Description

Multiple-imputation data analysis in Stata is similar to standard data analysis. The standard syntax applies, but you need to remember the following for MI data analysis:

1. The data must be declared as mi data.

   If you already have multiply imputed data (saved in Stata format), use `mi import` to import it into mi; see `[MI] mi import`.

   If you do not have multiply imputed data, use `mi set` (see `[MI] mi set`) to declare your original data to be mi data and use `mi impute` (see `[MI] mi impute`) to fill in missing values.

2. After you have declared mi data, commands such as `svyset`, `stset`, and `xtset` cannot be used. Instead use `mi svyset` to declare survey data, use `mi stset` to declare survival data, and use `mi xtset` to declare panel data. See `[MI] mi XXXset`.

3. Prefix the estimation commands with `mi estimate` (see `[MI] mi estimate`).

   The following estimation commands support the `mi estimate` prefix.

<table>
<thead>
<tr>
<th>Command</th>
<th>Entry</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression models</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>regress</code></td>
<td>[R] <code>regress</code></td>
<td>Linear regression</td>
</tr>
<tr>
<td><code>cnsreg</code></td>
<td>[R] <code>cnsreg</code></td>
<td>Constrained linear regression</td>
</tr>
<tr>
<td><code>mvreg</code></td>
<td>[MV] <code>mvreg</code></td>
<td>Multivariate regression</td>
</tr>
<tr>
<td>Binary-response regression models</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>logistic</code></td>
<td>[R] <code>logistic</code></td>
<td>Logistic regression, reporting odds ratios</td>
</tr>
<tr>
<td><code>logit</code></td>
<td>[R] <code>logit</code></td>
<td>Logistic regression, reporting coefficients</td>
</tr>
<tr>
<td><code>probit</code></td>
<td>[R] <code>probit</code></td>
<td>Probit regression</td>
</tr>
<tr>
<td><code>cloglog</code></td>
<td>[R] <code>cloglog</code></td>
<td>Complementary log-log regression</td>
</tr>
<tr>
<td><code>binreg</code></td>
<td>[R] <code>binreg</code></td>
<td>GLM for the binomial family</td>
</tr>
<tr>
<td>Count-response regression models</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>poisson</code></td>
<td>[R] <code>poisson</code></td>
<td>Poisson regression</td>
</tr>
<tr>
<td><code>nbreg</code></td>
<td>[R] <code>nbreg</code></td>
<td>Negative binomial regression</td>
</tr>
<tr>
<td><code>gnbreg</code></td>
<td>[R] <code>gnbreg</code></td>
<td>Generalized negative binomial regression</td>
</tr>
<tr>
<td>Ordinal-response regression models</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>ologit</code></td>
<td>[R] <code>ologit</code></td>
<td>Ordered logistic regression</td>
</tr>
<tr>
<td><code>oprobit</code></td>
<td>[R] <code>oprobit</code></td>
<td>Ordered probit regression</td>
</tr>
<tr>
<td>Categorical-response regression models</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>mlogit</code></td>
<td>[R] <code>mlogit</code></td>
<td>Multinomial (polytomous) logistic regression</td>
</tr>
<tr>
<td><code>mprobit</code></td>
<td>[R] <code>mprobit</code></td>
<td>Multinomial probit regression</td>
</tr>
<tr>
<td><code>clogit</code></td>
<td>[R] <code>clogit</code></td>
<td>Conditional (fixed-effects) logistic regression</td>
</tr>
</tbody>
</table>
Fractional-response regression models
fracreg \[R\] fracreg Fractional response regression

Quantile regression models
qreg \[R\] qreg Quantile regression
iqreg \[R\] qreg Interquantile range regression
sqreg \[R\] qreg Simultaneous-quantile regression
bsqreg \[R\] qreg Bootstrapped quantile regression

Survival regression models
stcox \[ST\] stcox Cox proportional hazards model
streg \[ST\] streg Parametric survival models
stcrreg \[ST\] stcrreg Competing-risks regression

Other regression models
glm \[R\] glm Generalized linear models
areg \[R\] areg Linear regression with a large dummy-variable set
rreg \[R\] rreg Robust regression
truncreg \[R\] truncreg Truncated regression

Descriptive statistics
mean \[R\] mean Estimate means
proportion \[R\] proportion Estimate proportions
ratio \[R\] ratio Estimate ratios
total \[R\] total Estimate totals

Panel-data models
xtreg \[XT\] xtreg Fixed-, between- and random-effects, and population-averaged linear models
xtrc \[XT\] xtrc Random-coefficients model
xtlogit \[XT\] xttlogit Fixed-effects, random-effects, and population-averaged logit models
xtprobit \[XT\] xttprobit Random-effects and population-averaged probit models
xtcloglog \[XT\] xttcloglog Random-effects and population-averaged cloglog models
xtpoisson \[XT\] xttpoisson Fixed-effects, random-effects, and population-averaged Poisson models
xtnbreg \[XT\] xtnbreg Fixed-effects, random-effects, and population-averaged negative binomial models
xtgee \[XT\] xtgee Fit population-averaged panel-data models by using GEE

Multilevel mixed-effects models
meqrlogit \[ME\] meqrlogit Multilevel mixed-effects logistic regression (QR decomposition)
meqrpoisson \[ME\] meqrpoisson Multilevel mixed-effects Poisson regression (QR decomposition)
mixed \[ME\] mixed Multilevel mixed-effects linear regression

Survey regression models
svy: \[SVY\] svy Estimation commands for survey data (excluding commands that are not listed above)
Also see

[MI] mi estimate — Estimation using multiple imputations
[MI] mi estimate postestimation — Postestimation tools for mi estimate
[MI] mi import — Import data into mi
[MI] mi impute — Impute missing values
[MI] mi set — Declare multiple-imputation data
[MI] workflow — Suggested workflow
[MI] intro — Introduction to mi
[MI] intro substantive — Introduction to multiple-imputation analysis
[MI] Glossary