

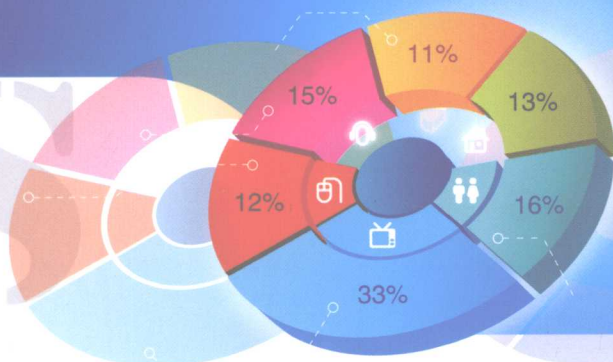


高等学校统计分析软件系列教材

统计软件SPSS

在医学中的应用实例教程

孔晓荣 张星光 主编



清华大学出版社

<http://www.tup.tsinghua.edu.cn>



北京交通大学出版社

<http://press.bjtu.edu.cn>

中国医药统计学会统计软件应用专业委员会

统计软件SPSS

在医学中的应用实例教程

张松林 王增武 主编

中国医药出版社

SPSS

SPSS

中国医药出版社

中国医药出版社

◇高等学校统计分析软件系列教材

统计软件 SPSS 在医学中的应用实例教程

孔晓荣 张星光 主编

清华大学出版社
北京交通大学出版社

·北京·

内 容 简 介

SPSS (Statistical Product and Service Solution) 是世界上著名的统计分析软件之一, 具有界面友好、操作简便易学、统计功能强大等特点, 广泛应用于社会科学和自然科学的各个领域。本书基于 SPSS 15.0 for Windows 版本内容, 从医学统计实际问题出发, 详细介绍了 SPSS 在医学数据统计分析中的具体使用方法, 包括根据实际问题选择合适的统计方法, 多种统计方法的软件操作步骤及对统计分析结果的医学解释等内容。本书具有内容充实、语言简练、图文并茂、深入浅出等特点, 重点突出实用性, 通过大量医学数据统计实例, 为读者详细讲述 SPSS 在医学数据处理方面的应用方法。

本书适用于医学高等院校相关专业的师生, 医疗系统科研人员及医务工作者等, 具有很好的实用价值。

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

统计软件 SPSS 在医学中的应用实例教程/孔晓荣, 张星光主编. —北京: 清华大学出版社; 北京交通大学出版社, 2009.9

(高等学校统计分析软件系列教材)

ISBN 978-7-81123-735-1

I. 统… II. ①孔… ②张… III. 医学统计-统计分析-软件包, SPSS-高等学校-教材
IV. R195.1-39

中国版本图书馆 CIP 数据核字 (2009) 第 153289 号

责任编辑: 解 坤

出版发行: 清华大学出版社 邮编: 100084 电话: 010-62776969

北京交通大学出版社 邮编: 100044 电话: 010-51686414

印刷者: 北京交大印刷厂

经 销: 全国新华书店

开 本: 185×260 印张: 14 字数: 350 千字

版 次: 2009 年 9 月第 1 版 2009 年 9 月第 1 次印刷

书 号: ISBN 978-7-81123-735-1/R·2

印 数: 1~4 000 册 定价: 29.00 元

本书如有质量问题, 请向北京交通大学出版社质监组反映。对您的意见和批评, 我们表示欢迎和感谢。

投诉电话: 010-51686043, 51686008; 传真: 010-62225406; E-mail: press@bjtu.edu.cn。

编委会成员名单

主 编 孔晓荣 张星光

编 委 苏 俐 王学梅 郭淑怡 杨树伟

前言

随着信息时代的发展,社会各领域都存在大量的数据和信息需要人们收集和处理,在医学这一特殊领域中,大量的临床数据资料和医学科研中的实验数据等各种数据信息都需要进行科学的统计与分析,才能得到正确的统计分析结果和科学的结论。因此,对医学数据信息的采集、存储、整理和统计分析这一系列处理过程就显得非常重要,而 SPSS 正是一款能够进行数据信息处理的、功能强大的统计分析软件。

SPSS (Statistical Product and Service Solution) 是世界上著名的统计分析软件之一,具有界面友好、操作简便易学、统计功能强大等特点,广泛应用于社会科学和自然科学的各个领域。本书从医学统计实际应用角度出发,针对临床医务人员及医学科研人员在工作中经常遇到的医学统计学问题,详细介绍了 SPSS 软件在医学数据统计分析中的具体使用方法,包括根据实际问题选择合适的统计方法、多种统计方法的软件操作步骤及对统计分析结果的医学解释等内容。本书具有内容充实、语言简练、图文并茂、深入浅出等特点,具有很强的实用性和参考价值。

本书适用于医学高等院校相关专业的师生、医疗系统科研人员及医务工作者等,特别适合作为医学类本科生、研究生的辅助教材或医务工作者的培训教材和继续教育用书。建议读者应具备一定的医学统计知识基础,以便更好地利用 SPSS 的强大统计分析功能。

本书基于 SPSS 15.0 for Windows 版本编写,全书共分 13 章,通过大量医学统计具体案例详细介绍了在 SPSS 软件系统环境下医学统计分析中常用方法的具体操作步骤。各章节的主要内容如下。第 1 章为概述部分,概括介绍 SPSS 软件特点及界面。第 2 章主要介绍了 SPSS 软件中数据文件的建立、编辑及处理方法。第 3~12 章为本书的核心部分,主要介绍 SPSS 中常用的统计功能模块在医学统计应用中的具体操作步骤,其中包括 Frequencies、Descriptives、Explore 三种基本统计分析方法和 t 检验、方差分析、协方差分析、卡方检验、秩和检验、相关与回归分析、Logistic 回归分析、生存分析及聚类分析与判别分析等统计方法。第 13 章介绍了统计图的制作与编辑方法。书中实例的数据文件可通过出版社网站下载获得。

本书采用图文并茂的方式,重点突出实用性,力求为从事医务工作的读者提供一本简明易懂且非常切合实际需要的工具书,能够帮助读者迅速掌握统计软件 SPSS 在医学统计方面的应用方法。

本书由孔晓荣、张星光主编,在编写过程中参阅了大量的相关参考文献,还参阅了近年来国内医学统计学文献,在此向相关作者表示衷心的感谢。

由于编者水平有限,书中难免有疏漏与不足之处,还望读者提出宝贵意见,批评指正,以利提高。

编 者

2009 年 8 月

目 录

第 1 章 统计软件 SPSS 概述	1
1.1 SPSS 的特点与新增功能	1
1.2 SPSS 的运行环境	2
1.3 SPSS 的界面介绍	3
1.3.1 SPSS for Windows 的启动与界面介绍	3
1.3.2 SPSS for Windows 系统的退出	6
◇ 习题 1	7
第 2 章 数据文件的处理	8
2.1 数据文件的建立	8
2.1.1 定义变量	10
2.1.2 录入数据	13
2.1.3 保存数据	14
2.2 数据文件的编辑	14
2.2.1 查看信息	15
2.2.2 数据的定位	16
2.2.3 变量的插入与删除	16
2.2.4 观察量的插入与删除	17
2.3 数据文件的预处理	18
2.3.1 按照某变量的值排序	18
2.3.2 数据文件的拆分	18
2.3.3 数据文件的合并	19
◇ 习题 2	21
第 3 章 基本统计分析	22
3.1 Frequencies 统计分析	23
3.2 Descriptives 统计分析	27
3.3 Explore 统计分析	29
◇ 习题 3	33
第 4 章 t 检验	34
4.1 单个样本 t 检验	35
4.2 两独立样本 t 检验	37
4.3 配对样本 t 检验	40

◇ 习题 4	42
第 5 章 方差分析	44
5.1 方差分析简介	44
5.1.1 方差分析中的常见术语	44
5.1.2 方差分析的原理	45
5.1.3 方差分析的条件	45
5.2 单因素方差分析	45
5.3 随机区组设计的方差分析	54
5.4 析因设计的方差分析	62
5.5 重复测量设计的方差分析	65
◇ 习题 5	71
第 6 章 协方差分析	73
6.1 协方差分析简介	73
6.1.1 基本思想	73
6.1.2 应用条件	73
6.2 完全随机设计的协方差分析	73
6.3 随机区组设计的协方差分析	78
◇ 习题 6	81
第 7 章 χ^2 检验	83
7.1 行 $\times$ 列资料的 χ^2 检验	83
7.2 四格表资料的 χ^2 检验	88
7.3 配对资料的 χ^2 检验	93
◇ 习题 7	97
第 8 章 秩和检验	99
8.1 两个独立样本比较的秩和检验	100
8.2 多个独立样本比较的秩和检验	103
8.3 两个相关样本比较的秩和检验	105
8.4 多个相关样本比较的秩和检验	108
◇ 习题 8	110
第 9 章 相关与回归分析	112
9.1 相关分析	112
9.1.1 双变量相关分析	112
9.1.2 偏相关分析	119
9.2 回归分析	122
9.2.1 线性回归分析	122
9.2.2 曲线拟合	136
◇ 习题 9	141
第 10 章 Logistic 回归分析	143

10.1 非条件 Logistic 回归分析.....	143
10.2 条件 Logistic 回归分析.....	153
◇ 习题 10	155
第 11 章 生存分析	157
11.1 生存分析方法简介.....	157
11.1.1 基本概念	157
11.1.2 生存分析的主要内容与研究方法	158
11.2 Life Tables 过程	158
11.3 Kaplan-Meier 过程	162
11.4 Cox 回归模型.....	169
◇ 习题 11	175
第 12 章 聚类分析与判别分析	178
12.1 聚类分析.....	178
12.1.1 K-Means 聚类法	178
12.1.2 Hierarchical 聚类法	182
12.2 判别分析.....	187
◇ 习题 12	193
第 13 章 统计图	195
13.1 统计图的绘制.....	195
13.1.1 条形图.....	196
13.1.2 饼图	199
13.1.3 线图	201
13.1.4 散点图.....	204
13.1.5 箱图	204
13.1.6 面积图.....	207
13.1.7 误差条图	209
13.1.8 直方图.....	211
13.2 统计图的编辑.....	213
◇ 习题 13	214
参考文献	216

第 1 章

统计软件 SPSS 概述

社会科学统计软件包 SPSS (Statistical Package for Social Science) 是一种集成化的计算机数据统计应用软件。1968 年, 美国斯坦福大学的三位学生开发了最早的 SPSS 统计软件, 并于 1975 年在芝加哥成立 SPSS 公司, 至今已有 40 多年的发展历史。SPSS 广泛应用于商业、金融、医疗卫生、市场研究、体育、农林业、科研、教育等多个行业, 应用范围遍及了自然科学、技术科学及社会科学的各个领域。随着应用领域的不断扩大, SPSS 改名为 Statistical Product and Service Solution, 即统计产品与服务解决方案, 简称仍为 SPSS。

SPSS 与 SAS、SYSTAT 并称为国际上最有影响的三大优秀统计软件。相比而言, SAS 具有完善的数据管理和统计分析功能, 是熟悉统计学并擅长编程的专业人员的首选, 而 SPSS 具有操作简便、易学易用、分析结果清晰直观的特点, 是非统计学专业人员的首选。而且 SPSS 主要针对社会科学研究领域开发, 因而更适合应用于教育科学研究, 是教育科研人员必备的科研工具。

1.1 SPSS 的特点与新增功能

SPSS for Windows 具有如下特点。

1. 操作简单

自从 1995 年 SPSS 公司与微软公司合作开发 SPSS 界面后, SPSS 界面变得越来越友好, 操作方法也变得更为简单。SPSS for Windows 界面完全是菜单式, 除了数据录入及部分命令程序等少数输入工作需要键盘输入外, 大多数操作通过鼠标单击菜单命令、相关按钮、对话框中的选项等即可完成。

2. 无须编程

具有第四代语言的特点, 只要了解统计分析的原理, 无须精通统计方法的各种算法, 即可得到需要的统计分析结果。对于常见的统计方法, SPSS 的命令语句、子命令及功能项的选择不必花费很多时间输入而直接可以通过相关对话框的操作来完成。

3. 功能强大

集数据录入、资料编辑、数据管理、统计分析、报表制作、图形绘制于一体。SPSS 软件统计功能包括了统计学中的绝大多数项目, 如相关分析、回归分析、方差分析、卡方检验、t 检验和非参数检验等; 也包括近期发展的多因素统计分析方法技术, 如多元回归分析、聚类分析、判别分析、主成分分析、因子分析、聚类分析、非线性回归和 Logistic 回归

等方法,并能显示正态分布图、直方图、散点图等各种统计图表。

4. 方便的数据接口

可以直接读取及输出多种格式的文件。例如,由 FoxPro 产生的 *.dbf 文件、由文本编辑器生成的 ASCII 数据文件及 Excel 文件 (*.xls) 等均可转换成可供分析的 SPSS 数据文件,而由 SPSS 软件生成的数据文件也可以方便地转化为相应的其他数据文件。

SPSS 统计软件从 1968 年开发以来多次推出新的版本,2006 年 9 月推出了第 15 版。SPSS 15.0 for Windows 版本除了保持以往版本的特点以外,在许多功能方面都得到了扩展和加强,本书是以 SPSS 15.0 for Windows 版本为基础编写的,主要介绍如何利用计算机统计软件 SPSS 对医学临床、实验及医学科研中得到的数据资料进行统计处理,根据统计结果进行科学的分析与解释。

下面对 SPSS 15.0 for Windows 版本的新增功能作一简要介绍。

① 和以往版本相比,SPSS 15.0 for Windows 除了可以定义变量的标准属性外,还能创建变量的自定义属性。它允许利用变量集控制数据编辑窗口和对话框中的变量,而以往版本只可在对话框的变量列表框中实现对变量集的控制,此新增功能能够更方便地使用含有多变量的数据文件。

② 新增了数据库导出向导功能,可以用此功能创建新数据库表,还可以对指定域值进行替换,对已存在的表中添加变量等,无须编写 SQL 语句。SPSS 15.0 for Windows 支持将导出的数据文件保存为 CSV 格式文件,即逗号分割值数据文件,这是一种常用的数据文件格式。

③ SPSS 15.0 for Windows 的图表功能变化较大,图表界面更加友好,还可以定制个性化的图表,创建通用图表的模式也更加多样化,如输出结果文件支持 PDF 格式。SPSS 15.0 for Windows 还允许自定义控制图形的规则,帮助快速识别不可控制的点。

④ SPSS 15.0 for Windows 进一步扩展了图形生成器的功能,可创建直方图、箱图、金字塔图、高低图等多种类型的图形,还提供了以往版本中没有的图形类型(如对偶图等)。SPSS 15.0 for Windows 改进了一些统计模块的功能,如有序结果变量回归、广义估计方程。

1.2 SPSS 的运行环境

1. SPSS 15.0 for Windows 的硬件运行环境

SPSS 15.0 for Windows 的硬件运行环境要求不是很高,但由于涉及的数据量比较多,为了使用方便也需要一定程度的硬件运行环境。

① CPU 的要求在 Pentium II 以上。

② 软件运行时的内存空间至少为 256 MB 以上。

③ 不同版本的 SPSS 及同一版本所选的模块多少所需要的硬盘空间也是不同的。一般来讲,如果安装 SPSS 全部模块,至少需要 500 MB 左右的硬盘剩余空间。

④ CD-ROM 光盘驱动器,这是用光盘安装 SPSS 的基本要求,如果在网上安装 SPSS 则不需要光盘驱动器。

⑤ S-VAD 显示器和 Windows 98/2000/XP/2003 兼容的图形适配卡。

⑥ 支持 TCP/IP 网络协议的网络适配卡,用于访问 SPSS 公司的服务器,以获得服务和

软件的升级。

2. SPSS 15.0 for Windows 的软件运行环境

SPSS 15.0 for Windows 可以在 Windows 98/2000/2003/XP 操作系统环境下使用，可以使用中文输入变量名等资料数据，但是分析结果还不能汉化。

需要安装 SPSS Data Access Pack，该软件提供不同的数据类型和不同数据库共享的解决方案。一般还需要安装 Adobe Reader 软件，便于阅读 PDF 格式的帮助用户文件和 SPSS 相关文件。

1.3 SPSS 的界面介绍

由于 SPSS 公司与微软公司合作开发了 SPSS 界面，SPSS 界面和其他基于 Windows 操作系统环境下开发的应用程序一样，具有一致化的 Windows 应用程序界面，操作方法也变得更为简单。

1.3.1 SPSS for Windows 的启动与界面介绍

SPSS 软件安装好后，在【开始】主菜单中的【程序】选项中可以找到 SPSS 15.0 for Windows 图标，单击其级联菜单中的【SPSS 15.0 for Windows】启动快捷图标，即可启动 SPSS 15.0 for Windows 程序。首先出现的是 SPSS 文件对话框，如图 1-1 所示。

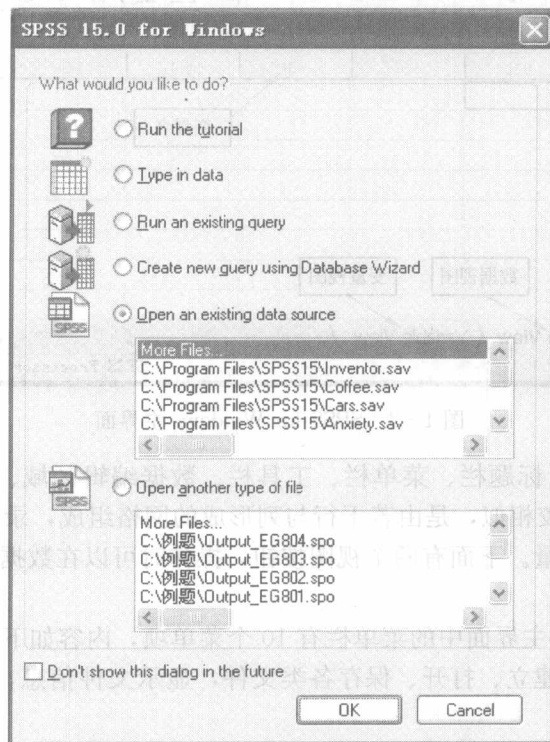


图 1-1 SPSS 文件对话框

对话框中的 What would you like to do 选项组内有 6 个单选按钮，选择不同的选项可以打开不同类型的文件。若想在以后打开 SPSS 15.0 for Windows 时不再显示此对话框，则选

择此对话框最下端的 Don't show this dialog in the future 复选框。

对话框中的 6 个单选按钮和含义现说明如下。

① Run the tutorial: 运行操作指导。若选择此项,则可以浏览操作指导。

② Type in data: 在数据窗口输入数据选项。选择此项则显示数据编辑窗口,用户可输入新的数据来建立数据文件。

③ Run an existing query: 运行一个已存在的问题文件。若选择此项则要求选择一个扩展名为 .sqp 的文件。

④ Create new query using Database Wizard: 使用数据库向导创建一个新的数据文件。

⑤ Open an existing data source: 打开一个已存在的数据文件。使用该选项能打开一个扩展名为 .sav 的文件。在列表框内显示了所有的数据文件列表及近期曾打开过的数据文件,可直接在列表中选择文件。

⑥ Open another type of file: 打开一个其他类型的文件。

当选择新建或打开一个扩展名为 .sav 的数据文件时,则显示 SPSS 软件的主界面窗口,如图 1-2 所示。

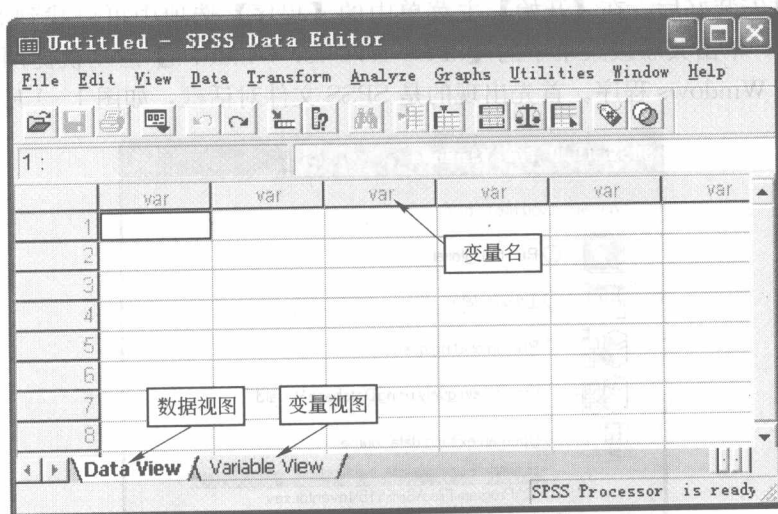


图 1-2 SPSS for Windows 主界面

SPSS 的主界面包括标题栏、菜单栏、工具栏、数据编辑区域、视图按钮、状态栏等。数据编辑区和 Excel 比较相似,是由若干行与列形成的网格组成,录入时一般每行对应一条记录,每列对应一个变量。下面有两个视图按钮,操作时可以在数据视图、变量视图两种视图之间切换,非常方便。

SPSS for Windows 主界面中的菜单栏有 10 个菜单项,内容如下。

① File: 主要包括建立、打开、保存各类文件,显示文件信息,打印文件,退出 SPSS 程序等命令。

② Edit: 主要包括剪切、复制、粘贴、删除、查找文本等命令。

③ View: 主要包括状态栏、工具栏、数据视图和变量视图的显示等命令。

④ Data: 主要包括数据文件的建立和编辑等命令。

⑤ Transform: 主要包括数据转换操作等命令。

- ⑥ Analyze: 主要包括统计分析操作等命令。
- ⑦ Graphs: 主要包括建立和编辑统计图形的操作命令。
- ⑧ Utilities: 主要包括实用工具的使用命令。
- ⑨ Window: 主要包括窗口的操作命令。
- ⑩ Help: 主要包括 SPSS 的使用帮助及 SPSS 的相关信息。

SPSS 主要有 3 个窗口: 数据编辑窗口 (Data Editor)、结果输出窗口 (Output Viewer) 和程序编辑窗口 (Syntax Editor)。启动 SPSS 时首先进入的是数据编辑窗口, 可以进行数据录入、编辑、存储等操作, 如图 1-3 所示; 结果输出窗口中可以对输出结果进行编辑、保存、传输等操作, 如图 1-4 所示; 程序编辑窗口中可以对 SPSS 程序进行编辑、存储和提交执行操作, 适合专业人员使用, 如图 1-5 所示。

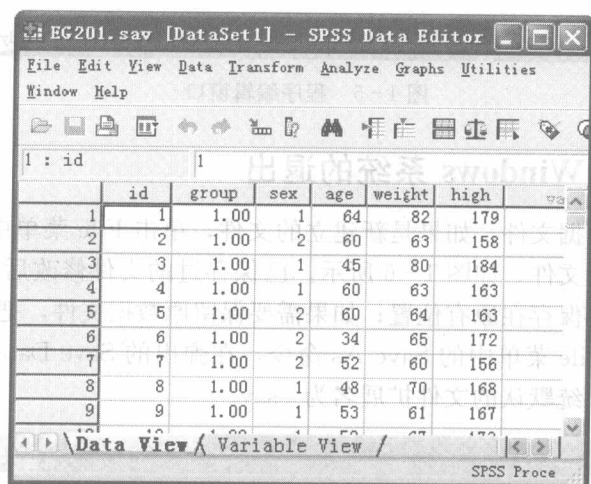


图 1-3 数据编辑窗口

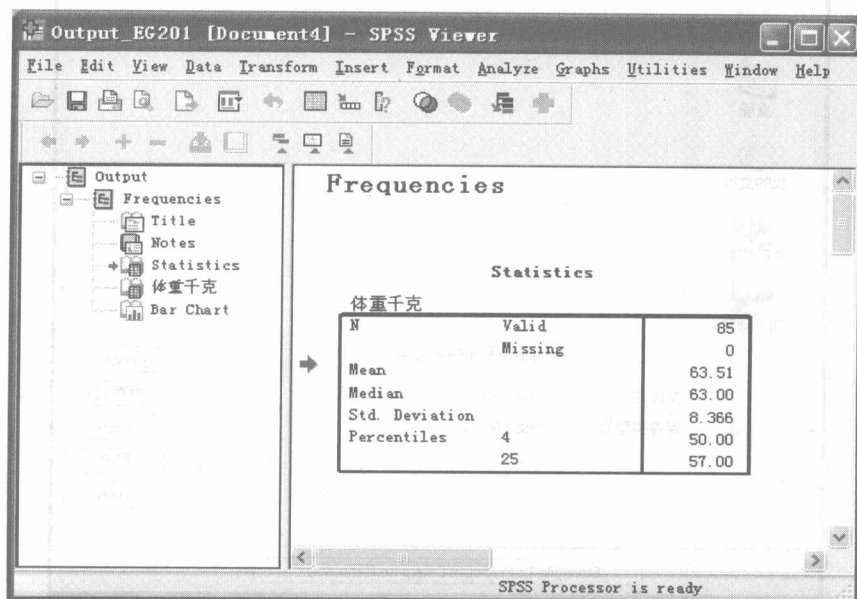


图 1-4 结果输出窗口

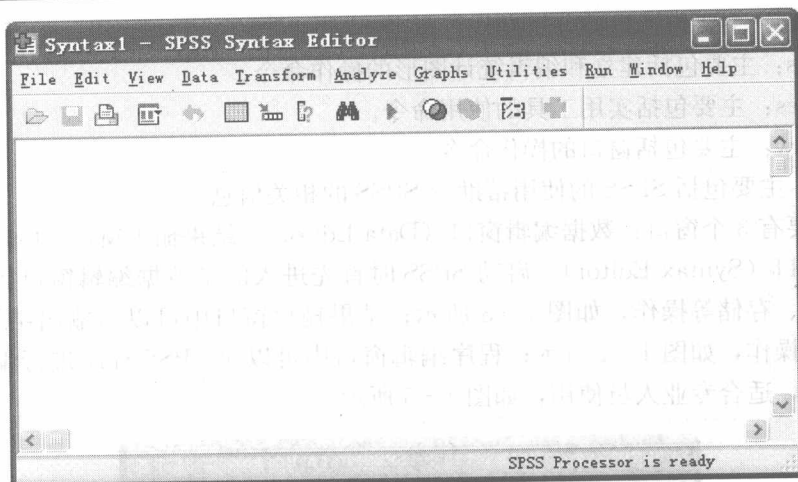


图 1-5 程序编辑窗口

1.3.2 SPSS for Windows 系统的退出

首先要保存相关数据文件，如果是新建的文件，单击 File 菜单中的 Save 命令，在弹出的保存对话框中保存文件，如图 1-6 所示。已保存过的文件修改后可以直接单击工具栏中【保存】按钮，文件保存在原有位置；如果需要保留原数据文件，把修改后的文件保存成另一个文件，则单击 File 菜单中的 Save As 命令，在弹出的 Save Data As 对话框中保存文件，数据文件保存时系统默认的文件扩展名为 .sav。

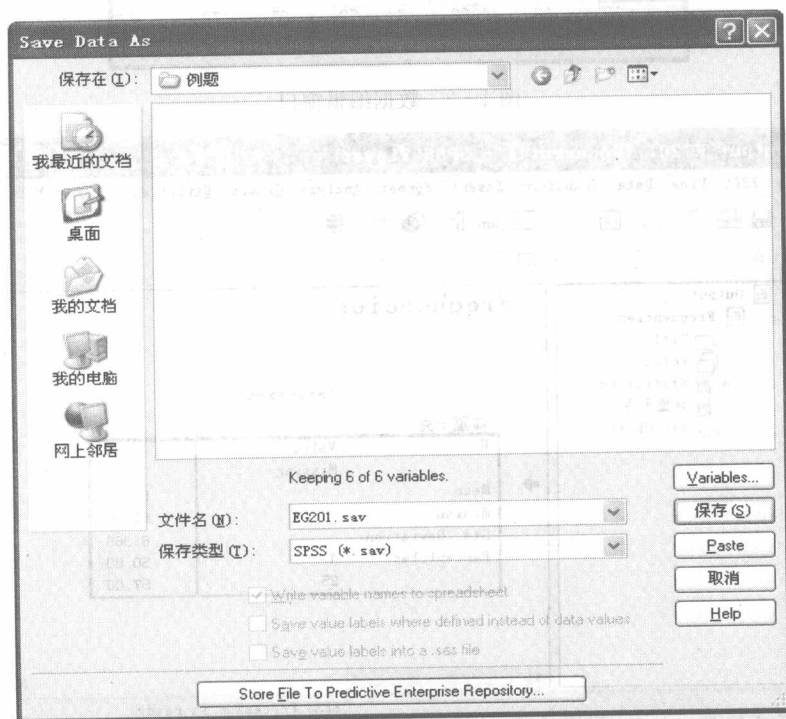


图 1-6 文件保存对话框

保存好文件后,如果要退出 SPSS 系统,单击 File 菜单中的 Exit 命令,即可退出 SPSS 系统;也可单击 SPSS 窗口右上角的关闭按钮退出系统。



习题 1

1. 了解并熟悉 SPSS 软件的主界面,熟悉 10 个菜单项中的主要命令。
2. 观察 Analyze 菜单,试统计 SPSS 软件能够完成统计学中的哪些统计分析方法?
3. 打开随书所附例题文件中 EG201.sav 数据文件,将文件另存到硬盘中。

第 2 章

数据文件的处理

在对医学研究中收集的数据利用 SPSS 软件进行统计分析前，首先要输入数据建立数据文件，可以随时对数据文件进行存取、修改和统计分析等处理工作。在输入数据时要遵循“配对资料、相关资料直接按照分组情况输入数据；完全随机分组资料必须输入分组数据，即所有原始数据输入一列（因变量），再输入一列或多列分组数据（因素变量）”的原则输入数据，为今后进行正确的统计分析做好数据准备工作，具体输入方法在后面的各种统计方法中再作详细介绍。

2.1 数据文件的建立

在建立数据文件之前，首先要进行数据编码，下面通过一个例子详细介绍数据文件的建立过程。

实例

【例 2-1】 现测得 85 例人员的有关年龄、体重、身高等数据，分成对照组和实验组两个组，其原始数据如表 2-1 所示。

表 2-1 85 例人员的有关年龄、体重、身高等数据

编 号	分 组	性 别	年 龄	体重/千克	身高/厘米
1	1.00	1	64	82	179
2	1.00	2	60	63	158
3	1.00	1	45	80	184
4	1.00	1	60	63	163
5	1.00	2	60	64	163
6	1.00	2	34	65	172
7	1.00	2	52	60	156
8	1.00	1	48	70	168
9	1.00	1	53	61	167
10	1.00	1	58	67	173
11	1.00	1	49	71	174
12	1.00	2	63	67	158