

The background features a complex geometric design. On the left, there are large, semi-transparent pink arrows that curve and point in various directions. On the right, there are several dark gray, three-dimensional geometric shapes, including cubes and rectangular prisms, arranged in a way that suggests depth and perspective. The overall color palette is a mix of soft pinks, grays, and white.

SAS

编程技术与 金融数据处理

朱世武 著

清华大学出版社

SAS

编程技术与金融数据处理

朱世武 著

清华大学出版社
北 京

内 容 简 介

本书内容由中外金融数据库、SAS 编程技术和金融工具收益计算三大模块组成。

中外金融数据库模块介绍了中外著名金融数据库的数据结构及本书所用的数据库。

SAS 编程技术模块分为五个部分,系统介绍了基于 SAS 系统的数据加工整理、数据访问、数据管理和数据展现技术,以及 SAS 系统的开发工具 SAS/IML 软件。

金融工具收益计算模块给出了基于 SAS 系统的简单金融建模和金融计算。

本书的结构及具体内容都经过仔细设计,具有系统性强、层次明确、语言简练的特点。通过大量的例程,让读者轻松掌握数据处理和金融计算的技术。所附光盘中的教学数据库及全套程序集,为本书的学习提供了可靠的实验条件。

本书适合多层次多专业的人士阅读,如数学、统计学、经济和金融等专业的本科生、研究生及相应部门的专业人员。

书 名: SAS 编程技术与金融数据处理

作 者: 朱世武 著

出 版 者: 清华大学出版社

<http://www.tup.com.cn>

社 总 机: 010-62770175

地 址: 北京清华大学学研大厦

邮 编: 100084

客户服务: 010-62776969

责任编辑: 江 娅

版式设计: 肖 米

印 刷 者: 国防工业出版社印刷厂

发 行 者: 新华书店总店北京发行所\清华大学出版社出版发行

开 本: 787×1092 1/16 印张: 32.25 字数: 731 千字

版 次: 2003 年 7 月第 1 版 2003 年 7 月第 1 次印刷

书 号: ISBN 7-89494-124-7

印 数: 1~5000

定 价: 52.00 元

序

言

现代社会中定量分析方法越来越显示出它的重要地位,而进行定量分析需要掌握很好的数据处理技术。

本书涵盖内容包括中外金融数据库、SAS 编程技术和金融工具收益计算三大模块。中外金融数据库模块挑选了国际上最著名的静态金融数据库证券价格研究中心(CRSP)证券市场数据库的数据结构,动态金融数据服务系统路透(REUTER)的使用方法,以及作者对中国证券市场的数据结构设计。这既展拓读者的眼界,又为本书的学习提供了数据源。SAS 编程技术模块系统介绍了基于 SAS 系统的数据加工整理、数据访问、数据管理和数据展现技术,以及 SAS 系统的开发工具 SAS/IML 软件。金融工具收益计算模块给出了基于 SAS 系统的简单金融建模和金融计算。

SAS 软件早已全球著名,目前,在具有工业用途的三个统计分析软件 SAS, S-plus 和 SPSS 中,SAS 是最好的。SAS 系统已由最初的统计分析软件发展成为大型集成应用软件系统,具有完备的数据访问、管理、分析、展现及应用开发等功能。在数据处理与统计分析领域,SAS 系统已成为国际上的标准软件。

我同意作者对掌握 SAS 技术应用水平的三层次分类,即:

第一个层次:会使用 SAS 菜单以及一些菜单界面的 SAS 模块,如 INSIGHT, ANALYST 等,了解初步的 SAS BASE 语句,能用 SAS 系统作简单的数据加工处理和分析,具有一定的数理统计知识,掌握一定的 SAS STAT 过程。

第二个层次:精通 SAS BASE,能用 SAS 语言编写复杂的 SAS 程序,能用 SAS BASE 进行大型的、复杂的数据加工整理和展现,掌握 SAS 和外部数据文件的接口,会进行复杂的统计建模和分析等,初步掌握一种基于 SAS 系统的开发工具。

第三个层次:在一、二层面基础上,进一步掌握 SAS/AF, SAS/IML 等,能开发基于 SAS 系统的数据管理和分析模块。

本书的目的是让读者的 SAS 技术达到第二层面水平,即能充分利用 SAS 的强大功能进行金融数据加工处理、数据整合、数据展现及简单的金融建模和计算。所以,读者一定要学好本书的 SAS 编程技术,它才是 SAS 软件的精华,只有利用 SAS 语言编写程序才能更有效地发挥 SAS 软件的强大功能。

特别指出的是作者多年来一直负责清华大学金融工程实验室和

金融数据中心的技术工作,率先在清华大学金融专业开设了金融数据库课程,使学生系统了解金融数据库结构、掌握数据处理和金融计算技术,为以后的学习、研究和工作打下了良好的基础。作者有深厚的数理统计、金融工程专业的知识背景,有多年应用 SAS 语言编程解决实际问题的经验;相信此书的含金量一定很高,值得阅读和作为手册使用。为此,我诚挚地将它推荐给读者。

宋逢明教授

2003 年 4 月于清华园

本书由中外金融数据库、SAS 编程技术和金融工具收益计算三大模块组成。中外金融数据库模块介绍了中外著名金融数据库的数据结构及本书所用的数据库。SAS 编程技术模块分为五个部分,系统介绍了基于 SAS 系统的数据加工整理、数据访问、数据管理和数据展现技术,以及 SAS 系统的开发工具 SAS/IML 软件。金融工具收益计算模块给出了基于 SAS 系统的简单金融建模和金融计算。

全书共分 7 大部分 20 章。第 1 章构成第一部分“中外金融数据库”,是相对独立的一部分内容,主要介绍了证券价格研究中心(CRSP)证券市场数据库和本书金融数据库的结构,路透(REUTER)动态金融数据库的使用方法。第二部分(第 2~5 章)为 SAS 语言的编程基础,介绍了 SAS 的编程环境和基本概念,怎样建立 SAS 数据集,怎样访问外部数据及 SAS 函数的使用方法;第三部分(第 6~9 章)为利用 SAS 数据步进行数据加工整理技术,包括有关管理 SAS 数据步产生文件技术,修改和选择数据集观测技术,控制程序循环和转移技术,控制变量及相关信息技术;第四部分(第 10~14 章)为 SAS 高级编程技术,包括 SAS 的一些通用语句、输出控制技术、输出输送系统(ODS)技术、输入输出格式控制技术和宏功能;第五部分(第 15~17 章)为 SAS 数据管理与展现技术,包括数据管理、统计量计算和数据展现技术;第 18 章构成第六部分,为 SAS/IML 软件编程技术;第七部分(第 19~20 章)为基于 SAS 的金融工具计算技术,包括金融资产收益计算和固定收益计算技术。

本书有以下特点:

- 编程技术与实际问题相结合。本书的三大模块内容充分体现了这一特点。

- 倡导编程技术,充分发挥 SAS 系统的优势。选用当今最好的应用软件平台 SAS 系统,倡导编程处理数据,使读者充分体会到 SAS 系统的强大功能,实现复杂的数据处理。

- 语言简练、准确。一般情况下,一句话完成对 SAS 语句、过程相关选项的解释,从而避免了目前一些翻译的 SAS 使用手册中出现的大段解释说明文字却不能讲清问题的情况。

- 内容编写便于学习。书中大量的可实现例程,对理解和掌握相应的内容起了关键的作用,使本书的学习变得容易和高效率,本书所附光盘提供的教学数据库和程序集为学好本书的内容提供了可靠的保证。这些例程不仅适合阅读和理解内容,也为读者在今后实际工

作中提供了大量可参考程序,可以作为有关 SAS 编程技术和金融计算的手册使用。

- 信息量大。大部分输出结果要求读者通过运行本书光盘中的相应程序得到,这样既便于巩固学习,又节省了大量篇幅,从而大大增加本书的信息量。

本书适合多层次多专业的人士阅读,如数学、统计、经济和金融等专业的本科生、研究生及相应部门的专业人员。

本书少部分内容得到了以下人士的帮助:陈建恒参与了第 2 章的写作,侯梧参与了第 20 章的写作。

感谢作者的两位恩师张尧庭教授和宋逢明教授多年来的教导和帮助,长期跟随他们进行教学和研究的经验是写作本书的必要条件之一。

感谢清华大学出版社的支持和帮助。

限于作者的水平,书中一定会存在不少不足之处,敬请读者提出宝贵建议并对有错误的地方进行指正,以便以后再版时加以改进。此外,作者还计划在清华大学出版社出版《基于 SAS 系统的金融计算》和《基于 SAS 系统的金融建模》两本图书,以便读者能对这一领域形成一个更全面、更深入的认识。

本书得到国家社会科学基金项目(02BJY132)和国家自然科学基金项目(70273016)的资助,一并致谢!

朱世武

2003 年 4 月于清华园

E-mail:zhushw@em. tsinghua. edu. cn

第一部分 中外金融数据库

第 1 章 金融数据库介绍	3
1.1 CRSP	3
1.1.1 CRSP 简介	3
1.1.2 股票数据库结构	4
1.1.3 股票数据文件描述	4
1.1.4 指数数据库结构图	10
1.1.5 指数数据文件描述	10
1.2 路透(Reuter)	13
1.2.1 路透终端	13
1.2.2 终端屏幕	13
1.2.3 报价	14
1.2.4 财经快讯	15
1.2.5 图表	16
1.2.6 其他	17
1.3 其他数据库	19
1.3.1 COMPUSTAT	19
1.3.2 PACAP	20
1.3.3 DATASTREAM	21
1.3.4 IBES	22
1.3.5 SDC	22
1.3.6 GovPX	23
1.3.7 BLOOMBERG	24
1.3.8 巨灵	24
1.4 本书教学金融数据	25
1.4.1 教学数据库(FDATA)	26
1.4.2 证券基础数据库(STOBASE)	26
1.4.3 个股数据库(STOINDIV)	29
1.4.4 复权个股数据库(STOINDIF)	30
1.4.5 股票高频数据库(HFDATA)	30

第二部分 SAS 编程基础

第 2 章 SAS 系统快速入门	35
2.1 SAS 系统简介	35
2.1.1 概况	35
2.1.2 SAS 功能模块	35
2.1.3 SAS 模块功能分类	38
2.1.4 SAS 系统特点	39
2.1.5 SAS 技术应用层次	39
2.1.6 本章目的	39
2.2 SAS 系统工作环境	40
2.2.1 SAS 工作界面	40
2.2.2 SAS 窗口	40
2.2.3 其他界面	43
2.2.4 浏览文件	45
2.2.5 数据库操作	45
2.2.6 创建文件快捷方式	47
2.2.7 功能键	48
2.3 存取与管理数据	49
2.3.1 浏览和编辑数据	49
2.3.2 导出数据	53
2.3.3 导入数据	53
2.3.4 创建查询	53
2.4 展现数据	60
2.4.1 产生 HTML 输出	60
2.4.2 以报表形式显示数据	61
2.4.3 用汇总报表显示数据	64
2.5 分析数据	65
2.5.1 一维数据分析	67
2.5.2 盒形图	67
2.5.3 马赛克图	68
2.5.4 二维数据分析	68
2.5.5 三维数据分析	69
2.5.6 回归分析	69
第 3 章 数据步创建 SAS 数据集	71
3.1 SAS 语言	71
3.1.1 SAS 语句	71



3.1.2	SAS 名与关键词	72
3.1.3	SAS 语句格式	72
3.2	数据步与过程步	73
3.2.1	数据步	73
3.2.2	过程步	73
3.3	数据集与变量	73
3.3.1	SAS 数据集	73
3.3.2	SAS 变量	74
3.4	SAS 程序	75
3.4.1	程序举例	75
3.4.2	书写规则	76
3.4.3	变量列表及其缩写方法	76
3.5	程序执行与输出	77
3.5.1	程序执行	77
3.5.2	程序输出	77
3.6	SAS 表达式	78
3.6.1	SAS 常数	78
3.6.2	SAS 算符	79
3.7	数值与字符之间的转换	82
3.7.1	字符转换为数值	82
3.7.2	数值转换为字符	84
3.8	错误类型与处理	85
3.8.1	句法错	86
3.8.2	词义错	87
3.8.3	运行错	88
3.8.4	数据错	88
3.9	SAS 文件系统	89
3.9.1	逻辑库	89
3.9.2	文件类型	90
3.10	数据步创建 SAS 数据集	91
3.10.1	数据存于非电子介质	91
3.10.2	数据存于文本文件	92
3.10.3	数据来自其他 SAS 数据集	93
3.10.4	输出外部文件	94
第 4 章 访问外部数据文件		95
4.1	概述	95
4.1.1	访问两类外部文件	95

4.1.2 访问外部数据文件方法	96
4.2 通过 IMPORT 过程	97
4.2.1 句法与选项说明	97
4.2.2 应用举例	98
4.3 通过 LIBNAME 语句和库引擎	99
4.3.1 读入其他版本或分析软件的数据集	99
4.3.2 读入流行数据库	99
4.4 通过 ACCESS 过程	100
4.4.1 创建访问描述器	100
4.4.2 创建数据视窗	102
4.4.3 由数据视窗创建数据集	103
4.5 通过 ODBC	104
4.5.1 创建 ODBC 数据源	104
4.5.2 创建 ODBC 引擎逻辑库	104
4.6 创建 SAS 数据集方法总结	105
第 5 章 SAS 函数及其应用	106
5.1 SAS 函数定义	106
5.1.1 函数定义	106
5.1.2 函数用法	106
5.2 SAS 函数自变量与结果	107
5.2.1 函数自变量	107
5.2.2 函数结果	108
5.2.3 显示函数值的简单方法	108
5.3 SAS 函数分类	109
5.4 日期时间函数	110
5.4.1 日期时间函数	110
5.4.2 应用举例	111
5.5 概率分布函数	114
5.5.1 标准正态分布	114
5.5.2 卡方分布	114
5.5.3 伽马分布	115
5.5.4 贝塔分布	115
5.5.5 F 分布	115
5.5.6 t 分布	115
5.5.7 二项分布	116
5.5.8 泊松分布	116
5.5.9 负二项分布	116

5.5.10	超几何分布	117
5.6	分位数函数	117
5.6.1	卡方分布分位数	117
5.6.2	贝塔分布分位数	117
5.6.3	F 分布分位数	118
5.6.4	t 分布分位数	118
5.6.5	正态分布分位数	118
5.6.6	伽马分布分位数	118
5.7	样本统计函数	119
5.7.1	均值	119
5.7.2	最大值	119
5.7.3	最小值	119
5.7.4	非缺失数据个数	119
5.7.5	缺失数据个数	119
5.7.6	求和	119
5.7.7	方差	119
5.7.8	标准差	119
5.7.9	标准误	120
5.7.10	变异系数	120
5.7.11	极差	120
5.7.12	校正平方和	120
5.7.13	未校正平方和	120
5.7.14	偏斜度	120
5.7.15	峰度	121
5.8	随机数函数	121
5.8.1	正态分布	121
5.8.2	均匀分布	122
5.8.3	二项分布	122
5.8.4	伽马分布	123
5.8.5	泊松分布	123
5.8.6	贝塔分布	124
5.8.7	指数分布	124
5.8.8	几何分布	125
5.8.9	极值分布	125
5.8.10	随机数函数的自变量 SEED	125
5.9	SAS CALL 子程序	126
5.9.1	CALL 子程序的类型	126
5.9.2	随机数子程序	126

第三部分 SAS 数据加工整理技术

第 6 章 数据步文件管理	131
6.1 DATA 语句	131
6.1.1 语句格式	131
6.1.2 选项说明	132
6.1.3 特殊数据集名	133
6.1.4 一个 DATA 语句下多个数据集名	133
6.2 INPUT 语句	134
6.2.1 语句格式	134
6.2.2 选项说明	135
6.2.3 列输入方式	135
6.2.4 列表方式输入	136
6.2.5 格式化输入方式	137
6.2.6 命名输入方式	138
6.3 CARDS 与 CARDS4 语句	139
6.3.1 CARDS 语句	139
6.3.2 CARDS4 语句	140
6.4 PUT 语句	140
6.4.1 语句格式	140
6.4.2 选项说明	141
6.4.3 应用举例	141
6.4.4 指针控制	142
6.4.5 列方式输出	143
6.4.6 列表方式输出	143
6.4.7 格式化输出	144
6.5 BY 语句	145
6.5.1 语句格式	145
6.5.2 选项说明	145
6.5.3 有关 BY 语句的概念	145
6.5.4 FIRST. 变量和 LAST. 变量	145
6.6 SET 语句	146
6.6.1 语句格式	146
6.6.2 选项说明	146
6.6.3 应用举例	146
6.7 MERGE 语句	149
6.7.1 语句格式	150
6.7.2 选项说明	150

6.7.3 应用举例	150
6.8 UPDATE 语句	151
6.8.1 语句格式	151
6.8.2 选项说明	151
6.8.3 应用举例	151
6.9 MODIFY 语句	152
6.9.1 语句格式	152
6.9.2 选项说明	152
6.9.3 应用举例	153
6.10 FILE 语句	154
6.10.1 语句格式	154
6.10.2 应用举例	155
6.11 INFILE 语句	156
6.11.1 语句格式	156
6.11.2 应用举例	157
第7章 数据加工整理:修改与选择观测	159
7.1 赋值语句	159
7.1.1 语句格式	159
7.1.2 结果变量类型	160
7.1.3 结果变量长度	160
7.2 累加语句	161
7.2.1 语句格式	161
7.2.2 应用举例	162
7.3 DELETE 语句与 LOSTCARD 语句	162
7.3.1 DELETE 语句	162
7.3.2 LOSTCARD 语句	163
7.4 STOP 语句与 ABORT 语句	165
7.4.1 STOP 语句	165
7.4.2 ABORT 语句	166
7.5 WHERE 语句	168
7.5.1 语句格式	168
7.5.2 WHERE 表达式特殊算符	168
7.5.3 应用举例	169
7.5.4 WHERE 和子集 IF 语句的比较	170
7.6 OUTPUT 语句	170
7.6.1 语句格式	170
7.6.2 应用举例	171

7.7	REMOVE 语句与 REPLACE 语句	172
7.7.1	REMOVE 语句	172
7.7.2	REPLACE 语句	173
7.8	MISSING 语句	174
7.8.1	语句格式	174
7.8.2	应用举例	175
7.9	其他语句	175
7.9.1	LIST 语句	175
7.9.2	PUT 语句与 LIST 语句比较	176
7.9.3	CALL 语句	176
7.9.4	CALL 语句调用的子程序	176
7.9.5	NULL 语句	177
7.9.6	ERROR 语句	178
第 8 章	数据加工整理: 循环与转移控制	180
8.1	DO 语句	180
8.1.1	简单 DO 语句	181
8.1.2	循环 DO 语句	181
8.1.3	DO OVER 语句	184
8.1.4	DO WHILE 语句	184
8.1.5	DO UNTIL 语句	185
8.2	END 语句	185
8.2.1	语句格式	185
8.2.2	应用举例	185
8.3	SELECT 语句	186
8.3.1	语句格式	186
8.3.2	应用举例	186
8.4	IF 语句	187
8.4.1	IF-THEN 与 IF-THEN/ELSE 语句	187
8.4.2	子集 IF 语句	188
8.5	GO TO 语句与语句标号	188
8.5.1	GO TO 语句	188
8.5.2	语句标号	190
8.6	LINK 语句	190
8.6.1	语句格式	190
8.6.2	LINK 语句与 GOTO 语句的差别	192
8.7	RETURN 语句	192
8.7.1	语句格式	192

8.7.2 应用举例	192
8.8 CONTINUE 语句与 LEAVE 语句	193
8.8.1 CONTINUE 语句	193
8.8.2 LEAVE 语句	194
8.8.3 LEAVE 语句与 CONTINUE 语句的差别	195
第9章 数据加工整理：变量与信息控制	196
9.1 ARRAY 语句	196
9.1.1 显式下标数组语句	196
9.1.2 引用显式下标数组元素	198
9.1.3 隐含下标数组语句	201
9.1.4 引用隐含数组元素	201
9.2 INFORMAT 语句与 FORMAT 语句	204
9.2.1 INFORMAT 语句	204
9.2.2 FORMAT 语句	204
9.3 LENGTH 语句与 LABEL 语句	206
9.3.1 LENGTH 语句	206
9.3.2 控制变量长度的方法	207
9.3.3 数值变量长度控制	207
9.3.4 字符变量长度控制	208
9.3.5 LENGTH 语句位置的重要性	208
9.3.6 LABEL 语句	210
9.4 ATTRIB 语句	211
9.4.1 语句格式	211
9.4.2 可以规定的变量属性	211
9.4.3 应用举例	211
9.5 DROP 语句与 KEEP 语句	212
9.5.1 DROP 语句	212
9.5.2 KEEP 语句	212
9.5.3 DROP 和 KEEP 语句的使用规则	213
9.6 RENAME 语句与 RETAIN 语句	214
9.6.1 RENAME 语句	214
9.6.2 RETAIN 语句	214

第四部分 SAS 高级编程技术

第10章 过程步通用语句	219
10.1 PROC 语句	219
10.1.1 语句格式	220

10.1.2	应用举例	220
10.2	VAR 语句与 MODLE 语句	221
10.2.1	VAR 语句	221
10.2.2	MODEL 语句	221
10.3	ID 语句与 WHERE 语句	221
10.3.1	ID 语句	221
10.3.2	WHERE 语句	222
10.4	CLASS 语句与 BY 语句	223
10.4.1	CLASS 语句	223
10.4.2	BY 语句	224
10.5	OUTPUT 语句与 QUIT 语句	224
10.5.1	OUTPUT 语句	224
10.5.2	QUIT 语句	225
10.6	FORMAT 语句与 ATTRIB 语句	225
10.6.1	FORMAT 语句	225
10.6.2	ATTRIB 语句	226
10.7	LABEL 语句	226
第 11 章	全程通用语句	227
11.1	注释语句	227
11.1.1	语句格式	227
11.1.2	应用举例	228
11.2	DM 语句与 X 语句	228
11.2.1	DM 语句	228
11.2.2	X 语句	229
11.3	TITLE 语句与 FOOTNOTE 语句	230
11.3.1	TITLE 语句	230
11.3.2	FOOTNOTE 语句	230
11.4	RUN 语句与 ENDSAS 语句	231
11.4.1	RUN 语句	231
11.4.2	ENDSAS 语句	232
11.5	LIBNAME 语句	232
11.5.1	语句格式	232
11.5.2	选项说明	232
11.5.3	应用举例	233
11.6	FILENAME 语句	233
11.6.1	语句格式	234
11.6.2	选项说明	234