# Title

**xtmelogit postestimation —** Postestimation tools for xtmelogit

# Description

The following postestimation commands are of special interest after `xtmelogit`:

| 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 |

For information about these commands, see below.

The following standard postestimation commands are also available:

| Command | Description |
|---|---|
| `contrast` | contrasts and ANOVA-style joint tests of estimates |
| `estat` | AIC, BIC, VCE, and estimation sample summary |
| `estimates` | cataloging estimation results |
| `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 |

See the corresponding entries in the *Base Reference Manual* for details.

## Special-interest postestimation commands

`estat group` reports 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 `tabulated` 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 zero. They are not available for crossed-effects models.

# Syntax for predict

*Syntax for obtaining estimated random effects or their standard errors*

predict [*type*] {*stub**|*newvarlist*} [*if*] [*in*], {reffects|reses}
[level(*levelvar*)]

*Syntax for obtaining other predictions*

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

| *statistic* | Description |
|---|---|
| Main | |
| mu | the predicted mean; the default |
| xb | linear prediction for the *fixed* portion of the model only |
| stdp | standard error of the fixed-portion linear prediction |
| pearson | Pearson residuals |
| deviance | deviance residuals |
| anscombe | Anscombe residuals |

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

# Menu

Statistics > Postestimation > Predictions, residuals, etc.

# Options for predict

Main

reffects calculates posterior modal estimates of the random effects. By default, estimates for all random effects in the model are calculated. However, if the level(*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*, reffects level(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 *stub*1, *stub*2, ..., *stubq* for you.

`reses` calculates standard errors for the random-effects estimates obtained by using the `reffects` option. By default, standard errors for all random effects in the model are calculated. However, if the `level(`*levelvar*`)` option is specified, then standard errors for only level *levelvar* in the model are calculated. For example, if `classes` are nested within `schools`, then typing

```
. predict se*, reses level(school)
```

would yield standard errors 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 *stub*`1`, *stub*`2`, ..., *stubq* for you.

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 `xtmelogit`. Still, examining the variable labels of the generated variables (using the `describe` command, for instance) can be useful in deciphering which variables correspond to which terms in the model.

`level(`*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 `_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 zero.

`xb` calculates the linear prediction $\mathbf{x}\boldsymbol{\beta}$ 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 zero.

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

`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, residuals that 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.

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

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

# Syntax for estat group

`estat group`

# Menu

Statistics > Postestimation > Reports and statistics

# Syntax for estat recovariance

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

# Menu

Statistics > Postestimation > Reports and statistics

# Options for estat recovariance

`level(`*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 variable describing the grouping at that level or `_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) are displayed; see [P] **matlist** for a list of what is available.

# Syntax for estat icc

`estat icc [, level(#)]`

# Menu

Statistics > Postestimation > Reports and statistics

# 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

Various predictions, statistics, and diagnostic measures are available after fitting a logistic mixed-effects model with `xtmelogit`. 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 using `xtmelogit`:

```
. use http://www.stata-press.com/data/r12/toenail
(Onycholysis severity)
. xtmelogit outcome treatment month trt_month || patient:, intpoints(30)
  (output omitted)
Mixed-effects logistic regression               Number of obs      =      1908
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            |
                   sd(_cons) |   4.008164   .3813917      3.326216    4.829926
------------------------------------------------------------------------------
LR test vs. logistic regression: chibar2(01) =   565.22 Prob>=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 [XT] **xtmelogit**, we represented the probability of contraceptive use among Bangladeshi women by using the model (stated with slightly different notation here)

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

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

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

```
. use http://www.stata-press.com/data/r12/bangladesh
(Bangladesh Fertility Survey, 1989)
. generate byte rural = 1 - urban
. xtmelogit c_use rural urban age child*, nocons || district: rural urban,
> nocons cov(unstructured)
  (output omitted)
Mixed-effects logistic regression               Number of obs      =      1934
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 | -.5616962 |
| 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       |
                   sd(rural) |   .6242947   .1035136      .4510794    .8640251
                   sd(urban) |   .4942636    .146751      .2762039    .8844789
           corr(rural,urban) |  -.0523099   .3384599     -.6153876    .5461173
------------------------------------------------------------------------------
LR test vs. logistic regression:     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 standard deviations and correlations as above, we can instead see them as variance–covariances in matrix form; that is, we can see $\widehat{\boldsymbol{\Sigma}}$

```
. estat recovariance
Random-effects covariance matrix for level district

             |     rural      urban
-------------+----------------------
       rural |   .3897439
       urban |  -.0161411   .2442965
```

or we can see $\widehat{\boldsymbol{\Sigma}}$ as a correlation matrix

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

             |     rural      urban
-------------+----------------------
       rural |          1
       urban |  -.0523099          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

```
. predict re_rural re_urban, reffects
```

and their corresponding standard errors

```
. predict se_rural se_urban, reses
```

The order in which we specified the variables to be generated corresponds to the order in which the variance components are listed in `xtmelogit` 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 [XT] **xtmelogit**, we noted that not all districts contained both urban and rural areas. This fact is somewhat demonstrated by the random effects that are nearly zero 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 zero in these cases is 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-effect 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 zero.

❑

## ❑ Technical note

The estimated standard errors produced above using the `reses` option are conditional on the values of the estimated model parameters: $\boldsymbol{\beta}$ and the components of $\boldsymbol{\Sigma}$. 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 random effects due to district. Specifying the `fixedonly` option gives predictions that set the random effects to their prior mean of zero. 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 `xtmelogit`, 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 that 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 zero 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, which both have random-effects covariates equal to zero. 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 zero baseline values of random-effects covariates. For example, mean centering of the covariates is often used to make a zero 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 zero. 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 [XT] **xtmelogit**, we represented the probability of contraceptive use among Bangladeshi women using 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 using `xtmelogit`:

```
. use http://www.stata-press.com/data/r12/bangladesh
(Bangladesh Fertility Survey, 1989)

. xtmelogit c_use urban age child* || district: urban, covariance(unstructured)
> variance
  (output omitted )
Mixed-effects logistic regression               Number of obs      =      1934
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 regression:     chi2(3) =    58.42   Prob > chi2 = 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 zero:

```
. 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 [XT] **xtmelogit**, 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 using `xtmelogit`:

```
. use http://www.stata-press.com/data/r12/towerlondon
(Tower of London data)
. xtmelogit dtlm difficulty i.group || family: || subject:
  (output omitted)
Mixed-effects logistic regression               Number of obs      =       677
```

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

```
                                                Wald chi2(3)       =     74.89
Log likelihood = -305.12043                     Prob > chi2        =    0.0000
```

| dtlm | Coef. | Std. Err. | z | P>\|z\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| difficulty | -1.648506 | .1932139 | -8.53 | 0.000 | -2.027198 | -1.269814 |
| group | | | | | | |
| 2 | -.24868 | .3544065 | -0.70 | 0.483 | -.943304 | .445944 |
| 3 | -1.0523 | .3999896 | -2.63 | 0.009 | -1.836265 | -.2683349 |
| _cons | -1.485861 | .2848469 | -5.22 | 0.000 | -2.04415 | -.9275709 |

```
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
family: Identity             |
                   sd(_cons) |   .7544415   .3457249      .3072983    1.852213
-----------------------------+------------------------------------------------
subject: Identity            |
                   sd(_cons) |   1.066739   .3214235      .5909884    1.925472
------------------------------------------------------------------------------
LR test vs. logistic regression:     chi2(2) =    17.54   Prob > chi2 = 0.0002

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

We can use `estat icc` to estimate the residual intraclass correlation (conditional on the difficulty level and the individual's category) between the latent responses of subjects within the same family or between the latent responses of the same subject and family:

```
. estat icc

Residual intraclass correlation

------------------------------------------------------------------------------
                       Level |        ICC   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
                      family |   .1139052   .0997976      .0181741    .4716556
              subject|family |   .3416289   .0889531      .1929133    .5297405
------------------------------------------------------------------------------
```

`estat icc` reports two intraclass correlations for this three-level nested model. The first is the level-3 intraclass correlation at the family level, the correlation between latent measurements of the cognitive ability in the same family. The second is the level-2 intraclass correlation at the subject-within-family level, the correlation between the latent measurements of cognitive ability in the same subject and family.

There is not a strong correlation between individual realizations of the latent response, even within the same subject.

◁

# Saved results

`estat recovariance` saves the last-displayed random-effects covariance matrix in `r(cov)` or in `r(corr)` if it is displayed as a correlation matrix.

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

Scalars

| | |
|---|---|
| `r(icc#)` | level-# intraclass correlation |
| `r(se#)` | standard errors of level-# intraclass correlation |
| `r(level)` | confidence level of confidence intervals |

Macros

| | |
|---|---|
| `r(label#)` | label for level # |

Matrices

| | |
|---|---|
| `r(ci#)` | 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 [XT] **xtmelogit**, 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 `xtmelogit` parameters, $(\widehat{\boldsymbol\beta}, \widehat{\boldsymbol\Sigma})$, a profile likelihood in $\mathbf{u}_j$ is derived from the joint distribution $f(\mathbf{y}_j, \mathbf{u}_j)$ as

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

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

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

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

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

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

For any $i$ observation in the $j$ 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 the $\mathbf{z}^{(p)}$ 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

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

Anscombe residuals are calculated as

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

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 (2007, 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 $i = 1, \ldots, n_j$ and level-2 groups $j = 1, \ldots, M$. Here $\beta$ 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 zero and variance $\sigma_1^2 = \pi^2/3$; random intercepts are assumed to be normally distributed with mean zero and variance $\sigma_2^2$ and independent of error terms.

The intraclass correlation for this model is

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

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:

$$y^*_{ijk} = \beta + u^{(2)}_{jk} + u^{(3)}_{k} + \epsilon^{(1)}_{ijk}$$

for measurements $i = 1, \ldots, n_{jk}$ and level-2 groups $j = 1, \ldots, M_{1k}$ nested within level-3 groups $k = 1, \ldots, M_2$. Here $u^{(2)}_{jk}$ is a level-2 random intercept, $u^{(3)}_{k}$ is a level-3 random intercept, and $\epsilon^{(1)}_{ijk}$ is a level-1 error term. The error terms have a logistic distribution with mean zero and variance $\sigma_1^2 = \pi^2/3$. The random intercepts are assumed to be normally distributed with mean zero and variances $\sigma_2^2$ and $\sigma_3^2$, 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

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

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

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

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 zero baseline values for the random-effects covariates, and labels them as conditional intraclass correlations.

Intraclass correlations are estimated using the delta method and 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.

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

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

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. 2007. *Generalized Linear Models and Extensions*. 2nd 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

[XT] **xtmelogit** — Multilevel mixed-effects logistic regression

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