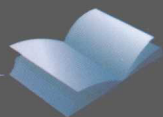


SAS 实验设计 与 统计分析



主 编 胡良平

副主编 胡纯严 周诗国

柳伟伟 高 辉



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

SAS 实验设计 与统计分析

主 编 胡良平

副主编 胡纯严 周诗国 柳伟伟 高 辉

编 委 (以姓氏笔画为序)

王 琪 (军事医学科学院)

毛 玮 (军事医学科学院)

人 民 卫 生 出 版 社

图书在版编目 (CIP) 数据

SAS 实验设计与统计分析/胡良平主编. —北京:
人民卫生出版社, 2010.6

ISBN 978-7-117-12712-7

I. ①S… II. ①胡… III. ①统计分析-应用软件,
SAS IV. ①C812

中国版本图书馆 CIP 数据核字(2010)第 062178 号

门户网: www.pmph.com	出版物查询、网上书店
卫人网: www.ipmph.com	护士、医师、药师、中医师、卫生资格考试培训

版权所有, 侵权必究!

SAS 实验设计与统计分析

主 编: 胡良平

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

购书热线: 010-67605754 010-65264830

010-59787586 010-59787592

印 刷: 保定市中画美凯印刷有限公司

经 销: 新华书店

开 本: 787×1092 1/16 印张: 37 插页: 2

字 数: 888 千字

版 次: 2010 年 6 月第 1 版 2010 年 6 月第 1 版第 1 次印刷

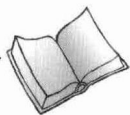
标准书号: ISBN 978-7-117-12712-7/R · 12713

定 价: 72.00 元

打击盗版举报电话: 010-59787491 E-mail: WQ@pmph.com

(凡属印装质量问题请与本社销售中心联系退换)

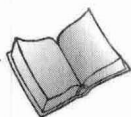
作者简介



胡良平，教授，博士生导师，现任军事医学科学院生物医学统计学咨询中心主任、国际一般系统论研究会中国分会概率统计系统专业理事会常务理事、第八届中国现场统计研究会理事、中国生物医学统计学会副会长、《中华医学杂志》等 10 余种杂志编委、北京大学口腔医学院客座教授、国家食品药品监督管理局评审专家；主编统计学专著 16 部，参编统计学专著 8 部，发表第一作者学术论文 150 余篇，发表合作论文 80 余篇，获军队科技成果和省部级科技成果多项；在从事统计学工作的 20 余年中，为几千名研究生、医学科研人员、临床医生和杂志编辑讲授生物医学统计学，在全国各地

作统计学学术报告 100 余场，举办 10 余期全国统计学培训班，培养多名统计学专业硕士和博士研究生；近几年来，参加国家级新药和医疗器械项目评审数十项；归纳并提炼出“八性”和“八思维”的统计学思想；独创了逆向统计学教学法和统计学三型理论。

内容提要

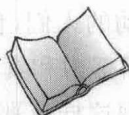


本书内容涉及面十分宽泛,由 SAS 软件基础、SAS 非编程模块用法介绍、SAS 编程法用法介绍、SAS 高级编程技术及其应用和 SAS 语言基础 5 篇组成,涵盖了 SAS 软件及其语言的基础和高级用法,实验设计、统计表达与描述和统计分析的主要内容,以及 SAS 实现。

第 1 篇 SAS 软件基础(包括 SAS 用法入门和 SAS 语言简介),介绍了使用 SAS 软件不可缺少的基本概念、基本方法和操作步骤,读者可以在很短的时间内学会运用 SAS 软件和查找所需信息的技能;第 2 篇 SAS 非编程模块用法介绍(包括 3 个 SAS 非编程模块简介、SAS/ASSIST 用法、SAS/INSIGHT 用法和 SAS/ADX 用法)和第 3 篇 SAS 编程法用法介绍(包括用编程法绘制统计图、常见情况下的样本含量估计与检验效能分析、单因素设计定量资料假设检验、单因素设计定性资料假设检验、简单相关与回归分析、多重回归分析),这两篇的应用内容均包括实验设计、随机化、样本含量估计、常用统计分析和部分多元统计分析;第 4 篇 SAS 高级编程技术及其应用(包括 SAS 宏及其应用、SAS SQL 及其应用、SAS ODS 及其应用、SAS 数组及其应用、SAS/IML 及其应用和 SAS 高级编程的综合应用),这是本书的亮点,涵盖了 SAS 软件中最重要的编程技术;第 5 篇 SAS 语言基础(包括 SAS 语句简介、常用 SAS 函数简介和 SAS 过程简介,以及 SAS 软件平台简介),比较详细但又提纲挈领地介绍了 SAS 语言中的绝大部分内容。

本书既适合于未学过 SAS 软件的新用户,也适合于应用 SAS 软件实际问题多年的老用户;既适合于未学过统计学的新读者,也适合于从事统计学科科研、教学、咨询和培训多年的老读者。总之,本书适合于一切需要学习和运用 SAS 软件解决各种实验设计、统计表达与描述和统计分析问题的本科生、研究生、博士生、科研和管理工作者、临床医生和杂志编辑。

前言



对 SAS 软件和统计分析稍有了解的人都知道, SAS 说明书非常多, 用英文和中文出版的关于如何使用 SAS 软件 and 如何用 SAS 进行统计分析的有关书籍更是不计其数。在这些书中不乏精品之作, 但经过作者深思熟虑且反复提炼出来, 便于学习 SAS 软件并能帮助读者正确而又巧妙地运用 SAS 进行实验设计、统计表达和各种统计分析的书籍并不多。

笔者在大量的统计咨询、项目评审和杂志审稿中看到: 无数实际工作者(科技人员、管理人员、临床医生、杂志编辑、本科生、研究生和博士生)很想快速学会使用 SAS 软件, 并能正确地将其用于自己的科研课题及学位论文的科研设计和统计分析之中。但他们中的很多人几乎都是盲目地从书海中随机挑选一两本, 即便他们很有毅力地从头至尾学完一本 SAS 统计分析书籍, 但一部分人仍不知如何使用 SAS 解决常见的统计学问题, 更谈不上用 SAS 的高级编程技术去完成自己具有创新性的科研课题; 仍不知正确运用统计学的要领是什么、不知统计学内容的合理划分和正确选用及 SAS 的巧妙实现方法。他们耗费了大量宝贵的时间和精力, 学到的却是许多零零星星、支离破碎和一知半解的知识和技能, 就如同一个饥饿的人看到“墙上画着令人垂涎欲滴的烧饼”一样, 只能解解“眼馋”, 而无法实现“充饥”的目的。本书旨在帮助广大读者快速提高科研设计和统计分析的水平, 以及用 SAS 解决这些问题的能力。

本书前两章由胡良平、刘惠刚等编写, 提纲挈领地介绍了 SAS 的基本用法和 SAS 语言要素; 第 3 章由高辉编写, 概略地介绍了 SAS/ASSIST、SAS/ANALYST 和 SAS/INSIGHT 3 个非编程模块的基本情况和用法; 第 4 章由周诗国编写, 详细地介绍了 SAS/ASSIST 模块的使用方法; 第 5 章由郭晋编写, 详细地介绍了 SAS/INSIGHT 模块的使用方法; 第 6 章由高辉编写, 详细地介绍了 SAS/ADX(实验设计)模块的使用方法; 第 7 章由柳伟伟编写, 详细地介绍了用编程法绘制常用统计图的方法; 第 8 章由周诗国编写, 详细地介绍了用 SAS 实现样本含量估计和检验效能分析的方法; 第 9~11 章分别由王琪、毛玮和陶丽新编写, 全面地介绍了单因素设计定量资料和定性资料的统计分析方法、简单相关与回归分析方法及 SAS 实现; 第 12 章由柳伟伟编写, 全面地介绍了多重线性回归分析和多重 logistic 回归分析方法及其应用; 第 13~18 章由胡纯严编写, 系统而又全面地介绍了 SAS 软件的五大高级编程技术(即 SAS 宏、SAS SQL、SAS ODS、SAS 数组、SAS/IML 及其综合应用), 这 6 章是本



书的一大亮点,言简意赅地呈现出 SAS 软件的精华;第 19~21 章分别由程瑞专、郭晋和赵铁牛编写,概括介绍了 SAS 语句、SAS 函数和 SAS 过程,为读者学习 SAS 编程方法提供了必要的技术支持;第 22 章由宋卫平编写,概括介绍了 SAS 软件的发展历史和技术平台。

在本书即将出版之际,笔者真挚地感谢为本书作出过突出贡献的教授、副教授和青年学者;还要感谢所有为本书付出过辛勤劳动的人们,包括 2009 级研究生鲍晓蕾、贾元杰、关雪和孟冰。正是由于他们的积极参与、不懈努力和真心奉献,才使这部专著能够问世!

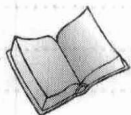
由于笔者水平有限,书中难免会出现这样或那样的不妥,甚至错误之处,恳请广大读者不吝赐教,以便再版时修正。

主编 胡良平

于北京军事医学科学院生物医学统计学咨询中心

2010 年 4 月

目 录



第 1 篇 SAS 软件基础

第 1 章 SAS 用法入门	3
1.1 SAS 软件简介	3
1.2 SAS 界面简介	3
1.3 SAS 产品简介	4
1.4 SAS 过程简介	7
1.5 SAS 程序简介	7
1.6 SAS 用法简介	8
1.7 SAS Pal 软件简介	17
第 2 章 SAS 语言简介	19
2.1 SAS 语句简介	19
2.2 SAS 函数简介	21
2.3 SAS 过程简介	30

第 2 篇 SAS 非编程模块用法介绍

第 3 章 3 个 SAS 非编程模块简介	45
3.1 SAS/ASSIST 模块简介	45
3.2 SAS/ANALYST 模块简介	54
3.3 SAS/INSIGHT 模块简介	65
第 4 章 SAS/ASSIST 模块用法	71
4.1 绘制常用统计图	71
4.2 定量资料统计描述	86
4.3 定量资料统计分析	92
4.4 定性资料统计分析	97
4.5 简单相关与回归分析	101
4.6 多重线性回归分析	104



4.7	主成分分析	108
4.8	典型相关分析	114
第5章	SAS/INSIGHT 模块用法	118
5.1	绘制常用统计图	118
5.2	分布检验	125
5.3	曲线拟合	127
5.4	多元统计分析	131
第6章	SAS/ADX 模块用法	138
6.1	ADX 菜单驱动系统界面简介	138
6.2	ADX 菜单驱动系统操作方法简介	138
6.3	ADX 菜单驱动系统应用示例	141
第3篇 SAS 编程法用法介绍		
第7章	用编程法绘制统计图	171
7.1	绘制条图	171
7.2	绘制圆图	177
7.3	绘制箱式图	178
7.4	绘制直方图	180
7.5	绘制散布图	182
7.6	绘制普通线图	183
7.7	绘制半对数线图	185
7.8	绘制 P-P 图和 Q-Q 图	187
第8章	常见情况下的样本含量估计与检验效能分析	190
8.1	确定样本含量的意义	190
8.2	确定样本含量时应具备的条件	190
8.3	估计总体均数时样本含量的估计	191
8.4	估计总体概率时样本含量的估计	192
8.5	单组、配对或交叉设计定量资料统计分析时样本含量估计	193
8.6	成组设计定量资料统计分析时样本含量估计	195
8.7	单因素多水平设计定量资料方差分析时样本含量估计	200
8.8	两因素析因设计定量资料方差分析时样本含量估计	202
8.9	重复测量设计定量资料方差分析时样本含量估计	205
8.10	两样本频率比较时样本含量的估计	206
8.11	两样本生存频率比较时样本含量的估计	207
8.12	多个样本频率比较时样本含量的估计	208
8.13	配对设计四格表资料统计分析时样本含量估计	209



8.14	队列研究设计四格表资料统计分析时样本含量估计	211
8.15	简单直线相关或回归分析时样本含量的估计	212
8.16	单因素二水平设计定量资料的等效性检验时样本含量的估计	214
8.17	单因素二水平设计定量资料的非劣效性检验时样本含量的估计	216
8.18	单因素二水平设计定性资料的等效性检验时样本含量的估计	217
8.19	单因素二水平设计定性资料的非劣效性检验时样本含量的估计	218
8.20	检验效能的计算概述	219
8.21	单组、配对或交叉设计定量资料假设检验时检验效能的计算	219
8.22	单因素二水平设计样本均数比较时检验效能的计算	220
8.23	单因素多水平设计定量资料的方差分析时检验效能的计算	222
8.24	两因素析因设计定量资料方差分析时检验效能的计算	223
8.25	例数相等的两组样本频率比较时检验效能的计算	225
8.26	例数不相等的两组样本频率比较时检验效能的计算	228
8.27	单因素二水平设计定量资料的等效性检验时检验功效的分析	229
8.28	单因素二水平设计定量资料的非劣效性检验时检验功效的分析	230
8.29	单因素二水平设计定性资料的等效性检验时检验功效的分析	231
8.30	单因素二水平设计定性资料的非劣效性检验时检验功效的分析	232
8.31	用统计软件的界面操作方式实现样本含量估计和检验功效分析	233
第 9 章	单因素设计定量资料假设检验	238
9.1	单组设计定量资料假设检验	238
9.2	配对设计定量资料假设检验	242
9.3	成组设计定量资料假设检验	249
9.4	单因素多水平设计定量资料假设检验	255
第 10 章	单因素设计定性资料假设检验	266
10.1	单组设计定性资料假设检验	266
10.2	配对设计定性资料假设检验	267
10.3	成组设计定性资料假设检验	272
10.4	单因素多水平设计定性资料假设检验	280
第 11 章	简单相关与回归分析	287
11.1	重要知识点概述	287
11.2	简单相关分析	288
11.3	简单回归分析	291
11.4	可直线化曲线回归分析	297
第 12 章	多重回归分析	306
12.1	多重线性回归分析	306



12.2	结果变量为二值变量的多重 logistic 回归分析	311
12.3	结果变量为多值有序变量的多重 logistic 回归分析	315
12.4	结果变量为多值名义变量的多重 logistic 回归分析	319

第 4 篇 SAS 高级编程技术及其应用

第 13 章	SAS 宏及其应用	327
13.1	概述	327
13.2	宏变量	330
13.3	宏/宏函数	333
13.4	宏语句	338
13.5	宏的本质和几点使用技巧	340
13.6	宏实例详解	346
第 14 章	SAS SQL 及其应用	352
14.1	概述	352
14.2	SQL 的本质	353
14.3	重要 SQL 语句的使用及其与相应功能 data 步的对比	354
14.4	实例分析	360
第 15 章	SAS ODS 及其应用	368
15.1	概述	368
15.2	SAS ODS 的分类与重点	369
15.3	重要 ODS 语句的使用	370
15.4	实例分析	383
第 16 章	SAS 数组及其应用	388
16.1	SAS 数组的概况	388
16.2	SAS 数组的语法结构	389
16.3	实例分析	394
第 17 章	SAS/IML 及其应用	402
17.1	SAS/IML 概况	402
17.2	IML 过程的语法结构	403
17.3	实例分析	416
第 18 章	SAS 高级编程的综合应用	422
18.1	SAS 与 MVC	422
18.2	实例分析	422



第 5 篇 SAS 语言基础

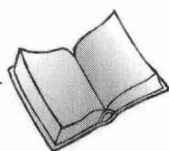
第 19 章 SAS 语句简介	443
19.1 SAS 程序(SAS Program)	443
19.2 SAS 程序的构成和书写格式	445
19.3 SAS 语句(SAS Statements)概念	448
19.4 SAS 数据步语句简介	461
19.5 数据步文件处理语句	463
19.6 数据步动作语句	486
19.7 数据步控制语句	500
19.8 数据步信息语句	510
19.9 窗体显示语句	518
19.10 全局语句简介	519
19.11 全局数据存取语句	521
19.12 全局日志控制语句	525
19.13 全局环境控制语句	525
19.14 全局输出控制语句	526
19.15 全局程序控制语句	527
第 20 章 常用 SAS 函数简介	532
20.1 日期时间函数	532
20.2 算术函数	534
20.3 数组函数	535
20.4 截取函数	536
20.5 数学函数	536
20.6 三角函数和双曲函数	538
20.7 概率函数	538
20.8 分位数函数	541
20.9 样本统计函数	541
20.10 随机数函数	542
20.11 SAS call 子程序	544
第 21 章 SAS 过程简介	547
21.1 SAS 过程步的一般形式	547
21.2 SAS 过程步中常用的语句	547
21.3 SAS/BASE 模块中几个实用过程应用举例	549
21.4 常用 SAS 过程分类及功能简介	556
第 22 章 SAS 软件平台简介	565



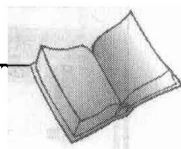
目 录

22.1	SAS 软件的历史	565
22.2	SAS 软件平台简介	568
22.3	SAS 软件平台的组成部件	569
22.4	SAS 软件平台的架构	572
22.5	SAS 软件平台的综合优势	572
附录 胡良平统计学专著及配套软件简介		574

第 1 篇



SAS 软件基础



1.1 SAS 软件简介

SAS 软件内容非常丰富,它采取模块式结构,每个模块可被称为一个 SAS 产品。有些 SAS 产品中仅有具体的 SAS 过程(即编译后的 SAS 程序),如 SAS/STAT;有些 SAS 产品中不仅有 SAS 过程,还有其他内容,如 SAS 语言、SAS 窗口、SAS 宏、SAS SQL 等,见 SAS 帮助窗口中有关 SAS/BASE 模块的内容介绍;有些 SAS 产品本身就是一个功能比较齐全的软件,可以完成一系列相关的功能,如 SAS/ASSIST 模块、SAS/ANALYST 模块和 SAS/INSIGHT 模块等;还有些 SAS 产品是其他 SAS 产品的开发工具,如 SAS/AF 等。

SAS 不仅可以用来解决统计学方面的问题(如实验设计、绘制统计图和编制统计表、完成简单和复杂的统计分析),还可以用来解决许多其他领域中的问题,如质量控制(SAS/QC)、矩阵运算(SAS/IML)和数据挖掘(SAS Enterprise Miner)等,参见本章“1.3 SAS 产品简介”一节。

1.2 SAS 界面简介

不同版本的 SAS 软件,其界面不尽相同。比较新的 SAS 软件界面见图 1-1。

由图 1-1 可知,在 SAS 主界面上,顶部有两行,分别为“菜单栏”和“工具栏”;左边有一个“SAS 资源管理器”窗口,与它重叠的一个窗口叫“结果”窗口;右边有两个窗口,上方的窗口叫“日志”窗口,下方的窗口叫“程序编辑器”窗口,通过顶部菜单栏中的“窗口”选项,可以切换到其他窗口,如“输出”窗口。

可以说,基本的 SAS 窗口有“SAS 资源管理器”、“结果”、“程序编辑”、“日志”和“输出”窗口,但另外还有 30 多个窗口可供处理打印和微调 SAS 会话之类的操作。这些窗口的名称和窗口命令的详细列表从略,若用户想获得此列表,可通过下面的方法实现:先选中主界面第二行工具栏最后一个选项(图标为一本书),即帮助(help);然后按照“帮助(help)→SAS 产品→Base SAS→SAS 窗口引用→SAS 窗口索引”的步骤即可显示全部 SAS 窗口列表。

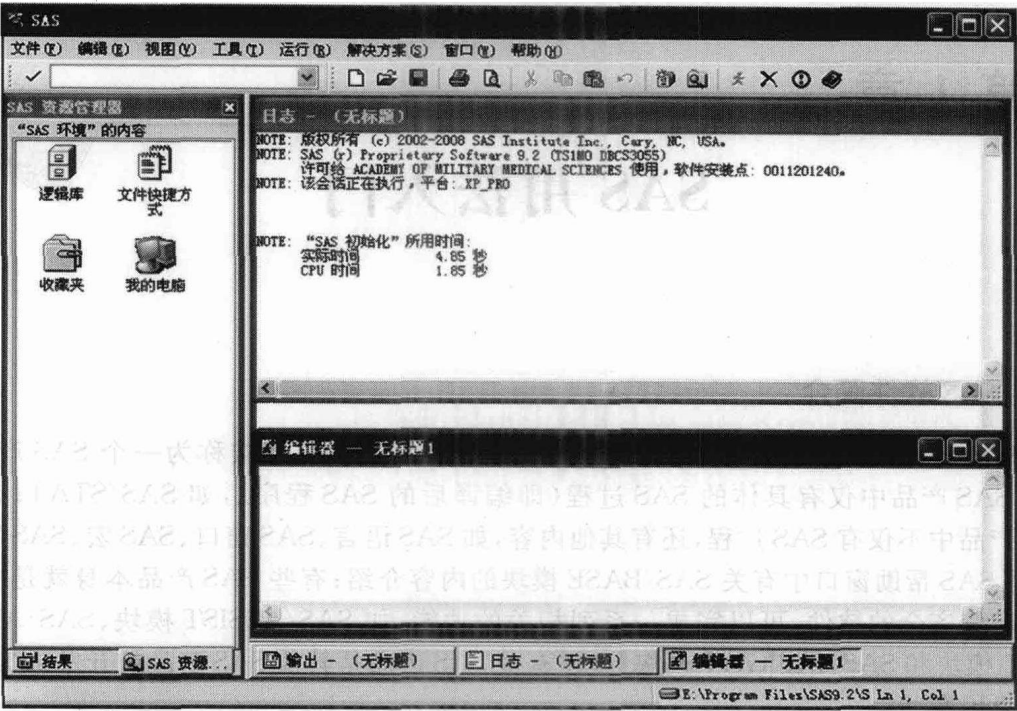


图 1-1 进入 SAS 9.2 后的主界面

1.3 SAS 产品简介

SAS 9.1.3 软件中包含了哪些产品呢? 见表 1-1。

表 1-1 SAS 软件产品目录

产品 编号	产品名称	产品 编号	产品名称
1	Base SAS	18	SAS Integration Technologies
2	Enterprise Miner	19	SAS/IntrNet
3	SAS/ACCESS	20	SAS/Inventory Replenishment planning
4	SAS/AF	21	SAS IT Resource Management
5	SAS/ASSIST	22	SAS/LAB
6	SAS/CONNECT	23	SAS National Language Support(NLS); User's Guide
7	SAS Data Quality Server	24	SAS OLAP Server
8	SAS e-DATA ETL	25	SAS ODBC(Open Data Base Connectivity)
9	SAS/EIS	26	SAS/OR
10	SAS/ETS	27	SAS/QC
11	SAS/FSP	28	SAS Risk Dimensions
12	SAS/Genetics	29	SAS/SHARE
13	SAS/GIS	30	SAS/SPECTRAVIEW
14	SAS/GRAPH	31	SAS/STAT
15	SAS High-Performance Forecasting	32	SAS/TOOLKIT
16	SAS/IML	33	SAS/Warehouse Administrator
17	SAS/INSIGHT		