

<http://www.phei.com.cn>

SAS

统计分析应用

(第2版)

© 董大钧 主编



统计分析教材



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

014043965

C819
83-2

SAS

统计分析应用

(第2版)

◎ 董大钧 主编
◎ 申笑颜 董丽 李茜 副主编



统计分析教材



北航 C1731313

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

C819
83-2

014043882

内 容 简 介

本书基于 SAS 9.2 中文版编写,介绍了 Base SAS、SAS\STAT 的用法。全书以统计分析方法为主线,通过大量实例,详细介绍了 SAS 程序设计方法及各种统计过程适用条件和使用方法,并对统计过程实例的输出结果做了详尽的解释。本书中的例题和分析数据均放在华信教育资源网 www.hxedu.com.cn,供读者下载使用。

本书既可作为高等院校本科生、研究生信息统计分析相关课程的教材,也可作为从事数据分析和数据管理的人员的参考用书。

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,侵权必究。

图书在版编目(CIP)数据

SAS 统计分析应用/董大钧主编. —2 版. —北京:电子工业出版社,2014.5
统计分析教材
ISBN 978-7-121-22793-6

I. ①S… II. ①董… III. ①统计分析-应用软件-高等学校-教材 IV. ①C819

中国版本图书馆 CIP 数据核字(2014)第 062627 号

策划编辑:秦淑灵

责任编辑:秦淑灵

印 刷:三河市双峰印刷装订有限公司

装 订:三河市双峰印刷装订有限公司

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本:787×1092 1/16 印张:28.5 字数:804 千字

印 次:2014 年 5 月第 1 次印刷

印 数:3 000 册 定价:55.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

质量投诉请发邮件至 zlt@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

前 言

在信息时代的今天，人们在工作实践中会获取大量的信息，存储、整理和分析处理它们是一项极其重要的工作。由于对数据的分析大都是基于基本的统计原理进行的，国内外学者多年来编制了许多统计软件包，如 BMDP、SAS、SPSS 等，其中 SAS 软件包则是诸多统计软件包中的佼佼者。

SAS 系统全称为 Statistics Analysis System，是用于决策支持的大型集成信息系统，经过多年的发展，SAS 已被全世界 120 多个国家和地区的六万多家机构采用，遍及金融、医药卫生、生产、运输、通信、政府和教育科研等领域。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，是建立数据库的首选产品，堪称统计软件界的巨无霸。

SAS 系统具有灵活的功能扩展接口和强大的功能模块，在 Base SAS 的基础上，还可以增加不同的模块而实现不同的功能，如 SAS/STAT(统计分析模块)、SAS/GRAPH(绘图模块)、SAS/QC(质量控制模块)、SAS/ETS(经济计量学和时间序列分析模块)、SAS/OR(运筹学模块)、SAS/IML(交互式矩阵程序设计语言模块)、SAS/FSP(快速数据处理的交互式菜单系统模块)、SAS/AF(交互式全屏幕软件应用系统模块)，等等。SAS 有一个智能型绘图系统，能绘制各种统计图。SAS 提供了许多统计过程，每个过程均含有丰富的任选项。用户还可以通过对数据集的一连串加工，实现更为复杂的统计分析。此外，SAS 还提供了各类概率分析函数、分位数函数、样本统计函数和随机数生成函数，使用户能方便地实现特殊统计目的。

自 1993 年由董大钧教授主编的全国第一本高校 SAS 统计软件应用教材出版以来，SAS 软件被迅速推广，它大大地提高了用户的计算机应用水平、信息处理的工作效率和工作质量。

本书为高等院校计算机数据处理、统计分析课程的教科书，也可以作为 SAS 软件应用的指导书。本书介绍了 SAS 9.2 中文版的 Base SAS 和 SAS/STAT 的主要内容；通过大量实例详细阐述如何使用 SAS 语言编写程序，以及 SAS 统计软件包中许多基本统计过程的使用方法和结果的分析。本书的例题和分析数据均可通过华信教育资源网(www.hxedu.com.cn)下载。

本书第 1 章和第 2 章由董丽编写，第 3 章由申笑颜编写，第 4 章、第 5 章由李茜编写。其余各章由董大钧编写。

由于我们知识和经验有限，加之时间仓促，本书难免存在各种问题，望读者提出宝贵的意见和建议。

编 者

目 录

绪论	(1)	习题	(31)
第 1 章 SAS 系统概述	(3)	第 2 章 建立数据集	(33)
1.1 SAS 窗口系统	(3)	2.1 SAS 数据集	(33)
1.1.1 SAS 系统的启动和退出	(3)	2.2 利用工具建立数据集	(34)
1.1.2 SAS 窗口环境	(4)	2.2.1 用 Viewtable 建立和编辑数据集	(34)
1.1.3 程序编辑器(Editor)窗口	(6)	2.2.2 使用分析数据表	(36)
1.1.4 日志(LOG)窗口	(8)	2.3 数据步建立数据集的基本语句	(38)
1.1.5 输出(OUTPUT)窗口	(8)	2.3.1 DATA 语句	(38)
1.1.6 结果(Results)窗口	(9)	2.3.2 INPUT 语句	(40)
1.1.7 改变窗口显示	(9)	2.3.3 INFILE 语句	(48)
1.1.8 打印和保存 SAS 活动窗口中的 内容	(9)	2.3.4 DATALINES 和 DATALINES4 语句	(59)
1.1.9 KEYS 窗口	(10)	2.4 数据步流程	(61)
1.2 SAS 数据集	(11)	2.5 提供信息的几个语句	(61)
1.2.1 SAS 数据集	(11)	2.5.1 LENGTH 语句	(61)
1.2.2 数据集名	(11)	2.5.2 LABEL 语句	(62)
1.2.3 SAS 逻辑库概述	(11)	2.5.3 LOSTCARD 语句	(62)
1.2.4 创建逻辑库	(12)	2.5.4 INFORMAT 语句与 FORMAT 语句及 SAS 输入格式与输出格式	(64)
1.2.5 用 LIBNAME 语句指定逻辑库 引用名	(13)	2.5.5 MISSING 语句	(70)
1.2.6 创建逻辑库成员	(13)	2.5.6 注释语句	(71)
1.2.7 删除逻辑库	(14)	2.5.7 OPTIONS 语句	(71)
1.3 SAS 程序的输入与运行	(14)	2.6 建立数据集的另外几种途径	(72)
1.3.1 SAS 程序结构	(14)	2.6.1 从外部文本数据文件中读入 数据	(72)
1.3.2 SAS 程序的输入、修改和运行	(15)	2.6.2 从已建立的数据集中读入数据	(72)
1.3.3 中断 SAS 会话	(20)	2.6.3 从数据库文件中导入数据	(73)
1.3.4 终止 SAS 过程	(20)	习题	(79)
1.3.5 使用 X 语句或 X 指令运行 Windows 指令	(20)	第 3 章 整理数据集	(83)
1.4 SAS 编程基础	(20)	3.1 数据集集中的数据格式	(83)
1.4.1 观测	(20)	3.2 建立、删除与保留变量	(84)
1.4.2 变量	(20)	3.2.1 赋值语句建立变量	(84)
1.4.3 常量	(22)	3.2.2 求和语句	(85)
1.4.4 函数	(24)	3.2.3 RETAIN 语句	(86)
1.4.5 操作符	(29)	3.2.4 DROP 语句删除变量	(87)
1.4.6 SAS 表达式	(30)	3.2.5 KEEP 语句保留变量	(87)

3.3 条件语句	(88)	3.17.4 拆分数据集	(134)
3.3.1 条件 IF 语句	(88)	3.18 DATASETS 过程	(135)
3.3.2 求子集 IF 语句	(93)	3.18.1 PROC DATASETS 过程语法	(135)
3.4 SELECT 语句	(94)	3.18.2 PROC DATASETS 语句	(136)
3.4.1 根据读入的值进行不同的计算	(94)	3.18.3 APPEND 语句	(136)
3.4.2 SELECT 语句中使用 DO 组	(95)	3.18.4 CHANGE 语句	(137)
3.4.3 SELECT 判断用复合表达式	(96)	3.18.5 COPY 语句	(137)
3.5 WHERE 语句	(97)	3.18.6 SELECT 语句	(137)
3.6 删除观测	(98)	3.18.7 CONTENTS 语句	(138)
3.6.1 用 DELETE 语句删除观测	(98)	3.18.8 MODIFY 语句	(139)
3.6.2 删除 SAS 数据集中的重复值	(99)	3.18.9 FORMAT 语句	(139)
3.7 GOTO 语句	(100)	3.18.10 INFORMAT 语句	(140)
3.8 RETURN 语句	(101)	3.18.11 LABEL 语句	(140)
3.9 LINK 语句	(102)	3.18.12 RENAME 语句	(140)
3.10 循环	(103)	习题	(141)
3.10.1 利用 GOTO 实现循环	(103)	第 4 章 SAS 过程步中的常用语句	(148)
3.10.2 DO/END 循环	(103)	4.1 PROC 语句	(148)
3.10.3 DO WHILE 循环	(104)	4.2 BY 语句	(148)
3.10.4 DO UNTIL 循环	(105)	4.3 CLASS 语句	(149)
3.11 数组	(106)	4.4 VAR 语句	(149)
3.11.1 数组与下标变量	(106)	4.5 WEIGHT 语句	(150)
3.11.2 在 DO 循环中使用数组	(109)	4.6 FREQ 语句	(150)
3.11.3 多维数组	(110)	4.7 ID 语句	(151)
3.12 数据排序	(111)	4.8 OUTPUT 语句	(152)
3.12.1 PROC SORT 过程	(111)	4.9 TITLE 语句	(153)
3.12.2 分组控制	(113)	习题	(153)
3.13 数据集连接	(114)	第 5 章 SAS 实用过程	(154)
3.13.1 SET 语句	(114)	5.1 PRINT 过程	(154)
3.13.2 利用 MERGE 语句进行数据集的 横向合并	(117)	5.1.1 语句说明	(154)
3.14 输出信息	(120)	5.1.2 应用举例	(155)
3.14.1 PUT 语句格式	(121)	5.2 FORMAT 过程	(157)
3.14.2 输出形式	(121)	5.2.1 概述	(157)
3.15 FILENAME 语句	(124)	5.2.2 语法说明	(158)
3.16 FILE 语句	(127)	5.2.3 举例	(161)
3.17 OUTPUT 语句	(131)	5.3 TRANSPOSE 过程	(166)
3.17.1 在一个数据步中创建一个以上的 数据集	(131)	5.3.1 概述	(166)
3.17.2 从一个输入行中创建几个 观测	(132)	5.3.2 语句说明	(167)
3.17.3 从几个记录合并信息	(133)	5.4 查询数据	(172)
		5.4.1 概述	(172)
		5.4.2 PROC SQL 语句	(174)

5.4.3	CREATE TABLE 语句	(174)	7.3.3	列联表卡方检验	(213)
5.4.4	用 INSERT INTO 语句向表中 插入数据	(175)	7.3.4	举例	(213)
5.4.5	用 SELECT 语句进行查询	(176)	7.4	TABULATE 过程	(215)
5.4.6	CREATE VIEW 语句	(178)	7.4.1	简介	(215)
5.4.7	DELETE 语句	(178)	7.4.2	用菜单制作汇总表	(217)
5.4.8	DROP 语句	(178)	7.4.3	语句说明	(219)
5.4.9	举例	(178)	7.4.4	举例	(221)
5.5	PROC PRINTTO 过程	(180)	7.5	UNIVARIATE 过程	(224)
	习题	(182)	7.5.1	简介	(224)
第 6 章	ODS 输出交付系统	(185)	7.5.2	语句说明	(224)
6.1	ODS 概述	(185)	7.5.3	使用输出交付系统 ODS	(226)
6.1.1	ODS LISTING 语句	(186)	7.5.4	应用举例	(227)
6.1.2	ODS OUTPUT 语句	(187)	7.6	PLOT 过程	(231)
6.1.3	ODS TRACE 语句	(188)	7.6.1	简介	(231)
6.1.4	选择或排除部分输出	(189)	7.6.2	语句说明	(232)
6.1.5	ODS HTML 语句	(190)	7.6.3	举例	(233)
6.1.6	ODS PDF 语句	(193)	7.7	CHART 过程	(235)
6.1.7	ODS RTF 语句	(194)	7.7.1	简介	(235)
6.1.8	ODS PRINTER 语句	(195)	7.7.2	语句说明	(236)
6.1.9	ODS _all_CLOSE 语句	(195)	7.7.3	选项说明	(238)
6.2	PROC DOCUMENT 过程	(195)	7.7.4	举例	(240)
6.3	TEMPLATE 过程简介	(198)	7.8	RANK 过程	(242)
	习题	(199)	7.8.1	简介	(242)
第 7 章	描述性统计过程	(201)	7.8.2	语句说明	(242)
7.1	统计概念简介	(201)	7.8.3	应用	(243)
7.1.1	总体和样本	(201)	7.8.4	举例	(244)
7.1.2	定位测量	(201)		习题	(246)
7.1.3	离散程度测量	(201)	第 8 章	假设检验	(248)
7.1.4	形状测量	(202)	8.1	假设检验简介	(248)
7.1.5	正态分布	(203)	8.1.1	两样本均数差别的统计意义 检验	(248)
7.1.6	样本均值的分布	(203)	8.1.2	t 检验及类型	(248)
7.1.7	置信度和置信区间	(203)	8.1.3	关于非正态分布资料均数差别的 检验	(249)
7.1.8	统计关键字	(203)	8.1.4	双侧检验和单侧检验	(249)
7.1.9	单变量统计对数据的要求	(204)	8.2	用 MEANS 过程进行 t 检验	(250)
7.2	MEANS 过程	(205)	8.2.1	已知均值比较的 t 检验	(250)
7.2.1	语句说明	(205)	8.2.2	配对比较的 t 检验	(250)
7.2.2	举例	(206)	8.3	用 TTEST 过程进行 t 检验	(251)
7.3	FREQ 过程	(208)	8.3.1	TTEST 过程格式	(251)
7.3.1	简介	(208)	8.3.2	单样本 t 检验	(253)
7.3.2	语句说明	(211)			

8.3.3 成组实验的 t 检验	(255)	10.3.1 CANCELL 过程介绍	(295)
8.3.4 配对比较 t 检验	(256)	10.3.2 应用举例	(296)
8.4 NPAR1WAY 过程	(258)	10.4 REG 过程	(302)
8.4.1 语句说明	(258)	10.4.1 REG 过程使用的语句	(302)
8.4.2 举例	(259)	10.4.2 举例	(304)
习题	(261)	10.5 STEPWISE 过程	(308)
第9章 方差分析	(263)	10.5.1 STEPWISE 过程简介	(308)
9.1 概述	(263)	10.5.2 STEPWISE 过程使用的语句	(308)
9.2 均衡数据的方差分析 (ANOVA 过程)	(264)	10.5.3 举例	(309)
9.2.1 ANOVA 过程的主要语句	(264)	10.6 非线性回归 NLIN 过程	(311)
9.2.2 语句说明	(265)	10.6.1 非线性回归概述	(311)
9.2.3 输出结果	(266)	10.6.2 过程说明	(312)
9.3 ANOVA 过程应用实例	(267)	10.6.3 可能出现的问题	(313)
9.3.1 完全随机设计资料的方差分析 (单因素方差分析)	(267)	10.6.4 输出内容	(314)
9.3.2 随机区组设计资料的方差分析 (两因素方差分析)	(270)	10.6.5 举例	(314)
9.3.3 拉丁方资料的方差分析	(272)	10.7 LOGISTIC 过程	(316)
9.3.4 正交实验设计的方差分析	(274)	10.7.1 Logistic 回归简介	(316)
9.3.5 析因实验设计——三因素实验 分析	(275)	10.7.2 模型的选择	(317)
9.3.6 R * C 析因设计资料的方差分析 (有交互因素的方差分析)	(277)	10.7.3 LOGISTIC 过程使用的语句	(318)
9.4 GLM 过程	(278)	10.7.4 使用注意	(320)
9.4.1 概述	(278)	10.7.5 LOGISTIC 过程的应用	(321)
9.4.2 过程使用的语句	(278)	习题	(326)
9.4.3 四种类型的可估计函数	(280)	第11章 判别分析	(328)
9.4.4 举例	(280)	11.1 判别分析概述	(328)
9.5 协方差分析	(283)	11.2 DISCRIM 过程	(329)
9.5.1 协方差分析简介	(283)	11.2.1 概述	(329)
9.5.2 举例	(283)	11.2.2 DISCRIM 过程使用的语句	(330)
习题	(287)	11.2.3 DISCRIM 过程的输出	(332)
第10章 相关与回归	(289)	11.2.4 应用举例	(332)
10.1 概述	(289)	11.3 STEPDISC 过程	(340)
10.1.1 相关分析	(289)	11.3.1 概述	(340)
10.1.2 回归分析	(290)	11.3.2 STEPDISC 过程使用的语句	(341)
10.2 CORR 过程	(290)	11.3.3 STEPDISC 过程应用举例	(342)
10.2.1 语句说明	(290)	习题	(345)
10.2.2 举例	(291)	第12章 聚类分析	(348)
10.3 CANCELL 过程	(295)	12.1 概述	(348)
		12.1.1 聚类方法	(348)
		12.1.2 SAS 提供的聚类过程	(349)
		12.1.3 聚类过程的输出及聚类类数的 确定	(349)
		12.1.4 散点图和树形图	(350)

12.2 CLUSTER 过程	(350)	14.1.2 生存分析的过程	(401)
12.2.1 简介	(350)	14.1.3 生存分析中使用的术语	(401)
12.2.2 语句说明	(350)	14.2 LIFEREG 过程	(401)
12.2.3 举例	(352)	14.2.1 简介	(401)
12.3 FASTCLUS 过程	(356)	14.2.2 LIFEREG 过程使用的语句	(402)
12.3.1 简介	(356)	14.2.3 补充说明	(403)
12.3.2 过程语句说明	(356)	14.2.4 生存回归应用举例	(403)
12.3.3 过程输出数据集说明	(357)	14.3 LIFETEST 过程	(408)
12.3.4 举例	(358)	14.3.1 简介	(408)
12.4 VARCLUS 过程	(362)	14.3.2 LIFETEST 过程用的语句	(409)
12.4.1 简介	(362)	14.3.3 举例	(410)
12.4.2 VARCLUS 过程使用的语句	(362)	14.3.4 单变量 Wilcoxon 和 logrank	
12.4.3 举例	(364)	检验	(413)
12.4.4 输出结果的进一步说明	(367)	14.3.5 多变量 Wilcoxon 和 logrank	
12.5 TREE 过程	(370)	检验	(417)
12.5.1 简介	(370)	14.4 PHREG 过程	(420)
12.5.2 过程使用的语句	(370)	14.4.1 PHREG 过程概述	(420)
12.5.3 补充说明	(371)	14.4.2 PHREG 过程使用的语句	(421)
12.5.4 举例	(371)	14.4.3 实例	(423)
习题	(374)	习题	(424)
第 13 章 因子分析	(377)	第 15 章 分类资料分析	(427)
13.1 因子分析的简单原理	(377)	15.1 概述	(427)
13.2 FACTOR 过程	(378)	15.2 CATMOD 过程的基本原理	(427)
13.2.1 FACTOR 过程所用的语句	(378)	15.2.1 单总体	(427)
13.2.2 输出内容	(380)	15.2.2 多总体	(428)
13.2.3 举例	(382)	15.2.3 参数估计及假设检验	(428)
13.3 PRINCOMP 过程	(389)	15.3 CATMOD 过程语句	(429)
13.3.1 语句格式说明	(390)	15.3.1 CATMOD 过程语句	(429)
13.3.2 举例	(390)	15.3.2 关于设计矩阵	(431)
习题	(393)	15.4 举例	(432)
第 14 章 生存分析	(400)	习题	(442)
14.1 概述	(400)	参考文献	(445)
14.1.1 什么是生存分析	(400)		

绪 论

1. 什么是 SAS

SAS(Statistical Analysis System)是一个用于决策支持的大型集成信息系统,用来管理、分析数据和编写报告的组合软件系统。该软件系统最早的功能限于统计分析,至今,统计分析功能也仍是它的重要组成部分和核心功能。经过多年的发展,SAS已被全世界120多个国家和地区的三万多家机构所采用,直接用户超过三百万人,遍及金融、医药卫生、生产、运输、通信、政府和教育科研等领域。在数据处理和统计分析领域,SAS系统被誉为国际上的标准统计软件系统,堪称统计软件界的巨无霸。对数据处理要求最严格的美国食品和药物管理局FDA,在进行新药报批时,实验数据只有使用SAS软件处理才接受申请。

SAS是唯一一家将领先的数据仓库、分析和传统的商业智能(BI)应用完全集成在一起的供应商,可从海量数据中创造智能。SAS提供了世界级的分析、数据操纵和报表展现能力,它具有可伸缩性、高度的互用性和易管理性。

SAS是一种模块化软件,SAS系统由三十几个模块组成,其分析功能散布在几乎所有的模块之中,较为集中的具有统计分析功能的是BASE SAS、SAS/STAT、SAS/QC、SAS/INSIGHT、SAS/ETS等一些模块,通过编程可以调用各种分析功能。

对于常用的一些统计分析方法,SAS系统中的如下三种方法可以达到同样的目的:

- SAS/INSIGHT(“交互式数据分析”)该方法直观。
- Analyst(“分析家”)在属性数据分析和功效函数计算方面功能强。
- 直接编程功能最强,对一些特殊的分析,只能用编程方法处理。

Base SAS软件是SAS提供的各种数据管理和分析软件组件的核心基础,它是访问、管理、分析和显示数据的世界技术领先的信息传输系统。Base SAS软件包括第四代编程语言和称为过程的现成程序。这些过程可处理数据、信息存储和检索、统计分析和报表编写等。

SAS/STAT可用于高级统计分析和应用,包括方差分析、回归分析、属性数据分析、多元分析、生存分析、聚类分析、判断分析及非参数分析等。

SAS/ACCESS用于从数据库管理系统(DBMS),以及PC文件格式读取和写入数据。

SAS/GRAPH以各种彩色图表展示信息,完成各种用于业务管理和统计分析的中、外文图形报表,包括直方图、饼形图、曲线图、三维图等。

SAS/ASSIST提供任务导向式、菜单驱动用户友好界面,利用图形式的选项画面,自动生成完整的SAS程序。

SAS/OR可做线性规划、项目管理、资源分配和决策支持。

SAS/IML交互式矩阵语言,可利用其内置的功能进行科学计算和统计分析。

SAS/QC提供用于质量管理的各种专门工具,可做实验设计、质量管理和过程控制,并提供完整的实验设计和质量管理的菜单系统。

SAS/ETS建立计量经济及财务模型,分析和预测各种时间序列资料。

SAS/INSIGHT为交互式统计分析及动态图形展示工具,进行交互式动态资料分析和展示。

SAS/EM智能型的数据挖掘集成工具,SAS/EM的图形化界面、可视化操作引导用户按SEMMA

原则成功地进行数据挖掘。SAS/EM 可实现与数据仓库和数据库、商务智能及报表工具的无缝集成。

2. 利用 SAS 系统处理数据的过程

SAS 系统提供了一种完善的编程语言。属于第四代计算机高级语言,在 SAS 中,把大部分常用的复杂数据计算的算法作为标准过程,SAS 用户只需要熟悉其命令、语句及简单的语法规则就可以实现对数据的管理和分析处理工作。这使得 SAS 编程十分简单。SAS 系统也提供了了图形化的软件模块,使数据集的建立像使用电子表格软件那样方便,处理数据只需要回答对话框中的问题提交即可,使对数据处理更简便。然而在预处理数据时,就不像使用 SAS 语言编程那样灵活。因此,掌握 SAS 编程技术是使用 SAS 的关键环节。

SAS 编程处理数据分为两个步骤,数据步和过程步。在数据步建立起数据集、这是对数据能否处理成功的关键。

可以用各种数据步的语句对数据进行预处理,整理数据,以使数据符合数据处理过程的需要,整理好的数据放在数据集中。在过程步,利用各种过程对指定的数据集中的数据进行处理、分析,并输出结果

数据步(DATA Step)以 DATA 语句开始,以 RUN 语句结束。

过程步(PROC Step)以 PROC 语句开始,以 RUN 语句结束。也有些过程步要用 QUIT 语句结束。

3. 随书附带光盘的使用

本书中所有例题、数据及某些参考资料保存在华信教育资源网 www.hxedu.com.cn 上,在使用本书时请将其中的 SASDT、SASPRG、SASLX、SASLX1、SASLX2 和 SASEX 等文件夹复制到你的计算机 E:\ 下。

由于有极个别的例子中所用的命令是 SAS 9.2 后新增加的,故在 SAS 早期版本上运行时,会出现命令或函数找不到的错误。

第 1 章 SAS 系统概述

1.1 SAS 窗口系统

1.1.1 SAS 系统的启动和退出

1. 启动

SAS 软件有租期限制，只有机器时间在软件有效期内才能启动 SAS。

若计算机安装有 SAS 系统，打开“开始”菜单，选择 SAS 9.2，或“程序”→“The SAS System”→“The SAS System for Windows 9.2”，即进入 SAS 系统，打开“SAS”窗口，如图 1-1 所示。

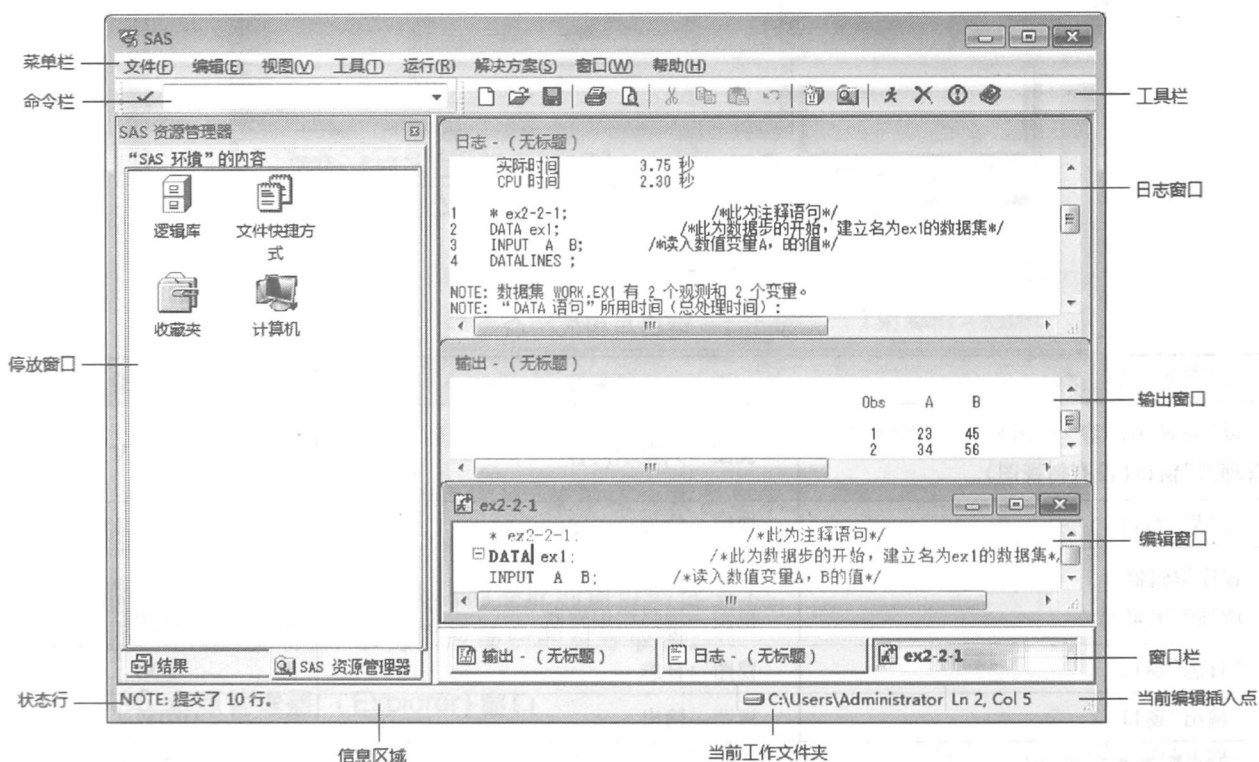


图 1-1 “SAS”窗口

2. 退出 SAS 系统

在工作结束之后要退出 SAS 系统，可使用下列方法之一：

- (1) 单击“SAS”窗口的“关闭”按钮；
- (2) 在窗口的命令栏上输入“BYE”后按回车键；
- (3) 在程序编辑器的语句行上输入“ENDSAS;”按回车或工具栏上的提交(SUBMIT)键提交它。

1.1.2 SAS 窗口环境

当启动 SAS 时,将打开 5 个主窗口:“SAS 资源管理器”窗口、“结果”窗口、“程序编辑器”窗口、“日志”窗口和“输出”窗口。另外还有 30 多个窗口可供使用。

1. 打开窗口

要打开主窗口,可以从“视图”菜单选择相应的命令项或在命令栏中使用相应的命令,如图 1-2 和表 1-1 所示。

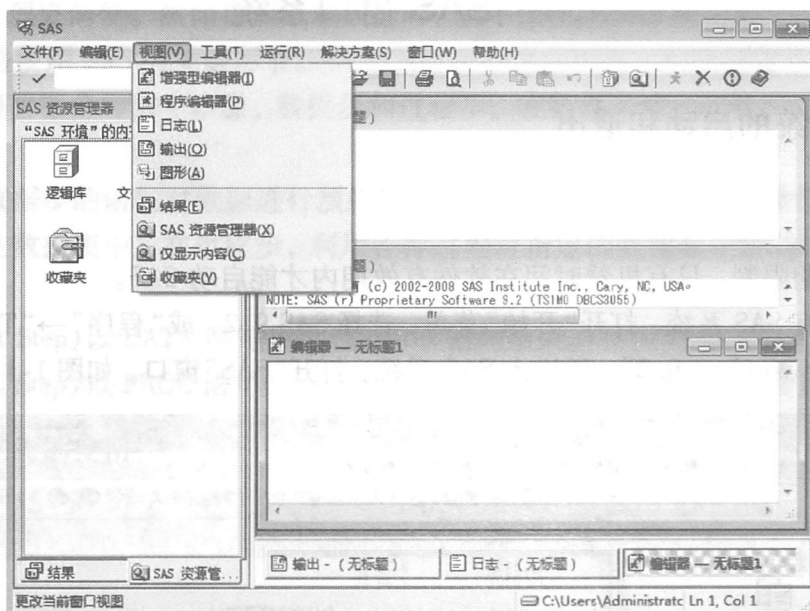


图 1-2 “视图”菜单

表 1-1 窗口与相应的命令

窗 口	菜 单 路 径	命 令
以“只显示内容”视图显示的“SAS 资源管理器”窗口(没有树视图)	视图→只显示内容	SASENV
“结果”窗口	视图→结果	ODSRESULTS
程序编辑器 增强型编辑器	视图→程序编辑器 视图→增强型编辑器	PGM WEDIT
“日志”窗口	视图→日志	LOG
“输出”窗口	视图→输出	OUTPUT
“SAS 资源管理器”窗口	视图→SAS 资源管理器	EXPLORER

SAS 命令无须完整输入,只要前面的字符能与其他命令区分开即可。在命令栏中可以输入多个窗口命令,命令之间须用分号分隔。例如,以下命令可打开“脚注”和“查找”窗口:

```
footnote;expfind;
```

单击窗口切换按钮,或在“窗口”菜单中选中某窗口,使该窗口成为当前窗口。当前窗口标题栏颜色深并有光标闪烁。SAS 只对当前窗口进行操作。

2. “停放”视图

工作区的左侧区域为停放视图区。其他任何打开的窗口将不会遮挡停放视图。默认情况下,“SAS 资源管理器”窗口和“结果”窗口为停放窗口。每个停放窗口的停放区域底部都有一个选项卡,单击选项卡,即可在停放的窗口间切换。当停放窗口数量较多时,下部将出现向左和向右箭头,可用以浏览各停放窗口。

(1) 停放窗口

- ① 选择窗口以使其成为当前窗口。
- ② 选择“窗口”→“停放”,如图 1-3 所示。



图 1-3 窗口菜单中的“停放”项

注意:如果某窗口不能停放,则“停放”菜单项为灰色显示。

(2) 取消停放窗口

- ① 在停放区域的底部,单击要取消停放的窗口的选项卡。
- ② 选择“窗口”→“停放”,将停放项前的对号取消。

3. “SAS 资源管理器”(Explorer)窗口

“SAS 资源管理器”窗口的作用类似于 Windows 操作系统下的资源管理器,是管理、访问数据的中心。这些数据包括目录、表(数据集)、逻辑库和主机文件数据,也可用于执行基本的 SAS 任务。

“SAS 资源管理器”的使用方法与 Windows 资源管理器的使用方法几乎相同。在命令栏使用“视图”,“SAS 资源管理器”命令,可以打开“SAS 资源管理器”。

“SAS 资源管理器”有两个窗格。左窗格为“SAS 环境视图”(也称为树视图)显示 SAS 的逻辑库及 Windows 资源管理器的大部分内容。右窗格显示在左窗格中所选中的项目的内容,如果在“视图”菜单中单击“仅显示内容”,则打开的是“SAS 环境”的内容视图。

利用视图菜单，可以一次打开多个“SAS 资源管理器”窗口。注意，当前窗口不同，视图菜单也不同，如图 1-4 所示。



图 1-4 SAS 资源管理器窗口和查看菜单

可以选择“视图”→“显示树状结构”，或输入 TREE 命令，切换显示“SAS 资源管理器”视图和“只显示内容”视图。

除使用 SAS 管理器(SAS Explorer)窗口直接管理 SAS 数据库及其文件外，也可通过其他窗口来实现对数据库的管理：

用“LIBNAME”命令进入 Libname 窗口；

用“DIR”命令进入 Dir 窗口；

用“VAR”命令进入 Var 窗口（默认为显示临时库中数据集的属性）。

1.1.3 程序编辑器(Editor)窗口

程序编辑器窗口是 SAS 的文字编辑器，程序编辑器使用方法像 Word 等文字编辑软件一样。在此窗口可以编辑任何文本文件和程序，可用“文件”(FILE)菜单的“保存”(Save)命令存储此窗口中的内容。编辑好的程序可利用工具栏的提交按钮或在命令窗口用“SUBMIT”命令提交给 SAS 系统去执行。

1. 窗口分类

程序编辑器分为两种。一种仅用于 Windows 操作环境，能利用彩色分出程序的各部分，可展开并折叠程序段，自动缩排使程序清晰易懂；用虚线隔开程序的各步，可记录宏，称为增强型程序“编辑器”，如图 1-5 所示。

在命令栏使用“WEDIT”命令可打开“增强型程序编辑器”窗口。

如果对“编辑器”窗口中的内容进行了修改，标题中将增加一个星号。

除了增强型编辑器外，SAS 还提供了可用于所有 SAS 平台的普通的“程序编辑器”(program editor)窗口，它拥有与其他文本编辑器类似的文本编辑功能，可以作为读入、编辑和提交 SAS 程序的窗口。在命令栏使用“PGM”命令可打开普通程序编辑器窗口，如图 1-6 所示。

在这两种程序编辑器窗口都能输入和编辑 SAS 程序，并能将程序提交给 SAS 系统。

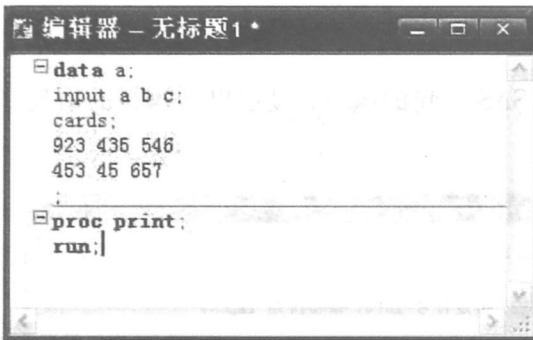


图 1-5 增强型程序“编辑器”

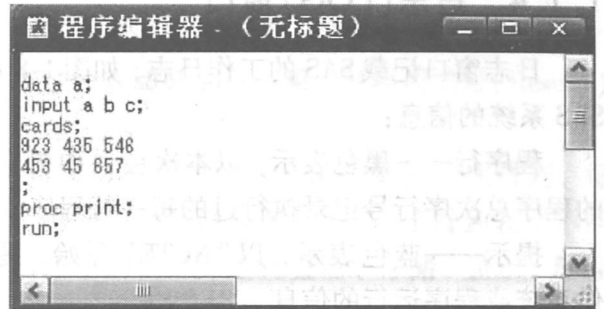


图 1-6 程序编辑器

2. 窗口内容的存取

1) 调入外部文件

利用“文件”菜单的“打开”命令，找到 SAS 程序文件，即可将其调入程序编辑器窗口，或在程序编辑器窗口的命令栏上，利用“INCLUDE”命令将文本文件的内容调入该窗口。例如，将存于 E 盘 SASPG 文件夹的 ex2-1-1. sas 文件调入程序编辑器窗口，在命令栏上使用 INCLUDE 'E:\SASPG\ex2-1-1. sas'，即可将该文件调入。

2) 保存当前窗口内容

这两种程序编辑器窗口都能输入和编辑 SAS 程序。保存程序时，选择“文件”→“保存”，指定路径，给出文件名确定即可。程序文件后缀为 . sas。

3) 当前窗口内容的删除

如想将当前窗口中的内容全部删除，可以按 Ctrl + E 组合键或在命令栏上使用“CLEAR”命令或在工具栏上单击删除按钮。

3. 窗口设置

要显示程序的行号或在提交程序后清除提交的程序，在程序编辑器为当前窗口时，选择“工具”→“选项”→“程序编辑器”，在弹出的“程序编辑器选项”对话框中进行设置，如图 1-7 所示。

也可选择“工具”→“选项”→“参数选择”，在弹出的“参数选择”对话框中进行设置。

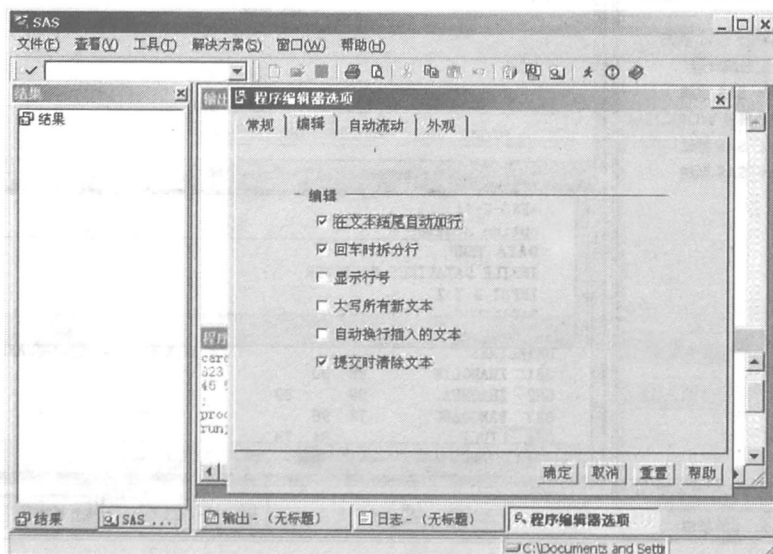


图 1-7 “程序编辑器选项”对话框

1.1.4 日志(LOG)窗口

日志窗口记载 SAS 的工作日志,如图 1-8 所示。随着 SAS 语句的执行,显示出已执行的语句和 SAS 系统的信息:

程序行——黑色表示,以本次会话中提交的程序总次序行号记录执行过的每一行程序。

提示——蓝色表示,以“NOTE”开始,提供系统或程序运行的信息。

警告——绿色表示,以“WARNING”开始,系统能自动纠正程序中的小错误,如将 INPUT 输入成 IMPUT 会提示错误语句,并自动纠正错误。当内存不足或软件即将到期也会给出警告信息。

错误——红色显示,以“ERROR”开始,表示该程序存在错误,即使程序显示出结果,结果必是错误的。应仔细查看此部分信息,找出错误原因,要找出最早的错误所在,因为该错误会影响后面的程序行的执行。

日志窗口中的内容在保存时被存为扩展名为“.log”的纯文本文件。

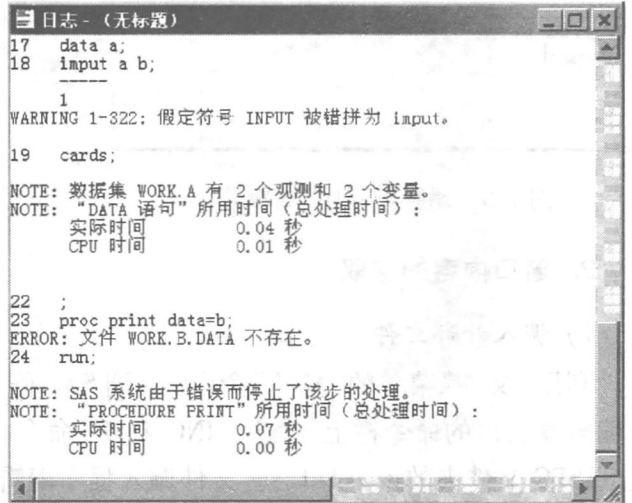


图 1-8 LOG 窗口

1.1.5 输出(OUTPUT)窗口

显示由 SAS 过程所输出的结果。在查看日志窗口确认程序运行无误后,该窗口显示的内容才是程序运行的结果。输出窗口中的内容在保存时被存为扩展名为“.lst”的纯文本文件。使用“output”命令或按 F6 键可打开输出窗口,如图 1-9 所示。

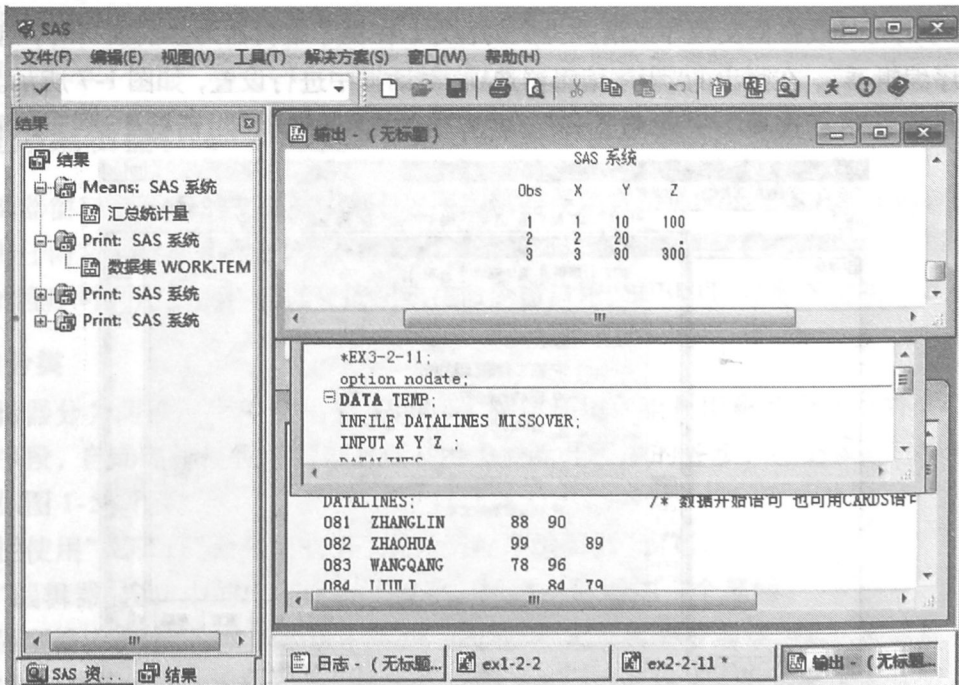


图 1-9 结果与输出窗口