Tables for Multilevel Models in Stata

Andy Grogan-Kaylor

2024-03-28

1 Introduction

Stata has a number of ways of making tables. Here are some alternative commands, and some *tweaks* that may be especially useful for multilevel models.

2 Data Source

The data used in this example are derived from the R package *Functions and Datasets for "Forest Analytics with R"*.

According to the documentation, the source of these data are: "von Guttenberg's Norway spruce (Picea abies [L.] Karst) tree measurement data."

Figure 1: Old Tjikko, a 9,550 Year Old Norway Spruce in Sweden

The documentation goes on to further note that:

> "The data are measures from 107 trees. The trees were selected as being of average size from healthy and well stocked stands in the Alps."

3 Setup

```
clear all // clear workspace

use "gutten.dta", clear // use tree data as example

describe // describe the data
```

```
Contains data from gutten.dta
 Observations:         1,200
    Variables:             9                  19 Feb 2020 08:23
```

```
------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
------------------------------------------------------------------------------
site            long    %9.0g      site       site
location        long    %9.0g      location   location
tree            long    %9.0g                 tree
age_base        long    %9.0g                 age.base
height          double  %9.0g                 height
dbh_cm          double  %9.0g                 dbh.cm
volume          double  %9.0g                 volume
age_bh          double  %9.0g                 age.bh
tree_ID         long    %9.0g      tree_ID    tree.ID
------------------------------------------------------------------------------
Sorted by:
```

4 Estimate a Multilevel Model 

```
mixed height age_base i.site || tree_ID: // mixed model

est store M1 // store the estimates (this would work with multiple stored
estimates)
```

```
Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood = -3051.1192
Iteration 1:  Log likelihood = -3051.1192

Computing standard errors ...

Mixed-effects ML regression                         Number of obs    =   1,200
Group variable: tree_ID                             Number of groups =     107
                                                    Obs per group:
                                                                 min =       5
                                                                 avg =    11.2
                                                                 max =      15
                                                    Wald chi2(5)     = 8651.66
Log likelihood = -3051.1192                         Prob > chi2      =  0.0000

------------------------------------------------------------------------------
      height | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    age_base |   .2144446   .0023691    90.52   0.000     .2098014    .2190879
             |
```

```
        site |
          2  |  -3.316408   .4738969    -7.00   0.000    -4.245229   -2.387587
          3  |  -8.094846   .5358151   -15.11   0.000    -9.145024   -7.044667
          4  |  -11.50985   .5291215   -21.75   0.000    -12.54691   -10.47279
          5  |  -15.86582   .7116202   -22.30   0.000    -17.26057   -14.47107
             |
       _cons |   8.233362   .4092147    20.12   0.000     7.431316    9.035408
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
tree_ID: Identity            |
                  var(_cons) |   2.170508   .4004445      1.511891    3.116037
-----------------------------+------------------------------------------------
               var(Residual) |   8.392966   .3586298      7.718693     9.12614
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 135.90        Prob >= chibar2 = 0.0000
```

5 Use `estimates table`

```
estimates table M1, b(%9.3f) star // nicely formatted table of results
```

```
    Variable |      M1
-------------+---------------
height       |
    age_base |     0.214***
             |
        site |
          2  |    -3.316***
          3  |    -8.095***
          4  |   -11.510***
          5  |   -15.866***
             |
       _cons |     8.233***
-------------+---------------
lns1_1_1     |
       _cons |     0.387***
-------------+---------------
lnsig_e      |
       _cons |     1.064***
-----------------------------
Legend: * p<0.05; ** p<0.01; *** p<0.001
```

6 Use `estimates store` With `, variance post`

Frustratingly, as you can see in Section 5, with multilevel models, the default behavior of `estimates table` is to report the *ln* of the random effects. Below, I use the `, variance post` option to post the *variance* rather than the *logarithm of the variance*.

> Notice how `, variance post` essentially *replays* the results, but with the random effects as variances, rather than as the logarithm of the standard deviation.

```
mixed height age_base i.site || tree_ID: // mixed model

estat sd, variance post // post results as variance scale rather than log scale

est store M2 // store the estimates (this would work with multiple stored
estimates)
```

```
Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood = -3051.1192
Iteration 1:  Log likelihood = -3051.1192

Computing standard errors ...

Mixed-effects ML regression                     Number of obs    =   1,200
Group variable: tree_ID                         Number of groups =     107
                                                Obs per group:
                                                              min =       5
                                                              avg =    11.2
                                                              max =      15
                                                Wald chi2(5)     = 8651.66
Log likelihood = -3051.1192                     Prob > chi2      =  0.0000

------------------------------------------------------------------------------
      height | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    age_base |   .2144446   .0023691    90.52   0.000     .2098014    .2190879
             |
        site |
          2  |  -3.316408   .4738969    -7.00   0.000    -4.245229   -2.387587
          3  |  -8.094846   .5358151   -15.11   0.000    -9.145024   -7.044667
          4  |  -11.50985   .5291215   -21.75   0.000    -12.54691   -10.47279
          5  |  -15.86582   .7116202   -22.30   0.000    -17.26057   -14.47107
             |
       _cons |   8.233362   .4092147    20.12   0.000     7.431316    9.035408
```

```
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
tree_ID: Identity            |
                  var(_cons) |   2.170508   .4004445      1.511891    3.116037
-----------------------------+------------------------------------------------
               var(Residual) |   8.392966   .3586298      7.718693     9.12614
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 135.90        Prob >= chibar2 = 0.0000

------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
height       |
    age_base |   .2144446   .0023691    90.52   0.000     .2098014    .2190879
             |
        site |
          2  |  -3.316408   .4738969    -7.00   0.000    -4.245229   -2.387587
          3  |  -8.094846   .5358151   -15.11   0.000    -9.145024   -7.044667
          4  |  -11.50985   .5291215   -21.75   0.000    -12.54691   -10.47279
          5  |  -15.86582   .7116202   -22.30   0.000    -17.26057   -14.47107
             |
       _cons |   8.233362   .4092147    20.12   0.000     7.431316    9.035408
-------------+----------------------------------------------------------------
tree_ID      |
   var(_cons)|   2.170508   .4004445                      1.511891    3.116037
-------------+----------------------------------------------------------------
Residual     |
       var(e)|   8.392966   .3586298                      7.718693     9.12614
------------------------------------------------------------------------------
```

7 Use `estimates table` To Compare These Approaches

> ⚠️ We Usually Use `estimates table` for *Different* Models
>
> When used with multiple sets of estimates, we usually use `estimates table` to present the results of *different* models, rather than the same model presented in different ways. Below, however, for the sake of illustration, we present the *same* model in two different ways.

```
* nicely formatted table of results
```

```
estimates table M1 M2, b(%9.3f) star ///
  title("M1 and M2 are the Same Model Presented Differently")
```

```
M1 and M2 are the Same Model Presented Differently

----------------------------------------------
    Variable |      M1             M2
-------------+--------------------------------
height       |
    age_base |     0.214***       0.214***
             |
        site |
          2  |    -3.316***      -3.316***
          3  |    -8.095***      -8.095***
          4  |   -11.510***     -11.510***
          5  |   -15.866***     -15.866***
             |
       _cons |     8.233***       8.233***
-------------+--------------------------------
lns1_1_1     |
       _cons |     0.387***
-------------+--------------------------------
lnsig_e      |
       _cons |     1.064***
-------------+--------------------------------
tree_ID      |
   var(_cons)|                    2.171***
-------------+--------------------------------
Residual     |
       var(e)|                    8.393***
----------------------------------------------
    Legend: * p<0.05; ** p<0.01; *** p<0.001
```

8 Use `etable`

`etable` is a newer Stata command that is very useful for making nicely formatted tables. `etable` works with one estimate or multiple estimates.

```
etable, estimates(M1) /// use these estimate(s)
novarlabel /// variable names only
cstat(_r_b) /// beta's only
showstars showstarsnote ///
column(estimate) // column is modelname
```

```
                           M1
---------------------------------
age_base                0.214 **
site
  2                    -3.316 **
  3                    -8.095 **
  4                   -11.510 **
  5                   -15.866 **
_cons                   8.233 **
var(_cons)              2.171
var(e)                  8.393
Number of observations   1200
---------------------------------
** p<.01, * p<.05
```

There is also a very helpful `export` option for exporting these tables to a variety of ouput formats. See `help etable` in Stata for more information.

9 Add One More Set of Estimates for Illustration

9.1 Multiple Estimates With `estimates table`

```
mixed height age_base i.site i.location || tree_ID: // mixed model

estat sd, variance post // post results as variance scale rather than log scale

est store M3 // store the estimates (this would work with multiple stored
estimates)

est table M2 M3, b(%9.3f) star
```

```
Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood = -3047.8267
Iteration 1:  Log likelihood = -3047.8267

Computing standard errors ...

Mixed-effects ML regression                     Number of obs    =   1,200
Group variable: tree_ID                         Number of groups =     107
                                                Obs per group:
                                                              min =       5
                                                              avg =    11.2
```

```
                                                              max =      15
                                                Wald chi2(11)    = 8700.21
Log likelihood = -3047.8267                     Prob > chi2      =  0.0000

------------------------------------------------------------------------------
      height | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    age_base |   .2143854   .0023822    89.99   0.000     .2097163    .2190544
             |
        site |
          2  |  -2.994348   .5335979    -5.61   0.000     -4.04018   -1.948515
          3  |  -7.764809    .563856   -13.77   0.000    -8.869947   -6.659672
          4  |  -10.84402   .6356708   -17.06   0.000    -12.08991    -9.59813
          5  |  -15.17887   .7953014   -19.09   0.000    -16.73763    -13.6201
             |
    location |
          2  |  -.3215123   1.246019    -0.26   0.796    -2.763665    2.120641
          3  |   .4745482   .6385101     0.74   0.457    -.7769087    1.726005
          4  |   .0598813   .7092946     0.08   0.933    -1.330311    1.450073
          5  |  -.4502186   .5372169    -0.84   0.402    -1.503144    .6027071
          6  |  -.2549412   .7074584    -0.36   0.719    -1.641534    1.131652
          7  |  -1.453754   .7466009    -1.95   0.052    -2.917065    .0095567
             |
       _cons |   8.180898   .5441571    15.03   0.000      7.11437    9.247426
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
tree_ID: Identity            |
                  var(_cons) |   1.981234   .3765076      1.365137    2.875382
-----------------------------+------------------------------------------------
               var(Residual) |   8.396723   .3589345      7.721889    9.130533
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 118.04        Prob >= chibar2 = 0.0000

------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
height       |
    age_base |   .2143854   .0023822    89.99   0.000     .2097163    .2190544
             |
        site |
          2  |  -2.994348   .5335979    -5.61   0.000     -4.04018   -1.948515
          3  |  -7.764809    .563856   -13.77   0.000    -8.869947   -6.659672
          4  |  -10.84402   .6356708   -17.06   0.000    -12.08991    -9.59813
          5  |  -15.17887   .7953014   -19.09   0.000    -16.73763    -13.6201
             |
```

```
    location |
          2  |  -.3215123   1.246019    -0.26   0.796    -2.763665    2.120641
          3  |   .4745482   .6385101     0.74   0.457    -.7769087    1.726005
          4  |   .0598813   .7092946     0.08   0.933    -1.330311    1.450073
          5  |  -.4502186   .5372169    -0.84   0.402    -1.503144    .6027071
          6  |  -.2549412   .7074584    -0.36   0.719    -1.641534    1.131652
          7  |  -1.453754   .7466009    -1.95   0.052    -2.917065    .0095567
             |
       _cons |   8.180898   .5441571    15.03   0.000      7.11437    9.247426
-------------+------------------------------------------------------------------
tree_ID      |
   var(_cons)|   1.981234   .3765076                      1.365137    2.875382
-------------+------------------------------------------------------------------
Residual     |
       var(e)|   8.396723   .3589345                      7.721889    9.130533
------------------------------------------------------------------------------



----------------------------------------------
    Variable |      M2             M3
-------------+--------------------------------
height       |
    age_base |     0.214***       0.214***
             |
        site |
          2  |    -3.316***      -2.994***
          3  |    -8.095***      -7.765***
          4  |   -11.510***     -10.844***
          5  |   -15.866***     -15.179***
             |
    location |
          2  |                   -0.322
          3  |                    0.475
          4  |                    0.060
          5  |                   -0.450
          6  |                   -0.255
          7  |                   -1.454
             |
       _cons |     8.233***       8.181***
-------------+--------------------------------
tree_ID      |
   var(_cons)|     2.171***       1.981***
-------------+--------------------------------
Residual     |
       var(e)|     8.393***       8.397***
----------------------------------------------
    Legend: * p<0.05; ** p<0.01; *** p<0.001
```

9.2 Multiple Estimates With `etable`

```
mixed height age_base i.site i.location || tree_ID: // mixed model

est store M4

etable, estimates(M1 M4) /// use these estimate(s)
novarlabel /// variable names only
cstat(_r_b) /// beta's only
showstars showstarsnote ///
column(estimate) // column is modelname
```

```
Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood = -3047.8267
Iteration 1:  Log likelihood = -3047.8267

Computing standard errors ...

Mixed-effects ML regression                     Number of obs    =   1,200
Group variable: tree_ID                         Number of groups =     107
                                                Obs per group:
                                                              min =       5
                                                              avg =    11.2
                                                              max =      15
                                                Wald chi2(11)    = 8700.21
Log likelihood = -3047.8267                     Prob > chi2      =  0.0000

------------------------------------------------------------------------------
      height | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    age_base |   .2143854   .0023822    89.99   0.000     .2097163    .2190544
             |
        site |
          2  |  -2.994348   .5335979    -5.61   0.000     -4.04018   -1.948515
          3  |  -7.764809    .563856   -13.77   0.000    -8.869947   -6.659672
          4  |  -10.84402   .6356708   -17.06   0.000    -12.08991    -9.59813
          5  |  -15.17887   .7953014   -19.09   0.000    -16.73763    -13.6201
             |
    location |
          2  |  -.3215123   1.246019    -0.26   0.796    -2.763665    2.120641
          3  |   .4745482   .6385101     0.74   0.457    -.7769087    1.726005
          4  |   .0598813   .7092946     0.08   0.933    -1.330311    1.450073
          5  |  -.4502186   .5372169    -0.84   0.402    -1.503144    .6027071
          6  |  -.2549412   .7074584    -0.36   0.719    -1.641534    1.131652
          7  |  -1.453754   .7466009    -1.95   0.052    -2.917065    .0095567
```

```
             |
       _cons |   8.180898   .5441571    15.03   0.000      7.11437    9.247426
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
tree_ID: Identity            |
                  var(_cons) |   1.981234   .3765076      1.365137    2.875382
-----------------------------+------------------------------------------------
               var(Residual) |   8.396723   .3589345      7.721889    9.130533
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 118.04        Prob >= chibar2 = 0.0000



---------------------------------------------
                           M1         M4
---------------------------------------------
age_base                0.214 **   0.214 **
site
  2                    -3.316 **  -2.994 **
  3                    -8.095 **  -7.765 **
  4                   -11.510 ** -10.844 **
  5                   -15.866 ** -15.179 **
location
  2                              -0.322
  3                               0.475
  4                               0.060
  5                              -0.450
  6                              -0.255
  7                              -1.454
_cons                   8.233 **   8.181 **
var(_cons)              2.171      1.981
var(e)                  8.393      8.397
Number of observations    1200       1200
---------------------------------------------
** p<.01, * p<.05
```