Workshop on Multilevel Modeling

Andy Grogan-Kaylor

15 Feb 2023 13:18:15

Navigation

When this document is presented in slide show format, some slides may be long, and you may need to scroll down to see the full slide. In slide show format `b` makes text bigger, and `s` makes text smaller.

Cross Sectional Model

Get Data

```
. use "../multilevel-thinking/simulate-and-analyze-multilevel-data/simulated_multilevel_data.d
> ta", clear
```

The Equation

$$\text{outcome}_{ij} = \beta_0 + \beta_1 \text{parental warmth} + \beta_2 \text{physical punishment} + \beta_3 \text{time} +$$

$$\beta_4 \text{group}_2 + \beta_5 HDI +$$

$$u_{0j} + u_{1j} \times \text{parental warmth} + e_{ij}$$

Descriptive Statistics

```
. summarize // descriptive statistics
```

Variable	Obs	Mean	Std. dev.	Min	Max
country	3,000	15.5	8.656884	1	30
HDI	3,000	64.76667	17.24562	33	87
family	3,000	50.5	28.87088	1	100
id	0				
group	3,000	1.496	.5000674	1	2
physical_p~t	3,000	1.516	1.884744	-2	5
warmth	3,000	2.543667	2.431336	-2	7
outcome	3,000	53.45039	6.884502	25.02363	81.63657

Spaghetti Plot

```
. spagplot outcome warmth, id(country) scheme(s1color)
.
. graph export spagplot1.png, width(1000) replace
file /Users/agrogan/Desktop/GitHub/multilevel-workshop/spagplot1.png saved as PNG format
```

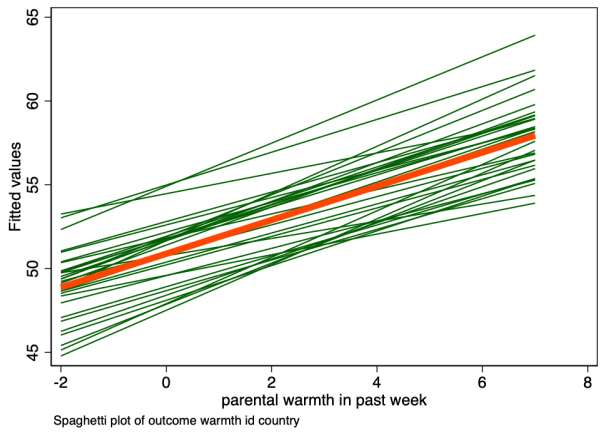

Figure 1: Spaghetti Plot of Outcome by Warmth by Country

Unconditional Model

Model

```
. mixed outcome || country: // unconditional model
Performing EM optimization:
Performing gradient-based optimization:
Iteration 0:   log likelihood = -9956.6096
Iteration 1:   log likelihood = -9956.6096
Computing standard errors:
Mixed-effects ML regression                     Number of obs     =      3,000
Group variable: country                         Number of groups  =         30
                                                Obs per group:
                                                              min =        100
                                                              avg =      100.0
                                                              max =        100
                                                Wald chi2(0)      =          .
Log likelihood = -9956.6096                     Prob > chi2       =          .
```

outcome	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
_cons	53.45039	.3702932	144.35	0.000	52.72463	54.17615

Random-effects parameters	Estimate	Std. err.	[95% conf.	interval]
country: Identity				
var(_cons)	3.676471	1.062168	2.086944	6.476667
var(Residual)	43.70413	1.134121	41.53688	45.98446

```
LR test vs. linear model: chibar2(01) = 175.05        Prob >= chibar2 = 0.0000
```

ICC

```
. estat icc
Intraclass correlation
```

Level	ICC	Std. err.	[95% conf.	interval]
country	.0775944	.0207813	.0454528	.129384

Full Model

```
. mixed outcome warmth physical_punishment i.group HDI || country: warmth // multilevel model
Performing EM optimization:
Performing gradient-based optimization:
Iteration 0:   log likelihood = -9616.8876
Iteration 1:   log likelihood = -9616.3536
Iteration 2:   log likelihood = -9616.3475
Iteration 3:   log likelihood = -9616.3475
Computing standard errors:
Mixed-effects ML regression                     Number of obs     =      3,000
Group variable: country                         Number of groups  =         30
                                                Obs per group:
                                                              min =        100
                                                              avg =      100.0
                                                              max =        100
                                                Wald chi2(4)      =     764.27
Log likelihood = -9616.3475                     Prob > chi2       =     0.0000
```

outcome	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
warmth	.9826773	.0444598	22.10	0.000	.8955377	1.069817
physical_punishment	-.9239791	.0573291	-16.12	0.000	-1.036342	-.8116161
2.group	.7280691	.2163084	3.37	0.001	.3041125	1.152026
HDI	.0075692	.0206019	0.37	0.713	-.0328098	.0479482
_cons	51.50019	1.392584	36.98	0.000	48.77077	54.2296

Random-effects parameters	Estimate	Std. err.	[95% conf.	interval]
country: Independent				
var(warmth)	5.38e-17	4.71e-16	1.90e-24	1.53e-09
var(_cons)	3.43782	.9775981	1.968931	6.002548
var(Residual)	34.7837	.9026366	33.0588	36.5986

```
LR test vs. linear model: chi2(2) = 210.87                Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
.
. est store crosssectional // store estimates
```

Longitudinal Model

Setup

```
. use "../multilevel-thinking/simulate-and-analyze-multilevel-data/simulated_multilevel_longit
> udinal_data.dta", clear
```

The Equation

$$\text{outcome}_{ij} = \beta_0 + \beta_1 \text{parental warmth} + \beta_2 \text{physical punishment} + \beta_3 \text{time} +$$

$$\beta_4 \text{group}_2 + \beta_5 HDI +$$

$$u_{0j} + u_{1j} \times \text{parental warmth} +$$

$$v_{0i} + v_{1i} \times t + e_{ij}$$

Descriptive Statistics

```
. summarize // descriptive statistics
```

Variable	Obs	Mean	Std. dev.	Min	Max
country	9,000	15.5	8.655922	1	30
HDI	9,000	64.76667	17.2437	33	87
family	9,000	50.5	28.86767	1	100
id	0				
group	9,000	1.496	.5000118	1	2
t	9,000	2	.8165419	1	3
physical_p_t	9,000	1.517111	1.884289	-2	5
warmth	9,000	2.533778	2.449075	-2	7
outcome	9,000	54.43846	7.019933	25.02363	81.63657

Alternate Plot

```
. encode id, generate(idNUMERIC) // numeric version of id
.
. * spagplot outcome t if idNUMERIC <= 10, id(idNUMERIC) scheme(s1color)
.
. twoway (lfit outcome t) (scatter outcome t) if idNUMERIC <= 10, by(idNUMERIC) scheme(s1color
> )
.
. graph export spagplot2.png, width(1000) replace
file /Users/agrogan/Desktop/GitHub/multilevel-workshop/spagplot2.png saved as PNG format
```

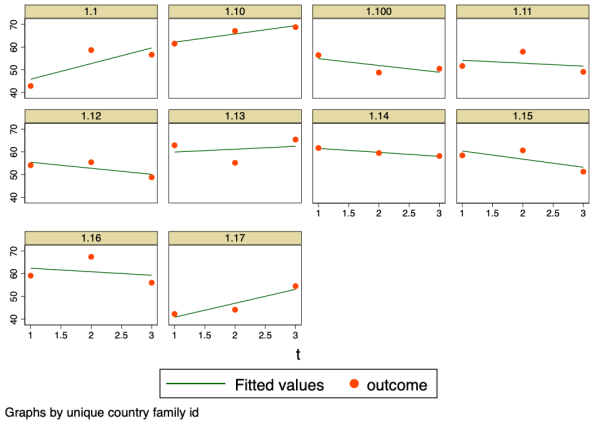

Figure 2: Alternate Plot of Outcome by Time by Individual; First 10 Observations

Unconditional Model

Model

```
. mixed outcome || country: || id: // unconditional model
Performing EM optimization:
Performing gradient-based optimization:
Iteration 0:   log likelihood = -29398.984
Iteration 1:   log likelihood = -29398.984
Computing standard errors:
Mixed-effects ML regression                     Number of obs     =      9,000
        Grouping information
```

Group variable	No. of groups	Observations per group Minimum	Average	Maximum

country	30	300	300.0	300
id	3,000	3	3.0	3

Wald chi2(0) = .
Log likelihood = -29398.984 Prob > chi2 = .

outcome	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
_cons	54.43846	.3767998	144.48	0.000	53.69995	55.17698

Random-effects parameters	Estimate	Std. err.	[95% conf.	interval]
country: Identity				
var(_cons)	3.995172	1.099853	2.329182	6.85279
id: Identity				
var(_cons)	16.98591	.7068169	15.65556	18.42931
var(Residual)	28.29352	.5165663	27.29897	29.3243

LR test vs. linear model: chi2(2) = 1819.49 Prob > chi2 = 0.0000

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

ICC

```
. estat icc
```

Intraclass correlation

Level	ICC	Std. err.	[95% conf.	interval]
country	.0810797	.0205569	.0488675	.1315879
id|country	.4257992	.0163912	.3940284	.4581946

Full Model

```
. mixed outcome t warmth physical_punishment i.group HDI || country: warmth || id: t // multil
> evel model
```

Performing EM optimization:

Performing gradient-based optimization:

```
Iteration 0:   log likelihood = -28546.535
Iteration 1:   log likelihood = -28524.928
Iteration 2:   log likelihood = -28524.635
Iteration 3:   log likelihood = -28524.601
Iteration 4:   log likelihood = -28524.598
Iteration 5:   log likelihood = -28524.598
```

Computing standard errors:

Mixed-effects ML regression Number of obs = 9,000

Grouping information

Group variable	No. of groups	Observations per group Minimum	Average	Maximum
country	30	300	300.0	300
id	3,000	3	3.0	3

Wald chi2(5) = 1818.96
Log likelihood = -28524.598 Prob > chi2 = 0.0000

outcome	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
t	.9929535	.0658203	15.09	0.000	.8639481	1.121959
warmth	1.047045	.0338001	30.98	0.000	.9807983	1.113292

physical_punishment	-.9377711	.0381761	-24.56	0.000	-1.012595	-.8629473
2.group	.8219777	.1530957	5.37	0.000	.5219157	1.12204
HDI	.0047772	.0205645	0.23	0.816	-.0355285	.0450829
_cons	50.50391	1.389611	36.34	0.000	47.78032	53.2275

Random-effects parameters	Estimate	Std. err.	[95% conf.	interval]
country: Independent				
var(warmth)	.0071126	.0086595	.0006542	.0773303
var(_cons)	3.560166	.9807369	2.074844	6.108788
id: Independent				
var(t)	3.01e-10	2.17e-10	7.29e-11	1.24e-09
var(_cons)	8.722256	.4792014	7.831839	9.713906
var(Residual)	25.98996	.4745951	25.07622	26.93699

LR test vs. linear model: chi2(4) = 1331.93 Prob > chi2 = 0.0000

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

```
.
. est store longitudinal // store estimates
```

Nice Table of Results

```
. est table crosssectional longitudinal, ///
> b(%9.3f) star stats(N ll chi2) ///
> varwidth(20) modelwidth(15)
```

Variable	crosssectional	longitudinal
outcome		
warmth	0.983***	1.047***
physical_punishment	-0.924***	-0.938***
group		
2	0.728***	0.822***
HDI	0.008	0.005
t		0.993***
_cons	51.500***	50.504***
lns1_1_1		
_cons	-18.731***	-2.473***
lns1_1_2		
_cons	0.617***	0.635***
lnsig_e		
_cons	1.775***	1.629***
lns2_1_1		
_cons		-10.963***
lns2_1_2		
_cons		1.083***
Statistics		
N	3000	9000
ll	-9616.347	-2.85e+04
chi2	764.268	1818.962

Legend: * p<0.05; ** p<0.01; *** p<0.001

QUESTIONS???