

SPSS

13.0

在生物统计中的应用

张力 编著



厦门大学出版社
XIAMEN UNIVERSITY PRESS

SPSS 13.0 在生物统计中的应用

张 力 编著

厦门大学出版社

图书在版编目(CIP)数据

SPSS 13.0 在生物统计中的应用/张力编著. —厦门:厦门大学出版社,2006.5
ISBN 7-5615-2574-5

I. S… II. 张… III. 生物统计-统计分析-软件包, SPSS 13.0 IV. Q-332

中国版本图书馆 CIP 数据核字(2006)第 045527 号

厦门大学出版社出版发行

(地址:厦门大学 邮编:361005)

<http://www.xmupress.com>

xmup@public.xm.fj.cn

厦门昕嘉莹印刷有限公司印刷

2006 年 6 月第 1 版 2006 年 6 月第 1 次印刷

开本:787×1092 1/16 印张:12

字数:304 千字 印数:0001—3000 册

定价:18.00 元

如有印装质量问题请与承印厂调换

前 言

SPSS 是世界上通用的优秀统计软件包之一,广泛应用于社会科学、自然科学的各个领域。本书以简明、实用的方式,应用大量的实例介绍了 SPSS 13.0 在生物统计中的常用分析方法,包括次数分布表和常用统计图的编制、 t 检验、不同试验设计方法的方差分析、协方差分析、 χ^2 检验、相关分析、回归分析、二项分布检验、聚类分析、主成分分析、半数效量的计算等,并对输出结果作出统计学的分析与推断。同时,本书还以实例简要介绍了 Excel 电子表格在生物统计中的应用,包括 t 检验、方差分析、线性相关回归分析和次数分布表与直方图的编制。

本书可作为农林院校生物学、动物医学、动物科学、农学、水产和蜂学等专业本科生学习生物统计课程的补充教材,并可供论文写作时进行数据统计之用,也可作为农业科研人员进行数据统计分析的参考资料。

在本书的编写过程中,承蒙福建农林大学周以飞教授仔细审阅修改,热情鼓励与指导,并提出了许多具有建设性的宝贵意见。徐秀琼老师在本书的资料整理和图表编制上做了大量的工作。此外,本书还参阅了相关的教材、著作和论文。在此一并表示真挚的感谢。

在编写中,编者虽然已尽心竭力,但限于水平与经验,书中错误与不当之处在所难免,敬请读者指正。

张 力
于福建农林大学
2006 年 3 月

目 录

前 言

第一章 SPSS for Windows 基本知识	(1)
一、SPSS 13.0 for Windows 的启动	(1)
二、数据编辑窗	(2)
三、数据文件的建立	(5)
第二章 资料的整理	(8)
一、次数分布表的编制	(8)
二、常用统计图	(14)
三、动物完全随机分组	(28)
第三章 t 检验	(31)
一、样本平均数与总体平均数差异显著性检验	(31)
二、非配对设计两样本均数差异显著性检验	(33)
三、配对设计两样本均数差异显著性检验	(37)
第四章 方差分析	(42)
一、完全随机设计的单因素方差分析	(42)
二、随机单位组设计的两因素无重复观察值方差分析	(48)
三、交叉分组的两因素有重复观察值方差分析	(54)
四、系统分组的两因素有重复观察值方差分析	(58)
五、拉丁方设计的方差分析	(62)
六、交叉设计的方差分析	(65)
七、正交设计的方差分析	(68)
八、方差分析中的数据转换	(78)
第五章 协方差分析	(81)
一、单向分组资料的协方差分析	(81)
二、双向分组资料的协方差分析	(86)
第六章 卡方(χ^2)检验	(91)
一、 2×2 列联表的独立性检验	(91)
二、 $R \times C$ 列联表的独立性检验	(95)
三、配对 χ^2 检验	(100)
第七章 线性相关分析	(104)
一、两个变量间的相关分析	(105)

二、两个等级(秩)变量间的相关分析	(108)
三、多个变量间的相关分析	(112)
四、偏相关分析	(114)
第八章 回归分析	(120)
一、一元线性回归分析	(121)
二、多元线性回归分析	(125)
三、逐步回归分析	(129)
四、曲线回归分析	(133)
第九章 二项分布检验	(137)
第十章 半数效量的计算	(140)
第十一章 聚类分析	(146)
一、指标(或变量)聚类	(146)
二、样品聚类	(154)
第十二章 主成分分析	(159)
第十三章 Excel 常用生物统计功能简介	(165)
一、分析工具库	(165)
二、 t 检验	(166)
三、方差分析	(169)
四、线性相关回归分析	(173)
五、次数分布表与直方图的编制	(179)
参考文献	(183)

第一章 SPSS for Windows 基本知识

SPSS (Statistical Package for Social Science) 统计软件全称为社会科学统计软件包,是目前世界上公认的用户最多、功能最强大的优秀通用统计软件之一。它不仅适用于社会科学,同样可应用于生物学、动物医学、动物科学、作物学、水产和蜂学等自然科学的各个领域。

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 14.0 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。在 SPSS 中,无须编写程序,只需粗通统计分析原理,就完全可以在 Windows 下通过菜单的选择、对话框的操作得到统计分析结果。SPSS 分析方法丰富,提供了从简单描述统计分析到多因素统计分析的方法,能够基本满足生物统计分析工作的需要,是非专业统计人员的首选统计软件。

SPSS 公司的网址为 www.spss.com,其中国公司的网址为 www.spssbj.com.cn,可在该网站的下载专区下载到 SPSS 相关软件,并了解该软件的功能,或在其论坛向同行专家请教。

本书是以 SPSS 13.0 为基础编写的,书中所有统计分析内容均可在 SPSS 10.0~14.0 版本中应用。

一、SPSS 13.0 for Windows 的启动

SPSS 13.0 可在 Windows 98/2000/me/xp 操作系统下使用。

在 SPSS 13.0 软件安装完成后,单击 Windows 下的“开始”按钮,将光标移到“程序”选项上,便出现“SPSS for Windows”菜单,在下拉式“SPSS 13.0 for Windows”及“SPSS 13.0 Production Mode Facility”菜单中,选中并单击“SPSS 13.0 for Windows”便可启动 SPSS 13.0(图 1-1)。

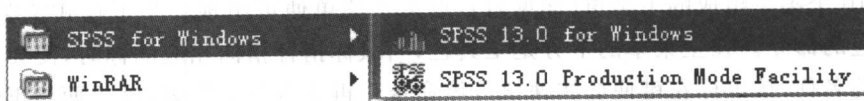


图 1-1 启动 SPSS 13.0 的方法

SPSS 启动后,在屏幕上显示 SPSS for Windows 主界面的导航对话框,如图 1-2 所示。常用的三个命令菜单为:Type in data(输入数据)、Open an existing data source(打开已存在的数据文件)、Run the tutorial(运行指导)。用鼠标选中相应命令,便可进入有关操作,也可点击 Cancel(取消)按钮,直接进入 SPSS 数据编辑窗口(图 1-3)进行相关操作。

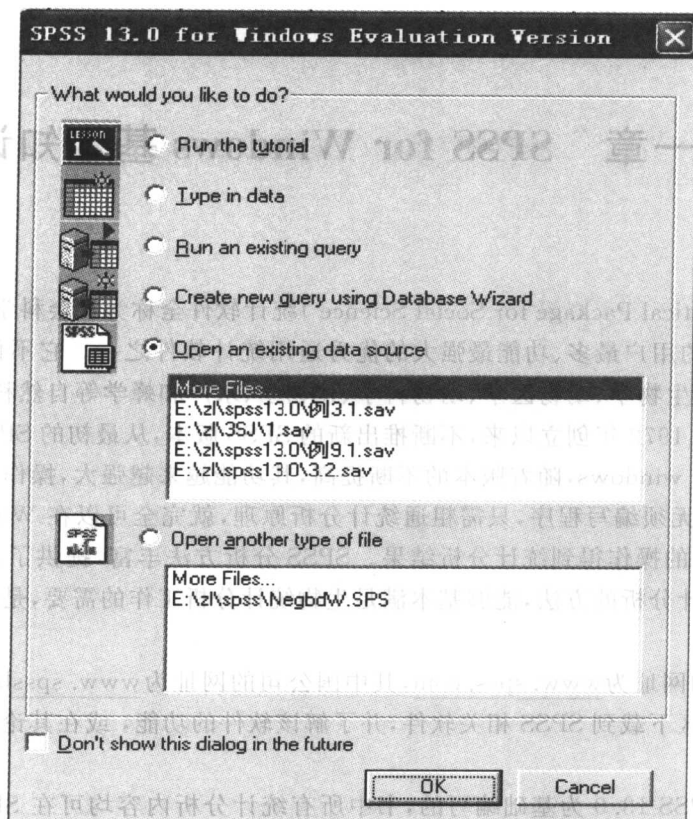


图 1-2 SPSS 启动后所弹出的导航对话框

二、数据编辑窗

1. 数据编辑窗简介

SPSS 数据编辑窗(SPSS Data Editor)主要作用是编辑数据文件,它分为两个工作表:数据视图(Data View)工作表和变量视图(Variable View)工作表(图 1-3)。编辑窗最上方一行是由 10 个菜单组成的主菜单(详见后),点击主菜单便可出现一个下拉式菜单,在其中可选取相应的命令。在主菜单的下方是工具栏,当鼠标指针指向相应图标时,便会自动显示工具按钮的名称,按鼠标左键便会快速执行该命令。再下方是工作表,该界面与 Excel 极为相似,在工作表中可以编辑各种数据,其上方为列标栏,左边为行标栏。在列标栏中可选相应的变量名进行编辑。工作表最下方是状态栏(Data View/Variable View),可选择点击进入数据视图或变量视图编辑窗。

数据视图(Data View)工作表主要用于数据输入、修改及查看资料中各变量的内容。变量视图(Variable View)工作表主要有以下功能:命名(Name)变量,从 SPSS 12.0 起变量名的最大长度可多达 64 个字符(32 个汉字);定义变量的类型(Type)、宽度(Width)以及小数

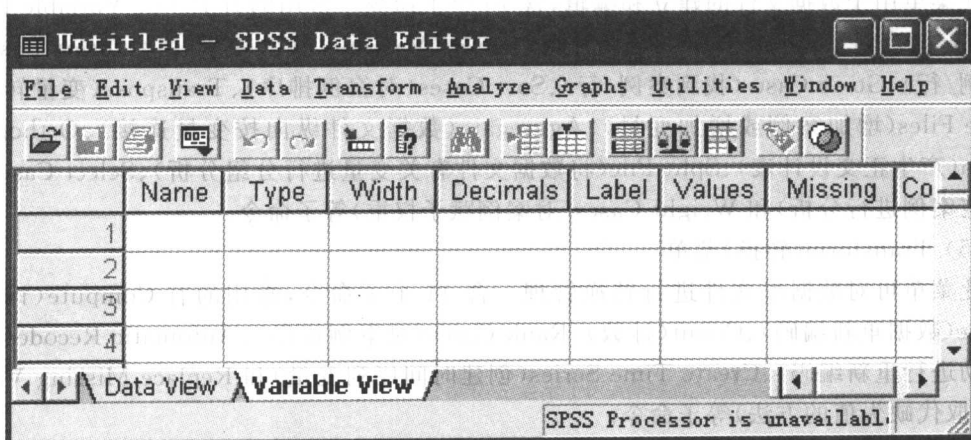
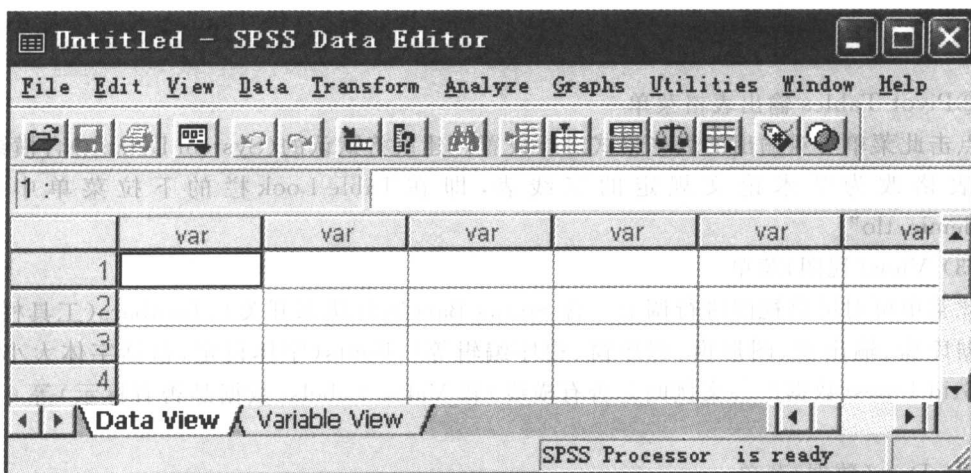


图 1-3 数据视图(Data View)和变量视图(Variable View)工作表

点的位数(Decimals);对变量名进行标记(Label)等。

2. 数据编辑窗 10 个主菜单内容简介

(1) File(文件)菜单

此菜单可对数据库文件(*.sav)或程序文件(*.sps)进行打开、保存、打印等管理。含 16 个子命令,常用的有 New(建立新数据库或程序文件)、Open(打开各种类型文件)、Save(保存文件)、Save As(另保存为)、Display Data File Information(显示所有数据库文件信息)、Print(打印文件)和 Exit(退出 SPSS 系统)等命令。

(2) Edit(编辑)菜单

此菜单可对数据库文件进行编辑管理,含 Undo(撤消/恢复)、Cut(剪切)、Copy(复制)、Paste(粘贴)、Clear(清除)、Find(查找)和 Options(选择项)等 9 项子命令。其中 Options 子命令又含有 10 个选择项设置,其中较重要、建议重新设置的是:

① Viewer 视图菜单

点击此菜单可对输出字体及大小进行设置。建议将标题字体(Title Font)设置为“Ari-

al”,大小为 14 左右;将输出文本字体(Text Output Font)设置为“Courier New”,大小为 12 左右。

②Pivot Tables 输出表格菜单

点击此菜单可对输出的表格格式进行设置。建议将默认的(System Default)含横线、竖线的表格改为学术论文规定的三线表,即在 Table Look 栏的下拉菜单中选择“Academic. tlo”

(3) View(视图)菜单

此菜单可对屏幕视图进行调节。含 Status Bar(运行状态开关)、Toolbars(工具栏。用于数据库窗、输出窗、图形窗、程序窗、程序编辑等)、Fonts(字体设定,包括字体大小与样式)、Grid Lines(数据库各案例间是否有横线)和 Value Labels(数据是否有标示)等 6 个子命令。

(4) Data(数据)菜单

此菜单用于数据文件的建立和编辑,含 16 个子命令。常用的有 Define Variable Properties(定义变量)、Define Dates(定义时间)、Insert Variable(插入变量/列)、Insert Case(插入案例/行)、Go to Case(找到案例/行)、Sort Cases(对案例排序)、Transpose(变量转换)、Merge Files(增加案例或增加变量)、Aggregate(数据文件纵向按变量连接)、Orthogonal Design(产生正交设计表)、Split File(将数据文件有关变量进行分组分析)、Select Cases(选择特定案例进行分析)和 Weight Cases(对案例赋予权重)等子命令。

(5) Transform(转换)菜单

此菜单可对数据库文件进行转换管理。含 11 个子命令,常用的有 Compute(计算)、Recode(数据重新编码)、Count(计数)、Rank Cases(对案例排序)、Automatic Recode(对变量自动进行重新编码)、Create Time Series(创建时间序列变量)和 Replace Missing Values(选用取代缺失值的方法)等子命令。

(6) Analyze(统计分析)菜单

此菜单包含 SPSS 的最重要部分,它提供了功能强大、齐全、应有尽有的统计分析方法,含 18 个主命令、80 多种统计分析方法。

(7) Graphs(绘图)菜单

此菜单拥有约 50 余种绘图功能,含有 21 个子菜单。最常用的有 Bar...(直条图)、Line...(线图)、Pie...(圆形图/饼图)、Histogram...(直方图)等。

(8) Utilities(实用命令)菜单

此菜单包含 Variables(变量信息)、Define Sets(定义变量集)、Use Sets(使用中的变量集)、Run Script(运行程序)和 Menu Editor(主菜单编辑器)等 8 个子命令。

(9) Windows(窗口)菜单

此菜单包含 Minimize SPSS(将 SPSS 最小化)子命令和已打开的数据库文件名。

(10) Help(帮助)菜单

此菜单下含 Topic(检索用主题)、Tutorial(自学教程)、Statistics Coach(SPSS 各统计分析方法指导)、SPSS Home Page(点击此命令可进入 spss.com 主页)和 About...(关于 SPSS 此版本的信息)等 8 个子命令。

三、数据文件的建立

在利用统计软件进行统计分析之前,必须建立可分析的数据文件,即把生产科研工作过程中采集的各种信息、数据录入到 SPSS 数据编辑窗,并进行保存。数据文件是由变量和变量的数据组成的,因此,编辑数据文件无非是要定义变量和输入数据。下面分别介绍。

1. 定义变量

在输入数据之前,为了计算及结果输出阅读方便,应首先定义数据的名称(变量名)、类型、大小和性质。

启动 SPSS 进入 SPSS 数据编辑窗(系统默认为 Data View 工作表),单击工作表左下方的 Variable View 进入变量视图(定义变量)窗口(图 1-3),可以对变量进行定义。变量视图工作表主要有以下功能:

(1) Name(变量名)

用于设定变量名,SPSS 中变量名可用英文或中文(SPSS 官方公布 10.07 以后版本 100%兼容中文)。从 SPSS 12.0 版起,变量名最多可以长达 64 个字节,完全可以满足各种情况下的需要。变量名首字符应该是英文或中文,其后可为数字等。对变量命名的目的只是为了在数据处理时有助于记忆,可根据资料的具体情况进行命名,以简明扼要、便于记忆为原则。

(2) Type(变量类型)

点击 Type,在其下出现一个矩形框,单击它会弹出定义变量类型的对话框,如图 1-4 所示。

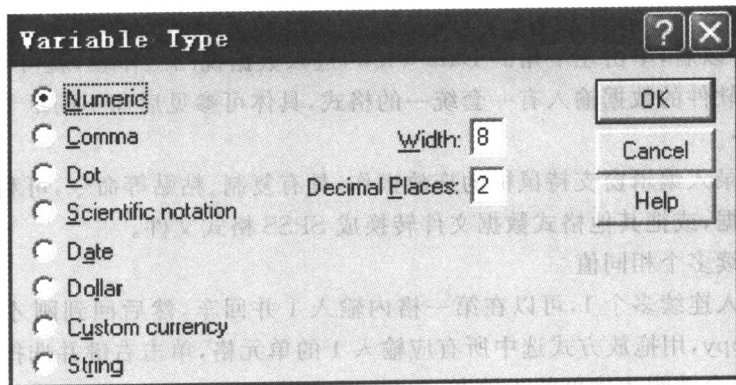


图 1-4 定义变量类型对话框

在图 1-4 的左边出现 8 种可供选择的变量类型,如表 1-1 所示。

系统默认变量为数值型(Numeric),一般情况下默认的数值型可以满足大多数资料的分析。

(3) Width(宽度)

设置变量运算宽度,系统默认为 8,一般不用更改。

表 1-1 8 种可供选择的变量类型

Numeric	标准数值型
Comma	带逗点的数值型(财务型),如 345,618.92
Dot	逗点作小数点的数值型,如 345.618,92
Scientific notation	科学记数法,如 1.3 E-0.7
Date	日期型
Dollar	带有美元符号的数值型
Custom currency	自定义型
String	字符串型

(4)Decimals(小数位数)

设置小数位,系统默认为 2,可依分析数值的具体情况更改。这里的小数点位并不影响运算过程的精度,只改变输入数据的显示宽度。

(5)Label(变量标签)

对变量名进行标记,以便在结果输出时方便阅读。

(6)Values(变量值标签)

用于注释变量值,尤其方便分类变量的输入和显示。

(7)Missing(定义变量缺失值)、Columns(定义显示列宽)、Align(定义显示对齐方式)、Measure(定义变量的测量尺度)等项

这几项均可遵从 SPSS 系统默认设置,一般情况下不必修改。

定义变量的具体操作可参见后面不同章节统计分析的数据输入相关部分。

2. 数据输入

变量定义好以后,单击左下角的 Data View,进入数据视图工作表,就可以输入数据了。

SPSS 统计软件的数据输入有一套统一的格式,具体可参见后面不同章节统计分析的数据输入相关部分。

SPSS 数据录入编辑窗支持鼠标的拖放操作,具有复制、粘贴等命令,可利用这些功能方便、快速输入数据,或把其他格式数据文件转换成 SPSS 格式文件。

(1)输入连续多个相同值

例:若需输入连续多个 1,可以在第一格内输入 1 并回车,然后回到刚才的单元格并单击右键,选择 Copy,用拖放方式选中所有应输入 1 的单元格,单击右键并选择 Paste,所有选中的单元格就都会被刚才拷贝的 1 填充。

(2)将 Excel 文件(*.xls)、数据库文件(*.dbf)调入 SPSS

①利用复制、粘贴功能调入

打开 Excel 数据库文件,选中所有的数据(不包括变量名),选择复制命令,然后切换到 SPSS,最好使第 1 行、第 1 列单元格成为当前单元格,然后执行粘贴命令,数据就会全部转入 SPSS,再定义相应的变量名即可。

②直接调入

启动 SPSS 后,在导航对话框(参见图 1-2)选中 Open an existing data source(打开已存在的数据文件),按 **OK** 按钮,弹出打开数据文件的对话框,如图 1-5 所示。

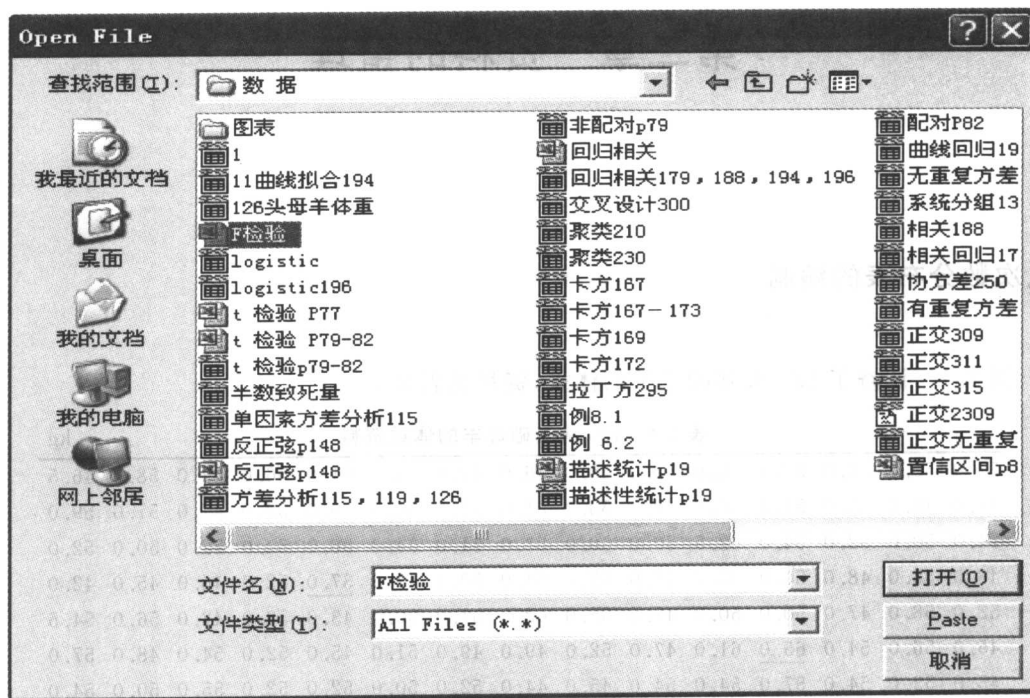


图 1-5 打开数据文件的对话框

在图 1-5 中的“查找范围(I)”右面的矩形中,选择指定文件所在的路径。在图下方“文件类型(T)”右面的矩形框中选择“All Files (*.*)”,在“查找范围(I)”下面的矩形框中,寻找所要读入的文件的文件名,并单击之,则指定文件名会出现在“文件名(N)”框中,点击 **打开** 按钮,弹出对话框,再按 **OK** 按钮,指定的数据文件就被读入 SPSS。

(3) 将 Word 中的数据表格调入 SPSS

如果 Word 表格中全部都是数值,则可以选中整个表格,选择拷贝命令,然后切换到 SPSS,再执行粘贴命令,数据就会全部转入 SPSS,并且原来的单元格会自动对应为 SPSS 中的单元格。此时再定义相应的变量名即可。

3. 数据文件的保存

单击菜单 File → Save(或 Save As),选择保存文件的地址和文件名填入对话框中。通常 SPSS 默认的数据文件类型为 *.sav(后缀的“.sav”不必输入,系统自动生成)。同时也可以保存为 Excel (*.xls)、数据库文件 (*.dbf) 和多种 SAS 文件,可在单击“Save As(另存为)”后弹出的对话框中单击“保存类型(T)”,在下拉列表菜单中选择不同数据文件名的后缀,并键入数据文件名后,按 **OK** 按钮,便可把数据存储为相应类型的数据文件。

第二章 资料的整理

一、次数分布表的编制

例 2.1 调查了 126 头基础母羊的体重,资料见表 2-1。

表 2-1 126 头基础母羊的体重资料 kg

53.0	50.0	51.0	57.0	56.0	51.0	48.0	46.0	62.0	51.0	61.0	56.0	62.0	58.0	46.5
48.0	46.0	50.0	54.5	56.0	40.0	53.0	51.0	57.0	54.0	59.0	52.0	47.0	57.0	59.0
54.0	50.0	52.0	54.0	62.5	50.0	50.0	53.0	51.0	54.0	56.0	50.0	52.0	50.0	52.0
43.0	53.0	48.0	50.0	60.0	58.0	52.0	64.0	50.0	47.0	37.0	52.0	46.0	45.0	42.0
53.0	58.0	47.0	50.0	50.0	45.0	55.0	62.0	51.0	50.0	43.0	53.0	42.0	56.0	54.5
45.0	56.0	54.0	65.0	61.0	47.0	52.0	49.0	49.0	51.0	45.0	52.0	54.0	48.0	57.0
45.0	53.0	54.0	57.0	54.0	54.0	45.0	44.0	52.0	50.0	52.0	52.0	55.0	50.0	54.0
43.0	57.0	56.0	54.0	49.0	55.0	50.0	48.0	46.0	56.0	45.0	45.0	51.0	46.0	49.0
48.5	49.0	55.0	52.0	58.0	54.5									

1. 数据输入

(1)启动 SPSS,点击 Variable View 进入定义变量工作表,用 Name 命令定义变量“体重”,小数位(Decimals)依题意定义为 1。如图 2-1 所示。

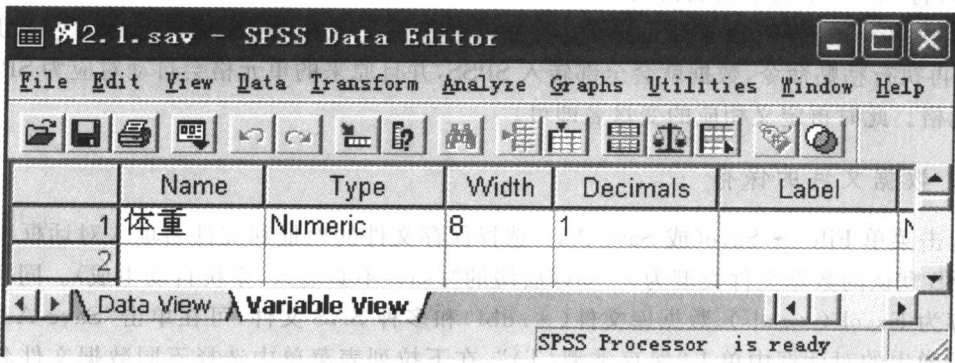


图 2-1 例 2.1 资料的变量命名

(2)点击工作表下方的 Data View 命令,进入“数据视图”工作表,按图 2-2 的格式将 126 头基础母羊的体重数据分别输入到变量名为“体重”的各个单元格内。

2. 求均数 (Mean)、标准差 (Std. Deviation)、最大值 (Maximum)、最小值 (Minimum) 和全距 (Range)

(1) 单击主菜单 Analyze(分析)→ Descriptive Statistics(描述性统计)→ Descriptives(描述), 则弹出对话框如图 2-3 所示, 选中变量“体重”, 按三角按钮将其置入 Variable[s]框内, 按 Options ... 按钮, 则弹出如图 2-4 所示对话框。

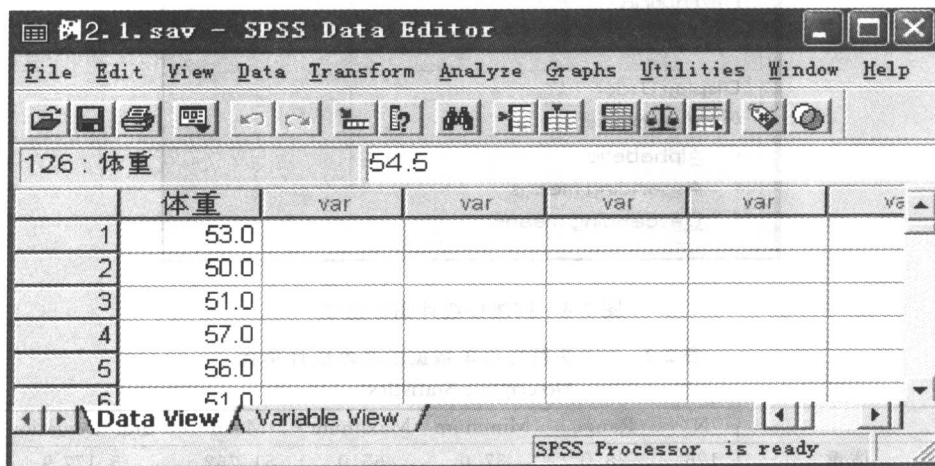


图 2-2 例 2.1 数据输入格式

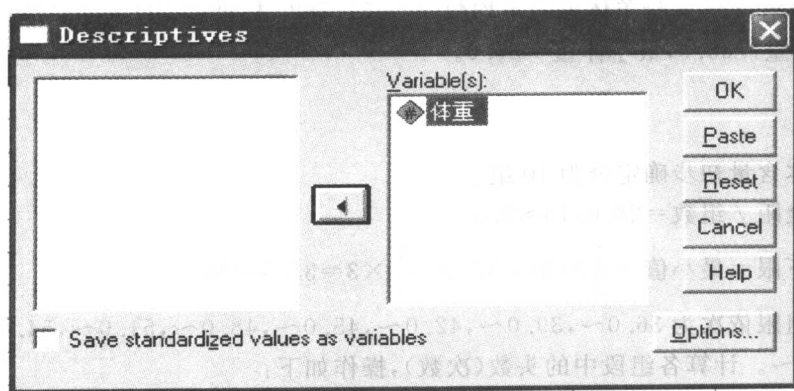


图 2-3 求“体重”的描述性指标

(2) 用鼠标选中 Mean(均数)、Std. deviation(标准差)、Maximum(最大值)、Minimum(最小值)和 Range(全距)。

(3) 点击 Continue 按钮, 返回图 2-3, 点击 OK 按钮则输出表 2-2 所示结果。

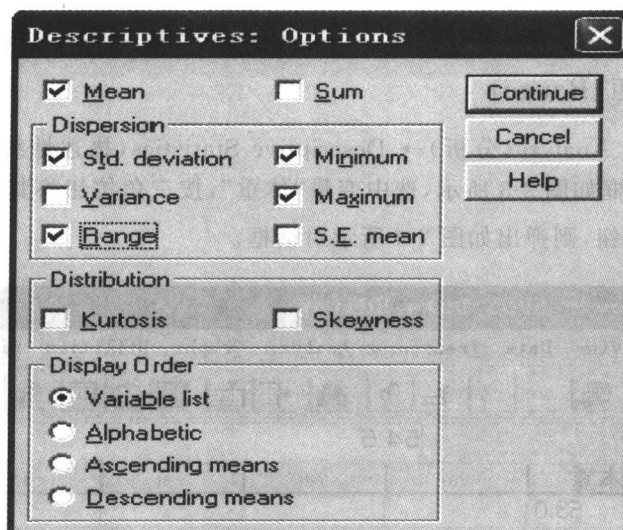


图 2-4 描述性统计指标选项

表 2-2 126 头基础母羊数据的基本统计指标

Descriptive Statistics						
	N	Range	Minimum	Maximum	Mean	Std. Deviation
体重	126	28.0	37.0	65.0	51.762	5.177 9
Valid N (listwise)	126					

表 2-2 表示:126 头母羊体重的平均数 $\bar{x} = 51.762$, 标准差 $SD = 5.177 9$, 全距 $R = 28.0$, 最大体重 = 65.0, 最小体重 = 37.0。

3. 分组

根据样本含量初步确定分为 10 组。

组距 = 全距 / 组数 = $28.0 / 10 \approx 3.0$

第一组下限 = 最小值 - $\frac{1}{2}$ 组距 = $37.0 - \frac{1}{2} \times 3 = 35.5 \approx 36$

分组的组限依次为 36.0~, 39.0~, 42.0~, 45.0~, 48.0~, 51.0~, 54.0~, 57.0~, 60.0~, 63.0~。计算各组段中的头数(次数), 操作如下:

(1) 单击主菜单 Transform(转换) → Compute(计算), 则弹出对话框(图 2-5)。

(2) 在 Target Variable(目标变量)栏键入变量名“次数”(用于标识观察值所属的组段), 在 Numeric Expression(数式表达)框内输入第一组代码“1”。

(3) 按图 2-5 左下方 If ... (如果) 按钮, 则弹出如图 2-6 所示对话框, 选中“Include if case satisfies condition”, 意思是对符合下列条件的观察值进行操作。

(4) 选中“体重”变量, 按三角按钮, 将其置入条件表达框内, 点击图 2-6 中的键盘键, 依次输入“ ≥ 36.0 & 体重 < 39.0 ”, 再按 Continue 按钮, 返回图 2-5, 按 OK 按钮。此时便会在数据编辑窗工作表变量名“体重”右边产生一个变量名为“次数”的新变量名。凡是体重在 36.0~39.0(不包含 39.0)的记录, 其相应的“次数”变量列的值为 1(代码)。即凡是观察值

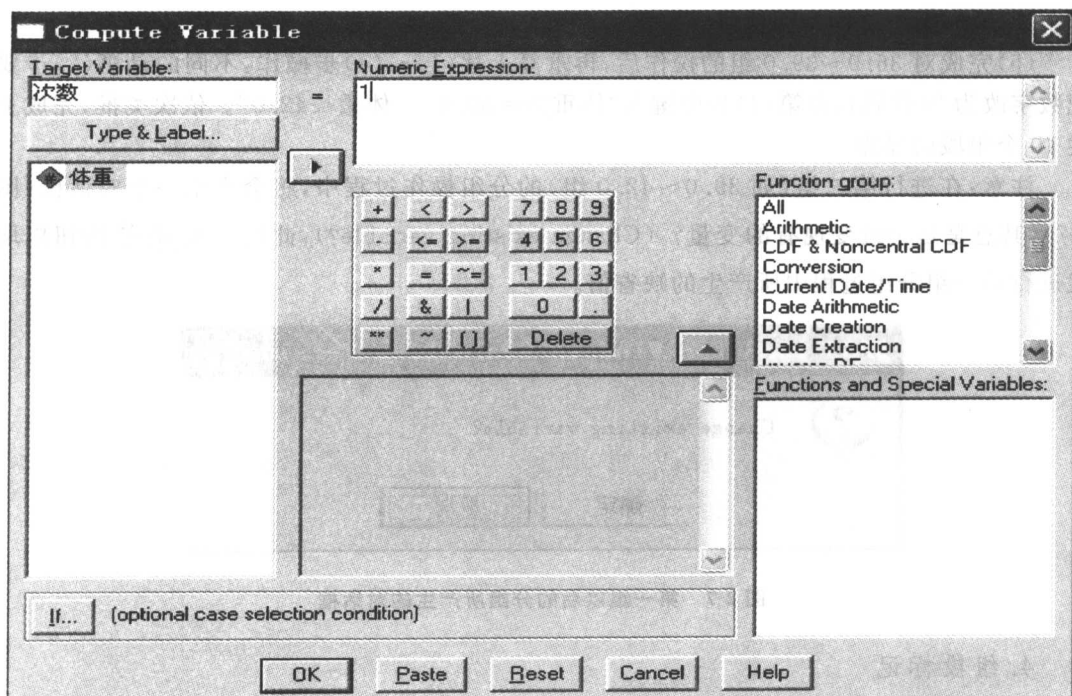


图 2-5 变量转换对话框

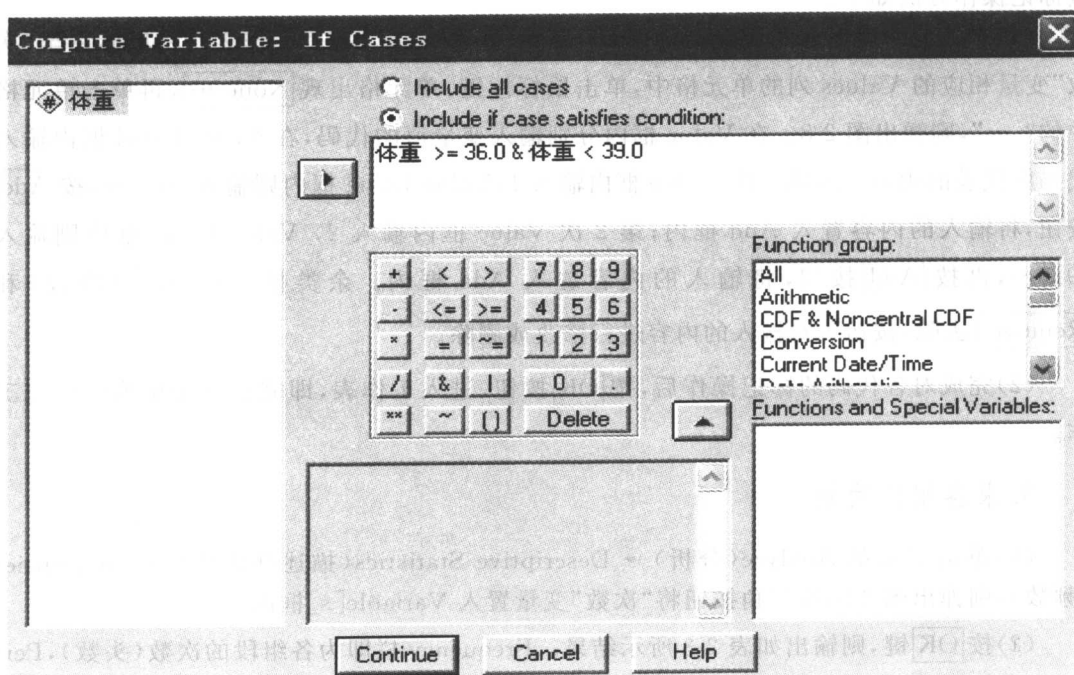


图 2-6 按图 2.5 中 If ... 按钮所弹出的条件语句对话框