Statistics/Data Analysis

Title

runmixregls —— Run the MIXREGLS mixed-effects location scale software from within Stata

Syntax

runmixregls *depvar* [*varlist*] [*if*] [*in*] [, *options*]

where *varlist* specifies variables in the mean function.

options	Description
Model	
noconstant	suppress constant term in the mean function
between(*varlist* [, **noconstant**]**)**	specify variables in between-group variance function
within(*varlist* [, **noconstant**]**)**	specify variables in within-group variance function
association(*atype***)**	specify the group-level association between the (log of the) within-group variance and the random-location effects; default is **association(linear)**
Random effects/Residuals	
reffects(*newvar1 newvar2***)**	retrieve standardized random-location and random-scale effects
residuals(*newvar***)**	retrieve standardized residual errors
Integration	
noadapt	do not perform adaptive Gaussian quadrature
intpoints(#)	set the number of integration (quadrature) points; default is **intpoints(11)**
Maximization	
iterate(#)	maximum number of iterations; default is **iterate(200)**
tolerance(#)	tolerance; default is **tolerance(0.0005)**
standardize	standardize all covariates
ridgein(#)	initial value for ridge; default is **ridgein(0)**
Reporting	
level(#)	set confidence level; default is **level(95)**
display options	control column formats, row spacing, line width, and display of omitted variables and base and empty cells
noheader	suppress table header
notable	suppress coefficient table
coeflegend	display legend instead of statistics

atype	Description
none	none
linear	linear; the default
quadratic	quadratic

A panel variable must be specified. Use **xtset**.

Description

runmixregls runs the MIXREGLS mixed-effects location scale software (Hedeker and Nordgren 2013) from within Stata.

The mixed-effects location scale model extends the standard two-level random-intercept mixed-effects model for continuous responses (**xtreg, mle**) in three ways.

(1) The (log of the) within- and between-group variances are further modeled as functions of the covariates.

(2) A new random effect, referred to as the random-scale effect, is then entered into the within-group variance function to account for any unexplained group differences in the residual variance. The existing random-intercept effect is now referred to as the random-location effect.

(3) A group-level association between the location and the scale may be allowed for by entering the random-location effect into the within-group variance function using either a linear or quadratic functional form. The regression coefficients of these linear and quadratic terms are then estimated.

The distributions of the random-location and random-scale effects are assumed to be Gaussian.

Options

Model

noconstant; see **[R] estimation options**.

between(*varlist* [, **noconstant**]**)** specifies the variables in the between-group variance function.

within(*varlist* [, **noconstant**]**)** specifies the variables in the within-group variance function.

association(*atype***)**, where *atype* is

none|**linear**|**quadratic**

specifies the group-level association between the (log of the) within-group variance and the random-location effects. The default is **association(linear)**.

Random effects/Residuals

reffects(*newvar1 newvar2***)** retrieves the best linear unbiased predictions (BLUPs) of the standardized random effects from MIXREGLS. BLUPs are also known as empirical Bayes estimates. The standardized random-location effects are placed in *newvar1* while the standardized random-scale effects are placed in *newvar2* The associated standard errors are placed in *newvar1*_se and *newvar2*_se.

residuals(*newvar***)** retrieves the standardized residual errors from MIXREGLS.

Integration

noadapt prevents MIXREGLS from using adaptive Gaussian quadrature. MIXREGLS will use ordinary Gaussian quadrature instead.

intpoints(#) sets the number of integration points for (adaptive) Gaussian quadrature. The default is **intpoints(11)**. The more points, the more accurate the approximation to the log likelihood. However, computation time increases with the number of quadrature points. When models do not converge properly, increasing the number of quadrature points can sometimes lead to convergence.

Maximization

iterate(#) specifies the maximum number of iterations. The default is **iterate(200)**. You should seldom have to use this option.

tolerance(#) specifies the convergence tolerance. The default is **tolerance(0.0005)**. You should seldom have to use this option.

standardize standardizes all covariates in all functions during optimization. This ensures all covariates are on the same numerical scale with mean 0 and variance 1. This can be helpful if the model "blows up" or does not converge to the solution.

ridgein(#) specifies the initial value for the ridge parameter. The default is **ridgein(200)**. This is a numeric value that adds to the diagonal of the second derivative matrix, which can aid in convergence of the solution; usually set to 0 or some small fractional value.

Reporting

level(#); see **[R] estimation options**.

display_options : **noomitted**, **vsquish**, **noemptycells**, **baselevels**, **allbaselevels**, **cformat(%*fmt*)**, **pformat(%*fmt*)**, **sformat(%*fmt*)**, and **nolstretch**; see **[R] estimation options**.

noheader suppresses the display of the summary statistics at the top of the output; only the coefficient table is displayed.

notable suppresses display of the coefficient table.

coeflegend; see **[R] estimation options**.

Remarks

Remarks are presented under the following headings:

Remarks on the mixed-effects location scale model
Remarks on getting runmixregls working for the first time
Remarks on MIXREGLS estimation
Remarks on runmixregls output

Remarks on the mixed-effects location scale model

The mixed-effects location scale model fitted by MIXREGLS consists of three functions

(1) the mean function,

(2) the between-group variance function,

(3) the within-group variance function.

These three functions can be written as

```
y_ij = b_1*x1_ij + sigma_u_ij*theta1_j +
e_ij,   i=1,...,n_j;   j=1,...,J,

log(sigma2_u_ij) = b_2*x2_ij,
```

log(sigma2_e_ij) = b_3*x3_ij + a_l*theta1_j + a_q*theta1_j^2 + sigma_v*theta2_j,

where

theta1_j ~ N(0,1),

theta2_j ~ N(0,1),

e_ij ~ N(0,sigma2_e_ij),

and

y_ij is the continuous response variable,

x1_ij, x2_ij and x3_ij are vectors of observation- and group-level covariates,

b_1, b_2 and b_3 are vectors of parameters to be estimated,

theta1_j are the unobserved standardized random-location effects,

theta2_j are the unobserved standardized random-scale effects,

a_l and a_q are scalar parameters to be estimated,

e_ij are the observation-specific errors.

See Hedeker and Nordgren 2013 for further details on the mixed-effects location scale model.

Remarks on getting runmixregls working for the first time

runmixregls can be installed from the Statistical Software Components (SSC) archive by typing the following from a net-aware version of Stata

```
. ssc install runmixregls
```

If you have already installed **runmixregls** from the SSC, you can check that you are using the latest version by typing the following command:

```
. adoupdate runmixregls
```

Remarks on MIXREGLS estimation

MIXREGLS uses maximum likelihood estimation, utilizing both the EM algorithm and a Newton-Raphson solution. Because the log likelihood for this model has no closed form, it is approximated by adaptive Gaussian quadrature. Estimation of the random effects is accomplished using empirical Bayes methods. The full model is estimated in three sequential stages:

(1) Standard random-intercept model + between-group variance function regression coefficients

(2) Stage 1 model + within-group variance function regression coefficients

(3) Stage 2 model + group-level association between the (log of the) within-group variance and the random-location effects + random-scale effects

Prior to Stage 1, 20 iterations are performed of the EM algorithm to estimate the parameters of a standard random-intercept model (regression coefficients, between-group variance, within-group variance, and random-location effects). These estimates are then used as starting values for Stage 1 Estimates at each stage are used as starting values for the next stage, which improves the convergence of the final model. This also provides a way of assessing the statistical significance of the additional parameters in each stage via likelihood-ratio tests. The results of each stage as well as these likelihood-ratio tests are provided in the saved results.

See Hedeker and Nordgren 2013 for further details on the MIXREGLS estimation.

Remarks on runmixregls output

The **runmixregls** output displays five different sets of parameters

Mean: Mean function regression coefficients

Between: Between-group variance function regression coefficients (log scale)

Within: Within-group variance function regression coefficients (log scale)

Association: Group-level association parameters between the (log of the) within-group function and the random-location effects

Scale: Random-scale standard deviation

Example: Replicate Hedeker and Nordgren 2013 (pages 10-18)

Load the data

```
. use http://www.bristol.ac.uk/cmm/media/runmixregls/reisby, clear
```

Recode missing values in hamdep from -9 to Stata system missing

```
. recode hamdep (-9 = .)
```

Declare panel variable to be id

```
. xtset id
```

Fit the mixed-effects location scale model

```
. runmixregls hamdep week endog endweek, between(endog) within(week endog)
```

Refit the model, this time retrieving the BLUPs of the standardized random-location and random-scale effects, and their associated standard errors

```
. runmixregls hamdep week endog endweek, between(endog) within(week endog) reffects(theta1 theta2)
```

Examine a scatter plot of the BLUPs of the standardized random-scale effects against the standardized random-location effects

```
. scatter theta2 theta1
```

Refit the model removing the group-level linear association between the (log of the) within-group variance and the intercept

```
. runmixregls hamdep week endog endweek, between(endog) within(week endog) association(none)
```

Saved results

runmixregls saves the following in e():

Scalars

e(N)	number of observations
e(N_g)	number of groups
e(g_min)	smallest group size
e(g_avg)	average group size
e(g_max)	largest group size
e(k)	number of parameters
e(k_1)	number of parameters, stage 1 model
e(k_2)	number of parameters, stage 2 model
e(k_3)	number of parameters, stage 3 model
e(ll)	log likelihood
e(ll_1)	log likelihood, stage 1 model
e(ll_2)	log likelihood, stage 2 model
e(ll_3)	log likelihood, stage 3 model
e(deviance_1)	deviance, stage 1 model
e(deviance_2)	deviance, stage 2 model
e(deviance_3)	deviance, stage 3 model
e(iterations)	number of iterations
e(iterations_1)	number of iterations, stage 1 model
e(iterations_2)	number of iterations, stage 2 model
e(iterations_3)	number of iterations, stage 3 model
e(time)	estimation time (seconds)
e(chi2_1vs2)	chi-squared, stage 1 model vs. stage 2 model
e(chi2_1vs3)	chi-squared, stage 1 model vs. stage 3 model
e(chi2_2vs3)	chi-squared, stage 2 model vs. stage 3 model
e(p_1vs2)	p-value, stage 1 model vs. stage 2 model
e(p_1vs3)	p-value, stage 1 model vs. stage 3 model
e(p_2vs3)	p-value, stage 2 model vs. stage 3 model

Macros

e(cmd)	runmixregls
e(cmdline)	command as typed
e(depvar)	name of dependent variable
e(ivar)	variable denoting groups
e(adapt)	adaptive Gaussian quadrature
e(n_quad)	number of integration points
e(iterate)	maximum number of iterations
e(tolerance)	tolerance
e(ridgein)	initial ridge
e(standardize)	standardized variables
e(properties)	b V

Matrices

e(b)	coefficient vector
e(V)	variance-covariance matrix of the estimators
e(b_1)	coefficient vector, stage 1 model
e(b_2)	coefficient vector, stage 2 model
e(b_3)	coefficient vector, stage 3 model
e(V_1)	variance-covariance matrix of the estimators, stage 1 model
e(V_2)	variance-covariance matrix of the estimators, stage 2 model
e(V_3)	variance-covariance matrix of the estimators, stage 3 model

Functions

e(sample)	marks estimation sample

Citation of runmixregls and MIXREGLS

runmixregls is not an official Stata command. It is a free contribution to the research community, like a paper. Please cite it as such:

Leckie, G. 2014. **runmixregls** - A Program to Run the MIXREGLS Mixed-effects Location Scale Software from within Stata. *Journal of Statistical Software* , 59 (Code Snippet 2), 1-41. URL: http://www.jstatsoft.org/v59/c02.

Similarly, please also cite the MIXREGLS software:

Hedeker, D. R. and Nordgren. 2013. MIXREGLS: A Program for Mixed-effects Location Scale Analysis. *Journal of Statistical Software* , 52, 12, 1-38. URL: http://www.jstatsoft.org/v52/i12.

Authors

George Leckie
Centre for Multilevel Modelling
University of Bristol
g.leckie@bristol.ac.uk
http://www.bristol.ac.uk/cmm/team/leckie.html

Chris Charlton
Centre for Multilevel Modelling
University of Bristol

Acknowledgments

The development of this command was funded under the LEMMA3 project, a node of the UK Economic and Social Research Council's National Centre for Research Methods (grant number RES-576-25-0035).

Disclaimer

runmixregls comes with no warranty.

References

Hedeker, D. R. and Nordgren. 2013. MIXREGLS: A Program for Mixed-effects Location Scale Analysis. *Journal of Statistical Software* , 52, 12, 1-38. URL: http://www.jstatsoft.org/v52/i12.

Also see

Manual: **[XT] xtreg [ME] mixed**

Online: **[XT] xtreg**, **[ME] mixed**, **runmlwin**, **gllamm**