



普通高等教育“十一五”国家级规划教材

SAS与现代经济统计分析

SAS YU XIANDAI JINGJI TONGJI FENXI

► 岳朝龙 黄永兴 / 编著

中国科学技术大学出版社



普通高等教育“十一五”国家级规划教材

SAS与现代经济统计分析

岳朝龙 黄永兴 / 编著

中国科学技术大学出版社

内 容 简 介

本书涵盖了大部分常用的现代统计分析方法,是使用 SAS 进行经济统计分析的一本非常实用的书籍。全书分三篇 15 章:第一篇 3 章,系统介绍 SAS 系统的特点、SAS 数据库和 SAS 数据集的创建,其他格式数据文件与 SAS 数据集的相互转换、SAS 编程基础,并通过实例逐步引入 DATA 步和 PROC 步的基本句法,为介绍 SAS 在经济统计分析中的应用奠定基础。第二篇 5 章,通过统计学中常用的图、表、分布状况特征及假设检验等方法,运用 SAS 着力对经济问题进行描述和推断分析,包括:单变量过程、相关过程、频数过程、制表过程、图形过程、假设检验和方差分析。第三篇 7 章,主要介绍 SAS 在多变量统计分析中的应用,包括:线性回归分析、主成分分析、因子分析、聚类分析、判别分析、时间序列分析等。

本书基于 SAS9.1 中文版,从案例分析入手,侧重于应用 SAS 解决实际经济问题。书中大量的案例为读者提供了学习和应用 SAS 分析和解决实际经济问题的良好范例,每章适量的习题为读者巩固学习内容提供了方便。

本书可作为高等院校经济管理类专业本科高年级学生学习 SAS 软件的教材,也可作为经济管理类专业研究生学习应用统计的教材,还可作为高等院校教师、相关经济部门、统计部门、科研单位技术人员、计算机应用人员学习和应用 SAS 的参考资料。

图书在版编目(CIP)数据

SAS 与现代经济统计分析/岳朝龙,黄永兴编著. —合肥:中国科学技术大学出版社,2009.9

普通高等教育“十一五”国家级规划教材

ISBN 978-7-312-02568-6

I. S… II. ①岳… ②黄… III. 经济统计—统计分析—应用软件, SAS—高等学校—教材 IV. F222.1

中国版本图书馆 CIP 数据核字(2009)第 106586 号

责任编辑:张善金

出版者:中国科学技术大学出版社

地址:合肥市金寨路 96 号 邮编:230026

网 址: <http://www.press.ustc.edu.cn>

电 话:发行部 0551-3602905 邮购部 3602906

印刷者:合肥义兴印务有限责任公司

发 行 者:中国科学技术大学出版社

经 销 者:全国新华书店

开 本:710mm×960mm 1/16 印张:39 字数:743 千

版 次:2009 年 9 月第 1 版 2009 年 9 月第 1 次印刷

印 数:1—4000 册

定 价:52.00 元



前 言

随着知识经济时代的到来,人们在实践中会遇到越来越多的经济信息。如何系统、完整、高效地存储、整理和分析这些信息,是经济统计工作者面临的一项重要而艰巨的任务。要完成这一任务,未来的经济统计工作者除了需要掌握一套整理和分析经济问题的统计方法外,还必须学会并能熟练地使用统计分析软件。作为数据处理和统计分析的国际标准软件、国家统计局大力推荐的产品 SAS,以其功能齐全,灵活方便,可视性强而越来越受到用户的喜爱。

目前国内许多高校的经济管理专业或与此相关的专业都相继开设了与 SAS 软件有关的课程,也出版了一些与 SAS 软件有关的教材。这些教材要么侧重于介绍 SAS 软件,要么侧重于统计方法的介绍,把两者结合起来用于解决实际经济问题的并不多见。正是基于这一考虑,我们经过长时间的酝酿,在精心策划,认真组织的基础上,基于 SAS9.1 中文版编写了这本《SAS 与现代经济统计分析》。其目的就是要为经济管理专业教师、学生以及广大的在实际经济工作第一线工作的人士提供一本既介绍 SAS 软件,又介绍统计方法应用的实用性教科书。

鉴于知识经济时代对人才素质培养的要求,在本书的编写过程中,我们始终遵循以下指导思想:

(1) 教材建设应紧密结合“新世纪高等教育教学改革工程”中的本科教育教学改革目标,体现 21 世纪高等统计教育对人才培养的素质要求,使知识传授与素质教育得到有机结合。

(2) 教材建设应体现时代特征,要把由于社会发展、科学技术的不断进步所出现的新思想、新理论和新方法融入到新的教材中。



(3) 统计软件教材的编写应以应用为主, 避免统计软件教学与统计课程教学相互脱节, 使学生在实际应用中掌握统计软件的应用技巧、统计方法的应用背景和应用条件。

为此, 我们力求使本书具有以下特色:

第一, 突出应用, 兼顾方法。本书除第一篇外, 其他各篇都是从经济问题的实际入手, 先提出问题, 再介绍统计方法的背景、思想、模型和解决思路, 然后给出 SAS 求解程序和结果的详细说明 (有的输出结果由编者进行了加工整理)。

第二, 针对和提高相统一。本书针对一般本科院校经济统计学专业, 基本涵盖本专业应掌握的主要统计方法。在简单回顾方法的同时, 重点解决方法的应用问题, 从而把以前所学内容通过 SAS 软件和实际经济问题有机地结合起来, 使所学知识得到进一步的深化和提高。

第三, 通俗和适用相结合。本书在统计方法、原理的介绍上, 不追求理论或数学上的详细推导, 而强调方法的实际应用和结果的详细说明, 从而不至于使读者因晦涩难懂的数学证明望而却步。在 SAS 过程和语句的选择上, 以常用统计方法和实际应用情况而定, 对非常用语句和过程中非常用的选择项尽可能少讲或不讲, 从而不至于使读者因记忆困难而失去学习 SAS 软件勇气。另外, 本书在相关章节对常用的 SAS 过程还增加了菜单操作方法和步骤, 这为 SAS 初学者提供了极大的方便, 使本书不仅适用于已学习过 SAS 并想进一步提高的读者, 也适用于初次接触 SAS 的读者。

本书的大部分内容先后在本校统计学和金融学专业开设的“统计分析应用软件——SAS”课程和研究生开设的“多元统计分析 with SAS”课程中讲授过。学生们对于本书的内容及形式给予了充分的肯定, 并提出了许多宝贵的意见和建议, 相当一部分已得到采纳, 并在本书的最后统稿过程中得到体现。在一定程度上, 可以说本书的最终完成是教学相长的产物。

本书分三篇 15 章。第一篇 SAS 基础。该篇系统介绍了 SAS 的特点、SAS 数据库和 SAS 数据集的创建, 其他格式数据文件与 SAS 数据集的相互转换、SAS 编程基础及常用语句, 并通过实例逐步引入 DATA 步和 PROC 步的基本句法, 为后两部分 SAS 在经济统计分析

中的应用奠定了基础。第二篇通过统计学中常用图、表,分布状况特征及假设检验等方法,运用 SAS 着力对经济问题进行描述和分析,具体包括:单变量过程、相关过程、频数过程、制表过程、图形过程、假设检验和方差分析。第三篇主要介绍 SAS 在多变量统计分析中的应用,具体包括:线性回归分析、主成分分析、因子分析、聚类分析、判别分析、时间序列分析中的 ARIMA 模型和条件异方差模型等。其中,第 1, 2, 3, 7, 8, 9, 13 章由岳朝龙编写;第 10, 11, 12 章由黄永兴编写;第 4, 5, 6 章由周世军编写,第 14, 15 章由江海峰编写。

本书被列入“普通高等教育‘十一五’国家级规划教材”,在酝酿和编写过程中,得到了安徽工业大学经济学院、教务处等部门的大力支持,经济学院前任院长严忠教授为本书的出版给予了无私的帮助,并参与前期的准备工作;中国科学技术大学相关专家审阅了本书的全部内容,并对本书出版提出了宝贵的修改意见和建议;研究生李悦、储灿春、熊宏祥等同学也参与了资料及文稿的收集与整理工作,在此一并表示衷心感谢!

限于作者的水平,书中疏漏和不妥之处在所难免,恳请同行专家、学者以及广大的读者朋友不吝赐教,以便在将来再版时使本书内容更臻完美。

岳朝龙

2009年6月8日

目 录

前言 i

第一篇 SAS 系统基础

第 1 章 SAS 系统简介 3

 1.1 SAS 系统特点及主要功能 3

 1.2 SAS 系统的启动与关闭 5

 1.3 WINDOWS 下的 SAS 应用工作空间 6

 1.4 SAS 会话实例 10

 实验 1.1 SAS 系统应用工作空间 (AWS) 12

 实验 1.2 熟悉 SAS/ASSIST 软件 12

第 2 章 SAS 数据库与 SAS 数据集 16

 2.1 SAS 数据库 16

 2.2 SAS 数据集 18

 2.3 SAS 数据集的创建 20

 2.4 几个常用的 SAS 过程 36

 实验 2.1 SAS 数据集的创建 43

 实验 2.2 SAS/ACCESS 软件访问 44

 习题 2 46

第 3 章 SAS 编程基础 47

 3.1 SAS 语法基础 47

 3.2 DATA 步中的一些常用语句 54

 实验 3.1 SAS 系统编程语句练习 (一) 91



实验 3.2 SAS 系统编程语句练习 (二)	93
习题 3.....	94

第二篇 基础统计分析

第 4 章 统计描述与 SAS 过程	101
4.1 变量的数字特征与 MEANS 过程	101
4.2 单变量分析与 UNIVARIATE 过程.....	119
习题 4.....	141
附录 本章部分例题的菜单实现	145
第 5 章 统计图表与 SAS 过程	151
5.1 统计报表与 TABULATE 过程	151
5.2 统计图与 SAS 过程	166
习题 5.....	181
附录 本章部分例题的菜单实现	186
第 6 章 简单统计分析与 SAS 过程	193
6.1 假设检验与 SAS 过程	193
6.2 相关分析与 CORR 过程.....	216
习题 6.....	224
附录 本章部分例题的菜单实现	232
第 7 章 属性数据分析与 FREQ 过程	239
7.1 属性数据简介	239
7.2 FREQ 过程	244
习题 7.....	260
附录 本章部分例题的菜单实现	262
第 8 章 方差分析与 ANOVA 过程	268
8.1 方差分析概述	268
8.2 ANOVA 过程简介	276

8.3 ANOVA 过程的应用举例	279
习题 8.....	297
附录 本章部分例题的菜单实现	300

第三篇 高级统计分析

第 9 章 回归分析与 REG 过程	311
9.1 线性回归分析方法简介	311
9.2 REG 过程（回归分析过程）简介	320
9.3 REG 过程在一元线性回归分析中的应用	325
9.4 REG 过程在多元线性回归分析中的应用	347
9.5 二元选择模型	361
习题 9.....	375
附录 本章部分例题的菜单实现	380
第 10 章 主成分分析与 PRINCOMP 过程	399
10.1 主成分分析概述	399
10.2 PRINCOMP 过程及其应用	409
习题 10.....	420
附录 本章例题的菜单实现	424
第 11 章 因子分析与 FACTOR 过程	427
11.1 因子分析方法简介	428
11.2 因子分析方法与因子旋转方法	432
11.3 FACTOR 过程及其应用	434
习题 11.....	446
第 12 章 判别分析与 DISCRIM 过程	450
12.1 判别分析简介	450
12.2 SAS 系统中的判别分析过程简介	461
12.3 判别分析过程的应用	465
习题 12.....	482



第 13 章 聚类分析	486
13.1 聚类分析概述	486
13.2 系统聚类方法简介	492
13.3 样品聚类与 CLUSTER 过程.....	498
13.4 变量聚类与 VARCLUS 过程.....	514
习题 13.....	524
第 14 章 时间序列模型与 ARIMA 过程	529
14.1 时间序列分析简介	529
14.2 ARIMA 过程	541
14.3 ARIMA 过程的应用实例	546
习题 14.....	572
附录 本章部分例题的菜单实现	574
第 15 章 ARCH 模型簇与 AUTOREG 过程	583
15.1 自相关与条件异方差 (ARCH) 模型简介	583
15.2 单位根、协整检验及误差修正模型	590
15.3 AUTOREG 过程及实例	592
习题 15.....	608
附录 本章部分例题的菜单实现	609
后记	611
参考文献	612

第一篇

SAS 系统基础





第1章 SAS 系统简介

1.1 SAS 系统特点及主要功能

1.1.1 什么是 SAS 系统

SAS 系统是大型集成应用软件系统。所谓软件系统就是一组在一起作业的计算机程序，使用 SAS 系统的用户可以根据自己的需要随时做出明智的选择。由于 SAS 系统是一种组合软件系统，因此，它具有完备的数据访问、数据管理、数据分析和报告功能。它是美国 SAS 软件研究所经多年研制，于 1976 年推出的，是目前商务智能市场上最大的供应商，已被 120 多个国家和地区的 45000^①多个机构所采用。在 2008 年的世界《财富》全球 500 强企业前 100 家企业中有 96 家是 SAS 客户。2009 年 SAS 再次获“中国杰出雇主”殊荣，此次总计有 14 家企业荣获这项殊荣。其中包括拜耳集团、滨海高尔夫、道康宁、新元素、赢创德固赛、如家、灏讯中国、毕马威、朗盛化学、赛仕中国、上海日立、上海大众、迅达、天纳克。WINDOWS 环境下运行的 SAS 系统充分利用了 WINDOWS 操作系统良好的图形界面，以及与其他系统、数据良好的互联性，大大方便了程序编辑，数据操作与管理，受到用户的普遍好评和青睐。

1.1.2 SAS 系统的组成

SAS 系统是一种组合软件系统，它的基本部分是 Base/SAS 软件，此外，还有用于统计分析的 SAS/STAT 软件，用于高级绘图的 SAS/GRAPH 软件，用于矩阵运算的 SAS/IML 软件，用于运筹学和线性规划的 SAS/OR 软件，用于经济预测和时间序列分析的 SAS/ETS 软件，等等。

^① 45000 应为 45 000，但由于本书类似数据很多，且出现的场合各异，每 3 位分节后，容易与两个或两个以上的数字混淆，故本书对 3 位以上的数字一律不予分节表示。

1.1.3 SAS 系统的特点

SAS 系统具有可扩展性、交互操作性、可管理性及可用性，其特点主要有：

1. 使用灵活方便，功能齐全

(1) SAS 的宗旨是为所有需要进行数据处理、数据分析的非计算机专业人员提供一种易学易用，完整可靠的软件系统。

(2) 使用简单、操作方便。用户把要解决的问题，用 SAS 语言（近乎自然英语的非过程语言）表达出来，组成 SAS 程序，提交给 SAS 系统就可以解决用户提出的问题。

(3) 灵活。SAS 系统提供很多语句及选项，供用户灵活地使用某个统计方法。

例如，对变量 Y 关于 $X_1—X_{10}$ 做回归分析，可以用： $MODEL Y = X_1—X_{10}$ ；或者：

$MODEL Y = X_1—X_{10} / SELECTION = STEPWISE$

(4) 功能齐全。SAS 系统的模块结构，用户根据需要可灵活地选择使用。

2. SAS 语言是功能强大且简洁易学的非过程语言

(1) SAS 语言是 SAS 系统的基础，是用户与系统对话的语言。

(2) SAS 语言是功能强大的程序设计语言，类似于 C 语言，且综合了各种高级语言的功能和灵活的格式；有 176 个标准函数和大量编程语句可用于数据的加工处理等。

(3) SAS 语言的特点是用户不必告诉 SAS “怎样做”，只需告诉它你要“做什么”就行了。

(4) 宏功能。把需要重复做的类似工作利用宏功能定义为宏，从而大大简化了 SAS 程序的编写。

3. 把数据处理与统计分析融为一体

(1) SAS 程序的结构由两个基本步组合而成。这两个基本步分别为数据步（DATA 步）和过程步（PROC 步），其中，DATA 步用于对数据进行加工处理，而 PROC 步则是用于分析数据和编写报告。

(2) 对数据的连续处理。SAS 系统克服了通常软件或偏重于数据管理功能，或偏重于统计分析的弊病，把数据管理功能与统计分析功能有机地结合在一起。它不仅具有一整套从数据输入，加工处理，文件操作，直至打印输出等完备灵活的数据管理功能，而且还能对所存贮的数据连续地进行各种统计分

析。例如，某个过程产生的输出数据，可以作为另一个过程的输入数据做进一步的分析。

4. 适应性强，应用面广

(1) SAS 系统适用于任何类型的人员，包括初学者或有经验的用户。

(2) 适用于任何类型的数据。包括各种数据库生成的数据文件，例如，DBF 文件，SPSS 文件，EXL 文件，等等。

(3) 几乎适用于任何应用的需要。SAS 系统可用来解决自然科学和社会科学各个领域中的各种问题。例如，统计学、心理学、经济学、生物学、商业、金融、社会调查，等等。适当地组合 SAS 系统的模块，可用于图形显示与分析、数据输入、检索、管理、报表生成、统计计算、市场调查、分析与预测等许多方面。

5. 扩展性能强

SAS 系统是没有上限的软件系统。

SAS 现在可以在众多 64 位操作环境下运行，从而增加了 SAS 的内存处理能力。

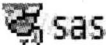
并行处理机制通过在可用 CPU 之间分配处理任务充分利用多个 CPU，因此对于下列两种类型的 SAS 处理，提高了性能：线程化 I/O 和线程化应用程序处理。

SAS 9.1 版本在以前版本的基础上又新增加了许多产品，主要有：SAS/ACCESS 与 Bloomberg Financial Data 的接口，SAS Data Surveyor 以及 SAS ETL Studio 等。

1.2 SAS 系统的启动与关闭

1.2.1 SAS 系统的启动

若计算机装有 SAS 系统，可以采用下面两种方式打开。本书以 SAS 9.1 中文版本为例。

(1) 在 WINDOWS 桌面上双击 SAS 快捷图标  就可以进入 SAS 应用工作空间 (AWS)。图 1.1 所示是刚进入 SAS 系统时的 AWS (Application Workspace) 的式样。

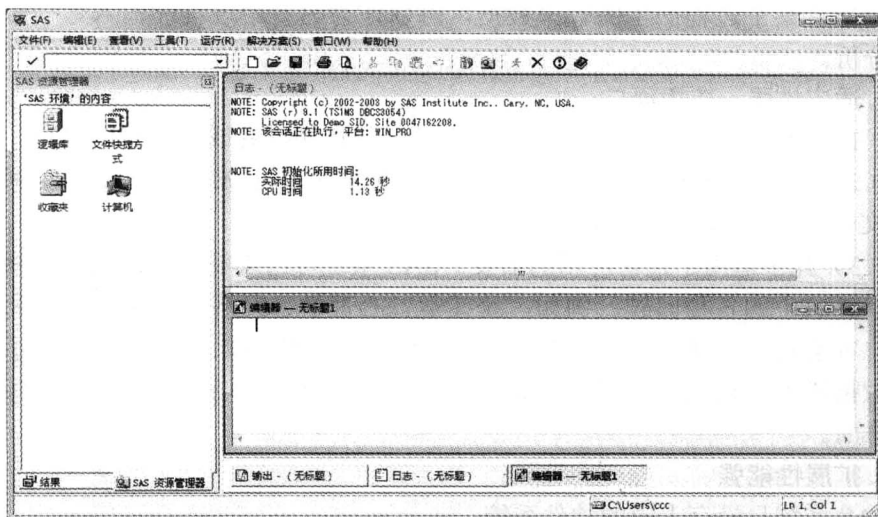


图 1.1 编辑窗口下 SAS AWS 应用工作空间

(2) 从 WINDOWS 程序菜单中选择“SAS” → 单击“SAS 9.1 (简体中文)”，同样可以进入 SAS 应用工作空间。

1.2.2 SAS 系统的退出

从“文件”菜单选“退出”，或关闭 SAS 的主窗口，即单击标题栏中的 **X**。当选择“退出”或单击图标 **X** 时，会出现一个对话框以便进一步确认是否退出。确认退出，单击“确定”，否则单击“取消”。

1.3 WINDOWS 下的 SAS 应用工作空间

SAS 应用工作空间 (AWS) 包含所有打开的 SAS 窗口及包括那些已被最小化的窗口，SAS AWS 的主要任务是为所有的 SAS 应用窗口提供基本框架。

1.3.1 主窗口

子窗口是 SAS AWS 中的各个单独的窗口。SAS 系统共有五个主窗口，分别为 SAS 资源管理器窗口，程序编辑窗口 (PROGRAM EDIT)，程序运行日志窗口 (LOG)，程序输出窗口 (OUTPUT)，图形编辑窗口 (GRAPH)。这些窗口像 WINDOWS 图形界面中的其他窗口一样，可被最大化、最小化、滚动



及重置大小，但因为它们是 SAS AWS 中的窗口，所以不能被移出 SAS AWS 窗口的边界。

SAS 资源管理器窗口用于查看并管理 SAS 文件以及创建非 SAS 格式文件的快捷方式。程序编辑窗口 (PROGRAM EDITOR) 用于输入程序，其编辑输入方法类似 WINDOWS 中的记事本的使用。运行日志窗口 (LOG) 用于记录用户程序的运行历史，其中包括用户已提交的程序代码列表，运行结果是成功的还是失败的。如果出错的话，则给出错误的位置及错误的具体信息，通常用棕红色来标记。输出窗口 (OUTPUT) 显示用户程序执行后的字符型程序结果。GRAPH 窗口显示高分辨率图形。这些窗口的操作方法与 WINDOWS 窗口操作基本一致。

1.3.2 窗口的组成

SAS AWS 中的每一窗口都由一个标准的窗口组成，主要有标题栏、菜单条、工具栏、命令栏和状态栏，其操作步骤与 WINDOWS 基本一致，所不同的是窗口中的内容。

注意，SAS 中窗口的内容会随着窗口的不同而有所改变。为使读者对菜单条有一个比较清楚的了解，以下对几个常用菜单主要功能做一简单介绍，以便读者更好地学习和运用 SAS。

在编辑窗口下，菜单条内容如图 1.2。



图 1.2 菜单条

1. “文件”菜单

在程序编辑窗口下的“文件”菜单列出了“新建程序”、“打开程序”、“关闭”、“追加”、“打开对象”、“保存”、“另存为”、“另存为对象”、“导入数据”、“导出数据”、“页面设置”、“打印预览”、“打印”和“关闭”选项。

新建程序：打开一个新的程序编辑器窗口，在 SAS 9.1 版本中可以同时运行多个程序编辑器窗口。

追加：将外部的已经存在的 SAS 编辑程序追加到现行编写的程序末尾，类似于合并程序，不同于“打开程序”（在一个新的编辑器窗口中打开已存的程序文件）。

打开对象：打开目前 SAS 逻辑库中的相关文件，可以是 LOG 条目、OUTPUT 条目、SAS 组件语言和 SOURCE 条目。