

# THE STATA NEWS

July/August/September 2010

Vol 25 No 3

## Stata makes a difference

Get an inside look into how Stata is used by the Health Policy Institute of Ohio.

p. 1

## In the spotlight

Learn about Stata 11's **margins** command.

p. 2

## New from Stata Press

Read about the latest three books published by Stata Press.

p. 6

## The Stata Journal: Have you subscribed?

p. 9

## 2010 Italian Stata Users Group meeting

Make plans to attend this meeting in Bologna, Italy.

p. 12

## Also in this issue

New from the Stata Bookstore ..... 8

Additional featured titles ..... 9

Public training courses ..... 10

Visit us at APHA 2010 ..... 11

Upcoming NetCourses ..... 11

*The Stata News*

Executive Editor: Karen Strope

Production Supervisor: Annette Fett

## Stata makes a difference at the Health Policy Institute of Ohio

The Health Policy Institute of Ohio (HPIO) is an independent, nonpartisan, statewide center that fosters sound health policy within the state by forecasting health trends, analyzing key health issues, and communicating current research to policymakers, state agencies, and other decision makers. HPIO promotes and facilitates health policy research among research centers, universities, and other organizations. It identifies gaps in health policy research and data; designs studies; leads the development of a statewide health policy research agenda; promotes collaboration among researchers; develops research projects to address health problems; and, as necessary, undertakes research directly. In addition, HPIO assists researchers in presenting important findings and serves as a network facilitator among health researchers and practitioners.

***“HPIO primarily depends upon Stata as its analytical weapon of choice.”***

Examples of research topics include the uninsured and underinsured, health systems capacities, health safety net capacities, determinants of health, health disparities, health care reform, public health systems, family violence prevention, poverty, community health status, health information technologies, and behavioral health. Findings from various research topics are readily available at HPIO's website, in public presentations, and by request. For all this work, HPIO primarily depends upon Stata as its analytical weapon of choice.

The main reasons HPIO uses Stata are its highly intuitive interface, its support for complex survey data, its epidemiology commands, and its support for various types of biostatistical, social-science, and econometric analyses. Some examples of how Stata has helped HPIO in its analytic needs are in analysis of the 2008 Ohio Family Health Survey (OFHS)—a complex, dual-framed survey of health systems, behaviors, and demographics of 50,944 Ohio adults—and in the Medicaid Atlas Project, which uses approximately 2,200,000 cases to examine Medicaid use in Ohio's 88 counties.

For both projects, the expanded datasets are very large—the OFHS is approximately 300 megabytes and the Medicaid dataset is approximately 1.3 gigabytes. In the 1990s, analyzing such datasets was difficult because of software and equipment limitations. With the prerequisite of needing to allocate a large amount of memory at program startup, Stata/MP 11.1 easily handles the analysis of such datasets. For the OFHS, programming code to model the uninsured in Ohio is easily done using ado-files and do-files. The OFHS is the main source of Ohio-specific population-based health system information provided to the state's legislators, agency heads, and health system stakeholders. Analysis of the OFHS provides Ohioans with information relating to how federal health reform will affect Ohio. Areas of interest include:

- the characteristics of Ohio's 1.3 million uninsured;
- the degree to which Ohio's uninsured will be eligible for various coverage expansions, including criteria of income, duration of being uninsured, family composition, chronic and extreme health conditions, etc.;

*Continued on p. 2*

- the potential costs of covering the newly insurance eligible in Ohio; and
- select issues, such as potential crowd-out, which occurs when currently insured individuals move their health insurance to a government-sponsored program.

Examination and modeling of these types of issues relies on Stata's survey commands, which allow us to incorporate the design characteristics of the survey.

The Medicaid Atlas Project analyzes Medicaid billing information to determine issues such as total Medicaid use per county and the number of physicians serving Medicaid patients in each county. The project also uses this data to monitor expenditures per Medicaid utilization category and to project the growth in average expenditure per category. Additionally, procedures are used to model relative-risk profiles of Medicaid enrollees versus nonenrollees and to model relative-risk profiles of Medicaid managed-care enrollees versus fee-for-service

enrollees. Because health policy stakeholders are large contributors of health services in the state, determining the overall population-based health impact of Ohio Medicaid is very important to them. For example, Ohio has experienced a prolonged economic downturn, having lost over 560,000 jobs since 2000.

**“The main reasons HPIO uses Stata are its highly intuitive interface, its support for complex survey data, its epidemiology commands, and its support for various types of biostatistical, social-science, and econometric analyses.”**

During this period, because of the State Children's Health Insurance Program (SCHIP), Ohio's rate of uninsured children actually decreased while the adult rate increased. Using internal data from Medicaid mixed with state-specific external data from surveys allowed us to estimate the risk buffering of children's access to health care that is attributable to Medicaid in hard economic times.

Finally, Stata's web-enabled interactive search capacities are often indispensable for figuring out complex data setup and analysis issues.

The Stata community, including researchers at universities, research institutes, and government agencies, is an excellent resource for figuring out problems. As an example, HPIO is participating in a project to test a concept for examining simulated benefit models for dual-frame surveys—surveys where samples are drawn independently from two overlapping sampling frames to cover the population of interest (e.g., respondents to a survey of households with landline telephones and households with both cell phones and landline telephones). The research team intends to develop a program in Stata that will enable survey researchers to determine whether to develop dual-frame or single-frame surveys rather than making sampling decisions based upon convenience.

In summary, Stata allows the Health Policy Institute of Ohio and its partners to keep an analytical edge on very complex health issues. The program is robust enough to handle very large datasets, fast enough in its MP versions to use high-end computers, and thorough enough to address epidemiology, social-science, and econometric analyses.

— Timothy R. Sahr,

Director of Research, The Health Policy Institute of Ohio

Director of Research, The Ohio Colleges of Medicine Government Resource Center

## In the spotlight: Margins of predicted outcomes

Stata 11's new `margins` command combines and extends the features available in the older `adjust` and `mfex` commands. The `margins` command can calculate

- Adjusted predictions
- Predictive margins
- Marginal effects
- Average marginal effects

Adjusted predictions were previously available via `adjust`, and `mfex` reported marginal effects. `margins` provides a unified syntax to obtain those statistics as well as predictive margins and average marginal effects,

which were previously unavailable without significant effort.

`margins` takes advantage of the new factor-variables features in Stata 11. In this spotlight, I show how `margins` can be used to compute marginal predictions from a linear regression that includes a quadratic term.

The venerable `auto.dta` dataset contains vehicle prices in dollars in the variable `price`. We first generate a new variable, `kprice`, that contains prices in thousands of dollars. This not only enhances the output by rescaling the corresponding regression coefficients, it also makes interpreting the numerical results easier.

```
. sysuse auto
(1978 Automobile Data)

. sum price
  Variable |      Obs      Mean   Std. Dev.   Min   Max
-----+-----
   price |       74   6165.257   2949.496   3291 15906

. gen kprice = price/1000
```

Now let's model `mpg` on `kprice` and its square, along with a few other variables.

```
. regress mpg c.kprice##c.kprice i.foreign i.rep78 turn trunk headroom
```

Source	SS	df	MS	Number of obs = 69		
Model	1606.17185	10	160.617185	F( 10, 58)	=	12.69
Residual	734.031053	58	12.6557078	Prob > F	=	0.0000
-----				R-squared	=	0.6863
Total	2340.2029	68	34.4147485	Adj R-squared	=	0.6323
-----				Root MSE	=	3.5575

mpg	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
kprice	-3.022153	1.072402	-2.82	0.007	-5.168798	-.8755078
c.kprice#						
c.kprice	.1447453	.0595132	2.43	0.018	.0256167	.263874
1.foreign	-.1969568	1.536553	-0.13	0.898	-3.272702	2.878788
rep78						
2	1.248398	2.989061	0.42	0.678	-4.734857	7.231653
3	.3856603	2.818224	0.14	0.892	-5.255628	6.026949
4	1.760669	2.910874	0.60	0.548	-4.066078	7.587416
5	4.635368	3.072096	1.51	0.137	-1.5141	10.78484
turn	-.6624524	.1685471	-3.93	0.000	-.9998362	-.3250686
trunk	-.0158062	.1776647	-0.09	0.929	-.3714406	.3398283
headroom	-.5875156	.7168267	-0.82	0.416	-2.0224	.8473689
_cons	60.08219	7.169182	8.38	0.000	45.73151	74.43286

To get `kprice` in the model in its linear and squared form, we typed

```
c.kprice##c.kprice
```

The `##` operator is really a shortcut notation for two operations. The first `#` tells Stata that this term is an interaction; the second `#` tells Stata to include the associated variables in addition to their interaction. Thus the above specification expands to

```
kprice c.kprice#c.kprice
```

The `c.` operator tells Stata that `kprice` is to be treated as a continuous variable. This overrides the default assumption that variables associated with the `#` operator are factor variables that should enter the model as indicator variables for each level of the variable.

The estimated regression coefficients on `kprice` and its square result in a parabola that is concave up with a minimum at approximately \$10,400. Entering our coefficient values into the expression for the minimum of a concave-up parabola, which is obtained by setting to zero the derivative of `mpg` with respect to `kprice` and by solving for `kprice`, we have

```
. di -_b[kprice]/(2*_b[kprice#kprice])
10.439552
```

Now we can use `margins` to estimate the expected value of `mpg` at various values of `kprice`. Because `kprice` ranges from about 3 to 15 (in thousands of dollars), let's predict `mpg` at all the integer values in this range. To do this, we will use the `at()` option and assign `kprice` the integer values from 3 to 15.

For each specified value, `margins` will call `predict` to generate a variable with the linear prediction and take the average of the predictions to get the predictive margin.

Here are the results from our call to `margins`:

```
. margins, at(kprice=(3/15)) vsquish post
```

```
Predictive margins                                Number of obs   =           69
Model VCE      : OLS
```

```
Expression   : Linear prediction, predict()
```

```
1._at       : kprice           =           3
2._at       : kprice           =           4
3._at       : kprice           =           5
4._at       : kprice           =           6
5._at       : kprice           =           7
6._at       : kprice           =           8
7._at       : kprice           =           9
8._at       : kprice           =          10
9._at       : kprice           =          11
10._at      : kprice           =          12
11._at      : kprice           =          13
12._at      : kprice           =          14
13._at      : kprice           =          15
```

	Delta-method					[95% Conf. Interval]	
	Margin	Std. Err.	z	P> z			
_at							
1	25.42282	1.313698	19.35	0.000	22.84802	27.99762	
2	23.41388	.7385698	31.70	0.000	21.96631	24.86145	
3	21.69444	.4723934	45.92	0.000	20.76856	22.62031	
4	20.26448	.620569	32.65	0.000	19.04819	21.48078	
5	19.12402	.8617617	22.19	0.000	17.435	20.81304	
6	18.27305	1.045006	17.49	0.000	16.22487	20.32122	
7	17.71157	1.146571	15.45	0.000	15.46433	19.9588	
8	17.43957	1.173328	14.86	0.000	15.13989	19.73926	
9	17.45707	1.153073	15.14	0.000	15.19709	19.71705	
10	17.76406	1.144364	15.52	0.000	15.52115	20.00698	
11	18.36054	1.240161	14.80	0.000	15.92987	20.79121	
12	19.24651	1.523544	12.63	0.000	16.26042	22.23261	
13	20.42198	2.012575	10.15	0.000	16.4774	24.36655	

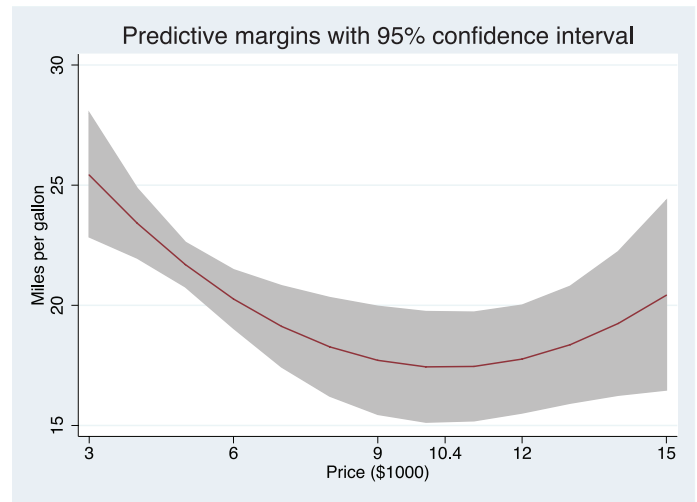
The `vsquish` option suppresses the extra vertical space in the legend for the `at()` option. We `posted` the `margins` results so that we could use Roger Newson's convenient `parmest` command to put the `margins` results in a dataset. (You can obtain `parmest` by typing `ssc install parmest` in Stata.) We want to graph the predictive margins, and `parmest` gets us most of the way there. All we need for our graph is a `kprice` variable with the values associated with the predictive margins. We will also give our graph some nice titles, nice labels, and a tick mark on the  $x$  axis at the minimum `mpg` that we computed earlier.

```
. parmest, norestore
. gen kprice = 3 + _n - 1
. label var kprice "Price (\$1000)"
. twoway rarea max95 min95 kprice, pstyle(ci) ||
    line estimate kprice,
    ytitle(Miles per gallon)
    title(Predictive margins with 95% confidence interval)
    xlabel(3(3)15 10.4) legend(off)
```

To the right is the graph of the predictive margins with a shaded 95% confidence region.

We can also use `margins` to compute the effect of `kprice` on the prediction of `mpg`. This is what economists refer to as the average marginal effect. We simply rerun our analysis with the addition of `dydx(kprice)` in the call to `margins`. This will compute the slope of the predictive margin curve with respect to `kprice`.

Below are the marginal effects of `kprice` at our list of values. Before running this `margins` command, we would need to restore the auto dataset and rerun our regression.



```
. margins, dydx(kprice) at(kprice=(3/15)) vsquish post
```

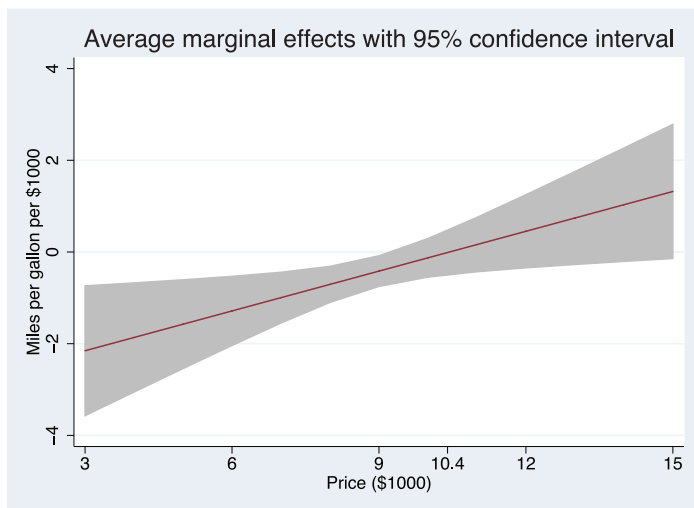
```
Average marginal effects          Number of obs   =          69
Model VCE      : OLS
```

```
Expression      : Linear prediction, predict()
dy/dx w.r.t.    : kprice
1._at           : kprice           =          3
2._at           : kprice           =          4
3._at           : kprice           =          5
4._at           : kprice           =          6
5._at           : kprice           =          7
6._at           : kprice           =          8
7._at           : kprice           =          9
8._at           : kprice           =         10
9._at           : kprice           =         11
10._at          : kprice           =         12
11._at          : kprice           =         13
12._at          : kprice           =         14
13._at          : kprice           =         15
```

		Delta-method				
		dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]
kprice						
	_at					
	1	-2.153681	.7220297	-2.98	0.003	-3.568833 - .7385285
	2	-1.86419	.6069789	-3.07	0.002	-3.053847 - .6745334
	3	-1.574699	.4938161	-3.19	0.001	-2.542561 - .6068377
	4	-1.285209	.3842133	-3.35	0.001	-2.038253 - .5321645
	5	-.9957181	.2823474	-3.53	0.000	-1.549109 - .4423275
	6	-.7062274	.2003865	-3.52	0.000	-1.098978 - .3134771
	7	-.4167368	.1700708	-2.45	0.014	-.7500693 - .0834042
	8	-.1272461	.2145413	-0.59	0.553	-.5477392 .2932471
	9	.1622446	.3024342	0.54	0.592	-.4305155 .7550047
	10	.4517353	.4064966	1.11	0.266	-.3449834 1.248454
	11	.741226	.5170562	1.43	0.152	-.2721856 1.754638
	12	1.030717	.6307054	1.63	0.102	-.2054433 2.266877
	13	1.320207	.7460336	1.77	0.077	-.1419917 2.782406



Here is a graph of the average marginal effects.

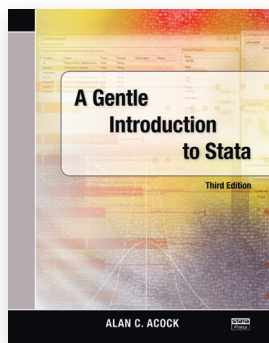


These are just two basic analyses that you can perform with `margins`. There are many others. Read [R] `margins` for more ideas.

— Jeff Pitblado, Associate Director, Stata Statistical Software

## New from Stata Press

### A Gentle Introduction to Stata, Third Edition



Author: Alan C. Acock  
 Publisher: Stata Press  
 Copyright: 2010  
 Pages: 393; paperback  
 ISBN-10: 1-59718-075-0  
 ISBN-13: 978-1-59718-075-7  
 Price: \$48.00

Alan C. Acock's *A Gentle Introduction to Stata, Third Edition* is aimed at new Stata users who want to become proficient in Stata. After reading this introductory text, new users not only will be able to use Stata well but also will learn new aspects of Stata easily.

Acock assumes that the user is not familiar with any statistical software. This assumption of a blank slate is central to the structure and contents of the book. Acock starts with the basics; for example, the portion of the book that deals with data management begins with a careful and detailed example of turning survey data on paper into a Stata-ready dataset on the computer. When explaining how to go about basic exploratory statistical procedures, Acock includes notes that will help the reader develop good work habits. This mixture of explaining good Stata habits and good statistical habits continues throughout the book.

Acock is quite careful to teach the reader all aspects of using Stata. He covers data management, good work habits (including the use of basic do-files), basic exploratory statistics (including graphical displays), and analyses using the standard array of basic statistical tools (correlation, linear and logistic regression, and parametric and nonparametric tests of location and dispersion). Acock teaches Stata commands by using the menus and dialog boxes while still stressing the value of do-files. In this way, he ensures that all types of users can build good work habits. Each chapter has exercises that the motivated reader can use to reinforce the material.

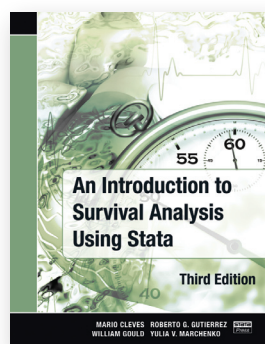
The tone of the book is friendly and conversational without ever being glib or condescending. Important asides and notes about terminology are set off in boxes, which makes the text easy to read without any convoluted twists or forward-referencing. Rather than splitting topics by their Stata implementation, Acock chose to arrange the topics as they would appear in a basic statistics textbook; graphics and postestimation are woven into the material in a natural fashion. Real datasets, such as the General Social Surveys from 2002 and 2006, are used throughout the book.

The focus of the book is especially helpful for those in psychology and the social sciences, because the presentation of basic statistical modeling is supplemented with discussions of effect sizes and standardized coefficients. Various selection criteria, such as semipartial correlations, are discussed for model selection.

The third edition of the book has been updated to reflect the new features included in Stata 11. An entire chapter is devoted to the analysis of missing data and the use of multiple-imputation methods. Factor-variable notation is introduced as an alternative to the manual creation of interaction terms. The new Variables Manager and revamped Data Editor are featured in the discussion of data management.

You can find the table of contents and online ordering information at [www.stata-press.com/books/acock3.html](http://www.stata-press.com/books/acock3.html).

### An Introduction to Survival Analysis Using Stata, Third Edition



Authors: Mario Cleves, William Gould, Roberto G. Gutierrez, and Yulia V. Marchenko  
 Publisher: Stata Press  
 Copyright: 2010  
 Pages: 412; paperback  
 ISBN-10: 1-59718-074-2  
 ISBN-13: 978-1-59718-074-0  
 Price: \$58.00

*An Introduction to Survival Analysis Using Stata, Third Edition* is the ideal tutorial for professional data analysts who want to learn survival analysis for the first time or who are well versed in survival analysis but are not as

dexterous in using Stata to analyze survival data. This text also serves as a valuable reference to those readers who already have experience using Stata's survival analysis routines.

The third edition has been updated for Stata 11, and it includes a new chapter on competing-risks analysis. This chapter describes the problems posed by competing events (events that impede the failure event of interest), and covers estimation of cause-specific hazards and cumulative incidence functions. Other enhancements include the handling of missing values by multiple imputation in Cox regression, a new-to-Stata-11 system for specifying categorical (factor) variables and their interactions, three additional diagnostic measures for Cox regression, and a more efficient syntax for obtaining predictions and diagnostics after Cox regression.

Survival analysis is a field of its own that requires specialized data management and analysis procedures. To meet this requirement, Stata provides the **st** family of commands for organizing and summarizing survival data. The authors of this text are also the authors of Stata's **st** commands.

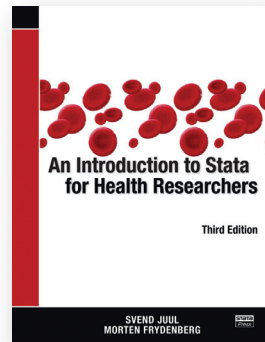
This book provides statistical theory, step-by-step procedures for analyzing survival data, an in-depth usage guide for Stata's most widely used **st** commands, and a collection of tips for using Stata to analyze survival data and to present the results. This book develops from first principles the statistical concepts unique to survival data and assumes only a knowledge of basic probability and statistics and a working knowledge of Stata.

The first three chapters of the text cover basic theoretical concepts: hazard functions, cumulative hazard functions, and their interpretations; survivor functions; hazard models; and a comparison of nonparametric, semiparametric, and parametric methodologies. Chapter 4 deals with censoring and truncation. The next three chapters cover the formatting, manipulation, **stsetting**, and error checking involved in preparing survival data for analysis using Stata's **st** analysis commands. Chapter 8 covers nonparametric methods, including the Kaplan–Meier and Nelson–Aalen estimators and the various nonparametric tests for the equality of survival experience.

Chapters 9–11 discuss Cox regression and include various examples of fitting a Cox model, obtaining predictions, interpreting results, building models, model diagnostics, and regression with survey data. The next four chapters cover parametric models, which are fit using Stata's **streg** command. These chapters include detailed derivations of all six parametric models currently supported in Stata and methods for determining which model is appropriate, as well as information on stratification, obtaining predictions, and advanced topics such as frailty models. Chapter 16 is devoted to power and sample-size calculations for survival studies. The final chapter covers survival analysis in the presence of competing risks.

You can find the table of contents and online ordering information at [www.stata-press.com/books/saus3.html](http://www.stata-press.com/books/saus3.html).

## An Introduction to Stata for Health Researchers, Third Edition



Authors: Svend Juul and Morten Frydenberg  
 Publisher: Stata Press  
 Copyright: 2010  
 Pages: 340; paperback  
 ISBN-10: 1-59718-077-7  
 ISBN-13: 978-1-59718-077-1  
 Price: \$52.00

Svend Juul and Morten Frydenberg's *An Introduction to Stata for Health Researchers, Third Edition* is distinguished in its careful attention to detail. The reader will learn how to use Stata for statistical analysis and the skills needed to make the analysis reproducible. The authors use a friendly, down-to-earth tone and include tips gained from a lifetime of collaboration and consulting. The book is based on the assumption that the reader has some basic knowledge of statistics but no knowledge of Stata.

The authors build the reader's abilities as a builder would build a house: laying a firm foundation in Stata; framing a general structure in which good work can be accomplished; adding the details that are particular to various types of statistical analyses; and finally, trimming with a thorough treatment of graphics.

Juul and Frydenberg start by teaching the reader how to communicate with Stata by demonstrating how Stata thinks about its basic building blocks. They show how Stata views data, thus allowing the reader to see the variety of possible data structures. They also show how to manipulate data to create a dataset that is well documented. When demonstrating analysis techniques, the authors show how to think of analysis in terms of estimation and postestimation.

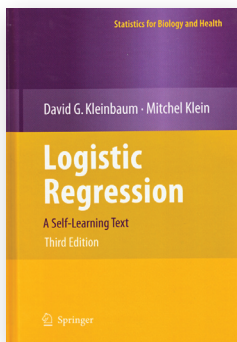
Once they introduce Stata to new users, Juul and Frydenberg fill in the details for performing analysis in Stata. They mostly demonstrate the statistical techniques that are common in biostatistics and epidemiology: case–control, matched case–control, and incidence-rate data analysis, which can be stratified or not; linear and generalized linear models, including logistic, Poisson, and binomial regression; survival analysis with proportional hazards; and classification using receiver operating characteristic curves. While presenting general estimation techniques, the authors also spend time with interactions and techniques for checking model assumptions.

While teaching Stata implementation, Juul and Frydenberg reinforce habits that allow reproducible research and graceful backtracking in case of errors. Early in the book, they introduce how to use do-files for creating sequences and log files for tracking work. At the end of the book, they introduce some useful programming techniques, such as loops and branching, that simplify repetitive tasks.

You can find the table of contents and online ordering information at [www.stata-press.com/books/ishr3.html](http://www.stata-press.com/books/ishr3.html).

## New from the Stata Bookstore

### Logistic Regression: A Self-Learning Text, Third Edition



Authors: David G. Kleinbaum and Mitchel Klein  
 Publisher: Springer  
 Copyright: 2010  
 Pages: 701; hardback  
 ISBN-10: 1-4419-1741-1  
 ISBN-13: 978-1-4419-1741-6  
 Price: \$78.00

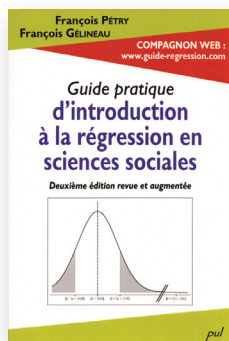
This book, the third edition of the text originally published in 1994, succeeds in demonstrating that one need not be a mathematician to fully understand the underpinnings of logistic regression in all its forms. Ideally suited for a graduate course for students in the medical sciences, the text has the look and feel of a course textbook; formulas, diagrams, and important points are set off in the side margins for emphasis. Also, each chapter contains a summary, detailed outline, objectives, practice exercises (with answers), and a chapter test.

The mathematics are kept to the most basic level, but nevertheless, because of its completeness in coverage of logistic regression, this text would be a good reference for even the most theoretical statistician. In fact, those readers already well versed in logistic regression methods will benefit from seeing advanced topics such as generalized estimating equations (GEE) explained from first principles.

The third edition features new chapters on modeling strategies (for example, dealing with influential observations), goodness of fit, and receiver operating characteristic (ROC) analysis. An appendix on computer software describes how to perform the analyses described in the text using Stata version 10.0.

You can find the table of contents and online ordering information at [www.stata.com/bookstore/lr.html](http://www.stata.com/bookstore/lr.html).

### Guide pratique d'introduction à la régression en sciences sociales



Authors: François Pétry and François Gélinau  
 Publisher: Université Laval Press  
 Copyright: 2009  
 Pages: 219; paperback  
 ISBN-10: 2-7637-8628-6  
 ISBN-13: 978-2-7637-8628-5  
 Price: \$30.00

*Guide pratique d'introduction à la régression en sciences sociales*, by François Pétry and François Gélinau, offers a brief introduction to exploratory

data analysis and applies examples to illustrate the empirical elements associated with linear and logistic regressions.

Chapters 1 and 2 review some of the aspects involved in describing data and preparing them for regression analysis. A number of graphs implemented in Stata are used to examine the behavior of the principal variables associated with the empirical examples.

In chapters 3 and 4, the authors work with data on infant mortality to illustrate the use of simple and multivariate linear regression models. In addition to presenting a series of regression outputs produced in Stata, Pétry and Gélinau describe diagnostic tools that are normally used in standard linear regression.

Chapter 5 provides a short description of the basic elements involved in time-series regression analysis. It discusses the detection of autocorrelation, along with the Prais–Winsten approach to fitting models with autocorrelated errors of order 1. The second part of the chapter describes a case in which government expenses are a function of gross domestic product (GDP). Through the case study, the authors introduce the basic aspects associated with the treatment of nonstationary series in regression analysis. Stata is used to produce most of the graphs and regression outputs contained in this chapter and those that follow.

Chapters 6 and 7 show a case that regards people's participation in general elections. In chapter 6, the authors fit a logistic model for a binary dependent variable that indicates whether the person voted. The application is extended in chapter 7 by fitting a multinomial model for a categorical dependent variable that is constructed on the basis of participation in the current and previous elections.

The empirical examples and the inclusion of graphs, regression outputs, and summary tables provide smooth reading for the introduction of the technical concepts contained in the text.

You can find the table of contents and online ordering information at [www.stata.com/bookstore/dirss.html](http://www.stata.com/bookstore/dirss.html).

### Cuadernos Metodológicos: Análisis de datos con Stata



Authors: Modesto Escobar Mercado, Enrique Fernández Macías, and Fabrizio Bernardi  
 Publisher: Centro de Investigaciones Sociológicas  
 Copyright: 2009  
 Pages: 512; paperback  
 ISBN-10: 84-7476-483-1  
 ISBN-13: 978-84-7476-483-3  
 Price: \$44.50

*Análisis de datos con Stata*, by Escobar, Fernández, and Bernardi, is an excellent resource for new and intermediate Stata users who would like to



quickly become familiar with data-management facilities that help prepare data for statistical analysis. The authors illustrate the use of Stata for descriptive statistics and regression analysis with examples that are mainly focused on social-science research but that are easy to follow for users with different backgrounds.

There are four chapters (1–3 and 5) explicitly devoted to the description of Stata tools related to loading, creating, and handling files, as well as to performing common data-management tasks. Chapter 6 provides a good presentation of a series of commands and options to produce graphs in Stata. It includes a variety of examples that allow the user to start with the simple syntax for unidimensional graphs and continue with more complex commands and options that can be used for more elaborate graphs. It also includes a section on the Graph Editor.

Chapters 4, 7, and 8 concentrate on the use of Stata for descriptive statistics and basic statistical inference. In chapters 9 and 10, the authors illustrate some of the model fit and regression diagnostic tools available for linear models. They show the command lines and output for a few different equation specifications for child mortality regressed on a couple of macro-economic variables and a few geographical dummy variables. The latter are introduced by using the factor-variables facilities added in Stata 11.

The book continues with a couple of chapters on logistic and multinomial logistic regression. After fitting models associated to individual political stance explained by certain demographic characteristics, a combination of official and user-written commands is used to compute predictions and regression diagnostic statistics.

The last two chapters use examples linked to employment to show some of the Stata commands that implement event and survey analysis.

The authors were able to combine a fairly complete introduction to Stata with theoretical readings on statistical analysis, which makes this book an outstanding tool for those who are starting their statistical analysis journey with Stata or even for those with a basic-intermediate knowledge about the software.

You can find the table of contents and online ordering information at [www.stata.com/bookstore/adcs.html](http://www.stata.com/bookstore/adcs.html).

## The Stata Journal: Have you subscribed?

The *Stata Journal* is a quarterly publication containing articles about statistics, data analysis, teaching methods, and effective use of Stata's language. The *Journal* publishes reviewed papers together with shorter notes and comments, regular columns, book reviews, and other material of interest to researchers applying statistics in a variety of disciplines.



Starting with the first quarter 2010 issue, the *Stata Journal* is indexed in Thomson Scientific's citation indexes—*Social Sciences Citation Index* and *Current Contents/Social & Behavioral Sciences*—joining the *Science Citation Index Expanded* and the *CompuMath Citation Index*, which began indexing the *Journal* in 2005.

### Subscription information (new or renewal)

Prices include electronic copy and printed copy mailed using standard shipping, unless otherwise specified. All prices are in U.S. dollars.

#### U.S. and Canada

3-year subscription	\$195
2-year subscription	\$135
1-year subscription	\$69
1-year student subscription	\$42
1-year university library subscription	\$89
1-year institutional subscription	\$195

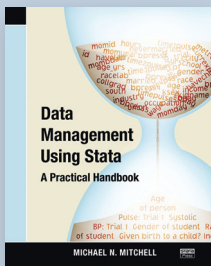
#### Elsewhere

3-year subscription	\$285
3-year subscription (electronic copy only)	\$185
2-year subscription	\$195
1-year subscription	\$99
1-year student subscription	\$69
1-year university library subscription	\$119
1-year institutional subscription	\$225

New subscriptions and individual issues may be ordered online at [www.stata-journal.com](http://www.stata-journal.com). Stata users outside the U.S. and Canada may also order subscriptions from our distributors. All prices are in U.S. dollars.

## Additional featured titles

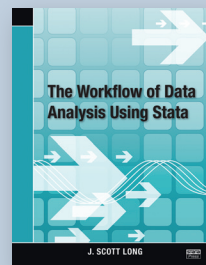
### Data Management Using Stata: A Practical Handbook



Author: Michael N. Mitchell  
 Publisher: Stata Press  
 Copyright: 2010  
 Pages: 387; paperback  
 ISBN-13: 978-1-59718-076-4  
 Price: \$48.00

Find out more at [www.stata-press.com/books/dmus.html](http://www.stata-press.com/books/dmus.html).

### The Workflow of Data Analysis Using Stata



Author: J. Scott Long  
 Publisher: Stata Press  
 Copyright: 2009  
 Pages: 379; paperback  
 ISBN-13: 978-1-59718-047-4  
 Price: \$52.00

Find out more at [www.stata-press.com/books/wdaus.html](http://www.stata-press.com/books/wdaus.html).

## Public training courses

Course	Dates	Location	Cost
Using Stata Effectively: Data Management, Analysis, and Graphics Fundamentals	October 5–6	New York, NY	\$950
	October 27–28	Boston, MA	\$950
Multilevel/Mixed Models Using Stata	October 7–8	New York, NY	\$1,295

### Multilevel/Mixed Models Using Stata

Instructor: Roberto G. Gutierrez, StataCorp's Director of Statistics and primary developer of Stata's official multilevel/mixed models features

This two-day course is an introduction to using Stata to fit multilevel/mixed models. Mixed models contain both fixed effects analogous to the coefficients in standard regression models and random effects not directly estimated, but instead summarized through the unique elements of their variance–covariance matrix. Mixed models may contain more than one level of nested random effects. Hence, these models are also referred to as multilevel or hierarchical models, particularly in the social sciences. Stata's approach to linear mixed models is to assign random effects to independent panels, where a hierarchy of nested panels can be defined for handling nested random effects.

#### Course topics

- Introduction to linear mixed models
- Random coefficients and hierarchical models
- Postestimation analysis
- Nonlinear models
- Advanced topics

For more information or to enroll, visit

[www.stata.com/training/mixed.html](http://www.stata.com/training/mixed.html).

### Stata now available from the GSA

For the convenience of U.S. federal, state, and local government users, Stata can now be purchased through the U.S. General Services Administration (GSA). New Stata/SE 11 and Stata/MP 11 software products are currently available.

Visit [www.stata.com/gsa/](http://www.stata.com/gsa/) for more details.

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

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

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
- Data management
- Workflow
- Analysis
- Graphics

For more information or to enroll, visit

[www.stata.com/training/eff\\_stata.html](http://www.stata.com/training/eff_stata.html).

Enrollment in public training courses 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.

For a complete schedule of upcoming training courses, visit

[www.stata.com/training/public.html](http://www.stata.com/training/public.html).



## Visit us at APHA 2010

### Denver, Colorado, November 6–10

The American Public Health Association will have its annual meeting in Denver from November 6 through 10. For more information, go to [www.apha.org/meetings/highlights/](http://www.apha.org/meetings/highlights/).

Stata representatives, including Roberto G. Gutierrez, Director of Statistics, will be available at the Stata booth to answer your questions about all things Stata. Stop by booth #1603 to visit with the people who develop and support the software.

.....

## Stay informed

For up-to-the-minute news about Stata, be sure to check our website:

[www.stata.com](http://www.stata.com)

There you will find announcements regarding updates to Stata, upcoming public training courses, Stata Conferences and Users Group meetings, Stata Press books, web resources for learning Stata, Statalist archives, more than 400 FAQs about Stata and statistics, and more. You can also subscribe to an RSS feed to have our news headlines delivered straight to your browser.

Prefer to receive an email alert for news that interests you? You can subscribe to our email alert service at

[www.stata.com/alerts/](http://www.stata.com/alerts/)

## Upcoming NetCourses®

Enroll by visiting [www.stata.com/netcourse/](http://www.stata.com/netcourse/).

### NC101: Introduction to Stata

An introduction to using Stata interactively.

Dates: October 15–November 26, 2010

Enrollment deadline: October 14, 2010

Price: \$95

Details: [www.stata.com/netcourse/nc101.html](http://www.stata.com/netcourse/nc101.html)

### NC151: Introduction to Stata Programming

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

Dates: October 15–November 26, 2010

Enrollment deadline: October 14, 2010

Price: \$125

Details: [www.stata.com/netcourse/nc151.html](http://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: October 8–November 26, 2010

Enrollment deadline: October 7, 2010

Price: \$150

Details: [www.stata.com/netcourse/nc152.html](http://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: October 8–November 26, 2010

Enrollment deadline: October 7, 2010

Price: \$295

Details: [www.stata.com/netcourse/nc461.html](http://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](http://www.stata.com/netcourse/ncnow.html).

## 2010 Italian Stata Users Group meeting

Dates: November 11–12, 2010

Venue: Grand Hotel Baglioni  
Via Indipendenza, 8  
40121 Bologna  
Italy

Cost:

Day 1 only (no training course)	€90
Day 1 + course (in Italian) "Analisi dei Rischi Competitivi"	€375
Day 1 + course (in English) "Programming an Estimation Command in Stata and Mata"	€375
Optional dinner: All'Osteria Bottega	€45

Register: [www.stata.com/meeting/italy10/](http://www.stata.com/meeting/italy10/)

TStat S.r.l., the distributor of Stata software in Italy, is pleased to announce the seventh Italian Stata Users Group meeting, to be held in Bologna on November 11–12, 2010. The objective of the meeting is to provide Stata users working in different research areas with the unique opportunity to exchange ideas, experiences, and information on new applications of the software.

### Preliminary program

#### Thursday, November 11: Presentations

9:00–9:20 Registration of participants / Welcome from the scientific committee

9:20–9:30 **Stata—diffusione in Italia dal 2000 al 2010**

9:30–10:45 **Session I — Invited speaker**

##### Extracting results from nonlinear models

Maarten Buis, *Institut für Soziologie Eberhard Karls Universität Tübingen*

10:45–11:00 Coffee break

11:00–12:45 **Session II—User-written commands and routines I**

12:45–2:15 Lunch

2:15–3:30 **Session III—Exploiting the potential of Stata 11**

##### Multiple imputation in Stata

Bill Rising, *StataCorp LP*

3:30–4:45 **Session IV—User-written commands and routines II**

4:45–5:00 Coffee break

5:00–5:30 **Session V—Report to users / Wishes and grumbles**

Bill Rising, *StataCorp LP*

### Friday, November 12: Courses

9:00–6:00 **Programming an Estimation Command in Stata and Mata**

9:00–6:00 **Analisi dei Rischi Competitivi**

### Scientific organizers

Una-Louise Bell, *TStat S.r.l.*

Rino Bellocco, *Karolinska Institutet*

Giovanni Capelli, *Università degli Studi di Cassino*

Marcello Pagano, *Harvard School of Public Health*

Maurizio Pisati, *Università degli Studi di Milano Bicocca*

### Logistics organizer

Monica Gianni

TStat S.r.l.

Via Rettangolo, 12/14

67039 Sulmona (AQ)

Italy

Tel: +39-0864 210101

Fax: +39-0864 206014

Email: [statausers@tstat.it](mailto:statausers@tstat.it)

### Contact us

StataCorp  
4905 Lakeway Dr.  
College Station, TX 77845  
USA

Phone 979-696-4600  
Fax 979-696-4601  
Email [service@stata.com](mailto:service@stata.com)  
Web [www.stata.com](http://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/](http://www.stata.com/worldwide/).