| 一、增删变量时的不同后果 .....                       | (228) |
| 二、自变量(预测变量)的最佳个数 .....                   | (229) |
| 第十四节 向前选择变量法 .....                       | (230) |
| 一、向前选择变量的两种标准 .....                      | (230) |
| 二、工作原理 .....                             | (231) |
| 三、结果分析 .....                             | (231) |
| 第十五节 自后淘汰变量(消元)法 .....                   | (232) |

|  |       |
|--|-------|
| 一、消元法的两种标准值 .....                      | (232) |
| 二、消元法的工作原理 .....                       | (232) |
| 三、结果分析 .....                           | (233) |
| 第十六节 逐步回归 .....                        | (234) |
| 一、两种标准值 .....                          | (234) |
| 二、工作原理 .....                           | (234) |
| 三、注意事项 .....                           | (234) |
| 四、结果分析 .....                           | (234) |
| 第十七节 多元回归中假设之检验 .....                  | (235) |
| 一、用 Regreesion 产生偏回归图 .....            | (236) |
| 二、用已生成的 Res_1 和 Res_2 偏残差变量画出散点图 ..... | (236) |
| 三、比较两幅可比性偏回归图 .....                    | (236) |
| 四、用变量的平方法修正线性度 .....                   | (237) |
| 第十八节 多元回归中影响点的探测 .....                 | (238) |
| 一、找出杠杆值 .....                          | (239) |
| 二、逐一排除异常个案法 .....                      | (239) |
| 三、探测之三:是否影响回归系数的方差 .....               | (239) |
| 四、保存协方差比的方法 .....                      | (240) |
| 第十九节 共线性诊断 .....                       | (240) |
| 一、共线性诊断之一:容许度 .....                    | (240) |
| 二、共线性诊断之二:方差膨胀因子 VIF .....             | (240) |
| 三、容许度的计算 .....                         | (242) |
| 四、共线性诊断之三:特征值和条件指标 .....               | (242) |
| 五、共线性诊断之四:方差比例 .....                   | (242) |
| 第二十节 解释回归模型 .....                      | (243) |
| 第二十一节 在 SPSS for Windows 中运行线性回归 ..... | (244) |
| 一、主对话框简析 .....                         | (244) |
| 二、WLS 选项 .....                         | (246) |
| 三、线性回归统计量的选择 .....                     | (246) |
| 四、线性回归图 .....                          | (247) |
| 五、保存回归的临时变量 .....                      | (248) |
| 六、线性回归可选项 .....                        | (249) |
| 七、用编程法弥补对话框的功能 .....                   | (250) |
| 第二十二节 线性回归的命令、子命令 .....                | (251) |
| 一、VARIABLE 子命令 .....                   | (252) |
| 二、DEPENDENT 子命令 .....                  | (253) |
| 三、METHOD 子命令 .....                     | (253) |
| 四、STATISTICS 子命令 .....                 | (254) |
| 五、CRITERIA 子命令 .....                   | (255) |
| 六、ORIGIN 子命令和 NOORIGIN 子命令 .....       | (256) |

|   |              |
|---|--------------|
| 七、DESCRIPTIVES 子命令 .....                  | (256)        |
| 八、SELECT 子命令 .....                        | (257)        |
| 九、缺省值(MISSING 子命令) .....                  | (257)        |
| 十、分析残差(使用 19 个临时变量之一) .....               | (257)        |
| 十一、画出残差图(RESIDUALS 子命令) .....             | (258)        |
| 十二、临时变量的 Casewise 图(用 Casewise 子命令) ..... | (259)        |
| 第二十三节 应用实例 .....                          | (260)        |
| 一、程序文件(919293.SPS) .....                  | (260)        |
| 二、用 REGRESSION 窗口的 Paste 命令剪贴的回归命令 .....  | (261)        |
| 习题九 .....                                 | (262)        |
| <b>第十章 多项选择的统计分析过程的应用 .....</b>           | <b>(265)</b> |
| 第一节 什么是多项选择 .....                         | (265)        |
| 一、多项二分法(Multiple dichotomy method) .....  | (265)        |
| 二、多项分类法(Multiple category method) .....   | (266)        |
| 第二节 多项选择的数据分析法 .....                      | (266)        |
| 第三节 首先均需选定多项选择的变量集 .....                  | (266)        |
| 第四节 统计多项选择时的频次 .....                      | (268)        |
| 第五节 多项选择时的交叉汇总表 .....                     | (270)        |
| 一、操作步骤 .....                              | (270)        |
| 二、图形解释 .....                              | (273)        |
| 第六节 在 Syntax 窗口输入命令(编程法) .....            | (275)        |
| 一、MULT RESPONSE 命令语句一览表 .....             | (275)        |
| 二、过程概述 .....                              | (276)        |
| 三、基本语句 .....                              | (276)        |
| 四、条件限制及命令的次序 .....                        | (276)        |
| 五、GROUPS 子命令 .....                        | (277)        |
| 六、VARIABLES 子命令 .....                     | (277)        |
| 七、FREQUENCIES 子命令 .....                   | (278)        |
| 八、TABLES 子命令 .....                        | (278)        |
| 九、TABLES 中的 PAIRED 关键词 .....              | (279)        |
| 十、CELLS 子命令 .....                         | (280)        |
| 十一、BASE 子命令 .....                         | (280)        |
| 十二、FORMAT 子命令 .....                       | (281)        |
| 十三、MISSING 子命令 .....                      | (281)        |
| 习题十 .....                                 | (282)        |
| <b>第十一章 偏相关分析(PARTIAL 过程的应用) .....</b>    | <b>(285)</b> |
| 第一节 计算偏相关系数 .....                         | (285)        |
| 一、偏相关系数的“阶” .....                         | (285)        |

|  |              |
|--|--------------|
| 二、显著性检验的公式 .....                         | (286)        |
| 第二节 辨别变量虚假相关 .....                       | (286)        |
| 一、选择一个控制的变量 .....                        | (287)        |
| 二、选择两个控制的变量 .....                        | (287)        |
| 第三节 如何确定被控制的变量 .....                     | (288)        |
| 一、首先要把握因果关系 .....                        | (288)        |
| 二、若有隐含相关,则应控制相应的变量 .....                 | (228)        |
| 三、产生图 11-3 的程序命令 .....                   | (289)        |
| 第四节 利用 Windows 下的 Partial 做偏相关分析 .....   | (289)        |
| 一、操作步骤 .....                             | (289)        |
| 二、偏相关主对话框 .....                          | (289)        |
| 三、基本选择 .....                             | (290)        |
| 四、显著性检验的选择(见 Test of significance) ..... | (290)        |
| 五、偏相关选择项 .....                           | (290)        |
| 第五节 在 Syntax 窗口可附加的功能命令 .....            | (291)        |
| 第六节 在 Syntax 窗口编程的命令语句 .....             | (291)        |
| 一、PARTIAL CORR 命令语句一览表 .....             | (291)        |
| 二、基本语句和条件限制 .....                        | (292)        |
| 三、VARIABLE 子命令 .....                     | (292)        |
| 四、STATISTICS 子命令 .....                   | (293)        |
| 五、SIGNIFICANCE 子命令 .....                 | (293)        |
| 六、FORMAT 子命令 .....                       | (293)        |
| 七、MISSING 子命令 .....                      | (294)        |
| 八、MATRIX 子命令 .....                       | (294)        |
| 习题十一 .....                               | (297)        |
| <b>第十二章 运用 PROXIMITY 过程做距离测量 .....</b>   | <b>(299)</b> |
| 第一节 区间以上的数据采用距离测量 .....                  | (299)        |
| 第二节 频率数据采用非相似测量 .....                    | (300)        |
| 第三节 双择数据采用相似性测量 .....                    | (300)        |
| 第四节 PROXIMITIES 过程命令表 .....              | (302)        |
| 一、命令的作用 .....                            | (303)        |
| 二、命令的最基本写法 .....                         | (303)        |
| 三、变量指定 .....                             | (303)        |
| 四、用 STANDARDIZE 标准化数据 .....              | (303)        |
| 五、MEASURE 子命令的用途 .....                   | (304)        |
| 六、连续数据的测量法 .....                         | (304)        |
| 七、频次数据的测量法 .....                         | (305)        |
| 八、对分数据(Binary data)的测量法 .....            | (305)        |
| 九、例子 .....                               | (307)        |

|      |                            |       |
|------|----------------------------|-------|
| 第五节  | MATRIX 子命令的功能 .....        | (307) |
|      | 一、用 OUT 关键词写一个矩阵数据文件 ..... | (308) |
|      | 二、矩阵数据文件的内容 .....          | (308) |
|      | 三、用 IN 关键词读取已有的矩阵文件 .....  | (308) |
|      | 四、文件裂变 .....               | (309) |
| 第六节  | 写/读矩阵的例子 .....             | (309) |
|      | 一、写矩阵的例子 .....             | (309) |
|      | 二、读取矩阵文件中的数据 .....         | (310) |
|      | 三、现存现读矩阵数据文件 .....         | (310) |
| 习题十二 | .....                      | (311) |



# 第一章 SPSS for Windows 快速入门知识

Windows 环境支持的 SPSS 与 DOS 环境支持的 SPSS/PC+ 相比,最大的优点是:界面清晰、人机对话简明快捷,不懂 SPSS 语句及语法知识的用户,也可随心所欲地用鼠标点出菜单命令及子命令、统计关键词;即使点错了也无妨,改对了命令、子命令及关键词之后,见到高亮度的 **OK** 框,说明符合过程命令的语法规则,便可点一下 **OK** 框,让 SPSS 执行命令、统计出结果。

因此,有人比喻说,即使是机盲,也能从 SPSS 的对话框中点出“鸳鸯谱”。

但是,对于 SPSS 所显示的统计结果,我们要用相应的统计学知识才能准确无误地分析出结果,并写出符合实际的科学论文。

一句话,学习 SPSS for Windows 并不难,难的是如何正确地分析、解释 SPSS 的统计结果。因此,除了需要在这本应用教程的指导下亲自上机实践外,还应该参加有关专家举办的技术培训,专家往往可为您“点石成金”,使我们对统计结果的分析不出笑话。

本章就如何快速入门 Windows 环境下的 SPSS 讲述以下几点看法。

## 第一节 用 Windows 环境下的 SPSS 做统计分析的三大步骤

在 DOS 环境下的 SPSS/PC+ 中,进行统计分析主要有以下两个步骤:

- 其一:编辑一个 SPSS/PC+ 可执行的程序(包括数据定义、数据和统计过程,即命令);
- 其二:从菜单中选择其它统计过程命令后运行。

但在 Windows 环境下的 SPSS 中,进行统计分析需要以下三大步骤:

- 其一:编辑一个 SPSS 可执行的程序(包括数据定义、数据和统计过程,即命令);然后,放在 OPEN 到 SPSS 系统中;详见(二)。
- 其二:从菜单中选择一个统计过程,如:

Statistics、Summarize►、Frequencies…

- 其三:从 SPSS 对话框中,点选变量清单及统计量。然后点 **OK** 执行命令。详见下文。

### 一、编辑 SPSS 的程序

SPSS 的程序指的是:①由 DATA LIST 命令定义的变量名、变量类型、栏目位置;②统计之前所需的数据(或数据文件);③统计分析所用的过程命令。

SPSS 程序的编辑工具很广泛,凡是文本文件的编辑器,都适用于 SPSS 的编程。

如①DOS 命令中的 EDIT 编辑命令、②WordStar、③WPS 的 N 命令、④中文之星系统中的 NPS 编辑器、⑤中文 Word 编辑器、⑥用 dBASE(或 FoxBASE)建立数据库数据、⑦用 SPSS 的 SYNTAX 窗口编辑程序等。

有人也许会问:Windows 下的 SPSS 的 NewData 窗口是为输入数据而特设的窗口,为什么还要另外在上述 7 种编辑器中编辑?

我们对此的回答是:一切为了省时简便、事半功倍!

虽说 NewData 窗口可用以输入数据,其实上册的第三章和本书(即中册)的第二章都不同