Multilevel Mixed-Effects Generalized Linear Models in STATA

Prof. Dr. Luiz Paulo Fávero
Prof. Dr. Matheus Albergaria
SUMMARY

- Theoretical Fundamentals of Multilevel Models.

- Estimation of Multilevel Mixed-Effects Generalized Linear Models in Stata.
Theoretical Fundamentals of Multilevel Models
Different researchers, from the same database, can estimate different models and, consequently, obtain different predicted values of the phenomenon under study. The objective is to estimate models that, although simplifications of reality, present the best possible adherence between real and fitted values.

A graph showing fitted values against real values. The x-axis represents real values ranging from 0 to 100, and the y-axis represents fitted values also ranging from 0 to 100. There is a dashed line representing a 45° angle and a solid line labeled 'Model 1'. The graph suggests a comparison between the expected (fitted) and actual (real) values.
Multilevel Mixed-Effects Generalized Linear Models (GLLAMM)

Multilevel Mixed-Effects Linear Models

Multilevel Mixed-Effects Non-Linear Models

Multilevel Mixed-Effects Logistic Models

Multilevel Mixed-Effects Count Data Models (Poisson + Negative Binomial)

Multilevel models are models that recognize nested structure in the data.

MULTILEVEL STRUCTURE
MULTILEVEL STRUCTURE

Level 1
Firm

Level 2
Country
MULTILEVEL MODEL
MULTILEVEL MODEL

Context 1: \[ Y_{i1} = \beta_{01} + \beta_{11}.X_{i1} + r_{i1} \]

Context 2: \[ Y_{i2} = \beta_{02} + \beta_{12}.X_{i2} + r_{i2} \]

Context 3: \[ Y_{i3} = \beta_{03} + \beta_{13}.X_{i3} + r_{i3} \]

Context 4: \[ Y_{i4} = \beta_{04} + \beta_{14}.X_{i4} + r_{i4} \]
**MULTILEVEL MODEL**

**Level 1**

\[ \eta_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij} \]

**Level 2**

\[ \beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j} \]
\[ \beta_{1j} = \gamma_{10} + \gamma_{11}W_j + u_{1j} \]

\[ \eta_{ij} = \left(\gamma_{00} + \gamma_{01}W_j + u_{0j}\right) + \left(\gamma_{10} + \gamma_{11}W_j + u_{1j}\right)X_{ij} + r_{ij} \]

intercept with random effects

slope with random effects
Traditional GLM models ignore interactions between variables in the fixed effects component and between error terms and variables in the random effects component.
- If variances of error terms $u_{0j}$ and $u_{1j}$ are statistically different from zero, traditional GLM estimations will not be adequate.


Barbara G. Tabachnick
California State University

Linda S. Fidell
California State University
The inclusion of *dummies representing groups do not capture contextual effects*, because this procedure do not allow the split between observable and unobservable effects over the outcome variable.
Multilevel models allow the development of new and more complex research constructs.

Within a model structure with a single equation, there seems to be no connection between individuals and the society in which they live. In this sense, the use of level equations allows the researcher to 'jump' from one science to another: students and schools, families and neighborhoods, firms and countries. Ignoring this relationship means elaborate incorrect analyzes about the behavior of the individuals and, equally, about the behavior of the groups. Only the recognition of these reciprocal influences allows the correct analysis of the phenomena.

Estimation of Multilevel Mixed-Effects Generalized Linear Models in STATA®
A THREE-LEVEL MIXED-EFFECTS COUNT DATA MODEL

**Level 1**

\[
\ln(\lambda_{ijk}) = \pi_{0jk} + \pi_{1jk} \cdot Z_{1jk} + \pi_{2jk} \cdot Z_{2jk} + \ldots + \pi_{Pjk} \cdot Z_{Pjk}
\]

**Level 2**

\[
\pi_{pjk} = b_{p0k} + \sum_{q=1}^{Q_p} b_{pqk} \cdot X_{qjk} + r_{pjk}
\]

**Level 3**

\[
b_{pqk} = \gamma_{pq0} + \sum_{s=1}^{S_{pq}} \gamma_{pqrs} \cdot W_{sk} + u_{pqk}
\]
Stata example

- Let’s look at the relationship between traffic accidents and alcohol consumption (Fávero and Belfiore, 2017).

- We want to estimate the relationship between the number of traffic accidents and the consumption alcohol per person/day in the district, considering differences in cities and states.
Data description (file “TrafficAccidents.dta”)

```
. desc

obs: 1,062  
vars: 5  
size: 15,930

---------------------------------------------------------------------------------------
<table>
<thead>
<tr>
<th>variable</th>
<th>storage</th>
<th>display</th>
<th>variable label</th>
</tr>
</thead>
<tbody>
<tr>
<td>state</td>
<td>str2</td>
<td>%2s</td>
<td>state k (level 3)</td>
</tr>
<tr>
<td>city</td>
<td>int</td>
<td>%8.0g</td>
<td>city j (level 2)</td>
</tr>
<tr>
<td>district</td>
<td>int</td>
<td>%8.0g</td>
<td>municipal district i (level 1)</td>
</tr>
<tr>
<td>accidents</td>
<td>byte</td>
<td>%8.0g</td>
<td>number of traffic accidents in the district over the last year</td>
</tr>
<tr>
<td>alcohol</td>
<td>float</td>
<td>%9.2f</td>
<td>average consumption of alcohol per person/day in the district (in grams)</td>
</tr>
</tbody>
</table>
---------------------------------------------------------------------------------------
```

Sorted by:
## Implementation

**Data tabulation**

```
. tab accidents

<table>
<thead>
<tr>
<th>number of traffic accidents</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>26</td>
<td>2.45</td>
<td>2.45</td>
</tr>
<tr>
<td>1</td>
<td>168</td>
<td>15.82</td>
<td>18.27</td>
</tr>
<tr>
<td>2</td>
<td>308</td>
<td>29.00</td>
<td>47.27</td>
</tr>
<tr>
<td>3</td>
<td>208</td>
<td>19.59</td>
<td>66.85</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>10.36</td>
<td>77.21</td>
</tr>
<tr>
<td>5</td>
<td>62</td>
<td>5.84</td>
<td>83.05</td>
</tr>
<tr>
<td>6</td>
<td>38</td>
<td>3.58</td>
<td>86.63</td>
</tr>
<tr>
<td>7</td>
<td>27</td>
<td>2.54</td>
<td>89.17</td>
</tr>
<tr>
<td>8</td>
<td>22</td>
<td>2.07</td>
<td>91.24</td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td>1.60</td>
<td>92.84</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td>1.41</td>
<td>94.26</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>0.94</td>
<td>95.20</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>0.94</td>
<td>96.14</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>0.47</td>
<td>96.61</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>0.38</td>
<td>96.99</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>0.38</td>
<td>97.36</td>
</tr>
<tr>
<td>16</td>
<td>6</td>
<td>0.56</td>
<td>97.93</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td>0.47</td>
<td>98.40</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>0.19</td>
<td>98.59</td>
</tr>
<tr>
<td>21</td>
<td>2</td>
<td>0.19</td>
<td>98.78</td>
</tr>
<tr>
<td>22</td>
<td>4</td>
<td>0.38</td>
<td>99.15</td>
</tr>
<tr>
<td>23</td>
<td>5</td>
<td>0.47</td>
<td>99.62</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>0.09</td>
<td>99.72</td>
</tr>
<tr>
<td>32</td>
<td>1</td>
<td>0.09</td>
<td>99.81</td>
</tr>
<tr>
<td>33</td>
<td>2</td>
<td>0.19</td>
<td>100.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1,062</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>
```
Histogram

```
. hist accidents, discrete freq
 (start=0, width=1)
```
Mean and variance (possible existence of overdispersion)

```
. tabstat accidents, stats(mean var)

<table>
<thead>
<tr>
<th>variable</th>
<th>mean</th>
<th>variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>accidents</td>
<td>3.812618</td>
<td>15.24007</td>
</tr>
</tbody>
</table>
```
Proposed Model

\[
\ln(\text{accidents}_{ijk}) = \pi_{0\,jk} + \pi_{1\,jk} \cdot \text{alcohol}_{jk}
\]

\[
\pi_{0\,jk} = b_{00\,k} + r_{0\,jk}
\]

\[
\pi_{1\,jk} = b_{10\,k}
\]

\[
b_{00\,k} = \gamma_{000} + u_{00\,k}
\]

\[
b_{10\,k} = \gamma_{100}
\]

\[
\ln(\text{accidents}_{ijk}) = \gamma_{000} + \gamma_{100} \cdot \text{alcohol}_{jk} + u_{00\,k} + r_{0\,jk}
\]
Multilevel Mixed-Effects Poisson Model

```
. meglm accidents alcohol || state: || city: , family(poisson) link(log) nolog

Mixed-effects GLM                  Number of obs =  1062
Family:                  Poisson
Link:                      log

<table>
<thead>
<tr>
<th>No. of</th>
<th>Observations per Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Variable</td>
<td>Groups Minimum Average Maximum</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>state</td>
<td>27          1       39.3         95</td>
</tr>
<tr>
<td>city</td>
<td>235          1        4.5         13</td>
</tr>
</tbody>
</table>

Integration method: mvaghermite
Integration points = 7
Wald chi2(1)       =  5.60
Log likelihood  = -2295.9047  Prob > chi2 =  0.0180

------------------------------------------------------------------------------
accidents            Coef.     Std. Err.     z  P>|z|     [95% Conf. Interval]
------------------------------------------------------------------------------
alcohol     |   .0478279    .020216     2.37   0.018     .0082053    .0874506
_cons      |   .7293659   .2638594     2.76   0.006     .2122111    1.246521
------------------------------------------------------------------------------
state      | var(_cons)|   .3857761   .12319                      .2063103    .7213563
------------------------------------------------------------------------------
state>city | var(_cons)|   .0829691   .0142976                       .059188    .1163053
------------------------------------------------------------------------------
LR test vs. Poisson regression:   chi2(2) =  1279.65  Prob > chi2 = 0.0000

. estimates store mepoisson
```
Multilevel Mixed-Effects Negative Binomial Model

```
. meglm accidents alcohol || state: || city: , family(nbinomial) link(log) nolog

Mixed-effects GLM Family: negative binomial
Link: log

Number of obs = 1062

Family: negative binomial
Link: log

Overdispersion: mean

<table>
<thead>
<tr>
<th>No. of Observations per Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Variable</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>state</td>
</tr>
<tr>
<td>city</td>
</tr>
</tbody>
</table>

Integration method: mvaghermite
Integration points = 7

Log likelihood = -2234.3721
Wald chi2(1) = 4.38
Prob > chi2 = 0.0363

| accidents | Coef. | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|-----------|-------|-----------|-------|------|----------------------------|
| alcohol   | .0466768 | .0222975  | 2.09  | 0.036 | .0029746    to    .0903791 |
| _cons     | .7538477 | .2843403  | 2.65  | 0.008 | .196551     to    1.311144 |

| /lnalpha  | -2.258241 | .1355339  | -16.66 | 0.000 | -2.523883    to    -1.9926 |

state var(_cons) | .3775391 | .1205934  | .2018698  | .7060775 |
state>city var(_cons) | .0613878 | .0138809  | .0394104  | .0956212 |

LR test vs. nbinomial regression: chi2(2) = 508.99
Prob > chi2 = 0.0000
```

estimates store menegbin
Likelihood-ratio test

```
. lrtest mepoisson menegbin

Likelihood-ratio test (Assumption: mepoisson nested in menegbin)
LR chi2(1) = 123.07
Prob > chi2 = 0.0000
```
<table>
<thead>
<tr>
<th>state</th>
<th>city</th>
<th>accidents</th>
<th>lambda</th>
<th>u00</th>
<th>r0</th>
</tr>
</thead>
<tbody>
<tr>
<td>MT</td>
<td>148</td>
<td>2</td>
<td>1.600369</td>
<td>-0.815816</td>
<td>-0.0064477</td>
</tr>
<tr>
<td>MT</td>
<td>148</td>
<td>2</td>
<td>1.63053</td>
<td>-0.815816</td>
<td>-0.0064477</td>
</tr>
<tr>
<td>MT</td>
<td>148</td>
<td>1</td>
<td>1.63053</td>
<td>-0.815816</td>
<td>-0.0064477</td>
</tr>
<tr>
<td>MT</td>
<td>148</td>
<td>1</td>
<td>1.585499</td>
<td>-0.815816</td>
<td>-0.0064477</td>
</tr>
<tr>
<td>MT</td>
<td>148</td>
<td>2</td>
<td>1.499133</td>
<td>-0.815816</td>
<td>-0.0064477</td>
</tr>
</tbody>
</table>

predict lambda

predict u00 r0, remeans

list state city accidents lambda u00 r0 if state=="MT", sepby(city)
quietly nbreg accidents alcohol

predict lambdatrad

graph twoway lfit accidents accidents ||
mspline lambda accidents || mspline lambdatrad accidents ||, legend(label(2 "GLLAMM Multilevel Negative Binomial")
label(3 "GLM Traditional Negative Binomial "))
Deep interactions

Methods of parameter estimations

Sample clustering

Estimation of models with better adjustment between real and fitted values!

Andrew Gelman
Multilevel Conference, 31 Out 2015, Columbia University, NYC.

- Stata 15 has a full command suite for the estimation of these models.

- Several research opportunities, both in theoretical and applied terms, in areas such as microeconomics, finance, transportation, real estate, leisure, ecology, education, and health.
REFERENCES


We must widen the circle of our love till it embraces the whole village; the village in its turn must take into its fold the district; the district the province; and so on, until the scope of our love becomes co-terminous with the world.
Thank you!

Multilevel Mixed-Effects Generalized Linear Models in Stata

Prof. Dr. Luiz Paulo Fávero - lpfavero@usp.br

Prof. Dr. Matheus Albergaria - matheus.albergaria@usp.br
*Stata do-file for the Presentation "Multilevel Mixed-Effects Generalized Linear Models in Stata", by Luiz Paulo Fávero and Matheus Albergaria
*2017 Brazilian Stata Users Group Meeting
*Universidade de São Paulo (USP), São Paulo, Brazil
*December 8th, 2017

*This do-file was written by Luiz Paulo Fávero and Matheus Albergaria
*The data file is "TrafficAccidents.dta". For more details, see:

+Open Dataset
use C:\TrafficAccidents.dta

+Data Description
desc

+Data Tabulation	
tab accidents
hist accidents, discrete freq
*Descriptive Statistics
  tabstat accidents, stats(mean var)

*Multilevel Mixed-Effects Count Models Estimation
  meglm accidents alcohol || state: || city: , family(poisson) link(log) nolog
  estimates store mepoisson
  meglm accidents alcohol || state: || city: , family(nbinomial) link(log) nolog
  estimates store menegbin

*Fitting Distinct Models
  lrtest mepoisson menegbin

predict lambda

predict u00 r0, remeans
  list state city accidents lambda u00 r0 if state == "MT", seby(city)
  quietly nbreg accidents alcohol
  predict lambdatrad
  graph twoway lfit accidents accidents || mspline lambda accidents || ///
  mspline lambdatrad accidents ||, ///
  legend(label(2 "GLLAMM Multilevel Negative Binomial") ///
  label(3 "GLM Traditional Negative Binomial"))