

STATA 初级视频教程（2010 版）

使用说明

连玉君

（中山大学 岭南学院 金融系）

arlionn@163.com

目 录

1 课程简介	1
2 课程特色	2
3 课程配套资料	2
4 配套资料的使用方法	2
5 讨论和建议	4
6 讲师介绍	4
7 报名咨询	4
8 培训优惠	4
附录A：STATA初级视频目录（时间节点）	5
第一讲 STATA简介	5
第二讲 数据处理	9
第三讲 Stata绘图	13
第四讲 矩阵操作	16
第五讲 STATA 编程初步	18
附录B：STATA高级视频教程简介	20

1 课程简介

工欲善其事，必先利其器。无论是经济学、管理学还是社会科学的其他学科，定量分析都变得越来越重要了。作为一个较为年轻的计量软件，STATA 自 1985 年问世以来，以其在数据处理、绘图、回归分析等方面的出色表现，赢得了越来越多的青睐。然而，相比于 SPSS、Eviews 等以菜单操作为主的软件，以命令操作见长的 STATA 软件门槛相对较高。由于进入国内的时间较短，相关的参考资料甚为有限，而 STATA 公司提供的近 10000 页的全英文使用手册更是令多数初学者望而生畏。这也成为阻碍多数国内同仁学习这款功能强大的计量软件的主要障碍。

鉴于上述情况，我们分别于 2007 年 11 月和 2008 年 10 月推出了“STATA 初级视频教程”和“STATA 高级视频教程”，内容涉及 STATA 的基本操作、数据处理、绘图、编程、常用计量模型的估计，以及 Bootstrap 和 Monte Carlo 模拟等内容。视频教学的直观性，加之课程的实用性导向，使这两套教程获得了广泛的好评。承蒙广大 STATA 视频教程学员的积极参与和反馈，我在过去两年多的时间里收集到了 100 余条修改建议，历经半年多的制作，最终得以为大家呈现这套新版 STATA 初级视频——“STATA 初级视频教程（2010 版）”。

STATA 初级视频教程（2010 版）共 5 讲，包含 36 个视频文件，总计 40 余个学时。内容涉及：STATA 入门、数据处理、绘图、矩阵以及编程。内容简介如下（详细目录见附录 A）：

第 1 讲从整体上介绍了 STATA 的基本架构，以便使学员在最短的时间内掌握 STATA 的精髓，包括：数据的导入导出、执行命令、修改和查验资料、log 文件和 do 文件的使用、STATA 与 Word（Excel、LaTeX）等软件的完美结合，以及 STATA 常用设定等内容。

第 2 讲通过大量的实例，介绍了各种数据处理技巧，是本课程最为核心、最有特色的内容。主要包括：复杂变量的创建；分位数；重复样本值、缺漏值和离群值的处理；资料的合并、追加和重新组合；文字变量、类别变量的处理；时间序列和面板资料的处理；以及数据的查验和对比等内容。这些内容的学习将大幅提高学员的数据处理能力。

第 3 讲介绍 STATA 绘图。为了达到举一反三的效果，我先从整体上介绍 STATA 绘图的基本知识，进而将绘图命令拆解成 8 类选项和 5 类元素，并最终通过 40 余个实例全面介绍了 15 类常用图形的绘制方法。

第 4 讲介绍 STATA 矩阵操作，包括矩阵的定义和管理、矩阵运算、矩阵解析等四个主题，为后续学习 STATA 编程知识奠定了扎实的基础。

第 5 讲介绍 STATA 编程的基本知识，包括：程序的定义和调用、单值、暂时性物件（暂元、暂时性变量、暂时性文件、暂时性矩阵等）、循环语句、条件语句，以及 STATA 返回值的引用等内容。通过本讲的学习，学员将能够通过独立编写 STATA 程序来提高数据处理和模型估计的效率，为后续学习 STATA 高级程序奠定基础。

以上各讲都以专题的形式进行讲解，其中不但包含了 STATA 官方提供的常用命令，还包含了大量外部命令（520 多个），为学员提供了强有力的实证分析工具。

2 课程特色

STATA 初级视频教程非常注重培养学员的实证分析能力。我们一方面在尽可能呈现 STATA 全貌的同时做到重点突出、结构清晰，另一方面则强调实际问题的解决能力。例如，传统教科书中很少会涉及数据的合并、离群值的处理、面板资料的筛选等问题，而这些往往又是实证分析过程中至关重要的环节。又如，在整理回归结果的过程中，多数学员主要通过“复制-粘贴”的方式，效率甚低。通过本课程的学习，这些问题都能得到很好的解决。本课程的主要特色包括：

(1) 视频教学，不受时间地点的限制，随时可以开始学习。

(2) 以专题的方式讲解，并辅以大量的实例，有助于学员快速掌握 STATA 数据处理、编程、绘图等方面的基本技巧，举一反三。

(3) 翔实的课程配套资料。本课程除提供视频教程外，还同时提供了大量的范例数据和外部命令、PDF 版 do 文档，以及具有“时间节点”的视频目录，便于学员课后练习和快速查找自己感兴趣的内容。详见“3 课程配套资料”和“4 配套资料的使用方法”。

(4) 提供多种答疑和交流方式。详见“5 讨论和建议”。

3 课程配套资料

(1) 本课程中使用的 do 文档和 ado 文档，便于学员在观看视频的同时进行操作练习。

(2) 视频中所用到的所有 do 文档的 PDF 打印版 ([Stata_A_dofiles](#))，附详细书签，便于快速查阅和定位。可以利用 adobe 的 PDF 文件搜索功能查找某个 stata 命令，以便了解该命令在不同情形下的使用方法。


(3) STATA11 电子手册全套（能与 STATA 帮助文件完美链接）、STATA Journal 全套（2001-2009）、STATA Technical Bulletin (STB) 全套 (Vol.1-Vol.61, 1991-2001)。这些资料统一存放于 [STATA_资料](#) 文件夹中。

(4) 100 余个范例数据，包括：STATA 官方范例数据包、中国宏观经济、中国上市公司范例数据等。

(5) 520 多个 STATA 外部命令。

4 配套资料的使用方法

Q1. 如何放置 STATA11 软件和相关资料？

A: 请将课程附赠的 stata11.rar 文件解压后放置于 D 盘根目录下（注意：是 D:\stata11，而不是 D:\stata11\stata11，另外，也请注意文件名称的大小写）。双击  图标即可打开 stata 软件（无需安装）。


Q2. 如何打开视频教程中使用的 do 文档？


A: 视频中使用的所有 do 文档统一存放于 D:\stata11\ado\personal\Net_course_A 文件夹下。若希望练习第一讲中的相关操作，可依次执行如下命令：

Step1: 在 STATA 命令窗口中输入 `cd D:\stata11\ado\personal\Net_course_A` 命令，定义当前工作路径（会显示在 STATA 屏幕左下角）；

Step2: 输入 `doedit A1_intro` 命令，即可打开第一讲的讲义 A1_intro.do 文件。

当然，我们也可以通过点击菜单的方式完成上述操作，步骤如下：

Step1: 在 STATA 主菜单中点击“New do-file Editor”图标 ；

Step2: 在第一步中弹出的“Do-file Editor”窗口中点击“Open”图标 ，然后到 D:\stata11\ado\personal\Net_course_A 文件夹下，双击“A1_intro”文件即可打开之。

Q3. 如何快速定位到自己希望观看的视频节点？

A: 你可以先浏览一下“附录 A: STATA 初级视频目录”，进而确定你希望观看的内容所在的视频文件，然后根据“时间节点目录”拖拽视频即可实现快速定位。如果想了解特定命令（如 `esttab`）的使用方法，只需在 `Stata_A_dofiles.pdf` 文档中执行全文搜索，由此便可确定该命令所在的视频文件，进而通过“时间节点目录”实现精确定位。此外，我们已经为 `Stata_A_dofiles.pdf` 文档制作了详细的书签，便于大家浏览整个教程的结构安排。


Q4. 如何修改 STATA 启动时的默认设定？

A: 在每次打开 STATA 时，它会自动执行 D:\stata11 安装目录下的 `profile.do` 文件，该文件设定了默认工作路径、STATA 内存的大小、STATA 结果窗口的最大显示行数，以及日志文件（log 文件）的自动存储方式等内容。你可以执行 `doedit D:\stata11\profile.do` 命令打开这个文件并修改其中的设定。具体设定方法参见视频文件 `▲A1_intro_02_intro` 中第“1.2.2 首次使用 STATA 的一些基本设定”小节介绍。

Q5. 如何查看每次执行 STATA 的 log 文件？

A: 我在 `profile.do` 文件中进行了设定，每次启动 STATA11 后，屏幕上呈现的所有结果都会被自动记录到一个日志文件中（详见视频文件 `▲A1_intro_06_log` 中“1.8 log 文件：记录你的分析过程”的介绍）。这些日志文件被自动保存在 D:\stata11\do 文件夹下，文件名称为打开 STATA 时的时间，例如，名称为 `s22Apr2010111857.txt` 的日志文件是我在 2010 年 4 月 22 日 11 点 18 分 57 秒打开 STATA11 时自动生成的。

Q6. 如何使用 STATA 官方提供的范例数据？

A: 多数情况下，教程中提供的资料已经足够大家练习所需了。当然，为了满足个别学员的需求，本教程还进一步提供了 STATA 官方范例数据，统一存放于 D:\stata11\ado\Example 文件夹中。例如，你可以输入 `use "D:\stata11\ado\Examples\XTFiles\abdata.dta", clear` 命令导入范例数据 `abdata.dta`。亦可点击 STATA 主界面中的  按钮（快捷键为 `Ctrl+o`），到上述目录下选择你想要练习的数据文件。

5 讨论和建议

(1) 人大经济论坛STATA专版: <http://www.pinggu.org/bbs/index.asp?boardid=67>

(2) 人大经济论坛“统计软件培训班VIP答疑区”。所有STATA初级视频教程的学员都将自动升级为人大论坛VIP会员,在答疑区可就视频教程中的疑难问题提问和讨论。授课老师会在24小时内给出解答。<http://www.pinggu.org/bbs/forum-114-1.html>。

(3) 连玉君的博客: <http://blog.cnfol.com/arlion>, 提供了大量STATA资源。

(4) E-mail: arlionn@163.com。在网络不畅或问题较为复杂的情况下,学员亦可通过E-mail与授课教师讨论问题。

6 讲师介绍

连玉君,经济学博士,2007年7月毕业于西安交通大学金禾经济研究中心,现为中山大学岭南学院金融系讲师。主讲课程为计量分析与STATA应用、实证金融、金融计量。目前已在《Global Finance Journal》、《世界经济》、《统计研究》、《经济学动态》、《经济学(季刊)》、《财经研究》等期刊发表论文20余篇,出版专著一部(《中国上市公司投资效率研究》,经济管理出版社,2009年8月)。目前主持教育部人文社科基金、高校基本业务费基金、广东自然科学基金项目各一项,并曾参与了多项国家自然科学基金、社会科学基金以及总理基金项目,具有丰富的数据处理、模型构建和模拟分析经验。目前已完成Panel VAR(1800余行)、Panel Threshold(1200余行)、Two-tier Stochastic Frontier(500余行)等复杂计量模型的STATA实现程序,并编写过几十个小程序,如xtbalance.ado、bdiff.ado、gqhet.ado等等。

(个人主页: <http://www.lingnan.net/intranet/teachinfo/dispuser.asp?name=lianyj>)

7 报名咨询

QQ号: 619492407

MSN: pinggu.org@hotmail.com

电话: (010)68472925

邮箱: pinggu.org@gmail.com

网址: <http://baoming.pinggu.org/Default.aspx?id=16>

8 培训优惠

(1) 培训班是目前业内最优惠的价格,欲报从速。同时报两个班及以上,9折优惠

(2) 所有参与培训的学员都将享受人大经济论坛VIP身份待遇一年

(3) 为学员在论坛开设“[统计软件培训班VIP答疑区](#)”,提供独享的疑难解答和经验交流。

附录 A: STATA 初级视频目录 (时间节点)

=====

第一讲 STATA 简介

=====

1.1 本课程简介	 A1_intro_01_outline
1.1.1 课程纲要	00:04:50
1.1.2 课程特点	00:07:25
1.1.3 课程配套资料	00:08:50
1.1.4 讨论和建议	00:25:35
1.2 STATA 概貌	 A1_intro_02_intro
1.2.1 stata 界面	00:01:32
1.2.2 首次使用 STATA 的一些基本设定	00:11:28
1.3 输入和导入数据	 A1_intro_03_inputdata
1.3.1 手动输入	00:01:38
1.3.2 从 .txt, excel 表格中粘贴	00:10:05
1.3.3 使用 stata 命令: infile, insheet, infix	00:12:19
1.3.3.1 以-tab-分隔的数据: -insheet- 命令	00:13:03
1.3.3.2 以 空格 分隔的数据: -infile- 命令	00:15:59
1.3.3.3 调入 STATA 格式的数据: -use- 命令	00:22:22
1.3.3.4 调入 Excel 格式的数据: -xmluse-命令	00:24:05
1.3.3.5 行列对调的数据	00:29:19
1.3.4 时间序列资料	00:31:11
1.3.5 面板资料	00:33:36
1.3.6 STATA 官方提供的资料	00:36:27
1.3.7 其它软件中的数据	00:39:19
1.4 存储和导出数据	00:40:53
1.4.1 存储数据	00:41:13
1.4.2 导出和转换	00:42:04
1.4.2.1 -outfile-命令:导出为 .raw 文本格式	00:42:09
1.4.2.2 -outsheet-命令:导出为 -Tab- 分隔的文本文件	00:44:02
1.4.2.3 -xmlsave-命令:导出为 XML 格式	00:44:59
1.4.2.4 -dataout-命令:导出为 Word,Excel,Tex	00:45:58
1.4.2.5 -outdat- 命令:导出为 .spss, .rats, .limdep 格式	00:51:12
1.5 浏览资料	 A1_intro_04_browse
1.5.1 变量的名称	00:01:18

1.5.2 查看资料的结构	00:02:49
1.5.2.1 更改变量的存储类型	00:09:02
1.5.2.2 -list- 命令的使用	00:12:11
1.5.2.3 定义变量的显示格式	00:15:14
1.5.2.4 数据和变量的标签	00:17:55
1.5.2.5 附加说明文字	00:25:26
1.5.2.6 搜索变量	00:27:11
1.5.3 基本统计量	00:30:30
1.5.3.1 -summarize-命令	00:30:43
1.5.3.2 -codebook-命令	00:33:20
1.5.3.3 -inspect-命令	00:35:24
1.5.3.4 列表统计(table, tabulate)	00:36:28
1.5.3.5 论文格式的统计表格(tabstat)	00:40:21
1.5.3.6 将统计结果输出到 txt 文档中	00:44:04
1.5.4 基本图形分析	00:48:52
1.5.4.1 直方图：样本的总体分布情况	00:49:04
1.5.4.2 密度函数图	00:54:34
1.5.4.3 散点图	00:55:53
1.5.4.4 相关系数矩阵	00:57:19
1.6 执行指令	 A1 intro 05 command
1.6.1 概览	00:00:15
1.6.2 命令的适用范围	00:06:03
1.6.2.1 列举多个变量	00:06:10
1.6.2.2 样本范围的限制	00:07:46
1.6.3 命令作用的增减：使用选项	00:09:45
1.7 修改资料	00:12:17
1.7.1 数学表达式	00:12:45
1.7.2 变量的创建和修改	00:15:07
1.7.2.1 变量的存储类型	00:15:18
1.7.2.2 创建新变量	00:17:22
1.7.2.3 修改旧变量	00:21:30
1.7.2.4 删除变量和样本值	00:29:45
1.7.2.5 移动变量窗口中变量的位置	00:36:15
1.7.2.6 克隆已有变量	00:39:07
1.7.2.7 拆分变量	00:41:12
1.7.3 样本值的排序	00:43:51
1.8 log 文件：记录你的分析过程	 A1 intro 06 log
1.8.1 log 文件简介	00:01:31
1.8.2 将 log 文件转换为网页	00:08:25

1.8.2.1 -log2html-命令：制作“单页”网页	00:08:41
1.8.2.2 -hyperlog-命令：制作“框架型”网页	00:12:01
1.8.2.3 其他命令	00:14:26
1.9 do 文档：高效快捷地执行命令	A1_intro_07_dofile
1.9.1 do 文档简介	00:01:16
1.9.1.1 打开 do 文档编辑器	00:01:59
1.9.1.2 保存和关闭	00:04:11
1.9.1.3 执行 do 文档	00:05:02
1.9.2 合理规划你的 do 文档	00:06:13
1.9.2.1 一些基本规则	00:06:29
1.9.2.2 注释语句	00:08:00
1.9.2.3 断行	00:08:56
1.9.2.4 大型 do 文档的设定	00:11:10
1.9.3 列印文字	00:12:14
1.9.3.1 -display-命令	00:12:34
1.9.3.2 列印的颜色	00:14:46
1.9.3.3 列印的位置	00:16:16
1.9.4 关于编辑器	00:19:36
1.9.5 do 文件的转换(制作网页教程)	00:20:49
1.10 stata 与 Excel、Word、LaTeX 的亲密接触	A1_intro_08_word_excel
1.10.1 统计表格、矩阵的输出	00:01:39
1.10.1.1 输出基本统计量	00:02:25
1.10.1.2 输出相关系数矩阵	00:05:03
1.10.1.3 输出矩阵	00:06:39
1.10.1.4 其它说明	00:07:49
1.10.2 估计结果的输出	00:10:04
1.10.2.1 -esttab-命令：回归结果的呈现	00:10:36
1.10.2.2 -logout-命令：输出 Excel、Word、TeX 文档	00:24:44
1.10.2.3 -xml_tab-命令：输出 Excel 文档	00:30:21
1.10.2.4 -outreg2-命令：输出 Word、Excel 文档	00:36:06
1.11 Stata 设定	A1_intro_09_set
1.11.1 Stata 帮助	00:01:12
1.11.2 文件目录	00:15:11
1.11.3 Stata 外部命令的获取	00:20:37
1.11.3.1 外部命令的存储路径	00:20:47
1.11.3.2 外部命令的获取方式	00:21:29
1.11.3.3 外部命令的管理和更新	00:34:07
1.11.4 Stata 的系统参数	00:39:12
1.11.5 文件和文件夹的操作	00:44:47

1.11.5.1 文件的基本操作：查找、查看、复制和删除	00:44:51
1.11.5.2 使用 stata 打开-.txt-, -Word-, -Excel-文件	00:49:58
1.11.5.3 文件夹的操作	00:55:06
1.11.6 每次启动时均需执行的命令(profile)	00:58:58
1.11.7 常用快捷键	01:02:20
1.11.8 退出 stata(exit)	01:03:28

2.1 创建变量的更多技巧	 A2_data_01_gen
2.1.1 <code>_n</code> 和 <code>_N</code>	00:02:12
2.1.1.1 <code>_n</code> 与 <code>_N</code>	00:02:20
2.1.1.2 <code>_n</code> 与 <code>_N</code> 的应用	00:06:48
2.1.2 虚拟变量的产生	00:16:48
2.1.2.1 基本方式	00:17:16
2.1.2.2 基于类别变量生成虚拟变量: <code>-tab-</code> 命令	00:19:19
2.1.2.3 基于类别变量生成虚拟变量: <code>-xi-</code> 命令	00:20:46
2.1.2.4 因子变量 (stata11 的一大亮点)	00:27:28
2.1.2.5 将连续变量转换为类别变量	00:46:12
2.1.2.6 利用条件函数产生虚拟变量	00:55:34
2.1.3 交乘项的产生	 A2_data_02 egen
2.1.4 <code>-egen-</code> 命令	00:05:27
2.1.4.1 <code>egen</code> 与 <code>gen</code> 的区别	00:05:58
2.1.4.2 产生等差数列: <code>seq()</code> 函数	00:09:26
2.1.4.3 填充数据: <code>fill()</code> 函数	00:11:51
2.1.4.4 产生组内均值和中位数	00:13:21
2.1.4.5 跨变量的比较和统计	00:20:47
2.1.4.6 变量的标准化	00:22:23
2.1.4.7 变量的平滑化 (Moving Average)	00:25:43
2.1.4.8 更多的 <code>egen()</code> 函数	00:29:03
2.2 分位数	 A2_data_03_quantile
2.2.1 分位数的基本概念	00:01:16
2.2.2 <code>-pctile-</code> 命令	00:09:28
2.2.3 <code>-xtile-</code> 命令	00:11:55
2.2.4 <code>-_pctile-</code> 命令	00:16:23
2.3 重复样本值的处理	 A2_data_04_duplicates
2.3.1 检查重复的样本组	00:01:11
2.3.2 标记和删除重复的样本组合	00:08:24
2.4 缺漏值的处理	 A2_data_05_omit
2.4.1 缺漏值简介	00:01:15
2.4.2 缺漏值的标记	00:04:51
2.4.3 查找/删除缺漏值	00:11:49

2.4.3.1 缺漏值的形态	00:12:07
2.4.3.2 删除缺漏值	00:18:44
2.4.4 填补空缺(gap)	00:27:33
2.4.5 多重补漏分析(multiple-imputation)	00:33:59
2.4.5.1 MI 简介	00:34:43
2.4.5.2 实例分析	00:39:35
2.4.5.3 MI impute regress 的假设条件	00:54:20
2.4.5.4 其它补漏方法	00:59:23
2.4.5.5 假设检验	01:00:28
2.5 离群值的处理	A2_data_06_outlier
2.5.1 离群值的影响	00:00:47
2.5.2 查找离群值	00:03:32
2.5.3 离群值的处理	00:14:08
2.5.3.1 删除	00:14:15
2.5.3.2 对数转换	00:15:52
2.5.3.3 缩尾处理	00:20:38
2.5.3.4 截尾处理	00:27:44
2.6 资料的合并和追加	A2_data_07_merge
2.6.1 横向合并：增加变量	00:01:40
2.6.1.1 一对一合并	00:02:51
2.6.1.2 多对一合并	00:14:14
2.6.1.3 一对多合并	00:18:24
2.6.1.4 一个例子	00:19:45
2.6.2 横向关联: -joinby-	00:22:20
2.6.3 纵向合并：追加样本	00:28:02
2.6.4 大型数据的处理	00:35:54
2.6.5 一些有用的外部命令	00:39:35
2.7 重新组合样本	A2_data_08_reshape
2.7.1 样本的转置	00:00:39
2.7.2 数据的纵横变换	00:02:03
2.7.3 样本的交叉组合	00:08:07
2.7.3.1 -fillin- 命令	00:08:17
2.7.3.2 -cross-命令	00:12:03
2.7.4 样本的堆砌	00:15:45
2.8 文字变量的处理	A2_data_09_string
2.8.1 将文字转换成数字	00:00:56
2.8.1.1 以文字类型存储的数字之转换	00:01:00

2.8.1.2 纯文字类别变量之转换	00:06:59
2.8.2 将数字转换成文字	00:13:19
2.8.3 文字样本值的分解	00:20:28
2.8.4 处理文字的函数	00:23:48
2.8.4.1 文字函数简介	00:24:08
2.8.4.2 例-1-: 上市公司日期、行业代码和所在地	00:31:47
2.8.4.3 例-2-: 银企关系数据中银行名称的提取	00:41:15
2.8.4.4 例-3-: 处理不规则的日期	00:45:42
2.9 类别变量的分析	A2_data_10_cat
2.9.1 类别数的统计	00:01:10
2.9.2 交叉类别变量的生成	00:04:41
2.9.3 分组统计量	00:08:17
2.9.3.1 单层分组统计量	00:08:42
2.9.3.2 二层次和三层次分组统计量	00:14:28
2.9.3.3 多层次分组统计量	00:18:44
2.9.4 计算分组统计量的其它方法	00:22:21
2.9.4.1 -egen-命令	00:22:35
2.9.4.2 转换原资料为分组统计量: -collapse-命令	00:23:48
2.9.5 图示分组统计量	00:30:50
2.9.5.1 柱状图	00:31:13
2.9.5.2 箱形图	00:38:48
2.10 时间序列资料的处理	A2_data_11 TimeS
2.10.1 简介	00:01:42
2.10.1.1 声明时间序列: tsset 命令	00:01:43
2.10.1.2 检查是否有断点	00:04:26
2.10.1.3 填充缺漏的日期	00:06:09
2.10.1.4 追加样本	00:07:57
2.10.2 时序变量的生成	00:11:29
2.10.2.1 滞后项、超前项和差分	00:11:43
2.10.2.2 产生增长率变量: 对数差分	00:14:41
2.10.2.3 日期变量的处理	00:15:56
2.11 面板资料的处理	A2_data_12 Panel a
2.11.1 声明面板资料: xtset 命令	00:01:41
2.11.2 公司数目和年度的统计	00:10:34
2.11.2.1 面板资料的基本描述: xtodes 命令	00:11:04
2.11.2.2 记录面板的资料形态: xtpattern 命令	00:16:00
2.11.2.3 统计公司数目: panels 命令	00:20:21
2.11.3 产生连续的公司代码	00:24:59


2.11.4	处理为平行面板	00:30:01
2.11.5	剔除 IPO 当年的数据	00:35:46
2.11.6	行业发生变更的公司	00:43:11
	♠ A2_data_13_Panel_b	
2.11.7	如何删除面板资料首尾的缺漏值?	00:02:45
2.11.8	仅保留连续 T 年以上可获得资料的公司	00:19:50
2.11.9	面板资料瘦身 I: 每隔 T 年保留一次资料	00:35:17
2.11.10	面板资料瘦身 II: 采用 P 年平均值进行估计	00:44:15
2.11.11	面板缺漏值的扩充	00:51:08
2.11.12	变量的“去均值”和标准化处理	01:00:28
2.11.13	面板资料处理的其他主题	01:03:36
2.12	数据的查验和比较	♠ A2_data_14_compare
2.12.1	查验变量	00:01:35
2.12.1.1	计数	00:01:38
2.12.1.2	条件确认	00:03:08
2.12.1.3	比较变量的大小	00:07:46
2.12.2	查验两组数据	00:09:05
2.12.2.1	查验两笔数据的观察值是否一致	00:09:08
2.12.2.2	查验两笔数据的变量是否一致	00:13:04


=====

第三讲 Stata 绘图

=====

3.1 简介	A3_graph_01_intro
3.1.1 Stata 图形的种类	00:02:12
3.1.2 二维图命令的基本结构	00:03:41
3.1.3 几种常用图形的简单示例	00:15:19
3.1.4 图形的管理	00:20:08
3.1.4.1 图形的保存	00:20:23
3.1.4.2 图形的导出	00:23:20
3.1.4.3 图形的调入	00:26:26
3.1.4.4 插入 Word	00:27:27
3.1.4.5 查询	00:31:31
3.1.4.6 重新显示图形	00:31:53
3.1.4.7 图形的合并	00:32:52
3.1.4.8 删除图形	00:34:54
3.1.5 图形的显示模式(绘图模板)	00:35:20
3.1.5.1 显示模式种类	00:35:26
3.1.5.2 中文投稿的黑白图	00:40:06
3.1.5.3 stata 用户提供的模板	00:42:40
3.1.5.4 创建自己的图形模板	00:46:51
3.2 二维图选项	A3_graph_02_option1
3.2.1 坐标类	00:02:15
3.2.1.1 坐标轴刻度(tick)及刻度标签(label)	00:02:20
3.2.1.2 坐标轴标题: ytitle() xtitle()	00:11:29
3.2.1.3 坐标结构: yscale() xscale()	00:13:45
3.2.1.4 双坐标系	00:23:26
3.2.2 标题类	00:29:25
3.2.2.1 标题的种类	00:29:33
3.2.2.2 示例	00:29:49
3.2.2.3 标题的位置	00:32:16
3.2.3 区域类	00:36:45
3.2.3.1 Stata 图形的区域划分	00:36:49
3.2.3.2 控制内区和外区的边距	00:38:19
3.2.3.3 控制图形的纵横比例	00:41:57
3.2.3.4 绘图区的显示模式	00:42:48
3.2.3.5 绘图区和全图区背景颜色的控制	00:43:34
3.2.4 图例类	00:43:58
3.2.4.1 自动产生的图例	00:44:02

3.2.4.2 重新定制图例		00:45:57
3.2.4.3 图例的位置		00:48:26
3.2.4.4 多个图例的重排		00:50:43
3.2.4.5 线型的控制		00:52:47
3.2.5 附加线类	 A3_graph_03_option2	
3.2.5.1 选项结构		00:01:22
3.2.5.2 附加线 <位置>		00:01:58
3.2.5.3 附加线 <风格>		00:03:00
3.2.5.4 附加线 <线宽>		00:04:50
3.2.5.4 附加线 <颜色>		00:06:11
3.2.5.5 附加线 <线型>		00:07:12
3.2.5.5 附加线属性的独立性		00:08:57
3.2.6 文字与文本框		00:09:54
3.2.6.1 选项类别		00:10:38
3.2.6.2 文字和文本框的整体风格		00:11:17
3.2.6.3 文本框属性		00:12:38
3.2.6.4 文字属性		00:17:15
3.2.7 图标类		00:23:47
3.2.7.1 简介		00:24:04
3.2.7.2 图标的位置		00:25:54
3.2.7.3 图标的大小		00:30:39
3.2.7.4 图标的角度		00:32:13
3.2.7.5 图标的颜色		00:34:09
3.2.8 其它选项		00:34:25
3.2.8.1 分组绘图		00:34:33
3.2.8.2 重新设置变量标签		00:44:11
3.2.8.3 重新设置变量显示格式		00:45:32
3.2.8.4 重设图形种类		00:46:57
3.3 元素代号	 A3_graph_04_element	
3.3.1 颜色代号		00:02:09
3.3.2 线 相关的代号		00:11:11
3.3.2.1 线型代号		00:11:59
3.3.2.2 线宽代号		00:13:00
3.3.2.3 连接方式代号		00:13:49
3.3.3 标记符号的代号		00:15:06
3.3.3.1 符号样式		00:16:16
3.3.3.2 符号的边界和填充		00:17:41
3.3.3.3 符号代号一览		00:18:44
3.3.4 文字相关的代号		00:20:17

3.3.4.1 文字大小代号	00:20:28
3.3.4.2 文字角度代号	00:20:48
3.3.4.3 文字对齐方式的代号	00:21:06
3.3.5 边距大小的代号	00:21:13
3.4 常用图形示例 (I)	 A3_graph_05_eg1
3.4.1 散点图	00:04:33
3.4.2 折线图	00:06:02
3.4.3 区域图	00:11:34
3.4.4 钉形图	00:13:55
3.4.5 直方图	00:21:15
3.4.6 密度函数图	00:35:37
3.4.7 累积分布函数图	00:47:25
3.4 常用图形示例 (II)	 A3_graph_06_eg2
3.4.8 线性/非线性 拟合图	00:01:17
3.4.9 矩阵图: 显示变量间的相关性	00:05:56
3.4.10 柱状图	00:13:18
3.4.10.1 一维柱状图	00:13:21
3.4.10.2 二维柱状图	00:29:23
3.4.11 点图	00:37:04
3.4.12 函数图	00:40:08
3.4.13 合图示例	00:49:17
3.4.14 三维图形	01:00:50
3.4.15 地 图	01:02:23
3.5 结语	01:10:46

=====
 第四讲 矩阵操作
 =====

4.1 矩阵的基本操作	 A4_matrix_01_intro
4.1.1 基本定义方式	00:02:26
4.1.2 矩阵的管理	00:05:32
4.1.2.1 矩阵的名称	00:05:42
4.1.2.2 列示矩阵	00:08:47
4.1.2.3 矩阵的行数和列数	00:23:30
4.1.2.4 查找/删除矩阵	00:25:01
4.1.2.5 查验矩阵中是否存在缺漏值	00:26:05
4.1.3 矩阵的行名和列名	00:27:32
4.1.4 选取部分矩阵	00:30:15
4.1.4.1 选取 1 个元素: 1*1 矩阵	00:30:45
4.1.4.2 选取子矩阵	00:31:30
4.1.4.3 矩阵元素的修改	00:33:10
4.1.5 更一般化的矩阵定义	00:34:51
4.1.6 常用矩阵的定义	00:37:00
4.1.6.1 单位矩阵	00:37:14
4.1.6.2 常数矩阵	00:37:37
4.1.6.3 元素为随机数的矩阵	00:43:40
4.1.6.4 对角矩阵	00:45:43
4.1.7 变量和矩阵的相互转换	00:48:06
4.1.7.1 变量—>矩阵	00:48:42
4.1.7.2 矩阵—>变量	00:54:24
4.1.8 用矩阵存储统计结果	00:59:51
4.1.8.1 以矩阵方式呈现 tabstat 命令的结果	01:00:20
4.1.8.2 更一般化的矩阵存储	01:02:14
4.1.9 采用变量的方式操作矩阵	01:09:22
4.1.9.1 对矩阵中的各列进行变换和运算	01:09:48
4.1.9.2 矩阵元素的数学变换	01:11:00
4.1.10 矩阵的保存和调入	01:12:13
4.1.10.1 将矩阵保存为 .dta 文档中	01:12:48
4.1.10.2 将矩阵保存到 txt, word, excel 文档中	01:16:18
4.2 矩阵运算	 A4_matrix_02_oper
4.2.1 矩阵的基本运算	00:02:19
4.2.1.1 加、减、乘	00:02:24
4.2.1.2 直乘	00:04:52
4.2.1.3 哈式乘法	00:08:27

4.2.1.4 矩阵元素的数学变换	00:09:42
4.2.1.5 矩阵与单值的运算	00:18:15
4.2.2 矩阵的转置	00:19:39
4.2.3 矩阵的逆矩阵	00:20:56
4.2.3.1 矩阵的行列式	00:21:11
4.2.3.2 矩阵求逆	00:23:16
4.2.4 矩阵的向量化	00:24:23
4.2.5 矩阵的对角值	00:31:11
4.2.6 交乘矩阵的定义	00:32:20
4.2.6.1 简单交乘矩阵	00:33:35
4.2.6.2 加权交乘矩阵	00:42:52
4.2.6.3 用户自行设定的权重	00:51:54
4.2.6.3 特殊加权交乘矩阵	00:57:02
4.3 矩阵的解析	♠A4 matrix_03 more
4.3.1 线性相关、线性独立和正交向量	00:02:45
4.3.2 矩阵的秩	00:04:29
4.3.3 特征根和特征向量	00:07:45
4.3.4 正定矩阵和负定矩阵	00:17:15
4.3.5 裘氏分解	00:20:49
4.4 关于矩阵的进一步说明	00:28:21
4.4.1 矩阵函数	00:28:36
4.4.2 返回系统中的矩阵	00:29:51
4.4.3 定义约束矩阵	00:33:21
4.4.4 矩阵与暂元的相关操作	00:38:41
4.4.5 矩阵对内存的需求	00:42:51

第五讲 STATA 编程初步

5.1 stata 程序简介	A5_prog_01_intro
5.1.1 Stata 程序的基本结构	00:03:30
5.1.2 程序的执行	00:07:35
5.1.2.1 第一种执行方式: ado 文档执行方式	00:07:40
5.1.2.2 第二种执行方式: run(Ctrl+R)	00:11:55
5.1.3 程序的管理	00:13:49
5.1.4 避免列印过多的结果	00:18:59
5.1.5 避免程序因错误而中断	00:21:07
5.1.6 避免数据在程序执行过后有所变动	00:30:38
5.2 单值(scalar)	A5_prog_02_local
5.2.1 存放数值	00:00:45
5.2.2 存放字符串	00:01:34
5.2.3 执行命令后的单值结果	00:04:41
5.2.4 单值的管理	00:09:07
5.3 暂元(local)	00:11:16
5.3.1 暂元的定义和引用	00:11:46
5.3.1.1 暂元的基本功能	00:11:51
5.3.1.2 数学运算符的处理	00:19:21
5.3.1.3 复合双引号: `"'`'`"'`	00:20:53
5.3.1.4 暂元中的暂元	00:23:13
5.3.1.5 暂元引用机制的简化	00:26:07
5.3.2 全局暂元	00:32:03
5.3.3 暂元的管理	00:36:06
5.4 其它暂时性物件	00:36:21
5.4.1 暂时性变量	00:36:47
5.4.2 暂时性矩阵和暂时性单值	00:41:51
5.4.3 暂时性文件	00:49:39
5.5 控制语句	A5_prog_03_loop
5.5.1 循环语句	00:00:41
5.5.1.1 条件循环: while 语句	00:00:42
5.5.1.2 forvalues 语句	00:20:55
5.5.1.3 foreach 语句	00:49:06
5.6 条件语句	A5_prog_04_if
5.6.1 if 语句	00:00:44

5.6.2 一些有用的条件函数	00:28:26
5.6 引用 Stata 命令的返回值	♣A5_prog_05_return
5.6.1 留存在内存中的结果	00:01:20
5.6.2 r-class	00:04:55
5.6.3 e-class	00:22:01
5.6.4 c-class	00:35:27

附录 B: STATA 高级视频教程简介

STATA 高级视频教程是初级视频教程的姊妹篇。整个教程共 50 多个学时, 包含 48 个视频文件, 内容涉及多种常用的估计方法(如普通最小二乘法、广义最小二乘法、非线性最小二乘法、最大似然估计、IV 估计和 GMM)、时间序列分析(如 ARMA 模型、VAR 模型、单位根检验、协整分析等)、面板数据模型(固定效应、随机效应、面板 IV-GMM、动态面板、面板单位根检验、面板协整分析等)。同时, 视频中还细致的讲述了 STATA 编程、Bootstrap 以及 Monte Carlo 模拟相关的内容。

STATA 高级视频教程非常强调计量理论和应用的结合。在讲解每一个计量模型的过程中, 我们会首先介绍该模型的理论基础和基本思想, 进而结合实例讲解该模型在 STATA 中的操作方法。课程的主要特色包括: (1) 以专题的方式讲解实证分析中常用的各种计量模型, 并辅以大量的实例, 有助于学员迅速理解各个模型的基本思想, 掌握其估计和检验方法; (2) 借助计算机模拟数据来理解不同估计方法的优劣, 这一方面有利于学员理解不同估计量的小样本性质, 另一方面也有助于学习模拟数据的产生方法。这也是最近几年国外计量经济学教学的一个主要发展方向。(3) 提供了实证分析过程中经常遇到的一些细节问题的处理方法, 这是传统教科书中不会涉及的, 但在实证分析过程中却又非常重要的问题。

STATA 高级视频的报名网址为: <http://baoming.pinggu.org/Default.aspx?id=25>

高级部分: 计量分析与 Stata 应用

1. 普通最小二乘法 (OLS)
 - 1.1 解读 OLS 回归结果
 - 1.2 残差分析与稳健型估计
 - 1.3 管理多个回归结果
2. 广义最小二乘法 (GLS)
 - 2.1 GLS 的基本思想
 - 2.2 异方差
 - 2.3 序列相关
 - 2.4 似无相关模型 (SUR)
3. 非线性最小二乘法 (NLS)
 - 3.1 NLS 的基本思想
 - 3.2 NLS 程序的编写
 - 3.3 范例: 估计动态部分调整模型
4. 最大似然估计 (MLE)
 - 4.1 MLE 的基本原理
 - 4.2 似然函数的设定
 - 4.3 程序的调试、起始值的设定和相关问题
 - 4.4 范例: 线性回归模型、Logit 模型、Probit 模型
5. 工具变量法与 GMM

- 5.1 内生性问题与工具变量法
- 5.2 两阶段最小二乘法 (2SLS)
- 5.3 广义矩估计法 (GMM)
- 5.4 过度识别检验 (Sargan 检验与 Hausman 检验)
- 5.5 弱工具变量问题
- 6. 时间序列分析**
 - 6.1 时间序列资料的处理
 - 6.2 ARIMA 模型
 - 6.3 向量自回归 (VAR) 模型: 估计和检验
 - 6.4 向量自回归 (VAR) 模型: 因果检定和冲击反应
 - 6.5 单位根检验
 - 6.6 协整分析和误差修正模型
 - 6.7 GARCH 模型 (GARCH, E-GARCH, T-GARCH)
- 7. 面板数据模型**
 - 7.1 静态面板模型: 固定效应 v.s. 随机效应
 - 7.2 时间效应、模型的筛选和常见问题
 - 7.3 异方差、序列相关和截面相关
 - 7.4 内生性问题 (面板 IV-GMM 估计)
 - 7.5 动态面板模型 (Difference GMM 和 System GMM)
 - 7.6 面板随机系数模型
 - 7.7 面板随机前沿模型
 - 7.8 面板单位根检验
 - 7.9 面板协整分析
- 8. STATA 高级程序**
 - 8.1 暂元的高级功能
 - 8.2 暂时性物件
 - 8.3 输入项
 - 8.4 输出项
 - 8.5 可分组执行的程序
 - 8.6 可重新显示结果的程序
 - 8.7 子程序
 - 8.8 程序勘误与调试
 - 8.9 帮助文件的编写
- 9. 模拟分析 (Simulation) 与自体抽样 (Bootstrap)**
 - 9.1 随机数的产生和常用分布
 - 9.2 Bootstrap
 - 9.3 Bootstrap 的基本原理
 - 9.4 Bootstrap 获得标准误
 - 9.5 Bootstrap 获得置信区间
 - 9.6 Bootstrap 检验组间差异
 - 9.7 Bootstrap 应用实例
 - 9.8 Jackknife (刀切法)
 - 9.9 Jackknife 的基本思想
 - 9.10 Jackknife 应用实例

- 9.11 Jackknife 与 Bootstrap 的差别
- 9.12 Permutation Tests (组合检验)
- 9.13 Monte Carlo Simulation (蒙特卡罗模拟分析)
- 9.14 MC 的基本思想
- 9.15 MC 应用实例
- 9.16 统计量的 Power 和 Size
- 9.17 模拟数据的产生