

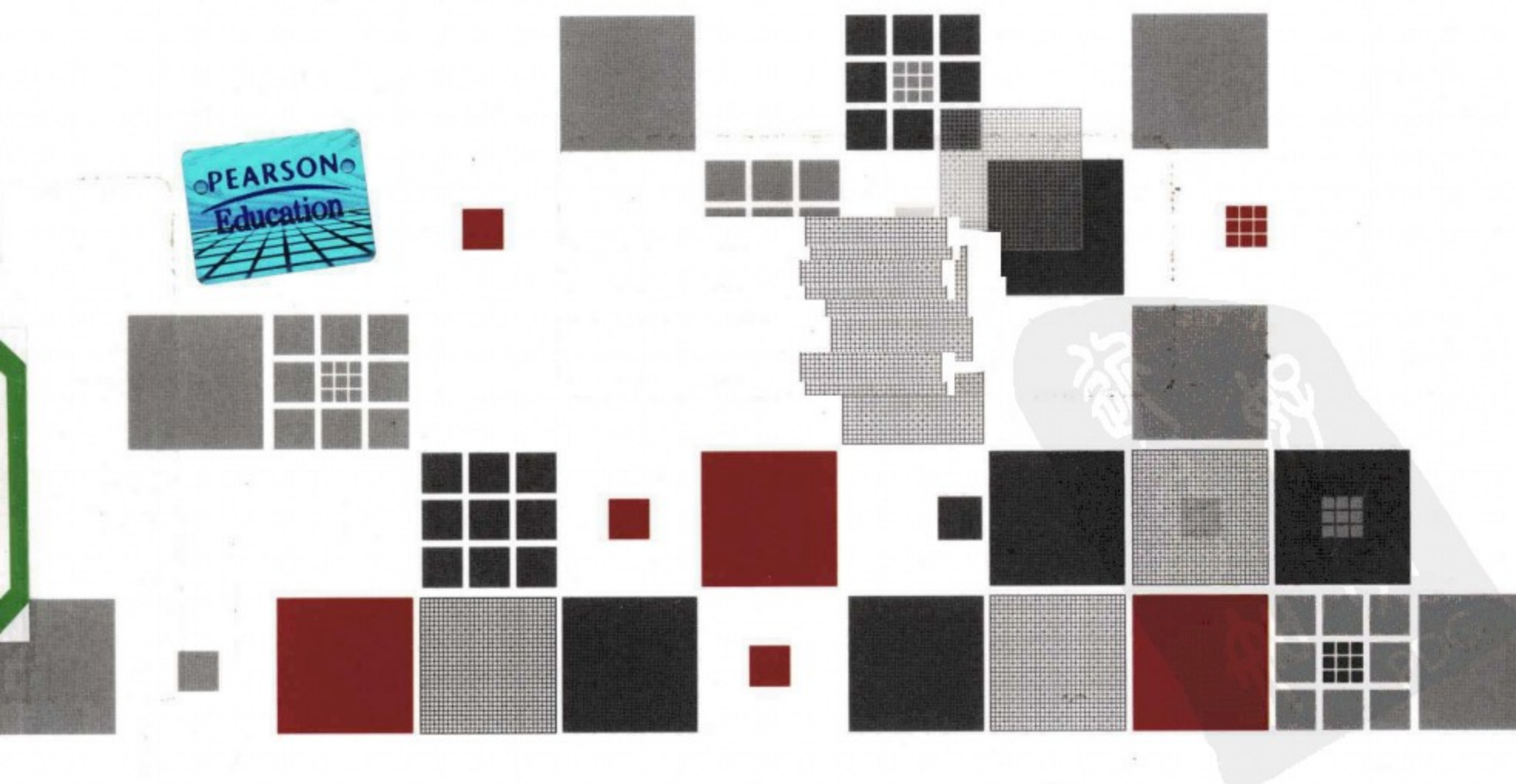


SPSS其实很简单

SPSS Demystified

罗纳德·D·约克奇 著
(Ronald D. Yockey)

刘超 吴铮 译
吴喜之 审校



SPSS其实很简单

SPSS Demystified

软件的发展，使统计从专业方法变成大众的游戏。只要输入格式无误的数据，就能得到漂亮的结果，然而最重要的问题——方法的选择以及结果的解读却被忽略。本书力图打破这种局面：

- ◎ 从实际问题入手，剥离出需要研究的问题，帮助读者理解如何选择恰当的统计方法。
- ◎ 从使用SPSS生成变量开始，到最终实现撰写APA（美国心理协会）格式的结果，提供SPSS每一操作步骤的截图，并对输出结果进行解读，帮助读者在面对大量输出结果时，快速有效地找到所需部分，并做出合理分析。
- ◎ 总结统计方法使用的前提假设和利用SPSS进行各种统计分析的程序步骤，带领读者理解统计方法的实质。

PEARSON

www.pearsonhighered.com

著作权合同登记号

图字：01-2010-0805

- 人大经管图书在线：www.rdjg.com.cn
- 读者信箱：rdcbsjg@crup.com.cn

ISBN 978-7-300-11797-3

上架指导 统计

ISBN 978-7-300-11797-3



9 787300 117973 >

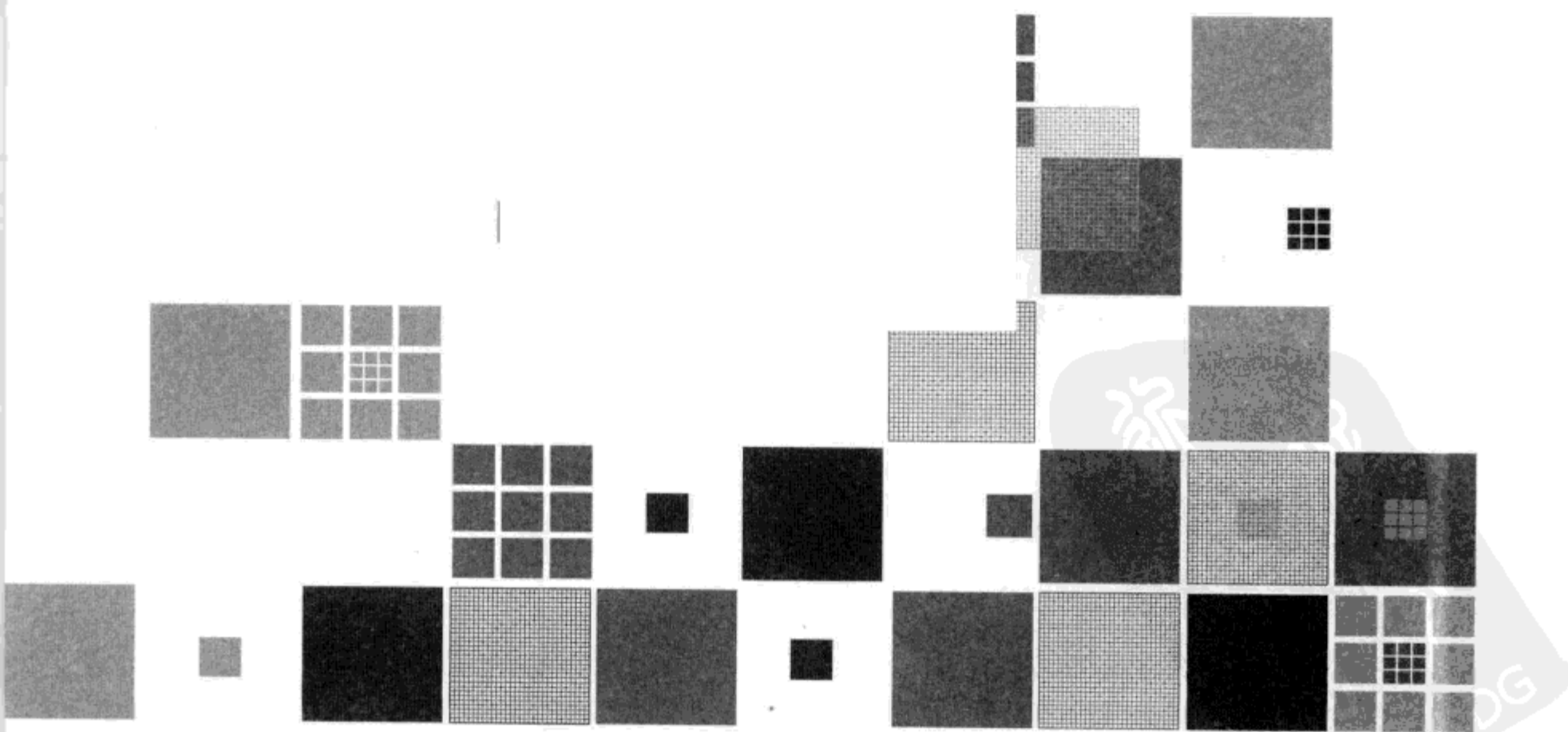
定价：39.00元

SPSS其实很简单

SPSS Demystified

罗纳德·D·约克奇 著
(Ronald D. Yockey)

刘超 吴铮 译
吴喜之 审校



中国人民大学出版社

·北京·

图书在版编目 (CIP) 数据

SPSS 其实很简单/约克奇著;刘超,吴铮译.
北京:中国人民大学出版社,2010
(管理者终身学习)
ISBN 978-7-300-11797-3

- I. ①S…
II. ①约…②刘…③吴…
III. ①统计分析-软件包, SPSS-教材
IV. ①C819

中国版本图书馆 CIP 数据核字 (2010) 第 035922 号

管理者终身学习
SPSS 其实很简单
罗纳德·D·约克奇 著
刘超 吴铮 译
吴喜之 审校
SPSS Qishi hen Jiandan

出版发行	中国人民大学出版社		
社 址	北京中关村大街 31 号	邮政编码	100080
电 话	010-62511242 (总编室)		010-62511398 (质管部)
	010-82501766 (邮购部)		010-62514148 (门市部)
	010-62515195 (发行公司)		010-62515275 (盗版举报)
网 址	http://www.crup.com.cn		
	http://www.ttrnet.com (人大教研网)		
经 销	新华书店		
印 刷	北京山润国际印务有限公司		
规 格	175 mm×250 mm 16 开本	版 次	2010 年 6 月第 1 版
印 张	21.25 插页 1	印 次	2010 年 6 月第 1 次印刷
字 数	336 000	定 价	39.00 元

版权所有 侵权必究 印装差错 负责调换

管 理 者 终 身 学 习

博学之，审问之，慎思之，明辨之，笃行之。

——《礼记·中庸》

新平知覺

PDG

为了JMJ，我的妻子米歇尔，我的儿子克里斯琴，塞缪尔，蒂莫西，威廉，斯蒂芬，你们是我生活中的喜悦和写作的灵感。感谢你，米歇尔，在我忙于写书的那些夜晚，照看我们的孩子们！

新平知聲
PDG

今天，在经济、财政、金融、营销、会计、管理及人文社会科学等领域都需要处理大量的信息。统计学能有效地处理各种信息问题，统计软件 SPSS 则为统计广泛而有效的使用提供了极大的便利。

作为一门应用性学科，多数人学习统计的目的是希望能把统计应用到与自身工作有关的数据分析中，但是在实践中，统计的使用出现泛滥的迹象，原因在于：统计软件的发展使得统计从统计学家的圈内游戏变成了大众的游戏，只要输入数据，选一些选项，马上就能得到令人惊叹的漂亮结果。只要数据格式无误、选项不矛盾、不用零作为除数，统计软件就一定会给你结果，而且几乎没有任何警告。另外，统计软件输出的结果太多。即使以同样的方法，不同软件输出的内容也有差异，甚至有时同样的内容名称也会不同。这就使得使用者大伤脑筋，即使是统计学家也不一定能解释所有的输出结果。因此，使用者需要特别留神，明确自己的目的，不要在得到一堆毫无意义的垃圾之后还沾沾自喜。

本书能够帮助读者有效地理解统计原理和学习利用 SPSS 进行数据分析。读者在学习本书后，既能根据实际工作需要选择合适的统计方法，又能利用 SPSS 软件整理、分析数据，解释数据分析结果，从而大大提高分析和解决问题的能力。

本书的最大特色就是应用性强。本书与注重数学证明的传统统计学著作和就操作谈操作的 SPSS 软件使用手册的不同之处在于：略去了复杂的数学推导过程，强调统计学在实践中的应用。每一章先引出需要研究的一个实际问题，再帮助读者理解哪种统计方法适用于该问题，并结合实际案例介绍在经济、管理、人文社会科学领域中常用的统计学原理、方法与技术。通过许多鲜活的案例分析，本书还统计应用以本来面目，将统计变成人人都能够理解和掌握的有用工具，让读者感受到统计其实很简单。

本书的另一个特色是详细解读 SPSS 统计分析结果。一般介绍 SPSS 的书往往对软件操作介绍得非常翔实，但忽略了对 SPSS 输出的统计结果的解释，而这正是读者在学习统计和实际应用时感到困难的地方。本书在这方面的内容足以使读者学会如何使用统计，体会到数据分析的乐趣。

本书设计成让读者自己动手学习。读者可以一边阅读，一边使用计算机进行分析。为了帮助读者始终走在正确的学习道路上，本书循序渐进地介绍了 SPSS 的使用。从用 SPSS 生成变量开始，到写出标准格式的结果为止。SPSS 操作的每一步都用截图显示，并在旁边的批注框中注明了输出结果中的重要信息，让读者在使用 SPSS 的过程中获得友好界面，直观地看到一些比较复杂的统计分析的实现过程，从而获得成功的体验，减少对数据分析的焦虑。在此基础上，对 SPSS 得出的结果进行了详细的分析，使得读者在面对大量输出结果时，能够快速有效地找到自己需要的部分，对实际问题做出合理的解释，手把手教会读者进行成功的数据分析。在每章的末尾还总结了统计方法使用的前提假设和利用 SPSS 进行各种统计分析的程序步骤，帮助读者理解统计方法的实质，更好地利用 SPSS 完成各种统计分析。这一切都表明：SPSS 其实很简单。

需要强调的一点是，本书虽然使用 SPSS 15.0，但是对 SPSS 的操作介绍和统计分析解读同样适用于 SPSS（以及 SPSS 更名后的 PASW）的最新版本，比如 SPSS 17.0。

本书既可以作为研究生、MBA 和本科生的教材，也可以供从事商务活动和经济分析的各类人员参考。

本书由刘超、吴铮、吴喜之翻译。刘超翻译了第 4 章到第 16 章，吴铮翻译了第 2 章、第 3 章和附录。特别感谢中国人民大学统计学院的吴喜之教授，他不仅翻译了前言和第 1 章，还为本书确定了吸引人的中文译名，而且审校了全书。

由于译者的水平有限，翻译中难免存在疏漏和错误，恳请读者批评指正。欢迎使用 SPSS 进行教学或数据分析的朋友们与我们一起交流、探讨，联系邮箱是：statchao@gmail.com。

刘超 吴铮

2010 年 1 月 25 日

毋庸置疑，对于学习社会学和行为学的学生来说，统计学是最令人头疼的课程之一。从第一次统计课开始，学生们就普遍体验到了对该课程的恐惧和高度焦虑，或者极其渴望了解统计的内容。因此，SPSS 虽然是统计软件中最易于使用的一种，但在焦虑的学生眼里，这意味着他们在学习统计之余，还要掌握一个新软件，这几乎成为不可能完成的任务。

了解到学生们对统计学的普遍畏惧（并且这种畏惧会影响学习的成绩），我在我的课程中加入了一些有助于缓解学习者畏惧情绪的特别布置，这些布置也可以帮助他们获得在 SPSS 数据分析上的成功经验，本书的特色如下。

本书的特征

首先，本书设计为动手实践型，读者可以在阅读每个章节时跟随指导同时在计算机上进行分析操作。为了帮助读者跟上进度，我们采取了循序渐进的方法：从使用 SPSS 生成变量开始，到最终实现撰写 APA（美国心理协会）格式的结果。本书提供 SPSS 每一个操作步骤的截图，并且在旁边的批注框中注明了输出结果中的重要信息。以上特征能够让 SPSS 更加容易被接受，增加软件学习过程的顺利程度，从而减轻学习者对该科目的畏惧。所有内容特征都在图表 1 中给出了详细的描述。

除了图表 1 中的特征以外，在每一章结尾都有练习，练习的答案列示在附录 C 中。我们鼓励学生通过练习来得到更加精通统计及使用 SPSS 所需的必要经验。

本书的覆盖面及结构

本书主要是为统计和研究方法的入门课程设计的。从这个目的出发，我

们介绍了在这些课程中最常用的方法（比如， t 检验、一维组间 ANOVA、一维组内 ANOVA、二维组间 ANOVA、卡方、相关、回归）。本书还适用于中级统计及研究生的开始课程，因为书中包括了多元回归、组间及组内 ANOVA 和可靠性。

本书分为两部分。第一部分引进 SPSS 软件程序（第 1 章），覆盖了描述统计（第 2 章），讨论如何用 SPSS 来生成各种图形（第 3 章），用一章的篇幅讲述了如何利用 α 系数来估计一个量表的内部一致可靠性（第 4 章）。第二部分覆盖推断统计学，包括 t 检验（第 5~7 章），方差分析方法（第 8~11 章），相关（第 12 章），简单回归和多元回归（第 13、14 章），卡方方法（第 15、16 章）。另外，数据变换和在 SPSS 导入 Excel 文件分别安排在附录 A 和附录 B 中。

图表 1 本书内容特征

特征	描述
数据分析四步骤	在每一章 ^a ，数据分析过程都分为四个容易跟进的步骤，包括： 步骤 1 生成变量。包括需要时引入变量值标签 步骤 2 输入数据。描述了数据文件的正确构成 步骤 3 分析数据。描述了如何利用 SPSS 的下拉菜单来运行正确的分析 步骤 4 解释结果。对每个输出表都进行讨论，每次一张表，APA 格式的结果输出在第 5~16 章提供
截图	每章都有几个截图，帮助读者不偏离正确的学习路径
批注框	批注框用来强调重要的信息，并且在 SPSS 中潜在的混淆区域提示读者
研究问题及原假设和对立假设	对每种方法提供研究问题以及原假设和对立假设，以帮助读者更好地把数据和研究问题以及感兴趣的假设结合起来（应用于第 5~16 章）
效应量	介绍了如何生成、报告、解释效应量（报告效应量是 APA 推荐的，并且是按几个杂志的投稿要求做的，应用于第 5~14 章和第 16 章）
假定	提供了每个推断方法的假定，以及违反假定对方法精确度的影响，以使读者来评估他们的数据是否符合一个感兴趣的统计方法的要求（应用于第 5~16 章）

注：a. 第 3 章不包括步骤 1 和步骤 2。

对于仅将本书用来辅助解决某些具体问题的 SPSS 新手来说，读完第 1 章

之后，可以直接翻到与自己相关的章节，按其指导程序输入和分析自己的数据。

SPSS 15.0 版（以及之前的版本）

虽然本书是基于 SPSS 15.0 版本的应用而写的，但是使用 SPSS 14.0、SPSS 13.0 和 SPSS 12.0 版本的用户也可以无障碍地根据本书的指导来分析数据。个别情况下，当本书中的命令出现不同版本之间的差异时，不同之处会有清楚的注明，并同时列出早期版本中该命令的各种形式（这种少量的差异出现在第 2 章的菜单选项和附录 A 中）。

本书也基本兼容 SPSS 11.5 和 SPSS 11.0 版本，SPSS 15.0 与之前版本的主要差别是 SPSS 11.5 和 SPSS 11.0 版本只允许不超过 8 个字符的变量名。如果你使用的是 SPSS 11.5 和 SPSS 11.0 版本，需要确保你在 SPSS 中生成的任何变量名都不超过 8 个字符。

致谢

本书的出版得益于很多人的帮助。首先我要感谢我所有的学生，多年来他们的问题和反馈让我在讲授统计学和 SPSS 的过程中，对“怎样教才能奏效”得到很多重要启发和宝贵信息。感谢培生出版公司的策划编辑 Jessica Mosher，在本书撰写之初她就表达了自己的信心。同时感谢培生出版公司的执行编辑 Jeff Marshall，他在本书撰写过程中提供了大量有价值的反馈信息。我还要感谢所有的审阅者，让我获得很多有益的评论和建议，他们是：Iona 大学的 Linda Refsland；北科罗拉多大学的 Thom Dunn；Adelphi 大学的 Carolyn Springer；St. John Fisher 大学的 Timothy Franz；休斯敦大学的 Joshua Priddy；Michigan-Dearborn 大学的 Dan Swift；佛罗里达大学的 Eve Brank；北卡罗来纳大学 Charlotte 分校的 Arnie Cann；Rochester 大学的 Rafael Klorman，得克萨斯大学的 Paul Westmeyer；佛罗里达大学的 Brian Stults；New Hampshire 大学的 John Sparrow；加利福尼亚州立大学 Sacramento 分校的 Qin Jianjian；以及得克

萨斯大学 San Antonio 分校的 James Dykes。最后我要感谢这些年来在很多方面给予我特殊影响并鼓励我写这本书的人们：Diana Divecha, Barbara Dodd, Steve Fitzpatrick, Rachel Fouladi, Barry Godolphin, Earl Jennings, Paul Kelley 以及 Bill Koch。

我的目标是通过本书揭开用 SPSS 进行数据分析的神秘面纱，最终还是要由你们来评价我的努力有没有白费。真心希望这本书能对你成功分析和解读数据有所帮助。最后，欢迎你将使用本书的经验和意见反馈给我，我的邮箱是 ryockey@csufresno.edu (请在邮件主题内写明 "SPSS Demystified")。

在了解了本书的表达模式和内容范围后，你可以尽情开始对 SPSS 的探索！

罗纳德·D·约克奇



第一部分 SPSS 介绍, 描述统计, 数据的图形展示, 利用系数 α 描述的可靠性

第 1 章 SPSS 概述 3

打开 SPSS 3

数据编辑窗口 5

在 SPSS 中生成数据文件 9

数据输入及分析 11

Viewer (输出) 窗口 18

保存文件 19

打印文件 23

练习 24

第 2 章 描述统计: 频数、集中趋势的度量以及 变异性的度量 27

频数 30

度量集中趋势与变异性 31

使用 Means 程序进行的组分析 32

执行 Frequencies 和 Means 程序的步骤摘要 40

练习 41

第3章 作图 45

条形图 46

直方图 47

散点图 49

盒形图 50

在 SPSS 中建立条形图、直方图、散点图和盒形图的步骤摘要 55

练习 56

第4章 可靠性 (用 α 系数度量) 59

例子 60

α 系数的目标和数据要求 61

在 SPSS 中输入数据及分析 62

结果的表达 67

在 SPSS 中执行 α 系数的步骤摘要 67

练习 68

第二部分 推断统计学

第5章 单样本 t 检验 77

例子 77

单样本 t 检验的目标和数据要求 77

原假设和对立假设 78

研究问题 78

在 SPSS 中输入数据及分析 79

效应量 82

APA 格式的结果表达 83

单样本 t 检验的假定 83

在 SPSS 中执行单样本 t 检验的步骤摘要 84

练习 84

第 6 章 独立样本 t 检验 87

例子 87

独立样本 t 检验的目标和数据要求 87

原假设和对立假设 88

研究问题 88

在 SPSS 中输入数据及分析 89

效应量 95

APA 格式的结果表达 95

独立样本 t 检验的假定 96

在 SPSS 中执行独立样本 t 检验的步骤摘要 96

练习 97

第 7 章 相依样本 t 检验 99

例子 99

相依样本 t 检验的目标和数据要求 99

原假设和对立假设 100

研究问题 100

在 SPSS 中输入数据及分析 101

效应量 105

APA 格式的结果表达	105
相依样本 t 检验的假定	106
在 SPSS 中执行相依样本 t 检验的步骤摘要	106
练习	107

第 8 章 一维组间方差分析 109

例子	109
一维组间方差分析的目标和数据要求	109
原假设和对立假设	110
研究问题	110
在 SPSS 中输入数据及分析	111
效应量	119
APA 格式的结果表达	120
一维组间方差分析的假定	120
在 SPSS 中执行一维组间方差分析的步骤摘要	121
练习	122

第 9 章 二维组间方差分析 125

例子	125
二维组间方差分析的目标和数据要求	126
原假设和对立假设	126
研究问题	127
在 SPSS 中输入数据及分析	128
当交互效应显著时分析主效应	140
效应量	141

APA 格式的结果表达	141
二维组间方差分析的假定	142
在 SPSS 中执行二维组间方差分析的步骤摘要	143
练习	144

第 10 章 一维组内方差分析 147

例子	147
一维组内方差分析的目标和数据要求	147
原假设和对立假设	148
研究问题	148
在 SPSS 中输入数据及分析	149
效应量	161
APA 格式的结果表达	161
一维组内方差分析的假定	162
在 SPSS 中执行一维组内方差分析的步骤摘要	163
练习	164

第 11 章 组间组内方差分析 167

例子	167
组间组内方差分析的目标和数据要求	168
原假设和对立假设	168
研究问题	169
在 SPSS 中输入数据及分析	170
效应量	185
APA 格式的结果表达	185

组间组内方差分析的假定 186

在 SPSS 中执行组间组内方差分析的步骤摘要 187

练习 188

第 12 章 皮尔逊 r 相关系数 191

例子 191

皮尔逊 r 相关系数的目标和数据要求 191

原假设和对立假设 192

研究问题 192

在 SPSS 中输入数据及分析 193

效应量 196

APA 格式的结果表达 196

皮尔逊 r 相关系数的假定 197

在 SPSS 中执行相依样本 t 检验的步骤摘要 197

练习 197

第 13 章 简单线性回归 201

例子 201

简单回归的目标和数据要求 201

原假设和对立假设 202

研究问题 202

在 SPSS 中输入数据及分析 203

效应量 209

APA 格式的结果表达 210

简单回归的假定 210

在 SPSS 中执行简单回归分析的步骤摘要 210

练习 211

第 14 章 多元线性回归 215

例子 215

多元回归的目标和数据要求 216

原假设和对立假设 216

研究问题 217

在 SPSS 中输入数据及分析 218

效应量 226

APA 格式的结果表达 226

多元回归的假定 227

在 SPSS 中执行多元回归分析的步骤摘要 227

练习 228

第 15 章 χ^2 拟合优度检验 231

例子 231

χ^2 拟合优度检验的目标和数据要求 231

原假设和对立假设 232

研究问题 232

在 SPSS 中输入数据及分析 235

APA 格式的结果表达 240

χ^2 拟合优度检验的假定 241

在 SPSS 中执行 χ^2 拟合优度检验的步骤摘要 241

练习 242

第16章 χ^2 独立性检验 245

例子 245

χ^2 独立性检验的目标和数据要求 246

原假设和对立假设 246

研究问题 246

在 SPSS 中输入数据及分析 247

效应量 255

APA 格式的结果表达 255

χ^2 独立性检验的假定 256

在 SPSS 中执行 χ^2 独立性检验的步骤摘要 256

练习 258

附录 A 数据变换及其他方法 261

重新编码方法 262

计算方法 267

选择案例方法 270

拆分文件方法 275

关闭拆分文件方法 277

附录 B 导入文件 279

附录 C 各章练习答案 285

注释 312

参考文献 318

Part 1

第一部分

SPSS 介绍，描述统计，数据的图形展示，利用系数 α 描述的可靠性

SPSS

在这一部分，第 1 章介绍了 SPSS 软件程序，包括生成变量、输入和分析数据、保存文件、打印结果。第 2 章描述了利用 SPSS 来计算一些不同的概括统计量，包括频数、集中趋势的度量、变异性的度量。第 3 章描述了如何在 SPSS 中生成一些不同的图形，包括条形图、直方图、散点图、盒型图。最后，第 4 章描述了如何通过 SPSS 来估计一个使用系数 α 表示的相合可靠性。


数字资源
PDG

SPSS 概述

在这一章，将介绍 SPSS 软件程序，包括打开 SPSS、生成变量、输入数据、利用下拉菜单执行基本的分析、保存文件、打印结果。让我们开始学习 SPSS 吧。


打开 SPSS

在打开 SPSS 之前，必须先看它是否在你的计算机里。按照计算机的配置，SPSS 应该至少下面的位置之一：


1. 桌面上的图标¹ () (桌面是计算机屏幕上工作空间的主要区域)。
2. 程序菜单中。
3. 快速运行工具条中 (快速运行工具条位于计算机屏幕下方“开始”按钮的右边)。

下面叙述分别从这些位置来打开 SPSS (只须选择其中一种方法来打开 SPSS，这样不会在计算机上重复打开 SPSS)。

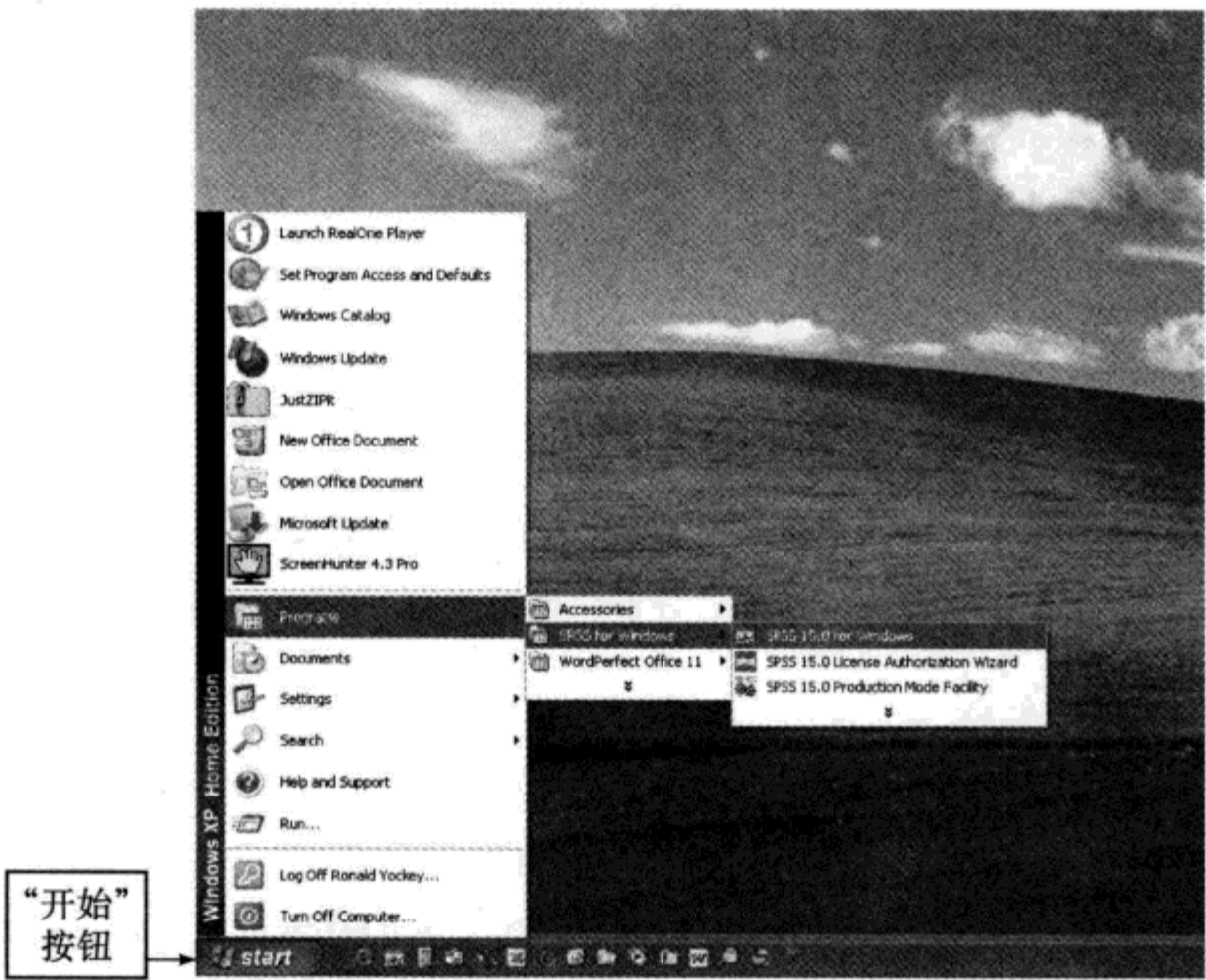
用屏幕上的图标打开 SPSS：

1. 找到桌面上的 SPSS 图标 ()。
2. 双击² SPSS 图标，打开 SPSS。

用程序菜单打开 SPSS：

1. 点击“开始”按钮 ()。
2. 单击程序 (或者所有程序)。


- 3. 选择 SPSS for Windows。
- 4. 选择 SPSS 15.0 for Windows（如果你没有 15.0 版本的 SPSS，选择你有的版本，如 14.0、13.0、12.0、11.5 版本），详见图表 1—1。^①



图表 1—1 用程序菜单打开 SPSS

说明：如果你的计算机屏幕和图中显示的不同，SPSS 应该在计算机的程序菜单里。

从快速运行工具条打开 SPSS：

- 1. 根据你的计算机配置，可能有一个 SPSS 图标（）在“开始”按钮的右边。如果你看见了它，可以单击它来打开 SPSS 程序（在图表 1—1 中，SPSS 图标是“开始”按钮右边的第二个图标）。

如果你还没有这样做，利用上面描述的方法打开 SPSS。

打开 SPSS 后，可能会出现一个对话框³，在该对话框的上部会出现“*What would you like to do*（你想要做什么）？”（见图表 1—2）。这个对话框是用来帮助

^① 需要强调的是，本书虽然使用 SPSS 15.0，但是对 SPSS 的操作介绍和统计分析解读基本上也适用于 SPSS（以及 SPSS 更名后的 PASW）的最新版本，比如 SPSS 17.0。——译者注

你开始运行 SPSS 的。我们在本书中不用该对话框（我们将学习如何自己来使用 SPSS），如果你在计算机上看到它，点击“Cancel”（取消）按钮关闭它。

数据编辑窗口

打开 SPSS 后，Data Editor（数据编辑）窗口显示在屏幕上（见图表 1—3）。数据编辑窗口是用来在 SPSS 中生成变量和输入数据的。在该窗口的最上端显示了 SPSS 文件名（Untitled 表示这个文件还没有命名）。在文件名下面是菜单栏，它包括几个不同的菜单选项（File，Edit，View，等等），可以用来完成各种 SPSS 任务（如保存文件、打印结果等）。紧接在菜单栏下面的是

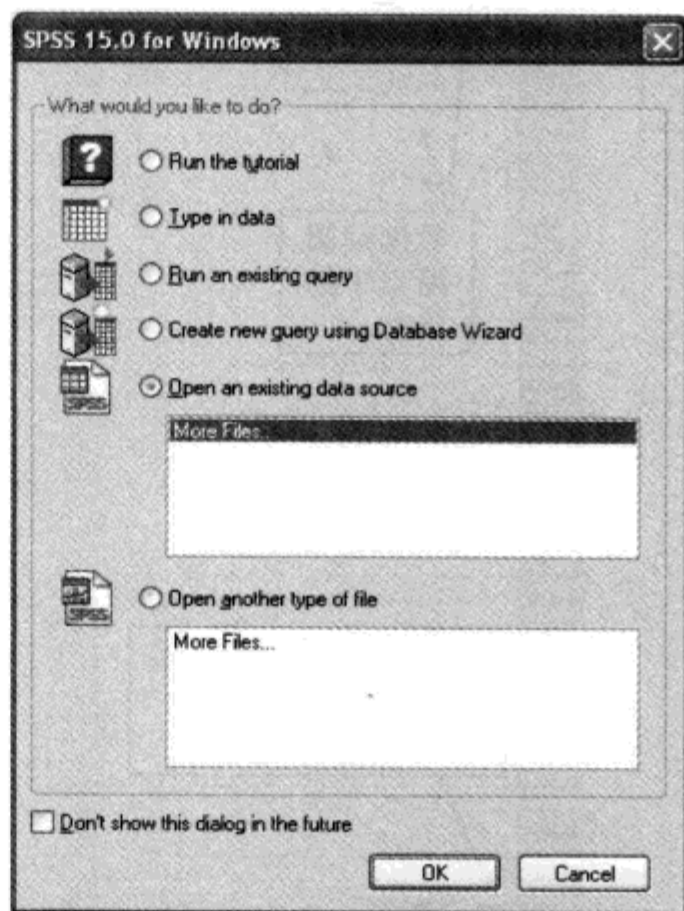
一些工具栏按钮，提供了 SPSS 中几个不同选项的快速通道。数据编辑窗口的主要区域由许多长方形空格组成（称为单元（cell）），用来输入数据。

数据编辑窗口的主要区域包含两个不同的窗口：Data View（数据查看）窗口和 Variable View（变量查看）窗口。每一个窗口都能够通过点击屏幕左下角相应的标签得到。显示白色背景的标签说明目前是打开的（注意图表 1—3 中 Data View 窗口目前是打开的）。如果 Data View 窗口目前不是打开的，点击“Data View”标签来打开它。

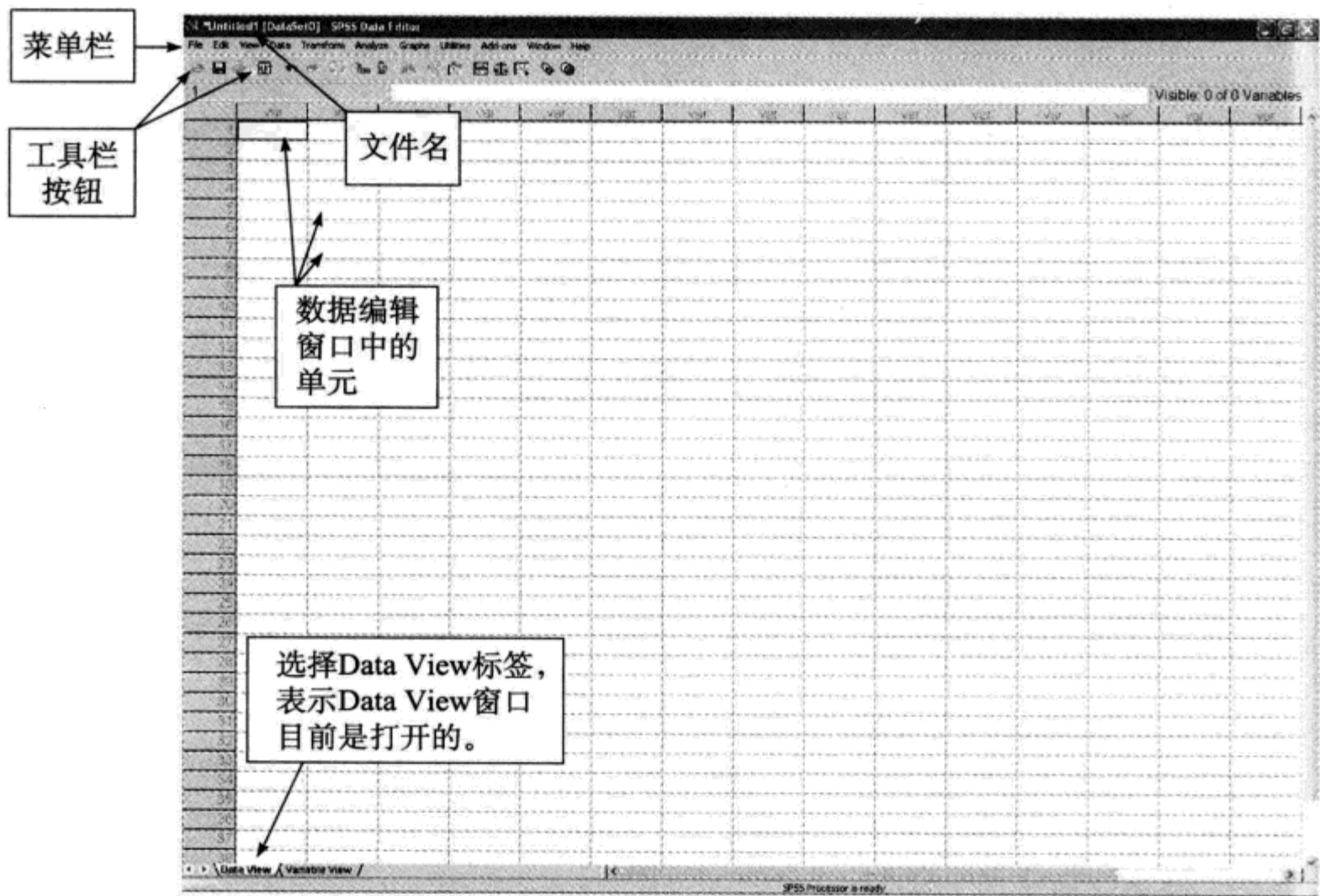
下面将讨论 Data View 窗口和 Variable View 窗口。

Data View 窗口

在图表 1—4 显示的 Data View 窗口中包含了一些单元，它们用于输入数据（数据通常包含数目，也可以是字母或符号）。⁴Data View 窗口的行是有数字标号



图表 1—2 SPSS 15.0 的“*What would you like to do*（你想要做什么）？”对话框



图表 1—3 SPSS 的数据编辑窗口 (选择 Data View 标签)

的 (图表 1—4 显示了第 1 行到第 38 行), 所有的列初始都显示 var。第一个单元 (左上角) 有一个深色边框, 意味着它是激活的或者准备接受输入。

Data View 窗口——SPSS 中用来输入数据的窗口。

Variable View 窗口

在 SPSS 中, Variable View 窗口用来生成变量及增加数据文件的信息。点击屏幕下方的 Variable View 标签可以打开 Variable View 窗口。因为 Data View 窗口目前是打开的, 所以需要点击 Variable View 标签来打开 Variable View 窗口。

- 1. 点击 Variable View 标签打开 Variable View 窗口。图表 1—5 显示的是 Variable View 窗口。

和 Data View 窗口一样, Variable View 窗口的每一行都有数字标号, 窗



图表 1—4 SPSS 的 Data View 窗口

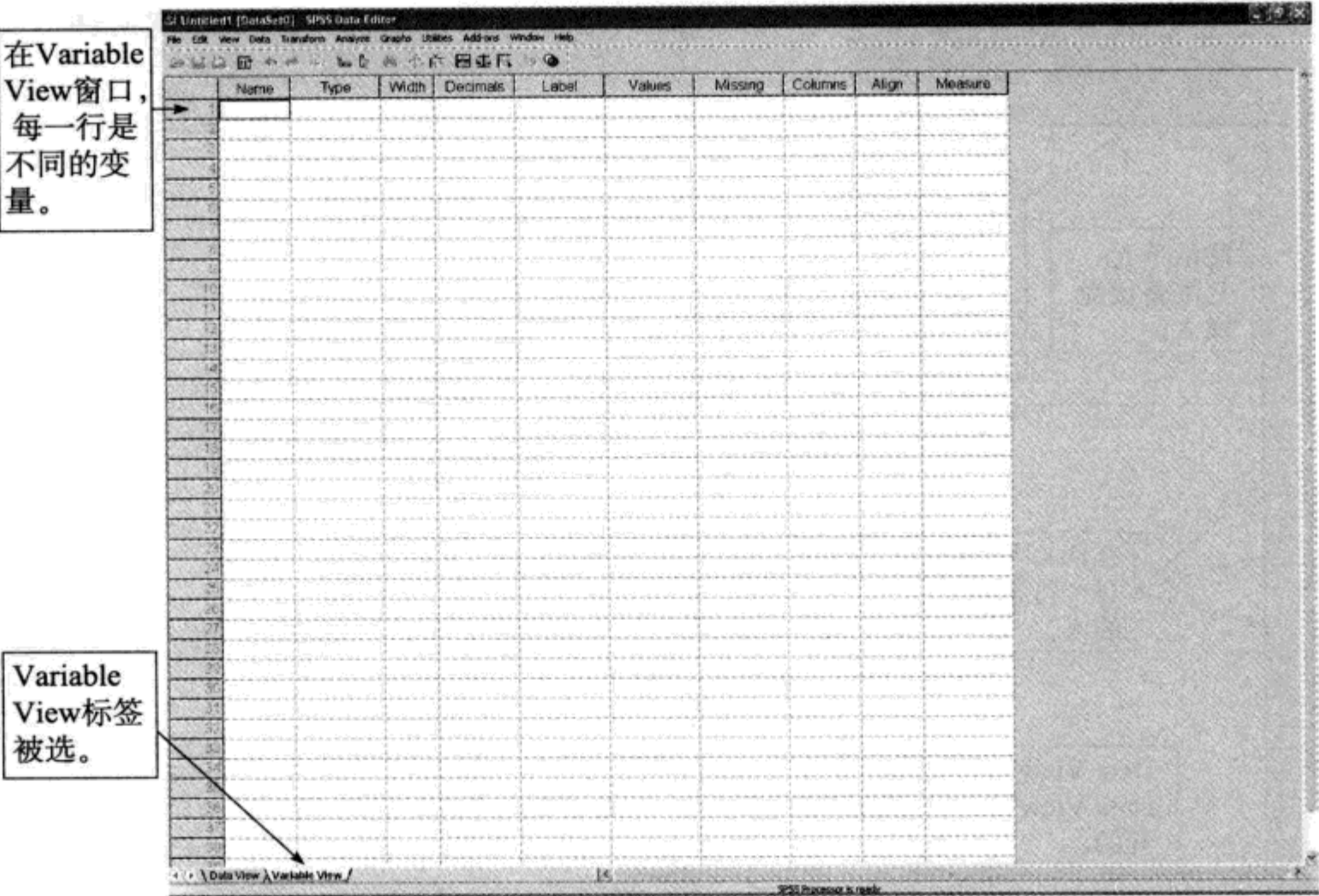
口的主要区域由一个个单元构成。然而，与 Data View 窗口不同，Variable View 窗口的每一列有不同的名字和功能。对 Variable View 窗口不同列的描述显示在图表 1—6 中。

在 Variable View 窗口列的属性中，本书将使用 Name 和 Values 选项（其他列用默认值⁵）。

Variable View 窗口——SPSS 中用来生成变量及增加数据文件信息的窗口。

数据编辑的关键特征已经引入，让我们在 SPSS 中生成一个数据文件。

记住：在数据编辑中，Variable View 是用来生成变量的，而 Data View 窗口是用来输入数据的。



图表 1—5 SPSS 的 Variable View 窗口

图表 1—6 Variable View 窗口不同列的名字和功能

列名	功能
Name (名字)	Name 用来给变量命名。在 SPSS 11.5 及以前的版本中，变量名的长度不能超过 8 个字符（一个字符是一个字母、数字或者符号）。在 SPSS 12.0 及以后的版本中，变量名的长度可以达到 256 个字符。每个变量应该以一个字母开始，而且两个变量不能是同样的名字。变量名中不允许有空格
Type (类型)	Type 表示变量的数据类型，包括：数值、逗号、句号、科学符号等。本书将用数值数据
Width (宽度)	宽度表示显示在 Data View 窗口的字符数目，默认值是 8
Decimals (小数位)	Decimals 表示在 Data View 窗口显示的小数位数，默认值是 2
Label (标签)	Label 是用来描述变量的。最多允许有 256 个字符。在 Label 列输入的信息将出现在输出中
Values (值)	Values 用来对分类变量编码（分类变量在下一部分讨论）。这个特征将在本书中反复出现

续前表

列名	功能
Missing (缺失)	Missing 表示所读的值是缺失的 (缺失数据意味着某些值没有提供或者对于一个变量来说是“缺失的”)。缺失数据的默认值为句号 “.”
Columns (列)	Columns 表示在 Data View 窗口的列的宽度, 默认是 8 个字符
Align (对齐)	Align 表示在 Data View 窗口的单元中把数据左对齐、右对齐, 或者居中对齐, 默认是右对齐
Measure (度量)	Measure 描述变量的度量水平。现有选项为 nominal (名义)、ordinal (定序), 或者 scale (数量)。默认值为数量

在 SPSS 中生成数据文件

一个 SPSS 数据文件 (data file) 是一个计算机文件, 包含一个或多个变量信息 (数据)。一个变量 (variable) 是一个属性 (attribute) 或者特征 (characteristic), 它可以取两个或更多的值。

数据文件——计算机文件, 包含一个或多个变量的信息。
变量——取两个或更多的值的属性或特征。

为了在 SPSS 中生成一个数据文件, 我们将使用在图表 1—7 中显示的数据。图表 1—7 包含了 5 个人的关于变量 **gender** (性别)、**age** (年龄)、**employment** (受雇) 和 **iq** (智商) (本书中变量的名字都用黑体表示) 的数据。注意在图表 1—7 中的每一行包含了一个不同的人在感兴趣的变量上的值。例如, 在第一行中, 给出了关于第一个人的值, 第一个人是 male (男性), 23 岁, employed (有工作), IQ (智商) 为 115。其余 4 个人的值在第 2~5 行中给出。

图表 1—7 输入 SPSS 的样本数据

Person	Gender	Age	Employment (Employment status)	IQ (IQ score)
Person 1	Male	23	Employed	115
Person 2	Male	19	Not Employed	90
Person 3	Female	32	Employed	120
Person 4	Female	28	Not Employed	90
Person 5	Male	18	Employed	116

说明: 图表中包含的 person 一列是为了描述说明, 不用输入 SPSS。

注意：在图表 1—7 中，变量 **age** 和 **iq** 是数值形式（它们的值是数值），而 **gender** 和 **employment** 为字符型（它们的值为字词）。要记住的一个重点是（对于本书中提到的大部分程序）：要在 SPSS 中对变量执行分析，必须用数值形式。因此，在把数据输入 SPSS 之前，需要把目前为字符型的变量（**gender** 和 **employment**）转换为数值型的，下面将说明这个过程。

把数值指定到 **gender** 和 **employment** 的类别

我们先把数值分配到 **gender** 的类别中。按照分配数值的规则，任何数值都可以分到一个变量的不同类别，只要每个类别被指定为不同的数目（你可以把这看成是一个度量的名义数值的例子）。为了描述这个过程，我们将指定 **male**（男性）为 1，而 **female**（女性）为 2。因此，对数据中的每个男性都输入 1，每个女性都输入 2。

对于 **employment**，将 **employed**（有工作的）指定为 1，**not employed**（失业的）指定为 2。

数据集（对于变量 **gender** 和 **employment** 输入了数值）显示在图表 1—8 中。

图表 1—8 修正的数据，对于 **gender** 和 **employment** 输入了数量值

Person	Gender	Age	Employment (Employment status)	IQ (IQ score)
Person 1	1	23	1	115
Person 2	1	19	2	90
Person 3	2	32	1	120
Person 4	2	28	2	90
Person 5	1	18	1	116

说明：对于 **gender**，1 = male，2 = female；对于 **employment**，1 = employed，2 = not employed。

诸如 **gender** 和 **employment** 这样的变量称为分类变量。分类变量（categorical variable）取有限数目的值，仅仅取两个值的变量（比如 **gender** 和 **employment**）称为二分变量（dichotomous variable）。变量 **age** 和 **iq** 为连续变量（continuous variable）。连续变量取大量不同的值。

- 分类变量——取有限数目值的变量。
- 二分变量——仅有两个值的分类变量。
- 连续变量——取大量不同值的变量。

数据输入及分析

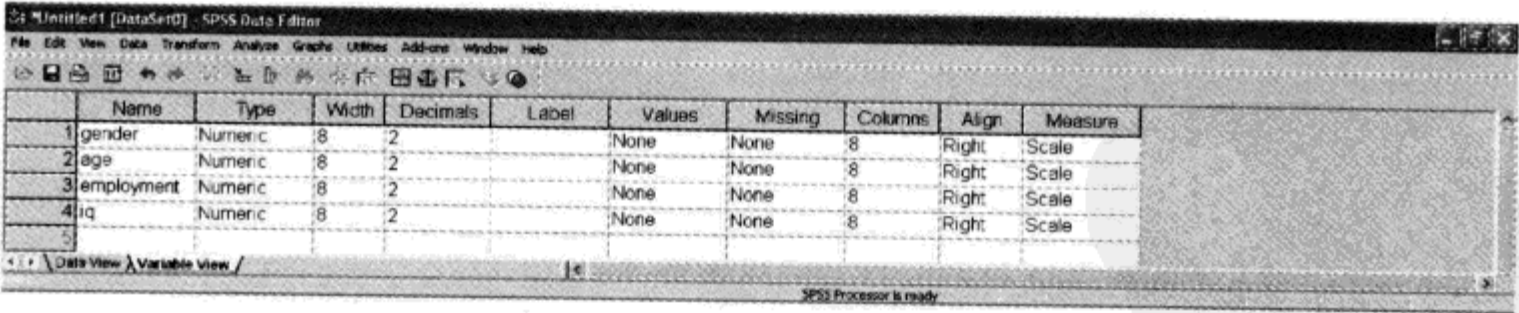
在所有变量都成为数值型的之后，现在我們能够在 SPSS 中输入数据和分
析数据。在本书中，数据输入和分析过程将分成下面四个步骤：（1）生成变
量；（2）输入数据；（3）分析数据；（4）解释结果。其中每一步都利用图表
1—8 的数据集来描述。

步骤 1：生成变量

我们将从生成变量 **gender**，**age**，**employment**⁶ 和 **iq** 开始。为了在 SPSS 中
生成变量，应遵从下面的指导。

在 SPSS 中生成变量：

- 1. 确认 Variable View 窗口是打开的。如果没有打开，点击屏幕左下角的
Variable View 标签。
- 2. Variable View 窗口左上角的第一个单元应该是激活的。如果没有激活，
点击它。
- 3. 在 Variable View 窗口的第 1 行，输入名字 **gender**，并点击向下箭头
(↓)。注意，所有右边的单元都自动填入默认值（Label 的默认值是空单元）。
- 4. 在第 2 行，输入名字 **age** 并点击向下箭头 (↓)。
- 5. 在第 3 行，输入名字 **employment** 并点击向下箭头 (↓)。
- 6. 在第 4 行，输入名字 **iq** 并点击向下箭头 (↓)。现在四个变量在 SPSS
中生成了。详见图表 1—9。



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	gender	Numeric	8	2		None	None	8	Right	Scale
2	age	Numeric	8	2		None	None	8	Right	Scale
3	employment	Numeric	8	2		None	None	8	Right	Scale
4	iq	Numeric	8	2		None	None	8	Right	Scale


图表 1—9 具有生成的变量 **gender**，**age**，**employment** 和 **iq** 的 Variable View 窗口

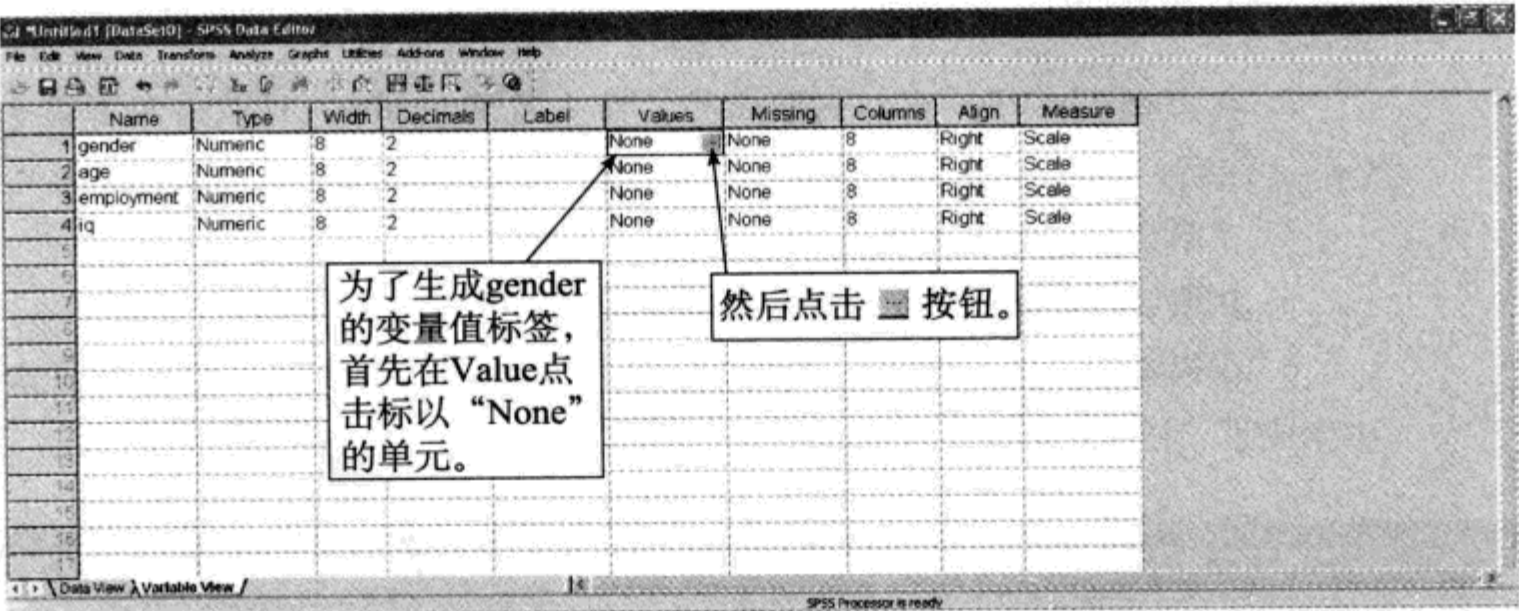
变量值标签

回顾我们把数值（如 1，2）指定到 **gender** 和 **employment** 不同的类别是因为它们需要改变为数值形式，以便在 SPSS 中进行分析。你可能记得，我们指定的数值是任意的（可以选择任意两个数值），只要它们能够区别变量的两个类别即可。为了帮助我们更容易地记住指定的数值，我们将用不同数值表示的类别输入 SPSS（1 为男性，2 为女性），这称为生成变量值标签（value labels）的过程。我们将对 **gender** 和 **employment** 生成变量值标签。

让我们先生成 **gender** 的变量值标签。

为 **gender** 生成变量值标签：

- 1. 确认 Variable View 窗口是打开的。在 Variable View 窗口的第一行（**gender** 行），点击在 “Value” 行的单元（那里显示的是 “None”）。在点击了标以 “None” 的单元之后，在该单元的右侧出现一个省略号（...）⁷。
- 2. 点击  按钮。详见图表 1—10。



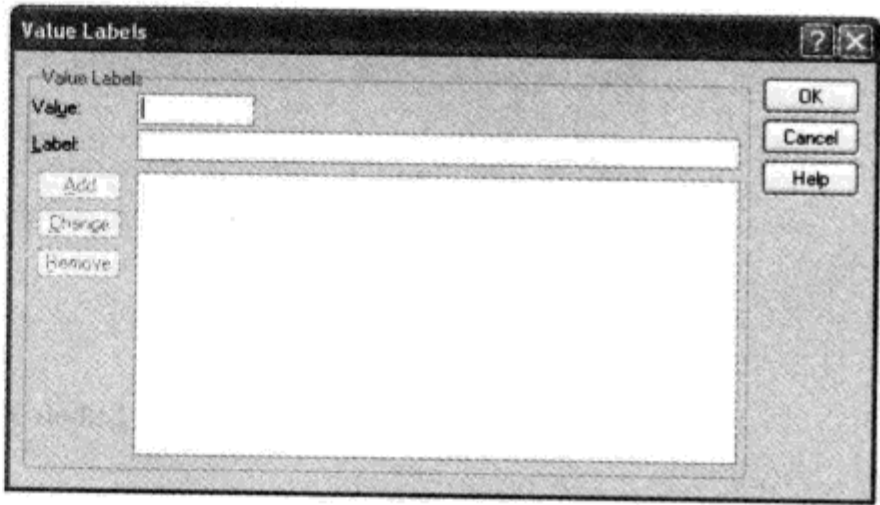
图表 1—10 选择 gender 的 Value 单元

- 3. Value Labels 对话框打开（见图表 1—11）。
- 4. 首先在 SPSS 中设定 male 的代码 1。输入 “1” 到 “Value” 的右边，输入 “male” 到 “Label” 的右边。详见图表 1—12。
- 5. 点击 “Add”，现在 “Add” 按钮右边显示 1.00 = “male”。详见图表 1—13。

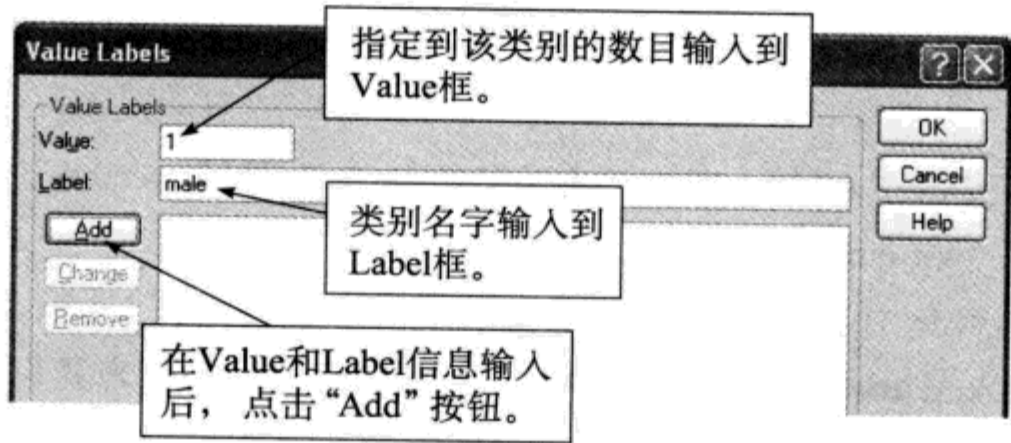
6. 下面在 SPSS 中设定 female 的代码 2。输入“2”到“Value”的右边，输入“female”到“Label”的右边。

7. 点击“Add”，现在“Add”按钮右边显示 1.00 = “male” 及 2.00 = “female”。见图表 1—14。

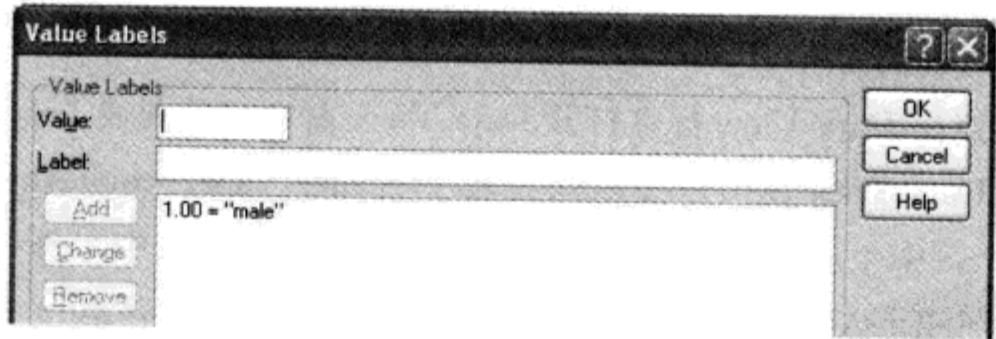
8. 点击“OK”。



图表 1—11 Value Labels 对话框



图表 1—12 Value Labels 对话框（续一）

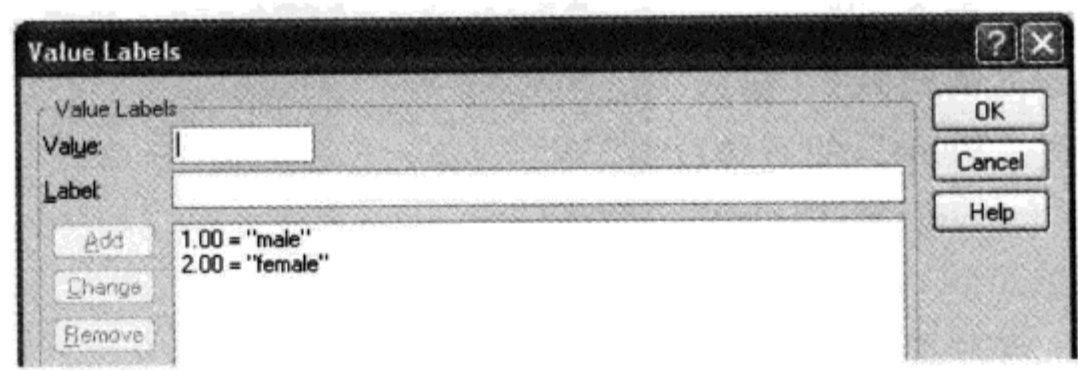


图表 1—13 Value Labels 对话框（续二）


输入完 gender 的变量值标签，我们将输入 employment 的变量值标签。有工作的指定为 1，失业者则指定为 2。

为 employment 生成变量值标签：

1. 在 Variable View 窗口的第 3 行（employment 行），点击在“Value”行的单元。



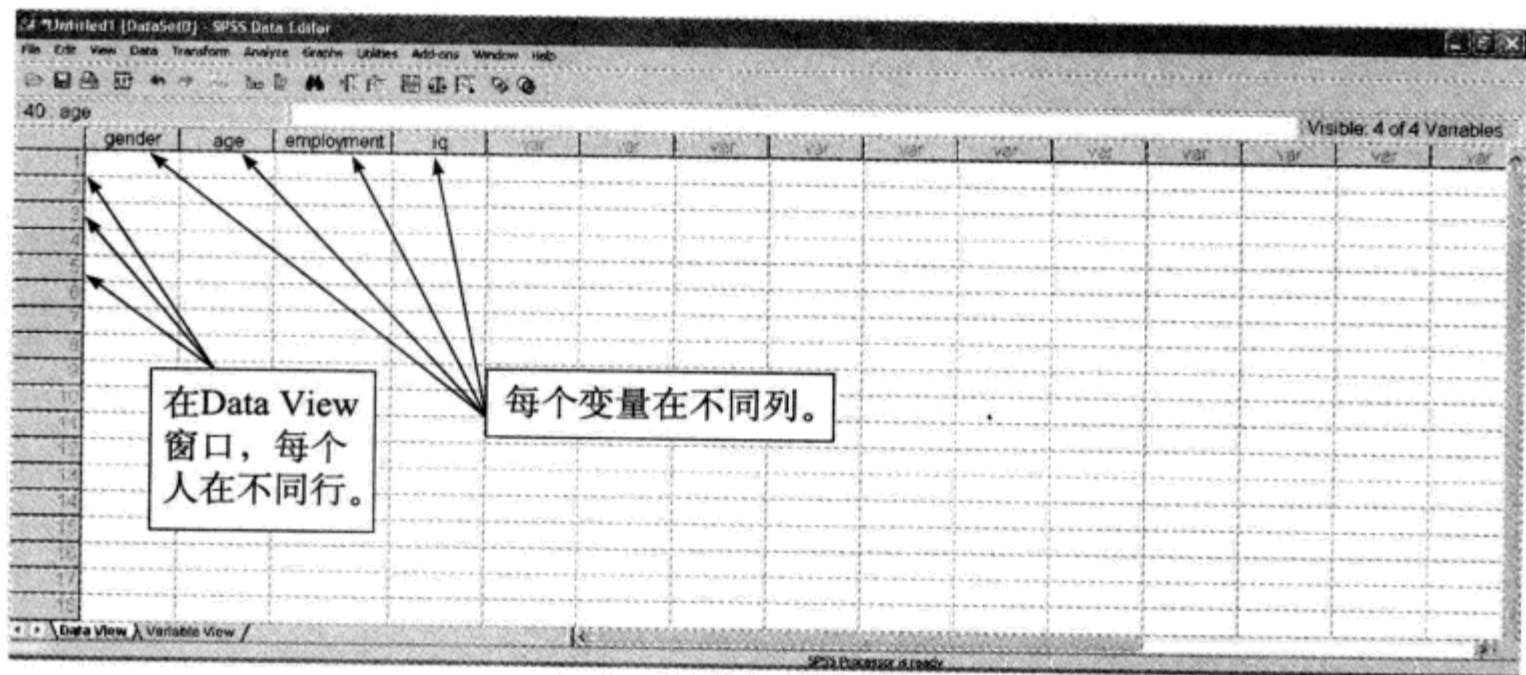
图表 1—14 Value Labels 对话框 (续三)

2. 点击  按钮。
 3. Value Labels 对话框打开，输入 “1” 到 “Value” 的右边，输入 “employed” 到 “Label” 的右边。
 4. 点击 “Add”，现在 “Add” 按钮右边显示 1.00 = “employed”。
 5. 输入 “2” 到 “Value” 的右边，输入 “not employed” 到 “Label” 的右边。
 6. 点击 “Add”，现在 “Add” 按钮右边显示 1.00 = “employed” 及 2.00 = “not employed”。
 7. 点击 “OK”。
- 输入完 **gender** 和 **employment** 的变量值标签，接下来我们将把数据输入 SPSS。

步骤 2：输入数据

为把数据输入 SPSS：

1. 点击 Data View 标签打开 Data View 窗口（见图表 1—15）。
在 Data View 窗口，头四列为名字 **gender**，**age**，**employment** 和 **iq**，对应于在 SPSS 中生成的四个变量。回忆在 Data View 窗口中，每一行对应于数据集中不同的个人，因此，当数据输入之后，第一行含有第一个人的值，第二行含有第二个人的值，等等。
2. 参见图表 1—8（数据），我们将为每个人输入四个感兴趣变量的值。为输入第一个参加者的值，点击 Data View 窗口第一行的第一个单元，为变量 **gender**，**age**，**employment** 和 **iq** 分别输入值 1，23，1，115（对一个变量输入数据的有效方法是在输入一个变量数据之后点击键盘的向右箭头键（→）。比如，对第一个人，你为 **gender** 输入 1，然后点击向右箭头键，为 **age** 输入



图表 1—15 SPSS 的 Data View 窗口，头四列为变量 gender，age，employment 和 iq

- 23，再点击向右箭头键，如此操作下去）。
3. 为输入第二个参加者的值，点击 Data View 窗口第二行的第一个单元，为变量 gender，age，employment 和 iq 分别输入值 1，19，2，90。
 4. 为剩下的三个参加者输入数据。完成的数据文件显示在图表 1—16 中。

	gender	age	employment	iq	var1	var2	var3	var4	var5	var6	var7	var8	var9	var10	var11	var12	var13	var14
1	1.00	23.00	1.00	115.00														
2	1.00	19.00	2.00	90.00														
3	2.00	32.00	1.00	120.00														
4	2.00	28.00	2.00	90.00														
5	1.00	18.00	1.00	118.00														

图表 1—16 SPSS 中完成的数据文件

现在数据已经输入，我们将在 SPSS 中执行基本分析。

步骤 3：分析数据

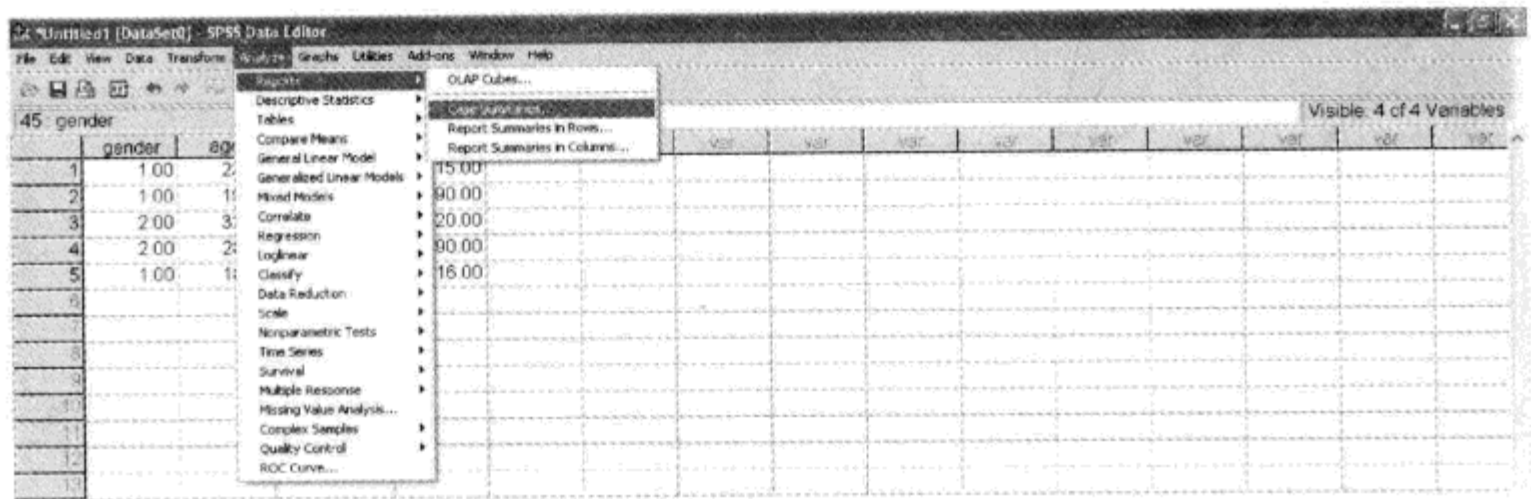
本书一直都将利用 SPSS 的下拉菜单来执行数据的统计分析。对于第一个分析，我们将从 SPSS 的一个基本方法开始——利用 Case Summaries 方法产生一个关于每个变量的概括报告。

为了在 SPSS 中执行 Case Summaries 方法：

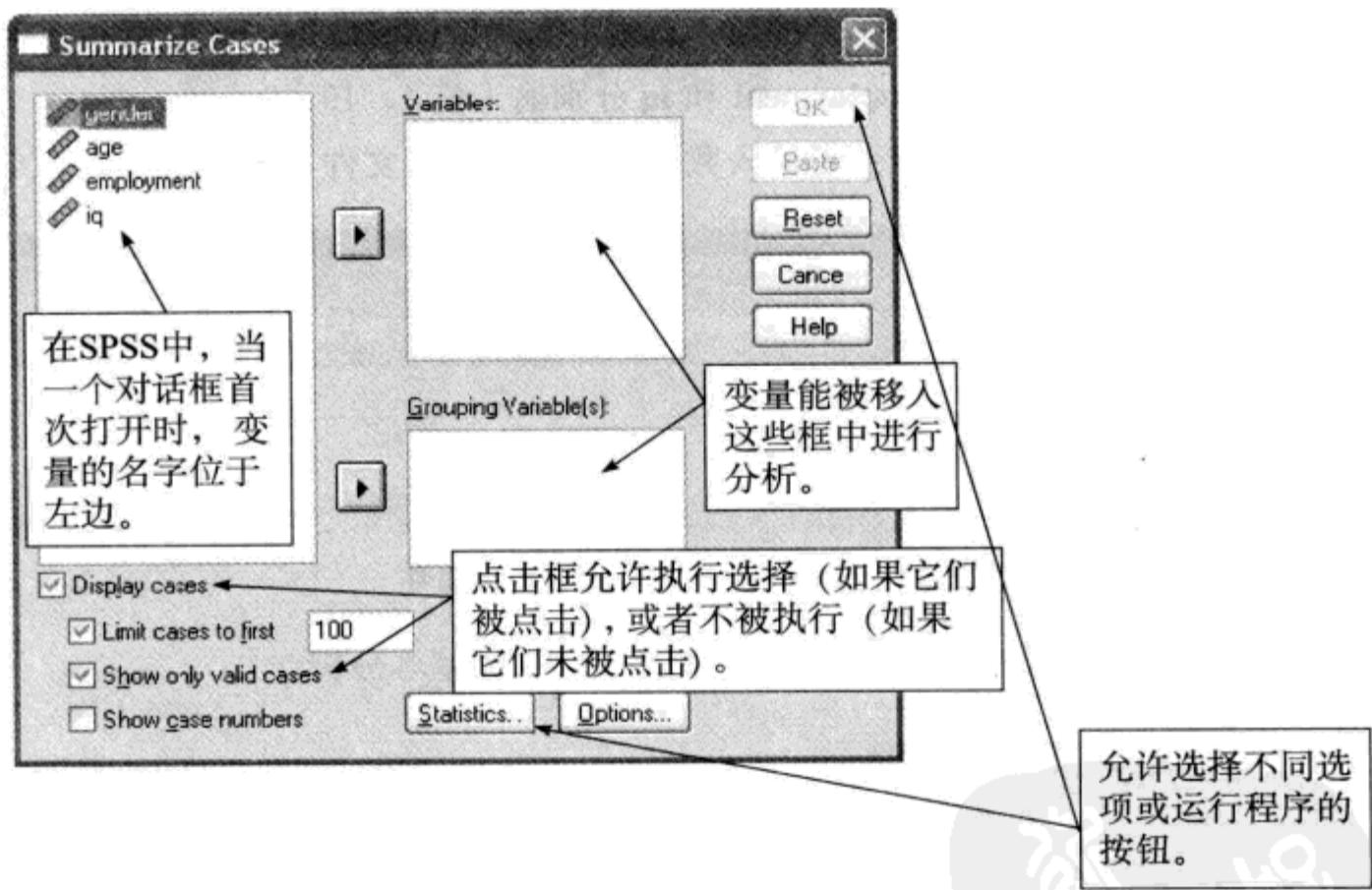
1. 从菜单栏中选择 **Analyze>Reports>Case Summaries...**（见图表 1—17）（当提供了菜单命令时，记号“>”意味着要做的下一个选择。**Analyze>Re-**

ports>Case Summaries 读作“选择 Analyze, 然后选择 Reports, 再选择 Case Summaries”)。

将出现一个 Summarize Cases 对话框（见图表 1—18）。



图表 1—17 SPSS 中的 Case Summaries 方法的菜单命令



图表 1—18 Summarize Cases 对话框

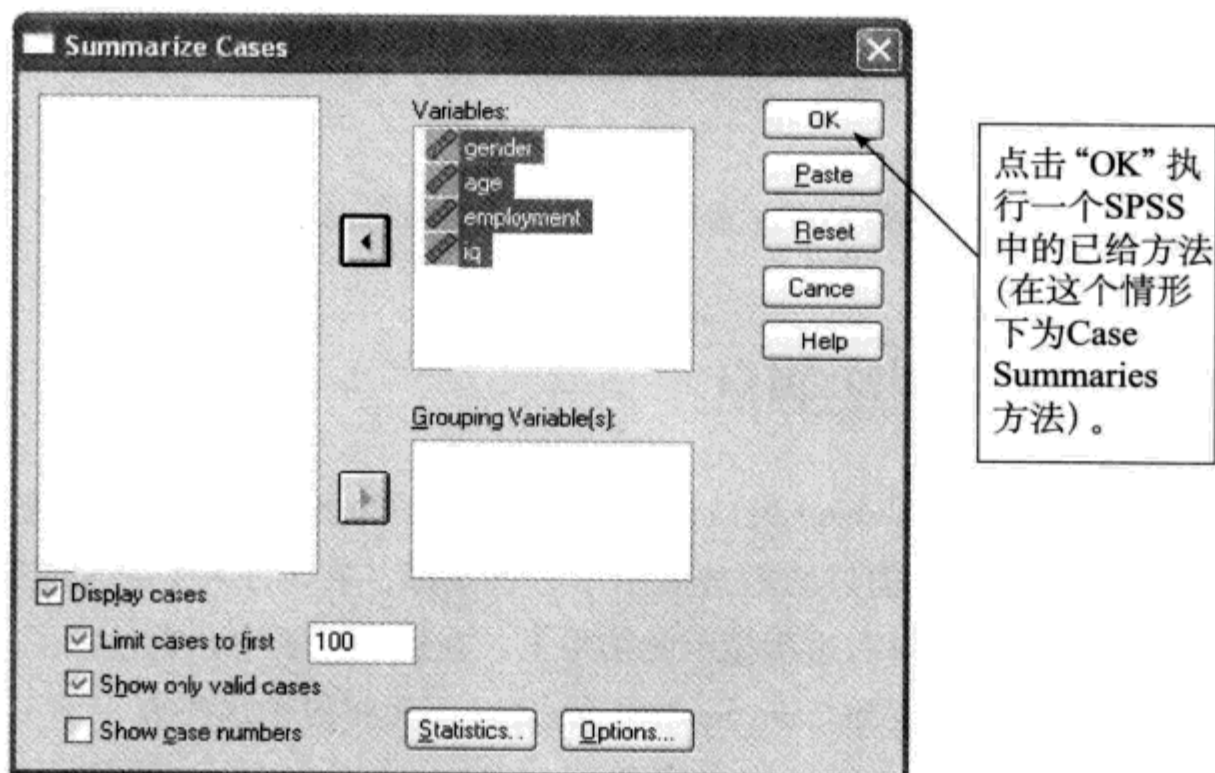
图表 1—18 中所示的 Summarize Cases 对话框是我们在 SPSS 中将要遇到的许多对话框的代表。首先注意到，当对话框打开时，变量 **gender**，**age**，**employment** 和 **iq** 出现在对话框左边，对话框中间为两个向右箭头（▶），表示左

边框中的变量能被移入右边的框以进行分析。还有一些其他按钮 (“Statistics”, “Options”, “OK”, “Paste”, “Reset”, 等等), 它们在 SPSS 中执行不同的操作, 点击它们允许某些操作执行或不执行。在对话框打开时就已经勾了的框为默认设置。

让我们把每个变量移到 Variables 框, 对它们执行 Case Summaries 方法。

2. 为将变量移动到 Variables 框, 选择 **gender**, 按着鼠标键把光标 (屏幕上的箭头) 拖到变量 **age**, **employment** 和 **iq** 上。所有四个变量都应该被选上。点击向右箭头 (►) 将变量移动到 Variables 框。⁸ 详见图表 1—19。

3. 点击 “OK”。

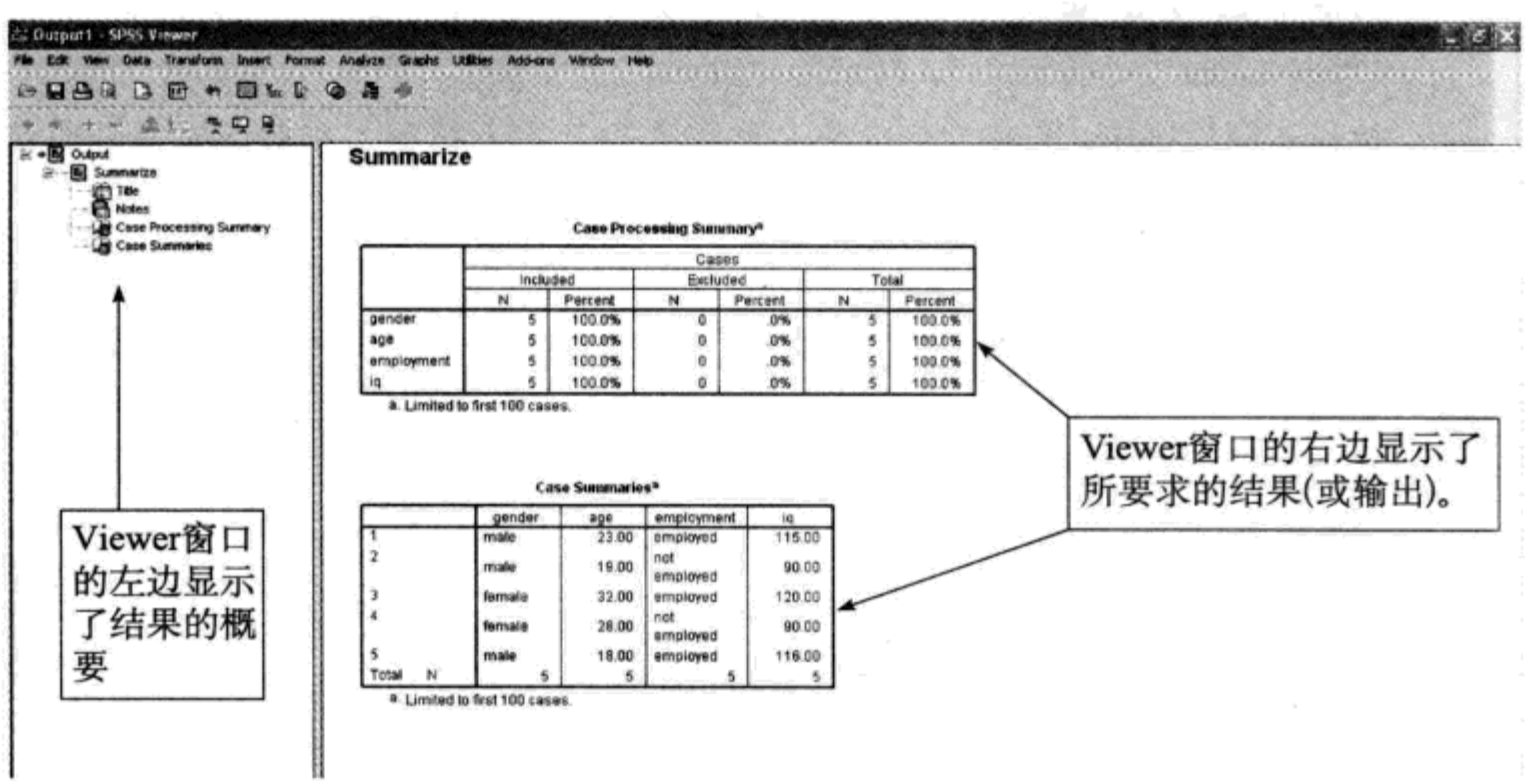


图表 1—19 Summarize Cases 对话框 (续)

SPSS 打开一个新的包含输出的窗口 (称为 Viewer 窗口), 它显示了 Case Summaries 方法的结果。下面将讨论这个输出。

步骤 4: 解释结果

Case Summaries 方法的输出显示在图表 1—20 中。



图表 1—20 SPSS Viewer 窗口及展示的 Case Summaries 方法的结果

Viewer (输出) 窗口

在 SPSS 中，Viewer 窗口分成两个不同的部分，窗口的左边包含了所要求的分析的概要，窗口的右边展示了分析结果。在本书中，我们将把注意力集中在 Viewer 窗口右边所显示的结果上。在这个例子中，它包含了 Case Processing Summary 表和 Case Summaries 表。下面将讨论这些表。

Case Processing Summary 表和 Case Summaries 表展示在图表 1—21 中。

概括

概括 (Summarize)

在输出的顶部是标题 “Summarize”，它表示了我们在 SPSS 中选择的方法 (Case Summaries 方法)。

案例处理摘要 (Case Processing Summary)

结果的第一个表 Case Processing Summary 展示了数据文件中的每个变量的参加者 (如，样本点) 数目。第一列 Included 对每个变量报告了 N 的值为 5 (N 对应于参加者或样本点的数目)，说明对每个变量都有 5 个人包含在分

Case Processing Summary ^a						
	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
gender	5	100.0%	0	.0%	5	100.0%
age	5	100.0%	0	.0%	5	100.0%
employment	5	100.0%	0	.0%	5	100.0%
iq	5	100.0%	0	.0%	5	100.0%

^a. Limited to first 100 cases.

注意标签 (male/female/
employed/not employed)是为
gender和employment展示的
(而不是那些1和2), 这是在
SPSS中生成变量值标签的好
处之一。

Case Summaries ^a				
	gender	age	employment	iq
1	male	23.00	employed	115.00
2	male	19.00	not employed	90.00
3	female	32.00	employed	120.00
4	female	28.00	not employed	90.00
5	male	18.00	employed	116.00
Total N	5	5	5	5

^a. Limited to first 100 cases.

图表 1—21 Case Summaries 方法在 SPSS 中的输出结果

析之中。在 Excluded 列对每个变量报告了 N 的值为 0，表示没有参加者被排除 (exclude) 在分析之外 (每个人都有所有四个变量的值)。最后一列 Total 展示了数据中参加者的总数，它等于 5。

案例摘要 (Case Summaries)

第二个表 Case Summaries 展示了参加者在四个感兴趣变量上的取值。注意，我们为 gender 和 employment 所生成的变量值标签展示在表中，而不是原先输入 SPSS 的数值 (1 和 2)，这便于阅读结果。

这就结束了 SPSS 中的 Case Summaries 方法的讨论。下面将讨论如何把你在 SPSS 中的工作保存起来。

保存文件

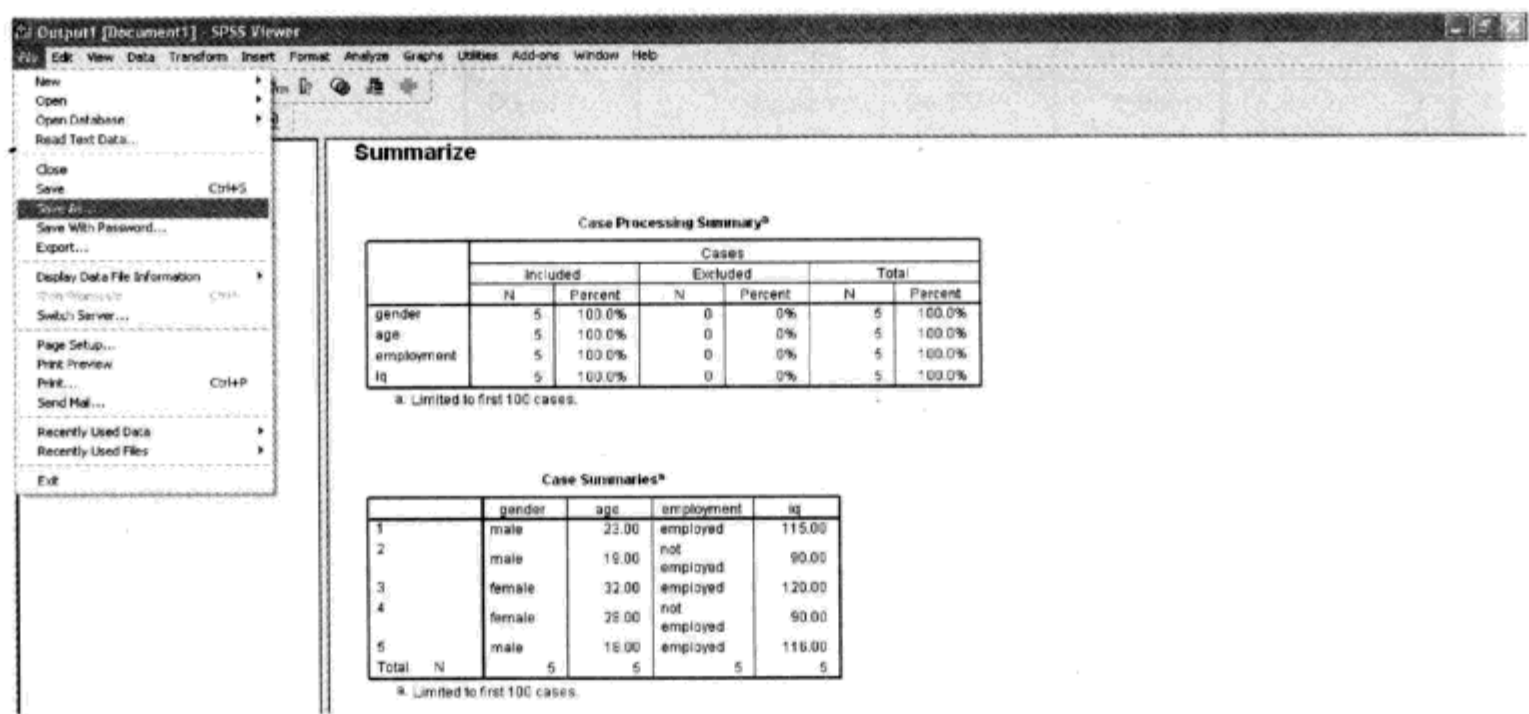
在 SPSS 中，输出文件和数据文件是用不同的扩展名分别保存的。文件扩

展名通常在文件名后包括三四个字母，用句点分开（例如 “.doc”，“.html”，“.mp3”）。在 SPSS 中，数据文件的扩展名为 “.sav”，输出文件的扩展名则为 “.spo”。

我们将在 SPSS 中实践保存数据文件和输出文件。从保存输出文件开始。为了保存输出文件，首先要确认 Viewer 窗口是激活的（如果 Viewer 窗口是激活的，你应该能在屏幕上看见它，有一个深蓝色的条在其窗口的上面）。如果 Viewer 窗口没有激活，点击它。

为保存输出文件：

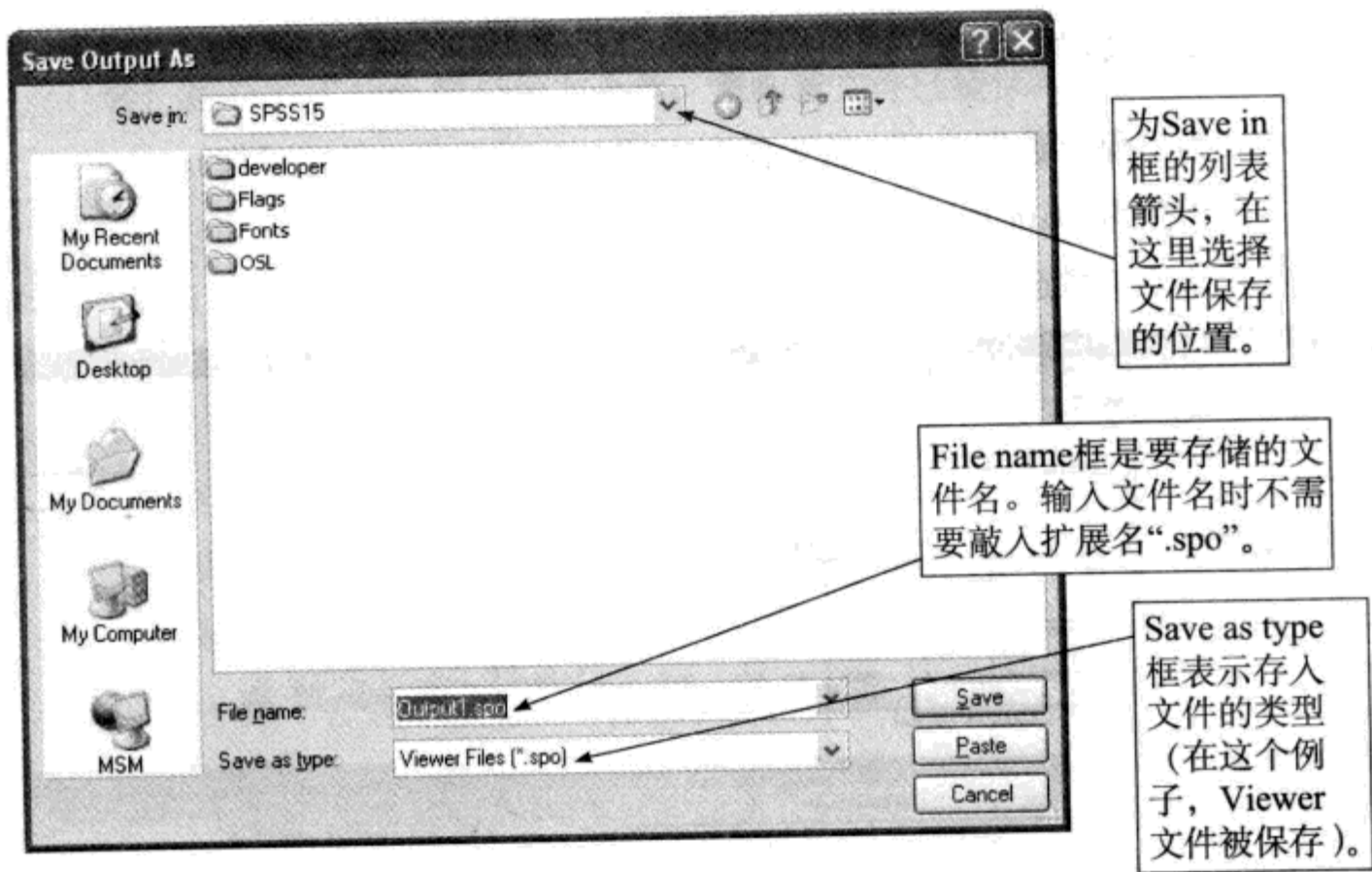
- 1. 当 Viewer 窗口是激活的，选择 **File>Save As...**（见图表 1—22）。
- 2. Save Output As 对话框打开（见图表 1—23）。



图表 1—22 保存文件的菜单命令

在 Save Output As 对话框底部的是 File name（文件名）和 Save as type（保存类型）框（见图表 1—23）。在 File name 框中，显示默认名字 Output1.spo，而在 Save as type 框中显示 Viewer 文件 (*.spo) 以确认 Viewer 窗口的结果（输出）要被保存。注意，Output1.spo 的存在意味着文件名字还处于编辑模式，并将被键盘输入的名字所代替（如果你没有在计算机上看到扩展名 “.spo”，说明你的计算机没有配置为显示文件扩展名，用不着担心。如果是这样，请遵从说明来保存文件，忽略所有关于扩展名的说明）。

让我们把文件命名为 Introduction output.



图表 1—23 Save Output As 对话框

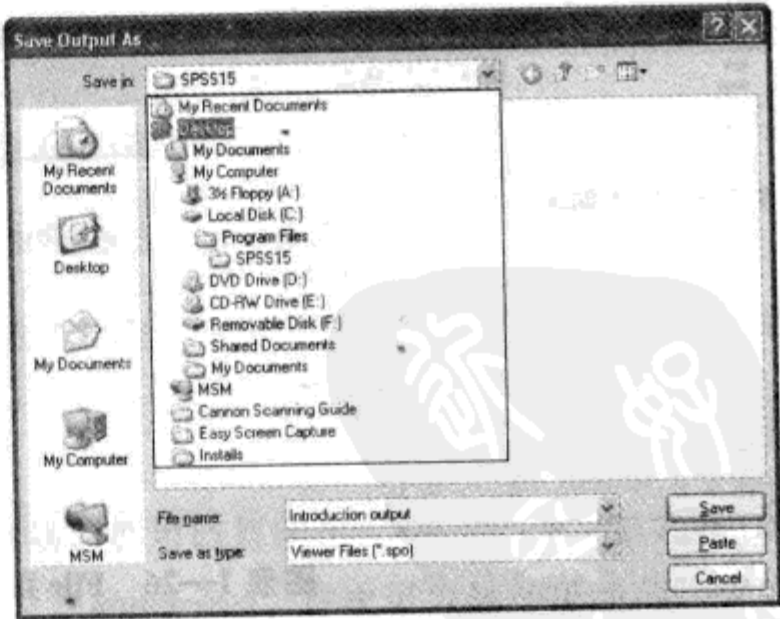
3. 输入名字 Introduction output（你现在应该在 File name 文本框中看到 Introduction output）。

下面将选择一个保存文件的位置。让我们把文件保存在桌面。
为保存文件到桌面：

4. 点击 Save in 框右边的列表箭头（▼），一个文件夹表和/或硬盘将显示。

5. 选择 “Desktop”（桌面）（见图表 1—24）（注意：如果你想要将该文件保存到其他地方，选择你所要存的位置）。

6. 点击 “Save”，你将看到名字 Introduction output.spo（无论是否有文件扩展名）在 Viewer 窗口的左侧，意味着文件已经保存。

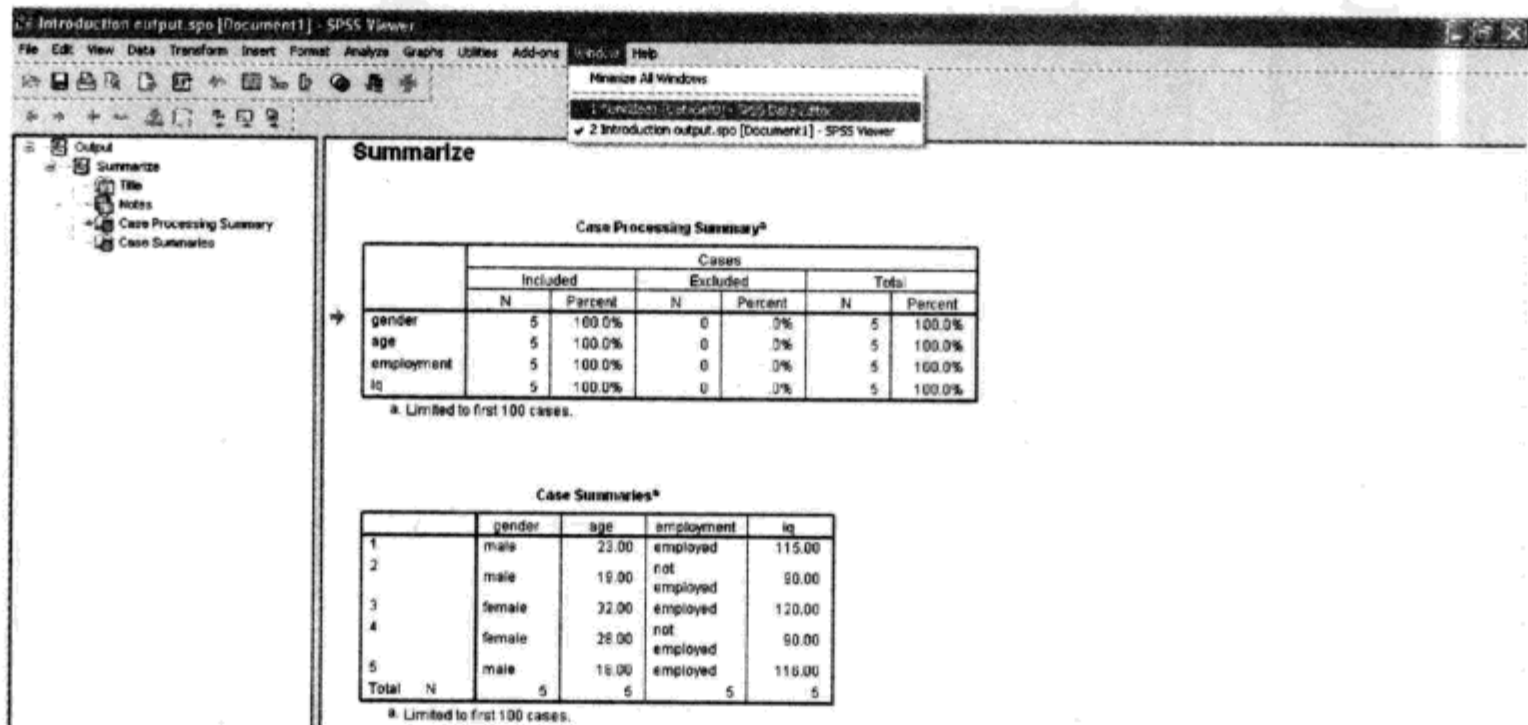


图表 1—24 Save Output As 对话框（续）

保存数据文件

为了保存数据文件，首先需要激活 Data Editor（数据编辑）。
为激活 Data Editor：

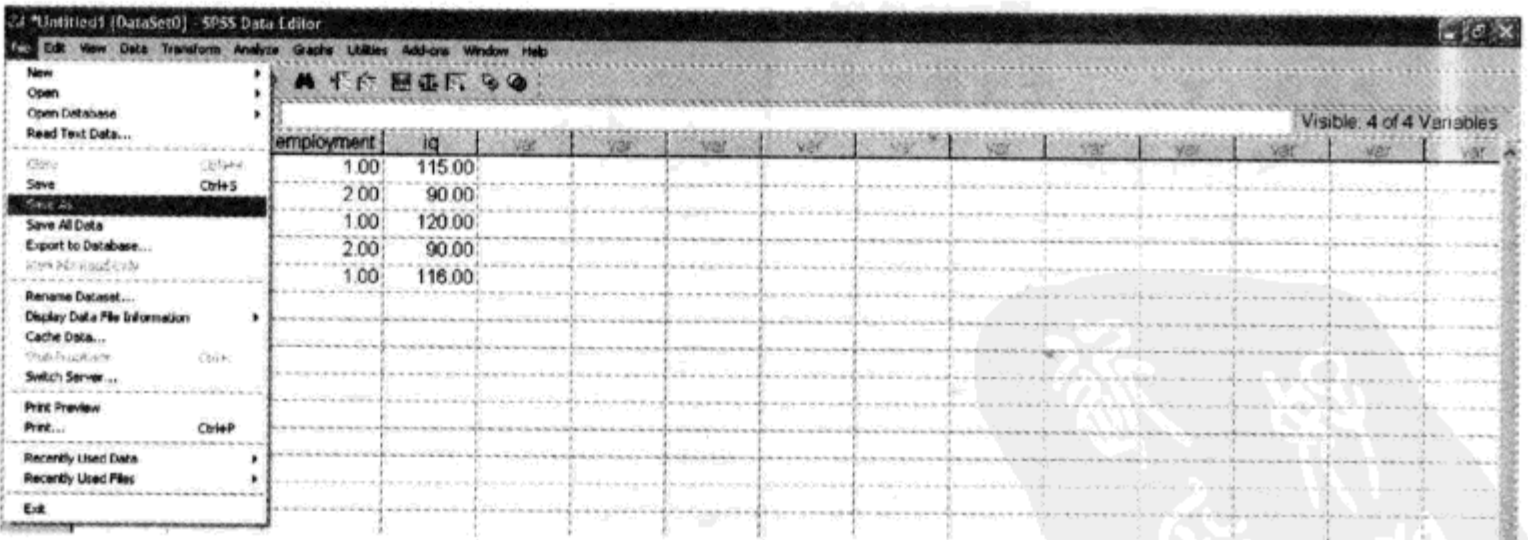
从菜单栏选择 **Window>Untitled-SPSS Data Editor**（见图表 1—25）。



图表 1—25 使 Data Editor 被激活的菜单命令

为保存数据文件：

- 1. 在 Data Editor 激活时，选择 **File>Save As...**（见图表 1—26）。



图表 1—26 File 菜单的 Save as 命令

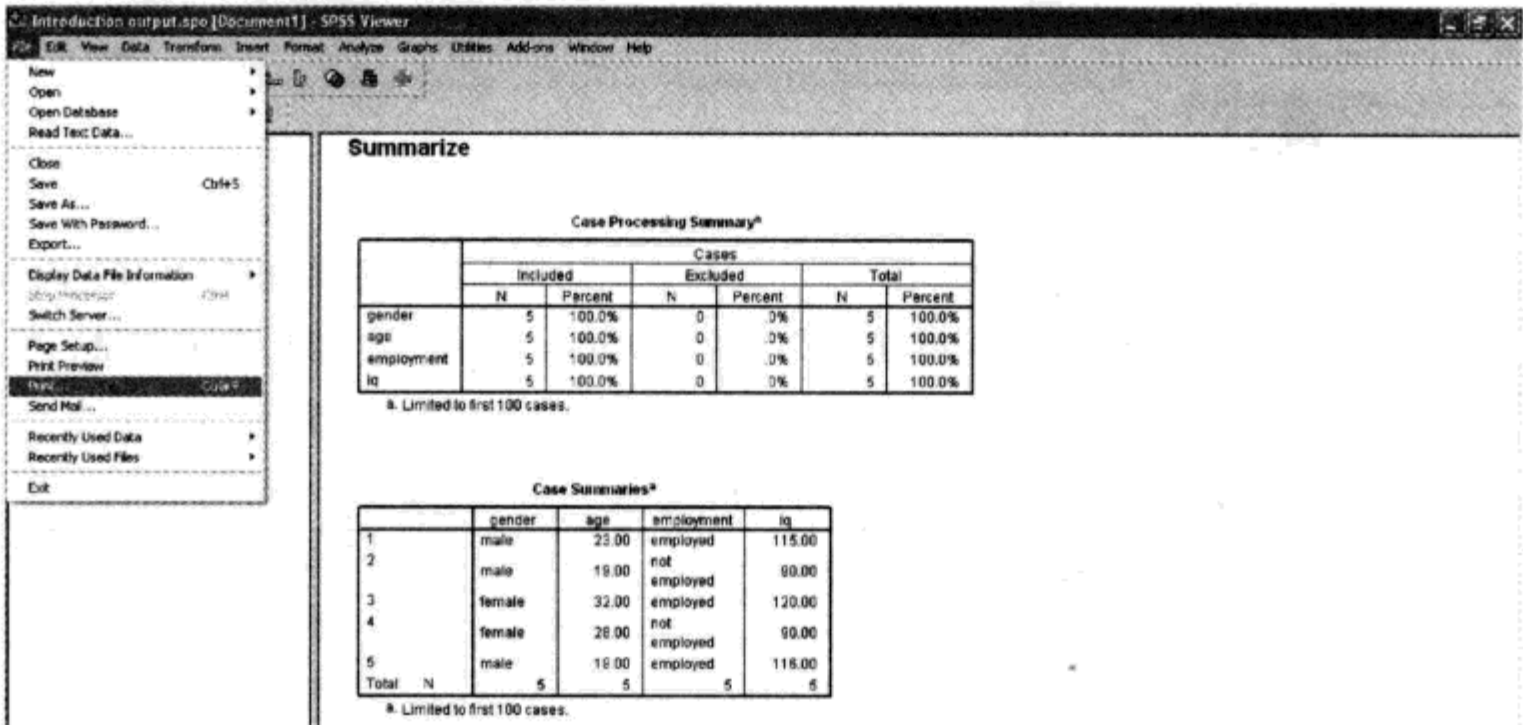
- 2. Save Data As 对话框打开。
- 3. 在 File name 框，输入名字 Introduction Data。
- 4. 点击 Save in 框右边的列表箭头（▼）以选择希望文件保存的位置。

5. 点击“Save”，你应看到名字 Introduction data.sav（无论是否有扩展名）在 Data Editor 窗口的左上角，意味着文件已经保存。

打印文件

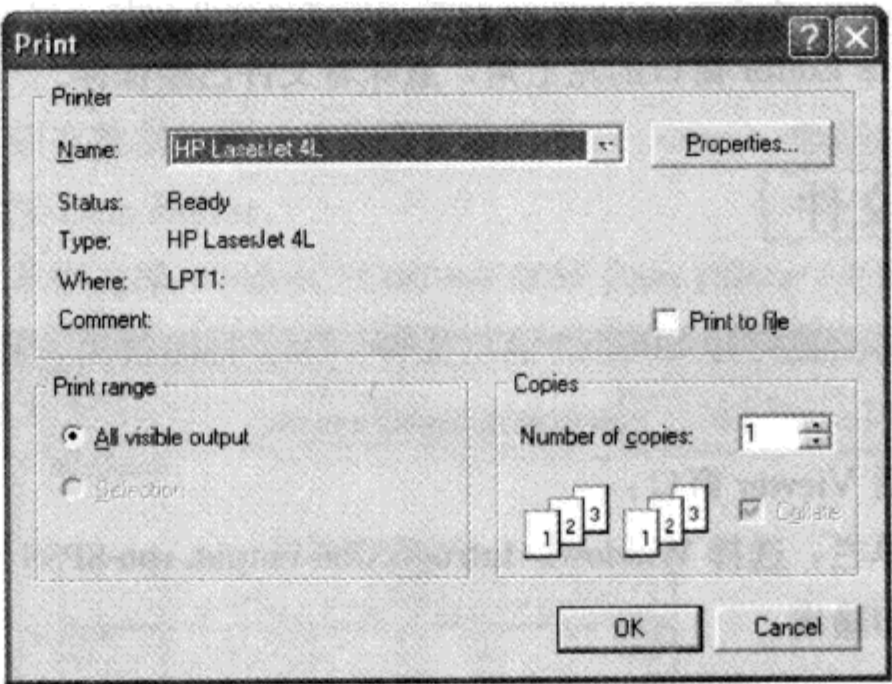
下面将描述打印 SPSS 的输出文件。为了打印输出，我们首先需要激活 Viewer 窗口。

- 为激活 Viewer 窗口：
- 从菜单栏，选择 **Window>Introduction output. spo-SPSS Viewer**。
- 为打印输出：
1. 在 Viewer 窗口激活状态下，选择 **File>Print...**（见图表 1—27）。



图表 1—27 打印 SPSS 输出的菜单命令

2. Print 对话框打开（详见图表 1—28）（你的计算机屏幕上的 Print 对话框可能与图表 1—28 显示的不同）。
 3. 点击“OK”。
- 假设你有一台打印机连接着电脑，就可以打印出 Case Processing Summary 表和 Case Summaries 表（标题为 Summarize）。你也可以选择单独的一张表（用鼠标点击它），然后从 **File** 菜单选择 **Print** 进行打印（激活 Data Editor 窗口（选择 Data View），从 **File** 菜单选择 **Print** 就可以打印数据文件）。
- 到这里，对 SPSS 的概述就结束了。



图表 1—28 SPSS 中的 Print 对话框

练习

1. 在图表 1—29 中显示了 7 个住疗养院的老人的数据，包括他们的年龄 (age)、性别 (gender)、健康状况 (well-being，用 1~10 的尺度来度量，高分意味着较健康)、每周的活动量 (number of activities)。

图表 1—29 疗养院数据

Person	Age	Gender	Well-being	Number of Activities
1	86	Male	4	2
2	72	Female	7	6
3	59	Female	6	5
4	86	Female	8	7
5	92	Female	4	1
6	68	Male	2	3
7	73	Male	8	5

注：person 变量包含在数据中，但不用输入 SPSS。

把数据输入 SPSS，并执行必要的分析来回答下面的问题。把变量分别命名为 age, gender, wellbeing 和 activities。

- a. 为 gender 生成变量值标签，指定男性值为 1，标签为 male，女性值为 2，标签为 female。

- b. 把文件保存到你选择的位置, 命名为 Nursing home data。
- c. 对数据运行 Case Summaries 方法并打印结果。
- d. 在 Case Summaries 表中, 在 SPSS 输出中是数值 (1 和 2) 还是变量值标签 (male 和 female)? 为什么?

2. 一个研究人员想要比较对儿童自尊心和控制愤怒能力的两种不同类型的疗法。被研究的疗法为非指导式游戏治疗 (儿童自主游戏) 和指导式游戏治疗 (治疗员指导儿童进行计划的游戏活动)。研究对象总共包括 12 名儿童, 每种疗法用于 6 名儿童 (3 个男孩, 3 个女孩)。在 6 周治疗后, 对儿童的自尊心 (用 10~50 的尺度来度量) 及控制愤怒能力 (用 5~25 的尺度来度量) 进行了评估。数据位于网站 www.prenhall.com/yockey 上的文件夹 Chapter 1 中的文件 Chapter 1 _ Exercise 2. sav 之中 (变量名字为 **therapy**, **gender**, **self-esteem**, **angermanage**)。打开文件并用 SPSS 进行必要的分析来回答下面的问题。

- a. 为 **therapy** 和 **gender** 生成变量值标签。对 **therapy**, 1 = “nondirective”, 2 = “directive”; 对 **gender**, 1 = “male”, 2 = “female”。

- b. 对数据运行 Case Summaries 方法并打印结果。

3. 该网站上的 Chapter 1 _ Exercise 3. sav 数据文件包含了 10 名学生的下面三个变量值: 性别 (**gender**), 注册的课程数目 (**numberclasses**), 每周工作小时数 (**hoursworked**)。在 SPSS 中打开该文件并实施必要的分析来回答下面的问题。

- a. 为 **gender** 生成变量值标签。对于 **gender**, 1 = “male”, 2 = “female”。
- b. 对数据运行 Case Summaries 方法并打印结果。

描述统计：频数、集中趋势的度量 以及变异性的度量

本章我们将用 SPSS 计算一系列的描述统计量，包括频数、集中趋势的度量（比如均值、中位数和众数）以及变异性的度量（比如标准差和方差）。下面将一一介绍各个统计量。

频数 (Frequencies) 是指变量在每一类中的观测次数。例如，在一个数据集中，变量性别 (**gender**) 的观测值为有 4 位男性和 6 位女性。男性的频数就是 4，女性的频数就是 6。经常在分类变量里可以得到频数。

集中趋势的度量用来描述得分集合的中心，包括均值、中位数、众数。**均值 (mean)** 就是算术平均数（得分的总和除以得分的总次数）。**中位数 (median)** 就是位于中间的得分（假设得分从最低到最高进行了排序）。**众数 (mode)** 是在数据集中出现最多次的得分。

变异性的度量用来描述分集合的离散性或变异性。**标准差 (standard deviation)** 和 **方差 (variance)** 是两个最常用的度量变异性的方法。标准差是度量得分平均偏离均值有多大的方法，方差等于标准差的平方。度量变异性的其他方法有极差 (**range**，最大值和最小值的差) 和四分位数间距 (**interquartile range**，分布的 75% 分位数和 25% 分位数的差)。

我们将使用图表 2—1 中给出的数据说明在 SPSS 中如何计算频数、集中趋势的度量以及变异性的度量。

数据

图表 2—1 样本数据集

Participant	Gender	Mathexam	Collage	Satquant
1	1	24	1	570
2	1	18	2	450
3	2	34	1	600
4	2	27	2	450
5	1	15	1	580
6	1	26	1	550
7	1	42	2	480
8	2	25	1	520
9	2	31	1	450
10	2	44	2	550

说明：变量 Participant 包含在数据中，但不用输入 SPSS。

在图表 2—1 中记录了每一个参与者的性别（gender）、数学入学考试分数（mathexam）、在家乡或外地上大学（college）、SAT 中的数学分数排名（satquant）。Gender 和 College 是分类变量，Mathexam 和 Satquant 是连续变量，得分越高表明他们在这两个方面表现越好（Mathexam 得分范围是 0~50，Satquant 得分范围是 200~800）。对于性别（gender），男性记为“1”，女性记为“2”；对于大学所在地（college），在本地读大学的记为“1”，在外地读大学的记为“2”。

概述了上面提供的数据集之后，现在我们开始将数据输入 SPSS。

下面的步骤 1 和步骤 2 描述了在 SPSS 中怎样输入数据。数据文件在网址 www.prenhall.com/yockey 上的文件夹 Chapter 2 中的文件 descriptive statisticsl.sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中将生成四个变量，四个变量分别命名为 gender，mathexam，college 和 satquant。

3. 使用第1章描述的方法，在 Variable View 窗口前四行分别输入变量名 **gender**，**mathexam**，**college** 和 **satquant**（见图表 2—2）。

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	gender	Numeric	8	2		None	None	8	Right	Scale
2	mathexam	Numeric	8	2		None	None	8	Right	Scale
3	college	Numeric	8	2		None	None	8	Right	Scale
4	satquant	Numeric	8	2		None	None	8	Right	Scale

图表 2—2 输入变量 **gender**，**mathexam**，**college** 和 **satquant** 的 Variable View 窗口

在输入数据之前，我们将在 SPSS 中对分类变量 **gender** 和 **college** 创建变量值标签。

4. 用第1章描述的方法给 **gender** 和 **college** 创建变量值标签。对于 **gender**，1=“男性”，2=“女性”。对于 **college**，1=“在本地上大学”，2=“在外地上大学”。

步骤 2：输入数据

接下来，我们将把数据输入 SPSS。

为输入数据：

1. 点击 Data View 标签。变量 **gender**，**mathexam**，**college** 和 **satquant** 出现在 Data View 窗口的前四列。

2. 参照图表 2—1，为每个参与者输入感兴趣的四个变量的数据。对第一个参与者，分别在变量 **gender**，**mathexam**，**college** 和 **satquant** 下输入 1，24，1 和 570。依次输入全部 10 个参与者的数据。图表 2—3 中给出了完整的数据文件。

	gender	mathexam	college	satquant
1	1.00	24.00	1.00	570.00
2	1.00	18.00	2.00	450.00
3	2.00	34.00	1.00	600.00
4	2.00	27.00	2.00	450.00
5	1.00	15.00	1.00	580.00
6	1.00	26.00	1.00	550.00
7	1.00	42.00	2.00	480.00
8	2.00	25.00	1.00	520.00
9	2.00	31.00	1.00	450.00
10	2.00	44.00	2.00	550.00

图表 2—3 10 名参与者的完整数据文件

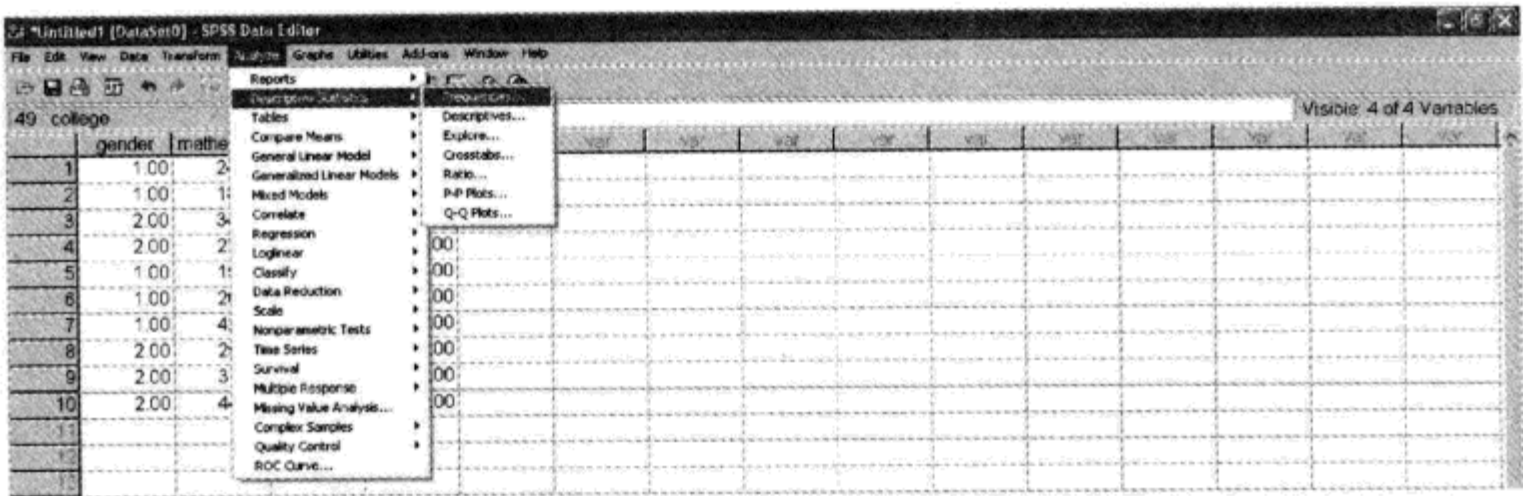
步骤 3：分析数据

在分析数据时，我们首先计算分类变量 **gender** 和 **college** 的频数，然后计算连续变量 **mathexam** 和 **satquant** 的集中趋势和变异性的度量。

频数

为了计算 **gender** 和 **college** 的频数，在菜单栏中选择：

- 1. **Analyze>Descriptive Statistics> Frequencies...**（见图表 2—4）。



图表 2—4 频数程序的菜单命令

打开 **Frequencies** 对话框，变量 **gender**，**mathexam**，**college** 和 **satquant** 出现在对话框的左边（见图表 2—5）。

- 2. 选择变量 **gender** 并按住 **Ctrl** 键（**Ctrl** 键位于键盘的左下角）。再选择变量 **college**（**gender** 和 **college** 现在被选中了），点击向右箭头按钮（▶），把两个变量移到 **Variable (s)** 框中。¹ 详见图表 2—6。

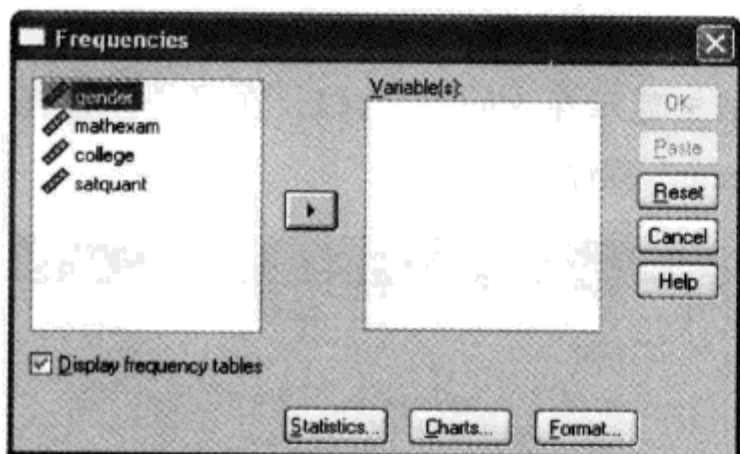
- 3. 点击 “OK”。

运行频数程序，运行结果显示在 **Viewer** 窗口。在讨论 **gender** 和 **college** 的频数程序的运行结果之前，我们将会得到关于 **mathexam** 和 **satquant** 的描述统计量。

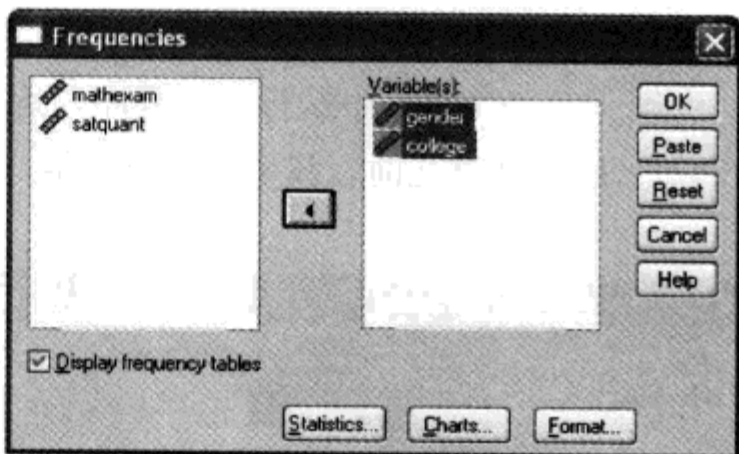
为保持全书的一致性，我们在 **Data Editor** 窗口运行程序。由于目前 **Viewer** 窗口是激活的，我们将会在运行下一个程序之前激活 **Data Editor** 窗口。

为激活 **Data Editor** 窗口：

在菜单栏中选择 **Window>Untitled-SPSS Data Editor**。现在 **Data Editor** 窗



图表 2—5 频数对话框



图表 2—6 频数对话框（续一）

口应该被激活了。（如果你是从网站上打开的数据，要在菜单栏中选择 **Window>descriptive statistics. sav-SPSS Data Editor**。如果直接从网站上打开数据，那么这些说明将贯穿本章。）

度量集中趋势与变异性

接下来，我们将用频数程序获得连续变量 **mathexam** 和 **satquant** 的集中趋势与变异性的度量结果。

为获得集中趋势与变异性的度量：

1. 从菜单栏中选择 **Analyze>Descriptive Statistics> Frequencies...**（见图表 2—4）。²

2. 打开 **Frequencies** 对话框，点击“Reset”（复位）按钮清除 Variable(s) 框中现有的变量。选择变量 **mathexam** 并按住 Ctrl 键。再选择变量 **satquant**（**mathexam** 和 **satquant** 现在被选中了），点击向右箭头按钮（▶），把两个变量移到 Variable(s) 框中（详见图表 2—7）。

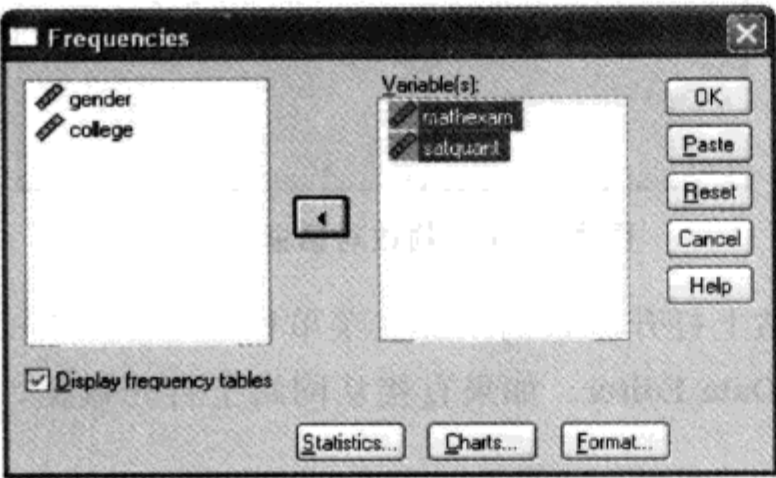
3. 点击“Statistics”按钮，打开 Frequencies: Statistics 窗口，在 Central Tendency 下选择 Mean, Median 和 Mode。在 Dispersion 下选择 Std. deviation, Variance, Range, Minimum 和 Maximum。详见图表 2—8。

4. 点击“Continue”。

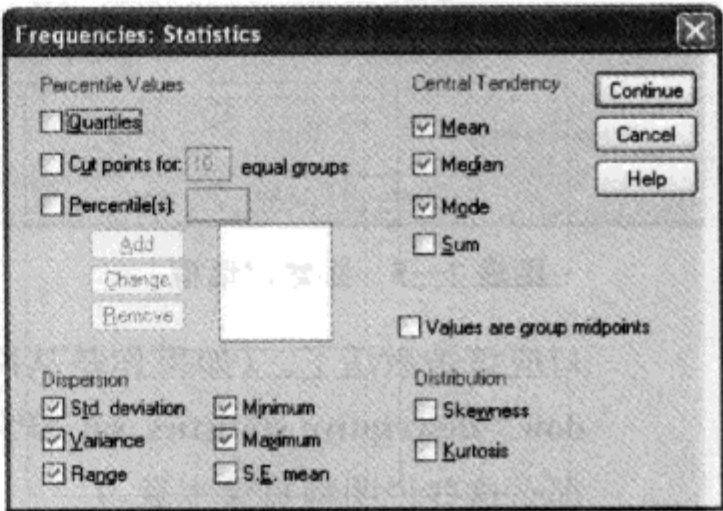
5. 打开 **Frequencies** 对话框，单击 Display frequency tables 左边的复选标记，关闭选项。现在方框内是空的，表明将不会生成变量 **mathexam** 和 **satquant** 的频数表（通常，连续变量不需要频数表，因为集中趋势和变异性度量的信息已经很充分了）。详见图表 2—9。

6. 点击“OK”。

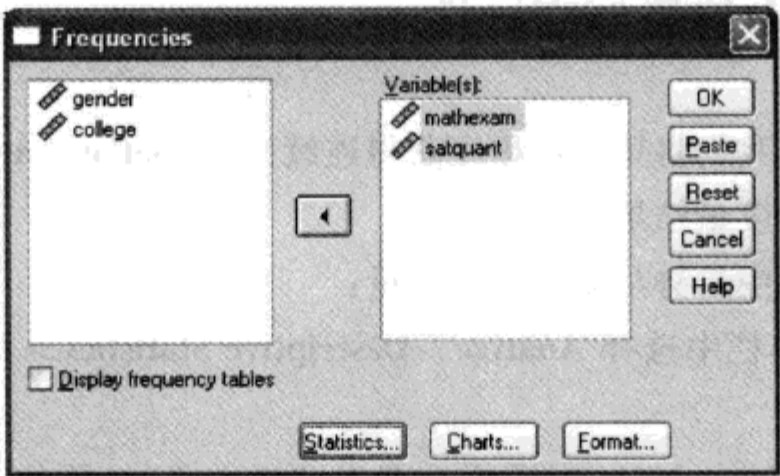
在 SPSS 中运行频数程序，运行结果显示在 Viewer 窗口。在讨论频数程序的结果之前，我们先介绍均值（Means）程序。



图表 2—7 频数对话框（续二）



图表 2—8 Frequencies: Statistics 对话框



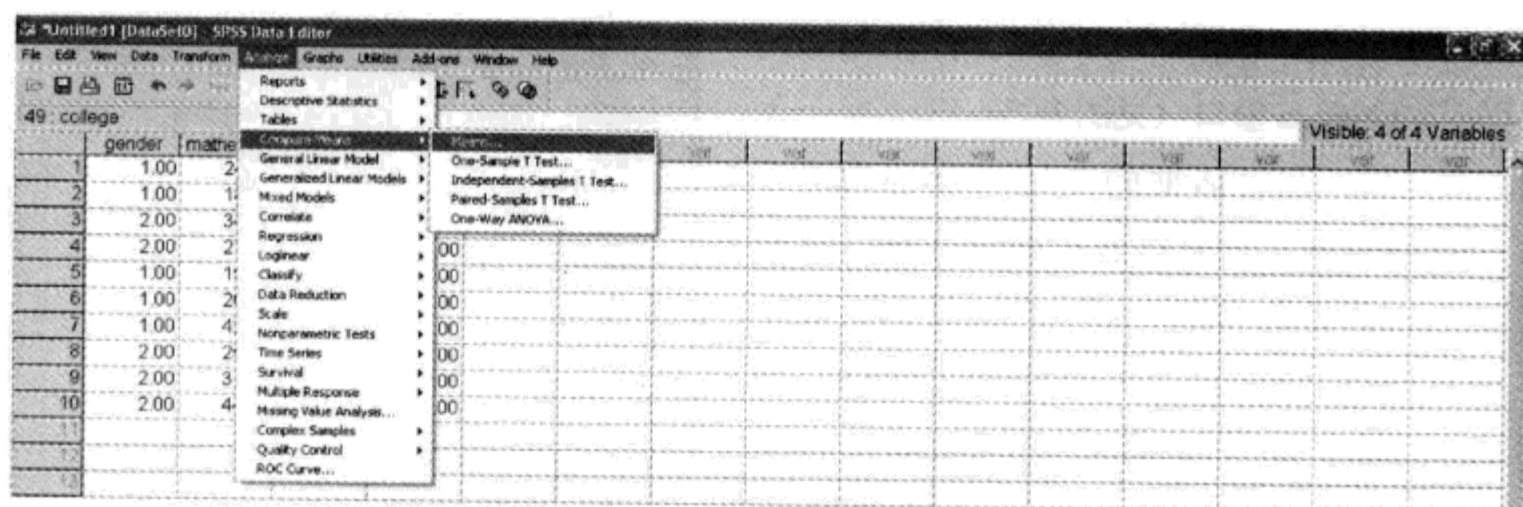
图表 2—9 频数对话框（续三）
(Display frequency tables 检查框没有选中)

使用 Means 程序进行的组分析

这个 Means 程序是用来获得感兴趣的一个变量（或多个变量）的各组单独的描述统计量。比如：不是得到所有参与者的 **satquant** 的均值，而是分别得到在本地上大学 and 在外地上大学的这些人的 **satquant** 的均值。我们将通过对不同 college 组的人分别获得 **satquant** 的均值和标准差来介绍 Means 程序。

为得到 Means 程序：

1. 激活 Data Viewer 窗口（选择 **Window>Untitled-SPSS Data Editor**）。
2. 从菜单栏中选择 **Analyze>Compare Means>Means...**（详见图表 2—10）。



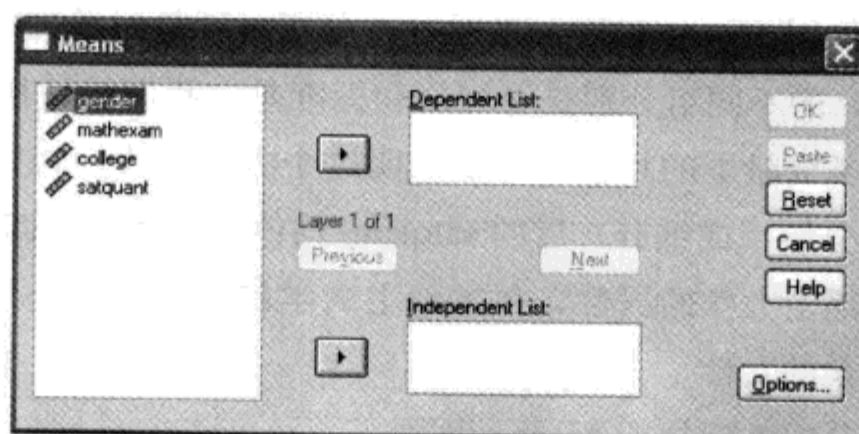
图表 2—10 Means 程序的菜单命令

3. 打开 Means 对话框（见图表 2—11）。

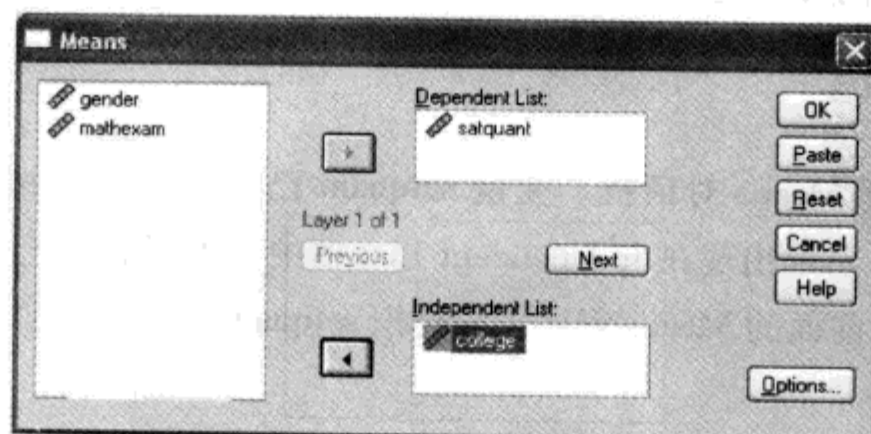
把希望得到均值和标准差的变量（**satquant**）移到 Dependent List 框中，把包含了不同组别的变量（**college**）移到 Independent List 框中。

4. 选择 **satquant**，点击上面的向右箭头按钮（▶），把变量移到 Dependent List 框中。

5. 选择 **college**，点击下面的向右箭头按钮（▶），把变量移到 Independent List 框中（见图表 2—12）。



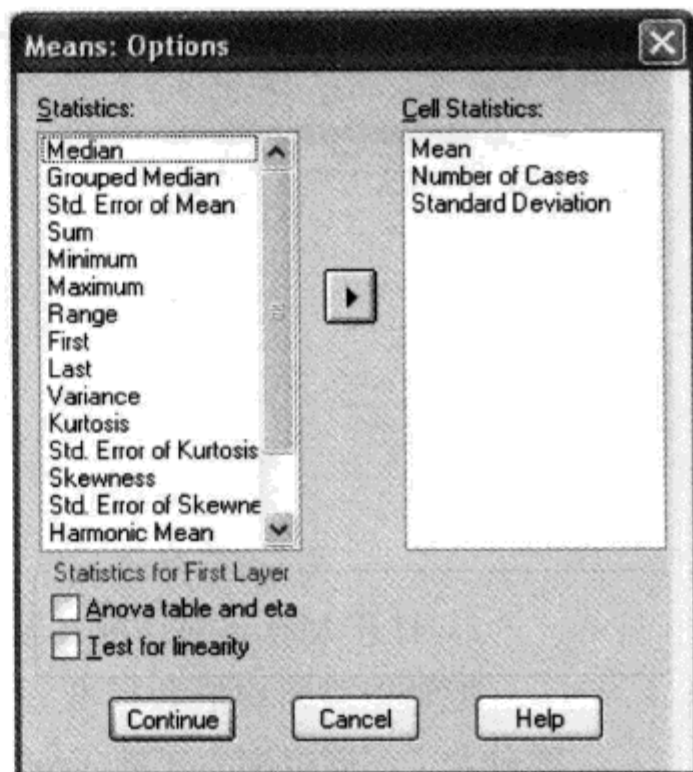
图表 2—11 Means 对话框



图表 2—12 Means 对话框（续）

6. 单击“Options”。打开 Means: Options 窗口（见图表 2—13）。

在对话框的左边（标记为 Statistics）是可以选择的统计量，右边（标记为 Cell Statistics）是默认要计算的统计量。因为均值（Mean）和标准差（Standard Deviation）已经在右边了，这里我们保留默认项（Number of Cases 可以移到 Statistics 框中，选择它，然后点击箭头按钮，但是我们把它保留在 Cell Statistics 框中，因为它不影响我们的结果）。



图表 2—13 Means: Option 对话框

7. 单击“Continue”。

8. 单击“OK”。

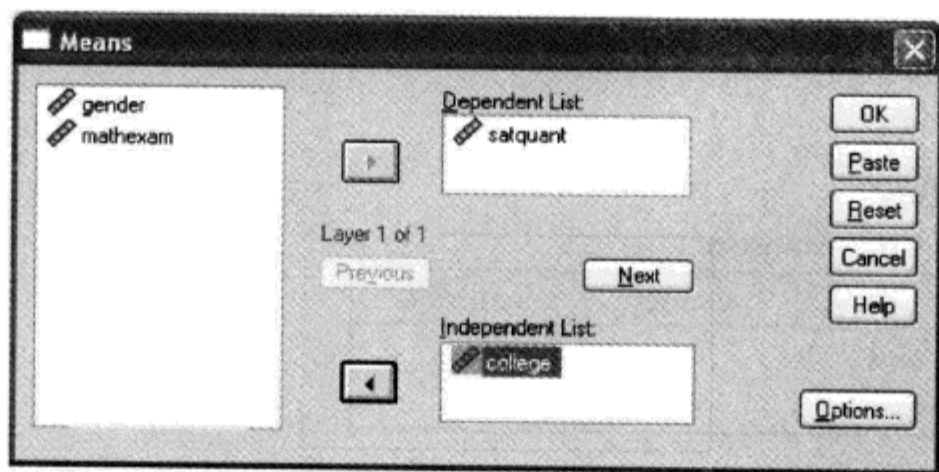
在 SPSS 中运行 Means 程序，结果显示在 Viewer 窗口中。在讨论结果之前，我们将使用 Means 程序做最后一个分析。不是对单独一个变量（这里使用了 **college**）的不同组分别获得均值和标准差，我们将使用 **layered** 选项获得对于两个分类变量下的均值。为了说明这个方法，我们将得到 **college** 的不同组被 **gender** 分层（或细分）后的 **satquant** 均值。这将得到四个不同的 **satquant** 均值：在本地上大学的男性，在本地上大学的女性，在外地上大学的男性和在外地上大学的女性。

为使用分层选项执行 Means 程序：

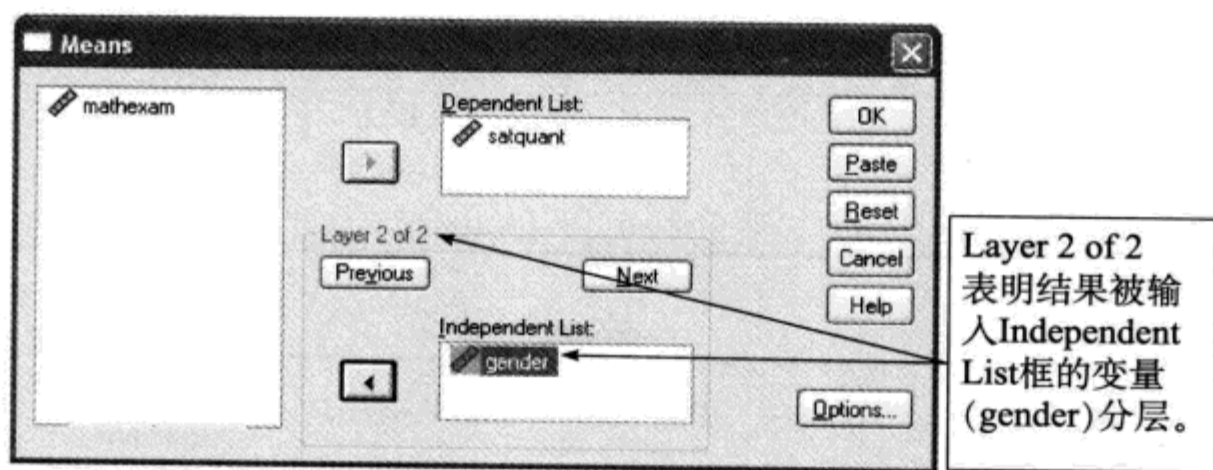
1. 激活 Data Viewer 窗口（选择 **Window>Untitled-SPSS Data Editor**）。
2. 从菜单栏中选择 **Analyze>Compare Means>Means...**（详见图表 2—10）。
3. 打开 Means 对话框，变量 **satquant** 已经出现在 Dependent List 框中，变量 **college** 已经出现在 Independent List 框中（见图表 2—14）。（注意：如果你没有执行前面的 Means 程序，需要把 **satquant** 和 **gender**^① 移到相应的框中。）

① 此处似有误，应为 **college**。——译者注

4. 为了对结果用 **gender** 分层，在对话框的 Layer 1 of 1 部分，点击“Next”（再点击“Next”就会出现“Layer 2 of 2”），选择 **gender**，点击下面的向右箭头按钮（▶），把变量移到 Independent List 框中（见图表 2—15）。
5. 点击“OK”。



图表 2—14 Means 对话框



图表 2—15 使用分层选项的 Means 对话框

在 SPSS 中运行 Means 程序，结果显示在 Viewer 窗口中。我们将讨论本章做的四个分析的结果。这四个分析概述如下：

1. 变量 **college** 和 **gender** 的频数。
2. 变量 **satquant** 和 **mathexam** 的集中趋势和变异性的度量。
3. **satquant** 在不同 **college** 组中的均值和标准差。
4. 不同的 **college** 组被 **gender** 分层后的 **satquant** 的均值和标准差。

步骤 4：结果解释

这四个分析的结果见图表 2—16 至图表 2—19。（如果你在阅读本章时，

在自己的电脑上进行了这些分析，滚动到第一组结果的 Viewer（输出）窗口的最上端就是 college 和 gender 的频数。）

分析 1：Frequencies 程序——Gender 和 College

我们做的第一个分析的结果——变量 college 和 gender 的频数，见图表 2—16。

Frequencies

		Statistics	
		gender	college
N	Valid	10	10
	Missing	0	0

Frequency Table

		gender			
		Frequency	Percent	Valid Percent	Comulative Percent
Valid	male	5	50.0	50.0	50.0
	female	5	50.0	50.0	100.0
	Total	10	100.0	100.0	

		college			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	home	6	60.0	60.0	60.0
	away	4	40.0	40.0	100.0
	Total	10	100.0	100.0	

图表 2—16 对 college 和 gender 用 Frequencies 程序得到的输出结果

统计量 (Statistics)

Statistics 表显示了变量 college 和 gender 的样本量。该表表明变量 college 和 gender 都有 10 个有效观测值（所有的 10 个人在两个变量上都有值）。在 Missing 行的值为 0 表明任何一个变量都没有缺失值（例如，变量 gender 的缺失值将表明有一个参与者的性别没有被记录）。

频数表——Gender

Gender 的频数表 (Frequency Table) 表明 gender 的每一类的频数或者说

参与者的人数。在表中有 5 位男性和 5 位女性，每个性别在数据中各占 50%。

频数表——College

College 的频数表表明 college 的每一类的频数。该表说明 6 个学生在本地上大学，4 个学生在外地上大学，比例分别是 60%和 40%。

分析 2：Frequencies 程序——Mathexam 和 Satquant

我们做的第二个分析的结果——变量 mathexam 和 satquant 的频数，见图 2—17。

Frequencies

		Statistics	
		mathexam	satquant
N	Valid	10	10
	Missing	0	0
Mean		28.6000	520.0000
Median		26.5000	535.0000
Mode		15.00 ^a	450.00
Std. Deviation		9.38320	58.30952
Variance		88.044	3400.000
Range		29.00	150.00
Minimum		15.00	450.00
Maximum		44.00	600.00

^a Multiple modes exist. The smallest value is shown

图表 2—17 对 mathexam 和 satquant 用 Frequencies 程序得到的输出结果

Statistics 表显示了连续变量 mathexam 和 satquant 的频数程序的结果（回忆我们要求对这些变量的集中趋势和变异性的度量）。mathexam 的均值和标准差分别是 28.60 和 9.38（四舍五入，保留两位小数）。satquant 的均值和标准差分别为 520.00 和 58.31。minimum 和 maximum 分别表明数据中的最低值和最高值，有助于保证所有的数据在可能的范围内取值（超过这个范围的取值很有可能表明数据输入有错误）。我们确定数据中可能的排列变化。mathexam 的最小值和最大值分别是 15.00 和 44.00。satquant 的最小值和最大值是 450.00 和 600.00。说明两个变量都没有超出取值范围（正如前面表明的，mathexam 的可能得分范围是 0~50，satquant 的是 200~800）。³

分析 3: Means 程序——Satquant 被 College 分组

我们做的第三个分析的结果——对不同的 college 组的 satquant 均值，见图表 2—18。

Means

Case Processing Summary						
	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
satquant* college	10	100.0%	0	0.0%	10	100.0%

Report

satquant			
college	Mean	N	Std. Deviation
home	545.0000	6	53.94442
away	482.5000	4	47.16991
Total	520.0000	10	58.30952

图表 2—18 Means 程序关于 satquant 在不同 college 组（home 和 away）下的输出结果

案例处理摘要

Case Processing Summary 表表明了分析中的观测数量（10），已排除的观测数量（0）以及数据集中的全部观测数量（10）。该表说明所有 10 名参与者都包括在分析之中（所有 10 个人在 satquant 和 college 上都有取值）。

报告（Report）

Report 表显示了不同 College 组别的均值、样本量（N）和标准差，包括本地（home）、外地（away）和总计（Total）（总计等于本地与外地之和）。检查两组数据的均值说明在本地上大学的人的 satquant 均值（545.00）高于在外地上大学的人的 satquant 均值（482.50）。10 个学生中有 6 个人在本地上大学，4 个人在外地上大学（我们在前面第一个分析里的 college 的频数表中看到过这个结果）。那些在本地上大学的学生的 satquant 得分的波动性略高于在外地上大学的人（本地的标准差是 53.94，而外地的标准差是 47.17）。整个样本的 satquant 的均值是 520.00，标准差是 58.31（这与第二个分析中得到

的结果是一致的)。

分析 4：Means 程序——College 被 Gender 分层下的 Satquant

我们做的最后一个分析的结果——college 被 gender 分层下的 satquant 的均值，见图表 2—19。

Means

	Case Processing Summary					
	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
satquant* college* gender	10	100.0%	0	0.0%	10	100.0%

Report

satquant				
college	gender	Mean	N	Std. Deviation
home	male	566.6667	3	15.27525
	female	523.3333	3	75.05553
	Total	545.0000	6	53.94442
away	male	465.0000	2	21.21320
	female	500.0000	2	70.71068
	Total	482.5000	4	47.16991
Total	male	526.0000	5	57.70615
	female	514.0000	5	65.03845
	Total	520.0000	10	58.30952

图表 2—19 Means 程序得到 College 被 Gender 分层下的 Satquant 的输出结果

案例处理摘要

Case Processing Summary 表表明了分析中的观测数量（10），已排除的观测数量（0）以及数据集中的全部观测数量（10）。该表说明所有 10 名参与者都包括在分析之中（所有 10 个人在 satquant, college 和 gender 上都有取值）。

报告

Report 表显示了 College 被 Gender 分层下的 Satquant 的均值、样本量

(N) 和标准差。该表的第一部分显示了在本地上大学的男性和女性的均值和标准差。注意：在 satquant 上，男性的均值比女性高（566.67 对 523.33），女性有较大的标准差（75.06 对 15.28）。该表的第二部分显示了在外地上大学的男性和女性的均值和标准差。对于这一组学生，在 satquant 上，女性的均值比男性高（500.00 对 465.00），标准差也较大（70.71 对 21.21）。最后，该表的最后一部分是男性和女性在本地上大学 and 在外地上大学合并起来时的均值和标准差。当 college 组合并起来时，在 Satquant 上，男性的均值高于女性的均值（526.00 对 514.00），女性的标准差更大（65.04 对 57.71）。

至此结束了对 SPSS 中 Frequencies 和 Means 程序的讨论。

执行 Frequencies 和 Means 程序的步骤摘要

I. 数据输入

1. 通过在 Variable View 窗口输入变量名称而在 SPSS 中生成变量。
2. 给每一个分类变量建立变量值标签。在 Variable View 窗口输入合适的数值和标签。点击“OK”。
3. 在 Data View 窗口输入每一个参与者的数据。

II. 数据分析

使用 Frequencies 程序获得频数：

1. 从菜单栏中选择 **Analyze>Descriptive Statistics> Frequencies...**
2. 把要分析的分类变量移到 Variable (s) 框中。
3. 点击“OK”。

使用 Frequencies 程序获得描述统计量：

1. 从菜单栏中选择 **Analyze>Descriptive Statistics> Frequencies...**
2. 把要分析的分类变量移到 Variable (s) 框中。
3. 点击“Statistics”。
4. 选择感兴趣的统计量。
5. 点击“Continue”。
6. 在 Frequencies 对话框，不选择 Display frequency tables。
7. 点击“OK”。

使用 Means 程序获得关于一个或更多分类变量的描述统计量：

- 1. 从菜单栏中选择 **Analyze>Compare Means>Means...**。
- 2. 把要得到描述统计量的变量移到 **Dependent List** 框中，把分类变量移到 **Independent List** 框中。
- 3. 如果你想用第二个分类变量对结果分层，点击 “**Layered**” 按钮，把分类变量移到 **Independent List** 框中（如果不想对结果分层就忽略该步骤）。
- 4. 点击 “**OK**”。

练习

1. 一位足球教练想评价一场足球比赛当中两种不同踢法（脚尖踢和脚后跟踢）的效果。4 个男孩和 4 个女孩使用其中一种方法踢球，记录了他们踢球的距离（英尺）和准确度（范围为 1~10），也确定了小孩的经验水平（值越高表明经验越多）。一半的小孩（两个男孩和两个女孩）使用脚尖踢，一半的小孩使用脚后跟踢。图表 2—20 列出了数据。对于 **gender**，1=“男孩”，2=“女孩”。对于 **method**，1=“脚尖踢”（toe kick），2=“脚后跟踢”（heel kick）。

图表 2—20 8 个孩子在足球踢法调查中的数据

Gender	Method	Experience	Distance	Accuracy
1	1	5	50	5
1	1	6	45	4
1	2	2	35	6
1	2	9	40	9
2	1	2	42	3
2	1	7	39	6
2	2	4	28	7
2	2	2	32	6

在 SPSS 中输入数据，进行必要分析以回答下列问题（使用图表 2—20 中的变量名）。为 **gender** 和 **method** 建立合适的变量值标签。

- a. 报告 **gender** 和 **method** 的频数。
- b. 报告全部样本的 **experience**，**distance** 和 **accuracy** 的均值和标准差。
- c. 分别报告男孩和女孩的 **experience**，**distance** 和 **accuracy** 的均值和标准差。哪一个组在三个变量上有较高的均值？哪一个组有较大的标准差？

d. 分别报告脚尖踢和脚后跟踢的 **experience**, **distance** 和 **accuracy** 的均值和标准差。哪一个组在三个变量上有较高的均值? 哪一个组有较大的标准差?

e. 报告 **gender** 被 **method** 分层后的 **experience**, **distance** 和 **accuracy** 的均值和标准差。四种情况下, 哪一个组有最低的 **experience** 和 **accuracy** 的均值? 哪一个组有最高的平均踢球距离?

2. 一个学生收集了美国三个不同城市的水质量感知的数据。她从中西部、西海岸和东海岸的每个主要城市中抽取了 10 个人 (5 名男性和 5 名女性), 每个人品尝了他们各自城市的水 (水由研究者提供), 对水的味道和透明度进行了 1~10 的排序 (排名越高表明水的味道和透明度越好)。数据位于网站 www.prenhall.com/yockey 中文件夹 Chapter 2 中的文件 Chapter2 _ Exercise 2. sav 之中, 数据文件包括了 30 个参与者在 **location**, **gender**, **taste** 和 **clarity** 上的取值。对于 **location**, 1=“西海岸”, 2=“中西部”, 3=“东海岸”。对于 **gender**, 1=“男性”, 2=“女性”。在 SPSS 中打开文件, 进行必要分析以回答下列问题。

a. 报告 **location** 和 **gender** 的频数。

b. 报告全部样本的 **taste** 和 **clarity** 的均值和标准差。

c. 分别报告不同城市的 **taste** 和 **clarity** 的均值和标准差。哪一个城市在两个变量上有最高的均值? 哪一个城市有最大的标准差?

d. 分别报告男性和女性的 **taste** 和 **clarity** 的均值和标准差。哪一个组在两个变量上有最高的均值? 哪一个组有最大的标准差?

e. 报告 **gender** 被 **location** 分层后的 **taste** 和 **clarity** 的均值和标准差。六种情况下, 哪一个组有最低的 **taste** 的均值? 哪一个组有最低的 **clarity** 的均值? 哪一个组有最高的 **taste** 和 **clarity** 的均值?

3. 一个研究者收集了 30 个小孩每天的阅读能力和看电视的小时数的数据, 也记录了每个参与者的性别和社会经济状况 (SES)。数据在网站上文件夹 Chapter 2 中的文件 Chapter2 _ Exercise 3. sav 之中。数据文件包括了变量 **gender**, **ses**, **hourstv** 和 **readingscores**。对于 **gender**, 1=“男孩”, 2=“女孩”。对于 **ses**, 1=“低”, 2=“中”, 3=“高”。阅读得分范围为 0~50, 得分越高表明阅读能力越好。在 SPSS 中打开文件, 进行必要分析以回答下列问题。

a. 报告 **gender** 和 **ses** 的频数。

b. 报告全部样本的 **hourstv** 和 **readingscores** 的均值和标准差。

c. 分别报告男性和女性的 **hourstv** 和 **readingscores** 的均值和标准差。哪一个组在每一个变量上有较高的均值？哪一个组有较大的标准差？

d. 报告 **gender** 被 **ses** 分层后的 **hourstv** 和 **readingscores** 的均值和标准差。平均来说，哪一个组看电视较多？哪一个组有较高的平均阅读得分？

资源分享
PDG



第3章

作图

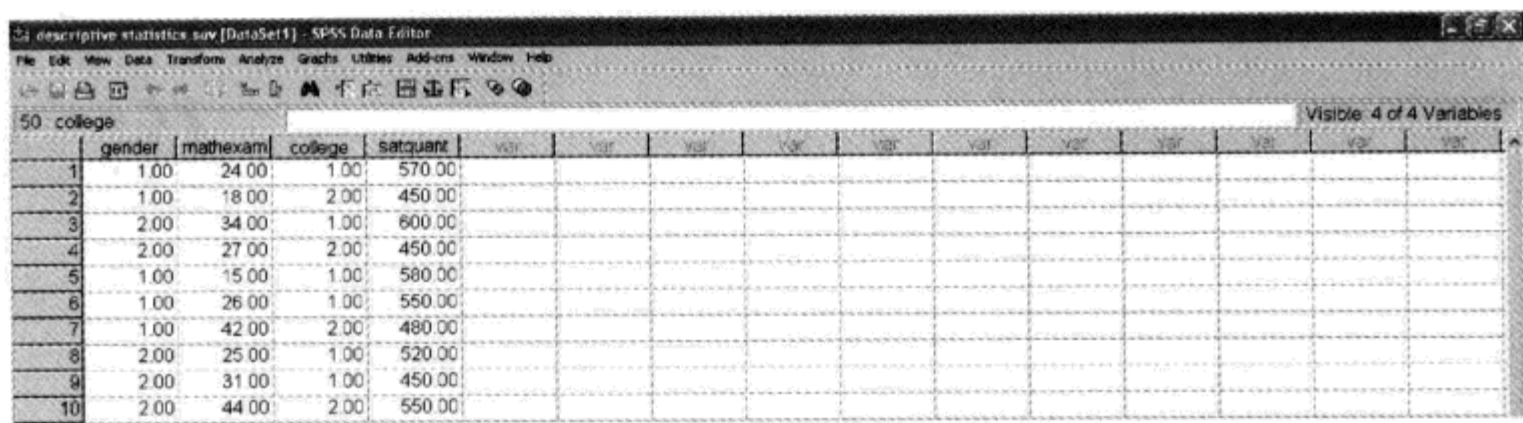
本章我们将在 SPSS 中建立一些不同的图表。通过建立图表，你能很快地确定数据分布的一般形状，数据的整体变异性，检测异常值或那些在数据集中不具代表性的数据值。我们将用 SPSS 建立一些常用的图表，包括条形图、直方图、散点图和盒形图。

为了建立本章的图表，我们将使用在第 2 章中建立的数据文件。该文件的复件位于网站 www.penhall.com/yockey 上文件夹 Chapter 3 之中，文件名为 descriptive statistics.sav。我们从打开这个数据文件开始（注意：下面的命令假定本书的数据文件已经从网上下载到你的电脑中）。

为从电脑上打开文件：

1. 打开 SPSS。
2. 从菜单栏中选择 **File>Open>Data...**。
3. Open Data（打开数据）对话框打开了（如果你在用 SPSS 14.0 或更早的版本，对话框的名字是 Open File（打开文件））。点击下拉列表的箭头，选择本书（SPSS Demystified）的数据文件在你电脑上所处的位置。
4. 双击打开文件夹 Chapter 3。你应该可以看到叫 descriptive statistics.sav 的文件（文件的扩展名 .sav 可能不会出现在你的电脑上）。双击打开这个文件。稍后，这个文件应该会在 SPSS 中的数据编辑窗口处打开。该数据文件显示在图表 3—1 中（注意：如果在你的电脑屏幕上没有看到该文件，从菜单栏选择 **Window>descriptive statistics.sav-SPSS Data Editor（SPSS 数据统**

计编辑器)，以便激活数据编辑窗口)。



	gender	mathexam	college	satquant	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
1	1.00	24.00	1.00	570.00										
2	1.00	18.00	2.00	450.00										
3	2.00	34.00	1.00	600.00										
4	2.00	27.00	2.00	450.00										
5	1.00	15.00	1.00	580.00										
6	1.00	26.00	1.00	550.00										
7	1.00	42.00	2.00	480.00										
8	2.00	25.00	1.00	520.00										
9	2.00	31.00	1.00	450.00										
10	2.00	44.00	2.00	550.00										

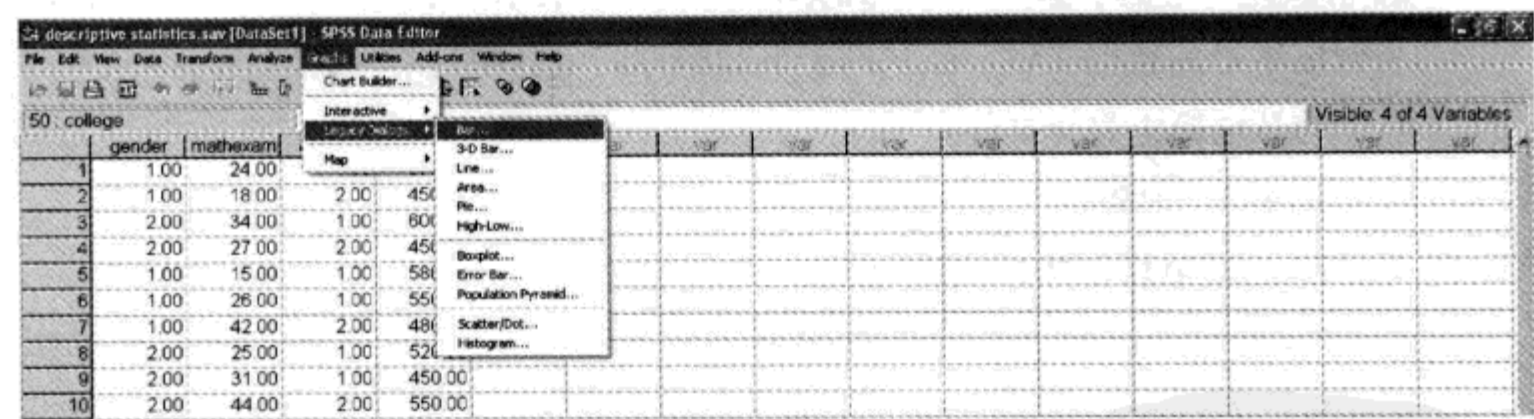
图表 3—1 数据文件 descriptive statistics.sav

条形图

我们在 SPSS 中要建立的第一个图表是条形图 (bar chart)。条形图能表明变量每个类别的频数或观测的数量 (条形图通常是用分类数据建立的)。我们将为变量 college 建立一个条形图。

为在 SPSS 中建立一个条形图：

- 1. 从菜单栏中选择 Graphs>Legacy Dialogs>Bar... (见图表 3—2)。



图表 3—2 条形图的菜单命令

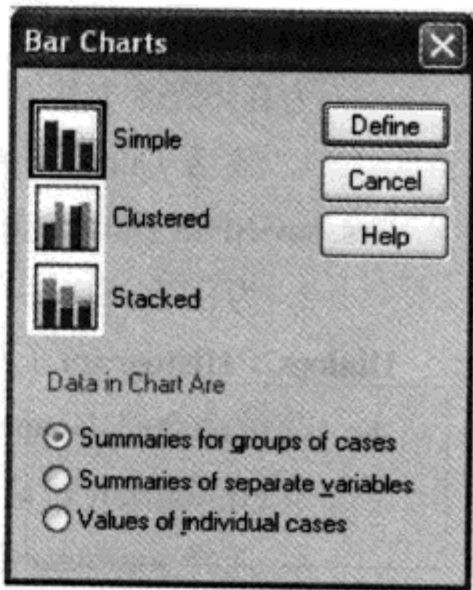
(注意：如果你在用 SPSS 14.0 或更早的版本，菜单命令是 Graphs>Bar，其他的命令是一样的。)

- 2. 打开 Bar Charts 对话框。为了建立 college 的条形图，将使用默认选项 Simple。点击 “Define” (见图表 3—3)。

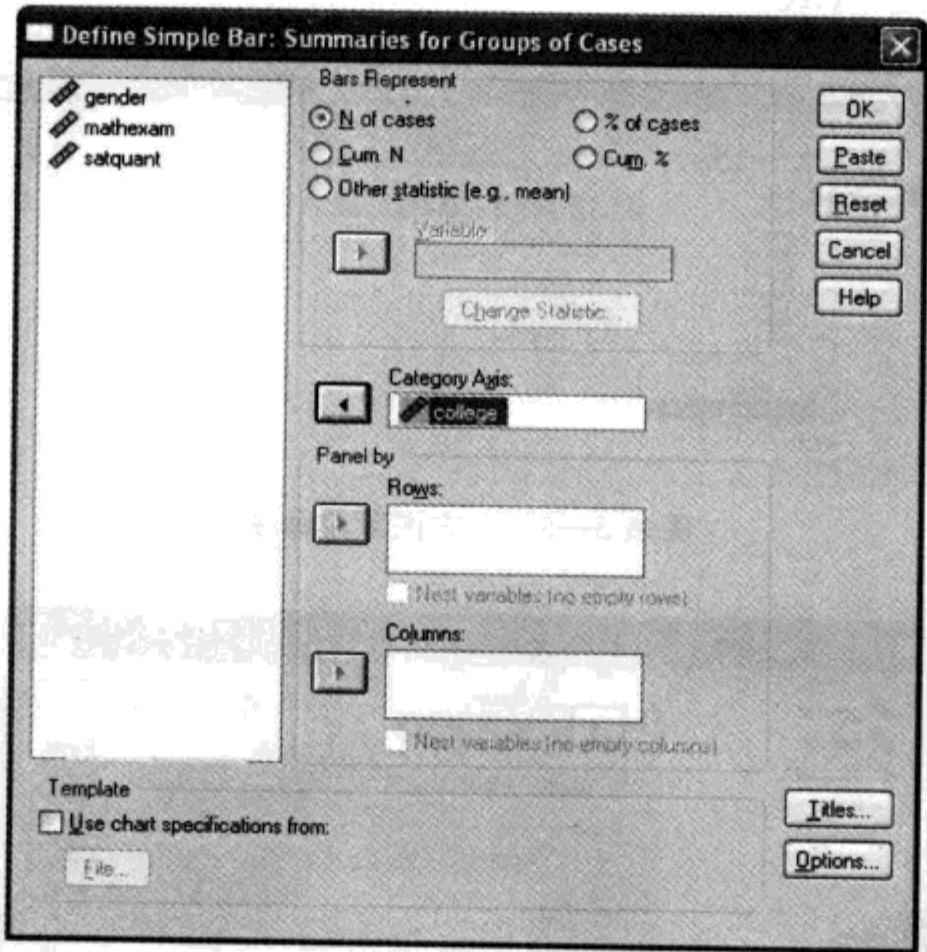
3. 打开 Define simple Bar: Summaries for Groups of Cases (定义简单的条形图: 案例分组摘要) 对话框。选择 **college** 再点击第二个向右箭头按钮 (►), 使其从顶部移到 Category Axis (分类轴) 框中。详见图表 3—4。

4. 点击 “OK”。

college 的条形图建立在 Viewer 窗口中 (在 SPSS 中建立所有的图形后, 再讨论条形图)。



图表 3—3 条形图对话框



图表 3—4 Define Simple Bar: Summaries for Groups of Cases 对话框

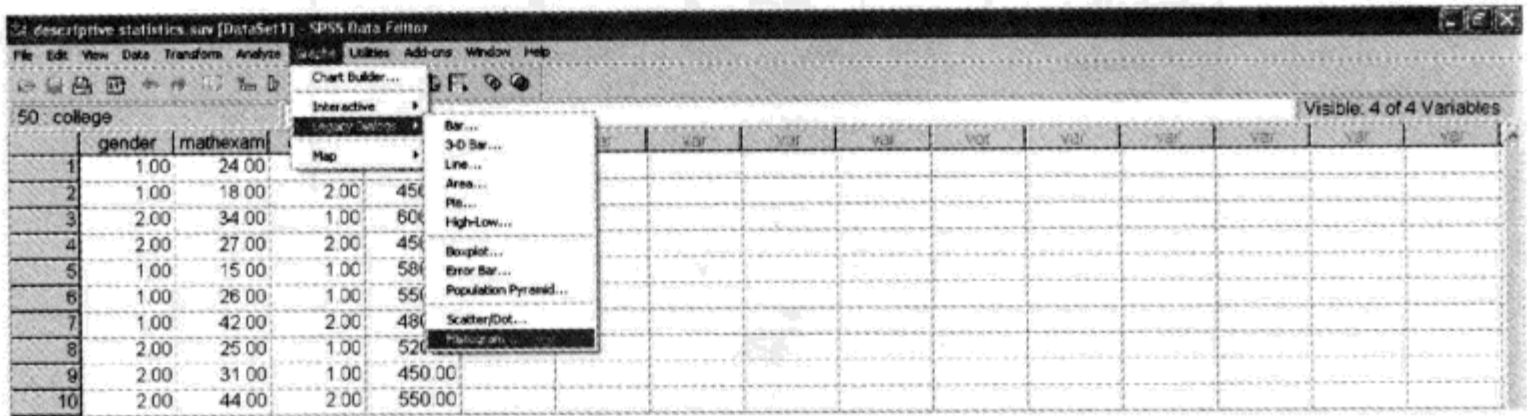
直方图

接下来我们将在 SPSS 中建立一个直方图 (histogram)。直方图能反映连续变量区间上的频数或观测的数量。我们将为变量 **mathexam** 建立一个直

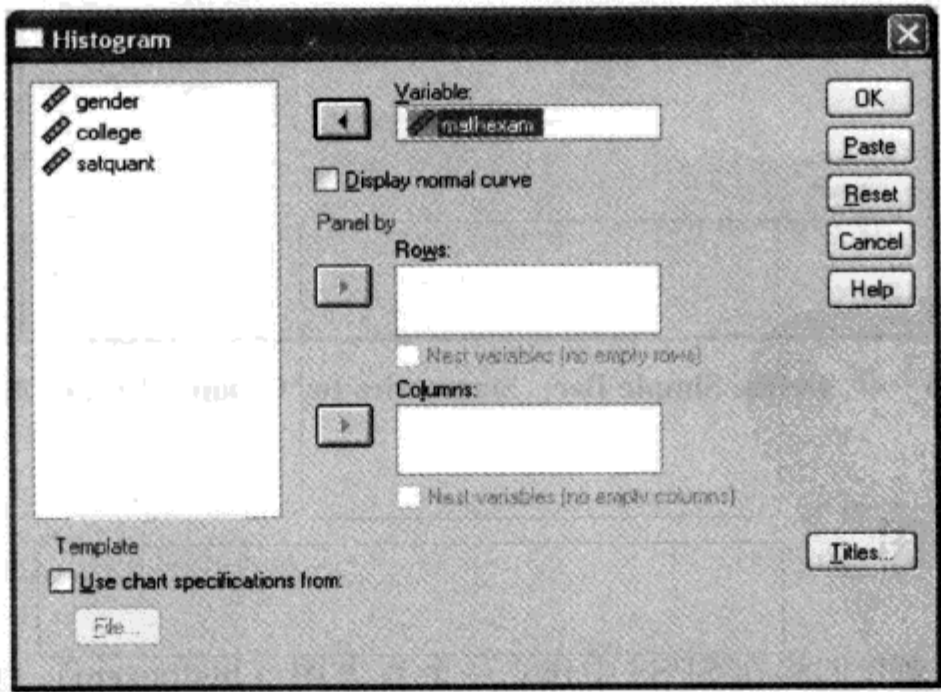
方图。

为在 SPSS 中建立一个直方图：

- 1. 激活 Data Editor 窗口，从菜单栏中选择 **Window > descriptive statistics.sav-SPSS Data Editor**。
- 2. 为了在 SPSS 中建立一个直方图，从菜单栏中选择 **Graphs > Legacy Dialogs > Histogram**（见图表 3—5）（注意：如果你在用 SPSS 14.0 或更早的版本，菜单命令是 **Graphs > Histogram...**。其他的命令是一样的）。
- 3. Histogram 对话框打开了。
- 4. 选择 **mathexam**，点击上面的向右箭头（▶）按钮，使其移到 Variable（变量）框中（见图表 3—6）。
- 5. 点击“OK”。



图表 3—5 直方图菜单命令



图表 3—6 直方图对话框

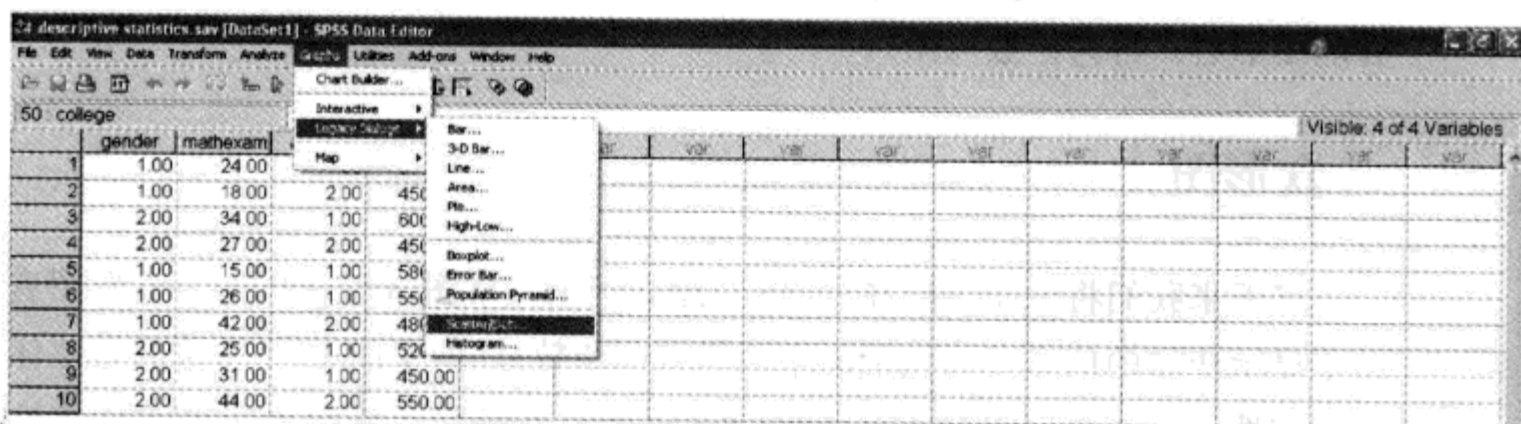
mathexam 的直方图建立在 Viewer 窗口中（在 SPSS 中建立所有的图形后，再讨论直方图）。

散点图

下一步我们将在 SPSS 中建立一个散点图（scatterplots）。散点图把参与者对两个变量的反应画在坐标图上（坐标是两变量相关点的数值），在计算相关系数（相关性将在第 12 章讨论）时经常使用散点图。我们将用变量 **mathexam** 和 **satquant** 建立一个散点图。

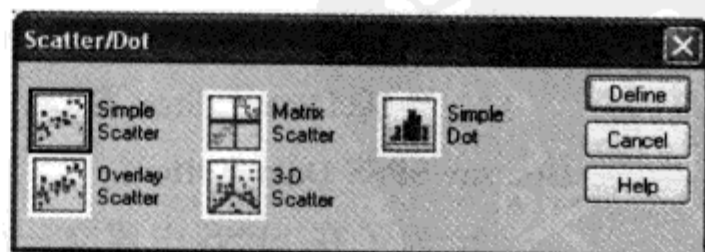
为在 SPSS 中建立一个散点图：

1. 激活 Data Editor 窗口，从菜单栏中选择 **Window > descriptive statistics.sav-SPSS Data Editor**。
2. 为了在 SPSS 中建立一个散点图，从菜单栏中选择 **Graphs > Legacy Dialogs > Scatter/Dot...**（见图表 3—7）（注意：如果你在用 SPSS 14.0 或更早的版本，菜单命令是 **Graphs > Scatter/Dot**。其他的命令是一样的）。
3. Scatter/Dot 对话框打开了（见图表 3—8）。为了建立一个关于 **mathexam** 和 **satquant** 散点图，使用默认选项 Simple Scatter。点击“Define”。

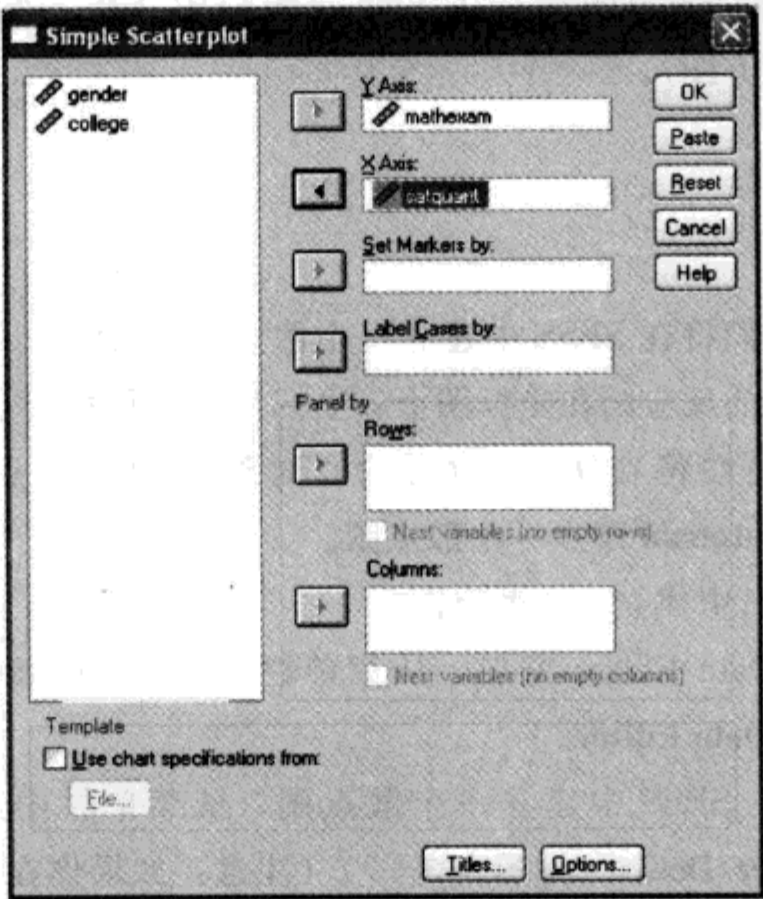


图表 3—7 SPSS 中散点图菜单命令

4. Simple Scatterplot（简单散点图）对话框打开了。选择 **mathexam**，点击上面的向右箭头（▶）按钮，使其移到 Y Axis 框中。选择 **satquant**，点击第二个向右箭头（▶）按钮，使其移到 X Axis 框中（见图表 3—9）。



图表 3—8 散点图对话框



图表 3—9 简单散点图对话框

5. 点击“OK”。

mathexam 和 **satquant** 的散点图建立在 Viewer 窗口中（在 SPSS 中建立所有的图形后，再讨论散点图）。

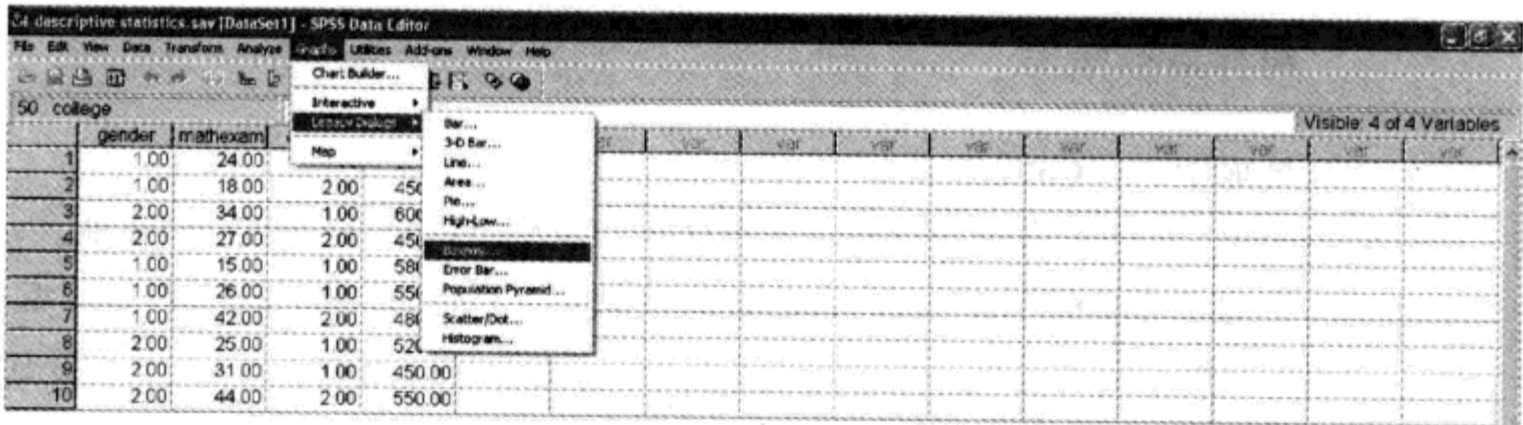
盒形图

接下来我们将说明怎样在 SPSS 中建立一个盒形图（Boxplots）。盒形图有助于显示变量的信息，这包括集中位置（中位数），即中间 50% 的数据，整体的散布性，以及数据中是否有异常值（异常值是指那些在数据集里无代表性的点）。我们将为变量 **mathexam** 建立盒形图。

为在 SPSS 中建立一个盒形图：

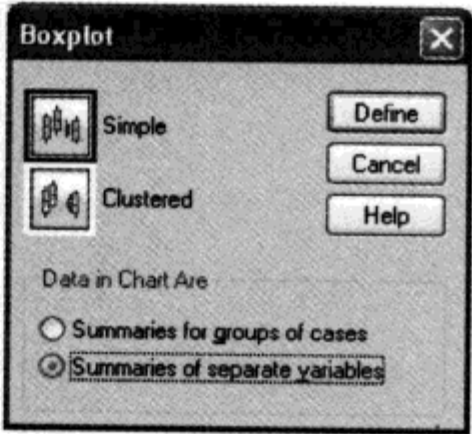
1. 激活 Data Editor 窗口，从菜单栏中选择 **Window > descriptive statistics. sav-SPSS Data Editor**。
2. 为了在 SPSS 中建立一个盒形图，从菜单栏中选择 **Graphs > Legacy Dialogs > Boxplot...**（见图表 3—10）（注意：如果你在用 SPSS 14.0 或更早的版本，菜单命令是 **Graphs > Boxplot**。其他的命令是一样的）。

3. Boxplot 对话框打开了。保持默认选项 Simple。在对话框的 Data in Chart Are 中选择 Summaries of separate variables (见图表 3—11)。
4. 点击 “Define”。

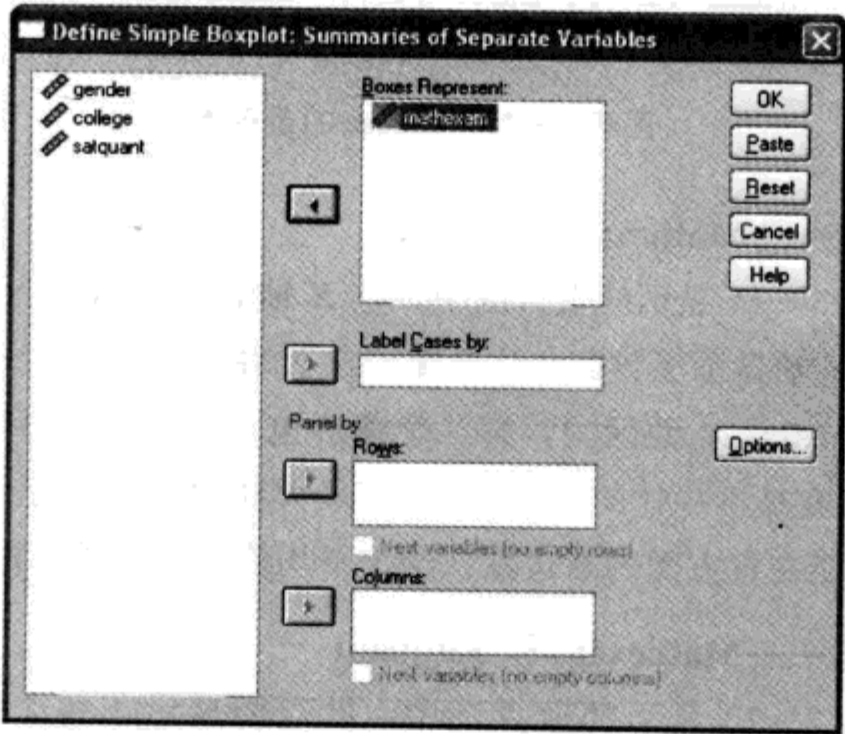


图表 3—10 盒形图菜单命令

5. 打开 Define Simple Boxplot: Summaries of Separate Variables (定义简单盒形图：独立变量摘要) 对话框，选择变量 **mathexam**，再点击上面的向右箭头 (►) 按钮，以便把它移到 Boxes Represent 框中 (见图表 3—12)。
6. 点击 “OK”。



图表 3—11 盒形图对话框



图表 3—12 Define Simple Boxplot: Summaries of Separate Variables 对话框

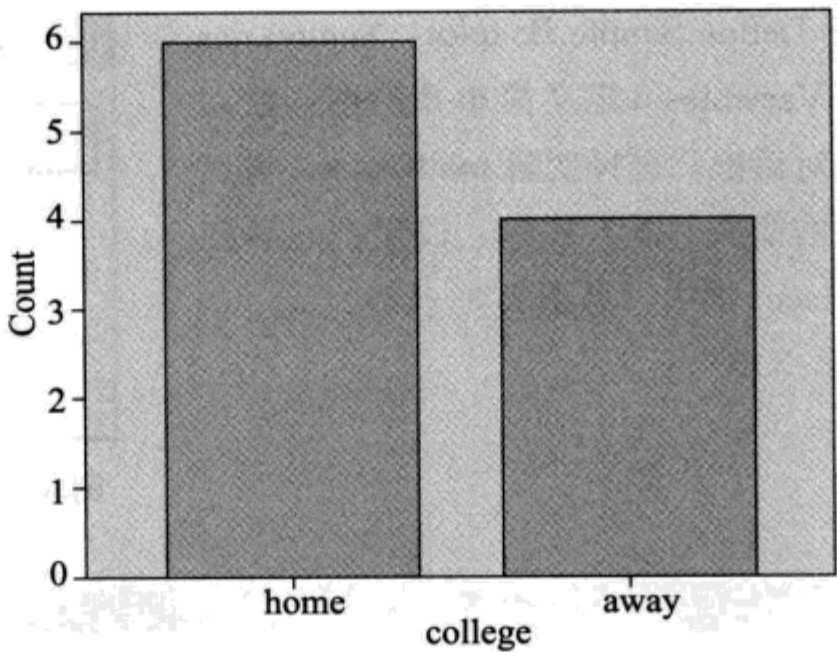
mathexam 的盒形图建立在 Viewer 窗口中。

接下来，我们描述建立的四个不同的图表——条形图、直方图、散点图和盒形图——的结果。

解释这些结果

条形图——College

图表 3—13 显示的 **college** 的条形图中，组别（本地（home）和外地（away））显示在横轴（X 轴），频数显示在纵轴（Y 轴）。对于给定的每一组，条形的高度对应着每组的频数。条形图表明 6 个学生在本地上大学，4 个学生在外地上大学。



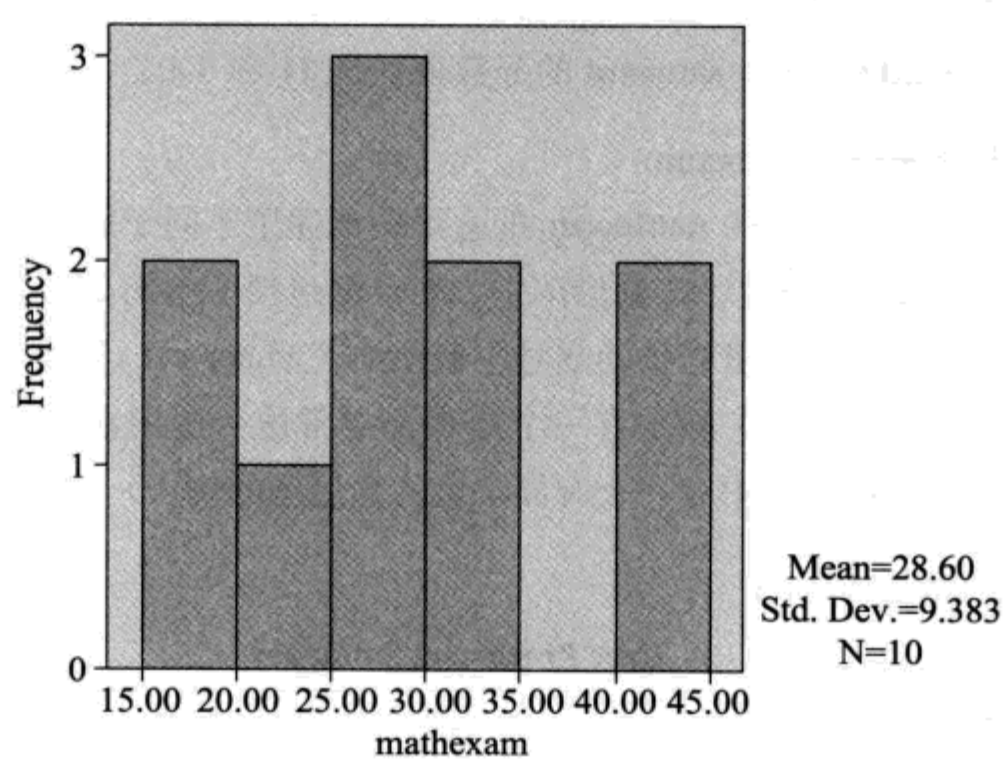
图表 3—13 college 的条形图

直方图——Mathexam

在图表 3—14 显示的直方图中，在 X 轴显示了 **mathexam** 的值（从最小到最大），在 Y 轴显示了频数。注意图中在 X 轴上每个条形的尺度是 5 分：第一个条形包含 15~20 的区间，第二个条形包含 20~25 的区间，等等（不同图中区间的尺度是不一样的）。图表中有最大频数的区间是 25~30，频数为 3。注意 SPSS 能在直方图的右边默认给出平均值、标准差和样本量。

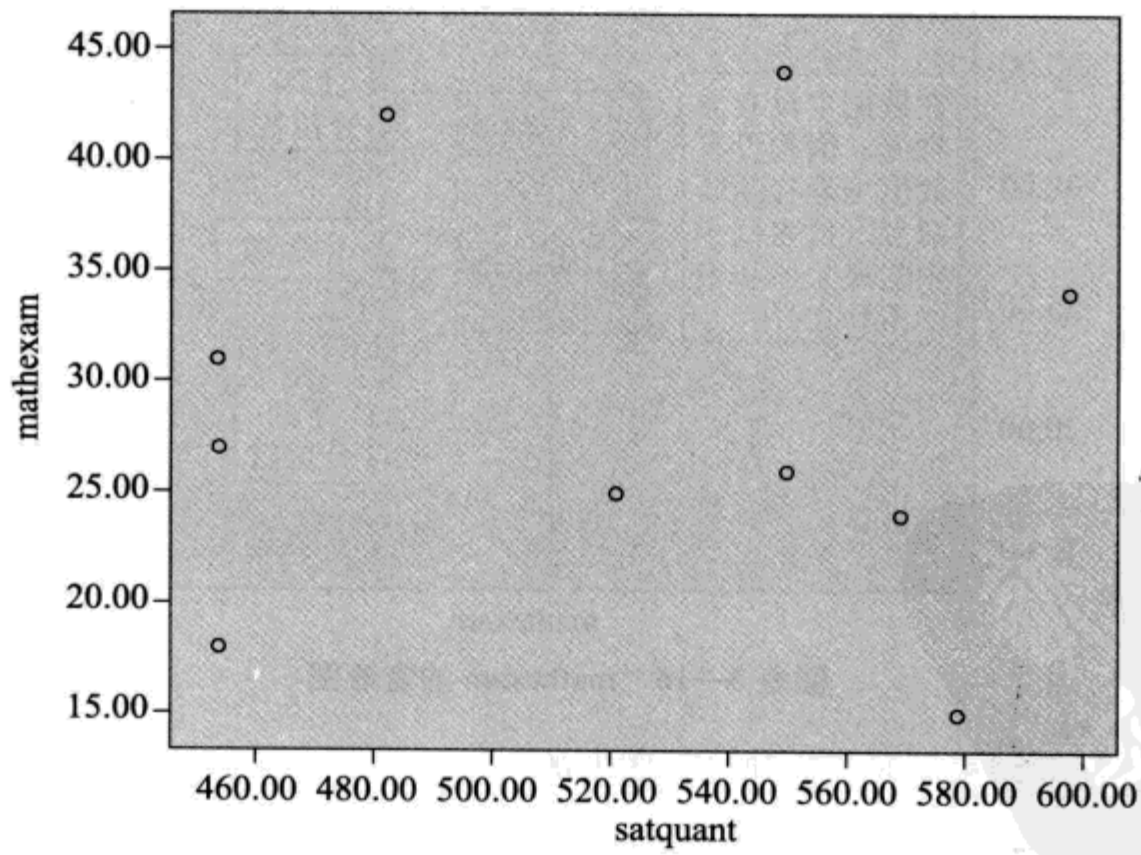
散点图——Mathexam 和 Satquant

在图表 3—15 显示的散点图中，散点图显示了每个参与者在两个变量



图表 3—14 mathexam 的直方图

mathexam 和 **satquant** 上的坐标（每个坐标是图中的圆圈）。正如较早前我们在 Simple Scatterplot 对话框里指定的，变量 **satquant** 的值显示在 X 轴上，把变量 **mathexam** 的值显示在 Y 轴上。在图中有 10 个不同的坐标，每个坐标显示



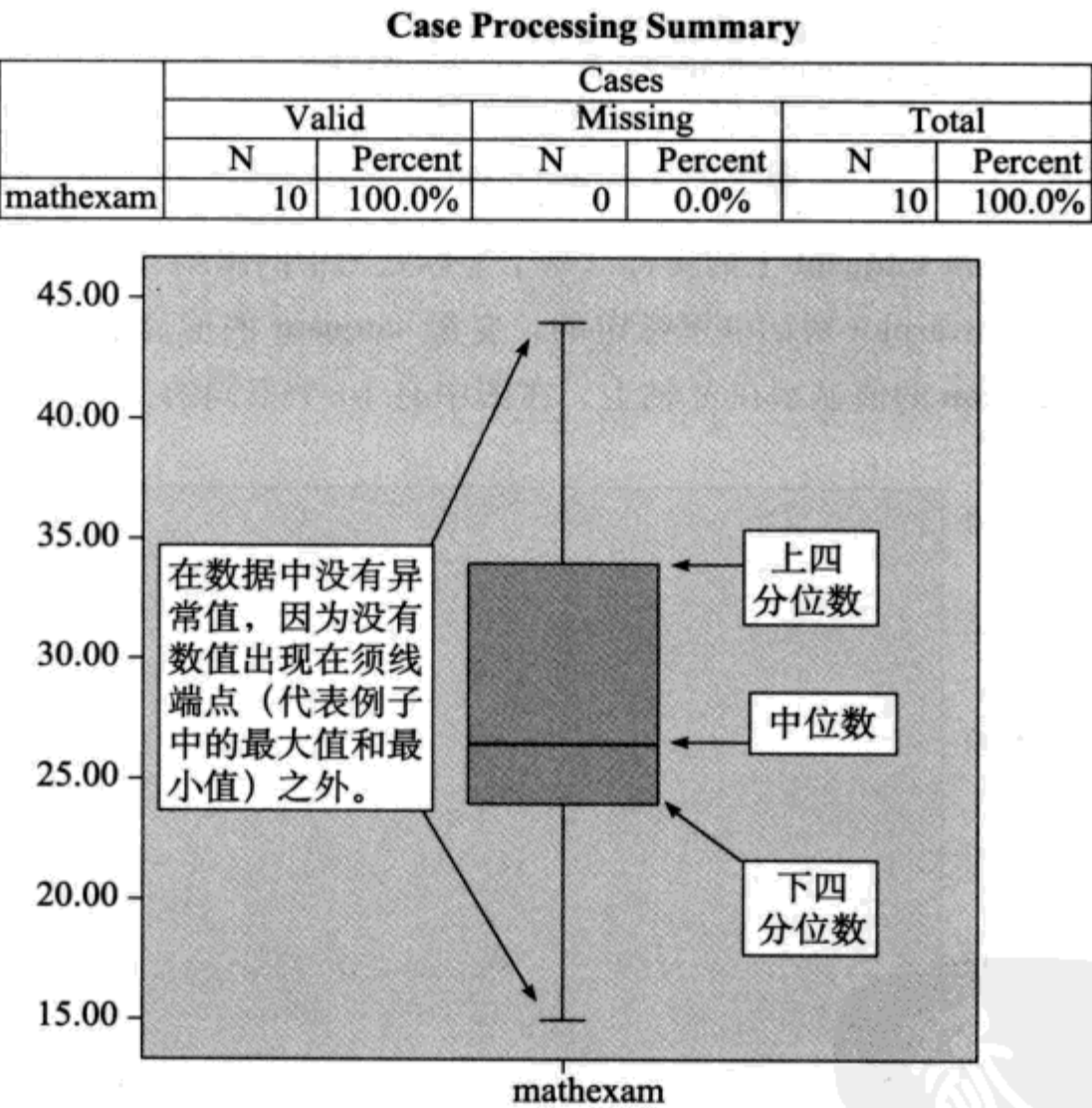
图表 3—15 mathexam 和 satquant 的散点图

了每个参与者在两个变量上的得分。例如，图中右边较远的坐标表示第三个参与者在 **mathexam** 和 **satquant** 的分数分别是 34 和 600。

盒形图——Mathexam

图表 3—16 显示的 **mathexam** 的盒形图概述了不同于较早前直方图中计算得到的 **mathexam** 分数。在盒形图中，矩形框包括了中间 50% 的数据，框内的线等于中位数。从框里延长的线对应着须线（whisker）。须线上的两个端点代表了最大值和最小值，除非有一个或多个异常值，这用在须线之外的星号表示（在本章末的练习中有一个有异常值的盒形图的例子）。

Boxplot-mathexam



图表 3—16 mathexam 的盒形图

至此结束了对我们建立的图形的讨论。

在 SPSS 中建立条形图、直方图、散点图和盒形图的步骤摘要

I. 数据输入

1. 在 Variable View 窗口中输入变量名，以便在 SPSS 中生成变量。
2. 为每个分类变量建立变量值标签。在 Variable View 窗口，输入适当的数值和标签。点击“OK”。
3. 输入每个参与者的数据。

II. 数据分析

为获得一个条形图：

1. 从菜单栏中选择 **Graphs>Legacy Dialogs>Bar...**（注意：如果你在用 SPSS 14.0 或更早的版本，选择 **Graphs>Bar**）。
2. 在 Bar Charts（条形图）对话框里，保留默认选项 Simple。点击“Define”。
3. 在 Define Simple Bar: Summaries for Groups of Cases 对话框中，把感兴趣的变量移到 Category Axis（分类轴）对话框。
4. 点击“OK”。

为获得一个直方图：

1. 从菜单栏中选择 **Graphs>Legacy Dialogs>Histogram**（注意：如果你在用 SPSS 14.0 或更早的版本，菜单命令是 **Graphs>Histogram**）。
2. 在 Histogram 对话框里，把感兴趣的变量移到 Variable 框中。
3. 点击“OK”。

为获得一个散点图：

1. 从菜单栏中选择 **Graphs>Legacy Dialogs>Scatter/Dot...**（注意：如果你在用 SPSS 14.0 或更低的版本，菜单命令是 **Graphs>Scatter/Dot**）。
2. 在 Scatter/Dot 对话框里，保持默认选项 Simple Scatter。点击“Define”。
3. 在 Simple Scatterplot 对话框里，把一个变量移到 Y Axis 框中，把另一个变量移到 X Axis 框中。
4. 点击“OK”。

为获得一个盒形图：

1. 从菜单栏中选择 **Graphs > Legacy Dialogs > Boxplots...** (注意：如果你在用 SPSS 14.0 或更早的版本，菜单命令是 **Graphs > Boxplot**)。
2. 在 Boxplot 对话框里，保持默认选项 Simple。在 Data in Chart Are 中选择 Summaries of separate variables。点击 “Define”。
3. 在 Define Simple Boxplot: Summaries of Separate Variables 对话框中，把感兴趣的变量移到 Boxes Represent 框中。
4. 点击 “OK”。

练习

1. 数据来自当地大学的 10 名二年级学生，包括他们的性别 (**gender**, 1=男孩, 2=女孩)，他们每天的学习时间 (**hoursstudied**)，他们每天喝咖啡的杯数 (**cupscoffee**) 和他们的平均成绩 (**gpa**)。数据在网站 www.prenhall.com/yockey 上文件夹 Chapter 3 内的文件 Chapter 3 _ Exercise 1. sav 之中。在 SPSS 中打开文件并进行适当分析以回答下列问题。

a. 为变量 **gender** 建立一个条形图并且打印结果。在这个数据中分别有多少男学生和女学生？

b. 为变量 **hoursstudied** 建立一个直方图并且打印结果。每天学习时间的平均值是多少？

c. 用变量 **gpa** 和 **hoursstudied** 建立一个散点图并且打印结果 (把 **gpa** 放在 Y 轴)。

d. 为变量 **cupscoffee** 建立一个盒形图并且打印结果。在图中有异常值吗？

2. 数据文件 Chapter 3 _ Exercise 2. sav 中包括 15 个人的数据，包括他们是否有过敏症 (**allergy**; 1= “yes”, 2= “no”), 去年见医生的次数 (**doctorvisits**) 和去年没有上班的天数 (**daysmissed**)。在 SPSS 中打开数据并进行适当分析以回答下列问题 (这个文件位于网站上的文件夹 Chapter 3 内)。

a. 为变量 **allergy** 建立一个条形图并打印结果。多少人有过敏症？多少人没有过敏症？

b. 为变量 **daysmissed** 建立一个直方图并打印结果。每年没有上班天数的平均值是多少？

- c. 为变量 **doctervisits** 建立一个盒形图并打印结果。在图中有异常值吗？
3. 数据来自 7 个孩子，包括每周平均吃快餐 (**fastfood**) 的时间和喜欢玉米 (**corn**)、胡萝卜 (**carrots**) 和卷心菜 (**cabbage**) 的程度 (尺度为 1~10, 高的分数表明更高的爱好等级)。这些数据位于网站上的文件夹 Chapter 3 中的文件 Chapter 3 _ Exercise 3. sav 之中。在 SPSS 中打开文件并进行适当分析以回答下列问题。
- a. 为变量 **fastfood** 建立一个直方图并打印结果。
- b. 为变量 **corn** 建立一个直方图并打印结果。**corn** 的爱好等级的平均值是多少？
- c. 用变量 **fastfood** 和 **carrots** 建立一个散点图并打印结果 (把 **carrots** 放在 Y 轴)。
- d. 为变量 **cabbage** 建立一个盒形图并打印结果。在图中有异常值吗？



可靠性 (用 α 系数度量)

可靠性指某些感兴趣的测量结果的一致性 or 可重复性。例如，对一组参与者在两周内进行两次信心的 10 个项目的测量，如果参与者对这些项目的反应相同或很相近，这个测量的可靠性就很高；如果反应有很大的不同，这个测量的可靠性就很低。

尽管有几种不同的方法用来估计可靠性，但是大部分都可以归为两种形式：基于对一个测量多次执行之上的可靠性和基于对一个测量单独执行之上的可靠性。¹ 评估一个测量多次执行的可靠性包括在两个不同时刻执行量表以及测量参与者在两个执行内对这些项目的反应的一致性。对于这种类型的可靠性，重测信度 (test-retest reliability, 重复测量两次) 和复本信度 (alternate form reliability, 执行两次非常类似的测量，每次一个版本) 是两种常用的方法。

评估一个测量单独执行的可靠性包括执行一次量表以及测量参与者对测量的所有项目反应的一致性。对一个测量单独执行的两种不同的可靠性是折半信度 (split-half reliability) 和 α 系数 (也称为 Cronbach's α)。折半信度是将受测题目分成两半，然后估计参与者对这两部分反应的一致性。此外， α 系数不是对测量的单独划分进行计算，而是对测量所有可能的花费的均值。当使用折半系数或 α 系数对测量的可靠性进行估计时，重要的是量表内所有项目测量同样的性质，否则内部的一致性将大大减弱。在本章，我们将把重点放在 α 系数上，这是估计可靠性的一种最常用的方法。

α 系数的取值范围通常是 0~1，取值越高表明一系列项目的内部一致性越强。² 图表 4—1 显示的是 α 系数不同取值范围的正确性的近似准则。

下面将介绍一个 α 系数应用的例子。

图表 4—1 对不同 α 系数值的内部一致可靠性估计的正确性

α 系数	正确性
0.90 及以上	优秀
0.80~0.89	好
0.70~0.79	一般
0.60~0.69	边缘
0.59 及以下	糟糕

例子

作为研究项目的一部分，一个学生想估计生活意义中 5 个项目的测量值的内部一致可靠性。关于生活意义量表的项目显示在图表 4—2 中。

图表 4—2 生活意义量表的项目

项目	强烈 不同意	不同意	既不是同意 也不是不同意	同意	强烈同意
1. 我对我的人生方向感觉很好	1	2	3	4	5
2. 我的生活有意义	1	2	3	4	5
3. 通常我觉得是在我的生活轨道上	1	2	3	4	5
4. 我在生活中没有迷失过	1	2	3	4	5
5. 目前我正努力工作，以实现自己的生活目标	1	2	3	4	5

说明：对每个项目，反应选项是 1（强烈不同意），2（不同意），3（既不是同意也不是不同意），4（同意），5（强烈同意）。每个参与者被要求对每个问题仅选择一个反应选项。

生活意义量表的每个项目的反应范围是 1~5，1=强烈不同意，5=强烈同意。把 5 个项目的反应加在一起可以得到量表的总分数，对每个参与者来说，总得分的取值范围是 5~25。量表中的得分越高表明生活意义水平越高。

20 个大学生回答了这 5 个项目的测量。

如果关于生活意义量表的项目有一个较高的内部一致性水平 (量表是可靠的), 那么参与者需要在量表的每个项目上都有一致性的反应。例如, 对图表 4—2 中的项目, 在生活意义上有较高得分的那些人应该同意或强烈同意这些项目, 而在生活意义上有较低得分的那些人应该不同意或强烈不同意这些项目。参与者的这种一致性反应模式将导致一个高的系数值, 因为参与者在量表上有一致性的反应 (例如, 一个人同意所有的项目, 另一个人不同意这些项目, 等等)。该规则的一个例外是当量表中包括一个或更多个负向项目时, 这是下一步要讨论的。

负向项目和反向编码

在图表 4—2 中给出了所有关于生活意义量表的项目, 因此, 有较高生活意义的人倾向于对它们表示同意或强烈同意 (假定量表是可靠的)。然而, 一些测量设计使用了一些混合项目, 其中, 在感兴趣的特征上有较高得分的人同意一些项目, 并且不同意其他的项目。那些在感兴趣的特征上有较高得分的人不同意的项目被认为是负向项目 (negative items) (如果在图表 4—2 中的项目 2 被重写为“我的生活几乎没有意义”, 这就是一个负向项目的例子。有较高生活意义的人最有可能不同意或强烈不同意这些负向项目)。如果一个量表里包含一个或更多的负向项目, 在估计量表的内部一致性之前, 这些项目都需要反向编码。反向编码是这样一个过程: 对负向项目的反应是按照字面反向的, 因此 1 变成了 5, 2 变成了 4, 等等。在附录 A 里说明了负向项目的逆向编码。

α 系数的目标和数据要求

α 系数		
目标	数据要求	例子
在测量中估计反应的内部一致性	需要至少两个项目 (变量) 的得分 (尽管推荐超过两个项目)。所有的项目应该测量同样的特点或特征	生活意义量表包括 5 个项目 (变量)。将确定参与者在 5 个项目上的反应的内部一致性

数据

在图表 4—3 中列出了 20 个参与者对生活意义量表中 5 个项目的反应。表中每一行包括一个参与者的反应，每一列代表了量表的不同项目。**meaning1** (意义 1) 对应于量表中的第一个项目的反应，**meaning2** (意义 2) 对应于量表中第二个项目的反应，等等。

图表 4—3 参与者在生活意义量表中 5 个项目的反应

Participant	meaning1	meaning2	meaning3	meaning4	meaning5
1	5	5	4	5	4
2	2	3	2	1	3
3	1	2	2	1	2
4	5	5	5	4	5
5	3	4	4	2	3
6	5	5	2	4	3
7	5	4	4	3	4
8	4	4	4	4	4
9	5	4	4	4	3
10	4	5	5	4	3
11	5	5	5	5	5
12	3	2	2	3	3
13	1	1	1	2	1
14	5	5	5	5	5
15	5	4	4	4	5
16	4	4	3	4	3
17	4	4	4	4	3
18	5	5	5	5	5
19	5	4	5	5	4
20	4	5	5	5	4

说明：变量 participant 包含在数据中，但不用输入 SPSS。

在 SPSS 中输入数据及分析

下面的步骤 1 和步骤 2 描述了如何在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 4 中的文件 meaning.sav 之中。

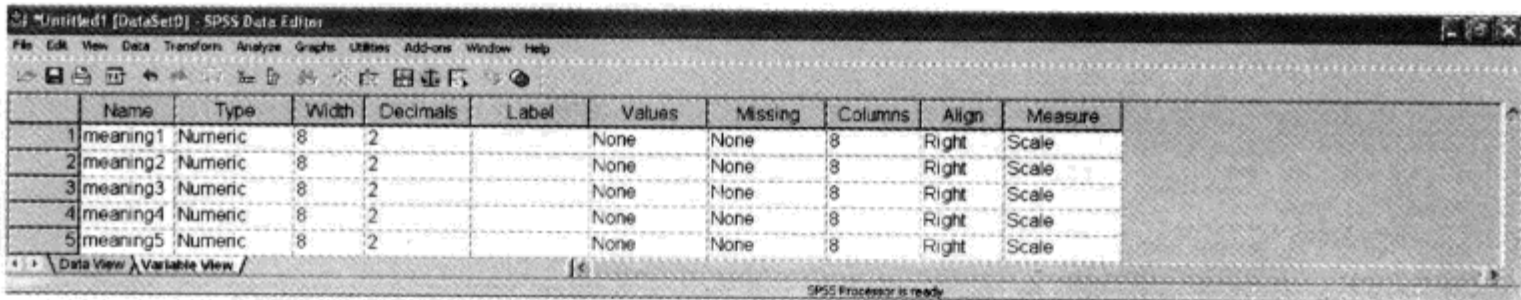
如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中将生成 5 个变量，每个变量对应于生活意义量表中的一个项目。变量将分别命名为 **meaning1** (意义 1)，**meaning2** (意义 2)，**meaning3** (意义 3)，**meaning4** (意义 4) 和 **meaning5** (意义 5)。

- 3. 在 Variable View 窗口前五行输入变量名称 **meaning1**，**meaning2**，**meaning3**，**meaning4** 和 **meaning5** (详见图表 4—4)。



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	meaning1	Numeric	8	2		None	None	8	Right	Scale
2	meaning2	Numeric	8	2		None	None	8	Right	Scale
3	meaning3	Numeric	8	2		None	None	8	Right	Scale
4	meaning4	Numeric	8	2		None	None	8	Right	Scale
5	meaning5	Numeric	8	2		None	None	8	Right	Scale

图表 4—4 输入变量 meaning 1~meaning 5 的 Variable View 窗口

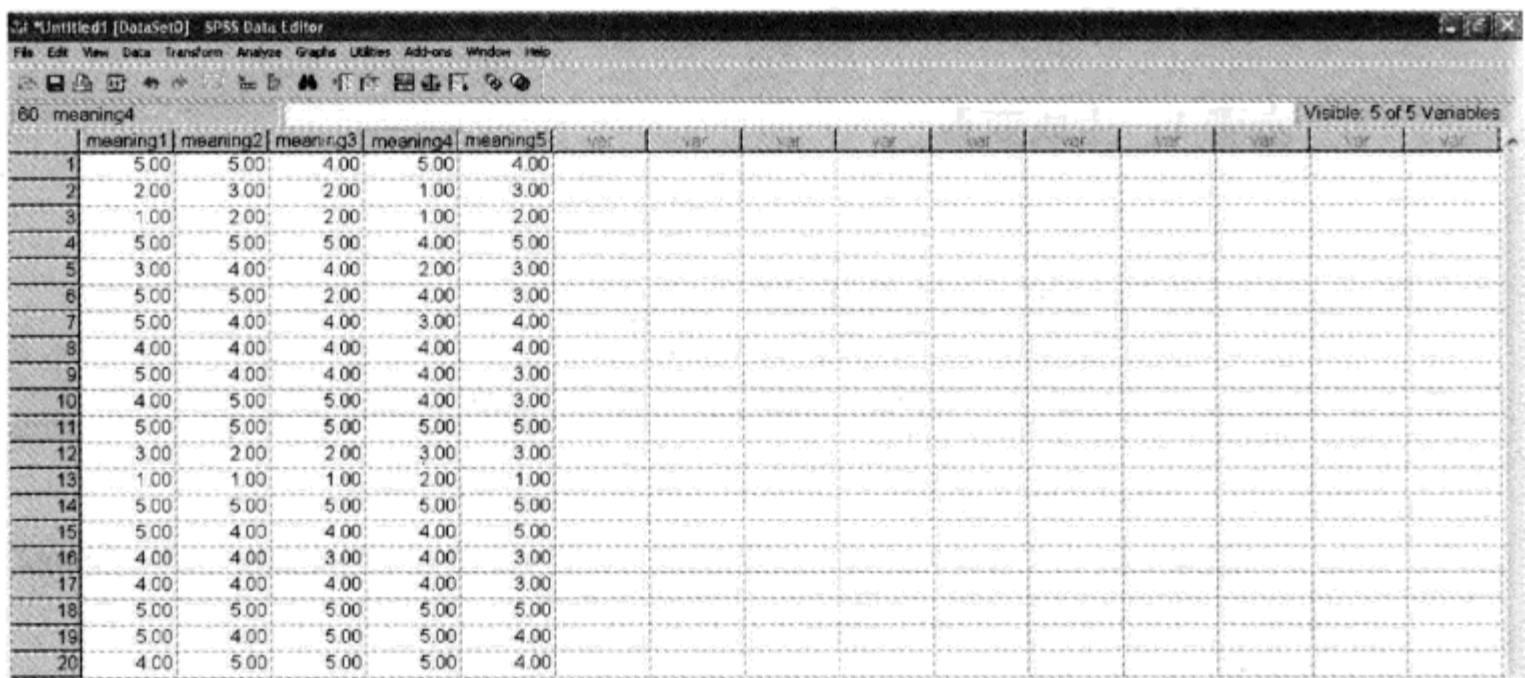
步骤 2：输入数据

- 1. 点击 Data View 标签。变量 **meaning1**，**meaning2**，**meaning3**，**meaning4** 和 **meaning5** 出现在 Data View 窗口的前五列。
- 2. 参照图表 4—3，为每个参与者输入 5 个变量的数据。对第一个参与者，为变量 **meaning1**，**meaning2**，**meaning3**，**meaning4** 和 **meaning5** 分别输入 5、5、4、5 和 4。依次输入全部 20 个参与者的数据。图表 4—5 给出了完整的数据文件。

步骤 3：分析数据

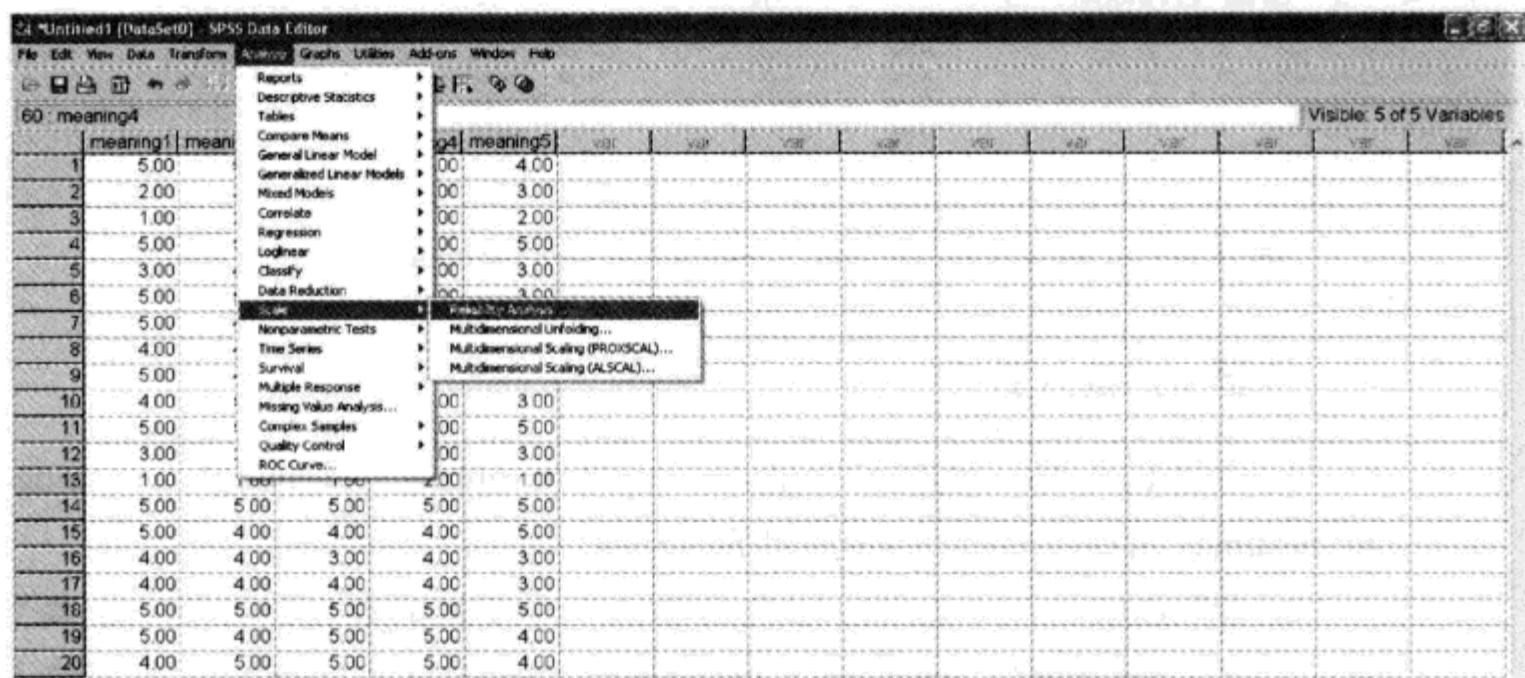
- 1. 从菜单栏中选择 **Analyze>Scale>Reliability Analysis** (可靠性分析)... (见图表 4—6)。
- 2. 打开 Reliability Analysis 对话框，变量 **meaning1** ~ **meaning5** 出现在对话框的左边 (见图表 4—7)。

SPSS 其实很简单

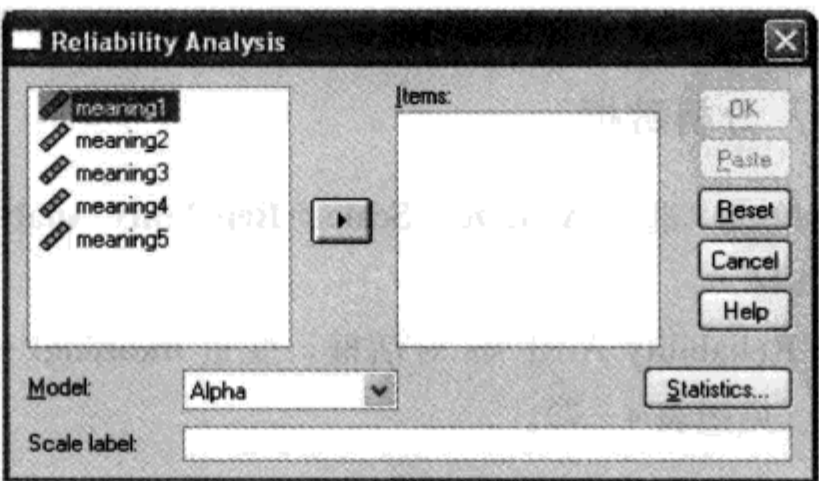


	meaning1	meaning2	meaning3	meaning4	meaning5										
1	5.00	5.00	4.00	5.00	4.00										
2	2.00	3.00	2.00	1.00	3.00										
3	1.00	2.00	2.00	1.00	2.00										
4	5.00	5.00	5.00	4.00	5.00										
5	3.00	4.00	4.00	2.00	3.00										
6	5.00	5.00	2.00	4.00	3.00										
7	5.00	4.00	4.00	3.00	4.00										
8	4.00	4.00	4.00	4.00	4.00										
9	5.00	4.00	4.00	4.00	3.00										
10	4.00	5.00	5.00	4.00	3.00										
11	5.00	5.00	5.00	5.00	5.00										
12	3.00	2.00	2.00	3.00	3.00										
13	1.00	1.00	1.00	2.00	1.00										
14	5.00	5.00	5.00	5.00	5.00										
15	5.00	4.00	4.00	4.00	5.00										
16	4.00	4.00	3.00	4.00	3.00										
17	4.00	4.00	4.00	4.00	3.00										
18	5.00	5.00	5.00	5.00	5.00										
19	5.00	4.00	5.00	5.00	4.00										
20	4.00	5.00	5.00	5.00	4.00										

图表 4—5 系数 α 的例子中的完整数据文件

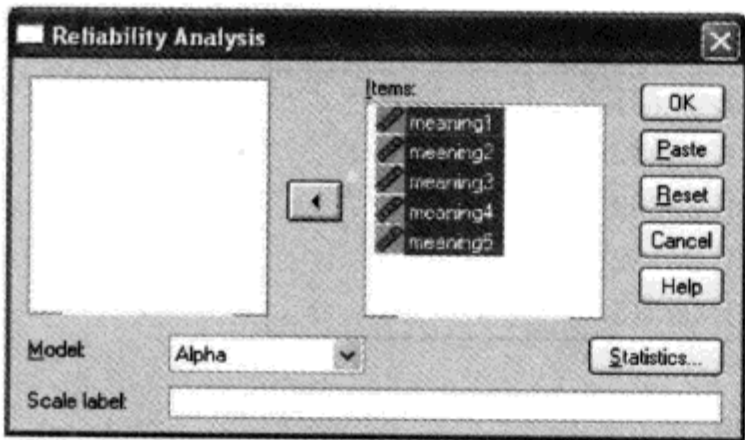


图表 4—6 可靠性分析的菜单命令

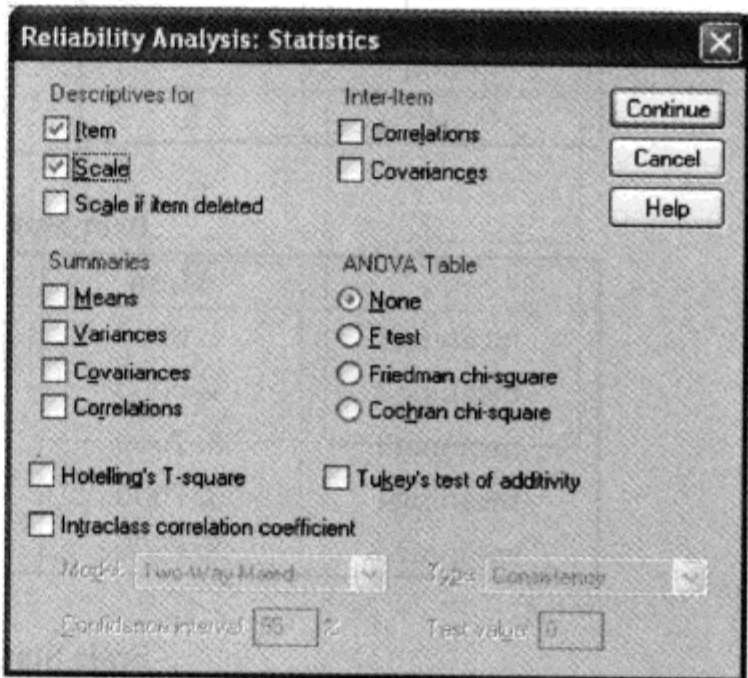


图表 4—7 可靠性分析对话框

- 3. 选择变量 **meaning1**，拖拽鼠标选择剩余四个变量。所有 5 个变量都应该被选择。点击向右箭头按钮 (►)，把变量都移到 Items 框中 (见图表 4—8)。
- 4. 点击 “Statistics”。Reliability Analysis: Statistics (可靠性分析: 统计) 对话框打开了。在 Descriptives for 下选择 Item 和 Scale (见图表 4—9)。
- 5. 点击 “Continue”。
- 6. 点击 “OK”。



图表 4—8 可靠性分析对话框 (续)



图表 4—9 可靠性分析: 统计对话框

在 SPSS 中运行可靠性程序，结果显示在 Viewer 窗口中。

步骤 4: 解释结果

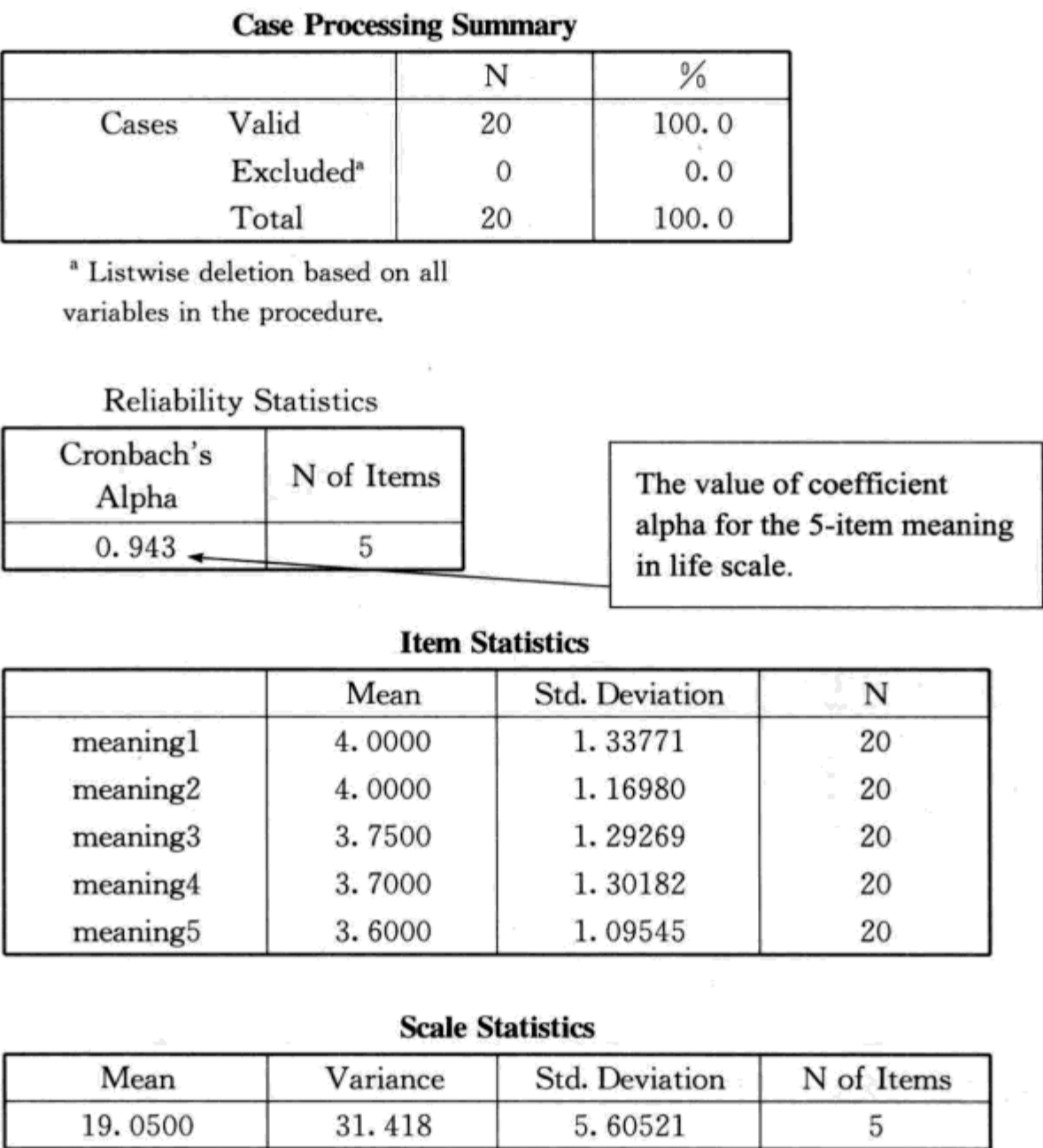
可靠性分析的输出结果显示在图表 4—10 中。

案例处理摘要

第一个表格 Case Processing Summmary 显示了计算量表可靠性所需案例 (参与者) 的数量。表中显示 N 为 20，这表明所有的 20 个参与者包括在分析中 (在所有的 5 个变量中他们有完整的数据)。

可靠性统计 (Reliability Statistics)

第二个表格 Reliability Statistics 显示了 5 个项目量表的系数值 (也叫 Cronbach's α)。系数值是 0.94 (四舍五入至小数点后两位)，这表明量表中的



图表 4—10 可靠性步骤的输出结果

5 个项目有很高的内部一致性。

项目统计量 (Item Statistics)

下一个表格 Item Statistics 提供了量表的每个项目的均值、标准差和样本量。如果你要检查 5 个项目的均值，就会注意到参与者在项目 1 (**meaning1**) 和项目 2 (**meaning2**) 上有较高排名，平均反应为 4.00，这相当于对量表的“同意”的反应。剩下的三个项目上的平均反应在 3.60~3.75 之间。项目 1 是变化最大的，近似标准差为 1.34 (5 个项目的标准差是非常类似的)。总体上，参与者在 5 个项目上的反应表明他们发现自己的生活是相当有意义的，对每个项目的平均反应要么等于“同意”，要么在“同意”的方向上 (值为 4.00)。

尺度统计量 (Scale Statistics)

最后一个表格 Scale Statistics 提供了总量表的均值、标准差和样本量 (总量表等于 5 个项目的和)。5 个项目的量表的均值是 19.05 (回想一下总量表可能的范围是 5~25), 标准差是 5.61。因此, 在这个量表上, 参与者在个人意义上有相当高的排名。量表的均值是评估参与者落在感兴趣的特性之中的一个替代的估计方法, 它等于 5 个项目的均值之和 ($4.00 + 4.00 + 3.75 + 3.70 + 3.60 = 19.05$)。

结果的表达

在本章的开始, 给出了每个方法的简明书面结果。在写可靠性分析结果时, 通常需要提供总量表的 α 系数值、均值和标准差 (如果需要, 单个项目的均值和标准差也可以显示在表格中)。下面给出了结果的简单书面例子。

我们对有 5 个项目的意义量表的内部一致性进行了研究。该量表的 α 系数是 0.94, 显示这个量表的项目有很高的内部一致性。每个项目的均值取值范围是 3.65~4.00, 总量表的均值是 19.05 ($SD=5.61$)。该量表显示参与者的生活有很高的意义。图表 4—11 提供了在生活意义量表里项目的均值和标准差。³

图表 4—11 在生活意义量表里的项目的均值和标准差

Item	M	SD
1	4.00	1.34
2	4.00	1.17
3	3.75	1.29
4	3.70	1.30
5	3.60	1.10

在 SPSS 中执行 α 系数的步骤摘要

I. 输入数据和分析

- 1. 在 SPSS 中建立与量表中项目一样多的变量。
- 2. 输入参与者对量表里每个项目的反应。

- 3. 选择 **Analyze>Scale>Reliability Analysis...**。
- 4. 把量表里的每一个项目移到 Items 框中。
- 5. 点击 “Statistics”。在 Descriptives for 下选择 Item 和 Scale。点击 “Continue”。
- 6. 点击 “OK”。

II. 结果解释

- 1. 注意 Reliability Statistics 表中的 (Cronbach) α 系数的值。
- 2. 观察 Items Statistics 表和 Scale Statistics 表，分别确定项目和总量表的平均反应。
- 3. 写出量表的 α 系数值、均值和标准差（如果需要，每个项目的均值和标准差也可以显示在表格里）。

练习

1. 一个研究人员负责 7 个项目的居民满意度为 20 的养老院的测量，他们的反应显示在图表 4—12 中。对于测量的每个项目，1 对应强烈不同意，7 对应强烈同意。量表里分数越高表明居民对养老院满意度越高。

图表 4—12 居民满意量表的 7 个项目

item1	item2	item3	item4	item5	item6	item7
4	3	5	4	4	3	4
1	2	1	1	5	1	2
4	5	2	3	1	1	1
4	3	2	1	1	4	2
6	5	4	2	2	3	1
4	2	6	6	4	2	3
4	2	1	1	5	4	5
4	3	3	2	5	4	6
4	1	1	2	4	4	2
5	4	6	4	5	7	7
5	6	4	2	2	4	7
2	4	6	5	7	7	5
2	4	6	5	4	4	7
1	1	4	2	2	4	2

续前表

item1	item2	item3	item4	item5	item6	item7
2	3	4	4	5	4	2
4	6	2	3	5	5	4
2	4	3	4	2	3	4
2	4	5	6	4	4	5
2	4	6	6	5	4	5
6	4	5	5	6	4	5

在 SPSS 中录入数据，进行适当分析以回答下列问题。变量分别命名为 item1 ~ item7。

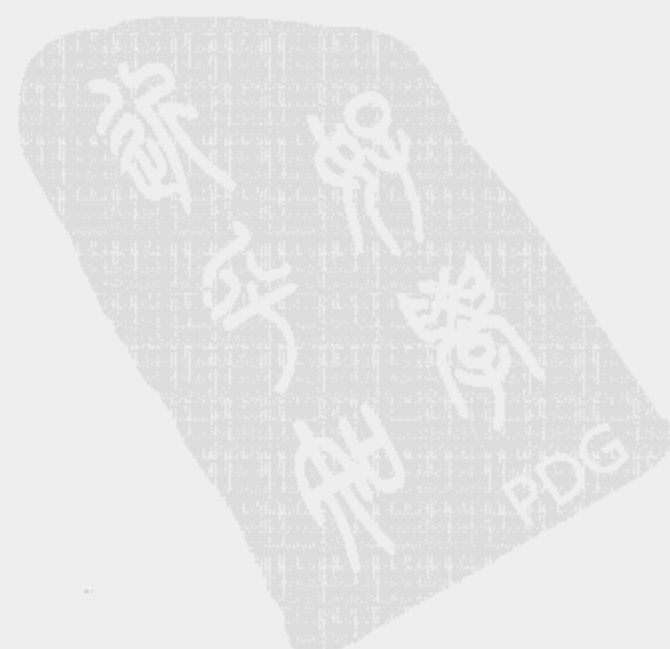
- a. 量表可靠性 (α 系数) 是多少?
- b. 怎样描述可靠性?
- c. 写下你的结果，表明量表的整体可靠性，还要包括总量表的均值和标准差。

2. 30 个居民在 10 个项目的恢复力测量中的反应显示在网站 www.prenhall.com/yockey 上的文件夹 Chapter 4 中的 Chapter 4 _ Exercise 2. sav 之中 (变量分别命名为 **resilience1** (恢复力 1) 到 **resilience10** (恢复力 10))。关于测量项目的反应取值范围是从 1 (强烈不同意) 到 7 (强烈同意)，分值越大表明恢复力越强。

在 SPSS 中打开文件并进行适当分析以回答下列问题。

- a. 量表可靠性 (α 系数) 是多少?
- b. 怎样描述可靠性?
- c. 写下你的结果，表明量表的整体可靠性，还要包括总量表的均值和标准差。

3. 一名学生对一个研究项目使用了两个不同的量表，一个量表测量自信，另一个量表测量手工灵巧。他计算了联合测量的 α 系数，发现总体可靠系数是 0.41。为什么可靠性这么低?



Part 2

第二部分

推断统计学

第二部分是本书的核心部分，包括许多常用的统计推断方法。第二部分包括以下各章：单样本 t 检验（第 5 章），独立样本 t 检验（第 6 章），相依样本 t 检验（第 7 章），一维组间方差分析（第 8 章），二维组间方差分析（第 9 章），一维组内方差分析（第 10 章），组间组内方差分析（第 11 章），皮尔逊 r 相关系数（第 12 章），简单线性回归（第 13 章），多元线性回归（第 14 章）， χ^2 拟合优度检验（第 15 章）和 χ^2 独立性检验（第 16 章）。

在开始第 5 章（单样本 t 检验）之前，回顾与假设检验有关的一些术语是有益的。如果你是初次接触统计，查阅一本统计入门教材中关于假设检验的那一章是很有帮助的。

样本与总体

在统计中，我们经常谈到样本和总体。一个**总体**（population）包括感兴趣事件的全部，而一个**样本**（sample）是总体的一小部分（或子集）。假设我们想对美国的手机使用情况进行研究，从美国的各个城市中随机选择了 200 个手机使用者。美国所有拥有手机的人构成了总体，而选择进行研究的 200 个人就构成了样本。在实践中，由于总体要么成本太高，要么太难以获得，因此大多数时候我们使用样本进行分析。

推断统计学

推断统计学（inferential statistics）是使用样本归纳总体的一种统计方法。例如，想预测总统选举中的胜者。对于相对小的样本

(例如, 2 000 个人), 使用推断统计学得到的对所有美国投票者的总体陈述具有相当高的准确度 (假设从总体中选择的样本对总体具有代表性)。推断统计学非常有用, 因为它允许我们基于有限的信息 (样本) 对总体得出结论。

假设检验

当我们使用样本对总体进行推断时, 这个过程为假设检验 (hypothesis testing)。在假设检验中, 通常要陈述两个假设: **原假设 (null hypothesis)** 和 **对立假设 (alternative hypothesis)**。原假设通常陈述处理没有效果, 而对立假设陈述处理有效果。我们看一个例子。假设要评估一种新的治疗方法以确定它能否帮助人们减轻抑郁。原假设就是治疗方法对抑郁没有效果 (它不能改变人们的感受), 而对立假设则是治疗方法对抑郁有效果 (它能改变人们的感受)。在假设检验中, 我们要么拒绝原假设, 要么不拒绝原假设。如果基于研究的结果, 原假设看起来是合理的, 我们不能拒绝它, 但是, 如果原假设不合理, 我们就拒绝它 (并且接受对立假设)。

单边检验和双边检验

可以使用**单边检验 (one-tailed test)** 或者**双边检验 (two-tailed test)** 对假设进行评价。在评估处理要看是否对任一方向有影响 (了解得分是更高还是更低) 时使用双边检验, 而在目的仅仅是调查单一方向 (仅仅是更高或者仅仅是更低) 时使用单边检验。看一个例子, 调查两种疗法确定哪一种对帮助人们克服恐惧更有效 (我们把两种疗法称为疗法 A 和疗法 B)。双边检验将考虑到任一疗法更有效的可能性 ($A < B$ 或者 $B < A$), 而单边检验要事先确定单一的方向进行调查 (比如, $A < B$)。单边检验的优势在于有更大的机会发现有效 (假定有效与假设的方向一致); 单边假设的不利之处在于如果有效与期望的方向相反, 就不能做出断言。实际上, 研究中常用的是双边检验, 基于这个原因, 在本书中将强调该检验。

第一类错误和第二类错误

回忆在假设检验中, 使用样本对总体进行推断。因为样本是总体的不完整的“图像”, 所以在假设检验过程中就可能犯错误。有两类错误会发生: **第一类错误 (Type I error)** 和 **第二类错误 (Type II error)**。如果在原假设是真

实的情况下拒绝了原假设（如果它是真实的，就不能被拒绝），就发生了第一类错误。如果在原假设是错误的情况下没有拒绝原假设（如果它是错误的，就应该被拒绝），就发生了第二类错误。这都是假设检验中的错误，因为从假设检验中得到的结论是与真实情况相反的。

功效

尽管在假设检验中会犯第一类错误和第二类错误，但是功效是与做正确的决策相关的。**功效（power）**等于原假设错误时拒绝原假设的概率（如果原假设是错误的，它也被拒绝了，就是做了一个正确的决策）。功效的取值范围是0~1，数值越大，功效就越大。例如，功效为0.8意味着在进行研究之前，有80%的机会拒绝原假设（换句话说，如果研究进行了许多次，功效为0.8就意味着有80%的机会拒绝原假设——正确的决策，20%的机会不能拒绝原假设——第二类错误）。

抽样误差

抽样误差是推断统计学中一个非常重要的概念，如果能理解它，将使得下面的假设检验的逻辑变得更简单。考虑一个比较记忆单词的两种策略的问题。假设两种策略之间没有差异，使用哪一种策略都可以，记住的单词数量在总体上是相等的。假定选择了10个人的样本，给他们第一种策略，再选择另一个包括10个人的样本，给他们第二种策略，然后比较每一组记住的单词数量。尽管两种策略可以使记住的单词数量在总体上相等，但是极少出现两个组记住的单词数量在样本上相等的情况。这是因为样本是总体的子集，仅仅是从总体中抽取的一部分，所以表现不会那么完美。一般来说，样本越小，样本与总体的差异就越大。样本与总体的差异就是**抽样误差（sampling error）**。这是统计中的正常部分，重要的是在后面的章节中记住这一点。当从总体中抽取不同的样本时，这些样本通常是不相同的（例如，它们没有相同的均值）。

p -值

继续前面单词策略的例子，如果从总体中抽取的样本通常是不相同的，那么我们如何确定这些样本之间存在有意义的差异（例如，一个策略总体确

实比其他策略更好)还是由于抽样误差导致差异(例如,这些策略总体是相同的)?可以基于从SPSS的输出结果中的 p 值进行决策。 p 值表明在原假设为真时获得特定结果(或更极端结果)的概率。例如,在前面讨论的记忆单词策略问题中,让我们假定两个策略的两个单词记住数量对应的 p 值为0.03。0.03的 p 值表明如果原假设为真,仅有3%的机会得到这些组的单词记住数量之间有差异。在假设检验中,检验的 p 值是和预先确定的数值进行比较,该值用符合表示。基于比较的结果,对原假设进行决策。在本书中,我们以0.05为水平(这是在社会和行为科学中常用的数值),来评价 p 值。

评价 p 值的过程如下:
如果 p 值小于或等于0.05 (α), 拒绝原假设(假定策略之间有差异)。
如果 p 值大于0.05 (α), 不能拒绝原假设(没有假定策略之间有差异)。
在SPSS中, p 值报告为“sig.”。
假设检验的决策过程概况在下面的表中。

p 值 (SPSS 中为 “sig.”)	决策
如果 $p \leq 0.05$	拒绝原假设
如果 $p > 0.05$	不能拒绝原假设

考虑下表中的例子。第一个例子表明 p 值为0.020,对于0.05的 α 水平,对于原假设可以做什么样的决策呢?因为0.020小于或者等于0.05,所以拒绝原假设。在例2和例3中,原假设不能被拒绝,因为两个 p 值(0.080和0.521)都大于0.05。

例子	p 值 (SPSS 中为 “sig.”)	决策 (α 水平为 0.05)	解释
1	0.020	拒绝原假设	支持对立假设
2	0.080	不能拒绝原假设	支持原假设
3	0.521	不能拒绝原假设	支持原假设

本书中使用的 α 水平为0.05,如果你愿意,可以使用另外的值(比如,0.01或者0.001),你所要做的就是调整你的决策规则。例如,使用0.01的 α 水平,前面表中的三个 p 值将导致原假设不能被拒绝,因为所有的值都大于0.01。正如你在例子中所看到的,不同的 α 值(0.05或0.01)可以对原假设得到不同的结论(无论是拒绝还是不拒绝)。

对本部分的每一章，我们将检查 p 值，并且使用上面详细介绍的决策过程，我们要么拒绝，要么不拒绝原假设。如果拒绝原假设，结果通常称为“显著”或“统计显著”；如果不拒绝原假设，结果称为“不显著”。

效应量

尽管假设检验是对总体进行推断的一个强大的工具，但重要的是知道假设检验能够告诉我们什么，不能告诉我们什么。再次考虑我们的单词策略问题，假设检验允许我们在合理的保证程度上得到结论，即无论如何两个策略总体上有差异，然而，假设检验未能说明这些组有何不同（假设检验表明这些组有差异，但是没有标明差异是小、中等还是大）。可以计算效应量来描述这些组之间的差异程度。**效应量**（**effect sizes**）表明我们研究的结果的大小。我们将在第 5 章进一步描述效应量，将在第二部分描述的几种不同的方法中对它加以举例说明。

至此结束了对第二部分的介绍。接下来我们将在第 5 章开始介绍单样本 t 检验。



单样本 t 检验

当把一个样本均值和一些已知或估计的总体均值进行比较时，我们可以使用单样本 t 检验。从某种意义上说，这些样本通常不同于总体，而且感兴趣的问题是：感兴趣的因变量的样本均值是否显著不同于总体均值。下面给出一个单样本 t 检验的例子。

例子

一家知名会计师事务所雇用的一名工业心理学家想知道这家公司每周的平均工作小时数是否显著不同于每周 52 个小时的国家平均水平。工业心理学家随机抽取来自公司不同部门的 16 个人，记录他们在三个月期间的每周平均工作小时数。该研究中的因变量是每周工作小时数。

单样本 t 检验的目标和数据要求

单样本 t 检验		
目标	数据要求	例子
确定样本均值是否显著不同于某个已知或估计的总体均值	参与者的一个样本 因变量 连续	知名会计师事务所中的会计师们 因变量 • 每周工作小时数

原假设和对立假设

原假设 (H_0) 表明会计师事务所中每周工作小时数等于全国平均水平 52 小时:

$$H_0: \mu=52$$

对立假设 (H_1) 表明会计师事务所中每周工作小时数不等于全国平均水平 52 小时:

$$H_1: \mu\neq 52$$

原假设评价

单样本 t 检验对会计师事务所中每周工作小时数等于全国平均水平 52 小时的原假设进行了检验。如果 t 检验产生的结果在原假设正确时看起来不可能 (结果发生的可能性小于 5%), 那么拒绝原假设。如果 t 检验产生的结果在原假设正确时看起来正确 (结果发生的可能性大于 5%), 那么不拒绝原假设。

研究问题

在研究中基本的兴趣问题也可以用研究问题的方式表示, 例如,

“知名会计师事务所中的雇员每周工作小时数不同于国家平均水平 52 小时吗?”

数据

在图表 5—1 中列出了 16 名参与者的数据。

图表 5—1 知名会计师事务所中的 16 名雇员的每周工作小时数

Participant	Hours worked per week	Participant	Hours worked per week
1	54	4	53
2	48	5	60
3	68	6	45

续前表

Participant	Hours worked per week	Participant	Hours worked per week
7	57	12	63
8	62	13	68
9	71	14	64
10	60	15	56
11	55	16	60

说明：变量 participant 包含在数据中，但不用输入 SPSS。

在 SPSS 中输入数据及分析

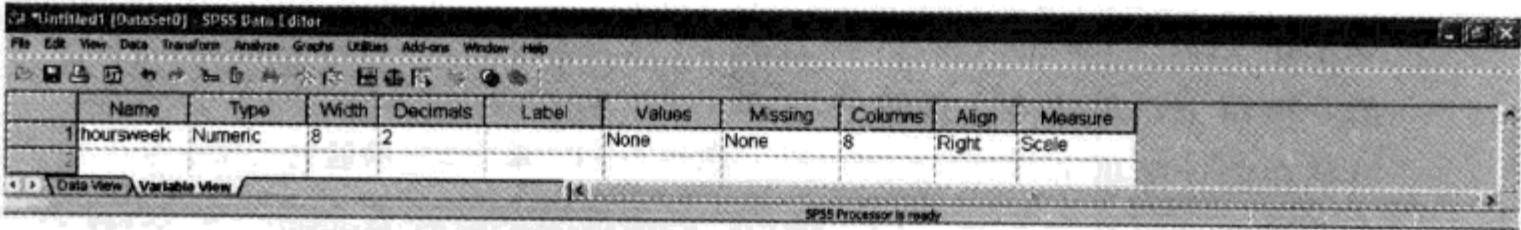
下面的步骤 1 和步骤 2 描述了怎样在 SPSS 中输入数据。数据在 www.prenhall.com/yockey 上的文件夹 Chapter 5 中的文件 hours worked. sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中可以为雇员在会计师事务所中的每周工作小时数生成一个变量，该变量被命名为 **hoursweek**。

- 3. 输入名字 **hoursweek**，按向下箭头键（↓）。详见图表 5—2。



图表 5—2 输入变量 hoursweek 的 Variable View 窗口

步骤 2：输入数据

- 1. 点击 Data View 标签。变量 **hoursweek** 出现在 Data View 窗口的第 1 列。
- 2. 为输入变量 **hoursweek** 的数据，确认窗口左上角的第一个单元是激活的

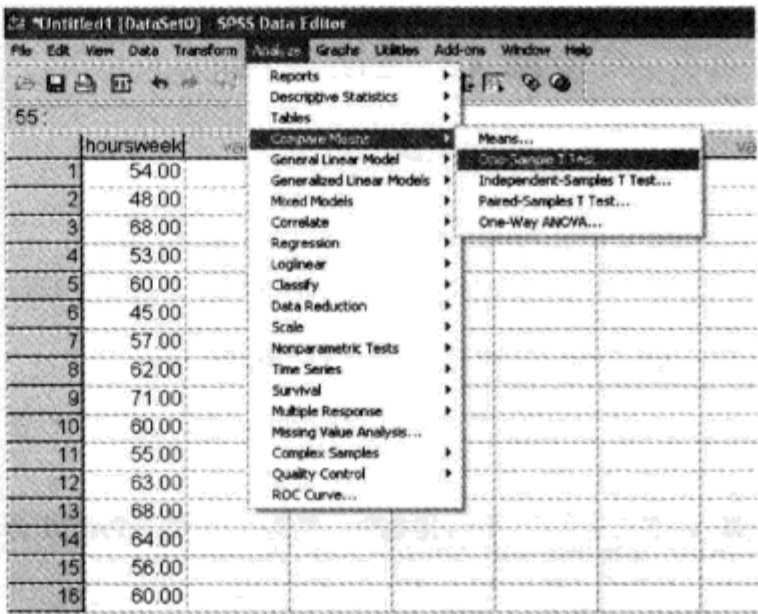
(如果不是激活的, 点击它)。输入每个参与者的工作小时数, 以第一个参与者的值 (54) 开始, 直到输入最后一个参与者的值 (60) (输入一个参与者的值后, 按向下箭头 (↓) 或回车按钮, 以便移动到 Data View 窗口的下一行)。图表 5—3 中给出了完整的数据文件。

步骤 3: 分析数据

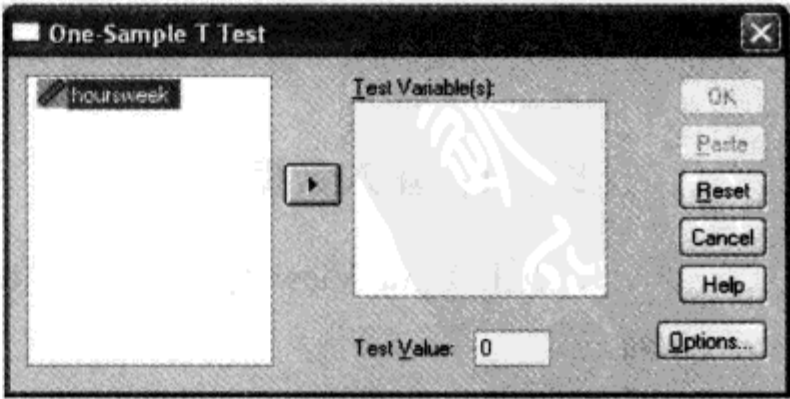
- 1. 从菜单栏中选择 **Analyze > Compare Means > One-Sample T test...** (见图表 5—4)。
- 2. 打开 One-Sample T Test 对话框, 变量 **hoursweek** 出现在对话框的左边 (见图表 5—5)。
- 3. 选择因变量 **hoursweek**, 点击向右箭头按钮 (▶), 把变量移到 Test Variable(s) (检验变量) 框。
- 4. 在 Test Value 框中输入 52, 这是原假设中的指定值, 详见图表 5—6。(注意: 这一步很关键, 经常被忽略。如果在 Test Value 框中默认值 0 被留下, 单样本 *t* 检验就是在会计师事务所中雇员的工作小时数是否显著不同于每周 0 小时。(那是理所当然的!) 当在 SPSS 中实施一个单样本 *t* 检验时, 确保在 Test Value 框中输入原假设中的指定值 (在本例中为 52)。)
- 5. 点击 “OK”。

hoursweek	
1	54.00
2	48.00
3	68.00
4	53.00
5	60.00
6	45.00
7	57.00
8	62.00
9	71.00
10	60.00
11	55.00
12	63.00
13	68.00
14	64.00
15	56.00
16	60.00

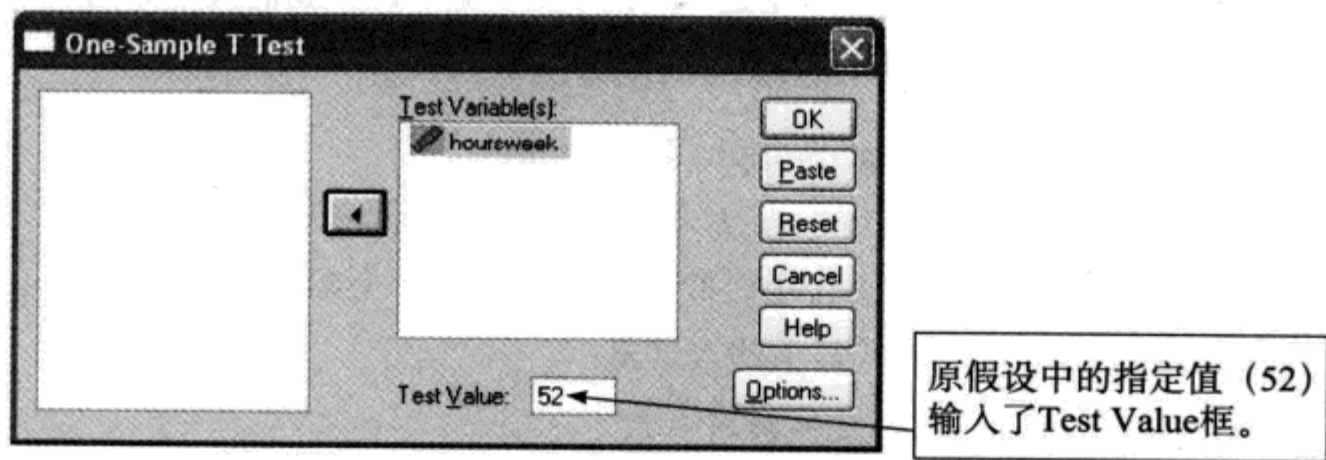
图表 5—3 单样本 *t* 检验例子中的完整数据文件



图表 5—4 单样本 *t* 检验的菜单命令



图表 5—5 单样本 *t* 检验的对话框



图表 5—6 单样本 t 检验的对话框 (续)

在 SPSS 中运行单样本 t 检验程序，结果显示在 Viewer 窗口中。

步骤 4：解释结果

单样本 t 检验的输出结果显示在图表 5—7 中。

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
hoursweek	16	59.0000	7.14609	1.78652

One-Sample Test						
Test Value=52						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
hoursweek	3.918	15	0.001	7.00000	3.1921	10.8079

均值差7等于在知名会计师事务所中每周平均工作小时数 (59) 与原假设中 (会计师总体) 的指定值 52的差。

因为p-值0.001小于0.05，因此拒绝原假设。

图表 5—7 单样本 t 检验的输出结果

单样本统计量表 (One-Sample Statistics Table)

输出的第一个表格 One-Sample Statistics 显示了样本量 (N)，平均值，标准差和样本均值的标准误差。注意在会计师事务所中每周工作小时数的平

均值是 59，它比全国平均水平的 52 小时多了 7 小时。我们接下来考虑这个差异（7 小时）对具有统计显著性而言是否足够大。

单样本检验表 (One-Sample Test Table)

下一个表格 One-Sample Test 提供了我们研究问题的答案，问题是这家会计师事务所每周的平均工作小时数是否显著不同于全国平均水平 52 小时。原假设的检验是以 t 的形式显示的，这里

$$t = \frac{\text{样本均值与总体均值之差}}{\text{均值的标准误差}}$$

从表格 One-Sample Test 中代入适当的值（均值差，均值标准误差）得到 t 值是

$$t = \frac{7}{1.786\ 52} = 3.918$$

这与 One-Sample Test 表格中 t 列的值一致。有 16 个参与者，自由度 (df) 等于 15（参与者数目-1），对应的 p 值为 0.001。因为 p 值小于 0.05，所以会计师事务所中每周工作小时数等于全国平均水平 52 小时的原假设被拒绝。因为检验是显著的，从 One-Sample Statistics 表格中检查 16 个参与者的平均值，发现会计师事务所每周的平均工作小时数（59）高于全国平均水平 52 小时。

效应量

正如第二部分引言所描述的那样，假设检验表明在群组间是否有一个显著的不同（回答是或不是的问题），而效应量提供了对结果的“大小”的表示（回答了一个什么程度的问题）。在当前的例子中，使用假设检验得到的结论是当地会计师事务所雇员的平均工作小时数与全国平均水平存在差异。我们将要计算的效应量能表达这个差异有多大。

单样本 t 检验的常用效应量统计量是 Cohen's d ，这里

$$d = \frac{\text{样本均值与总体均值之差}}{\text{标准差}}$$

从图表 5—7 代入适当的值（均值差，标准差）得到 d 值为

$$d = \frac{7}{7.15} = 0.98$$

雅各布·科恩 (Jacob Cohen, 1988) 是一位推动效应量度量的先驱, 他提供了在实践中规定什么构成小、中和大的效应量的准则。当科恩的准则被广泛应用时 (也用在本书里), 应该注意它们仅仅是近似 (科恩强调的一种看法), 在不同的学科中是变化的。

科恩准则规定单样本 t 检验的小、中和大的效应量分别对应着 0.20、0.50 和 0.80。这些值表明可以用标准差表示样本均值和总体均值的差异有多大。0.20 表示组别之间有 $1/5$ 的标准差, 0.50 表示群组间有一半的标准差, 0.80 表示群组间有 $8/10$ 的标准差。利用科恩的结论, d 值 0.98 对应实践中一个大的效应, 表明雇员在当地会计师事务所与国家规定的工作时间的差异基本上是 1 倍标准差 (雇员在会计师事务所的工作时间更长)。¹

APA 格式的结果表达

对于第二部分的每一章, 将以美国心理协会 (APA) 的格式提供书面结果。在美国心理协会的出版手册 (2001) 上可以找到 APA 格式特定的规则。

在写结果时, 除了报告样本均值和样本标准差, 还要报告假设检验的结论, 自由度, t 值, p -值和效应量。接下来我们将展示一个 APA 格式的简单书面例子。

在知名会计师事务所的雇员 ($M=59.0$, $SD=7.15$) 每周工作时间显著多于 52 个小时的全国平均水平, $t(15) = 3.92$, $p < 0.05$, $d = 0.98$ 。

单样本 t 检验的假定

第二部分的几章, 将介绍每个统计方法的假定条件。假定是很重要的, 因为如果不满足假定条件, 特定的统计方法的结果就不具可信性 (p -值不准确)。然而, 一个检验是否由于违反假定条件而无效取决于违反的特定假定 (与其他假定相比, 一些假定更糟糕而且更易违反) 和假定条件不满足的程度。下面将描述单样本 t 检验的假定。

1. 观测是独立的。

该假定能通过设计你的研究使得参与者在任何情况下都不相互影响得到满足。违反这个假定可能会严重破坏单样本 t 检验的准确性。如果我们有理由

相信已经违反了独立性假定，就不应该再用单样本 t 检验了。

2. 因变量总体服从正态分布。

该假定意味着每周的工作小时数的总体服从正态分布。对于中等到较大的样本量 ($N \geq 30$)，绝大部分的非正态分布趋向于对 t 检验的精确性没有多少影响，尽管一些非正态分布可能不利于 t 检验的功效。对非正态数据的一种修正办法就是对数据进行非线性变换。然而，非线性变换超出了本书的范围（有兴趣的读者可以参考 Tabachnick and Fidell (2007) 获得更多非线性变换的信息）。

在 SPSS 中执行单样本 t 检验的步骤摘要

I. 数据输入和分析

1. 在 SPSS 中生成一个变量。
2. 输入数据。
3. 从菜单栏中选择 **Analyze > Compare Means > One-Sample T Test...**
4. 把因变量移到 Test Variable (s) 框。
5. 在 Test Value 框输入原假设中的指定值。
6. 点击 “OK”。

II. 结果解释

1. 检查 One-Sample Test 表中的 p -值 (“sig.” (2-tailed))。
 - 如果 $p \leq 0.05$ ，拒绝原假设。比较样本均值和总体均值，表明样本均值是大于还是小于总体均值，写下结果。
 - 如果 $p > 0.05$ ，不拒绝原假设。表明样本均值和总体均值间没有显著差异，写下结果。

练习

1. 在研究项目中，一个学生想检验宣称有能力选择获胜足球队的人在挑选优胜球队时是否与一般人的水平不同（选择获胜者的正确率或高于或低于 50%）。她选出 15 个声称有能力选择获胜者的人并记录下每个人在整个足球赛季的选择正确的百分比。这 15 名预言者选择正确次数的百分率显示在图表

5—8 中。

图表 5—8 15 名预言者选择正确次数的百分率

Prognosticator	Percentage of correct picks	Prognosticator	Percentage of correct picks
1	45	9	53
2	46	10	51
3	47	11	52
4	52	12	50
5	51	13	48
6	43	14	47
7	47	15	51
8	38		

在 SPSS 中输入数据，进行必要分析以回答下列问题（变量命名 **success**（成功））。（提示：对于这个问题，用 50 作为 Test Value，录入数据时不包括百分号。）

- a. 陈述原假设和对立假设。
 - b. 陈述数据的一个研究问题。
 - c. 预言者选择获胜者的正确率不同于 50% 吗？在 $\alpha=0.05$ 时检验。
 - d. 效应量是多少？你能区分效应量是小、中还是大吗？
 - e. 将研究结果写成适当的 APA 格式。
2. 在全美国四年级学生总体中，标准化数学能力考试的平均分数是 50 分（考试中更高的分数显示更强的能力）。在当地校区使用了一个新的数学技巧培训项目，行政官员负责评估在新项目中孩子的表现是否不同于全国平均水平（可取表现高于全国平均水平，但是两种可能性都要检验）。25 名四年级学生被指导使用了数学技巧并进行了标准化的数学测试。学生的考试成绩展示在网站 www.prenhall.com/yockey 上的文件夹 Chapter 5 中的文件 Chapter 5 _ Exercise 2. sav（变量名为 **mathscore**）。在 SPSS 中输入数据，进行必要分析回答下列问题。
- a. 陈述原假设和对立假设。
 - b. 陈述数据的一个研究问题。
 - c. 当地校区四年级学生的数学考试成绩不同于全国平均水平吗？在 $\alpha=0.05$ 时检验。

- d. 效应量是多少？你能区分效应量是小、中还是大吗？
 - e. 将研究结果写成适当的 APA 格式。
3. 研究者想调查学生是支持还是反对总统候选人。对 50 名学生进行了询问，以确定下面哪一个声明最好地反映了他们对候选人的态度。

声明	数据文件中的数值
我非常支持这个候选人	1
我有点儿支持这个候选人	2
我既不支持也不反对这个候选人	3
我有点儿反对这个候选人	4
我非常反对这个候选人	5

50 个参与者的反应可以从网站上文件夹 Chapter 5 中的文件 Chapter 5 _ Exercise 3. sav（变量命名为 **score**）中获得。在 SPSS 中打开文件，进行必要分析以回答下列问题。（提示：为了调查学生们对于候选人的态度，检验看看他们的态度是否不同于中立反应——这意味着 Test Value 框里的值是 3。）

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 学生是支持还是反对总统候选人呢？在 $\alpha=0.05$ 时检验。
- d. 效应量是多少？你能区分效应量是小、中还是大吗？
- e. 将研究结果写成适当的 APA 格式。



独立样本 t 检验

当对两个独立分组中感兴趣的一个连续因变量的均值进行比较时，我们可以使用独立样本 t 检验。下面给出一个独立样本 t 检验的例子。

例子

临床心理学家想调查认知行为治疗和精神分析治疗对抑郁症的相对有效性。30 名患有抑郁症的病人随机分配接受两个疗法，其中 15 人接受行为治疗，另外 15 人接受精神分析治疗。经过两个月的治疗后，记录下每个病人抑郁症得分。在本研究中，自变量是治疗方法（认知行为治疗与精神分析治疗），因变量是抑郁症，较高的分数表示更高的抑郁水平（抑郁水平的分数变化范围为 10~70）。

独立样本 t 检验的目标和数据要求

独立样本 t 检验		
目标	数据要求	例子
检验两个组别中关于某些感兴趣的因变量的均值是否存在显著差异	具有两个不同组别的一个自变量 一个连续因变量	自变量 • 治疗类型（认知行为治疗与精神分析治疗） 因变量 • 抑郁症

原假设和对立假设

原假设指出两组的抑郁症分数均值在总体上是相等的：

$H_0: \mu_{\text{精神分析}} = \mu_{\text{认知行为}}$

对立假设指出两组的抑郁症分数均值在总体上是不等的：

$H_1: \mu_{\text{精神分析}} \neq \mu_{\text{认知行为}}$

原假设评价

独立样本 *t* 检验对两组的总体均值相等的原假设进行了检验。如果 *t* 检验产生的结果在原假设正确时看起来不可能（结果发生的可能性小于 5%），那么拒绝原假设。如果 *t* 检验产生的结果在原假设正确时看起来正确（结果发生的可能性大于 5%），那么不拒绝原假设。

研究问题

在研究中基本的兴趣问题也可以用研究问题的方式表示，例如，

“在接受认知行为治疗与精神分析治疗的病人中，抑郁症水平的均值是否存在差异呢？”

数据

在图表 6—1 中列出了 30 个参与者的数据。接受精神分析治疗的参与者标记为“1”，接受认知行为治疗的标记为“2”。

图表 6—1 独立样本 *t* 检验例子的数据

Participant	Therapy	Depression	Participant	Therapy	Depression
1	1	57	8	1	62
2	1	61	9	1	41
3	1	67	10	1	36
4	1	63	11	1	55
5	1	51	12	1	57
6	1	55	13	1	70
7	1	45	14	1	62

续前表

Participant	Therapy	Depression	Participant	Therapy	Depression
15	1	58	23	2	49
16	2	47	24	2	61
17	2	42	25	2	43
18	2	59	26	2	47
19	2	37	27	2	49
20	2	35	28	2	37
21	2	42	29	2	41
22	2	38	30	2	48

说明：变量 participant 包含在数据中，但不用输入 SPSS。

在 SPSS 中输入数据及分析

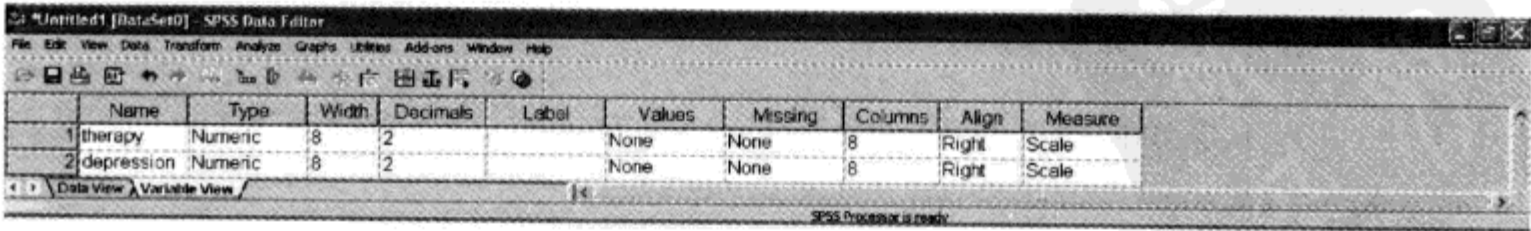
下面的步骤 1 和步骤 2 描述了如何在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 6 中的文件 therapy.sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中将生成两个变量，一个是不同治疗方法的组别（自变量），另一个是抑郁症分数（因变量）。这些变量将各自被命名为 **therapy**（治疗方法）和 **depression**（抑郁症）。

- 3. 在 Variable View 窗口前两行分别输入变量名称 **therapy** 和 **depression**（详见图表 6—2）。



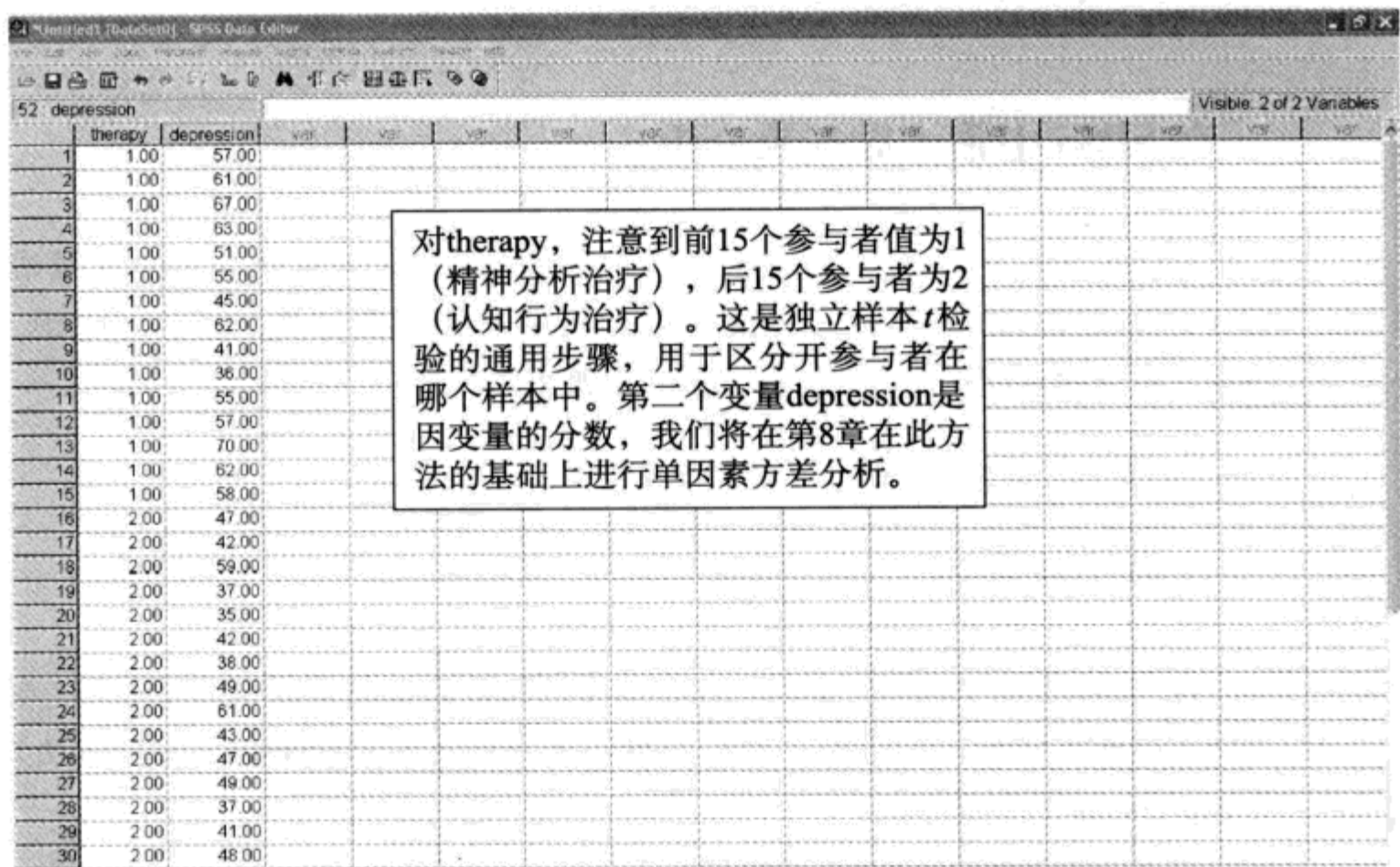
图表 6—2 输入变量 therapy 和 depression 的 Variable View 窗口

- 4. 使用第 1 章描述的过程，为变量 **therapy** 建立变量值标签，1=“精神

分析治疗”，2= “认知行为治疗”。

步骤 2：输入数据

- 1. 点击 Data View 标签。变量 **therapy** 和 **depression** 出现在 Data View 窗口的前两列。
- 2. 参照图表 6—1，为每个参与者输入两个变量的数据。对第一个参与者，为变量 **therapy** 和 **depression** 分别输入数值 1 和 57。依次输入全部 30 个参与者的数据。图表 6—3 中给出了完整的数据文件。



对therapy，注意到前15个参与者值为1（精神分析治疗），后15个参与者为2（认知行为治疗）。这是独立样本t检验的通用步骤，用于区分参与者在哪个样本中。第二个变量depression是因变量的分数，我们将在第8章在此方法的基础上进行单因素方差分析。

	therapy	depression
1	1.00	57.00
2	1.00	61.00
3	1.00	67.00
4	1.00	63.00
5	1.00	51.00
6	1.00	55.00
7	1.00	45.00
8	1.00	62.00
9	1.00	41.00
10	1.00	36.00
11	1.00	55.00
12	1.00	57.00
13	1.00	70.00
14	1.00	62.00
15	1.00	58.00
16	2.00	47.00
17	2.00	42.00
18	2.00	59.00
19	2.00	37.00
20	2.00	35.00
21	2.00	42.00
22	2.00	38.00
23	2.00	49.00
24	2.00	61.00
25	2.00	43.00
26	2.00	47.00
27	2.00	49.00
28	2.00	37.00
29	2.00	41.00
30	2.00	48.00

图表 6—3 独立样本 t 检验例子中的完整数据文件

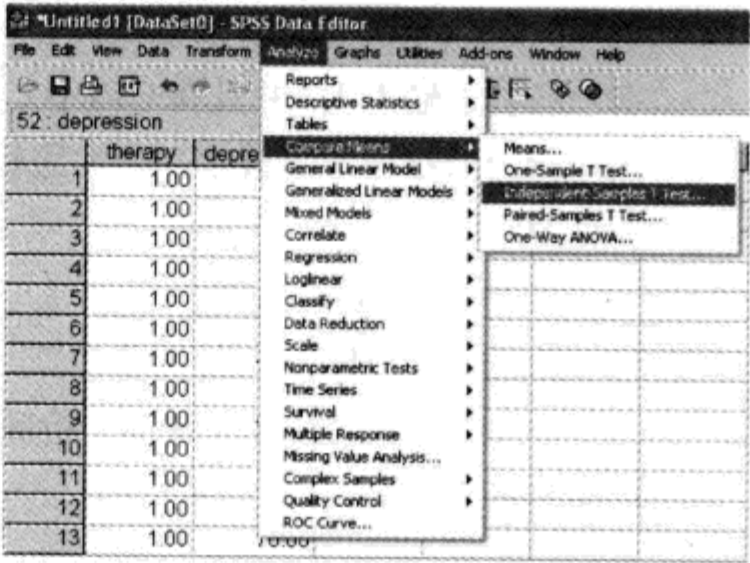
步骤 3：分析数据

- 1. 从菜单栏中选择 **Analyze>Compare Means>One-Sample T Test...**（见图表 6—4）。
打开 Independent-Samples T Test 对话框，变量 **therapy** 和 **depression** 出现在对话框的左边（见图表 6—5）。
- 2. 选择因变量 **depression**，点击向右箭头按钮（▶），把变量移到 Test

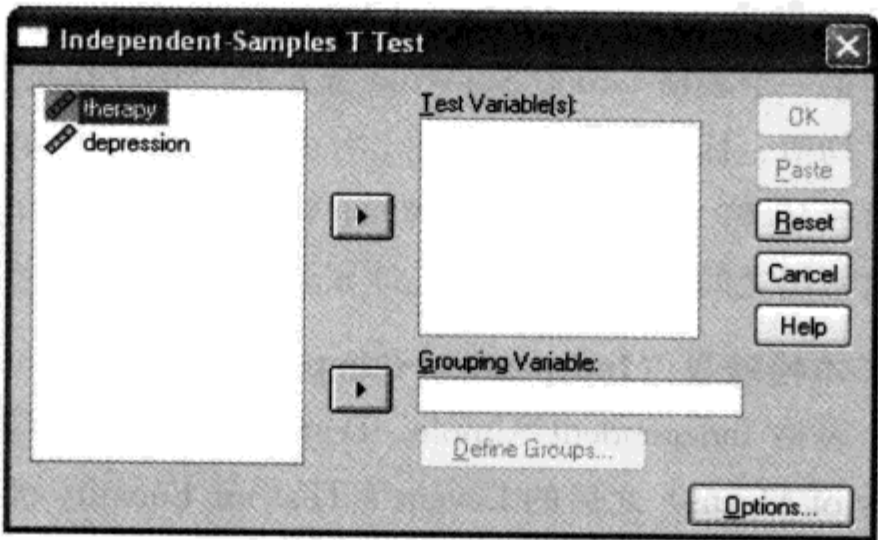
Variable (s) 框。

3. 选择自变量 **therapy**，点击向右箭头按钮 (►)，把变量移到 Grouping Variable (分组变量) 框中。

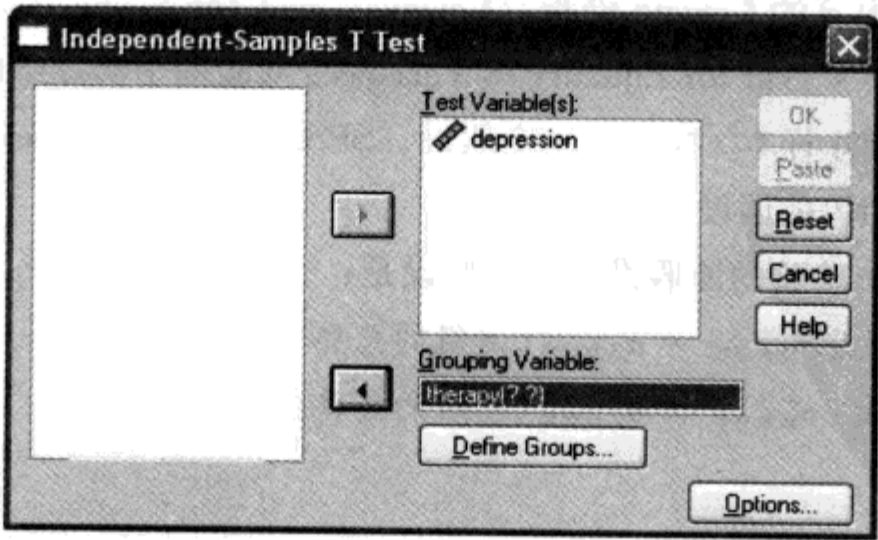
在 Grouping Variable 框中，两个在括号内的问号出现在 **therapy** 的右边 (见图表 6—6)。这些问号表示原先的数字分配到两个治疗样本中 (也就是 1, 2)，这些数字需要通过点击“Define Groups”来输入。



图表 6—4 独立样本 t 检验的菜单命令



图表 6—5 独立样本 t 检验对话框



图表 6—6 独立样本 t 检验对话框 (续)

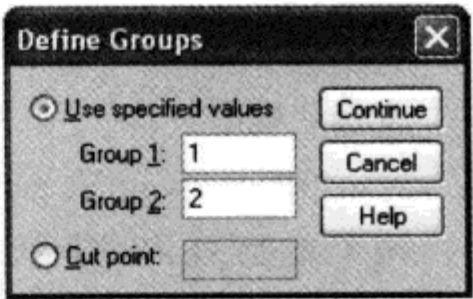
4. 点击 “Define Groups”。

5. Define Groups 对话框被打开，在 Group 1（表示精神分析治疗样本的数字）的右边输入 1，并且在 Group 2（表示认知行为治疗样本的数字）的右边输入 2。详见图表 6—7。

6. 点击 “Continue”。

7. 点击 “OK”。

在 SPSS 中运行独立样本 t 检验程序，结果显示在 Viewer 窗口中。



图表 6—7 Define Groups 对话框

步骤 4：解释结果

独立样本 t 检验的输出结果显示在图表 6—8 中。

分组统计量表 (Group Statistics Table)

输出的第一个表格 Group Statistics 显示每个治疗组的描述统计量，包括样本量、平均值、标准差和标准误差。注意到认知行为治疗样本的抑郁分数均值（均值=45.00）比精神分析治疗样本（均值=56.00）的低。我们稍后将会考虑这两组之间的差异对具有统计显著性而言是否足够大。

独立样本检验表 (Independent-Samples Test Table)

第二个表格 Independent-Samples Test 显示在均值相等的 t 检验 (t-test for Equality of Means) 之后的 Levene’s Test for Equality of Variances 的结果。这些检验会在后面单独讨论。

方差相等的 Levene 检验 (Levene’s Test for Equality of Variances)

Levene’s Test for Equality of Variances 检验两个治疗组的总体方差是否相等，这是独立样本 t 检验的一个假设。SPSS 使用一个由 Levene 开发的方法来检验总体相等的假设。

Levene 检验的原假设和对立假设是：

$$H_0: \sigma^2_{\text{精神分析}} = \sigma^2_{\text{认知行为}} \text{ (两组的总体方差相等)}$$

$$H_1: \sigma^2_{\text{精神分析}} \neq \sigma^2_{\text{认知行为}} \text{ (两组的总体方差不相等)}$$

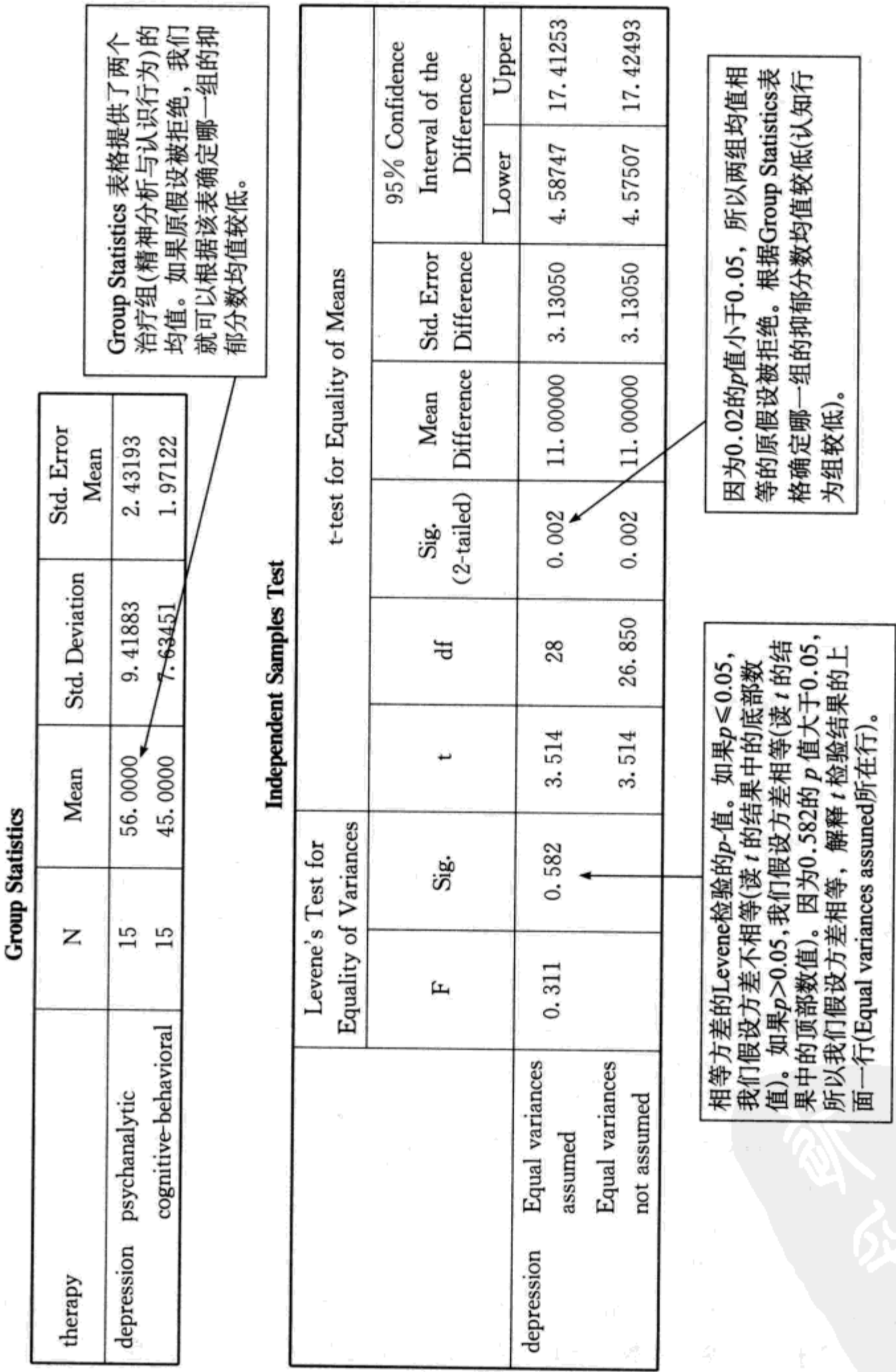


图 表 6—8 独立样本t检验的输出结果

通过检查输出结果中 Levene's Test for Equality of Variances 的检验报告的 p 值 (sig.) 来评价方差相等的假设。如果 $p \leq 0.05$, 拒绝原假设, 表明总体方差不相等; 如果 $p > 0.05$, 不拒绝原假设, 表明总体方差相等。

通过 Independent Samples Test 表格我们看到, Levene 检验得到 F 值为 0.311 (F 检验将在第 8 章讨论) 和 p 值为 0.582。因为 0.582 的 p 值大于 0.05, 接受方差相等的原假设。因此, 基于 Levene 的检验结果, 我们推断两个样本的总体方差相等。评价相等方差假设的决策规则概述于图表 6—9 中。

图表 6—9 评价方差相等假设的决策规则

评价方差相等的假设		
Levene 检验结果	决策	结论
$p > 0.05$	不能拒绝 H_0	假设总体方差相等
$p \leq 0.05$	拒绝 H_0	假设总体方差不相等
我们的数据		
$p = 0.582; 0.582 > 0.05$	不能拒绝 H_0	假设总体方差不相等

均值相等的 t 检验 (t-test for Equality of Means)

表格的下一部分 t-test for Equality of Means 为我们研究的问题提供了答案, 即两个治疗组的抑郁程度是否不同。在 SPSS 中的两个结果里, 一个假设组间方差相等, 另一个假设组间方差不相等。由于 Levene 的检验不显著 (Levene 检验中 p 值 > 0.05), 我们将使用标有 Equal variances assumed 第一行的结果 (如果 Levene 的检验是显著的, 即 $p \leq 0.05$, 我们将使用标有 Equal variances not assumed 第二行的结果)。

均值相等的原假设的检验是以 t 的形式显示的, 这里

$$t = \frac{\text{两样本间均值之差}}{\text{两均值之差的标准误差}}$$

从表格 Independent Samples Test 代入适当的值 (均值差、均值标准误差) 得到 t 值是

$$t = \frac{11}{3.1305} = 3.514$$

这与表格 Independent Samples Test 中 t 列的值一致。有 30 个参与者, 自由度 (df) 等于 28 (参与者数目-1), 对应的 p 值为 0.002。因为 p 值小于 0.05,

所以均值相等的原假设被拒绝，结论为两个治疗组的抑郁分数有显著差异。由于检验是显著的，从 Group Statistics 表格中检查均值，发现认知行为治疗样本的抑郁分数均值（均值 = 45.00）比精神分析治疗样本（均值 = 56.00）的低（如果 t 检验不显著，我们可以认为 Group Statistics 表格中均值的任何差异都是由于抽样误差导致的）。

效应量

对独立样本 t 检验的效应量的估计通常由 d 给出，其中

$$d = t \sqrt{\frac{N_1 + N_2}{N_1 N_2}}$$

N_1 和 N_2 是样本 1 和样本 2 的样本量， t 是图表 6—8 中 t 统计给出的值。¹

从图表 6—8 代入适当的值（均值差，标准差）得到效应量为

$$d = 3.514 \sqrt{\frac{15 + 15}{15 \times 15}}$$

$$d = 1.28$$

科恩（Cohen, 1988）给出了 d 值的估计 0.2, 0.5 和 0.8 在行为科学中分别对应于小、中和大的效应量。根据科恩准则，在实践中效应量 1.28 被认为非常大，表明认知行为治疗组的抑郁分数比精神分析组低，差值为 1.28 倍标准差。

APA 格式的结果表达

在写独立样本 t 检验的结果时，除了报告样本均值和样本标准差，还要报告假设检验的结论，自由度， t 值， p -值和效应量。接下来我们将展示一个 APA 格式的简单书面例子。

接受认知行为治疗（ $M=45.00$, $SD=7.63$ ）显著低于接受精神分析治疗组的抑郁分数（ $M=56.00$, $SD=9.42$ ）。 $t(28) = 3.51$, $p < 0.05$, $d=1.28$ 。

独立样本 t 检验的假定

1. 观测是独立的。

违反这一假定将严重损坏独立样本 t 检验的准确性。这个假定可以通过设计研究中的所有参与者在任何情况下都不相互影响得到满足（参与者们共同合作回答关于抑郁标准的问题是违反独立假定的一个例子）。如果我们有理由相信违反了独立假定，就不应该使用独立样本 t 检验。

2. 每组因变量总体服从正态分布。

这个假定指的是两个治疗样本的抑郁分数的总体服从正态分布（画图时应类似于钟形曲线）。对于中等到较大的样本量，绝大部分的非正态分布趋向于对 t 检验的精确性没有多少影响，尽管一些非正态分布可能对 t 检验的功效有不利影响。

3. 每组总体方差相等。

违反方差相等假定将会损害独立样本 t 检验的准确性，尤其在样本量不相等时。解释 SPSS 中 Levene 的检验结果，从输出的结果中读出合适的结论来强调这个假定。

在 SPSS 中执行独立样本 t 检验的步骤摘要

I. 数据输入和分析

1. 在 SPSS 中生成两个变量（一个是自变量，另一个是因变量）。
2. 输入数据。
3. 为自变量创建变量值标签。在 Value Labels 对话框，适当输入数值和标签。单击“OK”。
4. 选择 **Analyze > Compare Means > Independent-Samples T Test...**
5. 将因变量移到 Test Variable (s) 框，自变量移到 Grouping Variable 框内。
6. 点击“Define Groups”。为 1 组输入 1，2 组输入 2。点击“Continue”。
7. 点击“OK”。

II. 结果解释

1. 检查 Levene 检验中的平等性差异。

● 如果 Levene 检验中 $p > 0.05$ ，假设两组具有相等的总体方差。按第一行数据解释 *t* 检验，即 Equal variances assumed。

● 如果 Levene 检验中 $p \leq 0.05$ ，不能假设两组具有相等的总体方差。按第二行数据解释 *t* 检验，即 Equal variances not assumed。

2. 检查 Independent Samples Test 表格中 t-test for Equality of Means 部分里的 *p*-值 (sig.)。

● 如果 $p \leq 0.05$ ，拒绝原假设。检验表格 Group Statistics 中的均值并给出结果中组间的性质差异。

● 如果 $p > 0.05$ ，不拒绝原假设。给出结果标志，其中组间没有显著差异。

练习

1. 最近外科技术的发展已经增加了对微创外科 (MIS) 的使用，MIS 解决了外科手术中产生一个或更多小的切口 (使用小型器械) 的问题，而传统外科手术需要切取适当大的切口 (使用更大型器械)。MIS 的一项声明称可以减少病人由外科手术带来的痛苦。为了检验这个声明，对 30 个预定接受心血管外科手术的人进行研究，其中 15 人接受 MIS，另外 15 人进行传统外科手术。记录了手术后 24 小时每一个病人的疼痛级别，见图表 6—10。疼痛程度范围为 1~10，分数越高表示疼痛指数越高。

图表 6—10 两组的术后疼痛指数

程序 (Procedure)	疼痛 (Pain)
微创手术	5, 4, 7, 2, 3, 5, 4, 6, 2, 5, 6, 8, 4, 5, 5
传统手术	6, 7, 5, 8, 9, 7, 8, 6, 6, 7, 8, 8, 6, 3, 9

在 SPSS 中输入数据，进行必要分析以回答下列问题。变量命名为 **procedure** (程序) 和 **pain** (疼痛)。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 检验方差相等的假设。数据是否表明组间的方差不相等？

- d. 组间的疼痛指数是否存在明显差异？在 $\alpha=0.05$ 时检验。
- e. 效应量是多少？你能区分效应量是小、中还是大吗？
- f. 将研究结果写成适当的 APA 格式。

2. 为了对两种不同的帮助人们战胜蛇恐惧症的方法进行比较，选取 20 个患有过度的蛇恐惧症的人，每人随机接受其中一种治疗方法：系统脱敏疗法（小而逐步地帮助战胜蛇恐惧症的步骤）或者内爆疗法（参与者过度充斥恐惧，目的是克服它）。² 经过三周的治疗，使用一种蛇恐惧测量方式来得到每个参与者对蛇紧张程度的分数（更高的分数表示对蛇更高的恐惧程度）。数据位于网站 www.prenhall.com/yockey 中文件夹 Chapter 6 中的文件 Chapter 6 _ Exercise 2. sav 之中（文件中变量名命名为 **therapy**（治疗方法）和 **snakefear**（蛇恐惧）；对于 **therapy**，1 = “系统脱敏疗法”，2 = “内爆疗法”）。在 SPSS 中打开文件，进行必要分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 检验方差相等的假设。数据是否表明组间的方差不相等？
- d. 组间的蛇恐惧指数是否存在明显差异？在 $\alpha=0.05$ 时检验。
- e. 效应量是多少？你能区分效应量是小、中还是大吗？
- f. 将研究结果写成适当的 APA 格式。

3. 一个研究员想要调查能够接触宠物的居民与不能接触宠物的居民对疗养院满意程度是否存在差别。满意程度在 5~25 范围内，高的数值代表高的满意度。60 个疗养院居民参与了调查，其中 30 人来自有宠物的家庭，30 人没有宠物。居民满意度的数据在网站上文件夹 Chapter 6 中的文件 Chapter 6 _ Exercise 3. sav 的文件之中（变量命名为 **group**（组）和 **satisfaction**（满意度）；对于 **group**，1 = “有宠物”，2 = “没有宠物”）。在 SPSS 中打开文件，进行必要分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 检验方差相等的假设。数据是否表明组间的方差不相等？
- d. 居民间的满意程度是否存在明显差异？在 $\alpha=0.05$ 时检验。
- e. 效应量是多少？你能区分效应量是小、中还是大吗？
- f. 将研究结果写成适当的 APA 格式。

相依样本 t 检验

在对某种程度上相关的两个样本的均值进行比较时，我们可以使用相依样本 t 检验（也称为配对样本 t 检验，重复测量 t 检验，匹配样本 t 检验等）。在相依样本 t 检验中，两个样本可能包含同一个人在两个不同时刻进行测量或者两个有联系的人分别测量的结果（例如，双胞胎的 IQ，妻子与丈夫的沟通质量）。准确定义相依样本 t 检验的关键在于记住两样本间要在某方面存在自然联系。下面给出一个相依样本 t 检验的例子。

例子

一个国家选举机构的工作人员负责通过民意调查来决定经济和国家安全哪个议题对于选民更重要。有 25 个选民被调查以确定两个议题的重要性等级，每个议题用 1~7 的等级表示（1=一点也不重要，7=极其重要）。自变量是投票议题（经济、国家安全），因变量是重要性等级。¹

相依样本 t 检验的目标和数据要求

相依样本 t 检验		
目标	数据要求	例子
检验两个相关组别中关于某些感兴趣的因变量的均值是否存在显著差异	具有两个有联系的组或类别的一个自变量 一个连续因变量	自变量 • 投票议题（经济、国家安全） 因变量 • 重要性等级（1~7）

原假设和对立假设

原假设表明经济和国家安全性等级之差在总体上为 0：

$$H_0: \mu_{\text{经济}} - \mu_{\text{国家安全}} = 0$$

对立假设表明经济和国家安全性等级之差在总体上不为 0：

$$H_1: \mu_{\text{经济}} - \mu_{\text{国家安全}} \neq 0$$

原假设评价

相依样本 *t* 检验对两总体均值差为 0 的原假设进行了检验。如果 *t* 检验产生的结果在原假设正确时看起来不可能（结果发生的可能性小于 5%），那么拒绝原假设。如果 *t* 检验产生的结果在原假设正确时看起来正确（结果发生的可能性大于 5%），那么不拒绝原假设。

研究问题

在研究中基本的兴趣问题也可以用研究问题的方式表示，例如，

“对选民来说，经济重要性等级和国家安全是否存在不同？”

数据

在图表 7—1 中列出了 25 个参与者的数据。

图表 7—1相依样本 *t* 假设的数据

Participant	Economy	National Security	Participant	Economy	National Security
1	5	7	9	7	3
2	6	4	10	5	4
3	5	2	11	4	6
4	6	3	12	6	4
5	5	5	13	7	5
6	7	7	14	4	7
7	7	6	15	6	4
8	4	5	16	7	5

续前表

Participant	Economy	National Security	Participant	Economy	National Security
17	7	5	22	3	1
18	6	7	23	5	2
19	6	4	24	7	6
20	6	2	25	7	7
21	6	4			

注：变量 participant 包含在数据中，但不用输入 SPSS。

在 SPSS 中输入数据及分析

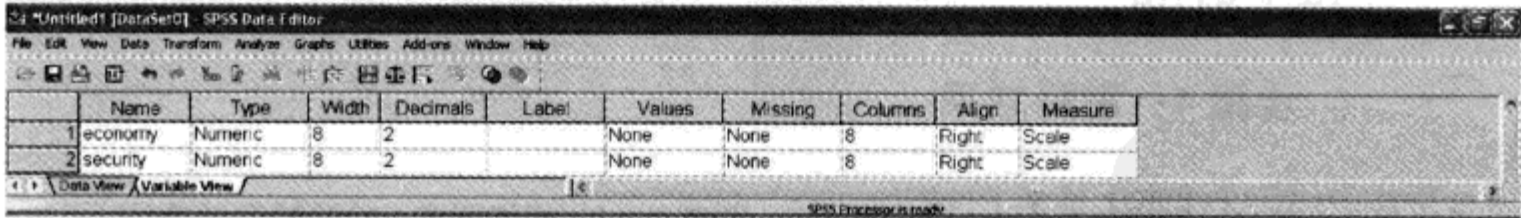
下面的步骤 1 和步骤 2 描述了怎样在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 7 中的文件 opinion.sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中将生成两个变量，分别用于经济等级和国家安全。两个变量分别命名为 **economy** 和 **security**。

- 3. 在 Variable View 窗口前两行分别输入变量名称 **economy** 和 **security**（见图表 7—2）。



图表 7—2 输入变量 economy 和 security 的 Variable View 窗口

步骤 2：输入数据

- 1. 点击 Data View 标签。变量 **economy** 和 **security** 出现在 Data View 窗口

的前两列。

2. 参照图表 7—1，为每个参与者输入两个变量的数据。对第一个参与者，为变量 **economy** 和 **security** 分别输入等级 5 和 7。依次输入全部 25 个参与者的数据。图表 7—3 中给出了完整的数据文件。

步骤 3：分析数据

1. 从菜单栏中选择 **Analyze > Compare Means > Paired-Samples T test...**（见图表 7—4）。

打开 Paired-Samples T Test 对话框，变量 **economy** 和 **security** 出现在对话框的左边（见图表 7—5）。

2. 选择变量 **economy** 和 **security**，点击向右箭头按钮（▶），把变量移到 Paired Variables 框中（见图表 7—6）。

3. 点击“OK”。

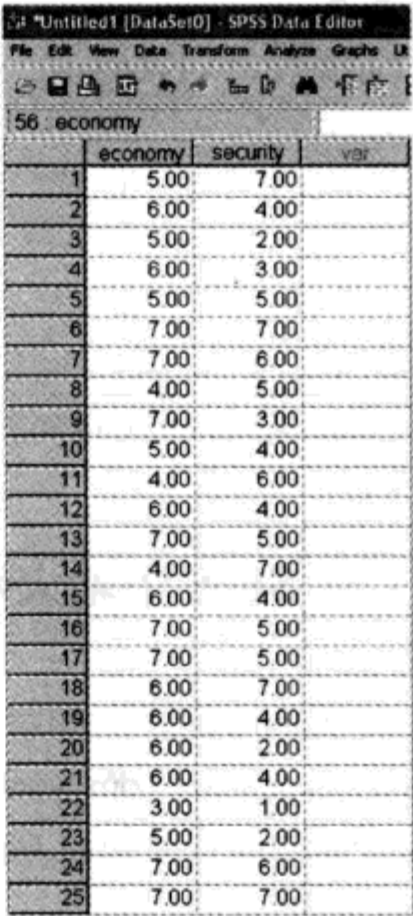
在 SPSS 中运行相依样本 *t* 检验程序，结果显示在 Viewer 窗口中。

步骤 4：解释结果

相依样本 *t* 检验的输出结果显示在图表 7—7 中。

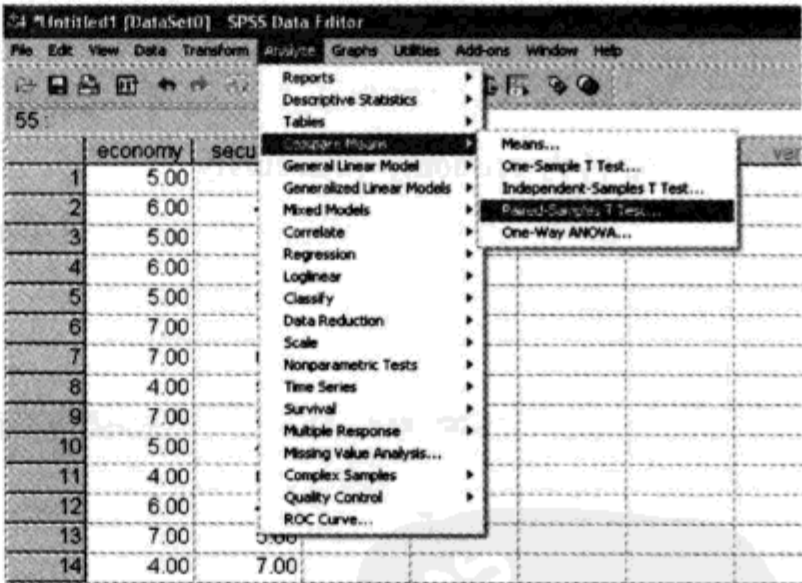
配对样本统计量 (Paired Samples Statistics)

输出的第一个表格 Paired Samples Statistics 显示了 **economy** 和 **security** 的描述统计量、包括样本量、平均值，标准差和标准误差。请注意，经济的平均重要性等级（均值=5.76）比国家安全（均值=4.60）的高。我们稍后将

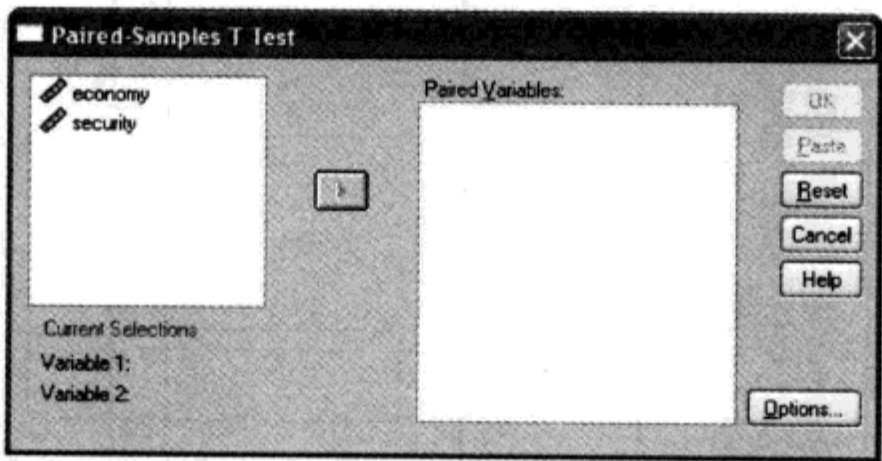


	economy	security	V01
1	5.00	7.00	
2	6.00	4.00	
3	5.00	2.00	
4	6.00	3.00	
5	5.00	5.00	
6	7.00	7.00	
7	7.00	6.00	
8	4.00	5.00	
9	7.00	3.00	
10	5.00	4.00	
11	4.00	6.00	
12	6.00	4.00	
13	7.00	5.00	
14	4.00	7.00	
15	6.00	4.00	
16	7.00	5.00	
17	7.00	5.00	
18	6.00	7.00	
19	6.00	4.00	
20	6.00	2.00	
21	6.00	4.00	
22	3.00	1.00	
23	5.00	2.00	
24	7.00	6.00	
25	7.00	7.00	

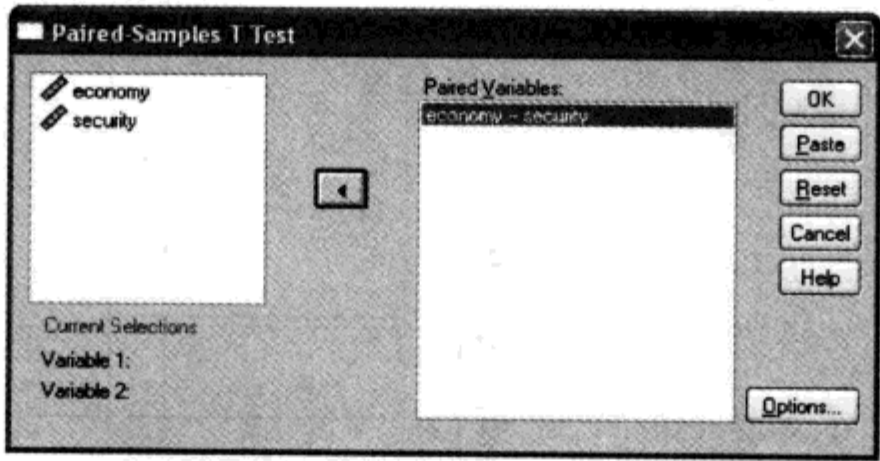
图表 7—3 相依样本 *t* 检验例子中的完整数据文件



图表 7—4 相依样本 *t* 检验的菜单命令



图表 7—5 Paired-Samples T Test 对话框



图表 7—6 Paired-Samples T Test 对话框（续）

会考虑这两个平均等级之间的差异（5.76 对 4.60）是否大到足以具有统计显著性。

配对样本相关性 (Paired Samples Correlations)

表格 Paired Samples Correlations 除了提到这个相关性等于 25 个参与者对于经济和国家安全的等级之间的皮尔逊相关系数（皮尔逊相关系数将在第 12 章讨论）外，对于解释配对样本 t 检验不是重要的，因此将不讨论。

配对样本检验 (Paired Samples Test)

表格 Paired Samples Test 为我们的研究问题提供了答案，就是经济和国家安全的重要性等级间是否存在差异。原假设的检验是以 t 的形式显示的，这里

$$t = \frac{\text{经济和安全的等级之间的均值差}}{\text{分数差的均值标准误差}}$$

从表格 Paired Samples Test 中代入适当的值（均值、均值标准误差）得

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 economy	5.7600	25	1.16476	0.23295
security	4.6000	25	1.77951	0.35590

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 economy &. security	25	0.253	0.222

表格Paired Samples Statistics提供了两个变量 (economy和security) 的均值。如果原假设被拒绝, 我们将查阅这个表格以决定哪个变量有更高的均值。

Paired Samples Test

	Paired Differences				t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference			
				Lower			
Pair 1 economy-security	1.16000	1.86369	0.37274	0.39071	1.92929	24	0.005

均值1.16等于economy和security均值间的差 (5.76-4.60=1.16)。

因为p-值为0.005, 小于0.05, 原假设被拒绝。

图表 7—7 相依样本 t 检验的输出结果

到 t 值是

$$t = \frac{1.16}{0.37274} = 3.112$$

这与表格 Paired Samples Test 中 t 列的值一致。有 25 个参与者，自由度 (df) 等于 24 (参与者数目-1)，对应的 p 值为 0.005。

因为 p 值 0.005 小于 0.05，所以均值相等的原假设被拒绝，结论为经济和国家的重要性等级有显著差异。因为检验是显著的，从 Paired Samples Statistics 表格中检查均值，发现经济的重要性等级均值 (均值=5.76) 比国家安全 (均值=4.60) 的高。(如果 t 检验不显著，我们应该假设任何表格 Paired Samples Statistics 中的均值差异都会导致抽样错误。)

效应量

对相依样本 t 检验的效应量的估计通常由 d 给出，其中

$$d = \frac{\text{均值差}}{\text{分数差的标准差}}$$

从表格 Paired Samples Test 中代入适当的值 (均值、标准差) 得到效应量为

$$d = \frac{1.16}{1.86} = 0.62$$

科恩 (Cohen, 1988) 约定对于相依样本 t 检验的小、中和大的效应量分别对应着 0.2, 0.5 和 0.8。根据科恩准则，在实践中效应量 0.62 被认为对应中等效应量，表明经济的重要性等级比国家安全高，差值为 0.62 倍标准差。

APA 格式的结果表达

在写结果时，除了报告样本均值和样本标准差，还要报告假设检验的结论，自由度， t 值， p -值和效应量。接下来我们将展示一个 APA 格式的简单书面例子。

选民认为经济 ($M=5.76$, $SD=1.16$) 显著重要于国家安全 ($M=4.60$, $SD=1.78$), $t(24) = 3.11$, $p < 0.05$, $d = 0.62$ 。

相依样本 t 检验的假定

下面提供了相依样本 t 检验的假定。第二个假定是涉及分数差而不是原来显示在图表 7—1 中的原始分数。分数差等于每个参与者的两个度量之间的差（每个人对于经济和国家安全等级评分的差）。例如，图表 7—1 中显示的第一个参与者的数据得到等级差为 -2 ($5-7=-2$)。

1. 组内观测是独立的。

违反组内独立性这一假定将严重损坏相依样本 t 检验的准确性。该假定能通过设计研究使得参与者在任何情况下都不相互影响得到满足（参与者们共同合作确定经济重要性等级是违反独立假定的一个例子）。如果有理由相信已经违反了组内独立性假定，那么就不应该再用相依样本 t 检验了。重要的是，组间的独立性不是相依样本 t 检验的假定（我们期望同一个人的两个分数间存在相依性，这是在开始我们使用相依样本 t 检验的原因）。

2. 因变量总体服从正态分布。

这个假定意味着等级差的总体应该服从正态分布。对于中等到较大的样本量，绝大部分的非正态分布趋向于对 t 检验的精确性没有多少影响，尽管一些非正态分布可能对 t 检验的功效有不利影响。

在 SPSS 中执行相依样本 t 检验的步骤摘要

I. 数据输入和分析

1. 在 SPSS 中生成两个变量。
2. 输入数据。
3. 选择 **Analyze>Compare Means>Paired-Samples T Test...**
4. 将两个变量移到 Paired Variables 框中。
5. 点击“OK”。

II. 结果解释

1. 检查 Paired Samples Test 表格中右边栏中的 p 值（“sig.”）：
 - 如果 $p \leq 0.05$ ，拒绝原假设。检查表格 Paired Samples Statistics 中的均值，并写出表示两组间存在差异本质的结果。

● 如果 $p > 0.05$ ，不拒绝原假设。写出表示两组间不存在显著差异的结果。

练习

1. 调查竞选电影对政治候选人态度上的影响。在放影片之前，使用政治好感度量 15 个人对候选人的态度。完成度量后，给参与者们播放 20 分钟的候选人的竞选电影，之后再由他们再一次完成好感度评价。图表 7—8 中列出了参与者两个阶段的好感度评分（好感度分数范围是 20~100 分，分数越高，表示对候选人的好感度越高）。

图表 7—8 观看电影前后对政治候选人的态度倾向

参与者	观影前	观影后	参与者	观影前	观影后
1	75	80	9	52	59
2	25	26	10	46	52
3	50	58	11	48	52
4	45	65	12	48	53
5	48	59	13	57	61
6	88	93	14	55	59
7	55	52	15	42	41
8	78	92			

在 SPSS 中输入数据，进行必要分析以回答下列问题。变量命名为 beforefilm（观影前）和 afterfilm（观影后）。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 观看电影后对候选人的态度发生了改变吗？在 $\alpha=0.05$ 时检验。
- d. 效应量是多少？你能区分效应量是小、中还是大吗？
- e. 将研究结果写成适当的 APA 格式。

2. 空中交通管制员经常经历高强度的工作压力。为努力帮助空管员应对他们的压力，调查放松训练的影响。从国内一些最繁忙的机场选取了 30 名空管员，对他们实施了一项关于工作相关压力的调查（范围为 10~50，高的分数表示高压力程度）。在完成压力等级调查之后，对空管员进行 4 周的放松训练，然后再一次测量他们的工作相关压力指数。数据位于网站 www.pren-

hall.com/yockey 上文件夹 Chapter 7 中的文件 Chapter7 _ Exercise 2. sav 之中 (变量命名为 **prerelax** (放松前) 和 **postrelax** (放松后))。在 SPSS 中打开文件, 进行必要分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 在进行放松训练后压力程度发生了改变吗? 在 $\alpha=0.05$ 时检验。
- d. 效应量是多少? 你能区分效应量是小、中还是大吗?
- e. 将研究结果写成适当的 APA 格式。

3. 一个研究者进行了接受运动心理学家的训练对于校园网球选手打球准确度影响的调查。分别记录了 15 名网球运动员在接受 8 周的运动心理学家的训练前后的打球准确度 (准确度分数范围为 0~100 分, 高的分数代表高的准确度)。数据在网站上文件夹 Chapter 7 中的文件 Chapter7 _ Exercise 3. sav 之中 (变量命名为 **beforetraining** (训练前) 和 **aftertraining** (训练后))。在 SPSS 中打开文件, 进行必要分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 在进行训练后网球选手的准确度跟以前相比发生了改变吗? 在 $\alpha=0.05$ 时检验。
- d. 效应量是多少? 你能区分效应量是小、中还是大吗?
- e. 将研究结果写成适当的 APA 格式。



一维组间方差分析

当对感兴趣的因变量在两个或者更多独立组中的均值进行比较时，我们可以使用一维组间方差分析（ANOVA^①）。在一维组间方差分析中，自变量是组间因素，每个参与者都仅得到因素的一个水平（就是说，每个人都在一个单独的组内）。¹ 下面给出一个一维组间方差分析的例子。

例子

作为班级需要的研究项目的一部分，一个学生想调查三种学习策略（策略 A，B 和 C）在单词记忆方面有没有差异。30 个学生同意参与这个研究，每个人都随机分配接受三种策略中的一种（每 10 个学生在一个策略组）。在学习了感兴趣的策略后，给每个学生 15 个单词看 5 分钟，要求使用新策略记住尽可能多的单词。片刻之后，每个学生要尽可能写出他能够记住的单词，同时研究者记录下每个参与者回忆起的正确的单词数量。在这项研究中，自变量是学习策略的类型（A，B 和 C），因变量是回忆起的正确单词的数量。

一维组间方差分析的目标和数据要求

一维组间方差分析		
目标	数据要求	例子
检验感兴趣的因变量在两个或者更多的组的均值是否有显著差异	自变量 <ul style="list-style-type: none"> • 组间因素有两个或更多单独的组或类别 因变量 <ul style="list-style-type: none"> • 连续 	自变量 <ul style="list-style-type: none"> • 学习策略（策略 A，B 和 C） 因变量 <ul style="list-style-type: none"> • 正确回忆的单词数量

① ANOVA: Analysis of Variance (方差分析)。——译者注

原假设和对立假设

原假设是三个组回忆起的单词数量的总体均值相等：

$$H_0: \mu_{策略A} = \mu_{策略B} = \mu_{策略C}$$

为了说明原假设是错误的，没有必要要求所有组之间彼此互不相同（即使这是一种可能性），所需要满足的是原假设在某些方面错误（某些组间会存在不同）。因此，不需要产生所有可能引起原假设错误的可能性（这将会随组数的增加而变得乏味），常用的陈述是均值在某些方面有不同。

$$H_1: \text{其中至少一个均值与其他均值不同}$$

原假设评价

一维组间方差分析对三个组的总体均值相等的原假设进行了检验。如果检验产生的结果在原假设正确时看起来不可能（结果发生的可能性小于 5%），那么拒绝原假设。如果检验产生的结果在原假设正确时看起来正确（结果发生的可能性大于 5%），那么不拒绝原假设。

研究问题

在研究中基本的兴趣问题也可以用研究问题的方式表示，例如，
“回忆起来的单词数量是否依赖于学习策略的类型？”

数据

在图表 8—1 中列出了 30 个参与者的数据。学习策略 A 的参与者记为“1”，学习策略 B 的参与者记为“2”，学习策略 C 的参与者记为“3”。

图表 8—1 一维组间方差分析的数据

Participant	Strategy	Word recall	Participant	Strategy	Word recall	Participant	Strategy	Word recall
1	1	8	3	1	9	5	1	9
2	1	10	4	1	7	6	1	7

续前表

Participant	Strategy	Word recall	Participant	Strategy	Word recall	Participant	Strategy	Word recall
7	1	10	15	2	9	23	3	5
8	1	8	16	2	8	24	3	6
9	1	12	17	2	12	25	3	4
10	1	9	18	2	10	26	3	2
11	2	12	19	2	10	27	3	5
12	2	10	20	2	12	28	3	3
13	2	9	21	3	6	29	3	7
14	2	10	22	3	4	30	3	5

说明：变量 participant 包含在数据中，但不用输入 SPSS。

在 SPSS 中输入数据及分析

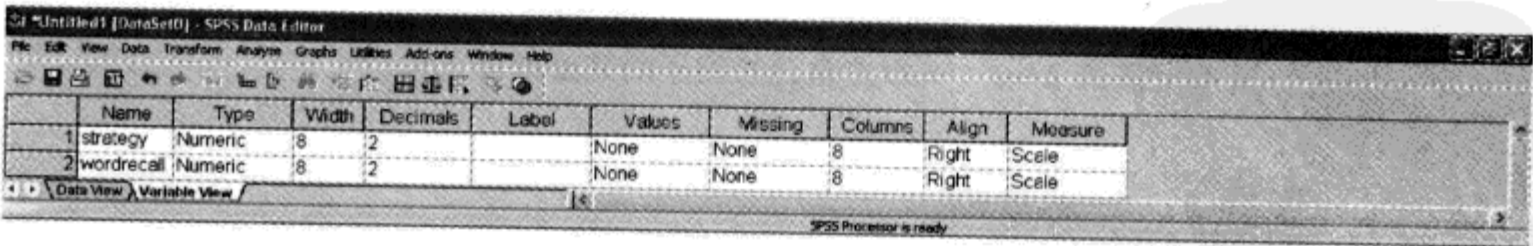
下面的步骤 1 和步骤 2 描述了怎样在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 8 中的文件 word recall. sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中，将生成两个变量，分别用于不同策略组（自变量）和回忆起的单词量（因变量）。两个变量分别命名为 **strategy** 和 **wordrecall**。

- 3. 在 Variable View 窗口前两行输入变量 **strategy** 和 **wordrecall**（见图表 8—2）。



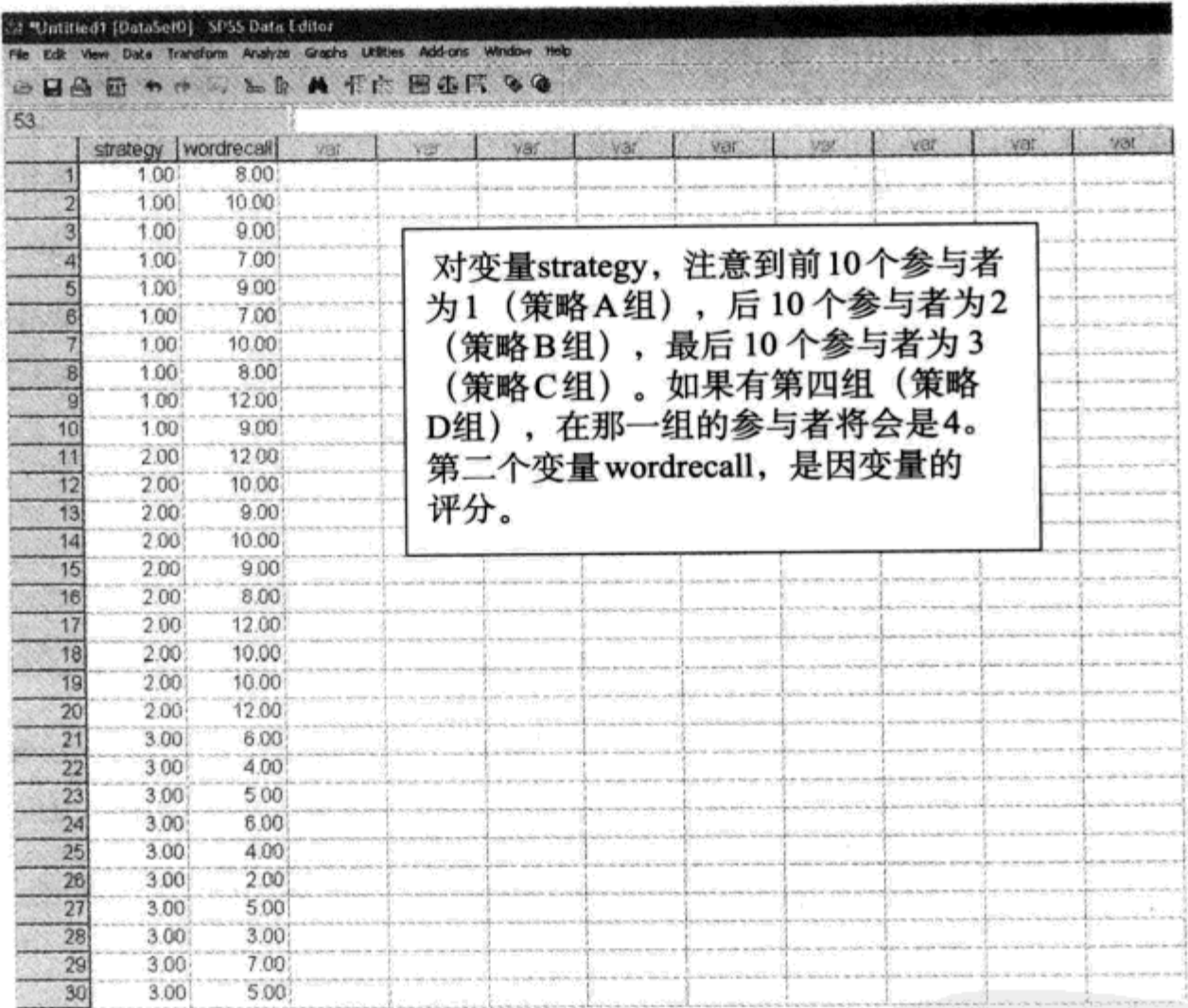
图表 8—2 输入了变量 strategy 和 wordrecall 的 Variable View 窗口

- 4. 使用在第 1 章描述的过程，为变量 **strategy** 建立变量值标签。对于

strategy, 1= “策略 A”, 2= “策略 B”, 3= “策略 C”。

步骤 2：输入数据

- 1. 点击 Data View 标签。变量 strategy 和 wordrecall 出现在 Data View 窗口的前两列。
- 2. 参照图表 8—1，为每个参与者输入两个变量的数据。对第一个参与者，为变量 strategy 和 wordrecall 分别输入数值 1 和 8。依次输入全部 30 个参与者的数据。图表 8—3 给出了完整的数据文件。



对变量strategy，注意到前10个参与者为1（策略A组），后10个参与者为2（策略B组），最后10个参与者为3（策略C组）。如果有第四组（策略D组），在那一组的参与者将会是4。第二个变量wordrecall，是因变量的评分。

	strategy	wordrecall	var	var	var	var	var	var	var	var	var
1	1.00	8.00									
2	1.00	10.00									
3	1.00	9.00									
4	1.00	7.00									
5	1.00	9.00									
6	1.00	7.00									
7	1.00	10.00									
8	1.00	8.00									
9	1.00	12.00									
10	1.00	9.00									
11	2.00	12.00									
12	2.00	10.00									
13	2.00	9.00									
14	2.00	10.00									
15	2.00	9.00									
16	2.00	8.00									
17	2.00	12.00									
18	2.00	10.00									
19	2.00	10.00									
20	2.00	12.00									
21	3.00	6.00									
22	3.00	4.00									
23	3.00	5.00									
24	3.00	6.00									
25	3.00	4.00									
26	3.00	2.00									
27	3.00	5.00									
28	3.00	3.00									
29	3.00	7.00									
30	3.00	5.00									

图表 8—3 一维方差分析例子的完整数据文件

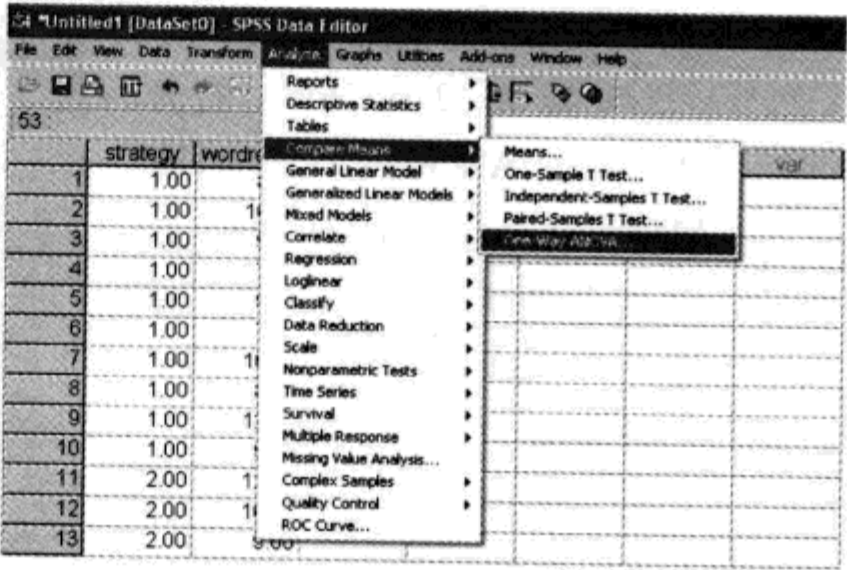
步骤 3：分析数据

- 1. 从菜单栏中选取 Analyze>Compare Means>One-Way ANOVA...（见图表 8—4）。

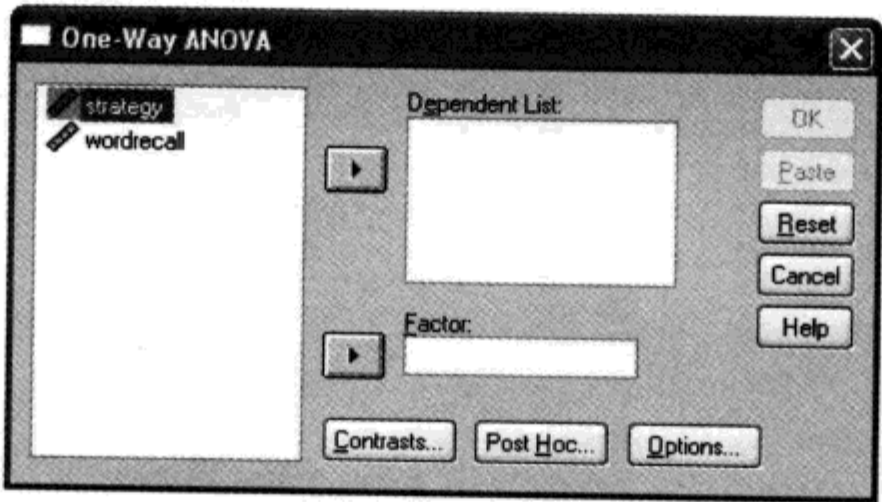
打开 One-Way ANOVA 对话框，变量 **strategy** 和 **wordrecall** 出现在对话框的左边（见图表 8—5）。

2. 选择因变量 **wordrecall**，点击上面的向右箭头按钮（▶），把因变量移动到 Dependent List 框中。

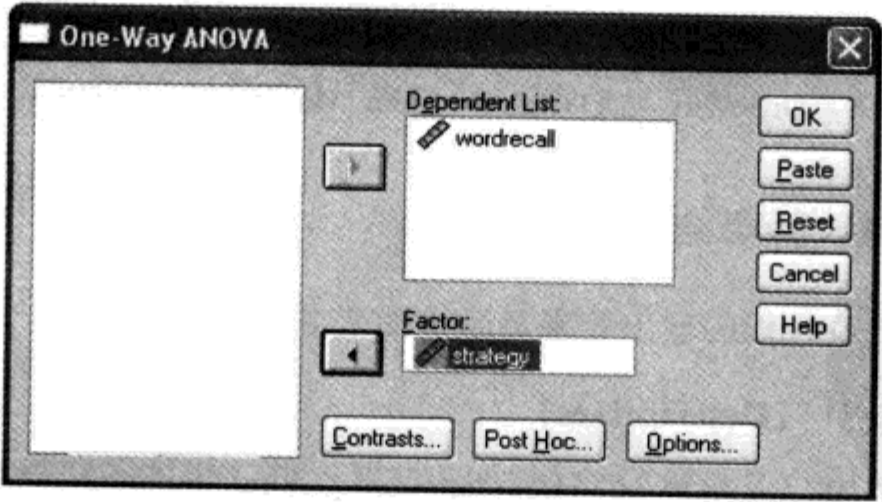
3. 选择自变量 **strategy**，点击下面的向右箭头按钮（▶），把自变量移动到 Factor 框（见图表 8—6）。



图表 8—4 One-Way ANOVA 程序的菜单命令



图表 8—5 One-Way ANOVA 对话框

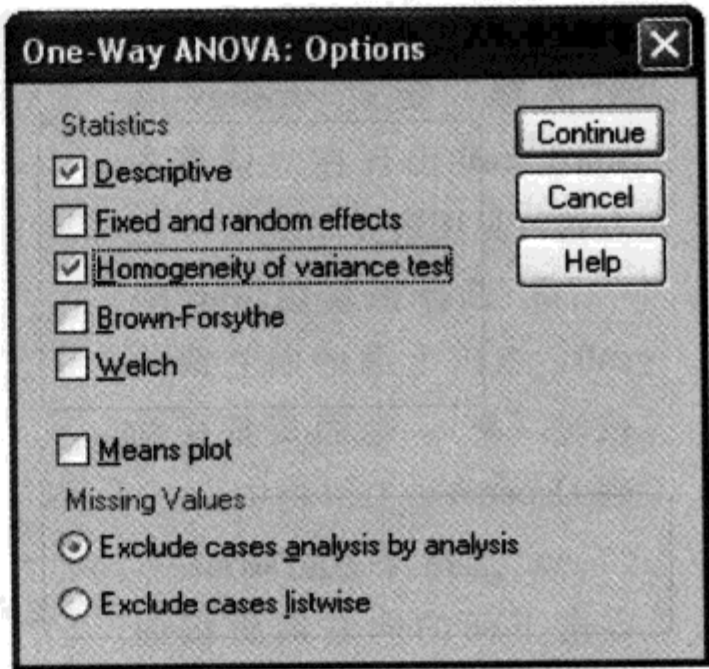


图表 8—6 One-Way ANOVA 对话框（续）

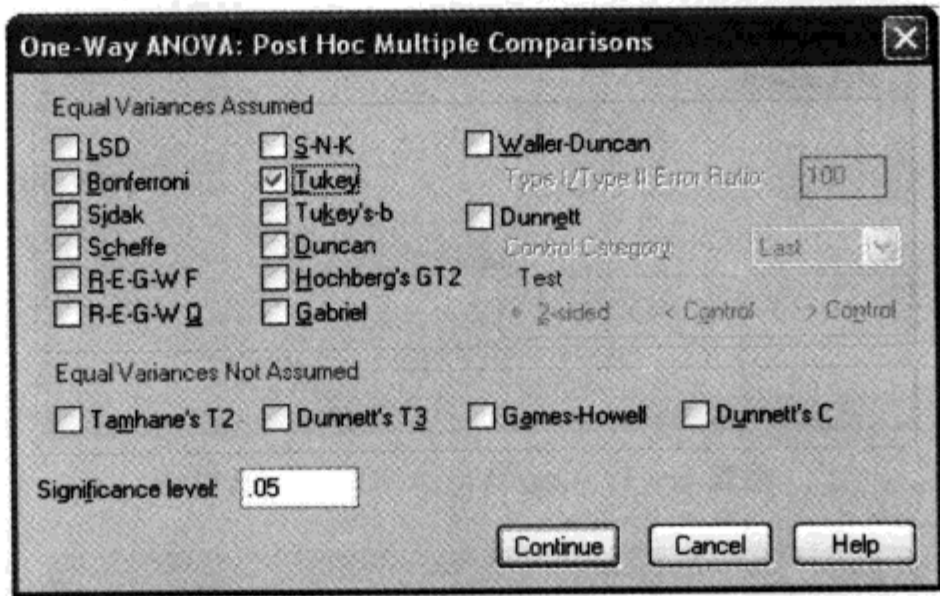
4. 点击“Options”，打开 One-Way ANOVA: options 对话框，在 Statistics 下面选择 Descriptive 和 Homogeneity of variance test²（见图表 8—7）。

- 5. 点击 “Continue”。
- 6. 点击 “Post Hoc”，打开 One-Way ANOVA: Post Hoc Multiple Comparisons 对话框，在 Equal Variance Assumed 的下面选择 Tukey（不是 Tukey's-b）³（见图表 8—8）。
- 7. 点击 “Continue”。
- 8. 点击 “OK”。

在 SPSS 中运行一维方差分析程序，结果显示在 Viewer 窗口中。



图表 8—7 One-Way ANOVA: Options 对话框



图表 8—8 One-Way ANOVA: Post Hoc Multiple Comparisons 对话框

步骤 4：解释结果

一维方差分析的输出结果显示在图表 8—9 中。

描述统计 (Descriptives)

输出的第一个表格 Descriptives 为每一组（和全部样本）列出了描述统计量。检查不同策略的回忆单词的均值可以发现，策略 B 的均值最高（均值 = 10.20），其次是策略 A（均值 = 8.9），最低的是策略 C（均值 = 4.70）。均值之间的差异是否大到足以具有统计显著性，我们很快将考虑这个问题。

One way

Descriptives

wordrecall

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Strategy A	10	8.9000	1.52388	0.48189	7.8099	9.9901	7.00	12.00
Strategy B	10	10.2000	1.39841	0.44222	9.1996	11.2004	8.00	12.00
Strategy C	10	4.7000	1.49443	0.47258	3.6309	5.7691	2.00	7.00
Total	30	7.9333	2.77841	0.50727	6.8959	8.9708	2.00	12.00

Test of Homogeneity of Variances

wordrecall

Levene Statistic	df1	df2	Sig.
0.021	2	27	0.980

ANOVA 中的一个假设——方差相等检验的 p -值。因为0.980大于0.05，不能拒绝原假设，假定三组策略方差相等。

ANOVA

wordrecall

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	165.267	2	82.633	38.073	0.000
Within Groups	58.600	27	2.170		
Total	223.867	29			

总体均值是否相等检验中的 p -值。因为 p -值小于0.05，表示三个策略组均值相等的原假设被拒绝。

图表 8—9 一维组间方差分析的输出结果

Post Hoc Tests

图 8-9 一维组间方差分析结果(续)
Multiple Comparisons

Dependent Variable: wordrecall

Tukey HSD

(I) strategy	(J) strategy	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Strategy A	Strategy B	-1.30000	0.65884	0.138	-2.9335	0.3335
	Strategy C	4.20000*	0.65884	0.000	2.5665	5.8335
Strategy B	Strategy A	1.30000	0.65884	0.138	-0.3335	2.9335
	Strategy C	5.0000*	0.65884	0.000	3.8665	7.1335
Strategy C	Strategy A	-4.20000*	0.65884	0.000	-5.8335	-2.5665
	Strategy B	-5.50000*	0.65884	0.000	-7.1335	-3.8665

* The mean difference is significant at the 0.05 level.

Homogeneous Subsets

图 8-9 一维组间方差分析结果(续)
wordrecall

Tukey HSD*

Strategy	N	Subset for alpha=0.05	
		1	2
Strategy C	10	4.7000	
Strategy A	10		8.9000
Strategy B	10		10.2000
Sig.		1.000	0.138

Means for groups in homogeneous subsets are displayed.
* Uses Harmonic Mean Sample Size=10.000.

因为策略C不与策略A和策略B在一列，所以它显著不同于(低于)策略A和策略B。

因为策略A和策略B在一列，所以策略A和策略B无显著差异。

图 8-9 一维组间方差分析的结果(续)

方差齐性检验 (Test of Homogeneity of Variances)

表格 Test of Homogeneity of Variances 用以检验三组的方差是否相等，它是组间方差分析的一个假设。SPSS 使用 Levene 建立的程序来检验方差齐性假设。

Levene 检验的原假设和对立假设是：

$H_0 : \sigma^2_{策略A} = \sigma^2_{策略B} = \sigma^2_{策略C}$ （三组的总体方差相等）

H_1 ：其中至少一个方差与其他方差不相等

方差齐性检验由表格 Test of Homogeneity of Variances 中输出的 p -值来评定。如果 $p \leq 0.05$ ，拒绝原假设，表明总体方差不相等。如果 $p > 0.05$ ，不拒绝原假设，假定三组的总体方差相等。

对于我们的数据，Levene 检验得到 $F=0.021$ （SPSS 表中标记为 Levene Statistic 的值）和 $p\text{-值}=0.980$ （见图表 8—9）。由于 0.980 大于 0.05，不拒绝原假设。因此，基于 Levene 检验的结果，我们假定三组的总体方差相等。方差齐性假设的判断规则总结在图表 8—10 中。

图表 8—10 评价方差齐性假设的规则

Levene 结果	决策	结论
$p > 0.05$	不拒绝 H_0	总体方差相等
$p \leq 0.05$	拒绝 H_0	总体方差不相等
对于我们的数据， $p\text{-值} = 0.980, 0.980 > 0.05$	不拒绝	总体方差相等

ANOVA：检验三种策略均值相等的原假设

表格 ANOVA 给出了我们研究问题的答案，即三个策略组的单词准确回忆数是否不同。ANOVA 进行了 F 检验，即两方差的比率，而且每个方差在输出结果中表示为均方（MS）的形式：

$$F = \frac{\text{组间均方}}{\text{组内均方}}$$

从图表 8—9 的 ANOVA 表格中代入适当的值，得到 F 值为

$$F = \frac{82.633}{2.170} = 38.073$$

这与 ANOVA 表中 F 列给出的值一致。

该检验中有两个自由度，组间自由度（组数目-1）和组内自由度（样本

容量总数一组的数目), 由 ANOVA 表得到组间和组内自由度分别是 2 和 27。
在 ANOVA 表格 “sig.” 列下的 p 值为 0.000 (读作 “小于 0.001” 而不是 “0”)。由于 p 值小于 0.05, 拒绝均值相等的原假设, 并得出结论, 其中至少一个方差与其他方差不相等。

事后检验：评价哪些组不同

尽管三组的均值相等的原假设被拒绝了, 但是对立假设并不明确, 仅仅说明均值在某些方面不同。为了找出这些组如何不同, 需要更深入地检验。
一种常见的做法就是检验所有可能的成对的组, 通常称为检验配对比较。SPSS 可以提供很多不同的配对比较的检验, 最为常见的一种检验是 Tukey’s post hoc 方法 (post hoc 表示 “在此之后”; Tukey 的检验一般在总体 ANOVA 得到显著结果之后进行解释, 也就是说, 在原假设被排除后)。在当前三个组的例子中, Tukey 程序将进行三个独立检验: A 和 B, A 和 C, B 和 C。⁴

SPSS 默认为 Tukey 检验提供两个不同的输出表格, Multiple Comparisons 和 Homogeneous Subsets。尽管其中任一个表格都可以用于解释配对比较的结果, 但是通常使用后面的表格 Homogeneous Subsets 来描述结果。在图表 8—9 中的表格 Homogeneous Subsets 中, 给出了记为 “1” 和 “2” 的两个不同的列。在解释表格 Homogeneous Subsets 产生的结果时, 共享同一列的组间没有显著差异。例如, 策略 A 和策略 B 共享同一列 (列 2), 它们之间没有显著差异 (策略 A 与策略 B 间的任何差异将被认为是抽样误差导致的)。因为策略 C 与策略 A、策略 B 不共享同一列, 策略 C 与策略 A 和策略 B 之间有显著差异。检查表格中给出的三组单词回忆数目均值可以看到, 策略 C (均值=4.7) 低于策略 A (均值=8.9) 和策略 B (均值=10.2)。

下面的表格概括了 Homogeneous Subsets 表中 Tukey 检验的结果。

检验	组间是否共享同一列	结果
A 和 B	是	不显著; 推测 A 与 B 相等
A 和 C	否	显著; A 比 C 能回忆更多单词
B 和 C	否	显著; B 比 C 能回忆更多单词

作为选择, 配对比较的结果可以用 Multiple Comparisons 表解释。在图表 8—9 中给出的表格 Multiple Comparisons 中, 在前两栏列出用于检验的两个组, 随后是组间的均值差, 标准误差, p 值 (sig.) 和 95% 置信区间。从该表

的第一行读起,检验策略 A 与策略 B,得到均值差为-1.3 (策略 A 均值-策略 B 均值) 和 p 值为 0.138。因为 p 值大于 0.05,策略 A 与策略 B 之间无显著差异。下一组配对比较 (沿着对角线读) 是策略 A 和策略 C,得到均值差为 4.2 (策略 A 均值-策略 C 均值) 和小于 0.001 的 p 值 (SPSS 中由舍入而表示成 0.000)。由于 p 值小于 0.05,策略 A 和策略 C 显著不同。检查图表 8—9 里 Descriptives 表中的均值,显示用过策略 A (均值=8.9) 的人平均比用过策略 C (均值=4.7) 的人能回忆起更多单词。

Multiple Comparisons 中潜在的混淆部分在于每一个配对比较出现了两次。在 A 与 C 的检验之后, A 与 B 的检验又一次出现,唯一的不同之处在于 B 出现在前。由于 B 和 A 的检验与 A 和 B 的完全相同 (注意到两个检验中 p 值完全相等,为 0.138),给出的两个检验没有差别。表格 Multiple Comparisons 总会产生这种冗余,每个配对比较的结果出现两次。因此,在解释结果时,请确认对每一配对比较只报告一次。概述目前的结果,我们报告了 A 和 B 与 A 和 C 间的检验结果,还剩 B 和 C 的检验。沿着下一行的对角线看,我们发现 B 和 C 显著不同, p 值小于 0.001 (依然由于 SPSS 中的舍入关系报告为 0.000)。表格 Descriptives 中的均值显示出策略 B (均值=10.2) 显著高于策略 C (均值=4.7)。总的来说,两种表格的结果相同: A 和 B 之间无显著差异, A 和 B 都显著高于 C。

效应量

ANOVA 中的效应量通常使用 η^2 度量,其中

$$\eta^2 = \frac{\text{组间平方和}}{\text{总平方和}}$$

组间平方和与总平方和可以从图表 8—9 的 ANOVA 表格中找到。⁵ 从该表中代入适当的值得到 η^2 为

$$\eta^2 = \frac{165.267}{223.867} = 0.74$$

科恩 (Cohen, 1988) 约定小、中和大的效应量的 η^2 值分别对应着 0.01, 0.06 和 0.14。效应量 η^2 可以用因变量的方差被自变量解释的百分比的方式来表示。在实践中效应量 0.74 被认为对应一个非常大的效应,表明学习策略解释了单词回忆的方差的 74%。

APA 格式的结果表达

在写结果时，除了报告三个策略组的均值和标准差（如果需要，均值和标准差可以单独成一张表），还要报告一维组间方差分析的假设检验的结论、自由度 (df)、 F 值、 p 值和效应量。接下来将展示一个 APA 格式的简单书面例子。

由学习策略的使用引起单词回忆数量的变化， $F(2, 27) = 38.07$ ， $p < 0.05$ ， $\eta^2 = 0.74$ 。Tukey 的事后检验程序表明使用策略 A 的人 ($M = 8.90$ ， $SD = 1.52$) 和使用策略 B 的人 ($M = 10.20$ ， $SD = 1.40$) 能够比使用策略 C 的人 ($M = 4.70$ ， $SD = 1.49$) 回忆更多单词。使用策略 A 与使用策略 B 之间在单词回忆数量上没有显著差异。

一维组间方差分析的假定

1. 观测是独立的。

违反这一假定将严重影响 ANOVA 检验的准确度。该假定能通过设计你的研究使参与者不会以任何方式相互影响得到满足（参与者们在一起讨论单词表是违反独立假定的一个例子）。如果我们有理由相信已经违反了独立性假定，那么就不应再用组间方差分析。

2. 每组因变量的总体服从正态分布。

该假定意味着每一个策略组的单词回忆数量的总体应服从正态分布。对于中等到较大的样本量，绝大部分的非正态分布趋向于对 ANOVA 检验的精确性没有多少影响，尽管一些非正态分布可能对 ANOVA 检验的功效有不利影响。

3. 每组的总体方差相等。

违反方差齐性假定将影响到 ANOVA 检验的准确度，特别是当各组的样本量不相等时。解释 SPSS 中的 Levene 检验的结果，如果方差不相等（需要的话接下来要进行合适的事后检验，例如 Dunnett's T3），选择并报告另一种程序（Brown-Forsythe 检验或 Welch 检验）来满足该假定。

在 SPSS 中执行一维组间方差分析的步骤摘要

I. 数据输入和分析

1. 在 SPSS 中生成两个变量（一个是自变量，另一个是因变量）。
2. 输入数据。
3. 为自变量建立变量值标签，在 Value Labels 对话框中输入数据和合适的标签，点击“OK”。
4. 选择 **Analyze > Compare Means > One-Way ANOVA...**。
5. 将因变量移动到 Dependent List 框中，将自变量移动到 Factor 框中。
6. 点击“Options”，选择 Descriptive 和 Homogeneous of Variance Test，点击“Continue”。
7. 点击“Post Hoc”，选择 Tukey，点击“Continue”。
8. 点击“OK”。

II. 结果解释

1. 检查方差齐性的 Levene 检验。
 - a. 如果 Levene 检验中 $p > 0.05$ ，总体方差相等。解释 ANOVA F 。
 - 如果 ANOVA 显著 ($p \leq 0.05$)，解释 Tukey 事后程序的结果。
 - 如果 ANOVA 不显著 ($p > 0.05$)，则立即停止。写出组间无显著差异的结果。
 - b. 如果 Levene 检验中 $p \leq 0.05$ ，总体方差不等。选择 Brown-Forsythe 或者 Welch 程序（点击“Options”）来重新分析，并通过点击“Post Hoc”选择其中一个不假设组间方差相等的事后检验（例如，Dunnett's T3）。与 ANOVA 和 Tukey 检验不同，这些检验不需要方差齐性假设。
 - 如果全部检验（Brown-Forsythe 或者 Welch）显著 ($p \leq 0.05$)，解释事后程序给出的结果。
 - 如果全部检验（Brown-Forsythe 或者 Welch）不显著 ($p > 0.05$)，停止。写出组间无显著差异的结果。

练习

1. 一位药物研究员想调查不同的止痛药对于偏头痛患者的效果。21 个近期因偏头痛就医的患者随即接受三种药物中的一种：药物 A、药物 B 或安慰剂。当吃下相应的药物后，每天三次记录参与者报告的疼痛指数，持续一周（疼痛分为 10 个尺度，高的分数代表高度疼痛）。图表 8—11 列出了一周内的患者平均疼痛指数。

图表 8—11		三组的疼痛指数
Drug	Pain	
A	5.2, 4.1, 5.8, 6.85, 4.75, 1.75, 4	
B	3.05, 6.15, 5.5, 6.15, 1.85, 6.4, 3.1	
C	8.15, 7.15, 6.2, 7.85, 9.45, 9.25, 6.3	

在 SPSS 中输入数据，进行必要分析以回答下列问题。变量命名为 **drug** 和 **pain**。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 检验相等差异假设，组间的数据是否存在不等差异？
- d. 组间的疼痛指数是否存在明显差异？在 $\alpha=0.05$ 时检验。
- e. ANOVA 的效应量是多少？你能区分效应量是小、中还是大吗？
- f. 如果所有 ANOVA 显著，简要归纳出 Tukey 的事后程序结果。
- g. 将研究结果写成适当的 APA 格式（确认如果全部 ANOVA 显著，包含了 Turkey 检验的结果）。

2. 一个研究者想知道高的石油价格对不同类型的交通工具使用者的影响等级。随着油价的猛涨，研究员询问了 30 个人的受挫程度（受挫分数为 1~10 分，高的分数代表高的受挫程度）。其中用于交通运输的拥有摩托车的有 10 个人（数据文件中编号 1），拥有混合驱动汽车的有 10 人（数据文件中编号 2），拥有非混合驱动汽车的有 10 人（数据文件中编号 3）。数据在网站 www.prenhall.com/yockey 上 Chapter 8 中的文件 Chapter 8 _ Exercise 2. sav 之中（变量命名为 **vehicle**（交通工具）和 **frustration**（受挫程度））。

在 SPSS 中输入数据，进行必要分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 检验相等差异假设，组间的数据是否存在不等差异？在 $\alpha=0.05$ 时检验。
- d. 交通工具拥有者组间的受挫指数是否存在明显差异？
- e. ANOVA 的效应量是多少？你能区分效应量是小、中还是大吗？
- f. 如果所有 ANOVA 显著，简要归纳出 Tukey 的事后程序结果。
- g. 将研究结果写成适当的 APA 格式（如果总体 ANOVA 显著，确保包括 Turkey 检验的结果）。

3. 一个市场研究员受雇于救援队以调查公司内对紧急呼叫的反应程度。3 个龙头公司参与这项研究。每个公司对 15 个随机选择的警报的反应时间（以秒为单位）被记录下来并放在网站上 Chapter 8 文件夹中的 Chapter 8 _ Exercise 3. sav 文件之中（变量名为 **company**（公司）和 **time**（时间），对 **company**, 1=“公司 A”，2=“公司 B”，3=“公司 C”）。

在 SPSS 中输入数据，进行必要分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 检验相等差异假设，组间的数据是否存在不等差异？在 $\alpha=0.05$ 时检验。
- d. 公司警报的反应时间是否存在明显差异？在 $\alpha=0.05$ 时检验。
- e. ANOVA 的效应量是多少？你能区分效应量是小、中还是大吗？
- f. 如果所有 ANOVA 显著，简要归纳出 Tukey 的事后程序结果。
- h. 将研究结果写成适当的 APA 格式（如果总体 ANOVA 显著，确保包括 Turkey 检验的结果）。



二维组间方差分析

二维组间方差分析应用于两个自变量估计一个连续因变量的情况。在二维组间方差分析中，两个自变量都是包含两个或更多水平的组间因素，这里每个参与者只能接受每个因素的一个水平。下面给出一个二维组间方差分析的例子。

例子

一个研究者想调查物理治疗和放松锻炼对治疗背伤的效果。来自一个当地外科整形诊所的 24 名遭受严重背伤的病人被招募参加这次研究。物理治疗和放松锻炼分别有两种不同方式。对于物理治疗，一半的参与者（12 名）使用拉伸锻炼，另一半的参与者（12 名）使用力量锻炼。对于放松锻炼，一半参与者（12 名）使用肌肉放松（从头到脚的放松和松弛肌肉），另一半的参与者（12 名）使用引导意象（一个结构化的可视技术用于放松）。

研究计划（如图表 9—1 所示）就是一个物理治疗水平和放松锻炼水平之间的交叉图。这个研究设计导致四种情形（每种情形分别有 6 名参与者）：拉伸和肌肉放松（单元 a），拉伸和引导意象（单元 b），力量和肌肉放松（单元 c）及力量和引导意象（单元 d）。由于物理治疗有两个水平，放松锻炼也有两个水平，于是得到四种情形，因此这种设计通常被称作 2×2 方差分析。

物理治疗	放松锻炼			合计
	拉伸	肌肉放松	引导意象	
		6 名参与者 (单元 a)	6 名参与者 (单元 b)	
	力量	6 名参与者 (单元 c)	6 名参与者 (单元 d)	
合计		12	12	24

图表 9—1 二维组间方差分析研究实例

这个研究持续六周，每个参与者每周定期接受特定的训练。在六周结束时，每个参与者完成一个评价他们目前疼痛水平的调查问卷（调查问卷分数为 0~60，其中 0 表示没伤，60 表示严重背伤）。在这个研究中，六周后的疼痛水平是因变量，物理治疗和放松锻炼是自变量。

二维组间方差分析的目标和数据要求

目标	数据要求	例子
1. 检验主效应	自变量	自变量
• 拉伸和力量锻炼对背伤的影响有差异吗？	• 两个及以上水平的组间因素	• 物理治疗（拉伸和力量）
• 肌肉放松和引导意象对背伤的影响有差异吗？	因变量	• 放松锻炼（肌肉放松和引导意象）
2. 检验交互效应	• 连续	因变量
• 物理治疗对背伤的影响依赖于放松锻炼的类型吗？		• 六周之后的疼痛水平

原假设和对立假设

在二维组间方差分析中将检验三个不同的原假设。一个原假设被用来检验每个自变量（主效应检验），一个原假设用来检验两个自变量的混合效应（交互效应检验）。假设描述如下。

假设 1. 物理治疗检验：拉伸和力量锻炼

物理治疗的原假设指出拉伸和力量锻炼带来的平均疼痛水平在总体上是同样的：

$$H_0: \mu_{\text{拉伸}} = \mu_{\text{力量}} \quad (H. 1)$$

对立假设指出两个总体均值不相等:

$$H_1: \mu_{\text{拉伸}} \neq \mu_{\text{力量}}$$

假设 2. 放松锻炼检验: 肌肉放松和引导意象

放松锻炼的原假设指出肌肉放松和引导意象带来的平均疼痛水平在总体上是一样的:

$$H_0: \mu_{\text{肌肉放松}} = \mu_{\text{引导意象}} \quad (H. 2)$$

对立假设指出两个总体均值不相等:

$$H_1: \mu_{\text{肌肉放松}} \neq \mu_{\text{引导意象}}$$

假设 3. 物理治疗和放松锻炼交互效应的检验

原假设指出这两个自变量没有交互效应:

$$H_0: \text{没有物理治疗} \times \text{放松锻炼的交互效应} \quad (H. 3)$$

对立假设指出两者有交互效应:

$$H_1: \text{有物理治疗} \times \text{放松锻炼的交互效应}$$

原假设评价

二维组间方差分析对上面的三个原假设进行了检验。如果检验产生的结果在原假设正确时看起来不可能 (结果发生的可能性小于 5%), 那么拒绝原假设。如果检验产生的结果在原假设正确时看起来正确 (结果发生的可能性大于 5%), 那么不拒绝原假设。

研究问题

在二维方差分析研究中基本的兴趣问题也可以用研究问题的方式表示。

对于物理治疗

“使用拉伸锻炼和力量锻炼的人们的疼痛水平是否不同?”

对于放松锻炼

“使用肌肉放松和引导意象的人们的疼痛水平是否不同?”

对于物理治疗和放松锻炼的交互效应

“物理治疗对于背伤的影响是否依赖于肌肉放松或者引导意象?”

数据

在图表 9—2 中列出了 24 个参与者的数据。对于物理治疗，接受拉伸锻炼的记为 1，接受力量锻炼的记为 2。对于放松锻炼，接受肌肉放松的记为 1，接受引导意象的记为 2。

图表 9—2 二维组间方差分析例子的数据

Participant	Physica therapy	Relaxation exercise	Pain level	Participant	Physica therapy	Relaxation exercise	Pain level
1	1	1	30	13	2	1	40
2	1	1	22	14	2	1	50
3	1	1	25	15	2	1	38
4	1	1	28	16	2	1	52
5	1	1	20	17	2	1	45
6	1	1	20	18	2	1	50
7	1	2	50	19	2	2	50
8	1	2	45	20	2	2	55
9	1	2	35	21	2	2	50
10	1	2	40	22	2	2	45
11	1	2	30	23	2	2	47
12	1	2	45	24	2	2	43

说明：变量 participant 包含在数据中，但不用输入 SPSS。

在 SPSS 中输入数据及分析

下面的步骤 1 和步骤 2 描述了怎样在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 9 中的文件 back pain.sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

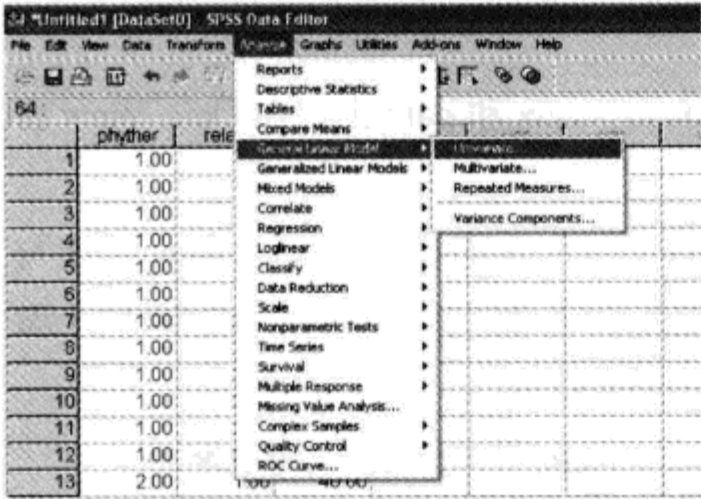
在 SPSS 中将生成三个变量，一个是物理治疗（自变量），一个是放松锻炼（自变量），一个是疼痛水平（因变量）。这些变量将被分别命名为 **phyther**，**relax** 和 **pain**。

步骤 3：分析数据

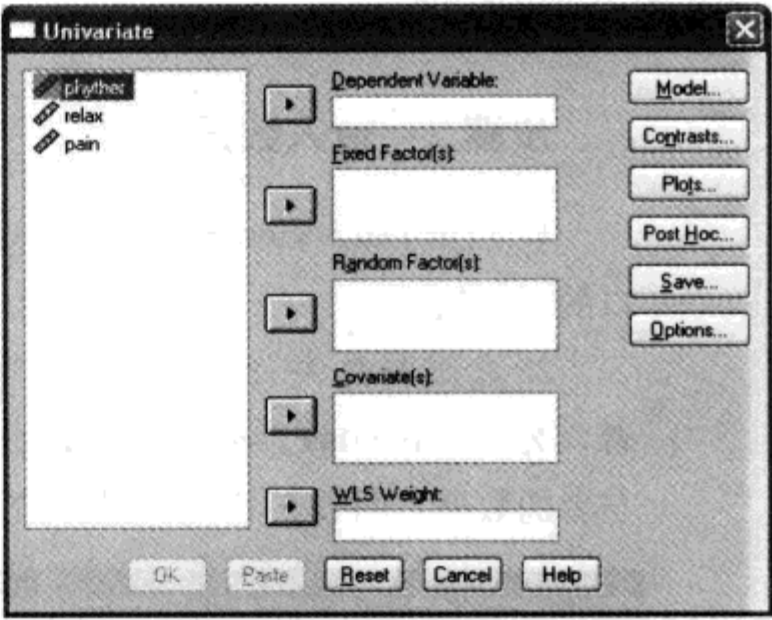
1. 从菜单栏中选择 **Analyze>General Linear Model>Univariate...**（见图表 9—5）。

打开 Univariate（单变量）对话框，变量 **phyther**，**relax** 和 **pain** 出现在对话框的左边（见图表 9—6）。

2. 选择因变量 **pain**，点击向右箭头按钮（▶），把变量移到 Dependent Variable 框中。



图表 9—5 二维组间方差分析的菜单命令



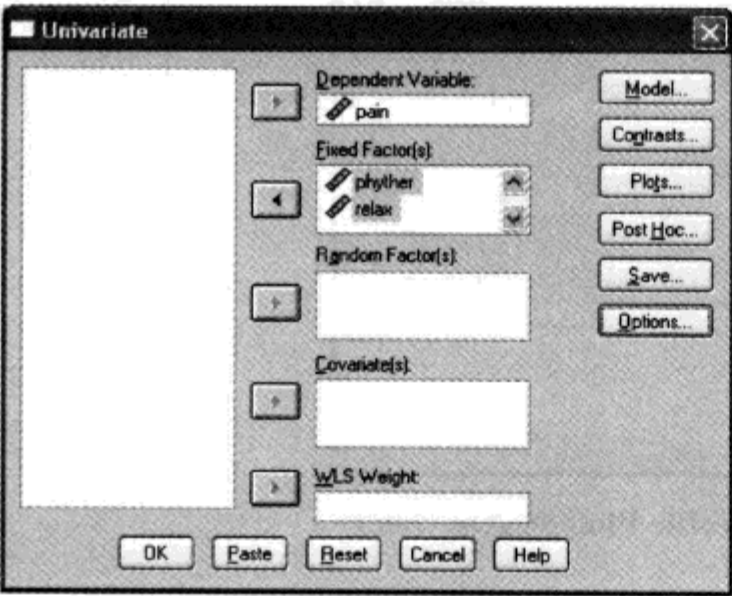
图表 9—6 Univariate 对话框

3. 选择自变量 **phyther** 和 **relax**，然后点击第二个向右箭头按钮（▶），把变量移到 Fixed Factor(s) 工具箱¹（见图表 9—7）。

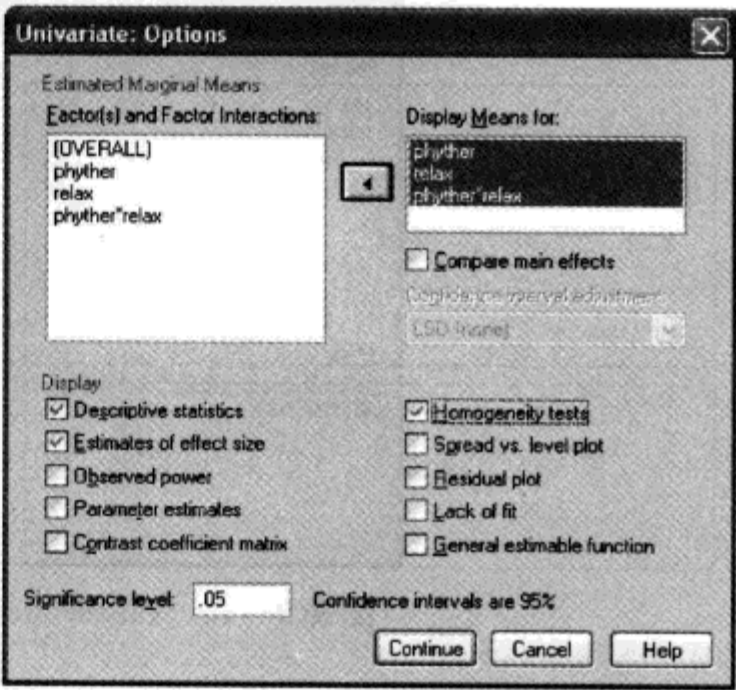
4. 点击“Options”，打开 **Univariate: Options** 对话框。在 Factor (s) and Factor Interactions 之下选择变量 **phyther**，**relax** 和 **pain**（不要选择 **OVERALL**），然后点击向右箭头按钮（▶）使它们移到 Display Means for 框中。在 Display 之下，选择 Descriptive statistics, Estimates of effect size 和 Homogeneity tests（见图表 9—8）。

5. 点击“Continue”。

6. 点击“Plots”，打开 **Univariate: Profile Plots** 对话框。选择 **phyther** 并

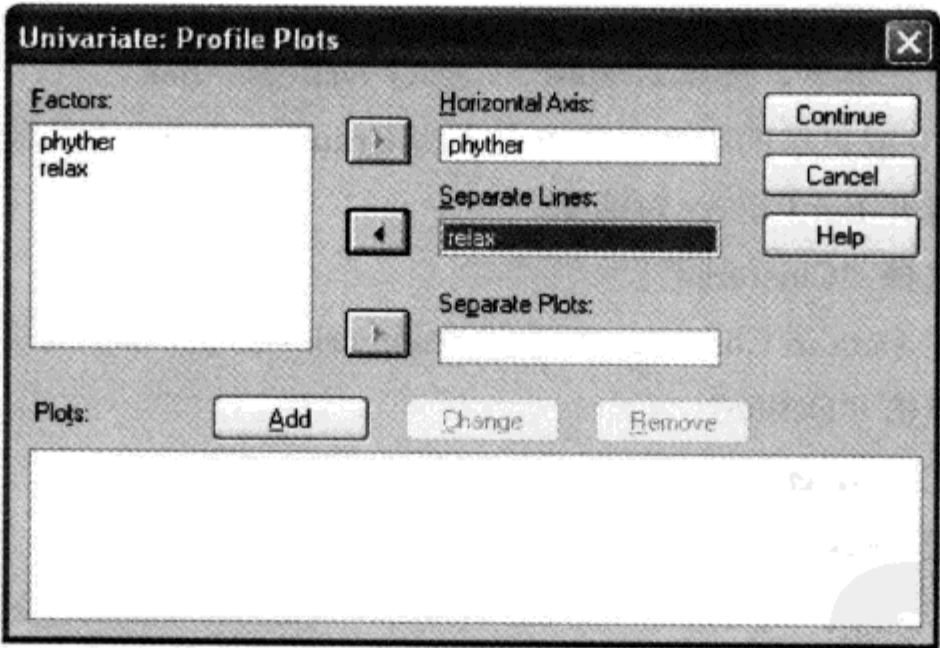


图表 9—7 Univariate 对话框 (续)



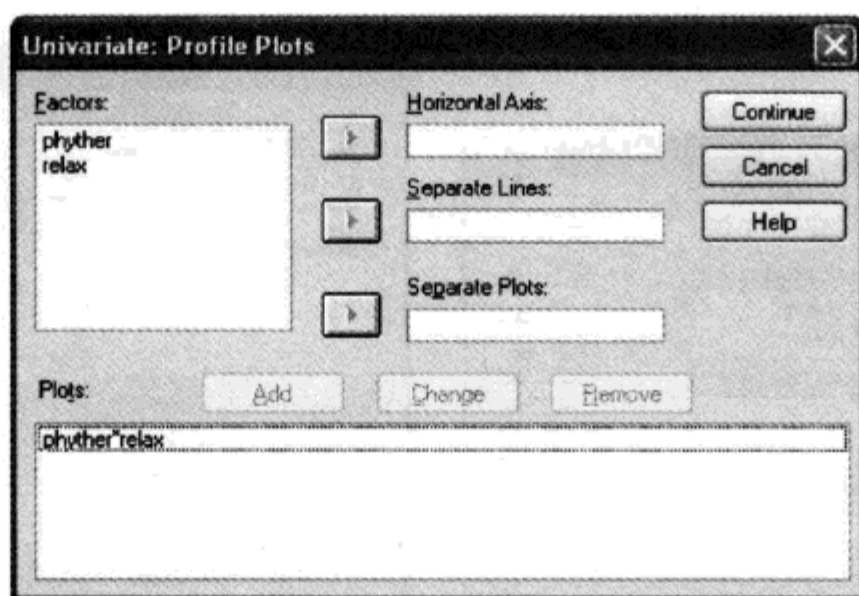
图表 9—8 Univariate: Options 对话框

点击向右箭头按钮 (►) 使它移到 Horizontal Axis 框中。选择 **relax** 并点击中间的向右箭头按钮 (►) 使它移到 Separate Lines 框中 (见图表 9—9)。



图表 9—9 Univariate: Profile Plots 对话框

7. 点击 “Add”。交互效应 **phythe**×**relax** 就显示在 Plots 对话框中 (见图表 9—10)。
8. 点击 “Continue”。
9. 点击 “OK”。



图表 9—10 Univariate: Profile Plots 对话框 (续)

在 SPSS 中运行二维组间方差分析程序, 结果显示在 Viewer 窗口中。

在讨论 ANOVA 结果之前, 将生成一个条形图。条形图是用来替代显示交互效应的轮廓图 (profile plot)。下面将给出在二维方差分析中生成条形图的命令 (详见第 3 章)。

为生成一个条形图:

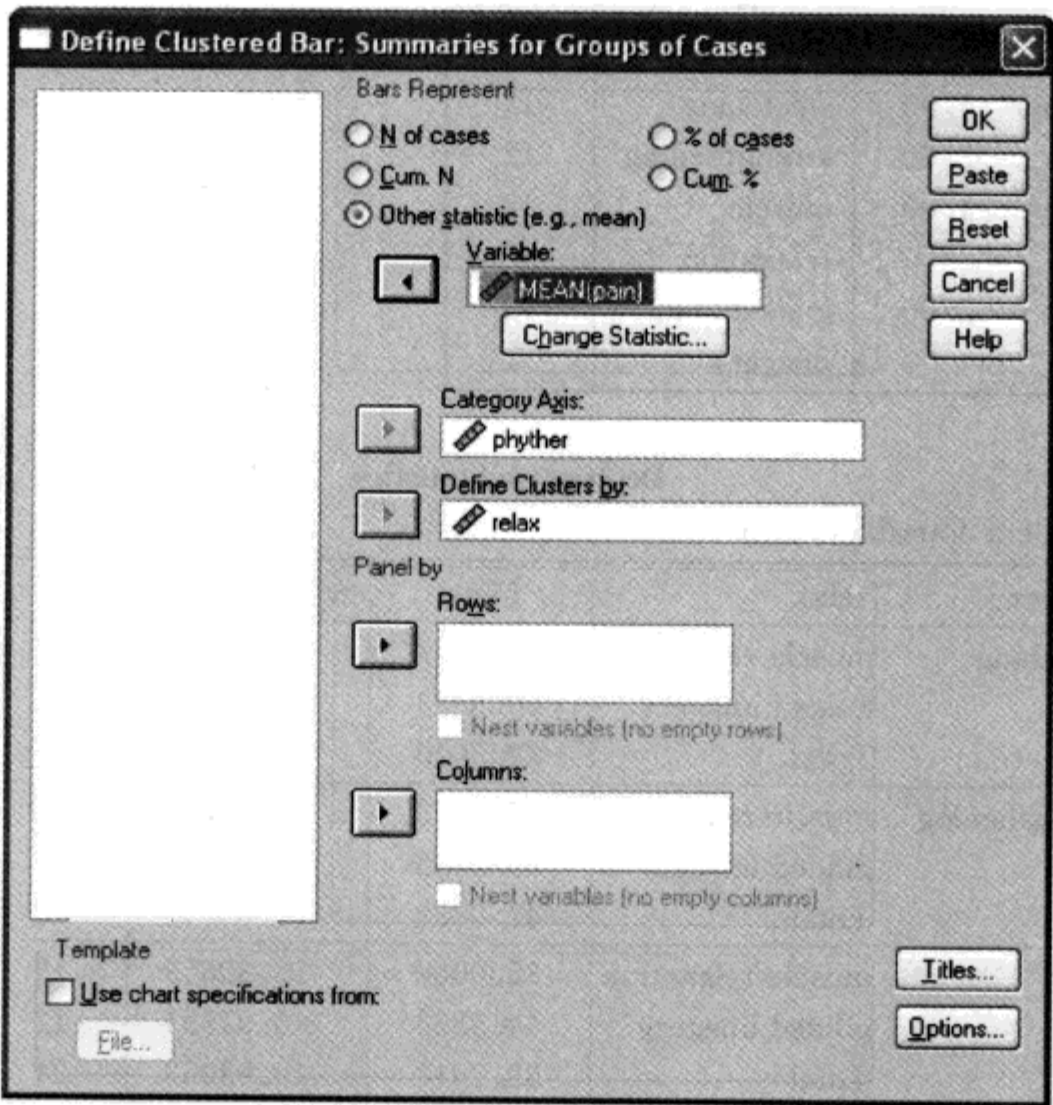
1. 打开菜单, 选择 **Graphs>Legacy Dialogs>Bar...** (注意: 如果你使用 SPSS 14.0 或更早的版本, 菜单命令为 **Graphs>Bar...**, 其他命令一样)。
2. 打开 Bar Charts 对话框。
3. 选择 “Clustered”。
4. 在 Data in Chart are 下确保 Summaries as groups of cases 被选中。
5. 点击 “Define”。
6. 把 **phyther** 移到 Category Axis 框中。
7. 把 **relax** 移到 Define Clusters by 框中。
8. 在 Bars Represent 下, 选择 Other statistic (例如, 均值), 然后把 **pain** 移到 Variable 框中 (见图表 9—11)。
9. 点击 “OK”。

在 SPSS 中运行条形图程序, 结果显示在 Viewer 窗口中。

步骤 4: 结果解释

二维组间方差分析的输出结果显示在图表 9—12 中, 条形图的结果显示

在图表 9—13 中。



图表 9—11 Define Clustered Bar: Summaries for Groups of Cases 对话框

组间因素 (Between-Subjects Factors)

Between-Subjects Factors 表显示研究的所有因素（自变量），每个因素的水平数目，变量值标签，变量每个水平的样本量。注意到 **phyther** 和 **relax** 因素的每个水平有 12 名参与者，这与我们的研究设计是一致的。

描述统计量 (Descriptive Statistics)

Descriptive Statistics 表显示研究中每种情况（和每个因素的每个水平）的均值、标准差和样本量。我们以后将集中注意 Estimated Marginal Means（估计边际均值）表的结果来解释输出结果中的组间均值差。在写出结果时，我们将使用该表中的标准差。

Univariate Analysis of Variance
Between-Subjects Factors

		Value Label	N
phyther	1.00	stretching	12
	2.00	strengthening	12
relax	1.00	muscle	12
		relaxation	
	2.00	guided imagery	12

Descriptive Statistics

Dependent Variable: pain

phyther	relax	Mean	Std. Deviation	N
stretching	muscle relaxation	24.1667	4.21505	6
	guided imagery	40.8333	7.35980	6
	Total	32.5000	10.41415	12
strengthening	muscle relaxation	45.8333	5.81091	6
	guided imagery	48.3333	4.27395	6
	Total	47.0833	5.03548	12
Total	muscle relaxation	35.0000	12.30669	12
	guided imagery	44.5833	6.94731	12
	Total	39.7917	10.93053	24

Levene's Test of Equality of Error Variances^a

Dependent Variable: pain

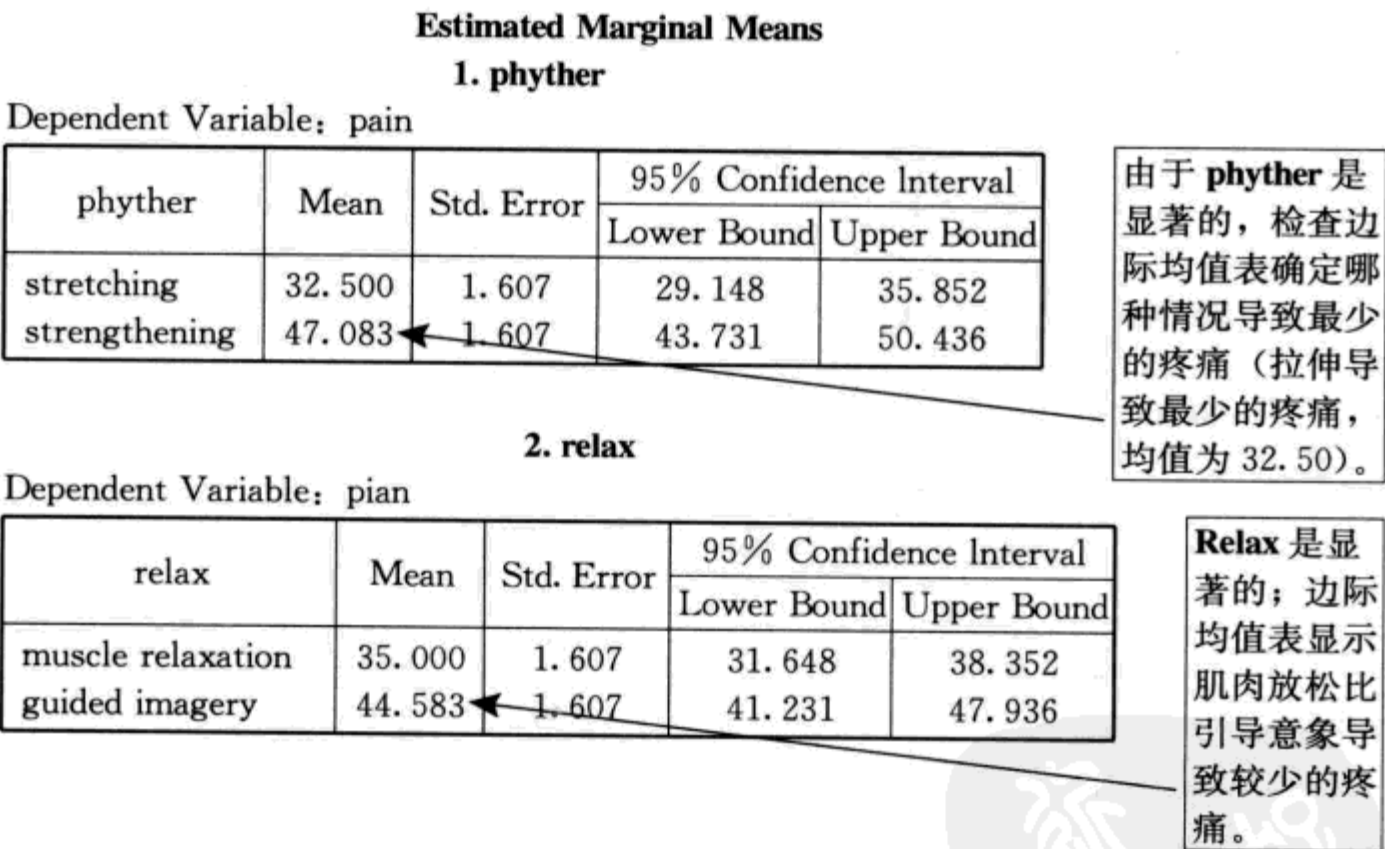
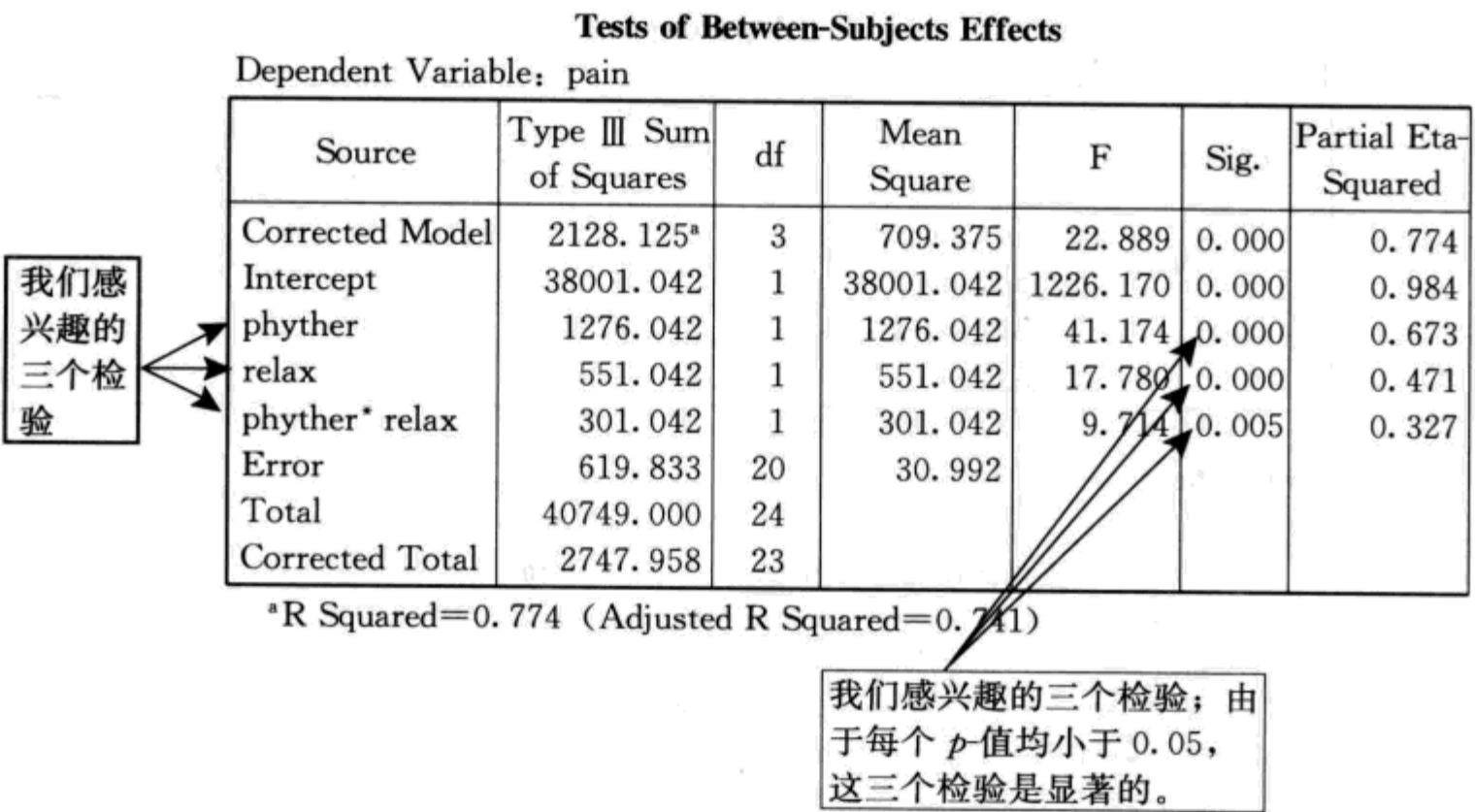
F	df1	df2	Sig.
1.238	3	20	0.322

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

^aDesign: Intercept+phyther+relax+phyther*relax

这是方差齐性检验的 p -值 (sig.), 方差分析的一个假设。由于 0.322 比 0.05 大许多, 不能拒绝原假设, 认为四个单元组的方差一致。

图表 9—12 二维组间方差分析的输出结果



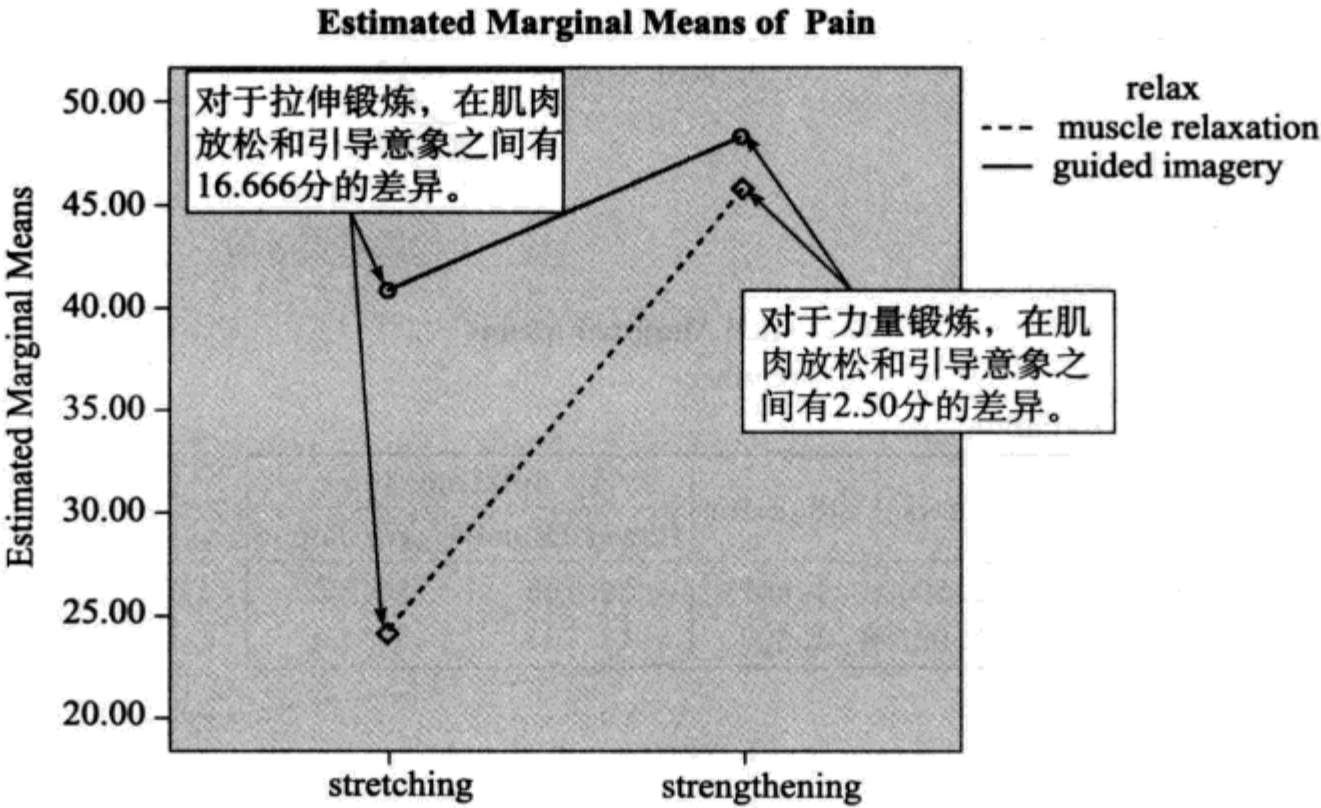
图表 9—12 二维组间方差分析的输出结果（续一）

3. phyther* relax

Dependent Variable: pian

phyther	relax	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
stretching	muscle relaxation	24.167	2.273	19.426	28.907
	guided imagery	40.833	2.273	36.093	45.574
strengthening	muscle relaxation	45.833	2.273	41.093	50.574
	guided imagery	48.333	2.273	43.593	53.074

Profile Plots



图表 9—12 二维组间方差分析的输出结果 (续二)

误差方差齐性 Levene 检验 (Levene's Test of Equality of Error Variances)

表 Levene's Test of Equality of Error Variances, 对我们研究中的四个单元 (情况) 的方差是否相等提供了检验, 这是二维组间方差分析的一个假设 (详见第 8 章方差相等的假设)。

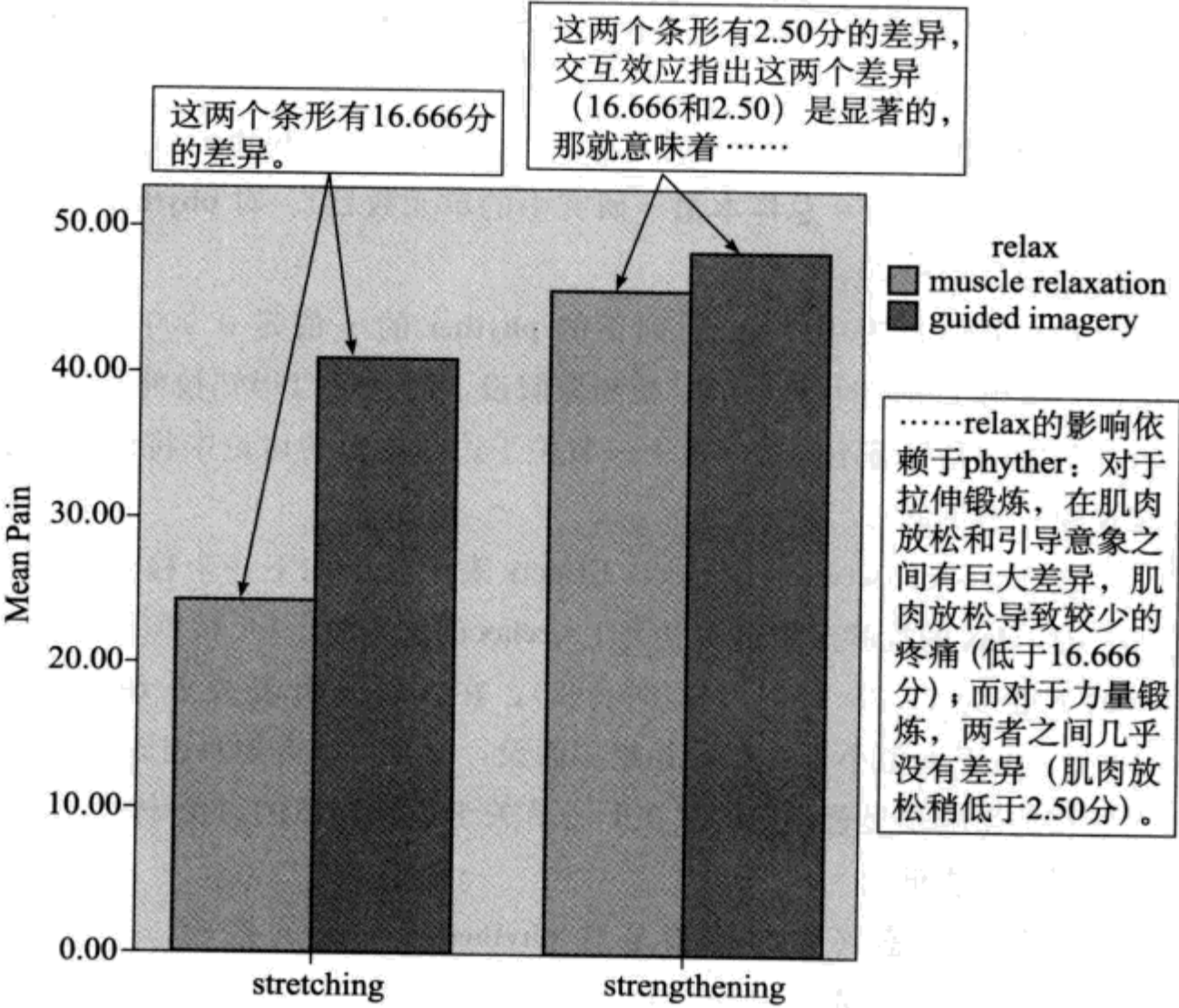
Levene 检验的原假设和对立假设为:

$H_0: \sigma^2_{1,1} = \sigma^2_{2,1} = \sigma^2_{2,2}$ (四个单元的方差总体相等)
 H_1 : 至少有一个方差与其他不等。

通过检查输出结果中表格 Levene's Test of Equality of Error Variances 中的 p 值 (sig.) 来评价方差相等假设。如果 $p \leq 0.05$, 拒绝原假设, 认为总体

方差不等。若 $p>0.05$ ，不能拒绝原假设，假设研究中的四个单元方差相等。

Levene 检验得到 F 值为 1.238， p 值为 0.322（见图表 9—12）。由于 0.322 远大于 0.05，不能拒绝方差相等的原假设，假设研究中的四个单元方差相等。



图表 9—13 条形图输出结果

组间效应检验 (Test of Between-Subjects Effects)

表 Test of Between-Subjects Effects 显示了主效应 (phyther 和 relax) 和交互效应 (phyther×relax) 的结果。² 在二维方差分析中，对每个主效应和交互效应进行了独立的 F 检验。如同第 8 章讨论的一维方差分析一样， F 检验是两个方差的比值，每个方差在输出中表示为均方 (MS)：

$$F = \frac{MS \text{ Effect}}{MS \text{ Error}}$$

式中，MS Effect 表示感兴趣检验的均方，MS Error 表示 Test of Between-

Subjects Effects 表中的均方误差值。为计算感兴趣检验的 F 值，在式中代入具体的 MS 值。例如：**phyther** 的 MS 值为 1 276.042， MS Error 为 30.992，则 **phyther** 的 F 值为：

$$F = \frac{1\,276.042}{30.992} = 41.174$$

这与图表 9—12 中的 Tests of Between-Subjects Effects 表中给出的 **phyther** 的 F 值一致。

这个检验产生了 **phyther** 的自由度 (df) (df 为 **phyther** 的水平数-1) 和误差的自由度 (df =总样本量-研究中的单元数目)。对 **phyther** 的检验，自由度分别为 1 和 20。

在方差分析表中“sig.”列下的 **phyther** 的 p -值是 0.000 (应读作小于 0.001)。由于 p -值小于 0.05，拒绝原假设，而且可以得知拉伸和力量锻炼对背伤的影响是显著不同的 (在下一节关于边际均值的讨论中将讨论对 **phyther** 来说哪一组更低)。

在 Tests of Between-Subjects Effects 表中给出的下一个检验是关于 **relax** 的。对 **relax** 的检验产生自由度为 1 (**relax** 的水平数-1) 和 20 (误差自由度) 上的 F 值为 17.78 (551.042/30.992)。在方差分析表关于 **relax** 的 p -值是 0.000。由于 p -值小于 0.05，拒绝原假设，而且可以得知肌肉放松和引导意象对背伤的影响是显著不同的 (在下一节关于边际均值的讨论中将讨论对 **relax** 来说哪一组更低)。

最后一个感兴趣的检验是对 **phyther** × **relax** 的检验，在自由度为 1 ((**phyther** 的水平数-1) × (**relax** 的水平数-1)) 和 20 (误差自由度) 上的 F 值为 9.714 (301.042/30.992)。交互检验的 p -值为 0.005，由于其低于 0.05，我们拒绝原假设，可以得知 **phyther** 和 **relax** 有显著的交互效应。

估计边际均值 (Estimated Marginal Means)

Estimated Marginal Means 产生一系列表格，分别对应每个因素和交互效应 (均值边际表将不随检验是否显著而显示)。均值边际表对显著结果的方向性有好的解释。由于三个检验都是显著的，下面每个均值边际表都将被讨论 (在一个检验不是显著的情况下，任何边际均值之间的差别将被认为是出于样本误差，而且将不被描述)。

第一个表 **phyther** 显示了两种物理锻炼方法的边际均值。由于 **phyther** 是

显著的，我们将考察平均值去看看哪一组有较低的疼痛水平。边际均值表显示拉伸情况报告的疼痛水平（32.5）显著低于力量锻炼情况（47.08）。

第二个表 **relax** 显示了两种放松锻炼方法的边际均值。由于 **relax** 是显著的，我们将考察平均值去看看哪一组有较低的疼痛水平。边际均值表显示肌肉放松情况报告的疼痛水平（35.0）显著低于引导意象情况（44.58）。

边际均值的最后一个表 **phyther**×**relax** 显示了显著交互效应的均值。研究中的四个单元均值均被显示，每组的均值代表感兴趣的单元的六个参与者的平均疼痛水平。为了更准确地确定交互效应，这个表被稍加修改，成为如图表 9—14 所示的表格。

图表 9—14 该研究中四种条件下的单元均值，两种放松锻炼方法下的拉伸、力量的均值差和放松治疗方法的边际均值

	拉伸	力量	放松锻炼的边际均值
肌肉放松	24.167	45.833	35.000
引导意象	40.833	48.333	44.583
均值差	16.666	2.500	9.583

图表 9—14 第一列的值表明对那些处于拉伸锻炼的人来说，肌肉放松带来的疼痛水平（24.167）平均比引导意象（40.833）低 16.666 分。下一列给出力量锻炼情况的均值，表明肌肉放松带来的疼痛水平仅仅比引导意象低 2.50 分（45.833 对 48.333）。这些均值差说明存在交互效应：对于拉伸锻炼，肌肉放松和引导意象之间的差距很大（16.666），而对于力量锻炼，肌肉放松和引导意象之间的差距较小（2.50）。一个明显的交互效应指出这些差距（16.666 对 2.50）说明它们之间存在显著不同。比较这四种情况下的相对平均疼痛水平，从图表 9—14 可以明显看出肌肉放松和拉伸锻炼一起进行的患者有较低的疼痛水平（24.167 是所有参与者中最低的疼痛水平）。

交互效应的图像显示

在我们通过比较所有单元的均值来检查交互效应后，很有必要画个图来显示这些结果。我们将介绍两种不同图像来表示交互效应，一个是轮廓图，另一个是条形图。图像分别见图表 9—12 和图表 9—13。

轮廓图 (Profile Plots)

轮廓图（见图表 9—12）在研究中展示四种情况的单元均值。在图像中，

Phyther 的水平在 x 轴上，拉伸在左，力量在右。图中的线表示 **relax** 的不同水平，实线代表引导意象，虚线代表肌肉放松。在图的左侧，两点的差值表示在拉伸锻炼下肌肉放松和引导意象的差异（差值是 16.666 分）。在图的右侧，两点的差值表示在力量锻炼下肌肉放松和引导意象的差异（差值是 2.50 分）。回忆交互效应，表明这两个差异（16.666 对 2.50）显著不相等。当对均值作图时，这个差异导致（明显）不平行的两条线，这是交互的另一种表现形式。

条形图

图表 9—13 中的条形图用来替代轮廓图，以描述显著交互效应的图像。轮廓图和条形图描述同样的信息——研究中的四种情况的单元均值，它们只是表示同样信息的两种不同方式。在图表 9—13 的条形图中，**phyther** 在横轴（ x 轴），**relax** 的各个水平被画成不同的条形。四种情况中的每一种的平均疼痛水平等于纵轴（ y 轴）上的条形的高度。在左边的前两个条形表明在拉伸情况下肌肉放松的疼痛得分低于引导意象的得分（差值是 16.666 分），结果表现为对拉伸情况而言，图中相邻条形的高度有显著差异。右边两个条形是关于力量锻炼的。注意到这两个条形在高度上非常接近，肌肉放松稍低于引导意象，在两种情况下疼痛的差距非常小（2.50 分）。从条形图上看，一个显著的交互效应表明 x 轴上所有水平的相邻条形的高度是不一样的（这是表示 16.666 明显不同于 2.50 的另一种方式）。

那么使用哪一种图像呢？当至少有一个自变量是区间或比例变量（即使在研究中该变量已被分类）时，我们使用轮廓图，当两个自变量是名义变量或有序变量时使用条形图。因为我们的两个自变量是名义变量（它们没有内在的顺序），所以我们使用条形图。

当交互效应显著时分析主效应

当交互效应是显著时，显著的主效应可能被误解（取决于结果的性质），因此应当小心理解结果。这一点在图表 9—14 中进行了说明。在图表 9—14 的最右边列的前两个平均值对应于 **relax** 的边际均值。**relax** 的显著主效应说明肌肉放松和引导意象是显著不同的，肌肉放松的疼痛水平比引导意象低 9.583 分（35.000 对 44.583）。这是 **relax** 主效应预测的差，可以通过把拉伸和力量

两种情况下的均值差进行平均得到 $((16.666 + 2.50) / 2 = 9.583)$ 。然而，显著的交互效应说明放松锻炼的差异依赖于物理治疗类型：对于拉伸锻炼，这个差异大于其主效应的预测值（16.666 对预测值 9.583），然而对于力量锻炼，差异较小（2.500 对预测值 9.583）。因此，使用主效应描述放松锻炼的差异错误地表示两个值，低估了拉伸锻炼的差异，高估了力量锻炼的差异。³ 然而如果没有显著交互效应，主效应将恰当地表现组之间的差异。

简单效应 (Simple Effects)

当出现显著的交互效应时，接下来可能使用简单效应分析进行进一步的效应检验。简单效应分析是在其他因素处于某个单一水平时比较一个因素的效应（例如在引导意象情形下研究拉伸和力量锻炼的差异）。简单效应分析可以在 SPSS 中使用 LMATRIX 命令来执行（详见 Green and Salkind, 2005），或者对感兴趣的检验进行独立样本 t 检验。由于篇幅所限，本章不介绍简单效应，在第 11 章有一个简单效应的例子。

效应量

二维组间方差分析的效应量通常使用偏 η^2 度量。为了计算偏 η^2 ，使用 Tests of Between-Subjects Effects 表中的平方和 (SS)（在输出结果中以 Type III Sum of Squares 形式表示）。偏 η^2 的计算公式如下：

$$\text{偏 } \eta^2 = \frac{SS_{\text{Effect}}}{SS_{\text{Effect}} + SS_{\text{Error}}}$$

SS_{Effect} 对应感兴趣效应的平方和， SS_{Error} 对应误差的平方和。为了计算给定的感兴趣的偏 η^2 ，把图表 9—12 中合适的 SS 值代入上面的公式。例如，**phyther** 的 SS 值为 1 276.04， SS_{Error} 为 619.83。把这些值代入公式得到偏 η^2 ：

$$\text{偏 } \eta^2 = \frac{1\,276.04}{1\,276.04 + 619.83} = 0.67$$

这个结果与图表 9—12 中 **phyther** 的值是一致的。

偏 η^2 的取值范围为 0~1，偏 η^2 值越大，表示因变量的方差被效应解释得越多。还没有提供常规的偏 η^2 的效应量度量。

APA 格式的结果表达

在写二维组间方差的结果时，除了报告均值和标准差（如果需要，均值

和标准差可以单独成一张表), 还要报告假设检验的结论、自由度 (df)、 F 值、 p 值和效应量。如果交互效应显著, 通常还需要提供一个条形图或者轮廓图。接下来将展示一个简单的 APA 格式的书面例子。

执行了 2×2 组间方差分析, 背伤为因变量, 物理治疗 (拉伸/力量) 和放松锻炼 (肌肉放松/引导意象) 为自变量。结果显示对物理治疗有显著主效应, $F(1, 20) = 41.17, p < 0.05$, 偏 $\eta^2 = 0.67$, 那些进行拉伸锻炼 ($M = 32.5, SD = 10.41$) 的人的疼痛显著小于那些进行力量锻炼的人 ($M = 47.08, SD = 5.04$)。对于放松锻炼也有显著的主效应, $F(1, 20) = 17.78, p < 0.05$, 偏 $\eta^2 = 0.47$, 那些进行肌肉放松 ($M = 35.0, SD = 12.31$) 的人的疼痛显著小于那些使用引导意象的人 ($M = 44.58, SD = 6.95$)。还有一个显著的物理治疗和放松锻炼的交互效应, $F(1, 20) = 9.71, p < 0.05$, 偏 $\eta^2 = 0.33$ 。对于力量锻炼, 肌肉放松和引导意象在背伤方面的影响没有多大区别, 而对于拉伸锻炼, 肌肉放松和引导意象在背伤方面的影响有着很大区别, 其中肌肉放松和拉伸锻炼一起大大减小背伤疼痛 (详见图表 9—13)。总之, 接受肌肉放松和拉伸锻炼的参与者报告的疼痛要比其他情况少很多。

二维组间方差分析的假定

1. 观测是独立的。

违反独立性假定会严重影响方差分析结果的准确性。该假定能通过设计你的研究使参与者不会以任何方式相互影响得到满足。如果我们有理由相信已经违反了独立性假定, 那么就不应再用组间方差分析。

2. 每个单元的因变量总体服从正态分布。

这个假定意味着研究中的每个单元的因变量总体应该服从正态分布。对于中等到较大的样本量, 绝大部分的非正态分布对于结果的准确性没有多少影响, 尽管一些非正态分布可能影响方差分析的功效。

3. 每个单元的总方差相等。

这个假定意味着每个单元的总方差应该相等。违反方差齐性假定会影响方差分析结果的准确性, 特别是当每组的样本量不相等时。对 SPSS 中的 Levene 检验结果的解释满足了这个假定。如果违反了方差相等的假定, 而且每个单元的样本量也不相等, 那么不能使用二维方差分析 (不像第 8 章的一

维方差分析, SPSS 目前还没有内置程序可以运行方差不等的二维方差分析)。

在 SPSS 中执行二维组间方差分析的步骤摘要

I. 数据输入和分析

1. 在 SPSS 中生成三个变量 (对应于自变量和因变量)。
2. 给每个自变量建立变量值标签。在 Value Labels 对话框中输入数值和合适的标签, 点击 “OK”。
3. 输入数据。
4. 选择 **Analyze>General Linear Model>Univariate...**。
5. 把因变量移到 Dependent Variable 框中, 把自变量移到 Fixed Factor(s) 框中。
6. 点击 “Options”。把因素和交互效应移到 Display Means for 框中。选择 Descriptive statistics, Estimates of effect size 和 Homogeneity tests, 点击 “Continue”。
7. 点击 “Plots”。把有最大水平数的因素 (如果有关) 移到 Horizontal Axis 框中, 其他因素移到 Separate Lines 框中。点击 “Add”, 点击 “Continue” (可以生成一个条形图)。
8. 如果哪个因素有三个或三个以上的水平, 点击 “Post Hoc”, 把这个变量移到 Post Hoc Tests for 框中。选择 Tukey (不是 Tukey's-b)。点击 “Continue”。
9. 点击 “OK”。

II. 结果解释

1. 检验误差方差相等的 Levene 检验结果。
 - a. 如果 Levene 检验的 $p > 0.05$, 则假设方差相等, 继续解释方差分析结果。
 - 解释每个因素的结果, 指出有两个水平的显著主效应的差异的本质。
 - 对于有三个及三个以上水平的显著主效应, 解释 Tukey 检验的结果。
 - 如果交互效应显著, 描述这两个变量关系的本质。
 - b. 如果 $p \leq 0.05$, 方差相等假设被拒绝。SPSS 没有一个容许方差不等的二维方差分析, 但如果样本量相等或近似相等, 可以考虑运行普通的二维方差分析, 并且产生相对精确的 p -值 (精确程度取决于样本量和方差的不相等程度)。如果样本量的差异是中等到较大, 则可以考虑两个独立的一维方差分析 (每个因素有三个或三个以上水平) 或者是 t 检验 (因素有两个水平), 检

验每个因素的方差是否相等（如果进行两个独立的检验，将不能检验交互效应）。对于有两个水平的因素，从表格 t-test for Equality of Means 中合适的行读结果，解释结果（详见第 6 章）。对于有三个或更多水平以及方差相等的因素，在运行普通的一维方差分析后进行 Tukey 检验。对于有三个或更多水平以及方差不相等的因素，可以使用 Brown-Forsythe 或者 Welch 检验，然后选择一个不用假定方差相等的事后检验（详见第 8 章）。

- 对于有两个水平的显著因素（如 *t* 检验），在 Group Statistics 表中检查均值来描述组间差异的本质。对于三个或更多水平的因素，对感兴趣的因素解释事后程序的结果。
- 对于不显著因素，不要描述任何组间均值差异。直接写出对感兴趣的因素组间没有明显不同的结果。

练习

1. 一个研究者想调查性别和使用手机对驾驶技术的影响。在封闭道路开车的 24 个人（12 男和 12 女）参与了该项研究，评估了他们的驾驶技术，用 0~50 分的分数进行度量，分数越高表示驾驶技术越好。在一段封闭的道路上驾驶，一半的参与者（6 男和 6 女）使用手机，另外一半没有使用。数据见图表 9—15。

输入数据并且进行适当的分析，回答下面的问题。变量分别取名为 **gender**，**cellphone** 和 **drivingscore**。

图表 9—15 练习 1 的数据

Gender	Cellphone	Driving score	Gender	Cellphone	Driving score
1	1	34	2	1	35
1	1	29	2	1	32
1	1	38	2	1	27
1	1	34	2	1	26
1	1	33	2	1	37
1	1	30	2	1	24
1	2	45	2	2	48
1	2	44	2	2	47
1	2	46	2	2	40
1	2	42	2	2	46
1	2	47	2	2	50
1	2	40	2	2	39

说明：对于性别，1 为男，2 为女；对于手机，1 为使用手机，2 为不使用手机。

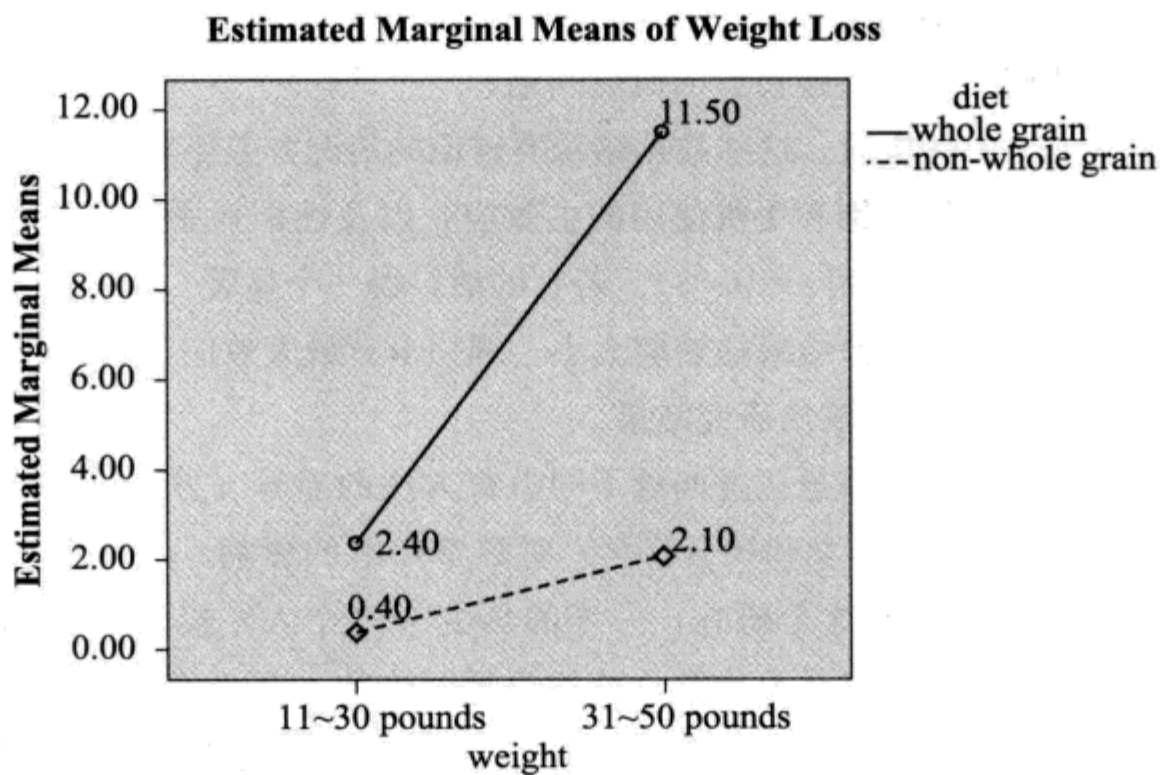
- a. 对每个感兴趣的检验陈述原假设。
- b. 描述一个感兴趣检验的研究问题。
- c. 方差齐性检验。这些数据暗示要讨论的情况方差不等?
- d. 分别检验性别和手机使用的主效应, 以及性别 \times 手机使用的交互效应 (对每一个检验, 使用 $\alpha=0.05$)。若有的话, 哪一个显著? 哪一个不显著?
- e. 输出每个检验结果的效应大小。哪一个有最大效应?
- f. 以 APA 格式输出研究结果。

2. 一个心理学家想调查两种不同教育方法的差异 (方法 A 和方法 B) 和大学的专业 (自然科学和社会科学) 对数学焦虑的影响。40 名大学生 (每个专业 20 个人) 加入这个研究, 一半的学生 (10 个人) 接受 A 方法, 另一半的学生 (10 个人) 接受 B 方法。使用合适的方法, 在 6 周之后, 调查对数学焦虑的得分。分数范围为 20~80 分, 得分越高, 焦虑越大。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 9 内的文件 Chapter 9 _ Exercise 2. sav 之中。在这个数据文件中, 变量分别取名为 **method** (1 为方法 A, 2 为方法 B), **major** (1 为自然科学, 2 为社会科学) 和 **mathanxiety** (对数学焦虑)。打开 SPSS 文件, 进行恰当分析以回答下面的问题。

- a. 对每个感兴趣的检验陈述原假设。
- b. 描述一个感兴趣检验的研究问题。
- c. 方差齐性检验。这些数据暗示要讨论的情况方差不等?
- d. 分别检验 **method**, **major** 的主效应, 以及 **method** \times **major** 的交互效应 (对每一个检验, 使用 $\alpha=0.05$)。若有的话, 哪一个显著? 哪一个不显著?
- e. 输出每个检验结果的效应大小。哪一个有最大效应?
- f. 以 APA 格式输出研究结果。

3. 图表 9—16 是一个研究饮食 (全谷类和非全谷类) 和研究开始时的超重 (超重 11~30 磅, 超重 31~50 磅) 对于减肥效果的显著交互效应的图。对于饮食, 除了一个人食用全谷类 (例如, 全麦面), 另一个人不食用全谷类 (例如, 白面食) 外, 两个人的饮食相似。因变量是研究中的减肥量, 用一个正值表示体重的减少 (在图表 9—16 中分别展示了四种情况中的均值)。

使用图像的信息, 回答下列问题。



图表 9—16 weight×diet 交互效应图像

- a. 在研究中哪种方式减掉的体重最多？哪一种最少？
- b. 描述显著的交互效应图。确保你的描述类似于目前给出的交互效应（例如，不要限制你自己去描述一个较重的人通常如何减掉更多体重，而那些体重较轻的人如何减去稍少的体重，因为这是一个主效应的描述）。



一维组内方差分析

一维组内方差分析应用于一个自变量估计一个连续因变量的情况。在一维组内方差分析中，自变量是一个包含两个或更多水平的组内因素，而且每个参与者接受自变量的所有水平。¹ 下面给出一个一维组内方差分析的例子。

例子

一位学校心理学者想评估一个全新的社交能力训练计划的潜在效果。这个计划旨在教孩子们与其伙伴建立良好的朋友关系。25 名有交友障碍的四年级学生参加这次为期 16 周的训练。在社交能力训练之前、第 8 周时和结束时测量这 25 名孩子的社交能力（由老师排名）。社交等级被用来表示学生的交友能力。这个等级分数为 6~60 分不等，当然分数越高，表示社交能力越强。在该研究中，自变量是时间（开始时、第 8 周、第 16 周），因变量是社交能力得分。

一维组内方差分析的目标和数据要求

一维组内方差分析		
目标	数据要求	例子
检验两个及更多相关组的均值关于同一个因变量是否差异显著	自变量 • 有两个及更多水平组内因素 因变量 • 连续变量	自变量 • 时间（开始，8 周，以后） 因变量 • 社交能力得分

原假设和对立假设

原假设指出在各个时刻的平均社交能力得分在总体上是一样的：

$$H_0: \mu_{\text{之前}} = \mu_{\text{8周}} = \mu_{\text{之后}}$$

为了说明原假设是错误的，没有必要指出这三个均值之间是互不相同的（当然这也是一种情况），只要说明原假设在某一点是错误的（例如，这些组之间某一些不同）。因此，与其产生原假设错误的所有可能方式（但随着组数增多，这样比较烦琐），不如找到一个不同即可。

$$H_1: \text{至少有一个均值与其他两个均值不同}$$

原假设评价

一维组内方差分析对三次总体均值是一致的原假设进行了检验。如果检验产生的结果在原假设正确时看起来不可能（结果发生的可能性小于 5%），那么拒绝原假设。如果检验产生的结果在原假设正确时看起来正确（结果发生的可能性大于 5%），那么不拒绝原假设。

研究问题

在研究中基本的兴趣问题也可以用研究问题的方式表示，例如，
“在 16 周社交能力训练之前、之中和之后的社交能力有差异吗？”

数据

在图表 10—1 中列出了 25 个参与者的数据。开始时，第 8 周和之后这三个时刻分别对应着社交能力训练之前、之中和之后的社交能力得分。

图表 10—1 一维组内方差分析例子的数据

Participant	Before	8 weeks	After	Participant	Before	8 weeks	After
1	20	24	26	5	24	25	30
2	25	26	25	6	30	28	31
3	28	31	30	7	18	22	25
4	18	16	17	8	14	17	21

续前表

Participant	Before	8 weeks	After	Participant	Before	8 weeks	After
9	12	15	16	18	31	28	31
10	9	11	10	19	14	14	18
11	15	15	19	20	13	10	14
12	17	16	20	21	15	14	16
13	35	32	31	22	31	32	32
14	32	30	31	23	35	32	36
15	34	32	36	24	28	29	31
16	22	25	29	25	26	25	25
17	16	19	22				

说明：变量 participant 包含在数据中，但不用输入 SPSS。

在 SPSS 中输入数据及分析

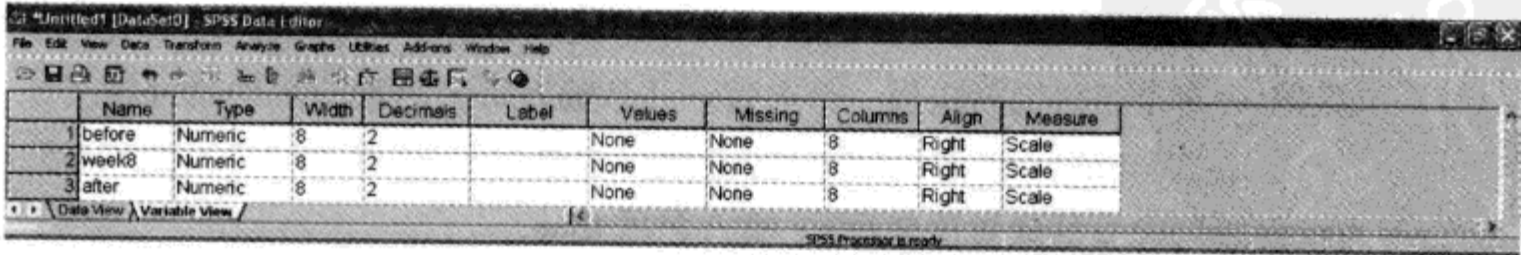
下面的步骤 1 和步骤 2 描述了怎样在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 10 中的文件 social skills.sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中将生成三个变量，将测量每一个变量对应不同时刻社交能力得分。这些变量将被分别命名为 before, week8 和 after。

- 3. 参照图表 10—2，在第一个 Variable View 窗口前三行分别输入变量名称 before, week8 和 after。



图表 10—2 输入变量 before, week8 和 after 的 Variable View 窗口

步骤 2：输入数据

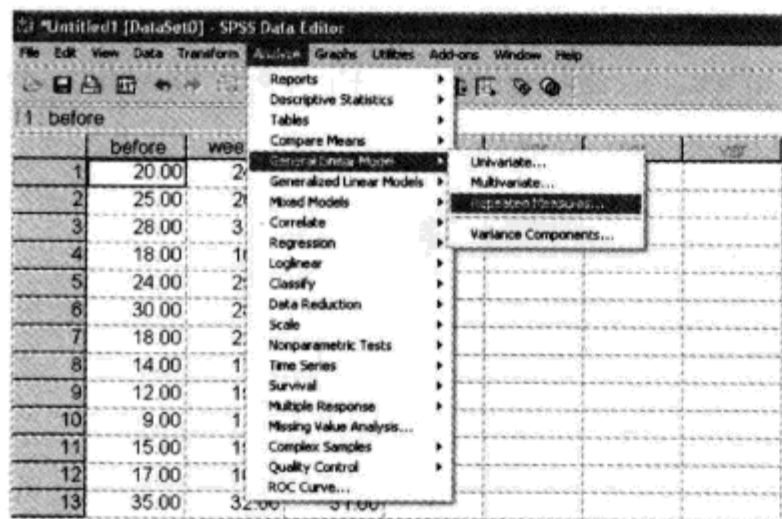
- 1. 点击 Data View 标签。变量 **before**, **week8** 和 **after** 出现在 Data View 窗口的前三列。
- 2. 参照图表 10—1，为每个参与者输入三个变量的数据。对第一个参与者，为变量 **before**, **week8** 和 **after** 分别输入数据 20, 24 和 26。依次输入全部 25 个参与者的数据。图表 10—3 中给出了完整的数据文件。

	before	week8	after						
1	20.00	24.00	26.00						
2	25.00	26.00	25.00						
3	28.00	31.00	30.00						
4	18.00	16.00	17.00						
5	24.00	25.00	30.00						
6	30.00	28.00	31.00						
7	18.00	22.00	25.00						
8	14.00	17.00	21.00						
9	12.00	15.00	16.00						
10	9.00	11.00	10.00						
11	15.00	15.00	19.00						
12	17.00	16.00	20.00						
13	35.00	32.00	31.00						
14	32.00	30.00	31.00						
15	34.00	32.00	36.00						
16	22.00	25.00	29.00						
17	16.00	19.00	22.00						
18	31.00	28.00	31.00						
19	14.00	14.00	18.00						
20	13.00	10.00	14.00						
21	15.00	14.00	16.00						
22	31.00	32.00	32.00						
23	35.00	32.00	36.00						
24	28.00	29.00	31.00						
25	26.00	25.00	25.00						

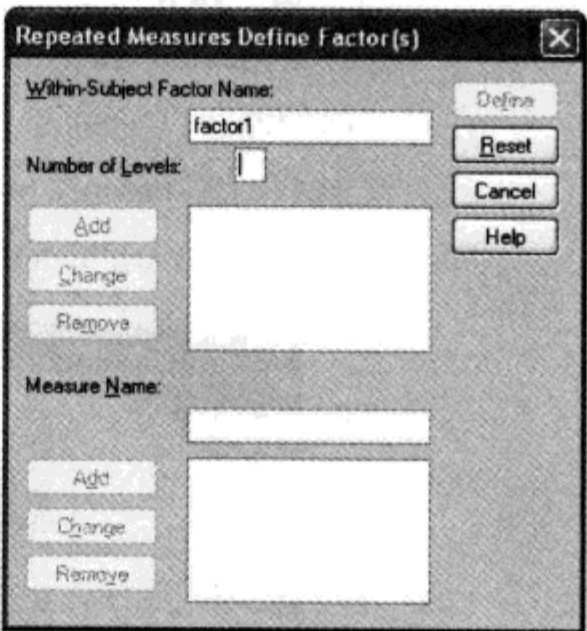
图表 10—3 一维组内方差分析例子的数据文件

步骤 3：分析数据

- 1. 从菜单栏中选择 **Analyze>General Linear Model>Repeated Measures...** (见图表 10—4)。
打开 Repeated Measures Define Factor (s) 对话框 (见图表 10—5)。这个对话框为每个组内因素提供一个名称，并且输入每个因素的水平数目。
- 2. 在 Within Subject Factor Name 框中双击 “factor1” (factor1 是 SPSS 中为组内因素建立的缺省值名称)，输入名称 **time**。

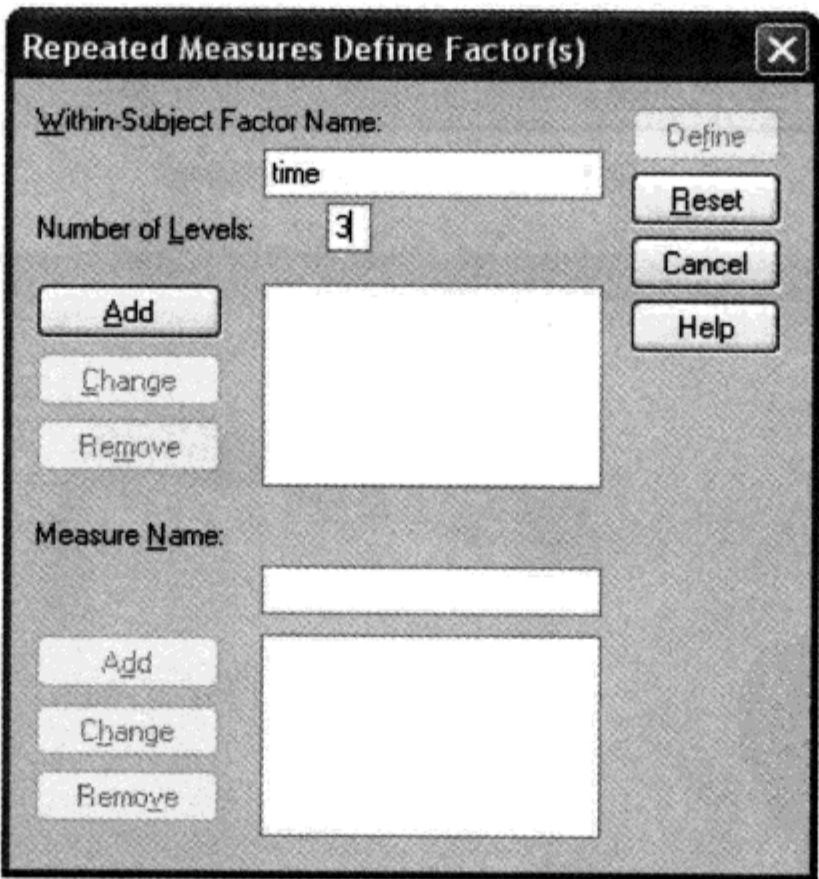


图表 10—4 一维组内方差分析的菜单命令



图表 10—5 重复测量定义因素对话框

3. 在 Number of Levels 右边的文本框中输入 3。这主要是对应于组内因素的三个水平（before, week8, after），详见图表 10—6。



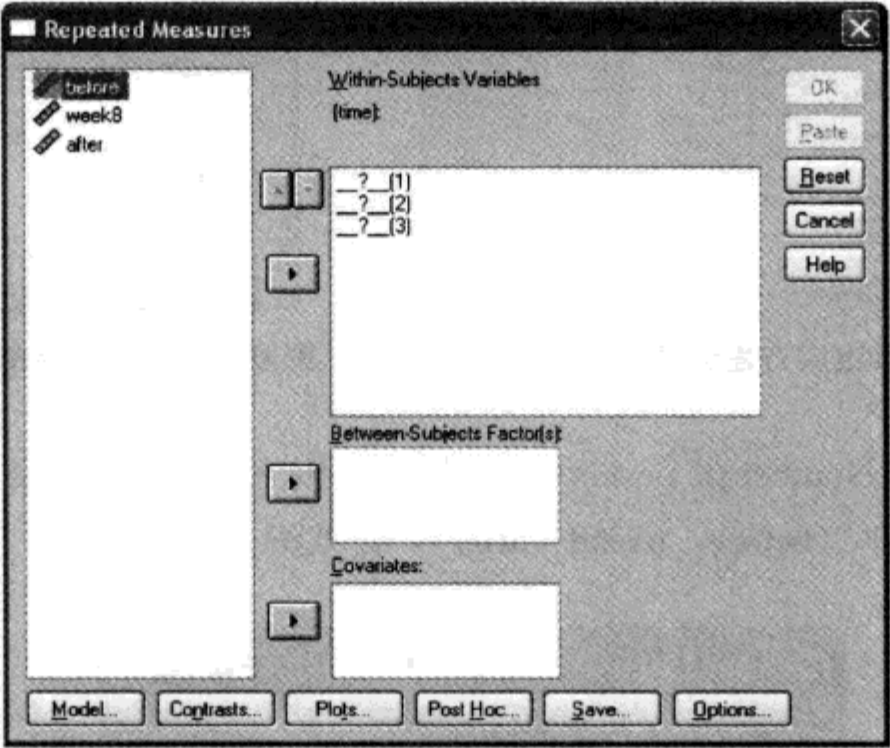
图表 10—6 重复测量定义因素对话框（续）

4. 点击 “Add”。

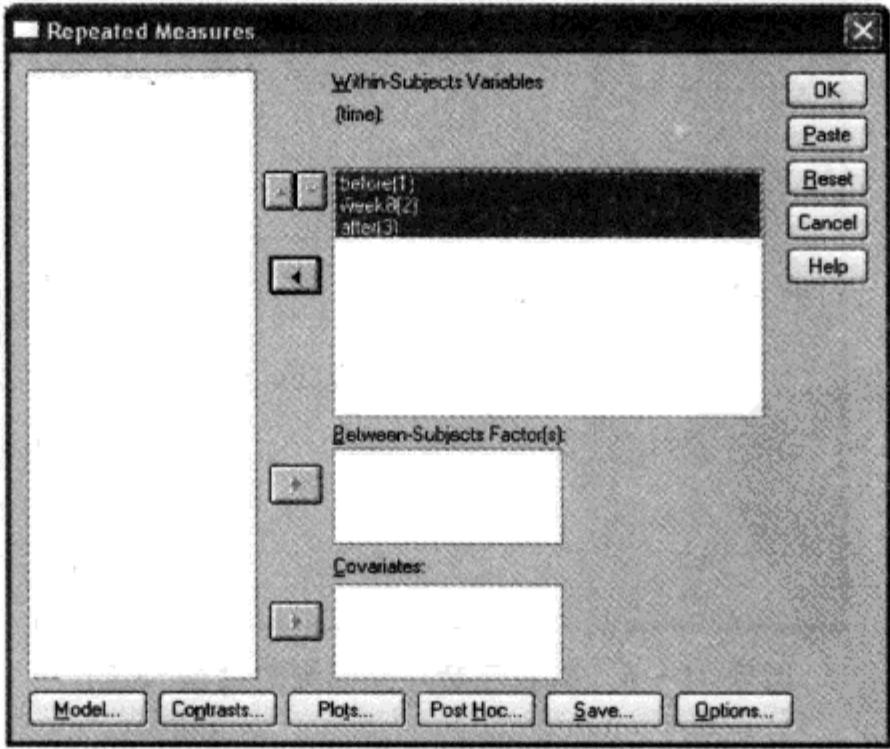
5. 点击 “Define”。

打开 Repeated Measures 对话框，三个时刻（before, week8, after）出现在对话框的左边（见图表 10—7）。

6. 选择时刻 before, week8 和 after。点击向右箭头按钮（▶）使它们移到 Within Subjects Variables 框中（见图表 10—8）。²

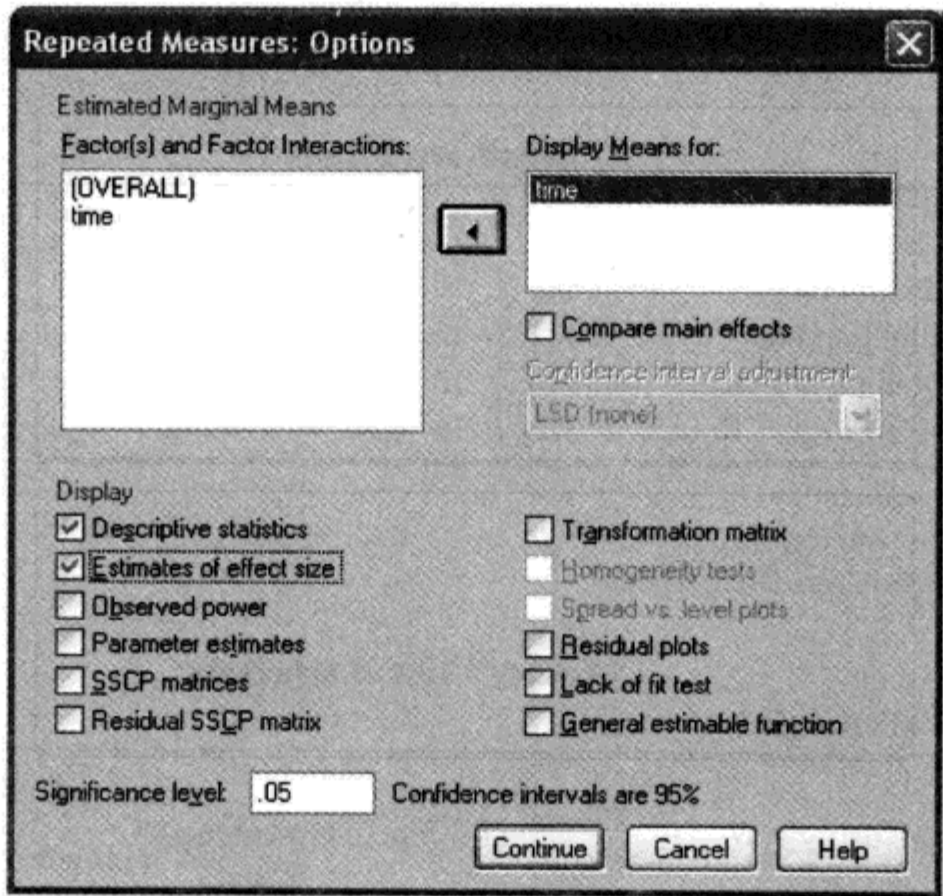


图表 10—7 重复测量对话框



图表 10—8 重复测量对话框（续）

- 7. 点击“Options”。在 Factor (s) and Factor Interactions 下选择 **time** 并点击向右箭头按钮 (►) 使它们移到 Display Means for 框中。在 Display 下, 选择 Descriptive statistics and Estimates of effect size (见图表 10—9)。
- 8. 点击“Continue”。
- 9. 点击“OK”。



图表 10—9 重复测量：选项对话框

在 SPSS 中运行一维组内方差分析程序，结果显示在 Viewer 窗口中。

步骤 4：结果描述

一维组内方差分析的输出结果显示在图表 10—10 中。

组内因素 (Within-Subjects Factors)

第一张表格 Within-Subjects Factors 显示了度量参与者的三个时刻，包括训练之前 (before)，训练 8 周 (week8) 和训练结束时 (after)。

General Linear Model

Within-Subjects Factors
Measure: MEASURE_1

time	Dependent Variable	Descriptive Statistics		
		Mean	Std. Deviation	N
1	before	22.4800	8.10411	25
2	week8	22.7200	7.35142	25
3	after	24.8800	7.20139	25

Multivariate Tests^b

Effect	Value	F	Hypothesis df	Error df	Sig.
time Pillai's Trace	0.569	15.155 ^a	2.000	23.000	0.000
Wilks' Lambda	0.431	15.155 ^a	2.000	23.000	0.000
Hotelling's Trace	1.318	15.155 ^a	2.000	23.000	0.000
Roy's Largest Root	1.318	15.155 ^a	2.000	23.000	0.000

^aExact statistic
^bDesign: Intercept
Within Subjects Design: time

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^a		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
time	0.779	5.751	2	0.056	0.819	0.871	0.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.
^aMay be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.
^bDesign: Intercept Within Subjects Design: time

图表 10—10 一维组内方差分析的输出结果

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
time	Sphericity Assumed	87.360	2	43.680	14.298	0.000	0.373
	Greenhouse-Geisser	87.360	1.638	53.343	14.298	0.000	0.373
	Huynh-Feldt	87.360	1.741	50.166	14.298	0.000	0.373
	Lower-bound	87.360	1.000	87.360	14.298	0.001	0.373
Error (time)	Sphericity Assumed	146.640	48	3.055			
	Greenhouse-Geisser	146.640	39.305	3.731			
	Huynh-Feldt	146.640	41.794	3.509			
	Lower-bound	146.640	24.000	6.110			

检验用的 p -值显示在这三个时间段的社交得分是一致的。使用 Greenhouse-Geisser 方法的值或者使用球形假设值 (两者的 p -值均小于 0.05)。

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
time	Linear	72.000	1	72.000	16.457	0.000	0.407
	Quadratic	15.360	1	15.360	8.853	0.007	0.269
Error (time)	Linear	105.000	24	4.375			
	Quadratic	41.640	24	1.735			

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	40 926.720	1	40 926.720	247.336	0.000	0.912
Error	3 971.280	24	165.470			

Estimated Marginal Means

time

Measure: MEASURE_1

time	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	22.480	1.621	19.135	25.825
2	22.720	1.470	19.685	25.755
3	24.880	1.440	21.907	27.853

图表 10—10 一维组内方差分析的输出结果 (续)

描述统计量 (Descriptive Statistics)

表格 Descriptive Statistics 显示三个时刻中的每一个时刻的均值和标准差。注意到社交得分最低的是对应 **before** 的 22.48 分，其次是对应 **week8** 的 22.72 分，对应 **after** 的 24.88 分。我们很快会考虑这些均值的差异是否大到足以具有统计显著性。

多元检验 (Multivariate Tests)

三个时刻的均值在总体上相等的原假设可以使用单变量检验 (ANOVA) 或者多变量检验 (MANOVA)。Multivariate Tests 表提供了四种不同的多变量检验的结果。因为多元检验的结果可以自动给出，而且多元检验的程序已经超出本章研究内容，所以我们这里不讨论。感兴趣的读者可以参考 Maxwell and Delaney (2004) 或者 Stevens (2002) 以了解详细信息。

Mauchly 球形检验 (Mauchly's Test of Sphericity)

表格 Mauchly's Test of Sphericity 提供了对球形假设的检验，当组内因素有三个或更多水平时³，这是组内方差分析的一个假设。该假设的检验是由 Mauchly's Test of Sphericity 表提供的 (检验的 p -值是 0.056)。这个检验是不精确的 (详细信息可以参考 Howell (2007) 或者 Maxwell and Delaney (2004))，因此不必考虑 (将提供一个替代的方法来检验球形假设)⁴。

如果球形假设不满足，则标准方差分析 F 检验 (在 Tests of Within-Subjects Effects 表中记为 Sphericity Assumed) 将不准确，产生的结果将比保证的更多地拒绝一个真实的原假设。当违反球形假设时，由于 F 检验的不精确性，有几个替代的 F 检验被调整适用于缺乏球形假设的情况。在 Tests of Within-Subjects Effects 表中提供了三个替代程序：Greenhouse-Geisser, Huynh-Feldt 和 Lower-bound。在违反球形假设时，每个方法都试图调整标准方差分析 F 检验的不精确性。由于 Mauchly 球形检验不精确，我们绕过这个方法，在 SPSS 中小心地选择一个调整方法来得到结果 (也就是，我们已经假设球形假设被违背，所以这样做)，在这三种方法中，Lower-bound 方法由于很少拒绝原假设，显得过于谨慎 (称之为保守)，因此不推荐这种方法，只有在剩下的两种方法中选择。两者都可以提供相对精确的效果，但 Greenhouse-Geisser 方法略显保守 (它经常比确保的要较少拒绝原假设)，而

Huynh-Feldt 方法稍微自由些（经常拒绝原假设）。我们将采用更严谨的方法并且解释 Greenhouse-Geisser 调整方法（尽管这将在功效上有点儿损失）。当违反球形假设时，Greenhouse-Geisser 调整方法提供了一个比标准（球形假设） F 值检验更精确的结果，由于手算比较烦琐，只有依靠计算机帮助。因此，那些想用 SPSS 确认手算结果的人将使用球形假设（Sphericity Assumed）值。

组内效应检验 (Tests of Within-Subjects Effects)

表格 Tests of Within-Subjects Effects 给我们的研究问题提供了答案，即三个时刻的社交能力得分是否不同。方差分析产生一个 F 检验， F 检验是两个方差的比值，每个方差在输出中表示为均方 (MS)：

$$F = \frac{MS \text{ Time}}{MS \text{ Error (time)}}$$

把表 Tests of Within-Subjects Effects 中在 Greenhouse-Geisser 行的值代入公式得到 F 值为：

$$F = \frac{53.343}{3.731} = 14.298$$

这与方差分析表中对于 Greenhouse-Geisser 解的 F 值一致（球形假设的 F 值 (Sphericity Assumed F) 也为 14.298⁵；作为一个提醒，如果你想把这个结果与手算的作比较，将使用球形假设值）。

对于 Sphericity Assumed 值，方差检验产生了 **time** 的自由度 (df = 水平数 - 1) 和误差的自由度 (df = (总样本量 - 1) × (水平数 - 1))，自由度分别为 2 和 48。为获得 Greenhouse-Geisser 方法的自由度，一个调整的方法是应用球形假设自由度来补偿数据中球形的缺乏。事实上，如果你看前面的表 Mauchly's Test of Sphericity，你将会看到 Greenhouse-Geisser 的 epsilon 的统计值为 0.819。这个值再乘以自由度 (2, 48) 将产生（有舍入误差的）自由度 1.638 和 39.305，这些在 Greenhouse-Geisser 检验中有记录。

在“Sig.”列中的 Greenhouse-Geisser 的 p 值是 0.000（应读作小于 0.001）。由于 p 值小于 0.05，则拒绝三个时刻均值相等的原假设（使用球形假设值也可以拒绝原假设）。在讨论输出结果的其他三个表之后，将进行其他检验来检查这三个时刻差异的本质。

组内对照检验 (Test of Within-Subjects Contrasts) 和组间效应检验 (Tests of Between-Subjects Effects)

接下来的两张表 Test of Within-Subjects Contrasts 和 Tests of Between-Subjects Effects 将不会出现在一维组内方差分析的讨论中, 然而 Test of Within-Subjects Contrasts 表可以用于对 **time** 的某些进一步检验中, 我们将进行相依样本 t 检验, 因为它们不需要球形假设。

Tests of Between-Subjects Effects 表给出了研究中任一组间因素的结果。因为我们在研究中不使用组间因素, 所以这个表在这里不相关 (将在第 11 章讨论该表)。

估计边际均值 (Estimated Marginal Means)

Estimated Marginal Means 表对于三个时刻分别提供了均值、标准误差和 95% 置信区间。虽然前面的 Descriptive Statistics 表也提供了每组的均值, 两个表之间的区别是 Estimated Marginal Means 表还提供了每个组的标准误差和 95% 置信区间, 而 Descriptive Statistics 表则提供标准差和样本大小。

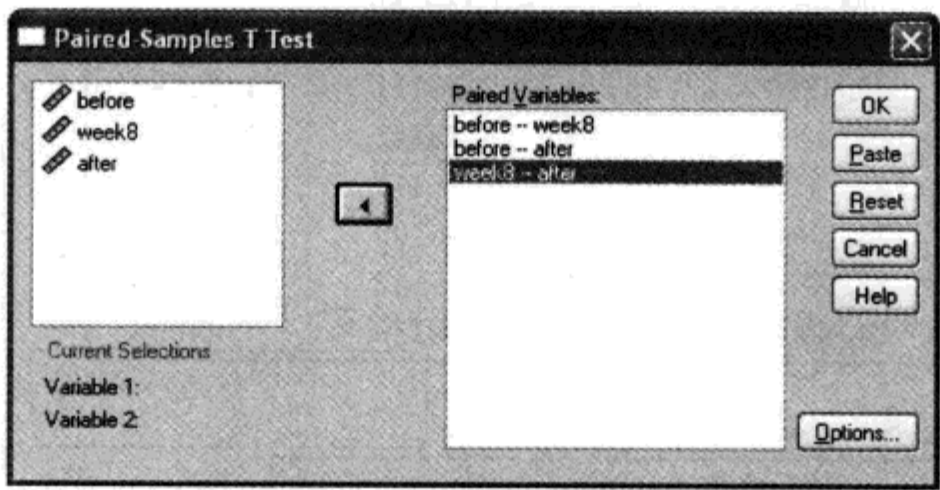
在总体方差分析得到显著结果之后的进一步检验

由于 **time** 的检验是显著的 (复习一下 Tests of Within-Subjects Effects 部分), 关于三个时刻的社交能力得分相等的原假设被拒绝。然而由于对立假设不明确, 仅仅说明均值在某些程度上不一致, 需要进一步检验来确定哪一个时刻显著不同于其他时刻。我们进一步检验采取的方法类似于第 8 章的一维组间方差分析。回想一下组间方差分析, 如果全部检验是显著的, 所有的配对比较将使用 Tukey 检验法来检验。对于组内方差分析, 所有的配对比较将使用相依样本 t 检验, 这是组内方差分析的一个合适的进一步程序。这将导致 3 个配对 t 检验: before 对 week8, before 对 after 和 week8 对 after。在 SPSS 中进行相依样本 t 检验的命令总结如下 (相依样本 t 检验的详细内容见第 7 章)。

进行三个相依样本 t 检验:

1. 从菜单栏中选择 **Analyze > Compare Means > Paired-Samples T Test...**。
2. 打开 Paired-Samples T Test 对话框, 选择变量 **before** 和 **week8**, 点击向右箭头按钮 (►) 使它们移到 Paired Variables 框中。
3. 对于 **before** 和 **after**、**week8** 和 **after** 重复这个操作。所有三对变量应当

出现在 Paired Variables 框中 (见图表 10—11)。



图表 10—11 配对样本 t 检验对话框

4. 点击 “OK”。

在 SPSS 中运行配对样本 t 检验程序，结果显示在 Viewer 窗口中。独立样本 t 检验的结果见图表 10—12。

T 检验：三个时刻的配对比较

在对所有可能的配对比较进行 t 检验时，检验下面的三个原假设：

$$H_0: \mu_{\text{之前}} = \mu_{8\text{周}}$$

$$H_0: \mu_{\text{之前}} = \mu_{\text{之后}}$$

$$H_0: \mu_{8\text{周}} = \mu_{8\text{之后}}$$

在进行进一步检验时一个习惯做法是调整每个检验的 α 水平使得整个进一步检验的 α 值不超过 0.05（每次检验使用的 α 值被当作每次比较的 α 水平）。为了求出每次检验的 α 水平，设置整个检验的总体 α 水平（0.05），然后再除以进一步检验执行的次数。对于三次检验以及 α 水平为 0.05，每次比较的 α 水平是 $0.05/3=0.016$ （只舍不入是为了保证三个检验的总和不超过 0.05）。这就意味着代替了每次用 α 水平 0.05 来评价检验，而是用每次检验的 p -值与 α 水平 0.016 相比较。如果 p -值小于或等于 0.016，则拒绝感兴趣比较的原假设；如果 p -值大于 0.016，则不能拒绝原假设。这样将保证整个进一步的检验犯第一类错误的概率不超过 0.05（如果不这样调整，每个进一步检验都用 0.05 来执行，那么一系列检验的总体 α 水平将接近三个检验的 α 水平之和，即 0.15）。

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	before	22.4800	25	8.10411	1.62082
	week8	22.7200	25	7.35142	1.47028
Pair 2	before	22.4800	25	8.10411	1.62082
	after	24.8800	25	7.20139	1.44028
Pair 3	week8	22.7200	25	7.35142	1.47028
	after	24.8800	25	7.20139	1.44028

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	before & week8	25	0.956	0.000
Pair 2	before & after	25	0.932	0.000
Pair 3	week8 & after	25	0.964	0.000

Paired Samples Test

		Paired Differences					t	df	sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	before-week8	−0.24000	2.40278	0.48056	−1.23182	0.75182	−0.499	24	0.622
Pair 2	before-after	−2.40000	2.95804	0.59161	−3.62102	−1.17898	−4.057	24	0.000
Pair 3	week8-after	−2.16000	1.95107	0.39021	−2.96536	−1.35464	−5.535	24	0.000

对于配对比较 before 和 after, week8 和 after 都是显著的, 因为 p -值小于 0.016 (但 before 和 week8 之间不显著, 因为其 p -值为 0.622, 大于 0.016)。

图表 10—12 独立样本 t 检验的输出结果

三个配对比较的结果在 Paired Samples Test 表中显示。这些结果指出第一个检验 before 和 week8 之间不显著, 这是由于其 p -值为 0.622, 大于 0.016。其余两个检验 before 和 after 以及 week8 和 after 都是显著的, p -值为 0.000 ($p < 0.001$)。由于后两者差异显著, 我们将考察在 Paired-Samples Statistics 中的均值来确定哪一个时刻有较高的社交能力得分。Paired-Samples Statistics 表显示 after 时刻的社交能力得分(24.88)明显高于 before 时刻(22.48)和 week8 时刻(22.72)。

配对比较的结果总结在图表 10—13 中。

图表 10—13 三个配对比较结果的总结

检 验	组之间是否有显著差异 (<i>p</i> -值是否小于或等于 0.016)?	结 论
Before 对 8 weeks	没有	训练之前和训练 8 周的社交能力得分没有显著差异
Before 对 after	有	显著; 孩子们在训练结束时的社交能力得分高于训练之前的社交能力得分
8 weeks 对 after	有	显著; 孩子们在训练结束时的社交能力得分高于训练 8 周时的社交能力得分

效应量

一维组内方差分析的效应量通常使用偏 η^2 度量。为了计算偏 η^2 ，使用表格 Tests of Within-Subjects Effects 中的平方和（在输出中称为 Type III Sum of Squares）。偏 η^2 的计算公式如下：

$$\text{偏 } \eta^2 = \frac{SS_{\text{Effect}}}{SS_{\text{Effect}} + SS_{\text{Error}}}$$

对于 **time**， SS_{Effect} 和 SS_{Error} 的值在图表 10—10 中的 Tests of Within-Subjects Effects 表中为 SS_{Time} 和 $SS_{\text{Error}(\text{time})}$ 。从图表 10—10 中代入适当的值得到 η^2 为

$$\text{偏 } \eta^2 = \frac{87.360}{87.360 + 146.640} = 0.373$$

这个结果与表格 Tests of Within-Subjects Effects 中对 **time** 报告的值是一致的。

还没有提供常规的对偏 η^2 的小、中和大效应量的度量方法。偏 η^2 值越大，因变量的方差被效应解释得越多。

APA 格式的结果表达

在写一维组内方差分析的结果时，除了报告每个时刻社交能力得分的均

值和标准差（如果需要，均值和标准差可以单独成一张表），还要报告假设检验的结论，自由度（ df ）， F 值， p -值，效应量和配对比较的结果（假定总体方差分析是显著的）。接下来将展示一个简单的 APA 格式的书面例子。

儿童的社交能力的得分在各个时刻是显著不同的，Greenhouse-Geisser 调整 $F(1.64, 39.31) = 14.30, p < 0.05$ ，偏 $\eta^2 = 0.37$ 。独立样本 t 检验用来说明哪一个时刻是不同于其他的，每个检验的 α 水平为 0.016。结果表明在训练结束时的社交能力得分（ $M = 24.88, SD = 7.20$ ）明显高于训练 8 周时的得分（ $M = 22.72, SD = 7.35$ ）， $t(24) = -5.54, p < 0.016$ ，而且比较了训练结束时的得分与训练开始时的得分（ $M = 22.48, SD = 8.10$ ）， $t(24) = -4.06, p < 0.016$ ，但是在训练开始和训练 8 周时没有显著差异， $t(24) = -0.50, p < 0.016$ 。

（注意：如果你正在使用球形假设值，用值（2，48）替换 Geisser-Greenhouse 的自由度（1.64，39.31），并且从结果中删去“Geisser-Greenhouse 调整”，其他所有结果仍然保持不变。）

一维组内方差分析的假定

1. 参与者间的观测是独立的。

违反该假定会严重影响组内方差分析结果的准确性（教师一起工作确定学生的社交能力得分的排名是违反独立假定的一个例子）。如果我们有理由相信已经违反了独立性假定，那么不能使用组内方差分析。

2. 在自变量每个水平上的因变量总体服从正态分布。

这个假定意味着社交能力对于每个研究的单元在总体上要服从正态分布（例如，社交能力得分应该在 before, week8 和 after 时刻服从正态分布）。对于中等到较大的样本量，绝大部分的非正态分布对于方差分析检验的结果的精确性没有多少影响，尽管一些非正态分布可能影响方差分析检验的功效。

3. 球形假定。

球形假定要求组内因素的所有配对水平上的差值得分的总体方差都是相等的。例如，如果你计算每个参与者组内因素的头两个水平的差异（before-week8），然后计算这些得分的方差（方差为标准差的平方），对所有水平组合再重复这个程序（before-after 和 week8-after），球形假定要求这三者的方

差一致:

$$\sigma_{\text{before-week8}}^2 = \sigma_{\text{before-after}}^2 = \sigma_{\text{week8-after}}^2$$

违反这个假定将影响方差分析的精确性, 导致原假设较多地被拒绝。使用一个替代的方法 (如 Greenhouse-Geisser) 适用于球形存在的情况, 这强调了对这个假定的违反。^① 由于多元方差分析 (MANOVA) 不需要球形假定, 这是一个可行的选择, 即使讨论它已经超出我们这节的内容 (详见 Maxwell and Delaney (2004) 或者 Stevens (2002) 关于多元方差分析的内容)。

在 SPSS 中执行一维组内方差分析的步骤摘要

I. 数据输入和分析

1. 在 SPSS 中生成数量等于组内因素水平数的变量。
2. 输入数据。
3. 选择 **Analyze > General Linear Model > Repeated Measures...**
4. 在 Within-Subject Factor Name 中输入组内因素的名称, 然后在 Number of Levels 框中输入水平数。点击 “Define”。
5. 把变量 (组内因素的所有水平) 移到 Within-Subjects Variables 框中。
6. 点击 “Options”。把组内因素移到 Display Means for 框中。在 Display 下选择 Descriptive statistics 和 Estimates of effect size, 点击 “Continue”。
7. 点击 “OK”。

II. 结果解释

通过对 Greenhouse-Geisser 下或者在 Sphericity Assumed 行中的 p -值进行检查来解释整个方差分析结果 (如果与手算结果进行比较, 可以使用球形假设值)。

a. 如果 ANOVA 是显著的 ($p \leq 0.05$), 对所有的配对比较执行独立样本 t 检验。在 α 值等于 0.05 除以配对比较检验的数量下对每个进一步的 t 检验进行评价。在你写出的结果中包括总体方差分析和独立样本 t 检验的结果。

b. 如果 ANOVA 不是显著的 ($p > 0.05$), 则立即停止 (不要进行 t 检

^① 原书该句似乎有问题。——译者注

验)。写出结果表明这些组之间没有显著差异。

练习

1. 一个学校社工想了解预防逃课计划对当地学校学生逃课数的影响，15 名学生参加了该研究，对每个孩子记录了 9 个月的逃课数，其中，每隔 3 个月记录一次数值。数据见图表 10—14。

图表 10—14 参加该计划的学生（每 3 个月）的逃课数

Month3	Month6	Month9
30	28	20
36	30	24
39	37	34
45	40	35
21	24	17
36	44	20
30	28	17
36	34	28
33	36	27
42	37	26
36	39	38
30	35	36
21	20	10
24	18	20
36	30	30

在 SPSS 中输入数据，进行必要分析以回答下列问题。变量命名为 month3，month6 和 month9。

- a. 陈述原假设和对立假设。
 - b. 陈述数据的一个研究问题。
 - c. 三个时间段内学生的逃课数有明显不同吗？在 $\alpha=0.05$ 时检验。
 - d. 总体方差分析的效应量是多少？
 - e. 如果总体方差分析是显著的，执行合适的进一步检验，调整每个检验的 α 使其一系列检验的水平 α 总和不超过 0.05。
 - f. 将研究结果写成适当的 APA 格式。
2. 一个医院心理学家想调查心理分析疗法的影响。15 个人同意参加这次

研究,集中起来进行观察。每个参与者接受为期8周的心理治疗。为了测量其疗效,分别在开始,第4周,第8周进行心理治疗的测试。健康度量的分数为10~50分,分数越高,表示疗效越好。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 10 中的文件 Chapter10 _ Exercise 2. sav 之中(变量命名为 **beforetherapy**, **fourweeks** 和 **eightweeks**)。在 SPSS 中打开文件,进行必要分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 整个治疗期间健康得分是否显著不同? 在 $\alpha=0.05$ 时检验。
- d. 总体方差分析的效应量是多少?
- e. 如果总体方差分析是显著的,执行合适的进一步检验,调整每个检验的 α 使其一系列检验的水平 α 总和不超过 0.05。
- f. 将研究结果写成适当的 APA 格式。

3. 一个市场研究员想调查提供不同的无线上网服务对在快餐店就餐的影响。调查了在提供三种不同服务(免费的无线上网服务,花钱的无线上网服务,没有无线上网服务)的快餐店吃饭的可能性。10个人(这些人至少每月吃一次快餐)参加此次调查,他们被要求用1~10表示对服务的满意程度(1表示最不满意,10表示最满意)。数据位于网站上文件夹 Chapter 10 中的文件 Chapter10 _ Exercise 3. sav 之中(变量命名为 **freewifi**, **payforwifi** 和 **nowifi**)。在 SPSS 中打开文件,进行必要分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 在不同无线上网服务类型的快餐店吃饭的可能性是否显著不同? 在 $\alpha=0.05$ 时检验。
- d. 总体方差分析的效应量是多少?
- e. 如果总体方差分析是显著的,执行合适的进一步检验,调整每个检验的 α 使其一系列检验的水平 α 总和不超过 0.05。
- f. 将研究结果写成适当的 APA 格式。



组间组内方差分析

组间组内方差分析应用于两个自变量估计一个感兴趣的因变量的情况。在组间组内方差分析中，其中一个自变量是组间因素，另一个自变量是组内因素。组间因素包含两个或更多的水平，每个参与者只接受处理的一个水平。组内因素也包含两个或更多水平，每个参与者接受处理的所有水平。¹下面给出一个组间组内方差分析的例子。

例子

一个研究者想调查为新中学老师设计的导师计划是否可以有效地减轻压力。20名新老师参加了这次研究，其中10名老师接受导师计划，另外10名为控制组（也就是不接受任何导师指导）。在这个导师计划中，这些新老师被指派了一名导师来讨论他们在刚开始做老师时遇到的一些困难（如教室管理等）。在同意参加这次研究之后，这些新老师分别在正式执教之前1周、执教4周和执教8周后记录一个压力评分。压力评分为20~80分，分数越高表示相应的执教压力越大。在这个研究中，指导（support，两水平——有导师或者没有导师）是组间因素，时间（time，三水平——计划开始前，进行4周后和进行8周后）是组内因素，而因变量是这些新老老师的压力评分。

组间组内方差分析的目标和数据要求

组间组内方差分析		
目标	数据要求	例子
检验主效应 • 有导师和没有导师的老师的压力水平有差异吗? • 计划开始前, 进行 4 周后和进行 8 周后的压力水平有差异吗?	自变量 • 一个组间因素 (有两个或更多水平) • 一个组内因素 (有两个或更多水平)	自变量 • 指导 (有导师, 没有导师) • 时间 (计划开始前, 进行 4 周后和进行 8 周后)
检验交互效应 • 指导 (有导师, 没有导师) 的影响依赖于时间 (计划开始前, 进行 4 周后和进行 8 周后) 吗?	因变量 • 连续	因变量 • 教学压力得分

原假设和对立假设

在一维组间组内方差分析中使用了三个不同的原假设。一个原假设用来检验每个自变量 (常说的主效应检验), 一个原假设用来检验两个自变量的混合效应 (常说的交互效应检验)。这些假设如下面描述。

假设 1. 对指导进行检验: 有导师对没有导师

指导的原假设指出有导师和没有导师的老师的平均压力在总体上是一样的:

$$H_0: \mu_{\text{有导师}} = \mu_{\text{无导师}}$$

对立假设指出有导师和没有导师的老师的平均压力在总体上是不同的:

$$H_1: \mu_{\text{有导师}} \neq \mu_{\text{无导师}}$$

假设 2. 对时间进行检验: 计划开始前, 进行 4 周后和进行 8 周后

时间的原假设指出计划开始前, 进行 4 周后和进行 8 周后的平均压力总体上是一样的:

$$H_0: \mu_{\text{指导前}} = \mu_{\text{4周}} = \mu_{\text{8周}}$$

当一个因素有三个或更多水平时, 对立假设指出因素的不同水平之间有

差异:

H_1 : 至少有一个总体均值与其他均值不同²

假设 3. 指导和时间的交互效应检验

原假设指出指导和时间之间没有交互效应:

H_0 : 没有指导 \times 时间的交互效应

对立假设指出两个变量之间有交互效应:

H_1 : 有指导 \times 时间的交互效应

原假设评价

组间组内方差分析对上面三个原假设进行了检验。如果检验产生的结果在原假设正确时看起来不可能(结果发生的可能性小于 5%), 那么拒绝原假设。如果检验产生的结果在原假设正确时看起来正确(结果发生的可能性大于 5%), 那么不拒绝原假设。

研究问题

在组间组内方差分析研究中基本的兴趣问题也可以用研究问题的方式表示, 例如:

对于指导

“有导师指导和没有导师指导的老师报告的压力水平是否不同?”

对于时间

“计划开始前, 进行 4 周后和进行 8 周后报告的压力水平是否不同?”

对于指导和时间的交互效应

“报告的压力水平随时间变化是否依赖于接受导师的指导?”

数据

在图表 11—1 中列出了 20 个参与者的数据。对于指导, 接受导师计划的记为 1, 没有接受导师计划的记为 2。对于时间, before, week4 和 week8 分别对应于在计划开始前、进行 4 周后和进行 8 周后的老师压力评分。

图表 11—1 组间组内方差分析的数据

Participant	Support	Before	Week4	Week8	Participant	Support	Before	Week4	Week8
1	1	40	39	35	11	2	38	44	42
2	1	45	44	42	12	2	47	45	42
3	1	42	44	44	13	2	41	48	45
4	1	38	36	30	14	2	39	41	43
5	1	46	44	38	15	2	44	44	41
6	1	40	37	25	16	2	42	39	35
7	1	42	38	35	17	2	42	48	46
8	1	39	37	29	18	2	40	46	42
9	1	35	33	31	19	2	38	45	43
10	1	43	44	38	20	2	41	40	35

说明：变量 participant 包含在数据中，但不用输入 SPSS。

在 SPSS 中输入数据及分析

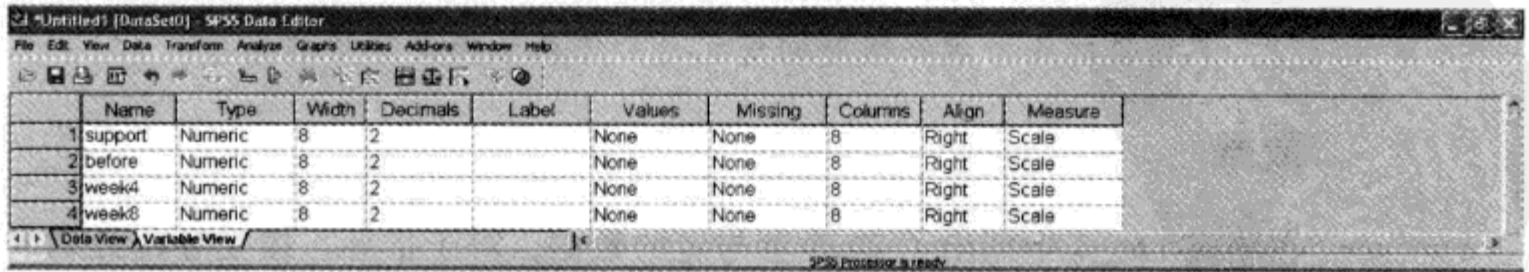
下面的步骤 1 和步骤 2 描述了怎样在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 11 中的文件 teaching stress.sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中将生成四个变量，一个是给 Support 组的（有导师、无导师），另一个是三个不同的时间段。这些变量将被分别命名为 support, before, week4 和 week8。

- 3. 在 Variable View 窗口前四行分别输入变量名称 support, before, week4 和 week8（见图表 11—2）。

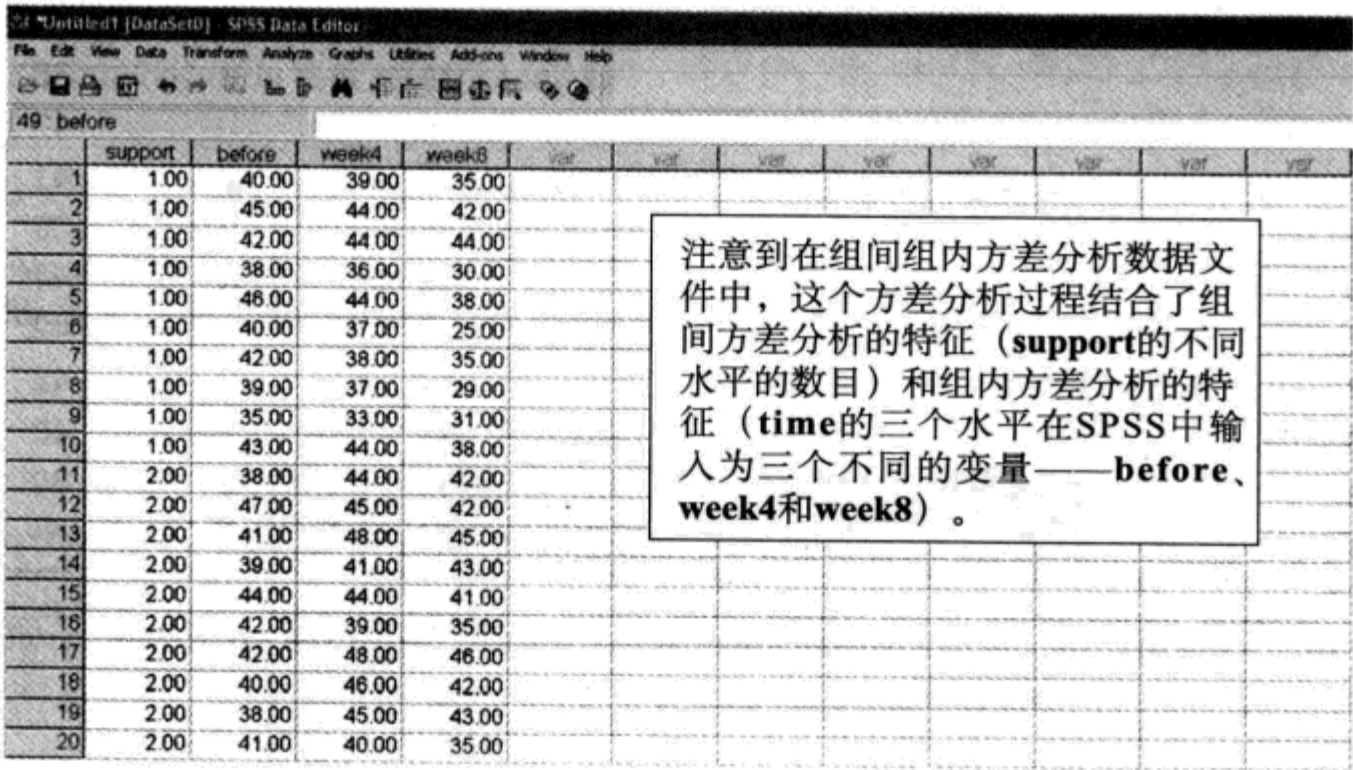


图表 11—2 输入变量 support、before、week4 和 week8 的 Variable View 窗口

4. 使用第 1 章描述的过程，为 **support** 建立变量值标签。如对于 **support**，1=“有导师”，2=“无导师”。

步骤 2：输入数据

- 1. 点击 Data View 标签。变量 **support**、**before**、**week4** 和 **week8** 出现在 Data View 窗口的前四列。
- 2. 参照图表 11—1，为每个参与者输入四个变量的数据。对第一个参与者，为变量 **support**、**before**、**week4** 和 **week8** 分别输入数据 1，40，39 和 35。依次输入全部 20 个参与者的数据。图表 11—3 给出了完整的数据文件。

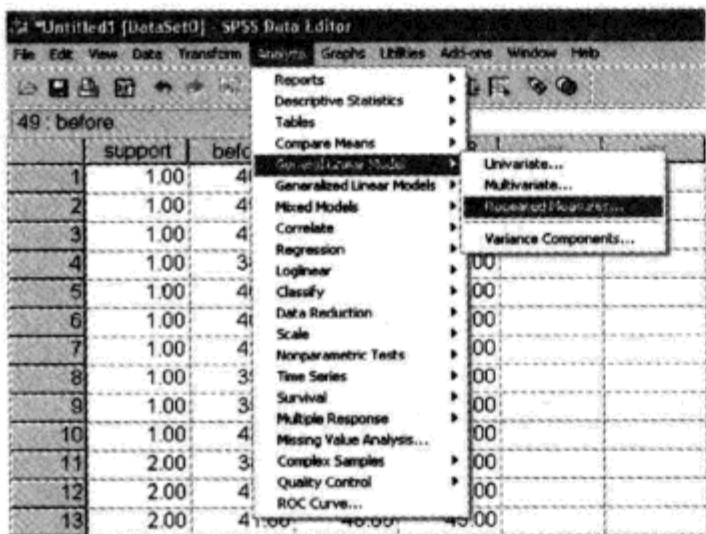


	support	before	week4	week8
1	1.00	40.00	39.00	35.00
2	1.00	45.00	44.00	42.00
3	1.00	42.00	44.00	44.00
4	1.00	38.00	36.00	30.00
5	1.00	46.00	44.00	38.00
6	1.00	40.00	37.00	25.00
7	1.00	42.00	38.00	35.00
8	1.00	39.00	37.00	29.00
9	1.00	35.00	33.00	31.00
10	1.00	43.00	44.00	38.00
11	2.00	38.00	44.00	42.00
12	2.00	47.00	45.00	42.00
13	2.00	41.00	48.00	45.00
14	2.00	39.00	41.00	43.00
15	2.00	44.00	44.00	41.00
16	2.00	42.00	39.00	35.00
17	2.00	42.00	48.00	46.00
18	2.00	40.00	46.00	42.00
19	2.00	38.00	45.00	43.00
20	2.00	41.00	40.00	35.00

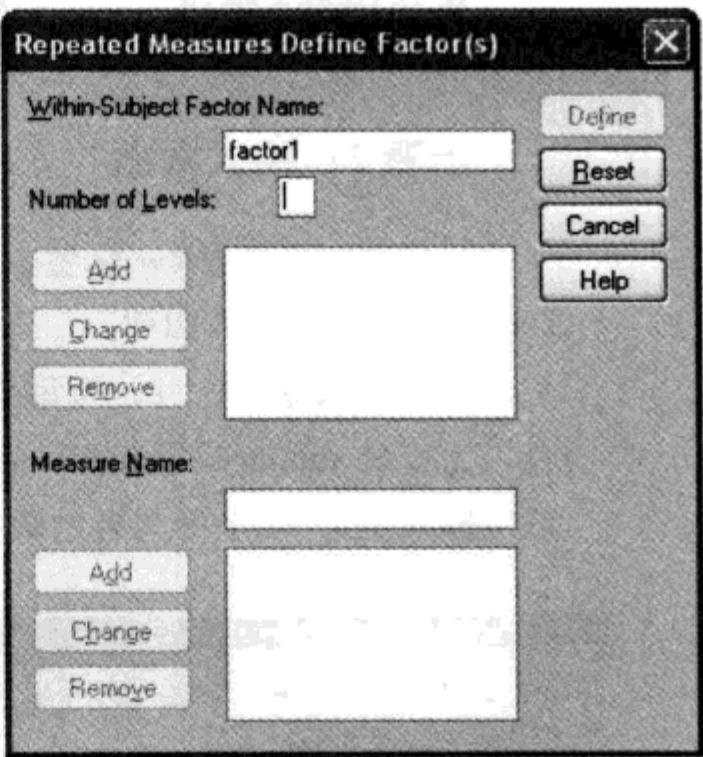
图表 11—3 组间组内方差分析的完整数据文件

步骤 3：分析数据

- 1. 从菜单栏中选择 **Analyze>General Linear Model>Repeated Measures...**（见图表 11—4）。
打开 Repeated Measures Define Factor (s) 对话框（见图表 11—5）。这个对话框为每个组内因素提供一个名字并且输入每个因素的水平数目。³
- 2. 在 Within Subject Factor Name 框中双击“factor1”（factor1 是 SPSS 中为组内因素建立的缺省值名字），输入名字 **time**。

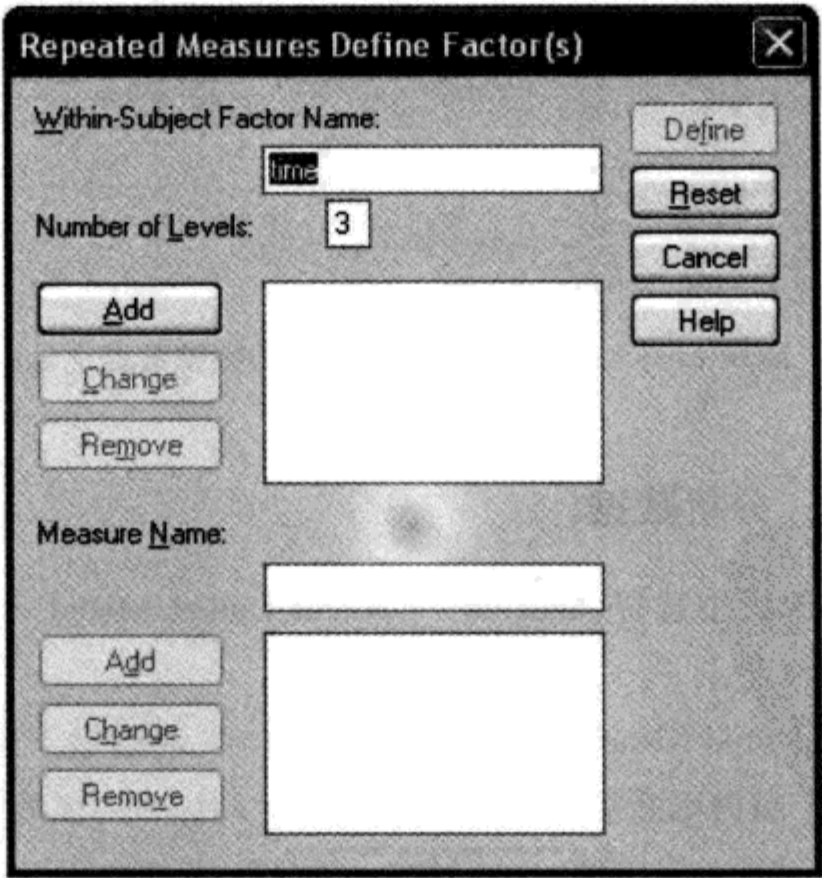


图表 11—4 组间组内方差分析的菜单命令



图表 11—5 重复度量定义因素对话框

3. 在 Number of Levels 的右边框中输入 3。这是对应于组内因素 **time** 的三个水平 (before, week4, week8)，详见图表 11—6。



图表 11—6 重复度量定义因素对话框 (续)

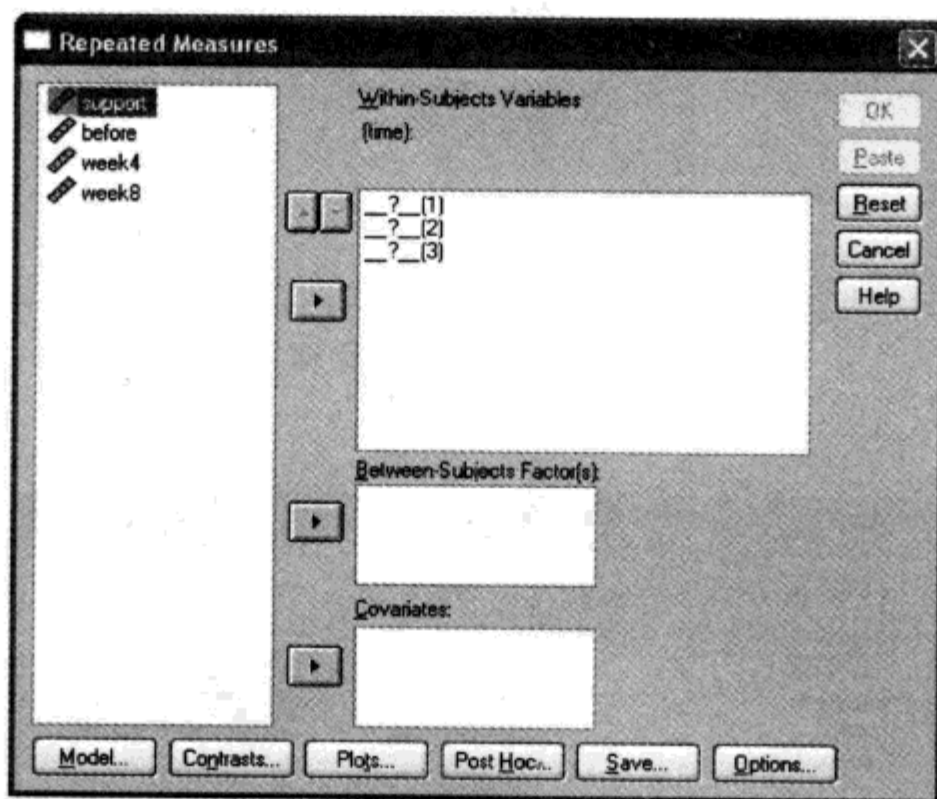
4. 点击 “Add”。

5. 点击 “Define”。

伴随着 Repeated Measures 对话框的打开, **support** 和三个时间点 **before**, **week4**, **week8** 出现在对话框的左边 (见图表 11—7)。

6. 选择组间变量 **support**, 点击右上方向右箭头按钮 (►) 移到 Between-Subjects Factor (s) 框中。

7. 选择时间点 **before**, **week4** 和 **week8**。点击向右箭头按钮 (►) 使它们移到 Within-Subjects Variables 框中 (见图表 11—8)。

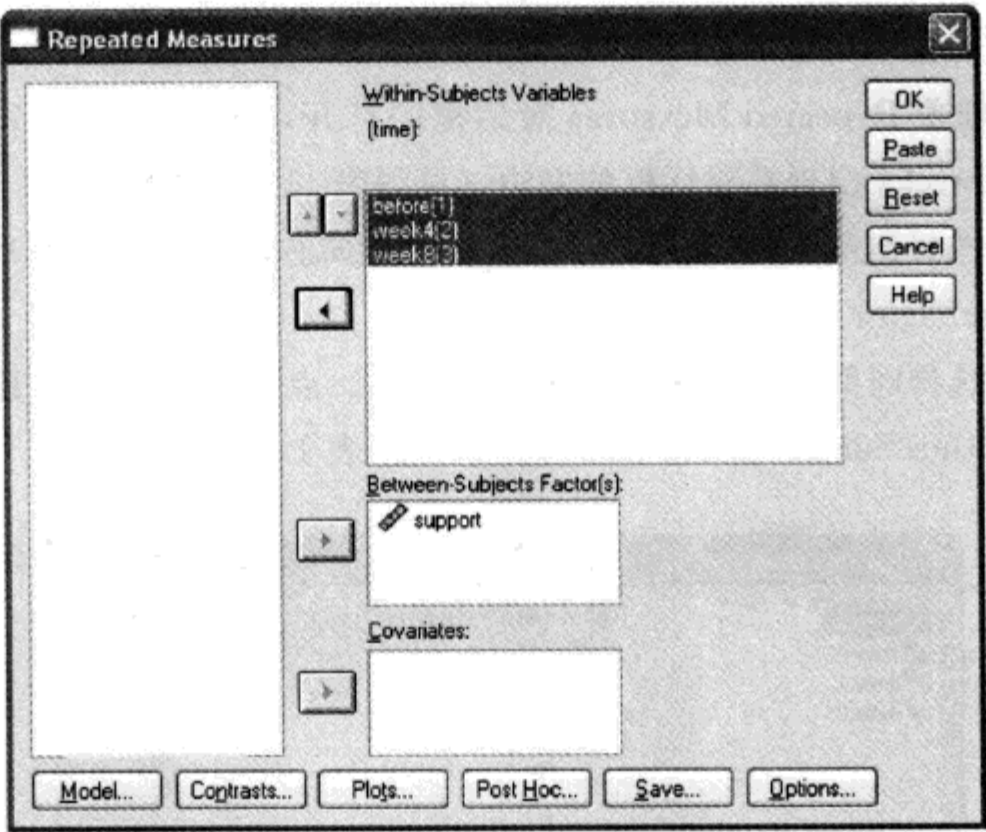


图表 11—7 重复度量对话框

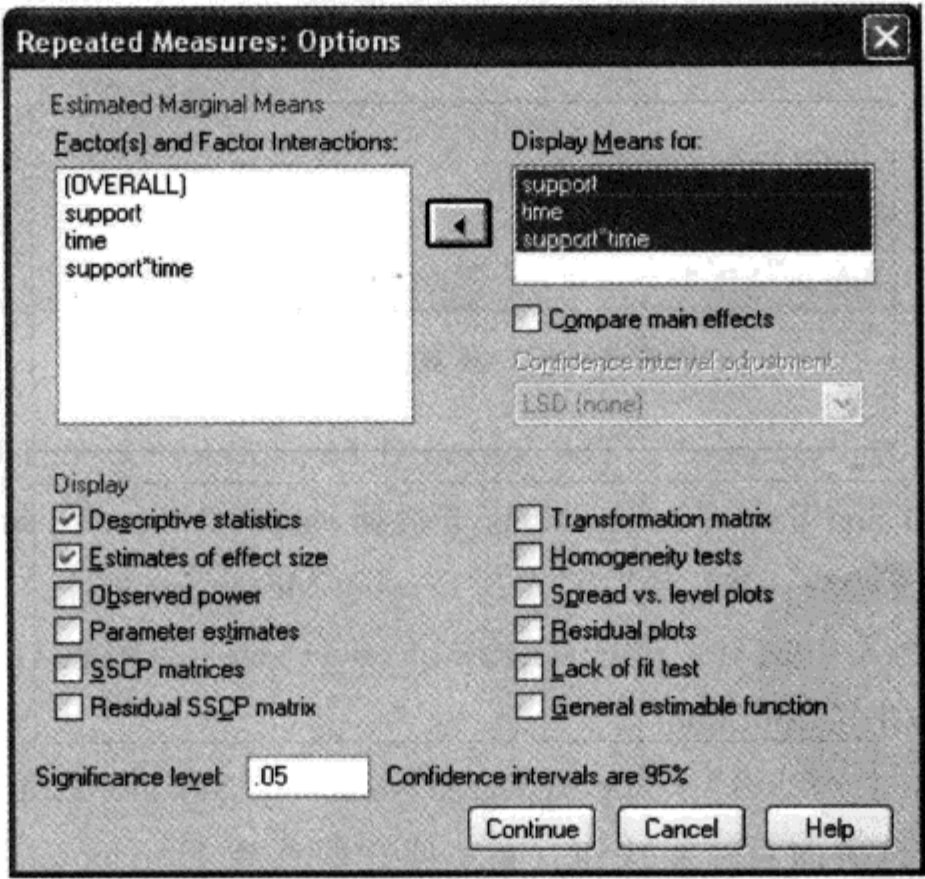
8. 点击 “Options”。The Repeated Measures: Options 对话框打开。在 Factor (s) and Factor Interactions 下选择 **support**, **time** 和 **support** × **time** 并点击向右箭头按钮 (►) 使它们移到 Display Means for 框中。在 Display 下, 选择 Descriptive statistics 和 Estimates of effect size (见图表 11—9)。

9. 点击 “Continue”。

10. 点击 “Plots”。The Repeated Measures: Profile Plots 对话框打开。选择 **time** 并点击向右箭头按钮 (►) 使它们移到 Horizontal Axis 框中。选择 **support** 并点击向右箭头按钮 (►) 使它们移到 Separate Lines 框。详见图表 11—10。



图表 11—8 重复度量对话框（续）



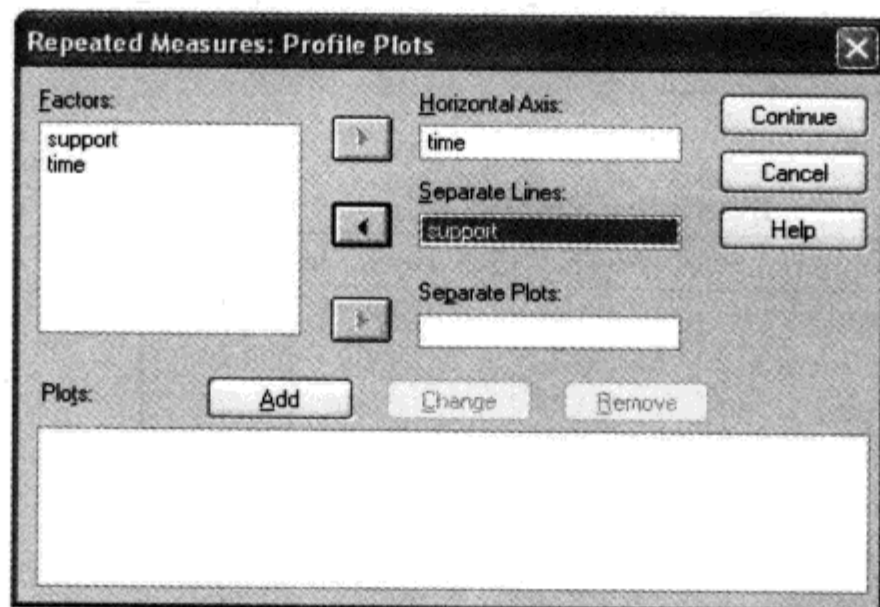
图表 11—9 重复度量：选项对话框

11. 点击“Add”。交互效应 **time**×**support** 就显示在 Plots 对话框中（详见图表 11—11）。

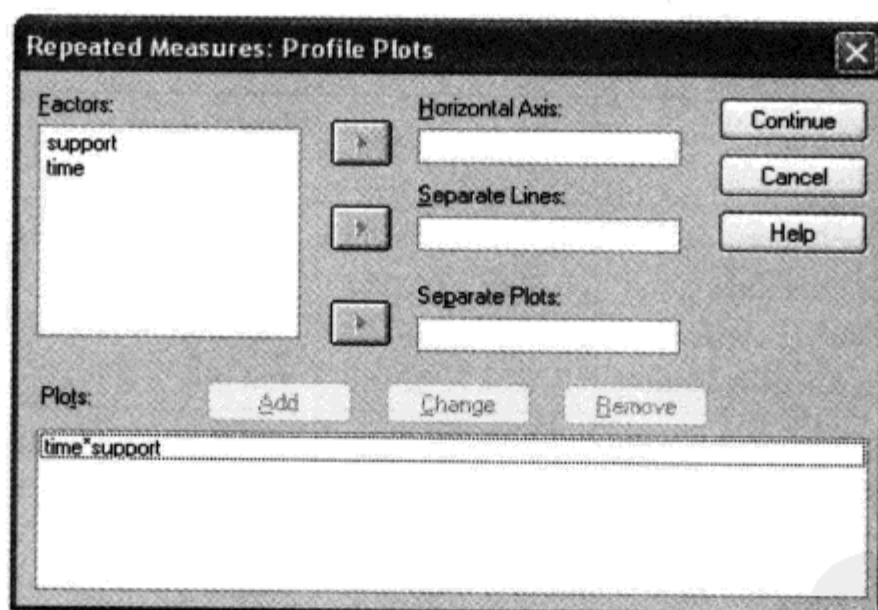
12. 点击“Continue”。

13. 点击“OK”。

在 SPSS 中运行组间组内方差分析程序，结果显示在 Viewer 窗口中。



图表 11—10 重复度量：轮廓图对话框



图表 11—11 重复度量：轮廓图对话框（续）

步骤 4：结果描述

组间组内方差分析的输出结果显示在图表 11—12 中。

General Linear Model

Within-Subjects Factors

Measure: MEASURE_1

time	Dependent Variable
1	before
2	week4
3	week8

Between-Subjects Factors

	Value Label	N
support 1.00	mentor	10
2.00	no mentor	10

Descriptive Statistics

	Support	Mean	Std. Deviation	N
before	mentor	41.0000	3.29983	10
	no mentor	41.2000	2.78089	10
	Total	41.1000	2.97180	20
week4	mentor	39.6000	4.08792	10
	no mentor	44.0000	3.12694	10
	Total	41.8000	4.20025	20
week8	mentor	34.7000	6.00093	10
	no mentor	41.4000	3.68782	10
	Total	38.0500	5.94249	20

Multivariate Tests^b

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
time	Pillai's Trace	0.704	20.183 ^a	2.000	17.000	0.000	0.704
	Wilks' Lambda	0.296	20.183 ^a	2.000	17.000	0.000	0.704
	Hotelling's Trace	2.375	20.183 ^a	2.000	17.000	0.000	0.704
	Roy's Largest Root	2.375	20.183 ^a	2.000	17.000	0.000	0.704
time* support	Pillai's Trace	0.365	4.878 ^a	2.000	17.000	0.021	0.365
	Wilks' Lambda	0.635	4.878 ^a	2.000	17.000	0.021	0.365
	Hotelling's Trace	0.574	4.878 ^a	2.000	17.000	0.021	0.365
	Roy's Largest Root	0.574	4.878 ^a	2.000	17.000	0.021	0.365

^aExact statistic

^bDesign: Intercept+support
Within Subjects Design: time

图表 11—12 组间组内方差分析的输出结果

Mauchly's Tests of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^a		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
time	0.545	10.310	2	0.006	0.687	0.767	0.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

^aMay be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

^bDesign: Intercept+support
Within Subjects Design: time

Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	E	Sig.	Partial Eta Squared
time	Sphericity Assumed	159.033	2	79.517	12.078	0.000	0.402
	Greenhouse-Geisser	159.033	1.375	115.676	12.078	0.001	0.402
	Huynh-Feldt	159.033	1.534	103.699	12.078	0.000	0.402
	Lower-bound	159.033	1.000	159.033	12.078	0.003	0.402
time* support	Sphericity Assumed	108.633	2	54.317	8.251	0.001	0.314
	Greenhouse-Geisser	108.633	1.375	79.016	8.251	0.004	0.314
	Huynh-Feldt	108.633	1.534	70.836	8.251	0.003	0.314
	Lower-bound	108.633	1.000	108.633	8.251	0.010	0.314
Error (time)	Sphericity Assumed	237.000	36	6.583			
	Greenhouse-Geisser	237.000	24.727	9.577			
	Huynh-Feldt	237.000	27.605	8.585			
	Lower-bound	237.000	18.000	13.167			

time 检验和 time×support 检验的 p 值。可以使用 Greenhouse-Geisser 方法或 Sphericity Assumed 的值（在本例中，使用任何一种方法检验，time 和 time×support 都是显著的，因为它们的 p 值均小于 0.05）。

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
time	Linear	93.025	1	93.025	8.463	0.009	0.320
	Quadratic	66.008	1	66.008	30.349	0.000	0.628
time* support	Linear	105.625	1	105.625	9.610	0.006	0.348
	Quadratic	3.008	1	3.008	1.383	0.255	0.071
Error (time)	Linear	197.850	18	10.992			
	Quadratic	39.150	18	2.175			

图表 11—12 组间组内方差分析的输出结果（续一）

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	97526.017	1	97526.017	2852.101	0.000	0.994
support	212.817	1	212.817	6.224	0.023	0.257
Error	615.500	18	34.194			

Support (指导) 是显著的, 因为 $p < 0.05$ 。

Estimated Alarginal Means

1. support

Measure: MEASURE_1

support	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
mentor	38.433	1.068	36.190	40.676
no mentor	42.200	1.068	39.957	44.443

2. time

Measure: MEASURE_1

time	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	41.100	0.682	39.667	42.533
2	41.800	0.814	40.090	43.510
3	38.050	1.114	35.710	40.390

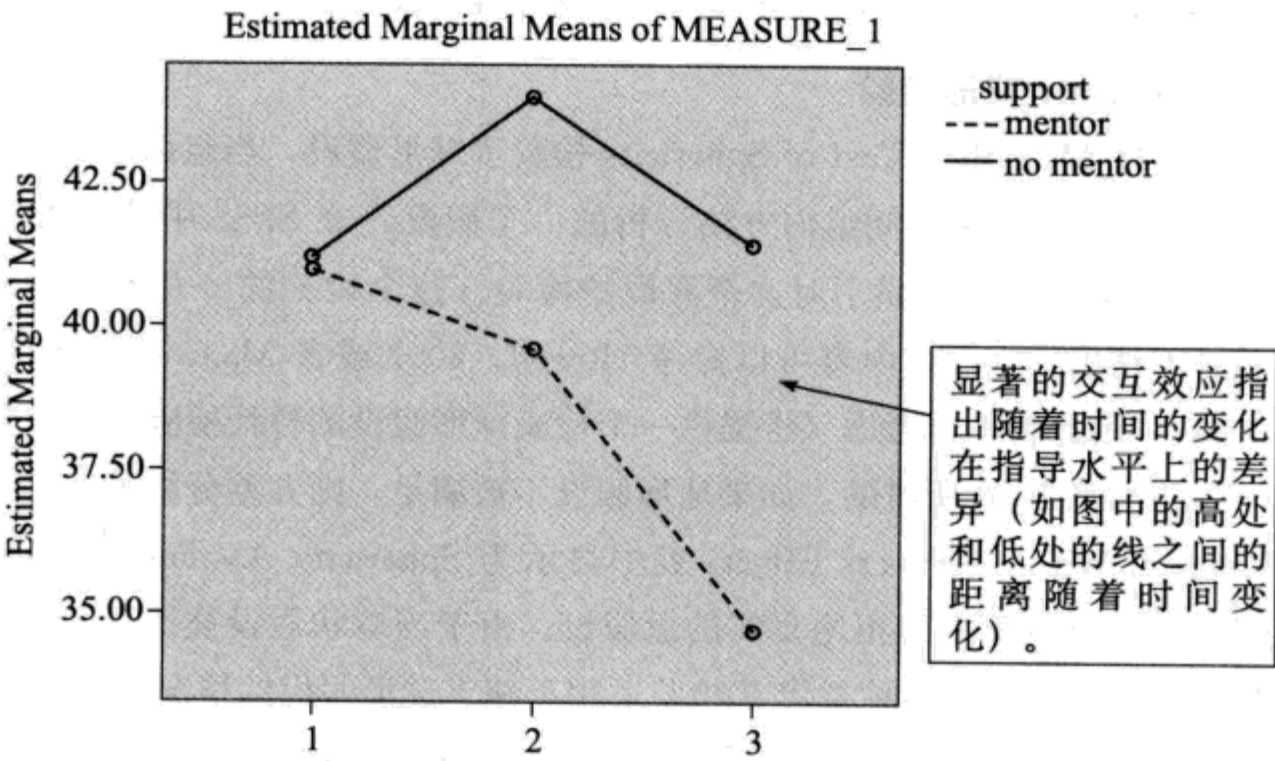
3. support* time

Measure: MEASURE_1

support	time	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
mentor	1	41.000	0.965	38.973	43.027
	2	39.600	1.151	37.182	42.018
	3	34.700	1.575	31.391	38.009
no mentor	1	41.200	0.965	39.173	43.227
	2	44.000	1.151	41.582	46.418
	3	41.000	1.575	38.091	44.709

图表 11—12 组间组内方差分析的输出结果 (续二)

Profile Plots



图表 11—12 组间组内方差分析的输出结果（续三）

组内因素 (Within-Subjects Factors)

第一张表格 Within-Subjects Factors 列出了度量参与者的三个时刻，包括计划开始前 (before)，进行 4 周后 (week4) 和进行 8 周后 (week8)。

组间因素 (Between-Subjects Factors)

表格 Between-Subjects Factors 显示了组间因素 support，support 的变量值标签和每组的样本量。

描述统计量

表格 Descriptive Statistics 显示了研究中每种情况的平均值、标准差和样本量 (time 的水平显示在表中的 Total 栏下)。我们将以后聚焦于输出结果中的表格 Estimated Marginal Means (估计边际均值) 解释均值差，而且我们将在写结果中使用表中的标准差。

多元检验

对于 time 和交互项 time×support 的原假设可以使用单变量检验或者多变量检验。表格 Multivariate Tests 对于 time 和 time×support 提供了四种不同的检验方法。⁴ 因为多元检验的结果可以自动给出，而且多元检验的程序已经超

出本章研究内容，所以我们这里不讨论。感兴趣的读者可以参考 Maxwell and Delaney (2004) 或者 Stevens (2002) 以了解详细信息。

Mauchly 球形检验

表格 Mauchly's Test of Sphericity 检验了球形假设，当组内因素有三个或更多水平时，这是组间组内方差分析的一个假设。⁵ 在 SPSS 中的表 Mauchly's Test of Sphericity 提供了对这个假设的检验（这个检验的 p -值为 0.006），检验是不精确的（详细内容可以参考 Howell, 2007 或者 Maxwell and Delaney, 2004），因此将不再考虑（将提供一个检验球形假设的替代检验方法）。

正如第 10 章所讲的，如果球形假设不被满足，则方差分析的 F 检验（在 Tests of Within Subjects Effects 表中显示为 Sphericity Assumed）将不准确，得到的结果将更多地拒绝真实的原假设。由于当球状假设被违背时 F 检验不精确，已经提出了几种替代 F 检验适用于缺乏球形假设的情况。在表格 Tests of Within-Subjects Effects 中有三个调整方法：Greenhouse-Geisser, Huynh-Feldt 和 Lower-bound。鉴于第 10 章中讨论的原因，我们将评价输出结果中的 Greenhouse-Geisser F 。当球形假设被违背时，Greenhouse-Geisser 调整方法提供了一个更精确的结果，由于手算比较烦琐，只有依靠计算机帮助，因此，那些想用 SPSS 确认手算结果的人将使用球形假设值。

组内效应检验

表格 Tests of Within-Subjects Effects 提供了我们所研究问题的答案，即三个时刻的压力得分是否不同以及 **time** 和 **support** 之间是否存在交互效应。正如第 8 章一维方差分析所说的那样， F 检验是两个方差的比值，每个方差在输出中表示为均方 (MS)：

$$F = \frac{MS \text{ Effect}}{MS \text{ Error}}$$

在组间组内方差分析中，包含组内因素（如 **time** 和 **time** × **support**）的检验有个误差不同于组间因素（如 **support**）的检验。在 Tests of Within-Subjects Effects 表中，**time** 和 **time** × **support** 有相同的误差项 MS Error (time)。对于 **time**，其 F 值为：

$$F = \frac{MS \text{ Time}}{MS \text{ Error (time)}}$$

把图表 11—12 中在 Greenhouse-Geisser 行的值代入公式得到 F 值为：

$$F = \frac{115.676}{9.577} = 12.078$$

这与图表 11—12 中的 Tests of Within-Subjects Effects 表中的 F 值一致（球形假设的 F 值也为 12.078 (79.517/6.583)。虽然这四个检验有相同的 F 值，但它们的 df 和 p -值不同）。

对于球形假设值，对 **time** 的检验为时间和误差各产生一个自由度 (df)，分别是 2 和 36（提醒一下，如果你手算 F 值，可以将结果与球形假设值进行比较）。为获得 Greenhouse-Geisser 方法的自由度值，一个调节方法被应用于球形假设的自由度，以此来补偿在数据上球形的缺失。如果现在看看 Mauchley's Test of Sphericity 表，你将会看到叫做 epsilon 的统计值为 0.687。这个值再乘以自由度 (2, 36)，将产生 1.375 和 24.747 的自由度，这些在 Greenhouse-Geisser 表中有记录。

在“Sig.”列中对时间 (**time**) 的 Greenhouse-Geisser 的 p -值是 0.001。由于 p -值小于 0.05，拒绝三个时刻均值相等的原假设，而且可以推断出至少有一个时刻与其他时刻不相等。

转移到 **time** × **support** 检验，Greenhouse-Geisser F 值是 8.251 (79.016/9.577)，其 p -值为 0.004。由于 p -值小于 0.05，拒绝原假设，也就是说存在显著的 **time** × **support** 的交互效应（本章后面将提供书面结果）。

组内对照检验

表格 Tests of Within-Subjects Contrasts 能够对组内主效应和交互效应进行进一步检验。这张表将不用在组间组内方差分析中考虑，因为我们将马上讨论采用一个替代的方法来进行进一步检验。

组间效应检验

表格 Tests of Between-Subjects Effects 显示了组间因素 **support** 检验的结果。对 **support** 检验的 F 值为 6.224 (MS Support/MS Error = 212.817/34.194)，相应的 p -值为 0.023。⁶ 因为 p -值小于 0.05，拒绝有导师和没有导师的老师的平均压力水平相等的原假设（本章后面将提供关于 **support** 的书面结果）。

估计边际均值

表格 Estimated Marginal Means 给出了每个因素的所有水平和交互效应的均值。第一张表 **support** 列出了有导师和没有导师两个组的均值。由于 **support**

是显著的，我们将检查均值来确定哪一个组有较低的压力得分。指导的边际均值表明有导师的那些人的平均压力水平（均值为 38.43）要低于没有导师的那些人（42.20）。

下一张表格 **time** 给出了 **time** 的所有水平的均值。由于 **time** 有三个水平，不进一步检验的话，我们不能准确区别出哪一个时刻明显不同（回想有三个或更多组的对立假设的非特异性本质）。（在本章练习前的“在 SPSS 中执行组间组内方差分析的步骤摘要”里提供了对组内因素的进一步检验的指令。）

最后一张表格 **support* time** 给出了研究中的六种情况的均值（3 **time** × 2 **support**）。为了解释交互效应的显著意义，我们将把注意力放在轮廓图上，那里用图形显示了这六种情况的均值。

轮廓图

轮廓图显示了研究中的 6 种情况的均值。就像我们前面在 SPSS 中指定的那样，**time** 在横轴，**support** 的各个水平表示为不同的直线（在轮廓图中，1, 2, 3 分别对应着时间的 **before**, **week4** 和 **week8**）。我们考察均值，发现在计划开始前，有导师组和无导师组的压力得分几乎相等（41 对 41.2），但在计划进行 4 周后和进行 8 周后，有导师指导的老师的压力水平明显低于没有导师指导的老师（在计划进行 4 周后是 39.6 对 44.0，在进行 8 周后是 34.7 对 41.4）。显著的交互效应表明这些直线不平行，证实有导师和没有导师的组之间的差异随着时间变化。

交互效应检验：简单效应分析

尽管交互效应指出这些直线明显不平行，但它并不能指出哪一个点明显不同于其他的点。⁷ 为了确定哪一个点明显不同于其他点，将进行简单效应检验。简单效应检验用于固定其他因素的某一个水平而对因素的差异进行检验。对目前的例子，进行了几个不同的简单效应检验，包括在不同的时间对指导（有导师对无导师）进行检验（对轮廓图中不同的线进行检验）以及在指导的不同水平上对时间（**before**, **week4** 和 **week8**）进行检验（对轮廓图中每条线里面进行检验）。

为了说明组间组内方差分析的简单效应检验，我们将检查在计划开始前、进行 4 周后和进行 8 周后，有导师和没有导师指导的老师的压力水平是否一样（查看交互效应图中的 1、2、3）。这就需要进行三次不同的独立样本 *t* 检

验，每次对时间的每一个水平检验。由于 t 检验是进一步的检验，它们将在 α 水平为 0.016 ($0.05/3$) 下进行以保证三个独立样本 t 检验的试验误差比例不超过 0.05 (详见第 10 章中关于进一步检验的调整 α 水平)。

在每个时刻检验有导师和没有导师的压力水平意味着下面的三个原假设：

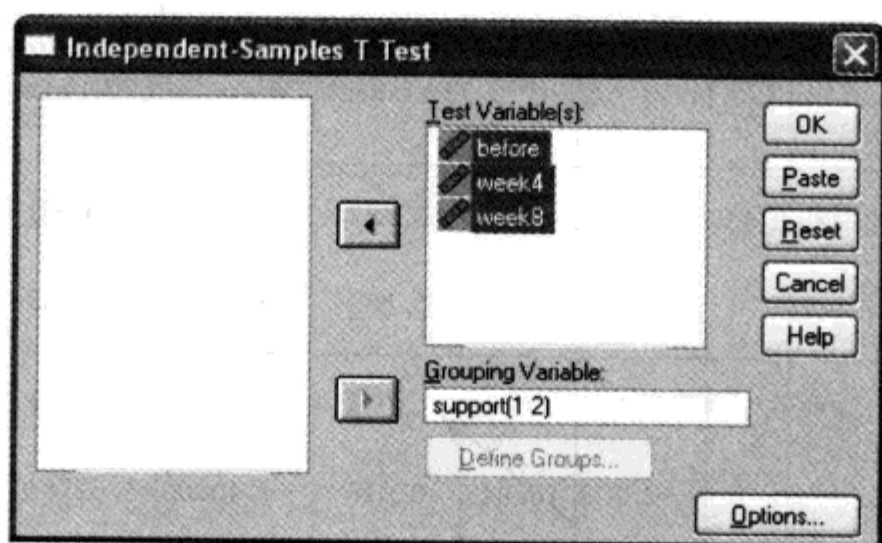
$$H_0: \mu_{\text{mentored_before}} = \mu_{\text{not mentored_before}}$$

$$H_0: \mu_{\text{mentored_4weeks}} = \mu_{\text{not mentored_4weeks}}$$

$$H_0: \mu_{\text{mentored_8weeks}} = \mu_{\text{not mentored_8weeks}}$$

下面提供了独立样本 t 检验的命令 (详见第 7 章独立样本 t 检验)。

1. 从菜单中选择 **Analyze > Compare Means > Independent-Samples T Test**。
2. 把 **support** 移到 Grouping Variable 框中。
3. 点击 “Define Groups”。将 a1 赋给 Group1, a2 赋给 Group2。
4. 选择 **before**, **week4** 和 **week8**, 然后点击右上方向右箭头按钮 (►), 把它们移到 Test Variable(s) 框中 (见图表 11—13)。
5. 点击 “OK”。



图表 11—13 独立样本 t 检验对话框

简单效应分析结果：在时间的每个水平检验指导

在图表 11—14 中显示了独立样本 t 检验的结果。独立样本 t 检验表显示在进行 4 周后 ($p=0.015$) 和进行 8 周后 ($p=0.008$) 的指导都很显著，但是在计划开始前 ($p=0.885$) 不显著 (回想每个检验的 α 水平是 0.016)。在 Group Statistics 表中的平均值指出在进行 4 周后和进行 8 周后有导师指导的新

老师比那些没有得到导师指导的新老师有较低的压力水平。这些结果将一并写到本章结尾处的整个书面结果里。

T-Test

Group Statistics

support		N	Mean	Std. Deviation	Std. Error Mean
before	mentor	10	41.0000	3.29983	1.04350
	no mentor	10	41.2000	2.78089	0.87939
week 4	mentor	10	39.6000	4.08792	1.29271
	no mentor	10	44.0000	3.12694	0.98883
week 8	mentor	10	34.7000	6.00093	1.89766
	no mentor	10	41.4000	3.68782	1.16619

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
before	Equal variances assumed	0.483	0.496	-147	18	0.885	-0.20000	1.36463	-3.06699	2.66699
	Equal variances not assumed			-0.147	17.498	0.885	-0.20000	1.36463	-3.07290	2.67290
week 4	Equal variances assumed	1.984	0.176	-2.703	18	0.015	-4.40000	1.62754	-7.81934	-0.98066
	Equal variances not assumed			-2.703	16.846	0.015	-4.40000	1.62754	-7.83621	-0.96379
week 8	Equal variances assumed	2.693	0.118	-3.008	18	0.008	-6.70000	2.22736	-11.37950	-2.02050
	Equal variances not assumed			-3.008	14.949	0.009	-6.70000	2.22736	-11.44890	-1.95110

在调整 α 水平 0.016 下，week 4 和 week 8 是显著的，其 p -值 < 0.016 (before 不是显著的，由于其 p -值 > 0.016)。

图表 11—14 独立样本 t 检验输出结果 (在显著交互效应之后的简单效应检验)

还可以进行简单效应的其他检验，比如检验不同指导小组在时间 (time)

上的差异。例如，我们可以检验那些有导师指导的新老师在不同时间下的压力水平，这意味着是一个一维组内方差分析。如果这个方差分析是显著的，那么正如第10章所说，所有的时间的配对比较都可以用独立样本 t 检验进行检验（如1对2，1对3，2对3）。由于篇幅所限我们将不在这里一一介绍这些检验，在本章的练习3中有个对有导师的老师关于时间的检验（答案在附录C中给出）。

效应量

组间组内方差分析的效应量通常使用偏 η^2 度量。为了计算偏 η^2 ，使用方差分析中合适的表（要么是 Tests of Within-Subjects Effects 表，要么是 Tests of Between-Subjects Effects 表）的平方和。偏 η^2 的计算公式如下：

$$\text{偏 } \eta^2 = \frac{SS_{\text{Effect}}}{SS_{\text{Effect}} + SS_{\text{Error}}}$$

SS_{Effect} 和 SS_{Error} 的值在图表 11—12（对于 **time** 和 **time** × **support**， SS_{Error} 是以 $SS_{\text{Error}(\text{time})}$ 形式出现）的方差分析表中以 Type III Sum of Squares 形式存在。对于 **time**，在公式中代入适当的值得到

$$\text{偏 } \eta^2 = \frac{159.033}{159.033 + 237.000} = 0.402$$

这个结果与组内因素效应表结果相吻合。

偏 η^2 值越大，因变量的方差被效应解释得越多。常规度量方法还没有提供对小、中和大效应量的偏的效应量度量。

APA 格式的结果表达

在写组间组内方差分析的结果时，除了报告组间组内方差分析均值和标准差（如果需要，均值和标准差可以单独成一张表），还要报告假设检验的结论，自由度 (df)， F 值， p 值和效应量。如果交互效应显著，将需要作一个直方图或者均值轮廓图。下面展示一个简单的 APA 格式的书面例子。

对教学压力进行了 2×3 的组间组内方差分析，以指导（有导师指导，无导师指导）作为组间因素，以时间（before, week 4, week 8）为组内因素。这些结果显示指导有显著主效应， $F(1, 18) = 6.22, p < 0.05$ ，偏 $\eta^2 =$

0.26, 而且时间也有显著主效应, Greenhouse-Geisser adjusted $F(1.38, 24.75) = 12.08, p < 0.05$, 偏 $\eta^2 = 0.40$ 。对于指导, 有导师的新老师报告的压力得分显著低于无导师指导的老师。也有一个显著的指导 \times 时间交互效应, Greenhouse-Geisser adjusted $F(1.38, 24.75) = 8.25, p < 0.05$, 偏 $\eta^2 = 0.31$ 。在时间的每个水平上对指导进行了简单效应分析, 每个检验的 α 水平为 0.016。简单效应检验的结果指出在进行 4 周后和进行 8 周后, 有导师指导的新老师报告的压力得分显著低于无导师指导的老师, 在进行 4 周后, $t(18) = -2.70, p < 0.016$, 在进行 8 周后, $t(18) = -3.01, p < 0.016$ 。在计划开始前, 有导师指导和没有导师指导的新老师的压力得分几乎没有差异, $t(18) = -0.15, p > 0.016$ 。在图表 11—15 中显示有导师指导和没有导师指导的新老师在计划开始前, 进行 4 周后和进行 8 周后的压力得分的均值和标准差。

图表 11—15 计划开始前, 进行 4 周后和进行 8 周后有无导师计划时的均值和标准差

时间	有导师		没有导师	
	M	SD	M	SD
计划开始前	41.00	3.30	41.20	2.78
进行 4 周后	39.60	4.09	44.00	3.13
进行 8 周后	34.70	6.00	41.40	3.69

组间组内方差分析的假定

1. 组内的观测是独立的。

违反该假定会严重影响组间组内方差分析结果的准确性。如果我们有理由相信已经违反了独立性假定, 那么不能使用组间组内方差分析。详见第 10 章。

2. 正态。

这个假定意味着: (1) 对于组内因素的每个水平, 因变量总体服从正态分布; (2) 对于组间因素的每个水平, 组间的平均得分 (组内因素的平均水平) 应该服从正态分布。详见第 8 章的违反正态分布的后果。

3. 方差齐性。

这个假定意味着组间因素的每个水平上的总体方差应相等 (在组内因素的平均水平上计算每个参与者的方差)。详见第 8 章的违背这个假定的后果。

4. 球形假定。

球形假定要求组内因素的所有配对水平上的差值得分的总体方差是相等的。违反这个假定将影响方差分析检验的精确性，使用一个替代的方法（如 Greenhouse-Geisser, Huynh-Feldt）以适应球形的存在。详见第 10 章。

5. 方差—协方差矩阵的齐性。

这个假定意味着组间因素的不同水平对应的方差和协方差相等（协方差是度量在两个变量之间共有的变异性）。在当前的例子中，对组内因素的三个水平，这个假定意味着三个相应的方差（ var_{before} , var_{week4} , var_{week8} ）和三个协方差（ $cov_{(before, week4)}$, $cov_{(before, week8)}$, $cov_{(week4, week8)}$ ）对于两个组是相等的。当样本量相等或者近似相等时，方差—协方差相等的假定适度被违背是可以容忍的。如果样本量之间的不相等是从中等到较大，就不能使用组内组间方差分析。

在 SPSS 中执行组间组内方差分析的步骤摘要

I. 数据输入和分析

1. 对组间因素生成一个变量，对组内因素的每个水平生成一个独立的变量。
2. 给每个组间变量建立变量值标签。在 Value Labels 对话框中，输入恰当的数值类型，点击“OK”。
3. 输入数据。
4. 选择 **Analyze > General Linear Model > Repeated Measures...**
5. 在 Repeated Measures Define Factor(s) 对话框中，在 Within-Subject Factor Name 框中输入组内因素的名字，然后在 Number of Levels 框中输入水平数目。点击“Define”。
6. 在 Repeated Measures 对话框中，把组间因素移到 Between-Subjects Factor(s) 框中。把组内因素的使用水平移到 Within-Subjects Variables 框中。
7. 点击“Options”。把所有因素和交互效应变量移到 Display Means for 框中。在 Display 下选择 Descriptive statistics 和 Estimates of effect size，点击“Continue”。
8. 点击“OK”。

II. 结果解释

在组内因素效应的检验表中，要么在 Greenhouse-Geisser 检验下，要么在球形假定值栏下检查 p -值来解释组内因素检验的结果（如果与手算结果进行比较，可以使用球形假定值）。通过检查在组间因素的检验的 p -值来解释组间因素的结果。

- 如果组间因素是显著的，且对每个变量有两个水平，检查均值边际分布。对于有三个或者更多水平的，运行 Tukey 检验（通过点击 “Post Hoc” 和显著 Tukey）。如果组内因素是显著的，且有两个因素，检查均值轮廓图。对于有三个或者更多水平的，如果方差分析是显著的，在运行独立样本 t 检验之后继续运行一维方差分析（忽视组间因素）。如果交互效应是显著的，使用 α 水平为 0.05 来描述交互效应的本质或者执行简单效应检验（如果有需要）。
- 如果组间组内方差分析没有一个检验是显著的，直接书写结论，主效应和交互效应不显著。

练习

1. 一位临床心理学家想比较认知行为和心理分析随着时间变化的关系。20 个人同意参加这次研究，其中有 10 个人接受认知行为治疗，另外 10 个人接受心理分析治疗。所有参与者在治疗开始前，进行 8 周后和进行 16 周后分别记录得分情况。分数范围为 10~50 分。分数越高表示状况越好。数据见图表 11—16。

图表 11—16 练习 1 的数据

Therapy	Before	Week8	Week16	Therapy	Before	Week8	Week16
1	19	18	22	2	23	23	17
1	18	18	21	2	19	19	20
1	21	22	24	2	16	17	15
1	22	23	28	2	18	19	19
1	24	24	26	2	23	25	24
1	18	21	27	2	24	25	23
1	19	23	25	2	19	20	18
1	19	20	23	2	22	24	22
1	17	18	16	2	20	20	19
1	23	24	28	2	22	22	23

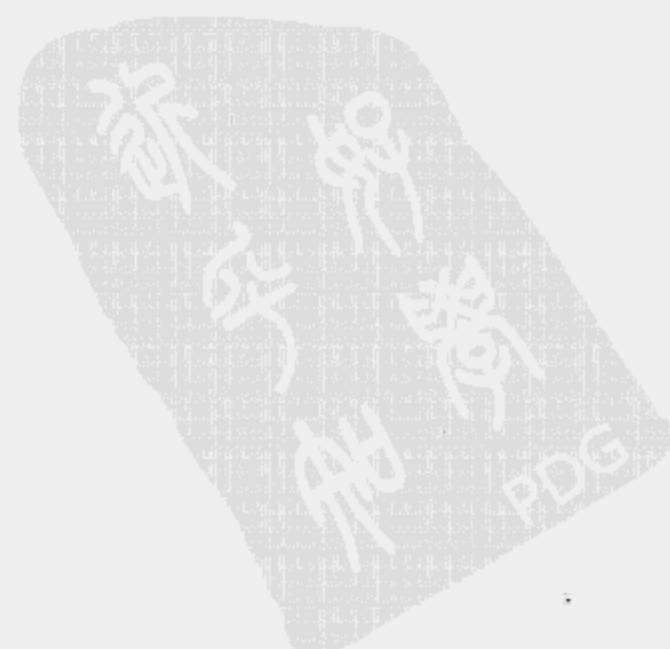
说明：对于 therapy，1=认知行为治疗，2=心理分析治疗。

在 SPSS 中输入数据，并且进行适当的分析，回答下面的问题。分别取变量名为 **therapy**, **before**, **week8** 和 **week16**。

- a. 陈述感兴趣的检验的原假设和对立假设。
 - b. 描述感兴趣的检验的一个研究问题。
 - c. 检验主效应和交互效应（对每个检验取 $\alpha=0.05$ ）。若有的话，哪些是显著的？哪些不显著？
 - d. 报告每个检验的效应量。哪一个检验效应量最大？
 - e. 将研究结果写成适当的 APA 格式。
2. 执行练习 1 的简单效应检验。检验三个时间点上的两个治疗方案是否有差异。对于每个检验取合适的 α 水平值，使其检验的 α 水平总和不超过 0.05。

3. 在本章讨论的问题中，在各个时间点对有无导师指导的老师进行了简单效应检验分析。另一个可能的分析是看这三个时间点上所对应的压力得分是否不同。在 SPSS 中运行这个检验并回答下面的问题。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter11 中的文件 **teaching stress.sav** 之中。在 SPSS 中输入数据，并且进行适当的分析，回答下面的问题（对于有导师的老师，**support** 记为 1）。（提示：为了执行这个检验，要么选择案例，要么拆分文件。这些程序详见附录 A。）

- a. 陈述感兴趣的检验的原假设和对立假设。
- b. 描述感兴趣的检验的一个研究问题。
- c. 有导师的老师在三个时间点上的压力得分是否显著不同？ $\alpha=0.05$ 。
- d. 如果时间是显著的，执行合适的进一步检验，调整每个检验的 α 水平使进一步检验的 α 水平总和不超过 0.05。若有的话，哪些是显著的？哪些不显著？
- e. 将研究结果写成适当的 APA 格式。



皮尔逊 r 相关系数

皮尔逊 r 相关系数度量两个变量的线性相关程度。相关程度用字母 r 表示； r 可能是正的（一个变量较高的得分与另一个变量较高的得分相关联）、负的（一个变量较高的得分与另一个变量较低的得分相关联）或 0（两变量间无相关性）。相关系数的范围可以从 -1.0 （一个完美的负相关）到 1.0 （一个完美的正相关）。下面给出一个皮尔逊 r 相关系数的例子。

例子

在一项研究中，一个学生想检查生活意义和心理健康是否相关。同意参与这项研究的 30 个学生测量了生活意义和心理健康。生活意义的得分范围是 10~70 分（更高的得分表示更强的生活意义），心理健康的得分范围是 5~35 分（更高的得分表示更健康的心理状态）。

皮尔逊 r 相关系数的目标和数据要求

皮尔逊 r 相关系数		
目标	数据要求	例子
度量两变量的线性相关程度	连续变量 ¹	变量 1：生活意义 变量 2：心理健康

原假设和对立假设

原假设表明两变量总体上没有相关性。相关系数的总体符号用希腊字母 ρ 表示：

$H_0: \rho=0$

对立假设表明两变量总体上有相关性：

$H_1: \rho \neq 0$

原假设的评价

皮尔逊 r 相关系数对生活意义和心理健康间不相关的原假设进行了检验。如果检验产生的结果在原假设正确时看起来不可能（结果发生的可能性小于 5%），那么拒绝原假设。如果检验产生的结果在原假设正确时看起来正确（结果发生的可能性大于 5%），那么不拒绝原假设。

研究问题

在研究中基本的兴趣问题也可以用研究问题的方式表示，例如，
“生活意义和心理健康相关吗？”

数据

在图表 12—1 中列出了 30 个参与者的数据。

图表 12—1			相关系数数据的例子		
Participant	Meaning in Life	Well-being	Participant	Meaning in Life	Well-being
1	35	19	9	25	12
2	65	27	10	58	21
3	14	19	11	30	18
4	35	35	12	37	25
5	65	34	13	51	19
6	33	34	14	50	25
7	54	35	15	30	29
8	20	28	16	70	31

续前表

Participant	Meaning in Lite	Well-being	Participant	Meaning in Life	Well-being
17	25	12	24	50	20
18	55	20	25	39	24
19	61	31	26	68	34
20	53	25	27	56	28
21	60	32	28	19	12
22	35	12	29	56	35
23	35	28	30	60	35

说明：变量 participant 包含在数据中，但不用输入 SPSS。

在 SPSS 中输入数据及分析

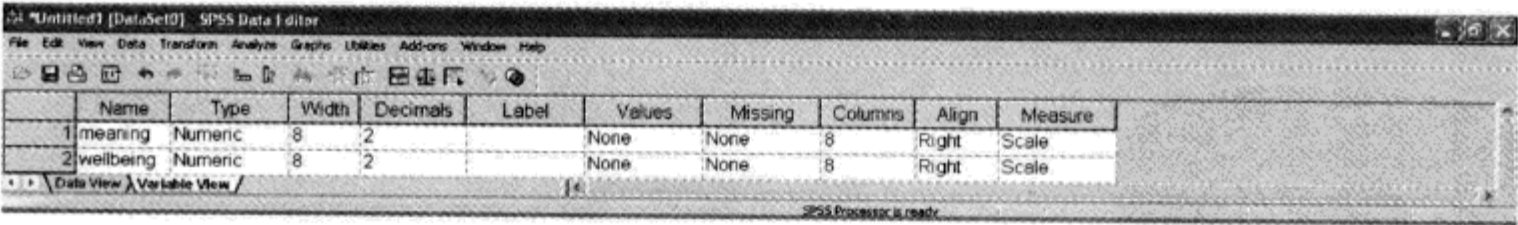
下面的步骤 1 和步骤 2 描述了怎样在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 12 中的文件 meaning.sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中将生成两个变量，一个是生活意义，另一个是心理健康。变量分别被命名为 **meaning** 和 **wellbeing**。

- 3. 在 Variable View 窗口前两行分别输入变量名称 **meaning** 和 **wellbeing** (见图表 12—2)。



图表 12—2 输入变量 meaning 和 wellbeing 的 Variable View 窗口

步骤 2：输入数据

- 1. 点击 Data View 标签。变量 **meaning** 和 **wellbeing** 出现在 Data View 窗口

的前两列。

2. 参照图表 12—1，为每个参与者输入两个变量的数据。对第一个参与者，为变量 **meaning** 和 **wellbeing** 分别输入 35 和 19。依次输入全部 30 个参与者的数据。图表 12—3 给出了完整的数据文件。

步骤 3：分析数据

1. 从菜单栏中选择 **Analyze>Correlate>Bivariate...**（见图表 12—4）。

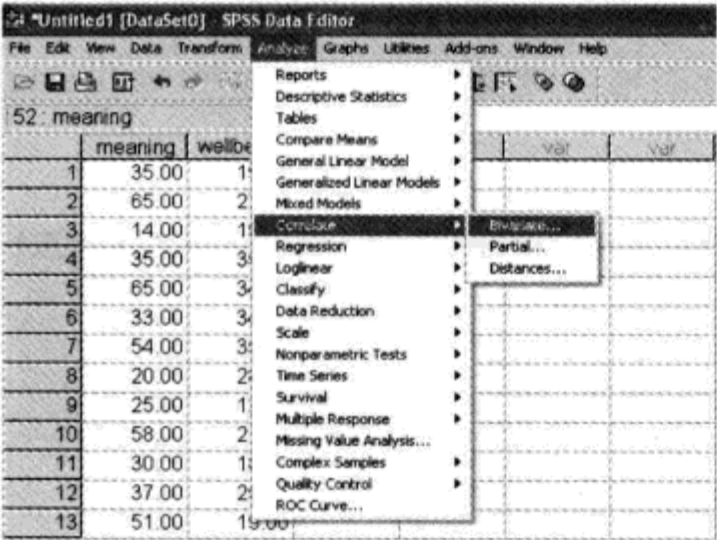
打开 Bivariate Correlations（二元相关性）对话框，变量 **meaning** 和 **wellbeing** 出现在对话框的左边（见图表 12—5）。

2. 选择变量 **meaning** 和 **wellbeing**，点击向右箭头按钮（▶），把变量移到 Variables 框中（见图表 12—6）。

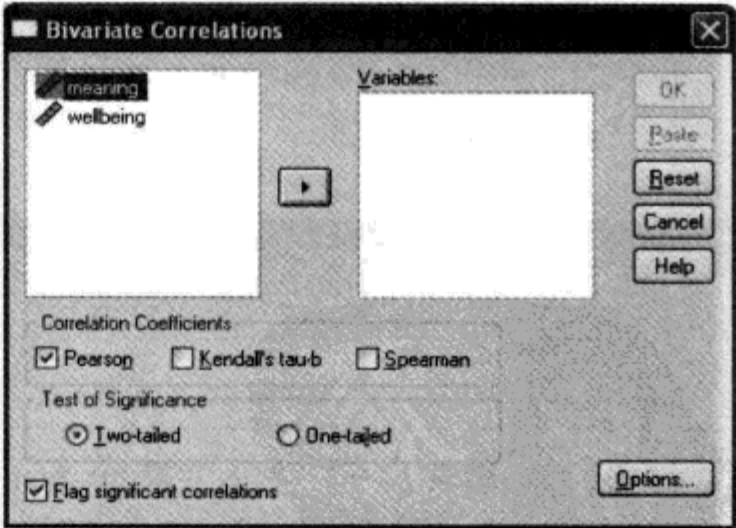
3. 点击“OK”。

52: meaning			
	meaning	wellbeing	var
1	35.00	19.00	
2	65.00	27.00	
3	14.00	19.00	
4	35.00	35.00	
5	65.00	34.00	
6	33.00	34.00	
7	54.00	35.00	
8	20.00	28.00	
9	25.00	12.00	
10	58.00	21.00	
11	30.00	18.00	
12	37.00	25.00	
13	51.00	19.00	
14	50.00	25.00	
15	30.00	29.00	
16	70.00	31.00	
17	25.00	12.00	
18	55.00	20.00	
19	61.00	31.00	
20	53.00	25.00	
21	60.00	32.00	
22	35.00	12.00	
23	35.00	28.00	
24	50.00	20.00	
25	39.00	24.00	
26	68.00	34.00	
27	56.00	28.00	
28	19.00	12.00	
29	56.00	35.00	
30	60.00	35.00	

图表 12—3 皮尔逊 r 相关系数例子的数据文件

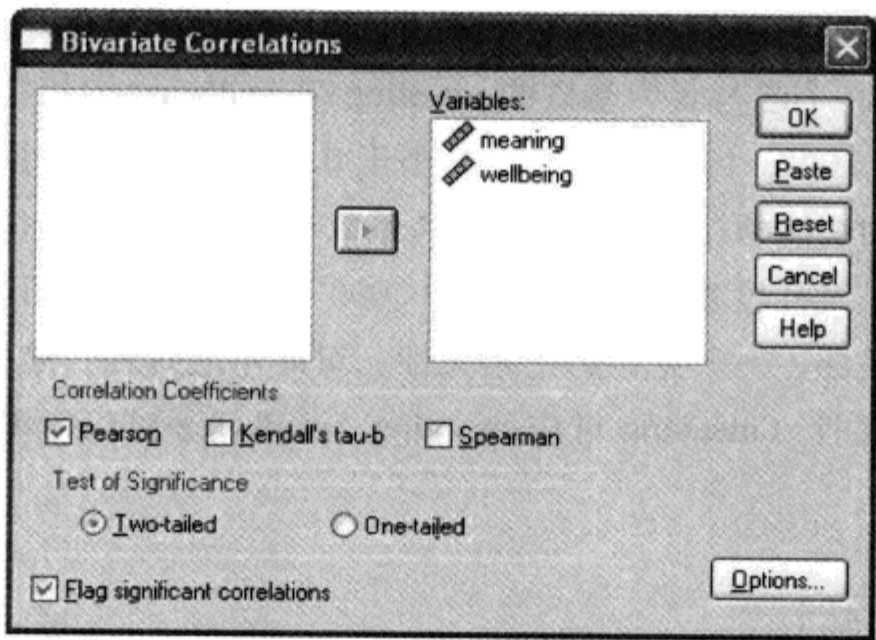


图表 12—4 皮尔逊 r 相关系数的菜单命令



图表 12—5 Bivariate Correlations 对话框

在 SPSS 中运行相关性程序，结果显示在 Viewer 窗口中。



图表 12—6 Bivariate Correlations 对话框 (续)

步骤 4：解释结果

二元相关性的输出结果显示在图表 12—7 中。

Correlations

Correlations		meaning	wellbeing
meaning	Pearson Correlation	1	0.549**
	Sig. (2-tailed)		0.002
	N	30	30
wellbeing	Pearson Correlation	0.549**	1
	Sig. (2-tailed)	0.002	
	N	30	30

同样的结果在相关性表格中出现了两次。

生活意义和心理
健康间的相关性。

相关性显著，因
为 p 值小于 0.05。

** Correlation is significant at the 0.01 level (2-tailed).

图表 12—7 皮尔逊 r 相关系数的输出结果

相关性 (Correlations)

SPSS 生成了一个输出表，标记为 Correlations，其中包括我们研究问题的答案，即变量 **meaning** 和 **wellbeing** 之间是否相关。

注意在表格 Correlations 中变量 **meaning** 和 **wellbeing** 出现了两次，一次在行，一次在列（这表明 SPSS 生成的相关性表格中出现了冗余）。相关系数值

和原假设检验的 p -值位于变量 **meaning** 和 **wellbeing** 相交处。表格 Correlations 显示了以 30 个参与者为基础的 **meaning** 和 **wellbeing** 的相关性是 0.549，相应的 p -值是 0.002。因为 p -值 0.002 小于 0.05，原假设被拒绝，在 **meaning** 和 **wellbeing** 的总体中存在正相关（相关系数右边的两个星号暗示了在 0.01 水平上相关性是统计显著的，因为 p -值 0.002 小于 0.01）。剩下的两个单元格显示了 1 的相关性，一个完美的正相关性。对此不应奇怪，这些值对应了变量与自身的相关性（**meaning** 对应 **meaning**，**wellbeing** 对应 **wellbeing**），它始终等于 1。

效应量

不像到目前为止我们考虑过的其他统计方法，相关系数本身就是一个常用的效应量的度量。科恩（Cohen, 1988）给出了 r 的估计值为 ± 0.1 ， ± 0.3 ， ± 0.5 ，分别对应小、中和大的效应量。基于科恩准则，相关性 0.549 对应于实际中一个大的规模效应，表明在生活意义和心理健康间有一个相当强的正相关。

APA 格式的结果表达

SPSS 在表格 Correlations 中显示了样本量 (N)，但是没有提供自由度，尽管自由度是基于这个结果的。自由度公式是：

$$df = (N - 2)$$

N 对应研究中总的样本量（参与者数量）。研究中的 30 个人，自由度等于 28。

在写结果时，要报告假设检验的结论， r 值，自由度 (df) 和 p -值。我们接下来将展示一个 APA 格式的简单书面例子。

在生活意义和心理健康间存在显著的正相关关系， $r(28) = 0.55$ ， $p < 0.05$ 。

当计算一个相关系数时，经常有兴趣去建立一个散点图去检查两变量间的线性（直线）关系和检查异常值。建立一个散点图的命令在第 3 章（检验线性（直线）关系和异常值的更多信息查阅本书推荐的统计入门书）。

皮尔逊 r 相关系数的假定

1. 参与者间的观测是独立的。

该假定意味着每个参与者的得分应该相对于其他人的得分是独立的（如果两个参与者一起进行心理健康测量是违反独立假定的一个例子）。如果违反了独立性假定，就不应该再用相关系数了。

2. 二元正态。

该假定意味着每个变量的总体都应该服从正态分布，对于每个变量的值，其他变量的得分应该服从正态分布。对于中等到较大的样本量，绝大部分的非正态分布趋向于对相关系数的检验的精确性没有多少影响。

在 SPSS 中执行相依样本 t 检验的步骤摘要

I. 数据输入和分析

1. 在 SPSS 中生成两个变量。
2. 输入数据。
3. 选择 **Analyze > Correlations > Bivariate...**
4. 将两个变量移到 Variables 框中。
5. 点击“OK”。

II. 结果解释

检查 Correlations 表格中的 p 值：

- 如果 $p \leq 0.05$ ，拒绝原假设。写出表示两变量间有显著关系（积极或消极）的结果。
- 如果 $p > 0.05$ ，不拒绝原假设。写出表示两变量间没有显著关系的结果。

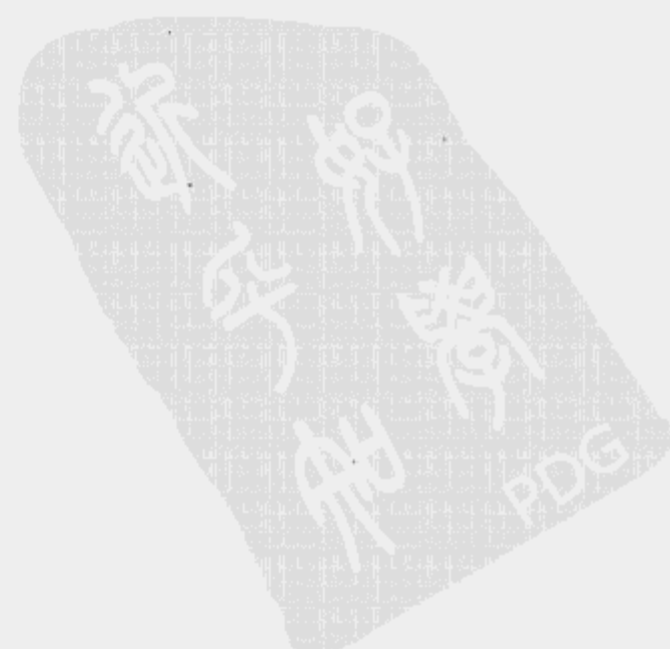
练习

1. 一个学生对检查学习时间（以分钟计）和考试分数（分数范围为 0~100 分）间的关系很感兴趣。25 个学生的学习时间（**examprep**）和最后考试

- b. 陈述数据的一个研究问题。
- c. 在 SPSS 中计算两变量的相关系数。相关性值是多少？
- d. 相关性显著吗？在 $\alpha=0.05$ 时检验。
- e. 该研究的效应量是多少？你能区分效应量是小、中还是大吗？
- f. 将研究结果写成适当的 APA 格式。

3. 一个研究者调查了一个孩子（3岁）听阅读的时间和后来在二级英语技能考试中的表现之间是否有关系。这里记录了30位父母为孩子朗读的时间（范围为1~10分）和英语技能的考试分数（范围为10~50分）（越高的分数分别表明越多的朗读时间和越好的考试成绩）。数据在网站上文件夹 Chapter 12 中的文件 Chapter12 _ Exercise 3. sav 之中（变量名为 **readingtime** 和 **exam-scores**）。在 SPSS 中打开文件，进行必要分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 在 SPSS 中计算两变量的相关系数。相关性值是多少？
- d. 相关性显著吗？在 $\alpha=0.05$ 时检验。
- e. 该研究的效应量是多少？你能区分效应量是小、中还是大吗？
- f. 将研究结果写成适当的 APA 格式。



简单线性回归

简单线性回归用于预测随另一个变量取值变动而变动的变量。¹ 在回归中，被预测的变量通常称作因变量，用于预测的变量通常称为自变量或预测变量。下面给出使用简单线性回归的例子。

例子

在一项研究中，一名学生想要在大学生中调查社会支持（一个人能向其他人寻求支持的程度）能否预测心理健康。同意参与这项研究的 25 名学生测量了社会支持和心理健康。社会支持的得分范围是 8~40 分，心理健康的范围是 10~70 分。得分越高表示社会支持和心理健康程度越高。

简单回归的目标和数据要求

简单线性回归		
目标	数据要求	例子
为了预测随另一个变量变动而变动的变量	因变量 • 连续 自变量 • 连续 ²	因变量 • 心理健康 自变量 • 社会支持

原假设和对立假设

原假设说明社会支持无法预测心理健康。为了评价原假设，建立回归方程（方程将在后面介绍），检验通常被称为贝塔权重（beta weight）的回归系数是否显著不等于 0。贝塔权重显著不等于 0 表示自变量是因变量的显著预测变量。回归方程中的总体符号用希腊字母 β 表示。

原假设规定贝塔权重等于 0：

$$H_0: \beta_{\text{SSI}} = 0 \text{ (社会支持的贝塔权重为零；社会支持不能预测心理健康)}$$

对立假设规定贝塔权重不等于 0：

$$H_1: \beta \neq 0 \text{ (社会支持的贝塔权重不为零；社会支持能预测心理健康)}$$

原假设评价

SPSS 中的线性回归程序可以对社会支持不能预测心理健康的原假设进行检验。如果检验产生的结果在原假设正确时看起来不可能（结果发生的可能性小于 5%），那么拒绝原假设。如果检验产生的结果在原假设正确时看起来正确（结果发生的可能性大于 5%），那么不拒绝原假设。

研究问题

在研究中基本的兴趣问题也可以用研究问题的方式表示，例如，

“在大学生里社会支持能否预测心理健康？”

数据

在图表 13—1 中列出了 25 名参与者的社会支持和心理健康数据。

图表 13—1 多元回归例子的数据

Participant	Social Support	Well-being	Participant	Social Support	Well-being
1	20	32	14	35	66
2	38	65	15	32	25
3	35	60	16	34	52
4	30	56	17	35	70
5	12	25	18	28	51
6	31	25	19	17	32
7	18	65	20	24	42
8	28	56	21	31	25
9	14	23	22	16	61
10	24	42	23	32	52
11	34	60	24	19	26
12	28	51	25	25	41
13	32	58			

说明：变量 participant 包含在数据中，但不用输入 SPSS。

在 SPSS 中输入数据及分析

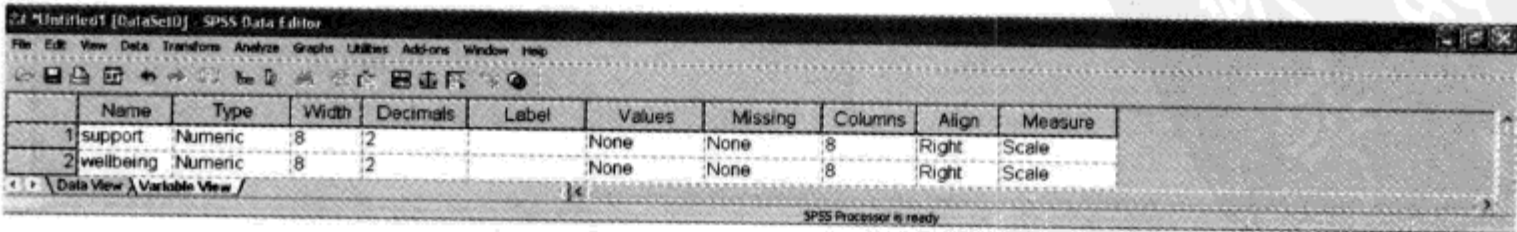
下面的步骤 1 和步骤 2 描述了怎样在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 13 中的文件 well being.sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中将生成两个变量，一个是社会支持，另一个是心理健康。变量分别命名为 **support**（支持）和 **wellbeing**（心理健康）。

- 3. 在 Variable View 窗口前两行输入变量名称 **support** 和 **wellbeing**（详见图表 13—2）。



图表 13—2 输入变量 support 和 wellbeing 的 Variable View 窗口

步骤 2：输入数据

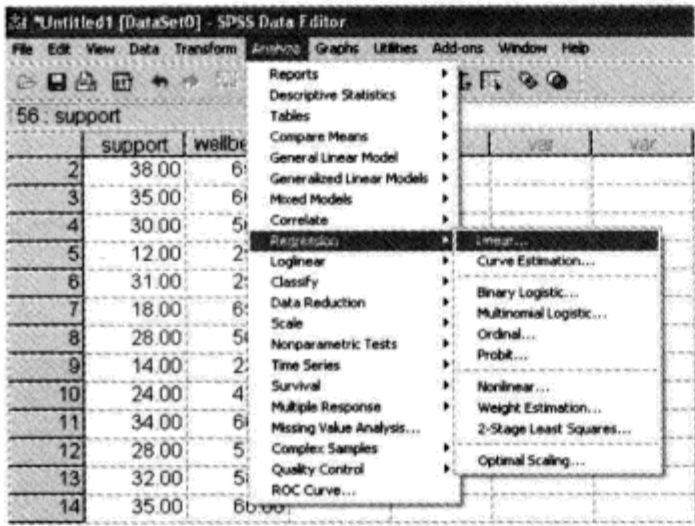
- 1. 点击 Data View 标签。变量 **support** 和 **wellbeing** 出现在 Data View 窗口前两列。
- 2. 参照图表 13—1，为每个参与者输入两个变量的数据。对第一个参与者，为变量 **support** 和 **wellbeing** 输入 20 和 32。依次输入全部 25 个参与者的数据。图表 13—3 给出了完整的数据文件。

步骤 3：分析数据

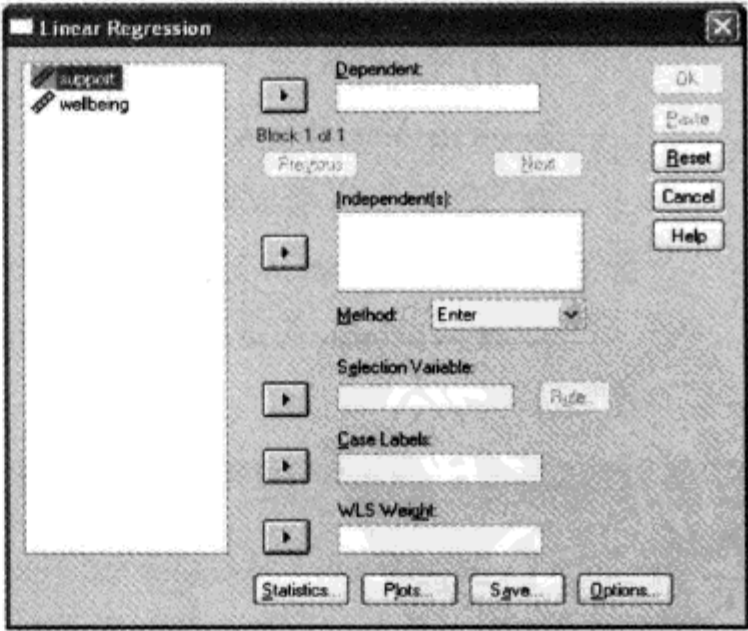
- 1. 从菜单栏中选择 **Analyze>Regression>Linear...**（见图表 13—4）。
- 打开 Linear Regression 对话框，变量 **support** 和 **wellbeing** 出现在对话框的左边（见图表 13—5）。
- 2. 选择变量 **wellbeing**，点击向右箭头按钮（▶），把变量移到 Dependent 框中。

54. wellbeing			
	support	wellbeing	var
1	20.00	32.00	
2	38.00	65.00	
3	35.00	60.00	
4	30.00	56.00	
5	12.00	25.00	
6	31.00	25.00	
7	18.00	65.00	
8	28.00	56.00	
9	14.00	23.00	
10	24.00	42.00	
11	34.00	60.00	
12	28.00	51.00	
13	32.00	58.00	
14	35.00	66.00	
15	32.00	25.00	
16	34.00	52.00	
17	35.00	70.00	
18	28.00	51.00	
19	17.00	32.00	
20	24.00	42.00	
21	31.00	25.00	
22	16.00	61.00	
23	32.00	52.00	
24	19.00	26.00	
25	25.00	41.00	

图表 13—3 线性回归例子的完整数据文件



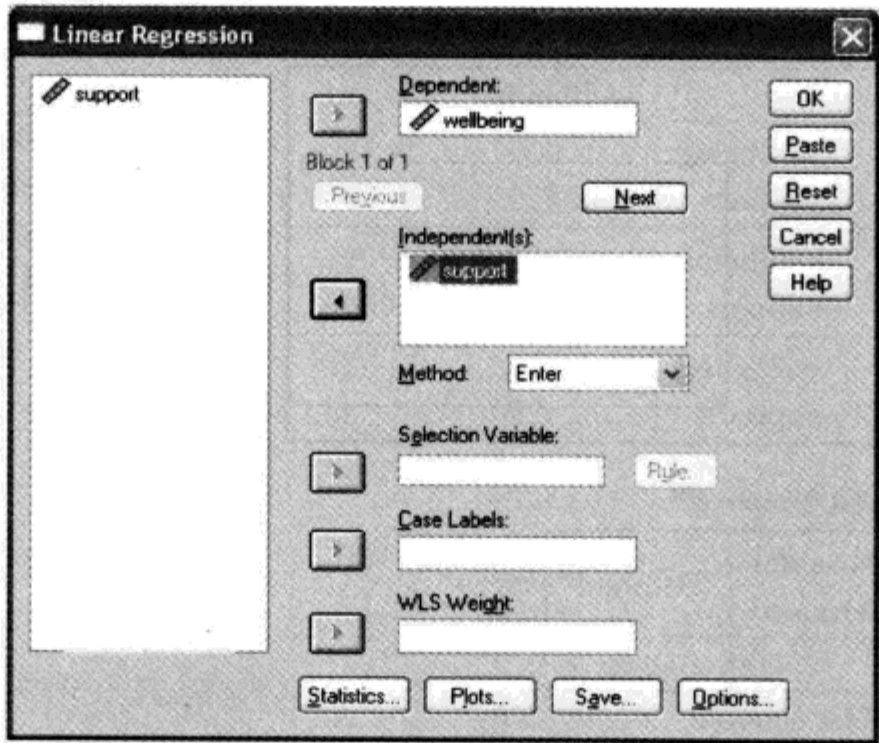
图表 13—4 线性回归方法的菜单命令



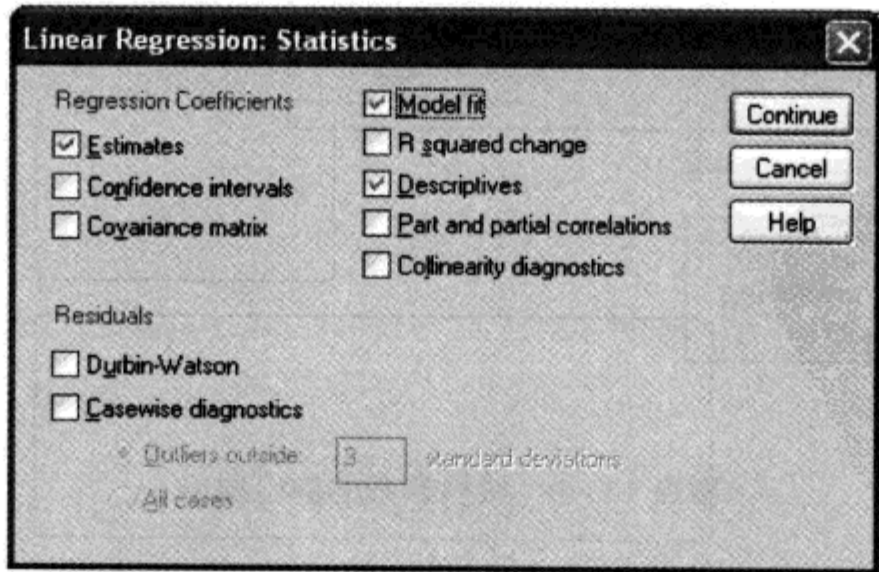
图表 13—5 线性回归对话框

- 3. 选择变量 **support**，点击第二个向右箭头按钮 (►)，把变量移到 Independent (s) 框中 (见图表 13—6)。
- 4. 点击 “Statistics”。选择 Descriptives (Estimates 和 Model fit 为已选中状态)。详见图表 13—7。
- 5. 点击 “Continue”。
- 6. 点击 “OK”。

在 SPSS 中运行线性回归程序，结果显示在 Viewer 窗口中。



图表 13—6 线性回归对话框 (续)



图表 13—7 线性回归：统计量对话框

步骤 4：解释结果

线性回归程序输出结果显示在图表 13—8 中。

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
wellbeing	46.4400	15.70318	25
support	26.8800	7.52950	25

Correlations

		wellbeing	support
Pearson Correlation	wellbeing	1.000	0.490
	support	0.490	1.000
Sig. (1-tailed)	wellbeing	0.000	0.006
	support	0.006	0.00
N	wellbeing	25	25
	support	25	25

wellbeing 和
social support
间的皮尔逊
相关系数

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	support ^a	0.000	Enter

^aAll requested variables entered.

^bDependent Variable: wellbeing

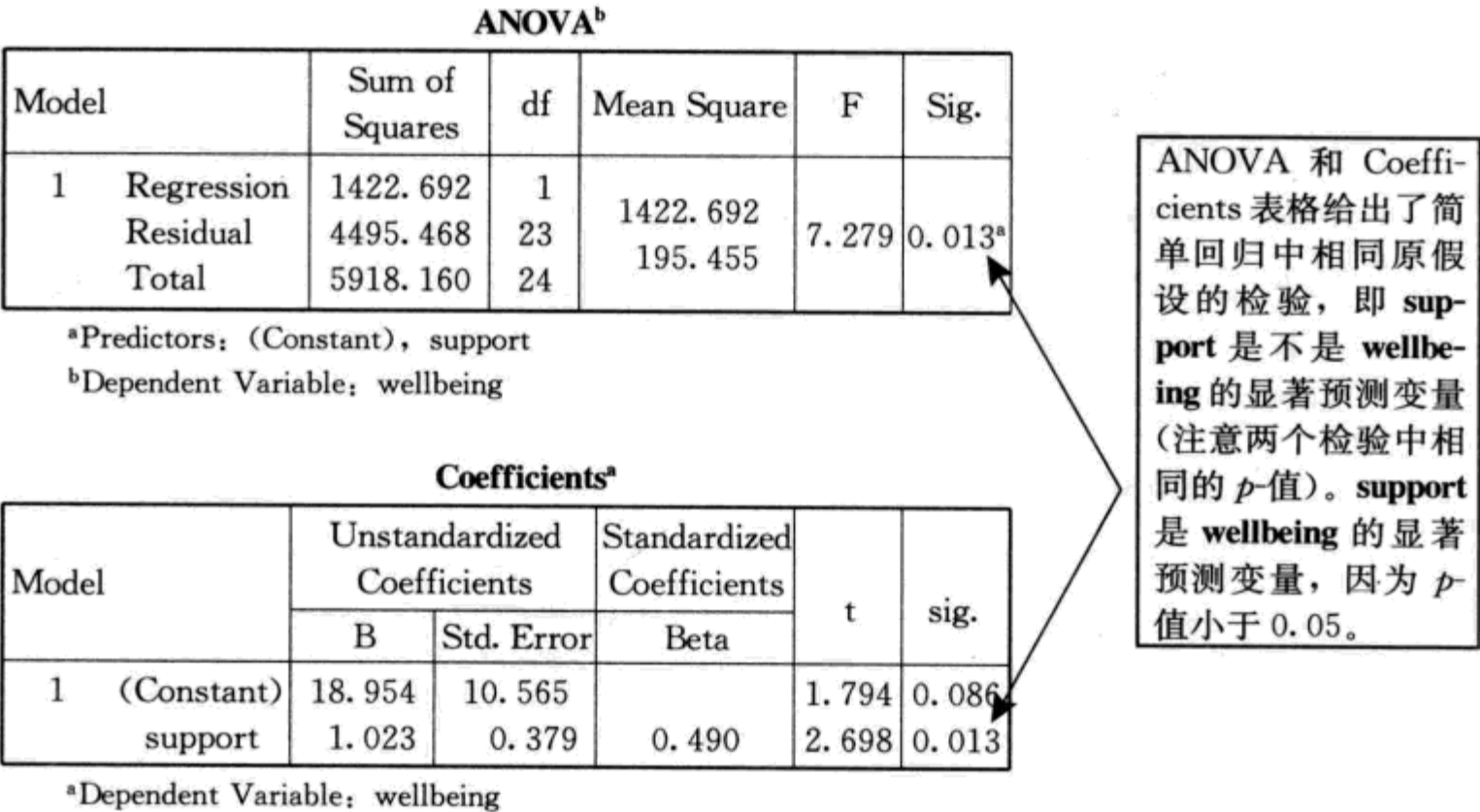
Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.490 ^a	0.240	0.207	13.98053

^aPredictors: (Constant), support

0.24 的 R^2 值表示社会支持解释了变量 wellbeing 分数中 24% 的变化。

图表 13—8 线性回归的输出结果



图表 13—8 线性回归的输出结果（续）

描述统计量

表格 Descriptive Statistics 列出了每个变量的均值、标准差和样本量。

相关性 (Correlations)

表格 Correlations 显示了社会支持和心理健康的相关性。两变量间的相关性为 0.490，单侧检验的 p 值为 0.0064（SPSS 保留小数点后三位为 0.006；双精度为 0.0064，有舍入误差，双侧检验的 p 值为 0.013）。不论使用单侧检验（ $0.006 < 0.05$ ）还是双侧检验（ $0.013 < 0.05$ ），两变量间的相关性都是显著的。在简单线性回归中，如果（双侧）相关性显著，那么回归结果也是显著的。

变量进入/移出 (Variables Entered/Removed)

表格 Variables Entered/Removed 显示了用于预测心理健康（wellbeing）得分的变量（support）。

模型概要 (Model Summary)

表格 Model Summary 给出了 R ， R^2 ，调整的 R^2 和估计标准误差。前三个值度量心理健康被社会支持预测的程度有多大，最后一个值度量心理健康不能被社会支持预测的程度有多大。下面对这些值进行讨论。

第一个值 R 是多元相关系数，它等于社会支持和心理健康间的皮尔逊相

关系数的绝对值 (R 取值范围为 $0 \sim 1$)。第二个值 R^2 是 R 的平方 ($0.490^2 = 0.24$)，当乘以 100% 时，可理解为因变量的总变异性中被自变量解释的百分比。在目前的例子中，社会支持解释了心理健康的总变异性的 24% ($0.24 \times 100\%$)。调整的 R^2 修正了 R^2 是为了对总体值做更好的估计 (基于样本计算的 R^2 容易高估总体值)。最后，估计标准误差表明自变量不能预测因变量值的程度。值 13.98 说明用社会支持预测心理健康，回归方程在预测心理健康时平均偏离大约 13.98 点。

方差分析

表格 ANOVA 检验社会支持是不是心理健康的显著预测变量。该检验使用方差分析，这已经在第 8 章讨论过。在简单回归中，ANOVA 表中 p 值小于或等于 0.05 表示自变量是因变量的显著预测变量。因为 p 值 0.013 小于 0.05 ，社会支持是心理健康的显著预测变量。

系数 (Coefficients)

最后一个表 Coefficients 提供了构造回归方程所必需的值以及检验社会支持是不是心理健康的显著预测变量的原假设 (由于 ANOVA 表已经指出社会支持是心理健康的显著预测变量，Coefficients 表内的检验是多余的)。³

首先将会讨论回归方程，其次讨论对社会支持的检验。

在线性回归中，方程以下面的形式建立：

$$\hat{Y} = a + bX$$

式中， \hat{Y} ——因变量的预测值，在我们的例子里， \hat{Y} 为心理健康预测值；

a —— Y 轴截距； $X=0$ 时 \hat{Y} 的值；

b ——回归直线的斜率；

X ——每一个参与者的自变量值，在我们的例子中， X 为社会支持得分。

回归方程中的 a (Y 轴截距) 和 b (斜率) 可以在 Coefficients 表中标为“B”的 Unstandardized Coefficients 栏中找到。在表中， Y 轴截距为 18.954 ，斜率为 1.023 。将这些数据代入回归方程中得到心理健康预测值的如下结果：

$$\hat{Y}_{\text{心理健康}} = 18.954 + 1.023 (\text{社会支持})$$

这是基于 25 名大学生得到的回归方程。对给定的社会支持值，可以得到一个心理健康预测值。例如，前两名参与者的社会支持值分别为 20 和 38。将这些数值代入上面的方程得到下面的心理健康预测值：

参与者 1 的心理健康预测值：

$$\hat{Y}_{\text{心理健康}} = 18.954 + 1.023 \quad (20)$$

$$\hat{Y}_{\text{心理健康}} = 39.41$$

参与者2的心理健康预测值:

$$\hat{Y}_{\text{心理健康}} = 18.954 + 1.023 \quad (38)$$

$$\hat{Y}_{\text{心理健康}} = 57.83$$

可以求出数据库中每个人的预测值。⁴ 预测值大多会有些误差(它们不会刚好等于真实值); R 值越高, 预测值越接近真实值, 两变量间完美的相关可以使得对每个个体有完美的预测。

社会支持显著性检验

Coefficients 表也提供了对社会支持的检验。在该表的最后两列中给出了社会支持检验的 t 和 p -值。社会支持检验得到 $t = 2.698$ 和相应的 p -值 = 0.013。因为 p -值 0.013 小于 0.05, 拒绝贝塔权重等于 0 的原假设, 并可推断社会支持是心理健康的显著预测变量。注意到 Coefficients 表中的社会支持的 p -值 (0.013) 等于 ANOVA 表中的 p -值 (0.013)。正如前面提到的, 这些 p -值相同的原因是在简单回归中它们检验相同的原假设, 即社会支持是不是心理健康的显著预测变量。

当 Unstandardized Coefficients 列中的数值用于寻找预测值时, Standardized Coefficients 值将会在结果中写出。标准化系数称为贝塔权重, 如果预测变量与因变量都是 z 得分形式, 它等于回归系数的值(就是说, 如果它们是标准化, 均值为 0 且标准差为 1)。⁵ 在标准化系数列中贝塔的值为 0.490, 这与 Correlations 表中给出的两变量间皮尔逊相关系数一致(在简单回归中贝塔总是等于皮尔逊 r 相关系数)。

在简单回归中通常对常数的检验不感兴趣, 它检验 Y 轴截距是否显著不等于 0。在本例中, 由于检验中 p -值 (0.086) 大于 0.05, 值 18.954 不显著不为零。

效应量

在简单回归中效应的度量是用 R^2 给出的。科恩 (Cohen, 1998) 定义简单回归中 R^2 的值 0.01、0.09 和 0.24 分别表示为小、中和大的效应量。在我

们的例子中, $R^2=0.24$ 表示心理健康得分的变异性的 24% 能被社会支持得分解释, 这表示为大效应。

APA 格式的结果表达

在写结果时, 要报告标准回归系数和预测变量 (社会支持) 检验的结果以及 R^2 值。接下来将展示一个简单的书面例子。

完成了以心理健康作为因变量和以社会支持作为预测变量的回归分析。社会支持是心理健康的显著预测变量。 $\beta=0.49$, $t(23)=2.70$, $p<0.05$, 解释了心理健康得分的变异性的 24% ($R^2=0.24$)。

(注意: 在简单回归中 t 的自由度等于 $N-2$ 。)

简单回归的假定

1. 参与者间的观测是独立的。

该假定意味着每个参与者的得分应该相对于其他人的得分是独立的 (如果两个参与者一起进行心理健康测量是违反独立假定的一个例子)。违反这一假定会严重影响到使用回归方法进行统计检验的准确性。如果有理由相信独立假定已被违反, 那么不能使用回归方法。

2. 二元正态。

该假定意味着每个变量的总体都应该服从正态分布, 对于每个变量的值, 其他变量的得分应该服从正态分布。对于中等到较大的样本量, 绝大部分的非正态数据类型对回归方法的精确性没有多少影响。

3. 方差齐性。

方差齐性是指对于自变量的不同水平等级, 因变量的方差总体上是相等的。在简单回归中, 轻度或中度违反该假定是可以容忍的。

在 SPSS 中执行简单回归分析的步骤摘要

I. 数据输入和分析

1. 在 SPSS 中生成两个变量, 一个为自变量, 一个为因变量。

- 2. 输入数据。
- 3. 选择 **Analyze>Regression>Linear...**。
- 4. 将自变量和因变量移到各自的框中。
- 5. 点击 “Statistics”。点击 “Descriptives” (Estimates 和 Model fit 已经选中)。点击 “Continue”。
- 6. 点击 “OK”。

II. 结果解释

本章已经描述了所有的表格 (Descriptive Statistics, Correlations, Variables Entered/Removed, Model Summary, ANOVA 和 Coefficients)。在这个摘要中, 只会描述 Model Summary 和 Coefficients 表的信息。

- 注意 Model Summary 表中 R^2 的值。
- 检查 Coefficients 表中预测变量的 p 值 (或者在 ANOVA 表中等价的这些检验)。
- 如果预测变量显著 ($p \leq 0.05$), 拒绝原假设。写出表示自变量是因变量的显著预测变量的结果。
- 如果预测变量不显著 ($p > 0.05$), 不拒绝原假设。写出表示自变量不是因变量的显著预测变量的结果。

练习

- 1. 一个研究员调查父亲的乐观度能否预测他的儿子在青年时期的乐观度。20 对父亲和孩子参与研究并且按照标度给出各自的乐观程度 (范围为 1~50 分, 高的分数代表高乐观度), 图表 13—9 给出了数据。

图表 13—9 20 对父亲和孩子的乐观分数

Father	Son	Father	Son
40	45	49	39
30	35	46	49
25	20	23	38
29	35	46	35
20	22	26	28
25	35	16	19
46	48	29	45

续前表

Father	Son	Father	Son
46	31	31	36
49	41	42	45
37	31	43	48

在 SPSS 中输入数据并进行适当分析以回答下列问题。变量分别命名为 **father** (父亲) 和 **son** (儿子)。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 预测变量是否显著？在 $\alpha=0.05$ 时检验。
- d. 效应量是多少？你能区分效应量是小、中还是大吗？
- e. 写出数据的回归方程。
- f. 将研究结果写成适当的 APA 格式。

2. 一个学校的雇员想知道七年级数学技能入学考试的成绩能不能预测七年级数学成绩（由期末成绩衡量）。参加数学入学考试与期末考试的 30 名学生参与了研究。入学考试分数范围为 20~100 分，期末考试分数范围为 0~100 分（更高的分数表示更好的成绩）。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 13 中的文件 Chapter13 _ Exercise 2. sav 之中（变量命名为 **mathexam**（数学考试）和 **grade**（成绩））。在 SPSS 中打开数据并进行适当分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 预测变量是否显著？在 $\alpha=0.05$ 时检验。
- d. 效应量是多少？你能区分效应量是小、中还是大吗？
- e. 写出数据的回归方程。
- f. 将研究结果写成适当的 APA 格式。

3. 一位工业心理学家想要调查能否根据人的易相处性（与人相处时可爱的或友善的）预测工作成功。选择了 35 名工人，并给出了易相处性测量（范围为 7~35 分，高的分数表示好的易相处性）。还获得了管理者对工人的满意度（评价分数为 0~10 分，高的分数代表高的满意度）。数据在网站上文件夹 Chapter 13 中的文件 Chapter13 _ Exercise 3. sav 之中（变量命名为 **agreeable-**

ness (易相处) 和 **satisfaction** (满意度))。在 SPSS 中打开数据并进行适当分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 预测变量是否显著? 在 $\alpha=0.05$ 时检验。
- d. 效应量是多少? 你能区分效应量是小、中还是大吗?
- e. 写出数据的回归方程。
- f. 将研究结果写成适当的 APA 格式。



多元线性回归

多元线性回归的目标是用两个或两个以上的不同变量值来预测一个变量值。¹在多元线性回归中，被预测的变量称为因变量，用来预测的变量称为自变量或预测变量。下面给出一个应用多元线性回归的例子加以说明。

例子

为了研究方法课上的一个项目，一位学生想在大学生中研究沟通能力、乐观积极和学术成功能否预测生活的意义。他对 30 名同意参加研究的学生做了四项度量，见图表 14—1。

图表 14—1 在多元回归例子中的度量说明

尺度	度量什么	尺度范围
生活意义 (meaning)	目前个人生活情况	10~70；数值越高表明生活越有意义
沟通能力 (connect)	个人与他人如何沟通；个人需要他人支持的程度	8~40；数值越高表明越能得到他人帮助
乐观积极 (optimism)	个人对自己生活的乐观程度	10~50；数值越高表明越乐观
学术成功 (success)	个人在学术上的成功	5~25；数值越高表明越大的学术成功

多元回归的目标和数据要求

多元线性回归		
目标	数据要求	例子
用两个或多个自变量预测因变量	因变量 • 连续的 自变量 • 连续的或确定的 ²	因变量 • 生活意义 自变量 • 乐观积极 • 沟通能力 • 学术成功

原假设和对立假设

在多元线性回归中，每个预测变量（自变量）都有独立的原假设和对立假设。建立一个回归方程（方程将在后面介绍），检验每个自变量的回归系数（称为贝塔权重）是否显著不为 0。一个贝塔权重显著不为 0 表明这个自变量是因变量的显著预测变量。回归方程中的总体符号用希腊字母 β 表示。

每个预测变量的原假设是贝塔权重等于 0：

- $H_0: \beta_{\text{沟通}} = 0$ (H. 1) (沟通的贝塔权重等于 0；沟通不能预测生活意义)
- $H_0: \beta_{\text{乐观}} = 0$ (H. 2) (乐观的贝塔权重等于 0；乐观不能预测生活意义)
- $H_0: \beta_{\text{成功}} = 0$ (H. 3) (成功的贝塔权重等于 0；学术成功不能预测生活意义)

每个预测变量的对立假设是贝塔权重不等于 0：

- $H_1: \beta_{\text{沟通}} \neq 0$ (沟通的贝塔权重不等于 0；沟通可以预测生活意义)
- $H_1: \beta_{\text{乐观}} \neq 0$ (乐观的贝塔权重不等于 0；乐观可以预测生活意义)
- $H_1: \beta_{\text{成功}} \neq 0$ (成功的贝塔权重不等于 0；学术成功可以预测生活意义)

除了检验单个变量的显著性，在多元回归中我们还需要对包含所有自变量的回归方程能否显著预测因变量做出检验。如果回归方程是显著的，则自变量（整体）能够解释生活意义的方差的很大比重（它们可以预测生活意义）。用 R^2 来衡量能解释方差的比重（取值范围是 $0 \sim 1$ ）， $R^2 = 0$ 表示自变量不能解释生活意义的方差， $R^2 = 1$ 表示自变量能解释全部的生活意义的方差（实际 R^2 通常介于 0 和 1 之间）。

原假设常用于自变量（作为整体）不能解释生活意义的方差（即它们不能预测生活意义）的情况：

$$H_0: R^2 = 0 \quad (\text{H. 4})$$

对立假设常用于自变量（作为整体）能解释的生活意义的方差（即它们可以预测生活意义）的情况：

$$H_1: R^2 \neq 0$$

原假设评价

SPSS 中的多元回归程序可以对每个自变量的原假设（H. 1, H. 2 和 H. 3）和自变量（作为整体）不能预测因变量的原假设（H. 4）进行检验。给定一个检验，如果检验产生的结果在原假设正确时看起来不可能（结果发生的可能性小于 5%），那么拒绝原假设。如果检验产生的结果在原假设正确时看起来正确（结果发生的可能性大于 5%），那么不拒绝原假设。

研究问题

在研究中基本的兴趣问题也可以用研究问题的方式表示，例如，

单独检验（H. 1, H. 2 和 H. 3）

- “沟通能力可以预测生活意义吗？”
- “乐观积极可以预测生活意义吗？”
- “学术成功可以预测生活意义吗？”

所有自变量的整体检验（H. 4）

- “沟通能力、乐观积极和学术成功作为整体可以解释生活意义吗？”

数据

在图表 14—2 中列出了 30 名参与者在生活意义（**meaning**）、沟通能力（**connect**）、乐观积极（**optimism**）和学术成功（**success**）上的得分值。

图表 14—2 多元回归例子的数据

Participant	Meaning	Connect	Optimism	Success	Participant	Meaning	Connect	Optimism	Success
1	34	25	14	12	16	54	36	29	18
2	62	41	35	20	17	68	37	42	22
3	54	38	40	18	18	53	29	46	15
4	59	36	35	17	19	33	20	23	14
5	28	18	32	15	20	45	21	27	11
6	31	28	15	18	21	25	29	32	7
7	64	25	35	22	22	61	38	39	22
8	57	22	45	24	23	53	21	30	17
9	26	19	32	21	24	22	21	18	11
10	44	27	23	12	25	40	25	30	16
11	62	31	47	17	26	50	25	41	18
12	54	25	50	20	27	32	16	17	12
13	59	24	35	14	28	55	30	34	21
14	68	28	31	16	29	35	19	22	7
15	36	31	17	19	30	52	21	42	13

说明：变量 participant 包含在数据中，但不用输入 SPSS。

在 SPSS 中输入数据及分析

下面的步骤 1 和步骤 2 描述了怎样在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 14 中的文件 meaning in life. sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

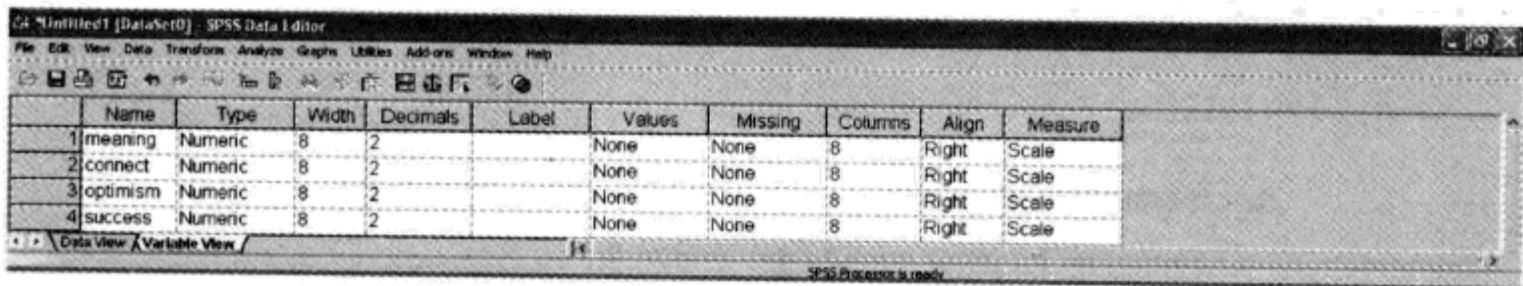
步骤 1：生成变量

- 1. 打开 SPSS。

2. 点击 Variable View 标签。

在 SPSS 中将生成四个变量，一个因变量（生活意义）和三个自变量（沟通能力、乐观积极和学术成功）。变量分别命名为 **meaning**, **connect**, **optimism** 和 **success**。

3. 在 Variable View 窗口前两行分别输入变量名称 **meaning**, **connect**, **optimism** 和 **success**（见图表 14—3）。



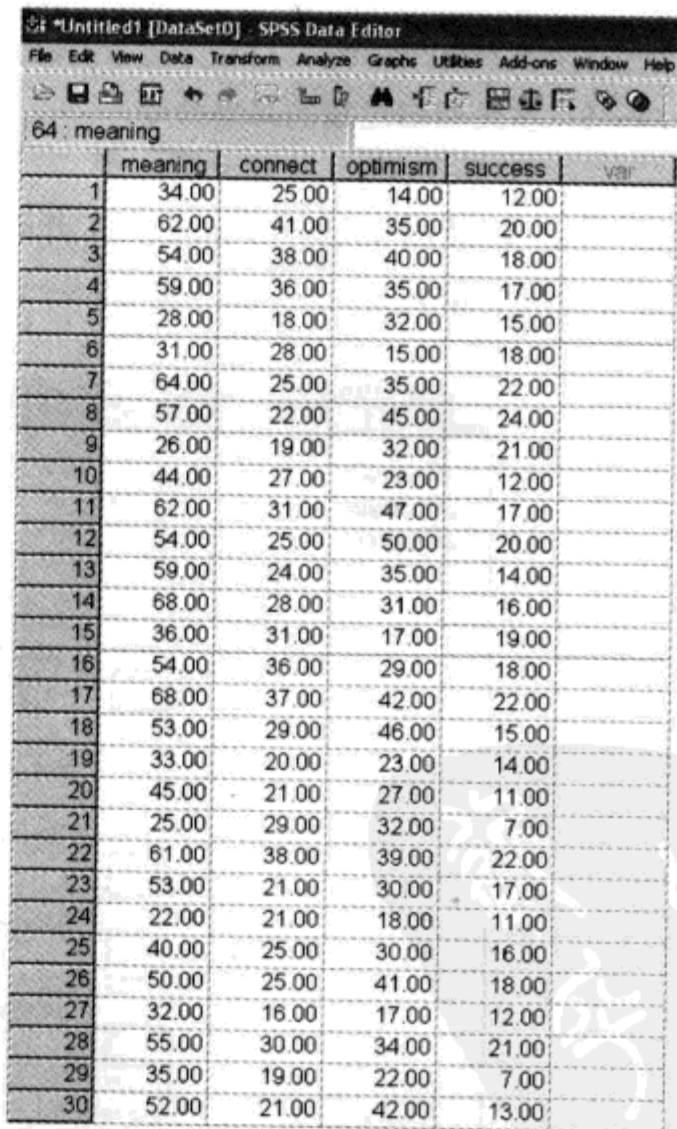
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	meaning	Numeric	8	2		None	None	8	Right	Scale
2	connect	Numeric	8	2		None	None	8	Right	Scale
3	optimism	Numeric	8	2		None	None	8	Right	Scale
4	success	Numeric	8	2		None	None	8	Right	Scale

图表 14—3 输入变量 **meaning**, **connect**, **optimism** 和 **success** 的 Variable View 窗口

步骤 2：输入数据

1. 点击 Data View 标签。变量 **meaning**, **connect**, **optimism** 和 **success** 出现在 Data View 窗口的前四列。

2. 参照图表 14—2，为每个参与者输入四个变量的数据。对第一个参与者，为变量 **meaning**, **connect**, **optimism** 和 **success** 分别输入数据 34、25、14 和 12。依次输入全部 30 名参与者的数据。图表 14—4 中给出了完整的数据文件。



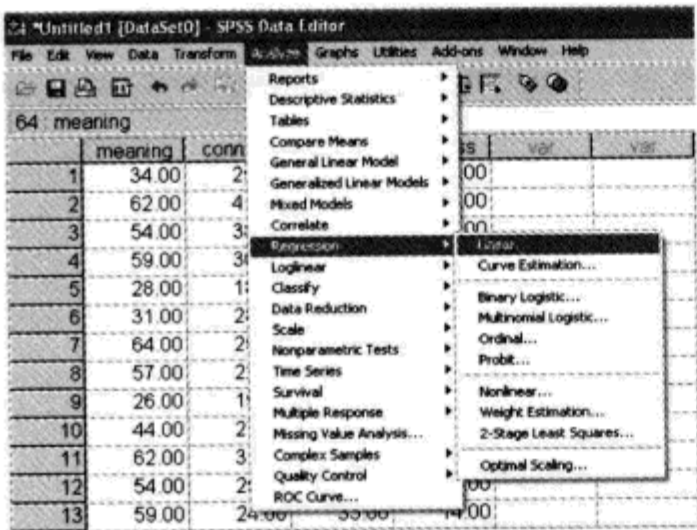
	meaning	connect	optimism	success	Var1
1	34.00	25.00	14.00	12.00	
2	62.00	41.00	35.00	20.00	
3	54.00	38.00	40.00	18.00	
4	59.00	36.00	35.00	17.00	
5	28.00	18.00	32.00	15.00	
6	31.00	28.00	15.00	18.00	
7	64.00	25.00	35.00	22.00	
8	57.00	22.00	45.00	24.00	
9	26.00	19.00	32.00	21.00	
10	44.00	27.00	23.00	12.00	
11	62.00	31.00	47.00	17.00	
12	54.00	25.00	50.00	20.00	
13	59.00	24.00	35.00	14.00	
14	68.00	28.00	31.00	16.00	
15	36.00	31.00	17.00	19.00	
16	54.00	36.00	29.00	18.00	
17	68.00	37.00	42.00	22.00	
18	53.00	29.00	46.00	15.00	
19	33.00	20.00	23.00	14.00	
20	45.00	21.00	27.00	11.00	
21	25.00	29.00	32.00	7.00	
22	61.00	38.00	39.00	22.00	
23	53.00	21.00	30.00	17.00	
24	22.00	21.00	18.00	11.00	
25	40.00	25.00	30.00	16.00	
26	50.00	25.00	41.00	18.00	
27	32.00	16.00	17.00	12.00	
28	55.00	30.00	34.00	21.00	
29	35.00	19.00	22.00	7.00	
30	52.00	21.00	42.00	13.00	

图表 14—4 多元回归例子的完整数据文件

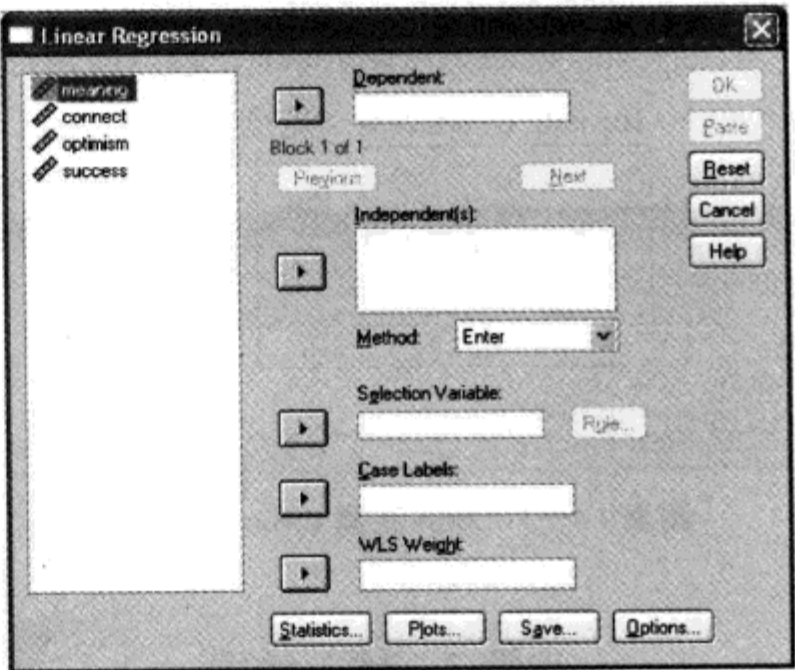
1. 从菜单栏中选择

Analyze>Regression>Linear... (见图表 14—5)。

打开 Linear Regression (线性回归) 对话框, 变量 **meaning**, **connect**, **optimism**和 **success** 出现在对话框的左边 (见图表 14—6)。

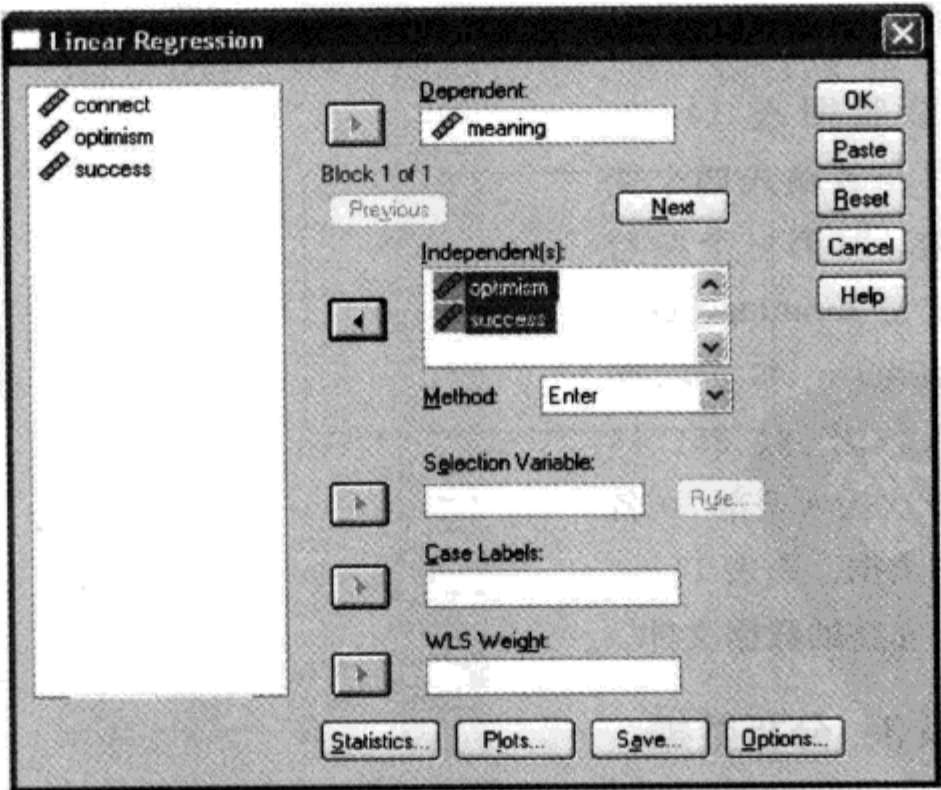


图表 14—5 多元回归过程的菜单命令



图表 14—6 线性回归对话框

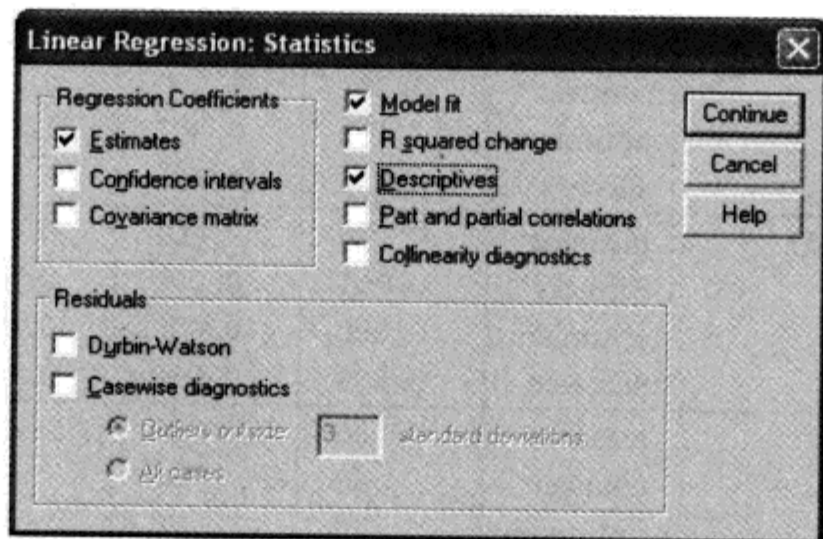
2. 选择因变量 **meaning**, 点击向右箭头按钮 (▶), 把变量移到 Dependent (因变量) 框中 (见图表 14—7)。



图表 14—7 线性回归对话框 (续)

3. 选择自变量 **connect**、**optimism** 和 **success**，单击第二个向右箭头按钮 (►)，把变量移到 Independent (s) (自变量) 框中，详见图表 14—7 (注：在图表 14—7 中，尽管只显示了两个自变量，但其实三个自变量都在 Independent (s) 框内)。

4. 单击 “Statistics” (统计量)。打开 Linear Regression: Statistics (线性回归：统计量) 对话框，选择 Descriptives (描述) (其中 Estimates (估计) 和 Model fit (模拟拟合度) 已被选择)。详见图表 14—8。



图表 14—8 线性回归：统计量对话框

5. 点击 “Continue”。

6. 点击 “OK”。

在 SPSS 中运行多元线性回归程序，结果显示在 Viewer 窗口中。

步骤 4：结果描述

多元回归过程的结果输出显示在图表 14—9。

描述统计量

表格 Descriptive Statistics 列出了每个变量的均值、标准差和样本量。

相关性

表格 Correlations 显示了这项研究中所有变量的二元相关性 (即两个变量之间的相互关系。关于相关性的介绍见第 12 章)。理想的情况是每个自变量都与因变量具有高度相关性 (较高的相关性有更好的预测趋势)，并且各个自变量之间没有高度相关性 (自变量之间相关性低，则它们可以对因变量有相对

Descriptive Statistics

	Mean	Std. Deviation	N
meaning	47.2000	13.90212	30
connect	26.8667	6.76570	30
optimism	31.9333	10.02044	30
success	16.3000	4.38768	30

自变量和因变量（生活意义）之间的相关性。

Correlations

	meaning	connect	optimism	success	
Pearson Correlation	meaning	1.000	0.561	0.663	0.555
	connect	0.561	1.000	0.307	0.425
	optimism	0.663	0.307	1.000	0.474
	success	0.555	0.425	0.474	1.000
Sig. (1-tailed)	meaning	0.000	0.001	0.000	0.001
	connect	0.001	0.000	0.050	0.010
	optimism	0.000	0.050	0.000	0.004
	success	0.001	0.010	0.004	0.000
N	meaning	30	30	30	30
	connect	30	30	30	30
	optimism	30	30	30	30
	success	30	30	30	30

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	success, connect, optimism ^a	.	Enter

^aAll requested variables entered.

^bDependent Variable: meaning

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.778 ^a	0.606	0.560	9.22060

^aPredictors: (Constant), success, connect, optimism

自变量解释了生活意义得分的方差的 61%。

图表 14—9 多元线性回归过程的输出结果

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3394.295	3	1131.432	13.308	0.000 ^a
	Residual	2210.505	26	85.019		
	Total	5604.800	29			

^aPredictors: (Constant), success, connect, optimism

^bDependent Variable: meaning

因为 p -值小于 0.05, 所以回归模型(包含所有自变量)是显著的。自变量整体解释了生活意义得分占显著比例的方差。

Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.971	8.281		-0.238	0.814
	connect	0.691	0.282	0.336	2.449	0.021
	optimism	0.653	0.196	0.470	3.334	0.003
	success	0.599	0.470	0.189	1.274	0.214

因为 **Connect** 和 **optimism** 的 p -值都小于 0.05,所以它们都是显著的。它们都解释了生活意义得分占显著比例的方差。

^aDependent Variable: meaning

图表 14—9 多元线性回归过程的输出结果 (续)

独立的贡献)。³ 表格 Correlations 显示了所有变量之间的相关性是显著的 (p 值均小于 0.05), 其中 **meaning** 和 **optimism** 之间的相关性最高 (为 0.66), **meaning** 和 **success** 之间的相关性最低 (为 0.56), 相关系数取值在 0.31 和 0.47 之间时, 自变量之间为中度相关。

变量进入/移出 (Variables Entered/Removed)

表格 Variables Entered/Removed 总结了预测生活意义的各个变量。注意到所有的自变量 (**success**、**connect** 和 **optimism**) 在 Variables Entered (输入的变量) 框中被列出。表格下方的标示 “all requested variables entered” (已输入所有要求的变量) 说明所有的变量同时包含在回归模型里, 备用程序允许先在回归模型中输入变量的一个子集, 然后输入其他自变量, 等等。这些程序中更常用的是分层回归 (hierarchical regression), 允许用户按照指定的顺序输入自变量, 用这种方法, 在预测因变量时, 一个或多个变量可以作为第一组变量, 一个或多个其他变量可以作为第二组变量, 依次类推。⁴

模型概要 (Model Summary)

表格 Model Summary 给出了 R , R^2 , 调整的 R^2 和估计标准误差。前三个

值度量生活意义可以被三个自变量预测的程度多大，最后一个值度量生活意义不能被三个自变量解释的程度。下面对这些值进行讨论。

第一个值 R 是多元相关系数，它等于生活意义的原始值和通过回归分析得到的生活意义的预测值之间的相关系数的绝对值 ($0 \leq R \leq 1$)。第二个值 R^2 是 R 的平方 ($0.778^2 = 0.61$)，当乘以 100% 时，可理解为因变量的总变异性中被自变量解释的百分比。在目前的例子中，自变量 **connect**，**optimism** 和 **success** 解释了生活意义的总变异性的 61% ($0.61 \times 100\%$)。调整的 R^2 修正了 R^2 是为了对总体值做更好的估计（基于样本计算的 R^2 容易高估总体值）。最后，估计标准误差表明自变量不能预测因变量值的程度。9.2206 说明用 **connect**，**optimism** 和 **success** 预测生活意义时，回归方程在预测生活意义时平均偏离大约 9.22。

方差分析——检验回归的整体显著性

表格 ANOVA 用来检验包含所有预测变量的回归模型预测生活意义的显著性（类似于检验 R^2 是否显著不为 0）。这个检验是用第 8 章讨论过的方差分析来进行的。在 ANOVA 表格中， p 值小于或等于 0.05 说明包含所有预测变量的回归模型能够显著地预测生活意义。因为 p 值 0.000（我们应该读为小于 0.001）是小于 0.05 的，所以原假设 $R^2 = 0$ 被拒绝，说明回归方程（包含三个预测变量）能够显著预测生活意义。

系数——检验单个预测变量的显著性

最后一个表格 Coefficients 提供了构建回归方程和检验每个预测变量显著性的必要值。

我们首先讨论回归方程，接下来说明 **connect**，**optimism** 和 **success** 是不是预测生活意义的显著预测变量。

在多元回归中，方程以下面的形式建立：

$$\hat{Y} = a + b_1 X_1 + b_2 X_2 + b_3 X_3$$

式中， \hat{Y} ——因变量的预测值，在我们的例子里， \hat{Y} 为生活意义的预测值；

a —— Y 轴截距，即当所有的 X_i 为 0 时 \hat{Y} 的值；

b_i ——第 i 个预测变量的回归系数。在这个例子中， i 的取值为 1、2、3，依次表示第一个变量（**connect**）、第二个变量（**optimism**）和第三个变量（**success**）；

X_i ——每一个参与者第 i 个自变量的取值。在我们的例子中, i 的取值为 1、2、3, 依次表示第一个自变量、第二个自变量和第三个自变量。

回归方程可以包含许多自变量, 例如, 如果包含第四个自变量, 可以在方程后添加 $+b_4 X_4$ 。

回归方程中的 a (Y 轴截距; 在 SPSS 中称为常数) 和 b_1, b_2, b_3 (每个自变量的回归系数) 可以在 Coefficients 表格中 Unstandardized Coefficients (非标准化系数) 一列的 “B” 标题下找到。 Y 轴截距的取值为 -1.971 , **connect**, **optimism** 和 **success** 的系数分别是 0.691 、 0.653 和 0.599 。将这些值代入回归方程中, 我们得到如下所示的预测生活意义的方程:

$$\hat{Y}_{\text{生活意义}} = -1.971 + 0.691 (\text{沟通}) + 0.653 (\text{乐观}) + 0.599 (\text{成功})$$

这是基于 30 名大学生的数据所得到的回归方程, 对这三个自变量分别给定一个取值, 可以得到一个生活意义的预测值。例如, 第一个参与者的 **connect**, **optimism** 和 **success** 取值分别为 25、14 和 12, 将这些数代入方程得到生活意义的一个预测值

$$\hat{Y}_{\text{生活意义}} = -1.971 + 0.691 (25) + 0.653 (14) + 0.599 (12)$$

$$\hat{Y}_{\text{生活意义}} = 31.634$$

使用这种方法可以求出数据库中每个人的预测值。⁵ 预测值大多会有些误差 (它们不会刚好等于真实值); R 值越大, 预测值越接近真实值, $R=1.0$ 得到完美的预测 (预测值与实际值完全一致)。

对回归系数方向的解释

每个预测变量的回归权重都是正值 (0.691 、 0.653 和 0.599) 的事实意味着 **connect** 的取值每增加 1 (例如从 25 到 26), 则生活意义的预测值将增加 0.691 (假设其他预测变量为常数保持不变)。一个负的回归权重则有相反的意义, 即预测变量的值增加, 将导致生活意义的预测值减少 (假设其他预测变量为常数保持不变)。强调回归系数的解释意义是建立在所有其他预测变量为不变的常数的假设下, 这是非常重要的, 但在实际中这是不可能的, 因为一般有较好的学术表现的人会更加乐观积极 (即预测变量是相关的)。

检验每个预测变量的显著性

在 Coefficients (系数) 表格里提供了对每个预测变量的显著性的检验。在该表的最后两列中给出了每个预测变量的 t 和 p 值。变量 **connect** ($t =$

2.449, $p=0.021$) 和 **optimism** ($t=3.334$, $p=0.003$) 都是显著的, 因为它们的 p -值都小于 0.05。然而, 变量 **success** 是不显著的, 因为它的 p -值 0.214 大于 0.05。

注意 Coefficients 表格里的 Standardized Coefficients (标准系数) 一列, 非标准化系数是用来预测得分的, 而标准系数则是用来对结果进行描述的。标准系数是贝塔权重, 它等于预测变量和因变量都是 z 得分 (z -score) 形式时的回归系数 (它们标准化为均值为 0, 标准差为 1)。^⑥

大多数情况下, 我们在多元回归建模时对常数项不感兴趣, 在 Coefficients 表格中报告的对常数项的检验是要看 Y 轴截距是否显著不为 0。在这个例子中, 由于其 p -值 0.814 大于 0.05, 因此该值 -1.971 并不是显著不为 0。

效应量

在回归中效应的度量是用 R^2 给出的。科恩 (Cohen, 1988) 定义回归中 R^2 的值 0.02、0.13 和 0.26 分别表示为小、中和大的效应量。在上面的例子中, $R^2=0.61$ 被认为在实际中有非常大的效应量, 这表明生活意义得分的变异性的 61% 能被预测变量解释。

APA 格式的结果表达

在写多元回归的结果时, 要报告对包含所有预测变量的回归模型的检验 (包括 R^2 和 ANOVA 检验结果) 和对单个预测变量的检验 (包括 β , t 值和 p -值)。接下来将展示一个简单的书面例子。

用沟通能力、乐观积极和学术成功三个变量建立一个多元回归模型来预测生活意义, 总的来说这个回归模型是显著的, $F(3, 26) = 13.31$, $p < 0.05$, $R^2 = 0.61$ 。在被调查的预测变量中, 沟通能力 ($\beta = 0.34$, $t(26) = 2.45$, $p < 0.05$) 和乐观积极 ($\beta = 0.47$, $t(26) = 3.33$, $p < 0.05$) 是显著的预测变量, 学术成功是不显著的预测变量, 其 $\beta = 0.19$, $t(26) = 1.27$, $p > 0.05$ 。^①

① 每个 t 检验的自由度等于 $N-p-1$, 其中 p 等于预测量的个数。

多元回归的假定

1. 参与者间的观测是独立的。

该假定意味着每个参与者的得分应该相对于其他任何人的得分是独立的（在整个测量过程中参与者不能相互影响）。违反这一假定会严重影响到使用多元回归方法进行统计检验的准确性。⁷ 如果有理由相信独立假定已被违反，那么不能使用多元回归方法。

2. 变量服从多元正态分布。

该假定意味着每个变量自身服从正态分布，而且对于其他变量的所有可能联合也服从正态分布（例如，对与生活意义、沟通能力和乐观积极的所有可能组合，学术成功得分服从正态分布）。对于中等到较大的样本量，绝大部分的非正态数据类型对回归方法的精确性没有多少影响。

3. 方差齐性。

方差齐性是指对于自变量所有可能的水平组合，因变量的方差总体上是相等的。在多元回归中，轻度或中度违反该假定是可以容忍的。

在 SPSS 中执行多元回归分析的步骤摘要

I. 数据输入和分析

1. 在 SPSS 中生成每个自变量和因变量。
2. 输入数据。
3. 选择 **Analyze > Regression > Linear...**
4. 将自变量和因变量移到它们各自的框内。
5. 点击 “Statistics” （统计量）。选择 Descriptives （其中 Estimates 和 Model fit 已被选择）。点击 “Continue”。
6. 点击 “OK”。

II. 结果解释

在本章中，已经描述了所有的表格（Descriptive Statistics, Correlations, Variables Entered/Removed, Model Summary, ANOVA 和 Coefficients）。在这个摘要中，只描述 Model Summary, ANOVA 和 Coefficients 表的信息。

- a. 在 Model Summary 表格中说明了 R^2 的值；在 ANOVA 表格中，如果 $p \leq 0.05$ ，则原假设 $R^2 = 0$ 被拒绝，说明回归模型（包含所有预测变量）能解释因变量方程的很大的比例。
- b. 在 Coefficients 表格中，检验每个预测变量的 p 值：
 - 如果一个或多个自变量是显著的 ($p \leq 0.05$)，那么拒绝自变量的 β 值为 0 的原假设。写出表示自变量是因变量的显著预测变量的结果。同理也可说明任何不显著的自变量。
 - 如果没有任何一个自变量是显著的（在 Coefficients 表格中每个自变量都是 $p > 0.05$ ），那么对每个自变量的原假设都不拒绝。写出表示自变量不是因变量的显著预测变量的结果。

练习

1. 一位工业心理学家调查了不同公司的 30 个雇员的工作满意度的预测变量。调查的预测变量包括工作重要性 (**importance**)、在公司中晋升的机会 (**advance**) 和向老板表达观点的能力 (**express**)，因变量是工作满意度 (**satisfaction**)。每个变量得分值越大说明感兴趣的特征越高（例如，满意度的分值越大说明对工作越满意）。图表 14—10 提供了相应数据。

图表 14—10 练习 1 的数据

Satisfaction	Importance	Advance	Express	Satisfaction	Importance	Advance	Express
35	20	12	18	41	21	29	17
48	24	22	20	33	15	24	17
22	10	17	12	40	21	23	17
38	8	25	12	48	19	21	25
33	20	10	17	45	18	13	22
45	15	30	15	23	10	22	18
30	12	20	19	45	15	10	20
38	22	9	17	27	22	21	21
45	21	18	23	12	12	21	8
23	12	14	11	12	6	18	17
13	8	16	15	41	22	31	15

续前表

Satisfaction	Importance	Advance	Express	Satisfaction	Importance	Advance	Express
29	17	25	15	42	22	19	21
42	19	21	19	46	19	15	17
30	20	27	15	25	8	21	19
39	24	19	18	41	17	12	23

把数据输入 SPSS 并且进行必要的分析以回答下面问题。分别定义变量名为 **satisfaction**, **importance**, **advance** 和 **express**。

- a. 写出每个变量的原假设和对立假设。
- b. 写出整个回归模型的原假设和对立假设。
- c. 分别为每个变量和整个回归模型写出研究问题。
- d. 这个模型的 R^2 是什么？整个回归模型是显著的吗？
- e. 哪个预测变量是显著的？哪个预测变量是不显著的？
- f. 整个回归模型的效应量是多少？你能区分效应量是小、中还是大吗？
- g. 为这组数据写出回归方程。
- h. 将结果写成适当的 APA 格式。

2. 一位研究员对用变量“宽恕”和“社会支持”来预测幸福指数感兴趣，他对 25 个同意参与这项研究的人采取了图表 14—11 所描述的三种度量。

练习 2 的变量		
尺度	度量什么	尺度范围
幸福 (happiness)	一个人近期在生活中的幸福程度	7~35；分值越大表示越幸福
宽恕 (forgiveness)	一个人可以宽恕他人的程度	10~50；分值越大表示越大的宽恕意愿
社会支持 (support)	一个人觉得可以得到他人支持的程度	5~30；分值越大表示越强的支持网络

数据在网站 www.prenhall.com/yockey 的文件夹 Chapter14 中的文件 Chapter14 _ Exercise2. sav 之中。在 SPSS 中打开这个文件并且进行必要的分析以回答下列问题。

- a. 写出每个变量的原假设和对立假设。
- b. 写出整个回归模型的原假设和对立假设。
- c. 分别为每个变量和整个回归模型写出研究问题。

- 这个模型的 R^2 是什么？整个回归模型是显著的吗？
- 哪个预测变量是显著的？哪个预测变量是不显著的？
- 整个回归模型的效应量是多少？你能区分效应量是小、中还是大吗？
- 为这组数据写出回归方程。
- 将研究结果写成适当的 APA 格式。

3. 一位研究员对用变量“幸福感”和“生活意义”来预测对死亡的恐惧指数感兴趣。他对 50 个同意参加这项研究的人采取了图表 14—12 所描述的三种措施。

图表 14—12

尺度	度量什么	尺度范围
对死亡的恐惧指数 (feardeath)	一个人面临死亡时的恐惧程度	8~48; 分值越大表示对死亡越恐惧
幸福感 (wellbeing)	一个人近期生活的幸福程度	5~40; 分值越大表示越幸福
生活意义 (meaning)	一个人近期生活的意义	10~70; 分值越大表示生活越有意义

数据在网站上文件夹 Chapter 14 中的文件 Chapter14 _ Exercise3. sav 之中。在 SPSS 中打开这个文件并且进行必要的分析以回答下列问题。

- 写出每个变量的原假设和对立假设。
- 写出整个回归模型的原假设和对立假设。
- 分别为每个变量和整个回归模型写出研究问题。
- 这个模型的 R^2 是什么？整个回归模型是显著的吗？
- 哪个预测变量是显著的？哪个预测变量是不显著的？
- 整个回归模型的效应量是多少？你能区分效应量是小、中还是大吗？
- 为这组数据写出回归方程。
- 将研究结果写成适当的 APA 格式。

χ² 拟合优度检验

χ² 拟合优度检验对变量的每个类别里的案例比例（或频数）是否与原假设下指定的比例（或频数）一致提供了检验。下面给出一个应用 χ² 拟合优度检验的例子。

例子

一位对社会心理学感兴趣的学生想调查认知责任感是否随团体大小而变化（即人们责任感的不同是否基于所在团体的大小）。他对参与这个项目的 90 名大学生描述了一位妇女需要帮助的三个场景，这三个场景的不同之处在于附近可以帮助这位妇女的人数有所不同，在这三个场景中团体大小分别为 1 个人、5 个人和 25 个人。在这三个场景中，参与者被告知没有人帮助那位妇女。当描述场景时，参与者被要求在每一组找出一个人，并且指出在三种情形中的哪种情况，这个人为无法帮助女士负最多的责任（他们只能选择三个场景中的一个）。

χ² 拟合优度检验的目标和数据要求

χ² 拟合优度检验		
目标	数据要求	例子
检验样本中的比例（或频数）是否显著不同于原假设下指定的比例（或频数）	一个有两种或更多类别的分类变量 （根据落入哪一类别）对每个参与者提供一个频数	变量：团体大小（1 人、5 人、25 人）

原假设和对立假设

原假设表明认知责任感不会随团体的大小而变化，这意味着在这三个场景下有一个相等的比例（或者人数）：

H_0 ：认知责任感不会随团体的大小而变化（在总体中，选择每个场景的人数或所占比例是相同的）

对立假设表明认知责任感会随团体的大小而变化，这意味着在这三个场景下比例（或者人数）是不相等的：

H_1 ：认知责任感会随团体的大小而变化（在总体中，选择每个场景的人数或所占比例是不相同的）

原假设评价

χ^2 拟合优度检验对每个类别里的案例比例与原假设下指定的比例是否一致进行了检验。如果检验产生的结果在原假设正确时看起来不可能（结果发生的可能性小于 5%），那么拒绝原假设。如果检验产生的结果在原假设正确时看起来正确（结果发生的可能性大于 5%），那么不拒绝原假设。

研究问题

在研究中基本的兴趣问题也可以用研究问题的方式表示，例如，

“认知责任感是否随团体的大小而变化？”

数据

在图表 15—1 中列出了 90 个参与者的数据。在 90 个参与者中，42 个人选择了 1 人场景最有责任，30 个人选择了 5 人场景最有责任，18 个人选择了 25 人场景最有责任。由于这些数据是在研究中观测到的，因此称为观测频数（observed frequencies）。

图表 15—1 选择 1 人场景、5 人场景和 25 人场景作为最有责任的参与者的数量（观测频数）

1 人场景	5 人场景	25 人场景
42	30	18

χ^2 拟合优度检验将对研究中的观测频数（图表 15—1 所列出的那些）与在原假设下频数的期望值（称为期望频数）相比较。为了确定期望频数，我们首先需要找出在原假设为真的条件下选择每个类别的人数的比例。因为原假设表明认知责任感不会随团体的大小而变化，所以每个类别都应该有 1/3 的参与者选择，如图表 15—2 所示。

图表 15—2 在原假设为真的条件下参与者选择每个类别的期望比例

1 人场景	5 人场景	25 人场景
1/3	1/3	1/3

在研究中一旦确定了每个类别的比例，用总样本量（N）乘以每个比例就可以得到期望频数。

计算每个类别的期望频数的公式如下：

某个类别的期望频数 = 在这个类别中的期望比例 \times N

将这个公式应用到每个类别得到如下期望频数：

“1 人场景”下的期望频数 = $(1/3) \times 90 = 30$

“5 人场景”下的期望频数 = $(1/3) \times 90 = 30$

“25 人场景”下的期望频数 = $(1/3) \times 90 = 30$

将每个类别下的期望频数总结到图表 15—3 中。

图表 15—3 在原假设为真的条件下选择每个类别的期望人数（期望频数）

1 人场景	5 人场景	25 人场景
30	30	30

χ^2 拟合优度检验将图表 15—1 中的观测频数与图表 15—3 中的期望频数相比较。一般来说，观测频数与期望频数的差异越大，原假设被拒绝的可能性越大。¹

在把数据输入 SPSS 之前，我们将讨论 χ^2 检验过程的两种不同的数据输入方法。这两种方法是加权（weight cases）和个体观测值（individual obser-

vations) 输入方法, 不同之处在于如何构造数据文件。下面将描述这两种方法。

案例加权方法

用案例加权方法 (Weight Cases Method) 来输入数据, 要求 Data View 窗口的行数等于变量的类别数量 (这个例子中对于三个不同的团体大小有 3 行数据)。在 SPSS 中需要两个变量: 一个是不同的类别, 一个是选择每个类别的人数 (即频数)。当每个类别的频数已经计算出来时 (如图表 15—1 中的数据), 加权方法是一种常用的方法。

个体观测值方法

用个体观测值方法 (Individual Observations Method) 来输入数据, 要求 Data View 窗口的行数等于研究中的人数 (这个例子中为 90 个参与者设有 90 行数据)。在 SPSS 中只需要一个变量: 每个类别的参与者给定一个值 (如: 1、2 或 3)。当没有计算出所有类别的频数而是直接输入一个文件或调查表的形式时, 这个方法最有用。

图表 15—4 总结了加权方法和个体观测值方法的数据输入特征。

图表 15—4 χ^2 拟合优度检验过程中加权和个体观测值方法的数据输入特征

方法	SPSS 中需要的变量个数	Data View 窗口中需要的行数	何时应用
加权	两个——一个是不同类别, 一个是频数	等于变量的类别数量 (我们的例子中为 3)	每个类别的频数已经计算出来 (如图表 15—1)
个体观测值	一个——不同类别	等于研究中的人数 (我们的例子中为 90)	没有计算出所有类别的频数 (如从调查表中输入数据)

由于已经计算出了每个类别的频数, 我们采用加权方法输入数据, 这在 SPSS 中需要两个变量 (对于更喜欢采用个体观测值方法输入数据的人, 图表 15—7 描述了如何用这种方法构造数据文件, 本章末的练习题 2 是一个例子)。

在 SPSS 中输入数据及分析

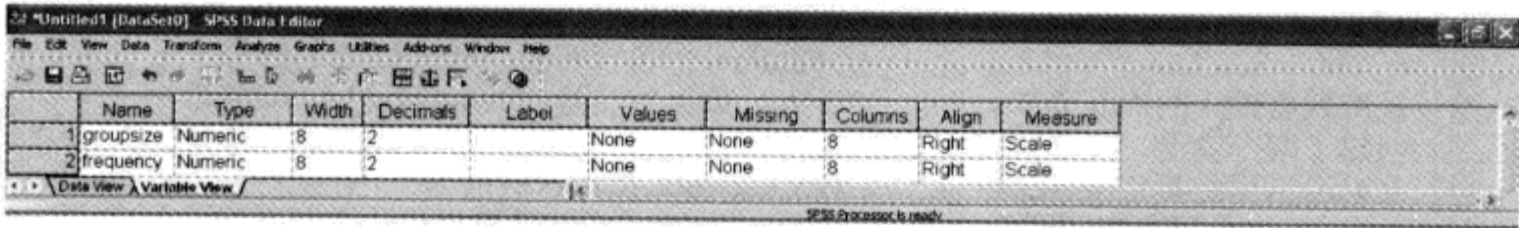
下面的步骤 1 和步骤 2 描述了怎样在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 15 中的文件 no help.sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中将生成两个变量，一个变量反映不同团体大小的类别（1 人、5 人和 25 人），另一个变量是频数（选择认为最有责任感的每个团体的人数）。这两个变量分别命名为 **groupsize** 和 **frequency**。

- 3. 在 Variable View 窗口前两行输入变量名称 **groupsize** 和 **frequency**（详见图表 15—5）。



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	groupsize	Numeric	8	2		None	None	8	Right	Scale
2	frequency	Numeric	8	2		None	None	8	Right	Scale

图表 15—5 输入变量 **groupsize** 和 **frequency** 的 Variable View 窗口

接下来，我们为研究中的不同类别建立变量值标签。我们定义“1 人类别”为 a1，“5 人类别”为 a2，“25 人类别”为 a3。

- 4. 用第 1 章中描述的过程，为变量 **groupsize** 建立变量值标签。对于 **groupsize**，1=“1 人”，2=“5 人”，3=“25 人”。

步骤 2：输入数据

点击 Data View 标签。变量 **groupsize** 和 **frequency** 出现在 Data View 窗口的前两列。

由于采用了数据输入的加权方法，我们需要在数据文件里为研究中三个团体大小类别的每一个创建独立的一行。用加权方法建立的数据文件结构如图表 15—6 所示。

图表 15—6 用数据输入的加权方法得到的 groupsize 和 frequency 的值

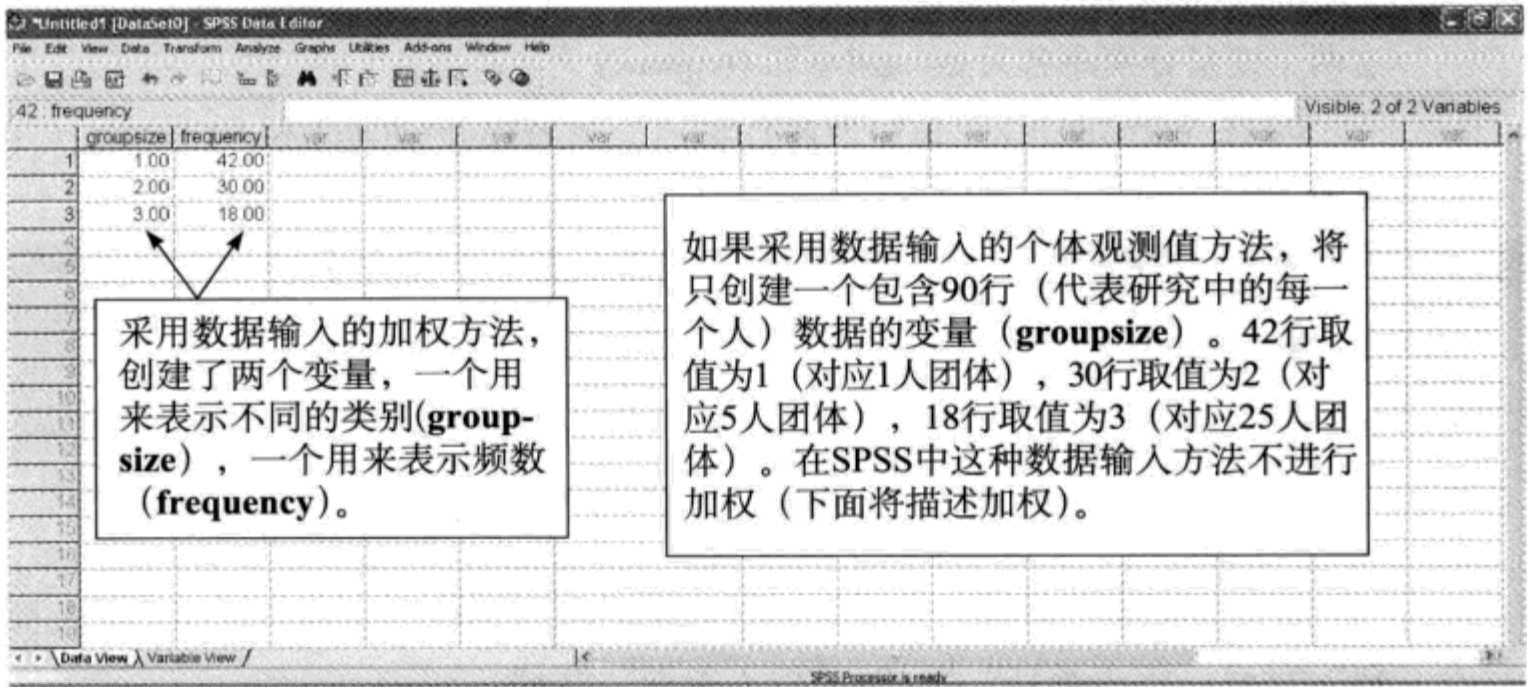
Groupsize	Frequency
1	42
2	30
3	18

说明：对于 groupsize，1 对应 “1 人类别”，2 对应 “5 人类别”，3 对应 “25 人类别”。

按照图表 15—6，在 SPSS 中我们将输入三行数据，每一行对应于研究中的不同类别。

输入数据

在 Data View 窗口的第一行，分别对应变量 **groupsize** 和 **frequency** 输入 1 和 42。在第二行，分别对应 **groupsize** 和 **frequency** 输入 2 和 30。在第三行，输入 3 和 18。图表 15—7 给出了完整的数据文件。



图表 15—7 χ^2 拟合优度例子的完整数据文件

步骤 3：数据分析

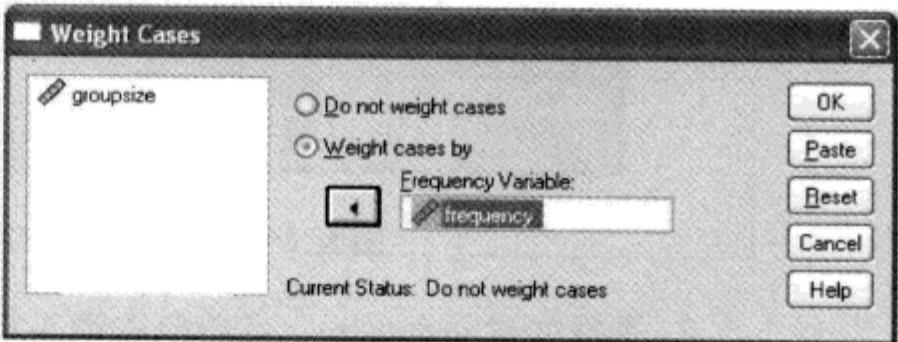
在运行 χ^2 分析之前，我们首先需要对 **frequency** 进行加权。² 对案例加权表明给定变量的值表示观测总次数，而不仅仅是一个分数值。例如，对 **frequency** 进行加权时，第一个类别的取值 42 对应于 42 个人，而不是分数为 42。

对 frequency 进行加权

- 1. 在菜单栏中选择 **Data>Weight Cases...**（见图表 15—8）。
- 2. 打开 Weight Cases 对话框。选择 Weight cases by 并选择变量 **frequency**。点击向右箭头按钮（▶）使 **frequency** 移到 Frequency Variable 框中（见图表 15—9）。



表 15—8 加权过程的菜单命令

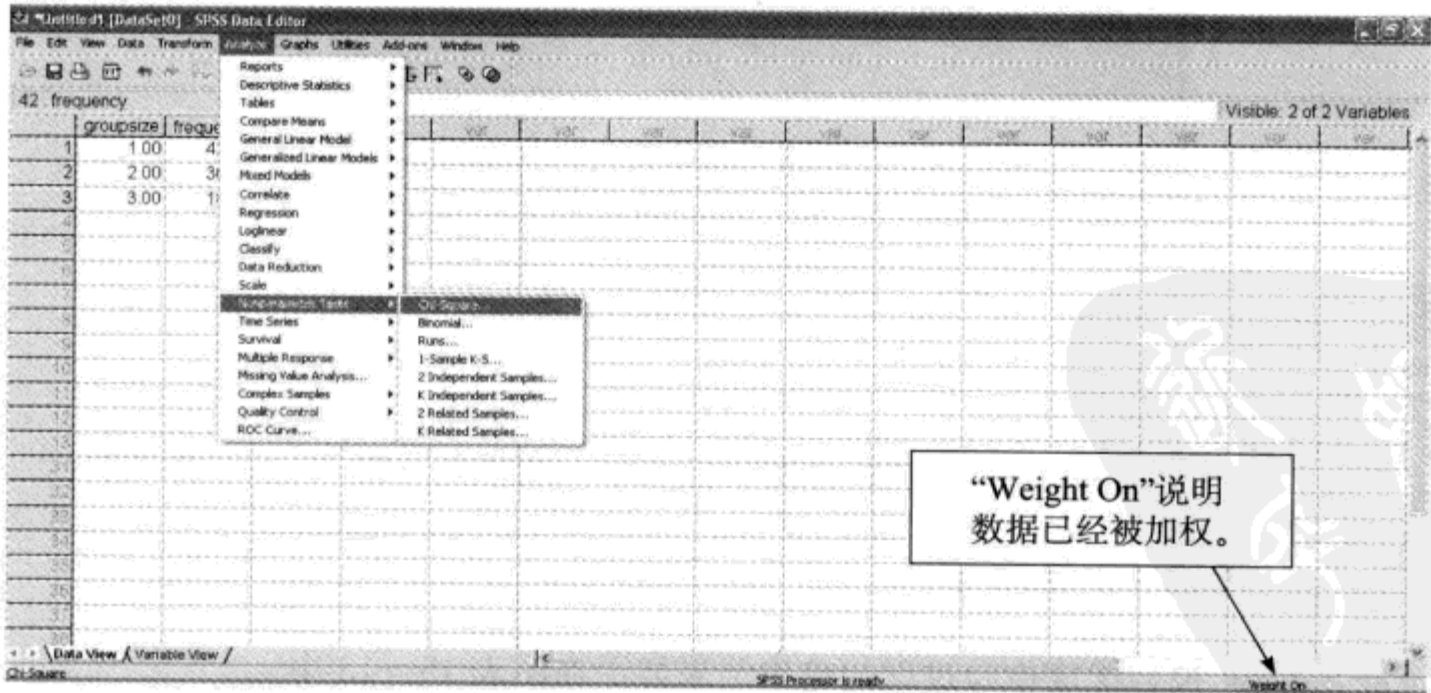


图表 15—9 加权的对话框

3. 点击“OK”。这表示在每个类别中频数的取值（42、30、18）对应于观测总次数，而不是一个分数。³

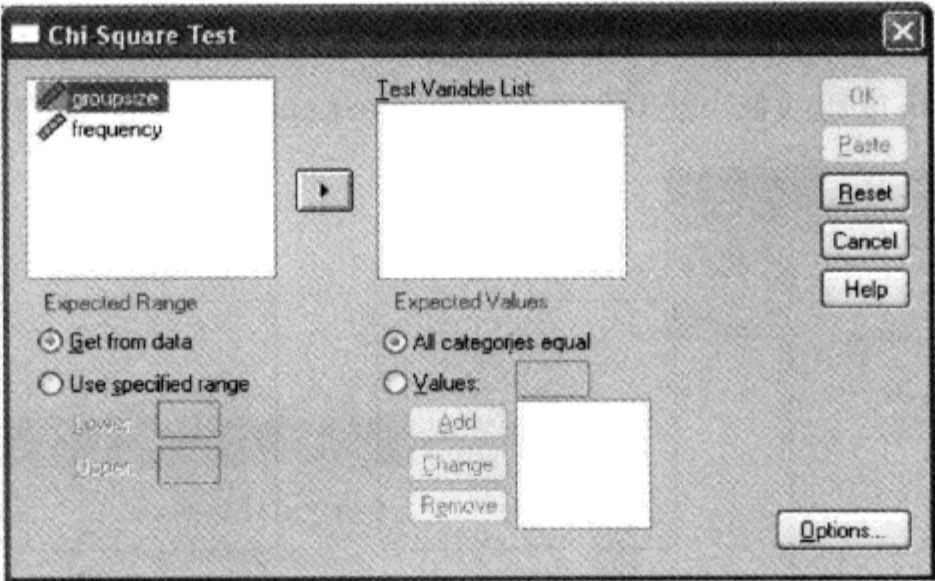
通过对 **frequency** 进行加权，现在我们在 SPSS 中执行 χ^2 拟合优度检验。

- 1. 在菜单栏中选择 **Analyze>Nonparametric Test（非参数检验）>Chi-Square（卡方检验）...**（见图表 15—10）。



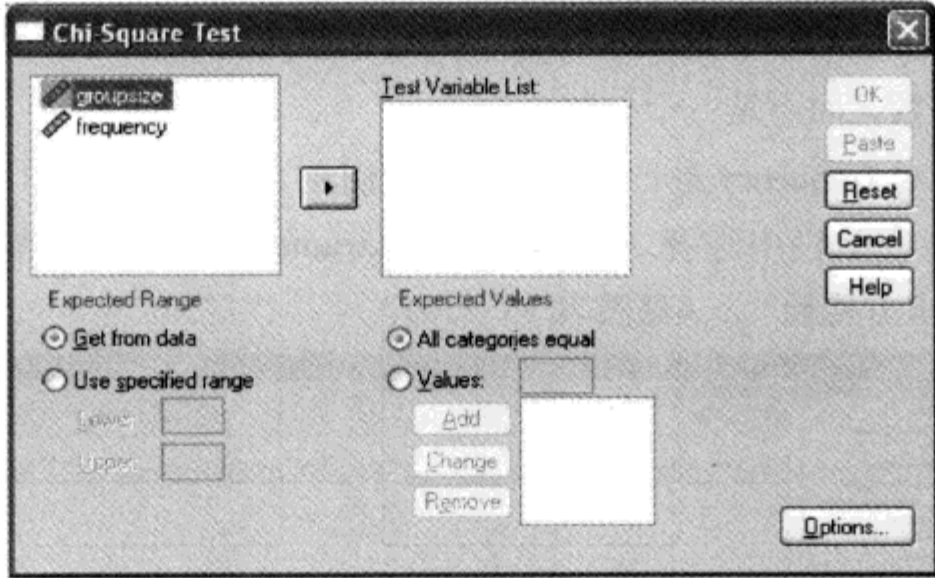
图表 15—10 χ^2 拟合优度检验过程的菜单操作

打开 χ^2 检验对话框，变量 **groupsize** 和 **frequency** 出现在对话框的左边（见图表 15—11）。



图表 15—11 χ^2 检验对话框

2. 选择变量 **groupsize**，点击向右箭头按钮（►）使其移到 Test Variable List 框中（见图表 15—12）。（注意：**frequency** 仍然在对话框的左边。）



图表 15—12 χ^2 检验对话框（续）

3. 点击 “OK”。

在 SPSS 中运行 χ^2 拟合优度检验程序，结果显示在 Viewer 窗口中。

步骤 4：结果描述

χ^2 拟合优度检验过程的输出结果显示在图表 15—13 中。

Npar Tests
Chi-Square Test
Frequencies

groupsize			
	Observed N	Expected N	Residual
1 person	42	30.0	12.0
5 person	30	30.0	0.0
25 person	18	30.0	-12.0
Total	90		

Test Statistics

	groupsize
Chi-Square ^a	9.600
df	2
Asymp. Sig	0.008

^a0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 30.0.

图表 15—13 χ^2 拟合优度检验的输出结果

团体大小 (Groupsize)

第一个表格 groupsize 报告了观测频数 (Observed N), 期望频数 (Expected N) 和残差, 或者是在 1 人、5 人和 25 人团体中观测频数与期望频数值的差异。请注意, 在 1 人和 25 人团体中, 观测频数和期望频数值之间的差异都为 12 (在 1 人团体中 Observed N > Expected N, 而在 25 人团体中 Observed N < Expected N)。下面我们会考虑观测频数和期望频数之间的差异是否大到足以具有统计显著性。

检验统计量 (Test Statistics)

表格 Test Statistics 为我们的研究问题提供了答案, 也就是认知责任感是否随团体大小而变化。 χ^2 拟合优度检验产生了一个 χ^2 统计量, 由以下公式给出:

$$\chi^2 = \sum \frac{(\text{观测频数} - \text{期望频数})^2}{\text{期望频数}} \text{①}$$

① 原书为 $\chi^2 = \sum \frac{(\text{观测频数} - \text{期望频数})^2}{\text{期望频数}^2}$, 有误。——译者注

其中，观测频数和期望频数分别对应于表格 groupsize 中的 Observed N 和 Expected N。把表格 groupsize 中的观测频数和期望频数代入公式得到一个卡方 (χ^2) 的值为

$$\chi^2 = \frac{(42-30)^2}{30} + \frac{(30-30)^2}{30} + \frac{(18-30)^2}{30}$$

$$\chi^2 = 9.60$$

这个值与表格 Test Statistics 中的 Chi-Square (卡方) 值是相等的。卡方检验的自由度为 2 (df =类别的数目-1)，其 p -值渐近 (Asymp. Sig.) 为 0.008。因为 p -值 < 0.05 ，所以原假设三个团体频数相等被拒绝，说明认知责任感随团体大小不同而不同。

由于结果是显著的，我们将检验观测频数和期望频数在表格 groupsize 中的差异，以确定哪个团体因没有进行救助而负有最大 (或最小) 责任。对于 1 人团体，观测频数是 42 而期望频数是 30，这表明选择 1 人团体的人比预测的多 12 个。对于 5 人团体，观测频数和期望频数没有不同 (观测频数是 30，期望频数也是 30)。对于 25 人团体，选择该组别的人比预测的少 12 个 (其观测频数是 18，而期望频数是 30)。总的说来，选择 1 人团体的人比预期的多而选择 25 人团体的比预期的少 (对于 5 人团体则没有不同)。这说明与处于较大的团体中相比，人们在独处时会有更高的责任感去帮助别人。

APA 格式的结果表达

在写结果时，要报告原假设的结论、自由度、样本量、 χ^2 值以及 p -值。对显著性检验结果还要描述观测频数和期望频数之间的关系。下面给出一个用 APA 格式做的简要描述例子。

未能帮助一个需要帮助的人的认知责任感随团体大小而变化， χ^2 (2, $N=90$) = 9.6, $p < 0.05$ 。对于未能帮助一个需要帮助的人的认知责任感，参与者认为一个人时的责任感最高，而处于 25 人团体中的那些人责任感最低。认为 1 人团体，5 人团体或 25 人团体责任感最高的人所占比例分别为 46.7%，33.3% 和 20.0%。⁴

(注：在上述结果中位于 χ^2 右侧的 2 是指自由度。在括号中标明了样本量，由于与本书中其他方法不同，并不能由自由度来决定 χ^2 检验的样本量。)

χ^2 拟合优度检验的假定

1. 观测是独立的。

独立性假定意味着每个单元的观测（即参与者）是彼此独立的。对一个人记录不同的两个单元（例如，每个参与者被要求提供责任感最重的前两个）是违反这个假定的一个例子。违反了独立性假定会严重影响卡方检验的准确性。如果我们有理由相信已经违反了独立性假定，那么就不应该再用 χ^2 拟合优度检验。

2. 期望单元频数的规模。

对于小的表格（有4个单元或更少），建议每个单元的频数为5或更多。对于大的表格，有一个或几个单元的期望频数小于5，则很可能不会使得 χ^2 拟合优度检验无效，在样本量很小时尽管效果会有损失。

在 SPSS 中执行 χ^2 拟合优度检验的步骤摘要

已经给出了在 SPSS 中用数据输入的加权方法和个体观测值方法执行一个 χ^2 拟合优度检验的说明（前面已提到，用个体观测值方法创建数据文件的说明在本章末的练习题2中）。

Ia. 数据输入和分析——加权方法

1. 在 SPSS 中生成两个变量（一个是分类变量，一个是频数变量）。
2. 为分类变量建立变量值标签。在 Value Labels 对话框中输入适当的数值和标签。点击“OK”。
3. 输入数据（对于数据输入的加权方法，数据文件中的行数等于类别变量数目）。
4. 通过选择 **Data > Weight Cases...**，对频数变量进行加权处理。
5. 选择 Weight cases by，把频数变量移到 Frequency Variable 框中。点击“OK”。
6. 选择 **Analyze > Nonparametric Tests > Chi-Square...**。
7. 把类别变量移到 Test Variable List 框中（频数变量仍然在对话框的左侧）。

8. 点击 “OK”。

Ib. 数据输入和分析——个体观测值方法

- 1. 在 SPSS 中生成一个（类别）变量。
- 2. 为这个变量建立变量值标签。在 Value Labels 对话框中输入适当的数值和标签。点击 “OK”。
- 3. 输入数据（对于数据输入的个体观测值方法，数据文件的行数等于研究中参与者的数量）。
- 4. 选择 **Analyze>Nonparametric Tests>Chi-Square...**。
- 5. 把变量移到 Test Variable List 框中。
- 6. 点击 “OK”。

II. 结果描述（不管采用哪种数据输入方法，解释是一样的）

在表格 Test Statistics 中，检查 p -值（以 Asymp. Sig 列出）。

- 如果 $p \leq 0.05$ ，拒绝原假设。写出表示观测频数和期望频数之间关系的结果。
- 如果 $p > 0.05$ ，不拒绝原假设。写出表示观测频数和期望频数之间没有显著不同的结果（例如，任何一个类别都没有优先权）。

练习

1. 一位研究员对调查眼睛大小对吸引力的影响感兴趣。在研究中，用到两个不同人的照片，在进行研究前对照片进行了处理，使得其中一个人的眼睛看起来足够大。让 80 个人看了这两张照片并要求他们说出这两张中的哪一个人更有吸引力。图表 15—14 给出了数据。

图表 15—14 选择更有吸引力的照片的参与者人数

大眼睛	小眼睛
60	20

把数据输入 SPSS，进行必要分析以回答下列问题。变量名定义为 **eyesize** 和 **frequency**（在分析数据之前先进行加权）。对于 **eyesize**，大眼睛编码为 1，小眼睛为 2。

- a. 陈述原假设和对立假设。

- b. 陈述数据的一个研究问题。
- c. 其中一张照片有更大的吸引力吗？在 $\alpha=0.05$ 时检验。
- d. 将研究结果写成适当的 APA 格式。

2. 一位市场调研员要调查人们对等离子电视机和 LCD 电视机的画面质量是否有不同的感觉。他用等离子电视机和 LCD 电视机给当地的购物者放映了一场电影，并且要求他们说明两台电视机中哪一台画面质量更好。图表 15—15 提供了研究中的 65 位购物者的选择。

图表 15—15 分别认为等离子电视机和 LCD 电视机画面质量最好的人数

等离子	LCD
35	30

数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter15 中的文件 Chapter15 _ Exercise 2. sav 之中（变量命名为 **tv**，其中等离子 **tv** 编码为 1，LCD **tv** 编码为 2）。数据文件已经用个体观测值方法输入了（即这个例子不需要进行加权）。在 SPSS 中打开文件，进行必要分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 人们对等离子电视机和 LCD 电视机的画面质量感受有不同吗？在 $\alpha=0.05$ 时检验。
- d. 将研究结果写成适当的 APA 格式。

3. 一个营销公司进行一项研究以评估顾客对不同价格的咖啡的喜欢程度。选择了 3 美元、6 美元和 10 美元三种咖啡（12 盎司的咖啡粉）的代表性品牌。150 名参与者品尝了三种咖啡（在不知道咖啡价格的情况下），并且说明他们更喜欢哪种咖啡。图表 15—16 描述了 150 名参与者的选择情况。

图表 15—16 选择 3 美元、6 美元和 10 美元咖啡的人数

3 美元咖啡	6 美元咖啡	10 美元咖啡
30	62	58

数据在网站上的文件夹 Chapter15 中的文件 Chapter 15 _ Exercise3. sav 之中（变量命名为 **cost** 和 **frequency**，其中对于 **cost**，3 美元赋值为 1，6 美元赋值为 2，10 美元赋值为 3）。在 SPSS 中打开文件，进行必要分析以回答下列问题。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 有一种或多种咖啡受到偏爱吗？在 $\alpha=0.05$ 时检验。
- d. 将研究结果写成适当的 APA 格式。



χ² 独立性检验

χ² 独立性检验对两个分类变量是否有关系提供了检验，其中每个变量包括一个或多个分类。下面给出一个用到 χ² 独立性检验的例子。

例子

一位研究员想调查性格类型（个性内向的人、个性外向的人）和休闲运动的选择（逛游乐园、休息一天）是否有关系。他对 100 名答应参与这项研究的人做了性格测试，并且基于测试的分值把他们分为性格内向的人和性格外向的人，然后要求每个参与者在逛游乐园和休息一天两者之中选择更喜欢的休闲方式。图表 16—1 描述了每个参与者的性格类型和选择的休闲方式。

图表 16—1 100 名参与者的分类表格

		喜欢的休闲方式		
		逛游乐园	休息	合计
性格类型	内向	12	28	40
	外向	43	17	60
合计		55	45	100

在图表 16—1 中，性格类型在表格的行中，而休闲方式的选择显示在表格的列中。在参与研究的 100 人中，有 40 人是性格内向的，60 人是性格外向的。在性格内向的人中，有 12 人更喜欢逛游乐园，28 人更喜欢休息。在性格外向的人中，有 43 人更喜欢逛游乐园，17 人更喜欢休息。这项研究中分类变

量是性格类型和休闲方式的选择，对数据按照总数或频数进行分析。因为性格类型和休闲方式都有两个水平，得到四个单元，所以当前的例子为 2×2 （读为 2 乘 2）卡方表。

χ^2 独立性检验的目标和数据要求

χ^2 独立性检验		
目标	数据要求	例子
确定两个分类变量之间是否有关系	频数或计数 具有两个，包含两个或多个类别的分类变量	频数（每个单元中的人数） 分类变量 ● 性格类型（内向、外向） ● 休闲方式（逛游乐园、休息）

原假设和对立假设

χ^2 独立性检验的原假设和对立假设如下所述：

- H_0 ：性格类型和休闲方式的选择之间没有联系
- H_1 ：性格类型和休闲方式的选择之间存在某种联系

原假设的评价

χ^2 独立性检验对性格类型和休闲方式的选择之间没有联系的原假设进行了检验。如果检验产生的结果在原假设正确时看起来不可能（结果发生的可能性小于 5%），那么拒绝原假设。如果检验产生的结果在原假设正确时看起来正确（结果发生的可能性大于 5%），那么不拒绝原假设。

研究问题

在研究中基本的兴趣问题也可以用研究问题的方式表示，例如，
“性格类型和休闲方式的选择之间存在联系吗？”

在 SPSS 中输入数据及分析

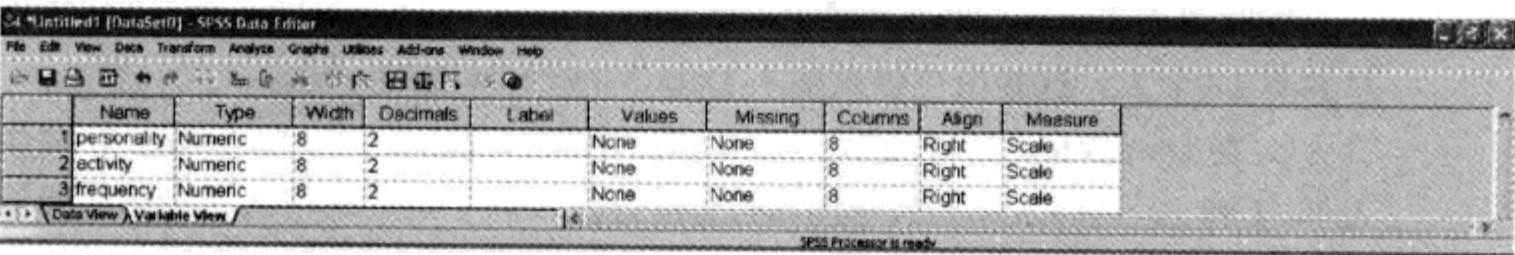
下面的步骤 1 和步骤 2 描述了怎样在 SPSS 中输入数据。数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 16 中的文件 choice.sav 之中。如果你更喜欢从网站上打开文件，跳到步骤 3。

步骤 1：生成变量

- 1. 打开 SPSS。
- 2. 点击 Variable View 标签。

在 SPSS 中将生成三个变量，一个是不同的性格类型，一个是休闲方式，一个是频数。这三个变量分别命名为 **personality**，**activity** 和 **frequency**。

- 3. 在 Variable View 窗口前三行分别输入变量名称 **personality**，**activity** 和 **frequency**（详见图表 16—2）。



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	personality	Numeric	8	2		None	None	8	Right	Scale
2	activity	Numeric	8	2		None	None	8	Right	Scale
3	frequency	Numeric	8	2		None	None	8	Right	Scale

图表 16—2 输入变量 **personality**，**activity** 和 **frequency** 的 Variable View 窗口

- 4. 使用第 1 章描述的过程，为分类变量 **personality** 和 **activity** 建立变量值标签，对于 **personality**，1 = “内向”，2 = “外向”。对于 **activity**，1 = “逛游乐园”，2 = “休息”。

步骤 2：输入数据

接下来，我们在 SPSS 中输入数据。 χ^2 独立性检验有两种不同的数据输入方法：加权方法和个体观测值方法（见第 15 章对这两种方法的描述）。如第 15 章所述，当数据在每个单元的频数统计出来时，应采用加权方法。由于在我们的例子中，单元中的频数已经被统计出来（如图表 16—1），我们将采用加权方法来输入数据（图表 16—4 描述了如何用个体观测值方法来创建数据文件）。

在我们的例子中，内向性格和外向性格的人可以选择逛游乐园和休息中

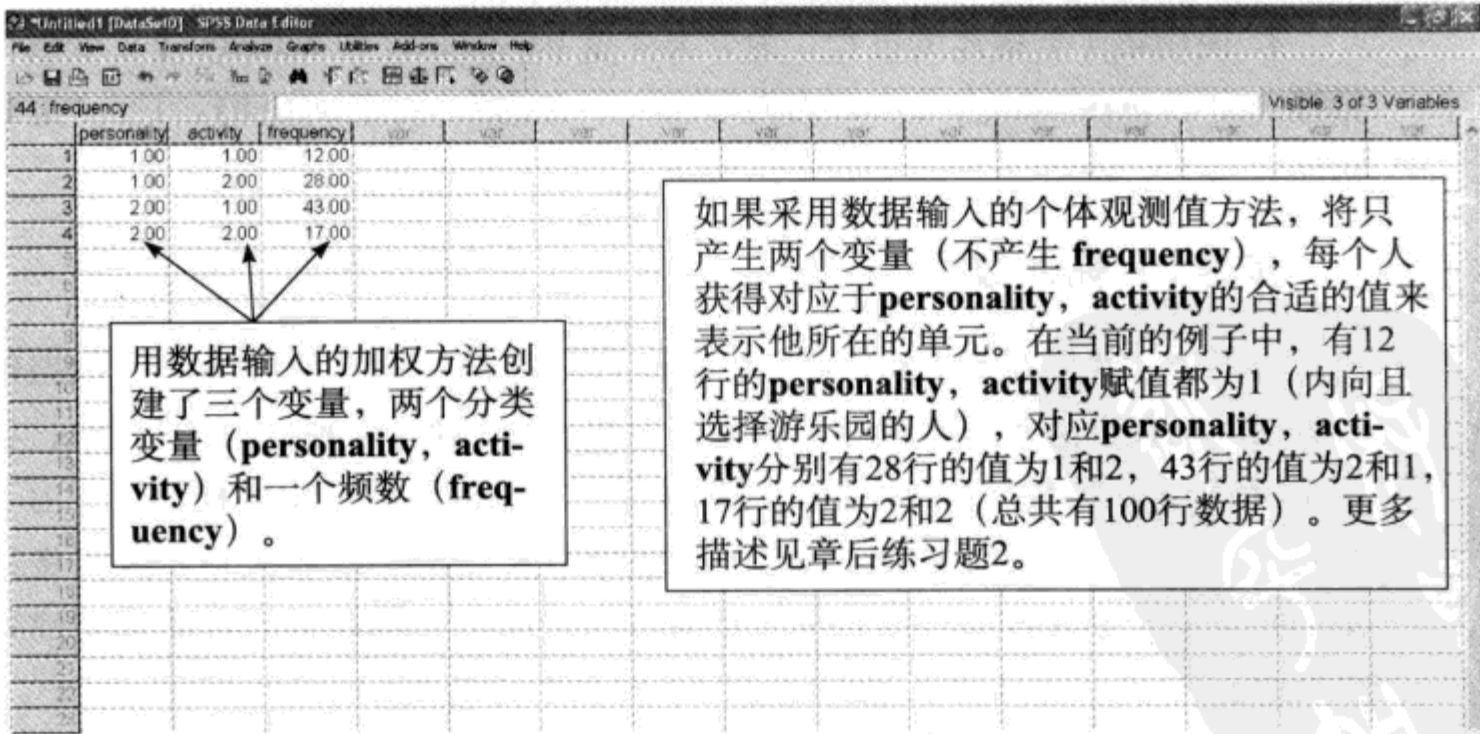
的一个，于是产生了四种不同情况（内向/逛游乐园、内向/休息、外向/逛游乐园、外向/休息）。由于我们采用加权方法来输入数据，我们需要在 Data View 窗口为这四种情况的每一种创建单独的一行。用加权方法建立的数据文件结构如图表 16—3 所示。

图表 16—3 研究中设定的四种情况（对于 personality，内向赋值为“1”，外向赋值为“2”。对于 activity，逛游乐园赋值为“1”，休息赋值为“2”）

Personality	Activity	Frequency
1	1	12
1	2	28
2	1	43
2	2	17

输入数据

- 1. 点击 Data View 标签。变量 personality，activity 和 frequency 出现在 Data View 窗口的前三列。
- 按照图表 16—3，第一种情况对应于内向（1）且选择逛游乐园（1）的人总共有 12 个人，这些值应该被输入 Data View 窗口的第一行。
- 2. 在 Data View 窗口的第一行对 personality，activity 和 frequency 分别输入 1，1 和 12，在 Data View 窗口的 2~4 行输入剩下的三种情况（在第 2 行输入 1，2 和 28，在第 3 行输入 2，1 和 43，在第 4 行输入 2，2 和 17）。图表 16—4 中给出了完整的数据文件。



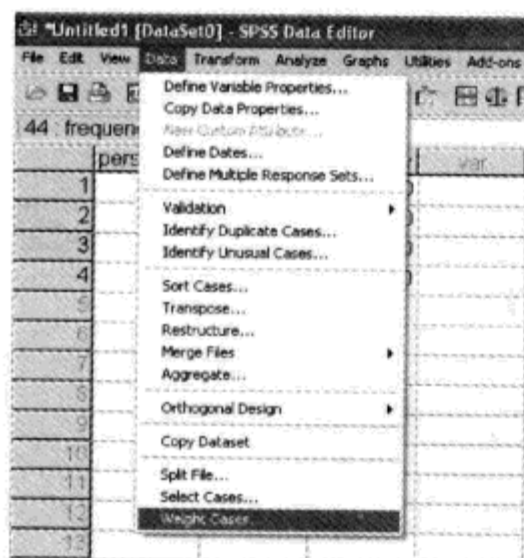
图表 16—4 χ^2 独立性检验例子的完整数据文件

步骤 3: 分析数据

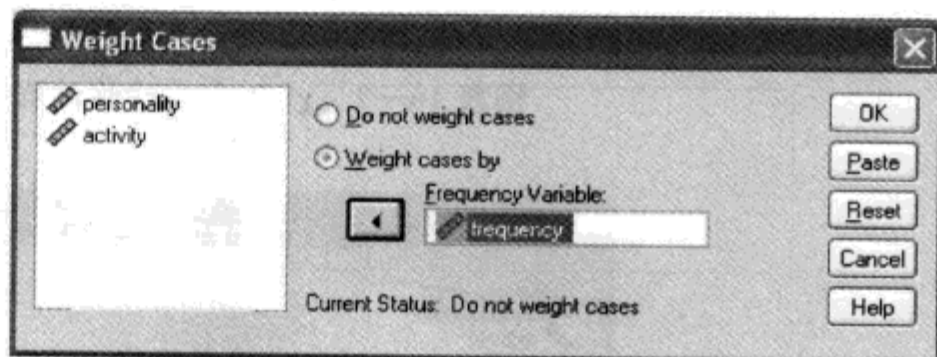
在执行 χ^2 检验之前, 我们首先需要对 **frequency** 进行加权。加权表明给定变量的值表示观测总次数, 而不仅仅是一个分数值。例如, 对 **frequency** 进行加权时, **frequency** 取值为 12 代表 12 个人, 而不是分数为 12。

对 frequency 进行加权

1. 在菜单栏中选择 **Data > Weight Cases...** (见图表 16—5)。
2. 打开 Weight Cases 对话框。选择 Weight cases by 并选择变量 **frequency**。点击向右箭头按钮 (►), 把 **frequency** 移到 Frequency Variable (加权变量) 框中 (如图表 16—6)。
3. 点击 “OK”。这表示在每个类别中频数的取值 (12, 28, 43 和 17) 对应于每个单元的所有参与者, 而不仅仅是一个分数。¹



图表 16—5 在 SPSS 中案例加权过程的菜单操作



图表 16—6 加权的对话框

通过对 **frequency** 进行加权, 现在我们可以执行 χ^2 独立性检验。

执行 χ^2 独立性检验

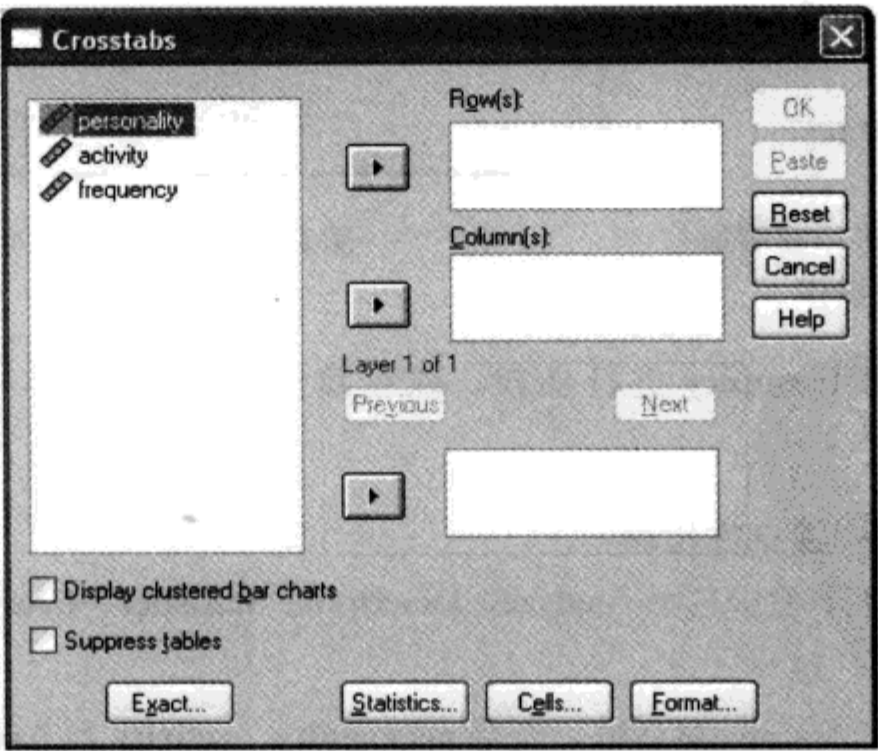
1. 在菜单栏中选择 **Analyze > Descriptive Statistics > Crosstabs** (交叉表) ... (见图表 16—7)。

打开 Crosstabs 对话框, 变量 **personality**, **activity** 和 **frequency** 出现在对话框的左侧 (见图表 16—8)。

- 2. 选择 **personality**，点击向右箭头按钮 (▶)，把变量移到 Row(s) 框中。
- 3. 选择 **activity**，点击向右箭头按钮 (▶)，把变量移到 Column(s) 框中，详见图表 16—9 (注：**frequency** 仍然在 Crosstabs 对话框的左侧)。

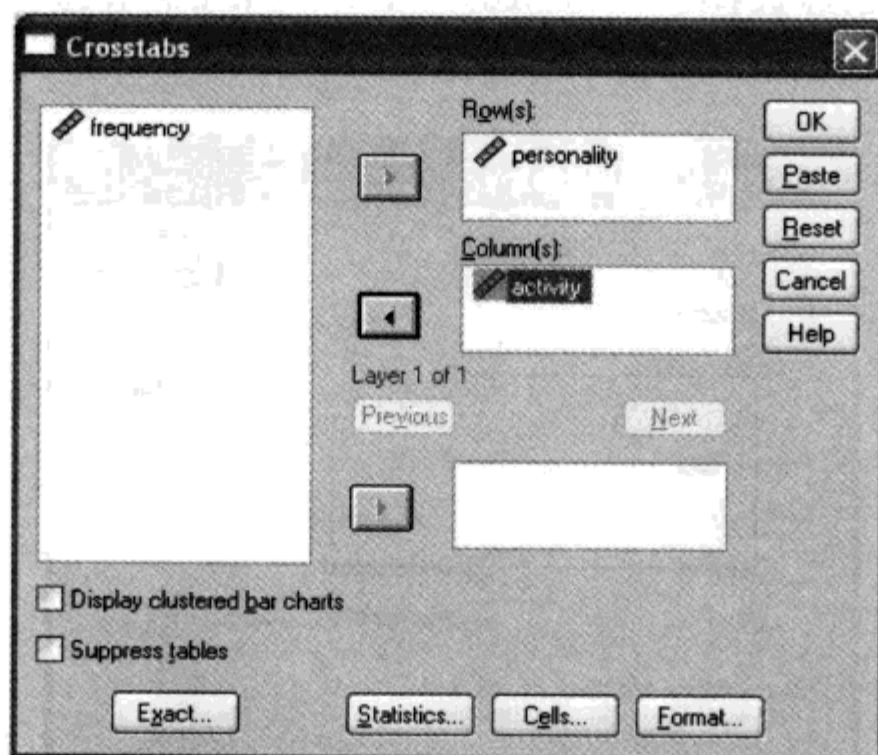


图表 16—7 χ^2 独立性检验过程的菜单操作

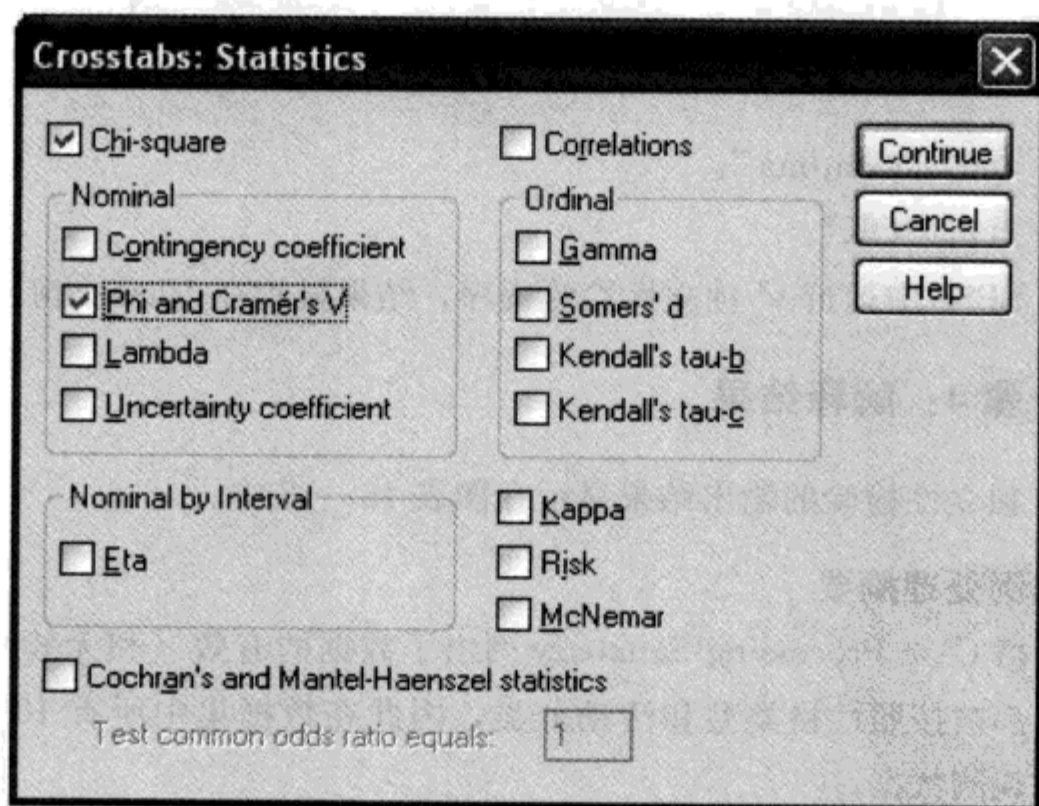


图表 16—8 Crosstabs (交叉表) 对话框

4. 点击 “Statistics”。打开 Crosstabs: Statistics 对话框, 选择 Chi-square 和 Phi and Cramer's V (见图表 16—10)。



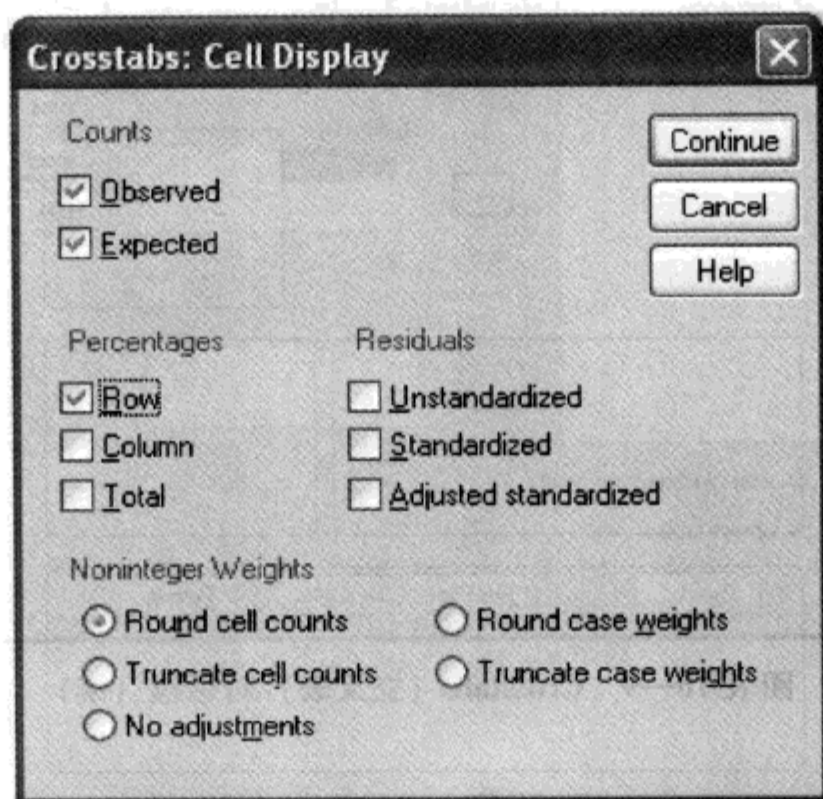
图表 16—9 Crosstabs (交叉表) 对话框 (续)



图表 16—10 Crosstabs: Statistics 对话框

5. 点击“Continue”。

6. 点击“Cells”。打开 Crosstabs: Cell Display 对话框。在 Counts 下选择 Expected (已经选择了 Observed)，在 Percentages 下选择 Row。详见图表 16—11。



图表 16—11 Crosstabs: Cell Display (单元列表) 对话框

7. 点击“Continue”。

8. 点击“OK”。

在 SPSS 中运行 χ^2 独立性检验程序，结果显示在 Viewer 窗口中。

步骤 4：解释结果

χ^2 独立性检验的输出结果显示在图表 16—12 中。

案例处理摘要

表格 Case Processing Summary 列出了数据的有效（和无效）案例。由于每个人都被按照性格类型和活动分类，因此在数据集中所有 100 个参与者都有有效观测数据。

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
personality * activity	100	100.0%	0	0.0%	100	100.0%

personality * activity Crosstabulation

		activity		Total	
		amusement park	retreat		
Personality	introvert	Count	12	28	40
		Expected Count	22.0	18.0	40.0
		%within personality	30.0%	70.0%	100.0%
	extrovert	Count	43	17	60
		Expected Count	33.0	27.0	60.0
		%within personality	71.7%	28.3%	100.0%
Total	Count	55	45	100	
	Expected Count	55.0	45.0	100.0	
	%within personality	55.0%	45.0%	100.0%	

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	16.835 ^b	1	0.000		
Continuity Correction ^a	15.194	1	0.000		
Likelihood Ratio	17.230	1	0.000		
Fisher's Exact Test				0.000	0.000
Linear-by-Linear Association	16.667	1	0.000		
N of Valid Cases	100				

^a Computed only for a 2x2 table.
^b 0 cells (0.0%) have expected count less than 5. The minimum expected count is 18.00.

Symmetric Measures

		Value	Approx. Sig.
Nominal by	Phi	-.410	0.000
Nominal	Cramer's V	0.410	0.000
N of Valid Cases		100	

^a Not assuming the null hypothesis.
^b Using the asymptotic standard error assuming the null hypothesis.

图表 16—12 χ^2 独立性检验的输出结果

性格×活动交叉表 (Personality * activity Crosstabulation)

表格 Personality * activity Crosstabulation 为观测频数和期望频数提供了重要信息。表格中标为“Count”（总数）的值包括我们输入 SPSS 的原始观测频数（12、28、43 和 17），标为“Expected Count”（期望总数）的值是原假设为真的条件下（包含抽样误差）的期望频数。如果 χ^2 检验是显著的，则观测频数和期望频数之差就可以确定两个变量的本质关系。

卡方检验 (Chi-Square Tests)

表格 Chi-Square Tests 提供了一些不同检验的结果，其中最常用的是 Pearson Chi-Square（皮尔森卡方）。在第 15 章就已给出的皮尔逊卡方统计量 (Pearson Chi-Square Statistic) 是

$$\chi^2 = \sum \frac{(\text{观测频数} - \text{期望频数})^2}{\text{期望频数}}$$

其中，观测频数和期望频数分别对应于表格 personality * activity Crosstabulation 中的 Count 和 Expected Count。把表格中的观测频数和期望频数代入公式得到 χ^2 值为：

$$\begin{aligned}\chi^2 &= \frac{(12-22)^2}{22} + \frac{(28-18)^2}{18} + \frac{(43-33)^2}{33} + \frac{(17-27)^2}{27} \\ \chi^2 &= 16.835\end{aligned}$$

这个值与 Chi-Square Tests 表格中的 Pearson Chi-Square 值是相等的。卡方检验的自由度为 1 ($df = (\text{性格类型组别的数目} - 1) \times (\text{休闲方式组别的数目} - 1)$)，其 p -值 (Asymp. Sig.) 为 0.000（读为小于 0.001）。因为 p -值 < 0.05 ，所以拒绝原假设，这说明性格类型和休闲方式的选择有关系。

由于结果是显著的，我们将检验观测频数和期望频数（总数）在表格 personality * activity Crosstabulation 中的差异，以确定这两个变量的本质关系。检查表格中的第一个单元（性格内向且选择逛游乐园的人），我们期望有 22 个人选择逛游乐园，然而实际只有 12 个人选择，这说明内向的人选择逛游乐园的概率比预期的小。再看休息这一列，28 个内向的人选择了休息，而我们预期的只有 18 个人，这说明内向的人选择休息的概率比预期的多。对于外向的人，出现了相反的情况：外向的人选择逛游乐园的比预期的概率大（观测值是 43 而预期值是 33），而选择休息的概率比预期的小（观测值为 17 而预期值为 27）。图表 16—13 对此作了总结。

图表 16—13 χ^2 独立性检验的结果摘要

	逛游乐园	休息	结论
性格内向	选择的人数少于期望数（观测到 12 人选择，期望 22 人选择）	选择的人数多于期望数（观测到 28 人选择，期望 18 人选择）	性格内向的人更喜欢休息
性格外向	选择的人数多于期望数（观测到 43 人选择，期望 33 人选择）	选择的人数少于期望数（观测到 17 人选择，期望 27 人选择）	性格外向的人更喜欢逛游乐园

总结结果，性格内向的人更喜欢休息，而性格外向的人更喜欢逛游乐园。

对称性度量 (Symmetric Measures)

我们将在下面的效应量中讨论对称性度量表格。

效应量

常用 Cramér's V 对 χ^2 独立性检验的效应量进行测量，其公式是

$$V = \sqrt{\frac{\chi^2}{N(k-1)}}$$

其中 χ^2 是表格 Chi-Square Tests 中 Pearson Chi-Square 的值，N 是研究中的样本量，k 是类别最少的变量的水平数。

代入图表 16—12 中相应的值，得到 V 值为

$$V = \sqrt{\frac{16.835}{100(2-1)}} = 0.41$$

这与表格 Symmetric Measures 中的 Cramér's V 值相等。

科恩 (Cohen, 1998) 约定小、中和大的效应量分别对应着 Cramér's V 值为 0.10、0.30 和 0.50 (这些效应量仅仅描述至少有一个变量具有两个类别的表格，即 2×2、2×3、2×4 的表格，等等)。因此，值 0.41 被认为在实际中有中等的效果，这表示两个变量之间有中度相关关系。

APA 格式的结果表达

在写结果时，要报告原假设的结论、自由度、样本量、 χ^2 值、p 值，以

及效应量。对显著性检验结果还要描述预测频数和期望频数之间的关系。下面给出一个用 APA 格式做的简要描述的例子。

性格类型与休闲方式的选择有显著相关性, $\chi^2(1, N=100) = 16.84$, $p < 0.05$, Cramer's $V = 0.41$ 。对于选择休息还是逛游乐园, 性格内向的人更喜欢休息 (70% 性格内向的人选择休息), 而性格外向的人更喜欢逛游乐园 (72% 性格外向的人选择逛游乐园)。

χ^2 独立性检验的假定

1. 观测值是独立的。

独立性假定意味着每个单元的观测值 (即参与者) 是彼此独立的。对一个人记录不同的两个单元 (例如, 允许一个人同时选择两种活动) 是违反这个假定的一个例子。违反独立性假定会严重影响卡方检验的准确性。如果我们有理由相信已经违反了独立性假定, 那么就不应该再用 χ^2 独立性检验。

2. 期望单元频数的规模。

对于小的表格 (有 4 个单元或更少), 建议每个单元的频数为 5 或更多。对于大的表格, 有一个或几个单元的期望频数小于 5, 则很可能不会使得 χ^2 独立性检验无效, 在样本量很小时尽管效果会有损失。

在 SPSS 中执行 χ^2 独立性检验的步骤摘要

已经给出了在 SPSS 中用数据输入的加权方法和个体观测值方法执行一个 χ^2 独立性检验的说明 (前面的注释已提到, 用个体观测值方法创建数据文件的说明在本章末的练习题 2 中)。

Ia. 数据输入和分析——加权方法

1. 在 SPSS 中生成三个变量 (两个分类变量, 一个频数变量)。
2. 为分类变量建立变量值标签。在 Value Labels 对话框中输入适当的数值。点击 “OK”。
3. 输入数据 (对于数据输入的加权方法, 数据文件中的行数等于类别变量数目)。

4. 通过选择 **Data>Weight Cases...**，对频数变量进行加权处理。
5. 选择 **Weight cases by**，把频数变量移到 **Frequency Variable** 框中。点击“OK”。
6. 选择 **Analyze>Descriptive Statistics>Crosstabs...**。
7. 把类别变量移到 **Row (s)** 框中，其他变量移到 **Column (s)** 框中（频数变量仍然在对话框的左侧）。
8. 点击“Statistics”。选择 **Chi-square** 和 **Phi and Cramer's V**。点击“Continue”。
9. 点击“Cells”。在 **Counts** 下选择 **Expected**（已经选择了 **Observed**），在 **Percentages** 下选择 **Row**。点击“Continue”。
10. 点击“OK”。

Ib. 数据输入和分析——一个体观测值方法

1. 在 SPSS 中生成两个变量（每个类别中一个）。
2. 为变量建立变量值标签。在 **Value Labels** 对话框中输入适当的数值和标签。点击“OK”。
3. 输入数据（对于数据输入的个体观测值方法，数据文件的行数等于研究中参与者的数量）。
4. 选择 **Analyze>Descriptive Statistics>Crosstabs...**。
5. 把一个分类变量移到 **Row (s)** 框中，其他变量移到 **Cloumn (s)** 框中。
6. 点击“Statistics”。选择 **Chi-square** 和 **Phi and Cramer's V**。点击“Continue”。
7. 点击“Cells”。在 **Counts** 下选择 **Expected**（已经选择了 **Observed**），在 **Percentages** 下选择 **Row**。点击“Continue”。
8. 点击“OK”。

II. 结果描述（不管采用哪种数据输入方法，解释是一样的）

在表格 **Chi-Square Tests** 中，检验 **Pearson Chi-Square** 的 p -值（以 **Asymp. Sig.** 列出）。

- 如果 $p \leq 0.05$ ，拒绝原假设。写出表示两个变量之间本质关系的结果。

- 如果 $p > 0.05$ ，不拒绝原假设。写出表示两个变量之间没有关联的结果。

练习

1. 一位研究员调查一个婴儿吸收的营养种类（按照是否为母乳喂养（breastfed））与他一年级时的体重（按照是否超重（overweight）分）是否有关系，他确定了 300 名一年级学生的喂养方式的育儿报告和身体质量指数（BMI）。基于 BMI 取值，一年级学生们被按是否超重分类，图表 16—14 给出了这组数据。

图表 16—14 母乳喂养研究中 300 名儿童的数据

		Overweight		Total
		Yes	No	
Breastfed	Yes	16	84	100
	No	56	144	200
Total		72	228	300

把数据输入 SPSS，进行必要分析以回答下列问题。变量命名为 **breastfed**，**overweight** 和 **frequency**（确保在分析之前对数据进行加权）。对于 **breastfed** 和 **overweight**，是赋值为“1”，否赋值为“2”。

- a. 陈述原假设和对立假设。
- b. 陈述数据的一个研究问题。
- c. 一个婴儿的吸收营养方式和他在一年级时的体重有显著性关系吗？在 $\alpha = 0.05$ 时检验。
- d. 效应量是多少？你能区分效应量是小、中还是大吗？
- e. 将研究结果写成适当的 APA 格式。

2. 一个学生想检验性别与电影爱好之间是否有关联。他询问了 160 名进入录像带店的人（95 名女性和 65 名男性）是喜欢动作电影还是喜剧电影（假设他们只能在这两种类型中选择）。图表 16—15 给出了参与者的选择。

图表 16—15 160 名参与者的性别和电影爱好

		Film		Total
		Action	Drama	
Gender	Females	15	80	95
	males	35	30	65
Total		50	110	160

数据在网站 www.prenhall.com/yockey 上的文件夹 Chapter 16 中的文件 Chapter16 _ Exercise 2. sav 之中，数据文件已经用个体观测值方法输入了。在文件中，变量命名为 **gender**（其中女性赋值为“1”，男性赋值为“2”）和 **film**（其中动作电影赋值为“1”，喜剧电影赋值为“2”）。在 SPSS 中打开文件，进行必要分析以回答下列问题（在练习之前的“在 SPSS 中执行 χ^2 独立性检验的步骤摘要”一节中有数据分析的描述）。

- 陈述原假设和对立假设。
- 陈述数据的一个研究问题。
- 性别和电影爱好之间有显著性关系吗？在 $\alpha=0.05$ 时检验。
- 效应量是多少？你能区分效应量是小、中还是大吗？
- 将研究结果写成适当的 APA 格式。

3. 一个医药研究员想调查老年人运动（是/否）和患心脏病（是/否）是否有关系。他提问了 405 位老年人的运动爱好和是否曾经患过心脏病，图表 16—16 给出了研究参与者的回答。

图表 16—16 运动调查中 405 位参与者的数据

		Heart attack		Total
		Yes	No	
Exercise	Yes	10	140	150
	No	35	220	255
Total		45	360	405

数据在网站文件夹 Chapter 16 中的文件 Chapter16 _ Exercise3. sav 之中（变量命名为 **exercise**，**heartattack** 和 **frequency**，其中对于 **exercise** 和 **heartattack**，1 表示是，2 表示否）。在 SPSS 中打开文件，进行必要分析以回答下列问题（在这个问题中，分析数据前数据加权）。

- 陈述原假设和对立假设。
- 陈述数据的一个研究问题。

- c. 运动和是否曾患心脏病之间有显著性关系吗？在 $\alpha=0.05$ 时检验。
- d. 效应量是多少？你能区分效应量是小、中还是大吗？
- e. 将研究结果写成适当的 APA 格式。



数据变换及其他方法

在数据分析中，常常需要预先对一个或多个变量执行特定的操作。例如，一个常用的必要操作是在 SPSS 中将一些变量相加来得到总分。本附录会举例说明 SPSS 中一些常用的方法，包括 Recode（重新编码）、Compute（计算）、Select Cases（选择案例）以及 Split File（拆分文件）方法。

为了举例说明每一种方法，我们使用图表 A—1 中给出的数据（数据文件可以马上从你的电脑里打开）。这些数据包括了 10 个参与者在 4 个变量上的取值，这 4 个变量分别为：**gender**（1 = “男性”，2 = “女性”），**meaning1**，**meaning2** 和 **meaning3**。后面三个变量是度量生活意义的尺度。

图表 A—1

样本数据

Participant	Gender	Meaning1	Meaning2	Meaning3
1	1	4	2	5
2	1	1	4	2
3	1	5	1	5
4	1	3	3	3
5	1	4	2	4
6	2	4	1	5
7	2	5	2	4
8	2	5	1	5
9	2	1	4	1
10	2	5	1	5

说明：变量 participant 包含在数据中，但不用输入 SPSS。

图表 A—2 中给出了度量生活意义的尺度。

图表 A—2 生活意义尺度表

项目	生活意义尺度	强烈反对	反对	既不反对也不同意	同意	非常同意
meaning1	我对生活方向感觉很好	1	2	3	4	5
meaning2	我的生活没有意义	1	2	3	4	5
meaning3	一般而言，我感觉生活步入正轨	1	2	3	4	5

说明：对于每一项有：1（强烈反对）、2（反对）、3（既不反对也不同意）、4（同意）及 5（非常同意）。规定每位参与者对于每个问题只能选择一个回应选项。

重新编码方法

对于图表 A—2 中生活意义尺度的项目，参与者每读一条陈述，就选择一种反应——从非常同意到强烈反对——最能表达出对于这条陈述的感觉。通读这些项，注意到对于 meaning1 和 meaning3 项，那些认为生活很有意义的人最可能回答“非常同意”（选择“5”），而对于 meaning2，这样的人最可能回答“强烈反对”（选择“1”）。meaning2 是一个负向题目的例子，一个认为生活很有意义的人对此会做出相反的回答。因此，在这个例子中，一个认为生活很有意义的人最可能对 meaning1、meaning2 及 meaning3 分别选择 5、1、5。通常要分析这些反应，需要将这些项的反应值相加起来计算总分。在这个例子中，一个认为生活很有意义的人，得分是 11（5+1+5）。然而，因为最终的设计是要使得较高分代表较高水平的生活意义，那么一个认为生活很有意义的人应该得 15 分（5+5+5）。因为 meaning2 是负向题目，所以不能满足这个设计要求。为了纠正它，给出了一种叫逆向编码的方法，使得对负向题目（meaning2）在程度上的反应逐个颠倒，使得 1 变为 5、2 变为 4、3 变为 3、4 变为 2、5 变为 1。通过 meaning2 的逆向编码，对于一个认为生活很有意义的人，总分不再是 11（5+1+5），而是得到总分 15（5+5+5），与设计的目的达到一致。通过提供合理的逆向编码，我们在 SPSS 中对 meaning2 进行重新编码。

首先我们打开数据文件（注：以下的说明是假设 SPSS Demystified 文件

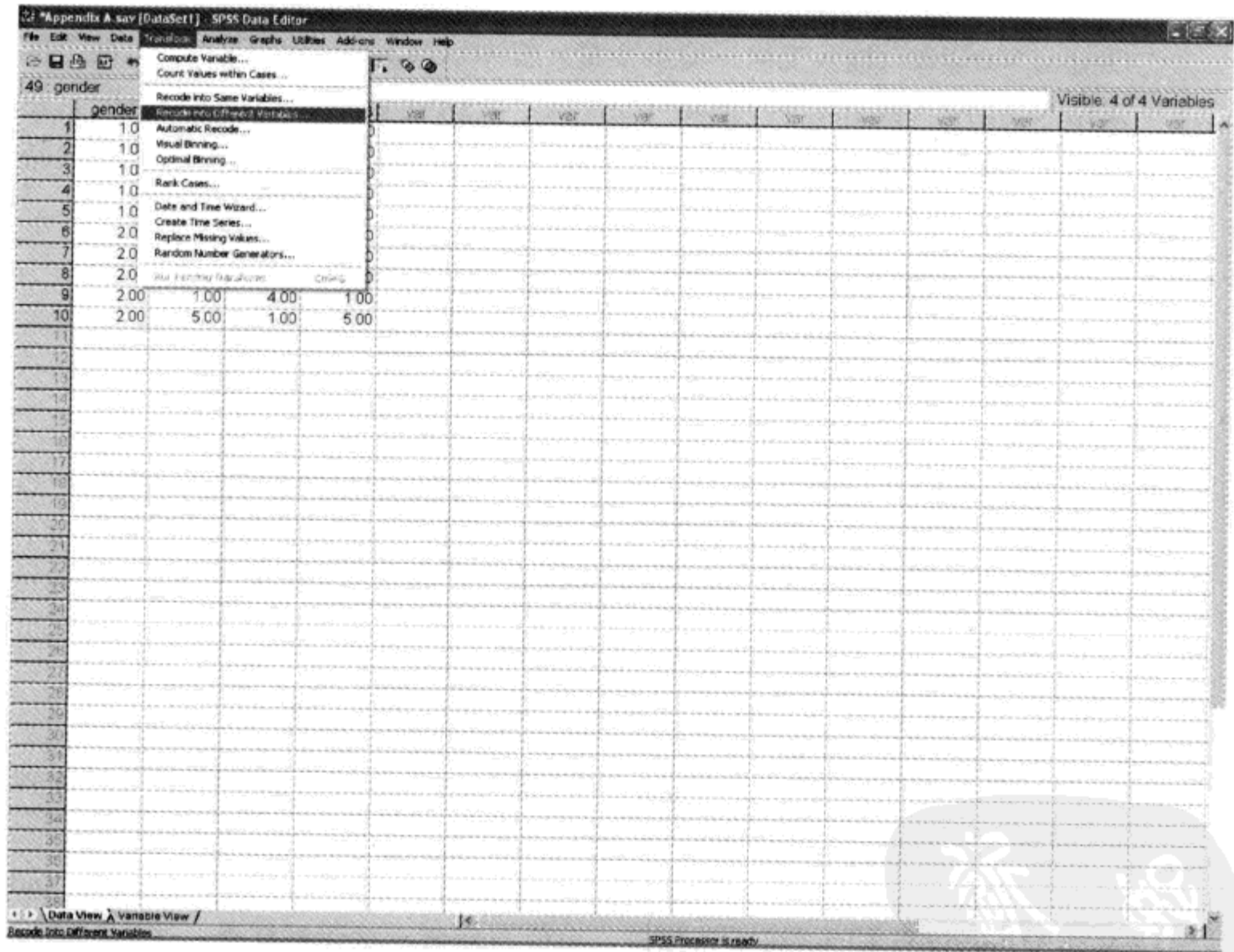
已经从网上下载到你的电脑里。如果你的电脑里没有该文件，可以从网上下载，网址是 www.prenhall.com/yockey）。

从你的电脑里打开文件：

查找你电脑里 SPSS Demystified 文件的位置。在 Appendix A 文件夹中，打开文件 Appendix A. sav。

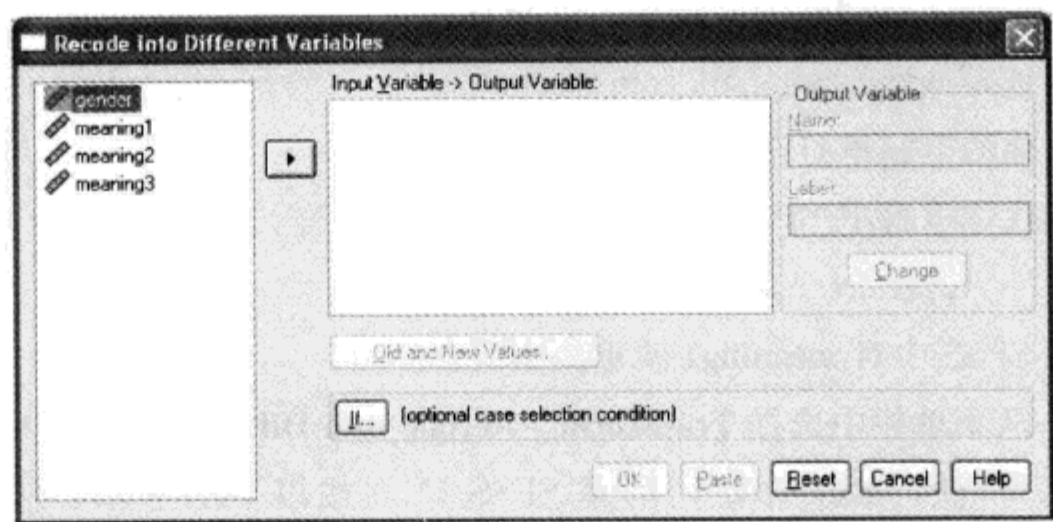
在 SPSS 中对 **meaning2** 项重新编码：

1. 从菜单栏中选择 **Transform>Recode into Different Variables...**（见图表 A—3）。（注：如果你使用的是 SPSS 14.0 或以下的版本，从菜单栏选择 **Transform>Recode>into different variables**。所有其他的命令是相同的。）



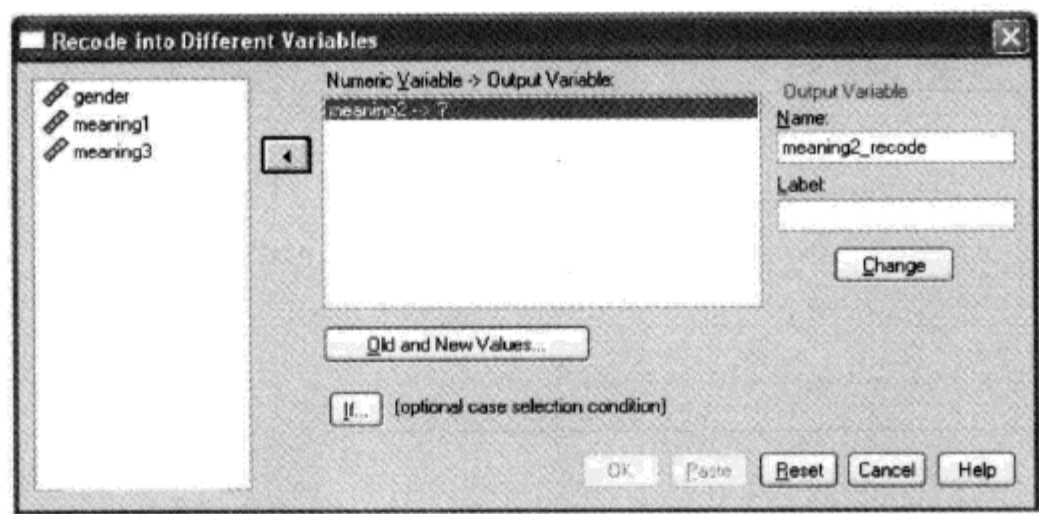
图表 A—3 Recode into Different Variables 方法的菜单命令

2. Recode into Different Variables（为不同变量重新编码）对话框打开（见图表 A—4）。
3. 选中 **meaning2**，点击向右箭头按钮（▶），将它移到 Input Variable→



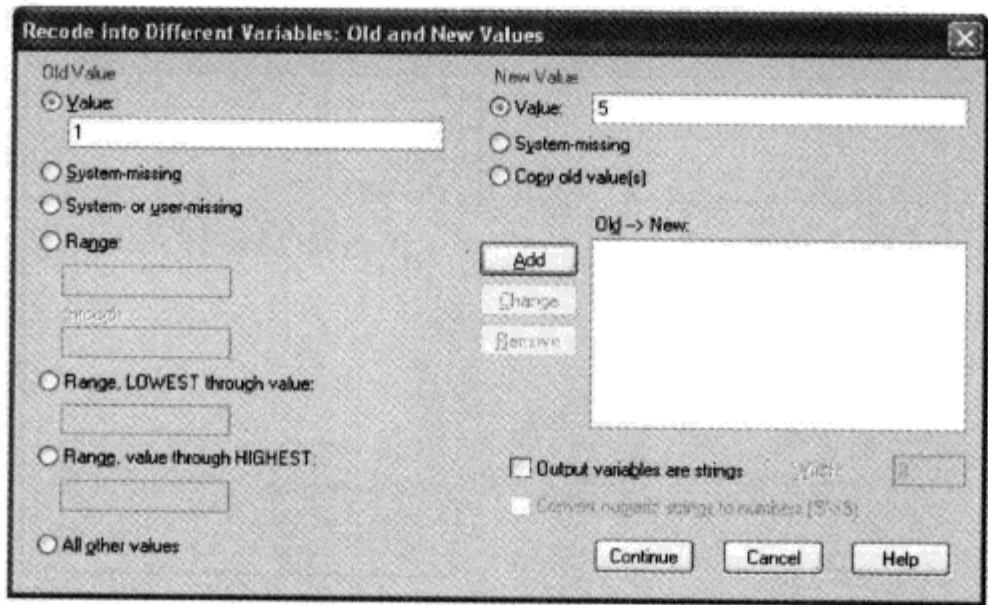
图表 A—4 Recode into Different Variables 对话框

Output Variable（输入变量→输出变量）框中（当 **meaning2** 移入时，这个框读作 Numeric Variable→Output Variable（数值变量→输出变量）框）。在 Name 下输入 **meaning2 _ recode**（见图表 A—5）。这就是包含相反编码值的新变量的名字。

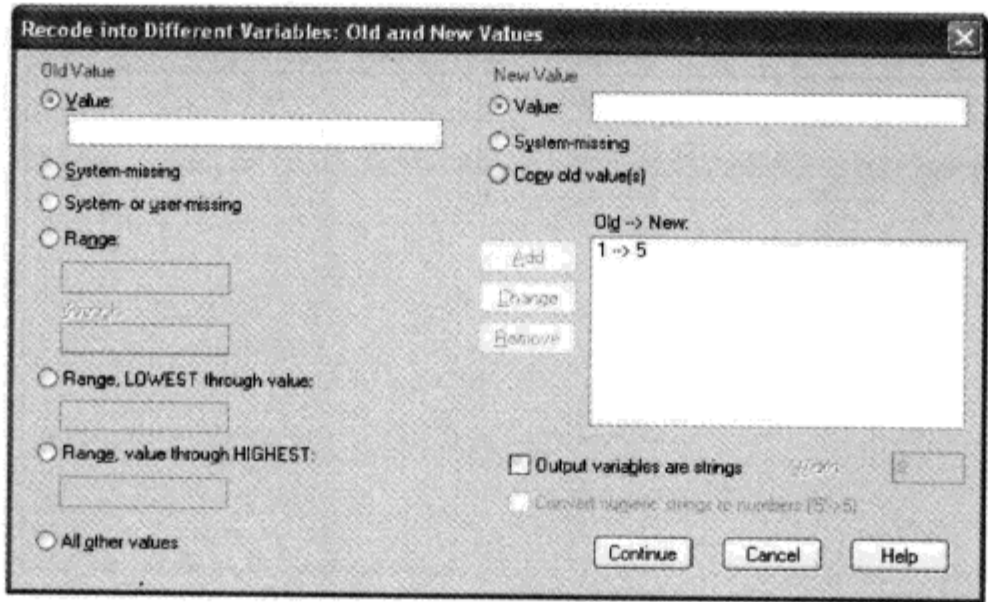


图表 A—5 Recode into Different Variables 对话框（续）

- 4. 点击“Change”。在 Numeric Variable→Output Variable 框中将看到 meaning2→meaning2 _ recode。
- 5. 点击“Old and New Values”，Recode into Different Variables: Old and New Values（为不同变量重新编码：旧值与新值）对话框打开。这个对话框用来重新编码 **meaning2** 的值，使得 1 变为 5，2 变为 4，依次类推。
- 6. 在 Old Value（旧值）下 Value 文本框中输入 1。在 New Value（新值）下 Value 文本框中输入 5（见图表 A—6）。
- 7. 点击“Add”。在 Old→New 框中会显示 1→5（见图表 A—7）。



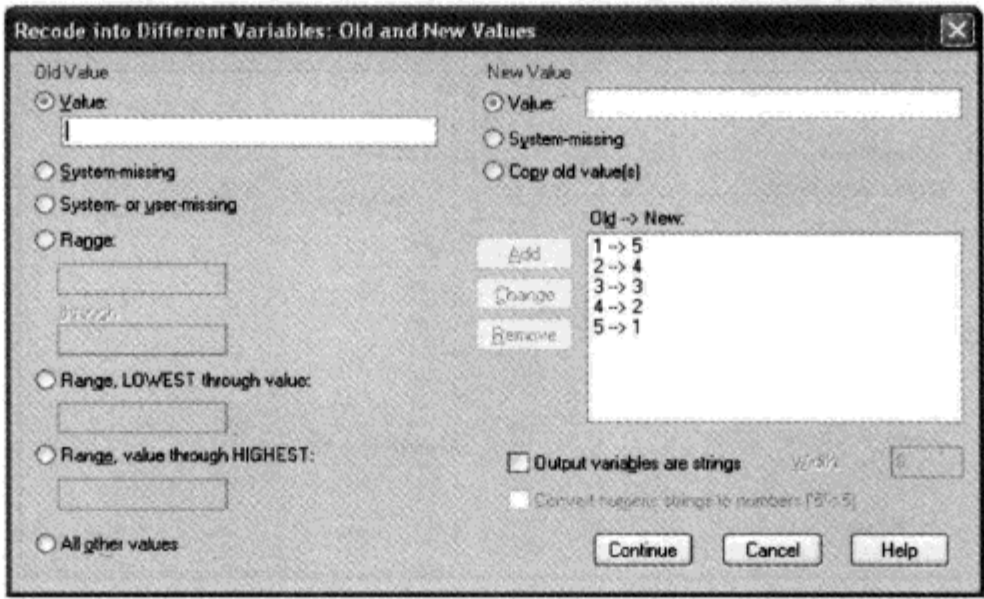
图表 A—6 Recode into Different Variables: Old and New Values 对话框



图表 A—7 Recode into Different Variables: Old and New Values 对话框 (续)

8. 输入剩余的值进行重新编码 (2→4, 3→3, 4→2, 5→1), 点击 “Add” 输入 Old→New 框。当所有的 5 对值都输入进去后, Old→New 框应包含如图表 A—8 所示的重新编码的值。
9. 点击 “Continue”。
10. 点击 “OK”。

在 SPSS 的 Data View 窗口, 你可以在数据文件中看到新的变量 `meaning2_recode`, 它包含了与 `meaning2` 相反编码的项 (见图表 A—9)。



图表 A—8 Recode into Different Variables: Old and New Values 对话框（续）

SPSS Data Editor window showing the Data View for a dataset named "Appendix A.sav". The table has 5 columns: gender, meaning1, meaning2, meaning3, and meaning2_recode. The first 10 rows of data are shown.

	gender	meaning1	meaning2	meaning3	meaning2_recode
1	1.00	4.00	2.00	5.00	4.00
2	1.00	1.00	4.00	2.00	2.00
3	1.00	5.00	1.00	5.00	5.00
4	1.00	3.00	3.00	3.00	3.00
5	1.00	4.00	2.00	4.00	4.00
6	2.00	4.00	1.00	5.00	5.00
7	2.00	5.00	2.00	4.00	4.00
8	2.00	5.00	1.00	5.00	5.00
9	2.00	1.00	4.00	1.00	2.00
10	2.00	5.00	1.00	5.00	5.00

图表 A—9 包含重新编码变量 meaning2_recode 的 Data Editor 窗口²

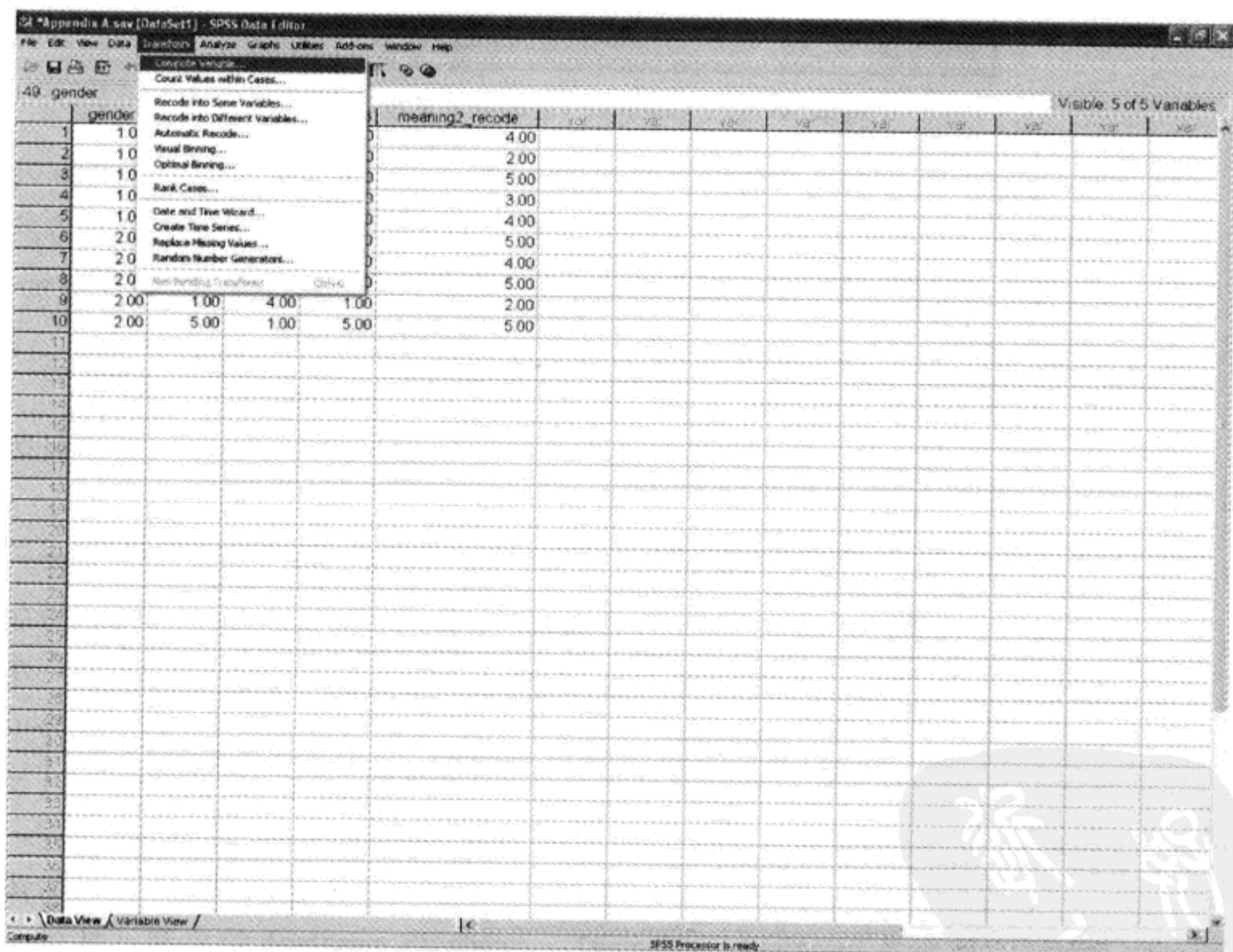
计算方法

正如在本附录开始指出的，SPSS 中一个常用的执行操作是将一些变量相加起来得到总分。在 SPSS 中，我们使用 Compute 方法来计算变量总和。

我们使用 Compute 方法将生活意义尺度的项相加产生一个总分。

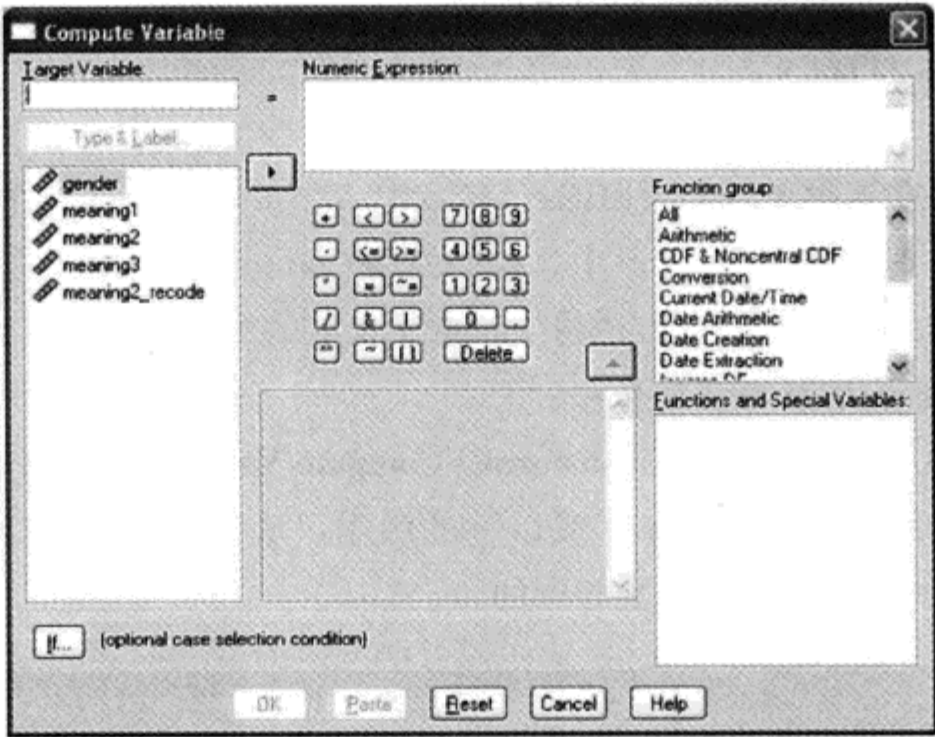
在 SPSS 中计算变量总和：

- 1. 从菜单栏中选择 Transform>Compute Variable（见图表 A—10）。（注：如果你使用的是 SPSS 14.0 或以下的版本，从菜单栏选择 Transform>Compute…。所有其他的命令是相同的。）



图表 A—10 Compute 方法的菜单命令

- 2. Compute Variable（计算变量）对话框打开（见图表 A—11）。



图表 A—11 Compute Variable 对话框

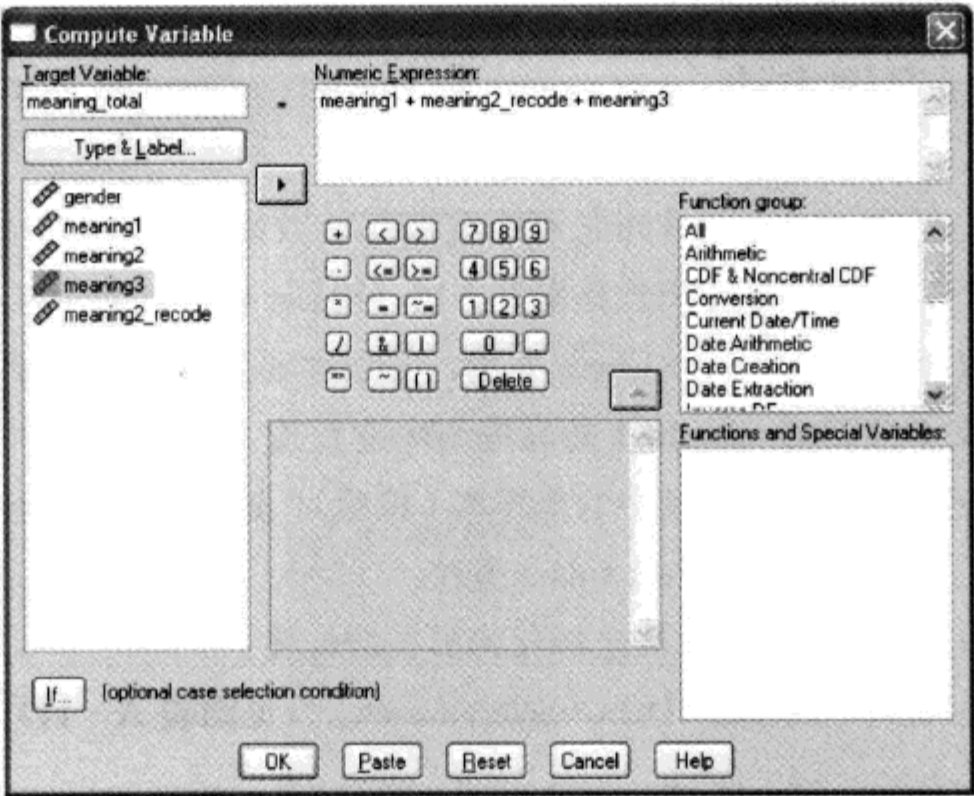
为计算生活意义的总分，我们需要将 **meaning1**，**meaning2 _ recode** 及 **meaning3** 这些项相加。³ 为了在 SPSS 中建立一个与这三个变量的总和相等的新变量，首先需要给出表示总分的这个新变量的名字，我们将它命名为 **meaning _ total**。

1. 在 Target Variable（目标变量）框中，输入名字 **meaning _ total**。

为了得到总分，我们将适当的变量移到 Numeric Expression（数字表达式）框中，并将它们用加号（+）连接。⁴

2. 选中第一个变量 **meaning1**，点击向右箭头按钮（▶）将其移到 Numeric Expression 框中。点击 Numeric Expression 框下面的加号（+）（或在键盘上按 Shift 与 + 键）。再将 **meaning2 _ recode** 移到 Numeric Expression 框中，并点击加号（+）。再将 **meaning3** 移到 Numeric Expression 框中。此时在 Numeric Expression 框中，显示为 **meaning1 + meaning2 _ recode + meaning3**（见图表 A—12）。

3. 点击“OK”。在 Data View 窗口中生成了一个名为 **meaning _ total** 的新变量，它是变量 **meaning1**、**meaning2 _ recode** 及 **meaning3** 的总和（见图表 A—13）。



图表 A—12 Compute Variable 对话框 (续)

SPSS Data Editor window showing the Data View tab. The dataset is named "Appendix A.sav [DataSet1]". The table contains 37 rows and 13 columns. The first 7 columns are: gender, meaning1, meaning2, meaning3, meaning2_recode, meaning_total, and a blank column. The last 6 columns are labeled V1 through V6. The data is as follows:

	gender	meaning1	meaning2	meaning3	meaning2_recode	meaning_total	V1	V2	V3	V4	V5	V6
1	1.00	4.00	2.00	5.00	4.00	13.00						
2	1.00	1.00	4.00	2.00	2.00	5.00						
3	1.00	5.00	1.00	5.00	5.00	15.00						
4	1.00	3.00	3.00	3.00	3.00	9.00						
5	1.00	4.00	2.00	4.00	4.00	12.00						
6	2.00	4.00	1.00	5.00	5.00	14.00						
7	2.00	5.00	2.00	4.00	4.00	13.00						
8	2.00	5.00	1.00	5.00	5.00	15.00						
9	2.00	1.00	4.00	1.00	2.00	4.00						
10	2.00	5.00	1.00	5.00	5.00	15.00						
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												
24												
25												
26												
27												
28												
29												
30												
31												
32												
33												
34												
35												
36												
37												

图表 A—13 包含添加入数据文件的新变量 meaning_total 的 Data Editor 窗口

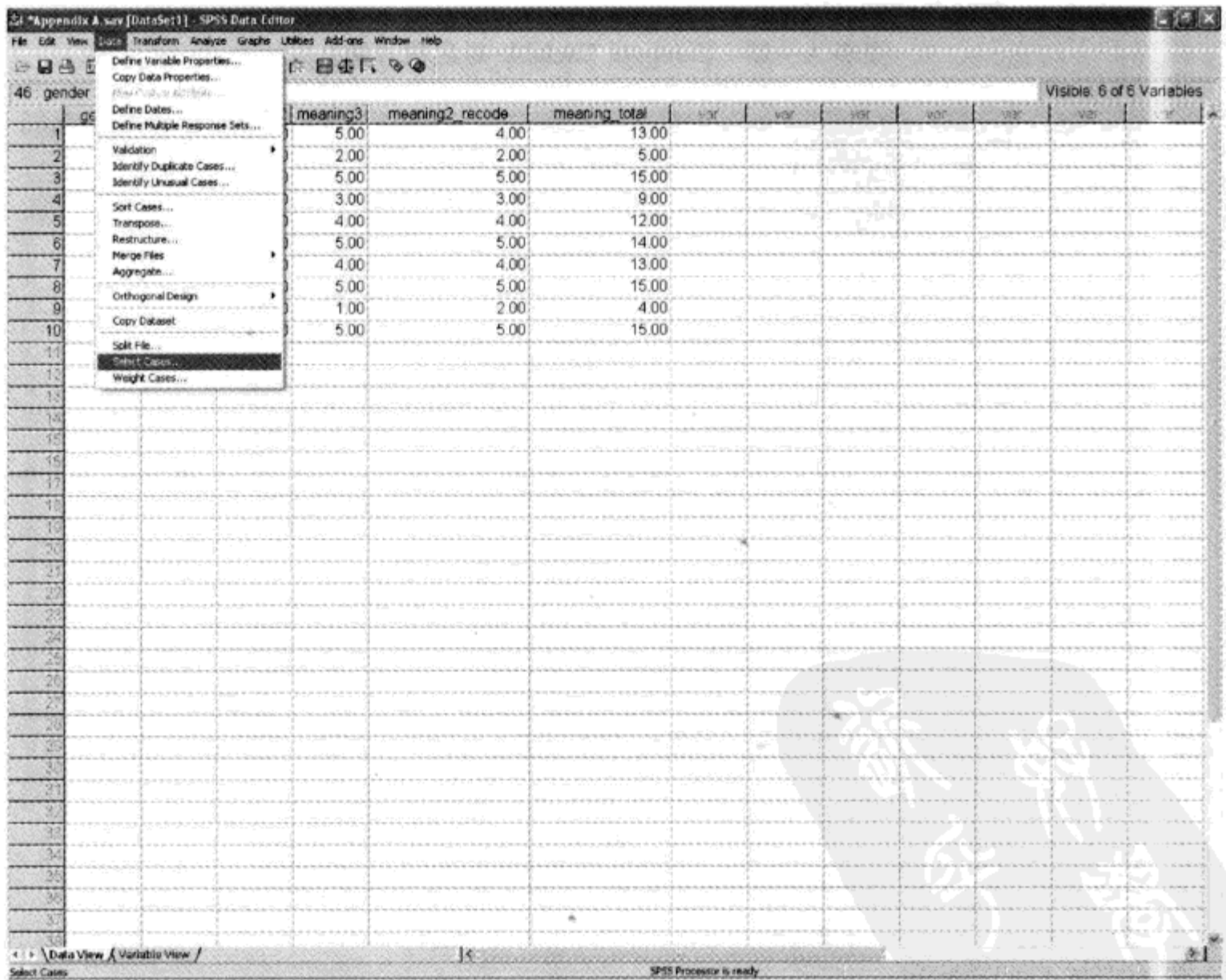
选择案例方法

当分析数据时，某些问题可能要求仅对数据文件的部分执行分析。在这种情况下，我们可以使用选择案例方法（Select Cases）。

为了举例说明 Select Cases 方法，假设你要进行一个关于男性的单样本 t 检验，来看他们对 **meaning1** 的反应与中立反应（对应值为 3）是否有明显的不同（更多信息见第 5 章中的单样本 t 检验），这就要求选择性别仅为男性的数据集，这时要使用 Select Cases 方法。

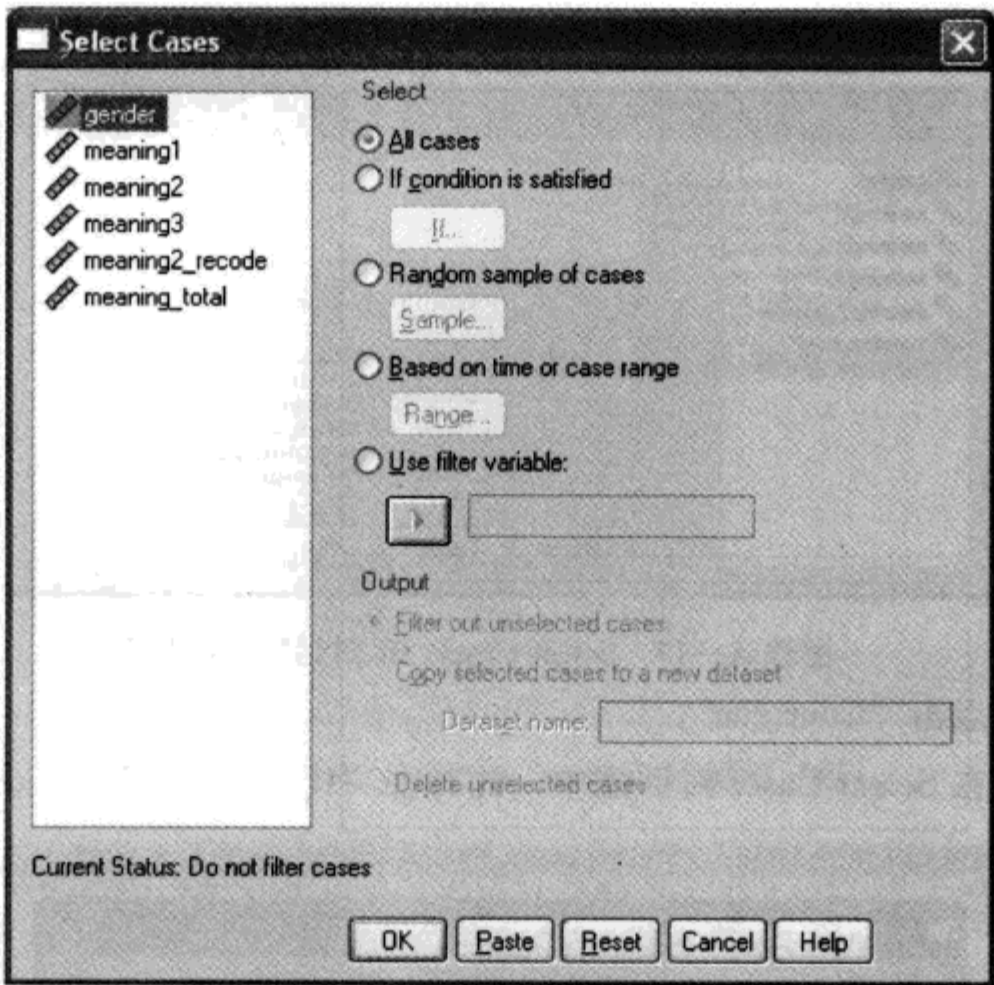
使用 Select Cases 方法选择仅为男性的情况：

- 1. 从菜单栏中选择 **Data>Select Cases...**（见图表 A—14）。



图表 A—14 Select Cases 方法的菜单命令

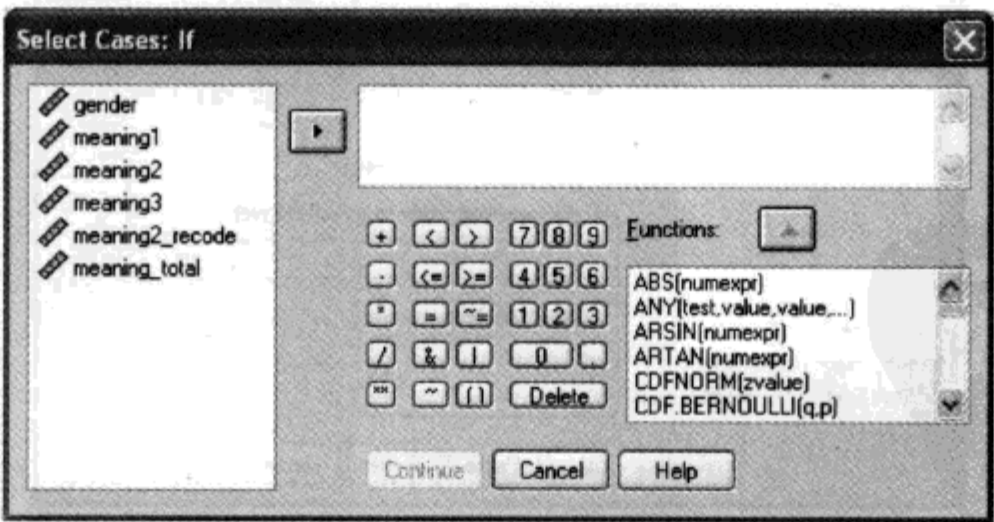
2. Select Cases 对话框打开（见图表 A—15）。



图表 A—15 Select Cases 对话框

3. 在 Select 下，选中 If condition is satisfied（是否满足条件）。

4. 点击 “If”（在 If condition is satisfied 下面）。Select Cases: If 对话框打开（见图表 A—16）。

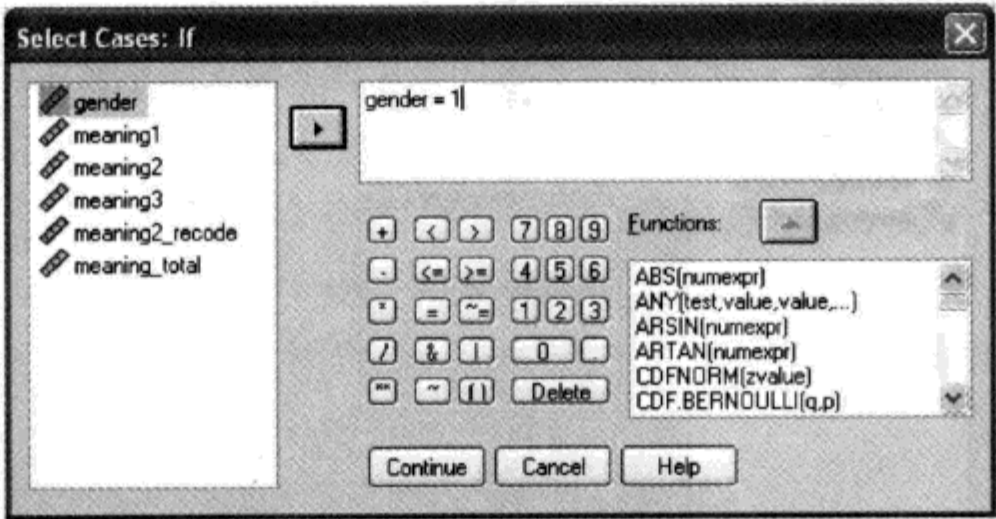


图表 A—16 Select Cases: If 对话框

5. 选中变量 **gender**，点击右箭头按钮（▶）将其移到右边的框中。

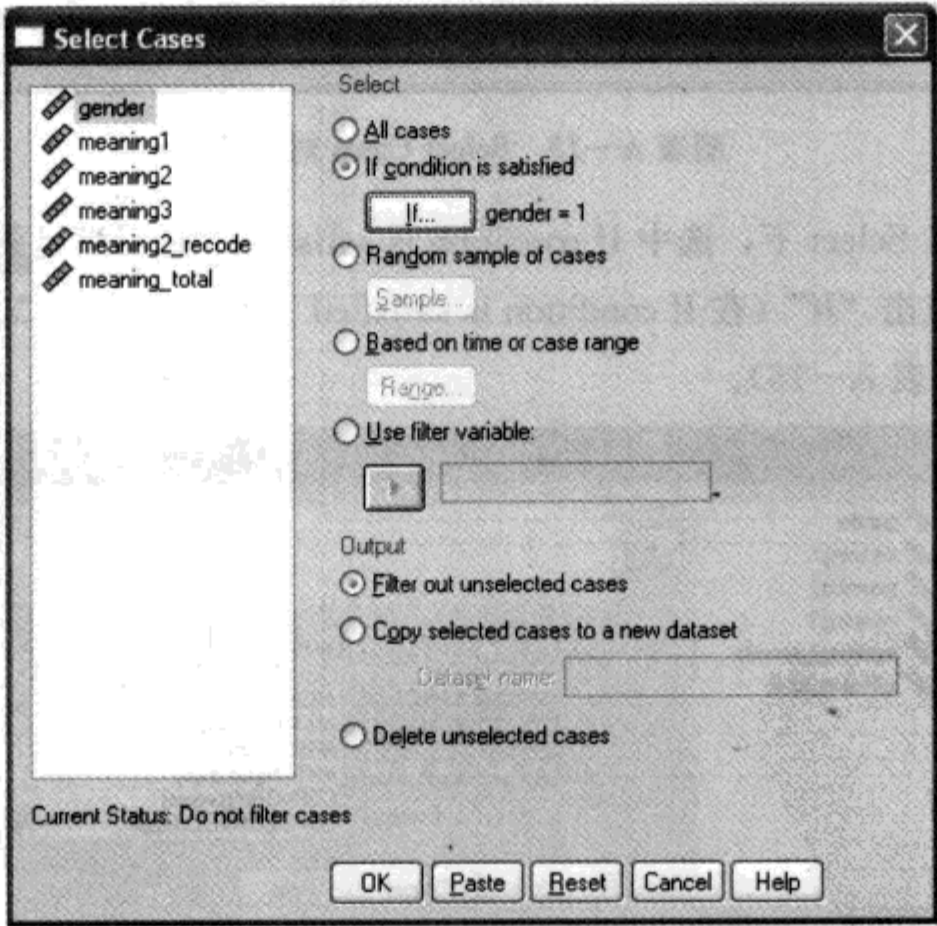
对于这个例子我们要选择仅为男性的情况。代表男性的编码为 1，因此在

右边输入“=1”（输入内容不包括引号）（见图表 A—17）。这说明 SPSS 将只选择男性的情况。



图表 A—17 Select Cases: If 对话框（续）

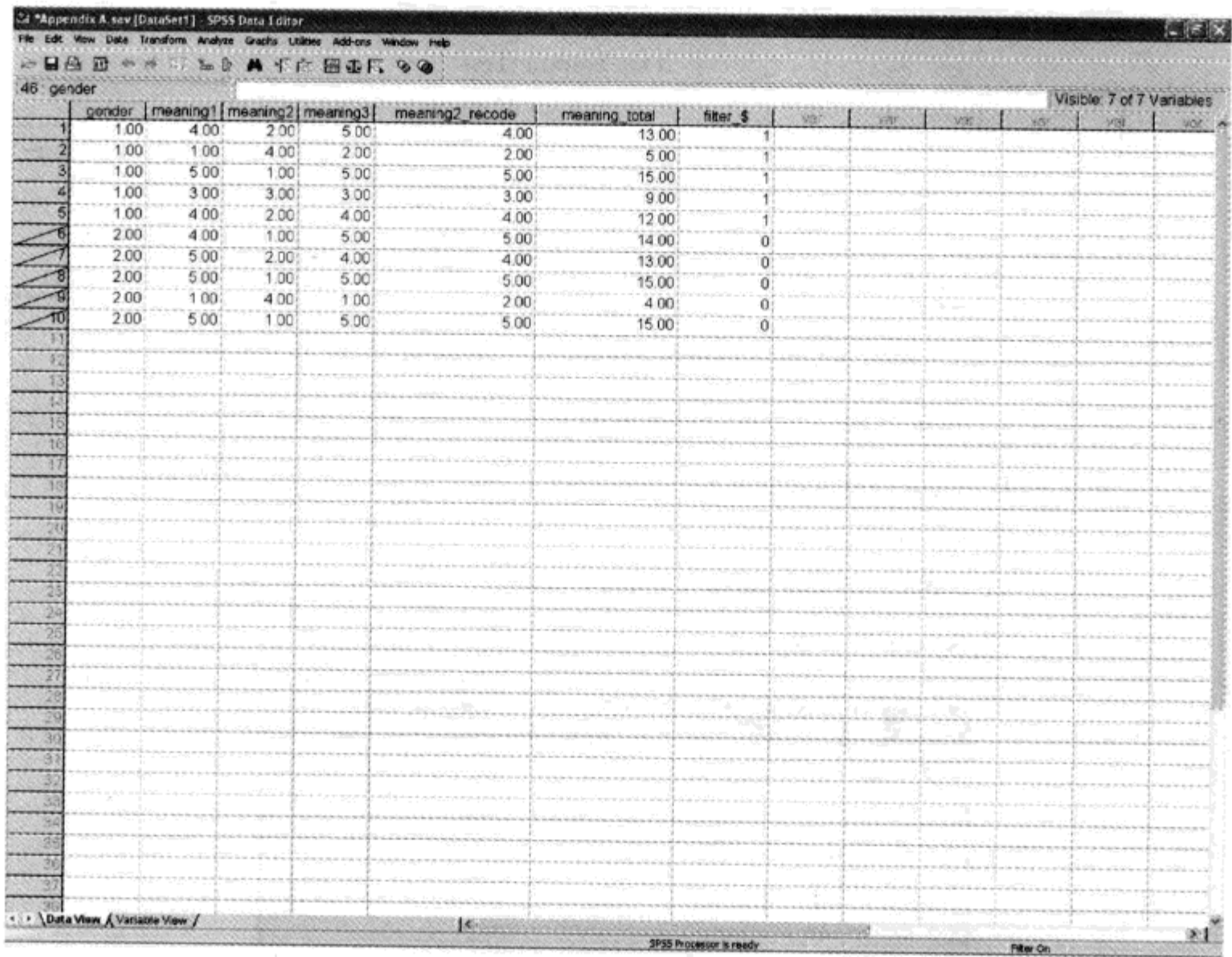
- 6. 点击“Continue”。
- 7. 在 Select Cases 对话框中，此时在“if”的右边有 gender=1（见图表 A—18）。



图表 A—18 Select Cases 对话框

- 8. 点击“OK”。
- 注意图表 A—19 所示的 Data View 窗口，在观察对象 6~10 行的行头有

斜线，说明女性不包括在分析的范围。在数据集中还创建了一个筛选变量，来说明哪些情况是被选择的（“1”表示被选择的，“0”表示没有被选择的）。



图表 A—19 选择男性的数据文件

使用 Select Case 方法解答开始的问题，在 **meaning1** 上对男性的情况进行单样本 *t* 检验。单样本 *t* 检验的结果如图表 A—20 所示。

注意图表 A—20 所示的结果，没有表明男性被选择了。One-Sample Statistics 表显示 *N* 为 5，说明检验仅仅对于数据文件中的 5 个参与者进行。（结果中，因为 *p* 值（“sig.”）为 0.587 大于 0.05，说明在 **meaning1** 上男性的得分与中立反应的分值 “3” 没有明显的不同。）⁵

下面我们将说明如何使用 Split File 方法，它是另一种对 SPSS 中数据集的单独的部分执行分析的方法。在使用 Split File 方法之前，我们需要关闭 Se-

T-Test

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
meaning1	5	3.4000	1.51658	0.67823

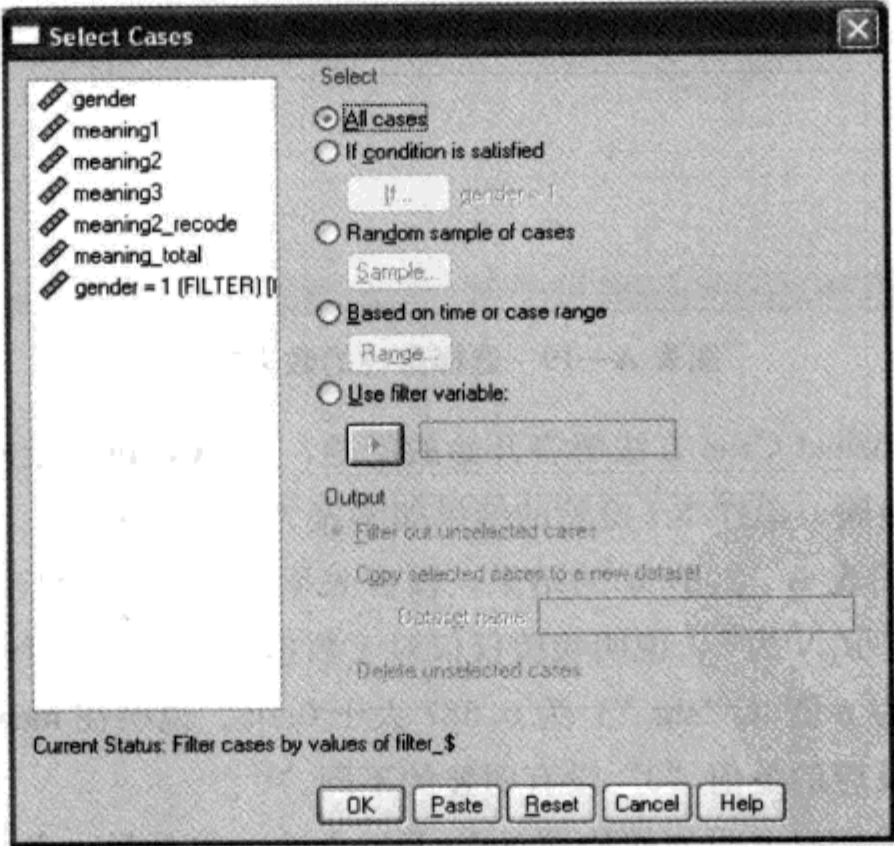
One-Sample Test						
	Test Value=3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
meaning1	0.590	4	0.587	0.40000	-1.4831	2.2831

图表 A—20 对男性的 meaning1 的单样本 t 检验的结果

lect Cases 方法以便可以访问完整的数据集。如果没有关闭 Select Cases 方法，那么所有后面的分析都是在男性的情况下。

关闭 Select Case 方法：

- 1. 从菜单栏中选择 Data>Select Cases...。
- 2. 在 Select 下，选中 All Cases（见图表 A—21）。



图表 A—21 Select Cases 对话框

3. 点击 “OK”。

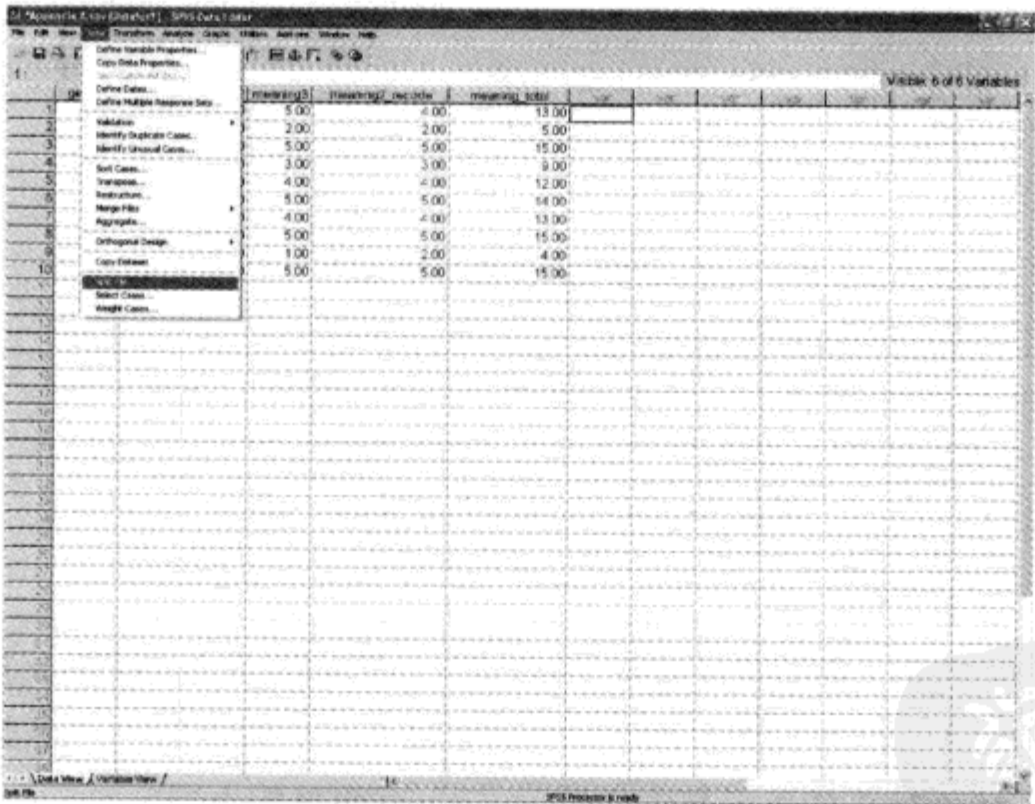
在 Data View 窗口中，此时所有的情况都应该被选择（不再只是观察对象 6~10 行）。尽管筛选变量仍旧存在于数据文件，但是它不再起作用。筛选变量可以通过以下方式删除（如果想要删除），单击 Data View 窗口中 **filter_** \$ 列的列头，并按下键盘上的 Delete 键。

拆分文件方法

假设不只需要对男性的情况做单样本 t 检验，而是要对男性和女性的情况分别在 **meaning1** 上进行单样本 t 检验，这种情况下，如果用 Select Cases 方法，需要执行两次分析（一次对男性，一次对女性）。与这种方法相比，一种更有效率的方法是使用拆分文件方法（Split File），它可以自动地对所关注变量的每一组执行独立分析。下面我们来举例说明 Split File 方法。

使用 Split File 方法：

- 1. 从菜单栏中选择 **Data>Split File...**（见图表 A—22）。

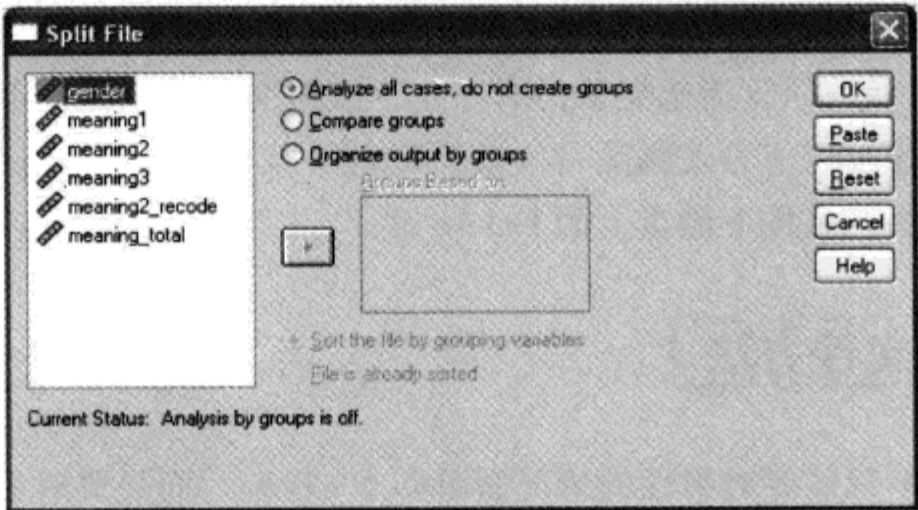


图表 A—22 Split File 方法的菜单命令

- 2. Split File 对话框打开（见图表 A—23）。

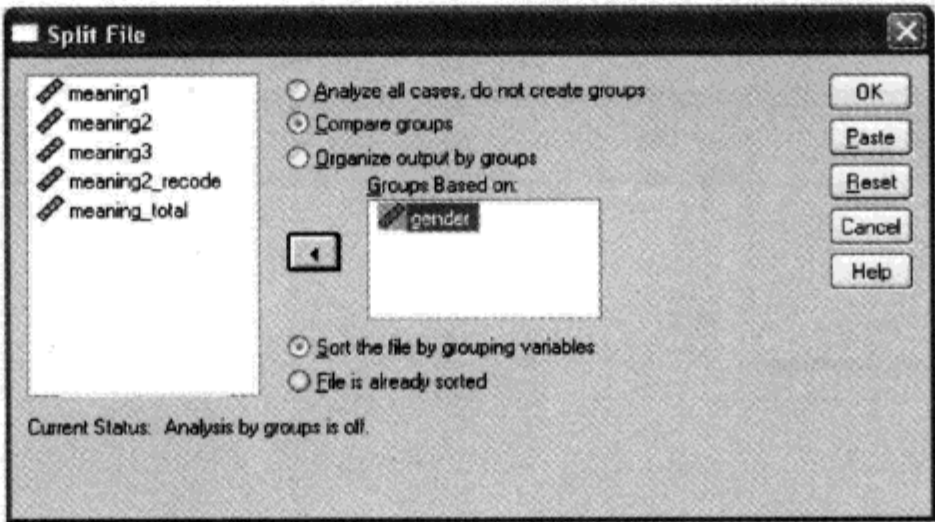
对于拆分文件有两种不同的方法：Compare groups 和 Organize output by groups。Compare groups 是在同一个表中按组显示结果，而 Organize output by

groups 是分组在不同的表中显示结果。下面我们使用 Compare groups 方法。



图表 A—23 Split File 对话框

- 3. 选择 Compare groups。
- 4. 选中 **gender**。点击向右箭头按钮 (►) 将 **gender** 移到 Groups Based on 框中 (见图表 A—24)。



图表 A—24 Split File 对话框 (续)

- 5. 点击 “OK”。
- 点击 “OK” 后，可以看到 “Split File On” 显示在 Data Editor 窗口的右下角。
- 图表 A—25 给出单样本 *t* 检验的结果（一个男性的，一个女性的），评价是否对 **meaning1** 的反应与中立反应 “3” 有明显的不同。
- 注意单样本 *t* 检验的输出在同一个表中显示了男性和女性的结果。如果选择 Organize output by groups，对男性和女性的结果会显示在不同的表中（男性的有两个表，女性的有两个表）。（顺便说一下，结果是否说明男性或女性在 **meaning1** 上的平均得分与中立反应 “3” 有显著不同呢？不是的，因为 *p* 值

分别为 0.587 和 0.266，都大于 0.05。)

T-Test

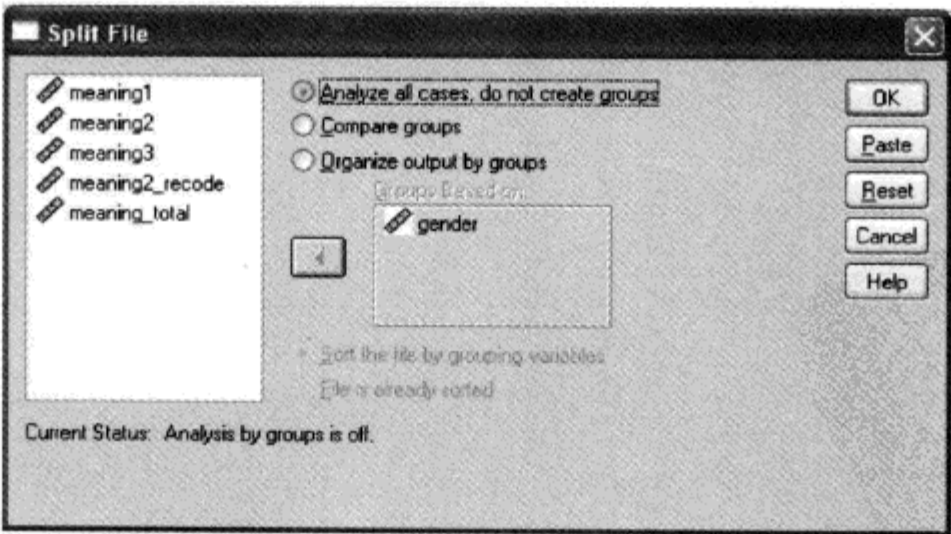
One-Sample Statistics				
gender	N	Mean	Std. Deviation	Std. Error Mean
male meaning1	5	3.4000	1.51658	0.67823
female meaning1	5	4.0000	1.73205	0.77460

One-Sample Test						
gender	Test Value=3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
male meaning1	0.590	4	0.587	0.40000	-1.4831	2.2831
female meaning1	1.291	4	0.266	1.00000	-1.1506	3.1506

图表 A—25 男性和女性在 meaning1 上的单样本 t 检验的结果

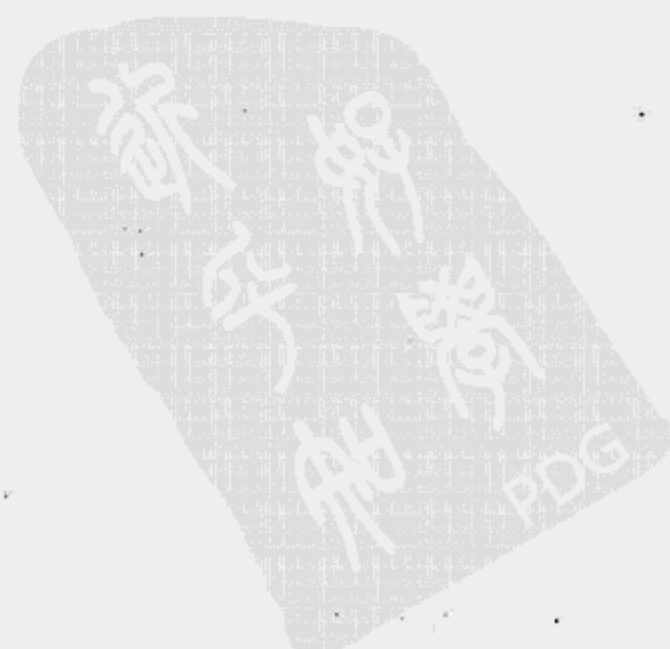
关闭拆分文件方法

正如之前在 Select Cases 方法的案例中讨论的那样，如果你想要以完整的例子执行分析，那么首先应该关闭 Split File 方法。为关闭 Split File 方法，在 Split File 对话框中选择 Analyze all cases, do not create groups，并点击“OK”（见图表 A—26）。（如果在使用 Split File 或 Select Cases 方法结束后关闭了 SPSS，当 SPSS 再次打开时会默认已关闭了该方法。）



图表 A—26 关闭 Split File 方法

至此结束了对 Recode，Compute，Select Cases 及 Split File 方法的讨论。



附录 B

导入文件

本附录举例说明了如何向 SPSS 中导入 Excel 文件。为演示这个方法，我们把附录 A 中的数据做成 Excel 文件，如图表 B—1 所示。

图表 B—1 Excel 文件：Excel File _ Appendix B. xls

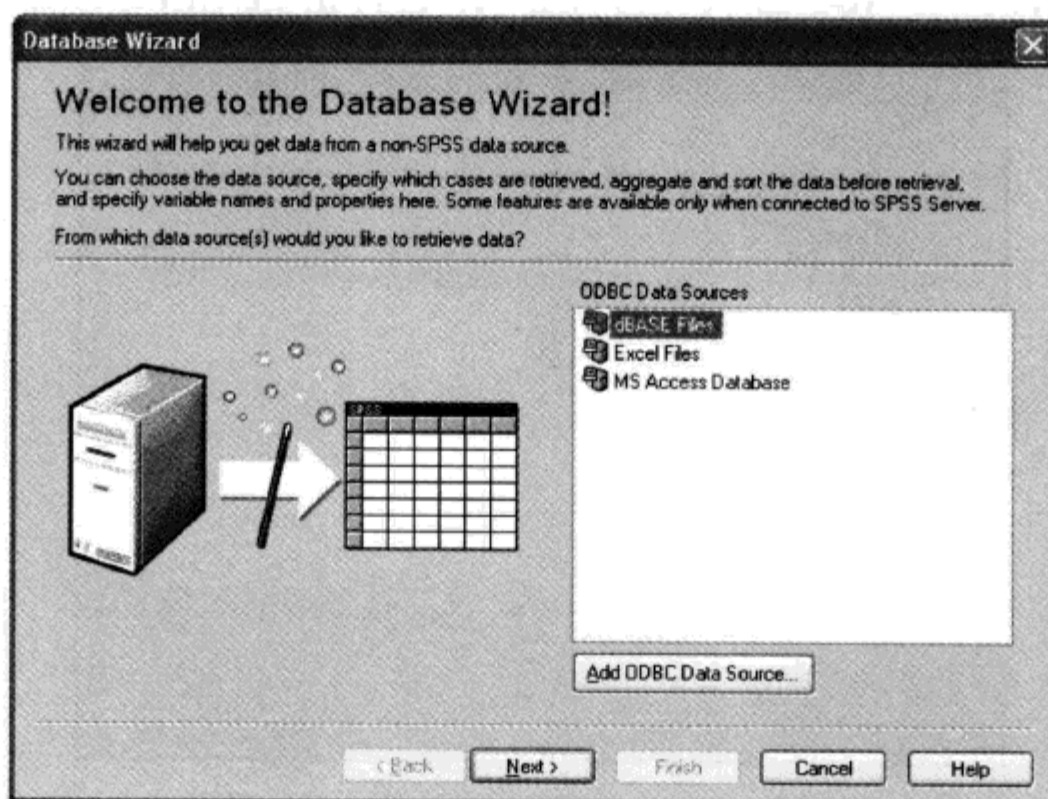
	A	B	C	D
1	gender	meaning1	meaning2	meaning3
2	1	4	2	5
3	1	1	4	2
4	1	5	1	5
5	1	3	3	3
6	1	4	2	4
7	2	4	1	5
8	2	5	2	4
9	2	5	1	5
10	2	1	4	1
11	2	5	1	5

为了把 Excel 文件导入 SPSS，我们使用 SPSS Database Wizard（SPSS 数据库向导），它的作用是向 SPSS 导入文件时提供帮助。

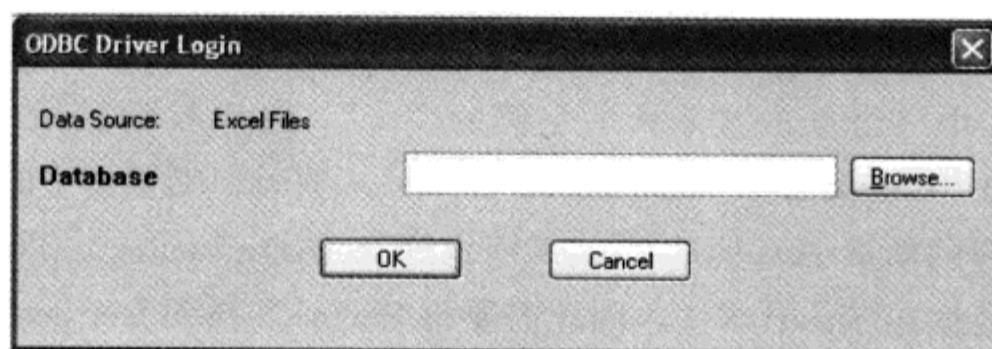
进入 Database Wizard：

- 1. 从菜单栏中选择 **File>Open Database>New Query...**（见图表 B—2）。
- 2. Database Wizard 对话框打开了（见图表 B—3）。

在 Database Wizard 对话框中列出了三种不同的文件类型，包括 dBASE

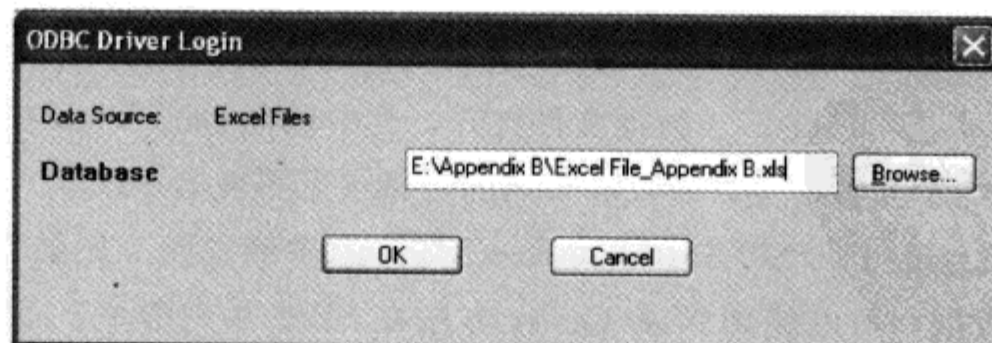


图表 B—3 Database Wizard 对话框



图表 B—4 ODBC Driver Login 对话框

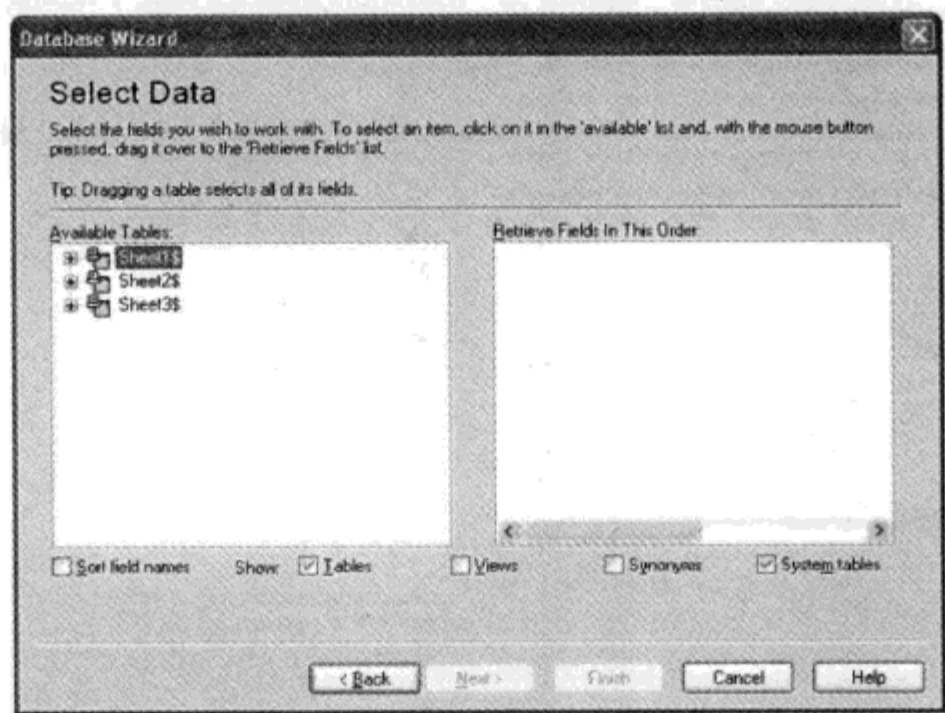
5)。(注意：图表 B—5 中所示文件在 E 盘，但是该文件可能在你的电脑中不同的位置。)



图表 B—5 ODBC Driver Login 对话框 (续)

8. 点击“OK”。

9. Database Wizard: Select Data 对话框打开（见图表 B—6）。



图表 B—6 Database Wizard: Select Data 对话框

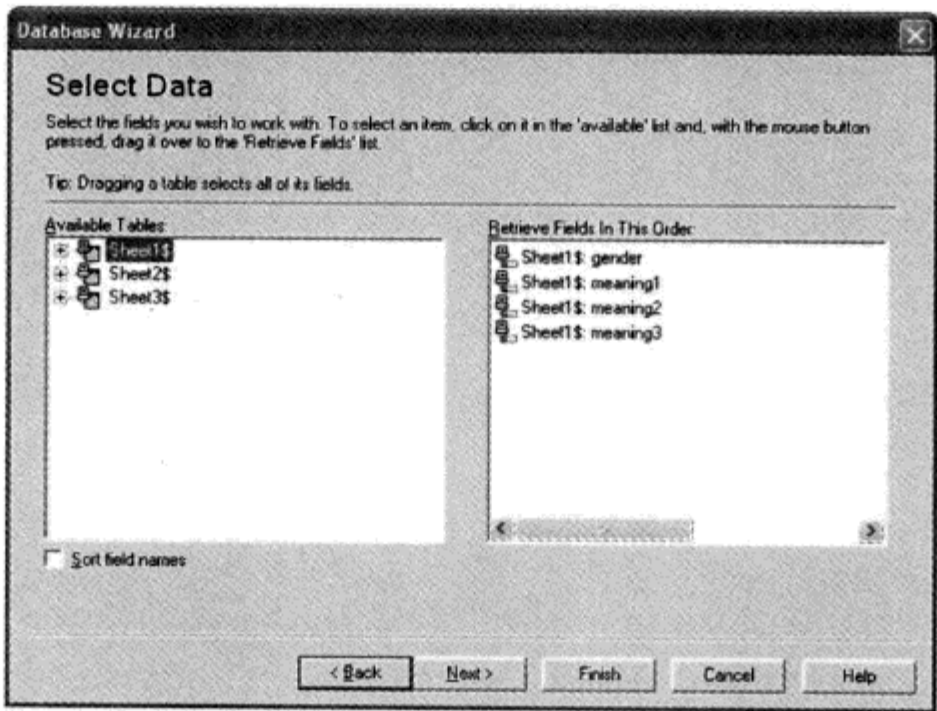
三个不同的表（Sheet1 \$，Sheet2 \$，Sheet3 \$）出现在 Available Tables 框，对应三个为 Excel 文件默认生成的工作表。我们的数据在第一个工作表 Sheet1 \$ 中（当点击 Sheet1 \$ 左侧的“+”，可以看到显示的变量 **gender**、**meaning1**、**meaning2** 及 **meaning3**）。Database Wizard 对话框中的指示说明需要把感兴趣的项从 Available Tables 框拖到 Retrieve Fields in This Order 框中。我们的数据在 Sheet1 \$ 中，因此需要将 Sheet1 \$ 拖到 Retrieve Fields in This Order 框中。

10. 选择 Sheet1 \$，并将它拖到 Retrieve Fields in This Order 框中（见图表 B—7）

Retrieve Fields in This Order 框中显示了数据文件中的四个变量，**gender**、**meaning1**、**meaning2** 及 **meaning3**，确认我们选择的是正确的文件。

11. 点击“Finish”。数据文件打开，其中包含 Data View 窗口的第 1 行到第 10 行的参与者得分，以及前 4 列的变量（见图表 B—8）。

至此结束了向 SPSS 中导入 Excel 文件的举例说明。如果你有 Ms Access 保存的数据，可以使用 New Query 方法在 SPSS 中打开它。如果你有数据值被逗号、空格或制表符分离的形式保存的数据文件，该文件需要使用 SPSS 中一种选择性方法打开。为打开这些类型的文件，从菜单栏中选择 **File>Read Text Data...**，并根据屏幕上的指示操作。



图表 B—7 Database Wizard: Select Data 对话框 (续)

SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

48 meaning1

	gender	meaning1	meaning2	meaning3	V01	V02	V03	V04	V05	V06	V07	V08	V09	V10
1	1.00	4.00	2.00	5.00										
2	1.00	1.00	4.00	2.00										
3	1.00	5.00	1.00	5.00										
4	1.00	3.00	3.00	3.00										
5	1.00	4.00	2.00	4.00										
6	2.00	4.00	1.00	5.00										
7	2.00	5.00	2.00	4.00										
8	2.00	5.00	1.00	5.00										
9	2.00	1.00	4.00	1.00										
10	2.00	5.00	1.00	5.00										
11														
12														
13														
14														
15														
16														
17														
18														
19														
20														
21														
22														
23														
24														
25														
26														
27														
28														
29														
30														
31														
32														
33														
34														
35														
36														
37														
38														
39														
40														
41														
42														
43														
44														
45														
46														
47														
48														

Visible: 4 of 4 Variables

Data View Variable View

SPSS Processor is ready

图表 B—8 输入 SPSS 的 Excel 文件



各章练习答案

第 1 章 SPSS 概述

1. c. Case Processing Summary^a

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
age	7	100.0%	0	0.0%	7	100.0%
gender	7	100.0%	0	0.0%	7	100.0%
wellbeing	7	100.0%	0	0.0%	7	100.0%
activities	7	100.0%	0	0.0%	7	100.0%

a. Limited to first 100 cases.

Case Summaries^a

	age	gender	wellbeing	activities
1	86.00	male	4.00	2.00
2	72.00	female	7.00	6.00
3	59.00	female	6.00	5.00
4	86.00	female	8.00	7.00
5	92.00	female	4.00	1.00
6	68.00	male	2.00	3.00
7	73.00	male	8.00	5.00
Total N	7	7	7	7

a. Limited to first 100 cases.

d. 在 Case Summaries 表中输出变量值标签 (male/female)。当变量值标签生成后，它们被默认打印在结果里，代替原来输入 SPSS 中的数值 (1, 2)。这样有利于阅读结果，避免弄混男性和女性 (万一忘记哪组被赋值 1，哪组被赋值 2)。

2. b. Case Processing Summary^a

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
therapy	12	100.0%	0	0.0%	12	100.0%
gender	12	100.0%	0	0.0%	12	100.0%
selfesteem	12	100.0%	0	0.0%	12	100.0%
angermanage	12	100.0%	0	0.0%	12	100.0%

a. Limited to first 100 cases.

Case Summaries^a

	therapy	gender	selfesteem	angermanage
1	nondirective	male	25.00	12.00
2	nondirective	male	30.00	15.00
3	nondirective	male	34.00	13.00
4	nondirective	female	28.00	19.00
5	nondirective	female	39.00	21.00
6	nondirective	female	42.00	15.00
7	directive	male	37.00	9.00
8	directive	male	29.00	19.00
9	directive	male	26.00	22.00
10	directive	female	38.00	17.00
11	directive	female	43.00	11.00
12	directive	female	26.00	21.00
Total N	12	12	12	12

a. Limited to first 100 cases.

3. b. Case Summaries^a

	gender	number classes	hourworked
1	male	3.00	20.00
2	male	5.00	10.00

续前表

	gender	number classes	hourworked
3	female	5.00	20.00
4	female	4.00	14.00
5	female	6.00	6.00
6	male	2.00	0.00
7	female	1.00	40.00
8	female	3.00	20.00
9	male	5.00	15.00
10	female	6.00	25.00
Total N	10	10	10

a. Limited to first 100 cases.

第 2 章 描述统计

1. a. 就性别而言有 4 个男孩和 4 个女孩，就方法而言有 4 个用脚尖踢球和 4 个用脚后跟踢球。

b. Statistics

	experience	distance	accuracy
Mean	4.6250	38.8750	5.7500
Std. Deviation	2.61520	7.10005	1.83225

c. Report

gender		experience	distance	accuracy
boys	Mean	5.5000	42.5000	6.0000
	Std. Deviation	2.88675	6.45497	2.16025
girls	Mean	3.7500	35.2500	5.5000
	Std. Deviation	2.36291	6.39661	1.73205

男孩在所有 3 个变量上都有更高的均值。男孩也在所有 3 个变量上有更大的标准差。

d.

		Report		
method		experience	distance	accuracy
toe kick	Mean	5.0000	44.0000	4.5000
	Std. Deviation	2.16025	4.69042	1.29099
heel kick	Mean	4.2500	33.7500	7.0000
	Std. Deviation	3.30404	5.05800	1.41421

用脚尖踢球者在踢球上有更多的经验（5.00）和更长的距离（44.00），而用脚后跟踢球者有更高的准确度（7.00）。用脚后跟踢球者在所有 3 个变量上有更大的标准差。

e.

		Report		
gender	method	experience	distance	accuracy
boys	toe kick	Mean	5.5000	47.5000
		Std. Deviation	0.70711	3.53553
	heel kick	Mean	5.5000	37.5000
		Std. Deviation	4.94975	3.53553
girls	toe kick	Mean	4.5000	40.5000
		Std. Deviation	3.53553	2.12132
	heel kick	Mean	3.0000	30.0000
		Std. Deviation	1.41421	2.82843

用脚后跟踢球的女孩得到最低的平均经验值（3.00），而用脚尖踢球的男孩得到最低的平均准确度得分（4.5）。用脚尖踢球的男孩获得所有四种情况中最高的平均踢球距离（47.50）。

2. a. 就位置而言，每个地区有 10 人（西海岸、中西部及东海岸）。数据集里有 15 个男性和 15 个女性。

b.

Statistics		
	taste	clarity
Mean	6.8667	8.1333
Std. Deviation	1.45586	1.16658

c.

Report

location		taste	clarity
West Coast	Mean	6.9000	7.8000
	Std. Deviation	1.28668	1.03280
Midwest	Mean	7.8000	8.7000
	Std. Deviation	0.91894	1.15950
East Coast	Mean	5.9000	7.9000
	Std. Deviation	1.52388	1.19722

中西部在 **taste** (7.80) 和 **clarity** (8.70) 上都有最高的均值。东海岸在 **taste** (1.52) 和 **clarity** (1.20) 上都有最大的标准差。

d.

Report

gender		taste	clarity
male	Mean	6.6667	8.0667
	Std. Deviation	1.58865	1.27988
female	Mean	7.0667	8.2000
	Std. Deviation	1.33452	1.08233

女性在 **taste** (7.07) 和 **clarity** (8.20) 上都有最高的均值。男性在 **taste** (1.59) 和 **clarity** (1.28) 上都有最大的标准差。

e.

Report

gender	location		taste	clarity
male	West Coast	Mean	6.6000	8.0000
		Std. Deviation	1.14018	1.22474
	Midwest	Mean	8.0000	8.8000
		Std. Deviation	0.70711	1.30384
	East Coast	Mean	5.4000	7.4000
		Std. Deviation	1.67332	1.14018
female	West Coast	Mean	7.2000	7.6000
		Std. Deviation	1.48324	0.89443
	Midwest	Mean	7.6000	8.6000
		Std. Deviation	1.14018	1.14018
	East Coast	Mean	6.4000	8.4000
		Std. Deviation	1.34164	1.14018

东海岸的男性在 **taste** (5.40) 和 **clarity** (7.40) 上都有最低的均值。中西部的男性在 **taste** (8.00) 和 **clarity** (8.80) 上都有最高的均值。

3. a. 有 15 个男性和 15 个女性。每个 **SES** 组中有 10 人 (10 个低的、10 个中等、10 个高的)。

b.

Statistics

	hourstv	readingscores
Mean	3.1500	28.2667
Std. Deviation	1.37922	12.34262

c.

Report

gender		hourstv	readingscores
male	Mean	3.3367	25.9333
	Std. Deviation	1.52976	11.90718
female	Mean	2.9633	30.6000
	Std. Deviation	1.23512	12.73241

男孩在 **hourstv** 上有较高的均值 (3.34)，而女孩在 **readingscores** 上有较高的均值 (30.60)。男孩在 **hourstv** (1.53) 和 **readingscores** (11.91) 上都有最大的标准差。

d.

Report

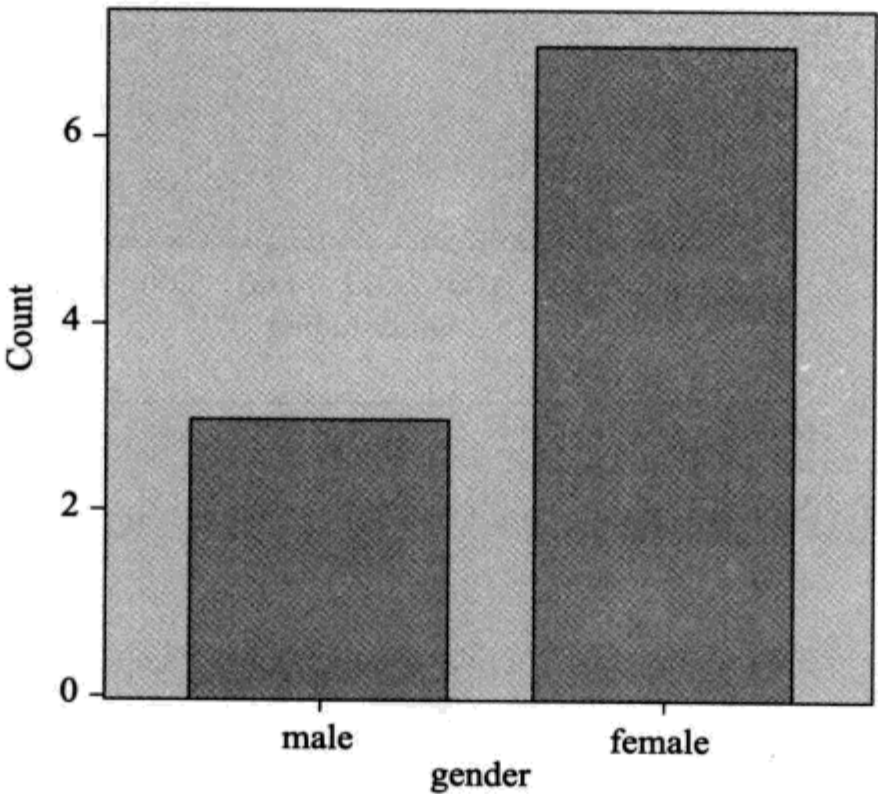
gender	ses		hourstv	readingscores
male	low	Mean	3.3400	18.0000
		Std. Deviation	1.47792	9.87421
	middle	Mean	3.9000	23.8000
		Std. Deviation	1.03983	6.30079
	high	Mean	2.7700	36.0000
		Std. Deviation	2.03150	12.16553
female	low	Mean	3.3400	18.0000
		Std. Deviation	1.47792	9.87421
	middle	Mean	3.6000	36.4000
		Std. Deviation	0.51841	11.32696
	high	Mean	1.9500	37.4000
		Std. Deviation	0.95851	6.58027

中等的 **SES** 组男孩看电视最多，均值为 3.90。较高的 **SES** 组女孩在

readingscores 上有最高的均值 (37.40)。

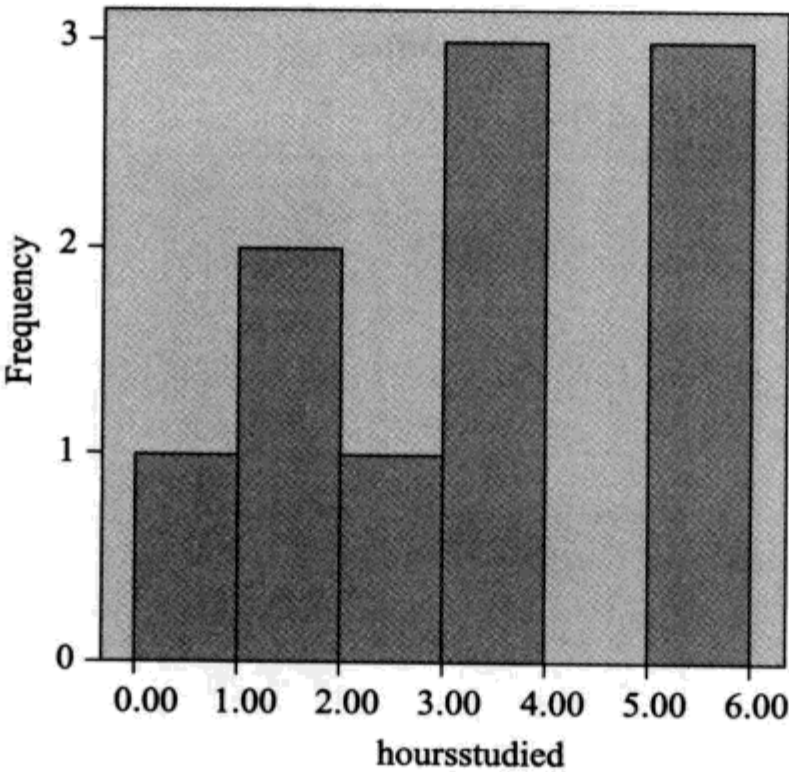
第 3 章 作图

1. a.



有 3 个男孩和 7 个女孩。

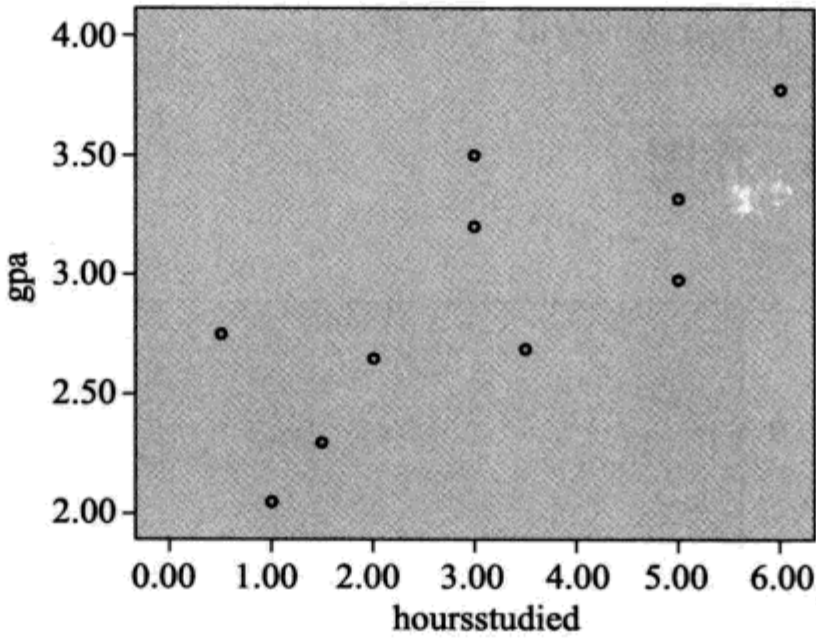
b.



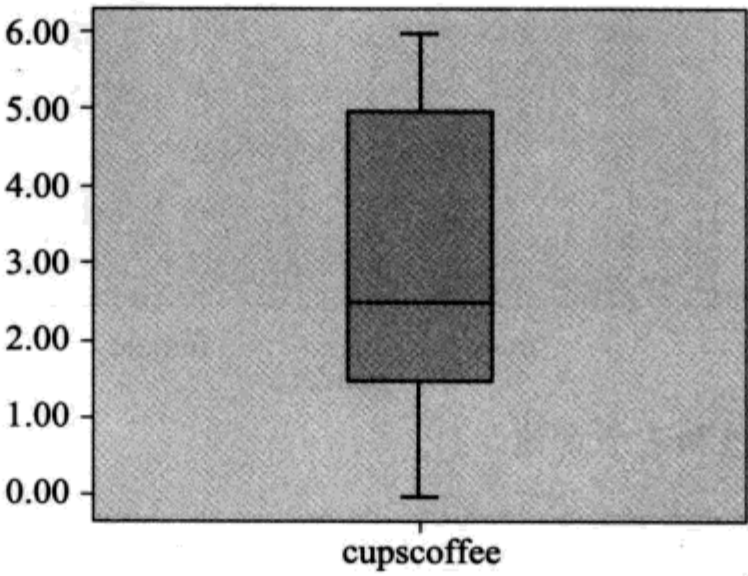
Mean=3.05
Std.Dev.=1.848
N=10

学习时间 (hoursstudied) 的均值为 3.05。

c.

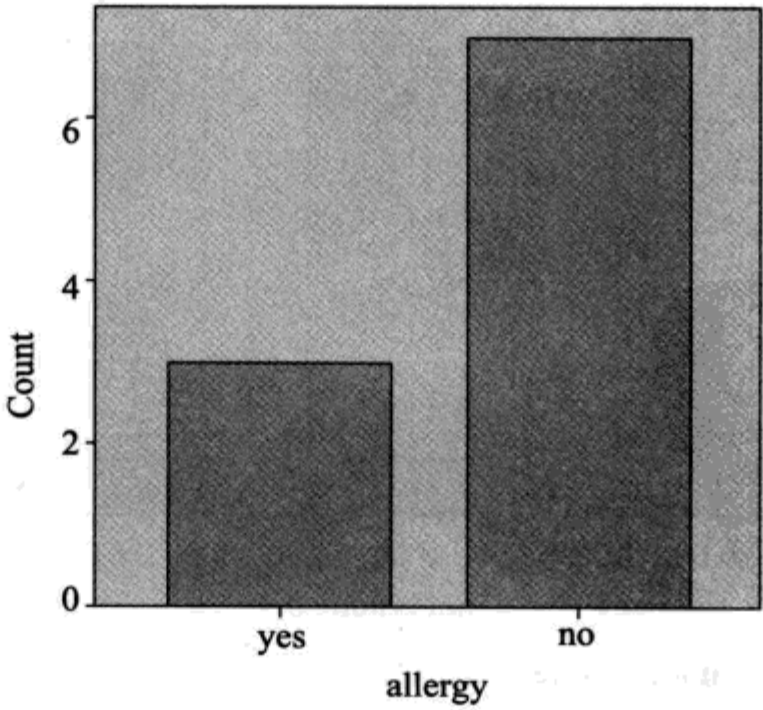


d.

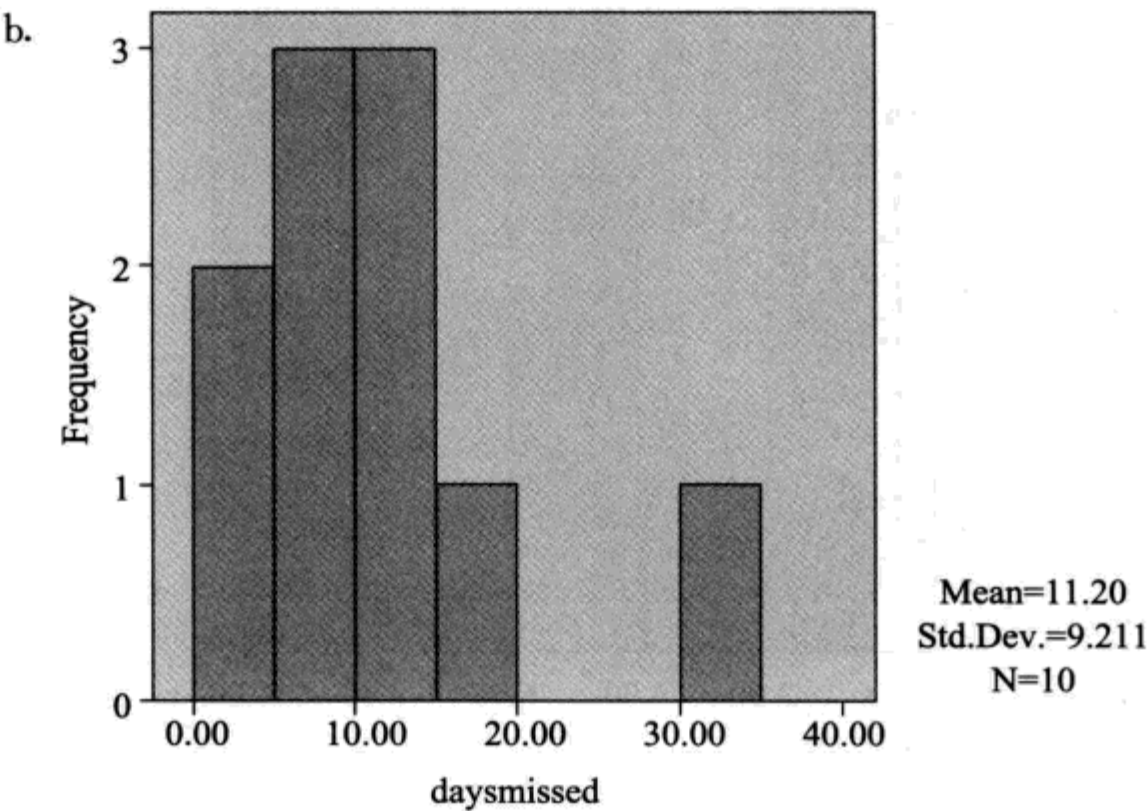


没有异常值在盒形图之外。

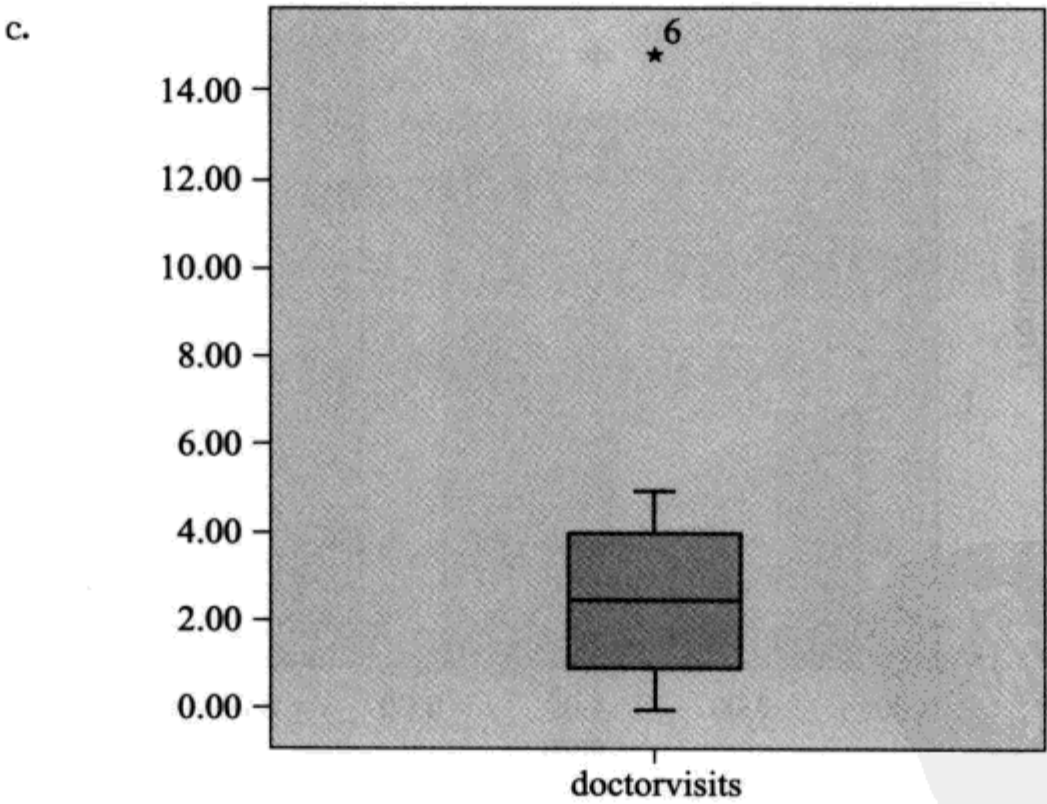
2. a.



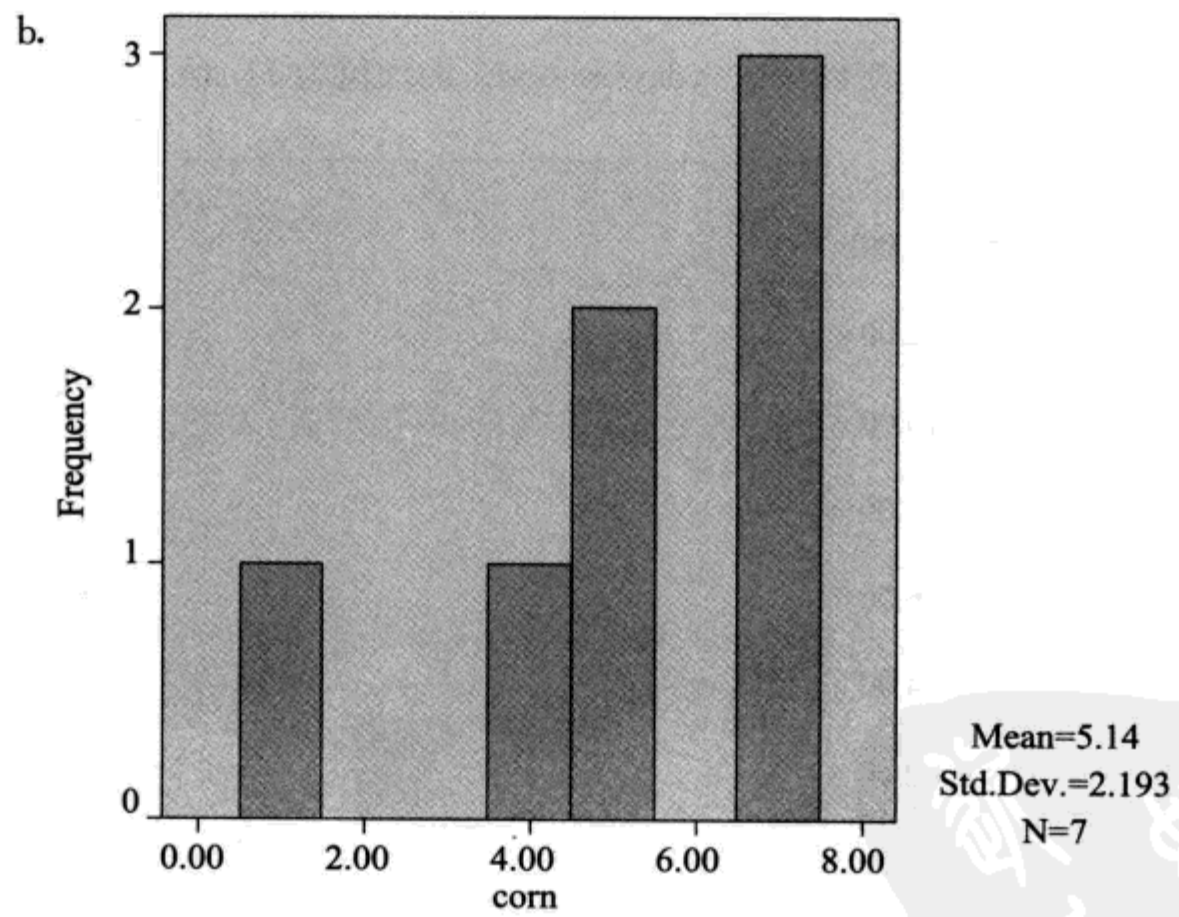
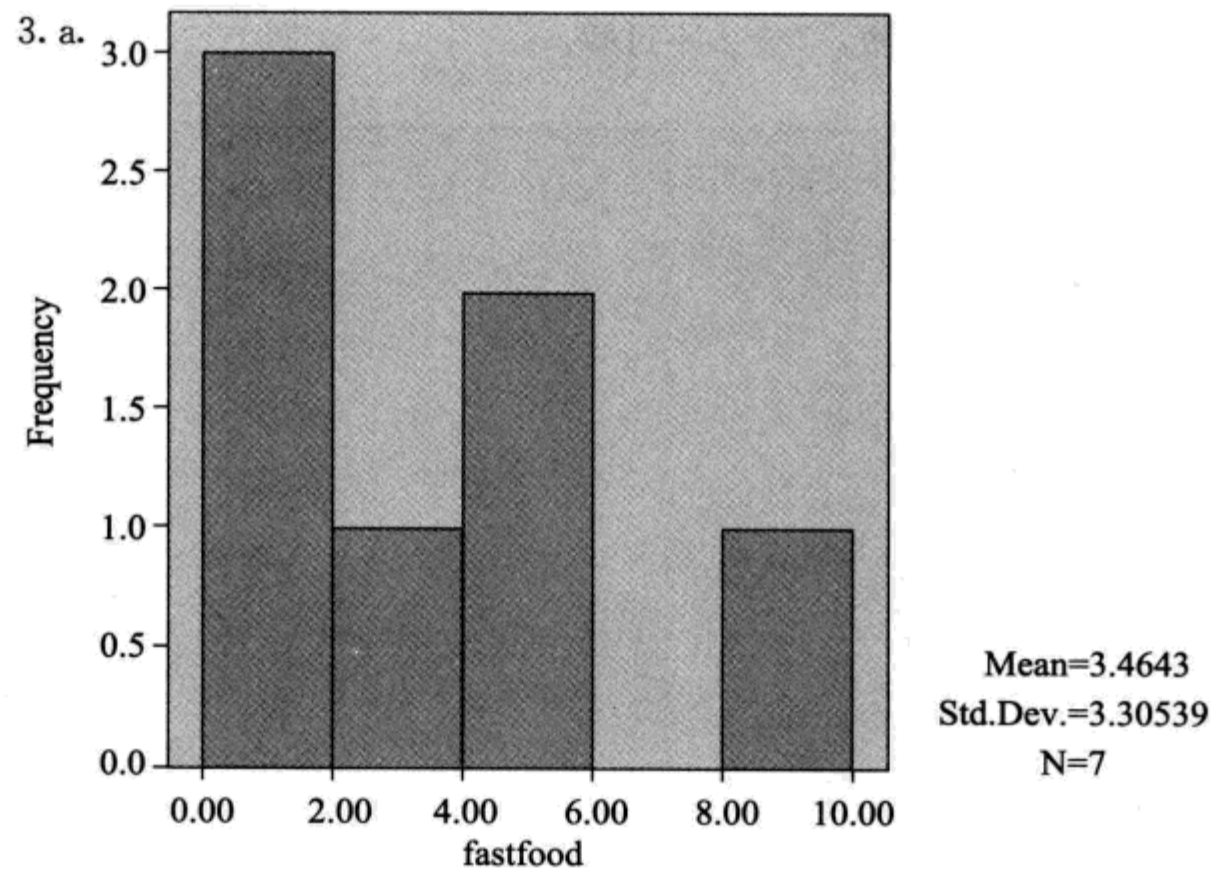
有 3 个人过敏，7 个人不过敏。



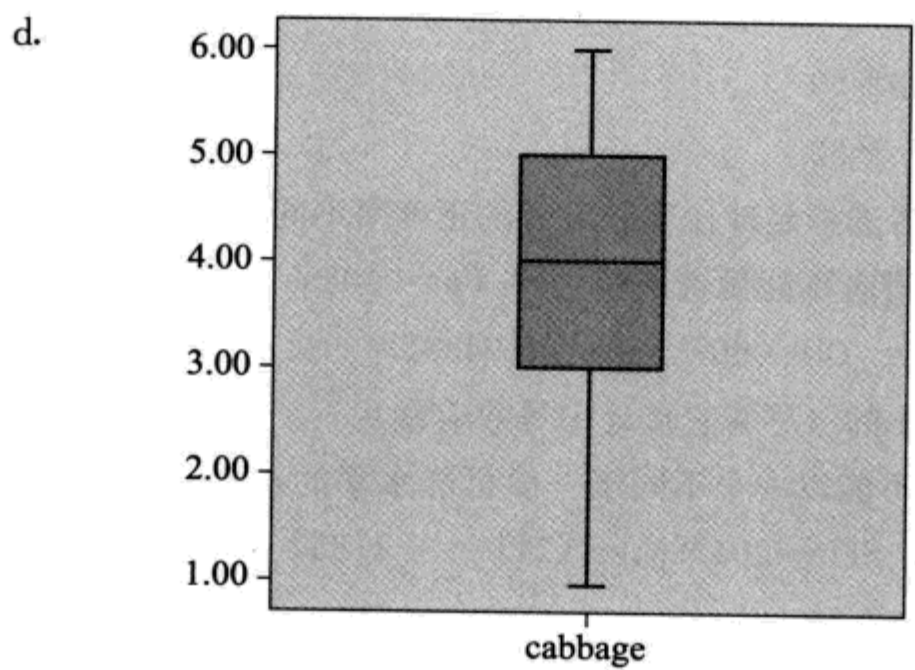
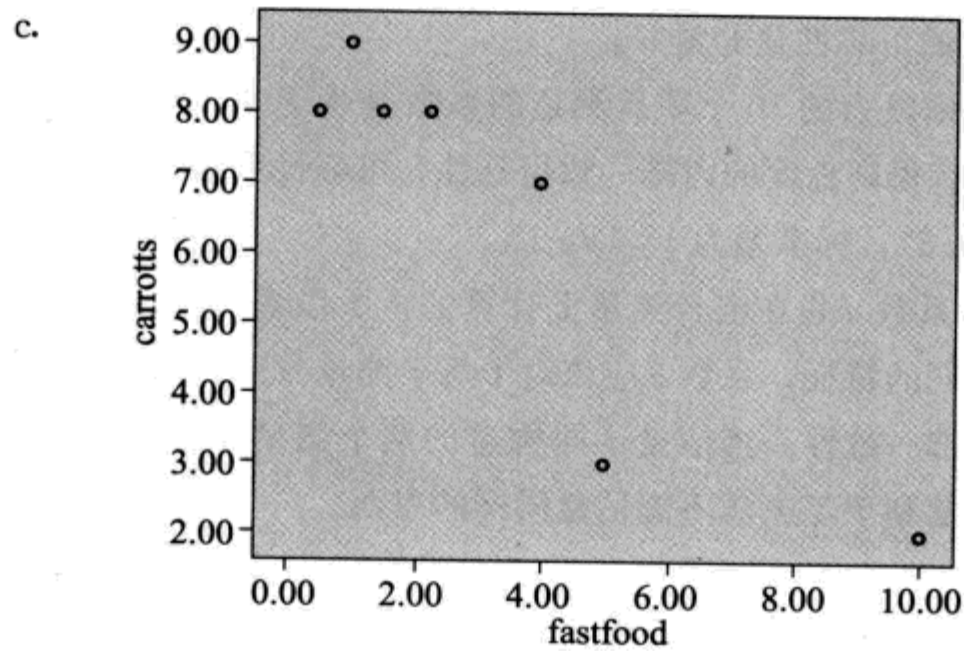
每年没有上班的天数（daysmissed）的均值为 11.20。



第 6 个观测是在去年见医生的 15 人中的异常值。



玉米的均值为 5.14。



没有异常值在盒形图之外。

第 4 章 可靠性 (用 α 系数度量)

1. a. 系数 $\alpha=0.72$ 。
b. 测量的可靠性是不错的。
c. 表示居民满意度的 7 个项目估量的系数 α 是基于 20 个家庭主妇的反应计算得到的。测量可靠性是 fair (中下等), 系数 $\alpha=0.72$ 。总测量上平均得分是 25.75, 标准差为 7.18。
2. a. 系数 $\alpha=0.95$ 。

b. 测量的可靠性是非常好。

c. 表示恢复力的 10 个项目测量的系数是基于 30 个参与者的反应计算得到的。测量有很高程度的内部一致可靠性，可靠性系数 $\alpha=0.95$ 。总测量上平均得分是 50.37，标准差为 13.59。

3. 因为系数 α 是在组合测量上计算的，所以系数 α 的估计较低。当两个测量估量不同的特征，系数 α 应该对于每个测量分别计算。回忆系数 α 估量项的集合的内部一致性。当估量不同构造的两个测量联合时，内部一致性会降低，因为自信和手工灵巧不能估量同样的东西。

第 5 章 单样本 t 检验

1. a. $H_0: \mu=50$

$H_1: \mu \neq 50$

b. “预言者选择足球比赛获胜者的正确率不同于一般可能性的 50% 吗？”

c. 不。预言者不能挑选出获胜者 ($p=0.085$)。

d. 效应量 = $(48.0667 - 50) / 4.04381 = -0.48$ (或 0.48)。用科恩的标准，效应量是小的 (尽管它接近中等效应量)。

e. 预言者不能以一个不同于一般可能水平的正确率挑选出获胜的足球队 ($M=48.07\%$, $SD=4.04\%$), $t(14) = -1.85$, $p > 0.05$, $d=0.48$ 。

2. a. $H_0: \mu=50$

$H_1: \mu \neq 50$

b. “当地校区四年级的数学考试成绩不同于全国平均水平 (50) 吗？”

c. 是的。学生们在数学考试中的成绩比全国平均水平更好 ($p=0.018$)。

d. 效应量 = $(55.44 - 50) / 10.73577 = 0.51$ 。用科恩的标准，这表示一个中等效应，当地校区学生的考试成绩高于全国平均水平，高出值约为 0.5 倍标准差。

e. 拒绝 H_0 。使用了数学技巧项目的学生的成绩明显高于 ($M=55.44$, $SD=10.74$) 全国平均水平 (50), $t(24) = 2.53$, $p < 0.05$, $d=0.51$ 。

3. a. $H_0: \mu=3$

$H_1: \mu \neq 3$

b. “学生们是支持还是反对候选人？”

- c. 都不是。学生们既不支持也不反对候选人 ($p=0.290$)。
- d. 效应量 = $(3.16-3) / 1.056\ 76=0.15$ 。用科恩的标准, 效应量非常小。
- e. 学生们既不明显支持也不反对候选人 ($M=3.16$, $SD=1.06$), $t(49)=1.07$, $p>0.05$, $d=0.15$ 。

第6章 独立样本 t 检验

1. a. $H_0: \mu_{\text{微创手术}} = \mu_{\text{传统手术}}$
 $H_1: \mu_{\text{微创手术}} \neq \mu_{\text{传统手术}}$
- b. “微创外科手术 (MIS) 与传统外科手术导致报告的疼痛有不同吗?”
- c. 方差没有显著不同。Levene's $F=0.005$, $p>0.05$ 。
- d. 是 ($p=0.001$)。
- e. $d=-3.58 \sqrt{\frac{15+15}{15 \times 15}} = -1.31$ 或 1.31 。效应量非常大, MIS 病人报告的疼痛指数低于接受传统外科手术的病人的平均水平, 差值为 1.31 倍标准差。
- f. 微创外科手术 (MIS) 与传统外科手术的疼痛指数有明显不同, $t(28)=-3.58$, $p<0.05$, $d=1.31$ 。接受 MIS 的病人 ($M=4.73$, $SD=1.67$) 比起那些接受传统外科手术的病人 ($M=6.87$, $SD=1.60$) 明显疼痛更轻。
2. a. $H_0: \mu_{\text{系统脱敏疗法}} = \mu_{\text{内爆疗法}}$
 $H_1: \mu_{\text{系统脱敏疗法}} \neq \mu_{\text{内爆疗法}}$
- b. “接受系统脱敏疗法的人和接受内爆疗法的人的蛇恐惧指数有不同吗?”
- c. 方差没有显著不同, Levene's $F=2.29$, $p>0.05$ 。
- d. 是 ($p=0.010$)。
- e. $d=-2.86 \sqrt{\frac{10+10}{10 \times 10}} = -1.28$ 或 1.28 。效应量很大, 接受了内爆疗法的人恐惧指数低于那些接受系统脱敏疗法的人, 差值为 1.28 倍标准差。
- f. 接受系统脱敏疗法和内爆疗法后报告的蛇恐惧指数有明显的不同, $t(18)=-2.86$, $p<0.05$, $d=1.28$ 。接受系统脱敏疗法的人对蛇恐惧的程度 ($M=47.10$, $SD=6.82$) 比起接受内爆疗法的人 ($M=59.30$, $SD=11.64$) 明显更低。

$$3. a. H_0: \mu_{\text{有宠物的家庭}} = \mu_{\text{没有宠物的家庭}}$$

$$H_1: \mu_{\text{有宠物的家庭}} \neq \mu_{\text{没有宠物的家庭}}$$

b. “对于有宠物的疗养院和没有宠物的疗养院居民满意度有不同吗?”

c. 方差没有显著不同, Levene's $F=0.241$, $p>0.05$ 。

d. 否 ($p=0.284$)。

$$e. d = -1.08 \sqrt{\frac{30+30}{30 \times 30}} = -0.28 \text{ 或 } 0.28. \text{ 效应量较小。}$$

f. 能够接触宠物和不能接触宠物的疗养院居民的满意度没有明显不同,
 $t(58) = -1.08$, $p>0.05$, $d=0.28$ 。

第7章 相依样本 t 检验

$$1. a. H_0: \mu_{\text{观影前}} - \mu_{\text{观影后}} = 0$$

$$H_1: \mu_{\text{观影前}} - \mu_{\text{观影后}} \neq 0$$

b. “观看电影后对候选人的态度与之前相比有不同吗?”

c. 是 ($p=0.001$)。

d. $d = (-6/5.73212) = -1.05$ 或 1.05 。效应量是 1.05 , 表明在观看竞选电影后对候选人的态度好于观看电影前, 差值超过 1 倍标准差。效应量是大的。

e. 在观看竞选电影后对政治候选人的态度 ($M=60.13$, $SD=17.53$) 比起之前 ($M=54.13$, $SD=15.68$) 明显更好, $t(14) = -4.05$, $p<0.05$, $d=1.05$ 。

$$2. a. H_0: \mu_{\text{放松训练前的压力}} - \mu_{\text{放松训练后的压力}} = 0$$

$$H_1: \mu_{\text{放松训练前的压力}} - \mu_{\text{放松训练后的压力}} \neq 0$$

b. “空中交通管制员在进行放松训练前后压力水平有不同吗?”

c. 是 ($p=0.047$)。

d. $d = (1.4/3.69156) = 0.38$, 效应量是 0.38 , 指出进行放松训练后压力水平以 0.38 倍标准差低于放松训练前。效应量是小的 (尽管接近中等效应)。

e. 在进行了放松训练后空中交通管制员的压力水平 ($M=34.83$, $SD=8.78$) 比起之前 ($M=36.23$, $SD=9.76$) 明显更低, $t(29) = 2.08$, $p<$

0.05, $d=0.38$ 。

3. a. $H_0: \mu_{\text{训练前}} - \mu_{\text{训练后}} = 0$

$H_1: \mu_{\text{训练前}} - \mu_{\text{训练后}} \neq 0$

b. “网球选手打球的准确度在接受运动心理学家训练之前与训练之后相比有不同吗?”

c. 是 ($p=0.002$)。

d. $d = (-4.2/4.212\ 23) = -1.00$ 或 1.00 , 效应量是 1.00 , 表明在接受 8 周的训练后打球准确度高于训练之前, 差值为 1 倍的标准差。效应量是大的。

e. 在接受了运动心理学家的训练后校园网球运动员的准确度分数 ($M=89.27$, $SD=5.78$) 比起之前 ($M=85.07$, $SD=6.46$) 明显更高, $t(14) = -3.86$, $p < 0.05$, $d=1.00$ 。

第8章 一维组间方差分析

1. a. $H_0: \mu_{\text{药物A}} = \mu_{\text{药物B}} = \mu_{\text{安慰剂}}$

H_1 : 至少其中一个均值与其他不同

b. “偏头痛患者的疼痛指数与使用的药物类型有关系吗?”

c. 方差没有显著不同, $F(2, 18) = 1.27$, $p > 0.05$ 。

d. 是 ($p=0.002$)。

e. $\eta^2 = \frac{46.205}{93.057} = 0.50$ 。效应量较大。

f. 药物 A 和药物 B 对应的疼痛指数小于安慰剂。药物 A 与药物 B 的效果没有明显的差异。

g. 偏头痛患者的疼痛指数因使用的药物类型不同而不同, $F(2, 18) = 8.88$, $p < 0.05$, $\eta^2 = 0.50$ 。Tukey 的事后方法显示, 吃下药物 A 的人 ($M=4.64$, $SD=1.61$) 和吃下药物 B 的人 ($M=4.60$, $SD=1.87$) 报告的疼痛明显低于那些吃下安慰剂的人 ($M=7.76$, $SD=1.30$)。药物 A 和药物 B 的疼痛水平没有显著不同。

2. a. $H_0: \mu_{\text{摩托车}} = \mu_{\text{混合动力汽车}} = \mu_{\text{非混合动力汽车}}$

H_1 : 至少其中一个均值与其他不同

b. “不同类型交通工具的使用者对石油价格的受挫程度不同吗?”

c. 方差没有显著不同, $F(2, 27) = 0.05, p > 0.05$ 。

d. 是 ($p < 0.001$)。

e. $\eta^2 = \frac{101.267}{169.367} = 0.60$ 。效应量较大。

f. 混合驱动汽车使用者的受挫程度低于摩托车使用者和非混合驱动汽车使用者。摩托车使用者的受挫程度低于非混合驱动汽车使用者。

g. 不同的交通工具使用者的受挫程度不同, $F(2, 27) = 20.08, p < 0.05, \eta^2 = 0.60$ 。Tukey 的事后方法显示, 混合驱动汽车使用者 ($M = 3.50, SD = 1.43$) 和摩托车使用者 ($M = 5.80, SD = 1.69$) 报告的受挫程度显著低于非混合驱动汽车使用者 ($M = 8.00, SD = 1.63$)。混合驱动汽车使用者报告的受挫程度显著低于摩托车使用者。

3. a. $H_0: \mu_{\text{公司A}} = \mu_{\text{公司B}} = \mu_{\text{公司C}}$

H_1 : 至少其中一个均值与其他不同。

b. “对紧急呼叫的反应时间因公司不同而不同吗?”

c. 方差没有显著不同, $F(2, 42) = 0.001, p > 0.05$ 。

d. 否 ($p = 0.974$)。

e. $\eta^2 = \frac{132.844}{104933.2} = 0.001$ 。效应量几乎为 0。

f. ANOVA 不显著。不需要 Tukey 检验。

g. 不同公司对紧急呼叫的反应时间没有显著不同, $F(2, 42) = 0.03, p > 0.05, \eta^2 = 0.001$ 。

第 9 章 二维组间方差分析

1. a. $H_0: \mu_{\text{男性}} = \mu_{\text{女性}}; H_1: \mu_{\text{男性}} \neq \mu_{\text{女性}}$

$H_0: \mu_{\text{使用手机}} = \mu_{\text{不使用手机}}; H_1: \mu_{\text{使用手机}} \neq \mu_{\text{不使用手机}}$

H_0 : 不存在性别×手机使用的交互效应

H_1 : 存在性别×手机使用的交互效应

b. “男性和女性的驾驶表现不同吗?”

“开车时使用手机的人和不使用手机的人的驾驶表现不同吗?”

“性别和使用手机之间有交互效应吗?”

c. 方差没有显著不同, $F(3, 20) = 2.47, p > 0.05$ 。

d. 性别的效应不显著 ($p=0.584$); 手机的效应显著 ($p<0.001$); 性别×手机使用的交互效应不显著 ($p=0.258$)。

e.

Test	Partial eta-square
Gender	0.02
Cell phone	0.76
Gender×cell phone	0.06

手机有最大效应。

f. 以驾驶表现为因变量、性别和手机使用为自变量执行 2×2 组间方差分析。手机效应显著, $F(1, 20)=61.63, p<0.05$, 偏 $\eta^2=0.76$, 使用手机的人 ($M=31.58, SD=4.42$) 驾驶技术明显比不使用手机的人 ($M=44.50, SD=3.53$) 表现差。性别效应不显著, $F(1, 20)=0.31, p>0.05$, 偏 $\eta^2=0.02$, 交互效应也不显著, $F(1, 20)=1.36, p>0.05$, 偏 $\eta^2=0.06$ 。

2. a. $H_0: \mu_{\text{方法A}} = \mu_{\text{方法B}}; H_1: \mu_{\text{方法A}} \neq \mu_{\text{方法B}};$

$H_0: \mu_{\text{自然科学}} = \mu_{\text{社会科学}}; H_1: \mu_{\text{自然科学}} \neq \mu_{\text{社会科学}}$

H_0 : 不存在方法×专业交互效应; H_1 : 存在方法×专业交互效应。

b. “不同教育方法下的数学焦虑有不同吗?”

“自然科学和社会科学专业对数学的焦虑不同吗?”

“教育方法和学校专业之间有交互效应吗?”

c. 方差没有显著不同, $F(3, 36)=2.26, p>0.05$ 。

d. 方法的主效应 ($p=0.002$)、专业的主效应 ($p=0.001$)、方法×专业交互效应 ($p=0.008$) 都显著。

e.

Test	Partial eta-square
Method	0.24
Major	0.27
Method×major	0.18

专业有最大效应。

f. 以数学焦虑为因变量、教育方法和学校专业为自变量执行 2×2 组间方差分析。教育方法的效应显著, $F(1, 36)=11.35, p<0.05$, 偏 $\eta^2=0.24$, 接受方法 B 指导的人对数学的焦虑 ($M=35.55, SD=8.44$) 明显低于那些接

受方法 A 的人 ($M=45.75$, $SD=14.29$)。专业的效应显著, $F(1, 36) = 13.20$, $p < 0.05$, 偏 $\eta^2 = 0.27$, 自然科学专业的人对数学的焦虑 ($M=35.15$, $SD=9.20$) 低于社会科学专业的人 ($M=46.15$, $SD=13.48$)。方法 \times 专业交互效应也显著, $F(1, 36) = 7.88$, $p < 0.05$, 偏 $\eta^2 = 0.18$ 。在接受方法 B 的人中, 自然科学专业的人 ($M=34.30$) 与社会科学专业的人 ($M=36.80$) 之间对数学焦虑的差异小。但是, 在接受方法 A 的人中, 社会科学专业的人 ($M=55.50$) 比自然科学专业的人 ($M=36.00$) 有更高程度的数学焦虑。

3. a. 全谷类饮食的 31~50 磅超重的组减掉最多的体重 (11.5 磅)。非全谷类饮食的 31~50 磅超重的组减掉最少的体重 (0.40 磅)。

b. 对于超重 11~30 磅的组, 全谷类饮食的人减掉的体重稍微多于非全谷类饮食的人 (多 2 磅), 但是, 对于超重 31~50 磅的组, 全谷类饮食的人减掉的体重多于非全谷类饮食的人 (9.4 磅的差距)。交互说明了全谷类饮食且超重 31~50 磅在减少体重上的组合效用, 该组比其他组明显减少更多的体重。全谷类饮食看起来对于那些超重的人 (31~50 磅) 尤其有效。

第 10 章 一维组内方差分析

1. a. $H_0: \mu_{\text{前}} = \mu_{\text{中}} = \mu_{\text{后}}$

H_1 : 至少其中一个均值与其他不同

b. “在进行预防逃课计划中逃课数随着时间变化吗?”

c. 是 ($p < 0.001$ — Greenhouse = Geisser; $p < 0.001$ — sphericity assumed)。

d. 偏 $\eta^2 = 0.52$ 。

e. 每个检验以水平 0.016 (0.05/3) 执行。

f. 拒绝 H_0 。逃课率随时间有显著不同, Greenhouse-Geisser adjusted $F(1.71, 23.98) = 15.20$, $p < 0.05$, 偏 $\eta^2 = 0.52$ 。相依样本 t 检验说明学生在 9 个月时 ($M=25.47$, $SD=8.18$) 逃课明显少于在 6 个月时 ($M=32.00$, $SD=7.46$), $t(14) = 3.92$, $p < 0.016$, 也少于在 3 个月时 ($M=33.00$, $SD=7.08$), $t(14) = 4.75$, $p < 0.016$ 。6 个月时的逃课数和 3 个月时的逃课数没有显著不同, $t(14) = 0.87$, $p > 0.016$ 。

(注:如果你更愿意用球形假定值,那么用2和28代替 Greenhouse-Geisser 的自由度 df (球形假定的 df),并删除该句“Greenhouse-Geisser adjusted”。)

2. a. $H_0: \mu_{\text{治疗组}} = \mu_{\text{对照组}} = \mu_{\text{八周}}$

H_1 : 至少其中一个均值与其他不同

b. “接受心理分析疗法时健康度随着时间变化吗?”

c. 否 ($p = 0.330$ — Greenhouse = Geisser; $p = 0.341$ — sphericity assumed)。

d. 偏 $\eta^2 = 0.07$ 。

e. 因为总体方差分析不显著,所以不能执行进一步检验。

f. 接受心理分析疗法的人的健康得分不随时间变化, Greenhouse-Geisser adjusted $F(1.50, 20.94) = 1.12$, $p > 0.05$, 偏 $\eta^2 = 0.07$ 。

(注:如果你更愿意用球形假定值,那么用2和28代替 Greenhouse-Geisser df (球形假定的 df),并删除该句“Greenhouse-Geisser adjusted”。)

3. a. $H_0: \mu_{\text{免费上网}} = \mu_{\text{付费上网}} = \mu_{\text{不能上网}}$

H_1 : 至少其中一个均值与其他不同

b. “在不同无线上网服务类型的快餐店吃饭的可能性不同吗?”

c. 是 ($p < 0.001$ — Greenhouse = Geisser; $p < 0.001$ — sphericity assumed)。

d. 偏 $\eta^2 = 0.70$ 。

e. 每个检验以水平 0.016 (0.05/3) 执行。

f. 在不同无线上网服务类型的快餐店吃饭的可能性不同, Greenhouse-Geisser adjusted $F(1.78, 16.00) = 20.53$, $p < 0.05$, 偏 $\eta^2 = 0.70$ 。相依样本 t 检验说明,比起提供花钱的无线上网服务的餐厅 ($M = 3.80$, $SD = 2.30$), $t(9) = 4.95$, $p < 0.016$, 比起没有无线上网服务的餐厅 ($M = 3.40$, $SD = 1.96$), $t(9) = 5.25$, $p < 0.016$, 人们更喜欢在提供免费无线上网服务的餐厅吃饭 ($M = 7.00$, $SD = 2.58$)。人们去提供花钱的无线上网服务的餐厅与没有无线上网服务的餐厅吃饭的可能性没有显著不同, $t(9) = 0.80$, $p > 0.016$ 。

(注:如果你更愿意用球形假定值,那么用2和18代替 Greenhouse-Geisser df (球形假定的 df),并删除该句“Greenhouse-Geisser adjusted”。)

第 11 章 组间组内方差分析

1. a. $H_0: \mu_{\text{认知行为}} = \mu_{\text{心理分析}}; H_1: \mu_{\text{认知行为}} \neq \mu_{\text{心理分析}}$
 $H_0: \mu_{\text{治疗前}} = \mu_{\text{8周}} = \mu_{\text{16周}}$
 H_1 : 至少其中一个均值与其他不同
 H_0 : 不存在时间 \times 治疗方法交互效应
 H_1 : 存在时间 \times 治疗方法交互效应
- b. “接受心理分析治疗和接受认知行为治疗的健康得分不同吗?”
“健康得分随时间变化而不同吗?”
“接受心理分析治疗和认知行为治疗的健康得分变化率不同吗?”
- c. 时间 ($p = 0.007 - \text{Greenhouse} = \text{Geisser}; p = 0.002 - \text{sphericity assumed}$) 和时间 \times 治疗方法 ($p < 0.001 - \text{Greenhouse} = \text{Geisser}; p = 0.001 - \text{sphericity assumed}$) 是显著的。治疗方法不是显著的 ($p = 0.388$)。
- d.

Test	Partial eta-square
Time	0.29
therapy	0.04
time \times therapy	0.49

时间 \times 治疗方法有最大效应。

- e. 执行组间组内方差分析，以不同时间 (before, 8 weeks, 16 weeks) 的健康为组内因素，以治疗类型 (cognitive-behavioral, psychoanalytic) 为组间因素。时间的效应显著，Greenhouse-Geisser adjusted $F(1.36, 24.44) = 7.46, p < 0.05$ ，偏 $\eta^2 = 0.29$ 。治疗方法的效应不显著， $F(1, 18) = 0.78, p > 0.05$ ，偏 $\eta^2 = 0.04$ 。时间 \times 治疗方法交互效应显著，Greenhouse-Geisser adjusted $F(1.36, 24.44) = 17.02, p < 0.05$ ，偏 $\eta^2 = 0.49$ 。交互点显示在治疗前和 8 周两个治疗组的健康得分没有不同，但是，在 16 周，进行认知行为治疗的人比起进行心理分析治疗的人有更高的健康得分。
- (注：如果你更愿意用球形假定值，那么用 2 和 36 代替 Greenhouse-Geisser df (球形假定的 df)，并删除该句 “Greenhouse-Geisser adjusted”。)
2. 简单效应分析的结果表明，在 16 周，进行认知行为治疗的人 ($M =$

24.00, $SD=3.71$) 的健康得分显著高于进行心理分析治疗的人 ($M=20.00$, $SD=2.94$) $t(18)=2.67$, $p<0.016$ (检验的 p 值约为 0.0156, 通过双击 Independent Samples Test 表格, 然后双击值 0.016 可以得到这个 p 值)。在治疗前两种治疗的健康得分没有显著不同 ($t(18)=-0.54$, $p>0.016$), 在 8 周两种治疗的健康得分也没有显著不同 ($t(18)=-0.25$, $p>0.016$)。

3. a. $H_0: \mu_{\text{前}} = \mu_{\text{中}} = \mu_{\text{后}}$

H_1 : 至少其中一个均值与其他不同

b. “有导师指导的老师在计划开始前、进行 4 周后、进行 8 周后的压力水平有不同吗?”

c. 是的, 时间是显著的 ($p=0.001$ —Greenhouse=Geisser; $p<0.001$ —sphericity assumed)。

d. 计划进行 8 周后的压力水平明显低于计划进行 4 周后和计划开始前的压力。计划开始前的压力和计划进行 4 周后没有显著不同。

e. 执行简单效应分析来评估是否三个时点老师的压力得分不同。一维组内方差分析的结果显示, 时间是显著的, Greenhouse-Geisser adjusted $F(1.17, 10.56)=18.39$, $p<0.05$, 偏 $\eta^2=0.67$ 。执行进一步相依样本 t 检验, 每个检验以 0.016 的 α 水平执行。结果指出, 与计划开始前, $t(9)=4.41$, $p<0.016$, 和计划进行 4 周后, $t(9)=4.46$, $p<0.016$ 相比, 计划进行 8 周后有导师指导的老师压力水平更低。在计划开始前和计划进行 4 周后的压力得分没有显著不同, $t(9)=2.49$, $p>0.016$ 。每个时间水平两组的均值和标准差在图表 C—1 中给出。

(注: 如果你更愿意用球形假定值, 那么用 2 和 18 代替 Greenhouse-Geisser df (球形假定的 df), 并删除该句 “Greenhouse-Geisser adjusted”。)

图表 C—1 在计划开始前、进行 4 周后、进行 8 周后老师的均值和标准差

Time	Mentored	
	<i>M</i>	<i>SD</i>
Before	41.00	3.30
4 weeks	39.60	4.09
8 weeks	34.70	6.00

第 12 章 皮尔逊 r 相关系数

1. a. $H_0: \rho=0$
 $H_1: \rho \neq 0$
b. “学习的时间量和考试成绩有关系吗?”
c. $r=0.68$ 。
d. 是 ($p<0.001$)。
e. 效应量较大 ($r=0.68$)。
f. 学习的时间量和考试成绩之间有显著正相关关系, $r(23)=0.68$, $p<0.05$ 。
2. a. $H_0: \rho=0$
 $H_1: \rho \neq 0$
b. “婚姻满意度和移情倾向有关系吗?”
c. $r=0.56$ 。
d. 是 ($p=0.004$)。
e. 效应量较大 ($r=0.56$)。
f. 婚姻满意度和移情倾向之间有显著正相关关系, $r(23)=0.56$, $p<0.05$ 。
3. a. $H_0: \rho=0$
 $H_1: \rho \neq 0$
b. “3 岁时听阅读的时间和英语二级考试的得分有关系吗?”
c. $r=0.52$
d. 是 ($p=0.003$)。
e. 效应量较大 ($r=0.52$)。
f. 3 岁时听阅读的时间量和英语二级考试的得分之间有显著正相关关系, $r(28)=0.52$, $p<0.05$ 。

第 13 章 简单线性回归

1. a. $H_0: \beta_{x_2} = 0$

$$H_1: \beta_{\text{父亲}} \neq 0$$

- b. “父亲的乐观度能否预测他的儿子在青年时的乐观度?”
 c. 是 ($p=0.001$)。
 d. $R^2=0.44$, 效应量较大。
 e. $\hat{Y}_{\text{儿子}}=16.313+0.571$ (父亲)。
 f. 父亲的乐观度分数可以显著预测他的儿子的乐观度水平, $\beta=0.66$, $t(18)=3.75$, $p<0.05$, $R^2=0.44$ 。

$$2. a. H_0: \beta_{\text{数学考试}} = 0$$

$$H_1: \beta_{\text{数学考试}} \neq 0$$

- b. “七年级数学入学考试的分数是否可以预测期末数学考试成绩?”
 c. 是 ($p<0.001$)。
 d. $R^2=0.38$, 效应量较大。
 e. $\hat{Y}_{\text{成绩}}=47.92+0.506$ (数学考试)。
 f. 入学考试可以显著预测七年级的最终数学成绩, $\beta=0.62$, $t(28)=4.17$, $p<0.05$, $R^2=0.38$ 。

$$3. a. H_0: \beta_{\text{易相处}} = 0$$

$$H_1: \beta_{\text{易相处}} \neq 0$$

- b. “工人的易相处性可以用来预测管理者对他们工作的满意度吗?”
 c. 否 ($p=0.093$)。
 d. $R^2=0.12$, 中等效应量。
 e. $\hat{Y}_{\text{满意度}}=5.332+0.097$ (易相处)。
 f. 工人的易相处性不能显著预测管理者对他们工作的满意度, $\beta=0.34$, $t(23)=1.75$, $p>0.05$, $R^2=0.12$ 。

第 14 章 多元线性回归

$$1. a. H_0: \beta_{\text{importance}}=0; H_0: \beta_{\text{advance}}=0; H_0: \beta_{\text{express}}=0$$

$$H_1: \beta_{\text{importance}} \neq 0; H_1: \beta_{\text{advance}} \neq 0; H_1: \beta_{\text{express}} \neq 0$$

$$b. H_0: R^2=0$$

$$H_1: R^2 \neq 0$$

c. 单个预测变量:

“工作重要性可以预测工作满意度吗?”

“在公司晋升的机会可以预测工作满意度吗?”

“向老板表达观点的能力可以预测工作满意度吗?”

所有预测变量:

“当所有变量一起考虑时,工作重要性、在公司晋升的机会和向老板表达观点的能力可以预测工作满意度吗?”

d. $R^2=0.57$ 。是的,整个回归模型是显著的, $p<0.001$ 。

e. 变量 importance ($p=0.001$) 和 express ($p=0.018$) 是显著的。advance ($p=0.403$) 是不显著的。

f. $R^2=0.57$, 效应量较大。

g. $\hat{Y}_{\text{工作满意度}} = -5.756 + 1.051(\text{importance}) + 0.205(\text{advance}) + 1.069(\text{express})$ 。

h. 执行多元线性回归,通过变量工作重要性、在公司晋升的机会和向老板表达观点的能力预测工作满意度。总的说来,回归是显著的, $F(3, 26) = 11.51$, $p<0.05$, $R^2=0.57$ 。在调查的预测变量中,工作重要性是工作满意度的显著预测变量, $\beta=0.52$, $t(26) = 3.65$, $p<0.05$, 同样向老板表达观点的能力也是显著的 $\beta=0.37$, $t(26) = 2.53$, $p<0.05$ 。在公司晋升的机会不是工作满意度的显著预测变量, $\beta=0.11$, $t(26) = 0.85$, $p>0.05$ 。

2. a. $H_0: \beta_{\text{宽恕}} = 0$; $H_0: \beta_{\text{社会支持}} = 0$

$H_1: \beta_{\text{宽恕}} \neq 0$; $H_1: \beta_{\text{社会支持}} \neq 0$

b. $H_0: R^2 = 0$

$H_1: R^2 \neq 0$

c. 单个预测变量:

“宽恕可以预测幸福指数吗?”

“社会支持可以预测幸福指数吗?”

所有预测变量:

“当所有变量一起考虑时,宽恕和社会支持可以预测幸福指数吗?”

d. $R^2=0.46$ 。是的,整个回归模型是显著的, $p=0.001$ 。

e. 宽恕 ($p<0.001$) 是显著的,社会支持 ($p=0.246$) 是不显著的。

f. $R^2=0.46$, 效应量较大。

g. $\hat{Y}_{\text{幸福}} = 5.798 + 0.399 (\text{宽恕}) + 0.299 (\text{社会支持})$ 。

h. 执行多元线性回归, 通过变量宽恕和社会支持预测幸福指数。总的说来, 回归是显著的, $F(2, 22) = 9.32, p < 0.05, R^2 = 0.46$ 。在调查的预测者中, 宽恕是显著的, $\beta = 0.66, t(22) = 4.22, p < 0.05$, 社会支持不是显著的, $\beta = 0.19, t(22) = 1.19, p > 0.05$ 。

3. a. $H_0: \beta_{\text{幸福}} = 0; H_0: \beta_{\text{生活意义}} = 0$

$H_1: \beta_{\text{幸福}} \neq 0; H_1: \beta_{\text{生活意义}} \neq 0$

b. $H_0: R^2 = 0$

$H_1: R^2 \neq 0$

c. 单个预测变量:

“幸福感可以预测对死亡的恐惧指数吗?”

“生活意义可以预测对死亡的恐惧指数吗?”

所有预测变量:

“当所有变量一起考虑时, 幸福感和生活意义可以预测对死亡的恐惧指数吗?”

d. $R^2 = 0.13$ 。是的, 整个回归模型是显著的, $p < 0.05$ 。

e. 生活意义 ($p = 0.010$) 是显著的。幸福感 ($p = 0.176$) 是不显著的。

f. $R^2 = 0.13$, 中等效应量。

g. $\hat{Y}_{\text{对死亡的恐惧}} = 36.434 + 0.290 (\text{幸福感}) - 0.349 (\text{生活意义})$ 。

h. 执行多元线性回归, 通过变量幸福感和生活意义预测对死亡的恐惧指数。总的说来, 回归是显著的, $F(2, 47) = 3.60, p < 0.05, R^2 = 0.13$ 。在调查的预测变量中, 生活意义 (meaning) 是显著的, $\beta = -0.41, t(47) = -2.68, p < 0.05$, 但是幸福感 (wellbeing) 不是显著的, $\beta = 0.21, t(47) = 1.38, p > 0.05$ 。

第 15 章 χ^2 拟合优度检验

1. a. H_0 : 眼睛大小和吸引力没有关系 (没有偏爱其中一张照片)

H_1 : 眼睛大小和吸引力有关系 (偏爱其中一张照片)

b. “眼睛大小和吸引力有关系吗?”

c. 是 ($p < 0.001$)。

d. 对大眼睛的照片有明显偏爱, $\chi^2 (1, N=80) = 20.0, p < 0.05$ 。对于两个类型的照片, 75%的参与者认为有大眼睛的人更有吸引力, 25%的参与者认为小眼睛的人更有吸引力。

2. a. H_0 : 人们对等离子电视和 LCD 电视在画面质量感受上没有不同

H_1 : 人们对等离子电视和 LCD 电视在画面质量感受上有不同

b. “人们对等离子电视和 LCD 电视的画面质量感受有不同吗?”

c. 没有 ($p=0.535$)。

d. 人们对等离子电视和 LCD 电视的画面质量感受没有不同, $\chi^2 (1, N=65) = 0.39, p > 0.05$ 。

3. a. H_0 : 对于三种咖啡没有偏爱

H_1 : 对于三种咖啡有偏爱

b. “在三种咖啡中有被偏爱的吗?”

c. 是 ($p=0.002$)。

d. 在三种咖啡中有偏爱的, $\chi^2 (2, N=150) = 12.16, p < 0.05$ 。对于三种不同价格的咖啡, 20%的参与者更喜欢 3 美元的咖啡, 41%的参与者更喜欢 6 美元的咖啡, 39%的参与者更喜欢 10 美元的咖啡。与 3 美元的咖啡相比, 6 美元和 10 美元的咖啡被认为味道更好。

第 16 章 χ^2 独立性检验

1. a. H_0 : 婴儿时吸收的营养种类与他一年级时的体重没有关系

H_1 : 婴儿时吸收的营养种类与他一年级时的体重有关系

b. “婴儿时吸收的营养种类与他一年级时的体重有关系吗?”

c. 是 ($p=0.022$)。

d. 效应量 = -0.13 (或 0.13), 效应量较小。

e. 一个孩子婴儿时吸收的营养种类与他一年级时是否超重有显著关系, $\chi^2 (1, N=300) = 5.26, p < 0.05$, Cramer's $V=0.13$ 。母乳喂养的孩子比不是母乳喂养的孩子 (28%的超重率) 有明显更低的超重率 (16%的超重率)。

2. a. H_0 : 性别与电影爱好之间没有关系

H_1 : 性别与电影爱好之间有关系

- b. “性别与电影爱好之间有关系吗?”
- c. 是 ($p < 0.001$)。
- d. 效应量 = -0.40 (或 0.40)，中等效应量。
- e. 性别与电影爱好之间有显著关系， $\chi^2(1, N=160) = 26.02$ ， $p < 0.05$ ，Cramer's $V = 0.40$ 。对动作电影和喜剧电影，女性更喜欢喜剧电影 (16%对 84%)，而男性更喜欢动作电影 (54%对 46%)。

3. a. H_0 : 老年人运动和患有心脏病之间没有关系



H_1 : 老年人运动和患有心脏病之间有关系

- b. “老年人运动和患有心脏病之间有关系吗?”
- c. 是 ($p = 0.029$)。
- d. 效应量 = -0.11 (或 0.11)，效应量较小。
- e. 老年人运动和患有心脏病之间有显著关系， $\chi^2(1, N=405) = 4.77$ ， $p < 0.05$ ，Cramer's $V = 0.11$ 。没有运动经历的人患心脏病的概率 (13.7%) 高于那些运动的人 (6.7%)。

(注：即使按照 Cohen 标准，效应量还是小的，注意到没有运动经历的人心脏病患病率是有运动经历的人的两倍 ($13.7/6.7 = 2.04$)，说明一个小的效应量有时有很大的实际重要性。)

注 释

第 1 章

1. 一个计算机图标就是你的计算机里的图画。例如，和.
2. “双击”鼠标意味着快速地连按左键两次；“单击”鼠标意味着按左键一次（指一次点击）；右击鼠标意味着按右键一次。
3. 对话框是一个允许做一些操作的窗口，比如选择选项、点击按钮、移动变量，等等。
4. 单词“数据（data）”是复数，指两个或更多的值。“数据（datum）”是单数，指一个值。
5. 默认值（default values）是 SPSS 中的一个设置，在 SPSS 程序启动时就是这样的状态。
6. SPSS 11.5 版本和早期的版本要求变量名不超过 8 个字符。如果你使用的是 SPSS 11.5 或者更早的版本，变量名表示为 **employ** 而不是 **employment**。
7. 如果你点击 Values 单元的右边（点击单元后省略键就会出现），Value Labels（变量值标签）对话框就会立即打开（仅用一步就可以打开对话框，而不是两步）。
8. 如果你喜欢一个个地移动变量，选择 **gender**，点击向右箭头按钮（▶），然后选择 **age**，再点击向右箭头按钮（▶），继续这个过程直到所有的变量移到 Variables 框。

第 2 章

1. 两个变量可以一个个地移动，选择 **gender**，点击向右箭头按钮（▶），然后选择 **college**，点击向右箭头按钮（▶）。
2. 集中趋势和变异性的度量也可以使用 SPSS 中的 Descriptives 程序得到。为了运行 Descriptives 程序，从菜单栏中选择 **Analyze>Descriptives Statistics>Descriptives**，然后把变量移到 Variable(s) 框中。在 Descriptives 程序中，均值、标准差、极小值和极大值是默认选择的。想选择其他的统计量可以点击“Options”按钮。

3. 超出范围的值是在变量极小值和极大值之外的一个数值（例如，50 就是 **satquant** 的一个超出范围的数值）。没有超出范围的值并不能确保没有数据输入错误（在数据的范围内也会有错误），简单的检查可以避免常见的数据输入时打字错误。为了避免数据输入错误，小心输入数据，至少要复查所有输入的数据。

第 4 章

1. 评判间信度 (interrater reliability) 是可靠性的一种，是没有落到这些类中的一个。评判间信度度量了一些感兴趣特征的估价人或裁判的一致性。

2. 系数 α 可能是负的，这典型地表明计算时产生了误差，例如包括了一个或更多在测量时没有逆向编码的负项（负向题目稍后讨论）。

3. M 和 SD 分别是均值和标准差的简写。使用了美国心理协会 (APA) 的格式。

第 5 章

1. 本例中另一个效应的度量方法是简单地表明顶尖会计师事务所中员工的工作小时数与全国平均水平工作小时数的差异（顶尖公司的员工平均每周多工作 7 小时）。如果因变量直觉上是有意义的，这是一个合理的练习（尽管 d 仍然也被报告）。然而，当因变量没有直观意义时（比如用 10~50 的尺度度量自信），那么一个标准化的度量比如 d 就更好。

第 6 章

1. 另一种 d 的方法是将两个组的均值差除以合并的标准差。尽管这个公式直觉上更吸引人，但是这里没有提供该公式，因为 SPSS 的输出结果中不提供合并的标准差。

2. 研究者的责任是确保遵守美国心理协会的指南，人道地对待研究的参与者。

第 7 章

1. 本书使用的术语自变量并不意味着这个变量被研究者控制，但是有助于识别每个统计程序的特点。不论该变量是自变量、准实验的变量还是非实验的变量，都不会影响 t 检验的结果，然而，它将影响从研究中得出的结论的类型。

第 8 章

1. “因素”是 ANOVA 中自变量的另一个名称。“水平”是类别或组别的另一个名称。例如，变量性别有两个水平：男性和女性。

2. 当组间方差不相等时, Brown-Forsythe 和 Welch (见图表 8—7) 是 ANOVA 的备择检验。如果 Levene 检验表明方差不相等, 那么这其中的任一个检验就可以替代标准的一维方差分析 (很快将讨论 Levene 检验和方差相等的假定)。

3. 如果 Levene 检验表明方差不相等, 在 Equal Variances Not Assumed 下使用事后检验 (代替 Tukey 检验)。

4. Tukey 事后方法有一个内置的控制使得组合的三个检验的整体 α (第 1 类错误率) 不超过 0.05。在第 10 章和第 11 章将介绍更多关于控制进一步检验的整体 α 的内容。

5. 在 SPSS 中能够使用 General Linear Model——Univariate 方法计算 η^2 。在第 9 章中将介绍 General Linear Model 程序 (使用它计算 η^2)

第 9 章

1. 固定因素是一个自变量, 其中水平是特别为研究选择的, 没有故意一般化为其他可能的水平。此外, 随机因素包含的一个自变量的水平是随机从较大的潜在水平中选择的, 有意概括了水平更大的总体结果。在大部分的研究中, 通常固定因素比随机效应更常见, 尽管随机因素是在某些特定的情况下有规律出现的。

2. 在 Tests of Between-Subjects Effects 表中, 修正的总平方和 (corrected total) 等于 **phyther**, **relax**, **phyther**×**relax** 和误差的平方和。修正的模型平方和 (corrected model) 等于 **phyther**, **relax** 和 **phyther**×**relax** 的平方和。通常在二维 ANOVA 中对截距 (intercept) 和总计 (total) 不感兴趣。

3. 显著的交互效应的存在并不必定意味着主效应, 但是表明它们可能被误解。因此, 如果它们可以被解释, 应该根据交互来完成。

第 10 章

1. 组内因素通常包括同样的人在不同时刻的测量, 也可以包括相关的人 (如丈夫和妻子) 同时被测量。

2. 在 SPSS 中输入 **before**, **week8** 和 **affter** 作为 variables, 在一维组内方差分析中是组内因素 **time** 的水平。

3. 如果组内因素仅有两个水平, 可以不需要球形假定。

4. 这里不能评价球形的 Mauchley 检验, 评价方法如下: 原假设是数据总体是球形的; 如果 Mauchley 检验的 $p \leq 0.05$, 拒绝原假设, 球形假设不满足。

如果 $p < 0.05$, 原假设不能被拒绝, 球形假设得到满足。

5. 在 Tests of Within-Subjects Effects 表中的四个程序得到了同样的 F 值, 它们常常有不同的自由度 (以及不同的 p 值)。

第 11 章

1. 大多数组内因素包括了同一个人在不同时刻的多次测量, 也可以包括相关人一次测量结果 (例如, 同辈)。

2. 因为对立假设没有特别指定, 如果 **time** 是显著的, 将要求进一步检验来评价差异在哪里。把这与仅有两个组的 **support** 的假设进行比较: 如果 **support** 显著, 这两个组的均值仅需要检查以确定哪一组的压力较小 (不需要进一步的检验)。

3. 为了这一点, 我们在 SPSS 中没有为组内因素创建名称 (我们已经为组内因素——**before**、**week4** 和 **week8**——的不同水平创建了名称)。

4. Multivariate Tests 表检验了组内因素 (**time**) 和任一个包含了组内因素的交互项。该表没有检验组间因素 (它们自己)。

5. 当组内因素只有两个水平时, 不需要球形假定。

6. 因为球形假定没有应用到组间因素, 对这个检验没有不同的调整方法。

7. 例如, 有导师对没有导师可能仅在 **time3** 显著差异或者在 **time2** 和 **time3** 上显著差异。

第 12 章

1. 尽管大部分的相关是对两个连续变量计算得到的, 皮尔逊相关系数也可以对一个分类变量和连续变量进行计算 (称为列相关) 或者两个分类变量 (称为 phi 相关) 进行计算。

第 13 章

1. 线性回归与相关 (已经在第 12 章讨论了相关) 有关系; 事实上, 在线性回归中为了使得一个变量是另一个变量的显著预测变量, 要求两个变量之间有显著相关。然而, 两个方法的目标有些差异。相关描述了两个变量之间的关系, 而简单线性回归的目标是为了评价一个变量是不是另一个变量的显著预测变量。

2. 尽管在简单回归中通常使用连续预测变量, 但是预测变量也可能是分类变量, 如果它仅仅包含了两个值 (它是二分变量)。

3. 尽管在简单回归中, ANOVA 和 Coefficients 表提供了相同的信息, 但

是在多元回归中，它们并不相同，将在第 14 章进行说明。

4. 在 SPSS 中通过在 Linear Regression 对话框中点击“Save”（见图表 13—6），再点击对话框中的“Unstandardized in the Predicted Values”就可以得到预测得分。预测得分将出现在数据文件中的变量 **wellbeing** 的右边。

5. 当回归系数是标准化时，Y 轴截距为 0，并且将它从 Coefficients 表中删除了（这就是为什么在 Standardized Coefficients 列中没有报告 Constant 值的原因）。

第 14 章

1. 多元回归是一个相当复杂的主题，这里只做一些入门介绍。Stevens (2002) 是一个较好的资源，可以从中获得较多的多元回归信息。

2. 有两个水平的分类变量（例如，性别）可以直接输入 SPSS 作为一个预测变量；有三个或更多水平的分类变量在输入回归方程之前必须重新编码成为多元预测变量（预测变量数等于类别数-1）。参见 Cohen, Cohen, West and Aiken (2003) 以获得更多关于三个或更多水平的分类变量进行编码的信息。

3. 预测变量之间的高相关性可能导致一个称之为多重共线性的问题，这会导致对回归方程中预测变量的不稳定的估计。参见 Stevens (2002) 获得更多关于多重共线性的信息。

4. 在一个分层回归中，预测变量的一个子集首先进入模型，然后运行回归程序；预测变量的另一个子集再添加到模型，然后再运行回归程序，等等。分层回归的一个主要目的是看后面添加的预测变量解释显著的变异性数量是否大于较早前添加的预测变量。

5. 在 SPSS 中通过在 Linear Regression 对话框中点击“Save”（见图表 14—6），再点击对话框中的“Unstandardized in the Predicted Values”就可以得到得分值。预测得分将出现在数据文件中的变量 **success** 的右边。

6. 当回归系数是标准化时，Y 轴截距为 0，并且将它从 Coefficients 表中删除了（这就是为什么在 Standardized Coefficients 列中没有报告 Constant 值的原因）。

7. 如果模型中包含了一个或更多个分类变量，对于中等到较大的不相等的样本量，同方差性的存在能够折中多元回归方法的准确性。

第 15 章

1. 严格说，卡方统计量的值也受到期望频数的绝对大小的影响。
2. 当数据输入使用案例加权方法时，如果在执行卡方程序之前，没有加权，那么结果将不正确。
3. 要注意一件重要的事情，保持加权直到案例加权的选项关掉（在 Weight Cases 对话框里选择 Do not weight cases，点击“OK”），或者直到 SPSS 结束。如果除了卡方检验还要做其他的分析，未关闭 weight cases 选项可能导致不正确的分析或者得到错误的信息。
4. 每个类别的人的百分比可以在 SPSS 中对 **groupsize** 使用 Frequencies 程序计算得到（假设频数已被加权）。在 SPSS 中使用 Frequencies 程序的更多信息见第 2 章。

第 16 章

1. 要注意一件重要的事情，保持加权直到案例加权的选项关掉（在 Weight Cases 对话框里选择 Do not weight cases，点击“OK”），或者直到 SPSS 结束。如果除了卡方检验还要做其他的分析，未关闭 weight cases 选项可能导致不正确的分析或者得到错误的信息。

附录 A

1. 尽管重新编码步骤 3→3 看起来似乎不必要，但是如果省略了，那么在数据集中将丢失所有 3 的值。
2. 只要所有的编码变量有相同的尺度值（例如，1、2、3、4 和 5），那么一次就可以重新编码很多变量。为了增加其他变量，把它们移到 Numeric Variable→Output Variable 框（见图表 A—5），给编码变量一个合适的名称。例如，如果有第 4 项，**meaning4** 存在且是负向题目，把 **meaning4** 移动到 Numeric Variable→Output Variable 框，在 Output Variable Name 框中输入变量名 **meaning4 _ recode**。因为已经给 **meaning2 _ recode** 输入了编码值，所以不需要再次输入。
3. 因为 **meaning2** 是负向题目，在计算总得分时使用 **meaning2 _ recode**。
4. 一起增加变量的另一个方法就是在 SPSS 中使用和函数。SPSS 中内置的函数命令可以对变量执行一些操作。
5. 尽管使用 *t* 检验说明了 Select Cases 程序，但是仅有 5 个参与者的样本量不可能是显著的（因为功效太低）。

参考文献

- American Psychological Association (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Green, S. B., & Salkind, N. J. (2005). *Using SPSS for Windows and Macintosh: Analyzing and understanding data*. Upper Saddle River, NJ: Prentice Hall.
- Howell, D. C. (2007). *Statistical methods for psychology* (6th ed.). Belmont, CA: Wadsworth.
- Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn and Bacon.

鄧平舟
解慶
PDG

中国人民大学出版社读者信息反馈表

尊敬的读者：

感谢您购买和使用中国人民大学出版社的_____一书，我们希望通过这张小小的反馈表来获得您更多的建议和意见，以改进我们的工作，加强我们双方的沟通和联系。我们期待着能为更多的读者提供更多的好书。

请您填妥下表后，寄回或传真回复我们，对我们的支持我们不胜感激！

1. 您是从何种途径得知本书的：

☐ 书店 ☐ 网上 ☐ 报刊 ☐ 朋友推荐

2. 您为什么决定购买本书：

☐ 工作需要 ☐ 学习参考 ☐ 对本书主题感兴趣

☐ 随便翻翻

3. 您对本书内容的评价是：

☐ 很好 ☐ 好 ☐ 一般 ☐ 差 ☐ 很差

4. 您在阅读本书的过程中有没有发现明显的专业及编校错误，如果有，它们是：_____

5. 您对哪些专业的图书信息比较感兴趣：_____

6. 如果方便，请提供您的个人信息，以便于我们和您联系（您的个人资料我们将严格保密）：

您供职的单位：_____

您教授的课程（教师填写）：_____

您的通信地址：_____

您的电子邮箱：_____

请联系我们：

电话：62515732 82501704

传真：62514775

E-mail: rdcbsjg@crup.com.cn http: //www. rdjg. com. cn

通讯地址：北京市海淀区中关村大街甲 59 号文化大厦 15 层 100872

中国人民大学出版社工商管理出版分社