Agenda

1. Building a 2-level model
2. Interpreting multi-level estimates
3. Specifying and and interpreting in R

Building a two-level model

Building a two-level model

Linear model of test scores

Scores depend on students' age, sex, race, and ethnicity.

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i + \beta_{2k} Female_i + \beta_{3k} Black_i$$

Building a two-level model

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i + \beta_{2k} Female_i + \beta_{3k} Black_i$$

Random intercept

Let each class have its own overall performance.

$$\beta_{0k} = \gamma_{00} + \eta_{0k}$$

Building a two-level model

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i + \beta_{2k} Female_i + \beta_{3k} Black_i$$

Random slopes

Let the “effect” of students’ characteristics differ from classroom to classroom

$$\beta_{0k} = \gamma_{00} + \eta_{0k}$$

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

$$\beta_{2k} = \gamma_{20} + \eta_{2k}$$

$$\beta_{3k} = \gamma_{30} + \eta_{3k}$$

Building a two-level model

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i + \beta_{2k} Female_i + \beta_{3k} Black_i$$

Intercept predictors

Does each classes' ***average score*** depend on classroom features? (number of students and teacher's experience)

$$\beta_{0k} = \gamma_{00} + \gamma_{01} Size_k + \gamma_{02} Exp_k + \eta_{0k}$$

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

$$\beta_{2k} = \gamma_{20} + \eta_{2k}$$

$$\beta_{3k} = \gamma_{30} + \eta_{3k}$$

Building a two-level model

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i + \beta_{2k} Female_i + \beta_{3k} Black_i$$

Slope predictors

Does the expected *score difference between older and younger students* depend on the size of the class?

$$\beta_{0k} = \gamma_{00} + \gamma_{01} Size_k + \gamma_{02} Exp_k + \eta_{0k}$$

$$\beta_{1k} = \gamma_{10} + \gamma_{11} Size_k + \eta_{1k}$$

$$\beta_{2k} = \gamma_{20} + \eta_{2k}$$

$$\beta_{3k} = \gamma_{30} + \eta_{3k}$$

Building a two-level model

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i + \beta_{2k} Female_i + \beta_{3k} Black_i$$

Group-level covariates on slope and intercept

In most cases, covariates on slope models should be included in intercept model as well

$$\begin{aligned}
\beta_{0k} = {} & \gamma_{00} + \gamma_{01} Size_k + \gamma_{02} Exp_k + \\
& \gamma_{03} TeacherFemale_k + \\
& \gamma_{04} TeacherBlack_k + \\
& \gamma_{05} PropBlack_k + \eta_{0k} \\
\beta_{1k} = {} & \gamma_{10} + \gamma_{11} Size_k + \eta_{1k} \\
\beta_{2k} = {} & \gamma_{20} + \gamma_{21} TeacherFemale_k + \eta_{2k} \\
\beta_{3k} = {} & \gamma_{30} + \gamma_{31} TeacherBlack_k + \\
& \gamma_{32} PropBlack_k + \eta_{3k}
\end{aligned}$$

Covariance model

Covariance structure

Covariance of the level-one coefficients modeled using variance Φ and correlations P.

$$\begin{bmatrix}\eta_{0k}\\ \eta_{1k}\\ \eta_{2k}\\ \eta_{3k}\end{bmatrix} \sim \text{MVNorm}\left(\begin{bmatrix}0\\0\\0\\0\end{bmatrix}, \Phi P \Phi\right)$$

$$\Phi = \begin{bmatrix}\phi_0 & 0 & 0 & 0\\ 0 & \phi_1 & 0 & 0\\ 0 & 0 & \phi_2 & 0\\ 0 & 0 & 0 & \phi_3\end{bmatrix}$$

$$P = \begin{bmatrix}1 & \rho_{01} & \rho_{02} & \rho_{03}\\ \rho_{01} & 1 & \rho_{12} & \rho_{13}\\ \rho_{02} & \rho_{12} & 1 & \rho_{23}\\ \rho_{03} & \rho_{13} & \rho_{23} & 1\end{bmatrix}$$

$$\mu_{ik} = \beta_{0k} + \beta_{1k}Age_i + \beta_{2k}Female_i + \beta_{3k}Black_i$$

Building a two-level model

Complete model

The full two-level model describes variation in test scores as a function of student- and class-level covariates.

Residual variation is attributed to within-class (σ) and between-class (ϕ) differences.

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i + \beta_{2k} Female_i + \beta_{3k} Black_i$$

$$\begin{aligned}
\beta_{0k} &= \gamma_{00} + \gamma_{01} Size_k + \gamma_{02} Exp_k + \\
&\quad \gamma_{03} TeacherFemale_k + \\
&\quad \gamma_{04} TeacherBlack_k + \\
&\quad \gamma_{05} PropBlack_k + \eta_{0k} \\
\beta_{1k} &= \gamma_{10} + \gamma_{11} Size_k + \eta_{1k} \\
\beta_{2k} &= \gamma_{20} + \gamma_{21} TeacherFemale_k + \eta_{2k} \\
\beta_{3k} &= \gamma_{30} + \gamma_{31} TeacherBlack_k + \\
&\quad \gamma_{32} PropBlack_k + \eta_{3k}
\end{aligned}$$

Building a two-level model

"Expanded" or "flat" model

To specify the model in R/brms, it is useful to substitute the second-level equations to make the cross-level interactions implicit in the hierarchical model clear

$$S_{ik} =$$

$$\gamma_{00} + \gamma_{01} Size_k + \gamma_{02} Exp_k + TeacherFemale_k +$$

$$\gamma_{04} TeacherBlack_k + \gamma_{05} PropBlack_k + \gamma_{10} Age_i$$

$$\gamma_{11} Age_i Size_k + \gamma_{20} Female_i + \gamma_{21} Female_i TeacherFemale_k +$$

$$\gamma_{30} Black_i + \gamma_{31} Black_i TeacherBlack_k + \gamma_{32} Black_i PropBlack_k +$$

$$\eta_{0k} + \eta_{1k} Age_i + \eta_{2k} Female_i + \eta_{3k} Black_i + \varepsilon_i$$

Interpreting estimates

Two-level model: estimates

$$S_{ik} \sim \text{Norm}(\mu_{ik}, \sigma)$$
$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i + \beta_{2k} Female_i + \beta_{3k} Black_i$$

$$\beta_{0k} = \boxed{\gamma_{00}} + \boxed{\gamma_{01}} Size_k + \boxed{\gamma_{02}} Exp_k +$$
$$\gamma_{03} TeacherFemale_k +$$
$$\gamma_{04} TeacherBlack_k +$$
$$\gamma_{05} PropBlack_k + \eta_{0k}$$
$$\beta_{1k} = \boxed{\gamma_{10}} + \boxed{\gamma_{11}} Size_k + \eta_{1k}$$
$$\beta_{2k} = \gamma_{20} + \gamma_{21} TeacherFemale_k + \eta_{2k}$$
$$\beta_{3k} = \gamma_{30} + \gamma_{31} TeacherBlack_k +$$
$$\gamma_{32} PropBlack_k + \eta_{3k}$$

		Mean	90% credible interval	
Expected score when all covariates are zero	γ_{00}	544.18	521.97	566.37
Baseline effect of class size	γ_{01}	–1.47	–2.02	–0.93
Baseline effect of teacher experience	γ_{02}	0.23	–0.01	0.47
Baseline effect of age	γ_{10}	–4.76	–5.87	–3.67
Effect of class size on age difference in scores	γ_{11}	0.10	–0.18	0.39

Two-level model: estimates

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i + \beta_{2k} Female_i + \beta_{3k} Black_i$$

$$\begin{aligned}
\beta_{0k} = {} & \gamma_{00} + \gamma_{01} Size_k + \gamma_{02} Exp_k + \\
& \boxed{\gamma_{03}} TeacherFemale_k + \\
& \gamma_{04} TeacherBlack_k + \\
& \gamma_{05} PropBlack_k + \eta_{0k} \\
\beta_{1k} = {} & \gamma_{10} + \gamma_{11} Size_k + \eta_{1k} \\
\beta_{2k} = {} & \boxed{\gamma_{20}} + \boxed{\gamma_{21}} TeacherFemale_k + \eta_{2k} \\
\beta_{3k} = {} & \gamma_{30} + \gamma_{31} TeacherBlack_k + \\
& \gamma_{32} PropBlack_k + \eta_{3k}
\end{aligned}$$

		Mean	***90% credible interval***	
Baseline gender difference in teacher effectiveness	γ_{03}	–12.73	–34.86	9.44
Expected score difference between female versus male students with male teacher	γ_{20}	10.61	–7.43	28.56
Expected score difference for female students with female versus male teachers	γ_{21}	–0.61	–18.72	17.43

Two-level model: estimates

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i + \beta_{2k} Female_i + \beta_{3k} Black_i$$

$$\beta_{0k} = \gamma_{00} + \gamma_{01} Size_k + \gamma_{02} Exp_k + \gamma_{03} TeacherFemale_k + \boxed{\gamma_{04}} TeacherBlack_k + \boxed{\gamma_{05}} PropBlack_k + \eta_{0k}$$

$$\beta_{1k} = \gamma_{10} + \gamma_{11} Size_k + \eta_{1k}$$

$$\beta_{2k} = \gamma_{20} + \gamma_{21} TeacherFemale_k + \eta_{2k}$$

$$\beta_{3k} = \boxed{\gamma_{30}} + \boxed{\gamma_{31}} TeacherBlack_k + \boxed{\gamma_{32}} PropBlack_k + \eta_{3k}$$

		Mean	90% credible interval	
Race difference in teacher effectiveness (white students)	γ_{04}	–7.96	–17.80	1.74
Effect of class racial composition (white students)	γ_{05}	–25.99	–38.44	–13.41
Baseline race difference in student scores	γ_{30}	–23.73	–29.22	–18.09
Added effect of having a Black teacher for Black (versus white) students	γ_{31}	13.97	4.37	23.78
Added effect of class racial composition for Black (versus white) students	γ_{32}	3.97	–10.53	18.72

Two-level model: estimates