# Linear Mixed-Effect Models

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + Z\underset{\sim}{u} + \underset{\sim}{e}$$

$X$ $n \times p$ matrix of known constants

$\underset{\sim}{\beta} \in \mathbb{R}^p$ an unknown parameter vector

$Z$ $n \times q$ matrix of known constants

$\underset{\sim}{u}$ $q \times 1$ random vector

$\underset{\sim}{e}$ $n \times 1$ vector of random errors

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

The elements of $\underset{\sim}{u}$ are called "random effects."

Because the model includes both fixed and random effects, 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) $E(\underset{\sim}{y}|\underline{u}) = X\underline{\beta} + Z\underline{u}$, a linear function of fixed and random effects.

We assume that

$$E(\underset{\sim}{e}) = \underset{\sim}{0} \qquad Var(\underset{\sim}{e}) = R$$

$$E(\underset{\sim}{u}) = \underset{\sim}{0} \qquad Var(\underset{\sim}{u}) = G$$

$$Cov(\underset{\sim}{e}, \underset{\sim}{u}) = 0.$$

It follows that

$$E(\underset{\sim}{y}) = E(X\underset{\sim}{\beta} + Z\underline{u} + \underset{\sim}{e})$$

$$= X\underset{\sim}{\beta} + Z\,E(\underset{\sim}{u}) + E(\underset{\sim}{e})$$

$$= X\underset{\sim}{\beta} \qquad \text{and}$$

$$Var(\underset{\sim}{y}) = Var(X\underset{\sim}{\beta} + Z\underset{\sim}{u} + \underset{\sim}{e})$$

$$= Var(Z\underline{u} + \underset{\sim}{e})$$

$$= Var(Z\underset{\sim}{u}) + Var(\underset{\sim}{e})$$

$$= Z\,Var(\underline{u})\,Z' + R$$

$$= ZGZ' + R \equiv \Sigma.$$

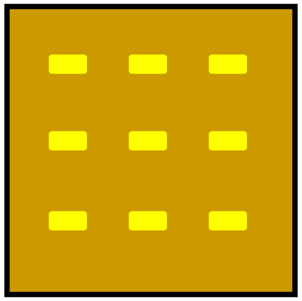
We usually consider the special case in which

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

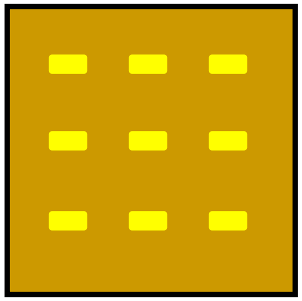
$$\Longrightarrow \quad \underset{\sim}{y} \sim N(X\underset{\sim}{\beta}, ZGZ' + R).$$

Example: Recall the seedling dry weight data. Suppose a weight was available for each individual seedling. Let $y_{ijk}$ denote the weight of the $k^{th}$ seedling from the $j^{th}$ tray of genotype $i$ ($i = 1, 2$; $j = 1, 2, 3, 4$; $k = 1, \ldots, n_{ij}$; where $n_{ij}$ = # of seedlings for genotype $i$ tray $j$).

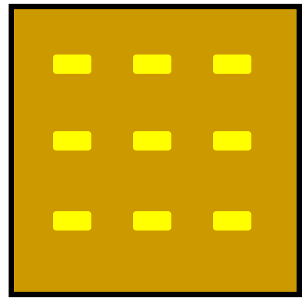
A



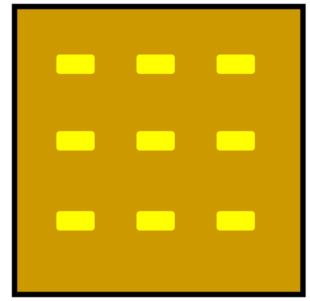
A



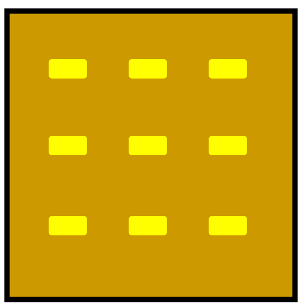
B



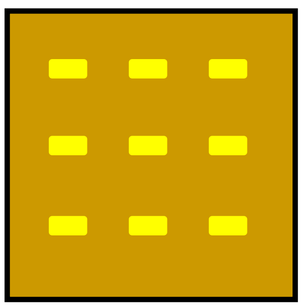
B



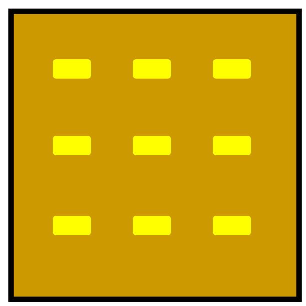
B



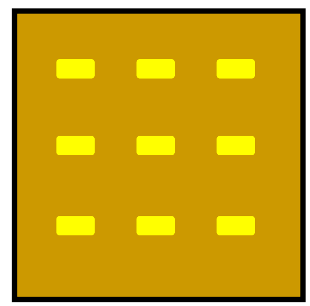
A



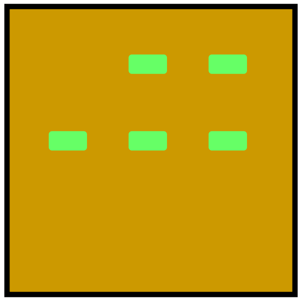
B



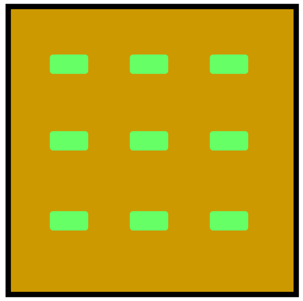
A



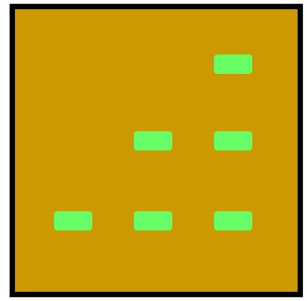
A



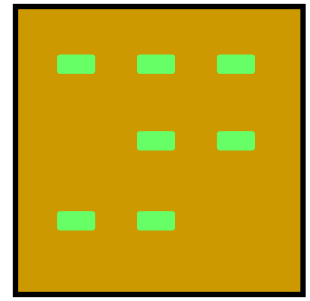
A



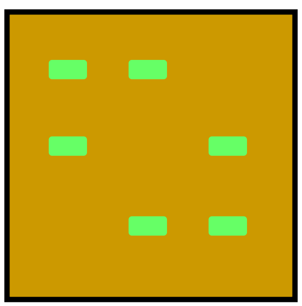
B



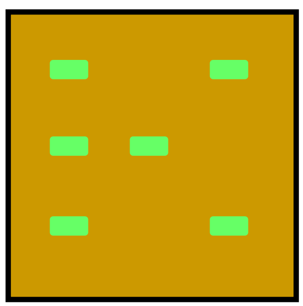
B



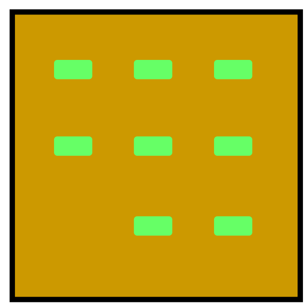
B



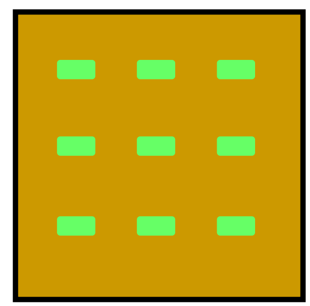
A



B



A



```
d=read.delim(
"http://www.public.iastate.edu/~dnett/S511/SeedlingDryWeight2.txt"
)
d
```

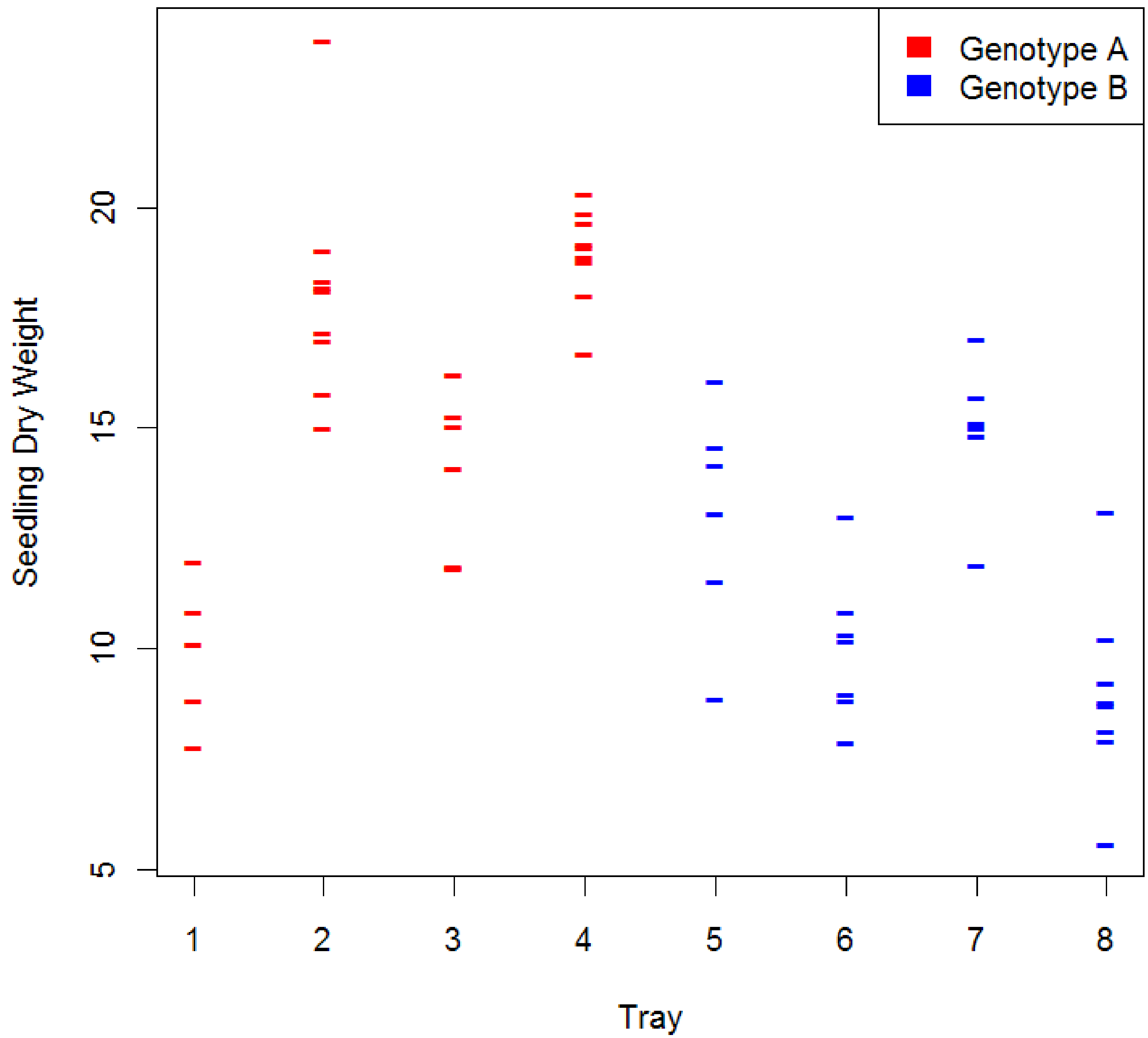
```
   Genotype Tray Seedling SeedlingWeight
1         A    1        1              8
2         A    1        2              9
3         A    1        3             11
4         A    1        4             12
5         A    1        5             10
6         A    2        1             17
7         A    2        2             17
8         A    2        3             16
9         A    2        4             15
10        A    2        5             19
11        A    2        6             18
12        A    2        7             18
13        A    2        8             18
14        A    2        9             24
15        A    3        1             12
16        A    3        2             12
17        A    3        3             16
```

| | | | | |
|---|---|---|---|---|
| 18 | A | 3 | 4 | 15 |
| 19 | A | 3 | 5 | 15 |
| 20 | A | 3 | 6 | 14 |
| 21 | A | 4 | 1 | 17 |
| 22 | A | 4 | 2 | 20 |
| 23 | A | 4 | 3 | 20 |
| 24 | A | 4 | 4 | 19 |
| 25 | A | 4 | 5 | 19 |
| 26 | A | 4 | 6 | 18 |
| 27 | A | 4 | 7 | 20 |
| 28 | A | 4 | 8 | 19 |
| 29 | A | 4 | 9 | 19 |
| 30 | B | 5 | 1 | 9 |
| 31 | B | 5 | 2 | 12 |
| 32 | B | 5 | 3 | 13 |
| 33 | B | 5 | 4 | 16 |
| 34 | B | 5 | 5 | 14 |
| 35 | B | 5 | 6 | 14 |
| 36 | B | 6 | 1 | 10 |
| 37 | B | 6 | 2 | 10 |
| 38 | B | 6 | 3 | 9 |

| | | | | |
|---|---|---|---|---|
| 39 | B | 6 | 4 | 8 |
| 40 | B | 6 | 5 | 13 |
| 41 | B | 6 | 6 | 9 |
| 42 | B | 6 | 7 | 11 |
| 43 | B | 7 | 1 | 12 |
| 44 | B | 7 | 2 | 16 |
| 45 | B | 7 | 3 | 17 |
| 46 | B | 7 | 4 | 15 |
| 47 | B | 7 | 5 | 15 |
| 48 | B | 7 | 6 | 15 |
| 49 | B | 8 | 1 | 9 |
| 50 | B | 8 | 2 | 6 |
| 51 | B | 8 | 3 | 8 |
| 52 | B | 8 | 4 | 8 |
| 53 | B | 8 | 5 | 13 |
| 54 | B | 8 | 6 | 9 |
| 55 | B | 8 | 7 | 9 |
| 56 | B | 8 | 8 | 10 |

```
plot(d[,2],d[,4]+rnorm(56,0,.2),
    xlab="Tray",ylab="Seedling Dry Weight",
    col=2*(1+(d[,1]=="B")),pch="-",cex=2)


legend("topright",c("Genotype A","Genotype B"),
       fill=c(2,4),border=c(2,4))
```


Genotype A
Genotype B
Seedling Dry Weight
5
10
15
20
1
2
3
4
5
6
7
8
Tray

Consider the model

$$y_{ijk} = \mu + \gamma_i + T_{ij} + e_{ijk}$$

$T_{ij}$'s $\overset{iid}{\sim} N(0, \sigma_T^2)$ independent of

$e_{ijk}$'s $\overset{iid}{\sim} N(0, \sigma_e^2)$.

This model can be written in the form $\underset{\sim}{y} = X\underset{\sim}{\beta} + Z\underset{\sim}{u} + \underset{\sim}{e}$.

$$\underset{\sim}{y} = \begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{114} \\ y_{115} \\ y_{121} \\ y_{122} \\ \vdots \\ y_{247} \\ y_{248} \end{bmatrix} \quad \underset{\sim}{\beta} = \begin{bmatrix} \mu \\ \gamma_1 \\ \gamma_2 \end{bmatrix} \quad \underset{\sim}{u} = \begin{bmatrix} T_{11} \\ T_{12} \\ T_{13} \\ T_{14} \\ T_{21} \\ T_{22} \\ T_{23} \\ T_{24} \end{bmatrix} \quad \underset{\sim}{e} = \begin{bmatrix} e_{111} \\ e_{112} \\ e_{113} \\ e_{114} \\ e_{115} \\ e_{121} \\ e_{122} \\ \vdots \\ e_{247} \\ e_{248} \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \end{bmatrix} \qquad Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & \vdots & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

```
X=cbind(1,d[,1]=="A",d[,1]=="B")
X
      [,1] [,2] [,3]
 [1,]    1    1    0
 [2,]    1    1    0
 [3,]    1    1    0
 [4,]    1    1    0
 [5,]    1    1    0
 [6,]    1    1    0
 [7,]    1    1    0
 [8,]    1    1    0
 [9,]    1    1    0
[10,]    1    1    0
[11,]    1    1    0
[12,]    1    1    0
[13,]    1    1    0
[14,]    1    1    0
[15,]    1    1    0
[16,]    1    1    0
[17,]    1    1    0
[18,]    1    1    0
```

```
[19,]    1    1    0
[20,]    1    1    0
[21,]    1    1    0
[22,]    1    1    0
[23,]    1    1    0
[24,]    1    1    0
[25,]    1    1    0
[26,]    1    1    0
[27,]    1    1    0
[28,]    1    1    0
[29,]    1    1    0
[30,]    1    0    1
[31,]    1    0    1
[32,]    1    0    1
[33,]    1    0    1
[34,]    1    0    1
[35,]    1    0    1
[36,]    1    0    1
[37,]    1    0    1
[38,]    1    0    1
[39,]    1    0    1
```

```
[40,]    1    0    1
[41,]    1    0    1
[42,]    1    0    1
[43,]    1    0    1
[44,]    1    0    1
[45,]    1    0    1
[46,]    1    0    1
[47,]    1    0    1
[48,]    1    0    1
[49,]    1    0    1
[50,]    1    0    1
[51,]    1    0    1
[52,]    1    0    1
[53,]    1    0    1
[54,]    1    0    1
[55,]    1    0    1
[56,]    1    0    1
```

```
Z=matrix(model.matrix(~0+as.factor(d[,2])),ncol=8)
Z
```

| | [,1] | [,2] | [,3] | [,4] | [,5] | [,6] | [,7] | [,8] |
|---|---|---|---|---|---|---|---|---|
| [1,] | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| [2,] | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| [3,] | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| [4,] | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| [5,] | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| [6,] | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| [7,] | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| [8,] | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| [9,] | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| [10,] | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| [11,] | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| [12,] | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| [13,] | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| [14,] | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| [15,] | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| [16,] | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| [17,] | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| [18,] | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |

```
[19,]    0    0    1    0    0    0    0    0
[20,]    0    0    1    0    0    0    0    0
[21,]    0    0    0    1    0    0    0    0
[22,]    0    0    0    1    0    0    0    0
[23,]    0    0    0    1    0    0    0    0
[24,]    0    0    0    1    0    0    0    0
[25,]    0    0    0    1    0    0    0    0
[26,]    0    0    0    1    0    0    0    0
[27,]    0    0    0    1    0    0    0    0
[28,]    0    0    0    1    0    0    0    0
[29,]    0    0    0    1    0    0    0    0
[30,]    0    0    0    0    1    0    0    0
[31,]    0    0    0    0    1    0    0    0
[32,]    0    0    0    0    1    0    0    0
[33,]    0    0    0    0    1    0    0    0
[34,]    0    0    0    0    1    0    0    0
[35,]    0    0    0    0    1    0    0    0
[36,]    0    0    0    0    0    1    0    0
[37,]    0    0    0    0    0    1    0    0
[38,]    0    0    0    0    0    1    0    0
[39,]    0    0    0    0    0    1    0    0
```

```
[40,]    0    0    0    0    0    1    0    0
[41,]    0    0    0    0    0    1    0    0
[42,]    0    0    0    0    0    1    0    0
[43,]    0    0    0    0    0    0    1    0
[44,]    0    0    0    0    0    0    1    0
[45,]    0    0    0    0    0    0    1    0
[46,]    0    0    0    0    0    0    1    0
[47,]    0    0    0    0    0    0    1    0
[48,]    0    0    0    0    0    0    1    0
[49,]    0    0    0    0    0    0    0    1
[50,]    0    0    0    0    0    0    0    1
[51,]    0    0    0    0    0    0    0    1
[52,]    0    0    0    0    0    0    0    1
[53,]    0    0    0    0    0    0    0    1
[54,]    0    0    0    0    0    0    0    1
[55,]    0    0    0    0    0    0    0    1
[56,]    0    0    0    0    0    0    0    1
```

$$G = Var(\underline{u}) = Var([T_{11}, \ldots, T_{24}]')$$

$$= \sigma_T^2 I_{8\times 8}$$

$$R = Var(\underset{\sim}{e}) = \sigma_e^2 I_{56\times 56}$$

$$Var(\underset{\sim}{y}) = ZGZ' + R = Z\sigma_T^2 I Z' + \sigma_e^2 I$$

$$= \sigma_T^2 ZZ' + \sigma_e^2 I.$$

$$
ZZ' = \begin{bmatrix} \underset{\sim}{1}\underset{\sim}{1}' & 0 & 0 & \cdots & 0 \\ 0 & \underset{\sim}{1}\underset{\sim}{1}' & 0 & \cdots & 0 \\ 0 & 0 & \underset{\sim}{1}\underset{\sim}{1}' & & \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & & & \underset{\sim}{1}\underset{\sim}{1}' \end{bmatrix}
$$

This is a block diagonal matrix with blocks of size $n_{ij} \times n_{ij}$ $i = 1, 2$, $j = 1, 2, 3, 4$.

# (Z%*%t(Z))[1:14,1:14]

```
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14]
 [1,]    1    1    1    1    1    0    0    0    0     0     0     0     0     0
 [2,]    1    1    1    1    1    0    0    0    0     0     0     0     0     0
 [3,]    1    1    1    1    1    0    0    0    0     0     0     0     0     0
 [4,]    1    1    1    1    1    0    0    0    0     0     0     0     0     0
 [5,]    1    1    1    1    1    0    0    0    0     0     0     0     0     0
 [6,]    0    0    0    0    0    1    1    1    1     1     1     1     1     1
 [7,]    0    0    0    0    0    1    1    1    1     1     1     1     1     1
 [8,]    0    0    0    0    0    1    1    1    1     1     1     1     1     1
 [9,]    0    0    0    0    0    1    1    1    1     1     1     1     1     1
[10,]    0    0    0    0    0    1    1    1    1     1     1     1     1     1
[11,]    0    0    0    0    0    1    1    1    1     1     1     1     1     1
[12,]    0    0    0    0    0    1    1    1    1     1     1     1     1     1
[13,]    0    0    0    0    0    1    1    1    1     1     1     1     1     1
[14,]    0    0    0    0    0    1    1    1    1     1     1     1     1     1
```

Thus, $Var(\underset{\sim}{y}) = \sigma_T^2 ZZ' + \sigma_e^2 I$ is also block diagonal. The first block is

$$Var\left(\begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{114} \\ y_{115} \end{bmatrix}\right) = \begin{bmatrix} \sigma_T^2+\sigma_e^2 & \sigma_T^2 & \sigma_T^2 & \sigma_T^2 & \sigma_T^2 \\ \sigma_T^2 & \sigma_T^2+\sigma_e^2 & \sigma_T^2 & \sigma_T^2 & \sigma_T^2 \\ \sigma_T^2 & \sigma_T^2 & \sigma_T^2+\sigma_e^2 & \sigma_T^2 & \sigma_T^2 \\ \sigma_T^2 & \sigma_T^2 & \sigma_T^2 & \sigma_T^2+\sigma_e^2 & \sigma_T^2 \\ \sigma_T^2 & \sigma_T^2 & \sigma_T^2 & \sigma_T^2 & \sigma_T^2+\sigma_e^2 \end{bmatrix}.$$

Other blocks are the same except that the dimensions are $(\#\text{ of seedlings per tray})^2$.

$$Var(y_{ijk}) = \sigma_T^2 + \sigma_e^2 \quad \forall\, i,j,k.$$

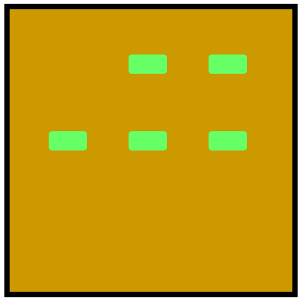
$$Cov(y_{ijk}, y_{ijl}) = \sigma_T^2 \quad \forall\, i,j, \text{ and } k \neq l.$$

$$Cov(y_{ijk}, y_{i'j'k'}) = 0 \quad \text{if } i \neq i' \text{ or } j \neq j'.$$

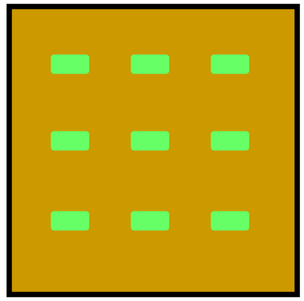
Any two observations from the same tray have covariance $\sigma_T^2$.

Any two observations from different trays are uncorrelated.

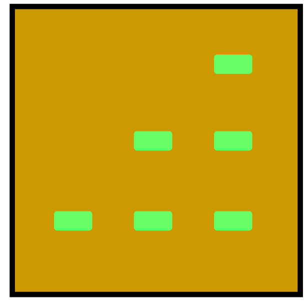
A



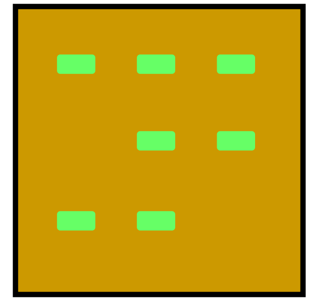
A



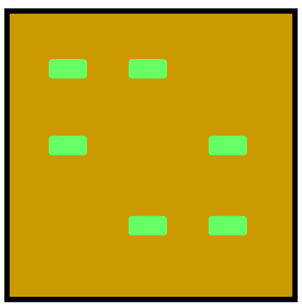
B



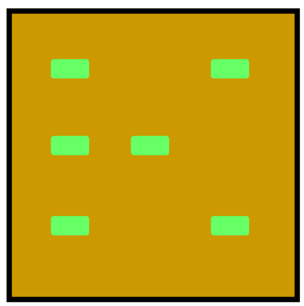
B



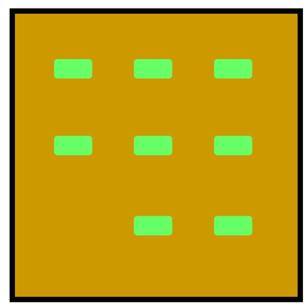
B



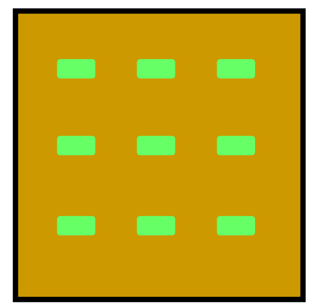
A



B



A



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

$$\begin{bmatrix} 1+\sigma_T^2/\sigma_e^2 & \sigma_T^2/\sigma_e^2 & \cdots & \sigma_T^2/\sigma_e^2 \\ \sigma_T^2/\sigma_e^2 & 1+\sigma_T^2/\sigma_e^2 & & \sigma_T^2/\sigma_e^2 \\ \vdots & & \ddots & \\ \sigma_T^2/\sigma_e^2 & \sigma_T^2/\sigma_e^2 & \cdots & 1+\sigma_T^2/\sigma_e^2 \end{bmatrix}.$$

Thus, if $\sigma_\tau^2/\sigma_e^2$ were known, we would have the Aitken model.

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underbrace{Z\underset{\sim}{u} + \underset{\sim}{e}}$$

$$= X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}, \quad \underset{\sim}{\varepsilon} \sim N(\underset{\sim}{0}, \sigma^2 V)$$

$$\sigma^2 \equiv \sigma_e^2.$$

We could use GLS to estimate estimable $C\underset{\sim}{\beta}$ by

$$C\hat{\underset{\sim}{\beta}}_V = C(X'V^{-1}X)^- X'V^{-1}\underset{\sim}{y}.$$

However, we seldom know $\sigma_T^2/\sigma_e^2$ or, more generally, $V$. Thus, our strategy usually involves estimating the unknown parameters in $V$ to obtain $\hat{V}$. Then inference proceeds based on $C\hat{\boldsymbol{\beta}}_{\hat{V}} = C(X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}\boldsymbol{y}$.

# Remaining Questions:

1. How do we estimate the unknown parameters in $V$ ?

2. What is the distribution of

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

3. How should we conduct inference regarding $C\boldsymbol{\beta}$ ?

Note that our previous analysis of the seedling dry weight tray means is not valid unless $\sigma_T^2 = 0$.

We assumed $Var(\bar{y}_{ij\cdot}) = \sigma^2 / n_{ij} \quad \forall\, i,j.$

However, under our mixed effects model,

$$Var(\bar{y}_{ij\cdot}) = Var\left(\frac{1}{n_{ij}} \underset{\sim}{1}'(y_{ij1}, \ldots, y_{ijn_{ij}})'\right)$$

$$= \frac{1}{n_{ij}^2} \underset{\sim}{1}'(\sigma_T^2 \underset{\sim}{1}\underset{\sim}{1}' + \sigma_e^2 I)\underset{\sim}{1}$$

$$= \frac{1}{n_{ij}^2}\left(\sigma_T^2 \, \underset{\sim}{1}' \underset{\sim}{1} \, \underset{\sim}{1}' \underset{\sim}{1} + \sigma_e^2 \, \underset{\sim}{1}' I \underset{\sim}{1}\right)$$

$$= \frac{1}{n_{ij}^2}\left(\sigma_T^2 \, n_{ij} \, n_{ij} + \sigma_e^2 \, n_{ij}\right)$$

$$= \sigma_T^2 + \sigma_e^2 / n_{ij} \qquad \forall \; i, j.$$