Title

estat wcorrelation — Display within-cluster correlations and standard deviations

Description

`estat wcorrelation` is for use after estimation with `menl` and `mixed`.

`estat wcorrelation` displays the overall correlation matrix for a given cluster calculated on the basis of the design of the random effects and their assumed covariance and the correlation structure of the residuals. This allows for a comparison of different multilevel models in terms of the ultimate within-cluster correlation matrix that each model implies.

Menu for estat

Statistics > Postestimation

Syntax

`estat wcorrelation` [, *options*]

options	Description
`at(`*at_spec*`)`	specify the cluster for which you want the correlation matrix; default is the first two-level cluster encountered in the data
`all`	display correlation matrix for all the data
`covariance`	display the covariance matrix instead of the correlation matrix
`list`	list the data corresponding to the correlation matrix
`nosort`	list the rows and columns of the correlation matrix in the order they were originally present in the data
`iterate(`*#*`)`	maximum number of iterations to compute random effects; default is `iterate(50)`; only for use after `menl`
`tolerance(`*#*`)`	convergence tolerance when computing random effects; default is `tolerance(1e-4)`; only for use after `menl`
`format(`*%fmt*`)`	set the display format; default is `format(%6.3f)`
matlist_options	style and formatting options that control how matrices are displayed

Options

`at(`*at_spec*`)` specifies the cluster of observations for which you want the within-cluster correlation matrix. *at_spec* is

relevel_var = *value* [, *relevel_var* = *value* ...]

For example, if you specify

```
. estat wcorrelation, at(school = 33)
```

you get the within-cluster correlation matrix for those observations in school 33. If you specify

```
. estat wcorrelation, at(school = 33 classroom = 4)
```

you get the correlation matrix for classroom 4 in school 33.

If `at()` is not specified, then you get the correlations for the first level-two cluster encountered in the data. This is usually what you want.

`all` specifies that you want the correlation matrix for all the data. This is not recommended unless you have a relatively small dataset or you enjoy seeing large $n \times n$ matrices. However, this can prove useful in some cases.

`covariance` specifies that the within-cluster covariance matrix be displayed instead of the default correlations and standard deviations.

`list` lists the model data for those observations depicted in the displayed correlation matrix. With linear mixed-effects models, this option is also useful if you have many random-effects design variables and you wish to see the represented values of these design variables.

`nosort` lists the rows and columns of the correlation matrix in the order that they were originally present in the data. Normally, `estat wcorrelation` will first sort the data according to level variables, by-group variables, and time variables to produce correlation matrices whose rows and columns follow a natural ordering. `nosort` suppresses this.

`iterate(`*#*`)` specifies the maximum number of iterations when computing estimates of the random effects. The default is `iterate(50)`. This option is only for use after `menl`.

`tolerance(`*#*`)` specifies a convergence tolerance when computing estimates of the random effects. The default is `tolerance(1e-4)`. This option is only for use after `menl`.

`format(`*%fmt*`)` sets the display format for the standard-deviation vector and correlation matrix. The default is `format(%6.3f)`.

matlist_options are style and formatting options that control how the matrix (or matrices) is displayed; see [P] **matlist** for a list of options that are available.

Remarks and examples

stata.com

▷ Example 1: Displaying within-cluster correlations for different clusters

Here we fit a model where different clusters have different within-cluster correlations, and we show how to display them for different clusters. We use the Asian children weight data from example 6 of [ME] **mixed**.

```
. use http://www.stata-press.com/data/r15/childweight
(Weight data on Asian children)

. mixed weight age || id: age, covariance(unstructured)

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0:   log likelihood = -344.37065
Iteration 1:   log likelihood = -342.83887
Iteration 2:   log likelihood = -342.71863
Iteration 3:   log likelihood = -342.71777
Iteration 4:   log likelihood = -342.71777

Computing standard errors:

Mixed-effects ML regression                     Number of obs     =        198
Group variable: id                              Number of groups  =         68

                                                Obs per group:
                                                              min =          1
                                                              avg =        2.9
                                                              max =          5

                                                Wald chi2(1)      =     755.27
Log likelihood = -342.71777                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
      weight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   3.459671   .1258877    27.48   0.000     3.212936    3.706406
       _cons |   5.110496   .1494781    34.19   0.000     4.817524    5.403468
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                    var(age) |    .202392   .1242868      .0607406    .6743838
                  var(_cons) |   .0970272   .1107998      .0103483    .9097447
              cov(age,_cons) |    .140134   .0566901      .0290234    .2512445
-----------------------------+------------------------------------------------
               var(Residual) |   1.357922   .1650502      1.070076    1.723198
------------------------------------------------------------------------------
LR test vs. linear model: chi2(3) = 27.38                 Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

We use `estat wcorrelation` to display the within-cluster correlations for the first cluster.

```
. estat wcorrelation, list
Standard deviations and correlations for id = 45:
Standard deviations:

         obs |     1       2       3       4       5
-------------+----------------------------------------
          sd | 1.224   1.314   1.448   1.506   1.771
Correlations:

         obs |     1       2       3       4       5
-------------+----------------------------------------
           1 | 1.000
           2 | 0.141   1.000
           3 | 0.181   0.274   1.000
           4 | 0.193   0.293   0.376   1.000
           5 | 0.230   0.348   0.447   0.477   1.000
Data:

     +-------------------------+
     | id   weight         age |
     |-------------------------|
  1. | 45    5.171     .136893 |
  2. | 45    10.86     .657084 |
  3. | 45    13.15     1.21834 |
  4. | 45     13.2     1.42916 |
  5. | 45    15.88     2.27242 |
     +-------------------------+
```

We specified the `list` option to display the data associated with the cluster. The next cluster in the dataset has ID 258. To display the within-cluster correlations for this cluster, we specify the `at()` option.

```
. estat wcorrelation, at(id=258) list
Standard deviations and correlations for id = 258:
Standard deviations:

         obs |     1       2       3       4
-------------+--------------------------------
          sd | 1.231   1.320   1.424   1.782
Correlations:

         obs |     1       2       3       4
-------------+--------------------------------
           1 | 1.000
           2 | 0.152   1.000
           3 | 0.186   0.270   1.000
           4 | 0.244   0.356   0.435   1.000
Data:

     +---------------------------+
     |  id   weight          age |
     |---------------------------|
  1. | 258      5.3       .19165 |
  2. | 258     9.74      .687201 |
  3. | 258     9.98      1.12799 |
  4. | 258    11.34      2.30527 |
     +---------------------------+
```

The within-cluster correlations for this model depend on age. The values for `age` in the two clusters are different, as are the corresponding within-cluster correlations.

◁

See example 1 of [ME] **mixed postestimation** for a model fit where each cluster had the same model-implied within-cluster correlations.

Stored results

`estat wcorrelation` stores the following in `r()`:

Matrices

`r(sd)`	standard deviations
`r(Corr)`	within-cluster correlation matrix
`r(Cov)`	within-cluster variance–covariance matrix
`r(G)`	variance–covariance matrix of random effects
`r(Z)`	model-based design matrix
`r(R)`	variance–covariance matrix of level-one errors

Results `r(G)`, `r(Z)`, and `r(R)` are available only after `mixed`.

Methods and formulas

Methods and formulas are presented under the following headings:

Linear mixed-effects model
Nonlinear mixed-effects model

Linear mixed-effects model

A two-level linear mixed model of the form

$$\mathbf{y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{u}_j + \boldsymbol{\epsilon}_j$$

implies the marginal model

$$\mathbf{y}_j = \mathbf{X}_j\boldsymbol{\beta} + \boldsymbol{\epsilon}_j^*$$

where $\boldsymbol{\epsilon}_j^* \sim N(\mathbf{0}, \mathbf{V}_j)$, $\mathbf{V}_j = \mathbf{Z}_j\mathbf{G}\mathbf{Z}_j' + \mathbf{R}$. In a marginal model, the random part is described in terms of the marginal or total residuals $\boldsymbol{\epsilon}_j^*$, and $\mathbf{V}_j$ is the covariance structure of these residuals.

`estat wcorrelation` calculates the marginal covariance matrix $\widetilde{\mathbf{V}}_j$ for cluster j and by default displays the results in terms of standard deviations and correlations. This allows for a comparison of different multilevel models in terms of the ultimate within-cluster correlation matrix that each model implies.

Calculation of the marginal covariance matrix extends naturally to higher-level models; see, for example, chapter 4.8 in West, Welch, and Gałecki (2015).

Nonlinear mixed-effects model

For nonlinear mixed-effects models, there is no closed-form expression for the marginal covariance matrix $\text{Cov}(\mathbf{y}_j)$. This is because it is expressed in terms of a q-dimensional integral (q is the number of random effects in the model), which, in general, is analytically intractable. Under the linear mixed-effects approximation, the marginal covariance matrix is estimated by $\widehat{\mathbf{V}}_j = \widehat{\mathbf{Z}}_j\widehat{\boldsymbol{\Sigma}}\widehat{\mathbf{Z}}_j' + \widehat{\sigma}^2\widehat{\boldsymbol{\Lambda}}_j$, where $\widehat{\mathbf{Z}}_j$, $\widehat{\boldsymbol{\Sigma}}$, and $\widehat{\boldsymbol{\Lambda}}_j$ are defined in *Methods and formulas* of [ME] **menl**.

estat wcorrelation calculates the estimated marginal covariance matrix $\widehat{\mathbf{V}}_j$ for cluster j and by default displays the results in terms of standard deviations and correlations.

Under the linear mixed-effects approximation, estimation of the marginal covariance matrix extends naturally to higher-level models; see, for example, chapter 4.8 in West, Welch, and Gałecki (2015).

Reference

West, B. T., K. B. Welch, and A. T. Gałecki. 2015. *Linear Mixed Models: A Practical Guide Using Statistical Software*. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC.

Also see

[ME] **menl** — Nonlinear mixed-effects regression

[ME] **mixed** — Multilevel mixed-effects linear regression

[U] **20 Estimation and postestimation commands**