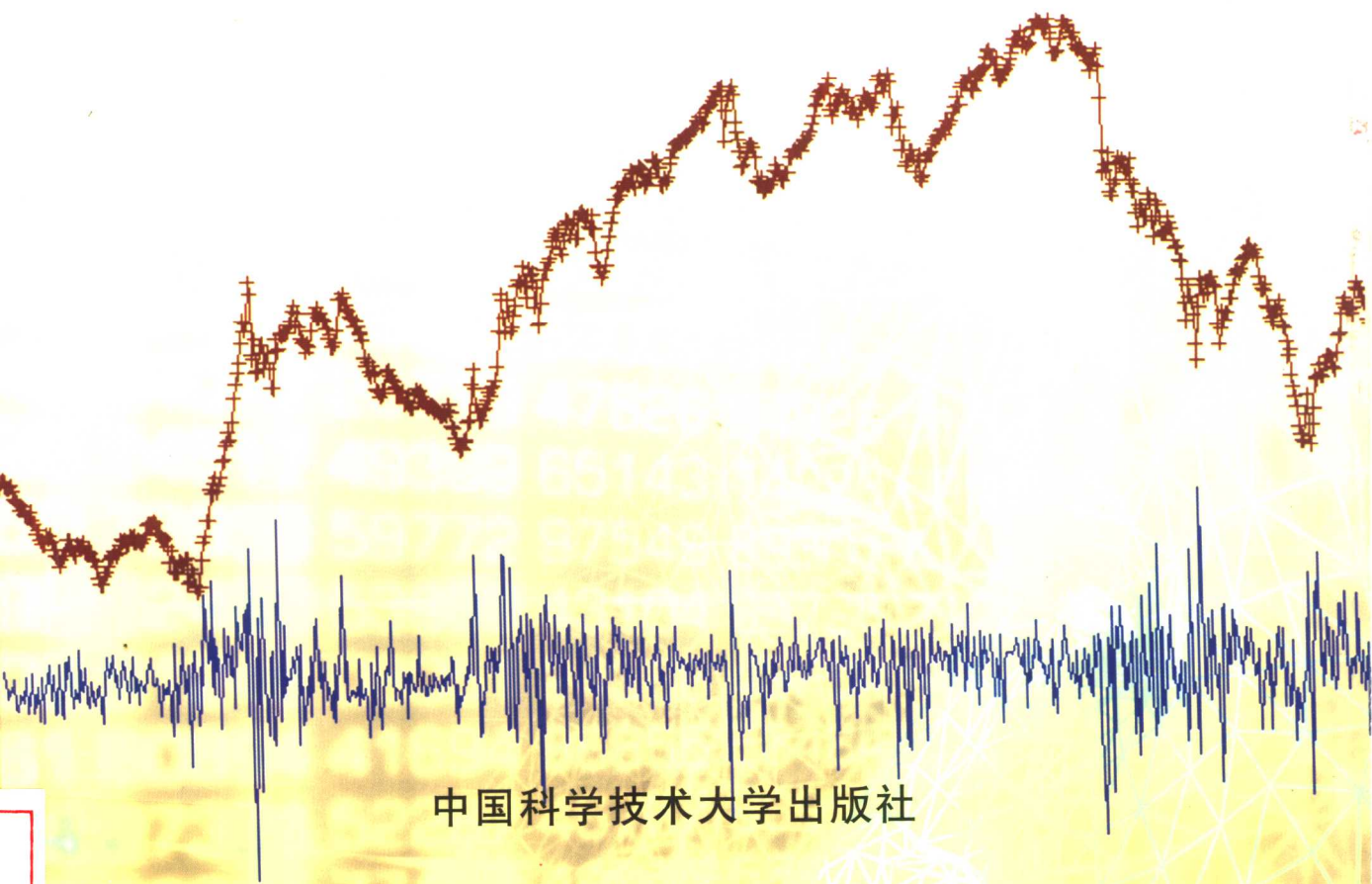




21 世纪高校规划教材 · 经济类

SAS系统与经济统计分析

岳朝龙 黄永兴 严 忠 编著



中国科学技术大学出版社

21 世纪高校规划教材·经济类

SAS 系统与经济统计分析

岳朝龙 黄永兴 严忠 编著

中国科学技术大学出版社

合 肥

内 容 简 介

本书涵盖了大部分常用的统计分析方法,是使用 SAS 系统进行经济统计分析的一本非常实用的书籍。全书分 3 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 编程基础及常用语句,通过实例逐步引入 DATA 步和 PROC 步的基本句法,为介绍 SAS 系统在经济统计分析中的应用奠定了基础。第二篇 5 章,通过统计学中常用的矩、图、表,分布状况及假设检验等方法,运用 SAS 系统着力对经济问题进行描述和分析,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 系统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、典型相关分析、时间序列分析等。

本书从案例分析入手,侧重于应用 SAS 系统解决实际经济问题。书中大量的案例为读者提供了学习和应用 SAS 系统分析和解决实际经济问题的良好范例,每章适量的习题和书末的参考答案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 系统的参考资料。

图书在版编目(CIP)数据

SAS 系统与经济统计分析/岳朝龙,黄永兴,严忠编著. —合肥:中国科学技术大学出版社, 2003. 7

21 世纪高校规划教材·经济类

ISBN 7-312-01603-0

I. S… II. ①岳… ②黄… ③严… 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(2003)第 062889 号

凡购买中国科大版图书,如有白页、缺页、倒页者,由承印厂负责调换。

中国科学技术大学出版社出版发行

(安徽省合肥市金寨路 96 号,邮编:230026,发行电话:0551—3602905,3602906)

中国科学技术大学印刷厂印刷

全国新华书店经销

开本:787mm×960mm 1/16 印张:33.25 彩插:1 页 字数:775 千

2003 年 7 月第 1 版 2003 年 7 月第 1 次印刷

印数:1—4000 册 定价:38.00 元

ISBN 7-312-01603-0/F·82

总 前 言

随着我国教育体制改革的不断深入,特别是高等教育的改革和发展,对21世纪人才培养提出了更高、更新的要求。不同类型的高等院校对人才培养的目标要体现各自的特色。作为理工科普通高等学校应侧重于应用型人才的培养。在体现宽专业、厚基础、重应用的基础上,要充分体现21世纪经济建设主战场对人才的需求。因此,需要在教学计划、教材、教学大纲以及教育教学方法等方面得到反映。对此,我校经济学院在教育部“新世纪高等教育教学改革工程”本科教育教学改革课题《一般工科院校经济学类本科生人才培养模式研究》的研究取得成果之后,组织编写了这套系列教材。

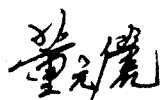
该系列教材包括《政治经济学新编》、《现代西方经济学》、《管理经济学》、《国际贸易》、《国际贸易实务》、《国际金融学》、《现代货币银行学》、《金融投资学》、《统计学》、《期货与期权交易》、《概率与数理统计》、《SAS系统与经济统计分析》等。

该系列教材的宗旨:面对21世纪经济社会发展对人才的需求,适应我国加入WTO及世界经济一体化的需求,体现一般院校经济管理类专业应用型人才培养目标。力求内容完整、重点突出、深入浅出;重点、难点内容叙述详细;注重应用和案例教学。同时,力求能够反映当今国内、国际经济科学发展前沿。

参加这套教材编写工作的作者都是长期在教学岗位上从事教学和研究工作的教授、学者,他们具有较高的学术造诣和宽厚的专业基础知识以及丰富的教学和科研经验。更为可贵的是他们思路开阔,接

受新知识、新事物的能力强,热爱本专业,具有远大的志向。当然,由于水平和经验的局限,这套教材也难免存在这样或那样的不足或欠妥之处,希望同行专家、学者和广大读者提出宝贵的意见和建议,以便他们在今后的教学和研究工作中总结、提高,从而使本书在以后再版时渐臻完美,为造就新一代高素质经济管理复合型人才做出更多的贡献。

安徽工业大学校长



2002 年 1 月 18 日

前 言

随着知识经济的到来,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。完成这一任务,未来的经济统计工作者除了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 系统,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,编写了这本《SAS 系统与经济统计分析》。其目的就是要为经济管理专业和广大的实际经济工作者提供一本既介绍 SAS 软件,又介绍统计方法应用的实用性教科书。

鉴于 21 世纪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而出现的新思想、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的现象,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。

根据以上指导思想,我们力求使本书具有以下特色:

(1) 突出应用,兼顾方法。本书除第一篇外,第二和第三篇都是从经济问题的实际着手,先提出问题,再介绍统计方法的背景、思想、模型和解决思路,最后给出 SAS 求解程序和结果的详细说明。

(2) 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。

(3) 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而望而却步。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 软件的勇气。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,这为广大的 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。为了方便读者,我们还将书中的例题以及每章的练习题中的数据和程序附在一张软盘上(见本书后记),读者可以直接拷贝上机运行。

本书的大部分内容曾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《多元统计分析与 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。值得一提的是,在本书的形成过程中,研究生江海峰、刘亚琴、肖转乔为本书提供了大量的数据资料以及书稿的校对和习题答案的验算工作,江海峰还参与了部分章节和全部附录的编写。这些支持对于本书的顺利出版起到了很大的促进作用。

本书分三篇 15 章。第一篇 SAS 系统基础。该篇系统介绍了 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 系统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的矩、图、表,分布状况及假设检验等方法运用 SAS 系统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。以上两篇共 8 章均由岳朝龙编写。第三篇主要介绍了 SAS 系统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、典型相关分析、时间序列等。其中,第 9,12,13 章由岳朝龙编写;第 10,11,14 章由黄永兴编写;第 15 章由严忠编写。

本书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,中国科学技术大学有关专家、教授对本书的出版也给予了大力支持,在此一并表示衷心感谢!

限于作者的水平和时间,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,批评指正。

岳朝龙

2003 年 6 月 18 日

目 次

总前言	(i)
前 言	(iii)

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	(7)
实验 1.1 SAS 系统应用工作空间(AWS)	(9)
第 2 章 SAS 数据库与 SAS 数据集	(10)
2.1 SAS 数据库	(10)
2.2 SAS 数据集的基本概念	(13)
2.3 SAS 数据集的创建	(14)
2.4 几个常用的 SAS 过程	(31)
实验 2.1 SAS 数据集的创建	(38)
实验 2.2 SAS/ACCESS 软件访问	(38)
习题二	(40)
第 3 章 SAS 编程基础	(41)
3.1 SAS 语法基础	(41)
3.2 DATA 步中的一些常用语句	(47)
实验 3.1 SAS 系统编程语句练习(一)	(80)
实验 3.2 SAS 系统编程语句练习(二)	(81)
习题三	(82)

第二篇 基础统计分析

第 4 章 统计描述与 SAS 过程	(89)
4.1 变量的数字特征与 MEANS 过程	(89)
4.2 单变量分析与 UNIVARIATE 过程	(102)
习题四	(119)
附录 第 4 章部分例题的菜单实现	(123)
第 5 章 统计图表与 SAS 过程	(128)
5.1 统计报表与 TABULATE 过程	(128)
5.2 统计图与 SAS 过程	(141)
习题五	(153)
附录 第 5 章部分例题的菜单实现	(157)
第 6 章 简单统计分析与 SAS 过程	(165)
6.1 假设检验与 SAS 过程	(165)
6.2 相关分析与 CORR 过程	(186)
习题六	(193)
附录 第 6 章部分例题的菜单实现	(200)
第 7 章 属性数据分析与 FREQ 过程	(205)
7.1 属性数据简介	(205)
7.2 FREQ 过程	(210)
习题七	(222)
附录 第 7 章部分例题的菜单实现	(224)
第 8 章 方差分析与 ANOVA 过程	(229)
8.1 方差分析概述	(229)
8.2 ANOVA 过程简介	(236)
8.3 ANOVA 过程的应用举例	(238)
习题八	(254)
附录 第 8 章部分例题的菜单实现	(257)

第三篇 高级统计分析

第 9 章 回归分析与 REG 过程	(265)
9.1 线性回归分析方法简介	(265)
9.2 REG 过程(回归分析过程)简介	(273)
9.3 REG 过程在一元线性回归分析中的应用	(277)
9.4 REG 过程在多元线性回归分析中的应用	(298)
9.5 逻辑斯缔回归与 LOGISTIC 过程	(311)
习题九	(319)
附录 第 9 章部分例题的菜单实现	(323)
第 10 章 主成分分析与 PRINCOM 过程	(339)
10.1 主成分分析概述	(339)
10.2 PRINCOMP(主成分分析)过程及其应用	(346)
习题十	(356)
附录 第 10 章例题的菜单实现	(358)
第 11 章 因子分析与 FACTOR 过程	(360)
11.1 因子分析方法简介	(360)
11.2 因子分析方法与因子旋转方法	(365)
11.3 FACTOR(因子分析)过程及其应用	(367)
习题十一	(378)
第 12 章 判别分析与 DISCRIM 过程	(381)
12.1 判别分析简介	(381)
12.2 SAS 系统中的判别分析过程简介	(389)
12.3 判别分析过程的应用	(393)
12.4 判别分析与逻辑斯缔回归分析比较	(406)
习题十二	(409)
第 13 章 聚类分析	(413)
13.1 聚类分析概述	(413)
13.2 系统聚类方法简介	(419)

13.3 样品聚类与 CLUSTER 过程	(424)
13.4 变量聚类与 VARCLUS 过程	(437)
习题十三	(446)
第 14 章 典型相关分析	(451)
14.1 引论	(451)
14.2 典型相关分析的基本原理	(451)
14.3 CANCORR(典型相关分析)过程	(453)
习题十四	(463)
附录 第 14 章例题的菜单实现	(467)
第 15 章 时间序列分析	(468)
15.1 时间序列综述	(468)
15.2 ARIMA 过程	(476)
15.3 自相关与条件异方差(ARCH)模型简介	(489)
15.4 AUTOREG 过程及其应用	(496)
习题十五	(509)
附录 第 15 章部分例题的菜单实现	(511)
后记	(520)
参考文献	(521)

第 一 篇

SAS 系 统 基 础

SEC 100-100

100-100-100

100-100-100

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，目前已被 120 个国家和地区的 29000^①多个机构所采用，直接用户超过 300 万。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。

1.1.3 SAS 系统的特点

一、使用灵活方便、功能齐全

(1) SAS 系统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。

^① 29000 应为 29 000，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，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分节表示。

(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 系统就可以解决你提出的问题。

(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。

如对变量 Y 关于 X_1-X_{10} 做回归分析, 可以用: $\text{MODEL } Y=X_1-X_{10}$; 或用: $\text{MODEL } Y=X_1-X_{10}/\text{SELECTION=STEPWISE}$;

(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。

二、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 SAS 程序的编写。

三、SAS 系统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骤任意组合而成。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, PROC 步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, 而且还能够对所存贮的数据连续地进行各种统计分析。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。

四、适应性强、应用面广

(1) SAS 系统适用于任何类型的人员, 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, 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。如统计学、心理学、经济学、生物学、商业、金融、社会调查等。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、分析与预测等等方面。

五、扩展性能强

SAS 系统是没有上限的软件系统。

1.2 SAS 系统的启动与关闭

1.2.1 SAS 系统的启动

(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。

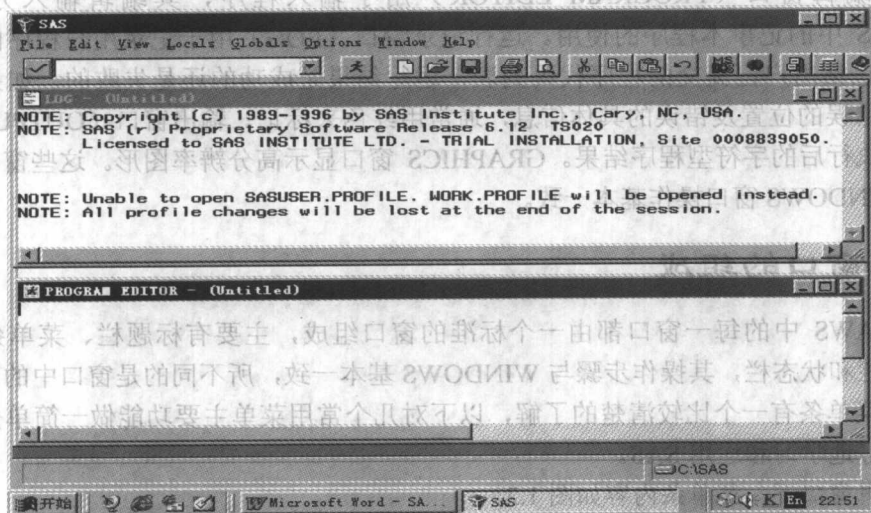


图 1.1 编辑窗口下 SAS AWS 应用工作空间

(2) 从 WINDOWS 程序菜单中选择 The SAS System→单击 The SAS System for Windows v6.12, 同样可以进入 SAS 应用工作空间。

1.2.2 SAS 系统的退出

从 File 菜单选 Exit, 或关闭 SAS 的主窗口, 即单击标题栏中的 。当选择 Exit 或单击图标  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 子窗口

子窗口是 SAS AWS 中的各个单独的窗口。SAS 系统共有四个窗口, 分别为程序编辑窗口 (PROGRAM EDIT), 程序运行记录窗口 (LOG), 程序输出窗口 (OUTPUT), 图形编辑窗口 (GRAPHICS)。这些窗口像 WINDOWS 图形界面中的其他子窗口一样, 可被最大化、最小化、滚动及重置大小, 但因为它们是 SAS AWS 中的子窗口, 所以不能被移出 SAS AWS 窗口的边界。

程序编辑窗口 (PROGRAM EDITOR) 用于输入程序, 其编辑输入方法类似 WINDOWS 中的记事本程序的使用。运行记录窗口 (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, 发现错误的位置及错误的具体信息, 通常用红色来标记。输出窗口 (OUTPUT) 显示用户程序执行后的字符型程序结果。GRAPHICS 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。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。

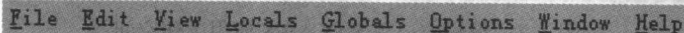


图 1.2 菜单条

1. FILE 菜单

NEW: 清空程序编辑窗口。

OPEN: 打开外部文件到程序编辑窗口。

IMPORT: 启动 SAS 的导入向导, 把其他格式的数据转换为 SAS 数据。

EXPORT: 启动 SAS 的导出向导, 把 SAS 数据转换为其他文件格式, 与 IMPORT 一起完成外部文件与 SAS 数据集之间的相互转换。