10.4 Hierarchical Models
10.5 Model Fitting and Inference for GLMMs

Melia Fernandez and Matthew Staley

May 12, 2015

Overview

1. Hierarchical (Multilevel) Models
   - What is hierarchical data?
   - How does it work?
   - Example of Multilevel Model
2. Model Fitting and Inference for GLMMs

What is hierarchical data?

- Data is hierarchical if it has nested categories that data falls into
- e.g.: A study on student performance on an exam measures whether a student passed.
  - Students nested within *schools* and a multilevel model can study variability among schools as well as just among students.
- These are also referred to as multilevel models.
  - Level 1 in the example above would be measurements at the student level.
  - Level 2 in this example would be measurement at the school level.

How does it work?

Suppose we want to look at student advancement in school from one grade to the next. For student t in school i, y_{it} = success (passing to the next grade) or fails.

At the student level, we have k explanatory variables: $\{x_{it1}, \ldots, x_{itk}\}$. Given student t in school i, x_{it1} could be gender, x_{it2} could be race, etc.

Student level

$$\operatorname{logit}[P(y_{it} = 1)] = \alpha_i + \beta_1 x_{it1} + \beta_2 x_{it2} + \ldots + \beta_k x_{itk}$$

How does it work? (continued)

At the school level i, we have ℓ explanatory variables: $\{w_{i1}, \ldots, w_{i\ell}\}$
Notice that these variables only vary at the school level, there is no t index on the variables.
For example, w_{i1} could represent expenditure per student for a given school, w_{i2} could represent mean socioeconomic status for a school, etc.

School level

$$\alpha_i = u_i + \alpha + \gamma_1 w_{i1} + \ldots + \gamma_\ell w_{i\ell}$$

The level two model provides a linear predictor for the level two term, α_i in the level one model.

How does it work? (continued)

We then substitute the level two model into the level one model to obtain

Multilevel model

$$\text{logit}[P(y_{it}=1)] = u_i + \alpha + \gamma_1 w_{i1} + \ldots + \gamma_\ell w_{i\ell} + \beta_1 x_{it1} + \beta_2 x_{it2} + \ldots + \beta_k x_{itk}$$

This is a logistic normal model with random intercept u_i.
Random effect can be included at either level, especially with higher number of observations at the lower level.
This level does not show interaction, but interaction can be included in these models as well.

Example of a Multilevel Model

The example here is based on a study involving Mathematical achievement data from Bryk and Raudenbush (1992) and Singer (1998) (as interpreted and presented by John Fox (2002)).

The individual level for this model, for individual j in school i is

$$\text{mathach}_{ij} = \alpha_{0i} + \alpha_{1i}\text{cses}_{ij}$$

At the school level, we will look at the dependence on sector (public vs. private Catholic) and average level of SES in the schools. So for the public sector school i coefficients and intercepts, we have

$$\alpha_{0i} = \gamma_{00} + \gamma_{01}\text{meanses}_i + \gamma_{02}\text{sector}_i + u_{0i}$$

and for Catholic sector, we have

$$\alpha_{1i} = \gamma_{10} + \gamma_{11}\text{meanses}_i + \gamma_{12}\text{sector}_i + u_{1i}$$

Example (continued)

The α_{0i} and α_{1i} are group dependent regression coefficients.
We will use them in the mathach$_{ij}$ formula in order to obtain a more advanced cross level interaction effects model.
It will result in a product interaction effect between the individual level and the sector of the data.

Example (continued)

We insert the second level equations into our individual level equation in order to find our mixed model.

$$\text{mathach}_{ij} = \gamma_{00} + \gamma_{01}\text{meanses}_i + \gamma_{02}\text{sector}_i + u_{0i} +$$

$$(\gamma_{10} + \gamma_{11}\text{meanses}_i + \gamma_{12}\text{sector}_i + u_{1j})\text{cses}_{ij}$$

Rearranging and expanding,

$$\text{mathach}_{ij} = \gamma_{00} + \gamma_{01}\text{meanses}_i + \gamma_{02}\text{sector}_i +$$

$$\gamma_{10}\text{cses}_{ij} + \gamma_{11}\text{meanses}_i\text{cses}_{ij} +$$

$$\gamma_{12}\text{sector}_i\text{cses}_{ij} + u_{0i} + u_{1j}\text{cses}_{ij}$$

Example (continued)

```
> Bryk[sample15, ]
school     ses mathach      meanses        cses   sector
142     1317  0.932   8.961  0.345333333  0.58666667 Catholic
646     1942 -0.568  20.514  0.682000000 -1.25000000   Public
684     1946 -0.758  16.863  0.004051282 -0.76205128   Public
1917    3152 -0.778   8.789  0.031038462 -0.80903846   Public
2391    3688  0.062  19.719  0.405023256 -0.34302326 Catholic
2439    3705  0.112  10.921  0.234666667 -0.12266667 Catholic
3051    4325  0.252  15.476 -0.049132075  0.30113208   Public
3056    4325  0.202  14.259 -0.049132075  0.25113208   Public
3726    5619  0.932  22.517  0.420333333  0.51166667 Catholic
3913    5720 -0.078  15.896  0.032566038 -0.11056604 Catholic
4037    5762 -1.118   1.083 -1.193945946  0.07594595   Public
4069    5783  0.492  17.287  0.173034483  0.31896552   Public
4090    5815  0.182   5.443 -0.680000000  0.86200000   Public
4109    5819  1.212  16.246  0.181600000  1.03040000   Public
4537    6415 -0.068  10.360 -0.187259259  0.11925926   Public
```

Example (continued)

```
> library(nlme)
>
> # catholic list
> cat.list <- lmList(mathach ~ ses | school, subset = sector =
>
> # public list
> pub.list <- lmList(mathach ~ ses | school, subset = sector =
> Bryk$sector = factor(Bryk$sector, levels = c('Public', 'Cath
> contrasts(Bryk$sector)
         Catholic
Public          0
Catholic        1
```

Example (continued)

```
> bryk.lme.1 = lme(mathach ~ meanses*cses + sector*cses, rando
> summary(bryk.lme.1)
Linear mixed-effects model fit by REML
 Data: Bryk
       AIC      BIC    logLik
  46523.66 46592.45 -23251.83

Random effects:
 Formula: ~cses | school
 Structure: General positive-definite, Log-Cholesky parametriz
            StdDev    Corr
(Intercept) 1.5426166 (Intr)
cses        0.3182103 0.391
Residual    6.0597952
```

Example (continued)

```
Fixed effects: mathach ~ meanses * cses + sector * cses
                         Value Std.Error   DF  t-value p-value
(Intercept)          12.127931 0.1992921 7022 60.85505   0e+00
meanses               5.332875 0.3691687  157 14.44563   0e+00
cses                  2.945041 0.1556008 7022 18.92690   0e+00
sectorCatholic        1.226579 0.3062736  157  4.00485   1e-04
meanses:cses          1.039229 0.2988976 7022  3.47687   5e-04
cses:sectorCatholic  -1.642674 0.2397803 7022 -6.85074   0e+00
 Correlation:
                     (Intr) meanss cses   sctrCt mnss:c
meanses               0.256
cses                  0.075  0.019
sectorCatholic       -0.699 -0.356 -0.053
meanses:cses          0.019  0.074  0.293 -0.026
cses:sectorCatholic  -0.052 -0.027 -0.696  0.077 -0.351
```

Example (continued)

```
Standardized Within-Group Residuals:
       Min         Q1        Med         Q3        Max
-3.1592610 -0.7231893  0.0170471  0.7544512  2.9582207

Number of Observations: 7185
Number of Groups: 160
```

Example (continued)

Now, referring back to the equation,

$$\text{mathach}_{ij} = \gamma_{00} + \gamma_{01}\text{meanses}_i + \gamma_{02}\text{sector}_i +$$

$$\gamma_{10}\text{cses}_{ij} + \gamma_{11}\text{meanses}_i\text{cses}_{ij} +$$

$$\gamma_{12}\text{sector}_i\text{cses}_{ij} + u_{0i} + u_{1j}\text{cses}_{ij}$$

and filling in the respective coefficients, we get

$$\text{mathach}_{ij} = 12.128 + 5.333\text{meanses}_i + 1.227\text{sector}_i +$$

$$2.945\text{cses}_{ij} + 1.039\text{meanses}_i\text{cses}_{ij}$$

$$-1.643\text{sector}_i\text{cses}_{ij} + 1.543 + 0.318\text{cses}_{ij}$$

Notice that the u_{0i} and u_{1i} come from the random effects.

Model Fitting

Fitting GLMMs are not always simple to fit.
Model fitting treats $\{y_{ij}\}$ as independent over i and t, conditional on $\{u_i\}$.
In practice, we do not know u_i and the model treats them as unobserved random variables.
Likelihood function for GLMM refers to fixed effect parameters and σ of the $N(0, \sigma)$ random effects distribution. Software eliminates $\{u_i\}$ by first forming the function as if $\{u_i\}$ values were known, then averages that function with respect to the $N(0, \sigma)$ distribution of $\{u_i\}$.
The numerical methods for averaging that function are intensive. Two main approaches are the Gauss-Hermite quuadrature and the Monte Carlo methods.

References

Alan Agresti (2007)
An Introduction to Categorical Data Analysis, 2nd Edition.
pp 313-318. John Wiley & Sons, Inc.

John Fox (2002)
Linear Mixed Models.
http://socserv.mcmaster.ca/jfox/Books/Companion-1E/appendix-mixed-models.pdf

The End