

# SAS 系统

## Base SAS 软件 使用手册

高惠璇等 编译

# SAS

OPTIONS WI  
HELP MENU KE  
VAR FILENAME LIBN  
DATA INPUT PUT OUT  
MERGE SET LENGTH LA  
UPDATE CARDS DO GOTO  
BY CLASS FILE INFILE WI  
IF-THEN / ELSE SELECT  
ARRAY ATTRIB DROP KEEP  
RENAME RETAIN PROC MO  
QUIT FORMAT INFORMAT X  
PRINT CORR MEANS UNIVA  
CHART TABULATE SUMMAR  
PLOT FORMS CALENDAR S  
COPY DATASETS CATALOG  
REPORT SQL CONTENTS  
PMENU COMPARE FORM  
CUT BYE LOG MACRO  
PROGRAM EDITOR D  
KEEP = DATA = OBS  
EXP LOG10 GAMM  
NORMAL UNI

中国统计出版社

# SAS 系统

## Base SAS 软件使用手册

高惠璇等 编译

中国统计出版社

# (京)新登字 041 号

版权所有。未经许可，本书的任何部分均不得以任何形式重印、复制、拷贝、翻译。

## 图书在版编目(CIP)数据

SAS 系统·Base SAS 软件使用手册/高惠璇等编译。

—北京：中国统计出版社，1997.4

ISBN 7-5037-2407-2

I. S…

II. 高…

III. 数据管理系统, SAS-手册

IV. TP317-62

## SAS 系统使用手册编委会

主 编：高惠璇

副主编：李贵斌 耿 直

编 委：(按姓氏拼音为序)

陈家鼎 陈莎莉 高惠璇 耿 直

何书元 蒋建成 李贵斌 李东风

刘大双 刘朝峰 栾贻会 栾世武

孙万龙 吴 岚 杨宝慧 张 平

中国统计出版社出版

(北京三里河月坛南街 75 号 100826)

新华书店经销

科伦克三来印务有限公司印刷

\*

787×1092 毫米 16 开本 37 印张 90 万字

1997 年 4 月第 1 版 1997 年 4 月北京第 1 次印刷

印数：1—3000

\*

定价：60.00 元

# 前 言

SAS 系统是大型集成应用软件系统,具有完备的数据访问、数据管理、数据分析和数据呈现功能。它是美国 SAS 软件研究所(SAS Institute Inc.)经多年的研制于 1976 年推出的。目前已被 120 个国家和地区的两万九千多个机构所采用,直接用户超过三百万人。SAS 系统广泛应用于金融、医疗卫生、生产、运输、通讯、政府、科研和教育等领域。它运用统计分析、时间序列分析、运筹决策等科学方法进行质量管理、财务管理、生产优化、风险管理、市场调查和预测等业务,并可将各种数据以灵活多样的各种报表、图形和三维透视的形式直观地表现出来。在数据处理和统计分析领域,SAS 系统被誉为国际上的标准软件系统,并且 1996 年《Datamation》评选 SAS 系统为建立数据仓库的首选产品。

北京大学概率统计系自 1987 年以来对 SAS 系统进行了近十年的开发应用,并于 1990 年至 1996 年成功地举办了多期 SAS 系统学习班,同时为学习班编写了六本《SAS 系统使用手册》、应用统计 SAS 系列丛书:《SAS 系统与基础统计分析》等资料,这些资料很受用户欢迎。

本书是 Base SAS 软件的使用手册。Base SAS 软件是 SAS 系统的基础和核心。它用于数据管理、基础统计计算、管理用户的使用环境、调用其它 SAS 软件产品等。Base SAS 软件可以单独使用,也可以和其它 SAS 产品共同组成一个功能强大的软件系统。本书共分三部分二十二章。第一部分介绍 SAS 系统初步,通过在 DOS 操作系统下和在 Windows 操作系统下运行 SAS 系统的简例介绍如何启动和使用 SAS 系统。第二部分介绍 SAS 语言,包括 SAS 表达式、SAS 函数、用于 DATA 步和 PROC 步的 SAS 语句、SAS 文件、SAS 显示管理系统等。第三部分介绍 Base SAS 软件提供的一些过程,包括用于基础统计计算、报表生成、得分计算和一些实用的过程。

本书是以北京大学概率统计系 1990 年 6 月首次编译出版的《SAS 统计分析系统(SAS/PC 入门和 SAS 语言)使用手册》(编译者:高惠璇、张平、栾世武)和 1992 年 4 月编译出版的《SAS 统计分析系统(Base SAS 软件:Base SAS 过程)使用手册》(编译者:高惠璇),以及 1993 年和 1995 年两次再版的使用手册为基础,根据 SAS 北京办事处提供的 SAS 系统 6.06 至 6.11 版本的大量资料重新编译而成的。全书由高惠璇负责组织编译和排版工作。耿直教授编译了 FREQ 过程;张利华、王海燕同志分别参加了 PMENU 过程和 SQL 过程的初译工作;蒋建成同志对 SAS 宏功能一章的编译作了部分工作;其余部分均由高惠璇编译。全书的审校工作由 SAS 北京办事处的王勇等先生完成。限于我们的水平,加上对 Base SAS 软件的全面功能的了解还不是很深透,书中难免出现错误或不妥之处,敬请读者批评指正。

本书在编译出版过程中得到了 SAS 软件研究所北京办事处的大力支持。他们提供了有关 Base SAS 软件的全部原版资料,并赠送给北京大学概率统计系一套 Windows 下 SAS 系统 6.11 版本软件。多年来 SAS 北京办事处对北京大学概率统计系编译出版的一系列中文 SAS 资料给予极大支持。在此我们表示衷心的感谢。

本书能得以顺利出版,我们还应感谢国家统计局培训部和中国统计出版社。

编译者

1997 年 1 月

# 本书的用法

《SAS 系统·Base SAS 软件使用手册》是参考美国 SAS 软件研究所出版的有关 Base SAS 软件的几本资料编写而成。主要的参考资料为

[1] SAS Language; Reference, Version 6, First Edition. 5th printing, September 1995.

[2] SAS Procedures Guides, Version 6, Third Edition. 4th printing, June 1994.

其它的参考资料为参考文献中的[3]~[10]。

## 一、本书的目的

《SAS 系统·Base SAS 软件使用手册》介绍 Base SAS 软件在 DOS 操作系统下的 6.03 或 6.04 版本、Windows 操作系统下的 6.11 版本或其它操作系统下的 6.06 及以后版本中 SAS 语句的功能和 Base SAS 过程的用法。书中所有例子均为使用 PC 机的 DOS 操作系统或 Windows 操作系统下得到的结果。

用户如果想找出您正在使用的 Base SAS 软件是哪一个版本,请查看您的 SAS 作业生成的 SAS 日志(LOG)的开头部分。如果您在 PC 机的 DOS 操作系统上使用 SAS 系统,您可以继续使用北京大学概率统计系编写的《SAS 系统使用手册(一)》和《SAS 系统使用手册(二)》两本资料。

您还需要读“本书的用法”的其余部分,以便了解本书关于读者的假设和关于 Base SAS 软件在出现的信息中使用的约定。

## 二、本书的读者

《SAS 系统·Base SAS 软件使用手册》帮助想要使用 SAS 系统完成各类分析的用户更有效地使用它。该书对 SAS 系统的新用户、有一定经验的 SAS 用户、熟悉其它编程语句而无 SAS 经验的用户等都是很有帮助的。根据用户的需要,参考本手册的有关部分,用户可以完成最终的使用、应用开发系统、或系统编程等工作。

## 三、预备知识

下表描述使用 Base SAS 软件的用户必须了解有关 SAS 系统的概念。

您必须知道如何做	参考
所使用的操作系统	用户使用的主机系统的资料
在您的计算机上安装 SAS 系统	由当地 SAS 软件顾问提供的说明

## 四、如何使用本书

下面介绍本书所包括信息的概况、其结构及如何使用。

### (1) 本书的结构

《SAS 系统·Base SAS 软件使用手册》分为三大部分。第一部分介绍 SAS 系统入门;第二部分介绍 SAS 语言;第三部分介绍 Base SAS 过程。

书中每一部分都提供一些例子。用户可以把这些例子中的 SAS 语句和数据输入到您的计算机上,然后运行这个 SAS 程序。另一方面,这些例子作为您编写自己程序的模式也是有用的。运行每个过程所必须的语句和句法在本书中采用一致的格式。

Base SAS 过程中每个过程的叙述分为以下几个主要部分:

- 简介 简要叙述每个过程是做什么的。介绍背景资料,包括一些定义和有必要引入的例子
- 语句说明 介绍该过程句法的参考部分。首先概括介绍所用语句的句法,然后叙述 PROC 语句,接着介绍所有其它的语句。每个语句的选项也依次被介绍
- 专题介绍 介绍该过程涉及到的一些专题,如统计背景、缺失值的处理、计算方法、输入和输出数据集及打印输出等
- 应用例子 给出使用本过程的一些例子,包括例子的数据、SAS 语句和打印输出。用户可以复制例子中的这些 SAS 语句和数据到您的计算机上,然后运行这个 SAS 程序。另外,这些例子中有许多都包含在 SAS 例子库中。可以通过 SAS 软件顾问来了解如何使用您的计算机上的 SAS 例子库

## (2) 用户需要读哪几部分

根据用户使用 Base SAS 软件的不同经验水平,你将要阅读本书的不同章节。下表列出对不同经验水平的用户建议阅读的部分,供参考。

如果您是…	建议您阅读…
有经验的 SAS 用户	不必系统阅读本书,根据当前任务的需要有选择地阅读本书的有关章节。
SAS 系统的新用户,但具有其它编程语言的经验	首先阅读第一部分和第十五章;然后根据当前任务的需要有选择地阅读第二和第三部分的有关章节。
SAS 系统的新用户,而且也没有使用计算机的经验	首先阅读第一部分和第四、五和十五章;然后根据当前任务的需要有选择地阅读第二和第三部分的有关章节。

## 五、约定

本节介绍这本书使用的一些约定,包括字体约定、句法约定、以及呈现在输出中使用的约定。

### (1) 正文字体的约定

本书中用到几种类型的字体,下面列出字体约定的概况:

- 宋体字 这是本书正文使用的基本字体
- 黑体 用于标题;在正文中用来表示一些很重要的位置
- 楷体 在正文中用于指出用户应该注意的内容
- 大写英文白正体 这是在正文中用于表示 SAS 语言的要素,如 PROC,REG,DATA 等。另在正文波及到的一些变量名、文件名等有时也用大写白正体
- 英文白斜体 用来指出您应提供的句法中的元素
- 小写英文白正体 用于介绍句法的一些例子或专用名词。在很多情况下,本书对 SAS 程序代码使用小写字母。其实您可以在您自己的 SAS 代码中用小写码、大写码、或两者混用。SAS 系统把您输入的变量名字改变为大写,但字符变量的值如果您已用小写码输入则必须保留用小写码。输入的标题、注释和脚注必须严格地按您希望在输出中出现的方式输入

## (2) 句法约定

句法约定用于给出一个 SAS 语句的基本格式。本书对句法使用下列约定：

大写英文黑体	表示函数、语句和宏名字的 SAS 关键词。这些名字必须严格按显示的形式拼写(例如, INPUT, PROC PRINT 等)
斜体	表示您应提供的元素(例如语句: VAR <i>variables</i> ; 和选项 END= <i>variable</i> 中由用户提供的变量名等)
大写英文白正体	表示语句中这些项您必须严格按显示的形式拼写(例如, DATA =, PAGE 等)
<>(角括号)	表示可选择的元素
(竖条)	表示从一组选项中选择一项。用竖条分开的这些项或者是相互不相容的或者是一些代号
…(省略号)	表示可以多次重复的一些项

另在语句中必须按显示的形式加上标点, 比如圆括号、逗号和分号。

下面例子说明这些句法约定:

```
FOOTNOTE<n> <'text' | "text">;
```

FOOTNOTE 用大写黑体, 因为它是一个 SAS 语句的名字; *n* 和 *text* (文本) 用斜体, 因为这是用户要提供的信息; 用户提供的 *text* (文本) 可以用单引号括起来, 也可以用双引号括起来, 但不能两者混用。

```
BY <DESCENDING> <GROUPFORMAT> variable-1 ...  
  <<DESCENDING> <GROUPFORMAT> variable-n >;
```

BY 用大写黑体, 因为它是一个 SAS 语句的名字; DESCENDING 和 GROUPFORMAT 用大写白正体, 因为他们是用户必须严格按显示的形式拼写的部分; *variable* 用斜体, 因为这是用户要提供的变量名; 用户至少必须提供一个变量名, 而 DESCENDING 和 GROUPFORMAT 可任选其中之一, 或两者。第二行用 <> 括起来的部分表示用户可以类似地列出多组变量及选项。

注: 以上这些约定是本书排版中使用的编写格式。特别是句法的约定, 目的是让用户阅读时更清楚。在 SAS 语言中, 除输入/输出的字符串外(一般均用引号括起来), 字母不分大小写。

## (3) 例子和输出的约定

在同一章里, 应用例子依次编号, 如第十九章的第 5 个例子编号为例 19.5。在同一章里的过程输出结果也依次编号, 如第二十章的第 15 个输出结果编号为输出 20.15。过程产生的每一页输出结果都用一个长方形框子围住。大多数例子的过程输出结果在方框的右上角从 1 开始依次编号, 而且每页输出都给出标题(缺省的或定制的)。另外本书的很多程序使用 SAS 系统的下列选项运行:

```
LINESIZE=78 或 132    设置正文每行的长度为 78 或 132 个字符(过程输出用 132)。  
PAGESIZE=60         设置正文每页的长度为 60 行。  
NODATE              在输出中不显示日期和时间。
```

在使用其它选项的地方, 通常与输出结果一起用 SAS 代码加以说明。有的时候, 用户运行了这些例子, 可能得到稍有差别的输出结果。这是由于您的计算机是否使用协处理器而产生的, 不是软件的问题。而且在很多情况下, 这种差别是非常小。

# 目 录

## 本书的用法

## 第一部分 SAS 系统入门

<b>第一章 SAS 系统概况</b> .....	(2)
§ 1 SAS 系统的特点 .....	(2)
§ 2 Base SAS 软件简介 .....	(4)
§ 3 SAS 系统的其它软件产品 .....	(5)
§ 4 SAS 系统的资料 .....	(7)
§ 5 安装 SAS 系统对环境的要求 .....	(8)
<b>第二章 SAS 系统会话实例</b> .....	(11)
§ 1 在 DOS 环境下的应用实例 .....	(11)
§ 2 SAS 系统的运行方式 .....	(16)
§ 3 在 Windows 环境下使用 SAS 系统 .....	(24)
§ 4 使用 SAS/ASSIST 通用菜单系统 .....	(29)
<b>第三章 SAS 数据集与 SAS 过程</b> .....	(36)
§ 1 SAS 数据集 .....	(36)
§ 2 SAS 过程 .....	(41)

## 第二部分 SAS 语言

<b>第四章 SAS 编程基础</b> .....	(54)
§ 1 使用 SAS 语言 .....	(54)
§ 2 SAS 表达式 .....	(61)
§ 3 SAS 函数 .....	(68)
§ 4 SAS CALL 子程序 .....	(84)
<b>第五章 DATA 步入门</b> .....	(87)
§ 1 DATA 步简介 .....	(87)
§ 2 DATA 步语句及其类型 .....	(89)
§ 3 DATA 步流程 .....	(92)
<b>第六章 用在 DATA 步的文件操作语句</b> .....	(95)
§ 1 DATA 语句(数据语句) .....	(95)
§ 2 INPUT 语句(输入语句) .....	(97)
§ 3 CARDS 与 CARDS4 语句及 DATALINES 与 DATALINES4 语句 .....	(107)
§ 4 PUT 语句(输出语句) .....	(108)
§ 5 BY 语句 .....	(112)
§ 6 SET 语句(组合语句) .....	(114)
§ 7 MERGE 语句(合并语句) .....	(119)

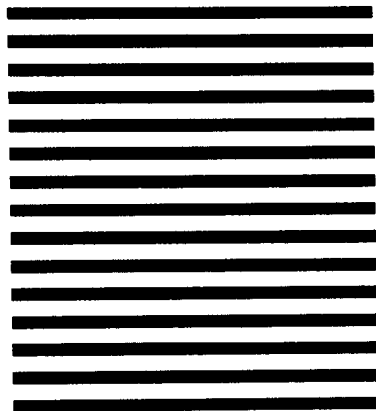


§ 8	UPDATE 语句(更新语句)	(121)
§ 9	MODIFY 语句(修改语句)	(123)
§ 10	FILE 语句(文件语句)	(128)
§ 11	INFILE 语句(打开文件语句)	(133)
§ 12	DISPLAY 语句(显示窗口语句)	(137)
<b>第七章</b>	<b>用在 DATA 步的运行语句</b>	(139)
§ 1	赋值语句和累加语句	(139)
§ 2	DELETE 语句和 LOSTCARD 语句	(141)
§ 3	STOP 语句和 ABORT 语句	(143)
§ 4	WHERE 语句	(146)
§ 5	OUTPUT 语句(输出到数据集语句)	(151)
§ 6	REMOVE 语句与 REPLACE 语句	(153)
§ 7	其它运行语句	(156)
<b>第八章</b>	<b>用在 DATA 步的控制语句</b>	(160)
§ 1	DO 语句(循环语句)	(160)
§ 2	END 语句	(164)
§ 3	SELECT 语句(选择语句)	(164)
§ 4	IF 语句(条件语句)	(166)
§ 5	GO TO 语句与语句标号	(168)
§ 6	LINK 语句(连接语句)	(170)
§ 7	RETURN 语句(返回语句)	(172)
§ 8	CONTINUE 语句与 LEAVE 语句	(173)
<b>第九章</b>	<b>用在 DATA 步的信息语句</b>	(175)
§ 1	ARRAY 语句(数组语句)	(175)
§ 2	INFORMAT 语句和 FORMAT 语句	(183)
§ 3	LENGTH 语句和 LABEL 语句	(186)
§ 4	ATTRIB 语句(属性语句)	(189)
§ 5	DROP 语句和 KEEP 语句	(190)
§ 6	RENAME 语句和 RETAIN 语句	(191)
§ 7	WINDOW 语句(窗口语句)	(194)
<b>第十章</b>	<b>PROC 步及在 PROC 步通用的语句</b>	(200)
§ 1	PROC 步入门	(200)
§ 2	PROC 语句(过程语句)	(201)
§ 3	VAR 语句和 MODEL 语句	(202)
§ 4	WEIGHT 语句和 FREQ 语句	(202)
§ 5	ID 语句和 WHERE 语句	(203)
§ 6	CLASS 语句和 BY 语句	(204)
§ 7	OUTPUT 语句和 QUIT 语句	(205)
§ 8	FORMAT 语句、LABEL 语句和 ATTRIB 语句	(206)
<b>第十一章</b>	<b>全程语句(用在任何地方的 SAS 语句)</b>	(208)
§ 1	注释语句(Comment 语句)	(208)
§ 2	DM 语句和 X 语句	(209)
§ 3	TITLE 语句和 FOOTNOTE 语句	(211)

§ 4	RUN 语句和 ENDSAS 语句 .....	(213)
§ 5	LIBNAME 语句(库名字语句) .....	(214)
§ 6	FILENAME 语句(文件名语句) .....	(216)
§ 7	%INCLUDE 语句、%RUN 语句和 %LIST 语句 .....	(219)
§ 8	MISSING 语句、PAGE 语句和 SKIP 语句 .....	(223)
§ 9	OPTIONS 语句 .....	(224)
<b>第十二章</b>	<b>SAS 输出和错误的处理</b> .....	(225)
§ 1	SAS 日志 .....	(225)
§ 2	SAS 程序的输出结果 .....	(228)
§ 3	改变日志和输出的地点 .....	(230)
§ 4	错误类型和错误的处理 .....	(231)
§ 5	影响显示输出外观的选项 .....	(237)
<b>第十三章</b>	<b>SAS 输入格式和输出格式及缺失值</b> .....	(238)
§ 1	输入格式(informat) .....	(238)
§ 2	输出格式(format) .....	(244)
§ 3	SAS 日期、时间和日期时间的输入格式和输出格式 .....	(250)
§ 4	缺失值 .....	(254)
<b>第十四章</b>	<b>SAS 文件</b> .....	(258)
§ 1	SAS 数据集 .....	(258)
§ 2	SAS 索引 .....	(266)
§ 3	SAS 目录 .....	(268)
§ 4	SAS 数据库 .....	(271)
§ 5	有关 SAS 文件的一些新概念 .....	(273)
<b>第十五章</b>	<b>SAS 显示管理系统</b> .....	(280)
§ 1	引言 .....	(280)
§ 2	SAS 窗口环境 .....	(280)
§ 3	命令 .....	(287)
§ 4	SAS 显示管理系统的任务 .....	(294)
§ 5	SAS 文件编辑器 .....	(303)
<b>第十六章</b>	<b>SAS 系统选项</b> .....	(312)
§ 1	SAS 系统选项的类型 .....	(312)
§ 2	改变 SAS 系统选项的设置 .....	(316)
§ 3	SAS 系统选项的描述 .....	(319)
<b>第十七章</b>	<b>SAS 宏功能</b> .....	(323)
§ 1	宏变量 .....	(323)
§ 2	宏 .....	(327)
§ 3	宏参数 .....	(329)
§ 4	具有分支的宏 .....	(331)
§ 5	宏工具中的输入和输出 .....	(335)
§ 6	DATA 步接口程序 .....	(336)
§ 7	宏程序语句和宏函数 .....	(338)

### 第三部分 Base SAS 过程

<b>第十八章 Base SAS 过程概况</b> .....	(342)
§ 1 统计概念综述 .....	(342)
§ 2 Base SAS 过程简介 .....	(350)
<b>第十九章 SAS 基础统计过程</b> .....	(352)
§ 1 MEANS 过程(均值过程) .....	(352)
§ 2 SUMMARY 过程(概括过程) .....	(359)
§ 3 UNIVARIATE 过程(单变量过程) .....	(362)
§ 4 CORR 过程(相关过程) .....	(369)
§ 5 FREQ 过程(频数过程) .....	(382)
§ 6 TABULATE 过程(制表过程) .....	(391)
<b>第二十章 SAS 报表过程</b> .....	(412)
§ 1 PRINT 过程(打印过程) .....	(412)
§ 2 PLOT 过程(图形过程) .....	(419)
§ 3 CHART 过程(图表过程) .....	(427)
§ 4 FORMS 过程(标签过程) .....	(439)
§ 5 CALENDAR 过程(日历过程) .....	(444)
§ 6 TIMEPLOT 过程(时间间隔图形过程) .....	(450)
<b>第二十一章 SAS 得分过程</b> .....	(456)
§ 1 STANDARD 过程(标准化过程) .....	(456)
§ 2 RANK 过程(秩得分过程) .....	(460)
<b>第二十二章 SAS 实用过程</b> .....	(465)
§ 1 SORT 过程(排序过程) .....	(465)
§ 2 FORMAT 过程(定义格式过程) .....	(469)
§ 3 TRANSPOSE 过程(转置过程) .....	(486)
§ 4 OPTIONS 过程(系统选项过程) .....	(493)
§ 5 PRINTTO 过程(输出地点过程) .....	(495)
§ 6 APPEND 过程(附加过程) .....	(499)
§ 7 CONTENTS 过程(内容过程) .....	(501)
§ 8 COPY 过程(拷贝过程) .....	(506)
§ 9 DATASETS 过程(数据库管理过程) .....	(509)
§ 10 CATALOG 过程(目录过程) .....	(518)
§ 11 COMPARE 过程(比较过程) .....	(523)
§ 12 CPORT 过程和 CIMPORT 过程 .....	(530)
§ 13 DBF 过程和 DIF 过程 .....	(535)
§ 14 DOWNLOAD 过程和 UPLOAD 过程 .....	(538)
§ 15 PMENU 过程(菜单过程) .....	(540)
§ 16 SQL 过程(结构查询语言过程) .....	(552)
§ 17 REPORT 过程(报表过程) .....	(567)
§ 18 其它过程 .....	(574)
<b>参考文献</b> .....	(580)



## 第一部分

# SAS 系统入门

---

第一章	SAS 系统概况 .....	(2)
第二章	SAS 系统会话实例 .....	(11)
第三章	SAS 数据集和 SAS 过程 .....	(36)

# 第一章 SAS 系统概况

SAS 系统是世界领先的信息系统,它由最初的用于统计分析经不断发展和完善而成为大型集成应用软件系统;具有完备的数据存取、管理、分析和显示功能。在数据处理和统计分析领域,SAS 系统被誉为国际上的标准软件系统。

SAS 系统于 1966 年由美国北卡罗连纳州(North Carolina)州立大学开始研制,1976 年成立美国 SAS 软件研究所公司,并开始对 SAS 系统进行维护、开发、销售和培训等工作。1985 年推出了 SAS/PC(6.02)版本。自 SAS 系统推出以来,它的版本更新很快,功能也不断增加。如自 1989 年推出 SAS/PC(6.04)版本后,目前已推出适用于多种操作系统的最新版本 6.11。

美国 SAS 软件研究所自创建以来,就以 SAS 系统的卓越功能和可靠的技术支持著称于世,逐步发展成为全球最大的独立软件开发商之一。目前 SAS 软件研究所世界六大洲的 60 个国家设有分公司及分支机构。世界上有 119 个国家和地区(包括中国)的 29000 多个机构在使用 SAS 系统,直接用户超过三百万人。国际上还有一个 SAS 协会 SUGI(SAS User Group International),每年都召开交流会对 SAS 系统的使用进行交流和研究。SAS 系统在世界范围内被广泛地应用于政府、科研、教育、生产和商业等不同领域,发挥着积极作用。

## § 1 SAS 系统的特点

### 一、什么是 SAS 系统

SAS 系统是一个模块化的集成软件系统。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,随时作出明智的选择。不论你从事简单还是复杂的工作,不论你是 SAS 系统的初学者还是老用户,SAS 系统都能满足你的需要。

SAS 系统是一种组合软件系统。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。

### 二、SAS 系统的特点

SAS 系统的最大特点是把数据管理和数据分析融为一体。具体地说,它具有如下几方面的特点:

#### 1. 使用灵活方便、功能齐全

- 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。
- 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。执行的情况和输出结果都在屏幕上显示出来。例如进行回归计算时使用以下简单的 SAS 程序:

```
proc reg data=a;
    model y=x1-x10;
run;
```

用户操作是在很友好的界面(显示管理系统)下进行的。

- 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。如以上例子中加上以下选项可进行逐步回归:

```
model y=x1-x10/selection=stepwise;
```

加上以下语句可以画出 Y 对 X1 的散布图:

```
plot y * x1=' * ';
```

- 功能齐全: 从 SAS 系统的模块结构, 用户根据需要可灵活地选择使用。SAS 系统提供的二十多个模块(产品)可完成各方面的实际问题, 功能非常齐全。
2. SAS 语言是编程能力强且简洁易学的非过程语言
    - SAS 语言是 SAS 系统的基础, 是用户与系统对话的语言。
    - 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; 约有 150 个函数和大量编程语句可用于数据的加工处理等。
    - SAS 语言是非过程语言(第四代语言)。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。如上例中用简单的几个语句告诉 SAS 你将数据集 A 中的数据建立 Y 与 X1 至 X10 的多元线性回归模型。
    - 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。
  3. SAS 系统把数据处理与统计分析融为一体
    - SAS 程序的结构由两个基本步骤任意组合而成。DATA 步用于对数据的加工处理; PROC 步用于分析数据和编写报告。
    - 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、加工处理直至文件操作, 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, 可以作为另一过程的输入数据(必要时用 DATA 步进行处理)做进一步地分析。
  4. 适用性强、应用面广
    - 人员: SAS 系统适用于任何经验的人员(初学者或有经验用户); SAS 系统可满足从生产第一线到管理部门各方面人员对信息的需求(SAS 提供多种界面供不同人员使用)。
    - 适用于任何类型的数据: SAS 系统可以访问各种数据资源。包括各种数据库生成的数据文件(如 dBASE 产生的 \*.DBF 文件), 其它软件(如 SPSS)生成的数据文件, 高级语言生成的数据文件等。
    - 适用于几乎是任何应用的需要: 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、考古学、医疗卫生、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、图形分析、数据输入、数据检索、数据管理、报表生成、统计计算、工程计算、市场研究、销售管理、供应管理、订货追踪、调查分析、建立预测模型、管理信息系统、行政信息系统等方面。
- SAS 系统提供的面向对象式开发工具使得用户可在 SAS 系统众多产品的基础上快速开发

出界面友好、功能强大的应用系统。

- 适用于任何计算机环境: SAS 系统具有“多硬件厂商结构”可支持多种硬件平台,它可以在大型计算机,小型计算机及微机上运行。对软件环境, SAS 系统可以在十几种不同的操作系统下运行(如 DOS、OS/2、Windows、Windows NT、UNIX、VSE、MVS、CMS、OpenVMS、PRIMOS、AOS/VS 等)。
  - 适用于今日和未来: 由 SAS 系统的模块结构,用户在任何时候都可根据需要选择最适合的产品组成自己的信息系统。将来当计算机环境发生变化时,用户仍可以用同样的方式工作,并发挥不同系统各自的优势。
5. 扩展性能强——没有上限的软件系统
- SAS 系统的模块式结构可以让用户灵活地按自己的需要去选择相应的模块与 Base SAS 软件一起完成各种应用的需要。
  - SAS/TOOLKIT 是 SAS 系统功能扩充工具。利用该产品用户不仅可扩充 SAS 系统的功能,又保护了用户已有的资源。
  - 用 SAS/IML 的编程语言编写特殊的应用程序。不论你的工作变得多么复杂, SAS 系统都可以伴你一起工作。

#### 6. 利用 SAS 系统可实现灵活强大的分布式计算环境

SAS 的应用逻辑分解及远端任务提交功能,远端数据库功能以及数据传输服务使用户不但可方便地实现 Client/Server 模式的应用,而且可实现更多样的分布式处理方案。

#### 7. SAS 系统可以作为用户唯一需要的最经济的软件系统

由以上的几个特点可以看出, SAS 系统是一个大型的集成软件系统。Base SAS 软件可以作为数据库来管理数据;系统提供的 120 多个可以灵活使用的过程,用户可直接使用,或稍加修改,以满足用户的特殊需要。用户只要学会 SAS 系统,它可以替代你学习其它程序设计语言或其它统计软件包。SAS 系统可以成为用户唯一需要的最经济的软件系统。

## § 2 Base SAS 软件简介

Base SAS 软件是 SAS 系统的核心。主要功能是数据管理和数据加工处理,并有报表生成和描述统计的功能。Base SAS 软件可以单独使用,也可以同其它软件产品一起组成一个用户化的 SAS 系统。Base SAS 软件提供以下几方面功能:

### 一、数据管理功能

#### 1. 信息存贮和检索

SAS 系统可用任何格式读入数据值,然后把数据组成 SAS 数据集。它具有很强的与外部文件交换信息的功能,可以用文件操作管理方法把不同数据集的数据组合在一起,供 SAS 过程分析处理,生成报表。SAS 数据集可用临时或永久数据集两种形式存贮,它具有自动生成文档的能力,即既包含数据值,又含有它们的描述信息。利用 SAS 数据库的特殊结构可使维护工作降到最低的限度。

#### 2. 数据修改与程序设计

SAS 系统为用户提供完备的 SAS 语句和函数用于数据加工处理;有些语句用于执行标准操

作,如建立新变量,累加求和及修改错误;另外还提供了 DO/END,IF-THEN/ELSE 等控制语句,构成一套完整的语言系统。由于 Base SAS 软件具有非常强的数据管理功能,所以许多人亦把 Base SAS 软件作为数据库管理系统。

### 3. 文件操作

数据分析常常需要从几个数据集中抽取一些变量和观测进行组合。Base SAS 软件中有用于对数据进行编辑、整理、连结、合并及更新的功能,还能同时处理多个输入文件,或者对一次输入的数据能够生成几种报表等。

## 二、基础统计计算功能

Base SAS 软件中有一些基本过程和 SAS 函数能够完成基础统计计算。它包括:

- 计算简单的描述统计量,如均值、标准差、极差、总和、平方和、偏度、峰度、分位数和相关系数等;
- 对数据进行标准化、求秩及有关统计量;
- 生成并分析列联表;
- SAS 函数可用于计算概率分布函数、分位数、样本统计量及产生随机数。

## 三、报表生成和图形显示功能

- Base SAS 软件输出数据与读入数据时同样可以采用任何格式。除了 SAS 过程产生的固定格式的报表外,用户可根据自己的需要来设计报表的输出格式,包括输出文件到磁盘上。
- SAS 的一些基本过程可绘制水平的或垂直的直方图、饼图、块图和星形图,还可以画散布图、曲线图、层次图、时间序列图等。

## § 3 SAS 系统的其它软件产品

SAS 系统是由众多产品组成的模块化的大型集成系统,其中 Base SAS 软件是 SAS 系统的基础核心。其它软件产品均是在 Base SAS 软件提供的环境中使用。用户根据自己的应用需要选择相应的产品或产品中的部分过程与 Base SAS 软件一起构成一个用户化的 SAS 系统。

SAS 系统除 Base SAS 软件外还包括几下一一些产品:

**SAS/STAT** ——这是一个完整可靠的统计分析软件包。包括回归分析、方差分析、属性数据分析,多变量分析、判别与聚类分析、残存分析、心理测验分析和非参数分析等八类方法共 40 多个过程。每个过程还提供多种不同的算法及选择,从而组成一个庞大而完整的统计分析方法集。该软件产品是国际统计分析领域的标准软件。

**SAS/ETS** ——用于计量经济与时间序列分析的专用软件。它是研究复杂系统和进行预测的有力工具。利用该软件可建立各种统计模型进行所关心系统的模拟和预测。SAS/ETS 包含全面的时间序列时域分析和谱域分析,如实用预测(逐步自回归、指数平滑、Winters 方法)、序列相关校正回归、分布滞后回归、ARIMA 模型、状态空间方法、谱分析和互谱分析等。还提供许多处理时间序列数据的实用程序,如时间频率转换和插值,X11 季节调整等。

**SAS/OR** ——用于运筹学和工程管理的专用软件。该软件提供全面的运筹学方法,是一种强有力的决策支持工具。它辅助人们实现对人力、时间以及其它各种资源的最佳利用。该软件包含通用



的线性规划、整数规划以及混合整数规划和非线性规划方法。还包含用于项目管理,时间安排和资源分配等问题的一整套方法。

**SAS/QC** —— 用于质量控制的专用软件。该软件提供全面质量管理的一系列工具。它研究的对象是过程,既可以是产品制作过程,也可以是某种特定的工作。它提供从发现和明确问题所在及进行试验设计到过程控制图和进行过程的能力分析。还提供一套全屏幕菜单系统引导用户进行标准的统计过程控制以及试验设计。

**SAS/IML** —— 该软件提供功能强大的面向矩阵运算的编程语言,它是用户研究新算法或解决 SAS 系统中没有现成方法的工具。SAS/IML 中处理的基本数据元素是一个矩阵;它允许用户直接用矩阵代数的记号来组成 IML 的程序语句。

**SAS/GRAPH** —— 这个强有力的图形软件包能够完成多种绘图功能,如生成等值线图、二维和三维曲线图、直方图、圆饼图、区块图、星形图、地理图及各种映象图。这些图形非常形象、直观地表现各变量之间的关系及数据的分布状态,对解决各种实际问题起着重要的辅助作用。SAS/GRAPH 还有一个全屏幕图形编辑器,用户可以在幅面上自由地绘制文字及图形元素,也可以对图形进行修改。

**SAS/FSP** —— 这是一个用来进行数据处理的交互式菜单系统。可用来进行全屏幕的数据录入、编辑、查询及数据文件的创建等,它也是一个开发工具。

**SAS/AF** —— 这是一个应用开发工具,利用 SAS/AF 的屏幕设计能力及 SCL 语言的处理能力可快速开发各种功能强大的应用系统。SAS/AF 采用先进的 OOP(面向对象编程)的技术,使用户可方便快速地实现各类具有图形用户界面(GUI)的应用系统。

**SAS/ASSIST** —— 该软件为 SAS 系统提供了面向任务的菜单驱动界面。它可免去用户学习 SAS 语言的负担,同时 SAS/ASSIST 生成的 SAS 程序既可辅助有经验的用户快速编写 SAS 程序,又可帮助用户学习 SAS 语言。该软件是利用 SAS/AF 开发的产品,用户还可以根据自己的需要调用 SAS/ASSIST 的不同部分或裁剪 SAS/ASSIST 的菜单来构成自己的应用系统。

**SAS/ACCESS** —— 该软件是对目前许多流行数据库的接口组成的接口集,它提供的与外部数据库的接口是透明和动态的。

**SAS/EIS** —— 该软件是 SAS 系统中采用新兴的面向对象的编程模式(OOP)的又一个开发工具。该产品也称为行政信息系统或每个人的信息系统。

**SAS/ENGLISH** —— 该软件为用户提供利用自然英语获取信息的能力。

**SAS/CALC** —— 它是一个功能完善的多维电子表格软件,具有财务分析、数据建模、数据整合及管理的能力。

**SAS/INSIGHT** —— 该软件为可视化数据探索工具。

**SAS/LAB** —— 这是菜单驱动面向任务的解释引导式数据分析软件。

**SAS/CPE** —— 这是计算机及网络系统运行状况分析评估软件。

**SAS/GIS** —— 这是集地理信息系统功能与空间数据的显示分析于一体的软件。

**SAS/SHARE** —— 这是 SAS 系统中进行数据库的并发性控制的专门软件。

**SAS/CONNECT** —— 在标准网络环境下,通过该软件可使各平台的 SAS 系统建立内在联系,实现分布处理,从而有效地利用各平台数据和机器资源。

**SAS/NVISION** —— 这是用于高标准影像制品的设计和制作的软件。

**SAS/SPECTRAVIEW** —— 这是进行多维数据观察、分析、研究的一个交互式立体可视化工