# Deterministic Models for EBV

Jack Dekkers

## Accuracy of Selection

**The Breeder's Equation for Response to Selection:**

$$\Delta G = i \, r \, \sigma_g$$

i = selection intensity $\quad\quad$ $\sigma_g$ = genetic St. Deviation

r = accuracy of selection

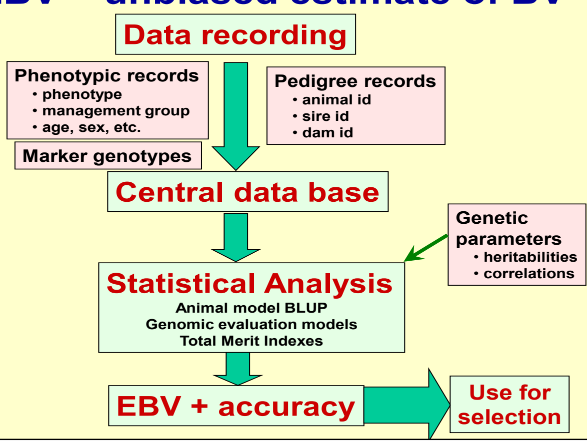
= corr(selection criterion, true breeding value)

= accuracy of EBV for single trait selection

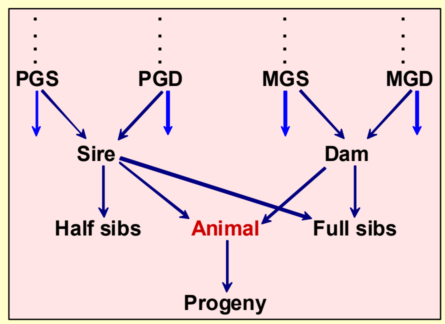
or of total merit index for multi-trait selection

- maximize to increase gain
- need to be able to model accuracy to be able to predict genetic gain and to determine the value of information

## EBV = unbiased estimate of BV




## Animal model EBV

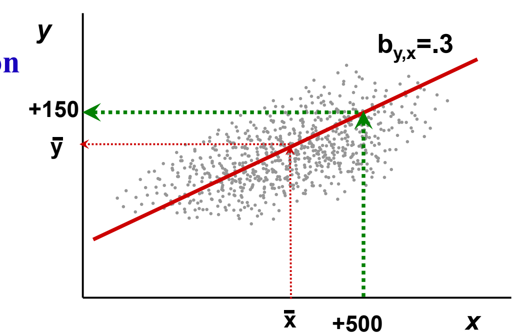



## Methods to Model Accuracy of EBV

1) EBV from own records
   - simple regression
2) EBV from records on a single type of relatives
   - simple regression
3) EBV from multiple sources of information
   - multiple regression - selection index
4) EBV from BLUP animal model

All derivations assume records are perfectly adjusted for systematic environmental effects (herd, season, sex, age, etc.)

**Linear regression theory:**




Model: $y = a + b_{yx} x + e$

Prediction: $\hat{y}_i = \bar{y} + b_{yx}(x_i - \bar{x})$

Regression coefficient: $b_{yx} = \dfrac{\sigma_{xy}}{\sigma_x^2} = r_{xy}\dfrac{\sigma_y}{\sigma_x}$

Accuracy of prediction: $r_{xy} = \dfrac{\sigma_{xy}}{\sqrt{\sigma_y^2 \sigma_x^2}}$

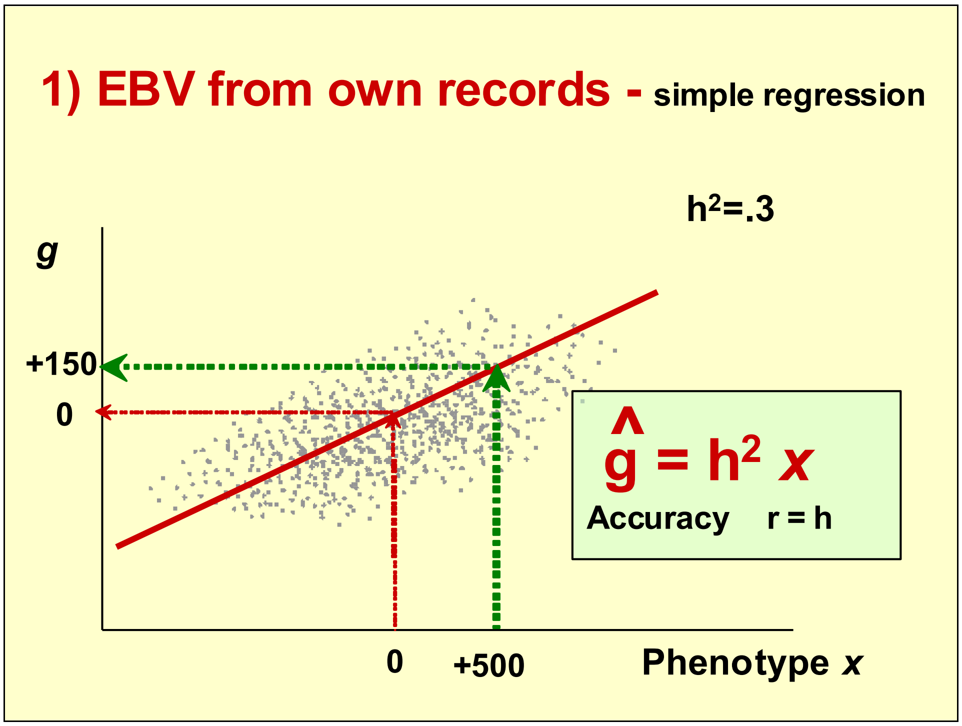
# 1) EBV from own records ($x$)

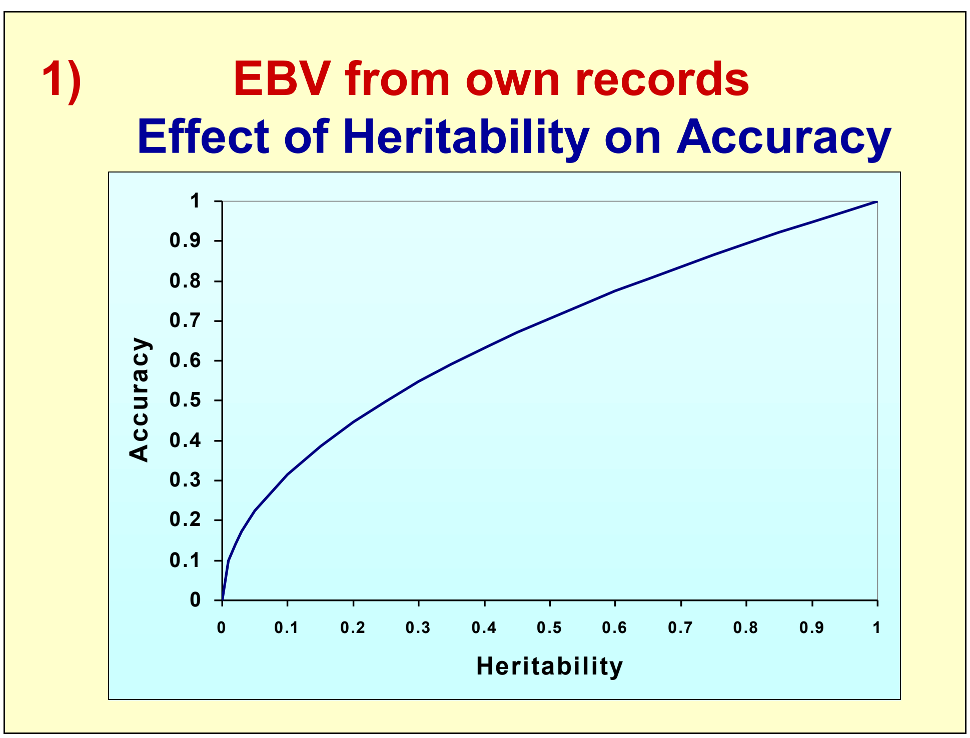
$$\hat{g}_i = b_{g,x}\, x_i = b_{g,x}\,(\text{phenotype of individual})$$

$$b_{g,x} = \sigma_{x_i g_i}/\sigma_p^2 = \sigma_{g_i+e_i, g_i}/\sigma_p^2 = \sigma_g^2/\sigma_p^2 = h^2$$

$$\hat{g}_i = h^2 x_i$$

$$\text{Accuracy} = r = r_{g,\hat{g}} = \sigma_{g_i,(h^2 x_i)}/\sigma_g \sigma_{h^2 x} = h$$



## EBV based on the Mean of Two or more Phenotypic Records

**Definition of Repeatability**

Repeated records on same individual: $x = g + pe + te$

$pe$ = permanent environment effect

$te$ = temporary environment effect

**Repeatability, $t$** = prop. of total phenotypic variance that is due to permanent effects (envir. + genetic)

$$t = \frac{\sigma_g^2 + \sigma_{pe}^2}{\sigma_p^2} \quad \text{or} \quad \frac{\sigma_g^2 + \sigma_{pe}^2}{\sigma_g^2 + \sigma_{pe}^2 + \sigma_{te}^2}$$

Cow $i$ has two lactation records, $x_{i1}$ and $x_{i2}$

$$x_{i1} = g_i + pe_i + te_{i1}$$
$$x_{i2} = g_i + pe_i + te_{i2}$$

Correlation between records on an individual is

$$r_{x_1 x_2} = \frac{\sigma_{x_1 x_2}}{\sqrt{\sigma_{x_1}^2 \sigma_{x_2}^2}}$$

and

$$\sigma_{x_1 x_2} = \sigma_{(g_i + pe_i + te_{i1},\, g_i + pe_i + te_{i2})} = \sigma_g^2 + \sigma_{pe}^2$$

Hence,

$$r_{x_1 x_2} = \frac{\sigma_g^2 + \sigma_{pe}^2}{\sigma_p^2} = t$$

**EBV from Repeated Records on a Single Trait**

Select on mean of $m$ records: $\hat{g}_i = b_{g.\bar{x}}\, \bar{x}_i$ where $\bar{x}_i = \sum_{j=1}^{m} x_{ij} / m = \sum_{j=1}^{m} (g_i + pe_i + te_{ij}) / m$

Then,

$$b_{g.\bar{x}} = \sigma_{g.\bar{x}} / \sigma_{\bar{x}}^2$$

The variance of $\bar{x}_i$ is:

$$\sigma_{\bar{x}}^2 = \sigma_g^2 + \sigma_{pe}^2 + \frac{\sigma_{te}^2}{m} = t\sigma_p^2 + \frac{(1-t)\sigma_p^2}{m} = \frac{(mt + 1 - t)\sigma_p^2}{m}$$
$$= \frac{((m-1)t + 1)\sigma_p^2}{m}$$

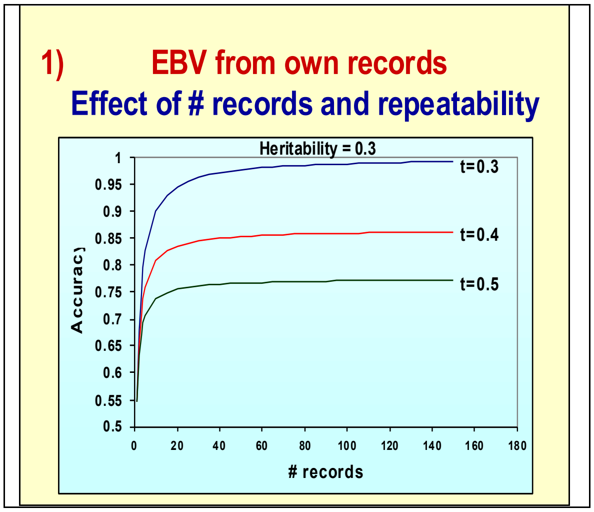
The covariance is: $\sigma_{g.\bar{x}} = \sigma_g^2$

Thus,

$$b_{g.\bar{x}} = \frac{m\sigma_g^2}{\sigma_p^2((m-1)t+1)} = \frac{mh^2}{(m-1)t+1}$$

Accuracy is:

$$r = \sqrt{\frac{mh^2}{(m-1)t+1} \frac{\sigma_g^2}{\sigma_g^2}} = \sqrt{\frac{mh^2}{(m-1)t+1}}$$

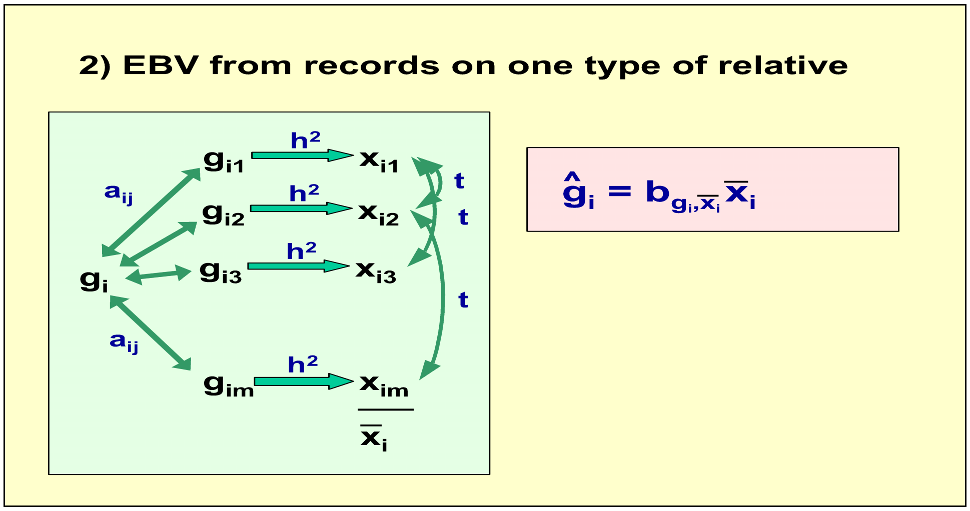


## 2) EBV from One Type of Relatives' Records

1 record on *m* relatives of individual *i*

$a_{ij}$ = additive genetic relationship of each relative with individual *i*.

$a_{jj'}$ = additive genetic relationship among relatives with records




$$\hat{g}_i = b_{g.\bar{x}}\,\bar{x}_i \qquad \text{where } \bar{x}_i = \sum_{j=1}^{m} x_{ij}/m$$

Then, $$b_{g.\bar{x}} = \sigma_{g.\bar{x}}/\sigma^2_{\bar{x}}$$

$t$ = **(intra-class) correlation** between phenotypic records on $j$ and $j'$:

$$t = r_{x_{ij}x_{ij'}} = \sigma_{x_{ij}x_{ij'}}/\sigma^2_p = \sigma_{(g_{ij}+e_{ij},\, g_{ij'}+e_{ij'})}/\sigma^2_p = (a_{jj'}\sigma^2_g + c^2\sigma^2_p)/\sigma^2_p = a_{jj'}h^2 + c^2$$

$c^2$ = **common environment correlation** between records $c^2 = \sigma_{e_{ij}e_{ij'}}/\sigma^2_p$

**Variance** of mean of *m* records with intra-class correlation *t*:

$$\sigma^2_{\bar{x}} = \mathrm{Var}\left(\sum_{j=1}^{m} x_{ij}/m\right) = \frac{m\sigma^2_p + m(m-1)t\sigma^2_p}{m^2} = \frac{1+(m-1)t}{m}\sigma^2_p$$

The **covariance** is: $$\sigma_{g.\bar{x}} = a_{ij}\sigma^2_g$$

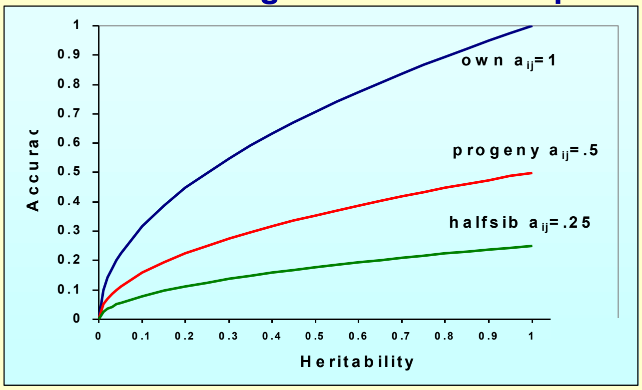
Thus, $$b_{g.\bar{x}} = \frac{m a_{ij}\sigma^2_g}{\sigma^2_p\left((m-1)t+1\right)} = a_{ij}\frac{mh^2}{(m-1)t+1}$$

**Accuracy** of selection: $$r = a_{ij}\sqrt{\frac{mh^2}{(m-1)t+1}}$$

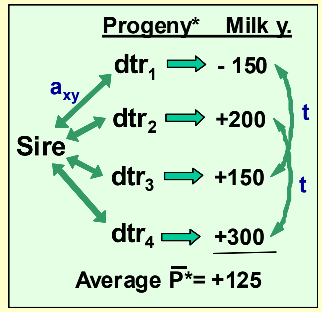
For repeated own records $a_{ij}$=1 and t = repeatability

## 2) EBV from single relative record

### Effect of degree of relationship




## 2) EBV from records on one type of relative




$h^2$=0.3 $a_{ij}$= 0.5

$a_{j,j'}$=0.25 t=(.25)(.3)=.075

b = 0.24

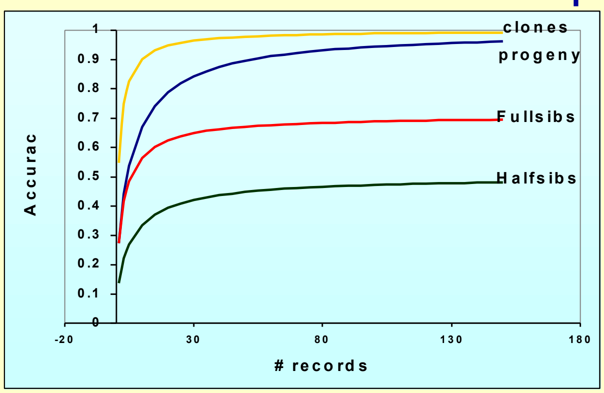
$\hat{g}$ = (0.69)(+125)=+30.6

r = 0.49

* each from a different dam

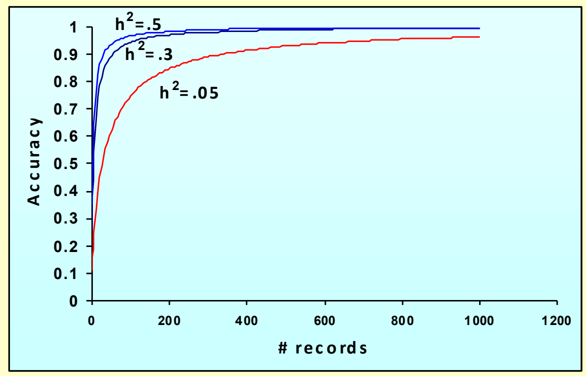
## 2) EBV from records on one type of relative

### Effect of # records and relationship



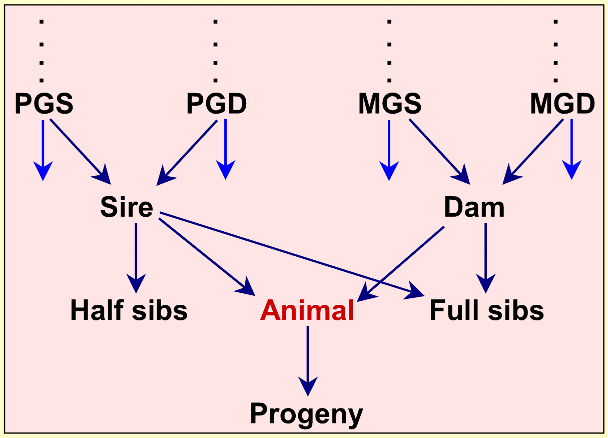

## 2) EBV from records on one type of relative

### Effect of heritability on progeny test accuracy



# 3) EBV from Multiple Sources - Selection Index

## 3) Estimating EBV from Multiple Sources of Information



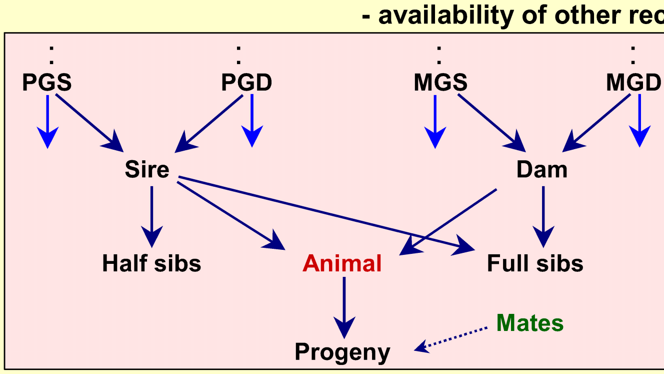

## Maximizing Accuracy of EBV

Optimize weight given to each record

$$\hat{g}_{Animal} = b_1x_{own} + b_2x_{HS} + b_3x_{dam} + b_4x_{prog} + \ldots\ldots$$

Weights by selection index theory = Multiple regression

Optimal weights depend on
- genetic relationships
- genetic parameters of trait
- availability of other records




**Selection index theory**: combining information from a variety of sources to obtain the most accurate predictor of genetic merit.

Two separate types of selection indexes:

1) **economic selection index**: predict genetic merit for overall economic value
2) **family selection index**: predict genetic merit for a single trait.

# Selection Index theory

**Breeding objective** - maximize improvement of economic merit.

**Aggregate genotype** or **breeding goal**

$H = v_1 g_1 + v_2 g_2 + \ldots + v_n g_n \quad = \mathbf{v'g}$

$\mathbf{v'} = [\, v_1 , v_2 , \ldots , v_n \,]$ — $v_i$ is economic weight for trait $i$

$\mathbf{g'} = [\, g_1 , g_2 , \ldots , g_n \,]$ — $g_i$ is true breeding value for trait $i$

**Selection index** $I = b_1 x_1 + b_2 x_2 + \ldots + b_m x_m \quad = \mathbf{b'x}$

$\mathbf{b'} = [\, b_1 , b_2 , \ldots , b_m \,]$ — vector of index weights

$\mathbf{x'} = [\, x_1 , x_2 , \ldots , x_m \,]$ — vector of records

Estimate $b_i$, such that:

$$\left.\begin{array}{l}\text{- selection on } I \text{ maximizes response in } H \\ \text{- } r_{I,H} \text{ is maximized} \\ \text{- prediction error variance} = \mathrm{Var}(H\text{-}I) \text{ minimized}\end{array}\right\} \text{Equivalent}$$

**Family selection index**: $H = g \quad \rightarrow \quad \mathbf{v} = [1]$

$I = \hat{g} = b_1 x_1 + b_2 x_2 + \ldots + b_m x_m$

## Derivation of index coefficients

Find index weights such that Var($H$-$I$) is minimized

$$E(H - I)^2 = E[I - H)'(I - H)] = E[I - H)'(I - H)'] = E[(\mathbf{b'x} - \mathbf{v'g})(\mathbf{x'b} - \mathbf{g'v})]$$

$$= E[(\mathbf{b'xx'b} - \mathbf{b'xg'v} - \mathbf{v'gx'b} + \mathbf{v'gg'v}]$$

$E(\mathbf{b'xx'b}) = \mathrm{Var}(\mathbf{b'x}) = \mathrm{Var}(I) = \mathbf{b'}\,\mathrm{Var}(\mathbf{x})\mathbf{b} = \mathbf{b'Pb}$ $\qquad \mathbf{P} = \mathrm{Var}(\mathbf{x})$

$E(\mathbf{b'xg'v}) = \mathrm{Cov}(\mathbf{b'x},\mathbf{g'v}) = \mathrm{Cov}(I,H) = \mathbf{b'}\,\mathrm{Cov}(\mathbf{x},\mathbf{g})\mathbf{v} = \mathbf{b'Gv}$ $\qquad \mathbf{G} = \mathrm{Cov}(\mathbf{x},\mathbf{g})$

$E(\mathbf{v'gx'b}) = \mathbf{v'G'b} = \mathbf{b'Gv}$

$E(\mathbf{v'gg'v}) = \mathrm{Var}(\mathbf{v'g}) = \mathrm{Var}(H) = \mathbf{v'}\,\mathrm{Var}(\mathbf{g})\mathbf{v} = \mathbf{v'Cv}$ $\qquad \mathbf{C} = \mathrm{Var}(\mathbf{g})$

**P** = $m$ x $m$ matrix of phenotypic covariances among the observations in $I$

**G** = $m$ x $n$ matrix of genetic covariances among $m$ observations in $I$ and the $n$ traits in $H$

**C** = $n$ x $n$ matrix of genetic covariances among the $n$ traits in $H$

To find index weights, minimize $M = \mathbf{b'Pb} - 2\mathbf{b'Gv} + \mathbf{v'Cv}$

$$\rightarrow \text{set first derivative} = 0 \rightarrow \frac{\delta M}{\delta \mathbf{b}} = 0 = 2\mathbf{Pb} - 2\mathbf{Gv} + 0 \rightarrow \quad \mathbf{Pb} = \mathbf{Gv}$$

⇨ Optimal index weights found from $\boxed{\mathbf{b} = \mathbf{P}^{-1}\mathbf{Gv}}$ = **SELECTION INDEX EQUATIONS**

## Accuracy of the index

$$r_{HI} = \frac{\sigma_{HI}}{\sigma_I \sigma_H}$$

$\sigma_I^2 = \mathrm{Var}(I) = \mathbf{b'Pb}$

$\sigma_H^2 = \mathrm{Var}(H) = \mathbf{v'Cv}$

$\sigma_{HI} = \mathrm{Cov}(H,I) = \mathbf{b'Gv}$

$$\boxed{r_{HI} = \frac{\sigma_{HI}}{\sigma_I \sigma_H} = \frac{\mathbf{b'Gv}}{\sqrt{\mathbf{b'Pb\, v'Cv}}}}$$

For the **optimal** index: $b_{HI} = 1 = \sigma_{HI}/\sigma_I^2 \;\rightarrow\; \sigma_{HI} = \sigma_I^2$

and $\mathbf{Pb} = \mathbf{Gv} \;\rightarrow\; \mathbf{b'Pb} = \mathbf{b'Gv}$

$$\rightarrow \quad r_{HI} = \frac{\sigma_I}{\sigma_H} = \sqrt{\frac{\mathbf{b'Pb}}{\mathbf{v'Cv}}} = \frac{\sigma_{HI}}{\sigma_H} = \boxed{\sqrt{\frac{\mathbf{b'Gv}}{\mathbf{v'Cv}}}}$$

# Accuracy for Family Selection Indexes

$H = g \qquad \mathbf{v} = [1] \qquad \sigma_H^2 = \sigma_g^2$

$\mathbf{b} = \mathbf{P}^{-1}\mathbf{Gv} \;\rightarrow\; \mathbf{b} = \mathbf{P}^{-1}\mathbf{G} \qquad r_{HI} = r_{g,\hat{g}} = \boxed{\sqrt{\frac{\mathbf{b'G}}{\sigma_{\mathbf{g}}^2}}}$

**Example Index of individual record and full-sib mean performance**

$x_1$ = individual's performance

$x_2$ = mean performance of that individual's 5 full sibs $\qquad h^2 = 0.5$

$$I = \hat{g} = b_1 x_1 + b_2 x_2$$

$$\mathbf{P} = \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1x_2} \\ \sigma_{x_1x_2} & \sigma_{x_2}^2 \end{bmatrix} \qquad \mathbf{G} = \begin{bmatrix} \sigma_{x_1 g} \\ \sigma_{x_2 g} \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} 1 & \tfrac{1}{2}h^2 \\ \tfrac{1}{2}h^2 & \frac{1 + (m-1)\tfrac{1}{2}h^2}{m} \end{bmatrix} \sigma_p^2 = \begin{bmatrix} 1 & .25 \\ .25 & .4 \end{bmatrix} \sigma_p^2$$

$$\mathbf{G} = \begin{bmatrix} h^2 \\ \tfrac{1}{2}h^2 \end{bmatrix} \sigma_p^2 = \begin{bmatrix} .5 \\ .25 \end{bmatrix} \sigma_p^2$$

$$\mathbf{b} = \mathbf{P}^{-1}\mathbf{G} = \begin{bmatrix} 1 & .25 \\ .25 & .4 \end{bmatrix}^{-1} \begin{bmatrix} .5 \\ .25 \end{bmatrix} = \begin{bmatrix} .4074 \\ .3704 \end{bmatrix}$$

$$I = \hat{g} = 0.4074\, x_1 + 0.3704\, x_2$$

$$r_{HI} = r_{g,\hat{g}} = \sqrt{\frac{\mathbf{b'G}}{\sigma_{\mathbf{g}}^2}} = \sqrt{\frac{\begin{bmatrix} .4074 \\ .3704 \end{bmatrix}' \begin{bmatrix} .5 \\ .25 \end{bmatrix} \sigma_p^2}{0.5\,\sigma_{\mathrm{p}}^2}} = 0.77$$

By adding the mean of 5 full sibs the accuracy of evaluation is increased from 0.71 to 0.77, i.e. by 8.9%.

**Correlation between relatives:** Correlation between index values of two relatives, *i* and *j*,

$$t_{i,j} = \text{corr}(I_i, I_j) = \text{corr}(\mathbf{b'x}_i, \mathbf{b'x}_j) = \frac{\mathbf{b'}\,\mathbf{cov}(\mathbf{x}_i, \mathbf{x}_j)\mathbf{b}}{\mathbf{b'Pb}} = \frac{\mathbf{b'Rb}}{\mathbf{b'Pb}}$$

$\mathbf{R}$ = $m$ x $m$ matrix with covariances between information sources on the relatives

## Summary of selection index formulae for any index and for the optimum index

| Parameters | | Any Index | Optimal Index |
|---|---|---|---|
| Index weights | $\mathbf{b}$ | Arbitrary | $\mathbf{b} = \mathbf{P}^{-1}\mathbf{Gv}$ |
| Index variance | $\sigma_I^2$ | $\mathbf{b'Pb}$ | $\mathbf{b'Pb} = \mathbf{b'Gv}$ |
| Breeding goal var. | $\sigma_H^2$ | $\mathbf{v'Cv}$ | $\mathbf{v'Cv}$ |
| Goal-index covar. | $\sigma_{HI}$ | $\mathbf{b'Gv}$ | $\mathbf{b'Gv}$ |
| Accuracy | $r_{HI}$ | $\dfrac{\mathbf{b'Gv}}{\sqrt{\mathbf{b'Pb\,v'Cv}}}$ | $\sqrt{\dfrac{\mathbf{b'Gv}}{\mathbf{v'Cv}}} = \sqrt{\dfrac{\mathbf{b'Pb}}{\mathbf{v'Cv}}} = \dfrac{\sigma_I}{\sigma_H}$ |

## General equations to derive elements of selection index matrices

$m$ = number of records within a group

$c^2$ = common environment component

$\sigma_{p_k}$ = phenotypic standard deviation of trait $k$

$\sigma_{g_k}$ = genetic standard deviation of trait $k$

$r_{p_{kl}}$ = phenotypic correlation between traits $k$ and $l$

$r_{g_{kl}}$ = genetic correlation between traits $k$ and $l$

$a$ = genetic relationship within a group

$a_{ij}$ = relationship between groups $i$ and $j$

$a_{hj}$ = additive genetic relationship between individual in breeding goal ($h$) and individuals in group $j$

### P-matrix

diagonal:

- Variance of $m$ records of a given type

$$\frac{1+(m-1)t}{m}\sigma_p^2 \qquad (=\sigma_p^2 \text{ for } m=1)$$

$t$ = repeatability for repeated records

$t = ah^2 + c^2$ for multiple individuals

off-diagonal:

- Covariance between mean of $m$ records on different traits ($k$ and $l$) for same group:

$$\frac{r_{p_{kl}}\sigma_{p_k}\sigma_{p_l} + (m-1)a\,r_{g_{kl}}\sigma_{g_k}\sigma_{g_l}}{m} \qquad ( = r_{p_{kl}}\sigma_{p_k}\sigma_{p_l} \text{ for } m=1)$$

- Covariance between (mean of) record(s) on same trait $k$ for different groups ($i$ and $j$):

$$(a_{ij}h_k^2 + c_k^2)\sigma_{p_k}^2$$

- Between records on different traits ($k$ and $l$) in different groups ($i$ and $j$):

$$a_{ij}r_{g_{kl}}\sigma_{g_k}\sigma_{g_l}$$

**G-matrix**

- Covariance of the genetic value for trait $k$ on the breeding goal animal ($h$) with records on trait $l$ for group $j$ $\quad a_{hj} r_{g_{kl}} \sigma_{g_k} \sigma_{g_l} \quad (= a_{hj} \sigma^2_{g_k}$ if $k=l)$

**C-matrix**

- diagonal: Variance of genetic value for trait $k$ $\quad \sigma^2_{g_k}$
- off-diagonal: Covariance between genetic values for traits $k$ and $l$ on breeding goal animal $\quad r_{g_{kl}} \sigma_{g_k} \sigma_{g_l}$

## 4) Selection Index and Animal Model BLUP – Pseudo BLUP EBV

Two approaches to modeling Animal model BLUP EBV using selection index:

1) Selection index based on relatives providing the greatest amount of information
2) Pseudo-BLUP: Selection index that includes parental EBV as sources of information, along with records on the individual itself, collateral relatives, and progeny.

Information sources:

$x_i$ = animal's own record,

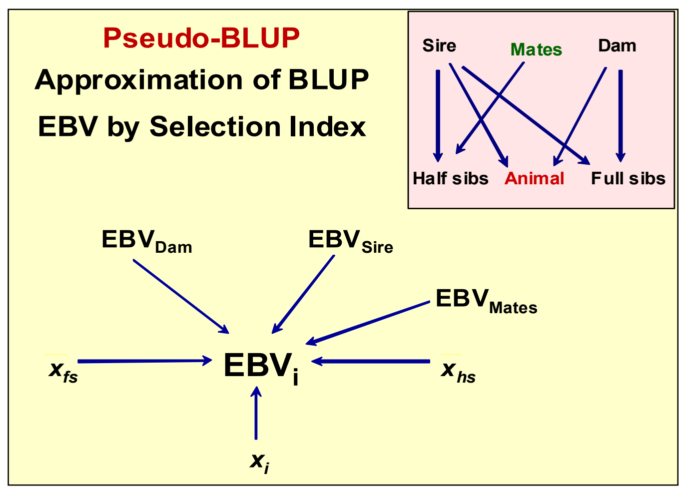
$x_{fs}$ = average of $n$-1 full sibs

$x_{hs}$ = average of $(m-1)n$ half sibs

$\hat{g}_s$ = EBV of the sire

$\hat{g}_d$ = EBV of the dam

$\bar{\hat{g}}_m$ = average EBV of ($m$-1) mates of the sire that produced the half sibs




Pseudo BLUP EBV = $\quad I_i = \hat{g}_i = b_1 x_i + b_2 x_{fs} + b_3 x_{hs} + b_4 \hat{g}_s + b_5 \hat{g}_d + b_6 \bar{\hat{g}}_m$

$$
\mathbf{P} = \begin{bmatrix}
\sigma^2_{x_i} & \sigma_{x_i x_{fs}} & \sigma_{x_i x_{hs}} & \sigma_{x_i \hat{g}_s} & \sigma_{x_i \hat{g}_d} & \sigma_{x_i \bar{\hat{g}}_m} \\
 & \sigma^2_{x_{fs}} & \sigma_{x_{fs} x_{hs}} & \sigma_{x_{fs} \hat{g}_s} & \sigma_{x_{fs} \hat{g}_d} & \sigma_{x_{js} \bar{\hat{g}}_m} \\
 & & \sigma^2_{x_{hs}} & \sigma_{x_{hs} \hat{g}_s} & \sigma_{x_{hs} \hat{g}_d} & \sigma_{x_{hs} \bar{\hat{g}}_m} \\
 & & & \sigma^2_{\hat{g}_s} & \sigma_{\hat{g}_s \hat{g}_d} & \sigma_{\hat{g}_s \bar{\hat{g}}_m} \\
 & & & & \sigma^2_{\hat{g}_d} & \sigma_{\hat{g}_d \bar{\hat{g}}_m} \\
 & & & & & \sigma^2_{\bar{\hat{g}}_m}
\end{bmatrix}
$$

$$\mathbf{P} = \begin{bmatrix} 1 & {}^1/_2\, h^2 + c^2 & {}^1/_4\, h^2 & {}^1/_2\, r_s^2 h^2 & {}^1/_2\, r_d^2 h^2 & 0 \\ & \dfrac{1 + (n-2)({}^1/_2\, h^2 + c^2)}{n-1} & {}^1/_4\, h^2 & {}^1/_2\, r_s^2 h^2 & {}^1/_2\, r_d^2 h^2 & 0 \\ & & {}^1/_4\, h^2 + \dfrac{{}^1/_4\, h^2 + c^2}{m-1} + \dfrac{1 - {}^1/_2\, h^2 - c^2}{n(m-1)} & {}^1/_2\, r_s^2 h^2 & 0 & \dfrac{{}^1/_2\, r_m^2 h^2}{m-1} \\ & & & r_s^2 h^2 & 0 & 0 \\ & & & & r_d^2 h^2 & 0 \\ & & & & & \dfrac{r_m^2 h^2}{m-1} \end{bmatrix} \sigma_p^2$$

$$\mathbf{G} = \begin{bmatrix} \sigma_{g_i x_i} & \sigma_{g_i x_{fs}} & \sigma_{g_i x_{hs}} & \sigma_{g_i \hat{g}_s} & \sigma_{g_i \hat{g}_d} & \sigma_{g_i \bar{\hat{g}}_m} \end{bmatrix}$$

$$\mathbf{G} = \begin{bmatrix} h^2 & {}^1/_2\, h^2 & {}^1/_4\, h^2 & {}^1/_2\, r_s^2\, h^2 & {}^1/_2\, r_d^2\, h^2 & 0 \end{bmatrix} \sigma_p^2$$

$$\mathbf{b} = \mathbf{P}^{-1}\mathbf{G} \qquad\qquad r_{g,\hat{g}} = \sqrt{\mathbf{b'}\,\mathbf{Pb}/\sigma_g^2}$$

With $$x_{hs} = \left( \sum_{k=1}^{m-1} \sum_{l=1}^{n} \frac{x_{kl}}{n} \right) /(m-1)$$

Where $$x_{kl} = {}^1/_2\, g_s + {}^1/_2\, g_{d_k} + g_{ms_{kl}} + c_{kl} + e_{kl}$$

Thus $$x_{hs} = {}^1/_2\, g_s + {}^1/_2 \frac{\sum_{k=1}^{m-1} \left( g_{d_k} + c_k \right)}{m-1} + \frac{\sum_{k=1}^{m-1} \sum_{l=1}^{n} \left( g_{ms_{kl}} + e_{kl} \right)}{n(m-1)}$$

And $$\sigma_{x_{hs}}^2 = {}^1/_4 \sigma_g^2 + \frac{{}^1/_4 \sigma_g^2 + c^2}{m-1} + \frac{{}^1/_2 \sigma_g^2 + \sigma_e^2}{n(m-1)}$$

Also, $$\sigma_{\hat{g}}^2 = r_{g,\hat{g}}^2\, \sigma_g^2$$

And $$\sigma_{x_i \hat{g}_s} = \sigma_{({}^1/_2 g_s + {}^1/_2 g_d + g_{m_i} + e_i,\, \hat{g}_s)} = \sigma_{({}^1/_2 g_s,\, \hat{g}_s)} = {}^1/_2 \sigma_{g_s, \hat{g}_s} = {}^1/_2\, r_s^2 \sigma_g^2$$

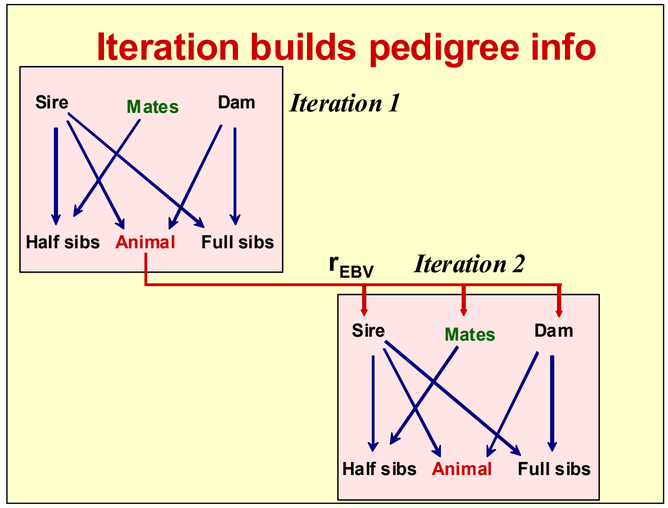
**Iterative Procedure to Compute Accuracy of EBV**
**Building-up pedigree information**

$$EBV_i = b' [x_i, x_{fs}, x_{hs}, EBV_{Dam}, EBV_{Sire}, EBV_{Mates}]'$$

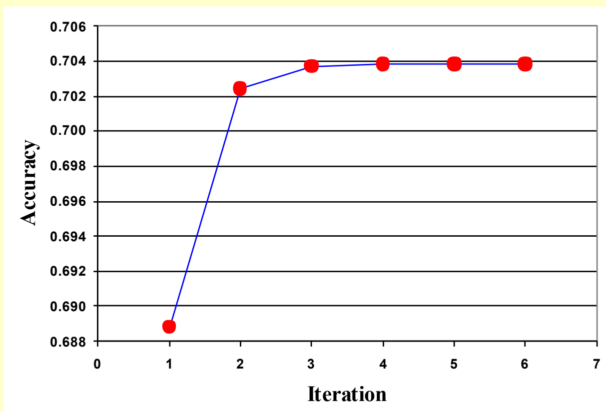
$$b = P^{-1}G \qquad r_{EBV_i} = \sqrt{b'Pb}/\sigma_g$$

1) Set accuracy of $EBV_{Sire}$=$EBV_{Dam}$=$EBV_{Mates}$= h (own record)
2) Set up index (P, G) and derive accuracy $r_{EBVi} = \sqrt{b'Pb}/\sigma_g$
3) Set accuracy of $EBV_{Sire}$=$EBV_{Dam}$=$EBV_{Mates}$ equal to $r_{EBV_i}$
4) Repeat steps 2 and 3 until accuracy converges

Needs adaptation if selection after progeny




**Example**




➔➔ BLUPEBV.XLS

➔➔ SELACTION growth sigp=.25 h2 = 0.25 own + 5 FS SIP_r =0.7698

# Some general properties of EBV

**Unbiased:** $E(g_i | \hat{g}_i) = \hat{g}_i$

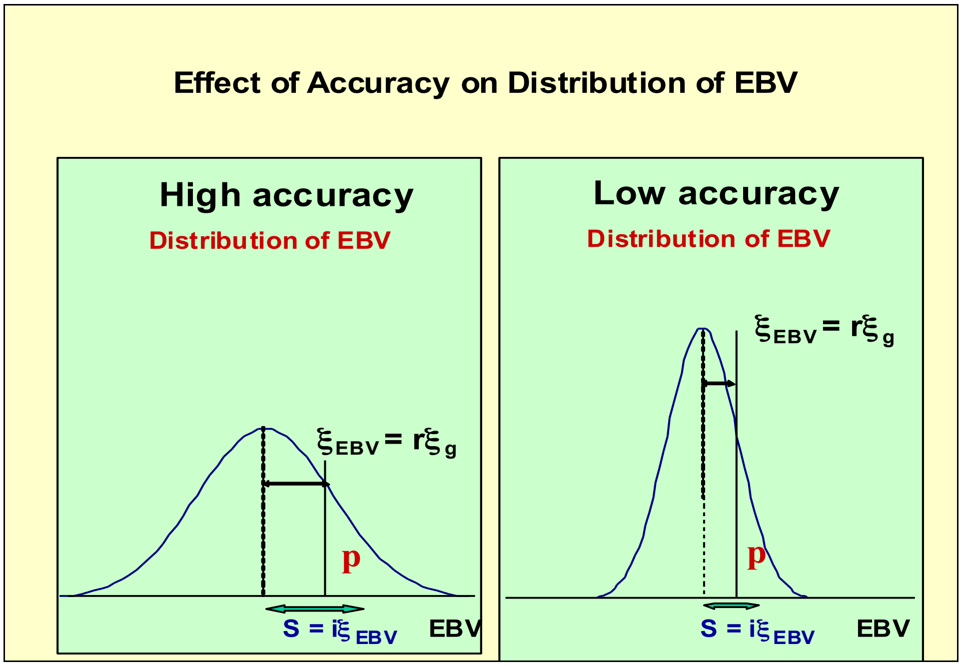
➔ selection on $\hat{g}$ maximizes $E(g)$ for the group of selected individuals

**Regression of true on EBV = 1:** $b_{g,\hat{g}} = 1$

**Accuracy of EBV:** $r = r_{g,\hat{g}} = b_{\hat{g},g} \dfrac{\sigma_{\hat{g}}}{\sigma_g} = \dfrac{\sigma_{\hat{g}}}{\sigma_g}$

**Covariance between true and EBV:** $\sigma_{g,\hat{g}} = r_{g,\hat{g}}\ \sigma_g\ \sigma_{\hat{g}} = \sigma^2_{\hat{g}}$

**Variance of EBV:** $\sigma^2_{\hat{g}} = r^2 \sigma^2_g$




**Prediction error:** $\varepsilon_i = g_i - \hat{g}_i$

**Variance of prediction errors:** $\sigma_\varepsilon^2 = \sigma_{\hat{g}}^2 + \sigma_g^2 - 2\sigma_{g,\hat{g}} = \sigma_{\hat{g}}^2 + \sigma_g^2 - 2\sigma_{\hat{g}}^2 = \sigma_g^2 - \sigma_{\hat{g}}^2 = \sigma_g^2 - r^2\sigma_g^2$

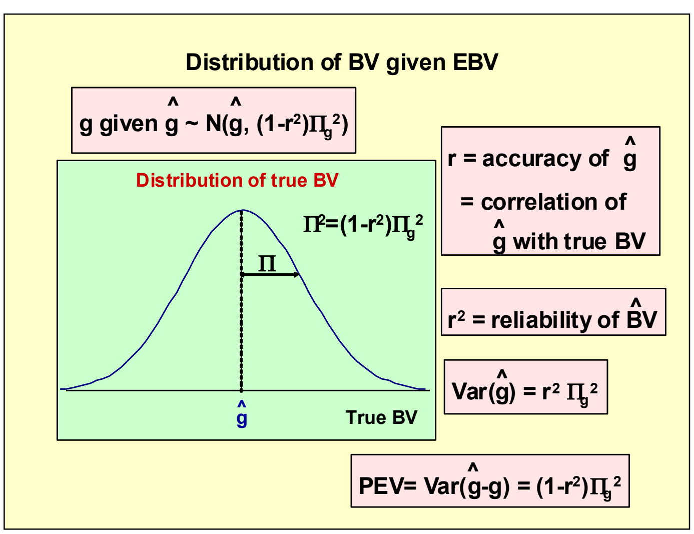
$= (1\text{-}r^2)\sigma_g^2$ Note that $\sigma_g^2 = \sigma_{\hat{g}}^2 + \sigma_\varepsilon^2$

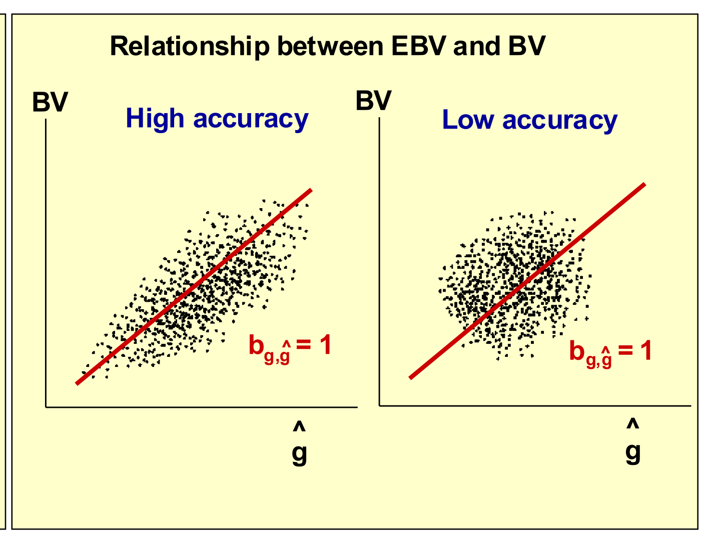
**Covariance between EBV and prediction errors:**

$$\sigma_{\hat{g},\varepsilon} = \sigma_{\hat{g},\hat{g}-g} = \sigma_{\hat{g}}^2 - \sigma_{g,\hat{g}} = \sigma_{\hat{g}}^2 - \sigma_{\hat{g}}^2 = 0$$

**Dist'n of true BV given EBV:** $g_i | \hat{g}_i \sim N(\hat{g}_i, (1\text{-}r^2)\sigma_g^2)$

**Dist'n prediction errors:** $\varepsilon_i \sim N(0, (1\text{-}r^2)\sigma_g^2)$



# Effect of Accuracy on Distribution of True BV for animals with a given $\hat{BV}$