Intuitive Joint Priors for Bayesian Multilevel Models

The R2D2M2 prior

University of Stuttgart
Cluster of Excellence in Data-integrated Simulation Science

Project Coordinator: Dr. Paul-Christian Bürkner
Independent Junior Research Group for Bayesian Statistics

Javier Aguilar
PN6

Multilevel Models

Grouped data

$$\mu_n = \sum_{i=1}^{p} x_{ni} b_i + \sum_{i=1}^{p} x_{ni} \left(\sum_{g \in G_i} u_{ig_{j[n]}} \right)$$

$$y_n \sim N(\mu_n, \sigma)$$

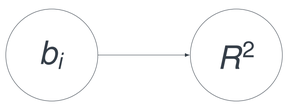
- b_i Overall coefficients
- u_{igj} Varying coefficients

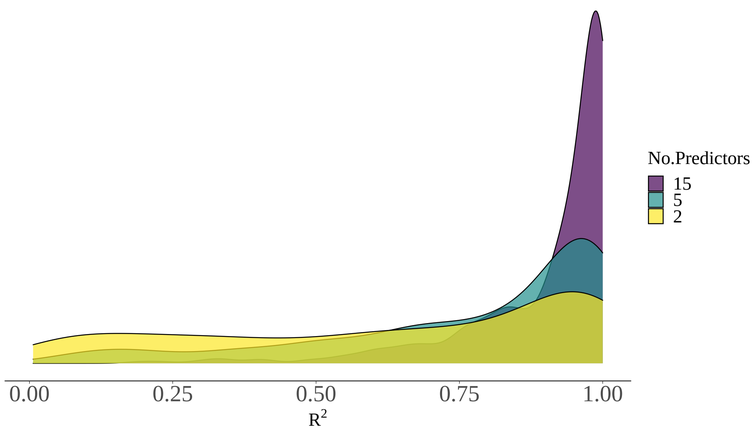
Common Inference Setting

Weakly informative priors

$$b_i \sim N(0, 1), \quad \sigma \sim \mathrm{E}(1)$$

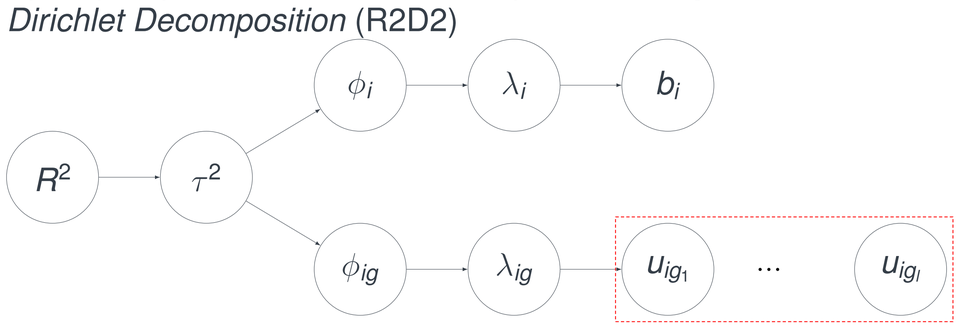
Effect on proportion of explained variance





An intuitive joint prior on R^2

Specify a prior on R^2 and decompose the **explained variance** via a *Dirichlet Decomposition* (R2D2)

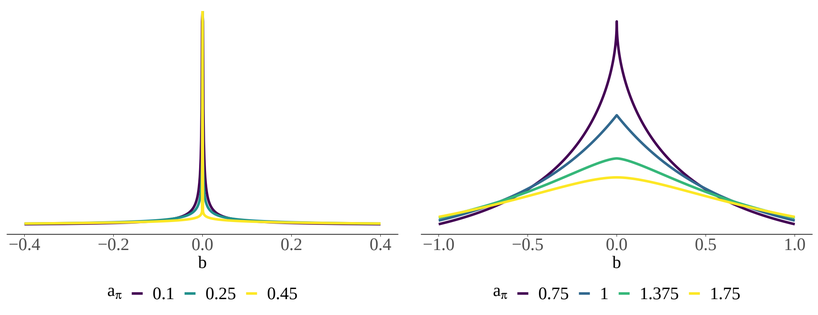


R^2 *D*irichlet *D*ecomposition *M*ultilevel *M*odels: R2-D2-M2 prior

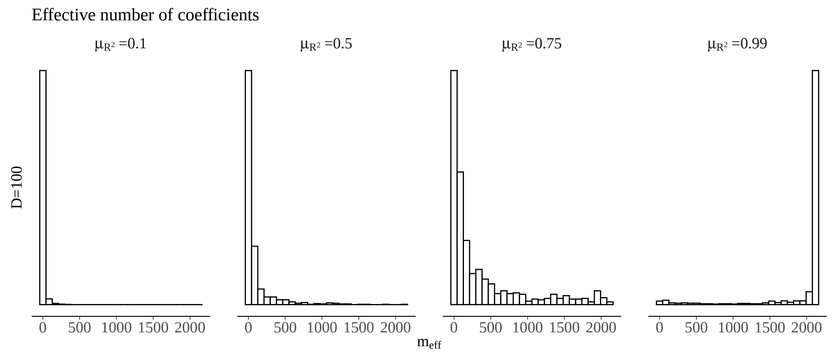
$$R^2 \sim B(\mu, \varphi), \phi \sim \mathrm{Dirichlet}(\alpha), b_i \sim N(0, \lambda_i), u_{ig} \sim N(0, \lambda_{ig}), \sigma \sim p(\sigma)$$

Properties

Marginal Priors

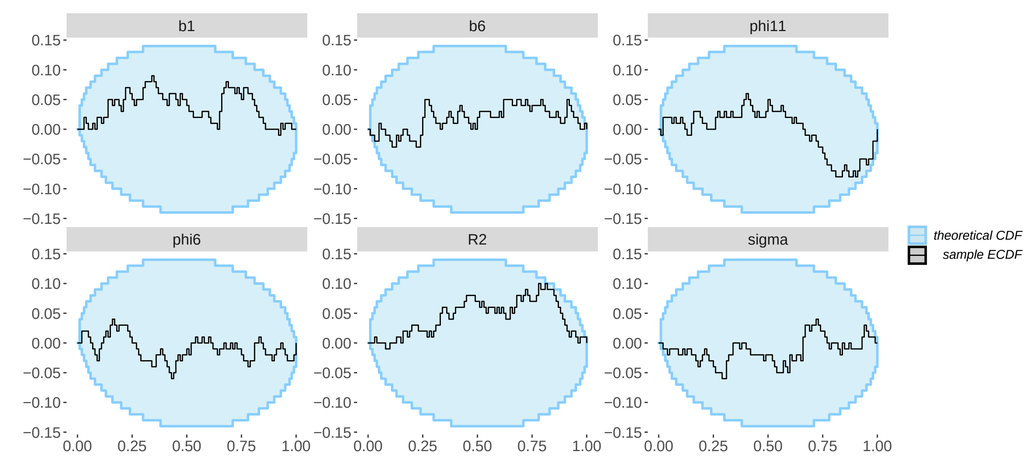


Implied shrinkage



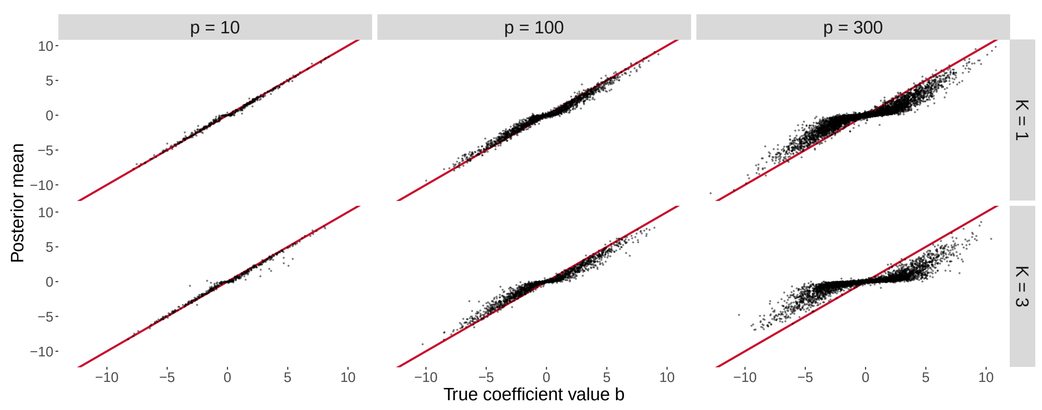
Experiments

Simulation Based Calibration

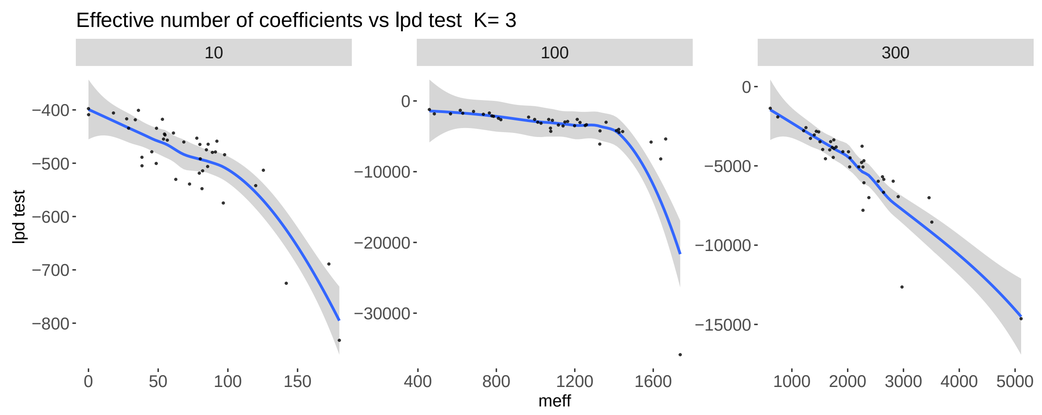


General Multilevel Model Simulation

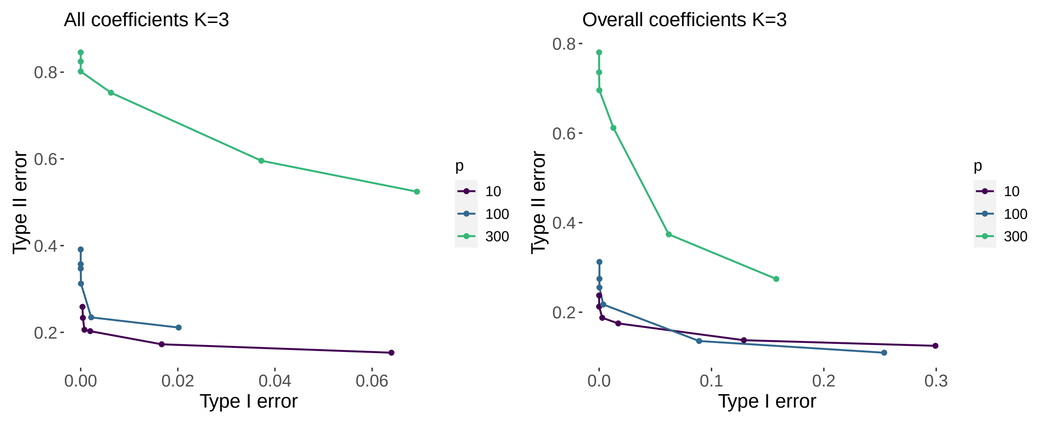
Posterior shrinkage



Out-of-sample Predictive Performance



Posterior coverage and Credibility Intervals



Conclusions

- Our model is well calibrated.
- Prior specification over the whole set of coefficients presents advantages.
- Out-of-sample predictive performance is related to shrinkage.
- Errors are properly controlled.

DFG Deutsche Forschungsgemeinschaft

SimTech