# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined subject table of contents</td>
<td>1</td>
</tr>
<tr>
<td>Acronym glossary</td>
<td>39</td>
</tr>
<tr>
<td>Glossary</td>
<td>47</td>
</tr>
<tr>
<td>Vignette index</td>
<td>143</td>
</tr>
<tr>
<td>Author index</td>
<td>145</td>
</tr>
<tr>
<td>Subject index</td>
<td>185</td>
</tr>
</tbody>
</table>
## Combined subject table of contents

This is the complete contents for all manuals. Every estimation command has a postestimation entry; however, not all postestimation entries are listed here.

### Getting started

#### Data manipulation and management

<table>
<thead>
<tr>
<th>Basic data commands</th>
<th>Reshaping datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creating and dropping variables</td>
<td>Labeling, display formats, and notes</td>
</tr>
<tr>
<td>Functions and expressions</td>
<td>Changing and renaming variables</td>
</tr>
<tr>
<td>Strings</td>
<td>Examining data</td>
</tr>
<tr>
<td>Dates and times</td>
<td>File manipulation</td>
</tr>
<tr>
<td>Loading, saving, importing, and exporting data</td>
<td>Miscellaneous data commands</td>
</tr>
<tr>
<td>Combining data</td>
<td>Multiple imputation</td>
</tr>
</tbody>
</table>

#### Utilities

<table>
<thead>
<tr>
<th>Basic utilities</th>
<th>Internet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error messages</td>
<td>Data types and memory</td>
</tr>
<tr>
<td>Stored results</td>
<td>Advanced utilities</td>
</tr>
</tbody>
</table>

#### Graphics

<table>
<thead>
<tr>
<th>Common graphs</th>
<th>Survival-analysis graphs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distributional graphs</td>
<td>Time-series graphs</td>
</tr>
<tr>
<td>Item response theory graphs</td>
<td>More statistical graphs</td>
</tr>
<tr>
<td>Multivariate graphs</td>
<td>Editing</td>
</tr>
<tr>
<td>Quality control</td>
<td>Graph utilities</td>
</tr>
<tr>
<td>Regression diagnostic plots</td>
<td>Graph schemes</td>
</tr>
<tr>
<td>ROC analysis</td>
<td>Graph concepts</td>
</tr>
<tr>
<td>Smoothing and densities</td>
<td></td>
</tr>
</tbody>
</table>

#### Statistics

<table>
<thead>
<tr>
<th>ANOVA and related</th>
<th>Multidimensional scaling and biplots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic statistics</td>
<td>Multilevel mixed-effects models</td>
</tr>
<tr>
<td>Bayesian analysis</td>
<td>Multiple imputation</td>
</tr>
<tr>
<td>Binary outcomes</td>
<td>Multivariate analysis of variance and related techniques</td>
</tr>
<tr>
<td>Categorical outcomes</td>
<td>Nonlinear regression</td>
</tr>
<tr>
<td>Censored and truncated regression models</td>
<td>Nonparametric statistics</td>
</tr>
<tr>
<td>Cluster analysis</td>
<td>Ordinal outcomes</td>
</tr>
<tr>
<td>Correspondence analysis</td>
<td>Other statistics</td>
</tr>
<tr>
<td>Count outcomes</td>
<td>Pharmacokinetic statistics</td>
</tr>
<tr>
<td>Discriminant analysis</td>
<td>Power and sample size</td>
</tr>
<tr>
<td>Do-it-yourself generalized method of moments</td>
<td>Quality control</td>
</tr>
<tr>
<td>Do-it-yourself maximum likelihood estimation</td>
<td>ROC analysis</td>
</tr>
<tr>
<td>Endogenous covariates</td>
<td>Rotation</td>
</tr>
<tr>
<td>Epidemiology and related</td>
<td>Sample selection models</td>
</tr>
<tr>
<td>Estimation related</td>
<td>Simulation/resampling</td>
</tr>
<tr>
<td>Exact statistics</td>
<td>Standard postestimation tests, tables, and other analyses</td>
</tr>
<tr>
<td>Factor analysis and principal components</td>
<td>Structural equation modeling</td>
</tr>
<tr>
<td>Fractional outcomes</td>
<td>Survey data</td>
</tr>
<tr>
<td>Generalized linear models</td>
<td></td>
</tr>
<tr>
<td>Indicator and categorical variables</td>
<td></td>
</tr>
</tbody>
</table>
Combined subject table of contents

<table>
<thead>
<tr>
<th>Item response theory</th>
<th>Survival analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression and related</td>
<td>Time series, multivariate</td>
</tr>
<tr>
<td>Logistic and probit regression</td>
<td>Time series, univariate</td>
</tr>
<tr>
<td>Longitudinal data/panel data</td>
<td>Transforms and normality tests</td>
</tr>
<tr>
<td>Mixed models</td>
<td>Treatment effects</td>
</tr>
</tbody>
</table>

Matrix commands

<table>
<thead>
<tr>
<th>Basics</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programming</td>
<td>Mata</td>
</tr>
</tbody>
</table>

Programming

<table>
<thead>
<tr>
<th>Basics</th>
<th>Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program control</td>
<td>Advanced programming commands</td>
</tr>
<tr>
<td>Parsing and program arguments</td>
<td>Special-interest programming commands</td>
</tr>
<tr>
<td>Console output</td>
<td>File formats</td>
</tr>
<tr>
<td>Commonly used programming commands</td>
<td>Mata</td>
</tr>
<tr>
<td>Debugging</td>
<td></td>
</tr>
</tbody>
</table>

Interface features

Getting started

<table>
<thead>
<tr>
<th>[GSM]</th>
<th>Getting Started with Stata for Mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>[GSU]</td>
<td>Getting Started with Stata for Unix</td>
</tr>
<tr>
<td>[GSW]</td>
<td>Getting Started with Stata for Windows</td>
</tr>
<tr>
<td>[U]</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>[U]</td>
<td>Chapter 4</td>
</tr>
<tr>
<td>[R]</td>
<td>help</td>
</tr>
<tr>
<td>[R]</td>
<td>search</td>
</tr>
</tbody>
</table>

Data manipulation and management

Basic data commands

<table>
<thead>
<tr>
<th>[D]</th>
<th>codebook</th>
</tr>
</thead>
<tbody>
<tr>
<td>[D]</td>
<td>data management</td>
</tr>
<tr>
<td>[D]</td>
<td>data types</td>
</tr>
<tr>
<td>[D]</td>
<td>datetime</td>
</tr>
<tr>
<td>[D]</td>
<td>describe</td>
</tr>
<tr>
<td>[D]</td>
<td>edit</td>
</tr>
<tr>
<td>[D]</td>
<td>format</td>
</tr>
<tr>
<td>[D]</td>
<td>insobs</td>
</tr>
<tr>
<td>[D]</td>
<td>inspect</td>
</tr>
<tr>
<td>[D]</td>
<td>label</td>
</tr>
<tr>
<td>[D]</td>
<td>list</td>
</tr>
<tr>
<td>[D]</td>
<td>missing values</td>
</tr>
<tr>
<td>[D]</td>
<td>rename</td>
</tr>
<tr>
<td>[D]</td>
<td>save</td>
</tr>
<tr>
<td>[D]</td>
<td>sort</td>
</tr>
<tr>
<td>[D]</td>
<td>use</td>
</tr>
<tr>
<td>[D]</td>
<td>varmanage</td>
</tr>
</tbody>
</table>

Describe data contents
Introduction to data management commands
Quick reference for data types
Date and time values and variables
Describe data in memory or in file
Browse or edit data with Data Editor
Set variables’ output format
Add or insert observations
Display simple summary of data’s attributes
Manipulate labels
List values of variables
Quick reference for missing values
Rename variable
Save Stata dataset
Sort data
Load Stata dataset
Manage variable labels, formats, and other properties
Creating and dropping variables

- Date and time functions
- Mathematical functions
- Matrix functions
- Programming functions
- Random-number functions
- Selecting time-span functions
- Statistical functions
- String functions
- Trigonometric functions

- clear: Clear memory
- compress: Compress data in memory
- drop: Drop variables or observations
- egen: Extensions to generate
- generate: Create or change contents of variable
- orthog: Orthogonalize variables and compute orthogonal polynomials

Functions and expressions

- Section 12.4.2.1: Unicode string functions
- Chapter 13: Functions and expressions
- Date and time functions
- Mathematical functions
- Matrix functions
- Programming functions
- Random-number functions
- Selecting time-span functions
- Statistical functions
- String functions
- Trigonometric functions

- egen: Extensions to generate

Strings

- Section 12.4: Strings
- Section 12.4.2: Handling Unicode strings
- Chapter 23: Working with strings
- String functions
- data types: Quick reference for data types
- unicode: Unicode utilities

Dates and times

- Section 12.5.3: Date and time formats
- Chapter 24: Working with dates and times
- bcal: Business calendar file manipulation
- datetime: Date and time values and variables
- datetime business calendars: Business calendars
- datetime business calendars creation: Business calendars creation
- datetime display formats: Display formats for dates and times
- datetime translation: String to numeric date translation functions
Combined subject table of contents

Loading, saving, importing, and exporting data

[GS] Chapter 6 (GSM, GSU, GSW)                      Using the Data Editor
[U]  Chapter 21                                      Entering and importing data
[D]  edit                                         Browse or edit data with Data Editor
[D]  export                                       Overview of exporting data from Stata
[D]  import                                       Overview of importing data into Stata
[D]  import delimited                               Import delimited text data
[D]  import excel                                  Import and export Excel files
[D]  import haver                                  Import data from Haver Analytics databases
[D]  import sasxport                                Import and export datasets in SAS XPORT format
[D]  infile (fixed format)                         Read text data in fixed format with a dictionary
[D]  infile (free format)                          Read unformatted text data
[D]  infix (fixed format)                           Read text data in fixed format
[D]  input                                         Enter data from keyboard
[D]  odbc                                          Load, write, or view data from ODBC sources
[D]  outfile                                       Export dataset in text format
[P]  putexcel                                      Export results to an Excel file
[P]  putexcel advanced                             Export results to an Excel file using advanced syntax
[D]  save                                          Save Stata dataset
[D]  sysuse                                        Use shipped dataset
[D]  use                                           Load Stata dataset
[D]  webuse                                        Use dataset from Stata website
[D]  xmlsave                                       Export or import dataset in XML format

Combining data

[U]  Chapter 22                                     Combining datasets
[D]  append                                        Append datasets
[MI] mi append                                     Append mi data
[D]  cross                                         Form every pairwise combination of two datasets
[D]  joinby                                        Form all pairwise combinations within groups
[D]  merge                                         Merge datasets
[MI] mi merge                                      Merge mi data

Reshaping datasets

[D]  collapse                                      Make dataset of summary statistics
[D]  contract                                      Make dataset of frequencies and percentages
[D]  expand                                        Duplicate observations
[D]  expandcl                                      Duplicate clustered observations
[D]  fillin                                        Rectangularize dataset
[D]  obs                                           Increase the number of observations in a dataset
[D]  reshape                                       Convert data from wide to long form and vice versa
[MI] mi reshape                                    Reshape mi data
[TS] rolling                                      Rolling-window and recursive estimation
[D]  separate                                      Create separate variables
[SEM] ssd                                         Making summary statistics data (sem only)
[D]  stack                                         Stack data
[D]  statsby                                       Collect statistics for a command across a by list
[D]  xpose                                         Interchange observations and variables
**Labeling, display formats, and notes**

<table>
<thead>
<tr>
<th>[GS]</th>
<th>Chapter 7 (GSM, GSU, GSW)</th>
<th>Using the Variables Manager</th>
</tr>
</thead>
<tbody>
<tr>
<td>[U]</td>
<td>Section 12.5</td>
<td>Formats: Controlling how data are displayed</td>
</tr>
<tr>
<td>[U]</td>
<td>Section 12.6</td>
<td>Dataset, variable, and value labels</td>
</tr>
<tr>
<td>[D]</td>
<td>format</td>
<td>Set variables’ output format</td>
</tr>
<tr>
<td>[D]</td>
<td>label</td>
<td>Manipulate labels</td>
</tr>
<tr>
<td>[D]</td>
<td>label language</td>
<td>Labels for variables and values in multiple languages</td>
</tr>
<tr>
<td>[D]</td>
<td>labelbook</td>
<td>Label utilities</td>
</tr>
<tr>
<td>[D]</td>
<td>notes</td>
<td>Place notes in data</td>
</tr>
<tr>
<td>[D]</td>
<td>varmanage</td>
<td>Manage variable labels, formats, and other properties</td>
</tr>
</tbody>
</table>

**Changing and renaming variables**

<table>
<thead>
<tr>
<th>[GS]</th>
<th>Chapter 7 (GSM, GSU, GSW)</th>
<th>Using the Variables Manager</th>
</tr>
</thead>
<tbody>
<tr>
<td>[U]</td>
<td>Chapter 25</td>
<td>Working with categorical data and factor variables</td>
</tr>
<tr>
<td>[D]</td>
<td>clonevar</td>
<td>Clone existing variable</td>
</tr>
<tr>
<td>[D]</td>
<td>destring</td>
<td>Convert string variables to numeric variables and vice versa</td>
</tr>
<tr>
<td>[D]</td>
<td>encode</td>
<td>Convert string into numeric and vice versa</td>
</tr>
<tr>
<td>[D]</td>
<td>generate</td>
<td>Create or change contents of variable</td>
</tr>
<tr>
<td>[D]</td>
<td>mvencode</td>
<td>Change missing values to numeric values and vice versa</td>
</tr>
<tr>
<td>[D]</td>
<td>order</td>
<td>Reorder variables in dataset</td>
</tr>
<tr>
<td>[D]</td>
<td>recode</td>
<td>Recode categorical variables</td>
</tr>
<tr>
<td>[D]</td>
<td>rename</td>
<td>Rename variable</td>
</tr>
<tr>
<td>[D]</td>
<td>rename group</td>
<td>Rename groups of variables</td>
</tr>
<tr>
<td>[D]</td>
<td>split</td>
<td>Split string variables into parts</td>
</tr>
<tr>
<td>[D]</td>
<td>varmanage</td>
<td>Manage variable labels, formats, and other properties</td>
</tr>
</tbody>
</table>

**Examining data**

<table>
<thead>
<tr>
<th>[GS]</th>
<th>Chapter 6 (GSM, GSU, GSW)</th>
<th>Using the Data Editor</th>
</tr>
</thead>
<tbody>
<tr>
<td>[D]</td>
<td>cf</td>
<td>Compare two datasets</td>
</tr>
<tr>
<td>[D]</td>
<td>codebook</td>
<td>Describe data contents</td>
</tr>
<tr>
<td>[D]</td>
<td>compare</td>
<td>Compare two variables</td>
</tr>
<tr>
<td>[D]</td>
<td>count</td>
<td>Count observations satisfying specified conditions</td>
</tr>
<tr>
<td>[D]</td>
<td>describe</td>
<td>Describe data in memory or in file</td>
</tr>
<tr>
<td>[D]</td>
<td>ds</td>
<td>List variables matching name patterns or other characteristics</td>
</tr>
<tr>
<td>[D]</td>
<td>duplicates</td>
<td>Report, tag, or drop duplicate observations</td>
</tr>
<tr>
<td>[D]</td>
<td>edit</td>
<td>Browse or edit data with Data Editor</td>
</tr>
<tr>
<td>[D]</td>
<td>gsort</td>
<td>Ascending and descending sort</td>
</tr>
<tr>
<td>[D]</td>
<td>inspect</td>
<td>Display simple summary of data’s attributes</td>
</tr>
<tr>
<td>[D]</td>
<td>isid</td>
<td>Check for unique identifiers</td>
</tr>
<tr>
<td>[D]</td>
<td>lookfor</td>
<td>Search for string in variable names and labels</td>
</tr>
<tr>
<td>[R]</td>
<td>lv</td>
<td>Letter-value displays</td>
</tr>
<tr>
<td>[R]</td>
<td>misstable</td>
<td>Tabulate missing values</td>
</tr>
<tr>
<td>[MI]</td>
<td>mi describe</td>
<td>Describe mi data</td>
</tr>
<tr>
<td>[MI]</td>
<td>mi misstable</td>
<td>Tabulate pattern of missing values</td>
</tr>
<tr>
<td>[D]</td>
<td>pctlile</td>
<td>Create variable containing percentiles</td>
</tr>
<tr>
<td>[ST]</td>
<td>stdescribe</td>
<td>Describe survival-time data</td>
</tr>
<tr>
<td>[R]</td>
<td>summarize</td>
<td>Summary statistics</td>
</tr>
<tr>
<td>[SVY]</td>
<td>svy: tabulate oneway</td>
<td>One-way tables for survey data</td>
</tr>
<tr>
<td>[SVY]</td>
<td>svy: tabulate twoway</td>
<td>Two-way tables for survey data</td>
</tr>
<tr>
<td>[P]</td>
<td>tabdisp</td>
<td>Display tables</td>
</tr>
<tr>
<td>[R]</td>
<td>table</td>
<td>Flexible table of summary statistics</td>
</tr>
</tbody>
</table>
Combined subject table of contents

Multiple imputation

[MI]    mi add ............................................ Add imputations from another mi dataset
[MI]    mi append ........................................ Append mi data
[MI]    mi convert ....................................... Change style of mi data
[MI]    mi copy ............................................ Copy mi flongsep data
[MI]    mi describe ...................................... Describe mi data
[MI]    mi erase .......................................... Erase mi datasets
[MI]    mi expand ......................................... Expand mi data
[MI]    mi export ......................................... Export mi data
[MI]    mi export ice ..................................... Export mi data to ice format
[MI]    mi export nhanes1 ................................ Export mi data to NHANES format
[MI]    mi extract ......................................... Extract original or imputed data from mi data
[MI]    mi import ......................................... Import data into mi
[MI]    mi import flong ................................... Import flong-like data into mi
[MI]    mi import flongsep ................................ Import flongsep-like data into mi
[MI]    mi import ice ...................................... Import ice-format data into mi
[MI]    mi import nhanes1 ................................ Import NHANES-format data into mi
[MI]    mi import wide .................................... Import wide-like data into mi

Miscellaneous data commands

[D]    cd .................................................... Change directory
[D]    cf .................................................... Compare two datasets
[D]    changeeol ......................................... Convert end-of-line characters of text file
[D]    checksum .......................................... Calculate checksum of file
[D]    copy ................................................ Copy file from disk or URL
[D]    dir ................................................... Display filenames
[D]    erase ................................................ Erase a disk file
[D]    filefilter .......................................... Convert ASCII or binary patterns in a file
[D]    mkdir .............................................. Create directory
[D]    rmdir ................................................ Remove directory
[D]    type ................................................ Display contents of a file
[D]    unicode convertfile ............................... Low-level file conversion between encodings
[D]    unicode translate ................................ Translate files to Unicode
[D]    zipfile ............................................. Compress and uncompress files and directories in zip archive format

File manipulation

[D]    cor2data .......................................... Create dataset with specified correlation structure
[D]    drawnorm .......................................... Draw sample from multivariate normal distribution
[R]    dydx ................................................ Calculate numeric derivatives and integrals
[D]    icd .................................................. Introduction to ICD commands
[D]    icd10 ............................................... ICD-10 diagnosis codes
[D]    icd9 ................................................ ICD-9-CM diagnosis and procedure codes
[D]    ipolate ............................................. Linearly interpolate (extrapolate) values
[D]    range ............................................... Generate numerical range
[D]    sample .............................................. Draw random sample

Multiple imputation

[MI]    mi add ............................................ Add imputations from another mi dataset
[MI]    mi append ........................................ Append mi data
[MI]    mi convert ....................................... Change style of mi data
[MI]    mi copy ............................................ Copy mi flongsep data
[MI]    mi describe ...................................... Describe mi data
[MI]    mi erase .......................................... Erase mi datasets
[MI]    mi expand ......................................... Expand mi data
[MI]    mi export ......................................... Export mi data
[MI]    mi export ice ..................................... Export mi data to ice format
[MI]    mi export nhanes1 ................................ Export mi data to NHANES format
[MI]    mi extract ......................................... Extract original or imputed data from mi data
[MI]    mi import ......................................... Import data into mi
[MI]    mi import flong ................................... Import flong-like data into mi
[MI]    mi import flongsep ................................ Import flongsep-like data into mi
[MI]    mi import ice ...................................... Import ice-format data into mi
[MI]    mi import nhanes1 ................................ Import NHANES-format data into mi
[MI]    mi import wide .................................... Import wide-like data into mi

File manipulation

[D]    cd .................................................... Change directory
[D]    cf .................................................... Compare two datasets
[D]    changeeol ......................................... Convert end-of-line characters of text file
[D]    checksum .......................................... Calculate checksum of file
[D]    copy ................................................ Copy file from disk or URL
[D]    dir ................................................... Display filenames
[D]    erase ................................................ Erase a disk file
[D]    filefilter .......................................... Convert ASCII or binary patterns in a file
[D]    mkdir .............................................. Create directory
[D]    rmdir ................................................ Remove directory
[D]    type ................................................ Display contents of a file
[D]    unicode convertfile ............................... Low-level file conversion between encodings
[D]    unicode translate ................................ Translate files to Unicode
[D]    zipfile ............................................. Compress and uncompress files and directories in zip archive format

Miscellaneous data commands

[D]    cor2data .......................................... Create dataset with specified correlation structure
[D]    drawnorm .......................................... Draw sample from multivariate normal distribution
[R]    dydx ................................................ Calculate numeric derivatives and integrals
[D]    icd .................................................. Introduction to ICD commands
[D]    icd10 ............................................... ICD-10 diagnosis codes
[D]    icd9 ................................................ ICD-9-CM diagnosis and procedure codes
[D]    ipolate ............................................. Linearly interpolate (extrapolate) values
[D]    range ............................................... Generate numerical range
[D]    sample .............................................. Draw random sample

Multiple imputation

[MI]    mi add ............................................ Add imputations from another mi dataset
[MI]    mi append ........................................ Append mi data
[MI]    mi convert ....................................... Change style of mi data
[MI]    mi copy ............................................ Copy mi flongsep data
[MI]    mi describe ...................................... Describe mi data
[MI]    mi erase .......................................... Erase mi datasets
[MI]    mi expand ......................................... Expand mi data
[MI]    mi export ......................................... Export mi data
[MI]    mi export ice ..................................... Export mi data to ice format
[MI]    mi export nhanes1 ................................ Export mi data to NHANES format
[MI]    mi extract ......................................... Extract original or imputed data from mi data
[MI]    mi import ......................................... Import data into mi
[MI]    mi import flong ................................... Import flong-like data into mi
[MI]    mi import flongsep ................................ Import flongsep-like data into mi
[MI]    mi import ice ...................................... Import ice-format data into mi
[MI]    mi import nhanes1 ................................ Import NHANES-format data into mi
[MI]    mi import wide .................................... Import wide-like data into mi
Utilities

Basic utilities

| MI  | mi merge               | Merge mi data               |
| MI  | mi misstable          | Tabulate pattern of missing values |
| MI  | mi passive            | Generate/replace and register passive variables |
| MI  | mi ptrace             | Load parameter-trace file into Stata |
| MI  | mi rename             | Rename variable               |
| MI  | mi replace0           | Replace original data         |
| MI  | mi reset              | Reset imputed or passive variables |
| MI  | mi reshape            | Reshape mi data               |
| MI  | mi set                | Declare multiple-imputation data |
| MI  | mi stsplit           | Stsplit and stjoin mi data    |
| MI  | mi update             | Ensure that mi data are consistent |
| MI  | mi varying            | Identify variables that vary across imputations |
| MI  | mi xeq                | Execute command(s) on individual imputations |
| MI  | mi XXXset             | Declare mi data to be svy, st, ts, xt, etc. |
| MI  | noupdate option      | The noupdate option           |
| MI  | styles                | Dataset styles                |
| MI  | workflow              | Suggested workflow            |

Error messages

| U   | Chapter 8             | Error messages and return codes |
| P   | error                 | Display generic error message and exit |
| R   | error messages        | Error messages and return codes |
| P   | rmsg                  | Return messages                 |
### Stored results

<table>
<thead>
<tr>
<th>Section 13.5</th>
<th>Accessing coefficients and standard errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 18.8</td>
<td>Accessing results calculated by other programs</td>
</tr>
<tr>
<td>Section 18.9</td>
<td>Accessing results calculated by estimation commands</td>
</tr>
<tr>
<td>Section 18.10</td>
<td>Storing results</td>
</tr>
<tr>
<td>return</td>
<td>Return c-class values</td>
</tr>
<tr>
<td>return</td>
<td>Post the estimation results</td>
</tr>
<tr>
<td>estimates</td>
<td>Save and manipulate estimation results</td>
</tr>
<tr>
<td>estimates describe</td>
<td>Describe estimation results</td>
</tr>
<tr>
<td>estimates for</td>
<td>Repeat postestimation command across models</td>
</tr>
<tr>
<td>estimates notes</td>
<td>Add notes to estimation results</td>
</tr>
<tr>
<td>estimates replay</td>
<td>Redisplay estimation results</td>
</tr>
<tr>
<td>estimates save</td>
<td>Save and use estimation results</td>
</tr>
<tr>
<td>estimates stats</td>
<td>Model-selection statistics</td>
</tr>
<tr>
<td>estimates store</td>
<td>Store and restore estimation results</td>
</tr>
<tr>
<td>estimates table</td>
<td>Compare estimation results</td>
</tr>
<tr>
<td>estimates title</td>
<td>Set title for estimation results</td>
</tr>
<tr>
<td>_return</td>
<td>Preserve stored results</td>
</tr>
<tr>
<td>return</td>
<td>Return stored results</td>
</tr>
<tr>
<td>stored results</td>
<td>Stored results</td>
</tr>
</tbody>
</table>

### Internet

| Chapter 28 | Using the Internet to keep up to date |
| adoupdate | Update user-written ado-files |
| checksum | Calculate checksum of file |
| copy | Copy file from disk or URL |
| net | Install and manage user-written additions from the Internet |
| net search | Search the Internet for installable packages |
| netio | Control Internet connections |
| sj | Report Stata news |
| ssc | Install and uninstall packages from SSC |
| update | Check for official updates |
| use | Load Stata dataset |

### Data types and memory

| Chapter 6 | Managing memory |
| Section 12.2.2 | Numeric storage types |
| Section 12.4 | Strings |
| Section 12.4.2 | Handling Unicode strings |
| Section 13.12 | Precision and problems therein |
| Chapter 23 | Working with strings |
| compress | Compress data in memory |
| data types | Quick reference for data types |
| matsize | Set the maximum number of variables in a model |
| memory | Memory management |
| missing values | Quick reference for missing values |
| recast | Change storage type of variable |
Advanced utilities

- `assert` .......................................................... Verify truth of claim
- `cd` ................................................................. Change directory
- `changepool` ....................................................... Convert end-of-line characters of text file
- `checksum` ........................................................... Calculate checksum of file
- `copy` ............................................................... Copy file from disk or URL
- `_datasignature` ................................................. Determine whether data have changed
- `datasignature` ..................................................... Determine whether data have changed
- `db` ................................................................. Launch dialog
- `dialog programming` ............................................ Dialog programming
- `dir` ................................................................. Display filenames
- `discard` ............................................................. Drop automatically loaded programs
- `erase` ............................................................... Erase a disk file
- `file` ............................................................... Read and write text and binary programs
- `filefilter` .......................................................... Convert ASCII or binary patterns in a file
- `hexdump` ............................................................ Display hexadecimal report on file
- `mkdir` ............................................................... Create directory
- `more` ............................................................... The —more— message
- `query` .............................................................. Display system parameters
- `quietly` ............................................................. Quietly and noisily perform Stata command
- `rmdir` ............................................................... Remove directory
- `set` ................................................................. Overview of system parameters
- `set cformat` ....................................................... Format settings for coefficient tables
- `set defaults` ....................................................... Reset system parameters to original Stata defaults
- `set emptycells` ................................................ Set what to do with empty cells in interactions
- `set locale_functions` .......................................... Specify default locale for functions
- `set locale_ui` .................................................. Specify a localization package for the user interface
- `set rng` ............................................................ Set which random-number generator (RNG) to use
- `set seed` .......................................................... Specify random-number seed and state
- `set showbaselevels` ........................................... Display settings for coefficient tables
- `shell` ............................................................... Temporarily invoke operating system
- `signestimationsample` ........................................... Determine whether the estimation sample has changed
- `smcl` .............................................................. Stata Markup and Control Language
- `sysdir` .............................................................. Query and set system directories
- `type` ............................................................... Display contents of a file
- `unicode collator` ............................................. Language-specific Unicode collators
- `unicode convertfile` .......................................... Low-level file conversion between encodings
- `unicode encoding` ............................................ Unicode encoding utilities
- `unicode locale` ................................................... Unicode locale utilities
- `which` ............................................................. Display location and version for an ado-file

Graphics

Common graphs

- `graph intro` ........................................................ Introduction to graphics
- `graph` .............................................................. The graph command
- `graph bar` .......................................................... Bar charts
- `graph box` .......................................................... Box plots
- `graph close` ........................................................ Close Graph windows
- `graph combine` .................................................... Combine multiple graphs
<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>graph copy</code></td>
<td>Copy graph in memory</td>
</tr>
<tr>
<td><code>graph describe</code></td>
<td>Describe contents of graph in memory or on disk</td>
</tr>
<tr>
<td><code>graph dir</code></td>
<td>List names of graphs in memory and on disk</td>
</tr>
<tr>
<td><code>graph display</code></td>
<td>Display graph stored in memory</td>
</tr>
<tr>
<td><code>graph dot</code></td>
<td>Dot charts (summary statistics)</td>
</tr>
<tr>
<td><code>graph drop</code></td>
<td>Drop graphs from memory</td>
</tr>
<tr>
<td><code>graph export</code></td>
<td>Export current graph</td>
</tr>
<tr>
<td><code>graph manipulation</code></td>
<td>Graph manipulation commands</td>
</tr>
<tr>
<td><code>graph matrix</code></td>
<td>Matrix graphs</td>
</tr>
<tr>
<td><code>graph other</code></td>
<td>Other graphics commands</td>
</tr>
<tr>
<td><code>graph pie</code></td>
<td>Pie charts</td>
</tr>
<tr>
<td><code>graph play</code></td>
<td>Apply edits from a recording on current graph</td>
</tr>
<tr>
<td><code>graph print</code></td>
<td>Print a graph</td>
</tr>
<tr>
<td><code>graph query</code></td>
<td>List available schemes and styles</td>
</tr>
<tr>
<td><code>graph rename</code></td>
<td>Rename graph in memory</td>
</tr>
<tr>
<td><code>graph replay</code></td>
<td>Replay multiple graphs</td>
</tr>
<tr>
<td><code>graph save</code></td>
<td>Save graph to disk</td>
</tr>
<tr>
<td><code>graph set</code></td>
<td>Set graphics options</td>
</tr>
<tr>
<td><code>graph twoway</code></td>
<td>Twoway graphs</td>
</tr>
<tr>
<td><code>graph twoway area</code></td>
<td>Twoway line plot with area shading</td>
</tr>
<tr>
<td><code>graph twoway bar</code></td>
<td>Twoway bar plots</td>
</tr>
<tr>
<td><code>graph twoway connected</code></td>
<td>Twoway connected plots</td>
</tr>
<tr>
<td><code>graph twoway contour</code></td>
<td>Twoway contour plot with area shading</td>
</tr>
<tr>
<td><code>graph twoway contourline</code></td>
<td>Twoway contour-line plot</td>
</tr>
<tr>
<td><code>graph twoway dot</code></td>
<td>Twoway dot plots</td>
</tr>
<tr>
<td><code>graph twoway dropline</code></td>
<td>Twoway dropped-line plots</td>
</tr>
<tr>
<td><code>graph twoway fpfit</code></td>
<td>Twoway fractional-polynomial prediction plots</td>
</tr>
<tr>
<td><code>graph twoway fpfitci</code></td>
<td>Twoway fractional-polynomial prediction plots with CIs</td>
</tr>
<tr>
<td><code>graph twoway function</code></td>
<td>Twoway line plot of function</td>
</tr>
<tr>
<td><code>graph twoway histogram</code></td>
<td>Histogram plots</td>
</tr>
<tr>
<td><code>graph twoway kdensity</code></td>
<td>Kernel density plots</td>
</tr>
<tr>
<td><code>graph twoway lfit</code></td>
<td>Twoway linear prediction plots with CIs</td>
</tr>
<tr>
<td><code>graph twoway lfitci</code></td>
<td>Twoway linear prediction plots</td>
</tr>
<tr>
<td><code>graph twoway line</code></td>
<td>Twoway line plots</td>
</tr>
<tr>
<td><code>graph twoway lowess</code></td>
<td>Local linear smooth plots</td>
</tr>
<tr>
<td><code>graph twoway lpoly</code></td>
<td>Local polynomial smooth plots with CIs</td>
</tr>
<tr>
<td><code>graph twoway lpolyci</code></td>
<td>Local polynomial smooth plots</td>
</tr>
<tr>
<td><code>graph twoway mband</code></td>
<td>Twoway median-band plots</td>
</tr>
<tr>
<td><code>graph twoway mspline</code></td>
<td>Twoway median-spline plots</td>
</tr>
<tr>
<td><code>graph twoway parrow</code></td>
<td>Paired-coordinate plot with arrows</td>
</tr>
<tr>
<td><code>graph twoway parrowi</code></td>
<td>Twoway parrow with immediate arguments</td>
</tr>
<tr>
<td><code>graph twoway pccapsym</code></td>
<td>Paired-coordinate plot with spikes and marker symbols</td>
</tr>
<tr>
<td><code>graph twoway pci</code></td>
<td>Twoway paired-coordinate plot with immediate arguments</td>
</tr>
<tr>
<td><code>graph twoway pscatter</code></td>
<td>Paired-coordinate plot with markers</td>
</tr>
<tr>
<td><code>graph twoway qfit</code></td>
<td>Twoway quadratic prediction plots</td>
</tr>
<tr>
<td><code>graph twoway qfitci</code></td>
<td>Twoway quadratic prediction plots with CIs</td>
</tr>
<tr>
<td><code>graph twoway rarea</code></td>
<td>Range plot with area shading</td>
</tr>
<tr>
<td><code>graph twoway rbar</code></td>
<td>Range plot with bars</td>
</tr>
<tr>
<td><code>graph twoway rcap</code></td>
<td>Range plot with capped spikes</td>
</tr>
<tr>
<td><code>graph twoway rcapsym</code></td>
<td>Range plot with spikes capped with marker symbols</td>
</tr>
<tr>
<td>[G-2] graph twoway rconnected</td>
<td>Range plot with connected lines</td>
</tr>
<tr>
<td>[G-2] graph twoway rline</td>
<td>Range plot with lines</td>
</tr>
<tr>
<td>[G-2] graph twoway rsscatter</td>
<td>Range plot with markers</td>
</tr>
<tr>
<td>[G-2] graph twoway rspike</td>
<td>Range plot with spikes</td>
</tr>
<tr>
<td>[G-2] graph twoway scatter</td>
<td>Scatter with immediate arguments</td>
</tr>
<tr>
<td>[G-2] graph twoway spike</td>
<td>Twoway spike plots</td>
</tr>
<tr>
<td>[G-2] graph twoway tsline</td>
<td>Twoway line plots</td>
</tr>
<tr>
<td>[G-2] graph use</td>
<td>Display graph stored on disk</td>
</tr>
<tr>
<td>[R] histrogram</td>
<td>Histograms for continuous and categorical variables</td>
</tr>
<tr>
<td>[R] marginsplot</td>
<td>Graph results from margins (profile plots, etc.)</td>
</tr>
<tr>
<td>[G-2] palette</td>
<td>Display palettes of available selections</td>
</tr>
</tbody>
</table>

**Distributional graphs**

- [R] cumul: Cumulative distribution
- [R] diagnostic plots: Distributional diagnostic plots
- [R] dotplot: Comparative scatterplots
- [R] histogram: Histograms for continuous and categorical variables
- [R] ladder: Ladder of powers
- [R] spikeplot: Spike plots and rootograms
- [R] sunflower: Density-distribution sunflower plots

**Item response theory graphs**

- [MV] biplot: Biplots
- [IRT] irtgraph icc: Item characteristic curve plot
- [IRT] irtgraph iif: Item information function plot
- [IRT] irtgraph tcc: Test characteristic curve plot
- [IRT] irtgraph tif: Test information function plot

**Multivariate graphs**

- [MV] biplot: Biplots
- [MV] ca postestimation: Postestimation tools for ca and camat
- [MV] ca postestimation plots: Postestimation plots for ca and camat
- [MV] cluster dendrogram: Dendrograms for hierarchical cluster analysis
- [MV] mca postestimation: Postestimation tools for mca
- [MV] mca postestimation plots: Postestimation plots for mca
- [MV] mds postestimation: Postestimation tools for mds, mdsmat, and mdslong
- [MV] mds postestimation plots: Postestimation plots for mds, mdsmat, and mdslong
- [MV] procrustes postestimation: Postestimation tools for procrustes
- [MV] scoreplot: Score and loading plots
- [MV] screeplot: Scree plot

**Quality control**

- [R] cusum: Cusum plots and tests for binary variables
- [R] qc: Quality control charts
- [R] serrbar: Graph standard error bar chart

**Regression diagnostic plots**

- [R] regress postestimation diagnostic plots: Postestimation plots for regress
12 Combined subject table of contents

ROC analysis

[R] estat classification ........................................... Classification statistics and table
[R] estat gof ........................................................ Pearson or Hosmer–Lemeshow goodness-of-fit test
[R] logistic postestimation ........................................... Postestimation tools for logistic
[R] lroc ............................................................... Compute area under ROC curve and graph the curve
[R] lsens ............................................................... Graph sensitivity and specificity versus probability cutoff
[R] roccomp .......................................................... Tests of equality of ROC areas
[R] rocfit postestimation ........................................... Postestimation tools for rocfit
[R] rocregplot ......................................................... Plot marginal and covariate-specific ROC curves after rocreg
[R] roctab ............................................................. Nonparametric ROC analysis

Smoothing and densities

[R] kdensity ............................................................ Univariate kernel density estimation
[R] lowess .............................................................. Lowess smoothing
[R] lpoly ............................................................. Kernel-weighted local polynomial smoothing

Survival-analysis graphs

[ST] ltable ............................................................. Life tables for survival data
[ST] stci .............................................................. Confidence intervals for means and percentiles of survival time
[ST] stcox PH-assumption tests ................................... Tests of proportional-hazards assumption
[ST] stcurve .......................................................... Plot survivor, hazard, cumulative hazard, or cumulative incidence function
[ST] strate ............................................................. Tabulate failure rates and rate ratios
[ST] sts graph ......................................................... Graph the survivor, hazard, or cumulative hazard function

Time-series graphs

[TS] corrgram ........................................................ Tabulate and graph autocorrelations
[TS] cump ............................................................. Cumulative spectral distribution
[TS] estat acplot ....................................................... Plot parametric autocorrelation and autocovariance functions
[TS] estat aroots ...................................................... Check the stability condition of ARIMA functions
[TS] fcast graph ....................................................... Graph forecasts after fcast compute
[TS] irf cgraph ......................................................... Combined graphs of IRFs, dynamic-multiplier functions, and FEVDs
[TS] irf graph ........................................................ Graphs of IRFs, dynamic-multiplier functions, and FEVDs
[TS] irf ograph ........................................................ Overlaid graphs of IRFs, dynamic-multiplier functions, and FEVDs
[TS] pergram .......................................................... Periodogram
[TS] tsline ............................................................. Plot time-series data
[TS] varstable ......................................................... Check the stability condition of VAR or SVAR estimates
[TS] vecstable ......................................................... Check the stability condition of VECM estimates
[TS] wntestb .......................................................... Bartlett’s periodogram-based test for white noise
[TS] xcorr ............................................................. Cross-correlogram for bivariate time series

More statistical graphs

[BAYES] bayesgraph .................................................. Graphical summaries and convergence diagnostics
[R] epitab ............................................................. Tables for epidemiologists
[R] fp postestimation ................................................ Postestimation tools for fp
[R] grmeanby ........................................................ Graph means and medians by categorical variables
[R] pkexamine ........................................................ Calculate pharmacokinetic measures
[R] pksumm .......................................................... Summarize pharmacokinetic data
[PSS] power, graph ...................................................... Graph results from the power command
Editing

[G-1]  graph editor  ..................................................  Graph Editor

Graph utilities

[G-2]  set graphics  ..................................................  Set whether graphs are displayed
[G-2]  set printcolor  .............................................  Set how colors are treated when graphs are printed
[G-2]  set scheme  ..................................................  Set default scheme

Graph schemes

[G-4]  schemes intro  .................................................  Introduction to schemes
[G-4]  scheme economist  ...........................................  Scheme description: economist
[G-4]  scheme s1  ....................................................  Scheme description: s1 family
[G-4]  scheme s2  ....................................................  Scheme description: s2 family
[G-4]  scheme sj  ......................................................  Scheme description: sj

Graph concepts

[G-4]  concept: lines  ...............................................  Using lines
[G-4]  concept: repeated options  ...............................  Interpretation of repeated options
[G-4]  text  .............................................................  Text in graphs

Statistics

ANOVA and related

[U]  Chapter 26  ..................................................  Overview of Stata estimation commands
[R]  anova  .......................................................  Analysis of variance and covariance
[R]  contrast  .....................................................  Contrasts and linear hypothesis tests after estimation
[R]  icc  ...........................................................  Intraclass correlation coefficients
[R]  loneway  .......................................................  Large one-way ANOVA, random effects, and reliability
[MV] manova  .......................................................  Multivariate analysis of variance and covariance
[ME] meglm  .......................................................  Multilevel mixed-effects generalized linear model
[ME] mixed  .......................................................  Multilevel mixed-effects linear regression
[R] oneway  .........................................................  One-way analysis of variance
[R] pkcross  .........................................................  Analyze crossover experiments
[R] pkshape  .......................................................  Reshape (pharmacokinetic) Latin-square data
[R] pwcompare  .....................................................  Pairwise comparisons
[R] regress  .........................................................  Linear regression
[XT] xtreg  .........................................................  Fixed-, between-, and random-effects and population-averaged linear models

Basic statistics

[R]  anova  .......................................................  Analysis of variance and covariance
[R]  bitest  .........................................................  Binomial probability test
[R]  ci  ..............................................................  Confidence intervals for means, proportions, and variances
[R]  correlate  .....................................................  Correlations (covariances) of variables or coefficients
[D]  egen  ............................................................  Extensions to generate
## Combined subject table of contents

<table>
<thead>
<tr>
<th>Subject</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian analysis</td>
<td></td>
</tr>
</tbody>
</table>
  - `bayes` | Introduction to commands for Bayesian analysis |
  - `bayesgraph` | Graphical summaries and convergence diagnostics |
  - `bayesmh` | Bayesian regression using Metropolis–Hastings algorithm |
  - `bayesmh evaluators` | User-defined evaluators with bayesmh |
  - `bayesmh postestimation` | Postestimation tools for bayesmh |
  - `bayesstats` | Bayesian statistics after bayesmh |
  - `bayesstats ess` | Effective sample sizes and related statistics |
  - `bayesstats ic` | Bayesian information criteria and Bayes factors |
  - `bayesstats summary` | Bayesian summary statistics |
  - `bayestest` | Bayesian hypothesis testing |
  - `bayestest interval` | Interval hypothesis testing |
  - `bayestest model` | Hypothesis testing using model posterior probabilities |

### Binary outcomes

<table>
<thead>
<tr>
<th>Subject</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 20</td>
<td>Estimation and postestimation commands</td>
</tr>
<tr>
<td>Section 26.7</td>
<td>Binary-outcome qualitative dependent-variable models</td>
</tr>
<tr>
<td><code>binreg</code></td>
<td>Generalized linear models: Extensions to the binomial family</td>
</tr>
<tr>
<td><code>biprobit</code></td>
<td>Bivariate probit regression</td>
</tr>
<tr>
<td><code>cloglog</code></td>
<td>Complementary log-log regression</td>
</tr>
<tr>
<td><code>eteffects</code></td>
<td>Endogenous treatment-effects estimation</td>
</tr>
<tr>
<td><code>exlogistic</code></td>
<td>Exact logistic regression</td>
</tr>
<tr>
<td><code>glm</code></td>
<td>Generalized linear models</td>
</tr>
<tr>
<td><code>heckprobit</code></td>
<td>Probit model with sample selection</td>
</tr>
<tr>
<td><code>hetprobit</code></td>
<td>Heteroskedastic probit model</td>
</tr>
<tr>
<td><code>irt 1pl</code></td>
<td>One-parameter logistic model</td>
</tr>
</tbody>
</table>
Censored and truncated regression models

- **[IRT]**  
  - `irt 2pl` ................................................ Two-parameter logistic model
  - `irt 3pl` .............................................. Three-parameter logistic model
  - `irt hybrid` ........................................ Hybrid IRT models
  - `ivprobit` ............................................. Probit model with continuous endogenous covariates
  - `logistic` ............................................. Logistic regression, reporting odds ratios
  - `logit` ................................................ Logit regression, reporting coefficients
  - `mecloglog` ......................................... Multilevel mixed-effects complementary log-log regression
  - `melogit` ............................................. Multilevel mixed-effects logistic regression
  - `meprobit` ........................................... Multilevel mixed-effects probit regression
  - `meqrllogit` ......................................... Multilevel mixed-effects logistic regression (QR decomposition)
  - `probit` ............................................... Probit regression
  - `rocfit` ............................................... Parametric ROC models
  - `rocreg` .............................................. Receiver operating characteristic (ROC) regression
  - `scobit` ............................................... Skewed logistic regression
  - `teffects aipw` ...................................... Augmented inverse-probability weighting
  - `teffects ipw` ........................................ Inverse-probability weighting
  - `teffects ipwra` ..................................... Inverse-probability-weighted regression adjustment
  - `teffects nnmatch` .................................... Nearest-neighbor matching
  - `teffects psmatch` .................................. Propensity-score matching
  - `xtcloglog` ......................................... Random-effects and population-averaged cloglog models
  - `xtlogit` ............................................. Fixed-effects, random-effects, and population-averaged logit models
  - `xtprobit` ............................................ Random-effects and population-averaged probit models

Categorical outcomes

- **[U]**  
  - Chapter 20 ............................................. Estimation and postestimation commands
  - Section 26.11 ......................................... Multiple-outcome qualitative dependent-variable models
  - `aslogit` ............................................. Alternative-specific conditional logit (McFadden’s choice) model
  - `asmprobit` .......................................... Alternative-specific multinomial probit regression
  - `clogit` ................................................ Conditional (fixed-effects) logistic regression
  - `irt nrm` ............................................... Nominal response model
  - `mlogit` .............................................. Multinomial (polytomous) logistic regression
  - `mprobit` ............................................. Multinomial probit regression
  - `nlogit` ............................................... Nested logit regression
  - `slogit` ............................................... Stereotype logistic regression

Censored and truncated regression models

- **[R]**  
  - `churdle` ............................................. Cragg hurdle regression
  - `cpoisson` .......................................... Censored Poisson regression
  - `heckman` ............................................. Heckman selection model
  - `heckoprobit` ....................................... Ordered probit model with sample selection
  - `heckprobit` ......................................... Probit model with sample selection
  - `intreg` ............................................. Interval regression
  - `mestreg` ............................................ Multilevel mixed-effects parametric survival models
  - `streg` ............................................... Parametric survival models
  - `stteffects` ......................................... Treatment-effects estimation for observational survival-time data
  - `tnbreg` ............................................. Truncated negative binomial regression
  - `tobit` ............................................... Tobit regression
  - `tpoisson` .......................................... Truncated Poisson regression
  - `truncreg` .......................................... Truncated regression
Cluster analysis

[XT] xtintreg ............................ Random-effects interval-data regression models
[XT] xttobit ................................ Random-effects tobit models

Cluster analysis

[U] Section 26.28 .......................... Multivariate and cluster analysis
[MV] cluster ................................ Introduction to cluster-analysis commands
[MV] cluster dendrogram ................. Dendrograms for hierarchical cluster analysis
[MV] cluster generate ..................... Generate summary or grouping variables from a cluster analysis
[MV] cluster kmeans and kmedians ...... Kmeans and kmedians cluster analysis
[MV] cluster linkage ....................... Hierarchical cluster analysis
[MV] cluster notes ........................ Place notes in cluster analysis
[MV] cluster programming subroutines ... Add cluster-analysis routines
[MV] cluster programming utilities ...... Cluster-analysis programming utilities
[MV] cluster stop ......................... Cluster-analysis stopping rules
[MV] cluster utility ....................... List, rename, use, and drop cluster analyses
[MV] clustermat ............................ Introduction to clustermat commands
[MV] measure_option ..................... Option for similarity and dissimilarity measures
[MV] multivariate .......................... Introduction to multivariate commands

Correspondence analysis

[MV] ca .................................... Simple correspondence analysis
[MV] mca .................................... Multiple and joint correspondence analysis

Count outcomes

[U] Chapter 20 ............................. Estimation and postestimation commands
[U] Section 26.13 .......................... Count dependent-variable models
[U] Section 26.20.5 ....................... Count dependent-variable models with panel data
[R] cpoisson ................................ Censored Poisson regression
[TE] eteffects ................................ Endogenous treatment-effects estimation
[TE] etpoisson ................................ Poisson regression with endogenous treatment effects
[R] expoisson ................................ Exact Poisson regression
[ME] menbreg .............................. Multilevel mixed-effects negative binomial regression
[ME] mepoisson ............................ Multilevel mixed-effects Poisson regression
[ME] meqrpoisson ......................... Multilevel mixed-effects Poisson regression (QR decomposition)
[R] nbreg .................................... Negative binomial regression
[R] poisson .................................. Poisson regression
[TE] teffects aipw .......................... Augmented inverse-probability weighting
[TE] teffects ipw ........................... Inverse-probability weighting
[TE] teffects ipwra ....................... Inverse-probability-weighted regression adjustment
[TE] teffects nmmatch ...................... Nearest-neighbor matching
[TE] teffects psmatch ...................... Propensity-score matching
[TE] teffects ra ............................ Regression adjustment
[R] tnbreg .................................. Truncated negative binomial regression
[R] tpoisson ................................ Truncated Poisson regression
[XT] xtnbreg Fixed-effects, random-effects, & population-averaged negative binomial models
[XT] xtpoisson ............................. Fixed-effects, random-effects, and population-averaged Poisson models
[R] zinb .................................... Zero-inflated negative binomial regression
[R] zip ..................................... Zero-inflated Poisson regression
<table>
<thead>
<tr>
<th>Discriminant analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>[MV] candisc .............................. Canonical linear discriminant analysis</td>
</tr>
<tr>
<td>[MV] discrim ......................... Discriminant analysis</td>
</tr>
<tr>
<td>[MV] discrim estat ..................... Postestimation tools for discrim</td>
</tr>
<tr>
<td>[MV] discrim knn ....................... kth-nearest-neighbor discriminant analysis</td>
</tr>
<tr>
<td>[MV] discrim lda ....................... Linear discriminant analysis</td>
</tr>
<tr>
<td>[MV] discrim logistic ................ Logistic discriminant analysis</td>
</tr>
<tr>
<td>[MV] discrim qda ....................... Quadratic discriminant analysis</td>
</tr>
<tr>
<td>[MV] scoreplot .......................... Score and loading plots</td>
</tr>
<tr>
<td>[MV] screeplot ......................... Scree plot</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Do-it-yourself generalized method of moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>[R] gmm ..................................... Generalized method of moments estimation</td>
</tr>
<tr>
<td>[P] matrix ................................ Introduction to matrix commands</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Do-it-yourself maximum likelihood estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[P] matrix ................................ Introduction to matrix commands</td>
</tr>
<tr>
<td>[R] ml ....................................... Maximum likelihood estimation</td>
</tr>
<tr>
<td>[R] mlexp .................................. Maximum likelihood estimation of user-specified expressions</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Endogenous covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>[U] Chapter 20 ............ Estimation and postestimation commands</td>
</tr>
<tr>
<td>[U] Chapter 26 ............. Overview of Stata estimation commands</td>
</tr>
<tr>
<td>[TE] eteffects ............... Endogenous treatment-effects estimation</td>
</tr>
<tr>
<td>[TE] etpoisson ............... Poisson regression with endogenous treatment effects</td>
</tr>
<tr>
<td>[TE] etregress ............... Linear regression with endogenous treatment effects</td>
</tr>
<tr>
<td>[TS] forecast .................. Econometric model forecasting</td>
</tr>
<tr>
<td>[R] gmm ........................... Generalized method of moments estimation</td>
</tr>
<tr>
<td>[R] ivpoisson ................... Poisson model with continuous endogenous covariates</td>
</tr>
<tr>
<td>[R] ivprobit ..................... Probit model with continuous endogenous covariates</td>
</tr>
<tr>
<td>[R] ivregress .................... Single-equation instrumental-variables regression</td>
</tr>
<tr>
<td>[R] ivtobit ...................... Tobit model with continuous endogenous covariates</td>
</tr>
<tr>
<td>[R] reg3 .......................... Three-stage estimation for systems of simultaneous equations</td>
</tr>
<tr>
<td>[XT] xtabond .................... Arellano–Bond linear dynamic panel-data estimation</td>
</tr>
<tr>
<td>[XT] xtdpd ....................... Linear dynamic panel-data estimation</td>
</tr>
<tr>
<td>[XT] xtdpdsys .................. Arellano–Bover/Blundell–Bond linear dynamic panel-data estimation</td>
</tr>
<tr>
<td>[XT] xthtaylor ................ Hausman–Taylor estimator for error-components models</td>
</tr>
<tr>
<td>[XT] xtiivreg .................. Instrumental variables and two-stage least squares for panel-data models</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Epidemiology and related</th>
</tr>
</thead>
<tbody>
<tr>
<td>[R] binreg .................. Generalized linear models: Extensions to the binomial family</td>
</tr>
<tr>
<td>[R] brier ...................... Brier score decomposition</td>
</tr>
<tr>
<td>[R] clogit ..................... Conditional (fixed-effects) logistic regression</td>
</tr>
<tr>
<td>[R] dstdize .................... Direct and indirect standardization</td>
</tr>
<tr>
<td>[R] epitab ...................... Tables for epidemiologists</td>
</tr>
<tr>
<td>[R] exlogistic ................ Exact logistic regression</td>
</tr>
<tr>
<td>[D] icd .......................... Introduction to ICD commands</td>
</tr>
<tr>
<td>[D] icd10 ....................... ICD-10 diagnosis codes</td>
</tr>
<tr>
<td>[D] icd9 .......................... ICD-9-CM diagnosis and procedure codes</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>R</td>
</tr>
</tbody>
</table>

Also see Treatment effects

**Estimation related**

<table>
<thead>
<tr>
<th>R</th>
<th>BIC note</th>
<th>Calculating and interpreting BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>constraint</td>
<td>Define and list constraints</td>
</tr>
<tr>
<td>R</td>
<td>eform_option</td>
<td>Displaying exponentiated coefficients</td>
</tr>
<tr>
<td>R</td>
<td>estimation options</td>
<td>Estimation options</td>
</tr>
<tr>
<td>R</td>
<td>fp</td>
<td>Fractional polynomial regression</td>
</tr>
<tr>
<td>R</td>
<td>maximize</td>
<td>Details of iterative maximization</td>
</tr>
<tr>
<td>R</td>
<td>mfp</td>
<td>Multivariable fractional polynomial models</td>
</tr>
<tr>
<td>R</td>
<td>mkspline</td>
<td>Linear and restricted cubic spline construction</td>
</tr>
<tr>
<td>R</td>
<td>stepwise</td>
<td>Stepwise estimation</td>
</tr>
<tr>
<td>R</td>
<td>vce_option</td>
<td>Variance estimators</td>
</tr>
<tr>
<td>XT</td>
<td>vce_options</td>
<td>Variance estimators</td>
</tr>
</tbody>
</table>

**Exact statistics**

<table>
<thead>
<tr>
<th>U</th>
<th>Section 26.14</th>
<th>Exact estimators</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>bitest</td>
<td>Binomial probability test</td>
</tr>
<tr>
<td>R</td>
<td>centile</td>
<td>Report centile and confidence interval</td>
</tr>
<tr>
<td>R</td>
<td>ci</td>
<td>Confidence intervals for means, proportions, and variances</td>
</tr>
<tr>
<td>R</td>
<td>dstdize</td>
<td>Direct and indirect standardization</td>
</tr>
<tr>
<td>R</td>
<td>epitab</td>
<td>Tables for epidemiologists</td>
</tr>
<tr>
<td>R</td>
<td>exlogistic</td>
<td>Exact logistic regression</td>
</tr>
<tr>
<td>R</td>
<td>expoisson</td>
<td>Exact Poisson regression</td>
</tr>
<tr>
<td>R</td>
<td>ksmirnov</td>
<td>Kolmogorov–Smirnov equality-of-distributions test</td>
</tr>
<tr>
<td>R</td>
<td>loneway</td>
<td>Large one-way ANOVA, random effects, and reliability</td>
</tr>
<tr>
<td>R</td>
<td>ranksum</td>
<td>Equality tests on unmatched data</td>
</tr>
<tr>
<td>R</td>
<td>roctab</td>
<td>Nonparametric ROC analysis</td>
</tr>
<tr>
<td>R</td>
<td>symmetry</td>
<td>Symmetry and marginal homogeneity tests</td>
</tr>
<tr>
<td>R</td>
<td>tabulate twoway</td>
<td>Two-way table of frequencies</td>
</tr>
<tr>
<td>R</td>
<td>tetrachoric</td>
<td>Tetrachoric correlations for binary variables</td>
</tr>
</tbody>
</table>
### Factor analysis and principal components
- **alpha** Compute interitem correlations (covariances) and Cronbach’s alpha
- **canon** Canonical correlations
- **factor** Factor analysis
- **pca** Principal component analysis
- **rotate** Orthogonal and oblique rotations after factor and pca
- **scoreplot** Score and loading plots
- **screenplot** Score and loading plots
- **rotatemat** Orthogonal and oblique rotations of a Stata matrix
- **tetrachoric** Tetrachoric correlations for binary variables

### Fractional outcomes
- **betareg** Beta regression
- **eteffects** Endogenous treatment-effects estimation
- **fracreg** Fractional response regression
- **teffects ipw** Nearest-neighbor matching
- **teffects nnmatch** Nearest-neighbor matching
- **teffects psmatch** Propensity-score matching

### Generalized linear models
- **binreg** Generalized linear models: Extensions to the binomial family
- **fracreg** Fractional response regression
- **glm** Generalized linear models
- **xtgee** Fit population-averaged panel-data models by using GEE

### Indicator and categorical variables
- **fvset** Declare factor-variable settings

### Item response theory
- **irt 1pl** One-parameter logistic model
- **irt 2pl** Two-parameter logistic model
- **irt 3pl** Three-parameter logistic model
- **irt grm** Graded response model
- **irt hybrid** Hybrid IRT models
- **irt nrm** Nominal response model
- **irt pcm** Partial credit model
- **irt rsm** Rating scale model
- **irtgraph icc** Item characteristic curve plot
- **irtgraph iif** Item information function plot
- **irtgraph tcc** Test characteristic curve plot
- **irtgraph tif** Test information function plot
### Linear regression and related

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>areg</td>
<td>Linear regression with a large dummy-variable set</td>
</tr>
<tr>
<td>cnreg</td>
<td>Constrained linear regression</td>
</tr>
<tr>
<td>constraint</td>
<td>Define and list constraints</td>
</tr>
<tr>
<td>eivreg</td>
<td>Errors-in-variables regression</td>
</tr>
<tr>
<td>etpoisson</td>
<td>Poisson regression with endogenous treatment effects</td>
</tr>
<tr>
<td>etregress</td>
<td>Linear regression with endogenous treatment effects</td>
</tr>
<tr>
<td>fp</td>
<td>Fractional polynomial regression</td>
</tr>
<tr>
<td>frontier</td>
<td>Stochastic frontier models</td>
</tr>
<tr>
<td>glm</td>
<td>Generalized linear models</td>
</tr>
<tr>
<td>heckman</td>
<td>Heckman selection model</td>
</tr>
<tr>
<td>ivpoisson</td>
<td>Poisson model with continuous endogenous covariates</td>
</tr>
<tr>
<td>ivregress</td>
<td>Single-equation instrumental-variables regression</td>
</tr>
<tr>
<td>ivtobit</td>
<td>Tobit model with continuous endogenous covariates</td>
</tr>
<tr>
<td>lpoly</td>
<td>Kernel-weighted local polynomial smoothing</td>
</tr>
<tr>
<td>meglm</td>
<td>Multilevel mixed-effects generalized linear model</td>
</tr>
<tr>
<td>mfp</td>
<td>Multivariable fractional polynomial models</td>
</tr>
<tr>
<td>mixed</td>
<td>Multilevel mixed-effects linear regression</td>
</tr>
<tr>
<td>mvreg</td>
<td>Multivariate regression</td>
</tr>
<tr>
<td>nestreg</td>
<td>Nested model statistics</td>
</tr>
<tr>
<td>newey</td>
<td>Regression with Newey–West standard errors</td>
</tr>
<tr>
<td>prais</td>
<td>Prais–Winsten and Cochrane–Orcutt regression</td>
</tr>
<tr>
<td>qreg</td>
<td>Quantile regression</td>
</tr>
<tr>
<td>reg3</td>
<td>Three-stage estimation for systems of simultaneous equations</td>
</tr>
<tr>
<td>regress</td>
<td>Linear regression</td>
</tr>
<tr>
<td>rocfit</td>
<td>Parametric ROC models</td>
</tr>
<tr>
<td>rreg</td>
<td>Robust regression</td>
</tr>
<tr>
<td>stcox</td>
<td>Cox proportional hazards model</td>
</tr>
<tr>
<td>stcrreg</td>
<td>Competing-risks regression</td>
</tr>
<tr>
<td>stepwise</td>
<td>Stepwise estimation</td>
</tr>
<tr>
<td>streg</td>
<td>Parametric survival models</td>
</tr>
<tr>
<td>sureg</td>
<td>Zellner’s seemingly unrelated regression</td>
</tr>
<tr>
<td>tnbreg</td>
<td>Truncated negative binomial regression</td>
</tr>
<tr>
<td>vwls</td>
<td>Variance-weighted least squares</td>
</tr>
<tr>
<td>xtabond</td>
<td>Arellano–Bond linear dynamic panel-data estimation</td>
</tr>
<tr>
<td>xtdpd</td>
<td>Linear dynamic panel-data estimation</td>
</tr>
<tr>
<td>xtdpdsys</td>
<td>Arellano–Bover/Blundell–Bond linear dynamic panel-data estimation</td>
</tr>
<tr>
<td>xtgee</td>
<td>Fit population-averaged panel-data models by using GEE</td>
</tr>
<tr>
<td>xtlgs</td>
<td>Fit panel-data models by using GLS</td>
</tr>
<tr>
<td>xthtaylor</td>
<td>Hausman–Taylor estimator for error-components models</td>
</tr>
<tr>
<td>xtitreg</td>
<td>Instrumental variables and two-stage least squares for panel-data models</td>
</tr>
<tr>
<td>xtpcse</td>
<td>Linear regression with panel-corrected standard errors</td>
</tr>
<tr>
<td>xtrc</td>
<td>Random-coefficients model</td>
</tr>
<tr>
<td>xtreg</td>
<td>Fixed-, between-, and random-effects and population-averaged linear models</td>
</tr>
<tr>
<td>xtregar</td>
<td>Fixed- and random-effects linear models with an AR(1) disturbance</td>
</tr>
<tr>
<td>xtstreg</td>
<td>Random-effects parametric survival models</td>
</tr>
</tbody>
</table>
### Logistic and probit regression

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>logit</td>
<td>Logistic regression, reporting coefficients</td>
</tr>
<tr>
<td>logistic</td>
<td>Logistic regression, reporting odds ratios</td>
</tr>
<tr>
<td>mlogit</td>
<td>Multinomial (polytomous) logistic regression</td>
</tr>
<tr>
<td>mprobit</td>
<td>Multinomial probit regression</td>
</tr>
<tr>
<td>nlogit</td>
<td>Nested logit regression</td>
</tr>
<tr>
<td>ologit</td>
<td>Ordered logistic regression</td>
</tr>
<tr>
<td>oprobit</td>
<td>Ordered probit regression</td>
</tr>
<tr>
<td>probit</td>
<td>Probit regression</td>
</tr>
<tr>
<td>rologit</td>
<td>Rank-ordered logistic regression</td>
</tr>
<tr>
<td>scobit</td>
<td>Skewed logistic regression</td>
</tr>
<tr>
<td>slogit</td>
<td>Stereotype logistic regression</td>
</tr>
<tr>
<td>xtclolog</td>
<td>Random-effects and population-averaged cloglog models</td>
</tr>
<tr>
<td>xtgee</td>
<td>Fit population-averaged panel-data models by using GEE</td>
</tr>
<tr>
<td>xtlogit</td>
<td>Fixed-effects, random-effects, and population-averaged logit models</td>
</tr>
<tr>
<td>xtololog</td>
<td>Random-effects ordered logistic models</td>
</tr>
<tr>
<td>xtoprobit</td>
<td>Random-effects ordered probit models</td>
</tr>
<tr>
<td>xtprobit</td>
<td>Random-effects and population-averaged probit models</td>
</tr>
</tbody>
</table>

### Longitudinal data/panel data

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 20</td>
<td>Estimation and postestimation commands</td>
</tr>
<tr>
<td>Chapter 26</td>
<td>Overview of Stata estimation commands</td>
</tr>
<tr>
<td>asclogit</td>
<td>Alternative-specific conditional logit (McFadden’s choice) model</td>
</tr>
<tr>
<td>asmprobit</td>
<td>Alternative-specific multinomial probit regression</td>
</tr>
<tr>
<td>asroprobit</td>
<td>Alternative-specific rank-ordered probit regression</td>
</tr>
<tr>
<td>biprobit</td>
<td>Bivariate probit regression</td>
</tr>
<tr>
<td>clogit</td>
<td>Conditional (fixed-effects) logistic regression</td>
</tr>
<tr>
<td>cloglog</td>
<td>Complementary log-log regression</td>
</tr>
<tr>
<td>exlogistic</td>
<td>Exact logistic regression</td>
</tr>
<tr>
<td>heckoprobit</td>
<td>Ordered probit model with sample selection</td>
</tr>
<tr>
<td>heckprobit</td>
<td>Probit model with sample selection</td>
</tr>
<tr>
<td>hetprobit</td>
<td>Heteroskedastic probit model</td>
</tr>
<tr>
<td>irt 1pl</td>
<td>One-parameter logistic model</td>
</tr>
<tr>
<td>irt 2pl</td>
<td>Two-parameter logistic model</td>
</tr>
<tr>
<td>irt 3pl</td>
<td>Three-parameter logistic model</td>
</tr>
<tr>
<td>irt grm</td>
<td>Graded response model</td>
</tr>
<tr>
<td>irt hybrid</td>
<td>Hybrid IRT models</td>
</tr>
<tr>
<td>irt nrm</td>
<td>Nominal response model</td>
</tr>
<tr>
<td>irt pcm</td>
<td>Partial credit model</td>
</tr>
<tr>
<td>irt rsm</td>
<td>Rating scale model</td>
</tr>
<tr>
<td>ivprobit</td>
<td>Probit model with continuous endogenous covariates</td>
</tr>
<tr>
<td>logistic</td>
<td>Logistic regression, reporting odds ratios</td>
</tr>
<tr>
<td>logit</td>
<td>Logistic regression, reporting coefficients</td>
</tr>
<tr>
<td>meologit</td>
<td>Multilevel mixed-effects logistic regression</td>
</tr>
<tr>
<td>meoprobit</td>
<td>Multilevel mixed-effects ordered probit regression</td>
</tr>
<tr>
<td>meprobit</td>
<td>Multilevel mixed-effects probit regression</td>
</tr>
<tr>
<td>meqrologit</td>
<td>Multilevel mixed-effects logistic regression (QR decomposition)</td>
</tr>
<tr>
<td>mlogit</td>
<td>Multinomial (polytomous) logistic regression</td>
</tr>
<tr>
<td>mprobit</td>
<td>Multinomial probit regression</td>
</tr>
<tr>
<td>nlogit</td>
<td>Nested logit regression</td>
</tr>
<tr>
<td>ologit</td>
<td>Ordered logistic regression</td>
</tr>
<tr>
<td>oprobit</td>
<td>Ordered probit regression</td>
</tr>
<tr>
<td>probit</td>
<td>Probit regression</td>
</tr>
<tr>
<td>rologit</td>
<td>Rank-ordered logistic regression</td>
</tr>
<tr>
<td>scobit</td>
<td>Skewed logistic regression</td>
</tr>
<tr>
<td>slogit</td>
<td>Stereotype logistic regression</td>
</tr>
<tr>
<td>xtclolog</td>
<td>Random-effects and population-averaged cloglog models</td>
</tr>
<tr>
<td>xtgee</td>
<td>Fit population-averaged panel-data models by using GEE</td>
</tr>
<tr>
<td>xtlogit</td>
<td>Fixed-effects, random-effects, and population-averaged logit models</td>
</tr>
<tr>
<td>xtololog</td>
<td>Random-effects ordered logistic models</td>
</tr>
<tr>
<td>xtoprobit</td>
<td>Random-effects ordered probit models</td>
</tr>
<tr>
<td>xtprobit</td>
<td>Random-effects and population-averaged probit models</td>
</tr>
</tbody>
</table>
### Mixed models

**ME**  
- **meprobit**  
- **meqppoisson**  
- **mixed**  
- **quadchk**  
- **xt**  
- **xtabond**  
- **xtcloglog**  
- **xtdata**  
- **xtdescribe**  
- **xtdpd**  
- **xtdpdys**  
- **xtfrontier**  
- **xtgeee**  
- **xtgls**  
- **xthaylor**  
- **xtintreg**  
- **xtivreg**  
- **xtlologit**  
- **xtlogit**  
- **xtprobit**  
- **xtreg**  
- **xtpcse**  
- **xtpoisson**  
- **xtpcset**  
- **xtrc**  
- **xtreg**  
- **xtset**  
- **xtstreg**  
- **xtsum**  
- **xttab**  
- **xtunitroot**  

**XT**  
- **xtregar**  
- **xtstreg**  
- **xtsum**  
- **xttab**  
- **xtunitroot**

**ME**  
- **me**  
- **meqreg**  
- **melogit**  
- **menbreg**  
- **melogit**  
- **meglm**  
- **mecloglog**  

**MV**  
- **manova**  

**U**  
- **Chapter 20**  
- **Section 26.21**

**R**  
- **anova**  
- **icc**  

**MV**  
- **manova**  

**ME**  
- **me**  
- **mecloglog**  
- **meqpmreg**  
- **meqlogit**  
- **meprobit**  
- **mepoisson**  
- **meqprobit**  

**XT**  
- **xtregar**  
- **xtstreg**  
- **xtsum**  
- **xttab**  
- **xtunitroot**
<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>meqrpoisson</td>
<td>Multilevel mixed-effects Poisson regression (QR decomposition)</td>
</tr>
<tr>
<td>mestreg</td>
<td>Multilevel mixed-effects parametric survival models</td>
</tr>
<tr>
<td>mixed</td>
<td>Multilevel mixed-effects linear regression</td>
</tr>
<tr>
<td>xtcloglog</td>
<td>Random-effects and population-averaged cloglog models</td>
</tr>
<tr>
<td>xttlogit</td>
<td>Fixed-effects, random-effects, and population-averaged logit models</td>
</tr>
<tr>
<td>xttologit</td>
<td>Random-effects ordered logistic models</td>
</tr>
<tr>
<td>xtprobit</td>
<td>Random-effects ordered probit models</td>
</tr>
<tr>
<td>xtreg</td>
<td>Fixed-, between-, and random-effects and population-averaged linear models</td>
</tr>
<tr>
<td>biplot</td>
<td>Biplots</td>
</tr>
<tr>
<td>mds</td>
<td>Multidimensional scaling for two-way data</td>
</tr>
<tr>
<td>mdslong</td>
<td>Multidimensional scaling of proximity data in long format</td>
</tr>
<tr>
<td>mdsmat</td>
<td>Multidimensional scaling of proximity data in a matrix</td>
</tr>
<tr>
<td>measure_option</td>
<td>Option for similarity and dissimilarity measures</td>
</tr>
<tr>
<td>mi estimate</td>
<td>Estimation using previously saved estimation results</td>
</tr>
<tr>
<td>mi estimate using</td>
<td>Estimation using multiple imputations</td>
</tr>
<tr>
<td>mi impute</td>
<td>Impute missing values</td>
</tr>
<tr>
<td>mi impute chained</td>
<td>Impute using chained equations</td>
</tr>
<tr>
<td>mi impute intreg</td>
<td>Impute using interval regression</td>
</tr>
<tr>
<td>mi impute logit</td>
<td>Impute using logistic regression</td>
</tr>
<tr>
<td>mi impute mlogit</td>
<td>Impute using multinomial logistic regression</td>
</tr>
<tr>
<td>mi impute monotone</td>
<td>Impute using multivariate normal regression</td>
</tr>
<tr>
<td>mi impute nbvn</td>
<td>Impute using negative binomial regression</td>
</tr>
</tbody>
</table>

**Multidimensional scaling and biplots**

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>me</td>
<td>Introduction to multilevel mixed-effects models</td>
</tr>
<tr>
<td>meclloglog</td>
<td>Multilevel mixed-effects complementary log-log regression</td>
</tr>
<tr>
<td>meglm</td>
<td>Multilevel mixed-effects generalized linear model</td>
</tr>
<tr>
<td>melogit</td>
<td>Multilevel mixed-effects logistic regression</td>
</tr>
<tr>
<td>menbreg</td>
<td>Multilevel mixed-effects negative binomial regression</td>
</tr>
<tr>
<td>meologit</td>
<td>Multilevel mixed-effects ordered logistic regression</td>
</tr>
<tr>
<td>meoprobit</td>
<td>Multilevel mixed-effects ordered probit regression</td>
</tr>
<tr>
<td>mepoisson</td>
<td>Multilevel mixed-effects Poisson regression</td>
</tr>
<tr>
<td>meprobit</td>
<td>Multilevel mixed-effects probit regression</td>
</tr>
<tr>
<td>meqrlogit</td>
<td>Multilevel mixed-effects logistic regression (QR decomposition)</td>
</tr>
<tr>
<td>meqrpoisson</td>
<td>Multilevel mixed-effects Poisson regression (QR decomposition)</td>
</tr>
<tr>
<td>mestreg</td>
<td>Multilevel mixed-effects parametric survival models</td>
</tr>
<tr>
<td>mixed</td>
<td>Multilevel mixed-effects linear regression</td>
</tr>
</tbody>
</table>

**Multiple imputation**

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mi estimate</td>
<td>Estimation commands for use with mi estimate</td>
</tr>
<tr>
<td>mi estimate using</td>
<td>Estimation using multiple imputations</td>
</tr>
<tr>
<td>mi estimate postestimation</td>
<td>Postestimation tools for mi estimate</td>
</tr>
<tr>
<td>mi impute</td>
<td>Impute missing values</td>
</tr>
<tr>
<td>mi impute chained</td>
<td>Impute missing values using chained equations</td>
</tr>
<tr>
<td>mi impute intreg</td>
<td>Impute using interval regression</td>
</tr>
<tr>
<td>mi impute logit</td>
<td>Impute using logistic regression</td>
</tr>
<tr>
<td>mi impute mlogit</td>
<td>Impute using multinomial logistic regression</td>
</tr>
<tr>
<td>mi impute monotone</td>
<td>Impute using multivariate normal regression</td>
</tr>
<tr>
<td>mi impute mvn</td>
<td>Impute using negative binomial regression</td>
</tr>
</tbody>
</table>
Multivariate analysis of variance and related techniques

Section 26.28 Multivariate and cluster analysis

Canonical correlations

Hotelling’s T-squared generalized means test

Multivariate analysis of variance and covariance

Multivariate tests of covariances

Multivariate tests of means

Nonlinear regression

Box–Cox regression models

Nonlinear least-squares estimation

Estimation of nonlinear systems of equations

Nonparametric statistics

Binomial probability test

Bootstrap sampling and estimation

Sampling with replacement

Report bootstrap results

Report centile and confidence interval

Cusum plots and tests for binary variables

Univariate kernel density estimation

Kolmogorov–Smirnov equality-of-distributions test

Kruskal–Wallis equality-of-populations rank test

Lowess smoothing

Kernel-weighted local polynomial smoothing

Test for trend across ordered groups

Tests of proportions

Quantile regression

Equality tests on unmatched data

Receiver operating characteristic (ROC) analysis

Tests of equality of ROC areas

Receiver operating characteristic (ROC) regression

Plot marginal and covariate-specific ROC curves after rocreg

Nonparametric ROC analysis

Test for random order

Equality tests on matched data

Monte Carlo simulations

Robust nonlinear smoother

Spearman’s and Kendall’s correlations

Symmetry and marginal homogeneity tests

Two-way table of frequencies
### Ordinal outcomes

| [U] | Chapter 20 | Estimation and postestimation commands |
| [R] | asoprobit | Alternative-specific rank-ordered probit regression |
| [R] | heckoprobit | Ordered probit model with sample selection |
| [IRT] | irt grm | Graded response model |
| [IRT] | irt pcm | Partial credit model |
| [IRT] | irt rsm | Rating scale model |
| [ME] | meologit | Multilevel mixed-effects ordered logistic regression |
| [ME] | meoprobit | Multilevel mixed-effects ordered probit regression |
| [R] | ologit | Ordered logistic regression |
| [R] | oprobit | Ordered probit regression |
| [R] | rologit | Rank-ordered logistic regression |
| [XT] | xtologit | Random-effects ordered logistic models |
| [XT] | xtoprob | Random-effects ordered probit models |

### Other statistics

| [MV] | alpha | Compute interitem correlations (covariances) and Cronbach’s alpha |
| [R] | amean | Arithmetic, geometric, and harmonic means |
| [R] | brier | Brier score decomposition |
| [R] | centile | Report centile and confidence interval |
| [R] | kappa | Interrater agreement |
| [MV] | mvtest correlations | Multivariate tests of correlations |
| [R] | pcorr | Partial and semipartial correlation coefficients |
| [D] | pctile | Create variable containing percentiles |
| [D] | range | Generate numerical percentiles |

### Pharmacokinetic statistics

| [U] | Section 26.29 | Pharmacokinetic data |
| [R] | pk | Pharmacokinetic (biopharmaceutical) data |
| [R] | pkcollapse | Generate pharmacokinetic measurement dataset |
| [R] | pkcross | Analyze crossover experiments |
| [R] | pkequiv | Perform bioequivalence tests |
| [R] | pkexamine | Calculate pharmacokinetic measures |
| [R] | pkshape | Reshape (pharmacokinetic) Latin-square data |
| [R] | pksumm | Summarize pharmacokinetic data |

### Power and sample size

<p>| [U] | Section 26.31 | Power and sample-size analysis |
| [PSS] | GUI | Graphical user interface for power and sample-size analysis |
| [PSS] | power | Power and sample-size analysis for hypothesis tests |
| [PSS] | power cmh | Power and sample size for the Cochran–Mantel–Haenszel test |
| [PSS] | power cox | Power analysis for the Cox proportional hazards model |
| [PSS] | power exponential | Power analysis for the exponential test |
| [PSS] | power logrank | Power analysis for the log-rank test |
| [PSS] | power mcc | Power analysis for matched case–control studies |
| [PSS] | power one-correlation | Power analysis for a one-sample correlation test |
| [PSS] | power one-mean | Power analysis for a one-sample mean test |
| [PSS] | power one-proportion | Power analysis for a one-sample proportion test |
| [PSS] | power one-variance | Power analysis for a one-sample variance test |
| [PSS] | power one-way | Power analysis for a one-way analysis of variance |</p>
<table>
<thead>
<tr>
<th>[PSS]</th>
<th>power pairedmeans</th>
<th>Power analysis for a two-sample paired-means test</th>
</tr>
</thead>
<tbody>
<tr>
<td>[PSS]</td>
<td>power pairedproportions</td>
<td>Power analysis for a two-sample paired-proportions test</td>
</tr>
<tr>
<td>[PSS]</td>
<td>power repeated</td>
<td>Power analysis for repeated-measures analysis of variance</td>
</tr>
<tr>
<td>[PSS]</td>
<td>power trend</td>
<td>Power analysis for the Cochran–Armitage trend test</td>
</tr>
<tr>
<td>[PSS]</td>
<td>power twocorrelations</td>
<td>Power analysis for a two-sample correlations test</td>
</tr>
<tr>
<td>[PSS]</td>
<td>power twomeans</td>
<td>Power analysis for a two-sample means test</td>
</tr>
<tr>
<td>[PSS]</td>
<td>power twoproportions</td>
<td>Power analysis for a two-sample proportions test</td>
</tr>
<tr>
<td>[PSS]</td>
<td>power twovariances</td>
<td>Power analysis for a two-sample variances test</td>
</tr>
<tr>
<td>[PSS]</td>
<td>power twoway</td>
<td>Power analysis for two-way analysis of variance</td>
</tr>
<tr>
<td>[PSS]</td>
<td>unbalanced designs</td>
<td>Specifications for unbalanced designs</td>
</tr>
</tbody>
</table>

**Quality control**

<table>
<thead>
<tr>
<th>[R]</th>
<th>cusum</th>
<th>Cusum plots and tests for binary variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>[R]</td>
<td>qc</td>
<td>Quality control charts</td>
</tr>
<tr>
<td>[R]</td>
<td>serrbar</td>
<td>Graph standard error bar chart</td>
</tr>
</tbody>
</table>

**ROC analysis**

<table>
<thead>
<tr>
<th>[U]</th>
<th>Section 26.8</th>
<th>ROC analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>[R]</td>
<td>roc</td>
<td>Receiver operating characteristic (ROC) analysis</td>
</tr>
<tr>
<td>[R]</td>
<td>roccomp</td>
<td>Tests of equality of ROC areas</td>
</tr>
<tr>
<td>[R]</td>
<td>rocfit</td>
<td>Parametric ROC models</td>
</tr>
<tr>
<td>[R]</td>
<td>rocfit postestimation</td>
<td>Postestimation tools for rocfit</td>
</tr>
<tr>
<td>[R]</td>
<td>rocreg</td>
<td>Receiver operating characteristic (ROC) regression</td>
</tr>
<tr>
<td>[R]</td>
<td>rocreg postestimation</td>
<td>Postestimation tools for rocreg</td>
</tr>
<tr>
<td>[R]</td>
<td>rocregplot</td>
<td>Plot marginal and covariate-specific ROC curves after rocreg</td>
</tr>
<tr>
<td>[R]</td>
<td>roctab</td>
<td>Nonparametric ROC analysis</td>
</tr>
</tbody>
</table>

**Rotation**

<table>
<thead>
<tr>
<th>[MV]</th>
<th>procrustes</th>
<th>Procrustes transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[MV]</td>
<td>rotate</td>
<td>Orthogonal and oblique rotations after factor and pca</td>
</tr>
<tr>
<td>[MV]</td>
<td>rotatemat</td>
<td>Orthogonal and oblique rotations of a Stata matrix</td>
</tr>
</tbody>
</table>

**Sample selection models**

<table>
<thead>
<tr>
<th>[U]</th>
<th>Chapter 20</th>
<th>Estimation and postestimation commands</th>
</tr>
</thead>
<tbody>
<tr>
<td>[U]</td>
<td>Section 26.18</td>
<td>Models with endogenous sample selection</td>
</tr>
<tr>
<td>[TE]</td>
<td>etpoisson</td>
<td>Poisson regression with endogenous treatment effects</td>
</tr>
<tr>
<td>[TE]</td>
<td>etregress</td>
<td>Linear regression with endogenous treatment effects</td>
</tr>
<tr>
<td>[R]</td>
<td>heckman</td>
<td>Heckman selection model</td>
</tr>
<tr>
<td>[R]</td>
<td>heckoprobit</td>
<td>Ordered probit model with sample selection</td>
</tr>
<tr>
<td>[R]</td>
<td>heckprobit</td>
<td>Probit model with sample selection</td>
</tr>
</tbody>
</table>

**Simulation/resampling**

<table>
<thead>
<tr>
<th>[R]</th>
<th>bootstrap</th>
<th>Bootstrap sampling and estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[R]</td>
<td>bsample</td>
<td>Sampling with replacement</td>
</tr>
<tr>
<td>[R]</td>
<td>jackknife</td>
<td>Jackknife estimation</td>
</tr>
<tr>
<td>[R]</td>
<td>permute</td>
<td>Monte Carlo permutation tests</td>
</tr>
<tr>
<td>[R]</td>
<td>simulate</td>
<td>Monte Carlo simulations</td>
</tr>
</tbody>
</table>
Standard postestimation tests, tables, and other analyses

- Section 13.5: Accessing coefficients and standard errors
- Chapter 20: Estimation and postestimation commands
- contr: Contrasts and linear hypothesis tests after estimation
- correlate: Correlations (covariances) of variables or coefficients
- estat: Postestimation statistics
- estat ic: Display information criteria
- estat summarize: Summarize estimation sample
- estat vce: Display covariance matrix estimates
- estimates: Save and manipulate estimation results
- estimates describe: Describe estimation results
- estimates for: Repeat postestimation command across models
- estimates notes: Add notes to estimation results
- estimates replay: Redisplay estimation results
- estimates save: Save and use estimation results
- estimates stats: Model-selection statistics
- estimates store: Store and restore estimation results
- estimates table: Compare estimation results
- estimates title: Set title for estimation results
- forecast: Econometric model forecasting
- forecast adjust: Adjust a variable by add factoring, replacing, etc.
- forecast clear: Clear current model from memory
- forecast coefvector: Specify an equation via a coefficient vector
- forecast create: Create a new forecast model
- forecast describe: Describe features of the forecast model
- forecast drop: Drop forecast variables
- forecast estimates: Add estimation results to a forecast model
- forecast exogenous: Declare exogenous variables
- forecast identity: Add an identity to a forecast model
- forecast list: List forecast commands composing current model
- forecast query: Check whether a forecast model has been started
- forecast solve: Obtain static and dynamic forecasts
- hausman: Hausman specification test
- lincom: Linear combinations of parameters
- linktest: Specification link test for single-equation models
- lrtest: Likelihood-ratio test after estimation
- margins: Marginal means, predictive margins, and marginal effects
- margins, contrast: Contrasts of margins
- margins, pwcompare: Pairwise comparisons of margins
- marginsplot: Graph results from margins (profile plots, etc.)
- mvtest: Multivariate tests
- ncom: Nonlinear combinations of estimators
- postest: Postestimation Selector
- predict: Obtain predictions, residuals, etc., after estimation
- predictnl: Obtain nonlinear predictions, standard errors, etc., after estimation
- pwcompare: Pairwise comparisons
- suet: Seemingly unrelated estimation
- test: Test linear hypotheses after estimation
- testnl: Test nonlinear hypotheses after estimation
### Structural equation modeling

- **Section 26.4** Structural equation modeling (SEM)
- **Builder** SEM Builder
- **Builder, generalized** SEM Builder for generalized models
- **estat eform** Display exponentiated coefficients
- **estat eqgof** Equation-level goodness-of-fit statistics
- **estat eqtest** Equation-level test that all coefficients are zero
- **estat framework** Display estimation results in modeling framework
- **estat gdf** Group-level goodness-of-fit statistics
- **estat ginvariant** Tests for invariance of parameters across groups
- **estat gof** Goodness-of-fit statistics
- **estat mindices** Modification indices
- **estat residuals** Display mean and covariance residuals
- **estat scoretests** Score tests
- **estat stable** Check stability of nonrecursive system
- **estat stdize** Test standardized parameters
- **estat summarize** Report summary statistics for estimation sample
- **estat teffects** Decomposition of effects into total, direct, and indirect
- **example 1** Single-factor measurement model
- **example 2** Creating a dataset from published covariances
- **example 3** Two-factor measurement model
- **example 4** Goodness-of-fit statistics
- **example 5** Modification indices
- **example 6** Linear regression
- **example 7** Nonrecursive structural model
- **example 8** Testing that coefficients are equal, and constraining them
- **example 9** Structural model with measurement component
- **example 10** MIMIC model
- **example 11** estat framework
- **example 12** Seemingly unrelated regression
- **example 13** Equation-level Wald test
- **example 14** Predicted values
- **example 15** Higher-order CFA
- **example 16** Correlation
- **example 17** Correlated uniqueness model
- **example 18** Latent growth model
- **example 19** Creating multiple-group summary statistics data
- **example 20** Two-factor measurement model by group
- **example 21** Group-level goodness of fit
- **example 22** Testing parameter equality across groups
- **example 23** Specifying parameter constraints across groups
- **example 24** Reliability
- **example 25** Creating summary statistics data from raw data
- **example 26** Fitting a model with data missing at random
- **example 27g** Single-factor measurement model (generalized response)
- **example 28g** One-parameter logistic IRT (Rasch) model
- **example 29g** Two-parameter logistic IRT model
- **example 30g** Two-level measurement model (multilevel, generalized response)
- **example 31g** Two-factor measurement model (generalized response)
- **example 32g** Full structural equation model (generalized response)
- **example 33g** Logistic regression
| SEM     | example 34g                      | Combined models (generalized responses) |
| SEM     | example 35g                      | Ordered probit and ordered logit         |
| SEM     | example 36g                      | MIMIC model (generalized response)       |
| SEM     | example 37g                      | Multinomial logistic regression          |
| SEM     | example 38g                      | Random-intercept and random-slope models (multilevel) |
| SEM     | example 39g                      | Three-level model (multilevel, generalized response) |
| SEM     | example 40g                      | Crossed models (multilevel)              |
| SEM     | example 41g                      | Two-level multinomial logistic regression (multilevel) |
| SEM     | example 42g                      | One- and two-level mediation models (multilevel) |
| SEM     | example 43g                      | Tobit regression                         |
| SEM     | example 44g                      | Interval regression                      |
| SEM     | example 45g                      | Heckman selection model                  |
| SEM     | example 46g                      | Endogenous treatment-effects model       |
| SEM     | gsem                             | Generalized structural equation model estimation command |
| SEM     | gsem estimation options          | Options affecting estimation             |
| SEM     | gsem family-and-link options     | Family-and-link options                  |
| SEM     | gsem model description options   | Model description options                |
| SEM     | gsem path notation extensions    | Command syntax for path diagrams         |
| SEM     | gsem postestimation              | Postestimation tools for gsem            |
| SEM     | gsem reporting options           | Options affecting reporting of results   |
| SEM     | intro 1                          | Introduction                             |
| SEM     | intro 2                          | Learning the language: Path diagrams and command language |
| SEM     | intro 3                          | Learning the language: Factor-variable notation (gsem only) |
| SEM     | intro 4                          | Substantive concepts                     |
| SEM     | intro 5                          | Tour of models                           |
| SEM     | intro 6                          | Comparing groups (sem only)              |
| SEM     | intro 7                          | Postestimation tests and predictions     |
| SEM     | intro 8                          | Robust and clustered standard errors     |
| SEM     | intro 9                          | Standard errors, the full story          |
| SEM     | intro 10                         | Fitting models with survey data          |
| SEM     | intro 11                         | Fitting models with summary statistics data (sem only) |
| SEM     | intro 12                         | Convergence problems and how to solve them |
| SEM     | lincom                           | Linear combinations of parameters        |
| SEM     | lrtest                           | Likelihood-ratio test of linear hypothesis |
| SEM     | methods and formulas for gsem    | Methods and formulas for gsem            |
| SEM     | methods and formulas for sem     | Methods and formulas for sem             |
| SEM     | nlcom                            | Nonlinear combinations of parameters     |
| SEM     | predict after gsem               | Generalized linear predictions, etc.     |
| SEM     | predict after sem                | Factor scores, linear predictions, etc.  |
| SEM     | sem                              | Structural equation model estimation command |
| SEM     | sem and gsem option constraints() | Specifying constraints                   |
| SEM     | sem and gsem option covstructure() | Specifying covariance restrictions     |
| SEM     | sem and gsem option from()       | Specifying starting values               |
| SEM     | sem and gsem option reliability() | Fraction of variance not due to measurement error |
| SEM     | sem and gsem path notation       | Command syntax for path diagrams         |
| SEM     | sem and gsem syntax options      | Options affecting interpretation of syntax |
| SEM     | sem estimation options           | Options affecting estimation             |
| SEM     | sem group options                | Fitting models on different groups       |
| SEM     | sem model description options    | Model description options                |
| SEM     | sem option method()              | Specifying method and calculation of VCE |
| SEM     | sem option noxconditional        | Computing means, etc., of observed exogenous variables |
Combined subject table of contents

Survey data

[SEM] sem option select() .......................... Using sem with summary statistics data
[SEM] sem path notation extensions .................. Command syntax for path diagrams
[SEM] sem postestimation .......................... Postestimation tools for sem
[SEM] sem reporting options ........................ Options affecting reporting of results
[SEM] sem ssd options ........................ Options for use with summary statistics data
[SEM] ssd ........................................ Making summary statistics data (sem only)
[SEM] test ........................................ Wald test of linear hypotheses
[SEM] testnl ........................................ Wald test of nonlinear hypotheses

[SVY] Chapter 20 .............................. Estimation and postestimation commands
[SVY] Section 26.26 .............................. Survey data
[U] Chapter 20 .............................. Estimation and postestimation commands
[U] Section 26.26 .............................. Survey data
[SVY] bootstrap_options ......................... Introduction to survey commands
[SVY] brr_options .............................. More options for BRR variance estimation
[SVY] direct standardization ........................ Direct standardization of means, proportions, and ratios
[SVY] estat ........................................ Postestimation statistics for survey data
[SVY] jackknife_options ......................... More options for jackknife variance estimation
[SVY] ml for svy ................................ Maximum pseudolikelihood estimation for survey data
[SVY] poststratification .......................... Poststratification for survey data
[P] _robust ...................................... Robust variance estimates
[SVY] sdr_options .............................. More options for SDR variance estimation
[SVY] subpopulation estimation .................. Subpopulation estimation for survey data
[SVY] svy ......................................... The survey prefix command
[SVY] svy bootstrap ............................ Bootstrap for survey data
[SVY] svy brr ..................................... Balanced repeated replication for survey data
[SVY] svy estimation ............................ Estimation commands for survey data
[SVY] svy jackknife ............................. Jackknife estimation for survey data
[SVY] svy postestimation ....................... Postestimation tools for svy
[SVY] svy sdr ..................................... Successive difference replication for survey data
[SVY] svy: tabulate oneway ........................ One-way tables for survey data
[SVY] svy: tabulate twoway ........................ Two-way tables for survey data
[SVY] svydesc ..................................... Describe survey data
[SVY] svymarkout ................................ Mark observations for exclusion on the basis of survey characteristics
[SVY] svyset .................................... Declare survey design for dataset
[M] mi XXXset .................................. Declare mi data to be svy, st, ts, xt, etc.
[SVY] variance estimation ...................... Variance estimation for survey data

Survival analysis

[U] Chapter 20 .............................. Estimation and postestimation commands
[U] Section 26.20.6 .............................. Survival models with panel data
[U] Section 26.22 .............................. Survival-time (failure-time) models
[U] Section 26.23 .............................. Treatment-effect models
[U] Section 26.31 .............................. Power and sample-size analysis
[ST] survival analysis .......................... Introduction to survival analysis
[ST] ct ............................................ Count-time data
[ST] ctset ....................................... Declare data to be count-time data
[ST] cttost ..................................... Convert count-time data to survival-time data
[ST] discrete ................................... Discrete-time survival analysis
[ST] ltable ....................................... Life tables for survival data
[ME] mestreg .................................. Multilevel mixed-effects parametric survival models
snapspan .......................... Convert snapshot data to time-span data
st .................................. Survival-time data
st_is ............................... Survival analysis subroutines for programmers
stbase ................................ Form baseline dataset
stci ................................. Confidence intervals for means and percentiles of survival time
stcox ................................ Cox proportional hazards model
stcox PH-assumption tests ........ Tests of proportional-hazards assumption
stcrreg ............................. Competing-risks regression
stcurve ............................. Plot survivor, hazard, cumulative hazard, or cumulative incidence function
stdescribe .......................... Describe survival-time data
stepwise ............................ Stepwise estimation
stfill ............................... Fill in by carrying forward values of covariates
stgen ................................ Generate variables reflecting entire histories
stir .................................. Report incidence-rate comparison
stptime ............................. Calculate person-time, incidence rates, and SMR
strate ................................ Graph rates and rate ratios
streg ................................ Parametric survival models
sts .................................. Generate, graph, list, and test the survivor and cumulative hazard functions
sts generate ....................... Create variables containing survivor and related functions
sts graph ............................ Graph the survivor, hazard, or cumulative hazard function
sts list .............................. List the survivor or cumulative hazard function
sts test .............................. Test equality of survivor functions
sts .................................. Declare data to be survival-time data
mi XXXset ........................... Declare mi data to be svy, st, ts, xt, etc.
stteffects iwpr ................................ Survival-time inverse-probability weighting
stteffects ipwra ..................... Survival-time inverse-probability-weighted regression adjustment
stteffects ra .......................... Survival-time regression adjustment
stteffects ipw ................................ Survival-time inverse-probability weighting
sttimestr .................... Convert survival-time data to time-span data
stvar ................................ Report variables that vary over time
xt .................................. Random-effects parametric survival models

Also see Power and sample size

Time series, multivariate
Section 11.4.4 .......................... Time-series varlists
Section 13.10 .......................... Time-series operators
Chapter 20 ........................... Estimation and postestimation commands
Section 26.19 .......................... Models with time-series data
time series .......................... Introduction to time-series commands
dfactor ................................ Dynamic-factor models
fcast compute ........................ Compute dynamic forecasts after var, svar, or vec
fcast graph .......................... Graph forecasts after fcast compute
forecast .............................. Econometric model forecasting
forecast adjust ....................... Adjust a variable by add factoring, replacing, etc.
forecast clear ....................... Clear current model from memory
forecast coefvector ................. Specify an equation via a coefficient vector
forecast create ..................... Create a new forecast model
forecast describe .......................... Describe features of the forecast model
forecast drop .................................. Drop forecast variables
forecast estimates ............................. Add estimation results to a forecast model
forecast exogenous ........................... Declare exogenous variables
forecast identity ..................... Add an identity to a forecast model
forecast list ............................. List forecast commands composing current model
forecast query .................................. Check whether a forecast model has been started
forecast solve .................................. Obtain static and dynamic forecasts
irf ................................. Create and analyze IRFs, dynamic-multiplier functions, and FEVDs
irf add .................................. Add results from an IRF file to the active IRF file
irf cgraph ........... Combined graphs of IRFs, dynamic-multiplier functions, and FEVDs
irf create .................. Obtain IRFs, dynamic-multiplier functions, and FEVDs
irf table ...................... Combined tables of IRFs, dynamic-multiplier functions, and FEVDs
irf describe .................................. Describe an IRF file
irf drop .................................. Drop IRF results from the active IRF file
irf graph ..................... Graphs of IRFs, dynamic-multiplier functions, and FEVDs
irf ograph ........... Overlaid graphs of IRFs, dynamic-multiplier functions, and FEVDs
irf rename ..................... Rename an IRF result in an IRF file
irf set .................................. Set the active IRF file
irf table ...................... Tables of IRFs, dynamic-multiplier functions, and FEVDs
mgarch ccc ................. Constant conditional correlation multivariate GARCH models
mgarch dcc ................... Dynamic conditional correlation multivariate GARCH models
mgarch dvech .............. Diagonal vech multivariate GARCH models
mgarch vcc ................. Varying conditional correlation multivariate GARCH models
rolling ......................... Rolling-window and recursive estimation
sspace .................................. State-space models
tsappend .................................. Add observations to a time-series dataset
tsfill .................................. Fill in gaps in time variable
tsline ............................. Report time-series aspects of a dataset or estimation sample
tssset .................................. Time-series operator programming command
var intro ..................... Introduction to vector autoregressive models
var svar ..................... Structural vector autoregressive models
var ..................................... Vector autoregressive models
varbasic ................................. Fit a simple VAR and graph IRFs or FEVDs
vargranger .................. Perform pairwise Granger causality tests after var or svar
varlmar ...................... Perform LM test for residual autocorrelation after var or svar
varnorm ............................. Test for normally distributed disturbances after var or svar
varsoc ..................... Obtain lag-order selection statistics for VARs and VECMs
varstable .................. Check the stability condition of VAR or SVAR estimates
varwle ............................. Obtain Wald lag-exclusion statistics after var or svar
vec intro ..................... Introduction to vector error-correction models
vec ..................................... Vector error-correction models
vecclmar .................. Perform LM test for residual autocorrelation after vec
vecnorm ............................. Test for normally distributed disturbances after vec
vecrank ..................... Estimate the cointegrating rank of a VECM
vecstable .................. Check the stability condition of VECM estimates
xcorr .......................... Cross-correlogram for bivariate time series
Time series, univariate

[U] Section 11.4.4  Time-series varlists
[U] Section 13.10  Time-series operators
[U] Chapter 20  Estimation and postestimation commands
[U] Section 26.19  Models with time-series data
[TS] time series  Introduction to time-series commands
[TS] arch  Autoregressive conditional heteroskedasticity (ARCH) family of estimators
[TS] arfima  Autoregressive fractionally integrated moving-average models
[TS] arima  ARIMA, ARMAX, and other dynamic regression models
[TS] corrgam  Tabulate and graph autocovariances
[TS] cumsp  Cumulative spectral distribution
[TS] dfgls  DF-GLS unit-root test
[TS] dfuller  Augmented Dickey–Fuller unit-root test
[TS] estat acplot  Parametric autocorrelation and autocovariance functions
[TS] estat aroots  Check the stability condition of ARIMA estimates
[TS] estat sbknown  Test for a structural break with a known break date
[TS] estat sbsingle  Test for a structural break with an unknown break date
[TS] forecast  Econometric model forecasting
[TS] forecast adjust  Adjust a variable by add factoring, replacing, etc.
[TS] forecast clear  Clear current model from memory
[TS] forecast coefvector  Specify an equation via a coefficient vector
[TS] forecast create  Create a new forecast model
[TS] forecast describe  Describe features of the forecast model
[TS] forecast drop  Drop forecast variables
[TS] forecast estimates  Declare exogenous variables
[TS] forecast exogenous  Add an identity to a forecast model
[TS] forecast list  List forecast commands composing current model
[TS] forecast query  Check whether a forecast model has been started
[TS] forecast solve  Obtain static and dynamic forecasts
[TS] mswitch  Markov-switching regression models
[TS] newey  Regression with Newey–West standard errors
[TS] pergram  Periodogram
[TS] pperron  Phillips–Perron unit-root test
[TS] prais  Prais–Winsten and Cochrane–Orcutt regression
[TS] psdensity  Parametric spectral density estimation after arima, arfima, and ucm
[R] regress postestimation time series  Postestimation tools for regress with time series
[TS] rolling  Rolling-window and recursive estimation
[TS] sspace  State-space models
[TS] tsappend  Add observations to a time-series dataset
[TS] tsfill  Fill in gaps in time variable
[TS] tsfilter  Filter a time-series, keeping only selected periodicities
[TS] tsfilter bk  Baxter–King time-series filter
[TS] tsfilter bw  Butterworth time-series filter
[TS] tsfilter cf  Christiano–Fitzgerald time-series filter
[TS] tsfilter hp  Hodrick–Prescott time-series filter
[TS] tsline  Plot time-series data
[TS] tsreport  Report time-series aspects of a dataset or estimation sample
[TS] tsrevar  Time-series operator programming command
[TS] tsset  Declare data to be time-series data
[TS] tssmooth  Smooth and forecast univariate time-series data
Combined subject table of contents

[TS] tssmooth dexpontial .......................................................... Double-exponential smoothing
[TS] tssmooth exponential .......................................................... Single-exponential smoothing
[TS] tssmooth hwinters ............................................................. Holt–Winters nonseasonal smoothing
[TS] tssmooth ma ................................................................. Moving-average filter
[TS] tssmooth nl ................................................................. Nonlinear filter
[TS] tssmooth shwinters ......................................................... Holt–Winters seasonal smoothing
[TS] ucm .......................................................... Unobserved-components model
[TS] wntestb ................................................................. Bartlett’s periodogram-based test for white noise
[TS] wntestq ........................................................... Portmanteau (Q) test for white noise
[TS] xcorr .......................................................... Cross-correlogram for bivariate time series

Transforms and normality tests

[R] boxcox .......................................................... Box–Cox regression models
[R] fp ................................................................. Fractional polynomial regression
[R] ladder ............................................................... Ladder of powers
[R] lnskew0 .......................................................... Find zero-skewness log or Box–Cox transform
[R] mfp ................................................................. Multivariable fractional polynomial models
[MV] mvtest normality ........................................................... Multivariate normality tests
[R] sktest .......................................................... Skewness and kurtosis test for normality
[R] swilk .......................................................... Shapiro–Wilk and Shapiro–Francia tests for normality

Treatment effects

[U] Section 26.23 .......................................................... Treatment-effect models
[TE] eteffects .......................................................... Endogenous treatment-effects estimation
[TE] etpoisson ........................................................... Poisson regression with endogenous treatment effects
[TE] etregress .......................................................... Linear regression with endogenous treatment effects
[TE] stteffects .......................................................... Treatment-effects estimation for observational survival-time data
[TE] stteffects intro .......................................................... Introduction to treatment effects for observational survival-time data
[TE] stteffects ipw .......................................................... Survival-time inverse-probability weighting
[TE] stteffects ipwra .......................................................... Survival-time inverse-probability-weighted regression adjustment
[TE] stteffects ra .......................................................... Survival-time regression adjustment
[TE] stteffects wra .......................................................... Survival-time weighted regression adjustment
[TE] tebalance .......................................................... Check balance after teffects or stteffects estimation
[TE] tebalance box .......................................................... Covariate balance box
[TE] tebalance density .......................................................... Covariate balance density
[TE] tebalance overid .......................................................... Test for covariate balance
[TE] tebalance summarize .......................................................... Covariate-balance summary statistics
[TE] teffects .......................................................... Treatment-effects estimation for observational data
[TE] teffects aipw .......................................................... Augmented inverse-probability weighting
[TE] teffects intro .......................................................... Introduction to treatment effects for observational data
[TE] teffects intro advanced .......................................................... Advanced introduction to treatment effects for observational data
[TE] teffects ipw .......................................................... Inverse-probability weighting
[TE] teffects ipwra .......................................................... Inverse-probability-weighted regression adjustment
[TE] teffects multivalued .......................................................... Multivalued treatment effects
[TE] teffects nnmatch .......................................................... Nearest-neighbor matching
[TE] teffects overlap .......................................................... Overlap plots
[TE] teffects psmatch .......................................................... Propensity-score matching
[TE] teffects ra .......................................................... Regression adjustment
[TE] treatment effects .......................................................... Introduction to treatment-effects commands
## Matrix commands

### Basics
- **Chapter 14**. Matrix expressions
- **matlist**. Display a matrix and control its format
- **matrix**. Introduction to matrix commands
- **matrix define**. Matrix definition, operators, and functions
- **matrix utility**. List, rename, and drop matrices

### Programming
- **ereturn**. Post the estimation results
- **matrix accum**. Form cross-product matrices
- **matrix rownames**. Name rows and columns
- **matrix score**. Score data from coefficient vectors
- **ml**. Maximum likelihood estimation

### Other
- **makecns**. Constrained estimation
- **matrix dissimilarity**. Compute similarity or dissimilarity measures
- **matrix eigenvalues**. Eigenvalues of nonsymmetric matrices
- **matrix get**. Access system matrices
- **matrix mkmat**. Convert variables to matrix and vice versa
- **matrix svd**. Singular value decomposition
- **matrix symeigen**. Eigenvalues and eigenvectors of symmetric matrices

### Mata
- **putmata**. Put Stata variables into Mata and vice versa

## Programming

### Basics
- **Chapter 18**. Programming Stata
- **Section 18.3**. Macros
- **Section 18.11**. Ado-files
- **comments**. Add comments to programs
- **fvexpand**. Expand factor varlists
- **macro**. Macro definition and manipulation
- **program**. Define and manipulate programs
- **return**. Return stored results

### Program control
- **Section 18.11.1**. Version
- **capture**. Capture return code
- **continue**. Break out of loops
- **error**. Display generic error message and exit
- **foreach**. Loop over items
- **forvalues**. Loop over consecutive values
<table>
<thead>
<tr>
<th>Description</th>
<th>Section/Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parsing and program arguments</td>
<td></td>
</tr>
<tr>
<td>if programming command</td>
<td>Section 18.4</td>
</tr>
<tr>
<td>Version control</td>
<td>Version</td>
</tr>
<tr>
<td>Looping</td>
<td></td>
</tr>
<tr>
<td>Confirm</td>
<td>Argument verification</td>
</tr>
<tr>
<td>Low-level parsing</td>
<td></td>
</tr>
<tr>
<td>Levels of variable</td>
<td></td>
</tr>
<tr>
<td>Parse numeric lists</td>
<td></td>
</tr>
<tr>
<td>Parse Stata syntax</td>
<td></td>
</tr>
<tr>
<td>Divide strings into tokens</td>
<td></td>
</tr>
<tr>
<td>Console output</td>
<td></td>
</tr>
<tr>
<td>Section 12.4.2 Handling Unicode strings</td>
<td></td>
</tr>
<tr>
<td>Dialog programming</td>
<td></td>
</tr>
<tr>
<td>Display</td>
<td>Display strings and values of scalar expressions</td>
</tr>
<tr>
<td>Stata Markup and Control Language</td>
<td></td>
</tr>
<tr>
<td>Display tables</td>
<td></td>
</tr>
<tr>
<td>Unicode utilities</td>
<td></td>
</tr>
<tr>
<td>Commonly used programming commands</td>
<td></td>
</tr>
<tr>
<td>byable</td>
<td>Make programs byable</td>
</tr>
<tr>
<td>Change delimiter</td>
<td></td>
</tr>
<tr>
<td>Exit</td>
<td>Exit from a program or do-file</td>
</tr>
<tr>
<td>Factor-variables operator programming command</td>
<td></td>
</tr>
<tr>
<td>Mark observations for inclusion</td>
<td></td>
</tr>
<tr>
<td>Introduction to matrix commands</td>
<td></td>
</tr>
<tr>
<td>Pause until key is pressed</td>
<td></td>
</tr>
<tr>
<td>Preserve until key is pressed</td>
<td></td>
</tr>
<tr>
<td>Preserve and restore data</td>
<td></td>
</tr>
<tr>
<td>Quietly and noisily perform Stata command</td>
<td></td>
</tr>
<tr>
<td>Scalar variables</td>
<td></td>
</tr>
<tr>
<td>Stata Markup and Control Language</td>
<td></td>
</tr>
<tr>
<td>Sort within programs</td>
<td></td>
</tr>
<tr>
<td>Time sections of code by recording and reporting time spent</td>
<td></td>
</tr>
<tr>
<td>Time-series operator programming command</td>
<td></td>
</tr>
<tr>
<td>Debugging</td>
<td></td>
</tr>
<tr>
<td>Program debugging command</td>
<td></td>
</tr>
<tr>
<td>Time sections of code by recording and reporting time spent</td>
<td></td>
</tr>
<tr>
<td>Debug Stata programs</td>
<td></td>
</tr>
<tr>
<td>Advanced programming commands</td>
<td></td>
</tr>
<tr>
<td>Sorting strings containing Unicode characters</td>
<td>Section 12.4.2.5</td>
</tr>
<tr>
<td>Create a PDF file</td>
<td></td>
</tr>
<tr>
<td>Generate Office Open XML (.docx) file</td>
<td></td>
</tr>
<tr>
<td>Automation</td>
<td></td>
</tr>
<tr>
<td>Suppress Break key</td>
<td></td>
</tr>
<tr>
<td>Characteristics</td>
<td></td>
</tr>
</tbody>
</table>
Special-interest programming commands

[R] bstat ............................................. Report bootstrap results
[MV] cluster programming subroutines ............. Add cluster-analysis routines
[MV] cluster programming utilities ................. Cluster-analysis programming utilities
[R] fvarextract .................................... Factor-variables operator programming command
[P] matrix dissimilarity ........................... Compute similarity or dissimilarity measures
[MI] mi select ....................................... Programmer’s alternative to mi extract
[ST] st_is ........................................... Survival analysis subroutines for programmers
[SVY] svymarkout . Mark observations for exclusion on the basis of survey characteristics
[MI] technical ........................................ Details for programmers
[TS] tsrevar ......................................... Time-series operator programming command

Projects

[P] Project Manager .................................. Organize Stata files
File formats

[P] file formats .dta ........................................ Description of .dta file format
[D] unicode convertfile ............................... Low-level file conversion between encodings
[D] unicode translate ................................. Translate files to Unicode

Mata


Interface features

[GS] Chapter 1 (GSM, GSU, GSW) ................. Introducing Stata—sample session
[GS] Chapter 2 (GSM, GSU, GSW) ...................... The Stata user interface
[GS] Chapter 3 (GSM, GSU, GSW) ....................... Using the Viewer
[GS] Chapter 6 (GSM, GSU, GSW) ....................... Using the Data Editor
[GS] Chapter 7 (GSM, GSU, GSW) ..................... Using the Variables Manager
[GS] Chapter 13 (GSM, GSU, GSW) ........ Using the Do-file Editor—automating Stata
[GS] Chapter 15 (GSM, GSU, GSW) ................ Editing graphs

[P] dialog programming .................................. Dialog programming
[R] doedit .................................................. Edit do-files and other text files
[D] edit ..................................................... Browse or edit data with Data Editor
[P] set locale_ui ........................................ Specify a localization package for the user interface
[P] sleep .................................................. Pause for a specified time
[P] smcl ................................................... Stata Markup and Control Language
[D] varmanage ............................................. Manage variable labels, formats, and other properties
[P] viewsource .......................................... Display open/save dialog box
[P] window manage ..................................... Manage window characteristics
[P] window push ........................................ Copy command into Review window
[P] window stopbox .................................... Display message box
### Acronym glossary

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2SIV</td>
<td>two-step instrumental variables</td>
</tr>
<tr>
<td>2SLS</td>
<td>two-stage least squares</td>
</tr>
<tr>
<td>3SLS</td>
<td>three-stage least squares</td>
</tr>
<tr>
<td>ADF</td>
<td>asymptotic distribution free</td>
</tr>
<tr>
<td>AF</td>
<td>attributable fraction for the population</td>
</tr>
<tr>
<td>AFE</td>
<td>attributable fraction among the exposed</td>
</tr>
<tr>
<td>AFT</td>
<td>accelerated failure time</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike information criterion</td>
</tr>
<tr>
<td>AIDS</td>
<td>almost-ideal demand system</td>
</tr>
<tr>
<td>AIPW</td>
<td>augmented inverse-probability weights</td>
</tr>
<tr>
<td>ANCOVA</td>
<td>analysis of covariance</td>
</tr>
<tr>
<td>ANOVA</td>
<td>analysis of variance</td>
</tr>
<tr>
<td>APE</td>
<td>average partial effects</td>
</tr>
<tr>
<td>AR</td>
<td>autoregressive</td>
</tr>
<tr>
<td>AR(1)</td>
<td>first-order autoregressive</td>
</tr>
<tr>
<td>ARCH</td>
<td>autoregressive conditional heteroskedasticity</td>
</tr>
<tr>
<td>ARFIMA</td>
<td>autoregressive fractionally integrated moving average</td>
</tr>
<tr>
<td>ARIMA</td>
<td>autoregressive integrated moving average</td>
</tr>
<tr>
<td>ARMA</td>
<td>autoregressive moving average</td>
</tr>
<tr>
<td>ARMAX</td>
<td>autoregressive moving-average exogenous</td>
</tr>
<tr>
<td>ASCII</td>
<td>American Standard Code for Information Interchange</td>
</tr>
<tr>
<td>ASE</td>
<td>asymptotic standard error</td>
</tr>
<tr>
<td>ASL</td>
<td>achieved significance level</td>
</tr>
<tr>
<td>ATE</td>
<td>average treatment effect</td>
</tr>
<tr>
<td>ATET</td>
<td>average treatment effect on the treated</td>
</tr>
<tr>
<td>AUC</td>
<td>area under the time-versus-concentration curve</td>
</tr>
<tr>
<td>BC</td>
<td>bias corrected</td>
</tr>
<tr>
<td>BCa</td>
<td>bias-corrected and accelerated</td>
</tr>
<tr>
<td>BCC</td>
<td>boundary characteristic curve</td>
</tr>
<tr>
<td>BE</td>
<td>between effects</td>
</tr>
<tr>
<td>BFGS</td>
<td>Broyden–Fletcher–Goldfarb–Shanno</td>
</tr>
<tr>
<td>BHHH</td>
<td>Berndt–Hall–Hall–Hausman</td>
</tr>
<tr>
<td>BIC</td>
<td>Bayesian information criterion</td>
</tr>
<tr>
<td>BLOB</td>
<td>binary large object</td>
</tr>
<tr>
<td>BLUP</td>
<td>best linear unbiased prediction</td>
</tr>
<tr>
<td>BRR</td>
<td>balanced repeated replication</td>
</tr>
<tr>
<td>Acronym</td>
<td>Glossary</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>CA</td>
<td>correspondence analysis</td>
</tr>
<tr>
<td>CCC</td>
<td>category characteristic curve</td>
</tr>
<tr>
<td>CCI</td>
<td>conservative confidence interval</td>
</tr>
<tr>
<td>CCT</td>
<td>controlled clinical trial</td>
</tr>
<tr>
<td>CD</td>
<td>coefficient of determination</td>
</tr>
<tr>
<td>CDC</td>
<td>Centers for Disease Control and Prevention</td>
</tr>
<tr>
<td>CDF</td>
<td>cumulative distribution function</td>
</tr>
<tr>
<td>CES</td>
<td>constant elasticity of substitution</td>
</tr>
<tr>
<td>CFA</td>
<td>confirmatory factor analysis</td>
</tr>
<tr>
<td>CFI</td>
<td>comparative fit index</td>
</tr>
<tr>
<td>CI</td>
<td>conditional independence</td>
</tr>
<tr>
<td>CIF</td>
<td>cumulative incidence function</td>
</tr>
<tr>
<td>CMI</td>
<td>conditional mean independence</td>
</tr>
<tr>
<td>CMLE</td>
<td>conditional maximum likelihood estimates</td>
</tr>
<tr>
<td>ct</td>
<td>count time</td>
</tr>
<tr>
<td>ctsum</td>
<td>cumulative sum</td>
</tr>
<tr>
<td>c.v.</td>
<td>coefficient of variation</td>
</tr>
<tr>
<td>DA</td>
<td>data augmentation</td>
</tr>
<tr>
<td>DDF</td>
<td>denominator degrees of freedom</td>
</tr>
<tr>
<td>DDFS</td>
<td>multiple denominator degrees of freedom</td>
</tr>
<tr>
<td>DEFF</td>
<td>design effect</td>
</tr>
<tr>
<td>DEFT</td>
<td>design effect (standard deviation metric)</td>
</tr>
<tr>
<td>DF</td>
<td>dynamic factor</td>
</tr>
<tr>
<td>df / d.f.</td>
<td>degree(s) of freedom</td>
</tr>
<tr>
<td>d.f.</td>
<td>distribution function</td>
</tr>
<tr>
<td>DFAR</td>
<td>dynamic factors with vector autoregressive errors</td>
</tr>
<tr>
<td>DFP</td>
<td>Davidon–Fletcher–Powell</td>
</tr>
<tr>
<td>DPD</td>
<td>dynamic panel data</td>
</tr>
<tr>
<td>EBCDIC</td>
<td>extended binary coded decimal interchange code</td>
</tr>
<tr>
<td>EGARCH</td>
<td>exponential GARCH</td>
</tr>
<tr>
<td>EGLS</td>
<td>estimated generalized least squares</td>
</tr>
<tr>
<td>EIM</td>
<td>expected information matrix</td>
</tr>
<tr>
<td>EM</td>
<td>expectation maximization</td>
</tr>
<tr>
<td>EPS</td>
<td>Encapsulated PostScript</td>
</tr>
<tr>
<td>ESS</td>
<td>error sum of squares</td>
</tr>
<tr>
<td>ESS</td>
<td>effective sample size</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>FCS</td>
<td>fully conditional specification</td>
</tr>
<tr>
<td>FD</td>
<td>first-differenced estimator</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>FE</td>
<td>fixed effects</td>
</tr>
<tr>
<td>FEVD</td>
<td>forecast-error variance decomposition</td>
</tr>
<tr>
<td>FGLS</td>
<td>feasible generalized least squares</td>
</tr>
<tr>
<td>FGNLS</td>
<td>feasible generalized nonlinear least squares</td>
</tr>
<tr>
<td>FIML</td>
<td>full information maximum likelihood</td>
</tr>
<tr>
<td>FIVE estimator</td>
<td>full-information instrumental-variables efficient estimator</td>
</tr>
<tr>
<td>flong</td>
<td>full long</td>
</tr>
<tr>
<td>flongsep</td>
<td>full long and separate</td>
</tr>
<tr>
<td>FMI</td>
<td>fraction of missing information</td>
</tr>
<tr>
<td>FP</td>
<td>fractional polynomial</td>
</tr>
<tr>
<td>FPC</td>
<td>finite population correction</td>
</tr>
<tr>
<td>GARCH</td>
<td>generalized autoregressive conditional heteroskedasticity</td>
</tr>
<tr>
<td>GEE</td>
<td>generalized estimating equations</td>
</tr>
<tr>
<td>GEV</td>
<td>generalized extreme value</td>
</tr>
<tr>
<td>GHK</td>
<td>Geweke–Hajivassiliou–Keane</td>
</tr>
<tr>
<td>GHQ</td>
<td>Gauss–Hermite quadrature</td>
</tr>
<tr>
<td>GLIM</td>
<td>generalized linear interactive modeling</td>
</tr>
<tr>
<td>GLLAMM</td>
<td>generalized linear latent and mixed models</td>
</tr>
<tr>
<td>GLM</td>
<td>generalized linear models</td>
</tr>
<tr>
<td>GLS</td>
<td>generalized least squares</td>
</tr>
<tr>
<td>GMM</td>
<td>generalized method of moments</td>
</tr>
<tr>
<td>GPCM</td>
<td>generalized partial credit model</td>
</tr>
<tr>
<td>GRM</td>
<td>graded response model</td>
</tr>
<tr>
<td>GSEM</td>
<td>generalized structural equation modeling/model</td>
</tr>
<tr>
<td>GUI</td>
<td>graphical user interface</td>
</tr>
<tr>
<td>HAC</td>
<td>heteroskedasticity- and autocorrelation-consistent</td>
</tr>
<tr>
<td>HR</td>
<td>hazard ratio</td>
</tr>
<tr>
<td>HRF</td>
<td>human readable form</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>IC</td>
<td>information criteria</td>
</tr>
<tr>
<td>ICC</td>
<td>item characteristic curve</td>
</tr>
<tr>
<td>ICD-9</td>
<td>International Classification of Diseases, Ninth Revision</td>
</tr>
<tr>
<td>ICD-10</td>
<td>International Classification of Diseases, Tenth Revision</td>
</tr>
<tr>
<td>ICU</td>
<td>International Components for Unicode</td>
</tr>
<tr>
<td>IIA</td>
<td>independence of irrelevant alternatives</td>
</tr>
<tr>
<td>i.i.d.</td>
<td>independent and identically distributed</td>
</tr>
<tr>
<td>IIF</td>
<td>item information function</td>
</tr>
<tr>
<td>IPW</td>
<td>inverse-probability weighting</td>
</tr>
<tr>
<td>IPWRA</td>
<td>inverse-probability-weighted regression adjustment</td>
</tr>
<tr>
<td>IQR</td>
<td>interquartile range</td>
</tr>
<tr>
<td>IR</td>
<td>incidence rate</td>
</tr>
<tr>
<td>IRF</td>
<td>impulse–response function</td>
</tr>
<tr>
<td>IRLS</td>
<td>iterated, reweighted least squares</td>
</tr>
<tr>
<td>IRR</td>
<td>incidence-rate ratio</td>
</tr>
<tr>
<td>IRT</td>
<td>item response theory</td>
</tr>
<tr>
<td>IV</td>
<td>instrumental variables</td>
</tr>
<tr>
<td>JAR</td>
<td>Java Archive file</td>
</tr>
<tr>
<td>JCA</td>
<td>joint correspondence analysis</td>
</tr>
<tr>
<td>JRE</td>
<td>Java Runtime Environment</td>
</tr>
<tr>
<td>LAPACK</td>
<td>linear algebra package</td>
</tr>
<tr>
<td>LAV</td>
<td>least absolute value</td>
</tr>
<tr>
<td>LDA</td>
<td>linear discriminant analysis</td>
</tr>
<tr>
<td>LIML</td>
<td>limited-information maximum likelihood</td>
</tr>
<tr>
<td>LM</td>
<td>Lagrange multiplier</td>
</tr>
<tr>
<td>LOO</td>
<td>leave one out</td>
</tr>
<tr>
<td>LOWESS</td>
<td>locally weighted scatterplot smoothing</td>
</tr>
<tr>
<td>LR</td>
<td>likelihood ratio</td>
</tr>
<tr>
<td>LSB</td>
<td>least-significant byte</td>
</tr>
<tr>
<td>MA</td>
<td>moving average</td>
</tr>
<tr>
<td>MAD</td>
<td>median absolute deviation</td>
</tr>
<tr>
<td>MANCOVA</td>
<td>multivariate analysis of covariance</td>
</tr>
<tr>
<td>MANOVA</td>
<td>multivariate analysis of variance</td>
</tr>
<tr>
<td>MAR</td>
<td>missing at random</td>
</tr>
<tr>
<td>MCA</td>
<td>multiple correspondence analysis</td>
</tr>
<tr>
<td>MCAGHQC</td>
<td>mode-curvature adaptive Gauss–Hermite quadrature</td>
</tr>
<tr>
<td>MCAR</td>
<td>missing completely at random</td>
</tr>
<tr>
<td>MCE</td>
<td>Monte Carlo error</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov chain Monte Carlo</td>
</tr>
<tr>
<td>MCSE</td>
<td>MCMC standard errors</td>
</tr>
<tr>
<td>MDES</td>
<td>minimum detectable effect size</td>
</tr>
<tr>
<td>MDS</td>
<td>multidimensional scaling</td>
</tr>
<tr>
<td>ME</td>
<td>multiple equation</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>MEFF</td>
<td>misspecification effect</td>
</tr>
<tr>
<td>MEFT</td>
<td>misspecification effect (standard deviation metric)</td>
</tr>
<tr>
<td>MFP</td>
<td>multivariable fractional polynomial</td>
</tr>
<tr>
<td>MI / mi</td>
<td>multiple imputation</td>
</tr>
<tr>
<td>midp</td>
<td>mid-$p$-value</td>
</tr>
<tr>
<td>MIMIC</td>
<td>multiple indicators and multiple causes</td>
</tr>
<tr>
<td>MINQUE</td>
<td>minimum norm quadratic unbiased estimation</td>
</tr>
<tr>
<td>MIVQUE</td>
<td>minimum variance quadratic unbiased estimation</td>
</tr>
<tr>
<td>ML</td>
<td>maximum likelihood</td>
</tr>
<tr>
<td>MLE</td>
<td>maximum likelihood estimate</td>
</tr>
<tr>
<td>MLMV</td>
<td>maximum likelihood with missing values</td>
</tr>
<tr>
<td>mlong</td>
<td>marginal long</td>
</tr>
<tr>
<td>MM</td>
<td>method of moments</td>
</tr>
<tr>
<td>MNAR</td>
<td>missing not at random</td>
</tr>
<tr>
<td>MNP</td>
<td>multinomial probit</td>
</tr>
<tr>
<td>MPL</td>
<td>modified profile likelihood</td>
</tr>
<tr>
<td>MS</td>
<td>mean square</td>
</tr>
<tr>
<td>MSAR</td>
<td>Markov-switching autoregression</td>
</tr>
<tr>
<td>MSB</td>
<td>most-significant byte</td>
</tr>
<tr>
<td>MSDR</td>
<td>Markov-switching dynamic regression</td>
</tr>
<tr>
<td>MSE</td>
<td>mean squared error</td>
</tr>
<tr>
<td>MSL</td>
<td>maximum simulated likelihood</td>
</tr>
<tr>
<td>MSS</td>
<td>model sum of squares</td>
</tr>
<tr>
<td>MUE</td>
<td>median unbiased estimates</td>
</tr>
<tr>
<td>MVAGHQ</td>
<td>mean–variance adaptive Gauss–Hermite quadrature</td>
</tr>
<tr>
<td>MVN</td>
<td>multivariate normal</td>
</tr>
<tr>
<td>MVREG</td>
<td>multivariate regression</td>
</tr>
<tr>
<td>NARCH</td>
<td>nonlinear ARCH</td>
</tr>
<tr>
<td>NHANES</td>
<td>National Health and Nutrition Examination Survey</td>
</tr>
<tr>
<td>NLS</td>
<td>nonlinear least squares</td>
</tr>
<tr>
<td>NPARCH</td>
<td>nonlinear power ARCH</td>
</tr>
<tr>
<td>NR</td>
<td>Newton–Raphson</td>
</tr>
<tr>
<td>NRM</td>
<td>nominal response model</td>
</tr>
<tr>
<td>ODBC</td>
<td>Open DataBase Connectivity</td>
</tr>
<tr>
<td>OIM</td>
<td>observed information matrix</td>
</tr>
<tr>
<td>OIRF</td>
<td>orthogonalized impulse–response function</td>
</tr>
<tr>
<td>OLE</td>
<td>Object Linking and Embedding (Microsoft product)</td>
</tr>
<tr>
<td>OLS</td>
<td>ordinary least squares</td>
</tr>
<tr>
<td>OPG</td>
<td>outer product of the gradient</td>
</tr>
<tr>
<td>OR</td>
<td>odds ratio</td>
</tr>
<tr>
<td>Acronym</td>
<td>Glossary</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>PA</td>
<td>population averaged</td>
</tr>
<tr>
<td>PARCH</td>
<td>power ARCH</td>
</tr>
<tr>
<td>PCA</td>
<td>principal component analysis</td>
</tr>
<tr>
<td>PCM</td>
<td>partial credit model</td>
</tr>
<tr>
<td>PCSE</td>
<td>panel-corrected standard error</td>
</tr>
<tr>
<td>p.d.f.</td>
<td>probability density function</td>
</tr>
<tr>
<td>PF</td>
<td>prevented fraction for the population</td>
</tr>
<tr>
<td>PFE</td>
<td>prevented fraction among the exposed</td>
</tr>
<tr>
<td>PH</td>
<td>proportional hazards</td>
</tr>
<tr>
<td>pk</td>
<td>pharmacokinetic data</td>
</tr>
<tr>
<td>p.m.f.</td>
<td>probability mass function</td>
</tr>
<tr>
<td>PMM</td>
<td>predictive mean matching</td>
</tr>
<tr>
<td>PNG</td>
<td>Portable Network Graphics</td>
</tr>
<tr>
<td>POM</td>
<td>potential-outcome means</td>
</tr>
<tr>
<td>PSS</td>
<td>power and sample size</td>
</tr>
<tr>
<td>PSU</td>
<td>primary sampling unit</td>
</tr>
<tr>
<td>QDA</td>
<td>quadratic discriminant analysis</td>
</tr>
<tr>
<td>QML</td>
<td>quasimaximum likelihood</td>
</tr>
<tr>
<td>RA</td>
<td>regression adjustment</td>
</tr>
<tr>
<td>rc</td>
<td>return code</td>
</tr>
<tr>
<td>RCT</td>
<td>randomized controlled trial</td>
</tr>
<tr>
<td>RE</td>
<td>random effects</td>
</tr>
<tr>
<td>REML</td>
<td>restricted (or residual) maximum likelihood</td>
</tr>
<tr>
<td>RESET</td>
<td>regression specification-error test</td>
</tr>
<tr>
<td>RMSE</td>
<td>root mean squared error</td>
</tr>
<tr>
<td>RMSEA</td>
<td>root mean squared error of approximation</td>
</tr>
<tr>
<td>RNG</td>
<td>random-number generator</td>
</tr>
<tr>
<td>ROC</td>
<td>receiver operating characteristic</td>
</tr>
<tr>
<td>ROP</td>
<td>rank-ordered probit</td>
</tr>
<tr>
<td>ROT</td>
<td>rule of thumb</td>
</tr>
<tr>
<td>RR</td>
<td>relative risk</td>
</tr>
<tr>
<td>RRR</td>
<td>relative-risk ratio</td>
</tr>
<tr>
<td>RSM</td>
<td>rating scale model</td>
</tr>
<tr>
<td>RSS</td>
<td>residual sum of squares</td>
</tr>
<tr>
<td>RUM</td>
<td>random utility maximization</td>
</tr>
<tr>
<td>RVI</td>
<td>relative variance increase</td>
</tr>
</tbody>
</table>
SAARCH: simple asymmetric ARCH
SARIMA: seasonal ARIMA
S.d.: standard deviation
SE / s.e.: standard error
SEM: structural equation modeling/model
SF: static factor
SFAR: static factors with vector autoregressive errors
SIF: Stata internal form
SIR: standardized incidence ratio
SJ: Stata Journal
SMCL: Stata Markup and Control Language
SMR: standardized mortality/morbidity ratio
SMSA: standard metropolitan statistical area
SOR: standardized odds ratio
SQL: Structured Query Language
SRD: standardized rate difference
SRMR: standardized root mean squared residual
SRR: standardized risk ratio
SRS: simple random sample/sampling
SRSWR: SRS with replacement
SSC: Statistical Software Components
SSCP: sum of squares and cross products
SSD: summary statistics data
SSU: secondary sampling unit
st: survival time
STB: Stata Technical Bulletin
STS: structural time series
SUR: seemingly unrelated regression
SURE: seemingly unrelated regression estimation
SUTVA: stable unit treatment value assumption
SVAR: structural vector autoregressive model
SVD: singular value decomposition
TAR: target acceptance rate
TARCH: threshold ARCH
TCC: test characteristic curve
TDT: transmission/disequilibrium test
TIF: test information function
TIFF: tagged image file format
TLI: Tucker–Lewis index
TSS: total sum of squares
UCA  Unicode Collation Algorithm
UCM  unobserved-components model
UI  user interface
UTF-8  Universal character set + Transformation Format—8-bit

VAR  vector autoregressive model
VAR(1)  first-order vector autoregressive
VARMA  vector autoregressive moving average
VARMA(1,1)  first-order vector autoregressive moving average
VCE  variance–covariance estimate
VECM  vector error-correction model
VIF  variance inflation factor

WLC  worst linear combination
WLF  worst linear function
WLS  weighted least squares
WNLS  weighted nonlinear least squares
wrt  with respect to

XML  Extensible Markup Language

ZINB  zero-inflated negative binomial
ZIP  zero-inflated Poisson
ZTNB  zero-truncated negative binomial
ZTP  zero-truncated Poisson
Glossary

1PL. See one-parameter logistic model.

1 : M matched case–control study. See matched study.

2PL. See two-parameter logistic model.

2 × 2 contingency table. A 2 × 2 contingency table is used to describe the association between a binary independent variable and a binary response variable of interest.

2 × 2 × K contingency table. See stratified 2 × 2 tables.

3PL. See three-parameter logistic model.

100% sample. See census.

a posteriori. In the context of Bayesian analysis, we use a posteriori to mean “after the sample is observed”. For example, a posteriori information is any information obtained after the data sample is observed. See posterior distribution, posterior.

a priori. In the context of Bayesian analysis, we use a priori to mean “before the sample is observed”. For example, a priori information is any information obtained before the data sample is observed. In a Bayesian model, a priori information about model parameters is specified by prior distributions.

ability. See latent trait.

accelerated failure-time model. A model in which everyone has, in a sense, the same survivor function, \( S(\tau) \), and an individual’s \(\tau_j \) is a function of his or her characteristics and of time, such as \(\tau_j = t \ast \exp(\beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j})\).

acceptance rate. In the context of the MH algorithm, acceptance rate is the fraction of the proposed samples that is accepted. The optimal acceptance rate depends on the properties of the target distribution and is not known in general. If the target distribution is normal, however, the optimal acceptance rate is known to be 0.44 for univariate distributions and 0.234 for multivariate distributions.

acceptance region. In hypothesis testing, an acceptance region is a set of sample values for which the null hypothesis cannot be rejected or can be accepted. It is the complement of the rejection region.

accrual period or recruitment period or accrual. The accrual period (or recruitment period) is the period during which subjects are being enrolled (recruited) into a study. Also see follow-up period.

actual alpha, actual significance level. This is an attained or observed significance level.

adaptation. In the context of the MH algorithm, adaptation refers to the process of tuning or adapting the proposal distribution to optimize the MCMC sampling. Typically, adaptation is performed periodically during the MCMC sampling. The bayesmh command performs adaptation every # of iterations as specified in option adaptation(every(#)) for a maximum of adaptation(maxiter()) iterations. In a continuous-adaptation regimes, the adaptation lasts during the entire process of the MCMC sampling. See [BAYES] bayesmh.

adaptation period. Adaptation period includes all MH adaptive iterations. It equals the length of the adaptation interval, as specified by adaptation(every()), times the maximum number of adaptations, adaptation(maxiter()).

adaptive iteration. In the adaptive MH algorithm, adaptive iterations are iterations during which adaptation is performed.
add factor. An add factor is a quantity added to an endogenous variable in a forecast model. Add factors can be used to incorporate outside information into a model, and they can be used to produce forecasts under alternative scenarios.

ADF, method(adf). ADF stands for asymptotic distribution free and is a method used to obtain fitted parameters for standard linear SEMs. ADF is used by sem when option method(adf) is specified. Other available methods are ML, QML, and MLMV.

administrative censoring. Administrative censoring is the right-censoring that occurs when the study observation period ends. All subjects complete the course of the study and are known to have experienced one of two outcomes at the end of the study: survival or failure. This type of censoring should not be confused with withdrawal and loss to follow-up. Also see censored, censoring, left-censoring, and right-censoring.

AFT, accelerated failure time. See accelerated failure-time model.

agglomerative hierarchical clustering methods. Agglomerative hierarchical clustering methods are bottom-up methods for hierarchical clustering. Each observation begins in a separate group. The closest pair of groups is agglomerated or merged in each iteration until all the data are in one cluster. This process creates a hierarchy of clusters. Contrast to divisive hierarchical clustering methods.

AIPW estimator. See augmented inverse-probability-weighted estimator.

Akaike information criterion, AIC. Akaike information criterion (AIC) is an information-based model-selection criterion. It is given by the formula $-2 \times \log \text{likelihood} + 2k$, where $k$ is the number of parameters. AIC favors simpler models by penalizing for the number of model parameters. It does not, however, account for the sample size. As a result, the AIC penalization diminishes as the sample size increases, as does its ability to guard against overparameterization.

allocation ratio. This ratio $n_2/n_1$ represents the number of subjects in the comparison, experimental group relative to the number of subjects in the reference, control group. Also see [PSS] unbalanced designs.

alpha. Alpha, $\alpha$, denotes the significance level.

alternative hypothesis. In hypothesis testing, the alternative hypothesis represents the counterpoint to which the null hypothesis is compared. When the parameter being tested is a scalar, the alternative hypothesis can be either one sided or two sided.

alternative value, alternative parameter. This value of the parameter of interest under the alternative hypothesis is fixed by the investigator in a power and sample-size analysis. For example, alternative mean value and alternative mean refer to a value of the mean parameter under the alternative hypothesis.

analysis of variance, ANOVA. This is a class of statistical models that studies differences between means from multiple populations by partitioning the variance of the continuous outcome into independent sources of variation due to effects of interest and random variation. The test statistic is then formed as a ratio of the expected variation due to the effects of interest to the expected random variation. Also see one-way ANOVA, two-way ANOVA, one-way repeated-measures ANOVA, and two-way repeated-measures ANOVA.

analysis time. Analysis time is like time, except that 0 has a special meaning: $t = 0$ is the time of onset of risk, the time when failure first became possible.

Analysis time is usually not what is recorded in a dataset. A dataset of patients might record calendar time. Calendar time must then be mapped to analysis time.
The letter $t$ is reserved for time in analysis-time units. The term \textit{time} is used for time measured in other units.

The \textit{origin} is the \textit{time} corresponding to $t = 0$, which can vary subject to subject. Thus $t = \text{time} - \text{origin}$.

\textbf{anchoring, anchor variable.} A variable is said to be the anchor of a latent variable if the path coefficient between the latent variable and the anchor variable is constrained to be 1. \textproc{sem} and \textproc{gsem} use anchoring as a way of normalizing latent variables and thus identifying the model.

\textbf{ANOVA denominator degrees of freedom (DDF) method.} This method uses the traditional ANOVA for computing DDF. According to this method, the DDF for a test of a fixed effect of a given variable depends on whether that variable is also included in any of the random-effects equations. For traditional ANOVA models with balanced designs, this method provides exact sampling distributions of the test statistics. For more complex mixed-effects models or with unbalanced data, this method typically leads to poor approximations of the actual sampling distributions of the test statistics.

\textbf{anti-image correlation matrix} or \textbf{anti-image covariance matrix}. The image of a variable is defined as that part which is predictable by regressing each variable on all the other variables; hence, the anti-image is the part of the variable that cannot be predicted. The anti-image correlation matrix $A$ is a matrix of the negatives of the partial correlations among variables. Partial correlations represent the degree to which the factors explain each other in the results. The diagonal of the anti-image correlation matrix is the Kaiser–Meyer–Olkin measure of sampling adequacy for the individual variables. Variables with small values should be eliminated from the analysis. The anti-image covariance matrix $C$ contains the negatives of the partial covariances and has one minus the squared multiple correlations in the principal diagonal. Most of the off-diagonal elements should be small in both anti-image matrices in a good factor model. Both anti-image matrices can be calculated from the inverse of the correlation matrix $R$ via

$$A = \{\text{diag}(R)\}^{-1}R\{\text{diag}(R)\}^{-1}$$

$$C = \{\text{diag}(R)\}^{-1/2}R\{\text{diag}(R)\}^{-1/2}$$

Also see \textit{Kaiser–Meyer–Olkin measure of sampling adequacy}.

\textbf{approximation denominator degrees of freedom (DDF) methods.} The Kenward–Roger and Satterthwaite DDF methods are referred to as approximation methods because they approximate the sampling distributions of test statistics using $t$ and $F$ distributions with the DDF specific to the method for complicated mixed-effects models and for simple mixed models with unbalanced data. Also see \textit{exact denominator degrees of freedom (DDF) methods}.

\textbf{arbitrary missing pattern.} Any missing-value pattern. Some imputation methods are suitable only when the pattern of missing values is special, such as a \textit{monotone-missing pattern}. An imputation method suitable for use with an arbitrary missing pattern may be used regardless of the pattern.

\textbf{ARCH model.} An autoregressive conditional heteroskedasticity (ARCH) model is a regression model in which the conditional variance is modeled as an autoregressive (AR) process. The ARCH($m$) model is

$$y_t = \mathbf{x}_t\beta + \epsilon_t$$

$$E(\epsilon_t^2|\epsilon_{t-1}^2, \epsilon_{t-2}^2, \ldots) = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \ldots + \alpha_m\epsilon_{t-m}^2$$

where $\epsilon_t$ is a white-noise error term. The equation for $y_t$ represents the conditional mean of the process, and the equation for $E(\epsilon_t^2|\epsilon_{t-1}^2, \epsilon_{t-2}^2, \ldots)$ specifies the conditional variance as an autoregressive function of its past realizations. Although the conditional variance changes over time, the unconditional variance is time invariant because $y_t$ is a stationary process. Modeling the conditional variance as an AR process raises the implied unconditional variance, making this model particularly appealing to researchers modeling fat-tailed data, such as financial data.
**Arellano–Bond estimator.** The Arellano–Bond estimator is a generalized method of moments (GMM) estimator for linear dynamic panel-data models that uses lagged levels of the endogenous variables as well as first differences of the exogenous variables as instruments. The Arellano–Bond estimator removes the panel-specific heterogeneity by first-differencing the regression equation.

**ARFIMA model.** An autoregressive fractionally integrated moving-average (ARFIMA) model is a time-series model suitable for use with long-memory processes. ARFIMA models generalize autoregressive integrated moving-average (ARIMA) models by allowing the differencing parameter to be a real number in (−0.5, 0.5) instead of requiring it to be an integer.

**arguments.** The values a function receives are called the function’s arguments. For instance, in `lud(A, L, U)`, `A`, `L`, and `U` are the arguments.

**ARIMA model.** An autoregressive integrated moving-average (ARIMA) model is a time-series model suitable for use with integrated processes. In an ARIMA\((p, d, q)\) model, the data is differenced \(d\) times to obtain a stationary series, and then an ARMA\((p, q)\) model is fit to this differenced data. ARIMA models that include exogenous explanatory variables are known as ARMAX models.

**ARMA model.** An autoregressive moving-average (ARMA) model is a time-series model in which the current period’s realization is the sum of an autoregressive (AR) process and a moving-average (MA) process. An ARMA\((p, q)\) model includes \(p\) AR terms and \(q\) MA terms. ARMA models with just a few lags are often able to fit data as well as pure AR or MA models with many more lags.

**ARMAX model.** An ARMAX model is a time-series model in which the current period’s realization is an ARMA process plus a linear function of a set of exogenous variables. Equivalently, an ARMAX model is a linear regression model in which the error term is specified to follow an ARMA process.

**array.** An array is any indexed object that holds other objects as elements. Vectors are examples of 1-dimensional arrays. Vector \(v\) is an array, and \(v[1]\) is its first element. Matrices are 2-dimensional arrays. Matrix \(X\) is an array, and \(X[1, 1]\) is its first element. In theory, one can have 3-dimensional, 4-dimensional, and higher arrays, although Mata does not directly provide them. See [M-2] subscripts for more information on arrays in Mata.

Arrays are usually indexed by sequential integers, but in associative arrays, the indices are strings that have no natural ordering. Associative arrays can be 1-dimensional, 2-dimensional, or higher. If \(A\) were an associative array, then \(A[“first”]\) might be one of its elements. See [M-5] asarray() for associative arrays in Mata.

**ASCII.** ASCII stands for American Standard Code for Information Interchange. It is a way of representing text and the characters that form text in computers. It can be divided into two sections: plain, or lower, ASCII, which includes numbers, punctuation, plain letters without diacritical marks, whitespace characters such as space and tab, and some control characters such as carriage return; and extended ASCII, which includes letters with diacritical marks as well as other special characters.

Before Stata 14, datasets, do-files, ado-files, and other Stata files were encoded using ASCII.

**at risk.** A subject is at risk from the instant the first failure event becomes possible and usually stays that way until failure, but a subject can have periods of being at risk and not at risk.

**ATE.** See average treatment effect.

**ATET.** See average treatment effect on the treated.

**attributable fraction.** An attributable fraction is the reduction in the risk of a disease or other condition of interest when a particular risk factor is removed.

**augmented inverse-probability-weighted estimator.** An augmented inverse-probability-weighted (AIPW) estimator is an inverse-probability-weighted estimator that includes an augmentation term that corrects the estimator when the treatment model is misspecified. When the treatment is correctly
specified, the augmentation term vanishes as the sample size becomes large. An AIPW estimator uses both an outcome model and a treatment model and is a doubly robust estimator.

**augmented regression.** Regression performed on the augmented data, the data with a few extra observations with small weights. The data are augmented in a way that prevents perfect prediction, which may arise during estimation of categorical data. See *The issue of perfect prediction during imputation of categorical data* under Remarks and examples of [MI] *mi impute*.

**autocorrelation function.** The autocorrelation function (ACF) expresses the correlation between periods $t$ and $t - k$ of a time series as function of the time $t$ and the lag $k$. For a stationary time series, the ACF does not depend on $t$ and is symmetric about $k = 0$, meaning that the correlation between periods $t$ and $t - k$ is equal to the correlation between periods $t$ and $t + k$.

**autoregressive process.** An autoregressive process is a time-series model in which the current value of a variable is a linear function of its own past values and a white-noise error term. A first-order autoregressive process, denoted as an AR(1) process, is $y_t = \rho y_{t-1} + \epsilon_t$. An AR($p$) model contains $p$ lagged values of the dependent variable.

An autoregressive processes can be extended to panel data. An AR(1) process in this is $y_{it} = \rho y_{i,t-1} + \epsilon_{it}$, where $i$ denotes panels, $t$ denotes time, and $\epsilon_{it}$ is white noise. In some applications, the parameter $\rho$ is written as $\rho_i$ and is allowed to differ across panels.

**average treatment effect.** The average treatment effect is the average among all individuals in a population.

**average treatment effect on the treated.** The average treatment effect on the treated is the average among those individuals who actually get the treatment.

**average-linkage clustering.** Average-linkage clustering is a hierarchical clustering method that uses the average proximity of observations between groups as the proximity measure between the two groups.

**balanced data.** A longitudinal or panel dataset is said to be balanced if each panel has the same number of observations. See also weakly balanced and strongly balanced.

**balanced design.** A balanced design represents an experiment in which the numbers of treated and untreated subjects are equal. For many types of two-sample hypothesis tests, the power of the test is maximized with balanced designs.

**balanced repeated replication.** Balanced repeated replication (BRR) is a method of variance estimation for designs with two PSUs in every stratum. The BRR variance estimator tends to give more reasonable variance estimates for this design than does the linearized variance estimator, which can result in large values and undesirably wide confidence intervals. The BRR variance estimator is described in [SVY] *variance estimation*.

**band-pass filter.** Time-series filters are designed to pass or block stochastic cycles at specified frequencies. Band-pass filters, such as those implemented in *tsfilter bk* and *tsfilter cf*, pass through stochastic cycles in the specified range of frequencies and block all other stochastic cycles.

**baseline.** In survival analysis, baseline is the state at which the covariates, usually denoted by the row vector $\mathbf{x}$, are zero. For example, if the only measured covariate is systolic blood pressure, the baseline survivor function would be the survivor function for someone with zero systolic blood pressure. This may seem ridiculous, but covariates are usually centered so that the mathematical definition of baseline (covariate is zero) translates into something meaningful (mean systolic blood pressure).

**baseline model.** A baseline model is a covariance model—a model of fitted means and covariances of observed variables without any other paths—with most of the covariances constrained to 0. That
is, a baseline model is a model of fitted means and variances but typically not all the covariances. Also see saturated model. Baseline models apply only to standard linear SEMs.

**batch means.** Batch means are means obtained from batches of sample values of equal size. Batch means provide an alternative method for estimating MCMC standard errors (MCSE). The batch size is usually chosen to minimize the correlation between different batches of means.

**Bayes factor.** Bayes factor is given by the ratio of the marginal likelihoods of two models, $M_1$ and $M_2$. It is a widely used criterion for Bayesian model comparison. Bayes factor is used in calculating the posterior odds ratio of model $M_1$ versus $M_2$,

$$
P(M_1|y) = \frac{P(y|M_1) P(M_1)}{P(y|M_2) P(M_2)}
$$

where $P(M_i|y)$ is a posterior probability of model $M_i$, and $P(M_i)$ is a prior probability of model $M_i$. When the two models are equally likely, that is, when $P(M_1) = P(M_2)$, the Bayes factor equals the posterior odds ratio of the two models.

**Bayes’s rule.** The Bayes’s rule is a formal method for relating conditional probability statements. For two (random) events $X$ and $Y$, the Bayes’s rule states that

$$
P(X|Y) \propto P(Y|X)P(X)
$$

that is, the probability of $X$ conditional on $Y$ is proportional to the probability of $X$ and the probability of $Y$ conditional on $X$. In Bayesian analysis, the Bayes’s rule is used for combining prior information about model parameters and evidence from the observed data to form the posterior distribution.

**Bayes’s theorem.** Bayes’s theorem states that the probability of an event, $A$, conditional on another event, $B$, is generally different from the probability of $B$ conditional on $A$, although the two are related. Bayes’s theorem is that

$$
P(A|B) = \frac{P(B|A)P(A)}{P(B)}
$$

where $P(A)$ is the marginal probability of $A$, and $P(A|B)$ is the conditional probability of $A$ given $B$, and likewise for $P(B)$ and $P(B|A)$.

**Bayesian analysis.** Bayesian analysis is a statistical methodology that considers model parameters to be random quantities and estimates their posterior distribution by combining prior knowledge about parameters with the evidence from the observed data sample. Prior knowledge about parameters is described by prior distributions and evidence from the observed data is incorporated through a likelihood model. Using the Bayes’s rule, the prior distribution and the likelihood model are combined to form the posterior distribution of model parameters. The posterior distribution is then used for parameter inference, hypothesis testing, and prediction.

**Bayesian hypothesis testing.** Bayesian hypothesis testing computes probabilities of hypotheses conditional on the observed data. In contrast to the frequentist hypothesis testing, the Bayesian hypothesis testing computes the actual probability of a hypothesis $H$ by using the Bayes’s rule,

$$
P(H|y) \propto P(y|H)P(H)
$$

where $y$ is the observed data, $P(y|H)$ is the marginal likelihood of $y$ given $H$, and $P(H)$ is the prior probability of $H$. Two different hypotheses, $H_1$ and $H_2$, can be compared by simply comparing $P(H_1|y)$ to $P(H_2|y)$. 
Bayesian information criterion, BIC. The Bayesian information criterion (BIC), also known as
Schwarz criterion, is an information based criterion used for model selection in classical statistics.
It is given by the formula 
\[-0.5 \times \log \text{likelihood} + k \times \ln n,\]
where \(k\) is the number of parameters and \(n\) is the sample size. BIC favors simpler, in terms of complexity, models and it is more
conservative than AIC.

BCC. See boundary characteristic curve.

Bentler’s invariant pattern simplicity rotation. Bentler’s (1977) rotation maximizes the invariant
pattern simplicity. It is an oblique rotation that minimizes the criterion function

\[c(\Lambda) = - \log[||\Lambda^2||^2] + \log[\text{diag}\{(\Lambda^2)'\Lambda^2\}]\]

See Crawford–Ferguson rotation for a definition of \(\Lambda\). Also see oblique rotation.

the results in a series of matrices organized around how results are calculated. See [SEM] estat
framework.

beta. Beta, \(\beta\), denotes the probability of committing a type II error, namely, failing to reject the null
hypothesis even though it is false.

between estimator. The between estimator is a panel-data estimator that obtains its estimates by
running OLS on the panel-level means of the variables. This estimator uses only the between-panel
variation in the data to identify the parameters, ignoring any within-panel variation. For it to
be consistent, the between estimator requires that the panel-level means of the regressors be
uncorrelated with the panel-specific heterogeneity terms.

between matrix and within matrix. The between and within matrices are SSCP matrices that measure
the spread between groups and within groups, respectively. These matrices are used in multivariate
analysis of variance and related hypothesis tests: Wilks’s lambda, Roy’s largest root, Lawley–
Hotelling trace, and Pillai’s trace.

Here we have \(k\) independent random samples of size \(n\). The between matrix \(H\) is given by

\[H = n \sum_{i=1}^{k} (\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet})(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet})' = \sum_{i=1}^{k} \frac{1}{n} y_{i\bullet}y_{i\bullet}' - \frac{1}{kn} y_{\bullet\bullet}y_{\bullet\bullet}'\]

The within matrix \(E\) is defined as

\[E = \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{i\bullet})(y_{ij} - \bar{y}_{i\bullet})' = \sum_{i=1}^{k} \sum_{j=1}^{n} y_{ij}y_{ij}' - \sum_{i=1}^{k} \frac{1}{n} y_{i\bullet}y_{i\bullet}'\]

Also see SSCP matrix.

between–within denominator degrees of freedom (DDF) method. See repeated denominator degrees
of freedom (DDF) method.

between-subjects design. This is an experiment that has only between-subjects factors. See
[PSS] power oneway and [PSS] power twoway.

between-subjects factor. This is a factor for which each subject receives only one of the levels.

binary 0. Binary 0, also known as the null character, is traditionally used to indicate the end of a
string, such as an ASCII or UTF-8 string.
Binary 0 is obtained by using `char(0)` and is sometimes displayed as \(\backslash 0\). See [U] 12.4.10 `strL` variables and binary strings for more information.

**binary item.** A binary item is an item that is scored as either 0 or 1.

**binary operator.** A binary operator is an operator applied to two arguments. In `2-3`, the minus sign is a binary operator, as opposed to the minus sign in `(−9)`, which is a unary operator.

**binary string.** A binary string is, technically speaking, any string that does not contain text. In Stata, however, a string is only marked as binary if it contains binary 0, or if it contains the contents of a file read in using the `fileread()` function, or if it is the result of a string expression containing a string that has already been marked as binary.

In Stata, `strL` variables, string scalars, and Mata strings can store binary strings. See [U] 12.4.10 `strL` variables and binary strings for more information.

**binomial test.** A binomial test is a test for which the exact sampling distribution of the test statistic is binomial; see [R] `bitest`. Also see [PSS] power oneproportion.

**biplot.** A biplot is a scatterplot which represents both observations and variables simultaneously. There are many different biplots; variables in biplots are usually represented by arrows and observations are usually represented by points.

**biquartimax rotation** or **biquartimin rotation.** Biquartimax rotation and biquartimin rotation are synonyms. They put equal weight on the varimax and quartimax criteria, simplifying the columns and rows of the matrix. This is an oblique rotation equivalent to an oblimin rotation with \(γ = 0.5\). Also see varimax rotation, quartimax rotation, and oblimin rotation.

**bisection method.** This method finds a root \(x\) of a function \(f(x)\) such that \(f(x) = 0\) by repeatedly subdividing an interval on which \(f(x)\) is defined until the change in successive root estimates is within the requested tolerance and function \(f(·)\) evaluated at the current estimate is sufficiently close to zero.

**BLOB.** BLOB is database jargon for binary large object. In Stata, BLOBs can be stored in `strL`s. Thus `strL`s can contain BLOBs such as Word documents, JPEG images, or anything else. See `strL`.

**blocking.** In the context of the MH algorithm, blocking refers to the process of separating model parameters into different subsets or blocks to be sampled independently of each other. MH algorithm generates proposals and applies the acceptance–rejection rule sequentially for each block. It is recommended that correlated parameters are kept in one block. Separating less-correlated or independent model parameters in different blocks may improve the mixing of the MH algorithm.

**BLUPs.** BLUPs are best linear unbiased predictions of either random effects or linear combinations of random effects. In linear models containing random effects, these effects are not estimated directly but instead are integrated out of the estimation. Once the fixed effects and variance components have been estimated, you can use these estimates to predict group-specific random effects. These predictions are called BLUPs because they are unbiased and have minimal mean squared errors among all linear functions of the response.

**bootstrap.** The bootstrap is a method of variance estimation. The bootstrap variance estimator for survey data is described in [SVY] variance estimation.

**bootstrap, vce(bootstrap).** The bootstrap is a replication method for obtaining variance estimates. Consider an estimation method \(E\) for estimating \(θ\). Let \(\widehat{θ}\) be the result of applying \(E\) to dataset \(D\) containing \(N\) observations. The bootstrap is a way of obtaining variance estimates for \(\widehat{θ}\) from repeated estimates \(\widehat{θ}_1, \widehat{θ}_2, \ldots\), where each \(\widehat{θ}_i\) is the result of applying \(E\) to a dataset of size \(N\) drawn with replacement from \(D\). See [SEM] sem option method( ) and [R] bootstrap.
vce(bootstrap) is allowed with \texttt{sem} but not \texttt{gsem}. You can obtain bootstrap results by prefixing the \texttt{gsem} command with \texttt{bootstrap:}, but remember to specify \texttt{bootstrap}'s \texttt{cluster()} and \texttt{idcluster()} options if you are fitting a multilevel model. See \cite{SEM intro 9}.

**boundary characteristic curve.** A boundary characteristic curve (BCC) expresses the probability of transitioning across a given boundary threshold that separates the ordered item categories into two groups as a function of the latent trait.

**boundary kernel.** A boundary kernel is a special kernel used to smooth hazard functions in the boundaries of the data range. Boundary kernels are applied when the \texttt{epan2}, \texttt{biweight}, or \texttt{rectangle kernel()} is specified with \texttt{stcurve}, \texttt{hazard} or \texttt{sts graph}, \texttt{hazard}.

**boundary solution** or **Heywood solution.** See \cite{Heywood case}.

**broad type.** Two matrices are said to be of the same broad type if the elements in each are numeric, are string, or are pointers. Mata provides two numeric types, real and complex. The term \textit{broad type} is used to mask the distinction within numeric and is often used when discussing operators or functions. One might say, “The comma operator can be used to join the rows of two matrices of the same broad type,” and the implication of that is that one could join a real to a complex. The result would be complex. Also see \textit{type}, \textit{eltype}, and \textit{orgtype}.

**BRR.** See \textit{balanced repeated replication}.

**Builder.** The Builder is Stata’s graphical interface for building \texttt{sem} and \texttt{gsem} models. The Builder is also known as the SEM Builder. See \cite{SEM intro 2}, \cite{SEM Builder}, and \cite{SEM Builder, generalized}.

**burn-between period.** The number of iterations between two draws of an MCMC sequence such that these draws may be regarded as independent.

**burn-in period.** The burn-in period is the number of iterations it takes for an MCMC sequence to reach stationarity.

**byte.** Formally, a byte is eight binary digits (bits), the units used to record computer data. Each byte can also be considered as representing a value from 0 through 255. Do not confuse this with Stata’s \texttt{byte} variable storage type, which allows values from −127 to 100 to be stored. With regard to strings, all strings are composed of individual characters that are encoded using either one byte or several bytes to represent each character.

For example, in \textit{UTF-8}, the encoding system used by Stata, byte value 97 encodes “a”. Byte values 195 and 161 in sequence encode “á”.

**CA.** See \textit{correspondence analysis}.

**calibration.** The procedure of estimating parameters of an IRT model.

**canonical correlation analysis.** Canonical correlation analysis attempts to describe the relationships between two sets of variables by finding linear combinations of each so that the correlation between the linear combinations is maximized.

**canonical discriminant analysis.** Canonical linear discriminant analysis is LDA where describing how groups are separated is of primary interest. Also see \textit{linear discriminant analysis}.

**canonical link.** Corresponding to each family of distributions in a generalized linear model is a canonical link function for which there is a sufficient statistic with the same dimension as the number of parameters in the linear predictor. The use of canonical link functions provides the GLM with desirable statistical properties, especially when the sample size is small.

**canonical loadings.** The canonical loadings are coefficients of canonical linear discriminant functions. Also see \textit{canonical discriminant analysis} and \textit{loading}.
canonical variate set. The canonical variate set is a linear combination or weighted sum of variables obtained from canonical correlation analysis. Two sets of variables are analyzed in canonical correlation analysis. The first canonical variate of the first variable set is the linear combination in standardized form that has maximal correlation with the first canonical variate from the second variable set. The subsequent canonical variates are uncorrelated to the previous and have maximal correlation under that constraint.

case–control studies. In case–control studies, cases meeting a fixed criterion are matched to noncases ex post to study differences in possible covariates. Relative sample sizes are usually fixed at 1:1 or 1:2 but sometimes vary once the survey is complete. In any case, sample sizes do not reflect the distribution in the underlying population.

casewise deletion. See listwise deletion.

categorical item. A categorical item is an item that is either ordinal or nominal.

category boundary curve. See boundary characteristic curve.

category boundary location. See difficulty.

category characteristic curve. A category characteristic curve (CCC) expresses the probability of a response in a given item category as a function of the latent trait.

category response function. See category characteristic curve.

cause-specific hazard. In a competing-risks analysis, the cause-specific hazard is the hazard function that generates the events of a given type. For example, if heart attack and stroke are competing events, then the cause-specific hazard for heart attacks describes the biological mechanism behind heart attacks independently of that for strokes. Cause-specific hazards can be modeled using Cox regression, treating the other events as censored.

CCC. See category characteristic curve.

c-conformability. Matrix, vector, or scalar $A$ is said to be c-conformable with matrix, vector, or scalar $B$ if they have the same number of rows and columns (they are p-conformable), or if they have the same number of rows and one is a vector, or if they have the same number of columns and one is a vector, or if one or the other is a scalar. $c$ stands for colon; c-conformable matrices are suitable for being used with Mata’s :op operators. $A$ and $B$ are c-conformable if and only if

<table>
<thead>
<tr>
<th></th>
<th>$A$</th>
<th>$B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r \times c$</td>
<td>$r \times c$</td>
<td>$r \times c$</td>
</tr>
<tr>
<td>$r \times 1$</td>
<td>$r \times c$</td>
<td>$r \times c$</td>
</tr>
<tr>
<td>$1 \times c$</td>
<td>$r \times c$</td>
<td>$r \times c$</td>
</tr>
<tr>
<td>$1 \times 1$</td>
<td>$r \times c$</td>
<td>$r \times 1$</td>
</tr>
<tr>
<td>$r \times c$</td>
<td>$1 \times c$</td>
<td>$r \times c$</td>
</tr>
<tr>
<td>$r \times c$</td>
<td>$1 \times 1$</td>
<td>$r \times 1$</td>
</tr>
</tbody>
</table>

The idea behind c-conformability is generalized elementwise operation. Consider $C = A:*B$. If $A$ and $B$ have the same number of rows and have the same number of columns, then $||C_{ij}|| = ||A_{ij}*B_{ij}||$. Now say that $A$ is a column vector and $B$ is a matrix. Then $||C_{ij}|| = ||A_i*B_{ij}||$: each element of $A$ is applied to the entire row of $B$. If $A$ is a row vector, each column of $A$ is applied to the entire column of $B$. If $A$ is a scalar, $A$ is applied to every element of $B$. And then all the rules repeat, with the roles of $A$ and $B$ interchanged. See [M-2] op_colon for a complete definition.
CCT. See *controlled clinical trial study*.

cell means. These are means of the outcome of interest within cells formed by the cross-classification of the two factors. See [PSS] *power twoway* and [PSS] *power repeated*.

cell-means model. A cell-means model is an ANOVA model formulated in terms of cell means.

censored, censoring, left-censoring, and right-censoring. An observation is left-censored when the exact time of failure is not known; it is merely known that the failure occurred before \( t_l \). Suppose that the event of interest is becoming employed. If a subject is already employed when first interviewed, his outcome is left-censored.

An observation is right-censored when the time of failure is not known; it is merely known that the failure occurred after \( t_r \). If a patient survives until the end of a study, the patient’s time of death is right-censored.

In common usage, censored without a modifier means right-censoring.

Also see *truncation, left-truncation, and right-truncation*.

census. When a census of the population is conducted, every individual in the population participates in the survey. Because of the time, cost, and other constraints, the data collected in a census are typically limited to items that can be quickly and easily determined, usually through a questionnaire.

centered data. Centered data has zero mean. You can center data \( x \) by taking \( x - \bar{x} \).

central posterior interval. See *equal-tailed credible interval*.

centroid-linkage clustering. Centroid-linkage clustering is a hierarchical clustering method that computes the proximity between two groups as the proximity between the group means.

CFA, CFA models. CFA stands for confirmatory factor analysis. It is a way of analyzing measurement models. CFA models is a synonym for measurement models.

chained equations. See *fully conditional specification*.

chi-squared test, \( \chi^2 \) test. This test for which either an asymptotic sampling distribution or a sampling distribution of a test statistic is \( \chi^2 \). See [PSS] *power onevariance* and [PSS] *power twoproportions*.

Cholesky ordering. Cholesky ordering is a method used to orthogonalize the error term in a VAR or VECM to impose a recursive structure on the dynamic model, so that the resulting impulse–response functions can be given a causal interpretation. The method is so named because it uses the Cholesky decomposition of the error-covariance matrix.

CI. CI is an abbreviation for confidence interval.

CI assumption. See *conditional-independence assumption*.

CIF. See *cumulative incidence function*.

class programming. See *object-oriented programming*.

classical scaling. Classical scaling is a method of performing MDS via an eigen decomposition. This is contrasted to modern MDS, which is achieved via the minimization of a loss function. Also see *multidimensional scaling* and *modern scaling*.

classification. Classification is the act of allocating or classifying observations to groups as part of discriminant analysis. In some sources, classification is synonymous with cluster analysis.

classification function. Classification functions can be obtained after LDA or QDA. They are functions based on Mahalanobis distance for classifying observations to the groups. See *discriminant function* for an alternative. Also see *linear discriminant analysis* and *quadratic discriminant analysis*.
classification table. A classification table, also known as a confusion matrix, gives the count of observations from each group that are classified into each of the groups as part of a discriminant analysis. The element at \((i, j)\) gives the number of observations that belong to the \(i\)th group but were classified into the \(j\)th group. High counts are expected on the diagonal of the table where observations are correctly classified, and small values are expected off the diagonal. The columns of the matrix are categories of the predicted classification; the rows represent the actual group membership.

clinical trial. A clinical trials is an experiment testing a medical treatment or procedure on human subjects.

clinally meaningful difference, clinically meaningful effect, clinically significant difference. Clinically meaningful difference represents the magnitude of an effect of interest that is of clinical importance. What is meant by “clinically meaningful” may vary from study to study. In clinical trials, for example, if no prior knowledge is available about the performance of the considered clinical procedure, a standardized effect size (adjusted for standard deviation) between 0.25 and 0.5 may be considered of clinical importance.

cluster. A cluster is a collection of individuals that are sampled as a group. Although the cost in time and money can be greatly decreased, cluster sampling usually results in larger variance estimates when compared with designs in which individuals are sampled independently.

cluster analysis. Cluster analysis is a method for determining natural groupings or clusters of observations.

cluster tree. See dendrogram.

clustered, vce(cluster clustvar). Clustered is the name we use for the generalized Huber/White/sandwich estimator of the VCE, which is the robust technique generalized to relax the assumption that errors are independent across observations to be that they are independent across clusters of observations. Within cluster, errors may be correlated.

Clustered standard errors are reported when sem or gsem option vce(cluster clustvar) is specified. The other available techniques are OIM, OPG, robust, bootstrap, and jackknife. Also available for sem only is EIM.

clustering. See cluster analysis.

Cochran–Armitage test. The Cochran–Armitage test is a test for a linear trend in a probability of response in a \(J \times 2\) contingency table. The test statistic has an asymptotic \(\chi^2\) distribution under the null hypothesis. See [PSS] power trend.


Cochrane–Orcutt estimator. This estimation is a linear regression estimator that can be used when the error term exhibits first-order autocorrelation. An initial estimate of the autocorrelation parameter \(\rho\) is obtained from OLS residuals, and then OLS is performed on the transformed data \(\tilde{y}_t = y_t - \rho y_{t-1}\) and \(\tilde{x}_t = x_t - \rho x_{t-1}\).

code pages. A code page maps extended ASCII values to a set of characters, typically for a specific language or set of languages. For example, the most commonly used code page is Windows-1252, which maps extended ASCII values to characters used in Western European languages. Code pages are essentially encodings for extended ASCII characters.

code point. A code point is the numerical value or position that represents a single character in a text system such as ASCII or Unicode. The original ASCII encoding system contains only 128 code points and thus can represent only 128 characters. Historically, the 128 additional bytes of extended ASCII have been encoded in many different and inconsistent ways to provide additional sets of 128 code points. The formal Unicode specification has 1,114,112 possible code points, of
which roughly 250,000 have been assigned to actual characters. Stata uses UTF-8 encoding for Unicode. Note that the UTF-8–encoded version of a code point does not have the same numeric value as the code point itself.

**coefficient of determination.** The coefficient of determination is the fraction (or percentage) of variation (variance) explained by an equation of a model. The coefficient of determination is thus like $R^2$ in linear regression.

**cohort studies.** In cohort studies, a group that is well defined is monitored over time to track the transition of noncases to cases. Cohort studies differ from incidence studies in that they can be retrospective as well as prospective.

**cointegrating vector.** A cointegrating vector specifies a stationary linear combination of nonstationary variables. Specifically, if each of the variables $x_1, x_2, \ldots, x_k$ is integrated of order one and there exists a set of parameters $\beta_1, \beta_2, \ldots, \beta_k$ such that $z_t = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$ is a stationary process, the variables $x_1, x_2, \ldots, x_k$ are said to be cointegrated, and the vector $\beta$ is known as a cointegrating vector.

**colon operators.** Colon operators are operators preceded by a colon, and the colon indicates that the operator is to be performed elementwise. $A : * B$ indicates element-by-element multiplication, whereas $A * B$ indicates matrix multiplication. Colons may be placed in front of any operator. Usually one thinks of elementwise as meaning $c_{ij} = a_{ij} <op> b_{ij}$, but in Mata, elementwise is also generalized to include c-conformability. See [M-2] op_colon.

**column stripes.** See row and column stripes.

**column-major order.** Matrices are stored as vectors. Column-major order specifies that the vector form of a matrix is created by stacking the columns. For instance,

```
: A
  1  2
  1  4
  2  5
  3  6
```

is stored as

```
  1  2  3  4  5  6
  1  2  3  4  5  6
```

in column-major order. The LAPACK functions use column-major order. Mata uses row-major order. See row-major order.

**colvector.** See vector, colvector, and rowvector.

**command language.** Stata’s sem and gsem command provide a way to specify SEMs. The alternative is to use the Builder to draw path diagrams; see [SEM] intro 2, [SEM] Builder, and [SEM] Builder, generalized.

**common factors.** Common factors are found by factor analysis. They linearly reconstruct the original variables. In factor analysis, reconstruction is defined in terms of prediction of the correlation matrix of the original variables.

**common odds ratio.** A measure of association in stratified $2 \times 2$ tables. It can be viewed as a weighted aggregate of stratum-specific odds ratios.

**communality.** Communality is the proportion of a variable’s variance explained by the common factors in factor analysis. It is also “1 – uniqueness”. Also see uniqueness.
comparison value. See alternative value.

competing risks. Competing risks models are survival-data models in which the failures are generated by more than one underlying process. For example, death may be caused by either heart attack or stroke. There are various methods for dealing with competing risks. One direct way is to duplicate failures for one competing risk as censored observations for the other risk and stratify on the risk type. Another is to directly model the cumulative incidence of the event of interest in the presence of competing risks. The former method uses stcox and the latter, stcrreg.

complementary log-log regression. Complementary log-log regression is a term for generalized linear response functions that are family Bernoulli, link cloglog. It is used for binary outcome data. Complementary log-log regression is also known in Stata circles as cloglog regression or just cloglog. See generalized linear response functions.

complete and incomplete observations. An observation in the \( m = 0 \) data is said to be complete if no imputed variable in the observation contains soft missing (.). Observations that are not complete are said to be incomplete.

complete data. Data that do not contain any missing values.

complete degrees of freedom. The degrees of freedom that would have been used for inference if the data were complete.

complete DF. See complete degrees of freedom.

complete-cases analysis. See listwise deletion.

completed data. See imputed data.

completed-data analysis. The analysis or estimation performed on the complete data, the data for which all values are observed. This term does not refer to analysis or estimation performed on the subset of complete observations. Do not confuse this with completed-data analysis.

completed-data analysis. The analysis or estimation performed on the made-to-be completed (imputed) data. This term does not refer to analysis or estimation performed on the subset of complete observations.

complete-linkage clustering. Complete-linkage clustering is a hierarchical clustering method that uses the farthest pair of observations between two groups to determine the proximity of the two groups.

complex. A matrix is said to be complex if its elements are complex numbers. Complex is one of two numeric types in Stata, the other being real. Complex is generally used to describe how a matrix is stored and not the kind of numbers that happen to be in it: complex matrix \( Z \) might happen to contain real numbers. Also see type, eltype, and orgtype.

component scores. Component scores are calculated after PCA. Component scores are the coordinates of the original variables in the space of principal components.

compound symmetry. A covariance matrix has a compound-symmetry structure if all the variances are equal and all the covariances are equal. This is a special case of the sphericity assumption.

Comrey’s tandem 1 and 2 rotations. Comrey (1967) describes two rotations, the first (tandem 1) to judge which “small” factors should be dropped, the second (tandem 2) for “polishing”.

Tandem principle 1 minimizes the criterion

\[
c(\Lambda) = \langle \Lambda^2, (\Lambda\Lambda')^2\Lambda^2 \rangle
\]
Tandem principle 2 minimizes the criterion

$$c(\Lambda) = \langle \Lambda^2, \{11' - (\Lambda\Lambda')^2\}\Lambda^2 \rangle$$

See Crawford–Ferguson rotation for a definition of $\Lambda$.

**concordant pairs.** In a $2 \times 2$ contingency table, a concordant pair is a pair of observations that are both either successes or failures. Also see discordant pairs and Introduction under Remarks and examples in [PSS] power pairedproportions.

**condition number.** The condition number associated with a numerical problem is a measure of that quantity’s amenability to digital computation. A problem with a low condition number is said to be well conditioned, whereas a problem with a high condition number is said to be ill conditioned. Sometimes reciprocals of condition numbers are reported and yet authors will still refer to them sloppily as condition numbers. Reciprocal condition numbers are often scaled between 0 and 1, with values near $\text{epsilon}(1)$ indicating problems.

**conditional conjugacy.** See semiconjugate prior.

**conditional fixed-effects model.** In general, including panel-specific dummies to control for fixed effects in nonlinear models results in inconsistent estimates. For some nonlinear models, the fixed-effect term can be removed from the likelihood function by conditioning on a sufficient statistic. For example, the conditional fixed-effect logit model conditions on the number of positive outcomes within each panel.

**conditional hazard function.** In the context of mixed- and random-effects survival models, the conditional hazard function is the hazard function computed conditionally on the random effects. Even within the same covariate pattern, the conditional hazard function varies among individuals who belong to different random-effects clusters.

**conditional hazard ratio.** In the context of mixed- and random effects survival models, the conditional hazard ratio is the ratio of two conditional hazard functions evaluated at different values of the covariates. Unless stated differently, the denominator corresponds to the conditional hazard function at baseline, that is, with all the covariates set to zero.

**conditional imputation.** Imputation performed using a conditional sample, a restricted part of the sample. Missing values outside the conditional sample are replaced with a conditional constant, the constant value of the imputed variable in the nonmissing observations outside the conditional sample. See Conditional imputation under Remarks and examples of [MI] mi impute.

**conditional independence.** The assumption that responses are not correlated after controlling for the latent trait.

**conditional mean.** The conditional mean expresses the average of one variable as a function of some other variables. More formally, the mean of $y$ conditional on $x$ is the mean of $y$ for given values of $x$; in other words, it is $E(y|x)$.

A conditional mean is also known as a regression or as a conditional expectation.

**conditional normality assumption.** See normality assumption, joint and conditional.

**conditional overdispersion.** In a negative binomial mixed-effects model, conditional overdispersion is overdispersion conditional on random effects. Also see overdispersion.

**conditional variance.** Although the conditional variance is simply the variance of a conditional distribution, in time-series analysis the conditional variance is often modeled as an autoregressive process, giving rise to ARCH models.
**conditional-independence assumption.** The conditional-independence assumption requires that the common variables that affect treatment assignment and treatment-specific outcomes be observable. The dependence between treatment assignment and treatment-specific outcomes can be removed by conditioning on these observable variables.

This assumption is also known as a selection-on-observables assumption because its central tenet is the observability of the common variables that generate the dependence.

**configuration.** The configuration in MDS is a representation in a low-dimensional (usually 2-dimensional) space with distances in the low-dimensional space approximating the dissimilarities or disparities in high-dimensional space. Also see multidimensional scaling, dissimilarity, and disparity.

**configuration plot.** A configuration plot after MDS is a (usually 2-dimensional) plot of labeled points showing the low-dimensional approximation to the dissimilarities or disparities in high-dimensional space. Also see multidimensional scaling, dissimilarity, and disparity.

**conformability.** Conformability refers to row-and-column matching between two or more matrices. For instance, to multiply $A \times B$, $A$ must have the same number of columns as $B$ has rows. If that is not true, then the matrices are said to be nonconformable (for multiplication).

Three kinds of conformability are often mentioned in the Mata documentation: $p$-conformability, $c$-conformability, and $r$-conformability.

**confounding.** In the analysis of epidemiological tables, factor or interaction effects are said to be confounded when the effect of one factor is combined with that of another. For example, the effect of alcohol consumption on esophageal cancer may be confounded with the effects of age, smoking, or both. In the presence of confounding, it is often useful to stratify on the confounded factors that are not of primary interest, in the above example, age and smoking.

**confusion matrix.** A confusion matrix is a synonym for a classification table after discriminant analysis. See classification table.

**conjugate.** If $z = a + bi$, the conjugate of $z$ is $\text{conj}(z) = a - bi$. The conjugate is obtained by reversing the sign of the imaginary part. The conjugate of a real number is the number itself.

**conjugate prior.** A prior distribution is conjugate for a family of likelihood distributions if the prior and posterior distributions belong to the same family of distributions. For example, the gamma distribution is a conjugate prior for the Poisson likelihood. Conjugacy may provide an efficient way of sampling from posterior distributions and is used in Gibbs sampling.

**conjugate transpose.** See transpose.

**constraints.** See parameter constraints.

**containment denominator degrees of freedom (DDF) method.** See ANOVA denominator degrees of freedom (DDF) method.

**continuous parameters.** Continuous parameters are parameters with continuous prior distributions.

**contrast or contrasts.** In ANOVA, a contrast in $k$ population means is defined as a linear combination

$$\delta = c_1 \mu_1 + c_2 \mu_2 + \cdots + c_k \mu_k$$

where the coefficients satisfy

$$\sum_{i=1}^{k} c_i = 0$$
In the multivariate setting (MANOVA), a contrast in $k$ population mean vectors is defined as

$$\delta = c_1 \mu_1 + c_2 \mu_2 + \cdots c_k \mu_k$$

where the coefficients again satisfy

$$\sum_{i=1}^{k} c_i = 0$$

The univariate hypothesis $\delta = 0$ may be tested with contrast (or test) after ANOVA. The multivariate hypothesis $\delta = 0$ may be tested with manovatest after MANOVA.

**control group.** A control group comprises subjects that are randomly assigned to a group where they receive no treatment or receives a standard treatment. In hypothesis testing, this is usually a reference group. Also see experimental group.

**controlled clinical trial study.** This is an experimental study in which treatments are assigned to two or more groups of subjects without the randomization.

**correlated uniqueness model.** A correlated uniqueness model is a kind of measurement model in which the errors of the measurements has a structured correlation. See [SEM intro 5].

**correlation structure.** A correlation structure is a set of assumptions imposed on the within-panel variance–covariance matrix of the errors in a panel-data model. See [XT xtgee] for examples of different correlation structures.

**correlogram.** A correlogram is a table or graph showing the sample autocorrelations or partial autocorrelations of a time series.

**correspondence analysis.** Correspondence analysis (CA) gives a geometric representation of the rows and columns of a two-way frequency table. The geometric representation is helpful in understanding the similarities between the categories of variables and associations between variables. CA is calculated by singular value decomposition. Also see singular value decomposition.

**correspondence analysis projection.** A correspondence analysis projection is a line plot of the row and column coordinates after CA. The goal of this graph is to show the ordering of row and column categories on each principal dimension of the analysis. Each principal dimension is represented by a vertical line; markers are plotted on the lines where the row and column categories project onto the dimensions. Also see correspondence analysis.

**costs.** Costs in discriminant analysis are the cost of misclassifying observations.

**counterfactual.** A counterfactual is an outcome a subject would have obtained had that subject received a different level of treatment. In the binary-treatment case, the counterfactual outcome for a person who received treatment is the outcome that person would have obtained had the person instead not received treatment; similarly, the counterfactual outcome for a person who did not receive treatment is the outcome that person would have obtained had the person received treatment.

Also see potential outcome.

**count-time data.** See ct data.

**covariance stationarity.** A process is covariance stationary if the mean of the process is finite and independent of $t$, the unconditional variance of the process is finite and independent of $t$, and the covariance between periods $t$ and $t-s$ is finite and depends on $t-s$ but not on $t$ or $s$ themselves. Covariance-stationary processes are also known as weakly stationary processes.

**covariance structure.** In a mixed-effects model, covariance structure refers to the variance–covariance structure of the random effects.
covariates. Covariates are the explanatory variables that appear in a model. For instance, if survival time were to be explained by age, sex, and treatment, then those variables would be the covariates. Also see time-varying covariates.

covarimin rotation. Covarimin rotation is an orthogonal rotation equivalent to varimax. Also see varimax rotation.

Crawford–Ferguson rotation. Crawford–Ferguson (1970) rotation is a general oblique rotation with several interesting special cases.

Special cases of the Crawford–Ferguson rotation include

<table>
<thead>
<tr>
<th>$\kappa$</th>
<th>Special case</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>quartimax / quartimin</td>
</tr>
<tr>
<td>$1/p$</td>
<td>varimax / covarimin</td>
</tr>
<tr>
<td>$f/(2p)$</td>
<td>equamax</td>
</tr>
<tr>
<td>$(f - 1)/(p + f - 2)$</td>
<td>parsimax</td>
</tr>
<tr>
<td>1</td>
<td>factor parsimony</td>
</tr>
</tbody>
</table>

$p = \text{number of rows of } \mathbf{A}.$

$f = \text{number of columns of } \mathbf{A}.$

Where $\mathbf{A}$ is the matrix to be rotated, $\mathbf{T}$ is the rotation and $\Lambda = \mathbf{A}\mathbf{T}$. The Crawford–Ferguson rotation is achieved by minimizing the criterion

$$c(\Lambda) = \frac{1}{4} - \frac{\kappa}{4} \langle \Lambda^2, \Lambda^2 (11' - I) \rangle + \frac{\kappa}{4} \langle \Lambda^2, (11' - I)\Lambda^2 \rangle$$

Also see oblique rotation.

credible interval. In Bayesian analysis, the credible interval of a scalar model parameter is an interval from the domain of the marginal posterior distribution of that parameter. Two types of credible intervals are typically used in practice: equal-tailed credible intervals and HPD credible intervals.

credible level. The credible level is a probability level between 0% and 100% used for calculating credible intervals in Bayesian analysis. For example, a 95% credible interval for a scalar parameter is an interval the parameter belongs to with the probability of 95%.

critical region. See rejection region.

critical value. In hypothesis testing, a critical value is a boundary of the rejection region.

cross-correlation function. The cross-correlation function expresses the correlation between one series at time $t$ and another series at time $t - k$ as a function of the time $t$ and lag $k$. If both series are stationary, the function does not depend on $t$. The function is not symmetric about $k = 0$: $\rho_{12}(k) \neq \rho_{12}(-k)$.

crossed variables or stacked variables. In CA and MCA crossed categorical variables may be formed from the interactions of two or more existing categorical variables. Variables that contain these interactions are called crossed or stacked variables.

crossed-effects model. A crossed-effects model is a mixed-effects model in which the levels of random effects are not nested. A simple crossed-effects model for cross-sectional time-series data would contain a random effect to control for panel-specific variation and a second random effect to control for time-specific random variation. Rather than being nested within panel, in this model a random effect due to a given time is the same for all panels.

crossed-random effects. See crossed-effects model.
crossing variables or stacking variables. In CA and MCA, crossing or stacking variables are the existing categorical variables whose interactions make up a crossed or stacked variable.

cross-sectional or prevalence studies. Cross-sectional studies sample distributions of healthy and diseased subjects in the population at one point in time.

cross-sectional data. Cross-sectional data refers to data collected over a set of individuals, such as households, firms, or countries sampled from a population at a given point in time.

cross-sectional study. This type of observational study measures various population characteristics at one point in time or over a short period of time. For example, a study of the prevalence of breast cancer in the population is a cross-sectional study.

cross-sectional time-series data. Cross-sectional time-series data is another name for panel data. The term cross-sectional time-series data is sometimes reserved for datasets in which a relatively small number of panels were observed over many periods. See also panel data.

crude estimate. A crude estimate has not been adjusted for the effects of other variables. Disregarding a stratification variable, for example, yields a crude estimate.

crude estimate. A crude estimate has not been adjusted for the effects of other variables. Disregarding a stratification variable, for example, yields a crude estimate.

cumulative hazard. See hazard, cumulative hazard, and hazard ratio.

cumulative incidence estimator. In a competing-risks analysis, the cumulative incidence estimator estimates the cumulative incidence function (CIF). Assume for now that you have one event of interest (type 1) and one competing event (type 2). The cumulative incidence estimator for type 1 failures is then obtained by

$$\hat{CIF}_1(t) = \sum_{j: t_j \leq t} \hat{h}_1(t_j) \hat{S}(t_{j-1})$$

with

$$\hat{S}(t) = \prod_{j: t_j \leq t} \left\{ 1 - \hat{h}_1(t_j) - \hat{h}_2(t_j) \right\}$$

The \( t_j \) index the times at which events (of any type) occur, and \( \hat{h}_1(t_j) \) and \( \hat{h}_2(t_j) \) are the cause-specific hazard contributions for type 1 and type 2, respectively. \( \hat{S}(t) \) estimates the probability that you are event free at time \( t \).

The above generalizes to multiple competing events in the obvious way.

cumulative incidence function. In a competing-risks analysis, the cumulative incidence function, or CIF, is the probability that you will observe the event of primary interest before a given time. Formally,

$$CIF(t) = P(T \leq t \text{ and event type of interest})$$

for time-to-failure, \( T \).

cumulative subhazard. See subhazard, cumulative subhazard, and subhazard ratio.

curse of dimensionality. The curse of dimensionality is a term coined by Richard Bellman (1961) to describe the problem caused by the exponential increase in size associated with adding extra dimensions to a mathematical space. On the unit interval, 10 evenly spaced points suffice to sample with no more distance than 0.1 between them; however a unit square requires 100 points, and a unit cube requires 1000 points. Many multivariate statistical procedures suffer from the curse of dimensionality. Adding variables to an analysis without adding sufficient observations can lead to imprecision.
curved path. See path.

cusum plot, CUSUM plot. The cusum (CUSUM) plot of an MCMC sample is a plot of cumulative sums of the differences between sample values and their overall mean against the iteration number. Cusum plots are useful graphical summaries for detecting early drifts in MCMC samples.

cyclical component. A cyclical component is a part of a time series that is a periodic function of time. Deterministic functions of time are deterministic cyclical components, and random functions of time are stochastic cyclical components. For example, fixed seasonal effects are deterministic cyclical components and random seasonal effects are stochastic seasonal components.

Random coefficients on time inside of periodic functions form an especially useful class of stochastic cyclical components; see [TS] ucm.

DA. See data augmentation.

data augmentation. An MCMC method used for the imputation of missing data.

data matrix. A dataset containing \( n \) observations on \( k \) variables in often stored in an \( n \times k \) matrix. An observation refers to a row of that matrix; a variable refers to a column. When the rows are observations and the columns are variables, the matrix is called a data matrix.

declarations. Declarations state the eltype and orgtype of functions, arguments, and variables. In

```plaintext
real matrix myfunc(real vector A, complex scalar B) {
    real scalar i
    ...
}
```

the real matrix is a function declaration, the real vector and complex scalar are argument declarations, and real scalar i is a variable declaration. The real matrix states the function returns a real matrix. The real vector and complex scalar state the kind of arguments myfunc() expects and requires. The real scalar i helps Mata to produce more efficient compiled code.

Declarations are optional, so the above could just as well have read

```plaintext
function myfunc(A, B) {
    ...
}
```

When you omit the function declaration, you must substitute the word function.

When you omit the other declarations, transmorphic matrix is assumed, which is fancy jargon for a matrix that can hold anything. The advantages of explicit declarations are that they reduce the chances you make a mistake either in coding or in using the function, and they assist Mata in producing more efficient code. Working interactively, most people omit the declarations.

See [M-2] declarations for more information.

defective matrix. An \( n \times n \) matrix is defective if it does not have \( n \) linearly independent eigenvectors.

DEFF and DEFT. DEFF and DEFT are design effects. Design effects compare the sample-to-sample variability from a given survey dataset with a hypothetical SRS design with the same number of individuals sampled from the population.

DEFF is the ratio of two variance estimates. The design-based variance is in the numerator; the hypothetical SRS variance is in the denominator.
DEFT is the ratio of two standard-error estimates. The design-based standard error is in the numerator; the hypothetical SRS with-replacement standard error is in the denominator. If the given survey design is sampled with replacement, DEFT is the square root of DEFF.

degree-of-freedom adjustment. In estimates of variances and covariances, a finite-population degree-of-freedom adjustment is sometimes applied to make the estimates unbiased.

Let’s write an estimated variance as $\hat{\sigma}_{ii}$ and write the “standard” formula for the variance as $\hat{\sigma}_{ii} = S_{ii}/N$. If $\hat{\sigma}_{ii}$ is the variance of observable variable $x_i$, it can readily be proven that $S_{ii}/N$ is a biased estimate of the variances in samples of size $N$ and that $S_{ii}/(N-1)$ is an unbiased estimate. It is usual to calculate variances using $S_{ii}/(N-1)$, which is to say, the “standard” formula has a multiplicative degree-of-freedom adjustment of $N/(N-1)$ applied to it.

If $\hat{\sigma}_{ii}$ is the variance of estimated parameter $\beta_i$, a similar finite-population degree-of-freedom adjustment can sometimes be derived that will make the estimate unbiased. For instance, if $\beta_i$ is a coefficient from a linear regression, an unbiased estimate of the variance of regression coefficient $\beta_i$ is $S_{ii}/(N-p-1)$, where $p$ is the total number of regression coefficients estimated excluding the intercept. In other cases, no such adjustment can be derived. Such estimators have no derivable finite-sample properties and one is left only with the assurances provided by its provable asymptotic properties. In such cases, the variance of coefficient $\beta_i$ is calculated as $S_{ii}/N$, which can be derived on theoretical grounds. SEM is an example of such an estimator.

SEM is a remarkably flexible estimator and can reproduce results that can sometimes be obtained by other estimators. SEM might produce asymptotically equivalent results, or it might produce identical results depending on the estimator. Linear regression is an example in which sem and gsem produce the same results as regress. The reported standard errors, however, will not look identical because the linear-regression estimates have the finite-population degree-of-freedom adjustment applied to them and the SEM estimates do not. To see the equivalence, you must undo the adjustment on the reported linear regression standard errors by multiplying them by $\sqrt{\frac{(N-p-1)}{N}}$.

delta. Delta, $\delta$, in the context of power and sample-size calculations, denotes the effect size.

delta method. See linearization.

dendrogram or cluster tree. A dendrogram or cluster tree graphically presents information about how observations are grouped together at various levels of (dis)similarity in hierarchical cluster analysis. At the bottom of the dendrogram, each observation is considered its own cluster. Vertical lines extend up for each observation, and at various (dis)similarity values, these lines are connected to the lines from other observations with a horizontal line. The observations continue to combine until, at the top of the dendrogram, all observations are grouped together. Also see hierarchical clustering.

dereference. Dereferencing is an action performed on pointers. Pointers contain memory addresses, such as 0x2a1228. One assumes something of interest is stored at 0x2a1228, say, a real scalar equal to 2. When one accesses that 2 via the pointer by coding $*p$, one is said to be dereferencing the pointer. Unary $*$ is the dereferencing operator.

design effects. See DEFF and DEFT.

deterministic trend. A deterministic trend is a deterministic function of time that specifies the long-run tendency of a time series.

deviance information criterion, DIC. The deviance information criterion (DIC) is an information based criterion used for Bayesian model selection. It is an analog of AIC and is given by the formula $D(\bar{\theta}) + 2 \times p_D$, where $D(\bar{\theta})$ is the deviance at the sample mean and $p_D$ is the effective complexity, a quantity equivalent to the number of parameters in the model. Models with smaller DIC are preferred.
DFBETA. A DFBETA measures the change in the regressor’s coefficient because of deletion of that subject. Also see partial DFBETA.

diagonal matrix. A matrix is diagonal if its off-diagonal elements are zero; $A$ is diagonal if $A[i,j] = 0$ for $i \neq j$. Usually, diagonal matrices are also square. Some definitions require that a diagonal matrix also be a square matrix.

diagonal of a matrix. The diagonal of a matrix is the set of elements $A[i,j]$.

dichotomous item. See binary item.

difference operator. The difference operator $\Delta$ denotes the change in the value of a variable from period $t - 1$ to period $t$. Formally, $\Delta y_t = y_t - y_{t-1}$, and $\Delta^2 y_t = \Delta(y_t - y_{t-1}) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}$.

difficulty. A level of the latent trait needed to pass an item or an item category.

dilation. A dilation stretches or shrinks distances in Procrustes rotation.

dimension. A dimension is a parameter or measurement required to define a characteristic of an object or observation. Dimensions are the variables in the dataset. Weight, height, age, blood pressure, and drug dose are examples of dimensions in health data. Number of employees, gross income, net income, tax, and year are examples of dimensions in data about companies.

diminishing adaptation. Diminishing adaptation of the adaptive algorithm is the type of adaptation in which the amount of adaptation decreases with the size of the MCMC chain.

direct, indirect, and total effects. Consider the following system of equations:

$$y_1 = b_{10} + b_{11}y_2 + b_{12}x_1 + b_{13}x_3 + e_1$$
$$y_2 = b_{20} + b_{21}y_3 + b_{22}x_1 + b_{23}x_4 + e_2$$
$$y_3 = b_{30} + b_{32}x_1 + b_{33}x_5 + e_3$$

The total effect of $x_1$ on $y_1$ is $b_{12} + b_{11}b_{22} + b_{11}b_{21}b_{32}$. It measures the full change in $y_1$ based on allowing $x_1$ to vary throughout the system.

The direct effect of $x_1$ on $y_1$ is $b_{12}$. It measures the change in $y_1$ caused by a change in $x_1$ holding other endogenous variables—namely, $y_2$ and $y_3$—constant.

The indirect effect of $x_1$ on $y_1$ is obtained by subtracting the total and direct effect and is thus $b_{11}b_{22} + b_{11}b_{21}b_{32}$.

direct standardization. Direct standardization is an estimation method that allows comparing rates that come from different frequency distributions.

Estimated rates (means, proportions, and ratios) are adjusted according to the frequency distribution from a standard population. The standard population is partitioned into categories called standard strata. The stratum frequencies for the standard population are called standard weights. The standardizing frequency distribution typically comes from census data, and the standard strata are most commonly identified by demographic information such as age, sex, and ethnicity.

directional test. See one-sided test.

disambiguation: characters, and bytes, and display columns. A character is simply the letter or symbol that you want to represent—the letter “a”, the punctuation mark “.”, or a Chinese logogram. A byte or sequence of bytes is how that character is stored in the computer. And, a display column is the space required to display one typical character in the fixed-width display used by Stata’s Results window and Viewer. Some characters are too wide for one display column. Each character is displayed in one or two display columns.
For plain ASCII characters, the number of characters always equals the number of bytes and equals the number of display columns.

For UTF-8 characters that are not plain ASCII, there are usually two bytes per character but there are sometimes three or even four bytes per character, such as for Chinese, Japanese, and Korean (CJK) characters. Characters that are too wide to fit in one display column (such as CJK characters) are displayed in two display columns.

In general, for Unicode characters, the relationship between the number of characters and the number of bytes and the relationship between the number of characters and the number of display columns is more ambiguous. All characters can be stored in four or fewer bytes and are displayed in Stata using two or fewer display columns.

See [U] 12.4.2.1 Unicode string functions and [U] 12.4.2.2 Displaying Unicode characters to learn how to deal with the distinction between characters, bytes, and display columns in your code.

discordant pairs. In a $2 \times 2$ contingency table, discordant pairs are the success-failure or failure-success pairs of observations. Also see concordant pairs and Introduction under Remarks and examples in [PSS] power pairedproportions.

discordant proportion. This is a proportion of discordant pairs or discordant sets. Also see Introduction under Remarks and examples in [PSS] power pairedproportions as well as Introduction under Remarks and examples in [PSS] power mcc.

discordant sets. In a matched study with multiple controls matched to a case, discordant sets are the sets in which there is any success–failure or failure–success match between the case and any matched control. Also see Introduction under Remarks and examples in [PSS] power mcc.

discrete parameters. Discrete parameters are parameters with discrete prior distributions.

discriminant analysis. Discriminant analysis is used to describe the differences between groups and to exploit those differences when allocating (classifying) observations of unknown group membership. Discriminant analysis is also called classification in many references.

discriminant function. Discriminant functions are formed from the eigenvectors from Fisher’s approach to LDA. See linear discriminant analysis. See classification function for an alternative.

discriminating variables. Discriminating variables in a discriminant analysis are analyzed to determine differences between groups where group membership is known. These differences between groups are then exploited when classifying observations to the groups.

discrimination. A measure of how well an item can distinguish between contiguous latent trait levels near the inflection point of an item characteristic curve.

disparity. Disparities are transformed dissimilarities, that is, dissimilarity values transformed by some function. The class of functions to transform dissimilarities to disparities may either be (1) a class of metric, or known functions such as linear functions or power functions that can be parameterized by real scalars or (2) a class of more general (nonmetric) functions, such as any monotonic function. Disparities are used in MDS. Also see dissimilarity, multidimensional scaling, metric scaling, and nonmetric scaling.

display column. A display column is the space required to display one typical character in the fixed-width display used by Stata’s Results window and Viewer. Some characters are too wide for one display column. Each character is displayed in one or two display columns.

All plain ASCII characters (for example, “M” and “9”) and many UTF-8 characters that are not plain ASCII (for example, “é”) require the same space when using a fixed-width font. That is to say, they all require a single display column.
Characters from non-Latin alphabets, such as Chinese, Cyrillic, Japanese, and Korean, may require two display columns.

See [U] 12.4.2.2 Displaying Unicode characters for more information.

dissimilarity, dissimilarity matrix, and dissimilarity measure. Dissimilarity or a dissimilarity measure is a quantification of the difference between two things, such as observations or variables or groups of observations or a method for quantifying that difference. A dissimilarity matrix is a matrix containing dissimilarity measurements. Euclidean distance is one example of a dissimilarity measure. Contrast to similarity. Also see proximity and Euclidean distance.

disturbance term. The disturbance term encompasses any shocks that occur to the dependent variable that cannot be explained by the conditional (or deterministic) portion of the model.

divisive hierarchical clustering methods. Divisive hierarchical clustering methods are top-down methods for hierarchical clustering. All the data begin as a part of one large cluster; with each iteration, a cluster is broken into two to create two new clusters. At the first iteration there are two clusters, then three, and so on. Divisive methods are very computationally expensive. Contrast to agglomerative hierarchical clustering methods.

doubly robust estimator. A doubly robust estimator only needs one of two auxiliary models to be correctly specified to estimate a parameter of interest.

Doubly robust estimators for treatment effects are consistent when either the outcome model or the treatment model is correctly specified.

drift. Drift is the constant term in a unit-root process. In

\[ y_t = \alpha + y_{t-1} + \epsilon_t \]

\( \alpha \) is the drift when \( \epsilon_t \) is a stationary, zero-mean process.

dropout. Dropout is the withdrawal of subjects before the end of a study and leads to incomplete or missing data.

dyadic operator. Synonym for binary operator.

dynamic forecast. A dynamic forecast uses forecast values wherever lagged values of the endogenous variables appear in the model, allowing one to forecast multiple periods into the future.

dynamic model. A dynamic model is one in which prior values of the dependent variable or disturbance term affect the current value of the dependent variable.

dynamic-multiplier function. A dynamic-multiplier function measures the effect of a shock to an exogenous variable on an endogenous variable. The \( k \)th dynamic-multiplier function of variable \( i \) on variable \( j \) measures the effect on variable \( j \) in period \( t + k \) in response to a one-unit shock to variable \( i \) in period \( t \), holding everything else constant.

EB. See empirical Bayes.

EE estimator. See estimating-equation estimator.

effect size. The effect size is the size of the clinically significant difference between the treatments being compared, often expressed as the hazard ratio (or the log of the hazard ratio) in survival analysis.

effective sample size, ESS. Effective sample size (ESS) is the MCMC sample size \( T \) adjusted for the autocorrelation in the sample. It represents the number of independent observations in an MCMC sample. ESS is used instead of \( T \) in calculating MCSE. Small ESS relative to \( T \) indicates high autocorrelation and consequently poor mixing of the chain.
**effect-size curve.** The effect-size curve is a graph of the estimated effect size or target parameter as a function of some other study parameter such as the sample size. The effect size or target parameter is plotted on the $y$ axis, and the sample size or other parameter is plotted on the $x$ axis.

**effect-size determination.** This pertains to the computation of an effect size or a target parameter given power, sample size, and other study parameters.

**efficiency.** In the context of MCMC, efficiency is a term used for assessing the mixing quality of an MCMC procedure. Efficient MCMC algorithms are able to explore posterior domains in less time (using fewer iterations). Efficiency is typically quantified by the sample autocorrelation and effective sample size. An MCMC procedure that generates samples with low autocorrelation and consequently high ESS is more efficient.

**eigenvalues and eigenvectors.** A scalar, $\lambda$, is said to be an eigenvalue of square matrix $A$: $n \times n$ if there is a nonzero column vector $x$: $n \times 1$ (called an eigenvector) such that

$$Ax = \lambda x$$

(1)

Equation (1) can also be written

$$(A - \lambda I)x = 0$$

where $I$ is the $n \times n$ identity matrix. A nontrivial solution to this system of $n$ linear homogeneous equations exists if and only if

$$\det(A - \lambda I) = 0$$

(2)

This $n$th-degree polynomial in $\lambda$ is called the characteristic polynomial or characteristic equation of $A$, and the eigenvalues $\lambda$ are its roots, also known as the characteristic roots.

The eigenvector defined by (1) is also known as the right eigenvector, because matrix $A$ is postmultiplied by eigenvector $x$. See [M-5] eigensystem() and left eigenvectors.

**EIM, vce(eim).** EIM stands for expected information matrix, defined as the inverse of the negative of the expected value of the matrix of second derivatives, usually of the log-likelihood function. The EIM is an estimate of the VCE. EIM standard errors are reported when `sem` option `vce(eim)` is specified. EIM is available only with `sem`. The other available techniques are OIM, OPG, robust, clustered, bootstrap, and jackknife.

**eltype.** See type, eltype, and orgtype.

**EM.** See expectation-maximization algorithm.

**empirical Bayes.** In generalized linear mixed-effects models, empirical Bayes refers to the method of prediction of the random effects after the model parameters have been estimated. The empirical Bayes method uses Bayesian principles to obtain the posterior distribution of the random effects, but instead of assuming a prior distribution for the model parameters, the parameters are treated as given.

In IRT models, empirical Bayes refers to the method of prediction of the latent trait after the model parameters have been estimated. The empirical Bayes method uses Bayesian principles to obtain the posterior distribution of the latent trait. However, instead of assuming a prior distribution for the model parameters, one treats the parameters as given.

**empirical Bayes mean.** See posterior mean.

**empirical Bayes mode.** See posterior mode.

**encodings.** An encoding is a way of representing a character as a byte or series of bytes. Examples of encoding systems are ASCII and UTF-8. Stata uses UTF-8 encoding.

For more information, see [U] 12.4.2.3 Encodings.
endogenous variable. An endogenous variable is a regressor that is correlated with the unobservable error term. Equivalently, an endogenous variable is one whose values are determined by the equilibrium or outcome of a structural model.

In the context of structural equation modeling and path diagrams, a variable, either observed or latent, is endogenous if any paths point to it.

Also see exogenous variable.

epsilon(1), etc. epsilon(1) refers to the unit roundoff error associated with a computer, also informally called machine precision. It is the smallest amount by which a number may differ from 1. For IEEE double-precision variables, epsilon(1) is approximately 2.22045e–16.

epsilon(x) is the smallest amount by which a real number can differ from x, or an approximation thereof; see [M-5] epsilon().

equal-allocation design. See balanced design.

equal-tailed credible interval. An equal-tailed credible interval is a credible interval defined in such a way that both tails of the marginal posterior distribution have the same probability. A \{100 \times (1 - \alpha)\}% equal-tailed credible interval is defined by the \(\alpha/2\)th and \((1 - \alpha)/2\)th quantiles of the marginal posterior distribution.

equamax rotation. Equamax rotation is an orthogonal rotation whose criterion is a weighted sum of the varimax and quartimax criteria. Equamax reflects a concern for simple structure within the rows and columns of the matrix. It is equivalent to oblimin with \(\gamma = p/2\), or to the Crawford–Ferguson family with \(\kappa = f/2p\), where \(p\) is the number of rows of the matrix to be rotated, and \(f\) is the number of columns. Also see orthogonal rotation, varimax rotation, quartimax rotation, oblimin rotation, and Crawford–Ferguson rotation.

error, error variable. The error is random disturbance \(e\) in a linear equation:

\[ y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + e \]

An error variable is an unobserved exogenous variable in path diagrams corresponding to \(e\). Mathematically, error variables are just another example of latent exogenous variables, but in \texttt{sem} and \texttt{gsem}, error variables are considered to be in a class by themselves. All (Gaussian) endogenous variables—observed and latent—have a corresponding error variable. Error variables automatically and inalterably have their path coefficients fixed to be 1. Error variables have a fixed naming convention in the software. If a variable is the error for (observed or latent) endogenous variable \(y\), then the residual variable’s name is \(e.y\).

In \texttt{sem} and \texttt{gsem}, error variables are uncorrelated with each other unless explicitly indicated otherwise. That indication is made in path diagrams by drawing a curved path between the error variables and is indicated in command notation by including \texttt{cov(e.name1*e.name2)} among the options specified on the \texttt{sem} command. In \texttt{gsem}, errors for family Gaussian, link log responses are not allowed to be correlated.

error-components model. The error-components model is another name for the random-effects model. See also random-effects model.

estimating-equation estimator. An estimating-equation (EE) estimator calculates parameters estimates by solving a system of equations. Each equation in this system is the sample average of a function that has mean zero.

These estimators are also known as \(M\) estimators or \(Z\) estimators in the statistics literature and as generalized method of moments (GMM) estimators in the econometrics literature.
estimation method. There are a variety of ways that one can solve for the parameters of an SEM. Different methods make different assumptions about the data-generation process, and so it is important that you choose a method appropriate for your model and data; see [SEM] intro 4.

Euclidean distance. The Euclidean distance between two observations is the distance one would measure with a ruler. The distance between vector $P = (P_1, P_2, \ldots, P_n)$ and $Q = (Q_1, Q_2, \ldots, Q_n)$ is given by

$$D(P, Q) = \sqrt{(P_1 - Q_1)^2 + (P_2 - Q_2)^2 + \cdots + (P_n - Q_n)^2} = \sqrt{\sum_{i=1}^{n} (P_i - Q_i)^2}$$

event. An event is something that happens at an instant in time, such as being exposed to an environmental hazard, being diagnosed as myopic, or becoming employed.

The failure event is of special interest in survival analysis, but there are other equally important events, such as the exposure event, from which analysis time is defined.

In st data, events occur at the end of the recorded time span.

event of interest. In a competing-risk analysis, the event of interest is the event that is the focus of the analysis, that for which the cumulative incidence in the presence of competing risks is estimated.

exact denominator degrees of freedom (DDF) methods. Residual, repeated, and ANOVA DDF methods are referred to as exact methods because they provide exact $t$ and $F$ sampling distributions of test statistics for special classes of mixed-effects models—linear regression, repeated-measures designs, and traditional ANOVA models—with balanced data. Also see [PSS] power oneprop and [PSS] power twoprop.

exact test. An exact test is one for which the probability of observing the data under the null hypothesis is calculated directly, often by enumeration. Exact tests do not rely on any asymptotic approximations and are therefore widely used with small datasets. See [PSS] power oneprop and [PSS] power twoprop.

exogenous variable. An exogenous variable is a regressor that is not correlated with any of the unobservable error terms in the model. Equivalently, an exogenous variable is one whose values change independently of the other variables in a structural model.

In the context of structural equation modeling and path diagrams, a variable, either observed or latent, is exogenous if paths only originate from it, or, equivalently, no paths point to it.

Also see endogenous variable.

exp. exp is used in syntax diagrams to mean “any valid expression may appear here”; see [M-2] exp.

expectation-maximization algorithm. In the context of MI, an iterative procedure for obtaining maximum likelihood or posterior-mode estimates in the presence of missing data.

experimental group. An experimental group is a group of subjects that receives a treatment or procedure of interest defined in a controlled experiment. In hypothesis testing, this is usually a comparison group. Also see control group.

experimental study. In an experimental study, as opposed to an observational study, the assignment of subjects to treatments is controlled by investigators. For example, a study that compares a new treatment with a standard treatment by assigning each treatment to a group of subjects is an experimental study.
exponential smoothing. Exponential smoothing is a method of smoothing a time series in which the smoothed value at period $t$ is equal to a fraction $\alpha$ of the series value at time $t$ plus a fraction $1 - \alpha$ of the previous period’s smoothed value. The fraction $\alpha$ is known as the smoothing parameter.

exponential test. The exponential test is the parametric test comparing the hazard rates, $\lambda_1$ and $\lambda_2$, (or log hazards) from two independent exponential (constant only) regression models with the null hypothesis $H_0$: $\lambda_2 - \lambda_1 = 0$ [or $H_0$: $\ln(\lambda_2) - \ln(\lambda_1) = \ln(\lambda_2/\lambda_1) = 0$].

exposure odds ratio. An odds ratio of exposure in cases relative to controls in a case–control study.

extended ASCII. Extended ASCII, also known as higher ASCII, is the byte values 128 to 255, which were not defined as part of the original ASCII specification. Various code pages have been defined over the years to map the extended ASCII byte values to many characters not supported in the original ASCII specification, such as Latin letters with diacritical marks, such as “á” and “Á”; non-Latin alphabets, such as Chinese, Cyrillic, Japanese, and Korean; punctuation marks used in non-English languages, such as “<”, complex mathematical symbols such as “±”, and more.

Although extended ASCII characters are stored in a single byte in ASCII encoding, UTF-8 stores the same characters in two to four bytes. Because each code page maps the extended ASCII values differently, another distinguishing feature of extended ASCII characters is that their meaning can change across fonts and operating systems.

external variable. See global variable.

f test. An $f$ test is a test for which a sampling distribution of a test statistic is an $F$ distribution. See [PSS] power twovariances.

factor. A factor is an unobserved random variable that is thought to explain variability among observed random variables.

factor analysis. Factor analysis is a statistical technique used to explain variability among observed random variables in terms of fewer unobserved random variables called factors. The observed variables are then linear combinations of the factors plus error terms.

If the correlation matrix of the observed variables is $R$, then $R$ is decomposed by factor analysis as

$$R = \Lambda \Phi \Lambda' + \Psi$$

$\Lambda$ is the loading matrix, and $\Psi$ contains the specific variances, for example, the variance specific to the variable not explained by the factors. The default unrotated form assumes uncorrelated common factors, $\Phi = I$.

factor loading plot. A factor loading plot produces a scatter plot of the factor loadings after factor analysis.

factor loadings. Factor loadings are the regression coefficients which multiply the factors to produce the observed variables in the factor analysis.

factor parsimony. Factor parsimony is an oblique rotation, which maximizes the column simplicity of the matrix. It is equivalent to a Crawford–Ferguson rotation with $\kappa = 1$. Also see oblique rotation and Crawford–Ferguson rotation.

factor scores. Factor scores are computed after factor analysis. Factor scores are the coordinates of the original variables, $x$, in the space of the factors. The two types of scoring are regression scoring (Thomson 1951) and Bartlett (1937, 1938) scoring.

Using the symbols defined in factor analysis, the formula for regression scoring is

$$\hat{f} = \Lambda' R^{-1} x$$
In the case of oblique rotation the formula becomes

$$\hat{f} = \Phi \Lambda' R^{-1} x$$

The formula for Bartlett scoring is

$$\hat{f} = \Gamma^{-1} \Lambda' \Psi^{-1} x$$

where

$$\Gamma = \Lambda' \Psi^{-1} \Lambda$$

Also see factor analysis.

**failure event.** Survival analysis is really time-to-failure analysis, and the failure event is the event under analysis. The failure event can be death, heart attack, myopia, or finding employment. Many authors—including Stata—write as if the failure event can occur only once per subject, but when we do, we are being sloppy. Survival analysis encompasses repeated failures, and all of Stata’s survival analysis features can be used with repeated-failure data.

**family distribution.** See generalized linear response functions.

**FCS.** See fully conditional specification.

**feasible initial value.** An initial-value vector is feasible if it corresponds to a state with a positive posterior probability.

**fictional data.** Fictional data are data that have no basis in reality even though they might look real; they are data that are made up for use in examples.

**finite population correction.** Finite population correction (FPC) is an adjustment applied to the variance of a point estimator because of sampling without replacement, resulting in variance estimates that are smaller than the variance estimates from comparable with-replacement sampling designs.

**first- and second-order latent variables.** If a latent variable is measured by other latent variables only, the latent variable that does the measuring is called first-order latent variable, and the latent variable being measured is called the second-order latent variable.

**first-, second-, and higher-level (latent) variables.** Consider a multilevel model of patients within doctors within hospitals. First-level variables are variables that vary at the observational (patient) level. Second-level variables vary across doctors but are constant within doctors. Third-level variables vary across hospitals but are constant within hospitals. This jargon is used whether variables are latent or not.

**Fisher–Irwin’s exact test.** See Fisher’s exact test.

**Fisher’s exact test.** Fisher’s exact test is an exact small-sample test of independence between rows and columns in a $2 \times 2$ contingency table. Conditional on the marginal totals, the test statistic has a hypergeometric distribution under the null hypothesis. See [PSS] power twoproportions and [R] tabulate twoway.

**Fisher’s z test.** This is a $z$ test comparing one or two correlations. See [PSS] power onecorrelation and [PSS] power twocorrelations. Also see Fisher’s $z$ transformation.

**Fisher’s z transformation.** Fisher’s $z$ transformation applies an inverse hyperbolic tangent transformation to the sample correlation coefficient. This transformation is useful for testing hypothesis concerning Pearson’s correlation coefficient. The exact sampling distribution of the correlation coefficient is complicated, while the transformed statistic is approximately standard normal.
fixed effects. In the context of multilevel mixed-effects models, fixed effects represent effects that are constant for all groups at any level of nesting. In the ANOVA literature, fixed effects represent the levels of a factor for which the inference is restricted to only the specific levels observed in the study. See also fixed-effects model in [XT] Glossary.

fixed-effects model. The fixed-effects model is a model for panel data in which the panel-specific errors are treated as fixed parameters. These parameters are panel-specific intercepts and therefore allow the conditional mean of the dependent variable to vary across panels. The linear fixed-effects estimator is consistent, even if the regressors are correlated with the fixed effects. See also random-effects model.

flong data. See style.

flongsep data. See style.

FMI. See fraction of missing information.

follow-up period or follow-up. The (minimum) follow-up period is the period after the last subject entered the study until the end of the study. The follow-up defines the phase of a study during which subjects are under observation and no new subjects enter the study. If $T$ is the total duration of a study, and $R$ is the accrual period of the study, then follow-up period $f$ is equal to $T - R$. Also see accrual period.

follow-up study. See cohort study.

forecast-error variance decomposition. Forecast-error variance decompositions measure the fraction of the error in forecasting variable $i$ after $h$ periods that is attributable to the orthogonalized shocks to variable $j$.

forward operator. The forward operator $F$ denotes the value of a variable at time $t + 1$. Formally, $F y_t = y_{t+1}$, and $F^2 y_t = F y_{t+1} = y_{t+2}$.

FPC. See finite population correction.

fraction of missing information. The ratio of information lost due to the missing data to the total information that would be present if there were no missing data.

An equal FMI test is a test under the assumption that FMI’s are equal across parameters.

An unrestricted FMI test is a test without the equal FMI assumption.

fractional polynomial. A polynomial that may include logarithms, noninteger powers, and repeated powers.

Each time a power repeats in a fractional polynomial of $x$, it is multiplied by another $\ln(x)$.

We write a fractional polynomial in $x$ as

$$x^{(p_1:p_2\ldots:p_m)}/\beta$$

A fractional polynomial in $x$ with powers $(-1, 0, 0.5, 3, 3)$ and coefficients $\beta$ has the following form:

$$x^{(-1,0.5,3,3)}/\beta = \beta_0 + \beta_1 x^{-1} + \beta_2 \ln(x) + \beta_3 x^5 + \beta_4 x^3 + \beta_5 x^3 \ln(x)$$

The notation $x^{(-2,3)}$, for example, means the variable $x^{-2}$ and the variable $x^3$.

frailty. In survival analysis, it is often assumed that subjects are alike—homogeneous—except for their observed differences. The probability that subject $j$ fails at time $t$ may be a function of $j$’s covariates and random chance. Subjects $j$ and $k$, if they have equal covariate values, are equally likely to fail.
Frailty relaxes that assumption. The probability that subject $j$ fails at time $t$ becomes a function of $j$’s covariates and $j$’s unobserved frailty value, $\nu_j$. Frailty $\nu$ is assumed to be a random variable. Parametric survival models can be fit even in the presence of such heterogeneity.

Shared frailty refers to the case in which groups of subjects share the same frailty value. For instance, subjects 1 and 2 may share frailty value $\nu$ because they are genetically related. Both parametric and semiparametric models can be fit under the shared-frailty assumption.

**frequency-domain analysis.** Frequency-domain analysis is analysis of time-series data by considering its frequency properties. The spectral density and distribution functions are key components of frequency-domain analysis, so it is often called spectral analysis. In Stata, the `cumsp` and `pergram` commands are used to analyze the sample spectral distribution and density functions, respectively. `psdensity` estimates the spectral density or the spectral distribution function after estimating the parameters of a parametric model using `arfima`, `arima`, or `ucm`.

**frequentist analysis.** Frequentist analysis is a form of statistical analysis where model parameters are considered to be unknown but fixed constants and the observed data are viewed as a repeatable random sample. Inference is based on the sampling distribution of the data.

**full conditionals.** A full conditional is the probability distribution of a random variate conditioned on all other random variates in a joint probability model. Full conditional distributions are used in Gibbs sampling.

**full Gibbs sampling.** See *Gibbs sampling, Gibbs sampler.*

**full joint and conditional normality assumption.** See *normality assumption, joint and conditional.*

**fully conditional specification.** Consider imputation variables $X_1, X_2, \ldots, X_p$. Fully conditional specification of the prediction equation for $X_j$ includes all variables except $X_j$; that is, variables $X_{-j} = (X_1, X_2, \ldots, X_{j-1}, X_{j+1}, \ldots, X_p)$.

**function.** The words *program* and *function* are used interchangeably. The programs that you write in Mata are in fact functions. Functions receive arguments and optionally return results.

Examples of functions that are included with Mata are `sqrt()`, `ttail()`, and `substr()`. Such functions are often referred to as the built-in functions or the library functions. Built-in functions refer to functions implemented in the C code that implements Mata, and library functions refer to functions written in the Mata programming language, but many users use the words interchangeably because how functions are implemented is of little importance. If you have a choice between using a built-in function and a library function, however, the built-in function will usually execute more quickly and the library function will be easier to use. Mostly, however, features are implemented one way or the other and you have no choice.

Also see *underscore functions.*

For a list of the functions that Mata provides, see `[M-4] intro`.

**future history.** Future history is information recorded after a subject is no longer at risk. The word history is often dropped, and the term becomes simply future. Perhaps the failure event is cardiac infarction, and you want to know whether the subject died soon in the future, in which case you might exclude the subject from analysis.

Also see *past history.*

**gain (of a linear filter).** The gain of a linear filter scales the spectral density of the unfiltered series into the spectral density of the filtered series for each frequency. Specifically, at each frequency, multiplying the spectral density of the unfiltered series by the square of the gain of a linear filter yields the spectral density of the filtered series. If the gain at a particular frequency is 1, the filtered and unfiltered spectral densities are the same at that frequency and the corresponding stochastic
cycles are passed through perfectly. If the gain at a particular frequency is 0, the filter removes all the corresponding stochastic cycles from the unfiltered series.

**gamma regression.** Gamma regression is a term for generalized linear response functions that are family gamma, link log. It is used for continuous, nonnegative, positively skewed data. Gamma regression is also known as log-gamma regression. See *generalized linear response functions.*

**gaps.** Gaps refers to gaps in observation between entry time and exit time; see *under observation.*

**GARCH model.** A generalized autoregressive conditional heteroskedasticity (GARCH) model is a regression model in which the conditional variance is modeled as an ARMA process. The GARCH($m, k$) model is

\[ y_t = x_t \beta + \epsilon_t \]
\[ \sigma_t^2 = \gamma_0 + \gamma_1 \epsilon_{t-1}^2 + \cdots + \gamma_m \epsilon_{t-m}^2 + \delta_1 \sigma_{t-1}^2 + \cdots + \delta_k \sigma_{t-k}^2 \]

where the equation for $y_t$ represents the conditional mean of the process and $\sigma_t$ represents the conditional variance. See [TS arch] or Hamilton (1994, chap. 21) for details on how the conditional variance equation can be viewed as an ARMA process. GARCH models are often used because the ARMA specification often allows the conditional variance to be modeled with fewer parameters than are required by a pure ARCH model. Many extensions to the basic GARCH model exist; see [TS arch] for those that are implemented in Stata. See also *ARCH model.*

**Gauss–Hermite quadrature.** In the context of generalized linear mixed models, Gauss–Hermite quadrature is a method of approximating the integral used in the calculation of the log likelihood. The quadrature locations and weights for individual clusters are fixed during the optimization process.

In the context of IRT models, Gauss–Hermite quadrature (GHQ) is a method of approximating the integral used in the calculation of the log likelihood. The quadrature locations and weights for individuals are fixed during the optimization process.

**Gaussian regression.** Gaussian regression is another term for linear regression. It is most often used when referring to generalized linear response functions. In that framework, Gaussian regression is family Gaussian, link identity. See *generalized linear response functions.*

**generalized eigenvalues.** A scalar, $\lambda$, is said to be a generalized eigenvalue of a pair of $n \times n$ square numeric matrices $A, B$ if there is a nonzero column vector $x$: $n \times 1$ (called a generalized eigenvector) such that

\[ Ax = \lambda Bx \]

Equation (1) can also be written

\[ (A - \lambda B)x = 0 \]

A nontrivial solution to this system of $n$ linear homogeneous equations exists if and only if

\[ \det(A - \lambda B) = 0 \] (2)

In practice, the generalized eigenvalue problem for the matrix pair $(A, B)$ is usually formulated as finding a pair of scalars $(w, b)$ and a nonzero column vector $x$ such that

\[ wAx = bBx \]
The scalar $w/b$ is a generalized eigenvalue if $b$ is not zero.

Infinity is a generalized eigenvalue if $b$ is zero or numerically close to zero. This situation may arise if $B$ is singular.

The Mata functions that compute generalized eigenvalues return them in two complex vectors, $w$ and $b$ of length $n$. If $b[i] = 0$, the $i$th generalized eigenvalue is infinite, otherwise the $i$th generalized eigenvalue is $w[i]/b[i]$.

generalized estimating equations (GEE). The method of generalized estimating equations is used to fit population-averaged panel-data models. GEE extends the GLM method by allowing the user to specify a variety of different within-panel correlation structures.

generalized least-squares estimator. A generalized least-squares (GLS) estimator is used to estimate the parameters of a regression function when the error term is heteroskedastic or autocorrelated. In the linear case, GLS is sometimes described as “OLS on transformed data” because the GLS estimator can be implemented by applying an appropriate transformation to the dataset and then using OLS.

generalized linear mixed-effects model. A generalized linear mixed-effects model is an extension of a generalized linear model allowing for the inclusion of random deviations (effects).

generalized linear model. The generalized linear model is an estimation framework in which the user specifies a distributional family for the dependent variable and a link function that relates the dependent variable to a linear combination of the regressors. The distribution must be a member of the exponential family of distributions. The generalized linear model encompasses many common models, including linear, probit, and Poisson regression.

generalized linear response functions. Generalized linear response functions include linear functions and include functions such as probit, logit, multinomial logit, ordered probit, ordered logit, Poisson, and more.

These generalized linear functions are described by a link function $g(\cdot)$ and statistical distribution $F$. The link function $g(\cdot)$ specifies how the response variable $y_i$ is related to a linear equation of the explanatory variables, $x_i\beta$, and the family $F$ specifies the distribution of $y_i$:

$$g\{E(y_i)\} = x_i\beta, \quad y_i \sim F$$

If we specify that $g(\cdot)$ is the identity function and $F$ is the Gaussian (normal) distribution, then we have linear regression. If we specify that $g(\cdot)$ is the logit function and $F$ the Bernoulli distribution, then we have logit (logistic) regression.

In this generalized linear structure, the family may be Gaussian, gamma, Bernoulli, binomial, Poisson, negative binomial, ordinal, or multinomial. The link function may be the identity, log, logit, probit, or complementary log-log.

gsem fits models with generalized linear response functions.

generalized method of moments. Generalized method of moments (GMM) is a method used to obtain fitted parameters. In this documentation, GMM is referred to as ADF, which stands for asymptotic distribution free and is available for use with sem. Other available methods for use with sem are ML, QML, ADF, and MLMV.

The SEM moment conditions are cast in terms of second moments, not the first moments used in many other applications associated with GMM.

generalized partial credit model. The generalized partial credit model (GPCM) is an IRT model for ordinal responses. The categories within each item vary in their difficulty and share the same discrimination parameter.
**generalized SEM.** Generalized SEM is a term we have coined to mean SEM optionally allowing generalized linear response functions or multilevel models. `gsem` fits generalized SEMs.

**GHQ.** See *Gauss–Hermite quadrature.*

**Gibbs sampling, Gibbs sampler.** Gibbs sampling is an MCMC method, according to which each random variable from a joint probability model is sampled according to its full conditional distribution.

**GLM.** See *generalized linear model.*

**GLME model.** See *generalized linear mixed-effects model.*

**GLMM.** Generalized linear mixed model. See *generalized linear mixed-effects model.*

**global variable.** Global variables, also known as external variables and as global external variables, refer to variables that are common across programs and which programs may access without the variable being passed as an argument.

The variables you create interactively are global variables. Even so, programs cannot access those variables without engaging in another step, and global variables can be created without your creating them interactively.

To access (and create if necessary) global external variables, you declare the variable in the body of your program:

```plaintext
function myfunction(...) 
{
    external real scalar globalvar
    ...
}
```

See *Linking to external globals* in [M-2] declarations.*

There are other ways of creating and accessing global variables, but the declaration method is recommended. The alternatives are `crexternal()`, `findexternal()`, and `rmexternal()` documented in [M-5] `findexternal()` and `valofexternal()` documented in [M-5] `valofexternal()`.

**GMM.** See *generalized method of moments.*

**goodness-of-fit statistic.** A goodness-of-fit statistic is a value designed to measure how well the model reproduces some aspect of the data the model is intended to fit. SEM reproduces the first- and second-order moments of the data, with an emphasis on the second-order moments, and thus goodness-of-fit statistics appropriate for use after `sem` compare the predicted covariance matrix (and mean vector) with the matrix (and vector) observed in the data.

**GPCM.** See *generalized partial credit model.*

**graded response model.** The graded response model (GRM) is an extension of the two-parameter logistic model to ordinal responses. The categories within each item vary in their difficulty and share the same discrimination parameter.

**Granger causality.** The variable $x$ is said to Granger-cause variable $y$ if, given the past values of $y$, past values of $x$ are useful for predicting $y$.

**Greenhouse–Geisser correction.** See *nonsphericity correction.*

**GRM.** See *graded response model.*

`gsem`. `gsem` is the Stata command that fits generalized SEMs. Also see `sem`.

**guessing.** The guessing parameter incorporates the impact of chance on an observed response. The parameter lifts the lower asymptote of the item characteristic curve above zero.
GUI. See *Builder*.

**Hadamard matrix.** A Hadamard matrix is a square matrix with \( r \) rows and columns that has the property

\[
H_r' H_r = r I_r
\]

where \( I_r \) is the identity matrix of order \( r \). Generating a Hadamard matrix with order \( r = 2^p \) is easily accomplished. Start with a Hadamard matrix of order 2 (\( H_2 \)), and build your \( H_r \) by repeatedly applying Kronecker products with \( H_2 \).

**hard missing and soft missing.** A hard missing value is a value of .a, .b, . . ., .z in \( m = 0 \) in an imputed variable. Hard missing values are not replaced in \( m > 0 \).

A soft missing value is a value of . in \( m = 0 \) in an imputed variable. If an imputed variable contains soft missing, then that value is eligible to be imputed, and perhaps is imputed, in \( m > 0 \).

Although you can use the terms hard missing and soft missing for passive, regular, and unregistered variables, it has no special significance in terms of how the missing values are treated.

**hashing, hash functions, and hash tables.** Hashing refers to a technique for quickly finding information corresponding to an identifier. The identifier might be a name, a Social Security number, fingerprints, or anything else on which the information is said to be indexed. The hash function returns a many-to-one mapping of identifiers onto a dense subrange of the integers. Those integers, called hashes, are then used to index a hash table. The selected element of the hash table specifies a list containing identifiers and information. The list is then searched for the particular identifier desired. The advantage is that rather than searching a single large list, one need only search one of \( K \) smaller lists. For this to be fast, the hash function must be quick to compute and produce roughly equal frequencies of hashes over the range of identifiers likely to be observed.

**hazard, cumulative hazard, and hazard ratio.** The hazard or hazard rate at time \( t \), \( h(t) \), is the instantaneous rate of failure at time \( t \) conditional on survival until time \( t \). Hazard rates can exceed 1. Say that the hazard rate were 3. If an individual faced a constant hazard of 3 over a unit interval and if the failure event could be repeated, the individual would be expected to experience three failures during the time span.

The cumulative hazard, \( H(t) \), is the integral of the hazard function \( h(t) \), from 0 (the onset of risk) to \( t \). It is the total number of failures that would be expected to occur up until time \( t \), if the failure event could be repeated. The relationship between the cumulative hazard function, \( H(t) \), and the survivor function, \( S(t) \), is

\[
S(t) = \exp\{-H(t)\}
\]

\[
H(t) = -\ln\{S(t)\}
\]

The hazard ratio is the ratio of the hazard function evaluated at two different values of the covariates: \( h(t \mid x)/h(t \mid x_0) \). The hazard ratio is often called the relative hazard, especially when \( h(t \mid x_0) \) is the baseline hazard function.

**hazard contributions.** Hazard contributions are the increments of the estimated cumulative hazard function obtained through either a nonparametric or semiparametric analysis. For these analysis types, the estimated cumulative hazard is a step function that increases every time a failure occurs. The hazard contribution for that time is the magnitude of that increase.

Because the time between failures usually varies from failure to failure, hazard contributions do not directly estimate the hazard. However, one can use the hazard contributions to formulate an estimate of the hazard function based on the method of smoothing.
Hermitian matrix. Matrix $A$ is Hermitian if it is equal to its conjugate transpose; $A = A'$; see transpose. This means that each off-diagonal element $a_{ij}$ must equal the conjugate of $a_{ji}$, and that the diagonal elements must be real. The following matrix is Hermitian:

$$
\begin{bmatrix}
2 & 4 + 5i \\
4 - 5i & 6
\end{bmatrix}
$$

The definition $A = A'$ is the same as the definition for a symmetric matrix, although usually the word symmetric is reserved for real matrices and Hermitian, for complex matrices. In this manual, we use the word symmetric for both; see symmetric matrices.

Hessenberg decomposition. The Hessenberg decomposition of a matrix, $A$, can be written as

$$
Q' A Q = H
$$

where $H$ is in upper Hessenberg form and $Q$ is orthogonal if $A$ is real or unitary if $A$ is complex. See [M-5] hessenbergd().

Hessenberg form. A matrix, $A$, is in upper Hessenberg form if all entries below the first subdiagonal are zero: $A_{ij} = 0$ for all $i > j + 1$.

A matrix, $A$, is in lower Hessenberg form if all entries above the first superdiagonal are zero: $A_{ij} = 0$ for all $j > i + 1$.

Heywood case or Heywood solution. A Heywood case can appear in factor analysis output; this indicates that a boundary solution, called a Heywood solution, was produced. The geometric assumptions underlying the likelihood-ratio test are violated, though the test may be useful if interpreted cautiously.

hierarchical clustering and hierarchical clustering methods. In hierarchical clustering, the data is placed into clusters via iterative steps. Contrast to partition clustering. Also see agglomerative hierarchical clustering methods and divisive hierarchical clustering methods.

hierarchical model. A hierarchical model is one in which successively more narrowly defined groups are nested within larger groups. For example, in a hierarchical model, patients may be nested within doctors who are in turn nested within the hospital at which they practice.

higher ASCII. See extended ASCII.

highest posterior density credible interval, HPD credible interval. The highest posterior density (HPD) credible interval is a type of a credible interval with the highest marginal posterior density. An HPD interval has the shortest width among all other credible intervals. For some multimodal marginal distributions, HPD may not exists. See highest posterior density region, HPD region.

highest posterior density region, HPD region. The highest posterior density (HPD) region for model parameters has the highest marginal posterior probability among all domain regions. Unlike an HPD credible interval, an HPD region always exist.

high-pass filter. Time-series filters are designed to pass or block stochastic cycles at specified frequencies. High-pass filters, such as those implemented in tsfilter bw and tsfilter hp, pass through stochastic cycles above the cutoff frequency and block all other stochastic cycles.

Holt–Winters smoothing. A set of methods for smoothing time-series data that assume that the value of a time series at time $t$ can be approximated as the sum of a mean term that drifts over time, as well as a time trend whose strength also drifts over time. Variations of the basic method allow for seasonal patterns in data, as well.

Hotelling’s T-squared generalized means test. Hotelling’s T-squared generalized means test is a multivariate test that reduces to a standard $t$ test if only one variable is specified. It tests whether one set of means is zero or if two sets of means are equal.
**hybrid MH sampling, hybrid MH sampler.** A hybrid MH sampler is an MCMC method in which some blocks of parameters are updated using the MH algorithms and other blocks are updated using Gibbs sampling.

**hybrid model.** A hybrid IRT model is a model that performs a single calibration of an instrument consisting of different response formats.

**hyperparameter.** In Bayesian analysis, hyperparameter is a parameter of a prior distribution, in contrast to a model parameter.

**hyperprior.** In Bayesian analysis, hyperprior is a prior distribution of hyperparameters. See hyperparameter.

**hypothesis.** A hypothesis is a statement about a population parameter of interest.

**hypothesis testing, hypothesis test.** This method of inference evaluates the validity of a hypothesis based on a sample from the population. See Hypothesis testing under Remarks and examples in [PSS] intro.

**hypothesized value.** See null value.

**ICC.** See item characteristic curve.

**ID variable.** An ID variable identifies groups; equal values of an ID variable indicate that the observations are for the same group. For instance, a stratification ID variable would indicate the strata to which each observation belongs.

When an ID variable is referred to without modification, it means subjects, and usually this occurs in multiple-record st data. In multiple-record data, each physical observation in the dataset represents a time span, and the ID variable ties the separate observations together:

<table>
<thead>
<tr>
<th>idvar</th>
<th>t0</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>7</td>
</tr>
</tbody>
</table>

ID variables are usually numbered 1, 2, ..., but that is not required. An ID variable might be numbered 1, 3, 7, 22, ..., or −5, −4, ..., or even 1, 1.1, 1.2, ....

**identification.** Identification refers to the conceptual constraints on parameters of a model that are required for the model’s remaining parameters to have a unique solution. A model is said to be unidentified if these constraints are not supplied. These constraints are of two types: substantive constraints and normalization constraints.

Normalization constraints deal with the problem that one scale works as well as another for each latent variable in the model. One can think, for instance, of propensity to write software as being measured on a scale of 0 to 1, 1 to 100, or any other scale. The normalization constraints are the constraints necessary to choose one particular scale. The normalization constraints are provided automatically by sem and gsem by anchoring with unit loadings.

Substantive constraints are the constraints you specify about your model so that it has substantive content. Usually, these constraints are zero constraints implied by the paths omitted, but they can include explicit parameter constraints as well. It is easy to write a model that is not identified for substantive reasons; See [SEM] intro 4.

**idiosyncratic error term.** In longitudinal or panel-data models, the idiosyncratic error term refers to the observation-specific zero-mean random-error term. It is analogous to the random-error term of cross-sectional regression analysis.
**ignorable missing-data mechanism.** The missing-data mechanism is said to be ignorable if missing data are **missing at random** and the parameters of the data model and the parameters of the missing-data mechanism are distinct; that is, the joint distribution of the model and the missing-data parameters can be factorized into two independent marginal distributions of model parameters and of missing-data parameters.

**i.i.d. sampling assumption.** See *independent and identically distributed sampling assumption.*

**IIF.** See *item information function.*

**improper prior.** A prior is said to be improper if it does not integrate to a finite number. Uniform distributions over unbounded intervals are improper. Improper priors may still yield proper posterior distributions. When using improper priors, however, one has to make sure that the resulting posterior distribution is proper for Bayesian inference to be invalid.

**impulse–response function.** An impulse–response function (IRF) measures the effect of a shock to an endogenous variable on itself or another endogenous variable. The $k$th impulse–response function of variable $i$ on variable $j$ measures the effect on variable $j$ in period $t + k$ in response to a one-unit shock to variable $i$ in period $t$, holding everything else constant.

**imputed, passive, and regular variables.** An imputed variable is a variable that has missing values and for which you have or will have imputations.

A passive variable is a **varying variable** that is a function of imputed variables or of other passive variables. A passive variable will have missing values in $m = 0$ and varying values for observations in $m > 0$.

A regular variable is a variable that is neither imputed nor passive and that has the same values, whether missing or not, in all $m$.

Imputed, passive, and regular variables can be registered using the *mi register* command; see [mi] mi set. You are required to register imputed variables, and we recommend that you register passive variables. Regular variables can also be registered. See *registered and unregistered variables.*

**imputed data.** Data in which all missing values are imputed.

**incidence and incidence rate.** Incidence is the number of new failures (for example, number of new cases of a disease) that occur during a specified period in a population at risk (for example, of the disease).

Incidence rate is incidence divided by the sum of the length of time each individual was exposed to the risk.

Do not confuse incidence with prevalence. Prevalence is the fraction of a population that has the disease. Incidence refers to the rate at which people contract a disease, whereas prevalence is the total number actually sick at a given time.

**incidence studies, longitudinal studies, and follow-up studies.** Whichever word is used, these studies monitor a population for a time to track the transition of noncases into cases. Incidence studies are prospective. Also see *cohort studies.*

**incomplete observations.** See *complete and incomplete observations.*

**independent a posteriori.** Parameters are considered independent a posteriori if their marginal posterior distributions are independent; that is, their joint posterior distribution is the product of their individual marginal posterior distributions.

**independent a priori.** Parameters are considered independent a priori if their prior distributions are independent; that is, their joint prior distribution is the product of their individual marginal prior distributions.
**independent and identically distributed.** A series of observations is independent and identically distributed (i.i.d.) if each observation is an independent realization from the same underlying distribution. In some contexts, the definition is relaxed to mean only that the observations are independent and have identical means and variances; see Davidson and MacKinnon (1993, 42).

**independent and identically distributed sampling assumption.** The independent and identically distributed (i.i.d.) sampling assumption specifies that each observation is unrelated to (independent of) all the other observations and that each observation is a draw from the same (identical) distribution.

**indicator variables, indicators.** The term “indicator variable” has two meanings. An indicator variable is a 0/1 variable that contains whether something is true. The other usage is as a synonym for measurement variables.

**indirect effects.** See *direct, indirect, and total effects.*

**individual-level treatment effect.** An individual-level treatment effect is the difference in an individual’s outcome that would occur because this individual is given one treatment instead of another. In other words, an individual-level treatment effect is the difference between two potential outcomes for an individual.

For example, the blood pressure an individual would obtain after taking a pill minus the blood pressure an individual would obtain had that person not taken the pill is the individual-level treatment effect of the pill on blood pressure.

**ineligible missing value.** An ineligible missing value is a missing value in a to-be-imputed variable that is due to inability to calculate a result rather than an underlying value being unobserved. For instance, assume that variable *income* had some missing values and so you wish to impute it. Because *income* is skewed, you decide to impute the log of *income*, and you begin by typing

```
    . generate lnincome = log(income)
```

If *income* contained any zero values, the corresponding missing values in *lnincome* would be ineligible missing values. To ensure that values are subsequently imputed correctly, it is of vital importance that any ineligible missing values be recorded as hard missing. You would do that by typing

```
    . replace lnincome = .a if lnincome==. & income!=.
```

As an aside, if after imputing *lnincome* using *mi impute* (see [MI] *mi impute*), you wanted to fill in *income*, *income* surprisingly would be a passive variable because *lnincome* is the imputed variable and *income* would be derived from it. You would type

```
    . mi register passive income
    . mi passive: replace income = cond(lnincome==.a, 0, exp(lnincome))
```

In general, you should avoid using transformations that produce ineligible missing values to avoid the loss of information contained in other variables in the corresponding observations. For example, in the above, for zero values of *income* we could have assigned the log of *income*, *lnincome*, to be the smallest value that can be stored as *double*, because the logarithm of zero is negative infinity:

```
    . generate lnincome = cond(income==0, mindouble(), log(income))
```

This way, all observations for which *income*==0 will be used in the imputation model for *lnincome*.

**inertia.** In CA, the inertia is related to the definition in applied mathematics of “moment of inertia”, which is the integral of the mass times the squared distance to the centroid. Inertia is defined as the total Pearson chi-squared for the two-way table divided by the total number of observations or the sum of the squared singular values found in the singular value decomposition.
total inertia $= \frac{1}{n} \chi^2 = \sum_k \lambda_k^2$

In MCA, the inertia is defined analogously. In the case of the indicator or Burt matrix approach, it is given by the formula

$$\text{total inertia} = \left(\frac{q}{q-1}\right) \sum \phi_t^2 - \frac{(J-q)}{q^2}$$

where $q$ is the number of active variables, $J$ is the number of categories and $\phi_t$ is the $t$th (unadjusted) eigenvalue of the eigen decomposition. In JCA the total inertia of the modified Burt matrix is defined as the sum of the inertias of the off-diagonal blocks. Also see correspondence analysis and multiple correspondence analysis.

**information.** Precision with which an item or an instrument measures the latent trait; also see item information function and test information function.

**informative prior.** An informative prior is a prior distribution that has substantial influence on the posterior distribution.

**initial values.** See starting values.

**instance and realization.** Instance and realization are synonyms for variable, as in Mata variable. For instance, consider a real scalar variable $X$. One can equally well say that $X$ is an instance of a real scalar or a realization of a real scalar. Authors represent a variable this way when they wish to emphasize that $X$ is not representative of all real scalars but is just one of many real scalars. Instance is often used with structures and classes when the writer wishes to emphasize the difference between the values contained in the variable and the definition of the structure or the class. It is confusing to say that $V$ is a class $C$, even though it is commonly said, because the reader might confuse the definition of $C$ with the specific values contained in $V$. Thus careful authors say that $V$ is an instance of class $C$.

**instrument.** A collection of items, usually called a test, a survey, or a questionnaire.

**instrumental variables.** Instrumental variables are exogenous variables that are correlated with one or more of the endogenous variables in a structural model. The term instrumental variable is often reserved for those exogenous variables that are not included as regressors in the model.

**instrumental-variables (IV) estimator.** An instrumental variables estimator uses instrumental variables to produce consistent parameter estimates in models that contain endogenous variables. IV estimators can also be used to control for measurement error.

**integrated process.** A nonstationary process is integrated of order $d$, written $I(d)$, if the process must be differenced $d$ times to produce a stationary series. An $I(1)$ process $y_t$ is one in which $\Delta y_t$ is stationary.

**interaction effects.** Interaction effects measure the dependence of the effects of one factor on the levels of the other factor. Mathematically, they can be defined as the differences among treatment means that are left after main effects are removed from these differences.

**intercept.** An intercept for the equation of endogenous variable $y$, observed or latent, is the path coefficient from `_cons` to $y$. `_cons` is Stata-speak for the built-in variable containing 1 in all observations. In SEM-speak, `_cons` is an observed exogenous variable.

**interval data.** Interval data are data in which the true value of the dependent variable is not observed. Instead, all that is known is that the value lies within a given interval.

**interval hypothesis testing.** Interval hypothesis testing performs interval hypothesis tests for model parameters and functions of model parameters.
interval test. In Bayesian analysis, an interval test applied to a scalar model parameter calculates the marginal posterior probability for the parameter to belong to the specified interval.

intraclass correlation. In the context of mixed-effects models, intraclass correlation refers to the correlation for pairs of responses at each nested level of the model.

invariance. When an IRT model fits the data exactly in the population, then the estimated item parameters should be the same, within sampling error, regardless of what sample the data were derived from, and the estimated person latent traits should be the same regardless of what items they are based on.

inverse-probability-weighted estimators. Inverse-probability-weighted (IPW) estimators use weighted averages of the observed outcome variable to estimate the potential-outcome means. The weights are the reciprocals of the treatment probabilities estimated by a treatment model.

inverse-probability-weighted regression-adjustment estimators. Inverse-probability-weighted regression-adjustment (IPWRA) estimators use the reciprocals of the estimated treatment probability as weights to estimate missing-data-corrected regression coefficients that are subsequently used to compute the potential-outcome means.

IPW estimators. See inverse-probability-weighted estimators.

IPWRA estimators. See inverse-probability-weighted regression-adjustment estimators.

IRT. See item response theory.

istmt. An istmt is an interactive statement, a statement typed at Mata’s colon prompt.

item. An item is a single question or task on a test or an instrument.

item characteristic curve. An item characteristic curve (ICC) expresses the probability for a given response to a binary item as a function of the latent trait.

item information function. An item information function (IIF) indicates the precision of an item along the latent trait continuum.

item location. Location of an item on the difficulty scale.

item response function. See item characteristic curve.

item response theory. Item response theory (IRT) is a theoretical framework organized around the concept of the latent trait. IRT encompasses a set of models and associated statistical procedures that relate observed responses on an instrument to a person’s level of the latent trait.

iterated principal-factor method. The iterated principal-factor method is a method for performing factor analysis in which the communalities \( \hat{h}_i^2 \) are estimated iteratively from the loadings in \( \hat{\Lambda} \) using

\[
\hat{h}_i^2 = \sum_{j=1}^{m} \hat{\lambda}_{ij}^2
\]

Also see factor analysis and communality.

J × 2 contingency table. A J × 2 contingency table is used to describe the association between an ordinal independent variable with J levels and a binary response variable of interest.

\( J(\text{r, c, value}) \). \( J() \) is the function that returns an r × c matrix with all elements set to value; see [M-5] J(). Also, J() is often used in the documentation to describe the various types of void matrices; see void matrix. Thus the documentation might say that such-and-such returns \( J(0, 0, .) \) under certain conditions. That is another way of saying that such-and-such returns a 0 × 0 real matrix.
When \( r \) or \( c \) is 0, there are no elements to be filled in with \textit{value}, but even so, \textit{value} is used to determine the type of the matrix. Thus \( J(0, 0, 1i) \) refers to a \( 0 \times 0 \) complex matrix, \( J(0, 0, ")" \) refers to a \( 0 \times 0 \) string matrix, and \( J(0, 0, \text{NULL}) \) refers to a \( 0 \times 0 \) pointer matrix.

In the documentation, \( J() \) is used for more than describing \( 0 \times 0 \) matrices. Sometimes, the matrices being described are \( r \times 0 \) or \( 0 \times c \). Say that a function \texttt{example(X)} is supposed to return a column vector; perhaps it returns the last column of \( X \). Now say that \( X \) is \( 0 \times 0 \). Function \texttt{example()} still should return a column vector, and so it returns a \( 0 \times 1 \) matrix. This would be documented by noting that \texttt{example()} returns \( J(0, 1, .) \) when \( X \) is \( 0 \times 0 \).

\textbf{jackknife.} The jackknife is a data-dependent way to estimate the variance of a statistic, such as a mean, ratio, or regression coefficient. Unlike BRR, the jackknife can be applied to practically any survey design. The jackknife variance estimator is described in \textit{[SVY] variance estimation}.

\textbf{jackknife, vce(jackknife).} The jackknife is a replication method for obtaining variance estimates. Consider an estimation method \( E \) for estimating \( \theta \). Let \( \hat{\theta} \) be the result of applying \( E \) to dataset \( D \) containing \( N \) observations. The jackknife is a way of obtaining variance estimates for \( \hat{\theta} \) from repeated estimates \( \hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N \), where each \( \hat{\theta}_i \) is the result of applying \( E \) to \( D \) with observation \( i \) removed. See \textit{[SEM] sem option method(\() and \textit{[R] jackknife}. vce(jackknife) is allowed with \texttt{sem} but not \texttt{gsem}. You can obtain jackknife results by prefixing the \texttt{gsem} command with \texttt{jackknife:}, but remember to specify \texttt{jackknife}'s \texttt{cluster()} and \texttt{idcluster()} options if you are fitting a multilevel model. See \textit{[SEM] intro 9}.

\textbf{jackknifed standard error.} See \textit{Monte Carlo error}.

\textbf{JCA.} An acronym for joint correspondence analysis; see \textit{multiple correspondence analysis}.

\textbf{Jeffreys prior.} The Jeffreys prior of a vector of model parameters \( \theta \) is proportional to the square root of the determinant of its Fisher information matrix \( I(\theta) \). Jeffreys priors are locally uniform and, by definition, agree with the likelihood function. Jeffreys priors are considered noninformative priors that have minimal impact on the posterior distribution.

\textbf{joint correspondence analysis.} See \textit{multiple correspondence analysis}.

\textbf{joint normality assumption.} See \textit{normality assumption, joint and conditional}.

\textbf{Kaiser–Meyer–Olkin measure of sampling adequacy.} The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy takes values between 0 and 1, with small values meaning that the variables have too little in common to warrant a factor analysis or PCA. Historically, the following labels have been given to values of KMO (Kaiser 1974):

- 0.00 to 0.49 unacceptable
- 0.50 to 0.59 miserable
- 0.60 to 0.69 mediocre
- 0.70 to 0.79 middling
- 0.80 to 0.89 meritorious
- 0.90 to 1.00 marvelous

\textbf{Kalman filter.} The Kalman filter is a recursive procedure for predicting the state vector in a state-space model.

\textbf{Kaplan–Meier product-limit estimate.} This is an estimate of the survivor function, which is the product of conditional survival to each time at which an event occurs. The simple form of the calculation, which requires tallying the number at risk and the number who die and at each time, makes accounting for censoring easy. The resulting estimate is a step function with jumps at the event times.
Kenward–Roger denominator degrees of freedom (DDF) method. This method implements the Kenward and Roger (1997) method, which is designed to approximate unknown sampling distributions of test statistics for complex linear mixed-effects models. This method is supported only with restricted maximum-likelihood estimation.

**kmeans.** Kmeans is a method for performing partition cluster analysis. The user specifies the number of clusters, \(k\), to create using an iterative process. Each observation is assigned to the group whose mean is closest, and then based on that categorization, new group means are determined. These steps continue until no observations change groups. The algorithm begins with \(k\) seed values, which act as the \(k\) group means. There are many ways to specify the beginning seed values. Also see partition clustering.

**kmedians.** Kmedians is a variation of kmeans. The same process is performed, except that medians instead of means are computed to represent the group centers at each step. Also see kmeans and partition clustering.

**KMO.** See Kaiser–Meyer–Olkin measure of sampling adequacy.

**KNN.** See \(k\)th nearest neighbor.

**Kruskal stress.** The Kruskal stress measure (Kruskal 1964; Cox and Cox 2001, 63) used in MDS is given by

\[
\text{Kruskal}(\hat{D}, E) = \left\{ \frac{\sum (E_{ij} - \hat{D}_{ij})^2}{\sum E_{ij}^2} \right\}^{1/2}
\]

where \(D_{ij}\) is the dissimilarity between objects \(i\) and \(j\), \(1 \leq i, j \leq n\), and \(\hat{D}_{ij}\) is the disparity, that is, the transformed dissimilarity, and \(E_{ij}\) is the Euclidean distance between rows \(i\) and \(j\) of the matching configuration. Kruskal stress is an example of a loss function in modern MDS. After classical MDS, estat stress gives the Kruskal stress. Also see classical scaling, multidimensional scaling, and stress.

**kth nearest neighbor.** \(k\)th-nearest-neighbor (KNN) discriminant analysis is a nonparametric discrimination method based on the \(k\) nearest neighbors of each observation. Both continuous and binary data can be handled through the different similarity and dissimilarity measures. KNN analysis can distinguish irregular-shaped groups, including groups with multiple modes. Also see discriminant analysis and nonparametric methods.

**lag operator.** The lag operator \(L\) denotes the value of a variable at time \(t - 1\). Formally, \(Ly_t = y_{t-1}\), and \(L^2y_t = Ly_{t-1} = y_{t-2}\).

**Lagrange multiplier test.** Synonym for score test.

**LAPACK.** LAPACK stands for Linear Algebra PACKage and forms the basis for many of Mata’s linear algebra capabilities; see [M-1] LAPACK.

**Laplacian approximation.** Laplacian approximation is a technique used to approximate definite integrals without resorting to quadrature methods. In the context of mixed-effects models, Laplacian approximation is as a rule faster than quadrature methods at the cost of producing biased parameter estimates of variance components.

**latent growth model.** A latent growth model is a kind of measurement model in which the observed values are collected over time and are allowed to follow a trend. See [SEM] intro 5.

**latent space.** Number of latent traits that are measured by an instrument. All IRT models described in this manual assume a unidimensional latent space or, in other words, that a single latent trait explains the response pattern.

**latent trait.** A variable or construct that cannot be directly observed.
latent variable. A variable is latent if it is not observed. A variable is latent if it is not in your dataset but you wish it were. You wish you had a variable recording the propensity to commit violent crime, or socioeconomic status, or happiness, or true ability, or even income accurately recorded. Latent variables are sometimes described as imagined variables.

In the software, latent variables are usually indicated by having at least their first letter capitalized. Also see first- and second-order latent variables, first-, second-, and higher-level (latent) variables, and observed variables.

Lawley–Hotelling trace. The Lawley–Hotelling trace is a test statistic for the hypothesis test $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ based on the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_s$ of $E^{-1}H$. It is defined as

$$U(s) = \text{trace}(E^{-1}H) = \sum_{i=1}^{s} \lambda_i$$

where $H$ is the between matrix and $E$ is the within matrix, see between matrix.

LDA. See linear discriminant analysis.

leave one out. In discriminant analysis, classification of an observation while leaving it out of the estimation sample is done to check the robustness of the analysis; thus the phrase “leave one out” (LOO). Also see discriminant analysis.

left eigenvectors. A vector $x : n \times 1$ is said to be a left eigenvector of square matrix $A : n \times n$ if there is a nonzero scalar, $\lambda$, such that $xA = \lambda x$.

left-censoring. See censored, censoring, left-censoring, and right-censoring.

left-truncation. See truncation, left-truncation, and right-truncation.

life table. Also known as a mortality table or actuarial table, a life table is a table that shows for each analysis time the fraction that survive to that time. In mortality tables, analysis time is often age.

likelihood displacement value. A likelihood displacement value is an influence measure of the effect of deleting a subject on the overall coefficient vector. Also see partial likelihood displacement value.

likelihood-ratio test. The likelihood-ratio (LR) test is one of the three classical testing procedures used to compare the fit of two models, one of which, the constrained model, is nested within the full (unconstrained) model. Under the null hypothesis, the constrained model fits the data as well as the full model. The LR test requires one to determine the maximal value of the log-likelihood function for both the constrained and the full models. See [PSS] power twoproportions and [R] lrtest.

linear discriminant analysis. Linear discriminant analysis (LDA) is a parametric form of discriminant analysis. In Fisher’s (1936) approach to LDA, linear combinations of the discriminating variables provide maximal separation between the groups. The Mahalanobis (1936) formulation of LDA assumes that the observations come from multivariate normal distributions with equal covariance matrices. Also see discriminant analysis and parametric methods.

linear filter. A linear filter is a sequence of weights used to compute a weighted average of a time series at each time period. More formally, a linear filter $\alpha(L)$ is

$$\alpha(L) = \alpha_0 + \alpha_1 L + \alpha_2 L^2 + \cdots = \sum_{\tau=0}^{\infty} \alpha_\tau L^\tau$$
where $L$ is the lag operator. Applying the linear filter $\alpha(L)$ to the time series $x_t$ yields a sequence of weighted averages of $x_t$:

$$\alpha(L)x_t = \sum_{\tau=0}^{\infty} \alpha_{\tau} L^\tau x_{t-\tau}$$

**linear mixed model.** See *linear mixed-effects model.*

**linear mixed-effects model.** A linear mixed-effects model is an extension of a linear model allowing for the inclusion of random deviations (effects).

**linear regression.** Linear regression is a kind of SEM in which there is a single equation. See [SEM] intro 5.

**linearization.** Linearization is short for Taylor linearization. Also known as the delta method or the Huber/White/robust sandwich variance estimator, linearization is a method for deriving an approximation to the variance of a point estimator, such as a ratio or regression coefficient. The linearized variance estimator is described in [SVY] variance estimation.

**link function.** See *generalized linear response functions.*

**linkage.** In cluster analysis, the linkage refers to the measure of proximity between groups or clusters.

**listwise deletion, casewise deletion.** Omitting from analysis observations containing missing values.

**LMAX value.** An LMAX value is an influence measure of the effect of deleting a subject on the overall coefficient vector and is based on an eigensystem analysis of efficient score residuals. Also see partial LMAX value.

**LME model.** See *linear mixed-effects model.*

**loading.** A loading is a coefficient or weight in a linear transformation. Loadings play an important role in many multivariate techniques, including factor analysis, PCA, MANOVA, LDA, and canonical correlations. In some settings, the loadings are of primary interest and are examined for interpretability. For many multivariate techniques, loadings are based on an eigenanalysis of a correlation or covariance matrix. Also see eigenvalues and eigenvector.

**loading plot.** A loading plot is a scatter plot of the loadings after LDA, factor analysis or PCA.

**local independence.** See *conditional independence.*

**locale.** A locale is a code that identifies a community with a certain set of rules for how their language should be written. A locale can refer to something as general as an entire language (for example, “en” for English) or something as specific as a language in a particular country (for example, “en_HK” for English in Hong Kong).

A locale specifies a set of rules that govern how the language should be written. Stata uses locales to determine how certain language-specific operations are carried out. For more information, see [U] 12.4.2.4 Locales in Unicode.

**logistic discriminant analysis.** Logistic discriminant analysis is a form of discriminant analysis based on the assumption that the likelihood ratios of the groups have an exponential form. Multinomial logistic regression provides the basis for logistic discriminant analysis. Because multinomial logistic regression can handle binary and continuous regressors, logistic discriminant analysis is also appropriate for binary and continuous discriminating variables. Also see *discriminant analysis.*

**logit regression.** Logit regression is a term for generalized linear response functions that are family Bernoulli, link logit. It is used for binary outcome data. Logit regression is also known as logistic regression and also simply as logit. See *generalized linear response functions.*

**longitudinal data.** Longitudinal data is another term for panel data. See also *panel data.*
**Glossary**

**long-memory process.** A long-memory process is a stationary process whose autocorrelations decay at a slower rate than a short-memory process. ARFIMA models are typically used to represent long-memory processes, and ARMA models are typically used to represent short-memory processes.

**LOO.** See *leave one out.*

**loss.** Modern MDS is performed by minimizing a loss function, also called a loss criterion. The loss quantifies the difference between the disparities and the Euclidean distances.

Loss functions include Kruskal’s stress and its square, both normalized with either disparities or distances, the strain criterion which is equivalent to classical metric scaling when the disparities equal the dissimilarities, and the Sammon (1969) mapping criterion which is the sum of the scaled, squared differences between the distances and the disparities, normalized by the sum of the disparities.

Also see *multidimensional scaling, Kruskal stress, classical scaling,* and *disparity.*

**loss to follow-up.** Subjects are lost to follow-up if they do not complete the course of the study for reasons unrelated to the event of interest. For example, loss to follow-up occurs if subjects move to a different area or decide to no longer participate in a study. Loss to follow-up should not be confused with administrative censoring. If subjects are lost to follow-up, the information about the outcome these subjects would have experienced at the end of the study, had they completed the study, is unavailable. Also see *withdrawal, administrative censoring,* and *follow-up period or follow-up.*

**lower ASCII.** See *plain ASCII.*

**lower asymptote.** See *guessing.*

**lower one-sided test, lower one-tailed test.** A lower one-sided test is a one-sided test of a scalar parameter in which the alternative hypothesis is lower one sided, meaning that the alternative hypothesis states that the parameter is less than the value conjectured under the null hypothesis. Also see *One-sided test versus two-sided test* under Remarks and examples in [PSS] intro.

**lval.** *lval* stands for left-hand-side value and is defined as the property of being able to appear on the left-hand side of an equal-assignment operator. Matrices are *lvals* in Mata, and thus

\[ x = \ldots \]

is valid. Functions are not *lvals;* thus, you cannot code

\[ \text{substr(mystr,1,3)} = "abc" \]

*lvals* would be easy to describe except that *pointers* can also be lvals. Few people ever use pointers. See [M-2] op_assignment for a complete definition.

**M, m.** \( M \) is the number of imputations. \( m \) refers to a particular imputation, \( m = 1, 2, \ldots, M \). In \( \text{mi} \), \( m = 0 \) is used to refer to the original data, the data containing the missing values. Thus \( \text{mi} \) data in effect contain \( M + 1 \) datasets, corresponding to \( m = 0, m = 1, \ldots, \) and \( m = M \).

**machine precision.** See epsilon(1), etc.

**Mahalanobis distance.** The Mahalanobis distance measure is a scale-invariant way of measuring distance. It takes into account the correlations of the dataset.

**Mahalanobis transformation.** The Mahalanobis transformation takes a Cholesky factorization of the inverse of the covariance matrix \( S^{-1} \) in the formula for Mahalanobis distance and uses it to transform the data. If we have the Cholesky factorization \( S^{-1} = L' L \), then the Mahalanobis transformation of \( x \) is \( z = L x \), and \( z' z = D_M^2 (x) \).
**main effects.** These are average, additive effects that are associated with each level of each factor. For example, the main effect of level $j$ of a factor is the difference between the mean of all observations on the outcome of interest at level $j$ and the grand mean.

**MANCOVA.** MANCOVA is multivariate analysis of covariance. See *multivariate analysis of variance.*

**manifest variables.** Synonym for observed variables.

**MANOVA.** multivariate analysis of variance.

**Mantel–Haenszel test.** The Mantel–Haenszel test evaluates whether the overall degree of association in stratified $2 \times 2$ tables is significant assuming that the exposure effect is the same across strata. See [PSS] power cmh.

**MAR.** See missing at random.

**marginal distribution.** In Bayesian context, a distribution of the data after integrating out parameters from the joint distribution of the parameters and the data.

**marginal homogeneity.** Marginal homogeneity refers to the equality of one or more row marginal proportions with the corresponding column proportions. Also see Introduction under Remarks and examples in [PSS] power pairedproportions.

**marginal likelihood.** In the context of Bayesian model comparison, a marginalized over model parameters $\theta$ likelihood of data $y$ for a given model $M$, $P(y|M) = m(y) = \int P(y|\theta, M)P(\theta|M)d\theta$. Also see Bayes factor.

**marginal posterior distribution.** In Bayesian context, a marginal posterior distribution is a distribution resulting from integrating out all but one parameter from the joint posterior distribution.

**marginal proportion.** This represents a ratio of the number of observations in a row or column of a contingency table relative to the total number of observations. Also see Introduction under Remarks and examples in [PSS] power pairedproportions.

**Markov chain.** Markov chain is a random process that generates sequences of random vectors (or states) and satisfies the Markov property: the next state depends only on the current state and not on any of the previous states. MCMC is the most common methodology for simulating Markov chains.

**mass.** In CA and MCA, the mass is the marginal probability. The sum of the mass over the active row or column categories equals 1.

**.mata file.** By convention, we store the Mata source code for function function() in file function.mata; see [M-1] source.

**matched case–control study.** Also known as a retrospective study, a matched case–control study is a study in which persons with positive outcomes are each matched with one or more persons with negative outcomes but with similar characteristics.

**matched study.** In a matched study, an observation from one group is matched to one or more observations from another group with respect to one or more characteristics of interest. When multiple matches occur, the study design is $1 : M$, where $M$ is the number of matches. Also see paired data, also known as $1 : 1$ matched data.
matching coefficient. The matching similarity coefficient is used to compare two binary variables. If $a$ is the number of observations that both have value 1, and $d$ is the number of observations that both have value 0, and $b, c$ are the number of $(1, 0)$ and $(0, 1)$ observations, respectively, then the matching coefficient is given by

$$\frac{a + d}{a + b + c + d}$$

Also see similarity measure.

matching configuration. In MDS, the matching configuration is the low dimensional configuration whose distances approximate the high-dimensional dissimilarities or disparities. Also see multidimensional scaling, dissimilarity, and disparity.

matching configuration plot. After MDS, this is a scatter plot of the matching configuration.

matching estimator. An estimator that compares differences between the outcomes of similar—that is, matched—individuals. Each individual that receives a treatment is matched to a similar individual that does not get the treatment, and the difference in their outcomes is used to estimate the individual-level treatment effect. Likewise, each individual that does not receive a treatment is matched to a similar individual that does get the treatment, and the difference in their outcomes is used to estimate the individual-level treatment effect.

matrix. The most general organization of data, containing $r$ rows and $c$ columns. Vectors, column vectors, row vectors, and scalars are special cases of matrices.

matrix model parameter. A matrix model parameter is any model parameter that is a matrix. Matrix elements, however, are viewed as scalar model parameters.

Matrix model parameters are defined and referred to within the bayesmh command as $\{param, \text{matrix}\}$ or $\{eqname: \text{param, matrix}\}$ with the equation name $eqname$. For example, $\{\text{Sigma, matrix}\}$ and $\{\text{Scale:Omega, matrix}\}$ are matrix model parameters. Individual matrix elements cannot be referred to within the bayesmh command, but they can be referred within postestimation commands accepting parameters. For example, to refer to the individual elements of the defined above, say, $2 \times 2$ matrices, use $\{\text{Sigma}_1, 1\}$, $\{\text{Sigma}_2, 1\}$, $\{\text{Sigma}_1, 2\}$, $\{\text{Sigma}_2, 2\}$ and $\{\text{Scale:Omega}_1, 1\}$, $\{\text{Scale:Omega}_2, 1\}$, $\{\text{Scale:Omega}_1, 2\}$, $\{\text{Scale:Omega}_2, 2\}$, respectively. See [BAYES] bayesmh.

matrix parameter. See matrix model parameter.

maximum likelihood factor method. The maximum likelihood factor method is a method for performing factor analysis that assumes multivariate normal observations. It maximizes the determinant of the partial correlation matrix; thus, this solution is also meaningful as a descriptive method for nonnormal data. Also see factor analysis.

MCA. See multiple correspondence analysis.

MCAGHQ. See mode-curvature adaptive Gauss–Hermite quadrature.

MCAR. See missing completely at random.

MCE. See Monte Carlo error.

MCMC, Markov chain Monte Carlo. MCMC is a class of simulation-based methods for generating samples from probability distributions. Any MCMC algorithm simulates a Markov chain with a target distribution as its stationary or equilibrium distribution. The precision of MCMC algorithms increases with the number of iterations. The lack of a stopping rule and convergence rule, however, makes it difficult to determine for how long to run MCMC. The time needed to converge to the target distribution within a prespecified error is referred to as mixing time. Better MCMC algorithms have faster mixing times. Some of the popular MCMC algorithms are random-walk Metropolis, Metropolis–Hastings, and Gibbs sampling.
MCMC sample. An MCMC sample is obtained from MCMC sampling. An MCMC sample approximates a target distribution and is used for summarizing this distribution.

MCMC sample size. MCMC sample size is the size of the MCMC sample. It is specified in bayesmh’s option mcmcsize(); see [BAYES] bayesmh.

MCMC sampling, MCMC sampler. MCMC sampling is an MCMC algorithm that generates samples from a target probability distribution.

MCMC standard error, MCSE MCSE is the standard error of the posterior mean estimate. It is defined as the standard deviation divided by the square root of ESS. MCSEs are analogs of standard errors in frequentist statistics and measure the accuracy of the simulated MCMC sample.

McNemar’s test. McNemar’s test is a test used to compare two dependent binary populations. The null hypothesis is formulated in the context of a $2 \times 2$ contingency table as a hypothesis of marginal homogeneity. See [PSS] power pairedproportions and the mcc command in [R] epitab.

MDES. See minimum detectable effect size.

MDS. See multidimensional scaling.

MDS configuration plot. See configuration plot.

mean contrasts. See contrasts.

mean–variance adaptive Gauss–Hermite quadrature. In the context of generalized linear mixed models, mean–variance adaptive Gauss–Hermite quadrature is a method of approximating the integral used in the calculation of the log likelihood. The quadrature locations and weights for individual clusters are updated during the optimization process by using the posterior mean and the posterior standard deviation.

In the context of IRT models, mean–variance adaptive Gauss–Hermite quadrature (MVAGHQ) is a method of approximating the integral used in the calculation of the log likelihood. The quadrature locations and weights for individuals are updated during the optimization process by using the posterior mean and the posterior standard deviation.

measure. A measure is a quantity representing the proximity between objects or method for determining the proximity between objects. Also see proximity.

measure, measurement, x a measurement of X, x measures X. See measurement variables.

measurement models, measurement component. A measurement model is a particular kind of model that deals with the problem of translating observed values to values suitable for modeling. Measurement models are often combined with structural models and then the measurement model part is referred to as the measurement component. See [SEM] intro 5.

measurement variables, measure, measurement, x a measurement of X, x measures X. Observed variable $x$ is a measurement of latent variable $X$ if there is a path connecting $x \leftarrow X$. Measurement variables are modeled by measurement models. Measurement variables are also called indicator variables.

median-linkage clustering. Median-linkage clustering is a hierarchical clustering method that uses the distance between the medians of two groups to determine the similarity or dissimilarity of the two groups. Also see cluster analysis and agglomerative hierarchical clustering methods.

MEFF and MEFT. MEFF and MEFT are misspecification effects. Misspecification effects compare the variance estimate from a given survey dataset with the variance from a misspecified model. In Stata, the misspecified model is fit without weighting, clustering, or stratification.

MEFF is the ratio of two variance estimates. The design-based variance is in the numerator; the misspecified variance is in the denominator.
MEFT is the ratio of two standard-error estimates. The design-based standard error is in the numerator; the misspecified standard error is in the denominator. MEFT is the square root of MEFF.

**method.** Method is just an English word and should be read in context. Nonetheless, method is used here usually to refer to the method used to solve for the fitted parameters of an SEM. Those methods are ML, QML, MLMV, and ADF. Also see technique.

**metric scaling.** Metric scaling is a type of MDS, in which the dissimilarities are transformed to disparities via a class of known functions. This is contrasted to nonmetric scaling. Also see multidimensional scaling.

**Metropolis–Hastings (MH) sampling, MH sampler.** A Metropolis–Hastings (MH) sampler is an MCMC method for simulating probability distributions. According to this method, at each step of the Markov chain, a new proposal state is generated from the current state according to a prespecified proposal distribution. Based on the current and new state, an acceptance probability is calculated and then used to accept or reject the proposed state. Important characteristics of MH sampling is the acceptance rate and mixing time. The MH algorithm is very general and can be applied to an arbitrary target distribution. However, its efficiency is limited, in terms of mixing time, and decreases as the dimension of the target distribution increases. Gibbs sampling, when available, can provide much more efficient sampling than MH sampling.

**mi data.** Any data that have been mi set (see [MI mi set]), whether directly by mi set or indirectly by mi import (see [MI mi import]). The mi data might have no imputations (have $M = 0$) and no imputed variables, at least yet, or they might have $M > 0$ and no imputed variables, or vice versa. An mi dataset might have $M > 0$ and no imputed variables, but the missing values have not yet been replaced with imputed values. Or mi data might have $M > 0$ and imputed variables and the missing values of the imputed variables filled in with imputed values.

**MIMIC.** See multiple indicators and multiple causes.

**minimum detectable effect size.** The minimum detectable effect size is the smallest effect size that can be detected by hypothesis testing for a given power and sample size.

**minimum detectable value.** The minimum detectable value represents the smallest amount or concentration of a substance that can be reliably measured.

**minimum entropy rotation.** The minimum entropy rotation is an orthogonal rotation achieved by minimizing the deviation from uniformity (entropy). The minimum entropy criterion (Jennrich 2004) is

$$c(\Lambda) = -\frac{1}{2} \langle \Lambda^2, \log \Lambda^2 \rangle$$

See Crawford–Ferguson rotation for a definition of $\Lambda$. Also see orthogonal rotation.

**misclassification rate.** The misclassification rate calculated after discriminant analysis is, in its simplest form, the fraction of observations incorrectly classified. See discriminant analysis.

**missing at random.** Missing data are said to be missing at random (MAR) if the probability that data are missing does not depend on unobserved data but may depend on observed data. Under MAR, the missing-data values do not contain any additional information given observed data about the missing-data mechanism. Thus the process that causes missing data can be ignored.

**missing completely at random.** Missing data are said to be missing completely at random (MCAR) if the probability that data are missing does not depend on observed or unobserved data. Under MCAR, the missing data values are a simple random sample of all data values, so any analysis that discards the missing values remains consistent, albeit perhaps inefficient.
missing not at random. Missing data are missing not at random (MNAR) if the probability that data are missing depends on unobserved data. Under MNAR, a missing-data mechanism (the process that causes missing data) must be modeled to obtain valid results.

misspecification effects. See MEFF and MEFT.

mixed design. A mixed design is an experiment that has at least one between-subjects factor and one within-subject factor. See [PSS] power repeated.

mixed model. See mixed-effects model.

mixed-effects model. A mixed-effects model contains both fixed and random effects. The fixed effects are estimated directly, whereas the random effects are summarized according to their (co)variances. Mixed-effects models are used primarily to perform estimation and inference on the regression coefficients in the presence of complicated within-subject correlation structures induced by multiple levels of grouping.

mixing of Markov chain. Mixing refers to the rate at which a Markov chain traverses the parameter space. It is a property of the Markov chain that is different from convergence. Poor mixing indicates a slow rate at which the chain explores the stationary distribution and will require more iterations to provide inference at a given precision. Poor (slow) mixing is typically a result of high correlation between model parameters or of weakly-defined model specifications.

ML, method(ml). ML stands for maximum likelihood. It is a method to obtain fitted parameters. ML is the default method used by sem and gsem. Other available methods for sem are QML, MLMV, and ADF. Also available for gsem is QML.

.mlib library. The object code of functions can be collected and stored in a library. Most Mata functions, in fact, are located in the official libraries provided with Stata. You can create your own libraries. See [M-3] mata mlib.

MLMV, method(mlmv). MLMV stands for maximum likelihood with missing values. It is an ML method used to obtain fitted parameters in the presence of missing values. MLMV is the method used by sem when the method(mlmv) option is specified; method(mlmv) is not available with gsem. Other available methods for use with sem are ML, QML, and ADF. These methods omit from the calculation observations that contain missing values.

mlong data. See style.

MNAR. See missing not at random.

.mo file. The object code of a function can be stored in a .mo file, where it can be later reused. See [M-1] how and [M-3] mata mosave.

mode-curvature adaptive Gauss–Hermite quadrature. In the context of generalized linear mixed models, mode-curvature adaptive Gauss–Hermite quadrature is a method of approximating the integral used in the calculation of the log likelihood. The quadrature locations and weights for individual clusters are updated during the optimization process by using the posterior mode and the standard deviation of the normal density that approximates the log posterior at the mode.

In the context of IRT models, mode-curvature adaptive Gauss–Hermite quadrature (MCAGHQ) is a method of approximating the integral used in the calculation of the log likelihood. The quadrature locations and weights for individuals are updated during the optimization process by using the posterior mode and the standard deviation of the normal density that approximates the log posterior at the mode.

model hypothesis testing. Model hypothesis testing tests hypotheses about models by computing model posterior probabilities.
**model parameter.** A model parameter refers to any (random) parameter in a Bayesian model. Model parameters can be scalars or matrices. Examples of model parameters as defined in `bayesmh` are `{mu}`, `{scale:s}`, `{Sigma,matrix}`, and `{Scale:Omega,matrix}`. See [BAYES] `bayesmh` and, specifically, *Declaring model parameters* and *Referring to model parameters* in that entry. Also see *Different ways of specifying model parameters* in [BAYES] `bayesmh postestimation`.

**model posterior probability.** Model posterior probability is probability of a model $M$ computed conditional on the observed data $y$,

$$P(M|y) = P(M)P(y|M) = P(M)m(y)$$

where $P(M)$ is the prior probability of a model $M$ and $m(y)$ is the marginal likelihood under model $M$.

**modern scaling.** Modern scaling is a form of MDS that is achieved via the minimization of a loss function that compares the disparities (transformed dissimilarities) in the higher-dimensional space and the distances in the lower-dimensional space. Contrast to classical scaling. Also see *dissimilarity, disparity, multidimensional scaling*, and *loss*.

**modification indices.** Modification indices are score tests for adding paths where none appear. The paths can be for either coefficients or covariances.

**moments (of a distribution).** The moments of a distribution are the expected values of powers of a random variable or centralized (demeaned) powers of a random variable. The first moments are the expected or observed means, and the second moments are the expected or observed variances and covariances.

**monadic operator.** Synonym for unary operator.

**monotone-missing pattern, monotone missingness.** A special pattern of missing values in which if the variables are ordered from least to most missing, then all observations of a variable contain missing in the observations in which the prior variable contains missing.

**Monte Carlo error.** Within the multiple-imputation context, a Monte Carlo error is defined as the standard deviation of the multiple-imputation results across repeated runs of the same imputation procedure using the same data. The Monte Carlo error is useful for evaluating the statistical reproducibility of multiple-imputation results. See *Example 6: Monte Carlo error estimates* under Remarks and examples of [MI] `mi estimate`.

**moving-average process.** A moving-average process is a time-series process in which the current value of a variable is modeled as a weighted average of current and past realizations of a white-noise process and, optionally, a time-invariant constant. By convention, the weight on the current realization of the white-noise process is equal to one, and the weights on the past realizations are known as the moving-average (MA) coefficients. A first-order moving-average process, denoted as an MA(1) process, is $y_t = \theta \epsilon_{t-1} + \epsilon_t$.

**multiarm trial.** A multiarm trial is a trial comparing survivor functions of more than two groups.

**multidimensional scaling.** Multidimensional scaling (MDS) is a dimension-reduction and visualization technique. Dissimilarities (for instance, Euclidean distances) between observations in a high-dimensional space are represented in a lower-dimensional space which is typically two dimensions so that the Euclidean distance in the lower-dimensional space approximates in some sense the dissimilarities in the higher-dimensional space. Often the higher-dimensional dissimilarities are first transformed to disparities, and the disparities are then approximated by the distances in the lower-dimensional space. Also see *dissimilarity, disparity, classical scaling, loss, modern scaling, metric scaling*, and *nonmetric scaling*.
multilevel models. Multilevel models are models that include unobserved effects (latent variables) for different groups in the data. For instance, in a dataset of students, groups of students might share the same teacher. If the teacher’s identity is recorded in the data, then one can introduce a latent variable that is constant within teacher and that varies across teachers. This is called a two-level model.

If teachers could in turn be grouped into schools, and school identities were recorded in the data, then one can introduce another latent variable that is constant within school and varies across schools. This is called a three-level (nested-effects) model.

In the above example, observations (students) are said to be nested within teacher nested within school. Sometimes there is no such subsequent nesting structure. Consider workers nested within occupation and industry. The same occupations appear in various industries and the same industries appear within various occupations. We can still introduce latent variables at the occupation and industry level. In such cases, the model is called a crossed-effects model.

The latent variables that we have discussed are also known as random effects. Any coefficients on observed variables in the model are known as the fixed portion of the model. Models that contain fixed and random portions are known as mixed-effects models.

multinomial logit regression. Multinomial logit regression is a term for generalized linear response functions that are family multinomial, link logit. It is used for categorical-outcome data when the outcomes cannot be ordered. Multinomial logit regression is also known as multinomial logistic regression and as mlogit in Stata circles. See generalized linear response functions.

multiple correlation. The multiple correlation is the correlation between endogenous variable $y$ and its linear prediction.

multiple correspondence analysis. Multiple correspondence analysis (MCA) and joint correspondence analysis (JCA) are methods for analyzing observations on categorical variables. MCA and JCA analyze a multiway table and are usually viewed as an extension of CA. Also see correspondence analysis.

multiple indicators and multiple causes. Multiple indicators and multiple causes is a kind of structural model in which observed causes determine a latent variable, which in turn determines multiple indicators. See [SEM] intro 4.

multiple-record st data. See st data.

multivalued treatment effect. A multivalued treatment refers to a treatment that has more than two values. For example, a person could have taken a 20 mg dose of a drug, a 40 mg dose of the drug, or not taken the drug at all.

multivariate analysis of covariance. See multivariate analysis of variance.

multivariate analysis of variance. Multivariate analysis of variance (MANOVA) is used to test hypotheses about means. Four multivariate statistics are commonly computed in MANOVA: Wilks’s lambda, Pillai’s trace, Lawley–Hotelling trace, and Roy’s largest root. Also see Wilks’s lambda, Pillai’s trace, Lawley–Hotelling trace, and Roy’s largest root.

multivariate GARCH models. Multivariate GARCH models are multivariate time-series models in which the conditional covariance matrix of the errors depends on its own past and its past shocks. The acute trade-off between parsimony and flexibility has given rise to a plethora of models; see [TS] mgarch.

multivariate regression. A multivariate regression is a linear regression model in which the regressand is vector valued. Equivalently, a multivariate regression is a linear regression model in which multiple left-hand-side variables are regressed on the same set of explanatory variables simultaneously, allowing the disturbance terms to be contemporaneously correlated. Multivariate regression is
a special case of \textit{seemingly unrelated regression} in which all equations share the same set of explanatory variables.

\textbf{MVAGHQ}. See \textit{mean–variance adaptive Gauss–Hermite quadrature}.

\textbf{NaN}. NaN stands for Not a Number and is a special computer floating-point code used for results that cannot be calculated. Mata (and Stata) do not use NaNs. When NaNs arise, they are converted into \texttt{(missing value)}.

\textbf{nearest neighbor}. See \textit{kth nearest neighbor}.

\textbf{nearest-neighbor matching}. Nearest-neighbor matching uses the distance between observed variables to find similar individuals.

\textbf{negative binomial regression}. Negative binomial regression is a term for generalized linear response functions that are family negative binomial, link log. It is used for count data that are overdispersed relative to Poisson. Negative binomial regression is also known as nbreg in Stata circles. See \textit{generalized linear response functions}.

\textbf{negative binomial regression model}. The negative binomial regression model is for applications in which the dependent variable represents the number of times an event occurs. The negative binomial regression model is an alternative to the Poisson model for use when the dependent variable is overdispersed, meaning that the variance of the dependent variable is greater than its mean.

\textbf{negative effect size}. In power and sample-size analysis, we obtain a negative effect size when the postulated value of the parameter under the alternative hypothesis is less than the hypothesized value of the parameter under the null hypothesis. Also see \textit{positive effect size}.

\textbf{nested random effects}. In the context of mixed-effects models, nested random effects refer to the nested grouping factors for the random effects. For example, we may have data on students who are nested in classes that are nested in schools.

\textbf{nested-effects models}. See \textit{multilevel models}.

\textbf{Newey–West covariance matrix}. The Newey–West covariance matrix is a member of the class of heteroskedasticity- and autocorrelation-consistent (HAC) covariance matrix estimators used with time-series data that produces covariance estimates that are robust to both arbitrary heteroskedasticity and autocorrelation up to a prespecified lag.

\textbf{nominal alpha}, \textbf{nominal significance level}. This is a desired or requested significance level.

\textbf{nominal item}. A nominal item is an item scored in categories that have no natural ordering.

\textbf{nominal response model}. The nominal response model (NRM) is an IRT model for nominal responses. The categories within each item vary in their difficulty and discrimination.

\textbf{noncentrality parameter}. In power and sample-size analysis, a noncentrality parameter is the expected value of the test statistic under the alternative hypothesis.

\textbf{nondirectional test}. See \textit{two-sided test}.

\textbf{noninformative prior}. A noninformative prior is a prior with negligible influence on the posterior distribution. See, for example, \textit{Jeffreys prior}.

\textbf{nonmetric scaling}. Nonmetric scaling is a type of modern MDS in which the dissimilarities may be transformed to disparities via any monotonic function as opposed to a class of known functions. Contrast to \textit{metric scaling}. Also see \textit{multidimensional scaling}, \textit{dissimilarity}, \textit{disparity}, and \textit{modern scaling}.

\textbf{nonparametric methods}. Nonparametric statistical methods, such as KNN discriminant analysis, do not assume the population fits any parameterized distribution.
nonrecursive (structural) model (system), recursive (structural) model (system). A structural model (system) is said to be nonrecursive if there are paths in both directions between one or more pairs of endogenous variables. A system is recursive if it is a system—it has endogenous variables that appear with paths from them—and it is not nonrecursive.

A nonrecursive model may be unstable. Consider, for instance,

\[
\begin{align*}
y_1 &= 2y_2 + 1x_1 + e_1 \\
y_2 &= 3y_1 - 2x_2 + e_2
\end{align*}
\]

This model is unstable. To see this, without loss of generality, treat \(x_1 + e_1\) and \(2x_2 + e_2\) as if they were both 0. Consider \(y_1 = 1\) and \(y_2 = 1\). Those values result in new values \(y_2 = 2\) and \(y_2 = 3\), and those result in new values \(y_1 = 6\) and \(y_2 = 6\), and those result in new values, \ldots. Continue in this manner, and you reach infinity for both endogenous variables. In the jargon of the mathematics used to check for this property, the eigenvalues of the coefficient matrix lie outside the unit circle.

On the other hand, consider these values:

\[
\begin{align*}
y_1 &= 0.5y_2 + 1x_1 + e_1 \\
y_2 &= 1.0y_1 - 2x_2 + e_2
\end{align*}
\]

These results are stable in that the resulting values converge to \(y_1 = 0\) and \(y_2 = 0\). In the jargon of the mathematics used to check for this property, the eigenvalues of the coefficients matrix lie inside the unit circle. Finally, consider the values

\[
\begin{align*}
y_1 &= 0.5y_2 + 1x_1 + e_1 \\
y_2 &= 2.0y_1 - 2x_2 + e_2
\end{align*}
\]

Start with \(y_1 = 1\) and \(y_2 = 1\) and that yields new values \(y_1 = 0.5\) and \(y_2 = 2\) and that yields new values \(y_1 = 1\) and \(y_2 = 1\), and that yields \(y_1 = 0.5\) and \(y_2 = 2\), and it will oscillate forever. In the jargon of the mathematics used to check for this property, the eigenvalues of the coefficients matrix lie on the unit circle. These coefficients are also considered to be unstable.

nonsphericity correction. This is a correction used for the degrees of freedom of a regular \(F\) test in a repeated-measures ANOVA to compensate for the lack of sphericity of the repeated-measures covariance matrix.

norm. A norm is a real-valued function \(f(x)\) satisfying

\[
\begin{align*}
f(0) &= 0 \\
f(x) &> 0 \quad \text{for all } x \neq 0 \\
f(cx) &= |c|f(x) \\
f(x+y) &\leq f(x) + f(y)
\end{align*}
\]

The word norm applied to a vector \(x\) usually refers to its Euclidean norm, \(p = 2\) norm, or length: the square root of the sum of its squared elements. The are other norms, the popular ones being \(p = 1\) (the sum of the absolute values of its elements) and \(p = \infty\) (the maximum element). Norms can also be generalized to deal with matrices. See [M-5] norm().
normality assumption, joint and conditional. The derivation of the standard, linear SEM estimator usually assumes the full joint normality of the observed and latent variables. However, full joint normality can replace the assumption of normality conditional on the values of the exogenous variables, and all that is lost is one goodness-of-fit test (the test reported by `sem` on the output) and the justification for use of optional method MLMV for dealing with missing values. This substitution of assumptions is important for researchers who cannot reasonably assume normality of the observed variables. This includes any researcher including, say, variables age and age-squared in his or her model.

Meanwhile, the generalized SEM makes only the conditional normality assumption.

Be aware that even though the full joint normality assumption is not required for the standard linear SEM, `sem` calculates the log-likelihood value under that assumption. This is irrelevant except that log-likelihood values reported by `sem` cannot be compared with log-likelihood values reported by `gsem`, which makes the lesser assumption.

See [SEM] intro 4.

normalization. Normalization presents information in a standard form for interpretation. In CA the row and column coordinates can be normalized in different ways depending on how one wishes to interpret the data. Normalization is also used in rotation, MDS, and MCA.

normalization constraints. See identification.

normalized residuals. See standardized residuals.

NRM. See nominal response model.

NULL. A special value for a pointer that means “points to nothing”. If you list the contents of a pointer variable that contains NULL, the address will show as 0x0. See pointer.

null hypothesis. In hypothesis testing, the null hypothesis typically represents the conjecture that one is attempting to disprove. Often the null hypothesis is that a treatment has no effect or that a statistic is equal across populations.

null value, null parameter. This value of the parameter of interest under the null hypothesis is fixed by the investigator in a power and sample-size analysis. For example, null mean value and null mean refer to the value of the mean parameter under the null hypothesis.

null-terminator. See binary 0.

numeric. A matrix is said to be numeric if its elements are real or complex; see type, eltype, and orgtype.

object code. Object code refers to the binary code that Mata produces from the source code you type as input. See [M-1] how.

objective prior. See noninformative prior.

object-oriented programming. Object-oriented programming is a programming concept that treats programming elements as objects and concentrates on actions affecting those objects rather than merely on lists of instructions. Object-oriented programming uses classes to describe objects. Classes are much like structures with a primary difference being that classes can contain functions (known as methods) as well as variables. Unlike structures, however, classes may inherit variables and functions from other classes, which in theory makes object-oriented programs easier to extend and modify than non–object-oriented programs.

oblimax rotation. Oblimax rotation is a method for oblique rotation which maximizes the number of high and low loadings. When restricted to orthogonal rotation, oblimax is equivalent to quartimax rotation. Oblimax minimizes the oblimax criterion
\[ c(\Lambda) = -\log(\langle \Lambda^2, \Lambda^2 \rangle) + 2 \log(\langle \Lambda, \Lambda \rangle) \]

See Crawford–Ferguson rotation for a definition of \( \Lambda \). Also see oblique rotation, orthogonal rotation, and quartimax rotation.

**oblimin rotation.** Oblimin rotation is a general method for oblique rotation, achieved by minimizing the oblimin criterion

\[ c(\Lambda) = \frac{1}{4} \langle \Lambda^2, \{I - (\gamma/p)11'\} \Lambda^2 (11' - I) \rangle \]

Oblimin has several interesting special cases:

<table>
<thead>
<tr>
<th>( \gamma )</th>
<th>Special case</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>quartimax / quartimin</td>
</tr>
<tr>
<td>1/2</td>
<td>biquartimax / biquartimin</td>
</tr>
<tr>
<td>1</td>
<td>varimax / covarimin</td>
</tr>
<tr>
<td>( p/2 )</td>
<td>equamax</td>
</tr>
</tbody>
</table>

\( p = \) number of rows of \( A \).

See Crawford–Ferguson rotation for a definition of \( \Lambda \) and \( A \). Also see oblique rotation.

**oblique rotation or oblique transformation.** An oblique rotation maintains the norms of the rows of the matrix but not their inner products. In geometric terms, this maintains the lengths of vectors, but not the angles between them. In contrast, in orthogonal rotation, both are preserved.

**observational data.** In observational data, treatment assignment is not controlled by those who collected the data; thus some common variables affect treatment assignment and treatment-specific outcomes.

**observational study.** In an observational study, as opposed to an experimental study, the assignment of subjects to treatments happens naturally and is thus beyond the control of investigators. Investigators can only observe subjects and measure their characteristics. For example, a study that evaluates the effect of exposure of children to household pesticides is an observational study.

**observations and variables.** A dataset containing \( n \) observations on \( k \) variables in often stored in an \( n \times k \) matrix. An observation refers to a row of that matrix; a variable refers to a column.

**observed level of significance.** See \( p \)-value.

**observed variables.** A variable is observed if it is a variable in your dataset. In this documentation, we often refer to observed variables by using \( x_1, x_2, \ldots, y_1, y_2, \) and so on; in reality, observed variables have names such as mpg, weight, testscore, etc.

In the software, observed variables are usually indicated by having names that are all lowercase. Also see latent variable.

**odds** and **odds ratio.** The odds in favor of an event are \( o = p/(1 - p) \), where \( p \) is the probability of the event. Thus if \( p = 0.2 \), the odds are 0.25, and if \( p = 0.8 \), the odds are 4.

The log of the odds is \( \ln(o) = \logit(p) = \ln\{p/(1 - p)\} \), and logistic-regression models, for instance, fit \( \ln(o) \) as a linear function of the covariates.

The odds ratio is a ratio of two odds: \( o_1/o_0 \). The individual odds that appear in the ratio are usually for an experimental group and a control group, or two different demographic groups.
offset variable and exposure variable. An offset variable is a variable that is to appear on the right-hand side of a model with coefficient 1:

\[ y_j = \text{offset}_j + b_0 + b_1 x_j + \cdots \]

In the above, \(b_0\) and \(b_1\) are to be estimated. The offset is not constant. Offset variables are often included to account for the amount of exposure. Consider a model where the number of events observed over a period is the length of the period multiplied by the number of events expected in a unit of time:

\[ n_j = T_j e(X_j) \]

When we take logs, this becomes

\[ \log(n_j) = \log(T_j) + \log\{e(X_j)\} \]

\(\ln(T_j)\) is an offset variable in this model.

When the log of a variable is an offset variable, the variable is said to be an exposure variable. In the above, \(T_j\) is an exposure variable.

**OIM, vce(oim).** OIM stands for observed information matrix, defined as the inverse of the negative of the matrix of second derivatives, usually of the log likelihood function. The OIM is an estimate of the VCE. OIM is the default VCE that \texttt{sem} and \texttt{gsem} report. The other available techniques are EIM, OPG, robust, clustered, bootstrap, and jackknife.

**one-at-a-time MCMC sampling.** A one-at-a-time MCMC sample is an MCMC sampling procedure in which random variables are sampled individually, one at a time. For example, in \textit{Gibbs sampling}, individual variates are sampled one at a time, conditionally on the most recent values of the rest of the variates.

**one-level model.** A one-level model has no multilevel structure and no random effects. Linear regression is a one-level model.

**one-parameter logistic model.** The one-parameter logistic (1PL) model is an IRT model for binary responses where items vary in their difficulty but share the same discrimination parameter.

**one-sample test.** A one-sample test compares a parameter of interest from one sample with a reference value. For example, a one-sample mean test compares a mean of the sample with a reference value.

**one-sided test, one-tailed test.** A one-sided test is a hypothesis test of a scalar parameter in which the alternative hypothesis is one sided, meaning that the alternative hypothesis states that the parameter is either less than or greater than the value conjectured under the null hypothesis but not both. Also see \textit{One-sided test versus two-sided test} under Remarks and examples in [PSS] intro.

**one-step-ahead forecast.** See \textit{static forecast}.

**one-way ANOVA, one-way analysis of variance.** A one-way ANOVA model has a single factor. Also see [PSS] power oneway.

**one-way repeated-measures ANOVA.** A one-way repeated-measures ANOVA model has a single within-subject factor. Also see [PSS] power repeated.

**operating characteristic curve.** See \textit{category characteristic curve}.

**operator.** An operator is \(+\), \(-\), and the like. Most operators are binary (or dyadic), such as \(+\) in \(A+B\) and \(\ast\) in \(C\ast D\). Binary operators also include logical operators such as \& and \| (\textquotedblright and\textquotedblright and \textquotedblright or\textquotedblright) in \(E\&F\) and \(G\|H\). Other operators are unary (or monadic), such as \(!\) (not) in \(!J\), or both unary and binary, such as \(-\) in \(-K\) and in \(L\!-\!M\). When we say \textquotedblright operator\textquotedblright without specifying which, we mean binary operator. Thus colon operators are in fact colon binary operators. See [M-2] exp.
OPG, vce(opg). OPG stands for outer product of the gradients, defined as the cross product of the observation-level first derivatives, usually of the log likelihood function. The OPG is an estimate of the VCE. The other available techniques are OIM, EIM, robust, clustered, bootstrap, and jackknife.

optimization. Mata compiles the code that you write. After compilation, Mata performs an optimization step, the purpose of which is to make the compiled code execute more quickly. You can turn off the optimization step—see [M-3] _mata set—but doing so is not recommended.

ordered complementary log-log regression. Ordered complementary log-log regression is a term for generalized linear response functions that are family ordinal, link cloglog. It is used for ordinal-outcome data. Ordered complementary log-log regression is also known as ocloglog in Stata circles. See _generalized linear response functions._

ordered logit regression. Ordered logit regression is a term for generalized linear response functions that are family ordinal, link logit. It is used for ordinal outcome data. Ordered logit regression is also known as ordered logistic regression, as just ordered logit, and as ologit in Stata circles. See _generalized linear response functions._

ordered probit regression. Ordered probit regression is a term for generalized linear response functions that are family ordinal, link probit. It is used for ordinal-outcome data. Ordered probit regression is also known as just ordered probit and known as oprobit in Stata circles. See _generalized linear response functions._

ordinal item. An ordinal item is an item scored on a scale where a higher score indicates a “higher” outcome.

ordination. Ordination is the ordering of a set of data points with respect to one or more axes. MDS is a form of ordination.

orgtype. See _type, eltype, and orgtype._

original data. Original data are the data as originally collected, with missing values in place. In _mi_ data, the original data are stored in _m_ = 0. The original data can be extracted from _mi_ data by using _mi extract_; see [MI] _mi extract._

orthogonal matrix and unitary matrix. A is orthogonal if A is square and _A′A_ = _I_. The word orthogonal is usually reserved for real matrices; if the matrix is complex, it is said to be unitary (and then transpose means conjugate-transpose). We use the word orthogonal for both real and complex matrices.

If A is orthogonal, then _det(A) = ±1._

orthogonal rotation or orthogonal transformation. Orthogonal rotation maintains both the norms of the rows of the matrix and also inner products of the rows of the matrix. In geometric terms, this maintains both the lengths of vectors and the angles between them. In contrast, oblique rotation maintains only the norms, that is, the lengths of vectors.

orthogonalized impulse–response function. An orthogonalized impulse–response function (OIRF) measures the effect of an orthogonalized shock to an endogenous variable on itself or another endogenous variable. An orthogonalized shock is one that affects one variable at time _t_ but no other variables. See [TS] _irf create_ for a discussion of the difference between IRFs and OIRFs.

outcome model. An outcome model is a model used to predict the outcome as a function of covariates and parameters.

overdispersion. In count-data models, overdispersion occurs when there is more variation in the data than would be expected if the process were Poisson.

overidentifying restrictions. The order condition for model identification requires that the number of exogenous variables excluded from the model be at least as great as the number of endogenous
regressors. When the number of excluded exogenous variables exceeds the number of endogenous regressors, the model is overidentified, and the validity of the instruments can then be checked via a test of overidentifying restrictions.

**overlap assumption.** The overlap assumption requires that each individual have a positive probability of each possible treatment level.

**paired data.** Paired data consist of pairs of observations that share some characteristics of interest. For example, measurements on twins, pretest and posttest measurements, before and after measurements, repeated measurements on the same individual. Paired data are correlated and thus must be analyzed by using a paired test.

**paired observations.** See paired data.

**paired test.** A paired test is used to test whether the parameters of interest of two paired populations are equal. The test takes into account the dependence between measurements. For this reason, paired tests are usually more powerful than their two-sample counterparts. For example, a paired-means or paired-difference test is used to test whether the means of two paired (correlated) populations are equal.

**panel data.** Panel data are data in which the same units were observed over multiple periods. The units, called panels, are often firms, households, or patients who were observed at several points in time. In a typical panel dataset, the number of panels is large, and the number of observations per panel is relatively small.

**panel-corrected standard errors (PCSEs).** The term panel-corrected standard errors refers to a class of estimators for the variance–covariance matrix of the OLS estimator when there are relatively few panels with many observations per panel. PCSEs account for heteroskedasticity, autocorrelation, or cross-sectional correlation.

**parameter constraints.** Parameter constraints are restrictions placed on the parameters of the model. These constraints are typically in the form of 0 constraints and equality constraints. A 0 constraint is implied, for instance, when no path is drawn connecting \( x \) with \( y \). An equality constraint is specified when one path coefficient is forced to be equal to another or one covariance is forced to be equal to another.

Also see identification.

**parameters, ancillary parameters.** The parameters are the to-be-estimated coefficients of a model. These include all path coefficients, means, variances, and covariances. In mathematical notation, the theoretical parameters are often written as \( \theta = (\alpha, \beta, \mu, \Sigma) \), where \( \alpha \) is the vector of intercepts, \( \beta \) is the vector of path coefficients, \( \mu \) is the vector of means, and \( \Sigma \) is the matrix of variances and covariances. The resulting parameters estimates are written as \( \hat{\theta} \).

Ancillary parameters are extra parameters beyond the ones just described that concern the distribution. These include the scale parameter of gamma regression, the dispersion parameter for negative binomial regression, and the cutpoints for ordered probit, logit, and cloglog regression, and the like. These parameters are also included in \( \theta \).

**parametric methods.** Parametric statistical methods, such as LDA and QDA, assume the population fits a parameterized distribution. For example, for LDA we assume the groups are multivariate normal with equal covariance matrices.

**parsimax rotation.** Parsimax rotation is an orthogonal rotation that balances complexity between the rows and the columns. It is equivalent to the Crawford–Ferguson family with \( \kappa = (f-1)/(p+f-2) \), where \( p \) is the number of rows of the original matrix, and \( f \) is the number of columns. See orthogonal rotation and Crawford–Ferguson rotation.
**partial autocorrelation function.** The partial autocorrelation function (PACF) expresses the correlation between periods $t$ and $t - k$ of a time series as a function of the time $t$ and lag $k$, after controlling for the effects of intervening lags. For a stationary time series, the PACF does not depend on $t$. The PACF is not symmetric about $k = 0$: the partial autocorrelation between $y_t$ and $y_{t-k}$ is not equal to the partial autocorrelation between $y_t$ and $y_{t+k}$.

**partial credit model.** The partial credit model (PCM) is an IRT model for ordinal responses. The categories across all items vary in their difficulty and share the same discrimination parameter.

**partial DFBETA.** A partial DFBETA measures the change in the regressor’s coefficient because of deletion of that individual record. In single-record data, the partial DFBETA is equal to the DFBETA. Also see DFBETA.

**partial likelihood displacement value.** A partial likelihood displacement value is an influence measure of the effect of deleting an individual record on the coefficient vector. For single-record data, the partial likelihood displacement value is equal to the likelihood displacement value. Also see likelihood displacement value.

**partial LMAX value.** A partial LMAX value is an influence measure of the effect of deleting an individual record on the overall coefficient vector and is based on an eigensystem analysis of efficient score residuals. In single-record data, the partial LMAX value is equal to the LMAX value. Also see LMAX value.

**partially specified target rotation.** Partially specified target rotation minimizes the criterion

$$c(\Lambda) = \|W \otimes (\Lambda - H)\|^2$$

for a given target matrix $H$ and a nonnegative weighting matrix $W$ (usually zero–one valued). See Crawford–Ferguson rotation for a definition of $\Lambda$.

**partition clustering and partition cluster-analysis methods.** Partition clustering methods break the observations into a distinct number of nonoverlapping groups. This is accomplished in one step, unlike hierarchical cluster-analysis methods, in which an iterative procedure is used. Consequently, this method is quicker and will allow larger datasets than the hierarchical clustering methods. Contrast to hierarchical clustering. Also see kmeans and kmedians.

**passive variable.** See imputed, passive, and regular variables.

**past history.** Past history is information recorded about a subject before the subject was both at risk and under observation. Consider a dataset that contains information on subjects from birth to death and an analysis in which subjects became at risk once diagnosed with a particular kind of cancer. The past history on the subject would then refer to records before the subjects were diagnosed. The word history is often dropped, and the term becomes simply past. For instance, we might want to know whether a subject smoked in the past.

Also see future history.

**path.** A path, typically shown as an arrow drawn from one variable to another, states that the first variable determines the second variable, at least partially. If $x \rightarrow y$, or equivalently $y \leftarrow x$, then $y_j = \alpha + \cdots + \beta x_j + \cdots + e.y_j$, where $\beta$ is said to be the $x \rightarrow y$ path coefficient. The ellipses are included to account for paths to $y$ from other variables. $\alpha$ is said to be the intercept and is automatically added when the first path to $y$ is specified.

A curved path is a curved line connecting two variables, and it specifies that the two variables are allowed to be correlated. If there is no curved path between variables, the variables are usually assumed to be uncorrelated. We say usually because correlation is assumed among observed exogenous variables and, in the command language, assumed among latent exogenous variables, and if some of the correlations are not desired, they must be suppressed. Many authors refer to
covariances rather than correlations. Strictly speaking, the curved path denotes a nonzero covariance. A correlation is often called a standardized covariance.

A curved path can connect a variable to itself and in that case, indicates a variance. In path diagrams in this manual, we typically do not show such variance paths even though variances are assumed.

**path coefficient.** The path coefficient is associated with a path; see path. Also see intercept.

**path diagram.** A path diagram is a graphical representation that shows the relationships among a set of variables using paths. See [SEM intro 2] for a description of path diagrams.

**path notation.** Path notation is a syntax defined by the authors of Stata’s sem and gsem commands for entering path diagrams in a command language. Models to be fit may be specified in path notation or they may be drawn using path diagrams into the Builder.

**PCA.** See principal component analysis.

**PCM.** See partial credit model.

**p-conformability.** Matrix, vector, or scalar $A$ is said to be p-conformable with matrix, vector, or scalar $B$ if $\text{rows}(A) == \text{rows}(B)$ and $\text{cols}(A) == \text{cols}(B)$. $p$ stands for plus; p-conformability is one of the properties necessary to be able to add matrices together. p-conformability, however, does not imply that the matrices are of the same type. Thus $(1,2,3)$ is p-conformable with $(4,5,6)$ and with ("this","that","what") but not with $(4\backslash 5\backslash 6)$.

**Pearson’s correlation.** Pearson’s correlation $\rho$, also known as the product-moment correlation, measures the degree of association between two variables. Pearson’s correlation equals the variables’ covariance divided by their respective standard deviations, and ranges between $-1$ and $1$. Zero indicates no correlation between the two variables.

**penalized log-likelihood function.** This is a log-likelihood function that contains an added term, usually referred to as a roughness penalty, that reduces its value when the model overfits the data. In Cox models with frailty, such functions are used to prevent the variance of the frailty from growing too large, which would allow the individual frailty values to perfectly fit the data.

**periodogram.** A periodogram is a graph of the spectral density function of a time series as a function of frequency. The pergram command first standardizes the amplitude of the density by the sample variance of the time series, and then plots the logarithm of that standardized density. Peaks in the periodogram represent cyclical behavior in the data.

**permutation matrix and permutation vector.** A permutation matrix is an $n \times n$ matrix that is a row (or column) permutation of the identity matrix. If $P$ is a permutation matrix, then $P \ast A$ permutes the rows of $A$ and $A \ast P$ permutes the columns of $A$. Permutation matrices also have the property that $P^{-1} = P'$.

A permutation vector is a $1 \times n$ or $n \times 1$ vector that contains a permutation of the integers $1, 2, \ldots, n$. Permutation vectors can be used with subscripting to reorder the rows or columns of a matrix. Permutation vectors are a memory-conserving way of recording permutation matrices; see [M-1] permutation.

**person location.** Location of a person on the latent trait scale.

**phase function.** The phase function of a linear filter specifies how the filter changes the relative importance of the random components at different frequencies in the frequency domain.

**Pillai’s trace.** Pillai’s trace is a test statistic for the hypothesis test $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ based on the eigenvalues $\lambda_1, \ldots, \lambda_s$ of $E^{-1}H$. It is defined as
\[ V^{(s)} = \text{trace}[(E + H)^{-1}H] = \sum_{i=1}^{s} \frac{\lambda_i}{1 + \lambda_i} \]

where \( H \) is the between matrix and \( E \) is the within matrix. See \textit{between matrix}.

**plain ASCII.** We use plain ASCII as a nontechnical term to refer to what computer programmers call lower ASCII. These are the plain Latin letters “a” to “z” and “A” to “Z”; numbers “0” through “9”; many punctuation marks, such as “!”; simple mathematical symbols, such as “+”; and whitespace and control characters such as space (“ “), tab, and carriage return.

Each plain ASCII character is stored as a single byte with a value between 0 and 127. Another distinguishing feature is that the byte values used to encode plain ASCII characters are the same across different operating systems and are common between ASCII and UTF-8.

Also see \textit{ASCII} and \textit{encodings}.

**point estimate.** A point estimate is another name for a statistic, such as a mean or regression coefficient.

**pointer.** A matrix is said to be a pointer matrix if its elements are pointers.

A pointer is the address of a variable. Say that variable \( X \) contains a matrix. Another variable \( p \) might contain 137,799,016 and, if 137,799,016 were the address at which \( X \) were stored, then \( p \) would be said to point to \( X \). Addresses are seldom written in base 10, and so rather than saying \( p \) contains 137,799,016, we would be more likely to say that \( p \) contains 0x836a568, which is the way we write numbers in base 16. Regardless of how we write addresses, however, \( p \) contains a number and that number corresponds to the address of another variable.

In our program, if we refer to \( p \), we are referring to \( p \)'s contents, the number 0x836a568. The monadic operator \(*\) is defined as “refer to the address” or “dereference”: \( *p \) means \( X \). We could code \( Y = *p \) or \( Y = X \), and either way, we would obtain the same result. In our program, we could refer to \( X[i,j] \) or \( (\ast p)[i,j] \), and either way, we would obtain the \( i, j \) element of \( X \).

The monadic operator \& is how we put addresses into \( p \). To load \( p \) with the address of \( X \), we code \( p = \&X \).

The special address 0 (zero, written in hexadecimal as 0 \times 0), also known as \texttt{NULL}, is how we record that a pointer variable points to nothing. A pointer variable contains \texttt{NULL} or it contains a valid address of another variable.

See [M-2] \texttt{pointers} for a complete description of pointers and their use.

**Poisson regression.** Poisson regression is a term for generalized linear response functions that are family Poisson, link log. It is used for count data. See \textit{generalized linear response functions}.

**Poisson regression model.** The Poisson regression model is used when the dependent variable represents the number of times an event occurs. In the Poisson model, the variance of the dependent variable is equal to the conditional mean.

**polytomous item.** See \textit{categorical item}.

**POMs.** See \textit{potential-outcome means}.

**pooled estimator.** A pooled estimator ignores the longitudinal or panel aspect of a dataset and treats the observations as if they were cross-sectional.

**population-averaged model.** A population-averaged model is used for panel data in which the parameters measure the effects of the regressors on the outcome for the average individual in the population. The panel-specific errors are treated as uncorrelated random variables drawn from a population with zero mean and constant variance, and the parameters measure the effects of the regressors on the dependent variable after integrating over the distribution of the random effects.
**Glossary**

**portmanteau statistic.** The portmanteau, or $Q$, statistic is used to test for white noise and is calculated using the first $m$ autocorrelations of the series, where $m$ is chosen by the user. Under the null hypothesis that the series is a white-noise process, the portmanteau statistic has a $\chi^2$ distribution with $m$ degrees of freedom.

**positive effect size.** In power and sample-size analysis, we obtain a positive effect size when the postulated value of the parameter under the alternative hypothesis is greater than the hypothesized value of the parameter under the null hypothesis. Also see negative effect size.

**posterior distribution, posterior.** A posterior distribution is a probability distribution of model parameters conditional on observed data. The posterior distribution is determined by the likelihood of the parameters and their prior distribution. For a parameter vector $\theta$ and data $y$, the posterior distribution is given by

$$P(\theta | y) = \frac{P(\theta) P(y | \theta)}{P(y)}$$

where $P(\theta)$ is the prior distribution, $P(y | \theta)$ is the model likelihood, and $P(y)$ is the marginal distribution for $y$. Bayesian inference is based on a posterior distribution.

**posterior independence.** See independent a posteriori.

**posterior interval.** See credible interval.

**posterior mean.** In generalized linear mixed-effects models, posterior mean refers to the predictions of random effects based on the mean of the posterior distribution.

In IRT models, posterior mean refers to the predictions of the latent trait based on the mean of the posterior distribution.

**posterior mode.** In generalized linear mixed-effects models, posterior mode refers to the predictions of random effects based on the mode of the posterior distribution.

In IRT models, posterior mode refers to the predictions of the latent trait based on the mode of the posterior distribution.

**posterior odds.** Posterior odds for $\theta_1$ compared with $\theta_2$ is the ratio of posterior density evaluated at $\theta_1$ and $\theta_2$ under a given model,

$$\frac{p(\theta_1 | y)}{p(\theta_2 | y)} = \frac{p(\theta_1) p(y | \theta_1)}{p(\theta_2) p(y | \theta_2)}$$

In other words, posterior odds are prior odds times the likelihood ratio.

**posterior predictive distribution.** A posterior predictive distribution is a distribution of unobserved (future) data conditional on the currently observed data. Posterior predictive distribution is derived by marginalizing the likelihood function with respect to the posterior distribution of model parameters.

**posterior probabilities.** After discriminant analysis, the posterior probabilities are the probabilities of a given observation being assigned to each of the groups based on the prior probabilities, the training data, and the particular discriminant model. Contrast to prior probabilities.

**poststratification.** Poststratification is a method for adjusting sampling weights, usually to account for underrepresented groups in the population. This usually results in decreased bias because of nonresponse and underrepresented groups in the population. Poststratification also tends to result in smaller variance estimates.

The population is partitioned into categories, called poststrata. The sampling weights are adjusted so that the sum of the weights within each poststratum is equal to the respective poststratum size. The poststratum size is the number of individuals in the population that are in the poststratum. The frequency distribution of the poststrata typically comes from census data, and the poststrata are most commonly identified by demographic information such as age, sex, and ethnicity.
postulated value. See alternative value.

potential outcome. The potential outcome is the outcome an individual would obtain if given a specific treatment.

For example, an individual has one potential blood pressure after taking a pill and another potential blood pressure had that person not taken the pill.

potential-outcome means. The potential-outcome means refers to the means of the potential outcomes for a specific treatment level.

The mean blood pressure if everyone takes a pill and the mean blood pressure if no one takes a pill are two examples.

The average treatment effect is the difference between potential-outcome mean for the treated and the potential-outcome mean for the not treated.

power. The power of a test is the probability of correctly rejecting the null hypothesis when it is false. It is often denoted as 1 − β in statistical literature, where β is the type II error probability. Commonly used values for power are 80% and 90%. Also see type I error and type II error. power and sample-size analysis. Power and sample-size analysis investigates the optimal allocation of study resources to increase the likelihood of the successful achievement of a study objective. See [PSS] intro.

power curve. A power curve is a graph of the estimated power as a function of some other study parameter such as the sample size. The power is plotted on the y axis, and the sample size or other parameter is plotted on the x axis. See [PSS] power, graph.

power determination. This pertains to the computation of a power given sample size, effect size, and other study parameters.

power function. The power functions is a function of the population parameter θ, defined as the probability that the observed sample belongs to the rejection region of a test for given θ. See Hypothesis testing under Remarks and examples in [PSS] intro.

power graph. See power curve.

pragma. “(Pragmatic information) A standardised form of comment which has meaning to a compiler. It may use a special syntax or a specific form within the normal comment syntax. A pragma usually conveys non-essential information, often intended to help the compiler to optimise the program.” See The Free On-line Dictionary of Computing, http://www.foldoc.org/, Editor Denis Howe. For Mata, see [M-2] pragma.

Prais–Winsten estimator. A Prais–Winsten estimator is a linear regression estimator that is used when the error term exhibits first-order autocorrelation; see also Cochrane–Orcutt estimator. Here the first observation in the dataset is transformed as \( \hat{y}_1 = \sqrt{1 - \rho^2} y_1 \) and \( \hat{x}_1 = \sqrt{1 - \rho^2} x_1 \), so that the first observation is not lost. The Prais–Winsten estimator is a generalized least-squares estimator.

predetermined variable. A predetermined variable is a regressor in which its contemporaneous and future values are not correlated with the unobservable error term but past values are correlated with the error term.

predictive margins. Predictive margins provide a way of exploring the response surface of a fitted model in any response metric of interest—means, linear predictions, probabilities, marginal effects, risk differences, and so on. Predictive margins are estimates of responses (or outcomes) for the groups represented by the levels of a factor variable, controlling for the differing covariate distributions across the groups. They are the survey-data and nonlinear response analogue to what are often called estimated marginal means or least-squares means for linear models.
Because these margins are population-weighted averages over the estimation sample or subsamples, and because they take account of the sampling distribution of the covariates, they can be used to make inferences about treatment effects for the population.

**Prevented fraction.** A prevented fraction is the reduction in the risk of a disease or other condition of interest caused by including a protective risk factor or public-health intervention.

**Prewhiten.** To prewhiten is to apply a transformation to a time series so that it becomes white noise.

**Primary sampling unit.** Primary sampling unit (PSU) is a cluster that was sampled in the first sampling stage; see cluster.

**Priming values.** Priming values are the initial, preestimation values used to begin a recursive process.

**Principal component analysis.** Principal component analysis (PCA) is a statistical technique used for data reduction. The leading eigenvectors from the eigen decomposition of the correlation or the covariance matrix of the variables describe a series of uncorrelated linear combinations of the variables that contain most of the variance. In addition to data reduction, the eigenvectors from a PCA are often inspected to learn more about the underlying structure of the data.

**Principal factor method.** The principal factor method is a method for factor analysis in which the factor loadings, sometimes called factor patterns, are computed using the squared multiple correlations as estimates of the communality. Also see factor analysis and communality.

**Prior distribution, prior.** In Bayesian statistics, prior distributions are probability distributions of model parameters formed based on some a priori knowledge about parameters. Prior distributions are independent of the observed data.

**Prior independence.** See independent a priori.

**Prior odds.** Prior odds for $\theta_1$ compared with $\theta_2$ is the ratio of prior density evaluated at $\theta_1$ and $\theta_2$ under a given model, $p(\theta_1)/p(\theta_2)$. Also see posterior odds.

**Prior probabilities.** Prior probabilities in discriminant analysis are the probabilities of an observation belonging to a group before the discriminant analysis is performed. Prior probabilities are often based on the prevalence of the groups in the population as a whole. Contrast to posterior probabilities.

**Probability of a type I error.** This is the probability of committing a type I error of incorrectly rejecting the null hypothesis. Also see significance level.

**Probability of a type II error.** This is the probability of committing a type II error of incorrectly accepting the null hypothesis. Common values for the probability of a type II error are 0.1 and 0.2 or, equivalently, 10% and 20%. Also see beta and power.

**Probability weight.** Probability weight is another term for sampling weight.

**Probit regression.** Probit regression is a term for generalized linear response functions that are family Bernoulli, link probit. It is used for binary outcome data. Probit regression is also known simply as probit. See generalized linear response functions.

**Procrustes rotation.** A Procrustes rotation is an orthogonal or oblique transformation, that is, a restricted Procrustes transformation without translation or dilation (uniform scaling).

**Procrustes transformation.** The goal of Procrustes transformation is to transform the source matrix $X$ to be as close as possible to the target $Y$. The permitted transformations are any combination of dilation (uniform scaling), rotation and reflection (that is, orthogonal or oblique transformations), and translation. Closeness is measured by residual sum of squares. In some cases, unrestricted Procrustes transformation is desired; this allows the data to be transformed not just by orthogonal or oblique rotations, but by all conformable regular matrices $A$. Unrestricted Procrustes transformation is equivalent to a multivariate regression.
The name comes from Procrustes of Greek mythology; Procrustes invited guests to try his iron bed. If the guest was too tall for the bed, Procrustes would amputate the guest’s feet, and if the guest was too short, he would stretch the guest out on a rack.

Also see orthogonal rotation, oblique rotation, dilation, and multivariate regression.

**production function.** A production function describes the maximum amount of a good that can be produced, given specified levels of the inputs.

**promax power rotation.** Promax power rotation is an oblique rotation. It does not fit in the minimizing-a-criterion framework that is at the core of most other rotations. The promax method (Hendrickson and White 1964) was proposed before computing power became widely available. The promax rotation consists of three steps:

1. Perform an orthogonal rotation.
2. Raise the elements of the rotated matrix to some power, preserving the sign of the elements. Typically the power is in the range $2 \leq \text{power} \leq 4$. This operation is meant to distinguish clearly between small and large values.
3. The matrix from step two is used as the target for an oblique Procrustean rotation from the original matrix.

**propensity score.** The propensity score is the probability that an individual receives a treatment.

**propensity-score matching.** Propensity-score matching uses the distance between estimated propensity scores to find similar individuals.

**proportional hazards model.** This is a model in which, between individuals, the ratio of the instantaneous failure rates (the hazards) is constant over time.

**proposal distribution.** In the context of the MH algorithm, a proposal distribution is used for defining the transition steps of the Markov chain. In the standard random-walk Metropolis algorithm, the proposal distribution is a multivariate normal distribution with zero mean and adaptable covariance matrix.

**prospective study.** In a prospective study, the population or cohort is classified according to specific risk factors, such that the outcome of interest, typically various manifestations of a disease, can be observed over time and tied in to the initial classification. Also see retrospective study.

Also known as a prospective longitudinal study, a prospective study is a study based on observations over the same subjects for a given period.

**proximity, proximity matrix, and proximity measure.** Proximity or a proximity measure means the nearness or farness of two things, such as observations or variables or groups of observations or a method for quantifying the nearness or farness between two things. A proximity is measured by a similarity or dissimilarity. A proximity matrix is a matrix of proximities. Also see similarity and dissimilarity.

**pseudoconvergence.** A Markov chain may appear to converge when in fact it did not. We refer to this phenomenon as pseudoconvergence. Pseudoconvergence is typically caused by multimodality of the stationary distribution, in which case the chain may fail to traverse the weakly connected regions of the distribution space. A common way to detect pseudoconvergence is to run multiple chains using different starting values and to verify that all of the chain converge to the same target distribution.

**pseudolikelihood.** A pseudolikelihood is a weighted likelihood that is used for point estimation. Pseudolikelihoods are not true likelihoods because they do not represent the distribution function for the sample data from a survey. The sampling distribution is instead determined by the survey design.
PSS analysis. See power and sample-size analysis.

PSS Control Panel. The PSS Control Panel is a point-and-click graphical user interface for power and sample-size analysis. See [PSS] GUI.

PSU. See primary sampling unit.

p-value. P-value is a probability of obtaining a test statistic as extreme or more extreme as the one observed in a sample assuming the null hypothesis is true.

QDA. See quadratic discriminant analysis.

QML, method(ml) vce(robust). QML stands for quasimaximum likelihood. It is a method used to obtain fitted parameters, and a technique used to obtain the corresponding VCE. QML is used by sem and gsem when options method(ml) and vce(robust) are specified. Other available methods are ML, MLMV, and ADF. Other available techniques are OIM, EIM, OPG, clustered, bootstrap, and jackknife.

QR decomposition. QR decomposition is an orthogonal-triangular decomposition of an augmented data matrix that speeds up the calculation of the log likelihood; see Methods and formulas in [ME] mixed for more details.

quadratic discriminant analysis. Quadratic discriminant analysis (QDA) is a parametric form of discriminant analysis and is a generalization of LDA. Like LDA, QDA assumes that the observations come from a multivariate normal distribution, but unlike LDA, the groups are not assumed to have equal covariance matrices. Also see discriminant analysis, linear discriminant analysis, and parametric methods.

quadrature. Quadrature is a method for performing numerical integration. gsem uses quadrature in any model including latent variables (excluding error variables). sem, being limited to linear models, does not need to perform quadrature.

quartimax rotation. Quartimax rotation maximizes the variance of the squared loadings within the rows of the matrix. It is an orthogonal rotation that is equivalent to minimizing the criterion

$$c(\Lambda) = \sum_i \sum_r \lambda^4_{ir} = -\frac{1}{4} \left< \Lambda^2, \Lambda^2 \right>$$

See Crawford–Ferguson rotation for a definition of \(\Lambda\).

quartimin rotation. Quartimin rotation is an oblique rotation that is equivalent to quartimax rotation when quartimin is restricted to orthogonal rotations. Quartimin is equivalent to oblimin rotation with \(\gamma = 0\). Also see quartimax rotation, oblique rotation, orthogonal rotation, and oblimin rotation.

random coefficient. In the context of mixed-effects models, a random coefficient is a counterpart to a slope in the fixed-effects equation. You can think of a random coefficient as a randomly varying slope at a specific level of nesting.

random effects. In the context of mixed-effects models, random effects represent effects that may vary from group to group at any level of nesting. In the ANOVA literature, random effects represent the levels of a factor for which the inference can be generalized to the underlying population represented by the levels observed in the study. See also random-effects model in [XT] Glossary. For a information related to Bayesian analysis, see random-effects parameters.

random intercept. In the context of mixed-effects models, a random intercept is a counterpart to the intercept in the fixed-effects equation. You can think of a random intercept as a randomly varying intercept at a specific level of nesting.
random walk. A random walk is a time-series process in which the current period’s realization is equal to the previous period’s realization plus a white-noise error term: $y_t = y_{t-1} + \epsilon_t$. A random walk with drift also contains a nonzero time-invariant constant: $y_t = \delta + y_{t-1} + \epsilon_t$. The constant term $\delta$ is known as the drift parameter. An important property of random-walk processes is that the best predictor of the value at time $t + 1$ is the value at time $t$ plus the value of the drift parameter.

random-coefficients model. A random-coefficients model is a panel-data model in which group-specific heterogeneity is introduced by assuming that each group has its own parameter vector, which is drawn from a population common to all panels.

random-effects linear form. A linear form representing a random-effects variable that can be used in substitutable expressions.

random-effects model. A random-effects model for panel data treats the panel-specific errors as uncorrelated random variables drawn from a population with zero mean and constant variance. The regressors must be uncorrelated with the random effects for the estimates to be consistent. See also fixed-effects model.

random-effects parameters. Parameters associated with a random-effects variable. Random-effects parameters are assumed to be conditionally independent across levels of the random-effects variable given all other model parameters. Often, random-effects parameters are assumed to be normally distributed with a zero mean and an unknown variance–covariance matrix.

random-effects variable. A variable identifying the group structure for the random effects at a specific level of hierarchy.

randomized controlled trial. In this experimental study, treatments are randomly assigned to two or more groups of subjects.

rank. Terms in common use are rank, row rank, and column rank. The row rank of a matrix $A$: $m \times n$ is the number of rows of $A$ that are linearly independent. The column rank is defined similarly, as the number of columns that are linearly independent. The terms row rank and column rank, however, are used merely for emphasis; the ranks are equal and the result is simply called the rank of $A$.

For a square matrix $A$ (where $m==n$), the matrix is invertible if and only if $\text{rank}(A)==n$. One often hears that $A$ is of full rank in this case and rank deficient in the other. See [M-5] $\text{rank}(\cdot)$.

rating scale model. The rating scale model (RSM) is an IRT model for ordinal responses. The categories within each item vary in their difficulty; however, the distances between adjacent difficulty parameters are constrained to be the same across the items. The categories across all items share the same discrimination parameter.

r-conformability. A set of two or more matrices, vectors, or scalars $A$, $B$, ..., are said to be r-conformable if each is c-conformable with a matrix of $\max(\text{rows}(A), \text{rows}(B), \ldots)$ rows and $\max(\text{cols}(A), \text{cols}(B), \ldots)$ columns.

r-conformability is a more relaxed form of c-conformability in that, if two matrices are c-conformable, they are r-conformable, but not vice versa. For instance, $A$: $1 \times 3$ and $B$: $3 \times 1$ are r-conformable but not c-conformable. Also, c-conformability is defined with respect to a pair of matrices only; r-conformability can be applied to a set of matrices.

r-conformability is often required of the arguments for functions that would otherwise naturally be expected to require scalars. See R-conformability in [M-5] $\text{normal}(\cdot)$ for an example. RCT. See randomized controlled trial.
real. A matrix is said to be a real matrix if its elements are all reals and it is stored in a real matrix.

Real is one of the two numeric types in Mata, the other being complex. Also see type, eltype, and orgtype.

recursive regression analysis. A recursive regression analysis involves performing a regression at time $t$ by using all available observations from some starting time $t_0$ through time $t$, performing another regression at time $t + 1$ by using all observations from time $t_0$ through time $t + 1$, and so on. Unlike a rolling regression analysis, the first period used for all regressions is held fixed.

reference prior. See noninformative prior.

reference value. See null value.

reflection. A reflection is an orientation reversing orthogonal transformation, that is, a transformation that involves negating coordinates in one or more dimensions. A reflection is a Procrustes transformation.

registered and unregistered variables. Variables in mi data can be registered as imputed, passive, or regular by using the mi register command; see [MI] mi set.

You are required to register imputed variables.

You should register passive variables; if your data are style wide, you are required to register them. The mi passive command (see [MI] mi passive) makes creating passive variables easy, and it automatically registers them for you.

Whether you register regular variables is up to you. Registering them is safer in all styles except wide, where it does not matter. By definition, regular variables should be the same across $m$. In the long styles, you can unintentionally create variables that vary. If the variable is registered, mi will detect and fix your mistakes.

Super-varying variables, which rarely occur and can be stored only in flong and flongsep data, should never be registered.

The registration status of variables is listed by the mi describe command; see [MI] mi describe.

regressand. The regressand is the variable that is being explained or predicted in a regression model. Synonyms include dependent variable, left-hand-side variable, and endogenous variable.

regression. A regression is a model in which an endogenous variable is written as a function of other variables, parameters to be estimated, and a random disturbance.

regression-adjustment estimators. Regression-adjustment estimators use means of predicted outcomes for each treatment level to estimate each potential-outcome mean.

regressor. Regressors are variables in a regression model used to predict the regressand. Synonyms include independent variable, right-hand-side variable, explanatory variable, predictor variable, and exogenous variable.

regular variable. See imputed, passive, and regular variables.

rejection region. In hypothesis testing, a rejection region is a set of sample values for which the null hypothesis can be rejected.

relative efficiency. Ratio of variance of a parameter given estimation with finite $M$ to the variance if $M$ were infinite.

relative risk. See risk ratio.

relative variance increase. The increase in variance of a parameter estimate due to nonresponse.

reliability. Reliability is the proportion of the variance of a variable not due to measurement error. A variable without measure error has reliability 1.
REML. See restricted maximum likelihood.

repeated denominator degrees of freedom (DDF) method. This method uses the repeated-measures ANOVA for computing DDF. It is used with balanced repeated-measures designs with spherical correlation error structures. It partitions the residual degrees of freedom into the between-subject degrees of freedom and the within-subject degrees of freedom. The repeated method is supported only with two-level models. For more complex mixed-effects models or with unbalanced data, this method typically leads to poor approximations of the actual sampling distributions of the test statistics.

repeated measures. Repeated measures data have repeated measurements for the subjects over some dimension, such as time—for example test scores at the start, midway, and end of the class. The repeated observations are typically not independent. Repeated-measures ANOVA is one approach for analyzing repeated measures data, and MANOVA is another. Also see sphericity.

replicate-weight variable. A replicate-weight variable contains sampling weight values that were adjusted for resampling the data; see [SVY] variance estimation for more details.

resampling. Resampling refers to the process of sampling from the dataset. In the delete-one jackknife, the dataset is resampled by dropping one PSU and producing a replicate of the point estimates. In the BRR method, the dataset is resampled by dropping combinations of one PSU from each stratum. The resulting replicates of the point estimates are used to estimate their variances and covariances.

residual. In this manual, we reserve the word “residual” for the difference between the observed and fitted moments of an SEM model. We use the word error for the disturbance associated with a (Gaussian) linear equation; see error. Also see standardized residuals.

residual denominator degrees of freedom (DDF) method. This method uses the residual degrees of freedom, \( n - \text{rank}(X) \), as the DDF for all tests of fixed effects. For a linear model without random effects with independent and identically distributed errors, the distributions of the test statistics for fixed effects are \( t \) or \( F \) distributions with the residual DDF. For other mixed-effects models, this method typically leads to poor approximations of the actual sampling distributions of the test statistics.

restricted maximum likelihood. Restricted maximum likelihood is a method of fitting linear mixed-effects models that involves transforming out the fixed effects to focus solely on variance–component estimation.

retrospective study. In a retrospective study, a group with a disease of interest is compared with a group without the disease, and information is gathered in a retrospective way about the exposure in each group to various risk factors that might be associated with the disease. Also see prospective study.

right-censoring. See censored, censoring, left-censoring, and right-censoring.

right-truncation. See truncation, left-truncation, and right-truncation.

risk difference. A risk difference is defined as the probability of an event occurring when a risk factor is increased by one unit minus the probability of the event occurring without the increase in the risk factor.

When the risk factor is binary, the risk difference is the probability of the outcome when the risk factor is present minus the probability when the risk factor is not present.

When one compares two populations, a risk difference is defined as a difference between the probabilities of an event in the two groups. It is typically a difference between the probability in the comparison group or experimental group and the probability in the reference group or control group.
risk factor. This is a variable associated with an increased or decreased risk of failure.

risk pool. At a particular point in time, this is the subjects at risk of failure.

risk ratio. In a log-linear model, this is the ratio of probability of survival associated with a one-unit increase in a risk factor relative to that calculated without such an increase, that is, \( R(x + 1)/R(x) \). Given the exponential form of the model, \( R(x + 1)/R(x) \) is constant and is given by the exponentiated coefficient.

robust, vce(robust). Robust is the name we use here for the Huber/White/sandwich estimator of the VCE. This technique requires fewer assumptions than most other techniques. In particular, it merely assumes that the errors are independently distributed across observations and thus allows the errors to be heteroskedastic. Robust standard errors are reported when the `vce(robust)` option is specified. The other available techniques are `OIM`, `EIM`, `OPG`, `clustered`, `bootstrap`, and `jackknife`.

robust standard errors. Robust standard errors, also known as Huber/White or Taylor linearization standard errors, are based on the sandwich estimator of variance. Robust standard errors can be interpreted as representing the sample-to-sample variability of the parameter estimates, even when the model is misspecified. See also `semirobust standard errors`.

rolling regression analysis. A rolling, or moving window, regression analysis involves performing regressions for each period by using the most recent \( m \) periods' data, where \( m \) is known as the window size. At time \( t \) the regression is fit using observations for times \( t - 19 \) through time \( t \); at time \( t + 1 \) the regression is fit using the observations for time \( t - 18 \) through \( t + 1 \); and so on.

rotation. A rotation is an orientation preserving orthogonal transformation. A rotation is a Procrustes transformation.

row and column stripes. Stripes refer to the labels associated with the rows and columns of a Stata matrix; see `Stata matrix`.

row-major order. Matrices are stored as vectors. Row-major order specifies that the vector form of a matrix is created by stacking the rows. For instance,

\[
\begin{bmatrix}
1 & 2 & 3 \\
1 & 2 & 3 \\
4 & 5 & 6 \\
\end{bmatrix}
\]

is stored as

\[
\begin{bmatrix}
1 & 2 & 3 & 4 & 5 & 6 \\
1 & 2 & 3 & 4 & 5 & 6 \\
\end{bmatrix}
\]

in row-major order. Mata uses row-major order. The LAPACK functions use column-major order. See `column-major order`.

rowvector. See `vector`, `colvector`, and `rowvector`.

Roy’s largest root. Roy’s largest root test is a test statistic for the hypothesis test \( H_0 : \mu_1 = \cdots = \mu_k \) based on the largest eigenvalue of \( E^{-1}H \). It is defined as

\[
\theta = \frac{\lambda_1}{1 + \lambda_1}
\]

Here \( H \) is the between matrix, and \( E \) is the within matrix. See `between matrix`. 
RSM. See rating scale model.

RVI. See relative variance increase.

Sammon mapping criterion. The Sammon (1969) mapping criterion is a loss criterion used with MDS; it is the sum of the scaled, squared differences between the distances and the disparities, normalized by the sum of the disparities. Also see multidimensional scaling, modern scaling, and loss.

Sample. A sample is the collection of individuals in the population that were chosen as part of the survey. Sample is also used to refer to the data, typically in the form of answered questions, collected from the sampled individuals.

Sample size. This is the number of subjects in a sample. See [PSS] intro to learn more about the relationship between sample size and the power of a test.

Sample-size curve. A sample-size curve is a graph of the estimated sample size as a function of some other study parameter such as power. The sample size is plotted on the y axis, and the power or other parameter is plotted on the x axis.

Sample-size determination. This pertains to the computation of a sample size given power, effect size, and other study parameters.

Sampling stage. Complex survey data are typically collected using multiple stages of clustered sampling. In the first stage, the PSUs are independently selected within each stratum. In the second stage, smaller sampling units are selected within the PSUs. In later stages, smaller and smaller sampling units are selected within the clusters from the previous stage.

Sampling unit. A sampling unit is an individual or collection of individuals from the population that can be selected in a specific stage of a given survey design. Examples of sampling units include city blocks, high schools, hospitals, and houses.

Sampling weight. Given a survey design, the sampling weight for an individual is the reciprocal of the probability of being sampled. The probability for being sampled is derived from stratification and clustering in the survey design. A sampling weight is typically considered to be the number of individuals in the population represented by the sampled individual.

Sampling with and without replacement. Sampling units may be chosen more than once in designs that use sampling with replacement. Sampling units may be chosen at most once in designs that use sampling without replacement. Variance estimates from with-replacement designs tend to be larger than those from corresponding without-replacement designs.

Satterthwaite denominator degrees of freedom (DDF) method. This method implements a generalization of the Satterthwaite (1946) approximation of the unknown sampling distributions of test statistics for complex linear mixed-effects models. This method is supported only with restricted maximum-likelihood estimation.

Satterthwaite’s t test. Satterthwaite’s t test is a modification of the two-sample t test to account for unequal variances in the two populations. See Methods and formulas in [PSS] power twomeans for details.

Saturated model. A saturated model is a full covariance model—a model of fitted means and covariances of observed variables without any restrictions on the values. Also see baseline model. Saturated models apply only to standard linear SEMs.

Scalar. A special case of a matrix with one row and one column. A scalar may be substituted anywhere a matrix, vector, column vector, or row vector is required, but not vice versa.
scalar model parameter. A scalar model parameter is any model parameter that is a scalar. For example, `{mean}` and `{hapt:alpha}` are scalar parameters, as declared by the `bayesmh` command. Elements of matrix model parameters are viewed as scalar model parameters. For example, for a 2 × 2 matrix parameter `{Sigma,matrix}`, individual elements `{Sigma_1_1}`, `{Sigma_2_1}`, `{Sigma_1_2}`, and `{Sigma_2_2}` are scalar parameters. If a matrix parameter contains a label, the label should be included in the specification of individual elements as well. See[BAYES] bayesmh.

scalar parameter. See scalar model parameter.

Schur decomposition. The Schur decomposition of a matrix, A, can be written as

\[ Q' A Q = T \]

where \( T \) is in Schur form and \( Q \), the matrix of Schur vectors, is orthogonal if \( A \) is real or unitary if \( A \) is complex. See[M-5] schurd().

Schur form. There are two Schur forms: real Schur form and complex Schur form.

A real matrix is in Schur form if it is block upper triangular with 1 × 1 and 2 × 2 diagonal blocks. Each 2 × 2 diagonal block has equal diagonal elements and opposite sign off-diagonal elements. The real eigenvalues are on the diagonal and complex eigenvalues can be obtained from the 2 × 24 diagonal blocks.

A complex square matrix is in Schur form if it is upper triangular with the eigenvalues on the diagonal.

cscore. A score for an observation after factor analysis, PCA, or LDA is derived from a column of the loading matrix and is obtained as the linear combination of that observation’s data by using the coefficients found in the loading.

cscore plot. A score plot produces scatterplots of the score variables after factor analysis, PCA, or LDA.

cscore test, Lagrange multiplier test. A score test is a test based on first derivatives of a likelihood function. Score tests are especially convenient for testing whether constraints on parameters should be relaxed or parameters should be added to a model. Also see Wald test.

cscores. Scores has two unrelated meanings. First, scores are the observation-by-observation first-derivatives of the (quasi) log-likelihood function. When we use the word “scores”, this is what we mean. Second, in the factor-analysis literature, scores (usually in the context of factor scores) refers to the expected value of a latent variable conditional on all the observed variables. We refer to this simply as the predicted value of the latent variable.

cscree plot. A scree plot is a plot of eigenvalues or singular values ordered from greatest to least after an eigen decomposition or singular value decomposition. Scree plots help determine the number of factors or components in an eigen analysis. Scree is the accumulation of loose stones or rocky debris lying on a slope or at the base of a hill or cliff; this plot is called a scree plot because it looks like a scree slope. The goal is to determine the point where the mountain gives way to the fallen rock.

SDR. See successive difference replication.

seasonal difference operator. The period-s seasonal difference operator \( \Delta_s \) denotes the difference in the value of a variable at time \( t \) and time \( t - s \). Formally, \( \Delta_s y_t = y_t - y_{t-s} \), and \( \Delta^2_s y_t = \Delta_s (y_t - y_{t-s}) = (y_t - y_{t-s}) - (y_{t-s} - y_{t-2s}) = y_t - 2y_{t-s} + y_{t-2s} \).

secondary sampling unit. Secondary sampling unit (SSU) is a cluster that was sampled from within a PSU in the second sampling stage. SSU is also used as a generic term unit to indicate any sampling unit that is not from the first sampling stage.

second-level latent variable. See first-, second-, and higher-order latent variables.
second-order latent variable. See first- and second-order latent variables.

seemingly unrelated regression. Seemingly unrelated regression is a kind of structural model in which each member of a set of observed endogenous variables is a function of a set of observed exogenous variables and a unique random disturbance term. The disturbances are correlated and the sets of exogenous variables may overlap. If the sets of exogenous variables are identical, this is referred to as multivariate regression.

selection-on-observables. See conditional-independence assumption.

SEM. SEM stands for structural equation modeling and for structural equation model. We use SEM in capital letters when writing about theoretical or conceptual issues as opposed to issues of the particular implementation of SEM in Stata with the sem or gsem commands.

dem. sem is the Stata command that fits standard linear SEMs. Also see gsem.

semiconjugate prior. A prior distribution is semiconjugate for a family of likelihood distributions if the prior and (full) conditional posterior distributions belong to the same family of distributions. For semiconjugacy to hold, parameters must typically be independent a priori; that is, their joint prior distribution must be the product of the individual marginal prior distributions. For example, the normal prior distribution for a mean parameter of a normal data distribution with an unknown variance (which is assumed to be independent of the mean a priori) is a semiconjugate prior. Semiconjugacy may provide an efficient way of sampling from posterior distributions and is used in Gibbs sampling.

semiparametric model. This is a model that is not fully parameterized. The Cox proportional hazards model is such a model:

\[ h(t) = h_0(t) \exp(\beta_1 x_1 + \cdots + \beta_k x_k) \]

In the Cox model, \( h_0(t) \) is left unparameterized and not even estimated. Meanwhile, the relative effects of covariates are parameterized as \( \exp(\beta_1 x_1 + \cdots + \beta_k x_k) \).

semirobust standard errors. Semirobust standard errors are closely related to robust standard errors and can be interpreted as representing the sample-to-sample variability of the parameter estimates, even when the model is misspecified, as long as the mean structure of the model is specified correctly. See also robust standard errors.

sensitivity analysis. Sensitivity analysis investigates the effect of varying study parameters on power, sample size, and other components of a study. The true values of study parameters are usually unknown, and power and sample-size analysis uses best guesses for these values. It is therefore important to evaluate the sensitivity of the computed power or sample size in response to changes in study parameters. See [PSS] power, table and [PSS] power, graph for details.

sequential limit theory. The sequential limit theory is a method of determining asymptotic properties of a panel-data statistic in which one index, say, \( N \), the number of panels, is held fixed, while \( T \), the number of time periods, goes to infinity, providing an intermediate limit. Then one obtains a final limit by studying the behavior of this intermediate limit as the other index (\( N \) here) goes to infinity.

serial correlation. Serial correlation refers to regression errors that are correlated over time. If a regression model does not contain lagged dependent variables as regressors, the OLS estimates are consistent in the presence of mild serial correlation, but the covariance matrix is incorrect. When the model includes lagged dependent variables and the residuals are serially correlated, the OLS estimates are biased and inconsistent. See, for example, Davidson and MacKinnon (1993, chap. 10) for more information.
serial correlation tests. Because OLS estimates are at least inefficient and potentially biased in the presence of serial correlation, econometricians have developed many tests to detect it. Popular ones include the Durbin–Watson (1950, 1951, 1971) test, the Breusch–Pagan (1980) test, and Durbin’s (1970) alternative test. See [R] regress postestimation time series.

shape parameter. A shape parameter governs the shape of a probability distribution. One example is the parameter \( p \) of the Weibull model.

Shepard diagram. A Shepard diagram after MDS is a 2-dimensional plot of high-dimensional dissimilarities or disparities versus the resulting low-dimensional distances. Also see multidimensional scaling.

sign test. A sign test is used to test the null hypothesis that the median of a distribution is equal to some reference value. A sign test is carried out as a test of binomial proportion with a reference value of 0.5. See [PSS] power oneproportion and [R] bitest.

significance level. In hypothesis testing, the significance level \( \alpha \) is an upper bound for a probability of a type I error. See [PSS] intro to learn more about the relationship between significance level and the power of a test.

similarity, similarity matrix, and similarity measure. A similarity or a similarity measure is a quantification of how alike two things are, such as observations or variables or groups of observations, or a method for quantifying that likeness. A similarity matrix is a matrix containing similarity measurements. The matching coefficient is one example of a similarity measure. Contrast to dissimilarity. Also see proximity and matching coefficient.

simple random sample. In a simple random sample (SRS), individuals are independently sampled—each with the same probability of being chosen.

single-linkage clustering. Single-linkage clustering is a hierarchical clustering method that computes the proximity between two groups as the proximity between the closest pair of observations between the two groups.

single-record st data. See st data.

singleton-group data. A singleton is a frailty group that contains only 1 observation. A dataset containing only singletons is known as singleton-group data.

singular value decomposition. A singular value decomposition (SVD) is a factorization of a rectangular matrix. It says that if \( M \) is an \( m \times n \) matrix, there exists a factorization of the form

\[
M = U \Sigma V^* \]

where \( U \) is an \( m \times m \) unitary matrix, \( \Sigma \) is an \( m \times n \) matrix with nonnegative numbers on the diagonal and zeros off the diagonal, and \( V^* \) is the conjugate transpose of \( V \), an \( n \times n \) unitary matrix. If \( M \) is a real matrix, then so is \( V \), and \( V^* = V' \).

size of test. See significance level.

slope. See discrimination.

smooth treatment-effects estimator. A smooth treatment-effects estimator is a smooth function of the data so that standard methods approximate the distribution of the estimator. The RA, IPW, AIPW, and IPWRA estimators are all smooth treatment-effects estimators while the nearest-neighbor matching estimator and the propensity-score matching estimator are not.

smoothing. Smoothing a time series refers to the process of extracting an overall trend in the data. The motivation behind smoothing is the belief that a time series exhibits a trend component as well as an irregular component and that the analyst is interested only in the trend component. Some smoother also account for seasonal or other cyclical patterns.
SMR. See *standardized mortality (morbidity) ratio*.

**snapshot data.** Snapshot data are those in which each record contains the values of a set of variables for a subject at an instant in time. The name arises because each observation is like a snapshot of the subject.

In snapshot datasets, one usually has a group of observations (snapshots) for each subject.

Snapshot data must be converted to st data before they can be analyzed. This requires making assumptions about what happened between the snapshots. See [ST] *snapspan*.

**source code.** Source code refers to the human-readable code that you type into Mata to define a function. Source code is compiled into object code, which is binary. See [M-1] *how*.

**spectral analysis.** See *frequency-domain analysis*.

**spectral density function.** The spectral density function is the derivative of the spectral distribution function. Intuitively, the spectral density function $f(\omega)$ indicates the amount of variance in a time series that is attributable to sinusoidal components with frequency $\omega$. See also *spectral distribution function*. The spectral density function is sometimes called the *spectrum*.

**spectral distribution function.** The (normalized) spectral distribution function $F(\omega)$ of a process describes the proportion of variance that can be explained by sinusoids with frequencies in the range $(0, \omega)$, where $0 \leq \omega \leq \pi$. The spectral distribution and density functions used in frequency-domain analysis are closely related to the autocorrelation function used in time-domain analysis; see Chatfield (2004, chap. 6) and Wei (2006, chap. 12).

**spectrum.** See *spectral density function*.

**spell data.** Spell data are survival data in which each record represents a fixed period, consisting of a begin time, an end time, possibly a censoring/failure indicator, and other measurements (covariates) taken during that specific period.

**sphericity.** Sphericity is the state or condition of being a sphere. In repeated measures ANOVA, sphericity concerns the equality of variance in the difference between successive levels of the repeated measure. The multivariate alternative to ANOVA, called MANOVA, does not require the assumption of sphericity. Also see *repeated measures*.

**square matrix.** A matrix is square if it has the same number of rows and columns. A $3 \times 3$ matrix is square; a $3 \times 4$ matrix is not.

**SRS.** See *simple random sample*.

**SSCP matrix.** SSCP is an acronym for the sums of squares and cross products. Also see *between matrix*.

**SSD, ssd.** See *summary statistics data*.

**SSU.** See *secondary sampling unit*.

**st data.** st stands for survival time. In survival-time data, each observation represents a span of survival, recorded in variables $t_0$ and $t$. For instance, if in an observation $t_0$ were 3 and $t$ were 5, the span would be $(t_0, t]$, meaning from just after $t_0$ up to and including $t$.

Sometimes variable $t_0$ is not recorded; $t_0$ is then assumed to be 0. In such a dataset, an observation that had $t = 5$ would record the span $(0, 5]$.

Each observation also includes a variable $d$, called the failure variable, which contains 0 or nonzero (typically, 1). The failure variable records what happened at the end of the span: 0, the subject was still alive (had not yet failed) or 1, the subject died (failed).
Sometimes variable \( d \) is not recorded; \( d \) is then assumed to be 1. In such a dataset, all time-span observations would be assumed to end in failure.

Finally, each observation in an \( st \) dataset can record the entire history of a subject or each can record a part of the history. In the latter case, groups of observations record the full history. One observation might record the period \((0,5]\) and the next, \((5,8]\). In such cases, there is a variable \( ID \) that records the subject for which the observation records a time span. Such data are called multiple-record \( st \) data. When each observation records the entire history of a subject, the data are called single-record \( st \) data. In the single-record case, the \( ID \) variable is optional.

See [ST] \texttt{stset}.

**stacked variables.** See crossed variables.

**stacking variables.** See crossing variables.

**standard linear SEM.** An SEM without multilevel effects in which all response variables are given by a linear equation. Standard linear SEM is what most people mean when they refer to just SEM. Standard linear SEMs are fit by \texttt{sem}, although they can also be fit by \texttt{gsem}; see generalized SEM.

**standard strata.** See direct standardization.

**standard weights.** See direct standardization.

**standardized coefficient.** In a linear equation \( y = \ldots bx + \ldots \), the standardized coefficient \( \beta \) is \((\frac{\hat{\sigma}_y}{\hat{\sigma}_x})b\). Standardized coefficients are scaled to units of standard deviation change in \( y \) for a standard deviation change in \( x \).

**standardized covariance.** A standardized covariance between \( y \) and \( x \) is equal to the correlation of \( y \) and \( x \), which is to say, it is equal to \( \frac{\sigma_{xy}}{\sigma_x\sigma_y} \). The covariance is equal to the correlation when variables are standardized to have variance 1.

**standardized data.** Standardized data has a mean of zero and a standard deviation of one. You can standardize data \( x \) by taking \((x - \bar{x})/\sigma\), where \( \sigma \) is the standard deviation of the data.

**standardized mortality (morbidity) ratio.** Standardized mortality (morbidity) ratio (SMR) is the observed number of deaths divided by the expected number of deaths. It is calculated using indirect standardization: you take the population of the group of interest—say, by age, sex, and other factors—and calculate the expected number of deaths in each cell (expected being defined as the number of deaths that would have been observed if those in the cell had the same mortality as some other population). You then take the ratio to compare the observed with the expected number of deaths. For instance,

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{Age} & \text{(1) Population of group} & \text{(2) Deaths per 100,000 in general pop.} & \text{(1) \times (2) Expected # of deaths} & \text{Observed deaths} \\
\hline
25–34 & 95,965 & 105.2 & 100.9 & 92 \\
34–44 & 78,280 & 203.6 & 159.4 & 180 \\
44–54 & 52,393 & 428.9 & 224.7 & 242 \\
55–64 & 28,914 & 964.6 & 278.9 & 312 \\
\hline
\text{Total} & & 763.9 & 826 & \\
\hline
\end{array}
\]

\[
\text{SMR} = \frac{826}{763.9} = 1.08
\]

**standardized residuals, normalized residuals.** Standardized residuals are residuals adjusted so that they follow a standard normal distribution. The difficulty is that the adjustment is not always possible. Normalized residuals are residuals adjusted according to a different formula that roughly follow a standard normal distribution. Normalized residuals can always be calculated.
starting values. The estimation methods provided by `sem` and `gsem` are iterative. The starting values are values for each of the parameters to be estimated that are used to initialize the estimation process. The `sem` software provides starting values automatically, but in some cases, these are not good enough and you must (1) diagnose the problem and (2) provide better starting values. See [SEM] intro 12.

Stata matrix. Stata itself, separate from Mata, has matrix capabilities. Stata matrices are separate from those of Mata, although Stata matrices can be gotten from and put into Mata matrices; see [M-5] st_matrix(). Stata matrices are described in [P] matrix and [U] 14 Matrix expressions.

Stata matrices are exclusively numeric and contain real elements only. Stata matrices also differ from Mata matrices in that, in addition to the matrix itself, a Stata matrix has text labels on the rows and columns. These labels are called row stripes and column stripes. One can think of rows and columns as having names. The purpose of these names is discussed in [U] 14.2 Row and column names. Mata matrices have no such labels. Thus three steps are required to get or to put all the information recorded in a Stata matrix: 1) getting or putting the matrix itself; 2) getting or putting the row stripe from or into a string matrix; and 3) getting or putting the column stripe from or into a string matrix. These steps are discussed in [M-5] st_matrix().

state-space model. A state-space model describes the relationship between an observed time series and an unobservable state vector that represents the “state” of the world. The measurement equation expresses the observed series as a function of the state vector, and the transition equation describes how the unobserved state vector evolves over time. By defining the parameters of the measurement and transition equations appropriately, one can write a wide variety of time-series models in the state-space form.

static forecast. A static forecast uses actual values wherever lagged values of the endogenous variables appear in the model. As a result, static forecasts perform at least as well as dynamic forecasts, but static forecasts cannot produce forecasts into the future if lags of the endogenous variables appear in the model.

Because actual values will be missing beyond the last historical time period in the dataset, static forecasts can only forecast one period into the future (assuming only first lags appear in the model); for that reason, they are often called one-step-ahead forecasts.

stationary distribution. Stationary distribution of a stochastic process is a joint distribution that does not change over time. In the context of MCMC, stationary distribution is the target probability distribution to which the Markov chain converges. When MCMC is used for simulating a Bayesian model, the stationary distribution is the target joint posterior distribution of model parameters.

steady-state equilibrium. The steady-state equilibrium is the predicted value of a variable in a dynamic model, ignoring the effects of past shocks, or, equivalently, the value of a variable, assuming that the effects of past shocks have fully died out and no longer affect the variable of interest.

stochastic equation. A stochastic equation, in contrast to an identity, is an equation in a forecast model that includes a random component, most often in the form of an additive error term. Stochastic equations include parameters that must be estimated from historical data.

stochastic trend. A stochastic trend is a nonstationary random process. Unit-root process and random coefficients on time are two common stochastic trends. See [TS] ucm for examples and discussions of more commonly applied stochastic trends.

stopping rules. Stopping rules for hierarchical cluster analysis are used to determine the number of clusters. A stopping-rule value (also called an index) is computed for each cluster solution, that is, at each level of the hierarchy in hierarchical cluster analysis. Also see hierarchical clustering.

str1, str2, ..., str2045. See strL.
**stratification.** The population is partitioned into well-defined groups of individuals, called strata. In the first sampling stage, PSUs are independently sampled from within each stratum. In later sampling stages, SSUs are independently sampled from within each stratum for that stage.

Survey designs that use stratification typically result in smaller variance estimates than do similar designs that do not use stratification. Stratification is most effective in decreasing variability when sampling units are more similar within the strata than between them.

**stratified $2 \times 2$ tables.** Describe the association between a binary independent variable and a binary response variable of interest. The analysis is stratified by a nominal (categorical) variable with $K$ levels.

**stratified model.** A stratified survival model constrains regression coefficients to be equal across levels of the stratification variable, while allowing other features of the model to vary across strata.

**stratified test.** A stratified test is performed separately for each stratum. The stratum-specific results are then combined into an overall test statistic.

**stress.** See *Kruskal stress* and *loss.*

**strict stationarity.** A process is strictly stationary if the joint distribution of $y_1, \ldots, y_k$ is the same as the joint distribution of $y_{1+\tau}, \ldots, y_{k+\tau}$ for all $k$ and $\tau$. Intuitively, shifting the origin of the series by $\tau$ units has no effect on the joint distributions.

**string.** A matrix is said to be a string matrix if its elements are strings (text); see *type, eltype, and orgtype.* In Mata, a string may be text or binary and may be up to 2,147,483,647 characters (bytes) long.

**strL.** `strL` is a storage type for string variables. The full list of string storage types is `str1`, `str2`, ..., `str2045`, and `strL`.

`str1`, `str2`, ..., `str2045` are fixed-length storage types. If variable `mystr` is `str8`, then 8 bytes are allocated in each observation to store `mystr`’s value. If you have 2,000 observations, then 16,000 bytes in total are allocated.

Distinguish between storage length and string length. If `myvar` is `str8`, that does not mean the strings are 8 characters long in every observation. The maximum length of strings is 8 characters. Individual observations may have strings of length 0, 1, ..., 8. Even so, every string requires 8 bytes of storage.

You need not concern yourself with the storage length because string variables are automatically promoted. If `myvar` is `str8`, and you changed the contents of `myvar` in the third observation to “Longer than 8”, then `myvar` would automatically become `str13`.

If you changed the contents of `myvar` in the third observation to a string longer than 2,045 characters, `myvar` would become `strL`.

`strL` variables are not necessarily longer than 2,045 characters; they can be longer or shorter than 2,045 characters. The real difference is that `strL` variables are stored as varying length. Pretend that `myothervar` is a `strL` and its third observation contains “this”. The total memory consumed by the observation would be $64 + 4 + 1 = 69$ bytes. There would be 64 bytes of tracking information, 4 bytes for the contents (there are 4 characters), and 1 more byte to terminate the string. If the fifth observation contained a 2,000,000-character string, then $64 + 2,000,000 + 1 = 2,000,069$ bytes would be used to store it.

Another difference between `str1`, `str2`, ..., `str2045`, and `strLs` is that the `str#` storage types can store only ASCII strings. `strL` can store ASCII or binary strings. Thus a `strL` variable could contain, for instance, the contents of a Word document or a JPEG image or anything else.

`strL` is pronounce `sturl.`
strongly balanced. A longitudinal or panel dataset is said to be strongly balanced if each panel has the same number of observations, and the observations for different panels were all made at the same times.

structural equation model. Different authors use the term “structural equation model” in different ways, but all would agree that an SEM sometimes carries the connotation of being a structural model with a measurement component, that is, combined with a measurement model.

structural model. A structural model is one that describes the relationship among a set of variables, based on underlying theoretical considerations. In particular, the parameters of a structural model are posited to quantify an actual causal relationship among the variables rather than a mere description of the variables’ correlations.

Structural models often have multiple equations and dependencies between endogenous variables, although that is not a requirement.

Structural models can be viewed in a structural equation modeling (SEM) framework and can thus be fitted by sem and gsem, though these commands are not limited to fitting just structural models. See [SEM] intro 5 and structural equation model.

Structural models are also used in econometric forecasting applications. See [TS] forecast for information about forecasting from structural models based on time-series data.

structure (programming version). A structure is an eltype, indicating a set of variables tied together under one name. struct mystruct might be

```plaintext
struct mystruct {
    real scalar n1, n2
    real matrix X
}
```

If variable a was declared a struct mystruct scalar, then the scalar a would contain three pieces: two real scalars and one real matrix. The pieces would be referred to as a.n1, a.n2, and a.X. If variable b were also declared a struct mystruct scalar, it too would contain three pieces, b.n1, b.n2, and b.X. The advantage of structures is that they can be referred to as a whole. You can code a.n1=b.n1 to copy one piece, or you can code a=b if you wanted to copy all three pieces. In all ways, a and b are variables. You may pass a to a subroutine, for instance, which amounts to passing all three values.

Structures variables are usually scalar, but they are not limited to being so. If A were a struct mystruct matrix, then each element of A would contain three pieces, and one could refer, for instance, to A[2,3].n1, A[2,3].n2, and A[2,3].X, and even to A[2,3].X[3,2].

See [M-2] struct.

structure (statistics version). Structure, as in factor structure, is the correlations between the variables and the common factors after factor analysis. Structure matrices are available after factor analysis and LDA. Also see factor analysis and linear discriminant analysis.

structured (correlation or covariance). See unstructured and structured (correlation or covariance).

style. Style refers to the format in which the mi data are stored. There are four styles: flongs, flong, mlong, and wide. You can ignore styles, except for making an original selection, because all mi commands work regardless of style. You will be able to work more efficiently, however, if you understand the details of the style you are using; see [MI] styles. Some tasks are easier in one style than another. You can switch between styles by using the mi convert command; see [MI] mi convert.
The flongsep style is best avoided unless your data are too big to fit into one of the other styles. In flongsep style, a separate .dta set is created for \( m = 0 \), for \( m = 1 \), \ldots, and for \( m = M \). Flongsep is best avoided because \( \text{mi} \) commands work more slowly with it.

In all the other styles, the \( M + 1 \) datasets are stored in one .dta file. The other styles are both more convenient and more efficient.

The most easily described of these .dta styles is flong; however, flong is also best avoided because mlong style is every bit as convenient as flong, and mlong is memorywise more efficient. In flong, each observation in the original data is repeated \( M \) times in the .dta dataset, once for \( m = 1 \), again for \( m = 2 \), and so on. Variable \( _\text{mi}_m \) records \( m \) and takes on values 0, 1, 2, \ldots, \( M \). Within each value of \( m \), variable \( _\text{mi}_\text{id} \) takes on values 1, 2, \ldots, \( N \) and thus connects imputed with original observations.

The mlong style is recommended. It is efficient and easy to use. Mlong is much like flong except that complete observations are not repeated.

Equally recommended is the wide style. In wide, each imputed and passive variable has an additional \( M \) variables associated with it, one for the variable’s value in \( m = 1 \), another for its value in \( m = 2 \), and so on. If an imputed or passive variable is named \( vn \), then the values of \( vn \) in \( m = 1 \) are stored in variable \( _1\_vn \); the values for \( m = 2 \), in \( _2\_vn \); and so on.

What makes mlong and wide so convenient? In mlong, there is a one-to-one correspondence of your idea of a variable and Stata’s idea of a variable—variable \( vn \) refers to \( vn \) for all values of \( m \). In wide, there is a one-to-one correspondence of your idea of an observation and Stata’s idea—physical observation 5 is observation 5 in all datasets.

Choose the style that matches the problem at hand. If you want to create new variables or modify existing ones, choose mlong. If you want to drop observations or create new ones, choose wide. You can switch styles with the \texttt{mi convert} command; see [MI] \texttt{mi convert}.

For instance, if you want to create new variable \( \text{ageXexp} \) equal to \( \text{age} \times \text{exp} \) and your data are mlong, you can just type \texttt{generate ageXexp = age*exp}, and that will work even if \texttt{age} and \texttt{exp} are imputed, passive, or a mix. Theoretically, the right way to do that is to type \texttt{mi passive}: \texttt{generate agexExp = age*exp}, but concerning variables, if your data are mlong, you can work the usual Stata way.

If you want to drop observation 20 or drop if \texttt{sex==2}, if your data are wide, you can just type \texttt{drop in 20} or \texttt{drop if sex==2}. Here the “right” way to do the problem is to type the \texttt{drop} command and then remember to type \texttt{mi update} so that \( \text{mi} \) can perform whatever machinations are required to carry out the change throughout \( m > 0 \); however, in the wide form, there are no machinations required.

\textbf{subhazard, cumulative subhazard, and subhazard ratio}. In a competing-risks analysis, the hazard of the subdistribution (or subhazard for short) for the event of interest (type 1) is defined formally as

\[
\bar{h}_1(t) = \lim_{\delta \to 0} \left\{ \frac{P(t < T \leq t + \delta \text{ and event type 1}) \mid T > t \text{ or } (T \leq t \text{ and not event type 1})}{\delta} \right\}
\]

Less formally, think of this hazard as that which generates failure events of interest while keeping subjects who experience competing events “at risk” so that they can be adequately counted as not having any chance of failing.

The cumulative subhazard \( \bar{H}_1(t) \) is the integral of the subhazard function \( \bar{h}_1(t) \), from 0 (the onset of risk) to \( t \). The cumulative subhazard plays a very important role in competing-risks analysis. The cumulative incidence function (CIF) is a direct function of the cumulative subhazard:
CIF\(_1(t) = 1 - \exp\{-\bar{H}_1(t)\}\)

The subhazard ratio is the ratio of the subhazard function evaluated at two different values of the covariates: \(\frac{\bar{h}_1(t|x)}{\bar{h}_1(t|x_0)}\). The subhazard ratio is often called the relative subhazard, especially when \(\bar{h}_1(t|x_0)\) is the baseline subhazard function.

**subjective prior.** See **informative prior.**

**subpopulation estimation.** Subpopulation estimation focuses on computing point and variance estimates for part of the population. The variance estimates measure the sample-to-sample variability, assuming that the same survey design is used to select individuals for observation from the population. This approach results in a different variance than measuring the sample-to-sample variability by restricting the samples to individuals within the subpopulation; see [SVY] subpopulation estimation.

**subsampling the chain.** See **thinning.**

**subscripts.** Subscripts are how you refer to an element or even a submatrix of a matrix.

Mata provides two kinds of subscripts, known as list subscripts and range subscripts.

In list subscripts, \(A[2,3]\) refers to the (2,3) element of \(A\). \(A[(2\,3), (4,6)]\) refers to the submatrix made up of the second and third rows, fourth and sixth columns, of \(A\).

In range subscripts, \(A[2,3]\) also refers to the (2,3) element of \(A\). \(A[2,3\,4,6]\) refers to the submatrix beginning at the (2,3) element and ending at the (4,6) element.

See [M-2] subscripts for more information.

**substantive constraints.** See **identification.**

**successive difference replication.** Successive difference replication (SDR) is a method of variance typically applied to systematic samples, where the observed sampling units are somehow ordered. The SDR variance estimator is described in [SVY] variance estimation.

**summary statistics data.** Data are sometimes available only in summary statistics form, as (1) means and covariances, (2) means, standard deviations or variances, and correlations, (3) covariances, (4) standard deviations or variances and correlations, or (5) correlations. SEM can be used to fit models using such data in place of the underlying raw data. The ssd command creates datasets containing summary statistics.

**super-varying variables.** See **varying and super-varying variables.**

**supplementary rows or columns or supplementary variables.** Supplementary rows or columns can be included in CA, and supplementary variables can be included in MCA. They do not affect the CA or MCA solution, but they are included in plots and tables with statistics of the corresponding row or column points. Also see correspondence analysis and multiple correspondence analysis.

**survey data.** Survey data consist of information about individuals that were sampled from a population according to a survey design. Survey data distinguishes itself from other forms of data by the complex nature under which individuals are selected from the population.

In survey data analysis, the sample is used to draw inferences about the population. Furthermore, the variance estimates measure the sample-to-sample variability that results from the survey design applied to the fixed population. This approach differs from standard statistical analysis, in which the sample is used to draw inferences about a physical process and the variance measures the sample-to-sample variability that results from independently collecting the same number of observations from the same process.
**survey design.** A survey design describes how to sample individuals from the population. Survey designs typically include stratification and cluster sampling at one or more stages.

**survival-time data.** See *st data.*

**survivor function.** Also known as the survivorship function and the survival function, the survivor function, \( S(t) \), is 1) the probability of surviving beyond time \( t \), or equivalently, 2) the probability that there is no failure event prior to \( t \), 3) the proportion of the population surviving to time \( t \), or equivalently, 4) the reverse cumulative distribution function of \( T \), the time to the failure event: \( S(t) = \Pr(T > t) \). Also see hazard.

**SVAR.** A structural vector autoregressive (SVAR) model is a type of VAR in which short- or long-run constraints are placed on the resulting impulse–response functions. The constraints are usually motivated by economic theory and therefore allow causal interpretations of the IRFs to be made.

**SVD.** See *singular value decomposition.*

**symmetric matrices.** Matrix \( A \) is symmetric if \( A = A' \). The word *symmetric* is usually reserved for real matrices, and in that case, a symmetric matrix is a square matrix with \( a_{ij} = a_{ji} \).

Matrix \( A \) is said to be Hermitian if \( A = A' \), where the transpose operator is understood to mean the conjugate-transpose operator; see *Hermitian matrix.* In Mata, the ‘ operator is the conjugate-transpose operator, and thus, in this manual, we will use the word *symmetric* both to refer to real, symmetric matrices and to refer to complex, Hermitian matrices.

Sometimes, you will see us follow the word *symmetric* with a parenthesized Hermitian, as in, “the resulting matrix is symmetric (Hermitian)”. That is done only for emphasis.

The inverse of a symmetric (Hermitian) matrix is symmetric (Hermitian).

**symmetriconly.** Symmetriconly is a word we have coined to refer to a square matrix whose corresponding off-diagonal elements are equal to each other, whether the matrix is real or complex. Symmetriconly matrices have no mathematical significance, but sometimes, in data-processing and memory-management routines, it is useful to be able to distinguish such matrices.

**symmetry.** In a \( 2 \times 2 \) contingency table, symmetry refers to the equality of the off-diagonal elements. For a \( 2 \times 2 \) table, a test of *marginal homogeneity* reduces to a test of symmetry.

**t test.** A *t* test is a test for which the sampling distribution of the test statistic is a Student’s *t* distribution.

A one-sample *t* test is used to test whether the mean of a population is equal to a specified value when the variance must also be estimated. The test statistic follows Student’s *t* distribution with \( N - 1 \) degrees of freedom, where \( N \) is the sample size.

A two-sample *t* test is used to test whether the means of two populations are equal when the variances of the populations must also be estimated. When the two populations’ variances are unequal, a modification to the standard two-sample *t* test is used; see Satterthwaite’s *t* test.

**target parameter.** In power and sample-size analysis, the target parameter is the parameter of interest or the parameter in the study about which hypothesis tests are conducted.

**target rotation.** Target rotation minimizes the criterion

\[
c(\Lambda) = \frac{1}{2} ||\Lambda - H||^2
\]

for a given target matrix \( H \).

See Crawford–Ferguson rotation for a definition of \( \Lambda \).
**taxonomy.** Taxonomy is the study of the general principles of scientific classification. It also denotes classification, especially the classification of plants and animals according to their natural relationships. Cluster analysis is a tool used in creating a taxonomy and is synonymous with numerical taxonomy. Also see *cluster analysis.*

**Taylor linearization.** See *linearization.*

**TCC.** See *test characteristic curve.*

**technique.** Technique is just an English word and should be read in context. Nonetheless, technique is usually used here to refer to the technique used to calculate the estimated VCE. Those techniques are OIM, EIM, OPG, robust, clustered, bootstrap, and jackknife.

Technique is also used to refer to the available techniques used with `ml`, Stata’s optimizer and likelihood maximizer, to find the solution.

**test characteristic curve.** A test characteristic curve (TCC) is the sum of item characteristic curves and represents the expected score on the instrument.

**test information function** A test information function (TIF) is the sum of item information functions and indicates the precision of the entire instrument along the latent trait continuum.

**test statistic.** In hypothesis testing, a test statistic is a function of the sample that does not depend on any unknown parameters.

**tetrachoric correlation.** A tetrachoric correlation estimates the correlation coefficients of binary variables by assuming a latent bivariate normal distribution for each pair of variables, with a threshold model for manifest variables.

**thinning.** Thinning is a way of reducing autocorrelation in the MCMC sample by subsampling the MCMC chain every prespecified number of iterations determined by the thinning interval. For example, the thinning interval of 1 corresponds to using the entire MCMC sample; the thinning interval of 2 corresponds to using every other sample value; and the thinning interval of 3 corresponds to using values from iterations 1, 4, 7, 10, and so on. Thinning should be applied with caution when used to reduce autocorrelation because it may not always be the most appropriate way of improving the precision of estimates.

**thrashing.** Subjects are said to thrash when they are censored and immediately reenter with different covariates.

**three-level model.** A three-level mixed-effects model has one level of observations and two levels of grouping. Suppose that you have a dataset consisting of patients overseen by doctors at hospitals, and each doctor practices at one hospital. Then a three-level model would contain a set of random effects to control for hospital-specific variation, a second set of random effects to control for doctor-specific random variation within a hospital, and a random-error term to control for patients’ random variation.

**three-parameter logistic model.** The three-parameter logistic (3PL) model is an IRT model for binary responses where items vary in their difficulty and discrimination and can share or have their own guessing parameter.

**ties.** After discriminant analysis, ties in classification occur when two or more posterior probabilities are equal for an observation. They are most common with KNN discriminant analysis.

**TIF.** See *test information function.*

**time-domain analysis.** Time-domain analysis is analysis of data viewed as a sequence of observations observed over time. The autocorrelation function, linear regression, ARCH models, and ARIMA models are common tools used in time-domain analysis.
time-series–operated variable. Time-series–operated variables are a Stata concept. The term refers to \texttt{op.varname} combinations such as \texttt{L.gnp} to mean the lagged value of variable \texttt{gnp}. Mata’s \texttt{M-5 st\_data()} function works with time-series–operated variables just as it works with other variables, but many other Stata-interface functions do not allow \texttt{op.varname} combinations. In those cases, you must use \texttt{M-5 st\_tsrevar()}.

time-varying covariates. Time-varying covariates appear in a survival model whose values vary over time. The values of the covariates vary, not the effect. For instance, in a proportional hazards model, the log hazard at time \( t \) might be \( b \times \text{age}_t + c \times \text{treatment}_t \). Variable \text{age} might be time varying, meaning that as the subject ages, the value of \text{age} changes, which correspondingly causes the hazard to change. The effect \( b \), however, remains constant.

Time-varying variables are either continuously varying or discretely varying.

In the continuously varying case, the value of the variable \( x \) at time \( t \) is \( x_t = x_o + f(t) \), where \( f() \) is some function and often is the identity function, so that \( x_t = x_o + t \).

In the discretely varying case, the value of \( x \) changes at certain times and often in no particular pattern:

<table>
<thead>
<tr>
<th>idvar</th>
<th>t0</th>
<th>t</th>
<th>hp</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>5</td>
<td>150</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>7</td>
<td>130</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>9</td>
<td>135</td>
</tr>
</tbody>
</table>

In the above data, the value of \( hp \) is 150 over the period \((0, 5]\), then 130 over \((5, 7]\), and 135 over \((7, 9]\).

titlecase, title-cased string, and Unicode title-cased string. In grammar, titlecase refers to the capitalization of the key words in a phrase. In Stata, titlecase refers to (a) the capitalization of the first letter of each word in a string and (b) the capitalization of each letter after a nonletter character. There is no judgment of the word’s importance in the string or whether the letter after a nonletter character is part of the same word. For example, “it’s” in titlecase is “It’S”.

A title-cased string is any string to which the above rules have been applied. For example, if we used the \texttt{strproper()} function with the book title \textit{Zen and the Art of Motorcycle Maintenance}, Stata would return the title-cased string \textit{Zen And The Art Of Motorcycle Maintenance}.

A Unicode title-cased string is a string that has had Unicode title-casing rules applied to Unicode words. This is almost, but not exactly, like capitalizing the first letter of each Unicode word. Like capitalization, title-casing letters is locale-dependent, which means that the same letter might have different titlecase forms in different locales. For example, in some locales, capital letters at the beginning of words are not supposed to have accents on them, even if that capital letter by itself would have an accent.

If you do not have characters beyond plain ASCII and your locale is English, there is no distinction in results. For example, \texttt{ustrtitle()} with an English \texttt{locale} locale also would return the title-cased string \textit{Zen And The Art Of Motorcyle Maintenance}.

Use the \texttt{ustrtitle()} function to apply the appropriate capitalization rules for your language (locale).

total characteristic curve. See test characteristic curve.

total effects. See direct, indirect, and total effects.

total inertia or total principal inertia. The total (principal) inertia in CA and MCA is the sum of the principal inertias. In CA, total inertia is the Pearson \( \chi^2/n \). In CA, the principal inertias are
the singular values; in MCA the principal inertias are the eigenvalues. Also see correspondence analysis and multiple correspondence analysis.

total information function. See test information function.

traceback log. When a function fails—either because of a programming error or because it was used incorrectly—it produces a traceback log:

```plaintext
: myfunction(2,3)
solve(): 3200 conformability error
mysub(): - function returned error
myfunction(): - function returned error
<istmt>: - function returned error
r(3200);
```

The log says that `solve()` detected the problem—arguments are not conformable—and that `solve()` was called by `mysub()` was called by `myfunction()` was called by what you typed at the keyboard. See [M-2] errors for more information.

transmorphic. Transmorphic is an `eltype`. A scalar, vector, or matrix can be transmorphic, which indicates that its elements may be real, complex, string, pointer, or even a structure. The elements are all the same type; you are just not saying which they are. Variables that are not declared are assumed to be transmorphic, or a variable can be explicitly declared to be transmorphic. Transmorphic is just fancy jargon for saying that the elements of the scalar, vector, or matrix can be anything and that, from one instant to the next, the scalar, vector, or matrix might change from holding elements of one type to elements of another.

See [M-2] declarations.

transpose. The transpose operator is written different ways in different books, including ′, superscript *, superscript T, and superscript H. Here we use the ′ notation: \( A' \) means the transpose of \( A \), \( A \) with its rows and columns interchanged.

In complex analysis, the transpose operator, however it is written, is usually defined to mean the conjugate transpose; that is, one interchanges the rows and columns of the matrix and then one takes the conjugate of each element, or one does it in the opposite order—it makes no difference. Conjugation simply means reversing the sign of the imaginary part of a complex number: the conjugate of \( 1+2i \) is \( 1-2i \). The conjugate of a real is the number itself; the conjugate of 2 is 2.

In Mata, ′ is defined to mean conjugate transpose. Since the conjugate of a real is the number itself, \( A' \) is regular transposition when \( A \) is real. Similarly, we have defined ′ so that it performs regular transposition for string and pointer matrices. For complex matrices, however, ′ also performs conjugation.

If you have a complex matrix and simply want to transpose it without taking the conjugate of its elements, see [M-5] transposeonly(). Or code \( \text{conj}(A') \). The extra \( \text{conj()} \) will undo the undesired conjugation performed by the transpose operator.

Usually, however, you want transposition and conjugation to go hand in hand. Most mathematical formulas, generalized to complex values, work that way.

treatment model. A treatment model is a model used to predict treatment-assignment probabilities as a function of covariates and parameters.

trend. The trend specifies the long-run behavior in a time series. The trend can be deterministic or stochastic. Many economic, biological, health, and social time series have long-run tendencies to increase or decrease. Before the 1980s, most time-series analysis specified the long-run tendencies as deterministic functions of time. Since the 1980s, the stochastic trends implied by unit-root processes have become a standard part of the toolkit.
**triangular matrix.** A triangular matrix is a matrix with all elements equal to zero above the diagonal or all elements equal to zero below the diagonal.

A matrix $A$ is *lower triangular* if all elements are zero above the diagonal, that is, if $A[i, j] = 0$, $j > i$.

A matrix $A$ is *upper triangular* if all elements are zero below the diagonal, that is, if $A[i, j] = 0$, $j < i$.

A *diagonal matrix* is both lower and upper triangular. That is worth mentioning because any function suitable for use with triangular matrices is suitable for use with diagonal matrices.

A triangular matrix is usually *square*.

The inverse of a triangular matrix is a triangular matrix. The determinant of a triangular matrix is the product of the diagonal elements. The eigenvalues of a triangular matrix are the diagonal elements.

**truncation, left-truncation, and right-truncation.** In survival analysis, truncation occurs when subjects are observed only if their failure times fall within a certain observational period of a study. Censoring, on the other hand, occurs when subjects are observed for the whole duration of a study, but the exact times of their failures are not known; it is known only that their failures occurred within a certain time span.

Left-truncation occurs when subjects come under observation only if their failure times exceed some time $t_l$. It is only because they did not fail before $t_l$ that we even knew about their existence. Left-truncation differs from left-censoring in that, in the censored case, we know that the subject failed before time $t_l$, but we just do not know exactly when.

Imagine a study of patient survival after surgery, where patients cannot enter the sample until they have had a post-surgical test. The patients’ survival times will be left-truncated. This is a “delayed entry” problem, one common type of left-truncation.

Right-truncation occurs when subjects come under observation only if their failure times do not exceed some time $t_r$. Right-truncated data typically occur in registries. For example, a cancer registry includes only subjects who developed a cancer by a certain time, and thus survival data from this registry will be right-truncated.

**two-independent-samples test.** See *two-sample test*.

**two-level model.** A two-level mixed-effects model has one level of observations and one level of grouping. Suppose that you have a panel dataset consisting of patients at hospitals; a two-level model would contain a set of random effects at the hospital level (the second level) to control for hospital-specific random variation and a random-error term at the observation level (the first level) to control for within-hospital variation.

**two-parameter logistic model.** The two-parameter logistic (2PL) model is an IRT model for binary responses where items vary in their difficulty and discrimination.

**two-sample paired test.** See *paired test*.

**two-sample test.** A two-sample test is used to test whether the parameters of interest of the two independent populations are equal. For example, two-sample means test, two-sample variances, two-sample proportions test, two-sample correlations test.

**two-sided test, two-tailed test.** A two-sided test is a hypothesis test of a parameter in which the alternative hypothesis is the complement of the null hypothesis. In the context of a test of a scalar parameter, the alternative hypothesis states that the parameter is less than or greater than the value conjectured under the null hypothesis.
two-way ANOVA, two-way analysis of variance. A two-way ANOVA model contains two factors. Also see [PSS] power twoway.

two-way repeated-measures ANOVA, two-factor ANOVA. This is a repeated-measures ANOVA model with one within-subject factor and one between-subjects factor. The model can be additive (contain only main effects of the factors) or can contain main effects and an interaction between the two factors. Also see [PSS] power repeated.

type, eltype, and orgtype. The type of a matrix (or vector or scalar) is formally defined as the matrix’s eltype and orgtype, listed one after the other—such as real vector—but it can also mean just one or the other—such as the eltype real or the orgtype vector.

eltype refers to the type of the elements. The eltypes are

- real numbers such as 1, 2, 3.4
- complex numbers such as 1+2i, 3+0i
- string strings such as "bill"
- pointer pointers such as &vargame
- struct structures
- numeric meaning real or complex
- transmorphic meaning any of the above

orgtype refers to the organizational type. orgtype specifies how the elements are organized. The orgtypes are

- matrix two-dimensional arrays
- vector one-dimensional arrays
- colvector one-dimensional column arrays
- rowvector one-dimensional row arrays
- scalar single items

The fully specified type is the element and organization types combined, as in real vector.

type I error or false-positive result. The type I error of a test is the error of rejecting the null hypothesis when it is true. The probability of committing a type I error, significance level of a test, is often denoted as $\alpha$ in statistical literature. One traditionally used value for $\alpha$ is 5%. Also see type II error and power.

type I error probability. See probability of a type I error.

type I study. A type I study is a study in which all subjects fail (or experience an event) by the end of the study; that is, no censoring of subjects occurs.

type II error or false-negative result. The type II error of a test is the error of not rejecting the null hypothesis when it is false. The probability of committing a type II error is often denoted as $\beta$ in statistical literature. Commonly used values for $\beta$ are 20% or 10%. Also see type I error and power.

type II error probability. See probability of a type II error.

type II study. A type II study is a study in which there are subjects who do not fail (or do not experience an event) by the end of the study. These subjects are known to be censored.

unary operator. A unary operator is an operator applied to one argument. In -2, the minus sign is a unary operator. In !(a==b | a==c), ! is a unary operator.

unbalanced data. A longitudinal or panel dataset is said to be unbalanced if each panel does not have the same number of observations. See also weakly balanced and strongly balanced.
unbalanced design. An unbalanced design indicates an experiment in which the numbers of treated and untreated subjects differ. Also see [PSS] unbalanced designs.

unconfoundedness. See conditional-independence assumption.

under observation. A subject is under observation when failure events, should they occur, would be observed (and so recorded in the dataset). Being under observation does not mean that a subject is necessarily at risk. Subjects usually come under observation before they are at risk. The statistical concern is with periods when subjects are at risk but not under observation, even when the subject is (later) known not to have failed during the hiatus.

In such cases, since failure events would not have been observed, the subject necessarily had to survive the observational hiatus, and that leads to bias in statistical results unless the hiatus is accounted for properly.

Entry time and exit time record when a subject first and last comes under observation, between which there may be observational gaps, but usually there are not. There is only one entry time and one exit time for each subject. Often, entry time corresponds to analysis time \( t = 0 \), or before, and exit time corresponds to the time of failure.

Delayed entry means that the entry time occurred after \( t = 0 \).

underscore functions. Functions whose names start with an underscore are called underscore functions, and when an underscore function exists, usually a function without the underscore prefix also exists. In those cases, the function is usually implemented in terms of the underscore function, and the underscore function is harder to use but is faster or provides greater control. Usually, the difference is in the handling of errors.

For instance, function `fopen()` opens a file. If the file does not exist, execution of your program is aborted. Function `_fopen()` does the same thing, but if the file cannot be opened, it returns a special value indicating failure, and it is the responsibility of your program to check the indicator and to take the appropriate action. This can be useful when the file might not exist, and if it does not, you wish to take a different action. Usually, however, if the file does not exist, you will wish to abort, and use of `fopen()` will allow you to write less code.

unequal-allocation design. See unbalanced design.

Unicode. Unicode is a standard for encoding and dealing with text written in almost any conceivable living or dead language. Unicode specifies a set of encoding systems that are designed to hold (and, unlike extended ASCII, to keep separate) characters used in different languages. The Unicode standard defines not only the characters and encodings for them, but also rules on how to perform various operations on words in a given language (locale), such as capitalization and ordering. The most common Unicode encodings are mUTF-8, UTF-16, and UTF-32. Stata uses UTF-8.

Unicode character. Technically, a Unicode character is any character with a Unicode encoding. Colloquially, we use the term to refer to any character other than the plain ASCII characters.

Unicode normalization. Unicode normalization allows us to use a common representation and therefore compare Unicode strings that appear the same when displayed but could have more than one way of being encoded. This rarely arises in practice, but because it is possible in theory, Stata provides the `ustrnormalize()` function for converting between different normalized forms of the same string.

For example, suppose we wish to search for “˜n” (the lowercase n with a tilde over it from the Spanish alphabet). This letter may have been encoded with the single code point U+00F1. However, the sequence U+006E (the Latin lowercase “n”) followed by U+0303 (the tilde) is defined by Unicode to be equivalent to U+00F1. This type of visual identicalness is called canonical equivalence. The one-code-point form is known as the canonical composited form, and the multiple-code-point form
is known as the canonical decomposed form. Normalization modifies one or the other string to the opposite canonical equivalent form so that the underlying byte sequences match. If we had strings in a mixture of forms, we would want to use this normalization when sorting or when searching for strings or substrings.

Another form of Unicode normalization allows characters that appear somewhat different to be given the same meaning or interpretation. For example, when sorting or indexing, we may want the code point U+FB00 (the typographic ligature “ff”) to match the sequence of two Latin “f” letters encoded as U+0066 U+0066. This is called compatible equivalence.

**Unicode title-cased string.** See *titlecase, title-cased string, and Unicode title-cased string.*

**unidimensionality.** See *latent space.*

**uniqueness.** In factor analysis, the uniqueness is the percentage of a variable’s variance that is not explained by the common factors. It is also “1 – communality”. Also see *communality.*

**unitary matrix.** See *orthogonal matrix.*

**unit-root process.** A unit-root process is one that is integrated of order one, meaning that the process is nonstationary but that first-differencing the process produces a stationary series. The simplest example of a unit-root process is the random walk. See Hamilton (1994, chap. 15) for a discussion of when general ARMA processes may contain a unit root.

**unit-root tests.** Whether a process has a unit root has both important statistical and economic ramifications, so a variety of tests have been developed to test for them. Among the earliest tests proposed is the one by Dickey and Fuller (1979), though most researchers now use an improved variant called the augmented Dickey–Fuller test instead of the original version. Other common unit-root tests implemented in Stata include the DF–GLS test of Elliott, Rothenberg, and Stock (1996) and the Phillips–Perron (1988) test. See [TS] `dfuller`, [TS] `dfgls`, and [TS] `pperron`.

Variants of unit-root tests suitable for panel data have also been developed; see [XT] `xtunitroot`.

**unregistered variables.** See *registered and unregistered variables.*

**unrestricted transformation.** An unrestricted transformation is a Procrustes transformation that allows the data to be transformed, not just by orthogonal and oblique rotations, but by all conformable regular matrices. This is equivalent to a multivariate regression. Also see *Procrustes transformation* and *multivariate regression.*

**unstandardized coefficient.** A coefficient that is not standardized. If $mpg = -0.006 \times weight + 39.44028$, then $-0.006$ is an unstandardized coefficient and, as a matter of fact, is measured in mpg-per-pound units.

**unstructured and structured (correlation or covariance).** A set of variables, typically error variables, is said to have an unstructured correlation or covariance if the covariance matrix has no particular pattern imposed by theory. If a pattern is imposed, the correlation or covariance is said to be structured.

**upper one-sided test, upper one-tailed test.** An upper one-sided test is a one-sided test of a scalar parameter in which the alternative hypothesis is upper one sided, meaning that the alternative hypothesis states that the parameter is greater than the value conjectured under the null hypothesis. Also see *One-sided test versus two-sided test* under *Remarks and examples* in [PSS] `intro`.

UTF-8. UTF-8 stands for Universal character set + Transformation Format—8-bit. It is a type of Unicode encoding system that was designed for backward compatibility with ASCII and is used by Stata 14.

**vague prior.** See *noninformative prior.*

**valid initial state.** See *feasible initial value.*
vanishing adaptation. See diminishing adaptation.

VAR. A vector autoregressive (VAR) model is a multivariate regression technique in which each dependent variable is regressed on lags of itself and on lags of all the other dependent variables in the model. Occasionally, exogenous variables are also included in the model.

variable. In a program, the entities that store values ($a, b, c, \ldots, x, y, z$) are called variables. Variables are given names of 1 to 32 characters long. To be terribly formal about it: a variable is a container; it contains a matrix, vector, or scalar and is referred to by its variable name or by another variable containing a pointer to it.

Also, variable is sometimes used to refer to columns of data matrices; see data matrix.

variance components. In a mixed-effects model, the variance components refer to the variances and covariances of the various random effects.

variance–covariance matrix of the estimator. The estimator is the formula used to solve for the fitted parameters, sometimes called the fitted coefficients. The VCE is the estimated variance–covariance matrix of the parameters. The diagonal elements of the VCE are the variances of the parameters or equivalent, the square root of those elements are the reported standard errors of the parameters.

variance estimation. Variance estimation refers to the collection of methods used to measure the amount of sample-to-sample variation of point estimates; see SVY variance estimation.

varimax rotation. Varimax rotation maximizes the variance of the squared loadings within the columns of the matrix. It is an orthogonal rotation equivalent to oblimin with $\gamma = 1$ or to the Crawford–Ferguson family with $\kappa = 1/p$, where $p$ is the number of rows of the matrix to be rotated. Also see orthogonal rotation, oblimin rotation, and Crawford–Ferguson rotation.

varying and super-varying variables. A variable is said to be varying if its values in the incomplete observations differ across $m$. Imputed and passive variables are varying. Regular variables are nonvarying. Unregistered variables can be either.

Imputed variables are supposed to vary because their incomplete values are filled in with different imputed values, although an imputed variable can be temporarily nonvarying if you have not imputed its values yet. Similarly, passive variables should vary because they are or will be filled in based on values of varying imputed variables.

VCE. See variance–covariance matrix of the estimator.

VECM. A vector error-correction model (VECM) is a type of VAR that is used with variables that are cointegrated. Although first-differencing variables that are integrated of order one makes them stationary, fitting a VAR to such first-differenced variables results in misspecification error if the variables are cointegrated. See The multivariate VECM specification in [TS] vec intro for more on this point.

vector, colvector, and rowvector. A special case of a matrix with either one row or one column. A vector may be substituted anywhere a matrix is required. A matrix, however, may not be substituted for a vector.

A colvector is a vector with one column.

A rowvector is a vector with one row.

A vector is either a rowvector or colvector, without saying which.

view. A view is a special type of matrix that appears to be an ordinary matrix, but in fact the values in the matrix are the values of certain or all variables and observations in the Stata dataset that is currently in memory. Its values are not just equal to the dataset’s values; they are the dataset’s values: if an element of the matrix is changed, the corresponding variable and observation in the Stata dataset also changes. Views are obtained by st_view() and are efficient; see [M-5] st_view.
**void function.** A function is said to be void if it returns nothing. For instance, the function [M-5] `printf()` is a void function; it prints results, but it does not return anything in the sense that, say, [M-5] `sqrt()` does. It would not make any sense to code `x = printf("hi there")`, but coding `x = sqrt(2)` is perfectly logical.

**void matrix.** A matrix is said to be void if it is $0 \times 0$, $r \times 0$, or $0 \times c$; see [M-2] void.

**Wald test.** A Wald test is a statistical test based on the estimated variance–covariance matrix of the parameters. Wald tests are especially convenient for testing possible constraints to be placed on the estimated parameters of a model. Also see score test.

**Ward’s linkage clustering.** Ward’s-linkage clustering is a hierarchical clustering method that joins the two groups resulting in the minimum increase in the error sum of squares.

**weakly balanced.** A longitudinal or panel dataset is said to be weakly balanced if each panel has the same number of observations but the observations for different panels were not all made at the same times.

**weighted least squares.** Weighted least squares (WLS) is a method used to obtain fitted parameters. In this documentation, WLS is referred to as ADF, which stands for asymptotic distribution free. Other available methods are ML, QML, and MLMV. ADF is, in fact, a specific kind of the more generic WLS.

**weighted-average linkage clustering.** Weighted-average linkage clustering is a hierarchical clustering method that uses the weighted average similarity or dissimilarity of the two groups as the measure between the two groups.

**weighted-regression-adjustment estimator.** Weighted-regression-adjustment estimators use means of predicted outcomes for each treatment level to estimate each potential-outcome mean. The weights are used to estimate censoring-adjusted regression coefficients.

**white noise.** A variable $u_t$ represents a white-noise process if the mean of $u_t$ is zero, the variance of $u_t$ is $\sigma^2$, and the covariance between $u_t$ and $u_s$ is zero for all $s \neq t$.

**wide data.** See style.

**Wilks’s lambda.** Wilks’s lambda is a test statistic for the hypothesis test $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ based on the eigenvalues $\lambda_1, \ldots, \lambda_s$ of $E^{-1}H$. It is defined as

$$\Lambda = \frac{|E|}{|E + H|} = \prod_{i=1}^{s} \frac{1}{1 + \lambda_i}$$

where $H$ is the between matrix and $E$ is the within matrix. See between matrix.

**Wishart distribution.** The Wishart distribution is a family of probability distributions for nonnegative-definite matrix-valued random variables (“random matrices”). These distributions are of great importance in the estimation of covariance matrices in multivariate statistics.

**withdrawal.** Withdrawal is the process under which subjects withdraw from a study for reasons unrelated to the event of interest. For example, withdrawal occurs if subjects move to a different area or decide to no longer participate in a study. Withdrawal should not be confused with administrative censoring. If subjects withdraw from the study, the information about the outcome those subjects would have experienced at the end of the study, had they completed the study, is unavailable. Also see loss to follow-up and administrative censoring.

**within estimator.** The within estimator is a panel-data estimator that removes the panel-specific heterogeneity by subtracting the panel-level means from each variable and then performing ordinary least squares on the demeaned data. The within estimator is used in fitting the linear fixed-effects model.
**within matrix.** See *between matrix.*

**within-subject design.** This is an experiment that has at least one within-subject factor. See [PSS] power repeated.

**within-subject factor.** This is a factor for which each subject receives several or all the levels.

**WLF.** See worst linear function.

**WLS.** See weighted least squares.

**worst linear function.** A linear combination of all parameters being estimated by an iterative procedure that is thought to converge slowly.

**Yule–Walker equations.** The Yule–Walker equations are a set of difference equations that describe the relationship among the autocovariances and autocorrelations of an autoregressive moving-average (ARMA) process.

**z test.** A z test is a test for which a potentially asymptotic sampling distribution of the test statistic is a normal distribution. For example, a one-sample z test of means is used to test whether the mean of a population is equal to a specified value when the variance is assumed to be known. The distribution of its test statistic is normal. See [PSS] power onemean, [PSS] power twomeans, and [PSS] power pairedmeans.

**Zellner's g-prior.** Zellner’s g-prior is a form of a weakly informative prior for the regression coefficients in a linear model. It accounts for the correlation between the predictor variables and controls the impact of the prior of the regression coefficients on the posterior with parameter g. For example, $g = 1$ means that prior weight is 50% and $g \to \infty$ means diffuse prior.

---

**References**


Aalen, O. O. (1947– ), [ST] st
Agnesi, M. G. (1718–1799), [R] dydx
Akaike, H. (1927–2009), [R] estatic
Arellano, M. (1957– ), [XT] xtabond
Bartlett, M. S. (1910–2002), [TS] wntestb
Bayes, T. (1701(?)–1761), [BAYES] intro
Berkson, J. (1899–1982), [R] logit
Bliss, C. I. (1899–1979), [R] probit
Bonferroni, C. E. (1892–1960), [R] correlate
Box, G. E. P. (1919–2013), [TS] arima
Breusch, T. S. (1953– ), [R] regress postestimation time series
Brier, G. W. (1913–1998), [R] brier
Cholesky, A.-L. (1875–1918), [M-5] cholesky( )
Cleveland, W. S. (1943– ), [R] lowess
Cohen, J. (1923–1998), [R] kappa
Cornfield, J. (1912–1979), [R] epitab
Cox, D. R. (1924– ), [ST] stcox
Cox, G. M. (1900–1978), [R] anova
Cronbach, L. J. (1916–2003), [MV] alpha
Cunliffe, S. (1917–2012), [R] ttest
David, F. N. (1909–1993), [R] correlate
Dickey, D. A. (1945– ), [TS] dfuller
Dunnett, C. W. (1921–2007), [FN] Statistical functions
Durbin, J. (1923–2012), [R] regress postestimation time series
Efron, B. (1938– ), [R] bootstrap
Engle, R. F. (1942– ), [TS] arch
Fourier, J. B. J. (1768–1830), [R] cumul
Fuller, W. A. (1931– ), [TS] dfuller
Gabriel, K. R. (1929–2003), [MV] biplot
Galton, F. (1822–1911), [R] regress
Gauss, J. C. F. (1777–1855), [R] regress
Gnanadesikan, R. (1932–2015), [MV] diagnostic plots
Godfrey, L. G. (1946– ), [R] regress postestimation time series
Gompertz, B. (1779–1865), [ST] streg
Gosset, W. S. (1876–1937), [R] ttest
Granger, C. W. J. (1934–2009), [TS] vargranger
Hartley, H. O. (1912–1980), [MI] mi impute
Harvey, A. C. (1947– ), [TS] ucm
Hastings, W. K. (1930– ), [BAYES] bayesmh
Hausman, J. A. (1946– ), [R] hausman
Hays, W. L. (1926–1995), [R] esize
Heckman, J. J. (1944– ), [R] heckman
Hens, L. O. (1825–1854), [M-5] optimize()
Hopper, G. M. (1906–1992), [P] trace
Nelder, J. A. (1924–2010), [R] glm
Nelson, W. B. (1936–), [ST] sts
Newton, I. (1643–1727), [M-5] optimize()
Neyman, J. (1894–1981), [R] ci
Orcutt, G. H. (1917–2006), [TS] prais
Pearson, K. (1857–1936), [R] correlate
Penrose, R. (1931–), [M-5] pinv()
Perron, P. (1959–), [TS] pperron
Poisson, S.-D. (1781–1840), [R] poisson
Prais, S. J. (1928–2014), [TS] prais
Raphson, J. (1648–1715), [M-5] optimize()
Rubin, D. B. (1943–), [MI] intro substantive
Scheffé, H. (1907–1977), [R] oneway
Schur, I. (1875–1941), [M-5] schur()
Schwarz, G. E. (1933–2007), [R] estatic
Scott, E. L. (1917–1988), [R] intro
scree, [MV] screeplot
Shapiro, S. S. (1930–), [R] swilk
Shepard, R. N. (1929–), [MV] mds postestimation
plots
Sewheart, W. A. (1891–1967), [R] qc
Šidák, Z. (1933–1999), [R] correlate
singular value decompositions, [M-5] svd()
Smirnov, N. V. (1900–1966), [R] ksmirnov
Sneath, P. H. A. (1923–2011), [MV] measure_option
Snow, J. (1813–1858), [R] epitab
Sokal, R. R. (1926–2012), [MV] measure_option
Spearman, C. E. (1863–1945), [R] spearman
Thiele, T. N. (1838–1910), [R] summarize
Tobin, J. (1918–2002), [R] tobit
Toepplitz, O. (1881–1940), [M-5] Toepplitz()
Vandermonde, A.-T. (1735–1796),
[M-5] Vandermonde()
Wald, A. (1902–1950), [TS] varwle
Wallis, W. A. (1912–1998), [R] kwallis
Ward, J. H., Jr. (1926–2011), [MV] cluster linkage
Watson, G. S. (1921–1998), [R] regress postestimation
time series
West, K. D. (1953–), [TS] newey
White, H. L., Jr. (1950–2012), [U] 20 Estimation and
postestimation commands
Whitney, D. R. (1915–2007), [R] ranksum
Wilcoxon, F. (1892–1965), [R] signrank
Wilks, S. S. (1906–1964), [MV] manova
Wilson, E. B. (1879–1964), [R] ci
Woolf, B. (1902–1983), [R] epitab
Zellner, A. (1927–2010), [R] sureg
Author index

A

Aalen, O. O., [ST] stcrreg postestimation, [ST] sts
Abayomi, K. A., [MI] intro substantive, [MI] mi impl
Abraira, V., [R] logit postestimation
Abraira-García, L., [R] anova
Abrams, K. R., [R] meta
Abraira, V., [R] intro substantive, [MI] mi impl
Abraira-García, L., [R] anova
Abramowitz, M., [FN] Mathematical functions
Acevedo, G., [M-1] LAPACK
Achen, C. H., [R] scobit
Achenback, T. M., [MV] mvtest
Agnesi, M. G., [R] dydx
Ahn, S. K., [TS] vec intro
Ahrens, J. H., [FN] Random-number functions
Aigner, D. J., [R] frontier, [XT] xfrontier
Aiken, L. S., [R] poorr
Aisbett, C. W., [ST] stcox, [ST] streg
Aitchison, J., [BAYES] intro, [R] ologit, [R] oprobit
Aitken, A. C., [R] reg3
Aitkin, M. A., [MV] mvtset correlations
Aivazian, S. A., [R] ksmirnov
Ajejo, J., [XT] xprobit, [XT] xtreg postestimation
Akmân, V. E., [BAYES] bayesmh
Albert, A., [MI] mi impl, [MV] discrim, [MV] discrim logistic
Albert, P. S., [XT] xtgee
Aldenderfer, M. S., [MV] cluster
Alderman, M. H., [PSS] intro, [PSS] power repeated
Aldrich, J. H., [R] logit, [R] probit
Alexander, J. T., [R] mlexp
Alexandersson, A., [R] regres
Alf, E., Jr., [R] rofit, [R] rocreg
Alfarro, R., [MI] intro
Algina, J., [R] esize
Allredge, J. R., [R] pk, [R] pkcross
Allen, M. J., [MV] alpha
ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group, [PSS] intro, [PSS] power repeated
Allison, M. J., [MV] manova
Aldás, I., [R] inequality
Alonzo, T. A., [R] rocreg, [R] rocreg postestimation, [R] rocregplot
Alvarez, J., [XT] xtabond
Alwin, D. F., [SEM] example 9
Amemiya, T., [R] glm
Anderson, T. W., [M-1] LAPACK
Anderson, B. D. O., [TS] sspace
Anderson, E. B., [R] logistic
Anderson, E. B., [R] logistic
Anderson, E. B., [R] logistic
Anderson, E. B., [R] logistic
Anderson, M. L., [ST] stcrreg
Anderson, R. E., [R] rologit
Anderson, R. L., [R] anova
Anderson, S., [R] pkequiv
Anderberg, M. R., [MV] cluster, [MV] measure_option
Andersen, E. B., [R] logit
Andersen, P. K., [R] glm, [ST] stcox, [ST] stcrreg
Andersen, B. D. O., [TS] sspace
Anderson, M. L., [ST] stcrreg
Anderson, R. E., [R] rologit
Anderson, R. L., [R] anova
Anderson, S., [R] pkequiv
Anderson, T. W., continued


Andersson, T. M.-L., [ST] stcox


Andrich, D., [IRT] irt rsm, [SEM] example 28g

Andrieu, C., [BAYES] intro, [BAYES] bayesmh


Anscombe, F. J., [R] binreg postestimation, [R] glm, [R] glm postestimation

Ansley, C. F., [TS] arima

Arbuthnott, J., [R] signrank


Arminger, G., [R] sued


Armstrong, R. D., [R] qreg

Arnold, B. C., [MI] intro substantive, [MI] mi impute chained

Arnold, S. F., [MV] manova

Arnon, W. S., [ME] mestreg

Arora, S. S., [XT] xtivreg, [XT] xtreg

Arseven, E., [MV] discrim lda

Arthur, M., [R] symmetry

Atchadé, Y. F., [BAYES] intro, [BAYES] bayesmh

Atella, V., [R] frontier, [XT] xfrontier

Athen, B., [XT] xtnitirrunt

Atkinson, A. C., [FN] Random-number functions, [R] boxcox, [R] nl

Austin, P. C., [TE] tebalance

Azen, S. P., [R] anova, [U] 20.25 References

Aznar, A., [TS] vecrank


Babin, B. J., [R] rologit

Babu, A. J. G., [FN] Random-number functions


Bailie, R. T., [TS] arfima

Baker, F. B., [IRT] irt, [IRT] irt nrm

Baker, M. J., [BAYES] bayes

Baker, R. J., [R] glm

Baker, R. M., [R] ivregress postestimation

Bakker, A., [R] mean

Balaam, L. N., [R] pkcross

Balakrishnan, N., [FN] Statistical functions

Baldi, W. P., [ST] sterreg

Balestra, P., [XT] xtivreg


Bancroft, T. A., [R] stepwise

Banerjee, A., [XT] xtnitirrunt

Bang, H., [TE] teffects intro advanced

Barbin, ´E., [M-5] cholesky()

Barlow, R. E., [BAYES] intro

Barnard, G. A., [R] spearman, [R] tttest


Barnett, A. G., [R] glm

Barnow, B. S., [TE] etregress

Baron, R. M., [SEM] example 42g

Barrett, J. H., [PSS] intro

Barrison, I. G., [R] spearman


Bartlett, J. W., [MI] mi impute


Barton, C. N., [PSS] power repeated

Bartus, T., [R] margins


Basilevsky, A. T., [MV] factor, [MV] pca

Basmann, R. L., [R] ivregress, [R] ivregress postestimation

Basu, A., [R] betareg, [R] glm
Batistatou, E., [PSS] power
Batteze, G. E., [XT] xfrontier
Bauldry, S., [R] ivregress, [SEM] intro 5
Bauwens, L., [TS] mgarch
Bayart, D., [R] qc
Bates, Y., [BAYES] intro
Beale, E. M. L., [R] stepwise, [R] test
Beall, G., [MV] mvtest, [MV] mvtest covariances
Bean, J. A., [PSS] power cmh
Beaton, A. E., [R] rreg
Beck, N. L., [XT] xtglm, [XT] xtpce
Becker, S. O., [TE] tteffects intro advanced
Beerstecher, E., [MV] manova
Beggs, S., [R] rologit
Belanger, A. J., [R] stest, [R] swilk
Belsley, D. A., [R] regress postestimation, [R] regress postestimation diagnostic plots, [U] 18.14 References
Beltrami, E., [M-5] svd()
Bendel, R. B., [R] stepwise
Benedetti, J. K., [R] tetrachoric
Bennett, K. J., [R] nbreg, [R] poisson
Bera, J., [TS] arfima, [TS] arfima postestimation
Bera, R. J., [R] regress postestimation time series
Berger, J. O., [BAYES] intro
Berger, R. L., [PSS] intro, [R] ci
Berglund, P. A., [SVY] survey, [SVY] subpopulation estimation
Berk, K. N., [R] stepwise
Berk, R. A., [R] rreg
Berkes, I., [TS] mgarch
Berkson, J., [R] logit, [R] probit
Berliner, L. M., [BAYES] intro
Bern, P. H., [R] nestreg
Bernards, C. A., [MV] rotatemat
Bernardo, J. M., [BAYES] intro
Bernasco, W., [R] tetrachoric
Berndt, E. R., [R] truncvec
Bernstein, I. H., [MV] alpha
Berry, D. A., [BAYES] intro, [BAYES] intro
Berry, K. J., [R] ranksum
Besag, J., [BAYES] intro
Best, D. J., [FN] Random-number functions
Best, N. G., [BAYES] bayesstats ic
Bewley, R., [R] reg3
Beyer, W. H., [R] qc
Beyersmann, J., [ST] sterreg
Bhargava, A., [XT] xtregar
Bianchi, G., [TS] tsfilter, [TS] tsfilter bw
Bickeblöcker, H., [R] symmetry
Bickel, P. J., [D] egen, [R] rreg
Birdsall, T. G., [R] lroc
<table>
<thead>
<tr>
<th>Name</th>
<th>Reference</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blackwell, J. L., III</td>
<td>[R]</td>
<td>betareg</td>
</tr>
<tr>
<td>Bishop, D. T.</td>
<td>[PSS]</td>
<td>intro</td>
</tr>
<tr>
<td>Black, F.</td>
<td>[TS]</td>
<td>arch</td>
</tr>
<tr>
<td>Black, H. R.</td>
<td>[PSS]</td>
<td>intro, [PSS] power repeated</td>
</tr>
<tr>
<td>Black, W. C.</td>
<td>[R]</td>
<td>rologit</td>
</tr>
<tr>
<td>Blashfield, R. K.</td>
<td>[MV]</td>
<td>cluster</td>
</tr>
<tr>
<td>Blasius, J.</td>
<td>[MV]</td>
<td>ca, [MV] mca</td>
</tr>
<tr>
<td>Blasnik, M.</td>
<td>[D]</td>
<td>donevar, [D] split, [D] statsby</td>
</tr>
<tr>
<td>Bleda, M.-J.</td>
<td>[MV]</td>
<td>alpha</td>
</tr>
<tr>
<td>Blevins, J. R.</td>
<td>[R]</td>
<td>hetprobit</td>
</tr>
<tr>
<td>Blake, P. D.</td>
<td>[R]</td>
<td>icc</td>
</tr>
<tr>
<td>Bliss, C. I.</td>
<td>[R]</td>
<td>probit</td>
</tr>
<tr>
<td>Bloch, D. A.</td>
<td>[R]</td>
<td>briar</td>
</tr>
<tr>
<td>Bloomfield, P.</td>
<td>[R]</td>
<td>qreg, [TS] arfima</td>
</tr>
<tr>
<td>Blossfeld, H.-P.</td>
<td>[ME]</td>
<td>mestreg</td>
</tr>
<tr>
<td>Blum, A. L.</td>
<td>[PSS]</td>
<td>power cmh</td>
</tr>
<tr>
<td>Bowker, A. H.</td>
<td>[R]</td>
<td>symmetry</td>
</tr>
<tr>
<td>Boyd, N. F.</td>
<td>[R]</td>
<td>kappa</td>
</tr>
<tr>
<td>Boyle, J. M.</td>
<td>[P]</td>
<td>matrix symeigen</td>
</tr>
<tr>
<td>Bottai, M.</td>
<td>[R]</td>
<td>epitab</td>
</tr>
<tr>
<td>Bowier, J. R.</td>
<td>[TS]</td>
<td>streg postestimation</td>
</tr>
<tr>
<td>Boyd, R. A.</td>
<td>[R]</td>
<td>signrank</td>
</tr>
<tr>
<td>Bock, R. D.</td>
<td>[IR]</td>
<td>intarm</td>
</tr>
<tr>
<td>Bofinger, E.</td>
<td>[R]</td>
<td>qreg</td>
</tr>
<tr>
<td>Boice, J. D., Jr.</td>
<td>[R]</td>
<td>hitest, [R] epitab</td>
</tr>
<tr>
<td>Boland, P. J.</td>
<td>[R]</td>
<td>ttest</td>
</tr>
<tr>
<td>Bolduc, D.</td>
<td>[R]</td>
<td>asprobit</td>
</tr>
<tr>
<td>Bond, T. G.</td>
<td>[IRT]</td>
<td>ttest, [SEM] example 28g</td>
</tr>
<tr>
<td>Bonett, D. G.</td>
<td>[R]</td>
<td>ci</td>
</tr>
<tr>
<td>Bonferroni, C. E.</td>
<td>[R]</td>
<td>correlate</td>
</tr>
<tr>
<td>Boos, D. D.</td>
<td>[TE]</td>
<td>tteffects aipw</td>
</tr>
<tr>
<td>Borenstein, M.</td>
<td>[R]</td>
<td>meta</td>
</tr>
<tr>
<td>Borg, Ø.</td>
<td>[ST]</td>
<td>stcrreg</td>
</tr>
<tr>
<td>Bornhorst, F.</td>
<td>[XT]</td>
<td>xtnunitroot</td>
</tr>
<tr>
<td>Borowczyk, J.</td>
<td>[M-5]</td>
<td>cholesky()</td>
</tr>
<tr>
<td>Bos, J. M.</td>
<td>[R]</td>
<td>betareg</td>
</tr>
<tr>
<td>Boswell, T. M.</td>
<td>[ST]</td>
<td>streg postestimation</td>
</tr>
<tr>
<td>Boswijk, H. P.</td>
<td>[TS]</td>
<td>vec</td>
</tr>
<tr>
<td>Bound, J.</td>
<td>[R]</td>
<td>ivregress postestimation</td>
</tr>
<tr>
<td>Bover, O.</td>
<td>[XT]</td>
<td>xtdpd, [XT] xtdpdsys</td>
</tr>
<tr>
<td>Box, J. F.</td>
<td>[R]</td>
<td>anova</td>
</tr>
<tr>
<td>Boyd, R. A.</td>
<td>[R]</td>
<td>signrank</td>
</tr>
<tr>
<td>Brady, R. A.</td>
<td>[R]</td>
<td>logistic plots, [R] wilk</td>
</tr>
<tr>
<td>Brady, A. R.</td>
<td>[PSS]</td>
<td>intro, [R] logistic, [R] spikeplot</td>
</tr>
<tr>
<td>Brady, T. D.</td>
<td>[D]</td>
<td>edit</td>
</tr>
<tr>
<td>Brand, J. P. L.</td>
<td>[MI]</td>
<td>intro substantive, [MI] mi impute chained</td>
</tr>
<tr>
<td>Brännäs, K.</td>
<td>[R]</td>
<td>cpoisson</td>
</tr>
<tr>
<td>Brent, R. P.</td>
<td>[MV]</td>
<td>mdsmat, [MV] mvtest means</td>
</tr>
<tr>
<td>Brion, G. W.</td>
<td>[R]</td>
<td>brier</td>
</tr>
<tr>
<td>Brillinger, D. R.</td>
<td>[R]</td>
<td>jackknife</td>
</tr>
</tbody>
</table>
Brockwell, P. J., [TS] corrgam, [TS] sspace
Brody, H., [R] epitab
Brook, R. H., [R] brier
Brooks, S., [BAYES] intro
Brown, B. W., [ST] stsfilter
Brown, C. C., [R] epitab
Brown, G. K., [TE] etregress, [TE] teffects intro advanced
Brown, H., [ME] mixed
Brown, J. D., [MV] manova
Brown, L. D., [R] ci
Brown, M. B., [R] sctest, [R] tetrachoric
Brown, S. E., [R] symmetry
Brown, T. A., [SEM] intro 4
Brown, W., [R] icc
Broyden, C. G., [TS] forecast solve
Bru, B., [R] poisson
Brückner, E., [ME] mestreg
Brzezinski, M., [R] swilk
Brzinsky-Fay, C., [G-2] graph twoway rbar
Buchholz, A., [ST] streg
Buchner, D. M., [R] ladder
Buja, A., [U] 20.25 References
Bunch, D. S., [R] asprobit
Buot, M.-L. G., [MV] mvtest means
Burden, R. L., [M-5] solvenl()
Burden Study Group, [D] icd10
Burke, W. J., [R] tobit
Burket, G. R., [IRT] irt 3pl
Burkhauser, R. V., [MI] intro substantive
Burns, J. C., [ME] mixed
Burr, I. W., [R] qc
Burwell, D. T., [ME] mestreg
Buskens, V., [R] tabstat
Busso, M., [TE] stteffects ipwra, [TE] teffects overlap
Butterworth, S., [TS] tsfilter, [TS] tsfilter bw

C
Caffo, B. S., [BAYES] bayesstats summary
Cai, T., [R] rocreg
Cai, T. T., [R] ci
Cailliez, F., [MV] mdsmat
Cain, G. G., [TE] etregress
Caines, P. E., [TS] sspace
Caliendo, M., [TE] teffects intro advanced
Califf, R. M., [ST] stcox postestimation
Calinski, T., [MV] cluster, [MV] cluster stop
Camilli, G., [IRT] dif
Campbell, D. T., [SEM] example 17
Canner, J., [D] icd10
Cardell, S., [R] rologit
Carey, R. B., [D] icd10
Caria, M. P., [XT] xtgee
Carle, A. C., [ME] mixed
Carlile, T., [R] kappa
Carnes, B. A., [ST] streg
Carroll, J. B., [MV] rotatemat
<table>
<thead>
<tr>
<th>Author</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carson, R. T.,</td>
<td>[R] tnreg, [R] tpoisson</td>
</tr>
<tr>
<td>Carter, S. L.,</td>
<td>[ME] me, [ME] melogit,</td>
</tr>
<tr>
<td></td>
<td>[ME] meoprobit, [ME] mepoisson,</td>
</tr>
<tr>
<td></td>
<td>[ME] meqreglogit, [ME] meqroproportion,</td>
</tr>
<tr>
<td></td>
<td>[ME] mestreg, [R] frontier, [R] lrtest,</td>
</tr>
<tr>
<td>Casagrande, J. T.,</td>
<td>[PSS] intro, [PSS] power twoproportions</td>
</tr>
<tr>
<td>Casals, J.,</td>
<td>[TS] sspace</td>
</tr>
<tr>
<td>Casella, G.,</td>
<td>[BAYES] intro, [ME] me, [ME] meclolog,</td>
</tr>
<tr>
<td></td>
<td>[ME] meglm, [ME] melogit, [ME] menbreg,</td>
</tr>
<tr>
<td></td>
<td>[ME] meologit, [ME] meoprobit,</td>
</tr>
<tr>
<td></td>
<td>[ME] mepoisson, [ME] meprobit, [ME] mestreg,</td>
</tr>
<tr>
<td></td>
<td>[ME] mixed, [PSS] intro, [R] ci</td>
</tr>
<tr>
<td>Castellani, M.,</td>
<td>[R] betareg</td>
</tr>
<tr>
<td>Castillo, E.,</td>
<td>[MI] intro substantive, [MI] mi impute chained</td>
</tr>
<tr>
<td>Castro, L. M.,</td>
<td>[IRT] irt 3pl</td>
</tr>
<tr>
<td>Cattaneo, M. D.,</td>
<td>[R] gmm, [TE] etefffects,</td>
</tr>
<tr>
<td></td>
<td>[TE] stteffects intro, [TE] stteffects ipw,</td>
</tr>
<tr>
<td></td>
<td>[TE] tebalance, [TE] tebalance box,</td>
</tr>
<tr>
<td></td>
<td>[TE] tebalance density, [TE] tebalance overid,</td>
</tr>
<tr>
<td></td>
<td>[TE] tebalance summarize, [TE] tefffects intro,</td>
</tr>
<tr>
<td>Cattelan, A. M.,</td>
<td>[R] betareg</td>
</tr>
<tr>
<td>Cattell, R. B.,</td>
<td>[MV] factor postestimation,</td>
</tr>
<tr>
<td></td>
<td>[MV] pca postestimation, [MV] procustes,</td>
</tr>
<tr>
<td></td>
<td>[MV] screeplot</td>
</tr>
<tr>
<td>Caudill, S. B.,</td>
<td>[R] frontier, [XT] xfrontend</td>
</tr>
<tr>
<td>Caulcutt, R.,</td>
<td>[R] qc</td>
</tr>
<tr>
<td>Cefalu, M. S.,</td>
<td>[ST] stcox postestimation, [ST] stcurve,</td>
</tr>
<tr>
<td></td>
<td>[ST] ssts graph</td>
</tr>
<tr>
<td>Center for Human Resource Research,</td>
<td>[SEM] example 38g, [SEM] example 46g,</td>
</tr>
<tr>
<td></td>
<td>[XT] xt</td>
</tr>
<tr>
<td>Centers for Disease Control and Prevention,</td>
<td>[D] icd,</td>
</tr>
<tr>
<td></td>
<td>[D] icd9</td>
</tr>
<tr>
<td>Cerulli, G.,</td>
<td>[TE] etefffects, [TE] etpoisson,</td>
</tr>
<tr>
<td></td>
<td>[TE] etregress, [TE] tefffects intro advanced,</td>
</tr>
<tr>
<td></td>
<td>[TE] tefffects ipw</td>
</tr>
<tr>
<td>Chabert, J.-L.,</td>
<td>[M-5] cholesky()</td>
</tr>
<tr>
<td>Chadwick, J.,</td>
<td>[R] poisson</td>
</tr>
<tr>
<td>Chakraborti, S.,</td>
<td>[R] ksmirnov</td>
</tr>
<tr>
<td>Chaloner, K.,</td>
<td>[BAYES] intro</td>
</tr>
<tr>
<td>Chamberlain, G.,</td>
<td>[R] clogit, [R] gmm, [R] qreg</td>
</tr>
<tr>
<td></td>
<td>[R] grmeanby, [R] lowess, [U] 1.4 References</td>
</tr>
<tr>
<td>Chang, I. M.,</td>
<td>[R] margins</td>
</tr>
<tr>
<td>Chang, Y.,</td>
<td>[TS] sspace</td>
</tr>
<tr>
<td>Chang, Y.-J.,</td>
<td>[XT] xtivreg, [XT] xtreg</td>
</tr>
<tr>
<td>Chao, E. C.,</td>
<td>[ME] me, [ME] meqreglogit,</td>
</tr>
<tr>
<td></td>
<td>[ME] meqroproportion, [ME] meqroproportion postestimation</td>
</tr>
<tr>
<td>Charlett, A.,</td>
<td>[R] fp</td>
</tr>
<tr>
<td></td>
<td>[TS] corrg, [TS] pergrm, [TS] tsmsmooth,</td>
</tr>
<tr>
<td></td>
<td>[TS] tsmsmooth dexpontial, [TS] tsmsmooth exponential, [TS] tsmsmooth hwinters,</td>
</tr>
<tr>
<td></td>
<td>[TS] tsmsmooth ma, [TS] tsmsmooth shwinters,</td>
</tr>
<tr>
<td></td>
<td>[TS] Glossary</td>
</tr>
<tr>
<td>Chatfield, M.,</td>
<td>[R] anova</td>
</tr>
<tr>
<td>Chatfield, M. D.,</td>
<td>[D] merge</td>
</tr>
<tr>
<td>Chatterjee, S.,</td>
<td>[R] poisson, [R] regress, [R] regress postestimation, [R] regress postestimation diagnostic plots</td>
</tr>
<tr>
<td>Chávez Juárez, F. W.,</td>
<td>[R] inequality</td>
</tr>
<tr>
<td>Chen, H.,</td>
<td>[TS] mswitch</td>
</tr>
<tr>
<td>Chen, M.,</td>
<td>[D] drawnorm</td>
</tr>
<tr>
<td>Chen, X.,</td>
<td>[ME] mixed, [PSS] power oneproportion,</td>
</tr>
<tr>
<td></td>
<td>[PSS] power twoproportions, [R] logistic,</td>
</tr>
<tr>
<td></td>
<td>[R] logistic postestimation, [R] logit</td>
</tr>
<tr>
<td>Chernozhukov, V.,</td>
<td>[BAYES] intro</td>
</tr>
<tr>
<td>Cheung, Y. B.,</td>
<td>[ST] stcox</td>
</tr>
<tr>
<td>Cheung, Y.-W.,</td>
<td>[TS] dfgls</td>
</tr>
<tr>
<td>Chiang, C. L.,</td>
<td>[ST] ttable</td>
</tr>
<tr>
<td>Chib, S.,</td>
<td>[BAYES] intro</td>
</tr>
<tr>
<td>Chiburis, R.,</td>
<td>[R] heckman, [R] heckoprobit,</td>
</tr>
<tr>
<td></td>
<td>[R] heckprobit, [R] oprobit</td>
</tr>
<tr>
<td>Choi, B. C. K.,</td>
<td>[R] rocrf, [R] rocreg postestimation,</td>
</tr>
<tr>
<td></td>
<td>[R] rocregplot, [R] roctab</td>
</tr>
<tr>
<td>Choi, I.,</td>
<td>[XT] xtunitroot</td>
</tr>
<tr>
<td>Choi, M.-D.,</td>
<td>[M-5] Hilbert()</td>
</tr>
<tr>
<td>Choi, S. C.,</td>
<td>[MV] discm kdn</td>
</tr>
<tr>
<td>Cholesky, A.-L.,</td>
<td>[M-5] cholesky()</td>
</tr>
<tr>
<td>Chou, R. Y.,</td>
<td>[TS] arch</td>
</tr>
<tr>
<td>Chow, G. C.,</td>
<td>[R] contrast, [TS] estat sbknown</td>
</tr>
<tr>
<td>Chow, S.-C.,</td>
<td>[PSS] intro, [PSS] power onemean,</td>
</tr>
<tr>
<td></td>
<td>[PSS] power towmeans, [PSS] power pairedmeans, [PSS] power oneproportion,</td>
</tr>
<tr>
<td></td>
<td>[PSS] power exponential, [R] pk, [R] pkcross,</td>
</tr>
<tr>
<td></td>
<td>[R] pkequiv, [R] pkexamine, [R] pkshape</td>
</tr>
<tr>
<td>Christakis, N.,</td>
<td>[R] rologit</td>
</tr>
<tr>
<td>Christensen, W. F.,</td>
<td>[MV] biplot, [MV] ca,</td>
</tr>
<tr>
<td></td>
<td>[MV] candisc, [MV] canon, [MV] canon postestimation, [MV] cluster, [MV] discrim,</td>
</tr>
<tr>
<td></td>
<td>[MV] discrim estat, [MV] discrim knn,</td>
</tr>
<tr>
<td></td>
<td>[MV] discrim lda, [MV] discrim lda postestimation, [MV] discrim logistic,</td>
</tr>
<tr>
<td></td>
<td>[MV] discrim qda, [MV] discrim qda postestimation, [MV] factor, [MV] manova,</td>
</tr>
<tr>
<td></td>
<td>[MV] mca, [MV] mvtest, [MV] mvtest correlations, [MV] mvtest covariances,</td>
</tr>
<tr>
<td></td>
<td>[MV] mvtest means, [MV] mvtest normality,</td>
</tr>
<tr>
<td></td>
<td>[MV] pca, [MV] screeplot</td>
</tr>
</tbody>
</table>
Cox, N. J., continued


Cozad, B. J., [MV] DISCRIM IDA

Cragg, J. G., [R] churdle, [R] ivregress postestimation

Craig, A. S., [D] icd10

Cramer, E. M., [MV] PROCUSTES

Cramér, H., [R] tabulate twoway

Cramer, J. S., [R] LOGIT


Creel, M. D., [R] psoisson

Cribari-Neto, F., [R] betareg

Crichtley, F., [MV] mdsmat

Cronbach, L. J., [MV] alpha, [R] ice

Crouchley, R., [ME] mestreg

Croux, C., [R] reg


Crowder, M. J., [ST] stcrreg, [ST] streg

Crowe, P. R., [G-2] graph box


Crowther, M. J., [PSS] intro, [ST] streg

Cudeck, R., [SEM] estat gof, [SEM] methods and formulas for sem


Cumming, G., [R] esize, [R] regress postestimation


Cummings, T. H., [R] nbreg, [R] poisson

Cunliffe, S., [T] ttest

Curtis, J. T., [MV] clusterreg

Curtis-García, J., [R] smooth

Cushman, W. C., [PSS] intro, [PSS] power repeated

Cutler, J. A., [PSS] intro, [PSS] power repeated

Cutler, S. J., [ST] itable

Cuzick, J., [R] kappa, [R] nptrend

Czeckanowski, J., [MV] measure_option

D

D’Agostino, R. B., [MV] mvtest normality, [R] sktest, [R] swilk

D’Agostino, R. B., Jr., [R] sktest, [R] swilk

Daidone, S., [XT] xtabond

Danahy, D. T., [ME] mestreg

Daniel, C., [R] diagnostic plots, [R] oneway


Danuso, F., [R] nl

Dardanoni, V., [MI] intro substantive

Das, S., [XT] xtabond2

DasGupta, A., [R] ci

Davey, P. G., [D] icd10

Davey Smith, G., [R] meta

David, F. N., [R] correlate


David, J. S., [TS] arima

Davidon, W. C., [M-5] optimize()

Davidson, J., [TS] mswitch postestimation


Davies, R. B., [TS] estat shsmodel

Davis, B. R., [PSS] intro, [PSS] power repeated

Davis, G., [TS] arima

Davis, R. A., [TS] correlogram, [TS] sspace

Davison, A. C., [R] bootstrap


Day, W. H. E., [MV] cluster


De Backer, M., [ME] measqlogit postestimation
De Hoyos, R. E., [XT] xtrreg
de Irala-Estévez, J., [R] logistic
De Keyser, J., [ME] meqrlogit postestimation
de Kraker, M. E. A., [D] icd10
de Leeuw, J., [MV] ca postestimation
De Stavola, B. L., [ST] stcox, [ST] stset, [TE] ttests intro advanced
De Vroey, C., [ME] meqrlogit postestimation
de Wolf, I., [R] rologit
Deady, S., [R] betareg
Dearden, L., [TE] teffects intro advanced, [TE] teffects multivalued
Deaton, A. S., [R] nlsur, [U] 20.25 References
Deb, P., [R] churdle, [R] nbreg, [R] tobit
Debarsy, N., [R] ipoly
Dehon, C., [R] correlate
Deistler, M., [TS] sspace
del Rio, A., [TS] tsplit hp
DeMaris, A., [R] regress postestimation
Demnati, A., [SVY] direct standardization, [SVY] poststratification, [SVY] variance estimation
Denis, D., [G-2] graph twoway scatter
Desbordes, R., [R] ivregress
Desmarais, B. A., [R] zinb, [R] zip
Desu, M. M., [PSS] power exponential
Devroye, L., [FN] Random-number functions
Dewey, M. E., [R] correlate
Dey, D. D., [BAYES] intro
Dey, D. K., [BAYES] intro
Dice, L. R., [MV] measure_option
Dickens, R., [TS] prais
Dickson, E. R., [ST] stcrreg
Dicle, M. F., [D] import
Didelez, V., [R] ivregress
Diebold, F. X., [TS] arch
Dieter, U., [FN] Random-number functions
Diogby, P. G. N., [R] tetrachoric
Dijkstra, G. B., [MV] procures
Ding, Z., [TS] arch
Dobbin, K., [PSS] power
Dobson, A. J., [R] glm
Dodd, L. E., [R] rocreg
Dohoo, I., [R] epitab, [R] regress
Doll, R., [R] epitab, [R] poisson
Donald, S. G., [R] ivregress postestimation
Dongarra, J. J., [M-1] LAPACK
Driver, H. E., [MV] eivreg
Duan, N., [R] boxcox postestimation, [R] heckman, [TS] forecast estimates
Dubes, R. C., [MV] cluster
Duda, R. O., [MV] cluster, [MV] cluster stop
Dumyati, G., [D] icd10
Duncan, A. J., [R] qc
Duncan, O. D., [SEM] example 7
Ferrari, S. L. P., [R] betareg
Ferri, H. A., [R] kappa
Festinger, L., [R] ranksum
Fibrinogen Studies Collaboration, [ST] stcox
postestimation
Fidell, L. S., [MV] discrim, [MV] discrim lda
Field, A., [MI] mi estimate, [MI] mi impute,
[XT] xtgee
Field, C. A., [R] Bootstrap
Fielier, E. C., [R] pkequiv
Fienberg, S. E., [BAYES] intro, [R] kwallis,
[R] tabulate twoway
Filon, L. N. G., [R] correlate
Filoso, V., [R] regress
Finch, S., [R] esize
Findley, D. F., [R] estat ic
Findley, T. W., [R] ladder
Fine, J. P., [ST] stcrreg
Finlay, K., [R] ivprobit, [R] ivregress, [R] ivtobit
Finney, D. J., [IRT] irt 3pl, [R] probit, [R] tabulate
twoway
Fiocco, M., [ST] streg, [ST] streg postestimation
Fiorentini, G., [TS] mgarch
Fiorio, C. V., [R] kdensity
Fischer, G. H., [IRT] irt, [SEM] example 28g
Fiser, D. H., [R] estat gof, [R] lroc
Fishell, E., [R] kappa
Fisher, L. D., [MV] factor, [MV] pca, [PSS] intro,
[PSS] power twomaneov, [PSS] power oneway,
[PSS] power twoway, [R] anova, [R] stdize,
[R] oneway
Fisher, M. R., [XT] xtclustermat, [XT] xtgee,
[XT] xintreg, [XT] xlogit, [XT] xtologit,
[XT] xtprob, [XT] xtprob, [XT] xttobit
Fisher, N. I., [R] regress postestimation time series
Fisher, R. A., [R] Glossary, [MV] clustermat,
[MV] discrim, [MV] discrim estat,
[MV] discrim lda, [MV] Glossary, [P] levelsof,
[PSS] intro, [PSS] power twoproportions,
[PSS] power onecorrelation, [PSS] power
twocorrelations, [R] anova, [R]anova, [R] esize,
[R] ranksum, [R] signrank, [R] tabulate
twoway, [ST] streg
Fiske, D. W., [SEM] example 17
Fitzgerald, T. J., [TS] tfilter, [TS] tsfilter cf
Fitmaurice, G. M., [ME] mixed
Fix, E., [MV] discrim knn
twoway contour, [M-5] solvenl(), [P] matrix
symeigen, [R] dydx, [R] vswls, [TS] arch,
[TS] arima
Flay, B. R., [ME] me, [ME] meglm, [ME] meologit,
[ME] meoprobit, [XT] xlogit, [XT] xtprob
Flegal, J. M., [BAYES] bayesstats summary
Fleiss, J. L., [PSS] intro, [PSS] power oneproportion,
[PSS] power twoproportions, [R] stdize,
[R] epitab, [R] icc, [R] kappa
Fletcher, K., [R] rocreg, [R] rocreg postestimation,
[R] rocregplot
Fletcher, R., [M-5] optimize()
Flood, S., [R] mlexp
Flynn, Z. L., [R] gmm
Folsom, R. C., [R] rocreg, [R] rocreg postestimation,
[R] rocregplot
Ford, C. E., [PSS] intro, [PSS] power repeated
Ford, J. M., [R] frontier, [XT] xtfrontier
Forsythe, A. B., [R] sdttest
Forthofer, R. N., [R] stdize
Fosheim, G. E., [D] icd10
Foster, A., [R] regress
Fouladi, R. T., [R] esize
Foulkes, M. A., [PSS] power cox, [PSS] power
exponential
Fourier, J. B. J., [R] cumul
Fox, C. M., [IRT] irt, [SEM] example 28g
Fox, J., [R] kdensity, [R] Iv
Fox, W. C., [R] lroc
Francia, R. S., [R] swilk
Francis, C., [PSS] intro, [PSS] power repeated
Frank, M. W., [XT] xtabond, [XT] xtdpd,
[XT] xtdpdsys
estimation, [U] 20.25 References
Franklin, C. H., [D] cross
Franzese, R. J., Jr., [XT] xtpcse
Franzini, L., [XT] xtregr
Fraser, M. W., [TE] stteffects intro, [TE] stteffects
ipw, [TE] stteffects ipwra, [TE] stteffects
postestimation, [TE] stteffects ra, [TE] stteffects
wra, [TE] tebalance
Frechette, G. R., [XT] xtprobit
Freedman, L. S., [PSS] intro, [PSS] power cox,
[PSS] power exponential, [PSS] power logrank
Freeman, D. H., Jr., [SVY] svy: tabulate twoway
Freeman, E. H., [SEM] estat stable
Freeman, J. L., [R] epitab, [SVY] svy: tabulate
twoway
Freeman, J. R., [TS] time series, [TS] arima,
Freese, J., [R] asprobit, [R] clogit, [R] cloglog,
[R] logistic, [R] logit, [R] mlogit, [R] mprobit,
[R] nbreg, [R] ologit, [R] oprobit, [R] poisson,
[R] probit, [R] regress, [R] regress
postestimation, [R] nbreg, [R] tpoisson,
[R] zinb, [R] zip, [U] 20.25 References
Fridkin, S. K., [D] icd10
Friedman, J. H., [MV] discrim knn
Friedman, M., [TS] arima
Friendly, M., [G-2] graph twoway scatter
Frölich, M., [R] qreg, [TE] tteffects multivalued
Frome, E. L., [R] qreg
Frühwirth-Schnatter, S., [TS] msimul
Frydenberg, M., [R] stdize, [R] roccomp, [R] roctab
Fu, V. K., [R] ologit
Gallup, J. L., [M-5]
Gallant, A. R., [R]
estat gof
stcrreg
Garcia, R., [TS]
tsway
Genest, C., [R]
Genadek, K., [R]
Genover, D., [BAYES]
tstest
sts test
Gijbels, I., [R]
tetrachoric
Gillispie, C. C., [R]
tobit
Gill, R. D., [ST] stcox
PH-assumption tests
Gillham, N. W., [R] regress
Gilpin, C. C., [R]
glm
Glass, G. V., [R] esize
Glass, R. I., [R] epibas

Gray, R. J., [ST] streg
Green, B. F., [MV] cca, [MV] mca, [MV] mca postestimation, [SEM] example 35g, [SEM] example 36g
Greenberg, E., [BAYES] intro
Greenhouse, J. B., [R] epitable
Greenhouse, S. W., [PSS] power repeated, [R] anova, [R] epitable
Greenwood, M., [ST] itable, [ST] sts
Greenwood, P., [MI] intro substantive
Gregoire, A., [R] kappa
Grieve, R., [R] bootstrap, [R] bstat
Griffith, J. L., [R] brier
Griffith, R., [R] gmm
Grimm, R. H., [PSS] intro, [PSS] power repeated
Grimmett, G., [M-5] halton()
Grisetti, R., [R] betareg
Grisson, R. J., [R] esize, [R] regress postestimation
Grizzle, J. E., [R] vwls
Grogger, J. T., [R] nbreg, [R] tpoisson
Gronau, R., [R] heckman, [SEM] example 45g
Groothuis-Oudshoorn, C. G. M., [MI] intro substantive, [MI] mi impute chained
Gross, A. J., [ST] itable
Grundmann, H., [D] iced10
Grzebyk, M., [ST] sts
Guan, W., [R] bootstrap
Guenthner, W. C., [PSS] intro
Guerry, A.-M., [G-2] graph twoway histogram
Guidolin, M., [TS] msieffect
Guilkey, D. K., [XT] xprobit
Guillelet, M., [M-5] cholesky()
Guimarães, P., [XT] xtnbreg
Guimu, S., [R] poisson

H

Haan, P., [R] asmprobit, [R] mlogit, [R] mprobit
Haario, H., [BAYES] intro, [BAYES] bayesmh
Haas, K., [M-5] moptimize()
Hadamard, J. S., [FN] Matrix functions
Hadi, A. S., [R] poisson, [R] regress, [R] regress postestimation, [R] regress postestimation diagnostic plots
Hadorn, D. C., [R] brier
Hadri, K., [XT] xtnonparametric
Hahn, G. J., [M-5] moptimize()
Hahn, J., [R] ivregress postestimation
Hair, J. F., Jr., [R] rologit
Hajian-Tilaki, K. O., [R] rocreg
Hajivassiliou, V. A., [R] asmprobit
Hakkio, C. S., [D] egen
Hald, A., [R] qreg, [R] regress, [R] signrank, [R] summarize
Haldane, J. B. S., [R] epitab, [R] ranksum
Hall, A. D., [R] frontier
Hall, N. S., [R] anova
Hall, P., [R] bootstrap, [R] qreg, [R] regress postestimation time series
Hall, W. J., [MV] biplot, [R] roccomp, [R] rocfit, [R] roctab
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, B. H., [M-5] optimize()
Hall, B. H., [M-5] optimize()
Hall, N. S., [R] anova
Hall, W. J., [MV] biplot, [R] roccomp, [R] rocfit, [R] roctab
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Hall, A. D.
Harville, D. A., [ME] meglm, [ME] mixed
Hassell, J. F., [ST] stst
Hassler, U., [TS] irf create
Hastie, T. J., [MV] discrim kNN, [R] grmeanby, [R] slogit
Hastings, W. K., [BAYES] intro, [BAYES] bayesmh
Hastorf, A. H., [R] epitable
Haughton, J. H., [R] inequality
Hauser, M. A., [TS] arfima
Havnes, T., [R] inequality
Hawkins, C. M., [PSS] intro, [PSS] power repeated
Hayashi, F., [R] gmm, [R] ivpoisson, [R] ivregress, [R] ivreg postestimation
Hayes, R. J., [R] permute
Hays, W. L., [R] esize, [R] regress postestimation
He, X., [ST] stcox PH-assumption tests
Hédelin, G., [ST] sts
Hedges, L. V., [R] esize, [R] meta
Hedley, D., [ST] streg, [ST] stcrreg postestimation
Heeringa, S. G., [SVY] survey, [SVY] subpopulation estimation
Heidelberger, P., [BAYES] intro
Heinecke, K., [P] matrix mkmat
Heinonen, O. P., [R] epitable
Heiss, F., [R] nlogit
Heitjan, D. F., [MI] intro substantive, [MI] mi impute
Heller, G., [ST] stcox postestimation
Hemming, K., [PSS] intro
Hempel, S., [R] epitable
Henderson, B. E., [R] symmetry
Henderson, C. R., [ME] me, [ME] mixed
Hendrickx, J., [R] mlogit, [R] xi
Hennevogl, W., [ME] me
Henry-Amar, M., [ST] ttable
Hensher, D. A., [R] nlogit
Henzé, N., [MV] mvtest, [MV] mvtest normality
Hermite, C., [M-5] issymmetric()
Herr, J. L., [TE] tteffects intro advanced, [TE] tteffects nmnmatch
Herrero, F. J., [ME] mixed
Herrin, J., [U] 18.14 References
Herriot, J. G., [M-5] spline3()
Hertz, S., [ST] stsplit
Hesse, L. O., [M-5] moptimize()
Hessenberg, K. A., [M-5] hessenbergd()
Heston, A., [XT] xtunitroot
Heyde, C. C., [U] 1.4 References
Hickam, D. H., [R] brier
Higbee, K. T., [D] clonevar, [D] ds
Higdon, D., [BAYES] intro
Higgins, J. E., [R] anova
Higgins, J. P. T., [R] meta
Higgins, M. L., [TS] arch
Hilbe, J. M., [FN] Random-number functions,
Hildreth, C., [TS] prais
Hilferty, M. M., [MV] mvtest normality
Hilgard, E. R., [R] epitable
Hill, A. B., [R] epitable, [R] poisson
Hill, J., [ME] me
Hill, W. G., [R] epitable
Hills, S. E., [BAYES] intro, [BAYES] bayesmh
Hinkley, D. V., [R] bootstrap
J

Jaccard, P., [MV] measure_option
Jackman, R. W., [R] regress postestimation
Jackson, J. E., [MV] pca, [MV] pca postestimation
Jacobi, C. G. J., [M-5] deriv()
Jacobs, K. B., [R] symmetry
Jacobs, M., [D] duplicates
Jacoby, W. G., [MV] biplot
Jaeger, A., [TS] tfilter, [TS] tfilter hp
Jaeger, D. A., [R] ivregress postestimation
Jagannathan, R., [TS] arch
Jain, A. K., [MV] cluster
James, B. R., [R] rocreg, [R] rocreg postestimation
James, G. S., [MV] mvtest, [MV] mvtest means
James, K. L., [R] rocreg, [R] rocreg postestimation
Jang, D. S., [SVY] variance estimation
Jarque, C. M., [R] sktest, [TS] varnorm, [TS] vnecnorm
Jarrett, R. G., [BAYES] bayesmh
Jeanetteau, T., [TS] mgarch
Jenkins, B., [M-5] hash1()
Jensen, A. R., [MV] rotate
Jensen, D. R., [MV] mvtest, [MV] mvtest means
Jerez, M., [TS] sspace
Jewell, N. P., [R] epitab
Jick, H., [R] epitab
Johnson, M. E., [R] sdttest
Johnson, M. M., [R] sdttest
[J] 1.4 References
Johnson, S., [R] epitab
Johnson, V. E., [BAYES] intro
Johnston, W., [MI] intro substantive, [SVY] survey
Johnston, J., [XT] xtrc
Johnston, J. E., [R] ranksum
Jolliffe, D., [R] inequality, [R] qreg, [R] regress
Jolliffe, I. T., [MV] biplot, [MV] pca, [R] brier
Jones, A., [R] heckman, [R] logit, [R] probit
Jones, B. D., [TS] mswitch
Jones, B. S., [ST] stcox, [ST] streg
Jones, D. R., [R] meta
Jones, G. L., [BAYES] intro, [BAYES] bayesstats summary
Jones, M. C., [R] kdensity, [R] lpoly
Jones, P. S., [M-5] Vandermonde()
Jordan, C., [M-5] svd()
Jöreskog, K. G., [MV] factor postestimation, [SEM] estat residuals
Jorgensen, R. A., [ST] stcox
Jorner, U., [G-1] graph intro
Joyce, R., [TS] arima
K

Kachitvichyanukul, V., [FN] Random-number functions
Kackar, R. N., [ME] mixed
Kalman, R. E., [TS] arima
Kalmijn, M., [R] tetrachoric
Kamphuis, J. H., [TS] mswitch
Kang, J. D. Y., [TE] stteffects intro advanced
Kantor, D., [D] cf, [FN] Programming functions
Karim, M. R., [ME] meglm
Karlin, S., [TS] mswitch
Kaufman, J., [D] ds
Kaufman, R. L., [U] 20.25 References
Keane, M. P., [R] asmprob
Keeler, E. B., [R] brier
Keiding, N., [ST] stcrreg, [ST] stsplit
Kelley, K., [R] esize, [R] regress postestimation
Kelly, S., [IRT] irt
Kemp, A. W., [FN] Random-number functions, [R] nbreg, [R] poisson
Kempthorne, P. J., [R] regress postestimation
Kendall, D. G., [MV] mds
Kendall, M. G., [MV] measure_option, [R] centile, [R] spearman, [R] tabulate twoway
Kenny, D. A., [SEM] intro 4, [SEM] example 42g
Kerlinger, F. N., [R] esize, [R] regress postestimation
Kesarlinski, H. J., [R] esize
Kettenring, J. R., [R] diagnostic plots
Keynes, J. M., [R] ameans
Khan, M. R., [R] epitab
Khan, S., [R] hetprobit
Khandker, S. R., [R] inequality
Khanti-Akom, S., [XT] xthtaylor
Khare, M., [MI] intro substantive
Khanti-Akom, S., [XT] xthtaylor
Khuri, A. I., [ME] mixed
Kiernan, M., [R] kappa
Kieser, M., [PSS] intro
Kilian, L., [TS] forecast solve
Kim, C.-J., [TS] mswitch, [TS] mswitch postestimation
Kim, H.-J., [TS] estat sbingle
Kim, J. J., [R] esize, [R] regress postestimation
Kim, J. O., [MV] factor
Kim, S., [BAYES] intro, [IRT] irt
Kimber, A. C., [ST] streg
Kimbrough, J. W., [MV] discrim knn
Kinderman, A. J., [FN] Random-number functions
King, A. A., [M-2] intro
King, J., [IRT] irt
King, M., [R] mlexp
King, M. L., [TS] prais
Kirk, R. E., [R] esize, [R] regress postestimation
Kirkwood, B. R., [R] dstdize, [R] summarize
Kitagawa, G., [R] BIC note
Kiviet, J. F., [XT] xtabond
Klar, J., [R] estat gof
Klecka, W. R., [MV] discrim, [MV] discrim lda
Kleiber, C., [R] inequality
Lunt, M., [R] ologit, [R] slogit, [TE] tteffects multivalued
Lurie, M. B., [MV] manova
Lynfield, R., [D] icd10

M

Ma, G., [R] roccomp, [R] rocfit, [R] roctab
Maas, B., [BAYES] bayesmh
Mack, T. M., [R] symmey
MacKinnon, D. P., [SEM] example 42g
MacLaren, M. D., [FN] Random-number functions
MacMahon, B., [R] epitab
MacRae, K. D., [R] binreg
MacRury, T. E., [XT] xthealower
Madansky, A., [R] runtest
Madigan, D., [ST] st's
Magnus, J. R., [TS] var svar
Maguire, B. A., [BAYES] bayesmh
Mair, C. S., [ME] membreg, [ME] mepoisson, [ME] meqrpoisson, [SEM] example 39g
Makles, A., [MV] cluster kmeans and kmedians
Malitz, F., [IRT] irt
Mallick, B. K., [BAYES] intro

Mallows, C. L., [R] regress postestimation diagnostic plots
Manca, A., [R] betareg
Manchul, L., [ST] streg, [ST] stcrreg postestimation
Mandelbrot, B. B., [TS] arch
Mangel, M., [TS] varwle
Manjón, M., [R] nbreg postestimation, [R] poisson postestimation, [R] zinb postestimation
Manly, B. F. J., [MV] discm qda postestimation
Mann, H. B., [R] kwallis, [R] ranksum
Manski, C. F., [R] gmm
Mansuy, R., [ST] stcox postestimation
Mantel, H., [SVY] ssv bootstrap, [SVY] variance estimation
Maravall, A., [TS] tfilter hp
Marcellino, M., [XT] xtunitroot
Marchenko, Y. V., [ME] me, [ME] meglm,

[ME] melogit, [ME] meprobit,
[ME] mepoisson, [ME] meqrlogit,
[ME] meqrpoisson, [ME] mestreg, [ME] mixed,
[MI] intro substantive, [MI] mi estimate,
[MI] mi impute, [PSS] power exponential,
[PSS] power logrank, [R] anova, [R] loneway,
[R] oneway, [R] sktest, [ST] survival analysis,
[ST] sttest, [ST] stsplit, [ST] stvary,
[TE] tteffects intro, [XT] xtreg
Marden, J. I., [R] rologit
Mardia, K. V., [MI] mi impute mvd, [MV] discrim,
[MV] discrim lda, [MV] factor, [MV] manova,
[MV] matrix dissimilarity, [MV] mds,
[MV] mds postestimation, [MV] mdslong,
[MV] mdsmat, [MV] mvttest, [MV] mvttest means, [MV] mvttest normality, [MV] pca,
[MV] procrustes, [P] matrix dissimilarity
Maris, G., [IRT] irt 3pl
Mark, D. B., [ST] stcox postestimation
Markel, H., [R] epitab
Markov, A., [BAYES] intro
Markowski, C. A., [R] sdtest
Markowski, E. P., [R] sdtest
Marks, H. M., [ST] st's
Marr, J. W., [ST] stsplit
Marsaglia, G., [FN] Random-number functions
Marschak, J., [R] ivregress
Marsh, H. W., [SEM] example 19
Marsh, J., [PSS] intro
Martin, M. E., [SVY] ssv: tabulate oneway
Martin, W., [R] epitab, [R] regress
Martínez, M. A., [R] logistic
Martínez, O., [R] nbreg postestimation, [R] poisson postestimation, [R] zinb postestimation, [R] zip postestimation
Martínez-Beneito, M. A., [TS] mswitch
Mascher, K., [R] rocreg, [R] rocregpostestimation, [R] rocregplot
Master, I. M., [R] exlogistic
Masters, G. N., [IRT] intro, [IRT] itrt pcm
Mastrucci, M. T., [R]
Matsumoto, M., [FN] Random-number functions, [R] set seed
Mátyás, L., [R] gmm
Mayer, K. U., [ME] mestreg
Mazliak, L., [ST] stcox postestimation
Mazúa, V. G., [FN] Matrix functions
McCabe, S. E., [SVY] estat
McCabbey, D. F., [R] areg, [XT] xtreg
McCleary, S. J., [R] regress postestimation diagnostic plots
McClish, D. K., [R] rocreg
McCrory, J., [TE] stteffects ipwra, [TE] stteffects overlap
McCullough, B. D., [TS] corrgam
McDonald, A., [ME] menbreg, [ME] mepoisson, [ME] meqrpoisson, [SEM] example 39g
McDonald, J. A., [R] sunflower
McDonald, J. F., [R] tobit, [R] tobit postestimation
McDonald, R. P., [IRT] itrt
McDougal, L. K., [D] icd10
McDowell, A. W., [R] sureg, [TS] arima
McGilchrist, C. A., [ST] stcox, [ST] streg
McGill, R., [R] sunflower
McGinnis, R. E., [R] symmetry
McGraw, K. O., [R] icc
McGuire, T. J., [R] dstdize
McKelvey, R. D., [R] ologit
McLain, A. C., [R] nbreg, [R] poisson
McLeod, A. I., [TS] arima, [TS] ucm
McNeil, D., [R] poisson, [ST] stcrreg
McNemar, Q., [PSS] intro, [R] epitab
Mead, R., [M-5] optimize()
Mealli, F., [MI] intro substantive
Meeusen, W., [R] frontier, [XT] xtfrontier
Mehta, P. D., [SEM] example 30g
Meijering, E., [D] ipolate
Meiselman, D., [TS] arima
Melly, B., [R] greg, [TE] teffects multivalued
Mendenhall, W., III, [SVY] survey
Mensing, R. W., [R] anova postestimation
Mergoupis, T., [TE] etregress, [TE] teffects intro advanced
Merryman, S., [XT] xtunitroot
Metropolis, N., [BAYES] intro, [BAYES] bayesmh
Metz, C. E., [R] Iroc
Meulders, M., [MI] intro substantive, [MI] mi impute
Meyer, B. D., [ST] discrete
Miao, W., [R] sdttest
Micali, N., [MI] mi estimate, [MI] mi impute, [XT] xtgee
Michener, C. D., [MV] measure_option
Mickey, M. R., [MV] discrim estat
Midhune, D., [SVY] estat, [SVY] svy estimation
Mielke, P. W., Jr., [R] brier, [R] ranksym
Miettinen, O. S., [R] epitab
Mihaly, K., [R] areg, [XT] xtreg
Miller, A. B., [R] kappa
Miller, H. W., [SVY] survey, [SVY] svy estimation
Miller, J. I., [TS] sspace
Milliff, R. F., [BAYES] intro
Milligan, G. W., [MV] cluster, [MV] cluster programming subroutines, [MV] cluster stop
Milosevic, M., [ST] stcrreg, [ST] stcrreg postestimation
Min, C.-K., [BAYES] intro
Miquel, J., [BAYES] intro
Mitchell, C., [R] exlogistic
Miura, H., [U] 14.11 Reference
Modica, S., [MI] intro substantive
Moffatt, P. G., [R] churdle
Moffitt, R. A., [R] Tobit, [R] tobit postestimation
Mogstad, M., [R] inequality
Molenaar, I. W., [IRT] irt, [SEM] example 28g
Moler, C. B., [P] matrix symeigen
Monahan, J. F., [FN] Random-number functions
Monshouwer, K., [MV] mvtest
Monson, R. S., [R] bitest, [R] epitab
Montes-Rojas, G., [XT] xtreg, [XT] xtreg postestimation
Montoya, D., [R] rocreg, [R] rocreg postestimation, [R] rocregplot
Mood, A. M., [R] centile
Moon, H. R., [XT] xtunitroot
Moore, E. H., [M-5] pinv( )
Moore, J. B., [TS] sspace
Moore, R. J., [FN] Statistical functions
Moran, J. L., [R] dstdize
Morgenstern, H., [R] epitab, [R] epitab
Mori, M., [ST] stcrreg
Morris, C., [R] bootstrap
Morris, J. N., [ST] stsplit
Morris, N. F., [R] binreg
Morris, T. P., [MI] mi impute, [MI] mi impute pmm
Morrison, M. A., [D] icd10
Morrow, A., [R] epitab
Mortimore, P., [MI] mi estimate
Mosier, C. I., [MV] procrustes
Moskowitz, M., [R] kappa
Mosteller, C. F., [R] jackknife, [R] regress, [R] regress postestimation diagnostic plots, [R] rreg
Moulton, L. H., [R] permute
Muller, C. W., [MV] factor
Muirhead, R. J., [MV] gmm, factor
Nagel, R. W., [MV] discrim lda
Nagler, J., [R] scobit
Naiman, D. Q., [R] qreg
Nam, J., [PSS] power cmh, [PSS] power trend
Nannicini, T., [TE] etregress
Nardi, G., [R] epitab
Narendranathan, W., [XT] xtregar
Narula, S. C., [R] qreg
Nash, J. C., [G-2] graph box
Nash, J. D., [DI] infile (fixed format), [D] merge
National Center for Health Statistics, [D] icd, [D] icd9
Navarro-Lozanno, S., [TE] tteffects intro advanced
Neal, R. M., [BAYES] intro
Neale, M. C., [SEM] example 30g
Neath, R., [BAYES] bayesstats summary
Nee, J. C. M., [R] kappa
Neff, R. K., [R] epitab
Neumann, H., [MV] mdsmat
Newel, D. G., [MV] mvtest, [MV] mvtest means
Nelson, F. D., [R] logit, [R] probit
Nelson, W., [ST] sterreg postestimation, [ST] sts
Nelson, W. C., [MV] mvtest correlations
Neter, J., [PSS] power oneway, [R] pkcross, [R] pkequiv, [R] pkshape, [R] regress postestimation
Neudecker, H., [TS] var svar
Nevels, K., [MV] procrustes
Newbold, P., [TS] arima, [TS] vec intro


Newton, I., [M-5] optimize()


Neumann, J., [R] ei

Ng, E. S.-W., [ME] me, [ME] meglm, [ME] melogit, [ME] meqreg2logit, [R] bootstrap, [R] bstat

Ng S., [TS] dfgls

Nicewander, W. A., [R] correlate


Nielsen, B., [TS] varsoc, [TS] vec intro

Nightingale, F, [G-2] graph pie

Nishimura, T., [FN] Random-number functions, [R] set seed

Nolan, D., [R] diagnostic plots

Nordlund, D. J., [MV] discrim lda

Norton, E. C., [R] churdle, [R] tobit


Nunnally, J. C., [MV] alpha

O

O'Brien, R. G., [PSS] power oneway


O'Connell, P. G. J., [XT] xtunitroot


O'Donnell, C. J., [XT] xtfrontier

O'Donnell, O., [SVY] svy estimation, [SVY] svyset

O'Fallon, W. M., [R] logit

O'Hara, B., [BAYES] bayesmh

O'Neill, D., [R] gmm, [R] inequality


Olstead, R., [XT] xtunitroot

Ochiai, A., [MV] measure_option

Odum, E. P., [MV] clustermat

Ochert, G. W., [R] ncom, [R] rocreg postestimation, [R] rocregplot

Oh, K.-Y., [XT] xtunitroot


Olivier, D., [R] expoisson

Olkin, I., [MV] hotelling, [R] kwallis, [TS] wntestb

Olsen, M. K., [MI] intro substantive

Olshansky, S. J., [ST] streg

Olson, J. M., [R] symmetry

Omb, M., [TS] arima

Oparil, S., [PSS] intro, [PSS] power repeated

Orcutt, G. H., [TS] prais


Osbat, C., [XT] xtunitroot

Osterlind, S. J., [IRT] dif

Osterwald-Lenum, M. G., [TS] arima

Ostle, B., [R] anova postestimation

Ott, R. L., [SVY] survey

Over, M., [R] regress, [XT] xtivreg

Ovgaard, M., [ST] stcox

Owen, A. L., [TS] forecast

P

Pacheco, J. M., [R] dstdize


Palma, W., [TS] arima, [TS] arima postestimation, [TS] estat acplot


Pampel, F. C., [R] logit, [R] logit, [R] probit

Pan, N., [R] epitab

Panis, C., [R] mkspline

Pantazis, N., [ME] meglm, [ME] mixed

Paolino, P., [R] betareg

Papke, L. E., [R] fraeag

Parent, E., [BAYES] intro

Park, J. Y., [R] boxcox, [R] margins, [R] nlcov,
[R] predictnl, [R] rocreg postestimation,
intro, [TS] vec, [TS] vecrank
Parks, W. P., [R] exlogistic
Parmar, M. K. B., [PSS] intro, [PSS] power cox,
[ST] stcox, [ST] streg
Parmigiani, G., [BAYES] intro
Parnet, E. T., [R] glm, [ST] stcox
Parzen, E., [R] estat ic, [R] kdensity
Pasquini, J., [R] epitab, [R] vvls
Patel, N. R., [R] exlogistic, [R] exlogistic
postestimation, [R] exposionn, [R] tabulate
twoway
Paterson, L., [ME] meqrologit
Patterson, H. D., [R] pkcross
Patterson, K., [XT] xtnunitroot
Pattitoni, P., [R] betareg
Paul, C., [R] logistic
Paulsen, J., [TS] varsoc, [TS] vec intro
Pawitan, Y., [TE] teffects ra
Pearce, M. S., [R] epitab, [R] logistic
Pearl, J., [BAYES] intro
Pearson, E. S., [BAYES] bayesmh, [R] ci, [R] ttest
Pearson, K., [G-2] graph twoway histogram,
[MV] mds, [MV] measure_option, [MV] pca,
[R] correlate, [R] esize, [R] tabulate twoway
Pechivanoglou, P., [R] betareg
Pedace, R., [R] logit, [R] probit, [R] regress,
[R] regress postestimation diagnostic plots,
[U] 20.25 References
Peen, C., [MV] procrustes
Pellock, I. M., [BAYES] bayesmh
Pendegast, J. F., [XT] xtologlog, [XT] xttgee,
[XT] xtitreg, [XT] xtologit, [XT] xtologitm,
[XT] xtprobit, [XT] xtpoblit, [XT] xtobit
Penfield, R. D., [IRT] dif, [R] esize
Peng, J., [PSS] intro, [PSS] power oneprop
Penrose, R., [M-5] pinn()
Pepe, M. S., [R] roc, [R] roccomp, [R] rocfit,
[R] rocreg, [R] rocreg postestimation,
[R] rocregplot, [R] roctab, [ST] stcrreg
Peracchi, F., [MI] intro substantive, [R] regress,
[R] regress postestimation
Pérez-Amaral, T., [U] 20.25 References
Pérez-Hernández, M. A., [R] kdensity
Pérez-Hoyos, S., [R] lrtest
Pérez-Santiago, M. I., [R] epitab
Pericchi, L. R., [BAYES] intro
Perkins, A. M., [R] ranksum
Perotti, V., [R] heckprobop, [R] heckprob, [R]
oprobit
postestimation, [R] lincom, [R] mlogit,
[R] mprobit, [R] mprobit postestimation,
[R] predictnl, [R] logit, [SEM] example 37g
Perry, H. M., [PSS] intro, [PSS] power repeated
Persson, R., [G-1] graph intro
Pesaran, M. H., [XT] xtsunitroot
Pesaran, F., [R] tabulate twoway
Peterson, B., [R] ologit
Peterson, W. W., [R] Iroc
Petit, S., [D] iced10
Petitclerc, M., [R] kappa
Petkova, E., [R] suest
Peto, J., [ST] stst test
Petrin, A. K., [R] frontier
Pevehouse, J. C. W., [TS] time series, [TS] arima,
Pfeffer, R. I., [R] symmetry
Pfeffermann, D., [ME] mixed
Pflueger, C. E., [R] ivregress postestimation
Phillips, A., [IR] difm
[R] margins, [R] nlcov, [R] predictnl,
[R] regress postestimation time series,
[R] rocreg postestimation, [R] rocregplot,
[R] testnl, [TS] pperron, [TS] vargranger,
[TS] vec intro, [TS] vec, [TS] vecrank,
[TS] Glossary, [XT] xtsunitroot
Piantadosi, S., [P] _robust, [U] 20.25 References
Pickles, A., [ME] me, [ME] mepoisson,
[ME] meqrologit, [ME] meqropoisson,
[ME] mestreg, [MV] cluster dendrogram,
[R] glm, [R] glm, [SEM] Acknowledgments,
[SEM] intro 2, [SEM] example 29g,
[SEM] methods and formulas for gsem,
[TE] teffects multivalued, [XT] xtggee,
[XT] xtrg
Pickup, M., [TS] time series, [TS] arch, [TS] arima,
[TS] vec
Pierce, D. A., [ME] me, [ME] meqrologit,
[ME] meqropoisson, [TS] wntestq
Pierson, R. A., [ME] mixed
Pike, M. C., [PSS] intro, [PSS] power twoproportions,
Pillai, K. C. S., [MV] canon, [MV] manova
Pindyck, R. S., [R] biprobit, [R] heckprob
Pinheiro, J. C., [ME] me, [ME] meglm,
[ME] meqrologit, [ME] meqrologit postestimation,
[ME] meqropoisson, [ME] meqropoisson
postestimation, [ME] mixed, [ME] mixed
postestimation
Pintilie, M., [ST] xtrcreg, [ST] xtrcreg postestimation
Pisati, M., [TS] time series
Pischke, J.-S., [R] ivregress, [R] ivregress
postestimation, [R] qreg, [R] regress,
[TE] stteffects ipw, [TE] stteffects ipwra,
[TE] stteffects postestimation, [TE] stteffects
ra, [TE] stteffects wra, [TE] teffects
intro advanced, [U] 20.25 References
Pitarakis, J.-Y., [TS] vecrank

Solaaga, I., [R] inequality

Song, F., [R] meta

Song, S. H., [ME] mixed

Soon, T. W., [R] qc


Sørensen, T. J., [MV] measure_option

Sorrentino, R., [TS] tsfilter, [TS] tsfilter bw

Sosa-Escudero, W., [XT] xtreg, [XT] xtreg postestimation, [XT] xtrregar

Sotoca, S., [TS] sspace

Sowell, F., [TS] arfima


Sparks, A. T., [SEM] example 41g

Späth, H., [MV] cluster

Spearman, C. E., [MV] factor, [R] icc, [R] spearman

Speed, F. M., [R] margins

Speed, T., [R] diagnostic plots

Spence, I., [G-2] graph pie


Spiegelhalter, D. J., [BAYES] bayesstats ic, [R] brier

Spieldman, R. S., [R] symmetry

Spiessens, B., [ME] me, [ME] meqrlogit postestimation

Spitzer, J. J., [R] boxcox

Spizzichino, F., [BAYES] intro

Sprent, P., [R] ranksum, [R] signrank


Staelin, R., [R] rologit

Stahl, W. A., [D] egen

Stahl, D., [MV] cluster, [MV] cluster stop

Staiger, D. O., [R] ivregress postestimation

Stangl, D. K., [BAYES] intro

Starmer, C. F., [R] vws

Startz, R., [R] ivregress postestimation, [TS] mswitch

Stefanski, L. A., [TE] tfeects aipw


Steiger, J. H., [R] esize

Steiger, W., [R] qreg

Stein, C., [R] bootstrap

Steinberg, L., [IRRT] irt grm

Stephenson, D. B., [MV] pca, [R] brier

Stephenska, K. A., [R] nptrend


Stevens, E. H., [MV] mvtest

Stevenson, R. E., [R] frontier


Stewart, J., [ST] xtable


Stigler, S. M., [R] ameans, [R] ci, [R] correlate,

Stillman, S., [R] ivregress, [R] ivregress postestimation

Stine, R., [R] bootstrap

Stock, J. H., [I] Glossary, [R] areg postestimation,

Stoll, B. J., [R] epitab

Stoll, L., [MI] mi estimate

Stolley, P. D., [R] epitab

Stone, M. H., [IRT] irt

Storger, B. E., [ST] sterreg

Stork, D. G., [MV] cluster, [MV] cluster stop

Stoto, M. A., [R] svtest

Stover, L., [R] signrank

Stower, L., [R] rocreg, [R] rogreg postestimation, [R] rogregplot

Støvring, H., [M-2] pointers

Straathof, B., [D] insobs

Stram, D. O., [ME] me

Street, J. O., [R] rreg

Stroup, W. W., [ME] me

Stryhn, H., [R] epitab, [R] regress

Stuart, A., [R] centile, [R] mean, [R] proportion, [R] qreg, [R] ratio, [R] summarize,

Student, see Gosset, W. S.

Stuetzle, W., [R] sunflower

Sturdivant, R. X., [PSS] intro, [PSS] power mcc,


Sørensen, T. J., [MV] factor postestimation


Speed, F. M., [R] margins

Speed, T., [R] diagnostic plots


Spiegelhalter, D. J., [BAYES] bayesstats ic, [R] brier

Spieldman, R. S., [R] symmetry

Spiessens, B., [ME] me, [ME] meqrlogit postestimation

Spitzer, J. J., [R] boxcox

Spizzichino, F., [BAYES] intro

Sprent, P., [R] ranksum, [R] signrank


Staelin, R., [R] rologit

Stahl, W. A., [D] egen

Stahl, D., [MV] cluster, [MV] cluster stop

Staiger, D. O., [R] ivregress postestimation

Stangl, D. K., [BAYES] intro

Starmer, C. F., [R] vws

Startz, R., [R] ivregress postestimation, [TS] mswitch

Stefanski, L. A., [TE] tfeects aipw


Sturdivant, R. X., continued  
[R] stepwise, [SEM] example 33g,  
[SEM] example 34g, [XT] xtgee  
Sturtz, S., [BAYES] bayesmh  
Suárez, C., [R] heckoprobit, [R] heckprobit  
Sued, M., [TE] teffects intro advanced  
Suen, H. K., [R] icc  
Sullivan, G., [P] _robust, [R] regress,  
[SVY] svy: tabulate twoway  
Summers, G. F., [SEM] example 9  
Summers, R., [XT] xtnitroot  
Sun, W., [MI] intro substantive  
Sussman, S., [ME] me, [ME] meglm, [ME] meologit,  
Sutton, A. J., [R] meta  
Svennerholm, A. M., [R] epitab  
Swagel, P. L., [U] 21.6 Reference  
Swaminathan, H., [IRT] itrt, [IRT] diflogistic  
Swamy, P. A. V. B., [XT] xtivreg, [XT] xtrc,  
[XT] xtreg  
Swanson, S. A., [MI] mi estimate, [MI] mi impute,  
[XT] xtgee  
Swed, F. S., [R] runtest  
Sweeting, T. J., [ST] streg  
Sweetman, O., [R] gmm, [R] inequality  
Swensson, B., [SVY] variance estimation  
Swets, J. A., [R] Iroc  
Sykes, R. C., [IRT] itrt 3pl  
Sylvester, J. J., [M-5] svd()  
Szroeter, J., [R] regress postestimation  

T  

Tabachnick, B. G., [MV] discrim, [MV] discrim lda  
Taka, M. T., [R] pkcross  
Tamhane, A. C., [FN] Statistical functions,  
[PSS] intro, [PSS] power onemean, [R] oneway,  
[R] ztest  
Tamminen, J., [BAYES] intro, [BAYES] bayesmh  
Tan, B. S., [PSS] intro, [PSS] power logrank  
Tan, S. H., [PSS] intro, [PSS] power logrank  
Tan, Z., [TE] teffects intro advanced, [TE] teffects alpw  
Tanimoto, T. T., [MV] measure_option  
Taniuchi, T., [R] kdensity  
Tanner, M. A., [BAYES] intro, [MI] intro substantive,  
[MI] mi impute mvn  
Tanner, W. P., Jr., [R] Iroc  
Tanur, J. M., [R] kwallis  
Tapia, R. A., [R] kdensity  
Tarlov, A. R., [MV] alpha, [MV] factor, [MV] factor postestimation,  
[R] lincom, [R] mlogit,  
[R] mprobit, [R] mprobit postestimation,  
[R] predictnl, [R] slogsit, [SEM] example 37g  
Tarone, R. E., [R] epitab, [ST] sts test  
Tastan, H., [TS] vargranger  
Taub, A. J., [XT] xtreg  

Tauchmann, H., [R] frontier, [R] heckman  
Taylor, C., [R] gllamm, [R] glm, [XT] xtgee,  
[XT] xtrc  
Taylor, H. M., [TS] mswitch  
Taylor, W. E., [XT] xthtaylor  
Teller, A. H., [BAYES] intro, [BAYES] bayesmh  
Teller, E., [BAYES] intro, [BAYES] bayesmh  
ten Berge, J. M. F., [MV] procures  
 ter Bogt, T., [MV] mvtest  
Teräsvirta, T., [TS] mgarch, [TS] mgarch ccc  
Terza, J. V., [R] eipoisson, [TE] etpoisson  
Teukolsky, S. A., [FN] Statistical functions,  
[PI] matrix symeigen, [R] dydx, [R] vwl,  
[TS] arch, [TS] arima  
Thall, P. F., [ME] mepoisson, [ME] meqrpoisson  
Thayer, D. T., [IRT] difmhh  
PH-assumption tests, [ST] stcox postestimation,  
[ST] stcrreg  
Thiele, T. N., [R] summarize  
Thissen, D., [IRT] itrt grm  
Thomas, A., [BAYES] bayesmh  
Thomas, D. C., [ST] sttocc  
Thomas, D. G., [R] epitab  
Thomas, D. R., [SVY] svy: tabulate twoway  
Thompson, B., [MV] canon postestimation, [R] esize,  
[R] regress postestimation  
Thompson, D. J., [TE] teffects intro advanced  
Thompson, J., [BAYES] intro, [BAYES] bayes,  
[BAYES] bayesmh  
Thompson, J. C., [R] diagnostic plots  
Thompson, J. R., [R] kdensity, [R] poisson,  
[ST] stptime  
Thompson, M. L., [R] rocreg  
Thompson, S. G., [ME] me  
Thompson, S. K., [BAYES] intro, [SVY] survey  
Thompson, W. A., Jr., [ME] me, [ME] mixed  
Thoms, J., [BAYES] bayesmh  
Thornridge, F., [R] poisson  
Thurstone, L. L., [MV] rotate, [R] rologit  
Tibshirani, R. J., [MV] discrim knn, [R] bootstrap,  
[R] qreg  
Tidmarsh, A. B., [R] fp  
Tierney, L., [BAYES] intro, [ME] me, [ME] meqrllogit,  
[ME] meqrpoisson  
Tilford, J. M., [R] estat gof, [R] Iroc  
Tilling, K., [ME] meqrllogit, [ME] meqrpoisson,  
[ME] mixed, [ST] stcox  
Timms, A., [MV] manova  
Ting Lee, M.-L., [ST] stcox PH-assumption tests  
Tippett, L. H. C., [ST] streg
1.4 References

Tukey, J. W., continued
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
[ST] regress
test diagnostic plots,
<table>
<thead>
<tr>
<th>Author</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van der Linden, W. J.</td>
<td>[IRT] i, [SEM] example 28g, [SEM] example 29g</td>
</tr>
<tr>
<td>Van der Merwe, C. A.</td>
<td>[MV] mvtest, [MV] mvtest means</td>
</tr>
<tr>
<td>Van der Reyden, D.</td>
<td>[R] ranksum</td>
</tr>
<tr>
<td>van Doorslaer, E.</td>
<td>[SVY] svy estimation, [SVY] svyset</td>
</tr>
<tr>
<td>Van Mechelen, I.</td>
<td>[MI] intro substantive, [MI] mi impute</td>
</tr>
<tr>
<td>Van Pragg, B. M. S.</td>
<td>[R] biprobit, [R] heckoprobit, [R] heckprobit</td>
</tr>
<tr>
<td>Varadharajan-Krishnakumar, J.</td>
<td>[XT] xivreg</td>
</tr>
<tr>
<td>Vella, F.</td>
<td>[ME] me, [TE] etregress</td>
</tr>
<tr>
<td>Velleman, P. F.</td>
<td>[R] regress postestimation, [R] smooth</td>
</tr>
<tr>
<td>Venables, W.</td>
<td>[R] esize</td>
</tr>
<tr>
<td>Verbeek, M.</td>
<td>[ME] me, [TE] etregress</td>
</tr>
<tr>
<td>Verdinelli, I.</td>
<td>[BAYES] intro</td>
</tr>
<tr>
<td>Verdurnen, J.</td>
<td>[MV] mvtest</td>
</tr>
<tr>
<td>Verkuilen, J.</td>
<td>[R] betarreg</td>
</tr>
<tr>
<td>Vermandele, C.</td>
<td>[R] summarize</td>
</tr>
<tr>
<td>Vick, R.</td>
<td>[R] mlexp</td>
</tr>
<tr>
<td>Vidakovic, B.</td>
<td>[BAYES] intro</td>
</tr>
<tr>
<td>Vidmar, S.</td>
<td>[R] means, [R] epitab</td>
</tr>
<tr>
<td>Vigfusson, R. J.</td>
<td>[TS] forecast solve</td>
</tr>
<tr>
<td>Vinten-Johansen, P.</td>
<td>[R] epitab</td>
</tr>
<tr>
<td>Vohr, B. R.</td>
<td>[R] rocreg, [R] rocreg postestimation, [R] rocregplot</td>
</tr>
<tr>
<td>Vollebergh, W. A. M.</td>
<td>[MV] mvtest</td>
</tr>
<tr>
<td>von Bortkiewicz, L.</td>
<td>[R] poisson</td>
</tr>
<tr>
<td>von Eye, A.</td>
<td>[R] correlate</td>
</tr>
<tr>
<td>von Neumann, J.</td>
<td>[BAYES] intro</td>
</tr>
<tr>
<td>Von Storch, H.</td>
<td>[R] brier</td>
</tr>
<tr>
<td>Vondrácek, J.</td>
<td>[R] correlate</td>
</tr>
<tr>
<td>Vuong, Q. H.</td>
<td>[R] ivprobit, [R] zinb, [R] zip</td>
</tr>
</tbody>
</table>

**W**

<table>
<thead>
<tr>
<th>Author</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wacholder, S.</td>
<td>[R] binreg</td>
</tr>
<tr>
<td>Wagner, H. M.</td>
<td>[R] qreg</td>
</tr>
<tr>
<td>Wagner, M.</td>
<td>[XT] xtunitroot</td>
</tr>
<tr>
<td>Wagner, T.</td>
<td>[MV] mvtest</td>
</tr>
<tr>
<td>Wagstaff, A.</td>
<td>[SVY] svy estimation, [SVY] svyset</td>
</tr>
<tr>
<td>Vandenbroucke, P.</td>
<td>[R] kwallsis</td>
</tr>
<tr>
<td>Wald, A.</td>
<td>[TS] varwle</td>
</tr>
<tr>
<td>Walker, A. M.</td>
<td>[R] epitab</td>
</tr>
<tr>
<td>Walker, S.</td>
<td>[ST] sts test</td>
</tr>
<tr>
<td>Wallgren, A.</td>
<td>[G-1] graph intro</td>
</tr>
<tr>
<td>Wallgren, B.</td>
<td>[G-1] graph intro</td>
</tr>
<tr>
<td>Wallis, W. A.</td>
<td>[R] kwallsis</td>
</tr>
<tr>
<td>Walstrum, T.</td>
<td>[TE] etregress</td>
</tr>
<tr>
<td>Walters, S. J.</td>
<td>[R] ci, [R] kappa, [R] tabulate twoway</td>
</tr>
<tr>
<td>Wang, H.-J.</td>
<td>[R] frontier, [XT] xtfrontier</td>
</tr>
<tr>
<td>Wang, J.-L.</td>
<td>[ST] sts graph</td>
</tr>
<tr>
<td>Wang, J. W.</td>
<td>[ST] streg</td>
</tr>
<tr>
<td>Wang, Q.</td>
<td>[R] ivregress, [TS] arima, [TS] newey</td>
</tr>
<tr>
<td>Wang, S.</td>
<td>[R] ivregress postestimation</td>
</tr>
<tr>
<td>Wang, Y.</td>
<td>[R] asmprob</td>
</tr>
<tr>
<td>Ward, J. H. Jr.</td>
<td>[MV] cluster, [MV] cluster linkage</td>
</tr>
<tr>
<td>Warren, K.</td>
<td>[R] epitab</td>
</tr>
<tr>
<td>Waterson, E. J.</td>
<td>[R] binreg</td>
</tr>
</tbody>
</table>
Williams, B., [SVY] survey
Williams, B. K., [MV] discrim lda
Williams, G. W., [PSS] intro
Williams, T. O., Jr., [SEM] example 2
Williams, W. T., [MV] cluster
Wilson, D. B., [BAYES] intro
Wilson, E. B., [MV] mvtest normality, [R] ci
Wilson, S. R., [R] bootstrap
Wingood, G. M., [R] nbreg, [R] poisson
Winkelmann, R., [ME] menbreg, [R] cpoisson
Winsten, C. B., [TS] prais
Wis, M., [MV] mds, [MV] mdslong, [MV] mdsmat
Wittes, J., [PSS] intro
Wolfe, F., [D] ds, [R] correlate, [R] spearman
Wolfe, R., [R] ologit, [R] oprobit, [R] tabulate twoway
Wolfinger, R. D., [ME] me
Wolfowitz, J., [TS] varwle
Wolfson, C., [R] kappa
Wolk, A., [R] epitab
Wolkewitz, M., [D] icd10
Wolpert, R. L., [BAYES] intro, [BAYES] intro
Wolpin, K. I., [R] asprobit
Wong, S. P., [R] icc
Wood, F. S., [R] diagnostic plots
Woodard, D. E., [MV] manova, [R] contrast
Woodward, M., [R] epitab
Wooldridge, J. M., continued
Woolf, B., [R] epitab
Woolson, R. F., [PSS] power cmh
Working, H., [R] rocomp, [R] rocfit, [R] roctab
World Health Organization, [D] icd, [D] icd10
World Health Organization Mortality Data Base (Cause of Death Query online; accessed December 11, 2014), [D] icd10
Wretman, J., [SVY] variance estimation
Wright, B. D., [IRT] irt
Wright, D. B., [SEM] example 41g
Wright, J. H., [R] ivregress, [R] ivregress postestimation, [XT] xthtaylor
Wright, J. T., [R] binreg
Wright, J. T., Jr, [PSS] intro, [PSS] power repeated
Wright, P. G., [R] ivregress
Wu, A. W., [IRT] irt
Wu, D.-M., [R] ivregress postestimation
Wu, N., [R] ivregress, [TS] arima, [TS] newey
Wu, P. X., [XT] xreg
Wu, S., [XT] xtunitroot
Wynn, A. H. A., [BAYES] bayesmh

X
Xiao, T., [ST] stcox PH-assumption tests
Xie, Y., [R] logit, [R] probit
Xu, Y., [ST] stcox

Y
Yang, K., [MV] mds
Yang, M., [ME] me
Yang, Z., [R] poisson
Zabell, S. L., [R] kwallis
Zakoian, J. M., [TS] arch
Zamora, M., [R] heckoprobit, [R] heckprobit
Zappasodi, P., [MV] manova
Zavoina, W., [R] ologit
Zeh, J., [D] egen
Zelen, M., [R] ttest, [R] ztest
Zell, E. R., [D] icd10
Zelterman, D., [R] tabulate twoway
Zeng, D., [TS] mswitch
Zhao, H., [ME] mecloglog, [ME] melogit, [ME] meprobit
Zimmerman, F., [R] regress
Zirkler, B., [MV] mvtest, [MV] mvtest normality
Zlotnik, A., [R] logit postestimation
Zubin, J., [MV] measure_option
### Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>!</td>
<td>(not), see logical operators</td>
</tr>
<tr>
<td>!=</td>
<td>(not equal), see relational operators</td>
</tr>
<tr>
<td>alpha</td>
<td>[PSS] Glossary</td>
</tr>
<tr>
<td>beta</td>
<td>[PSS] Glossary</td>
</tr>
<tr>
<td>delta</td>
<td>[PSS] Glossary</td>
</tr>
<tr>
<td>&amp;</td>
<td>(and), see logical operators</td>
</tr>
<tr>
<td>*</td>
<td>abbreviation character, see abbreviations</td>
</tr>
<tr>
<td>*</td>
<td>comment indicator, [P] comments</td>
</tr>
<tr>
<td>~</td>
<td>(not), see logical operators</td>
</tr>
<tr>
<td>=&gt;</td>
<td>(less than or equal), see relational operators</td>
</tr>
<tr>
<td>==</td>
<td>(equality), see relational operators</td>
</tr>
<tr>
<td>&gt;</td>
<td>(greater than), see relational operators</td>
</tr>
<tr>
<td>&gt;=</td>
<td>(greater than or equal), see relational operators</td>
</tr>
<tr>
<td>?</td>
<td>(or), see logical operators</td>
</tr>
<tr>
<td>~</td>
<td>(not), see logical operators</td>
</tr>
<tr>
<td>^=</td>
<td>(not equal), see relational operators</td>
</tr>
<tr>
<td>100% sample</td>
<td>[SVY] Glossary</td>
</tr>
<tr>
<td>1:M matched design</td>
<td>[PSS] power mcc, [PSS] Glossary</td>
</tr>
<tr>
<td>1pl, irt subcommand</td>
<td>[IRT] 1pl, [IRT] irt 1pl postestimation</td>
</tr>
<tr>
<td>2x2 K contingency table</td>
<td>[PSS] Glossary</td>
</tr>
<tr>
<td>2x2 contingency table</td>
<td>[PSS] Glossary</td>
</tr>
<tr>
<td>2PL, see two-parameter logistic model</td>
<td></td>
</tr>
<tr>
<td>2pl, irt subcommand</td>
<td>[IRT] irt 2pl, [IRT] irt 2pl postestimation</td>
</tr>
<tr>
<td>3PL, see three-parameter logistic model</td>
<td></td>
</tr>
<tr>
<td>3pl, irt subcommand</td>
<td>[IRT] irt 3pl, [IRT] irt 3pl postestimation</td>
</tr>
</tbody>
</table>

### Abbreviations

addgroup, ssd subcommand, [SEM] ssd
adding
  fits, see fits, adding
  lines, see lines, adding
text, see text, adding
addition across
  observations, [D] egen
variables, [D] egen
addition operator, see arithmetic operators
addplot() option, [G-3] addplot_option
ADF, see asymptotic distribution free
adf, see sem option method()
adjust, forecast subcommand, [TS] forecast adjust
adjusted
  Kaplan–Meier survivor function, [ST] sts
  means, [R] margins, [R] marginsplot
  partial residual plot, [R] regress postestimation
diagnostic plots
administrative censoring, [PSS] power cox,
  [PSS] power exponential, [PSS] power logrank,
  [PSS] Glossary
ado
  command, [R] net
describe command, [R] net
dir command, [R] net
uninstall command, [R] net
ado.d, view subcommand, [R] view
.ado file, [U] 11.6 Filenaming conventions
ado,
clear subcommand, [D] clear
dir subcommand, [R] view
  [U] 3.4 The Stata Journal, [U] 17 Ado-files,
  [U] 18.11 Ado-files
adding comments to, [P] comments
debugging, [P] trace
downloading, see files, downloading
editing, [R] doedit
installing, [R] net, [R] sj, [R] ssc, [U] 17.6 How do I install an addition?
location, [U] 17.5 Where does Stata look for ado-files?
location of, [R] which
long lines, [P] #delimit, [U] 18.11.2 Comments and long lines in ado-files
official, [R] update, [U] 28 Using the Internet to keep up to date
searching for, [R] search, [R] ssc
updating user-written, [R] adoupdate
adopath
  + command, [P] syssdir
  ++ command, [P] syssdir
  - command, [P] syssdir
command, [P] syssdir, [U] 17.5 Where does Stata look for ado-files?
ado-path, [M-5] adosubdir()
adosize, set subcommand, [P] syssdir, [R] set,
  [U] 18.11 Ado-files
adosubdir macro extended function, [P] macro
adosubdir() function, [M-5] adosubdir()
adoupdate command, [R] adoupdate
AFT, see accelerated failure-time model
agglomerative hierarchical clustering methods,
aggregate
  functions, [D] egen
  statistics, dataset of, [D] collapse
  agreement, interrater, [R] kappa
AIC, see Akaike information criterion
AIPW, see augmented inverse-probability weighting
aipw,ffects subcommand, [TE] tffects aipw
Akaike information criterion, [BAYES] bayesstats ic,
  [BAYES] Glossary, [R] BIC note, [R] estat,
  [R] estat ic, [R] estimates stats, [R] glm,
  [R] lrtest, [SEM] estat gof, [SEM] example 4,
  [SEM] methods and formulas for sem,
  [ST] streg
algebraic expressions, functions, and operators,
  [P] matrix define, [U] 13 Functions and expressions
alignment of text, [G-3] textbox_options
alignmentstyle, [G-4] alignmentstyle
_all, [U] 11.1.1 varlist
_all,
clear subcommand, [D] clear
  update subcommand, [R] update
_all() function, [M-5] all()
al macro extended function, [P] macro
allmissing option, see sem option allmissing
allocation ratio, [PSS] power twomeans, [PSS] power twoproportions, [PSS] power twovariances,
  [PSS] power twocorrelations, [PSS] power cmh,
  [PSS] power exponential, [PSS] power logrank,
  [PSS] unbalanced designs, [PSS] Glossary
allof() function, [M-5] all()
alpha, see α
alpha coefficient, Cronbach’s, [MV] alpha
alpha command, [MV] alpha
alphabetizing
  observations, [D] gsort, [D] sort
  variable names, [D] order
  variables, [D] sort
  alphanumeric variables, see string variables, parsing, see string variables
alternative
correlation, [PSS] power, [PSS] power onecorrelation
  hypothesis, [PSS] Glossary, also see null hypothesis
  and alternative hypothesis
mean, [PSS] power, [PSS] power onemean,
  [PSS] unbalanced designs
alternative, continued
mean difference, [PSS] power, [PSS] power pairedmeans
parameter, [PSS] Glossary, also see alternative value proportion, [PSS] power, [PSS] power oneproportion
scenarios, [TS] forecast, [TS] forecast adjust,
[TS] forecast clear, [TS] forecast coeftest, [TS] forecast create,
[TS] forecast solve
standard deviation, [PSS] power, [PSS] power onevariance
value, [PSS] Glossary, also see postulated value variance, [PSS] power, [PSS] power onevariance
alternatives, estat subcommand, [R] asclogit postestimation, [R] asmprobit postestimation,
[R] asroprobit postestimation, [R] nlogit postestimation
alternative-specific
conditional logit (McFadden’s choice) model, [R] asclogit
multinomial probit regression, [R] asmprobit
rank-ordered probit regression, [R] asroprobit
ameans command, [R] ameans
American Standard Code for Information Interchange, see ASCII
analysis of covariance, [R] anova
analysis of variance, [PSS] power, [PSS] power oneway, [PSS] power twoway, [PSS] Glossary,
[R] anova, [R] contrast, [R] psm, [R] one-way,
[R] one-way Kruskal–Wallis, [R] kwallis
plots, [R] marginsplot
repeated measures, [PSS] power repeated,
[R] anova
analysis step, [MI] intro substantive, [MI] mi estimate, also see estimation
analysis-of-variance test of normality, [R] swilk
analytic weight, [U] 11.1.6 weight,
[U] 20.23.2 Analytic weights
anchoring, see constraints, normalization
ANCOVA, see analysis of covariance
and operator, [U] 13.2.4 Logical operators
Anderberg coefficient similarity measure,
[MV] measure_option
angle of text, [G-4] anglestyle
anglestyle, [G-4] anglestyle
angular similarity measure, [MV] measure_option
ANOVA, see analysis of variance
anova command, [R] anova, [R] anova postestimation
ANOVA DDF, see denominator degrees of freedom, ANOVA
anova, estat subcommand, [MV] discrim lda postestimation
Anscombe residual, [ME] mecloglog postestimation,
[ME] meglm postestimation, [ME] melogit postestimation, [ME] mnbreg postestimation,
[ME] mepoisson postestimation, [ME] meprobit postestimation, [ME] meqlogit postestimation,
[ME] meqrpoisson postestimation,
[ME] mestreg postestimation
anti, estat subcommand, [MV] factor postestimation
anti-image
correlation matrix, [MV] factor postestimation,
[MV] pca postestimation, [MV] Glossary
covariance matrix, [MV] factor postestimation
[MV] pca postestimation, [MV] Glossary
any() function, [M-5] all()
anycount(), egen function, [D] egen
anymatch(), egen function, [D] egen
anyof() function, [M-5] all()
anyvalue(), egen function, [D] egen
A-PARCH, see asymmetric power autoregressive conditional heteroskedasticity
append command, [D] append, [U] 22 Combining datasets
append, mi subcommand, [MI] mi append
append variable, [D] append
appending data, [D] append, [MI] mi append,
[U] 22 Combining datasets
appending rows and columns to matrix, [P] matrix define
apply recording, [G-2] graph play
approximating Euclidean distances, [MV] mds postestimation
AR, see autoregressive
arbitrary pattern of missing values, [MI] mi impute chained, [MI] mi impute mvn, [MI] Glossary, also see pattern of missingness
arccosine, arcsine, and arctangent functions, [FN] Trigonometric functions
ARCH, see autoregressive conditional heteroskedasticity
arch command, [TS] arch, [TS] arch postestimation
archlm, estat subcommand, [R] regress postestimation time series
area, graph twoway subcommand, [G-2] graph twoway area
area under the curve, [R] lroc, also see pharmacokinetic data, also see receiver operating characteristic analysis
areas, [G-4] colorstyle, also see fill, areas, dimming and brightening, also see fill, color, setting
areaestyle, [G-4] areaestyle
areg command, [R] areg, [R] areg postestimation
Arellano–Bover estimator, [XT] xtreg, [XT] xtgedsys
ARFIMA, see autoregressive fractionally integrated moving-average model
arfima command, [TS] arfima, [TS] arfima postestimation
arg() function, [M-5] sin()
args command, [P] syntax
args() function, [M-5] args()
arguments,
  values returned in, [M-1] returnedargs
ARIMA, see autoregressive integrated moving-average model
arima command, [TS] arima, [TS] arima postestimation
arithmetic operators, [M-2] op_arith, [M-2] op_colon,
  [P] matrix define, [U] 13.2.1 Arithmetic operators
ARMA, see autoregressive moving average
ARMAX, see autoregressive moving average with exogenous inputs
aroots, estat subcommand, [TS] estat aroots
array, [M-6] Glossary
arrays, class, [P] class
  .Arrdropall built-in class modifier, [P] class
  .Arrdropel built-in class modifier, [P] class
  .arrindexof built-in class function, [P] class
  .arrnels built-in class function, [P] class
arrays, [G-2] graph twoway parrow
  .Arrpop built-in class modifier, [P] class
  .Arrpush built-in class modifier, [P] class
as error, display directive, [P] display
as input, display directive, [P] display
as result, display directive, [P] display
as text, display directive, [P] display
as txt, display directive, [P] display
asarray() function, [M-5] asarray()
asarray_contains() function, [M-5] asarray()
asarray_contents() function, [M-5] asarray()
asarray_create() function, [M-5] asarray()
asarray_elements() function, [M-5] asarray()
asarray_first() function, [M-5] asarray()
asarray_key() function, [M-5] asarray()
asarray_keys() function, [M-5] asarray()
asarray_next() function, [M-5] asarray()
asarray_notfound() function, [M-5] asarray()
asarray_remove() function, [M-5] asarray()
ASCII, [D] unicode, [D] unicode translate,
  [1] Glossary
  codes, [M-5] asciic()
  encoding conversion, [D] unicode convertfile, [D] unicode translate
  asciic() function, [M-5] asciic()
asclogit command, [R] asclogit, [R] asclogit postestimation
asin() function, [FN] Trigonometric functions,
  [M-5] sin()
asinh() function, [FN] Trigonometric functions,
  [M-5] sin()
  _asis, display directive, [P] display
as is print color mapping, [G-2] set printcolor
asmprob() function, [M-5] asmprob()
autocorrelation, continued
[TS] ucm, [TS] var, [TS] varlm,
[TS] Glossary, also see HAC variance estimate
dynamic model, [XT] xtabond, [XT] xtdpd,
[XT] xtdpdsys
residual, [XT] xtggee, [XT] xtglas, [XT] xtpcse,
[XT] xtregr
test, [XT] xtabond, [XT] xtabond postestimation,
[XT] xtdp postestimation, [XT] xtdpdsys,
[XT] xtdpdsys postestimation
autocovariance, [TS] arfima, [TS] arima,
[TS] corgram, [TS] estat acplot, [TS] psdensity
automatic print color mapping, [G-2] set printcolor
Automation, [P] automation
autoregressive, [TS] arch, [TS] arfima, [TS] arima,
[TS] dfactor, [TS] sspace, [TS] ucm
conditional heteroskedasticity
effects, [TS] arch
effects, testing for, [R] regress postestimation
time series
model, [TS] arch, [TS] arch postestimation,
[TS] Glossary, also see multivariate GARCH model
test, [R] regress postestimation time series
fractionally integrated moving-average model,
[TS] arfima, [TS] arfima postestimation,
integrated moving-average model, [TS] arima,
[TS] arima postestimation, [TS] estat acplot,
model, [TS] dfactor, [TS] estat acplot,
[TS] mswitch, [TS] psdensity, [TS] sspace,
[TS] ucm
moving average, [TS] arch, [TS] arfima,
[TS] arima, [TS] sspace, [TS] ucm,
[TS] Glossary
moving average with exogenous inputs, [TS] arfima,
[TS] arima, [TS] dfactor, [TS] sspace,
[TS] ucm, [TS] Glossary
process, [TS] Glossary, [XT] xtabond, [XT] xtdpd,
[XT] xtdpdsys, [XT] Glossary
autotabgraphs, set subcommand, [R] set
available area, [G-3] region_options
available-case analysis, [MI] intro substantive
average
marginal effects, [R] margins, [R] marginsplot
partial effects (APEs), [R] margins, [R] marginsplot
predictions, [R] margins, [R] marginsplot
RVI, [MI] mi estimate, [MI] Glossary
treatment effect, [TE] teffects, [TE] teffects intro,
[TE] teffects intro advanced, [TE] teffects aipw,
[TE] teffects ipw, [TE] teffects ipwra,
[TE] teffects multivalued, [TE] teffects
mmatch, [TE] teffects psmatch, [TE] teffects
ra, [TE] Glossary
comparing, [TE] teffects intro advanced
survival time, [TE] stteffects intro,
[TE] stteffects ipw, [TE] stteffects ipwra,
[TE] stteffects ra, [TE] stteffects wra
average, continued
treatment effect on treated, [TE] teffects intro,
[TE] teffects intro advanced, [TE] teffects ipw,
[TE] teffects ipwra, [TE] teffects multivalued,
[TE] teffects nmatch, [TE] teffects psmatch,
comparing, [TE] teffects intro advanced
survival time, [TE] stteffects intro,
[TE] stteffects ipw, [TE] stteffects ipwra,
[TE] stteffects ra, [TE] stteffects wra
averagelinkage,
cluster mat subcommand, [MV] cluster linkage
cluster subcommand, [MV] cluster linkage
average-linkage clustering, [MV] cluster,
[MV] clustermat, [MV] cluster linkage,
[MV] Glossary
averages, see means
avplot and avplots commands, [R] regress
postestimation diagnostic plots
[aweight=exp] modifier, [U] 11.1.6 weight,
[U] 20.23.2 Analytic weights
axes
multiple scales, [G-3] axis_choice_options
setting offset between and plot region,
[G-3] region_options
suppressing, [G-3] axis_scale_options
axis
labeling, [G-3] axis_label_options,
[G-3] axis_options
line, look of, [G-3] axis_scale_options,
[G-3] cat_axis_label_options,
[G-3] cat_axis_line_options
log, [G-3] axis_scale_options
overall look, [G-4] axisstyle
range, [G-3] axis_scale_options
reversed, [G-3] axis_scale_options
scale, [G-3] axis_options,
[G-3] axis_scale_options
selection of, [G-3] axis_choice_options
suppressing, [G-3] axis_scale_options
ticking, [G-3] axis_label_options
tiling, [G-3] axis_options,
[G-3] axis_title_options
suppressing, [G-3] axis_title_options
axisstyle, [G-4] axisstyle

B

b [] , [U] 13.5 Accessing coefficients and standard errors
btitle() option, [G-3] title_options
b2title() option, [G-3] title_options
backed up message, [R] maximize
background color, [G-4] schemes intro
setting, [G-3] region_options
balance
standardized differences, [TE] tebalance summarize
variance ratios, [TE] tebalance summarize
balance, tebalance subcommand, [TE] tebalance
box, [TE] tebalance density, [TE] tebalance
overid, [TE] tebalance summarize
balanced
data, [XT] Glossary
design, [PSS] power twomeans, [PSS] power
twoproportions, [PSS] power
twovariances, [PSS] power
twocorrelations, [PSS] power
oneway, [PSS] power
twoway, [PSS] power
repeated, [PSS] power
cmh, [PSS] power
trend, [PSS] power
exponential, [PSS] power
logrank, [PSS] unbalanced
designs, [PSS] Glossary
repeated replication, [SVY] brr_options,
[Svy] svy br, [SVY] variance estimation,
[SVY] Glossary
repeated replication standard errors, [SVY] svy br,
[SVY] variance estimation
band-pass filters, [TS] tsfilter bk, [TS] tsfilter cf
band-pass filters, [TS] Glossary
bar
graph subcommand, [G-2] graph bar
twoway subcommand, [G-2] graph twoway
bar
bar charts, [G-2] graph bar
twoway parrow
labeling, [G-3] blabel_option
look of, [G-3] barlook_options
Bartlett scoring, [MV] factor postestimation
Bartlett’s
bands, [TS] corrgram
periodogram test, [TS] wntestb
test for equal variances, [R] oneway
base
conversion, [M-5] inbase()
level, [U] 11.4.3 Factor variables
plottypes, [G-3] advanced_options
base, fvset subcommand, [R] fvset
BASE directory, [P] sysdir, [U] 17.5 Where does Stata
look for ado-files?
baseline, [ST] Glossary
comparisons, [SEM] estat gof, [SEM] example 4
dataset, [ST] stbase
hazard and survivor functions, [ST] stcox, [ST] stcox
PH-assumption tests, [ST] stcrreg
model, [SEM] estat gof, [SEM] example 4,
[SEM] methods and formulas for sem,
[SEM] Glossary
baseline suboption, [G-4] alignmentstyle
baseopts option, see sem option baseopts()
basis, orthonormal, [P] matrix svd
batch means, [BAYES] bayesmh, [BAYES] bayesstats
summary, [BAYES] Glossary
Battese–Coelli parameterization, [XT] xfrontier
Bayes factor, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesstats ic,
[BAYES] Glossary
Bayes’s rule, [BAYES] intro, [BAYES] bayesmh,
[BAYES] Glossary, see Bayes’s theorem
Bayes’s theorem, [MV] Glossary
bayesgraph
command, [BAYES] bayesgraph
matrix command, [BAYES] bayesgraph
Bayesian
analysis, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesmh
postestimation, [BAYES] bayesgraph,
[BAYES] bayesstas, [BAYES] bayesstas ess,
[BAYES] bayesstas ic, [BAYES] bayesstas
summary, [BAYES] bayestest,
[BAYES] bayestest interval, [BAYES] bayestest
analysis
concepts, [BAYES] intro, [BAYES] bayesmh,
[MII intro substantive
estimation, [BAYES] bayes, [BAYES] bayesmh,
[BAYES] bayesmh evaluators,
[BAYES] bayesstas ic
feasible initial values, [BAYES] bayesmh,
[BAYES] bayesmh evaluators,
[BAYES] Glossary
user-written evaluators, [BAYES] bayesmh
evaluators
graphical summaries, [BAYES] intro,
[BAYES] bayes, [BAYES] bayesmh,
[BAYES] bayesmh postestimation,
[BAYES] bayesgraph
hypothesis testing, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh postestimation,
[BAYES] bayestest, [BAYES] Glossary
interval, [BAYES] intro, [BAYES] bayesmh
postestimation, [BAYES] bayestest interval
model, [BAYES] intro, [BAYES] bayesmh
postestimation, [BAYES] bayestest model
information criterion, [BAYES] intro,
[BAYES] bayesmh, [BAYES] bayesstats ic,
[BAYES] Glossary, [R] BIC note, [R] estat,
[R] estat ic, [R] estimates stats, [R] glm,
[R] lrtest, [SEM] estat gof, [SEM] example 4,
[SEM] methods and formulas for sem
model comparison, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesmh
postestimation, [BAYES] bayesstas,
[BAYES] bayesstas ic, [BAYES] bayestest,
model parameters, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesmh
evaluators, [BAYES] bayesmh postestimation,
[BAYES] bayesstas, [BAYES] Glossary
Bayesian, continued

bayesmh command, [BAYES] bayesmh
bayestest
command, [BAYES] bayestest
ess command, [BAYES] bayestest
ic command, [BAYES] bayestest
summary command, [BAYES] bayestest
bayesstats
command, [BAYES] bayesstats
ess command, [BAYES] bayesstats
ic command, [BAYES] bayesstats
summary command, [BAYES] bayesstats
bayestest
interval command, [BAYES] bayestest interval
model command, [BAYES] bayestest model
cbal
check command, [D] bcal
create command, [D] bcal
describe command, [D] bcal
dir command, [D] bcal
load command, [D] bcal
BCC, see boundary characteristic curve bcskew0 command, [R] inskew0
Bentler–Raykov squared multiple-correlation coefficient, [SEM] estat eggof
Bentler’s invariant pattern simplicity rotation, [MV] rotatemat, [MV] Glossary
beta coefficients, [R] regress
density,
central, [FN] Statistical functions
noncentral, [FN] Statistical functions
distribution,
cumulative, [FN] Statistical functions
cumulative noncentral, [FN] Statistical functions
inverse cumulative, [FN] Statistical functions
inverse cumulative noncentral, [FN] Statistical functions
inverse reverse cumulative, [FN] Statistical functions
reverse cumulative, [FN] Statistical functions
complement to incomplete, [FN] Statistical functions
beta function, continued
incomplete, [FN] Statistical functions,
[M-5] normal()
regression, [R] betareg, [SVY] svy estimation,
[U] 26.10 Fractional-outcome dependent-variable models
betaden() function, [FN] Statistical functions,
[M-5] normal()
betareg command, [R] betareg, [R] betareg postestimation
between estimators, [XT] xtivreg, [XT] xreg,
[XT] Glossary
between matrix, [MV] Glossary
between–within DDF, see denominator degrees of freedom, repeated
between-cell means and variances, [XT] xtdescribe, [XT] xsum
between-group variance, [PSS] power oneway
between-imputation variability, [MI] mi estimate,
[MI] mi predict
between-subjects
design, [PSS] power oneway, [PSS] power twoway,
[PSS] power repeated, [PSS] Glossary
factor, [PSS] power repeated, [PSS] Glossary
variance, [PSS] power repeated
BFGS algorithm, see Broyden–Fletcher–Goldfarb–Shanno algorithm
bgodfrey, estat subcommand, [R] regress postestimation time series
BHHH algorithm, see Berndt–Hall–Hall–Hausman algorithm
bias corrected and accelerated, [R] bootstrap postestimation, [R] bstat
BIC, see Bayesian information criterion
Bickenböller test statistic, [R] symmetry
bin() option, [G-2] graph twoway histogram
binary
files, writing and reading, [P] file
I/O, [M-5] bufio( )
item, [IRT] Glossary
operator, [M-6] Glossary
outcome, [PSS] power
outcomes, [SEM] intro 5,
[SEM] example 27g, [SEM] example 28g,
[SEM] example 29g, [SEM] example 30g,
[SEM] example 31g, [SEM] example 32g,
[SEM] example 33g, [SEM] example 34g, see
subject index 191
string, [I] Glossary
variable imputation, see imputation, binary
binary 0, [I] Glossary
binomial
distribution,
confidence intervals, [R] ci
cumulative, [FN] Statistical functions
inverse cumulative, [FN] Statistical functions
inverse reverse cumulative, [FN] Statistical functions
biquartimin rotation, [MV] rotate, [MV] rotatemat
bisection method, see iteration, bisection method
bittest and bittesti commands, [R] btest
bivariate normal function, [FN] Statistical functions
bivariate probit regression, [R] biprobit
bivariate probit regression, [FN] Statistical functions
binormal() function, [FN] Statistical functions, [M-5] normal()
binormalp() function, [FN] Statistical functions, [M-5] normal()
binormalaltail() function, [FN] Statistical functions, [M-5] normal()
binormal() function, [FN] Statistical functions, [M-5] normal()
binreg command, [R] binreg, [R] binreg postestimation
bioequivalence test, [BAYES] bayesmh, [R] pk, [R] pkequiv
biopharmaceutical data, see pharmacokinetic data
biplot command, [MV] biplot
biprobit command, [R] biprobit, [R] biprobit postestimation
biquartimin rotation, [MV] rotate, [MV] rotatemat
bisection method, see iteration, bisection method
bittest and bittesti commands, [R] btest
bivariate normal function, [FN] Statistical functions
bivariate probit regression, [R] biprobit, [SVY] svy estimation
bias weight regression estimates, [R] rreg
biyearly() function, [U] 25 Working with categorical data and factor variables
bk, tsfilter subcommand, [TS] tsfilter bk
blanks, removing from strings, [FN] String functions
block
  diagonal covariance, [MV] mvtest covariances
diagonal matrix, [M-5] blockdiag()
exogeneity, [TS] vargranger
blog, see Stata Blog
Blundell–Bond estimator, [XT] xtder, [XT] xtdpdysys
bofed() function, [D] datetime business calendars, [FN] Date and time functions, [M-5] date()
bold font, [G-4] text
Bonferroni’s multiple-comparison adjustment, see multiple comparisons, Bonferroni’s method
bootstrap, [SEM] Glossary
bootstrap_options, [SVY] bootstrap_options
bootstrap prefix command, [R] bootstrap, [R] bootstrap postestimation
bootstrap, estat subcommand, [R] bootstrap postestimation
border around plot region, suppressing, [G-3] region_options
borders
Boston College archive, see Statistical Software Components archive
bottom suboption, [G-4] alignmentstyle
boundary
  characteristic curve, [IRT] rtgraph icc, [IRT] Glossary
kernel, [ST] Glossary
solution, [MV] Glossary
Box–Cox
  power transformations, [R] Inskew0 regression, [R] boxcox
box, graph subcommand, [G-2] graph box
Box M test, [MV] mvtest covariances
box plots, [G-2] graph box
boxcox command, [R] boxcox, [R] boxcox postestimation
Box’s conservative epsilon, [R] anova
break, [M-2] break
break command, [P] break
Break key, [U] 9 The Break key, [U] 16.1.4 Error handling in do-files
breakkey() function, [M-5] setbreakinr()
breakkeyreset() function, [M-5] setbreakinr()
Breitung test, [XT] xtunitroot
breitung, xtunitroot subcommand, [XT] xtunitroot
Breusch–Godfrey test, [R] regress postestimation time series
Breusch–Pagan Lagrange multiplier test, [XT] xtreg postestimation
Breusch–Pagan test, [MV] mvreg, [R] sureg
Breusch–Pagan/Cook–Weisberg test for heteroskedasticity, [R] regress postestimation
brier command, [R] brier
Brier score decomposition, [R] brier
Subject index

...
c(traceexpand) c-class value, [P] creturn, [P] trace
c(tracehilite) c-class value, [P] creturn, [P] trace
c(traceindent) c-class value, [P] creturn, [P] trace
c(tracecnumber) c-class value, [P] creturn, [P] trace
c(tracesep) c-class value, [P] creturn, [P] trace
c(type) c-class value, [D] generate, [P] creturn
c(update_interval) c-class value, [P] creturn
c(update_prompt) c-class value, [P] creturn
c(update_query) c-class value, [P] creturn
c(username) c-class value, [P] creturn
c(userversion) c-class value, [P] creturn
c(varabbrev) c-class value, [P] creturn
c(varkeyboard) c-class value, [P] creturn
c(version) c-class value, [P] creturn, [P] version
c(Wdays) c-class value, [P] creturn
c(Weekdays) c-class value, [P] creturn
c(screen) c-class value, [P] creturn

c apology)

CAT, see correspondence analysis
ca command, [MV] ca, [MV] ca postestimation,
[MV] ca postestimation plots
cabiplot command, [MV] ca postestimation plots
calculator, [R] display
calendars, [D] bcal, [D] datetime business calendars,
[D] datetime business calendars creation,
[TS] intro
calibration, [IRT] Glossary
Califans and Harabasz index stopping rules,
[MV] cluster stop
_caller() pseudofunction, [FN] Programming
functions
callersversion() function, [M-5] callersversion()
camat command, [MV] ca, [MV] ca postestimation,
[MV] ca postestimation plots

canberra dissimilarity measure, [MV] measure_option
candisc command, [MV] candisc, [MV] discrim estat,
[MV] discrim qda postestimation
canon command, [MV] canon, [MV] canon
postestimation

Canonical

correlation analysis, [MV] Glossary
correlations, [MV] canon, [MV] canon
postestimation
discriminant analysis, [MV] candisc, [MV] Glossary
loadings, [MV] canon, [MV] canon postestimation,
[MV] Glossary
variate set, [MV] canon, [MV] canon
postestimation, [MV] Glossary

canon test, estat subcommand, [MV] discrim lda
postestimation
capped spikes, [G-3] rcap_options
caprejection command, [MV] ca postestimation
plots
caption() option, [G-3] title_options
capture command, [P] capture
carryover effects, [R] pk, [R] pkcross, [R] pkshape
case–cohort data, [ST] sttocc

case–control, [PSS] Glossary
data, [R] clogit, [R] epitab, [R] logistic, [R] rocreg,
[R] symmetry, [ST] sttocc
study, [PSS] power, [PSS] power mecc, [R] epitab
caseement displays, [G-3] by_option
casewise deletion, [D] egen, [P] mark, see listwise
deletion
cat command, [D] type
cat() function, [M-5] cat()
categorical, also see factor variables
axis, look of
labels, [G-3] cat_axis_label_options
line, [G-3] cat_axis_line_options
contrasts after anova, [R] contrast
covariates, [R] anova
data, [D] egen, [D] recode, [MV] ca, [MV] manova,
[MV] mca, [R] epitab, [SVY] svy estimation,
[SVY] svy: tabulate oneway,
[SVY] svy: tabulate twoway
data, agreement, measures for, [R] kappa
graphs, [R] grmeanby, [R] spikeplot
item, [IRT] Glossary
outcomes, see outcomes, binary, see outcomes,
categorical, see outcomes, ordinal, also see outcomes,
binary, also see outcomes, ordinal
regression, also see outcomes subentry
absorbing one categorical variable, [R] areg
tabulations, [R] table, [R] tabstat, [R] tabulate
oneway, [R] tabulate twoway, [R] tabulate,
summarize()
variable creation, [R] tabulate oneway, [R] xi
variable imputation, see imputation, categorical
variables, [U] 25.1.2 Converting continuous
variables to categorical variables
category
boundary curve, see boundary characteristic curve
boundary location, [IRT] Glossary
characteristic curve, [IRT] irtgraph icc,
[IRT] Glossary
response function, [IRT] irtgraph icc,
[IRT] Glossary

cause–specific hazard, [ST] Glossary
cc command, [R] epitab
CCC, see category characteristic curve
cce, mgarch subcommand, [TS] mgarch cce
cchart command, [R] qc
cci command, [R] epitab
class command, [P] creturn
CCT, see controlled clinical trial study
CD, see coefficient of determination
cd command, [D] cd
cd, net subcommand, [R] net
Cdumps() function, [D] datetime, [FN] Date and time
functions, [M-5] date()
cd, classutil subcommand, [P] classutil
ceil() function, [FN] Mathematical functions,
[M-5] trunc()
Cholesky decomposition, [M-5] cholesky(), [P] matrix define
ordering, [TS] Glossary
cholesky() function, [M-5] cholesky()
cholesky() function, [FN] Matrix functions,
[M-5] cholesky(), [P] matrix define
_cholinv() function, [M-5] cholinv()
cholinv() function, [M-5] cholinv()
/_cholsolve() function, [M-5] cholsolve()
cholsolve() function, [M-5] cholsolve()
 chute function, [FN] Programming functions
Chow test, [R] anova, [R] contrast, [R] lrtest,
[TS] estat sbsknown
churdle
_command, [R] churdle, [R] churdle postestimation
exponential command, [R] churdle
linear command, [R] churdle
CI, see confidence interval
assumption, see conditional-independence assumption
ci
_means command, [R] ci
_proportions command, [R] ci
_variances command, [R] ci
CIF, see cumulative incidence function
cii
_means command, [R] ci
_proportions command, [R] ci
_variances command, [R] ci
class
definition, [P] class
instance, [P] class
programming utilities, [P] classutil
class, [M-2] class
class exit command, [P] class exit
classes, [M-2] class
classfunctions, estat subcommand, [MV] discrim
lidapostestimation
classical scaling, [MV] Glossary
classification, see cluster analysis, see discriminant analysis
data, see receiver operating characteristic analysis
classification, continued
function, [MV] discrim, [MV] discrim lda,
[MV] discrim lda postestimation, [MV] discrim
qda, [MV] discrim qda postestimation,
[MV] Glossary
intrater agreement, [R] kappa
table, [MV] candisc, [MV] discrim, [MV] discrim
_estat, [MV] discrim knn, [MV] discrim
knn postestimation, [MV] discrim lda,
[MV] discrim lda postestimation, [MV] discrim
logistic, [MV] discrim logistic postestimation,
[MV] discrim qda, [MV] discrim qda
postestimation, [MV] Glossary, [R] estat
classification
classutil
cdir command, [P] classutil
describe command, [P] classutil
dir command, [P] classutil
drop command, [P] classutil
which command, [P] classutil
classwide variable, [P] class
clean,
icd10 subcommand, [D] icd10
icd9 subcommand, [D] icd9
icd9p subcommand, [D] icd9
clear
* command, [D] clear
ado command, [D] clear
all command, [D] clear
command, [D] clear
matrix command, [D] clear
option, [U] 11.2 Abbreviation rules
programs command, [D] clear
results command, [D] clear
clear,
datasignature subcommand, [D] datasignature
ereturn subcommand, [P] ereturn, [P] return
_estimates subcommand, [P] _estimates
estimates subcommand, [R] estimates store
forecast subcommand, [R] return
[TS] forecast clear
fvset subcommand, [R] fvset
mata subcommand, [M-3] mata clear
ml subcommand, [R] ml
postutil subcommand, [P] postfile
putexcel subcommand, [P] putexcel, [P] putexcel
advanced
return subcommand, [P] return
clear, continued
serset subcommand, [P] serset
sreturn subcommand, [P] program, [P] return
timer subcommand, [P] timer
clearing estimation results, [P] ereturn, [P] _estimates,
[R] estimates store
clearing memory, [D] clear
clegend() option, [G-3] clegend_option,
[G-3] legend_options
clevel, set subcommand, [BAYES] set clevel, [R] set
cl臨inal trial, [BAYES] bayesmh, [PSS] Glossary
clinically
meaningful difference, [PSS] Glossary, also see \( \delta \)
meaningful effect, [PSS] Glossary, also see clinically meaningful difference
significance difference, see clinically meaningful difference
significant difference, see clinically meaningful difference
Clip() function, [FN] Programming functions
Clock() function, [D] datetime, [D] date and time translation, [FN] Date and time functions, [M-5] date()
clock() function, [D] datetime, [D] date and time translation, [FN] Date and time functions, [M-5] date()
clock position, [G-4] clockposstyle
clock time, [TS] tset
clockposstyle, [G-4] clockposstyle
cloglog command, [R] cloglog, [R] cloglog postestimation
cloglog option, see gsem option cloglog
closevar command, [D] closevar
close,
cmdlog subcommand, [R] log
file subcommand, [P] file
close subcommand, [G-2] graph close
log subcommand, [R] log
close graphs, [G-2] graph close
cls command, [R] cls
clstyle() option, [G-3] connect_options
cluster, continued
kmeans command, [MV] cluster kmeans and kmedians
kmedians command, [MV] cluster kmeans and kmedians
dset command, [MV] cluster kmeans and kmedians
list command, [MV] cluster kmeans and kmedians
list command, [MV] cluster programming utilities
medianlinkage command, [MV] cluster linkage
notes command, [MV] cluster notes
parsesistance command, [MV] cluster programming utilities
query command, [MV] cluster programming utilities
rename command, [MV] cluster utility
renamevar command, [MV] cluster utility
set command, [MV] cluster programming utilities
singlelinkage command, [MV] cluster linkage
stop command, [MV] cluster stop
use command, [MV] cluster utility
wardslinkage command, [MV] cluster linkage
waveragelinkage command, [MV] cluster linkage
cluster analysis, [MV] cluster
cluster dendrogram, [MV] cluster generate,
[MV] cluster kmeans and kmedians,
[MV] cluster linkage, [MV] cluster stop,
[MV] cluster utility, [MV] Glossary,
[U] 26.28 Multivariate and cluster analysis
dendrograms, [MV] cluster dendrogram
dropping, [MV] cluster utility
hierarchical, [MV] cluster, [MV] clustermat,
[MV] cluster linkage
kmeans, [MV] cluster kmeans and kmedians
kmedians, [MV] cluster kmeans and kmedians
listing, [MV] cluster utility
notes, [MV] cluster notes
programming, [MV] cluster programming subroutines, [MV] cluster programming utilities
renaming, [MV] cluster utility
stopping rules, [MV] cluster, [MV] cluster stop
tree, [MV] cluster dendrogram, [MV] Glossary
using, [MV] cluster utility
cluster estimator of variance, [P] _robust,
[R] vce_option, [XT] vce_options
alternative-specific
conditional logit model, [R] asclogit
multinomial probit regression, [R] asmprobit
rank-ordered probit regression, [R] asprobit
beta regression, [R] betareg
censored Poisson regression, [R] cpoisson
competing-risks regression, [ST] stcrreg
complementary log-log regression, [R] cloglog
Cox proportional hazards model, [ST] stcox
exponential regression, hurdle, [R] churdle
fixed-effects models,
linear, [XT] xtreg
Poisson, [XT] xtpoisson
cluster estimator of variance, continued

fractional response regression, [R] fracreg
generalized linear models, [R] glm
  for binomial family, [R] binreg
generalized method of moments, [R] gmm,
  [R] ivpoisson
hekman selection model, [R] heckman
hurdle regression, [R] churdle
instrumental-variables regression, [R] ivregress
interval regression, [R] interg
linear dynamic panel-data models, [XT] xtabond,
  [XT] xtdp, [XT] xtdpdsys
linear regression, [R] regress
  constrained, [R] cnreg
  hurdle, [R] churdle
  truncated, [R] truncreg
  with dummy-variable set, [R] areg
logistic regression, [R] logistic, [R] logit, also see
  logit regression subentry
  conditional, [R] clogit
  multinomial, [R] mlogit
  ordered, [R] ollogit
  rank-ordered, [R] rologit
  skewed, [R] scobit
  stereotype, [R] slogit
logit regression, [R] logit, also see logistic regression subentry
  nested, [R] nlogit
maximum likelihood estimation, [R] ml, [R] mlexp
multilevel mixed-effects models, [ME] meclag, [ME] meglm, [ME] melogit, [ME] menbreg,
  [ME] mixed
multinomial
  logistic regression, [R] mlogit
  probit regression, [R] mprobit
negative binomial regression
  truncated, [R] nbreg
  zero-inflated, [R] zinb
nonlinear
  least-squares estimation, [R] nl
  systems of equations, [R] nlsur
parametric survival models, [ST] streg
Poisson regression, [R] poisson
  censored, [R] cpoisson
  truncated, [R] tpoisson
  with endogenous covariates, [R] ivpoisson
  zero-inflated, [R] zinb
population-averaged models, [XT] xtggee
  complementary log-log, [XT] xtclag
  logit, [XT] xtlogit
  negative binomial, [XT] xtnbreg
Poisson, [XT] xtpoisson
  probit, [XT] xtpoibt
  truncreg
Prais–Winsten and Cochrane–Orcutt regression,
  [TS] prais
  with heteroskedastic, [R] hetreg
  multinomial, [R] mprobit
  ordered, [R] oprobit
  order=heckman selection model,
  [R] heckprobit
  with endogenous covariates, [R] ivprobit
  with sample selection, [R] heckprobit
random-effects models
  complementary log-log, [XT] xtclag
  linear, [XT] xtreg
  logistic, [XT] xtlogit, [XT] xtolreg
  parametric survival, [XT] xtstreg
Poisson, [XT] xtpoisson
  probit, [XT] xtpoibt, [XT] xtpoibt
  structural equation modeling, [SEM] intro 8,
  [SEM] sem option method()
summary statistics,
  mean, [R] mean
  proportion, [R] proportion
  ratio, [R] ratio
  total, [R] total
tobit model, [R] tobit
  with endogenous covariates, [R] ivtobit
treatment-effects model, [TE] eteffects,
  [TE] etpoisson, [TE] etregr
truncated
  negative binomial regression, [R] tnreg
  Poisson regression, [R] tpoisson
  regression, [R] truncreg
  with endogenous covariates,
  Poisson regression, [R] ivpoisson
  probit model, [R] ivprobit
tobit model, [R] ivtobit
  with endogenous regressors,
  instrumental-variables regression, [R] ivregress
  zero-inflated
  negative binomial regression, [R] zinb
  Poisson regression, [R] zip
cluster sampling, [P] _robust, [ST] stcox, [ST] streg,
  [R] bootstrap, [R] bsample, [R] jackknife
clustered, [SEM] Glossary
clustering, see cluster analysis
clustermat, [MV] clustermat
  averagelinkage command, [MV] clustermat
  centroidlinkage command, [MV] clustermat
  completelinkage command, [MV] clustermat
  medianlinkage command, [MV] clustermat
  singlelinkage command, [MV] clustermat
  stop command, [MV] clustermat
  wardslinkage command, [MV] clustermat
  waverelinkage command, [MV] clustermat
clusters, duplicating, [D] expandcl
cofd() function, [D] datetime, [FN] Date and time functions, [M-5] date()
cohort studies, [ST] ifile
cohort study, [PSS] intro, [PSS] power,
            [PSS] Glossary, [R] epitab, [ST] stcox,
cointegration, [TS] fcast compute, [TS] fcast graph,
            [TS] vec intro, [TS] vec, [TS] vecmlr,
            [TS] vecnorm, [TS] vcrnorm, [TS] vcestable,
            [TS] Glossary
coleq macro extended function, [P] macro
coleq, matrix subcommand, [P] matrix rrownames
colfullnames macro extended function, [P] macro
collapse command, [D] collapse
            _collapse() function, [M-5] sort()
collatorlocale() function, [FN] String functions
collarversion() function, [FN] String functions
collect statistics, [D] statsby
collinear option, see gsem option collinear
collinear variables, removing, [P] _rmcoll
collinearity,
display of omitted variables, [R] set showbaselevels
            handling by regress, [R] regress
            retaining collinear variables, [R] estimation options,
            [R] orthog
            variance inflation factors, [R] regress postestimation
colmax() function, [M-5] minmax()
colmaxabs() function, [M-5] minmax()
colmin() function, [M-5] minmax()
colminmax() function, [M-5] minmax()
colmissing() function, [M-5] missing()
colnames macro extended function, [P] macro
colnames, matrix subcommand, [P] matrix rrownames
colnomissing() function, [M-5] missing()
colnumb() function, [FN] Matrix functions,
            [P] matrix define
            background, [G-4] schemes intro
dimming and brightening, [G-2] graph twoway
            histogram, [G-2] graph twoway kdensity,
            [G-4] colorstyle
foreground, [G-4] schemes intro
intensity adjustment, [G-2] graph twoway
            histogram, [G-2] graph twoway kdensity,
            [G-4] colorstyle
            of bars, [G-3] barlook_options
            of connecting lines, [G-3] connect_options
            of markers, [G-3] marker_options
            of pie slices, [G-2] graph pie
            of text, [G-3] textbox_options
setting background and fill, [G-3] region_options
color() option, [G-2] graph twoway histogram,
            [G-2] graph twoway kdensity
color, palette subcommand, [G-2] palette
colors, specifying in programs, [P] display
            colorstyle, [G-4] colorstyle
cols() function, [M-5] rows()
colscalefactors() function, [M-5] _equilre()
colshape() function, [M-5] rowshape()
colsof() function, [FN] Matrix functions, [P] matrix define
colsum() function, [M-5] sum()
_column(#), display directive, [P] display
column of matrix, selecting, [M-5] select()
column stripes, [M-6] Glossary
column-join operator, [M-2] op_

column-major order, [M-6] Glossary
column-join operator, [M-2] op_
columns in graphs, [PSS] Glossary
columns of matrix,
appending to, [P] matrix define
rownames, [U] 14.2 Row and column names
operators on, [P] matrix define
combinatorial function, [M-5] comb()
combinatorials, calculating, [FN] Mathematical functions
combine, graph subcommand, [G-2] graph combine
combining
data, [MI] mi add, [MI] mi append, [MI] mi merge
datasets, [D] append, [D] cross, [D] joinby,
[D] merge, [U] 22 Combining datasets
graphs, [G-2] graph combine
command
arguments, [P] gettoken, [P] syntax, [P] tokenize,
[U] 18.4 Program arguments
language, [SEM] Glossary
line, launching dialog box from, [R] db
parsing, [P] gettoken, [P] syntax, [P] tokenize,
[U] 18.4 Program arguments
timings, [U] 8 Error messages and return codes
commands,
abbreviating, [U] 11.2 Abbreviation rules
aborting, [P] continue, [U] 9 The Break key,
[U] 10 Keyboard use
editing and repeating, [U] 10 Keyboard use
immediate, [U] 19 Immediate commands
repeating automatically, [D] by, [P] byable,
[P] continue, [P] foreach, [P] forvalues,
[P] while
reviewing, [R] #review
unabbreviating names of, [P] unabcmd
commas, reading data separated by, [D] import
delimited, [D] infile (fixed format), [D] infile
(free format)
comments, [M-2] comments
adding to programs, [P] comments
comments, continued
in programs, do-files, etc., [U] 16.1.2 Comments
and blank lines in
do-files, [U] 18.11.2 Comments and long lines
in ado-files
with data, [D] notes
common, estat subcommand, [MV] factor
postestimation
common factors, [MV] Glossary
common odds ratio, [PSS] power cmh, [PSS] Glossary
communality, [MV] factor, [MV] factor
postestimation, [MV] Glossary
commutation matrix, [M-5] Kmatrix()
comparative fit index, [SEM] estat gof, [SEM] methods
and formulas for sem
comparative scatterplot, [R] dotplot
compare command, [D] compare
compare, estat subcommand, [MV] procrustes
postestimation
comparing two
files, [D] cf, [D] checksum
variables, [D] compare
comparison
group, see experimental group
test between nested models, [R] nestreg
value, see alternative value
compassdirstyle, [G-4] compassdirstyle
compatibility of Stata programs across releases,
[P] version
complementary log-log regression, [ME] meclolog,
[R] cloglog, [R] glm, [SEM] Glossary,
[SVY] svy estimation, [XT] xtcloglog,
[XT] xtgee
complete
data, [MI] Glossary
degrees of freedom for coefficients, [MI] mi estimate, [MI] Glossary
observations, [MI] Glossary
complete-cases analysis, [MI] Glossary
complete-data analysis, [MI] Glossary
completed data, [MI] Glossary
completed-data analysis, [MI] intro substantive,
[MI] mi estimate, [MI] Glossary
completelinkage,
clustermat subcommand, [MV] cluster linkage
cluster subcommand, [MV] cluster linkage
complete-linkage clustering, [MV] cluster,
[MV] clustermat, [MV] cluster linkage,
[MV] Glossary
completely determined outcomes, [R] logit
complex, [M-2] declarations, [R] logit
complex, [M-2] declarations, [R] logit
component
analysis, [MV] factor, [MV] pca, [MV] rotate,
[MV] rotatemat
loading plot, [MV] scoreplot
plot, [MV] scoreplot
scores, [MV] Glossary
component-plus-residual plot, [G-2] graph other,
[R] regress postestimation diagnostic plots
components of PSS analysis
clinically meaningful difference, see clinically meaningful difference
effect size, see δ
power, see power
sample size, see sample-size
significance level, see significance level
statistical method, see test
compound double quotes, [P] macro
compound symmetric
correlation matrix, [MV] mvtest correlations
covariance matrix, [MV] mvtest covariances
compound symmetry, [PSS] Glossary
compress command, [D] Glossary
Glossary
component-plus-residual plot, [G-2] graph other,
[R] regress postestimation diagnostic plots
components of PSS analysis
clinically meaningful difference, see clinically meaningful difference
effect size, see δ
power, see power
sample size, see sample-size
significance level, see significance level
statistical method, see test
compound double quotes, [P] macro
compound symmetric
correlation matrix, [MV] mvtest correlations
covariance matrix, [MV] mvtest covariances
compound symmetry, [PSS] Glossary
compress command, [D] compress
compress files, [D] zipfile
compute, fcast subcommand, [TS] fcast compute
Comrey’s tandem 1 and 2 rotations, [MV] rotate, [MV] rotatemat, [MV] Glossary
cat(), egen function, [D] egen
concatenating strings, [U] 13.2.2 String operators
correlation matrix, [MV] mvtest correlations
covariance matrix, [MV] mvtest covariances
cond() function, [FN] Programming functions, [M-5] cond()
condition statement, [P] if
conditional
conjugacy, see semiconjugate prior
fixed-effects model, [XT] Glossary
(fixed-effects) logistic regression, [SVY] svy estimation
imputation, see imputation, conditional
independence, [IRT] Glossary
marginal effects, [R] margins, [R] marginsplot
margins, [R] margins, [R] marginsplot
mean, [TE] Glossary
mean independence assumption, [TE] teffects intro advanced
confidence interval, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [SEM] Glossary,
[SVY] variance estimation, [U] 20.7 Specifying the width of confidence intervals
for bioequivalence, [R] pkequiv
for bootstrap statistics, [R] bootstrap postestimation, [R] rocreg, [R] rocreg postestimation
for combinations of coefficients, linear, [R] lincom
nonlinear, [R] nlcom
for contrasts, [R] contrast
for counts, [R] ci
for cumulative hazard function, [ST] sts list
for false-positive rates, [R] rocregplot
for hazard ratios, [ST] stcox, [ST] streg
for incidence-rate ratios, [R] eipoisson,
[R] expoisson, [R] glm, [R] nbreg, [R] poisson,
[R] tnbreg, [R] tpoisson, [R] zinb, [R] zip,
[ST] stir, [TE] etpoisson, [XT] xtgee,
[XT] xtnbreg, [XT] xtpoisson
for intragroup correlations, [R] loneway
for linear combinations, [SVY] svy postestimation
for margins, [R] margins
for means, [R] ci, [R] ameans, [R] esize, [R] mean,
[R] ttest, [R] ztest
for means and percentiles of survival time, [ST] stci
for medians and percentiles, [R] centile
for odds and risk ratios, [R] epitab
for odds ratios, [R] exlogistic, [R] glm, [R] logistic,
[R] logit, [R] ologit, [R] scobit, [XT] xtcloglog,
[XT] xtgee, [XT] xtologit, [XT] xtstreg
for proportions, [R] ci, [R] proportion
for ratios, [R] ratio
for relative-risk ratios, [R] mlogit
for ROC area, [R] rocomp, [R] rocfit, [R] rocreg,
[R] roctab
for ROC values, [R] rocregplot
for standard deviations, [R] ci
for standardized mortality ratios, [R] dstdize,
[ST] stptime, [ST] strate
for subhazard ratios, [ST] stcrreg
for survival rates, [ST] sttable
for survivor function, [ST] stlist
for tabulated proportions, [SVY] svy: tabulate twoway
for totals, [R] total
for variances, [R] ci
confidence interval, set default, [R] level
confidence levels, [R] level
config, estat subcommand, [MV] mds postestimation
configuration, [MV] Glossary
configuration plot, [MV] mds postestimation plots,
[MV] Glossary
conditional-independence assumption, [TE] teffects intro advanced,
constrained estimation, continued
fractional response regression, [R] fracreg
GARCH model, [TS] mgarch ccc, [TS] mgarch
dcc, [TS] mgarch dvech, [TS] mgarch vcc
generalized linear models, [R] glm
for binomial family, [R] binreg
generalized negative binomial regression, [R] nbreg
heckman selection model, [R] heckman,
[R] heckoprobit
hurdle
regression, [R] churdle
interval regression, [R] intreg
linear regression, [R] cnreg
hurdle, [R] churdle
seemingly unrelated, [R] sureg
stochastic frontier, [R] frontier
three-stage least squares, [R] reg3
truncated, [R] truncreg
logistic regression, [R] logistic, [R] logit, also see
logit regression subentry
conditional, [R] clogit
multinomial, [R] mlogit
ordered, [R] ologit
skewed, [R] scobit
stereotype, [R] slogit
logit regression, [R] logit, also see logistic regression
subentry
nested, [R] nlogit
Markov-switching model, [TS] mswitch
maximum likelihood estimation, [R] ml, [R] mlexp
multilevel mixed-effects, [ME] mestreg
mle, [R] mepoisson, [ME] meprobit, [ME] mestreg
logit regression, [R] logit, also see logistic regression
subentry
nested, [R] nlogit
Markov-switching model, [TS] mswitch
maximum likelihood estimation, [R] ml, [R] mlexp
multilevel mixed-effects, [ME] mestreg
mle, [R] mepoisson, [ME] meprobit, [ME] mestreg
parametric survival models, [ST] streg
Binomial, [R] poisson
censored, [R] cpoisson
truncked, [R] tnbreg
zero-inflated, [R] zip
parametric survival models, [ST] streg
Poission regression, [R] poisson
censored, [R] cpoisson
truncked, [R] tnbreg
zero-inflated, [R] zip
probit regression, [R] probit
bivariate, [R] biprobit
heteroskedastic, [R] hetprob
multinomial, [R] mprobit
ordered, [R] oprobit
with endogenous covariates, [R] ivprobit
with sample selection, [R] heckprobit
programming, [P] makecns
random-effects models
complementary log-log, [XT] xtlogit
time, [XT] xtintreg
logit, [XT] xtlogit, [XT] xtlogit
constrained estimation, random-effects models, continued
negative binomial, [XT] xtnbreg
parametric survival, [XT] xstreg
Poisson, [XT] xtpoisson
probit, [XT] xtprobit, [XT] xtprobit
tobit, [XT] xtobit
state-space model, [TS] sspace
stochastic frontier models for panel data, [XT] xfrontier
structural vector autoregressive models, [TS] var

constraint
command, [R] constraint
define command, [R] constraint
dir command, [R] constraint
drop command, [R] constraint
free command, [R] constraint
get command, [R] constraint
list command, [R] constraint
macro extended function, [P] macro
constraint matrix, creating and displaying, [P] makecns
constraints, [SEM] sem and gsem option constraints(),
[SEM] Glossary
across groups, [SEM] intro 6
normalization, [SEM] intro 4, [SEM] gsem,
[SEM] sem, [SEM] Glossary
relaxing, [SEM] intro 6, [SEM] sem and gsem path notation,
[SEM] sem path notation extensions
specifying, [SEM] intro 4, [SEM] intro 6,
[SEM] sem and gsem option constraints(),
[SEM] sem and gsem option constrainstructure(),
[SEM] sem and gsem path notation, [SEM] sem path notation extensions
constraints() option, see gsem option constraints() see sem option constraints()
constructor, [M-2] class
containers, [M-5] asarray()
containment DDF, see denominator degrees of freedom,
ANOVA
contents of data, [D] codebook, [D] describe, [D] ds,
[D] labelbook
context, class, [P] class
contingency table, [MV] ca, [PSS] power, [PSS] power
paired proportions, [PSS] power cmh,
[PSS] power mec, [PSS] power trend,
[R] epitab, [R] roctab, [R] symmetry, [R] table,
[R] tabulate twoway, [SVY] ssv: tabulate twoway
continue command, [P] continue
continuity correction, [PSS] power cmh, [PSS] power trend
continuous outcomes, see outcomes, continuous
parameters, [BAYES] bayestest interval,
[BAYES] Glossary
variable imputation, see imputation, continuous
contour, graph twoway subcommand, [G-2] graph twoway contour
contour plot, [G-2] graph twoway contour,
[G-3] clegend_option
contour-line plot, [G-2] graph twoway contourline
contourplot, graph twoway subcommand,
[G-2] graph twoway contourline
contract command, [D] contract
contrast command, [R] anova postestimation,
[R] contrast, [R] contrast postestimation,
[R] margins, contrast, [SEM] intro 7,
[SVY] svy postestimation, [U] 20.18 Obtaining
contrasts, tests of interactions, and main effects
contrasts, [MV] intro, [MV] manova postestimation,
[R] margins, contrast, [R] marginsplot,
[U] 20.18 Obtaining contrasts, tests of
interactions, and main effects
graphing, [U] 20.19 Graphing margins, marginal
effects, and contrasts
correlation, see correlation, control-group
mean, see means, control-group
proportion, see proportions, control-group
sample size, see sample-size
standard deviation, see standard deviations, control-group
variance, see variances, control-group
controlled clinical trial study, [PSS] power,
[PSS] Glossary
convergence, [SEM] intro 12, [SEM] sem, [SEM] sem
and gsem option from()
criteria, [R] maximize
do MCMC, see Markov chain Monte Carlo,
convergence
conversion, file, [D] changeext, [D] filefilter
convert, mi subcommand, [MI] mi convert
converting between styles, [MI] mi convert
convolve() function, [M-5] fft()
Cook–Weisberg test for heteroskedasticity, [R] regress postestimation
Cook’s D, [R] glm postestimation, [R] regress postestimation coordinates, estat subcommand, [MV] ca postestimation, [MV] mca postestimation
copy,
  graph subcommand, [G-2] graph copy
  label subcommand, [D] label
  mi subcommand, [MI] mi copy, [MI] styles
  ssc subcommand, [R] ssc
copy and paste, [D] edit
  .copy built-in class function, [P] class
copy command, [D] copy
copy graph, [G-2] graph copy
copy macro extended function, [P] macro
copycolor, set subcommand, [G-2] set printcolor,
[R] set
copying variables, [D] clonevar, [D] edit
copyright
  Apache, [R] copyright apache
  boost, [R] copyright boost
  icd10, [R] copyright icd10
  icu, [R] copyright icu
  lapack, [R] copyright lapack
  libhara, [R] copyright libhara
  libpng, [R] copyright libpng
  Mersenne Twister, [R] copyright mersennetwister
  MiG Layout, [R] copyright miglayout
  scintilla, [R] copyright scintilla
  symbol, [G-4] text
  ttf2pt1, [R] copyright ttf2pt1
  zlib, [R] copyright zlib
copyright command, [R] copyright
Cornfield confidence intervals, [R] epitab
Corr() function, [M-5] corr()
  _corr() function, [M-5] corr()
  corr() function, [FN] Matrix functions, [M-5] corr(),
[P] matrix define
corr2data command, [D] corr2data
  correcting data, see editing data
correlate command, [R] correlate
  correlated errors, see robust, Huber/White/sandwich estimator of variance, also see autocorrelation
  correlated uniqueness model, [SEM] intro 5,
  [SEM] example 17, [SEM] Glossary
[PSS] power, [R] correlate
  between paired observations, [PSS] power onemean,
  [PSS] power pairedmeans, [PSS] power pairedproportions
  binary variables, [R] tetrachoric
  canonical, [MV] canon
  coefficient of exposure, [PSS] power mce
  compound symmetric, [MV] mvtest correlations
  continuous variables, [R] correlate
  control-group, [PSS] power twocorrrelations
  data generation, [D] corr2data, [D] drawnorm
correlation, continued
  experimental-group, [PSS] power twocorrrelations
  factoring of, [MV] factor
  independent, see correlation, two-sample
  interitem, [MV] alpha
  intraclass, see intraclass correlation
  intracluster, [R] loneway
  Kendall’s rank, [R] spearman
  matrices, [MV] mvtest correlations, [P] matrix define,
  [R] correlate, [R] estat, [R] estat vce
  matrix, anti-image, [MV] factor postestimation,
  [MV] pca postestimation
  model, [SEM] intro 5, [SEM] Glossary
  one-sample, [PSS] power onecorrrelation
  pairwise, [R] correlate
  partial and semipartial, [R] pcorr
  principal components of, [MV] pca
  serial, [R] runtest
  similarity measure, [MV] measure_option
  Spearman’s rank, [R] spearman
  structure, [R] asprobit, [R] aspropobit, [R] reg3,
  [XT] xtcloglog, [XT] xtgee, [XT] xtglm,
  [XT] xtolgit, [XT] xtnbreg, [XT] xtpoisson,
  [XT] xtpoisson, [XT] xtprobib, [XT] xtnreg,
  [XT] xtstreg, [XT] Glossary
  testing equality, [MV] mvtest correlations
  tests of, [SEM] estat stdize, [SEM] example 16
  tetrachoric, [R] tetrachoric
  two-sample, [PSS] power twocorrrelations
  correlation, estat subcommand, [R] asprobit postestimation, [R] aspropobit postestimation
  correlation() function, [M-5] mean()
correlations,
  estat subcommand, [MV] ca, [MV] mca,
  [MV] Glossary
  correspondence analysis, [MV] ca, [MV] mca,
  [MV] Glossary
  correspondence analysis projection, [MV] ca
  postestimation plots, [MV] Glossary
correlogram, [G-2] graph other, [TS] corrgram,
[TS] Glossary
  corrgram command, [TS] corrgram
cos() function, [FN] Trigonometric functions,
[M-5] sin()
cosh() function, [FN] Trigonometric functions,
[M-5] sin()
  cosine function, [FN] Trigonometric functions
  cosine kernel function, [R] kdensity, [R] lpoly,
  [R] qreg, [TE] tebalance density, [TE] tebalance overlap,
  [TE] tefffects overlap
cost frontier model, [R] frontier, [XT] xtfpeak
  costs, [MV] Glossary
  count command, [D] count
count data,
  confidence intervals for counts, [R] ci
count data, continued
count, [ME] glm
R glm
[U] 26.13 Count dependent-variable models
graphs, [R] histogram, [R] kdensity, [R] spikeplot
imputation, see imputation, count data
interrater agreement, [R] kappa
summary statistics of, [R] table, [R] tabstat, [R] tabulate oneway, [R] tabulate twoway,
[R] tabulate, summarize()
symmetry and marginal homogeneity tests,
[R] symmetry

count(), egen function, [D] egen
count model, [SEM] intro 5, [SEM] example 34g, [SEM] example 39g
count outcome model, see outcomes, count
count, ml subcommand, [R] ml
counterfactual, [TE] Glossary, also see potential outcome
counts, making dataset of, [D] collapse
courses about Stata, [U] 3.6.2 NetCourses
covariance, [SEM] intro 4, [SEM] Glossary
analysis of, [R] anova
assumptions, [SEM] gsem, [SEM] sem
matrix,
anti-image, [MV] factor postestimation,
[MV] pca postestimation
block diagonal, [MV] mtest covariances
spherical, [MV] mtest covariances
testing equality, [MV] mtest covariances
of variables or coefficients, [R] correlate
principal components of, [MV] pca
stationarity, [TS] Glossary
structure, [ME] me, [ME] Glossary
covariance, estat subcommand, [MV] discrim
lda postestimation, [MV] discrim qda
postestimation, [R] asmprobit postestimation,
[R] asprobit postestimation
covariance() option, see gsem option
covariance(), see sem option covariance()
covariances, mtest subcommand, [MV] mtest covariances
covariances, creating dataset from, see summary statistics data
covariate
class, [D] duplicates
patterns, [R] logistic postestimation, [R] logit postestimation, [R] probit postestimation
covariates, [ST] Glossary
covarimargin rotation, [MV] rotate, [MV] rotatemat,
[MV] Glossary
COVRATIO, [R] regress postestimation
covstructure() option, see gsem option
covstructure(), see sem option covstructure()
cox, power subcommand, [PSS] power cox
Cox proportional hazards model, [PSS] power cox,
[ST] stcox, [SVY] svy estimation
test of assumption, [ST] stcox, [ST] stcox PH-assumption tests, [ST] stcox postestimation,
[ST] stsplit
Cox–Snell residual, [ST] stcox postestimation,
[ST] streg postestimation
cpoisson command, [R] cpoisson, [R] cpoisson postestimation
cprplot command, [R] regress postestimation
diagnostic plots
crude residual, [R] churdle
Cramér’s V, [R] tabulate twoway
crawford-ferguson rotation, [MV] rotate,
[MV] rotatemat, [MV] Glossary
create,
beal subcommand beal subcommand, [D] beal
forecast subcommand, [TS] forecast create
irf subcommand, [TS] irf create
serset subcommand, [P] serset
create_cspine, serset subcommand, [P] serset
create_xmedians, serset subcommand, [P] serset
critical
credible interval, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesmh postestimation, [BAYES] bayesstats summary,
[BAYES] Glossary
set default, [BAYES] set clevel
credible level, [BAYES] intro, [BAYES] bayesmh,
creturn list command, [P] creturn
creexternal() function, [M-5] findexternal()
crossing variables, [MV] Glossary
crossover designs, [R] pk, [R] pkcross, [R] pkshape
cross-sectional data, [XT] Glossary
cross-tabulations, see tables

crude estimate, [R] epitab

cs command, [R] epitab

csi command, [R] epitab

.csv filename suffix, [D]
cut() function, [M-5]
cvpermute() function, [M-5]
cummul command, [R] cummul
cumulative distribution functions, [FN] Statistical functions
distribution, empirical, [R] cummul
hazard ratio, see hazard ratio incidence

data, [R] epitab


incidence data, [R] poisson
spectral distribution, empirical, [TS] cumsp, [TS] psdensity

current data, [P] creturn
curse of dimensionality, [MV] Glossary
curved path, [SEM] Glossary
custom prediction equations, [MI] mi impute chained, [MI] mi impute monotone
cusum

test, [R] cumsum
cusum command, [R] cumsum
CUSUM plot, see cumplot
cut(), egen function, [D] egen
cutil, see classutil
cv. estat subcommand, [SVY] estat
cvpermute() function, [M-5]
cvpermutesetup() function, [M-5]

crossing variables, [MV] Glossary
crossover designs, [R] pk, [R] pkcross, [R] pkshape
cross-sectional data, [XT] Glossary
cross-tabulations, see tables
crude estimate, [R] epitab, [ST] Glossary
cs command, [R] epitab
csi command, [R] epitab

.csv filename suffix, [D]
cut() function, [M-5]
cvpermute() function, [M-5]
cummul command, [R] cummul
cumulative distribution functions, [FN] Statistical functions
distribution, empirical, [R] cummul
hazard ratio, see hazard ratio incidence
data, [R] epitab


incidence data, [R] poisson
spectral distribution, empirical, [TS] cumsp, [TS] psdensity

current data, [P] creturn
curse of dimensionality, [MV] Glossary
curved path, [SEM] Glossary
custom prediction equations, [MI] mi impute chained, [MI] mi impute monotone
cusum

test, [R] cumsum
cusum command, [R] cumsum
CUSUM plot, see cumplot
cut(), egen function, [D] egen
cutil, see classutil
cv. estat subcommand, [SVY] estat
cvpermute() function, [M-5] cvpermute()

crossing variables, [MV] Glossary
crossover designs, [R] pk, [R] pkcross, [R] pkshape
cross-sectional data, [XT] Glossary
cross-tabulations, see tables
crude estimate, [R] epitab, [ST] Glossary
cs command, [R] epitab
csi command, [R] epitab

.csv filename suffix, [D]
cut() function, [M-5]
cvpermute() function, [M-5]
cummul command, [R] cummul
cumulative distribution functions, [FN] Statistical functions
distribution, empirical, [R] cummul
hazard ratio, see hazard ratio incidence
data, [R] epitab


incidence data, [R] poisson
spectral distribution, empirical, [TS] cumsp, [TS] psdensity

current data, [P] creturn
curse of dimensionality, [MV] Glossary
curved path, [SEM] Glossary
custom prediction equations, [MI] mi impute chained, [MI] mi impute monotone
cusum

test, [R] cumsum
cusum command, [R] cumsum
CUSUM plot, see cumplot
cut(), egen function, [D] egen
cutil, see classutil
cv. estat subcommand, [SVY] estat
cvpermute() function, [M-5] cvpermute()
data, continued

inputting, see importing data, see inputting data interactively, see reading data from disk
labeling, see labeling data
large, dealing with, see memory
listing, see listing data
loading, see importing data, see inputting data interactively, see using data
matched case-control, see matched case-control data
missing values, see missing values
mlong, see mlong
multiple-failure st, see multiple-failure st data
multiple-record st, see multiple-record st data
nested case-control, see nested case-control data
observational, see observational data
preserving, see preserving data
range of, see range of data
ranking, see ranking data
reading, see importing data, see loading data, see reading data from disk
recoding, see recoding data
rectangularizing, see rectangularize dataset
reorganizing, see reorganizing data
restoring, see restoring data
sampling, see sampling
saving, see exporting data, see saving data
single-failure st, see survival analysis
single-record st, see survival analysis
stacking, see stacking data
strings, see string variables
summarizing, see summarizing data
survey, see survey data
survival-time, see survival analysis
time-series, see time-series analysis
time-span, see time-span data
transposing, see transposing data
verifying, see certifying data
wide, see wide

Data Browser, see Data Editor
Data Editor, [D] edit

copy and paste, [D] edit
data label macro extended function, [P] macro
data, label subcommand, [D] label
data-have-changed flag, [M-5] st_updata() database, reading data from, [D] odbc
other software, [U] 21.4 Transfer programs
dataset,

adding notes to, [D] notes
comparing, [D] cf, [D] checksum
creating, [D] corr2data, [D] drawnorm
example, [U] 1.2.2 Example datasets
loading, see importing data, see inputting data interactively, see using data
rectangularize, [D] fillin
saving, see exporting data, see saving data
dataset labels, [D] label, [D] label language, [D] notes
determining, [D] codebook, [D] describe
managing, [D] varmanage
datasignature
clear command, [D] datasignature
command, [D] datasignature, [SEM] example 25,
[SEM] ssd
confirm command, [D] datasignature
report command, [D] datasignature
set command, [D] datasignature
_dataasignature command, [P] _datasignature
date
and time stamp, [D] describe
functions, [D] datetime, [D] datetime translation,
[FN] Date and time functions, [M-5] date()
date() function, [D] datetime, [D] datetime translation,
[FN] Date and time functions, [M-5] date()
date,

displaying, [U] 12.5.3 Date and time formats,
[U] 24.3 Displaying dates and times
formats, [U] 12.5.3 Date and time formats,
[U] 24.3 Displaying dates and times
functions, [U] 24.5 Extracting components of dates and times
inputting, [U] 24.2 Inputting dates and times
variables, [U] 24 Working with dates and times
datelist, [U] 11.1.9 datelist
dates,
business, see business calendars
Excel, [D] datetime
OpenOffice, [D] datetime
R, [D] datetime
SAS, [D] datetime
SPSS, [D] datetime

Davidson–Fletcher–Powell algorithm,
day() function, [D] datetime, [FN] Date and time functions, [M-5] date(), [U] 24.5 Extracting components of dates and times
db command, [R] db
dBASE, reading data from, [U] 21.4 Transfer programs
dcc, mgarch subcommand, [TS] mgarch dcc
dct, m garch subcommand, [TS] mgarch dcc
dct file, [U] 11.6 Filenaming conventions
dct filename suffix, [D] import, [D] infile (fixed format), [D] infx (fixed format), [D] outfile
DF, see denominator degrees of freedom
decimal symbol, setting, [D] format
.declare built-in class modifier, [P] class
declare, class, [P] class
debug, command, [P] debug
decode command, [D] encode
decode() function, [M-5] fft()
decrement operator, [M-2] op
default settings of system parameters, [P] set
locale_functions, [P] set locale_ui, [R] query,
[R] set_defaults
defective matrix, [M-6] Glossary
DEFF, see design effects
define,
char subcommand, [P] char
constraint subcommand, [R] constraint
label subcommand, [D] label
matrix subcommand, [P] matrix define
program subcommand, [P] program, [P] program properties
scalar subcommand, [P] scalar
transmap subcommand, [R] translate
DEFT, see design effects
degree-to-radian conversion, [FN] Mathematical functions
degree-of-freedom adjustment, [SEM] Glossary
degrees of freedom, [MI] mi estimate, [MI] mi predict
for coefficients, complete, see complete degrees of freedom for coefficients, also see estimation, degrees of freedom for coefficients
delete, [M-5] unlink()
delete, cluster subcommand, [MV] cluster programming utilities
deleting
case-wise, [D] egen
files, [D] erase
variables or observations, [D] drop
#delimit command, [M-2] semicolons, [P] #delimit
delimited,
export subcommand, [D] import delimited
import subcommand, [D] import delimited
delimiter
for comments, [P] comments
for lines, [P] #delimit
delta, see δ
delta, continued
method, [R] margins, [R] nlcom, [R] predictnl,
[R] testnl, [SEM] estat residuals, [SEM] estat ttests, [SVY] variance estimation,
[SVY] Glossary
dendrogram, [G-2] graph other, [MV] cluster,
[MV] cluster dendrogram, [MV] Glossary
dendrogram, cluster subcommand, [MV] cluster dendrogram
denominator degrees of freedom
ANOVA, [ME] mixed, [ME] Glossary
repeated, [ME] mixed, [ME] Glossary
residual, [ME] mixed, [ME] Glossary
Satterthwaite, [ME] mixed, [ME] Glossary
density
estimation, kernel, [R] kdensity
smoothing, [G-2] graph other
density option, [G-2] graph twoway histogram
density-distribution sunflower plot, [R] sunflower
dereference, [M-6] Glossary
._deriv() function, [M-5] deriv()
deriv() function, [M-5] deriv()
derivative of incomplete gamma function,
[FN] Mathematical functions, [FN] Statistical functions
derivatives, [M-5] deriv()
derivatives, numeric, [R] dydx, [R] testnl
derived plottypes, [G-3] advanced_options
deriv_init() functions, [M-5] deriv()
deriv_init-*(*) functions, [M-5] deriv()
deriv_query() function, [M-5] deriv()
deriv_result-*(*) functions, [M-5] deriv()
describe,
ado subcommand, [R] net
bcal subcommand, [D] bcal
classutil subcommand, [P] classutil
estimates subcommand, [R] estimates describe
forecast subcommand, [TS] forecast describe
graph subcommand, [G-2] graph describe
irf subcommand, [TS] irf describe
mata subcommand, [M-3] mata describe
mi subcommand, [MI] mi describe
net subcommand, [R] net
putexcel subcommand, [P] putexcel, [P] putexcel advanced
ssc subcommand, [R] ssc
ssd subcommand, [SEM] ssd
describe command, [D] describe, [U] 12.6 Dataset, variable, and value labels
describing graph, [G-2] graph describe
describing mi data, [MI] mi describe
descriptive statistics,
CIs for means, proportions, and variances, [R] ci correlations, [R] correlate, [R] pcorr,
[R] tetrachoric
description statistics, continued
  creating dataset containing, [D] collapse
  creating variables containing, [D] egen
displaying, [D] codebook, [D] pctile, [R] grmeanby,
estimation, [R] mean, [R] proportion, [R] ratio,
  [R] total
means, [R] ameans, [R] summarize
percentiles, [R] centile
pharmacokinetic data,
  make dataset of, [R] pkcollapse
  summarize, [R] pksumm
reporting, [SEM] estat summarize
tables, [R] table, [R] tabstat, [R] tabulate oneway,
  [R] tabulate twoway, [R] tabulate, summarize()
design, fnset subcommand, [R] fnset
design effects, [R] loneway, [SVY] estat,
  [SVY] svy: tabulate oneway,
design matrix, [M-5] designmatrix()
  designmatrix() function, [M-5] designmatrix()
destring function, [M-2] class
destroy() function, [M-2] class
destructors, class, [P] class
  [P] matrix define
determinant of matrix, [M-5] det(), [P] matrix define
deterministic trend, [TS] Glossary
dettriangular() function, [M-5] det()
deviance information criterion, [BAYES] bayes,
  [BAYES] bayesstats ic, [BAYES] Glossary
deviance residual, [ME] mecloglog postestimation,
  [ME] meglm postestimation, [ME] melogit
  postestimation, [ME] membreg postestimation,
  [ME] mepoisson postestimation, [ME] meprobit
  postestimation, [ME] meprobitpostestimation,
  [ME] meqprobit postestimation,
  [ME] mestreg postestimation, [R] binreg
  postestimation, [R] fp postestimation, [R] glm
  postestimation, [R] logistic postestimation,
  [R] logit postestimation, [R] probit
  postestimation, [ST] stcox postestimation,
  [ST] streg postestimation
deviation cross product, [M-5] crossdev(),
  [M-5] quadcross()
dexpowonential, tsmooth subcommand,
  [TS] tsmooth exponential
df, estat subcommand, [ME] mixed postestimation
dfactor command, [TS] dfactor, [TS] dfactor
  postestimation
DFBETA, [R] regress postestimation, [ST] stcox
  postestimation, [ST] stcrreg postestimation,
  [ST] Glossary
dbeta command, [R] regress postestimation
dfgls command, [TS] dfgls
DFITS, [R] regress postestimation
DFP algorithm, [R] ml
dfuller command, [TS] dfuller
dgammadada() function, [FN] Statistical functions,
  [M-5] normal()
dgammadada() function, [FN] Statistical functions,
  [M-5] normal()
dgammadadex() function, [FN] Statistical functions,
  [M-5] normal()
dgammadadex() function, [FN] Statistical functions,
  [M-5] normal()
dgammadadex() function, [FN] Statistical functions,
  [M-5] normal()
dhms() function, [D] datet ime, [FN] Date and time
  functions, [M-5] date()
  _diag() function, [M-5] _diag()
  [P] matrix define
diagccnt() function, [FN] Matrix functions,
  [M-5] diagccnt(), [P] matrix define
diagnostic codes, [D] icd, [D] icd9, [D] icd10
  diagnostic plots, [G-2] graph other, [R] diagnostic
  postestimation, [R] regress
  postestimation diagnostic plots
diagnostics, regression, see regression diagnostics
  [M-6] Glossary
  vech model, [TS] mgarch, [TS] mgarch dvech
diagonal() function, [M-5] diagonal()
diagonals of matrices, [P] matrix define
dialog
  box, [P] dialog programming, [P] window
  programming, [P] window fopen, [P] window
  manage, [P] window menu, [P] window push,
  [P] window stopbox, [R] db
  programming, [P] dialog programming, [P] window
  programming, [P] window fopen, [P] window
  manage, [P] window menu, [P] window push,
  [P] window stopbox
DIF, see deviance information criterion
Dice coefficient similarity measure,
  [MV] measure_option
dichotomous item, [IRT] Glossary
dichotomous outcome model, see outcomes, binary
  Dickey–Fuller test, [TS] dfgls, [TS] dfuller
dictionaries, [D] export, [D] import, [D] infile (fixed
  format), [D] infiles, [D] outfiles, [M-5] asarray()
  DIF, see differential item functioning
  diff(), egen function, [D] egen
difference of estimated coefficients, see linear
  combinations of estimators
difference operator, [TS] Glossary, [U] 11.4.4 Time-
  series varlists
differences of two means test, [SVY] svy
  postestimation
differential item functioning, [IRT] dif, [IRT] dilogistic,
differentiation, [M-5] deriv()
difficult option, [R] maximize
difficulty, [IRT] Glossary
diflogistic command, [IRT] dif, [IRT] diflogistic
difmh command, [IRT] dif, [IRT] difmh
digamma() function, [FN] Mathematical functions,
[M-5] factorial()
digitally signing data, see datasignature command
digits, controlling the number displayed, [D] format,
[U] 12.5 Formats: Controlling how data are displayed
dilation, [MV] procrustes, [MV] Glossary
dimension, [MV] Glossary
diminishing adaptation, [BAYES] Glossary
discount

discrim

\texttt{knn} command, [MV] discrim, [MV] discrim estat, [MV] discrim knn, [MV] discrim knn postestimation
lda command, [MV] discrim, [MV] discrim estat, [MV] discrim lda, [MV] discrim lda postestimation
logistic command, [MV] discrim, [MV] discrim estat, [MV] discrim logistic, [MV] discrim logistic postestimation

discriminant function, [MV] discrim, [MV] discrim lda, [MV] discrim lda postestimation

discriminating variables, [MV] Glossary
discrimination, [IRT] Glossary
disparity, [MV] Glossary
dispersion, measures of, [D] petile, [XT] xtsum, see percentiles, displaying, see standard deviations, displaying, see variance, displaying, see range of data
display

as error, [M-5] displayas(), [M-5] erprintf()
as text, as result, etc., [M-5] displayas()
column, [I] Glossary
formats, [D] describe, [D] display,
graph, [G-2] graph display
settings, [R] set showbaselevels
width and length, [R] log
display

as error, [M-5] displayas(), [M-5] erprintf()
as text, as result, etc., [M-5] displayas()
column, [I] Glossary
formats, [D] describe, [D] display,
graph, [G-2] graph display
settings, [R] set showbaselevels
width and length, [R] log
display

as error, [M-5] displayas(), [M-5] erprintf()
as text, as result, etc., [M-5] displayas()
column, [I] Glossary
formats, [D] describe, [D] display,
graph, [G-2] graph display
settings, [R] set showbaselevels
width and length, [R] log

display

as error, [M-5] displayas(), [M-5] erprintf()
as text, as result, etc., [M-5] displayas()
column, [I] Glossary
formats, [D] describe, [D] display,
graph, [G-2] graph display
settings, [R] set showbaselevels
width and length, [R] log
display

as error, [M-5] displayas(), [M-5] erprintf()
as text, as result, etc., [M-5] displayas()
column, [I] Glossary
formats, [D] describe, [D] display,
graph, [G-2] graph display
settings, [R] set showbaselevels
width and length, [R] log

display

as error, [M-5] displayas(), [M-5] erprintf()
as text, as result, etc., [M-5] displayas()
column, [I] Glossary
formats, [D] describe, [D] display,
graph, [G-2] graph display
settings, [R] set showbaselevels
width and length, [R] log

display

as error, [M-5] displayas(), [M-5] erprintf()
as text, as result, etc., [M-5] displayas()
column, [I] Glossary
formats, [D] describe, [D] display,
displayas() function, [M-5] displayas()
displayflush() function, [M-5] displayflush()
displaying, also see printing, logs (output)

contents, [D] describe
data, [D] edit, [D] list
directories, [D] type
long strings, see string variables, long
macros, [P] macro
matrix, [P] matrix utility

output, [P] display, [P] quietly, [P] smcl,
[P] table
previously typed lines, [R] #review
scalar expressions, [P] display, [P] scalar
stored results, [R] stored results
dissimilarity, [MV] Glossary

matrix, [MV] matrix dissimilarity, [MV] Glossary,
[P] matrix dissimilarity
measures,
[MV] cluster, [MV] cluster programming utilities, [MV] matrix dissimilarity, [MV] nds,
[MV] measure_option, [P] matrix dissimilarity
Bray and Curtis, [MV] clustermat
Canberra, [MV] measure_option
Euclidean, [MV] measure_option
Gower, [MV] measure_option
maximum value, [MV] measure_option
Minkowski, [MV] measure_option
dissimilarity, matrix subcommand, [MV] matrix dissimilarity,
[P] matrix dissimilarity
distance matrices, [MV] matrix dissimilarity,
[P] matrix dissimilarity
distances, see dissimilarity measures
distances, estat subcommand, [MV] ca postestimation
distribution functions, [M-5] normal()
distributional diagnostic plots, [G-2] graph other
distributions,
examining, [D] pctile, [R] ameans, [R] centile,
[R] kdensity, [R] mean, [R] pksumm,
[R] summarize, [R] total
income, [R] inequality
plots, [R] cumul, [R] cusum, [R] diagnostic plots,
[R] dotplot, [R] histogram, [R] kdensity,
[R] ladder, [R] Iv, [R] spikeplot, [R] stem
standard population, [R] dstdize
testing equality of, [R] ksmirnov, [R] kwallis,
[R] ranksum, [R] signrank
testing for normality, [MV] mvtest normality,
[R] sktest, [R] swilk
transformations
to achieve normality, [R] boxcox, [R] ladder
to achieve zero skewness, [R] lnskew0
disturbance term, [XT] Glossary
division operator, see arithmetic operators
divisive hierarchical clustering methods, [MV] cluster,
[MV] Glossary
DLL, [P] plugin
Dmatrix() function, [M-5] Dmatrix()
do command, [R] do, [U] 16 Do-files
.do file, [U] 11.6 Filenaming conventions
dockable, set subcommand, [R] set
dockingguides, set subcommand, [R] set
documentation, [U] 1 Read this—it will help
keyword search on, [R] search, [U] 4 Stata’s help and search facilities
documenting data, [D] codebook, [D] labelbook,
[D] notes
_.docx*() functions, [M-5] _.docx*()
doedit command, [R] doedit
dojb() function, [D] datetime business calendars,
[FN] Date and time functions, [M-5] date()
dojc() function, [D] datetime, [FN] Date and time functions, [M-5] date()
doje() function, [D] datetime, [FN] Date and time functions, [M-5] date()
dojh() function, [D] datetime, [FN] Date and time functions, [M-5] date()
[U] 16 Do-files, [U] 18.2 Relationship between a program and a do-file
adding comments to, [P] comments
editing, [R] doedit
long lines, [P] #delimit, [U] 18.11.2 Comments and long lines in ado-files
dofm() function, [D] datetime, [FN] Date and time functions, [M-5] date()
dofq() function, [D] datetime, [FN] Date and time functions, [M-5] date()
dofw() function, [D] datetime, [FN] Date and time functions, [M-5] date()
dofy() function, [D] datetime, [FN] Date and time functions, [M-5] date()
domain sampling, [MV] alpha
Doornik–Hansen normality test, [MV] mvtest normality

dose–response models, [R] binreg, [R] glm, [R] logistic
dose–response trend, [PSS] power, [PSS] power trend
dot
dot plots, [G-2] graph dot
dot graph twoway subcommand, [G-2] graph twoway dot
dot plots, [G-2] graph dot, [G-2] graph twoway dot,
[G-3] area_options, [G-3] line_options
dotplot command, [R] dotplot
dotted lines, [G-4] linepatternstyle
double, [D] data types, [U] 12.2.2 Numeric storage types
double-exponential smoothing, [TS] tssmooth dexpnential
double-precision floating point number,
[U] 12.2.2 Numeric storage types
double quotes, [P] macro
doublebuffer, set subcommand, [R] set
doubly robust estimator, [TE] teffects intro,
[TE] teffects intro advanced, [TE] teffects aipw,
[TE] teffects ipwra, [TE] Glossary
dow() function, [D] datetime, [FN] Date and time
functions, [M-5] date(), [U] 24.5 Extracting
components of dates and times
doy() function, [D] datetime, [FN] Date and time
functions, [M-5] date()
dp, set subcommand, [D] format, [R] set
drawnorm command, [D]
draw() function, [M-5] drawnorm
function, [D]
dow() function, [D] datetime, [FN] Date and time
functions, [M-5] date()
dp, set subcommand, [D] format, [R] set
drawnorm command, [D] drawnorm
drift, [TS] Glossary
drop,

duplicates subcommand, [D] duplicates
label subcommand, [D] label
classutil subcommand, [P] classutil
cluster subcommand, [MV] cluster utility
constraint subcommand, [R] constraint
_estimates subcommand, [P] _estimates
estimates subcommand, [R] estimates store
forecast subcommand, [TS] forecast drop
graph subcommand, [G-2] graph drop
irf subcommand, [TS] irf drop
macro subcommand, [P] macro
data subcommand, [M-3] data drop
matrix subcommand, [P] matrix utility
notes subcommand, [D] notes
program subcommand, [P] program
_return subcommand, [P] _return
scalar subcommand, [P] scalar
serset subcommand, [P] serset
drop command, [D] drop
dropline, graph twoway subcommand, [G-2] graph
twoway dropline
dropout, [PSS] Glossary
dropping

graphs, [G-2] graph drop
programs, [P] discard
variables and observations, [D] drop
ds command, [D] ds
dstdize command, [R] dstdize
.dta file, [U] 11.6 Filenaming conventions
.dta file extension, technical description, [P] file
formats .dta
.dtasig file, [U] 11.6 Filenaming conventions
dual scaling, [MV] ca
Duda and Hart index stopping rules, [MV] cluster stop
dummy variables, see indicator variables, see indicators
Duncan’s multiple-comparison adjustment, see multiple
comparisons, Duncan’s method
dunnettprob() function, [FN] Statistical functions,
[M-5] normal()

Dunnett’s multiple comparison adjustment, see multiple
comparisons, Dunnett’s method
Dunnett’s multiple range distribution,
cumulative, [FN] Statistical functions
inverse cumulative, [FN] Statistical functions
_dup(#), display directive, [P] display
duplicate observations,
dropping, [D] duplicates
identifying, [D] duplicates
duplicates

drop command, [D] duplicates
drawnorm command, [D] duplicates
list command, [D] duplicates
report command, [D] duplicates
tag command, [D] duplicates
duplicating
clustered observations, [D] expandcl
observations, [D] expand
duplication matrix, [M-5] Dmatrix()
duration analysis, see survival analysis
Durbin–Watson statistic, [R] regress postestimation
time series, [TS] prais
durbinalt, estat subcommand, [R] regress
postestimation time series
Durbin’s alternative test, [R] regress postestimation
time series
dvech, mgarch subcommand, [TS] mgarch dvech
dwatson, estat subcommand, [R] regress
postestimation time series
dydx, command, [R] dydx
dynamic

correlation model, [TS] mgarch,
[TS] mgarch dec
factor model, [TS] dfactor, [TS] dfactor
postestimation, also see state-space model
forecast, [TS] arch, [TS] arfima, [TS] fcast
compute, [TS] fcast graph, [TS] forecast,
[TS] forecast adjust, [TS] forecast clear,
[TS] forecast coevector, [TS] forecast
create, [TS] forecast describe, [TS] forecast
drop, [TS] forecast estimates, [TS] forecast
exogenous, [TS] forecast identity, [TS] forecast
list, [TS] forecast query, [TS] forecast solve,
forecasts and simulations
model, [XT] Glossary
panel-data regression, [XT] xtabond, [XT] xtdpd,
[XT] xtdpdsys
regression model, [TS] arfima, [TS] arima,
[TS] var
structural simultaneous equations, [TS] var svar
dynamic-multiplier function, [TS] irf, [TS] irf cghraph,
[TS] irf create, [TS] irf ctable, [TS] irf ograph,
.dynamcmev built-in class function, [P] class
estimates

option

E-class command, [P]
program

economist

see

EBCDIC files, [D]

edittozerotol() see

EE estimator,

editvalue() function, [M-5]

macro extended function, [P]

edittoint() function, [M-5]

edittointtol() function, [M-5]

function, [M-5]

function, [M-5]

e() function, [FN] Programming functions, [M-5]

stored results, [P] ereturn, [P] _estimates,
[P] return, [R] stored results,
[U] 18.8 Accessing results calculated by other programs, [U] 18.9 Accessing results calculated by estimation commands, [U] 18.10.2 Storing results in e()

e(functions) macro extended function, [P] macro

e(macros) macro extended function, [P] macro

e(matrices) macro extended function, [P] macro

e(sample) function, [FN] Programming functions, [P] ereturn, [P] return

e(sample), resetting, [R]

estimates save

e(scalars) macro extended function, [P] macro

EB, see empirical Bayes

EBCDIC files, [D] filefilter, [D] infile (fixed format),
[U] 21.2.9 If you have EBCDIC data
e-class command, [P] program, [P] return, [R] stored results,
[U] 18.8 Accessing results calculated by other programs
economist
e
data

doedit commands, [U] 10 Keyboard use
data, [D] edit, [D] generate, [D] merge, [D] recode
files while in Stata, [R] doedit
graphs, [G-1] graph editor
output, [U] 15 Saving and printing output—log files

__editmissing() function, [M-5] editmissing()

editmissing() function, [M-5] editmissing()

__edittoint() function, [M-5] edittoint()

edittoint() function, [M-5] edittoint()

__edittointtol() function, [M-5] edittointtol()

edittointtol() function, [M-5] edittointtol()

__edittozero() function, [M-5] edittozero()

edittozero() function, [M-5] edittozero()

__edittozero() function, [M-5] edittozero()

edittozero() function, [M-5] edittozero()

__editvalue() function, [M-5] editvalue()

editvalue() function, [M-5] editvalue()

EE estimator, see estimating-equation estimator
effect
detection of, see minimum detectable effect size
minimum detectable, see minimum detectable effect size

size, [PSS] Glossary, [ST] Glossary, also see \( \delta \)
treatment, see treatment effects
effective sample size, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesstats ess,
[BAYES] Glossary
effects, estat subcommand, [SVY] estat
effects,
direct, [SEM] estat tffects, [SEM] example 7,
[SEM] example 42g, [SEM] methods and formulas for sem, [SEM] Glossary
indirect, [SEM] estat tffects, [SEM] example 7,
[SEM] example 42g, [SEM] methods and formulas for sem, [SEM] Glossary
total, [SEM] estat tffects, [SEM] example 7,
[SEM] example 42g, [SEM] methods and formulas for sem, [SEM] Glossary
effect-size
curve, [PSS] Glossary
determination, [PSS] intro, [PSS] power,
[PSS] power onemean, [PSS] power twomeans,
[PSS] power pairedmeans, [PSS] power oneprop, [PSS] power twoprop,
[PSS] power pairedproportions, [PSS] power onevariance, [PSS] power twovariances,
[PSS] power onecorrelation, [PSS] power twocorrelations, [PSS] power oneway,
[PSS] power twoway, [PSS] power repeated,
[PSS] power cmh, [PSS] power mecc,
[PSS] power cox, [PSS] power exponential,
efficiency of Markov chain Monte Carlo,
[BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesgraph,
[BAYES] bayesstats ess
efficiency of MCMC, [BAYES] Glossary
efficiency, query subcommand, [R] query
efform, estat subcommand, [SEM] estat efform
eform_option, [R] efform_option
EGARCH, see exponential generalized autoregressive conditional heteroskedasticity
egen command, [D] egen, [MI] mi passive, [MI] mi

seq

EGLS, see estimated generalized least squares
__eigen_la() function, [M-5] eigensystem()
__eigensystem() function, [M-5] eigensystem()
eigensystem() function, [M-5] eigensystem()
__eigensystemselect*() functions,
[M-5] eigensystemselect()
eigensystemselect*() functions,
[M-5] eigensystemselect()
eigenvalue stability condition, [TS] estat arroots,
[TS] varstable, [TS] vecstable
eigenvalue stability index, [SEM] estat stable
eigenvalues, [M-5] eigensystem( ), [M-6] Glossary,
[MV] factor, [MV] factor postestimation,
[MV] pca, [MV] rotate, [MV] rotatemat,
__eigenvalues() function, [M-5] eigensystem()
eigenvalues() function, [M-5] eigensystem()
eigenvalues, matrix subcommand, [P] matrix eigenvalues
[MV] factor, [MV] factor postestimation
[MV] pca, [MV] rotate, [MV] rotatemat
[P] matrix symeigen
EIM, see expected information matrix
eim, see sem option vce()
eivreg command, [R] eivreg, [R] eivreg postestimation
el() function, [FN] Matrix functions, [P] matrix define
elimination matrix, [M-5] Lmatrix()
elipsis, [G-4] text
else command, [P] if
eltype() function, [M-5] eltype()
EM, see expectation-maximization algorithm
empirical Bayes, [IRT] irt 1pl postestimation, [IRT] irt
dexample, [SEM] intro 7, [SEM] methods and formulas for gsem, [SEM] predict after gsem
empirical cumulative distribution function, [R] cumul
emptycells, set subcommand, [R] set, [R] set emptycells
Encapsulated PostScript, [G-2] graph export,
encode command, [D] encode, [U] 23.2 Categorical string variables
encodings, [D] unicode, [D] unicode encoding,
[1] Glossary
end command, [M-3] end
end-of-line characters, [D] changeeol
ending a Stata session, [P] exit, [R] exit
endless loop, see loop, endless
endogeneity test, [R] ivregress postestimation
endogenous
[TE] treat, [TE] eteffects, [TE] etpoisson,
[TE] etregress
treatment-effects model, [SEM] example 46g
variable, [SEM] intro 4, [SEM] Glossary,
[SVY] svy estimation, [TS] Glossary,
[XT] Glossary
endogenous, estat subcommand, [R] ivregress postestimation
ends(), egen function, [D] egen
Engle’s LM test, [R] regres postestimation time series
Enhanced Metafile, [G-2] graph export
ensuring mi data are consistent, [MI] mi update
entering data, see importing data, see inserting data
interactively, see reading data from disk
environment macro extended function, [P] macro
environment variables (Unix), [P] macro
eolchar, set subcommand, [R] set
Epanechnikov kernel function, [G-2] graph twoway
kdensity, [G-2] graph twoway lpoly,
[R] kdensity, [R] lpoly, [R] qreg, [TE] tebalance
density, [TE] tebalance overid, [TE] teffects
overlap
epidemiological tables, [R] tabulate twoway
epidemiology, [R] epitab
epidemiology, [R] epitab, [ST] strate
epidemiology and related
Brier score decomposition, [R] brier
interrater agreement, [R] kappa
meta-analysis, [R] meta
pharmacokinetic data, see pharmacokinetic data
ROC analysis, see receiver operating characteristic analysis
standardization, [R] dstdize
symmetry and marginal homogeneity tests,
[R] symmetry
tables, [R] tabulate twoway
epsilondouble() function, [FN] Programming functions
epsfloat() function, [FN] Programming functions
egof, estat subcommand, [SEM] estat egof
eqtest, estat subcommand, [SEM] estat eqtest
equal FMI test, [MI] mi estimate, [MI] mi test, [MI] Glossary
equal-allocation design, see balanced design
equality of means tests, [MV] hotelling, [MV] manova,
[MV] mvtest means
equality operator, [U] 13.2.3 Relational operators
equality test of
binomial proportions, [R] bitest
coefficients, [R] pwcompare, [R] sureg, [R] test,
[R] testnl, [SVY] svy postestimation
distributions, [R] ksmirnov, [R] kwallis,
[R] ranksum, [R] signrank
margins, [R] margins, [R] pwcompare
means, [R] contrast, [R] esize, [R] pwmean,
[R] test, [R] ztest, [SVY] svy postestimation
medians, [R] ranksum
proportions, [R] bitest, [R] prtest
ROC areas, [R] roccomp, [R] rocreg
survivor functions, [ST] sts test
variances, [R] sdttest
equal-tailed credible interval, [BAYES] intro,
[BAYES] bayes, [BAYES] bayesmh,
[BAYES] bayesstats summary,
[BAYES] Glossary
estat, continued

eqtest command, [SEM] intro 7, [SEM] estat eqtest, [SEM] example 13
errorrate command, [MV] discrim estat,
[MV] discrim knn postestimation,
[MV] discrim lda postestimation, [MV] discrim logistic postestimation, [MV] discrim qda postestimation
esize command, [R] regress postestimation
factors command, [MV] factor postestimation
facweights command, [R] asprobit postestimation, [R] asprobit postestimation
firststage command, [R] ivregress postestimation
framework command, [SEM] intro 7, [SEM] estat framework, [SEM] example 11
ggof command, [SEM] intro 7, [SEM] estat ggof, [SEM] example 21
ginvariant command, [SEM] intro 7, [SEM] estat ginvaintar, [SEM] example 22
grdistances command, [MV] discrim lda postestimation, [MV] discrim qda postestimation
grmeans command, [MV] discrim lda postestimation


hbettest command, [R] regress postestimation
ic command, [R] estat, [R] estat ic
imtest command, [R] regress postestimation
inertia command, [MV] ca postestimation
kmo command, [MV] factor postestimation, [MV] pca postestimation
lceffects command, [SVY] estat

estat, continued

manova command, [MV] discrim lda postestimation
mfx command, [R] asclglogit postestimation, [R] asprobit postestimation
mindices command, [SEM] intro 7, [SEM] estat mindices, [SEM] example 5, [SEM] example 9
mvreg command, [MV] procstars postestimation
nproc command, [R] rocreg postestimation
overid command, [R] gmm postestimation, [R] ipoission postestimation, [R] ivregress postestimation
ovtest command, [R] regress postestimation
pairwise command, [MV] mds postestimation
period command, [TS] ucm postestimation
phtest command, [ST] stcgc PH-assumption tests
predict command, [R] exlogistic postestimation
profiles command, [MV] ca postestimation
quantiles command, [MV] mds postestimation
recovariance command, [ME] meqrlogit postestimation, [ME] meqrpoisson postestimation, [ME] mixed postestimation
report command, [IRT] estat report
rotate command, [MV] canon postestimation
rotatecompare command, [MV] canon postestimation, [MV] factor postestimation, [MV] pca postestimation
sargan command, [XT] xtabond postestimation, [XT] xtdpdp postestimation, [XT] xtdpdsys postestimation
sbknown command, [TS] estat sbknown
sbsingle command, [TS] estat sbsingle
scoretests command, [SEM] intro 7, [SEM] estat scoretests, [SEM] example 8
sd command, [SVY] estat
se command, [R] exlogistic postestimation, [R] exproison postestimation
size command, [SVY] estat

cmc command, [MV] factor postestimation, [MV] pca postestimation
stable command, [SEM] intro 7, [SEM] estat stable, [SEM] example 7
stdize: prefix command, [SEM] estat stdize, [SEM] example 16
strata command, [SVY] estat
stress command, [MV] mds postestimation
structure command, [MV] discrim lda postestimation, [MV] factor postestimation
subinertia command, [MV] mca postestimation
estat, continued

summarize command, [MV] ca postestimation,
    [MV] discrim estat, [MV] discrim
    knn postestimation, [MV] discrim lda
    postestimation, [MV] discrim logistic
    postestimation, [MV] discrim qda
    postestimation, [MV] factor postestimation,
    [MV] mca postestimation, [MV] mds
    postestimation, [MV] pca postestimation,
    [MV] procrustes postestimation, [R] estat,
    [R] estat summarize, [SEM] estat summarize
svyset command, [SVY] estat
tseffects command, [SEM] estat tseffects,
    [SEM] example 7, [SEM] example 42g
vce command, [R] estat, [R] estat vce, [SVY] estat
vif command, [R] regress postestimation
wcorrelation command, [ME] mixed
    postestimation, [XT] xtgls
estimate linear combinations of coefficients, see linear
    combinations of estimators

estimates mi subcommand, [MI] mi estimate, [MI] mi
    estimate using
estimated generalized least squares, [XT] xtgls,
    [XT] xtivreg, [XT] xtreg
_estimates
    clear command, [P] _estimates
    dir command, [P] _estimates
    drop command, [P] _estimates
    hold command, [P] _estimates
    unhold command, [P] _estimates
estimates
    clear command, [R] estimates store
    command, [R] suest, [SVY] svy postestimation
        introduction, [R] estimates
describe command, [R] estimates describe
dir command, [R] estimates store
drop command, [R] estimates store
esample command, [R] estimates save
for command, [R] estimates for
notes command, [R] estimates notes
query command, [R] estimates store
replay command, [R] estimates replay
restore command, [R] estimates store
save command, [R] estimates save
stats command, [R] estimates stats
store command, [R] estimates store
table command, [R] estimates table
title command, [R] estimates title
use command, [R] estimates save
estimates, forecast subcommand, [TS] forecast
    estimates
estimating-equation estimator, [TE] tseffects aipw,
    [TE] tseffects ipw, [TE] tseffects ipwra,

estimation
    allowed estimation commands, [MI] estimation
    Bayesian, see Bayesian estimation
commands, [P] ereturn, [P] _estimates,
    [U] 18.9 Accessing results calculated by
    estimation commands, [U] 26 Overview of
    Stata estimation commands
    allowing constraints in, [P] makecns
    eliminating stored information from, [P] discard
    obtaining predictions after, [P] _predict
    obtaining robust estimates, [P] _robust
    saving results from, [P] _estimates
degrees of freedom for coefficients, [MI] mi
    estimate
method, [SEM] Glossary
options, [R] estimation options, [SEM] gsem
    estimation options, [SEM] sem estimation
    options
postestimation dialog boxes, [R] postest
Posting VCE, [MI] mi estimate
predictions after, see predictions, obtaining after
    estimation
results,
    clearing, [P] ereturn, [P] _estimates,
    [R] estimates store
    listing, [P] ereturn, [P] _estimates
    saving, [P] _estimates
    storing, [P] ereturn
    storing and restoring, [R] estimates store
tables of, [R] estimates table
    sample, summarizing, [R] estat, [R] estat
    summarize
test after, [MI] mi estimate, [MI] mi test,
    [SVY] svy postestimation
estimators,
    covariance matrix of, [P] ereturn, [P] matrix
        get, [R] correlate, [R] estat, [R] estat vce,
        [U] 20.9 Obtaining the variance–covariance
        matrix
    linear combinations, [U] 20.13 Obtaining linear
        combinations of coefficients
    linear combinations of, [R] lincom
    nonlinear combinations of, [R] nlcom
eteffects command, [TE] eteffects, [TE] eteffects
    postestimation
etiologic fraction, [R] epitab
etpoisson command, [TE] etpoisson, [TE] etpoisson
    postestimation
etregress command, [TE] etregress, [TE] etregress
    postestimation
Euclidean dissimilarity measure,
    [MV] measure_option
Euclidean distance, [MV] Glossary
event, [ST] Glossary
    history analysis, see survival analysis
    of interest, [ST] Glossary
    probability, see failure probability
    Ex, [SEM] sem and gsem option covstructure( )
exact binomial test, see binomial test
exact DDF, see denominator degrees of freedom
exact statistics, [U] 26.14 Exact estimators
  - binary confidence intervals, [R] ci, [R] exlogistic, [R] roctab
  - centiles, [R] centile
  - confidence intervals for variances, [R] ci
  - indirect standardization, [R] dstdize
  - one-way analysis of variance, [R] oneway
  - regression, [R] exlogistic, [R] expoisson test,
    - binomial probability, [R] bitest
    - equality of distributions, [R] ksmirnov
    - equality of medians, [R] ranksum
  - Fisher’s, [R] tabulate twoway
  - symmetry and marginal homogeneity, [R] symmetry
  - tetrachoric correlations, [R] tetrachoric

exact test, [PSS] Glossary
example datasets, [U] 1.2.2 Example datasets
examples, duplicates subcommand, [D] duplicates
Excel, [U] 21 Entering and importing data
dates, [D] datetime
  - Microsoft, see Microsoft Excel
  - Microsoft, reading data from, [D] import excel,
    [D] odbc, [D] xmlsave, also see spreadsheets,
    transferring
  - Microsoft, write results to, [P] putexcel,
    [P] putexcel advanced
excel,
  - export subcommand, [D] import excel
  - import subcommand, [D] import excel
excess fraction, [R] epítab
exec(), odbc subcommand, [D] odbc
existence, confirm subcommand, [P] confirm
exit class program, [P] class exit
exit, class subcommand, [P] class exit
exit command, [P] capture, [P] exit, [R] exit,
  [U] 16.1.4 Error handling in do-files
exit() function, [M-5] exit()
exit Mata, [M-3] end
exiting Stata, see exit command
exlogistic command, [R] exlogistic, [R] exlogistic
  postestimation
exogeneity test, see endogeneity test
exogenous, forecast subcommand, [TS] forecast
  exogenous
exogenous variable, [SEM] intro 4, [SEM] Glossary,
#exp, [U] 11 Language syntax
exp() function, [FN] Mathematical functions,
  [M-5] exp()
exp_list, [SVY] svy bootstrap, [SVY] svy brr,
  [SVY] svy jackknife, [SVY] svy sdr,
  [TS] rolling
expand command, [D] expand
  - for mi data, [MI] mi expand
expand factor varlists, [P] fexpand
expand, mi subcommand, [MI] mi expand
expandcl command, [D] expandcl
expectation-maximization algorithm, [MI] mi impute
  mvm, [MI] Glossary
  - parameter trace files, [MI] mi ptrace
expected information matrix, [SEM] Glossary
experimental data, [MV] manova, [R] anova,
  [R] contrast, [R] correlate, [R] epitab,
  [R] kwallis, [R] logit, [R] mean, [R] regress,
  [R] summarize, [R] tabulate oneway,
  [R] tabulate twoway, [R] ttest, [R] ztest,
  [U] 12 Data, [U] 20 Estimation and
  postestimation commands, [U] 26.4 Structural
equation modeling (SEM), [U] 26.21 Multilevel
  mixed-effects models, [U] 26.22 Survival-time
  (failure-time) models
experimental group, [PSS] Glossary
correlation, see correlation, experimental-group
  mean, see means, experimental-group
  proportion, see proportions, experimental-group
  sample size, see sample-size
  standard deviation, see standard deviations,
  experimental-group
  variance, see variances, experimental-group
experimental study, [PSS] power, [PSS] Glossary
exploded logit model, [R] rologit
expoisson command, [R] expoisson, [R] expoisson
  postestimation
exponential
  - density, [FN] Statistical functions, [M-5] normal()
  - distribution, [FN] Statistical functions,
    [M-5] normal(), [ST] streg
  - function, [FN] Mathematical functions
    - generalized autoregressive conditional
      heteroskedasticity, [TS] arch
  - notation, [U] 12.2 Numbers
  - smoothing, [TS] tssmooth, [TS] tssmooth
    - exponential, [TS] Glossary
  - survival
    - regression, [ST] streg
  - test, [PSS] power exponential, [PSS] Glossary
exponential,
  - churdle subcommand, [R] churdle
  - power subcommand, [PSS] power exponential
  - tssmooth subcommand, [TS] tssmooth
  - exponential()
  - function, [FN] Statistical functions,
    [M-5] normal()
  - exponentialden() function, [FN] Statistical
    functions, [M-5] normal()
  - exponentialtail() function, [FN] Statistical
    functions, [M-5] normal()
  - exponentiated coefficients, [R] eform_option,
    [SEM] estat eform
export
delimited command, [D] import delimited
excel command, [D] import excel
sasxport command, [D] import sasxport
export, graph subcommand, [G-2] graph export
export, mi subcommand, [MI] mi export, [MI] mi export ice, [MI] mi export nhanes1
exporting results, [P] putexcel, [P] putexcel advanced
exposure odds ratio, [PSS] power mcc, [PSS] Glossary
variable, [ST] Glossary
exposure() option, see gsem option exposure()
Expression Builder, [U] 13.8 Using the Expression Builder
expressions, [M-2] exp, [P] matrix define,
[U] 13 Functions and expressions
extended
ASCII, [D] unicode, [D] unicode translate,
[1] Glossary
encoding conversion, [D] unicode convertfile,
[D] unicode translate
encodings, [D] unicode encoding
external, [M-2] declarations
[M-6] Glossary
extract diagonal, [M-5] diagonal(), [M-5] diag()
extract, mi subcommand, [MI] mi extract, [MI] mi replace0
extracting m=## data from mi data, [MI] mi extract,
[MI] mi select
extracting original data from mi data, [MI] mi extract extrapolation, [D] ipolate

F
density,
central, [FN] Statistical functions
noncentral, [FN] Statistical functions
distribution,
cumulative, [FN] Statistical functions
cumulative noncentral, [FN] Statistical functions
inverse cumulative, [FN] Statistical functions
inverse reverse cumulative, [FN] Statistical functions
inverse reverse cumulative noncentral, [FN] Statistical functions
F distribution, continued
reverse cumulative, [FN] Statistical functions
reverse cumulative noncentral, [FN] Statistical functions
noncentrality parameter, [FN] Statistical functions
test, [PSS] Glossary
F() function, [FN] Statistical functions,
[M-5] normal()
Facebook, see Stata on Facebook
analysis, [MV] alpha, [MV] canon, [MV] factor,
[MV] factor postestimation, [MV] Glossary, see confirmatory factor analysis
loading plot, [MV] scoreplot, [MV] Glossary
loadings, [MV] Glossary
model, [TS] dfactor
parsimony rotation, [MV] rotate, [MV] rotatemat,
[MV] Glossary
score plot, [MV] scoreplot
scores, [MV] factor postestimation, [MV] Glossary,
[SEM] intro 7, [SEM] example 14,
[SEM] methods and formulas for sem, [SEM] predict after sem
variables, [P] fvexpand, [P] matrix rownames,
[P] _rmcoll, [P] syntax, [P] unah,
[PSS] Glossary, [R] fvrefvar, [R] fsvset,
[U] 11.4.3 Factor variables, [U] 13.9 Indicator values for levels of factor variables,
[U] 14.2.2 Two-part names, [U] 20.11 Accessing estimated coefficients, [U] 25 Working with categorical data and factor variables
factor command, [MV] factor, [MV] factor postestimation
factorial, [U] 11.4.3 Factor variables
design, [MV] manova, [R] anova
function, [FN] Mathematical functions
factorial() function, [M-5] factorial()
factormat command, [MV] factor, [MV] factor postestimation
factors, estat subcommand, [MV] factor postestimation
factor-variable
notation, [SEM] intro 3
settings, [R] fsvset
facweights, estat subcommand, [R] asmprobit
postestimation, [R] asprobit postestimation
failure
probability, [PSS] power exponential, [PSS] power logrank
tables, [ST] ltable
time, see survival analysis
failure–success proportion, [PSS] power pairedproportions
failure-time model, see survival analysis
false-negative result, see type II error
false-positive rate, [R] estat classification, [R] roc,
[R] rocreg, [R] rocreg postestimation,
[R] rocregplot
false-positive result, see type I error

family
Bernoulli, [SEM] methods and formulas for gsem
beta, [SEM] methods and formulas for gsem
binomial, [SEM] methods and formulas for gsem
distribution, [SEM] Glossary
exponential, [SEM] methods and formulas for gsem
gamma, [SEM] methods and formulas for gsem
Gaussian, [SEM] methods and formulas for gsem
loglogistic, [SEM] methods and formulas for gsem
lognormal, [SEM] methods and formulas for gsem
multinomial, [SEM] methods and formulas for gsem
negative binomial, [SEM] methods and formulas for gsem
ordinal, [SEM] methods and formulas for gsem
Poisson, [SEM] methods and formulas for gsem
Weibull, [SEM] methods and formulas for gsem

option, see gsem option family()

FAQs, [U] 3.2.1 The Stata website (www.stata.com)
search, [R] search, [U] 4.8.4 FAQ searches

fastscroll, set subcommand, [R] set
favorspeed() function, [M-5] favorspeed()
obufget() function, [M-5] bufio()
obufput() function, [M-5] bufio()
fcast compute command, [TS] fcast compute
fcast graph command, [TS] fcast graph
_fclose() function, [M-5] fopen()
fclose() function, [M-5] fopen()
FCS, see fully conditional specification
Fden() function, [FN] Statistical functions,
[M-5] normal()
feasible generalized least squares, [R] reg3, [R] sureg,
[SEM] intro 4, [TS] dfgls, [TS] prais, [TS] var,
[XT] xtreg, [XT] xtivreg, [XT] xtdfe
feasible generalized nonlinear least squares, [R] nlsw
feedback loops, [SEM] estat stable, [SEM] estat
teffects
fences, [R] lv
ferrotext() function, [M-5] ferrotext()
FEVD, see forecast-error variance decomposition
_fft() function, [M-5] fft()
fft() function, [M-5] fft()
_fget() function, [M-5] fopen()
getic() function, [M-5] fopen()
_fgetmatrix() function, [M-5] fopen()
geticmatrix() function, [M-5] fopen()
_fgetnl() function, [M-5] fopen()
geticnl() function, [M-5] fopen()
FGLS, see feasible generalized least squares
FGNLS, see feasible generalized nonlinear least squares
fictional data, [SEM] Glossary
file
conversion, [D] changeeol, [D] filefilter
format, Stata, [P] file formats .dta
modification, [D] changeeol, [D] filefilter

file, continued
translation, [D] changeeol, [D] filefilter

file
close command, [P] file
open command, [P] file
query command, [P] file
read command, [P] file
seek command, [P] file
set command, [P] file

file, confirm subcommand, [P] confirm
file, find in path, [P] findfile
fileexists() function, [FN] Programming functions,
[M-5] fileexists()
filefilter() function, [FN] Programming functions
filename manipulation, [M-5] adosubdir(),
[M-5] pathjoin()
filenames, displaying, [D] dir
fileread() function, [FN] Programming functions
filereaderror() function, [FN] Programming functions
files,
checksum of, [D] checksum
comparison, [D] cf
compressing, [D] zip

filefilter() function, [FN] Programming functions
copying and appending, [D] copy
display contents of, [D] type
downloading, [D] checksum, [D] copy,
[ R] adoupdate, [R] net, [R] sj, [R] ssc,
[ R] update, [U] 28 Using the Internet to keep
up to date
erasing, [D] erase
exporting, see exporting data
extensions, [U] 11.6 Filenaming conventions
importing, see importing data
loading, [D] use
names, [U] 11.6 Filenaming conventions,
[U] 18.3.11 Constructing Windows filenames
by using macros
opening, [P] window programming, [P] window
fopen
reading text or binary, [P] file
saving, [D] save, [P] window programming,
[P] window fopen
uncompressing, [D] zip
writing text or binary, [P] file

filewrite() function, [FN] Programming functions
file
areas, dimming and brightening, [G-2] graph
twoway histogram, [G-2] graph twoway
kdensity, [G-4] colorstyle
fill, continued
color, setting, [G-3] region_options
fill(), egen function, [D] egen
fillin command, [D] fillin
filling in values, [ST] stfill
_fillmissing() function, [M-5] _fillmissing()
_filters, [TS] tfilter, also see smoothers
Baxter–King, [TS] tfilter bk
Butterworth, [TS] tfilter bw
Christiano–Fitzgerald, [TS] tfilter cf
Hodrick–Prescott, [TS] tfilter hp
final, [M-2] class
findexternal() function, [M-5] findexternal()
findfile command, [P] findfile
findfile() function, [M-5] findfile()
finding file in path, [P] findfile
finding variables, [D] lookfor
first-order latent variables, [SEM] Glossary
firststage, estat subcommand, [R] ivregress postestimation
fisher, xtunitroot subcommand, [XT] xtunitroot
z test, [PSS] power onecorrelation, [PSS] power twocorrelations, [PSS] Glossary
z transformation, [PSS] power onecorrelation, [PSS] power twocorrelations, [PSS] Glossary
Fisher-type test, [XT] xtunitroot
fixed effects, [PSS] Glossary
F-keys, [U] 10 Keyboard use
flat prior, see noninformative prior
flat.prior() suboption, [BAYES] bayesmh evaluators
flexible functional form, [R] boxcox, [R] fp, [R] mfp
flist command, [D] list
float, [D] data types, [U] 12.2.2 Numeric storage types, [U] 13.12 Precision and problems therein
floatround() function, [M-5] floatround()
floatwindows, set subcommand, [R] set flong
_flopin() function, [M-5] lapack()
_floput() function, [M-5] lapack()
FML, see fraction missing information
%formats, [D] format, [U] 12.5 Formats: Controlling how data are displayed
fmtwidth() function, [FN] Programming functions, [M-5] fmtwidth()
folders, see directories
follow-up, [PSS] Glossary
studies, see incidence studies study, see cohort study fonts, in graphs, [G-4] text
footnote, ml subcommand, [R] ml _fopen() function, [M-5] fopen()
fopen() function, [M-5] fopen()
fopen, window subcommand, [P] window programming, [P] window fopen
for, estimates subcommand, [R] estimates for forcecorrelations option, see sem option forcecorrelations
forcecoanchor option, see gsem option forcecoanchor, see sem option forcecoanchor
forceconditional option, see sem option forceconditional
foreach command, [P] foreach
cast, [G-2] graph other
forecast, [TS] forecast
adjust command, [TS] forecast adjust
clear command, [TS] forecast clear
coevector command, [TS] forecast coevector
create command, [TS] forecast create
describe command, [TS] forecast describe
drop command, [TS] forecast drop
estimates command, [TS] forecast estimates
forecast, continued

exogenous command, [TS] forecast exogenous
identity command, [TS] forecast identity
list command, [TS] forecast list
query command, [TS] forecast query
solve command, [TS] forecast solve

forecast,
ARCH model, [TS] arch postestimation
ARFIMA model, [TS] arfima postestimation
ARIMA model, [TS] arima postestimation
dynamic-factor model, [TS] dfactor postestimation
[U] 20.20 Dynamic forecasts and simulations
Markov-switching model, [TS] mswitch postestimation
MGARCH model, see multivariate GARCH postestimation
standard error of, [R] regress multivariate GARCH postestimation
state-space model, [TS] sspace postestimation
structural vector autoregressive model, [TS] var svar postestimation
unobserved-components model, [TS] ucm postestimation
vector autoregressive model, [TS] var postestimation
vector error-correction model, [TS] vec postestimation
foreground color, [G-4] schemes intro
format command, [D] format
format, confirm subcommand, [P] confirm
format macro extended function, [P] macro format settings, [R] set cformat
format width, [M-5] fmtwidth( )
formats, [D] datetime, [D] describe, [D] format, [D] varmanage, [U] 12.5 Formats: Controlling how data are displayed, [U] 20.8 Formatting the coefficient table, [U] 24.3 Displaying dates and times
formatted data, reading, see importing data
formatting contents of macros, [P] macro formatting statistical output, [D] format

forum, [U] 3.2.4 The Stata forum
forvalues command, [P] forvalues
forward operator, [TS] Glossary
fourfold tables, [R] epitab
Fourier transform, [M-5] fft( )
FoxPro, reading data from, [U] 21.4 Transfer programs
fp
  generate command, [R] fp
  plot command, [R] fp postestimation
  predict command, [R] fp postestimation
  prefix command, [R] fp, [R] fp postestimation
FPC, see finite population correction
fpfit, graph twoway subcommand, [G-2] graph twoway fpfit
fpfitci, graph twoway subcommand, [G-2] graph twoway fpfitci
  _fput() function, [M-5] fopen( )
  fput() function, [M-5] fopen( )
  _fputmatrix() function, [M-5] fopen( )
  fpputmatrix() function, [M-5] fopen( )
fracplot command, [R] mfp postestimation
fracpred command, [R] mfp postestimation
fracreg command, [R] fracreg, [R] fracreg postestimation
fraction defective, [R] qc
fraction option, [G-2] graph twoway histogram
fractional polynomial regression, [R] fp
  multivariable, [R] mfp
fractional sample size, see sample-size
fractionally integrated autoregressive moving-average model, [TS] estat acplot, [TS] psdensity
 frailty, see shared frailty
framework, estat subcommand, [SEM] estat framework
  _fread() function, [M-5] fopen( )
  fread() function, [M-5] fopen( )
freduse command, [TS] areg postestimation free, constraint subcommand, [R] constraint frequencies,
  creating dataset of, [D] collapse, [D] contract
  graphical representation, [R] histogram, [R] kdensity
frequency option, [G-2] graph twoway histogram
frequency table, [XT] xttab
frequency weight, [U] 11.1.6 weight,
  [U] 20.23.1 Frequency weights
functions, continued
combinatorial, [FN] Mathematical functions
creating dataset of, [D] collapse, [D] obs
cumulative distribution, [R] cumul
date, [U] 24.5 Extracting components of dates and times
date and time, [FN] Date and time functions
derivatives and integrals of, [R] dydx
estimable, [R] margins
evaluator program, [R] gmm, [R] nl, [R] nsur
extended macro, [P] char, [P] display, [P] macro,
[P] macro lists, [P] mels
fractional polynomial, [R] fp, [R] mfp
graphing, [D] range, [G-2] graph twoway function
index, [R] logistic postestimation, [R] logit
postestimation, [R] probit postestimation
kernel, [R] kdensity, [R] ipoly
link, [R] betareg, [R] glm
mathematical, [FN] Mathematical functions
matrix, [FN] Matrix functions, [P] matrix define,
[U] 14.8 Matrix functions
maximizing likelihood, [R] maximize, [R] ml
 naming convention, [M-1] naming
obtaining help for, [R] help
orthogonalization, [R] orthog
parameters, [R] nlcom
passing to functions, [M-2] ftot
 piecewise cubic and piecewise linear, [R] mkspline
prediction, [R] predict, [R] predictnl
production and cost, [R] frontier
programming, [FN] Programming functions
random-number, [D] generate, [FN] Random-number functions
[r] set rng, [R] set seed
statistical, [FN] Statistical Functions
string, [FN] String functions
time-series, [FN] Selecting time-span functions
trigonometric, [FN] Trigonometric functions
underscore, [M-6] Glossary
variance, [R] glm
fvexpand command, [P] fvexpand
fvlabel, set subcommand, [R] set, [R] set
showbaselevels
fvrevar command, [R] fvrevar
fvset
 base command, [R] fvset
clear command, [R] fvset
design command, [R] fvset
report command, [R] fvset
command for mi data, [MI] mi XXXset
mi subcommand, [MI] mi XXXset
fvstandard option, see gsem option fvstandard
fvunab command, [P] unab
fvwrap() option, see sem option fvwrap()
fvwrap, set subcommand, [R] set, [R] set
showbaselevels
fvwrapon() option, see sem option fvwrapon()
fvwrapon, set subcommand, [R] set, [R] set showbaselevels
[fweight=exp] modifier, [U] 11.1.6 weight, [U] 20.23.1 Frequency weights
__fwrite() function, [M-5] fopen()
fwrite() function, [M-5] fopen()
fxsize() option, [G-2] graph combine
fysize() option, [G-2] graph combine

G

g-prior, see Zellner’s g-prior
g2 inverse of matrix, [P] matrix define, [P] matrix svd
gamma
density function, [FN] Statistical functions
   incomplete, [FN] Statistical functions
distribution
cumulative, [FN] Statistical functions
   inverse cumulative, [FN] Statistical functions
   inverse reverse cumulative, [FN] Statistical functions
   reverse cumulative, [FN] Statistical functions
regression, [SEM] intro 5, [SEM] Glossary
gamma() function, [M-5] factorial()
gamma option, see gsem option gamma

gammaden() function, [FN] Statistical functions, [M-5] normal()
gammap() function, [FN] Statistical functions, [M-5] normal()
gammaptail() function, [FN] Statistical functions, [M-5] normal()
gap() option, [G-2] graph twoway histogram

GARCH, see generalized autoregressive conditional heteroskedasticity
Gauss, reading data from, [U] 21.4 Transfer programs
Gauss–Hermite quadrature, [IRT] Glossary, see quadrature, Gauss–Hermite

Gaussian kernel function, [G-2] graph twoway
density, [TE] tebalance overid, [TE] t_eff

Gaussian regression, [SEM] Glossary
GEE, see generalized estimating equations
__geigen_la() function, [M-5] geigensystem()
  _geigenselect*() functions, [M-5] geigensystem()
geigensystem() function, [M-5] geigensystem()
  _geigensystem_la() function, [M-5] geigensystem()
geigensystemselect*() functions, [M-5] geigensystem()
generalized
  autoregressive conditional heteroskedasticity, [TS] arch, [TS] Glossary
generalized, continued
eigensystem, [M-5] geigensystem()
eigenvalues, [M-6] Glossary
gamma survival regression, [ST] streg
Hessenberg decomposition, [M-5] ghessenbergd()
   inverse of matrix, [P] matrix define, [P] matrix svd
least squares,
estimated, see estimated generalized least squares feasible, see feasible generalized least squares
least-squares estimator, [TS] prais, [TS] Glossary
linear latent and mixed models, [R] gllamm
linear mixed model, [ME] me, [ME] Glossary
linear models, [R] binreg, [R] fracreg, [R] glm,
linear response functions, [SEM] Glossary
   method of moments, [P] matrix accum,
   [XT] xtdp, [XT] xtdpsys, see gmm command
negative binomial regression, [R] nbreg, [SVY] svy estimation
partial credit model, [IRT] Glossary
responses, combined, [SEM] example 34g
Schur decomposition, [M-5] gschur()
SEM, [SEM] Glossary

generate,
   cluster subcommand, [MV] cluster generate
   icd10 subcommand, [D] icd10
   icd9 subcommand, [D] icd9
   icd9p subcommand, [D] icd9
   sts subcommand, [ST] sts generate

generate command, [D] generate, [MI] mi passive,
   [MI] mi xeq

generate functions, adding, [MV] cluster programming subroutines
generating data, [D] egen, [D] generate
   generating variables, [ST] stgen, [ST] sts generate
   get,
      constraint subcommand, [R] constraint
      net subcommand, [R] net

getmata command, [D] putmata
getting started, [U] 1 Read this—it will help
Getting Started with Stata manuals, [U] 1.1 Getting Started with Stata
   keyword search of, [U] 4 Stata’s help and search facilities
gettoken command, [P] gettoken
ggof, estat subcommand, [SEM] estat ggof
ghalton() function, [M-5] halton()
  _ghessenberg() function, [M-5] ghessenberg()
  ghessenberg() function, [M-5] ghessenberg()
  _ghessenberg_la() function, [M-5] ghessenberg()
ghk() function, [M-5] ghk()
ghkfast() function, [M-5] ghkfast()
ghkfast_i() function, [M-5] ghkfast()
ghkfast_init() function, [M-5] ghkfast()
ghkfast_init_()() function, [M-5] ghkfast()
ghkfast_query_()() function, [M-5] ghkfast()
ghk_init() function, [M-5] ghk()
ghk_init_()() function, [M-5] ghk()
ghk_query_npts() function, [M-5] ghk()
GHW, see quadrature, Gauss–Hermite, see Gauss–Hermite quadrature
Gibbs sampler, see Gibbs sampling
Gibbs sampling, [BAYES] intro, [BAYES] bayesmh,
[BAYES] Glossary
ginvvariant, estat subcommand, [SEM] estat
ginvvariant
ginvvariant() option, see sem option ginvvariant()
GJR, see threshold autoregressive conditional heteroskedasticity
gladder command, [R] ladder
GLLAMM, see generalized linear latent and mixed models
gllamm command, [R] gllamm
GLM, see generalized linear models
glm command, [R] glm, [R] glm postestimation
GLME, see generalized linear mixed-effects model
GLMM, see generalized linear mixed model
Global, class prefix operator, [P] class
global command, [P] macro, [U] 18.3.2 Global macros,
[U] 18.3.10 Advanced global macro manipulation
global variable, [M-2] declarations,
glsaccum, matrix subcommand, [P] matrix accum
GMM, see generalized method of moments
gmm command, [R] gmm, [R] gmm postestimation
gnbreg command, [R] nbreg, [R] nbreg postestimation
gof, estat subcommand, [R] estat gof, [R] poisson postestimation,
  [SEM] estat gof, [SVY] estat
Gompertz survival regression, [ST] streg
Gönen and Heller’s K, [ST] stcox postestimation
Goodman and Kruskal’s gamma, [R] tabulate
goodness of fit, [R] brier, [R] diagnostic plots,
  [R] estat gof, [R] ksmirnov, [R] linktest,
  [R] logistic postestimation, [R] lrtest,
  [R] poisson postestimation, [R] regr
  postestimation, [SEM] intro 7, [SEM] estat
gogof, [SEM] estat ggof, [SEM] estat gof,
goodness of fit, continued
  [SEM] example 3, [SEM] example 4,
  [SEM] Glossary, [SVY] estat, also see deviance residual, also see normal distribution
  and normality, test for
goto, [M-2] goto
Gower coefficient similarity measure,
  [MV] measure_option
GPCM, see generalized partial credit model
gpcm, irt subcommand, [IRT] irt pcm
  .gph file, [U] 11.6 Filenaming conventions
  gph files
    describing contents, [G-2] graph describe
    graded response model, [IRT] Glossary
gradient option, [R] maximize
grammar, [M-2] syntax
Granger causality, [TS] vargranger, [TS] Glossary
  graph
    bar command, [G-2] graph bar
    box command, [G-2] graph box
    close command, [G-2] graph close
    combine command, [G-2] graph combine
    command, [G-2] graph
copy command, [G-2] graph copy
describe command, [G-2] graph describe
dir command, [G-2] graph dir
display command, [G-2] graph display
dot command, [G-2] graph dot,
  [G-3] area_options, [G-3] line_options
drop command, [G-2] graph drop
elexport command, [G-2] graph export
hbar command, [G-2] graph hbar
hbox command, [G-2] graph hbox
matrix command, [G-2] graph matrix
pie command, [G-2] graph pie
play command, [G-2] graph play
print command, [G-2] graph print,
  [G-3] pr_options
query command, [G-2] graph query
rename command, [G-2] graph rename
replay command, [G-2] graph replay
save command, [G-2] graph save
set command, [G-2] graph set
set print command, [G-2] graph set
stow command, [G-2] graph stow
stow area command, [G-2] graph stow area
stow bar command, [G-2] graph stow bar
stow command, [G-2] graph stow
stow connected command, [G-2] graph stow
  connected
twoway contour command, [G-2] graph twoway
twoway contourline command, [G-2] graph twoway contourline
twoway dot command, [G-2] graph twoway dot
twoway dropline command, [G-2] graph twoway dropline
twoway fpfit command, [G-2] graph twoway fpfit
graph, continued

twoway fpfitci command, [G-2] graph twoway
fpfitci
twoway function command, [G-2] graph twoway
function
twoway histogram command, [G-2] graph twoway
histogram
twoway kdensity command, [G-2] graph twoway
kdensity
twoway lfit command, [G-2] graph twoway lfit
twoway lfitci command, [G-2] graph twoway
lfitci
twoway line command, [G-2] graph twoway line
twoway lowess command, [G-2] graph twoway
lowess
twoway lpoly command, [G-2] graph twoway
lpoly
twoway lpolyci command, [G-2] graph twoway
lpolyci
twoway mband command, [G-2] graph twoway
mband
twoway mspline command, [G-2] graph twoway
mspline
twoway pcarrow command, [G-2] graph twoway
pcarrow
twoway pcarrowi command, [G-2] graph twoway
pcarrowi
twoway pcbarrow command, [G-2] graph twoway
pcbarrow
twoway pccapsym command, [G-2] graph twoway
pccapsym
twoway pci command, [G-2] graph twoway
pci
twoway pcscatter command, [G-2] graph twoway
pcscatter
twoway pcspike command, [G-2] graph twoway
pcspike
twoway qfit command, [G-2] graph twoway qfit
twoway qfitci command, [G-2] graph twoway
qfitci
twoway rarea command, [G-2] graph twoway
rarea
twoway rbar command, [G-2] graph twoway rbar
twoway rcap command, [G-2] graph twoway rcap
twoway rcapsym command, [G-2] graph twoway
rcapsym
twoway rconnected command, [G-2] graph twoway
rconnected
twoway rline command, [G-2] graph twoway
rline
twoway rscatter command, [G-2] graph twoway
rscatter
twoway rspike command, [G-2] graph twoway
rspike
twoway scatter command, [G-2] graph twoway
scatter
twoway scatteri command, [G-2] graph twoway
scatteri
twoway spike command, [G-2] graph twoway
spike

diagnostic plots
adjusted Kaplan–Meier survivor curves, [ST] sts
baseline hazard and survivor, [ST] stcox, [ST] sts
cumulative hazard function, [ST] stcurve, [ST] sts
graph hazard function, [ST] ltable, [ST] stcurve, [ST] sts
graph Kaplan–Meier survivor curves, [ST] stcox
PH-assumption tests, [ST] sts
log-log curve, [ST] stcox PH-assumption tests
survivor function, [ST] stcurve, [ST] sts
graphical user interface, [IRT] Control Panel,
[P] dialog programming, [PSS] GUI,
[PSS] power, [SEM] Builder, [SEM] Builder,
generalized, [SEM] Glossary
examples of, [U] 2 A brief description of Stata
graphics,
graphregion() option, [G-3] region_options
graphs,
added-variable plot, [R] regress postestimation
diagnostic plots
adjusted partial residual plot, [R] regress
postestimation diagnostic plots
augmented component-plus-residual plot, [R] regress
postestimation diagnostic plots
augmented partial residual plot, [R] regress
postestimation diagnostic plots
autocorrelations, [TS] corrgram
Bayesian, [BAYES] bayesgraph
binary variable cumulative sum, [R] csum
biplot, [MV] biplot, [MV] ca postestimation plots
box, [TE] tebalance box
CA dimension projection, [MV] ca postestimation
plots
cluster tree, see graphs, dendrogram
component-plus-residual, [R] regress postestimation
diagnostic plots
contrasts, see subentry margins
correlogram, [TS] corrgram
cross-correlogram, [TS] xcorr
graphs, continued

cross-sectional time-series data, [XT] xtdata, [XT] xline
cumulative distribution, [R] cumul
cumulative spectral density, [TS] cumsp

density-distribution sunflower, [R] sunflower
derivatives, [R] dydx
dotplot, [R] dotplot
diagnostic, [R] diagnostic plots
dotplot, [R] dotplot
diagnostic, [R] diagnostic plots
dotplot, [R] dotplot
eigenvalue
after discrimin lda, [MV] discrimin lda postestimation, [MV] screeplot
after factor, [MV] factor postestimation, [MV] screeplot
after manova, [MV] screeplot
after mca, [MV] screeplot
after mds, [MV] screeplot
after pca, [MV] pca postestimation, [MV] screeplot

error-bar charts, [R] serrbar
forecasts, [TS] fcast graph
fractional polynomial, [R] fp postestimation
functions, [D] obs, [D] range
histograms, [R] histogram, [R] kdensity
integrals, [R] dydx
interaction plots, [R] marginesplot
ladder-of-power histograms, [R] ladder
letter-value display, [R] lv
leverage-versus-(squared)-residual, [R] regress postestimation diagnostic plots
loading
after canncid, [MV] canncid, [MV] scoreplot
after discrimin lda, [MV] discrimin lda, [MV] discrimin lda postestimation, [MV] scoreplot
after factor, [MV] factor postestimation, [MV] scoreplot
after mca, [MV] scoreplot
after mds, [MV] scoreplot
after pca, [MV] pca postestimation, [MV] scoreplot
logistic diagnostic, [R] logistic postestimation, [R] lsens
loess smoothing, [R] loess
margins, [U] 20.19 Graphing margins, marginal effects, and contrasts
margins plots, [R] marginsplot
MDS configuration, [MV] mds postestimation plots
means and medians, [R] grmeanby

graphs, continued

normal probability, [R] diagnostic plots
overall look of, [G-4] schemes intro
parameterized curves, [D] range
parametric autocorrelation, [TS] estat acplot
parametric autocovariance, [TS] estat acplot
partial correlogram, [TS] corgram
partial residual, [R] regress postestimation diagnostic plots
partial-regression leverage, [R] regress postestimation diagnostic plots
periodogram, [TS] pergram
power and sample size, [PSS] power, graph
procrustes overlay, [MV] procrustes postestimation profile plots, [R] marginsplot
quality control, [R] qc
quantile, [R] diagnostic plots
quantile–normal, [R] diagnostic plots
quantile–quantile, [R] diagnostic plots
regression diagnostic, [R] regress postestimation diagnostic plots
residual versus fitted, [R] regress postestimation diagnostic plots
residual versus predictor, [R] regress postestimation diagnostic plots
rootograms, [R] spikeplot
saving, [G-3] saving_option
score
after canncid, [MV] canncid, [MV] scoreplot
after discrimin lda, [MV] discrimin lda, [MV] discrimin lda postestimation, [MV] scoreplot
after factor, [MV] factor postestimation, [MV] scoreplot
after mca, [MV] scoreplot
after mds, [MV] scoreplot
after pca, [MV] pca postestimation, [MV] scoreplot
scree
after canon, [MV] scoreplot
after ca, [MV] scoreplot
after discrimin lda, [MV] discrimin lda postestimation, [MV] scoreplot
after factor, [MV] factor postestimation, [MV] scoreplot
after manova, [MV] scoreplot
after mca, [MV] scoreplot
after mds, [MV] scoreplot
after pca, [MV] pca postestimation, [MV] scoreplot
Shepard diagram, [MV] mds postestimation plots
smoothing, [R] kdensity, [R] loess, [R] lpoly
spike plot, [R] spikeplot
term-and-leaf, [R] stem
time-series sunflower, [R] sunflower
suppressing, [G-3] nodraw_option
symmetry, [R] diagnostic plots
time-versus-concentration curve, [R] pk, [R] pksent
graphs, continued
  treatment-effects balance, [TE] tebalance box,
  [TE] tebalance density, [TE] tebalance overid
  treatment-effects overlap, [TE] teffects overlap
white-noise test, [TS] wntestb
gdistances, estat subcommand, [MV] discrim
  lds postestimation, [MV] discrim qda
  postestimation
greater than (or equal) operator, [U] 13.2.3 Relational
  operators
.csv file, [U] 11.6 Filenaming conventions
Greek letters, [G-4] text
Greenhouse–Geisser epsilon, [R] anova
Greenhouse–Geisser correction, see nonsphericity
  correction
Greenwood confidence intervals, [ST] st
  grid
definition, [G-4] gridstyle
  lines, [G-3] axis_label_options
  without ticks, [G-4] tickstyle
ggridstyle, [G-4] gridstyle
GRM, see graded response model
gm, irt subcommand, [IRT] irt grm, [IRT] irt grm
  postestimation
grmeanby command, [R] grmeanby
grmeans, estat subcommand, [MV] discrim lda
  postestimation
group(), egen function, [D] egen
group, estat subcommand, [ME] mecloglog
  postestimation, [ME] melogit postestimation,
  [ME] meprobit postestimation, [ME] mepoisson postestimation,
  [ME] meprobit postestimation, [ME] meqlogit postestimation,
  [ME] meqroissson postestimation,
  [ME] mestreg postestimation, [ME] mixed
  postestimation
group invariance test, [SEM] methods and formulas
  for sem
group() option, see sem option group()
group weights, [PSS] power trend
group-data regression, [R] intreg
grouping variables, generating, [MV] cluster generate
groups, graphs by, [G-3] by_option
groupvar, [U] 11.4 varlists
gsummarize, estat subcommand, [MV] discrim
  estat, [MV] discrim kkn postestimation,
  [MV] discrim lda postestimation, [MV] discrim qda
  postestimation
gs1 print color mapping, [G-2] set printcolor
gs2 print color mapping, [G-2] set printcolor
gs3 print color mapping, [G-2] set printcolor
  _gchurd() function, [M-5] gchurd()
gchurd() function, [M-5] gchurd()
  _gchurdgroupby() function, [M-5] gchurd()
  _gchurdgroupby_la() function, [M-5] gchurd()
gsem option, continued

inormal, [SEM] gsem family-and-link options
logit, [SEM] gsem family-and-link options
loglogistic, [SEM] gsem family-and-link options
lognormal, [SEM] gsem family-and-link options
maximize_options, [SEM] intro 12, [SEM] gsem estimation options
means(), [SEM] gsem model description options, [SEM] sem and gsem path notation
method(), [SEM] intro 8, [SEM] intro 9, [SEM] gsem estimation options
mlogit, [SEM] gsem family-and-link options
nbreg, [SEM] gsem family-and-link options
noanchor, [SEM] gsem model description options
noasis, [SEM] gsem model description options
nocapslatent, [SEM] gsem estimation options
noconstant, [SEM] gsem model description options
noestestimate, [SEM] gsem estimation options
nodevheader, [SEM] gsem reporting options
noheader, [SEM] gsem reporting options
notable, [SEM] gsem reporting options
ocloglog, [SEM] gsem family-and-link options
offset(), [SEM] gsem family-and-link options
ologit, [SEM] gsem family-and-link options
oprobit, [SEM] gsem family-and-link options
poisson, [SEM] gsem family-and-link options
probit, [SEM] gsem family-and-link options
pweights(), [SEM] gsem estimation options
regress, [SEM] gsem family-and-link options
reliability(), [SEM] intro 12, [SEM] gsem model description options, [SEM] sem and gsem option reliability()
startgrid(), [SEM] intro 12, [SEM] gsem estimation options
startvalues(), [SEM] intro 12, [SEM] gsem estimation options
variance(), [SEM] gsem model description options, [SEM] sem and gsem path notation
vce(), [SEM] intro 8, [SEM] intro 9, [SEM] gsem estimation options
weibull, [SEM] gsem family-and-link options
gsem postestimation commands, [SEM] intro 7
gsort command, [D] gsort
guessing, [IRT] Glossary
guessing parameter, [IRT] irt 3pl

GUI, see graphical user interface

H

hadamard() function, [FN] Matrix functions,
[P] matrix define
Hadamard matrix, [SVY] svy brr, [SVY] Glossary

Hadri Lagrange multiplier stationarity test, [XT] xtunitroot
hadri, xtunitroot subcommand, [XT] xtunitroot
half option, [G-2] graph matrix
halfyear() function, [D] datetime, [FN] Date and time functions, [M-5] date()
halfyearly() function, [D] datetime, [D] datetime translation, [FN] Date and time functions, [M-5] date()
_halton() function, [M-5] halton()
halton() function, [M-5] halton()
Halton set, [M-5] halton()

Hamann coefficient similarity measure, [MV] measure_option

Hammersley set, [M-5] halton()

Hansen’s J statistic, [R] gmm, [R] gmm postestimation

Hausman specification test, [R] hausman

Hausman–Taylor estimator, [XT] xthtaylor
Haver Analytics databases, reading data from, [D] import haver

haver import subcommand, [D] import haver

haverdir, set subcommand, [D] import haver, [R] set

hazard, [TE] etregress

corrections, [ST] Glossary

goof of, [ST] ltable, [ST] stcurve, [ST] sts graph

rate, [PSS] power exponential, [PSS] power logrank


tables, [ST] ltable

hazard-rate difference, [PSS] power exponential

hazards,

control-group, [PSS] power exponential,
[PSS] power logrank

experimental-group, [PSS] power exponential,
[PSS] power logrank

two-sample, [PSS] power exponential, [PSS] power logrank
hetprobit command, [R] hetprobit, [R] hetprobit postestimation
hettest, estat subcommand, [R] regress postestimation
hexadecimal report, [D] hexdump
hexdump command, [D] hexdump
Heywood case, [MV] Glossary
solution, [MV] Glossary
cluster analysis, [MV] cluster, [MV] clustermat, [MV] cluster linkage
clustering, [MV] Glossary
regression, [R] nestreg, [R] stepwise
samples, [R] anova, [R] glm, [R] lowneway, [R] areg
higher ASCII, see extended ASCII
higher-order models, see confirmatory factor analysis
region, [BAYES] intro, [BAYES] Glossary
Hilbert() function, [M-5] Hilbert()
Hildreth–Lu regression, [TS] prais
HILO, [M-5] byteorder() histogram command, [R] histogram
histogram, graph twoway subcommand, [G-2] graph twoway histogram
histograms, [G-2] graph twoway histogram, [R] histogram
dotplots, [R] dotplot
kernel density estimator, [R] kdensity
ladder-of-powers, [R] ladder
of categorical variables, [R] histogram
teeroots, [R] spikeplot
stem-and-leaf, [R] stem
histories, [G-2] graph bar, [G-2] graph box,
pie, [G-2] graph twoway histogram,
[G-3] by_option
hms() function, [D] datetime, [FN] Date and time
functions, [M-5] date()
Hodrick–Prescott filter, [TS] tsfilter, [TS] tsfilter hp
hofd() function, [D] datetime, [FN] Date and time
functions, [M-5] date()
hold,
  _estimates subcommand, [P] _estimates
  _return subcommand, [P] _return
Holm’s multiple-comparison adjustment, see multiple
comparisons, Holm’s method
dexpontential, [TS] tssmooth exponential,
[TS] tssmooth hwinters, [TS] tssmooth
shwinters, [TS] Glossary
homogeneity of variances, [R] oneway
homoskedasticity tests, [R] sdtest
Horst normalization, see Kaiser normalization
Hosmer–Lemeshow
delta chi-squared influence statistic, see delta chi-
squared influence statistic
delta deviance influence statistic, see delta deviance
influence statistic
goodness-of-fit test, [R] estat gof, [SVY] estat
hot, ssc subcommand, [R] ssc
hotelling command, [MV] hotelling
Hotelling’s
generalized T-squared statistic, [MV] manova
T-squared, [MV] hotelling, [MV] mvtest means,
[MV] Glossary
hours() function, [D] datetime, [FN] Date and time
functions, [M-5] date()
hp, tsfilter subcommand, [TS] tsfilter hp
HPD
credible interval, see highest posterior density
credible interval
  _hqr() function, [M-5] qrd()
_hqr() function, [M-5] qrd()
_hqrmultq() function, [M-5] qrd()
_hqrmultq1t() function, [M-5] qrd()
  _hqrd() function, [M-5] qrd()
_hqrd() function, [M-5] qrd()
  _hqrdp_1a() function, [M-5] qrd()
_hqrdq() function, [M-5] qrd()
_hqrdq1() function, [M-5] qrd()
_hqrdz() function, [M-5] qrd()
_hqrdz1() function, [M-5] qrd()
HRF, see human readable form
ht, xtunitroot subcommand, [XT] xtunitroot
http://www.stata.com, [U] 3.2.1 The Stata website
http://www.stata.com
httpproxy, set subcommand, [R] netio, [R] set
httpproxyauth, set subcommand, [R] netio, [R] set
httpproxyhost, set subcommand, [R] netio, [R] set
httpproxypport, set subcommand, [R] netio, [R] set
httpproxypw, set subcommand, [R] netio, [R] set
Huber weighting, [R] rreg
Huber/White/sandwich estimator of variance, see robust,
Huber/White/sandwich estimator of variance
human readable form, [D] datetime, [D] datetime
display formats, [D] datetime translation
hurdle regression, [R] churdle
Huynh–Feldt epsilon, [R] anova
hwinters, tssmooth subcommand, [TS] tssmooth
hwinters
hybrid
MH sampler, see hybrid MH sampling
MH sampling, [BAYES] intro
model, [IRT] Glossary
hybrid, irt subcommand, [IRT] irt hybrid, [IRT] irt
hybrid postestimation
hyperbolic functions, [FN] Statistical functions,
[M-5] sin()
hypergeometric() function, [FN] Statistical
functions, [M-5] normal()
hypergeometric,
cumulative distribution, [FN] Statistical functions
probability mass function, [FN] Statistical functions
hypergeometricp() function, [FN] Statistical
functions, [M-5] normal()
hyperparameters, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesmh
evaluators, [BAYES] bayesgraph,
[BAYES] Glossary
hyperprior, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesmh
evaluators, [BAYES] bayesgraph,
[BAYES] Glossary
hypertext help, [R] help, [U] 4 Stata’s help and search
facilities, [U] 18.11.6 Writing system help
hypothesis, [PSS] Glossary, also see null hypothesis
and alternative hypothesis
test, [PSS] Glossary, [SEM] test, [SEM] testnl, also
see null hypothesis and alternative hypothesis
testing, Bayesian, see Bayesian, hypothesis testing
hypothesized value, see null value

I()
I() function, [FN] Matrix functions, [M-5] I(),
[P] matrix define
i.i.d. assumption, see independent and identically
distributed sampling assumption
ibeta() function, [FN] Statistical functions,
[M-5] normal()
ibetatail() function, [FN] Statistical functions,
[M-5] normal()
ic, bayesstats subcommand, [BAYES] bayesstats
ic, estat subcommand, [R] estat, [R] estat ic
ICC, see item characteristic curve
icc command, [R] ice
imputation, continued
continuous, [MI] mi impute pmm, [MI] mi impute regress
with a limited range, [MI] mi impute intreg, [MI] mi impute truncreg
count data, [MI] mi impute nbreg, [MI] mi impute poisson
diagnostics, [MI] mi impute
interval regression, [MI] mi impute intreg
interval-censored data, [MI] mi impute intreg
linear regression, [MI] mi impute regress
logistic regression, [MI] mi impute logit
modeling, [MI] mi impute
monotone, [MI] mi impute, [MI] mi impute chained, [MI] mi impute monotone
multinomial logistic regression, [MI] mi impute mlogit
multiple, [MI] intro substantive
multivariate,
chained equations, [MI] mi impute, [MI] mi impute chained
normal, [MI] mi impute, [MI] mi impute mvn
negative binomial regression, [MI] mi impute nbreg
on subsamples, [MI] mi impute
ordered logistic regression, [MI] mi impute ologit
overdispersed count data, [MI] mi impute nbreg
passive, [MI] mi impute, [MI] mi impute chained
perfect prediction, [MI] mi impute
Poisson regression, [MI] mi impute poisson
predictive mean matching, [MI] mi impute, [MI] mi impute pmm
regression, [MI] mi impute, [MI] mi impute regress
semiparametric, [MI] mi impute pmm
step, [MI] intro substantive, [MI] mi estimate
transformations, [MI] mi impute
truncated data, [MI] mi impute truncreg
truncated regression, [MI] mi impute truncreg
user-defined, [MI] mi impute usermethod
imputations, recommended number of, [MI] intro substantive, [MI] mi estimate
imputed data, [MI] Glossary
imputed variables, see variables, imputed
intest, estat subcommand, [R] regress postestimation
in range qualifier, [P] syntax, [U] 11 Language syntax
in smcl, display directive, [P] display
inbase() function, [M-5] inbase()
incidence, [ST] Glossary
incidence rate, [ST] Glossary
estimation,
negative binomial regression, [R] nbreg, [R] tnreg, [R] zinb
postestimation, [R] contrast, [R] expoisson postestimation, [R] lincom
postestimation,
negative binomial regression, [R] nbreg postestimation, [R] tnreg postestimation, [R] zinb postestimation
Poisson regression, [R] poisson postestimation, [R] tpoisson postestimation, [R] zip postestimation
include_bitmap, set subcommand, [R] set
include command, [P] include
income distributions, [R] inequality
income tax rate function, [D] egen
incomplete
beta function, [FN] Statistical functions, [M-5] normal()
gamma function, [FN] Statistical functions, [M-5] normal()
observations, [MI] Glossary, see dropout
increment operator, [M-2] op_increment
independent of irrelevant alternatives, assumption, [R] clogit, [R] mlogit
relaxing assumption, [R] asclogit, [R] asmpoisit, [R] asprobit, [R] nlogit
test for, [R] hausman, [R] nlogit, [R] suest

independent and identically distributed, [TS] Glossary

index search, [R] search, [U] 4 Stata’s help and search facilities

index, mathematical functions, [M-4] statistical

matrix functions, [M-4] utility
statistical functions, [M-4] statistical
stopping rules, see stopping rules
utility functions, [M-4] utility

indexnot() function, [FN] String functions, [M-5] indexnot()

indicator variables, [R] tabulate oneway, [R] xi, [SEM] Glossary, also see factor variables

indicators, [U] 11.4.3 Factor variables

indirect standardization, [R] dstdize

individual-level treatment effect, [TE] Glossary

ineligible missing values, [MI] mi impute,
[MJ] Glossary

inequality measures, [R] inequality

inertia, [MV] Glossary, also see total inertia

inertia, estat subcommand, [MV] ca postestimation

infile command, [D] infile (fixed format), [D] infile (free format)

infix command, [D] infix (fixed format)

influence statistics, see delta beta influence statistic, see delta chi-squared influence statistic, see delta deviance influence statistic, see DFBETA, see LMAX value

`%infmt.`, [D] infile (fixed format)

information, [IRT] Glossary

criteria, see Akaike information criterion, see Bayesian information criterion

matrix, [P] matrix get, [R] correlate, [R] maximize

matrix test, [R] regress postestimation


Informix, reading data from, [U] 21.4 Transfer programs

inheritance, [M-2] class, [P] class

init, ml subcommand, [R] ml
init, ssc subcommand, [SEM] ssc

initial values, [SEM] Glossary, see starting values

initialization, class, [P] class

inlist() function, [FN] Programming functions

inner fence, [R] lq
innovation accounting, [TS] irf

input command, [D] input

input, matrix subcommand, [P] matrix define

input, obtaining from console in programs, see console, obtaining input from

input/output functions, [M-4] io

inputting data
from a file, see importing data, see reading data from disk
interactively, [D] edit, [D] input, also see editing data, also see importing data

irange() function, [FN] Programming functions

insert, odbc subcommand, [D] odbc
insobs command, [D] insobs
inspect command, [D] inspect
install,
net subcommand, [R] net
ssc subcommand, [R] ssc

installation
of official updates, [R] update, [U] 28 Using the Internet to keep up to date
of user-written commands (updating), [R] adoupdate
instance, [M-6] Glossary

instance, class, [P] class

instancemv built-in class function, [P] class

instance-specific variable, [P] class

instrument, [IRT] Glossary

instrumental-variables, [XT] Glossary

estimator, [XT] Glossary


int, [D] data types, [U] 12.2.2 Numeric storage types

int() function, [FN] Mathematical functions

integ command, [R] dydx

integer truncation function, [FN] Mathematical functions

integers, [M-5] trunc()

integrals, numeric, [R] dydx

integrated autoregressive moving-average model, [TS] estat acplot, [TS] psdensity

integrated process, [TS] Glossary


intensitystyle, [G-4] intensitystyle


effects, [PSS] Glossary
interaction, continued

expansion, [R] xi
plots, [R] marginsplot

intercept, [SEM] intro 4, [SEM] Glossary, also see
constraints, specifying

interface, query subcommand, [R] query

internal consistency test, [MV] alpha

International Components for Unicode, [D] unicode

Internet,
commands to control connections to, [R] netio
installation of updates from, [R] adoupdate, [R] net,
[R] sj, [R] update, [U] 28 Using the Internet to keep up to date
search, [R] net search
Stata, [U] 3.2.1 The Stata website (www.stata.com)
Stata Journal, [U] 3.4 The Stata Journal
Stata Press, [U] 3.3 Stata Press

interpolation, [D] interp

interquartile range, [R] lv

interrupting command execution, [U] 10 Keyboard use

interval
censoring, see imputation, interval-censored data
data, [XT] xintreg, [XT] Glossary

hypothesis test, [BAYES] Glossary

hypothesis testing, see Bayesian, hypothesis testing, interval

regression, [R] intreg, [SVY] svy estimation

regression model, [SEM] example 44g

interval regression, random-effects, [XT] xintreg
test, [BAYES] Glossary, see Bayesian, hypothesis testing, interval

intmethod() option, see gsem option intmethod()

intpoints() option, see gsem option intpoints()

intracluster correlation, [ME] Glossary, [R] icc, also see
estat icc command

intracluster correlation, see correlation, intracluster

intreg command, [R] intreg, [R] intreg

postestimation

inv() function, [FN] Matrix functions, [P] matrix define

invariance, [IRT] Glossary

invbinomial() function, [FN] Statistical functions,
[M-5] normal()

invbinomialtaill() function, [FN] Statistical functions,
[M-5] normal()

invchi2() function, [FN] Statistical functions,
[M-5] normal()

invchi2taill() function, [FN] Statistical functions,
[M-5] normal()

invchi2tail() function, [FN] Statistical functions,
[M-5] normal()

invdunnnettprob() function, [FN] Statistical functions,
[M-5] normal()

inverse

cumulative

beta distribution, [FN] Statistical functions

binomial function, [FN] Statistical functions

chi-squared distribution function, [FN] Statistical functions

exponential distribution, [FN] Statistical functions

F distribution, [FN] Statistical functions

incomplete gamma function, [FN] Statistical functions

Weibull distribution, [FN] Statistical functions

hyperbolic tangent transformation, see Fisher’s z transformation


noncentral

beta distribution, [FN] Statistical functions

chi-squared distribution function, [FN] Statistical functions

F distribution, [FN] Statistical functions

normal distribution function, [FN] Statistical functions

of matrix, [P] matrix define, [P] matrix svd

reverse cumulative

beta distribution, [FN] Statistical functions

binomial function, [FN] Statistical functions

chi-squared distribution function, [FN] Statistical functions

exponential distribution, [FN] Statistical functions

F distribution, [FN] Statistical functions

t distribution function, [FN] Statistical functions

Weibull distribution, [FN] Statistical functions

inverse-probability weighting, [TE] teffects intro,
[TE] teffects intro advanced, [TE] teffects ipw,
[TE] Glossary

inverse-probability-weighted regression adjustment,
[TE] teffects intro, [TE] teffects intro advanced,
[TE] teffects ipwra, [TE] Glossary

invexponential() function, [FN] Statistical functions,
[M-5] normal()

invexponentialtaill() function, [FN] Statistical functions,
[M-5] normal()

invF() function, [FN] Statistical functions,
[M-5] normal()

invF() function, [FN] Statistical functions,
[M-5] normal()

invF() function, [FN] Statistical functions,
[M-5] normal()

invFtaill() function, [FN] Statistical functions,
[M-5] normal()

invfft() function, [M-5] fft()

invfft() function, [M-5] fft()

invFtail() function, [FN] Statistical functions,
[M-5] normal()

invgammap() function, [FN] Statistical functions,
[M-5] normal()
invgammaptail() function, [FN] Statistical functions, [M-5] normal()
invhilbert() function, [M-5] Hilbert()
invibeta() function, [FN] Statistical functions, [M-5] normal()
invibetatail() function, [FN] Statistical functions, [M-5] normal()
invlogistic() function, [FN] Statistical functions, [M-5] normal()
invlogisticctail() function, [FN] Statistical functions, [M-5] normal()
invnbinomial() function, [FN] Statistical functions, [M-5] normal()
invnbinomialctail() function, [FN] Statistical functions, [M-5] normal()
invnchi2() function, [FN] Statistical functions, [M-5] normal()
invnchictail() function, [FN] Statistical functions, [M-5] normal()
invnF() function, [FN] Statistical functions, [M-5] normal()
invnFtail() function, [FN] Statistical functions, [M-5] normal()
invnF() function, [FN] Statistical functions, [M-5] normal()
invnchi2() function, [FN] Statistical functions, [M-5] normal()
invnchictail() function, [FN] Statistical functions, [M-5] normal()
invF() function, [FN] Statistical functions, [M-5] normal()
invFtail() function, [FN] Statistical functions, [M-5] normal()
invnibeta() function, [FN] Statistical functions, [M-5] normal()
invnibetatail() function, [FN] Statistical functions, [M-5] normal()
invnormal() function, [FN] Statistical functions, [M-5] normal()
invnt() function, [FN] Statistical functions, [M-5] normal()
invntail() function, [FN] Statistical functions, [M-5] normal()
invorder() function, [M-5] invorder()
invpoisson() function, [FN] Statistical functions, [M-5] normal()
invpoissonctail() function, [FN] Statistical functions, [M-5] normal()
_invsym() function, [M-5] invsym()
invt() function, [FN] Statistical functions, [M-5] normal()
invtokens() function, [M-5] invtokens()
invttail() function, [FN] Statistical functions, [M-5] normal()
invTukeyprob() function, [FN] Statistical functions, [M-5] normal()
invvech() function, [M-5] vec()
invvechull() function, [FN] Statistical functions, [M-5] normal()
invechullph() function, [FN] Statistical functions, [M-5] normal()
invechullphctail() function, [FN] Statistical functions, [M-5] normal()
invechulltail() function, [FN] Statistical functions, [M-5] normal()
I/O functions, [M-4] io
ipolate command, [D] ipolate
ips, xtunitroot subcommand, [XT] xtunitroot
IPW, see inverse-probability weighting
ipw, stteffects subcommand, [TE] stteffects
ipw, teffects subcommand, [TE] teffects
IPWRA, see inverse-probability-weighted regression adjustment
ipwra, stteffects subcommand, [TE] stteffects
ipwra
IRF, see impulse–response functions
irf, [TS] irf
   add command, [TS] irf add
cgraph command, [TS] irf cgraph
create command, [TS] irf create
catable command, [TS] irf ctable
describe command, [TS] irf describe
drop command, [TS] irf drop
graph command, [TS] irf graph
ogrph command, [TS] irf ograph
rename command, [TS] irf rename
set command, [TS] irf set
table command, [TS] irf table
.irf file, [U] 11.6 Filenaming conventions
iri command, [R] epitab
IRLS, see iterated, reweighted least squares
IRR, see incidence-rate ratio
IRT, see item response theory
irt
   1pl command, [IRT] irt 1pl, [IRT] irt 1pl postestimation
   2pl command, [IRT] irt 2pl, [IRT] irt 2pl postestimation
   3pl command, [IRT] irt 3pl, [IRT] irt 3pl postestimation
   command, [IRT] irt, [IRT] Control Panel
gpcm command, [IRT] irt gpcm
grm command, [IRT] irt grm, [IRT] irt grm postestimation
hybrid command, [IRT] irt hybrid, [IRT] irt hybrid postestimation
nrm command, [IRT] irt nrm, [IRT] irt nrm postestimation
pcm command, [IRT] irt pcm, [IRT] irt pcm postestimation
rsm command, [IRT] irt rsm, [IRT] irt rsm postestimation
IRT Control Panel, [IRT] Control Panel
irtgraph
   icc command, [IRT] irtgraph icc
   iif command, [IRT] irtgraph iif
tcc command, [IRT] irtgraph tcc
tif command, [IRT] irtgraph tif
.isa built-in class function, [P] class
iscale() option, [G-2] graph matrix  
iscomplex() function, [M-5] isreal()  
isdiagonal() function, [M-5] isdiagonal()  
isfleeting() function, [M-5] isfleeting()  
isid command, [D] isid  
isofclass built-in class function, [P] class  
isolines, [G-2] graph twoway contourline  
ispointer() function, [M-5] isreal()  
isreal() function, [M-5] isreal()  
isrealvalues() function, [M-5] isrealvalues()  
isstring() function, [M-5] isreal()  
isymmetric() function, [M-5] issymmetric()  
isymmetry() function, [P] matrix define  
isymmetrically() function, [M-5] issymmetric()  
isdistribute command, [R] dstdize  
isview() function, [M-5] isview()  
italics, [G-4] text  
item, [IRT] Glossary  
characteristic curve, [IRT] irtgraph icc, [IRT] Glossary  
information function, [IRT] irtgraph iif, [IRT] Glossary  
location, [IRT] Glossary  
response function, [IRT] irt, [IRT] Glossary  
iterate() option, [R] maximize, see gsem option maximize_options, see sem option maximize_options  
iterated, reweighted least squares, [R] binreg, [R] glm, [R] reg3, [R] sureg  
iteration,  
bisection method, [PSS] power, [PSS] Glossary  
Newton’s method, power, [PSS] power  
iterations, controlling the maximum number, [R] maximize  
ivpoisson command, [R] ivpoisson, [R] ivpoisson postestimation  
ivprobit command, [R] ivprobit, [R] ivprobit postestimation  
ivregress command, [R] ivregress, [R] ivregress postestimation  
ivtobit command, [R] ivtobit, [R] ivtobit postestimation  

J  
$J \times 2$ contingency table, [PSS] power trend, [PSS] Glossary  
Jaccard coefficient similarity measure, [MV] measure_option  
jackknife, [SEM] Glossary  
jackknife_options, [SVY] Jackknife_options  
jackknife prefix command, [R] jackknife, [R] jackknife postestimation  
jackknifed residuals, [R] regress postestimation  
jackknifed standard error, see Monte Carlo error  
Jarque–Bera statistic, [TS] varnorm, [TS] vecnorm  
JCA, see joint correspondence analysis  
Jeffreys noninformative prior, [MI] mi impute mvn  
jeffreys, prior() suboption, [BAYES] bayesmh evaluators  

join operator, [M-2] op_join  
joinby command, [D] joinby, [U] 22 Combining datasets  
joining datasets, see combining datasets  
joining time-span records, [ST] stsplit  
joint correspondence analysis, [MV] mca, [MV] mca postestimation, [MV] Glossary  
joint normality, see normality, joint  
jumble() function, [M-5] sort()  
jumble() function, [M-5] sort()  
justification of text, [G-3] textbox_options  
justificationstyle, [G-4] justificationstyle  

K  
forecast, [TS] dfactor postestimation, [TS] sspace postestimation, [TS] ucm postestimation  
smoothing, [TS] dfactor postestimation, [TS] sspace postestimation, [TS] ucm postestimation  
kap command, [R] kappa
Kaplan–Meier
  survivor function, [ST] ltable, [ST] stcox PH-assumption tests, [ST] sts
kappa command, [R] kappa
kapvalue command, [R] kappa
kdensity command, [R] kdensity
kdensity, graph twoway subcommand, [G-2] graph twoway kdensity
keep command, [D] drop
  keeping variables or observations, [D] drop
Kendall’s tau, [R] spearman
Kenward–Roger DDF, see denominator degrees of freedom, Kenward–Roger
kernel density estimator, [R] kdensity
kernel density smoothing, [G-2] graph other
kernel-weighted local polynomial estimator, [R] lpoly
keyboard
  entry, [U] 10 Keyboard use
  search, [U] 4 Stata’s help and search facilities
Kish design effects, [R] loneway, [SVY] estat
kiss32, see random-number generator
Kmatrix() function, [M-5] Kmatrix()
kmeans, [MV] Glossary
  kmeans, cluster subcommand, [MV] cluster kmeans and kmedians
  kmeans clustering, [MV] cluster, [MV] cluster kmeans and kmedians
kmedians, [MV] Glossary
  kmedians, cluster subcommand, [MV] cluster kmeans and kmedians
  kmedians clustering, [MV] cluster, [MV] cluster kmeans and kmedians
KMO, see Kaiser–Meyer–Olkin sampling adequacy
kmo, estat subcommand, [MV] factor postestimation, [MV] pca postestimation
KNN, see k-th-nearest neighbor
knn, discrim subcommand, [MV] discrim knn
Kolmogorov–Smirnov test, [R] ksmart
KR-20, [MV] alpha
Kruskal stress, [MV] mds postestimation, [MV] Glossary
Kruskal–Wallis test, [R] kwallis
ksmirnov command, [R] ksmart
ktau command, [R] spearman
k-th-nearest neighbor, [MV] discrim knn, [MV] Glossary
Kuder–Richardson Formula 20, [MV] alpha
Kuleczynski coefficient similarity measure, [MV] measure_option
kurt(), egen function, [D] egen
kwallis command, [R] kwallis

L
L1-norm models, [R] qreg
ltitle() option, [G-3] title_options
l2title() option, [G-3] title_options
label
  copy command, [D] label
  data command, [D] label, [U] 12.6 Dataset, variable, and value labels
define command, [D] label, [U] 12.6 Dataset, variable, and value labels
dir command, [D] label
drop command, [D] label
language command, [D] label language, [U] 12.6 Dataset, variable, and value labels
list command, [D] label, [U] 12.6 Dataset, variable, and value labels
save command, [D] label
values command, [D] label, [U] 12.6 Dataset, variable, and value labels
variable command, [D] label, [U] 12.6 Dataset, variable, and value labels
macro extended function, [P] macro
label, snapshot subcommand, [D] snapshot
label values, [P] macro, [U] 12.6 Dataset, variable, and value labels, [U] 13.11 Label values
labelbook command, [D] labelbook
labeling data in other languages, [U] 12.6.4 Labels in other languages
labels,
  axis, [G-3] axis_label_options
  creating, [D] edit, [D] varmanage
  editing, [D] edit, [D] varmanage
  marker, [G-3] marker_label_options
LAD regression, [R] qreg
ladder command, [R] ladder
ladder of powers, [G-2] graph other, [R] ladder
lag operator, [TS] Glossary, [U] 11.4.4 Time-series varlists
lag-exclusion statistics, [TS] varwle
lagged values, [U] 11.4.4 Time-series varlists, [U] 13 Functions and expressions, [U] 13.7 Explicit subscripting,
  [U] 13.10.1 Generating lags, leads, and differences
Lagrange multiplier test, [PSS] Glossary, [R] regress
postestimation time series, [SEM] estat
ginv, [SEM] estat mindices, [SEM] estat
coretests, [SEM] Glossary, [TS] varlmav,
[TS] vecmav, also see score test
Lance and Williams’s formula, [MV] cluster
language, [D] unicode locale
syntax, [P] syntax, [U] 11 Language syntax
language, label subcommand, [D] label language
languages, multiple, [D] label language
LAPACK, [M-1] LAPACK, [M-5] cholesky()
[M-5] cholinv()
[M-5] cholesolve()
[M-5] eigensystem()
[M-5] eigensystemselect()
[M-5] fullsvd()
[M-5] ghessenbergd()
[M-5] lapack()
[M-5] lud()
[M-5] luinv()
[M-5] lusolve()
[M-5] qrd()
[M-5] qrinv()
[M-5] qrsolve()
[M-5] svd()
[M-5] svsolve() Glossary
Laplacian approximation, [ME] me, [ME] meclolog,
[ME] meglm, [ME] melogit, [ME] menbreg,
[ME] meologit, [ME] meprobit,
[ME] mepoisson, [ME] meprobity, [ME] mestreg,
[ME] Glossary, [SEM] methods and formulas for gsem
latent
growth model, [SEM] intro 5, [SEM] example 18,
[SEM] Glossary
roots, [M-5] eigensystem()
space, [IRT] Glossary
trait, [IRT] Glossary
variable, [SEM] intro 4, [SEM] Glossary
latent() option, see gsem option latent(), see sem
option latent()
Latin-square designs, [MV] manova, [R] anova,
[R] pkshape
LAV regression, [R] qreg
Lawley–Hotelling trace statistic, [MV] canon,
[MV] manova, [MV] mvtetest means,
[MV] Glossary
lceffects, estat subcommand, [SVY] estat
lcolor() option, [G-3] connect_options,
[G-3] rspike_options
LDA, see linear discriminant analysis
lda, discrim subcommand, [MV] discrim lda
operator, [U] 11.4.4 Time-series varlists
values, see lagged values
leap seconds, [TS] tsset
least absolute
deviations, [R] qreg
residuals, [R] qreg
value regression, [R] qreg
least squared deviations, see linear regression
least squares, see linear regression
generalized, see feasible generalized least squares
least-squares means, [R] margins, [R] marginsplot,
[U] 20.15.1 Obtaining estimated marginal
means
leave one out, [MV] discrim, [MV] discrim
estat, [MV] discrim ktn, [MV] discrim
testat, [MV] discrim ktn postestimation, [MV] discrim
lda, [MV] discrim lda postestimation, [MV] discrim
qda, [MV] discrim qda postestimation,
[MV] Glossary
left eigenvectors, [M-5] eigensystem()
[M-6] Glossary
left suboption, [G-4] justificationstyle
left-censoring, [ST] Glossary, [TE] Glossary, see
imputation, interval-censored data
left-truncation, [TE] Glossary, see imputation, truncated
data, see truncation
lefteigensystem() function, [M-5] eigensystem()
lefteigensystem() function, [M-5] eigensystem()
lefteigensystemselect*() functions,
[M-5] eigensystemselect()
leftgeigensystem() function, [M-5] geigensystem()
leftgeigensystemselect*() function,
[M-5] geigensystem()
legend() option, [G-3] legend_options
problems, [G-3] legend_options
use with by(), [G-3] by_option,
legendstyle, [G-4] legendstyle
length() function, [M-5] rows()
length of string function, [FN] String functions
less than (or equal) operator, [U] 13.2.3 Relational
operators
letter values, [R] lv
level command and value, [P] macro
level() option, see gsem option level(), see sem
option level()
level, set subcommand, [R] level, [R] set
levels, [U] 11.4.3 Factor variables
levelsof command, [P] levelsof
Levene’s robust test statistic, [R] sdtest
leverage, [R] logistic postestimation, [R] regress
postestimation diagnostic plots
leverage plots, [G-2] graph other
leverage-versus-(squared)-residual plot, [R] regress
postestimation diagnostic plots
Levin–Lin–Chu test, [XT] stunitroot
\_LeEx, [SEM] sem and gsem option covstructure()
lexis command, [ST] stsplit
lexis diagram, [ST] stsplit
lfit, graph twoway subcommand, [G-2] graph
twoway lif
lfitci, graph twoway subcommand, [G-2] graph
twoway lifci
which
license, [R] about
likelihood, see maximum likelihood estimation
collation, [D] unicode collator
locale_functions, set subcommand, [P] set
locale_functions, [R] set
locale_ui, set subcommand, [P] set locale_ui,
[R] set
localization, [D] unicode locale
locally weighted smoothing, [R]
locale
localization, [D] unicode locale
locale
locale
locksplitters, set subcommand, [R] set
log
  close command, [R] log
  command, [R] log, [R] view, [U] 15 Saving and
  printing output—log files, [U] 16.1.2 Comments
  and blank lines in do-files
  off command, [R] log
  on command, [R] log
  query command, [R] log
  using command, [R] log
.log file, [U] 11.6 Filenaming conventions
log files, see log command
  printing, [R] translate
log() function, [FN] Mathematical functions,
  [M-5] exp() 
log hazard-ratio, [PSS] power cox, [PSS] power
  exponential, [PSS] power logrank
  one-sample, [PSS] power cox
log hazard-rate, [PSS] power exponential, [PSS] power
  logrank
log hazards
  control-group, [PSS] power exponential,
  [PSS] power logrank
  experimental-group, [PSS] power exponential,
  [PSS] power logrank
  two-sample, [PSS] power exponential, [PSS] power
  logrank
log likelihood, [BAYES] intro, [BAYES] bayes,
  [BAYES] bayesmh, [SEM] methods and
  formulas for gsem, [SEM] methods and
  formulas for sem
log or nolog option, [R] maximize
log scales, [G-3] axis_scale_options
log transformations, [R] boxcox, [R] inskew0
log10() function, [FN] Mathematical functions,
  [M-5] exp() 
logical operators, [M-2] op_logical, [U] 13.2.4 Logical
logistic
  density,
    mean \( \mu \), scale \( s \), [FN] Statistical functions
    standard, [FN] Statistical functions
  discriminant analysis, [MV] discrim
  logistic
distribution,
  cumulative, [FN] Statistical functions
  inverse cumulative, [FN] Statistical functions
logistic, continued
  regression imputation, see imputation, logistic
  regression
logistic and logit regression, [IRT] irt 1pl, [IRT] irt
  2pl, [IRT] irt 3pl, [IRT] irt hybrid, [R] logistic,
  [R] logit, [SEM] intro 5, [SEM] example 33g,
  [SEM] example 34g, [SEM] Glossary,
  [SVY] svy estimation
complementary log-log, [R] cloglog
conditional, [R] ascllogit, [R] clogit, [R] rologit
exact, [R] exlogistic
fixed-effects, [R] ascllogit, [R] clogit
fixed-effects, [XT] xtlogit, [XT] xtsreg
fractional polynomial, [R] fp
generalized estimating equations, [XT] xtgee
generalized linear model, [R] glm
mixed-effects, [ME] melogit, [ME] meqralogit, also
see ordered logistic regression
multinomial, [IRT] irt nrm, [IRT] irt hybrid,
  [R] ascllogit, [R] clogit, [R] mlogit
  nested, [R] nlogit
ordered, [IRT] irt grm, [IRT] irt pcm, [IRT] irt
  rsm, [IRT] irt hybrid, [R] ollogit
polytomous, [IRT] irt nrm, [IRT] irt hybrid,
  [R] mlogit
  population-averaged, [XT] xtgee, [XT] xtlogit,
  [XT] xtsreg
random-effects, [XT] xtlogit, [XT] xtslogit,
  [XT] xtsreg
rank-ordered, [R] rologit
skewed, [R] scobit
stereotype, [R] slocit
logistic command, [R] logistic, [R] logistic
  postestimation
logistic, discr subcommand, [MV] discr
logistic
logistic discriminant analysis, [MV] Glossary
logistic() function, [FN] Statistical functions,
  [M-5] normal()
logisticden() function, [FN] Statistical functions,
  [M-5] normal()
logisticdtail() function, [FN] Statistical functions,
  [M-5] normal()
logit command, [R] logit, [R] logit
  postestimation
logit() function, [M-5] logit()
logit function, [FN] Mathematical functions
logit option, see gsem option logit
logit regression, see logistic and logit regression
log-linear model, [R] cpoisson, [R] expoisson, [R] glm,
  [R] tpoisson, [R] poisson, [R] tpoisson, [R] zip,
  [SVY] svy estimation, [TE] etpoisson
log-log plot, [ST] stcox PH-assumption tests
loglogistic survival regression, [ST] streg
lognormal survival regression, [ST] streg
logrank, power subcommand, [PSS] power logrank
log-rank test, [PSS] power logrank
logtype, set subcommand, [R] log, [R] set
LOHI, [M-5] byteorder( )
loneway command, [R] loneway
long, [D] data types, [U] 12.2.2 Numeric storage types
long lines in ado-files and do-files, [P] #delimit,
[U] 18.11.2 Comments and long lines in ado-files
long reshape subcommand, [D] reshape
long strings, see string variables, long
longitudinal
data, [MI] mi estimate, [XT] Glossary, also see panel data
studies, see incidence studies
survey data, [SVY] svy estimation
LOO, see leave one out
look of areas, [G-3] area_options,
[G-3] fitarea_options
lookfor command, [D] lookfor
lookup,
icd10 subcommand, [D] icd10
icd9 subcommand, [D] icd9
icd9p subcommand, [D] icd9
continuing, [M-2] continue
endless, see endless loop
exiting, [M-2] break
use of semicolons in, [M-2] semicolons
looping, [P] continue, [P] foreach, [P] forvalues,
[P] while
Lorenz curve, [R] inequality
loss, [MV] Glossary
loss to follow-up, [PSS] power exponential,
[PSS] power logrank, [PSS] Glossary
Lotus 1-2-3, reading data from, see spreadsheets
lower
asymptote, [IRT] Glossary
one-sided test, [PSS] Glossary, also see one-sided test
one-tailed test, [PSS] Glossary, also see one-sided test
lower ASCII, see plain ASCII
lowercase-string function, [FN] String functions
lowertriangle() function, [M-5] lowertriangle()
lowertriangle() function, [M-5] lowertriangle()
lower-triangular matrix, see triangular matrix
lowess, see locally weighted smoothing
lowess command, [R] lowess
lowess, graph twoway subcommand, [G-2] graph twoway lowess
lowess smoothing, [G-2] graph other
lpattern() option, [G-3] connect_options,
[G-3] rspiike_options
lpoly command, [R] lpoly
lpoly, graph twoway subcommand, [G-2] graph twoway lpoly
lpoly, graph twoway subcommand, [G-2] graph twoway lpolyci
L-R plots, [G-2] graph other, [R] regress
postestimation diagnostic plots
LRECLs, [D] infile (fixed format)
lroc command, [R] lroc
lrtest command, [R] lrtest, [SEM] example 10,
[SEM] example 39g, [SEM] lrtest
ls command, [D] dir
lsens command, [R] lsens
lstat command, see estat classification command
lstretch, set subcommand, [R] set
lstyle() option, [G-3] rspiike_options
ltable command, [ST] table
ltolerance() option, [R] maximize
LU decomposition, [M-5] lud()
lud() function, [M-5] lud()
lud() function, [M-5] lud()
lud() function, [M-5] lud()
lud() function, [M-5] lud()
lud() function, [M-5] lud()
lud() function, [M-5] lud()
lud() function, [M-5] lud()
luinv() function, [M-5] luinv()
luinv() function, [M-5] luinv()
luinv() function, [M-5] luinv()
luinv() function, [M-5] luinv()
lusolve() function, [M-5] lusolve()
lusolve() function, [M-5] lusolve()
lusolve() function, [M-5] lusolve()
lusolve() function, [M-5] lusolve()
lv command, [R] lv
lvalue, class, [P] class
lvr2plot command, [R] regress postestimation diagnostic plots
lwidth() option, [G-3] connect_options,
[G-3] rspiike_options
M
M, [MI] mi impute, [MI] Glossary
size recommendations, [MI] intro substantive,
[MI] mi estimate
m, [MI] Glossary
MA, see moving average model
ma, tssmooth subcommand, [TS] tssmooth ma
Mac,
keyboard use, [U] 10 Keyboard use
pause, [P] sleep
specifying filenames, [U] 11.6 Filenaming conventions
macro
dir command, [P] macro
drop command, [P] macro
list command, [P] macro
shift command, [P] macro
macro substitution, [P] macro
class, [P] class
[U] 18.3 Macros, also see e() stored results
macval() macro expansion function, [P] macro
mad(), egen function, [D] egen
MAD regression, [R] qreg
Mahalanobis

distance, [MV] Glossary

transformation, [MV] Glossary

main effects, [MV] manova, [PSS] Glossary, [R] anova

makecns command, [P] makecns

_makesymmetric() function, [M-5] makesymmetric()
makesymmetric() function, [M-5] makesymmetric()

man command, [R] help

manage, window subcommand, [P] window

programming, [P] window manage

MANCOVA, see multivariate analysis of covariance

mangle option, [G-2] graph twoway pcarrow

manifest variables, [SEM] Glossary

manipulation commands, [G-2] graph manipulation

Mann–Whitney two-sample statistics, [R] ranksum

MANOVA, see multivariate analysis of variance

manova command, [MV] manova, [MV] manova

postestimation

manova, estat subcommand, [MV] discrim lda

postestimation

manovatest command, [MV] manova postestimation

Mantel–Cox method, [ST] strate

Mantel–Haenszel

method, [ST] strate


mapping strings to numbers, [D] destring, [D] encode,

[D] label, also see real() function

maps, [M-5] asarray()

MAR, see missing at random, see missing values

marginal

distribution, Bayesian, [BAYES] intro,

[BAYES] bayesmh, [BAYES] Glossary

effects, [R] margins, [R] marginsplot,

[U] 20.15 Obtaining marginal means, adjusted predictions, and predictive margins,

[U] 20.19 Graphing margins, marginal effects, and contrasts

homogeneity, [PSS] power, [PSS] power

pairedproportions, [PSS] power mcc,

[PSS] Glossary

homogeneity, test of, [R] symmetry

likelihood, Bayesian, [BAYES] intro,

[BAYES] bayesmh, [BAYES] bayesstats ic,


means, [R] contrast, [R] margins, [R] margins,

contrast, [R] margins, pwcompare,

[R] marginsplot, [R] pwcompare,

[U] 20.15 Obtaining marginal means, adjusted predictions, and predictive margins

posterior distribution, Bayesian, [BAYES] intro,

[BAYES] bayes, [BAYES] bayesmh,

[BAYES] bayesgraph, [BAYES] bayesstats

ess, [BAYES] bayesstats summary,

[BAYES] bayestest interval, [BAYES] Glossary

proportion, see proportions, marginal

parameter trace files, [MI] mi ptrace

sample, [BAYES] Glossary

standard error, see Monte Carlo standard error

Markov-switching model, [TS] mswitch

marksample command, [P] mark

Marquardt algorithm, [M-5] moptimize(),

[M-5] optimize()

martingale residual, [ST] stcox postestimation,

[ST] streg postestimation
mass, [MV] Glossary
Mata, [D] putmata
  commands, [M-3] intro
  error messages, [M-5] error(), also see traceback
  log
mata
  clear command, [M-3] mata clear
  describe command, [M-3] mata describe
  drop command, [M-3] mata drop
  help command, [M-3] mata help
  invocation command, [M-3] mata
  matdescribe command, [M-3] mata matsave
  matsave command, [M-3] mata matsave
  maturse command, [M-3] mata matsave
  memory command, [M-3] mata memory
  mlib add command, [M-3] mata mlib
  mlib create command, [M-3] mata mlib
  mlib index command, [M-3] mata mlib
  mlib query command, [M-3] mata mlib
  mosave command, [M-3] mata mosave
  query command, [M-3] mata set, [R] set
  rename command, [M-3] mata rename
  set matacache command, [M-3] mata set, [R] set
  set matalibs command, [M-3] mata set, [R] set
  set matalnum command, [M-3] mata set, [R] set
  set matamofirst command, [M-3] mata set, [R] set
  set mataoptimize command, [M-3] mata set, [R] set
  set matastrict command, [M-1] ado,
  stata command, [M-3] mata stata
  which command, [M-3] mata which
mata, clear subcommand, [D] clear
mata, query subcommand, [R] query
.mata source code file, [M-1] source, [M-3] mata
  conventions
matched
  2×2 tables, [PSS] power mcc
  case–control data, [R] asclogit, [R] clogit,
    [R] epitab, [R] symmetry, [ST] sttrace
  case–control study, [PSS] power, [PSS] power mcc
  study, [PSS] power, [PSS] power mcc,
    [PSS] Glossary
 matched‐pairs tests, [R] signrank, [R] ttest, [R] ztest
matching
  1: M, [PSS] power
  coefficient, [MV] Glossary
  coefficient similarity measure,
    [MV] measure_option
  configuration, [MV] Glossary
  estimator, [TE] teffects intro, [TE] teffects intro
    advanced, [TE] teffects nmatch, [TE] teffects
    psmatch, [TE] Glossary
matcproc command, [P] makecns
matdescribe, mata subcommand, [M-3] mata
  matsave
  _matexpssym() function, [M-5] mata expssym()
  matexpssym() function, [M-5] mata expssym()
  math symbols, [G-4] text
mathematical functions, [M-4] mathematical,
  [M-4] standard
mathematical functions and expressions,
  [FN] Mathematical functions, [P] matrix define,
    [U] 13.3 Functions
Matlab, reading data from, [U] 21.4 Transfer programs
matlist command, [P] matlist
  _matlogssym() function, [M-5] matexpssym()
  matlogssym() function, [M-5] matexpssym()
  matmissing() function, [FN] Matrix functions,
    [P] matrix define
matname command, [P] matrix mkmat
  _matpowersym() function, [M-5] matpowersym()
  matpowersym() function, [M-5] matpowersym()
  mat_put_rr command, [P] matrix get
  [U] 14 Matrix expressions
  accessing internal, [P] matrix get
  accumulating, [P] matrix accum
  appending rows and columns, [P] matrix define
  Cholesky decomposition, [P] matrix define
  coefficient matrices, [P] ereturn
  column names, see matrices, row and column names
  constrained estimation, [P] makecns
  copying, [P] matrix define, [P] matrix get,
    [P] matrix mkmat
  correlation, [MV] pca, [P] matrix define
  covariance, [MV] pca
  covariance matrix of estimators, [P] ereturn,
    [P] matrix get
  cross-product, [P] matrix accum
  determinant, [P] matrix define
  diagonals, [P] matrix define
  displaying, [P] matlist, [P] matrix utility
  dissimilarity, [MV] matrix dissimilarity,
    [MV] Glossary, [P] matrix dissimilarity
  distances, [MV] matrix dissimilarity, [P] matrix
    dissimilarity
  dropping, [P] matrix utility
  eigenvalues, [P] matrix eigenvalues, [P] matrix
    symeigen
  eigenvectors, [P] matrix symeigen
  elements, [P] matrix define
  equation names, see matrices, row and column
  names
  estimation results, [P] ereturn, [P] _estimates
  functions, [FN] Matrix functions,
    [M-4] manipulation, [M-4] matrix,
  identity, [P] matrix define
  input, [P] matrix define, [U] 14.4 Inputting
  matrices by hand
matrices, continued
   inversion, [P] matrix define, [P] matrix svd
Kronecker product, [P] matrix define
   labeling rows and columns, see matrices, row and column names
   linear combinations with data, [P] matrix score
listing, [P] matlist, [P] matrix utility
   namespace and conflicts, [P] matrix, [P] matrix define
norm, [M-5] norm()
   number of rows and columns, [P] matrix define
   operators such as addition, [P] matrix define,
   [U] 14.7 Matrix operators
   orthonormal basis, [P] matrix svd
   partitioned, [P] matrix define
   performing constrained estimation, [P] makecns
posting estimation results, [P] ereturn,
   [P] _estimates
   renaming, [P] matrix utility
   row and column names, [P] ereturn, [P] matrix define,
   [P] matrix mkmat, [P] matrix rownames,
   [U] 14.2 Row and column names
   rows and columns, [P] matrix define
   saving matrix, [P] matrix mkmat
   scoring, [P] matrix score
   similarity, [MV] matrix dissimilarity, [P] matrix dissimilarity
   store variables as matrix, [P] matrix mkmat
   submatrix extraction, [P] matrix define
   submatrix substitution, [P] matrix define
   subscripting, [P] matrix define,
   [U] 14.9 Subscripting
   sweep operator, [P] matrix define
   temporary names, [P] matrix
   trace, [P] matrix define
   transposing, [P] matrix define
   variables, make into matrix, [P] matrix mkmat
   zero, [P] matrix define
matrix, [M-2] declarations
   accum command, [P] matrix accum
coleq command, [P] matrix rownames
colnames command, [P] matrix rownames
   commands, introduction, [P] matrix
define command, [P] matrix define
dir command, [P] matrix utility
dissimilarity command, [MV] matrix dissimilarity,
   [P] matrix dissimilarity
drop command, [P] matrix utility
eigenvalues command, [P] matrix eigenvalues
glsaccum command, [P] matrix accum
   input command, [P] matrix define
list command, [P] matrix utility
opaccum command, [P] matrix accum
rename command, [P] matrix utility
roweq command, [P] matrix rownames
rownames command, [P] matrix rownames
score command, [P] matrix score
matrix, continued
   svd command, [P] matrix svd
   symeigen command, [P] matrix symeigen
   vecaccum command, [P] matrix accum
matrix.
   clear subcommand, [D] clear
confirm subcommand, [P] confirm
   ereturn subcommand, [P] ereturn, [P] return
graph subcommand, [G-2] graph matrix
   return subcommand, [P] return
matrix graphs, [G-2] graph matrix
   matrix model parameter, [BAYES] Glossary, see
   Bayesian, model parameters
matrix() function, [FN] Programming functions,
   [P] matrix define
matsave, mata subcommand, [M-3] mata matsave
   matsize command, [U] 14 Matrix expressions
   matsize, set subcommand, [R] matsize, [R] set
   matuniform() function, [FN] Matrix functions,
   [P] matrix define
matuse, mata subcommand, [M-3] mata matsave
   max(),
      built-in function, [FN] Mathematical functions
egen function, [D] egen
max() function, [M-5] minmax()
   max_memory, set subcommand, [R] set
   maxbyte() function, [FN] Programming functions
   maxdb, set subcommand, [R] db, [R] set
   maxdouble() function, [FN] Programming functions,
   [M-5] mindouble()
   maxes() option, [G-2] graph matrix
   maxfloat() function, [FN] Programming functions
   maximization technique explained, [R] maximize,
   [R] ml
   maximize, ml subcommand, [R] ml
 maximize_options, see gsem option maximize_options,
   see sem option maximize_options
maximum
   function, [D] egen, [FN] Programming functions
   length of string, [M-1] limits
   likelihood, [SEM] intro 4, [SEM] methods and
   formulas for gsem, [SEM] methods and
   formulas for sem, [SEM] Glossary
   with missing values, [SEM] example 26,
   [SEM] Glossary
   likelihood estimation, [MV] factor, [R] maximize,
   [R] ml, [R] mlexp
   likelihood factor method, [MV] Glossary
   limits, [R] limits
   number of observations, [D] memory
   number of variables, [D] describe, [D] memory
   number of variables and observations,
   [U] 6 Managing memory
   number of variables in a model, [R] matsize
   pseudolikelihood estimation, [SVY] ml for svy,
   [SVY] variance estimation
maximum, continued
  restricted likelihood, [ME] mixed
  size of dataset, [U] 6 Managing memory
  size of matrix, [M-1] limits
  value dissimilarity measure, [MV] measure_option
maxima and minima, [M-5] minindex()
maxcreating dataset of, [D] collapse
functions, [D] egen, [FN] Mathematical functions,
[FN] Programming functions
maxreporting, [R] iv, [R] summarize, [R] table
maxindex() function, [M-5] minindex()
maxint() function, [FN] Programming functions
maxiter, set subcommand, [R] maximize, [R] set
maxlong() function, [FN] Programming functions
max_memory, set subcommand, [D] memory
maxvar, set subcommand, [D] memory, [R] set
mband, graph twoway subcommand, [G-2] graph
twoway mband
MCA, see multiple correspondence analysis
mca command, [MV] mca, [MV] mca postestimation,
[MV] mca postestimation plots
MCAGH, see quadrature, mode-curvature adaptive
  Gauss–Hermite
MCAGHQ, see mode-curvature adaptive Gauss–Hermite
  quadrature
mcaplot command, [MV] mca postestimation,
[MV] mca postestimation plots
mcaposition command, [MV] mca postestimation,
[MV] mca postestimation plots
MCAR, see missing completely at random
mcc command, [R] epitab
mcc, power subcommand, [PSS] power mcc
mcci command, [R] epitab
MCE, see Monte Carlo error
McFadden’s choice model, [R] asclogit
MCMC, see Markov chain Monte Carlo
McNemar’s chi-squared test, [R] clogit, [R] epitab
McNemar’s test, [PSS] Glossary
mcolor() option, [G-3] marker_options
MCSE, see Monte Carlo standard error
md command, [D] mkdir
MDES, see minimum detectable effect size
mdev(), egen function, [D] egen
MDS, see multidimensional scaling
mds command, [MV] mds, [MV] mds postestimation,
[MV] mds postestimation plots
mdsconfig command, [MV] mds, [MV] mds
  postestimation plots
mdslog command, [MV] mds postestimation,
[MV] mds postestimation plots, [MV] mdslog
mdsplot command, [MV] mds postestimation,
[MV] mds postestimation plots, [MV] mdsplot
mds Shepard command, [MV] mds postestimation
  plots
mdy() function, [D] datetime, [FN] Date and time
  functions, [M-5] date()
mdyhm() function, [D] datetime, [FN] Date and time
  functions, [M-5] date()
measurement, continued
model, [SEM] intro 5, [SEM] example 1,
[SEM] example 3, [SEM] example 20,
[SEM] example 27g, [SEM] example 30g,
[SEM] example 31g, [SEM] Glossary
variables, [SEM] Glossary
measures, cluster subcommand, [MV] cluster
programming utilities
measures of
association, [R] tabulate twoway
central tendency, see means, see medians
dispersion, see percentiles, displaying, see standard
deviations, displaying, see variance, displaying,
see range of data
inequality, [R] inequality
location, [R] lv, [R] summarize
median(), egen function, [D] egen
median regression, [R] qreg
median test, [R] ranksum
medianlinkage,
clustermat subcommand, [MV] cluster linkage
cluster subcommand, [MV] cluster linkage
median-linkage clustering, [MV] cluster,
[MV] clustermat, [MV] cluster linkage,
[MV] Glossary
medians,
creating
dataset of, [D] collapse
variable containing, [D] egen
displaying, [D] pctl, [R] centile, [R] lv,
[R] summarize, [R] table, [R] tabstat
graphing, [R] grmeanby
testing equality of, see equality test of medians
mediation model, [SEM] intro 5, [SEM] example 42g
MEFF, see misspecification effects
MEFT, see misspecification effects
meglm command, [ME] meglm
melogit command, [ME] melogit
member
function, [M-2] class
program, [P] class
variable, [M-2] class, [P] class
memory
graphs, describing contents, [G-2] graph describe
matsize, see matsize, set subcommand
requirements, estimating for flongsep, [MI] mi convert
settings, [P] creturn
utilization, [M-1] limits, [M-3] mata memory
memory, continued
managing, [U] 6 Managing memory
reducing utilization, [D] compress, [D] encode,
[D] recast, [P] discard
memory command, [D] memory, [U] 6 Managing memory
menbreg command, [ME] menbreg
menu, window subcommand, [P] window
programming, [P] window menu
menus, programming, [P] dialog programming,
[P] window programming, [P] window fopen,
[P] window manage, [P] window menu,
[P] window push, [P] window stopbox
meologit command, [ME] meologit
meoprobit command, [ME] meoprobit
mepoisson command, [ME] mepoisson
meprobit command, [ME] meprobit
mqrlogit command, [ME] mqrlogit
mqrpoisson command, [ME] mqrpoisson
merge command, [D] merge, [U] 22 Combining datasets
merge, mi subcommand, [MI] mi merge
_merge variable, [D] merge
merging data, [MI] mi merge, see combining datasets
Mersenne Twister, see random-number generator
messages and return codes, see error messages and
return codes
mestreg command, [ME] mestreg
meta-analysis, [R] meta
method() option, see gsem option method(), see sem
option method()
metric scaling, [MV] Glossary
Metropolis–Hastings
algorithm, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesmh evaluators,
[BAYES] Glossary
sampler, see Metropolis–Hastings sampling
sampling, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesmh evaluators,
[BAYES] Glossary
mfcolor() option, [G-3] marker_options
mfp prefix command, [R] mfp, [R] mfp postestimation
mfx, estat subcommand, [R] asclogit postestimation,
[R] asprobit postestimation, [R] asprobit
postestimation
MGARCH, see multivariate GARCH model
mgarch
ccc command, [TS] mgarch ccc, [TS] mgarch ccc
postestimation
dcc command, [TS] mgarch dcc, [TS] mgarch dcc
postestimation
dvech command, [TS] mgarch dvech, [TS] mgarch
dvech postestimation
mgarch, continued

  vcc command, [TS] mgarch vcc, [TS] mgarch vcc postestimation

MH

  algorithm, see Metropolis–Hastings algorithm
  sampler, see Metropolis–Hastings sampling
  sampling, see Metropolis–Hastings sampling

mhodds command, [R] epitable

mi

  add command, [MI] mi add
  append command, [MI] mi append
  command, [MI] intro, [MI] styles, [MI] workflow
  convert command, [MI] mi convert
  copy command, [MI] mi copy, [MI] styles
describe command, [MI] mi describe
  erase command, [MI] mi erase, [MI] styles
  estimate command, [MI] mi estimate,
    [MI] mi estimate using, [MI] mi estimate postestimation, [MI] mi test
  estimate postestimation, [MI] mi estimate postestimation,
    [MI] mi predict, [MI] mi test
  expand command, [MI] mi expand
  export command, [MI] mi export, [MI] mi export
  ice, [MI] mi export nhanes1
  extract command, [MI] mi extract, [MI] mi replace0
  fvset command, [MI] mi XXXset
  import command, [MI] mi import, [MI] mi import
    flong, [MI] mi import flongsep, [MI] mi import
    ice, [MI] mi import nhanes1, [MI] mi import
    wide
  impute command, [MI] mi impute, [MI] mi impute
    chained, [MI] mi impute intreg, [MI] mi impute
    logit, [MI] mi impute mlogit, [MI] mi impute
    monotone, [MI] mi impute mvn, [MI] mi impute
    nbreg, [MI] mi impute ologit, [MI] mi impute
    pmm, [MI] mi impute poisson, [MI] mi impute
    regress, [MI] mi impute truncreg, [MI] mi impute
    usermethod
  merge command, [MI] mi merge
  misstable command, [MI] mi misstable
  passive command, [MI] mi passive
  predict command, [MI] mi estimate
    postestimation, [MI] mi predict
  predictnl command, [MI] mi estimate
    postestimation, [MI] mi predict
  ptrace command, [MI] mi ptrace
  query command, [MI] mi describe
  register command, [MI] mi set
  rename command, [MI] mi rename
  replace0 command, [MI] mi replace0
  reset command, [MI] mi reset
  reshape command, [MI] mi reshape
  select command, [MI] mi select, also see mi
  extract command
  set command, [MI] mi set
  st command, [MI] mi XXXset
  stjoin command, [MI] mi stsplit

mi, continued

  stset command, [MI] mi XXXset
  stsplit command, [MI] mi XXXset
  svyset command, [MI] mi XXXset
test command, [MI] mi estimate postestimation,
  [MI] mi test
testtransform command, [MI] mi estimate
  postestimation, [MI] mi test
tsset command, [MI] mi XXXset
  unregister command, [MI] mi set
  unset command, [MI] mi set
  update command, [MI] mi update, [MI] nouptdate
    option
  varying command, [MI] mi varying
  xeq command, [MI] mi xeq
  xtset command, [MI] mi XXXset

mi data, [MI] Glossary

mi() function, [FN] Programming functions

MICE, see multivariate imputation, chained equations

Microsoft

  Access, reading data from, [D] odbc,
    [U] 21.4 Transfer programs
  Excel, [M-5] xl()
  Excel, reading data from, [D] import excel,
    [D] odbc
  Excel, write results to, [P] putexcel, [P] putexcel
    advanced
  SpreadsheetML, [D] xmlsave
  Windows, see Windows
  Word, [M-5] _docx*( )

middle suboption, [G-4] alignmentstyle

midsummaries, [R] lv

mild outliers, [R] lv

Mills’s ratio, [R] heckman, [R] heckman
postestimation

MIMIC models, see multiple indicators and multiple
causes model

min(), egen function, [D] egen

min() function, [FN] Mathematical functions,
  [M-5] minmax()
min_memory, set subcommand, [R] set

minbyte() function, [FN] Programming functions

mindices, estat subcommand, [SEM] estat
mindices
mindouble() function, [FN] Programming functions,
  [M-5] mindouble()

minfloat() function, [FN] Programming functions


minimum

  absolute deviations, [R] qreg
detectable effect size, [PSS] power, [PSS] Glossary
detectable value, [PSS] Glossary

test rotation, [MV] rotate, [MV] rotatemat,
  [MV] Glossary

squared deviations, [R] areg, [R] cnstreg, [R] nl,
  [R] regress, [R] regress postestimation
minindex() function, [M-5] minindex()

minint() function, [FN] Programming functions

Minkowski dissimilarity measure,

[MV] measure_option

minlong() function, [FN] Programming functions

minmax() function, [M-5] minmax()

min_memory, set subcommand, [D] memory

minutes() function, [D] datetime, [FN] Date and time functions, [D] date()

misclassification rate, [MV]

missing at random, [MI] intro substantive

mi impute, [MI] Glossary

missing completely at random, [MI] intro substantive,

[MI] Glossary

missing data, [MI] intro substantive


monotone pattern, [MI] intro substantive, [MI] mi impute, [MI] mi impute chained, [MI] mi impute monotone, [MI] mi impute mvn,

[MI] Glossary

missing() function, [FN] Programming functions,

[M-5] missing()

missing not at random, [MI] intro substantive,

[MI] Glossary

missing observations, see dropout

missing values, [M-5] missing(), [M-5] missingof(),


[D] missing values, [R] misstable,

[SEM] example 26, [U] 12.2.1 Missing values,

[U] 13 Functions and expressions

counting, [D] codebook, [D] inspect

encoding and decoding, [D] mvencode

extended, [D] mvencode

hard and soft, [MI] Glossary

ineligible, [MI] Glossary

pattern of, [MI] mi misstable

replacing, [D] merge

missingness, pattern, see pattern of missingness

missingof() function, [M-5] missingof()

mis specification effects, [SVY] estat, [SVY] Glossary

misstable

for mi data, [MI] mi misstable

nested command, [R] misstable

patterns command, [R] misstable

summarize command, [R] misstable

tree command, [R] misstable

misstable, mi subcommand, [MI] mi misstable

mixed model, continued


mixed command, [ME] mixed

mixed-effects model, see multilevel model

mkdir command, [D] mkdir

 mkdir() function, [M-5] chdir()

mkdir() function, [M-5] chdir()

mkmat command, [P] matrix mkmat

mkspline command, [R] mkspline

ML, see maximum likelihood

ml, see gsem option method(), see sem option method()

check command, [R] ml

clear command, [R] ml

command, [SVY] ml for svy

count command, [R] ml

display command, [R] ml

footnote command, [R] ml

graph command, [R] ml

init command, [R] ml

maximize command, [R] ml

model command, [R] ml

plot command, [R] ml

query command, [R] ml

report command, [R] ml

score command, [R] ml

search command, [R] ml

trace command, [R] ml

mlabangle() option, [G-3] marker_label_options

mlabcolor() option, [G-3] marker_label_options

mlabel() option, [G-3] marker_label_options

mlabgap() option, [G-3] marker_label_options

mlabposition() option, [G-3] marker_label_options

mlabsize() option, [G-3] marker_label_options

mlabstyle() option, [G-3] marker_label_options

mlabtextstyle() option,

[G-3] marker_label_options

mlabvposition() option,

[G-3] marker_label_options

mlcolor() option, [G-3] marker_options

mleval command, [R] ml

mlexp command, [R] mlexp, [R] mlexp postestimation

mlib

add, mata subcommand, [M-3] mata mlib

create, mata subcommand, [M-3] mata mlib

index, mata subcommand, [M-3] mata mlib

query, mata subcommand, [M-3] mata mlib

.mlib library file, [M-1] how, [M-3] mata describe,


mlmatbysum command, [R] ml

mlmatsum command, [R] ml

MLMV, see maximum likelihood with missing values

mlmv, see sem option method()

mlogit command, [R] mlogit, [R] mlogit postestimation
mlmtest option, see gsem option mlogit

mlogit

data style, [MI] styles, [MI] Glossary

technical description, [MI] technical

mlpattern() option, [G-3] marker_options

mlstyle() option, [G-3] marker_options

mlsum command, [R] ml

mvce command, [R] ml

mlwidth() option, [G-3] marker_options


mode(), egen function, [D] egen

mode-curve adaptive Gauss–Hermite quadrature, [IRT] Glossary, see quadrature, mode-curve adaptive Gauss–Hermite

model

coefficients test, [R] lrtest, [R] test, [R] testnl, [SVY] svy postestimation

comparison, Bayesian, see Bayesian, model comparison

hypothesis testing, see Bayesian, hypothesis testing identification, [SEM] intro 4, [SEM] intro 12, [SEM] Glossary


simplification test, [SEM] example 8, [SEM] example 10

specification test, see specification test

model, ml subcommand, [R] ml

model-implied covariances and correlations, [SEM] example 11

modeling

fractions, [R] betareg

ratios, [R] betareg

models, maximum number of variables in, [R] matsize

modern scaling, [MV] Glossary

modification, file, [D] filefilter


modifying data, [D] generate, also see editing data modulus function, [FN] Mathematical functions

module transformations, [R] boxcox


monotone imputation, see imputation, monotone

monotone-missing pattern, [MI] mi impute monotone, [MI] Glossary, [R] mtable

monotonicity, see pattern of missingness

Monte Carlo

evaluate(), [D] evaluate, [D] evaluate

evaluate()


month() function, [D] datatime, [FN] Date and time functions, [M-5] date()

MNAR, see missing not at random

MNP, see outcomes, multinomial


moptimize() function, [M-5] moptimize()

moptimize() function, [M-5] moptimize()

moptimize_ado_cleanup() function, [M-5] moptimize()

moptimize_evaluate() function, [M-5] moptimize()

moptimize_evaluate() function, [M-5] moptimize()

moptimize_init() function, [M-5] moptimize()

moptimize_init_*() functions, [M-5] moptimize()

moptimize_query() function, [M-5] moptimize()

moptimize_result_*() functions, [M-5] moptimize()

moptimize_util_*() functions, [M-5] moptimize()

more command and parameter, [P] macro, [P] more, [R] more, [U] 7 --more-- conditions, [U] 16.1.6 Preventing --more-- conditions

more() function, [M-5] more()

more, set subcommand, [R] more, [R] set

mortality table, see life tables

mosave, mata subcommand, [M-3] mata mosave

moving average


process, [TS] Glossary

smoother, [TS] tssmooth, [TS] tssmooth ma

mprobit command, [R] mprobit, [R] mprobit

postestimation


mreldifre() function, [M-5] reldif()

mreldifsym() function, [M-5] reldif()

mreldifre() function, [M-5] reldif()

msize() option, [G-3] marker_options, [G-3] rcap_options

mosofhours() function, [D] datatime, [FN] Date and time functions, [M-5] date()

mosofminutes() function, [D] datatime, [FN] Date and time functions, [M-5] date()
multinomial
logistic regression imputation, see imputation, multinomial logistic regression
outcome model, see outcomes, multinomial
probit regression, [SVY] svy estimation
Duncan’s method, [R] pwcompare, [R] pwmean
Holm’s method, [R] anova postestimation, [R] correlate, [R] test, [R] testnl
multiple-range multiple-comparison adjustment, see multiple comparisons, continued
Studentized-range method, see Tukey’s method
Tukey’s method, [R] pwcompare, [R] pwmean
multiple correlation, [SEM] Glossary
multiple correspondence analysis, [MV] Glossary
imputation step, [MI] intro substantive, [MI] mi impute, [MI] mi impute usermethod
inference, [MI] intro substantive
pooling step, [MI] intro substantive, [MI] mi estimate, [MI] mi estimate using
prediction, [MI] mi predict
theory, [MI] intro substantive
multiple indicators and multiple causes model, [SEM] intro 5, [SEM] example 10, [SEM] Glossary
multiple indicators multiple causes model, [SEM] example 36g
multiple languages, [D] label language
multiple regression, see linear regression
multiple-range multiple-comparison adjustment, see multiple comparisons, Dunnett’s method
multiple-sample
means, see means, multiple-sample study, [PSS] power oneway, [PSS] power twoway, [PSS] power repeated
multiple-sample, continued
- test, [PSS] power oneway, [PSS] power twoway, [PSS] power repeated
- independent samples, [PSS] power oneway, [PSS] power twoway, [PSS] power repeated
- means, [PSS] power oneway, [PSS] power twoway, [PSS] power repeated

multiplication operator, see arithmetic operators

multiplicative heteroskedasticity, [TS] arch

multistage clustered sampling, [SVY] svy

multiplicative heteroskedasticity, [TS] ts

see multiplication operator, [G-3]

teffects aipw


multivariable fractional polynomial regression, [R] mfp

- bivariate probit, [R] biprobit
- three-stage least squares, [R] reg3
- Zellner’s seemingly unrelated, [R] nlsur, [R] sureg

analysis of covariance, [MV] manova, [MV] Glossary

analysis of variance, [MV] manova, [MV] Glossary

Behrens–Fisher problem, [MV] mvtest means

- constant conditional correlation, [TS] mgarch ccc
diagonal vech, [TS] mgarch dvech
dynamic conditional correlation, [TS] mgarch dce
- varying conditional correlation, [TS] mgarch vce

GARCH postestimation
- constant conditional correlation, [TS] mgarch ccc postestimation
diagonal vech, [TS] mgarch dvech postestimation
dynamic conditional correlation, [TS] mgarch dce postestimation
- varying conditional correlation, [TS] mgarch vce postestimation

imputation, see imputation, multivariate

imputation using chained equations, see imputation, multivariate, chained equations

kurtosis, [MV] mvtest normality

logistic variable imputation, see imputation, multivariate

normal, [MV] mvtest normality

normal imputation, see imputation, multivariate normal


[SEM] example 12, [SEM] Glossary, also see seemingly unrelated regression

multivariate, continued
- regression imputation, see imputation, multivariate
- skewness, [MV] mvtest normality

test, [MV] mvtest

time-series estimators,
- structural vector autoregressive models, [TS] var svar
time-series estimators,
- dynamic-factor models, [TS] dfactor
MGARCH models, see multivariate GARCH model
- state-space models, [TS] sspace
- vector autoregressive models, [TS] var,
- [TS] varbasic
- vector error-correction models, [TS] vec

MVAGH, see quadrature, mean–variance adaptive Gauss–Hermite

MVAGHQ, see mean–variance adaptive Gauss–Hermite quadrature

mvdecode command, [D] mvencode

mvencode command, [D] mvencode

MVN imputation, see imputation, multivariate normal

mvreg command, [MV] mvreg, [MV] mvreg postestimation

mvreg, estat subcommand, [MV] procrustes postestimation

mvtest, [MV] mvtest correlations command, [MV] mvtest correlations

covariances command, [MV] mvtest covariances

means command, [MV] mvtest means

normality command, [MV] mvtest normality

N

_`n` and _`N` built-in variables, [U] 13.4 System variables (_variables_), [U] 13.7 Explicit subscripting

name() option, [G-3] name_option

nameexternal() function, [M-5] findexternal()

namelists, [M-3] namelists

names, [U] 11.3 Naming conventions
- matrix row and columns, [P] ereturn, [P] matrix define,
- [P] matrix rownames

names, confirm subcommand, [P] confirm

namespace and conflicts, matrices and scalars,
- [P] matrix, [P] matrix define

naming
- convention, [M-1] naming
- groups of variables, [D] rename group

variables, [D] rename

NaN, [M-6] Glossary

NARCH, see nonlinear autoregressive conditional heteroskedasticity

NARCHK, see nonlinear autoregressive conditional heteroskedasticity with a shift

natural log function, [FN] Mathematical functions,
[FN] Statistical functions
natural splines, [R] mkspline

\texttt{nbetaden()} function, [FN] Statistical functions, [M-5] \texttt{normal()}
\texttt{nbetad() function, [FN] Statistical functions, [M-5] \texttt{normal()}}
\texttt{nbetadlp()} function, [FN] Statistical functions, [M-5] \texttt{normal()}

\texttt{nbetadtail()} function, [FN] Statistical functions, [M-5] \texttt{normal()}
\texttt{nbetad command, [R] nbetad, [R] nbetad postestimation}
\texttt{nbetad option, see \texttt{gsem} option \texttt{nbetad}}
\texttt{nchi2()} function, [FN] Statistical functions, [M-5] \texttt{normal()}
\texttt{nchi2den()} function, [FN] Statistical functions, [M-5] \texttt{normal()}
\texttt{nchi2() function, [FN] Statistical functions, [M-5] \texttt{normal()}}
\texttt{nbreg option, see \texttt{gsem} option \texttt{nbreg}}
\texttt{nchi2() function, [FN] Statistical functions, [M-5] \texttt{normal()}}
\texttt{nchi2den() function, [FN] Statistical functions, [M-5] \texttt{normal()}}

\texttt{n-class command, [P] \texttt{program}, [P] \texttt{return}}
\texttt{ndots() option, [G-2] graph twoway dot}
\texttt{nearest neighbor, [MI] \texttt{mi} impute pmm, [MV] discrim knn, [MV] Glossary}
\texttt{needle plot, [R] spikeplot}
\texttt{\_negate()} function, [M-5] \texttt{\_negate()}
\texttt{negation matrix, [M-5] \texttt{\_negate()}}
\texttt{negation operator, see arithmetic operators}
\texttt{negative binomial, [SEM] example 39g}
\texttt{distribution, cumulative, [FN] Statistical functions}
\texttt{inverse cumulative, [FN] Statistical functions}
\texttt{inverse reverse cumulative, [FN] Statistical functions}
\texttt{reverse cumulative, [FN] Statistical functions}
\texttt{probability mass function, [FN] Statistical functions}
\texttt{fixed-effects, [XT] xtnbreg}
\texttt{generalized linear models, [R] glm}
\texttt{mixed-effects, [ME] menbreg}
\texttt{population-averaged, [XT] xtggee, [XT] xtnbreg}
\texttt{random-effects, [XT] xtnbreg}
\texttt{truncated, [R] tnreg}
\texttt{zero-inflated, [R] znb}

\texttt{negative effect size, [PSS] power, [PSS] Glossary}
\texttt{Nelder–Mead algorithm, [M-5] \texttt{moptimize()}, [M-5] \texttt{optimize()}}
\texttt{Nelson–Aalen cumulative hazard, [ST] \texttt{sts}, [ST] \texttt{sts graph}, [ST] \texttt{sts list}}
\texttt{nested case–control data, [ST] \texttt{sttoce}}
\texttt{designs, [MV] manova, [R] anova}
\texttt{effects, [MV] manova, [R] anova}
\texttt{logit, [R] nlogit}
\texttt{model statistics, [R] nestreg}
\texttt{number list, [PSS] power}

\texttt{nested, continued}
\texttt{regression, [R] nestreg}
\texttt{nested, misstable subcommand, [R] misstable}
\texttt{nested-effects model, [SEM] Glossary}
\texttt{nestreg prefix command, [SEM] Glossary}
\texttt{net}
\texttt{cd command, [R] net}
\texttt{describe command, [R] net}
\texttt{from command, [R] net}
\texttt{get command, [R] net}
\texttt{install command, [R] net}
\texttt{link command, [R] net}
\texttt{query command, [R] net}
\texttt{search command, [R] net search}
\texttt{set ado command, [R] net set other command, [R] net}
\texttt{sj command, [R] net}
\texttt{stb command, [R] net}
\texttt{net_d, view subcommand, [R] view}
\texttt{net, view subcommand, [R] view}
\texttt{NetCourseNow, [U] 3.6.2 NetCourses}
\texttt{NetCourses, [U] 3.6.2 NetCourses}
\texttt{network, query subcommand, [R] query}
\texttt{.new() function, [M-2] class}
\texttt{new lines, data without, [D] \texttt{infix (fixed format)}}
\texttt{\texttt{new}, \texttt{ssc} subcommand, [R] ssc}
\texttt{newey command, [TS] newey, [TS] newey postestimation}
\texttt{Newey–West}
\texttt{covariance matrix, [TS] Glossary}
\texttt{postestimation, [TS] newey postestimation}
\texttt{regression, [TS] newey}
\texttt{standard errors, [P] matrix accum, [R] glm}
\texttt{\_newline(#). display directive, [P] display}
\texttt{news command, [R] news}
\texttt{news, view subcommand, [R] view}
\texttt{newsletter, [U] 3 Resources for learning and using Stata}
\texttt{Newton–Raphson algorithm, [M-5] \texttt{moptimize()}, [M-5] \texttt{optimize()}, [R] ml}
\texttt{Newton–Raphson method, [M-5] \texttt{solvenl()}}
\texttt{Newton’s method, see \texttt{iteration}, Newton’s method}
\texttt{Neyman allocation, [SVY] \texttt{estat}}
\texttt{\texttt{nF()} function, [FN] Statistical functions, [M-5] \texttt{normal()}}
\texttt{\texttt{nFden()} function, [FN] Statistical functions, [M-5] \texttt{normal()}}
\texttt{\texttt{nFtail()} function, [FN] Statistical functions, [M-5] \texttt{normal()}}
noncentrality parameter, [PSS] power onemean,
[PSS] power twomeans, [PSS] power
pairedmeans, [PSS] power oneway,
[PSS] power twoway, [PSS] power repeated,
[PSS] Glossary
nonconformities, quality control, [R] qc
nonconstant variance, see robust, Huber/White/sandwich estimator of variance
noninformative model, see nonrecursive model
nondirectional test, see two-sided test
noninformative prior, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh, [BAYES] bayesstats ic,
[BAYES] Glossary
nonlinear
- autoregressive conditional heteroskedasticity, [TS] arch
- autoregressive conditional heteroskedasticity with a shift, [TS] arch
- Bayesian regression, [BAYES] bayesmh
- combinations of estimators, [R] nlcom
- combinations, predictions, and tests, [SVY] svi
- hypothesis test after estimation, [R] Ixtest,
  [R] margins, [R] margins, contrast,
  [R] margins, pwcompare, [R] nlcom,
  [R] predictnl, [R] testnl
- least squares, [R] nl, [SVY] svi estimation
- power autoregressive conditional heteroskedasticity, [TS] arch
- prediction, see multiple imputation, prediction regression, [R] boxcox, [R] nl, [R] nlsm
  [TE] teffects ra
- smoothing, [TS] tsmooth nl
- test, see estimation, test after time-series model, [TS] mswitch
- nonmetric scaling, [MV] mds, [MV] mdslong,
  [MV] mdsmat, [MV] Glossary
- nonmissing() function, [M-5] missing()
- nonmonotonic power, see saw-toothed power function
- nonnormed fit index, see Tucker–Lewis index
- nonparametric analysis,
- hypothesis tests,
  agreement, [R] kappa
  association, [R] spearman, [R] tabulate twoway
cusum, [R] cusum
equality of distributions, [R] ksmirnov,
[R] kwallis, [R] ranksum, [R] signrank
  medians, [R] ranksum
  proportions, [R] bitest, [R] prtest
  random order, [R] runtest
trend, [R] nptrend
- percentiles, [R] centile
- quantile regression, [R] qreg
noncentral
- beta density, [FN] Statistical functions
- beta distribution, [FN] Statistical functions
- chi-squared distribution, [FN] Statistical functions
- F density, [FN] Statistical functions
- F distribution, [FN] Statistical functions
- Student’s t density, [FN] Statistical functions
- Student’s t distribution, [FN] Statistical functions
- alpha, [PSS] Glossary, also see significance level
- item, [IRT] Glossary
- power, see power
- response model, [IRT] Glossary
- sample size, see sample-size
- significance level, [PSS] Glossary, see significance level
- nonadaptive Gauss–Hermite quadrature, see quadrature, Gauss–Hermite
- nonoverlapping confidence intervals, see confidence intervals, nonoverlapping
- non-parallel regression, see parallel regression, non-
- normal distribution, see Gaussian distribution, normal
- normal density, [FN] Statistical functions
- normal distribution, see Gaussian distribution, normal
- normal density, [FN] Statistical functions
- normal distribution, see Gaussian distribution, normal
- normal density, [FN] Statistical functions
- normal distribution, see Gaussian distribution, normal
- normal density, [FN] Statistical functions
- normal distribution, see Gaussian distribution, normal
- normal density, [FN] Statistical functions
- normal distribution, see Gaussian distribution, normal
- normal density, [FN] Statistical functions
- normal distribution, see Gaussian distribution, normal
- normal density, [FN] Statistical functions
- normal distribution, see Gaussian distribution, normal
- normal density, [FN] Statistical functions
nonparametric analysis, continued

 ROC analysis, [R] roc
 estimation, [R] rocreg
graphs, [R] rocregplot
test equality of areas, [R] roccomp
without covariates, [R] roctab
smoothing, [R] kdensity, [R] lowess, [R] lpoly,
[R] smooth
survival analysis,
Kaplan–Meier curves, [ST] stsvs
log rank and other tests of equality, [ST] ststest
Nelson–Aalen curves, [ST] stsglm

treatment effect, [TE] teffects nnmatch,
[TE] teffects psmatch
nonparametric methods, [MV] discrim knn,
[MV] Glossary
nonrecursive model, [SEM] Glossary

stability of, [SEM] estat stable,
[SEM] example 7
nonstationary time series, [TS] dfgls,
[TS] perron, [TS] vec intro, [TS] vec

nonzero null, [PSS] power onemean,
[PSS] power paireddmeans, [PSS] power oneproportion,
[PSS] power onevariance, [PSS] power onecorrelation

nopreserve option, [P] nopreserve option
	norm, [M-6] Glossary

norm() function, [M-5] norm()
normal distribution and normality,
examining distributions for, [R] diagnostic plots,
[R] inv

generating multivariate data, [D] drawnorm

probability and quantile plots, [R] diagnostic plots
test for, [R] sktest, [R] swilk
transformations to achieve, [R] boxcox, [R] ladder,
[R] inskew0

normal probability plots, [G-2] graph other

normal() function, [FN] Statistical functions,
[M-5] normal()
normal, distribution, continued

natural log of cumulative, [FN] Statistical functions
sample from multivariate, [FN] Statistical functions

normalden() function, [FN] Statistical functions,
[M-5] normal()
normality, mvtest subcommand, [MV] mvtest
normality
normality test, [MV] mvtest normality
after VAR or SVAR, [TS] varnorm
after VEC, [TS] vecnorm

normality,
conditional, [SEM] intro 4, [SEM] Glossary

joint, [SEM] intro 4, [SEM] Glossary

normalization, [MV] ca, [MV] mca, [MV] mds,
[MV] rotate, [MV] Glossary

normalization constraints, see constraints, normalization
normalized residuals, [SEM] estat residuals,
[SEM] methods and formulas for sem,
[SEM] Glossary

normally distributed random numbers, [FN] Random-number functions, [R] set rng,
[R] set seed
not concave message, [R] maximize

Not Elsewhere Classified, see Stata Blog
not equal operator, [U] 13.2.3 Relational operators

not operator, [U] 13.2.4 Logical operators

notable option, see gsem option notable, see sem option notable

note() option, [G-3] title_options

notes

command, [D] notes
drop command, [D] notes
list command, [D] notes
renumber command, [D] notes
replace command, [D] notes
search command, [D] notes

notes, cluster subcommand, [MV] cluster notes

notes on estimation results, [R] estimates notes

notes,

cluster analysis, [MV] cluster notes
creating, [D] notes, [D] varmanage
editing, [D] notes, [D] varmanage

notes, estimates subcommand, [R] estimates notes

notifyuser, set subcommand, [R] set

noupdate option, [MI] noupdate option

novarabbrev command, [P] varabbrev

noxconditional option, see sem option noxconditional

NPARCH, see nonlinear power autoregressive
continuous heteroskedasticity

npnchi2() function, [FN] Statistical functions,
[M-5] normal()
npnF() function, [FN] Statistical functions,
[M-5] normal()
npnt() function, [FN] Statistical functions,
[M-5] normal()
nproc, estat subcommand, [R] rocreg postestimation
null command, [R] nptrend
NR algorithm, [R] ml
NRM, see nominal response model
nrm, irt subcommand, [IRT] irt nrm, [IRT] irt nrm postestimation
nrtolerance() option, [R] maximize
nt() function, [FN] Statistical functions, [M-5] normal()
ndden() function, [FN] Statistical functions, [M-5] normal()
nntail() function, [FN] Statistical functions, [M-5] normal()
null correlation, [PSS] power hypothesis and alternative hypothesis, [PSS] power,
[PSS] power onemean, [PSS] power twomeans, [PSS] power pairedmeans, [PSS] power
oneproportion, [PSS] power twopropotions, [PSS] power pairedproportions, [PSS] power
onevariance, [PSS] power twovariances, [PSS] power onecorrelation, [PSS] power
twocorrelations, [PSS] power oneway, [PSS] power twoway, [PSS] power repeated,
[PSS] power cmh, [PSS] power mcc, [PSS] power trend, [PSS] power cox,
mean, [PSS] power, [PSS] power onemean, [PSS] power oneproportion, [PSS] power
onecorrelation, [PSS] unbalanced designs
mean difference, [PSS] power, [PSS] power
pairedmeans parameter, [PSS] Glossary, see null value
proportion, [PSS] power standard deviation, [PSS] power, [PSS] power
onevariance value, [PSS] power, [PSS] Glossary
variance, [PSS] power, [PSS] power onevariance
nullmat() function, [FN] Matrix functions, [P] matrix define
number of events, see number of failures
of failures, [PSS] power cox, [PSS] power exponential, [PSS] power
logrank to string conversion, see string functions
number, confirm subcommand, [P] confirm
number to string conversion, see string functions
numbered
numbers, [U] 12.2 Numbers
formatting, [D] format
mapping to strings, [D] encode, [D] label
numeric list, [P] numlist, [P] syntax,
[U] 11.1.8 numlist
numeric value labels, [D] labelbook
numerical precision, [U] 13.12 Precision and problems therein
numlabel command, [D] labelbook
numlist command, [P] numlist, [U] 11.1.8 numlist
N-way analysis of variance, [R] anova
N-way multivariate analysis of variance, [MV] manova

O
object, [P] class
objective prior, see noninformative prior
object-oriented programming, [M-2] class,
objects, size of, [G-4] relative size
.objkey built-in class function, [P] class
.objtype built-in class function, [P] class
oblimax rotation, [MV] rotate, [MV] rotatemat,
[MV] Glossary
oblimin rotation, [MV] rotate, [MV] rotatemat,
[MV] Glossary
oblique rotation, [MV] factor postestimation,
oblique transformation, see oblique rotation
obs parameter, [D] describe, [D] obs
obs, set subcommand, [D] obs, [R] set
observational data, [R] correlate, [R] heckman,
[R] ivregress, [R] logit, [R] mean, [R] regress,
[R] summarize, [R] tabulate oneway,
[R] tabulate twoway, [R] ttest, [R] ztest,
[TE] intro, [TE] treatment effects,
postestimation, [TE] etregress, [TE] etregress
postestimation, [TE] stteffects, [TE] stteffects
postestimation, [TE] tebalance, [TE] tebalance
box, [TE] tebalance density, [TE] tebalance
overid, [TE] tebalance summarize, [TE] ttests,
[TE] ttests intro, [TE] ttests intro
advanced, [TE] ttests aipw, [TE] ttests
ipw, [TE] ttests ipwra, [TE] ttests
mismatch, [TE] ttests overlap, [TE] ttests
postestimation, [TE] ttests pmatch,
[U] 20 Estimation and postestimation
commands, [U] 26.4 Structural equation
modeling (SEM), [U] 26.19 Models with
tests, [U] 26.21 Multilevel mixed-effects
tests, [U] 26.22 Survival-time (failure-time)
tests, [U] 26.23 Treatment-effect models,
and cluster analysis
observational study, [PSS] power, [PSS] Glossary
observations, adding, [D] insobs
built-in counter variable, [U] 11.3 Naming
conventions
complete and incomplete, [MI] Glossary
creating dataset of, [D] collapse
dropping, [D] drop
observations, continued
  dropping duplicate, [D] duplicates
  duplicating, [D] expand
  duplicating, clustered, [D] expandcl
  identifying duplicate, [D] duplicates
  increasing number of, [D] obs
  inserting, [D] insobs
  marking, [P] mark
  maximum number of, [D] memory, [U] 6 Managing
  memory
  ordering, [D] gsort, [D] sort
  transposing with variables, [D] xpose
observed information matrix, [R] ml, [R] vce_option,
  [SEM] Glossary, [XT] vce_options
observed level of significance, see p-value
observed variables, [SEM] intro 4, [SEM] Glossary
Ochiai coefficient similarity measure,
  [MV] measure_option
ocloglog option, see gsem option ocloglog
odbc
describe command, [D] odbc
exec() command, [D] odbc
insert command, [D] odbc
list command, [D] odbc
load command, [D] odbc
query command, [D] odbc
sqlfile() command, [D] odbc
ODBC data source, reading data from, [D] odbc,
  [U] 21.4 Transfer programs, [U] 21.5 ODBC
sources
odbcdriver, set subcommand, [D] odbc, [R] set
odbcmgr, set subcommand, [D] odbc, [R] set
odds ratio, [ME] meglm, [ME] meologit,
  [ME] meologit, [ME] meqrgolog, [PSS] power,
  [PSS] power twoproportions, [PSS] power
  pairedproportions, [PSS] power cmh,
  [PSS] power mcm, [PSS] Glossary,
  [R] eform_option, [R] epilog, [ST] Glossary,
  [SVY] svy estimation, [XT] xtloglog,
  [XT] xtgee, [XT] xlogit, [XT] xtablog,
  [XT] xtabreg
differences, [SVY] svy postestimation
estimation, [R] asclogit, [R] binreg, [R] clogit,
  [R] cloglog, [R] exlogistic, [R] glm, [R] logistic,
  [R] logit, [R] mlogit, [R] scobit
postestimation, [R] contrast, [R] exlogistic
  postestimation, [R] lincom
  _0Ex, [SEM] sem and gsem option covstructure()
off,
cmdlog subcommand, [R] log
log subcommand, [R] log
timer subcommand, [P] timer
Office Open XML, [M-5] _docx()()
Office, Microsoft, see Microsoft Office
offset between axes and data, setting,
  [G-3] region_options
offset() option, see gsem option offset()
one-sided test, continued
  [PSS] power oneway, [PSS] power cmh,
  [PSS] power mcc, [PSS] power trend,
  [PSS] power cox, [PSS] power exponential,
  [PSS] power logrank, [PSS] Glossary
one-step-ahead forecast, see static forecast
one-tailed test, [PSS] Glossary, also see one-sided test
onevariance, power subcommand, [PSS] power
  onevariance
one-way analysis of variance, [PSS] power,
  [PSS] power oneway, [PSS] Glossary,
  [R] kwallis, [R] lowney, [R] oneway
oneway command, [R] oneway
one-way repeated-measures ANOVA, [PSS] power
  repeated, [PSS] Glossary
oneway, power subcommand, [PSS] power oneway
online help, [U] 7 – more– conditions
opaccum, matrix subcommand, [P] matrix accum
open, file subcommand, [P] file
OpenOffice dates, [D] datetime
operating characteristic curve, [IRT] Glossary
operating system command, [D] ed, [D] copy, [D] dir,
  [D] erase, [D] mkdir, [D] rmdir, [D] shell,
  [D] type
operator, [M-2] op_arith, [M-2] op_assignment,
  [M-2] op_colon, [M-2] op_condition,
  [M-2] op_increment, [M-2] op_join,
  [M-2] op_kronecker, [M-2] op_logical,
  [M-2] op_range, [M-2] op_transpose,
  [M-6] Glossary, [P] matrix define,
  [U] 13.2 Operators
difference, [U] 11.4.4 Time-series varlists
lag, [U] 11.4.4 Time-series varlists
lead, [U] 11.4.4 Time-series varlists
order of evaluation, [U] 13.2.5 Order of evaluation, all operators
seasonal lag, [U] 11.4.4 Time-series varlists
OPG, see outer product of the gradient
oprobit command, [R] oprobit, [R] oprobit
  postestimation
oprobit option, see gsem option oprobit
oprobit regression, mixed-effects, [ME] meoprobit
_optimize() function, [M-5] optimize()
optimize() function, [M-5] optimize()
  _optimize_evaluate() function, [M-5] optimize()
  optimize_evaluate() function, [M-5] optimize()
  optimize_init() function, [M-5] optimize()
  optimize_init_() functions, [M-5] optimize()
  optimize_query() function, [M-5] optimize()
  optimize_result_() functions, [M-5] optimize()
options, [U] 11 Language syntax
  in a programming context, [P] syntax, [P] unab
  repeated, [G-4] concept: repeated options
  or operator, [U] 13.2.4 Logical operators
Oracle, reading data from, [D] odbc, [U] 21.4 Transfer programs
order command, [D] order
order() function, [M-5] sort()
order statistics, [D] egen, [R] lv
ordered
  complementary log-log regression, [SEM] Glossary
  logistic regression, [BAYES] bayesmh,
  estimation
  logistic regression imputation, see imputation, ordered logistic regression
  logit, [R] ologit, [SEM] example 35g
  probit, [R] heckoprobit, [R] oprobit,
    [SEM] example 35g, [SEM] example 36g
  probit regression, [BAYES] bayesmh,
  estimation
  probit with sample selection, [SVY] svy estimation
ordering
  observations, [D] gsort, [D] sort
  variables, [D] order, [D] sort
ordinal
  exposure, [PSS] power trend
  item, [IRT] Glossary
  model, [SEM] intro 5, [SEM] example 31g,
    [SEM] example 32g, [SEM] example 35g,
    [SEM] example 36g
  outcome, see outcomes, ordinal
  outcome model, see outcomes, ordinal
ordinary least squares, see linear regression
ordination, [MV] mds, [MV] Glossary
orgtype() function, [M-5] eltype()
orientationstyle, [G-4] orientationstyle
original data, [MI] Glossary
orthog command, [R] orthog
orthogonal
  matrix, [M-6] Glossary
  polynomial, [R] contrast, [R] margins, contrast,
  [R] orthog
  rotation, [MV] factor postestimation, [MV] rotate,
    [MV] rotatemat, [MV] Glossary
  transformation, see orthogonal rotation
orthogonalized impulse–response function, [TS] irf,
  [TS] var intro, [TS] vec intro, [TS] vec,
  [TS] Glossary
orthogonalized basis, [P] matrix svd
orthopoly command, [R] orthog
other graph commands, [G-2] graph other
other, query subcommand, [R] query
outcome model, [TE] eteffects, [TE] etpoisson,
  [TE] etregress, [TE] eteffects intro advanced,
  [TE] tteffects aipw, [TE] tteffects ipwra,
outcomes,
  binary,
  complementary log-log, [R] cloglog,
    [XT] xtloglog
generalized estimating equations, [XT] xtgee
outcomes, binary, continued

- glm for binomial family, [BAYES] bayesmh,
  - [R] binreg, [R] glm

  - [R] scobit, [XT] xtlogit, [XT] xtrtreg


- probit, [BAYES] bayesmh, [R] biprobit,
  - [R] heckprobit, [R] hetprobit, [R] ivprobit,
  - [R] probit, [XT] xtprob

- ROC analysis, [R] rocfit, [R] rocreg

categorical,

- logistic, [BAYES] bayesmh, [IRT] irt nrm, [IRT] irt hybrid, [R] ascllogit, [R] clogit,
  - [R] mlogit, [R] nlogit

- probit, [BAYES] bayesmh, [R] asprobit,
  - [R] mprobit


- count,
  - generalized estimating equations, [XT] xtggee


- negative binomial, [R] nbreg, [R] nbpoisson,
  - [R] zlnb, [XT] xtnbreg

- Poisson, [BAYES] bayesmh, [R] poisson,
  - [R] expoisson, [R] ivpoisson, [R] poisson,
  - [R] tpoisson, [R] zip, [TE] tpoisson,
  - [XT] xtpoisson
treatment effects, [TE] eteffects, [TE] etpoisson,

fractional,

- beta, [R] betareg

- fractional response, [R] fracreg

- multinomial, see categorical subentry, see ordinal subentry, see rank subentry

outcomes, continued

- ordinal,
  - logistic, [BAYES] bayesmh, [IRT] irt grm,

- multilevel mixed-effects, [ME] meologit,
  - [ME] meoprobit

- probit, [BAYES] bayesmh, [R] heckprobit,
  - [R] oprobit, [XT] xtprob

treatment effect, [TE] teffects multivalued

- polytomous, see categorical subentry, see ordinal subentry, see rank subentry

- rank,
  - logistic, [R] rologit

- probit, [R] asprobit

- survival,
  - competing risks, [ST] stcrreg
  - Cox, [ST] stcox
  - parametric, [ST] streg

- power and sample size, [PSS] power cox,
  - [PSS] power exponential, [PSS] power logrank
treatment effects, [TE] stteffects ipw, [TE] stteffects ipwra, [TE] stteffects ra,
  - [TE] stteffects wra

- outer
  - fence, [R] lv
  - product, [D] cross
  - product of the gradient, [R] ml, [R] vce_option,
  - [SEM] Glossary, [XT] vce_options

- outfile command, [D] outfile

- outliers, [R] lv, [R] qreg, [R] regress postestimation,
  - [R] rreg

- outlines, suppressing, [G-4] liststyle

- outlining regions, [G-3] region_options

- out-of-sample predictions, [R] predict, [R] predictnl,
  - [U] 20.10.3 Making out-of-sample predictions

- output,
  - query subcommand, [R] query
  - set subcommand, [P] quietly, [R] set

- output settings, [P] creturn

- output, coefficient table,
  - automatically widen, [R] set
display settings, [R] set showbaselevels
format settings, [R] set cformat
controlling the scrolling of, [R] more
displaying, [P] display, [P] smcl
formatting numbers, [D] format
printing, [R] translate, [U] 15 Saving and printing output

- output—log files
  - recording, [R] log
  - suppressing, [P] quietly

- outside values, [R] lv

- over() option, [G-2] graph bar, [G-2] graph box,
  - [G-2] graph dot

overid, estat subcommand, [R] gmm postestimation, [R] ivpoisson postestimation, [R] ivregress postestimation

overidentifying restrictions, [XT] Glossary


overlap, tteffects subcommand, [TE] tteffects overlap

overloading, class program names, [P] class

ovtest, estat subcommand, [R] regress postestimation

P

P charts, [G-2] graph other

P–P plot, [R] diagnostic plots

p-value, [SEM] Glossary

pac command, [TS] corrgram

pagesize, set subcommand, [R] more, [R] set

paging of screen output, controlling, [P] more, [R] more

paired

data, [PSS] Glossary
means, see means, paired
observations, see paired data
proportions, see proportions, paired
study, [PSS] power, [PSS] power mce
test, [PSS] Glossary


pairedmeans, power subcommand, [PSS] power pairedmeans

pairedproportions, power subcommand, [PSS] power pairedproportions

paired-sample test, [PSS] intro, [PSS] power, [PSS] power pairedmeans, [PSS] power pairedproportions

means, [PSS] power pairedmeans
proportions, [PSS] power pairedproportions

pairwise
combinations, [D] cross, [D] joinby
correlation, [R] correlate

pairwise, estat subcommand, [MV] mds postestimation

palette command, [G-2] palette


panel-corrected standard error, [XT] xtpcse, [XT] Glossary

panels, variable identifying, [XT] xtset

panelsetup() function, [M-5] panelsetup()

panelstats() function, [M-5] panelsetup()

panelsubmatrix() function, [M-5] panelsetup()

panelsubview() function, [M-5] panelsetup()

Paradox, reading data from, [U] 21.4 Transfer programs

parallel number list, [PSS] power parameter

constraints, [SEM] estat ginvariant, [SEM] Glossary

trace files, [MI] mi impute mvn, [MI] mi ptrace values, obtaining symbolic names, see gssem option coeflegend, see sem option coeflegend

parameterized curves, [D] range

parameters, [SEM] Glossary
combinations of, [D] range

system, see system parameters

parametric

methods, [MV] Glossary

spectral density estimation, [TS] psdensity

survival models, [ST] streg, [SVY] svy estimation

PARCh, see power autoregressive conditional heteroskedasticity

parsedistance, cluster subcommand, [MV] cluster programming utilities


partial

autocorrelation function, [TS] corrgram, [TS] Glossary
correlation, [R] pcorr
credit model, [IRT] Glossary

effects, [R] margins, [R] marginsplot
likelihood displacement value, [ST] Glossary

LMAX value, [ST] Glossary

regression leverage plot, [R] regress postestimation diagnostic plots

regression plot, [R] regress postestimation diagnostic plots

residual plot, [R] regress postestimation diagnostic plots


partition cluster-analysis methods, [MV] cluster kmeans and kmedians, [MV] Glossary

partition clustering, see partition cluster-analysis methods
permute prefix command, [R] permute
person location, [IRT] Glossary
personal command, [P] sysdir
PERSONAL directory, [P] sysdir, [U] 17.5 Where does Stata look for ado-files?

person-time, [ST] stptime
pformat, set subcommand, [R] set, [R] set cformat
pharmaceutical statistics, [R] pk, [R] pksum
pharmacokinetic data, [R] pk,
[R] pkcollapse, [R] pkcross, [R] pksumm

pharmacokinetic plots, [G-2] graph other
phase function, [TS] Glossary
Phillips–Perron test, [TS] pperron
phtest, estat subcommand, [ST] stcox PH-assumption tests

_pi built-in variable, [U] 11.3 Naming conventions
pi() function, [M-5] sin()
pi, value of, [U] 11.3 Naming conventions,
[U] 13.4 System variables (_variables)

pie charts, [G-2] graph pie
pie, graph subcommand, [G-2] graph pie
piece macro extended function, [P] macro

cubic functions, [R] mkspline
linear functions, [R] mkspline

Pillai’s trace statistic, [MV] canon, [MV] manova,
[MV] mvtest means, [MV] Glossary

pi

pinnable, set subcommand, [R] set
_pinv() function, [M-5] pinv()
pinv() function, [M-5] pinv()

pk, see pharmacokinetic data
pkcollapse command, [R] pkcollapse
pkcross command, [R] pkcross
pkexamine command, [R] pkexamine
.pk filename suffix, [R] net
pkshape command, [R] pkshape
pksumm command, [R] pksumm

Plackett–Lucy model, [R] rologit
plain ASCII, [I] Glossary
platforms for which Stata is available,

[U] 5.1 Platforms
play, graph subcommand, [G-2] graph play
play() option, [G-3] play_option
playsnd, set subcommand, [R] set
plegend() option, [G-3] legend_options
plot, definition, [G-4] pstyle
plot, ml subcommand, [R] ml
plot region, [G-3] region_options
suppressing border around, [G-3] region_options
plotregion() option, [G-3] region_options
plotregionstyle, [G-4] plotregionstyle
plottypes
base, [G-3] advanced_options
derived, [G-3] advanced_options

plugin option, [P] plugin, [P] program
plugin
Java, [P] java, [P] javacall
loading, [P] plugin
plural() function, [FN] String functions
PLUGINS directory, [P] sysdir, [U] 17.5 Where does Stata look for ado-files?

PMM imputation, see imputation, predictive mean matching

PNG, [G-3] png_options

pnorm command, [R] diagnostic plots
point estimate, [SVY] Glossary
point-and-click analysis, see graphical user interface

pointers, [M-2] pointers, [M-2] ftof,
points, connecting, [G-3] connect_options,
[G-4] connectstyle

poisson() function, [FN] Statistical functions
Poisson

distribution,
confidence intervals, [R] ci
cumulative, [FN] Statistical functions
inverse cumulative, [FN] Statistical functions
inverse reverse cumulative, [FN] Statistical functions

regression, see Poisson regression
reverse cumulative, [FN] Statistical functions
probability mass function, [FN] Statistical functions
regression, [R] nbreg, [R] poisson,

[SEM] example 34g, [SEM] example 39g,
Bayesian, [BAYES] bayesmh
censored, [R] cpoisson
fixed-effects, [XT] xtpoisson
generalized linear model, [R] glm
mixed-effects, [ME] mepoisson,
[ME] meqppoisson
model, [XT] Glossary
population-averaged, [XT] xtgee, [XT] xtpoisson
random-effects, [XT] xtpoisson
truncated, [R] tpoisson
zero-inflated, [R] zip

poisson command, [R] nbreg, [R] poisson,
[R] poisson postestimation

poisson() function, [M-5] normal()
poisson option, see gsem option poisson
poissonp() function, [FN] Statistical functions,
[M-5] normal()
poissontail() function, [FN] Statistical functions,
[M-5] normal()

polar coordinates, [D] range
polyadd() function, [M-5] polyeval()
polyderiv() function, [M-5] polyeval()
polydiv() function, [M-5] polyeval()
polyeval() function, [M-5] polyeval()
polyinteg() function, [M-5] polyeval()
polymorphism, [P] class
polymult() function, [M-5] polyeval()
polynomial smoothing, see local polynomial smoothing
polynomials, [M-5] polyeval()
  fractional, [R] fp, [R] mfp
  orthogonal, [R] orthog
  smoothing, see local polynomial
polyroots() function, [M-5] polyeval()
polysolve() function, [M-5] polyeval()
polytomous, [IRT] Glossary
polytomous logistic regression, [SVY] svy estimation
polytomous outcome model, see outcomes, polytomous
POMs, see potential-outcome means
pooled estimates, [R] epitab
pooled estimator, [XT] Glossary
pooling step, [MI] intro substantive
population error, [SEM]
population parameter, [PSS]
power
population pyramid, [G-2] graph twoway bar
population size, [PSS] power, [PSS] power onemean, [PSS] power pairedmeans
population standard deviation, see subpopulation, standard deviations of
populations,
  diagnostic plots, [R] diagnostic plots
  standard, [R] dstdize
  testing equality of, see distributions, testing equality of
testing for normality, [R] sktest, [R] swilk
positive effect size, [PSS] power, [PSS] Glossary
post command, [P] postfile
postclose command, [P] postfile
posterior
  interval, see credible interval
posterior, continued
  mean, [BAYES] intro, [BAYES] bayes,
  odds, [BAYES] intro, [BAYES] bayes,
  probabilities, [BAYES] intro, [BAYES] bayes,
postest command, [R] postest
postestimation
  command, [BAYES] bayesgraph,
postestimation, command, continued

postestimation, predicted values, [SEM] intro 7,
[SEM] example 14, [SEM] example 28g,
[SEM] predict after gsem, [SEM] predict after sem

PostScript, [G-2] graph export, [G-3] eps_options,

poststratification, [SVY] poststratification,
[SVY] Glossary

power, continued

autoregressive conditional heteroskedasticity,
[TS] arch

curve, [PSS] power, [PSS] power, graph,
[PSS] Glossary
determination, [PSS] power, [PSS] power
onemean, [PSS] power twomeans, [PSS] power
pairedmeans, [PSS] power oneproportion, [PSS] power
pairedproportions, [PSS] power onevariance, [PSS] power
twovariances, [PSS] power onecorrelation, [PSS] power
twocorrelations, [PSS] power oneway, [PSS] power twoway,
[PSS] power repeated, [PSS] power cmh, [PSS] power
mcc, [PSS] power trend, [PSS] power cox, [PSS] power exponential,
[PSS] power logrank, [PSS] unbalanced
designs, [PSS] Glossary

function, [PSS] Glossary

graph, see power curve
tabular output, [PSS] power, graph

transformations, [R] boxcox, [R] inskew

power

cmh command, [PSS] power cmh

cmd, [PSS] intro, [PSS] GUI, [PSS] power,
[PSS] power, graph, [PSS] power, table

cox command, [PSS] power cox

exponential command, [PSS] power exponential

logrank command, [PSS] power logrank

mcc command, [PSS] power mcc

onecorrelation command, [PSS] power

oneproportion command, [PSS] power

oneproportion

onevariance command, [PSS] power onevariance

oneway command, [PSS] power oneway

pairedmeans command, [PSS] power pairedmeans

pairedproportions command, [PSS] power pairedproportions

repeated command, [PSS] power repeated

trend command, [PSS] power trend

twocorrelations command, [PSS] power
twocorrelations

twomeans command, [PSS] power twomeans

twoproporotions command, [PSS] power
twoproporotions

twovariances command, [PSS] power
twovariances

twoway command, [PSS] power twoway

power, raise to, function, see arithmetic operators

pperron command, [TS] pperron


pragama unset, [M-2] pragama

pragama unused, [M-2] pragama

prais command, [TS] prais, [TS] prais postestimation

Prais–Winsten regression, [TS] prais, [TS] prais

postestimation, [TS] Glossary, [XT] xtpcse
precision, [U] 13.12 Precision and problems therein
pre determined variable, [XT] Glossary
predict command, [P] _predict
predict, estat subcommand, [R] exlogistic postestimation
predict, mi subcommand, [MI] mi predict
predicted values, see postestimation, predicted values
predictions, [R] predict, [R] predictnl, [SVY] svy postestimation, see multiple imputation, prediction
Bayesian, see Bayesian, prediction obtaining after estimation, [MI] mi predict, [P] _predict
standard error of, [R] glm, [R] predict, [R] regress postestimation
predictive margins, [SVY] Glossary, [U] 20.15 Obtaining marginal means, adjusted predictions, and predictive margins
predictive mean matching imputation, see imputation, predictive mean matching
predictnl command, [R] predictnl, [SVY] svy postestimation
predictnl, mi subcommand, [MI] mi predict
Pregibon delta beta influence statistic, see delta beta influence statistic
pre processor commands, [R] #review
preserve command, [P] preserve
preserving data, [P] preserve
preserving user’s data, [P] preserve
pretreatment mean, see means, pretreatment
prevalence studies, see case–control data
prevented fraction, [R] epitab
prewhiten, [XT] Glossary
primary sampling unit, [SVY] svsdescribe, [SVY] svsys, [SVY] Glossary
priming values, [TS] Glossary
principal component analysis, [MV] pca, [MV] Glossary
factors analysis, [MV] factor
print, graph subcommand, [G-2] graph print
print command, [R] translate
printf() function, [M-5] printf()
profiles, estat subcommand, [MV] ca
  postestimation

program
  define command, [P] plugin, [P] program,
  [P] program properties
  dir command, [P] program
  drop command, [P] program
  list command, [P] program
  program properties, [P] program properties
  programmer’s commands and utilities, [MI] mi select,
  [MI] styles, [MI] technical

programming, [P] syntax
  cluster analysis, [MV] cluster programming
  utilities
  cluster subcommands, [MV] cluster programming
  subroutines
  cluster utilities, [MV] cluster programming
  subroutines
  dialog, [P] dialog programming
  estat, [P] estat programming
  functions, [M-4] programming
  limits, [R] limits
  Mac, [P] window programming, [P] window
  fopen, [P] window manage, [P] window menu,
  [P] window push, [P] window stopbox
  menus, [P] window programming, [P] window
  menu
  rotations, [MV] rotate
  use, [M-1] ado
  Windows, [P] window programming, [P] window
  fopen, [P] window manage, [P] window menu,
  [P] window push, [P] window stopbox

programs, clear subcommand, [D] clear

programs,
  adding comments to, [P] comments
  debugging, [P] trace
  dropping, [P] discard
  looping, [P] continue
  user-written, see ado-files

Project Manager, [P] Project Manager
projection matrix, diagonal elements of, [R] binreg
  postestimation, [R] clogit postestimation,
  [R] glm postestimation, [R] logistic
  postestimation, [R] logit postestimation,
  [R] regress postestimation, [R] rege
  postestimation
  projection plot, [G-2] graph twoway contour,
  [G-2] graph twoway contourline
  projangrman command, [P] Project Manager
  promax power rotation, [MV] rotate, [MV] rotatemat,
  [MV] Glossary
  promax rotation, [MV] rotate
  propensity score, [TE] stteffects postestimation,
  [TE] tteffects intro, [TE] tteffects intro advanced,
  [TE] tteffects postestimation, [TE] tteffects
  psmatch, [TE] Glossary

propensity-score matching, [TE] tteffects intro,
  [TE] tteffects intro advanced, [TE] tteffects
  psmatch, [TE] Glossary

proper imputation method, [MI] intro substantive
proper values, [M-5] eigensystem(
  properties, [P] program properties
  properties macro extended function, [P] macro
  proportion command, [R] proportion,
  [R] proportion postestimation

proportional
  estimation, also see Cox proportional hazards
  model
  hazards models, see survival analysis
  odds assumption, [R] ologit
  relaxed, [R] slogit
  odds model, [R] ologit
  sampling, [D] sample, [R] bootstrap

proportions, [PSS] power
  confidence intervals for, [R] ci
  control-group, [PSS] power twopropor
  tions, [PSS] power cmh
  correlated, see proportions, paired
  discordant, [PSS] power pairedproportions
  estimating, [R] proportion
  experimental-group, [PSS] power twopropor
  tions, [PSS] power cmh
  independent, see proportions, two-sample
  marginal, [PSS] power pairedproportions,
  [PSS] Glossary
  of exposed cases, [PSS] power mccc
  of exposed controls, [PSS] power mccc
  one-sample, [PSS] power oneproportion
  paired, [PSS] power pairedproportions
  stratified test, [PSS] power cmh
  survey data, [SVY] svy estimation,
  [SVY] svy: tabulate oneway,
  [SVY] svy: tabulate twoway
  test of marginal homogeneity, [PSS] power mccc
  testing equality of, [R] bitest, [R] prtest
  two-sample, [PSS] power twopropor
  tions, [PSS] power pairedproportions

proportions,
  ci subcommand, [R] ci
  cii subcommand, [R] cii

proposal distribution, [BAYES] intro,
  [BAYES] bayesmh, [BAYES] bayesgraph,
  [BAYES] Glossary
  prospective study, [PSS] power, [PSS] Glossary,
  [R] epitab, also see incidence studies

protected, [M-2] class

proximity, [MV] Glossary
prtest command, [R] prtest
prtesi command, [R] prtest
psdensity command, [TS] psdensity
pseudo R-squared, [R] maximize
pseudoguessing parameter, [IRT]IRT pseudoinverse, [M-5] pinv()
pseudolikelihood, [SVY] Glossary
pseudosigmas, [R] lv
psi function, [FN] Mathematical functions
psmatch, teffects subcommand, [TE] teffects psmatch
PSS analysis, see power and sample-size analysis
psstyle, [G-4] pstyle
PSU, see primary sampling unit
.ptrace file, [MI] mi impute mvn, [MI] mi ptrace
ptrace, mi subcommand, [MI] mi ptrace
public, [M-2] class
push, window subcommand, [P] window programming, [P] window push
putexcel
clear command, [P] putexcel, [P] putexcel advanced
command, [P] putexcel, [P] putexcel advanced
describe command, [P] putexcel, [P] putexcel advanced
set command, [P] putexcel, [P] putexcel advanced
putmata command, [D] putmata
p-value, [PSS] Glossary
pwcorr command, [R] correlate
pwd command, [D] cd
pwd() function, [M-5] chdir()
pweight, see sampling weight
[pweight=exp] modifier, [U] 11.1.6 weight, [U] 20.23.3 Sampling weights
pwmean command, [R] pwmean, [R] pwmean postestimation
pyramid, population, [G-2] graph twoway bar
q
Q-Q plot, [R] diagnostic plots
Q statistic, see portmanteau statistic
qc charts, see quality control charts
qchi command, [R] diagnostic plots
QDA, see quadratic discriminant analysis
qda, discrim subcommand, [MV] discrim qda
qfit, graph twoway subcommand, [G-2] graph twoway qfit
qfitci, graph twoway subcommand, [G-2] graph twoway qfitci
qfod() function, [D] datetime, [FN] Date and time functions, [M-5] date()
qqplot command, [R] diagnostic plots
qrd() function, [M-5] qrd()
qrdp() function, [M-5] qrd()
qreg command, [R] qreg, [R] qreg postestimation
_quadinv() function, [M-5] qrd()
qrsolve() function, [M-5] qrsolve()
_qtolerance() option, [R] maximize
quadchk command, [XT] quadchk
quadcolsum() function, [M-5] sum()
quadcorrelation() function, [M-5] mean()
quadcross() function, [M-5] quadcross()
quadcrossdev() function, [M-5] quadcross()
quadmeanvariance() function, [M-5] mean()
quad() function, [M-5] sign()
quadratic discriminant analysis, [MV] discrim qda,
[MV] Glossary
quadratic terms, [SVY] svy postestimation
Gauss–Hermit, [IRT] irt 1pl, [IRT] irt 2pl,
IRT irt 3pl, [IRT] irt grm, [IRT] irt nrm,
IRT irt pcm, [IRT] irt rsm, [IRT] irt hybrid,
IRT Glossary, [ME] me, [ME] mecloglog,
[ME] meglm, [ME] meologit, [ME] menbreg,
ME meologit, [ME] meoprobit,
ME mepoisson, [ME] meprobit,
ME meqreg2, [ME] meqreg3,
ME meqregpoisson, [ME] mestreg, [ME] mixed,
ME Glossary, [SEM] methods and formulas for gsem,
[XT] quadchk
IRT Glossary, [ME] me, [ME] mecloglog, [ME] meglm,
[ME] meologit, [ME] menbreg, [ME] meologit,
[ME] meoprobit, [ME] mepoisson,
[ME] meprobit, [ME] meqreg2, [ME] meqreg3,
[ME] meqregpoisson, [ME] mestreg, [ME] mixed,
[ME] Glossary, [SEM] methods and formulas for gsem
[ME] me, [ME] mecloglog, [ME] meglm,
[ME] meologit, [ME] menbreg, [ME] meologit,
[ME] meoprobit, [ME] mepoisson,
[ME] meprobit, [ME] meqreg2, [ME] meqreg3,
quadrature, mode-curvature adaptive Gauss–Hermite, continued
[M] meqrpoisson, [ME] mestreg, [ME] mixed,
[M] Glossary, [SEM] methods and formulas for gsem
nonadaptive Gauss–Hermite, see quadrature, Gauss–Hermite
quadrovs() function, [M-5] sum()
_quadrunningsum() function, [M-5] runningsum()
runningsum() function, [M-5] runningsum()
quadsum() function, [M-5] sum()
quadvariance() function, [M-5]
quadratic dependent variables, [BAYES] bayesmh,
[IRT] irt 1pl, [IRT] irt 2pl, [IRT] irt 3pl,
[IRT] irt grm, [IRT] irt nrm, [IRT] irt pcm,
[IRT] irt rsm, [IRT] irt hybrid, [ME] meclogog,
[ME] meglm, [ME] melogit, [ME] meologit,
[ME] meoprobit, [ME] meprobit,
[ME] meqrologit, [ME] meqrobit, [R] ainlogit, [R] asprobit,
[R] asprobit, [R] binreg, [R] bioprobit,
[R] brier, [R] clogit, [R] clogolog,
[R] csum,
[R] exlogistic, [R] glm, [R] heckprobit,
[R] heckprob, [R] hetprob, [R] ivprobit,
[R] logistic, [R] logit, [R] mlogit, [R] mprobit,
[R] nlogit, [R] ologit, [R] oprobit,
[R] probit, [R] robit, [R] rocreg, [R] rologit,
[R] scobit, [R] slogit, [SVY] syv estimation,
[U] 26.7 Binary-outcome qualitative dependent-
variable models, [U] 26.11 Multiple-outcome qualitative dependent-variable models,
[U] 26.20.4 Qualitative dependent-variable models with panel data, [XT] xtcloglog,
[XT] xtgee, [XT] xlogit, [XT] xtologit,
[XT] xtprobit, [XT] xtprobit, [XT] xtstreg
quality control charts, [G-2] graph other, [R] qc,
[R] srsbar
quantile command, [R] diagnostic plots
quantile--normal plots, [R] diagnostic plots
quantile plots, [G-2] graph other, [R] diagnostic plots
quantile--quantile plots, [G-2] graph other,
[R] diagnostic plots
quantile regression, [R] qreg
quantiles, see percentiles, displaying, see percentiles
quantiles, estat subcommand, [MV] mds
postestimation
quarter() function, [D] datetime, [FN] Date and time functions,
[M-5] date()
quarterly() function, [D] datetime, [D] datetime
translation, [FN] Date and time functions,
[M-5] date()
quartimax rotation, [MV] rotate, [MV] rotatemat,
[MV] Glossary
quartimin rotation, [MV] rotate, [MV] rotatemat,
[MV] Glossary
quasimaximum likelihood, [SEM] Glossary
Quattro Pro, reading data from, [U] 21.4 Transfer programs
query
command, [R] query
efficiency command, [R] query
query, continued
interface command, [R] query
mata command, [R] query
memory command, [D] memory, [R] query
network command, [R] query
other command, [R] query
output command, [R] query
trace command, [R] query
unicode command, [R] query
update command, [R] query
query,
cluster subcommand, [MV] cluster programming
utilities
estimates subcommand, [R] estimates store
file subcommand, [P] file
forecast subcommand, [TS] forecast query
graph subcommand, [G-2] graph query
dcid10 subcommand, [D] icd10
icd9 subcommand, [D] icd9
icd9p subcommand, [D] icd9
log subcommand, [R] log
mi subcommand, [M] mi describe
ml subcommand, [R] ml
net subcommand, [R] net
odbc subcommand, [D] odbc
translator subcommand, [R] translate
transmap subcommand, [R] translate
update subcommand, [R] update
webuse subcommand, [D] webuse
querybreakintr() function, [M-5] breakintr()
quick reference, [D] data types, [D] missing values
quietly prefix, [P] quietly
quit Mata, [M-3] end
quitting Stata, see exit command
quotes
to delimit strings, [U] 18.3.5 Double quotes
to expand macros, [P] macro, [U] 18.3.1 Local
macros
R
r() function, [FN] Programming functions
r() stored results, [P] discard, [P] return, [R] stored results, [U] 18.8 Accessing results calculated by
other programs, [U] 18.10.1 Storing results in
r()
r(functions) macro extended function, [P] macro
r(macros) macro extended function, [P] macro
r(matrices) macro extended function, [P] macro
r scalars) macro extended function, [P] macro
R charts, [G-2] graph other
R dates, [D] datetime
R, reading data from, [U] 21.4 Transfer programs
rtitle() option, [G-3] title_options
R², [SEM] estat eqgof
r2title() option, [G-3] title_options
ra, stteffects subcommand, [TE] stteffects ra
ra, teffects subcommand, [TE] teffects ra
radians, [FN] Mathematical functions
raise to a power function, [U] 13.2.1 Arithmetic operators
Ramsey test, [R] regress postestimation
random
coefficient, [BAYES] bayesmh, [ME] Glossary
linear form, [BAYES] Glossary
parameters, [BAYES] Glossary
variable, [BAYES] Glossary
intercept, [BAYES] bayesmh, [ME] Glossary,
[SEM] example 38g
model parameter, [BAYES] intro, [BAYES] bayes,
[BAYES] bayesmh
order, test for, [R] runtest
sample, [D] sample, [R] bootstrap, [U] 21.3 If you run out of memory
slope, [SEM] example 38g
walk, [TS] Glossary
random-coefficients
linear regression, [XT] xtrc
model, [BAYES] bayesmh, [U] 26.20.7 Random-coefficients model with panel data,
[XT] Glossary
random-effects model, [BAYES] bayesmh,
[ME] Glossary, [R] anova, [R] loneway,
[SEM] example 38g, [SEM] Glossary,
[XT] xtabond, [XT] xttologit, [XT] xtddp,
[XT] xttdpsys, [XT] xtgee, [XT] xthtaylor,
[XT] xtinreg, [XT] xttivreg, [XT] xttlogit,
[XT] xtnbreg, [XT] xttobit, [XT] xtoprobit,
[XT] xtpoisson, [XT] xtprobit, [XT] xtregr,
[XT] xtrc, [XT] xstreg, [XT] xttobit,
[XT] Glossary
multilevel mixed-effects models, [ME] me,
[ME] mecloglog, [ME] meglm, [ME] melogit,
[ME] memlogit, [ME] meprobit, [ME] mepoisson,
randomized controlled trial study, [PSS] power,
[PSS] Glossary
random-number
function, [FN] Random-number functions,
[M-5] runiform(), [R] set rng, [R] set seed,
[D] generate
generator, [FN] Random-number functions,
[M-5] runiform(), [R] set seed
generator setting, [R] set rng
seed, [BAYES] bayesmh, [MI] mi impute, [R] set seed
random-order test, [R] runtest
range
chart, [R] qc
of data, [D] codebook, [D] inspect, [R] lv,
[R] stem, [R] summarize, [R] table, [R] tabstat,
[XT] xtsum
operators, [M-2] op_range
plots, [G-3] rcap_options
spikes, [G-3] rspike_options
subscripts, see subscripts
vector, [M-5] range()
range command, [D] range
range() function, [M-5] range()
rangen() function, [M-5] range()
rank correlation, [R] spearman
rank(), egen function, [D] egen
rank() function, [M-5] rank()
ranking data, [R] rologit
rank-order statistics, [D] egen, [R] signrank,
[R] spearman
rank-ordered logistic regression, see outcomes, rank
ranks of observations, [D] egen
ranksum command, [R] ranksum
Rao’s canonical-factor method, [MV] factor
rarea, graph twoway subcommand, [G-2] graph twoway rarea
Rasch models, see item response theory
rate ratio, [R] epitab, [ST] stir, [ST] stptime,
[ST] stsum, see incidence-rate ratio
rating scale model, [IRT] Glossary
ratio command, [R] ratio, [R] ratio postestimation
ratio of sample sizes, see allocation ratio
ratios, estimating, [R] ratio
ratios, survey data, [SVY] svy estimation,
[SVY] svy: tabulate twoway
RATS, reading data from, [U] 21.4 Transfer programs
raw data, [U] 12 Data
.raw file, [U] 11.6 Filenaming conventions
raw residuals, [SEM] methods and formulas for sem
rbar, graph twoway subcommand, [G-2] graph twoway rbar
rbeta() function, [FN] Random-number functions,
[M-5] runiform()
rbinomial() function, [FN] Random-number functions, [M-5] runiform()
r (return codes), see error messages and return codes
_r (built-in variable, [P] capture, [U] 13.4 System variables (_variables)
rcap, graph twoway subcommand, [G-2] graph twoway rcap
rcapssm, graph twoway subcommand, [G-2] graph twoway rcapssm
drchart command, [R] qc
rchi2() function, [FN] Random-number functions,
[M-5] runiform()
rc-class command, [P] program, [P] return,
[U] 18.8 Accessing results calculated by other programs
regression, continued
  graphing, [R] logistic, [R] regress postestimation
diagnostic plots
  grouped data, [R] intreg
  hurdle, [R] churdle
  increasing number of variables allowed, [R] matsize
  instrumental variables, [R] gmm, [R] ipoision,
[R] ivprobit, [R] ivregress, [R] ivtobit
  linear, see linear regression
  lines, see fits, adding
  system, [MV] mvreg, [R] gmm, [R] ipoision,
  truncated, [R] truncreg

regression (in generic sense), also see estimation
commands
  accessing coefficients and standard errors, [P] matrix
get, [U] 13.5 Accessing coefficients and
standard errors
  dummy variables, with, [XT] xtreg
  fixed-effects, [XT] xtreg
  instrumental variables, [XT] xtabond, [XT] xt, [XT] xtpd,
  random-effects, [XT] xtgee, [XT] xtvreg
regression adjustment, [TE] teffects intro, [TE] teffects
regression coefficient, [PSS] power trend, [PSS] power
cox
  one-sample, [PSS] power trend, [PSS] power cox
regression scoring, [MV] factor postestimation
regular expressions, [FN] String functions
regular variables, see variables, regular
rejection region, [PSS] Glossary
relational operators, [U] 13.2.3 Relational operators
relative
difference function, [FN] Mathematical functions
efficiency, [MI] mi estimate, [MI] mi predict,
[MI] Glossary
risk, [PSS] power, [PSS] power twoproportions,
[PSS] power pairedproportions, [PSS] Glossary,
[R] epstab
variance increase, [MI] mi estimate, [MI] mi
predict, [MI] Glossary
relative-risk ratio, [R] eform_option, [R] lincom,
[R] mlogit
relativesize, [G-4] relativesize
reldif() function, [FN] Mathematical functions,
[M-5] reldif()
release marker, [P] version
releases, compatibility of Stata programs across,
[P] version
reliability, [MV] alpha, [MV] factor, [R] brier,
[R] eivreg, [R] icer, [R] intreg, [R] loneaway,
[R] poisson, [SEM] intro 5. [SEM] intro 12,
[SEM] example 24, [SEM] gsem model
description options, [SEM] sem and gsem
option reliability(), [SEM] sem model
description options, [SEM] Glossary,
[ST] survival analysis, [ST] discrete,
reliability, continued
  [ST] Itable, [ST] st, [ST] stcox, [ST] streg,
[ST] streg
reliability option, see gsem option
reliability(), see sem option
reliability()
reliability theory, see survival analysis
remainder function, [FN] Mathematical functions
REML, see restricted maximum likelihood
removing
directories, [D] rmdir
files, [D] erase
r._En, [SEM] sem and gsem option covstructure()
rename,
  char subcommand, [P] char
  cluster subcommand, [MV] cluster utility
  graph subcommand, [G-2] graph rename
  irf subcommand, [TS] irf rename
  mata subcommand, [M-3] mata rename
  matrix subcommand, [P] matrix utility
  mi subcommand, [MI] mi rename
rename command, [D] rename, [D] rename group
rename for mi data, [MI] mi rename
rename graph, [G-2] graph rename
renamevar, cluster subcommand, [MV] cluster
utility
renaming variables, [D] rename, [D] rename group,
[MI] mi rename
renumber, notes subcommand, [D] notes
reordering data, [D] gsort, [D] order, [D] sort
reorganizing data, [D] reshape, [D] xpose
repair, ssd subcommand, [SEM] ssd
repeated DDF, see denominator degrees of freedom,
repeated
repeated measures, [MV] Glossary, [PSS] power
pairedmeans, [PSS] power pairedproportions
repeated options, [G-4] concept: repeated options
repeated, power subcommand, [PSS] power repeated
repeated-measures ANOVA, [PSS] power, [PSS] power
repeated, [R] anova
repeated-measures MANOVA, [MV] manova
repeating and editing commands, [R] #review,
[U] 10 Keyboard use
repeating commands, [D] by, [P] continue, [P] foreach,
[P] forvalues
replace command, [D] generate, [MI] mi passive,
[MI] mi xeq
replace, notes subcommand, [D] notes
replace option, [U] 11.2 Abbreviation rules
replace0, mi subcommand, [MI] mi replace0
replay,
  estimates subcommand, [R] estimates replay
  replay subcommand, [G-2] graph replay
replay() function, [FN] Programming functions,
[P] ereturn, [P] _estimates
replay graphs, [G-2] graph replay
replaying models, [SEM] intro 7

replicating clustered observations, [D] expandcl observations, [D] expand


restore, estimates subcommand, [R] estimates store _return subcommand, [P] _return snapshot subcommand, [D] snapshot


return codes, [P] rmsg, see error messages and return codes

return() function, [FN] Programming functions return value, [P] class returning results, [P] return class programs, [P] class


revkeyboard, set subcommand, [R] set


ringposstyle, [G-4] ringposstyle

risk, continued
ratio, [PSS] Glossary, [R] binreg, [R] epitab, also see relative risk
rline, graph twoway subcommand, [G-2] graph
twoway rline
rlogistic() function, [FN] Random-number functions, [M-5] runiform()
rm command, [D] erase
  _rmcoll command, [P] _rmcoll
  _rmdcoll command, [P] _rmdcoll
rmdir command, [D] rmdir
  _rmdir() function, [M-5] chdir()
rmdir() function, [M-5] chdir()
rmdcoll() function, [M-5] findexternal()
RMSEA, see root mean squared error of approximation
rmse, set subcommand, [P] rms, [R] set
rnbinomial() function, [FN] Random-number functions, [M-5] runiform()
rng set subcommand, [R] set, [R] set rng
rngstate() function, [M-5] runiform()
rngstate, set subcommand, [R] set, [R] set seed
rnormal() function, [FN] Random-number functions, [M-5] runiform()
robust, [SEM] Glossary
  regression, [R] betareg, [R] regress, [R] rreg, also see robust, Huber/White/sandwich estimator of variance
  standard errors, [XT] Glossary
test for equality of variance, [R] sptest
robust, see gsem option vce(), see sem option vce()
robust, Abadie–Imbens standard errors, [TE] tfeffects
  nmatch, [TE] tfeffects psmatch
robust, Huber/White/sandwich estimator of variance,
alternative-specific
  conditional logit model, [R] ascllogit
  multinomial probit regression, [R] asmpois
  rank-ordered probit regression, [R] asprobit
ARCH, [TS] arch
ARFIMA, [TS] arfima
ARIMA and ARMAX, [TS] arima
beta regression, [R] betareg
censored Poisson regression, [R] cpoisson
  competing-risks regression, [ST] stcrreg
  complementary log-log regression, [R] cloglog
  Cox proportional hazards model, [ST] stcox
dynamic-factor model, [TS] dfactor
exponential regression hurdle, [R] churdle
fixed-effects models,
  linear, [XT] xtreg
  Poisson, [XT] xtpoisson
fractional response regression, [R] fracreg
GARCH, [TS] arch
robust, Huber/White/sandwich estimator of variance, continued
  generalized linear models, [R] glm
    for binomial family, [R] binreg
generalized method of moments, [R] gmm,
  [R] ivpoisson
  heckman selection model, [R] heckman
hurdle regression, [R] churdle
instrumental-variables regression, [R] ivregress
interval regression, [R] intreg
linear dynamic panel-data estimation, [XT] xtabond,
  [XT] xtdpd, [XT] xtdpdpsys
linear regression, [R] regress
  constrained, [R] cnssreg
table, [R] churdle
truncated, [R] truncate
  with dummy-variable set, [R] areg
logistic regression, [IRT] irt 1pl, [IRT] irt 2pl,
  [IRT] irt 3pl, [IRT] irt hybrid, [R] logistic,
  [R] logit, also see logit regression subentry
table, [R] clogit
  unconditional, [R] clogit
  multinomial, [IRT] irt nrm, [IRT] irt hybrid,
    [R] mlogit
  ordered, [IRT] irt grm, [IRT] irt pcm, [IRT] irt
    rsm, [IRT] irt hybrid, [R] ologit
  rank-ordered, [R] rologit
  table, [R] scobit
table, [R] stereotype, [R] slogit
Markov-switching model, [TS] mswitch
maximum likelihood estimation, [R] ml, [R] mlexp
multilevel mixed-effects model, [ME] meclloglog,
  [ME] meglm, [ME] melogit, [ME] menbreg,
  [ME] melogit, [ME] menbreg,
  [ME] mepoisson, [ME] meprobit, [ME] mestreg,
  [ME] mixed
multinomial
  logistic regression, [IRT] irt 1pl, [IRT] irt 2pl,
    [IRT] irt 3pl, [IRT] irt hybrid, [R] logistic,
    [R] logit, also see logistic regression subentry
  nested, [R] nlogit
Markov-switching model, [TS] mswitch
maximum likelihood estimation, [R] ml, [R] mlexp
multilevel mixed-effects model, [ME] meclloglog,
  [ME] meglm, [ME] melogit, [ME] menbreg,
  [ME] melogit, [ME] menbreg,
  [ME] mepoisson, [ME] meprobit, [ME] mestreg,
  [ME] mixed
multinomial
  logistic regression, [IRT] irt 1pl, [IRT] irt 2pl,
    [IRT] irt 3pl, [IRT] irt hybrid, [R] logistic,
    [R] logit, also see logistic regression subentry
  nested, [R] nlogit
Markov-switching model, [TS] mswitch
maximum likelihood estimation, [R] ml, [R] mlexp
Markov-switching model, [TS] mswitch
maximum likelihood estimation, [R] ml, [R] mlexp
multilevel mixed-effects model, [ME] meclloglog,
  [ME] meglm, [ME] melogit, [ME] menbreg,
  [ME] melogit, [ME] menbreg,
  [ME] mepoisson, [ME] meprobit, [ME] mestreg,
  [ME] mixed
Markov-switching model, [TS] mswitch
maximum likelihood estimation, [R] ml, [R] mlexp
Markov-switching model, [TS] mswitch
maximum likelihood estimation, [R] ml, [R] mlexp
multilevel mixed-effects model, [ME] meclloglog,
  [ME] meglm, [ME] melogit, [ME] menbreg,
  [ME] melogit, [ME] menbreg,
robust, Huber/White/sandwich estimator of variance, continued
population-averaged models, [XT] xtggee
complementary log-log, [XT] xtcloglog
logit, [XT] xtologit
negative binomial, [XT] xtnbreg
Poisson, [XT] xtpoisson
probit, [XT] xtprob
Prais–Winsten and Cochrane–Orcutt regression,
[TS] prais
probit regression, [R] probit
bivariate, [R] biprobit
heteroskedastic, [R] hetprob
multinomial, [R] mprobit
ordered, [R] heckoprobit, [R] oprobit
with endogenous covariates, [R] ivprobit
with sample selection, [R] heckprobit
quantile regression, [R] qreg
random-effects model
complementary log-log, [XT] xtcloglog
linear, [XT] xtreg
logistic, [XT] xtologit, [XT] xtcloglog
parametric survival, [XT] xstreg
Poisson, [XT] xtpoisson
probit, [XT] xtprob
month, [XT] xtmeprobit
state-space model, [TS] sspace
structural equation modeling, [SEM] intro 8,
[SEM] sem option method( )
summary statistics,
mean, [R] mean
proportion, [R] proportion
ratio, [R] ratio
total, [R] total
tobit model, [R] tobit
with endogenous covariates, [R] ivtobit
treatment effect, [TE] eteffects, [TE] etpoisson,
[TE] etregress, [TE] tffects aipw, [TE] tffects ipw,
[TE] tffects ipwra, [TE] tffects ra
truncated
negative binomial regression, [R] tnreg
Poisson regression, [R] tpoisson
regression, [R] truncreg
unobserved-components model, [TS] ucm
with endogenous covariates,
Poisson regression, [R] iipoisson
probit regression, [R] ivprobit
tobit regression, [R] ivtobit
with endogenous regressors,
instrumental-variables regression, [R] ivregress
zero-inflated
negative binomial regression, [R] zinb
Poisson regression, [R] zp
robust, other methods of, [R] rreg, [R] smooth
**_robust** command, [P] _robust
robvar command, [R] sdtest
ROC, see receiver operating characteristic analysis
roccomp command, [R] roc, [R] roccomp
rocfit command, [R] rocfit, [R] rocfit postestimation
rocgold command, [R] roc, [R] roccomp
rocplot command, [R] rocfit postestimation
rocreg command, [R] rocreg, [R] rocreg postestimation, [R] rocregplot
rocregplot command, [R] rocregplot
roctab command, [R] roc, [R] roctab
Rogers and Tanimoto similarity measure,
[TS] measure_option
roh, [R] loneway
rolling command, [TS] rolling
rolling regression, [TS] rolling, [TS] Glossary
rologit command, [R] rologit, [R] rologit postestimation
root mean squared error of approximation, [SEM] estat
gof, [SEM] example 4, [SEM] methods and formulas for sem
rootograms, [G-2] graph other, [R] spikeplot
roots of polynomials, [M-5] polyeval( )
rotate command, [MV] factor postestimation,
[MV] pca postestimation, [MV] rotate
rotate, estat subcommand, [MV] canon postestimation
rotatecompare, estat subcommand, [MV] canon postestimation,
[MV] factor postestimation, [MV] pca postestimation
rotated
factor loadings, [MV] factor postestimation
principal components, [MV] pca postestimation
rotatemat command, [MV] rotatemat
rotation, [MV] factor postestimation, [MV] pca postestimation,
Bentler’s invariant pattern simplicity, see Bentler’s
invariant pattern simplicity rotation
biquartimax, see biquartimax rotation
biquartimin, see biquartimin rotation
Comrey’s tandem 1, see Comrey’s tandem 1 and 2
rotations
Comrey’s tandem 2, see Comrey’s tandem 1 and 2
rotations
covarimin, see covarimin rotation
Crawford–Ferguson, see Crawford–Ferguson rotation
equamax, see equamax rotation
factor parsimony, see factor parsimony rotation
minimum entropy, see minimum entropy rotation
oblimax, see oblimax rotation
oblimin, see oblimin rotation
oblique, see oblique rotation
orthogonal, see orthogonal rotation
parsimax, see parsimax rotation
partially specified target, see partially specified target
rotation
Procrustes, see Procrustes rotation
promax, see promax rotation
quartimax, see quartimax rotation
quartimin, see quartimin rotation
rotation, continued
toward a target, see toward a target rotation
varimax, see varimax rotation
round() function, [FN] Mathematical functions,
[M-5] trunc() 
row
of matrix, selecting, [M-5] select()
operators for data, [D] egen
stripes, [M-6] Glossary
roweq macro extended function, [P] macro
roweq, matrix subcommand, [P] matrix rownames
rowfirst(), egen function, [D] egen
rowfullnames macro extended function, [P] macro
row-join operator, [M-2] op_jo
rowlast(), egen function, [D] egen
row-major order, [M-6] Glossary
rowmax(), egen function, [D] egen
rowmax() function, [M-5] minmax()
rowmaxabs() function, [M-5] minmax()
rowmean(), egen function, [D] egen
rowmedian(), egen function, [D] egen
rowmin(), egen function, [D] egen
rowmin() function, [M-5] minmax()
rowminmax() function, [M-5] minmax()
rowmiss(), egen function, [D] egen
rowmissing() function, [M-5] missing()
rownames macro extended function, [P] macro
rownames, matrix subcommand, [P] matrix rownames
rownonmiss(), egen function, [D] egen
rownonmissing() function, [M-5] missing()
rownumb() function, [FN] Matrix functions,
[P] matrix define
rowpctile(), egen function, [D] egen
rows() function, [M-5] rows()
rows of matrix
appending to, [P] matrix define
operators, [P] matrix define
rowscalefactors() function, [M-5] _equilrc()
rowsd(), egen function, [D] egen
rowshape() function, [M-5] rowshape()
rowsf() function, [FN] Matrix functions, [P] matrix define
rowsum() function, [M-5] sum()
rowtotal(), egen function, [D] egen
Roy’s largest root test, [MV] canon, [MV] manova,
[MV] mvtest means, [MV] Glossary
union-intersection test, [MV] canon, [MV] mvtest means
union-intersection test, [MV] manova
rpoisson() function, [FN] Random-number functions,
[M-5] runiform()
rreg command, [R] rreg, [R] rreg postestimation
rscatter, graph twoway subcommand, [G-2] graph twoway rscatter
rseed() function, [M-5] runiform()
RSM, see rating scale model
rsm, irt subcommand, [IRT] irt rsm, [IRT] irt rsm postestimation
rspike, graph twoway subcommand, [G-2] graph twoway rspike
rt() function, [FN] Random-number functions,
[M-5] runiform()
Rubin’s combination rules, [MI] mi estimate, [MI] mi estimate using, [MI] mi predict
run command, [R] do, [U] 16 Do-files
runiform() function, [FN] Random-number functions,
[M-5] runiform(), [R] set seed
runiformint() function, [FN] Random-number functions,
[M-5] runiform()

runningsum() function, [M-5] runningsum()
runttest command, [R] runtest
Russell and Rao coefficient similarity measure,
[MV] measure_option
rvalue, class, [P] class
rvfplot command, [R] regress postestimation
diagnostic plots
RVI, see relative variance increase
rvrplot command, [R] regress postestimation
diagnostic plots
rweibull() function, [FN] Random-number functions,
[M-5] runiform()
rweibullph() function, [FN] Random-number functions,
[M-5] runiform()

S
s() function, [FN] Programming functions
s() stored results, [FN] Programming functions,
[P] return, [R] stored results,
[U] 18.10.3 Storing results in s()
s(macros) macro extended function, [P] macro
s1color scheme, [G-4] scheme s1
s1manual scheme, [G-4] scheme s1
s1mono scheme, [G-4] scheme s1
s1rcolor scheme, [G-4] scheme s1
s2color scheme, [G-4] scheme s2
s2gcolor scheme, [G-4] scheme s2
s2gmono scheme, [G-4] scheme s2
s2gmono scheme, [G-4] scheme s2
SAARCH, see simple asymmetric autoregressive
conditional heteroskedasticity
Sammon mapping criterion, [MV] Glossary
sample, [SVY] Glossary
sample command, [D] sample
sample, random, see random sample
analysis, see power and sample-size analysis
curve, [PSS] power, [PSS] Glossary
determination, [PSS] intro, [PSS] power,
[PSS] power onemean, [PSS] power twomeans,
[PSS] power pairedmeans, [PSS] power oneproportion, [PSS] power twoproportions,
[PSS] power pairedproportions, [PSS] power onevariance, [PSS] power twovariances,
[PSS] power onecorrelation, [PSS] power twocorrelations, [PSS] power oneway,
[PSS] power twoway, [PSS] power repeated,
[PSS] power cmh, [PSS] power mcc,
[PSS] power trend, [PSS] power cox,
[PSS] power exponential, [PSS] power logrank,
[PSS] unbalanced designs, [PSS] Glossary
rounding rules for, [PSS] unbalanced designs
sampling, [D] sample, [R] bootstrap, [R] bsample,
rate, [PSS] power, [PSS] power onemean,
[PSS] power pairedmeans
stage, [SVY] estat, [SVY] Glossary
unit, [SVY] survey, [SVY] Glossary, also see primary sampling unit
weight, [SVY] survey, [SVY] poststratification,
[SVY] Glossary, [U] 11.1.6 weight,
[U] 20.23.3 Sampling weights, also see survey data
with and without replacement, [SVY] Glossary
sandwich/Huber/White estimator of variance, see robust, Huber/White/sandwich estimator of variance
sargan, estat subcommand, [XT] xtabond postestimation, [XT] xtdpdsys postestimation
Sargan test, [XT] xtabond postestimation, [XT] xtdpdsys postestimation
SAS dates, [D] datetime
SAS, reading data from, [U] 21.4 Transfer programs
SAS XPORT format, [D] import sasxport
sasxport,
export subcommand, [D] import sasxport
import subcommand, [D] import sasxport
satopts() option, see sem option satopts()
Satterthwaite DDF, see denominator degrees of freedom, Satterthwaite
Satterthwaite’s t test, [PSS] power, [PSS] power twomeans, [PSS] Glossary
saturated model, [SEM] estat gof, [SEM] example 4,
[SEM] methods and formulas for sem, [SEM] Glossary
save,
label subcommand, [D] label
estimates subcommand, [R] estimates save
graph subcommand, [G-2] graph save
snapshot subcommand, [D] snapshot
save estimation results, [P] _estimates
save command, [D] save
saved results, see stored results
saveold command, [D] save
saving data, [D] import delimited, [D] outfile,
[D] save, [D] snapshot, also see exporting data
saving() option, [G-3] saving_option
saving results, [P] _estimates, [P] _return,
[R] estimates save
saw-toothed power function, [PSS] power oneproportion, [PSS] power twoproportions
sbknown, estat subcommand, [TS] estat sbknown
sbsingle, estat subcommand, [TS] estat sbsingle
confirm subcommand, [P] confirm
define command, [P] scalar
dir command, [P] scalar
drop command, [P] scalar
ereturn subcommand, [P] ereturn, [P] return
list command, [P] scalar
return subcommand, [P] return
scalar functions, [M-4] scalar
scalar model parameter, [BAYES] Glossary, see Bayesian, model parameters
scalar() function, [FN] Programming functions
scalar() pseudofunction, [P] scalar
scalars, [P] scalar
namespace and conflicts, [P] matrix, [P] matrix define
scale,
log, [G-3] axis_scale_options
range of, [G-3] axis_scale_options
reversed, [G-3] axis_scale_options
scale() option, [G-3] scale_option
scaling, [MV] mds, [MV] mds postestimation plots,
[MV] mdslong, [MV] mdsmat
scatter, graph twoway subcommand, [G-2] graph twoway scatter
scatteri, graph twoway subcommand, [G-2] graph twoway scatteri
scatterplot matrices, [G-2] graph matrix
scenarios, [TS] forecast, [TS] forecast adjust,
[TS] forecast clear, [TS] forecast coevector,
[TS] forecast create, [TS] forecast describe,
[TS] forecast drop, [TS] forecast estimates,
[TS] forecast exogenous, [TS] forecast identity, [TS] forecast list, [TS] forecast query,
[TS] forecast solve
Scheffé’s multiple-comparison adjustment, see multiple comparisons, Scheffé’s method
scheme() option, [G-3] scheme_option
scheme, set subcommand, [G-2] set scheme, [R] set schemes,
[G-2] set scheme, [G-3] play_option,
[G-3] scheme_option, [G-4] schemes intro,
[G-4] scheme economist, [G-4] scheme s1,
changing, [G-2] graph display
creating your own, [G-4] schemes intro
schemes, continued
  default, [G-2] set scheme
Schoenfeld residual, [ST] stcox PH-assumption
tests, [ST] stcox postestimation, [ST] stcrreg
postestimation
Schur
  form, [M-6] Glossary
_schurd() function, [M-5] schurd()
schurd() function, [M-5] schurd()
_schurdgroupby() function, [M-5] schurd()
schurdgroupby() function, [M-5] schurd()
_schurdgroupby_la() function, [M-5] schurd()
_schurdl() function, [M-5] schurd()
Schwarz information criterion, see Bayesian information
criterion
scientific notation, [U] 12.2 Numbers
s-class command, [P] program, [P] return, [R] stored
results, [U] 18.8 Accessing results calculated by
other programs
scobit command, [R] scobit, [R] scobit
postestimation
scope, class, [P] class
score, [MV] Glossary
score, matrix subcommand, [P] matrix score
score, ml subcommand, [R] ml
score plot, [MV] scoreplot, [MV] Glossary
score test, [PSS] power one-proportion,
ginvvariant, [SEM] estat mindices, [SEM] estat
coretests, [SEM] methods and formulas for
  sem, [SEM] Glossary
scoreplot command, [MV] discrim lda
  postestimation, [MV] factor postestimation,
  [MV] pca postestimation, [MV] scoreplot
scores, [R] predict, [SEM] Glossary
scores, obtaining, [U] 20.22 Obtaining scores
scoretests, estat subcommand, [SEM] estat
  scoretests
scoring, [MV] factor postestimation, [MV] pca
  postestimation, [P] matrix score
scree plot, [MV] scoreplot, [MV] Glossary
screeplot command, [MV] discrim lda
  postestimation, [MV] factor postestimation,
  [MV] pca postestimation, [MV] scoreplot
scrollbufsize, set subcommand, [R] set
scrolling of output, controlling, [P] more, [R] more
sd(), egen function, [D] egen
sd, estat subcommand, [SVY] estat
SDR, see successive difference replication
_sdr_options, [SVY] _sdr_options
sdtest command, [R] sdtest
sdtest command, [R] sdtest
se, estat subcommand, [R] exlogistic postestimation,
  [R] expoisson postestimation
_se[], [U] 13.5 Accessing coefficients and standard
  errors
search,
  icd9 subcommand, [D] icd9
  icd9p subcommand, [D] icd9
  ml subcommand, [R] ml
  net subcommand, [R] net
  notes subcommand, [D] notes
  view subcommand, [R] view
search command, [R] search, [U] 4 Stata’s help and
  search facilities
search_d, view subcommand, [R] view
search Internet, [R] net search
searchdefault, set subcommand, [R] search, [R] set
seasonal
  ARIMA, [TS] arima
difference operator, [TS] Glossary
  lag operator, [U] 11.4.4 Time-series varlists
  smoothing, [TS] tsmooth, [TS] tsmonth winters
secondary sampling unit, [SVY] Glossary
second-order latent variables, [SEM] Glossary
seconds() function, [D] datatime, [FN] Date and
time functions, [M-5] date()
seed, set subcommand, [R] set, [R] set seed
seek, file subcommand, [P] file
seemingly unrelated
  estimation, [R] surest
  regression, [R] nsur, [R] reg3, [R] sureg,
  [SEM] intro 5, [SEM] example 12,
  [SEM] Glossary, [TS] dfactor
segmentsize, set subcommand, [D] memory, [R] set
select() function, [M-5] select()
select() option, see sem option select()
select, mi subcommand, [MI] mi select
selectindex() function, [M-5] select()
selection models, [R] heckman, [R] heckoprobit,
  [R] heckprobit, [SVY] svy estimation
selection-order statistics, [TS] varsoc
selection-on-observables, see conditional-independence
  assumption
SEM, see structural equation modeling
sem command, [SEM] Builder, [SEM] example 1,
  [SEM] example 3, [SEM] example 6,
  [SEM] example 7, [SEM] example 8,
  [SEM] example 9, [SEM] example 10,
  [SEM] example 12, [SEM] example 15,
  [SEM] example 16, [SEM] example 17,
  [SEM] example 18, [SEM] example 20,
  [SEM] example 23, [SEM] example 24,
  [SEM] example 26, [SEM] example 42g,
  [SEM] methods and formulas for sem,
  [SEM] sem, [SEM] sem and gsem path
  notation, [SEM] sem model description
  options, [SEM] sem path notation extensions,
  [SEM] sem postestimation, [SEM] Glossary
missing values, [SEM] example 26
with constraints, [SEM] example 8
sem option
  allmissing, [SEM] sem estimation options
  baseopts(), [SEM] sem estimation options
  coeflegend, [SEM] example 8,
  [SEM] example 16, [SEM] sem reporting options
  constraints(), [SEM] sem and gsem option constraints(), [SEM] sem model description options
  covariance(), [SEM] sem and gsem path notation, [SEM] sem model description options,
  [SEM] path notation extensions
  covstructure(), [SEM] intro 5,
  [SEM] example 17, [SEM] sem and gsem option covstructure(), [SEM] sem model description options
  forcecorrelations, [SEM] sem ssd options
  forcenotanchor, [SEM] sem model description options
  forceconditional, [SEM] sem option noxconditional
  from(), [SEM] intro 12, [SEM] sem and gsem option from()(), [SEM] sem model description options
  fvwrap(), [SEM] sem reporting options
  fvwrapon(), [SEM] sem reporting options
  ginvariant(), [SEM] intro 6, [SEM] example 23,
  [SEM] sem group options
  group(), [SEM] intro 6, [SEM] example 20,
  [SEM] example 23, [SEM] sem group options,
  [SEM] sem option select()(), [SEM] sem path notation extensions
  latent(), [SEM] sem and gsem syntax options
  level(), [SEM] sem reporting options
  maximize_options(), [SEM] intro 12, [SEM] sem estimation options
  means(), [SEM] intro 5, [SEM] example 18,
  [SEM] sem and gsem path notation, [SEM] sem model description options,
  [SEM] sem path notation extensions
  method(), [SEM] intro 4, [SEM] intro 8,
  [SEM] intro 9, [SEM] example 26, [SEM] sem estimation options,
  [SEM] sem option method()(), [SEM] Glossary
  mmi(), [SEM] sem estimation options
  noanchor, [SEM] sem model description options
  nocapslatent, [SEM] sem and gsem syntax options
  nocnomreport, [SEM] sem reporting options
  noconstant, [SEM] sem model description options
  nodescribe, [SEM] sem reporting options
  noestimate, [SEM] sem estimation options
  nofootnote, [SEM] sem reporting options
  nofvlabel, [SEM] sem reporting options
  noheader, [SEM] sem reporting options
  noistart, [SEM] sem estimation options
  nomain, [SEM] sem model description options
  notable, [SEM] sem reporting options

sem option, continued
  noxconditional, [SEM] sem estimation options,
  [SEM] sem option noxconditional
  reliability(), [SEM] intro 12,
  [SEM] example 24, [SEM] sem and gsem option reliability(), [SEM] sem model description options
  satopts(), [SEM] sem estimation options
  select(), [SEM] sem option select()(), [SEM] sem ssd options
  showginvariant, [SEM] sem reporting options
  standardized, [SEM] sem reporting options
  variance(), [SEM] sem and gsem path notation, [SEM] sem model description options,
  [SEM] sem path notation extensions
  vce(), [SEM] intro 4, [SEM] intro 8,
  [SEM] intro 9, [SEM] sem estimation options,
  [SEM] sem option method()(), [SEM] Glossary

sem postestimation commands, [SEM] intro 7

sem colons, [M-2] semicolons

semiconjugacy, see semiconguate prior

semiconjugate prior, [BAYES] intro,
  [BAYES] bayesmh, [BAYES] Glossary

semiparametric imputation method, see imputation,
  predictive mean matching

semiparametric model, [ST] stcox, [ST] stcrreg,
  [ST] Glossary

semirobust standard errors, [XT] Glossary

sensitivity, [R] estat classification, [R] lroc, [R] lsens,
  also see receiver operating characteristic analysis,
  see receiver operating characteristic analysis

analysis, [PSS] power, [PSS] power graph,
  [PSS] power table, [PSS] power onemean,
  [PSS] power twomeans, [PSS] power pairedmeans, [PSS] power oneproportion,
  [PSS] power twopropportions, [PSS] power oneproportion,
  [PSS] power twovariances, [PSS] power onevariance,
  [PSS] power twocorrelations, [PSS] power onevariance,
  [PSS] power oneway, [PSS] power toway,
  [PSS] power repeated, [PSS] power cmh,
  [PSS] power mcc, [PSS] power trend,
  [PSS] power cox, [PSS] power exponential,
  [PSS] power logrank, [PSS] Glossary, also see
  Bayesian, sensitivity analysis

model, [R] regress postestimation, [R] rreg

separate command, [D] separate

separating string variables into parts, [D] split

seq(), egen function, [D] egen

sequential imputation, [MI] mi impute, [MI] mi impute chained, [MI] mi impute monotone

sequential limit theory, [XT] Glossary

sequential regression multivariate imputation, see
imputation, multivariate, chained equations

serial correlation, see autocorrelation

test, [TS] Glossary

serial independence test, [R] runtest

serrbar command, [R] serrbar
serset, [P] serset
  clear command, [P] serset
  create command, [P] serset
  create_cspline command, [P] serset
  create_xmedians command, [P] serset
dir command, [P] serset
drop command, [P] serset
reset_id command, [P] serset
set command, [P] serset
sort command, [P] serset
summarize command, [P] serset
use command, [P] serset

sersetread, file subcommand, [P] serset
sersetwrite, file subcommand, [P] serset

session, recording, [R] log, [U] 15 Saving and printing output—log files

set
  adosize command, [P] sysdir, [R] set,
    [U] 18.11 A.do-files
  autotabgraphs command, [R] set
  cformat command, [R] set, [R] set cformat
  charset command, [P] smcl
  checksum command, [D] checksum, [R] set
cleveld command, [BAYES] set clevel, [R] set
  coeftabresults command, [R] set
  command, [R] query, [R] set
conren command, [R] set
dockable command, [R] set
dockingguides command, [R] set
doublebuffer command, [R] set
dp command, [D] format, [R] set
emptycells command, [R] set, [R] set emptycells
eolchar command, [R] set
fastscroll command, [R] set
floatwindows command, [R] set
fvlabel command, [R] set, [R] set showbaselevels
fwrap command, [R] set, [R] set showbaselevels
fwrapon command, [R] set, [R] set showbaselevels
graphics command, [G-2] set graphics, [R] set
haverdir command, [D] import haver
haverdir command, [R] set
httpproxy command, [R] netio, [R] set
httpproxysitecommand, [R] netio, [R] set
httpproxyhost command, [R] netio, [R] set
httpproxyport command, [R] netio, [R] set
httpproxysitepw command, [R] netio, [R] set
httpproxyuser command, [R] netio, [R] set
include_bitmap command, [R] set
level command, [R] level, [R] set
linegap command, [R] set
linesize command, [R] log, [R] set
locale_functions command, [P] set
locale_functions, [R] set
locale_ui command, [P] set locale_ui, [R] set
locksplitters command, [R] set

set, continued
  logtype command, [R] log, [R] set
  lstretch command, [R] set
  matsize command, [R] matsize, [R] set
  maxdb command, [R] db, [R] set
  maxiter command, [R] maximize, [R] set
  max_memory command, [D] memory, [R] set
  maxvar command, [D] memory, [R] set
  min_memory command, [D] memory, [R] set
  more command, [P] more, [R] more, [R] set
  niceness command, [D] memory, [R] set
  notifyuser command, [R] set
  obs command, [D] obs, [R] set
  odbcdriver command, [D] odbc, [R] set
  odbcmgr command, [D] odbc, [R] set
  output command, [P] quietly, [R] set
  pagesize command, [R] more, [R] set
  pformat command, [R] set, [R] set cformat
  pinnable command, [R] set
  playsnd command, [R] set
  revkeyboard command, [R] set
  rmmsg command, [P] rmmsg, [R] set
  rng command, [R] set, [R] set rng
  rngstate command, [R] set, [R] set seed
  scrollbufsize command, [R] set
  searchdefault command, [R] search, [R] set
  seed command, [R] set, [R] set seed
  segmentsize command, [D] memory, [R] set
  sformat command, [R] set, [R] set cformat
  showbaselevels command, [R] set, [R] set showbaselevels
  showemptycells command, [R] set, [R] set showbaselevels
  showomitted command, [R] set, [R] set showbaselevels
  smoothfonts command, [R] set
  timeout1 command, [R] netio, [R] set
  timeout2 command, [R] netio, [R] set
  trace command, [P] trace, [R] set
  tracedepth command, [P] trace, [R] set
  traceexpand command, [P] trace, [R] set
  tracehltile command, [P] trace, [R] set
  traceindent command, [P] trace, [R] set
  tracenumbers command, [P] trace, [R] set
  tracesep command, [P] trace, [R] set
type command, [D] generate, [R] set
  update_interval command, [R] set, [R] set
  update update_promp command, [R] set, [R] set
  update update_query command, [R] set, [R] set
  varabbrev command, [R] set
  varkeyboard command, [R] set
set,  
class subcommand, [MV] cluster programming utilities  
datasignature subcommand, [D] datasignature  
file subcommand, [P] file  
graph subcommand, [G-2] graph set  
irf subcommand, [TS] irf set  
m subcommand, [MI] mi set  
pseudocel subcommand, [P] pseudocel, [P] pseudocel advanced  
serset subcommand, [P] serset  
ssd subcommand, [SEM] ssd  
sysdir subcommand, [P] sysdir  
translate subcommand, [R] translate  
webuse subcommand, [D] webuse  
set ado, net subcommand, [R] net  
set matacache, mata subcommand, [M-3] mata set, [R] set  
set matalib, mata subcommand, [M-3] mata set, [R] set  
set matalnum, mata subcommand, [M-3] mata set, [R] set  
set matamofirst, mata subcommand, [M-3] mata set, [R] set  
set mataoptimiz, mata subcommand, [M-3] mata set, [R] set  
set other net subcommand, [R] net  
seset breakintri() function, [M-5] seset breakintri()  
seset defaults command, [R] set set defaults  
seset() function, [M-5] more()  
seset moreonexit() function, [M-5] more()  
setting M, [MI] mi add, [MI] mi set  
setting mi data, [MI] mi set  
settings,  
display, [R] set showbaselevels  
efficiency, [P] creturn  
format, [R] set cformat  
graphics, [P] creturn  
interface, [P] creturn  
memory, [P] creturn  
network, [P] creturn  
output, [P] creturn  
program debugging, [P] creturn  
random-number generator, [R] set rng  
trace, [P] creturn  
sformat, set subcommand, [R] set, [R] set cformat  
sfrancia command, [R] swilk  
shade style, [G-4] shade style  
shading region, [G-3] region_options  
Shapiro–Francia test for normality, [R] swilk  
Shapiro–Wilk test for normality, [R] swilk  
shared object, [P] class, [P] plugin  
sheen command, [D] sheen  
Shepard diagram, [MV] mds postestimation plots, [MV] Glossary  
plot, [MV] mds postestimation plots  
showharten command, [G-3] qc  
shift macro subcommand, [P] macro  
showbaselevels, set subcommand, [R] set, [R] set showbaselevels  
showemptycells, set subcommand, [R] set, [R] set showbaselevels  
showinvariant option, see sem option showinvariant  
showrtolerance option, [R] maximize  
showomitted, set subcommand, [R] set, [R] set showbaselevels  
showstep option, [R] maximize  
showtolerance option, [R] maximize  
SHR, see subhazard ratio  
shwinters, tssmooth subcommand, [TS] tssmooth  
shwinters  
Šídák’s multiple-comparison adjustment, see multiple comparisons, Šídák’s method  
sign test, [PSS] power one proportion, [PSS] Glossary  
signestimationsample command, [P] signestimationsample  
observed, see p-value  
signing digitally data, see datasignature command  
signrank command, [R] signrank  
sign test command, [R] sign test  
sign function, [FN] Mathematical functions, [FN] Trigonometric functions  
similarity, [MV] Glossary  
matrices, [MV] matrix dissimilarity, [P] matrix dissimilarity
similarity, continued
measures, [MV] cluster, [MV] cluster
programming utilities, [MV] matrix
dissimilarity, [MV] measure_option, [P] matrix
dissimilarity
Anderberg coefficient, [MV] measure_option
angular, [MV] measure_option
correlation, [MV] measure_option
Dice coefficient, [MV] measure_option
Gower coefficient, [MV] measure_option
Hamann coefficient, [MV] measure_option
Jaccard coefficient, [MV] measure_option
Kulczyński coefficient, [MV] measure_option
matching coefficient, [MV] measure_option
Ochiai coefficient, [MV] measure_option
Pearson coefficient, [MV] measure_option
Rogers and Tanimoto coefficient,
[MV] measure_option
Russell and Rao coefficient,
[MV] measure_option
Sneath and Sokal coefficient,
[MV] measure_option
Yule coefficient, [MV] measure_option
simple asymmetric autoregressive conditional
heteroskedasticity, [TS] arch
simple random sample, [SVY] Glossary
Simpson’s rule, [PSS] power logrank
simulate prefix command, [R] simulate
simulation, [TS] forecast, [TS] forecast adjust,
[TS] forecast clear, [TS] forecast coeffvector,
[TS] forecast create, [TS] forecast describe,
[TS] forecast drop, [TS] forecast estimates,
[TS] forecast exogenous, [TS] forecast
identity, [TS] forecast list, [TS] forecast query,
[TS] forecast solve, [U] 20.20 Dynamic forecasts
and simulations
Markov chain Monte Carlo, [BAYES] intro,
[BAYES] bayes, [BAYES] bayesmh,
[BAYES] bayesmh evaluators
Monte Carlo, [P] postfile, [R] permute,
[R] simulate
simultaneous
quantile regression, [R] qreg
systems, [R] reg3
sin() function, [FN] Trigonometric functions,
[M-5] sin()
sine function, [FN] Trigonometric functions
single-failure st data, see survival analysis
single-imputation methods, [MI] intro substantive
single linkage,
clustermat subcommand, [MV] cluster linkage
cluster subcommand, [MV] cluster linkage
single-linkage clustering, [MV] cluster,
[MV] clustermat, [MV] cluster linkage,
[MV] Glossary
single-precision floating point number,
[U] 12.2.2 Numeric storage types
single-record st data, see survival analysis
singleton strata, [SVY] estat, [SVY] variance
estimation
singular value decomposition, [M-5] svd(),
sinh() function, [FN] Trigonometric functions,
[M-5] sin()
smoothing, [G-2] graph twoway lpoly, [R] lpoly,
[R] smooth
graphs, [R] kdensity, [R] lowess
smoothing graphs, [G-2] graph other
SMR, see standardized mortality ratio
snapshot, [D] snapshot

snapshot
erase command, [D] snapshot
label command, [D] snapshot
list command, [D] snapshot
restore command, [D] snapshot
save command, [D] snapshot
snapshot data, [ST] snappy, [ST] stset,
[ST] Glossary
snappy command, [ST] snappy
Spearman’s rho, [R] spearman
Spearman–Brown prophecy formula, [MV] alpha
spearman command, [R] spearman
Spearman’s r, [R] spearman
specialized graphs, [G-2] graph other
specification test, [R] gmm postestimation,
[R] hausman, [R] ivpoisson postestimation,
[R] ivregress postestimation, [R] linktest,
[R] lnkew0, [R] regress postestimation,
tests, [ST] stcox postestimation, [ST] stsplit,
[XT] xtreg postestimation
specificity, [MV] factor, [R] estat classification,
[R] Iroc, [R] Isens, also see receiver operating
classification analysis, see receiver operating
classification analysis
spectral
analysis, [TS] Glossary
density, [TS] psdensity, [TS] Glossary
distribution, [TS] cumsp, [TS] pergram,
[TS] psdensity, [TS] Glossary
plots, cumulative, [G-2] graph other
spectrum, [TS] psdensity, [TS] Glossary
spell data, [ST] Glossary
spherical covariance, [MV] mvtest covariances
sphericity, [MV] Glossary
assumption, [PSS] power repeated, [PSS] Glossary
Spiegelhalter’s Z statistic, [R] brier
spike, graph twoway subcommand, [G-2] graph
twoway spike
spike plot, [R] spikeplot
spikeplot command, [R] spikeplot
spline3() function, [M-5] spline3()
spline3eval() function, [M-5] spline3()
splines
linear, [R] mkspline
restricted cubic, [R] mkspline
split command, [D] split
split-plot designs, [MV] manova, [R] anova
splitting time-span records, [ST] stsplit
S-Plus, reading data from, [U] 21.4 Transfer programs
spread, [R] IV

spreadsheets, transferring
from Stata, [D] edit, [D] export, [D] import
delimited, [D] import excel, [D] import
haver, [D] odbc, [D] outfile, [D] xmlsave,
[U] 21.4 Transfer programs
into Stata, [D] edit, [D] import, [D] import
delimited, [D] import excel, [D] import
haver, [D] infil (fixed format), [D] infil (free format),
[D] odbc, [D] xmlsave, [U] 21.4 Entering and
importing data, [U] 21.4 Transfer programs

sprintf() function, [M-5] printf()
SPSS dates, [D] datetime
SPSS, reading data from, [U] 21.4 Transfer programs
SQL, [D] odbc
sqlfile(), odbc subcommand, [D] odbc
sqreg command, [R] qreg, [R] qreg postestimation
sqrl() function, [R] qreg
sqrt() function, [FN] Mathematical functions,
[M-5] sqrt()
square
matrix, [M-6] Glossary
root function, [FN] Mathematical functions
squared multiple correlation, [SEM] methods and formulas for sem
squared multiple correlations, [MV] factor postestimation
sreturn
clear command, [P] return
list command, [P] return, [R] stored results
local command, [P] return
SRMI, see imputation, multivariate, chained equations
SRMR, see standardized, root mean squared residual
SRS, see simple random sample
ss() function, [D] datetime, [FN] Date and time functions, [M-5] date()
ssc
copy command, [R] ssc
descr copy command, [R] ssc
describe command, [R] ssc
hot command, [R] ssc
install command, [R] ssc
new command, [R] ssc
type command, [R] ssc
uninstall command, [R] ssc
SSC archive, see Statistical Software Components archive
ssC() function, [D] datetime, [FN] Date and time functions, [M-5] date()
SSCP matrix, [MV] Glossary
SSD, see summary statistics data
ssd
addgroup command, [SEM] ssd
build command, [SEM] ssd
describe command, [SEM] ssd
init command, [SEM] ssd
list command, [SEM] ssd
repair command, [SEM] ssd
set command, [SEM] ssd
status command, [SEM] ssd
unaddgroup command, [SEM] ssd
sspace command, [TS] sspace, [TS] sspace postestimation
SSU, see secondary sampling unit
_st_addobs() function, [M-5] st_addobs()
_st_addobs() function, [M-5] st_addobs()
_st_addvar() function, [M-5] st_addvar()
_st_addvar() function, [M-5] st_addvar()
_st_addvar() function, [M-5] st_addvar()
_st_addvar() function, [M-5] st_addvar()
st command, [ST] stset
st commands for mi data, [MI] mi stsplit, [MI] mi XXXset
_st_ct, [ST] st_is
_st_data() function, [M-5] st_data()
st_data() function, [M-5] st_data()
_st_dir() function, [M-5] st_dir()
_st_dropobsif() function, [M-5] st_dropvar()
st_varvalue() function, [M-5] st_varformat()  
st_view() function, [M-5] st_view()  
st_viewobs() function, [M-5] st_viewvars()  
st_viewvars() function, [M-5] st_viewvars()  
st_vldrop() function, [M-5] st_vlexists()  
st_vlexists() function, [M-5] st_vlexists()  
st_vload() function, [M-5] st_vlexists()  
st_vlmapping() function, [M-5] st_vlexists()  
st_vlmmap() function, [M-5] st_vlexists()  
st_vlmmap() function, [M-5] st_vlexists()  
st_vlmsearch() function, [M-5] st_vlexists()  

stability, [TS] var intro, [TS] var [TS] var svar,  
[TS] vecstable  
  after ARIMA, [TS] estat aroots  
  after VAR or SVAR, [TS] varstable  
  after VEC, [TS] vec intro, [TS] vec  
  of nonrecursive models, see nonrecursive model,  
  stability of  

stable, estat subcommand, [SEM] estat stable  

stable unit treatment value assumption, [TE] teffects intro advanced  

stack command, [D] stack  
stacked variables, [MV] Glossary  
stacking data, [D] stack  
stacking variables, [MV] Glossary  

stair step, connecting points with, [G-4] connect style  

standard deviations, [PSS] power, [PSS] power onevariance  
  confidence intervals for, [R] ci  
  control-group, [PSS] power twovariances  
  creating  
  dataset of, [D] collapse  
  variable containing, [D] egen  
  displaying, [R] lv, [R] summarize, [R] table,  
  [R] tabstat, [R] tabulate, summarize(), [XT] xtsum  

experimental-group, [PSS] power twovariances  
  independent, see standard deviations, two-sample  
  one-sample, [PSS] power onevariance  

subpopulations, see subpopulation, standard  
  deviations of  
  testing equality of, [R] sdtest  
  two-sample, [PSS] power twovariances  

standard error  
  robust, see robust, Huber/White/sandwich estimator  
  of variance  

standard error bar charts, [G-2] graph other  

standard errors, see gsem option vce(), see sem option vce()  
  accessing, [P] matrix get, [U] 13.5 Accessing coefficients and standard errors  
  balanced repeated replication, see balanced repeated replication standard errors  
  bootstrap, see bootstrap standard errors  
  for general predictions, [R] predictnl  
  forecast, [R] predict, [R] regres postestimation  
  jackknife, see jackknife standard errors  
  MCMC, see MCSE  
  mean, [R] ci, [R] mean  

standard errors, continued  
  panel-corrected, see panel-corrected standard error  
  prediction, [R] glm, [R] predict, [R] regress postestimation  
  residuals, [R] predict, [R] regress postestimation  
  robust, see robust, Abadie–Imbens standard errors,  
  see robust, Huber/White/sandwich estimator of variance  
  semirobust, see semirobust standard errors  
  successive difference replication, see successive difference replication  

standard linear SEM, [SEM] Glossary  

standard strata, see direct standardization  
standard weights, see direct standardization  
standardized  
  coefficients, [SEM] example 3, [SEM] example 6,  
  [SEM] Glossary, also see standardized parameters  
  covariance, [SEM] Glossary  
  covariance residual, [SEM] methods and formulas for sem  
  data, [MV] Glossary  
  difference, [PSS] power, [PSS] power one mean,  
  [PSS] power twomeans, [PSS] power paired means  
  incidence ratio, [R] dstdize  
  margins, [R] margins  
  mean residual, [SEM] methods and formulas for sem  
  means, [R] mean  
  mortality ratio, [R] dstdize, [R] epitab,  
  option, [SEM] example 11  
  parameters, [SEM] estat stdize, [SEM] methods and formulas for sem  
  proportions, [R] proportion  
  rates, [R] dstdize, [R] epitab  
  ratios, [R] ratio  
  residuals, [R] binreg postestimation, [R] glm postestimation,  
  [R] logistic postestimation, [R] predict, [R] regress postestimation, [SEM] estat residuals,  
  [SEM] methods and formulas for sem, [SEM] Glossary  
  root mean squared residual, [SEM] estat ggof,  
  [SEM] estat gogf, [SEM] example 4,  
  [SEM] example 21, [SEM] methods and formulas for sem  

standardized option, see sem option standardized  
standardized, variables, [D] egen  

start() option, [G-2] graph twoway histogram  

startgrid() option, see gsem option startgrid()  

starting values, [SEM] intro 12, [SEM] sem and gsem  
  option from(), [SEM] sem and gsem path notation, [SEM] sem path notation extensions,  
  [SEM] Glossary  

startvalues() option, see gsem option startvalues()  

Stat/Transfer, [U] 21.4 Transfer programs
Stata

Blog, [U] 3.2.3 The Stata Blog—Not Elsewhere Classified
c-class results, [M-5] st_global(),
conference, [U] 3.6.1 Conferences and users group meetings
description, [U] 2 A brief description of Stata
documentation, [U] 1 Read this—it will help
error message, [M-5] error()
example datasets, [U] 1.2.2 Example datasets
execute command, [M-3] mata stata execute command,
for Unix, see Unix
for Windows, see Windows
forum, [U] 3.2.4 The Stata forum
internal form, [D] datetime, [D] datetime display formats, [D] datetime translation
limits, [R] limits, [U] 5 Flavors of Stata
logo, [G-2] graph print
[P] smcl
NetCourseNow, [U] 3.6.2 NetCourses
NetCourses, [U] 3.6.2 NetCourses
on Facebook, [U] 3.2.5 Stata on social media
on Google+, [U] 3.2.5 Stata on social media
on LinkedIn, [U] 3.2.5 Stata on social media
on Twitter, [U] 3.2.5 Stata on social media
op . varname, see Stata, time-series–operated variable platforms, [U] 5.1 Platforms
Press, [U] 3.3 Stata Press
Small, see Small Stata
Stata/IC, see Stata/IC
Stata/MP, see Stata/MP
Stata/SE, see Stata/SE
supplementary material, [U] 3 Resources for learning and using Stata
support, [U] 3 Resources for learning and using Stata
temporary filenames, [M-5] st_tempname()
names, [M-5] st_tempname()
training, [U] 3.6 Conferences and training

Stata, continued

updates, see updates to Stata
users group meeting, [U] 3.6.1 Conferences and users group meetings
variable formats, [M-5] st_varformat(),
labels, [M-5] st_varformat()
website, [U] 3.2.1 The Stata website (www.stata.com)
YouTube Channel, [U] 3.2.2 The Stata YouTube Channel

Stata,
data file format, technical description, [P] file formats .dta
exiting, see exit command
for Mac, see Mac
for Unix, see Unix
for Windows, see Windows
install an addition?, see Stata, time-series–operated variable
installation of, [R] net, [R] sj, [U] 17.6 How do I install an addition?
keyword search of, [R] search, [U] 4 Stata’s help and search facilities
limits, see limits
Stata News, [U] 3 Resources for learning and using Stata
Stata Technical Bulletin Reprints, [U] 3.4 The Stata Journal
Stata/IC, [R] limits, [U] 5 Flavors of Stata
Stata/MP, [R] limits, [U] 5 Flavors of Stata
Stata/SE, [R] limits, [U] 5 Flavors of Stata
Stata/MP, see Stata/MP
Stata/SE, see Stata/SE
Supplementary Material, see Supplementary Material
Supplemental Material, see Supplementary Material
Supplementary Material, see Supplementary Material
Supplementary Material, see Supplementary Material
Supplemental Material, see Supplementary Material
Supplemental Material, see Supplementary Material
supplementary material, see Stata, time-series–operated variable platforms, [U] 5.1 Platforms
SuperNetCourses, [U] 3.6.2 NetCourses
update, [U] 3.6.1 Conferences and users group meetings
updates, see updates to Stata
updates, see updates to Stata
users group meeting, [U] 3.6.1 Conferences and users group meetings
variable formats, [M-5] st_varformat(),
labels, [M-5] st_varformat()
statistical
density functions, [M-5] normal()
distribution functions, [M-5] normal()
inference, hypothesis testing, see hypothesis test
Statistical Software Components archive, [R] ssc
stats, estimates subcommand, [R] estimates stats
statsby prefix command, [D] statsby
status, sst subcommand, [SEM] sst
STB, see Stata Journal and Stata Technical Bulletin
stb, net subcommand, [R] net
stbase command, [ST] stbase
stci command, [ST] stci
stcox, fractional polynomials, [R] fp, [R] mfp
stcoxkm command, [ST] stcox PH-assumption tests
stcurve command, [ST] stcurve
std(), egen function, [D] egen
stdescribe command, [ST] stdescribe
stdize, estat subcommand, [SEM] estat stdize
steady-state equilibrium, [TS] Glossary
steepest descent (ascent), [M-5] optimize()

stem() command, [R] stem
stem-and-leaf displays, [R] stem
stepwise estimation, [R] stepwise
stepwise prefix command, [R] stepwise
stereotype logistic regression, [R] slogit, [SVY] svy estimation
stfill command, [ST] stfill
stgen command, [ST] stgen
.sthip file, [U] 4 Stata’s help and search facilities, [U] 11.6 Filenaming conventions, [U] 18.11.6 Writing system help
stir command, [ST] stir
stjoin command, [ST] stsplit
stjoin for mi data, [MI] mi stsplit
stjoin, mi subcommand, [MI] mi stsplit
stmc command, [ST] strate
stmh command, [ST] strate
stochastic
equation, [TS] Glossary
stop.
clustermat subcommand, [MV] cluster stop
cluster subcommand, [MV] cluster stop
stopbox, window subcommand, [P] window programming, [P] window stopbox
stopping command execution, [U] 10 Keyboard use
stopping rules, [MV] Glossary
adding, [MV] cluster programming subroutines
Caliński and Harabasz index, [MV] cluster,
[MV] cluster stop
Duda and Hart index, [MV] cluster, [MV] cluster stop
stepsize, [MV] cluster programming subroutines
storage types, [D] codebook, [D] compress,
[D] describe, [D] encode, [D] format,
[D] generate, [D] recast, [D] varmanage,
[U] 11.4 varlists, [U] 12.2.2 Numeric storage types, [U] 12.4 Strings
store, estimates subcommand, [R] stores
store estimation results, [P] ereturn
stored results, [P] _return, [P] return, [R] stored results, [SEM] intro 7, [U] 18.8 Accessing results calculated by other programs,
hidden or historical, [M-5] st_global(),
storing and restoring estimation results, [R] estimates store
stphplot command, [ST] stcox PH-assumption tests
.stpr file, [U] 11.6 Filenaming conventions
stptime command, [ST] stptime
.stptrace file, [U] 11.6 Filenaming conventions
str#, [D] data types, [U] 12.4 Strings
strata, estat subcommand, [SVY] estat
strata with one sampling unit, [SVY] variance estimation
strate command, [ST] strate
stratified
2×2 table, [PSS] power, [PSS] power cmh,
[PSS] Glossary
analysis, [PSS] power, [PSS] power cmh
graphs, [R] dotplot
models, [R] asclogit, [R] asmprobit, [R] asprobit,
[R] clogit, [R] exlogistic, [R] expoisson,
[R] rocreg, [R] rologit
resampling, [R] bootstrap, [R] bsample, [R] bstat,
[R] permute
standardization, [R] stdize
summary statistics, [R] mean, [R] proportion, [R] ratio, [R] total
stratified, continued

tables, [R] epitab
test, [R] epitab, [ST] stcox PH-assumption tests,
[ST] sttest, [ST] Glossary
stratum collapse, [SVY] svydes
strcat() function, [FN] String functions
strdup() function, [M-5] strdup()
strdup() function, [FN] String functions
stream I/O versus record I/O, [U]
stream I/O versus record I/O, [U] 21 Entering and importing data
streset command, [ST] stset
streset command for mi data, [MI] mi XXXset
streset, mi subcommand, [MI] mi XXXset
stress, estat subcommand, [MV] mds postestimation
stress, estat subcommand, [MV] mds postestimation
strict stationarity, [TS] Glossary
string, also see Unicode strings
concatenation, [M-5] invtokens()
duplication, [M-5] strdup()
functions, [FN] String functions, [M-4] string,
[U] 12.4 Strings, [U] 12.4.2.1 Unicode string functions,
[U] 23 Working with strings
pattern matching, [M-5] strmatch()
to real, convert, [M-5] strtoreal()
variables, [D] data types, [D] infile (free format),
[U] 12.4 Strings, [U] 23 Working with strings
converting to numbers, [FN] String functions
encoding, [D] encode
exporting, [D] export
formatting, [D] format
importing, [D] import
inputting, [D] edit, [D] input, [U] 21 Entering and importing data
long, [U] 12.4.13 How to see the full contents of a strL or a str# variable, also see strL
making from value labels, [D] encode
mapping to numbers, [D] destring, [D] encode,
[D] label, also see real() function
parsing, [P] gettoken, [P] tokenize
sort order, [U] 13.2.3 Relational operators
splitting into parts, [D] split
string() function, [FN] String functions
strings, [D] unicode
strtrim() function, [FN] String functions,
[M-5] strtrim()
strL, [D] data types, [U] 12.4 Strings
displaying, [U] 12.4.13 How to see the full contents of a strL or a str# variable
strlen macro extended function, [P] macro
strlen() function, [FN] String functions,
[M-5] strlen()
strlower() function, [FN] String functions,
[M-5] strupper()
strltrim() function, [FN] String functions,
[M-5] strtrim()
strmatch() function, [FN] String functions,
[M-5] strmatch()
strvreal() function, [FN] String functions,
[M-5] strvreal()
strongly balanced, [XT] Glossary
strpos() function, [FN] String functions,
[M-5] strpos()
strproper() function, [FN] String functions,
[M-5] strupper()
strreverse() function, [FN] String functions,
[M-5] strreverse()
strrpos() function, [FN] String functions,
[M-5] strrpos()
strtrim() function, [FN] String functions,
[M-5] strtrim()
struct, [M-2] struct
structural break,
known break date, [TS] estat sbknown
unknown break date, [TS] estat sbsingle
structural equation modeling, [MV] intro,
[SEM] methods and formulas for gsem,
[SEM] methods and formulas for sem,
[SEM] Glossary, [SVY] syv estimation,
[U] 26.4 Structural equation modeling (SEM)
structural model, [SEM] intro 5, [SEM] example 7,
[SEM] example 9, [SEM] example 32g,
structural time-series model, [TS] psdensity,
postestimation, [R] regres postestimation time series, [TS] feast compute, [TS] feast graph,
[TS] varnorm, [TS] varsc, [TS] varstable,
[TS] vars, [TS] varv
structure, [MV] Glossary
structure, estat subcommand, [MV] discrim lda
postestimation, [MV] factor postestimation
structured, [SEM] Glossary
structured (correlation or covariance), [SEM] Glossary
[M-6] Glossary
strupper() function, [FN] String functions,
[M-5] strupper()
sts generate command, [ST] sts, [ST] sts generate
sts graph command, [ST] sts, [ST] sts graph
sts list command, [ST] sts, [ST] sts list
sts test command, [ST] sts, [ST] sts test
.stsem file, [U] 11.6 Filenaming conventions
stset command, [ST] stset
stset command for mi data, [MI] mi XXXset
stset, mi subcommand, [MI] mi XXXset
stsplit command, [ST] stsplit
stsplit for mi data, [MI] mi stsplit
stsplit, mi subcommand, [MI] mi stsplit
stsum command, [ST] stsum
stteffects, [TE] stteffects, [TE] stteffects
postestimation
command, [TE] tebalance
ipw command, [TE] stteffects ipw
ipwra command, [TE] stteffects ipwra
ra command, [TE] stteffects ra
wra command, [TE] stteffects wra
sttoct command, [ST] sttoct
sttoct command, [ST] sttoct
Stuart–Maxwell test statistic, [R] symmetry
Studentized residuals, [R] predict, [R] regress
Studentized-range multiple-comparison adjustment, see multiple comparisons, Tukey’s method
Student–Newman–Keuls’s multiple-comparison adjustment, see multiple comparisons, Student–Newman–Keuls’s method
Student’s $t$
density, central, [FN] Statistical functions
noncentral, [FN] Statistical functions
distribution, also see $t$ distribution
cumulative, [FN] Statistical functions
cumulative noncentral, [FN] Statistical functions
inverse cumulative, [FN] Statistical functions
inverse cumulative noncentral, [FN] Statistical functions
inverse reverse cumulative, [FN] Statistical functions
reverse cumulative, [FN] Statistical functions
study,
   case–control, see case–control study
cohort, see cohort study
controlled clinical trial, see controlled clinical trial study
cross-sectional, see cross-sectional study
experimental, see experimental study
follow-up, see cohort study
matched, see matched study
multiple-sample, see multiple-sample study
observational, see observational study
one-sample, see one-sample study
paired, see paired study
prospective, see prospective study
randomized controlled trial, see randomized controlled trial study
retrospective, see retrospective study
two-sample, see two-sample study
stvary command, [ST] stvary
style
   added line, [G-4] addlinedestyle
   area, [G-4] areastyle
   axis, [G-4] tickstyle
   by-graphs, [G-4] bystyle
clock position, [G-4] clockposstyle
color, [G-4] colorstyle
compass direction, [G-4] compassdirstyle
connect points, [G-4] connectstyle
grid lines, [G-4] gridstyle
legends, [G-4] legendstyle
   [G-4] linewidthstyle
lists, [G-4] stylelists
margins, [G-4] marginstyle
mark labels, [G-4] markerlabelstyle,
markers, [G-4] symbolstyle
plot, [G-4] pstyle
text display angle, [G-4] anglestyle
text justification, [G-4] justificationstyle
textboxes, [G-4] orientationstyle,
   [G-4] textboxstyle
vertical alignment of text, [G-4] alignmentstyle
study, [MI] mi convert, [MI] styles, [MI] Glossary
style,
   flong, see flong
   flongsep, see flongsep
   mlong, see mlong
   wide, see wide
stylelist, [G-4] stylelists
subclass, [M-2] class
subdirectories, [U] 11.6 Filenaming conventions
subhazard ratio, [R] eform_option, [R] lincom,
   [ST] Glossary, also see cumulative subhazard function
subinertia, estat subcommand, [MV] mca
postestimation
subinstr macro extended function, [P] macro
subinstr() function, [FN] String functions,
   [M-5] subinstr()
subinword() function, [FN] String functions,
   [M-5] subinword()
subjective prior, see informative prior
   _sublowertriangle() function,
   [M-5] _sublowertriangle()
subowertriangle() function,
   [M-5] subowertriangle()
subpopulation
   differences, [SVY] survey, [SVY] svy
   postestimation
   estimation, [SVY] subpopulation estimation,
   means, [SVY] svy estimation
   proportions, [SVY] svy estimation,
   [SVY] svy: tabulate oneway,
   [SVY] svy: tabulate twoway
subpopulation, continued
subroutines, adding, [MV] cluster programming utilities
subsampling the chain, see thinning
subscripting matrices, [P] matrix define
subscripts in expressions, [U] 13.7 Explicit subscripting
substantive constraints, see constraints
_substr() function, [M-5] _substr()
substr() function, [FN] String functions, [M-5] substr()
substring function, [FN] String functions subtitle() option, [G-3] title_options
subtraction operator, see subtraction
success–failure proportion, [PSS] su
suest command, [R] suuest, [SVY] syy estimation .sum file, [U] 11.6 Filenaming conventions
sum() function, [FN] Mathematical functions, [M-5] sum()
sum of vector, [M-5] runningsum()
summarize,
estat subcommand, [R] estat, [R] estat summarize
misstable subcommand, [R] misstable
serset subcommand, [P] serset
summarize command, [D] format, [R] summarize, [R] tabulate, summarize()
summary, bayesstats subcommand, [BAYES] bayesstats
summary statistics, see Bayesian, summary statistics, see descriptive statistics
summary variables, generating, [MV] cluster generate summative (Likert) scales, [MV] alpha sums,
survival analysis, [ST] survival analysis, [ST] discrete,
[ST] ltable, [ST] snapspan, [ST] st, [ST] st_is,
[ST] stcurve, [ST] stdservice, [ST] stuff,
[ST] sts, [ST] sts generate, [ST] sts list, [ST] sts test,
[ST] stset, [ST] stsplit, [ST] stsum,
[U] 26.22 Survival-time (failure-time) models

competing-risks regression, [ST] stcrreg,
[ST] stcrreg postestimation
count-time data, [ST] ct, [ST] ctset, [ST] cttost,
[ST] sts graph

Cox proportional hazards model, [ST] stcox,
[ST] stcox PH-assumption tests, [ST] stcox postestimation
graphs, [ST] sts graph
interval regression, [R] intreg
logistic regression, [R] logistic
mixed-effects parametric model, [ME] mestreg
parametric survival model, [ST] streg, [ST] streg
postestimation
Poisson regression, [R] poisson
power and sample size, [PSS] power, [PSS] power
cox, [PSS] power exponential, [PSS] power
logrank
random-effects parametric model, [XT] xtmstreg
treatment effects, [TE] stteffects ipw, [TE] stteffects
ipwra, [TE] stteffects ra, [TE] stteffects wra
survival data, [MI] mi estimate, [MI] mi predict,
[TE] stteffects
survival model, see survival analysis
survival outcomes, see outcomes, survival
survival-time data, see survival analysis
survivor function, [G-2] graph other, [ST] sts,
[ST] sts generate, [ST] sts list, [ST] sts test,
graph of, [ST] stcurve, [ST] sts graph

SUTVA, see stable unit treatment value assumption
SVAR, see structural vector autoregressive
svvar command, [TS] var svar, [TS] var svar
postestimation

SVD, see singular value decomposition
_svd() function, [M-5] svd()
svd() function, [M-5] svd()
svd, matrix subcommand, [P] matrix svd
_svdsv() function, [M-5] svd()
svdsv() function, [M-5] svd()
svmat command, [P] matrix mkmat
_svsolve() function, [M-5] svsolve()
svsolve() function, [M-5] svsolve()
svy: biprobit command, [SVY] svy estimation
svy: clogit command, [SVY] svy estimation
svy: cloglog command, [SVY] svy estimation
svy: cnsreg command, [SVY] svy estimation
svy: cpoisson command, [SVY] svy estimation
svy: etpoisson command, [SVY] svy estimation
svy: etregress command, [SVY] svy estimation
svy: glm command, [SVY] svy estimation
svy: gnbreg command, [SVY] svy estimation
svy: gsem command, [SVY] svy estimation
svy: heckman command, [SVY] svy estimation
svy: heckprobit command, [SVY] svy estimation
svy: hetprobit command, [SVY] svy estimation
svy: intreg command, [SVY] svy estimation
svy: ivprobit command, [SVY] svy estimation
svy: ivregress command, [SVY] svy estimation
svy: ivtobit command, [SVY] svy estimation
svy: logistic command, [SVY] svy estimation,
[SVY] svy postestimation
svy: logit command, [SVY] svy estimation
svy: mean command, [SVY] survey, [SVY] estat,
[SVY] poststratification, [SVY] subpopulation
estimation, [SVY] svy, [SVY] svy estimation,
[SVY] svy postestimation, [SVY] sydyscribe,
[SVY] svyset
svy: mecloglog command, [SVY] svy estimation
svy: mglm command, [SVY] svy estimation
svy: menbreg command, [SVY] svy estimation
svy: meoprobit command, [SVY] svy estimation
svy: mepoisson command, [SVY] svy estimation
svy: meprobit command, [SVY] svy estimation
svy: mestreg command, [SVY] svy estimation
svy: mlogit command, [SVY] svy estimation
svy: mlogit command, [SVY] svy estimation
svy: mlogit command, [SVY] svy estimation
svy: mprobit command, [SVY] svy estimation
svy: mreg command, [SVY] svy estimation
svy: nbreg command, [SVY] svy estimation
svy: nl command, [SVY] svy estimation
svy: ologit command, [SVY] svy estimation,
[SVY] svy postestimation
svy: oprobit command, [SVY] svy estimation
svy: poisson command, [SVY] svy estimation
svy: probit command, [SVY] svy estimation
svy: proportion command, [SVY] svy estimation
svy: ratio command, [SVY] direct standardization,
[SVY] svy brr, [SVY] svy estimation,
[SVY] svy: tabulate twoway
svy: regress command, [SVY] survey, [SVY] svy,
[SVY] svy estimation, [SVY] svy estimation
svy: regsize command, [ST] survey, [SVY] svy,
[SVY] estimation, [SVY] svy jackknife,
[SVY] svy postestimation
svy: scobit command, [SVY] svy estimation
svy: sem command, [SVY] svy estimation
svy: sem command, [SVY] svy estimation
svy: sologit command, [SVY] svy estimation
svy: stcox command, [SVY] svy estimation
svy: stlogit command, [SVY] svy estimation
svy: streg command, [SVY] svy estimation
svy: streg command, [SVY] svy estimation
svy: streg command, [SVY] svy estimation
svy: stata command, [SVY] svy estimation
svy: tabulate command, [SVY] svy: tabulate
oneaway, [SVY] svy: tabulate twoway
svy: tnbreg command, [SVY] svy estimation
svy: tobit command, [SVY] svy estimation
svy: total command, [SVY] svy brr, [SVY] svy estimation
svy: tpoisson command, [SVY] svy estimation
svy: truncreg command, [SVY] svy estimation
svy: zinb command, [SVY] svy estimation
svy: zip command, [SVY] svy estimation
svy bootstrap prefix command, [SVY] svy bootstrap
svy brr prefix command, [SVY] svy brr
svy jackknife prefix command, [SVY] svy jackknife
svy prefix command, [SVY] svy
svy sdr prefix command, [SVY] svy sdr
svydesc command, [SVY] svysystem
svyset command, [SVY] svysystem

prefix command, [SVY] svymarkout

svyset command, [SVY] svymarkout

svyset command for mi data, [MI] mi XXXset
svymarkout command, [P] mark
svysystem select*() function, [M-5] symeigensystem()
svysystem() function, [M-5] symeigensystem()
svysystemselect*() function, [M-5] symeigensystemselect*()

svyset(), [SVY] svysystem
set system parameter settings, [R] set
gmm() command, [R] gmm
ivpoisson() command, [R] ivpoisson
ivregress() command, [R] ivregress
nsur() command, [R] nsur
reg3() command, [R] reg3
sureg() command, [R] sureg

26.17 Regression with systems of equations
also see generalized method of moments

cf, [R] ccf

set cformat() command, [R] set
cf() command, [R] cf
cformat() command, [R] cformat
cformat() command, [R] cf
cf() command, [R] cf
cf() command, [R] cf
cf() command, [R] ccf

cf() command, [U] 13.4 System variables _variables

sysuse command, [D] sysuse
dir command, [D] dir

szroeter() command, [R] regress
postestimation
Szoeter’s test for heteroskedasticity, [R] regress
postestimation

T

t() function, [FN] Statistical functions,
[M-5] normal()
t test, [PSS] Glossary
%t formats, [D] format
%t values and formats, [D] datetime
ttitle() option, [G-3] title_options
t2title() option, [G-3] title_options
tab characters, show, [D] type
tab expansion of variable names, [U] 10.6 Tab expansion of variable names
tab1 command, [R] tabulate oneway
tab2 command, [R] tabulate twoway
tabdisp command, [P] tabdisp
tabi command, [R] tabulate twoway
table, estat subcommand, [MV] ca postestimation
estimates subcommand, [R] estimates table
irf subcommand, [TS] irf table
table command, [R] table
table, frequency, see frequency table
table output, [PSS] power, table, [PSS] powercox,
[PSS] powerexponential
tables, [TS] irf table, [TS] irf table
N-way, [P] tabdisp
actuarial, see life tables
coeficient, display in exponentiated form, [R] eform_option
display settings, [R] estimation options, [R] set
showbaselevels
format settings, [R] set cformat
maximum likelihood display options, [R] ml
system parameter settings, [R] set

svy: zinb command, [SVY] svy estimation
svy: zip command, [SVY] svy estimation
svy bootstrap prefix command, [SVY] svy bootstrap
svy brr prefix command, [SVY] svy brr
svy jackknife prefix command, [SVY] svy jackknife
svy prefix command, [SVY] svy
svy sdr prefix command, [SVY] svy sdr
svydesc command, [SVY] svysystem
svyset command, [SVY] svysystem

prefix command, [SVY] svymarkout

svyset command, [SVY] svymarkout

svyset command for mi data, [MI] mi XXXset
svymarkout command, [P] mark
svysystem select*() function, [M-5] symeigensystem()
svysystem() function, [M-5] symeigensystem()
svysystemselect*() function, [M-5] symeigensystemselect*()

svyset(), [SVY] svysystem
set system parameter settings, [R] set
gmm() command, [R] gmm
ivpoisson() command, [R] ivpoisson
ivregress() command, [R] ivregress
nsur() command, [R] nsur
reg3() command, [R] reg3
sureg() command, [R] sureg

26.17 Regression with systems of equations
also see generalized method of moments

cf, [R] ccf

set cformat() command, [R] set
cf() command, [R] cf
cformat() command, [R] cformat
cformat() command, [R] cf
cf() command, [R] cf
cf() command, [R] ccf

cf() command, [U] 13.4 System variables _variables

sysuse command, [D] sysuse
dir command, [D] dir

szroeter() command, [R] regress
postestimation
Szoeter’s test for heteroskedasticity, [R] regress
postestimation

T

t() function, [FN] Statistical functions,
[M-5] normal()
t test, [PSS] Glossary
%t formats, [D] format
%t values and formats, [D] datetime
ttitle() option, [G-3] title_options
t2title() option, [G-3] title_options
tab characters, show, [D] type
tab expansion of variable names, [U] 10.6 Tab expansion of variable names
tab1 command, [R] tabulate oneway
tab2 command, [R] tabulate twoway
tabdisp command, [P] tabdisp
tabi command, [R] tabulate twoway
table, estat subcommand, [MV] ca postestimation
estimates subcommand, [R] estimates table
irf subcommand, [TS] irf table
table command, [R] table
table, frequency, see frequency table
table output, [PSS] power, table, [PSS] powercox,
[PSS] powerexponential
tables, [TS] irf table, [TS] irf table
N-way, [P] tabdisp
actuarial, see life tables
coeficient, display in exponentiated form, [R] eform_option
display settings, [R] estimation options, [R] set
showbaselevels
format settings, [R] set cformat
maximum likelihood display options, [R] ml
system parameter settings, [R] set

svy: zinb command, [SVY] svy estimation
svy: zip command, [SVY] svy estimation
svy bootstrap prefix command, [SVY] svy bootstrap
svy brr prefix command, [SVY] svy brr
svy jackknife prefix command, [SVY] svy jackknife
svy prefix command, [SVY] svy
svy sdr prefix command, [SVY] svy sdr
svydesc command, [SVY] svysystem
svyset command, [SVY] svysystem

prefix command, [SVY] svymarkout

svyset command, [SVY] svymarkout

svyset command for mi data, [MI] mi XXXset
svymarkout command, [P] mark
svysystem select*() function, [M-5] symeigensystem()
svysystem() function, [M-5] symeigensystem()
svysystemselect*() function, [M-5] symeigensystemselect*()

svyset(), [SVY] svysystem
set system parameter settings, [R] set
gmm() command, [R] gmm
ivpoisson() command, [R] ivpoisson
ivregress() command, [R] ivregress
nsur() command, [R] nsur
reg3() command, [R] reg3
sureg() command, [R] sureg

26.17 Regression with systems of equations
also see generalized method of moments

cf, [R] ccf

set cformat() command, [R] set
cf() command, [R] cf
cformat() command, [R] cformat
cformat() command, [R] cf
cf() command, [R] cf
cf() command, [R] ccf

cf() command, [U] 13.4 System variables _variables

sysuse command, [D] sysuse
dir command, [D] dir

szroeter() command, [R] regress
postestimation
Szoeter’s test for heteroskedasticity, [R] regress
postestimation

T

t() function, [FN] Statistical functions,
[M-5] normal()
t test, [PSS] Glossary
%t formats, [D] format
%t values and formats, [D] datetime
ttitle() option, [G-3] title_options
t2title() option, [G-3] title_options
tab characters, show, [D] type
tab expansion of variable names, [U] 10.6 Tab expansion of variable names
tab1 command, [R] tabulate oneway
tab2 command, [R] tabulate twoway
tabdisp command, [P] tabdisp
tabi command, [R] tabulate twoway
table, estat subcommand, [MV] ca postestimation
estimates subcommand, [R] estimates table
irf subcommand, [TS] irf table
table command, [R] table
table, frequency, see frequency table
table output, [PSS] power, table, [PSS] powercox,
[PSS] powerexponential
tables, [TS] irf table, [TS] irf table
N-way, [P] tabdisp
actuarial, see life tables
coeficient, display in exponentiated form, [R] eform_option
display settings, [R] estimation options, [R] set
showbaselevels
format settings, [R] set cformat
maximum likelihood display options, [R] ml
system parameter settings, [R] set
tables, continued
target, continued
contingency, [R] table, [R] tabulate twoway,
epidemiological, see epidemiological tables
estimation results, [R] estimates table

tabulate twoway, [SVY] svy: tabulate twoway

tabstat command, [R] tabulate, summarize()
hazard, see hazard tables
life, see life tables
missing values, [R] misstable
of statistics, [P] tabdisp

printing, [U] 15 Saving and printing output—log files
summary statistics, [R] table, [R] tabstat,
        [R] tabulate, summarize()
tabodds command, [R] epitab
tabstat command, [R] tabstat

tabulate
two-way, [SVY] svy: tabulate twoway


tabulate command, [R] tabulate oneway,
        [R] tabulate twoway
        summarize(), [R] tabulate, summarize()
tag, duplicates subcommand, [D] duplicates
tag(), egen function, [D] egen
tan() function, [FN] Trigonometric functions,
        [M-5] sin()
tangent function, [FN] Trigonometric functions
tanh() function, [FN] Trigonometric functions
        [M-5] sin()

TARCH, see threshold autoregressive conditional heteroskedasticity
target
between-group variance, [PSS] power oneway
correlation, [PSS] power, [PSS] power onecorrelation, [PSS] power twocorrelations
discordant proportions, [PSS] power, [PSS] power pairedproportions
effect variance, [PSS] power twoway, [PSS] power repeated
hazard difference, [PSS] power exponential
hazard ratio, [PSS] power exponential, [PSS] power logrank
log hazard-ratio, [PSS] power exponential,
        [PSS] power logrank
mean, [PSS] power, [PSS] power onemean,
        [PSS] power twomeans, [PSS] unbalanced designs
mean difference, [PSS] power, [PSS] power pairedmeans
odds ratio, [PSS] power cmh, [PSS] power mcc
parameter, [PSS] Glossary
tau, [R] spearman
taxonomy, [MV] Glossary

Taylor linearization, see linearized variance estimator
tC() pseudofunction, [D] datetime, [FN] Date and time functions
tc() pseudofunction, [D] datetime, [FN] Date and time functions

TCC, see test characteristic curve
tcc, irtgraph subcommand, [IRT] irtgraph tcc
td() pseudofunction, [D] datetime, [FN] Date and time functions
tden() function, [FN] Statistical functions,
        [M-5] normal()

TDT test, see transmission-disequilibrium test
tebalance
balance command, [TE] tebalance box,
        [TE] tebalance density, [TE] tebalance overid,
        [TE] tebalance summarize
command, [TE] tebalance
technical support, [U] 3.8 Technical support
technique, [SEM] Glossary
technique() option, [R] maximize
tffects
aipw command, [TE] teffects aipw
command, [TE] tebalance, [TE] teffects,
        [TE] teffects postestimation
ipw command, [TE] teffects ipw
ipwra command, [TE] teffects ipwra
nmmatch command, [TE] teffects nmmatch
overlap command, [TE] teffects overlap
psmatch command, [TE] teffects psmatch
ra command, [TE] teffects ra
tffects, estat subcommand, [SEM] estat tffects
tempfile command, [P] macro
tempfile macro extended function, [P] macro
tempname, class, [P] class
tempname macro extended function, [P] macro
        [U] 18.7.2 Temporary scalars and matrices
variables, [P] macro, [U] 18.7.1 Temporary variables
tempvar command, [P] macro
tempvar macro extended function, [P] macro
termcap(5), [U] 10 Keyboard use
test, equality of, continued
distributions, see distributions, testing equality of
means, see equality test of means
medians, see equality test of medians
proportions, see equality test of proportions
ROC areas, see equality test of ROC areas
survivor functions, see equality test, survivor functions
variances, see equality test of variances
equivalence, see equivalence test
exact, see exact test
exogeneity, see endogeneity test
\( F \), see \( F \) test
Fisher–Irwin’s exact, see Fisher–Irwin’s exact test
Fisher-type, see Fisher-type test
Fisher’s exact, see Fisher’s exact test
Fisher’s \( z \), see Fisher’s \( z \) test
goodness-of-fit, see goodness of fit
Granger causality, see Granger causality
group invariance, see group invariance test
Hadri Lagrange multiplier, see Hadri Lagrange multiplier stationarity test
Harris–Tzavalis, see Harris–Tzavalis test
Hausman specification, see Hausman specification test
heterogeneity, see heterogeneity test
heteroskedasticity, see heteroskedasticity test
homogeneity, see homogeneity test
hypothesis, see hypothesis test
Im–Pesaran–Shin, see Im–Pesaran–Shin test
independence, also see Breusch–Pagan test, see independence test
independence of irrelevant alternatives, see independence of irrelevant alternatives
information matrix, see information matrix test
internal consistency, see internal consistency test
intrarrater agreement, see intrarrater agreement
Kolmogorov–Smirnov, see Kolmogorov–Smirnov test
Kruskal–Wallis, see Kruskal–Wallis test
kurtosis, see kurtosis
Lagrange multiplier, see Lagrange multiplier test
Levin–Lin–Chu, see Levin–Lin–Chu test
likelihood-ratio, see likelihood-ratio test
linear hypotheses after estimation, see linear hypothesis test after estimation
log-rank, see log-rank test
Mantel–Haenszel, see Mantel–Haenszel test
marginal homogeneity, [PSS] power mcc, see marginal homogeneity, test of
medians, see margins test
McNemar’s, see McNemar’s test
McNemar’s \( z \) test, see McNemar’s \( z \) test
model
coefficients, see model coefficients test
simplification, see model simplification test
test, model, continued
  specification, see specification test
modification indices, see modification indices
multiple-sample, see multiple-sample test
multivariate, see multivariate test
nonlinear, see nonlinear test
nonlinear hypotheses after estimation, see nonlinear hypothesis test after estimation
normality, see normal distribution and normality, see normality test
omitted variables, see omitted variables test
one-sample, see one-sample test
one-sided, see one-sided test
overidentifying restrictions, see overidentifying restrictions, tests of
overlap assumption, see overlap assumption
paired-sample, see paired-sample test
permutation, see permutation test
quadrature, see quadrature
Ramsey, see Ramsey test
random-order, see random-order test
RESET, see RESET test
score, see score test
serial correlation, see autocorrelation
serial independence, see serial independence test
Shapiro–Francia, see Shapiro–Francia test for normality
Shapiro–Wilk, see Shapiro–Wilk test for normality
skewness, see skewness
structural break, see structural break
symmetry, see symmetry test
Szroeter’s, see Szroeter’s test for heteroskedasticity
t, see t test
TDT, see transmission-disequilibrium test
transmission-disequilibrium test, see transmission-disequilibrium test
trend, see trend, test for
two-sample, see two-sample test
two-sided, see two-sided test
unit-root, see unit-root test
variance-comparison, see variance-comparison test
Vuong, see Vuong test
Wald, see Wald test
weak instrument, see weak instrument test
z, see z test
test command, [R] anova postestimation, [R] test,
  [SEM] estat stdize, [SEM] example 8,
  [SEM] example 9, [SEM] example 16,
  [SEM] test, [SVY] survey, [SVY] svy postestimation,
  [U] 20.12 Performing hypothesis tests on the coefficients
test, mi subcommand, [MI] mi test
test, sts subcommand, [ST] sts test
testnl command, [R] testnl, [SEM] estat stdize,
  [SEM] testnl, [SVY] svy postestimation
testparm command, [R] testparm, [SEM] test, [SVY] svy postestimation
testtransform, mi subcommand, [MI] mi test
tetrachoric command, [R] tetrachoric
tetrachoric correlation, [MV] Glossary
text files, writing and reading, [P] file
text() option, [G-3] added_text_options,
  [G-3] aspect_option
text,
  adding, [G-3] added_text_options
  angle of, [G-4] anglestyle
  captions, [G-3] title_options
  exporting, see exporting data
  in graphs, [G-4] text
  note, [G-3] title_options
  reading data in, see importing data
  resizing, [G-3] scale_option
  running outside of borders,
  [G-3] added_text_options
  saving data in, see exporting data
  size of, [G-3] textbox_options
  subtitle, [G-3] title_options
  title, [G-3] title_options
  vertical alignment, [G-4] alignmentstyle
text and textboxes, relationship between, [G-4] textstyle
textboxes, [G-3] textbox_options
  orientation of, [G-4] orientationstyle
textboxstyle, [G-4] textboxstyle
  textsize, [G-4] textsizestyle
textstyle, [G-4] textstyle
th() pseudofunction, [D] datetime, [FN] Date and time functions
thickness of lines, [G-4] linewidthstyle
thinning, [BAYES] bayesmh, [BAYES] Glossary
Thomson scoring, [MV] factor postestimation
thrashing, [ST] Glossary
three-dimensional graph, [G-2] graph twoway contour,
  [G-2] graph twoway contourline
three-parameter logistic model, [IRT] irt 3pl,
  [IRT] Glossary
tick, [TS] arch
  definition, [G-4] tickstyle
  suppressing, [G-4] tickstyle
  ticksetstyle, [G-4] ticksetstyle
tickstyle, [G-4] tickstyle
ties, [MV] Glossary
TIF, see test information function
tif, irtgraph subcommand, [IRT] irtgraph tif
TIFF, [G-3] tif_options
time of day, [P] creturn
time-series–operated variable, [M-5] st_data(),
timestep, [D] describe
time variables and values, [D] datetime
time-domain analysis, [TS] arch, [TS] arfima,
  [TS] arima, [TS] Glossary
timeout1, set subcommand, [R] netio, [R] set
timeout2, set subcommand, [R] netio, [R] set
timer
clear command, [P] timer
list command, [P] timer
off command, [P] timer
on command, [P] timer
times and dates, [M-5] c(), [M-5] date()
time-series
analysis, [D] egen, [P] matrix accum, [R] regress
postestimation time series
estimation, [U] 26.19 Models with time-series data
filter, [TS] psdensity, [TS] ucm
formats, [D] format
functions, [FN] Selecting time-span functions
operators, [TS] tsset, [U] 13.10 Time-series operators
tline
unabbreviating varlists, [P] unab
varlists, [U] 11.4.4 Time-series varlists
time-span data, [ST] snapspan
time-varying covariates, [ST] Glossary
time-varying variance, [TS] arch
time-versus-concentration curve, [R] pk
timing code, [P] timer
tin() function, [FN] Selecting time-span functions
title, estimates subcommand, [R] estimates title
title() option, [G-3] title_options
title-cased string, [I] Glossary
titles, [G-3] title_options
of axis, [G-3] axis_title_options
tlabel() option, [G-3] axis_label_options
TLI, see Tucker–Lewis index
tm() pseudofunction, [D] datetime, [FN] Date and time functions
tmlabel() option, [G-3] axis_label_options
TMPDIR Unix environment variable, [P] macro
tmtick() option, [G-3] axis_label_options
tnbreg command, [R] tnbr2, [R] tnbr2 postestimation
tobit command, [R] tobit, [R] tobit postestimation
tobit regression, [R] ivtobit, [R] tobit,
[SEM] example 43g, [SVY] svy estimation,
also see intreg command, also see truncreg command
random-effects, [XT] xtreg
with endogenous covariates, [SVY] svy estimation
tobytes() function, [FN] String functions
.time filename suffix, [R] net
Toeplitz() function, [M-5] Toeplitz()
tokenallowhex() function, [M-5] tokenget()
tokenallownum() function, [M-5] tokenget()
tokenget() function, [M-5] tokenget()
tokengetall() function, [M-5] tokenget()
tokeninit() function, [M-5] tokenget()
tokeninitstata() function, [M-5] tokenget()
tokenize command, [P] tokenize
tokenoffset() function, [M-5] tokenget()
tokenpchars() function, [M-5] tokenget()
tokenpeek() function, [M-5] tokenget()
tokenqchars() function, [M-5] tokenget()
tokens() function, [M-5] tokens()
tokens() function, [M-5] tokenset()
tokenset() function, [M-5] tokenget()
tokenwchars() function, [M-5] tokenget()
tolerance() option, [R] maximize
tolerances, [M-1] tolerance, [M-5] solve_tol()
top() suboption, [G-4] alignmentstyle
tostring command, [D] destring
total
characteristic curve, see test characteristic curve
inertia, [MV] ca, [MV] ca postestimation,
[MV] mca, [MV] mca postestimation,
[MV] Glossary
information function, see test information function
principal inertia, [MV] ca, [MV] mca,
[MV] Glossary
sample size, see sample-size
total command, [R] total, [R] total postestimation
total(), egen function, [D] egen
totals, estimation, [R] total
totals, survey data, [SVY] svy estimation
toward a target rotation, [MV] procrustes, [MV] rotate,
[MV] rotatemat
tpoisson command, [R] tpoisson, [R] tpoisson
postestimation
trace(), pseudofunction, [D] datetime, [FN] Date and time functions
trace,
ml subcommand, [R] ml
query subcommand, [R] query
trace() function, [FN] Matrix functions,
[M-5] trace(), [P] matrix define
trace of matrix, [M-5] trace(), [P] matrix define
trace option, [R] maximize
[M-6] Glossary
tracing iterative maximization process, [R] maximize
trademark symbol, [G-4] text
training, [U] 3.6 Conferences and training
transferring data
  copying and pasting, [D] edit
  from Stata, [D] export, [U] 21.4 Transfer programs
  into Stata, [D] import, [U] 21 Entering and importing data, [U] 21.4 Transfer programs
transformations, [MV] procrustes
  log, [R] lnskew0
  modulus, [R] boxcox
  power, [R] boxcox, [R] lnskew0
  to achieve normality, [R] boxcox, [R] ladder
  to achieve zero skewness, [R] lnskew0
transformed coefficients, [MI] mi estimate, [MI] mi estimate using, [MI] mi test
translate
  files with Unicode, [D] edit
  logs, [R] translate
  translate command, [R] translate
  translation, file, [D] changeeol
translator
  query command, [R] translate
  reset command, [R] translate
  set command, [R] translate
transmap
  define command, [R] translate
  query command, [R] translate
transmission-disequilibrium test, [R] symmetry
transpose, [M-6] Glossary, also see conjugate transpose
  in place, [M-5] _transpose()
  operator, [M-2] op_transpose
  without conjugation, [M-5] transposeonly()
  _transpose() function, [M-5] _transpose()
  _transposeonly() function, [M-5] transposeonly()
  transposeonly() function, [M-5] transposeonly()
  transposing data, [D] xplose
transposing matrices, [P] matrix define
transposition, [M-2] op_transpose, [M-5] _transpose()
  [M-5] transposeonly()

treatment effects,
doubly robust estimators, [TE] teffects aipw, [TE] teffects ipwra
glossary, [TE] Glossary
inverse-probability weighting, [TE] stteffects ipw, [TE] teffects ipw
matching estimators, [TE] teffects nnmatch, [TE] teffects psmatch
overlap plots, [TE] teffects overlap
overview, [TE] intro, [TE] treatment effects,
treatment effects, continued
  postestimation, [TE] teffects postestimation
  power, [PSS] power, [PSS] power twomeans,
  [PSS] power pairedmeans, [PSS] power oneprop,
  [PSS] power logrank
  regression adjustment, [TE] stteffects, [TE] teffects
  survey data, [SVY] svy estimation
tree, misstable subcommand, [R] misstable
  trees, [MV] cluster, [MV] cluster dendrogram
trend, [TS] Glossary
  trend, power subcommand, [PSS] power trend
trend test, [PSS] power, [PSS] power trend,
  [R] epitetab, [ST] strate, [ST] sts test
trend, test for, [R] nptrend, [R] symmetry
  triangle kernel function, [R] kdensity, [R] lpoly,
  trigamma() function, [FN] Mathematical functions,
  [M-5] factorial()
  trigonometric functions, [FN] Trigonometric functions,
  [M-5] sin()
  trunc() function, [FN] Mathematical functions,
  [M-5] trunc()

truncated
  negative binomial regression, [R] tnbinreg, [SVY] svy estimation
  observations, [R] truncreg, also see censored observations
  Poisson regression, [R] tpoisson, [SVY] svy estimation
  regression, [R] truncreg, [SVY] svy estimation
  truncating
  real numbers, [FN] Mathematical functions
  strings, [FN] String functions
  also see imputation, truncated data
truncreg command, [R] truncreg, [R] truncreg postestimation
tssappend command, [TS] tssappend
tsscale, graph twoway subcommand, [G-2] graph twoway tslave
tslib command, [G-3] axis_scale_options
tsfilter command, [TS] tsfill
tsfilter, [TS] tsfilter
  bk command, [TS] tsfilter bk
  bw command, [TS] tsfilter bw
  cf command, [TS] tsfilter cf
  hp command, [TS] tsfilter hp
  tsline command, [TS] tsline
tsline, graph twoway subcommand, [G-2] graph
twoway tsline
tsnorm macro extended function, [P] macro
tswar command, [TS] tsreport
tswave command, [TS] tsvar
tswave command, [TS] tsline
tsrline, graph twoway subcommand, [G-2] graph
twoway tsline
tset command, [TS] tsset
tset command for mi data, [MI] mi XXXset
tset, mi subcommand, [MI] mi XXXset
tssmooth, [TS] tssmooth
dexpomential command, [TS] tssmooth
dexpomentional command, [TS] tssmooth exponential
dexpomential command, [TS] tssmooth exponential
dexpomential command, [TS] tssmooth exponential
hwinters command, [TS] tssmooth hwinters
ma command, [TS] tssmooth ma
ml command, [TS] tssmooth ml
shwinters command, [TS] tssmooth shwinters
tsunab command, [P] unab
ttail() function, [FN] Statistical functions, [M-5] normal()
test and ttesti commands, [R] ttest
test command, [MV] hotelling
ttick() option, [G-3] axis_label_options
ttitle() option, [G-3] axis_title_options
Tucker–Lewis index, [SEM] estat gof, [SEM] methods and formulas for sem
tukeyprob() function, [FN] Statistical functions, [M-5] normal()
Tukey’s multiple-comparison adjustment, see multiple comparisons, Tukey’s method
Studentized range distribution, cumulative, [FN] Statistical functions
inverscumulative, [FN] Statistical functions
tuning constant, [R] rreg
tutorials, [U] 1.2.2 Example datasets
tw() pseudofunction, [D] datetime, [FN] Date and time functions
twithin() function, [FN] Selecting time-span functions
Twitter, see Stata on Twitter
two-level model, [ME] me, [ME] Glossary
twocorrelations, power subcommand, [PSS] power
twocorrelations
two-independent-samples test, [PSS] Glossary
twomeans, power subcommand, [PSS] power
twomeans
twoproportions, power subcommand, [PSS] power
twoproportions
two-sample, continued
standard deviations, see standard deviations, two-sample
study, [PSS] power, [PSS] unbalanced designs
students, [PSS] intro, [PSS] power, [PSS] Glossary
correlations, [PSS] power twocorrelations
dependent samples, [PSS] power mcc
hazard functions, [PSS] power exponential, [PSS] power logrank
independent samples, [PSS] power twomeans,
[PSS] power twoproportions, [PSS] power twovariances, [PSS] power twocorrelations,
[PSS] power cmh, [PSS] power exponential,
[PSS] power logrank
log hazards, [PSS] power exponential,
[PSS] power logrank
log-rank, [PSS] power exponential, [PSS] power logrank
means, [PSS] power twomeans, [PSS] power
paired means, [PSS] unbalanced designs
proportions, [PSS] power twoproportions,
[PSS] power pairedproportions, [PSS] power
cmh, [PSS] power mcc
survivor functions, [PSS] power exponential,
[PSS] power logrank
variances, [PSS] power twovariances
variances, see variances, two-sample
two-sided test, [PSS] power, [PSS] power onemean,
[PSS] power twomeans, [PSS] power
paired means, [PSS] power oneproportion,
[PSS] power twoproportions, [PSS] power
pairedproportions, [PSS] power onevariance,
[PSS] power twovariances, [PSS] power
onecorrelation, [PSS] power twocorrelations,
[PSS] power oneway, [PSS] power repeated,
[PSS] power cmh, [PSS] power mcc,
[PSS] power trend, [PSS] power cox,
[PSS] power exponential, [PSS] power logrank,
[PSS] unbalanced designs, [PSS] Glossary
two-stage least squares, [R] ivregress, [SVY] svy
estimation, [XT] xthtaylor, [XT] xtivreg
two-tailed test, [PSS] ivregress, [SVY] svy
estimation, [XT] xthtaylor, [XT] xtivreg
two-way analysis of variance, [PSS] power, [PSS] power
twoway, [PSS] Glossary, [R] anova
multivariate analysis of variance, [MV] manova
repeated-measures ANOVA, [PSS] power repeated,
[PSS] Glossary
scatterplots, [R] lowess
twoway, power subcommand, [PSS] power twoway
type command, [D] type
macro extended function, [P] macro
parameter, [D] generate
type.

set subcommand, [D] generate, [R] set
ssc subcommand, [R] ssc
type, broad, [M-6] Glossary
type I error probability, see probability of a type I error
type I study, [PSS] Glossary
type II error probability, see type II error probability,
see type II error, [PSS] Glossary
type II study, [PSS] Glossary
type II error probability, see probability of a type II error

type II study, [PSS] Glossary

U

U statistic, [R] ranksum
UCA, see Unicode collation algorithm
uchar( ) function, [FN] String functions,
[M-5] uchar( )
UCM, see unobserved-components model
ucm command, [TS] ucm, [TS] ucm postestimation
uconv, [D] unicode convertfile
udstrlen macro extended function, [P] macro
udstrlen() function, [FN] String functions,
[M-5] ustrlen()
udsubstr() function, [FN] String functions,
[M-5] usubstr()
uisdigit() function, [FN] String functions
uisletter() function, [FN] String functions
unab command, [P] unab
unabbreviate
command names, [P] unabcmd
variable list, [P] syntax, [P] unab
unabcmd command, [P] unabcmd
unadgroup, ssd subcommand, [SEM] ssd
.uname built-in class function, [P] class
unary operator, [M-6] Glossary
unbalanced
data, [XT] Glossary
design, [PSS] power twomeans, [PSS] power twoportions, [PSS] power twovariances,
[PSS] power twocorrelations, [PSS] power oneway, [PSS] power twoay, [PSS] power repeated,
[PSS] power cmh, [PSS] power trend,
[PSS] power exponential, [PSS] power logrank,
[PSS] unbalanced designs, [PSS] Glossary
uncompress files, [D] zipfile
unconfoundedness, see conditional-independence assumption
under observation, [ST] cttost, [ST] st, [ST] stset,
[ST] Glossary
underlining in syntax diagram, [U] 11 Language syntax
underscore variables, [U] 13.4 System variables (...variables)
unequal-allocation design, see unbalanced design
unhold, _estimates subcommand, [P] _estimates

Unicode, [D] unicode, [I] Glossary
collation, [FN] String functions,
[M-5] ustrcompare(), [U] 12.4.2.5 Sorting strings containing Unicode characters
collation algorithm, [D] unicode collator
collators, [D] unicode collator
count, [M-6] Glossary
coding conversion, [D] unicode convertfile,
[D] unicode translate
codings, [D] unicode encoding,
[U] 12.4.2.3 Encodings
locales, [D] unicode locale, [P] set locale_functions,
[P] set locale_ui, [U] 12.4.2.4 Locales in Unicode
normalization, [I] Glossary
strings, [FN] String functions, [M-4] string,
[U] 12.4.2 Handling Unicode strings
title-cased string, [I] Glossary
unidecode
analyze command, [D] unicode translate
collator list command, [D] unicode collator
collation, [D] unicode
collation with collator alias, [D] unicode collator
encoding alias command, [D] unicode encoding
encoding collation command, [D] unicode encoding
encoding set command, [D] unicode encoding,
[D] unicode translate
erasebackups command, [D] unicode translate
locale list command, [D] unicode locale
restore command, [D] unicode translate
retranslate command, [D] unicode translate
translate command, [D] unicode translate
upackage list command, [D] unicode locale
uninstall
unique, query subcommand, [R] query
unidimensionality, [IRT] Glossary
uniform
accrual, [PSS] power exponential,
[PSS] power logrank
prior, [BAYES] bayes, [BAYES] bayesmh,
[M] mi impute mvn
uniformly distributed
random numbers, [M-5] runiform()
random variates, [M-5] runiform()
random-number function, [FN] Random-number functions, [R] set seed
uninstall,
net subcommand, [R] net
ssc subcommand, [R] ssc
uniqrows() function, [M-5] uniqrows()
unique options, [G-4] concept: repeated options
unique value labels, [D] labelbook
unique values,
counting, [D] codebook, [R] table, [R] tabulate
one way
determining, [D] inspect, [D] labelbook
uniqueness, [MV] factor, [MV] factor postestimation,
[MV] rotate, [MV] Glossary
unit loading, [SEM] intro 4
unit vectors, [M-5] e()  
unitary matrix, [M-6] Glossary  
unitcircle() function, [M-5] unitcircle()  
unit-root  
models, [TS] vec intro, [TS] vec  
process, [TS] Glossary  
univariate  
distributions, displaying, [R] cumul, [R] diagnostic plots, [R] histogram, [R] Iv, [R] stem  
imputation, see imputation, univariate  
kernel density estimation, [R] kdensity  
Unix,  
keyboard use, [U] 10 Keyboard use  
pause, [P] sleep  
specifying filenames, [U] 11.6 Filenaming conventions  
 unlink() function, [M-5] unlink()  
unlink() function, [M-5] unlink()  
unobserved-components model, [TS] psdensity  
model, [TS] ucm  
postestimation, [TS] ucm postestimation  
unorder() function, [M-5] sort()  
unregister, mi subcommand, [MI] mi set  
unregistered variables, see variables, unregistered  
unrestricted FMI test, [MI] mi estimate, [MI] mi test, [MI] Glossary  
unrestricted transformation, [MV] procstmates  
postestimation, [MV] Glossary  
unstandardized coefficient, [SEM] Glossary  
unstructured, [SEM] Glossary  
unzifile command, [D] zipfile  
update  
all command, [R] update  
command, [R] update  
from command, [R] update  
query command, [R] update  
update,  
mi subcommand, [MI] mi update, [MI] noupdate  
option  
query subcommand, [R] query  
view subcommand, [R] view  
update_d, view subcommand, [R] view  
update_interval, set subcommand, [R] set, [R] update  
update_prompt, set subcommand, [R] set, [R] update  
update_query, set subcommand, [R] set, [R] update  
updates to Stata, [R] adoupdate, [R] net, [R] sj, [R] update, [U] 3.4 The Stata Journal, [U] 3.5 Updating and adding features from the web, [U] 17.6 How do I install an addition?, [U] 28 Using the Internet to keep up to date  
upper  
one-sided test, [PSS] Glossary  
one-tailed test, [PSS] Glossary  
upercase-string function, [FN] String functions  
_uppertriangle() function, [M-5] lowertriangle()  
upper() function, [M-5] lowertriangle()  
upper-triangular matrix, see triangular matrix use,  
cluster subcommand, [MV] cluster utility  
estimates subcommand, [R] estimates save  
graph subcommand, [G-2] graph use  
serset subcommand, [P] sser  
use command, [D] use  
uselabel command, [D] labelbook  
user interface, [P] dialog programming  
language, [D] unicode locale  
localization package, [D] unicode locale  
user-written additions,  
installing, [R] net, [R] ssc  
searching for, [R] net search, [R] ssc  
using,  
cmdlog subcommand, [R] log  
log subcommand, [R] log  
using data, [D] sysuse, [D] use, [D] webuse, [P] syntax, also see importing data  
using graphs, [G-2] graph use  
ustrcompare() function, [FN] String functions, [M-5] ustrcompare()  
ustrcompareex() function, [FN] String functions, [M-5] ustrcompare()  
ustrfix() function, [FN] String functions, [M-5] ustrfix()  
ustrfrom() function, [FN] String functions, [M-5] ustrto()  
ustrinvalidcnt() function, [FN] String functions, [M-5] ustrlen()  
ustrleft() function, [FN] String functions  
ustrlen macro extended function, [P] macro  
ustrlen() function, [FN] String functions, [M-5] ustrlen()  
ustrlower() function, [FN] String functions, [M-5] ustrupper()  
ustrltrim() function, [FN] String functions, [M-5] ustrtrim()  
ustrnormalize() function, [FN] String functions, [M-5] ustrnormalize()  
ustrpos() function, [FN] String functions, [M-5] ustrpos()  
ustrregexm() function, [FN] String functions  
ustrregexra() function, [FN] String functions  
ustrregexrf() function, [FN] String functions  
ustrreverse() function, [FN] String functions, [M-5] ustrreverse()  
ustrright() function, [FN] String functions  
ustrnpos() function, [FN] String functions, [M-5] ustrpos()
variables, continued
registered, [MI] mi rename, [MI] mi set, [MI] Glossary
renaming, see renaming variables
reordering, [D] order
setting properties of, [D] varmanage
sorting, [D] gsort, [D] sort, [D] varmanage
standardizing, [D] egen
storage types, see storage types
string, see string variables
system, see system variables
tab expansion of, [U] 10.6 Tab expansion of variable names
temporary, [P] macro
transposing with observations, [D] xpose
unabbreviating, [P] syntax, [P] unab
unique values, [D] codebook, [D] duplicates, [D] inspect
unregistered, [MI] mi rename, [MI] mi set, [MI] Glossary

Variables Manager, [D] varmanage
variance
analysis, [MV] manova
components, [ME] Glossary, also see mixed model decompositions, see forecast-error variance decomposition
creating dataset of, [D] collapse
creating variable containing, [D] egen
Huber/White/sandwich estimator, see robust, Huber/White/sandwich estimator of variance inflation factors, [R] regress postestimation linearized, [SVY] variance estimation nonconstant, [SVY] variance estimation, see robust, see robust, Huber/White/sandwich estimator of variance stabilizing transformations, [R] boxcox
testing equality of, [R] sdttest


variance() function, [M-5] mean()

variance() option, see gsem option variance(), see sem option variance()

variance-comparison test, [MV] mvtest covariances, [R] sdttest

variances, ci subcommand, [R] ci
varkeyboard, set subcommand, [R] set
time series, [U] 11.4.4 Time-series varlists
varlmar command, [TS] varlmar
varmanage command, [D] varmanage
varnorm command, [TS] varnorm
varsoc command, [TS] varsoc
varstable command, [TS] varstable
varwle command, [TS] varwle
varying conditional-correlation model, [TS] mgarch

[T] mgarch vce
estimation sample, [MI] mi estimate
variables, [ST] stvar, see variables, varying and super varying
varying, mi subcommand, [MI] mi varying
vcc, mgarch subcommand, [TS] mgarch vce
VCE, see variance–covariance matrix of estimators vce, estat subcommand, [R] estat, [R] estat vce, [SVY] estat vce() option, [R] vce_option, [XT] vce_options, see gsem option vce(), also see sem option vce()

veclmar command, [TS] veclmar
VECM, see vector error-correction model
views, [M-6] Glossary

virtual memory, [D] memory


W

[SEM] estat eqtest, [SEM] estat ginvariant, [SEM] example 13, [SEM] example 22,
tests on the coefficients, [U] 20.12.4 Nonlinear Wald tests

wardslinkage,

clustermat subcommand, [MV] cluster linkage
class subcommand, [MV] cluster linkage

Ward’s linkage clustering, [MV] cluster,
[MV] clustermat, [MV] cluster linkage,
[MV] Glossary

Ward’s method clustering, [MV] cluster,
[MV] clustermat

warning messages, [M-2] pragma

waveragelinkage,

clustermat subcommand, [MV] cluster linkage
class subcommand, [MV] cluster linkage

wcorrelation, estat subcommand, [ME] mixed

postestimation, [XT] xtggee postestimation

weak instrument test, [R] ivregress postestimation

weakly balanced, [XT] Glossary

website,

stata.com, [U] 3.2.1 The Stata website
(www.stata.com)

stata-press.com, [U] 3.3 Stata Press

webuse command, [D] webuse

query command, [D] webuse

set command, [D] webuse

week() function, [D] datetime, [FN] Date and time
functions, [M-5] date( )

weekly() function, [D] datetime, [D] datetime

translation, [FN] Date and time functions,
Windows,
filenames, [U] 18.3.11 Constructing Windows filenames by using macros
keyboard use, [U] 10 Keyboard use
pause, [P] sleep
specifying filenames, [U] 11.6 Filenaming conventions
winexec command, [D] shell
Wishart distribution, [MV] Glossary
density, [FN] Statistical functions
prior, [BAYES] bayesmh, [BAYES] bayesmh evaluators
within matrix, [MV] Glossary
within-cell
means and variances, [XT] xtsum
variance, [PSS] power twoway
within-group variance, [PSS] power oneway
within-imputation variability, [MI] mi estimate, [MI] mi predict
within-subject
design, [PSS] power repeated, [PSS] Glossary
factor, [PSS] power repeated, [PSS] Glossary
variance, [PSS] power repeated
WLF, see worst linear function
WLS, see weighted least squares
wntestb command, [TS] wntestb
wntestq command, [TS] wntestq
wofd() function, [D] datetime, [FN] Date and time functions, [M-5] date()
Woolf confidence intervals, [R] epitab
word macro extended function, [P] macro
word() function, [FN] String functions
Word, Microsoft, see Microsoft Word
wordbreaklocale() function, [FN] String functions
wordcount() function, [FN] String functions
workflow, [MI] workflow
worst linear function, [MI] mi impute mvn, [MI] Glossary
wra, stteffects subcommand, [TE] stteffects wra
write, file subcommand, [P] file
writing and reading text and binary files, [P] file
writing data, see exporting data, see saving data
www.stata.com website, [U] 3.2.1 The Stata website (www.stata.com)
www.stata-press.com website, [U] 3.3 Stata Press

X
xaxis() suboption, [G-3] axis_choice_options
X-bar charts, [G-2] graph other
xchart command, [R] qc
xcommon option, [G-2] graph combine
xcorr command, [TS] xcorr
xeq, mi subcommand, [MI] mi xeq
xi prefix command, [R] xi
xl() function, [M-5] xl()
xlabel() option, [G-3] axis_label_options
xline() option, [G-3] added_line_options
XML, [D] xmlsave
xmlabel() option, [G-3] axis_label_options
xmlsave command, [D] xmlsave
xmluse command, [D] xmlsave
xntick() option, [G-3] axis_label_options
xpose command, [D] xpose
xscale() option, [G-3] axis_scale_options
xshell command, [D] shell
xtabond command, [XT] xtabond, [XT] xtabond postestimation
xtcloglog command, [XT] quadchk, [XT] xtcloglog, [XT] xtcloglog postestimation
xtdata command, [XT] xtdata
xtdescribe command, [XT] xtdescribe
xtdpd command, [XT] xtdpd, [XT] xtdpd postestimation
xtdpsys command, [XT] xtdpsys, [XT] xtdpsys postestimation
xtfroniter command, [XT] xtfroniter, [XT] xtfroniter postestimation
xtgee command, [XT] xttgee, [XT] xttgee postestimation
xtgls command, [XT] xttgls, [XT] xttgls postestimation
xthtaylor command, [XT] xthtaylor, [XT] xthtaylor postestimation
xtick() option, [G-3] axis_label_options
xtile command, [D] pctile
xtintreg command, [XT] quadchk, [XT] xttintreg, [XT] xttintreg postestimation
xtlongit command, [XT] xttlongit, [XT] xttlongit postestimation
xtprobit command, [XT] quadchk, [XT] xttprobit, [XT] xttprobit postestimation
xttreg command, [XT] xttreg, [XT] xttreg postestimation
xttncse command, [XT] xttncse, [XT] xttncse postestimation
xttngis command, [XT] xttngis, [XT] xttngis postestimation
xttprobit command, [XT] xttprobit, [XT] xttprobit postestimation
xtw command, [XT] xtw, [XT] xtw postestimation
xtwide command, [XT] xttwide, [XT] xttwide postestimation
xtwocode command, [XT] xttwocode, [XT] xttwocode postestimation
Z

z test, [PSS] Glossary
Zellner’s
g-prior, [BAYES] bayes, [BAYES] bayesmh,
[BAYES] Glossary
seemingly unrelated regression, [R] sureg, [R] reg3,
[R] suest
zero matrix, [P] matrix define
zero-altered, see zero-inflated
zero-inflated
negative binomial regression, [R] zinb, [SVY] svy estimation
Poisson regression, [R] zip, [SVY] svy estimation
zero-skewness transform, [R] lnskew0
zinb command, [R] zinb, [R] zinb postestimation
zip command, [R] zip, [R] zip postestimation
zipfile command, [D] zipfile
zlabel() option, [G-3] axis_label_options
zmlabel() option, [G-3] axis_label_options
zmtick() option, [G-3] axis_label_options
zscale() option, [G-3] axis_scale_options
ztest and ztseti commands, [R] ztest
ztick() option, [G-3] axis_label_options
ztitle() option, [G-3] axis_title_options

Y

yaxis() suboption, [G-3] axis_choice_options
ycommon option, [G-2] graph combine
year() function, [D] datetime, [FN] Date and time functions, [M-5] date()
yearly() function, [D] datetime, [D] datetime
translation, [FN] Date and time functions,
[M-5] date()
yh() function, [D] datetime, [FN] Date and time functions,
[M-5] date()
ylabel() option, [G-3] axis_label_options
yline() option, [G-3] added_line_options
ym() function, [D] datetime, [FN] Date and time functions,
[M-5] date()
ymlabel() option, [G-3] axis_label_options
ymtick() option, [G-3] axis_label_options
yofd() function, [D] datetime, [FN] Date and time functions,
[M-5] date()
YouTube Channel, see Stata YouTube Channel
yq() function, [D] datetime, [FN] Date and time functions,
[M-5] date()
yscale() option, [G-3] axis_scale_options
ysize() option, [G-2] graph display,
region_options
ytitle() option, [G-3] axis_title_options
Yule coefficient similarity measure,
measure_option
yvarformat() option, [G-3] advanced_options
yvarlabel() option, [G-3] advanced_options
yw() function, [D] datetime, [FN] Date and time functions,
[M-5] date()