

“十五”国家重点电子出版物规划项目·计算机知识普及和软件开发系列

专业程序员快速成长之路丛书 (5)

SAS 8.X

经济统计

北京希望电子出版社 总策划
樊欣 邵谦谦 编著



北京希望电子出版社
Beijing Hope Electronic Press
www.bhp.com.cn

“十五”国家重点电子出版物规划项目·计算机知识普及和软件开发系列

专业程序员快速成长之路丛书 (5)

SAS 8.X

经济统计

北京希望电子出版社 总策划
樊欣 邵谦谦 编著



北京希望电子出版社
Beijing Hope Electronic Press
www.bhp.com.cn

内 容 简 介

本书是一本关于 SAS 系统使用方法的实用指导书。SAS 系统是一个集大型数据库管理、统计分析、报表图形、信息系统开发等多种强大功能为一体的大型软件系统。本书通过 **191** 个经典实例，讲解了 SAS 的最基础、最重要的功能和模块。

全书由 13 章组成，主要内容包括：SAS 概述、SAS 基本概念、建立 SAS 数据集、数据集的浏览与修改、SAS 语言初步、SAS 语言中的常用函数、数组与矩阵、基于 SAS 语言的数据管理、报表以及图形输出、基本统计分析、多元统计分析、SAS INSIGHT 模块、SAS ANALYST 模块。

本书以“学以致用”为原则组织内容，使用实例典型、语言通俗简练、结构清晰、图文并茂、边讲边练。在阅读本书前，要求读者具有基础的统计知识，并且学过简单的计算机程序设计。本书适合 SAS 语言的初、中级读者，同时也可作为社会培训班教材。

读者在使用本书过程中的技术问题，请与 Xinfanapply@hotmail.com 联系。

本版 CD 为配套电子书。

盘书系列名： “十五”国家重点电子出版物规划项目•计算机知识普及和软件开发系列
专业程序员快速成长之路丛书（5）

盘 书 名： SAS 8.X 经济统计

总 策 划： 北京希望电子出版社

文本著作者： 樊欣 邵谦谦

责 任 编 辑： 杨如林

出版、发行者： 北京希望电子出版社

地 址： 北京市海淀区知春路甲63号卫星大厦三层 100080

网址: www.bhp.com.cn

E-mail: zwb@bhp.com.cn

电 话： 010-62520290,62521724,62528991,62630301,62524940,62521921,82610344

（发行）010-82675588-202（门市） 010-82675588-501,82675588-201（编辑部）

经 销： 各地新华书店、软件连锁店

排 版： 希望图书输出中心 张月岭

CD 生产者： 北京中新联光盘有限责任公司

文本印刷者： 北京媛明印刷厂

开本 / 规格： 787 毫米×1092 毫米 1/16 24 印张 557.3 千字

版次 / 印次： 2003 年 3 月第 1 版 2003 年 3 月第 1 次印刷

印 数： 0001~4000 册

本 版 号： ISBN 7-89498-096-X

定 价： 35.00 元（本版 CD）

说明：凡我社产品如有残缺，可执相关凭证与本社调换。

前言

处于当今信息化的时代,如何获得有用的信息、如何管理这些信息、如何从这些信息中分析抽象出有用的东西已经成为一个不容忽视的问题。归根到底,信息的载体是数据,所以如何获取数据、管理数据、分析数据就成了重要的问题。随着计算机应用的普及,出现了各种各样的软件来帮助我们管理数据、分析数据,大体可以分为两类:第一类是数据库软件,常见的如 Oracle, dBase, SQL sever 等。这一类软件的着眼点在于将数据管理计算机化,使得原本效率低下、成本昂贵的数据管理可以被方便、快捷地进行。这一类软件虽然解决了管理数据的问题,但是并没有对信息进行加工,只是采用另外的形式将它们组织、保存起来。企业如果仅仅停留在这一步,就失去了保留这些数据的真正意义。第二类软件是统计分析软件,常见的如 SAS, SPISS, S-plus 等。这一类软件着眼点在于从浩如烟海的信息中提取有用的数据(Data Mining),运用先进的统计方法对这些数据进行分析,从中获取事物的本质特征以及一般性的规律以支持我们做出决策。因此,从这一点来说数据库软件建立了信息化的基础,而统计分析软件则从原有信息中得到新的信息帮助我们决策。例如,使用数据库软件可以记录一家企业多年以来的经营情况,在此基础上使用统计分析软件可以分析出企业发展的趋势并对未来的经营情况做出预测。目前大多数国内的企业还停留在第一步上,没有充分认识到统计、分析数据的重要性。而在其他一些发达国家,统计分析已经成为企业经营决策的一个重要组成部分。可以预见,在不远的将来将会有大量的企业、科研机构需要具有熟练统计分析能力的人才,其中必然要求对统计软件的熟练使用。

SAS (Statistical Analysis System) 是由美国北卡来罗纳的 SAS institute 公司开发的一种统计软件,它可以被广泛应用在商业、科研领域。它不但具有强大的统计分析功能,也具有一般数据库软件的数据管理功能。在众多的统计软件中,它以运行稳定、功能强大而著称,近 20 年来一直占据统计软件的高端市场。在财富 500 强中,有 90% 的公司使用 SAS。而在财富 500 强的前 100 家企业中,更是有 98% 的公司使用 SAS。SAS 软件正在国内迅速地普及、发展,因此如何学会使用 SAS 已经成为当务之急。写作本书的目的也就是希望为读者提供一本入门的教材,帮助读者掌握 SAS 软件最基本、最重要的一些功能。

本书由樊欣执笔编写。此外,赵冰晶、李秀屹、程华遴、李冰、魏浩然、李晓东、王宏、李四琦、王怀瑾、吴子城、李小炎、刘大伟、赵华刚、朱育、赵晓燕、李晓阳、马军、朱玉成、成兵、林军、章浩、君华强、余波、杨俊、李大成、王豪梁、王军、程松、高荣益和葛艳玲等同志在整理材料方面给予了作者很大的帮助,在此,编者对他们表示衷心的感谢。

由于 SAS 软件涵盖的内容非常多,其中又有各种各样的应用模块,因此在介绍过程中难免有遗漏、不足之处,希望广大读者谅解。

编 者

目 录

第 1 章 SAS 概述.....	1	6.8 数学函数.....	76
1.1 SAS 可以做什么	1	6.9 概率与密度函数	77
1.2 SAS 系统的安装	8	6.10 分位数函数	77
1.3 SAS 系统的概貌	11	6.11 随机数函数	78
1.4 SAS 命令行和工具栏.....	19	6.12 样本统计函数	78
1.5 思考题.....	20	6.13 地区以及邮政编码函数	79
第 2 章 SAS 基本概念.....	21	6.14 三角函数	80
2.1 SAS 数据集	21	6.15 截断函数	80
2.2 SAS 数据库	24	6.16 变量信息函数	80
2.3 SAS 程序	26	6.17 日期时间函数	81
2.4 思考题.....	27	6.18 思考题	82
第 3 章 建立 SAS 数据集.....	28	第 7 章 数组与矩阵	83
3.1 编辑程序导入数据	28	7.1 数 组	83
3.2 利用 VIEWTABLE 新建数据集	31	7.2 SAS/IML 模块	89
3.3 将其他格式的数据文件导入数据集	33	7.3 思考题	127
3.4 利用 SAS ASSIST 创建数据集	39	第 8 章 基于 SAS 语言的数据管理	128
3.5 思考题.....	45	8.1 读入数据	128
第 4 章 数据集的浏览与修改	48	8.2 数据集的复制	131
4.1 利用 VIEWTABLE 方式	48	8.3 数据集的拆分	133
4.2 SAS INSIGHT	50	8.4 数据集的合并	134
4.3 SAS ASSIST	55	8.5 修改数据集	137
4.4 思考题.....	60	8.6 Proc SQL	138
第 5 章 SAS 语言初步.....	61	8.7 思考题	162
5.1 SAS 语言的组成	61	第 9 章 报表以及图形输出.....	164
5.2 SAS 用作一般高级语言	64	9.1 SAS 过程步简介	164
5.3 思考题.....	72	9.2 PRINT 过程.....	164
第 6 章 SAS 语言中的常用函数.....	73	9.3 TABULATE	170
6.1 算术函数.....	73	9.4 SORT 过程.....	172
6.2 逻辑函数.....	73	9.5 MEANS 过程	173
6.3 字符函数.....	74	9.6 UNIVARIATE 过程.....	174
6.4 字符串匹配函数.....	75	9.7 FREQ 过程.....	176
6.5 货币转换函数.....	75	9.8 CORR 过程	177
6.6 金融函数.....	75	9.9 GPLOT 过程	178
6.7 目录函数.....	76	9.10 GCHART 过程	181

9.11 G3D 过程与 GCONTOUR 过程	186	12.3 一维数据探索	285
9.12 思考题	188	12.4 二维数据探索	297
第 10 章 基本统计分析	190	12.5 三维数据探索	300
10.1 正态性检验	190	12.6 分布分析	305
10.2 单变量均值检验	194	12.7 曲线拟和	311
10.3 两独立样本的均值检验	195	12.8 多变量回归	316
10.4 成对总体均值检验	199	12.9 方差分析	319
10.5 回归分析	202	12.10 相关系数的计算	323
10.6 方差分析	217	12.11 主成分分析	325
10.7 列联表检验	235	12.12 变量转换	326
10.8 思考题	241	12.13 思考题	329
第 11 章 多元统计分析	242	第 13 章 SAS ANALYST 模块	330
11.1 主成分分析	242	13.1 基础知识	330
11.2 因子分析	254	13.2 数据集相关操作	333
11.3 聚类分析	259	13.3 绘制统计图	340
11.4 判别分析	270	13.4 统计分析与计算	347
11.5 思考题	280	13.5 假设检验	351
第 12 章 SAS INSIGHT 模块	282	13.6 ANOVA 过程	360
12.1 SAS INSIGHT 的启动	282	13.7 回归分析	370
12.2 SAS INSIGHT 菜单	283	13.8 思考题	376

第 1 章 SAS 概述

本章导读:

在我们日常的工作生活中,每天都要接触到大量的各种各样的数据。有很大一部分工作必须每天和大量的数据打交道,例如银行、保险公司这样的金融机构,需要时时刻刻清楚地了解自身的经营状况、客户的必要的信息;气象信息部门则需要处理实时的气象数据,同时也需要统计各种历史的气象数据,这样才能对天气做出较为准确的预报;大部分公司的市场销售部门需要详细分析市场和用户反馈的信息,这样才能不断改进产品提高产品的竞争力;生物制药公司需要处理大量的药品试用数据,这样才能最后得到安全可靠的药品。正因为生活中充满了不确定性,而人们总是试图通过一些实验数据来找出背后具有规律性的东西,所以统计学正被越来越广泛地应用到生活工作的每一个角落。

如何从这些纷繁复杂的数据中采集有用的信息,如何管理这些数据,如何对这些数据进行统计分析以得到可以帮助我们做出决策的信息就成了一个非常重要的问题。随着数据量的增大和人们对于统计要求的增加,要想单纯依靠传统的办法完成这项工作已经基本不太可能。

幸好,统计软件可以帮我们完成这项工作。目前,使用较为广泛的统计软件有 SAS, SPISS, GAUSS, S-plus 等等。这些统计软件可以在很大程度上帮助我们找到有用的数据,很好地管理这些数据,并且分析这些数据的内部关系和属性,为以后的决策提供预报。SAS 系统作为同类产品中的佼佼者,功能非常强大。在本章中,我们首先介绍一下 SAS 的概况、SAS 系统的主要界面和菜单。这些内容将成为以后我们讲解的基础。

1.1 SAS 可以做什么

SAS 是由美国的 SAS institute 开发的统计软件,它的全称是 Statistical Analysis System,即统计分析系统。从 1976 年 SAS 开发成功至今,SAS 的用户遍及 119 个国家,它已经成为同类产品中的领导者。在财富 500 强中,有 90% 的公司使用 SAS。而在财富 500 强的前 100 家企业中,有 98% 的公司使用 SAS。目前国内的很多企业也开始使用 SAS 系统,来分析处理数据。尽管 SAS 在国内的普及程度还远远不及美国和欧洲一些国家,但是相信随着加入 WTO 后对数据统计需求的增加,企业对数据统计分析将会越来越重视,会有越来越多的企业使用 SAS 系统。那么,SAS 系统到底可以帮我们作些什么呢?下面通过一个实例解释一下。

假设投资基金 A 的投资组合包括以下 36 只股票。表 1-1 给出了这 36 只股票的股票代码、股票名称、流通盘、每股收益、持股数量和价格。

表 1-1 基金投资组合

证券代码	证券名称	流通盘	每股收益	持股数量	价格
000096	广聚能源	8500	0.059	1000	13.27
000099	中信海直	6000	0.028	2000	14.2
000150	ST 麦科特	12600	-0.003	1500	7.12
000151	中成股份	10500	0.026	1300	10.08
000153	新力药业	2500	0.056	2000	22.75
000155	川化股份	13000	-0.009	5000	6.85
000156	安塑股份	3600	0.033	2000	14.95
000157	中联重科	10000	0.06	3000	12.65
000158	常山股份	10000	0.018	2400	8.38
000159	国际实业	7000	0.008	3100	12.15
000301	丝绸股份	15365	0.04	2000	7.31
000488	晨鸣纸业	7700	0.101	1000	13.26
000725	京东方	6000	0.044	2000	12.33
000835	隆源实业	1338	0.07	1000	22.58
000869	张裕 A	3200	0.194	1300	18.29
000877	天山股份	7800	-0.084	3000	12.55
000885	ST 春都	6000	-0.073	2000	12.48
000890	法尔胜	16934	0.031	1000	9.12
000892	长丰通信	12000	0.031	1000	7.88
000897	津滨发展	14166	0.002	5000	6.91
000900	现代投资	21423	0.058	3000	8.59
000901	航天科技	4800	0.005	1000	27.95
000902	中国服装	6500	-0.031	2000	10.92
000903	云内动力	6000	0.109	3000	11.79
000905	厦门路桥	9500	0.046	1000	9.29
000906	南方建材	6650	0.007	2000	14.47
000908	天一科技	8988	0.006	3000	8.28
000909	数源科技	6000	0.002	1000	9.99
000910	大亚股份	8000	0.036	2000	8.9
000911	南宁糖业	7280	0.067	2400	9.01

续 表

证券代码	证券名称	流通盘	每股收益	持股数量	价格
000912	泸天化	15000	0.112	1500	8.06
000913	钱江摩托	8450	0.062	3000	11.86
000915	山大华特	4599	0.001	1000	14.4
000916	华北高速	34000	0.038	5000	5.15
000917	电广传媒	11800	0.086	3000	16.23
000918	亚华种业	6000	-0.045	1000	10.12

在将这些数据读入 SAS 系统后, 这些数据就被称为一个 SAS 数据集, SAS 可以对它进行各种各样的统计分析。例如, 想了解一下这些股票流通盘大小的分布情况, 看看在投资组合中有多少只股票属于大盘股, 有多少只股票属于小盘股。最直观的方法就是画出它们的直方图, 使用 SAS 系统画出的流通盘的分布直方图如图 1-1 所示。

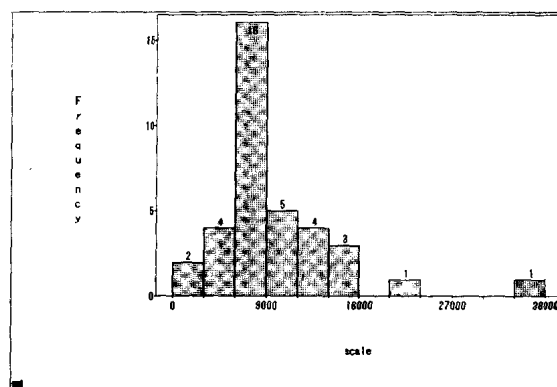


图 1-1 直方图

可以看到, 持有的股票的流通盘大部分在 18000 万以下, 有 34 家, 18000 万以上的只有两家。而在流通盘 18000 万以下的股票中, 位于 6000 万到 9000 万的股票最多。可见, 投资组合的主要是由流通盘适中的股票组成的, 小盘股、大盘股占的比例都不太大。

除了用直方图来表示流通盘大小的分布之外, 还可以使用盒形图来表示。盒形图是一种很常用的图示方法, 因为它不但可以反映分布情况, 还可以给出中位数、四分之一、四分之三分位数。盒形的中间有一条粗线, 这是数据分布的中位数的位置, 盒子的上边线是分布的四分之三分位数, 下边是分布的四分之一分位数, 盒子的上下边线包含了分布的 50% 的观测。盒子的长度叫做分布的四分位间距, 其作用类似于标准差, 可以反映数据分布的分散程度。从盒子边线向外画了两条线叫作触须线, 最长可以延长到四分位间距的 1.5 倍, 但是如果已经到了数据的最小值或最大值处就不再延伸。如果触须线没有达到数据的极端值, 则这些数据点用触须以外的点画出, 一般认为这样的点是异常点。图中就存在两个异常点。另外, 从盒形图可以看出数据的偏斜情况。如图 1-2 所示, 盒子的上半部分比下半部分长, 说明流通盘大小的分布略微右偏。

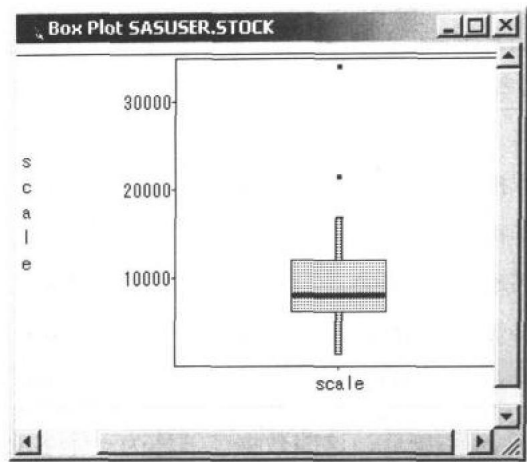


图 1-2 盒形图

在对数据有了直观的了解后，如果想进一步了解变量的一些统计数据，通过简单的操作，SAS 给出了流通盘这个变量的各种统计量，如图 1-3 所示。

Moments			
N	36.000000	Sum Wgts	36.000000
Mean	0.033083	Sum	1.191000
Std Dev	0.052345	Variance	0.002740
Skewness	0.402217	Kurtosis	1.945599
USS	0.135303	CSS	0.095901
CV	158.222309	Std Mean	0.008724

Quantiles			
100% Max	0.19400	99.0%	0.19400
75% Q3	0.05950	97.5%	0.19400
50% Med	0.03200	95.0%	0.11200
25% Q1	0.00350	90.0%	0.10100
0% Min	-0.08400	10.0%	-0.03100
Range	0.27800	5.0%	-0.07300
Q3-Q1	0.05600	2.5%	-0.08400
Mode	0.00200	1.0%	-0.08400

图 1-3 变量的常用统计量

图 1-3 中的表格是计算出的一些常用统计量，这些统计量的意义如下。

- N: 样本数量 n

$$\sum_{i=1}^n \omega_i y_i$$

- Sum Wgts: 加权和

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

- Mean: 均值

$$\sum_{i=1}^n y_i$$

- Sum: 总和

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}$$

- Std Dev: 标准偏差

- Variance: 方差
$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$
- Skewness: 偏度
$$\frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{y_i - \bar{y}}{s} \right)^3$$
- Kurtosis: 峰度
$$\frac{n(n+1)}{(n-1)(n-2)(n-3)} \frac{\sum_{i=1}^n (y_i - \bar{y})^4}{s^4} - \frac{3(n-1)^2}{(n-2)(n-3)}$$
- USS: 加权平方和
$$\sum_{i=1}^n \omega_i y_i^2$$
- CSS: 加权离差平方和
$$\sum_{i=1}^n \omega_i (y_i - \bar{y})^2$$
- CV: 变异系数
$$\frac{s}{\bar{y}} * 100$$
- Std Mean: 均值的标准误差
$$\frac{s}{\sqrt{n}}$$

分位数表中, Max 是最大值, Q3 是四分之三分位数, Med 是中位数 (反映数据的中心位置), Q1 是四分之一分位数, Min 是最小值, Range 是最大值减最小值, Q3-Q1 为四分位间距, 可以反映数据取值分散程度, Mode 是众数, 即出现次数最多的值。

除了了解单个变量的分布情况之外, 可能还比较关心股票价格到底和哪些变量有关系。在 SAS 中, 做出两个变量的折线图, 从折线图中可以看出大致的关系。然后使用 SAS 的拟和功能, 拟和出反映变量之间关系的曲线。在这个例子中, 我们认为股票价格可能和流通盘、每股收益有关系, 所以首先画出相应的折线图, 如图 1-4, 图 1-5 所示。

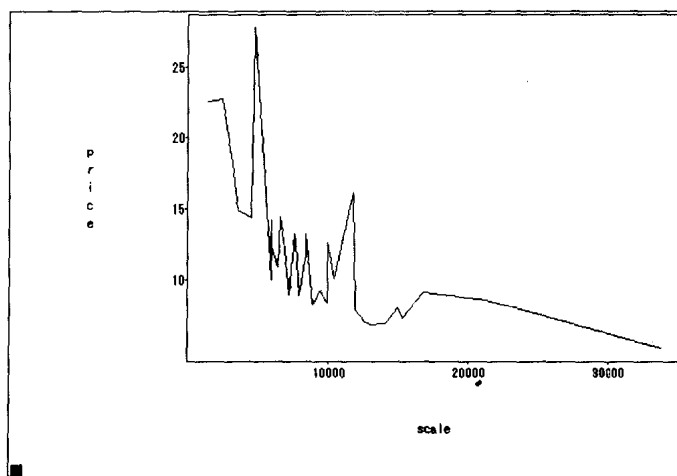


图 1-4 股票价格对流通盘的折线图

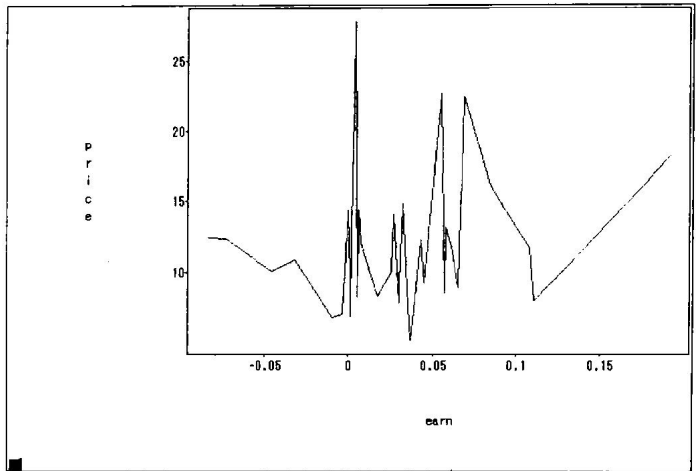


图 1-5 股票价格对每股收益的折线图

从折线图中，可以推测股票的价格与它的流通盘和每股收益有一定的关系。然后利用 SAS 的拟和功能，可以方便地得到股票价格与流通盘和每股收益的数学关系如图 1-6 所示。

从图 1-6 中可以看出，它们满足的关系式如下：

股票价格 (price) = 16.0976 - 0.0005 * 流通盘 (scale) + 18.3065 * 每股收益 (earn)。结果的其他部分是对拟和的检验结果。这个公式反应了投资组合中的股票价格与流通盘、每股收益这两个变量的关系，但是也可以用这个公式来粗略地判断组合外的一只股票的价格是否合适。方法很简单，就是把它的流通盘和每股收益这两个数据代入公式，然后就可以得到一个理论上的股票价格。利用这个理论上的股票价格，就可以判断这只股票在目前市场上的价格偏高还是偏低。

price = scale earn							
Response Distribution: Normal							
Link Function: Identity							
Model Equation							
price = 16.0976 - 0.0005 scale + 18.3065 earn							
Summary of Fit							
Mean of Response	11.9478	R-Square	0.4234				
Root MSE	3.8125	Adj R-Sq	0.3884				
Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Stat	Pr > F		
Model	2	352.2054	176.1027	12.12	0.0001		
Error	33	479.6576	14.5351				
C Total	35	831.8630					
Type III Tests							
Source	DF	Sum of Squares	Mean Square	F Stat	Pr > F		
scale	1	322.8364	322.8364	22.21	<.0001		
earn	1	32.1330	32.1330	2.21	0.1465		
Parameter Estimates							
Variable	DF	Estimate	Std Error	t Stat	Pr > t	Tolerance	Var Inflation
Intercept	1	16.0976	1.2556	12.82	<.0001		0
scale	1	-0.0005	0.0001	-4.71	<.0001	0.9998	1.0002
earn	1	18.3065	12.3123	1.49	0.1465	0.9998	1.0002

图 1-6 拟和结果

例如一只股票的流通盘是 8000 万，每股收益是 0.30 元，那么可以用拟和出的公式计算出它的理论价格：

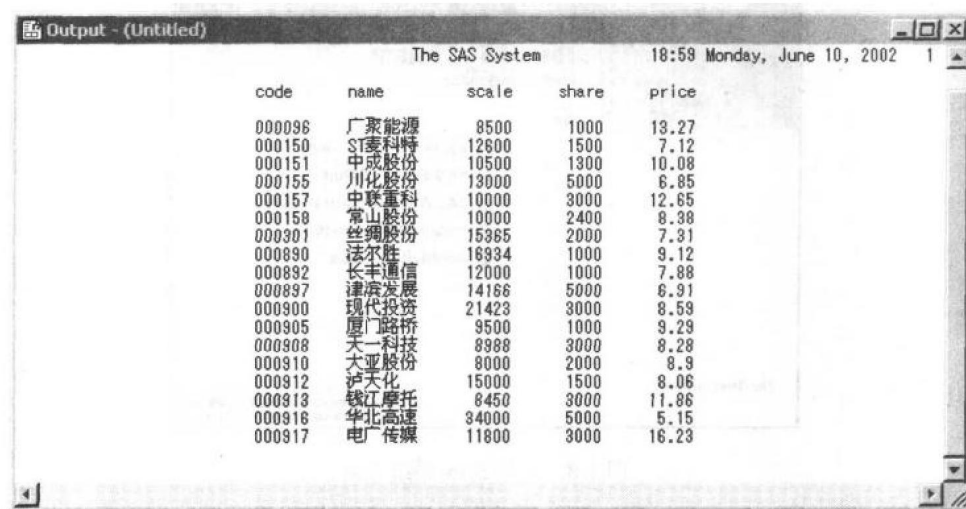
$$\begin{aligned}\text{股票价格 (price)} &= 16.0976 - 0.0005 * 8000 + 18.3065 * 0.30 \\ &= 17.58955\end{aligned}$$

当然我们计算出的价格和这只股票的实际价格会有一定的差异，但是从这个例子我们可以体会到统计在实际工作中的意义。首先从数据出发，通过一些统计方法来揭示一些规律。然后，可以利用这些规律来指导将来的或是类似的工作。

在某些时候，需要对数据集进行查询，寻找一些满足特殊条件的记录。例如，找出投资组合中流通盘大于 8000 万的股票，可以在编辑窗口中编写一个简单的 SQL 查询语句完成这个操作。

```
proc sql;
  select stock.code,name,scale,share,price
  from sasuser.stock
  where stock.scale>=8000;
run;
```

在输出窗口中可以得到查询结果，如图 1-7 所示。



code	name	scale	share	price
000096	广聚能源	8500	1000	13.27
000150	ST麦科特	12600	1500	7.12
000151	中成股份	10500	1300	10.08
000155	川化股份	13000	5000	6.85
000157	中联重科	10000	3000	12.65
000158	常山股份	10000	2400	8.38
000301	丝绸股份	15365	2000	7.31
000890	法尔胜	16934	1000	9.12
000892	长丰通信	12000	1000	7.88
000897	津滨发展	14166	5000	6.91
000900	现代投资	21423	3000	8.59
000905	厦门路桥	9500	1000	9.29
000908	天一科技	8988	3000	8.28
000910	大亚股份	8000	2000	8.9
000912	沪太化	15000	1500	8.06
000913	钱江摩托	8450	3000	11.86
000916	华北高速	34000	5000	5.15
000917	广电传媒	11800	3000	16.23

图 1-7 查询结果

本书采用的版本是 SAS 8.2，它的主要功能如下。

- 数据获取：SAS 可以从以各种形式存储的数据中获取有用的信息。数据的来源可以是长度不确定的记录、二进制的文件、自由格式的数据、甚至包括残缺数据的文件。并且可以通过 SAS 系统中的各种界面方便地访问这些数据。
- 数据管理：一旦将数据导入 SAS 系统后，用户可以根据自己地需求方便地管理这些数据。在 SAS 系统中可以使用 SQL 语言对数据库进行查询、创建、升级等操作。另外，SAS 系统还提供了专门的查询窗口。在这个窗口中，只要通过单击鼠标就可以完成对数据的查询。这样，即使一个完全不懂 SQL 语句的用户也可以方便地进行数据查询。

- 数据分析：SAS 系统提供了功能强大的数据分析工具。除了可以计算出一些常用的统计量（均值、总和、中值、方差、标准差、相关系数等等）之外，还可以进行大量的统计推断。
 - 数据表示：SAS 系统除了可以以列表形式给出数据外，还可以根据用户的需要给出各种各样的输出形式。很多种格式只需要简单的定义，并不用复杂的编程。
 - 应用软件开发：SAS 系统提供了灵活、方便的应用程序开发环境。在交互式的窗口内，可以通过简单易用的命令、特制的工具条和动态菜单方便地进行开发。
- 在以后各章中，将通过大量生动的实例来讲解如何使用 SAS 的这些功能。

1.2 SAS 系统的安装

在本节中，我们介绍 SAS 系统的安装过程。SAS 系统的完整安装大概需要 1G 左右的磁盘空间。SAS 的安装与其他软件的安装大同小异，将安装光盘放入光驱后，会自动打开安装界面，如图 1-8 所示。

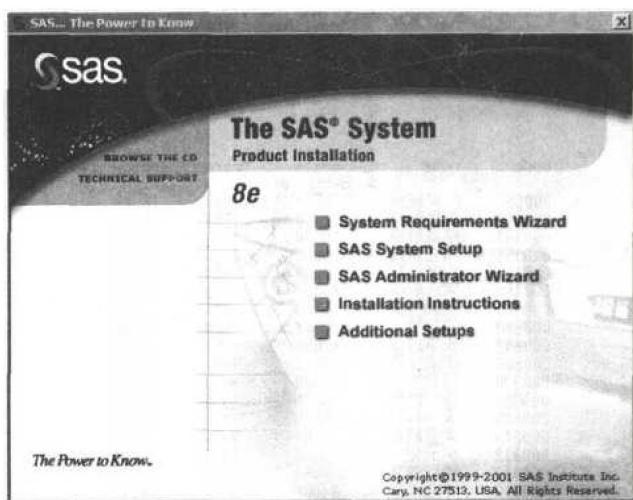


图 1-8 开始 SAS 系统安装

安装页面上有 5 个选项，它们的功能如下。

- System Requirement Wizard：将操作系统设置为满足 SAS 运行的最低系统要求。
- SAS System Setup：安装 SAS 系统。
- SAS Administrator Wizard：在网络驱动器上创建 SAS 安装文件，可供其他使用者安装。
- Installation Instructions：安装 SAS 系统的说明。
- Additional Setups：安装光盘上的其他 SAS 组件。

选择 SAS System Setup 开始安装。系统提示开始安装 SAS 系统，单击 Next，弹出如图 1-9 所示的窗口。

指定光盘中的 SAS 安装许可文件的位置，指定到 sasinst.sss 这个文件的位置即可，单击 Next 按钮。

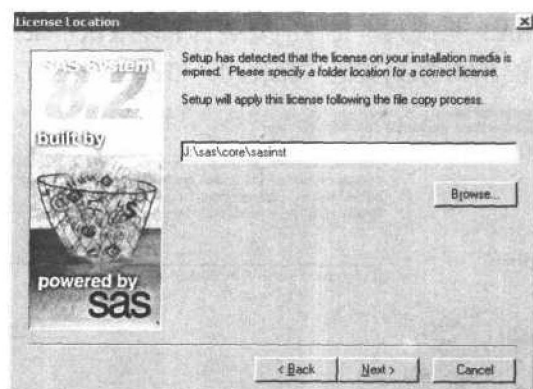


图 1-9 指定安装许可文件的位置

在图 1-10 中选择 SAS 系统使用的语言，选择 English，单击 Next 按钮。

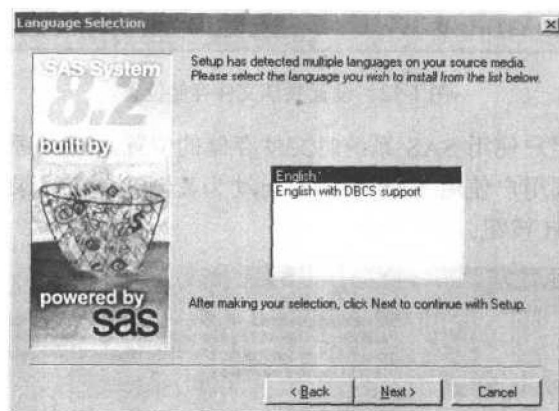


图 1-10 选择 SAS 系统语言种类

选择安装的类型，是完全安装还是选择性安装，在图 1-11 中选择完全安装 Complete，单击 Next 按钮。

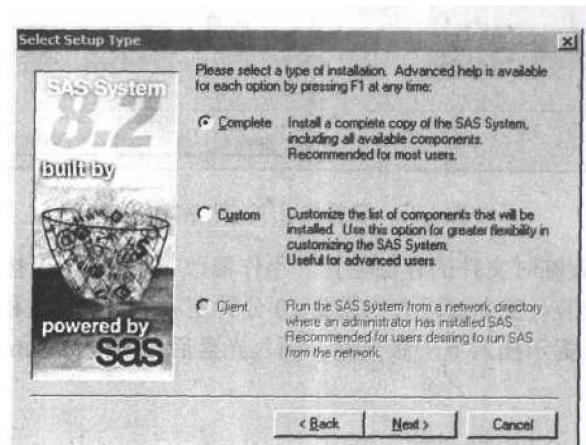


图 1-11 选择安装类型

在图 1-12 中设置 SAS 系统共享文件的安装目录，默认目录是 c:\Program Files\SAS Institute。共享文件可以被 SAS 系列其他产品使用。修改完路径后，单击 Next 按钮。

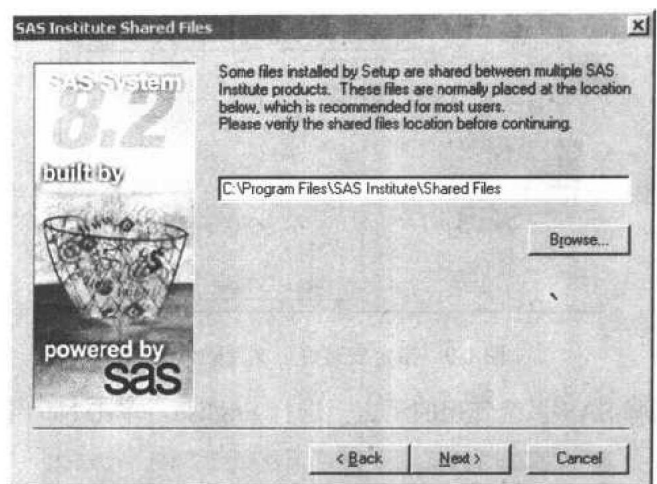


图 1-12 设置共享文件存储位置

系统要求设置多用户使用 SAS 系统时文件存储的位置。SAS 系统会根据 Windows 的用户配置文件，在不同用户使用 SAS 系统时自动为其创建一个目录，如图 1-13 所示。修改完路径后，单击 Next 按钮。

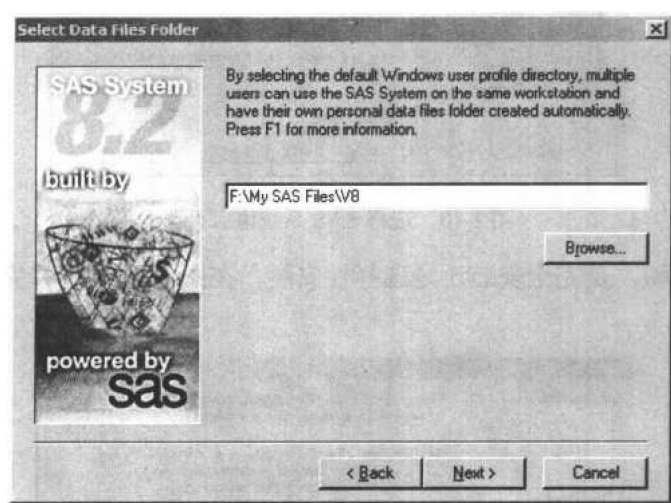


图 1-13 设置多用户文件存储位置

在图 1-14 中设置临时文件的存储位置。不作修改，单击 Next 按钮。然后就可以从光盘上复制文件开始安装，整个过程大约需要 10 分钟左右。在安装过程中，当一张光盘上的内容安装完毕后，会提示插入下一张光盘。插入光盘后，单击 Resume 按钮即可。

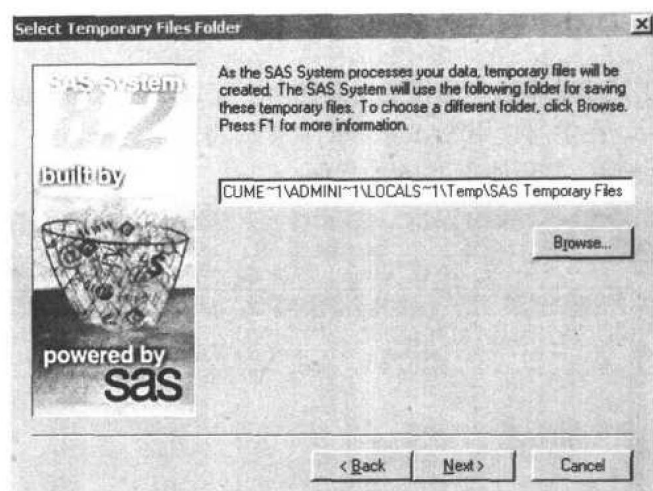


图 1-14 设置临时文件存储位置

在图 1-15 中单击 Finish 按钮，就完成了整个 SAS 系统的安装。

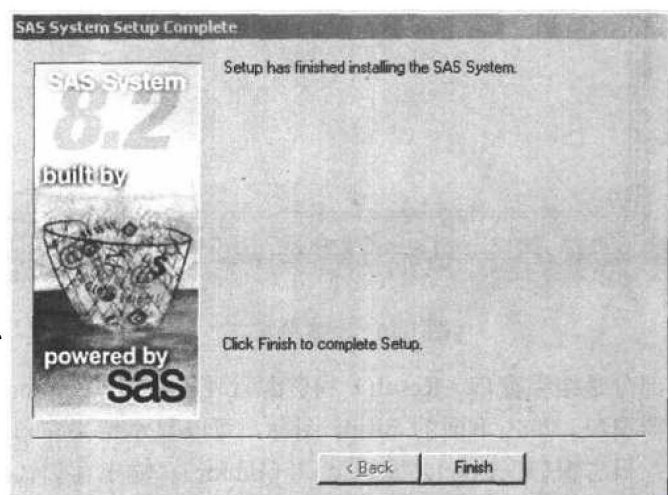


图 1-15 SAS 安装完成界面

1.3 SAS 系统的概貌

在本节中，将介绍 SAS 系统的启动、主要窗口、文件管理等基本概念。

1.3.1 SAS 系统的启动

SAS 系统安装完毕以后，就可以启动 SAS 系统了，方法有 2 种。

- 双击桌面上的 SAS 快捷方式。
- 单击屏幕左下角的“开始”菜单，在菜单中选择“程序”，在程序中选择 The SAS System，最后选择 The SAS System for Windows V8。