

Tables for Multilevel Models in Stata

Andy Grogan-Kaylor

2024-03-21

1 Introduction

Stata has a number of ways of making tables. There 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 name	Storage type	Display format	Value 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

Group variable: tree_ID

Number of obs = 1,200

Number of groups = 107

Obs per group:

min = 5

avg = 11.2

max = 15

Wald chi2(5) = 8651.66

Prob > chi2 = 0.0000

Log likelihood = -3051.1192

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 Table With , variance post

Frustratingly, with multilevel models, the default behavior of `estimates` table is to report the *ln* of the random effects. 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)

* nicely formatted table of results

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

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

M1 and M2 are the Same Model Presented Differently

Variable		

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***
-----+			
lnsl_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

7 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 output formats. See `help etable` in Stata for more information.