

Agenda

1. More on multilevel R formulas
2. Multilevel generalized linear models (MLGLMs)
3. MLGLMs in R
4. Final project help

More on multilevel R formulas

Building a two-level model

**A simple
two-level
model**

$$S_{ik} \sim \text{Norm}(\mu_{ik}, \sigma)$$

$$\mu_{ik} = \beta_{0k} + \beta_1 \text{Age}_i + \beta_2 \text{Female}_i + \beta_{3k} \text{Black}_i$$

$$\beta_{0k} = \gamma_{00} + \gamma_{01} \text{TeacherBlack}_k + \gamma_{02} \text{PropBlack}_k + \eta_{0k}$$

$$\beta_{3k} = \gamma_{30} + \gamma_{31} \text{TeacherBlack}_k + \gamma_{32} \text{PropBlack}_k + \eta_{3k}$$

Expanded notation

$$\mu_{ik} = \beta_{0k} + \beta_1 \text{Age}_i + \beta_2 \text{Female}_i + \beta_{3k} \text{Black}_i$$

$$\beta_{0k} = \gamma_{00} + \gamma_{11} \text{TeacherBlack}_k + \gamma_{051} \text{PropBlack}_k + \eta_{0k}$$

$$\beta_{3k} = \gamma_{30} + \gamma_{31} \text{TeacherBlack}_k + \gamma_{32} \text{PropBlack}_k + \eta_{3k}$$

$$\begin{aligned} \mu_{ik} = & (\gamma_{00} + \gamma_{01} \text{TeacherBlack}_k + \gamma_{02} \text{PropBlack}_k + \eta_{0k}) \\ & + \beta_1 \text{Age}_i \\ & + \beta_2 \text{Female}_i \\ & + (\gamma_{30} + \gamma_{31} \text{TeacherBlack}_k + \gamma_{32} \text{PropBlack}_k + \eta_{3k}) \text{Black}_i \end{aligned}$$

Expanded notation

$$\begin{aligned}\mu_{ik} = & (\gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k + \eta_{0k}) \\ & + \beta_1Age_i \\ & + \beta_2Female_i \\ & + (\gamma_{30} + \gamma_{31}TeacherBlack_k + \gamma_{32}PropBlack_k + \eta_{0k})Black_i\end{aligned}$$

$$\begin{aligned}s_{ik} = & \gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k \\ & + \beta_1Age_i + \beta_2Female_i + \gamma_{30}Black_i \\ & + \gamma_{31}TeacherBlack_kBlack_i + \gamma_{32}PropBlack_kBlack_i \\ & + \eta_{0k} + \eta_{3k}Black_i + \varepsilon_i\end{aligned}$$

Expanded notation

Fixed effects

Explained variation in outcome variable.

Describes the way that outcome and predictor variables co-vary.

$$\begin{aligned} S_{ik} = & \gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k \\ & + \beta_1Age_i + \beta_2Female_i + \gamma_{30}Black_i \\ & + \gamma_{31}TeacherBlack_kBlack_i + \gamma_{32}PropBlack_kBlack_i \\ & + \eta_{0k} + \eta_{3k}Black_i + \varepsilon_i \end{aligned}$$

Random effects

Unexplained variation in outcome variable.

Described in terms of individual variability and different types of group variability.

Building an R formula

With level-two covariates,
interactions need to be specified

```
student_reading_score ~  
  student_age_s + student_female  
  student_re_black*teacher_re_black +  
  student_re_black*class_prop_black +  
  (1 + student_re_black | teacher_id)
```

$$\begin{aligned} S_{ik} = & \gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k \\ & + \beta_1Age_i + \beta_2Female_i + \gamma_{30}Black_i \\ & + \gamma_{31}TeacherBlack_kBlack_i + \gamma_{32}PropBlack_kBlack_i \\ & + \eta_{0k} + \eta_{3k}Black_i + \varepsilon_i \end{aligned}$$

Building an R formula

Outcome variable

```
student_reading_score ~  
  student_age_s + student_female  
  student_re_black*teacher_re_black +  
  student_re_black*class_prop_black +  
  (1 + student_re_black | teacher_id)
```

$$\begin{aligned} S_{ik} &= \gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k \\ &+ \beta_1Age_i + \beta_2Female_i + \gamma_{30}Black_i \\ &+ \gamma_{31}TeacherBlack_kBlack_i + \gamma_{32}PropBlack_kBlack_i \\ &+ \eta_{0k} + \eta_{3k}Black_i + \varepsilon_i \end{aligned}$$

Building an R formula

Global intercept
(included automatically)

```
student_reading_score ~  
  student_age_s + student_female  
  student_re_black*teacher_re_black +  
  student_re_black*class_prop_black +  
  (1 + student_re_black | teacher_id)
```

$$\begin{aligned} S_{ik} = & \gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k \\ & + \beta_1Age_i + \beta_2Female_i + \gamma_{30}Black_i \\ & + \gamma_{31}TeacherBlack_kBlack_i + \gamma_{32}PropBlack_kBlack_i \\ & + \eta_{0k} + \eta_{3k}Black_i + \varepsilon_i \end{aligned}$$

Building an R formula

Interactions (*) automatically include standalone terms

```
student_reading_score ~
  student_age_s + student_female
  student_re_black*teacher_re_black +
  student_re_black*class_prop_black +
  (1 + student_re_black | teacher_id)
```

$$\begin{aligned} S_{ik} = & \gamma_{00} + \gamma_{01} \text{TeacherBlack}_k + \gamma_{02} \text{PropBlack}_k \\ & + \beta_1 \text{Age}_i + \beta_2 \text{Female}_i + \gamma_{30} \text{Black}_i \\ & + \gamma_{31} \text{TeacherBlack}_k \text{Black}_i + \gamma_{32} \text{PropBlack}_k \text{Black}_i \\ & + \eta_{0k} + \eta_{3k} \text{Black}_i + \varepsilon_i \end{aligned}$$

Building an R formula

Interactions (*) automatically include standalone terms

```
student_reading_score ~  
  student_age_s + student_female  
  student_re_black*teacher_re_black +  
  student_re_black*class_prop_black +  
  (1 + student_re_black | teacher_id)
```

$$\begin{aligned} S_{ik} = & \gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k \\ & + \beta_1Age_i + \beta_2Female_i + \gamma_{30}Black_i \\ & + \gamma_{31}TeacherBlack_kBlack_i + \gamma_{32}PropBlack_kBlack_i \\ & + \eta_{0k} + \eta_{3k}Black_i + \varepsilon_i \end{aligned}$$

Building an R formula

Redundant terms (`student_re_black`)
are not added twice

```
student_reading_score ~  
  student_age_s + student_female  
  student_re_black*teacher_re_black +  
  student_re_black*class_prop_black +  
  (1 + student_re_black | teacher_id)
```

$$\begin{aligned} S_{ik} = & \gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k \\ & + \beta_1Age_i + \beta_2Female_i + \gamma_{30}Black_i \\ & + \gamma_{31}TeacherBlack_kBlack_i + \gamma_{32}PropBlack_kBlack_i \\ & + \eta_{0k} + \eta_{3k}Black_i + \varepsilon_i \end{aligned}$$

Building an R formula

Random effects use pipe notation (|)

```
student_reading_score ~  
  student_age_s + student_female  
  student_re_black*teacher_re_black +  
  student_re_black*class_prop_black +  
  (1 + student_re_black | teacher_id)
```

$$\begin{aligned} S_{ik} = & \gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k \\ & + \beta_1Age_i + \beta_2Female_i + \gamma_{30}Black_i \\ & + \gamma_{31}TeacherBlack_kBlack_i + \gamma_{32}PropBlack_kBlack_i \\ & + \eta_{0k} + \eta_{3k}Black_i + \varepsilon_i \end{aligned}$$

Building an R formula

Grouping elements after the pipe

```
student_reading_score ~  
  student_age_s + student_female  
  student_re_black*teacher_re_black +  
  student_re_black*class_prop_black +  
  (1 + student_re_black | teacher_id)
```

$$\begin{aligned} S_{ik} = & \gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k \\ & + \beta_1Age_i + \beta_2Female_i + \gamma_{30}Black_i \\ & + \gamma_{31}TeacherBlack_kBlack_i + \gamma_{32}PropBlack_kBlack_i \\ & + \eta_{0k} + \eta_{3k}Black_i + \varepsilon_i \end{aligned}$$

Building an R formula

Random intercepts indicated
with constant (1)

```
student_reading_score ~  
  student_age_s + student_female  
  student_re_black*teacher_re_black +  
  student_re_black*class_prop_black +  
  (1 + student_re_black | teacher_id)
```

$$\begin{aligned} S_{ik} = & \gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k \\ & + \beta_1Age_i + \beta_2Female_i + \gamma_{30}Black_i \\ & + \gamma_{31}TeacherBlack_kBlack_i + \gamma_{32}PropBlack_kBlack_i \\ & + \eta_{0k} + \eta_{3k}Black_i + \varepsilon_i \end{aligned}$$

Building an R formula

Random-slope variables
included in grouping expression

```
student_reading_score ~  
  student_age_s + student_female  
  student_re_black*teacher_re_black +  
  student_re_black*class_prop_black +  
  (1 + student_re_black | teacher_id)
```

$$\begin{aligned} S_{ik} = & \gamma_{00} + \gamma_{01}TeacherBlack_k + \gamma_{02}PropBlack_k \\ & + \beta_1Age_i + \beta_2Female_i + \gamma_{30}Black_i \\ & + \gamma_{31}TeacherBlack_kBlack_i + \gamma_{32}PropBlack_kBlack_i \\ & + \eta_{0k} + \eta_{3k}Black_i + \varepsilon_i \end{aligned}$$

Generalized multilevel models

Generalized multilevel models

Generalized multilevel linear models

Simply add a link function and change the outcome distribution.

E.g. modeling whether a student did better on the math test than the reading test (M_{ik}).

$$M_{ik} \sim \text{Binomial}(1, p_{ik})$$

$$\text{logit}(p_{ik}) = \beta_{0k} + \beta_{1k} \text{Age}_i + \beta_{2k} \text{Female}_i$$

$$\beta_{0k} = \gamma_{00} + \gamma_{01} \text{TExp}_k + \gamma_{02} \text{TFemale}_k + \eta_{0k}$$

$$\beta_{1k} = \gamma_{10} + \eta_{1k}$$

$$\beta_{2k} = \gamma_{20} + \gamma_{21} \text{TFemale}_k + \eta_{2k}$$

Generalized multilevel models

This is where the
'generalized' part
of the model lives

$$M_{ik} \sim \text{Binomial}(1, p_{ik})$$
$$\text{logit}(p_{ik}) = \beta_{0k} + \beta_{1k} \text{Age}_i + \beta_{2k} \text{Female}_i$$

$$\beta_{0k} = \gamma_{00} + \gamma_{01} \text{TExp}_k + \gamma_{02} \text{TFemale}_k + \eta_{0k}$$

$$\beta_{1k} = \gamma_{10} + \eta_{1k}$$

$$\beta_{2k} = \gamma_{20} + \gamma_{21} \text{TFemale}_k + \eta_{2k}$$

Generalized multilevel models

**This is the same as
a standard
(Gaussian)
multilevel model**

$$M_{ik} \sim \text{Binomial}(1, p_{ik})$$

$$\text{logit}(p_{ik}) = \beta_{0k} + \beta_{1k} \text{Age}_i + \beta_{2k} \text{Female}_i$$

$$\beta_{0k} = \gamma_{00} + \gamma_{01} \text{TExp}_k + \gamma_{02} \text{TFemale}_k + \eta_{0k}$$

$$\beta_{1k} = \gamma_{10} + \eta_{1k}$$

$$\beta_{2k} = \gamma_{20} + \gamma_{21} \text{TFemale}_k + \eta_{2k}$$

Generalized multilevel models

Generalized multilevel linear models

Simply add a link function and change the outcome distribution.

E.g. modeling whether a student did better on the math test than the reading test (M_{ik}).

All coefficients and parameters are affected by link function

$$M_{ik} \sim \text{Binomial}(1, p_{ik})$$

$$\text{logit}(p_{ik}) = \beta_{0k} + \beta_{1k} \text{Age}_i + \beta_{2k} \text{Female}_i$$

$$\beta_{0k} = \gamma_{00} + \gamma_{01} \text{TEExp}_k + \gamma_{02} \text{TFemale}_k + \eta_{0k}$$

$$\beta_{1k} = \gamma_{10} + \eta_{1k}$$

$$\beta_{2k} = \gamma_{20} + \gamma_{21} \text{TFemale}_k + \eta_{2k}$$

Interpretation requires careful thinking about the ways that coefficients affect p_{ik} .

Interpreting *direction* of effect (positive vs. negative) is still straightforward. E.g. a strong positive estimate on γ_{02} would suggest that female teachers do a relatively better job of teaching mathematics.