

OHP Picture 1

Recap: Prediction of Breeding Values

□ Own performance

| Animal | Phenotype Weight |
|--------|------------------|
| 1      | 275 kg           |
| ⋮      | ⋮                |
| N      | 312 kg           |

► Predicted breeding value  $\hat{u}_i$  for animal  $i$

$$\hat{u}_i = h^2 (y_i - \mu)$$

Annotations:  $h^2$  → heritability,  $(y_i - \mu)$  → measurement - population mean

► Accuracy:

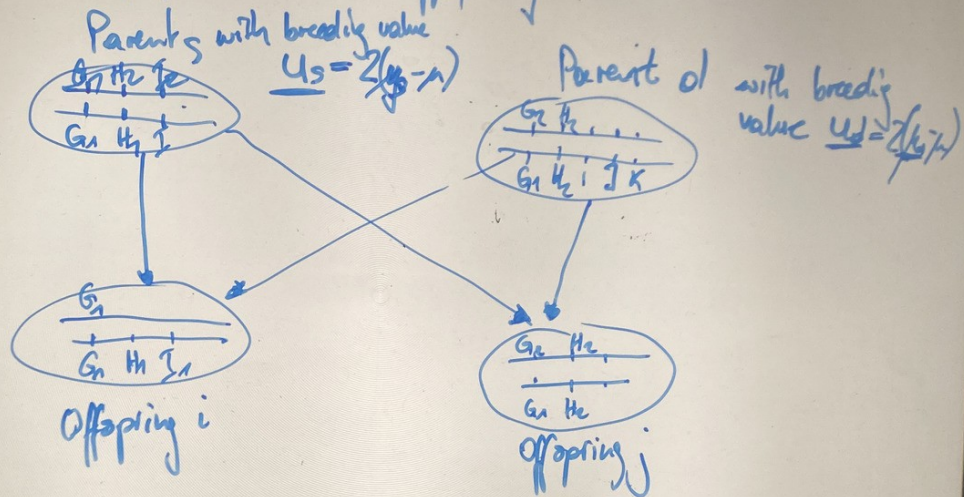
Genauigkeit

$$r(u_i, y_i) = \frac{\text{cov}(u_i, y_i)}{\sqrt{\text{var}(u_i) \cdot \text{var}(y_i)}} = h$$

Reliability:  $r^2(u_i, y_i) = h^2$   
(Bestimmtheitskoeffizient)

### Decomposition

□ Relationship between breeding values of parents and offspring



$$\begin{aligned}
 u_i &= \frac{1}{2} U_s + \frac{1}{2} U_d + m_i \\
 u_j &= \frac{1}{2} U_s + \frac{1}{2} U_d + m_j
 \end{aligned}$$

→ mendelian sampling

OHP Picture 3

Selection response

- Assume: Selection on  $\hat{u}_i$
- What is expected selection response per generation?

- $R = i \cdot r_{yy}^z \cdot \sigma_y$

$R$ : Selection response  
 $i$ : selection intensity  
 $r_{yy}^z$ : Reliability  
 $\sigma_y$ : phenotypic standard deviation =  $\sqrt{\text{var}(y)}$

- For Breeding programs:  $\left| \frac{R}{L} \right|$  where  $L$  = Generation interval  
 selection response per year

## Pros and Cons of Own Performance

Pro: - simple to compute

Con: - only animals with phenotypic records get predicted breeding values  
 ⇒ sex-limited traits (dairy)  
 - Not taking all information into account.  
 e.g. the same trait recorded several times.

### Repeated Records:

| Animal | Weight<br>t1 | t2  | t3 | ... | tm | mean<br>$\bar{y}_i = \frac{1}{m} \sum_{k=1}^m y_{ik}$ |
|--------|--------------|-----|----|-----|----|---|
| 1      | 225          | 277 |    |     |    | $y_{i1}$  |
| 2      | 305          | 305 |    |     |    | $y_{i2}$  |
| ⋮      |              |     |    |     |    |   |
| n      |              |     |    |     |    | $y_{in}$  |

$$\hat{u}_i = b_r(\bar{y}_i - \mu)$$

## Pros and Cons of Own Performance

- Pro: - simple to compute
- Con: - only animals with phenotypic records get predicted breeding values  
 ⇒ sex-limited traits (dairy)
- Not taking all information into account.  
 e.g. the same trait recorded several times.

### Repeated Records:

| Animal | Weight | t1 | t2  | t3 | ... | t <sub>m</sub> | mean  |
|--------|--------|----|-----|----|-----|----------------|---|
| 1      | 275    |    | 277 |    |     |                | $\bar{y}_i = \frac{1}{m} \sum_{k=1}^m y_{ik}$ |
| 2      | 305    |    | 305 |    |     |                |   |
| ...    |        |    |     |    |     |                |   |
| n      |        |    |     |    |     |                | $\bar{y}_n$                                   |

Annotations: Red arrows between t1 and t2, and t2 and t3 are labeled "different → t<sub>k</sub> different". Green arrows between 275 and 277, and 305 and 305 are labeled "const" and "cons". A green arrow labeled "pe" points from the 277 value to the 305 value.

$$\hat{u}_i = b_r(\bar{y}_i - \mu)$$

Regression Coefficient b<sub>r</sub>

OHP Picture 6

Regression Coefficient  $b_r$

$$b_r = \frac{\text{Cov}(u, \tilde{y})}{\text{var}(\tilde{y})}$$

□ For  $\text{var}(\tilde{y})$ , we have an additional variance component:

$\text{var}(y)$  using for  $y = \mu + u + e \rightarrow$  temp env (te)  
permanent environment (pe)

$$\begin{aligned}\text{var}(y) &= \text{var}(u + u + e) \\ &= \text{var}(\mu + u + pe + te) \\ &= \underbrace{\text{var}(u)}_{=0} + \text{var}(u) + \text{var}(pe) + \text{var}(te) \\ &= \text{var}(u) + \text{var}(pe) + \text{var}(te)\end{aligned}$$

□ Define Repeatability  $t$

$$t = \frac{\text{var}(u) + \text{var}(pe)}{\text{var}(y)} = \frac{\sigma_u^2 + \sigma_{pe}^2}{\sigma_y^2}$$

Predicted Breeding Value  $\hat{u}_i$  :

$$\hat{u}_i = b_r (\bar{y}_i - \mu)$$

$$b_r = \frac{\text{Cov}(u, \bar{y})}{\text{Var}(\bar{y})}$$

$$= \frac{\text{Cov}(u_i, \mu + u_i + p_{e_i} + \frac{1}{m} \sum_{j=1}^m t_{e_{ij}})}{\text{Var}(\bar{y})}$$

$$= \frac{\underbrace{\text{Cov}(u_i, \mu)}_{=0} + \underbrace{\text{Cov}(u_i, u_i)}_{=0} + \underbrace{\text{Cov}(u_i, p_{e_i})}_{=0} + \frac{1}{m} \sum_{j=1}^m \text{Cov}(u_i, t_{e_{ij}})}{\text{Var}(\bar{y})}$$

$$= \frac{\text{Cov}(u_i, u_i)}{\text{Var}(\bar{y})} = \frac{\text{Var}(u_i)}{\text{Var}(\bar{y})}$$

$$\text{Var}(\bar{y}) = \text{Var}\left(\mu + u_i + p_{e_i} + \frac{1}{m} \sum_{j=1}^m t_{e_{ij}}\right)$$

$$= \text{Var}(\mu) + \text{Var}(u_i) + \text{Var}(p_{e_i}) + \frac{1}{m^2} \sum_{j=1}^m \text{Var}(t_{e_{ij}})$$

$$= \text{Var}(u_i) + \text{Var}(p_{e_i}) + \frac{1}{m} \text{Var}(t_{e_i})$$

- Assumption:  $\text{Cov}(u, p_e) = \text{Cov}(u, t_e) = \text{Cov}(t_{e_{ij}}, t_{e_{ik}}) = 0$

- Use  $t = \frac{\text{Var}(u) + \text{Var}(p_e)}{\text{Var}(y)} \rightarrow t \cdot \text{Var}(y)$

$\bar{y}_i = \frac{1}{m} \sum_{j=1}^m y_{ji}$   
 $= \frac{1}{m} \sum_{j=1}^m (\mu + u_i + p_{e_i} + t_{e_{ij}})$   
 $= \frac{1}{m} \sum_{j=1}^m \mu + \frac{1}{m} \sum_{j=1}^m u_i + \frac{1}{m} \sum_{j=1}^m p_{e_i} + \frac{1}{m} \sum_{j=1}^m t_{e_{ij}}$   
 $= \mu + u_i + p_{e_i} + \frac{1}{m} \sum_{j=1}^m t_{e_{ij}}$

OHP Picture 8

$$b_r = \frac{\text{Cor}(u, \bar{y})}{\text{Var}(\bar{y})} = \frac{\text{var}(u)}{[t + \frac{1}{m}(1-t)] \text{var}(y)} =$$

$$= \frac{m \cdot \text{var}(u)}{[mt + (1-t)] \text{var}(y)} = \frac{m \cdot \text{var}(u)}{[(m-1)t + 1] \text{var}(y)}$$

$$= \frac{m \cdot h^2}{1 + (m-1)t}$$

$$\Rightarrow \hat{u}_i = \frac{m h^2}{1 + (m-1)t} (\bar{y}_i - \mu)$$


---

Progeny Records : Dairy Cattle  $\hat{u}_i = \frac{m h^2}{1 + (m-1)t}$   $\begin{matrix} \circ \\ \downarrow \\ \circ \end{matrix} \rightarrow y_m$

| Animal | offspring 1 | 2        | 3 | ... | n        | $\bar{y}_i$                                   |
|--------|-------------|----------|---|-----|----------|---|
| 1      | $y_{11}$    | $y_{12}$ | - | -   | $y_{1n}$ | $\bar{y}_1 = \frac{1}{n} \sum_{k=1}^n y_{1k}$ |
| 2      |             |          |   |     |          | $\bar{y}_2$                                   |
| ...    |             |          |   |     |          |   |
| N      |             |          |   |     |          | $\bar{y}_N$                                   |



$\begin{matrix} \bar{y}_2 \\ \vdots \\ \bar{y}_N \end{matrix}$

$N$

---

Goal: Predict  $\hat{u}_i$  based on  $\bar{y}_i$

$$\hat{u}_i = b (\bar{y}_i - \mu)$$

$$b = \frac{\text{Cov}(u, \bar{y})}{\text{Var}(\bar{y})}$$


---

- What is  $\bar{y}_i$ ?  $\bar{y}_i = \frac{1}{n} \sum_{k=1}^n y_{ik}$

- Use  $y_{ik} = \mu + u_k + e_k$   
 $= \mu + \frac{1}{2} u_i + \frac{1}{2} u_k + m_k + e_k$

$$\bar{y}_i = \frac{1}{n} \sum_{k=1}^n \left( \mu + \frac{1}{2} u_i + \frac{1}{2} u_k + m_k + e_k \right)$$

$$\mu + \frac{1}{2} u_i + \frac{1}{2n} \sum_{k=1}^n u_k + \frac{1}{n} \sum_{k=1}^n m_k + \frac{1}{n} \sum_{k=1}^n e_k$$

$$\text{Cov}(u_i, \bar{y}_i) = \text{Cov}\left(u_i, \left[ \mu + \frac{1}{2} u_i + \frac{1}{2n} \sum_{k=1}^n u_k + \frac{1}{n} \sum_{k=1}^n m_k + \frac{1}{n} \sum_{k=1}^n e_k \right]\right)$$

$$\begin{aligned}
 \text{cov}(u_i, \bar{y}_i) &= \text{cov}\left(u_i, \left[\mu + \frac{1}{2}u_i + \frac{1}{2n} \sum_{k=1}^n u_{ik} + \frac{1}{n} \sum_{k=1}^n w_k + \frac{1}{n} \sum_{k=1}^n e_k\right]\right) \\
 &= \underbrace{\text{cov}(u_i, \mu)}_{=0} + \text{cov}\left(u_i, \frac{1}{2}u_i\right) + \text{cov}\left(u_i, \frac{1}{2n} \sum_{k=1}^n u_{ik}\right) \\
 &\quad + \underbrace{\text{cov}\left(u_i, \frac{1}{n} \sum_{k=1}^n w_k\right)}_{=0} + \underbrace{\text{cov}\left(u_i, \frac{1}{n} \sum_{k=1}^n e_k\right)}_{=0} \\
 &= \text{cov}\left(u_i, \frac{1}{2}u_i\right) = \frac{1}{2} \text{cov}(u_i, u_i) = \frac{1}{2} \text{var}(u)
 \end{aligned}$$


---


$$\begin{aligned}
 \text{var}(\bar{y}_i) &= \text{var}\left(\mu + \frac{1}{2}u_i + \frac{1}{2n} \sum_{k=1}^n u_{ik} + \frac{1}{n} \sum_{k=1}^n w_k + \frac{1}{n} \sum_{k=1}^n e_k\right) \\
 &= \underbrace{\text{var}(\mu)}_{=0} + \text{var}\left(\frac{1}{2}u_i\right) \\
 &\quad + \text{var}\left(\frac{1}{2n} \sum_{k=1}^n u_{ik}\right) + \text{var}\left(\frac{1}{n} \sum_{k=1}^n w_k\right) \\
 &\quad + \text{var}\left(\frac{1}{n} \sum_{k=1}^n e_k\right) \\
 &= \boxed{\frac{1}{4} \text{var}(u_i)} + \text{var}\left(\frac{1}{2n} \sum_{k=1}^n u_{ik}\right) + \text{var}\left(\frac{1}{n} \sum_{k=1}^n w_k\right) \\
 &\quad + \text{var}\left(\frac{1}{n} \sum_{k=1}^n e_k\right)
 \end{aligned}$$

permanent

$$\begin{aligned}
 & \left[ t \cdot \text{var}(y) + \frac{1}{n}(1-t) \cdot \text{var}(y) \right] = \left[ \frac{t + (1-t)}{n} \right] \text{var}(y) \\
 t &= \frac{\frac{1}{4} \text{var}(u_i)}{\text{var}(y)} = \frac{h^2}{4} \quad \left[ \frac{1}{2} \cdot h^2 \cdot \text{var}(y) \right]
 \end{aligned}$$

OHP Picture 11

$$\begin{aligned}
 b &= \frac{\text{Cov}(u_i, \bar{y}_i)}{\text{Var}(\bar{y}_i)} = \frac{\frac{1}{2} \cdot \text{Var}(u_i)}{\left[ t + (1-t)/n \right] \cdot \text{Var}(y)} \\
 &= \frac{\frac{1}{2} h^2 \cdot \text{Var}(y)}{\left[ \frac{h^2}{4} + (1 - \frac{h^2}{4})/n \right] \cdot \text{Var}(y)} \\
 &= \frac{\frac{1}{2} h^2}{\frac{h^2}{4} + (1 - \frac{h^2}{4})/n} \\
 &= \frac{\frac{1}{2} n h^2}{\frac{nh^2}{4} + (1 - \frac{h^2}{4})} = \frac{2nh^2}{nh^2 + (4-h^2)} \\
 &= \frac{2n}{n + (4-h^2)/h^2} = \frac{2n}{n+k} \\
 &\quad k = \frac{4-h^2}{h^2} \\
 \hat{u}_i &= \frac{2n}{n+k} (\bar{y}_i - \mu)
 \end{aligned}$$

Goal: Find method that is able to use all available data and to predict breeding values for all animals in the population

Solution: Use BLUP

B: Best, i.e.  $\text{var}(u - \hat{u})$  is minimal  
prediction error variance

L: Linear, linear function of  $y$  as a predictor  $\hat{u}$  of true breeding values  $u$

U: unbiased,  $E(\hat{u}) = E(u)$

P: Prediction (Zuchtwertschätzung)  
(In English: • Prediction for random effect  
• Estimation for fixed effect)

• Estimation for fixed effect

Method for Prediction:

- Use a linear mixed effects model (Lme)
- Linear: Observations ( $y$ ) are taken as response variables in our statistical model, responses are modelled as linear function of unknown parameters
- Mixed: Model contains both fixed and random effects

estimates via best  
sevars

Fixed effect: • Example data, Herd is a fixed effect. disorder

- Regression covariate are also considered fixed, e.g. breast circumference

Random effects: Breeding values are random

Linear Mixed Effects Model as a ~~generalisation~~ generalisation of a fixed effect model:

Fixed effects:

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

$\mu$ : weaning weight of animal  $i$  in herd  $j$   
 $\text{herd}_j$ : fixed effect of herd  $j$  on weaning weight  
 $e_{ij}$ : random residual

Mixed effects:

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

$u_i$ : breeding value of animal  $i$  as random effect.

Insert information from data into model:

$$\begin{array}{l}
 y_{12,1,1}^{2.61} = \mu + \text{herd}_1 + u_{12} + e_{12,1,1} \\
 y_{13,1,1}^{2.51} = \mu + \text{herd}_1 + u_{13} + e_{13,1,1} \\
 \vdots \\
 y_{27,2,1}^{2.16} = \mu + \text{herd}_2 + u_{27} + e_{27,2,1}
 \end{array}$$

$\mu$ : unknown  
 $\text{herd}_1, \text{herd}_2$ : unknown  
 $u_{12}, u_{13}, u_{27}$ : known  
 $e_{12,1,1}, e_{13,1,1}, e_{27,2,1}$ : unknown