

SPSS

<http://www.phēi.com.cn>

(第3版)

统计应用与解析

SPSS for Intermediate Statistics
Use and Interpretation (Third Edition)

■ Nancy L. Leech

[美] Karen C. Barrett 著

■ George A. Morgan

■ 何丽娟 朱红兵 译

■ 卢纹岱 审校



含配书盘



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

SPSS 统计应用与解析

(第3版)

SPSS for Intermediate Statistics
Use and Interpretation (Third Edition)

Nancy L. Leech
[美] Karen C. Barrett 著
George A. Morgan

何丽娟 朱红兵 译
卢纹岱 审校

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING



内 容 简 介

本书通过对中级统计学课程所包括的各种统计学方法计算和演示,帮助读者学会使用 SPSS 分析和解释研究数据。本书虽然基于 SPSS 15.0 编写,但是同样适用于 SPSS 早期版本和 SPSS 将来的新版本。主要内容包括:数据编码和探索分析、可靠性测度方法、探索性因子分析和主成分分析、选择和解释推断统计、多元回归、逻辑斯谛回归和判别分析、多因素方差分析和协方差分析、重复测量和混合设计的方差分析、多元方差分析和典型相关、多水平线性模型/分层线性模型。每章都介绍几种相关的统计方法,并提供了如何运行这些方法和解释输出结果的指导。书中所有例题数据均保存在本书光盘中,同时本书可向采用本教材的教师免费提供教学支持和习题解答。

本书可作为本科高年级学生或研究生的中高级统计学课程的补充教材,也可供从事心理学、教育学、医疗卫生保健等领域统计工作者学习参考。

SPSS for Intermediate Statistics: Use and Interpretation, Third Edition

Nancy L. Leech, Karen C. Barrett, George A. Morgan

ISBN: 1-58488-508-4

Copyright © 2008 by Taylor & Francis Group, LLC

Authorized translation from the English language edition published by Lawrence Erlbaum Associates, Inc., part of Taylor & Francis Group LLC. All rights reserved.

本书英文版由 Taylor & Francis Group 出版集团旗下的 Lawrence Erlbaum Associates, Inc. 出版,并经其授权翻译出版,版权所有,侵权必究。

Publishing House of Electronics Industry is authorized to publish and distribute exclusively the Chinese (Simplified Characters) language edition. This edition is authorized for sale throughout Mainland of China. No part of the publication may be reproduced or distributed by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

本书中文简体版专有出版权由 Taylor & Francis Group, LLC 授予电子工业出版社,并限在中国大陆出版发行。未经出版者书面许可,不得以任何方式复制或发行本书的任何部分。

本书封底贴有 Taylor & Francis 公司防伪标签,无标签者不得销售。

版权贸易合同登记号 图字:01-2009-0643

图书在版编目(CIP)数据

SPSS 统计应用与解析:第3版/(美)利琪(Leech, N. L.), (美)巴瑞特(Barrett, K. C.), (美)摩根(Morgan, G. A.) 著;何丽娟、朱红兵译.—北京:电子工业出版社,2009.6

书名原文:SPSS for Intermediate Statistics Use and Interpretation, Third Edition

ISBN 978-7-121-08994-7

I. S… II. ①利… ②巴… ③摩… ④何… III. 统计分析—软件包, SPSS IV. C819
中国版本图书馆 CIP 数据核字(2009)第 089243 号

责任编辑:杨丽娟

印 刷:北京市天竺颖华印刷厂

装 订:三河市鑫金马印装有限公司

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本:787×980 印张:18.5 字数:405 千字

印 次:2009 年 6 月第 1 次印刷

印 数:4000 册 定价:38.00 元(含光盘 1 张)

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

质量投诉请发邮件至 zltz@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

译者序

SPSS 统计软件是目前国内外科学研究中最常用的统计分析软件之一,它集数据整理、分析功能于一身,广泛应用于自然科学、社会科学各个领域。其统计分析功能强大,操作简便。读者只要略知计算机操作基础和统计学基础,就能轻松通过 SPSS 完成统计分析和计算。由 Nancy L. Leech, Karen C. Barrett 和 George A. Morgan 编写的这本《SPSS 统计应用与解析》(第 3 版)是一本内容丰富、易于理解和掌握的 SPSS 统计分析实用教材。

本书的三位作者长期从事统计分析和研究工作,有着丰富的教学经验,熟知统计学专业知识,精通 SPSS 应用与统计分析。在本书出版之前已经出版了多本这一领域的教学用书,本书则是他们最近的又一力作,内容深入浅出,通过对中级统计学课程所包括的各种统计学方法计算和演示,帮助读者学会使用 SPSS 分析和解释研究数据。

本书阐述的 SPSS 统计方法,不仅涵盖国内教材中常见的探索分析、可靠性分析、探索性因子分析、判别分析、方差分析等方法,还包括不常见的典型相关和分层线性模型分析方法,在第 1 章和第 5 章详细介绍了统计基本概念和统计方法的选择。为了便于读者学习,在附录中对 SPSS 的基本操作、SPSS 基本统计分析方法等加以简述并提供部分习题答案。

这本书最值得一提的是,作者在介绍每种方法时,都是通过具体的实例,先提出问题,然后进行 SPSS 统计分析计算,计算的操作步骤清晰,对于输出结果中重要的内容都用插入方框的形式加以提示和说明,并提供输出结果的详尽而清晰的解释和撰写统计结论的实用样例。书中对输出结果效应量和功效的解释也是国内其他教材中很少涉及的,每章的问答题、拓展练习以及光盘所附数据极大地方便读者学习参考。可以说,本书既可作为本科高年级学生或研究生中高级统计学课程一本不可多得的补充教材,也可担当心理学、教育学、医疗卫生保健等领域统计工作者的良师益友。

本书由何丽娟和朱红兵翻译,其中第 1~3 章,第 8~11 章以及附录由何丽娟翻译,第 4~7 章由朱红兵翻译,卢纹岱教授负责全书的审校工作,全书由何丽娟负责统稿。在翻译过程中,得到了裴东光、茹秀英、白毅红、徐守森、刘坤、范娟、陈莹、殷小川、王勤、杨一虹、韩桂凤、解利辉、丁进国、费青松、李征、千青海等老师的帮助和支持,在此一并表示感谢。

由于译者水平有限,错误、疏漏之处在所难免,恳请广大读者批评指正。

译者

2009 年 5 月于学院路

前 言

本书通过对中级统计学知识的介绍来帮助读者学习分析和解释研究数据的方法。它可以作为行为科学或教育专业中级统计学课程的辅助教材，是对主干统计课程教材的补充。通过 SPSS 软件的学习，读者可轻松掌握其操作，从而在使用 SPSS 进行统计分析时变得简单易行，而不必通过正式的、专业性的计算实验室机构完成。当然，你需要安装 SPSS 及相关的 Windows 系统。虽然 SPSS 操作简单，但是由于它里面有覆盖范围广和强大的选项和统计功能，在使用中如何选择选项和解释输出结果还是非常困难的，本书则试图帮助大家解决这些难点。

1. SPSS 15.0 及早期版本

本书基于 SPSS 15.0 编写，除了图表内容外，其他内容与早期的 SPSS 10.0 ~ SPSS 14.0 版本相比只有小部分变化。实际上，本书所演示的过程，SPSS 7.0 ~ SPSS 15.0 版本也只是有细微不同。我们希望将来的 Windows 版本都不会有太大的变化。学生在学习这本书或早期版本的教材时，可以使用任何一个版本的 SPSS 软件，大部分的操作过程和输出结果都是非常相似的。

本书目标

本书将展示中级统计学的各种统计方法，并介绍如何使用这些方法进行研究工作。我们的目标是尽可能通过用非技术性的，且是专业表述描述各种统计方法的使用和解释。

本书的主要宗旨是：帮助读者选择适当的统计方法，合理解释输出结果，提高撰写统计结论的技能。其主要内容如下：

- 1) 如何选择适当的统计方法是研究设计的基础。
- 2) 如何使用 SPSS 回答所研究的问题。
- 3) 如何解释 SPSS 的输出结果。
- 4) 如何撰写研究论文的统计结论。

以上这些内容可提高研究问题的各个方面的技能，如研究设计、数据收集、数据整理、分析数据、解释输出结果以及撰写统计结论等。本书中所用的 HSB 数据实例也许与你所要做的论文、研究报告或研究设计非常相似。因此，我们认为该书很有可能对你的统计分析有帮助。本书附带的光盘里包含 HSB 数据文件和章后的 SPSS 拓展练习数据文件（当然，你自

己需要安装或购买 SPSS 软件)。本书实用性、可读性强, 书中没有列举太多的参考书目; 只是提供我们认为有用的一些书目和文章供读者参考。本书非常适合作为大多数的学生课堂学习的参考教材。

我们的系列教材, 由 Morgan, Leech, Gloeckner 和 Barrett 编写的《SPSS for Introductory Statistics: Use and Interpretation》(2007, 第 3 版) 一书也列在本书的参考书目中, 它是由劳伦斯·厄伦美厄协会出版发行的, 该书提供了基础统计学中关于如何解释和撰写统计分析结论的更广泛的讨论, 包括诸如 t 检验、 χ^2 检验和相关系数等内容。本书后的附录 B 中提供了有关基础统计学的主要内容回顾。

2. 本书特色

本书的特色主要表现在以下几个方面:

- 1) 能够清晰看到在进行统计分析时 SPSS 视窗的关键步骤, 有助于读者的“直观学习”。
- 2) 我们所做分析的全部输出, 可能就是你想要做的, 通过 SPSS 对输出结果适当的编辑后, 使输出效果更好地呈现出来。
- 3) 在插入的矩形框中详细解释 SPSS 输出中重点关注的内容, 帮助理解输出结果。
- 4) 特别增加了图表流程帮助选择适当的推断统计方法, 以及你如何解释统计显著性和效应量 (在第 5 章)。第 5 章中还扩充了例题, 提供如何阐述和撰写研究问题, 并列出一些研究问题以及对 t 检验和双变量回归的统计结论撰写的方法。
- 5) 本书的第 3、4 章和第 6~11 章, 通过举例介绍怎样撰写输出的统计结论以及毕业论文、长篇论文和学术论文中统计表格的制作。
- 6) 每章最后安排的问答题, 有助于掌握这一章的相关内容。
- 7) 除第 1 章和第 5 章外, 其余各章最后的 SPSS 拓展练习帮助巩固所学知识。
- 8) 附录 A 提供有关 SPSS 入门知识和其他有用的操作命令。
- 9) 附录 B 是对基础统计学的简要回顾。
- 10) 附录 C 提供对每章问答题奇数题号的答案。
- 11) 本书附带的光盘中是一些数据集, 这些真实数据便于学习和掌握每章内容、分析问题和解决难题。在光盘中还包括两个 pdf 文件, 提供了使用 SPSS 过程的快速向导指南 (QRG), 以及在 APA 格式中如何制作表格的制作图表的文件。

3. 章节内容的安排

撰写本书的整体思路是整本书中都贯穿使用真实的数据, 如 HSB 数据集, 介绍如何使用 SPSS 和如何分析结果。在开始使用 SPSS 前, 有两章的内容是基本知识的介绍。第 1 章是

研究设计介绍和回顾, 以及如何应用 HSB 数据集进行分析, 并简述测度水平和描述统计的内容。第 2 章讨论有关数据编码规则、探索性分析 (EDA) 以及研究假设的相关内容, 其中大部分内容为是为准备所要陈述的研究问题而应做的预分析。

第 3、4 章主要介绍评估数据的可靠性 (信度) 和效度分析的方法。其中, 第 3 章主要介绍如何计算克隆巴赫 α 系数、测试 - 再测试和观测者间的可靠性, 第 4 章讲述评估效度的方法, 包括探索性因子分析和主成分分析。

第 5 章提供研究设计的主要内容 (组间和组内)。通过流程图和表格呈现选择适当统计方法的过程, 并阐述统计分析过程中如何解释统计结果和撰写基础统计推断结论。此外包括假设检验中统计显著性的概念、功效和效应量测度的讨论以及如何解释输出结果等内容。

第 6 章到第 11 章介绍统计分析研究的主要方法。为了分析和解决实际问题, 这些章节主要介绍了一些中级统计学方法, 并且都可以通过 SPSS 完成统计计算。但愿通过本书的学习, 读者能够掌握统计分析研究的主要方法。同时, 我们也希望读者通过学习、研究输出结果、回答解释问题和完成 SPSS 拓展练习等环节的内容, 对统计计算过程有清晰的了解和掌握。

4. 有关研究问题、测度水平和选择统计分析方法

第 1 章和第 5 章的撰写与传统方法略有不同。这是基于学生对研究设计和统计方法选择方面存在困难而改变的。根据以往学生的学习情况看, 大部分学生对统计公式和用计算器或计算机进行计算很容易掌握。但是, 许多学生在如何使用统计方法和如何对统计结果进行解释方面有很大的困惑。他们头脑中没有一个画面或不能看到研究设计和测度水平如何影响数据分析。其中的一个主要原因是没有统一的表述。基于此原因, 在本书中将通过统一的表述和清晰的画面从三个方面的研究问题 (差异、关联和描述) 介绍研究设计的内容。我们知道发展和利用一致性的结构既是非传统性的, 又有些过分简化。但是, 就培养学生掌握利用统计学完成研究问题的能力而言, 结构性和统一性还是需要的。不认可这种知识结构安排的教师在使用本书时, 可以跳过第 1 章和第 5 章的第一部分内容, 直接让学生学习 SPSS 的使用方法。

5. 本书的主要变化和增加的内容

根据多年教学积累的经验、读者反馈意见和其他使用本教材的读者所提供的建议以及修订政策, 该书重新编排并增加了一些新内容, 这一做法已被 APA 统计推断专责小组 (1999)

和 APA 出版手册第 5 版 (2001) 指定为最佳做法。

1) 新的 SPSS 版本。本书中使用的 SPSS 版本由 SPSS 12.0 版本升级为 SPSS 15.0 版本。

2) 补充 SPSS 拓展练习。每章后都安排了 4 道 SPSS 拓展练习题。

3) 增加用 SPSS 进行分层线性模型 (HLM) 分析内容。在 SPSS 15.0 中增加了 HLM 分析, 因此, 本书也相应地增加了这部分内容的介绍。包括如何分析无条件层 1 重复测量模型; 带有层 2 预测变量的重复测量模型; 个体嵌套在学校的无条件模型; 个体嵌套在学校的层 1 协方差的有条件模型。

4) 后验多重比较 (Post Hoc) 分析。第 8 章的多因素方差分析和协方差分析中, 除了介绍对照比较 (contrast) 法外, 还增加了如何进行后验多重比较 (Post Hoc) 检验的内容。

5) 比比 (odds ratio)。在逻辑斯谛回归和判别分析中增加的效应量以及如何解释比比的内容。

6) 后验功效 (Post power)。第 5 章讨论功效内容中新增加的内容。并且对每种统计方法, 我们都说明如何计算和解释后验功效。

7) SPSS 入门知识。在附录 A 中增加了以下内容: (a) 从本书附赠的光盘中获取工作文件; (b) 打开和启动 SPSS 程序和逐渐熟悉 SPSS 数据编辑器; (c) 为打印语句或每个输出日志设置你的计算机; (d) 定义变量; (e) 定义变量水平; (f) 打印代码本; (g) 编辑输出; (h) 保存输出为 Excel、Word、HTML 或 PDF 文件; (i) 将变量转为标准化变量; (j) 选择观测; (k) 拆分文件; (l) 合并文件; (m) 使用 Means 命令; (n) 使用 Graphs 命令。

8) 如何进行常用的基础统计分析。在附录 B 中, 我们介绍了常用基础统计方法的操作步骤和例题, 包括 t 检验, 相关分析和 χ^2 检验等内容。

9) 假设。在如何撰写输出的统计结论内容中, 方括号中的内容都是关于检验假设的描述。对于撰写论文和学位论文的学生, 这部分内容为他们提供了如何讨论假设的模式。

6. 有关项目符号、箭头、粗体字和斜体字的说明

为帮助大家用 SPSS 分析问题, 我们沿用一些惯例。使用项目符号的部分显示的是有关 SPSS 视窗操作的步骤。例如:

- 选中 gender 和 math achievement。
- 单击箭头按钮将变量送入右手框中。
- 单击 Options, 进入如图 2.16 显示的对话框中。
- 选中 Mean, Std Deviation, Minimum 和 Maximum。
- 单击 Continue 按钮。

注意, 斜体字表示变量名, 粗体字是你在 SPSS 窗口会看到的字或词以及使用 SPSS 生成所要结果的关键内容。在本教材中, 这些粗体字正如你在 SPSS 窗口看到的一样, 是用大写

字母写的。在介绍、定义关键表述时，也使用粗体字，以表示它们的重要性。

在文中你也会看到，我们会用粗体字和粗体字间的箭头表示这些粗体字会出现在下拉菜单中。例如：

- 选择 **Analyze => Descriptive Statistics => Frequencies**

（上面的操作意味着在打开 **Analyze** 菜单后，将鼠标滑动到其下拉菜单中的 **Descriptive Statistics**，然后找到 **Frequencies**，并单击它。）

有时我们也使用下划线来强调要点或命令。

致 谢

本教材一直能够使用 Lawrence Erlbaum Associate 发行的 Gliner 和 Morgan (2000) 所著的《Research Methods in Applied Settings: An Integrated Approach to design and Analysis》，或者是 Mogan, Gliner 和 Harmon (2006) 所著的《Understanding and Evaluating Research in Applied and Clinical Settings》两本教材作为参考和补充。事实上，第 1 章和第 5 章中的一些内容只是对他们教材中的内容进行了细微的修改。在这里我们对这些书的合作者 Jeff Gliner 表示感谢。尽管 Orlando Griego 不是我们 SPSS 这本书的作者，但是她的学生也参与了教材的编写。

我们还要对教育学和人类发展班级中使用过这本教材早期版本的学生表示感谢，他们对这本教材的修改提供了很好的建议。我们还要感谢我们的技术顾问 Don Quick，文字处理 Alana Stewart 和从事这方面研究工作的学生。感谢 Laura Jensen, Andrea Fritz, James Lyall, Joan Anderson 和 Yasmine Andrews 对早期版本文字和编辑的帮助。感谢 Jeff Gliner, Jerry Vaske, Jim zumBrunnen, Laura Goodwin, David MacPhee, Gene Gloeckner, James O. Benedict, John Ruscio, Tim Urdan 和 Steve Knotek 对早期版本提供的回顾资料和对新版本的修改意见。感谢 Carolyn Springer, Jay Parkes, Joshua Watson 和 John Rugutt 对第 2 版编辑提供的帮助和对第 3 版提出的建议。感谢 Cress 和 Don Quick 对第 3 版附录提供的帮助。感谢 Bob Fetch 和 Ray Yang 阅读本书后提出反馈意见。最后，感谢我们的爱人（Grant, Terry 和 Hildy）和家人，让我们在没有太多家庭负担下顺利完成本书的撰写工作。

目 录

第 1 章 绪论	1
1.1 研究问题和变量	1
1.1.1 研究问题	1
1.1.2 变量(Variables)	1
1.2 研究假设和研究问题	4
1.2.1 差异研究	4
1.2.2 关联研究	4
1.2.3 描述研究	4
1.2.4 复杂研究	5
1.3 研究问题实例: 高中学校(HSB)研究	6
1.3.1 研究问题	6
1.3.2 HSB 的变量	7
1.3.3 SPSS 变量观察窗	7
1.3.4 HSB 数据中的变量名	8
1.3.5 HSB 原始数据和数据编辑	10
1.4 HSB 数据中的研究问题	11
1.5 频数分布	12
1.6 测度水平	13
1.6.1 名义变量(Nominal Variables)	14
1.6.2 二分变量	15
1.6.3 定序变量	15
1.6.4 近似正态(或尺度)变量	16
1.6.5 SPSS 中测度水平的分类	16
1.7 描述统计	18
1.7.1 集中趋势的测度	18
1.7.2 变异的测度	18
1.8 有关测度及其所对应的描述统计量	19
1.9 正态曲线	20
1.9.1 正态曲线的性质	20
1.9.2 非正态分布形状描述	21

1.9.3 正态曲线下的面积	21
问答题	22
第2章 数据编码和探索分析	23
2.1 数据编码规则	23
2.2 探索数据分析	25
2.2.1 什么是探索数据分析	25
2.2.2 如何做探索数据分析	26
2.2.3 验证假设条件	27
2.3 统计假设	27
2.4 对定序和尺度变量查找错误和验证假设	28
问题 2.1 定序变量和尺度变量的描述统计	29
问题 2.2 定序和尺度变量的箱图	32
问题 2.3 二分变量的箱图	34
问题 2.4 用表和图对二分变量和名义变量进行探索分析	37
问题 2.5 频数表的使用	38
问题 2.6 条形图	40
问题 2.7 直方图和频数多边形	41
问题 2.8 矩阵散点图	43
问题 2.9 变量的转换	45
问答题	47
SPSS 拓展练习	48
第3章 可靠性测度方法	49
问题 3.1 数学动机尺度的克隆巴赫 α 系数	50
问题 3.2 和 3.3 能力和兴趣尺度的克隆巴赫 α 系数	54
问题 3.4 使用相关计算测试 - 再测试可靠性	57
问题 3.5 名义数据的 Cohen' Kappa 系数	59
问答题	61
SPSS 拓展练习	62
第4章 探索性因子分析和主成分分析	63
问题 4.1 数学学习态度变量的因子分析	64
问题 4.2 在成绩变量上的主成分分析	71
问答题	75
SPSS 拓展练习	76
第5章 选择和解释推断统计	77
5.1 推断统计的选择	78

5.2	一般线性模型	83
5.3	解释统计检验结果	85
5.3.1	统计显著性	85
5.3.2	置信区间	86
5.3.3	效应量	86
5.3.4	功效(Power)	88
5.3.5	解释推断统计的步骤	90
5.4	如何选择和解释基础推断统计的回顾	91
5.5	撰写有关输出的回顾	97
	结束语	99
	问答题	99
第6章	多元回归	100
问题 6.1	使用同时引入方法计算多元回归	101
问题 6.2	同时回归校正多重共线性	108
问题 6.3	分批多元线性回归	115
	问答题	120
	SPSS 拓展练习	120
第7章	逻辑斯蒂回归和判别分析	122
问题 7.1	逻辑斯蒂回归	123
问题 7.2	分批逻辑斯蒂回归	128
问题 7.3	判别分析(DA)	132
	问答题	141
	SPSS 拓展练习	141
第8章	多因素方差分析和协方差分析	143
1.	多因素方差分析和协方差分析的假设	143
2.	协方差分析附加的假设条件	144
问题 8.1	析因(二维)方差分析	144
问题 8.2	交互效应的多重比较	150
8.2a	对新变量 cellcode 赋值的计算	150
8.2b	Post Hoc 检验(多重比较)的计算	153
8.2c	对照比较法的计算	156
问题 8.3	协方差分析(ANCOVA)	161
	问答题	167
	SPSS 拓展练习	167
第9章	重复测量和混合设计的方差分析	169

关于 Product data 数据集	170
问题 9.1 重复测量方差分析	171
问题 9.2 多个相关样本的 Friedman 非参数检验	177
问题 9.3 混合设计方差分析	181
问答题	185
SPSS 拓展练习	185
第 10 章 多元方差分析和典型相关	186
问题 10.1 GLM 单因素多元方差分析	187
问题 10.2 GLM 双因素多元方差分析	195
问题 10.3 混合多元方差分析	200
问题 10.4 典型相关	207
问答题	212
SPSS 拓展练习	213
第 11 章 多水平线性模型/分层线性模型	215
问题 11.1 无条件层 1 重复测量模型	217
问题 11.2 带有层 2 预测变量的重复测量	227
问题 11.3 个体嵌套在学校的无条件模型	234
问题 11.4 个体嵌套在学校的层 1 协方差的条件模型	239
问答题	245
SPSS 拓展练习	246
附录 A SPSS 入门知识和其他有用的操作命令	247
A.1 从 CD 光盘复制数据文件	247
A.2 打开和启动 SPSS 应用程序	248
A.3 为打印 SPSS 语句(日志)设置你的计算机	250
A.4 从 Excel 导入数据	252
A.5 定义变量	252
A.6 给变量值或变量水平加标签	253
A.7 打印代码本	254
A.8 编辑输出	254
A.9 保存输出结果为 Excel、Word 或 HTML 格式	255
A.10 将变量转换为标准化变量(z 分数)	256
A.11 数据选择	256
A.12 拆分文件	257
A.13 合并文件	257
A.14 Compute 命令和 Mean 函数	259

A.15	Graphs 命令	260
附录 B	基础统计知识回顾	261
B.1	单因素方差分析(ANOVA)	261
B.2	χ^2 、phi 和 Cramer' V	263
B.3	相关系数(Pearson 和 Spearman)	265
B.4	Mann-Whitney(曼 - 惠特尼)U 检验	267
B.5	比数比(和风险比)	269
B.6	回归分析	270
B.7	t 检验	270
B.8	威尔科克森符号秩检验	272
附录 C	问答题奇数题的答案	274

第1章 绪 论

本章主要介绍有关测量和描述统计的重要内容,并对 CD 光盘中所提供的高中及以上学校(HSB)的数据进行介绍,本章及以后各章节中的例题都会使用该数据。通过对该数据的分析,介绍统计方法的应用及对结果的解释。首先,我们简要叙述本教材中涉及的重要统计概念。

1.1 研究问题和变量

1.1.1 研究问题

研究过程始于研究者所感兴趣的论点或问题。研究问题是针对两个或两个以上变量之间关系的一种陈述。几乎所有的研究都会涉及两个以上的变量。

1.1.2 变量(Variables)

变量是研究问题的关键元素。所谓变量是指一项研究中对被试对象特征或环境特征的描述,它有若干个值。变量必须是可变的,有不同的取值或水平¹。*gender*(性别)是一个变量,它有两个取值,男和女。*Age*(年龄)是一个变量,它有多个不同的取值。如果处理或干预的类型(或者课程类型)是多于一种的处理或者有一个处理组和一个控制组,那么它们也是变量。学习天数或疾病恢复天数也是变量,类似地还有数学知识的数量,因为它的范围是从无到很多,所以也是变量。在某个特殊研究中,如果某个概念只有一个值,它就不是变量,而是一个常量。例如,如果所有的被试对象都是欧洲裔,则 *ethnic group*(种族)是常量。如果研究中所有被试对象都是女性,则 *gender* 就不是变量。

定量研究中,变量有不同的分类,通常将变量分为自变量(活动变量或属性变量)、因变量和外部变量。下面针对每个概念进行介绍。

1. 变量的操作定义

变量的操作定义是针对变量发生或测量它时使用的操作或技术而言。定量研究中,描述

1 为了帮助大家更好地掌握所讲内容,在本书中对于 SPSS 中的变量名或标签使用斜体字。有时使用斜体字是为了强调某个词。SPSS 中的表述,输出结果和其他重要表述则使用加粗字体(如 **SPSS Data Editor**)。使用下划线的地方,是为了强调其重要性。有关 SPSS 的操作过程则使用符号 • 来标注(如单击、选中等命令语句)。

一个变量时,是通过示范如何测量变量来指出变量含义的。年龄、性别、种族等人口统计学变量通常是很简单地通过让参与者从列表中选择恰当的选项得到。治疗(或课程)类型的操作定义通常是通过更深入地描述他们在治疗或新课程期间都做了什么。同样地,一些抽象概念,如数学知识,自我意识,或数学焦虑的操作定义,需要通过更详细地描述说明在特殊研究中是如何测量它们的。为了做到这些,研究者需要提供样例,附加实际的测量工具,或者提供一些尽可能找到的信息。

2. 自变量(Independent Variables)

自变量分为两类:活动自变量和属性自变量。讨论研究结果时,变量类型的区分是非常重要的。

活动或控制自变量。活动自变量是这样的变量,如某个工厂,某门新课程或某种干预,施加给被试对象的,或者研究过程中在指定的时间段中所施加的处理。例如,研究者想对新的治疗方法和传统治疗方法进行比较;又例如对一种新教学方法效果的研究,如合作学习中学生表现的评价。这两个例子中,我们要研究的变量是施加给被试对象的某个因素。因此,活动自变量是研究期间施加给被试对象的,但不一定是通过实验人员施加或控制的。自变量可能是通过诊所、学校或者某人施加的,而不是调查人员施加的,但它反应了被试对象的观点,这种情形是可控制的。治疗是在研究计划制定后施加的,因此有实验前测试和随机设计。如果没有实验前测试,研究的内部效度可能是有风险的。虽然我们所讲的活动自变量的概念与其他作者相似,但可能会有微小的差异。随机化试验和准试验研究都有活动自变量。一个活动自变量是必要的,但要得出因果结论,条件还不充分。只有当被试对象被随机分配到由实验者控制的几种条件下,才可能得出最清晰的因果结论。

属性或测定自变量。如果一个变量不能被控制,但它又是研究的重点,则被称为属性自变量。对于属性自变量,其自变量值是事先存在于人的属性或者他们所在环境中,在研究期间不能被改变。例如,性别、年龄、种族、智商和自尊都能够用来作为属性自变量。只对属性自变量进行的研究被称为非实验性研究。

为了与 SPSS 保持一致,本书与其他一些研究方法教材的不同之处是,没有把自变量限制为控制自变量或活动自变量,而是更广泛地定义了自变量,包括研究中的预测因素,前提条件,或假设的原因,或对观察的影响。被试对象的属性和活动自变量,都是符合这个定义的。对于社会科学和教育学,属性自变量特别重要。例如,瘫痪的类型或瘫痪的水平可能作为研究的核心,在研究中尽管它们没有被给出,但因为它可能有不同的值,所以瘫痪一定要作为一个变量。例如,脑瘫与唐氏综合症是不同的,与脊柱裂也不同,但所有这些都是瘫痪。对于同样的瘫痪也有不同的水平,人们已经定义了他们的特性和属性,并将它们放到不同的类别中。我们在开始研究时不同的瘫痪情况已经发生了,因此,我们可能感兴趣的研究变量不是研究期间施加或控制的变量,而是由其他人、学校或诊所施加或控制的。