

赢
在职场
第一步



Stata

统计分析与应用

职场训练
案例教学



1235分钟多媒体教学全程实录
课程实验+技术讲解+上机练习

周广肃 梁荣 田金秀 等编著

互动答疑
QQ: 16102537



机械工业出版社
China Machine Press

赢
在职场
第一步

Stata

统计分析与应用

职场训练
案例教学



1235分钟多媒体教学全程实录

课程实验+技术讲解+上机练习

周广肃 梁荣 田金秀 等编著

互动答疑

QQ: 16102537



机械工业出版社
China Machine Press

Stata软件是由Stata公司在1985年推出的,具有强大的数据处理和分析功能,被广泛地应用于统计学、经济学、生物学、医药学、社会学、人口学等领域,功能十分强大,现在普遍流行使用的版本为Stata 10.0。

本书主要讲解数据的基本处理、图形的绘制、统计分析、回归与建模分析、编程等方面的内容,讲解力求细致全面,从而使读者熟悉和掌握Stata 10.0的各种功能操作。另外,本书每章的后面附有习题,目的是培养读者的动手能力,使读者在实际练习的过程中能快速提高应用水平。本书作为实验教程,十分注重内容的实用性,不仅立足于典型案例进行教学安排,还补充了所有案例背后所对应的模型和原理,方便读者巩固理论知识。

本书内容丰富,结构清晰,层次分明,语言通俗易懂,是一本较为实用的Stata实验教程。本书面向大中专院校经济管理专业及相关的社会科学类学生,特别是具有一定统计学和计量经济学基础知识的学生,以及企事业单位和其他相关领域的科研工作人员。

封底无防伪标均为盗版

版权所有,侵权必究

本书法律顾问 北京市展达律师事务所

图书在版编目(CIP)数据

Stata统计分析与应用 / 周广肃等编著. - 北京: 机械工业出版社, 2011.9

(赢在职场第一步)

ISBN 978-7-111-35594-6

I. ①S… II. ①周… III. ①统计分析—应用软件, Stata IV. ①C819

中国版本图书馆CIP数据核字(2011)第156974号

机械工业出版社(北京市西城区百万庄大街22号 邮政编码100037)

责任编辑:夏非彼 迟振春

中国电影出版社印刷厂印刷

2011年9月第1版第1次印刷

188mm×260mm·30.5印张(含0.25印张彩插)

标准书号: ISBN 978-7-111-35594-6

ISBN 978-7-89433-076-5(光盘)

定价: 65.00元(附1DVD)

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

客服热线: (010) 88378991; 82728184

购书热线: (010) 68326294; 88379649; 68995259

投稿热线: (010) 82728184; 88379603

读者信箱: booksaga@126.com



多媒体光盘使用说明

将光盘放入光驱后，依次双击“我的电脑”、光盘驱动器，即可出现如图0.1所示的界面。

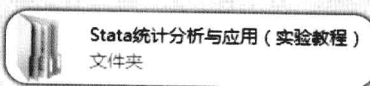


图0.1

Stata统计分析与应用（实验教程）文件夹包括光盘素材和正文文档两个子文件夹，如图0.2所示。

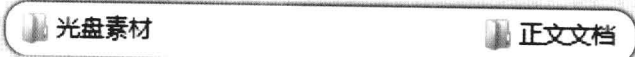


图0.2

正文文档文件夹中的文件编号与书中章号完全一致，如图0.3所示。

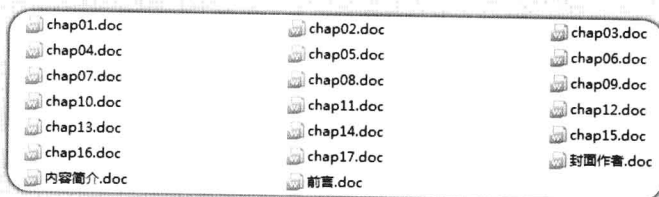


图0.3

光盘素材文件夹中的文件编号与书中素材的编号完全一致，包括data、PPT和video 3个子文件夹，以及Stata光盘使用说明文件，如图0.4所示。其中data文件夹包括教程数据和习题数据两个子文件夹。

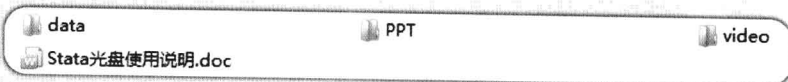


图0.4

video文件夹包含17章的视频文件压缩包，先解压缩，再根据学习需要双击视频文件，即可开始播放，界面如图0.5所示。

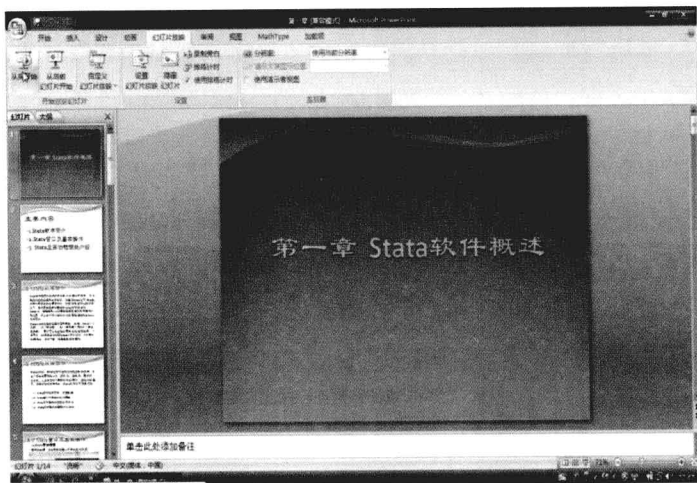
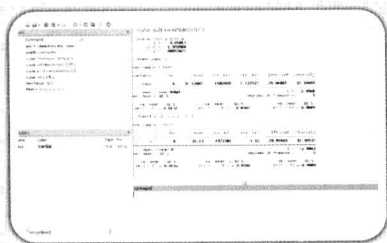
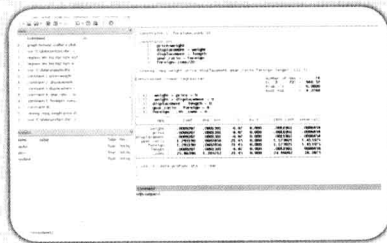


图0.5

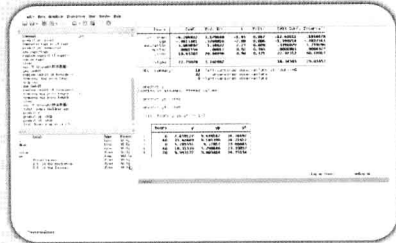
各实例多媒体预览



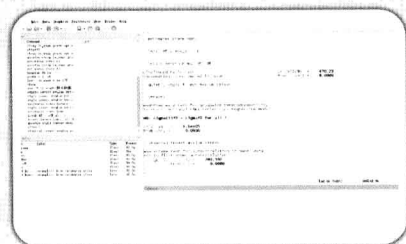
实验5.1



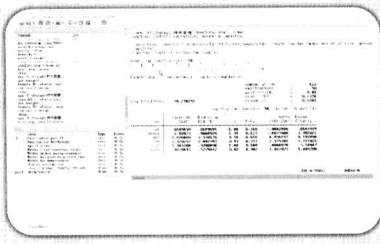
实验6.4



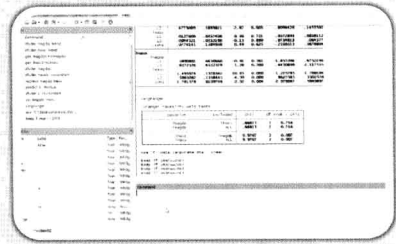
第11章习题



实验13.3



第15章习题



综合案例一

在当今信息化社会，计算机已经成为各个行业必不可少的应用工具。掌握各种软件的使用已然成为计算机从业人员最基本的技能，而行业软件的使用更是其生存之本。

行业应用软件与常用办公软件的基本操作类似，其主要特点是专业化的功能设置，该类软件的功能往往与相应行业技术紧密相连，主体功能完全按行业需求开发，所以行业软件的学习就有其固有的特点。本套丛书就专注于此类软件的典型应用。

行业软件的学习特点

软件风格决定软件学习的特点。在大学或职业教育学校的课程体系中，工程、财务、统计、数学等课程都会开设很多实验课，老师也会准备许多与实际工作比较接近的课程实验让学生上机练习，以此帮助他们提高实践能力。通过一个又一个的课程实验，学生不但能巩固所学的专业知识，还能掌握用相关软件解决实际问题的方法。

鉴于此，实践型教学模式就成为一种比较适合的教学方法。通过精心设计一系列接近实际工作的任务或者从实际工作中挑选一系列典型的工作案例，赋予读者一个模拟实际的工作场景，然后通过以下 4 个步骤来使其融会贯通，掌握行业软件的应用：

- 通过对一个实验基本原理的分析，让用户了解完成这个实验需要用到的专业背景知识及相关软件技术；
- 通过对实现过程的演示讲解，让用户掌握行业软件如何解决实际问题；
- 通过对运行结果的专业解释，让用户了解相关结果的具体含义，更深入地理解实验结果的内涵；
- 通过一些真实的应用案例进行上机练习，让用户巩固所学知识，从而达到举一反三的效果。

本套丛书的组成

行业应用软件最典型和最广泛的应用莫过于工程制图软件、财务软件、统计软件、工程项目管理软件、工程计算软件。工程制图软件中以 AutoCAD 为代表；财务软件中以用友 ERP、金蝶、Excel 为代表；统计软件中以 SPSS、EViews、SAS、Excel、Stata 为代表；工程项目管理软件中以 Project 为代表；工程计算软件中以 MATLAB、Fortran 为代表。

本套丛书选择以上几个行业应用最为广泛的软件进行介绍，首批推出以下产品：

- AutoCAD 2010 中文版电气设计与应用；



- AutoCAD 2010 中文版室内装潢设计与应用；
- Excel 在会计和财务管理中的应用；
- 用友 ERP-U8 财务管理与应用；
- 金蝶 K/3 财务管理与应用；
- SPSS 统计分析与应用；
- EViews 统计分析与应用；
- Excel 统计分析与应用；
- Stata 统计分析与应用；
- SAS 统计分析与应用；
- MATLAB 数值计算与应用。

丛书特色和目标读者

丛书特色如下：

- 完全通过有实际应用背景的“实验”案例进行教学，教学流程清晰完整，实验覆盖面和类型达岗位所需专业技能的 90% 以上，突出对读者就业能力的培养；
- 章后准备了大量与实验类似的项目，让读者自行上机演练，从而巩固提高，达到举一反三的效果；
- 针对不同的行业软件，设计相应的实验案例结构，AutoCAD 按照工程制图类型来组织图纸，其他软件的实验案例都按“实验基本原理→实验目的与要求→实验内容及数据来源→实验操作指导→实验结论”的思路组织内容；
- 所有实验和上机演练题都提供全程多媒体语音教学，方便自学，也有利于检验学习效果。

本丛书定位于使用相关行业软件的初中级用户、欲跨入相关行业的初学者以及学习相关课程的在校学生或技术人员。

丛书寄语

在这套丛书里，实验是最大的特色：通过实验演示软件使用技能；通过实验阐述行业应用方法；通过实验提高实践能力。实验是敲门砖，实验是您打开技能之门的钥匙。

我们真诚地希望本套丛书能够给即将走上工作岗位，或者正在进行专业技能学习的您提供帮助，帮助您走好职场的第一步，争取赢在第一步。

丛书编委会
2011 年 5 月

前 言

Stata 与 SAS、SPSS 统称为三大权威统计软件，它被广泛地应用于统计学、经济学、生物学、医药学、社会学、人口学等领域，功能十分强大。与其他软件相比，Stata 具有以下明显优势：操作简单，方便掌握；系统开放，及时更新；数据处理，功能强大；图形制作，精美优良；等等。Stata 软件具有五大功能，分别为数据的基本处理、图形的绘制、统计分析、回归与建模分析、编程等。该软件是由 Stata 公司在 1985 年推出的，到现在已经有二十多年的历史了。虽然现在最新的版本为 2009 年推出的 Stata 11.0，但是鉴于 11.0 版本还基本没有在中国推广和使用，所以本书所介绍的功能主要是通过 Stata 10.0 来实现的。

本书共分 17 章，编写时采用先讲解各个实验案例的理论依据和模型，然后讲解进行该种操作的通用命令模块，最后讲解本案例操作的具体命令语句和命令执行结果的思路。第 1 章是 Stata 软件概述，主要介绍 Stata 软件的窗口、基本操作及主要功能模块；第 2 章主要介绍数据的类型、压缩、转化、导入及整理；第 3 章主要介绍 Stata 的图形制作，主要包括图形制作的基本命令与相关操作，直方图、散点图、曲线标绘图、条形图、饼图、箱线图的绘制方法，以及图形的保存、合并和修改；第 4 章主要介绍单因素方差分析、双因素或多因素方差分析、协方差分析的相关实验内容；第 5 章主要介绍单个总体的假设检验和两个总体的假设检验等相关内容；第 6 章主要介绍小样本、大样本 OLS 估计的原理、操作方法和结果的解读，然后介绍约束回归和非线性最小二乘的相关实验内容；第 7 章主要介绍遗漏变量的检验、解释变量个数的选择、多重共线性与逐步回归法、极端数据的诊断与处理、虚拟变量的处理、经济结构变动的 Chow 检验等；第 8 章主要介绍异方差检验与处理，自相关和可行广义最小二乘法，内生性与 2SLS、GMM；第 9 章主要介绍二值选择模型、多值选择模型、排序数据模型、条件和嵌套 logit 模型等内容；第 10 章主要介绍泊松回归模型、负二项和广义负二项回归模型、零膨胀回归模型等；第 11 章主要介绍断尾回归模型、截取回归模型及样本选择模型；第 12 章主要介绍时间序列的定义与扩展，相关图绘制与白噪声检验，移动平均滤波与指数平滑法，ARIMA 模型、SARIMA 模型、ARIMAX 模型、单位根检验、向量自回归模型，协整与向量误差修正模型，ARCH 族模型；第 13 章主要介绍面板数据的基本操作、固定效应与随机效应模型、长面板模型、面板工具变量法、动态面板模型、面板数据的离散选择模型等内容；第 14 章主要介绍似不相关回归、多元回归模型、联立方程模型；第 15 章主要介绍随机数的生成、蒙特卡罗模拟、重复抽样及自助法；第 16 章是 Stata 编程基础，主要介绍基本概念与工具、程序文件的基本格式和程序控制语句；第 17 章是 Stata 综合案例分析，包括社会保障与经济增长关系实证研究，外部竞争环境不同的企业公司治理对绩效影响的敏感性分析，农民焚烧秸秆意愿的实证研究。

本书内容丰富，语言叙述清晰，结构简洁明了，通俗易懂，是一本较为实用的 Stata 实验教程。书中各章不仅通过详细的理论和案例讲解，较为全面地介绍了 Stata 的功能，而且还有与之相配套的 PPT 和操作录像供读者学习和使用，在每章之后还有一定数目的上机练习题供读者自行



练习和应用。因此，读者只需按照书中介绍的步骤一步步地实际操作，就能完全掌握本书的内容。

本书附带光盘中主要包括三部分内容，一部分为 data 文件下的数据，囊括了书中所有实验案例和课后练习题所用数据；一部分为 video 文件夹下的录像，囊括了书中所有实验案例和课后练习题的指导录像；一部分为 PPT 文件夹下的 PPT 文件，可以供老师教学使用，也可以供读者自行复习使用。

本书面向大中专院校经济管理专业及相关的社会科学类学生，特别是具有一定统计学和计量经济学基础知识的学生，以及企事业单位和其他相关领域的科研工作人员。由于本书编者尽量做到难易结合，所以本书既可以作为教材由老师负责讲解使用，也可以供读者自学使用。

本书主要由周广肃、梁荣、田金秀编写，参加本书编写工作的还有贾东永、凌佳、孙超逸、张楠、李龙、王华、李辉、刘峰、徐浩、李建国、马建军、唐爱华、苏小平、朱丽云、马淑娟、周毅、张浩、张乐、李大勇、许小荣、魏勇、王云等。本书在编写过程中吸收了大量的研究成果，在此对相关研究人员表示感谢。

编者力图使本书的知识性和实用性相得益彰，但由于水平有限，书中错误、纰漏之处在所难免，欢迎广大读者、同仁批评斧正。

编者
2011年6月

目 录

丛书序

前言

第 1 章 Stata 软件概述.....	1
1.1 Stata 软件简介.....	1
1.2 Stata 窗口及基本操作.....	2
1.2.1 Stata 窗口说明.....	2
1.2.2 Stata 帮助系统.....	5
1.2.3 Stata 语法和命令.....	6
1.3 Stata 主要功能模块.....	8
1.3.1 数据处理.....	8
1.3.2 绘图.....	8
1.3.3 统计分析.....	8
1.3.4 回归与建模分析.....	8
1.3.5 编程.....	8
复习与习题.....	9
本章回顾.....	9
习题.....	9
第 2 章 Stata 中的数据处理.....	10
2.1 数据的类型、压缩和转化.....	10
2.1.1 打开本地或网络数据文件.....	10
2.1.2 Stata 常用数据类型与压缩.....	12
2.1.3 数据类型的转化.....	14
2.2 数据的导入.....	17
2.2.1 创建新的数据库.....	17
2.2.2 使用已经保存的 Stata 数据.....	19
2.2.3 导入其他格式的数据.....	19
2.3 数据的整理.....	21
2.3.1 数据的标签与排序.....	21
2.3.2 数据的拆分.....	23



2.3.3 数据的合并	24
2.3.4 长宽数据的转换	25
复习与习题	27
本章回顾	27
习题	28
第 3 章 Stata 中的图形制作	30
3.1 图形制作的基本命令与相关操作	30
3.1.1 图形制作的基本命令	30
3.1.2 图形制作的菜单选项	31
3.1.3 与图形绘制相关的基本操作	32
3.2 直方图、散点图和曲线标绘图的绘制	34
3.2.1 直方图的绘制	34
3.2.2 散点图的绘制	36
3.2.3 曲线标绘图的绘制	39
3.3 条形图、饼图和箱线图的绘制	41
3.3.1 条形图的绘制	41
3.3.2 饼图的绘制	44
3.3.3 箱线图的绘制	45
3.4 图形的保存、合并及修改	46
3.4.1 图形的保存和已存图形的打开	47
3.4.2 图形的合并	47
3.4.3 图形的修改	48
复习与习题	49
本章回顾	49
习题	49
第 4 章 Stata 与方差分析	53
实验 4.1 单因素方差分析	53
实验 4.2 双因素或多因素方差分析	56
实验 4.3 协方差分析	59
复习与习题	62
本章回顾	62
习题	62
第 5 章 Stata 与假设检验	64
实验 5.1 单个总体的假设检验	64

实验 5.2 两个总体的假设检验	69
复习与习题	72
本章回顾	72
习题	73
第 6 章 基本回归分析	74
实验 6.1 小样本的普通最小二乘分析	74
实验 6.2 大样本的普通最小二乘分析	81
实验 6.3 约束回归	85
实验 6.4 非线性最小二乘	89
复习与习题	93
本章回顾	93
习题	93
第 7 章 Stata 与模型的设定	95
实验 7.1 遗漏变量的检验	95
实验 7.2 解释变量个数的选择	100
实验 7.3 多重共线性与逐步回归法	102
实验 7.4 极端数据的诊断与处理	109
实验 7.5 虚拟变量的处理	111
实验 7.6 经济结构变动的 Chow 检验	114
复习与习题	117
本章回顾	117
习题	117
第 8 章 Stata 与模型的修正	119
实验 8.1 异方差检验与处理	119
实验 8.2 自相关和可行广义最小二乘法	130
实验 8.3 内生性与 2SLS、GMM	145
复习与习题	154
本章回顾	154
习题	156
第 9 章 Stata 与离散被解释变量模型	157
实验 9.1 二值选择模型	157
实验 9.2 多值选择模型	168
实验 9.3 排序数据模型	173



实验 9.4 条件 logit 模型	177
实验 9.5 嵌套 logit 模型	182
复习与习题	189
本章回顾	189
习题	190
第 10 章 计数模型	192
实验 10.1 泊松回归模型	192
实验 10.2 负二项和广义负二项回归模型	198
实验 10.3 零膨胀回归模型	205
复习与习题	210
本章回顾	210
习题	210
第 11 章 受限被解释变量	212
实验 11.1 断尾回归模型	212
实验 11.2 截取回归模型	217
实验 11.3 样本选择模型	222
复习与习题	229
本章回顾	229
习题	229
第 12 章 基本时间序列分析	231
实验 12.1 时间序列的定义与扩展	231
实验 12.2 相关图绘制与白噪声检验	237
实验 12.3 移动平均滤波与指数平滑法	242
实验 12.4 ARIMA 模型	251
实验 12.5 SARIMA 模型	260
实验 12.6 ARIMAX 模型	266
实验 12.7 单位根检验	272
实验 12.8 向量自回归模型	277
实验 12.9 协整与向量误差修正模型	300
实验 12.10 ARCH 族模型	312
复习与习题	322
本章回顾	322
习题	324

第 13 章 面板数据分析	327
实验 13.1 面板数据的基本操作.....	327
实验 13.2 固定效应与随机效应模型.....	334
实验 13.3 长面板模型.....	344
实验 13.4 面板工具变量法.....	352
实验 13.5 动态面板模型.....	356
实验 13.6 面板数据的离散选择模型.....	363
实验 13.7 面板数据的计数模型.....	367
实验 13.8 随机效应 tobit 模型.....	371
复习与习题.....	374
本章回顾.....	374
习题.....	376
第 14 章 系统方程模型	378
实验 14.1 似不相关回归.....	378
实验 14.2 多元回归模型.....	383
实验 14.3 联立方程模型.....	386
复习与习题.....	392
本章回顾.....	392
习题.....	392
第 15 章 蒙特卡罗模拟和自助法	394
实验 15.1 随机数的生成.....	394
实验 15.2 蒙特卡罗模拟.....	397
实验 15.3 重复抽样.....	401
实验 15.4 自助法.....	409
复习与习题.....	415
本章回顾.....	415
习题.....	415
第 16 章 Stata 编程基础	417
实验 16.1 基本概念与工具.....	417
实验 16.2 程序文件的基本格式.....	425
实验 16.3 程序控制语句.....	440
复习与习题.....	448
本章回顾.....	448



习题	450
第 17 章 Stata 综合案例分析.....	451
17.1 综合案例一：社会保障与经济增长关系实证研究	451
17.1.1 问题背景和数据描述	451
17.1.2 统计方法与 Stata 实现	452
17.1.3 结论	455
17.2 综合案例二：外部竞争环境不同的企业公司治理对绩效影响的敏感性分析	456
17.2.1 问题背景和数据描述	456
17.2.2 统计方法与 Stata 实现	458
17.2.3 结论	467
17.3 综合案例三：农民焚烧秸秆意愿的实证研究.....	467
17.3.1 问题背景	467
17.3.2 相关假设	468
17.3.3 数据来源及描述	469
17.3.4 统计分析与 Stata 的实现	469
17.3.5 结论	473

第 1 章 Stata 软件概述

Stata 软件作为现阶段较为流行的统计软件，其应用已经深入到经济科学、社会科学、行为科学、生物统计等等各个学科领域。本章作为全书第一章，主要介绍 Stata 入门的相关知识：首先将介绍 Stata 软件的基本情况，让用户对将要学习的软件有一个较为宏观的认识；接下来，将介绍 Stata 软件的窗口和基本操作命令，以熟悉基本操作；最后简要介绍下 Stata 的主要功能模块，这里我们也将对本书的内容进行一个概览，对于本书的章节安排和 Stata 的具体功能读者可藉此有一个直观的了解。

1.1 Stata 软件简介

Stata 软件具有强大的数据处理和分析功能，是现今较为流行的统计计量分析软件，它由 Stata 公司在 1985 年研制开发成功之后面市，到现在已经有二十多年的历史了。在写作本书时其最新的版本为 2009 年刚推出的 Stata 11.0，但是鉴于 11.0 的版本还基本没有在中国推广和使用，所以本书所介绍的功能主要是通过 Stata 10.0 来实现的。

Stata 10.0 在安装时主要有四种版本，包括 Small（小型版）、IC（标准版）、SE（特殊版）和 MP（多处理器版）。用户可以在安装过程中自主进行选择，一般而言，SE 版已经能实现 Stata 的所有功能，MP 版与 SE 版相比，功能一致，但是运算速度更快。

Stata 与 SAS、SPSS 统称为三大权威统计软件，被广泛地应用于统计学、经济学、生物学、医药学、社会学、人口学等领域，功能十分强大。与其他软件相比，Stata 具有以下明显优势：

(1) Stata 操作较为简单，方便掌握。Stata 为窗口编辑和程序操作双操作方式的软件，其操作的实现既可以通过点击菜单选项完成，也可以通过输入命令完成。Stata 的命令语句简洁明快，逻辑清晰，灵活方便，用户可以结合自身实际情况熟练应用。

(2) Stata 是一个开放的软件系统。Stata 与其他统计软件相比，最大的优势就是可以不断地接收统计学和计量经济学的最新研究成果，保证与时俱进。众所周知，现在是一个知识爆炸的时代，计量和统计学发展迅速，新成果层出不穷。Stata 软件可以通过多种途径及时实现更新，保证用户可以将最新的理论成果应用于自身的实践。这种开放性主要体现在三个方面：

①用户可以到 Stata 官方网站 (<http://www.stata.com>) 上下载相关模块更新，或直接在命



令窗口中输入“update”，在联网的情况下也可以实现更新。

②使用 findit 命令，找到所需要的功能模块，并实现下载和安装。

③下载由其他用户编写的 Stata 模块，主要通过菜单选项 Help -> SJ and User-written Programs 进行下载更新。当然，用户也可以自己编写程序，从而实现所需要的功能。

(3) Stata 具有强大的数据分析功能。Stata 软件基本可以实现所有的统计与计量分析的功能。例如，在统计方面，可以实现数据的描述分析、方差分析、假设检验、主成分分析、聚类分析等；在计量分析方面，可以实现多种计量模型的应用，如单方程模型回归、联立方程模型回归、离散被解释变量模型、受限因变量模型、时间序列模型、面板数据模型、分位数回归模型等。Stata 软件这种统计与计量分析功能的有机结合，使得用户可以完成复杂性、综合性较强的各种分析和研究。

(4) Stata 具有强大的图形制作功能。图形分析现在仍然是各种分析必不可少的部分，具有强大的图形制作功能也是 Stata 被广泛接受的原因之一。利用 Stata 软件可以完成散点图、直方图、折线图、条形图、函数图等各种图形的制作，此外用户可以根据自身需要对图形进行修改，图形也可直接被图形处理软件和文字处理软件直接调用。这些功能使得 Stata 的画图功能方便有效，便于使用。

1.2 Stata 窗口及基本操作

本节介绍 Stata 的窗口组成和基本操作，读者在这里将逐渐熟悉 Stata 的操作界面与方法。

1.2.1 Stata 窗口说明

Stata 软件安装完成后，运行 Stata，将会看到如图 1.1 所示的操作界面。

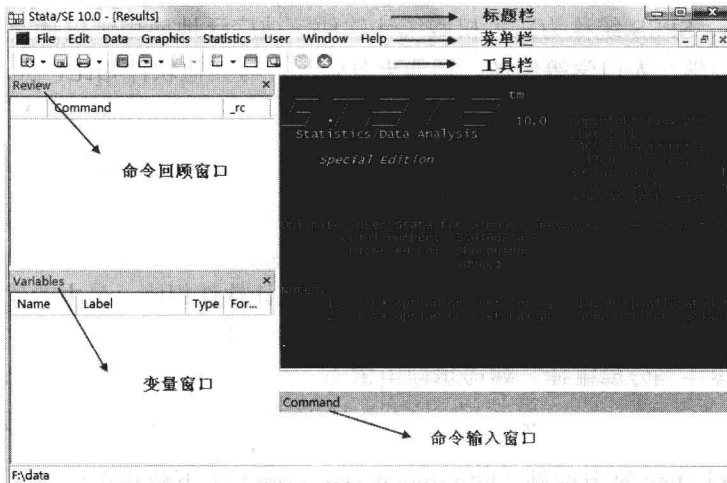


图 1.1 Stata 操作界面