

Multilevel Modeling in SAS

Daniel J. Bauer
University of North Carolina at Chapel Hill

Talk Outline

- ▶ When to use multilevel models.
- ▶ Overview of SAS Procedures for fitting multilevel models.
- ▶ Fitting multilevel linear models in SAS with PROC MIXED
 - ▶ Example: Educational achievement data clustered within schools
 - ▶ Example: Intensive short-term longitudinal data on psychophysiology.
- ▶ Fitting multilevel generalized linear models in SAS with PROC GLIMMIX
 - ▶ Example: Tracking vocabulary growth in a longitudinal study.

▶ 2

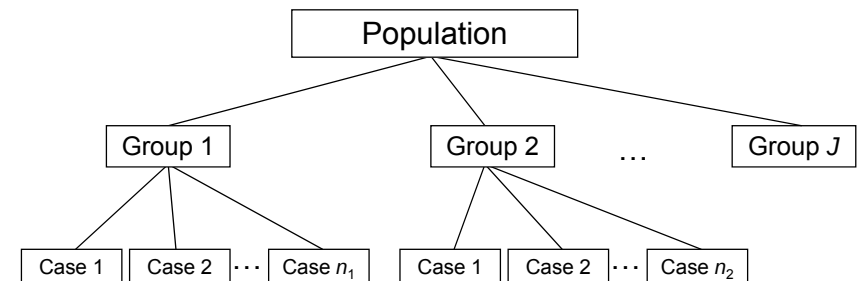
When to Use Multilevel Models

- ▶ Multilevel models (a subclass of mixed models) are employed when we have nested (or clustered) data.
- ▶ Nested data typically comes in one of two forms, hierarchical or longitudinal.
- ▶ Nested data structures are characterized by a lack of independence of observations, violating a key assumption of the general(ized) linear model

▶ 3

Hierarchical Data Structures

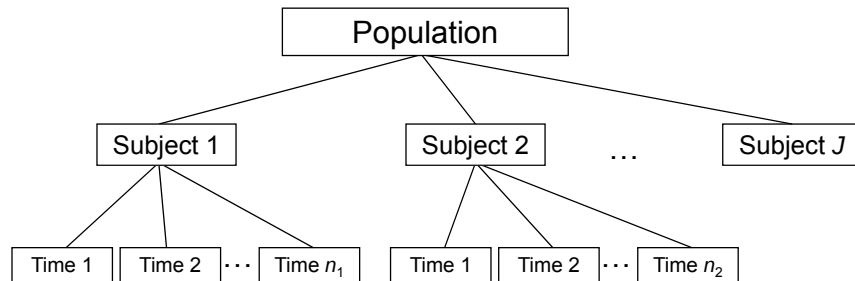
- ▶ Hierarchical data structures are those in which multiple observations are collected for each independent sampling unit, such as individuals within groups.



▶ 4

Longitudinal Data Structures

- ▶ Longitudinal data structures arise when the same units are sampled repeatedly over time.



▶ 5

Dependence in Nested Data

- ▶ Because multiple observations are collected within the same unit, this produces dependence in the data.
 - ▶ With school-based sampling, students attending the same school will have more similar academic outcomes than students attending different schools.
 - ▶ In a longitudinal psychophysiology study, some people will have consistently higher heart rates than others.
- ▶ Multilevel models provide one way to model this dependence.
 - ▶ Dependence is accounted for through the introduction of random effects (i.e., allowing certain effects to vary over units).

▶ 6

Modeling Dependence

- ▶ Suppose we sample multiple clusters j and within each cluster measure multiple individuals i on an outcome variable y and a predictor variable x .

- ▶ We can formulate a multilevel model as:

Level 1 Model (Within Cluster):

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + r_{ij} \quad r_{ij} \sim N(0, \sigma^2)$$

Level 2 Model (Between Cluster):

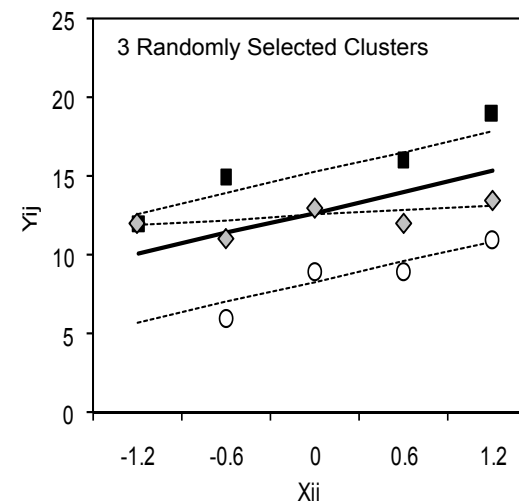
$$\begin{pmatrix} \beta_{0j} \\ \beta_{1j} \end{pmatrix} = \begin{pmatrix} \gamma_{00} \\ \gamma_{10} \end{pmatrix} + \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \quad \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{01} & \tau_{11} \end{pmatrix} \right]$$

Combined Model:

$$y_{ij} = (\gamma_{00} + \gamma_{10}x_{ij}) + (u_{0j} + u_{1j}x_{ij}) + r_{ij}$$

▶ 7

Visual Representation



Level 1 Model (Within)

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + r_{ij}$$

Level 2 Model (Between)

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

Combined Model

$$y_{ij} = (\gamma_{00} + \gamma_{10}x_{ij}) + (u_{0j} + u_{1j}x_{ij}) + r_{ij}$$

▶ 8

Modeling Dependence

Level 1 Model (Within Cluster):

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + r_{ij} \quad r_{ij} \sim N(0, \sigma^2)$$

Level 2 Model (Between Cluster):

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \end{aligned} \quad \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{01} & \tau_{11} \end{pmatrix} \right]$$

Combined Model:

$$y_{ij} = (\gamma_{00} + \gamma_{10}x_{ij}) + (u_{0j} + u_{1j}x_{ij}) + r_{ij}$$

Implied Variance/Covariance Structure:

$$\text{VAR}(y_{ij}) = \tau_{00} + 2\tau_{10}x_{ij} + \tau_{11}x_{ij}^2 + \sigma^2$$

$$\text{COV}(y_{ij}, y_{i'j}) = \tau_{00} + \tau_{10}(x_{ij} + x_{i'j}) + \tau_{11}x_{ij}x_{i'j} \quad \text{where } i \neq i'$$

▶ 9

Overview of SAS Procedures for Fitting Multilevel Models

- ▶ The MIXED procedure
 - ▶ Can be used to fit multilevel linear models
- ▶ The GLIMMIX procedure
 - ▶ Can be used to fit multilevel generalized linear models
 - ▶ Supersedes old %GLIMMIX macro
 - ▶ A downloadable addition to 9.1, standard in SAS/STAT with 9.2
 - ▶ 9.1 version includes only pseudo-(or quasi-)likelihood methods of estimation. 9.2 version adds ML using a Laplace approximation or quadrature.
- ▶ The NLMIXED procedure
 - ▶ Can be used to fit multilevel generalized linear models or multilevel *nonlinear* models.
 - ▶ Performs ML using Laplace or quadrature, but slower than GLIMMIX for multilevel generalized models.

▶ 10

The MIXED Procedure

```
PROC MIXED options;  
  CLASS classification variables;  
  MODEL outcome = fixed-effect predictors / options;  
  RANDOM random effects / options;  
  REPEATED / options;  
RUN;
```

- ▶ For PROC MIXED statement, an important option is METHOD, used to toggle between REML (default) and ML estimation.
- ▶ For MODEL statement, important options are SOLUTION (to get fixed effect estimates) DDFM (determines t- or F-dist for testing fixed effects) and OUTPRED (generates predicted values and residuals for observations)
- ▶ For RANDOM and REPEATED statements, important options are SUBJECT (to indicate the ID variable of the upper level units), and TYPE (to specify covariance matrix of random effects or Level 1 residuals, respectively).

▶ 11

The GLIMMIX Procedure

```
PROC GLIMMIX options;  
  CLASS classification variables;  
  MODEL outcome = fixed-effect predictors / options;  
  RANDOM random effects / options;  
  REPEATED / options;  
RUN;
```

- ▶ Similar options to those described for PROC MIXED, but several important additional options.
- ▶ For GLIMMIX statement, use METHOD option to choose between pseudo-likelihood estimators or ML estimators.
- ▶ For MODEL statement, important additional options are DIST (to designate response distribution) and LINK (to designate link function)

▶ 12

The NLMIXED Procedure

```

PROC NLMIXED options;
  PARMS list of parameters in model and start values;
  Program Statements use these to define the
  parameters of the response distribution in terms of
  model parameters;
  MODEL outcome ~ distribution(parameters);
  RANDOM random effects ~
    normal([means],[covariance matrix])
    subject = Level 2 ID variable;

RUN;
  
```

- ▶ Note: No CLASS statement, so data must be sorted by the Level 2 ID variable.

▶ 13

Example 1: Relationship between SES and Educational Achievement

- ▶ Sample is selected from the High School and Beyond (HSB) survey data set:
 - ▶ 7185 students from 160 schools.
 - ▶ 14 and 67 students per school.
- ▶ Interest is in:
 - ▶ Within-school relation between SES and Math Achievement (i.e., achievement inequities)
 - ▶ Whether this relation differs across schools at least partly as a function of school sector (0=Public, 1=Private).
- ▶ Data Features:
 - ▶ Continuous outcome
 - ▶ Classical clustered data structure
 - ▶ Reasonably large number of clusters (N=160)

▶ 14

Example 1: Relationship between SES and Educational Achievement

- ▶ Our model is

Level 1 Model (Within Cluster):

$$Math_{ij} = \beta_{0j} + \beta_{1j}cSES_{ij} + r_{ij} \quad r_{ij} \sim N(0, \sigma^2)$$

Level 2 Model (Between Cluster):

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}Sector_j + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}Sector_j + u_{1j} \end{aligned} \quad \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \\ & \tau_{11} \end{pmatrix} \right]$$

Combined Model:

$$\begin{aligned} Math_{ij} &= (\gamma_{00} + \gamma_{01}Sector_j + u_{0j}) + (\gamma_{10} + \gamma_{11}Sector_j + u_{1j})cSES_{ij} + r_{ij} \\ &= \underbrace{(\gamma_{00} + \gamma_{01}Sector_j + \gamma_{10}cSES_{ij} + \gamma_{11}Sector_j cSES_{ij})}_{\text{Fixed effects (MODEL)}} + \underbrace{(u_{0j} + u_{1j}cSES_{ij})}_{\text{Random effects (RANDOM)}} + r_{ij} \end{aligned}$$

▶ 15

Example 2: Race Differences in Cardiovascular Health

- ▶ Sample obtained from Laura Richman at Duke University:
 - ▶ 30 Caucasian Americans and 31 African Americans from Durham, NC
 - ▶ Starting in morning, 3 HR readings per hour while awake, then 2 per hour while asleep, with # assessments ranging from 6-60.
- ▶ Interest is in:
 - ▶ Whether HR level and changes over day and night differ between blacks and whites.
 - ▶ Whether race differences are maintained after accounting for differential experience of discrimination.
- ▶ Data Features:
 - ▶ Continuous outcome
 - ▶ Repeated measures data structure, obtained at many short time intervals with little measurement error
 - ▶ Modest number of independent sampling units (N=62)

▶ 16

Example 2, Model 1: Time Trends

Level 1 Model (Within Person):

$$HR_{ij} = \beta_{0j} + \beta_{1j}timewake_{ij} + \beta_{2j}timesleep_{ij} + r_{ij} \quad \mathbf{r}_j \sim N(\mathbf{0}, \Sigma_j)$$

$$\Sigma_j[i, i'] = \sigma^2 \rho^{(time_{ij} - time_{ij'})} \quad (\text{Continuous-time autoregressive error structure})$$

Level 2 Model (Between Person):

$$\begin{pmatrix} \beta_{0j} \\ \beta_{1j} \\ \beta_{2j} \end{pmatrix} = \begin{pmatrix} \gamma_{00} + u_{0j} \\ \gamma_{10} + u_{1j} \\ \gamma_{20} + u_{2j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & & \\ \tau_{01} & \tau_{11} & \\ \tau_{02} & \tau_{12} & \tau_{22} \end{pmatrix} \right]$$

Combined Model:

$$HR_{ij} = (\gamma_{00} + \gamma_{10}timewake_{ij} + \gamma_{20}timesleep_{ij}) + (u_{0j} + u_{1j}timewake_{ij} + u_{2j}timesleep_{ij}) + r_{ij}$$

▶ 17

Example 2, Model 2: Race Differences

Level 1 Model (Within Person):

$$HR_{ij} = \beta_{0j} + \beta_{1j}timewake_{ij} + \beta_{2j}timesleep_{ij} + r_{ij}$$

$$\Sigma_j[i, i'] = \sigma^2 \rho^{(time_{ij} - time_{ij'})} \quad (\text{Continuous-time autoregressive error structure})$$

Level 2 Model (Between Person):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Black_j + \gamma_{03}BMI_j + \gamma_{04}Male_j + \gamma_{05}Age_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Black_j + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}Black_j + u_{2j}$$

Combined Model: $HR_{ij} = (\gamma_{00} + \gamma_{01}Black_j + \gamma_{03}BMI_j + \gamma_{04}Male_j + \gamma_{05}Age_j + \gamma_{10}timewake_{ij} + \gamma_{11}Black_jtimewake_{ij} + \gamma_{20}timesleep_{ij} + \gamma_{21}Black_jtimesleep_{ij}) + (u_{0j} + u_{1j}timewake_{ij} + u_{2j}timesleep_{ij}) + r_{ij}$

▶ 18

Example 2, Model 3: Controlling for Perceived Discrimination

Level 1 Model (Within Person):

$$HR_{ij} = \beta_{0j} + \beta_{1j}timewake_{ij} + \beta_{2j}timesleep_{ij} + r_{ij}$$

$$\Sigma_j[i, i'] = \sigma^2 \rho^{(time_{ij} - time_{ij'})} \quad (\text{Continuous-time autoregressive error structure})$$

Level 2 Model (Between Person):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Black_j + \gamma_{02}PD_j + \gamma_{03}BMI_j + \gamma_{04}Male_j + \gamma_{05}Age_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Black_j + \gamma_{12}PD_j + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}Black_j + \gamma_{22}PD_j + u_{2j}$$

Combined Model:

BIG!

▶ 19

Example 3: Vocabulary Development

- ▶ Sample obtained from Huttenlocher et al. (1991) pools participants from two studies
 - ▶ Study 1 (N=11): infants observed for 5 hours beginning at 12 months of age every 2 months until 26 months old (Group=0)
 - ▶ Study 2 (N=11): infants observed for 3 hours beginning at 12 months of age every 4 months until 24 months old (Group=1)
- ▶ Interest is in:
 - ▶ Gender differences in vocabulary development and the effect of early language input from mother (# of unique words spoken to infant).
- ▶ Data Features:
 - ▶ Count outcome (# of unique words spoken by infant)
 - ▶ Differential 'exposure' across studies (5 hours versus 3 hours)
 - ▶ Small number of independent sampling units (N=22)

▶ 20

Example 3: Model 1, Time Trends

► Our model is

Level 1 Model (Within Person):

$$\ln(\mu_{ij}) = \ln(hours_{ij}) + \beta_{0j} + \beta_{1j}Age_{ij} + \beta_{2j}Age_{ij}^2$$

$$Vocab_{ij} \sim \text{Poisson}(\mu_{ij})$$

$$V(Vocab_{ij}) = \sigma^2 \mu_{ij}$$

Level 2 Model (Between Person):

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \\ \beta_{2j} &= \gamma_{20} \end{aligned} \quad \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \\ & \tau_{11} \end{pmatrix} \right]$$

Overdispersion
(RANDOM_residual_)

Combined Model:

$$\ln(\mu_{ij}) = \underbrace{\ln(hours_{ij})}_{\text{Offset (MODEL/OFFSET=)}} + \underbrace{(\gamma_{00} + \gamma_{10}Age_{ij} + \gamma_{20}Age_{ij}^2)}_{\text{Fixed effects (MODEL)}} + \underbrace{(u_{0j} + u_{1j}Age_{ij})}_{\text{Random effects (RANDOM)}}$$

Example 3: Model 2, Predictors

► Our model is

Level 1 Model (Within Person):

$$\ln(\mu_{ij}) = \ln(hours_{ij}) + \beta_{0j} + \beta_{1j}Age_{ij} + \beta_{2j}Age_{ij}^2$$

$$Vocab_{ij} \sim \text{Poisson}(\mu_{ij})$$

$$V(Vocab_{ij}) = \sigma^2 \mu_{ij}$$

Level 2 Model (Between Person):

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}Female_j + \gamma_{02} \ln(MomSpeak_j) + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \\ \beta_{2j} &= \gamma_{20} \end{aligned} \quad \begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \\ & \tau_{11} \end{pmatrix} \right]$$

Combined Model:

$$\ln(\mu_{ij}) = \ln(hours_{ij}) + \gamma_{00} + \gamma_{01}Female_j + \gamma_{02} \ln(MomSpeak_j) + \gamma_{10}Age_{ij} + \gamma_{20}Age_{ij}^2 + u_{0j} + u_{1j}Age_{ij}$$