

社会经济定量研究

与 SPSS 和 SAS 的应用

骆克任 编著



电子工业出版社
Publishing House of Electronics Industry

社会经济定量研究与 SPSS 和 SAS 的应用

骆克任 编著

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书较全面地介绍了在统计研究中,如何根据研究目的提出变量和问题、形成研究假设、设计研究方案、选择计量方法、搜集信息并用现代统计技术去分析,以及如何引出结论和撰写研究报告等一系列的定量研究方法。同时,结合定量研究的各个步骤,相应地介绍了 SPSS 和 SAS 统计软件的使用技术。该书不仅对定量分析研究起到科学和规范的引导作用,而且能对统计软件的使用有所指示。

本书广泛适用于市场、企业、政府管理与学术研究中的定量分析研究,并可供高等院校科研单位的研究人员、研究生、大学生以及企事业单位的管理人员和电脑技术人员作为教材或参考书使用。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,翻版必究。

图书在版编目(CIP)数据

社会经济定量研究与 SPSS 和 SAS 的应用/骆克任编著. —北京:电子工业出版社,2002.1

ISBN 7-5053-7129-0

I. 社… II. 骆… III. 经济统计—统计分析—软件包, SPSS、SAS IV. F222.1

中国版本图书馆 CIP 数据核字(2001)第 081838 号

书 名: 社会经济定量研究与 SPSS 和 SAS 的应用

编 著 者: 骆克任

责任编辑: 赵建宏

排版制作: 电子工业出版社计算机排版室

印 刷 者: 北京天宇星印刷厂

出版发行: 电子工业出版社 <http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036

经 销: 各地新华书店

开 本: 787×980 1/16 印张: 18.75 字数: 300 千字

版 次: 2002 年 1 月第 1 版 2002 年 1 月第 1 次印刷

书 号: ISBN 7-5053-7129-0
F·206

定 价: 24.00 元

凡购买电子工业出版社的图书,如有缺页、倒页、脱页、所附磁盘或光盘有问题者,请向购买书店调换;
若书店售缺,请与本社发行部联系调换。电话 68279077

前 言

当今世界仍以和平与发展为两大主题，各国经济进一步的国际化，促进了全球经济一体化的进程。同时，人类社会已进入信息及网络时代，知识经济正逐步成为现代经济的主流。我们面前的这个丰富多彩的世界，随时呈现着各种数量庞大的、无穷无尽的社会和经济信息。为了把握社会经济发展的现状、规律和未来，我们在对这些信息资料进行定性分析之外，还必须进行科学的定量分析研究。例如，在我们进行社会科学研究，编制社会与经济计划，以及进行市场调研、企业管理和营销抉择时，都需要以历史唯物论和辩证唯物论的观点，经过定量研究大量有关的社会经济信息数据，从事物的运动中分析事物之间的关系、矛盾和发展变化，从而认识事物发展的现状和趋势。为了科学地定量分析研究各种社会和经济信息，我们必须掌握结合社会学、经济学和数学等现代知识的先进定量研究方法。这些定量研究方法将包括根据研究的目的，提出拟研究的变量和方法、形成研究假设的方法、设计研究方案的方法、选择计量的方法、信息搜集的方法、现代统计分析的方法，以及引出结论和撰写研究报告的方法。同时，也需要会使用相关的现代统计和计算机辅助技术。掌握这些研究分析社会经济信息的跨学科的定量研究方法及其技术，将可以帮助我们多快好省地收集所需要的社会或经济信息，并对其进行去粗取精、去伪存真、由此及彼、由表及里的工作，使我们能准确地搜集和显示社会经济的客观情况，并对发生的现象总体进行概括性的推论，找出带有普遍性的规律。

科教兴国，人人有责。考虑到目前仍有广大读者和学生为了提高自己的工作水平，迫切希望掌握现代的社会经济的定量研究方法与技术，以及计算机应用软件版本不断有较大更新等原因，作者根据近年对《现代实用统计及计算机应用》一书教学和课题研究的体会，对原书内容进行了修改、补充，采用了最新版本的计算机统计软件，增加了 Windows 的屏幕界面。希望此本更新后的书，能对社会经济研究方面的有关教学改革有所帮助，并对我国社

会经济定量分析研究的科学化和规范化起到积极的作用，为我国社会经济的持续发展做出一些贡献。

本书在编写中特别注意以下两个方面：一、深入浅出地介绍，增强实用性和可操作性，避免烦琐的数理推导过程，使读者即使没有坚实的数理统计知识和扎实的计算机功底也可学会。在遇到需要定量研究的问题时，能够一步步地解决它。至于那些有兴趣学习更深奥的理论知识的人士，可以进一步另选专门的书籍。二、书中着重介绍目前国内社会经济统计书籍里缺少的定量研究过程（例如，介绍规范的科学研究原则和步骤），以及定量分析方法（例如，注重讨论不可以直接计量，只能用间接技术来反映的方法，这些方法常用于社会和市场经济中的很多方面，如对大众的看法或态度等进行计量）。

根据理论联系实际的原则，作者强调理解和应用书中的这些方法和技术，而不仅是阅读和死记硬背。书中每个部分都包括着用于解决重要的社会和经济研究课题的方法论方面的推理和逐步分析，以及现代统计知识和相应的计算机软件的应用技术。各部分的例题和图表，皆鼓励读者积极地思考和了解有关统计和定量研究方法问题的不同阐述过程。各部分的练习，是为了使读者能够实际应用讨论过和举例说明过的方法，以便进一步掌握和巩固所学知识。例题和练习使读者在有无计算机的情况下，都能进行统计和方法论的学习。当然，不熟悉计算机统计软件的部分读者，做一些有关计算机的练习，将是非常有益的。对于已经掌握了一定的计算机统计程序的读者，也会通过对现代统计方法及其两种著名的计算机统计分析软件包的学习，而更加大有可为。

本书的每个部分都包含着对定量研究方法的某个重要内容的讨论、举例和练习。作者努力将包含定量研究方法方面的基本内容，与使用国际上流行的著名计算机统计分析软件相结合，以期各类人员都能够同时简捷地掌握统计分析和计算机的相关技术。书中介绍的社会科学统计软件包 SPSS (Statistical Program for Social Sciences) for Windows 10.0 版，以及统计分析系统软件包 SAS (Statistical Analysis System) for Windows 6.12 版，将支持各种统计和定量研究方法的使用。这两个软件都是用于统计分析研究的著名软件，用这两个软件统计分析数据，其分析理论和算法都得到国际认可。

通过此书,作者希望达到下列目的:

- (1) 组织简要的、真实的有关社会经济的定量研究方法和技术的讨论。
- (2) 通过清楚的、有意义的例子,使读者理解计算机统计程序及运用。
- (3) 举例说明如何解释研究报告中的统计内容和其他定量分析结果。
- (4) 给出与章节内容直接有关的练习,使读者有机会去实际进行定量研究。
- (5) 使计算机用户和非计算机用户都能从本书的内容和练习中获益。

(6) 书中使用了某贫困地区的抽样调查数据集(以下简称 SQ 数据集)。该数据集与书中内容、例题和练习紧密配合,容易使用,并有实际意义。在社会调查课程及整个教学过程中, SQ 数据集都充当重要角色。SQ 数据集允许使用本书的教师进行讨论,并用它艺术地陈述所进行的研究,这个 SQ 数据集在全书有关基础统计及定量研究过程的讨论、例题和练习中,将作为实际有效的用具。通过附录 A 的说明,您可以清楚了解该数据集的内容。重要的是 SQ 数据集为个人计算机用户的 SPSS 和 SAS 统计分析程序,提供了必不可少的实用数据,使计算机的用法与统计学和方法论的教学融为一体(读者可在电子工业出版社的网站上下载此 SQ 数据集。网址为: <http://www.phei.com.cn/download.htm>)。

由于下列理由,本书新的教学法与传统的方法相比,将使读者获益更多。相信使用这本书的读者,能更深入地了解定量研究方法和相关技术。

(1) 与传统书籍中常见的冗长讨论相比,读者能更好地理解定量研究方法和统计讨论中的那些短小、简明的观点。

(2) 多数读者将乐意进行结合实际研究内容和研究方法的练习。

(3) 书中提供的有意义的变量及容易使用的数据集,将能引起读者更大的兴趣。

(4) 没有计算机知识的读者,只要按提示去做,即可容易地在个人计算机上运行 SPSS 或 SAS 软件包。

(5) 与仅仅是阅读和死记硬背统计和计量方法的某些方式相比,完成练习中的笔头和上机作业,可以使读者学会并牢记更多的知识。

书中内容安排如下:首先,在引导部分介绍计算机的准备知识,为学习下面各章的计算机辅助应用打下基础。对于基础数学知识,作者认为每个从事社会经济定量研究的人员都应该具备,所以不在此书开始时讲解。随后书中各章的编排是第一章至第三章介绍定量研究的基本步骤和目的以及统计描述;第四

章至第六章介绍假设的提出、概率的基本概念，以及讨论交叉分组列表、双变量间的分析；第七章至第九章介绍定量研究的程序，以及有关社会及经济问题所涉及的指标和指数等定量研究方法，每种定量方法都有相应的实例应用；第十章至第十七章讲述了必要的统计步骤和许多统计法则，定量研究人员应通过这些步骤和法则对数据资料进行收集、分析、理解及其深入的讨论，然后得出完善的解释。本书在内容上和结构上都可以满足一门既包括统计又包括定量研究方法的课程的需要。然而，它又是足够灵活的，每一个特定章节，都既可以用于定量研究方法，又可以用于统计的课程。因为每一个章节中的内容，都自成体系，包括内容、例题和练习（这些材料既可以用于非上机的学习，又可以用于上机的练习）。作者鼓励读者用不同的方法使用这本书，把它当做有广泛适用性的书籍，而不是固定不变的程序。

本书的全部目的在于，提高读者对必要的社会经济研究过程的理解，同时了解定量研究方法和相关统计技术是如何共同用于一个研究过程的；并且使有兴趣和远见的读者能够掌握计算机来协助研究的进行，使其在工作中如虎添翼。作者希望通过此书，使读者在实际运用所学知识去进行分析研究及解释研究结果时，能提高兴趣、增强信心，并在研究探讨社会与经济的问题中，使工作和研究水平不断提高。

作 者

2001 年 10 月

目 录

引导 计算机准备知识	1
一、计算机的基本结构	1
二、统计软件包	2
三、SPSS 和 SAS 统计软件包的功能介绍	2
四、SPSS 和 SAS 软件包的使用	9
第一章 变量及研究过程	17
一、变量的类型	17
二、变量在研究中的应用	19
三、变量及其研究步骤	20
第二章 单变量的频数分布	24
一、单变量频数分布表的构成	24
二、频数分布表及计算机软件的应用	26
第三章 单变量频数分布的统计计量和描述	34
一、比例、百分比、比和比率	34
二、集中趋势的计量	36
三、离散趋势的计量	42
四、单变量频数分布及计算机软件的应用	45
五、频数分布的图形描述及计算机软件的应用	52
第四章 假设的形成与检验	58
一、“假设”的概念	58
二、双变量假设	59
三、理论假设与工作假设的形成	60
四、概率基础知识	61
五、假设的检验及计算机软件的应用	64

第五章 双变量的频数分布：表及其统计过程	72
一、名义双变量的分布	72
二、用名义双变量表检验假设及计算机软件的应用	74
三、有序双变量的分布	85
四、用有序双变量表检验假设及计算机软件的应用	85
第六章 双变量表的详述	105
一、因果关系	105
二、因果关系的判别标准	107
三、双变量关系的确认和详述	108
四、计算机软件的应用与分析	113
五、双变量关系的研究步骤	120
第七章 计量	129
一、概念化单方面的理论陈述	129
二、可操作化单方面的概念	130
三、构造指标	130
四、指标可靠性和有效性的判断	130
第八章 复杂社会概念的指数解释及构成	135
一、指数的构成	135
二、指数体系中的指数项的保留与去除	138
三、构造指数变量及计算机软件的应用	140
四、有关标度的简短讨论	142
第九章 统计的一般指数方法	145
一、统计指数的概念	145
二、统计指数的分类	145
三、指数的编制	146
四、指数体系和因素分析	151
第十章 抽样过程	159
一、非概率抽样	159
二、概率抽样	160

第十一章 调查研究	168
一、正确选定调查单位	168
二、避免错误的推断	169
三、恰当进行分类	170
四、科学设计调查类型	170
五、掌握数据收集技术	172
六、高质量问题与问卷的设计准则	174
七、提高全面回答的比率	177
八、调查研究的道德准则	179
第十二章 推断统计的基础	184
一、正态曲线	184
二、正态曲线的应用	185
三、中心极限定理	190
四、均值置信区间的估计	191
五、成数的置信区间的估计	193
六、样本量的确定	194
七、计算机软件的应用	195
第十三章 推断统计：显著性检验	200
一、显著性检验的基础	200
二、单样本比例的 Z 检验	203
三、t 分布	204
四、单样本均值的 t 检验	205
五、双样本比例的 Z 检验	206
六、双样本均值的 t 检验及计算机软件的应用	207
七、前 / 后或成对样本的 t 检验及计算机软件的应用	210
八、方差的 F 检验	212
九、两个以上样本均值的单因素方差分析检验及计算机软件的应用	213
第十四章 精选的非参数统计	221
一、斯皮尔曼的等级顺序相关	221
二、卡方拟合优度检验及计算机软件的应用	223
三、曼-惠特尼 U 检验及计算机软件的应用	226

四、威尔科克森配对正负号等级检验	228
五、克罗斯考尔-瓦里斯 H 检验及计算机软件的应用	229
第十五章 双区间/比率变量的关联	234
一、线性回归	234
二、皮尔逊相关系数	236
三、 r 的显著性检验	238
四、预测及其误差	239
五、计算机软件的应用	240
六、回归与相关分析中应注意的问题	244
第十六章 多元线性回归和相关分析	247
一、多元线性回归和相关分析介绍	247
二、贝塔系数或贝塔权数	248
三、回归统计的显著性检验	249
四、计算机软件的应用与分析	249
第十七章 非线性回归分析	261
一、非线性回归模型的拟合	261
二、曲线类型的选择	265
三、计算机软件的应用	267
附录 A SQ 数据集的介绍	271
一、SQ 数据集中的变量说明	271
二、SQ 数据文件及其调用	272
附录 B SPSS 的分类命令	273
一、数据定义命令	273
二、数据传输命令	273
三、单位选择命令	274
四、其他常用命令	274
五、基本系统的命令	274
六、高级统计系统命令	275
七、表格系统命令	275
八、绘图系统	275

九、地图系统（略）	275
十、趋势分析（略）	275
附录 C 统计表格	276
参 考 文 献	285

引 导

计算机准备知识

社会的发展，导致了信息量的猛增，为了更有效地研究社会现象，当今的研究人员越来越感到，以往笔头和算盘一类的统计方法已远远不能适应统计工作的需要。为了解决这个问题，越来越多的研究人员想要或正在求助计算机这个现代的工具。因此，本书在介绍现代统计分析方法的同时，引导大家掌握一些有关的计算机辅助工作方法。在讲解正文之前，有必要先了解一些有关计算机的准备知识。

一、计算机的基本结构

对于以前未接触过计算机的读者，这里先简单介绍一些计算机的基本常识。计算机是由硬件和软件两部分组成的系统，缺一不可。硬件是指计算机的物理结构；软件指为了运行、管理和维修计算机所编制的各种程序和程序的有关说明资料。

硬件结构可以分为中央处理单元（CPU）、存储器、输入设备和输出设备四个部分。中央处理单元是由运算器和控制器组成，它是计算机的核心。存储器是计算机的存储部分，用来存储各种数据和程序。存储器又可分为内存和外存。内存是电路板上的一种电子芯片，外存则是磁盘、光盘、磁带等。输入设备是用于输入数据和程序的，常用的有键盘、鼠标器、磁盘驱动器、扫描仪、光电声读入器等。输出设备用于输出数据结果等，常用的有显示器、打印机、绘图仪和磁盘驱动器等。以上四部分加上电源和一些附加电路，并通过总线连接，就构成计算机硬件系统。人们通常将装在机箱中的中央处理单元和存储器称为主机，将输入和输出设备称为外部设备。

计算机不会自己去解决问题，而是人们通过使用计算机软件来让计算机硬件去解决问题的。软件主要有系统软件、计算机语言和应用软件。系统软件用于管理和控制计算机资源，组织计算机的工作流程。如用户要通过 DOS 或 Windows 操作系统来使用计算机。计算机语言是表达人的意志的程序语言，计算机语言经历了机器语言、汇编语言、高级语言（如 BASIC、FORTRAN、Pascal、COBOL、C 等等），直到现在开发的第四代程序设计语言。应用软件是已经编好的用于各种实际目的的程序。那种已形成标准化、模块化的解决不同类问题的组合应用程序被称为软件包，如目前流行的字处理软件包 WPS、Word，数据

库管理软件包 Foxpro 和 Oracle, 以及本书中介绍的统计软件包 SPSS 和 SAS 等等。

二、统计软件包

研究人员有时候会因为所研究问题的特殊复杂性, 自己编制一些特别的应用程序来解决问题。但在多数情况下, 要完成的统计是一些标准的和通用性的工作, 这时借助于一种权威可靠的统计软件包将是更明智的。统计软件包是针对各种不同统计任务而编制的专项应用程序的集合, 它是为适应统计手段现代化的要求而产生。这些统计软件包的功能全面, 且比较独立。统计软件包一般由数据和文件管理系统、统计命令及运行操作控制系统三部分组成, 三者联系密切, 但作用不同。要统计的数据和文件以及其结果, 要经过数据和文件管理系统来转换, 再输入、计算和输出。统计功能要靠命令程序来组织, 而数据和文件的管理以及统计命令的执行, 都要由运行操作控制系统来完成。

当今已有许多统计软件包。国外比较权威和著名的有社会科学统计软件包 SPSS (Statistical Program for Social Sciences)、统计分析系统软件包 SAS (Statistical Analysis System)、生物医学统计软件包 BMCP (Biomedical Computer Program)、数据处理组合软件包 IMPS (美国普查局开发的 CENTRY、CONCOR 和 CENTS 三个软件的组合) 等, 国内有关系式数据分析系统软件包 RDAS (Relational Data Analysis System)。下面将选择介绍 SPSS 和 SAS 软件包 (个人计算机版), 并在书中结合统计分析来讲解和使用。

三、SPSS 和 SAS 统计软件包的功能介绍

在利用计算机进行统计分析应用之前, 需要根据拟统计的数据量 (包括变量的数目、类型、单位记录的条数等) 以及统计的项目来选择计算机的硬件和软件。对于数据量大、统计计算工作量大 (如需要进行大矩阵运算和复杂预测) 的任务, 需要选择中央处理单元 CPU 的档次高、内存量大, 且具有大容量硬盘的计算机, 这样计算机的计算速度就快一些, 并可以处理大量的数据; 另外, 对于数据量大的分析资料, 也要选择能够处理相应数据量的软件包, 并且该软件包还要具有所需要的统计功能。

下面介绍 SPSS 和 SAS 软件包的功能, 使读者先全面了解两个软件包各自的功能, 以便今后进一步有针对性地去学习。

(一) SPSS 统计软件包的功能

自从 1966 年推出 SPSS 以来, 经过不断改进, SPSS 从开始针对 DOS 操作系统的软件, 到目前针对的 Windows 窗口软件, 版本不断更新到当前的第 10 版, 统计功能不断

完善。它几乎包括了所有的数据管理和统计分析功能, 并且命令简短, 具有方便的帮助 (HELP) 功能, 操作也更加方便。尤其是 Windows 版本, 甚至不必记忆统计命令即可通过人机对话的窗口菜单, 调用主要的统计功能, 用户只要通过菜单来选择所用的功能, 即可完成计划进行的工作, 使它更适应非统计专业的统计分析工作者使用。各种版本的 SPSS 的统计原理、语法规则、程序编写和结果分析等实质内容是相同的, 不同之处仅在于软件工作环境、操作方便程度和统计功能的多少。

下面按 SPSS 的基本和高级两部分功能, 分模块来简介其主要的功能过程。

1. 基本系统 (Base System)

基本系统模块提供以下过程:

(1) 数据文件管理 (Data Files 过程)。包括对有关文件进行创建、打开、查找、编辑和存储等工作。

(2) 数据的定义和操作 (Data Transformation 过程)。包括数据的重新编码和加权、创建新变量、对数据进行算术和逻辑的计算, 漏填数据的处理等; 并可以对样本数据进行分类排序, 读写数据文件及选择打印输出格式等。

(3) 数据汇总 (Frequencies 过程)。对给定变量进行频数分析, 并可以输出表征频数的条形图、直方图和分布图, 另外可以给出均值、众数、中位数、陡度和偏斜度等 14 种统计量。

(4) 描述性统计 (Descriptives 过程)。用于有序变量和区间/比率变量的数据汇总, 给出 12 种统计量, 特别是能计算概率度 (Z 值)。

(5) 交叉表和检测 (Crosstable 过程)。用于进行交叉汇总。完成两个或两个以上变量在分项不多的情况下的联表, 并可结合检验测量。

(6) 多维变量的描述统计 (Means 过程)。进行三维以上变量的描述统计。

(7) t 检验 (T-Test 过程)。用于检验两组或两组以上样本的均值, 从而进行组间显著性测试。

(8) 相关分析 (Correlation 过程)。计算多元相关系数矩阵。

(9) 多因子方差分析 (Anova of Variance 过程)。进行因素间的变异量分析。

(10) 多元线性回归 (Multiple Linear Regression Analysis 过程)。用于多变量的直线回归。

(11) 绘图专用系统 (Graphics)。可以用 SPSS 本身的 GRAPH 命令直接对 SPSS 数据文件的数据绘图, 也可以先综合 SPSS 的数据, 然后传送给 Microsoft Chart 专用制图软件来绘制高质量的图。它可以产生和输出 5 种基本图形和数 10 种派生图。

(12) 帮助功能 (包括联机帮助和教学程序 Tutorial)。在运行 SPSS 时, 只要键击 F1, 选择所要了解的题目, 便可以得到有关函数、语法及操作的信息。为了能学会 SPSS 的主要内容和数据分析的基本概念, 了解系统的运行环境, 以及掌握文件的编辑和操作, 命

令的输入、修改与执行,软件包中的电子教学程序 Tutorial 以交互的方式来由您选择各级菜单,以决定学习的内容。建议初学者使用该教学程序来自行学习。

2. 高级统计系统 (Advanced Statistics)

(1) 判别分析 (Discriminant 过程)。该过程是判别样本所属类型的一种多元统计分析方法。

(2) 快速聚类分析 (Quick Cluster 过程)。该过程是将一批样本或指标按照它们在性质上的紧密关系,在没有先验知识的情况下进行分类的方法。

(3) 因子分析 (Factor 过程)。该过程进行多种类型的主成分分析和因子分析。主成分分析是将多个指标化为少数几个指标,以舍弃观测值变化不大的变量的一种统计方法。因子分析是通过最少个数的不可观察的随机因素来说明出现在可观察随机向量的分布中的相关模型,以便尽可能合理地解释包含在原始变量之间的相关性,并用以简化变量的维数和结构。

(4) 多元方差分析 (Manova 过程)。该过程进行方差和协方差的最一般性分析,它执行单因变量和多因变量的方差分析。

(5) 非线性回归 (Nonlinear Regression 过程)。进行变量的曲线拟合。

(6) 残存统计 (Survival 过程)。计算残存表和考克斯回归等工作。

(7) 非参数检验 (Nonparametric Tests 过程)。用于在不知道变量概率分配情况下而要进行假设检验的工作。

(8) 趋势分析系统 (Trends)。主要用来对时间序列进行分析。它可以对数据进行绘图、平滑、分解、回归,以及频率分析和使用 ARIMA 方法建立模型。

(9) 制表专用系统 (Table)。该系统比基本系统能产生更为复杂的表格。它可以完成一维、二维和三维表;实现多个表的连接;确定不同形式的百分比输出;进行一些基本统计;进一步完善表格的内容。

(二) SAS 软件包的功能

SAS 是一种组合软件系统,其基本部分是 Base SAS 软件。此外,还有用于统计分析计算的 SAS/STAT 软件,用于经济预测或时间序列分析的 SAS/ETS 软件,用于绘图的 SAS/GRAPH 软件,用于矩阵运算的 SAS/IML 软件等等。SAS 系统的核心提供了 125 个灵活运用过程,这些过程可以直接使用,或者稍加修改,以便满足特殊的需要。

1. Base SAS 软件

Base SAS 软件包括 31 个 SAS 的基本过程,按照用途可以划分四大类。

(1) 基础统计计算类

- MEANS 均值过程——用于对数值变量产生单变量的简单统计描述。

- **SUMMARY** 概括过程——主要用于产生一个包含基本描述统计量的数据集。
- **UNIVARIATE** 单变量过程——该过程能完成类似 **MEANS** 过程的更多的基本统计计算。
- **CORR** 相关过程——用于计算变量间的相关系数。
- **FREQ** 频数过程——用于产生单向到 n 向的频率表和交叉表。
- **TABULATE** 制表过程——用于产生分类报表显示描述性统计量。
- **CHART** 图表过程——该过程用图表的方式来形象地表示变量的取值和两个以上变量之间的关系。

(2) 报表生成类

- **PRINT** 打印过程——用于输出 **SAS** 数据集中的数值。
- **PLOT** 图形过程——用于画散点图，以便观察数据的变化趋势和数据间的相关关系。
- **FORMS** 标签过程——用于产生一些信封标签、邮件标签、外部磁带标签、文件记录卡片及固定式样的其他打印形式。
- **CALENDAR** 日历过程——用于生成两种日历报表、总结表或计划表。
- **TIMEPLOT** 间隔图形过程——用于对一个或几个变量，绘制时间间隔的散点图。

(3) 得分类

- **STANDARD** 标准化过程——该过程将 **SAS** 数据集中的一些变量或所有变量按给定的均值和标准差进行标准化变换，并生成一个包含标准化值的新的 **SAS** 数据集。
- **RANK** 秩得分过程——该过程将数值先从最小到最大排列，再对最小值赋予秩 1，对第二个小的值赋予秩 2，直到赋予秩 n （即未丢失的观测值的数目），对于结值（数值相同的观测值）将赋予平均秩或最大秩或最小秩。

(4) 实用类

- **SORT** 排序过程——按一个或几个变量的大小顺序将观测值重新排序，并将结果存入新的数据集。
- **FORMAT** 格式过程——对字符和数值变量，定义用户自己的输入格式和输出格式。
- **TRANSPOSE** 转置过程——完成对数据矩阵的转置工作。
- **OPTIONS** 选择项过程——列出所有 **SAS** 系统选择项的当前值。
- **PRINTTO** 输出地点过程——用于定义 **SAS** 过程输出的地点，也可以用来改变 **SAS** 记录的目的地。
- **APPEND** 附加过程——将一个 **SAS** 数据集的观测附加到另一个 **SAS** 数据集的后面。
- **CONTENTS** 内容过程——用于打印输出 **SAS** 数据库中一个或几个文件的描述信