

教育心理多元统计学 与 SPSS 软件

梁荣辉 章 炼 封文波 编著

 北京理工大学出版社
BEIJING INSTITUTE OF TECHNOLOGY PRESS

教育心理多元统计 与 SPSS 软件

—— 第二版 ——

陈永发 王健 主编

教育心理多元统计学 与 SPSS 软件

梁荣辉 章 炼 封文波 编著

 北京理工大学出版社

BEIJING INSTITUTE OF TECHNOLOGY PRESS

内 容 简 介

本书结合教育、心理学科的特点,介绍在教育、心理研究领域中常用的多元统计方法,其中包括方差分析、相关分析、回归分析、聚类分析、判别分析、因子分析和对应分析,并将目前较为流行的SPSS软件的使用与多元统计方法的介绍有机地结合起来,方法的应用均给出了具体的计算实例。为了使读者容易理解多元统计分析方法并保持知识的连续性,对基本统计知识也进行了简要介绍。

本书可作为教育专业、心理专业本科生、研究生教材,也可供教育、心理学科的教师、科研工作者参考;本书虽以教育、心理学科作为实际背景,但对其他专业的学生、教师及科研工作者,也不失其参考价值。

版权专有 侵权必究

图书在版编目(CIP)数据

教育心理多元统计学与SPSS软件/梁荣辉,章炼,封文波编
著. —北京:北京理工大学出版社,2005.3

ISBN 7-5640-0494-0

I. 教… II. ①梁…②章…③封… III. 教育统计-多元分
析-应用软件, SPSS-高等学校-教材 IV. G40-051

中国版本图书馆CIP数据核字(2004)第135726号

出版发行/北京理工大学出版社

社 址/北京市海淀区中关村南大街5号

邮 编/100081

电 话/(010) 68914775(办公室) 68944990(发行部)

网 址/<http://www.bitpress.com.cn>

电子邮箱/chiefedit@bitpress.com.cn

经 销/全国各地新华书店

印 刷/北京圣瑞伦印刷厂

开 本/787毫米×1092毫米 1/16

印 张/14.5

字 数/338千字

版 次/2005年3月第1版 2005年3月第1次印刷

印 数/1~3000册

责任校对/陈玉梅

定 价/27.00元

责任印制/吴皓云

图书出现印装质量问题,本社负责调换

前 言

随着社会的进步与科学技术的飞速发展，要求人们掌握更多的知识与强有力的科研工具，更深入地探索、发现、研究新问题，揭示事物的内在规律性，更好地指导实践。对于从事教育、心理研究的工作者，教育统计无疑是必不可少的、强有力的研究工具，特别是多元统计，随着科学化、定量化程度要求的提高，其地位愈来愈重要。为了使广大从事教育、心理工作者更好地掌握这一工具，本书结合教育、心理学科介绍了多元统计的基本原理与方法。

多元统计的学习，最好与一种统计软件的学习相结合，以达到事半功倍的效果，避免使所学知识与应用脱节。目前出版的有关多元统计书籍主要有两类：一类着重介绍统计软件，附带统计知识；另一类只介绍统计知识，不介绍统计软件。从事多元统计教学、科研的工作者，需要的是系统地介绍多元统计知识、并辅以介绍合适的统计软件的书籍，本书正是为满足这一需要编写的。在统计内容的选择、安排上，主要介绍多元统计部分，但考虑到系统性、完整性，简单介绍了基本统计内容，但不作为重点；在统计软件的选择、安排上，考虑到先进性、实用性，介绍了目前较为流行的 SPSS 统计软件。

全书在叙述上力求科学性强、概念清楚，注重理论密切联系实际；在文字上做到精炼、通俗易懂；在例题的选配上考虑其代表性与实用性。

在本书的编写过程中，我们参考了相关统计学与 SPSS 软件著作，引用了其中的一些例题，在此，谨向这些作者表示谢意。

由于水平所限，书中不足之处在所难免，敬请读者批评、指正。

第六章 相关分析	84
第一节 简单相关分析	84
第二节 偏相关分析	87
第三节 距离分析	90
第七章 回归分析	91
第一节 线性回归分析	91
第二节 曲线回归分析	113
第三节 逻辑回归分析	124
第四节 其他回归分析	131
第八章 聚类分析	132
第一节 概述	132
第二节 系统聚类法	133
第三节 快速样品聚类法	152
第九章 判别分析	165
第一节 概述	165
第二节 SPSS 软件与应用实例	165
第十章 因子分析	198
第一节 概述	198
第二节 SPSS 软件与应用实例	199
第十一章 对应分析	221
第一节 概述	221
第二节 SPSS 软件与应用实例	221
参考文献	226

第一章 绪 论

教育心理多元统计学是将多元统计的理论与方法应用于教育、心理领域所形成的一门学科，所研究的对象是教育、心理研究领域中的随机现象，是教育、心理研究的工作者必不可少研究工具。

第一节 统计内容

本书将统计知识分为一般统计与多元统计两部分。前者包括统计资料的收集与整理、常用统计量、相对数与动态分析、正态分布的理论及其应用、假设检验（含参数与非参数检验）、参数估计、一元回归分析、方差分析、实验设计、统计图与统计表。这部分内容本书只介绍假设检验中常用的 U 检验与 T 检验、方差分析。后者包括相关分析、回归分析、聚类分析、判别分析、主成分分析、因子（素）分析、对应分析与典型相关分析。

典型相关分析是研究两组变量之间关系的多元统计分析方法。为了要找出两组变量之间的关系，要计算出 X 变量间的线性组合与 Y 变量间的线性组合，并使这两组的线性组合有最大的相关；X 与 Y 的线性组合是潜在的，无法直接观察，也是未知变量，称为“典型相关（Canonical variable）”。该方法的学习与使用可参考有关数理统计或统计软件书籍。

第二节 统计软件

目前常用的统计软件有：中科院的统计软件、STATISTICA 生物学用统计软件、SPSS 统计软件（Statistical Product and Service Solutions——统计产品与服务解决方案，原名为社会科学统计软件包）、SAS 统计软件、医学专用统计软件包、自编软件等。

本书选用目前较为流行的 SPSS 软件，它是目前世界上最优秀的统计分析软件之一，广泛应用于经济、金融、证券、生物、医药、教育、体育、农林等各个研究领域，深受广大用户的好评。

第三节 注意事项

收集的数据要准确、可靠。

注意每种统计方法的适用条件，不要生搬硬套。

所得到的结论不是百分之百的正确，有可能犯错误。

样本含量 n 要适当大。统计学科探索的是统计规律，太小时样本的代表性不高，有时反映的是极端情况，这样会使结论失真，因此， n 要适当大。

对于多元统计知识的学习，要求搞清楚方法的原理，知道哪类问题用哪一种方法，能正确使用 SPSS 软件进行计算，并能结合专业知识对计算结果进行分析，作出正确的结论。对

于复杂的计算、证明过程可以先略过，需要时再参考有关书籍。

大型研究课题要制定一个研究设计，主要包括实验的目的、对象、指标、方法、组织工作等内容。有时在实验过程中出现的需要花费很大气力才能解决的问题，在设计时略加改进就能解决或避免，因此，不能轻视研究设计。

本书的重点与难点，在于理解统计方法的原理、用法、结果分析，以及如何作出正确的结论等方面，需要读者花费一定的时间和精力。至于 SPSS 软件，会用即可，不必花费更多的时间和精力。

第二章 SPSS 软件基本功能简介

SPSS 软件经历了 for DOS、for Windows 6.0、7.0、8.0、9.0、10.0、11.0 几个版本（在书稿即将交付印刷时，12.0 版本问世，但尚未见到相关的书籍），随着版本的更新，功能愈来愈完善，操作愈来愈简便。本书选用了 11.0 版本（部分章节用到了 12.0 版本），较以前几个版本，它增强了部分模块，提高了运行速度，增加了一些新的统计方法。

SPSS 软件对硬件与操作系统的要求不高，目前市场上主流的计算机与操作系统均能满足要求。

第一节 启动 SPSS

一、启动 SPSS

进入 SPSS 系统，弹出如图 2.1.1 所示的界面，可根据需要进行选择。

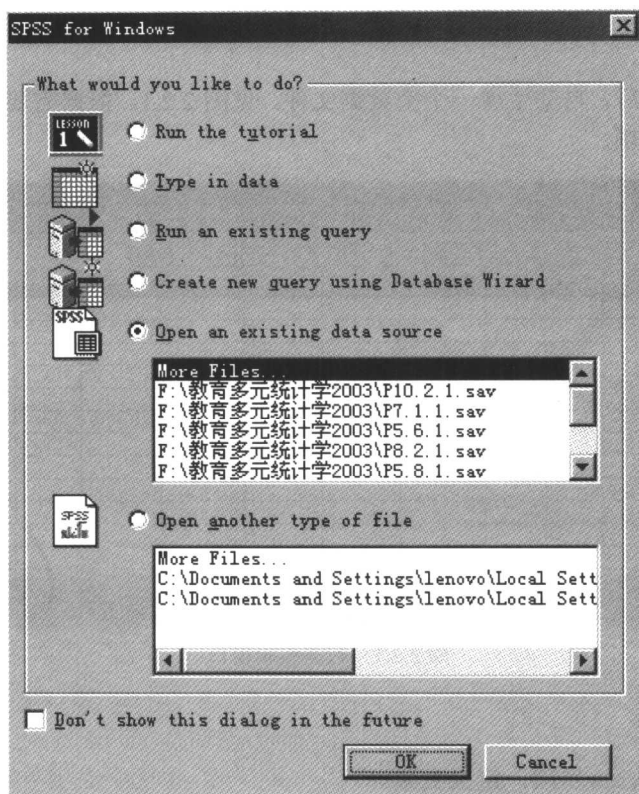


图 2.1.1 SPSS 系统界面

二、SPSS 系统界面

1. Run the tutorial

浏览操作指导。

2. Type in data

显示数据编辑窗口，建立新数据文件。

3. Run an existing query

运行一个已存在的文件。

4. Create new query using Database Wizard

使用数据库向导建立一个新的文件。

5. Open an existing data source

打开一个已存在的数据文件，可选择“.sav”文件。

6. Open another type of file

打开一个已存在的其他类型的文件。

7. Don't show this dialog in the future

再运行 SPSS 系统时，不显示图 2.1.1 系统界面，直接进入数据编辑窗口。

第二节 数据编辑

进入 SPSS 系统后，自动生成一个空数据文件，见图 2.2.1。根据需要，选择“Data View”或“Variable View”项。

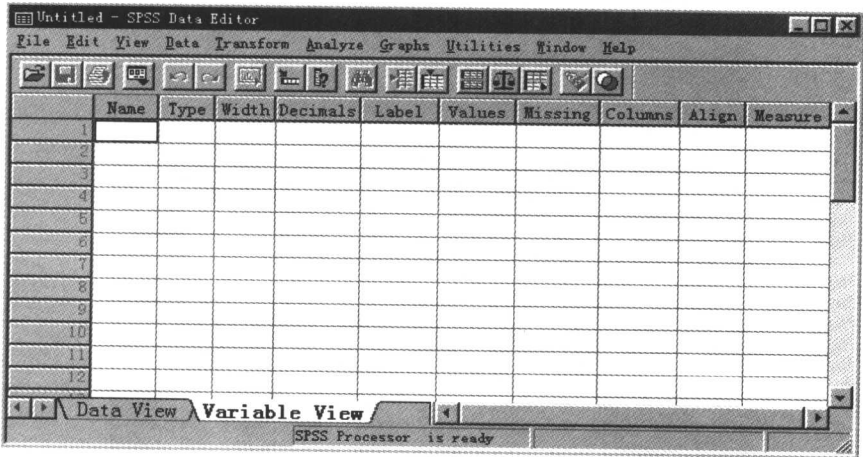


图 2.2.1 数据编辑窗口

一、定义变量

选择“Variable View”项，定义变量。

1. Name

变量名，不超过 8 个字符，首字符为字母或汉字。

2. Type

类型，共有 8 种类型供选择：

(1) Numeric

标准数值型。

(2) Comma

显示逗号的数值型。

(3) Dot

显示句点的数值型。

(4) Scientific notation

科学计数型。

(5) Date

日期型。

(6) Dollar

美元型。

(7) Custom currency

自定义型。

(8) String

字符串型。

3. Width

变量的长度。

4. Decimals

小数位数。

5. Label

变量名标签，起注释作用。

6. Values

变量值标签，对变量可能取值进一步描述。

7. Missing

定义缺失值。

8. Columns

显示变量的长度。

9. Align

变量的对齐方式：

Left（左对齐）、Right（右对齐）、Center（居中对齐）。

10. Measure

变量的测量尺度，按测量精度可以分为以下几种变量：

Nominal（定性或名义变量）、Ordinal（定序或有序变量）、Scale（间隔或比率变量）。

二、编辑数据

1. 输入数据

在空数据文件界面上选择“Data View”项。

输入的数据要及时存盘，避免丢失。

2. 常规修改

将光标移至错误数据处改正即可。

3. 利用剪切、复制与粘贴功能修改

使用“Edit”菜单或右键调出“剪切、复制与粘贴”功能，可以移动数据位置、拷贝相同的数据。

三、变量、观测值的插入与删除

1. 插入变量

选择“Data”→“Insert Variable”项，或者将光标移至列变量上，击右键选择“Insert Variable”项。

2. 删除变量

首先要选中被删除的变量，使用“Edit”菜单或右键中的“Cut”或“Clear”，或者用“Delete”键。

3. 插入观测值

选择“Data”→“Insert Case”项，或者将光标移至行序号上，击右键选择“Insert Case”项。

4. 删除观测值

首先要选中被删除的观测值，使用“Edit”菜单或右键中的“Cut”或“Clear”，或者用“Delete”键。

四、恢复删除或修改前的数据

选择“Edit”→“Undo”项，或者按工具栏中的“Undo”按钮。

第三节 数据整理

一、数据排序

在数据文件中，根据实际需要，可对一个或多个变量的值重新排序。

选择“Data”→“Sort Cases”项，将排序的变量移至“Sort by”框中，在“Sort Order”项中选择“Ascending”（升序）或“Descending”（降序）。

若排序的变量多于一个，则按选择的先后顺序依次排序。

二、分类汇总

在整个数据文件中，有时需要根据某个指定的变量进行分类汇总。如按性别、班级等进行分类汇总。

例 2.3.1 某年级有 100 名学生，男生 60 人分 3 个班，女生 40 人分 2 个班，试对其某一门课考试成绩分性别、班级计算平均数（见表 2.3.1）。

表 2.3.1 100 名学生考试成绩

性别	班级	考试成绩
男生	1	82, 72, 74, 86, 93, 74, 79, 83, 78, 81
		84, 75, 87, 78, 85, 74, 84, 78, 79, 83
	2	81, 77, 74, 80, 70, 74, 72, 90, 84, 87
		84, 79, 77, 69, 80, 79, 81, 80, 88, 78
	3	81, 65, 73, 77, 67, 76, 80, 81, 75, 76
		80, 75, 88, 76, 79, 80, 65, 89, 77, 68
女生	1	75, 73, 89, 72, 70, 74, 77, 70, 61, 91
		85, 82, 94, 85, 82, 86, 91, 83, 87, 77
	2	85, 83, 68, 79, 82, 87, 86, 81, 79, 83
		75, 84, 73, 96, 67, 88, 63, 89, 71, 94

用 x 表示考试成绩，xb 表示性别，“1”表示男生，“2”表示女生，bj 表示班级。
选择“Data”→“Aggregate”项，弹出如图 2.3.1 所示的对话框。

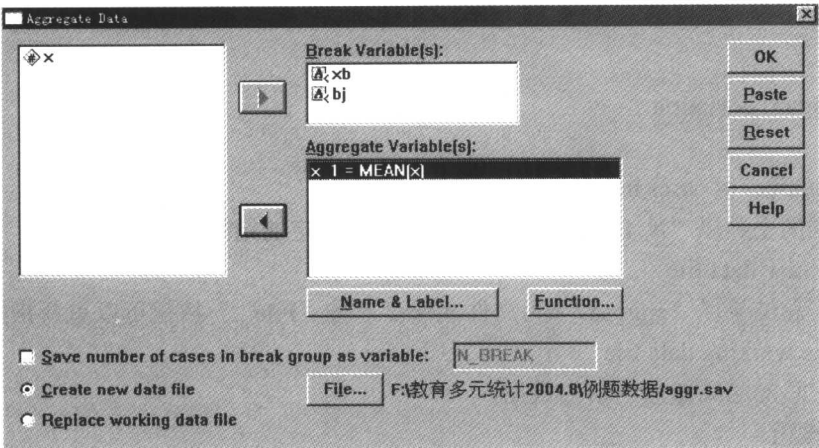


图 2.3.1 分类汇总对话框

1. Break Variables 框

存放分类变量，如 xb、bj。

2. Aggregate Variables 框

存放汇总变量，如 x。

3. Name & Label...按钮

重新指定结果文件中的变量名及其标签，若不指定，默认的变量名为原变量名后加“_1”。

4. Function...按钮

该对话框选项较多，见图 2.3.2，最常用的有以下几项：

(1) Mean

平均数。

(2) Median

中数。

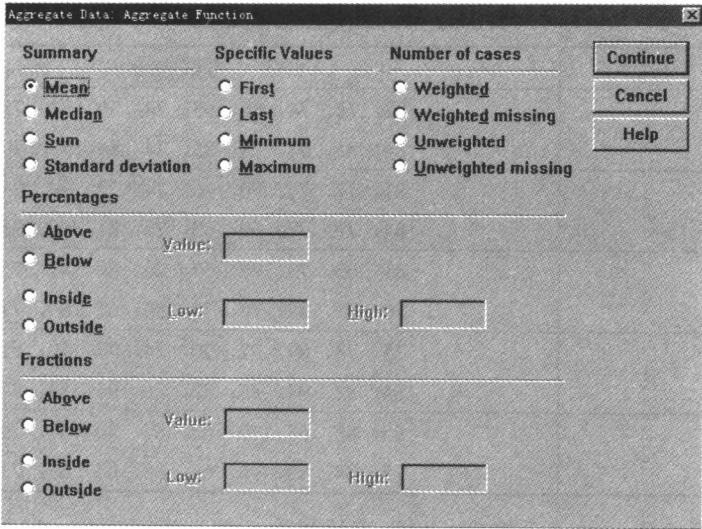


图 2.3.2 分类汇总函数对话框

(3) Sum

和数。

(4) Standard deviation

标准差。

5. Save number of cases in break group as variable

分组数据的个数以 “N_BREAK” 为变量名存入指定的结果文件中。

6. Create new data file

将分类汇总结果以 “aggr.sav” 为文件名保存，按 “File...” 按钮可以重新指定文件名。

7. Replace working data file

将分类汇总的结果覆盖当前的数据。

8. Paste 按钮

生成 SPSS 语句。

9. Reset 按钮

重新选择。

本例选中 “Save number of cases in break group as variable” 与 “Replace working data file”

项。

计算结果见表 2.3.2。

表 2.3.2 分类汇总结果

xb (性别)	bj (班级)	x_1 (平均数)	n_break (个数)
1	1	80.45	20
1	2	79.20	20
1	3	76.40	20
2	1	80.20	20
2	2	80.65	20

三、数据变换

例 2.3.2 10 名学生的两门期中考试成绩见表 2.3.3，按占总成绩的 40%计算总分。

表 2.3.3 10 名学生两门期中考试成绩

序号 (xu)	教育学 (x1)	心理学 (x2)	总成绩的 40% (x)
1	78	69	58.80
2	65	92	62.80
3	85	86	68.40
4	72	76	59.20
5	68	94	64.80
6	82	89	68.40
7	63	70	53.20
8	87	75	64.80
9	80	93	69.20
10	77	82	63.60

用 xu 表示序号，x1 表示教育学，x2 表示心理学。

选择 “Transform” → “Compute” 项，弹出如图 2.3.3 所示的对话框。

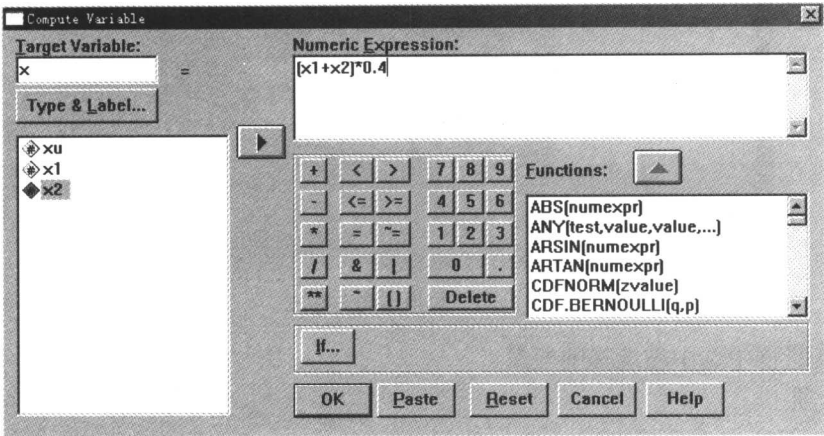


图 2.3.3 数据变换对话框

- 1. Target Variable 框
输入新变量名，如 x。
- 2. Numeric Expression 框
输入或选择表达式值，如 (x1+x2)*0.4。
计算结果见表 2.3.3。
- 3. Type&Label...按钮
输入新变量后，对该变量命名标签与设置类型。
- 4. Functions 项
系统预装的函数。

5. if...按钮

打开条件表达式对话框。

(1) Include all cases

包括所有观测值。

(2) Include if cases satisfies condition

包括符合条件的观测值。

四、数据选择

收集的数据应该尽量多而全，但针对某个实际问题进行计算时，可以根据需要对数据进行选择。

选择“Data”→“Select Cases”项，弹出如图 2.3.4 所示的对话框。

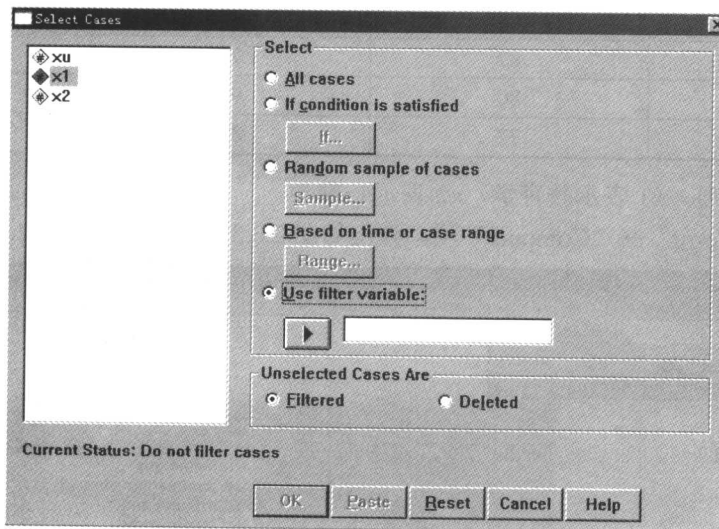


图 2.3.4 数据选择对话框

在左侧变量框中选中被选择的变量。

1. Select 项

(1) All cases

选择所有数据。

(2) If condition is satisfied

按指定条件选择。

(3) Random sample of cases

随机抽样。

① Approximately

大概抽样。

② Exactly

精确抽样。

(4) Based on time or case range

顺序抽样。

(5) Use filter variable

用指定的变量过滤, 对该变量值为 0 的数据作删除标记。

2. Unselected Cases Are 项

(1) Filtered

过滤 (不删除)。

(2) Deleted

删除有删除标记的数据。

第四节 文件编辑

一、打开 SPSS 文件

选择 “File” → “Open” 项。

1. Data

数据文件 (.sav)。

2. Syntax

源程序文件 (.sps)。

3. Output

结果输出文件 (.spo)。

4. Script

脚本文件 (.sbs)。

5. Other

其他文件。

二、打开其他类型文件

常用的是打开文本文件 (.txt)、“Excel” 文件 (.xls)。

1. 文本文件

选择 “File” → “Open” → “Data” 项, 文件类型选择 “Text”, 弹出对话框时取默认值。

2. Excel 文件

选择 “File” → “Open” → “Data” 项, 文件类型选择 “Excel”, 弹出对话框, 若文件中第一行是数据, 则不选 “Read variable names from the first row of data” 项。

三、文件合并

1. 纵向

首先打开第一个数据文件, 然后选择 “Data” → “Mergr Files” → “Add Cases” 项, 再打开第二个数据文件, 弹出如图 2.4.1 所示的对话框。