Generalized Linear Mixed Model (GLMM)

$$\eta_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i,$$

where $\boldsymbol{\beta}$ are parameters common to all subjects and $\mathbf{b}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Omega})$ are deviations for subject i.

If we choose $\mathbf{x}_{ij} = \mathbf{z}_{ij}$, then all the regression coefficients are assumed to vary for the different study subjects, with the amount of variance dependent on $\boldsymbol{\Omega}$.

If we choose $\mathbf{z}_{ij} = 1$, then only the intercept varies for the different study subjects (random intercept model)

The GLMM is a generalization of the Linear Mixed Effects model of Laird and Ware (1982, *Biometrics*), which is also referred to as a random-effect model.

The Linear Mixed Effects Model has the form,

$$y_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i + \epsilon_{ij},$$

where $\boldsymbol{\beta}$ are "fixed" effects, $\mathbf{b}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Omega})$ are "random" effects, and $\epsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$ is an error residual (which is assumed independent of $\mathbf{b}_i$).

The Linear Mixed Effects Model can be written as

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i,$$

where $\mathbf{X}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{i,n_i})'$, $\mathbf{Z}_i = (\mathbf{z}_{i1}, \ldots, \mathbf{z}_{i,n_i})'$, $\boldsymbol{\epsilon}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$, and $\boldsymbol{\Sigma} = \sigma^2\,\mathrm{I}_{n_i \times n_i}$

Hence, $\mathrm{E}(\mathbf{y}_i) = \mathbf{X}_i\boldsymbol{\beta} + 0 + 0$ and $\mathrm{V}(\mathbf{y}_i) = \mathbf{Z}_i\boldsymbol{\Omega}\mathbf{Z}_i' + \boldsymbol{\Sigma}$

So we have

$$\mathbf{y}_i \sim \mathrm{N}(\mathbf{X}_i\boldsymbol{\beta}, \mathbf{Z}_i\boldsymbol{\Omega}\mathbf{Z}_i' + \boldsymbol{\Sigma})$$

Bayesian Analysis for Linear Mixed Model

(Gilks et al., 1993)

Likelihood Function (Random-intercept case):

$$f(\mathbf{y} \mid b_i, \boldsymbol{\beta}, \sigma^2, \mathbf{X}) = \prod_{i=1}^{n} \prod_{j=1}^{n_i} (2\pi\sigma^2)^{-1/2} \exp\Big\{ -\frac{1}{2\sigma^2}(y_{ij} - \mathbf{x}'_{ij}\boldsymbol{\beta} - b_i)^2 \Big\},$$

where $b_i \sim \mathrm{N}(0, \psi^{-1})$.

Conditionally Conjugate Priors:

$$\boldsymbol{\beta} \sim \mathrm{N}(\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_\beta), \quad \sigma^{-2} \sim \mathrm{G}(c_{01}, d_{01}), \quad \text{and} \quad \psi \sim \mathrm{G}(c_{02}, d_{02}).$$

Gibbs Sampler

Alternately sample from the full conditional posterior distributions of each of the model unknowns

Full conditional distribution of $\boldsymbol{\beta}$:

$$\pi(\boldsymbol{\beta} \,|\, \mathbf{b}, \sigma^2, \psi, \mathbf{y}, \mathbf{X}) \stackrel{d}{=} \mathrm{N}(\widehat{\boldsymbol{\Sigma}}_\beta\{\boldsymbol{\Sigma}_\beta^{-1}\boldsymbol{\beta}_0 + \sigma^{-2}\sum_{i=1}^{n}\mathbf{X}_i'(\mathbf{y}_i - b_i)\}, \widehat{\boldsymbol{\Sigma}}_\beta),$$

where $\widehat{\boldsymbol{\Sigma}}_\beta = (\boldsymbol{\Sigma}_\beta^{-1} + \sigma^{-2}\mathbf{X}'\mathbf{X})^{-1}$.

Full conditional distribution of b_i, for $i = 1, \ldots, n$:

$$\pi(b_i \,|\, \boldsymbol{\beta}, \sigma^2, \psi, \mathbf{y}, \mathbf{X}) \stackrel{d}{=} \mathrm{N}\Big(\frac{\sigma^{-2}\sum_{j=1}^{n_i}(y_{ij} - \mathbf{x}_{ij}'\boldsymbol{\beta})}{\psi + \sigma^{-2}n_i}, \frac{1}{\psi + \sigma^{-2}n_i}\Big).$$

Full conditional distribution of σ^{-2}:

$$\pi(\sigma^{-2}\,|\,\mathbf{b},\boldsymbol{\beta},\psi,\mathbf{y},\mathbf{X}) \overset{d}{=} \mathrm{G}\Big(c_{01}+\frac{\sum_{i=1}^{n} n_i}{2}, d_{01}+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n_i}(y_{ij}-\mathbf{x}_{ij}'\boldsymbol{\beta}-b_i)^2\Big).$$

Full conditional distribution of ψ:

$$\pi(\psi\,|\,\mathbf{b},\boldsymbol{\beta},\sigma^2,\mathbf{y},\mathbf{X}) \overset{d}{=} \mathrm{G}\Big(c_{02}+\frac{n}{2}, d_{02}+\frac{1}{2}\sum_{i=1}^{n} b_i^2\Big).$$

The generalization to multiple random-effects is straightforward, by assuming a conditionally-conjugate Wishart prior for the random-effects precision.

Deriving these conditional distributions & their form should be second nature.

Probit Models for Multiple Binary Responses

We can use the linear mixed model machinery in combination with the Albert and Chib (1993) algorithm to account for dependency in multiple binary response data

Suppose that $\mathbf{y}_i = (y_{i1}, \ldots, y_{in_i})'$ consists of repeated binary responses for subject i, and we have

$$\Pr(y_{ij} = 1 \,|\, b_i, \mathbf{x}_{ij}) = \Phi(\mathbf{x}_{ij}'\boldsymbol{\beta} + b_i),$$

where $b_i \sim \mathrm{N}(0, \psi^{-1})$.

We introduce independent latent $z_{ij} \sim \mathrm{N}(\mathbf{x}_{ij}'\boldsymbol{\beta} + b_i, 1)$ underlying the observed binary response:

$$y_{ij} = 1 \text{ if } z_{ij} > 0 \quad y_{ij} = 0 \text{ if } z_{ij} \leq 0.$$

It is straightforward to implement the Gibbs sampler in this case.

After defining conditionally-conjugate priors,

$$\boldsymbol{\beta} \sim \mathrm{N}(\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_\beta) \quad \text{and} \quad \psi \sim \mathrm{G}(c_1, d_1),$$

we simply alternate between

1. Sample z_{ij} from its truncated normal full conditional distribution
2. Sample $\boldsymbol{\beta}$ from its normal full conditional distribution
3. Sample b_i from its normal full conditional distribution
4. Sample the variance component ψ from its gamma full conditional

Correlation between y_{ij} and $y_{ij'}$ under the normal linear random-intercept model,

$$\rho(y_{ij}, y_{ij'}) = \frac{\mathrm{E}[y_{ij}y_{ij'}] - \mathrm{E}[y_{ij}]\mathrm{E}[y_{ij'}]}{\sqrt{V[y_{ij}]V[y'_{ij}]}} = \frac{\psi^{-1}}{\psi^{-1} + \sigma^2}.$$

Thus, when the random-effect variance is ≈ 0, the correlation is low and the between-subject variance is high relative to the within-subject variance

For random-effects probit models, the correlation between y_{ij} and $y_{ij'}$ is not available in closed form. However, the correlation in z_{ij} and $z_{ij'}$ is often used as a surrogate,

$$\rho(z_{ij}, z_{ij'}) = \frac{\psi^{-1}}{\psi^{-1} + 1}.$$

Bayesian Variable Selection & Hypothesis Testing

Notes on Stochastic Search Variable Selection (SSVS), Bayes factors, and general approaches for hypothesis testing and variable selection will be written on board in class.