



全国高等农林院校“十三五”规划教材

SAS统计分析实验指导

赵洁 主编



 中国农业出版社

全国高等农林院校“十三五”规划教材

SAS统计分析实验指导

赵 洁 主编

中国农业出版社

图书在版编目 (CIP) 数据

SAS 统计分析实验指导 / 赵洁主编. —北京: 中国农业出版社, 2017. 7

全国高等农林院校“十三五”规划教材

ISBN 978-7-109-22850-4

I. ①S… II. ①赵… III. ①统计分析-应用软件-高等学校-教材 IV. ①C819

中国版本图书馆 CIP 数据核字 (2017) 第 071181 号

中国农业出版社出版

市朝阳区麦子店街 18 号楼)

(邮政编码 100125)

责任编辑 朱 雷

公司印刷 新华书店北京发行所发行

1 版 2017 年 7 月北京第 1 次印刷

开本: 787mm×1092mm 1/16 印张: 7.5

字数: 170 千字

定价: 18.00 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

SAS (Statistical Analysis System) 统计分析软件是一个集数据管理与数据处理功能于一体的大型软件系统。本书是与《SAS 统计分析》(赵洁主编) 配套的实验教材, 用于 SAS 统计分析相关课程的实践教学环节。本书共有 8 个实验, 主要包括 SAS 基本操作、SAS 数据步操作、SAS 数据集管理、过程步与常用过程步、SAS 统计图、方差分析、回归分析、SAS 多变量统计分析等, 在每个实验中重点介绍了在 SAS 系统环境下使用 SAS 过程的操作方法。

本书可作为理、工、农、林、医等专业相关课程的实验教材, 也可作为高校教师、研究生和高年级大学生从事数据统计分析的参考书。

编 写 人 员

主 编 赵 洁

副主编 郭 微 王会英

参 编 曹 然 高延英 杨宇姝

前 言

数据统计分析软件课程是我校计算机基础系列课程之一，通过本课程的学习使学生从理论和实践上掌握数据统计分析的方法及 SAS 统计分析软件的使用，为学生在今后实际工作中使用统计分析软件解决实际问题打下坚实的基础。本书是与《SAS 统计分析》（赵洁主编）配套的实验教材，用于 SAS 统计分析相关课程的实践教学环节。

全书共有 8 个实验，主要包括 SAS 基本操作、SAS 数据步操作、SAS 数据集管理、过程步与常用过程步、SAS 统计图、方差分析、回归分析、SAS 多变量统计分析等，在每个实验中重点介绍了在 SAS 系统环境下使用 SAS 过程的操作方法。其中，每个实验包括了知识要点、实验目的、参考题、实验题，实验中的“参考题”与“实验题”一一对应，难度略有提高。

本书由赵洁教授任主编，郭微、王会英任副主编。其中，王会英编写了实验 1 和实验 7；郭微编写了实验 5 和实验 6；赵洁编写了实验 3 和实验 8；曹然、高延英、杨宇姝编写了实验 2 和实验 4 并完成教材中材料的收集和部分试题的上机验证工作。全书由赵洁教授统稿。

本书主要为各院校非计算机专业学生编写。由于计算机技术发展速度快，SAS 软件也在不断更新，我们同样在不断学习和实践的过程中，再加上作者的写作和知识水平有限，书中难免存在不妥之处，恳请读者批评和指正。

作 者

2017 年 2 月

目 录

前言	
实验 1 SAS 基本操作	1
1.1 知识要点	1
1.2 实验目的	1
1.3 参考题	2
1.4 实验题	5
实验 2 SAS 数据步操作	6
2.1 知识要点	6
2.2 实验目的	8
2.3 参考题	8
2.4 实验题	15
实验 3 SAS 数据集管理	17
3.1 知识要点	17
3.2 实验目的	19
3.3 参考题	19
3.4 实验题	29
实验 4 过程步与常用过程步	31
4.1 知识要点	31
4.2 实验目的	33
4.3 参考题	33
4.4 实验题	43
实验 5 SAS 统计图	45
5.1 知识要点	45
5.2 实验目的	46
5.3 参考题	46
5.4 实验题	53
实验 6 方差分析	54
6.1 知识要点	54

6.2	实验目的	56
6.3	参考题	56
6.4	实验题	66
实验 7	回归分析	68
7.1	知识要点	68
7.2	实验目的	69
7.3	参考题	69
7.4	实验题	87
实验 8	SAS 多变量统计分析	91
8.1	知识要点	91
8.2	实验目的	92
8.3	参考题	93
8.4	实验题	108
参考文献	111

实验 1 SAS 基本操作

1.1 知识要点

1. SAS 的结构、功能和特点

SAS 是目前世界上公认的数据分析的标准软件之一。SAS 是一个用于数据管理和数据分析的组合软件系统。SAS 目前包括 30 多个功能模块，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。它具有操作简单、使用灵活方便、编程能力强且简单易学、数据处理和统计分析融为一体、功能强大、应用面广、实用性及扩展性强、结果专业等特点。

2. SAS 的开发环境

SAS 启动后默认状态下的主界面包括菜单栏、工具栏、命令行、状态栏、窗体条，以及 SAS 的几个主要窗口（SAS 浏览窗口、日志窗口、程序编辑器窗口、输出窗口）。

SAS 浏览窗口包括“结果”和“SAS 资源管理器”两个显示区域，可以方便地浏览和管理文件，同时在该窗体下还可以新建 SAS 文件，执行常用的文件操作；系统启动时，日志窗口显示 SAS 的版权和授权等信息。用户向系统提交任务后，日志窗口将给出 SAS 运行的状态；程序编辑器窗口用于 SAS 程序的编写；输出窗口在 SAS 程序提交运行后可将程序执行的详细结果显示出来，输出结果较多时，结果将以分页的形式展现；图形窗口显示所有由 SAS 的图形功能绘制的图形。

3. SAS 逻辑库

在 SAS 系统中的逻辑库可以分为临时逻辑库和永久逻辑库。SAS 每一次启动都将激活临时逻辑库 Work 和系统默认的永久逻辑库。临时逻辑库中存放的文件为临时文件，退出 SAS 后，该逻辑库下的文件将被清除。SAS 永久逻辑库中的数据文件为永久文件，每次启动 SAS 系统时都可以使用。

永久逻辑库即用户自定义逻辑库的创建有三种方法：

- (1) 菜单法。
- (2) 工具条法。
- (3) 用 LIBNAME 语句在程序编辑器窗口定义 SAS 逻辑库。

格式：LIBNAME 逻辑库名 “DOS 路径”；

其中：LIBNAME 是关键字；逻辑库名是用户给定的以字母或下划线开头的字符串；DOS 路径是一个存在的 DOS 目录路径。

4. SAS 语言基本概念

主要包括：SAS 语句、SAS 数据集、SAS 程序的书写格式、SAS 函数、SAS 表达式等。

1.2 实验目的

1. 熟悉了解 SAS 系统的结构、基本功能和特点。

2. 熟悉并掌握 SAS 的开发环境，学会使用 SAS 的几个主要窗口。
3. 熟练掌握 SAS 逻辑库的概念和分类，并熟练掌握创建永久逻辑库。
4. 了解和使用 SAS 软件中常用的窗口和常用命令。
5. 熟练掌握 SAS 语句、SAS 数据集的定义和使用；掌握 SAS 程序的书写格式、SAS 表达式、SAS 函数的使用；了解 SAS 程序的常见错误。

1.3 参考题

【参考题 1-1】利用 SAS 软件创建一个简单的 SAS 程序。程序功能为在临时逻辑库中新建一个数据表，数据表中的数据如表 1.1 所示。程序运行后在输出窗口查看输出结果。

表 1.1 学生成绩表一

Number	Name	Chinese	Math	English
A0215037	王亮	77	66	88
A0215038	李学	67	89	70
A0215039	张学平	89	68	95
A2015040	赵伟	98	87	67
A0215041	于心刚	90	88	67
A0215042	孙成庆	78	65	90

操作步骤：

(1) 启动 SAS 软件，在编辑器窗口输入如下程序：

```
DATA ck11;
INPUT Number $ Name $ Chinese Math English;
CARDS;
A0215037 王亮 77 66 88
A0215038 李学 67 89 70
A0215039 张学平 89 68 95
A2015040 赵伟 98 87 67
A0215041 于心刚 90 88 67
A0215042 孙成庆 78 65 90
;
PROC PRINT;
RUN;
```

(2) 鼠标单击“运行”菜单下的“提交”或单击 * 按钮，利用日志窗口查看 SAS 程序的运行状态，如果用户程序有错，系统在日志窗口以不同的颜色显示错误的大概位置及错误类型等信息。程序无错则在输出窗口有如下运行结果：

Obs	Number	Name	Chinese	Math	English
1	A0215037	王亮	77	66	88
2	A0215038	李学	67	89	70

3	A0215039	张学平	89	68	95
4	A2015040	赵伟	98	87	67
5	A0215041	于心刚	90	88	67
6	A0215042	孙成庆	78	65	90

(3) 点击 SAS 主窗口中的 SAS 资源管理器，在 SAS 资源管理器窗口中的逻辑库目录下的 Work 逻辑库下会出现 ck11 数据集。

【参考题 1-2】 分别利用菜单法和 LIBNAME 语句创建名称为 “why” 的 SAS 永久逻辑库，并把 “参考题 1-1” 中产生的 ck11 数据集放到 “why” 逻辑库中。

方法一：菜单条法。

(1) 启动 SAS 程序，在 SAS 主窗口中的 SAS 资源管理器窗口中逻辑库目录下单击鼠标右键，弹出如图 1.1 所示的菜单。

(2) 在弹出式菜单中单击 “新建” 菜单项，弹出 “新建逻辑库” 对话框，如图 1.2 所示。在对话框下逻辑库 “名称 (N)” 框内输入用户给定的逻辑库名如 “why”，选中 “启动时启用” 复选框（每次启动 SAS 时会启动新建的逻辑库），在逻辑库信息下的 “路径 (P)” 框内输入一个存放 SAS 数据文件的目录路径或鼠标单击 “浏览” 按钮指定 SAS 数据文件存放的目录路径，如 “E:\动医 1602 A0616012 张蕾”，单击 “确定” 按钮，完成永久逻辑库的创建。在 SAS 资源管理器窗口的逻辑库目录下产生一个名为 “why” 的逻辑库。

(3) 把 “参考题 1-1” 中的第一条语句改写为 DATA why.ck11，鼠标单击 “运行” 菜单下的 “提交” 或单击 * 按钮，数据集文件 ck11 就存放在 “why” 逻辑库下，同时在 “E:\动医 1602 A0616012 张蕾” 目录下也会出现 ck11 数据集文件。



图 1.1 逻辑库的创建

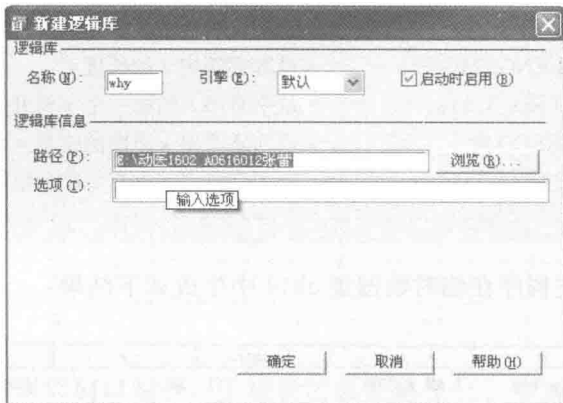


图 1.2 “新建逻辑库” 对话框

方法二：用 LIBNAME 语句创建名称为 “why” 的 SAS 永久逻辑库。

(1) 启动 SAS 程序，在编辑器窗口输入 “参考题 1-1” 所示的源程序。其中程序前两行语句更改为：

```
LIBNAME why "E:\动医 1602 A0616012 张蕾";
DATA why.ck11;
```

(2) 鼠标单击 “运行” 菜单下的 “提交” 或单击 * 按钮，数据集文件 ck11 就存放在

“why”逻辑库下，同时在“E:\动医 1602 A0616012 张蕾”目录下也会出现 ck11 数据集文件。

【参考题 1-3】数学函数的使用。

程序如下：

```
DATA ck13;
x1=MOD(8,3);
x2=SIN(30*3.1415/180);
x3=INT(6.24);          /* 截去数据小数部分取整 */
x4=FLOOR(2.45);       /* 返回小于等于数据的最大整数 */
x5=CEIL(2.33);        /* 返回大于等于数据的最小整数 */
x6=SUM(1,2,3);        /* 数据求和 */
x7=MEAN(1,2,3);       /* 数据求均值 */
RUN;
```

提交上述程序在临时数据集 ck13 中生成如下结果：

	x1	x2	x3	x4	x5	x6	x7
1	2	0.4999866265	6	2	3	6	2

【参考题 1-4】字符串函数的使用。

程序如下：

```
DATA ck14;
s=' I am a boy. ';
x=UPCASE(s);          /* 把字符串 s 转换成大写字母 */
Length=LENGTH(s);    /* 计算字符串 s 的长度 */
y=SUBSTR(s,3,2);     /* 从字符串 s 的第三个字符开始取两个字符 */
r=REVERSE(s);        /* 返回字符串 s 逆序的结果 */
m=REPEAT(s,2);       /* 返回字符串 s 重复 2 次的结果,包括原字符串一次 */
RUN;
```

提交上述程序在临时数据集 ck14 中生成如下结果：

	s	x	Length	y	r	m
1	I am a boy.	I AM A BOY.	11	am	.yob a ma I	I am a boy.I am a boy.I am a boy.

【参考题 1-5】在 SAS 程序编辑器窗口编程，实现对“参考题 1-2”生成的永久数据集 ck11 增加两个数据集变量，分别为每位同学各科总分和平均分。具体的程序如下：

程序 1：

```
DATA why.ck151;
SET why.ck11;        /* 把 why.ck11 数据集中的数据复制到 ck151 数据集中 */
Sum=Chinese+Math+English;
Aver=Sum/3;
RUN;
```

程序 2:

```
DATA why.ck152;
SET why.ck11;          /* 把 why.ck11 数据集中的数据复制到 ck152 数据集中 */
Sum=Sum(of Chinese Math English);
Aver=Mean(of Chinese Math English);
RUN;
```

提交上述程序在数据集 ck151、ck152 中生成如下结果:

	Number	Name	Chinese	Math	English	Sum	Aver
1	A0215037	王亮	77	66	88	231	77
2	A0215038	李学	67	89	70	226	75.333333333
3	A0215039	张学平	89	88	95	252	84
4	A2015040	赵伟	98	87	67	252	84
5	A0215041	于心刚	90	88	67	245	81.666666667
6	A0215042	孙成庆	78	65	90	233	77.666666667

1.4 实验题

【实验题 1-1】 利用 SAS 软件创建一个简单的 SAS 程序。程序功能是: 在 D 盘学生自己名字的文件夹下新建一个名称为 sy11 的数据集, 数据集数据如表 1.2 所示, 并在输出窗口输出。利用变量 Number 表示学号, Name 表示姓名, Sex 表示性别, Computer 表示计算机, English 表示英语, Math 表示高数。

表 1.2 学生成绩表二

Number	Name	Sex	Computer	English	Math
A01001	杜烨平	男	78	88	90
A01002	崔立伟	女	57	44	55
A01003	陈巍	女	80	80	79
A01004	杜庆怀	男	90	70	66
A01005	樊宇	女	44	64	50
A01006	高波	男	60	94	95

【实验题 1-2】 在 SAS 程序编辑器窗口编程, 生成新的数据集 sy12, 实现对“实验题 1-1”生成的 sy11 数据集里增加两个数据集变量 Sum 和 Aver, 分别为每位同学各科总分和平均分。程序运行结果如下:

	Number	Name	Sex	Computer	English	Math	Sum	Aver
1	A01001	杜烨平	男	78	88	90	256	85.333333333
2	A01002	崔立伟	女	57	44	55	156	52
3	A01003	陈巍	女	80	80	79	239	79.666666667
4	A01004	杜庆怀	男	90	70	66	226	75.333333333
5	A01005	樊宇	女	44	64	50	158	52.666666667
6	A01006	高波	男	60	94	95	249	83

实验 2 SAS 数据步操作

2.1 知识要点

1. 数据步的基本概念

数据步是产生 SAS 数据集的一组 SAS 语句，一个 SAS 程序中可以有若干个数据步。数据集包括变量、观测值等。

2. 数据步程序的三种基本结构

根据数据源的不同，数据步的程序有如下三种基本结构：

(1) 数据源是外部数据文件。

```
DATA 语句;  
[FILENAME 语句;]  
INFILE 语句;  
INPUT 语句;  
[其他数据步语句;]  
;
```

(2) 数据源在 SAS 程序中。

```
DATA 语句;  
INPUT 语句;  
[其他数据步语句;]  
CARDS;  
用空格分隔的多行数据(数据行)  
;
```

(3) 数据源是已存在的数据集。

```
DATA 语句;  
SET 或 MERGE 或 UPDATE 语句;  
[其他数据步语句;]  
;
```

3. 数据步的基本语句

SAS 的基本语句有：DATA 语句、数据块与 CARDS 语句、INPUT 语句、文件操作语句、空语句、赋值语句、累加语句、复合语句。

4. 数据步程序控制语句

数据步程序控制语句有：STOP 语句、IF 语句、多向选择语句、DELETE 语句。其中：

(1) IF 语句。

格式 1:

```
IF 表达式 THEN SAS 语句;
```

格式 2:

```
IF 表达式 THEN SAS 语句 1;
ELSE SAS 语句 2;
```

格式 3:

```
IF 表达式 THEN;
DO;
语句;
...
END;
ELSE [IF THEN];
DO;
语句;
...
END;
```

(2) 多向选择语句。

格式:

```
SELECT [(表达式 0)];
WHEN (表达式 1) 语句 1;
WHEN (表达式 2) 语句 2;
:
WHEN (表达式 n) 语句 n;
[OTHERWISE 语句 n+1;]
END;
```

(3) 循环语句。

步长型循环语句的格式:

```
DO 循环变量=初值 [TO 终值] [BY 增量];
循环体
END;
```

离散型循环语句的格式:

```
DO 循环变量=值 1, 值 2, ..., 值 i, ..., 值 n;
循环体
END;
```

当型循环语句格式：

```
DO WHILE (表达式);
    循环体
END;
```

直到型循环语句格式：

```
DO UNTIL (表达式);
    循环体
END;
```

5. 输出语句与输出文件语句

输出语句与输出文件语句有：PUT 语句、打开输出文件语句、OUTPUT 语句。

2.2 实验目的

1. 掌握 SAS 程序的基本结构，学会使用数据步的基本语句编写简单的 SAS 程序。
2. 掌握数据步程序控制语句和输出语句的使用，学会编写相关的 SAS 程序。

2.3 参考题

【参考题 2-1】我们有一个职工的基本信息表，包括姓名 (Name)、性别 (Sex)、年龄 (Age)、身高 (Height) 和体重 (Weight)，如表 2.1 所示。要求利用 INPUT 语句的列格式输入，产生一个 SAS 数据集。

表 2.1 职工的基本信息表

Name	Sex	Age	Height	Weight
杨柏松	男	20	1.85	77
周小雪	女	36	1.65	60
杨倩宇	女	45	1.75	68
周天乐	女	38	1.59	68
刘晓鑫	男	24	1.68	59
刘卫东	男	26	1.73	57
张丹	女	28	1.87	95

操作步骤：

首先，要确定 INPUT 语句的列格式，1~8 列是变量 Name \$，10~12 列是变量 Sex \$，13~15 列是变量 Age，18~21 列是变量 Height 的位置，23~25 列是变量 Weight 的位置，然后编写 SAS 程序。

程序如下：

```
DATA ck21;
INPUT Name$ 1-8 Sex$ 10-12 Age 13-15 Height 18-21 Weight 23-25;
CARDS;
杨柏松 男 20 1.85 77
周小雪 女 36 1.65 60
杨倩宇 女 45 1.75 68
周天乐 女 38 1.59 68
刘晓鑫 男 24 1.68 59
刘卫东 男 26 1.73 57
张丹 女 28 1.87 95
;
PROC PRINT;
RUN;
```

程序运行结果如下：

```
SAS 系统 2016年10月17日 星期一 下午02时24分32秒10
Obs Name Sex Age Height Weight
1 杨柏松 男 20 1.85 77
2 周小雪 女 36 1.65 60
3 杨倩宇 女 45 1.75 68
4 周天乐 女 38 1.59 68
5 刘晓鑫 男 24 1.68 59
6 刘卫东 男 26 1.73 57
7 张丹 女 28 1.87 95
```

【参考题 2-2】 我们已用文字处理软件建立了文件名为“ck21.txt”的文件，文件的内容是表 2.1 职工的基本信息表，文件保存在“E:\user\ck21.txt”，要求利用 SAS 数据步程序中相关的结构和文件操作语句，完成 SAS 永久的数据集的建立，假定数据集保存在“E:\user”目录下。

操作步骤：

首先，我们可以利用 LIBNAME 语句指定逻辑库，再利用数据在外部文件中 SAS 数据步的对应结构，使用 FILENAME 语句和 INFILE 语句，完成 SAS 数据集的建立。

编写 SAS 程序，以下两个程序是等价的。

程序 1：

```
LIBNAME xx 'E:\user';
FILENAME ck2 'E:\user\ck21.txt';
DATA xx.ck22;
INFILE ck2;
INPUT Name$ Sex$ Age Height Weight;
;
PROC PRINT;
RUN;
```