SOCI 620

March 21

1. Random slopes
2. Multivariate normal distribution
3. Jointly distributed random effects
4. Specifying formulas in R

Random slopes models

Random intercepts (refresher)

Random intercept model of test score

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

$$\mu_{ik} = \beta_{0k} + \beta_1 Age_i$$

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

η_{0k} allows each classroom to have a different average score (intercept)

$$\eta_{0k} \sim \text{Norm}(0, \phi_0)$$

(Fixed priors omitted)

Random slopes

Independent random coefficients

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i$$

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

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

η_{0k} and η_{1k} allow each classroom to have a different intercept *and* a different slope

$$\eta_{0k} \sim \text{Norm}(0, \phi_0)$$

$$\eta_{1k} \sim \text{Norm}(0, \phi_1)$$

(Fixed priors omitted)

Random slopes

Independent random coefficients

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i$$

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

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

$$\eta_{0k} \sim \text{Norm}(0, \phi_0)$$

$$\eta_{1k} \sim \text{Norm}(0, \phi_1)$$

η_{0k} and η_{1k} are *independent*: knowing one tells us nothing about the other.

(Fixed priors omitted)

Independence of random effects is rarely a realistic assumption.

The multivariate normal distribution

Multivariate normal distribution

Joint distribution over y_0 and y_1

Variables may not be independent

$$\begin{bmatrix} y_0 \\ y_1 \end{bmatrix} \sim \text{MVNorm}\left(\begin{bmatrix} \mu_0 \\ \mu_1 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix} \right)$$

μ_0 and μ_1 are mean of y_0 and y_1, respectively

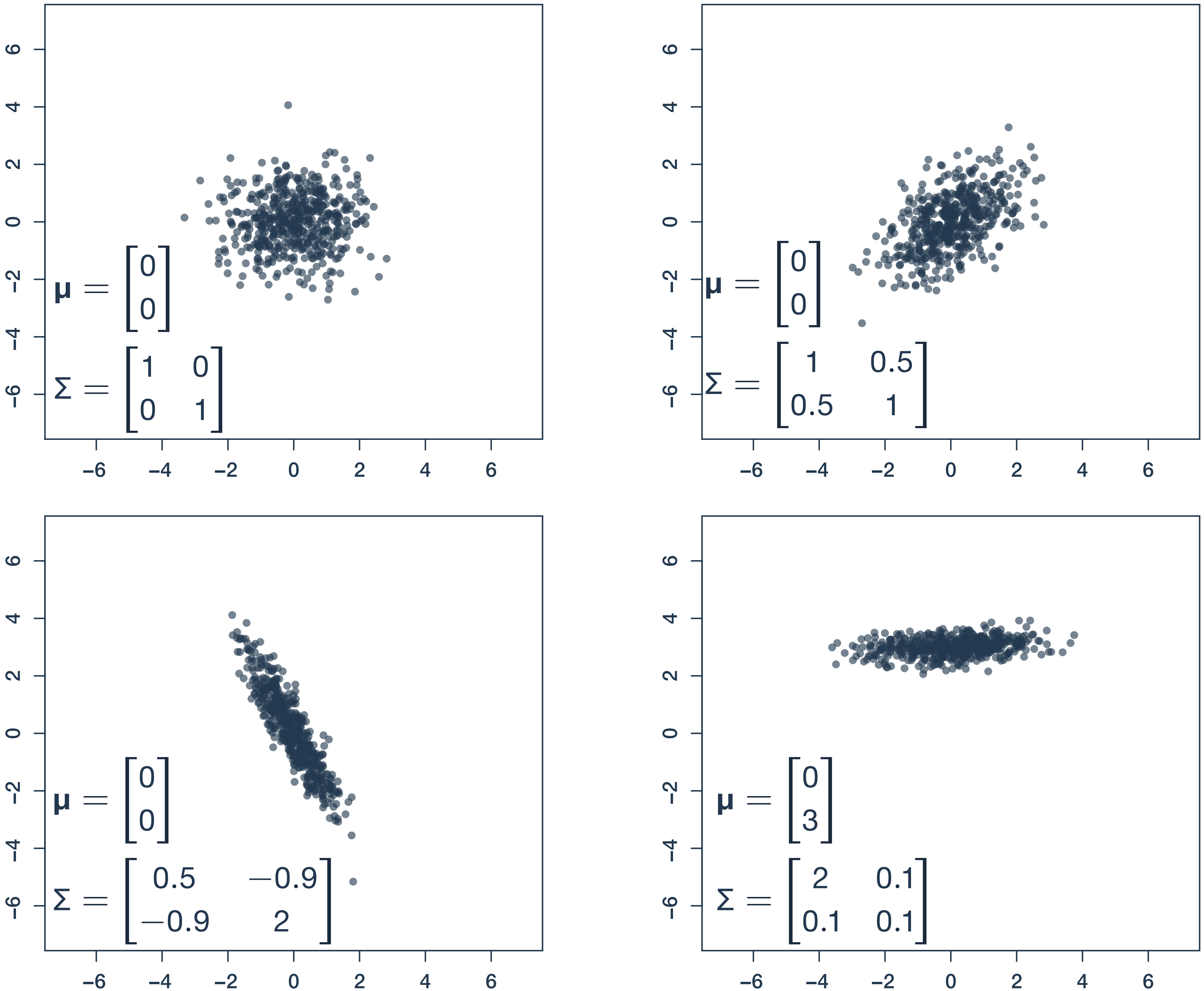
The *covariance matrix* describes the way that y_0 and y_1 inform each other

Covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix} = \begin{bmatrix} \sigma_0^2 & \sigma_0\sigma_1\rho_{01} \\ \sigma_0\sigma_1\rho_{01} & \sigma_1^2 \end{bmatrix}$$

Off-diagonal elements depend on the correlation between y_0 and y_1 (ρ_{01})

Multivariate normal distribution



Multivariate normal distribution

Multivariate normal in *k* dimensions:

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{bmatrix} \sim \text{MVNorm}\left(\begin{bmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_k \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} & \cdots & \sigma_{0k} \\ \sigma_{01} & \sigma_1^2 & \cdots & \sigma_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{0k} & \sigma_{1k} & \cdots & \sigma_k^2 \end{bmatrix} \right)$$

Jointly distributed random effects

Jointly distributed random effects

Independent	**Joint**
$S_{ik} \sim \text{Norm}(\mu_{ik}, \sigma)$	$S_{ik} \sim \text{Norm}(\mu_{ik}, \sigma)$
$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i$	$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i$
$\beta_{0k} = \gamma_{00} + \eta_{0k}$	$\beta_{0k} = \gamma_{00} + \eta_{0k}$
$\beta_{1k} = \gamma_{10} + \eta_{1k}$	$\beta_{1k} = \gamma_{10} + \eta_{1k}$
$\eta_{0k} \sim \text{Norm}(0, \phi_0)$	$\begin{bmatrix} \eta_{0k} \\ \eta_{1k} \end{bmatrix} \sim \text{MVNorm}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Phi\right)$
$\eta_{1k} \sim \text{Norm}(0, \phi_1)$	

(Fixed priors omitted)

Estimates

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

$$\beta_{0k} = \gamma_{00} + \eta_{0k}$$
$$\beta_{1k} = \gamma_{10} + \eta_{1k}$$

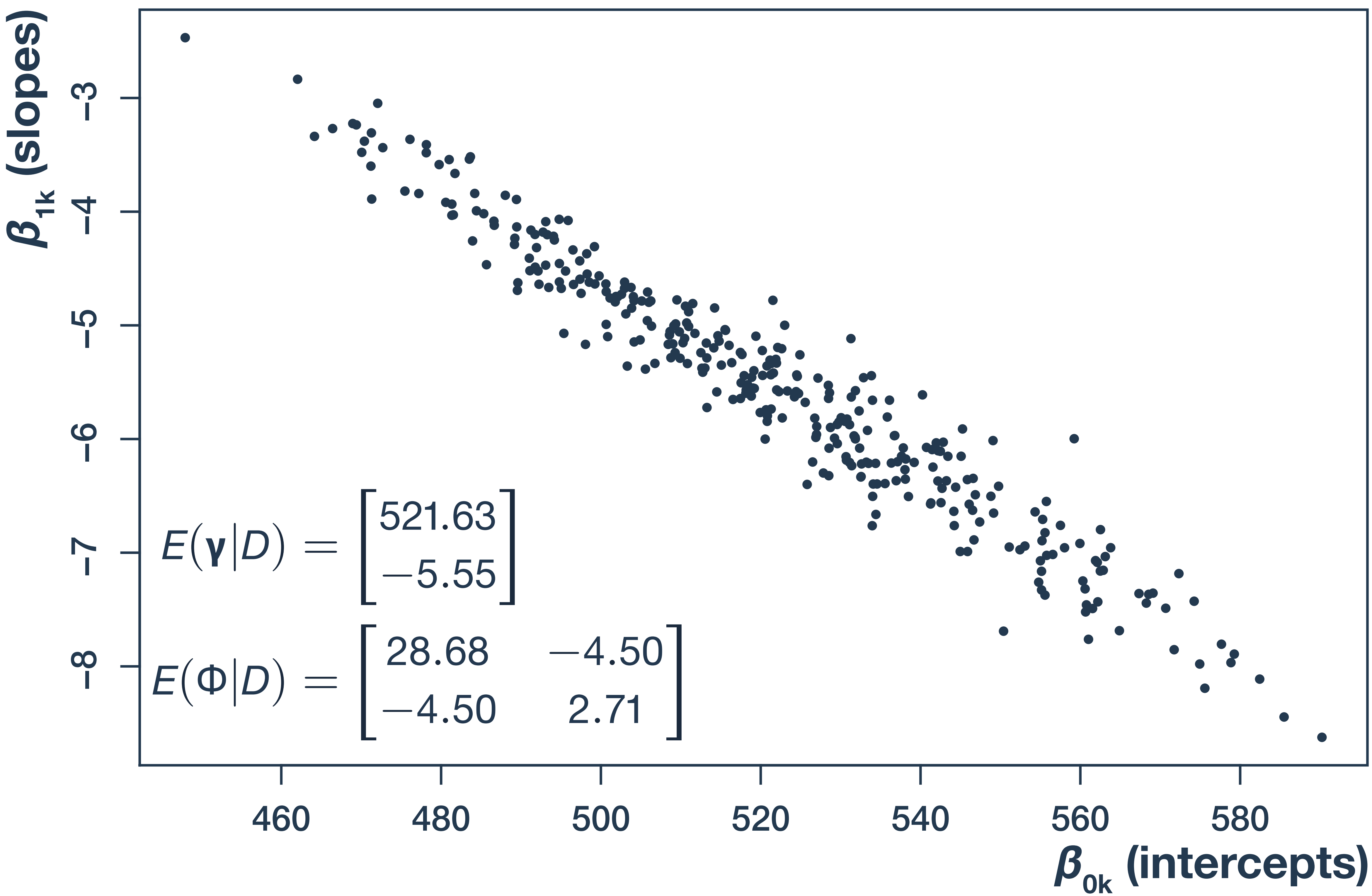
$$\begin{bmatrix} \eta_{0k} \\ \eta_{1k} \end{bmatrix} \sim \text{MVNorm}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Phi \right)$$

$$E(\gamma_{00}|D) = 521.63$$
$$E(\gamma_{10}|D) = -5.55$$
$$E(\sigma|D) = 46.96$$
$$E(\Phi|D) = \begin{bmatrix} 28.68 & -4.50 \\ -4.50 & 2.71 \end{bmatrix}$$

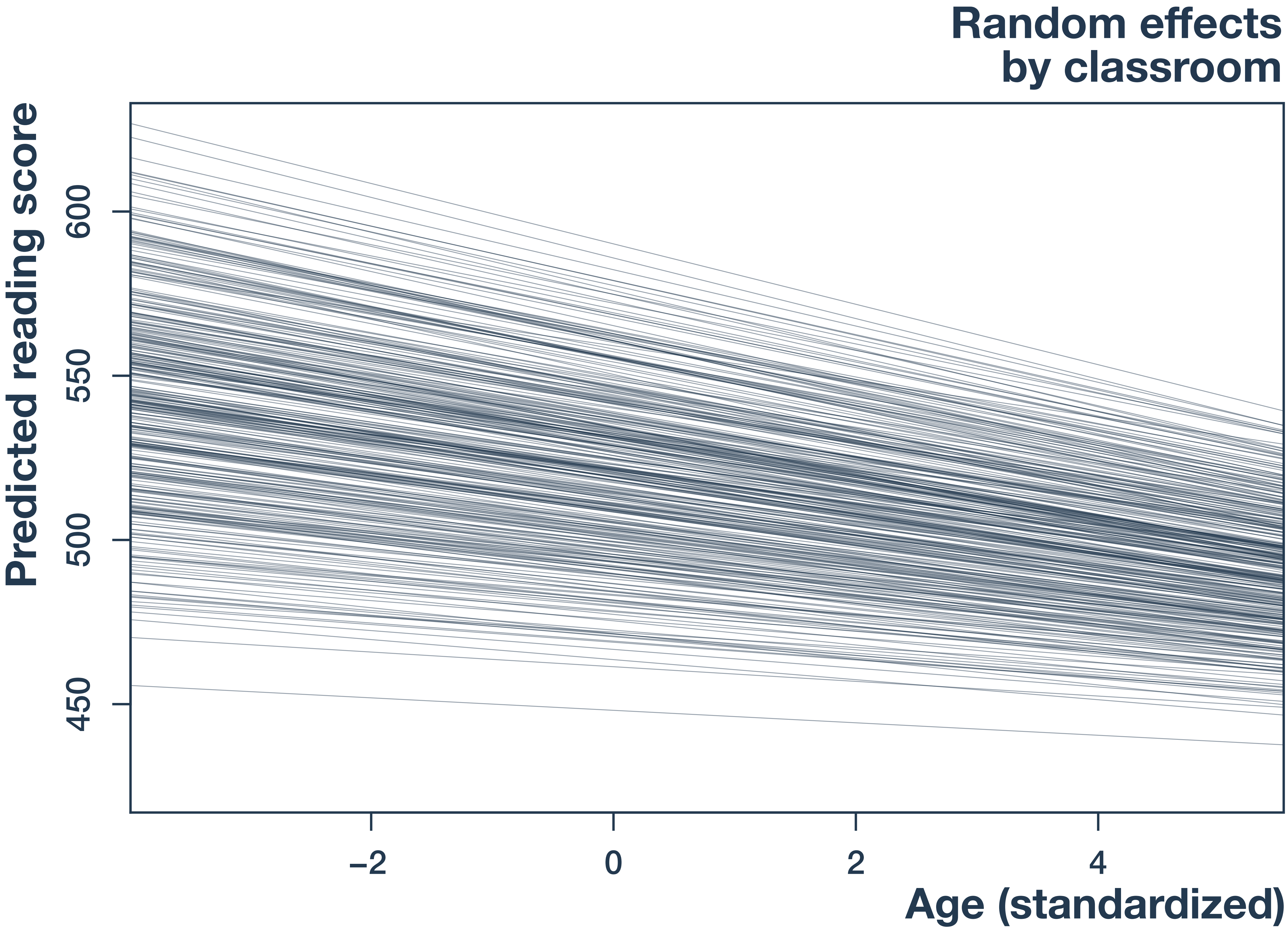
Covariance estimates indicate a relationship between classroom-specific intercepts and classroom-specific age effects

Class-level estimates

Random effects by classroom



Class-level predictions



Specifying formulas in R

Two representations of MLMs

Hierarchical

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} Age_i$$

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

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

Expanded

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

$$\mu_{ik} = \boxed{(\gamma_{00} + \eta_{0k})} + \boxed{(\gamma_{10} + \eta_{1k})} Age_i$$

β_{0k} β_{1k}

Two representations of MLMs

Hierarchical

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

$$\beta_{0k} = \gamma_{00} + \eta_{0k}$$
$$\beta_{1k} = \gamma_{10} + \eta_{1k}$$

Expanded

$$S_{ik} \sim \text{Norm}(\mu_{ik}, \sigma)$$
$$\mu_{ik} = (\gamma_{00} + \eta_{0k}) + (\gamma_{10} + \eta_{1k}) Age_i$$

$$S_{ik} = \gamma_{00} + \eta_{0k} + \gamma_{10} Age_i + \boxed{\eta_{1k} Age_i} + \varepsilon_i$$

This term interacts a covariate (Age_i) with an error term (η_{1k})

Expanded notation

"Fixed" effects

Explained variation in outcome variable.

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

$$S_{ik} = \gamma_{00} + \gamma_{10}Age_i + \eta_{0k} + \eta_{1k}Age_i + \varepsilon_i$$

"Random" effects

Unexplained variation in outcome variable.

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

Building an R formula

The "formula" notation we've been using in brms is used in other R packages like the frequentist "linear mixed effects" package lme4.
(extension of syntax for standard regressions)

```
student_reading_score ~ student_age_s +
        (1 + student_age_s | teacher_id)
```

$$S_{ik} = \gamma_{00} + \gamma_{10} Age_i + \eta_{0k} + \eta_{1k} Age_i + \varepsilon_i$$

Building an R formula

Outcome variable

```
student_reading_score ~ student_age_s +
        (1 + student_age_s| teacher_id)
```

$$S_{ik} = \gamma_{00} + \gamma_{10}Age_i + \eta_{0k} + \eta_{1k}Age_i + \varepsilon_i$$

Building an R formula

Global intercept (included automatically)

```
student_reading_score ~ student_age_s +
        (1 + student_age_s| teacher_id)
```

$$S_{ik} = \boxed{\gamma_{00}} + \gamma_{10} Age_i + \eta_{0k} + \eta_{1k} Age_i + \varepsilon_i$$

Building an R formula

"Fixed effects" covariates are included as usual

```
student_reading_score ~ student_age_s +
          (1 + student_age_s | teacher_id)
```

$$S_{ik} = \gamma_{00} + \gamma_{10} Age_i + \eta_{0k} + \eta_{1k} Age_i + \varepsilon_i$$

Building an R formula

Random effects use pipe notation (|)

```
student_reading_score ~ student_age_s +
        (1 + student_age_s | teacher_id)
```

$$S_{ik} = \gamma_{00} + \gamma_{10} Age_i + \eta_{0k} + \eta_{1k} Age_i + \varepsilon_i$$

Building an R formula

Grouping elements after the ‘pipe’ (“ | ”)

```
student_reading_score ~ student_age_s +
          (1 + student_age_s | teacher_id)
```

$$S_{ik} = \gamma_{00} + \gamma_{10} Age_i + \eta_{0k} + \eta_{1k} Age_i + \varepsilon_i$$

Building an R formula

Random intercepts indicated with constant (1)

```
student_reading_score ~ student_age_s +
         (1 + student_age_s| teacher_id)
```

$$S_{ik} = \gamma_{00} + \gamma_{10}Age_i + \eta_{0k} + \eta_{1k}Age_i + \varepsilon_i$$

Building an R formula

Random-slope variables included in grouping expression

```
student_reading_score ~ student_age_s +
        (1 + student_age_s | teacher_id)
```

$$S_{ik} = \gamma_{00} + \gamma_{10} Age_i + \eta_{0k} + \boxed{\eta_{1k} Age_i} + \varepsilon_i$$