

THE STATA NEWS

October/November/December 2010

Vol 25 No 4

In the spotlight

Learn about Stata 11's competing-risks regression capabilities.

p. 1

Stata now has an official blog

Read about the new Stata blog, *Not Elsewhere Classified*.

p. 2

New from Stata Press

Read about the latest book published by Stata Press.

p. 3

Stata Conference Chicago 2011

Announcement and call for presentations.

p. 6

2011 Mexican Stata Users Group meeting

Announcement and call for presentations.

p. 7

Also in this issue

New from the Stata Bookstore 3

Additional featured titles 4

Public training 5

Meet with Stata developers at ASSA 8

Stata is now on Facebook and Twitter 8

Upcoming NetCourses 8

The Stata News

Executive Editor: Karen Strope
Production Supervisor: Annette Fett

In the spotlight: Competing-risks regression

Competing-risks analysis generalizes standard survival analysis to cases in which subjects are exposed to more than one cause of failure. For example, consider a study in which we record the time from initial tumor removal to the recurrence of breast cancer in relation to factors such as drug treatment and age. Disease progression at an old tumor site (a local relapse) is of primary clinical interest, but progression to a new site (a distant relapse) can also occur. The two failure types, local relapses and distant relapses, are known as competing risks; they compete because only one of them can happen first and (presumably) alter treatment or other conditions of the study.

When it comes to dealing with competing events such as distant relapses, you can either treat them as censored observations or use competing-risks software such as the new `stcrreg` command in Stata 11. In our example, treating distant relapses as censored and using standard survival analysis software is suitable for examining the biological mechanism (the cause-specific hazard) behind local cancer relapses. However, suppose you are interested in the cumulative incidence function, the probability of a local relapse before a given time t . The probability of a local relapse not only is a function of the hazard for local relapses but also is a function of the hazard for distant relapses, because distant relapses impede local relapses as first events requiring treatment. As such, treating distant relapses as censored does not allow for direct estimation of the cumulative incidence function for local relapses.

This is where specialized software, such as the new `stcrreg` command for competing-risks regression, comes into play. `stcrreg` makes it easy to model and estimate cumulative incidence in the presence of competing risks. To see `stcrreg` in action, type

```
. use http://www.stata-press.com/data/cggm3/bc_compete, clear
(Breast cancer with competing risks)

. stset time, failure(status = 1) // Local relapse is the event of interest

. stcrreg drug age, compete(status = 2) // Distant relapse is the competing event

      failure _d: status == 1
      analysis time _t: time

Iteration 0:  log pseudolikelihood = -794.80611
Iteration 1:  log pseudolikelihood = -794.79099
Iteration 2:  log pseudolikelihood = -794.79099

Competing-risks regression              No. of obs      =      423
                                         No. of subjects =      423
Failure event : status == 1             No. failed      =      135
Competing event: status == 2           No. competing   =       78
                                         No. censored    =     210

                                         Wald chi2(2)    =       5.07
Log pseudolikelihood = -794.79099      Prob > chi2     =     0.0794
```

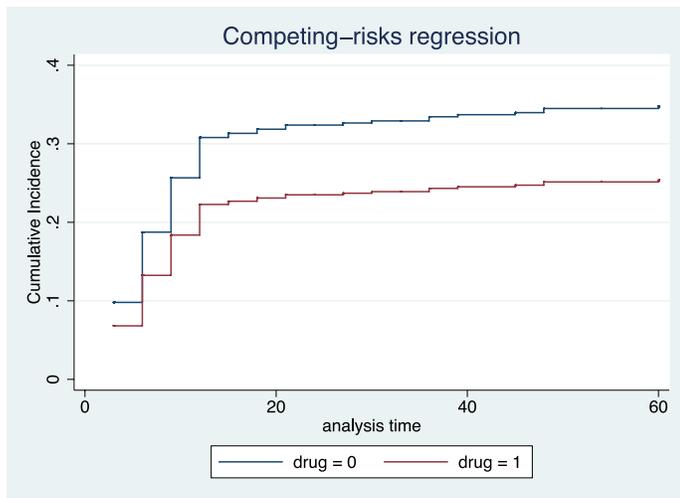
| _t | SHR | Robust Std. Err. | z | P> z | [95% Conf. Interval] | |
|------|----------|------------------|-------|-------|----------------------|----------|
| drug | .6845084 | .1215754 | -2.13 | 0.033 | .483278 | .9695282 |
| age | 1.006517 | .0107129 | 0.61 | 0.542 | .9857373 | 1.027734 |

Continued on p. 2

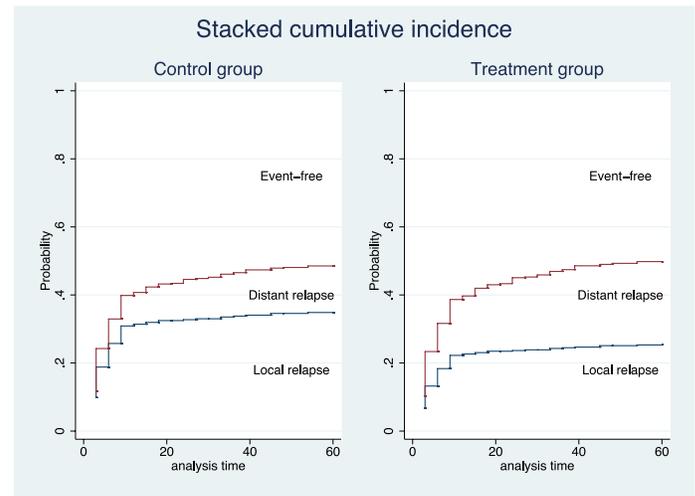
We modeled the cumulative incidence of local relapses (`status = 1`) in the presence of competing distant relapses (`status = 2`) with drug treatment and age as covariates. The parameter estimates under the SHR column are subhazard ratios and measure the effects of covariates on the cumulative incidence of local relapses. The subhazard for drug treatment (`drug = 1`) for a given age is 68% of the subhazard for the control group (`drug = 0`).

In other words, those subjects in the drug treatment group have a reduced incidence of local cancer relapse, and we can use `stcurve` after `stcrreg` to visualize that reduction, holding age at its mean value.

```
. stcurve, cif at1(drug = 0) at2(drug = 1)
```



The analyses you can perform with `stcrreg` and its postestimation commands can get quite sophisticated. For example, you can draw stacked cumulative incidence curves, which are common in the literature on competing risks.



For detailed instructions on constructing stacked cumulative incidence curves in Stata 11 and for more information on `stcrreg` and competing-risks analysis, see chapter 17 of Cleves et al. (2010, *An Introduction to Survival Analysis Using Stata, Third Edition* [Stata Press]).

— Roberto G. Gutierrez,
Director of Statistics, StataCorp

Stata now has an official blog

You are invited to visit the new Stata blog, *Not Elsewhere Classified (NEC)*, at <http://blog.stata.com>.

At *NEC* we will keep you up to date about all things related to Stata. That includes product announcements from StataCorp and from others, service announcements such as on-site and public training, timely tips and comments, and news and advice related to the use of Stata.

The articles appearing in the blog are individually signed. *NEC* is written by the same people who develop Stata, support Stata, and sell Stata. *NEC* is informal but useful, and even entertaining.

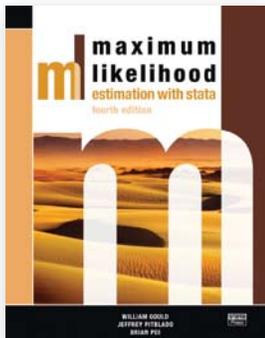
There are other fine blogs about Stata—take a look at the **Blogroll** on the right of the *NEC* page. Also on the right, check out the **Links** and especially **Links to others**. You can find resources for learning Stata, resources for adding new features to Stata, Stata examples and datasets, and sites that provide utilities and services to enhance the use of Stata. There's also a section on statistical (as opposed to Stata) resources, which will take you to other statistical software providers, statistical book publishers, statistical journals, and statistical organizations.

NEC is just getting started, but it already includes a fine article on multiprocessor computers and Stata, which was written by one of the principal developers of Stata/MP. Do be sure to visit and visit often.



New from Stata Press

Maximum Likelihood Estimation with Stata, Fourth Edition



Authors: William Gould, Jeffrey Pitblado, and Brian Poi
 Publisher: Stata Press
 Copyright: 2010
 Pages: 352; paperback
 ISBN-10: 1-59718-078-5
 ISBN-13: 978-1-59718-078-8
 Price: \$54.00

Maximum Likelihood Estimation with Stata, Fourth Edition, is the essential reference and guide for researchers in all disciplines who wish to write maximum likelihood (ML) estimators in Stata. Beyond providing comprehensive coverage of Stata's **m1** command for writing maximum-likelihood estimators, the book presents an overview of the underpinnings of maximum likelihood and how to think about ML estimation.

The book shows you how to take full advantage of the **m1** command's noteworthy features:

- linear constraints
- four optimization algorithms (Newton–Raphson, DFP, BFGS, and BHHH)
- observed information matrix (OIM) variance estimator
- outer product of gradients (OPG) variance estimator
- Huber/White/sandwich robust variance estimator
- cluster–robust variance estimator
- complete and automatic support for survey data analysis
- direct support of evaluator functions written in Mata

When appropriate options are used, many of these features are provided automatically by **m1** and require no special programming or intervention by the researcher writing the estimator.

The fourth edition has been updated to include new features introduced in Stata 11. Such features include new methods for handling scores, more consistent arguments for likelihood-evaluator programs, a general quadratic form evaluator for problems like nonlinear least squares and generalized methods of moments (GMM), and support for likelihood evaluators written in Mata (Stata's matrix programming language). The authors illustrate how to write your estimation command so that it fully supports factor-variable notation and the **svy** prefix for estimation with survey data. They have also restructured the chapters that introduce **m1** in a way that allows you to begin working with **m1** faster. This edition is essential for anyone using Stata 11.

In the final chapter, the authors illustrate the major steps required to get from log-likelihood function to fully operational estimation command. This is done using several different models: logit and probit, linear regression, Weibull regression, the Cox proportional hazards model, random-effects

regression, and seemingly unrelated regression.

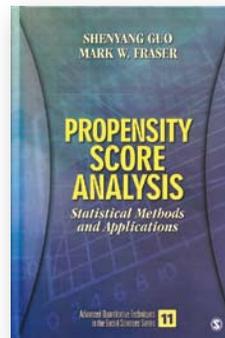
The authors provide extensive advice for developing your own estimation commands. With a little care and the help of this book, users will be able to write their own estimation commands—commands that look and behave just like the official estimation commands in Stata.

Whether you want to fit a special ML estimator for your own research or wish to write a general-purpose ML estimator for others to use, you need this book.

You can find the table of contents and online ordering information at www.stata-press.com/books/ml4.html.

New from the Stata Bookstore

Propensity Score Analysis: Statistical Methods and Applications



Authors: Shenyang Guo and Mark W. Fraser
 Publisher: Sage
 Copyright: 2010
 Pages: 370; hardback
 ISBN-10: 1-4129-5356-1
 ISBN-13: 978-1-4129-5356-6
 Price: \$39.75

Propensity Score Analysis by Shenyang Guo and Mark W. Fraser is an excellent book on estimating treatment effects from observational data. Researchers and graduate students interested in the analysis of observational data will find this book invaluable. This book is the first to provide step-by-step instructions for using Stata to estimate treatment effects by propensity-score analysis. The authors also cover Heckman's sample-selection estimator, nearest-neighbor matching estimators, propensity-score matching, and propensity-score nonparametric regression estimators.

After providing a useful history of these techniques, the authors discuss the issues involved in using observational data. The authors have consolidated the notation and theory found in the propensity-score literature and have carried this notation throughout the book. This book introduces the philosophy and methods of matching estimation in an approachable manner, even for those unfamiliar with the subject.

Readers embark on an interesting journey from the early formulation of the Neyman–Rubin counterfactual framework to a discussion of current methods and research issues. These ideas are discussed at theoretical and applied levels. The authors formally describe the assumptions underlying the analyses in a way that provides insight into the issues related to applying these techniques to practical problems.

Propensity-score analysis has its roots in both economics and statistics.

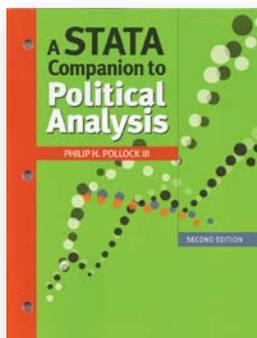
Continued on p. 4

Economists will find the discussion of Heckman's sample-selection estimator familiar, and statisticians will find the Rosenbaum and Rubin counterfactual framework familiar. The authors balance these two frameworks admirably and develop a unified presentation of the underlying issues.

While this book makes an excellent text for a graduate-level course on the analysis of observational data, it is readily accessible to researchers using observational data to estimate treatment effects. This book should be a required read for students in economics and statistics.

You can find the table of contents and online ordering information at www.stata.com/bookstore/psa.html.

A Stata Companion to Political Analysis, Second Edition



Author: Philip H. Pollock III
 Publisher: CQ Press
 Copyright: 2010
 Pages: 237; paperback
 ISBN-10: 1-60871-671-6
 ISBN-13: 978-1-60871-671-5
 Price: \$44.50

The second edition of Philip Pollock's *A Stata Companion to Political Analysis* can guide you whether you are taking your first political-science course or teaching one. The new edition has been updated for Stata 10 and Stata 11 and makes extensive use of the Graph Editor. Also new are line graphs, which are often clearer and more effective than bar graphs.

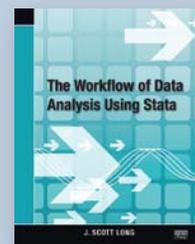
Each chapter is a tutorial with a rich set of exercises. The tutorials will teach you how to control Stata with either the command line or the graphical user interface—whichever mode fits the task. The book also surveys the statistical methods that professional political scientists use; the treatment of research methods deftly incorporates data management, graphical analysis, and statistics into the political-science domain. The thorough examples show how to complete each task with Stata while giving firsthand experience in political research.

You can find the table of contents and online ordering information at www.stata.com/bookstore/scpa.html.

Additional featured titles

The Workflow of Data Analysis Using Stata

Author: J. Scott Long
 Price: \$52.00



This text is an essential productivity tool for data analysts. Aimed at anyone who analyzes data, this book presents an effective strategy for designing and doing data-analytic projects.

Applied Survey Data Analysis

Authors: Steven G. Heeringa, Brady T. West, and Patricia A. Berglund
 Price: \$68.00



Applied Survey Data Analysis is an intermediate-level, example-driven treatment of current methods for complex survey data. It will appeal to researchers of all disciplines who work with survey data and have basic knowledge of applied statistical methodology for standard (nonsurvey) data.

Using Stata for Principles of Econometrics, Third Edition

Authors: Lee C. Adkins and R. Carter Hill
 Price: \$59.00



This text is a companion to the introductory econometrics textbook *Principles of Econometrics, Third Edition*. It shows how to use Stata to reproduce examples from the main textbook and how to interpret the output.

Veterinary Epidemiologic Research, Second Edition

Authors: Ian Dohoo, Wayne Martin, and Henrik Stryhn
 Price: \$119.00



Veterinary Epidemiologic Research, Second Edition, is a new edition of a popular graduate-level text on veterinary epidemiology. Although many of the examples relate to veterinary epidemiology, the principles apply equally to human epidemiology (except some of the diseases may not be familiar).

For more information or to order these and other titles, visit www.stata.com/bookstore/.

Public training

Using Stata Effectively: Data Management, Analysis, and Graphics Fundamentals

Instructor: Bill Rising, StataCorp's Director of Educational Services

| Dates | Location | Cost |
|--------------------|-------------------|-------|
| February 8–9, 2011 | San Diego, CA | \$950 |
| March 8–9, 2011 | New York City, NY | \$950 |
| April 6–7, 2011 | Washington, DC | \$950 |
| July 12–13, 2011 | Chicago, IL | \$950 |
| August 4–5, 2011 | Miami, FL | \$950 |

Become intimately familiar with all three components of Stata: data management, analysis, and graphics. This two-day course is aimed at new Stata users and at those who want to optimize their workflow and learn tips for efficient day-to-day usage of Stata. Upon completion of the course, you will be able to use Stata efficiently for basic analyses and graphics. You will be able to do this in a reproducible manner, making collaborative changes and follow-up analyses much simpler. You also will be able to make your datasets self-explanatory to your co-workers and to your future self.

Whether you currently own Stata 11 or you are considering an upgrade or a new purchase, this course will unquestionably make you more proficient with Stata's wide-ranging capabilities.

Course topics

- Stata basics
 - › Keeping organized
 - › Knowing how Stata treats data
 - › Using dialog boxes efficiently
 - › Using the Command window
 - › Saving time and effort while working
- Data management
 - › Reading in datasets of various standard formats, such as those from spreadsheets or databases
 - › Labeling variables and setting up encoded variables
 - › Generating new variables in an efficient fashion, including leading, lagging, generating statistics within groups, and working across variables
 - › Combining datasets by adding observations and by adding variables
 - › Reshaping datasets for repeated measurements

- Workflow
 - › Using menus and the Command window to work quickly
 - › Setting up Stata to one's liking
 - › Keeping complete records of what is done inside Stata
 - › Creating reproducible analyses, which are completely documented
 - › Finding, installing, and removing user-written extensions to Stata
 - › Customizing how Stata starts up and where it looks for files
- Analysis
 - › Using basic statistical commands
 - › Reusing results of Stata commands
 - › Using common postestimation commands, such as those that test hypotheses about linear or nonlinear combinations of coefficients, generate fitted values, or look at marginal effects
 - › Working with interactions and factor variables
- Graphics
 - › Making common, simple graphs
 - › Building up complex graphs
 - › Using the Graph Editor

For more information or to enroll, visit www.stata.com/training/eff_stata.html.

Enrollment is limited. Computers with Stata 11 installed are provided at all public training sessions. A continental breakfast, lunch, and an afternoon snack will also be provided. All training courses run from 8:30 AM to 4:30 PM each day. Participants are encouraged to bring a USB flash drive to all public training sessions; this is the safest and simplest way to save your work from the session.

Sign up to receive notifications about future public trainings at www.stata.com/alerts/.

Stata Conference



CHICAGO
2011
July 14–15

Call for presentations

The 2011 Stata Conference will be held at the University of Chicago Graduate School of Business's Gleacher Center. The Gleacher Center is located downtown on the bank of the Chicago River, just steps from Michigan Avenue and within walking distance of most downtown attractions.

The Stata Conference is enjoyable and rewarding for Stata users at all levels and from all disciplines. This year's program will consist of a mixture of user presentations, longer talks by invited presenters, and talks by StataCorp developers. In addition, the program will include the ever-popular "Wishes and grumbles" session in which users have an opportunity to share their comments and suggestions directly with developers from StataCorp.

All users are encouraged to submit abstracts for possible presentations. Presentations on any Stata-related topic will be considered, including (but not limited to) the following:

- new user-written commands, including commands for modeling and estimation, graphical analysis, data management, or reporting
- use or evaluation of existing Stata commands
- methods for teaching statistics with Stata or teaching the use of Stata
- case studies of Stata use in novel areas or applications
- surveys or critiques of Stata facilities in specific fields
- comparisons of Stata to other software or use of Stata together with other software

Each user presentation should be either 15 or 25 minutes long, and should be followed by 5 minutes for questions. Longer talks will be considered at the discretion of the scientific committee.

| | |
|-------------------------------|---|
| Dates | July 14–15, 2011 |
| Venue | Gleacher Center The University of Chicago Booth School of Business 450 North Cityfront Plaza Drive Chicago, IL 60611 312.464.8787 www.gleachercenter.com |
| Chair of Scientific Committee | Philip Schumm The University of Chicago Department of Health Studies |
| Details and registration | www.stata.com/chicago11 Submission deadline March 14, 2011 |

Submission guidelines

Please submit an abstract of no more than 200 words (ASCII text, no math symbols) by using the web submission form at <http://repec.org/chi11/chi11.php>. All abstracts must be received by **March 14, 2011**. Please include a short, informative title and indicate whether you wish to be considered for a short (15-minute) or long (25-minute) presentation. In addition, if your presentation has multiple authors, please identify the presenter. The conference registration fee will be waived for the presenter.

Presenters will be asked to provide the organizers with electronic materials (a copy of the presentation and any programs or datasets, where applicable) so that the materials can be posted on the StataCorp website and in the Stata Users Group RePEc archive.

2011 Mexican Stata Users Group meeting

Announcement and call for presentations

| | |
|----------------------|---|
| Date: | May 12, 2011 |
| Venue: | Institute for Economic Research, National Autonomous University of Mexico, Circuito Mario de la Cueva, Ciudad de la Investigación en Humanidades, Ciudad Universitaria, C.P.04510, México, D.F. |
| Cost: | To be announced (There will be reduced rates for presenters) |
| Submission deadline: | March 19, 2011 |
| Details: | www.stata.com/meeting/mexico11/ |

The Mexican Stata Users Group meeting is a one-day international conference about the use of Stata in a wide breadth of fields and environments, mixing theory and practice. The bulk of the conference is made up of selected submitted presentations. Together with the keynote address and a featured presentation by a member of StataCorp's technical staff, these sessions provide fertile ground for learning about statistics and Stata. All users are encouraged to submit abstracts for possible presentations.

Presentation topics might include (but are not limited to) the following:

- new user-written commands, including commands for modeling and estimation, graphical analysis, data management, or reporting
- use or evaluation of existing Stata commands
- methods for teaching statistics with Stata or teaching the use of Stata
- case studies of Stata use in novel areas or applications
- surveys or critiques of Stata facilities in specific fields
- comparisons of Stata to other software or use of Stata together with other software

We are delighted that Professor A. Colin Cameron (University of California–Davis) has accepted our invitation to present the keynote address. Professor Cameron is a world-known leader in the fields of econometrics and health economics. He is an author (jointly with Pravin K. Trivedi) of well-known textbooks in econometrics such as *Microeconometrics: Methods and Applications*, *Regression Analysis of Count Data*, and *Microeconometrics Using Stata*. He has authored more than 25 papers in highly regarded journals, including *Quarterly Journal of Economics*, *Journal of Public Economics*, *Journal of Econometrics*, *Review of Economic Studies*, *Journal of Business and Economic Statistics*, and *Journal of Applied Econometrics*, among others.

The meeting will include the usual “Wishes and grumbles” session, at which you may share your thoughts with Stata developers. After the meeting, there will be an optional informal meal (for an additional cost) at a Mexico City restaurant for further discussion of all things Stata.

The meeting is organized by MultiON Consulting SA de CV, distributor of Stata for Mexico and Central America. The meeting is open to all interested, and StataCorp will be represented.

Submissions

The deadline for submissions is **March 19, 2011**. The program will be announced by the end of the second week of April. Please send submissions by email to mexico11@stata.com. Indicate the length of the talk you wish to give:

- a 15-minute talk (followed by a 5-minute discussion)
- a 25-minute talk (followed by a 5-minute discussion)
- a longer review or a tutorial (about an hour long)

Presentations may be given in English or in Spanish, depending upon the preference of the speaker.

Contact the scientific committee if you have questions regarding the suitability of a potential contribution. For all other matters, contact the logistics organizer.

Registration

MultiON Consulting generously sponsors registration fee waivers for presentations (one fee waiver per presentation, regardless of the number of authors). Registration details will be available soon.

Organizers

Scientific committee:

Alfonso Miranda (chair). Institute of Education, University of London.
Email: A.Miranda@ioe.ac.uk

Armando Sánchez Vargas. Institute for Economic Research, National Autonomous University of Mexico.
Email: sva@economia.unam.mx

Graciela Teruel Belismelis. Economics Department, Iberoamerican University.
Email: graciela.teruel@s2.uia.mx

Logistics organizer:

MultiON Consulting SA de CV, distributor of Stata in Mexico and Central America

Victoria Leon
Email: vleon@multion.com.mx

Phone: +52 (55) 5559 4050 x 160

.....



Meet with Stata developers at ASSA: January 7–9

The Allied Social Science Association (ASSA) will have its annual meeting and exposition in Denver, Colorado, January 7–9, 2011. For more information, please visit www.vanderbilt.edu/AEA/Annual_Meeting/.

David Drukker, StataCorp's Director of Econometrics, will be available at the Stata booth (#317) to answer your questions about econometrics and Stata. Also attending will be Bill Rising, Stata's Director of Educational Services. Stop by and visit with the people who develop and support the software.

Stata is now on Facebook and Twitter

You can now keep up with the latest Stata-related happenings by following us on Facebook and Twitter. Click on the Facebook and Twitter icons on www.stata.com and connect with us today.



Find us on Facebook.



Follow us on Twitter.

Upcoming NetCourses

Enroll by visiting www.stata.com/netcourse/.

NC101: Introduction to Stata

This course is an introduction to using Stata interactively.

Dates: January 21–March 4, 2011

Enrollment deadline: January 20, 2011

Price: \$95

Details: www.stata.com/netcourse/nc101.html

NC151: Introduction to Stata Programming

This course is an introduction to Stata programming dealing with what most statistical software users mean by programming, namely, the careful performance of reproducible analyses.

Dates: January 21–March 4, 2011

Enrollment deadline: January 20, 2011

Price: \$125

Details: www.stata.com/netcourse/nc151.html

NC152: Advanced Stata Programming

This course teaches you how to create and debug new commands that are indistinguishable from those of official Stata. It is assumed that you know why and when to program and, to some extent, how. You will learn how to parse both standard and nonstandard Stata syntax by using the intuitive **syntax** command, how to manage and process saved results, how to process by-groups, and more.

Dates: January 21–March 11, 2011

Enrollment deadline: January 20, 2011

Price: \$150

Details: www.stata.com/netcourse/nc152.html

NC461: Introduction to Univariate Time Series with Stata

This course introduces univariate time-series analysis, emphasizing the practical aspects most needed by practitioners and applied researchers. The course is written to appeal to a broad array of users, including economists, forecasters, financial analysts, managers, and anyone who encounters time-series data.

Dates: January 21–March 11, 2011

Enrollment deadline: January 20, 2011

Price: \$295

Details: www.stata.com/netcourse/nc461.html

For a NetCourse® that you can take at your own pace and according to your own schedule, consider enrolling in a NetCourseNow™. To learn more or to enroll, visit www.stata.com/netcourse/ncnow.html.

Contact us

StataCorp
4905 Lakeway Dr.
College Station, TX 77845
USA

Phone 979-696-4600
Fax 979-696-4601
Email service@stata.com
Web www.stata.com

Please include your Stata serial number with all correspondence.



Copyright 2010 by StataCorp LP.

To locate a Stata international distributor near you, visit www.stata.com/worldwide/.