

## OHP Picture 1

Recap 2023-11-03

- Goal: Prediction of breeding values

$$y_{ij} = \mu + \text{herd}_j + e_{ij} \quad \left. \right\} \text{linear fixed effect model}$$

i: Tier

j: Herd

Insert data

$$\begin{aligned} y_{12,1} &= \mu + \text{herd}_1 + e_{12,1} \\ y_{26,1} &= \mu + \text{herd}_1 + e_{12,1} \end{aligned}$$

i=12

j=1

- Why mixed linear model and not fixed linear model?

$$y_{ij} = \mu + \text{herd}_j + u_i + e_{ij}$$

if herd<sub>j</sub> and u<sub>i</sub> both  
fix  $\Rightarrow$  fixed linear  
model

herd<sub>j</sub>: fix  
u<sub>i</sub>: random  $\left. \right\} \text{mixed linear model}$

## OHP Picture 2

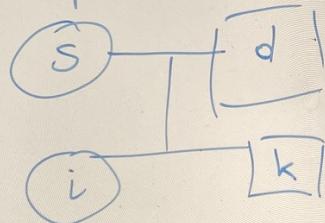
### Breeding Values

$$y_{ij} = \mu + \text{herd}_j + u_i + e_{ij}$$

• Observation: weight of animal i in herd j

•  $u_i$  must be random, because  $\text{cov}(u_i, u_k) \neq 0$   
if animals i and k are related

• Example: animals i and k are full-sibs :



For animal k:  $y_{kl} = \mu + \text{herd}_e + u_j + e_{ke}$

$\text{cov}(u_i, u_k) \neq 0$ , because of decomposition of breeding values

$$\begin{aligned} u_i &= \frac{1}{2} u_s + \frac{1}{2} u_d + m_i \\ u_k &= \frac{1}{2} u_s + \frac{1}{2} u_d + m_k \end{aligned} \quad \left\{ \text{cov}(u_i, u_k) \neq 0 \right.$$

$$\begin{aligned} \text{cov}(u_i, u_k) &= \text{cov}\left(\left[\frac{1}{2} u_s + \frac{1}{2} u_d + m_i\right], \left[\frac{1}{2} u_s + \frac{1}{2} u_d + m_k\right]\right) \\ &= \frac{1}{4} \text{var}(u_s) + \frac{1}{4} \text{var}(u_d) = \frac{1}{2} \sigma_u^2 \end{aligned}$$

provided s and d not related

### OHP Picture 3

(3)

□ Reason 2 why  $u_i$  is random :

- For fixed discrete effects, e.g. herd or breed, the factor level can be assigned to every record
- Breeding Value  $u_i$  is a sum of very many single locus breeding values

$\begin{array}{ c } \hline \bar{G}_i \\ \hline \end{array}$	$H_i$	$K_i$	}
$\begin{array}{ c } \hline G_j \\ \hline \end{array}$	$H_j$	$K_j$	
$BV_{ij}$	$+ BV + BU + \dots = u_i$	too many genotypes ⇒ $u_i$ random	

## OHP Picture 4

(4)

□ Regression

$$y_i = \mu + b \cdot \text{breast\_circumference}_i + e_i$$

herd  
 $u_i$

□ Mixed-linear Model

$$\underline{y_{ij} = \mu + \text{herd}_j + u_i + e_{ij}}$$

- Dataset
  - $y_{12,1} = \mu + \text{herd}_1 + u_{12} + e_{12,1}$
  - $y_{13,1} = \mu + \text{herd}_1 + u_{13} + e_{13,1}$
  - :
- Matrix-vector notation:
  - define vector  $y = \begin{bmatrix} 2.61 \\ 2.31 \\ \vdots \\ 3.16 \end{bmatrix}$
  - define vector  $\beta = \begin{bmatrix} \text{herd}_1 \\ \text{herd}_2 \end{bmatrix}$

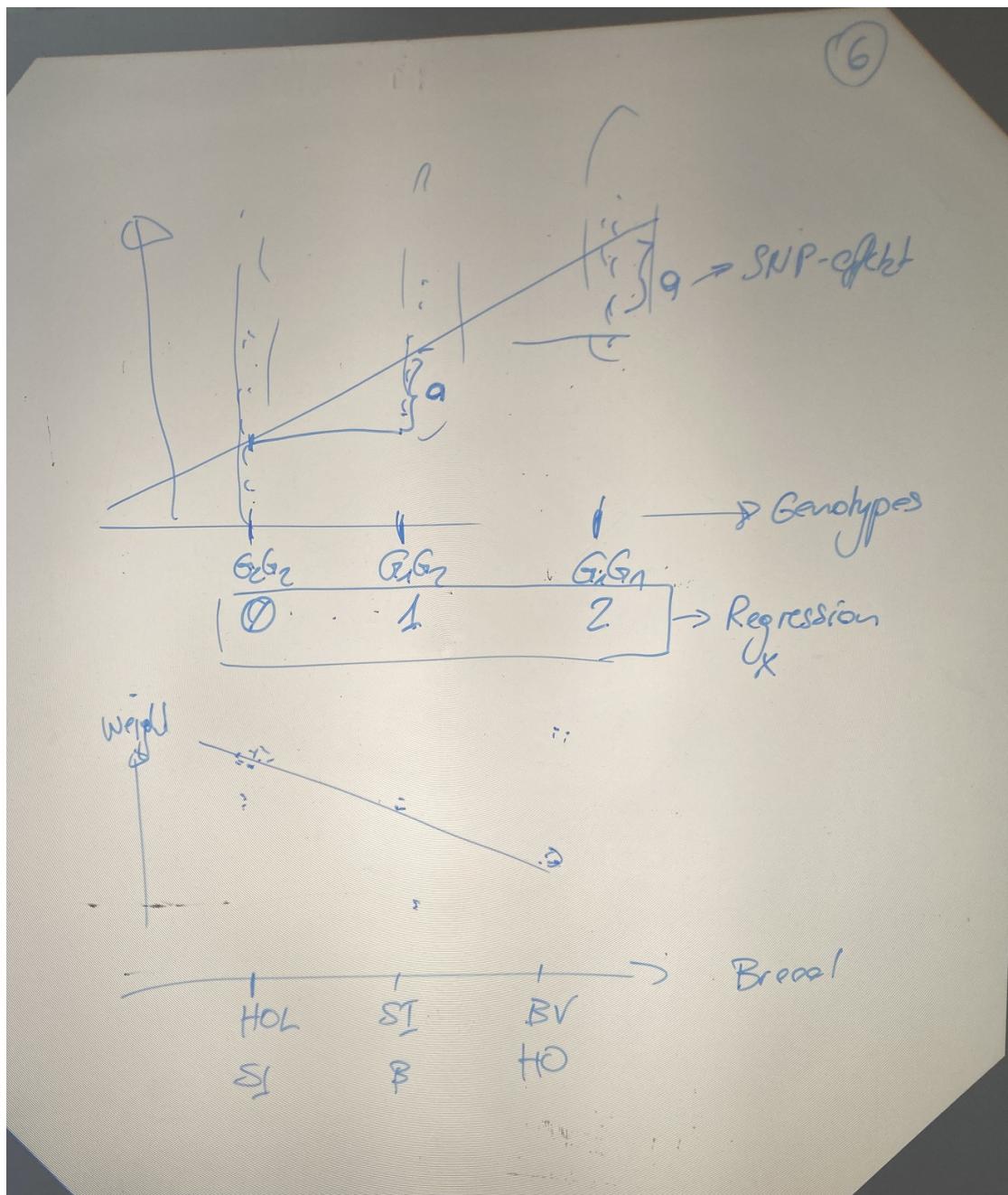
## OHP Picture 5

(5)

- Vector  $y$  : observations
- Vector  $\beta = \begin{bmatrix} \text{herd}_1 \\ \text{herd}_2 \end{bmatrix}$
- Vector  $u = \begin{bmatrix} u_1 \\ \vdots \\ u_q \end{bmatrix}$
- Vector  $e = \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix}$
- Matrices  $X$  and  $Z$  : design matrices
- Model:  $y = X\beta + Zu + e$

$$\rightarrow \begin{bmatrix} y \\ 2.61 \\ 2.31 \\ \vdots \\ 3.16 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \text{herd}_1 \\ \text{herd}_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

OHP Picture 6



## OHP Picture 7

- (1) Sire-model  $\Rightarrow$  only sires get breeding values  
(2) Animal-model  $\Rightarrow$  all animals get breeding values  
in pedigree

$$\text{Model : } y = X\beta + \mathbf{Z}u + e$$

Sire Model :  $u$  - vector of sire-breeding values

Sire Model :  $U = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix}$  with sires 1,2 and 3

- alternative notation:  $y = X\beta + \mathbb{I}s + e$

OHP Picture 8

(8)

Animal Model:  $y = X\beta + Z_u + e$

$$\Rightarrow \begin{bmatrix} y \\ 2.617 \\ \vdots \\ \vdots \end{bmatrix} = \begin{pmatrix} X\beta \\ \vdots \\ \vdots \end{pmatrix} + \begin{bmatrix} Z_u \\ \vdots \\ \vdots \\ \emptyset \end{bmatrix} + \begin{bmatrix} e \\ \vdots \\ \vdots \\ U_{27} \end{bmatrix}$$

- BLUP Method to get predictions for unknown breeding values  $u$  and estimates of fixed effects ( $\beta$ )
  - Model:  $y$ : known;  $X, \beta_0$ : known  
 $\beta, u, (e)$ : unknown
  - Goal: use model and data to get estimates  $\hat{\beta}$  for  $\beta$  and predictions  $\hat{u}$  for  $u$
  - Solution based on BLUP and on model assumptions

OHP Picture 9

Model Assumptions: (must be specified for mixed model) (9)

- Expected values

► vector  $u$  of breeding values, defined as deviations  $\Rightarrow E(u) = \emptyset \Leftrightarrow E(\begin{pmatrix} u \\ e \end{pmatrix}) = \begin{pmatrix} \emptyset \\ \emptyset \end{pmatrix}$

► vector  $e$  of residuals defined as deviations  $\Rightarrow E(e) = \emptyset$

$$\begin{aligned} \Rightarrow E(y) &= E[X\beta + Zu + e] \\ &= E(X\beta) + E(Zu) + \underbrace{E(e)}_{\emptyset} \\ &= X\beta + Z \cdot \underbrace{E(u)}_{\emptyset} \\ &= X\beta \end{aligned}$$

- Abbreviations:

$$E\begin{bmatrix} y \\ u \\ e \end{bmatrix} = \begin{bmatrix} X\beta \\ \emptyset \\ \emptyset \end{bmatrix}$$

- Variance:  $\text{var}(u) = G$ ;  $\text{var}(e) = R$ ;  $\text{var}(y) = V$

where  $G$  and  $R$  are known variance-covariance matrices.

$$\text{cov}(u, e) = \emptyset$$

## OHP Picture 10

(10)

- $\text{var}(u)$  : variance-covariance matrix

$$G = \text{var}(u) = \begin{bmatrix} \text{var}(u_1) & \text{cov}(u_1, u_2) & \text{cov}(u_1, u_3) & \dots \\ \text{cov}(u_2, u_1) & \text{var}(u_2) & \text{cov}(u_2, u_3) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

where  $u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{27} \end{bmatrix}$

- $\text{var}(e) = R = \begin{bmatrix} \text{var}(e_1) & \text{cov}(e_1, e_2) & \dots \\ \text{cov}(e_2, e_1) & \text{var}(e_2) & \dots \end{bmatrix}$
- $\text{var}(e_1) = \text{var}(e_2) = \dots = \bar{\sigma}_e^2$
- $\text{cov}(e_1, e_2) = \text{cov}(e_1, e_3) = \dots = \emptyset$
- $\Rightarrow R = \begin{bmatrix} \bar{\sigma}_e^2 & 0 & \emptyset & \dots & \emptyset \\ 0 & \bar{\sigma}_e^2 & & & \\ \vdots & & & & \end{bmatrix} = I \cdot \bar{\sigma}_e^2$

OHP Picture 11

$$\begin{aligned}
 \text{cov}(y, u) &= \text{cov}(X\beta + Zu + e, u^\top) \\
 &= \underbrace{\text{cov}(X\beta, u^\top)}_{=0} + \text{cov}(Zu, u^\top) + \underbrace{\text{cov}(e, u^\top)}_{=0} \\
 &= \underbrace{\text{cov}(Zu, u^\top)}_{=0} = \underbrace{\sum \text{cov}(u, u^\top)}_{\text{var}(u)} = \prod G
 \end{aligned}$$

$$\begin{aligned}
 \text{cov}(y, e^\top) &= \text{cov}(X\beta + Zu + e, e^\top) \\
 &= \text{cov}(X\beta, e^\top) + \text{cov}(Zu, e^\top) + \text{cov}(e, e^\top) \\
 &= 0 + \underbrace{2\text{cov}(u, e^\top)}_{=0} + \text{var}(e) \\
 &= R
 \end{aligned}$$

$$\text{var} \begin{bmatrix} y \\ u \\ e \end{bmatrix} = \begin{bmatrix} V & ZG & R \\ GZ & G & 0 \\ R & 0 & R \end{bmatrix}, \quad V = ZG^\top + R$$

## OHP Picture 12

(13)

- \* Using properties of BLUP:  
solutions for  $\hat{u}$  and  $\hat{\beta}$ :
$$\hat{u} = G Z^T V^{-1} (y - X \hat{\beta})$$

$$\hat{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} y$$

\* Practical Problem: Large data sets ( $10^7$  Mio records)  
 $V^{-1}$ : cannot be computed      ( $6 \cdot 10^7$  Mio)

□ Same solutions from system of equations:

$$\begin{bmatrix} X^T R^{-1} X & X^T R^{-1} B \\ Z^T R^{-1} X & Z^T R^{-1} Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X^T R^{-1} y \\ Z^T R^{-1} y \end{bmatrix}$$

-  $R = I / \sigma^2 \Rightarrow R^{-1} = I \cdot \frac{1}{\sigma^2}$