§ ❻ Hierarchical (Multi-Stage) Generalized Linear Models

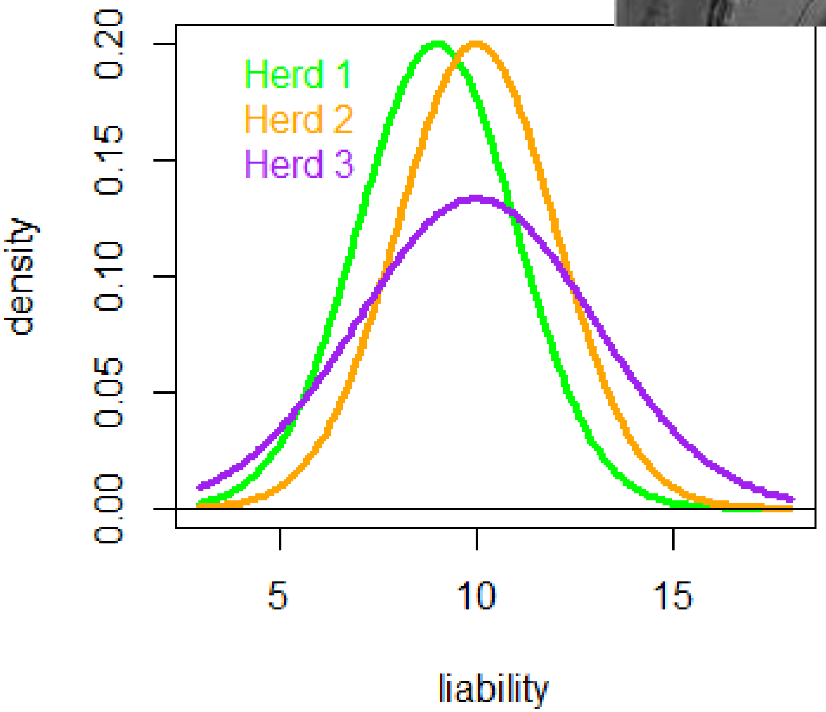
Robert J. Tempelman

Introduction

- Some inferential problems require non-classical approaches; e.g.
  - Heterogeneous variances and covariances across environments.
  - Different distributional forms (e.g. heavy-tailed or mixtures for residual/random effects).
  - High dimensional variable selection models
- Hierarchical Bayesian modeling provides some flexibility for such problems.

Heterogeneous variance models (Kizilkaya and Tempelman, 2005)

- Consider a study involving different subclasses (e.g. herds)
  - Mean responses are different.
  - But suppose residual variances are different too.
- Let's discuss in context of LMM (linear mixed model)



Recall linear mixed model

- Given:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

$$\mathbf{e} \sim N\left(\mathbf{0}, \mathbf{R}(\boldsymbol{\xi})\right)$$

$\mathbf{R}(\boldsymbol{\xi})$ has a certain "heteroskedastic" specification.

$$\mathbf{y} \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\xi} \sim p\left(\mathbf{y} \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\xi}\right) = N\left(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}, \mathbf{R}(\boldsymbol{\xi})\right)$$

- $\boldsymbol{\xi}$ determines the nature of heterogeneous residual variances

Modeling Heterogeneous Variances

- Suppose $\mathbf{e}' = \begin{bmatrix} \mathbf{e}'_{11} & \mathbf{e}'_{12} & \cdots & \mathbf{e}'_{st} \end{bmatrix}$

$$\mathbf{e}_{kl} \sim \mathrm{N}\left(\mathbf{0}, \mathbf{R}_{kl}\left(\boldsymbol{\xi}\right) = \mathbf{I}_{n_{kl}} \sigma^2_{e_{kl}}\right)$$

$$\sigma^2_{e_{kl}} = \sigma^2_e \gamma_k v_l;\ k = 1, 2, \ldots s;\ l = 1, 2, \ldots t.$$

  - with σ^2_e as a “fixed” intercept residual variance
  - $\gamma_k > 0$ k^{th} fixed scaling effect.
  - $v_l > 0$ l^{th} random scaling effect.

Subjective and Subjective Priors

- "intercept" variance $\sigma_e^2 \sim p\left(\sigma_e^2\right)$: subjective flat or conjugate vague inverted-gamma (IG) prior
- Invoke typical constraints for "fixed effects"
  - Corner parameterization: γ_s= 1.
  - Flat or vague IG prior p(γ_k); k=1,2,..,s
- Structural prior for "random effects"

$$p(v_l \mid \alpha_e) = \frac{(\alpha_e - 1)^{\alpha_e}}{\Gamma(\alpha_e)} (v_l)^{-(\alpha_e+1)} \exp\left(-\frac{(\alpha_e - 1)}{v_l}\right)$$

  - i.e., $v_l \sim$ IG(α_e, α_e -1).
    - E(v_l)=1; $\sigma_v^2 = \mathrm{Var}(v_l) = \frac{1}{\alpha_e - 2}$

α_e functions like a "variance component" for residual variances.-> hyperparameter

$$CV(v_l \mid \alpha_e) = \sqrt{\frac{1}{\alpha_e - 2}}$$

Remaining priors

- "Classical" random effects

$$\mathbf{u}|\boldsymbol{\varphi} \sim p(\mathbf{u}|\boldsymbol{\varphi}) = \mathrm{N}\left(\mathbf{0},\mathbf{G}(\boldsymbol{\varphi})\right)$$

- "Classical" fixed effects

$$\boldsymbol{\beta} \sim p(\boldsymbol{\beta})$$

- "Classical" random effects VC

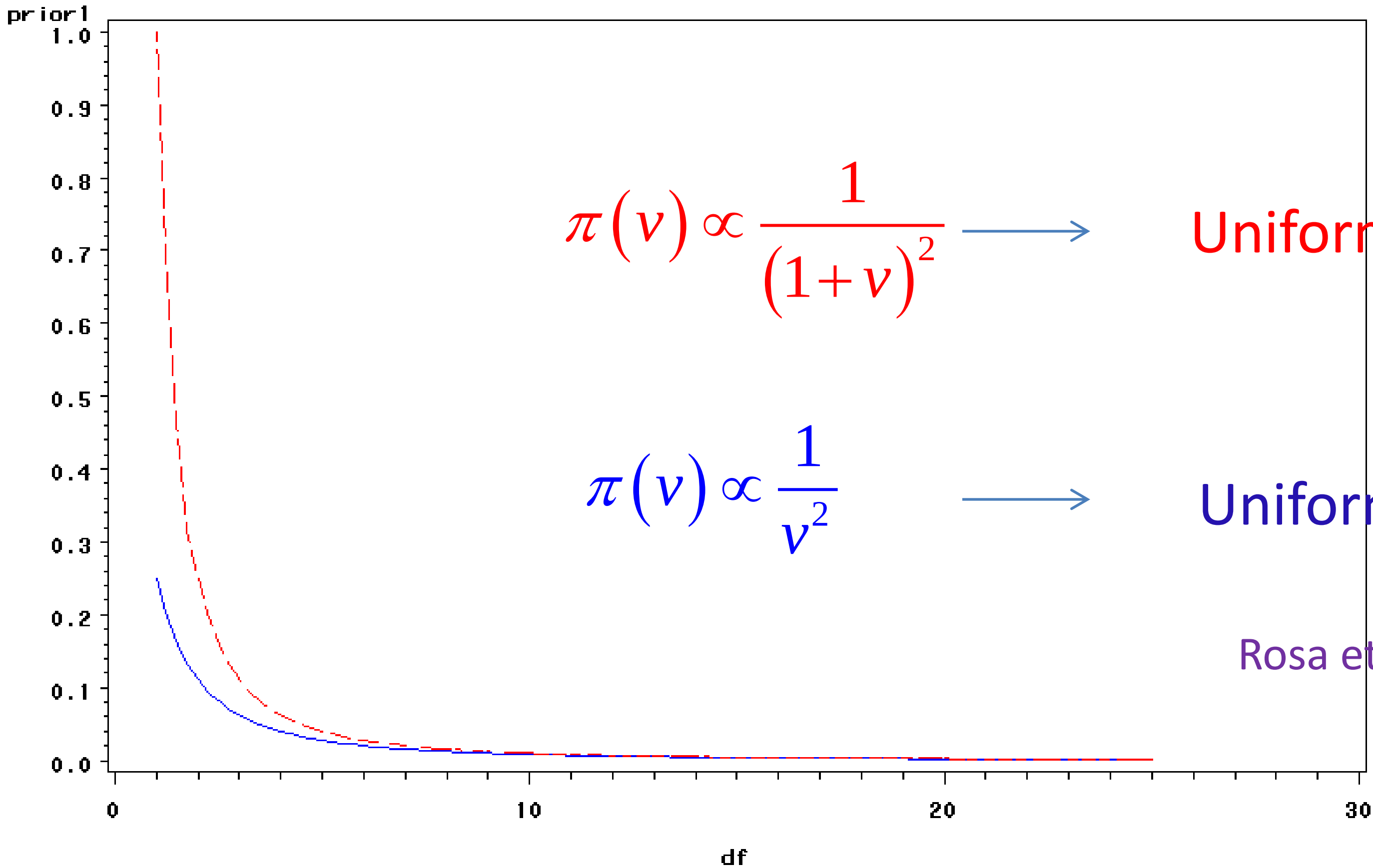
$$\boldsymbol{\varphi} \sim p(\boldsymbol{\varphi})$$

- Hyperparameter (Albert, 1988)

$$\alpha_e \sim p(\alpha_e) = \frac{1}{(1+\alpha_e)^2}$$

SAS PROC MCMC doesn't seem to handle this…prior can't be written as function of corresponding parameter

What was the last prior again???



$$\pi(v) \propto \frac{1}{(1+v)^2} \longrightarrow \text{Uniform(0,1) on } \frac{1}{(1+v)}$$

$$\pi(v) \propto \frac{1}{v^2} \longrightarrow \text{Uniform(0,1) on } \frac{1}{v}$$

Rosa et al. (2004)

Different diffuse priors can have different impacts on posterior inferences!...if data info is poor

Joint Posterior Density

- LMM:

$$p\left(\boldsymbol{\beta}, \mathbf{u}, \sigma_e^2, \boldsymbol{\gamma}, \mathbf{v}, \boldsymbol{\varphi}, \alpha_v \mid \mathbf{y}\right)$$

$$\propto p\left(\mathbf{y} \mid \boldsymbol{\beta}, \mathbf{u}, \sigma_e^2, \boldsymbol{\gamma}, \mathbf{v}\right) p(\boldsymbol{\beta}) p(\mathbf{u} \mid \boldsymbol{\varphi}) p(\boldsymbol{\varphi})$$

$$p\left(\sigma_e^2\right)\left(\prod_{k=1}^{s} p\left(\gamma_k\right)\right)\left(\prod_{l=1}^{t} p\left(v_l \mid \alpha_e\right)\right) p\left(\alpha_e\right)$$

Details on FCD

- All provided by Kizilkaya and Tempelman (2005)
  - All are recognizeable except for α_v:

$$p\left(\alpha_e \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\varphi}, \boldsymbol{\gamma}, \mathbf{v}, \mathbf{y}, \mathbf{L}, \boldsymbol{\tau}\right)$$

$$\propto \frac{\left(\alpha_e - 1\right)^{\alpha_e t}}{\left(\Gamma\left(\alpha_e\right)\right)^t} \exp\left(-\left(\alpha_e - 1\right)\sum_{l=1}^{t} v_l^{-1}\right)\prod_{l=1}^{t}\left(v_l\right)^{-\left(\alpha_e + 1\right)} p\left(\alpha_e\right)$$

  - Use Metropolis-Hastings random walk on $\psi_e = \log\left(\alpha_e\right)$ using normal as proposal density.
    - For MH, generally a good idea to transform parameters so that parameter space is entire real line…but don't forget to include Jacobian of transform.