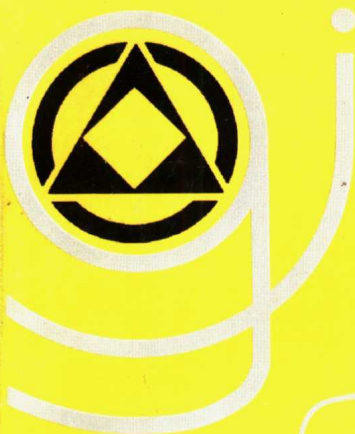
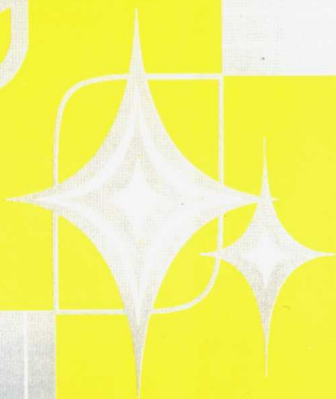




全国高等农业院校教材

全国高等农业院校教材指导委员会审定



SAS 及应用

● 各专业通用

● 裴喜春 薛河儒 主编

中国农业出版社

全国高等农业院校教材

SAS 及 应 用

裴喜春 薛河儒 主编

各 专 业 通 用

中 国 农 业 出 版 社

全国高等农业院校教材

SAS 及 应 用

裴喜春 薛河儒 主编

责任编辑 何致莹

出 版 中国农业出版社

(北京市朝阳区农展馆北路 2 号 100026)

发 行 新华书店北京发行所

印 刷 北京科技印刷厂

* * *

开 本 787mm×1092mm 16 开本

印 张 14.75 字数 336 千字

版、印次 1998 年 3 月第 1 版

1998 年 3 月北京第 1 次印刷

印 数 1~5,000 册 定价 20.00 元

书 号 ISBN 7-109-05219-2/TP·15

ISBN 7-109-05219-2



9 787109 052192 >

国家教委面向21世纪教改研究项目
中华农业科教基金 资助

内 容 摘 要

书中系统介绍了通用数据统计分析系统 SAS 软件的基本概念、程序结构、标准过程及方差分析、回归分析、聚类分析、主成分分析、因子分析、判别分析等多种多元统计分析的过程和应用方法。

本书可作为理、工、农、林、医、水文、地质、气象、经管、统计等有关学科的科技人员、大专院校教师、研究生和高年级大学生从事计算机数据统计分析的教材或参考书。

主 编 裴喜春 (内蒙古农牧学院)

副主编 薛河儒 (内蒙古农牧学院)

编 者 李永慈 (河北农业大学)

陈 勇 (内蒙古农牧学院)

主审人 梅建德 (中国农业大学)

序

试验统计是农林院校（包括医、理、工科类院校）多数专业中的一门重要专业基础课。多年来，由于信息量、计算量大等原因，使得课程中讲授的多数理论和方法不便引入专业课和在研究实践中得到应用。随着计算机技术的推广应用与普及，很多学者在这方面进行了不懈的探索和努力，以寻求解决这一问题的方法和途径。

几年来，国内外先后推出的关于统计计算程序书、程序库、软件包等已有多种。本书编著者选择了国际上公认的标准统计分析软件 SAS (Statistical Analysis System) 作为教学软件方案，起点高且具有很强的实用性和可行性。本教材已在教学中使用多遍，历时六年，几经修改最终定稿，这是一件值得庆贺的事情。

目前，我国教育界正处在面向 21 世纪的教育、教学改革广泛、深入地研究与实践中，以提高农林本科计算机素质为目标的计算机教育改革也和其它门类学科一样，教育改革研究与实践进行得如火如荼。国家教委立项的面向 21 世纪“高等农林院校本科计算机系列课程教学内容和课程体系改革的研究与实践”课题组，作为课题研究的阶段性成果提出农林院校计算机基础教育分三个层次进行，在第三个层次（应用基础）教育中，曾列出统计计算（统计电算化）等多门课程作为各院校根据本校实际选择的计算机应用基础课的课程。令人高兴的是，本书编著者们率先编撰了《SAS 及应用》一书，并成功地进行了教学实践，在农林院校计算机教学中的“应用基础”教育层次上迈出了可喜的第一步。教改是全方位的改革活动，既包括教学内容和课程体系的改革，也应在教学方法的改革上下功夫，本书的编著者们为配合本教材的使用，研制了配套课件 SAS CAI，经多次使用，证明该软件是成功的，在 1997 年度“全国高等农林院校计算机教育改革研究协作组”年会上被评为优秀软件。相信本教材的出版，将会对农林院校的计算机应用基础教育乃至整个计算机教育起到很大的推动作用。希望本书的编著者们继续努力，不断有新的力作问世，更希望其他院校的同仁，结合农业院校的特点，编撰出用于计算机教育的各种教材及开发配套的 CAI 课件，力争做到“起点高，立意新”以实际行动为推动计算机教育、教学改革作出贡献。

中国农业大学计算机与应用数学系 梅建德

一九九七年十月于北京

前 言

试验统计是农林医类专业重要的专业基础课。然而数据量大和计算量大一直是困扰从事统计计算科学工作者的一个问题。电子计算机的广泛应用,为统计计算提供了有力的手段。然而由于统计计算方法和程序设计的复杂性,也给非计算机专业人员利用计算机进行统计计算带来一定的困难。因此,在农林医类专业开设统计电算化课程势在必行。

目前,国内外推出很多统计电算化软件和程序库,由美国科学家开发的统计分析系统 SAS (Statistical Analysis System) 以其实用、可靠、通用和操作简便的特点称誉世界。SAS 软件的开发已有三十余年的历史,它是一个较成熟的系统。统计分析是 SAS 开发的最初目的,发展至今它还包括了绘图、编辑、运筹、质量控制、商业分析、预测和财务管理计算等方面的内容。

本书从农林医类专业的试验数据处理的统计分析电算化为目的,着重介绍了 SAS 的基本知识和统计分析功能,包括 SAS 的入门、标准过程和方差、回归、聚类、主成分、因子、典型相关、判别等统计分析过程。有了计算机基础知识与试验统计方法知识以后,本专科生可以用 40 学时左右讲授本书的前七章,学习过多元统计方法的高年级本科生和研究生,可以用 50 学时(包括上机)左右的时间学习本书的全部内容。

由于时间紧迫和编著者的水平有限,书中谬误难免,请读者指正。

书中引用多本国内外知名的统计分析著作中的数据与例题,在此向编著者深表谢意。

编 者

一九九七年九月

目 录

序

前言

第一章 SAS 概述	1
§ 1.1 SAS 的发展概况	1
§ 1.2 SAS 的结构、功能和特点	1
§ 1.3 SAS 系统的启动及简单操作	3
§ 1.4 程序编辑窗口的操作	4
§ 1.5 实例	6
第二章 SAS 语言基本概念	9
§ 2.1 SAS 语言概述	9
§ 2.2 SAS 语句	10
§ 2.3 SAS 数据集	11
§ 2.4 变量的属性	12
§ 2.5 SAS 程序的书写格式	13
§ 2.6 SAS 函数	14
§ 2.7 SAS 表达式	17
第三章 数据步与数据步语句	20
§ 3.1 数据步基本概念	20
§ 3.2 数据步程序的三种基本结构	20
§ 3.3 报告生成程序的基本结构	22
§ 3.4 数据步程序的执行流程	22
§ 3.5 数据步基本语句	23
§ 3.6 数据步程序控制语句	31
§ 3.7 循环语句	35
§ 3.8 数组概念及 ARRAY 语句	38
§ 3.9 输出语句与输出文件	40
第四章 过程步与过程步语句	44
§ 4.1 SAS 过程步的一般形式	44
§ 4.2 常用过程步语句	45
§ 4.3 通用 SAS 语句	46
第五章 通用标准过程	49
§ 5.1 绘图过程 CHART	49
§ 5.2 计算相关系数过程 CORR	57
§ 5.3 SAS 与 dBASE 间数据转换过程 DBF	62

§ 5.4 简单统计分析过程 MEANS	63
§ 5.5 二维散点图绘制过程 PLOT	67
§ 5.6 输出数据集过程 PRINT	72
§ 5.7 排序过程 SORT	73
§ 5.8 数据标准化过程 STANDARD	74
§ 5.9 数据集转置过程 TRANSPOSE	76
§ 5.10 两组数据均值比较过程 TTEST	78
第六章 方差分析	81
§ 6.1 平衡设计的方差分析过程 ANOVA	81
§ 6.2 不平衡设计的方差分析过程 GLM	103
§ 6.3 嵌套设计的方差分析过程 NESTED	112
§ 6.4 试验设计过程 PLAN	116
第七章 回归分析	121
§ 7.1 一般回归分析过程 REG	121
§ 7.2 多项式回归分析过程 RSREG	127
§ 7.3 逐步回归分析过程 STEPWISE	131
§ 7.4 非线性回归分析过程 NLIN	134
第八章 聚类分析	139
§ 8.1 系统聚类分析过程 CLUSTER	139
§ 8.2 逐步聚类分析过程 FASTCLUS	147
§ 8.3 变量聚类分析过程 VARCLUS	151
第九章 多变量统计分析	158
§ 9.1 主成分分析过程 PRINCOMP	158
§ 9.2 因子分析过程 FACTOR	163
§ 9.3 典型相关分析过程 CANCORR	171
第十章 判别分析	178
§ 10.1 贝叶斯判别分析过程 DISCRIM	178
§ 10.2 典型判别分析过程 CANDISC	187
§ 10.3 逐步判别分析过程 STEPDISC	196
附录 SAS 保留字集	204
参考文献	227

第一章 SAS 概述

§ 1.1 SAS 的发展概况

SAS (Statistical Analysis System) 是一个大型的数据管理与数据统计分析处理的软件包。1966 年, 由美国 North Carolina 州立大学开始研制 SAS。1976 年在美国成立了 SAS 研究所 (SAS Institute Inc, SAS Circle, Box 8000 Cary, NC 27512 - 8000)。SAS 研究所对商品化了的 SAS 系统进行经营、维护、培训和进一步开发等工作。早期的 SAS 只能在大型计算机上运行。1985 年推出了可以在 IBM PC 及其兼容机上运行的 SAS 版本。本书将主要介绍如何在 IBM PC 机上使用 SAS, 至于其它机型上的 SAS 版本与此大同小异, 读者有了本书的基础后, 可以查阅相关机型的 SAS 手册。

目前, SAS 已经发展成为一个功能齐全, 应用范围广泛和使用灵活方便的数据管理, 数据分析的标准软件系统, 有人称之为新一代计算机语言。其应用范围涉及到理、工、农、林、医、管理、商业和行政事务等各个领域。国际上有一个专门的 SAS 协会 SUGI (SAS User Group International), 每年有学术会议讨论研究有关 SAS 的问题。一些国家和地区的大学把 SAS 作为一门课程开设。近几年, 我国的一些科研机构 and 大学也引进并使用了 SAS。

§ 1.2 SAS 的结构、功能和特点

1.2.1 SAS 的结构

SAS 是一个用于数据管理和数据处理分析的组合软件系统, 它包括多个大的功能模块, 用户可以根据需要, 选取部分或全部 SAS 功能模块来组成一个运行系统。SAS 的微机版本主要有如下几个功能模块:

(1) SAS/BASE 是 SAS 的核心部分, 其它功能模块必须是在该模块的支持下运行。本模块提供的功能有: 信息存储与检索、程序和数据的编辑与修改、报告生成、简单统计计算和文件存储等。

(2) SAS/STAT 模块提供对数据进行统计分析的功能, 包括方差、相关、回归、判别、聚类、主成分和因子分析等多元统计分析方法。

(3) SAS/AF 模块提供了屏幕控制功能, 在本模块的支持下, 用户可以将其应用系统设计成为菜单驱动的形式, 通过在屏幕上的选择、填空以及调用有关信息实现对 SAS 的使用。

(4) SAS/FSP (Full Screen Procedures) 提供了较强的数据录入、编辑和检索等方面的功能。

(5) SAS/GRAPH 有很强的绘图功能。它支持以图形方式输出数据分析的结果。系统可以产生多种彩色的统计图形、地图、曲线图以及三维图形。

(6) SAS/ETS (Econometric and Time Series) 支持计量经济与时间序列分析研究。包括了计量经济学中的统计模型分析、时间序列分析及频数分析等方法。它是经济分析、预测、系统模型、金融等应用研究方面的工具。

(7) SAS/IML (Interactive Matrix Language) 是一种用于复杂数据分析和处理的交互式矩阵语言。它具有完善的编程环境和矩阵运算语言, 并且可以在语句级交互使用。

(8) SAS/OR (Operations Research) 是基于运筹学原理的一个程序包, 其主要功能有: 计划和管理大型项目、资源的调度分配、建立生产和网络流程模型等。

(9) SAS/QC 是一个用于产品质量控制管理的子系统, 本模块基于统计学方法, 用计算机图形显示技术来实现产品制造过程中的质量控制管理。

除了上述主要模块外, SAS 系统还有一些辅助性模块, 例如便于初学者使用的菜单化了的 SAS 调用系统 SAS/ASSIST, 全特性终端仿真器 SAS/RTERM 可以将 PC 机连接到主机上等。

本书将主要介绍 SAS/BASE、SAS/STAT 以及 SAS 入门。对 SAS 的其它功能感兴趣的读者, 有了本书的基础之后, 可以直接阅读 SAS 的软件说明书来使用它。

1.2.2 SAS 的特点

SAS 是一个实用性强、功能完善、使用方便、容易学习的计算机软件系统。它不仅具有一般数据管理系统的功能, 还提供了一个完善的可编程语言环境, 特别是以标准过程给出了常用的数据统计处理的算法。例如, 对于计算繁难的多元统计算法, 用户只需要指出过程名及其必要的参数提交系统, 就可以得到一张清晰的包括相应算法的全部计算结果和参数的输出表格。SAS 的这一特点极大的方便了非计算机专业人员。SAS 系统中典型的标准过程包括一般描述性统计分析和多元统计分析、经济预测与时间序列分析、多种运筹学算法、质量控制管理及绘制各种二维、三维图形等。

另外, SAS 具有极强的报表产生和绘图功能。用户可以得到清晰易读的表格和图形输出。SAS 还提供了与 Lotus 1-2-3、dBASE 数据等库软件数据交换接口。

1.2.3 SAS 运行环境

在 IBM PC 及其兼容机上的要求:

操作系统: PC DOS 或 MS DOS 2.0 版以上。

内存: $\geq 640\text{Kb}$

硬盘: $\geq 20\text{Mb}$

显示器: CGA、EGA、VGA

汉字环境: UC DOS 3.1 以上版本

协处理器可以提高系统的运算速度。

§ 1.3 SAS 系统的启动及简单操作

1.3.1 启动 SAS

假设 SAS 系统存储于 C 盘的名为 SAS 的子目录下, 那么当计算机用 DOS 启动后, 以如下操作启动 SAS 系统:

C> CD SAS 进入 SAS 子目录

C> SAS 启动 SAS

上述操作之后, 稍候片刻, 屏幕上将出现如图 1.1 所示的三个窗口, 这就进入了 SAS 的显示管理系统, 这是 SAS 的基本工作状态。在这种状态下, 用户可以与 SAS 会话, 当用户输入 SAS 命令或语句并将它们提交 SAS 系统去执行, 系统将给出运行结果和有关信息。

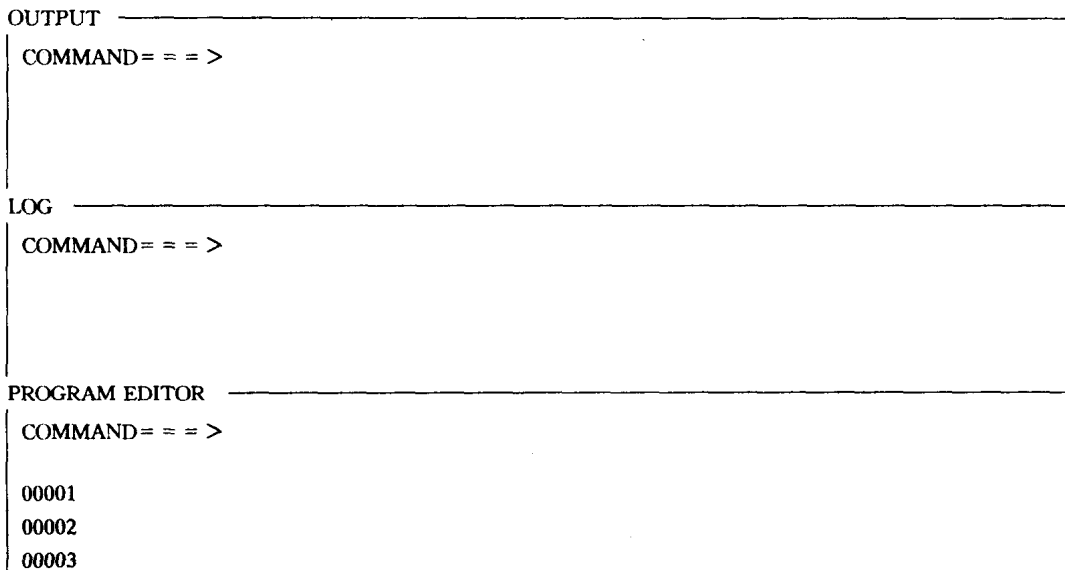


图 1.1 SAS 窗口

1.3.2 SAS 窗口

图 1-1 是 SAS 启动后屏幕的状态, 三个区分别为三个窗口。

(1) 输出 (OUTPUT) 窗口。输出窗口用于显示 SAS 过程的输出结果。

(2) 日志 (LOG) 窗口。日志窗口是显示 SAS 系统在执行用户提交的任务时的有关信息, 如运行时间, 出错信息等。

(3) 程序编辑 (PROGRAM EDITOR) 窗口。

该窗口是一个文本编辑器, 它像多数全屏编辑软件一样, 支持用户编辑输入程序、数据等文本文件。一般用户的 SAS 程序和数据文件都是通过这个窗口建立的。在该窗口的左侧, 系统自动提供五位整数的一列行号。在该窗口下, 系统提供了较强的插入、删除、移动、复制、字符串查找等有用的编辑功能。

在 SAS 里，还有多种窗口，用户可用有关 SAS 命令 (KEYS) 调用之。

1.3.3 SAS 窗口的特性

(1) 每一个窗口分为两个区：命令行区和正文区。窗口的第一行显示有 COMMAND 的行是命令行区，在该区用户可以输入 SAS 命令。命令行以下的区域是正文区。正文区将用于输入程序或显示信息。

(2) 通过光标移动键可以将屏幕上唯一的光标移至任一窗口的任意位置。

(3) 在一个窗口的命令行上通过输入命令 ZOOM 可以将该窗口放大到满屏或复原。

(4) 用光标移动键和翻页键 (PgUp 和 PgDn) 可以查看窗口的全部内容。

(5) 三个窗口的正文区内容都可以作为 DOS 文件存入在外存储器上，并且可以利用其它字处理软件对这些文件进行编辑处理。

1.3.4 窗口内容的存储与打印

(1) 窗口内容存盘命令。如下命令就可以实现对窗口内容以文件方式存盘：

COMMAND ===> FILE '文件名'

其中 FILE 是命令关键字，文件名是一个合法的 DOS 文件名，它可以是一个包括设备名、子目录、文件名和扩展名的完整文件标识。例如字符串 D: \ MYDIR \ MYFIL.DAT 是存储于 D 盘 MYDIR 子目录下一个文件，其文件名为 MYFIL 且扩展名为 DAT。

(2) 调用盘文件到窗口的命令。如下命令可以将指定文件调入窗口：

COMMAND ===> INCLUDE '文件名'

其中 INCLUDE 是命令关键字，文件名如 1.3.4.1 所述。这是一个已经存在指定路径下的文件。

(3) 窗口内容输出到打印机命令。命令格式如下：

COMMAND ===> FILE 'PRN:'

其中 FILE 是命令关键字，PRN: 是系统提供的打印机设备名。

(4) 清除窗口内容命令。随着 SAS 系统的运行，各个窗口的内容在不断的增加，这样会占用大量存储空间。所以经常使用 CLEAR 命令清扫窗口是必要的。

命令格式如下：

COMMAND ===> CLEAR

(5) 退出 SAS 返回 DOS。在任意窗口的命令行上按如下命令就可以退出 SAS，返回到 DOS 下：

COMMAND ===> BYE

或 COMMAND ===> ENDSAS

§ 1.4 程序编辑窗口的操作

程序编辑窗口是一个功能强、使用灵活的字处理器，在该窗口下用户可以编辑产生任

意的文本文件，对于编辑的文件可以存盘。如果编辑的是 SAS 程序，可以在完成编辑后直接提交 SAS 系统运行。SAS 提供了对文本的行与块的插入、删除、移动、复制及字符串查找等命令。

1.4.1 行 插 入

格式 1：在一行行号的五个位置上键入 I 并回车则在该行后出现一空行。

格式 2：在行号位置上按入 I n 并回车，则在该行后出现 n 行空行，其中 n 为正整数。

1.4.2 删 除

(1) 在一行的行号任一位置上键入 D 并回车则这一行被删除。

(2) 在 n1 行和 n2 行的行号上分别键入 DD 则从 n1 到 n2 行之间的所有行被删除。

1.4.3 复 制

在要被复制的源行行号上按入 C 或在被复制的连续多行的第一行和最后一行行号上按入 CC，这一步完成了复制源的定义。然后，将光标移到复制的目标位置行的行号上，若按入 A 则复制内容被复制到该行之后；若按入 B 则复制内容被复制到该行之前。

1.4.4 移 动

移动操作也分两步进行，首先要定义被移动内容，若移动单行，则在该行的行号位置上按入 M，若移动连续多行的一个块，则在该块的首末两行的行号上按入 MM。这一步叫定义移动源。第二步要进行的是移动，若要将已定义了的移动内容移至某一行的前边则在该行的行号上按入 B，否则按 A。

1.4.5 重 复

使用 R 或 RR 可以使一行或多行的一块重复出现一次。

例 1.1 假设要对如下一段出现在编辑窗口的内容进行编辑：

```
00001  LING A
00002  LINE B
00003  LINE C
00004  LINE D
00005  LINE E
00006  LINE F
```

使之变成为：

```
00001  LINE D
00002  LINE E
00003  LINE A
00004  LINE B
00005  LINE C
00006  LINE F
```

只要在原来显示的第一行和第三行的行号上分别按入 MM, 而后在第 5 行的行号上按入 A 就可以实现要求。操作如下:

```
MM001  A
00002  B
MM003  C
00004  D
a0005  E
00006  F
```

1.4.6 字符串搜索与自动替换

在进行文件编辑或查阅时, 依指定字符串将光标定位和依指定新字符串自动替换旧字符串的操作是十分有用的操作。SAS 提供了这一功能。这些命令用到“范围选项”的概念来指定搜索和替换的范围。范围选项共有五个可选词:

- ① FIRST 指文件中第一个指定的字符串。
- ② LAST 指文件中最后一个指定的字符串。
- ③ NEXT 指文件的光标当前位置之后的部分。
- ④ ALL 指整个文件。
- ⑤ PREV 指前一个字符串。

(1) 字符串搜索命令。

格式: COMMAND = = = > FIND“字符串” [范围选项]

功能: 执行命令的结果是将光标停在文件中指定范围内的第一个指定的“字符串”上。“范围选项”可以是上文给出的任一个。如果“字符串”中不含有空格, 可以不用引号括之。

(2) 字符串自动替换命令。

格式: COMMAND = = = > CHANGE“字符串 1”“字符串 2” [范围选项]

功能: 执行该命令的结果是将文件中的所有形如“字符串 1”的字符串都换成“字符串 2”, “范围选项”同上。

注意 FIND 与 CHANGE 命令可以用于日志、输出和编辑这三个窗口。而且是在窗口的命令行上输入。

§ 1.5 实 例

为了逐步了解 SAS 操作的过程, 下面给出一个用 SAS 处理的方差分析问题。例中数据源于南京农业大学主编的《田间试验和统计方法》一书表 5.9 (P102)。

例 1.1 研究 6 种氮肥施用法对小麦的效应, 每种施肥法种 5 盆小麦, 完全随机设计, 测得含氮量见表 1.1。

表 1.1 含氮量数据表

处 理	1	2	3	4	5	6
	2.9	4.0	2.6	0.5	4.6	4.0
	2.3	3.8	3.2	0.8	4.6	3.3
	2.2	3.8	3.4	0.7	4.4	3.7
	2.5	3.6	3.4	0.8	4.4	3.5
	2.7	3.6	3.0	0.5	4.4	3.7

SAS 程序:

```
DATA eg15;
DO rep= 1 TO 5;
    DO treat= 1 TO 6;
        INPUT x @@;
        OUTPUT;
    END;
END;
CARDS;
2.9 4.0 2.6 0.5 4.6 4.0
2.3 3.8 3.2 0.8 4.6 3.3
2.2 3.8 3.4 0.7 4.4 3.7
2.5 3.6 3.4 0.8 4.4 3.5
2.7 3.6 3.0 0.5 4.4 3.7
;
PROC ANOVA;
CLASS treat rep;
MODEL x = treat;
MEANS treat/t;
RUN;
```

SAS 处理后的结果见表 1.2。

本例的上机过程如下:

启动 SAS 后, 在编辑窗口下, 逐行将 SAS 程序录入。确认无误后, 在该窗口的命令行上键入 SUBMIT 命令, SAS 即开始执行该程序, 稍候片刻, 在 OUTPUT 窗口出现如上的输出结果表。由于窗口行数的限制, 用户可以用 Page Up 和 Page Dn 翻阅。

需要指出的是当键入 SUBMIT 命令 SAS 运行后, OUTPUT 窗口无显示并且在日志窗口出现出错提示 (红色字符) 时, 用户应当返回编辑窗口, 在命令行键入 RECALL 命令, 使录入的 SAS 程序再现, 而后改正错误, 再重新用 SUBMIT 命令提交系统执行。

表 1.2 例 1.1 的输出结果

Analysis of Variance Procedure					
Class Level Information					
Class	Levels	Values			
TREAT	6	1	2	3	4 5 6
REP	5	1	2	3	4 5
Number of observations in data set = 30					
Analysis of Variance Procedure					
Dependent Variable: X					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	44.46300000	8.89260000	164.17	0.0001
Error	24	1.30000000	0.05416667		
Corrected Total	29	45.76300000			
	R-Square	C.V.	Root MSE	X Mean	
	0.971593	7.681100	0.232737	3.03000000	
Source	DF	Anova SS	Mean Square	F Value	Pr > F
TREAT	5	44.46300000	8.89260000	164.17	0.0001
Analysis of Variance Procedure					
T tests (LSD) for variable: X					
NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.					
Alpha = 0.05 df = 24 MSE = 0.054167					
Critical Value of T = 2.06					
Least Significant Difference = 0.3038					
Means with the same letter are not significantly different.					
T Grouping	Mean	N TREAT			
A	4.480	5	5		
B	3.760	5	2		
B	3.640	5	6		
C	3.120	5	3		
D	2.520	5	1		
E	0.660	5	4		

第二章 SAS 语言基本概念

§ 2.1 SAS 语言概述

SAS 提供了一种完善的编程语言。类似于多数计算机高级语言，SAS 用户只需要熟悉其命令、语句及简单的语法规则就可以做数据管理和分析处理工作。因此，掌握 SAS 编程技术是学习 SAS 的关键环节。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调用，用户仅需要指出过程名及其必要的参数。这一特点使得 SAS 编程十分简单。

SAS 程序是 SAS 语句的有序集合。本书的以后章节中将逐一介绍 SAS 的主要语句。如果以程序段的功能将 SAS 程序划分成多个模块的话，那么，这些模块分为两类：数据准备模块——数据步（DATA Step）和数据分析处理模块——过程步（PROC Step）。在一份 SAS 程序中可能有多个数据步和多个过程步，通常的情形为一个数据步和一个过程步。

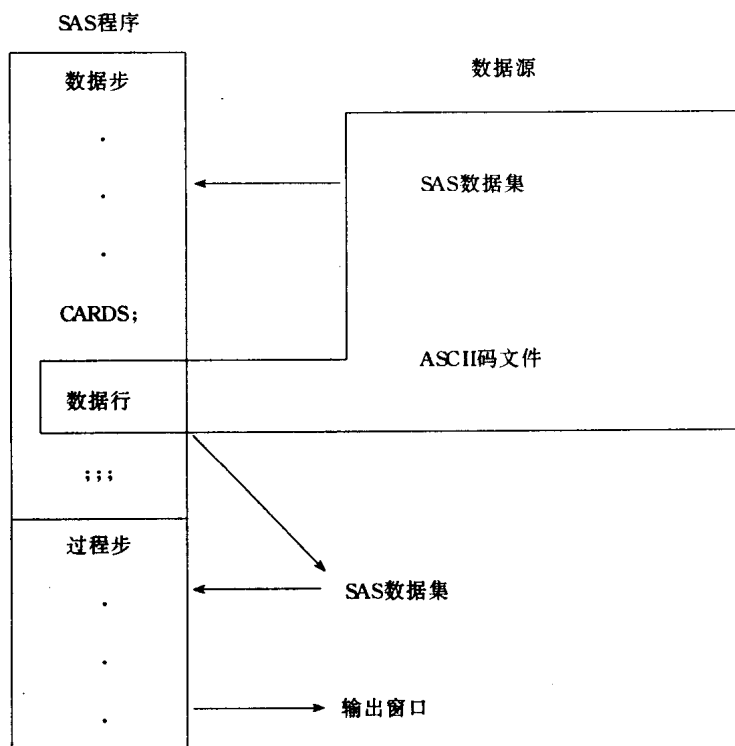


图 2.1 SAS 程序中的数据流向示意图