Hierarchical Bayes

...plus graph notation and a review
Assumptions of Linear Model

- Homoskedasticity
- No error in X variables
- No missing data
- Normally distributed error
- Error in Y variables is measurement error
- Observations are independent

Model variance
Errors in variables
Missing data model
GLM
Graph notation

- Focuses on relationships among parameters and data sets rather than distributions
- Can facilitate writing conditional distributions

\[ X \sim N(\mu, \sigma^2) \]
\[ \mu \sim N(\mu_0, V_\mu) \]
\[ \sigma^2 \sim IG(s_1, s_2) \]
Linear Regression

\[ \mathbf{\tilde{y}} \sim N(\mathbf{X} \hat{\mathbf{\beta}}, \sigma^2) \]
Heteroskedasticity

\[ y \sim N(\beta_1 + \beta_2 x, (\alpha_1 + \alpha_2 x)^2) \]
Errors in Variables

\[ \tilde{y} \sim N( X \tilde{\beta}, \sigma^2) \]
\[ x^{(o)} \sim N( x, \tau^2) \]
Missing Data Model

\[ \tilde{y} \sim N(\mathbf{X} \tilde{\mathbf{\beta}}, \sigma^2) \]
Logistic Regression

\[ \tilde{y} \sim \text{Binom}(1, \logit^{-1}(X\hat{\beta})) \]
Poisson Regression

\[ \tilde{y} \sim \text{Pois}(\exp(X\beta)) \]

\[ X \rightarrow Y \]

Data Model

\[ \beta \]

Process Model

\[ B_0, V_\beta \]

Parameter Model
Hierarchical Models

Common mean

Hierarchical

\[ \mu \]

Independent

\[ \mu_1 \quad \mu_2 \quad \mu_3 \]
Common Mean

\[ \tilde{y}_k \sim N(\mu, \sigma^2) \]
Hierarchical Mean, Common Variance

\[ Y_k \sim N\left( \mu_k, \sigma^2 \right) \]
\[ \mu_k \sim N\left( \mu, \tau^2 \right) \]
\[ \sigma^2 \sim IG\left( s_1, s_2 \right) \]
\[ \mu \sim N\left( \mu_0, V_\mu \right) \]
\[ \tau^2 \sim IG\left( t_1, t_2 \right) \]

At this point, this model is fitting each data set independently but assume the mean for each has the same prior.

For the hierarchical model, instead assume the prior contains unknown model parameters.

Then need to specify \textit{hyperpriors} on our prior.
Hierarchical Mean

Data Model

\[ Y \ldots Y \ldots Y \]

Process Model

\[ \mu_1, \mu_k, \mu_n, \sigma^2 \]

Parameter Model

\[ \mu, \tau^2, s_1, s_2 \]

Hyperparameters

\[ m_0, V_\mu, t_1, t_2 \]
Hierarchical Models

- Model variability in the parameters of a model
- Partition variability more explicitly into multiple terms
- Borrow strength across data sets

- Details usually in the SUBSCRIPTS
- Hierarchical with respect to parameters
Random Effects

- Common special case of Hierarchical models

\[ Y_k \sim N(\mu_k, \sigma^2) \]
\[ \mu_k \sim N(\mu, \tau^2) \]
\[ \sigma^2 \sim IG(s_1, s_\tau) \]
\[ \mu \sim N(\mu_\cdot, V_\mu) \]
\[ \tau^\tau \sim IG(t_\cdot, t_2) \]

\[ Y_k \sim N(\mu_g + \alpha_k, \sigma^2) \]
\[ \alpha_k \sim N(0, \tau^2) \]
\[ \sigma^2 \sim IG(s_1, s_2) \]
\[ \mu_g \sim N(\mu_0, V_\mu) \]
\[ \tau^2 \sim IG(t_1, t_2) \]
Random Effects

- Common special case of Hierarchical models

\[ Y_k \sim N(\mu_g + \alpha_k, \sigma^r) \]
\[ \alpha_k \sim N(0, \tau^2) \]
\[ \sigma^2 \sim IG(s_1, s_2) \]
\[ \mu_g \sim N(\mu_0, V_\mu) \]
\[ \tau^2 \sim IG(t_1, t_2) \]
Random Effects

\[ Y_k \sim N(\mu_g + \alpha_k, \sigma^2) \]
\[ \alpha_k \sim N(0, \tau^2) \]
\[ \sigma^2 \sim IG(s_\sigma, s_\tau) \]
\[ \mu_g \sim N(\mu, V_\mu) \]
\[ \tau^2 \sim IG(t_1, t_2) \]

- Random effects always have mean 0
- Random effects variance attributes a portion of uncertainty to a specific source
- Can be used to try an account for a lack of independence
Random Effects Mean

Data Model

Process Model

Parameter Model

Hyperparameters
What things can be random effects?

- Traditionally, random effects apply to aspects of the study that would not be the same if replicated
  - e.g. Plot, Block, Year, individual, etc.
  - Often used to account for a lack of independence
- Treatments and covariates of interest are usually treated as **fixed effects**
- Typically there is some degree of replication otherwise the random effect is not identifiably different from the residual “noise” term $\varepsilon \sim N(0, \sigma^2)$
### Random Effects Linear Model

<table>
<thead>
<tr>
<th>Component</th>
<th>Mathematical Formulation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Effects</td>
<td>$\mu_{i,k} = X_i \beta + \alpha_k + \epsilon_{i,k}$</td>
<td>Process model</td>
</tr>
<tr>
<td>Random Effect</td>
<td>$\epsilon_{i,k} \sim N(0, \sigma^2)$</td>
<td>Data model</td>
</tr>
<tr>
<td>Residual Error</td>
<td>$\alpha_k \sim N(0, \tau^2)$</td>
<td>Random effect</td>
</tr>
<tr>
<td>Error Variance Prior</td>
<td>$\sigma^2 \sim IG(s_1, s_2)$</td>
<td>Error variance prior</td>
</tr>
<tr>
<td>Fixed Effects Prior</td>
<td>$\beta \sim N(B_0, V_\beta)$</td>
<td>Fixed effects prior</td>
</tr>
<tr>
<td>Random Effects Variance</td>
<td>$\tau^2 \sim IG(t_1, t_2)$</td>
<td>Random effects variance prior</td>
</tr>
</tbody>
</table>
Random Effects Linear Model

Data Model

\[ X_1 \ldots X_k \ldots X_n \]

\[ Y_1 \ldots Y_k \ldots Y_n \]

Process Model

\[ \alpha_1 \alpha_k \alpha_n \beta, \sigma^2 \]

Parameter Model

\[ \tau^2 B_0, V \beta s_1, s_2 \]

Hyperparameters

\[ t_1, t_2 \]
Why bother? Impacts on inference...

FIGURE 8.5. Two simulated longitudinal data sets with $n = 10$ and the same total variance, but dominated by individual differences (a) or process error (b). Variance parameters in (a) are $\tau^2 = 0.09$ and $\sigma^2 = 0.01$ and in (b) are $\tau^2 = 0.01$ and $\sigma^2 = 0.09$. 
Figure 3 The impact of random individual effects (RITEs) on coexistence of two competing species. Two spatiotemporal and individual-based simulations were run using recruitment processes that are parameterized with data, summarized in Fig. 1. Panel (a) is the traditional approach having deterministic species differences and stochasticity in time, but no within-population heterogeneity, reflecting that fact the green species is the deterministic winner (Fig. 1a). Population heterogeneity in (b) means that green is not the deterministic winner, but rather both species win with some probability.
Explaining unexplained variance

- Random effects attempt to account for the unexplained variance associated with some group (plot, year, etc.) due to all the things that were not measured.
- May point to scales that need additional explanation.
- Adding covariates may explain some portion of this variance, but there's always something you didn't measure.
- Sometimes additional fixed effects not justified (model selection).
Example: Year effects

- Consider the number of new young produced per adult female from population of birds.
- Suppose adding a year effect shows significant year-to-year variability that is coherent through the whole population.
- Based on the estimates of the year effects, could look for additional covariates that correlate with these values (e.g. different climate variables) without having to rerun the whole model.
- Could refine the model to add additional drivers.
Modeling Uncertainty

• Overall take home message:

The proper accounting of uncertainty can be JUST AS IMPORTANT to making valid inference from your model as the process model and covariates

• Random effects are used to account for the impacts of unmeasured/unmeasurable covariates
Assumptions of Linear Model

- Homoskedasticity
- No error in X variables
- No missing data
- Normally distributed error
- Residual error in Y variables is measurement error
- Observations are independent

Model variance
Errors in variables
Missing data model
GLM
Hierarchical Models