



SPSS 统计分析实验指导

主 编 庄 致 冯运义
副主编 肖 伟 朱兴亮 刘 永



电子科技大学出版社

SPSS 统计分析实验指导

主 编 庄 致 冯运义
副主编 肖 伟 朱兴亮 刘 永

图书在版编目(CIP)数据

SPSS 统计分析实验指导 / 庄致, 冯运义主编. —成都:

电子科技大学出版社, 2011. 4

ISBN 978-7-5647-0794-1

I. ①S… II. ①庄… ②冯… III. ①统计分析—软件

包, SPSS—高等学校—教材 IV. ①C819

中国版本图书馆 CIP 数据核字 (2011) 第 038986 号

内 容 提 要

本书是为经济管理类本科学生《统计学》《管理统计学》及《应用统计学》专门编写的实验教材, 也可以作为其他学生学习统计分析(SPSS)的参考用书和补充教材, 具备一定的统计理论基础知识的读者均可阅读。

本书主要包括统计分析流程及建议、SPSS 简介、统计分析实验指导、统计分析方法分类及选择四部分。

本书主要内容为统计分析实验指导, 在此基础上, 汇总了统计应用过程中的方法选择原则, 帮助读者深刻理解统计思想, 切实把握统计方法原理和适用条件, 减少统计应用中的错误, 提升统计应用能力。

SPSS 统计分析实验指导

主 编 庄 致 冯运义

副主编 肖 伟 朱兴亮 刘 永

出 版: 电子科技大学出版社(成都市一环路东一段 159 号电子信息产业大厦 邮编: 610051)

策划编辑: 万晓桐

责任编辑: 万晓桐

主 页: www.uestcp.com.cn

电子邮件: uestcp@uestcp.com.cn

发 行: 新华书店经销

印 刷: 成都火炬印务有限公司

成品尺寸: 185mm×260mm 印张 9.25 字数 225 千字

版 次: 2011 年 4 月第一版

印 次: 2011 年 4 月第一次印刷

书 号: ISBN 978-7-5647-0794-1

定 价: 16.00 元

■ 版权所有 侵权必究 ■

◆ 本社发行部电话: 028-83202463; 本社邮购电话: 028-83208003。

◆ 本书如有缺页、破损、装订错误, 请寄回印刷厂调换。

目 录

第一章 统计分析流程及建议	1
第一节 统计问题及统计方法	1
第二节 统计分析流程	1
第三节 学习和应用统计的注意事项	3
特别提醒	4
第二章 SPSS 简介及使用	5
第一节 SPSS 简介	5
一、SPSS 软件的历史	5
二、SPSS 软件的优点	6
三、SPSS 软件的功能	6
第二节 SPSS 统计分析的一般流程	8
一、数据的输入和保存	8
二、数据预分析	9
三、按要求进行统计分析	11
四、保存和导出分析结果	12
第三节 SPSS 的数据整理	13
一、用 SPSS 输入和处理数据	13
二、用 SPSS 计算常用的描述统计指标	19
三、用 SPSS 制作常用统计图	22
第三章 统计分析及 SPSS 应用实验指导	31
试验一 SPSS 数据录入与整理	31
一、试验目的	31
二、试验内容	31
三、试验步骤和结果	32
四、实验总结	43
试验二 基本统计分析	43
一、试验目的	44
二、试验内容	44
三、试验步骤及结果	44
四、试验总结	57
试验三 SPSS 均数比较	57

一、试验目的	57
二、试验内容	57
三、试验步骤和结果	58
四、试验总结	62
试验四 非参数检验	62
一、试验目的	62
二、试验内容	62
三、试验步骤和结果	64
四、试验总结	73
试验五 方差分析	73
一、试验目的	74
二、试验内容	74
三、试验步骤和结果	76
四、试验总结	86
试验六 回归分析	86
一、试验目的	86
二、试验内容	87
三、试验步骤和结果	90
四、试验总结	111
试验七 聚类分析	112
一、试验目的	112
二、试验内容	112
三、试验步骤和结果	112
四、试验总结	116
试验八 判别分析	116
一、试验目的	116
二、试验内容	117
三、试验步骤和结果	117
四、试验总结	123
试验九 因子分析	123
一、试验目的	123
二、试验内容	123
三、试验步骤和结果	123
四、试验总结	127
试验十 对应分析	127
一、试验目的	127
二、试验内容	127
三、试验步骤和结果	127
四、试验总结	130

试验十一 信度分析 130

 一、试验目的 130

 二、试验内容 130

 三、试验步骤和结果 130

 四、试验总结 132

试验十二 典型相关分析 132

 一、试验目的 132

 二、试验内容 133

 三、试验步骤和结果 133

 四、试验总结 134

第四章 统计分析方法分类及选择 135

 第一节 统计方法分类 135

 一、根据统计分析目的不同进行分类 135

 二、根据变量特征的不同进行分类 136

 第二节 统计方法选择原则 136

主要参考文献 139

第一章 统计分析流程及建议

随着经济社会的快速发展和进步，统计对人们的影响越来越大，作用越来越强，人们学习和应用统计的兴趣也越来越浓厚，统计学因此也逐渐成为定量分析学科中的一门显学。

在学习和应用统计的过程中，首先必须清楚如下问题：什么是统计问题？统计分析应该经历怎样的过程？有哪些需要着重注意的事项？本章分三部分来对上述问题进行简单阐述。

第一节 统计问题及统计方法

需要通过收集、整理和分析数据（data）然后才能得到结论的问题都可以归结为统计问题。比如要获取某电视节目当时的收视率，就需要在该节目播出时，利用电话对看电视的人进行随机采访，询问他们正在收看的电视节目。在得到被采访的看电视的总人数和其中观看该节目的人数之后，就可以得知这部分观众中，收看该节目的比例，即粗略的收视率了。之后还要经过统计分析，评估这个收视率的可信度和代表性等。现实生活中，这样的例子也比比皆是。总之，统计问题和其他问题区别的关键要看是否需要收集数据、分析数据并从数据中得到结论。

统计方法则是收集、整理、展示和分析数据的系列方法的总和。从数据中得到结论是统计方法区别于其他分析方法的关键。

统计分析是指运用统计方法及与分析对象有关的知识，从定量与定性的结合上进行的研究活动。它是继统计设计、统计调查、统计整理之后的一项十分重要的工作，是在前几个阶段工作的基础上通过分析从而达到对研究对象更为深刻的认识。它又是在一定的选题下，集分析方案的设计、资料的搜集和整理而展开的研究活动。系统、完善的资料是统计分析的必要条件。运用统计方法、定量与定性的结合是统计分析的重要特征。

第二节 统计分析流程

任何一个数据分析项目，如果按照整个分析过程的流程结构来看，都可以被分解为大致7个阶段——计划阶段、数据收集、数据获取、数据准备、数据分析、结果报告和模型发布阶段。

（1）计划阶段。在分析项目的初始阶段，需要花费大量的时间来设计分析计划，以减少盲目分析，避免资源浪费。在该阶段，要对数据分析的各个行动步骤做好规划，主要任务是弄清楚以下几个问题：

①确定研究问题。从研究分析开始，就确立明确的分析目标是非常重要的。它可以帮助用户合理地计划人员、时间、资源分配，并能指导用户集中精力于研究性分析。

②建立项目预算。

③确定研究范围即确定研究总体和个体。

- ④确定样本的抽取方法。
- ⑤分析评估所需样本量。
- ⑥确定数据收集方式。
- ⑦确定与研究问题相关的数据即确定应该收集个体的哪些数据。
- ⑧确定研究问题的分析方法和分析工具。

(2) 数据收集阶段。如果手头已经有现成的数据,就可以不必再进行数据收集。数据收集的目标、方式完全取决于在上一步中所制订的计划。数据收集方式有很多种,可以是电话式访问,可以是面谈式收集,也可以是拦截式访问。如果是从头进行数据收集,则应当有一份标准问卷。问题的设计不仅要相关,还要能够从中得出有意义的结论。

(3) 数据获取阶段。该阶段的目的是将分散的、原始格式各不相同的数据读入分析工具中,以便分析工具可以对数据进行分析。

(4) 数据准备阶段。该阶段的主要任务是:

- ①清理数据以保证数据的准确性。数据准确性是数据分析结果正确的最基本前提。
- ②对数据进行必要的转换。如生成一些新的字段以供分析,将连续字段离散化、将字符型字段数值化等。目的是将数据结构转换成合适的形式。
- ③填充缺失数据。对各种缺失字段,利用适当的方法进行填补。
- ④对数据进行合并、汇总等。将数据文件进行合并,将个体数据进行汇总,生成各组数据。

(5) 数据分析阶段。利用各种数据分析方法对数据进行分析,得出结论。数据分析阶段又可以分为几个部分:

①预分析:包括概括性统计描述和探索性统计推断两部分,前者是使用统计图和统计表对数据进行更好的理解;而后者则基于对数据的理解开始尝试进行分析,以寻找最终分析模型的雏形。具体使用的方法可以是单因素分析,也可以是简单的多因素分析。

②精确分析:基于上一步得到的各种信息,开始尝试拟合最佳的统计模型,以寻求对数据中所蕴含信息最完美的解释。完成这一部分的工作往往需要统计知识和专业知识互相补充,而所使用的统计方法一般都是多变量方法,甚至是多元统计分析方法。

(6) 结果报告阶段。结果报告的目的是:即整个数据分析项目的结果以一种非学术化的方式表达出来,使得决策者(报告的阅读者)能够快速理解,并基于此分析结果做出决策。报告可以是文本文档、表格、图形或者是网页。

(7) 模型发布阶段。结果报告仅仅是对基于历史数据所建立的模型加以阐述,当需要利用该模型进行预测时,具体的做法可以是在分析软件中加以预测,也可以将生成的模型编译成单独运行的控件或程序,将其模型整合到应用平台中去。该阶段的目标是将分析阶段得到的模型、信息和知识带给机构决策者以便他们能为机构做出更好的规划。

在大多数分析过程中,不一定会经历所有的这7步。例如,根据分析的目的,所需的数据仅是日常工作产生的交易数据,那么就不必再经历“数据收集”阶段,而是直接进入“数据获取”阶段。另外,各阶段之间可能交叉进行。例如,有时在对原始数据进行分析之后,即进入了“数据分析”阶段,突然发现其他数据也是分析所必需的,所以不得不重新返回“数据收集”阶段。又如,在“数据分析”阶段中发现某个字段因其格式不能参与分析,所以需要再进行“数据准备”阶段所做的工作。

在一个数据分析项目结束后，可能因该项目中的新发现和对数据的新的理解，从而引发一个新数据分析项目。

由于统计分析工作通常不能一蹴而就，因此上述流程可能需要反复进行直到能满足实际需要为止。实际统计分析工作可以参照图 1-1 的流程进行。

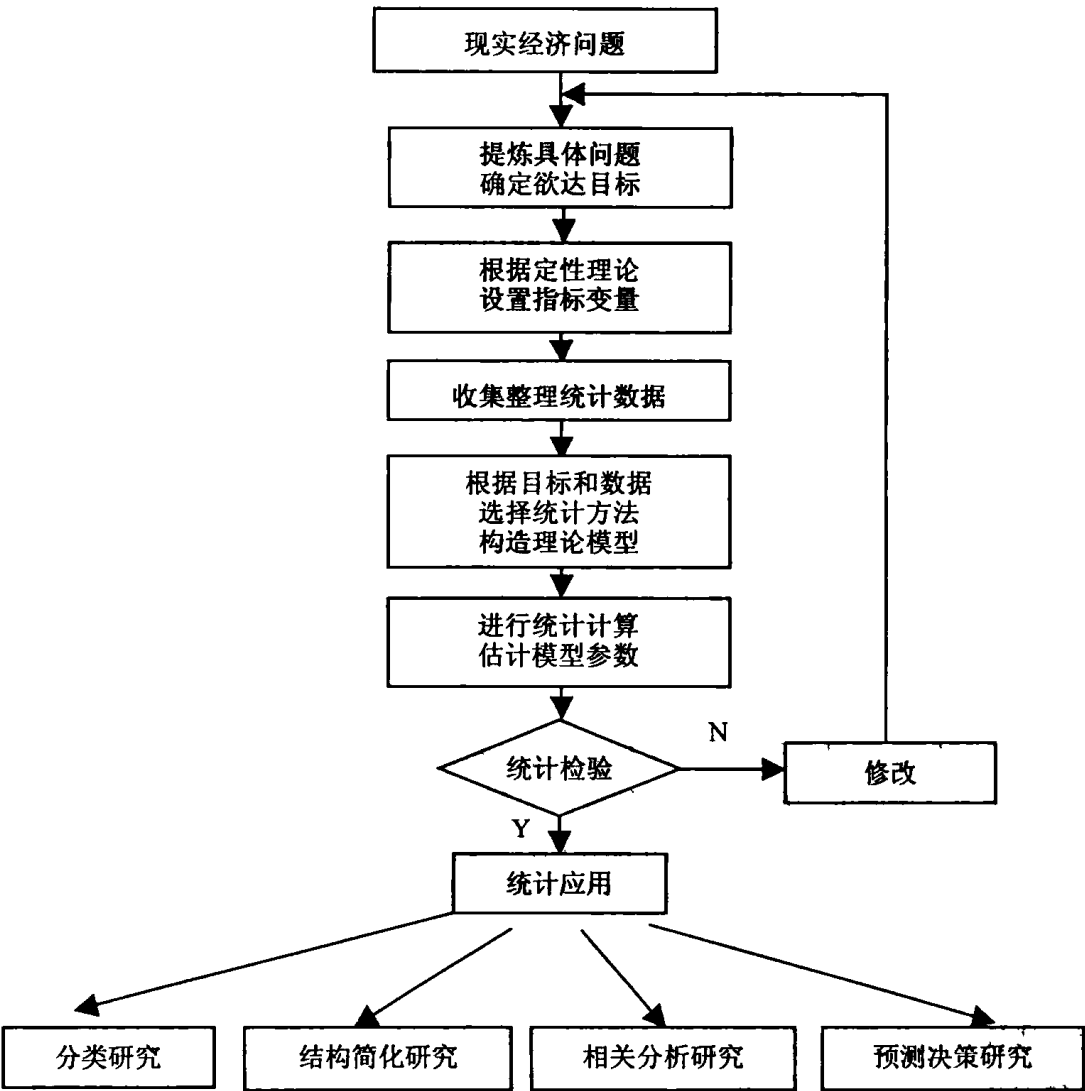


图 1-1 统计分析流程

第三节 学习和应用统计的注意事项

在学习和应用统计的过程中，需要特别避免一些常见的错误，注意学习方法，遵循正确的程序，这样才不至于学无所获，用无所长，进而怀疑统计的功用。

1. 先易后难，由简入繁

作为初学者，刚开始学习时，不要试图立刻就弄清楚软件的全部输出结果。因为有些内容可能已经大大超出了你的学科范围，这样甚至可能影响你的学习兴趣；即便你能够看懂全部输出结果，建议也不要这样做，因为先易后难，由简入繁是经过无数检验证实的最有效

的学习路径。

2. 在战争中学习战术，在应用中学习方法

兴趣是最好的老师，需要是最大的动力。在实践中学习，带着兴趣去学习，根据需要进行学习，这是学习统计分析方法和软件的最好方法。也只有在实践中才能更好把握方法的实质和应用的技巧，迅速取得良好的学习效果。

3. 贪多勿滥，理解为先

统计方法非常多，要正确使用这些方法得到可靠的结论，必须首先深刻理解这些方法。因此尽量不要使用不理解的统计方法，这很容易导致统计方法的误用，带来非常消极的后果；但必须努力学习新方法，因为新方法对所研究的问题可能更加适用。

4. 流沙里建不起高楼，可靠的结论来源于高质量的数据

可靠的结论来源于高质量的数据。全面、系统、准确、可靠和有针对性的数据是正确进行统计分析并进而得到可靠结论的基本前提。因此，必须确保数据质量，才能由此得出可靠的结论。

5. 透彻理解，正确运用

利用统计软件获得分析结果非常容易，但要获得正确的结论却非常困难。这不仅需要对相应统计方法有透彻的理解，对现实问题有深入的研究，而且需要能正确使用统计方法并确保处理程序正确。

6. 去粗取精，画龙点睛

统计软件可以输出很多结果，千万不要不加分析地把软件的全部输出结果直接复制到分析报告中，因为就所研究的问题而言，软件的大部分输出都是不必要的。而且，软件生成的图表往往也需要进行一些特别的处理才能使用。总之，在使用统计软件输出结果时，尽量遵循正确和简洁的原则，用精心选取的图表和数据说明问题，起到画龙点睛的作用。

特别提醒

很多错误的发生其实都源于学习（使用）者自身的问题，尤其是自己对统计思想和方法原理理解不深，对数据质量把握不好和对方法适用条件及其优缺点的忽视，这些都是统计应用的极大障碍。但很多人往往认识不到这一点，总是把问题归咎于工具的优劣和方法的复杂与简易上，执著于使用更复杂的工具，更先进的方法，更高深的理论，而不从自身寻找问题的根源。殊不知，过犹不及，事情的真相恰恰是：简单就是美，合适的才是最好的，需要决定一切。

理论、方法和工具仅仅只是手段，高效地解决问题、满足实际需要才是目的。手段也只能服从和服务于目的。如果我们还迷惘于工具的优劣、方法的简单与复杂和理论的高深与浅易时，请谨记“很多时候，我们在创造复杂的东西，往往代表着我们还缺乏足够的能力去找出深藏于 complexity 和 diversity 背后的 core simplicity”。学习是如此，应用也同样如此。

第二章 SPSS 简介及使用

实用起见，我们推荐大家使用易学好用、功能强大的 SPSS 软件来帮助学习统计分析。

第一节 SPSS 简介

SPSS 统计分析软件是在社会学研究、调查统计、市场研究、医学统计、政府和企业的数据分析应用中久享盛名的统计分析软件。在国际学术界的各个领域，其自动统计绘图、数据的深入分析、易学易用、功能齐全等方面得到了高度的评价和广泛的称赞。在国际学术界有条不紊的规定，即在国际学术交流中，凡是用 SPSS 软件完成的计算和统计分析，可以不必说明算法，由此可见其影响之大和信誉之高。

SPSS Statistics 软件原名 SPSS，是英文名称社会科学统计软件包（Statistical Package for the Social Sciences）首字母的缩写。随着 SPSS 公司产品服务领域的扩大和服务深度的增加，SPSS 公司整个产品线的名称都进行了调整，现在 SPSS 软件的名称全称为“Statistical Product and Service Solutions”，即“统计产品与服务解决方案”，虽然缩写仍为 SPSS，但这却标志着 SPSS 的战略方向做出了重大调整。

随着 SPSS 产品家族的丰富，原为大家熟知的统计分析软件 SPSS，现在全名为 SPSS Statistics，以区别于 SPSS 公司的其他产品，如 SPSS Data Collection 数据收集产品、SPSS Modeler 数据挖掘产品等。

同时，由于适应商业分析应用的要求，SPSS Statistics 产品已经发展为包括 Server 端和 Client 端等在内的整体 cs 架构的软件系统。我们通常熟悉的 SPSS 软件，应该说现在准确的全称为 SPSS Statistics Client。

一、SPSS 软件的历史

SPSS Statistics 统计分析软件是世界上最早的统计分析软件，由美国斯坦福大学的三位研究生于 1968 年研制，1984 年 SPSS 首先推出了世界上第一个统计分析软件微机版本 SPSS/Pc+，极大地扩充了它的应用范围，并使其能很快地应用于自然科学、技术科学、社会科学的各个领域。

迄今 SPSS Statistics 软件已有 40 余年的成长历史。全球约有 28 万家产品用户，它们分布于通信、医疗、银行、证券、保险、制造、商业、市场研究、科研教育等多个领域和行业，是世界上应用最广泛的专业统计软件。

在 2009 年 IBM 收购 SPSS 公司后，现在在国内市场上推出的最新产品：IBM SPSS Statistics19 多国语言版。

二、SPSS 软件的优点

Windows 风格，界面友好

SPSS Statistics for Windows 最突出的特点就是操作界面极为友好，输出结果美观漂亮。

易学易用

SPSS 易于操作、易于入门，结果易于阅读，对统计软件的学习不会冲淡统计的主题，这样研究人员就可以将精力集中在社会研究方法、市场研究方法、营销的业务问题上，而不是忙于编程和统计。

功能全面

SPSS 针对初学者、熟练者及精通者都比较适用。提供了数据获取、数据管理与准备、数据分析、结果报告这样一个数据分析的完整过程，因此非常全面地涵盖了数据分析的整个流程，特别适合设计调查方案、对数据进行统计分析，以及制作研究报告中的相关图表。

编程能力强大，支持二次开发

熟练或精通者可以通过编程，在 SPSS 命令语法语言的基础上提供与其他编程语言的结合功能，来实现更强大的功能。

支持丰富的数据源，具备强大数据访问和管理能力

SPSS 可以同时打开多个数据集，方便研究时对不同数据库进行比较分析和进行数据库转换处理。

三、SPSS 软件的功能

SPSS 统计分析软件是一款按照功能模块进行配置的软件产品，主要包括 SPSS Statistics Base 模块和其他一系列扩充功能模块，每个独立扩充功能模块均可在 SPSS Statistics Base 模块基础上，为其增加某方面的分析功能。SPSS Statistics 19 全模块软件包括如下主要 16 个功能模块：

SPSS Statistics Base

必需的基础模块，管理整个软件平台，管理数据访问、数据处理和输出，并能进行很多种常见基本统计分析。

SPSS Advanced Models

除基本的数据分析外，用于分析过程数据、嵌套数据、事件历史和持续时间数据得到更为精确预测模型所用的系列方法。

SPSS Bootstrapping

方便更有效使用小样本量的数据的方法。通过数据自身重采用的功能，让用户可以模拟大样本情况下的采样结果，从而对数据结构特征和偏差有更直接的认识。

SPSS Categories

优秀的对应分析程序，用启发性的二维图和感知图清晰展示数据关系，便于更完整地分

析数据。

SPSS ComDlex SamDles

如果使用了特别复杂的抽样方案，该模块可以计算复杂样本的统计数据，得到更精确的结果。

SPSS Conjoint

包含三个相互关联过程的一个系统，用于进行全特征联合分析。

SPSS Custom Tables

提供 35 种单元和摘要统计量，能够更方便地显示多重序列数据，它能串接所有的维度，以在同一表格中显示包含不同统计量的各种变量。

SPSS Data Preparation

简化数据准备过程。

SPSS Decision Trees

基于数据挖掘中发展起来的树结构模型对分类变量或连续变量进行预测，可以方便、快速地对样本进行细分。

SPSS Direct Marketing

主要用于处理市场直销中的一些分析需求，包括 RFM 客户评分、客户分群、目标客户轮廓概括、客户响应评分、不同营销行为响应测量等模型。

SPSS Exacl Tests

在数据结构不满足常规分析要求时，得到正确的 P 值，提供可信赖的结果。

SPSS Forecasting

功能强大的时间序列分析工具，是分析历史资料、建立模型与预测未来事件的强有力的工具，能帮助研究人员做更好的预测。

SPSS Mlssing Values

帮助研究人员在分析过程中排除数据中隐含的偏差，得出更精确的结论。

SPSS Neural Networks

神经网络是一种模拟人类大脑处理信息的方式的简化算法模型，帮助用户以全新的方式探索数据。

SPSS Programmability Extension

帮助用户使用外部语言来执行 SPSS 一连串的分析动作以达到自动化的目的。

SPSS Regression

大量的非线性建模工具、多维尺度分析功能等，帮助研究人员进行非线性回归分析。

第二节 SPSS 统计分析的一般流程

一、数据的输入和保存

当打开 SPSS 后，展现在我们面前的界面如图 2-1 所示。



图 2-1 数据管理

这是典型的 Windows 风格界面，有菜单栏、工具栏。特别的，工具栏下方的是数据栏，数据栏下方则是数据管理窗口的主界面。该界面和 Excel 极为相似，由若干行和列组成，每行对应了一条记录，每列则对应了一个变量。

SPSS 的变量管理窗口如图 2-2 所示。

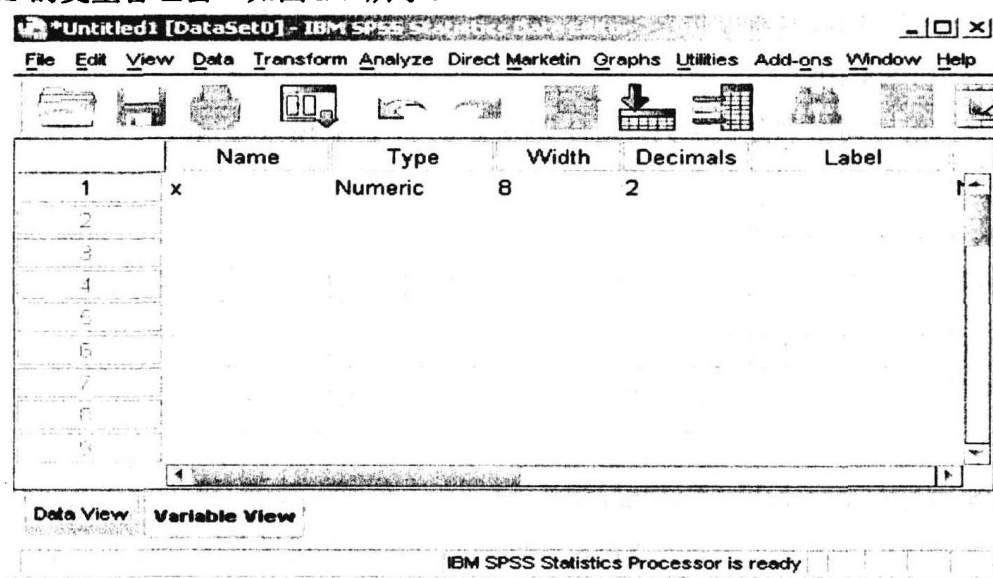


图 2-2 变量管理

要建立变量 X ，直接按要求输入（或选择）变量名、变量类型等内容即可，如图 2-2 所示。输入变量 X 的值，只需选中数值应在的位置，然后直接从键盘上输入即可，如图 2-3 所示。

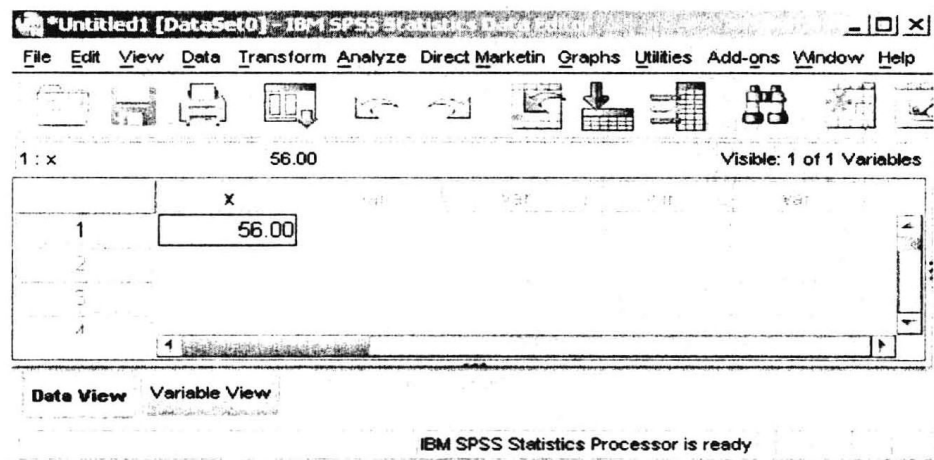


图 2-3 输入数据

选择菜单 File→Save，由于该数据从来没有被保存过，所以弹出 Save as 对话框如图 2-4 所示。

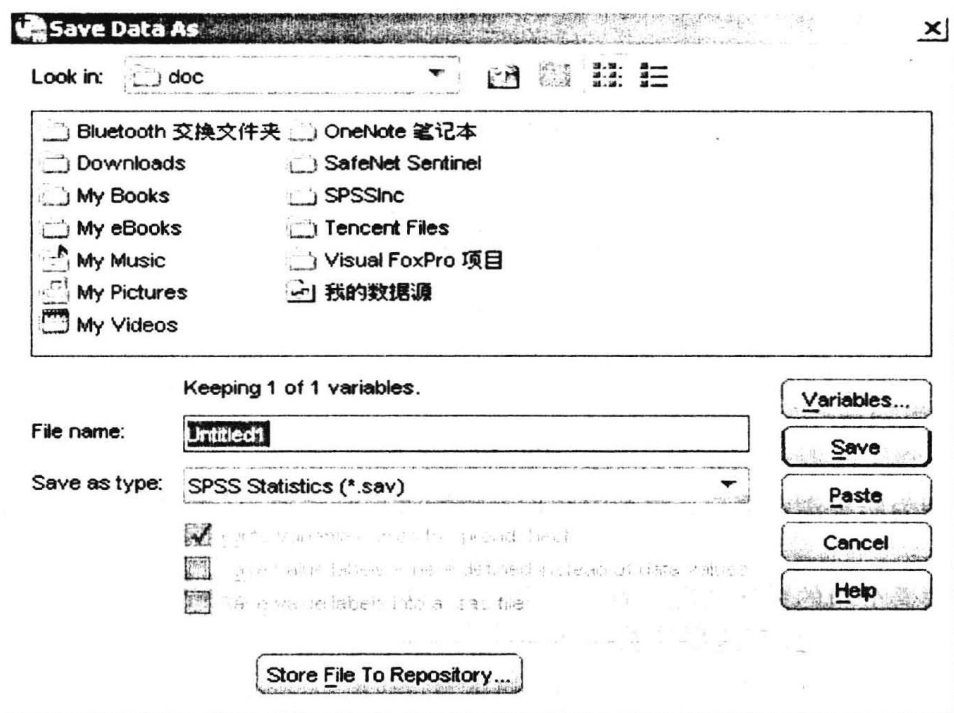


图 2-4 保存数据

单击保存类型列表框，可以看到 SPSS 所支持的各种数据类型，通常将其存为 SPSS 自己的数据格式 (*.sav 文件)。在文件名框内键入文件名并制定存放位置即可。

二、数据预分析

要知道数据的基本情况，如均数、标准差等，需选择 Analyze→Descriptive Statistics→Descriptives 菜单，系统弹出描述对话框如图 2-5 所示。

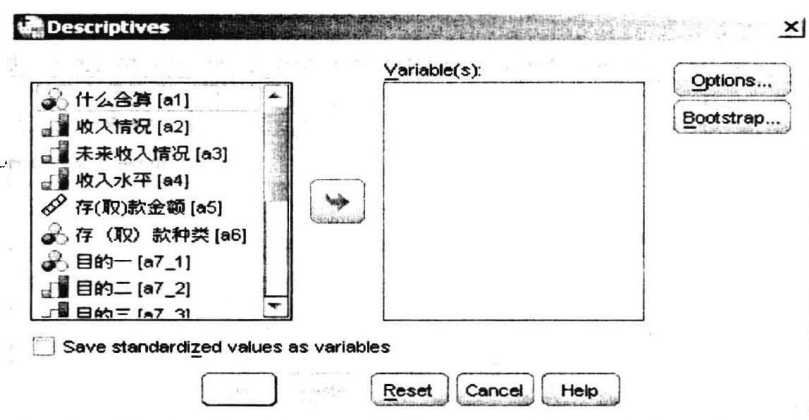


图 2-5 描述统计

把左侧的候选变量列表中需要统计的变量选入右侧即可进行分析，如图 2-6 所示。

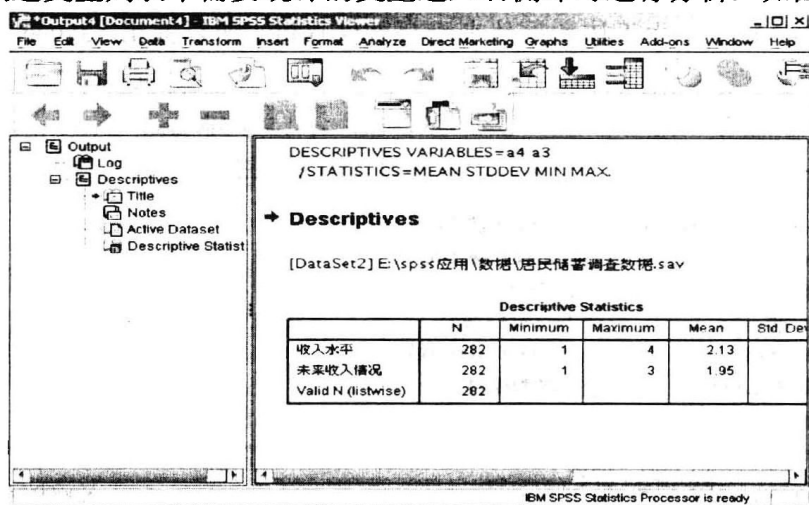


图 2-6 输出结果

该窗口即为结果浏览窗口，整个的结构和资源管理器类似，左侧为导航栏，右侧为具体的输出结果。

如果需要看分组的描述情况，就需要用文件分割功能，如图 2-7 所示。

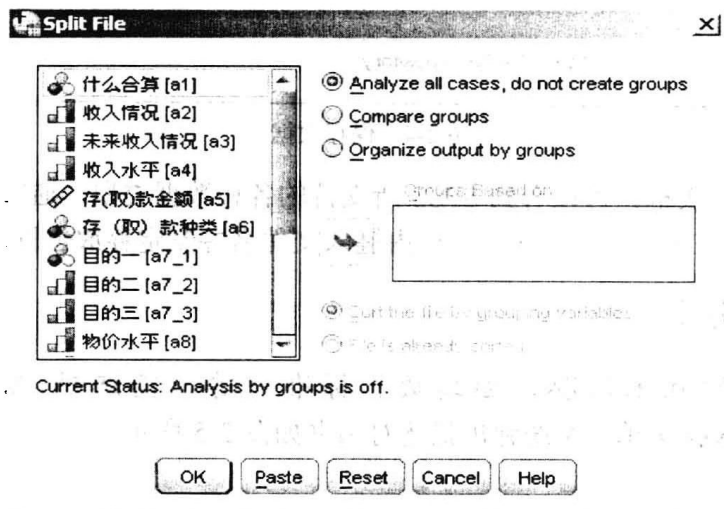


图 2-7 数据分组

选择 Organize output by groups，将变量分组变量选定再做描述分析即可。

如要绘制直方图，则需选择 Graphs→Legacy Dialogs→Histogram，并分别指定标题变量、行变量和列变量即可。系统会弹出绘制直方图对话框，如图 2-8 所示。

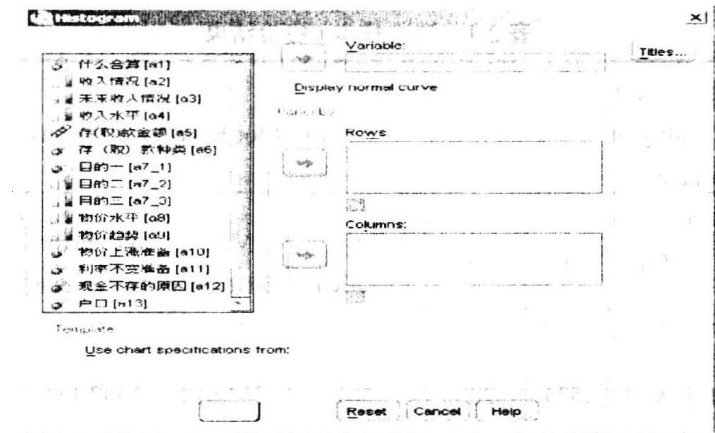


图 2-8 绘制直方图

三、按要求进行统计分析

如用 SPSS 来比较城镇居民和农村居民的存（取）款金额是否相同，方法如下：选择 Analyze→Compare Means→Independent-Samples T test，系统弹出独立样本 t 检验对话框，如图 2-9 所示。

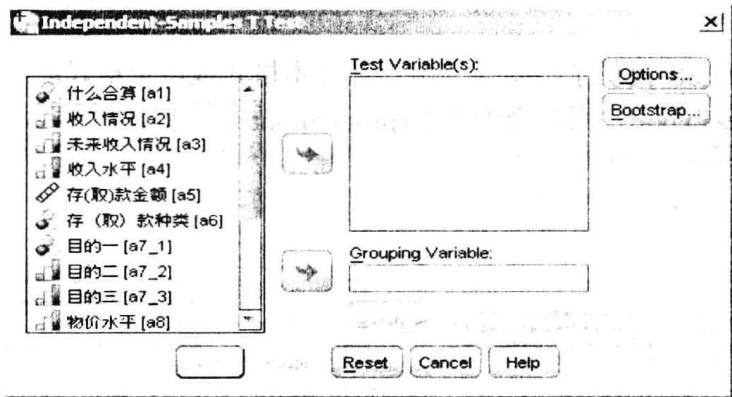


图 2-9 独立样本 t 检验

将变量存（取）款金额选入 test 变量框内，变量户口选入 grouping 变量框内，并确定相比较组变量取值即可，如图 2-10 所示。t 检验的结果如表 2-1 所示：

结果分为两部分：第一部分为 Levene's 方差齐性检验，用于判断两总体方差是否齐。检验结果为 $F = 0.001$ ， $p = 0.980$ ，方差是齐的；第二部分分别给出在总体方差齐和方差不齐时的 t 检验结果。由于前面的方差齐性检验结果为方差齐，第二部分就应选用方差齐时的 t 检验结果，即上面一行列出的

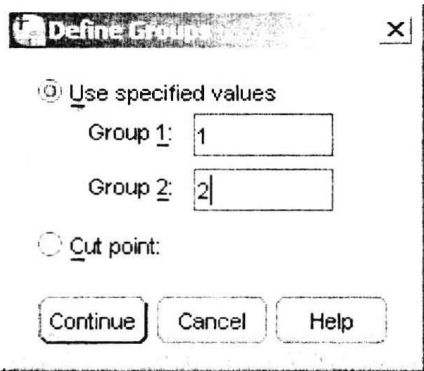


图 2-10 定义分组标志