

新手上路 即查即用

The SPSS logo is displayed in a stylized, bold, black font with a slight shadow effect, set against a dark, glowing rectangular background.

统计软件在体育中的应用

application of the SPSS statistical software in sports

郭秀文 编著

人民体育出版社

SPSS 统计软件在体育中的应用

郭秀文 编著

人民体育出版社

图书在版编目 (CIP) 数据

SPSS 统计软件在体育中的应用/郭秀文著.—北京:人民体育出版社,2007

ISBN 978-7-5009-3181-2

I. S… II. 郭… III. 体育统计—统计分析—软件包, SPSS IV. G80-32

中国版本图书馆 CIP 数据核字 (2007) 第 058958 号

人民体育出版社出版发行

北京嘉羽印务有限公司印刷厂印制

新华书店总销

*

787×960 16 开本 18.75 印张 385 千字

2007 年 5 月第 1 版 2007 年 5 月第 1 次印刷

印数: 1—3000 册

*

ISBN 978-7-5009-3181-2

定价: 28.00 元

社址: 北京市崇文区体育馆路 8 号 (天坛公园东门)

电话: 67151482 (发行部) 邮编: 100061

传真: 67151483 邮购: 67143708

(购买本社图书, 如遇有缺损页可与发行部联系)



作者简介

郭秀文，女，1968年生，祖籍山东，博士，副教授。1986年就读于西北师范大学数学系，1990年获得理学学士学位，同年留校从事体育教学工作。1993年至1996年，就读于成都体育学院，获硕士学位。2002年至2005年就读于北京体育大学，获体育教育与训练专业博士学位。艺术体操国家一级裁判员，健美操国家一级裁判员，国家级健美操指导员。主要从事体育统计学、体育测量与评价、运动训练学等理论课程及艺术体操、健美操等技术课程的教学工作。先后在《中国体育科技》《北京体育大学学报》等核心期刊上发表论文20余篇。承担过国家科技攻关计划项目《奥运优秀运动员科学选材的研究》的子课题《艺术体操优秀运动员选材的研究》的研究工作。

前 言

SPSS 是 Statistical Package for the Social Science 的缩写, 即社会科学统计软件包。SPSS 名为社会科学统计软件包, 是为了强调其社会科学应用的一面。其实, 它在自然科学、经济管理、商业金融、医疗卫生、体育等各个领域都能发挥巨大的作用。

统计学是一门应用科学, 正确地应用统计方法对所得到的结果进行合理的分析是使用者的愿望, 而统计学上那些晦涩难懂的数学公式及推导往往让使用者望而却步。本书就是了解到许多读者的这一想法, 因此, 在撰写过程中, 力求对必要的知识点进行深入浅出的讲解, 省去数学推导, 重在说明统计方法的适用条件或者说是使用范围。

本书以介绍 SPSS13.0 为主线, 对必要的知识点进行穿插讲解。为了使读者便于上机操作, 大部分内容以实例的形式出现, 读者只需按照书中的操作步骤进行学习, 就能很快地掌握 SPSS 统计软件。考虑到实际使用需要, 在撰写本书时, 尽量使每一部分内容具有相对的独立性, 这样更方便读者即查即用, 不必从头开始阅读。由于编写时间紧迫, 难免有错漏之处, 恳请读者指正。

编著者
2007 年春

目 录

第一章 数据文件的建立及整理 -----	1
第一节 数据文件的建立-----	1
第二节 数据文件的编辑与整理-----	12
第二章 描述统计过程 -----	35
第一节 基本概念介绍-----	35
第二节 频数分析过程-----	42
第三节 描述性统计分析过程-----	47
第四节 数据探索分析过程-----	49
第五节 列联表分析过程-----	57
第三章 均值比较和样本 T 检验 -----	65
第一节 均值比较-----	65
第二节 单个样本 T 检验-----	68
第三节 独立样本 T 检验-----	70
第四节 配对样本 T 检验-----	74
第四章 非参数检验 -----	79
第一节 卡方检验-----	79
第二节 二项检验-----	82
第三节 单个样本的 K-S 检验-----	85
第四节 两个独立样本的检验-----	87
第五节 多个独立样本的检验-----	91
第六节 两个相关样本的检验-----	95
第七节 多个相关样本的检验-----	99
第五章 方差分析 -----	103
第一节 单因素方差分析-----	103
第二节 单变量多因素方差分析-----	114

第六章 相关分析	125
第一节 二元变量的相关分析	125
第二节 偏相关分析	133
第七章 回归分析	137
第一节 一元线性回归分析	137
第二节 多元线性回归分析	144
第八章 聚类与判别分析	151
第一节 聚类分析	151
第二节 判别分析	169
第九章 因子分析	185
第十章 图形生成	197
第一节 条形图	197
第二节 线形图(曲线图)	216
第三节 饼形图	220
第四节 散点图	223
第五节 直方图	227
第六节 箱形图	229
第十一章 表格与图形的编辑及输出	235
附录: 统计术语中英文对照表	261

第一章 数据文件的建立及整理

统计工作中,对通过调查或试验得到的资料进行统计分析的前提条件是对得到的数据资料进行整理,如果要利用计算机对资料进行进一步的分析就需要将人工整理后的数据资料录入计算机,建立数据文件。我们知道数据资料质量的优劣将对统计分析结果直接影响产生,而按照实际调查或试验得来的数据资料往往是杂乱、不完整的,因此,为了提高统计分析的准确性,在对这些资料进行统计分析之前需要对它们进行归纳、整理。本章主要介绍如何建立数据文件,如何利用 SPSS 提供的功能对已建立的数据文件进行归纳和整理。

第一节 数据文件的建立

一、SPSS 主界面简介

在建立数据文件之前,让我们首先来了解 SPSS 主界面。在图 1-1 的最上方是标题栏,显示当前工作文件的名称。标题栏的下方是菜单栏,排列 SPSS 的所有菜单命令。菜单栏的下方是工具栏,排列各种图标按钮。工具栏的下方是数据编辑栏或称为数据输入栏,用户通过键盘录入的数据首先显示在此栏中。数据编辑栏的下方是数据显示区域,此区域是一个二维的表格,其中每一个矩形格称为一个单元格。当选定一个单元格时,该单元格的边框呈黑色。在图 1-1 的最下方是状态栏,它反映系统的工作状态,Date View 为数据窗口,Variable View 为变量定义窗口。

注意

在 SPSS 运行期间,不能同时打开两个或两个以上的数据窗口。即 SPSS 仅能进行一个数据文件的处理工作。

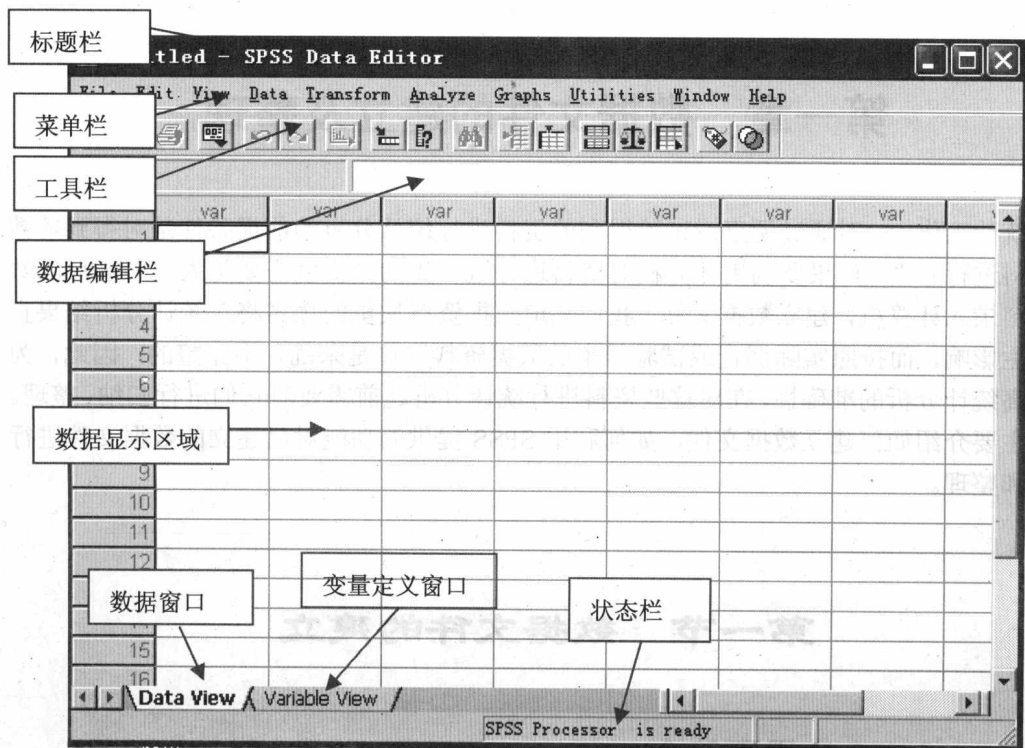


图 1-1 SPSS 主界面

二、变量、观测量概念介绍

(一) 变量

建立数据文件的第一步是定义变量，需要用户将数据窗口转换到变量定义窗口，鼠标左键单击变量定义窗口选项卡，数据文件将会由数据窗口转换为变量定义窗口（图 1-2）。

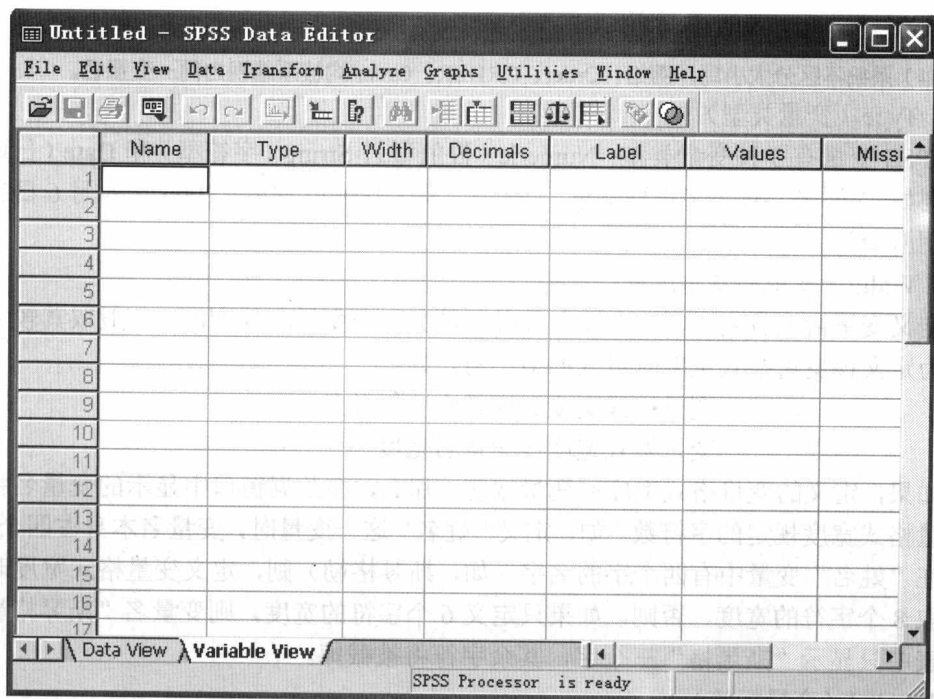


图 1-2 变量定义窗口

SPSS 变量主要由 Name(变量名)、Type (变量类型)、Width (变量格式宽度)、Decimal(小数点位数)、Label(变量标签)、Values(值标签)、Missing(缺失值)、Columns (单元格宽度)、Align (对齐格式) 和 Measure(测度水平)等属性构成。下面就变量的每一属性进行详细的介绍。

1. Name (变量名)

定义一个变量，首先应该为其命名。即给变量起个名字。

注意

- (1) 变量名的长度最多不超过 8 个字符。如果超出规定长度，则超出部分系统将不显示；
- (2) 第一个字符必须是字母或汉字，后面的可以是字母、符号、汉字或数字。例如，“order”、“序号”。但以下符号不能在变量名中使用：“!”、“*”、“?” 以及算术运算符，“.” 不能作为变量名的最后一个字符使用；
- (3) 变量名中不能使用 SPSS 的保留字，它们是：ALL、AND、BY、EQ、GE、GT、LE、LT、

NE、NOT、OR、TO、WITH;

(4) 系统不区分大小写, 例如: Order、ORDER、Order 均被系统视为同一变量名。

2. Type (变量类型)

SPSS 变量有 3 种变量类型: Numeric (数值型)、String (字符型) 和 Date (日期型)。例如, “姓名”属于字符型变量, “成绩”属于数值型变量, 2006 年 6 月 6 日属于日期型变量。

3. Width (变量格式宽度)

定义变量格式宽度时, 应当综合考虑变量宽度和变量名的宽度, 一般取其较大的一个作为定义该变量格式宽度时可取的最小值。即:

变量格式宽度 \geq 变量宽度, 并且

变量格式宽度 \geq 变量名宽度

如果, 定义的变量格式宽度 $<$ 变量宽度, 那么, 在数据窗口中显示的变量名将只显示变量格式宽度规定的字符数。如, 定义“姓名”这一变量时, 变量名本身占四个字符, 假如在“姓名”变量中有四个字的名字(如, 斯琴格勒)则, 定义变量格式宽度时就需至少取 8 个字符的宽度, 否则, 如果只定义 6 个字符的宽度, 则变量名“斯琴格勒”在数据表中只显示“斯琴格”三个字, 其余字符将被截掉。

4. Label (变量标签)

变量标签是对变量名的附加说明。由于 SPSS 中变量名的宽度不多于 8 个字符, 有时, 变量名不能充分表达变量的含义, 因此, 用户可以在变量标签中对变量进行标注。给变量加了标注之后, 当鼠标箭头指向一个变量的时候, 变量名下将立即显示出它的标签。SPSS 允许定义长达 255 个字节的变量标签。

5. Values (值标签)

变量的值标签是对变量的可能取值进行的附加说明。一般情况下, 仅对分类变量定义值标签。例如, “队别”我们可以把它定义为数值型变量, 也可以把它定义为字符型变量。如果将它定义为数值型变量, 则用户必须将各个队的代码记住, 因为, SPSS 数据表上仅显示各个队的代码, 时间长了可能用户自己都会忘记某一代码究竟代表哪个队, 从而造成统计结果的混乱; 如果将它定义为字符型变量, 用户的键盘输入工作量又较大, 造成时间的浪费。解决的最好办法就是将它定义为数值型变量, 并对它的可能取值情况进行附加说明, 即对该变量定义值标签, 这样做不仅节约用户的时间, 而且, 就算时间过去很久, 用户打开该数据文件的时候, 一样清楚各代码分别代表的是哪个队。

6. Missing (缺失值)

在数据的搜集、整理过程中, 由于笔误、口误、漏登等等原因难免会造成统计资料



的“缺”、“漏”、“误”。统计中将那些没有观测到、记录到，或者记录结果有明显错误的数值称为缺失值。

如果统计分析中使用了缺失值，就会影响分析结果的准确性。SPSS 提供了处理缺失值的功能，以便在统计分析中将缺失值排除。

7. Columns (单元格宽度)

Columns (单元格宽度) 是指数据窗口中变量所占的列数。Columns (单元格宽度) 与 Width (变量格式宽度) 初学者往往容易混淆，但实际上，它们是两个不同的概念，在定义变量中所表达的含义也不相同。在后面的实例中我们再来区分两者的不同。

8. Align (对齐格式)

在数据窗口中，变量值在单元格可以左、中、右的方式显示。一般情况下，系统默认的对齐方式如下：数值型为右对齐、字符型为左对齐。用户如对系统默认的对齐方式不是很满意，还可以根据自己的喜好选择变量值的对齐方式，如，所有数据都采用居中的格式。

9. Measure (测度水平)

测度水平有以下三种：

- **Scale** (定比测度，或称比率测度)：取值为连续型数值，可进行任何形式的数学运算。如：身高、体重、成绩等。

- **Ordinal** (定序测度，或称顺序测度)：无论是数值型变量还是字符型变量，只要资料具有某种内在的顺序关系，均可设为 **Ordinal** 测度。

- **Nominal** (定类测度，或称名义测度)：资料不具备某种内在顺序的字符型变量，可设置为 **Nominal** 测度。

(二) 观测量

SPSS 中用 Case(s) 表示观测量、样品或事件。在 SPSS 数据窗口的二维表格中，每一行存放的就是一个观测量，每一列存放的是观测量的某一变量值。例如，在图 1-3 所示的数据窗口中，第 m 行的运动员通过序号、年龄、专项、体重、身高、坐高等变量反映运动员的各项特征。

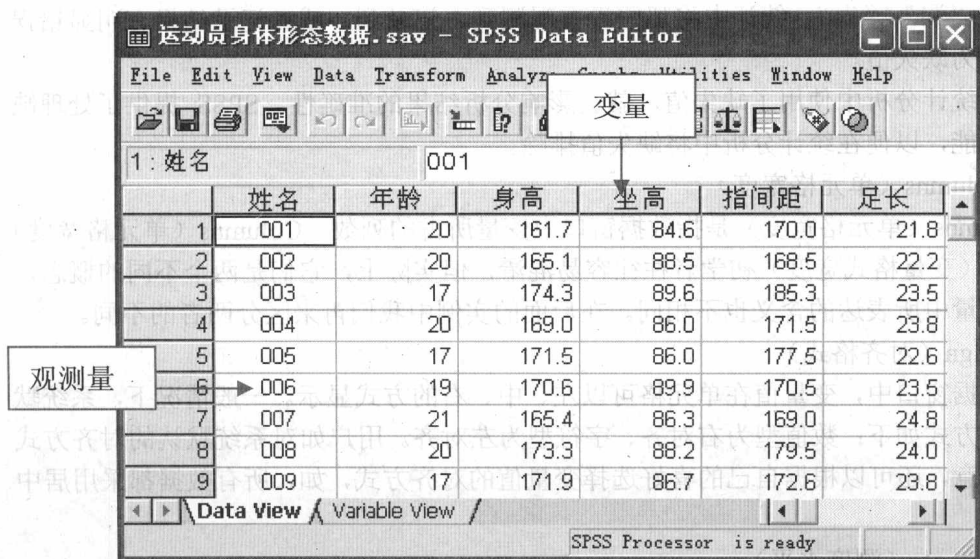


图 1-3 观测量、变量示意图

三、数据文件的建立


下面以具体实例来展示数据文件的建立步骤，将表 1-1 建立一个数据文件。需要说明的是读者不必非得按部就班的进行操作，此处只是为了讲解方便，故采用步骤的形式进行。当您已熟练掌握数据文件的建立方法后，您也可以根据自己的需要进行操作步骤先后顺序的安排。

通过对表中资料的分析，可以建立一个包含 5 个变量的数据文件，其中，“序号”“年龄”“跳高成绩”均是数值型变量，测度水平应定义为 Scale（比率测度），“序号”的测度水平也可定义为 Ordinal（顺序测度），“姓名”是字符型变量，测度水平定义为 Nominal（名义测度），“性别”属于类型变量，可以将它定义为数值型变量并附加值标签，测度水平定义为 Ordinal（顺序测度）。

表 1-1 某校学生跳高成绩

序号	姓名	性别	年龄	跳高成绩 (米)
1	刘 明	男	20	1.80
2	朱国良	男	18	1.79
3	林军亮	男	20	1.89
4	古志强	男	19	1.98
5	李 敏	女	21	1.69
6	刘晓枫	女	19	1.78
7	孙沁玲	女	18	1.83
8	黄静一	女	21	1.58
9	何永亮	男	20	1.72
10	赵志勇	男	19	1.77
11	张 勇	男	18	1.85
12	郭明民	男	20	1.75
13	胡志元	男	18	1.80
14	张倩倩	女	19	1.48

(一) 定义变量

下面进行定义变量的操作：首先，在桌面上找到 SPSS 的图标 ，并用鼠标左键双击该图标，进入 SPSS 的主界面（图 1-1）。然后将鼠标移至主界面最下方的 Variable View（变量定义窗口）选项卡，单击鼠标左键，进入变量定义窗口，开始定义变量。

第一步 定义变量名 (Name)

在 Name 下的单元格中输入变量名，如图 1-4 中的“序号”。

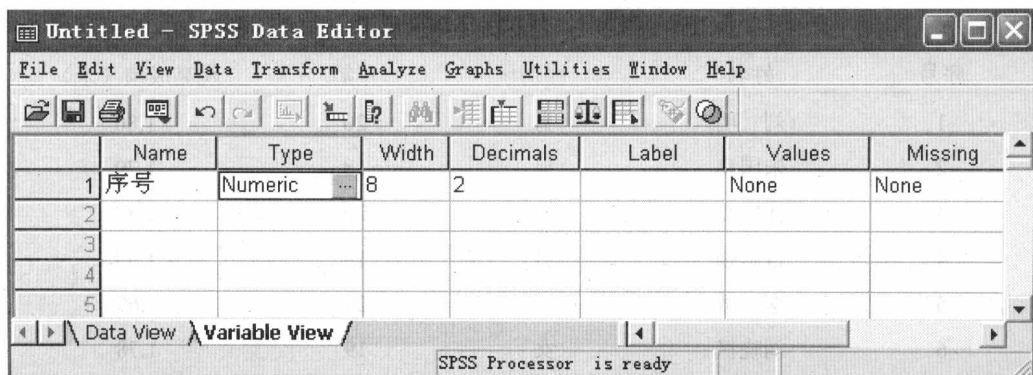
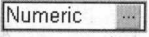


图 1-4 定义变量窗口

第二步 定义变量类型 (Type)

在 Type 下选择变量类型，单击 Type 下的单元格，出现图标 ，单击此图标右边的灰色按钮，打开如图 1-5 所示的变量类型对话框，从中选择变量类型。

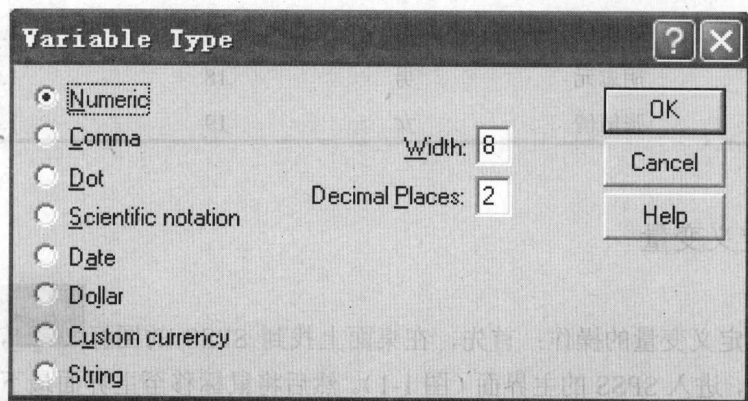
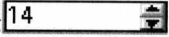


图 1-5 变量类型对话框

图 1-5 的左半部分为变量类型，其中，Numeric、Comma、Dot、Scientific notation、Dollar 为 5 种不同形式的数值型变量，Date 为日期型变量，String 为字符型变量。下面就 5 种数值型变量进行简要说明，见表 1-2。

表 1-2 5 种数值型变量的输入与显示格式

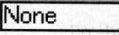
数值型变量	数据输入格式	数据显示格式示例
Numeric (标准型)	标准数值或科学记数法	标准数值格式: 1234.56
Comma (带逗号型)	带逗号型或科学记数法	带逗号格式: 1,234.56
Dot (带圆点型)	带圆点型或科学记数法	带圆点格式: 1.234,56
Scientific notation (科学记数法)	标准数值或科学记数法	科学记数法: 1.2346E+03
Dollar (带美元符号)	标准数值或科学记数法	带美元符号: \$ 1,234.56

第三步 定义变量格式宽度(Width)、小数点位数(Decimals)、单元格宽度(Columns)
单击 Width (或 Decimals、Columns) 下的单元格, 出现图标 , 用鼠标左键点击上下箭头进行变量格式宽度大小 (或小数点位数、单元格宽度) 的选择。

第四步 定义变量标签 (Label)

在 Label 下的单元格中, 输入变量标签, 如, “order”。

第五步 定义值标签

“序号”不需要定义值标签, 按系统默认取值, 单元格显示 None (无)。但是, 变量“性别”就需要定义值标签。将“性别”的变量类型定义为数值型, 定义性别值标签: 1=“男”, 2=“女”。单击 Values 下的单元格, 出现图标 , 单击此图标右边的灰色按钮, 打开如图 1-6 所示的值标签对话框。在 Value 栏中输入 1, 在 Value Label 栏中输入男, 此时, 在对话框最下方矩形框 (值标签清单) 左边的 Add (添加) 按钮被激活, 点击 Add 按钮, 值标签即被添加到值标签清单中。如上述操作定义 2=“女”的值标签。

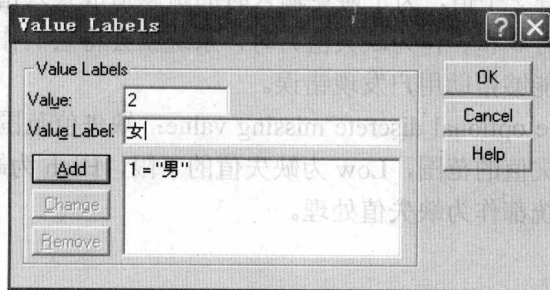
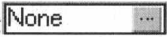


图 1-6 值标签对话框

如若对已经列入清单的值标签不满意或想进行修改, 可以将清单中的值标签选中, 然后单击清单框左边的第三个按钮 Remove 将该值标签移出。或在清单中选中需要

修改的值标签, 在 Value 和 Value Label 栏中将新的值标签录入, 这时, Change 按钮被激活, 单击此按钮, 旧的值标签就会被新值所取代。

第六步 定义缺失值

单击 Missing 下的单元格, 出现图标 , 单击此图标右边的灰色按钮, 打开如图 1-7 所示的缺失值对话框。

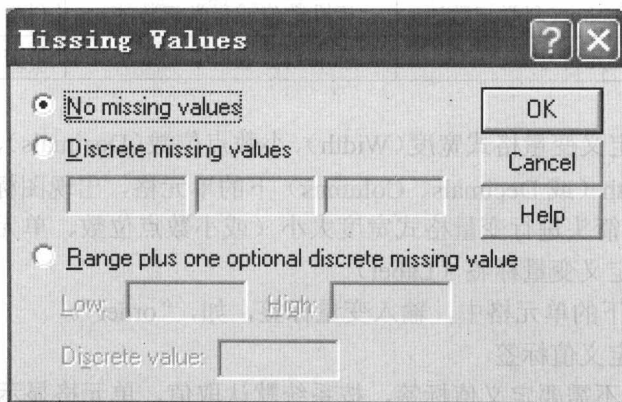


图 1-7 缺失值对话框

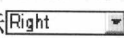
对话框中有三个选项:

- **No missing values:** 无缺失值。此选项为系统默认项。
- **Discrete missing values:** 离散缺失值。下面有 3 个矩形框, 当选中了此选项后, 这 3 个矩形框被激活, 可以将定义的缺失值输入矩形框中, 不必 3 个矩形框都输入数据。如, 变量“性别”共有 2 个值, 为了避免输入时的疏忽造成录入错误, 当输入的结果为定义的缺失值 (假如将 0 和 3 作为缺失值) 时, 系统就会将它们作为缺失值对待。定义缺失值的好处是系统能够帮助用户发现错误。
- **Range plus one optional discrete missing value:** 缺失值范围, 外加一个离散缺失值。此选项可定义缺失值的范围, Low 为缺失值的下限, High 为缺失值的上限。在上、下限之间的数据, 系统都作为缺失值处理。

注意

定义缺失值时一定不要把合法值定义进去, 尤其在定义缺失值范围时, 务必考虑周全, 否则会影响分析结果的准确性。

第七步 定义对齐格式

单击 Align 下的单元格, 出现图标 , 单击此图标右边的箭头按钮展开下拉