Linear Mixed-Effects Models
The Linear Mixed-Effects Model

\[ y = X\beta + Zu + e \]

- \( X \) is an \( n \times p \) design matrix of known constants
- \( \beta \in \mathbb{R}^p \) is an unknown parameter vector
- \( Z \) is an \( n \times q \) matrix of known constants
- \( u \) is a \( q \times 1 \) random vector
- \( e \) is an \( n \times 1 \) vector of random errors
The Linear Mixed-Effects Model

- $y = X\beta + Zu + e$

- The elements of $\beta$ are considered to be non-random and are called “fixed effects.”

- The elements of $u$ are random variables and are called “random effects.”

- The elements of the error vector $e$ are always considered to be random variables.
Because the model includes both fixed and random effects (in addition to the random errors), it is called a “mixed-effects” model or, more simply, a “mixed” model.

The model is called a “linear” mixed-effects model because (as we will soon see)

\[ E(y|u) = X\beta + Zu, \]

a linear function of fixed and random effects.
We assume that

\[ E(e) = 0 \quad \text{Var}(e) = R \]

\[ E(u) = 0 \quad \text{Var}(u) = G \]

\[ \text{Cov}(e, u) = 0. \]

It follows that
$$E(y) = E(X\beta + Zu + e)$$
$$= X\beta + ZE(u) + E(e)$$
$$= X\beta$$

$$\text{Var}(y) = \text{Var}(X\beta + Zu + e)$$
$$= \text{Var}(Zu + e)$$
$$= \text{Var}(Zu) + \text{Var}(e)$$
$$= Z\text{Var}(u)Z' + R$$
$$= ZGZ' + R \equiv \Sigma.$$
We usually consider the special case in which

\[
\begin{bmatrix}
  u \\
  e
\end{bmatrix} \sim N\left(\begin{bmatrix}
  0 \\
  0
\end{bmatrix}, \begin{bmatrix}
  G & 0 \\
  0 & R
\end{bmatrix}\right)
\]

\[\implies y \sim N(X\beta, ZGZ' + R).\]
The conditional moments, given the random effects, are

\[ E(y|u) = X\beta + Zu \]

\[ \text{Var}(y|u) = R. \]
Example 1

Suppose a study was conducted to compare two plant genotypes (genotype 1 and genotype 2). Suppose 10 seeds (5 of genotype 1 and 5 of genotype 2) were planted in a total of 4 pots. Suppose 3 genotype 1 seeds were planted in one pot, and the other 2 genotype 1 seeds were planted in another pot. Suppose the same planting strategy was used when planting genotype 2 seeds in the other two pots. The seeds germinated and emerged from the soil as seedlings. After a four-week growing period, each seedling was dried and weighed. Let $y_{ijk}$ denote the weight for genotype $i$, pot $j$, seedling $k$. Provide a linear mixed-effects model for the dry-weight data. Determine $y$, $X$, $\beta$, $Z$, $u$, $G$, $R$, and $\text{Var}(y)$. 

Copyright ©2012 Dan Nettleton (Iowa State University)
Consider the model

\[ y_{ijk} = \mu + \gamma_i + p_{ij} + e_{ijk} \]

\[ p_{11}, p_{12}, p_{21}, p_{22} \overset{i.i.d.}{\sim} N(0, \sigma_p^2) \]

independent of the \( e_{ijk} \) terms, which are assumed to be iid \( N(0, \sigma_e^2) \).

This model can be written in the form

\[ y = X\beta + Zu + e, \text{ where} \]
\[ y = \begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{213} \\ y_{221} \\ y_{222} \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \quad \beta = \begin{bmatrix} \mu \\ \gamma_1 \\ \gamma_2 \end{bmatrix}, \]
\[ Z = \begin{bmatrix}
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
\end{bmatrix} ,
\quad u = \begin{bmatrix}
p_{11} \\
p_{12} \\
p_{21} \\
p_{22} \\
\end{bmatrix} ,
\quad e = \begin{bmatrix}
e_{111} \\
e_{112} \\
e_{113} \\
e_{121} \\
e_{122} \\
e_{211} \\
e_{212} \\
e_{213} \\
e_{221} \\
e_{222} \\
\end{bmatrix} .\]
\[ G = \text{Var}(u) = \text{Var}([p_{11}, p_{12}, p_{21}, p_{22}]) = \sigma_p^2 I_{4\times4} \]

\[ R = \text{Var}(e) = \sigma_e^2 I_{10\times10} \]

\[ \text{Var}(y) = ZGZ' + R = Z\sigma_p^2 I Z' + \sigma_e^2 I = \sigma_p^2 ZZ' + \sigma_e^2 I. \]
\[ ZZ' = \begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
\end{bmatrix} \]
Thus, $\text{Var}(y) = \sigma_p^2 ZZ' + \sigma_e^2 I$ is a block diagonal matrix.

The first block is

$$
\text{Var} \begin{bmatrix}
  y_{111} \\
  y_{112} \\
  y_{113}
\end{bmatrix} =
\begin{bmatrix}
  \sigma_p^2 + \sigma_e^2 & \sigma_p^2 & \sigma_p^2 \\
  \sigma_p^2 & \sigma_p^2 + \sigma_e^2 & \sigma_p^2 \\
  \sigma_p^2 & \sigma_p^2 & \sigma_p^2 + \sigma_e^2
\end{bmatrix}.
$$
Note that

\[
\text{Var}(y_{ijk}) = \sigma_p^2 + \sigma_e^2 \quad \forall \ i, j, k.
\]

\[
\text{Cov}(y_{ijk}, y_{ijk^*}) = \sigma_p^2 \quad \forall \ i, j, \text{ and } k \neq k^*.
\]

\[
\text{Cov}(y_{ijk}, y_{i^*j^*k^*}) = 0 \quad \text{if } i \neq i^* \text{ or } j \neq j^*.
\]

Any two observations from the same pot have covariance \( \sigma_p^2 \).

Any two observations from different pots are uncorrelated.
Note that $\text{Var}(y)$ may be written as $\sigma_e^2 V$ where $V$ is a block diagonal matrix with blocks of the form

$$
\begin{bmatrix}
1 + \sigma_p^2 / \sigma_e^2 & \sigma_p^2 / \sigma_e^2 & \ldots & \sigma_p^2 / \sigma_e^2 \\
\sigma_p^2 / \sigma_e^2 & 1 + \sigma_p^2 / \sigma_e^2 & \ldots & \sigma_p^2 / \sigma_e^2 \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_p^2 / \sigma_e^2 & \sigma_p^2 / \sigma_e^2 & \ldots & 1 + \sigma_p^2 / \sigma_e^2
\end{bmatrix}
$$

Thus, if $\sigma_p^2 / \sigma_e^2$ were known, we would have the Aitken Model.

$$
y = X\beta + \varepsilon, \text{ where } \varepsilon = Zu + e \sim N(0, \sigma^2 V), \sigma^2 \equiv \sigma_e^2.
$$
Thus, if $\sigma_p^2/\sigma_e^2$ were known, we would use GLS to estimate any estimable $C\beta$ by $C\hat{\beta}_{\text{GLS}} = C(X'V^{-1}X)^{-1}X'V^{-1}y$.

However, we seldom know $\sigma_p^2/\sigma_e^2$ or, more generally, $\Sigma$ or $V$.

For the general problem where $\text{Var}(y) = \Sigma$ is an unknown positive definite matrix, we can rewrite $\Sigma$ as $\sigma^2 V$, where $\sigma^2$ is an unknown positive variance and $V$ is an unknown positive definite matrix.

As in our simple example, each entry of $V$ is usually assumed to be a known function of few unknown parameters.
Thus, our strategy for estimating an estimable $C\beta$ involves estimating the unknown parameters in $V$ to obtain

$$C\hat{\beta}_{\text{GLS}} = C(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y.$$ 

In general,

$$C\hat{\beta}_{\text{GLS}} = C(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y$$

is an nonlinear estimator that is an approximation to

$$C\hat{\beta}_{\text{GLS}} = C(X'V^{-1}X)^{-1}X'V^{-1}y,$$

which would be the BLUE of $C\beta$ if $V$ were known.
In special cases, $C\hat{\beta}_{\text{GLS}}$ may be a linear estimator.

For example, if there exists a matrix $Q$ such that $VX = XQ$, then we know that

$$C\hat{\beta}_{\text{GLS}} = C\hat{\beta} \quad \text{and} \quad C\hat{\beta}_{\text{GLS}} = C\hat{\beta},$$

which is a linear estimator of $C\beta$. 
However, even for our simple example involving seedling dry weight, $\hat{C}_G \hat{\beta}_{\text{GLS}}$ is a nonlinear estimator of $C\beta$ for

$C = [1, 1, 0] \iff C\beta = \mu + \gamma_1,$

$C = [1, 0, 1] \iff C\beta = \mu + \gamma_2,$ and

$C = [0, 1, -1] \iff C\beta = \gamma_1 - \gamma_2.$

Confidence intervals and tests for these estimable functions are not exact.
In our simple example involving seedling dry weight, there was only one random factor (pot).

When there are $m$ random factors, we can partition $Z$ and $u$ as

$$Z = [Z_1, \ldots, Z_m] \quad \text{and} \quad u = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix},$$

where $u_j$ is the vector of random effects associated with factor $j$ $(j = 1, \ldots, m)$. 
We can write $Zu$ as

$$\begin{bmatrix} Z_1, \ldots, Z_m \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} = \sum_{j=1}^{m} Z_j u_j.$$

We often assume that all random effects (including random errors) are mutually independent and that the random effects associated with the $j$th random factor have variance $\sigma_j^2$ ($j = 1, \ldots, m$).

Under these assumptions,

$$\text{Var}(y) = ZGZ' + R = \sum_{j=1}^{m} \sigma_j^2 Z_j Z_j' + \sigma_e^2 I.$$
Example 2

- Consider an experiment involving 4 litters of 4 animals each.
- Suppose 4 treatments are randomly assigned to the 4 animals in each litter.
- Suppose we obtain two replicate muscle samples from each animal and measure the response of interest for each muscle sample.
Let \( y_{ijk} \) denote the \( k \)th measure of the response for the animal from litter \( j \) that received treatment \( i \) \((i = 1, 2, 3, 4; j = 1, 2, 3, 4; k = 1, 2)\)

Suppose

\[
y_{ijk} = \mu + \tau_i + \ell_j + a_{ij} + e_{ijk},
\]

where

\[
\beta = [\mu, \tau_1, \tau_2, \tau_3, \tau_4]' \in \mathbb{R}^5
\]

is an unknown vector of fixed parameters,

\[
u = [\ell_1, \ell_2, \ell_3, \ell_4, a_{11}, a_{21}, a_{31}, a_{41}, a_{12}, \ldots, a_{34}, a_{44}]'
\]

is a vector of random effects, and
\[ e = \begin{bmatrix} e_{111}, e_{112}, e_{212}, \ldots, e_{411}, e_{412}, \ldots, e_{441}, e_{442} \end{bmatrix}' \]

is a vector of random errors.

With

\[ y = \begin{bmatrix} y_{111}, y_{112}, y_{212}, \ldots, y_{411}, y_{412}, \ldots, y_{441}, y_{442} \end{bmatrix}', \]

we can write the model as a linear mixed-effects model

\[ y = X\beta + Zu + e, \]

where
The matrix above repeated three more times.
We can write less and be more precise using Kronecker product notation.

\[ X = \begin{bmatrix} 1 \\ 4 \times 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 8 \times 1 \end{bmatrix}, \quad Z = \begin{bmatrix} I \\ 4 \times 4 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 2 \times 1 \end{bmatrix}, \quad Z = \begin{bmatrix} I \\ 8 \times 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 2 \times 1 \end{bmatrix}. \]

In this experiment, we have two random factors: litter and animal.

We can partition our random effects vector \( u \) into a vector of litter effects and a vector of animal effects:
\[ u = \begin{bmatrix} \ell \\ a \end{bmatrix}, \quad \ell = \begin{bmatrix} \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \end{bmatrix}, \quad a = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \\ a_{12} \\ \vdots \\ a_{44} \end{bmatrix}, \]

We make the usual assumption that

\[ u = \begin{bmatrix} \ell \\ a \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_\ell I \\ 0 & \sigma^2_a I \end{bmatrix} \right), \]

where \( \sigma^2_\ell, \sigma^2_a \in \mathbb{R}^+ \) are unknown parameters.
We can partition

$$Z = \left[ I_{4 \times 4} \otimes 1_{8 \times 1}, I_{16 \times 16} \otimes 1_{2 \times 1} \right]$$

$$= [Z_\ell, Z_a].$$

We have

$$Zu = [Z_\ell, Z_a] \begin{bmatrix} \ell \\ a \end{bmatrix}$$

$$= Z_\ell \ell + Z_a a$$

and
\[
\text{Var}(Zu) = ZGZ' \\
= [Z_\ell, Z_a] \begin{bmatrix} \sigma_\ell^2 I & 0 \\ 0 & \sigma_a^2 I \end{bmatrix} \begin{bmatrix} Z'_\ell \\ Z'_a \end{bmatrix} \\
= Z_\ell (\sigma_\ell^2 I) Z'_\ell + Z_a (\sigma_a^2 I) Z'_a \\
= \sigma_\ell^2 Z_\ell Z'_\ell + \sigma_a^2 Z_a Z'_a \\
= \sigma_\ell^2 I_{4 \times 4} \otimes 11'_{8 \times 8} + \sigma_a^2 I_{16 \times 16} \otimes 11'_{2 \times 2}. 
\]
We usually assume that all random effects and random errors are mutually independent and that the errors (like the effects within each factor) are identically distributed:

\[
\begin{bmatrix}
\ell \\
a \\
e
\end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix}
\sigma^2_{\ell} I & 0 & 0 \\
0 & \sigma^2_a I & 0 \\
0 & 0 & \sigma^2_e I
\end{bmatrix} \right).
\]

The unknown variance parameters \( \sigma^2_{\ell}, \sigma^2_a, \sigma^2_e \in \mathbb{R}^+ \) are called variance components.
In this case, we have \( R = \text{Var}(e) = \sigma_e^2 I. \)

Thus,

\[
\text{Var}(y) = ZGZ' + R = \sigma_\ell^2 Z_\ell Z_\ell' + \sigma_a^2 Z_a Z_a' + \sigma_e^2 I.
\]

This is a block diagonal matrix with a block as follows.

(To get a block to fit on one slide, let \( \ell = \sigma_\ell^2, a = \sigma_a^2, e = \sigma_e^2 \).)
\[
\begin{array}{ccccccccc}
\ell + a + e & \ell + a & \ell & \ell & \ell & \ell & \ell & \ell & \ell \\
\ell + a & \ell + a + e & \ell & \ell & \ell & \ell & \ell & \ell & \ell \\
\ell & \ell & \ell + a + e & \ell + a & \ell & \ell & \ell & \ell & \ell \\
\ell & \ell & \ell & \ell & \ell + a + e & \ell + a & \ell & \ell & \ell \\
\ell & \ell & \ell & \ell & \ell & \ell + a & \ell + a + e & \ell & \ell \\
\ell & \ell & \ell & \ell & \ell & \ell & \ell + a + e & \ell + a & \ell + a + e \\
\end{array}
\]