

实用SAS基础

吴令云 吴诚鸥·编著



SHIYONG SAS JICHU



东南大学出版社
SOUTHEAST UNIVERSITY PRESS

实用 SAS 基础

吴令云 吴诚鸥 编著

 东南大学出版社
SOUTHEAST UNIVERSITY PRESS

· 南京 ·

内 容 提 要

本书是学习 SAS 软件(第 9 版)的入门书。SAS 是特大规模集成软件系统,能优异地执行下列任务:数据的输入和获取,数据转换处理和管理,报表绘制和图形,统计和数学分析,商业规划,预测,运筹优化以及应用开发等。本书介绍 SAS 软件最基本的部分:SAS 系统简介;SAS 语言简介;常用的把外部数据文件输入成 SAS 数据集的方法;数据集的转换、处理、管理;输出数据的方法;制定报表和绘制图形举例;并介绍用 SAS 拟合和分析最常用的统计模型——线性和非线性回归模型的方法。每一内容和方法都举出实用例子并附一练习题加以巩固。通过学习这些内容,读者不仅能够用 SAS 完成一些简单的任务,而且根据需要,能够参照更专门的文献,更深入地学会使用 SAS 解决相应的问题,完成更复杂、更高级的任务。

本书可供应用气象、环境科学、遥感、统计、应用数学、信息、医药、生物、金融、经贸、管理、会计等方面的研究生,甚至高年级大学生作为教材或教学参考书。也可以供希望使用统计方法的实际工作者作为学习 SAS 软件的入门书籍。

图书在版编目(CIP)数据

实用 SAS 基础/吴令云,吴诚鸥编著. —南京:东南大学出版社,2016. 10

ISBN 978-7-5641-6791-2

I. ①实… II. ①吴… ②吴… III. ①统计分析—统计程序—教材 IV. ①C819

中国版本图书馆 CIP 数据核字(2016)第 240930 号

实用 SAS 基础

出版发行 东南大学出版社

出版人 江建中

社 址 南京市四牌楼 2 号

邮 编 210096

经 销 全国各地新华书店

印 刷 江苏凤凰数码印务有限公司

开 本 700 mm×1000 mm 1/16

印 张 13.25

字 数 260 千字

版 次 2016 年 10 月第 1 版

印 次 2016 年 10 月第 1 次印刷

书 号 ISBN 978-7-5641-6791-2

定 价 32.00 元

(本社图书若有印装质量问题,请直接与营销部联系。电话:025-83791830)

前 言

SAS 系统不仅是最强有力的大型统计分析软件,更是一个特大规模集成软件系统,能优异地执行下列任务:数据获取和输入、数据转换处理和管理、报表绘制和图形、统计和数学分析、商业规划、预测、运筹优化以及应用开发等,在科研和实用部门得到了广泛使用。本书通俗地介绍 SAS 软件最基本的用法,希望应用气象、环境科学、遥感、统计、应用数学、信息、医药、生物、金融、经贸、管理和会计等专业方向的研究生们,通过学习,能够在科研和实际工作中使用 SAS 软件,提高科研和实际工作能力,高质量完成毕业论文,适应就业单位需要。

本书起点不高,只要求读者们具有文科类线性代数、概率统计以及计算机的基础知识。由于 SAS 软件内容广泛,工具精深,学习起来很费时间、很困难,本书采用把 SAS“通俗化”和“实用化”的方法:不求理论精深严谨,但求容易掌握实用;不求内容全面和先进,但求掌握最最需要的基础知识。对于每一知识点和方法都举出实用例子加以说明,并且附一练习题加以巩固。好在已经有许多严谨、全面和先进的文献出现,尤其 SAS 自备的帮助系统详尽完备。希望读者们学完本书后能够从这些参考文献和资料中进一步学习提高。本书只求在读者学习 SAS 软件的过程中起到敲门砖的作用。预祝读者们能在学习本书的基础上学习更多有关 SAS 的书籍,能进一步学习 SAS 的更多知识和更好运用 SAS 软件。我们也计划按照本书风格编写一系列介绍 SAS 软件的书,以满足读者的需要。

本书共有 8 章,前 6 章介绍了 SAS 最基本的部分:SAS 系统、SAS 语言、构建 SAS 数据集、输出表格举例等。以上都属于 SAS 基础部分(BASE SAS),作为 SAS 进一步应用的例子,在第 7 章介绍了几个 SAS 作图的例子,第 8 章介绍了线性和非线性回归分析。

对于本书,自学者也可以根据需要暂不学习 3.2 节及第 6 章和第 7 章,等学完第 8 章再回过头来学习这些章节。虽然本书把 SAS 语法写在每一节的最前面,但

读者可以先重点阅读每节例题,初步了解这些语法如何使用,之后再回过头来细细阅读 SAS 语法。

本书介绍的内容是 Windows 环境下的 SAS 软件有关知识。SAS 软件以英文为主体,适当接受中文。本书主要介绍 SAS 9.1.3(纯英文版)的用法,适当介绍接受中文版 SAS 软件的用法。

本书由吴令云编写第 1~7 章,吴诚鸥编写第 8 章。由于我们水平有限,加上时间比较仓促,本书定有许多缺点及错误,衷心希望读者们批评指正,也希望读者们能够提供好的建议,特别是希望读者们能提供好的练习题或例题(要有数据)。我们将在本书再版时根据读者的建议加以修改和补充,并表示感谢。

感谢南京信息工程大学研究生院对本书的支持,将本书选为研究生院教材。

本书所用 SAS 代码(程序)和数据、配套数据库以及练习题参考答案可以在东南大学出版社网站上下载,以便于读者们练习。配套数据库请存于驱动器 D,以与书中程序一致。

编著者邮箱:Wulingyun@nau.edu.cn。

序 言

众所周知,SAS 软件是一个强有力的统计软件。SAS 不只有统计功能,它还是一个大规模集成软件(最新的 SAS 9.4 容量达到 14G),能执行许多方面的任务,如数据输入和获取、数据转换处理和管理、报表绘制和图形、统计和数学分析、商业规划、预测、运筹优化以及应用开发等。就其统计功能而言,它具有“新、全、深”的特点。“新”,是指它能很快地把最新的统计方法编成“过程”,便于统计工作者调用,避免了编程时遇到的各种困难,诸如零过多数据的统计分析、能筛选因子的非参数回归等较新模型,它都有相应过程,能方便地实施数据拟合和分析。“全”,一方面体现在它对各种统计方法都能编成“过程”;另一方面,它还有矩阵包、运筹包、画图包等近 30 个模块,在分析数据和写报告时能方便地调用。“深”,体现在它能深入使用统计理论对实际问题进行分析,例如对于线性回归的查找野点问题,SAS 提供近 10 个影响统计量进行判断。如今正值大数据时代,处理大数据更显示出 SAS 的优越性。学会运用 SAS 将使人们在实际应用与工作中如虎添翼。

初学 SAS 时可能会碰到一些困难,主要困难常出现在以下几个方面:(1) 软件太复杂(SAS 公司自编的介绍 SAS 软件的资料编成书得有一人多高),掌握起来比较困难;(2) 使用 SAS 时主要不是采用菜单式人机会话,而是要用 SAS 语言编写代码(SAS 程序)来向 SAS 软件发出指令;(3) SAS 的输出内容很多,初学者可能会觉得眼花缭乱,难以从中找出主要结果。

吴令云和吴诚鸥根据广大读者的需要编写了《实用 SAS 基础》,这是一本帮助初学者克服学习 SAS 中遇到的困难的书籍。该书具有起点低的优点,只要具有文科概率统计和线性代数基础知识以及计算机知识的人,都可以读懂此书;该书对一些较深的计算机、数学和统计知识,或浅显解释,或避而不谈,不影响读者初步掌握 SAS。该书精选内容,只介绍学习和使用 SAS 最必需的内容,诸如 SAS 软件的主要人机会话方式、输入数据建立 SAS 数据集、输出数据、整理和管理数据的基本方法,对于统计方法、制表和绘图都只介绍了其中一部分,而对于所选择的部分,较深

较细地介绍相关内容,对每一知识点通过一个易懂实例加以示范,再加一道练习题巩固。该书结合实际,列举了许多实际应用中的例题,诸如银行贷款的计算、建立工资表、避免醉酒驾车的方法等,既吸引读者兴趣,又启发读者解决实际问题。该书是学习 SAS 软件的敲门砖,既介绍 SAS 软件的基本用法,为读者深入学习 SAS 奠定基础,又介绍进一步学习 SAS 的参考文献(特别是 SAS 自带的帮助系统)。

希望统计、应用数学、信息、医药、生物、金融、经贸、管理、应用气象、环境科学、遥感等专业方向的研究生,和希望使用统计方法的实际工作者,通过阅读该书,能跨越掌握 SAS 软件的门槛,初步掌握 SAS 软件这一先进工具,解决一些实际问题,并且能进一步更深入、更细致地掌握 SAS 软件,特别是能较高水平地使用统计方法解决实际问题。

王金德

南京大学数学系

2016 年 7 月

目 录

第 1 章 SAS 软件简介	1
1.1 关于 SAS 软件	1
1.1.1 SAS 软件简介	1
1.1.2 如何学习 SAS 软件	3
1.2 SAS 系统人机会话窗口	5
1.3 SAS 中的中文字符	11
第 2 章 SAS 语言	13
2.1 SAS 语言简介	13
2.2 SAS 数据集	14
2.3 SAS 词、SAS 语句和 SAS 程序	16
第 3 章 构建 SAS 数据集	24
3.1 data 语句、input 语句和 datalines 语句配合数据体产生 SAS 数据集	25
3.2 用“表编辑器”窗口输入法形成 SAS 数据集	34
3.3 input 语句配合 infile 语句导入数据体产生 SAS 数据集	37
3.4 用“导入数据”窗口导入已定变量数据体形成 SAS 数据集	41
3.5 libname 语句配合数据引擎导入已定变量数据体形成 SAS 数据集	44
3.6 时间变量值的输入格式	46
第 4 章 预处理 SAS 数据集	48
4.1 用赋值语句增加变量	48
4.1.1 一般赋值语句	48
4.1.2 SAS 函数和随机数	50
4.2 用 drop 或 keep 语句剔除变量	53

4.3	用 set 语句的 firstobs 和 obs 选项或 set 语句配合 if 语句删减观测行	55
4.4	用 set 语句纵向拼接 SAS 数据集	58
4.5	用 sort 过程把数据集排序	61
4.6	用 merge 语句横向拼接 SAS 数据集	63
第 5 章	SAS 数据集的导出	67
5.1	SAS 数据集的导出方法	67
5.1.1	用 filename 语句、file 语句配合 put 语句导出数据体	68
5.1.2	用“导出数据”窗口导出 SAS 数据集	69
5.1.3	用 libname 语句配合数据引擎导出 SAS 数据集	71
5.2	SAS 数据集变量名和变量观测值输出的改进	72
5.2.1	用单独一条 format 语句按规定格式加强对输出时间变量和数值变量值的描述	73
5.2.2	用 format 过程自行创建替换格式	75
5.2.3	用 format 过程与 format 语句配合加强对输出变量值的描述	77
5.2.4	用 label 语句加强对输出变量名的描述	81
5.3	用 options 语句、title 语句、footnote 语句改变页面输出	83
5.3.1	用 options 语句改变输出的页面选择	83
5.3.2	用 title 语句为输出的图形或表格加标题	85
5.3.3	用 footnote 语句为输出的图形或表格加脚注	87
第 6 章	SAS 制表举例	89
6.1	将数据体的每一行写成一页报告	89
6.2	用 tabulate 过程输出报表	93
6.2.1	TABULATE 过程简介	93
6.2.2	用 tabulate 过程打印一维统计表	96
6.2.3	用 tabulate 过程打印二维统计表	99
6.2.4	用 tabulate 过程打印三维或更高维度的统计表	101
6.2.5	用 tabulate 过程打印统计表来显示分类变量各种交叉中样本的个数	106
6.2.6	对 tabulate 过程打印统计表的改进	107

第 7 章 SAS 制图举例	110
7.1 用 gplot 过程画时间序列图、散点图和气泡图	110
7.1.1 用 gplot 过程画散点图和时间序列图	110
7.1.2 用 gplot 过程画气泡图	113
7.2 用 gchart 过程画柱形图、条形图和饼形图	115
7.2.1 用 gchart 过程画柱形图	117
7.2.2 用 gchart 过程画条形图	119
7.2.3 用 gchart 过程画饼形图	119
第 8 章 回归分析	121
8.1 多元线性回归与 reg 过程	121
8.1.1 线性回归模型	121
8.1.2 用 reg 过程作多元线性回归	126
8.1.3 用 reg 过程作预报	130
8.1.4 选择预报因子	133
8.1.5 多项式回归	139
8.1.6 共线性问题、岭回归和主成分回归	144
8.1.7 野点查找和处理法	161
8.2 非线性回归	176
8.2.1 非线性回归模型	176
8.2.2 用 nlin 过程作非线性回归分析	179
8.2.3 优良初估计的获得	185
8.2.4 用 nlp 过程作 LAD 回归	191
参考文献	199

第 1 章 SAS 软件简介

SAS 是特大规模集成软件系统,其统计模块是最强有力的大型统计分析软件。它在发达国家被广泛使用,在我国的使用人数也迅猛增长。科研人员、研究生、本科生都需要它,大公司、金融机构也需要它;理工科学生需要它,文科学生也需要它。为了更好地掌握和使用 SAS 软件,本章将简单介绍 SAS 软件最常用的人机会话系统以及 SAS 软件的特点,还将介绍使用 SAS 软件的注意事项,尤其是 SAS 系统中使用中文字符的注意事项。

本章中,1.1 节简单介绍 SAS 软件,1.2 节介绍常用 SAS 人机会话窗口,1.3 节介绍 SAS 中使用中文字符的注意事项。

1.1 关于 SAS 软件

1.1.1 SAS 软件简介

SAS 是全球最强有力的大型统计分析软件,由美国 SAS 研究所(SAS Institute)研制,于 1966 年推向市场。该软件不断更新:第 1 版、第 2 版……第 9 版,同一版还再细分为不同发布(Release),如 8.0、8.1、8.2;9.0、9.1、9.2、9.3、9.4。不断出现的版本把最新的统计、数学和计算机的理论、方法和软件技术引入其中。SAS 具有以下优点:

(1) 内容最丰富。SAS 9.0 的容量达到 1.5G,SAS 9.3 和 SAS 9.4 的容量猛增至 14G;其统计模块包揽众多统计方法,可以说是统计方法大全。下面先举一个例子说明其广博程度。一般统计软件只有 1~2 个过程作线性回归,而 SAS 给出十多个过程作线性回归,每个过程都有其特色,其中有 3 个过程用作非参数回归。一般参数回归要求给出因变量的函数表达式,而 SAS 给出的这 3 个过程不要求给出因变量的函数表达式,就能作非参数回归,这就大大拓展了回归的使用范围。再举一个例子说明 SAS 的精深程度。回归分析在查找野点(反常点)时,一般统计软

件用残差或个别影响统计量来发现野点, SAS 则指出用残差来评估野点存在偏差, 必须通过影响分析来查找野点, 要计算影响统计量。一般统计软件只计算一两个影响统计量, 由这一两个影响统计量的大小来判断每个观测值是否是野点, 由于这些统计量带有随机性, 单独由一个统计量的值判断每一个观测值是否是野点作出的判断不全面, 易出偏差。SAS 则给出约 10 个影响统计量, 由这些统计量是否过大来判断每一个观测值是否是野点, 这样就保证查找野点的可靠性, 由此可见 SAS 的精深。除统计方法外, SAS 软件还包括质量管理、运筹学等数十个模块。

(2) 能利用最新科研成果和技术。最新的统计方法和计算技术等可以在 1~2 年内由 SAS 实现, 推荐给客户。科研和财经工作者、统计分析者可以避免费时、复杂的编程工作而直接使用最新统计方法和最新计算技术, 例如 SAS 9.4 的 `adaptivereg` 过程能让你执行非参数回归的模型选择功能, SAS 的 `genmod` 过程能让你方便地拟合零膨胀负二项分布模型。

(3) 能方便地从各种类型的数据库如 Dbase2、Oracle、Sybase 中导入、导出数据。

(4) 具有多种数据处理功能, 如 SQL 语言的功能都能用 SAS 轻松实现。

(5) 计算精度高。数据输入、输出采用双精度, 内部计算时使用三精度, 计算方法又很先进, 不易出现溢出现象。

(6) 容纳变量多, 只要计算机性能足够, 一般来说足够容纳你所需要的变量。

(7) 图形功能强, 有 25 种颜色可用于同一图形中的不同线型, 又有多种绘图功能, 如可以在地图上画出气泡图, 显示各地的 GDP 值。

(8) 输出结论最丰富。例如用 `freq` 过程做检验能输出 7 个检验统计量的检验结果, 而一般统计软件只输出 1~2 个检验统计量的检验结果。

(9) 便于写出内容丰富的报告。特别是便于分析云数据, 因而适用于大型企业、金融公司和科研部门对海量数据进行分析。

(10) SAS 语言是以 C 语言为工作母语而编出的更高级语言, 比 C 语言更接近人类自然语言。学习 SAS 语言有助于学习人工智能。

正因为有如此多的优点, SAS 是大数据时代读入数据、分析数据、写出分析报告的最有力工具, 是科研和实践领域的必备工具。许多 SCI 期刊要求论文必须用 SAS 计算数据, 欧美各发达国家的各大公司几乎都使用 SAS 软件, 以致 SAS 公司发行企业版。SAS 软件的使用是一些金融、经济、管理等专业研究生必修的课程, 能够使用 SAS 软件也是许多大公司招工的必要条件。某份美国雇主们对雇员掌握统计软件需求的调查报告指出: 需要雇员掌握 SAS 的雇主约有 3000 人, 需要掌握 MATLAB 的约 1000 人, 需要掌握 SPSS 的不到 800 人, 需要掌握 STATA 和 SPLUS 的不到 50 人。当前我国经济发展迅速, 数据分析的需要迅猛增长, 作为经

济大国必须通过 SAS 统计软件进行高效科学的统计分析,大公司、科研部门和政府机关对 SAS 软件的需求正不断增加,相应地,许多大学开设了用 SAS 软件分析数据的相关课程。

本书就是为需要使用高深统计工具的读者编写的。

1.1.2 如何学习 SAS 软件

SAS 软件用处很大,但是学习 SAS 软件存在困难。主要有以下几点:

(1) SAS 软件不断根据科学技术的发展更新内容,想要掌握甚至了解新技术都有困难。

(2) SAS 内容广博深厚,有近 20 个模块,成百个过程,每个过程能解决一大类问题,还有约千条 SAS 语句,每条语句都是用 C 语言作为工作母语编成的独特语句,每句又有许多选项,每个选项能指示 SAS 完成独特的任务,内容太多。

(3) 当给了实际问题后,SAS 加以计算和分析,输出的结果很多,有时有好几页,其中有许多统计术语,这样多的输出对于深入分析解决问题是很必要的,但是初学者见了这么多的图形表格确实会感到难以理解。

(4) SAS 软件自带的帮助系统全面系统地介绍 SAS 软件的功能,但该文献是英文写成,而且内容极其丰富,其全部内容若打印出来的话厚度超过一人高,初学者直接阅读的话很困难。

总之 SAS 软件是个精巧的巨无霸,学习起来自然感到困难,掌握它不容易。

对于初学者来说,一开始就全面系统地学习 SAS 软件必然要花大量时间和精力,而且感到枯燥难记。本书的目的是为读者们提供捷径,让他们快速迈过 SAS 的门槛。对本书的做法可以打个不恰当比方:德语句子有数以千句,一个未学过德语的成年人,想去德国学习某门技术,例如汽车制造,如果系统学完德语再去德国,不仅很费时间,而且感到索然无味;如果先学“常用德语 50 句”,会了就到德国去生活,在德国再根据实际需要不断学习德语知识,那样学习德语会感到收效快而且学得有意义。本书采用类似“常用德语 50 句”的方法,用较小篇幅,只介绍最基本、最必要的 SAS 知识,包括 50 句左右 SAS 语句的简单用法,这样初步学习使用 SAS 的方法就容易了,而且在初步使用 SAS 的过程中能够加深对 SAS 语句的理解和记忆。此外,SAS 具有强大的输出功能,基于高深的统计方法算出许多统计量,对此初学者也容易眼花缭乱,不知哪些是输出图表时所需要的。因此本书也介绍寻找和阅读主要 SAS 输出的方法,使读者们能够解决一些常用实际问题。掌握 SAS 软件的这些基本用法后,读者们可以根据需要,进一步查找有关 SAS 的文献,用 SAS 提供的精妙理论解决更多问题。就好像学完常用德语 50 句后,你就能在德国生活,根据你在德国所从事的专业,进一步学习有关德语。也许这样的学习 SAS 的

方法能更好地适应非统计学专业学习者的需求。

本书的特点是：

(1) 只讲最常用的 SAS 知识,不讲较少用的 SAS 知识;只讲简单易学的 SAS 知识,不讲复杂精巧但是费时较多的 SAS 知识。例如 array 语句能简化编程过程,但本着“学得少一点、精一点”的原则,本书暂时不介绍 array 语句,对 ods 语句和 sql 过程也是同样暂时不介绍。虽然 array 语句、ods 语句和 sql 过程能巧妙执行许多任务,但是本书不介绍这几部分,因而可能书中介绍的方法笨一点,繁琐一点,但是需要记忆的东西较少,读者更易接受和掌握。随着你使用 SAS 次数的增多,对于 SAS 了解到了一定程度,对于本书知识很熟悉后,再学习这两条语句和过程也就不难了。

(2) 以需要 SAS 完成的最常用任务为纲,不以 SAS 语法的系统为纲,就好像人们从常用德语 50 句学起,而不是按德汉字典顺序学起。

(3) 多举实例,以简单数据的例子代替复杂的例子。

(4) 只讲用法,少讲原理。

总之易学和实用是本书的宗旨,不求系统,不求严谨和面面俱到。

读者应当结合每个内容的例题学习相关知识(SAS 语法),并且通过练习题加以巩固。当然本书毕竟太基础,太简单了,而本书姊妹篇一时不能写成。对于掌握本书后想要进一步学习 SAS 者,建议首先参考 4 份文献:

(1) SAS 软件自带帮助系统。可通过“帮助(Help)”→“SAS 帮助与文档(SAS Help and Documents)”进入,该文献详尽而权威,但是它用英文写成,又有许多专业术语,可能只在深入钻研某一过程时可以查询。

(2) 夏坤庄等人编写的《深入解析 SAS》,该书是由 SAS 公司在中国的员工写的,是唯一一本全面介绍 SAS 软件的书。

(3) 朱世武编写的《SAS 编程技术教程》,该书全面广泛介绍一般的 SAS 编程技术。

(4) Lora D. Dlwiche 与 Susan J. Slaughter 所著的 *The Little SAS Book*。

此外读者们还可以通过学习其他文献掌握 SAS 过程的应用:例如想学习用 SAS 作经济(金融)数据挖掘者,可以学习姚志勇编写的《SAS 编程与数据挖掘商业案例》;想学习用 SAS 作统计分析者,可以学习高惠璇等人编写的《实用统计方法与 SAS 系统》和他们有关 SAS 的一系列著作,以及岳朝龙等人编写的《SAS 系统与现代表经济统计分析》;想学习用 SAS 作属性数据分析者,可以学习官琳琳等人编写的《基于 SAS 的属性数据分析》。有关 SAS 的文献很多,不再一一列举。应当注意后 3 份文献成书较早,所用 SAS 版本均在 SAS 9.3 之前,其中的程序虽然也可以在 SAS 9.4 上运行,但是得到的 SAS 9.4 的输出比这 3 本书中介绍的 SAS 输

出多了许多内容。

本书介绍的方法能用来分析云数据。但为了让读者们更直观理解处理 SAS 数据集的真实情况,本书列举的 SAS 数据集不是云数据,相反,变量很少,观测次数也很少。读者们应当知晓:可以用本书介绍的分析简单数据的方法同样去做云数据的数据处理和数据分析工作,最好能参照朱世武编写的《SAS 编程技术教程》一书,使用瑞思公司的金融数据网站(www.resset.cn)提供的数据作分析。

SAS 也有菜单式人机会话界面(例如 SAS/ASSIST),但是菜单式会话有局限性,不如使用 SAS 语言编程序,更便于施展 SAS 的巨大功能。本书仅在极个别处介绍菜单式人机会话的用法。

SAS 软件有企业版,可安装在 UNIX 等系统上。本书采用的是 Windows 环境下的 SAS 窗口环境。

安装 SAS 软件时可以选择各国的文字字符,对中国的读者们来说,兼容中文字符是必须的,但是许多用户使用的是纯英文软件,不兼容中文字符,所以本书以纯英文的 SAS 9.1.3 为基本教学软件,必要时也介绍兼容中文字符的其他版本 SAS 软件的用法。

本书所用程序和数据、配套数据库、练习题参考答案,读者们可以在东南大学出版社网站下载,以便于练习。

1.2 SAS 系统人机会话窗口

本书采用 Microsoft Windows 环境下的 SAS 人机会话窗口环境。一般第 9 版 SAS 软件以倒四棱锥为快捷启动图标(图 1-1),直接点击它就启动 SAS 系统,进入 Microsoft Windows SAS 下会话默认界面。



图 1-1 SAS 软件快捷启动图标

SAS 9.1.3 会话默认界面的顶部有菜单栏、命令框和工具栏,以及程序编辑器(Editor)窗口、日志(Log)窗口、输出(Output)窗口,左侧长方形窗口中左上角有缩小的结果(Results)窗口,右下角有缩小的浏览器(Explorer)窗口(图 1-2)。

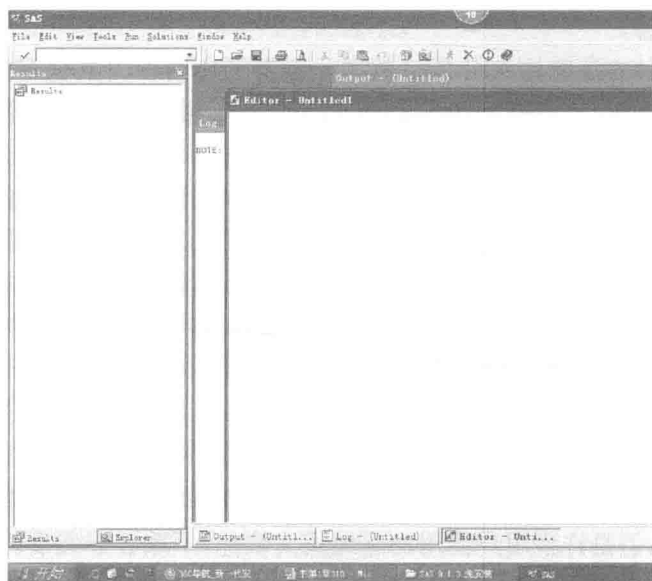


图 1-2 Microsoft Windows 环境下默认的 SAS 会话界面

SAS 9.1.3 中的人机会话主要通过 SAS 系统的 6 个窗口以及菜单栏和工具栏进行。最常用的 4 个窗口是：程序编辑器(Editor)窗口、日志(Log)窗口、输出(Output)窗口和图形(Graph)窗口。其余的还有结果(Results)窗口和浏览器(Explorer)窗口，如果窗口不全，在菜单栏的“查看”(英文为“View”，SAS 9.4 中译为“视图”)菜单内点击这些窗口的名字可以显示它们。

(1) 程序编辑器(Editor)窗口：用于存放给计算机的命令(一般为 SAS 程序)，工具栏中跑步的小人为“提交(Submit)”图标，F3 键是“提交”快捷键。

(2) 日志(Log)窗口：用于追踪显示程序运行过程，当程序有错误时，提交程序时日志窗口将错误语句以红字标出，往往还给出错误原因和修改建议。有时 SAS 能智能猜测用户的错误，它会说明用户的原意是什么，SAS 按用户的正确原意执行。当然日志(Log)窗口的用途很多，不只是查找错误，还随时提供计算中的信息，诸如变量个数、观测值个数、所用 CPU 时间等。

(3) 输出(Output)窗口：用于存放计算机的计算结果。SAS 的计算结果很多，在输出窗口形成许多数表，每个数表用表头区分。注意：从输出窗口复制计算结果时，不要用鼠标选中整行，而是在最后一行漏选几个字符，这样才能复制成功。

(4) 图形(Graph)窗口：用于输出图形结果，具有较高分辨率。复制图形时可以利用图中复制功能。

(5) 结果(Results)窗口：用于把计算机计算所得的数表名和图形名排成树形图以便查找。

(6) 浏览器 (Explorer) 窗口: 包含 4 个子窗口, 其中最常用的是逻辑库 (Library) 子窗口, 它可以用来查找和建立逻辑库。所有的 SAS 文件都用以“.”分开的两级符号命名,“.”前(第一级)是逻辑库名,“.”后(第二级)是在逻辑库中用于辨别文件的成员名。例如某文件全名为 work.sale, 则第一级“work”是逻辑库名, 第二级“sale”是在逻辑库中用于辨别文件的成员名。构造逻辑库名和成员名的规则是, 以字母或下划线开头, 并且名称中只能包含英文字母、数字和下划线。逻辑库名不能超过 8 个字符, 而成员名却可以有多达 32 个字符。Library 子窗口中有多个逻辑库。有的 SAS 文件的名称中只有成员名, 没有逻辑库名和“.”, 这样的文件称为临时文件, 当结束 SAS 系统后, 这些文件就自动消失了。SAS 自动指定这样的临时文件的逻辑库名是“work”, 所有只有成员名的文件构成的临时逻辑库名为 work。

以下举例介绍这几个窗口的应用。

例 1.1 假如某人准备在 2018 年 1 月购房, 房款 100 万元, 需要贷款 70 万元, 分 30 年还清, 每月还同样的数目, 年利率是 4%, 那么每月需要还多少钱?

解 SAS 有个 loan 过程, 专门用来计算有关贷款的问题。本例中每月还同样的数目, 属于 loan 过程可用方法中的固定还款额问题。这个问题只有 3 个参数: 贷款总额(本例 700000 元), 年利率(本例 4%), 还款期限(本例 30 年, 等于 360 个月), 只要采用如下 SAS 程序即可:

```
proc loan start=2018:1; /* 调用 loan 过程, 在 2018 年 1 月借款 */
fixed amount=700000 rate=4 life=360 out=house; /* 还款额固定, 总借款额为 70 万,
利率为每年 4%, 还期为 360 个月, 还款细节存数据集 house */
proc print data=house; /* 将数据集 house 的内容打印到 Output 窗口 */
run;
```

将程序代码拷贝到 Editor 窗口, 点击提交图标或按 F3 键, 则在 Output 窗口得到 2 张数表, 其中第 1 张表如下:

Downpayment	0.00	Principal Amount	700000.00
Initialization	0.00	Points	0.00
Total Interest	503085.52	Nominal Rate	4.0000%
Total Payment	1203085.52	Effective Rate	4.0742%
Pay Interval	MONTHLY	Compounding	MONTHLY
No. of Payments	360	No. of Compoundings	360
Start Date	JAN2018	End Date	JAN2048

由前 2 列可见全部利息为 50.308 552 万元, 共付 120.308 552 万元, 按月计算, 共还贷 360 个月, 2018 年 1 月开始。由后 2 列可见贷款额为 700000.00 元, 名