

422分钟视频教学+34个统计分析类型+94个专业统计案例
+35个专业统计分析练习

Stata

统计分析与应用

博 刘 荣 丁维岱 段美霞 等编著

(修订版)

分钟多媒体视频教学，读者可以通过观看光盘学习Stata软件

- 34个统计分析类型，基本涵盖Stata统计分析的各个领域
- 94个专业统计案例，既涉及不同统计分析类型，又涉及不同的行业
- 35个专业统计分析练习，给读者提供足够的上手练习机会



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

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

Stata 统计分析与应用

(修订版)

胡 博 刘 荣 丁维岱 段美霞 等编著

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

Stata 是近年来兴起的一款短小精悍、功能强大的统计计量软件, 由于它操作灵活、简单、易学易用, 现在已越来越受到人们的重视和欢迎, 并且和 SAS、SPSS 一起, 被称为新的三大权威统计软件。

本书对 Stata 的介绍是: Stata 使用基础、Stata 与回归分析、Stata 与多元统计分析、Stata 编程入门, 共计 14 章。这些内容中包括了 Stata 的数据管理、图形绘制、各种基本和高级的计量回归分析、常用的统计分析方法等, 最后还简明地介绍了编程入门知识。在讲解过程中, 本书使用 Stata12 这个最新版本, 穿插了几百个实例加以诠释, 并配有多媒体视频光盘进行讲解, 非常方便读者的理解和学习。

本书的内容和实例满足金融、经济、生物医疗、卫生保健、社会人文、心理学等多学科的需要, 可供高等院校相关专业本科生、研究生以及从事统计分析的研究者参考使用, 也可作为 Stata 软件培训和自学的教材。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有, 侵权必究。

图书在版编目 (CIP) 数据

Stata 统计分析与应用 / 胡博等编著. —修订本. —北京: 电子工业出版社, 2013.4
ISBN 978-7-121-20095-3

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

中国版本图书馆 CIP 数据核字 (2013) 第 064828 号

策划编辑: 祁玉芹

责任编辑: 鄂卫华

印 刷: 三河市鑫金马印装有限公司

装 订: 三河市鑫金马印装有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×1092 1/16 印张: 25 字数: 640 千字

印 次: 2013 年 4 月第 1 次印刷

定 价: 59.80 元 (含光盘 1 张)

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

前言

PREFACE

Stata 是近年来兴起的一款短小精悍、功能强大的统计计量软件，相比于功能强大但是显得臃肿的 SAS，它操作灵活、简单、易学易用，是一个非常有特色的统计分析软件。由于 Stata 本身的统计方法非常先进，现在已越来越受到人们的重视和欢迎，并且和 SAS、SPSS 一起，被称为新的三大权威统计软件。其普遍使用的 10.1、11.0、12.0 版安装文件只有 100MB 左右，但已经包含了全部的统计分析、数据管理和绘图等功能，尤其是它的统计分析功能极为全面，比起 1GB 以上大小的 SAS 系统也毫不逊色。

本书的内容及知识体系

本书基于 Stata 而编写，分为 Stata 使用基础、Stata 与回归分析、Stata 与多元统计分析、Stata 编程入门共 14 章。内容涵盖了 Stata 从入门的数据管理、图形绘制到各种高级的统计、回归分析以及编程入门知识，不仅详细地讲解了 Stata 各种命令的相关参数，还列举了大量的实例加以说明，非常方便读者的理解和学习。

第 1 章首先带领读者了解 Stata 的发展里程，了解 Stata 的基本功能，熟悉 Stata 的界面和各种相关的学习资料。

第 2 章切入正题，介绍 Stata 的数据管理功能，包括如何创建和导入数据集、如何对变量进行相关的操作、如何对两个数据集进行合并等，这一部分是 Stata 使用的基础。

第 3 章和第 4 章详细讲解了 Stata 的绘图功能。第 3 章以散点图为例，介绍了 Stata 绘图的各种选项的使用，第 4 章进一步介绍了曲线标绘图、连线标绘图、拟合图、条形图等，读者一旦掌握了这些内容，就可以对 Stata 绘图运用自如了。

第 5 章到第 7 章介绍了描述性统计分析、列联分析、方差分析三种基本的统计分析手段在 Stata 中的实现。

第 5 章首先回顾了统计学的基本知识，然后向读者展示了如何使用 Stata 进行描述性统计分析、探测异常值、计算相关系数以及进行数据的正态性转换。

第 6 章是列联分析的内容，包括如何利用 Stata 进行列联分析、如何生成包含描述性统计量的表格等。

第 7 章介绍方差分析的内容，包括 t 检验、单因素方差分析、多因素方差分析、协方差分析等。

第 8 章到第 10 章介绍了 Stata 的回归分析功能，这也是 Stata 最擅长的工作。

第 8 章介绍了经典假设下的横截面数据单方程线性回归模型的 Stata 实现，包括单方程回归分析的基础知识、使用 Stata 进行回归的各种命令、如何进行变量变换来更好地拟合方程。

第 9 章是非经典假设、线性方程组、面板数据估计的 Stata 实现，包括存在内生性、异方差、多重共线性时如何使用 Stata 进行相关的处理，还包括对似不相关方程组、联立方程组以及面板数据进行回归分析的命令。

第 10 章是非线性回归分析及回归分析专题，这是计量分析中较高级的内容，包括了二值响应模型、多值响应模型、角点解和样本选择模型，最具特色的是我们专门介绍了如何在 Stata



中进行回归诊断。

第 11 章介绍 Stata 的时间序列分析功能,包括基本时间序列模型的估计、ARIMA 模型的估计、单位根与协整、VAR 与 VEC 的估计、ARCH 与 GARCH 的估计,囊括了时间序列分析中各种常见的内容。

第 12 章和第 13 章是多元统计分析中的内容,这两章在回顾了相关的基础知识后,详细介绍了聚类分析、主成分分析、因子分析的 Stata 命令。

最后,第 14 章简要介绍了 Stata 编程中的基础知识,内容包括全局宏与局部宏、标量与矩阵、循环语句,以及 Stata 计算结果的提取。

本书有何特色

(1) 配有多媒体语音视频进行讲解

为了方便读者更加直观地学习 Stata 软件,作者为本书专门录制了大量的多媒体语音视频进行讲解,视频适合使用 Stata 10.1、11.0 或 12.0 版本的用户,相信通过这些视频读者可以更快地掌握本书的内容。

(2) 内容全面,讲解详细,即用即查

本书是一本百科全书式的 Stata 著作,详细讲解了 Stata 的大部分常用功能,对参数的介绍非常细致,本书既可作为很好的入门教程,也可以作为即用即查的参考手册。

(3) 进行了详尽的理论知识讲解

本书对每一个知识点都进行了详尽的回顾,在诸如回归分析、时间序列、方差分析等内容上达到了高级计量的程度,在这个基础上介绍 Stata 命令更加顺理成章,读者使用起来也更加游刃有余。

(4) 列举了几百个典型实例

本书对每一个命令都配备了一个以上的实例,对于常用命令则在全书中反复使用,示例中的命令大都进行了详细的解释,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。

本书的对象

本书的内容和实例满足金融、经济、生物、医疗卫生保健、社会人文、心理学等多学科的需要,可供高等院校相关专业本科生、研究生以及从事统计分析的研究者参考使用,也可作为 Stata 软件培训和自学的教材。

本书作者

本书是集体智慧的结晶,由胡博、刘荣、丁维岱、段美霞主持编写,参加本书编写工作的还有李龙、王华、李辉、刘峰、徐浩、李建国、唐爱华、陈胜可、贾东升、魏勇、陈杰等,在此,编者对以上人员致以诚挚的谢意!本书在编写过程中吸收了前人的研究成果,在此一并表示感谢。

由于作者水平有限,书中的缺点甚至错误在所难免,恳请广大读者批评指正。

作 者

2013 年 2 月



目录

CONTENTS

第 1 章 Stata 概述	1
1.1 Stata 的历史和特点	1
1.2 Stata 的使用界面	2
1.3 Stata 命令——help、search 命令	3
1.4 Stata 学习资源	8
第 2 章 数据管理	9
2.1 变量和变量的取值	9
2.1.1 变量的命名	9
2.1.2 变量的取值类型	10
2.1.3 变量的显示	11
2.1.4 变量的标签	13
2.2 创建一个新的数据集	14
2.2.1 关于数据集操作的基本命令	14
2.2.2 举例应用：创建新的数据集 auto.dta	16
2.3 导入已创建的数据集	19
2.3.1 一般的原则	19
2.3.2 读取格式为 .dta 的数据	20
2.3.3 利用 Excel 复制数据进入 Stata 系统中	20
2.4 Stata 中的表达式	21
2.4.1 算术符号	21
2.4.2 关系符号	22
2.4.3 逻辑符号	22
2.5 Stata 中的常用函数	22
2.6 使用 in、if 和 by 语句定义数据子集	24
2.6.1 in 的使用	24
2.6.2 if 的使用	24
2.6.3 by 语句的使用	24
2.7 变量的相关操作	25
2.7.1 建立新的变量——generate	25
2.7.2 更改已有的变量——replace	27
2.7.3 egen 命令	28
2.8 数值和字符串的转换	32



2.8.1	encode 和 decode 命令	32
2.8.2	real 函数	34
2.9	生成分类变量和虚拟变量	34
2.9.1	生成虚拟变量	35
2.9.2	生成分类变量	36
2.10	数据的整理	41
2.10.1	数据的横向合并	41
2.10.2	数据的纵向合并	44
2.10.3	数据的交叉合并	47
2.10.4	数据的抽取	50
2.11	Stata 操作习题	53
第 3 章	图形绘制基础	55
3.1	Stata 绘图简介	55
3.1.1	主要的图形类型	56
3.1.2	图形的组成部分与制图命令的结构	56
3.1.3	寻求帮助	56
3.2	绘制散点图	56
3.2.1	绘制散点图的命令和最基本的使用	56
3.2.2	散点显示选项 (marker_options) 的设定	61
3.2.3	散点标签选项 (marker_label_options) 的设定	67
3.2.4	连线选项 (connect_options) 的设定	73
3.2.5	振荡选项 (jitter_options) 的设定	74
3.3	二维绘图选项	75
3.3.1	坐标轴尺度选项组 (axis_scale_options) 的设定	76
3.3.2	坐标轴刻度选项组 (axis_label_options) 的设定	80
3.3.3	坐标轴标题选项组 (axis_title_options) 的设定	87
3.3.4	标题选项组 (title_options) 的设定	89
3.3.5	图例选项 (legend_option) 的设定	91
3.3.6	by 选项的设定	96
3.3.7	scheme 选项的设定	102
3.3.8	轴线选择选项 (axis_choice_options) 的设定	102
3.3.9	增加线选项 (added_line_options) 的设定	104
3.3.10	scale 选项的设定	105
3.3.11	图形保存选项	107
3.3.12	图形输出选项	108
3.4	Stata 操作习题	109
第 4 章	其他图形绘制	111
4.1	绘制曲线标绘图和连线标绘图	111

4.1.1 绘制曲线标绘图	111
4.1.2 绘制连线标绘图	115
4.2 绘制拟合图形	116
4.2.1 绘制一次拟合图形	116
4.2.2 绘制二次拟合图形	117
4.2.3 绘制 lowess 拟合图形	119
4.3 绘制条形图	121
4.3.1 关于分类变量的讲解	123
4.3.2 关于条形图外观的讲解	124
4.4 Stata 操作习题	129
第 5 章 描述性统计分析	131
5.1 描述性统计的原理	131
5.1.1 定性变量	131
5.1.2 定量变量	131
5.2 描述性统计量的 Stata 实现	134
5.3 探测异常值	141
5.3.1 计算 z 得分	142
5.3.2 箱线图	142
5.4 数据的正态性检验和数据转换	146
5.4.1 正态性检验的原理	146
5.4.2 正态性检验的 Stata 实现	151
5.4.3 改变数据的分布	153
5.5 相关系数	155
5.5.1 相关系数概述	156
5.5.2 相关系数在 Stata 中的实现	157
5.6 Stata 操作习题	161
第 6 章 列联表分析	163
6.1 列联表分析	163
6.1.1 列联表概述	163
6.1.2 独立性检验统计量	163
6.1.3 列联表中的相关测量统计量	164
6.2 Stata 的列联表分析——table 和 tabulate 命令	165
6.2.1 使用 table 命令生成列联表	165
6.2.2 使用 tabulate 命令进行列联表分析	169
6.3 利用 Stata 生成包含描述性统计量的列表	174
6.3.1 tabstat	174
6.3.2 tabulate, summarize ()	176
6.4 Stata 操作习题	178



第7章 方差分析	179
7.1 t 检验	179
7.1.1 单样本 t 检验的基本思想与理论	179
7.1.2 双样本 t 检验的基本思想与理论	179
7.1.3 t 检验的 Stata 基本命令	181
7.2 单因素方差分析	186
7.2.1 单因素方差分析原理	186
7.2.2 单因素方差分析 Stata 实现	189
7.3 双因素和多因素方差分析	192
7.3.1 双因素方差分析原理	192
7.3.2 多因素方差分析原理	196
7.3.3 双因素和多因素方差分析 Stata 实现	196
7.4 协方差分析	198
7.4.1 协方差分析原理	198
7.4.2 协方差分析 Stata 实现	199
7.5 Stata 操作习题	201
第8章 经典假设下的横截面数据单方程线性回归模型的 Stata 实现	203
8.1 线性回归分析	203
8.1.1 回归分析简介	203
8.1.2 线性回归分析简介	204
8.2 横截面数据	205
8.3 经典假设及其性质	207
8.3.1 经典假设	207
8.3.2 经典假设下线性模型的基本性质	208
8.4 Stata 的回归分析——regress、predict、test 命令	209
8.4.1 使用 regress 命令——因变量对自变量的回归	209
8.4.2 使用 predict 命令——计算拟合值和残差	216
8.4.3 使用 test 命令——进行读者指定的检验	219
8.5 sw regress 基本命令及其选项——逐步回归	221
8.6 对解释变量和被解释变量做变换——更好地拟合数据	223
8.7 习题	226
第9章 非经典假设、线性方程组、面板数据估计的 Stata 实现	227
9.1 非经典假设下的回归分析的 Stata 实现	227
9.1.1 多重共线性的检验和处理	227
9.1.2 内生性的检验与处理	231
9.1.3 异方差的检验与处理	236
9.2 线性方程组的回归分析——Stata 实现	238

9.2.1 似不相关模型	238
9.2.2 联立方程组模型	241
9.3 面板数据的 Stata 处理	243
9.3.1 固定效应的面板数据 Stata 实现	245
9.3.2 随机效应的面板数据 Stata 实现	246
9.4 练习题	247
第 10 章 非线性回归分析及回归诊断基础	249
10.1 非线性回归分析	249
10.1.1 非线性回归的 Stata 实现——nl 命令	249
10.2 二值响应模型——使用 probit、logit、dprobit、logistic 命令	251
10.2.1 probit、dprobit 命令的使用方法	251
10.2.2 logit、logistic 命令的使用方法	254
10.3 多值响应模型——使用 mlogit、ologit 命令	256
10.3.1 无序响应模型——mlogit 命令	256
10.3.2 有序响应模型——ologit 命令	259
10.4 角点解模型——Tobit 命令的使用方法	261
10.5 样本选择模型——heckman 命令的使用方法	263
10.6 回归诊断	265
10.7 练习题	269
第 11 章 时间序列分析	271
11.1 基本时间序列模型的估计	271
11.1.1 趋势分析与指数平衡	271
11.1.2 平稳性检验	273
11.1.3 趋势分析与指数平滑的 Stata 实现	274
11.2 ARIMA 模型的估计、单位根与协整	278
11.2.1 ARIMA 模型的估计	278
11.2.2 单位根过程及其检验	281
11.2.3 协整检验	283
11.2.4 ARIMA 模型的 Stata 实现	284
11.3 VAR 与 VEC 的估计及解释	293
11.3.1 普通 VAR 模型的估计	293
11.3.2 Granger 因果分析、IRF 与方差分解	295
11.3.3 Johansen 协整检验和 VEC 模型的估计	298
11.3.4 VAR 模型的 Stata 实现	300
11.4 ARCH 与 GARCH 的估计及解释	307
11.4.1 ARCH 模型	307
11.4.2 GARCH 模型	308
11.4.3 ARCH 模型的 Stata 实现	309



11.5 Stata 操作习题.....	312
第 12 章 聚类分析	315
12.1 聚类分析的基本思想与理论.....	315
12.1.1 聚类分析的基本思想	315
12.1.2 聚类分析的相似性测度	316
12.1.3 聚类分析的典型方法	318
12.1.4 聚类分析的步骤	321
12.2 聚类分析的基本命令	322
12.3 Stata 操作习题.....	333
第 13 章 主成分分析和因子分析.....	335
13.1 主成分分析.....	335
13.1.1 主成分分析的基本思想与理论	335
13.1.2 主成分分析基本命令	338
13.1.3 Stata 操作案例	348
13.2 因子分析.....	351
13.2.1 因子分析的基本思想与理论	352
13.2.2 因子分析基本命令	357
13.2.3 Stata 操作案例	360
13.3 Stata 操作习题.....	363
第 14 章 Stata 编程基础	365
14.1 do 文件和 Log 文件	365
14.1.1 do 文件的编写	365
14.1.2 运行 do 文件	366
14.1.3 log 文件.....	366
14.2 局部宏与全局宏.....	367
14.2.1 局部宏	368
14.2.2 全局宏	372
14.2.3 一些扩展函数以及列表函数	372
14.3 标量简介.....	374
14.4 循环结构.....	376
14.4.1 forvalues 语句	376
14.4.2 foreach 语句	378
14.5 矩阵简介.....	380
14.6 使用 Stata 命令的结果.....	381
14.6.1 r 类命令.....	381
14.6.2 e 类命令	385
14.7 Stata 操作习题.....	389



第 1 章 Stata 概述

在本书第 1 章自然要介绍 Stata 软件的入门知识。但作者希望不限于此，而是在介绍 Stata 基础知识的同时，能使读者感受到 Stata 的软件文化。Stata 一个最大的特点就是用户可以简单地编写自己的 Stata 应用程序，然后上传到相关论坛与其他用户交流。如果用户的自编程序拥有良好的应用前景，Stata 公司甚至会将你的程序加入到 Stata 标准版本中，让世界各地的用户分享你的成果。我们将在本章中介绍 Statalist，它提供了一个极有价值的供 Stata 活跃用户联系的在线论坛。

本章将从以下四个部分简单地介绍 Stata。

在第一部分介绍 Stata 的历史和特点，这有助于读者在后续的学习中更好地理解 Stata 命令。第二部分介绍 Stata 的使用界面，使读者对 Stata 的外观有一个感性的认识。第三部分介绍 Stata 中最基本的 help 和 search 命令。第四部分将介绍 Stata 的各种学习资源，Statalist 的申请与退订，便于读者在阅读完本书后可以进一步深入研究 Stata 的使用方法。

1.1 Stata 的历史和特点

Stata 最初由美国计算机资源中心 (Computer Resource Center) 研制，现在为 Stata 公司的产品，其最新版本为 12.0 版。相比于功能强大但略显臃肿的 SAS，它操作灵活、简单、易于易用，是一个非常轻便的统计分析软件。由于在 Stata 软件中能便捷地实现多种先进统计方法，它越来越受到用户的推崇，并和 SAS、SPSS 一起被称为新的三大权威统计软件。自 1985 推出至今，Stata 不断更新、日趋完善。

Stata 最突出的特点是短小精悍、功能强大。最新的 12.0 版安装文件只有 100 MB 左右，但包含了全部的统计分析、数据管理和绘图等功能。此外，由于 Stata 在分析时将数据全部读入内存，在计算全部完成后才和磁盘交换数据，因此运算速度极快。

Stata 不仅操作方式简捷，它的数据格式简单，分析结果输出简洁明快，易于阅读，这使得 Stata 成为极其适用于统计教学的软件。

Stata 的许多高级统计模块是编程人员用宏语言写成的程序文件 (ADO 文件)，这些文件可以自行修改、添加和下载。用户可随时到 Stata 网站搜索并下载最新的升级文件。Stata 的这一特点使得它始终处于统计分析方法发展的最前沿，用户总能很快找到最新统计算法的 Stata 程序版本，而这也使得 Stata 成了几大统计软件中升级最多、最频繁的一个。

Stata 同时具有数据管理软件、统计分析软件、绘图软件、矩阵计算软件和程序语言的特点，而且可用于制作精美的图形，图形可直接被图形处理软件或字处理软件 (如 Word 等) 调用。

对 Stata 的特点进行了大量文字描述后，接下来介绍 Stata 的使用界面。



1.2 Stata 的使用界面

首先以 Stata/SE 12.0 为例介绍 Stata 的使用界面。双击 Stata 图标将显示如图 1-1 所示的 Stata 启动界面。除了 Windows 版本的软件惯有的菜单栏、工具栏、状态栏等外, Stata 的界面主要由以下 4 个窗口构成。

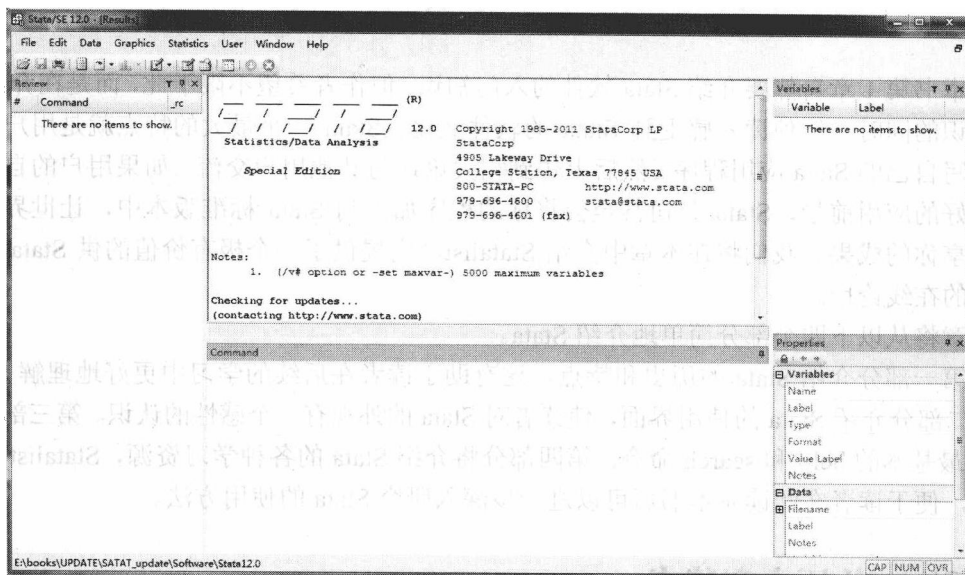


图 1-1 Stata 的启动界面

(1) 结果窗口: 位于界面上部 (见图 1-2), 软件运行中的所有信息, 如所执行的命令、执行结果和出错信息等均列于此。窗口中不同颜色的文本有不同的含义, 如黑色表示命令、蓝色代表链接、红色表示错误信息。

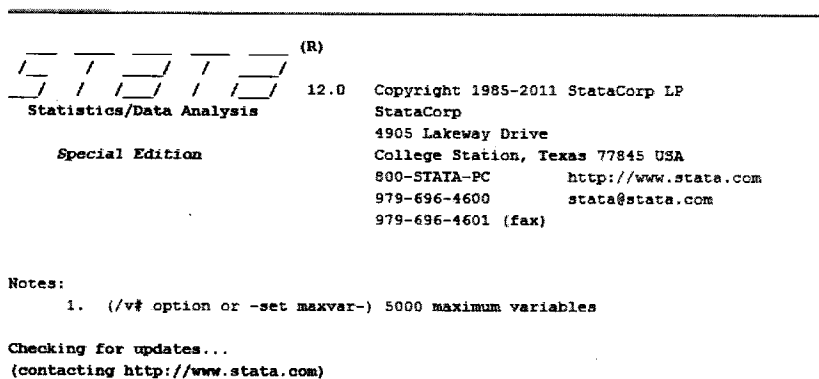


图 1-2 Stata 图形界面的结果窗口

(2) 命令窗口: 位于结果窗口下方 (见图 1-3), 相当于 DOS 软件中的命令行, 此处用于键入需要执行的命令, 按 Enter 键后即开始执行, 相应的结果会呈现在结果窗口中。

图 1-3 Stata 图形界面的命令窗口

(3) 命令回顾窗口：如图 1-4 所示，即 **Review** 窗口，位于界面左上方，所有执行过的命令会依次在该窗口中列出，单击后命令即被自动复制到命令窗口中；如果需要重复执行，双击相应的命令即可。

(4) 变量名窗口：如图 1-5 所示，位于界面右上方，列出当前数据中的所有变量名称。

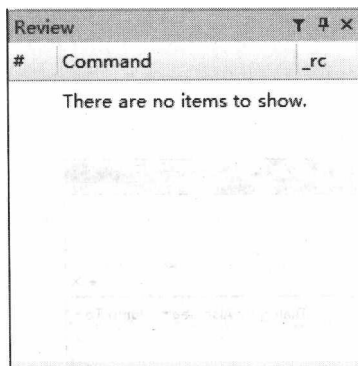


图 1-4 Stata 图形界面的命令回顾窗口

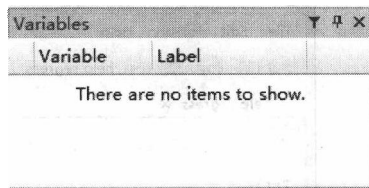


图 1-5 Stata 图形界面的变量窗口

除以上 4 个默认打开的窗口外，在 Stata 中还有数据编辑窗口、程序文件编辑窗口、帮助窗口、绘图窗口和 Log 窗口等，如果需要使用，可以用 **Window** 或 **Help** 菜单将其打开。

除了各种窗口以外，Stata 的菜单栏也是很重要的。我们在后面的章节中对 Stata 各种常用命令的介绍，都是通过示范如何使用命令语句实现的。但是，这些命令中有很多也可以使用菜单栏而不需要输入相关命令来实现。

Stata 菜单栏主由有 **File**、**Edit**、**Data**、**Graphics**、**Statistics**、**User**、**Window**、**Help** 组成。其中 **File**、**Edit** 与多数 Windows 软件使用方法一致。**Data** 主要用来导入导出数据，可以完成数据管理一章的部分任务。**Graphics** 可以用来实现 Stata 强大的作图功能，它可以实现 Stata 有关图形绘制章节的部分功能。**Statistics** 则与本书的介绍回归部分的相关章节有关，可以实现线性、非线性回归的绝大部分功能。**User** 用来管理用户的数据、图片、回归结果。**Window** 可以帮助我们管理以上所介绍的四个窗口和编辑、查看窗口的设置。**Help** 用来为用户提供帮助信息。

1.3 Stata 命令——help、search 命令

Stata 中最重要的两个命令莫过于 **help** 和 **search** 了。即使经常使用 Stata 的人也很难，也没必要记住常用命令的每一个细节，更不用说那些不常用的了。所以，在遇到困难又没有免费专家咨询时，使用 Stata 自带的帮助文件就是最佳选择。Stata 的帮助文件十分详尽，面面俱到，这既是好处也是麻烦。当你看到长长的帮助文件时，是不是对迅速找到相关信息感到没有信心？



`help` 和 `search` 都是查找帮助文件的命令，它们之间的区别在于 `help` 用于查找精确的命令名，而 `search` 是模糊查找。如果你知道某个命令的名字，并且想知道它的具体使用方法，只须在 Stata 的命令行窗口中输入 `help` 空格加上这个名字。按 Enter 键后屏幕上就会显示出这个命令的帮助文件的全部内容。如果你想在 Stata 下做某个估计或某种计算，而不知道具体该如何实现，就需要用 `search` 命令了。使用的方法和 `help` 类似，只需把准确的命令名改成某个关键词。按 Enter 键后窗口会给出所有和这个关键词相关的帮助文件名和链接列表。在列表中寻找最相关的内容，单击后在弹出的查看窗口中会给出相关的帮助文件。耐心寻找，反复实验，通常可以较快地找到用户需要的内容。

举个例子来说，如果我们要寻找关于基本回归命令 `regress` 命令（将在第 7 章中具体介绍）的使用方法，我们只需要在前面介绍过的命令窗口输入：

```
help regress
```

Stata 会显示如图 1-6 所示的窗口。

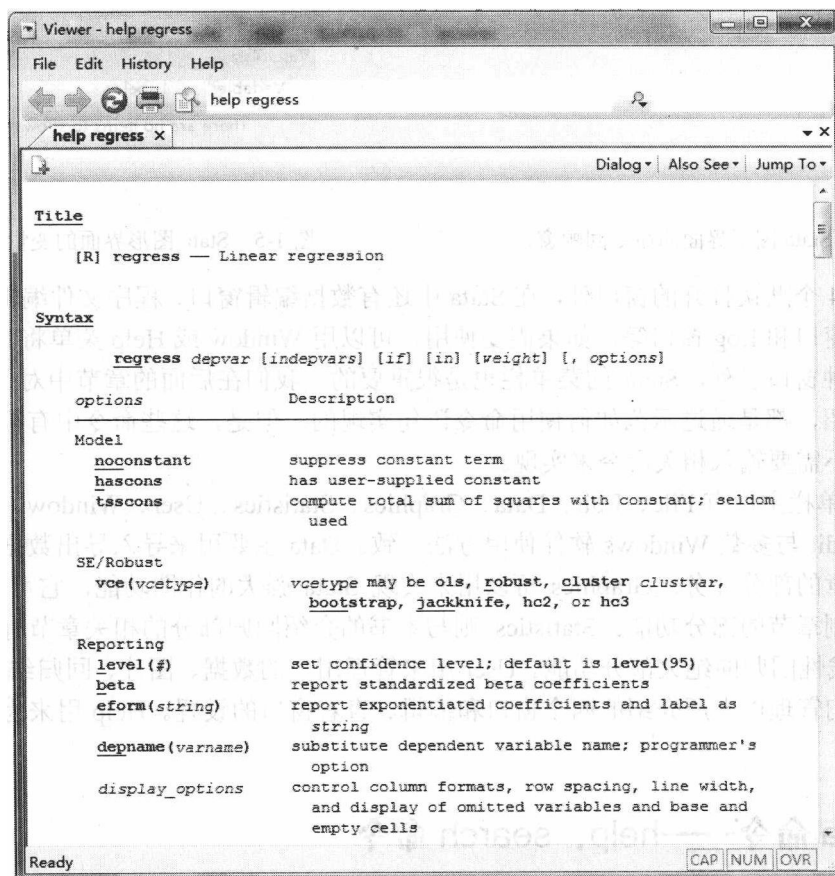


图 1-6 help 命令的结果

这个窗口详尽地展示了 `regress` 命令的信息。从命令名 `regress` 到语法 (syntax)，到对语法的解释 (description)，其后还有例子说明和参考内容（限于篇幅，我们只列出了 `help` 结果的一部分）。可以说，Stata 的帮助文件各种统计软件中最完备、思路最清晰的帮助文件了。

当使用 `help` 命令时, Stata 默认读者是进行的是精确查询, 如果我们对某个命令不太确定, 就可以使用 `search` 命令来进行模糊查询。

举个例子, 我们同样要查询关于 `regress` 命令的信息, 但是我们记不清 `regress` 命令的全名, 只记得 `regress` 的前半部分 `reg` 的话, 我们可以在 Stata 命令窗口中输入:

```
search reg
```

Stata 会输出如下结果:

```
. search reg
```

Keyword search

Keywords: reg

Search: (1) Official help files, FAQs, Examples, SJs, and STBs

Search of official help files, FAQs, Examples, SJs, and STBs

```
[U] Chapter 20 . . . . . Estimation and postestimation commands
(help estcom, postest, weights)

[U] Chapter 25 . . . . . Working with categorical data and factor variables
(help generate, fvvarlist)

[U] Chapter 26 . . . . . Overview of Stata estimation commands
(help estcom)

[R] regress . . . . . Linear regression
(help regress)

[R] regress postestimation . . . . . Postestimation tools for regress
(help regress postestimation)

[R] regress postestimation time series Postest. regress with time series
(help regress postestimations)
```

限于篇幅, 仅列出一部分。

我们可以比较一下 `help` 和 `search` 命令的区别。相比于 `help` 命令的精确, `search` 命令输出所有与 `reg` 关键词相关的帮助结果。这一点是 `search` 命令的缺点, 它往往会输出太多内容, 使用户的搜索相当困难。

除了输入 `help`、`search` 命令来查看帮助文件外, 我们还可以使用工具栏中的 `help` 标签。对于初学者, 使用帮助文件的最简单方法是选择 `help| contents` 命令, 则将显示如图 1-7 所示的窗口。

Content 窗口中按各种功能将命令的帮助文件分类, 读者可以逐级寻找需要的帮助文件。以上窗口中的蓝色部分都是可以继续单击的, 我们通过单击自己感兴趣的部分向下逐级寻找具体内容。所有你希望寻找的答案, 都可以通过这个页面进入。比如, 最简单的描述统计 `sum`, 就可以通过单击: `Top > Statistics > Summary statistics & tests > Summary statistics` 找到。这一过程用以下方式总结:

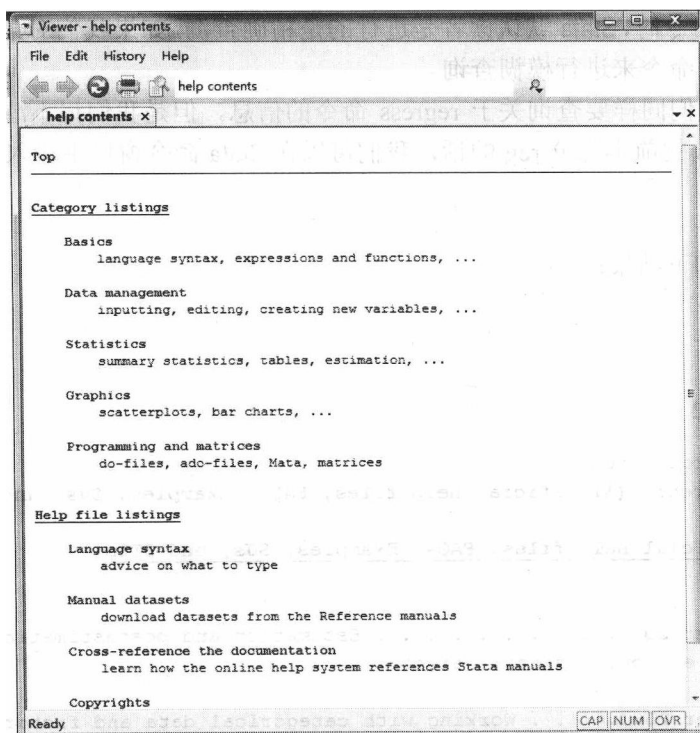


图 1-7 help 菜单栏下 contents 的内容

在我们单击 statistics 之后，得到如图 1-8 所示的结果。

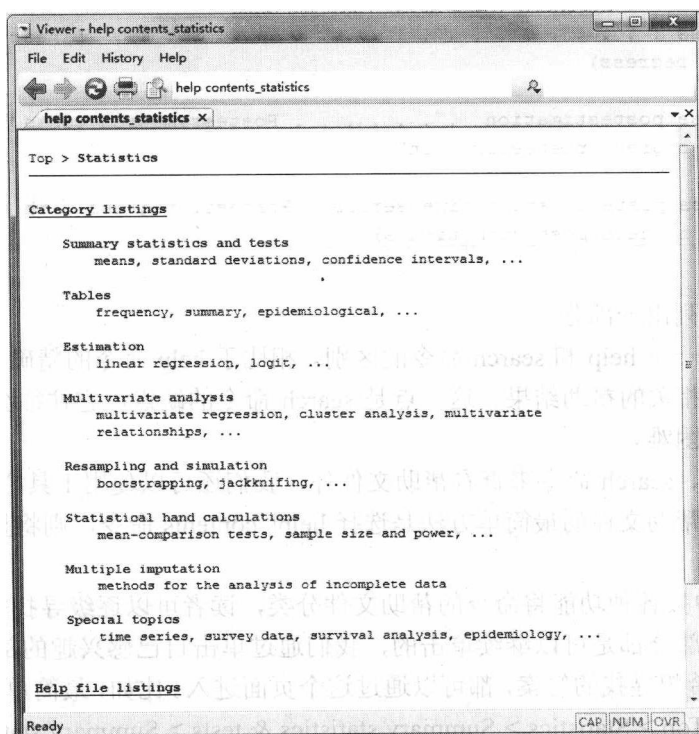


图 1-8 Help\Contents\Statistics 的内容