

SPSS

<http://www.phei.com.cn>

统计分析简明教程 (第10版)

SPSS for Windows Step by Step
A Simple Study Guide and Reference, Tenth Edition

— [加] Darren George 著
Paul Mallery

— 何丽娟 李 征 韦 玉 译

— 卢纹岱 审校



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

SPSS 统计分析简明教程

(第10版)

SPSS for Windows Step by Step
A Simple Study Guide and Reference

Tenth Edition

[加] Darren George Paul Mallery 著

何丽娟 李 征 韦 玉 译

卢纹岱 审校

電子工業出版社·

Publishing House of Electronics Industry

北京·BEIJING

译者序

SPSS 软件是在社会科学研究中被广为使用的一个统计分析工具之一。使用它可以实现多种统计分析过程,对所研究的数据进行统计分析,从而得出统计结论。本书作者对 SPSS 的统计分析过程和操作步骤进行了详细而清晰地介绍,随着本书各个章节的学习,读者可以很容易掌握 SPSS 中的主要分析过程。本书着重统计分析的目的、各种统计分析的基本假设,以及如何实现分析过程。《SPSS 统计分析简明教程》让读者对统计方法不再感到恐惧,让复杂的统计分析操作变得简单易学。读者可以借助每一章的操作步骤部分,完成对自己所收集的数据的分析。对于学生学习统计方法和研究者从事科学研究,该书具有很强的实用性。

该书已经是第 10 版,SPSS 软件已经更新为 17.0 版本,本书介绍的方法适用于以前的版本和今后将要升级的版本。

该书的特点是简洁、全面。虽然每个章节的内容不多,但涵盖的内容较为广泛。全书共有 28 章,包括统计分析的基本内容和高级统计分析方法。前五章介绍在 SPSS 中如何创建数据和管理数据。从第 6 章到第 22 章主要讲解统计分析的基本模块,包括频数分析、描述统计、 t 检验、相关分析、方差分析、回归分析、非参数检验、信度分析、多维尺度分析、因子分析、聚类分析和判别分析。最后的第 23 章到第 28 章主要介绍多变量的方差分析和协方差分析、重复测量的多变量方差分析、Logistic 回归分析、对数线性模型以及残差分析。

在阅读本书时,注意每一章的基本结构。每一章的概述部分主要是对要介绍的统计方法的基本介绍。操作步骤环节主要介绍在 SPSS 中如何实现分析过程的操作步骤,如何得出要分析的统计结果。读者可以借助这部分操作,使用自己的数据完成分析过程。输出部分主要是对输出结果的解释和说明,告诉你如何分析输出结果。

本书的作者 Darren George 是 Canadian University College 的教授,Paul Mallery 是 La Sierra University 的教授,他们都是心理学博士,并从事 SPSS 教学多年,统计分析的经验非常丰富,他们特别重视研究方法和实验方法,并熟知心理学研究的数据分析方法。为了方便读者学习,本书另提供配套的数据文件,读者可以从 wps.ablongman.com/ab_george_windows_10/ 或从 www.hxedu.com.cn 下载获得。

本书由何丽娟、李征、韦玉翻译,其中第 1~5 章由韦玉翻译,第 6~10 章由李征翻译,第 11~28 章以及附录由何丽娟翻译。卢纹岱教授负责该书全面的审校工作,全书由何丽娟负责统稿。在翻译过程中,得到卢纹岱、朱红兵关于统计方法的帮助,解利辉、费青松对软件操作上的基本概念给予了解答,杨一虹、王勤、千青海、陈文倩提供了文字阅读的帮助,

前 言

SPSS 是一个功能强大的统计分析软件，它具备对社会科学、自然科学以及商业等各个领域的数据进行分析的功能。如果说数学是科学的语言，那么数据分析则是调查研究的语言。许多领域的调查研究对人类进步是至关重要的，只要有调查研究，一定会涉及数据分析。本书力图使数据分析更易理解，尽可能避免枯燥乏味的内容。

在教学中，经常遇到这样的情况，一些教授很高兴地告诉学生“用 SPSS 分析这些数据；如果你不知道怎么做，去看手册。”于是，学生们放弃学习这门课程，而不是继续努力学习这门课程，让学生遭受损失。本书的构思过程是痛苦的。达伦·乔治曾在高中教授数学，保罗·马勒在学院为学生培训计算机的使用和编程，在着手编写本书时我们发现让原本复杂的事情变得尽可能简单清晰的这一挑战过程是令人非常愉快的。我们的最终目标是让人们通过本书的学习对 SPSS 过程有一个清晰的了解。

当本书初步成形时，我们又有了第二个目标。除了让初学者对 SPSS 过程了解，还想创建一个工具，任何人进行数据分析时都可以用它作为参考。这涉及将原有概念扩展到所有主要的统计过程，这些过程包含在 SPSS 覆盖的基本模块、高级模块以及回归模块中。

第 10 版与以前的版本比较，变化相对较小。“变化较小”并不是说不重要。该书每个版本首要的任务就是要反映新的 SPSS 版本出来后的变化。第 9 版使用的是 SPSS 16.0，而第 10 版使用 SPSS 17.0 进行介绍。对于非商业专业，一个基本的“独家销售主张”使教材总是最新的。一旦 SPSS 升级，我们就要编写一本新的版本以反映其变化。在 12 年中有 10 个版本。对每个版本，作者之一（通常是 Darren）都通读全文（结合读者的建议），并对每个步骤进行操作。这可以保证这本书总是最新的，每个步骤是准确的。新版本有时需要对某个章节进行完善或增加新的一章。幸亏这次两者都没有涉及。以后的版本可能在名称上会有所变化，如 PASW Statistics，PASW 是预测分析软件（Predictive Analytics Software）的简称。

本书的前 16 章是基本内容，它无需太多的统计知识背景，而后 12 章的内容逐步需要更牢固的统计知识背景，更具有挑战性。在每一章的开始，我们都用几页的篇幅描述要讲授的内容。但是，如何充分描述，例如，因子分析或判别分析用 5 页或 6 页的篇幅就是充分描述了吗？答案很简单：作为常识、概念性的过程一定要描述，避免过分详细和过分强调计算。然而，编写这些内容并非都是这样简单。本书中最费尽心机的是每章的概述部分。尽管我们知道大多数过程的解释缺乏更详细的内容，但是我们认为我们已经在每个环节上都企图做得更充分。即在有限的空间，在最大程度上成功地使所讲内容清晰易懂？最终的判断者是你们——读者。

SPSS 公司为了试图完整地描述软件包中的所有内容，已经出版了几个版本手册指南。精美打印的手册有 2000 页之多，是非常全面的。手册中对 SPSS 所能实现的功能一一进行了叙述。对于专业研究员，需要拥有这样的手册。而我们不可能将近 2000 页的 SPSS 手册涵盖在本教材中。然而，我们确信在科学或商业领域中曾使用的分析方法，有 95% 通过这本教材中介绍的方法可以完成提交资料的分析。额外的 5% 我们常常建议读者参考手册中的内容。

作者简介和联系方式

达伦·乔治 (Darren George): 心理学教授, 主讲人格心理学、社会心理学以及研究方法等课程。1982 年获得加利福尼亚州立大学硕士学位; 1980 ~ 1989 年在加拿大的 Mark Keppel 高中和 Mountain View 高中教授数学; 1992 年在 UCLA 大学获得心理学博士学位, 研究方向重点是行为心理学、社会心理学和心理测量学。其联系方式如下:

Canadian University College
5415 College Avenue
Lacombe, AB, T4L 2E5
403 - 782 - 3381, Ext. 4082
dgeorge@cauc.ca

保罗·马勒 (Paul Mallerry): 心理学教授, 主讲社会心理学及其相关课程, 并讲授实验方法 (包括 SPSS 的应用)。1994 年在 UCLA 获得社会心理学博士学位, 研究方向重点是统计学和政治心理学。保罗原先的工作是计算机专家, 既从事编程工作又讲授计算机课程。其联系方式如下:

La Sierra University
4500 Riverwalk Parkway
Riverside, CA, 92515
951 - 785 - 2528
pmallery@lasierra.edu

致 谢

回顾过去几年富有创造性的工作，衷心地感谢从未见过面的一些人。他们独具慧眼，审阅书稿，提供有价值的修改建议。我们还要感谢 John Brown 大学的 Richard Froman、Marist 学院的 Michael A. Britt、South Florida 大学的 Marc L. Carter 以及 Ouachita Baptist 大学的 Randolph A. Smith。然后还要感谢一直鼓励我们完成本书编写的家人和朋友，尤其是 Elizabeth George 和 Suzanne Mallery 对我们的支持和鼓励，最后感谢我们的孩子们。

目 录

第 1 章 绪论	1
1.1 必要的技能	1
1.2 覆盖范围	1
1.3 概述	3
1.4 各章结构	3
1.5 书中使用例子的介绍	4
1.6 印刷及格式约定	5
第 2 章 SPSS 窗口操作	8
2.1 鼠标	8
2.2 任务栏和开始菜单	9
2.3 常用按钮	10
2.4 数据窗口和其他常用窗口	11
2.4.1 一个关于统计过程对话框的例子	15
2.4.2 键盘操作、复选框以及单选按钮	16
2.5 输出窗口	17
2.6 打印输出	22
2.7 “Options…选项”: 修改格式	24
第 3 章 数据文件的创建和编辑	26
3.1 数据文件的研究要点和结构	26
3.2 创建数据文件的操作步骤	27
3.2.1 变量名	29
3.2.2 变量类型	30
3.2.3 小数位数	30
3.2.4 变量宽度	31
3.2.5 标签	31
3.2.6 变量值	32
3.2.7 缺失值	33
3.2.8 列宽	33
3.2.9 对齐方式	34
3.2.10 测度水平	34
3.3 输入数据	34

6.5	百分位数	83
6.6	操作步骤	83
6.7	打印结果	89
6.8	输出	89
	练习题	91
第7章	描述统计	93
7.1	统计显著性	93
7.2	正态分布	94
7.3	集中趋势统计量	95
7.4	离中趋势统计量	95
7.5	正态偏差统计量	96
7.6	对分布大小的度量	97
7.7	稳定性的度量: 标准误	97
7.8	操作步骤	97
7.9	输出	101
	练习题	101
第8章	交叉列联表和卡方分析	103
8.1	交叉列联表	103
8.2	卡方独立性检验	104
8.3	操作步骤	105
8.4	输出	110
8.5	术语定义	111
	练习题	112
第9章	均值过程	114
9.1	子总体差异	114
9.2	操作步骤	114
9.3	输出	118
9.4	术语	119
	练习题	120
第10章	双变量的相关分析	121
10.1	什么是相关?	121
10.2	线性相关与曲线相关	123
10.3	显著性概率	123
10.4	因果关系	124
10.5	偏相关	124
10.6	操作步骤	125

10.7 输出	129
练习题	130
第 11 章 t 检验过程	131
11.1 什么是 t 检验	131
11.2 独立样本 t 检验	131
11.3 配对样本 t 检验	131
11.4 单一样本的 t 检验	132
11.5 显著性检验	132
11.6 操作步骤	132
11.7 输出	137
练习题	139
第 12 章 单因素方差分析	141
12.1 单因素方差分析概述	141
12.2 操作步骤	142
12.3 输出	147
练习题	149
第 13 章 双因素方差分析	151
13.1 方差分析模型概述	151
13.2 统计功效	151
13.3 双因素方差分析及其主效应与交互效应	152
13.4 操作步骤	153
13.5 输出	156
练习题	158
第 14 章 三因素方差分析	159
14.1 方差分析模型概述	159
14.2 三因素方差分析	159
14.3 协变量的影响	160
14.4 操作步骤	160
14.5 输出	164
14.6 交互效应图	170
练习题	172
第 15 章 简单线性回归	174
15.1 简单线性回归方程概述	174
15.2 预测值和回归方程	174
15.3 简单回归和可解释的方差比例	176
15.4 曲线关系的检验	176

15.5	操作步骤	179
15.6	输出	183
	练习题	185
第 16 章	多元回归分析	188
16.1	多元回归分析概述	188
16.2	回归方程	188
16.3	可解释的方差百分比	190
16.4	曲线趋势和模型建构	190
16.5	操作步骤	191
16.6	输出	197
16.7	加入新变量时值的变化量	198
	练习题	201
第 17 章	非参数检验	203
17.1	非参数检验概述	203
17.2	操作步骤	204
17.3	曼-惠特尼秩和检验	205
17.3.1	操作步骤	206
17.3.2	输出	206
17.4	符号检验	207
17.4.1	操作步骤	207
17.4.2	输出	208
17.5	威尔科克逊配对符号秩的检验	208
17.5.1	操作步骤	209
17.5.2	输出	209
17.6	游程检验	209
17.6.1	操作步骤	210
17.6.2	输出	210
17.7	二项分布检验	210
17.7.1	操作步骤	211
17.7.2	输出	211
17.8	柯尔莫戈洛夫-斯米尔诺夫单样本检验	211
17.8.1	操作步骤	212
17.8.2	输出	212
17.9	单样本的卡方检验	212
17.9.1	操作步骤	213
17.9.2	输出	213

17.10	弗里德曼—维方差分析	213
17.10.1	操作步骤	214
17.10.2	输出	214
17.11	K 个样本的中位数检验	214
17.11.1	操作过程	215
17.11.2	输出	215
第 18 章	信度分析	217
18.1	信度分析概述	217
18.2	克隆巴赫 α 系数	218
18.3	分半信度	218
18.4	操作步骤	219
18.5	输出	223
	练习题	228
第 19 章	多维尺度分析	229
19.1	多维尺度分析概述	229
19.2	不对称方阵（社会关系图的例子）	230
19.3	从数据中创建距离对称方阵（测验分数的例子）	230
19.4	个体差异模型（电视节目的例子）	231
19.5	操作步骤	231
19.6	输出	237
第 20 章	因子分析	240
20.1	因子分析概述	240
20.2	计算相关系数矩阵	240
20.3	提取因子	241
20.4	因子的选取与旋转	241
20.5	对输出的解释	243
20.6	操作步骤	244
20.7	输出	249
第 21 章	聚类分析	253
21.1	分层聚类分析	253
21.2	聚类分析和因子分析的比较	253
21.3	聚类分析的步骤	254
21.4	操作步骤	256
21.5	输出	262
第 22 章	判别分析	267
22.1	判别分析概述	267

22.2	举例：获准毕业	268
22.3	判别分析的步骤	269
22.4	操作步骤	270
22.5	输出	276
第 23 章	多因变量方差分析和协方差分析	283
23.1	多因变量方差分析和多因变量协方差分析概述	283
23.2	操作步骤	284
23.3	模型选项对话框	286
23.4	轮廓图对话框	287
23.5	多重比较检验	288
23.6	附加选项	290
23.7	输出	292
	练习题	297
第 24 章	重复测量多元方差分析	299
24.1	重复测量多元方差分析概述	299
24.2	操作步骤	300
24.3	选择模型	301
24.4	输出	305
	练习题	308
第 25 章	逻辑斯蒂回归分析	310
25.1	逻辑斯蒂回归分析概述	310
25.2	逻辑斯蒂回归分析的数学模型	310
25.3	操作步骤	312
25.4	输出	316
第 26 章	分层对数线性模型	321
26.1	对数线性模型概述	321
26.2	模型选择对数线性过程	322
26.3	操作步骤	323
26.4	输出	328
第 27 章	非分层对数线性模型	334
27.1	非分层与分层对数线性模型	334
27.2	对数线性模型中的协变量	334
27.3	Logit 模型	335
27.4	关于模型选择的一些话题	335
27.5	本章未包括的模型类型	335
27.6	操作步骤	336

27.7 输出	341
第 28 章 残差分析	344
28.1 关于残差	344
28.2 案例研究之一：线性回归中的残差分析	345
28.3 案例研究之二：一般对数线性模型中的残差分析	347
28.4 关于 SPSS 中的残差分析	350
术语	353
参考文献	365

1.3 概 述

本章将介绍本书要讨论的一些主要概念，并对本书的组织结构以及将要用到的基本工具的使用方法做一个简要的概述。

如果你要执行一个统计过程，并且以前使用过 SPSS，知道要做什么分析，那么你应该阅读本章的章节组织结构、编辑和格式约定的部分，再学习本书后面的相关章节（第 6 至第 28 章）。那些章节会告诉你需要执行什么样的步骤来获取你所需要的输出结果。

如果你是 SPSS 新手，那么这一章将为你提供使用本书所需的重要而实用的背景知识。

1.4 各章结构

本书描述了分析数据的重要概念。与数据分析相关的三个基本任务是：

- A. 你必须将数据输入计算机中，并要建立数据结构和格式化数据，使得你和计算机都能轻易识别数据；
- B. 必须告诉 SPSS 你要进行什么类型的分析；
- C. 你还要能够解释 SPSS 输出结果的含义。

本章绪论之后，第 2 章将介绍一些基本操作，比如 SPSS 的窗口类型，工具栏和菜单的使用，保存，视图及编辑输出，打印输出等。虽然本章是面向初学者设计编写的，但其中有很多 SPSS 的特殊知识对每个人都很有实用。第 3 章介绍前面提到的第一步——对一个数据文件进行创建、编辑和格式化。SPSS 数据编辑器（SPSS data editor）是一个使得数据文件的创建、组织和格式化非常简单而清晰的工具。

第 4、第 5 章交待两个重要的问题——数据的修改和转换（第 4 章），以及图表的创建（第 5 章）。第 4 章具体介绍不同类型的数据操作，例如为分析创建新变量、排序、重组、合并文件、或选取数据子集。第 5 章介绍制作一部分图表所使用的基本过程；而其他的图表则在后面的章节中进行全面的描述。

第 6 ~ 28 章介绍步骤 B 和 C——分析数据并解释输出结果。需要注意每部分内容都可以从其所在的章节中找到相应的说明。比如，初学者要对某个数据做 t 检验，则可以从第 11 章找到完整的执行过程说明。在基本操作部分，步骤 1（Step 1）总是“打开 SPSS 程序”，如果对此有任何问题，建议读者参考第 2 章的内容。步骤 2（Step 2）一般都是“创建一个数据文件或者对一个已有文件进行编辑（如果必要的话）”；如果需要的话，读者可以查看第 3 章的相关指导。接下来的步骤就是对如何执行 t 检验进行完整的讲解。

正如前面提到的，本书覆盖了 SPSS 的三个模块：**PASW** 基本统计，**PASW** 高级统计以及 **PASW** 回归。由于一些高校的电脑中可能没有所有的三个模块（基本模块一般都有），本书按照 SPSS 的结构进行编排：覆盖几乎所有的过程，包括基本模块并从更复杂的高级和回

归模块中选择部分过程进行讲解。第 6~22 章介绍基本模块中的过程。第 23~27 章介绍高级和回归模块中的过程。第 28 章介绍的残差分析涉及到这三个模块的内容。

PASW 基本统计：第 6~10 章描述了最基本的数据分析方法，包括频数、条形图、直方图以及百分位数（第 6 章）；描述统计，如均值、中位数、众数、偏度和极差（第 7 章）；列联表和独立性卡方检验（第 8 章）；均值过程（第 9 章）；还有变量间的相关关系（第 10 章）。

下一组章节（第 11~17 章）介绍了给定数据中不同组别之间的检验方法，或通过使用 t 检验（第 11 章），ANOVA（第 12、第 13、第 14 章）；线性、曲线回归、多元回归分析（第 15、第 16 章）展示一个因变量和一个或多个自变量之间的相关强度；而非参数检验的最普遍的使用方法则在第 17 章进行讨论。

信度分析（第 18 章）是研究中使用的一种标准度量方法，它涉及多元响应的测度；多维尺度分析用来对一组不相似数据进行识别，以及构建模型结构和维度（第 19 章）；而因子分析（第 20 章）、聚类分析（第 21 章）以及判别分析（第 22 章）也是在科学家的研究中占据稳定和重要地位的方法。

PASW 高级统计和 PASW 回归：接下来的章节介绍的是涉及多个因变量的分析（SPSS 称这些过程为一般线性模型（General Linear Model）；它们也常被称为多元方差分析（MANOVA）或多元协方差分析（MANCOVA）。在一般线性模型下又包括简单的和一般的因子模型和多变量模型（第 23 章），以及重复测量或组内因素模型。

在接下来的三章中介绍的是一些不常用的过程，但是当需要使用这些过程的时候，它们是不可缺少的，因此这里也对它们进行介绍。第 25 章介绍逻辑斯蒂克回归分析，第 26 和第 27 章分别介绍分层和非分层对数线性模型。如前所述，以第 28 章的残差分析结束本书。

1.5 书中使用例子的介绍

在本书前 19 章中，有 17 章都是以一个简单数据文件为例的。对复杂的过程有必要选择不同的数据文件来反映相关过程。例题数据文件很实用，因为通常很多看起来难以理解的 SPSS 文档，在看到其相关例子是如何执行之后就显得清晰易懂。虽然此处用于描述的都只是些最常用的简单数据文件，但是本书一共提供了 12 组数据集用于演示说明这些过程，除此之外还有些数据用于练习题中。这些数据文件可以从 www.pearsonhighered.com/george 网站上下载。这些数据的极大好处在于为读者在练习中消除了输入处理数据的负担。我们建议读者将这些数据在各个不同的过程中充分地使用并将得到的结果与不同章节中输出部分的结果进行对比。

例子是经过设计的，以便于对这里讲述的大多数统计过程进行说明。这个简单数据文件是由一名教师对来自 3 个区域，每个区域约有 35 名学生进行教学所收集的数据。对于每一位学生，文件记录如下信息：

- 学号
- 姓名
- 性别
- 种族
- 在校时间
- 高低年级
- 先前的平均成绩（GPA）
- 区域
- 是否参加过复习单元或做过额外努力
- 在 5 个 10 分制的小测验和 1 个 75 分制期末考试中的成绩

第 4 章中我们将讲授如何创建 4 个新变量。而在后续的部分均包含这 4 个变量：

- 获得的总分
- 期末百分制成绩
- 期末评出的成绩等级
- 该生是否通过本门课程

这个数据文件（整个数据将显示在第 3 章结尾处）也将作为例子在介绍性章节（第 2 ~ 5 章）使用。如果你自行手动输入数据并按这几章描述的过程执行，你将会获得一个和本书前半部分所使用的完全相同的数据文件。当然，它与下载的数据文件一模一样，不过你可以由此练习数据输入、格式化以及使用该数据集进行特定数据的操作。如果你有自己的数据可用，那就更好了。

最后一点要注意的：**grades** 文件中的所有数据都是完全虚构的，因而对它的任何结果都是由我们在创建文件的时候设计好的。

1.6 印刷及格式约定

章节结构：第 2 ~ 5 章讲述了 SPSS 的 Windows 界面格式和过程，覆盖的内容决定了各章节的组织结构。第 6 ~ 28（分析章节）章中，除了个别例外，其他的结构完全相同。具体格式包括：

1. 概述是对要介绍的过程所做的短小而简洁的描述。这些介绍根据所描述分析方法的复杂程度长短不一。
2. 基本操作部分将完成特定分析的必要步骤展示出来。下面绝大多数对印刷和格式约定的描述都与基本操作部分相关。
3. 输出部分，负责将前述分析结果展示出来——通常被压缩。文字部分阐述的是输出的含义，并且定义所有关键的输出术语。

界面：基于 SPSS 的可视化特征，本书的每一章都包含操作过程的相关界面图片以及在