

主编 金丕焕 苏炳华 贺 佳

# 医用 SAS 统计分析

YIYONG  
SAS  
TONGJIFENXI



上海医科大学出版社



# 医用 SAS 统计分析

主编 金丕焕 苏炳华 贺 佳

编写者(以姓氏笔画排序)

邓 伟	苏炳华	何清波	汪 涛
陆 健	范思昌	金丕焕	贺 佳
曹 阳	曹素华		

上海医科大学出版社



## 图书在版编目(CIP)数据

医用 SAS 统计分析/金丕焕,苏炳华,贺 佳主编.-  
上海:上海医科大学出版社,2000.12

ISBN 7-5627-0610-7

I. 医… II. ①金… ②苏… ③贺… III. 医学  
统计-应用软件,SAS IV. R195.1

中国版本图书馆 CIP 数据核字 (2000) 第 48953 号

责任编辑 沈彬源

## 医用 SAS 统计分析

主编 金丕焕 苏炳华 贺 佳

---

上海医科大学出版社出版发行

上海市医学院路 138 号

邮政编码:200032

新华书店上海发行所经销

昆山亭林印刷总厂印刷

开本 787×1092 1/16 印张 13.5 字数 328 000

2000 年 12 月第 1 版 2000 年 12 月第 1 次印刷

印数:1-3 000

---

ISBN 7-5627-0610-7/R·579

---

定价:27.00 元

如遇印、装质量问题,请直接与印刷厂联系调换  
(地址:江苏省昆山市中山路 293 号 邮政编码:215300)

# 前

# 言

# 言

SAS (Statistical Analysis System) 是国际上非常流行的统计分析软件, 目前已发展为一个全面的信息系统开发软件。其功能完善, 内容庞大。本书所介绍的微机用 6.12 版的部分内容即占磁盘空间达几百兆, 其指导资料也极为丰富。对于初学者来说有如临汪洋大海而无从入手的感觉。为了解决初学者的困难我们编写了本书。目的是让读者能在边学习、边实践的过程中, 比较轻松地掌握 SAS 统计分析的基本内容。掌握了这些基本内容, 读者就可以有能力进行目前常用的统计工作, 再进一步学习也就不那么困难了。

SAS 程序主要有数据步 (Data Step) 和过程步 (Proc Step)。一般 SAS 书籍都先讲述数据步, 再讲述过程步而且是从语句出发, 不易理解。本书以解决实际统计问题为中心, 从例子和程序出发, 同时并进, 使读者便于理解, 更易于掌握。

本书基本上是对《医用统计方法》(金丕焕主编, 上海医科大学出版社, 1993 年版) 一书中的大部分例子按顺序地逐个进行分析讨论。读者可参考该书复习统计理论和进一步理解 SAS 的输出内容。在讨论中并分析了 SAS 中的各重要语句的用法, 介绍了在没有现成的过程 (proc) 可用时用这些语句编写程序的方法。

SAS 内容十分丰富, 本书仅限于 Base 及 Stat 两部分的主要内容。本书扩展了 logistic 回归及卡方检验部分, 充实了新的内容。由于这些内容在国内的同类统计书中较少介绍, 使本书更加具有参考价值。书末附有习题及答案, 可供读者练习使用。

本书由上海医科大学、上海第二医科大学、第二军医大学有关教研室具有丰富教学经验的专家联合编写, 使本书的质量得以大大提高。在本书正式出版前曾作为研究生教材, 在试用过程中经不断进行改进和补充, 已修订了 4 次。学生在上课及上机学习之后, 在上机考试中都取得了优秀的成绩, 取得了良好的效果。

汪涛老师和邓伟老师在本书的编写和不断改进中付出了大量的精力。本书的主要输入工作皆由陈春巍、石桂芳老师完成。对于他们如此认真和耐心的工作谨表衷心的感谢。

限于编者水平和经验, 本书难免还存在不足之处, 恳切希望读者提出宝贵的意见和建议, 以便再版时进一步提高。

金丕焕

2000 年 8 月 2 日

# 目 录

# 录

<b>1 概论</b> .....	( 1 )
1.1 SAS 的安装 .....	( 1 )
1.2 SAS 的启动 .....	( 1 )
1.3 SAS 的运行 .....	( 1 )
1.4 一个简单的程序 .....	( 3 )
1.5 程序修改、调用及存储 .....	( 5 )
1.6 寻求帮助信息 .....	( 6 )
<b>2 SAS 数据集</b> .....	( 7 )
2.1 用 input 和 cards 语句建立 SAS 数据集 .....	( 7 )
2.2 用 infile 和 input 语句建立 SAS 数据集 .....	( 8 )
2.3 从已建立的 SAS 数据集中读入数据建立新的 SAS 数据集 ..	( 9 )
2.4 建立永久 SAS 数据集 .....	( 10 )
2.5 数据的横行输入 .....	( 11 )
2.6 缺失值的表示 .....	( 12 )
2.7 INPUT 语句的列输入格式 .....	( 13 )
2.8 其他建立 SAS 数据集的方法简介 .....	( 14 )
2.9 SAS 运算符 .....	( 14 )
2.10 SAS 常用函数 .....	( 15 )
<b>3 单变量的统计描述</b> .....	( 17 )
3.1 一般统计描述的 means 过程 .....	( 17 )
3.2 详细统计描述的 univariate 过程 .....	( 18 )
3.3 正态性检验 .....	( 20 )
3.4 频数分布表数据的输入 .....	( 22 )
3.5 几何均数的计算 .....	( 23 )

# 目 录

# 录

<b>4</b>	<b>t 检验</b> .....	(26)
4.1	配对比较 t 检验 .....	(26)
4.2	两样本均数差别的 t 检验 .....	(28)
<b>5</b>	<b>方差分析(一)</b> .....	(33)
5.1	完全随机设计资料的方差分析 .....	(33)
5.2	随机区组设计资料的方差分析 .....	(38)
5.3	拉丁方设计资料的方差分析 .....	(40)
5.4	GLM 过程与 ANOVA 过程 .....	(42)
<b>6</b>	<b>方差分析(二)</b> .....	(43)
6.1	$2 \times 2$ 析因试验设计 .....	(43)
6.2	三因素析因试验分析 .....	(45)
6.3	次级组样本含量不等资料的析因试验设计的方差分析 .....	(46)
6.4	两水平无重复正交试验设计资料的方差分析 .....	(48)
6.5	有重复两水平正交试验设计的方差分析 .....	(50)
6.6	一个 $L_9 3^4$ 正交试验设计的方差分析 .....	(50)
6.7	$L_{16} 4^3 \times 2^6$ 正交试验的方差分析 .....	(52)
6.8	标题 title .....	(54)
<b>7</b>	<b>线性回归与相关分析</b> .....	(55)
7.1	直线回归分析 .....	(55)
7.2	大样本资料的线性回归分析 .....	(56)
7.3	model 语句后面的选择项 .....	(59)
7.4	相关 corr 过程 .....	(61)

7.5 多元回归.....	(63)
7.6 简单相关系数及偏相关系数.....	(65)
7.7 筛选自变量的多元回归分析.....	(67)
<b>8 非线性回归.....</b>	<b>(71)</b>
8.1 非线性回归——对数曲线.....	(71)
8.2 多项式回归——抛物线.....	(74)
8.3 非线性回归——指数曲线拟合.....	(77)
<b>9 Logistic 回归.....</b>	<b>(80)</b>
9.1 两分类反应变量的 Logistic 回归.....	(80)
9.2 多分类无序自变量的 Logistic 回归——例 9-3.....	(86)
9.3 连续型数值自变量的 Logistic 回归——例 9-4.....	(89)
9.4 多分类有序反应变量的 Logistic 回归.....	(91)
9.5 多分类无序反应变量的 Logistic 回归.....	(95)
9.6 1:1 配对资料的条件 Logistic 回归.....	(99)
9.7 1:m 配对资料的条件 Logistic 回归.....	(101)
9.8 m:n 配对资料的条件 Logistic 回归.....	(104)
<b>10 卡方检验.....</b>	<b>(106)</b>
10.1 四格表资料的卡方检验.....	(106)
10.2 行×列表资料的卡方检验.....	(110)
<b>11 协方差分析.....</b>	<b>(119)</b>
11.1 一元完全随机设计资料的协方差分析.....	(119)

# 目 录

# 录

- 11.2 随机区组设计资料的协方差分析 ..... (122)
- 11.3 多元完全随机设计的协方差分析 ..... (125)

## 12 二项分布与 Poisson 分布 ..... (129)

- 12.1 二项分布函数 ..... (129)
- 12.2 用二项分布检验治疗效果 ..... (130)
- 12.3 求二项分布的概率分布 ..... (131)
- 12.4 Poisson 分布概率分布函数 ..... (132)
- 12.5 求概率分布 ..... (132)
- 12.6 样本计数与总体均数差别的统计意义检验 ..... (133)

## 13 非参数统计 ..... (135)

- 13.1 配对资料的符号等级检验 ..... (135)
- 13.2 两样本成组比较 ..... (136)
- 13.3 中位数法的非参数检验 ..... (138)
- 13.4 对实例 13-1 的 t 检验 ..... (138)
- 13.5 完全随机化设计资料的非参数检验 ..... (139)
- 13.6 Spearman 等级相关 ..... (140)
- 13.7 Kendall 等级相关系数 ..... (142)

## 14 多元统计分析 ..... (144)

- 14.1 判别分析 ..... (144)
- 14.2 逐步判别分析 ..... (147)
- 14.3 主成分分析 ..... (148)
- 14.4 因子分析 ..... (149)
- 14.5 指标聚类分析(R 型聚类分析) ..... (153)
- 14.6 由相关阵出发的指标聚类分析 ..... (160)
- 14.7 样本聚类(Q 型聚类) ..... (164)



<b>15 生存分析</b> .....	(168)
15.1 单变量 Wilcoxon 和 logrank 检验 .....	(168)
15.2 多变量的 Wilcoxon 和 logrank 检验 .....	(173)
15.3 生存资料的参数拟合 .....	(175)
15.4 Cox 回归分析 .....	(178)
15.5 比例风险模型有效性的检验 .....	(180)
附录一 练习题 .....	(182)
附录二 练习题答案主要部分 .....	(197)

# 1 概 论

本书将以 SAS 6.12 版为依据,将其安装、启动和操作作一简单介绍,并通过一个简单程序的输入和运行来说明 SAS 程序的基本结构;进入和退出 SAS 显示管理系统(Display Management System—DMS)的方法;程序的基本内容以及程序的编辑、运行、存储、调用及打印等过程。

## 1.1 SAS 的安装

我们将主要介绍在 Windows 98 下安装 SAS 时用户所要了解的基本内容。如果用户购买了 SAS 软件,则可按照安装指南进行。这里只作简单的介绍。

安装时在光驱中放入光盘,自动播放(Autoplay)功能可以直接启动 SAS 的安装界面。如果 SAS 的安装界面没有出现,可打开窗口底端任务栏的“开始(Start)”菜单,选择“运行(Run)”命令,在对话框中键入“光碟号:\SAS\setup”,确认后即可启动安装。

不同版本的 SAS 系统以及安装不同数目模块所需要的内存容量及外存容量都不相同。对于 SAS 6.12 版,一般都需外存容量数百兆,按安装模块数决定;内存需 12 兆。在安装过程中 SAS 会提示所需容量的数字。对于 SAS 在统计方面的应用,最基本的是 SAS/BASE 和 SAS/STAT,如果有 FSP、GRAPH 及 ACCESS 模块则比较理想。安装过程中 SAS 会提出是要典型安装或是要按需要安装。本书内容用典型安装即可。

## 1.2 SAS 的启动

### ● 由程序栏启动

SAS 安装成功后,在“程序(Program)”栏中会自动生成 SAS 组项目(group item)‘SAS’(或取名‘SAS 6.12’等)。包括 SAS 运行和更新两个图标(icon)。在启动 Windows 后用鼠标打开“开始(Start)”菜单,选择“程序(Program)”栏中 SAS 组项目中运行 SAS 的图标,双击鼠标即可启动 SAS 系统。

### ● 由 run 命令启动

打开“开始(Start)”菜单,选择“运行(Run)”命令,在 run 命令的对话框中输入子目录及 SAS.EXE 文件名,如 D:\SAS612\SAS.EXE 即可启动 SAS。

### ● 在文件名中直接选择启动

从资源管理器中直接找到 SAS.EXE 文件,用鼠标器激活即可启动。

## 1.3 SAS 的运行

### 1.3.1 显示管理系统

SAS 在启动后一般即进入显示管理系统屏幕。如图 1 所示。

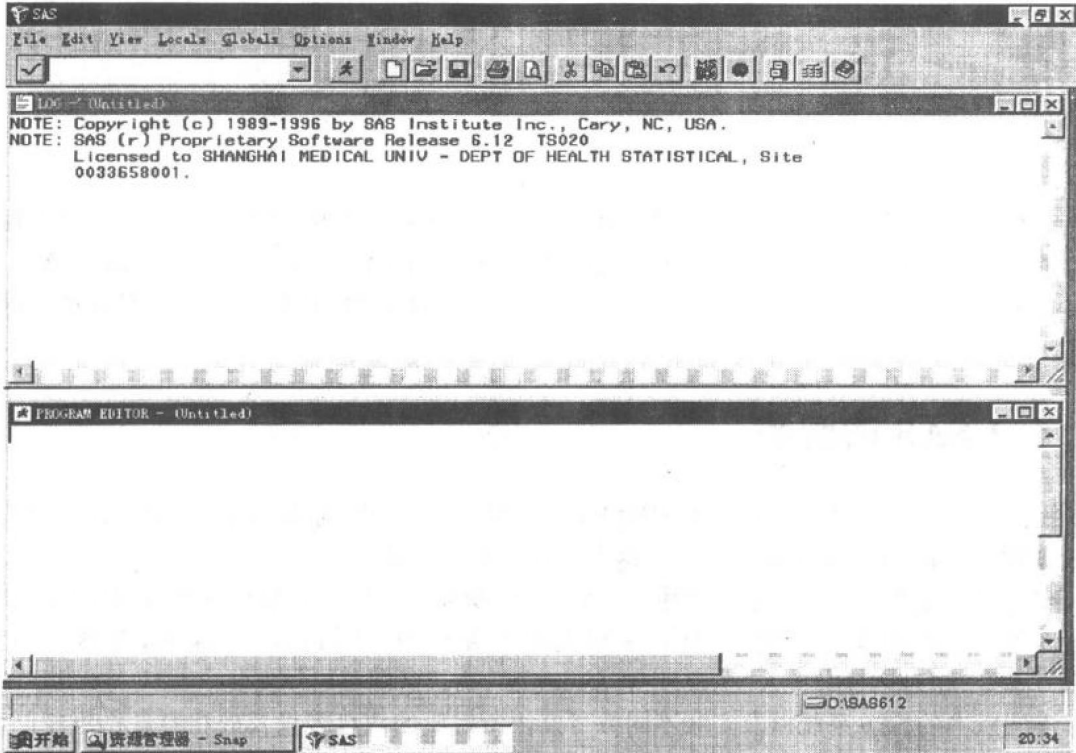


图1 SAS 6.12 版显示管理系统屏幕

在 SAS 的显示管理系统屏幕上通常有三个窗口:程序编辑窗口 (PROGRAM EDITOR)、记录窗口 (LOG) 及输出窗口 (OUTPUT)。6.12 版则显示除输出窗口外的两个窗口 (图 1)。由 SAS 菜单 (SAS 标题列下的一列) 中的 Windows 或 Globals 子菜单中可选择要显示的窗口。程序编辑窗口用于编辑 SAS 程序。程序编辑窗口中程序运行后如无错误则会显示输出窗口及窗口中输出的结果。

进入 SAS 系统后窗口中的内容根据设置而各不相同,可由菜单中 Options 项子菜单中 Preference 项中选择,在选择 Preference 后可出现一对话框。对话框中有两个小框 Commands 及 Tools 与显示内容有关。

### 1.3.2 显示管理系统的内容及设置

下面我们将逐项介绍 DMS 屏幕内容。

屏幕上第一列为标题列,显示的是 'SAS'。图 1 中各处图标的用法,如窗口的放大、缩小与退出等,与 Windows 中相同这里不再介绍。

标题除 'SAS' 外 LOG, PROGRAM EDITOR 及 OUTPUT 标题中的 [Untitled] 在有相应程序文件时将换成文件标识。

'SAS' 标题下第一列为 SAS 运行菜单,其中 File 为文件调用、存储、打印等内容。Edit 为编辑,用鼠标器如 Windows 中那样编辑则十分方便。Local 为本机文件运行所必需,其中尤为有用的是 Submit 及 Recall,前者为运行程序,后者为调回前面用过的程序。单击 Globals 可选择窗口。这些为运行 SAS 程序所必需,最后的 Help 为帮助,其他内容读者可逐步熟悉。

菜单下方一列为工具条(Tool bar),在 Tools 对话框中激活 Tool bar 即可出现。如激活 Tool box 则在屏幕右上方出现较大的工具条。如激活 Tool Tips, 则当鼠标在工具条某图标时显示该图标的作用,如光标在自左至右的各图标上时依次可显示 Run(运行)、Open(打开文件)、Save(存储文件)、Print(打印)、Cut(裁剪)、Mark(标记)、Paste(粘贴)、Undo(恢复)、Library(库)、Assist(应用 SAS/ASSIST 模块)及 Help(帮助)。在各窗口左上方的 Command = = = > 是供在其后输入命令用的。在 preference 对话框中激活 Command line 即可出现。如在命令中输入 Submit 则可运行程序,Recall 则可调回用过的程序,Clear 则清除屏幕等。

在程序编辑窗口中如需如 6.04 版中的行号可由在 'Command = = = >' 后键入 num 并回车即可出现,键入 num off 则消失,键入 bye 则退出 SAS 系统。在行号中可以进行编辑,这对熟悉 6.04 版而不熟悉 Windows 版的操作者十分有利,读者可参考其他 SAS 6.04 版的书籍。熟悉 Windows 的可用鼠标器编辑,或用菜单中 Edit 内容进行编辑。

用功能键也可进行简便操作,但各版本各功能键设置的作用并不相同,读者应根据自己版本使用。功能键也可自己设置,设成自己熟悉的用法。

以上介绍主要根据 6.12 版,其他版本会有些不同。

### 1.3.3 调用其他模块

在工具条中激活 SAS/ASSIST 即可调用 SAS/ASSIST 模块,在屏幕上出现菜单,可调用其他各模块,如 ACCESS 模块等。此外,在 DMS 的 Command = = = > 行输入相应命令也可调用。如输入 FSEDIT 命令则可进入 FSP 的编辑状态。

## 1.4 一个简单的程序

在程序编辑窗口输入一个程序,运行后即可得到显示在输出窗口的结果。下面用一个简单的例子,以说明程序的结构。

### 1.4.1 例 1-1

12 名健康成年男子血液第一小时末红细胞沉降率分别为 3,9,8,6,5,5,7,3,10,8,10,4mm,试计算其均数。

对于此例可在程序编辑窗口编写一简单程序如下:

程序 1-1

```
00001 data esr;  
00002 input x;  
00003 cards;  
00004 3  
00005 9  
00006 8  
00007 6  
00008 5  
00009 5  
00010 7
```

```
00011 3
00012 10
00013 8
00014 10
00015 4
00016 ;
00017 proc means;
00018 run;
```

如果未用 num 命令则没有前面的行号。

### 1.4.2 程序结构说明

SAS 程序由数据步 (Data step) 与过程步 (Proc step) 组成本例从 data esr; 语句到 proc means; 语句之前 (语句标号 00001 ~ 00016) 皆为数据步。后面一个语句 (语句标号 00017) 开始为过程步。proc 为英文单词 “procedure” 的缩写, 读音为 prok。数据步与过程步均由一个或几个语句组成, 每个语句都以一个关键词开始, 以分号 (“;”) 结束。最后一个语句 run; 指示数据步或过程步结束, 运行程序。

数据步的作用是输入数据并建立 SAS 数据集。对于本例, 程序运行后即生成一名为 esr.sd2 的 SAS 数据集。SAS 数据集有两点需要说明: ① SAS 数据集的后缀名一律为 .sd2; ② SAS 数据集的后缀名一律不出现在程序中。另外, SAS 系统只能分析 SAS 数据集中的数据。SAS 的数据步中有许多专用于数据步的语句, 如上例的 input 和 cards 语句。其中, input 语句用来生成变量, cards 语句用来指明数据输入的开始。

过程步的作用是对已形成的 SAS 数据集通过调用现成的 SAS 过程进行统计分析、打印等处理。对于上例: proc means; 语句即是调用 means 过程以计算默认 SAS 数据集某个变量的最大值、最小值、均数和标准差等统计量。再例如下面的程序即是调用 print 过程输出指定 SAS 数据集 esr 中的数据。

程序 1-2

```
00001 proc print data = esr;
00002 run;
```


执行以上程序即可在输出窗口列出 SAS 数据集 esr 的数据如下:

OBS	x
1	3
2	9
3	8
4	6
5	5
6	5
7	7
8	3
9	10

10	8
11	10
12	4

proc print 语句后的 data = esr 是 SAS print 过程的一个选择项,该选择项适用于 SAS 的所有过程,其作用在于指定将要被处理的 SAS 数据集,其书写格式为 data = SAS 数据集名。如缺省则 SAS 过程对最新生成的 SAS 数据集(默认的 SAS 数据集)进行处理。

### 1.4.3 程序运行

在程序编辑窗口输入程序后可在命令行键入 submit 或单击 SAS 运行菜单中  图标即可运行程序。如程序正确即可在输出窗口显示结果;如程序有错则在记录窗口用红字显示出错信息。

对于程序 1-2,运行正确时在输出窗口显示出如下结果:

Analysis Variable : X				
N Obs	Minimum	Maximum	Mean	Std Dev
12	3.0000000	10.0000000	6.5000000	2.5405797

在记录窗口显示如下内容:

```

1   data esr;
2   input x;
3   cards;
16  ;

```

NOTE: The data set WORK.ESR has 12 observations and 1 variables.

NOTE: The DATA statement used 2.00 seconds.

```

17  proc means;
18  run;

```

NOTE: The PROCEDURE MEANS used 2.00 seconds

## 1.5 程序修改、调用及存储

### 1.5.1 全屏幕编辑命令

在程序编辑窗口编写程序时难免需作各种改动。为此,SAS 提供了全屏幕编辑。这使得程序编辑窗口变成了一个文本编辑软件(实际上就是)。其编辑方法与 Windows 中相同,用 Edit 进行,这里不再介绍。另外,在任一窗口中均可在命令行上用以下屏幕翻卷命令或功能键:

top 或功能键 Ctrl + PageUp	光标移至程序第一行
bottom 或功能键 Ctrl + PageDown	光标移至程序最后一行
n	光标移至程序第 n 行

left                                   使窗口左移  
right                                  使窗口右移

此外,可用 < PageUp > 键或 < PageDown > 键使窗口向前或向后翻一页,用 < Home > 键使光标回到命令行,以便输入命令。

### 1.5.2 其他命令


在任一窗口命令行键入 clear 并回车即可清除屏幕内容。

在运行后程序编辑窗口内的程序即消失,为调回原程序(进行修改或再运行程序)可先进入编辑窗口再用菜单单击 Locals - Recall text 即可。每一次即调回更前面的内容,直至全部调出。

用进入编辑窗口再用菜单单击 File - Save 或用 File - Save As 或在命令行键入 file 命令及文件路径和文件名即可存储文件。如:

Command = = = > file 'c:\book\mean.prg' 则可把程序编辑窗口内的所有内容以取名为 mean.prg 的文件存入 C 盘的名为 book 的子目录中。如不用任何路径则存入当前子目录。同样可以存储输出窗口和记录窗口的内容。

对于已存储的程序可以进入编辑窗口再用菜单单击 File - Open 或在命令行键入:include(或 inc) '文件路径及文件名'。如:include 'c:\book\mean.prg' 或 inc 'c:\book\mean.prg'。即可调入名为 mean.prg 的程序,如不用任何路径则意味着从当前子目录中调入文件。

进入编辑窗口再单击菜单中的  图标或在命令行用 file 'pm' 命令即将该窗口中(不仅限于显示于屏幕部分)的所有内容输出至打印机。同样适用于输出窗口和记录窗口的内容。必要时也可以把某一窗口先存储在盘上,以后退出 SAS 后再用一般打印文本文件的方法输出至打印机。

### 1.5.3 定义功能键

应用功能键可以执行以上所述的功能,快速而方便。各版本的功能键不同,但功能键均可由我们自己修改并重新定义。我们可以在 KEYS 窗口浏览这些功能键的初始定义。进入 KEYS 窗口的办法就是在任一窗口的命令行上键入命令 KEYS。进入 KEYS 窗口后,我们可以根据自己的喜好修改功能键的定义或者增加一些功能键。但要注意在完成后,在 KEYS 窗口的命令行上键入 SAVE 命令。而且,下回再进入 SAS 时,要使用重新定义的功能键,必须从同一子目录进入方可。

## 1.6 寻求帮助信息

进入编辑窗口再用菜单单击 Help 图标,或在任何窗口的命令行键入 help 命令或在任何时候用帮助功能键(一般为 <F1>)均可进入 help 窗口。在该窗口中可显示对所有过程、命令、语句或函数等的简要解说,犹如一本全面而简明的说明书。读者可根据屏幕提示用选择菜单的方式调用下一个帮助面板或用 end 命令返回前一个帮助面板直至退出 help 窗口。

SAS 数据集是使用 SAS 软件的最基本也是最重要的一环,要想灵活、方便地使用 SAS 软件,必须掌握 SAS 数据集的建立和调用。本章将介绍建立 SAS 数据集的几种方法及对 SAS 数据集操作的常用语句。

## 2.1 用 input 和 cards 语句建立 SAS 数据集

当数据较少时,用这种方法建立 SAS 数据集较好。例 1-1 的 SAS 数据集就是用这种方法建立的。为更详细地介绍这种方法,我们再举一个例子。

### 2.1.1 例 2-1

收集 30~178 个月的正常儿童的有关健康指标如表 2.1(截选)。

表 2.1 30~178 个月的正常儿童的有关健康指标

编号 ID	性别 X1	月龄 X2	身高 X3	体重 X4	胸围 X5	心像面积 X6
1	m	32	95.5	14.0	53.5	49.64
2	m	35	92.0	13.0	52.0	41.61
3	m	33	89.0	12.5	53.5	35.81
(中间省略)						
254	m	176	168.0	53.5	82.0	100.14
255	f	30	91.0	11.0	48.0	35.39
256	f	33	91.0	11.5	47.0	44.98
521	f	178	163.0	51.0	79.0	87.42

用上表数据建立 SAS 数据集,取名 child.sd2,变量名如表格所示。

### 2.1.2 程序

程序 2-1

```
data child;
  input id x1 $ x2 x3 x4 x5 x6;
cards;
  1      m      32      95.5      14.0      53.5      49.64
  2      m      35      92.0      13.0      52.0      41.61
```



3	m	33	89.0	12.5	53.5	35.81
---	---	----	------	------	------	-------

(中间省略)

254	m	176	168.0	53.5	82.0	100.14
255	f	30	91.0	11.0	48.0	35.39
256	f	33	91.0	11.5	47.0	44.98
521	f	178	163.0	51.0	79.0	87.42

;  
run;

第一句 data child; 是要求建立一个 SAS 数据集, 数据集将被放在当前目录的子目录 \SASWORK 中, 名为 child.sd2。SASWORK 子目录是 SAS 自动建立的子目录, 非用户所能控制。数据集放在此子目录中, 只要不退出 SAS, 可随时调用。但一旦退出 SAS 系统, SAS 会自动删除 SASWORK 子目录下的所有文件。建立的数据集亦不复存在, 所以这样的数据集称为临时性数据集。

第二句 input id x1 x2 x3 x4 x5 x6; 是要求建立的 SAS 数据集中生成名为 id、x1、x2、x3、x4、x5 和 x6 的七个变量。x1 变量后的 '\$' 表示该变量为字符型变量。

第三句 cards; 表明数据行的开始。之所以用“cards”表示是因为早期计算机的数据是用穿孔卡片(punched cards)输入的。cards; 语句下面即为要输入的数据, 数据之间用一个或几个空格分开, 至分号“;”为数据步结束。注意: 分号“;”必须另起一行单独写, 不能写在最后一行数据的后面。

## 2.2 用 infile 和 input 语句建立 SAS 数据集

用这种方法时, 首先要建立一个数据文件。你可以用任何文字处理软件来建立它。但要记住, 一定要用纯文本格式保存。如我们用 WPS 或 WORD 来建立上例的数据文件如下:

1	m	32	95.5	14.0	53.5	49.64
2	m	35	92.0	13.0	52.0	41.61
3	m	33	89.0	12.5	53.5	35.81

(中间省略)

254	m	176	168.0	53.5	82.0	100.14
255	f	30	91.0	11.0	48.0	35.39
256	f	33	91.0	11.5	47.0	44.98
521	f	178	163.0	51.0	79.0	87.42

注意到数据文件是不包含变量名的, 你必须知道哪一列数据是哪一个变量的值。数据