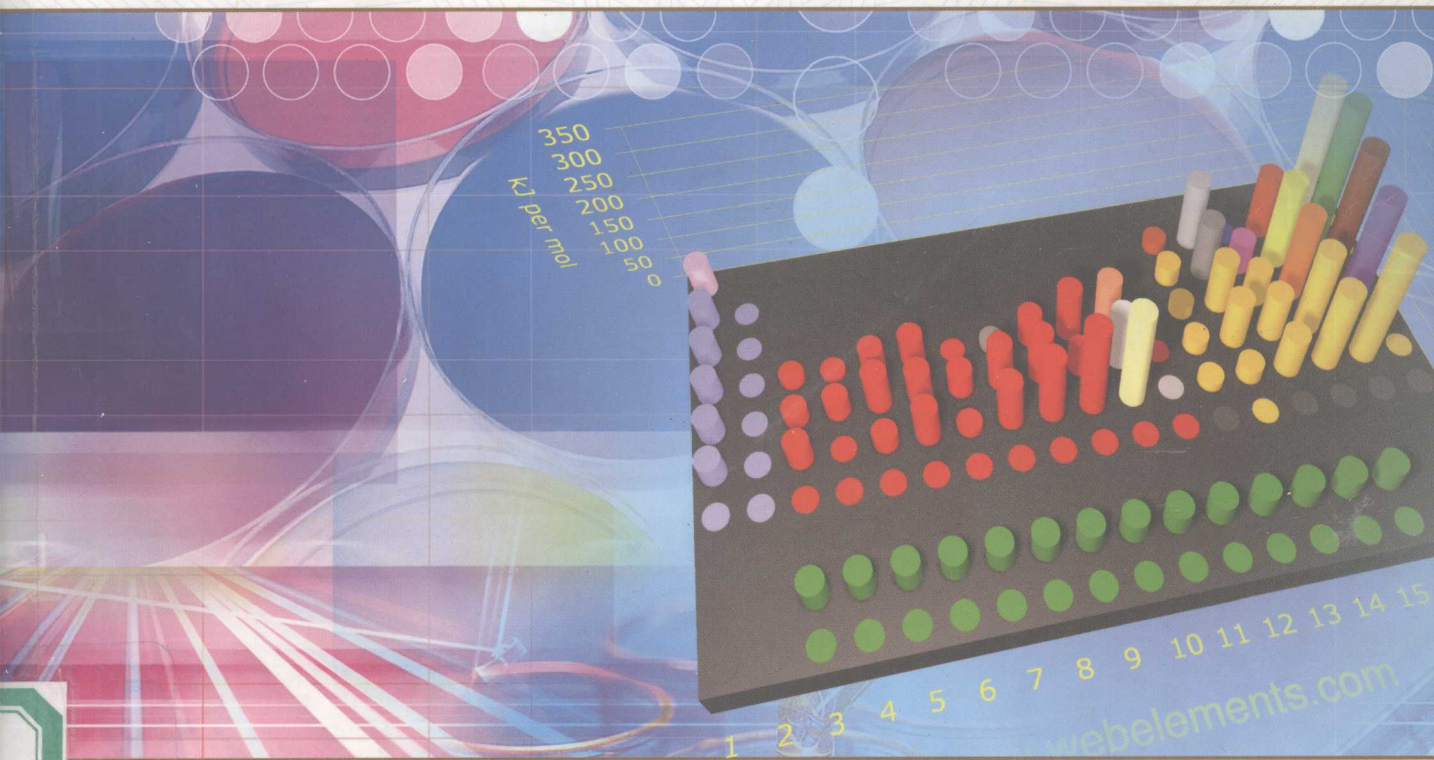




教育科学“十五”国家规划课题研究成果

SAS 统计分析

主编 沈其君



高等教育出版社
HIGHER EDUCATION PRESS

教育科学“十五”国家规划课题研究成果

SAS 统计分析

主编 沈其君

高等教育出版社

内容提要

本书是教育科学“十五”国家规划课题研究成果。

本书以 Windows 操作系统下的 8.1 版 SAS 软件为基础,以目前医学院校本科生和研究生使用的《医学统计学》、《卫生统计学》和《多元统计分析》教材中的统计方法为基本内容介绍 SAS 程序的编写与结果解释。本书的特点是以统计方法为主线介绍 SAS 程序的编写,并将统计方法与 SAS 软件有机结合;每章在简要介绍各种统计方法的基础上重点叙述 SAS 程序各个过程及选项的应用和结果解释。内容主要包括 SAS 系统简介、定量资料的统计描述与推断(包括各种设计下方差分析、协方差分析和重复测量资料的统计方法)、分类资料的统计分析方法、秩和检验、相关与回归、常用的多元统计分析方法、生存分析、logistic 回归、Poisson 回归、试验设计、SAS 数据集和探索性数据分析等共二十一章。

本书可作为高等医药院校各专业研究生和本科生的教材,也可作为非医药类专业本科生和研究生学习 SAS 统计软件的教材,同时适用于有关科研人员参考。

图书在版编目(CIP)数据

SAS 统计分析/沈其君主编. —北京:高等教育出版社,
2005.8

ISBN 7-04-016615-1

I. S... II. 沈... III. 统计分析 - 应用软件, SAS -
高等学校 - 教材 IV. C812

中国版本图书馆 CIP 数据核字(2005)第 065685 号

| | | | |
|------|----------------|------|---|
| 出版发行 | 高等教育出版社 | 购书热线 | 010-58581118 |
| 社 址 | 北京市西城区德外大街 4 号 | 免费咨询 | 800-810-0598 |
| 邮政编码 | 100011 | 网 址 | http://www.hep.edu.cn |
| 总 机 | 010-58581000 | | http://www.hep.com.cn |
| 经 销 | 北京蓝色畅想图书发行有限公司 | 网上订购 | http://www.landaco.com |
| 印 刷 | 国防工业出版社印刷厂 | | http://www.landaco.com.cn |
| 开 本 | 850×1168 1/16 | 版 次 | 2005 年 8 月第 1 版 |
| 印 张 | 23.75 | 印 次 | 2005 年 8 月第 1 次印刷 |
| 字 数 | 580 000 | 定 价 | 27.30 元 |

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 16615-00

SAS 统计分析编写人员

主 编 沈其君(东南大学,宁波大学)

副 主 编 贺 佳(第二军医大学)

李晓松(四川大学)

编 委 (按姓氏笔画顺序)

王 彤(山西医科大学)

尹 平(华中科技大学)

刘 沛(东南大学)

毕育学(西安交通大学)

李晓松(四川大学)

余红梅(山西医科大学)

沈其君(东南大学,宁波大学)

沈 毅(浙江大学)

闵 捷(东南大学)

陈长生(第四军医大学)

贺 佳(第二军医大学)

曹 阳(第二军医大学)

曾 庆(重庆医科大学)

学术秘书 陈炳为

前 言

本教材是教育科学“十五”国家规划课题研究成果。

SAS 统计软件是目前世界上应用最广泛的统计软件之一,许多向政府和企业提交的试验设计和统计分析文本都是用 SAS 统计软件分析和处理后得到的结果。我国高等医药类院校各专业的本科生和研究生的医学统计学教学中自 20 世纪 90 年代以来相继使用 SAS 软件或单独开设“SAS 统计软件及其应用”的课程。目前该课程已是医药类各专业学生的一门重要专业基础课。

东南大学公共卫生学院卫生统计教研室在 20 世纪 90 年代初为研究生开设“SAS 统计软件”课程,其后在本科生中也开设了这门课程,我尊敬的老师陈启光教授为这门课程教材体系与结构的建立和教学经验的积累奠定了基础,并在他的指导下于 2001 年出版了《SAS 统计分析》教材。这次 9 所高等医学院校的统计教师在研讨和总结各校教学经验的基础上,重新编写《SAS 统计分析》教材,是前述教材和讲义的继承与发展,在结构与内容上均作了较大的调整。

本教材以 Windows 操作系统下的 8.1 版 SAS 软件为基础,以目前医药院校使用的《医学统计学》、《卫生统计学》和《多元统计分析》等教材中统计方法为基本内容,介绍 SAS 软件的程序编写和结果解释。对学生已学过的基本统计方法,采用复习方式,提纲挈领地进行回顾,重点叙述其在 SAS 软件中的实现;对较深入的内容如协方差分析、重复测量资料的统计分析方法、多元统计分析方法、logistic 回归、Poisson 回归和生存分析等,在统计方法和 SAS 程序两方面都作了详细叙述。这样,既便于教师根据教学需要选择相应章节进行教学,也便于有一定统计学知识的学生和科研工作者自学、提高和参考。

本书的特点是以统计方法为主线介绍 SAS 软件,每章在简要介绍统计方法的基础上结合实例叙述 SAS 各个过程、选项的应用及结果的详细解释,并注重统计方法与选项和输出结果的有机结合;为使教材更加符合教学规律,考虑到学生的可接受性和 SAS 软件的特点,将 SAS 数据集一章移到最后作为总结和深入介绍,避免学生在开始学习 SAS 时就接触复杂的 SAS 数据集问题。考虑到 SAS 8.1 版本已有统计分析的菜单功能,因此在第二十一章介绍了数据探索性分析与 SAS/Insight 模块。

本书可作为医药类院校各专业本科生与研究生的教材,也可作为非医药类高等院校的学生学习 SAS 软件的教材,同时也适用于高等院校和科研机构的科研人员参考。

本教材的编写和出版得到高等教育出版社、东南大学、宁波大学领导和有关各参编院校领导的大力支持。感谢《医学统计学》(供研究生用,人民卫生出版社出版,2002.8)和《医学统计学》(高等教育出版社,2004.1)教材的主编孙振球教授与徐勇勇教授允许本教材作者引用两书的统计实例。感谢所有参加编写的老师和高等教育出版社的编辑为本教材付出的辛勤劳动。陈炳为、张宏、蒋敏、孙爱珊、贾斌和李运明等老师和研究生参与了本书的编写工作。共同的切磋和创作使我们增进了友谊,其间也留下了许多美好的回忆。

由于我们的知识与经验有限,教材中肯定存在许多不足之处,诚恳希望广大师生提出进一步的修改意见。

沈其君

2004年9月于宁波

郑重声明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人给予严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

反盗版举报电话：(010) 58581897/58581896/58581879

传 真：(010) 82086060

E - mail: dd@hep.com.cn

通信地址：北京市西城区德外大街4号

高等教育出版社打击盗版办公室

邮 编：100011

购书请拨打电话：(010)58581118

| | |
|------|-----|
| 策划编辑 | 李蕊 |
| 责任编辑 | 高尚华 |
| 封面设计 | 王凌波 |
| 责任绘图 | 黄建英 |
| 版式设计 | 范晓红 |
| 责任校对 | 王超 |
| 责任印制 | 杨明 |

目 录

| | | | |
|---------------------------------------|----|----------------------------|-----|
| 第一章 SAS 系统简介 | 1 | UNIVARIATE 过程 | 71 |
| 第一节 SAS 系统的特点 | 1 | 第二节 两个独立样本秩和检验与 | |
| 第二节 SAS 显示管理系统 | 3 | NPARIWAY 过程 | 74 |
| 第三节 一个简单的 SAS 程序 | 7 | 第三节 多个独立样本秩和检验与 | |
| 第四节 程序修改、调用和存储 | 10 | NPARIWAY 过程 | 78 |
| 练习题 | 12 | 第四节 配伍组设计资料秩和检验与 FREQ | |
| 第二章 定量资料的简单统计描述 | 13 | 过程 | 80 |
| 第一节 定量资料常用统计指标 | 13 | 练习题 | 82 |
| 第二节 MEANS 过程 | 14 | 第六章 方差分析 | 84 |
| 第三节 UNIVARIATE 过程 | 18 | 第一节 方差分析的基本思想、应用条件及 | |
| 第四节 CHART 过程 | 26 | 常见设计类型 | 84 |
| 第五节 PLOT 过程 | 29 | 第二节 多样本的正态性检验和方差齐性 | |
| 第六节 SAS/GRAPH 模块 | 31 | 检验 | 84 |
| 练习题 | 35 | 第三节 完全随机设计资料的方差分析(含 | |
| 第三章 t 检验与 u 检验 | 37 | 多重比较) | 88 |
| 第一节 t 检验与 u 检验方法简介 | 37 | 第四节 随机区组设计资料方差分析 | 91 |
| 第二节 UNIVARIATE 过程与 t 检验 | 39 | 第五节 拉丁方设计资料的分析 | 93 |
| 第三节 TTEST 过程 | 42 | 第六节 交叉试验设计资料的方差分析 | 95 |
| 练习题 | 48 | 第七节 GLM 过程与 ANOVA 过程 | 98 |
| 第四章 分类资料的统计方法 | 50 | 练习题 | 102 |
| 第一节 分类资料统计描述与统计推断 | | 第七章 多因素试验方差分析 | 104 |
| 简介 | 50 | 第一节 析因试验设计的方差分析 | 104 |
| 第二节 四格表资料的 χ^2 检验与 FREQ | | 第二节 正交试验设计资料的方差分析 | 107 |
| 过程 | 50 | 第三节 裂区试验设计资料的方差分析 | 111 |
| 第三节 行 \times 列表资料 χ^2 检验与 FREQ | | 第四节 系统分组(嵌套)设计资料的方差 | |
| 过程 | 56 | 分析 | 114 |
| 第四节 病例-对照与队列研究资料的 | | 练习题 | 118 |
| FREQ 过程分析 | 63 | 第八章 协方差分析 | 120 |
| 第五节 二项分布与 PROBBNML() 函数 | 66 | 第一节 协方差分析的基本思想 | 120 |
| 第六节 Poisson 分布与 POISSON() 函数 | 67 | 第二节 完全随机设计资料的协方差 | |
| 练习题 | 68 | 分析 | 123 |
| 第五章 秩和检验 | 71 | 第三节 随机区组设计资料的协方差 | |
| 第一节 配对设计资料的秩检验与 | | 分析 | 130 |

| | | | |
|---------------------------|-----|------------------------------|-----|
| 练习题 | 134 | 过程 | 235 |
| 第九章 多元方差分析 | 136 | 第四节 Cox 回归模型与 PHREG 过程 | 240 |
| 第一节 多变量组间比较的统计方法 | | 练习题 | 244 |
| 概述 | 136 | 第十五章 判别分析 | 246 |
| 第二节 多变量组间比较的 GLM 过程 | 138 | 第一节 判别分析的基本概念 | 246 |
| 练习题 | 144 | 第二节 两类 Fisher 判别与 DISCRIM | |
| 第十章 重复测量资料的统计分析 | 145 | 过程 | 247 |
| 第一节 重复测量设计的基本概念 | 145 | 第三节 定量资料的 Bayes 多类判别与 | |
| 第二节 重复测量资料的方差分析与 GLM | | DISCRIM 过程 | 251 |
| 过程 | 147 | 第四节 逐步判别分析 | 256 |
| 第三节 重复测量资料的轮廓分析 | 157 | 练习题 | 260 |
| 练习题 | 163 | 第十六章 主成分分析与因子分析 | 262 |
| 第十一章 相关与回归分析 | 165 | 第一节 主成分分析 | 262 |
| 第一节 相关分析与 CORR 过程 | 165 | 第二节 PRINCOMP 过程 | 263 |
| 第二节 直线回归与 REG 过程 | 169 | 第三节 因子分析 | 265 |
| 第三节 多元线性回归与自变量筛选 | 175 | 第四节 FACTOR 过程 | 268 |
| 第四节 非线性回归与 NLIN 过程 | 183 | 练习题 | 272 |
| 练习题 | 188 | 第十七章 聚类分析 | 275 |
| 第十二章 logistic 回归模型 | 191 | 第一节 聚类分析简介 | 275 |
| 第一节 logistic 回归模型简介 | 191 | 第二节 样品系统聚类分析与 CLUSTER | |
| 第二节 两分类反应变量的 LOGISTIC | | 过程 | 276 |
| 过程 | 197 | 第三节 指标聚类分析与 VARCLUS | |
| 第三节 有序反应变量的 LOGISTIC | | 过程 | 280 |
| 过程 | 205 | 第四节 动态样品聚类与 FASTCLUS | |
| 第四节 无序多分类反应变量 logistic 模型 | | 过程 | 286 |
| 与 CATMOD 过程 | 207 | 练习题 | 291 |
| 第五节 匹配资料的条件 logistic 回归 | | 第十八章 典则相关分析 | 295 |
| 模型 | 210 | 第一节 典则相关分析概述 | 295 |
| 练习题 | 215 | 第二节 CANCELL 过程 | 297 |
| 第十三章 Poisson 回归模型 | 219 | 练习题 | 302 |
| 第一节 Poisson 回归模型简介 | 219 | 第十九章 试验设计的统计过程 | 304 |
| 第二节 Poisson 回归模型的 GENMOD | | 第一节 PLAN 过程 | 304 |
| 过程 | 222 | 第二节 常见试验设计方法的样本量 | |
| 练习题 | 228 | 估计 | 310 |
| 第十四章 生存分析 | 230 | 练习题 | 315 |
| 第一节 生存分析的基本概念 | 230 | 第二十章 数据探索性分析与 SAS/Insight | |
| 第二节 生存率估计与组间比较 | 231 | 模块 | 316 |
| 第三节 生存率估计与组间比较的 LIFETEST | | 第一节 探索性分析的基本概念 | 316 |

| | | | | | |
|-------|-----------------------|-----|---------|-------------------------|-----|
| √ 第二节 | SAS/Insight 模块简介 | 316 | 第二十一章 | SAS 数据集 | 332 |
| 第三节 | 奇异值的探索性分析——盒型图 | 319 | 第一节 | 用 INPUT 和 CARDS 语句建立数据集 | 332 |
| 第四节 | 数据分布的探索性分析——直方图和正态性检验 | 321 | 第二节 | 建立数据集的另外几种途径 | 335 |
| 第五节 | 方差齐性检验的探索性分析——菱形图 | 324 | 第三节 | 建立永久性数据集 | 343 |
| 第六节 | 直线相关和回归的探索性分析——散点图 | 325 | 第四节 | 输入与输出格式 | 344 |
| 第七节 | 曲线拟合中模型选择的探索性分析方法 | 326 | 第五节 | SAS 数据集的编辑、整理与浏览 | 346 |
| 第八节 | 分类数据的探索性分析——马赛克图 | 329 | 第六节 | SAS 数据集的排序、连接与合并 | 350 |
| 练习题 | | 330 | 第七节 | SAS 运算符与 SAS 常用函数 | 354 |
| | | | 练习题 | | 357 |
| | | | 英汉名词对照表 | | 359 |
| | | | 主要参考文献 | | 368 |

第一章 SAS 系统简介

SAS(statistical analysis system)意为统计分析系统,是由 SAS 公司研发的模块化、集成化的应用软件系统,具备完备的数据访问、数据管理、数据分析和数据呈现功能。该系统于 1964 年由美国北卡罗莱纳州立大学开始研制,1976 年成立 SAS 软件研究所(SAS Institute Inc.)并推出该软件,1985 年推出 SAS/PC(6.02)版本,目前最新版本为 Windows 操作系统下的 9.1 版。SAS 软件在世界范围内被广泛应用于政府、科研、教学、医疗卫生、工业和商业等不同领域的的数据管理与处理。全球有 42 000 多家企业,包括财富 500 强中 90% 的企业都在使用 SAS 商业智能解决方案。SAS 是将世界领先的数据仓库技术、数据分析技术和传统的商业智能应用完全整合。在统计分析方面,SAS 软件也是世界范围内最权威的统计分析软件,许多向政府或企业提交的试验设计和统计分析文本都用 SAS 软件分析和处理数据。本书主要介绍 8.1 版 SAS 软件的数据分析与数据管理功能。

本章主要介绍 SAS 软件的特点、启动与退出、显示管理系统、SAS 程序的基本结构以及程序的调用、修改和存储。

第一节 SAS 系统的特点

一、SAS 系统组成

SAS 系统是由三十多个专用程序模块组成的大型集成应用软件系统,有四大部分构成:

1. SAS 数据库部分。

主要由 BASE 模块提供常见的数据库管理功能,所有其他模块必须与之结合起来使用。

2. SAS 分析核心。

是 SAS 系统的灵魂,具有对数据进行分析与决策的功能,这部分主要模块有 STAT、QC、ETS、IML、INSIGHT、PH-CLINICAL 等十大模块。

3. SAS 开发及呈现工具。

主要有 AF、EIS、与 GRAPH 模块提供便捷的面向对象的开发工具,支持客户/服务器的应用开发。

4. SAS 分布处理模式及数据仓库设计。

主要有 ACCESS、CONNECT、SHARE 和 WA 模块。

二、SAS 系统的基本功能

SAS 系统主要完成以数据为中心的四大功能。

1. 数据管理功能。

SAS 系统可手工输入数据值,然后把数据组成 SAS 数据集;它能够方便地将流行的数据格

式,如 Excel、Lotus、DBF、TXT 等转为 SAS 数据集。SAS 数据可用临时数据集或永久数据集两种形式储存。SAS 系统提供完备的 SAS 语句和函数用于数据加工处理,并具有对数据进行编辑、整理、连接和更新等功能。

2. 基础统计计算功能。

计算简单的描述统计量,多达 39 项;对数据进行标准化处理、编程及计算有关统计量;生成并分析列联表;计算概率分布函数、分位数及产生随机数。

3. 报表生成和图形显示功能。

固定格式报表、用户自定义报表;绘制二维与三维的基本的统计图,如直方图(水平或垂直)、圆形图、散点图、曲线图、时间序列图等。

4. 统计分析功能。

SAS 的核心功能。提供包括方差分析、相关与回归分析、属性数据分析、判别与聚类分析、非参数分析、生存分析、时间序列分析、质量控制、运筹学等统计分析过程。

三、SAS 系统的模块组成

与统计分析和数据管理有关的 SAS 主要常用模块及功能为:

1. BASE 模块。

进行数据管理和数据加工、处理。

2. STAT 模块。

完整的统计分析模块,可对数据进行单因素、多因素分析。

3. GRAPH 模块。

优异的图形软件包,能完成多种绘图功能。

4. INSIGHT 模块。

进行数据探索性分析。

5. ETS 模块。

用于计量经济与时间序列分析的专用软件,包括时间序列时域分析和谱域分析。如逐步自回归、指数平滑、序列相关校正分析、分布滞后回归、ARIMA 模型、谱分析等。

6. OR 模块。

用于运筹学和工程管理的专用软件。提供全面的运筹方法,是一个优异的决策支持工具。包括线性规划、整数规划、混合整数规划和非线性规划。

7. QC 模块。

用于质量控制的专用软件。提供全面质量管理工具,进行试验设计、流程控制图及执行流程的能力分析等。

8. IML 模块。

提供功能强大的矩阵运算,用户可根据需要编写程序,完成特定的矩阵运算。

一、SAS 的启动与退出

SAS 的启动常用两种方法:一是快捷方式。即按 Windows 下程序快捷方式的建立步骤,通过资源管理器,把 SAS 目录中的 SAS.EXE 拖放至桌面,形成快捷方式。用鼠标左键双击这一快捷方式,即可启动 SAS。

二是菜单方式。在“开始”=>“程序”菜单的“The SAS System”程序组中,选定其中的“The SAS System for Windows v8”项,即可将 SAS 启动。启动画面如图 1-1。

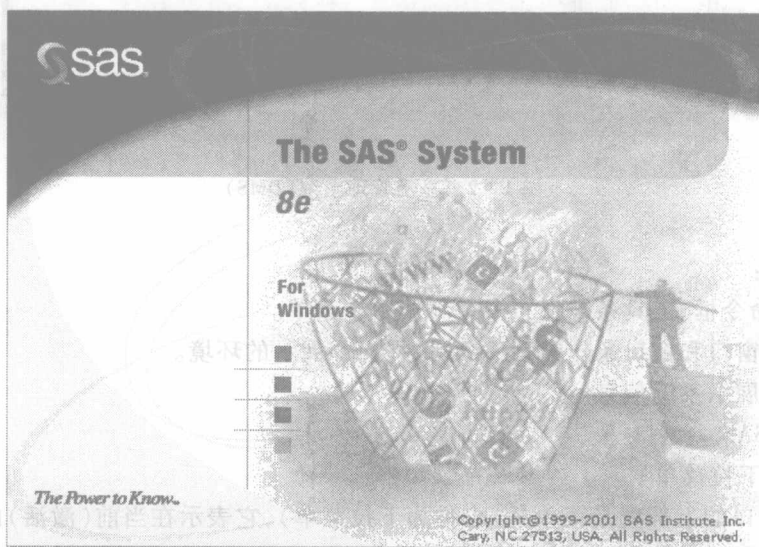


图 1-1 启动画面

然后,系统自动进入 SAS 的显示管理系统 (display manager system),见图 1-2。可以看出这是一个典型的 Windows 平台下的程序窗口,顶部依次为标题栏、菜单栏、命令与工具栏。

用鼠标左键单击标题栏右上角的关闭按钮 (如图 1-2),并在确认框中点击确认按钮,就可退出 SAS,返回 Windows 操作系统。

二、浏览 SAS 主窗口与发布命令

启动 SAS 后即展示如图 1-2 的 SAS 主窗口,它是交互方式下运行 SAS 的工作界面,称为应用工作空间 (application work space, AWS),也是 SAS 的显示管理界面 (系统) (display management system, DMS)。

在 SAS 系统中要完成各项应用任务,可以通过提交 SAS 程序或者通过 SAS 的一些子菜单系统提交任务来实现,对交互操作中一些简单的任务和动作,可以在交互环境下发布命令来实现。

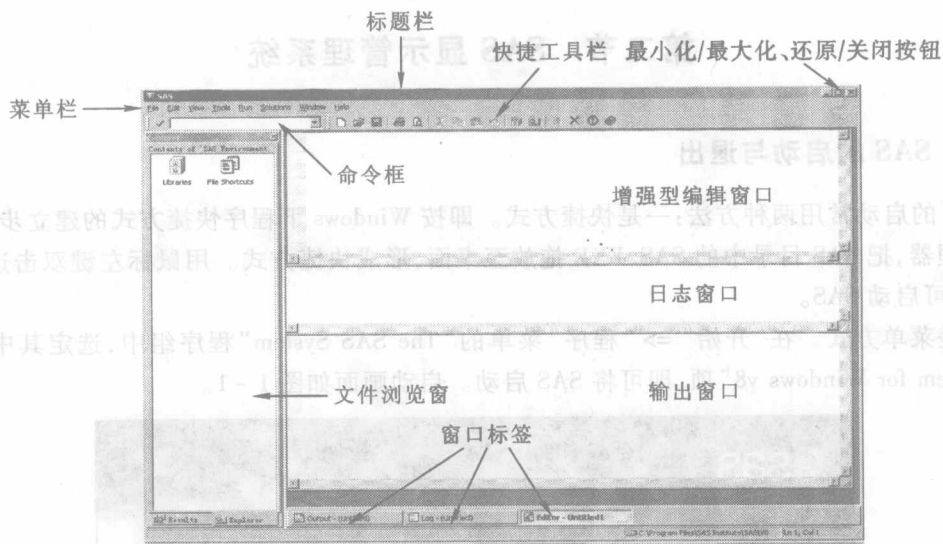


图 1-2 显示管理系统(DMS)

SAS 主窗口提供：

- (1) 菜单、命令框、工具条来发布命令和调用子菜单。
- (2) 多个子窗口和窗口条作为程序编写、提交和展示的环境。
- (3) 状态条展示系统信息。

(一) 浏览 SAS 主窗口

1. 菜单条(下拉菜单)。

在 SAS 主窗口的上部就是菜单条(也称为下拉菜单),它表示在当前(激活)的子窗口下可进行操作的菜单栏目。点击这些栏目就会伸展出可操作的栏目或子菜单。可以实现文件处理、编辑、查看、工具、窗口、帮助等一般 Windows 程序功能,同时还有运行和分析两个特有菜单。由于在不同的 SAS 子窗口可以实现不同的功能,所以激活不同的子窗口,菜单的内容也会改变,有些不能使用的功能会自动变灰。与菜单条功能相同的是弹出菜单,在当前子窗口点击鼠标右键就可弹出与菜单条相同的菜单。

2. 命令框。

命令框用于键入和提交 SAS 命令。例如键入“wpgm”并回车就可使增强型程序编辑窗口激活,键入“bye”就可退出 SAS 返回 Windows 操作系统。按下命令框右边的按钮就可显示其保留的最近曾提交的若干条命令并点击左边的钩号按钮就完成了提交命令。

3. 工具条。

在工具条中排列了若干按钮,每个按钮为一个常用命令,点击工具条中的按钮,用户可以快速、直接发布命令。当光标移至各个按钮时,按钮边上会弹出一个标签显示按钮的简单功能,同时在下方的信息条上显示按钮功能的说明。

工具条中的按钮个数、功能和采用的图标可以由用户自行设定。在菜单中选 Tool => Customize(客户化)。然后在弹出的对话框中进行设定。工具条常用按钮的功能如下：

□ 新建, 打开, 保存, 打印, 打印预览, 剪切, 复制, 粘贴, 撤销键入, 新建逻辑库, SAS 资源管理器, 提交运行, 全部清除, 中止, 帮助。

4. 子窗口工作空间。

在 SAS 主菜单的中部是一个供打开 SAS 子窗口的工作空间。常见的子窗口有浏览器 (explorer) 窗口、结果 (results) 窗口、增强型编辑器 (enhanced editor) 窗口、日志 (log) 窗口和输出 (output) 窗口。其中浏览器与结果窗口固定在工作空间的左侧, 其它三个窗口通常只有一个当前窗口, 通过点击窗口标签可以激活某个窗口 (详见后述)。

5. 窗口条与状态条。

在 SAS 子窗口工作空间的下方是窗口条, 包含了所有打开的子窗口标签。对非激活子窗口, 点击可激活该窗口为当前窗口, 对激活窗口, 点击标签可使它最小化而让其它窗口为当前窗口。

在 SAS 主窗口的最下方是状态条, 它的左部是信息显示区, 中部是运行 SAS 的工作文件夹名称, 右部是增强型编辑器光标所在的位置。

(二) 提交命令

SAS 的不同窗口有不同的特征可以帮助你工作, 要在各个子窗口实现各种功能以及实现窗口间的切换都是通过 SAS 的应用工作空间中发布命令来实现。通常可有如下几种方式来发布命令:

1. 在命令框中键入命令, 并按回车。
2. 使用菜单或弹出菜单。
3. 使用工具条。
4. 使用功能键。

各功能键的功能以及修改可在命令框中键入“Keys”, 并回车。在弹出的 Keys 窗口中查阅和修改。如按 F1 键就会弹出关于 SAS 系统的帮助菜单或其它与当前窗口相关的帮助菜单窗口。

三、SAS 的子窗口及其切换

要充分利用 SAS 的数据访问、管理、分析和展示功能的最常用做法是提交 SAS 程序。SAS 各个子窗口提供了如下功能: (1) 浏览与管理 SAS 文件, (2) 编写或调入 SAS 程序, (3) 提交 SAS 程序, (4) 展示程序运行的信息与结果。

SAS 系统是 Windows 操作系统的一个应用软件, 主窗口和每个子窗口都保持了 Windows 下应用窗口的一些共有功能。例如窗口间的复制、粘贴和相同的命令或热键, 窗口右上角的最小化、最大化和关闭按钮, 窗口的下拉菜单、弹出菜单和工具条等。此外, 在 SAS 编辑器、日志窗口和输出窗口都可用“clear”命令将窗口内容清空。

以下介绍几个主要子窗口的功能与特点。

1. 浏览器窗口。

SAS 系统中浏览器的作用类似于 Windows 操作系统下浏览器的作用。可以查看和管理 SAS 文件库和存放在 SAS 文件库中由 SAS 系统创建的数据文件和其它类型的 SAS 文件, 也可建立对非 SAS 文件的快捷方式。用这一窗口还可以指引建立新的 SAS 文件库和 SAS 数据文件。

为了激活左侧固定位置上的浏览器, 可以通过: 在菜单中选 View => Contents Only, 或在菜

单中选 Windows => Explorer,或点击窗口条左侧的浏览器的标签来实现。

此外,若要打开一个新的包含 SAS 文件库目录树的浏览器,可以通过:键入命令“explorer”并回车,或在菜单选 View => Explorer,或在工具栏点击相应图标来完成。

在浏览器窗口,可在 SAS 文件库和各种 SAS 文件上双击鼠标左键打开这些对象。在这些对象上点击右键可以在弹出的菜单中选择文件和文件库的各种操作,实现打开、移动、复制、删除和重命名文件等大部分管理文件的任务和浏览库及文件的属性。

2. 结果窗口。

结果窗口帮助您浏览和管理所提交 SAS 程序的输出结果,而非输出(output)窗口。在这个窗口中将整个 SAS 进程中提交 SAS 程序的输出结果依次排列为一个棵树,每次提交 SAS 程序后的输出结果都表示一个结点。展示这个结点可进一步显示表示不同内容的子结点。将鼠标右键点击每个结点,就可对这个输出结果进行察看、存储、打印或删除等操作。

缺省时结果窗口在浏览器背后,它在提交 SAS 程序创建输出以前一直是空白,在创建输出结果后,它移到显示的前台(窗口左侧)。使用窗口条中的标签或发布命令“odsresult”都可激活该窗口。

3. 增强型编辑器窗口。

在 Window 操作系统下,增强型编辑器窗口是 SAS 系统缺省提供的程序编辑器窗口。具有一般的文本输入和编辑功能。包括:

- (1) 输入和编辑文本。包括用 SAS 语言编写 SAS 程序、编辑 SAS 运行结果等。
- (2) 提交 SAS 程序。
- (3) 拷贝一个外部文件到编辑器窗口,或拷贝编辑器窗口内容到外部文件上。
- (4) 拷贝一个目录条目到编辑器窗口,或拷贝编辑器窗口内容到一个目录条目。

除此之外,还提供许多有用的编辑功能,它包括:

- (1) 对 SAS 程序不同语句用不同颜色显示并进行词法检查。
- (2) 程序段的展开和收缩功能,便于从宏观上了解程序的结构。
- (3) 程序输入时自动缩进功能。
- (4) 可对输入字符串自定义宏。
- (5) 支持键盘快捷方式(将 Alt 键或 Shift 键与其它键同时使用)。

点击窗口条中的相应标签或键入命令“wpgm”就可激活已打开的增强型编辑器窗口。

在菜单中选 View => Enhanced Editor 或在任何增强型编辑器窗口的菜单中选 File => New ,就可打开一个新的增强型编辑器窗口供直接键入或读入和提交 SAS 程序。

4. 程序编辑器窗口与 SAS 记事板。

除了增强型编辑器外, SAS 还提供了程序编辑器和 SAS 记事板。它们有与其他文本编辑器类似的文本编辑功能,都可以作为读入、编辑和提交 SAS 程序的窗口。程序编辑器是在 SAS 版本 6 中的各个操作平台中已被广泛使用的程序编辑器。在一个 SAS 进程中只能打开一个程序编辑器窗口,但在一个 SAS 进程中可以打开多个 SAS 记事板。

键入命令“pgm”,或在菜单中选 View => Program Editor,和键入“notepad”可分别打开程序编辑器和 SAS 记事板。

5. 日志窗口。

日志窗口随时显示 SAS 进程和提交的 SAS 程序运行信息。在进入 SAS 主窗口后就可从日志窗口看到启动 SAS 的信息。运行一个 SAS 程序后又可看到关于这个程序运行的信息。运行信息用不同颜色加以标识。提交执行的 SAS 语句(程序行)以黑色字表示;系统或程序运行的一些常规信息,诸如文件信息、程序运行时间(提示行)用蓝色字表示,并以 NOTE 引导;程序运行中遇到的一些问题,如系统能够自动更正的小的拼写错误;软件授权期限即将到期等(警告行)以绿色字表示,并以 WARNING 引导,这部分需要仔细阅读以判断是否会影响程序运行结果的准确性,一般情况下不会影响结果;错误提示即程序中有错误的位置以及可能的原因以红色字表示,并以 ERROR 引导。

发布命令“log”或使用热键 F6,或在菜单中选 View => Log,或点击窗口条中相应标签可激活日志窗口。

6. 输出窗口。

将 SAS 程序执行后,相继产生的结果追加到该窗口的内容中,是 SAS 运行结果的显示窗口。使用者可在该窗口对输出结果进行浏览。

缺省时该窗口位于其他窗口的背后。当程序有输出时,输出窗口自动移到显示的前台。发布命令“output”或使用热键 F7,或在菜单中选 View => Output,或点击窗口条中的相应标签可激活输出窗口。

若在 SAS 进程中提交了多个 SAS 程序,其每次的输出结果都按页依次存放,可在输出窗口依次浏览。为了方便地查阅某一次提交 SAS 程序的输出结果,可通过点击结果窗口的相应结点直接进行浏览。

7. 窗口切换。

在 SAS 操作环境下,要切换或激活的窗口除了直接点击要激活的窗口外,也可靠发布命令、使用菜单、窗口条或热键来实现。最方便的是点击窗口条中的相应窗口的标签。在下拉菜单中选 Window 或 View 进一步选窗口名可使某个窗口激活。

除此以外,显示管理系统还有一些其它窗口,例如

KEYS 窗口 查看及改变功能键(热键)的设置。

LIBNAME 窗口 查看已经存在的 SAS 数据库。

DIR 窗口 查看某个 SAS 数据库的内容。

VAR 窗口 查看 SAS 数据集的有关信息。

OPTIONS 窗口 查看及改变 SAS 的系统设置。

淘金式

第三节 一个简单的 SAS 程序

在应用 SAS 软件对数据作统计分析时,可以通过窗口菜单方式和在增强型编辑窗口编写相应的程序两种方式来实现,SAS 软件的特点是强大而灵活的编程功能。下面通过一个简单的实例来介绍程序的结构、程序编写的基本要求以及结果的解释。

一、程序结构与编写的基本要求

例 1-1 某单位测得 12 名女中学生体重 X_1 (kg)和胸围 X_2 (cm)资料如下: