

## Agenda

1. Administrative
2. Priors on covariance
3. Comparing random slopes to pooled (one-level) model
4. Modeling covariance in brms

## Grading

- ∴ Problem set 2 marked
- ∴ Problem set 4 and proposals by the end of the week

## Problem set 5

- ∴ Not due today (since it hasn't be released yet 😞)
- ∴ Will be a simple overview of multilevel models, released later this week

# Priors on covariance

# Priors on covariance

**Predicting reading score with random intercepts and random slopes**

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} \text{Age}_i$$

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

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

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

**$\Phi$  is a matrix of variance and covariance terms**

$$\Phi = \begin{bmatrix} \phi_0^2 & \phi_{01} \\ \phi_{01} & \phi_1^2 \end{bmatrix}$$

# Priors on covariance

## What is a reasonable prior for $\Phi$ ?

**Not every matrix is a covariance matrix.**

$$\Phi = \begin{bmatrix} 4 & 3 \\ 3 & 1 \end{bmatrix} \Rightarrow \text{Cor}(\eta_0, \eta_1) = 1.5$$

**Seemingly uninformative priors can be surprisingly restrictive.**

E.g. using an inverse-Wishart distribution induces a dependency between correlations and standard deviations.

# Priors on covariance

## LKJ correlation prior

Decompose the covariance matrix into matrices of standard deviations and correlations.

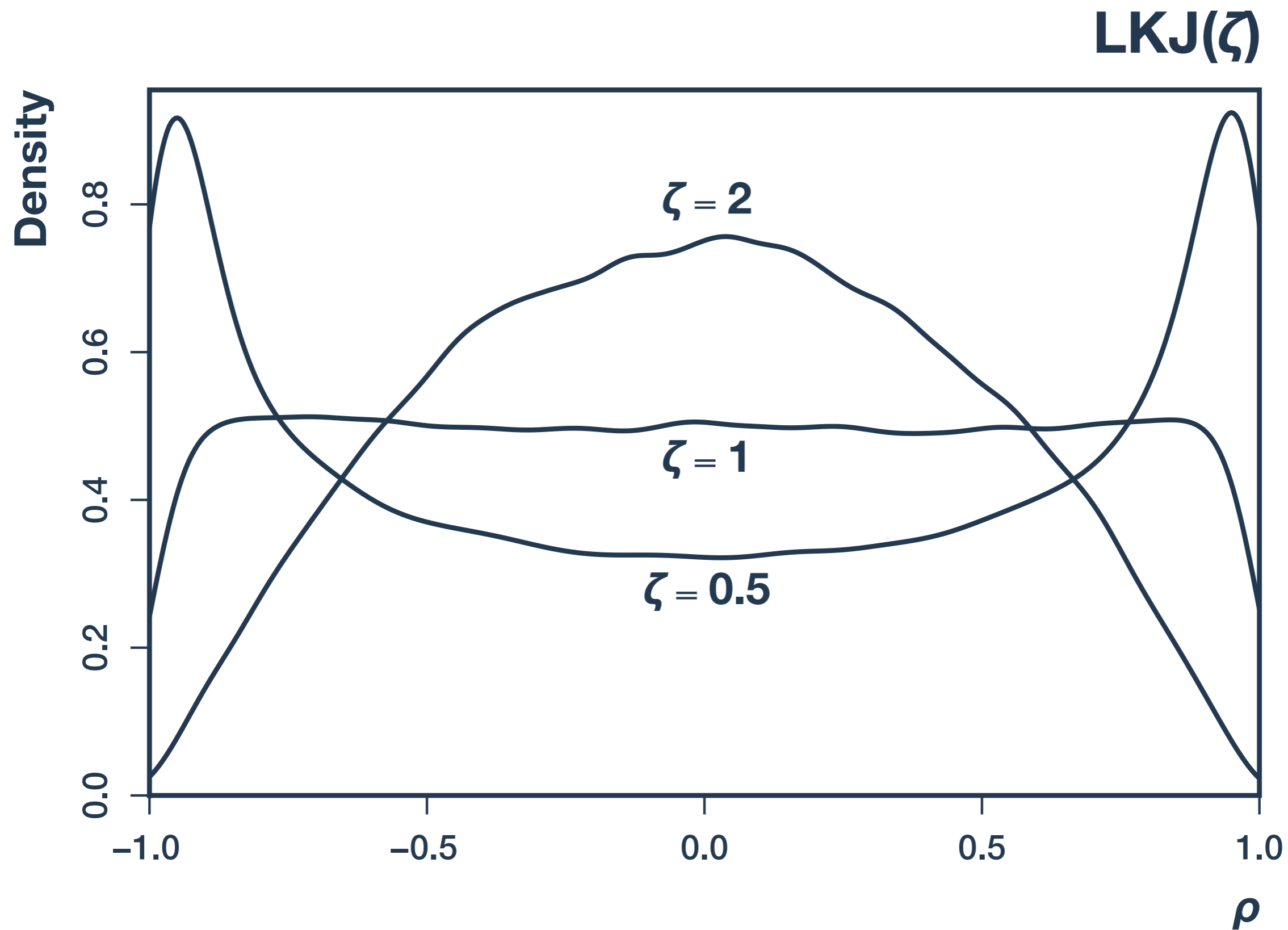
Lewandowski, Daniel, Dorota Kurowicka, and Harry Joe. "Generating Random Correlation Matrices Based on Vines and Extended Onion Method." *Journal of Multivariate Analysis* 100, no. 9 (October 1, 2009)

$$\begin{bmatrix} \phi_0^2 & \phi_{01} \\ \phi_{01} & \phi_1^2 \end{bmatrix} = \begin{bmatrix} \phi_0^2 & \phi_0 \phi_1 \rho_{01} \\ \phi_0 \phi_1 \rho_{01} & \phi_1^2 \end{bmatrix} \\ = \begin{bmatrix} \phi_0 & 0 \\ 0 & \phi_1 \end{bmatrix} \begin{bmatrix} 1 & \rho_{01} \\ \rho_{01} & 1 \end{bmatrix} \begin{bmatrix} \phi_0 & 0 \\ 0 & \phi_1 \end{bmatrix}$$

$\phi_0$  and  $\phi_1$  are the standard deviations of  $\eta_0$  and  $\eta_1$ , respectively (priors for these are straightforward).

The correlation matrix describes correlations for every pair of variables (in this case only  $\rho_{01}$ ).

# Priors on covariance



# Full model

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} \text{Age}_i$$

Student-level linear model

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

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

Class-level linear models

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

Joint distribution of class-level random effects

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

Covariance decomposition  
(standard deviation and correlation)

$$\sigma \sim \text{Unif}(0, 100)$$

$$\gamma_{00} \sim \text{Norm}(500, 100)$$

$$\gamma_{10} \sim \text{Norm}(0, 20)$$

$$\phi_0 \sim \text{HalfCauchy}(0, 40)$$

$$\phi_1 \sim \text{HalfCauchy}(0, 40)$$

$$R \sim \text{LKJ}(2, 2)$$

Priors



# Correlated versus independent

## Correlated random effects

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} \text{Age}_i$$

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

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

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

## Independent random effects

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} \text{Age}_i$$

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

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

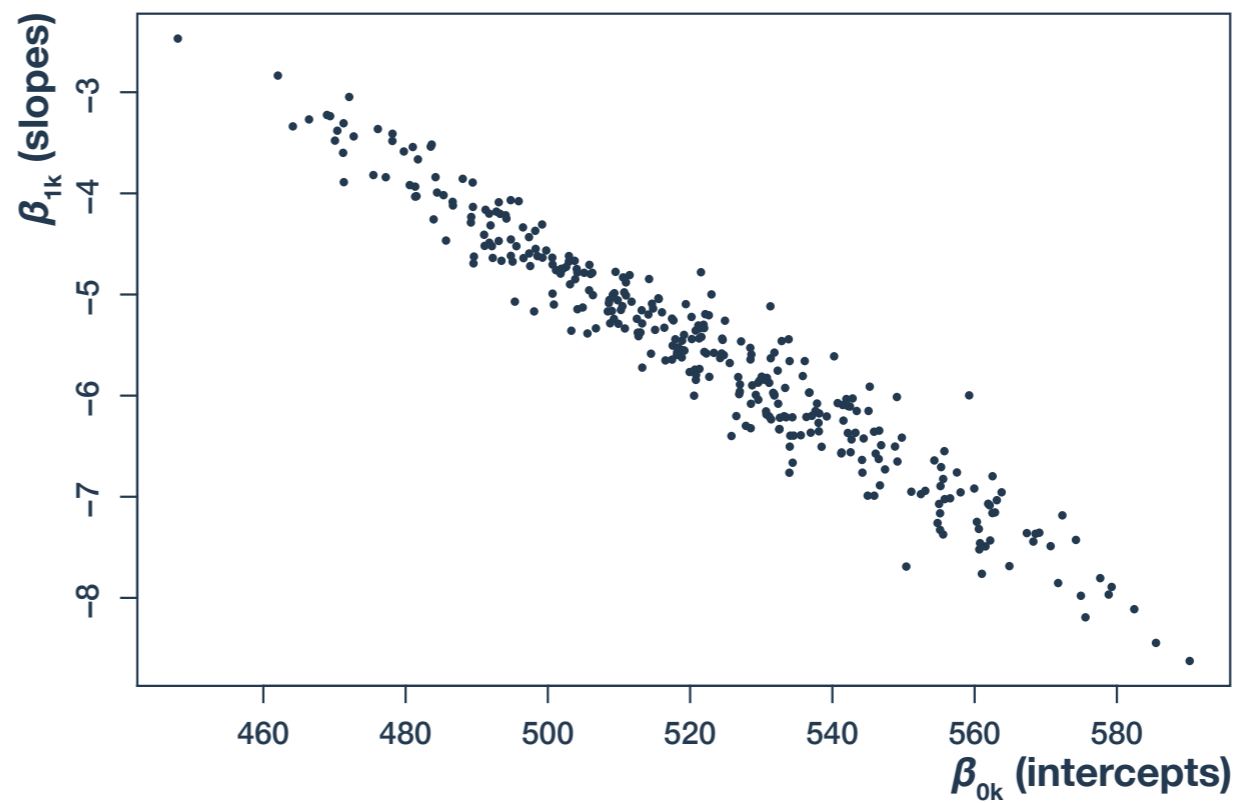
$$\eta_{0k} \sim \text{Norm}(0, \phi_0)$$

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

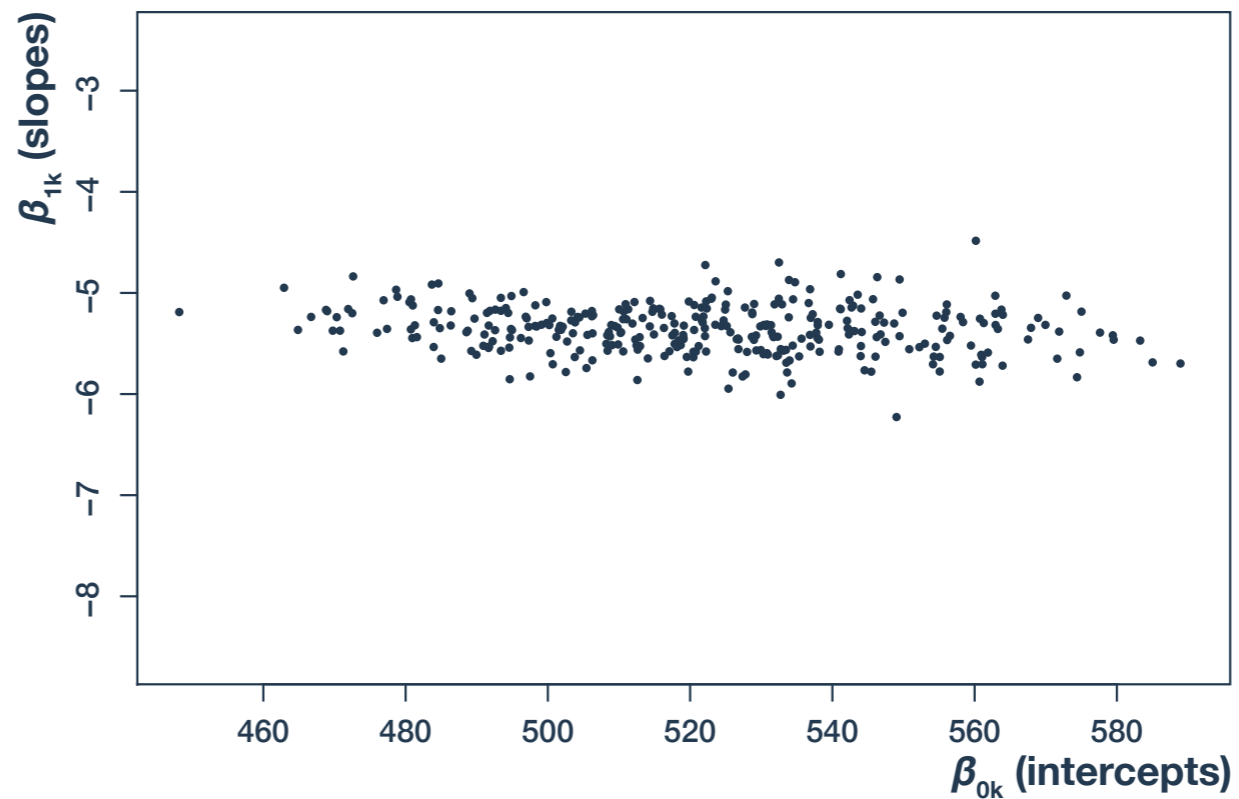
(Fixed priors omitted)

# Correlation of coefficients

**Correlated  
coefficients**



**Independent  
coefficients**



# Comparing to pooled model

# Partial versus total pooling

## Partial pooling (random effects)

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} \text{Age}_i$$

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

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

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

## No pooling (fixed effects)

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

$$\mu_{ik} = \beta_{0k} + \beta_{1k} \text{Age}_i$$

$$\beta_{0k} \sim \text{Norm}(500, 100)$$

$$\beta_{1k} \sim \text{Norm}(0, 20)$$

(Fixed priors omitted)

# Partial versus total pooling

## Unpooled versus partially pooled intercepts and slopes

