The R Statistical Computing Environment
Basics and Beyond
Mixed-Effects Models

John Fox

McMaster University

ICPSR/Berkeley 2016

The Linear Mixed-Effects Model

- The *Laird-Ware form* of the linear mixed model:

$$
\begin{aligned}
y_{ij} &= \beta_1 + \beta_2 x_{2ij} + \cdots + \beta_p x_{pij} + b_{1i} z_{1ij} + \cdots + b_{qi} z_{qij} + \varepsilon_{ij} \\
b_{ki} &\sim N(0, \psi_k^2), \mathrm{Cov}(b_{ki}, b_{k'i}) = \psi_{kk'} \\
& \quad b_{ki}, b_{k'i'} \text{ are independent for } i \neq i' \\
\varepsilon_{ij} &\sim N(0, \sigma^2 \lambda_{ijj}), \mathrm{Cov}(\varepsilon_{ij}, \varepsilon_{ij'}) = \sigma^2 \lambda_{ijj'} \\
& \quad \varepsilon_{ij}, \varepsilon_{i'j'} \text{ are independent for } i \neq i'
\end{aligned}
$$

The Linear Mixed-Effects Model

- where:
  - y_{ij} is the value of the response variable for the jth of n_i observations in the ith of M groups or clusters.
  - $\beta_1, \beta_2, \ldots, \beta_p$ are the fixed-effect coefficients, which are identical for all groups.
  - $x_{2ij}, \ldots, x_{pij}$ are the fixed-effect regressors for observation j in group i; there is also implicitly a constant regressor, $x_{1ij} = 1$.
  - $b_{1i}, \ldots, b_{qi}$ are the random-effect coefficients for group i, assumed to be multivariately normally distributed, independent of the random effects of other groups. The random effects, therefore, vary by group.
    - The b_{ik} are thought of as random variables, not as parameters, and are similar in this respect to the errors ε_{ij}.
  - $z_{1ij}, \ldots, z_{qij}$ are the random-effect regressors.
    - The z's are almost always a subset of the x's (and may include *all* of the x's).
    - When there is a random intercept term, $z_{1ij} = 1$.

The Linear Mixed-Effects Model

- and:
  - ψ_k^2 are the variances and $\psi_{kk'}$ the covariances among the random effects, assumed to be constant across groups.
    - In some applications, the ψ's are parametrized in terms of a smaller number of fundamental parameters.
  - ε_{ij} is the error for observation j in group i.
    - The errors for group i are assumed to be multivariately normally distributed, and independent of errors in other groups.
  - $\sigma^2 \lambda_{ijj'}$ are the covariances between errors in group i.
    - Generally, the $\lambda_{ijj'}$ are parametrized in terms of a few basic parameters, and their specific form depends upon context.
    - When observations are sampled independently within groups and are assumed to have constant error variance (as is typical in hierarchical models), $\lambda_{ijj} = 1$, $\lambda_{ijj'} = 0$ (for $j \neq j'$), and thus the only free parameter to estimate is the common error variance, σ^2.
    - If the observations in a "group" represent longitudinal data on a single individual, then the structure of the λ's may be specified to capture serial (i.e., over-time) dependencies among the errors.

Fitting Mixed Models in R

with the **nlme** and **lme4** packages

- In the **nlme** package (Pinheiro, Bates, DebRoy, and Sarkar):
  - `lme`: linear mixed-effects models with nested random effects; can model serially correlated errors.
  - `nlme`: nonlinear mixed-effects models.
- In the **lme4** package (Bates, Maechler, Bolker, and Walker):
  - `lmer`: linear mixed-effects models with nested or crossed random effects; no facility for serially correlated errors.
  - `glmer`: generalized-linear mixed-effects models.

A Mixed Model for the Exercise Data

Longitudinal Model

- A level-1 model specifying a linear "growth curve" for log exercise for each subject:

$$\text{log-exercise}_{ij} = \alpha_{0i} + \alpha_{1i}(\text{age}_{ij} - 8) + \varepsilon_{ij}$$

- Our interest in detecting differences in exercise histories between subjects and controls suggests the level-2 model

$$\alpha_{0i} = \gamma_{00} + \gamma_{01}\text{group}_i + \omega_{0i}$$
$$\alpha_{1i} = \gamma_{10} + \gamma_{11}\text{group}_i + \omega_{1i}$$

where group is a dummy variable coded 1 for subjects and 0 for controls.

A Mixed Model for the Exercise Data

Laird-Ware form of the Model

- Substituting the level-2 model into the level-1 model produces

$$\begin{aligned}\text{log-exercise}_{ij} &= (\gamma_{00} + \gamma_{01}\text{group}_i + \omega_{0i}) \\ &\quad + (\gamma_{10} + \gamma_{11}\text{group}_i + \omega_{1i})(\text{age}_{ij} - 8) + \varepsilon_{ij} \\ &= \gamma_{00} + \gamma_{01}\text{group}_i + \gamma_{10}(\text{age}_{ij} - 8) \\ &\quad + \gamma_{11}\text{group}_i \times (\text{age}_{ij} - 8) \\ &\quad + \omega_{0i} + \omega_{1i}(\text{age}_{ij} - 8) + \varepsilon_{ij}\end{aligned}$$

- in Laird-Ware form,

$$Y_{ij} = \beta_1 + \beta_2 x_{2ij} + \beta_3 x_{3ij} + \beta_4 x_{4ij} + \delta_{1i} + \delta_{2i} z_{2ij} + \varepsilon_{ij}$$

- Continuous first-order autoregressive process for the errors:

$$\text{Cor}(\varepsilon_{it}, \varepsilon_{i,t+s}) = \rho(s) = \phi^{|s|}$$

where the time-interval between observations, s, need not be an integer.

A Mixed Model for the Exercise Data

Specifying the Model in `lme`

- Using `lme` in the **nlme** package:

```
lme(log.exercise ~ I(age - 8)*group,
               random = ~ I(age - 8) | subject,
               correlation = corCAR1(form = ~ age |subject)
               data=Blackmoor)
```

A Mixed Model for the HSB Data

Hierarchical Model

- A "level-1" model for math achievement:

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

where $\text{cses}_{ij} = \text{ses}_{ij} - \overline{\text{ses}}_{i\cdot}$.

- Exploration of the data suggests the following "level-2" model:

$$\begin{aligned} \alpha_{0i} &= \gamma_{00} + \gamma_{01}\overline{\text{ses}}_{i\cdot} + \gamma_{02}\text{sector}_i + u_{0i} \\ \alpha_{1i} &= \gamma_{10} + \gamma_{11}\overline{\text{ses}}_{i\cdot} + \gamma_{12}\overline{\text{ses}}^2_{i\cdot} + \gamma_{13}\text{sector}_i + u_{1i} \end{aligned}$$

where sector is a dummy variable, coded 1 (say) for Catholic schools and 0 for public schools.

A Mixed Model for the HSB Data

Laird-Ware Form of the Model

- Substituting the school-level equation into the individual-level equation produces the *combined* or *composite model:*

$$\begin{aligned} \text{mathach}_{ij} &= (\gamma_{00} + \gamma_{01}\overline{\text{ses}}_{i\cdot} + \gamma_{02}\text{sector}_i + u_{0i}) \\ &\quad + \left(\gamma_{10} + \gamma_{11}\overline{\text{ses}}_{i\cdot} + \gamma_{12}\overline{\text{ses}}^2_{i\cdot} + \gamma_{13}\text{sector}_i + u_{1i}\right)\text{cses}_{ij} \\ &\quad + \varepsilon_{ij} \\ &= \gamma_{00} + \gamma_{01}\overline{\text{ses}}_{i\cdot} + \gamma_{02}\text{sector}_i + \gamma_{10}\text{cses}_{ij} \\ &\quad + \gamma_{11}\overline{\text{ses}}_{i\cdot} \times \text{cses}_{ij} + \gamma_{12}\overline{\text{ses}}^2_{i\cdot} \times \text{cses}_{ij} \\ &\quad + \gamma_{13}\text{sector}_i \times \text{cses}_{ij} \\ &\quad + u_{0i} + u_{1i}\text{cses}_{ij} + \varepsilon_{ij} \end{aligned}$$

A Mixed Model for the HSB Data

Laird-Ware Form of the Model

- Except for notation, this is a mixed model in Laird-Ware form, as we can see by replacing γ's with β's and u's with b's:

$$\begin{aligned} y_{ij} &= \beta_1 + \beta_2 x_{2ij} + \beta_3 x_{3ij} + \beta_4 x_{4ij} \\ &\quad + \beta_5 x_{5ij} + \beta_6 x_{6ij} + \beta_7 x_{7ij} \\ &\quad + b_{1i} + b_{2i} z_{2ij} + \varepsilon_{ij} \end{aligned}$$

A Mixed Model for the HSB Data

Laird-Ware Form of the Model

- Note that all explanatory variables in the Laird-Ware form of the model carry subscripts i for schools and j individuals within schools, even when the explanatory variable in question is constant within schools.
  - Thus, for example, $x_{2ij} = \overline{\text{ses}}_{i\cdot}$ (and so all individuals in the same school share a common value of school-mean SES).
- There is both a data-management issue here and a conceptual point:
  - With respect to data management, software that fits the Laird-Ware form of the model (such as the `lme` or `lmer` functions in R) requires that level-2 explanatory variables (here sector and school-mean SES, which are characteristics of schools) appear in the level-1 (i.e., student) data set.
  - The conceptual point is that the model can incorporate *contextual effects* — characteristics of the level-2 units can influence the level-1 response variable.

- Using `lmer` in the **lme4** package:

```
lmer(mathach ~ meanses + poly(meanses, 2, raw=TRUE):cses
          + sector*cses  + (cses | school), data=Bryk)
```

- Using `lme` in the **nlme** package:

```
lme(mathach ~ meanses + poly(meanses, 2, raw=TRUE):cses
                      + sector*cses
        random = ~ cses | school, data=Bryk)
```