11. Linear Mixed-Effects Models

The Linear Mixed-Effects Model

- $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}$
- $\boldsymbol{X}$ is an $n \times p$ matrix of known constants
- $\boldsymbol{\beta} \in \mathbb{R}^p$ is an unknown parameter vector
- $\boldsymbol{Z}$ is an $n \times q$ matrix of known constants
- $\boldsymbol{u}$ is a $q \times 1$ random vector
- $\boldsymbol{e}$ is an $n \times 1$ vector of random errors

The Linear Mixed-Effects Model

- $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}$

- The elements of $\boldsymbol{\beta}$ are considered to be non-random and are called *fixed effects*.

- The elements of $\boldsymbol{u}$ are random variables and are called *random effects*.

- The elements of the error vector $\boldsymbol{e}$ 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)

$$E(\boldsymbol{y}|\boldsymbol{u}) = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u},$$

  a linear function of fixed and random effects.

We assume that

$$E(\boldsymbol{e}) = \mathbf{0} \qquad \text{Var}(\boldsymbol{e}) = \boldsymbol{R}$$

$$E(\boldsymbol{u}) = \mathbf{0} \qquad \text{Var}(\boldsymbol{u}) = \boldsymbol{G}$$

$$\text{Cov}(\boldsymbol{e},\ \boldsymbol{u}) = \mathbf{0}.$$

It follows that

$$
\begin{aligned}
\boldsymbol{E}(\boldsymbol{y}) &= \boldsymbol{E}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}) \\
&= \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{E}(\boldsymbol{u}) + \boldsymbol{E}(\boldsymbol{e}) \\
&= \boldsymbol{X}\boldsymbol{\beta}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(\boldsymbol{y}) &= \mathrm{Var}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}) \\
&= \mathrm{Var}(\boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}) \\
&= \mathrm{Var}(\boldsymbol{Z}\boldsymbol{u}) + \mathrm{Var}(\boldsymbol{e}) \\
&= \boldsymbol{Z}\mathrm{Var}(\boldsymbol{u})\boldsymbol{Z}' + \boldsymbol{R} \\
&= \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' + \boldsymbol{R} \equiv \boldsymbol{\Sigma}.
\end{aligned}
$$

We usually consider the special case in which

$$\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{e} \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{G} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R} \end{bmatrix} \right)$$

$$\implies \boldsymbol{y} \sim N(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' + \boldsymbol{R}).$$

The conditional mean and variance, given the random effects, are

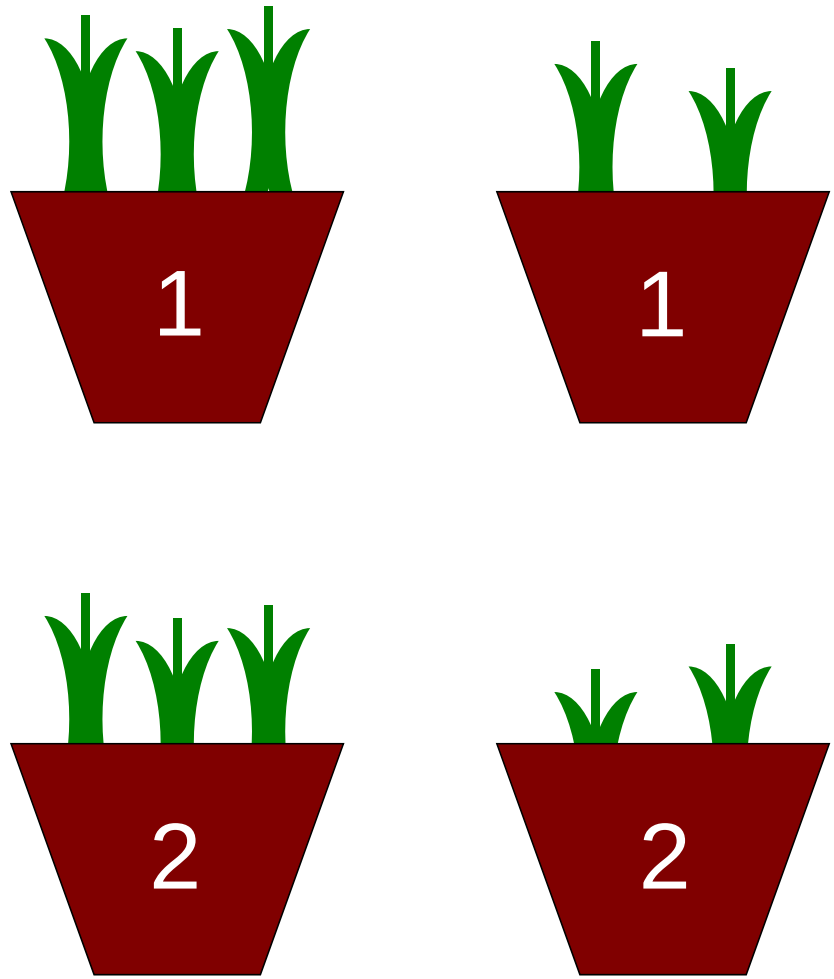
$$\boldsymbol{E}(\boldsymbol{y}|\boldsymbol{u}) = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u}$$

and

$$\mathrm{Var}(\boldsymbol{y}|\boldsymbol{u}) = \boldsymbol{R}.$$

Example 1

Suppose an experiment was conducted to compare the height of plants grown at two soil moisture levels (labeled 1 and 2). The soil moisture levels were randomly assigned to 4 pots with 2 pots per moisture level. For each moisture level, 3 seeds were planted in one pot and 2 seeds were planted in the other. After a four-week growing period, the height of each seedling was measured. Let y_{ijk} denote the height for soil moisture level i, pot j, seedling k.


1
1
2
2

Consider the model

$$y_{ijk} = \mu + \alpha_i + p_{ij} + e_{ijk}$$

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

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

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}, \text{ where}$$

$$\boldsymbol{y} = \begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{213} \\ y_{221} \\ y_{222} \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{bmatrix},$$

$$\boldsymbol{Z} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{u} = \begin{bmatrix} p_{11} \\ p_{12} \\ p_{21} \\ p_{22} \end{bmatrix}, \boldsymbol{e} = \begin{bmatrix} e_{111} \\ e_{112} \\ e_{113} \\ e_{121} \\ e_{122} \\ e_{211} \\ e_{212} \\ e_{213} \\ e_{221} \\ e_{222} \end{bmatrix}.$$

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{213} \\ y_{221} \\ y_{222} \end{bmatrix} = \begin{bmatrix} \mu + \alpha_1 \\ \mu + \alpha_1 \\ \mu + \alpha_1 \\ \mu + \alpha_1 \\ \mu + \alpha_1 \\ \mu + \alpha_2 \\ \mu + \alpha_2 \\ \mu + \alpha_2 \\ \mu + \alpha_2 \\ \mu + \alpha_2 \end{bmatrix} + \begin{bmatrix} p_{11} \\ p_{11} \\ p_{11} \\ p_{12} \\ p_{12} \\ p_{21} \\ p_{21} \\ p_{21} \\ p_{22} \\ p_{22} \end{bmatrix} + \begin{bmatrix} e_{111} \\ e_{112} \\ e_{113} \\ e_{121} \\ e_{122} \\ e_{211} \\ e_{212} \\ e_{213} \\ e_{221} \\ e_{222} \end{bmatrix}$$

$$
\begin{aligned}
y_{111} &= \mu + \alpha_1 + p_{11} + e_{111} \\
y_{112} &= \mu + \alpha_1 + p_{11} + e_{112} \\
y_{113} &= \mu + \alpha_1 + p_{11} + e_{113} \\
y_{121} &= \mu + \alpha_1 + p_{12} + e_{121} \\
y_{122} &= \mu + \alpha_1 + p_{12} + e_{122} \\
y_{211} &= \mu + \alpha_2 + p_{21} + e_{211} \\
y_{212} &= \mu + \alpha_2 + p_{21} + e_{212} \\
y_{213} &= \mu + \alpha_2 + p_{21} + e_{213} \\
y_{221} &= \mu + \alpha_2 + p_{22} + e_{221} \\
y_{222} &= \mu + \alpha_2 + p_{22} + e_{222}
\end{aligned}
$$

$$\boldsymbol{G} = \text{Var}(\boldsymbol{u}) = \text{Var}([p_{11}, p_{12}, p_{21}, p_{22}]') = \sigma_p^2 \boldsymbol{I}_{4\times 4}$$

$$\boldsymbol{R} = \text{Var}(\boldsymbol{e}) = \sigma_e^2 \boldsymbol{I}_{10\times 10}$$

$$\text{Var}(\boldsymbol{y}) = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' + \boldsymbol{R} = \boldsymbol{Z}\sigma_p^2\boldsymbol{I}\boldsymbol{Z}' + \sigma_e^2\boldsymbol{I} = \sigma_p^2\boldsymbol{Z}\boldsymbol{Z}' + \sigma_e^2\boldsymbol{I}.$$

$$ZZ' = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Thus, $\text{Var}(\boldsymbol{y}) = \sigma_p^2 \boldsymbol{Z}\boldsymbol{Z}' + \sigma_e^2 \boldsymbol{I}$ 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

$$\begin{aligned}
\text{Var}(y_{ijk}) &= \sigma_p^2 + \sigma_e^2 \qquad \forall\ i,j,k. \\
\text{Cov}(y_{ijk},\ y_{ijk^*}) &= \sigma_p^2 \qquad \forall\ i,j,\ \text{and}\ \ k \neq k^*. \\
\text{Cov}(y_{ijk},\ y_{i^*j^*k^*}) &= 0 \qquad \text{if } i \neq i^* \text{ or } j \neq j^*.
\end{aligned}$$

Any two observations from the same pot have covariance σ_p^2.

Any two observations from different pots are uncorrelated.

Alternative Derivation of Variances and Covariances

$$\begin{aligned}
\text{Var}(y_{ijk}) &= \text{Var}(\mu + \alpha_i + p_{ij} + e_{ijk}) = \text{Var}(p_{ij} + e_{ijk}) \\
&= \text{Var}(p_{ij}) + \text{Var}(e_{ijk}) + \text{Cov}(p_{ij}, e_{ijk}) + \text{Cov}(e_{ijk}, p_{ij}) \\
&= \sigma_p^2 + \sigma_e^2 + 0 + 0 = \sigma_p^2 + \sigma_e^2.
\end{aligned}$$

For $k \neq k^*$,

$$\begin{aligned}
\text{Cov}(y_{ijk},\ y_{ijk^*}) &= \text{Cov}(\mu + \alpha_i + p_{ij} + e_{ijk},\ \mu + \alpha_i + p_{ij} + e_{ijk^*}) \\
&= \text{Cov}(p_{ij} + e_{ijk},\ p_{ij} + e_{ijk^*}) \\
&= \text{Cov}(p_{ij},\ p_{ij}) + \text{Cov}(p_{ij},\ e_{ijk^*}) \\
&\quad + \text{Cov}(e_{ijk}, p_{ij}) + \text{Cov}(e_{ijk},\ e_{ijk^*}) \\
&= \text{Var}(p_{ij}) + 0 + 0 + 0 = \sigma_p^2.
\end{aligned}$$

- Note that $\mathrm{Var}(\boldsymbol{y})$ may be written as $\sigma_e^2 \boldsymbol{V}$ where $\boldsymbol{V}$ is a block diagonal matrix with blocks of the form

$$\begin{bmatrix} 1+\sigma_p^2/\sigma_e^2 & \sigma_p^2/\sigma_e^2 & . & . & . & \sigma_p^2/\sigma_e^2 \\ \sigma_p^2/\sigma_e^2 & 1+\sigma_p^2/\sigma_e^2 & . & . & . & \sigma_p^2/\sigma_e^2 \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ \sigma_p^2/\sigma_e^2 & \sigma_p^2/\sigma_e^2 & . & . & . & 1+\sigma_p^2/\sigma_e^2 \end{bmatrix}$$

- Thus, if σ_p^2/σ_e^2 were known, we would have the Aitken Model.

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \text{ where } \boldsymbol{\epsilon} = \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e} \sim N(\boldsymbol{0},\ \sigma^2 \boldsymbol{V}),\ \sigma^2 \equiv \sigma_e^2.$$

- Thus, if σ_p^2/σ_e^2 were known, we would use GLS to estimate any estimable $\boldsymbol{C\beta}$ by $\boldsymbol{C\hat{\beta}_V = C(X'V^{-1}X)^{-}X'V^{-1}y}$.
- However, we seldom know σ_p^2/σ_e^2 or, more generally, $\boldsymbol{\Sigma}$ or $\boldsymbol{V}$.
- For the general problem where $\mathrm{Var}(\boldsymbol{y}) = \boldsymbol{\Sigma}$ is an unknown positive definite matrix, we can rewrite $\boldsymbol{\Sigma}$ as $\sigma^2\boldsymbol{V}$, where σ^2 is an unknown positive variance and $\boldsymbol{V}$ is an unknown positive definite matrix.
- As in our simple example, each entry of $\boldsymbol{V}$ is usually assumed to be a known function of few unknown parameters.

- Thus, our strategy for estimating an estimable $\boldsymbol{C}\boldsymbol{\beta}$ involves estimating the unknown parameters in $\boldsymbol{V}$ to obtain

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

- In general,

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

is an nonlinear estimator that is an approximation to

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

which would be the BLUE of $\boldsymbol{C}\boldsymbol{\beta}$ if $\boldsymbol{V}$ were known.

- In special cases, $\boldsymbol{C}\hat{\boldsymbol{\beta}}_{\hat{\boldsymbol{V}}}$ may be a linear estimator.

- However, even for our simple example involving seedling height, $\boldsymbol{C}\hat{\boldsymbol{\beta}}_{\hat{\boldsymbol{V}}}$ is a nonlinear estimator of $\boldsymbol{C}\boldsymbol{\beta}$ for

$$\begin{aligned} \boldsymbol{C} = [1, 1, 0] &\iff \boldsymbol{C}\boldsymbol{\beta} = \mu + \alpha_1, \\ \boldsymbol{C} = [1, 0, 1] &\iff \boldsymbol{C}\boldsymbol{\beta} = \mu + \alpha_2, \text{ and} \\ \boldsymbol{C} = [0, 1, -1] &\iff \boldsymbol{C}\boldsymbol{\beta} = \alpha_1 - \alpha_2. \end{aligned}$$

- Confidence intervals and tests for these estimable functions are not exact.

- In our simple example involving seedling height, there was only one random factor (pot).
- When there are m random factors, we can partition $\boldsymbol{Z}$ and $\boldsymbol{u}$ as

$$\boldsymbol{Z} = [\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_m] \text{ and } \boldsymbol{u} = \begin{bmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_m \end{bmatrix},$$

where $\boldsymbol{u}_j$ is the vector of random effects associated with factor j ($j = 1, \ldots, m$).

- We can write $\boldsymbol{Zu}$ as

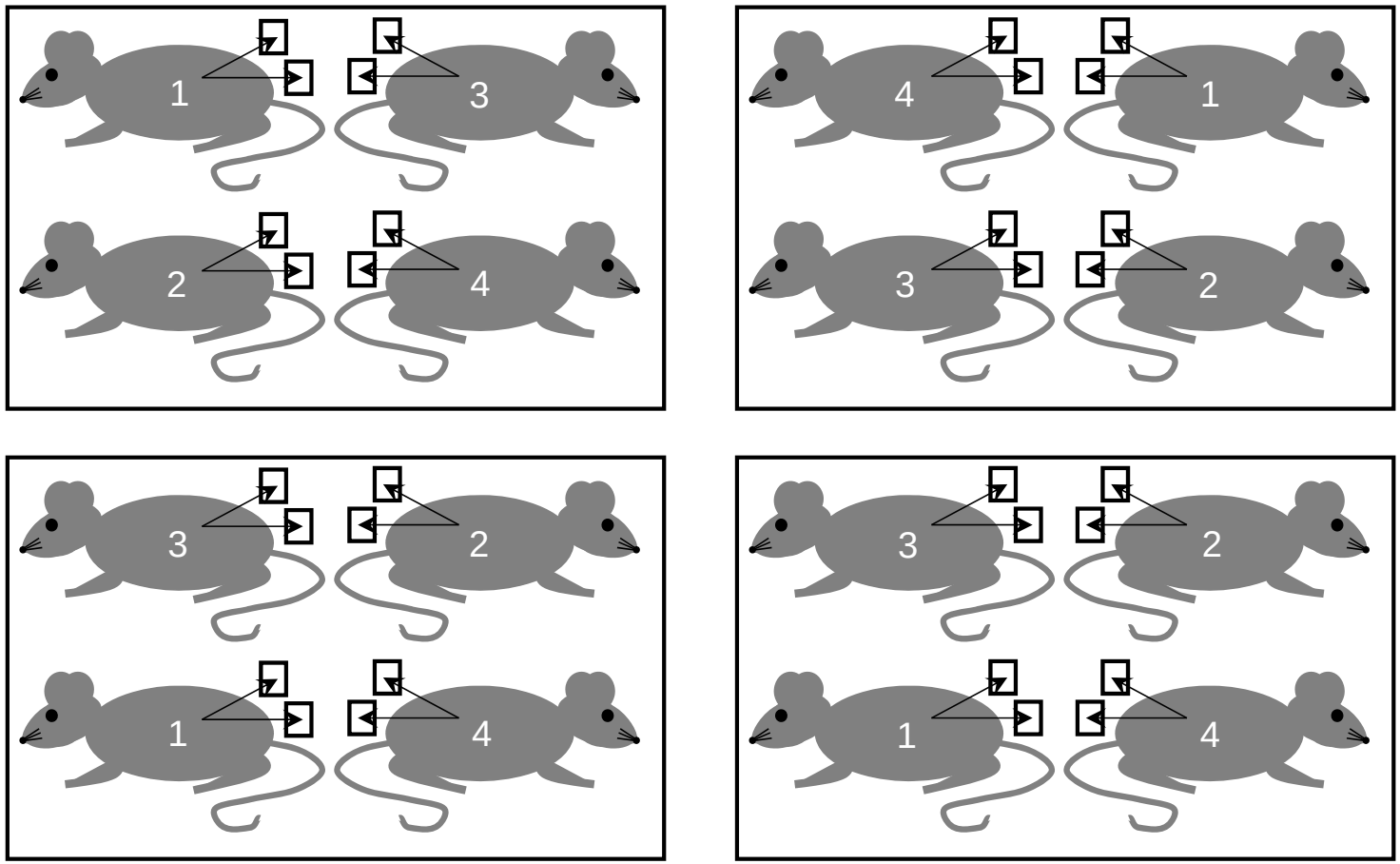
$$[\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_m] \begin{bmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_m \end{bmatrix} = \sum_{j=1}^{m} \boldsymbol{Z}_j \boldsymbol{u}_j.$$

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

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

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.


1
3
2
4
4
1
3
2
3
2
1
4
3
2
1
4

Let y_{ijk} denote the kth 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 $\boldsymbol{\beta} = [\mu, \tau_1, \tau_2, \tau_3, \tau_4]' \in \mathbb{R}^5$ is an unknown vector of fixed parameters,

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

is a vector of random effects, and

$$\boldsymbol{e} = [e_{111}, e_{112}, e_{211}, e_{212}, e_{311}, e_{312}, e_{411}, e_{412}, \ldots, e_{441}, e_{442}]'$$

is a vector of random errors.

With

$$\boldsymbol{y} = [y_{111}, y_{112}, y_{211}, y_{212}, y_{311}, y_{312}, y_{411}, y_{412}, \ldots, y_{441}, y_{442}]',$$

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

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e},$$

where

$$X = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ & & \text{The matrix} & & \\ & & \text{above repeated} & & \\ & & \text{three more} & & \\ & & \text{times.} & & \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Kronecker Product Notation

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \otimes \boldsymbol{B}$$

$$= \begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \cdots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{bmatrix}$$

We can write less and be more precise using Kronecker product notation.

$$\underset{4\times 1}{\boldsymbol{X}} = \underset{4\times 1}{\mathbf{1}} \otimes [\underset{8\times 1}{\mathbf{1}}, \underset{4\times 4}{\boldsymbol{I}} \otimes \underset{2\times 1}{\mathbf{1}}], \quad \underset{4\times 16}{\boldsymbol{Z}} = [\underset{4\times 4}{\boldsymbol{I}} \otimes \underset{8\times 1}{\mathbf{1}}, \underset{16\times 16}{\boldsymbol{I}} \otimes \underset{2\times 1}{\mathbf{1}}].$$

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

We can partition our random effects vector $\boldsymbol{u}$ into a vector of litter effects and a vector of animal effects:

$$\boldsymbol{u} = \begin{bmatrix} \boldsymbol{\ell} \\ \boldsymbol{a} \end{bmatrix}, \quad \boldsymbol{\ell} = \begin{bmatrix} \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \end{bmatrix}, \quad \boldsymbol{a} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \\ a_{12} \\ \vdots \\ a_{44} \end{bmatrix}.$$

We make the usual assumption that

$$\boldsymbol{u} = \begin{bmatrix} \boldsymbol{\ell} \\ \boldsymbol{a} \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_\ell^2 \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \sigma_a^2 \boldsymbol{I} \end{bmatrix} \right),$$

where $\sigma_\ell^2, \sigma_a^2 \in \mathbb{R}^+$ are unknown parameters.

We can partition

$$\begin{aligned} \boldsymbol{Z} &= \left[\underset{4\times 4}{\boldsymbol{I}} \otimes \underset{8\times 1}{\mathbf{1}}, \underset{16\times 16}{\boldsymbol{I}} \otimes \underset{2\times 1}{\mathbf{1}} \right] \\ &= [\boldsymbol{Z}_\ell, \boldsymbol{Z}_a]. \end{aligned}$$

We have

$$\begin{aligned} \boldsymbol{Zu} &= [\boldsymbol{Z}_\ell, \boldsymbol{Z}_a] \begin{bmatrix} \boldsymbol{\ell} \\ \boldsymbol{a} \end{bmatrix} \\ &= \boldsymbol{Z}_\ell \boldsymbol{\ell} + \boldsymbol{Z}_a \boldsymbol{a} \end{aligned}$$

and

$$
\begin{aligned}
\text{Var}(\boldsymbol{Zu}) &= \boldsymbol{ZGZ}' \\
&= [\boldsymbol{Z}_\ell, \boldsymbol{Z}_a] \begin{bmatrix} \sigma^2_\ell \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma^2_a \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{Z}'_\ell \\ \boldsymbol{Z}'_a \end{bmatrix} \\
&= \boldsymbol{Z}_\ell(\sigma^2_\ell \boldsymbol{I})\boldsymbol{Z}'_\ell + \boldsymbol{Z}_a(\sigma^2_a \boldsymbol{I})\boldsymbol{Z}'_a \\
&= \sigma^2_\ell \boldsymbol{Z}_\ell \boldsymbol{Z}'_\ell + \sigma^2_a \boldsymbol{Z}_a \boldsymbol{Z}'_a \\
&= \sigma^2_\ell \underset{4\times 4}{\boldsymbol{I}} \otimes \underset{8\times 8}{\boldsymbol{11}'} + \sigma^2_a \underset{16\times 16}{\boldsymbol{I}} \otimes \underset{2\times 2}{\boldsymbol{11}'}.
\end{aligned}
$$

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:

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

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 $\boldsymbol{R} = \text{Var}(\boldsymbol{e}) = \sigma_e^2 \boldsymbol{I}$.

Thus,

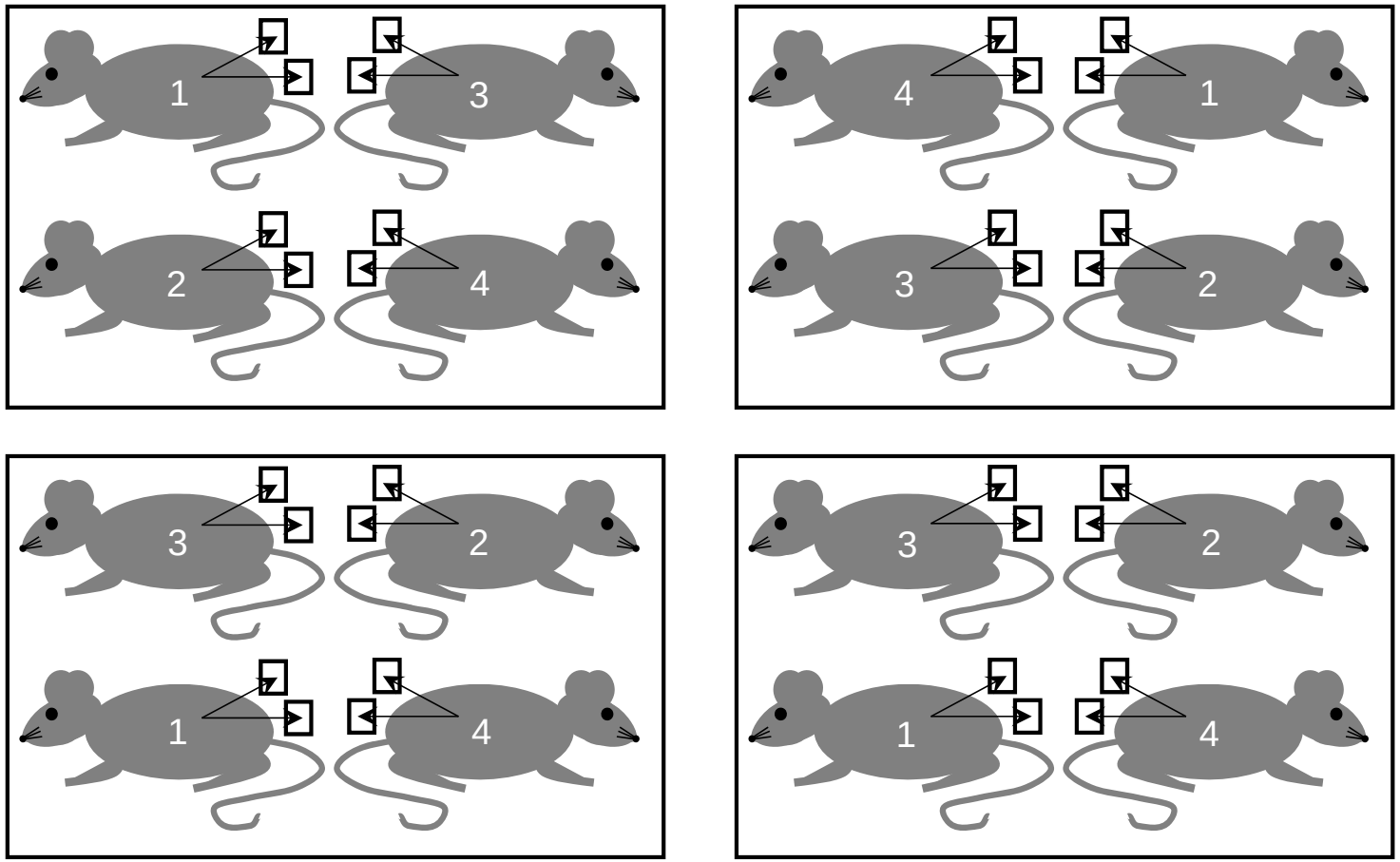
$$\begin{aligned} \text{Var}(\boldsymbol{y}) &= \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' + \boldsymbol{R} \\ &= \sigma_\ell^2 \boldsymbol{Z}_\ell \boldsymbol{Z}_\ell' + \sigma_a^2 \boldsymbol{Z}_a \boldsymbol{Z}_a' + \sigma_e^2 \boldsymbol{I}. \end{aligned}$$

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

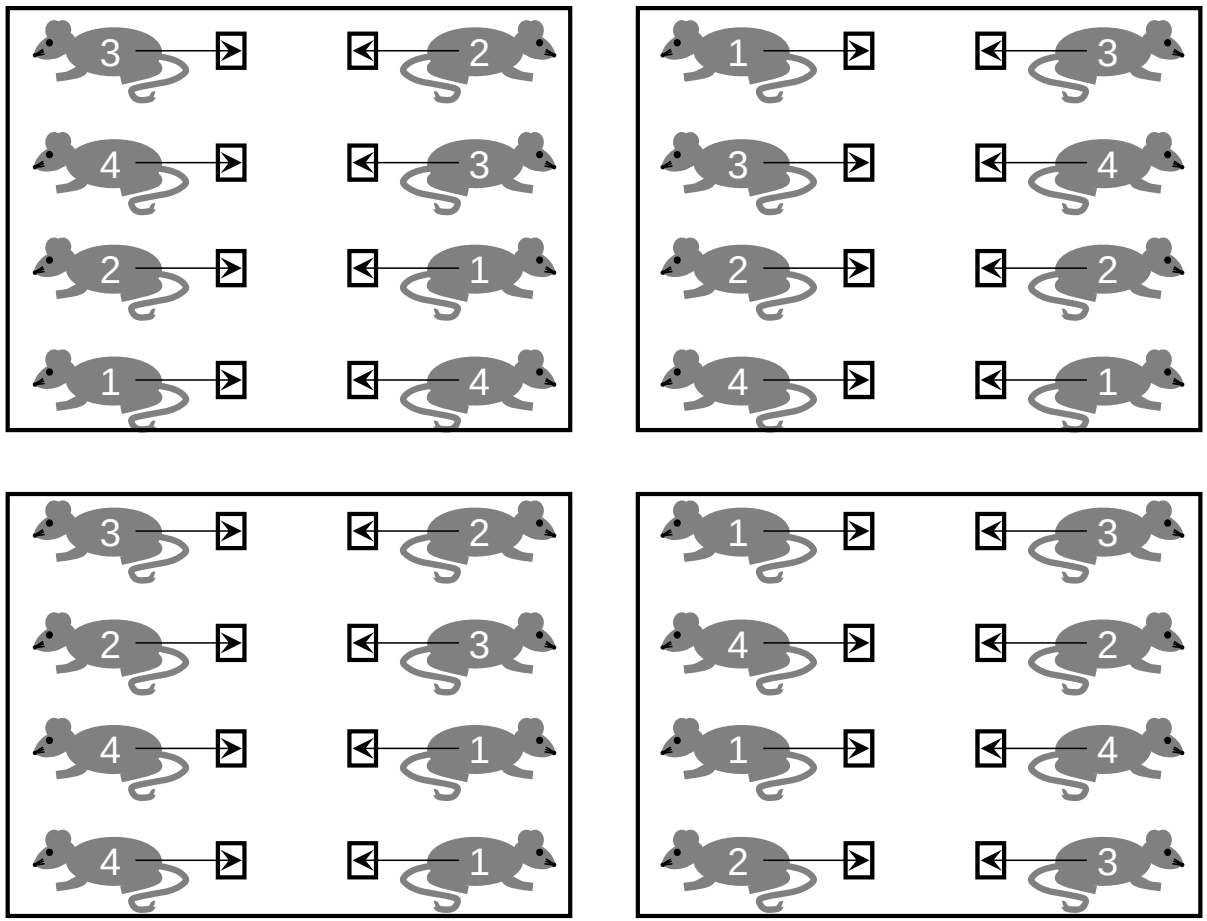
(To get a block to fit on one slide, let $\ell = \sigma_\ell^2, a = \sigma_a^2, e = \sigma_e^2$).

$$
\begin{array}{cccccccc}
\ell+a+e & \ell+a & \ell & \ell & \ell & \ell & \ell & \ell \\
\ell+a & \ell+a+e & \ell & \ell & \ell & \ell & \ell & \ell \\
\ell & \ell & \ell+a+e & \ell+a & \ell & \ell & \ell & \ell \\
\ell & \ell & \ell+a & \ell+a+e & \ell & \ell & \ell & \ell \\
\ell & \ell & \ell & \ell & \ell+a+e & \ell+a & \ell & \ell \\
\ell & \ell & \ell & \ell & \ell+a & \ell+a+e & \ell & \ell \\
\ell & \ell & \ell & \ell & \ell & \ell & \ell+a+e & \ell+a \\
\ell & \ell & \ell & \ell & \ell & \ell & \ell+a & \ell+a+e
\end{array}
$$

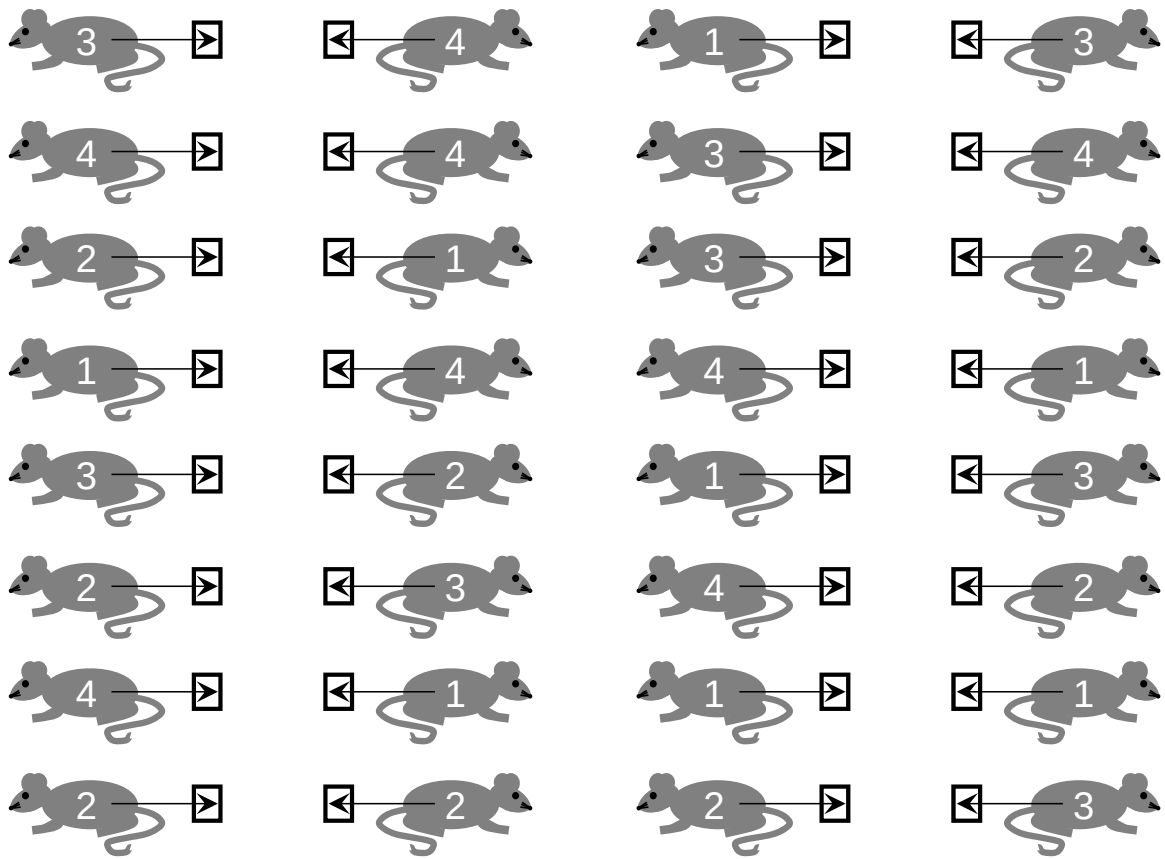
Random Effects Specify the Correlation Structure



Without the mouse random effects, our model would correspond to an RCBD with 2 mice per treatment per litter.



With no random effects, our model would correspond to a CRD with 8 mice per treatment.



Review of Experimental Design Terminology

Experiment – An investigation in which the investigator applies some treatments to experimental units and then observes the effect of the treatments on the experimental units by measuring one or more response variables.

Treatment – a condition or set of conditions applied to experimental units in an experiment.

Experimental Unit – the physical entity to which a treatment is randomly assigned and independently applied.

Response Variable – a characteristic of an experimental unit that is measured after treatment and analyzed to assess the effects of treatments on experimental units.

Observational Unit – the unit on which a response variable is measured.

There is often a one-to-one correspondence between experimental units and observational units, but that is not always true.

In our example involving plant heights and soil moisture levels, pots were the experimental units because soil moisture levels were randomly assigned to pots.

Seedlings were the observational units because the response was measured separately for each seedling.

Whenever there is more than one observational unit for an experimental unit or whenever the response is measured multiple times for an experimental unit, we say we have *multiple observations per experiment unit*.

This scenario is also referred to as *subsampling* or *pseudo-replication*.

Whenever an experiment involves multiple observations per experimental unit, it is important to include a random effect for each experimental unit.

Without a random effect for each experimental unit, a one-to-one correspondence between observations and experimental units is assumed.

Including random effects in a model is one way to account for a lack of independence among observations that might be expected based on the design of an experiment.

Completely Randomized Design (CRD) – experimental design in which, for given number of experiment units per treatment, all possible assignments of treatments to experimental units are equally likely.

Block – a group of experimental units that, prior to treatment, are expected to be more like one another (with respect to one or more response variables) than experimental units in general.

Randomized Complete Block Design (RCBD) – experimental design in which separate and completely randomized treatment assignments are made for each of multiple blocks in such a way that all treatments have at least one experimental unit in each block.

Blocking – grouping similar experimental units together and assigning different treatments within such groups of experimental units.

The experiment involving muscle samples from mice used blocking.

Each litter was a block in that experiment.

Each mouse was an experimental unit.

Each muscle sample was an observational unit.