ADVANCED BAYESIAN MODELING

Grouping and Random Effects

Situation:

Observations are divided into *groups* (or *batches*, or *classes*).

Observations in the same group have something in common, represented by sharing common parameter values.

Responses from observations within a group are often exchangeable.

Responses are *not* exchangeable overall.

Responses and Groups

Assume J groups, and scalar responses

$$y_{ij} \quad = \quad \text{response from observation } i \text{ in group } j \qquad\qquad j = 1, \ldots, J$$

Typically, index i runs within group:

$$i = 1, \ldots, n_j \qquad\qquad (n_j \text{ observations in group } j)$$

We consider normal-theory models:

$$y_{ij} \quad \sim \quad \mathrm{N}(\,\cdot\,,\, \sigma_y^2)$$

The y_{ij}s are not necessarily independent, except between groups.

Often useful to regard groups as random – as if sampled from a population of groups.

Why?

- To improve estimation of group parameters
- To allow generalization to other groups not in data

Example

$$y_{ij} \quad = \quad \text{test score of student } i \text{ in school } j$$

Students are grouped into schools.

There may be differences not only between students but between schools.

Regarding schools as a random sample helps to estimate their performance, compare them, and extend results to other schools.

A model should allow for variation in performance both within schools and between schools.

Hierarchical Model

$$
\begin{aligned}
y_{ij} \mid \beta_j, \sigma_y^2 \quad &\sim \quad \text{indep. } \mathrm{N}(\beta_j,\, \sigma_y^2) \\
\beta_j \mid \alpha, \sigma_\beta^2 \quad &\sim \quad \text{iid } \mathrm{N}(\alpha,\, \sigma_\beta^2)
\end{aligned}
$$

Parameters β_j are sometimes called **random effects**.

In the example,

$$
\begin{aligned}
\beta_j \quad &= \quad \text{effect due to school } j \\
\alpha \quad &= \quad \text{mean of school effect distribution} \\
\sigma_\beta^2 \quad &= \quad \text{variance of school effect distribution}
\end{aligned}
$$

Responses *within a group* are exchangeable (conditionally iid).

Random effects β_j are exchangeable.

Responses overall are *not* exchangeable: Either have different β_js (conditionally), or different correlations (marginally).

The normal distribution for the β_js is (partially) conjugate.

σ_y^2 and σ_β^2 are called **variance components**:

$$\begin{aligned} \sigma_y^2 &= \text{variance among observations within a group} \\ \sigma_\beta^2 &= \text{variance among groups} \end{aligned}$$

Marginally over the β_js, the correlation between two responses is

$$\begin{array}{ll} \dfrac{\sigma_\beta^2}{\sigma_\beta^2 + \sigma_y^2} & \text{for observations in the same group: } \textit{intraclass correlation} \\ 0 & \text{for observations in different groups} \end{array}$$

because only observations in the same group share a random effect.

An equivalent hierarchical formulation:

$$
\begin{aligned}
y_{ij} \mid \alpha, \beta_j, \sigma_y^2 &\sim \text{indep. N}(\alpha + \beta_j,\, \sigma_y^2) \\
\beta_j \mid \sigma_\beta^2 &\sim \text{iid N}(0,\, \sigma_\beta^2)
\end{aligned}
$$

Note: α becomes an **intercept**.

Sometimes called a *one-way random effect model*.

Priors

Typical noninformative (hyper)priors:

$$
\begin{aligned}
p(\sigma_y^2) &\propto (\sigma_y^2)^{-1} \qquad \sigma_y^2 > 0 \\
p(\sigma_\beta^2) &\propto (\sigma_\beta^2)^{-1/2} \qquad \sigma_\beta^2 > 0 \\
p(\alpha) &\propto 1
\end{aligned}
$$

Remark: Same as uniform prior on $(\log \sigma_y^2,\, \sigma_\beta,\, \alpha)$.

Note: The unusual prior on σ_β^2 is to avoid an improper posterior.

Matrix Formulations

$$
\begin{aligned}
y &= (y_{11}, \ldots, y_{n_1 1},\ y_{12}, \ldots y_{n_2 2}, \ldots\ \ldots,\ y_{1J}, \ldots, y_{n_J J})^T \\
\beta &= (\beta_1, \ldots, \beta_J)^T
\end{aligned}
$$

Then

$$
\begin{aligned}
y \mid \beta, \sigma_y^2, X &\sim \mathrm{N}(X\beta,\ \sigma_y^2 I) \\
\beta \mid \alpha, \sigma_\beta^2 &\sim \mathrm{N}(\mathbf{1}\alpha,\ \sigma_\beta^2 I)
\end{aligned}
$$

where $\mathbf{1}$ is a $J \times 1$ vector of ones.

X is a matrix of indicators for the group (column) of each observation (row):

$$
x_{ij,k} = 1 \text{ if } j = k, \quad 0 \text{ otherwise}
$$

The alternative matrix formulation (one-way random effect model):

$$\begin{aligned} y \mid \alpha, \beta, \sigma_y^2, X &\sim \mathrm{N}(\mathbf{1}\alpha + X\beta,\ \sigma_y^2 I) \\ \beta \mid \sigma_\beta^2 &\sim \mathrm{N}(0,\ \sigma_\beta^2 I) \end{aligned}$$

Advantage: Generalizes more readily to *mixed models* ...