

OHP Picture 1

- Recap: Prediction of Breeding Values
- Simple scenario: Own performance records (Eigenleistung)
 - Every animal i has one observation (y_i) of a phenotypic value of a trait of interest
 - birth weight
 - mastitis resistance
 - fertility

Animals	Observations y_i (numbers)
1	$y_1 = 52$
2	$y_2 = 48$
:	:
N	

- Predicted breeding value \hat{u}_i of animal i
- $\hat{u}_i = h^2(y_i - \mu)$ where h^2 : heritability of trait

μ : population mean

- Predictions (\hat{u}_i) are associated with errors:
 - Quantification of prediction error: accuracy r_{pred}
 - Reliability $r_{\text{pred}}^2 \rightarrow 8\%$ (Bestimmtheitsmaß)

OHP Picture 2

Accuracy of predicted breeding values is also important when quantifying the response to selection (R)

$$R = i \cdot \text{Tau}_g^2 \cdot \bar{y}_g = i \cdot h^2 \cdot F_g \rightarrow \text{phenotypic standard deviation}$$

\downarrow
selection intensity
 \downarrow
own performance
 \downarrow
Selection response per generation.

Breeders Equation

Repeated Records

More than one observation per animal

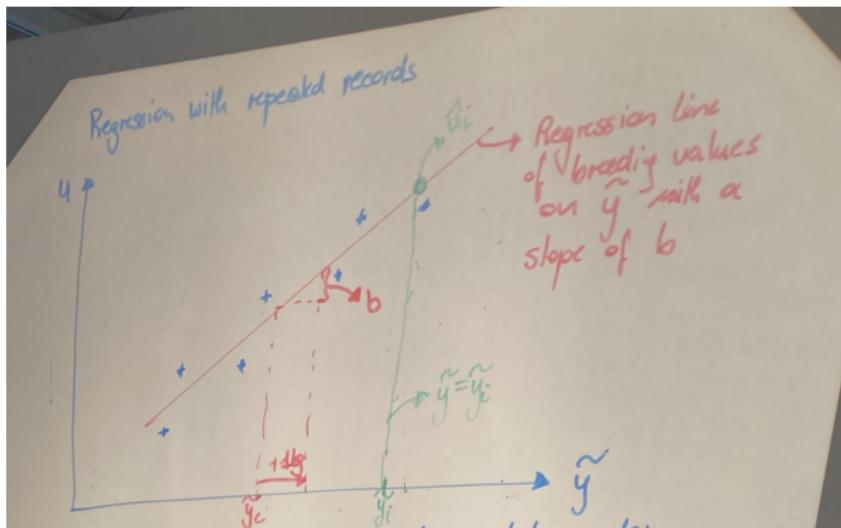
Animal i	Observation 1	Observation 2	... Observation k	Mean
1	$y_{i1} = 152$	$y_{i2} = 181$...	$y_{ik} = 545$ $\left(\sum_{j=1}^k y_{ij} \right) / k$
2				
⋮				
N				

$$\tilde{y}_i = \frac{1}{k} \sum_{j=1}^k y_{ij}$$

mean

Animal i	\tilde{y}_i
1	\tilde{y}_1
2	\tilde{y}_2
⋮	\tilde{y}_N

OHP Picture 3



- Prediction based on regression with repeated records:

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

$$b = \frac{\text{cor}(u, \tilde{y})}{\text{var}(\tilde{y})}$$

population mean

Prediction with a new value of \tilde{y}_i : search value on red regression line that intersects with the ordinate $\tilde{y} = \tilde{y}_i$

OHP Picture 4

Computation of regression coefficient $b = \frac{\text{cov}(y_i, g)}{\text{var}(g)}$

□ Genetic model:

$$y_{ij} = \mu + u_i + e_{ij}$$

□ Decompose e_{ij} into a permanent part and a temporary part
changes between different observations for animal i

$$e_{ij} = p_{ei} + t_{ej} \quad \text{constant over all observations for animal } i$$

$$\Rightarrow y_{ij} = \mu + u_i + p_{ei} + t_{ej}$$

□ Decomposition at the level of variance:

$$\text{var}(y_{ij}) = \text{var}(\mu + u_i + p_{ei} + t_{ej})$$

$$\begin{aligned} & \text{phenotypic variance of observations} \\ & = \text{var}(\mu) + \text{var}(u_i) + \text{var}(p_{ei}) + \text{var}(t_{ej}) \\ & = \bar{y}_j^2 - \left\{ + 2\text{cov}(\mu, u_i) + 2\text{cov}(\mu, p_{ei}) + 2\text{cov}(u_i, t_{ej}) \right\} + \dots \end{aligned}$$

$$= \underbrace{\text{var}(u_i) + \text{var}(p_{ei}) + \text{var}(t_{ej})}_{\text{genetic additive variance}} = \bar{y}_j^2$$

OHP Picture 5

Decomposition of \bar{y}_j^2 for repeated observations:

$$\text{var}(y_{ij}) = \underbrace{\text{var}(u_i)}_{\text{constant across all observations}} + \text{var}(p_{ei}) + \text{var}(t_{eij})$$

- Repeatability t : Tells us the ratio of variance components that are permanent compared to the total variance of all observations

$$t = \frac{\text{var}(u_i) + \text{var}(p_{ei})}{\text{var}(y_{ij})} \Rightarrow \underbrace{\text{var}(u_i) + \text{var}(p_{ei})}_{\text{constant}} = t \cdot \bar{y}_j^2$$

$$1-t = \frac{\text{var}(y_{ij})}{\text{var}(y_{ij})} - \frac{\text{var}(u_i) + \text{var}(p_{ei})}{\text{var}(y_{ij})}$$

$$= \frac{\text{var}(y_{ij}) - \text{var}(u_i) - \text{var}(p_{ei})}{\text{var}(y_{ij})} = \frac{\text{var}(t_{eij})}{\underbrace{\text{var}(y_{ij})}_{\Downarrow}}$$

$$\underbrace{\text{var}(t_{eij})}_{\Downarrow} = (1-t)\bar{y}_j^2$$

OHP Picture 6

Decomposition of \bar{y}_j^2 for repeated observations:

$$\text{var}(y_{ij}) = \underbrace{\text{var}(u_i)}_{\text{constant across all observations for animal } i} + \text{var}(p_{ei}) + \text{var}(e_{ij})$$

- Repeatability t : Tells us the ratio of variance components that are permanent compared to the total variance of all observations

$$t = \frac{\text{var}(u_i) + \text{var}(p_{ei})}{\text{var}(y_{ij})} \Rightarrow \underbrace{\text{var}(u_i) + \text{var}(p_{ei})}_{\text{var}(e_{ij})} = t \cdot \bar{y}_j^2$$

$$1-t = \frac{\text{var}(y_{ij})}{\text{var}(y_j)} - \frac{\text{var}(u_i) + \text{var}(p_{ei})}{\text{var}(y_j)}$$

$$= \frac{\text{var}(y_{ij}) - \text{var}(u_i) - \text{var}(p_{ei})}{\text{var}(y_j)} = \underbrace{\frac{\text{var}(e_{ij})}{\text{var}(y_j)}}_{\Downarrow}$$

$$\underbrace{\text{var}(e_{ij})}_{\text{var}(e_{ij})} = (1-t)\bar{y}_j^2$$

OHP Picture 7

Computation of b_i :

$$\text{cov}(u_i, \tilde{y}_i) = \text{cov}(u_i, [\mu + u_i + t_{e,i} + \frac{1}{k} \sum_{j=1}^k t_{e,j}])$$
$$= \frac{1}{k} \sum_{j=1}^k u_{ij} = \frac{1}{k} \sum_{j=1}^k [\mu + u_i + p_{e,i} + t_{e,j}]$$

\downarrow genetic model

$$= \frac{1}{k} \sum_{j=1}^k \mu + \frac{1}{k} \sum_{j=1}^k u_i + \frac{1}{k} \sum_{j=1}^k p_{e,i} + \frac{1}{k} \sum_{j=1}^k t_{e,j}$$
$$\frac{1}{k} \cdot k \cdot \mu = \mu$$

$$\text{cov}(u_i, \tilde{y}_i) = \underbrace{\text{cov}(u_i, \mu)}_{=0} + \underbrace{\text{cov}(u_i, u_i)}_{=0} + \underbrace{\text{cov}(u_i, p_{e,i})}_{=0}$$

$$+ \text{cov}(u_i, \frac{1}{k} \sum_{j=1}^k t_{e,j})$$
$$= 0$$

$$= \text{cov}(u_i, u_i) = \text{var}(u_i) = \bar{u}^2$$

genetic additive variance

OHP Picture 8

$$\begin{aligned}\text{var}(\tilde{y}_i) &= \text{var}\left(\frac{1}{k} \sum_{j=1}^k (\mu + u_i + p_{ij} + t e_{ij})\right) \\ &= \text{var}(\mu + u_i + p_{ij} + \frac{1}{k} \sum_{j=1}^k t e_{ij}) \\ &= \underbrace{\text{var}(\mu)}_0 + \text{var}(u_i) + \text{var}(p_{ij}) \\ &\quad + \text{var}\left(\frac{1}{k} \sum_{j=1}^k t e_{ij}\right) + \underbrace{2 \text{cov}(\mu, u_i) + \dots}_{=0}\end{aligned}$$

$$t \cdot \bar{e}_j^2 = \text{var}(u_i) + \text{var}(p_{ij}) + \frac{1}{k} \text{var}(t e_{ij}) \quad (1-t) \bar{F}_y^2$$

using repeatability relations

$$\Rightarrow \text{var}(\tilde{y}_i) = t \cdot \bar{e}_j^2 + \frac{1}{k} (1-t) \bar{F}_y^2$$

$$= \frac{1}{k} [k \cdot t + (1-t)] \bar{F}_y^2$$

$$= \frac{1 + (k-1)t}{k} \bar{F}_y^2$$

$$\begin{aligned}\Rightarrow \text{Regression coefficient } b &= \frac{\text{cov}(u_i, \tilde{y}_i)}{\text{var}(u_i, \tilde{y}_i)} = \frac{\bar{F}_u^2}{\frac{1 + (k-1)t}{k} \bar{F}_y^2} \\ &= \frac{k \bar{F}_u^2}{1 + (k-1)t \bar{F}_y^2} = \frac{k b_u^2}{1 + (k-1)t}\end{aligned}$$

OHP Picture 9

$$\begin{aligned}\text{var}(\tilde{y}_{ij}) &= \text{var}\left(\frac{1}{k} \sum_{j=1}^k \left[\mu + u_i + p_{ei} + t e_{ij}\right]\right) \\ &= \text{var}\left(\mu + u_i + p_{ei} + \frac{1}{k} \sum_{j=1}^k t e_{ij}\right) \\ &= \underbrace{\text{var}(\mu)}_0 + \text{var}(u_i) + \text{var}(p_{ei}) \\ &\quad + \text{var}\left(\frac{1}{k} \sum_{j=1}^k t e_{ij}\right) + \underbrace{2\text{cov}(\mu, u_i)}_{=0} + \underbrace{2\text{cov}(u_i, t e_{ij})}_{=0} \\ &= \text{var}(u_i) + \text{var}(p_{ei}) + \frac{1}{k} \text{var}(t e_{ij})\end{aligned}$$

using repeatability relations

$$\begin{aligned}\Rightarrow \text{var}(\tilde{y}_i) &= t \cdot \bar{v}_y^2 + \frac{1}{k} (1-t) \bar{v}_y^2 \\ &\quad - \frac{1}{k} [k \cdot t + (1-t)] \bar{v}_y^2 \\ &= \frac{1 + (k-1)t}{k} \bar{v}_y^2\end{aligned}$$

$$\begin{aligned}\Rightarrow \text{Regression coefficient } b &= \frac{\text{cov}(u_i, \tilde{y}_i)}{\text{var}(u_i, \tilde{y}_i)} = \frac{\bar{v}_u^2}{\frac{1 + (k-1)t}{k} \bar{v}_y^2} \\ &= \frac{k \bar{v}_u^2}{1 + (k-1)t \bar{v}_y^2} = \frac{k h^2}{1 + (k-1)t}\end{aligned}$$

OHP Picture 10

1. Own Performance
2. Repeated Records
3. Progeny Records } same animal, limited use in
comes from breeding programs in
dairy cattle

daughters - half-sibs

Dataset			
Animal i (Bull)	offspring 1	offspring 2	... offspring k
1	y_{11}		
2			
:			
N			\bar{y}_N

Average across offspring

\bar{y}_1
 \bar{y}_2

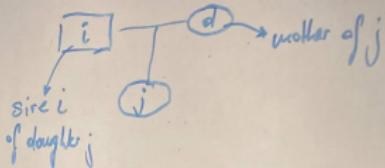
- Regression to predict breeding values:
- Predicted breeding value of animal i :
- $\hat{u}_i = b(\bar{y}_i - \mu)$; $b = \frac{\text{cov}(u_i, \bar{y}_i)}{\text{var}(\bar{y}_i)}$
- where $\bar{y}_i = \frac{1}{k} \sum_{j=1}^k y_{ij}$

OHP Picture 11

Decomposition of $\bar{y}_i = \frac{1}{k} \sum_{j=1}^k y_{ij}$

- Genetic model: $y_{ij} = \mu + u_j + e_j$

- Pedigree



- Breeding value u_j of daughter j :

$$u_j = \frac{1}{2} u_i + \frac{1}{2} u_d + m_j$$

→ Insert into genetic model:

$$y_{ij} = \mu + \frac{1}{2} u_i + \frac{1}{2} u_d + m_j + e_j ; \text{ do that for all daughters of sire } i$$

$$\rightarrow \bar{y}_i = \frac{1}{k} \sum_{j=1}^k [\mu + \frac{1}{2} u_i + \frac{1}{2} u_d + m_j + e_j]$$

$$= \mu + \frac{1}{2} u_i + \frac{1}{k} \sum_{j=1}^k \frac{1}{2} u_d + \frac{1}{k} \sum_{j=1}^k m_j + \frac{1}{k} \sum_{j=1}^k e_j //$$

OHP Picture 12

$$\begin{aligned}
 * \text{cov}(u_i, \bar{y}_i) &= \text{cov}(u_i, \mu + \frac{1}{2}u_i + \frac{1}{k} \sum_{j=1}^k u_{ij} + \frac{1}{k} \sum_{j=1}^k w_j) \\
 &\quad + \frac{1}{k} \sum_{j=1}^k e_j \\
 &= \underbrace{\text{cov}(u_i, \mu)}_{=0} + \text{cov}(u_i, \frac{1}{2}u_i) + \text{cov}(u_i, \frac{1}{k} \sum_{j=1}^k u_{ij}) \\
 &\quad + \underbrace{\text{cov}(u_i, \frac{1}{k} \sum_{j=1}^k w_j)}_{=0} + \underbrace{\text{cov}(u_i, \frac{1}{k} \sum_{j=1}^k e_j)}_{=0}
 \end{aligned}$$

Assume that
sire i and
daughters j are
unrelated

$$-\frac{1}{2} \text{cov}(u_i, u_i) = \frac{1}{2} \text{var}(u_i) = \frac{1}{2} \bar{u}^2$$

$$\Rightarrow \text{cov}(u_i, u_{ij}) = 0$$

$$\text{var}(\bar{y}_i) : \text{using } \bar{y}_i = \mu + \frac{1}{2}u_i + \frac{1}{k} \sum_{j=1}^k u_{ij} + \frac{1}{k} \sum_{j=1}^k w_j + \frac{1}{k} \sum_{j=1}^k e_j$$

$$\text{var}(\bar{y}_i) = \text{var}\left(\frac{1}{2}u_i\right) + \text{var}\left(\frac{1}{k} \sum_{j=1}^k \frac{1}{2}u_{ij}\right) + \text{var}\left(\frac{1}{k} \sum_{j=1}^k w_j\right) + \text{var}\left(\frac{1}{k} \sum_{j=1}^k e_j\right)$$

For a given sire i:
permanent part: $\text{var}\left(\frac{1}{2}u_i\right)$
across all daughters

OHP Picture 13

Computation with Variances:

Random variable X (continuous)

$$E[X] = \int x f(x) dx \quad \text{with } f(x) \text{ density of } X$$

$$\text{Var}[X] = \int (x - E(x))^2 f(x) dx$$

$$\text{Var}[a \cdot X] = \int a^2 (x - E(x))^2 f(x) dx$$

$$= a^2 \int (x - E(x))^2 f(x) dx$$

$$= a^2 \cdot \text{Var}(x)$$

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{cov}(X, Y)$$

$$\text{cov}(X, Y+Z) = \text{cov}(X, Y) + \text{cov}(X, Z)$$

OHP Picture 14

$$\begin{aligned}\text{var}(\bar{y}_i) &= \text{var}\left(\frac{1}{k} u_i\right) + \text{var}\left(\frac{1}{k} \sum_{j=1}^k y_j u_{ij}\right) \\ &\quad + \text{var}\left(\frac{1}{k} \sum_{j=1}^k w_j\right) + \text{var}\left(\frac{1}{k} \sum_{j=1}^k e_j\right) \\ &= \frac{1}{4} \text{var}(u_i) + \frac{1}{k^2} \text{var}\left(\sum_{j=1}^k y_j u_{ij}\right) \\ &= t \cdot \bar{u}_i^2 + \frac{1}{k} (1-t) \bar{y}_i^2 \quad \text{with } t = \frac{h^2}{4}\end{aligned}$$

Regression coefficient:

$$b = \frac{\text{cov}(u_i, \bar{y}_i)}{\text{var}(\bar{y}_i)} = \frac{\frac{1}{2} \bar{u}_i^2}{\left(t + \frac{(1-t)}{k}\right) \bar{y}_i^2}$$

$$= \frac{\frac{1}{2} h^2 \bar{y}_i^2}{\left(\frac{h^2}{4} + (1 - \frac{h^2}{4})\right) \bar{y}_i^2}$$

$$= \frac{\frac{1}{2} h^2 k}{\frac{h^2}{4} + (1 - \frac{h^2}{4})} = \frac{2kh^2}{k h^2 + (4-h^2)}$$

$$= \frac{2k}{k + (4-h^2)/h^2} = \frac{2k}{k + \beta} \quad \text{with } \beta = \frac{4-h^2}{h^2}$$

$$\hat{u}_i = b(\bar{y}_i - u) = \frac{2k}{k+\beta} (\bar{y}_i - u)$$

OHP Picture 15

$$\begin{aligned}\text{var}(\bar{y}_i) &= \text{var}\left(\frac{1}{k} u_i\right) + \text{var}\left(\frac{1}{k} \sum_{j=1}^k y_j u_{ij}\right) \\ &\quad + \text{var}\left(\frac{1}{k} \sum_{j=1}^k u_j\right) + \text{var}\left(\frac{1}{k} \sum_{j=1}^k e_j\right) \\ &= \frac{1}{k^2} \text{var}(u_i) + \frac{1}{k^2} \text{var}\left(\sum_{j=1}^k y_j u_{ij}\right) \\ &= t \cdot \bar{u}_i^2 + \frac{1}{k} (1-t) \bar{y}^2 \quad \text{with } t = \frac{h^2}{4}\end{aligned}$$

Regression coefficient:

$$\begin{aligned}b &= \frac{\text{cov}(u, \bar{y}_i)}{\text{var}(\bar{y}_i)} = \frac{\frac{1}{2} \bar{u}_i^2}{\left(t + \frac{(1-t)}{k}\right) \bar{y}^2} \\ &= \frac{\frac{1}{2} h^2 \bar{u}_i^2}{\left(\frac{h^2}{4} + (1 - \frac{h^2}{4})\right) / k} \bar{y}^2 \\ &= \frac{\frac{1}{2} h^2 k}{\frac{h^2}{4} + (1 - \frac{h^2}{4})} = \frac{\frac{1}{2} k h^2}{\frac{h^2}{4} + (4 - h^2)} \\ &= \frac{2k}{k + (4-h^2)/h^2} = \frac{2k}{k + \beta} \quad \text{with } \beta = \frac{4-h^2}{h^2} \\ \hat{u}_i &= b (\bar{y}_i - \mu) = \frac{2k}{k+\beta} (\bar{y}_i - \mu)\end{aligned}$$

OHP Picture 16

Reading Data into R:

- traditional : `read.table(...)` \rightarrow `data.frame`
- new : `readr package : readcsv()`
function

tibble
(modern `data frame`)

- writing data to a file

```
readr::write_csv(tbl_weight, file = "weight.data.csv")
```

\rightarrow column separator by `";"`
decimal delimiter : `"."` } US-style

column sep by `";"`
decimal : `,`