

Statistics with Spa ows II

Many models, matrices, and magic

Julia Schroeder

Julia.schroeder@imperial.ac.uk

Learning aims

- Understand mixed models
- Know when and why to use LMMs
- Know when to use random and fixed effects
- Know how to interpret random and fixed effects

Hand Out files: <https://bit.ly/2tdCTeb>

Linear models

$$y_i = b_0 + b_1 x_i + \varepsilon_i$$

- Find solution: these parameter estimates (scalars) that minimise the left-over error residuals (vector)

Linear mixed models have 2 parts

- Fixed part
- Random part

Linear mixed models have 2 parts

- Fixed part
- How mean change with predictors
- Random part
- Partitions variances between groups

Linear mixed models have 2 parts

- Fixed part
 - How mean change with predictors
 - Mean
- Random part
 - Partitions variances between groups
 - Variance

Linear mixed models have 2 parts

- Fixed part
- How mean change with predictors
- Random part
- Partitions variances between groups

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

- Mean
- Variance

Linear mixed models have 2 parts

- Fixed part
- How mean change with predictors
- Random part
- Partitions variances between groups

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

• Mean

• Variance

Linear mixed models have 2 parts

- Fixed part
 - How mean change with predictors
- Random part
 - Partitions variances between groups

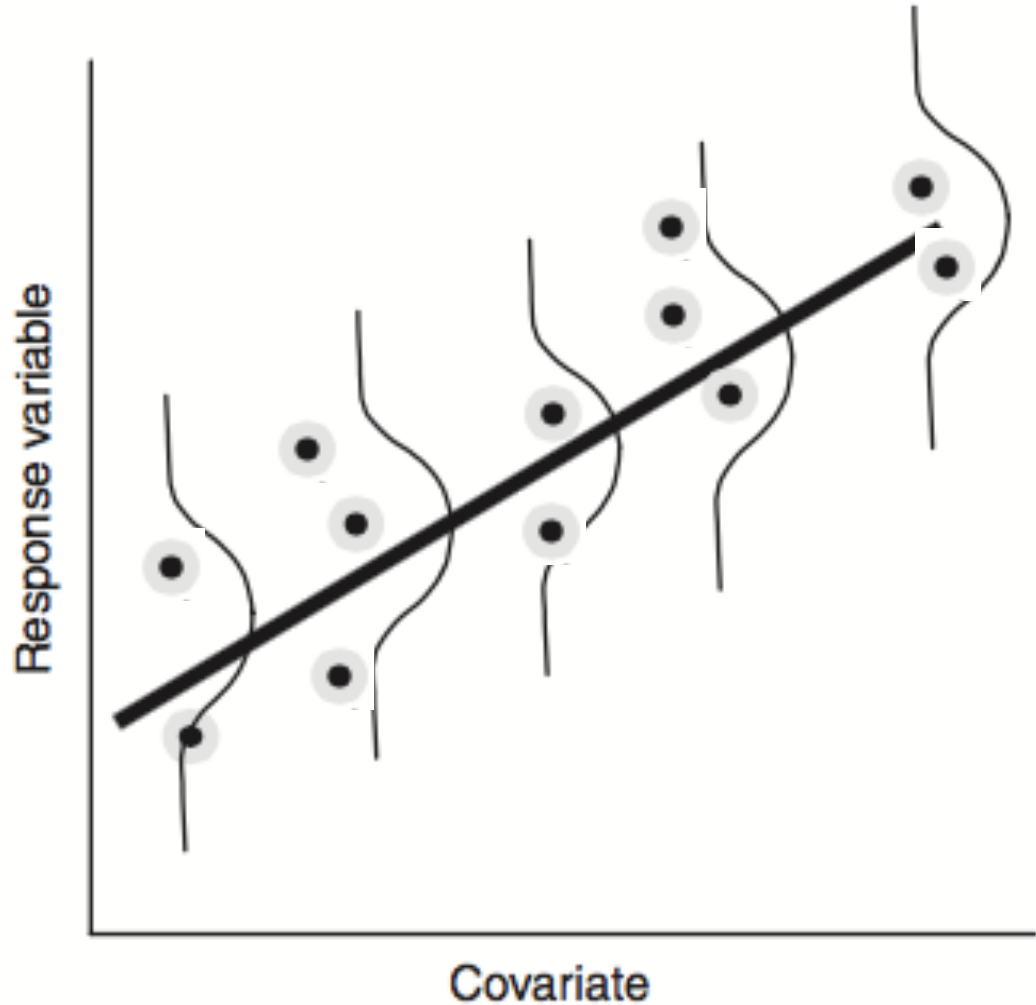
$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

• Mean

• Variance

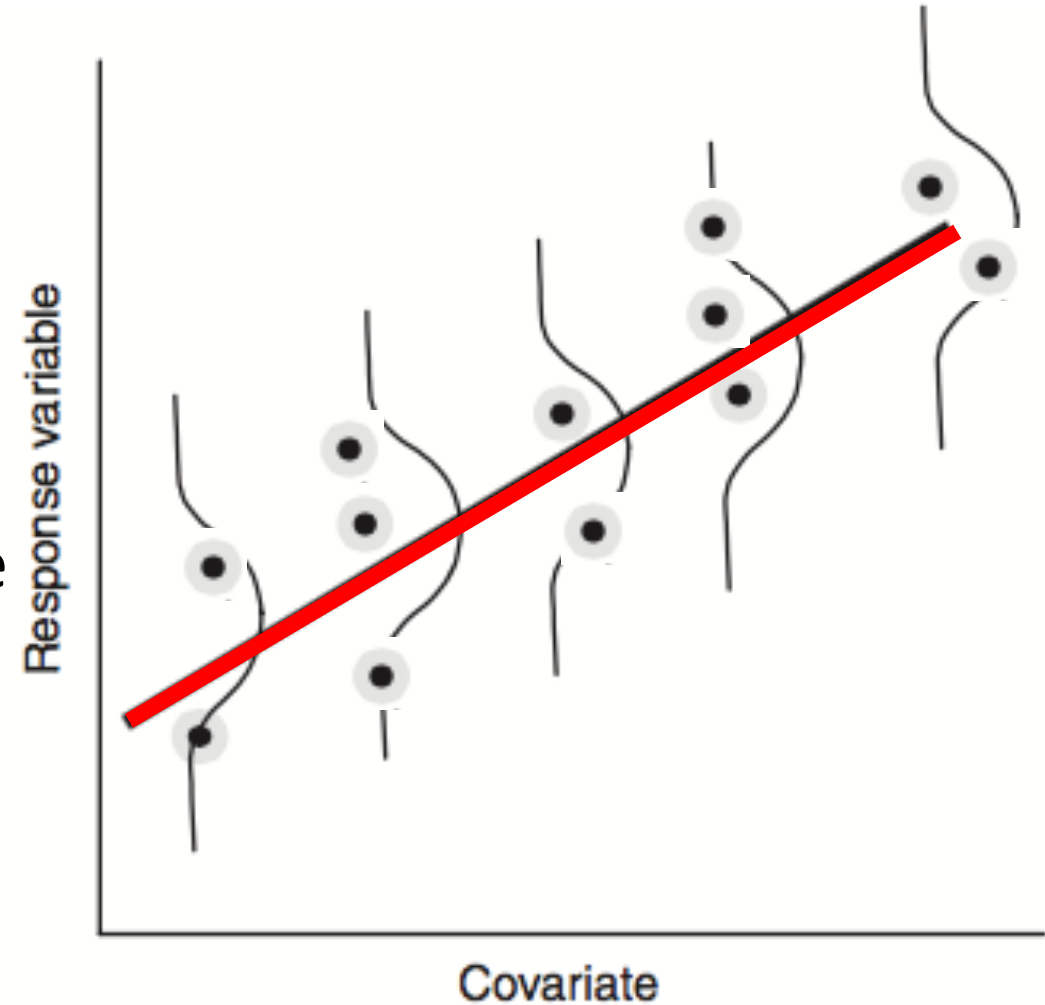
LMMs

- Estimate variance components and fixed parameter estimates simultaneously



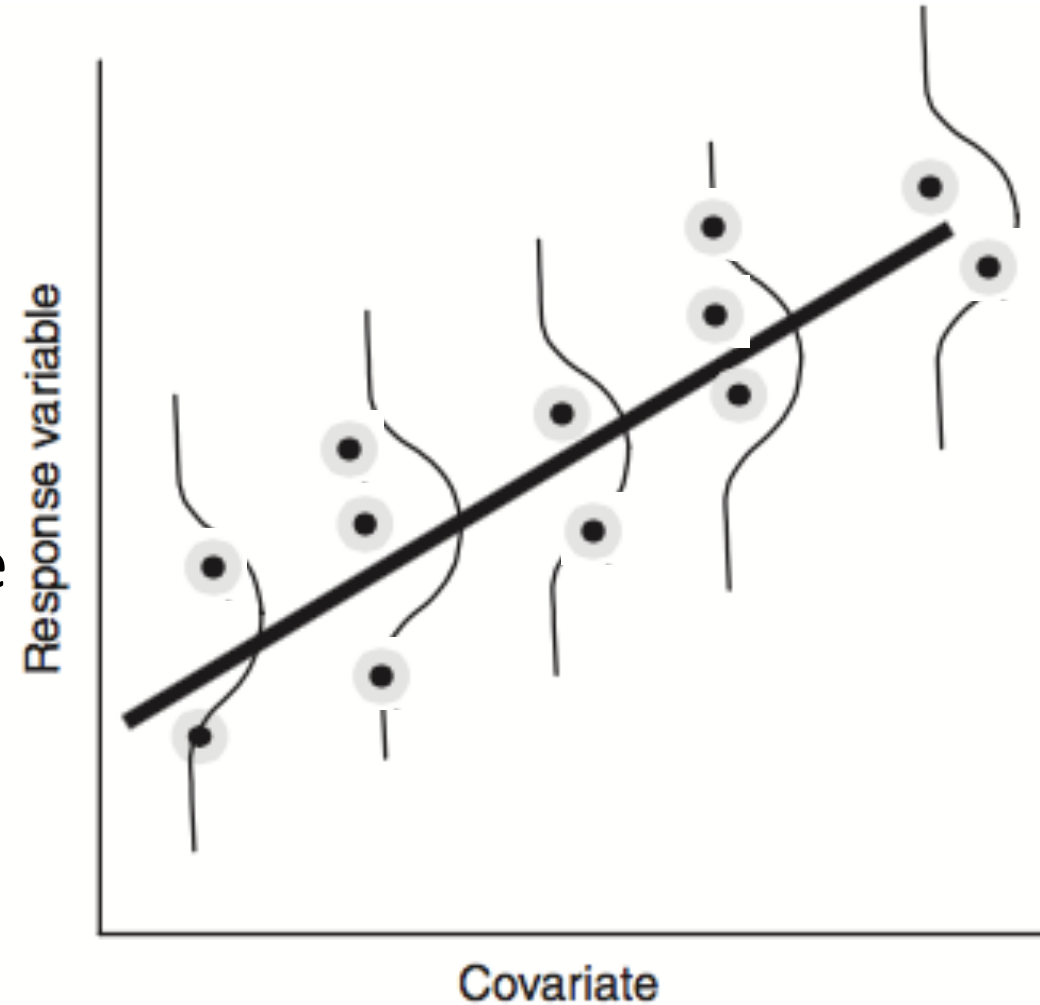
LMMs

- Estimate variance components and fixed parameter estimates simultaneously
- Fixed part: description of red line



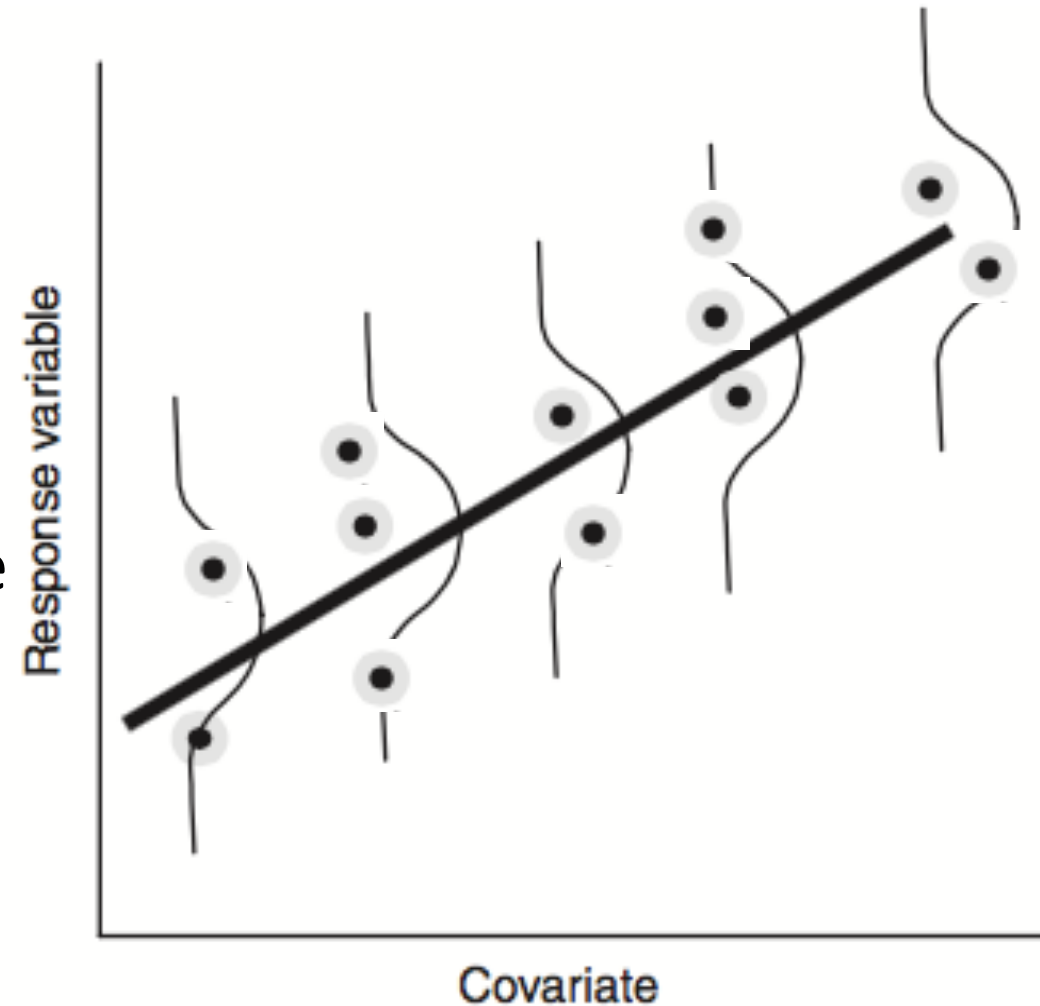
LMMs

- Estimate variance components and fixed parameter estimates simultaneously
- Fixed part: description of red line
- Random part: variance within, and between groups



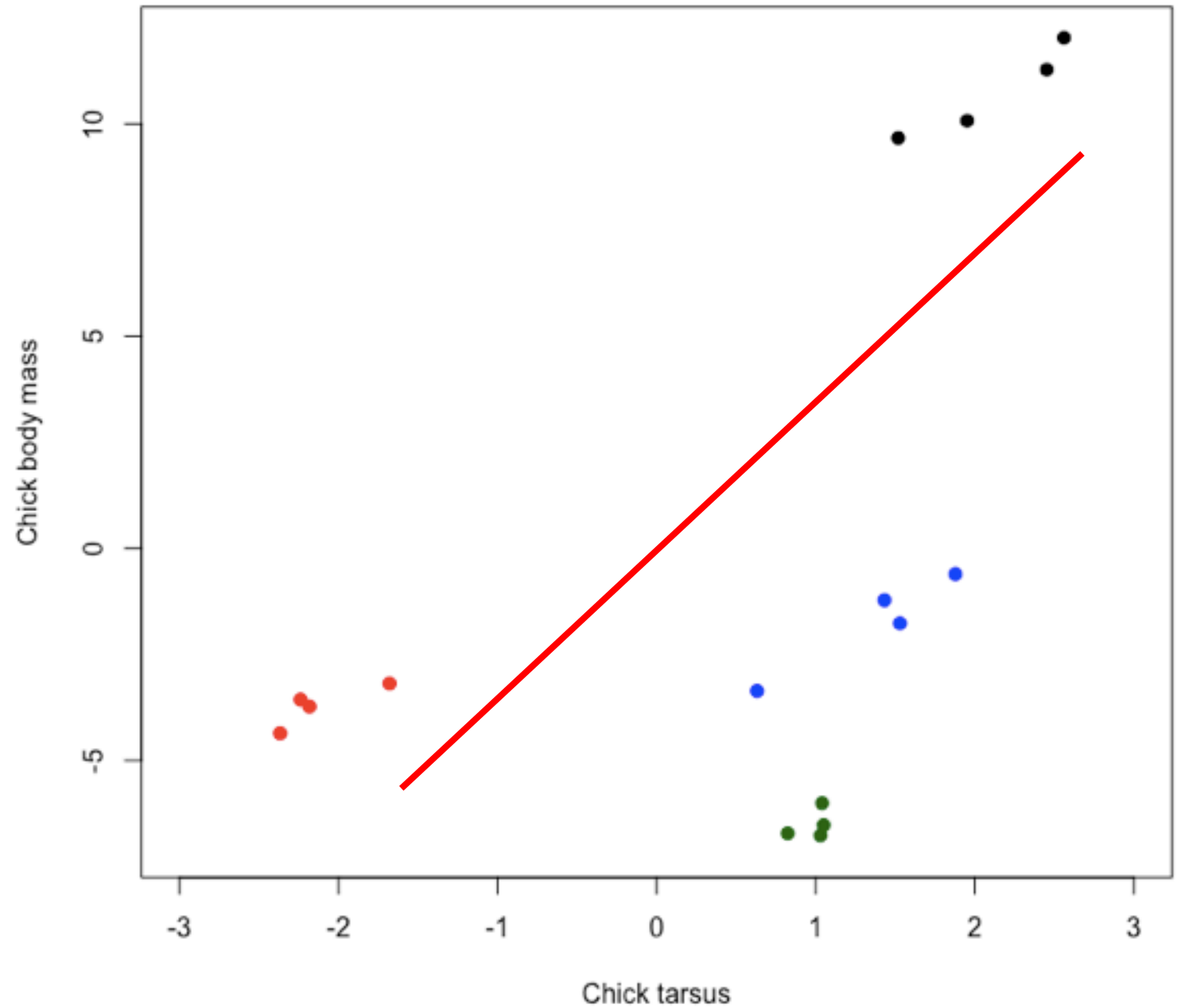
LMMs

- Estimate variance components and fixed parameter estimates simultaneously
- Fixed part: description of red line
- Random part: variance within, and between groups



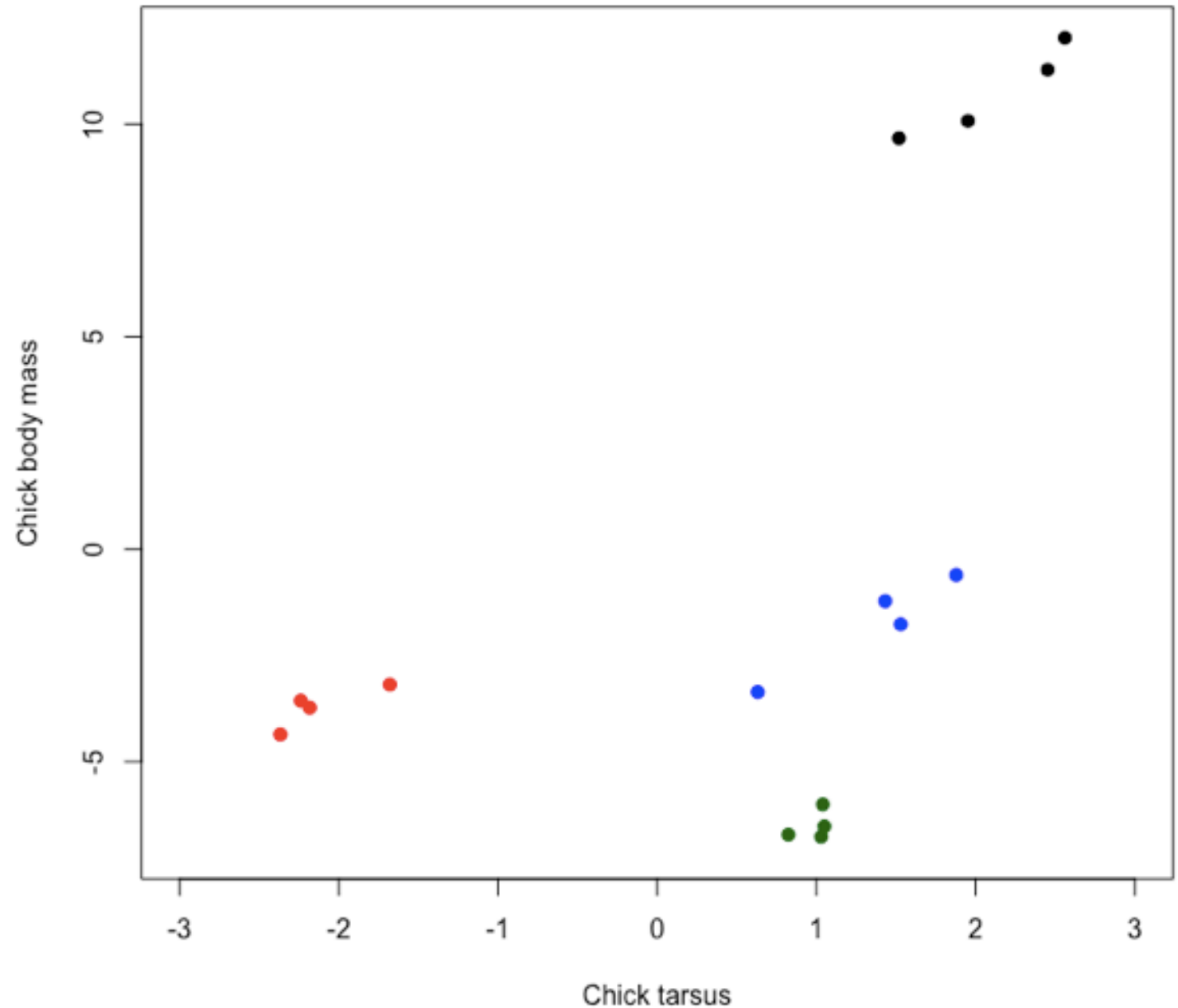
Mixed models

Fixed part



Mixed models

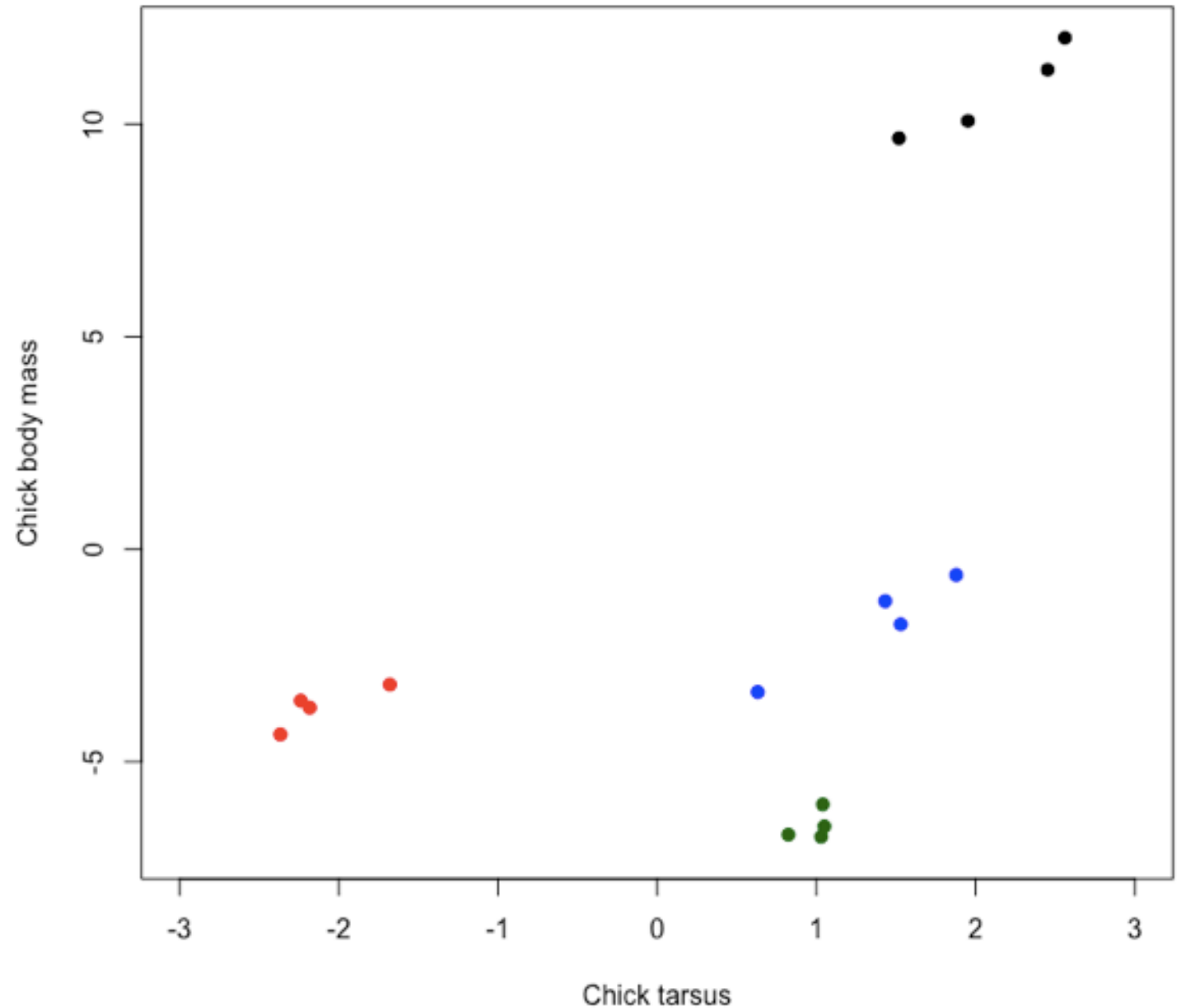
Random part
Groups?



Mixed models

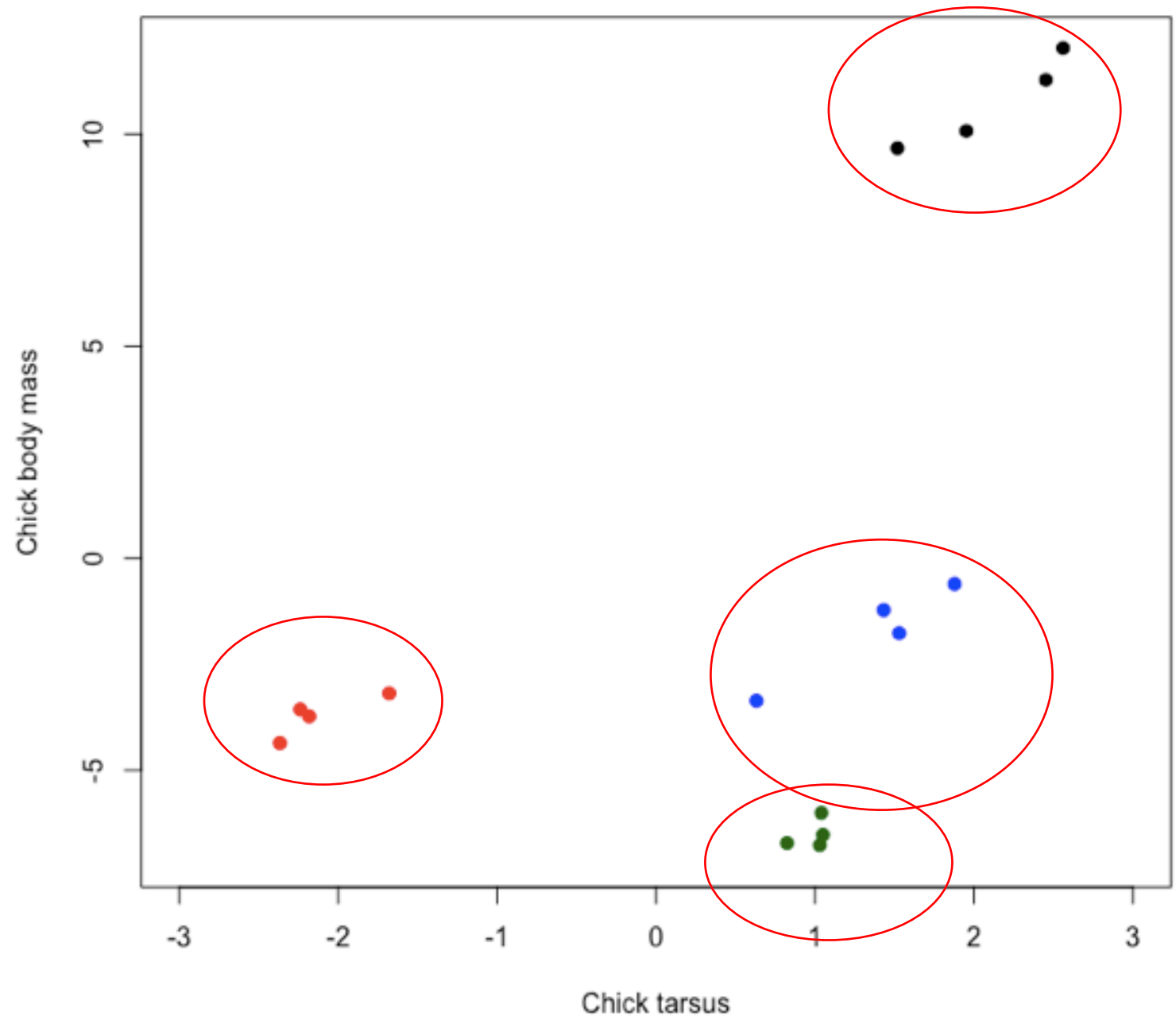
Random part

Groups = broods!



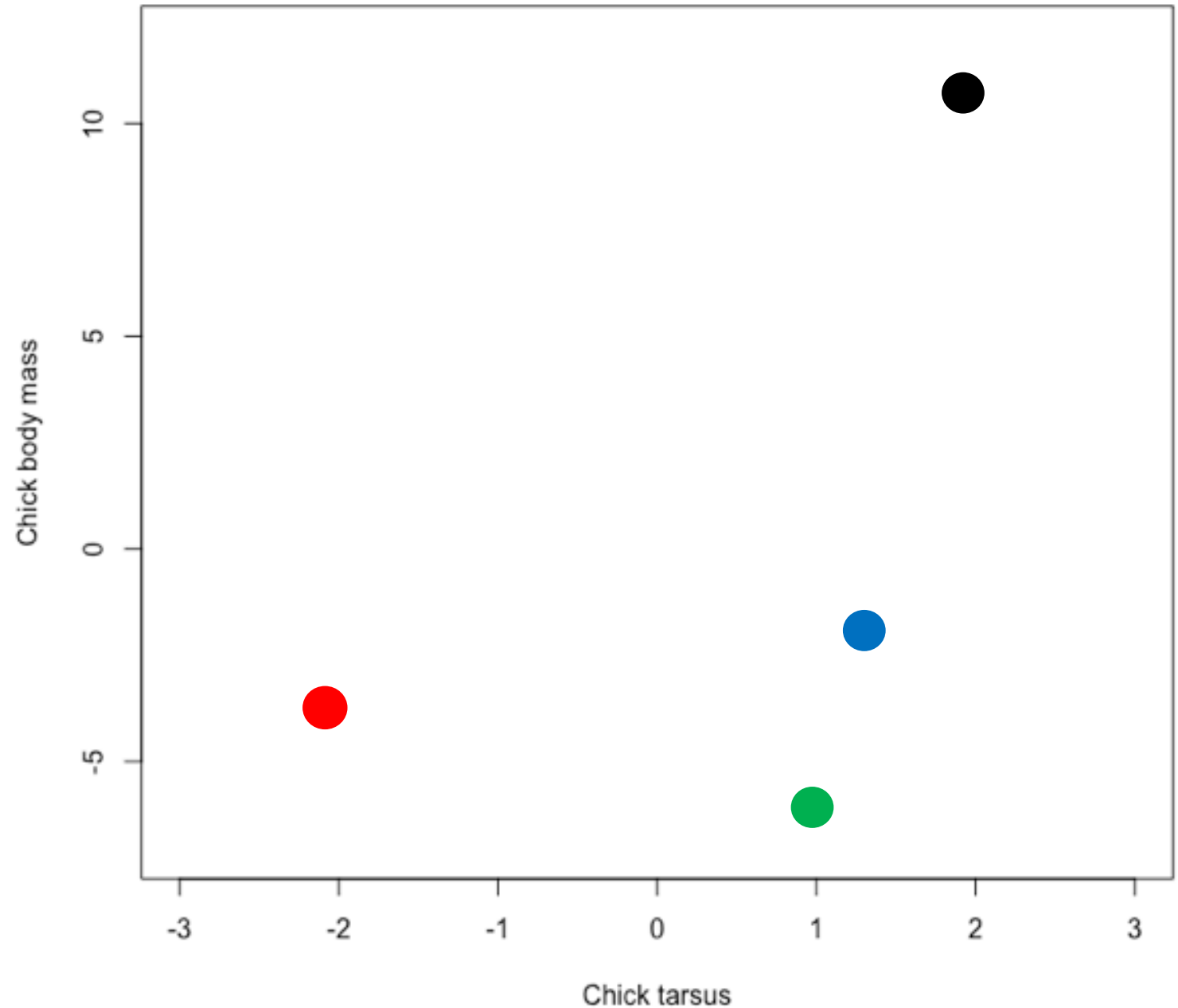
Mixed models

Estimate variances
between groups



Mixed models

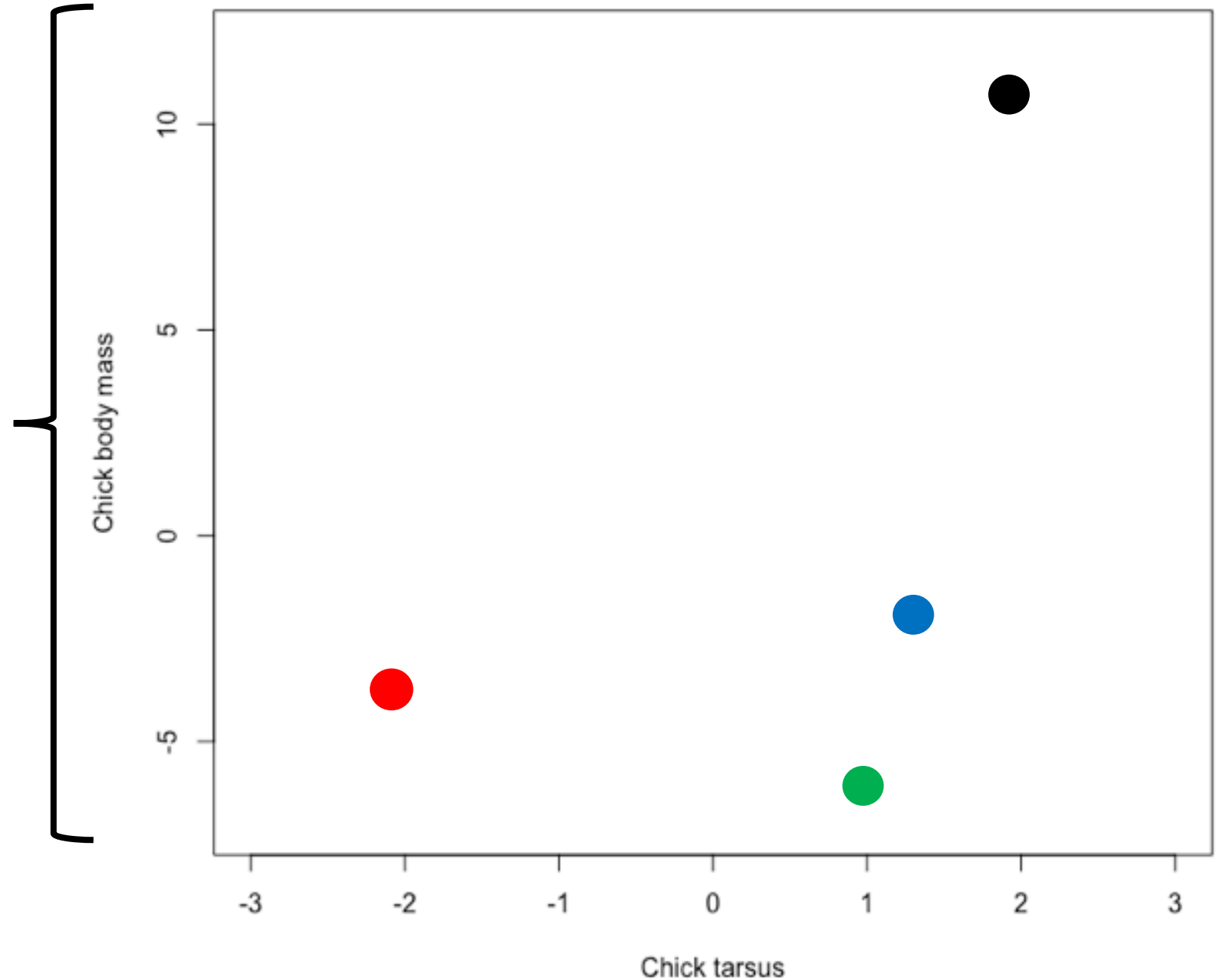
Estimate variances
between groups



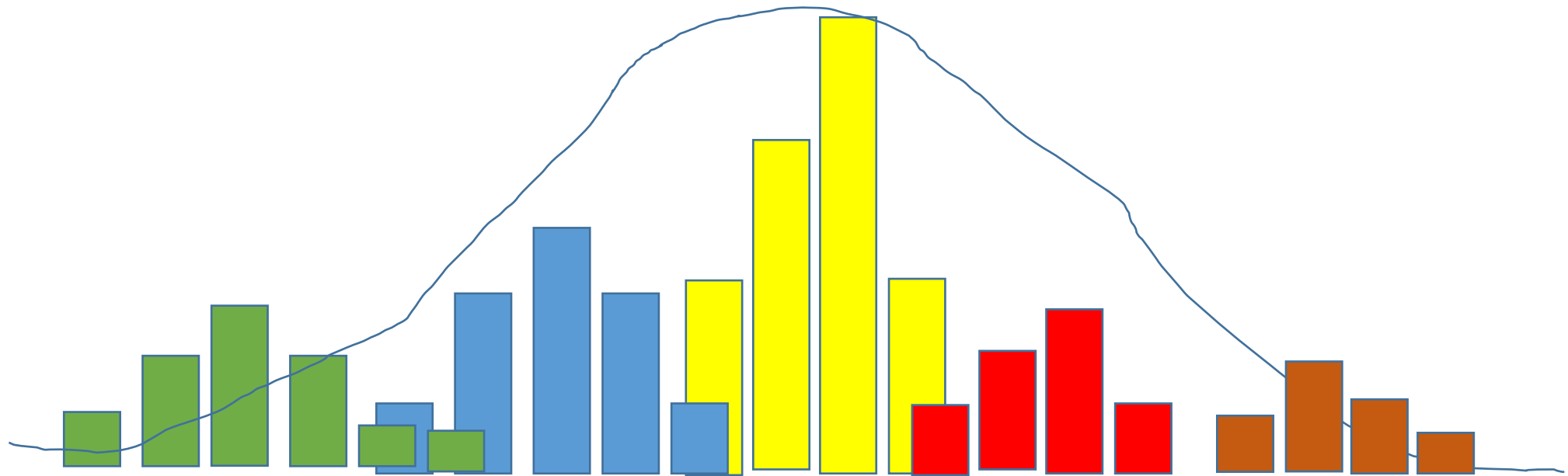
Mixed models

Estimate variances
between groups

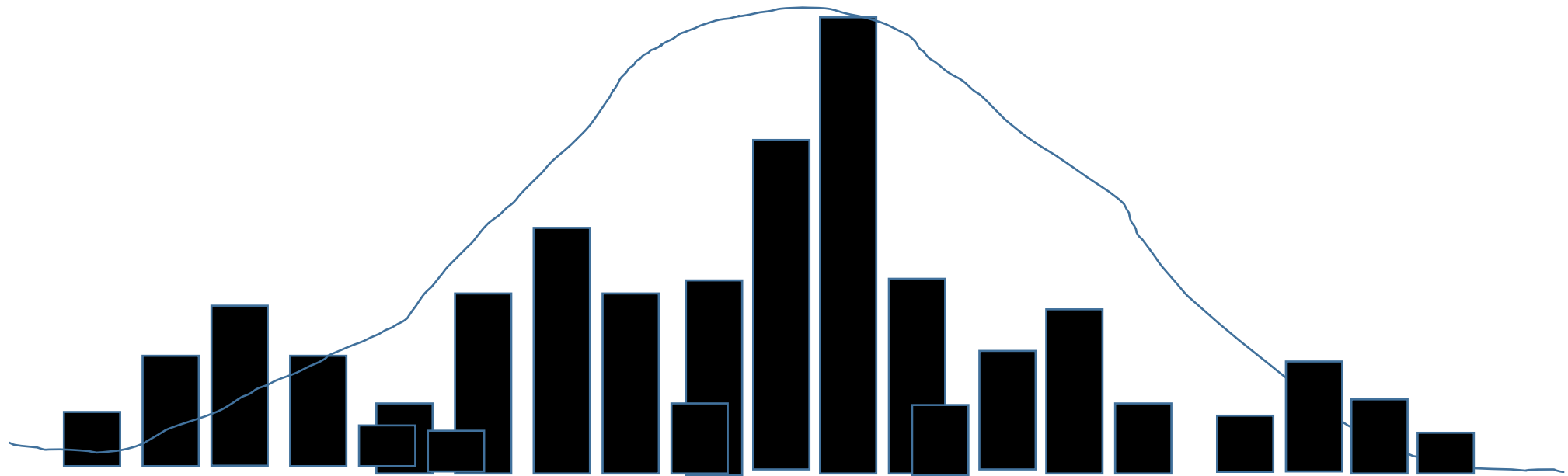
α_i



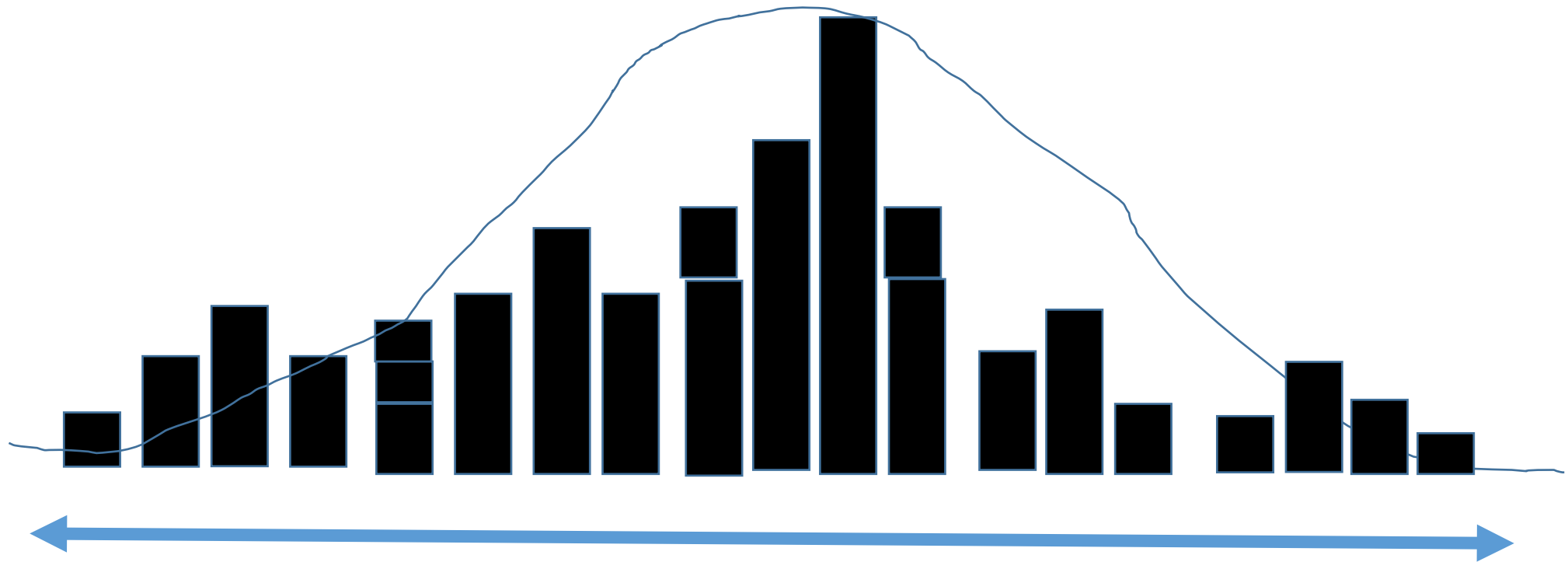
Variances and groups



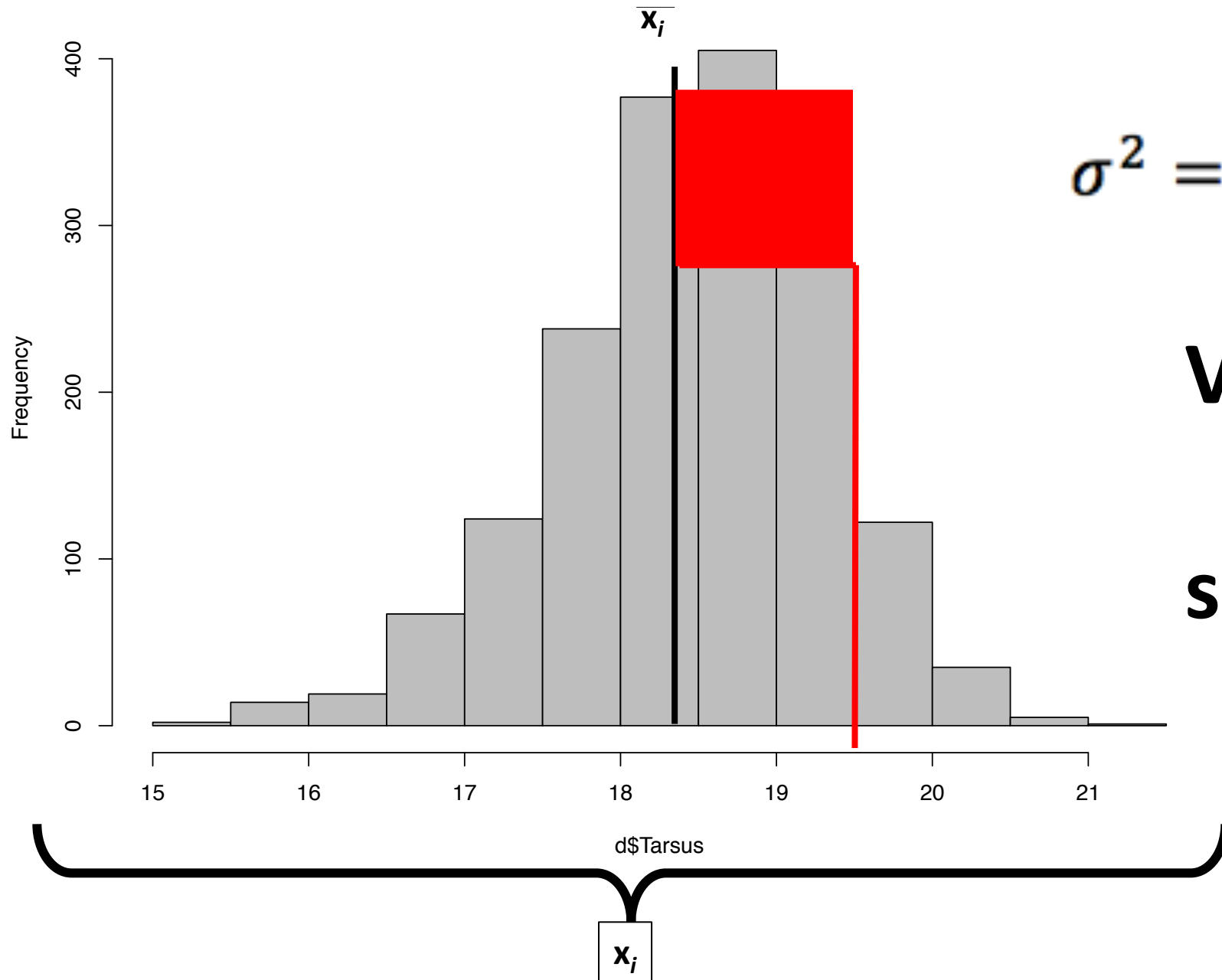
Variances and groups



Variances and groups



Total variance

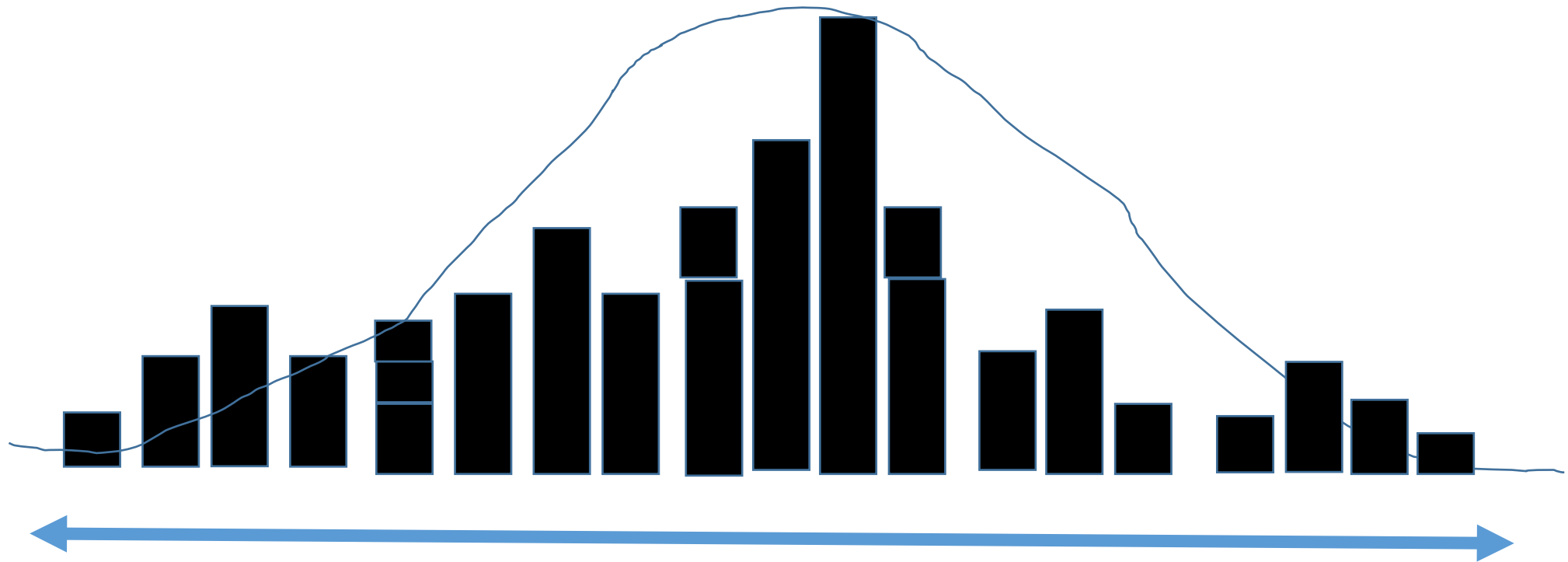


$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

VARIANCE

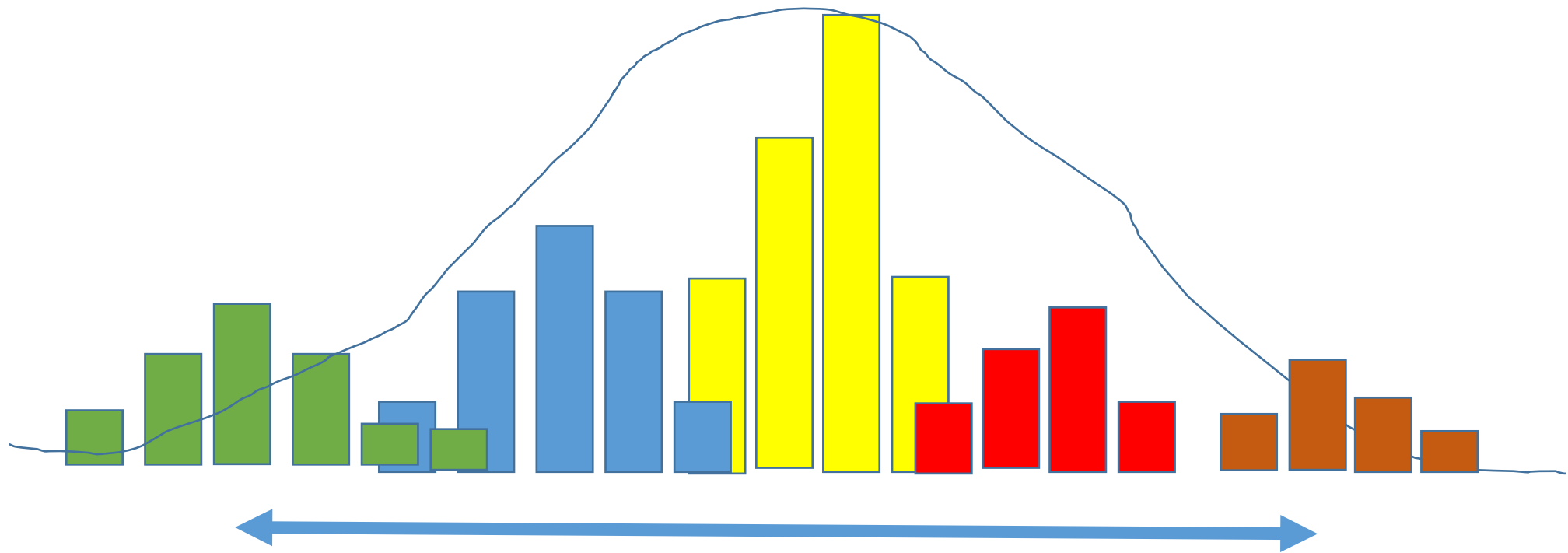
sd²

Variances and groups



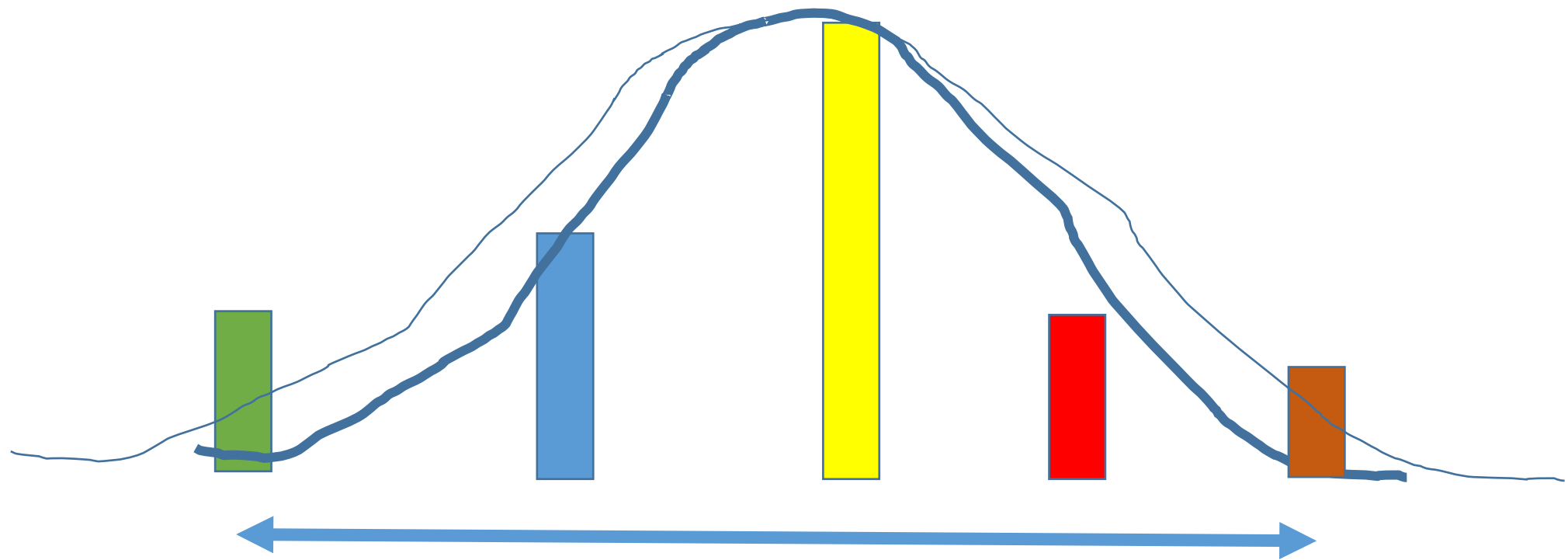
Total variance

Variances and groups



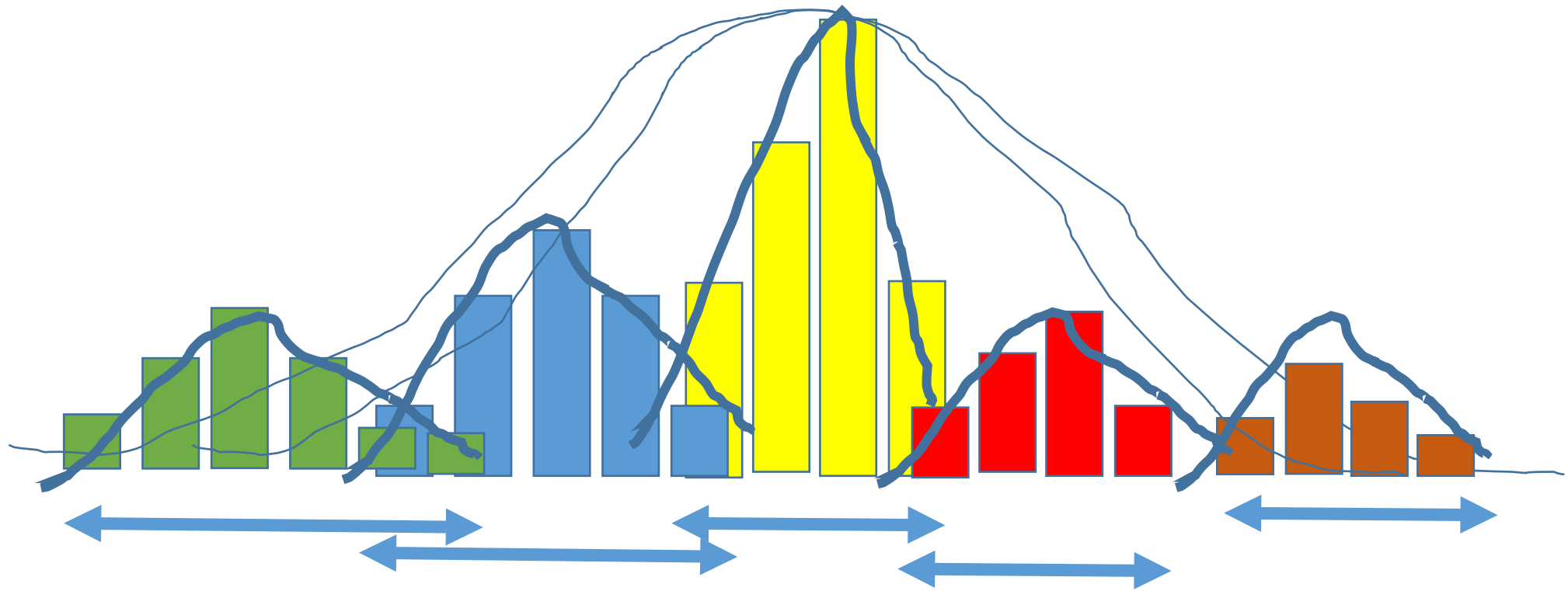
Between-group variance

Variances and groups



Between-group variance

Variances and groups



Within-group variance

Variances and groups

1,1,2,1,1,1 5,5,4,5,5,1 3,3,4,3,3,3

Variances and groups

k = number of groups
n = sample size in group
N = total sample size
i = row
j = column
 \bar{x} = group mean
 $\bar{\bar{x}}$ = grand total mean

1,1,2,1,1,1 5,5,4,5,5,1 3,3,4,3,3,3

Variances and groups

3 = k = number of groups
6 = n = sample size in group
18 = N = total sample size
1:18 i = individual counter
1:3 j = group counter
 \bar{x} = group mean
51 $\bar{\bar{x}}$ = grand total mean

1,1,2,1,1,1 5,5,4,5,5,1 3,3,4,3,3,3

Variances and groups

3 = k = number of groups
6 = n = sample size in group
18 = N = total sample size
1:18 i = individual counter
1:3 j = group counter
 \bar{x} = group mean
51 $\bar{\bar{x}}$ = grand total mean

j: 1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3

1,1,2,1,1,1 5,5,4,5,5,1 3,3,4,3,3,3

Variances and groups

3 = k = number of groups
6 = n = sample size in group
18 = N = total sample size
1:18 i = individual counter
1:3 j = group counter
 \bar{x} = group mean
51 $\bar{\bar{x}}$ = grand total mean

j: 1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3

i: 1,2,3,4,5,6,1,2,3,5,6

1,1,2,1,1,1 5,5,4,5,5,1 3,3,4,3,3,3

Variances and groups

3 = k = number of groups
6 = n = sample size in group
18 = N = total sample size
1:18 i = individual counter
1:3 j = group counter
 \bar{x} = group mean
51 $\bar{\bar{x}}$ = grand total mean

j: 1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3

i: 1,2,3,4,5,6,1,2,3,5,6

1,1,2,1,1,1 5,5,4,5,5,1 3,3,4,3,3,3

$$Y_{3,2} = 4$$

Variances and groups

3 = k = number of groups
6 = n = sample size in group
18 = N = total sample size
1:18 i = individual counter
1:3 j = group counter
 \bar{x} = group mean
51 $\bar{\bar{x}}$ = grand total mean

j: 1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3
i: 1,2,3,4,5,6,1,2,3,5,6
1,1,2,1,1,1 5,5,4,5,5,1 3,3,4,3,3,3

$$Y_{3,2} = 4$$

Variances and groups

j: 1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3

i: 1,2,3,4,5,6,1,2,3,5,6

$Y_{i,j}$: 1,1,2,1,1,1,5,5,4,5,5,1 3,3,4,3,3,3

$$\text{Between group variance} = \frac{SSG}{k - 1} =$$

3 = k = number of groups
6 = n_j = sample size in group
18 = N = total sample size
1:6 i = individual counter
1:3 j = group counter
 \bar{x}_j = group mean
51 $\bar{\bar{x}}_{i,j}$ = grand total mean

Variances and groups

j: 1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3

i: 1,2,3,4,5,6,1,2,3,5,6

$Y_{i,j}$: 1,1,2,1,1,15,5,4,5,5,1 3,3,4,3,3,3

$$\text{Between group variance} = \frac{SSG}{k - 1} = \frac{\sum_{j=1}^k n_j (\bar{x}_j - \bar{\bar{x}}_{i,j})^2}{k - 1}$$

3 = k = number of groups

6 = n_j = sample size in group

18 = N = total sample size

1:6 i = individual counter

1:3 j = group counter

\bar{x}_j = group mean

51 $\bar{\bar{x}}_{i,j}$ = grand total mean

Variances and groups

j: 1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3

i: 1,2,3,4,5,6,1,2,3,5,6

$Y_{i,j}$: 1,1,2,1,1,1,5,5,4,5,5,1 3,3,4,3,3,3

3 = k = number of groups
6 = n_j = sample size in group
18 = N = total sample size
1:6 i = individual counter
1:3 j = group counter
 \bar{x}_j = group mean
51 $\bar{\bar{x}}_{i,j}$ = grand total mean

$$\text{Within group variance} = \frac{SSE}{N - k} =$$

Variances and groups

j: 1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3

i: 1,2,3,4,5,6,1,2,3,5,6

$Y_{i,j}$: 1,1,2,1,1,1,5,5,4,5,5,1 3,3,4,3,3,3

3 = k = number of groups
 6 = n_j = sample size in group
 18 = N = total sample size
 1:6 i = individual counter
 1:3 j = group counter
 \bar{x}_j = group mean
 51 $\bar{\bar{x}}_{i,j}$ = grand total mean

$$\text{Within group variance} = \frac{SSE}{N - k} = \frac{\sum_{j=1}^k \sum_{i=1}^{n_j} (x_{i,j} - \bar{x}_j)^2}{N - k}$$

Variances and groups

j: 1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3

i: 1,2,3,4,5,6,1,2,3,5,6

$Y_{i,j}$: 1,1,2,1,1,15,5,4,5,5,1 3,3,4,3,3,3

3 = k = number of groups
 6 = n_j = sample size in group
 18 = N = total sample size
 1:6 i = individual counter
 1:3 j = group counter
 \bar{x}_j = group mean
 51 $\bar{\bar{x}}_{i,j}$ = grand total mean

$$\text{Within group variance} = \frac{SSE}{N - k} = \frac{\sum_{j=1}^k \sum_{i=1}^{n_j} (x_{i,j} - x_j)^2}{N - k}$$

$$\text{Between group variance} = \frac{SSG}{k - 1} = \frac{\sum_{j=1}^k n_j (\bar{x}_j - \bar{\bar{x}}_{i,j})^2}{k - 1}$$

$$\text{Total variance} = \frac{SST}{n - 1}$$

Variances and groups

j: 1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3

i: 1,2,3,4,5,6,1,2,3,5,6

$Y_{i,j}$: 1,1,2,1,1,15,5,4,5,5,1 3,3,4,3,3,3

3 = k = number of groups
 6 = n_j = sample size in group
 18 = N = total sample size
 1:6 i = individual counter
 1:3 j = group counter
 \bar{x}_j = group mean
 51 $\bar{\bar{x}}_{i,j}$ = grand total mean

$$\text{Within group variance} = \frac{SSE}{N - k} = \frac{\sum_{j=1}^k \sum_{i=1}^{n_j} (x_{i,j} - x_j)^2}{N - k}$$

$$\text{Between group variance} = \frac{SSG}{k - 1} = \frac{\sum_{j=1}^k n_j (\bar{x}_j - \bar{\bar{x}})^2}{k - 1}$$

$$\text{Total variance} = \frac{SST}{n - 1}$$

$$SST = SSG + SSE$$

LMMs

- So what does a mixed model look like?

LMMs

- So what does a mixed model look like?

LMMs

- So what does a mixed model look like?

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

Exam

Chick ID	Chick Mass	Tarsus	Nest				
A	5	3	L1				
B	3	1	L1				
C	6	4	L1				
D	10	8	S5				
E	4	2	S5				

LMMs

- Or estimate both. Let's look at the components of a LMM just like we did with the lm:

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i
5	1 A
3	2 B
6	3 C
10	4 D
4	5 D

$b_0 = ?$

$b_1 = ?$

x	i
3	1 A
1	2 B
4	3 C
8	4 D
2	5 D

ε	i
?	1 A
?	2 B
?	3 C
?	4 D
?	5 D

LMMs

- First things first. What's with the j ?

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1 A	1 L1
3	2 B	1 L1
6	3 C	1 L1
10	4 D	2 S5
4	5 D	2 S5

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1 L1
1	2	1 L1
4	3	1 L1
8	4	2 S5
2	5	2 S5

ε	i	j
?	1	1 L1
?	2	1 L1
?	3	1 L1
?	4	2 S5
?	5	2 S5

LMMs

- First things first. What's with the j ?
- J is a grouping factor. BirdID.

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

ε	i	j
?	1	1
?	2	1
?	3	1
?	4	2
?	5	2

LMMs

- First things first. What's with the j ?
- J is a grouping factor. BirdID. **Observer**.

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

ε	i	j
?	1	1
?	2	1
?	3	1
?	4	2
?	5	2

LMMs

- First things first. What's with the j ?
- J is a grouping factor. BirdID. Observer. NestID.

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

ε	i	j
?	1	1
?	2	1
?	3	1
?	4	2
?	5	2

LMMs

- First things first. What's with the j ?
- J is a grouping factor. BirdID. Observer. NestID. It's a factor in your data. $Year$.

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

ε	i	j
?	1	1
?	2	1
?	3	1
?	4	2
?	5	2

LMMs

- First things first. What's with the j ?
- J is a grouping factor. BirdID. Observer. NestID. It's a factor in your data. Year.
- It's **categorical**

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

ε	i	j
?	1	1
?	2	1
?	3	1
?	4	2
?	5	2

LMMs

- Ok. That only leaves α_j . What's that?

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

ε	i	j
?	1	1
?	2	1
?	3	1
?	4	2
?	5	2

LMMs

- Ok. That only leaves α_j . What's that?

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

Variance of
data
grouped
by j

ε	i	j
?	1	1
?	2	1
?	3	1
?	4	2
?	5	2

LMMs

- Ok. That only leaves α_j .
- So it's one number, with a measure of precision, like the bs!

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

Variance of
data
grouped
by j

ε	i	j
?	1	1
?	2	1
?	3	1
?	4	2
?	5	2

LMMs

- Ok. That only leaves **alpha-j**.
- So it's one number, with a measure of precision , like the bs!
- And we color it red because we want to estimate it!

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

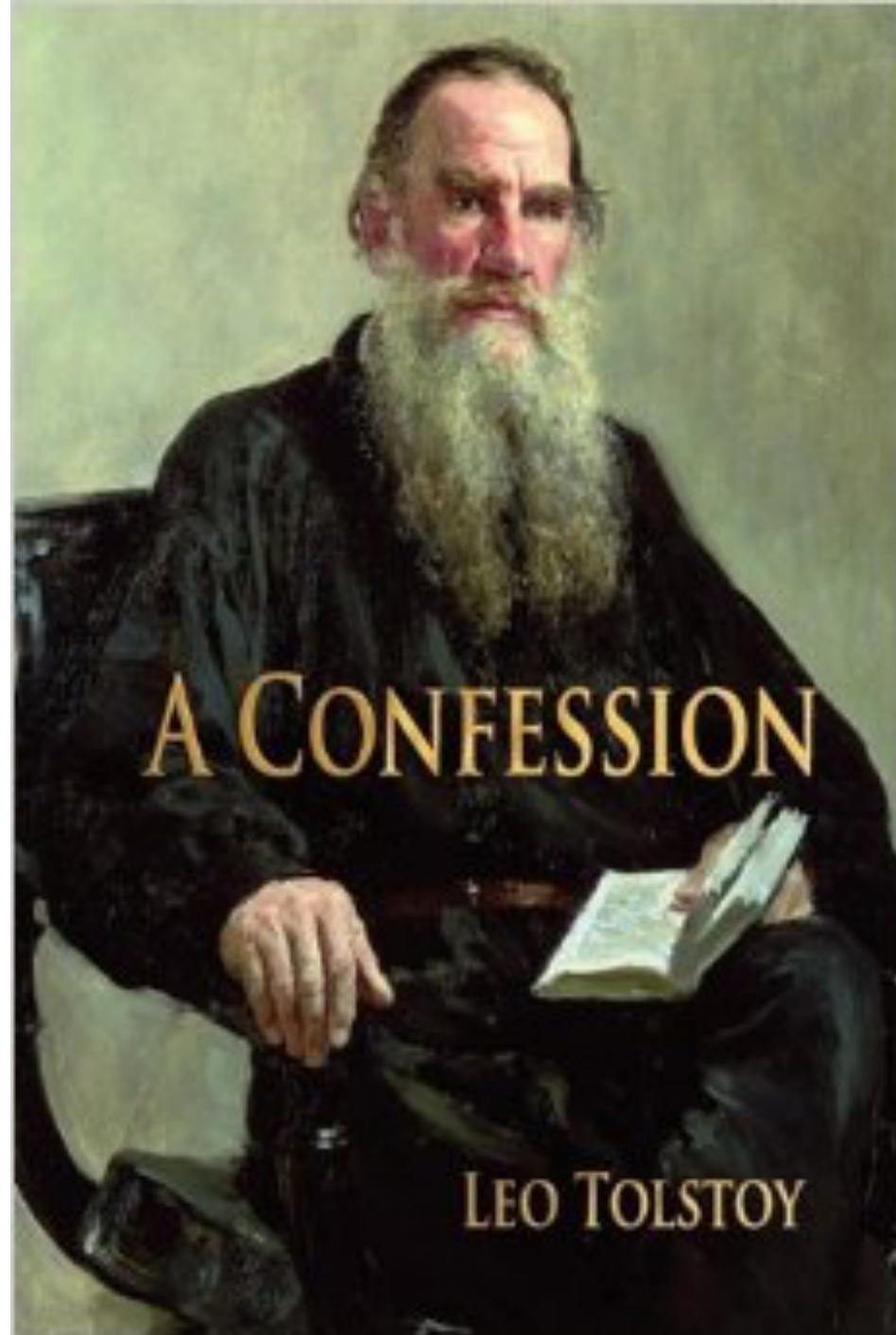
$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

y	j	Group
5	1	variance of
3	1	data y
6	1	grouped
10	2	by j
4	2	

ε	i	j
?	1	1
?	2	1
?	3	1
?	4	2
?	5	2



A CONFESSION

LEO TOLSTOY

LMMs – a confession

typically, we report for e i,j only
the variance - so it's a number.

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

y	j	Group variance of data y grouped by j
5	1	
3	1	
6	1	
10	2	
4	2	

ε	i	j	Variance of residuals
?	1	1	
?	2	1	
?	3	1	
?	4	2	
?	5	2	

LMMs – a confession

- Ok, cool. Let's revisit that group variance thing.

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

y	j	Group variance of data y grouped by j
5	1	
3	1	
6	1	
10	2	
4	2	

ε	i	j	Variance of residuals
?	1	1	
?	2	1	
?	3	1	
?	4	2	
?	5	2	

Between-group variance



$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$


x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

y	j	Group variance of data y grouped by j
5	1	
3	1	
6	1	
10	2	
4	2	

ε	i	j	Variance of residuals
?	1	1	
?	2	1	
?	3	1	
?	4	2	
?	5	2	

Between-group variance

Within-group variance


$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$


x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

y	j	Group variance of data y grouped by j
5	1	
3	1	
6	1	
10	2	
4	2	

ε	i	j	Variance of residuals
?	1	1	
?	2	1	
?	3	1	
?	4	2	
?	5	2	

Between-group variance

Within-group variance
Residual variance


$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

y	j	Group variance of data y grouped by j
5	1	
3	1	
6	1	
10	2	
4	2	


ε	i	j	Variance of residuals
?	1	1	
?	2	1	
?	3	1	
?	4	2	
?	5	2	

Between-group variance

Within-group variance

Residual variance

Error variance


$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

y	j	Group variance of data y grouped by j
5	1	
3	1	
6	1	
10	2	
4	2	

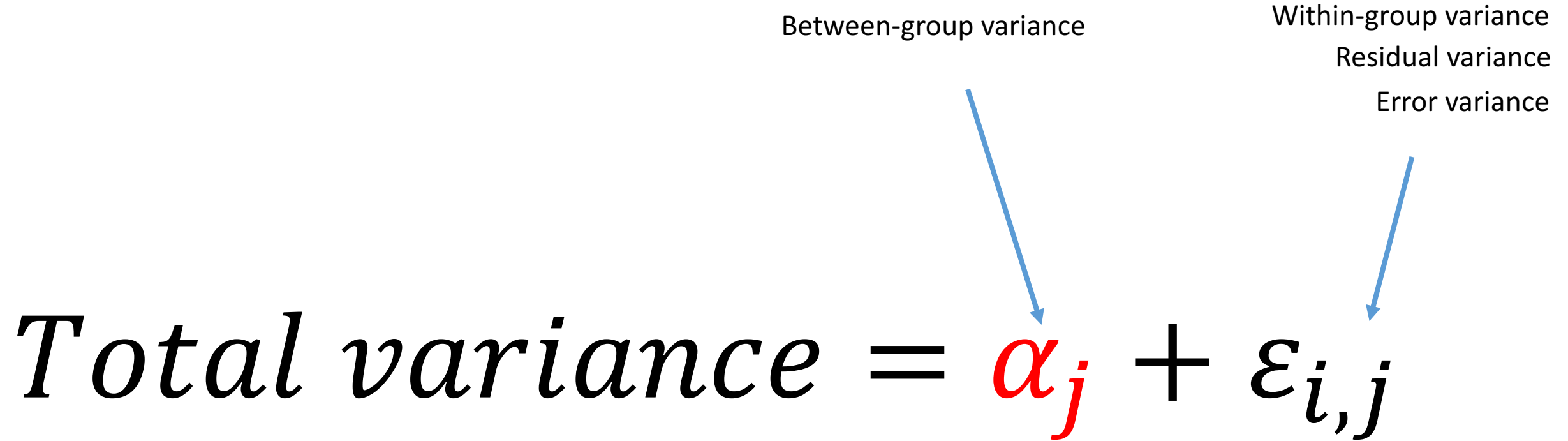
ε	i	j	Variance of residuals
?	1	1	
?	2	1	
?	3	1	
?	4	2	
?	5	2	

Between-group variance

Within-group variance

Residual variance

Error variance



The diagram illustrates the decomposition of total variance into two components. The equation $Total\ variance = \alpha_j + \varepsilon_{i,j}$ is shown in an italicized font. A blue arrow points from the text 'Between-group variance' to the term α_j , which is highlighted in red. Another blue arrow points from the text 'Within-group variance', 'Residual variance', and 'Error variance' to the term $\varepsilon_{i,j}$.

$$Total\ variance = \alpha_j + \varepsilon_{i,j}$$

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

y	j	
5	1	Group variance of data y grouped by j
3	1	
6	1	
10	2	
4	2	

Between-group variance

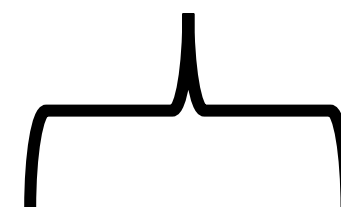
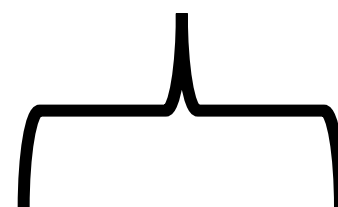
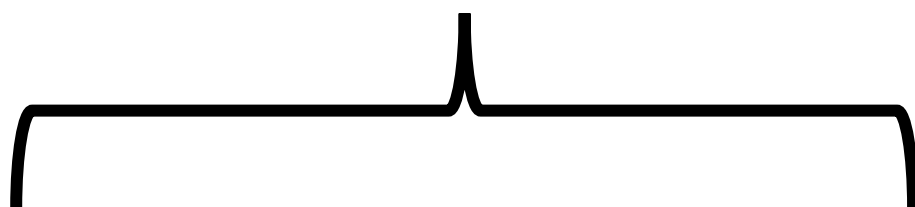
ε	i	j	
?	1	1	Variance of residuals
?	2	1	
?	3	1	
?	4	2	
?	5	2	

Residual variance

Fixed

Random

Error



$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

y	i	j
5	1	1
3	2	1
6	3	1
10	4	2
4	5	2

$b_0 = ?$

$b_1 = ?$

x	i	j
3	1	1
1	2	1
4	3	1
8	4	2
2	5	2

y	j	
5	1	Group variance of data y grouped by j
3	1	
6	1	
10	2	
4	2	

Between-group variance

ε	i	j	
?	1	1	Variance of residuals
?	2	1	
?	3	1	
?	4	2	
?	5	2	

Residual variance

LMMs

- You can choose to only estimate one part and set the other one to fixed:

LMMs

- You can choose to only estimate one part

$$y_{i,j} = b_0 + b_1 x_{i,j} + \alpha_j + \varepsilon_{i,j}$$

- You could set the random part to zero (always keep the error!)

LMMs

- You can choose to only estimate one part

$$y_{i,j} = 1 + \alpha_j + \varepsilon_{i,j}$$

- You could set the fixed part to one (always have an intercept is a good idea)

Mixed models

- Estimate variance components simultaneously to fixed terms

Mixed models

- Estimate variance components simultaneously to fixed terms
- Allow to account for nested structure in data

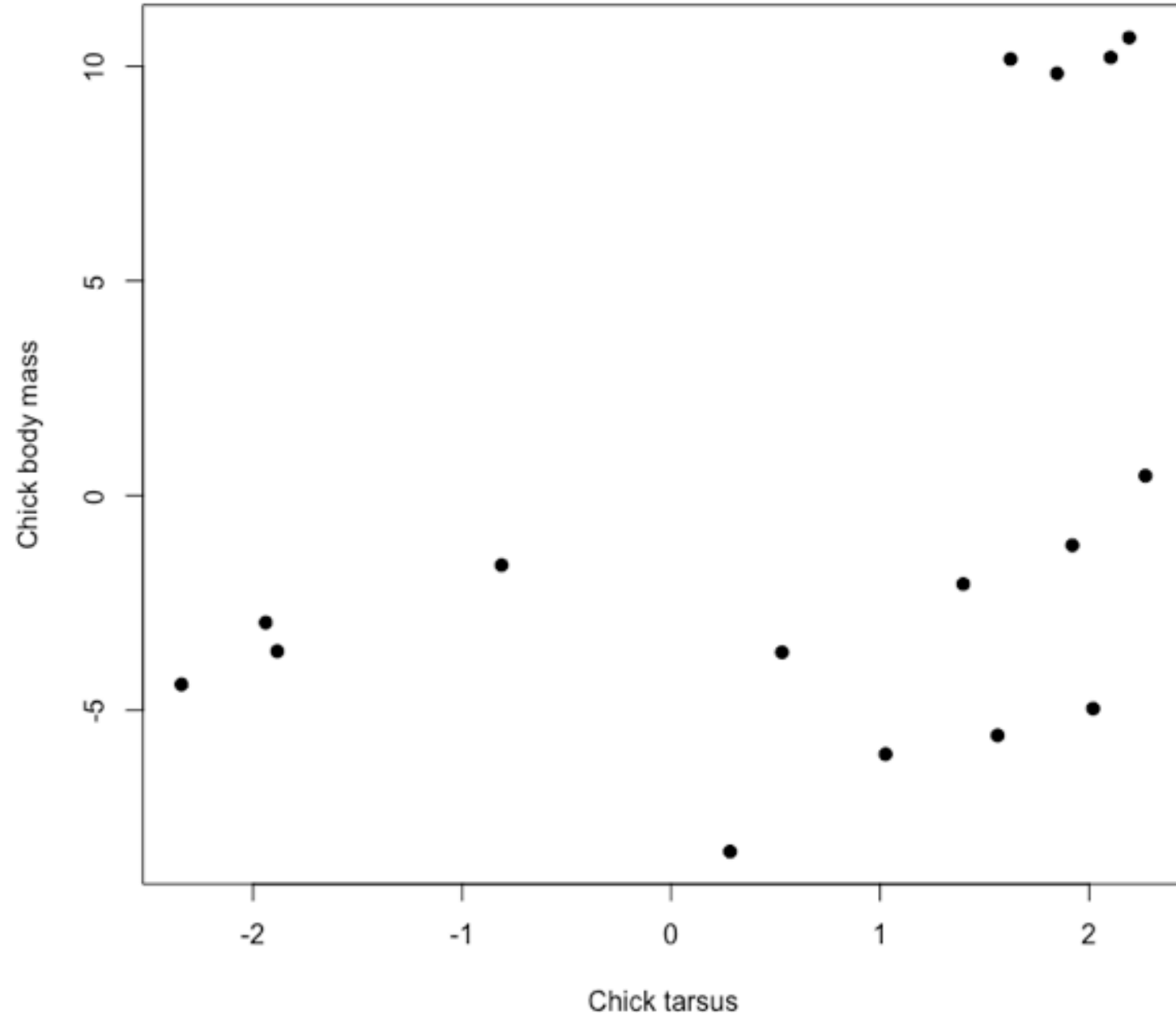
Mixed models

- Estimate variance components simultaneously to fixed terms
- Allow to account for nested structure in data
- Some things are more similar than others



Mixed models

- Estimate variance components simultaneously to fixed terms
- Allow to account for nested structure in data
- Some things are more similar than others



Mixed models

```
> m<-(lm(bm~tarsus))  
> summary(m)
```

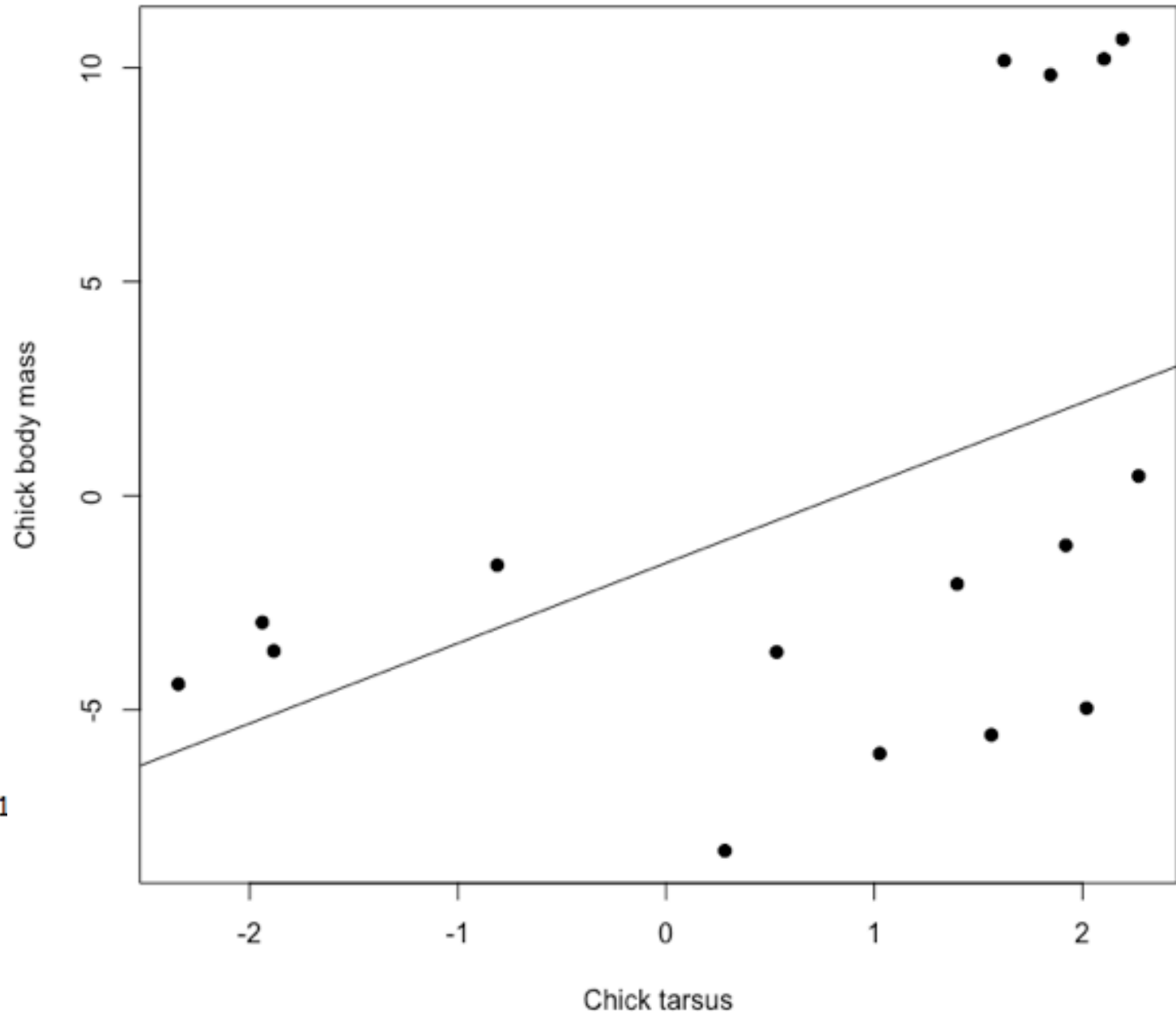
```
Call:  
lm(formula = bm ~ tarsus)
```

```
Residuals:  
    Min       1Q   Median       3Q      Max  
-7.2535 -3.9865 -0.3775  3.6447  8.6954
```

```
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)  -1.570      1.662  -0.944   0.3610  
tarsus         1.875      0.965   1.943   0.0725 .  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 6.008 on 14 degrees of freedom  
Multiple R-squared:  0.2123,    Adjusted R-squared:  0.1561  
F-statistic: 3.774 on 1 and 14 DF,  p-value: 0.07246
```

```
> |
```



Mixed models

```
> m<-(lm(bm~tarsus))  
> summary(m)
```

```
Call:  
lm(formula = bm ~ tarsus)
```

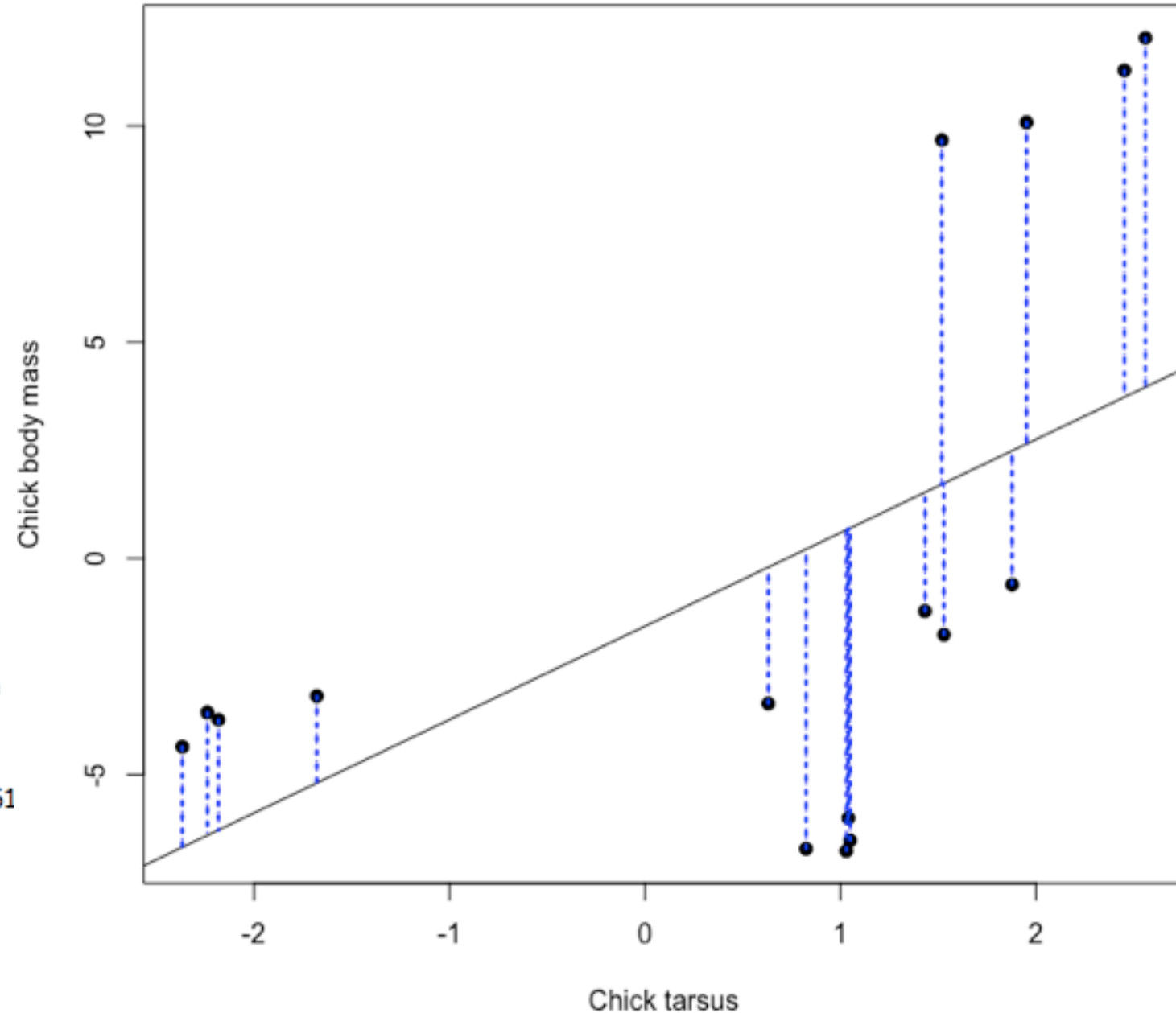
```
Residuals:  
    Min       1Q   Median       3Q      Max  
-7.2535 -3.9865 -0.3775  3.6447  8.6954
```

```
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)  -1.570      1.662  -0.944   0.3610  
tarsus         1.875      0.965   1.943   0.0725 .
```

```
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 6.008 on 14 degrees of freedom  
Multiple R-squared:  0.2123,    Adjusted R-squared:  0.1561  
F-statistic: 3.774 on 1 and 14 DF,  p-value: 0.07246
```

```
> |
```



Mixed models

```
> m<-(lm(bm~tarsus))  
> summary(m)
```

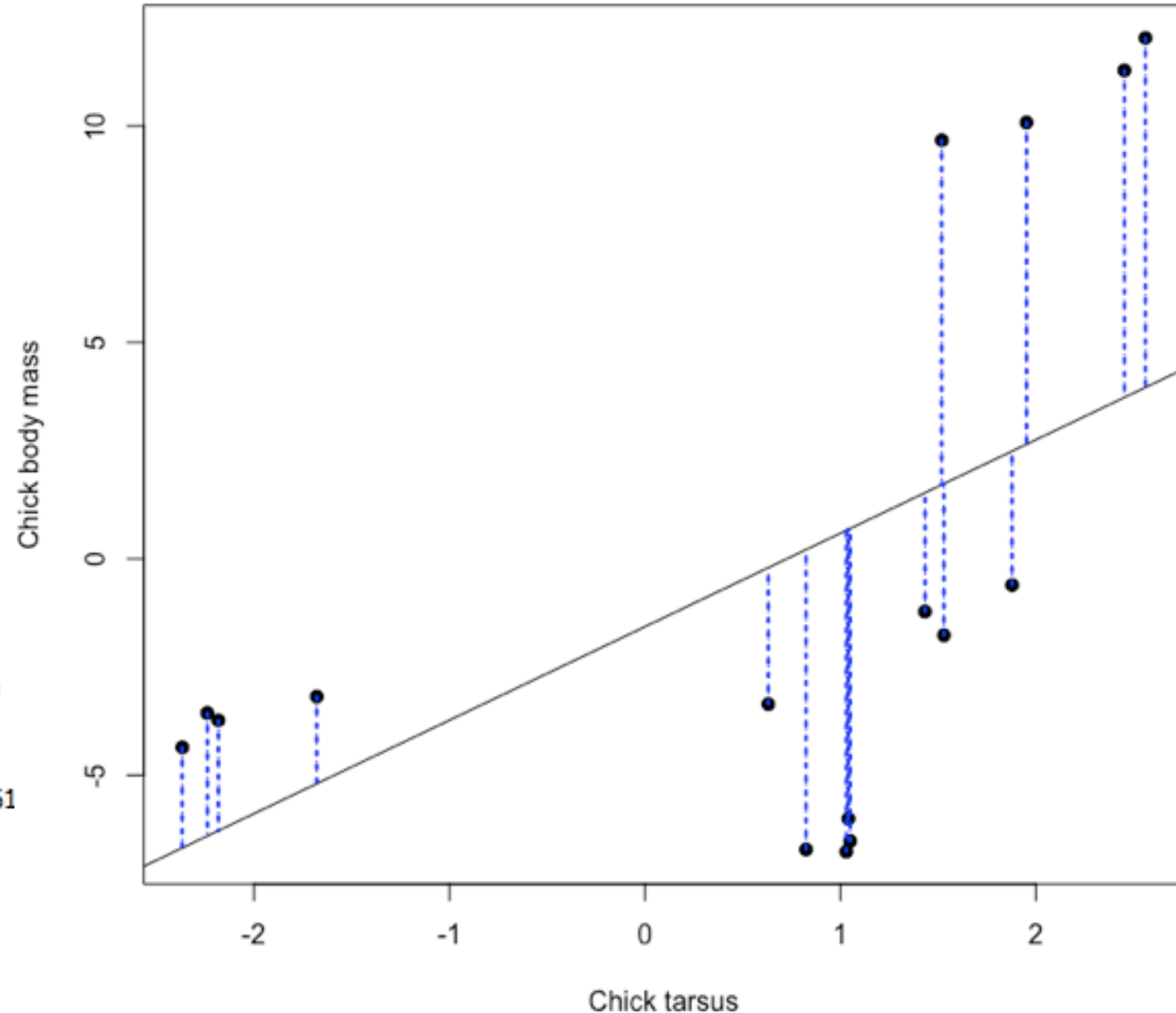
```
Call:  
lm(formula = bm ~ tarsus)
```

```
Residuals:  
    Min       1Q   Median       3Q      Max  
-7.2535 -3.9865 -0.3775  3.6447  8.6954
```

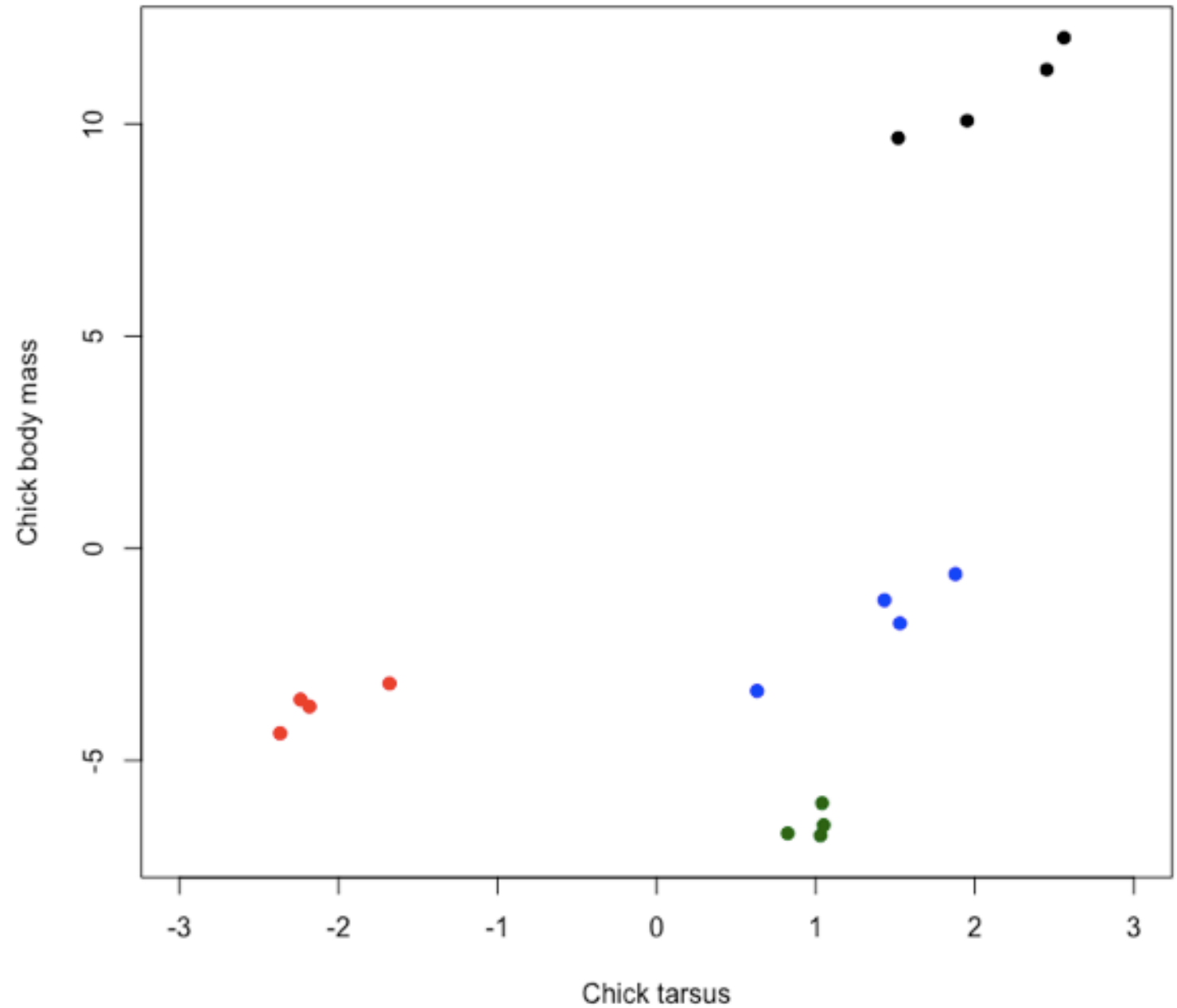
```
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)  
(Intercept)  -1.570      1.662   -0.944   0.3610  
tarsus         1.875      0.965    1.943   0.0725 .  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 6.008 on 14 degrees of freedom  
Multiple R-squared:  0.2123,    Adjusted R-squared:  0.1561  
F-statistic: 3.774 on 1 and 14 DF,  p-value: 0.07246
```

```
> |
```

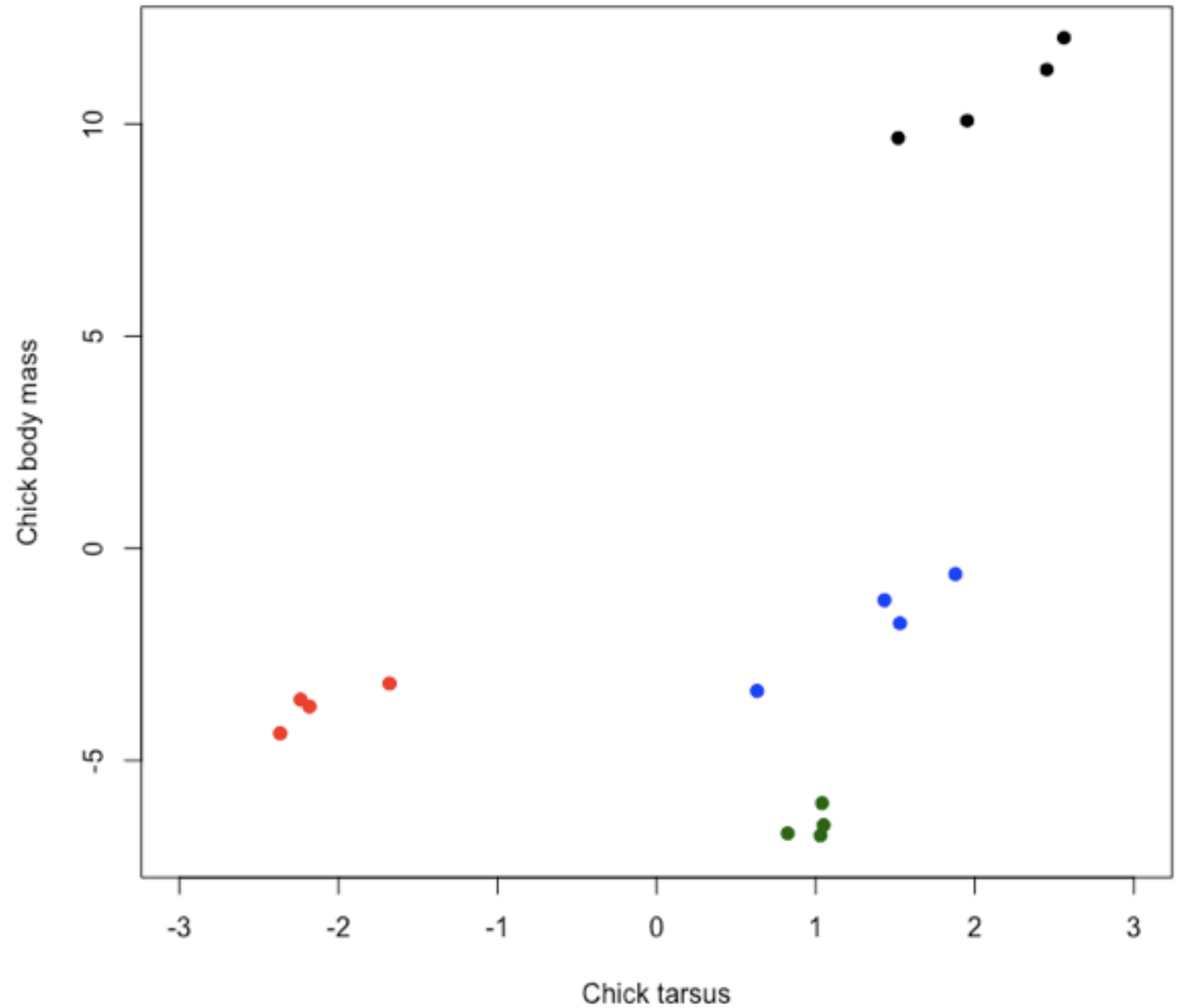


Mixed models



Mixed models

```
> require(lme4)  
> mm<-(lmer(bm~tarsus+(1|nest)))
```



Mixed models

```
> require(lme4)
> mm<-(lmer(bm~tarsus+(1|nest)))
> summary(mm)
```

Linear mixed model fit by REML ['lmerMod']
Formula: bm ~ tarsus + (1 | nest)

REML criterion at convergence: 34.4

Scaled residuals:

	Min	1Q	Median	3Q	Max
	-1.08739	-0.60062	-0.05266	0.58491	1.18030

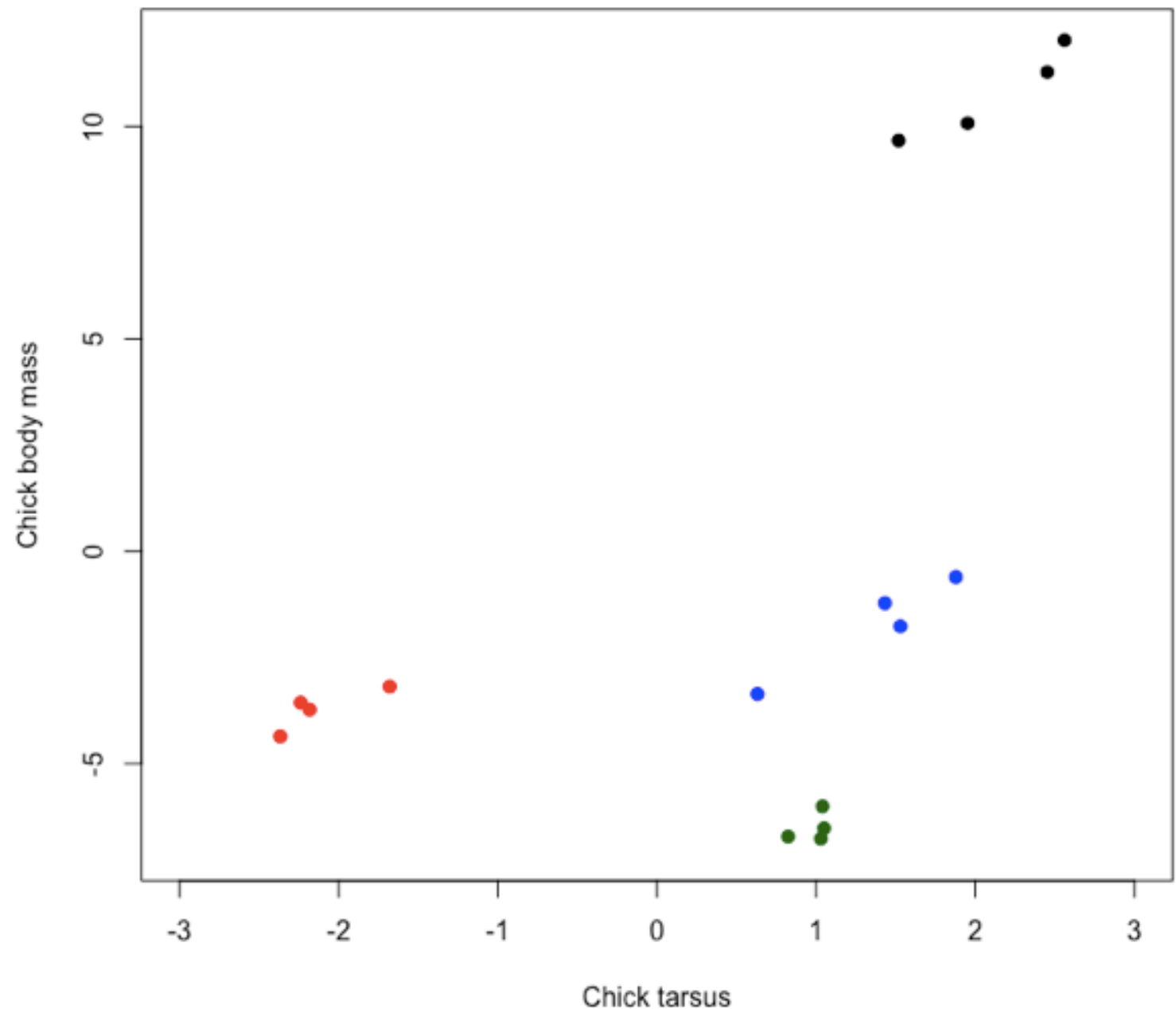
Random effects:

Groups	Name	Variance	Std.Dev.
nest	(Intercept)	42.1465	6.492
Residual		0.1115	0.334

Number of obs: 16, groups: nest, 4

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-1.4939	3.2503	-0.46
tarsus	2.0325	0.2446	8.31



Mixed models

```
> require(lme4)
> mm<-(lmer(bm~tarsus+(1|nest)))
> summary(mm)
```

Linear mixed model fit by REML ['lmerMod']
Formula: bm ~ tarsus + (1 | nest)

REML criterion at convergence: 34.4

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.08739	-0.60062	-0.05266	0.58491	1.18030

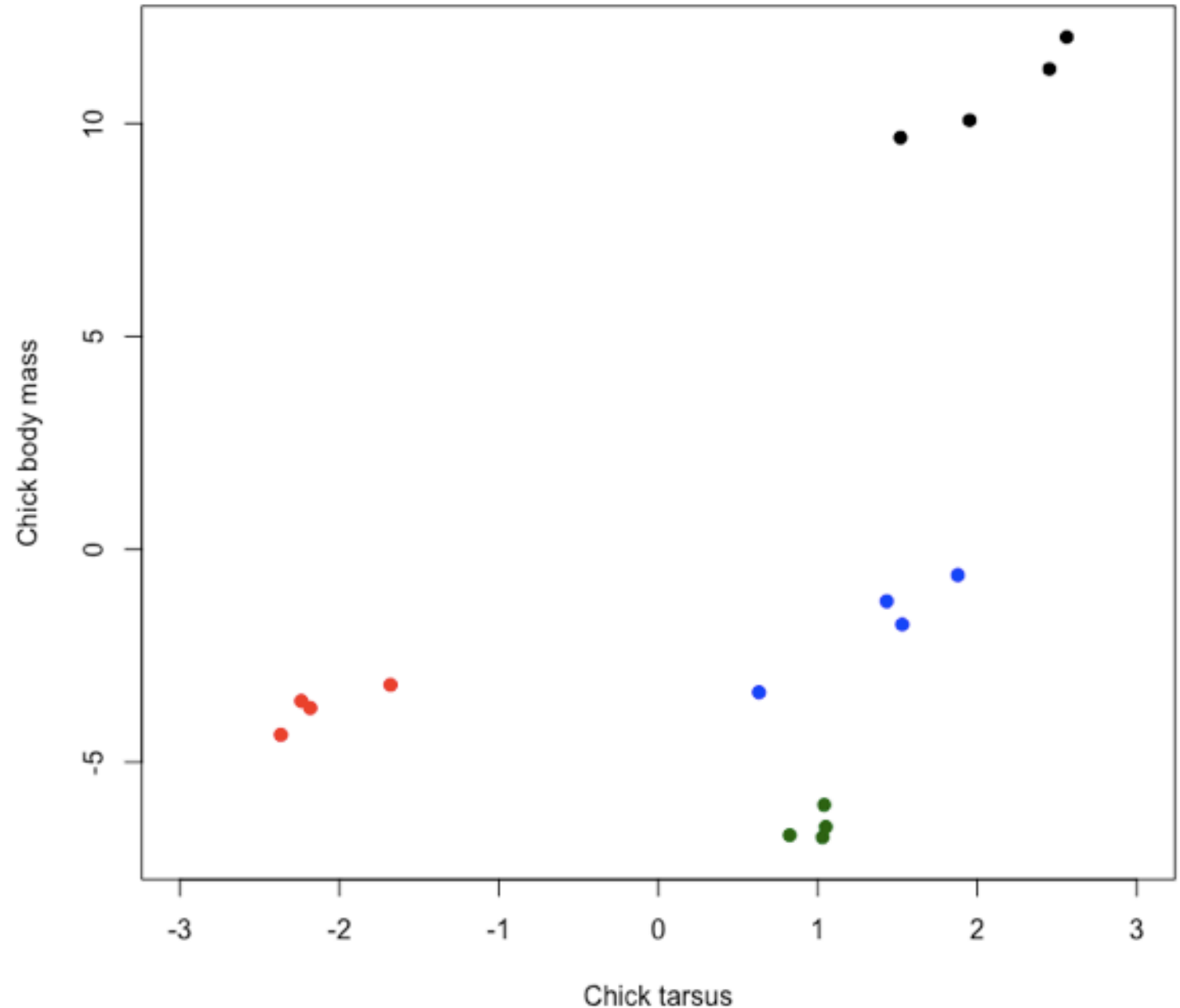
Random effects:

Groups	Name	Variance	Std.Dev.
nest	(Intercept)	42.1465	6.492
Residual		0.1115	0.334

Number of obs: 16, groups: nest, 4

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-1.4939	3.2503	-0.46
tarsus	2.0325	0.2446	8.31



Mixed models

```
> require(lme4)
> mm<-(lmer(bm~tarsus+(1|nest)))
> summary(mm)
```

Linear mixed model fit by REML ['lmerMod']
Formula: bm ~ tarsus + (1 | nest)

REML criterion at convergence: 34.4

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.08739	-0.60062	-0.05266	0.58491	1.18030

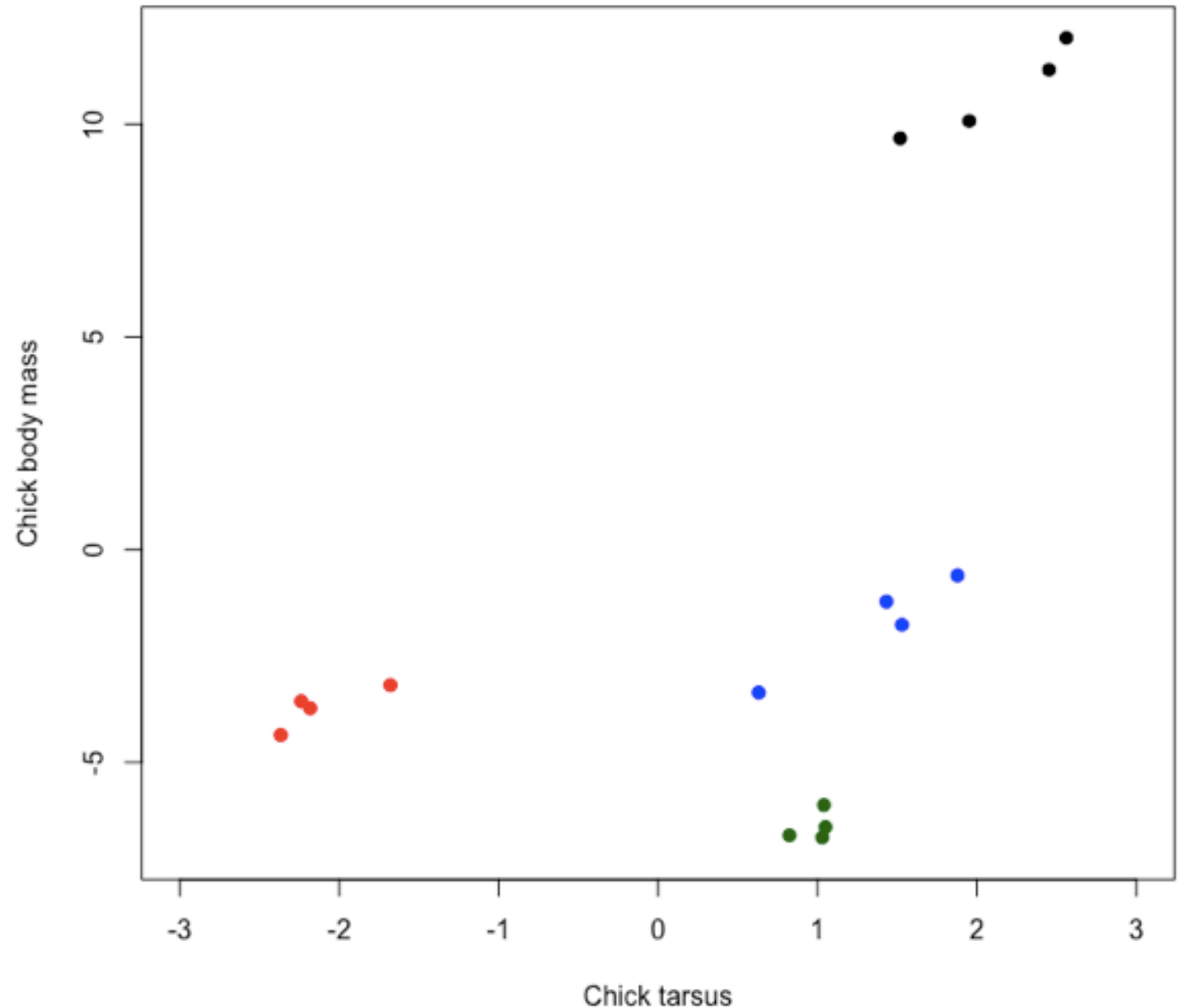
Random effects:

Groups	Name	Variance	Std.Dev.
nest	(Intercept)	42.1465	6.492
Residual		0.1115	0.334

Number of obs: 16, groups: nest, 4

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-1.4939	3.2503	-0.46
tarsus	2.0325	0.2446	8.31



Mixed models

```
> require(lme4)
> mm<-(lmer(bm~tarsus+(1|nest)))
> summary(mm)
```

Linear mixed model fit by REML ['lmerMod']
Formula: bm ~ tarsus + (1 | nest)

REML criterion at convergence: 34.4

Scaled residuals:

	Min	1Q	Median	3Q	Max
	-1.08739	-0.60062	-0.05266	0.58491	1.18030

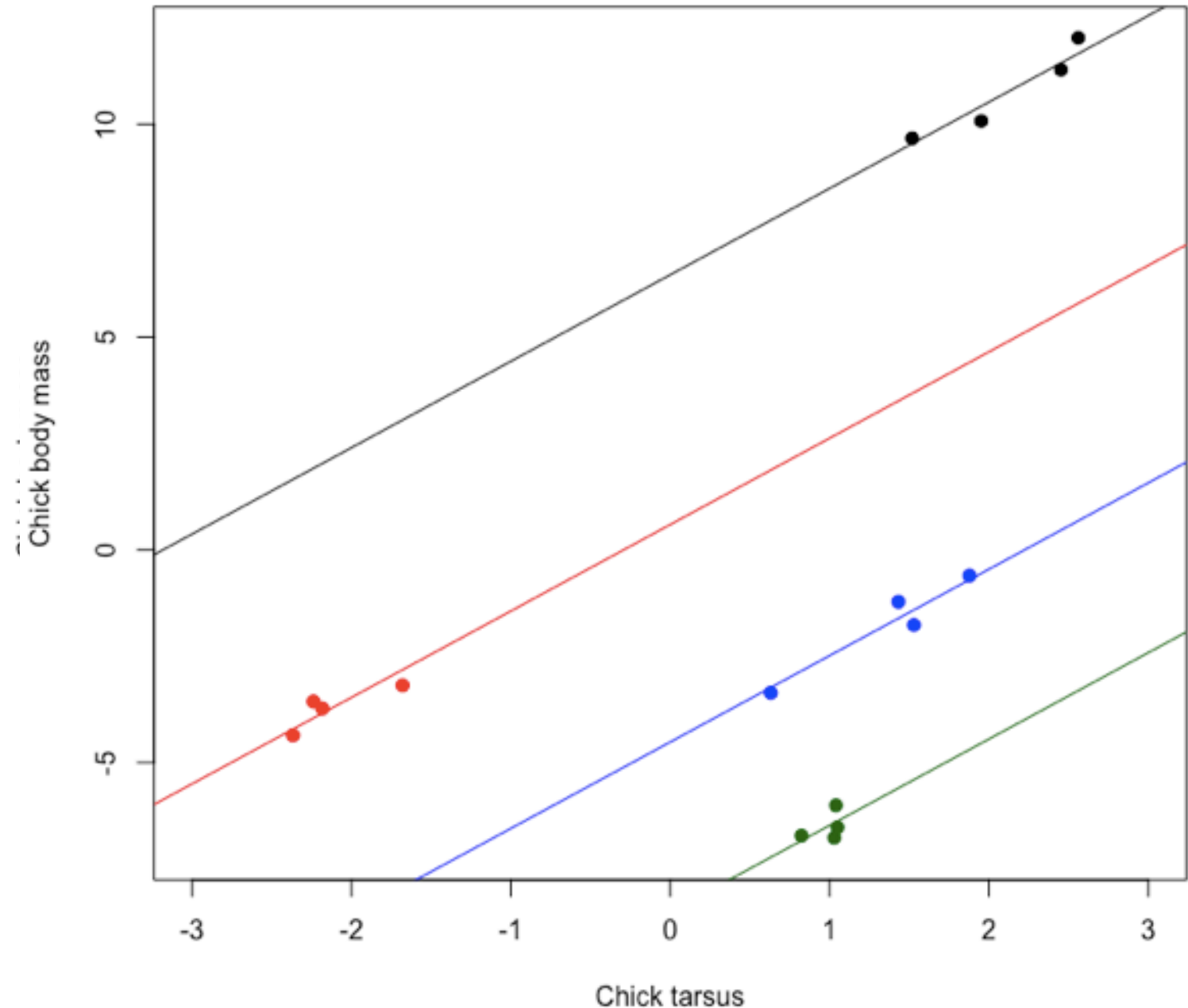
Random effects:

Groups	Name	Variance	Std.Dev.
nest	(Intercept)	42.1465	6.492
Residual		0.1115	0.334

Number of obs: 16, groups: nest, 4

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-1.4939	3.2503	-0.46
tarsus	2.0325	0.2446	8.31



Mixed models

```
> require(lme4)
> mm<-(lmer(bm~tarsus+(1|nest)))
> summary(mm)
```

Linear mixed model fit by REML ['lmerMod']
Formula: bm ~ tarsus + (1 | nest)

REML criterion at convergence: 34.4

Scaled residuals:

	Min	1Q	Median	3Q	Max
	-1.08739	-0.60062	-0.05266	0.58491	1.18030

Random effects:

Groups	Name	Variance	Std.Dev.
nest	(Intercept)	42.1465	6.492
Residual		0.1115	0.334

Number of obs: 16, groups: nest, 4

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-1.4939	3.2503	-0.46
tarsus	2.0325	0.2446	8.31

```
> m<-(lm(bm~tarsus))
> summary(m)
```

Call:
lm(formula = bm ~ tarsus)

Residuals:

	Min	1Q	Median	3Q	Max
	-7.2535	-3.9865	-0.3775	3.6447	8.6954

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.570	1.662	-0.944	0.3610
tarsus	1.875	0.965	1.943	0.0725 .

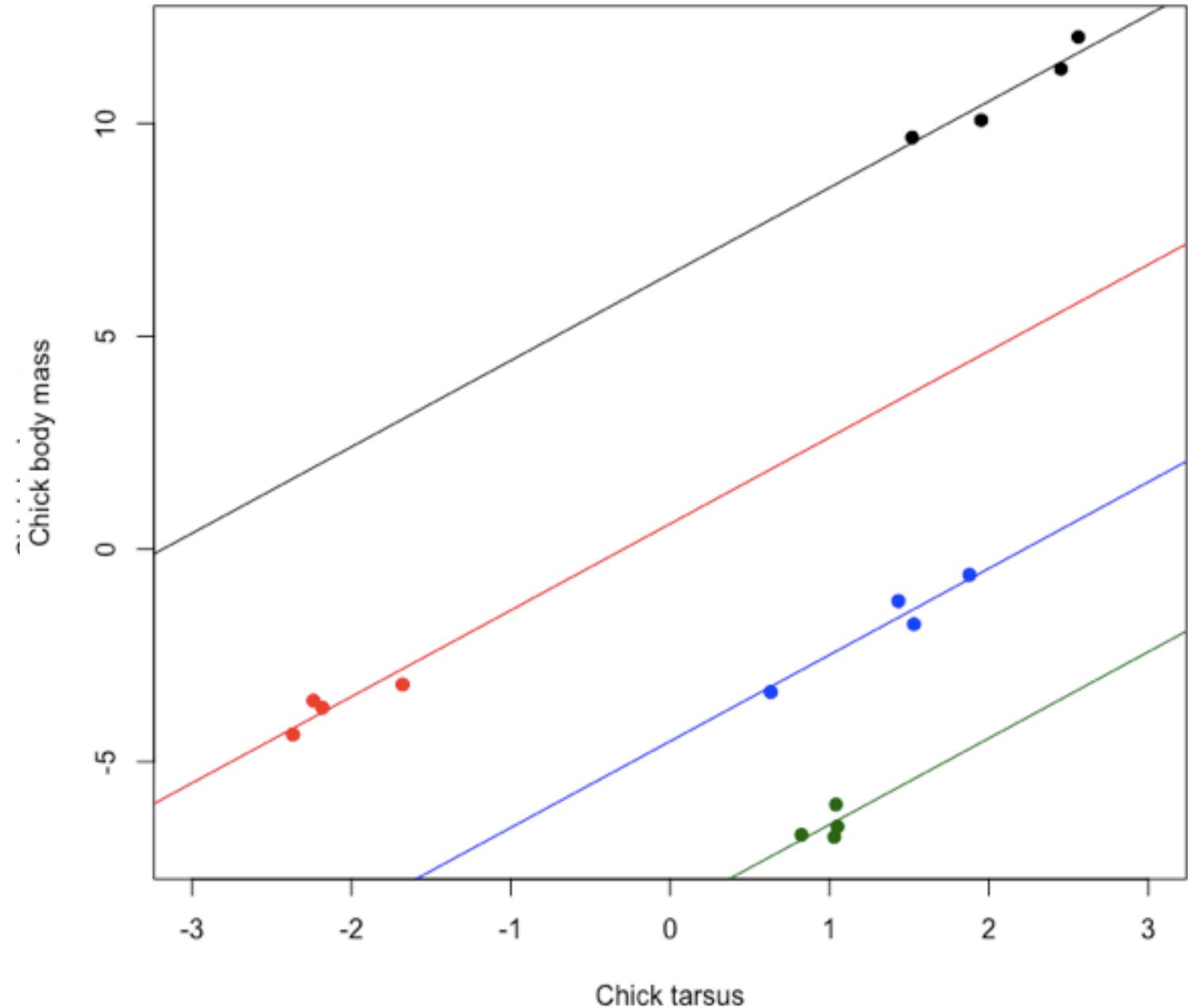
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.008 on 14 degrees of freedom
Multiple R-squared: 0.2123, Adjusted R-squared: 0.1561
F-statistic: 3.774 on 1 and 14 DF, p-value: 0.07246

```
> |
```

Mixed models

- Ok, so why don't we add nest as fixed factor instead?



Nest: Random

```
> require(lme4)
> mm<-(lmer(bm~tarsus+(1|nest)))
> summary(mm)
Linear mixed model fit by REML ['lmerMod']
Formula: bm ~ tarsus + (1 | nest)

REML criterion at convergence: 34.4

Scaled residuals:
    Min       1Q   Median       3Q      Max
-1.08739 -0.60062 -0.05266  0.58491  1.18030

Random effects:
 Groups   Name      Variance Std.Dev.
 nest     (Intercept) 42.1465   6.492
 Residual                    0.1115   0.334
Number of obs: 16, groups:  nest, 4

Fixed effects:
            Estimate Std. Error t value
(Intercept)  -1.4939     3.2503   -0.46
tarsus         2.0325     0.2446    8.31
```

Fixed

```
> summary(lm(bm~tarsus+nest))

Call:
lm(formula = bm ~ tarsus + nest)

Residuals:
    Min       1Q   Median       3Q      Max
-0.36367 -0.20454 -0.01581  0.18941  0.39257

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   6.4606     0.5490   11.767 1.42e-07 ***
tarsus         2.0305     0.2466    8.234 4.96e-06 ***
nestB        -5.8707     1.0715   -5.479 0.000192 ***
nestC       -10.9770     0.3005  -36.525 7.81e-13 ***
nestD       -14.9654     0.3663  -40.858 2.29e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3342 on 11 degrees of freedom
Multiple R-squared:  0.9983,    Adjusted R-squared:  0.9976
F-statistic: 1581 on 4 and 11 DF,  p-value: 4.275e-15
```

Nest: Random

```
> require(lme4)
> mm<-(lmer(bm~tarsus+(1|nest)))
> summary(mm)
Linear mixed model fit by REML ['lmerMod']
Formula: bm ~ tarsus + (1 | nest)

REML criterion at convergence: 34.4

Scaled residuals:
    Min       1Q   Median       3Q      Max
-1.08739 -0.60062 -0.05266  0.58491  1.18030

Random effects:
 Groups   Name      Variance Std.Dev.
 nest    (Intercept) 42.1465   6.492
 Residual                0.1115   0.334
Number of obs: 16, groups: nest, 4

Fixed effects:
              Estimate Std. Error t value
(Intercept)  -1.4030    3.2503   -0.46
tarsus        2.0325    0.2446    8.31
```

Fixed

```
> summary(lm(bm~tarsus+nest))

Call:
lm(formula = bm ~ tarsus + nest)

Residuals:
    Min       1Q   Median       3Q      Max
-0.36367 -0.20454 -0.01581  0.18941  0.39257

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   6.4606     0.5490   11.767 1.42e-07 ***
tarsus         2.0305     0.2466    8.234 4.90e-06 ***
nestB         -5.8707     1.0715   -5.479 0.000192 ***
nestC        -10.9770     0.3005  -36.525 7.81e-13 ***
nestD        -14.9654     0.3663  -40.858 2.29e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3342 on 11 degrees of freedom
Multiple R-squared:  0.9983,    Adjusted R-squared:  0.9976
F-statistic: 1581 on 4 and 11 DF,  p-value: 4.275e-15
```


Nest: Random

```
> require(lme4)
> mm<-(lmer(bm~tarsus+(1|nest)))
> summary(mm)
Linear mixed model fit by REML ['lmerMod']
Formula: bm ~ tarsus + (1 | nest)

REML criterion at convergence: 34.4

Scaled residuals:
    Min       1Q   Median       3Q      Max
-1.08739 -0.60062 -0.05266  0.58491  1.18030

Random effects:
 Groups   Name      Variance Std.Dev.
 nest     (Intercept) 42.1465   6.492
 Residual                    0.1115   0.334
Number of obs: 16, groups: nest, 4

Fixed effects:
              Estimate Std. Error t value
(Intercept)  -1.4939     3.2503    -0.46
tarsus         2.0325     0.2446     8.31
```

Fixed

```
> summary(lm(bm~tarsus+nest))

Call:
lm(formula = bm ~ tarsus + nest)

Residuals:
    Min       1Q   Median       3Q      Max
-0.36367 -0.20454 -0.01581  0.18941  0.39257

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    6.4606     0.5490   11.767 1.42e-07 ***
tarsus         2.0305     0.2466    8.234 4.96e-06 ***
nestB        -5.8707     1.0715   -5.479 0.000192 ***
nestC       -10.9770     0.3005  -36.525 7.81e-13 ***
nestD       -14.9654     0.3663  -40.858 2.29e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

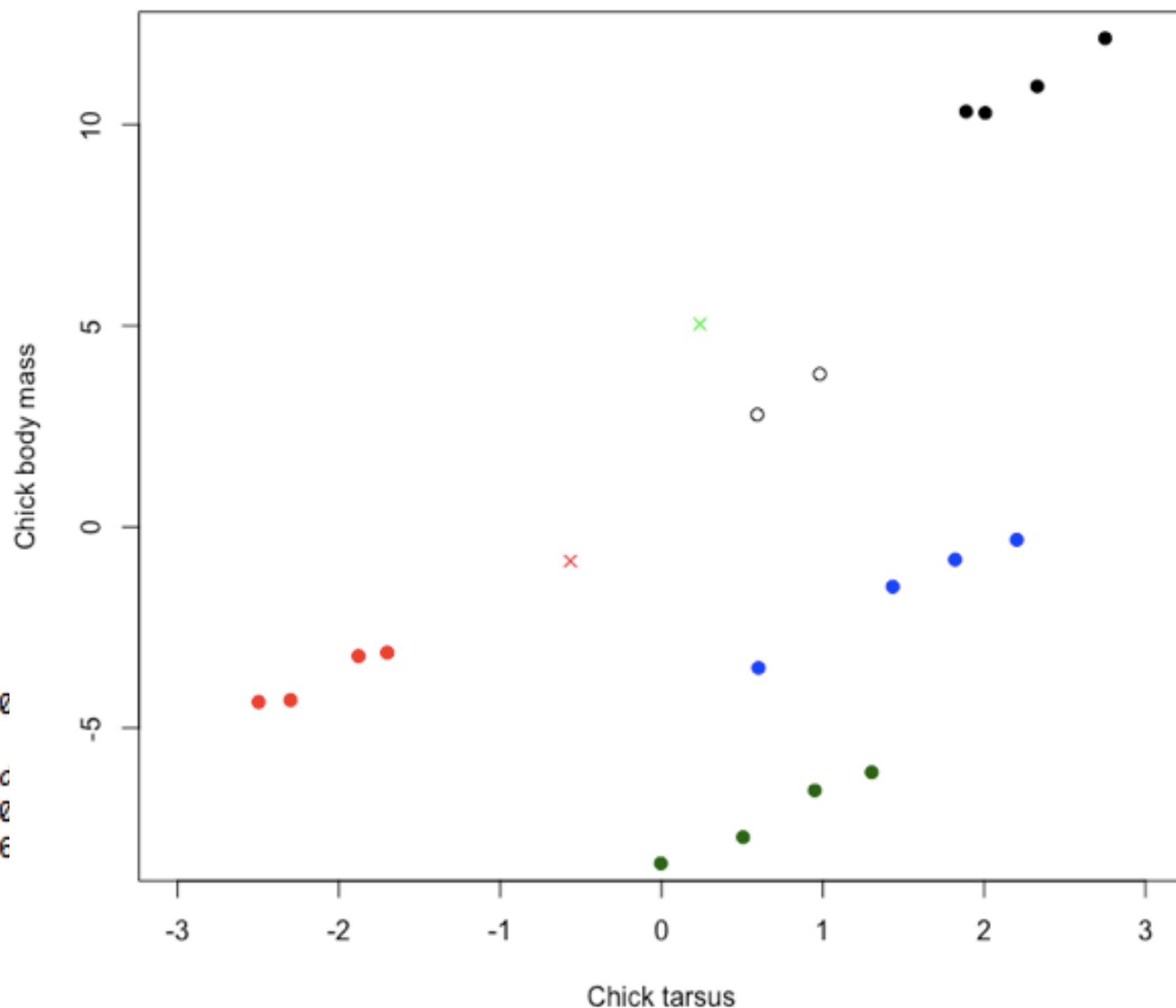
Residual standard error: 0.3312 on 11 degrees of freedom
Multiple R-squared:  0.9983,    Adjusted R-squared:  0.9976
F-statistic: 1581 on 4 and 11 DF,  p-value: 4.275e-15
```

```
Call:
lm(formula = bm ~ tarsus + nest)

Residuals:
    Min       1Q   Median       3Q      Max
-0.1731 -0.1472  0.0000  0.1178  0.2166

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   6.5022     0.2523  25.773 7.10e-12 ***
tarsus         1.9725     0.1043  18.919 2.67e-10 ***
nestB        -6.1191     0.4716 -12.974 2.02e-08 ***
nestC       -11.0200     0.1539 -71.617  < 2e-16 ***
nestD       -15.0427     0.2102 -71.571  < 2e-16 ***
nestE        -4.7637     0.2233 -21.331 6.57e-11 ***
nestF        -6.2356     0.3612 -17.262 7.72e-10 ***
nestG        -1.9313     0.2974  -6.494 2.96e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0

Residual standard error: 0.1892 on 12 degrees of freedom
Multiple R-squared:  0.9995,    Adjusted R-squared:  0
F-statistic: 3230 on 7 and 12 DF,  p-value: < 2.2e-16
```



Mixed models

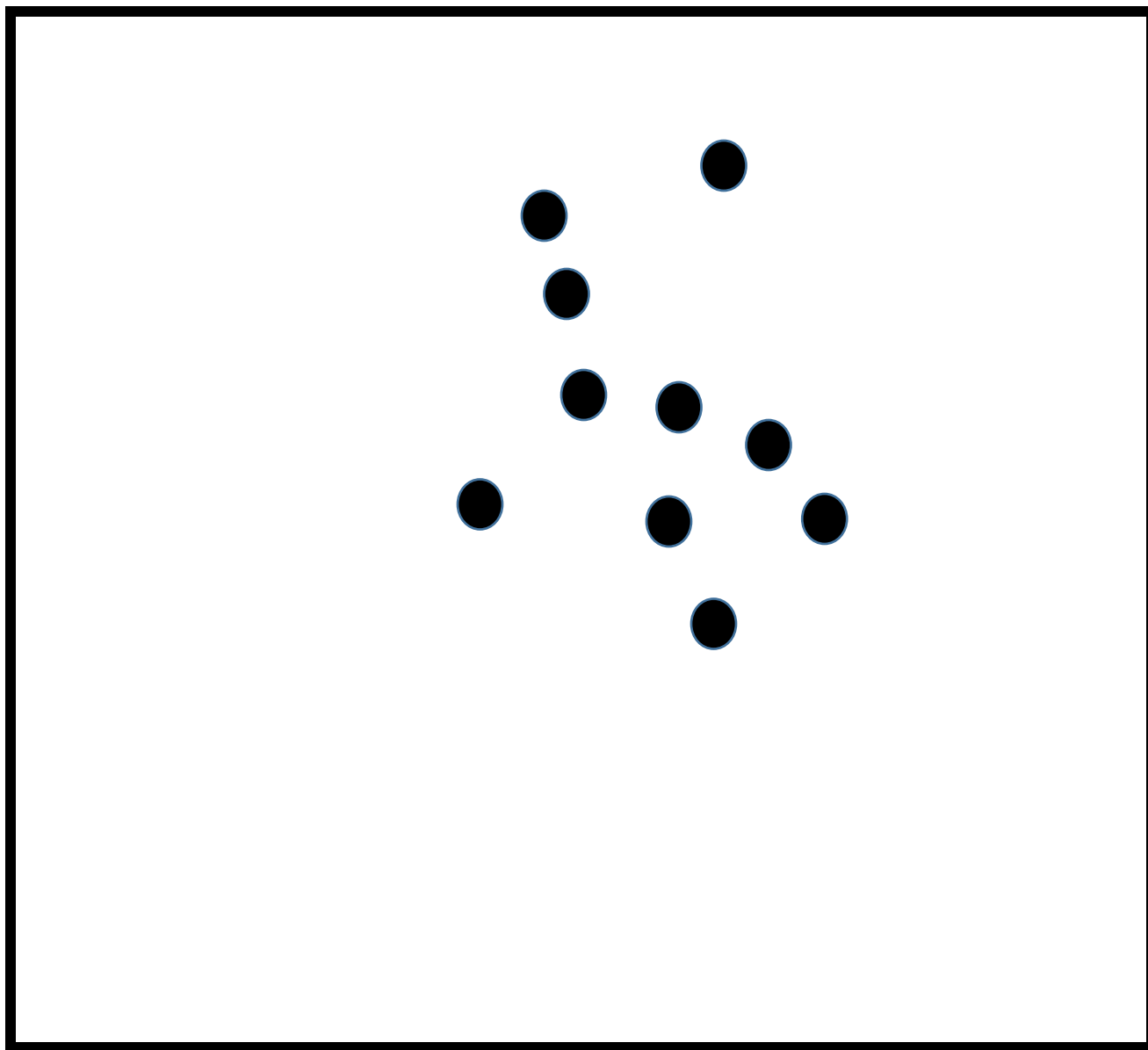
- Estimates variance components simultaneously to fixed terms
- Allow to account for nested structure in data with lots of levels

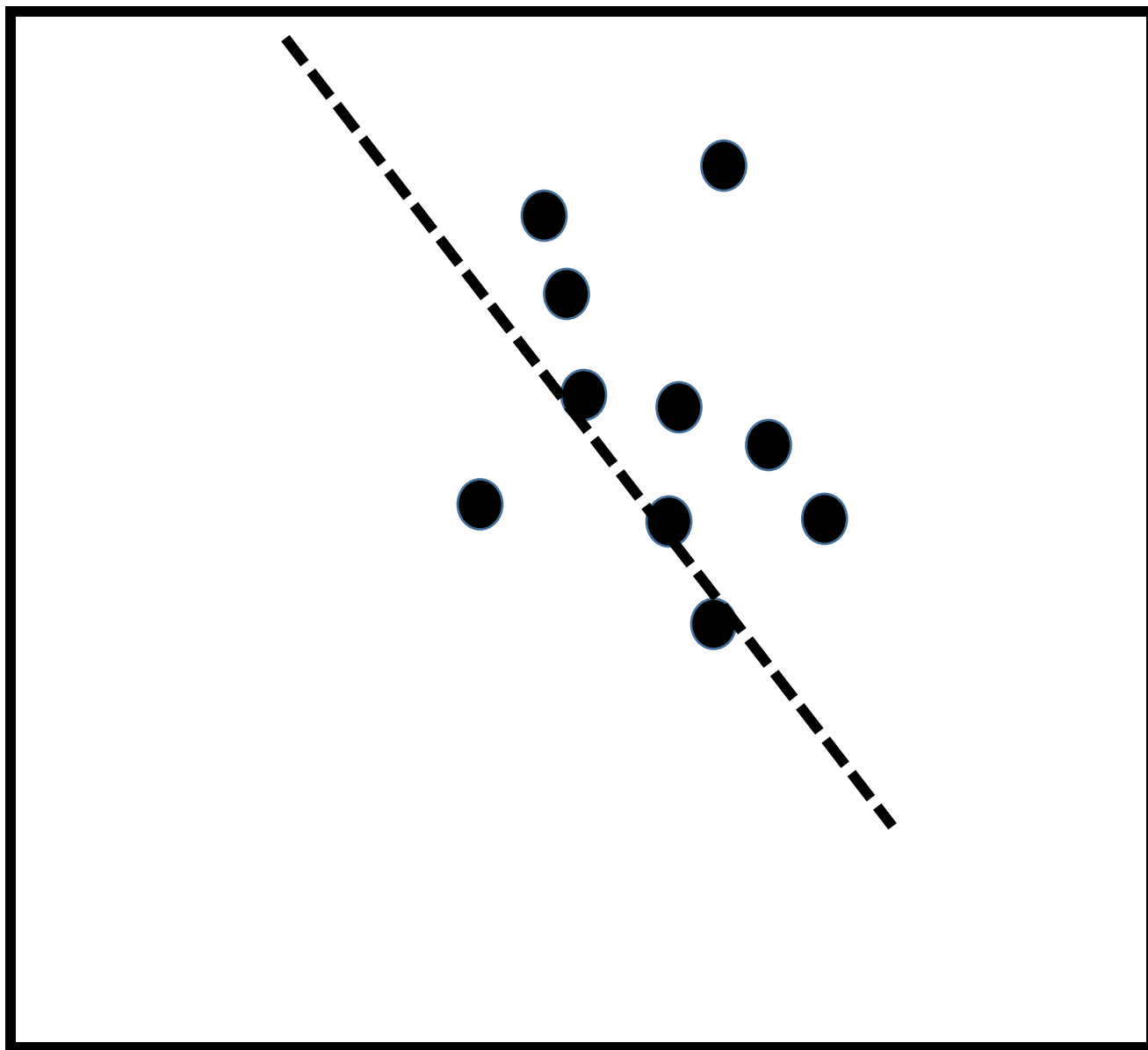
Mixed models

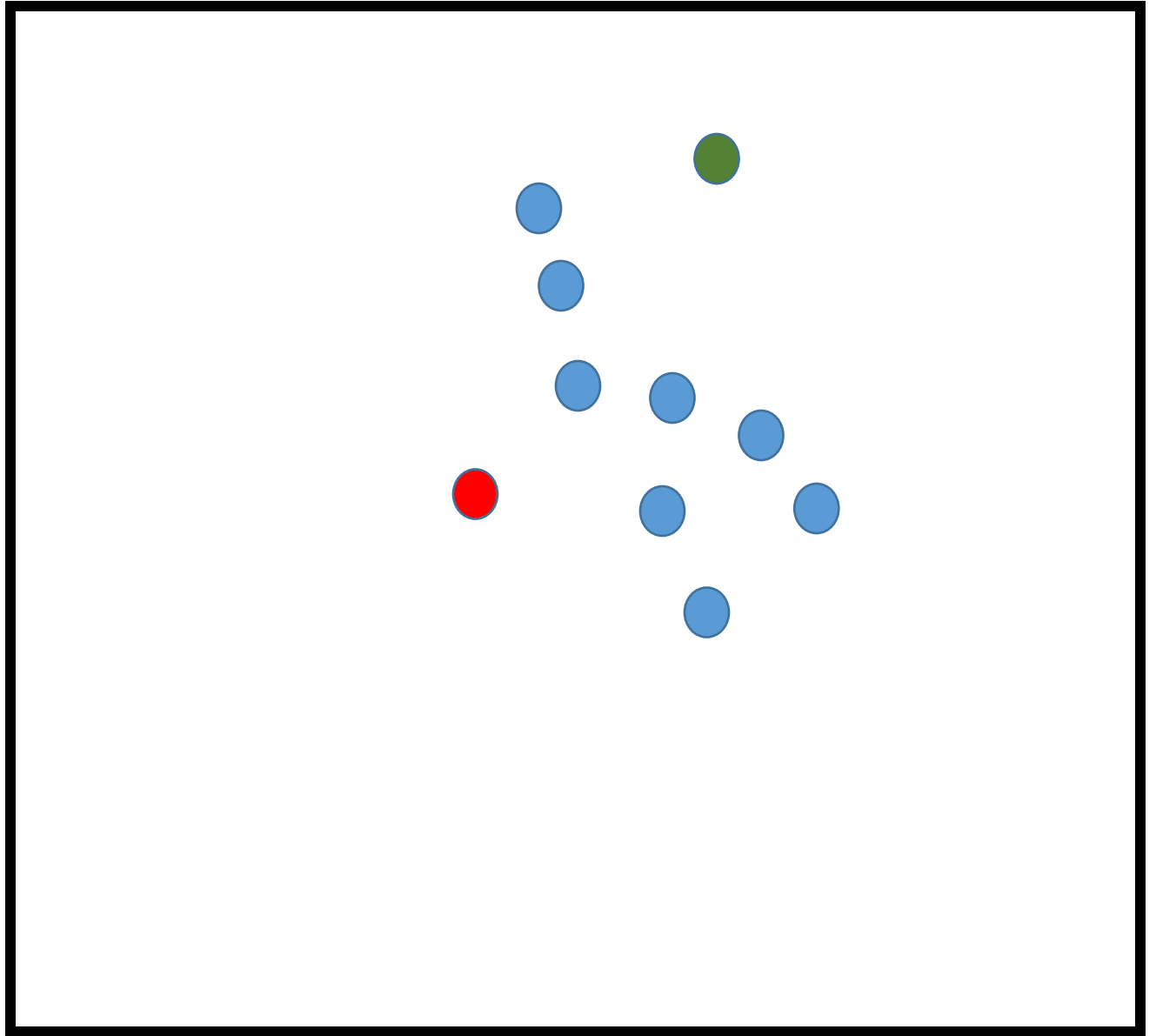
- Estimates variance components simultaneously to fixed terms
- Allow to account for nested structure in data with lots of levels
- Is robust against heterogeneous data

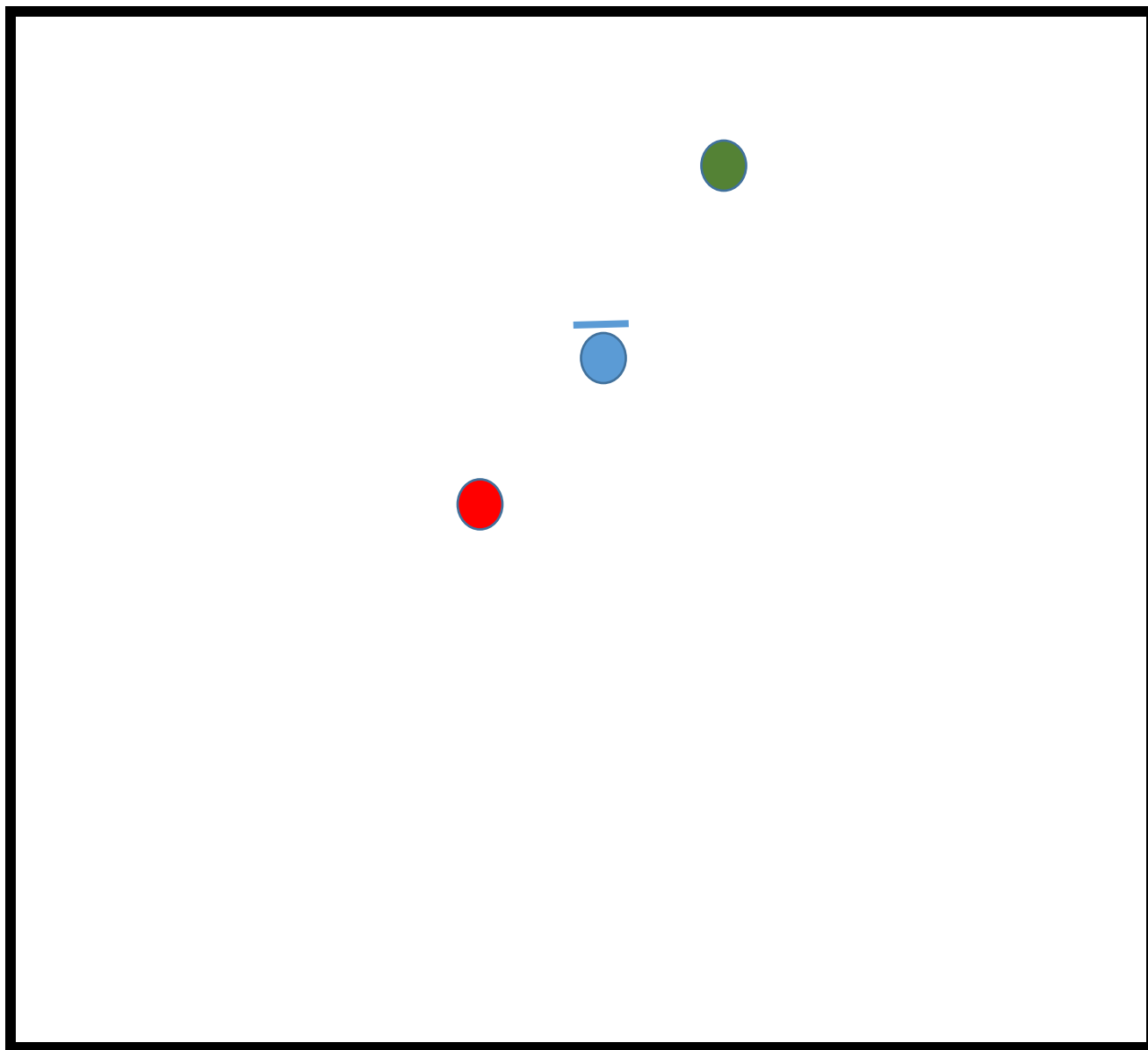
Mixed models

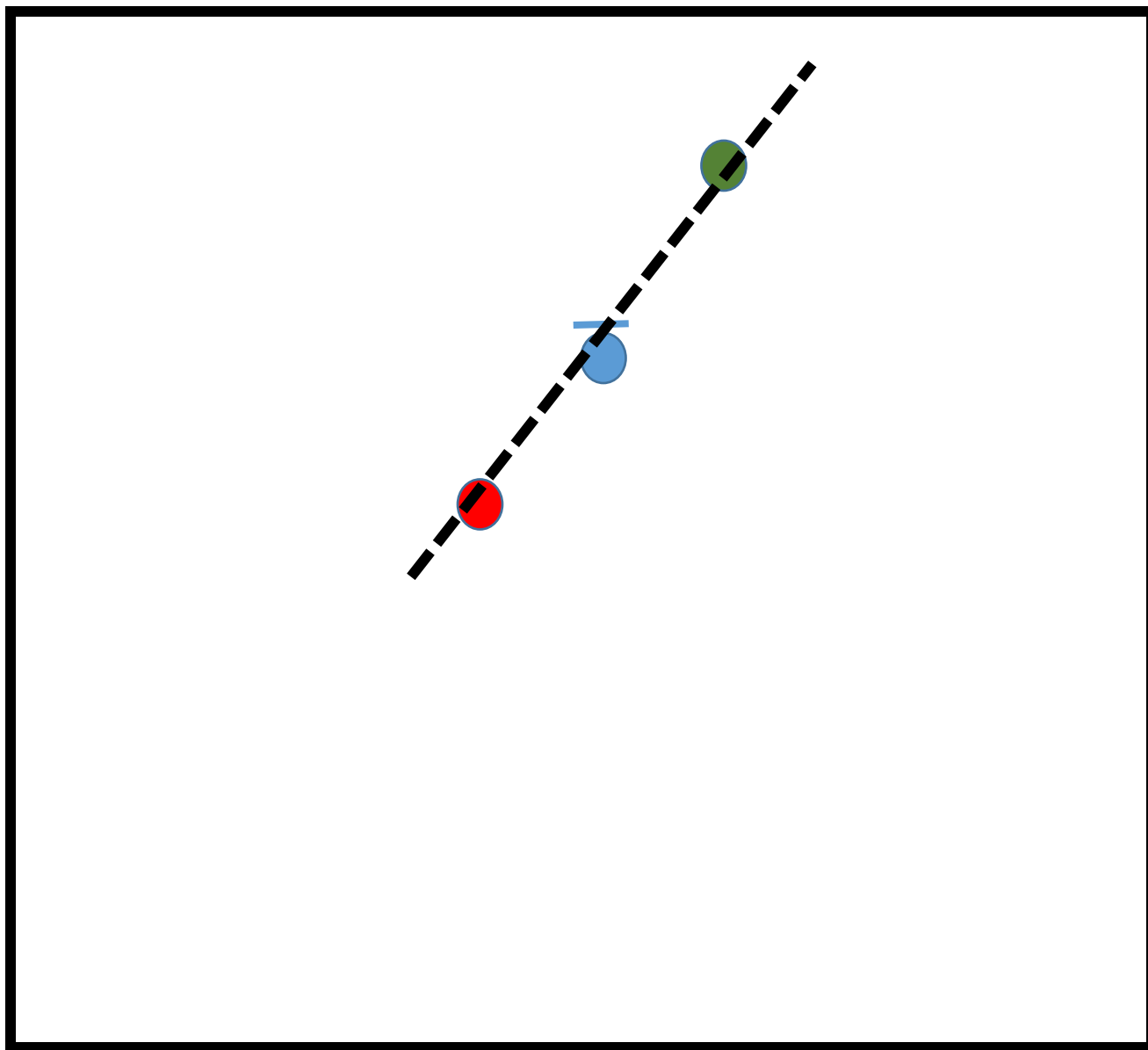
- Estimates variance components simultaneously to fixed terms
- Allow to account for nested structure in data with lots of levels
- Is robust against heterogeneous data
- Allow to account for pseudoreplication

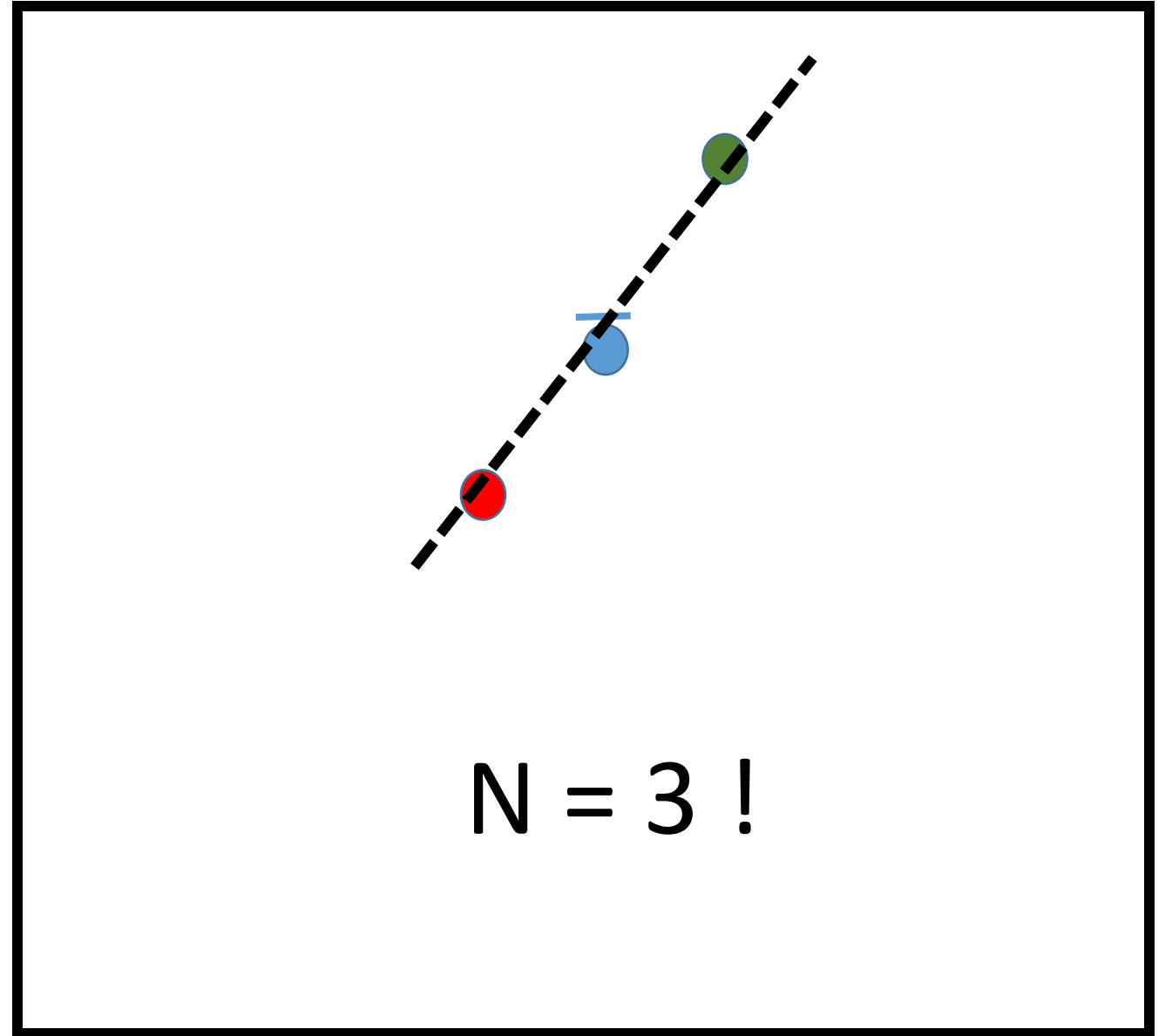


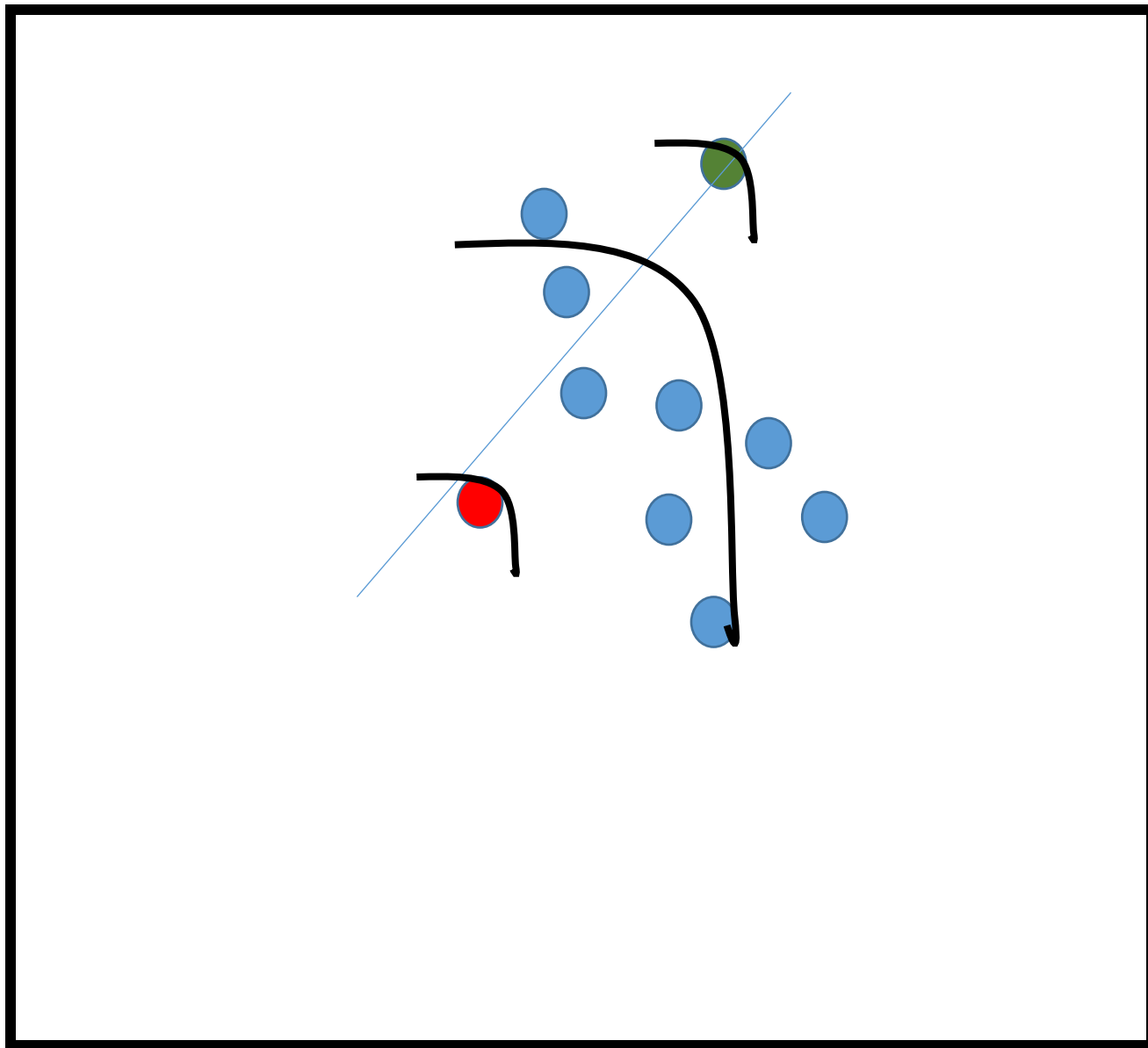












Mixed models

- Estimates variance components simultaneously to fixed terms
- Allow to account for nested structure in data with lots of levels
- Is robust against heterogeneous data
- Allow to account for pseudoreplication

Mixed models

- Estimates variance components simultaneously to fixed terms
 - Allow to account for nested structure in data with lots of levels
 - Is robust against heterogeneous data
 - Allow to account for pseudoreplication
-
- Get information about variance components (e.g. to inform about repeatability)

When is something random factor and when fixed?

- Rules of thumb:
- Random effects are factors!
- Are you interested in means (fixed) or variance (random)?
- Do you want to correct for a factorial effect but it's not in your questions specifically? -> random
- More than 5 levels: random
- LMM use only with large N (>50)

Mixed models – step by step

- Consider your hypothesis!
- Assume you want to know about fixed effects

Mixed models – step by step

- Consider your hypothesis!
- Assume you want to know about fixed effects, and control for pseudo-replication (BirdID) and nested structure (Nest)

Mixed models – step by step

- Consider your hypothesis!
- Assume you want to know about fixed effects, and control for pseudo-replication (BirdID) and nested structure (Nest)
- 1) decide which random effects you want to have in your model

Mixed models – step by step

- Consider your hypothesis!
- Assume you want to know about fixed effects, and control for pseudo-replication (BirdID) and nested structure (Nest)
- 1) decide which random effects you want to have in your model
- 2) decide which fixed effects you want to have

Mixed models – step by step

- Consider your hypothesis!
- Assume you want to know about fixed effects, and control for pseudo-replication (BirdID) and nested structure (Nest)
- 1) decide which random effects you want to have in your model
- 2) decide which fixed effects you want to have
- Run model with all fixed + random effects

Mixed models – step by step

- Consider your hypothesis!
- Assume you want to know about fixed effects, and control for pseudo-replication (BirdID) and nested structure (Nest)
- 1) decide which random effects you want to have in your model
- 2) decide which fixed effects you want to have
- Run model with all fixed + random effects
- Use common sense to reduce model if needed

Mixed models – step by step

- Consider your hypothesis!
- Assume you want to know about fixed effects, and control for pseudo-replication (BirdID) and nested structure (Nest)
- 1) decide which random effects you want to have in your model
- 2) decide which fixed effects you want to have
- Run model with all fixed + random effects
- Use common sense to reduce model if needed
- Use likelihood ratio test or DIC

Likelihood ratio test

- Needs two models, with 1 parameter difference

Likelihood ratio test

- Needs two models, with 1 parameter difference
- Calculates the logL ratio of both models

Likelihood ratio test

- Needs two models, with 1 parameter difference
- Calculates the logL ratio of both models
- Uses chi-square test to test which model is better

Likelihood ratio test

- Needs two models, with 1 parameter difference
- Calculates the logL ratio of both models
- Uses chi-square test to test which model is better

$$2(\log L_{model} - \log L_{reduced})$$

Likelihood ratio test

- Needs two models, with **1 parameter difference**
- Calculates the logL ratio of both models
- Uses chi-square test to test which model is better

$$2(\log L_{\text{model}} - \log L_{\text{reduced}})$$

Df for chi square = 1

LogL in R

```
> m0<-lm(y~1)
```

LogL in R

```
> m0<-lm(y~1)
> m1<-lm(y~x)
```

LogL in R

```
> m0<-lm(y~1)
> m1<-lm(y~x)
> lrtest(m0,m1)
```

LogL in R

```
> m0<-lm(y~1)
> m1<-lm(y~x)
> lrtest(m0,m1)
Likelihood ratio test

Model 1: y ~ 1
Model 2: y ~ x
  #Df  LogLik Df  Chisq Pr(>Chisq)
1    2 -17.488
2    3  -5.684  1 23.609  1.181e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
~
```

Mixed models – step by step

- Consider your hypothesis!
- Assume you want to know whether or not to include random effects

Mixed models – step by step

- Consider your hypothesis!
- Assume you want to know whether or not to include random effects
- Decide on fixed effects structure (best keep it to what's known)

Mixed models – step by step

- Consider your hypothesis!
- Assume you want to know whether or not to include random effects
- Decide on fixed effects structure (best keep it to what's known)
- Add all the random effects

Mixed models – step by step

- Consider your hypothesis!
- Assume you want to know whether or not to include random effects
- Decide on fixed effects structure (best keep it to what's known)
- Add all the random effects
- Compare with reduced model (always only drop one!), use LRT
- OR
- When using Bayesian (MCMCglmm) use BIC (equiv. to AIC)

Mixed models

- Are hard

Mixed models

- Are hard
- Not straightforward

Mixed models

- Are hard
- Not straightforward
- But very useful
- Ask for help

Mixed models

- Are hard
- Not straightforward
- But very useful
- Ask for help

Mixed models - resources

- <http://glmm.wikidot.com>
- Gelman and Hill 2006 Data Analysis Using Regression and Multilevel/Hierarchical Models
- Zuur et al.: Mixed Effects Models and Extensions in Ecology 2009
- Bolker et al. Trends Ecol Evol 2009 Generalized linear mixed models: a practical guide for ecology and evolution

Mixed models

Hand Out files: <https://bit.ly/2tdCTeb>

