



# 数据统计与分析

## —SPSS应用教程

Data Statistics and Analysis  
A Course in Applied SPSS



阮桂海 ◎主编  
蔡建平 刘爱玉 等 ◎编著



北京大学出版社  
PEKING UNIVERSITY PRESS



# 数据统计与分析

## —SPSS应用教程

Data Statistics and Analysis  
A Course in Applied SPSS



阮桂海 ⊙主编

蔡建平 刘爱玉 等 ⊙编著



北京大学出版社  
PEKING UNIVERSITY PRESS

## 图书在版编目(CIP)数据

数据统计与分析: SPSS 应用教程/阮桂海主编; 蔡建平, 刘爱玉等编著. —北京: 北京大学出版社, 2005. 7

(21世纪社会学系列教材)

ISBN 7 - 301 - 09268 - 7

I . 数… II . ①阮… ②蔡… ③刘… III . 统计分析 - 软件包, SPSS - 教材

IV . C819

中国版本图书馆 CIP 数据核字(2005)第 065975 号

书 名: 数据统计与分析——SPSS 应用教程

著作责任者: 阮桂海 主编 蔡建平 刘爱玉 等编著

责任编辑: 周 婧 耿协峰

标准书号: ISBN 7 - 301 - 09268 - 7/C · 0350

出版发行: 北京大学出版社

地 址: 北京市海淀区成府路 205 号 100871

网 址: <http://cbs.pku.edu.cn>

电 话: 邮购部 62752015 发行部 62750672 编辑部 62753121

电子信箱: ss@pup.pku.edu.cn

排 版 者: 北京高新特打字服务社 82350640

印 刷 者: 世界知识印刷厂

经 销 者: 新华书店

730 毫米 × 980 毫米 16 开本 17.5 印张 290 千字

2005 年 7 月第 1 版 2005 年 7 月第 1 次印刷

定 价: 29.00 元(含光盘)

---

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



# 21世纪社会学系列教材

Textbooks of Sociology in 21st Century



# 前 言

自 1987 年起,北京大学社会学系在全国高校率先进行 SPSS 和 SAS 的教学及科研应用,并率先出版了《SPSS/PC<sup>†</sup> 4.0 高级教程》,如今 18 年过去了,随着 SPSS 功能的不断创新以及应用领域的不断拓宽,有关 SPSS 的书也多了起来。作者认为,书固然要写,对话框也应该教,但必要的命令语句绝对不能丢,同时还应该在 SPSS 统计方法和分析方法的准确性方面把关,以免误导读者。

作者写 SPSS 和 SAS 教程喜欢“竹筒倒豆子”似地敬奉给读者,这就难免内容全而价位高,为此,“而今迈步从头越”,特改编一些普及型的 SPSS 及 SAS 教材。本书就是这样对已版《SPSS 实用教程》、《SPSS for Windows 高级应用教程》、《统计分析应用教程》等书进行锤炼而成的,以关照 SPSS 初学者,并尽量为 SPSS 朦胧者把关定向。因此本书内容简要,不受版本限制;若读者用惯了某个版本的 SPSS,则可以在本书的指导下大胆地继续使用该版本的 SPSS。

本书在操作方面以教对话框为主,同时尽可能地附上一条相应的命令进行对照,便于读者由表及里地了解 SPSS 的真谛与内涵。

按常用的 SPSS 模块,本书分为 17 章,有“\*”号的章节是研究生的教学内容,本科生也可选读。这些选读内容与附录 1 ~ 附录 3 的习题答案以及程序数据等是作为教学课件单独放入一张光盘中随书发行的。程序数据有版权,勿扩散他用。

第 1 章及附录 1 由刘爱玉编写,详细介绍了目前流行的 5 种数据输入法。

第 2 章 ~ 第 4 章包括从建立 SPSS 文件(即简易的程序)到数据的抽样及变换。由蔡建平、阮开林、刘云、刘萍执笔。

第 5 章 ~ 第 7 章是描述统计,包括对 Custom Table 过程、Frequencies 过程、Descriptives 过程、Crosstabs 过程、Explore 过程等的应用。由阮潮海、蔡楠、曲庆云、谢力丹、孟峥、严康敏执笔。

第 8 章是 Compare Mean(总体均值差异性的比较),包括 Means、T-test 以及 ANOVA 等过程的应用。由阮潮霞、郭崇德、张才彬、陈耀辉执笔。

第 9 章是 Correlation (相关分析) 模块, 包括双变量相关(零阶相关)和一阶以上的偏相关。由朱志海、阮开顺执笔。

第 10 章是 Multiple Response(多项选择, 即多选项) 的统计分析篇, 人人必学。由蔡建瓴、陈丹执笔。

第 11 章是 Nonparametric Test(非参数检验), 用于数据非正态分布时的检验。由阮桂海、钟小芳执笔。

第 12 章是 Linear Regression(一元和多元线性回归)。由蔡文英、秦小峰执笔。

第 13 章是非线性回归, 包括 Curve Estimation(曲线估计)、Binary Logistic Regression(二分逻辑斯蒂克回归)等回归分析。由蔡建琼、罗安娜执笔。

第 14 章是聚类分析过程。由朱志洪执笔。

第 15 章是因素分析与品牌对应分析。由綦淑娟、郝正元执笔。

第 16 章运用结合分析研究品牌战略。由杨武栓、顾林枝执笔。

第 17 章是复方差分析(GLM: Multivariate) 的应用。由阮桂海、李欣执笔。

特别值得一提的是: 本书第 1 章 ~ 第 3 章都涉及数据的输入法, 目的是让读者朋友将“面上了解”与“重点掌握”相结合, 进而重视学习第 3 章所介绍的 DATA LIST 语法知识和快速输入数据的技巧, 这是 SPSS 的精髓。只有通过 DATA LIST 命令输入原始数据和运行数据, 才是多快好省的基本功。特别是 SPSS 的新读者, 一开始就要防止浮躁, 养成勤奋的习惯, 除了一定要实践 DATA LIST 命令的强大功能外, 还要用好 Compute 命令、Recode 命令、IF 命令。这些命令学会了, 数据就能正确快捷地进入 SPSS 系统, 并方便地进行各种数据变换。

本书每章后面都有习题压阵, 便于读者加深理解与应用, 因此本书是初学者最好的“敲门砖”。本书随盘赠送程序数据 20 例, 但更多的程序或数据还可以在参加新知识培训期间再次获赠(详见附录 3)。

在编著过程中, 教育部教学指导委员会与北京大学出版社给予了宝贵的指导和有力的支持, 北京大学社会学系杨善华教授提供了全套正版 SPSS 原著作为参考书, 本书的责任编辑周婧、耿协峰老师始终一丝不苟、严把质量关, 在此一并表示由衷的感谢!

阮桂海

于北京大学社会学系

2005 年 1 月

rgih114716@sina.com 或 rghai@sohu.com

# 目 录

<b>第 1 章 SPSS 数据的输入法 .....</b>	(1)
1.1 问卷设计及数据编码 .....	(1)
1.1.1 问卷及编码样例 .....	(1)
1.1.2 编码应注意事项 .....	(2)
1.2 在 SPSS 数据窗口直接输入数据 .....	(3)
1.2.1 首先在 Data View(数据浏览)窗口输入数据 .....	(3)
1.2.2 然后定义变量及其数据的类型属性 .....	(4)
1.3 在中文 Word 中输入数据 .....	(8)
1.4 在 Excel 中输入数据 .....	(10)
◆ 习题 1 .....	(11)
<b>第 2 章 两种常用的数据输入法及数据的抽样加权 .....</b>	(12)
2.1 SPSS 的 Data Editor 窗口 .....	(12)
2.2 几十个 Cases 的数据输入法 .....	(13)
2.3 成百上千个 Cases 的数据输入法 .....	(14)
2.4 个案排序 .....	(15)
* 2.5 数据的行列互换 .....	(16)
2.6 对限选个案 Cases 进行局部统计 .....	(18)
2.6.1 限选某个子总体进行统计 .....	(18)
2.6.2 提取随机样本进行统计 .....	(20)
2.6.3 限选前 n 个 Cases .....	(22)
2.7 个案加权 .....	(23)
2.8 Data 中的其他菜单命令 .....	(25)
◆ 习题 2 .....	(26)

<b>第3章 SPSS最正规最快捷的数据输入法</b>	.....	(28)
3.1 SPSS的简单命令	.....	(28)
3.2 再谈对问卷中的数据编码	.....	(29)
3.3 定义SPSS变量的属性	.....	(31)
3.4 SPSS数据的固定格式	.....	(32)
3.5 SPSS数据的自由格式	.....	(33)
3.6 建立SPSS的命令文件	.....	(33)
3.7 SPSS命令文件的4种格式	.....	(34)
3.8 编程后立即进行统计分析	.....	(35)
3.9 调用老命令文件(程序)进行统计分析	.....	(36)
3.10 正确编辑SPSS原始数据小结	.....	(36)
3.11 统计分析的最佳方案	.....	(36)
◆习题3	.....	(38)
<b>第4章 数据变换</b>	.....	(40)
4.1 创建新变量	.....	(40)
4.1.1 点击Compute对话框创建新变量	.....	(41)
4.1.2 采用Compute命令创建新变量	.....	(42)
4.1.3 数学函数	.....	(42)
4.1.4 缺失值函数	.....	(43)
4.2 采用IF命令进行条件变换及逻辑校验	.....	(44)
4.2.1 IF命令格式	.....	(44)
4.2.2 关系表达式中的关系符	.....	(44)
4.2.3 逻辑表达式中的运算符	.....	(45)
4.2.4 两种缺失值的逻辑函数	.....	(45)
4.2.5 缺失值的处理	.....	(45)
4.2.6 IF命令的对话框	.....	(46)
4.3 数据重新编码	.....	(47)
4.3.1 点击对话框重新编码数据的实例	.....	(47)
4.3.2 Recode命令的格式	.....	(50)
4.3.3 Recode在命令文件中的位置	.....	(51)
4.4 计算有相同情况的次数	.....	(52)
4.5 SPSS的运算次序	.....	(52)
◆习题4	.....	(53)

<b>第 5 章 频次分布与描述统计</b>	.....	(55)
5.1 选用 Frequencies 过程进行频次统计	.....	(55)
5.1.1 Frequencies 过程的预备知识	.....	(55)
5.1.2 每日温差的频次统计	.....	(60)
5.1.3 另加一个完整的频率表	.....	(62)
5.1.4 Frequencies 过程的命令表	.....	(63)
*5.2 采用 Custom Table 过程制表	.....	(64)
5.2.1 Table 过程的对话框	.....	(64)
5.2.2 另一个双变量的表格	.....	(65)
5.3 采用 Descriptives 过程描述数据	.....	(67)
5.3.1 Descriptives 过程的最佳功能	.....	(67)
5.3.2 存储标准分(Z 值)	.....	(67)
5.3.3 在对话框中进行 Descriptives 的描述统计	.....	(68)
5.3.4 Descriptives 过程的子命令及其关键词	.....	(70)
◆ 习题 5	.....	(70)
<b>第 6 章 应用 Explore 过程检测数据</b>	.....	(75)
6.1 Explore 过程可检测数据的哪些错误	.....	(75)
6.1.1 数据错误的原因	.....	(75)
6.1.2 数据检测的重要性	.....	(76)
6.2 用于 Explore 过程的检测数据	.....	(77)
6.3 图形检测之一:直方图	.....	(77)
6.4 图形检测之二:茎-叶图	.....	(78)
6.5 图形检测之三:框图	.....	(80)
6.6 统计分析前要预先检验假设	.....	(83)
6.6.1 方差相等的检验法	.....	(83)
6.6.2 正态性检验	.....	(86)
*6.7 集中趋势分布的三种较佳平稳测度	.....	(87)
6.8 在对话框中进行 Explore 检测	.....	(90)
6.8.1 采用 Explore 对话框进行统计	.....	(90)
6.8.2 Plots 选项用于画图探测数据	.....	(91)
6.9 在 Syntax 窗口编程	.....	(92)
◆ 习题 6	.....	(93)

<b>第 7 章 用 Crosstabs 统计交叉汇总表及结合测量</b>	.....	(95)
7.1 妇女的年龄与文化水平的交叉汇总	.....	(95)
7.2 由变量“是否在业”控制的女性年龄与文化的关系	.....	(97)
7.3 定比一定比型变量用皮尔逊 Corr 系数	.....	(99)
7.4 定类一定距型变量用 Eta 系数	.....	(101)
7.5 定类一定序型以下变量用 Lambda 系数	.....	(103)
7.6 相对风险率估计	.....	(105)
7.6.1 研究范例	.....	(105)
7.6.2 相对风险率估计法	.....	(106)
7.7 在 Crosstabs 的对话框中进行交叉汇总和结合测量	.....	(107)
7.8 Crosstabs 过程的命令总表	.....	(109)
◆ 习题 7	.....	(109)
<b>第 8 章 两个总体均值的比较</b>	.....	(114)
8.1 Means 过程的应用	.....	(115)
8.1.1 寻找相关方面	.....	(115)
8.1.2 论证所用的数据及程序	.....	(115)
8.1.3 高考成绩与父亲职业的关系	.....	(116)
8.1.4 分析不同地区的考试成绩	.....	(116)
8.1.5 采用 Means 对话框进行统计	.....	(119)
8.1.6 过程 Means 的命令、子命令及其用法	.....	(120)
8.2 独立样本 T-test 的应用	.....	(122)
8.3 配对样本 T-test 的应用	.....	(124)
8.3.1 采用对话框进行配对样本 T-test	.....	(124)
8.3.2 分析配对样本 T-test 的结果	.....	(126)
8.3.3 采用命令法进行配对样本 T-test	.....	(127)
8.4 单因素方差分析	.....	(128)
8.5 多因素方差分析	.....	(131)
8.5.1 ANOVA 过程的应用之一:描述统计	.....	(131)
8.5.2 ANOVA 过程的应用之二:方差分析	.....	(133)
8.5.3 对主效应的检验	.....	(135)
8.5.4 ANOVA 过程的命令一览表	.....	(135)
8.5.5 采用对话框进行 ANOVA 分析	.....	(137)
◆ 习题 8	.....	(137)

<b>第 9 章 相关分析.....</b>	(141)
9.1 双变量相关分析 .....	(141)
9.1.1 数据 .....	(142)
9.1.2 从相关矩阵看“SARS 疫情”与气温的关系 .....	(143)
9.1.3 缺失值的剔除法对相关矩阵的影响 .....	(144)
9.1.4 Correlations 的对话框及其解释 .....	(146)
9.1.5 运行 Correlations 过程命令 .....	(148)
9.2 偏相关分析 .....	(149)
9.2.1 计算偏相关系数的原理与偏相关的“阶” .....	(150)
9.2.2 如何辨别变量之间的虚假相关 .....	(150)
9.2.3 找回隐含相关 .....	(152)
9.2.4 通过对话框进行偏相关分析 .....	(152)
9.2.5 偏相关主对话框的应用说明 .....	(153)
9.2.6 将对话框中已选择的命令粘贴到 Syntax 窗口 .....	(154)
9.2.7 Partial Correlations 命令表 .....	(155)
◆ 习题 9 .....	(156)
<b>第 10 章 多选项的统计技术 .....</b>	(158)
10.1 多选项的计算机编码 .....	(158)
10.2 多选项中二分法与分类法的区别 .....	(159)
10.2.1 多选项分类法 .....	(159)
10.2.2 多选项二分法 .....	(160)
10.3 多选项的数据及其程序 .....	(161)
10.4 多选项频次分布表及其分析 .....	(162)
10.5 多选项的交叉汇总表 .....	(164)
10.5.1 交叉汇总表的统计法 .....	(164)
10.5.2 输出的交叉汇总表 .....	(165)
10.5.3 交叉汇总表分析 .....	(165)
◆ 习题 10 .....	(166)
<b>第 11 章 非参数检验 .....</b>	(168)
11.1 非参数检验的命令菜单和数据要求 .....	(168)
11.2 单样本卡方检验 .....	(170)
11.3 二项分布检验 .....	(171)
11.4 Runs 检验 .....	(172)
11.5 单样本 K-S 检验 .....	(173)

11.6 双样本 K-S 检验.....	(175)
11.7 K 个独立样本的 Kruskal-Wallis 检验 .....	(176)
11.8 两个相关样本的检验.....	(177)
11.8.1 两个相关样本的 Wilcoxon Signed Ranks 检验 .....	(178)
11.8.2 两个相关样本的 Sign 检验 .....	(179)
11.8.3 两个相关样本的 McNemar 检验 .....	(180)
11.9 K 个相关样本没有差异的检验 .....	(181)
11.9.1 K 个相关样本没有差异的 Friedman 检验 .....	(181)
11.9.2 K 个相关样本没有差异的 Kendall 检验 .....	(182)
11.9.3 K 个相关样本没有差异的 Cochran 的 Q 检验 .....	(183)
❖ 习题 11 .....	(184)
<b>第 12 章 多元线性回归 .....</b>	<b>(187)</b>
12.1 多元线性回归模型的拟合优度.....	(188)
12.2 多元回归的基本条件.....	(190)
12.3 设定异常值和影响点.....	(194)
12.3.1 只有 1 个预测变量时的异常值测量法 .....	(194)
12.3.2 有多个预测变量时的异常值测量法 .....	(195)
12.4 数据不符合回归假设时的处理法 .....	(196)
12.5 向前选择变量法.....	(196)
12.6 自后淘汰变量法 .....	(198)
12.7 逐步回归法.....	(199)
12.8 多元共线性诊断.....	(201)
12.9 解释回归模型.....	(202)
❖ 习题 12 .....	(203)
<b>第 13 章 曲线估计及逻辑斯蒂克回归 .....</b>	<b>(206)</b>
13.1 曲线估计.....	(206)
13.1.1 曲线估计的数据及其统计法 .....	(207)
13.1.2 曲线估计的分析 .....	(208)
13.1.3 存储临时变量 .....	(209)
13.2 二分的逻辑斯蒂克回归.....	(210)
13.2.1 二分逻辑斯蒂克回归的对话框及数据例子 .....	(210)
13.2.2 设置二分逻辑斯蒂克回归模型 .....	(211)
13.2.3 二分逻辑斯蒂克回归的输出与分析 .....	(211)
13.2.4 二分逻辑斯蒂克回归系数的假设检验 .....	(213)

◆ 习题 13 .....	(213)
* 第 14 章 聚类分析 .....	(216)
14.1 二阶聚类 .....	(216)
14.1.1 二阶聚类的特点 .....	(216)
14.1.2 二阶聚类的数据 .....	(217)
14.1.3 二阶聚类的对话框例子 .....	(219)
14.1.4 二阶聚类的结果分析 .....	(220)
14.1.5 二阶聚类小结 .....	(224)
14.2 谱系聚类 .....	(225)
14.3 个案 Q 聚类 .....	(225)
14.3.1 从冰柱图看聚类结果 .....	(226)
14.3.2 平均连接法的图表 .....	(228)
14.3.3 从树形图分析聚类成员 .....	(228)
14.4 变量 R 聚类 .....	(229)
◆ 习题 14 .....	(230)
* 第 15 章 因素分析与品牌对应分析 .....	(232)
15.1 因素分析所用的数据 .....	(232)
15.2 因素分析的对话框设置法 .....	(233)
15.3 输出结果及其分析 .....	(234)
15.3.1 因素的初始统计量(共通性) .....	(234)
15.3.2 因素抽取 (Factor Extraction) .....	(235)
15.4 用主成分法抽取前 5 个因素 .....	(238)
15.5 通过最大似然度抽取因素 .....	(239)
15.6 删 除 Loading $\leqslant 0.5$ 的小载荷量 .....	(240)
15.7 转轴前后的对比 .....	(241)
15.8 对转轴效果的验证 .....	(243)
15.9 因素分析 .....	(243)
15.10 市场分析与社会调研中的对应分析 .....	(244)
◆ 习题 15 .....	(247)
* 第 16 章 品牌研究中的结合分析 .....	(249)
16.1 结合分析的方法 .....	(249)
16.2 通过正交设计建立卡片 .....	(249)
16.3 精品购物的问卷调查例子 .....	(251)
16.4 如何做好结合分析 .....	(251)

16.5 市场预测与决策.....	(253)
◆ 习题 16 .....	(254)
<b>第 17 章 复方差分析的应用 .....</b>	<b>(255)</b>
17.1 什么是复方差分析.....	(255)
17.2 假设与检验.....	(256)
17.2.1 假设的必要条件 .....	(256)
17.2.2 假设与检验 .....	(256)
17.3 因变量是二水平(双样本)模型的假设检验 .....	(259)
17.3.1 因变量之间相关性检验.....	(259)
17.3.2 因变量联合分布为正态性的假设与检验 .....	(260)
17.3.3 方差齐性检验 .....	(261)
17.4 双因变量单因素的 Hotelling $T^2$ 检验 .....	(262)
17.4.1 单因变量的 Hotelling $T^2$ 检验 .....	(263)
17.4.2 参数估计 .....	(264)
17.5 因变量是多水平(多样本)的复方差线性模型 .....	(265)
◆ 习题 17 .....	(266)

# 第 1 章

## SPSS 数据的输入法

数据输入是研究者利用 SPSS 进行统计分析的必要前提,这一工作实际上是把通过问卷调查获得的资料转变为 SPSS 能够识别的数据文件,为各种统计分析作准备。本章主要介绍数据输入的各种方法。第 2 章还要进一步概括性地点评数据输入法,以便读者各取所需。

### 1.1 问卷设计及数据编码

#### 1.1.1 问卷及编码样例

下面是一研究者为了解礼物赠送情况进行调查而获得的一份问卷,我们对该问卷进行了编码。

表 1-1 问卷及编码样例

问卷号:0001(栏位 1—4 列)

注意:数字填在横线上,选答题请画圈“○”		[ 编码 ]	[ 栏位 ]
1. 您的性别:	(1) 男 (2) 女	<input type="text" value="2"/>	(5)
2. 您的年龄:	35 岁	<input type="text" value="35"/>	(6—7)
3. 您的文化程度:	(1) 没上过学 (2) 小学 (3) 初中 (4) 普通高中 (5) 中专、技校或职高 (6) 大专及以上	<input type="text" value="4"/>	(8)
4. 您的婚姻状况:	(1) 未婚 (2) 在婚 (3) 丧偶 (4) 离异 (5) 分居	<input type="text" value="2"/>	(9)
5. 您目前的职业是:	(1) 高级专业技术人员 (2) 中低级专业技术人员 (3) 国家机关、群众团体干部或文职人员 (4) 企业高层经营管理人员	<input type="text" value="05"/>	(10—11)

## 数据统计与分析

(续表)

注意:数字填在横线上,选答题请画圈“○”		[编码]	[栏位]
(5) 企业中低层经营管理人员      (6) 生产工人、运输工人及相关人员			
(7) 军人      (8) 农、林、牧、渔劳动者      (9) 商业、服务业人员			
(10) 内退或下岗      (11) 离、退休      (12) 失业      (13) 其他			
6. 从 2004 年 7 月到 2005 年 7 月这一年,您的收入大约是(包括 12 个月的工资、奖金、津贴、补助和各种实物折款) <u>28000</u> 元		028000	(12—17)
7. 您一般是在什么情况下送礼(选择最经常送礼的两个场合): (编码时拆分为 2 个变量→)		01	(18—19)
		06	(20—21)
(1) 婚嫁      (2) 生日      (3) 搬迁新居      (4) 中国传统节日			
(4) 探访亲友      (5) 回乡探亲      (6) 探病      (7) 西方节日			
(8) 生子      (9) 毕业      (10) 其他(旅行、出国等)			
第一经常场合 <u>(1)</u>		第二经常场合 <u>(6)</u>	
8. 在最经常送礼的两个场合,您最经常选择的礼品是(选择两项): (编码时拆分为 2 个变量→)		07	(22—23)
		10	(24—25)
(1) 礼券      (2) 水果、糖果      (3) 贺卡      (4) 装饰品、工艺品			
(5) 日用品      (6) 烟酒      (7) 钱      (8) 家用电器			
(9) 衣物      (10) 书籍      (11) 补品      (12) 其他			
9. 您一般送礼的对象是(可多选):      (编码时拆分为 7 个变量→) <u>2 2 1 1 2 2 2</u> (26—32)			
(1) 家人      (2) 亲戚      (3) 朋友      (4) 同学			
(5) 同事      (6) 领导      (7) 其他 <u>      </u>			
10. 您送礼是为了(可以选择两项):      (编码时拆分为 2 个变量→)		1	(33)
		4	(34)
(1) 想让对方开心      (2) 礼貌      (3) 习惯      (4) 维持来往			
(5) 业务需要      (6) 希望对方帮忙办事      (7) 感谢对方的帮忙办事      (8) 其他 <u>      </u>			
11. 您一般在一次送礼上的花费大约是 <u>100</u> 元		0100	(35—38)
13. 从 2004 年 7 月到 2005 年 7 月,您用于礼物的花费大约是 <u>2500</u> 元		02500	(39—43)

### 1.1.2 编码应注意事项

#### 1. 多项选择题(即多选项)的处理

如第 9 题“您一般送礼的对象”是一种多项选择题,这种情况在编码时,可以将多项选择题的 7 个选项作 7 个变量处理,采用“1,2”编码,“1”表示“是”,即选择了这一项,“2”表示“否”,即没有选择这一项。上述问卷样例中,被调查者送礼对象选择了“朋友”和“同学”,而其他的选项没有被选择,因此,它的编码是 2 2 1 1 2 2 2。如果仅选择了“家人”,其他的选项没有选择,则编码为 1 2 2 2 2 2 2。

注意:关于多项选择题的详细知识可参阅第 10 章。

## 2. 排序题的处理

对于排序题的编码,其处理方法是:可以把需要排序的那些选项作为几个不同的变量来处理,如第7题,选择两个最经常送礼的场合,可以将“第一经常场合”和“第二经常场合”作为两个变量来处理,被调查者“第一经常场合”的选择为“婚嫁”,编码为“01”,“第二经常场合”的选择为“探病”,编码为“06”,该选项的最后编码为01 06。

## 3. 缺失值的处理

问卷中,若遇到被调查者不回答的一些选项,则应该当作缺失值。缺失值又称为用户缺失值(User Missing Value)。如果百岁以内(两个栏位)的年龄有缺失值,则可编码为99(三个栏位时可编码为999)。又比如,6个栏位的年收入的缺失值可编码为999999。如果可以选择的项目有9项,但被调查者正好选择了第9个选项,则以“0”表示缺失值。当然,缺失值也可用研究者自己能够识别的其他数字来表示,如“0”或“9”为用户缺失值。用户缺失值与系统缺失值(System Missing Value)的含义不同,系统缺失值主要是指计算机默认的缺失方式,如果在输入数据时空缺了某些数据或输入了非法的字符,计算机就把其界定为缺失值,这时的数据标记为“·”,而用户界定的缺失值则不会在数据显示时出现“·”。

## 4. 对“不适用情况”的处理

当碰到被调查者不适用的选项时,被调查者不需要对这些题进行回答,这种情况下,研究者可以用“8”、“98”等值来赋值表示“不适用”,若答案正好是8,可以以0替补。

# 1.2 在 SPSS 数据窗口直接输入数据

如果样本不大,变量不多,可以直接在SPSS窗口输入数据。输入方法如下。

## 1.2.1 首先在 Data View(数据浏览)窗口输入数据

方法:单击SPSS→File→New Data→Data View(数据浏览界面)命令→输入数据(见图1-1)→存成\*.sav(见图1-2~图1-3)。