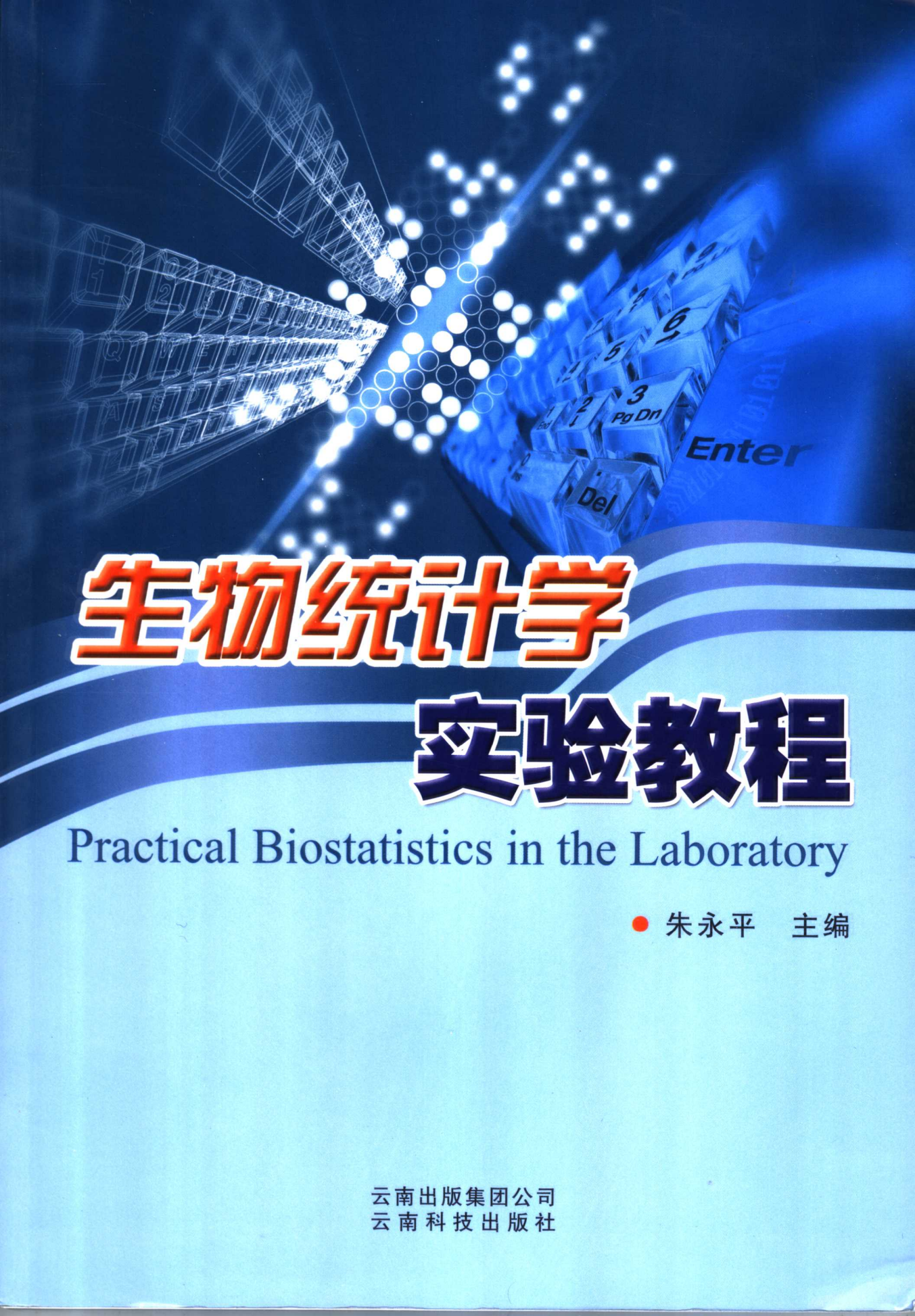




生物统计学

实验教程

Practical Biostatistics in the Laboratory

● 朱永平 主编

云南出版集团公司
云南科技出版社

An abstract graphic at the top of the cover features a perspective view of a computer keyboard. The keys are rendered in a 3D wireframe style, with some keys like '1', '2', '3', '4', '5', '6', '7', '8', '9', '0', 'Pg Dn', 'Del', and 'Enter' clearly visible. A bright, glowing beam of light or data points originates from the center of the keyboard and extends upwards and outwards, creating a sense of dynamic energy and technological advancement. The background is a dark, textured surface with some faint, glowing circular patterns.

生物统计学 实验教程

Practical Biostatistics in the Laboratory

主 编：朱永平（云南农业大学）

副主编：吴渝生（云南农业大学）

毛孝强（云南农业大学）

徐绍忠（云南农业大学）

杨永建（云南农业大学）

云南出版集团公司
云南科技出版社
· 昆明 ·

图书在版编目 (CIP) 数据

生物统计学实验教程 / 朱永平主编. —昆明: 云南科技出版社, 2006. 9

ISBN 7 - 5416 - 2428 - 4

I. 生... II. 朱... III. 生物统计—高等学校—教材 IV. Q-332

中国版本图书馆 CIP 数据核字 (2006) 第 119466 号

云南出版集团公司

云南科技出版社出版发行

(昆明市环城西路 609 号云南新闻出版大楼 邮政编码: 650034)

昆明市五华区教育委员会印刷厂印刷 全国新华书店经销

开本: 787mm × 1092mm 1/16 印张: 13.75 字数: 338 千字

2006 年 11 月第 1 版 2006 年 11 月第 1 次印刷

印数: 1 ~ 3000 册 定价: 33.00 元

前 言

生物统计学是运用数理统计的原理和方法来分析和解释生物界各种现象的试验调查数据的一门科学。随着生物学的不断发展,对试验结果不再局限于定性的描述,而是需要从大量试验数据中,应用统计学方法,分析和解释其数量上的变化,对试验结果进行科学的分析及符合科学实际的推断。生物学是一门试验科学,不论你从事的是生物学的哪一个分支,都必须进行试验,而不能只进行逻辑推理。因试验所得的结果都存在试验误差,使得试验结果无例外地带有或多或少的不确定性,在这种情况下,必须用统计学方法得出正确的结论。所以生物统计学是现代生物研究不可缺少的工具,是生物、农学类专业一门重要专业基础课,其教学质量的提高,将直接影响到科学技术人员的科研试验设计水平、运用统计方法作出结论的科学性,熟练地应用现代试验设计方法及统计分析技术,将直接改善科技人员的业务素质。

SAS (Statistical Analysis System) 系统软件是世界公认的权威性统计分析软件之一,是一个大型集成信息分析系统。SAS 系统集数据访问、数据管理、数据分析、数据挖掘功能于一体,是一个功能强大且被国际公认的国际标准统计分析软件。随着生物学的发展,数据管理与数据分析在各领域显得非常重要,SAS 系统已成为研究人员、学者发表学术论文的首选统计分析工具。

本书以 SAS 系统软件 6.12、9.0 版的主要统计分析功能,从实际应用角度出发,每一章都通过大量实用例子详细讲解了 SAS 统计分析软件中的相应分析过程的使用方法,对每个实例的统计输出结果和统计量的实际意义做了详细的解释,并给出了分析思路、具体应用方法。对本科生、研究生及科技工作者应用 SAS 系统软件可以提供很好的帮助。

本书凝聚了作者多年的一线教学经验,内容贴近读者,讲解透彻而实用,提供分析思路,给出解决方案及输出内容解释,通过掌握实例做到触类旁通,引导读者灵活应用所学。本书共分 17 章,概括了 11 个方面的内容:①SAS 软件基本操作;②利用 SAS 程序进行试验资料的整理和特征数的计算;③统计假设检验;④卡平方 (χ^2) 检验;⑤方差分析(单因素试验结果的方差分析、多因素试验结果的方差分析、裂区设计试验结果的方差分析、简单格子设计的方差分析、正交设计试验资料的方差分析、品种区域试验结果的方差分析);⑥线性相关与回归分析;⑦多元回归分析;⑧非线性回归分析;⑨聚类分析;⑩主成分分析;⑪通径分析。

本书适合生物技术、农学、植保、烟草、药鉴、种子科学与工程、检疫、食品等专业作生物统计学实验教材,也可供研究生、科技工作者阅读。对本科生、研究生毕业论文试验数据处理、分析说明、试验结果解释及论文撰写有较好的帮助。

由于作者水平的限制和涉猎资料有限,本书难免有错误和不确切之处,敬请读者批评指正,以便日后修改。

云南农业大学 朱永平

目 录

第 1 章 SAS 软件基本操作	1
1.1 SAS 系统概述	1
1.2 SAS 系统的启动	1
1.3 显示管理系统	1
1.3.1 显示管理系统的几个主要窗口	1
1.3.2 几个常用的显示管理系统命令	3
1.3.3 使用 pull down menu (下拉菜单)	3
1.3.4 使用功能键和工具栏	7
1.4 SAS 程序的编写	7
1.5 数据步	7
1.5.1 由外部数据文件建立 SAS 数据集	7
1.5.2 在作业流中输入数据	7
1.6 过程步	8
1.7 执行程序	9
1.7.1 提交程序	9
1.7.2 修改错误语句	9
1.7.3 保 存	9
1.8 SAS 常用功能键	10
第 2 章 试验资料的整理和特征数的计算	12
2.1 利用 SAS 程序绘制连续性变异资料的柱形图 (直方图)	12
2.2 样本平均数、标准差、标准误及变异系数的计算	13
第 3 章 统计假设检验	14
3.1 单个样本的检验 (样本平均数与总体平均数差异的假设检验)	14
3.2 配对数据检验	14
3.3 成组数据 t 检验	17
第 4 章 卡平方 (χ^2) 检验	20
4.1 适合性检验	20
4.1.1 适合性检验的意义	20
4.1.2 适合性检验的方法	20
4.1.3 资料分布类型的适合性检验	23
4.2 独立性检验	26
4.2.1 独立性检验的意义	26
4.2.2 独立性检验的方法	27
4.3 独立性检验的 SAS 程序 (FREQ 过程)	31

4.3.1	2×2 表的独立性检验	31
4.3.2	r×c 表的独立性检验	33
第 5 章	单因素试验结果的方差分析	35
5.1	完全随机设计试验结果的方差分析	35
5.2	单因素随机区组试验结果的方差分析	39
5.3	配伍组设计试验结果的方差分析	48
5.4	平衡不完全区组设计试验结果的方差分析	50
5.5	拉丁方设计试验结果的方差分析	55
5.6	希腊拉丁方设计试验结果的方差分析	62
第 6 章	多因素试验结果的方差分析	69
6.1	二因素完全随机设计试验结果的方差分析	69
6.1.1	两因素双向分组 (交叉分组) (组内无重复) 完全随机设计 试验结果的方差分析	69
6.1.2	两因素双向分组 (交叉分组) (组内有重复) 完全随机设计 试验结果的方差分析	75
6.1.3	两因素系统分组完全随机设计试验结果的方差分析	86
6.2	二因素随机区组试验结果的方差分析	89
6.3	三因素随机区组试验结果的方差分析	103
6.4	裂区设计试验结果的方差分析	105
第 7 章	简单格子设计试验结果的方差分析	121
第 8 章	正交设计试验资料的方差分析	126
8.1	单个观测值正交试验资料的方差分析	126
8.2	有重复观测值正交试验资料的方差分析	135
第 9 章	品种区域试验结果的方差分析	145
9.1	一年多点品种区域试验结果的方差分析	145
9.2	多年多点品种区域试验结果的方差分析	147
第 10 章	线性相关与回归分析	152
10.1	一元线性回归分析	152
10.2	一元线性回归的图形描述	154
10.3	相关系数的计算	156
第 11 章	多元回归分析	157
11.1	多元回归方程计算	157
11.2	多元回归关系的假设检验	158
11.3	偏回归系数的假设检验	158
第 12 章	协方差分析	159
12.1	单向分组资料的协方差分析	159
12.2	两向分组资料的协方差分析	162
第 13 章	非线性回归分析	164
13.1	倒数函数曲线	164

13.2	指数函数曲线	166
13.3	对数函数曲线	168
13.4	拟合抛物线（多项式回归分析）	171
13.5	Logistic 生长曲线（S 形曲线）	174
第 14 章	聚类分析	177
14.1	样品（Q 型）聚类分析	177
14.2	指标（R 型）聚类分析	182
第 15 章	主成分分析	190
第 16 章	通径分析	197
附录		204
附表 1	SAS 程序中有效关键字及其所表示的统计量	204
附表 2	SAS 程序在数据步中所用的语句及作用	205
附表 3	Excel 对数据进行统计处理的函数	206
附表 4	Excel 常用的统计分布函数	208
附表 5	Excel 用于统计检验的函数	209
参考文献		211

第1章 SAS 软件基本操作

1.1 SAS 系统概述

SAS 是“统计分析系统”(Statistical Analysis System)的缩写。该系统是由北卡罗来纳州立大学统计系的两位教授, A. J. Barr 和 J. H. Goodnight 于 20 世纪 60 年代末开发的。最初是以统计分析和线性统计模型为主, 至今已开发成为功能强大的集成应用软件系统。包括 30 多个工具模块, 广泛应用于实用统计、运筹学、质量控制、大型矩阵计算等, SAS 是目前公认的著名的国际上最流行并具有权威性的数据统计分析软件之一。本实验教程以 Windows 98、2000、XP 操作系统下的版本 6.12 (英文版) 及最新版本 9.0 (中文版), 介绍 SAS 统计功能的基本使用方法。

1.2 SAS 系统的启动

单击开始菜单“程序”项, 光标移到“The SAS System”程序项处, 屏幕上会显示“The SAS System for Windows v6.12”和“Update SAS License Information”, 选择“The SAS System for Windows v6.12”, 也可以在桌面上建立快捷方式, 选择桌面上“The SAS System for Windows v6.12”双击即启动 SAS。

1.3 显示管理系统

1.3.1 显示管理系统的几个主要窗口

SAS 启动后在屏幕上出现的是显示管理系统 (display manager) (见图 1.1)。

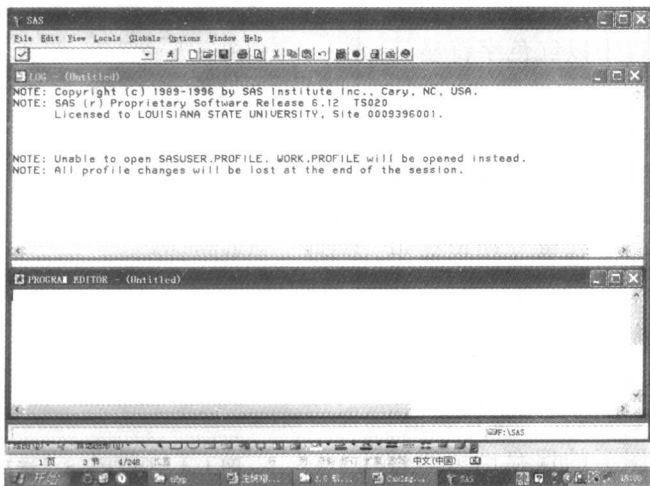


图 1.1 显示管理系统的 LOG (逻辑) 窗口和 PGM (编辑) 窗口

显示管理系统主要有三个窗口，其中两个窗口在启动后可直接看到，一个是逻辑记录窗口（LOG），另一个是程序编辑窗口（PROGRAM EDITOR）（PGM）。屏幕的左上角是命令框，在命令框中键入 OUTPUT 命令，屏幕上会显示出输出窗口（OUTPUT）（见图 1.2）

除这三个主要窗口外，还有其他一些窗口，如 KEYS 窗口（见图 1.3）、OPTIONS 窗口（见图 1.4）等。上述这些窗口的主要功能如下：

PROGRAM EDITOR（程序编辑窗口）：输入 SAS 程序和数据。编辑窗口支持全窗口编辑功能，支持 Windows 系统常规编辑操作，如复制、剪切、粘贴等功能操作。SAS V8、V9 提供了智能编辑功能，它可根据用户输入的 SAS 关键字、语句选项、变量名、数据、标记等不同内容显示不同的颜色，当用户输入的字串不正确时，对应颜色不对，并显示红色警告使用者有错误发生。

LOG（逻辑记录窗口）：运行逻辑记录窗口用于显示和记录 SAS 程序的运行情况，包括程序语句、注释、警告、错误显示（错误所在的列、出错原因及可能的修改方法）等信息，当程序出现错误时，LOG 记录窗口中以红色字符显示错误信息，以绿色显示一些警告信息。提示使用者修改程序中的错误。

OUTPUT（输出窗口）：显示程序执行的结果。输出窗口分页显示 SAS 程序运行所产生的文本输出结果（图形输出通过 GRAPHICS 窗口显示）。对于输出到 GRAPHICS 窗口的图形，可以使用菜单 File/Export as Image 将图形导出保存在磁盘中，图形文件格式可以在“保存类型”下拉列表中选择。图形的保存也可通过 edit→copy 粘贴于 photoshop，然后对图形进行编辑和保存。

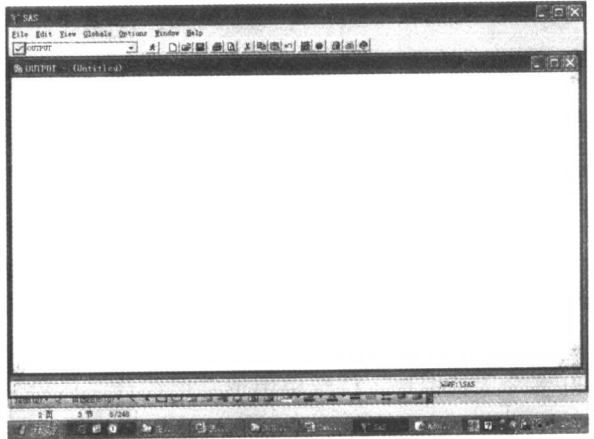


图 1.2 OUTPUT（输出）窗口

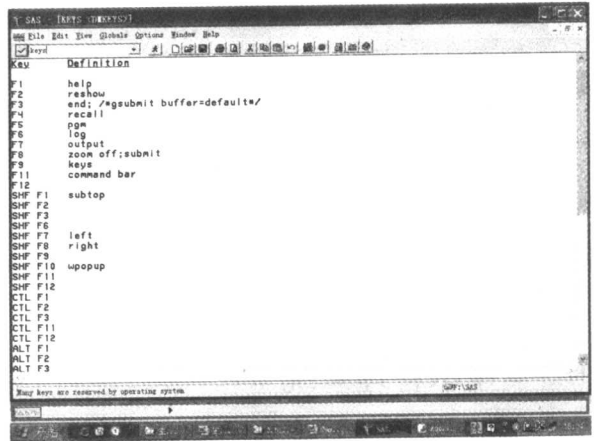


图 1.3 KEY（功能键）窗口



图 1.4 OPTIONS（选项）窗口

KEYS: 显示功能键的位置。

OPTIONS: 查看 SAS 系统设置。

1.3.2 几个常用的显示管理系统命令

不同窗口之间的转换, 可以通过在命令框内键入以下命令来实现:

KEYS: 进入 KEYS 窗口。

LOG: 进入 LOG 窗口。

OPTIONS: 进入 OPTIONS 窗口。

OUTPUT: 进入 OUTPUT 窗口。

PROGRAM: 进入 PROGRAM EDITOR 窗口。

键入 END 命令后, 退出当前窗口, 回复到 LOG 和 PGM 窗口。

除以上窗口之间转换命令外, 以下一些命令也是很有用的。

BYE: 退出 SAS 系统。

CLEAR: 清除当前工作窗口内容。

FILE (文件名): 将当前工作窗口中的内容储存到指定文件夹中。

HELP: 帮助。

NUMS: 显示 PGM 窗口数字区 (见图 1.5)。

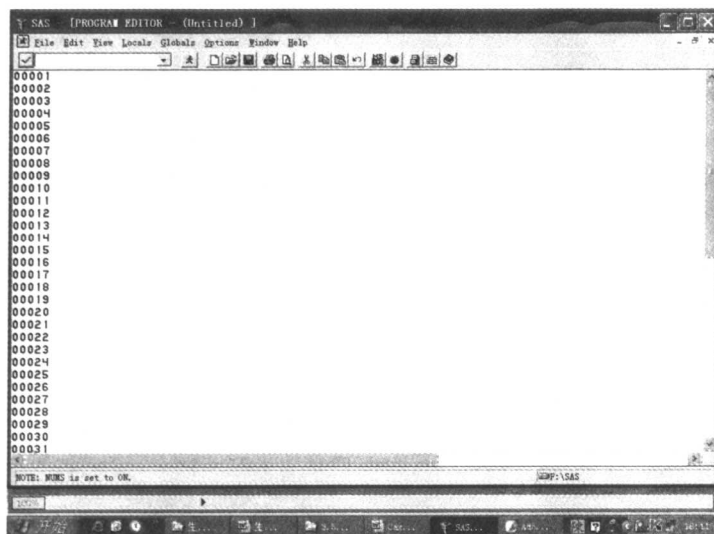


图 1.5 PGM (编辑) 窗口的数字区

RECALL: 调出上次执行过的程序, 以便校正、补充。

SUBMIT: 在 PGM 窗口输入程序后, 键入命令, 提交 SAS 执行这一程序。

ZOOM: 放大当前工作窗口, 使窗口充满整个屏幕, 重复执行命令, 回到原先显示的屏幕。

1.3.3 使用 pull down menu (下拉菜单)

启动 SAS 后, 在屏幕的上方有一个主窗口菜单栏 (图 1), 上述的各种命令, 可以通过选择菜单中相应项目来实现, 下面予以介绍。

BYE：将鼠标光标移至菜单栏中的“File”（文件）项，左键单击“File”，拉出“File”的子菜单，单击“Exit”（退出）（见图 1.6）。

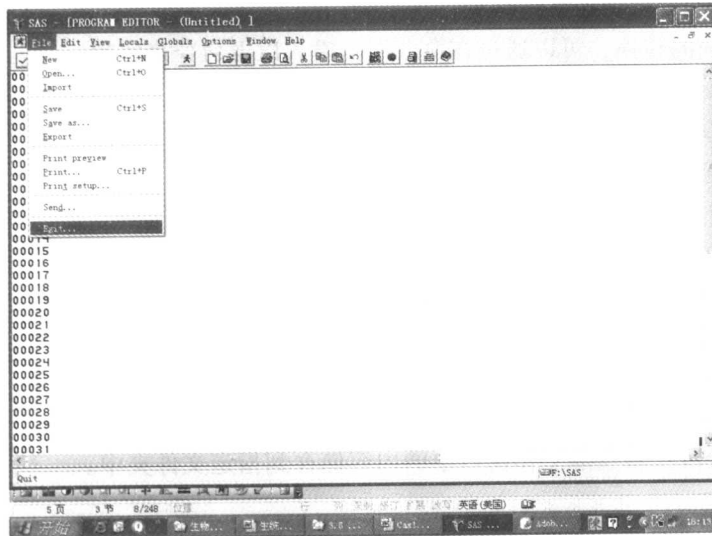


图 1.6 退出 SAS

CLEAR：切换到欲清除内容的窗口（将鼠标光标移至窗口内的任意一处，单击左键），单击菜单栏中的“Edit”（编辑），在下拉菜单中单击“Clear text”（见图 1.7）。

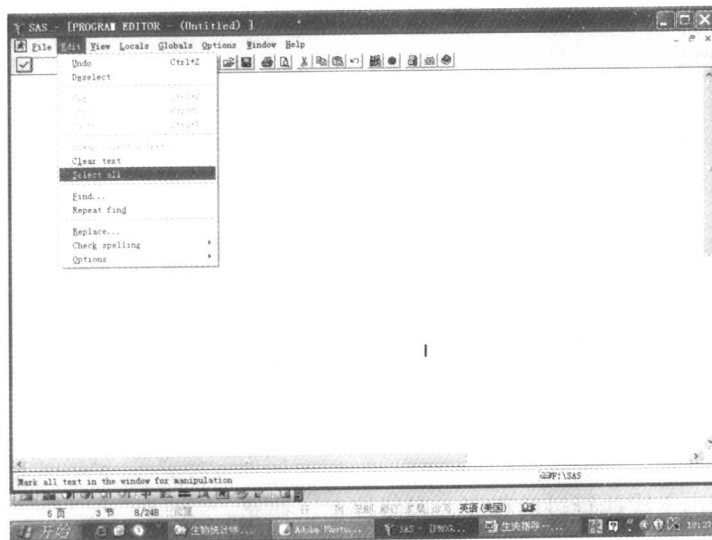


图 1.7 清除窗口内容

FILE：切换到欲存盘的窗口，单击“FILE”（文件），在下拉菜单中单击“Save as”（另存为），选定文件夹和文件存为类型，在对话框内键入文件名，再按“保存”键，窗口内容即被保存在指定的文件中（图见 1.8）。

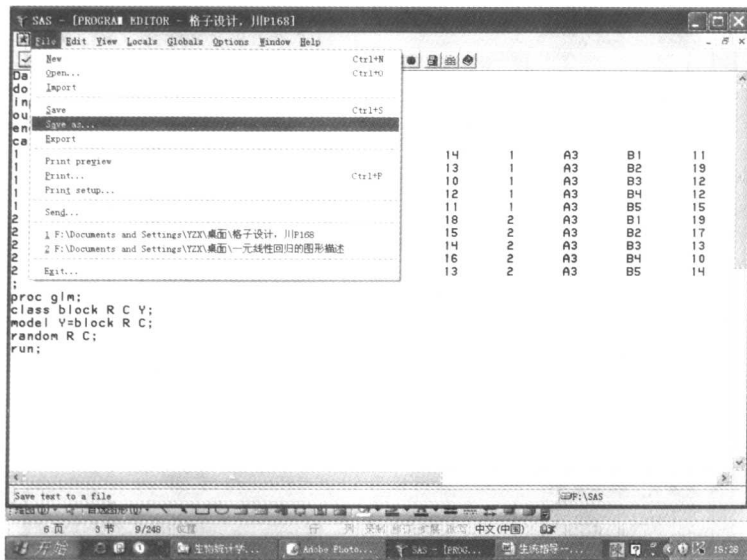


图 1.8 文件保存

HELP: 单击菜单栏中的“Help”（帮助）。

KEYS: 单击“Help”，在下拉菜单中单击“Keys”。

LOG: 单击“Globals”，在下拉菜单中单击“Log”（逻辑）（见图 1.9）。

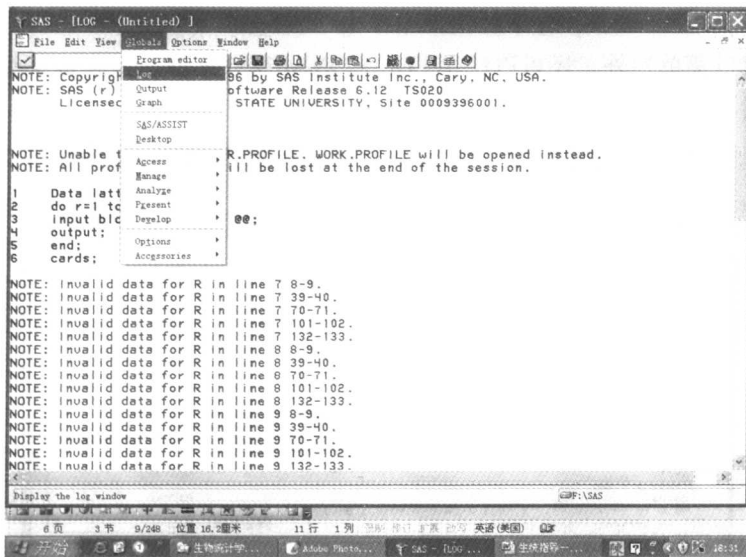


图 1.9 用下拉式菜单显示 LOG 窗口

NUMS: 切换到 PGM 窗口（数据编辑窗口），在菜单栏中单击“Edit”，选“Options”菜单下单击“Numbers”（见图 1.10）。

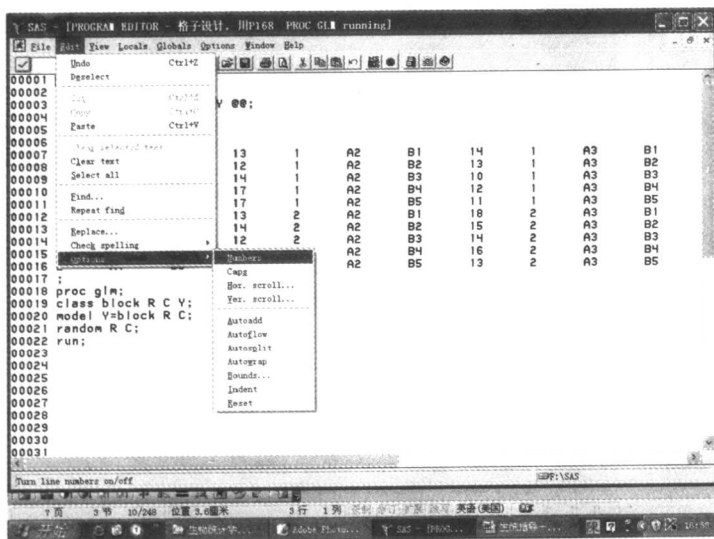


图 1.10 用下拉式菜单显示 PGM 窗口的数字区

OPTIONS: 单击菜单栏中的“Options”（选项）。

OUTPUT: 单击“Globals”，在下拉菜单中单击“Output”（数据输出）（见图 1.9）。

PROGRAM: 单击“Globals”，在下拉菜单中单击“Program editor”（数据编辑）（见图 1.9）。

RECALL: 切换到 PGM 窗口，单击“Locals”，在下拉菜单中单击“Recall text”。

SUBMIT: 单击“Locals”，在下拉菜单中单击“Submit”（提交）。

除使用主窗口菜单栏外，还可以使用 LOG 和 PGM 窗口菜单。方法是鼠标光标移至指定的窗口内，单击右键，即可拉出该窗口菜单（见图 1.11）。

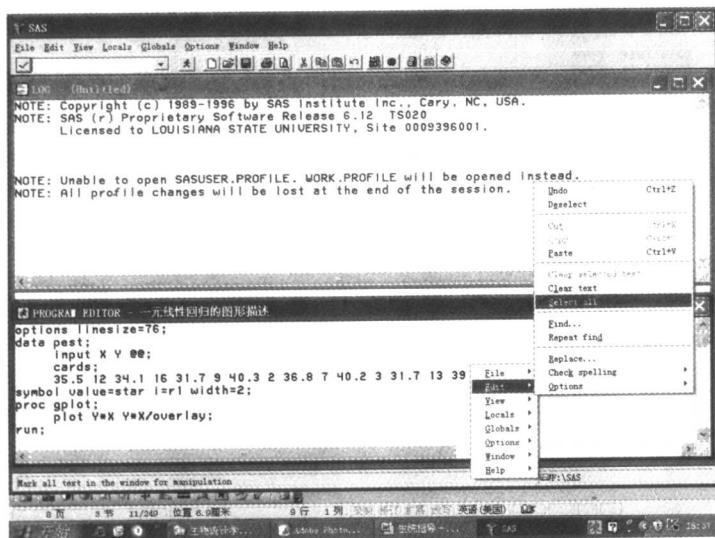


图 1.11 PGM 窗口和 LOG 窗口菜单

1.3.4 使用功能键和工具栏

除在命令框键入命令和使用下拉菜单操作外，还可以使用功能键和工具栏操作。关于各功能键的功能，可以从“KEYS”窗口查出，这里不再细述。在显示管理系统屏幕上，命令框的右边为工具栏，工具栏上有 15 个按钮，为了了解每一个按钮的功能，把鼠标光标移至按钮处，按钮的下方就会出现功能提示，读者可根据提示进行操作。

1.4 SAS 程序的编写

SAS 程序可划分成多个模块的话，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）。

SAS 程序由两类“操作步”构成，一种叫数据步（Data Step），一种叫过程步（Proc Step），数据步以 DATA 语句起始，过程步以 PROC 语句开始。下面表示一个 SAS 程序所包含的步骤：

数据步（Data Step）（数据准备模块）

在作业流中输入数据（直观、方便）或建立外部数据文件

将外部数据文件转换为 SAS 数据集

过程步（Proc Step）（数据分析处理模块）

调用一个或多个标准过程，由适当的过程语句分析处理数据

在一 SAS 程序中也可是有多个数据步和多个过程步。不论是 DATA 步还是 PROC 步，每一 SAS 语句之后都应以“;”结束。

以下逐项予以介绍。

1.5 数据步

输入所要分析的数据有两方法：（1）由外部数据文件建立 SAS 数据集；（2）在作业流中输入数据。

1.5.1 由外部数据文件建立 SAS 数据集

外部数据文件可以通过许多途径产生，可以用 Windows 中的记事本、写字板生成。但最方便的是在 SAS 系统的 PGM（数据编辑）窗口生成。但由于“由外部数据文件建立 SAS 数据集”，使用不方便且初学者不易掌握，故这里就不叙述。

1.5.2 在作业流中输入数据

若试验数据没有事先做成外部数据文件，也可以在用 SAS 分析处理数据的程序中输入。例如，在 5 种不同温度下研究一种微生物（*Diplodia natalensis* Evans）的生长和温度的关系，于接种后不同天数测量其生长速度，获得数据如下表 1.1，试进行方差分析。

表 1.1 不同温度与不同接种后天数某种微生物的生长速度

温 度	接种后天数			
	B ₁ (1 天)	B ₂ (2 天)	B ₃ (3 天)	B ₄ (4 天)
A ₁ (17.5)	0.3	1.3	2.6	3.5

续表 1.1

温 度	接种后天数			
	B ₁ (1 天)	B ₂ (2 天)	B ₃ (3 天)	B ₄ (4 天)
A ₂ (21.0)	0.3	1.7	2.9	4.0
A ₃ (24.5)	0.9	3.0	6.6	7.5
A ₄ (27.5)	1.7	4.8	9.0	9.0
A ₅ (30.5)	1.2	2.7	5.2	7.4

SAS 程序中分别用 A 代表温度, B 代表接种后天数, Y 为生长速度, SAS 程序如下:

```
options linesize = 76;
data micro;
input A $ B $ Y @@ ;
cards;
A1 B1 0.3 A1 B2 1.3 A1 B3 2.6 A1 B4 3.5
A2 B1 0.3 A2 B2 1.7 A2 B3 2.9 A2 B4 4.0
A3 B1 0.9 A3 B2 3.0 A3 B3 6.6 A3 B4 7.5
A4 B1 1.7 A4 B2 4.8 A4 B3 9.0 A4 B4 9.0
A5 B1 1.2 A5 B2 2.7 A5 B3 5.2 A5 B4 7.4
run;
proc anova;
class A B;
model Y = A B;
means A/duncan alpha =0.05;
means A/duncan alpha =0.01;
means B/duncan alpha =0.05;
means B/duncan alpha =0.01;
run;
```

这里的 DATA 语句和 INPUT 语句的功能前面已述及, CARDS 语句表示数据是作业流中的一部分。在作业流中输入数据后, 该程序可取一个文件名存在于磁盘中 (如在 D: \ 方差分析文件夹中存入“二因素双向分组 (交叉分组) 完全随机设计”文件), 以便下次调用该程序。

1.6 过程步

过程步是 SAS 用来处理分析数据的。处理不同类型统计问题需要不同的过程步 (PROC 步)。过程步 (PROC 步) 是以 PROC 语句为开头一个或一些 SAS 语句。每一个过程语句实际上是一个已经编好的一组程序的名字, 执行该语句即执行了这一组程序。下面仍以上例交叉分组试验设计的方差分析为例, 说明 PROC 步的操作。

切换到 PGM 窗口, 在 File→open 找到并打开“D: \ 方差分析 \ 二因素双向分组 (交叉分组) 完全随机设计”文件。整个程序输入格式见图 1.12。

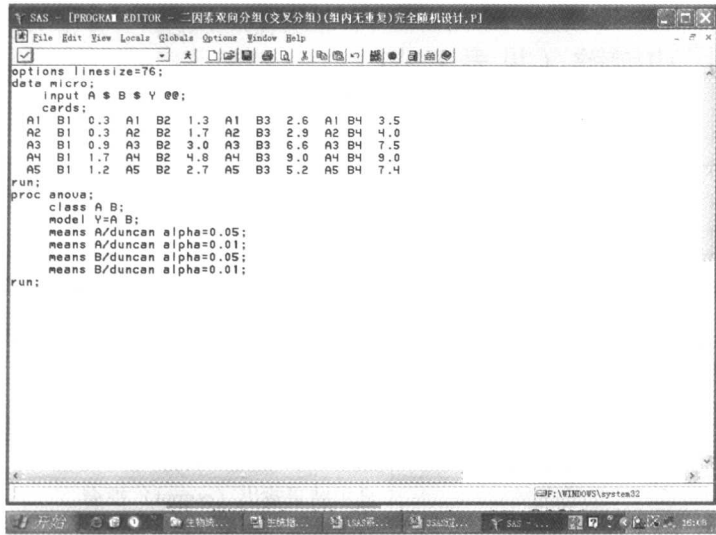


图 1.12 二因素双向分组（交叉分组）（组内无重复）完全随机设计方差分析 SAS 程序

为了不使行宽度太宽，超出屏幕范围不便阅览，采用语句 options linesize = 76，则把行宽限为 76。

1.7 执行程序

1.7.1 提交程序

单击工具栏上功能键  提交或用下拉菜单单击 submit 提交。

1.7.2 修改错误语句

如果在语句中出现其错误，在运行过程中 LOG 窗口即出现错误信息，例如，在上例程序中在 data micro 后忘记键入“;”，运行后在 LOG 窗口会出现以下信息（图 1.13）。

这时应调回原输入程序进行修改。调回的方法如下，切换到 PGM 窗口，在命令框键入 RECALL 命令，或使用功能键，先按 F5 再按 F4，调出原输入程序，检查产生错误原因，修改后，保存，关闭程后再重新打开程序，以清除原来错误的结果，重新运行。

1.7.3 保存

运行的结果和编写的程序可以储存起来。单击下拉菜单 file→save as（另存为）存入指定的文件夹中。

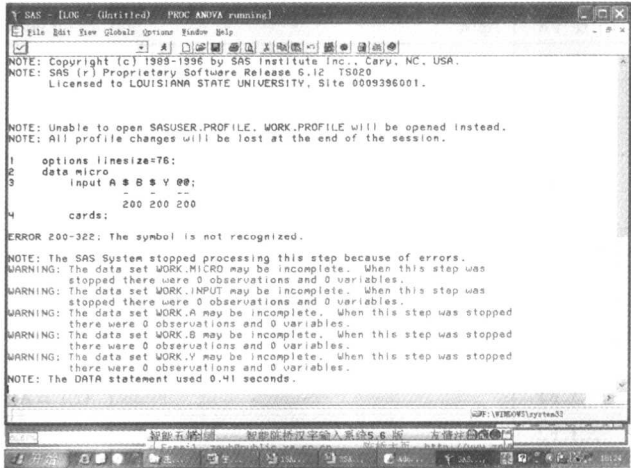


图 1.13 LOG 窗口显示错误信息

1.8 SAS 常用功能键（见表 1.2）

表 1.2 SAS for Windows 预设功能键

功能键 (key)	定义 (Definition)	说 明
F1	help	提供联机帮助, 打开 [help] 视窗
F2	reshow	再显示
F3	end; /* gsubmit buffer = default */	结束该视窗
F4	recall	调回上一次已运行的程序
F5	pgm	显示程序编辑 (program editor) 视窗
F6	log	显示记录 (log) 视窗
F7	output	显示结果输出 (output) 视窗
F8	zoom off; submit	放大或缩小视窗
F9	keys	打开功能键 (keys) 视窗
F10	command bar	打开命令 (指令) 条 (command bar)
F11		尚未定义
F12		
SHF F1	subtop	SHF = shift, SHF + F1
SHF F2		
SHF F3		
SHF F4		
SHF F5		
SHF F6		
SHF F7	left	数据左移 40 列
SHF F8	right	数据右移 40 列
SHF F9		
SHF F10	wpopup	弹出视窗
SHF F11		
SHF F12		
CTL F1		CLT = Ctrl, CLT F1 = ctrl + F1
CTL F2		
CTL F3		
CTL F11		
CTL F12		
ALT F1		ALT = Alt, ALT F1 = ALT + F1
ALT F2		
ALT F3		
ALT F11		
ALT F12		
CTL A		CTL A = CTL + A
CTL B	libname	打开库名 (libname) 视窗