



国家卫生和计划生育委员会“十二五”规划教材  
全国高等医药教材建设研究会“十二五”规划教材  
科研人员核心能力提升导引丛书  
供研究生及科研人员用

第**3**版

# SAS 统计软件应用

*Practical Applications of SAS*

主 编 贺 佳  
副主编 尹 平

 人民卫生出版社



国家卫生和计划生育委员会“十二五”规划教材  
全国高等医药教材建设研究会“十二五”规划教材  
科研人员核心能力提升导引丛书  
供研究生及科研人员用

# SAS 统计软件应用

## Practical Applications of SAS

### 第 3 版

主 编 贺 佳

副主编 尹 平

编 者 (以姓氏笔画为序)

马 骏 (天津医科大学)	李婵娟 (第四军医大学)
王 玖 (滨州医学院)	杨士保 (中南大学)
王 彤 (山西医科大学)	时松和 (郑州大学)
方 亚 (厦门大学)	陆 健 (第二军医大学)
尹 平 (华中科技大学)	贺 佳 (第二军医大学)
邓 伟 (复旦大学)	钱 聪 (中国医科大学)
石武祥 (桂林医学院)	隋 虹 (哈尔滨医科大学)
史周华 (山东中医药大学)	景学安 (泰山医学院)
刘军祥 (泸州医学院)	曾 庆 (重庆医科大学)
闫小妍 (北京大学)	谭旭辉 (南方医科大学)
李 强 (西安交通大学)	

秘 书 (以姓氏笔画为序)

张新佶 (第二军医大学)	蒋红卫 (华中科技大学)
高青斌 (第二军医大学)	



人民卫生出版社  
PEOPLE'S MEDICAL PUBLISHING HOUSE

图书在版编目(CIP)数据

SAS 统计软件应用 / 贺佳主编. —3 版. —北京: 人民卫生出版社, 2014.4

ISBN 978-7-117-18602-5

I. ①S… II. ①贺… III. ①医学统计—统计分析—应用软件—医学院校—教材 IV. ①R195.1-39

中国版本图书馆 CIP 数据核字(2014)第 031636 号

人卫社官网	<a href="http://www.pmph.com">www.pmph.com</a>	出版物查询, 在线购书
人卫医学网	<a href="http://www.ipmph.com">www.ipmph.com</a>	医学考试辅导, 医学数据库服务, 医学教育资源, 大众健康资讯

版权所有, 侵权必究!

SAS 统计软件应用

(第 3 版)

主 编: 贺 佳

出版发行: 人民卫生出版社(中继线 010-59780011)

地 址: 北京市朝阳区潘家园南里 19 号

邮 编: 100021

E - mail: [pmph@pmph.com](mailto:pmph@pmph.com)

购书热线: 010-59787592 010-59787584 010-65264830

印 刷: 潮河印业有限公司

经 销: 新华书店

开 本: 850×1168 1/16 印张: 14

字 数: 423 千字

版 次: 2006 年 9 月第 1 版 2014 年 4 月第 3 版

2014 年 4 月第 3 版第 1 次印刷(总第 4 次印刷)

标准书号: ISBN 978-7-117-18602-5/R·18603

定价(含光盘): 59.00 元

打击盗版举报电话: 010-59787491 E-mail: [WQ@pmph.com](mailto:WQ@pmph.com)

(凡属印装质量问题请与本社市场营销中心联系退换)

---

## 主编简介

---



贺佳, 教授, 博士生导师。现任第二军医大学卫生统计学教研室主任; 上海市杨浦区政协委员; 国际生物统计学会中国分会副理事长、中国卫生信息学会卫生统计教育专业委员会副主任委员、中国现场统计研究会生物医学统计学会副理事长、中国人民解放军医学科学技术委员会卫生信息学专业委员会副主任委员、上海市临床流行病学专业委员会副主任委员等。国际知名杂志 *The Lancet* 的统计学审稿专家, 《中国卫生统计》等十余种杂志的常务编委和编委。

从事教学工作近三十年, 负责的课程获国家精品课程、上海市精品课程、军队优质课程。主编、副主编教材 18 部, 其中 7 部由人民卫生出版社出版, 2 部由高等教育出版社出版。负责国家自然科学基金、上海基础研究重点项目、军队科技攻关等 30 余项科研基金课题, 以第一作者或通讯作者发表 SCI 论文 63 篇, 特别是在 *The Lancet* 上发表评述。自主研发软件 9 套, 获国家软件著作权。以第一完成人获上海市科技进步一等奖、二等奖和上海医学科技二等奖。为上海市领军人才、上海市优秀学科带头人、上海市三八红旗手标兵、上海市三八红旗手、上海市高校优秀青年教师。

# 全国高等学校医学研究生规划教材

## 第二轮修订说明

为了推动医学研究生教育的改革与发展,加强创新人才培养,自2001年8月全国高等医药教材建设研究会和原卫生部教材办公室启动医学研究生教材的组织编写工作开始,在多次大规模的调研、论证的前提下,人民卫生出版社先后于2002年和2008年分两批完成了第一轮五十余种医学研究生规划教材的编写与出版工作。

为了进一步贯彻落实第二次全国高等医学教育改革工作会议精神,推动“5+3”为主体的临床医学教育综合改革,培养研究型、创新性、高素质的卓越医学人才,全国高等医药教材建设研究会、人民卫生出版社在全面调研、系统分析第一轮研究生教材的基础上,再次对这套教材进行了系统的规划,进一步确立了以“解决研究生科研和临床中实际遇到的问题”为立足点,以“回顾、现状、展望”为线索,以“培养和启发研究生创新思维”为中心的教材创新修订原则。

修订后的第二轮教材共包括5个系列:①科研公共学科系列:主要围绕研究生科研中所需要的基本理论知识,以及从最初的科研设计到最终的论文发表的各个环节可能遇到的问题展开;②常用统计软件与技术介绍了SAS统计软件、SPSS统计软件、分子生物学实验技术、免疫学实验技术等常用的统计软件以及实验技术;③基础前沿与进展:主要包括了基础学科中进展相对活跃的学科;④临床基础与辅助学科:包括了临床型研究生所需要进一步加强的相关学科内容;⑤临床专业学科:通过对疾病诊疗历史变迁的点评、当前诊疗中困惑、局限与不足的剖析,以及研究热点与发展趋势探讨,启发和培养临床诊疗中的创新。从而构建了适应新时期研究型、创新性、高素质、卓越医学人才培养的教材体系。

该套教材中的科研公共学科、常用统计软件与技术学科适用于医学院校各专业的研究生及相应的科研工作者,基础前沿与进展主要适用于基础医学和临床医学的研究生及相应的科研工作者;临床基础与辅助学科和临床专业学科主要适用于临床型研究生及相应学科的专科医师。

# 全国高等学校第二轮医学研究生规划教材目录

1	医学哲学	主 编 柯 杨 张大庆 副主编 赵明杰 段志光 罗长坤 刘 虹
2	医学科研方法学(第2版)	主 编 刘 民 副主编 陈 峰
3	医学统计学(第4版)	主 编 孙振球 徐勇勇
4	医学实验动物学(第2版)	主 编 秦 川 副主编 谭 毅 张连峰
5	实验室生物安全(第2版)	主 审 余新炳 主 编 叶冬青
6	医学科研课题设计、申报与实施(第2版)	主 审 龚非力 主 编 李卓娅 副主编 李宗芳
7	医学信息搜集与利用(第2版)	主 编 代 涛 副主编 赵文龙 张云秋
8	医学实验技术原理与选择(第2版)	主 编 魏于全 副主编 向 荣 郭亚军 胡 汛 徐宁志
9	统计方法在医学科研中的应用	主 编 李晓松 副主编 李 康
10	医学科研论文撰写与发表(第2版)	主 编 张学军 副主编 王征爱 吴忠均
11	IBM SPSS 统计软件应用(第3版)	主 编 陈平雁 黄浙明 副主编 安胜利 欧春泉 陈莉雅
12	SAS 统计软件应用(第3版)	主 编 贺 佳 副主编 尹 平

- |    |                  |                                       |
|----|------------------|---------------------------------------|
| 13 | 医学分子生物学实验技术(第3版) | 主 编 药立波<br>副主编 韩 骅 焦炳华 常智杰            |
| 14 | 医学免疫学实验技术(第2版)   | 主 编 柳忠辉 吴雄文<br>副主编 王全兴 吴玉章 储以微        |
| 15 | 组织病理技术(第2版)      | 主 编 李甘地                               |
| 16 | 组织和细胞培养技术(第3版)   | 主 审 宋今丹<br>主 编 章静波<br>副主编 张世馥 连小华     |
| 17 | 组织化学与细胞化学技术(第2版) | 主 编 李 和 周 莉<br>副主编 周德山 周国民 肖 岚        |
| 18 | 人类疾病动物模型(第2版)    | 主 审 施新猷<br>主 编 刘恩岐<br>副主编 李亮平 师长宏     |
| 19 | 医学分子生物学(第2版)     | 主 审 刘德培<br>主 编 周春燕 冯作化<br>副主编 药立波 何凤田 |
| 20 | 医学免疫学            | 主 编 曹雪涛<br>副主编 于益芝 熊思东                |
| 21 | 基础与临床药理学(第2版)    | 主 编 杨宝峰<br>副主编 李学军 李 俊 董 志            |
| 22 | 医学微生物学           | 主 编 徐志凯 郭晓奎<br>副主编 江丽芳 龙北国            |
| 23 | 病理学              | 主 编 来茂德<br>副主编 李一雷                    |
| 24 | 医学细胞生物学(第3版)     | 主 审 钟正明<br>主 编 杨 恬<br>副主编 易 静 陈誉华 何通川 |
| 25 | 分子病毒学(第3版)       | 主 编 黄文林<br>副主编 徐志凯 董小平 张 辉            |
| 26 | 医学微生态学           | 主 编 李兰娟                               |
| 27 | 临床流行病学(第4版)      | 主 审 李立明<br>主 编 黄悦勤                    |
| 28 | 循证医学             | 主 编 李幼平<br>副主编 杨克虎                    |

- |    |             |   |
|----|-------------|---|
| 29 | 断层影像解剖学     | 主 编 刘树伟<br>副主编 张绍祥 赵 斌                        |
| 30 | 临床应用解剖学     | 主 编 王海杰<br>副主编 陈 尧 杨桂姣                        |
| 31 | 临床信息管理      | 主 编 崔 雷<br>副主编 曹高芳 张 晓 郑西川                    |
| 32 | 临床心理学       | 主 审 张亚林<br>主 编 李占江<br>副主编 王建平 赵旭东 张海音         |
| 33 | 医患沟通        | 主 编 周 晋<br>副主编 尹 梅                            |
| 34 | 实验诊断学       | 主 编 王兰兰 尚 红<br>副主编 尹一兵 樊绮诗                    |
| 35 | 核医学(第2版)    | 主 编 张永学<br>副主编 李亚明 王 铁                        |
| 36 | 放射诊断学       | 主 编 郭启勇<br>副主编 王晓明 刘士远                        |
| 37 | 超声影像学       | 主 审 张 运 王新房<br>主 编 谢明星 唐 杰<br>副主编 何怡华 田家玮 周晓东 |
| 38 | 呼吸病学(第2版)   | 主 审 钟南山<br>主 编 王 辰 陈荣昌<br>副主编 代华平 陈宝元         |
| 39 | 消化内科学(第2版)  | 主 审 樊代明 刘新光<br>主 编 钱家鸣<br>副主编 厉有名 林菊生         |
| 40 | 心血管内科学(第2版) | 主 编 胡大一 马长生<br>副主编 雷 寒 韩雅玲 黄 峻                |
| 41 | 血液内科学(第2版)  | 主 编 黄晓军 黄 河<br>副主编 邵宗鸿 胡 豫                    |
| 42 | 肾内科学(第2版)   | 主 编 谌贻璞<br>副主编 余学清                            |
| 43 | 内分泌内科学(第2版) | 主 编 宁 光 周智广<br>副主编 王卫庆 邢小平                    |



- |    |            |     |                    |
|----|------------|-----|--------------------|
| 44 | 风湿内科学(第2版) | 主 编 | 陈顺乐 邹和健            |
| 45 | 急诊医学(第2版)  | 主 编 | 黄子通 于学忠            |
|    |            | 副主编 | 吕传柱 陈玉国 刘 志        |
| 46 | 神经内科学(第2版) | 主 编 | 刘 鸣 谢 鹏            |
|    |            | 副主编 | 崔丽英 陈生弟 张黎明        |
| 47 | 精神病学(第2版)  | 主 审 | 江开达                |
|    |            | 主 编 | 马 辛                |
|    |            | 副主编 | 施慎逊 许 毅            |
| 48 | 感染病学(第2版)  | 主 编 | 李兰娟 李 刚            |
|    |            | 副主编 | 王宇明 陈士俊            |
| 49 | 肿瘤学(第4版)   | 主 编 | 曾益新                |
|    |            | 副主编 | 吕有勇 朱明华 陈国强<br>龚建平 |
| 50 | 老年医学(第2版)  | 主 编 | 张 建 范 利            |
|    |            | 副主编 | 华 琦 李为民 杨云梅        |
| 51 | 临床变态反应学    | 主 审 | 叶世泰                |
|    |            | 主 编 | 尹 佳                |
|    |            | 副主编 | 洪建国 何韶衡 李 楠        |
| 52 | 危重症医学      | 主 编 | 王 辰 席修明            |
|    |            | 副主编 | 杜 斌 于凯江 詹庆元<br>许 媛 |
| 53 | 普通外科学(第2版) | 主 编 | 赵玉沛 姜洪池            |
|    |            | 副主编 | 杨连粤 任国胜 陈规划        |
| 54 | 骨科学(第2版)   | 主 编 | 陈安民 田 伟            |
|    |            | 副主编 | 张英泽 郭 卫 高忠礼<br>贺西京 |
| 55 | 泌尿外科学(第2版) | 主 审 | 郭应禄                |
|    |            | 主 编 | 杨 勇 李 虹            |
|    |            | 副主编 | 金 杰 叶章群            |
| 56 | 胸心外科学      | 主 编 | 胡盛寿                |
|    |            | 副主编 | 孙立忠 王 俊 庄 建        |
| 57 | 神经外科学(第2版) | 主 审 | 周良辅                |
|    |            | 主 编 | 赵继宗 周定标            |
|    |            | 副主编 | 王 硕 毛 颖 张建宁<br>王任直 |

- |    |                |   |
|----|----------------|---|
| 58 | 血管淋巴管外科学(第2版)  | 主 编 汪忠镐<br>副主编 王深明 俞恒锡                    |
| 59 | 小儿外科学(第2版)     | 主 审 王 果<br>主 编 冯杰雄 郑 珊<br>副主编 孙 宁 王维林 夏慧敏 |
| 60 | 器官移植学          | 主 审 陈 实<br>主 编 刘永锋 郑树森<br>副主编 陈忠华 朱继业 陈江华 |
| 61 | 临床肿瘤学          | 主 编 赫 捷<br>副主编 毛友生 沈 铿 马 骏                |
| 62 | 麻醉学            | 主 编 刘 进<br>副主编 熊利泽 黄宇光                    |
| 63 | 妇产科学(第2版)      | 主 编 曹泽毅 乔 杰<br>副主编 陈春玲 段 涛 沈 铿<br>王建六 杨慧霞 |
| 64 | 儿科学            | 主 编 桂永浩 申昆玲<br>副主编 毛 萌 杜立中                |
| 65 | 耳鼻咽喉头颈外科学(第2版) | 主 编 孔维佳 韩德民<br>副主编 周 梁 许 庚 韩东一            |
| 66 | 眼科学(第2版)       | 主 编 崔 浩 王宁利<br>副主编 杨培增 何守志 黎晓新            |
| 67 | 灾难医学           | 主 审 王一镗<br>主 编 刘中民<br>副主编 田军章 周荣斌 王立祥     |
| 68 | 康复医学           | 主 编 励建安<br>副主编 毕 胜                        |
| 69 | 皮肤性病学          | 主 编 王宝玺<br>副主编 顾 恒 晋红中 李 岷                |
| 70 | 创伤、烧伤与再生医学     | 主 审 王正国 盛志勇<br>主 编 付小兵<br>副主编 黄跃生 蒋建新     |

# 全国高等学校第二轮医学研究生规划教材 评审委员会名单

## 顾 问

韩启德 桑国卫 陈 竺 赵玉沛

## 主任委员

刘德培

## 副主任委员 (以汉语拼音为序)

曹雪涛 段树民 樊代明 付小兵 郎景和 李兰娟 王 辰  
魏于全 杨宝峰 曾益新 张伯礼 张 运 郑树森

## 常务委员 (以汉语拼音为序)

步 宏 陈安民 陈国强 冯晓源 冯友梅 桂永浩 柯 杨  
来茂德 雷 寒 李 虹 李立明 李玉林 吕兆丰 瞿 佳  
田勇泉 汪建平 文历阳 闫剑群 张学军 赵 群 周学东

## 委 员 (以汉语拼音为序)

毕开顺 陈红专 崔丽英 代 涛 段丽萍 龚非力 顾 晋  
顾 新 韩德民 胡大一 胡盛寿 黄从新 黄晓军 黄悦勤  
贾建平 姜安丽 孔维佳 黎晓新 李春盛 李 和 李小鹰  
李幼平 李占江 栗占国 刘树伟 刘永峰 刘中民 马建辉  
马 辛 宁 光 钱家鸣 乔 杰 秦 川 尚 红 申昆玲  
沈志祥 谌贻璞 石应康 孙 宁 孙振球 田 伟 汪 玲  
王 果 王兰兰 王宁利 王深明 王晓民 王 岩 谢 鹏  
徐志凯 杨东亮 杨 恬 药立波 尹 佳 于布为 余祥庭  
张奉春 张 建 张祥宏 章静波 赵靖平 周春燕 周定标  
周 晋 朱正纲

# 前 言

当今,我们正快步进入“大数据”时代,数据的收集、整理、分析、利用都离不开统计分析软件。SAS 软件是国际上最著名、最常用的统计分析软件之一,可以满足广泛的统计分析需求,是科研工作者和从事日常数据处理工作人员的好帮手。随着我国医疗卫生事业的发展与壮大,广大医学科研工作者及研究生对统计分析软件的使用需求越来越大,本书的编撰正是顺应了这种需要,并期望能为其提供帮助。

本书为国家卫生和计划生育委员会“十二五”规划教材,供医学研究生及科研人员使用。本书的前身为教育部学位管理与研究生教育司推荐研究生教学用书《医学统计学》的配套教材,之后在该书基础上,对其中绝大部分实例的软件实现方法给出了详细讲解,并分别于 2006 年和 2010 年出版了第 1 版和第 2 版。本书以 SAS 9.2 版本为基础,增加了 SAS 菜单操作、统计图形绘制等章节,内容更加丰富,知识更加全面。

本书共包含 20 章,前三章主要介绍了 SAS 的基础知识;第四章至第十四章介绍了常用统计方法的软件实现,包括:计量资料的描述性统计、两个样本均数比较的  $t$  检验、多个样本均数比较的方差分析、相关和回归分析、 $\chi^2$  检验、二项分布、Poisson 分布和负二项分布的资料分析、非参数统计分析方法、协方差分析、logistic 回归分析和对数线性模型、生存分析以及多元统计分析方法等内容;第十五章至第十七章介绍了随机抽样、随机化分组和样本量估计等科研设计方面内容的软件实现;第十八章介绍了缺失数据填补方法的软件实现;第十九章介绍了 SAS 的菜单操作;第二十章介绍了运用 SAS 软件绘制统计图形的方法。

本书采用“实例简介—SAS 程序—程序说明—运行结果—结果解释”的模式介绍统计方法的 SAS 实现,重点阐述运用 SAS 软件进行统计分析的过程,并从 SAS 运行结果中选取有价值的信息。本书不仅可以作为高等医学院校研究生软件学习的教材,也可以作为医疗卫生事业科研人员的参考书或工具书。

我在此对编委们所付出的辛勤工作、无私奉献表示衷心的感谢。

由于编者水平所限,书中难免存在不足之处,恳请广大师生和同仁不吝赐教,多提宝贵意见和建议。

贺 佳

2013 年 10 月

# 目 录

第一章 SAS 的概述 .....	1	第七节 嵌套设计资料的方差分析 .....	50
第一节 SAS 的特点和运行环境 .....	1	第八节 裂区设计资料的方差分析 .....	51
第二节 SAS 的启动和退出 .....	3	第九节 重复测量资料的方差分析 .....	53
第三节 SAS 的视窗管理系统 .....	3	第十节 GLM 过程常用选项和语句 .....	58
第四节 SAS 程序 .....	6	第七章 相关和回归分析 .....	60
第五节 SAS 帮助 .....	8	第一节 直线相关分析 .....	60
第二章 数据库与数据集 .....	9	第二节 直线回归分析 .....	61
第一节 数据库 .....	9	第三节 秩相关 .....	65
第二节 数据集 .....	9	第四节 加权直线回归 .....	66
第三节 数据集的建立 .....	10	第五节 指数曲线回归 .....	67
第四节 数据集的整理 .....	14	第六节 对数曲线回归 .....	69
第三章 SAS 运算符和函数 .....	21	第七节 多元线性回归 .....	70
第一节 SAS 运算符 .....	21	第八节 CORR 过程和 REG 过程 常用选项和语句 .....	76
第二节 SAS 常用函数 .....	21	第八章 $\chi^2$ 检验 .....	77
第三节 SAS 运算符和函数举例 .....	23	第一节 四格表资料的 $\chi^2$ 检验 .....	77
第四章 计量资料的单变量描述 .....	25	第二节 $R \times C$ 表资料的 $\chi^2$ 检验 .....	80
第一节 频数表的编制 .....	25	第三节 FREQ 过程常用选项和语句 .....	84
第二节 单变量描述 .....	26	第九章 二项分布、Poisson 分布和负二项 分布 .....	85
第三节 正态性检验 .....	29	第一节 二项分布 .....	85
第四节 几何均数的计算 .....	30	第二节 Poisson 分布 .....	87
第五节 MEANS 及 UNIVARIATE 过程 常用选项和语句 .....	30	第三节 负二项分布中的参数估计 .....	89
第五章 两样本均数的比较 .....	32	第四节 拟合优度检验 .....	89
第一节 单样本均数的 $t$ 检验 .....	32	第五节 Poisson 回归和负二项回归 .....	91
第二节 配对资料两样本均数比较的 $t$ 检验 .....	32	第十章 非参数统计方法 .....	95
第三节 两样本均数比较的 $t$ 检验 .....	34	第一节 配对资料的比较 .....	95
第四节 TTEST 过程常用选项和语句 .....	35	第二节 成组资料的比较 .....	96
第六章 多个样本均数比较的方差分析 .....	36	第三节 两组等级资料的比较 .....	97
第一节 完全随机设计资料的方差分析 .....	36	第四节 完全随机设计资料的比较 .....	98
第二节 随机区组设计资料的方差分析 .....	39	第五节 随机区组设计资料的比较 .....	100
第三节 拉丁方设计资料的方差分析 .....	41	第六节 NPAR1WAY 过程常用选项和语句 .....	101
第四节 两阶段交叉设计资料的方差分析 .....	43	第十一章 协方差分析 .....	102
第五节 析因设计资料的方差分析 .....	44	第一节 完全随机设计资料的协方差分析 .....	102
第六节 正交设计资料的方差分析 .....	48	第二节 随机区组设计资料的协方差分析 .....	104

<b>第十二章</b>	<b>logistic 回归分析和对数线性模型</b> ··· 106	第二节	随机区组分组	170	
第一节	二分类 logistic 回归	106	第三节	分段随机分组	171
第二节	多个自变量的二分类 logistic 回归	109	<b>第十七章</b>	<b>样本含量的估计</b>	173
第三节	1:M 配对资料的条件二分类 logistic 回归	113	第一节	单个样本均数 $t$ 检验的样本含量估计	173
第四节	有序 logistic 回归	119	第二节	两个样本均数 $t$ 检验的样本含量估计	173
第五节	多分类 logistic 回归	120	第三节	多个样本均数比较的样本含量估计	174
第六节	logistic 过程常用选项和语句	122	第四节	单个样本率比较的样本含量估计	175
第七节	对数线性模型	122	第五节	两个独立样本率比较的样本含量估计	175
<b>第十三章</b>	<b>生存分析</b>	124	第六节	直线相关分析的样本含量估计	176
第一节	生存率的计算	124	第七节	两生存曲线比较的样本含量估计	177
第二节	生存曲线比较的 log-rank 检验	128	第八节	POWER 过程常用选项和语句	178
第三节	Cox 回归分析	129	<b>第十八章</b>	<b>缺失数据的多重填补</b>	180
<b>第十四章</b>	<b>多元统计分析</b>	133	第一节	多重填补	180
第一节	判别分析	133	第二节	MI 过程常用选项和语句	183
第二节	聚类分析	140	<b>第十九章</b>	<b>SAS 菜单操作</b>	185
第三节	主成分分析	143	第一节	建立 SAS 数据集	185
第四节	因子分析	145	第二节	数据整理	187
第五节	典型相关分析	148	第三节	数据统计分析	191
第六节	多水平统计模型	153	<b>第二十章</b>	<b>SAS 作图</b>	197
第七节	结构方程模型	159	第一节	直条图、圆图与直方图的绘制	197
<b>第十五章</b>	<b>随机抽样</b>	166	第二节	散点图的绘制	200
第一节	单纯随机抽样	166	第三节	箱式图的绘制	201
第二节	系统抽样	167	第四节	线图的绘制	202
第三节	分层抽样	167			
第四节	SURVEYSELECT 过程常用选项和语句	168			
<b>第十六章</b>	<b>随机化分组</b>	169			
第一节	完全随机分组	169			

# 第一章 SAS 的概述

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之一,由美国北卡罗来纳州州立大学的 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 可以完成数据管理、统计分析、运筹决策等工作。目前(2013 年)SAS 的最高版本为 SAS 9.4,自 SAS 8.0 版开始,SAS 已经有了正式的中文版。本书主要介绍 Windows XP 操作系统下的 SAS 9.2 中文版的统计分析功能。

## 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 可以用多种格式读入数据值,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;whereif-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删除和保留原数据集中的变量;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 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

和偏度系数等多达 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、曲线拟和图,以及时间序列图等。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 不仅可以数据集中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,它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 模块,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以外,还是 SAS 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 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,还可以通过增加不同的模块来增加不同的功能。SAS 有一个智能型绘图系统,不仅能绘制各种统计图,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 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、生存分析、心理测验分析和非参数统计分析等方法共 60 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 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 SAS/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,便于用户发现异常值。

4. SAS/ASSIST 菜单驱动模块。SAS/ASSIST 集成了 SAS 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,在菜单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。有经验的用户还可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 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、曲线图和三维曲面图等高线图以及地理图等,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的中西文矢量图形字体,用户可以在幅面上自由地绘制文字及图形元素,对图形进行修改,添加图形标记,对多幅图形进行任意地拼接组合。

6. SAS/ACCESS 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:IML-/I、SQL/DS、DB2、ADABAS、Rdb、ORACLE、Sybase、INGRES、Informix、DBF/DIF、ODBC 等。

7. SAS/ETS 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 运筹学模块,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 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 IML (Interactive Matrix Language)。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 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、查询和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 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 SCL 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 可以在 Windows 操作系统下使用,运行的环境如下。

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、Windows XP、Windows Vista 非家庭版或 Win7 (包括 32 位和 64 位)。



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

## 第二节 SAS 的启动和退出

### 一、SAS 的启动

SAS 9.2 版的启动方法有以下几种。

1. 快捷方式 在安装了 SAS 9.2 版后, 安装程序会自动在应用程序项中创建 SAS 启动的快捷方式, 用户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“所有程序”, 在显示的应用程序项中会出现“SAS”, 将鼠标移动到该项上, 就会看见“SAS 9.2 (简体中文)”快捷方式(图 1-1), 点击该项, 就可启动 SAS 9.2 中文版。如果在安装时同时安装了其他语言的版本, 会出现其他语言运行 SAS 9.2 的快捷方式。

2. 运行可执行文件 安装 SAS 9.2 中文版时, 会将有关文件安装在硬盘上的某个目录中, 如“C:\SAS”。运行可执行文件即可启动 SAS。具体方法: ①进入“资源管理器”中, 找到安装 SAS 的目录“C:\SAS”, 进入子目录“C:\SAS\SASFoundation\9.2”, 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\SASFoundation\9.2\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。

### 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 系统的方法有以下几种。

1. 菜单操作 点击“文件”, 选择“退出”, 或者同时按下“ALT”和“F4”, 或者点击“×”, 将会出现“退出”对话框, 对话框中写着“确实要结束该 SAS 会话吗?”, 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“退出”对话框。

## 第三节 SAS 的视窗管理系统

启动了 SAS 后, 就进入 SAS 的视窗管理系统 (display management system, DMS), 见图 1-2。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。



图 1-1 快捷方式启动 SAS 9.2 的界面