

关键词

统计分析

SPSS 软件应用



高效办公 Express to
Office Efficiency
“职”通车

SPSS 在统计分析中的应用

朱建平 殷瑞飞 编著

针对性 比较全面地覆盖了统计分析中的各种问题，针对性较强

系统性 系统地探讨了SPSS与统计分析相结合的各种方法和技术

实用性 介绍基于SPSS的统计分析方法，对企事业单位的统计工作实践有一定的指导意义



清华大学出版社

SPSS 在统计分析中的应用

朱建平 殷瑞飞 编著

清华大学出版社

北 京

内 容 简 介

本书以统计思想为主线,从统计分析的实际过程出发,根据计算机统计分析流行软件 SPSS 自身的特点,将统计分析方法的基本思想与计算机软件的实际操作有机结合,旨在帮助读者理解统计分析方法思想的同时,学习 SPSS 软件的操作方法,将统计分析软件灵活运用于实际的统计分析工作中。

本书特别强调实用性和可操作性,尽量避免各种统计方法中繁琐的数学证明过程。即使数学功底不坚实的读者,通过对本书的学习,也能成功完成各种统计分析任务。同时,本书注重理论联系实际,配有丰富的案例。各章均按照“理论知识——操作步骤——实例分析”的顺序来安排结构,书末附录中还提供了 SPSS 函数对照表和常用统计词汇英汉对照表。本书编排合理、便于检索,可帮助读者轻松掌握 SPSS 在统计分析中的应用。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13501256678 13801310933

图书在版编目(CIP)数据

SPSS 在统计分析中的应用/朱建平,殷瑞飞编著. —北京:清华大学出版社,2007.1

ISBN 978-7-302-14066-5

I. S… II. ①朱… ②殷… III. 统计分析-软件包, SPSS IV. C819

中国版本图书馆 CIP 数据核字(2007)第 126732 号

责任编辑:吴颖华 赵玉琨

封面设计:张 岩

版式设计:李永梅

责任校对:焦章英

责任印制:孟凡玉

出版发行:清华大学出版社 地 址:北京清华大学学研大厦 A 座

<http://www.tup.com.cn> 邮 编:100084

c-service@tup.tsinghua.edu.cn

社 总 机:010-62770175 邮购热线:010-62786544

投稿咨询:010-62772015 客户服务:010-62776969

印 刷 者:清华大学印刷厂

装 订 者:三河市漂源装订厂

经 销:全国新华书店

开 本:185×260 印 张:14.25 字 数:325 千字

版 次:2007 年 1 月第 1 版 印 次:2007 年 1 月第 1 次印刷

印 数:1~5000

定 价:20.00 元

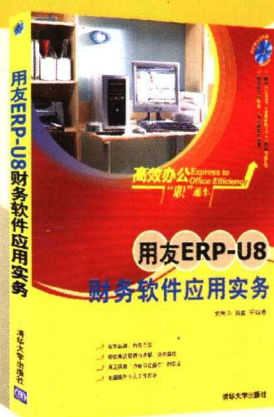
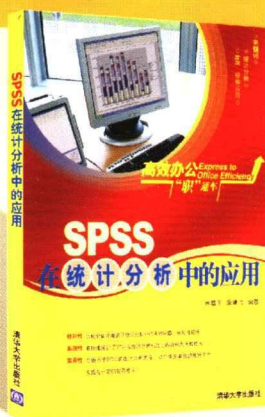
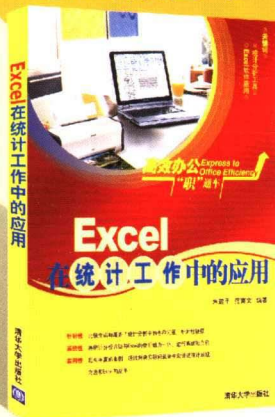
本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010)62770177 转 3103 产品编号:019821-01

高效办公 Express to Office Efficiency “职”通车

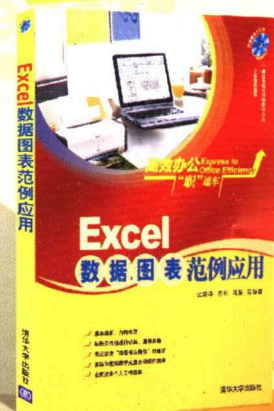
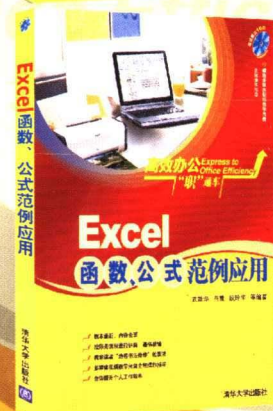
案例丰富易于学习
配套光盘全程指导

专业人士精心打造
全面提升工作效率

已出版图书



系列后续好书
正在出版中，
敬请期待……



清华大学出版社

地址：北京清华大学学研大厦
邮购热线：(010)62786544

邮编：100084
客户服务：(010)62791976-269

丛 书 序

随着计算机技术的全面迅速发展,人们将在行政办公、财务、统计、审计等众多领域面对计算机的应用和管理。掌握计算机在这些领域的应用,一方面可以极大地提高这些工作人员的工作效率,另一方面也可以提高他们的业务水平。信息时代,许多行业都要求工作者有很强的计算机操作技能,做到运用自如,熟练而且深入地掌握软件的应用。而要做到这一点,必须从软件的实际应用入手。

正是在这一大背景下,我们策划了本套丛书,精选了应用领域较广泛、较常用的一些软件,如 Excel、Access、SPSS、用友财务软件、金蝶财务软件等,从全新的实例角度出发,按照“基本知识点讲解——实践应用——常见问题解答”的结构,全面介绍了这些软件在日常工作中的应用,旨在帮助广大办公人员、财务人员、统计分析人员、审计人员及相关专业的学生快速掌握这些软件的应用,用以解决实际工作中的问题,提高自身的应用水平。

内容安排

本丛书强调软件与职业应用相结合,以实例为载体,着重介绍常用软件的操作功能和实践应用技巧。本套丛书包括:

第一批

《用友 ERP-U8 财务软件应用实务》

《金蝶 K/3 财务软件应用实务》

《SPSS 在统计分析中的应用》

《Excel 在统计工作中的应用》

第二批

《Excel 在会计工作中的应用》

《Excel 在财务管理中的应用》

《Excel 在审计分析中的应用》

《Excel 函数、公式范例应用》

《Excel 数据图表范例应用》

《Excel VBA 办公范例应用》

《Access 公司管理范例应用》

丛书特色

1. 软件与职业应用相结合,实用性强。深入浅出地讲述了财务、统计、审计各职业领域的关键知识,系统介绍了相应的软件应用方法及技巧,对实际工作有极大的帮助和指导意义。

2. 内容丰富,案例典型。本套书每章都有实践案例,广大读者可以根据自己的情况进行取舍,直接应用于具体的工作之中。

3. 结构合理, 逻辑清晰。从全新的实例角度出发, 按照“基本知识点讲解——实践应用——常见问题解答”的结构, 全面介绍了这些软件在日常工作中的应用。符合读者的学习思路, 可以使广大读者在最短的时间内学习并利用应用软件的各种强大功能, 少走弯路, 迅速提升专业技能和提高工作效率。

4. 光盘特色。本套丛书大部分都配有光盘, 汇集了书中所用的应用软件、实例素材, 及应用实例的视频, 极大地方便了读者的学习。

读者定位

1. 适合作为行政办公、财务、统计、审计等领域在职工作人员提高自身业务水平的参考用书。

2. 适合作为高校财务或经济管理等相关专业的教材或学习用书。

3. 适合作为各相关领域应用培训或职业培训的教学用书。

售后服务

如果读者在阅读图书的过程中有什么问题或需要帮助, 可以登录本丛书的信息支持网站 <http://www.thjd.com.cn> 或通过 zzfangcn@vip.163.com 联系, 也可以在 <http://www.thjd.com.cn> 的读者留言栏目留言, 我们将尽快给您提供帮助与支持。

前 言

统计学是研究如何测定、收集、整理、归纳和分析反映客观现象总体数量的数据，以便给出正确认识的方法论科学。现代统计学已经渗透到理、工、农、医、经济管理各类人文社会科学领域，并由此产生许多新的交叉学科，改变了原有单一学科发展的思路，对科学的发展具有极大的支撑作用。然而，在统计理论的学习以及统计方法的应用过程当中，进行大规模的数据处理和计算是必不可少的，这就必然需要有计算机及统计软件的支持。作为一名统计工作者，在熟悉统计理论的同时，掌握一种统计软件的操作也是十分必要的。SPSS (Statistics Package for Social Science) for Windows 是在 SPSS/PC (for DOS) 基础上发展起来的，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，是当今世界上最为流行的统计软件之一。因此，我们编写了《SPSS 在统计分析中的应用》一书。

本书的编写力求以统计思想为主线，并使广大的经济和管理工作者能够针对其所面临的实际问题，利用 SPSS 软件强大的统计功能快速而准确地解决问题。本书所有章节均按照“理论知识——操作步骤——实例分析”的顺序来安排结构，以使读者能够在理解统计方法的基础上读懂软件的输出结果，并对输出结果作出合理的解释。其基本框架是：第一部分包括第 1 章和第 2 章，具体介绍 SPSS 软件的入门基础知识；第二部分包括第 3 章至第 9 章，内容包括经典统计方法与 SPSS 的实现，具体介绍统计描述、统计推断、方差分析、非参数检验、相关分析和回归分析等；第三部分包括第 10 章至第 15 章，内容包括现代统计分析方法与 SPSS 的实现，具体介绍聚类分析、判别分析、因子分析、主成分分析、相应分析和典型相关分析等。本书在最后的附录中提供了 SPSS 函数对照表和常用统计词汇英汉对照表，以供读者在需要进行查询。

在本书的编写过程中，我们根据经济和管理人员在应用统计分析方法的实际要求，突出以下特点：

1. 充分体现“少而精”的原则。本书用较少的篇幅介绍 SPSS 软件在统计分析中的应用，对于统计基础薄弱的读者，可以在脱离任何一本统计理论书籍帮助的情况下，也能够理解 SPSS 软件每一个常用选项的含义，并读懂每一项重要的 SPSS 输出结果。本书的编写并没有简单地罗列软件菜单和机械地讲解计算机操作步骤，而是在书中适当地加入了丰富的统计理论知识和实际案例。

2. 充分贯穿统计分析过程。本书不像同类书籍那样，按照 SPSS 菜单的前后上下顺序安排章节，而是按照统计学的理论框架和统计分析的实际过程来安排章节。这样安排的好处是利于读者根据面对的实际问题，按照合理的步骤，由简至难，边做边学，直至顺利得到分析结果，而不至于被计算机软件菜单割裂统计理论和统计分析过程的内在连贯性。

3. 充分加强理论与实践的结合。本书的绝大部分实例都是来自与人们生活、工作息息相关的社会、经济和管理等方面的实例。这样将 SPSS 软件的学习和案例分析有机结合，

不仅能使得读者在实践运用中学习 SPSS 软件的操作方法，而且还使读者对统计分析的意义有深入的体会。

在本书的编写过程中，参考了国内外相关文献资料，书后列出了主要参考文献。本书的出版得到了清华大学出版社的大力支持，在此表示衷心感谢！尽管我们做了不少的努力，想奉献给读者一本满意的书，但由于水平有限，书中难免有疏漏或错误之处，恳请读者多提宝贵意见，以便今后进一步修改与完善。

本书的编写得到了国家教育部“新世纪优秀人才支持计划”（Program for New Century Excellent Talents in University, NCET）的资助。

为了方便读者高效、便捷地使用本书，特免费提供本书所有实例的原始数据、源文件，请登录清华大学出版社网站（www.tup.tsinghua.edu.cn）下载。

作 者

2006 年 12 月于厦门大学

目 录

第 1 章 SPSS 软件概述	1
1.1 SPSS 软件的基本特点和功能	1
1.2 SPSS 软件的安装、启动与退出	2
1.2.1 SPSS 软件的安装	2
1.2.2 SPSS 软件的启动	2
1.2.3 SPSS 软件的退出	3
1.3 SPSS 操作环境介绍	3
1.3.1 SPSS 软件的 3 个常用窗口	3
1.3.2 SPSS 菜单和工具栏	5
1.3.3 SPSS 对话框的基本操作方式	5
第 2 章 SPSS 数据文件管理	7
2.1 SPSS 数据文件的结构	7
2.1.1 SPSS 数据文件的特点	7
2.1.2 SPSS 变量的属性	7
2.2 建立一个数据文件	10
2.3 读取外部数据	11
2.3.1 读取 Excel 文件	12
2.3.2 读取 ASCII 码文件	12
2.4 SPSS 数据的编辑和保存	15
2.4.1 Edit 菜单中的数据编辑功能	15
2.4.2 Data 菜单中的数据编辑功能	16
2.4.3 SPSS 数据的保存	16
第 3 章 数据整理	17
3.1 数据排序	17
3.2 数据排秩	18
3.3 数据转置	19
3.4 选择观测的子集	20
3.5 数据分类汇总	22
3.6 合并数据文件	23
3.6.1 纵向合并 (Add Cases)	23
3.6.2 横向合并 (Add Variables)	24

3.7 数据拆分	26
3.8 计算新变量	28
3.9 数据重新编码	30
3.10 数据分组	31
3.11 数据标准化	32
第4章 统计描述	34
4.1 基本概念和原理	34
4.1.1 频数分布	34
4.1.2 集中趋势指标	34
4.1.3 离散程度指标	35
4.1.4 反映分布形态的描述性指标	35
4.2 频数分析	36
4.2.1 操作步骤	36
4.2.2 实例结果分析	38
4.3 描述性统计量	39
4.3.1 操作步骤	39
4.3.2 实例结果分析	40
4.4 探索性数据分析	41
4.4.1 操作步骤	42
4.4.2 实例结果分析	44
4.4.3 方差齐性检验的实例	46
第5章 统计推断	47
5.1 统计推断概述	47
5.1.1 参数估计	47
5.1.2 假设检验	48
5.2 单样本 t 检验	49
5.2.1 理论与方法	49
5.2.2 操作步骤	49
5.2.3 实例结果分析	50
5.3 两独立样本 t 检验	51
5.3.1 理论与方法	51
5.3.2 操作步骤	51
5.3.3 实例结果分析	52
5.4 配对样本 t 检验	53
5.4.1 理论与方法	53
5.4.2 操作步骤	54

5.4.3 实例结果分析	54
第6章 方差分析	56
6.1 方差分析概述	56
6.2 单因素单变量方差分析	56
6.2.1 理论和方法	56
6.2.2 操作步骤	58
6.2.3 实例结果分析	60
6.3 多因素单变量方差分析	62
6.3.1 理论与方法	62
6.3.2 固定效应、随机效应和协变量	64
6.3.3 操作步骤	65
6.3.4 实例结果分析	67
6.3.5 不考虑交互效应的多因素方差分析	70
6.3.6 引入协变量的多因素方差分析	70
第7章 非参数检验	72
7.1 非参数检验概述	72
7.2 卡方检验 (χ^2 检验)	72
7.2.1 理论与方法	72
7.2.2 操作步骤	73
7.2.3 实例结果分析	74
7.3 二项分布检验	75
7.3.1 理论与方法	75
7.3.2 操作步骤	75
7.3.3 实例结果分析	76
7.4 游程检验	77
7.4.1 理论与方法	77
7.4.2 操作步骤	77
7.4.3 实例结果分析	78
7.5 单样本 K-S 检验	79
7.5.1 理论与方法	79
7.5.2 操作步骤	79
7.5.3 实例结果分析	80
7.6 两独立样本检验	80
7.6.1 理论与方法	80
7.6.2 操作步骤	82
7.6.3 实例结果分析	82

7.7 多独立样本检验	84
7.7.1 理论与方法	84
7.7.2 操作步骤	85
7.7.3 实例结果分析	85
7.8 两配对样本检验	87
7.8.1 理论与方法	87
7.8.2 操作步骤	88
7.8.3 实例结果分析	88
7.9 多配对样本检验	90
7.9.1 理论与方法	90
7.9.2 操作步骤	91
7.9.3 实例结果分析	92
第 8 章 相关分析	93
8.1 相关分析的理论与方法	93
8.1.1 连续变量的相关分析	93
8.1.2 定序变量的相关分析	94
8.1.3 定类变量的相关分析	95
8.2 连续变量相关分析实例	96
8.2.1 操作步骤	96
8.2.2 实例结果分析	96
8.3 离散变量相关分析实例	98
8.3.1 操作步骤	98
8.3.2 实例结果分析	100
第 9 章 回归分析	103
9.1 线性回归	103
9.1.1 线性回归模型	103
9.1.2 线性回归模型的评价和检验	104
9.1.3 操作步骤	109
9.1.4 实例 1: 家庭住房支出与年收入的回归模型	112
9.1.5 实例 2: 中国电信业务总量的计量模型	115
9.2 二项 Logistic 回归	121
9.2.1 二项 Logistic 回归模型的形式	121
9.2.2 二项 Logistic 模型的估计	122
9.2.3 二项 Logistic 模型回归系数的意义	122
9.2.4 二项 Logistic 模型的评价和检验	123
9.2.5 操作步骤	125

9.2.6 实例结果分析	128
第 10 章 聚类分析	131
10.1 系统聚类法的理论与方法	131
10.1.1 系统聚类的基本思想	131
10.1.2 个体之间距离的度量方法	132
10.1.3 类之间距离的度量方法	134
10.2 K-均值聚类法的理论与方法	135
10.3 K-均值聚类法的实例	135
10.3.1 操作步骤	135
10.3.2 实例结果分析	137
10.4 系统聚类法的实例	140
10.4.1 操作步骤	140
10.4.2 实例结果分析	142
10.4.3 利用系统聚类法对不同地区产值进行分类	143
第 11 章 判别分析	145
11.1 Fisher 判别法的理论与方法	145
11.2 贝叶斯 (Bayes) 判别法的理论与方法	146
11.2.1 Bayes 判别法的基本思想	146
11.2.2 Bayes 线性判别函数	147
11.3 判别分析实例	147
11.3.1 操作步骤	147
11.3.2 实例结果分析	150
第 12 章 因子分析	155
12.1 因子分析的理论与方法	155
12.1.1 因子分析的基本思想	155
12.1.2 因子分析的数学模型	155
12.1.3 因子载荷阵的求解	156
12.1.4 因子载荷阵的统计意义	156
12.1.5 因子命名与因子旋转	158
12.1.6 因子得分	158
12.1.7 因子分析的出发点	159
12.2 因子分析的实例	159
12.2.1 操作步骤	160
12.2.2 实例结果分析	163
12.2.3 因子分析在市场研究中的应用	169

第 13 章 主成分分析	171
13.1 主成分分析的理论与方法	171
13.1.1 主成分分析的基本思想	171
13.1.2 主成分分析的数学模型	171
13.1.3 主成分的求解	172
13.1.4 主成分的方差贡献率	172
13.1.5 主成分的几何意义	173
13.1.6 主成分分析的出发点	173
13.1.7 利用因子分析的结果计算主成分	174
13.2 主成分分析的实例	174
13.2.1 操作步骤	175
13.2.2 对财务指标进行主成分分析	177
第 14 章 相应分析	179
14.1 相应分析的理论与方法	179
14.1.1 相应分析的基本思想	179
14.1.2 相应分析的基本步骤	179
14.2 相应分析的实例	181
14.2.1 操作步骤	181
14.2.2 实例结果分析	184
第 15 章 典型相关分析	189
15.1 典型相关分析的理论与方法	189
15.1.1 典型相关分析的基本思想	189
15.1.2 典型相关变量及典型相关系数的求解步骤	190
15.1.3 典型相关系数的显著性检验	191
15.1.4 从相关矩阵出发计算典型相关	192
15.2 典型相关的实例	192
15.2.1 操作步骤	192
15.2.2 实例结果分析	193
15.2.3 一个经典的典型相关分析的例子	196
附录	198
参考文献	215

第 1 章 SPSS 软件概述

社会科学统计软件包 (Statistics Package for Social Science, SPSS) 于 20 世纪 60 年代由美国斯坦福大学的 3 位研究生研制开发。20 世纪 80 年代以前, SPSS 软件主要应用于企事业单位。1984 年, SPSS 中心推出了基于 DOS 系统的微机版本。20 世纪 90 年代以后, 随着 Windows 系统的逐渐盛行, SPSS 也适时地推出了基于 Windows 操作平台的新版本。如今, SPSS 软件已经作为国际上最有影响力的统计软件之一, 广泛应用于社会学、经济学、生物学、教育学、心理学等各个领域。

1.1 SPSS 软件的基本特点和功能

SPSS 统计软件具有如下主要特点:

(1) SPSS 的命令语句、子命令及各种选项绝大部分都包含在各种菜单和对话框中, 因此, 用户无须花大量时间记忆繁杂的命令、过程、选项等。在 SPSS 中, 大多数操作可以通过菜单和对话框来完成, 因此操作简便, 易于学习和使用。

(2) 虽然大部分统计分析方法可以通过菜单和对话框来完成, 但是, 对于熟悉 SPSS 语言的用户, 也可以在语句窗口中直接编写程序语句, 从而更为灵活地完成各种复杂的统计分析任务。另外, 用对话框指定命令、子命令和选项之后, 通过单击 Paste 按钮可以把与选择对应的语句自动置于语句窗口中, 并可以文件形式保存。因此 SPSS for Windows 同时适用于 SPSS 的新老用户。

(3) 具有第四代语言的特点, 只要通过菜单的选择以及对话框的操作告诉系统要做什么, 而无须告之怎样做。只要粗通统计分析原理, 无须通晓统计分析的各种算法, 即可得到统计分析结果。

(4) 具有完善的与其他软件的数据转换接口。其他软件生成的数据文件, 如 Excel 文件、Access 文件、关系数据库生成的 DBF 文件、用文本编辑软件生成的 ASCII 码数据文件等均可方便地转换成可供分析的 SPSS 数据文件。

(5) 统计分析方法丰富, 提供了从简单的单变量描述分析到复杂的多变量分析的多种统计方法。

(6) 具有强大的图形功能, 不但可以得到数字结果, 还可以得到直观、漂亮的统计图, 形象地显示分析结果。

1.2 SPSS 软件的安装、启动与退出

1.2.1 SPSS 软件的安装

作为 Windows 操作系统下的应用软件，SPSS 软件的安装与一般的 Windows 应用软件的安装步骤大体相同。其步骤如下：

1. 在 SPSS 安装光盘的目录中找到安装文件 `setup.exe`，双击运行。
2. 在 Software License Agreement 对话框中选择接受软件使用协议。
3. 在 Choose Destination Location 对话框中指定软件安装目录。
4. 在 Use Information 对话框中输入用户基本信息及软件序列号。
5. 在 Setup Type 对话框中选择安装模式，通常选择典型安装 (Typical)。
6. 在 Personal or Shared installation 对话框中选择单机安装还是网络安装，一般选择系统默认的单机安装 (Personal installation)。
7. 在 Product license codes 对话框中输入软件许可证号码。
8. 在 Select Options 对话框中列出了各个 SPSS 软件模块供用户选择，用户可以按自己的需求自行选定需要安装的模块。
9. Ready To Install Files 对话框给出了一些安装信息，要求用户确认。直接单击 Next 按钮跳到下一个窗口。
10. Setup Complete 对话框的两个复选框要求用户选择是否在安装完成后直接运行 SPSS 程序以及是否阅读 SPSS 自述文件。单击 Finish 按钮，完成安装。

1.2.2 SPSS 软件的启动

启动 SPSS 软件的步骤如下：

1. 在 Windows “开始”菜单的“程序”菜单中找到 SPSS for Windows，在它的次级菜单中单击 SPSS 11.5 for Windows 即可启动 SPSS 软件，进入 SPSS for Windows 对话框，如图 1-1 所示。
2. 在 SPSS for Windows 对话框中各选项的意义分别为：
 - Run the tutorial: 运行操作向导。
 - Type in data: 在数据窗口中输入数据，选择此选项会直接打开一个空的数据编辑窗口，供用户自行输入数据。
 - Run an existing query: 运行一个已经存在的查询。
 - Create new query using Database Wizard: 用数据库处理工具建立新的数据文件。
 - Open an existing data source: 打开一个已经存在的数据文件，这是系统默认选项，在下面的列表中有用户最近使用过的 SPSS 数据文件供选择。

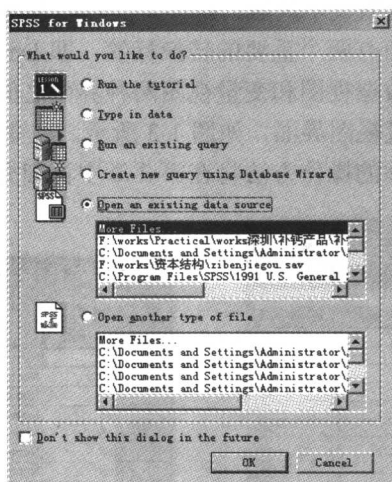


图 1-1 SPSS for Windows 对话框

- Open another type of file: 打开其他类型文件，如 SPSS 结果输出文件、语句文件等。
- Don't show this dialog in the future 复选框：选中之后，下次启动 SPSS 时将不会再显示该对话框，而直接进入数据编辑窗口。

1.2.3 SPSS 软件的退出

SPSS 软件的退出方法与其他 Windows 应用程序相同，有两种常用的退出方法：

- (1) 按 File→Exit 的顺序使用菜单命令退出程序。
- (2) 直接单击 SPSS 窗口右上角的“关闭”按钮，回答完系统提出的是否存盘的问题之后即可安全退出 SPSS 程序。

1.3 SPSS 操作环境介绍

1.3.1 SPSS 软件的 3 个常用窗口

1. 数据编辑窗口

启动 SPSS 后看到的第一个窗口便是数据编辑窗口，如图 1-2 所示。在数据编辑窗口中可以进行数据的录入、编辑以及变量属性的定义和编辑，是 SPSS 的基本界面。窗口上方是主菜单栏，包含了 SPSS 从文件管理到数据整理、分析的几乎所有功能。主菜单栏下方是工具栏，提供一些常用的 SPSS 功能，使某些操作更为快捷。工具栏下方是当前数据栏，左边部分显示了当前活动单元格对应的变量名和观测序号，右边部分显示了当前活动单元