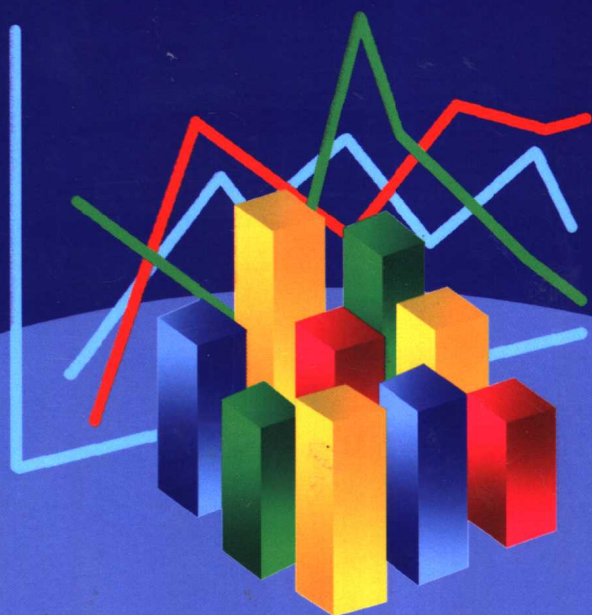




The SAS® System



医学统计学中的 SAS统计分析

SAS Analysis

In Medical Statistics

贺佳 陆健 主编

曹阳 贺宪民 副主编



第二军医大学出版社

医学统计学中的 SAS 统计分析

主 编：贺 佳 陆 健

副主编：曹 阳 贺宪民

编 者：陆 健 贺 佳 曹 阳
贺宪民 于长春 徐 蕾
孙亚林 范思昌 张罗漫
孟 虹 熊林平 马修强
吴 骋

第二军医大学出版社

内 容 简 介

SAS (statistical analysis system) 是全球著名的统计软件, 其数据管理和数据处理功能已在医学、金融、财经、工业、农业、社会科学等领域得到广泛应用, 被誉为权威性的统计软件。本书主要介绍了 SAS 在医学科研中的应用, 结合科研设计方案和医学实际例子, 说明数据集建立、整理和分析的过程, 从编写程序、程序解释、显示结果和结果解释四个方面介绍各种统计方法, 让读者轻松掌握 SAS 处理数据的方法。本书还介绍了数据填充和数据挖掘方法及 SAS 的实现过程, 为统计工作者进行较深入的研究提供参考。

本书可作为研究生《医学统计学》的上机实习教材, 也可以作为广大医学科研工作者自学 SAS 的参考书。

图书在版编目(CIP)数据

医学统计学中的 SAS 统计分析/贺佳, 陆健主编.-上海: 第二军医大学出版社, 2002.8

ISBN 7-081060-247-0

I. 医… II. ①贺…②陆… III. 医学统计-统计分析软件-应用软件, SAS IV. R195.1-39

中国版本图书馆 CIP 数据核字(2002)第 046890 号

医学统计学中的 SAS 统计分析

主编: 贺 佳 陆 健

责任编辑: 高敬泉

第二军医大学出版社出版发行

(上海翔殷路 818 号 邮政编码: 200433)

全国各地新华书店经销

上海第二教育学院印刷厂印刷

开本: 787×1092 1/16 印张: 13.25 字数: 350.0 千字

2002 年 8 月第 1 版 2002 年 8 月第 1 次印刷

印数: 4 500 册

ISBN 7-81060-247-0/R·192

定价: 28.00 元

前言

医学统计学是医学科研工作中必不可少的重要工具，也是医学硕士研究生的公共必修课，备受广大科研工作者的重视。随着计算机技术的发展及统计学软件的开发利用，使许多复杂的统计运算变得简单易行。SAS (statistical analysis system) 软件就是目前较为常用的一种统计软件。

SAS 处理数据的能力非常强大且方法齐全，通过不同的功能模块能够完成几乎所有的统计运算，得到统计结果，满足科研需要，因而它是医学工作者理想的工具。同时，SAS 还可以被当作一个统计平台，许多新的统计理论和方法都可以通过 SAS 来付诸实践。

面对庞大的 SAS 软件，如何充分利用这个工具却不是一件容易的事。一方面，医学工作者缺乏足够的统计学知识；另一方面，SAS 提供的功能十分丰富，挑选合适的方法无疑是大海捞针，而且，正确理解 SAS 的输出结果又是一个令人头疼的问题。为此，我们编写了这本书，希望能让医学工作者通过医学实例，了解医学科研中常用的统计设计方法，掌握 SAS 处理过程，挑选合适的输出结果，循序渐进地充分利用数据处理这一手段。

第二军医大学卫生勤务学系卫生统计学教研室是 SAS 6.12 版的合法用户，在科研、教学工作中使用 SAS 有多年的经历，总结了不少运用 SAS 的经验。我们根据以往的教学实践经验，以 SAS 6.12 版为蓝本编写本书，目的是为了将医学统计学的理论与电脑实践有机地结合起来，使医学工作者能够方便地学习统计学知识，掌握解决实际问题的能力和技巧，提高科研水平。

另外,本书作者中有关人员在世界卫生组织总部工作期间,对 SAS 8 中的数据挖掘和缺失数据填补技术进行了学习和应用。在 SAS 8 中,以往版本的数据挖掘功能得到了进一步的整合和提高,许多过程被综合在同一个操作环境中,界面十分友好,功能更加强大,已经成为统计分析人员和管理决策者必不可少的统计分析利器。SAS 8 中新创建的专门针对缺失数据的统计分析过程是 MI 过程和 MIANALYZE 过程,它们不仅可以单独使用,对数据集中的缺失数据进行填补和做出综合统计推断,还可以与其他常用统计分析过程相结合,帮助统计分析人员充分利用数据资源,并做出更加合理的统计分析。在本书中,将对这些新的功能进行介绍。

本书以卫生部医学专业研究生规划教材《医学统计学》为基础,依据该书中的大部分实例,重点介绍 SAS 实现统计运算的过程和结果解释。在卫生部规划教材中,SAS 统计分析章节内容由我校负责编写。本书既可以作为规划教材的辅助教材,又可以作为广大医学工作者学习 SAS 软件的自学教材。

本书共有 19 章,第一至第三章介绍了 SAS 6.12 的基础知识;第四至第十六章介绍了各种常用统计方法;第十七章介绍统计结果的输出,由于 SAS 计算结果输出较多,如何将输出结果提炼编辑,将在此章进行介绍;第十八、十九章为 SAS 8 中关于缺失数据多重填充和数据挖掘方面的 SAS 处理过程。

由于 SAS 软件自身的设定,使得程序与输出结果中的符号为特定的格式,故本书中的有关内容均保持原有格式排版,未能按编辑要求修改。由于我们水平有限,本书难免存在不足之处,希望大家提出宝贵的意见和建议。

编 者

2002 年 6 月 6 日

目 录

第一章 SAS 概述.....	1	二、新变量的产生.....	21
第一节 SAS 的特点和运行环境.....	1	三、数据集的排序.....	22
一、SAS 的基本功能.....	1	四、数据集的连接.....	22
二、SAS 的功能模块.....	2	五、数据集的合并.....	24
三、SAS 的运行环境.....	3	六、数据集的输出.....	25
第二节 SAS 的启动和退出.....	4	第三章 SAS 运算符和函数.....	28
一、SAS 的启动.....	4	第一节 SAS 运算符.....	28
二、SAS 的退出.....	4	第二节 SAS 函数.....	28
第三节 SAS 的显示管理系统.....	4	第四章 计量资料的单变量描述.....	32
一、显示管理系统的窗口.....	4	第一节 频数表的编制.....	32
二、常用指令.....	6	第二节 单变量描述.....	33
三、功能键.....	6	一、MEANS 过程.....	33
第四节 SAS 程序.....	7	二、UNIVARIATE 过程.....	35
一、简单程序示例.....	8	第三节 正态性检验.....	37
二、程序结构.....	8	第四节 几何均数的计算.....	38
三、程序语法规范.....	8	第五章 两均数的比较.....	40
四、程序运行.....	8	第一节 均数可信区间的估计.....	40
五、程序修改.....	9	一、单一总体均数的可信区间.....	40
六、程序保存.....	9	二、两个总体均数相差的可信区间.....	40
七、程序调用.....	10	第二节 单样本均数的 t 检验.....	41
第二章 数据库与数据集.....	11	第三节 配对资料两样本均数比较的 t 检验.....	42
第一节 数据库.....	11	一、MEANS 过程.....	43
一、库标记.....	11	二、UNIVARIATE 过程.....	43
二、数据库类型.....	11	第四节 成组资料两样本均数比较的 t 检验.....	43
第二节 数据集.....	11	第五节 两样本均数的等效性检验.....	44
一、数据集的类型.....	11	第六章 多个样本均数比较的方差分析.....	46
二、变量类型.....	11	第一节 完全随机设计资料的方差分析.....	46
三、数据集的命名.....	12	第二节 随机区组设计资料的方差分析.....	49
第三节 数据集的建立.....	12		
一、创建 SAS 数据集.....	12		
二、其他格式文件转换成 SAS 数据集.....	15		
第四节 数据集的整理.....	20		
一、数据集的浏览与修改.....	20		

第三节 拉丁方设计资料的方差分析	51
第四节 两阶段设计资料的方差分析	52
第五节 析因设计资料的方差分析	54
一、两因素两水平的析因分析	54
二、两因素三水平的析因分析	55
三、三因素不同水平的析因分析	56
第六节 正交设计资料的方差分析	58
第七节 嵌套设计资料的方差分析	59
第八节 裂区设计资料的方差分析	61
第九节 重复测量资料的方差分析	62
一、两因素两水平的分析	62
二、两因素多水平的分析	65
第七章 两变量直线相关和回归分析	68
第一节 直线相关分析	68
第二节 直线回归分析	69
一、直线回归方程的确定	69
二、两条回归直线的比较	71
第三节 秩相关	73
第八章 非线性回归分析	75
第一节 指数曲线回归	75
第二节 对数曲线回归	76
第九章 多变量回归分析	78
第一节 多元线性回归	78
第二节 筛选变量	79
第十章 χ^2 检验	84
第一节 四格表资料的 χ^2 检验	84
一、四格表专用公式	84
二、连续性校正公式	85
三、配对四格表资料的 χ^2 检验	87
第二节 $R \times C$ 表资料的 χ^2 检验	87
一、双向无序资料	88
二、单项有序资料	89
三、双向有序资料	90
四、分层资料	91
第十一章 二项分布、Poisson 分布和负二项分布	94
第一节 二项分布	94

一、阳性事件发生的概率	94
二、样本率与总体率的比较	94
三、两个样本率比较的 u 检验	95
第二节 Poisson 分布	95
一、样本均数与总体均数比较	95
二、两个样本均数比较	97
第三节 负二项分布中的参数估计	98
第四节 拟合优度检验	99
第十二章 非参数统计方法	101
第一节 配对资料的比较	101
第二节 成组资料的比较	102
第三节 两组等级资料的比较	103
第四节 完全随机设计资料的比较	104
第五节 随机区组设计资料的比较	105
第十三章 协方差分析	107
第一节 完全随机设计资料的协方差分析	107
第二节 随机区组设计资料的协方差分析	108
第十四章 Logistic 回归分析	110
第一节 两个自变量的 Logistic 回归分析	110
第二节 多个自变量的 Logistic 逐步回归分析	112
第三节 1:M 配对资料的条件 Logistic 回归分析	115
第十五章 生存分析	121
第一节 生存率估计	121
一、乘积极限法	121
二、寿命表法	122
第二节 生存曲线比较的 log-rank 检验	125
第三节 Cox 回归分析	127
第十六章 多元统计分析	131
第一节 判别分析	131
一、两类判别	131
二、Bayes 准则下的判别分析	133
三、逐步判别分析	134
第二节 聚类分析	138

一、样本聚类·····	138	项的详细说明·····	172
二、指标聚类·····	140	第三节 多重填补过程应用举例·····	177
第三节 主成分分析·····	141	一、基于 EM 法则的极大似然估计·····	177
第四节 因子分析·····	143	二、趋势得分法·····	178
第五节 典型相关分析·····	146	三、回归方法·····	179
第十七章 统计结果输出·····	152	四、MCMC 方法·····	179
第一节 MEANS 过程·····	152	五、检查 MCMC 过程中的收敛性·····	180
第二节 UNIVARIATE 过程·····	153	第十九章 数据挖掘及 SAS/Enterprise	
第三节 FREQ 过程·····	155	Miner 模块·····	182
一、频数表内容的输出·····	155	第一节 数据挖掘简介·····	182
二、关于行列变量间关系结果的输出·····	155	第二节 SAS/Enterprise Miner 模块	
第四节 ANOVA 过程·····	157	介绍·····	183
第五节 NPAR1WAY 过程·····	158	一、Enterprise Miner 界面的进入	
第六节 CORR 过程·····	159	·····	184
第七节 REG 过程·····	161	二、界面简介·····	184
第十八章 多重填补及其 SAS 过程·····	164	三、基本操作·····	185
第一节 多重填补的原理·····	164	第三节 SAS/Enterprise Miner 实现	
一、概述·····	164	过程·····	189
二、数据缺失的特征·····	165	一、计划的创建·····	189
三、数据填补的方法·····	166	二、输入数据集的定义·····	190
四、多重填补的综合统计推断·····	167	三、应变量的设置·····	190
五、应用多重填补的注意事项·····	169	四、训练集和验证集的创建·····	191
第二节 SAS 中的多重填补过程·····	169	五、变量转换·····	192
一、多重填补过程简介·····	169	六、创建模型·····	193
二、多重填补过程中各语句和选择		七、模型评价·····	198
		参考文献·····	200

第一章 SAS 概述

SAS 为 statistical analysis system 的缩写,即统计分析系统,是当今国际上最著名的数据分析软件,由美国北卡罗纳州州立大学的 A.J.Barr 和 J.H.Goodnight 两位教授于 1966 年开始研制,并于 1976 年正式推出。SAS 可以用于完成统计分析、时间序列分析和运筹决策等工作。本书主要介绍 Windows 操作系统下的 SAS 6.12 版,以及 SAS 8 版中关于数据填充和数据挖掘的功能。

第一节 SAS 的特点和运行环境

一、SAS 的基本功能

SAS 的最大特点就是数据管理和数据分析融为一体。完成以数据为中心的操作主要有以下几方面的功能:

1. 数据交换 SAS 可以用任何格式读入数据值,然后将数据转换成 SAS 数据集。它具有很强的与外部文件交换信息的功能,可以用文件操作管理方法把不同数据库的数据组合在一起,供 SAS 过程分析处理,也可以将 SAS 数据集的数据转换成其他格式的数据文件,供其他软件处理。

2. 数据管理 SAS 为用户提供完备的 SAS 语句和函数用于数据加工处理,有些语句用于执行标准操作,如建立新变量、数据查询、累加求和及修改错误等;DO/END、IF-THEN/ELSE 等控制语句,可用于选择满足条件的数据构成新的数据子集;DROP、KEEP 等信息语句,用于选择在新数据子集中被删除和保留的原数据集中的变量;SET、MERGE 等文件操作语句,可以进行数据集的合并、拼接,从而构成了一套完整的语言系统。其不仅可以同时处理多个数据文件,而且可以将一个数据集拆分成几个数据子集分别处理。

3. 数据分析 SAS 系统可以进行多种统计分析,包括:①计算简单的描述统计量,如均数、标准差、标准误、总和、平方和、极差、相关系数、峰度系数、偏度系数等,多达 40 项;②计算概率分布函数、分位数和产生随机数;③对数据进行标准化、编秩及计算其统计量;④产生并分析列联表;⑤进行方差分析、相关与回归分析、线性模型拟合、属性数据分析、多变量数据的判别和聚类分析、非参数统计分析、生存分析、时间序列分析、实用预测、质量控制、运筹学统计分析等过程;⑥绘制二维与三维的基本统计图,如条图、直方图(水平或垂直)、圆图、散点图、等差和等比线图、曲线拟合图、时间序列图等,从数据中获得有价值的信息,便于指导实践研究。

4. 数据呈现 SAS 不仅可以数据集的数据和统计分析结果打印输出,还可以将某个过程产生的数据输出到另外的数据集中,用另一个过程进行处理。还可以将多个过程产生的数据组合成新的数据集(有时需使用宏语言),归纳总结后一起输出或再分析。分析结果可以通过列表报告和汇总报告输出,还可以根据用户自定义的报表输出。

二、SAS 的功能模块

SAS 是一个模块化的组合软件系统,由多个功能模块组合而成,其基本部分是 BASE SAS 模块。

1. BASE SAS 该模块是 SAS 系统的核心,承担着主要的数据管理任务,并管理用户使用环境,进行用户语言的处理,调用其他 SAS 模块和产品。也就是说, SAS 的运行,首先必须启动 BASE SAS,它除了本身所具有的数据管理、程序设计及描述统计计算功能外,还是 SAS 系统的中央调度室。它除可单独存在外,也可与其他产品或模块共同构成一个完整的系统。各模块的安装及更新都可通过其安装程序非常方便地进行。

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增加不同的功能。SAS 有一个智能型绘图系统,不仅能绘各种统计图,还能绘出地图。SAS 提供多个统计过程,每个过程均含有极丰富的任选项。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要求。

2. SAS/STAT 为统计分析模块。该模块包括回归分析、方差分析、定性数据分析、多变量分析、判别和聚类分析、残存分析、心理测验分析和非参数统计分析等方法共 40 多个过程。每个过程还提供多种不同的算法及选择,从而组成一个庞大而完整的统计分析方法集。SAS/STAT 还为主成分分析、典型相关分析、判别分析和因子分析提供了许多专用过程。它是国际上统计分析领域中的标准软件。

3. SAS/INSIGHT 为交互数据分析模块。SAS/INSIGHT 为用户提供了一个进行交互式数据探索和分析的工具,运用鼠标和系统提供的下拉菜单,用户可以同时打开多个窗口对数据和图像进行比较、探索和分析。强有力的图像表现是 INSIGHT 的重要特点,用户对表中的数据进行检测时,可以同时将这些数据显示在直方图、散点图和三维旋转图上,便于用户发现奇异值。利用 INSIGHT 可以分析单变量分布和多变量分布,用方差分析、回归分析和广义线性模型拟合要说明的模型。

4. SAS/ASSIST 为菜单驱动模块。SAS/ASSIST 集成了 SAS 系统其他模块的各种功能,提供了一个菜单驱动、任务导向的用户界面。借助它用户不需编程,只要根据处理数据的需要,用鼠标在屏幕上指定选项就可方便地使用 SAS 提供的各种功能,免除了用户学习 SAS 语言的困扰,并可帮助用户学习 SAS 语言。有经验的用户还根据自己的需要调用 SAS/ASSIST 的不同部分或裁剪 SAS/ASSIST 的菜单构成自己的应用系统。

5. SAS/GRAPH 为绘图模块。SAS/GRAPH 能够完成多种绘图功能,如直方图、圆图、星形图、散点图、线图、曲线图、三维曲面图、等高线图和地理图等,这些图形可以非常形象、直观地表现各变量之间的关系及数据的分布状态,对解决各种实际问题起着重要的辅助作用。SAS/GRAPH 还有一个全屏幕图形编辑器和丰富地中西文矢量图形字体,用户可以在图面上自由地绘制文字及图形元素,对图形进行修改,填加图形标记;还可对多幅图形进行任意的拼接组合。

6. SAS/ACCESS 为数据库接口模块。SAS/ACCESS 提供了与目前许多流行数据库软件的接口,利用 SAS/ACCESS,可建立外部其他数据库的一个统一的公共数据界面。SAS/ACCESS 提供的接口是双向的,既可将数据读入 SAS,也可在 SAS 中更新外部数据或

将 SAS 数据加载到外部数据库中。SAS/ACCESS 支持的数据库主要有：IMS-DLI, SQL/DS, DB2, ADABAS, Rdb, ORACLE, Sybase, INGRES, Informix, DBF/DIF, ODBC 等。

7. **SAS/ETS** 为经济计量学和时间序列分析模块。它是研究复杂系统和进行预测的有力工具。该模块包含全面的时间序列时域分析和谱域分析，如实用预测（逐步自回归、指数平滑、Winters 方法）、序列相关校正回归、分布滞后回归、ARIMA 模型、状态空间方法、谱分析和互谱分析等，还提供许多处理时间序列数据的实用程序，如时间频率转换和插值、季节调整等。

8. **SAS/OR** 为运筹学模块。SAS/OR 提供全面的运筹学方法，是一个优秀的决策支持工具。该软件包含通用的线性规划、整数规划以及混合整数规划和非线性规划方法，还包含用于项目管理、时间安排和资源分配等问题的一系列方法。

9. **SAS/IML** 为交互式矩阵程序设计语言模块。SAS/IML 提供了一套完整的面向矩阵的交互式矩阵编程语言（interactive matrix language, IML）。该语言处理的基本数据元素是数据矩阵，数据可以是数值型的，也可以是字符型的。用这种语言可方便地处理各种复杂的矩阵运算，进而在控制语句的帮助下实现许多复杂的算法。

10. **SAS/FSP** 为快速数据处理的交互式菜单系统模块。SAS/FSP 是一个用来进行数据处理的交互式菜单系统，具有全屏幕数据录入、编辑和查询以及数据文件创建等功能。可以对一条记录进行操作，也可以在一个屏幕上操作多个记录。同时，它也是一个开发工具。

11. **SAS/AF** 为交互式全屏幕软件应用系统模块。SAS/AF 是一个应用开发工具。用户使用 SAS/AF 可将包含众多功能的 SAS 软件作为方法库，利用 SAS/AF 的屏幕设计能力以及 SQL 语言的处理能力来快速开发各种功能强大的应用系统。SAS/AF 也采用了 OOP（面向对象编辑）技术，使用户可方便快速开发各类具有图形用户界面(GUI)的应用系统。

三、SAS 的运行环境

SAS 6.12 已经可以在 Windows 操作系统下使用，为纯 32 位软件，其运行环境为：

1. **操作系统要求** Windows 95、Windows 98、Windows/NT Version 3.51 以上、Windows 2000 非服务器版以及 Windows XP。

2. **硬件要求** CPU: PENTIUM 100 以上；内存：16 MB 以上；显示器：SVGA；硬盘：350 MB 空闲硬盘；其他：光驱、鼠标等。

3. **SAS 6.12 各个模块需要的硬盘空间** 见表 1.1。

表 1.1 SAS 6.12 各模块运行所需硬盘空间

模 块	硬盘空间 (KB)
SAS/AF	1 127
SAS/ASSIST	7 961
BASE SAS	16 944
CORE OF THE SAS SYSTEM	62 697
SAS/ACCESS	2 040
SAS/ETS	11 177
SAS/FSP	425
SAS/GRAPH	17 331
SAS/IML	2 906
SAS/INSIGHT	2 789
SAS/OR	19 345
SAS/STAT	11 413

第二节 SAS 的启动和退出

一、SAS 的启动

SAS 系统的启动可以通过以下几种方法:

1. 快捷方式 在安装了 SAS 6.12 后,会自动在应用程序项中创建 SAS 启动的快捷方式,可以直接通过快捷方式启动 SAS。具体方法:打开电脑,进入 Windows 操作系统,用鼠标左键点击“开始”,将鼠标移动到“程序”,在出现的应用程序项中会出现“The SAS System”。将鼠标移动到该项上,会出现两个快捷方式,其中一个是“The SAS System for Windows v6.12”,点击该项,就可启动 SAS。

2. 运行可执行文件 安装 SAS 6.12 时,会将有关文件安装在硬盘上的某个子目录中,如“C:\SAS”。可以执行可执行文件启动 SAS。具体方法为:①进入“资源管理器”中,找到安装 SAS 的目录“C:\SAS”,进入该子目录,找到 SAS.EXE 文件,用鼠标双击它就可以启动 SAS。②点击“开始”,在菜单中点击“运行”,在运行对话框中直接键入“C:\SAS.EXE”,按下“确定”即可启动 SAS。或者点击“浏览”,进入浏览对话框,找到 SAS.EXE 所在的位置,双击该文件;或单击该文件再按“打开”,则返回运行对话框,再按下“确定”也可以启动 SAS。

另外,可以在桌面上建立“SAS.EXE”的快捷方式,具体方法为:进入“资源管理器”,进入安装 SAS 的目录,将 SAS.EXE 文件拖到桌面上;或者点击“开始”,在“程序”中进入“The SAS System”,用鼠标右键点击“The SAS System for Windows v6.12”,在出现的菜单中点击“发送到(T)”,再点击“桌面快捷方式”,则在桌面上出现启动 SAS 的快捷方式。双击桌面上的快捷方式可启动 SAS。

二、SAS 的退出

当我们用 SAS 完成了统计分析后,可以退出 SAS 系统。退出 SAS 系统可以用以下几种方法:

1. 菜单操作 点击“File”,选择“Exit...”,或者同时按下“ALT”和“F4”,或者点击“×”,将会出现“Exit”对话框,对话框中显示“Are you sure you really want to terminate the SAS session?”,点击“确定”即可退出 SAS 系统。如点击“取消”则返回 SAS 系统。

2. 键入命令 在命令框中键入“BYE”或“ENDSAS”可直接退出 SAS 系统,而不会出现上述的“Exit”对话框。

第三节 SAS 的显示管理系统

启动了 SAS 后,就进入 SAS 的显示管理系统(display management system, DMS),见图 1.1。在 DMS 中可以进行 SAS 程序的编辑、运行、存储、调用、结果输出及打印等过程。

一、显示管理系统的窗口

DMS 是 SAS 系统主要工作界面,用户可以灵活运用不同的窗口反复进行编程运算,直

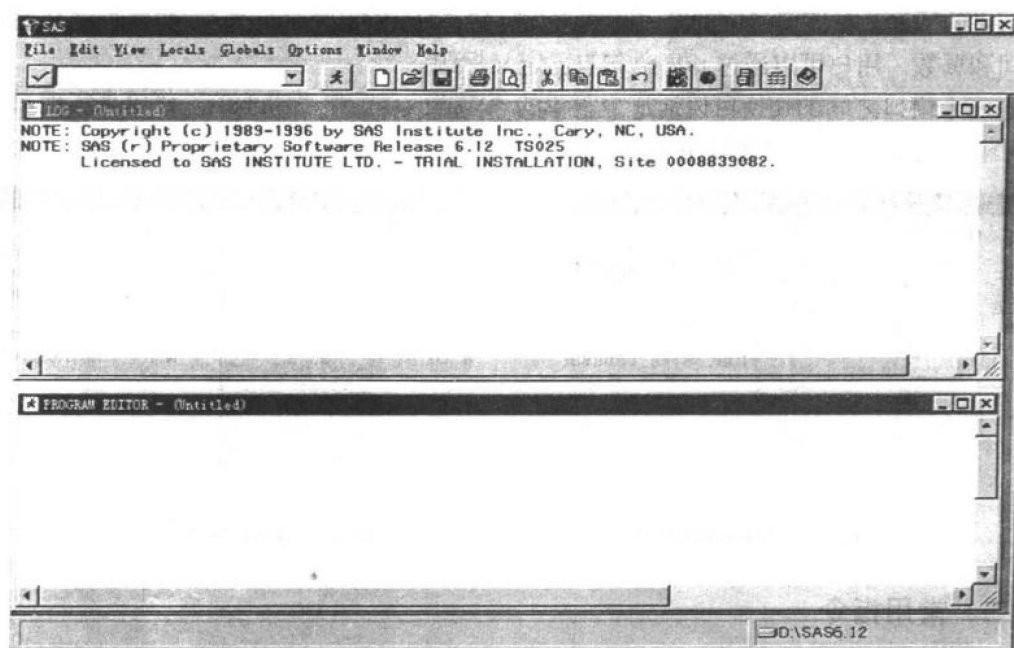


图 1.1 SAS 6.12 版视窗管理系统的界面

至完成统计分析为止。DMS 提供了一系列窗口，其中 PROGRAM EDIT 窗口、LOG 窗口和 OUTPUT 窗口是主要窗口，这些窗口是启动 SAS 就存在的，也是我们重点介绍的 3 个窗口。图 1.1 中显示了 2 个窗口，即 PROGRAM EDIT 窗口（下）和 LOG 窗口（上），而 OUTPUT 窗口被这两个窗口覆盖了。另外，DMS 还提供了多个辅助窗口，如 HELP、KEYS、LIBNAME 等。这些窗口可根据需要打开或关闭。

1. PROGRAM EDIT 窗口（编辑窗口） 该窗口的基本形式见图 1.1 下方的窗口，用户也可将该窗口最大化，进行全屏幕编辑。该窗口的主要功能是编辑 SAS 程序语句，并将程序语句提交系统执行。由于程序语句都是纯文本格式，所以该窗口又称为文本编辑器。在编辑程序语句时，每行语句前面可以显示语句标号（如 00001~00016），该部分被称为数字区。需要显示语句标号可以在主菜单下方的命令框中输入“NUMS”，然后点击前面的“√”或直接按“ENTER”键就可以了。如果不想显示语句标号，则键入“NUMS OFF”。

2. LOG 窗口（日志窗口） 该窗口的基本形式见图 1.1 上方的窗口，主要作用是显示运行程序后的有关信息。显示的信息内容包括建立了哪个数据集；建立的数据集包括多少变量和观测；执行了什么过程；执行过程运行了多少时间；语句中有什么错误等等。如果语句过程没有错误，提示的信息一般用蓝色字体表示，每个提示信息开头用 NOTE 表示；如果语句中有错误，而该错误是 SAS 系统能够纠正的，则显示的提示信息为绿色字体，开头用 WARNING 表示；如果语句中的错误是 SAS 系统无法纠正的，则用红色字体显示提示信息，开头用 ERROR 表示。

3. OUTPUT 窗口（输出窗口） 该窗口在启动 SAS 后没有直接显示出来，而是被 PROGRAM EDITOR 窗口和 LOG 窗口覆盖，只有运行了某个过程后，才显示出来。该窗口的主要作用是显示程序运行的结果。在运行 SAS 系统期间，相继产生的输出结果都附加在上

一次结果的后面。在该窗口中的第一行通常为标题“The SAS System”，后面是程序开始运行的时间和页数。用户可以通过 title 语句自己定义标题。

这三个窗口之间的切换可以通过主菜单的“Windows”和“Globals”菜单来实现，如图 1.2 和图 1.3 所示。也可以通过指令来实现。

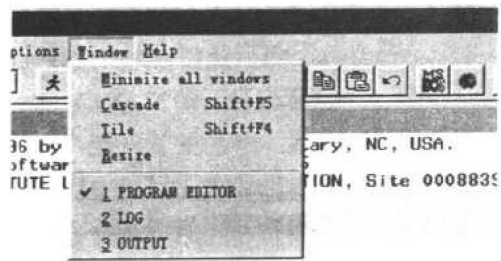


图 1.2 Windows 菜单

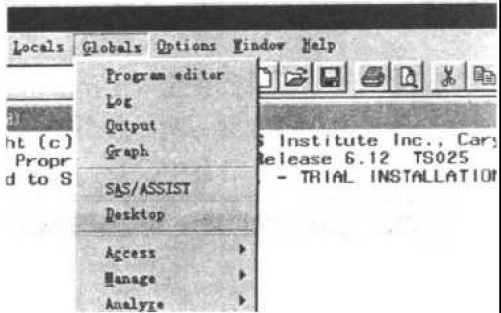


图 1.3 Globals 菜单

二、常用指令

在显示管理系统的主菜单的下方，是一个执行 SAS 指令的工具 ☒ ☐。在该工具的文本输入框中可以输入一套指令，用于实现对窗口的管理功能。有些指令只能在指定窗口执行，而有些指令是可以用于大多数窗口，称为窗口通用指令。输入一个指令，点击前面的“√”或按下“ENTER”键，就可以执行该指令。表 1.2 是几条比较常用的指令。

表 1.2 SAS 常用指令及其意义

命 令	意 义
BYE	退出 SAS 系统
CLEAR	清除当前窗口中的内容
END	退出当前窗口，返回编辑窗口
ENDSAS	退出 SAS 系统
FILE 'filename'	将当前窗口的内容存储到指定文件
HELP	进入帮助窗口
INCLUDE 'filename'	调用指定文件
KEYS	进入快捷键定义窗口
LIBNAME	确定 SAS 数据库的内容
LOG	进入日志窗口
NUMS	打开编辑窗口的数字区
NUMS OFF	关闭编辑窗口的数字区
OPTIONS	进入参数定义窗口
OUTPUT	进入输出口
PROGRAM	进入编辑窗口
RECALL	调用上次执行的程序
SUBMIT	提交编辑窗口编辑的程序代码

三、功能键

上述的有些指令可以通过 SAS 定义的功能键来完成，这些功能键都是 SAS 预先设定的（见表 1.3）。在指令工具中键入“KEYS”，点击“√”或按“ENTER”，或按“F9”，就可以查看和自定义功能键。表 1.3 中没有定义功能键的地方，允许用户自己定义功能，同时用户可以根据自己的习惯，改变功能键的位置。

表 1.3 SAS 预先设定的功能键

功能键	对应的指令	功 能
F1	Help	打开帮助对话框
F2	Reshow	重新建立被显示的窗口
F3	End; /*gsubmit buffer=default*/	关闭当前窗口, 返回编辑窗口
F4	Recall	调用上次执行的程序
F5	Pgm	进入编辑窗口
F6	Log	进入日志窗口
F7	Output	进入输出窗口
F8	Zoom off;submit	恢复编辑窗口的大小; 提交程序代码
F9	Keys	进入快捷键定义窗口
F11	Command bar	打开指令工具条
F12		
SHF F1	Subtop	执行编辑窗口编辑的第一行程序
SHF F2		
SHF F3		
SHF F6		
SHF F7	Left	向左翻页
SHF F8	Right	向右翻页
SHF F9		
SHF F10	Wpopup	相当于单击鼠标右键, 显示小菜单
SHF F11		
SHF F12		
CTL F1		
CTL F2		
CTL F3		
CTL F11		
CTL F12		
ALT F1		
ALT F2		
ALT F3		
ALT F11		
ALT F12		
CTL A		
CTL B	Libname	确定 SAS 数据库的内容
CTL D	Dir	进入 DIR 窗口
CTL E	Clear	清除当前窗口中的内容
CTL F	Footnote	进入 FOOTNOTE 窗口
CTL G		
CTL H	Help	进入帮助窗口
CTL I	Options	进入参数定义窗口
CTL J		
CTL K	Cut	将选择的的内容剪切
CTL L	Log	进入日志窗口
CTL M	Mark	标记选择的内容
CTL Q	Filename	进入 FILENAME 窗口
CTL R	Rfind	继续查找
CTL T	Title	进入 TITLE 窗口
CTL U	Unmark	取消标记
CTL W	Access	进入 ACCESS 窗口
CTL Y		

第四节 SAS 程序

SAS 程序是在 PROGRAM EDITOR 窗口中编辑的一段 SAS 语句, 提交后可以在 LOG 窗口中显示有关信息和提示, 在 OUTPUT 窗口中显示运行过程的结果。下面通过一个简单的例子, 来说明程序的结构。

一、简单程序示例

例 1.1 12 份肝炎病人血清 γ 谷氨酰转移酶(mmol/s)的含量分别为 60, 142, 195, 80, 242, 220, 190, 25, 212, 38, 236, 95, 试计算其均数。

程序 1.1

```
data prg1_1;
  input x @@;
cards;
60 142 195 80 242 220 190 25 212 38 236 95
;
proc means;
  var x;
run;
```

二、程序结构

一个完整的 SAS 程序一般由数据步 (DATA STEP) 和过程步 (PROC STEP) 两部分组成。数据步以关键词 “DATA” 开头, 过程步以 “PROC” 开头, 以 “RUN” 结束。PROC 为英文单词 PROCEDURE 的缩写。数据步的作用为指定数据集的名称, 定义数据集的变量 (如变量名称、变量类型等) 和读入原始数据。本例数据步从 “data prg1_1;” 开始到数据下面的分号 “;” 为止结束, 建立了名为 prg1_1 的数据集。过程步的作用是调用现有的 SAS 过程对指定的数据集进行统计分析。本例过程步执行的是 MEANS 过程, 计算默认数据集 (即数据集 prg1_1) 中数据的例数、均数、标准差、最小值和最大值等统计量。过程步从 “proc means;” 开始, 到 “run;” 结束。

三、程序语法规范

SAS 程序由语句组成, 每个语句以 “;” 作为结束符号。同一行中可以有多个语句, 一个语句也可分几行编写, 不过一般不宜超过三行。为方便检查和修改, 每行可输入一个语句, 每个语句中各个元素以一个或几个空格分隔。输入程序语句时, 可在光标闪烁处逐个字母输入。一行语句结束后, 按 “ENTER” 换行, 继续输入。值得注意的是, CARDS 语句后面的数据必须另起一行输入, 数据输入完毕后, 必须另起一行, 输入分号 “;” 表示数据输入结束。

四、程序运行

当程序语句被确认无误后, 可以将程序提交系统运行。提交程序的方法有以下几种:

- (1) 在执行指令的文本框中键入 “SUBMIT”, 然后点击 “√” 或按 “ENTER”。
- (2) 点击主菜单中的 “Locals”, 再点击 “Submit”, 如图 1.4 所示。

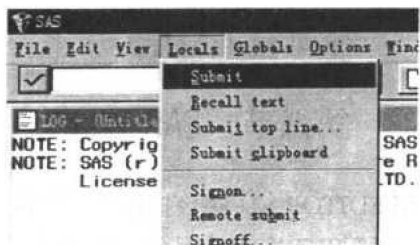


图 1.4 Locals 菜单

当 SAS 程序运行后, PROGRAM EDITOR 窗口中的程序语句就消失。如果程序编写无错误, 在 OUTPUT 窗口中就会出现如下程序运行结果:

Analysis Variable : X				
N	Mean	Std Dev	Minimum	Maximum
12	144.5833333	80.9797487	25.0000000	242.0000000

在 LOG 窗口中显示如下的程序运行信息:

```
1 data prg1_1;
2 input x @@;
3 cards;
NOTE: SAS went to a new line when INPUT statement reached past the end of a line.
NOTE: The data set WORK.PRG1 has 12 observations and 1 variables.
NOTE: The DATA statement used 0.05 seconds.
5 ;
6 proc means;
7 run;
NOTE: The PROCEDURE MEANS used 0.05 seconds.
```

五、程序修改

通常情况下, 在程序运行完毕后, 要先检查 LOG 窗口中的日志, 看程序语句有无错误。如果程序语句编写有误, 而且该错误能被 SAS 系统纠正, 则在 LOG 窗口中会出现红色下划线和错误标记, 并用绿色字体提示错误, 而程序照常运行。

```
8 proc meanss;
-----
1
9 run;
WARNING 1-322: Assuming the symbol MEANS was misspelled as MEANSS.
NOTE: The PROCEDURE MEANSS used 0.05 seconds.
```

如果所犯的错误无法被 SAS 系统纠正, 则在 LOG 窗口中会出现红色字体提示错误, 而且程序无法运行。

```
10 proc mean;
ERROR: Procedure MEAN not found.
11 run;
NOTE: The SAS System stopped processing this step because of errors.
NOTE: The PROCEDURE MEAN used 0.05 seconds.
```

此时需修改程序语句, 才能完成运算。修改程序语句, 首先将窗口切换到 PROGRAM EDITOR 窗口, 将原来编辑的程序语句调出来。调用被提交的程序语句可以通过输入指令“RECALL”, 或按功能键“F4”, 还可以通过“Locals”菜单中“Recall text”完成(图 1.4)。

调出程序后, 在原来有错误的地方修改程序语句, 然后再提交运行。有时需反复几次, 直到 LOG 窗口不再出现错误提示为止。

六、程序保存

程序语句编辑无误后, 可以将编辑好的程序以文件的形式保存下来, 以备以后检查或修改。保存程序可键入指令“file ‘路径+文件名’”, 该处路径为绝对路径, 文件名的后缀必须是“.sas”, 而且“路径+文件名”必须用单引号括起来, 如“file ‘c:\sas\sasuser\prg1_1.sas’”。也可以通过“File”菜单中的“Save”选项来保存程序文件, 如图 1.5 所示。在出现的“Save”对话框中选择好路径, 再键入文件名, 此时不必加后缀, 因为系统将默认后缀名为“.sas”。也可以通过工具栏上的快捷方式保存程序文件, 只需点击工具栏上的, 就会出现“Save”对