New in

Lasso—variable selection, prediction, inference

All the tools you expect for lasso machine learning
- Lasso, square-root lasso, and elastic net
- Cross-validation
- Adaptive lasso
- Knot analysis
- Coefficient paths

Alongside cutting-edge inferential methods
- Robust to mistakes in variable selection
- Proper inference for coefficients of interest
- Double selection
- Partialing out
- Cross-fit partialing out
- Double machine learning

Select predictors for continuous, binary, and count outcomes

Lasso with selection via cross-validation
- `lasso linear y x1-x1000`
- `lasso logit y x1-x1000`
- `lasso probit y x1-x1000`
- `lasso poisson y x1-x1000`

Adaptive lasso
- `lasso linear y x1-x1000, selection(adaptive)`

Selection via plugin method
- `lasso linear y x1-x1000, selection(plugin)`

Elastic net with selection via cross-validation
- `elasticnet linear y x1-x1000`
- `elasticnet logit y x1-x1000`
- `elasticnet probit y x1-x1000`
- `elasticnet poisson y x1-x1000`

Square-root lasso
- `sqrtlasso y x1-x1000`

Examine the results

View selected variables
- `lassoknots`
- `lassoinfo`
- `lassocoef`

Plot cross-validation function
- `cvplot`

Plot coefficient path
- `coefpath`

Obtain predictions
- `use newdata`
- `predict yhat`

Evaluate fit
- `lassogof`
Lasso for inference

With lasso inferential methods, you can estimate coefficients, standard errors, test statistics, and confidence intervals for variables of interest while using lassos to select from a potentially large number of control variables.

Double-selection method; estimate coefficients for \( x_1 \) and categorical \( x_2 \); selection of controls via plugin
\[
\text{dsregress } y \ x_1 \ i.x_2, \text{controls}(c1-c1000)
\]

Logit model for binary outcome; estimate odds ratios for \( x_1 \) and \( x_2 \)
\[
\text{dslogit } y \ x_1 \ i.x_2, \text{controls}(c1-c1000)
\]

Poisson model for count outcome; estimate incidence-rate ratios for \( x_1 \) and \( x_2 \)
\[
\text{dspoisson } y \ x_1 \ i.x_2, \text{controls}(c1-c1000)
\]

Selection of controls via cross-validation
\[
\text{dsregress } y \ x_1 \ i.x_2, \text{controls}(c1-c1000) \\text{ selection(cv)}
\]

Partialing-out method
\[
\text{poregress } y \ x_1 \ i.x_2, \text{controls}(c1-c1000)
\]

Cross-fit partialing-out method (double machine learning)
\[
\text{xporegress poregress } y \ x_1 \ i.x_2, \text{controls}(c1-c1000)
\]

Evaluate results using Stata’s standard tools

Perform tests on coefficients
\[
\text{test } x_1=1
\]

Estimate contrasts such as differences across levels
\[
\text{contrast ar.x2}
\]

Explore underlying lassos

View selected controls in the lasso for \( y \)
\[
\text{lassocoef\ (\., for(y))}
\]

Plot coefficient path in the lasso for \( y \)
\[
\text{coefpath, for(y)}
\]