



全国高等农林院校“十一五”规划教材

SAS 及应用

第二版

裴喜春 主编

 中国农业出版社

全国高等农林院校“十一五”规划教材

第二版

SAS及应用

■ 裴喜春 主编

中国农业出版社

图书在版编目 (CIP) 数据

SAS 及应用/裴喜春主编. —2 版. —北京: 中国农业出版社, 2007. 7

全国高等农林院校“十一五”规划教材

ISBN 978-7-109-11826-3

I. S… II. 裴… 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 (2007) 第 107205 号

010-62000830 (010) 62000830 : 客服电话

010-62000830 (010) 62000830 : 传真

E-mail: wixvay@163.com

地址: 北京市朝阳区农展馆北路 2 号

邮编: 100026

010-62000830 (010) 62000830 : 客服电话

客服电话

本图由北京农业出版社出版

中国农业出版社出版

(北京市朝阳区农展馆北路 2 号)

(邮政编码 100026)

责任编辑 朱 雷 甘敏敏

中国农业出版社印刷厂印刷 新华书店北京发行所发行

1998 年 3 月第 1 版 2007 年 7 月第 2 版

2007 年 7 月第 2 版北京第 1 次印刷

开本: 820mm×1080mm 1/16 印张: 18

字数: 417 千字

定价: 26.00 元

(凡本版图书出现印刷、装订错误, 请向出版社发行部调换)

内 容 简 介

.....

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、优化计算等多种 SAS 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。

第二版编写人员名单

.....

主 编 裴喜春

副 主 编 薛河儒

编写人员 (按姓氏笔画排序)

王玉顺 姚贵平 高利萍

裴喜春 薛河儒

第一版编写人员名单

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)

第二版前言

近年来,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受到用户的欢迎。

《SAS 及应用》作为全国高等农林院校教材,由裴喜春教授主编,于1998年3月出版第一版。第一版面世后,被许多高校选为教材或主要参考书。该教材以其简明扼要、易学易懂、注重适用的特点受到同行和读者的重视。由于 SAS 版本和理论的不断更新,以及教学需求的发展,编者在听取各方面意见后,参阅了国内外近年资料,对第一版内容做了调整和修订,力求突出 SAS 课程体系的先进性、适用性和启发性。

本次修订是第一版的继承和发展,在保留第一版体系结构的基础上,修订的重点为:

- (1) 将 SAS 软件版本提升为 SAS 9.0,教材中的讲述都是以 SAS 9.0 为蓝本。
- (2) 增加了 SAS 9.0 基本操作、交互式使用方法、GCHART、GPLOT、3D 绘图、多变量方差分析过程 MANOVA、析因试验设计过程 FACTEX、优化计算等内容。
- (3) 制作了与本教材配套的教学课件、实验指导等电子教学资料。

本教材面向理、工、农、林、医、管理等类的本科生及研究生,着重介绍了试验数据的统计分析计算方法,内容包括 SAS 操作及编程入门、标准化过程、方差分析、回归分析、聚类分析、主成分分析、因子分析、典型相关分析、判别分析及优化计算等 SAS 分析过程。编者在结合大量国内外统计分析应用成果及案例的基础上,选取了近百个实例作为教学例题,以案例驱动教学,力求起到举一反三、触类旁通的作用。书中的大量实例,不仅给出了解决实际问题的 SAS 编程方法,而且对计算结果进行了分析说明。

本教材根据授课对象的数学及统计基础,设计了不同的教学方案。对于具有计算机基础知识与试验统计方法知识的本专科生,可以用 20~30 学时讲授本书的第 1~7 章;对于具有多元统计分析知识的高年级本科生和研究生,可以用 40~50 学时(包括上机)学习本书的全部内容。书中带“*”号的内容可根据教学时数及要求灵活选用,也可作为读者阅读参考材料。

为便于教学,特制作了与本教材配套的教学课件、实验指导的电子资料。电子材料可向中国农业出版社教材中心索取。

在本教材的编写过程中,得到了国家教育部面向 21 世纪教改研究项目和中华农业科教基金的资助。研究生张春芳、刘俊丽同学在本书的程序调试、书稿整理等方面做了很多工作。在此谨向每一位曾关心和支持本教材编写工作的人士表示衷心的感谢!

由于编著者的水平有限,书中不当之处在所难免,请读者批评指正,使本书在使用中不断得到改进,日臻完善。编者 E-mail: xuehr@imau.edu.cn。

编 者

2007 年 1 月



第一版前言

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,但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 30 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、预测和财政管理计算等方面的内容。

本书从农林医类专业的实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 7 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

第二版前言

第一版前言

第 1 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
1.2.1 SAS 的结构	1
1.2.2 SAS 的特点	3
1.3 SAS 的启动与退出	3
1.3.1 启动 SAS	3
1.3.2 退出 SAS	3
1.4 SAS 主界面	3
1.4.1 SAS 的主要窗口	4
1.4.2 菜单栏	5
1.4.3 SAS 工具栏	8
1.4.4 SAS 逻辑库	8
1.5 实例	9
第 2 章 SAS 语言基础	12
2.1 SAS 语言概述	12
2.2 SAS 语句	13
2.2.1 SAS 关键字 (SAS Keywords)	13
2.2.2 SAS 名 (SAS Names)	13
2.2.3 本书的语句描述记号	13
2.3 SAS 数据集	14
2.3.1 数据值 (Data Values)	14
2.3.2 观测 (Observations)	14
2.3.3 变量 (Variables)	14
2.3.4 SAS 数据集存储空间计算	14
2.4 变量的属性	15
2.4.1 变量类型 (Types)	15

2.4.2	变量的长度 (Length)	15
2.4.3	变量的格式标号 (Labels)	15
2.4.4	数据缺值 (Missing Values)	15
2.5	SAS 程序的书写格式	16
2.5.1	空格	16
2.5.2	注释	16
2.5.3	变量的缩写记号	16
2.6	SAS 函数	16
2.6.1	函数的参数	17
2.6.2	函数值	17
2.6.3	函数的几点说明	17
2.6.4	算术函数	17
2.6.5	数学函数	18
2.6.6	截尾函数	18
2.6.7	三角和双曲函数	18
2.6.8	样本统计函数	19
2.6.9	概率函数	19
2.7	SAS 表达式	20
2.7.1	常量	20
2.7.2	操作符	21
2.7.3	操作符的执行优先顺序	22
2.7.4	表达式的几点说明	22
第 3 章	数据步与数据步语句	24
3.1	数据步基本概念	24
3.1.1	程序变量与数据集变量	24
3.1.2	数据向量	24
3.1.3	当前观测	24
3.2	数据步程序的 3 种基本结构	24
3.2.1	数据源是外部数据文件	25
3.2.2	数据源在源程序中	25
3.2.3	数据源是已存在的数据集	25
3.3	报告生成程序的基本结构	26
3.4	数据步程序的执行流程	26
3.4.1	当程序中有 INPUT 语句时	26
3.4.2	当程序中无 INPUT 语句时	27
3.5	数据步基本语句	27
3.5.1	DATA 语句	27
3.5.2	数据块与 CARDS 语句	28
3.5.3	INPUT 语句	28

12	3.5.4 文件操作语句	33
22	3.5.5 空语句	34
22	3.5.6 赋值语句	35
22	3.5.7 累加语句	35
22	3.5.8 复合语句	35
22	3.6 数据步程序控制语句	36
16	3.6.1 GOTO 语句与语句标号	36
16	3.6.2 STOP 语句	36
20	3.6.3 IF 语句	36
20	3.6.4 多向选择语句	38
20	3.6.5 DELETE 语句	39
22	3.7 循环语句	39
20	3.7.1 步长型循环语句	39
20	3.7.2 离散型循环语句	40
10	3.7.3 当型循环语句	40
10	3.7.4 直到型循环语句	40
10	3.8 数组概念及 ARRAY 语句	42
20	3.8.1 ARRAY 数组说明语句	42
20	3.8.2 关于数组的几点说明	43
20	3.8.3 SAS 数组应用举例	43
20	3.9 输出语句与输出文件	44
20	3.9.1 PUT 语句	44
20	3.9.2 打开输出文件语句	46
20	3.9.3 OUTPUT 语句	46
20	第 4 章 过程步与过程步语句	48
20	4.1 SAS 过程步的一般形式	48
20	4.2 常用过程步语句	49
20	4.2.1 VAR 语句	49
20	4.2.2 ID 语句	49
20	4.2.3 BY 语句	49
20	4.2.4 CLASS 语句	49
20	4.2.5 LABEL 语句	50
20	4.2.6 FREQ 语句	50
20	4.2.7 WEIGHT 语句	50
20	4.2.8 MODEL 语句	50
20	4.3 通用 SAS 语句	50
20	4.3.1 ENDSAS 语句	51
20	4.3.2 TITLE 语句	51
20	4.3.3 FOOTNOTE 语句	51

4.3.4	OPTIONS 语句	51
4.3.5	RUN 语句	52
4.3.6	X 语句	52
4.4	交互方式使用 SAS 的功能	52
4.4.1	数据集的建立与维护	52
4.4.2	数据分析的简单操作	57
第 5 章	通用标准过程	61
5.1	绘制平面图过程 GCHART (CHART)	61
5.1.1	频数条形图	62
5.1.2	百分数条形图	62
5.1.3	累积频数图	62
5.1.4	累积百分数图	63
5.1.5	分组合计条形图	63
5.1.6	分组均值条形图	63
5.1.7	再分组条形图	64
5.1.8	分组统计图	64
5.1.9	立体图	64
5.1.10	饼图	65
5.1.11	星图	65
5.1.12	GCHART 过程的一般形式	66
5.2	计算相关系数过程 CORR	69
5.2.1	CORR 过程格式	69
5.2.2	举例	70
5.3	SAS 与 dBASE 间数据转换过程 DBF	73
5.3.1	过程格式	73
5.3.2	说明	73
5.4	简单统计分析过程 MEANS	74
5.4.1	MEANS 过程的格式	75
5.4.2	举例	76
5.4.3	利用 MEANS 过程进行成对数据比较	76
5.5	二维散点图绘制过程 GPLOT (PLOT)	77
5.5.1	引例	77
5.5.2	GPLOT 过程的格式	78
5.5.3	等高线绘制过程 GCONTOUR	85
5.6	输出数据集过程 PRINT	86
5.6.1	PRINT 过程格式	86
5.6.2	举例	87
5.7	排序过程 SORT	87
5.7.1	SORT 过程的格式	87

5.7.2 举例	88
5.8 数据标准化过程 STANDARD	88
5.8.1 STANDARD 过程格式	88
5.8.2 举例	89
5.9 数据集转置过程 TRANSPOSE	90
5.9.1 TRANSPOSE 过程格式	90
5.9.2 举例	90
5.10 两组数据均值比较过程 TTEST	91
5.10.1 TTEST 过程格式	92
5.10.2 输出结果说明	92
5.10.3 举例	93
5.11 绘制三维图形过程 G3D	94
5.11.1 G3D 的格式	94
5.11.2 举例	95
第 6 章 方差分析	97
6.1 平衡设计的方差分析过程 ANOVA	97
6.1.1 概述	97
6.1.2 过程说明	97
6.1.3 使用说明	98
6.1.4 输出说明	99
6.1.5 举例	100
6.2 不平衡设计的方差分析过程 GLM	119
6.2.1 概述	119
6.2.2 过程说明	119
6.2.3 使用说明	120
6.2.4 输出说明	121
6.2.5 举例	121
6.3 多因变量方差分析 MANOVA	129
6.3.1 概述	129
6.3.2 过程说明	129
6.3.3 使用说明	130
6.3.4 输出说明	130
6.3.5 举例	130
6.4 嵌套设计的方差分析过程 NESTED	147
6.4.1 概述	147
6.4.2 语句说明	147
6.4.3 输出说明	148
6.4.4 举例	148
6.5 试验设计过程 PLAN	150

88	6.5.1 概述	150
88	6.5.2 语句说明	151
88	6.5.3 输出说明	152
88	6.5.4 举例	152
88	6.6 析因试验设计过程 FACTEX	156
88	6.6.1 概述	156
88	6.6.2 程序和语句说明	156
88	6.6.3 输出说明	158
88	6.6.4 举例	158
	第 7 章 回归分析	167
88	7.1 一般回归分析过程 REG	167
88	7.1.1 概述	167
88	7.1.2 过程说明	167
88	7.1.3 使用说明	168
88	7.1.4 输出说明	168
88	7.1.5 举例	169
88	7.2 多项式回归分析过程 RSREG	174
88	7.2.1 概述	174
88	7.2.2 过程说明	174
88	7.2.3 输出说明	175
88	7.2.4 举例	175
88	7.3 逐步回归分析过程 STEPWISE	177
88	7.3.1 概述	177
88	7.3.2 过程说明	177
88	7.3.3 输出说明	178
88	7.3.4 举例	179
88	7.4 非线性回归分析过程 NLIN	181
88	7.4.1 概述	181
88	7.4.2 过程说明	181
88	7.4.3 输出说明	183
88	7.4.4 举例	183
	第 8 章 聚类分析	186
88	8.1 系统聚类分析过程 CLUSTER	186
88	8.1.1 概述	186
88	8.1.2 过程说明	186
88	8.1.3 说明	187
88	8.1.4 输出说明	187
88	8.1.5 举例	189
88	8.2 逐步聚类分析过程 FASTCLUS	193

8.2.1	概述	193
8.2.2	过程说明	193
8.2.3	说明	195
8.2.4	输出说明	195
8.2.5	举例	195
8.3	变量聚类分析过程 VARCLUS	198
8.3.1	概述	198
8.3.2	过程说明	199
8.3.3	说明	200
8.3.4	输出说明	200
8.3.5	举例	201
第 9 章	多变量统计分析	205
9.1	主成分分析过程 PRINCOMP	205
9.1.1	概述	205
9.1.2	过程说明	205
9.1.3	说明	206
9.1.4	输出说明	206
9.1.5	举例	207
9.2	因子分析过程 FACTOR	210
9.2.1	概述	210
9.2.2	过程说明	210
9.2.3	说明	212
9.2.4	输出说明	213
9.2.5	举例	214
9.3	典型相关分析过程 CANCORR	218
9.3.1	概述	218
9.3.2	过程说明	218
9.3.3	输出说明	219
9.3.4	举例	220
第 10 章	判别分析	226
10.1	贝叶斯判别分析过程 DISCRIM	226
10.1.1	概述	226
10.1.2	过程说明	226
10.1.3	说明	228
10.1.4	输出说明	229
10.1.5	举例	229
10.2	典型判别分析过程 CANDISC	235
10.2.1	概述	235
10.2.2	过程说明	235

10.2.3	说明	236
10.2.4	输出说明	237
10.2.5	举例	238
10.3	逐步判别分析过程 STEPDISC	243
10.3.1	概述	243
10.3.2	过程说明	243
10.3.3	补充说明	244
10.3.4	输出说明	245
10.3.5	举例	245
第 11 章	SAS 优化计算	251
11.1	线性规划 LP	251
11.1.1	概述	251
11.1.2	过程说明	252
11.1.3	输出说明	253
11.1.4	举例	254
11.2	非线性规划和 NLP 过程	260
11.2.1	概述	260
11.2.2	过程说明	261
11.2.3	举例	262
参考文献		268