Approximate Bayesian inference for Structural Equation Models (SEM): The INLA approach

Psychoco 2026 @ Università di Padova

Haziq Jamil

Research Specialist

BAYESCOMP @ CEMSE-KAUST

Håvard Rue

Professor of Statistics

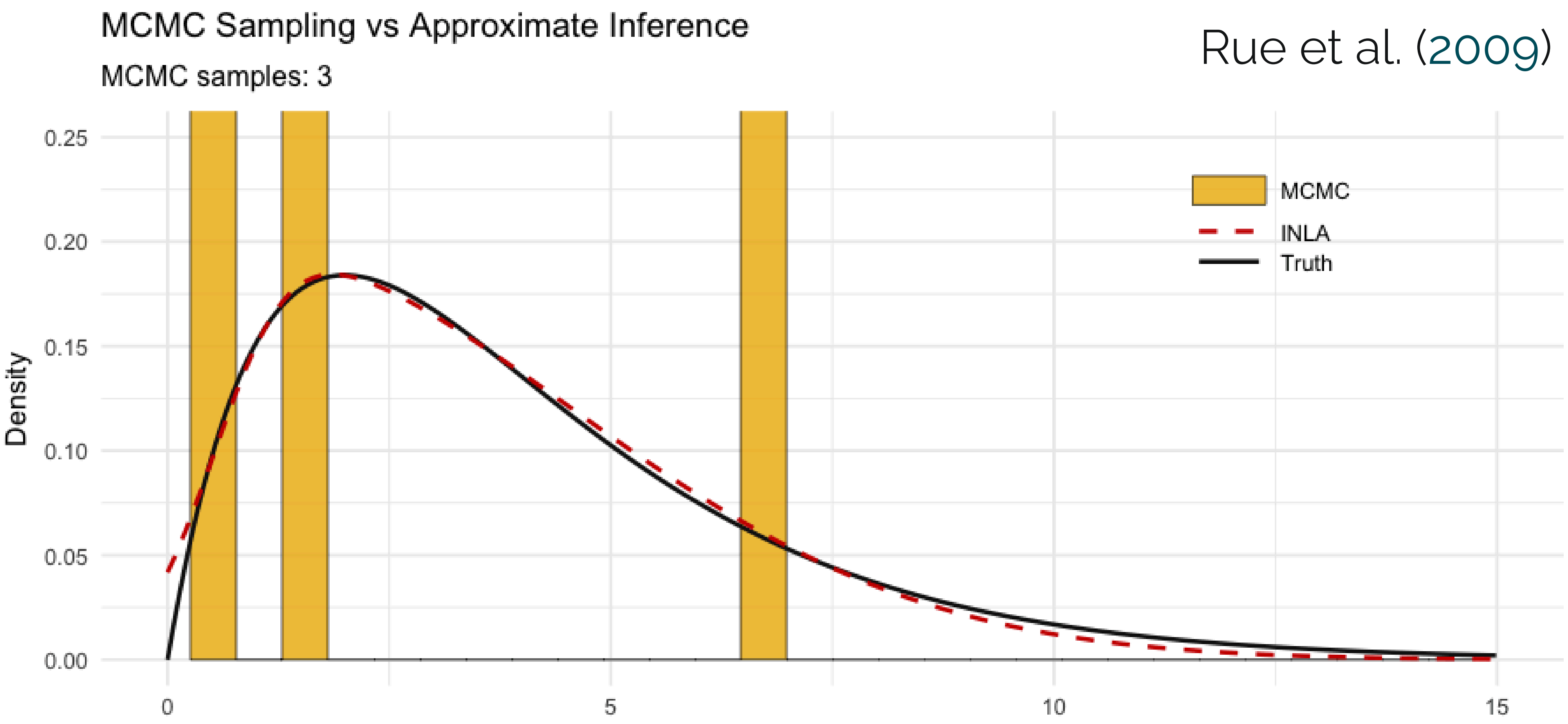
BAYESCOMP @ CEMSE-KAUST

`https://haziqj.ml/psychoco26` | PDF

February 5, 2026

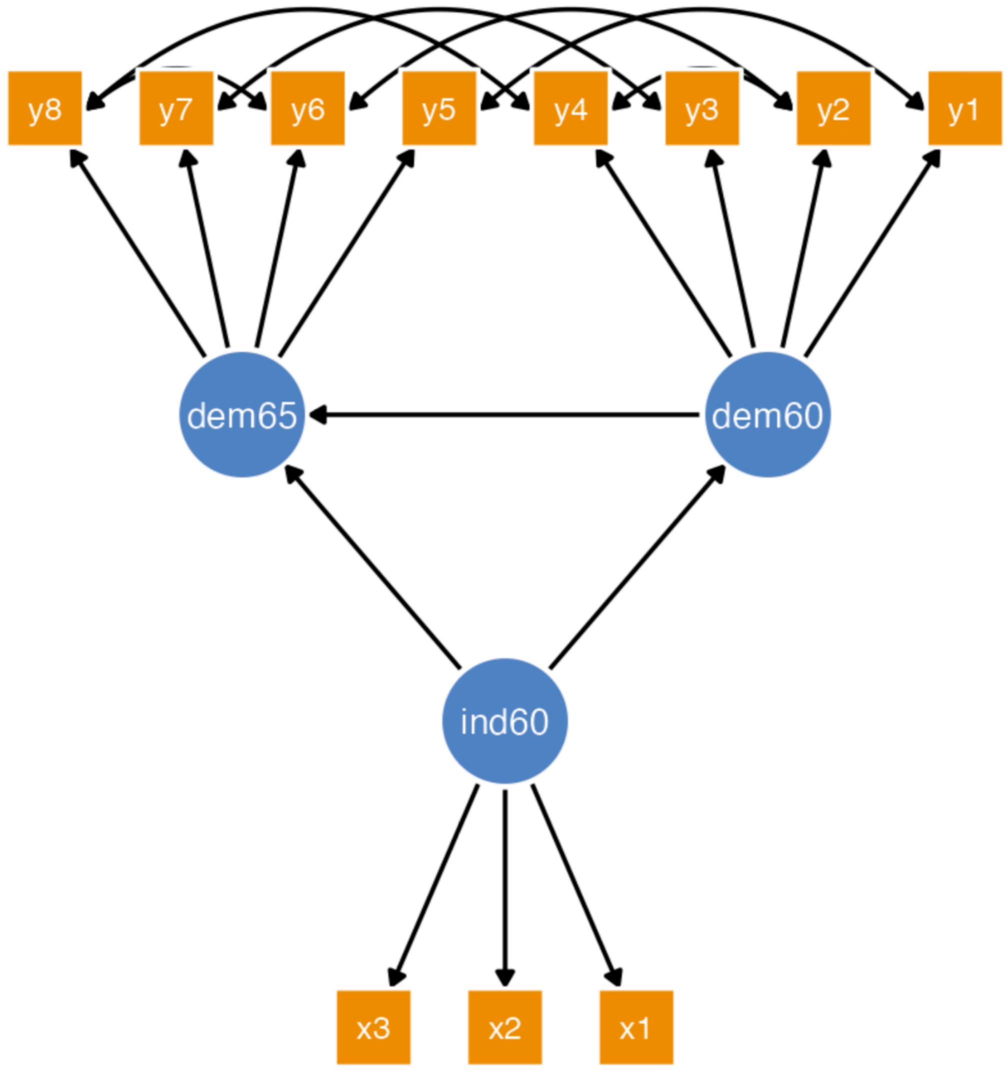
Integrated Nested Laplace Approximations

Rue et al. (2009)



INLA provides a fast, deterministic alternative to MCMC for performing accurate Bayesian inference on a *wide* class of **latent Gaussian models**.

Structural Equation Models using INLA



Bollen's (1989) political democracy 3-factor SEM.

- For $s = 1{:}n$, MVN-SEM equations are

$$y_s = \nu + \Lambda\eta_s + \epsilon_s$$
$$\eta_s = \alpha + B\eta + \zeta_s$$

  with assumptions $\epsilon_s \sim \mathrm{N}(0, \Theta)$, $\eta_s \sim \mathrm{N}(0, \Psi)$, $\mathrm{Cov}(\epsilon_s, \eta_s) = 0$.

- Additionally, Bayesian: $\mathbb{R}^m \ni \vartheta \sim \pi(\vartheta)$.

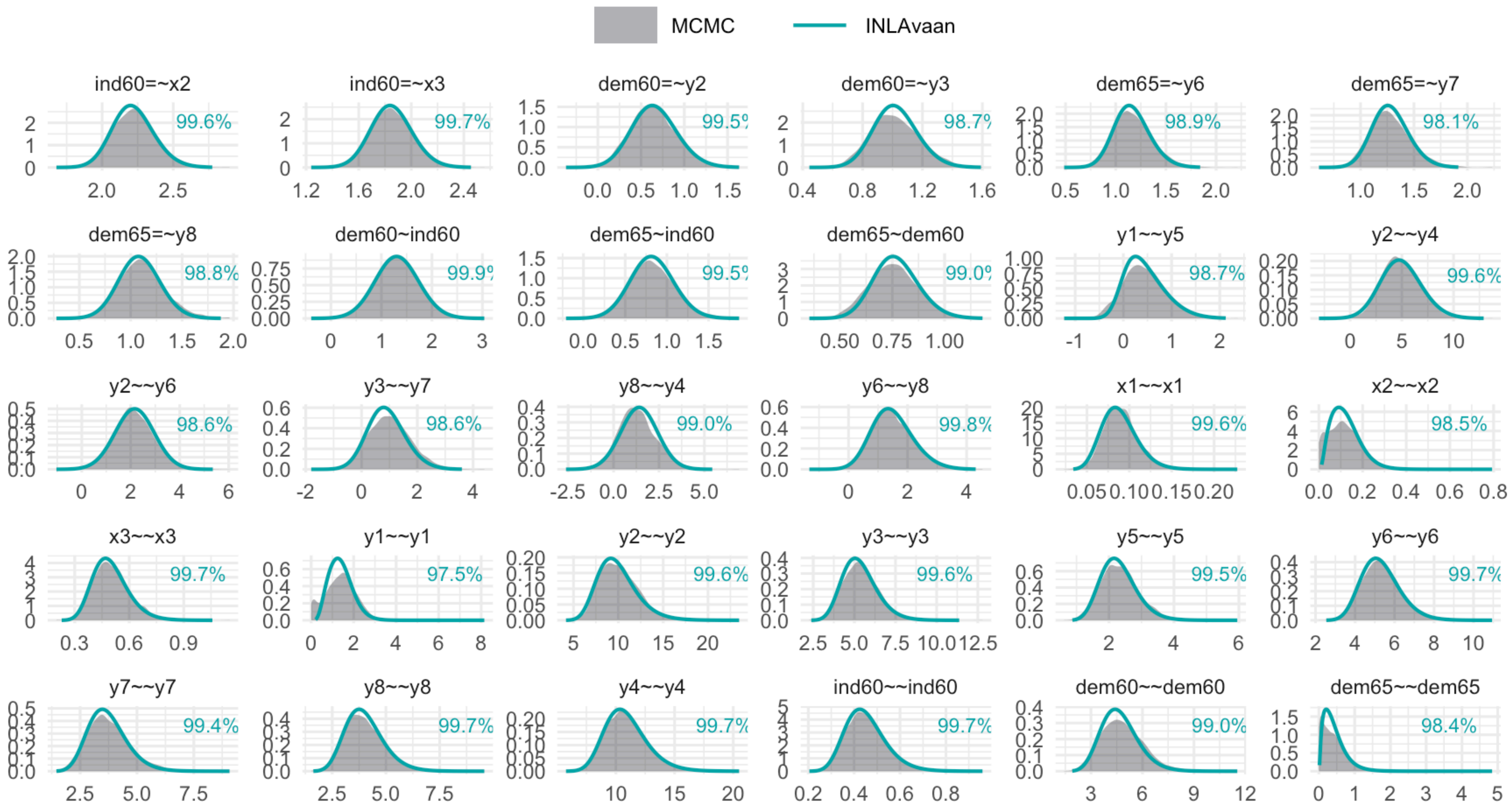
INLAvaan lifecycle experimental CRAN 0.2.3

```r
mod <- "
  ind60 =~ x1 + x2 + x3
  dem60 =~ y1 + y2 + y3
  dem65 =~ y5 + y6 + y7 + y8

  dem60 ~ ind60
  dem65 ~ ind60 + dem60

  y1 ~~ y5
  y2 ~~ y4 + y6
  y3 ~~ y7
  y4 ~~ y8
  y6 ~~ y8
"
fit <- INLAvaan::asem(
  model = mod,
  data = PoliticalDemocracy
)
```

```
ℹ Finding posterior mode.
✔ Finding posterior mode. [41ms]
ℹ Computing the Hessian.
✔ Computing the Hessian. [96ms]
ℹ Performing VB correction.
✔ VB correction; mean |δ| = 0.053σ. [101ms]
⠏ Fitting skew normal to 0/30 marginals.
✔ Fitting skew normal to 30/30 marginals. [673ms]
ℹ Sampling covariances and defined parameters.
✔ Sampling covariances and defined parameters.
[56ms]
⠏ Computing ppp and DIC.
✔ Computing ppp and DIC. [172ms]
```

- Ground up INLA implementation
  - (b)lavaan syntax and methods.
  - `asem()`, `acfa()`, `agrowth()` with `dp = blavaan::dpriors()` option.
  - Very fast—complexity ↑ with m only!

Comparison with MCMC



MCMC ran on 1 chain, 2000 samples after 5000 burn-in, using `{blavaan}` (Stan). % are JS errors.

Global Laplace & VB correction

- Consider the (unnormalised) posterior density function $\pi(\vartheta \mid y) \propto \pi(y \mid \vartheta)\, \pi(\vartheta)$. From a 2nd order Taylor expansion about $\vartheta^* = \mathbf{argmax}_\vartheta \log \pi(\vartheta \mid y)$,

$$\log \pi(\vartheta \mid y) \approx \log \pi(\vartheta^* \mid y) + \frac{1}{2}(\vartheta - \vartheta^*)^\top H(\vartheta - \vartheta^*),$$

clearly able to approximate the posterior by $\mathrm{N}(\vartheta^*, -H^{-1})$, where

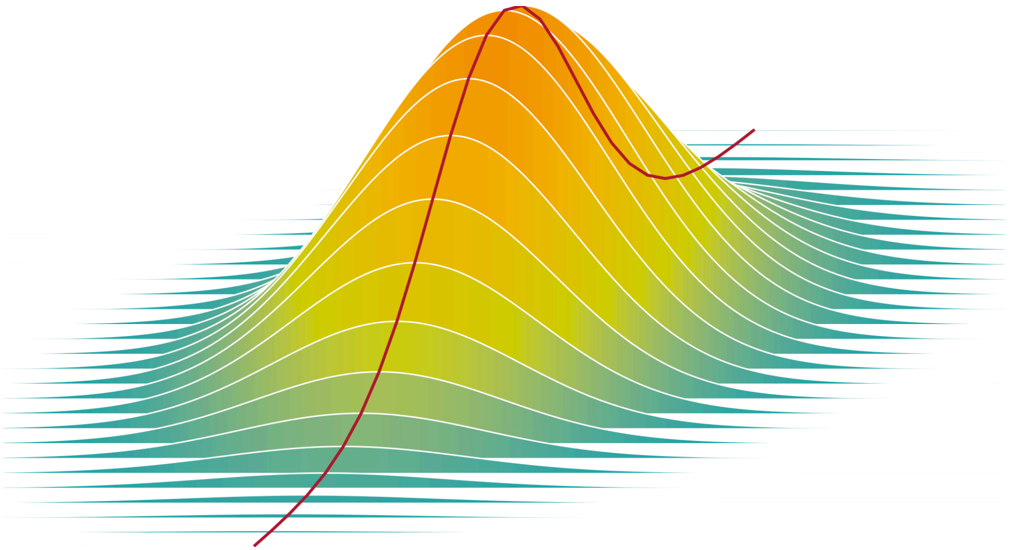
$$H = \nabla^2_\vartheta \log \pi(\vartheta \mid y)\big|_{\vartheta=\vartheta^*}.$$

Approximate posterior marginals

- Use the Laplace approximation again, this time to approximate an integral:

$$\pi(\vartheta_j \mid y) = \int e^{\log \pi(\vartheta_j, \vartheta_{-j} \mid y)} \approx K \times \overbrace{\pi(\vartheta_j, \hat{\vartheta}_{-j} \mid y)}^{\text{height}} \times \overbrace{|H_{-j}|^{-1/2}}^{\text{width}}$$

- This is **expensive** to compute, as for every ϑ_j evaluation, need
    1. An optimisation of ϑ_{-j} at this location; and
    2. The (log) determinant of the Hessian there.



Approximate posterior marginals (cont.)

Lemma 1 (Conditional Mean Path) = **No need for reoptimisation**

The set $\mathcal{C}_j = \{\vartheta \in \mathbb{R}^m \mid \vartheta_{-j} = \mathrm{E}_{\pi_G}[\vartheta_{-j} \mid \vartheta_j]\}$ is the sufficient integration path for the marginal, i.e. $\pi(\vartheta_j) \propto \pi(\vartheta)\big|_{\vartheta \in \mathcal{C}_j}$. Under Gaussianity, the path is linear:

$$\vartheta(t) = \vartheta^* + \Sigma_{\cdot j}\Sigma_{jj}^{-1} t.$$

Lemma 2 (Efficient volume correction) = **No need to factor dense Hessian**

Let L be a whitening matrix for $-H$. For the jth component along the cond. mean path, $\log|-H_{-j}(\vartheta(t))| \approx \mathbf{const.} + \gamma_j t$, where

$$\gamma_j = \sum_{k \neq j} L_{\cdot k}^{\top} \frac{d}{dt}\left(\frac{d}{dL_{\cdot k}}\left[\nabla_{\vartheta} \log \pi(\vartheta(t) \mid y)\right]\right)\bigg|_{t=0}.$$

Gaussian copula sampling

$$\underbrace{z \sim N(0, R)}_{\text{Gaussian correlation}} \xrightarrow{\Phi(\cdot)} \underbrace{u}_{\text{Uniform}} \xrightarrow{F_{\text{SN}}^{-1}(\cdot)} \underbrace{\vartheta^{(b)}}_{\text{SN Marginals}} \xrightarrow{h(\cdot)} \textbf{trans. params.}$$

- Reconstruct joint posterior samples $h(\vartheta^{(b)})$ using a Copula (Nelsen, 2006):

  1. **Generate Latent Dependence**. Draw $z \sim N(0, R)$, where R is the correlation matrix from the inverse Hessian at the mode, $-H^{-1} = SRS$.
  2. **Probability Integral Transform**. Get correlated uniform variates via $u = \Phi(z)$.
  3. **Inverse CDF Mapping**. Apply the inverse CDF of the fitted skew normal marginals.
  4. **Parameter Transformation**. Map to desired transformation using $h(\cdot)$.
  5. **Inference**. Use samples to compute estimates, credible intervals, etc.

- Get estimates for covariances, derived parameters (e.g. direct/indirect effects), fit indices (PPP & DIC), factor scores, etc. using this method.

(b)lavaan S4 methods

```
print(fit)
```

```
INLAvaan 0.2.3.9001 ended normally after 82 iterations

  Estimator                                      BAYES
  Optimization method                           NLMINB
  Number of model parameters                        30

  Number of observations                            75

Model Test (User Model):

   Marginal log-likelihood                   -1688.569
   PPP (Chi-square)                              0.000
```

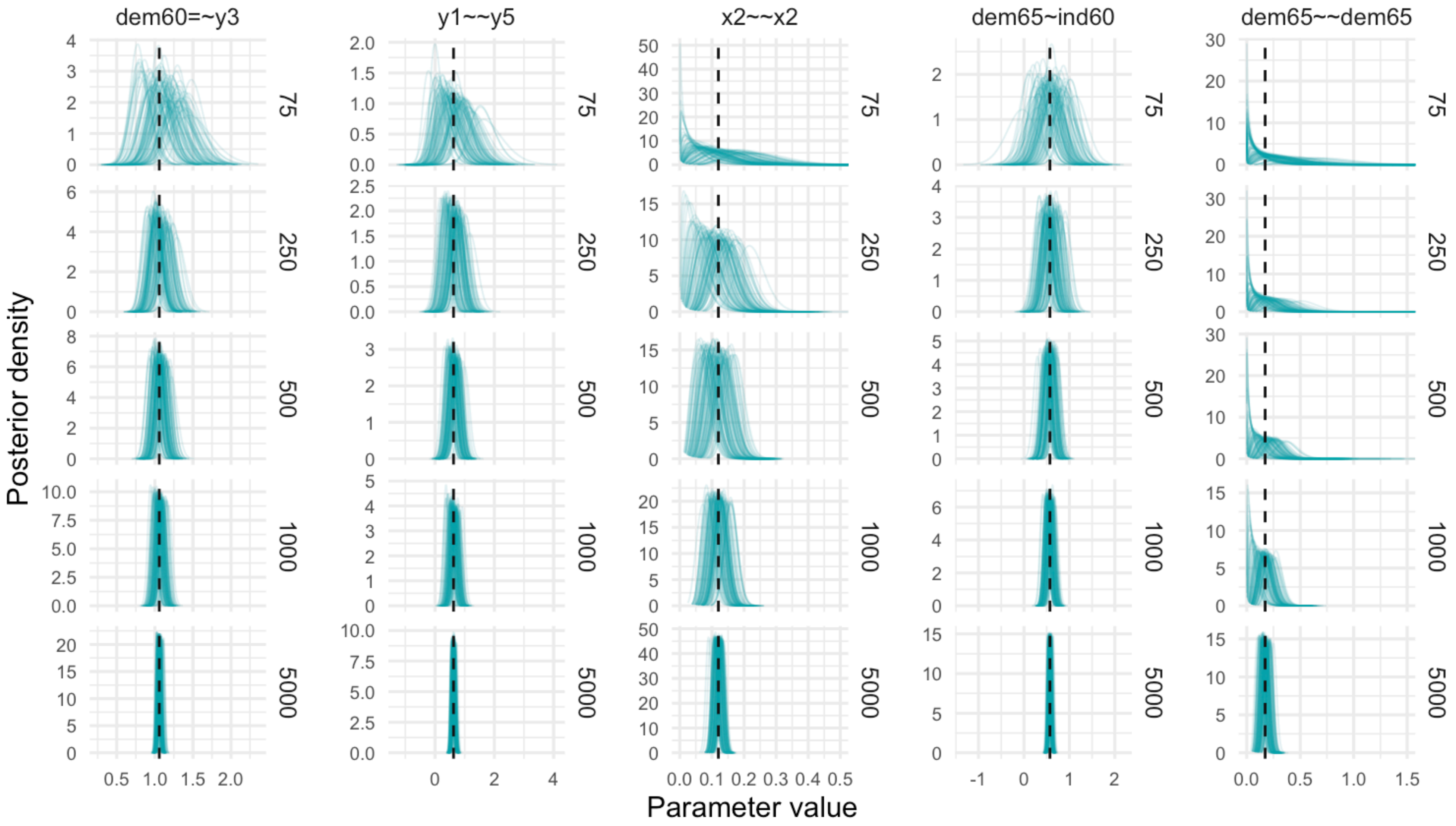
```
coef(fit)
```

```
   ind60=~x2    ind60=~x3    dem60=~y2    dem60=~y3
       2.219        1.851        0.656        1.021
   dem65=~y6    dem65=~y7    dem65=~y8  dem60~ind60
       1.163        1.281        1.086        1.310
 dem65~ind60  dem65~dem60       y1~~y5       y2~~y4
       0.823        0.768        0.224        0.444
      y2~~y6       y3~~y7       y8~~y4       y6~~y8
       0.311        0.207        0.253        0.305
```

Also available:

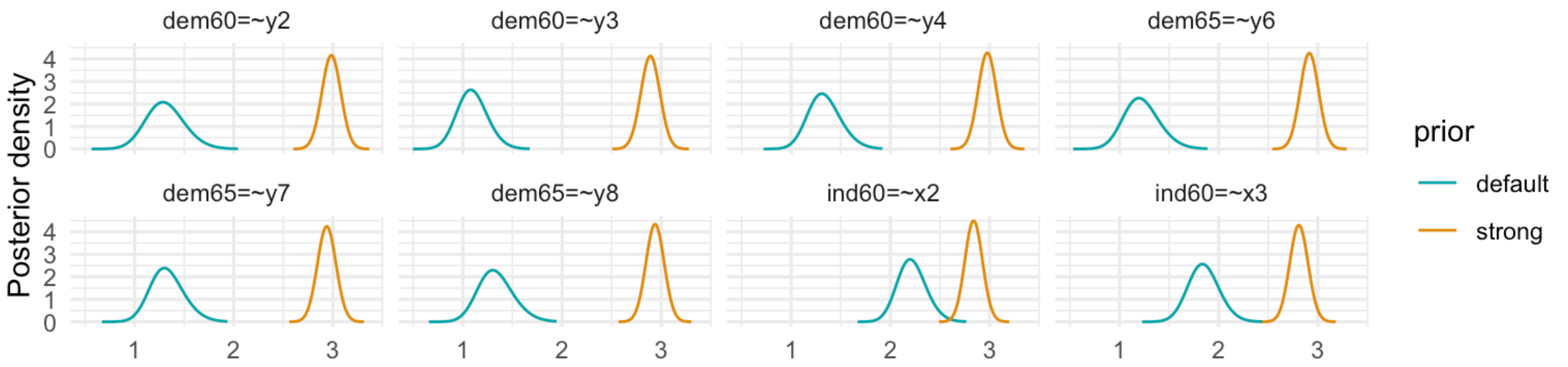
- `summary()`
- `fitmeasures()`
- `predict()`
- `standardisedsolution()`
- `vcov()`
- `plot()`

Parameter recovery from simulated data

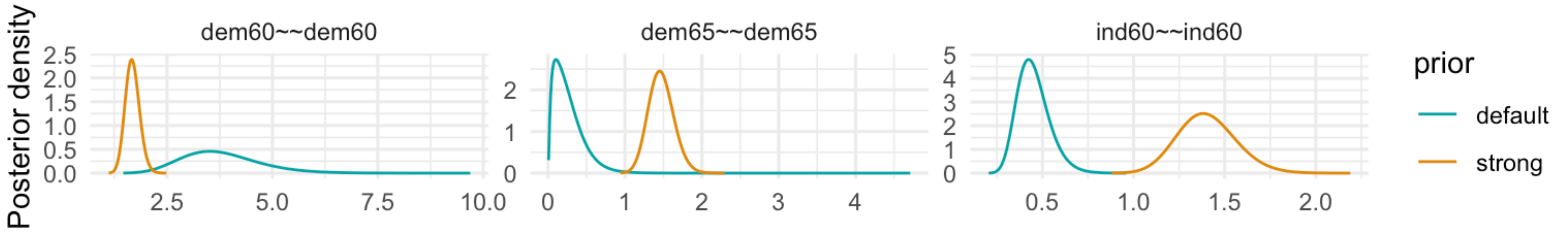


Effects of prior distributions

- Loadings: $\lambda \sim \mathrm{N}(0, 10^2) \rightsquigarrow \lambda \sim \mathrm{N}(3, 0.1^2)$



- Latent variances: $\psi^{1/2} \sim \Gamma(1, 0.5) \rightsquigarrow \psi \sim \Gamma(100, 50)$



Outro

- 🚀 INLA methodology offers a rapid alternative for SEM, delivering Bayesian results at near the speed of frequentist estimators.
  - Not aiming to replace MCMC; use case is model development, testing, and exploratory analysis.
- Forward looking
  - Exploratory factor analysis
  - Non-Gaussian outcomes (e.g. categorical, count)
  - Group-/cluster-aware predictions (for LOOCV)
  - Better prior support (priors for correlation matrices Freni-Sterrantino et al., 2025; penalised complexity priors Simpson et al., 2017)
  - Spatiotemporal extensions (Krainski et al., 2018)
- Vignettes and tutorials available on https://inlavaan.haziqj.ml/: Equality constraints, defined parameters, multigroup, multilevel, and missing data.


شكراً جزيلاً
https://haziqj.ml/psychoco26

References

Bollen, K. A. (1989). *Structural equations with latent variables* (pp. xiv, 514). John Wiley & Sons. https://doi.org/10.1002/9781118619179

Freni-Sterrantino, A., Rustand, D., Van Niekerk, J., Krainski, E., & Rue, H. (2025). A graphical framework for interpretable correlation matrix models for multivariate regression. *Statistical Methods & Applications*, *34*(3), 409–447. https://doi.org/10.1007/s10260-025-00788-y

Krainski, E., Gómez-Rubio, V., Bakka, H., Lenzi, A., Castro-Camilo, D., Simpson, D., Lindgren, F., & Rue, H. (2018). *Advanced spatial modeling with stochastic partial differential equations using R and INLA*. Chapman and Hall/CRC.

Nelsen, R. B. (2006). *An introduction to copulas* (2nd ed). Springer.

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian Inference for Latent Gaussian models by using Integrated Nested Laplace Approximations. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *71*(2), 319–392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors. *Statistical Science*, *32*(1), 1–28. https://doi.org/10.1214/16-STS576

van Niekerk, J., & Rue, H. (2024). Low-rank Variational Bayes correction to the Laplace method. *Journal of Machine Learning Research*.

Zellner, A. (1988). Optimal Information Processing and Bayes's Theorem. *The American Statistician*, *42*(4), 278–280. https://doi.org/10.2307/2685143