

meqrlogit postestimation — Postestimation tools for meqrlogit

Postestimation commands	predict	margins
estat	Remarks and examples	Stored results
Methods and formulas	References	Also see

Postestimation commands

The following postestimation commands are of special interest after meqrlogit:

Command	Description
estat group	summarize the composition of the nested groups
estat recovariance	display the estimated random-effects covariance matrix (or matrices)
estat icc	estimate intraclass correlations

The following standard postestimation commands are also available:

Command	Description
contrast	contrasts and ANOVA-style joint tests of estimates
estat ic	Akaike's and Schwarz's Bayesian information criteria (AIC and BIC)
estat summarize	summary statistics for the estimation sample
estat vce	variance–covariance matrix of the estimators (VCE)
estimates	cataloging estimation results
hausman	Hausman's specification test
lincom	point estimates, standard errors, testing, and inference for linear combinations of coefficients
lrtest	likelihood-ratio test
margins	marginal means, predictive margins, marginal effects, and average marginal effects
marginsplot	graph the results from margins (profile plots, interaction plots, etc.)
nlcom	point estimates, standard errors, testing, and inference for nonlinear combinations of coefficients
predict	predictions, residuals, influence statistics, and other diagnostic measures
predictnl	point estimates, standard errors, testing, and inference for generalized predictions
pwcompare	pairwise comparisons of estimates
test	Wald tests of simple and composite linear hypotheses
testnl	Wald tests of nonlinear hypotheses

predict

Description for predict

`predict` creates a new variable containing predictions such as mean responses; linear predictions; standard errors; and Pearson, deviance, and Anscombe residuals.

Menu for predict

Statistics > Postestimation

Syntax for predict

Syntax for obtaining estimated random effects and their standard errors

```
predict [type] newvarsspec [if] [in], reflects [reses(newvarsspec)
    relevel(levelvar) ]
```

Syntax for obtaining other predictions

```
predict [type] newvar [if] [in] [, statistic nooffset fixedonly]
```

newvarsspec is *stub** or *newvarlist*.

<i>statistic</i>	Description
------------------	-------------

Main	
<u>mu</u>	mean response; the default
<u>xb</u>	linear predictor for the fixed portion of the model only
<u>stdp</u>	standard error of the fixed-portion linear prediction
<u>pearson</u>	Pearson residuals
<u>deviance</u>	deviance residuals
<u>anscombe</u>	Anscombe residuals

These statistics are available both in and out of sample; type `predict ... if e(sample) ... if` wanted only for the estimation sample.

Options for predict

Main

`reflects` calculates posterior modal estimates of the random effects. By default, estimates for all random effects in the model are calculated. However, if the `relevel(levelvar)` option is specified, then estimates for only level *levelvar* in the model are calculated. For example, if `classes` are nested within `schools`, then typing

```
. predict b*, reflects relevel(school)
```

would yield random-effects estimates at the school level. You must specify *q* new variables, where *q* is the number of random-effects terms in the model (or level). However, it is much easier to just specify *stub** and let Stata name the variables *stub1*, *stub2*, ..., *stubq* for you.

`reses` (*newvarsspec*) calculates standard errors for the random-effects estimates. By default, standard errors for all random effects in the model are calculated. However, if the `relevel`(*levelvar*) option is specified, then standard errors for only level *levelvar* in the model are calculated; see the `reffects` option.

You must specify q new variables, where q is the number of random-effects terms in the model (or level). However, it is much easier to just specify *stub** and let Stata name the variables *stub1*, *stub2*, ..., *stubq* for you. The new variables will have the same storage type as the corresponding random-effects variables.

The `reffects` and `reses`() options often generate multiple new variables at once. When this occurs, the random effects (or standard errors) contained in the generated variables correspond to the order in which the variance components are listed in the output of `meqrlogit`. Still, examining the variable labels of the generated variables (with the `describe` command, for instance) can be useful in deciphering which variables correspond to which terms in the model.

`relevel`(*levelvar*) specifies the level in the model at which predictions for random effects and their standard errors are to be obtained. *levelvar* is the name of the model level and is either the name of the variable describing the grouping at that level or is `_all`, a special designation for a group comprising all the estimation data.

`mu`, the default, calculates the predicted mean. By default, this is based on a linear predictor that includes both the fixed effects and the random effects, and the predicted mean is conditional on the values of the random effects. Use the `fixedonly` option (see [below](#)) if you want predictions that include only the fixed portion of the model, that is, if you want random effects set to 0.

`xb` calculates the linear prediction $\mathbf{x}\beta$ based on the estimated fixed effects (coefficients) in the model. This is equivalent to fixing all random effects in the model to their theoretical (prior) mean value of 0.

`stdp` calculates the standard error of the fixed-effects linear predictor $\mathbf{x}\beta$.

`pearson` calculates Pearson residuals. Pearson residuals large in absolute value may indicate a lack of fit. By default, residuals include both the fixed portion and the random portion of the model. The `fixedonly` option modifies the calculation to include the fixed portion only.

`deviance` calculates deviance residuals. Deviance residuals are recommended by [McCullagh and Nelder \(1989\)](#) as having the best properties for examining the goodness of fit of a GLM. They are approximately normally distributed if the model is correctly specified. They may be plotted against the fitted values or against a covariate to inspect the model's fit. By default, residuals include both the fixed portion and the random portion of the model. The `fixedonly` option modifies the calculation to include the fixed portion only.

`anscombe` calculates Anscombe residuals, which are designed to closely follow a normal distribution. By default, residuals include both the fixed portion and the random portion of the model. The `fixedonly` option modifies the calculation to include the fixed portion only.

`nooffset` is relevant only if you specified `offset`(*varname*) with `meqrlogit`. It modifies the calculations made by `predict` so that they ignore the offset variable; the linear prediction is treated as $\mathbf{X}\beta + \mathbf{Z}\mathbf{u}$ rather than $\mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \text{offset}$.

`fixedonly` modifies predictions to include only the fixed portion of the model, equivalent to setting all random effects equal to 0; see the `mu` option.

margins

Description for margins

`margins` estimates margins of response for linear predictions.

Menu for margins

Statistics > Postestimation

Syntax for margins

```
margins [marginlist] [, options]
margins [marginlist] , predict(statistic ...) [options]
```

<i>statistic</i>	Description
<code>xb</code>	linear predictor for the fixed portion of the model only; the default
<code><u>reflects</u></code>	not allowed with <code>margins</code>
<code>mu</code>	not allowed with <code>margins</code>
<code>stdp</code>	not allowed with <code>margins</code>
<code><u>pearson</u></code>	not allowed with <code>margins</code>
<code><u>deviance</u></code>	not allowed with <code>margins</code>
<code><u>anscombe</u></code>	not allowed with <code>margins</code>

Statistics not allowed with `margins` are functions of stochastic quantities other than $e(b)$.

For the full syntax, see [R] [margins](#).

estat

Description for estat

`estat group` reports the number of groups and minimum, average, and maximum group sizes for each level of the model. Model levels are identified by the corresponding group variable in the data. Because groups are treated as nested, the information in this summary may differ from what you would get if you used the `tabulate` command on each group variable individually.

`estat recovariance` displays the estimated variance–covariance matrix of the random effects for each level in the model. Random effects can be either random intercepts, in which case the corresponding rows and columns of the matrix are labeled as `_cons`, or random coefficients, in which case the label is the name of the associated variable in the data.

`estat icc` displays the intraclass correlation for pairs of latent linear responses at each nested level of the model. Intraclass correlations are available for random-intercept models or for random-coefficient models conditional on random-effects covariates being equal to 0. They are not available for crossed-effects models.

Menu for estat

Statistics > Postestimation

Syntax for estat

Summarize the composition of the nested groups

```
estat group
```

Display the estimated random-effects covariance matrix (or matrices)

```
estat recovariance [ , relevel(levelvar) correlation matlist_options ]
```

Estimate intraclass correlations

```
estat icc [ , level(#) ]
```

Options for estat recovariance

`relevel(levelvar)` specifies the level in the model for which the random-effects covariance matrix is to be displayed and returned in `r(cov)`. By default, the covariance matrices for all levels in the model are displayed. *levelvar* is the name of the model level and is either the name of the variable describing the grouping at that level or is `_all`, a special designation for a group comprising all the estimation data.

`correlation` displays the covariance matrix as a correlation matrix and returns the correlation matrix in `r(corr)`.

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.

Option for estat icc

`level(#)` specifies the confidence level, as a percentage, for confidence intervals. The default is `level(95)` or as set by `set level`; see [U] [20.7 Specifying the width of confidence intervals](#).

Remarks and examples

Various predictions, statistics, and diagnostic measures are available after fitting a logistic mixed-effects model with `meqrlogit`. For the most part, calculation centers around obtaining estimates of the subject/group-specific random effects. Random effects are not provided as estimates when the model is fit but instead need to be predicted after estimation. Calculation of intraclass correlations, estimating the dependence between latent linear responses for different levels of nesting, may also be of interest.

► Example 1

Following [Rabe-Hesketh and Skrondal \(2012, chap. 10\)](#), we consider a two-level mixed-effects model for onycholysis (separation of toenail plate from nail bed) among those who contract toenail fungus. The data are obtained from [De Backer et al. \(1998\)](#) and were also studied by [Lesaffre and Spiessens \(2001\)](#). The onycholysis outcome is dichotomously coded as 1 (moderate or severe onycholysis) or 0 (none or mild onycholysis). Fixed-effects covariates include treatment (0: itraconazole; 1: terbinafine), the month of measurement, and their interaction.

We fit the two-level model with `meqrlogit`:

```
. use http://www.stata-press.com/data/r14/toenail
(Onycholysis severity)
. meqrlogit outcome treatment month trt_month || patient:, intpoints(30)

Refining starting values:
Iteration 0:  log likelihood = -749.95893
Iteration 1:  log likelihood = -630.0759
Iteration 2:  log likelihood = -625.69415

Performing gradient-based optimization:
Iteration 0:  log likelihood = -625.69415
Iteration 1:  log likelihood = -625.39739
Iteration 2:  log likelihood = -625.39709
Iteration 3:  log likelihood = -625.39709

Mixed-effects logistic regression      Number of obs    =    1,908
Group variable: patient                Number of groups =     294

Obs per group:
      min =         1
      avg =         6.5
      max =         7

Integration points = 30                Wald chi2(3)     =    150.52
Log likelihood = -625.39709           Prob > chi2      =     0.0000
```

outcome	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
treatment	-.1609377	.584208	-0.28	0.783	-1.305964	.984089
month	-.3910603	.0443957	-8.81	0.000	-.4780744	-.3040463
trt_month	-.1368073	.0680235	-2.01	0.044	-.270131	-.0034836
_cons	-1.618961	.4347771	-3.72	0.000	-2.471108	-.7668132

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
patient: Identity				
var(_cons)	16.06538	3.057361	11.06371	23.32818

LR test vs. logistic model: $\text{chibar2}(01) = 565.22$ Prob \geq $\text{chibar2} = 0.0000$

It is of interest to determine the dependence among responses for the same subject (between-subject heterogeneity). Under the latent-linear-response formulation, this dependence can be obtained with the intraclass correlation. Formally, in a two-level random-effects model, the intraclass correlation corresponds to the correlation of latent responses within the same individual and also to the proportion of variance explained by the individual random effect.

In the presence of fixed-effects covariates, `estat icc` reports the residual intraclass correlation, which is the correlation between latent linear responses conditional on the fixed-effects covariates.

We use `estat icc` to estimate the residual intraclass correlation:

```
. estat icc
```

```
Residual intraclass correlation
```

Level	ICC	Std. Err.	[95% Conf. Interval]	
patient	.830027	.026849	.7707981	.8764046

Conditional on treatment and month of treatment, we estimate that latent responses within the same patient have a large correlation of approximately 0.83. Further, 83% of the variance of a latent response is explained by the between-patient variability.

◀

▶ Example 2

In [example 3](#) of [\[ME\] meqrlogit](#), we represented the probability of contraceptive use among Bangladeshi women by using the model (stated with slightly different notation here)

$$\text{logit}(\pi_{ij}) = \beta_0 \text{rural}_{ij} + \beta_1 \text{urban}_{ij} + \beta_2 \text{age}_{ij} + \beta_3 \text{child1}_{ij} + \beta_4 \text{child2}_{ij} + \beta_5 \text{child3}_{ij} + a_j \text{rural}_{ij} + b_j \text{urban}_{ij}$$

where π_{ij} is the probability of contraceptive use, $j = 1, \dots, 60$ districts, $i = 1, \dots, n_j$ women within each district, and a_j and b_j are normally distributed with mean 0 and variance–covariance matrix

$$\Sigma = \text{Var} \begin{bmatrix} a_j \\ b_j \end{bmatrix} = \begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{bmatrix}$$

```

. use http://www.stata-press.com/data/r14/bangladesh
(Bangladesh Fertility Survey, 1989)
. generate byte rural = 1 - urban
. meqrlgit c_use rural urban age child*, noconstant || district: rural urban,
> noconstant cov(unstructured)
Refining starting values:
Iteration 0: log likelihood = -1208.3924
Iteration 1: log likelihood = -1204.1317
Iteration 2: log likelihood = -1200.6012
Performing gradient-based optimization:
Iteration 0: log likelihood = -1200.6012
Iteration 1: log likelihood = -1199.3332
Iteration 2: log likelihood = -1199.315
Iteration 3: log likelihood = -1199.315
Mixed-effects logistic regression           Number of obs   =       1,934
Group variable: district                   Number of groups =         60
                                           Obs per group:
                                           min =           2
                                           avg =          32.2
                                           max =          118
Integration points = 7                     Wald chi2(6)    =       120.24
Log likelihood = -1199.315                 Prob > chi2     =        0.0000

```

c_use	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
rural	-1.71165	.1605618	-10.66	0.000	-2.026345	-1.396954
urban	-.8958623	.1704961	-5.25	0.000	-1.230028	-.5616961
age	-.026415	.008023	-3.29	0.001	-.0421398	-.0106902
child1	1.13252	.1603285	7.06	0.000	.818282	1.446758
child2	1.357739	.1770522	7.67	0.000	1.010724	1.704755
child3	1.353827	.1828801	7.40	0.000	.9953882	1.712265

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
district: Unstructured				
var(rural)	.3897439	.1292459	.2034726	.7465393
var(urban)	.2442965	.1450674	.0762886	.782303
cov(rural,urban)	-.0161411	.1057469	-.2234012	.1911189

LR test vs. logistic model: chi2(3) = 58.42 Prob > chi2 = 0.0000

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

Rather than see the estimated variance components listed as variance and covariances as above, we can instead see them as correlations and standard deviations in matrix form; that is, we can see $\hat{\Sigma}$ as a correlation matrix:

```

. estat recovariance, correlation
Random-effects correlation matrix for level district

```

	rural	urban
rural	1	
urban	-.05231	1

The purpose of using this particular model was to allow for district random effects that were specific to the rural and urban areas of that district and that could be interpreted as such. We can obtain predictions of these random effects and their corresponding standard errors,


```
. predict re_rural re_urban, reffects reses(se_rural se_urban)
```

The order in which we specified the variables to be generated corresponds to the order in which the variance components are listed in meqrlogit output. If in doubt, a simple describe will show how these newly generated variables are labeled just to be sure.

Having generated estimated random effects and standard errors, we can now list them for the first 10 districts:

```
. by district, sort: generate tolist = (_n==1)
. list district re_rural se_rural re_urban se_urban if district <= 10 & tolist,
> sep(0)
```

	district	re_rural	se_rural	re_urban	se_urban
1.	1	-.9206641	.3129662	-.5551252	.2321872
118.	2	-.0307772	.3784629	.0012746	.4938357
138.	3	-.0149148	.6242095	.2257356	.4689535
140.	4	-.2684802	.3951617	.5760575	.3970433
170.	5	.0787537	.3078451	.004534	.4675104
209.	6	-.3842217	.2741989	.2727722	.4184852
274.	7	-.1742786	.4008164	.0072177	.493866
292.	8	.0447142	.315396	.2256406	.46799
329.	9	-.3561363	.3885605	.0733451	.4555067
352.	10	-.5368572	.4743089	.0222338	.4939776

◀

□ Technical note

When these data were first introduced in [ME] **meqrlogit**, we noted that not all districts contained both urban and rural areas. This fact is somewhat demonstrated by the random effects that are nearly 0 in the above. A closer examination of the data would reveal that district 3 has no rural areas, and districts 2, 7, and 10 have no urban areas.

The estimated random effects are not exactly 0 in these cases because of the correlation between urban and rural effects. For instance, if a district has no urban areas, it can still yield a nonzero (albeit small) random-effects estimate for a nonexistent urban area because of the correlation with its rural counterpart.

Had we imposed an independent covariance structure in our model, the estimated random effects in the cases in question would be exactly 0. □

□ Technical note

The estimated standard errors produced above with the `reses()` option are conditional on the values of the estimated model parameters: β and the components of Σ . Their interpretation is therefore not one of standard sample-to-sample variability but instead one that does not incorporate uncertainty in the estimated model parameters; see *Methods and formulas*.

That stated, conditional standard errors can still be used as a measure of relative precision, provided that you keep this caveat in mind. □

▷ Example 3

Continuing with [example 2](#), we can obtain predicted probabilities, the default prediction:

```
. predict p
(option mu assumed; predicted means)
```

These predictions are based on a linear predictor that includes both the fixed effects and the random effects due to district. Specifying the `fixedonly` option gives predictions that set the random effects to their prior mean of 0. Below we compare both over the first 20 observations:

```
. predict p_fixed, fixedonly
(option mu assumed; predicted means)
. list c_use p p_fixed age child* in 1/20
```

	c_use	p	p_fixed	age	child1	child2	child3
1.	no	.3579543	.4927183	18.44	0	0	1
2.	no	.2134724	.3210403	-5.56	0	0	0
3.	no	.4672256	.6044016	1.44	0	1	0
4.	no	.4206505	.5584864	8.44	0	0	1
5.	no	.2510909	.3687281	-13.56	0	0	0
6.	no	.2412878	.3565185	-11.56	0	0	0
7.	no	.3579543	.4927183	18.44	0	0	1
8.	no	.4992191	.6345999	-3.56	0	0	1
9.	no	.4572049	.594723	-5.56	1	0	0
10.	no	.4662518	.6034657	1.44	0	0	1
11.	yes	.2412878	.3565185	-11.56	0	0	0
12.	no	.2004691	.3040173	-2.56	0	0	0
13.	no	.4506573	.5883407	-4.56	1	0	0
14.	no	.4400747	.5779263	5.44	0	0	1
15.	no	.4794194	.6160359	-0.56	0	0	1
16.	yes	.4465936	.5843561	4.44	0	0	1
17.	no	.2134724	.3210403	-5.56	0	0	0
18.	yes	.4794194	.6160359	-0.56	0	0	1
19.	yes	.4637673	.6010735	-6.56	1	0	0
20.	no	.5001973	.6355067	-3.56	0	1	0

◀

□ Technical note

Out-of-sample predictions are permitted after `meqrlogit`, but if these predictions involve estimated random effects, the integrity of the estimation data must be preserved. If the estimation data have changed since the model was fit, `predict` will be unable to obtain predicted random effects that are appropriate for the fitted model and will give an error. Thus to obtain out-of-sample predictions that contain random-effects terms, be sure that the data for these predictions are in observations that augment the estimation data.

□

► Example 4

Continuing with [example 2](#), we can also compute intraclass correlations for the model.

In the presence of random-effects covariates, the intraclass correlation is no longer constant and depends on the values of the random-effects covariates. In this case, `estat icc` reports conditional intraclass correlations assuming 0 values for all random-effects covariates. For example, in a two-level model, this conditional correlation represents the correlation of the latent responses for two measurements on the same subject, both of which have random-effects covariates equal to 0. Similarly to the interpretation of intercept variances in random-coefficient models ([Rabe-Hesketh and Skrondal 2012](#), chap. 16), interpretation of this conditional intraclass correlation relies on the usefulness of the 0 baseline values of random-effects covariates. For example, mean centering of the covariates is often used to make a 0 value a useful reference.

Estimation of the conditional intraclass correlation in the Bangladeshi contraceptive study setting of [example 2](#) is of interest. In [example 2](#), the random-effects covariates `rural` and `urban` for the random level `district` are mutually exclusive indicator variables and can never be simultaneously 0. Thus we could not use `estat icc` to estimate the conditional intraclass correlation for this model, because `estat icc` requires that the random intercept is included in all random-effects specifications.

Instead, we consider an alternative model for the Bangladeshi contraceptive study. In [example 2](#) of [\[ME\] meqrlogit](#), we represented the probability of contraceptive use among Bangladeshi women with fixed-effects for urban residence (`urban`), age (`age`), and the number of children (`child1`–`child3`). The random effects for urban and rural residence are represented as a random slope for urban residence and a random intercept at the `district` level.

We fit the model with `meqrlogit`:

```
. use http://www.stata-press.com/data/r14/bangladesh, clear
(Bangladesh Fertility Survey, 1989)
. meqrlogit c_use urban age child* || district: urban, covariance(unstructured)

Refining starting values:
Iteration 0:  log likelihood = -1215.8594 (not concave)
Iteration 1:  log likelihood = -1204.0802
Iteration 2:  log likelihood = -1199.7994

Performing gradient-based optimization:
Iteration 0:  log likelihood = -1199.7994
(output omitted)
Iteration 4:  log likelihood = -1199.315

Mixed-effects logistic regression          Number of obs   =       1,934
Group variable: district                  Number of groups =         60

Obs per group:
      min =          2
      avg =        32.2
      max =        118

Integration points = 7                    Wald chi2(5)    =       97.50
Log likelihood = -1199.315                Prob > chi2     =       0.0000
```

c_use	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
urban	.8157872	.1715519	4.76	0.000	.4795516	1.152023
age	-.026415	.008023	-3.29	0.001	-.0421398	-.0106902
child1	1.13252	.1603285	7.06	0.000	.818282	1.446758
child2	1.357739	.1770522	7.67	0.000	1.010724	1.704755
child3	1.353827	.1828801	7.40	0.000	.9953882	1.712265
_cons	-1.71165	.1605617	-10.66	0.000	-2.026345	-1.396954

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
district: Unstructured				
var(urban)	.6663222	.3224715	.2580709	1.7204
var(_cons)	.3897434	.1292458	.2034723	.7465387
cov(urban,_cons)	-.4058846	.1755418	-.7499402	-.0618289

LR test vs. logistic model: $\chi^2(3) = 58.42$ Prob > $\chi^2 = 0.0000$

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

We use `estat icc` to estimate the intraclass correlation conditional on urban being equal to 0:

```
. estat icc
Conditional intraclass correlation
```

Level	ICC	Std. Err.	[95% Conf. Interval]	
district	.1059197	.0314044	.0582458	.1849513

Note: ICC is conditional on zero values of random-effects covariates.

This estimate suggests that the latent responses are not strongly correlated for rural residents (`urban == 0`) within the same district, conditional on the fixed-effects covariates.

◀

► Example 5

In [example 4](#) of [\[ME\] meqrlgit](#), we fit a three-level model for the cognitive ability of schizophrenia patients as compared with their relatives and a control. Fixed-effects covariates include the difficulty of the test, `difficulty`, and an individual's category, `group` (control, family member of patient, or patient). Family units (`family`) represent the third nesting level, and individual subjects (`subject`) represent the second nesting level. Three measurements were taken on all but one subject, one for each difficulty measure.

We fit the model with `meqrlgit`:

```
. use http://www.stata-press.com/data/r14/towerlondon
(Tower of London data)
. meqrlgit dtlm difficulty i.group || family: || subject:
Refining starting values:
Iteration 0: log likelihood = -310.28433
Iteration 1: log likelihood = -306.42785 (not concave)
Iteration 2: log likelihood = -305.25996
Performing gradient-based optimization:
Iteration 0: log likelihood = -305.25996
Iteration 1: log likelihood = -305.12097
Iteration 2: log likelihood = -305.12043
Iteration 3: log likelihood = -305.12043
Mixed-effects logistic regression                      Number of obs        =        677
```

Group Variable	No. of Groups	Observations per Group			Integration Points
		Minimum	Average	Maximum	
family	118	2	5.7	27	7
subject	226	2	3.0	3	7

Stored results

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

Scalars	
<code>r(relevels)</code>	number of levels
Matrices	
<code>r(Cov#)</code>	level-# random-effects covariance matrix
<code>r(Corr#)</code>	level-# random-effects correlation matrix (if option <code>correlation</code> was specified)

For a G -level nested model, # can be any integer between 2 and G .

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

Scalars	
<code>r(icc#)</code>	level-# intraclass correlation
<code>r(se#)</code>	standard errors of level-# intraclass correlation
<code>r(level)</code>	confidence level of confidence intervals
Macros	
<code>r(label#)</code>	label for level #
Matrices	
<code>r(ci#)</code>	vector of confidence intervals (lower and upper) for level-# intraclass correlation

For a G -level nested model, # can be any integer between 2 and G .

Methods and formulas

Methods and formulas are presented under the following headings:

Prediction
Intraclass correlations

Prediction

Continuing the discussion in *Methods and formulas* of [ME] `meqrlogit`, and using the definitions and formulas defined there, we begin by considering the prediction of the random effects \mathbf{u}_j for the j th cluster in a two-level model.

Given a set of estimated `meqrlogit` parameters $(\widehat{\beta}, \widehat{\Sigma})$, a profile likelihood in \mathbf{u}_j is derived from the joint distribution $f(\mathbf{y}_j, \mathbf{u}_j)$ as

$$\mathcal{L}_j(\mathbf{u}_j) = \exp\{c(\mathbf{y}_j, \mathbf{r}_j)\} (2\pi)^{-q/2} |\widehat{\Sigma}|^{-1/2} \exp\left\{g\left(\widehat{\beta}, \widehat{\Sigma}, \mathbf{u}_j\right)\right\} \quad (1)$$

The conditional maximum likelihood estimator of \mathbf{u}_j —conditional on fixed $(\widehat{\beta}, \widehat{\Sigma})$ —is the maximizer of $\mathcal{L}_j(\mathbf{u}_j)$ or, equivalently, the value of $\widehat{\mathbf{u}}_j$ that solves

$$\mathbf{0} = g'\left(\widehat{\beta}, \widehat{\Sigma}, \widehat{\mathbf{u}}_j\right) = \mathbf{Z}'_j \left\{ \mathbf{y}_j - \mathbf{m}(\widehat{\beta}, \widehat{\mathbf{u}}_j) \right\} - \widehat{\Sigma}^{-1} \widehat{\mathbf{u}}_j$$

Because (1) is proportional to the conditional density $f(\mathbf{u}_j | \mathbf{y}_j)$, you can also refer to $\widehat{\mathbf{u}}_j$ as the conditional mode (or posterior mode if you lean toward Bayesian terminology). Regardless, you are referring to the same estimator.

Conditional standard errors for the estimated random effects are derived from standard theory of maximum likelihood, which dictates that the asymptotic variance matrix of $\widehat{\mathbf{u}}_j$ is the negative inverse of the Hessian, which is estimated as

$$g'' \left(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\Sigma}}, \widehat{\mathbf{u}}_j \right) = - \left\{ \mathbf{Z}'_j \mathbf{V} \left(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{u}}_j \right) \mathbf{Z}_j + \widehat{\boldsymbol{\Sigma}}^{-1} \right\}$$

Similar calculations extend to models with more than one level of random effects; see [Pinheiro and Chao \(2006\)](#).

For any observation i in the j th cluster in a two-level model, define the linear predictor as

$$\widehat{\eta}_{ij} = \mathbf{x}_{ij} \widehat{\boldsymbol{\beta}} + \mathbf{z}_{ij} \widehat{\mathbf{u}}_j$$

In a three-level model, for the i th observation within the j th level-two cluster within the k th level-three cluster,

$$\widehat{\eta}_{ijk} = \mathbf{x}_{ijk} \widehat{\boldsymbol{\beta}} + \mathbf{z}_{ijk}^{(3)} \widehat{\mathbf{u}}_k^{(3)} + \mathbf{z}_{ijk}^{(2)} \widehat{\mathbf{u}}_{jk}^{(2)}$$

where $\mathbf{z}^{(p)}$ and $\mathbf{u}^{(p)}$ refer to the level p design variables and random effects, respectively. For models with more than three levels, the definition of $\widehat{\eta}$ extends in the natural way, with only the notation becoming more complicated.

If the `fixedonly` option is specified, $\widehat{\eta}$ contains the linear predictor for only the fixed portion of the model, for example, in a two-level model $\widehat{\eta}_{ij} = \mathbf{x}_{ij} \widehat{\boldsymbol{\beta}}$. In what follows, we assume a two-level model, with the only necessary modification for multilevel models being the indexing.

The predicted mean conditional on the random effects $\widehat{\mathbf{u}}_j$ is

$$\widehat{\mu}_{ij} = r_{ij} H(\widehat{\eta}_{ij})$$

Pearson residuals are calculated as

$$\nu_{ij}^P = \frac{y_{ij} - \widehat{\mu}_{ij}}{\{V(\widehat{\mu}_{ij})\}^{1/2}}$$

for $V(\widehat{\mu}_{ij}) = \widehat{\mu}_{ij}(1 - \widehat{\mu}_{ij}/r_{ij})$.

Deviance residuals are calculated as

$$\nu_{ij}^D = \text{sign}(y_{ij} - \widehat{\mu}_{ij}) \sqrt{\widehat{d}_{ij}^2}$$

where

$$\widehat{d}_{ij}^2 = \begin{cases} 2r_{ij} \log \left(\frac{r_{ij}}{r_{ij} - \widehat{\mu}_{ij}} \right) & \text{if } y_{ij} = 0 \\ 2y_{ij} \log \left(\frac{y_{ij}}{\widehat{\mu}_{ij}} \right) + 2(r_{ij} - y_{ij}) \log \left(\frac{r_{ij} - y_{ij}}{r_{ij} - \widehat{\mu}_{ij}} \right) & \text{if } 0 < y_{ij} < r_{ij} \\ 2r_{ij} \log \left(\frac{r_{ij}}{\widehat{\mu}_{ij}} \right) & \text{if } y_{ij} = r_{ij} \end{cases}$$

Anscombe residuals are calculated as

$$\nu_{ij}^A = \frac{3 \left\{ y_{ij}^{2/3} \mathcal{H}(y_{ij}/r_{ij}) - \widehat{\mu}_{ij}^{2/3} \mathcal{H}(\widehat{\mu}_{ij}/r_{ij}) \right\}}{2 \left(\widehat{\mu}_{ij} - \widehat{\mu}_{ij}^2/r_{ij} \right)^{1/6}}$$

where $\mathcal{H}(t)$ is a specific univariate case of the Hypergeometric2F1 function (Wolfram 1999, 771–772). For Anscombe residuals for binomial regression, the specific form of the Hypergeometric2F1 function that we require is $\mathcal{H}(t) = {}_2F_1(2/3, 1/3, 5/3, t)$.

For a discussion of the general properties of the above residuals, see [Hardin and Hilbe \(2012, chap. 4\)](#).

Intraclass correlations

Consider a simple, two-level random-intercept model, stated in terms of a latent linear response, where only $y_{ij} = I(y_{ij}^* > 0)$ is observed for the latent variable,

$$y_{ij}^* = \beta + u_j^{(2)} + \epsilon_{ij}^{(1)}$$

with $i = 1, \dots, n_j$ and level-2 groups $j = 1, \dots, M$. Here β is an unknown fixed intercept, $u_j^{(2)}$ is a level-2 random intercept, and $\epsilon_{ij}^{(1)}$ is a level-1 error term. Errors are assumed to be logistic with mean 0 and variance $\sigma_1^2 = \pi^2/3$; random intercepts are assumed to be normally distributed with mean 0 and variance σ_2^2 and to be independent of error terms.

The intraclass correlation for this model is

$$\rho = \text{Corr}(y_{ij}^*, y_{i'j}^*) = \frac{\sigma_2^2}{\pi^2/3 + \sigma_2^2}$$

It corresponds to the correlation between the latent responses i and i' from the same group j .

Now consider a three-level nested random-intercept model,

$$y_{ijk}^* = \beta + u_{jk}^{(2)} + u_k^{(3)} + \epsilon_{ijk}^{(1)}$$

for measurements $i = 1, \dots, n_{jk}$ and level-2 groups $j = 1, \dots, M_{1k}$ nested within level-3 groups $k = 1, \dots, M_2$. Here $u_{jk}^{(2)}$ is a level-2 random intercept, $u_k^{(3)}$ is a level-3 random intercept, and $\epsilon_{ijk}^{(1)}$ is a level-1 error term. The error terms have a logistic distribution with mean 0 and variance $\sigma_1^2 = \pi^2/3$. The random intercepts are assumed to be normally distributed with mean 0 and variances σ_2^2 and σ_3^2 , respectively, and to be mutually independent. The error terms are also independent of the random intercepts.

We can consider two types of intraclass correlations for this model. We will refer to them as level-2 and level-3 intraclass correlations. The level-3 intraclass correlation is

$$\rho^{(3)} = \text{Corr}(y_{ijk}^*, y_{i'j'k}^*) = \frac{\sigma_3^2}{\pi^2/3 + \sigma_2^2 + \sigma_3^2}$$

This is the correlation between latent responses i and i' from the same level-3 group k and from different level-2 groups j and j' .

The level-2 intraclass correlation is

$$\rho^{(2)} = \text{Corr}(y_{ijk}^*, y_{i'jk}^*) = \frac{\sigma_2^2 + \sigma_3^2}{\pi^2/3 + \sigma_2^2 + \sigma_3^2}$$

This is the correlation between latent responses i and i' from the same level-3 group k and level-2 group j . (Note that level-1 intraclass correlation is undefined.)

More generally, for a G -level nested random-intercept model, the g -level intraclass correlation is defined as

$$\rho^{(g)} = \frac{\sum_{l=g}^G \sigma_l^2}{\pi^2/3 + \sum_{l=2}^G \sigma_l^2}$$

The above formulas also apply in the presence of fixed-effects covariates \mathbf{X} in a random-effects model. In this case, intraclass correlations are conditional on fixed-effects covariates and are referred to as residual intraclass correlations. `estat icc` also uses the same formulas to compute intraclass correlations for random-coefficient models, assuming 0 baseline values for the random-effects covariates, and labels them as conditional intraclass correlations.

Intraclass correlations will always fall in $[0,1]$ because variance components are nonnegative. To accommodate the range of an intraclass correlation, we use the logit transformation to obtain confidence intervals. We use the delta method to estimate the standard errors of the intraclass correlations.

Let $\widehat{\rho}^{(g)}$ be a point estimate of the intraclass correlation and $\widehat{\text{SE}}(\widehat{\rho}^{(g)})$ be its standard error. The $(1 - \alpha) \times 100\%$ confidence interval for $\text{logit}(\rho^{(g)})$ is

$$\text{logit}(\widehat{\rho}^{(g)}) \pm z_{\alpha/2} \frac{\widehat{\text{SE}}(\widehat{\rho}^{(g)})}{\widehat{\rho}^{(g)}(1 - \widehat{\rho}^{(g)})}$$

where $z_{\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution and $\text{logit}(x) = \ln\{x/(1-x)\}$. Let k_u be the upper endpoint of this interval, and let k_l be the lower. The $(1 - \alpha) \times 100\%$ confidence interval for $\rho^{(g)}$ is then given by

$$\left(\frac{1}{1 + e^{-k_l}}, \frac{1}{1 + e^{-k_u}} \right)$$

References

- De Backer, M., C. De Vroey, E. Lesaffre, I. Scheys, and P. De Keyser. 1998. Twelve weeks of continuous oral therapy for toenail onychomycosis caused by dermatophytes: A double-blind comparative trial of terbinafine 250 mg/day versus itraconazole 200 mg/day. *Journal of the American Academy of Dermatology* 38: S57–S63.
- Hardin, J. W., and J. M. Hilbe. 2012. *Generalized Linear Models and Extensions*. 3rd ed. College Station, TX: Stata Press.
- Lesaffre, E., and B. Spiessens. 2001. On the effect of the number of quadrature points in a logistic random-effects model: An example. *Journal of the Royal Statistical Society, Series C* 50: 325–335.
- McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*. 2nd ed. London: Chapman & Hall/CRC.
- Pinheiro, J. C., and E. C. Chao. 2006. Efficient Laplacian and adaptive Gaussian quadrature algorithms for multilevel generalized linear mixed models. *Journal of Computational and Graphical Statistics* 15: 58–81.
- Rabe-Hesketh, S., and A. Skrondal. 2012. *Multilevel and Longitudinal Modeling Using Stata*. 3rd ed. College Station, TX: Stata Press.
- Wolfram, S. 1999. *The Mathematica Book*. 4th ed. Cambridge: Cambridge University Press.

Also see

- [ME] [meqrlogit](#) — Multilevel mixed-effects logistic regression (QR decomposition)
- [U] [20 Estimation and postestimation commands](#)