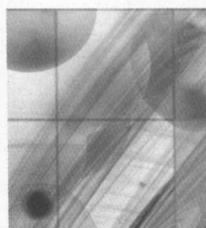


刘爱玉 著

SPSS 基础教程



世纪出版集团 上海人民出版社



刘爱玉 著

SPSS 基础教程

世纪出版集团 上海人民出版社

图书在版编目(CIP)数据

SPSS 基础教程/刘爱玉著. —上海:上海人民出版社,
2007

ISBN 978-7-208-06903-9

I. S… II. 刘… III. 统计分析-软件包 SPSS-教
材 IV. C819

中国版本图书馆 CIP 数据核字(2007)第 029978 号

出 品 人 施宏俊
责任编辑 姚映然
封面设计 王小阳



世纪文景

SPSS 基础教程

刘爱玉 著

出 版 世纪出版集团 上海人民出版社
(200001 上海福建中路 193 号 www.ewen.cc)
出 品 世纪出版集团 北京世纪文景文化传播有限公司
(100027 北京朝阳区幸福一村甲 55 号 4 层)
发 行 世纪出版集团发行中心
印 刷 北京中科印刷有限公司
开 本 787×1092 毫米 1/16
印 张 20
插 页 2
字 数 427,000
版 次 2007 年 7 月第 1 版
印 次 2007 年 7 月第 1 次印刷
ISBN 978-7-208-06903-9/C · 274
定 价 38.00 元

本书介绍如何运用SPSS对通过问卷或其他手段收集的定量资料进行描述和分析。作者力图以简明、扼要、直观的方式，展示SPSS分析的基本原理与操作，并在展示过程中将统计原理与SPSS数据分析相结合。本书有易学易用、步步到位、重操作和实务三大特点，是统计学初学者的必备参考书。

刘爱玉，北京大学社会学系副教授。主要研究领域为劳动社会学、组织社会学、人力资源开发与管理。主要学术成果有：《社会变迁过程中地区收入差异研究》、《社会变迁过程中的老年人家庭支持研究》、《政策、结构与人力资源》、《选择：国企变革与工人生存行动》、《劳动社会学教程》等。

前 言

SPSS 是英文 Statistical Package for the Social Science(社会科学统计软件包)的缩写。20 世纪 60 年代,美国斯坦福大学的三位研究生研制开发了最早的统计分析软件 SPSS,同时成立了 SPSS 公司,并于 1975 年在芝加哥组建了 SPSS 总部。1984 年 SPSS 总部首先推出了世界上第一个统计分析软件微机版本 SPSS/PC+。SPSS 最初的几个版本都是在 DOS 系统下运行的,虽然功能比较强,但在用户界面、输入、输出环境等方面并不十分理想。80 年代末,Microsoft 推出 Windows 后,SPSS 迅速向 Windows 移植。90 年代以来,由于 MS Windows 的普及,SPSS 公司于 1992、1993、1994 年相继推出了在 Windows 操作系统下的 5.0、6.0、6.1 版本。随着 Windows95 与 WindowsNT3.5 以上版本的 32 位操作系统的出现,1997 年以来又相继推出了 8.0 至 13.0 的各种版本。

随着 SPSS 版本的不断升级,其统计分析的功能也不断更新和完善,目前 SPSS 的主要功能有:样本数据的描述和预处理、描述性统计、均值比较、一般线性模型、相关分析、回归分析、对数线性模型、聚类分析、数据简化、假设检验(包括参数检验、非参数检验及其他检验)、生存分析、时间序列分析、多重响应等几大类,每类中又分好几个统计过程,比如回归分析中又分线性回归分析、曲线估计、Logistic 回归、Probit 回归、加权估计、两阶段最小二乘法、非线性回归等多个统计过程,而且每个过程中又允许用户选择不同的方法及参数。SPSS 也有专门的绘图系统,可以根据数据绘制各种图形。

SPSS 现已推广到各种操作系统的计算机上,它和 SAS、BMDP 并称为国际上最有影响的三大统计软件。目前 SPSS 已被广泛应用于社会学、经济学、政治学、心理学、管理学、市场营销、卫生统计等众多领域。世界上许多有影响的报刊杂志纷纷就 SPSS 的自动统计绘图、数据的深入分析、使用方便、功能齐全等方面给予了高度的评价与称赞。

SPSS 已越来越成为一种从事社会科学研究或调查,甚至从事自然科学研究的必备工具。

本书主要介绍的是如何运用 SPSS 对通过问卷或其他手段收集的定量资料进行描述和分析。主要内容安排如下:

内容一:SPSS 与社会科学研究。这是第一章的内容,主要介绍量化社会科学研究的一般步骤及其 SPSS 在量化分析中的作用。

内容二:数据分析前的准备。主要内容为第二章 SPSS 基本知识、第三章数据录入、第四章 SPSS 数据文件的编辑与管理、第七章数据变换。这些章节的内容将通过案例介绍如何打开文件、如何运行统计程序、检查结果、录入数据以及为进一步分析进行数据转换,书中还简要介绍了如何运用这些简单的操作来制作排行榜。

内容三:数据描述。主要内容为第五章用 Frequencies 作数据汇总、第六章描述性统计、第十五章统计图形绘制。统计分析的第一步是对数据进行展示,如多少人对工作感到满意?

他们的性别、年龄？可以对数据进行核查，如出现年龄大于 150 岁的人、每月收入超过几百万等一些看来明显的错误进行纠正。数据分析的一个重要功能是，用图示的方式来进行展示，在本书中，我们将用专门的一章来介绍一些主要的图形制作。

内容四：假设检验。很多情况下，我们不仅需要描述所调查群体描述性的一般性特征，还需要知道这些数据能否推论到更大的群体，比如对于药物使用的效果，我们不仅希望知道对于被研究对象（样本）的适用性，还需要知道对所有对象是否适用，能否推论到其他入。大学毕业的人会比其他人更快乐吗？我们可以对数据进行检验。第十章子总体均值比较与检验、第十四章非参数检验将主要介绍假设检验的操作与过程。

内容五：相关分析。有时我们的研究关心广告费用与销售额之间是一种什么样的关系？可以由出生率预知预期寿命吗？什么因素是对收入高低的较好预测？等问题。本书的第八章交互表与关联分析、第九章引进其他变量后的交互分析、第十一章方差分析、第十二章相关分析、第十三章线性回归分析，将通过实例来介绍不同变量之间的关系。

学习数据分析的最好方法是进行实际操作。在本书中，我们将提供一些真实的数据，你会看到：男性比女性工作更快乐吗？三大地区的收入差距怎么样？上大学的生活成本在不同性别/家庭背景的学生之间有什么样的差异，等等。

与其他同类教材或 SPSS 的操作类书籍相比，本书作者力图以简明、扼要、直观的方式，展示 SPSS 分析的基本原理与操作，并在展示过程中将统计原理与 SPSS 数据分析相结合。好的数据分析是常识、技术、好奇心等等的综合，如何提出好的问题及对其进行最好的解释非常重要。SPSS 作为一个强有力的统计分析工具，能够帮助我们更好地通过数据分析寻求对社会现象的解释，从数据得出更有说服力的结论。

本书的主要内容架构，得益于 2002 年起作者为北京大学社会学系本科生开设的数据分析课程实践。四年的课堂教育以及与学生的互动，使我受益良多，面对那些来自全国最优秀的学子，惟有不断改进和提高自己的教育水平，才是对得起他们的惟一选择。

感谢姚映然女士，是她的促动和帮助，让我有机会将四年多来的教育经验与大家分享，并接受检验与批评。

目 录

前言	1
----	---

第一章 SPSS 与社会科学调查研究	1
一、问题的提出	2
二、提出研究假设	3
三、设计研究方案	4
(一) 研究类型选择	4
(二) 研究方法选择	5
(三) 资料收集方法选择	6
(四) 研究层次和分析单位选择	6
(五) 概念和变量测量方法的确立	7
(六) 问卷、观察表格与访问提纲的制定	9
(七) 制定抽样方案	10
四、资料收集	10
(一) 试点调查	10
(二) 调查实施	10
五、SPSS 和资料的整理与分析	11
(一) 资料的校核与登录	11
(二) 统计分析与假设的检验	11
思考和练习	12

第二章 SPSS 基本知识	13
一、SPSS 的安装	14
二、SPSS 的启动和退出	16
(一) 启动 SPSS	16
(二) 退出 SPSS	18
三、SPSS 主要窗口介绍	18
(一) 数据编辑器窗口: Data Editor	18
(二) Syntax: 编辑窗口	20
(三) Output: 结果输出窗口	21
四、SPSS 的求助系统	23
(一) Topics 主题帮助子菜单	23

(二) Tutorial 在线帮助子菜单	24
(三) Case Studies 帮助子菜单	26
(四) Statistics Coach 统计分析帮助子菜单	28
思考 and 练习	28

第三章 数据录入 29

一、问卷设计及编码	30
(一) 问卷及编码样例	30
(二) 编码应注意事项	31
二、数据录入方法	31
(一) 在 SPSS 窗口录入数据	31
(二) 用 SPSS 命令程序录入数据	37
(三) 在 EXCEL 中录入数据	41
(四) 利用 EpiData 小程序进行录入	41
思考 and 练习	51

第四章 SPSS 数据文件的编辑与管理 53

一、增加新的变量(Insert Variable)	54
二、增加新的个案(Insert Case)	54
三、删除变量或个案	54
四、查找个案(Go To Case)	55
五、数据的排序(Sort Case)	55
六、数据的行列对换(Transpose)	57
七、数据文件的合并(Merge Files)	59
八、数据文件的分类汇总(Aggregate)	60
九、数据的选择(Select Cases)	67
十、对变量进行加权(Weight Cases)	69
十一、Data 中的其他菜单命令	72
思考 and 练习	72

第五章 用 Frequencies 做数据汇总 73

一、如何得到一张频次表	74
(一) 基本操作	74
(二) 频次表的解释	76
二、频次统计及统计量的计算	76
(一) 百分比计算	77
(二) 集中趋势测量	78
(三) 离散趋势测量	80

(四) 数据分布	81
(五) 组中值代表各组数据	82
三、条形图、饼图和直方图	82
(一) Bar Charts 条形图	82
(二) Pie Charts 饼图	83
(三) Histogram Charts 直方图	84
四、Frequencies 过程的命令	84
思考和练习	85

第六章 描述性统计——Descriptives 与 Explore 的应用 87

一、Descriptives 过程的应用	88
(一) Descriptives 对话框介绍	88
(二) 在对话框中进行 Descriptives 的描述统计	90
(三) 通过命令运行 Descriptive 过程	90
二、Explore 过程的应用	92
(一) 数据与变量介绍	92
(二) 打开 Explore 对话框	92
(三) 统计量的选择——Statistics 对话框介绍	93
(四) 图形的选择——Plots 对话框介绍	96
(五) 在 Syntax 窗口编辑	100
思考和练习	100

第七章 数据变换 102

一、用 Recode 对数据重新编码	103
(一) 运用对话框对数据进行重新编码	103
(二) 运用命令对数据进行重新编码	107
二、用 Compute 命令创建新变量	108
(一) 运用对话框创建新变量	108
(二) 运用命令对创建新变量	111
(三) 应用: 中国 31 个省市社会经济发展排行榜	112
三、用 Count 命令计算有相同情况的次数	117
(一) 在对话框中运行 Count 过程	118
(二) 用 Count 命令计次	120
四、变量值秩次确定	121
(一) Rank Cases 对话框介绍	121
(二) Rank Cases 实例操作	122
五、自动对变量重新赋值	123
六、缺少值的替代	124

思考和练习	125
-------	-----

第八章 交互表与关联分析(Crosstabs 的应用)	126
一、交互表的一般形式及其特点	127
二、如何获得交互表	127
(一) 基本操作	127
(二) 如何看一张交互表	129
三、期望值、百分比和残差	131
(一) 期望值	131
(二) 百分比	132
(三) 残差	135
四、定类一定类型变量的交互分析	136
(一) 交互表的检验——两个变量之间是否相关	136
(二) 交互表检验实例	137
五、两个定类变量之间相关的强度	138
(一) 相关系数介绍	138
(二) 选择最适合的相关系数进行测量	140
(三) 实例	141
六、两个定序变量之间相关的强度	141
(一) 相关系数介绍	141
(二) 实例	143
七、关于 Kappa 复选框的含义与操作	143
(一) Kappa 复选框的含义	143
(二) 操作方法	143
八、关于 Risk 框的含义及其操作	144
(一) Risk 框的含义	144
(二) 操作方法	145
九、Crosstabs 的总命令表	146
思考和练习	147

第九章 引进其他变量后的交互分析	149
一、因果分析	150
(一) 因果分析原理	150
(二) 用 SPSS 探寻因果关系	152
二、阐明分析	155
(一) 阐明分析原理	155
(二) 用 SPSS 进行阐明分析	156

三、条件分析	159
(一) 条件分析原理	159
(二) 用 SPSS 进行条件分析	159
思考和练习	162

第十章 子总体均值比较与检验 (Means 和 T Test 的应用) 163

一、子总体均值的比较 (Means 过程)	164
(一) Means 过程对话框介绍	164
(二) Means 过程窗口操作实例	166
二、单样本 T 检验 (One-Sample T Test)	167
(一) 关于检验的统计知识	167
(二) One-Sample T Test 窗口操作实例	168
三、独立样本 T 检验 (Independent-Sample T Test)	170
(一) 在对话框进行独立样本 T 检验	170
(二) 扩展案例	173
(三) 独立样本 T 检验的命令	175
四、配对样本 T 检验 (Paired-Sample T Test)	175
(一) 在对话框进行配对样本 T 检验	175
(二) 配对样本 T 检验的命令	177
思考和练习	177

第十一章 方差分析 (One Way ANOVA) 179

一、方差分析的原理	180
(一) 方差分解	180
(二) 方差分析的检验	180
(三) 方差分析表	181
(四) 方差齐性假设	181
二、用 SPSS 进行单因素方差分析	181
(一) 单因素方差分析对话框介绍	181
(二) 单因素方差分析实例操作	184
(三) 单因素方差分析的命令	187
三、双因素方差分析	188
(一) 双因素方差分析原理	188
(二) 双因素方差分析操作实例	190
(三) 双因素方差分析的命令	197
思考和练习	199

第十二章 相关分析(Correlate 过程分析)	201
一、双变量相关分析(Bivariate Correlation)	202
(一) Bivariate 过程对话框介绍	202
(二) Bivariate 过程窗口操作实例	204
(三) Bivariate 过程命令	206
二、净相关分析(Partial correlation)	206
(一) Partial 过程对话框介绍	207
(二) Partial 过程窗口操作实例	208
三、距离分析(Distance)	209
(一) Distance 过程对话框介绍	210
(二) 个案之间比较:Distance 过程窗口操作实例	214
(三) 变量之间比较:Distance 过程窗口操作实例	215
(四) Distance 过程命令	216
思考和练习	217
第十三章 线性回归分析(Linear Regression)	219
一、一元线性回归模型及其对变量的要求	220
(一) 一元线性回归模型	220
(二) 对变量的要求:回归分析的假定条件	220
(三) 数据要求	221
二、在对话框中做一元线性回归模型	221
(一) 对两个变量进行描述性分析	221
(二) SPSS 线性回归主对话框介绍	222
(三) 回归分析统计量选择	226
(四) 回归分析假定条件的检查	230
三、多元线性回归分析	233
(一) 数据与变量说明	234
(二) 虚拟变量的建立	234
(三) 回归方程的建立	235
(四) 在对话框中进行多元线性回归分析	235
(五) 线性回归分析命令	238
思考和练习	239
第十四章 非参数检验(Nonparametric Tests)	242
一、Chi-square Test:卡方检验	244
(一) 主要功能	244
(二) 卡方检验实例	244

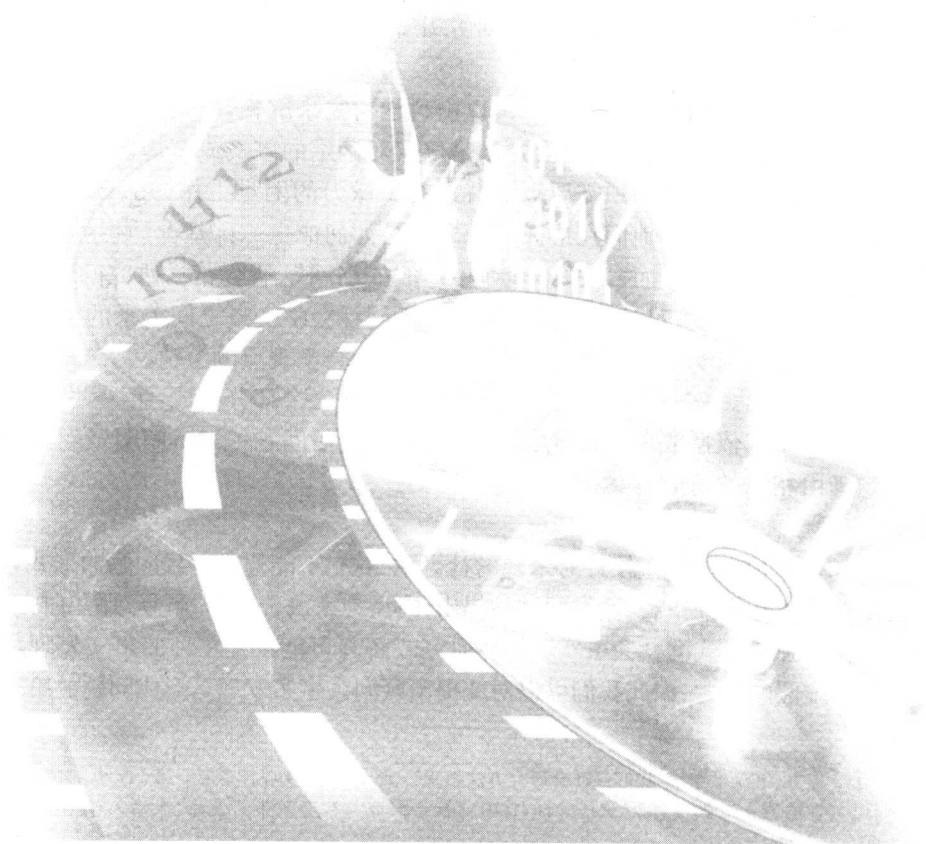
二、Binomial Test:二项分布检验	247
(一) 主要功能	247
(二) 二项分布检验实例	247
三、Runs Test:游程检验	248
(一) 主要功能	248
(二) 游程检验实例	249
四、1-Sample K-S Test:单样本 K-S 检验	251
(一) 主要功能	251
(二) 单样本 K-S 检验实例	251
五、2 Independent Samples Tests:双样本 K-S 检验	253
(一) 主要功能	253
(二) 双样本 K-S 检验实例	254
六、K Independent Samples Tests:多个独立样本的秩和检验	257
(一) 主要功能	257
(二) 多个独立样本秩和检验实例	257
七、2 Related Samples Tests:两个相关样本的检验	259
(一) 主要功能	259
(二) 两个相关样本检验实例	260
八、K Related Samples Tests:多个样本的秩和检验	262
(一) 主要功能	262
(二) 多个样本秩和检验操作实例	262
思考和练习	265

第十五章 SPSS 统计图形	268
一、条形图	269
(一) 条形图对话框介绍	269
(二) 制作不同类型的条形图	273
(三) 条形图的编辑	277
二、线图	280
(一) 线图对话框介绍	280
(二) 线图制作实例	281
三、面积图和饼图	284
(一) Area 面积图	284
(二) 饼图	285
四、高低图和帕雷托图	286
(一) 高低图	286
(二) 帕雷托图	289
五、控制图和箱图	290

(一) 控制图	290
(二) 箱图	293
六、直方图和散点图	295
(一) 直方图	295
(二) 散点图	297
七、P-P 概率图和 Q-Q 概率图	298
(一) P-P 概率图	298
(二) Q-Q 概率图	301
八、普通序列图和时间序列图	302
(一) 普通序列图	302
(二) 时间序列图	304
九、误差条图和 ROC Curve	306
(一) 误差条图	306
(二) ROC Curve	308
思考和练习	308

第一章

SPSS 与社会科学调查研究



社会研究的目的是对社会现象进行描述、解释和预测。为了达到上述目的可以有各种方法,当代社会科学主要采用科学方法,该方法建立在逻辑和观察两大支柱之上。社会科学调查研究的历程,大致可分为六个步骤:问题的提出、确定课题或假设、设计研究方案、收集资料、整理和分析资料、得出结论或解释。整个过程可以用一个简洁的图形来表达(见图 1-1)。

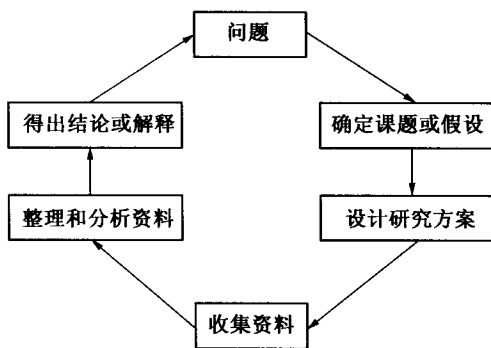


图 1-1 社会科学调查研究的一般步骤

一、问题的提出

所谓问题,是指在每一个科学研究开始时,必定存在的令研究者不解的一个或一系列事实。

所要研究的问题必须是:(1)可以由科学研究来解答,(2)具有一定的研究价值。^①

社会学家默顿指出,提出问题不是仅仅对确定的事实或事物常规地加上一个“为什么”,社会学问题的提出是一项需要训练有素的想像力的困难工作。那么,社会学家在提出一个问题时是如何做的?他们在形成问题之前是如何产生一种基本的想法的?

对于社会学研究而言,并不存在单一的提出问题的套路。研究问题的提出可以是对社会事实证据的怀疑,对社会生活领域中的不同社会学变量之间关系的分析、对一个曾经很有启发性但看来与所期望解释的事实有差异的社会学观点的澄清。

默顿指出,一个问题演化形成的过程主要由三个部分构成:第一,发问、陈述想解决的问题;第二,基本依据或理由,说明为什么要解答这一问题;第三,阐明疑问,即对提出的疑问找出与理由相符的可能答案。^②

(1) 发问

如果一种方法是直接针对未知的社会事实,它不能构成一种问题的形式,但却是问题形成的重要一步。默顿指出,这类问题的提出,典型地来源于试图“解释”那些研究者认为满意地提出的作为独特形式的社会事实。一些研究或解释可能只是单纯的事实发现,虚假的事实往往会导致虚假的问题,因此我们可以对此提出疑问。事实发现是形成问题的第一个阶段。

^① 袁方、王汉生主编,1997:《社会研究方法教程》,北京:北京大学出版社,第 117 页。

^② 罗伯特·K. 默顿著,林聚任等译,2001:《社会研究与社会政策》,北京:生活·读书·新知三联书店,第 19—47 页。