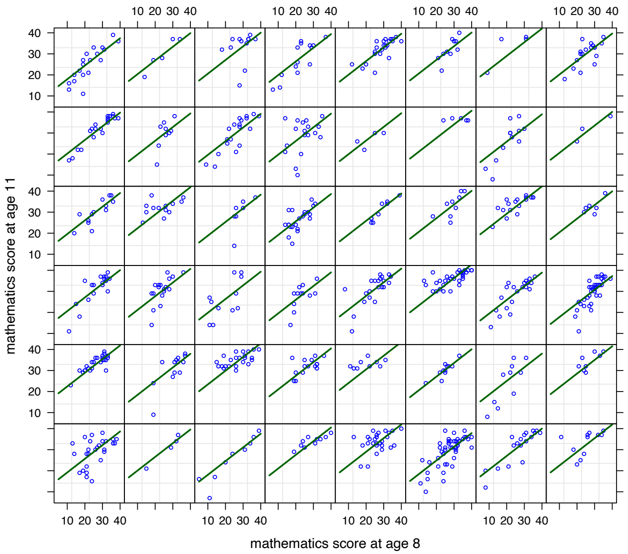
37457

Advanced Bayesian Methods

Bayesian Models for Grouped Data: Additional Aspects



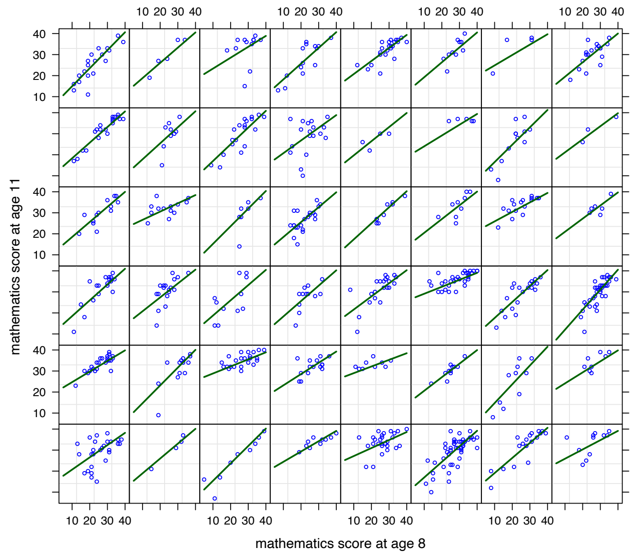
Random Intercepts Model Recap

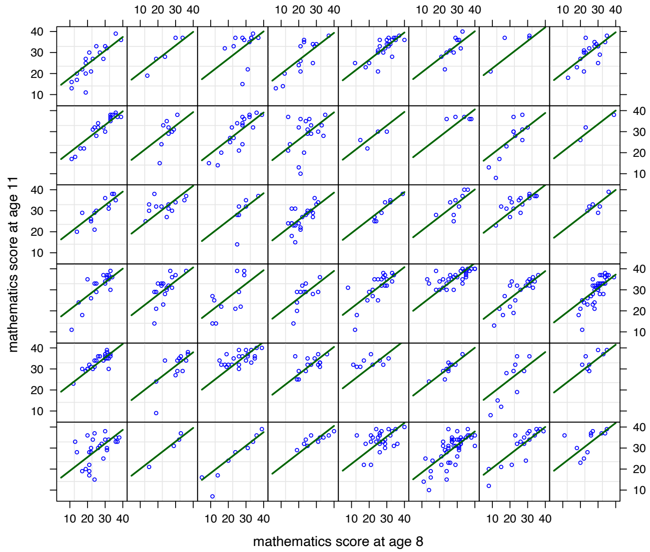
The previous slide fit the Bayesian random intercepts model

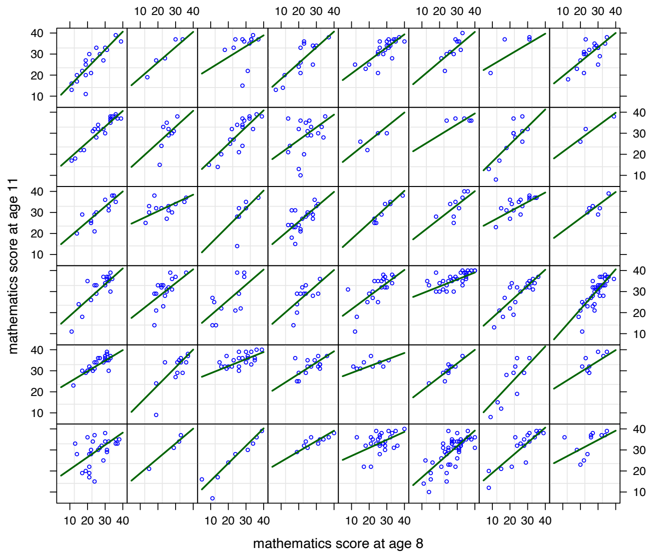
$$y_{ij}|\beta_0, \beta_1, u_i, \sigma_\varepsilon^2 \stackrel{\text{ind.}}{\sim} N\Big((\beta_0 + u_i) + \beta_1 x_{ij}, \sigma_\varepsilon^2\Big)$$

$$u_i|\sigma_u^2 \stackrel{\text{ind.}}{\sim} N\big(0, \sigma_u^2\big)$$

LIMITATION: The lines for each school are parallel (i.e. same slope).







Random Intercepts and Slopes Models

$$y_{ij}|\beta_0, \beta_1, u_{0i}, u_{1i}, \sigma_\varepsilon^2 \overset{\text{ind.}}{\sim} N\Big((\beta_0 + u_{0i}) + (\beta_1 + u_{1i})\,x_{ij}, \sigma_\varepsilon^2\Big),$$

$$\left[\begin{array}{c} u_{0i} \\ u_{1i} \end{array}\right] \Bigg| \sigma_{u0}^2, \sigma_{u1}^2, \rho_u \overset{\text{ind.}}{\sim} N\left(\left[\begin{array}{c} 0 \\ 0 \end{array}\right], \left[\begin{array}{cc} \sigma_{u0}^2 & \rho_u\sigma_{u0}\sigma_{u1} \\ \rho_u\sigma_{u0}\sigma_{u1} & \sigma_{u1}^2 \end{array}\right]\right).$$

Note that

$$\left[\begin{array}{cc} \sigma_{u0}^2 & \rho_u\sigma_{u0}\sigma_{u1} \\ \rho_u\sigma_{u0}\sigma_{u1} & \sigma_{u1}^2 \end{array}\right]$$

is a general 2×2 covariance matrix where

$-1 < \rho_u < 1$ is the correlation between the u_{0i} and u_{1i}.

Look at Stan code

HALF-CAUCHY PRIORS

FOR STANDARD DEVIATION

PARAMETERS

Variance Component Prior Debate (early 2000s)

Until about 2000 most Bayesian analysts used

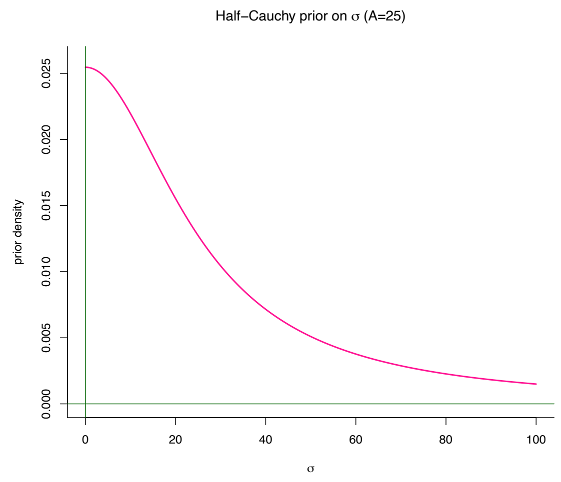
$$p(\sigma^2) \sim \text{Inverse-Gamma}(\varepsilon, \varepsilon) \quad (\varepsilon \text{ 'small'})$$

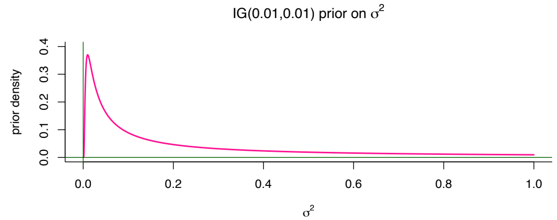
as a 'vague' prior for variance components – because of its conjugate status.

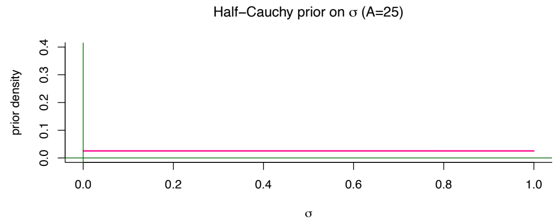
Recent literature (e.g. Gelman, 2006, *Bayesian Analysis*) has criticised this choice since the Inverse-Gamma$(\varepsilon, \varepsilon)$ distribution is not very vague.

Gelman's default recommendation is half-Cauchy with high scale parameter.

$$p(\sigma) = \frac{2A}{\pi(\sigma^2 + A^2)}, \quad \sigma^2 > 0, \quad A \text{ 'big'}$$







Look at Stan code

Covariance Matrix Priors

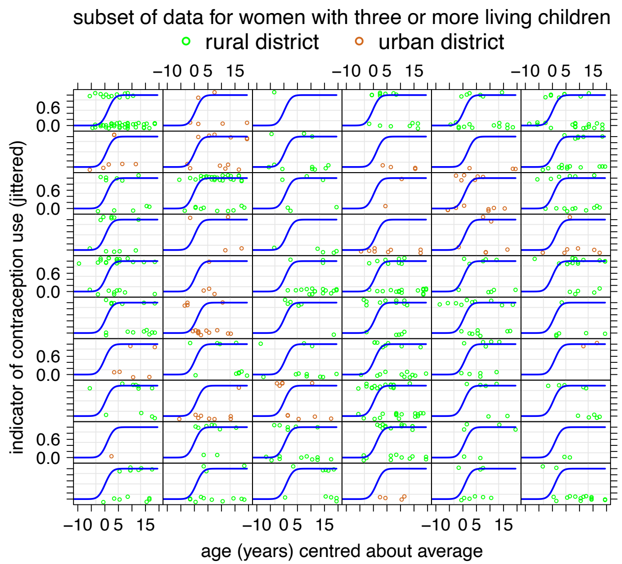
In 2012 (here in Ultimo, N.S.W.) Alan Huang and Matt Wand extended the Gelman Half-Cauchy prior idea to covariance matrix priors and published:

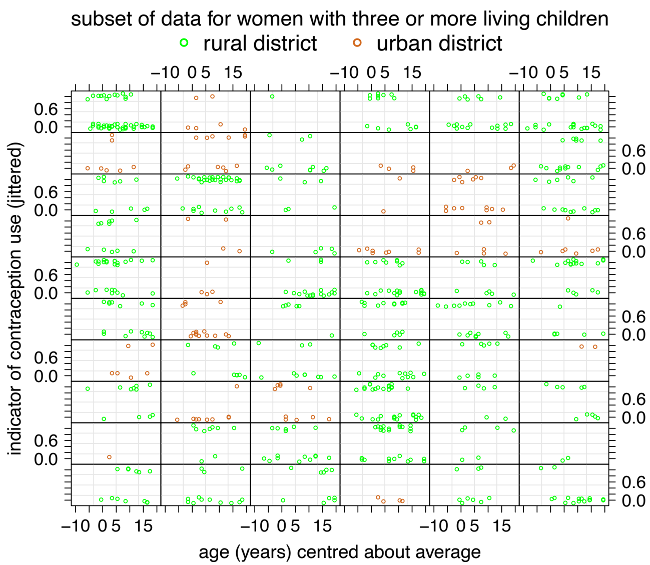
Huang, A. and Wand, M.P. (2013). Simple marginally noninformative prior distributions for covariance matrices. *Bayesian Analysis*, **8**, 439–452.

(Naturally) the class's Stan scripts use this prior – but technical details left aside (but Huang & Wand paper on lecturer's web-pages).

BINARY RESPONSE

GROUPED DATA





Logistic Mixed Models

An example is:

$$y_{ij}|\beta_0,\beta_1,\beta_2,\beta_3 \overset{\text{ind.}}{\sim} \text{Bernoulli}\left(\frac{1}{1+\exp[-\{\beta_0+u_{i0}+(\beta_1+u_{i1})x_{1ij}+\beta_2\,x_{2i}\}]}\right),$$

$$\begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix}\Bigg|\,\sigma^2_{u0},\sigma^2_{u1},\rho_u \overset{\text{ind.}}{\sim} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix},\begin{bmatrix} \sigma^2_{u0} & \rho_u\sigma_{u0}\sigma_{u1} \\ \rho_u\sigma_{u0}\sigma_{u1} & \sigma^2_{u1} \end{bmatrix}\right)$$

Assignment 6, Question 4 has an example of this type.

Let's look at coding in Stan...

Markov Chain Monte Carlo Making a Difference

Without Markov chain Monte Carlo, fitting logistic mixed models with accurate inference is VERY HARD due to intractable integrals (even in the non-Bayesian case).

The availability of Bayesian inference engines makes such analyses routine – but this is quite a recent development for such an important model.

♦ ♦ ♦