

本书附有磁片

SAS/PC

学习和使用微机统计分析系统



张绍勋
林秀娟 编著

学苑出版社

计算机实用技术系列丛书

计算机实用技术系列丛书

学习和使用微机 统计分析系统 SAS/PC

张绍勋 林秀娟 著

李 江 改编

亦 鸥 审

学苑出版社

1993

(京)新登字 151 号

内 容 简 介

本书全面、详细地介绍应用统计分析系统 SAS 的用法。全书共分为十三章和两个附录,内容包括描述统计、推论统计、实验设计、多变量统计等。书中给出了大量的例子,并在一些常用的程序之后给出了练习题。是一本学习和使用 SAS 入门书籍,也可以作为 SAS 的教材使用。本书可供计算机专业的科技人员、应用系统开发人员与用户学习、参考。

需要本书的用户,可与北京 8721 信箱联系,邮码 100080,电话 2562329。

版 权 声 明

本书繁体字中文版原名为《SAS/PC 统计分析与实务应用》,由松岗电脑图书资料股份有限公司出版。版权归松岗公司所有。本书简体字中文版权由松岗公司授予北京希望电脑公司,由北京希望电脑公司和学苑出版社独家出版、发行。未经出版者书面许可,本书的任何部分不得以任何形式或任何手段复制或传播。

计算机实用技术系列丛书

学习和使用微机统计分析系统 SAS/PC

编 著:张绍勋 林秀娟

改 编:李 汇

审 校:亦 鸥

责任编辑:徐建军

出版发行:学苑出版社 邮政编码:100032

社 址:北京市西城区成方街 33 号

排 版:北京天奥科技公司照排中心

印 刷:双青印刷厂

开 本:787×1092 1/16

印 张:46.5 字数:964 千字

印 数:1—5000 册

版 次:1993 年 12 月北京第 1 版第 1 次

ISBN 7-5077-0806-3/TP·17

本册定价:57.00 元(含盘)

学苑版图书印、装错误可随时退换

欢迎加入希望用户协会

当今的计算机技术发展迅猛,应用领域繁多,如何准确、有效、地为您提供服务,已成为我们迫切关心的问题。为此,我们决定创立希望用户协会,其宗旨在于加强与用户的联络,了解用户的各种信息与需求,为用户提供更完善更周到的服务。

本协会的成员将定期获得各种软件、资料与培训等讯息,优先得到有关技术服务。请您认真填写后面的会员登记表,其中的信息将用电脑进行管理。

作为美国Borland软件公司在中国的总代理,北京希望电脑公司经Borland公司授权,对其产品dBASE IV 2.0 进行系统级汉化。另外,对Borland C++ & AF 3.1、Turbo C++ for DOS 3.0、Turbo C++ for Windows、Turbo C++ Visual Edition for Windows四个产品进行了本地化。预计94年3月推出以上产品。

Borland 软件系列清单:

1. Borland C++ & AF 3.1
2. Borland C++ 3.1
3. Turbo C++ for DOS 3.0
4. Turbo C++ for Windows
5. Paradox for DOS 4.0
6. Paradox for Windows
7. Pdox Engine & Database Framework 3.0
8. Turbo Pascal for DOS 7.0
9. Turbo Pascal for Windows 1.5
10. dBASE IV 2.0
11. dBASE IV Compiler
12. Object Vision 2.1

软件与培训

北京希望电脑公司即将与Borland公司密切合作,经销Borland多种软件产品的同时,提供配套的函授与培训服务,以满足用户的不同要求。欢迎用户索取软件与培训资料。

联系方法

有关资料事宜,请与朱红小姐联系;有关软件事宜,请与周东先生联系;有关函授与培训事宜,请与宋明华先生联系。联系方法如下:

通信地址: 北京 8721 信箱 邮政编码: 100080

联系电话: (01)2562329 (01)2541992 (01)2579874

传真电话: (01)2561057

目 录

第零章 SAS 软件包	1
0.1 SAS 简介	1
0.2 SAS 系统产品及其应用功能	1
0.3 软件及硬件需求	2
0.4 SAS 系统的安装	2
0.5 启动 SAS 软件包	4
0.6 常用的程序管理系统指令	5
0.7 SAS 的画面及其功能	11
第一章 基本 SAS 数据集建立与数据转换	16
1.1 建立 SAS 数据集	16
1.2 数据集成的建立	16
1.3 数据输入方法及格式	20
1.4 数据集的合并及变量的选取	25
1.5 其他常用指令	29
1.6 中文配合 SAS 系统一起使用的方法	34
1.7 SAS 数据文件的种类	35
1.8 在 SAS 程序中任何位置均可使用的 SAS 指令	52
第二章 SAS 的函数	56
2.1 算术函数与数学函数	56
2.2 三角函数	57
2.3 概率函数	58
2.4 样本统计函数	60
2.5 截取(Truncation)函数	61
2.6 其他函数	61
2.7 SAS 运算符(Operator)与常数	63
第三章 基本数据的报表打印	65
3.1 日报表的制作:PROC CALENDAR	65
3.2 统计表格的制作:PROC CHART	69
3.3 在方格中印一组内容:PROC FORMS	77
3.4 一般二维坐标绘图:PROC PLOT	79
3.5 印出数据集内变量的内容:PROC PRINT	82
3.6 时间序列图的制作:PROC TIMEPLOT	85
3.7 印出变量的属性:PROC CONTENTS	94
3.8 变量观察值的输出格式:PROC FORMAT	96

3.9	数据集内容的转置:PROC TRANSPOSE	99
第四章	描述性统计分析.....	104
4.1	描述性统计分析:PROC MEANS	104
4.2	建立描述性统计数据集:PROC SUMMARY	108
4.3	描述性统计量的计算与绘图:PROC UNIVARIATE	112
4.4	PROC TABULATE	117
4.5	相关强度计算:PROC CORR	129
4.6	类变量的扫描:PROC FREQ	137
第五章	记分分析.....	148
5.1	观察值的排序:PROC BANK	148
5.2	变量值的线性组合:PROC SCORE	153
5.3	变量的标准化:PROC STANDARD	158
第六章	简单回归及复合回归.....	163
6.1	非线性回归分析:PROC NLIN	163
6.2	一般性回归分析:PROC REG	166
6.3	正交化回归分析:PROC ORTHOREG	189
6.4	多项式回归分析:PROC RSREG	192
6.5	逐步回归分析:PROC STEPWISE	204
6.6	用 R^2 (决定系数)决定最佳回归分析:PROC RSQUARE	220
第七章	类数据处理.....	227
7.1	类数据的线性模式:PROC CATMOD	227
7.2	类变量的扫描:PROC FREQ	264
第八章	变异数分析.....	275
8.1	平衡实验设计的变异数分析:PROC ANOVA	275
8.2	不平衡实验设计的变异数分析:PROC GLM	323
8.3	嵌套(层次)式变异数分析:PROC NESTED	378
8.4	无分母的单因子变异数分析:PROC NPAR1WAY	381
8.5	变异数分析的实验设计:PROC PLAN	386
8.6	两组平均数的比较:PROC TTEST	401
8.7	变异数成份的分析:PROC VARCOMP	412
第九章	多变量统计分析.....	422
9.1	典型相关分析:PROC CANCORR	422
9.2	因素分析:PROC FACTOR	440
9.3	主成份分析:PROC PRINCOMP	483
第十章	鉴别分析.....	505
10.1	分类鉴别分析:PROC DISCRIM	505
10.2	典型鉴别分析:PROC CANDISC	527
10.3	逐步鉴别分析:PROC STEPDISC	569
10.4	近邻鉴别分析:PROC NEIGHBOR	588

第十一章 集群分析	593
11.1 层次式集群分析:PROC CLUSTER	594
11.2 相斥式集群分析:PROC FASTCLUS	638
11.3 集群分析的树状图表示:PROC TREE	668
11.4 变量的集群分析:PROC VARCLUS	685
11.5 共轭变量估计值的集群分析法:PROC ACECLUS	702
第十二章 生存统计分析	709
12.1 PROC LIFEREG	709
附录 A 宏指令	733
附录 B 统计分析的应用技术	736

第零章 SAS 软件包

0.1 SAS 简介

SAS 的英文全名是 Statistical Analysis System, 中文译为统计分析系统。1960 年代末期由北卡州立大学(North Carolins Stats University)统计系的 A. J. Barr 与 J. H. Goodnight 两位教授开始开发。开发至今, SAS 在大型计算机上的主要版本已开发到第七版, 个人计算机的版本则到第 6.03 版, 本书介绍将以此最新版本为主。

SAS 之内部系统程序是以汇编语言撰写, 不需要编译。所以虽然它是软件包, 但处理速度并不逊于高级语言。而且经过多次修订后, SAS 除了增强原有的统计分析功能外, 还具备了数据库管理、高级语言程序及宏指令撰写、报告撰写及精密绘图等功能。

当前一般研究生从事研究时最关心和最困惑的问题包括: 一、在回答研究假设时, 应采用什么统计方法? 二、当统计方法选定后, 有关这个方法的知识在哪些教科书或文献上有详细记载? 三、如何操作使用效率愈来愈高的个人计算机? 四、有关 SAS 统计的计算机软件包如何使用? 五、如何理解 SAS 软件包所输出的统计数据?

根据笔者的经验, SAS 是一个一流的统计软件, 本书将针对以上问题, 将 SAS 介绍给研究生们, 供他们在数据处理及论文撰写时参考使用。

0.2 SAS 系统产品及其应用功能

1. SAS/AF

- (1) 专用于应用功能处理。
- (2) 可用以建立简单易懂的画面及选项, 由用户以简单指令控制系统的运行。
- (3) 应用范围覆盖企业管理、教育训练、医院管理、数据检索、物料控制等各方面。

2. SAS/BASICS

- (1) 为 SAS 系统中必备的设备。
- (2) 主要功能为数据管理及数据处理, 并具备报告打印及语句统计性分析的功能。

3. SAS/ETS

- (1) 专用于计量经济与时间序列(Econometric & Time Series)分析。
- (2) 在计量经济分析建立的联立模式, 以探讨联立结构关系, 应用以预测未来趋势。时间序列分析则运用独立或回归关系探讨序列特质并预测未来。
- (3) 应用范围不局限于经济分析, 并可扩展到医学、农业、社会、管理...等各方面。

4. SAS/FSP

- (1) 专用于全屏幕处理(Full Screen Process)。
- (2) 应用于 SAS 文件的数据登录、编辑、存取及撰写书信。可直接修订 SAS 数据组中的

任何数据。

5. SAS/GRAPHIC

(1)专用于精致美观的绘图(Graphic)。

(2)可绘制彩色、立体、地理图等,需配合相关硬件,此相关硬件可依自己需要再扩充。

6. SAS/IML

(1)为一个多元化、交互式矩阵语言(Interactive Matrix Language)。

(2)兼具程序语言与软件包弹性及简单的多阶(Multi-Level)特性,以矩阵形式处理一切数据及运算。

7. SAS/OR

(1)专用于作业研究(Operation Research)的处理,但 SAS/OR 产品开发至今仍尚未十分完备。

(2)备有各种线性规划及数学优化处理功能。

8. SAS/STAT

(1)专用于多变量统计(Statistics)分析,对象为固定特点下所汇集的横断面数据。

(2)具有回归分析、变异数分析、因素与主成分分析、集群、鉴别分析及典型相关分析等功能,可广泛应用于经济、农业、教育、社会心理各方面的研究。

0.3 软件及硬件需求

1. 软件:

(1)可选择 PE2 或 Wordstar 等其他较熟悉的文字处理的软件,来编辑所写的 SAS 程序;或用 SAS 自身的编辑器。

(2)MS-DOS 3.3 及以后版本。

(3)SAS/PC 系统 6.03 及以后版本。

2. 硬件:

(1)PC 286 或 PC 386 兼容的计算机,或大型计算机。

(2)PC 的硬盘需 20M bytes 容量以上,最好是 40M 硬盘。单色或彩色屏幕。

(3)EMS 卡,最好能购买此内存扩充卡,在中文配合运行 SAS 系统时,方便多了。

3. 若是从硬盘启动的个人计算机,则须用文字处理软件(如 PE2 等)建立一个 CONFIG.SYS 文件,存放于 C 磁盘的根目录中(即 C:\)。其内容如下:

```
c:\>type CONFIG.SYS
```

```
files=50
```

```
buffers=20
```

0.4 SAS 系统的安装

SAS/PC 系统安装步骤如下:

(1)个人计算机启动后,在硬盘 C 的根目录下,键入“a:sasload 等字,则屏幕出现图 0.1 画面,此时为了简化安装步骤,可选“A”。

```
C:\>A:ses1080

SAS SYSTEM RELEASE 6.03 INSTALLATION PROCEDURE
=====

NOTE: The Update to SAS/STAT Software, SAS/ASSIST Software and
      Micro-to-Host Link Enhancements can only be installed by
      FAST Install.

Please select the method of installation:

[A] FAST Install -- Install Complete Product.
[E] CUSTOM Install or AUTO Install.
```

图 0.1 SAS 安装步骤一

(2)经步骤一选 A 后,出现图 0.2 画面,此时若想将 SAS 系统安装到 C:\SAS 子目录,则键入“Y”。

```
The installation path is C:\SAS
Is this correct? (Y/N)
```

图 0.2 SAS 安装步骤二

(3)经步骤二选 Y 后,出现图 0.3 画面,此时依对 SAS 系统的需求,选择所需的 SAS 产品。例如想安装 SAS/BASIC,则键入“A”,直到将 BASE 的 5 片软盘全部安装完毕后,屏幕画面重回到图 0.3 画面。若又要继续安装 SAS/STAT 软件,则选“C”,直到 STAT 4 片软盘安装完毕后,画面又回到图 0.3。此时若想停止安装,则键入“Q”。

```
-----
Please select the SAS Product to install:

[A] Base SAS Software                [B] Base SAS Software with EMS Support
[C] SAS/STAT Software                [D] SAS/IML Software
[E] SAS/GRAPH Software               [F] SAS/GRAPH Software with MAPS
[G] SAS/FSP Software                 [H] SAS/AF Software
[I] SAS/ASSIST Software with Source  [J] SAS/ASSIST Software without Source
[K] Micro-to-Host Link Enhancements [L] Update to SAS/STAT Software
[M] Update to SAS/STAT Software without Menu System
[N] SAS/OR Software                  [O] SAS/ETS Software

Warning: Installing Update to SAS/STAT Software without menus will
         result in deleting the existing SAS/STAT Menu System.

[Q] Quit and return to DOS
```

图 0.3 SAS 安装步骤三

(4)将 PASSWORD 软盘放入 A 磁盘,键入“CD C:\SAS\SASINST”后,再键入

“SASBACK -RESTORE A; C; CORE.SET-SDIR”。

经过上列的几个安装步骤后,SAS 此时已存放在硬盘 C 上。若运行 pctools.exe,则会在屏幕上看到如图 0.4 所示的 SAS 文件结构:

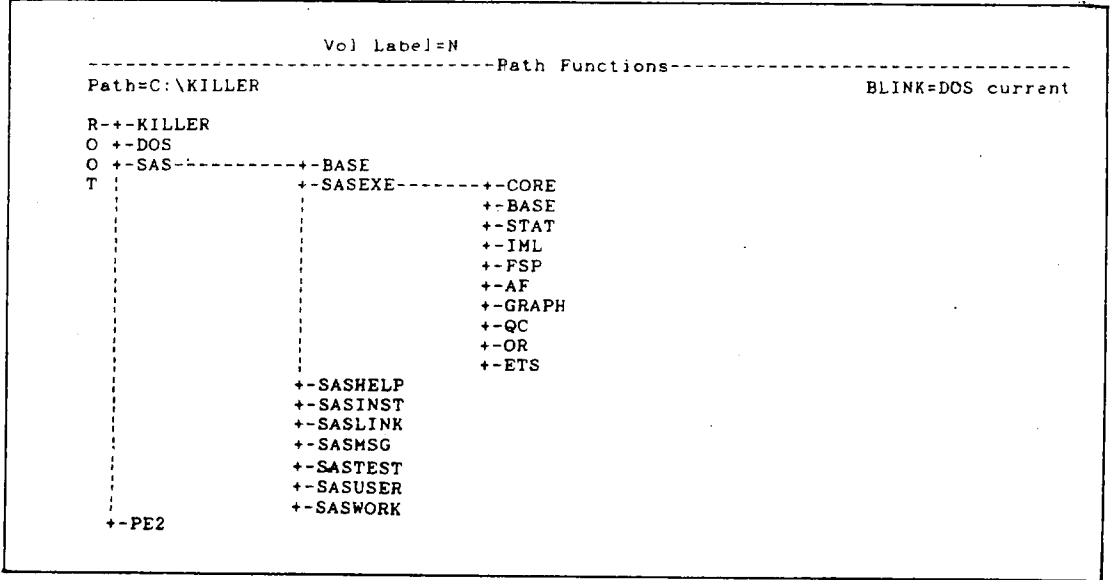


图 0.4 SAS 的树状结构

若出现以上树状结构,表示已成功地安装完毕。本书的介绍重点将以 BASICS 和 STAT 产品为主。至于 SAS 程序编辑,可选用自己较熟悉编辑软件来写 SAS 程序(如 PE2);或使用 SAS 系统所提供的 PROGRAM EDITOR 窗口来进行编辑,其用法如 0.6 节所示。

0.5 启动 SAS 软件包

在大型计算机启动 SAS 方法,因系统不同有不同的控制及编辑指令,可请教系统维护人员,或查阅相关的使用手册。

在个人计算机启动 SAS 之前,最好能建一个 Autoexec.bat 文件,存在硬盘的根目录下,其内容如下:

```
c:\>type autoexec.bat
path c:\;c:\dos;c:\sas;
prompt $p$g
```

在安装 SAS 系统的硬盘的提示符号(如 C:\>)右方键入 CD\SAS。

```
C:\>CD SAS
```

屏幕将出现“C:\SAS>”提示符号,再于其右方键入 SAS。

```
C:\SAS>SAS
```

若屏幕显现如图 0.5,则表示已成功进入 SAS 系统。

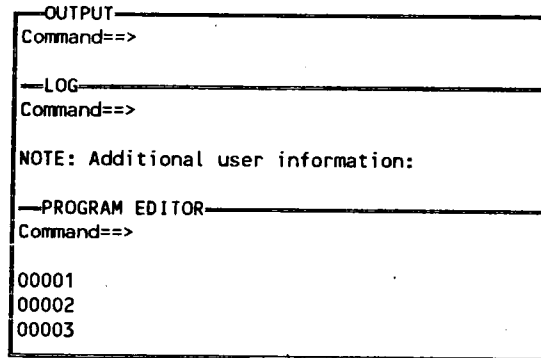


图 0.5 SAS 启动后显示的画面

其中 OUTPUT 窗口显示程序的运行结果;LOG 窗口显示程序的运行过程,包括是否有程序语法错误(ERROR 或 WARNING 信息)等;PROGRAM 窗口提供程序编辑及运行功能。

出现此三个窗口,表示已在 SAS 的 DISPLAY MANAGEMENT SYSTEM(简称 DMS)屏幕管理系统之中。

0.6 常用的程序管理系统指令

常用的 DISPLAY MANAGEMENT SYSTEM(DMS)指令有:

1. 一般 SAS 程序或数据编辑,可以选择自己较熟练的文字处理软件(如 PE2 等)来编辑程序,但也可选用 DMS 系统所提供的编辑窗口,即 PROGRAM EDITOR WINDOW 来编辑,其常用的指令如图 0.6。

(1)程序编辑指令:

其下达命令的位置为 PROGRAM EDITOR 窗口左侧列号(如 00002)之前二位数字。

命令	功能与作用
In	I 表示程序里某一行之后插入 n 个空行。例如在行号 00002 之后要插入一个空行,按 ↑,↓,←,→ 键,将光标移至 00002 行由左算起第一个 0 的位置,打入 II。
Dn	D (Delete) 表示删除所指定的特定行,例如在行号 00004 由左算起第一个 0 的位置,打入 D2 则删除了行号 00004 及行号 00005 的程序。
Rn	重复相同内容 n 列,例如 R2 则重复相同内容 2 列。
RR	重复某一范围的行数,例如要重复某一范围的行数可输入第一个 RR,及第二个 RR,就显示同样范围的行内容。
Ch	拷贝 (Copy) 指令。复制指定行的数据到 A (置后) 或 B (置前) 指令所指定的相关位置。例如 c2 则拷贝 2 行的数据到指定位置。
CC	拷贝此范围内的所有行,例如要把第 2 行到第 7 行的内容拷贝到第 1 行之前,则在行号 00002 之前 2 个零的位置上打 CC,然后移动光标到行号 00007,在前 2 个零的位置上打上 CC,然后移动光标到行号 00001,在第一个零的位置上输入 B,然后按 RETURN。
Mn	搬移指令。搬移指定的 n 行数据到 A (置后) 或 B (置前) 指令所指定的相对位置。
MM	搬移此范围内的所有行,MM MM 两次,其用法如同 CC 指令。
A	置后指令 (After)。将要拷贝 (copy) 或搬移 (Move) 的数据置于其所指定行的下方。和 B 指令相同,须和拷贝或搬移指令合用方能生效。
B	置前 (Before) 指令,将要拷贝或搬移的数据置于其所指定行的上方。

图 0.6 编辑常用的指令

(2) 功能键的使用如图 0.7:

有关更多的功能键说明,可在进入 SAS 系统后,按 F2 键,即可了解其他功能键的用法, SAS 允许更改此 KEYS 窗口内容。更改后内容,只须按 HOME 键,再键入“SAVE”即可。

功能键	使用结果
F1 (HELP)	提供联机帮助功能,打开帮助 (HELP) 窗口。
F2 (KEYS)	打开功能键 (KEYS) 窗口,用户可查询各键的缺省功能。
F3 (LOG)	显示 LOG 窗口,或使光标跳移窗口。
F4 (OUTPUT)	显示 OUTPUT 窗口,或使光标跳移该窗口。
F5 (NEXT)	显示或使光标跳移下一窗口。
F6 (PGM)	显示 PROGRAM EDITOR 窗口,或使光标跳移至该窗口。
F7 (ZOOM)	使光标所在的窗口放大或缩小。
F9 (RECALL)	调用上次已经运行的程序。
F10 (ZOOM, SUBMIT,)	运行程序并使屏幕回到 OUTPUT, LOG, PRRGRAM, EDITOR 三窗口的状态。
SHF F7 (LEFT)	使数据左移 40 列。
SHF F8 (RIGHT)	使数据右移 40 列。
SHF F9 (BACKWARD 1)	使数据下移一行。
SHF F10 (FORWARD 1)	使数据上移一行。

图 0.7 SAS 各功能键作用

(3)程序运行与管理指令

我们假定已用 PE2 或 SAS 的 PROGRAMEDITOR 窗口编辑一个 SAS 程序文件,名叫 DEMO.SAS。其内容如下:

```
DATA temp;
  Input height weight;
CARDS;
123 25
127 21
120 24
124 20
118 26
PROC PRINT;
PROC MEANS;
PROC PLOT;
  PLOT height * wight;
RUN;
```

INPUT 指令是 SAS 的保留字,允许定义数个变量。此处定义了变量 height 及变量 weight。CARDS 指令也是 SAS 的保留字,表示它之后为所研究的样本原始数据。PROC 是 PROCEDURE 的简写,是 SAS 的保留字,其后接 PRINT, MEANS 及 PLOT 程序,有关它们用法及时机,请看后几章的介绍。

最后将该文件存于 A 磁盘片,运行该 SAS 程序文件的步骤如下:

1. 键入 INC(Include)指令和欲运行的文件名。

INC 指令功能是将外部文件读进 SAS 系统,其用法如图 0.8。

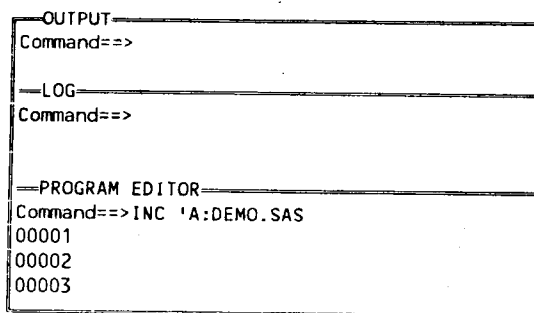


图 0.8

其中 A:表示驱动器代号,demo.sas 是所写的 SAS 程序。

2. 键入 submit 指令或按功能键 F10,即可运行 DEMO.SYS 程序,用法如图 0.9。

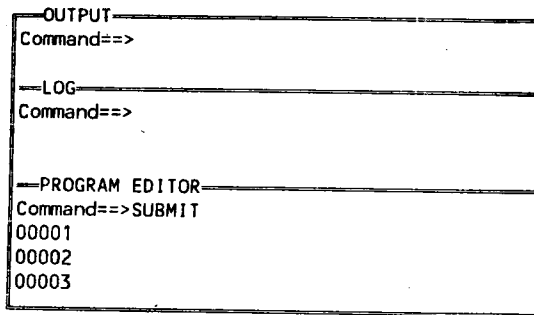


图 0.9

运行后的结果暂存于 OUTPUT 窗口,运行过程的记录暂存于 LOG 窗口。

3. 按 $\uparrow$, $\downarrow$, $\leftarrow$, $\rightarrow$ 键,移动光标至 LOG 窗口,并按功能键 F7 将窗口放大,即可看到运行记录的结果是否正确无误。若没有 ERROR 信息,表示该程序编辑正确。再按 F7 回到原来的三个窗口。

4. 移动光标至 OUTPUT 窗口(或按 F4 键),再按 F7 键将窗口放大,即可看到运行分析的最后结果。再按 F7 回到原来的三个窗口。

5. 在 OUTPUT 窗口键入 FILE 指令,即可存储当前在 OUTPUT 窗口上分析结果。其用法如图 0.10,表示已存储当前的所分析结果,A 驱动器中文件名叫 DEMO.OUT,其中. OUT 是自己命名的扩展文件名,当然可任选自己容易了解的扩展文件名。

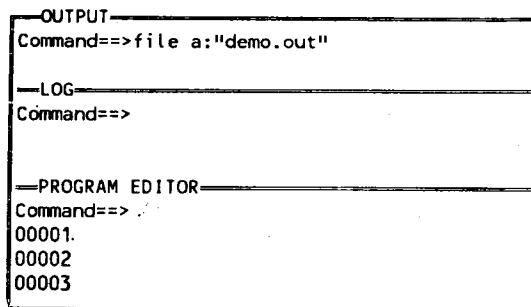


图 0.10

若欲将 LOG 窗口所记录的结果也存储起来,用法如图 0.11。

```
—OUTPUT—  
Command==>  
  
—LOG—  
Command==>file a:"demo.log"  
  
—PROGRAM EDITOR—  
Command==>  
00001  
00002  
00003
```

图 0.11

若欲将运行结果直接打印在打印机,其用法如图 0.12;或暂时先离开 SAS 系统,用 PE2 编辑先前所存储的 A:DEMO. OUT 文件,再用 PE2 所提供的 PRINT 指令,印出 DEMO. OUT 的内容至打印机上。

```
—OUTPUT—  
Command==>file 'prn'  
  
—LOG—  
Command==>  
  
—PROGRAM EDITOR—  
Command==>  
00001  
00002  
00003
```

图 0.12

6. 若欲清除任何窗口的数据,则键入 CLEAR 指令。例如欲清除 OUTPUT 窗口的数据,用法如图 0.13。每运行一次 SUBMIT 指令之后,如未将 OUTPUT 窗口及 LOG 窗口用 CLEAR 指令清除一遍,则下次的运行结果会接在上次的运行结果之后。


```
-----OUTPUT-----
Command==>clear

-----LOG-----
Command==>

-----PROGRAM EDITOR-----
Command==>
00001
00002
00003
```

图 0.13

7. 离开 SAS 系统的方法有二种:

(1)键入 X 指令,表示暂时离开 SAS 系统,回到 DOS 操作系统下。暂离时,原先 SAS 系统内各种状态均保持不变。欲由 DOS 再回 SAS,须在 DOS 提示符下键入 EXIT X 指令。用法如图 0.14。

```
-----OUTPUT-----
Command==>x

-----LOG-----
Command==>

-----PROGRAM EDITOR-----
Command==>
00001
00002
00003
```

图 0.14

(2)键入 BYE 指令或 ENDSAS 指令。

这样就可以永久地离开 SAS 系统。离开后,临时工作文件关闭,未存储至磁盘的数据将消失,如图 0.15。