

<<SAS统计分析与数据挖掘>>

图书基本信息

书名：<<SAS统计分析与数据挖掘>>

13位ISBN编号：9787121148880

10位ISBN编号：7121148889

出版时间：2012-1

出版时间：电子工业出版社

作者：谢龙汉、尚涛

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

<<SAS统计分析与数据挖掘>>

前言

SAS是当今国际最著名的数据分析软件。

本书从SAS编程出发，用案例的形式介绍SAS数据分析在各个领域的广泛应用。

本书是在结合作者多年的使用经验和工作经验的基础上编写的，在编写过程中，突出了以下特点：

(1) 直观易懂性。

全书以图解实例的形式介绍基础知识和实例操作，所有的知识块和案例分析都尽可能详细，直观易懂，使读者能够在最短的时间内获取最多的知识。

(2) 先进性。

以最新的SAS 9.2版为蓝本进行讲解，广泛吸收国内外优秀教材的成果进行内容设计，在系统介绍基本理论和基本方法的同时，注意介绍新的成熟内容，以及统计学在实际问题中的应用。

(3) 实用性。

全书采用了基础知识介绍和实例操作相结合的方法，互相补充，本书的实例大多来源于经济生活之中，使读者在学完本书后能够快速地将知识应用于实践。

(4) 结构清晰，讲解详尽。

全书采用基础知识 - 程序实现 - 综合实例分析的讲解方法，循序渐进地提高读者的SAS编程知识，而且每个知识点和实例都做了尽可能详细的讲解，使读者学习起来轻松自如。

(5) 配有全部的案例数据、程序与多媒体示范。

本书的配套光盘中提供了所有实例的数据、SAS程序、视频操作，读者可以在观看录像中增强对知识点的理解。

本书共分为21章，依次介绍了SAS 9.2的基本编程知识、基本统计分析、高级统计分析、SAS数据挖掘、SAS在金融中的应用，以及各章节中的案例分析等内容。

第1章 数据挖掘概述。

介绍数据挖掘的涵义，为什么要进行数据挖掘，数据挖掘的用途、过程等，最后介绍SAS软件系统在数据挖掘中的地位。

第2章 SAS模块概述。

详细介绍SAS软件系统和SAS各个常用的模块。

第3章 SAS程序设计基础。

详细介绍SAS编程中的各个步骤，包括SAS安装过程，SAS数据步和过程步，程序调试及SAS函数，掌握SAS编程基础。

第4章 数据预处理。

内容包括数据输入、数据整理、数据步变量控制、数据修改与选择，最后介绍SAS与SPSS软件之间的数据转换。

第5章 数据汇总与报表制作。

介绍最基本的数据报表生成过程，PROC PRINT和PROC TABULATE过程。

第6章 SAS绘图。

统计图是统计描述的重要工具，本章介绍SAS绘制图形的GPLOT、GCHART、G3D过程。

第7章 数据描述。

介绍利用统计图、统计量和数据分布进行数据描述的过程。

第8章 描述性统计分析。

叙述描述性统计分析中的平均数、中位数、众数等度量集中趋势，用极差、标准差、变异系数等度量。

第9章 ANALYST模块。

主要介绍利用ANALYST模块进行数据管理。

第10章 参数估计和假设检验。

介绍对于定量资料的统计描述和简单推断，主要有UNIV-ARIATE、MEANS、TTEST过程。

第11章 方差分析与协方差分析。

<<SAS统计分析与数据挖掘>>

介绍方差分析与协方差分析的基本原理及其SAS过程。

第12章 回归分析。

介绍线性回归、REG过程、多项式回归、逐步回归LOGISTIC过程及非线性回归。

第13章 主成分分析与因子分析。

介绍主成分分析与因子分析的原理、数学模型及SAS过程，然后应用于案例分析。

第14章 相关分析和对应分析。

介绍相关分析和对应分析的原理、数学模型及案例分析。

第15章 判别分析。

介绍判别分析的最大似然法、Fisher判别分析法、Bayes判别分析法、逐步判别分析法等，并应用于案例分析。

第16章 聚类分析。

介绍聚类分析的基本原理，以及CLUSTER、FASTCLUS、VARCLUS、TREE过程。

第17章 生存分析。

介绍SAS对生存分析的应用，并进行研究探讨。

第18章 时间序列分析。

介绍时间序列的数学模型，SAS的ARIMA过程及案例分析。

第19章 SAS数据挖掘应用。

主要介绍SAS数据挖掘方法论--SEMMA，SAS的企业数据挖掘套件SAS/EM及其案例分析。

第20章 SAS在数据预测中的应用。

介绍利用SAS软件系统对有关数据预测的一些案例进行分析研究，为读者展示数据预测的相关技术和方法。

第21章 SAS在金融数据分析中的应用。

介绍利用SAS软件对部分金融学中数学模型进行分析求解，展示SAS软件在金融数据挖掘中的应用。

本书主要由尚涛完成，香港中文大学谢龙汉博士提供技术支持和指导，并对本书进行了校对和完善。

参加本书编写和光盘开发的还有林伟、魏艳光、林木议、王悦阳、林伟洁、林树财、郑晓、吴苗、李翔、莫衍、朱小远、唐培培、耿煜、邓奕、张桂东、鲁力、刘文超、刘新东等，同时也非常感谢腾龙工作室其他成员的帮助和支持。

编著者

<<SAS统计分析与数据挖掘>>

内容概要

本书基于SAS

9.2版本编写，从SAS编程出发，用案例形式介绍SAS数据挖掘在各领域的广泛应用，全书分为SAS基础篇、提高篇及应用篇，每章均给出大量分析案例。

具体内容为SAS软件与数据挖掘简介，SAS编程基础，图形与报表制作，描述性分析，假设检验，回归分析，方差分析与因子分析，相关分析与对应分析，判别分析，聚类分析，生存分析，时间序列分析，以及SAS在具体数据挖掘项目中的应用等。

本书最大特点是抛弃了其他同类书籍中只说理论、缺少案例分析的弊病，全书给出大量数据挖掘分析案例，为读者展示SAS在数据整合、数据挖掘、商业智能、金融数据分析、金融风险管理等项目中的强大应用技术。

配套光盘中有实例的操作视频以及相关源程序文件。

<<SAS统计分析与数据挖掘>>

书籍目录

第1章 数据挖掘概述

1.1 数据挖掘简介

1.1.1 数据挖掘的含义

1.1.2 数据挖掘的起源

1.1.3 统计学与数据挖掘

1.1.4 数据挖掘相关的一些问题

1.2 数据挖掘用途

1.3 数据挖掘过程

1.3.1 数据挖掘用户

1.3.2 数据挖掘工具

1.3.3 数据挖掘步骤

1.4 SAS——数据挖掘领域的领导者

1.5 SAS在各种商业解决方案中的应用

1.5.1 SAS数据挖掘技术的实现

1.5.2 SAS在商业领域中的应用

第2章 SAS模块概述

2.1 SAS简介

2.1.1 SAS的设计思想

2.1.2 SAS的功能

2.1.3 SAS的特点

2.2 SAS软件安装、启动与退出

2.2.1 SAS软件的安装

2.2.2 SAS软件的启动

2.2.3 SAS软件的退出

2.3 SAS界面

2.3.1 Explorer窗口

2.3.2 Editor窗口

2.3.3 Results窗口

2.3.4 Log窗口

2.3.5 Output窗口

2.4 SAS模块介绍

2.4.1 SAS/BASE模块

2.4.2 SAS/ANALYSIS模块

2.4.3 SAS/ASSIST模块

2.4.4 SAS/INSIGHT模块

2.4.5 SAS/EM模块

第3章 SAS程序设计基础

3.1 SAS编程基础

3.1.1 SAS语言基础

3.1.2 SAS语言构成

3.1.3 SAS结构化编程语句

3.1.4 SAS程序编写规则

3.2 SAS程序的数据步

3.2.1 DATA语句

3.2.2 INPUT语句

<<SAS统计分析与数据挖掘>>

- 3.2.3 CARDS与CARDS4语句
- 3.2.4 INFILE语句
- 3.2.5 SET语句
- 3.2.6 MERGE语句
- 3.3 SAS数据步循环与转移控制
 - 3.3.1 IF语句
 - 3.3.2 SELECT语句
 - 3.3.3 DO语句
 - 3.3.4 GO TO语句
 - 3.3.5 RETURN语句
 - 3.3.6 CONTINUE语句与LEAVE语句
 - 3.3.7 如何跳出选择结构和循环体
- 3.4 SAS程序的过程步
 - 3.4.1 SAS过程步用法
 - 3.4.2 VAR与MODLE语句
 - 3.4.3 ID与WHERE语句
 - 3.4.4 BY与CLASS语句
 - 3.4.5 OUTPUT语句
 - 3.4.6 FERQ与WEIGHT语句
 - 3.4.7 LABEL与FORMAT语句
- 3.5 SAS函数
 - 3.5.1 数学函数
 - 3.5.2 数组函数
 - 3.5.3 日期时间函数
 - 3.5.4 概率分布函数
 - 3.5.5 分位数函数
 - 3.5.6 样本统计函数
 - 3.5.7 随机函数
- 第4章 数据预处理
 - 4.1 数据输入
 - 4.1.1 原始数据的读取
 - 4.1.2 数据导入
 - 4.2 数据整理
 - 4.2.1 数据集选项
 - 4.2.2 整理数据集
 - 4.2.3 缺失值处理
 - 4.2.4 UPDATE语句更新数据集
 - 4.2.5 数据清洗
 - 4.3 数据步变量控制
 - 4.3.1 ARRAY语句
 - 4.3.2 INFORMAT语句与FORMAT语句
 - 4.3.3 LABEL语句
 - 4.3.4 ATTRIB语句
 - 4.3.5 DROP语句与KEEP语句
 - 4.3.6 RENAME语句与RETAIN语句

<<SAS统计分析与数据挖掘>>

4.4 数据修改与选择

4.4.1 赋值语句

4.4.2 累加语句

4.4.3 DELETE语句与LOSTCARD
语句

4.4.4 STOP语句与ABORT语句

4.4.5 WHERE语句

4.4.6 REMOVE语句与REPLACE
语句

4.4.7 MISSING语句

第5章 数据汇总与报表制作

5.1 使用过程PRINT制作报表

5.1.1 基本用法

实例5-1 PROC PRINT操作
实例

5.1.2 使用中文列标题

实例5-2 修改标题实例

5.1.3 标题和脚注

实例5-3 修改标题实例

5.1.4 用BY语句分组处理

5.2 使用过程TABULATE制作汇
总报表

实例5-4 汇总报表实例

实例5-5 绘制统计量表格

第6章 SAS绘图

6.1 GPLOT过程

实例6-1 GPLOT过程绘制图形
编程操作

6.2 GCHART过程

实例6-2 GCHART过程绘制
条形图实例6-3 GCHART过程绘制GDP
数据的BLOCK图形

6.3 G3D过程

实例6-4 绘制二维正态分布曲面
图形实例6-5 绘制
函数的三维图形

第7章 数据描述

7.1 统计图

7.1.1 直方图

实例7-1 GCHART过程绘制
直方图

7.1.2 条形图

实例7-2 GCHART过程绘制
条形图

7.1.3 散点图

<<SAS统计分析与数据挖掘>>

实例7-3 GPLOT过程绘制散点图

7.1.4 饼图

实例7-4 GCHART过程绘制饼图

7.1.5 盒形图

实例7-5 BOXPLOT过程绘制
盒形图

7.1.6 茎叶图

实例7-6 UNIVARIATE过程绘制
茎叶图

7.1.7 时间序列图

实例7-7 TIMEPLOT过程绘制
时间序列图

7.2 统计量

7.2.1 集中趋势

实例7-8 利用MEAN函数求
平均数

7.2.2 离散程度

实例7-9 利用函数VAR和STD
求方差和标准差

7.2.3 分布状态

实例7-10 利用SKEWNESS 和
KURTOSIS函数求偏度
和峰度

7.3 数据分布

实例7-11 SAS中的部分概率分布
函数的应用

第8章 描述性统计分析

8.1 SAS编程进行统计分析

8.1.1 基本概念

8.1.2 FREQ过程

实例8-1 频数表的生成实例

实例8-2 绘制实验数据表格

8.1.3 MEANS过程

实例8-3 求平均增长率

实例8-4 利用MEANS过程求各种
统计量

8.1.4 UNIVARIATE过程

实例8-5 利用UNIVARIATE过程
求各种统计量

实例8-6 求样本的极差、上四分位
数和下四分位数

8.1.5 TABULATE过程

实例8-7 制作数据表格

8.2 其他描述性统计过程

8.2.1 产生描述性统计值的输出
文件：PROC SUMMARY

实例8-8 SUMMARY语句实例

<<SAS统计分析与数据挖掘>>

8.2.2 统计值的图形表示：PROC
CHART

实例8-9 绘制数据分布图形

实例8-10 利用CHART过程的
VBAR及HBAR命令
绘制条形图

8.2.3 一般制图：PROC PLOT

实例8-11 PLOT过程绘制图形

第9章 ANALYST模块

9.1 ANALYST模块概述

9.1.1 ANALYST模块简介

9.1.2 ANALYST菜单介绍

9.2 数据集的窗口操作

9.2.1 数据集输入

9.2.2 数据表修改

9.2.3 数据保存

9.3 绘制统计图

9.3.1 条形图

9.3.2 饼图

9.3.3 散点图

9.4 统计分析

第10章 参数估计与假设检验

10.1 参数估计和假设检验概述

10.1.1 参数估计

10.1.2 假设检验

10.2 假设检验的SAS过程

10.2.1 UNIVARIATE过程

10.2.2 MEANS过程

10.2.3 TTEST过程

10.3 不同类型的均值和方差的检验

10.3.1 单变量均值t检验

实例10-1 TTEST过程的实例数据
分析

实例10-2 总体均值检验

10.3.2 样本均数与总体均数差异的
t检验实例10-3 均值的显著性差别
检验

10.3.3 配对资料的t检验

实例10-4 乳酸饮料实验数据的
配对t检验

实例10-5 均值有无差异的检验

10.3.4 两样本均数比较的t检验

实例10-6 均数差别的显著性
检验实例10-7 数据比例的显著性
检验

<<SAS统计分析与数据挖掘>>

10.4 正态性检验

实例10-8 样本数据的正态性检验

实例1

实例10-9 样本数据的正态性检验

实例2

第11章 方差分析与协方差分析

11.1 方差分析的基本原理

11.1.1 自由度与平方和分解

11.1.2 F检验

11.2 单因素方差分析

11.2.1 单因素方差分析步骤

11.2.2 判断与结论

11.2.3 ANOVA过程

实例11-1 分析饲料营养效果是否有明显差异

实例11-2 分析不同实验室试制的纸张光滑度有无差异

实例11-3 研究6种棉花种子包衣剂对棉花生长的影响

11.3 双因素方差分析

11.3.1 只考虑主效应的多因素方差分析

11.3.2 存在交互效应的多因素方差分析

实例11-4 某药物对某癌细胞株增殖影响的研究

11.4 协方差分析

实例11-5 分析三种饲料的营养价值之间有无显著性差别

第12章 回归分析

12.1 线性回归

12.1.1 线性回归模型

12.1.2 回归方程的显著性检验

12.1.3 预测问题

12.2 REG过程

实例12-1 分析我国内地可支配收入和消费性支出之间的关系

实例12-2 利用多元线性回归分析学生肺活量及有关变量的关系

12.3 多项式回归

12.3.1 曲线回归的基本原理

12.3.2 RSREG过程

实例12-3 确定最佳经济用肥量的多项式回归模型

12.4 逐步回归

<<SAS统计分析与数据挖掘>>

实例12-4 人体血糖、胰岛素及生长素的多元线性回归关系

12.5 LOGISTIC回归

12.5.1 逻辑回归模型概述

12.5.2 LOGISTIC过程

实例12-5 对照研究单因素两暴露水平及多暴露水平资料的统计分析

12.6 非线性回归

12.6.1 非线性回归分析的基本原理

12.6.2 NLIN过程

实例12-6 酵母种群增长的拟合生长模型

实例12-7 最佳生长模型的LOGISTIC拟合

第13章 主成分分析与因子分析

13.1 主成分分析

13.1.1 主成分分析的数学原理

13.1.2 用PRINCOMP过程进行主成分分析

实例13-1 我国2006年经济发展情况的主成分分析

13.2 因子分析

13.2.1 因子分析的基本原理

13.2.2 因子分析的基本步骤和过程

13.2.3 利用FACTOR过程进行因子分析

实例13-2 中国房地产经济区的研究分析

13.3 主成分分析和因子分析的区别

第14章 相关分析和对应分析

14.1 相关分析

14.1.1 相关关系

14.1.2 相关图形和相关系数

14.1.3 简单相关分析的CORR过程

实例14-1 简单相关系数的计算

14.2 典型相关分析

14.2.1 典型相关分析的基本原理

14.2.2 典型相关分析的CANCORR过程

实例14-2 城市竞争力与基础设施的典型相关分析

实例14-3 城镇居民收入和支出的

<<SAS统计分析与数据挖掘>>

典型相关分析

14.3 对应分析

14.3.1 对应分析的基本原理

14.3.2 对应分析的CORRESP

过程

实例14-4 对应分析在市场细分中的应用

第15章 判别分析

15.1 判别分析的基本原理

15.1.1 判别分析的含义

15.1.2 判别分析的数学模型与判别方法

15.2 判别分析的SAS过程

15.2.1 DISCRIM过程

15.2.2 CANDISC过程

15.2.3 STEPDISC过程

15.3 综合实例

实例15-1 国内各省市农民家庭收支情况的研究

实例15-2 基于判别分析法的上市公司财务分析研究

第16章 聚类分析

16.1 聚类分析的基本原理

16.1.1 聚类的数学原理

16.1.2 SAS中的聚类过程

16.2 聚类分析的步骤和过程

16.2.1 CLUSTER过程(系统聚类过程)

实例16-1 中国城镇居民消费结构的聚类分析

16.2.2 FASTCLUS过程(快速聚类过程)

实例16-2 聚类分析在客户定位中的应用研究

16.2.3 VARCLUS过程(变量聚类过程)

实例16-3 变量聚类在多指标系统评价中的应用

16.2.4 TREE过程(画树状图过程)

实例16-4 对全球各国信息设施的发展情况进行聚类分析研究

第17章 生存分析

17.1 生存分析基本概述

17.1.1 生存分析的基本概念

17.1.2 生存资料的特点

<<SAS统计分析与数据挖掘>>

- 17.1.3 生存分析方法
- 17.2 生存分析的LIFETEST过程
 - 实例17-1 生存分析在医学课题研究中的应用
- 17.3 COX模型回归分析
 - 17.3.1 COX回归模型
 - 17.3.2 PHREG过程
 - 实例17-2 COX模型的分析应用
- 第18章 时间序列分析
 - 18.1 时间序列概述
 - 18.1.1 时间序列的组成部分
 - 18.1.2 时间序列的数学模型
 - 18.1.3 时间序列的因素分析
 - 18.1.4 随机时间序列分析
 - 18.1.5 时间序列的分析步骤
 - 18.2 SAS的ARIMA过程
 - 18.3 综合实例
 - 实例18-1 化工生产数据的时间序列分析
 - 实例18-2 国内金融及保险业每人每月平均薪资趋势分析
 - 实例18-3 运用ARIMA过程对上证指数日线数据进行拟合分析
- 第19章 SAS数据挖掘应用
 - 19.1 SAS数据挖掘
 - 19.2 SAS数据挖掘方法论——SEMMA
 - 19.2.1 数据取样
 - 19.2.2 数据探索
 - 19.2.3 问题明确化、数据调整和技术选择
 - 19.2.4 模型研发
 - 19.2.5 模型评估
 - 19.3 数据挖掘套件SAS/EM
 - 实例19-1 SAS/EM聚类分析
 - 实例19-2 购物篮问题分析
- 第20章 SAS在数据预测中的应用
 - 20.1 数据预测简介
 - 20.1.1 数据预测
 - 20.1.2 SAS中的预测分析模块
 - 20.2 数据预测案例分析
 - 实例20-1 国民生产总值的预测
 - 实例20-2 SAS/Time Series Forecasting System 模块应用

<<SAS统计分析与数据挖掘>>

第21章 SAS在金融数据分析中的应用

21.1 现金流贴现分析

实例21-1 现金流贴现的计算

实例21-2 企业现金流的贴现
计算

实例21-3 利用金融函数compound
计算复利率

21.2 股票分类

实例21-4 利用CLUSTER过程对
股票进行聚类分析

21.3 资本资产定价模型 (CAPM 模型)

实例21-5 CAPM模型实例研究

21.4 B-S模型期权定价

实例21-6 B-S期权定价的SAS
程序实现

章节摘录

版权页：插图：

<<SAS统计分析与数据挖掘>>

编辑推荐

《SAS统计分析与数据挖掘》SAS——全球最为著名的统计分析软件，SAS——金融、科研、数据挖掘领域必备之软件，基础知识-实训实例-工程实例，实例操作视频教学，轻松学习。

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>