How much genetic change?
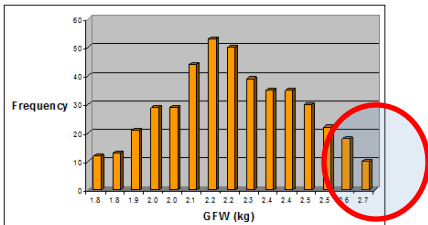
Breeder's Equation

Selection Response $= i \; r_{IA} \sigma_A$

Selection intensity

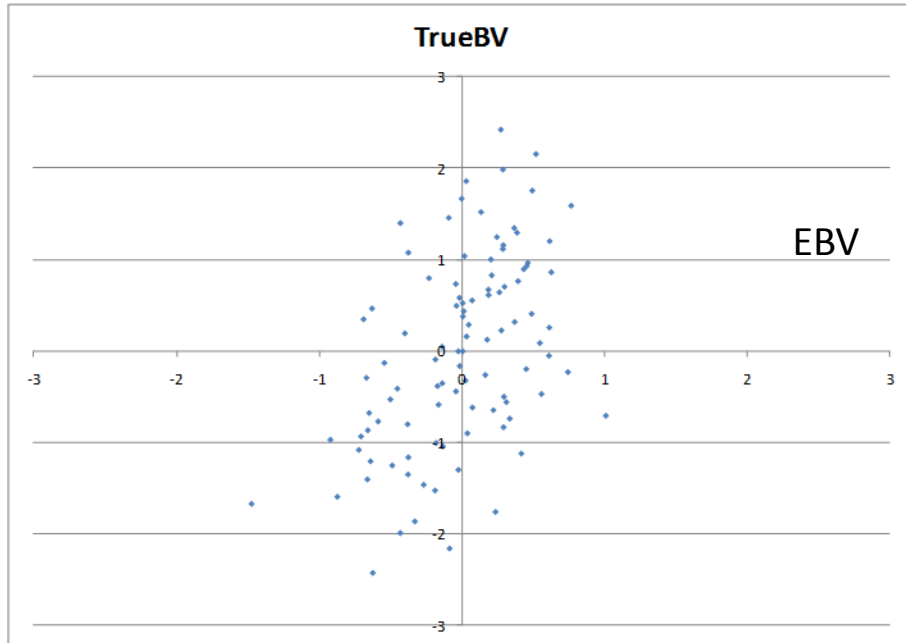Selection accuracy

Genetic variation

Selection Accuracy
## refers to the accuracy of selecting on breeding value

- Estimated Breeding Values *(EBVs)* are estimates of the True Breeding Values (TBVs), which cannot be observed directly

- Quality of EBV is measured by Accuracy

- Accuracy is correlation between EBV and TBV (Acc 0-1)

- The closer the EBV is to the true breeding value the higher the accuracy, the greater the selection response
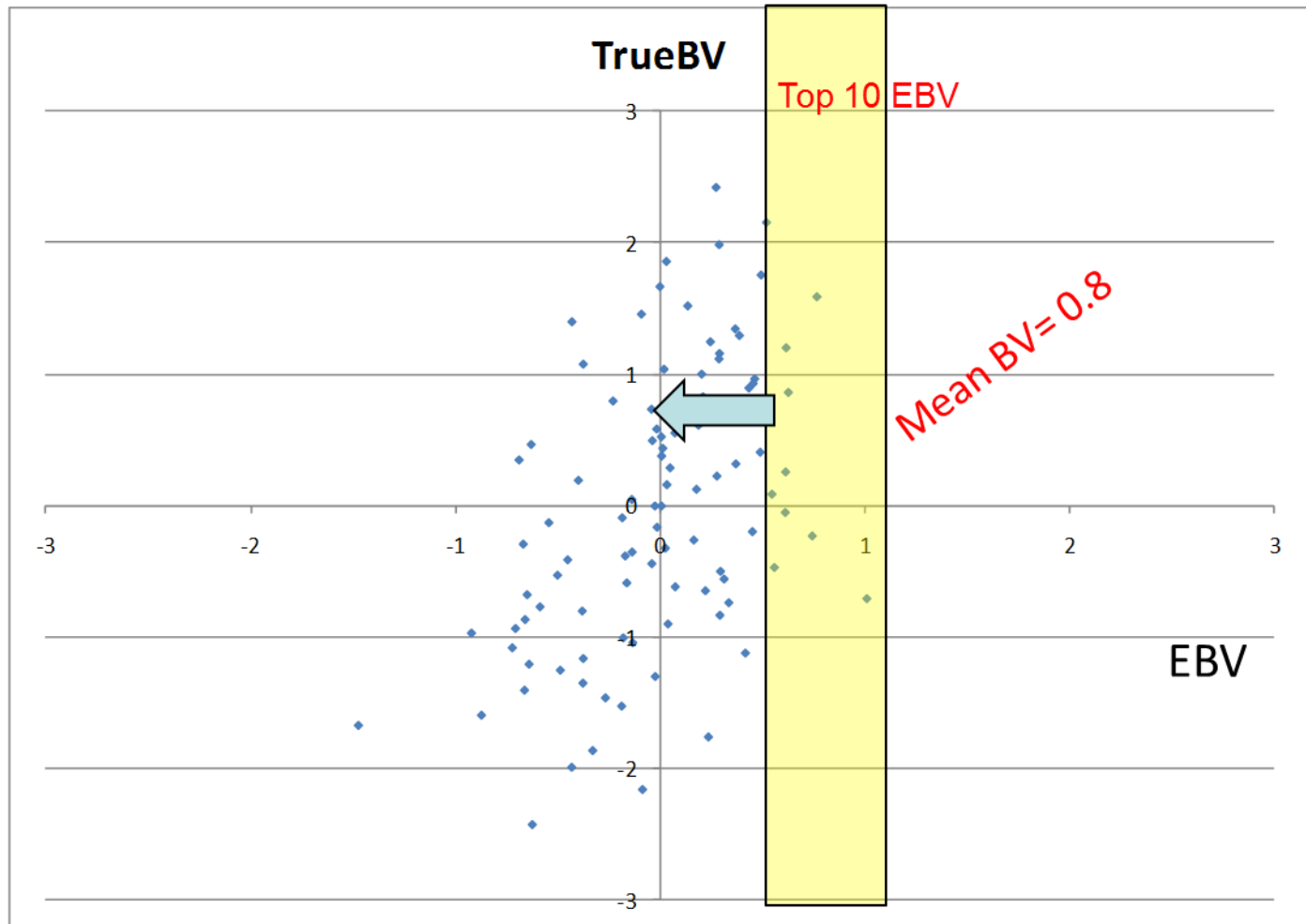
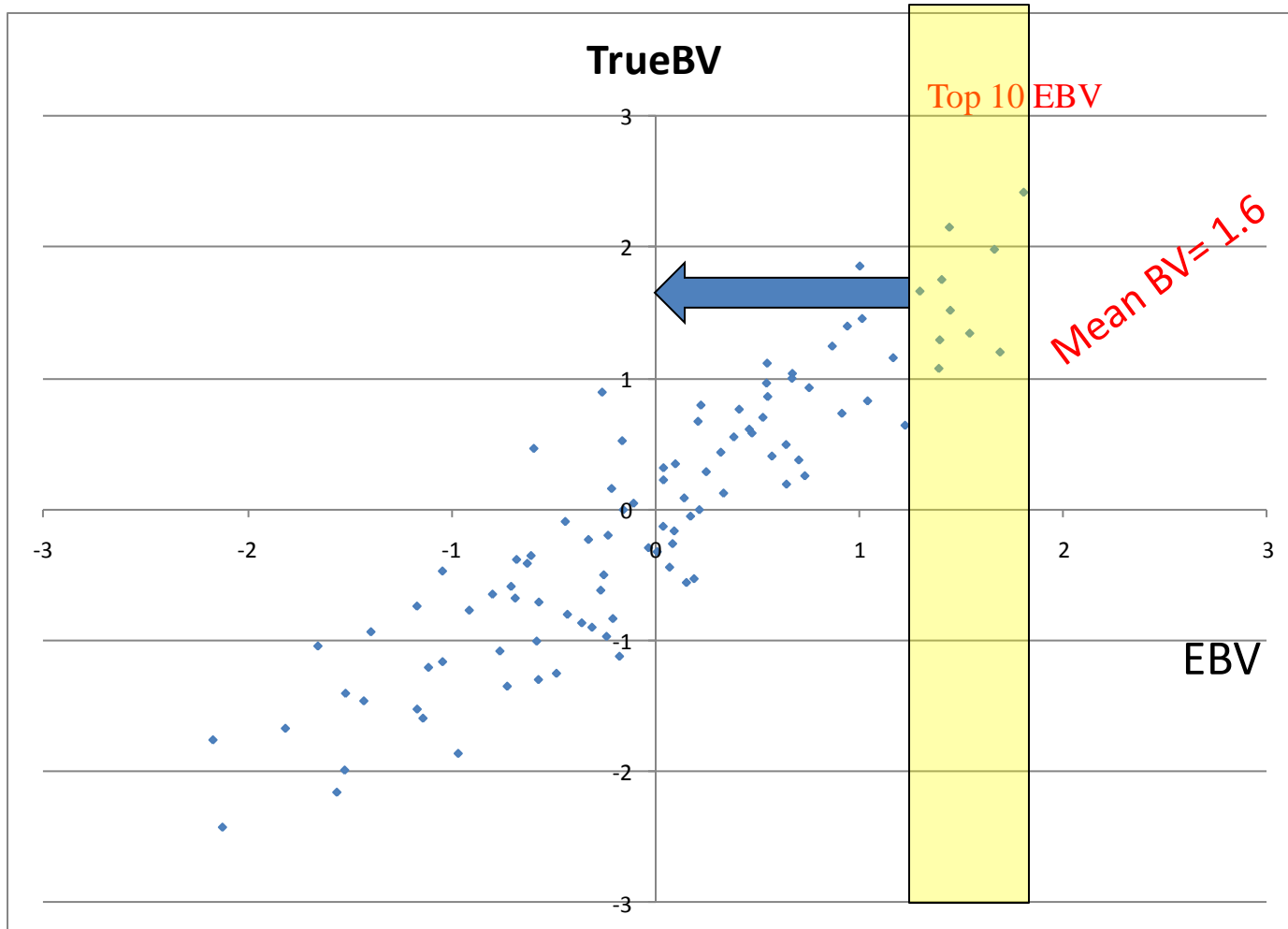# Accuracy of EBV = correlation with True BV

Accuracy = 45%
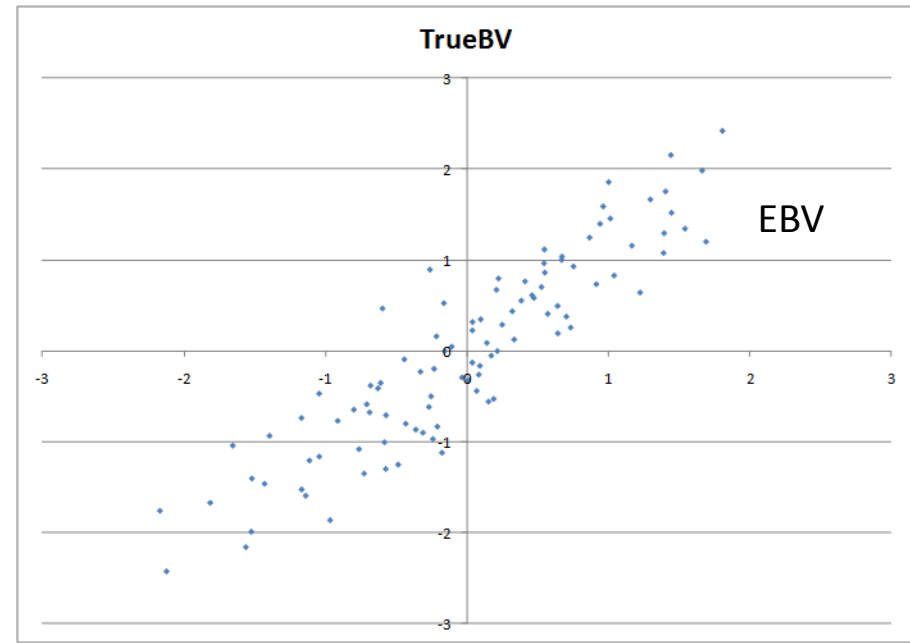
Accuracy = 90%

Select on EBV: accuracy related to response

Accuracy = 45%

Double accuracy gives double selection response!

Accuracy = 90%

Sources of information to get EBV

- <u>own phenotype</u>, as $EBV_i = h^2P$

- But we can have a number of sources of information, such <u>information from relatives</u>
  - *Performance on dam, sire, siblings,*
  - *Genomic test*

- More information results in a
  more accurate EBV

  - And therefore more selection response

Animal model EBV
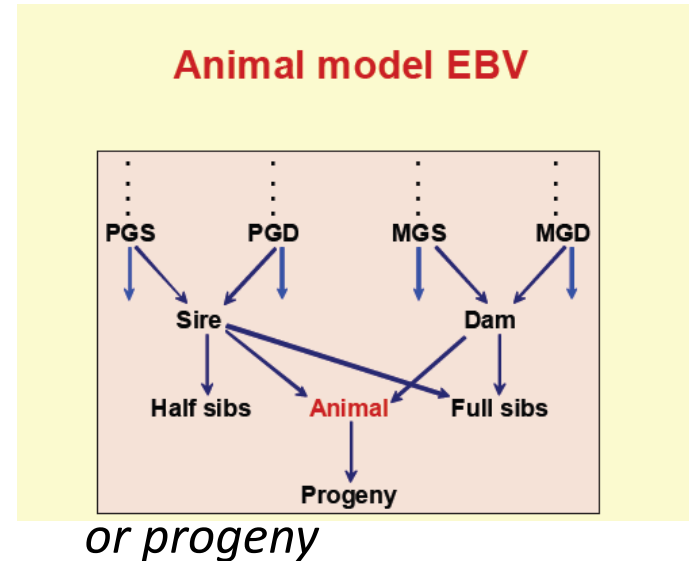
or progeny

Use Selection Index / BLUP

What part of observed phenotypic differences in parents is passed on to progeny?
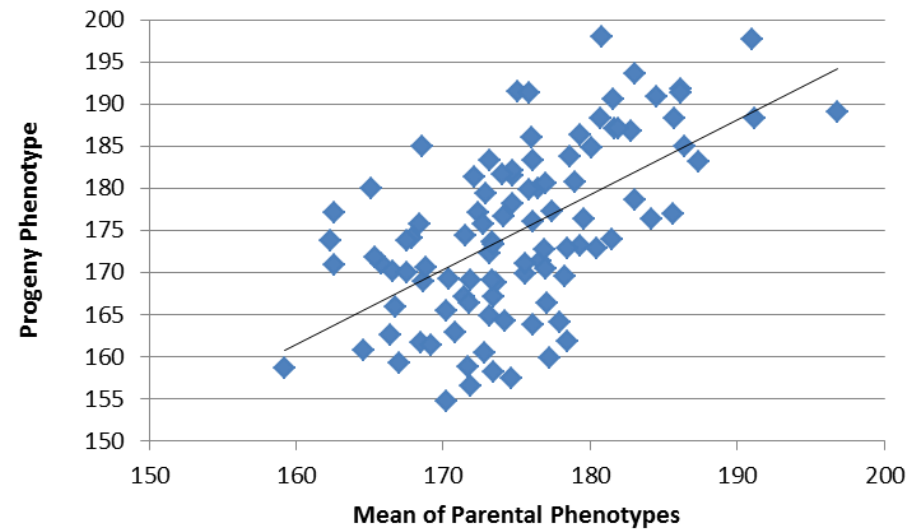
Parent has   *a phenotypic deviation (e.g. +10)*
Progeny gets   *0.5 * heritability x phenotypic deviation*

heritability = 5%   *heritability = 75%*

progeny gets + 0.25   *progeny gets + 3.75*
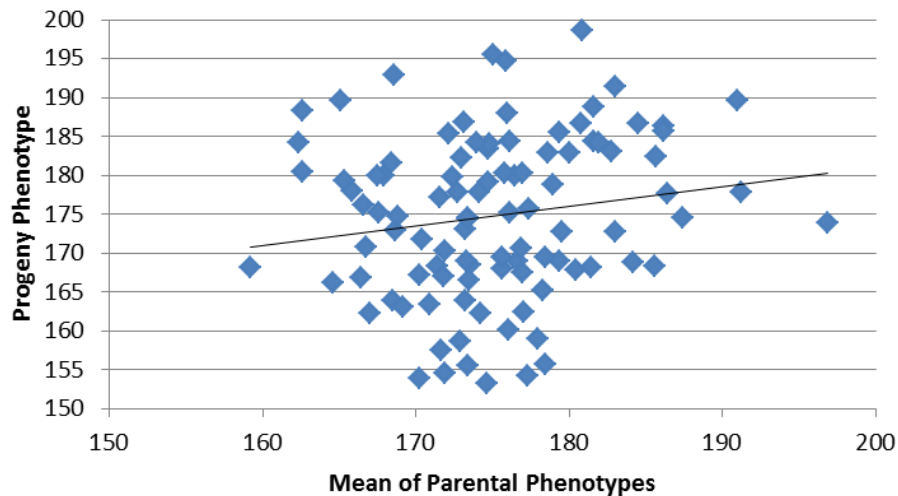
Breeding Value

What part of differences in parents is passed on to progeny?

Parent has   *a phenotypic deviation*   *e.g. +10*

Progeny gets   *0.5 * heritability x phenotypic deviation*

heritability = 5%   *heritability = 75%*

   parent EBV   =   + 0.50   *parent EBV   = + 7.50*

   parent EPD   =   + 0.25   *progeny EPD = + 3.75*
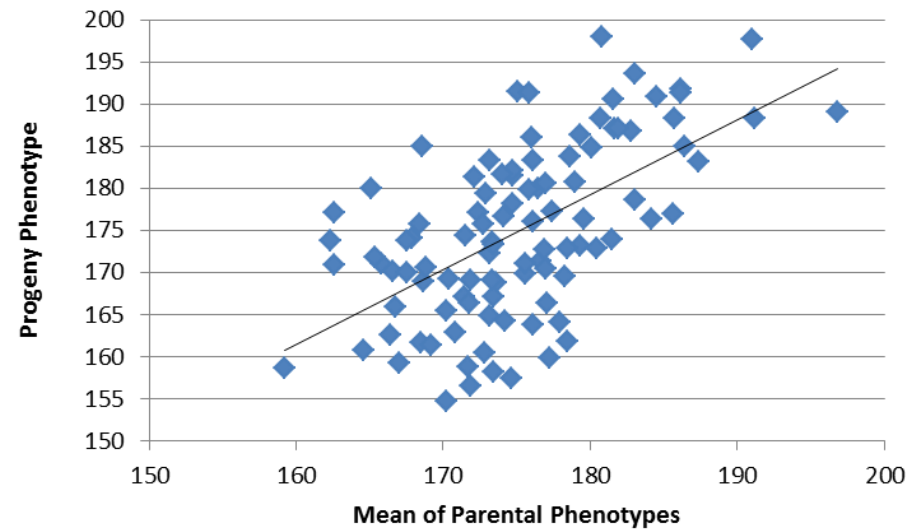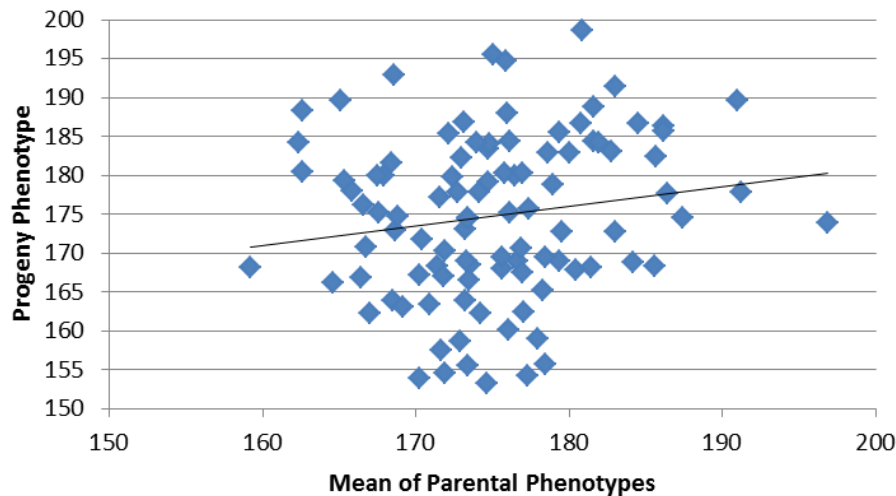
Estimating Breeding Values

- Principle
    - how is it estimated?
- Using information from relatives
- Properties
    - Accuracy
    - Variance
    - Prediction Error
- Selection Response
    - select on EBV

How to use information on an animals' own performance?

Animal	X1 own perform.	EBV
1	3.6	?
2	-8.3	?
3	3.5	?
4	16.7	?
5	-14.4	?
6	6.6	?
7	-1.4	?
8	-6.2	?
9	6.2	?
10	15.0	?

Principle of EBV estimation

Regression of breeding value on phenotype

How much do we expect the BV to be better if we know that P is one unit better

Slope = cov(A,P)/var(P)

A

P
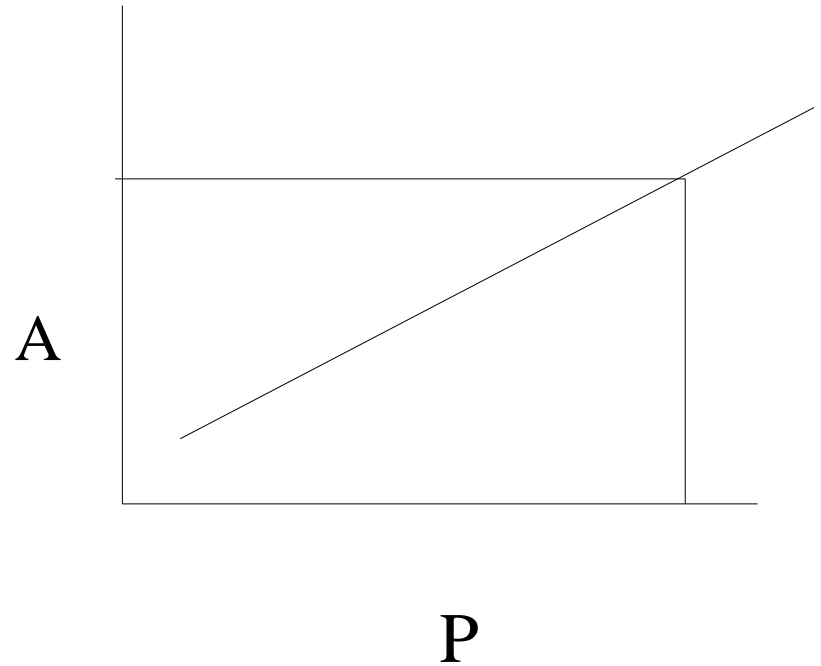
EBV estimation: regression

- If   P = Own Performance

Slope = $cov(A, P)/ var(P)$

$= V_A / V_P$

$= h^2$
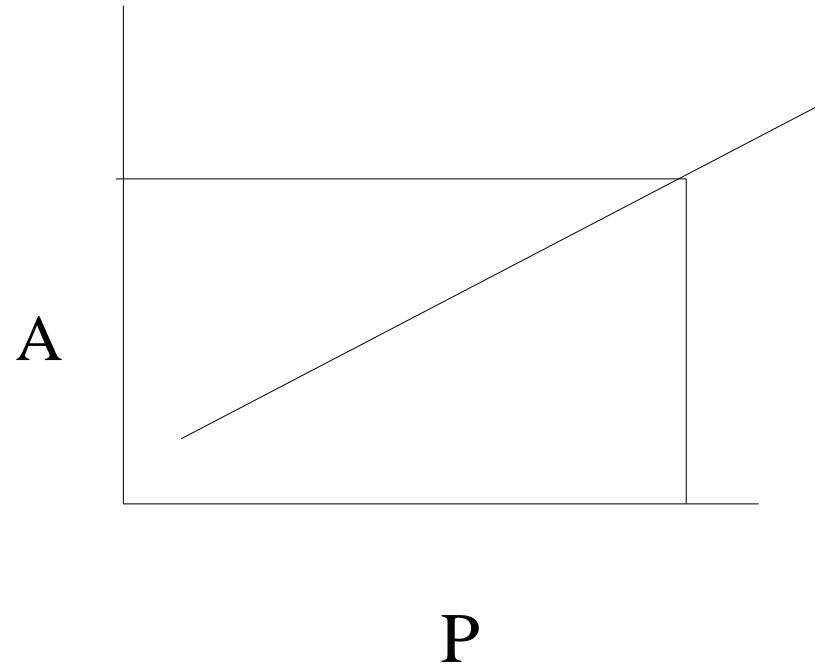
How to use information from relatives?

Animal	sire's perform.	EBV
1	9.2	?
2	3.2	?
3	-15.4	?
4	7.1	?
5	-4.9	?
6	16.1	?
7	-12.2	?
8	8.7	?
9	7.4	?
10	-4.0	?

EBV estimation: regression

- If P = Performance of Sire

Slope $= \text{cov}(A, P_{\text{sire}})/ \text{var}(P_{\text{sire}})$

$= \frac{1}{2} V_A / V_P$

$= \frac{1}{2} h^2$
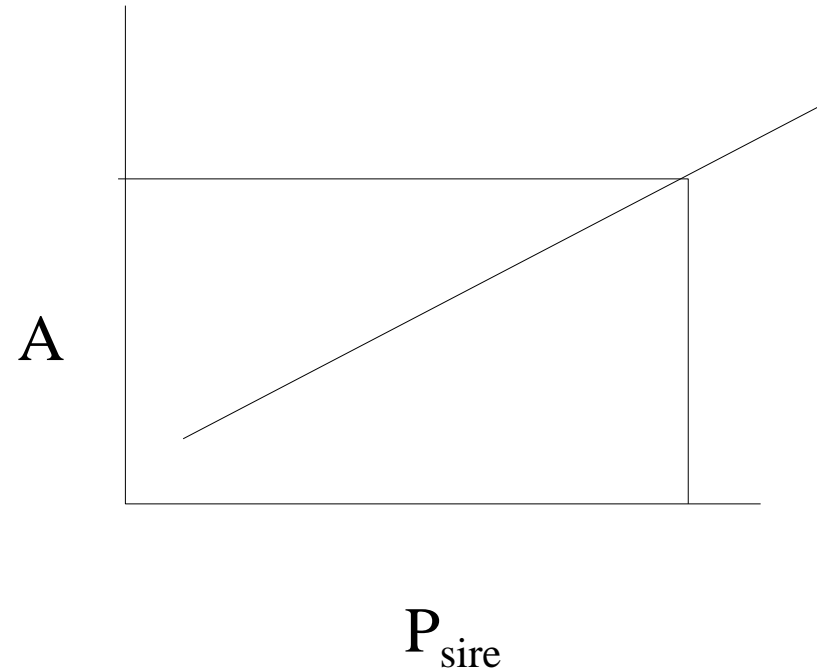
Here we need some basic Quantitative Genetic theory

$P = A + E$ → General Model

$Var(P) = var(A) + var(E) = V_A + V_E$

no cov. between A and E

$cov(A_i, P_i) = cov(A_i, A_i) + cov(A_i, E_i) = V_A$

if A same animal as P

$cov(A_i, P_i) = cov(A_j, A_i) + cov(A_j, E_i) = a_{ij}V_A$

a_{ij} = additive genetic relationship between i and j

$cov(P_i, P_j) = a_{ij}V_A$

as E's are uncorrelated if not same animal

Finding the optimal index weights

Regression of breeding value on a phenotype   **EBV = b . P**

Regression coefficient $= \dfrac{\operatorname{cov}(P_i, A)}{\operatorname{var}(P_i)}$

Examples:

if P = Own Performance $b = \dfrac{V_A}{V_A + V_E} = h^2$

if P = Sire Performance

$b = \dfrac{\frac{1}{2}V_A}{V_A + V_E} = \frac{1}{2}h^2$

What does this all mean?

Regression of x on y :

– If we know x, what does that tell us about y?

if we know an animal performed well, what does that tell us about his breeding value?

A: Depends on heritability

$$b = h^2$$

if we know an animal's sire performed well, what does that tell us about his breeding value?

A: It carries only half of Sire's genes + it depends on heritability

$$b = \frac{1}{2} h^2$$

if we know an animal's great granddam performed well, what does that tell us about his breeding value? A: It carries even less of great grandma's genes

$$b = 1/8 \, h^2$$

Some more basic Quantitative Genetics theory

Using information based on a mean

var(mean) =   common variance   +   specific/n

equal to all   *Specific bit is averaged out*

Common variance = t. $\sigma_p{}^2$ $\sigma_p{}^2$ = phenotypic variance

t =   intra class correlation

→ correlation among group members

e.g. between FS: t = ½ h² + c²

e.g. mean of n Full Sibs

Var(P_{FS}) =   ½ V_A + $V_{common\ env}$ +   { ½V_A + $V_{specific\ env}$ } / n

                 t. $\sigma_p{}^2$ +   (1- t) $\sigma_p{}^2$/n

Var(mean) = t. $\sigma_p{}^2$ + (1- t) $\sigma_p{}^2$/n   = $\dfrac{1+(n-1)t}{n} \sigma_p{}^2$

An easy way to workout variances of means

$$\begin{bmatrix} 1 & t & .. & t \\ t & 1 & & t \\ . & .. & ... & .. \\ t & t & .. & 1 \end{bmatrix} \sigma_p^{\,2}$$

Mean = Sum/n

Var(Mean) = var(Sum)/n²

Var-cov matrix of all elements that make up the mean

Var(Sum) = { n.1 + n(n-1).t } $\sigma_p^{\,2}$

Var(Mean) = { [n.1 + n(n-1)t] /n² } $\sigma_p^{\,2}$

= { [n.1 + n(n-1)t] /n² } $\sigma_p^{\,2}$

= { [1 + (n-1)t] /n } $\sigma_p^{\,2}$

An easy way to workout co-variances with means

Mean = Sum/n

CoVar(X, Mean) = CoVar(X, Sum)/n

$$\begin{matrix} t \\ . \\ t \end{matrix} \quad \sigma_p{}^2$$

CoVar(X, Mean) = { n.t} $\sigma_p{}^2$/n

CoVar(X, Mean) = t$\sigma_p{}^2$

Vector with covariances of a variable X with all elements that make up the mean, X could be anther phenotype, or a breeding value
Covar(X, P_i) = t$\sigma_p{}^2$

Covariance of a variable with the mean I same as covariance of a variable with each element that make up the mean

Some more basic Quantitative Genetics theory

Using information based on a mean (e.g. mean of sibs)

	FS	HS

$$\text{Var(mean P)} = \frac{1+(n-1)t}{n}\,\sigma_p^{\,2}$$

FS: $t = \frac{1}{2}\,h^2 + c^2$ HS: $t = \frac{1}{4}\,h^2$

$$\text{Cov(mean P, A)} = a_{ij}\sigma_a^{\,2}$$

FS: $a_{ij} = \frac{1}{2}\,h^2$ HS: $a_{ij} = \frac{1}{4}\,h^2$

a_{ij} = genetic relationship between members of group and subject (A)

Weight for Mean of		n Full Sibs	n Half Sibs
$\dfrac{\text{Cov}}{\text{Var}}$	=	$\dfrac{\frac{1}{2}.n.h^2}{1+(n-1)t_{FS}}$	$\dfrac{n.t}{1+(n-1)t_{HS}}$
			$= \dfrac{n}{n+(1-t_{HS})/t_{HS}}$

Some more Quantitative Genetics theory

Using a progeny mean

var(mean) =   common variance + specific/n

equal to all   Specific bit is averaged out

Mean performance of n progeny (= HS mean):

$P_{HS} = \quad$ ½ A_{sire} + {½ A_{dam} + *Mendelian Sampling term* + *Residual*} /n

$Var(P_{HS}) = \quad$ ¼ V_A + (¾ V_A + V_E)/n $\qquad = t. \, \sigma_p^2 + (1-t) \, \sigma_p^2/n \qquad t = $ ¼h^2

Weight for Mean of n Progeny

$$\underline{\text{Cov}} \qquad = \qquad \underline{\text{½ .n.}h^2} \qquad = \underline{2 \, nt} = \qquad \underline{2 \, n} \qquad = \frac{2n}{n + \lambda}$$
$$\text{Var} \qquad\qquad 1+(n-1)\text{¼}h^2 \qquad (1-t) + nt \qquad n+ (1-t)/t$$

alternatively

$$\frac{\underline{\text{Cov}}}{\text{Var}} \qquad = \frac{\frac{1}{2}V_A}{\frac{1}{4}V_A + (\frac{3}{4}V_A + V_E)/n} \qquad = \frac{2n}{n + (\frac{3}{4}V_A + V_E)/\frac{1}{4}V_A} \qquad = \frac{2n}{n + \lambda}$$

How to combine information from relatives?

Animal	X1 own perform.	X2 sire's perform.	EBV Index
1	3.6	9.2	?
2	-8.3	3.2	?
3	3.5	-15.4	?
4	16.7	7.1	?
5	-14.4	-4.9	?
6	6.6	16.1	?
7	-1.4	-12.2	?
8	-6.2	8.7	?
9	6.2	7.4	?
10	15.0	-4.0	?

Who should we select, animal 10 or animal 6?

Combining different sources of information

Selection Criteria ⬛⬛ ➡️ Breeding Objective

Own performance P_1

Performance on sire P_2

Breeding Value

Performance on sibs P_3
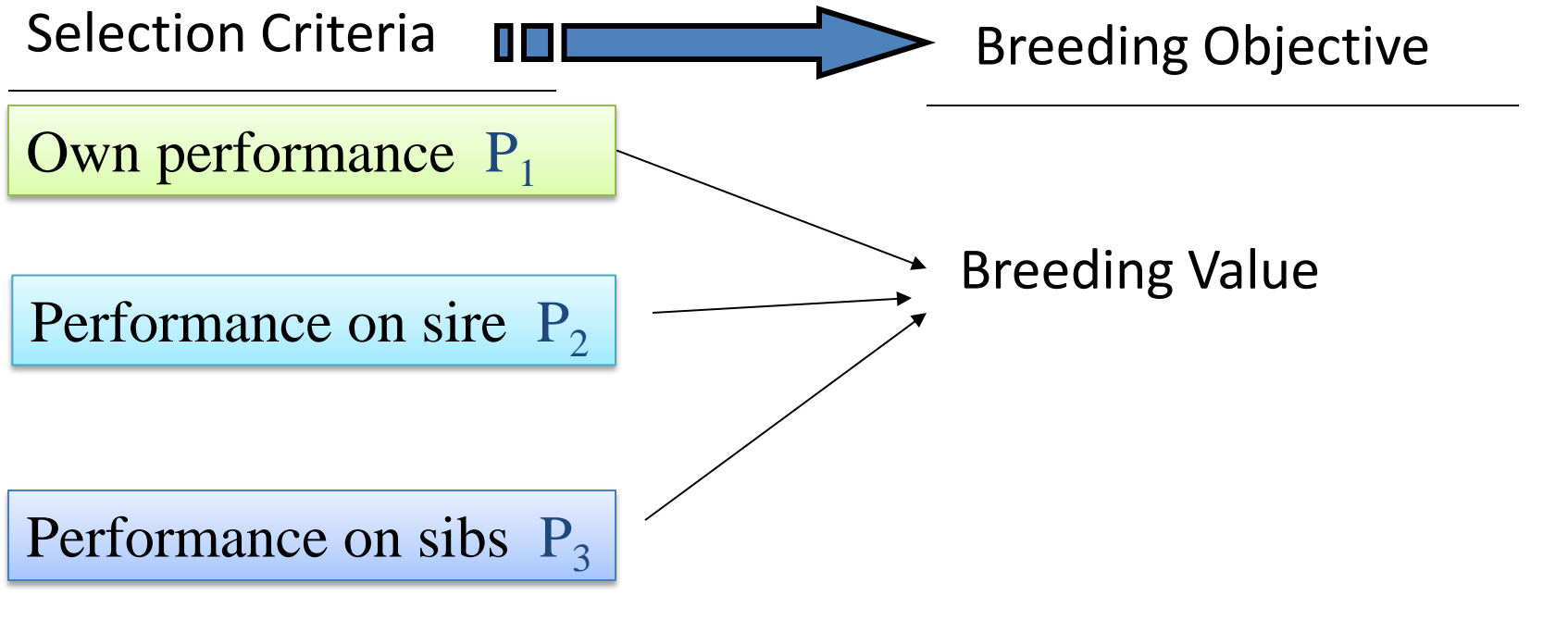
Selection Index (multiple regression)

$$EBV = Index = b_1P_1 + b_2P_2 + b_3P_3 + ... + b_nP_n$$

Selection index with more information sources

multiple regression: Index = $b_1 X_1 + b_2 X_2$

Need to account for covariance between various sources of information

X = vector with information sources

(phenotypic deviations, or P_i = Sel. Criteria, we call these now X = [X_1 X_2])

A = True breeding Value

$$\text{var(X) = P - matrix = P = } \begin{bmatrix} \text{var}(X_1) & \text{cov}(X_1, X_2) \\ \text{cov}(X_2, X_1) & \text{var}(X_2) \end{bmatrix}$$

$$\text{cov(X, A) = G - vector = G = } \begin{bmatrix} \text{cov}(X_1, A) \\ \text{cov}(X_2, A) \end{bmatrix}$$

Index weights $\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ = $P^{-1}G$ is like "cov(X,A)/var_X"

This is prediction of a single trait BV, tomorrow MT prediction

Example

X1 = Own Performance
X2 = Sire's Performance

				P		-1		G				b	
h2	0.3			100	15			30		=		0.284	
varP	100			15	100			15				0.107	
varA	30												
varE	70												

How do we use such weights?

weights	0.284	0.107

Animal	X1 own perform.	X2 sire's perform.	EBV Index
1	3.6	9.2	2.00
2	-8.3	3.2	-2.01
3	3.5	-15.4	-0.67
4	16.7	7.1	5.50
5	-14.4	-4.9	-4.63
6	6.6	16.1	3.60
7	-1.4	-12.2	-1.72
8	-6.2	8.7	-0.82
9	6.2	7.4	2.54
10	15.0	-4.0	3.83

Var-covar of actual x1 and x2

98	17
	104

This is like the P-matrix but based on realized values, usually we use population parameters (phenotypic (co-)variances)

e.g. use STEBVaccurcay.xls

Single Trait selection index calculation

Parameters	
Heritability	0.25
Repeatability of subsequent records	0.5
c-squared (among full sibs)	0

Information used	Nr.Records
nr of own records	1
nr. of records on dam	1
nr of records on sire	1
nr of full sib records	2
nr. of half sib records (excl. full sibs)	20
nr. of progeny	20

Run	Index weight
	0.111
	0.048
	0.025
	0.063
	0.179
	0.887

Why use selection index?

- BLUP does it 'automatically'

  but

- Can use selection index theory to work out:

  - Accuracy for a given set of information
  - The weighting of various information sources
  - The 'value' of various information sources
  - The weighting on own vs family information
  - Correlation between EBV of sibs

Single Trait selection index calculation

STEBVaccurcay.xls

Parameters	
Heritability	0.25
Repeatability of subsequent records	0.5
c-squared (among full sibs)	0

Run

Information used	Nr.Records
nr of own records	1
nr. of records on dam	1
nr of records on sire	1
nr of full sib records	2
nr. of half sib records (excl. full sibs)	20
nr. of progeny	20

Index weight		value of variate:
0.111	own	3%
0.048	dam	1%
0.025	sire	0%
0.063	FS	1%
0.179	HS	1%
0.887	Prog	22%

Accuracy of EBV **0.817**

correlation EBV FS	0.525
correlation EBV HS	0.418

Selection Index (single trait objective)

First summarize some definitions

$$I = b_1 x_1 + b_2 x_2 + ... + b_n x_n \qquad var(X) = P$$

Single trait breeding objective:

$H = A$ (breeding value)

$Cov(X, A) = G$ (a vector with ST objective)

Optimal weights are $b = P^{-1}G$

$Var(I) = var(b'X) = b'var(X)b = b'Pb = \sigma_I^2$

$Var(H) = \sigma_a^2$

$Cov(I, H) = cov(b'X, A) = b'cov(X, A) = b'G = b'Pb = \sigma_I^2$

$$b = P^{-1}G \rightarrow Pb = G$$

Accuracy of selection index (single trait)

r_{IA} = correlation between Index (=EBV) and A

$$= \frac{\text{cov}(I,A)}{\sigma_I \, \sigma_A} = \frac{\sigma^2_I}{\sigma_I \, \sigma_A} = \sigma_I / \sigma_A = \sqrt{(b'Pb / \sigma_a^2)}$$

Because cov(I,A) = var(I)

Selection Index = Best Linear Prediction BLP

Index (I) is best estimate of breeding value: I = E(A|X) = cov(X,A)/var(X)

Same as BLUP, but without fixed effects.

$\text{Var}(I) = \text{var}(EBV) = r_{IA}^2 \cdot \text{var}(BV) = r_{IA}^2 \, \sigma_a^2$ r_{IA}^2 also known as reliability

accuracy of EBV = correlation with True BV

TrueBV | **TrueBV**
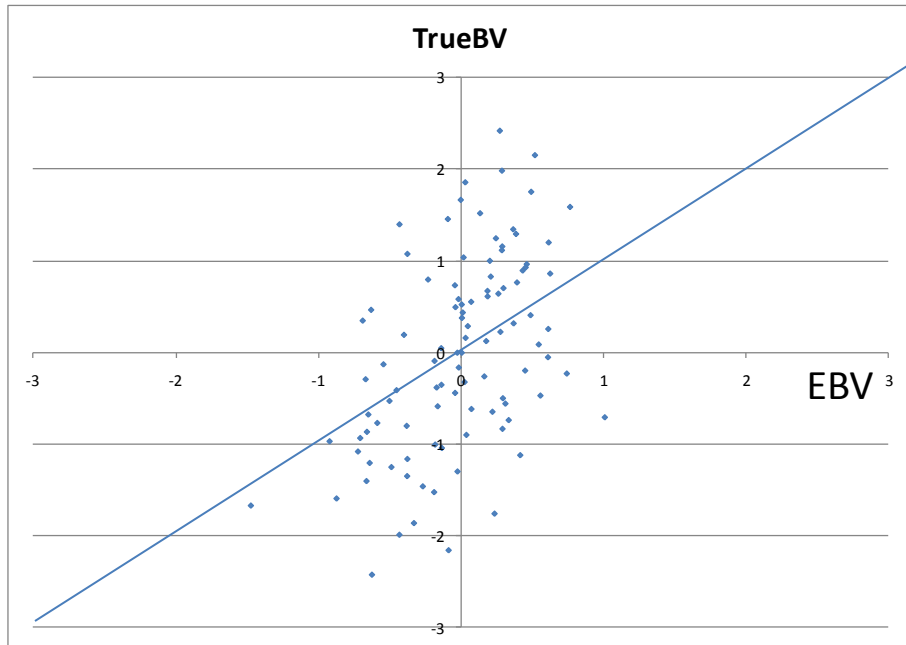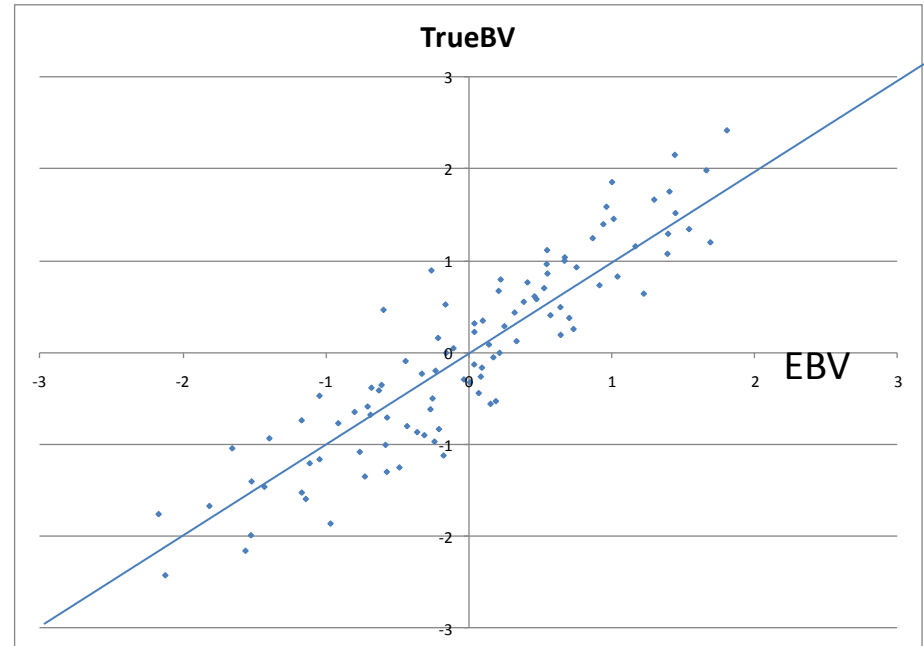
Accuracy = 45%   Accuracy = 90%

$$\text{Regression A on I} = \text{cov}(I,A) / \sigma^2_I = 1$$

Because cov(I,A) = var(I)

Suppose   EBV_A   +56   r = 0.50
                 EBV_B   +56   r = 0.95
                 select A or B?

Examples of accuracies

	h²=0.1	h²=0.3	
own information	0.32	0.55	*equal to sqrt h²*
mean of 10 half sib	0.23	0.33	
mean of 1000 half-sibs	0.49	0.50	*max is sqrt 0.25 = 0.5*
mean of 1000 full-sibs	0.70	0.71	*max is sqrt 0.5 = 0.71*
mean of 100 progeny	0.85	0.94	*max approaches 1.0*

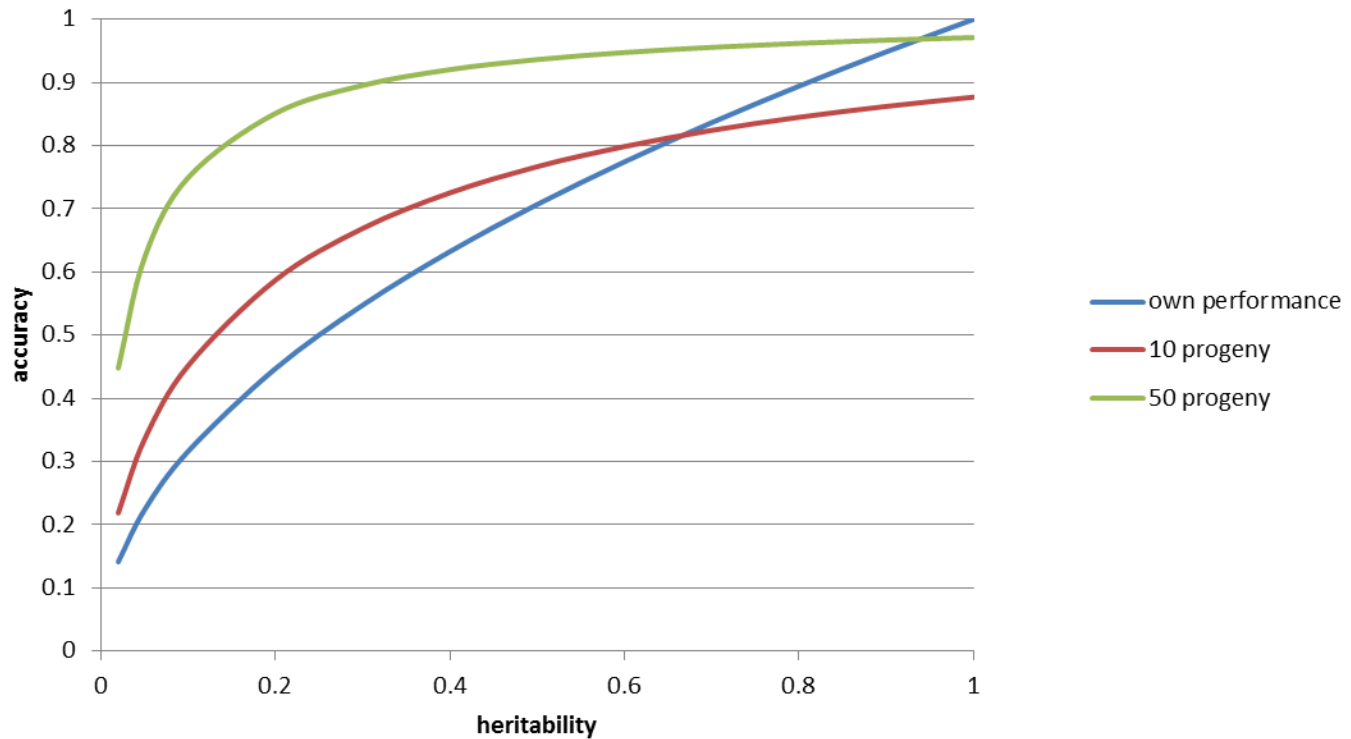Accuracies of animal increase as they get older → (more info)

Accuracy of predicting a breeding value
-            increases as an animal gets older
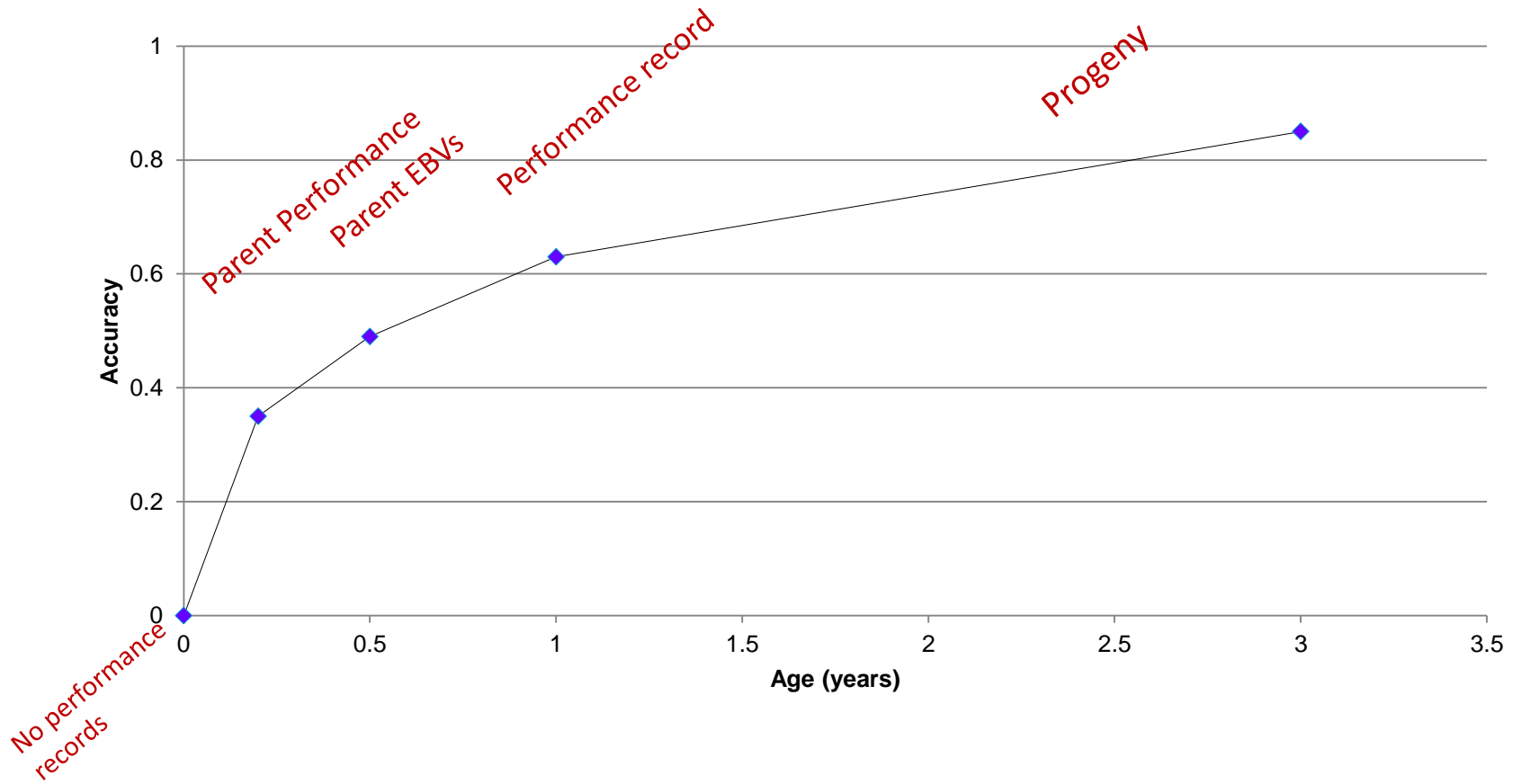
Assumed heritability =   25%

Accuracy of own performance vs progeny test

Progeny test gives usually more accurate EBV then own performance

EBV properties:

Variance of EBV's *- how much they differ*

- $\text{Var(EBV)} = r^2_{IA} \, \sigma_a^2$

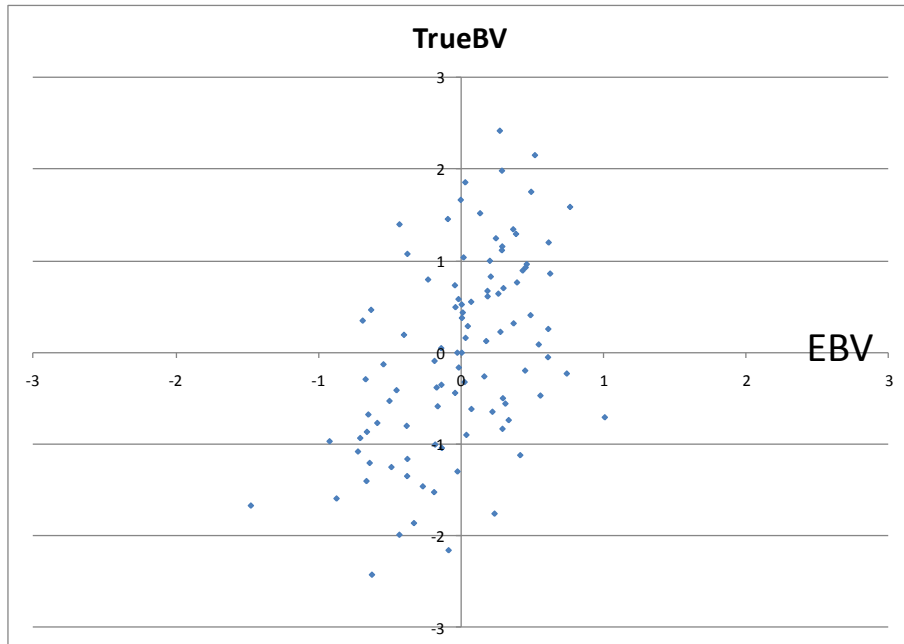- $\text{SD(EBV)} = r_{IA} \, \sigma_{A} =$ Standard deviation of EBVs

Simply a function of accuracy

Note the extremes of Var(EBV) if $r_{IA} = 0$ vs $r_{IA} = 1$
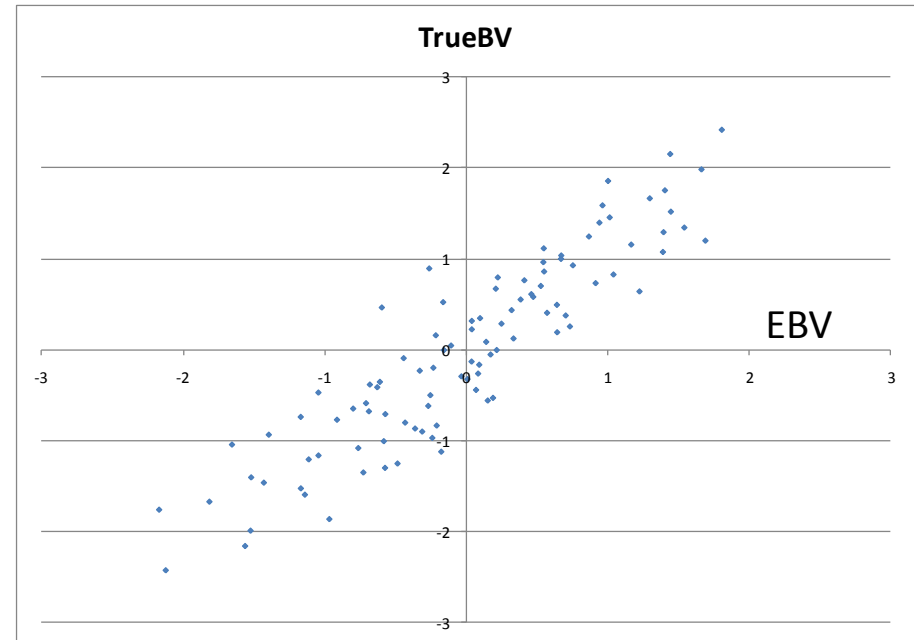
Important when selection across age classes

145   195   220   **proven sires**

Truncation Point

175   195   210   **young sires**
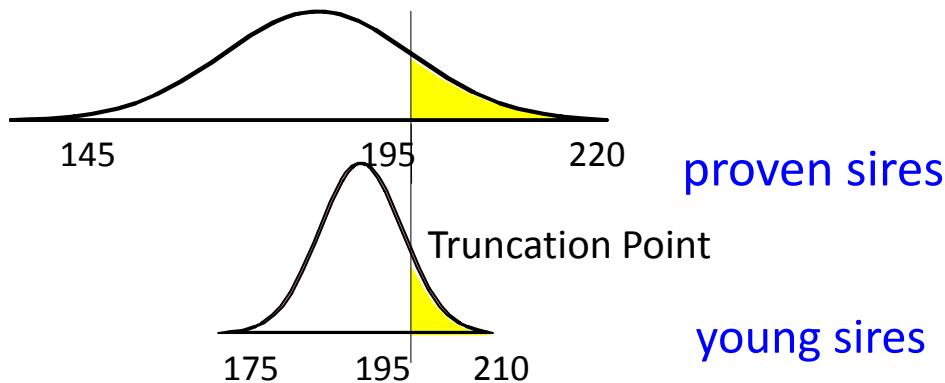
EBV properties

Accuracy = r   = 0.45
Reliability = r^2 = 0.2025

Var(EBV) = low = $r^2 V_A$

Regression BV on EBV = 1

Accuracy = r   = 0.90
Reliability = r^2 = 0.81

Var(EBV) = high = $r^2 V_A$

Regression BV on EBV = 1

EBV properties:

Prediction Error Variance *- how much they still may change*

- PEV = var(EBV-TBV) = $(1-r^2_{IA})V_A$ Prediction Error Variance

- SEP = sqrt(PEV) = $\sqrt{(1-r^2_{IA})}\sigma_A$ Standard Error of Prediction

- EBV$\pm$ 1.96.SEP   Confidence Interval
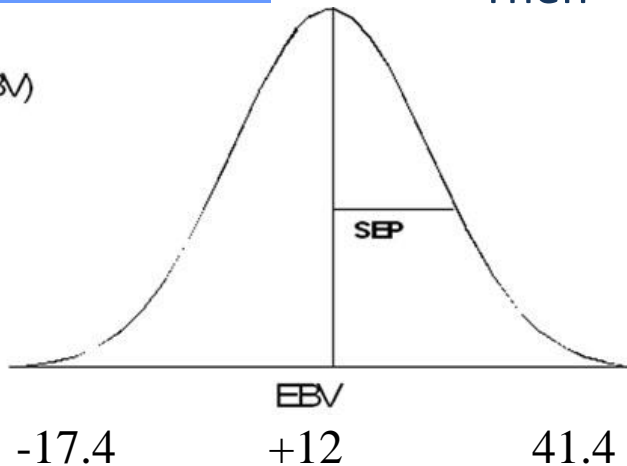
Let σ_a= 19 Kg, EBV = +12 Kg,

Probability density of TBV |EBV

Then

prob(TBV)

SEP

EBV

-17.4   +12   41.4

prob(TBV)

SEP

EBV

-4.6   +12   28.6
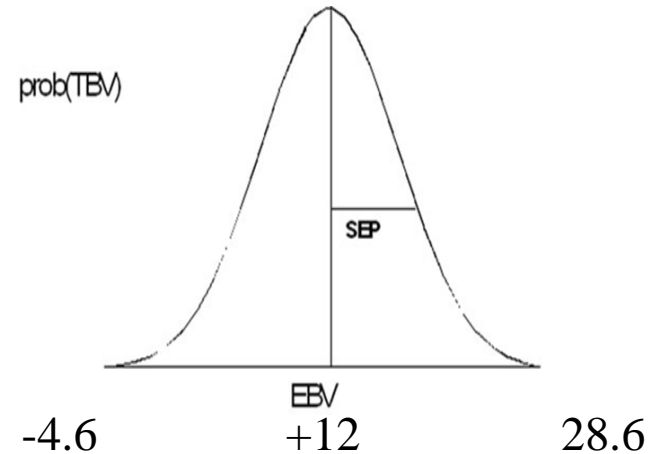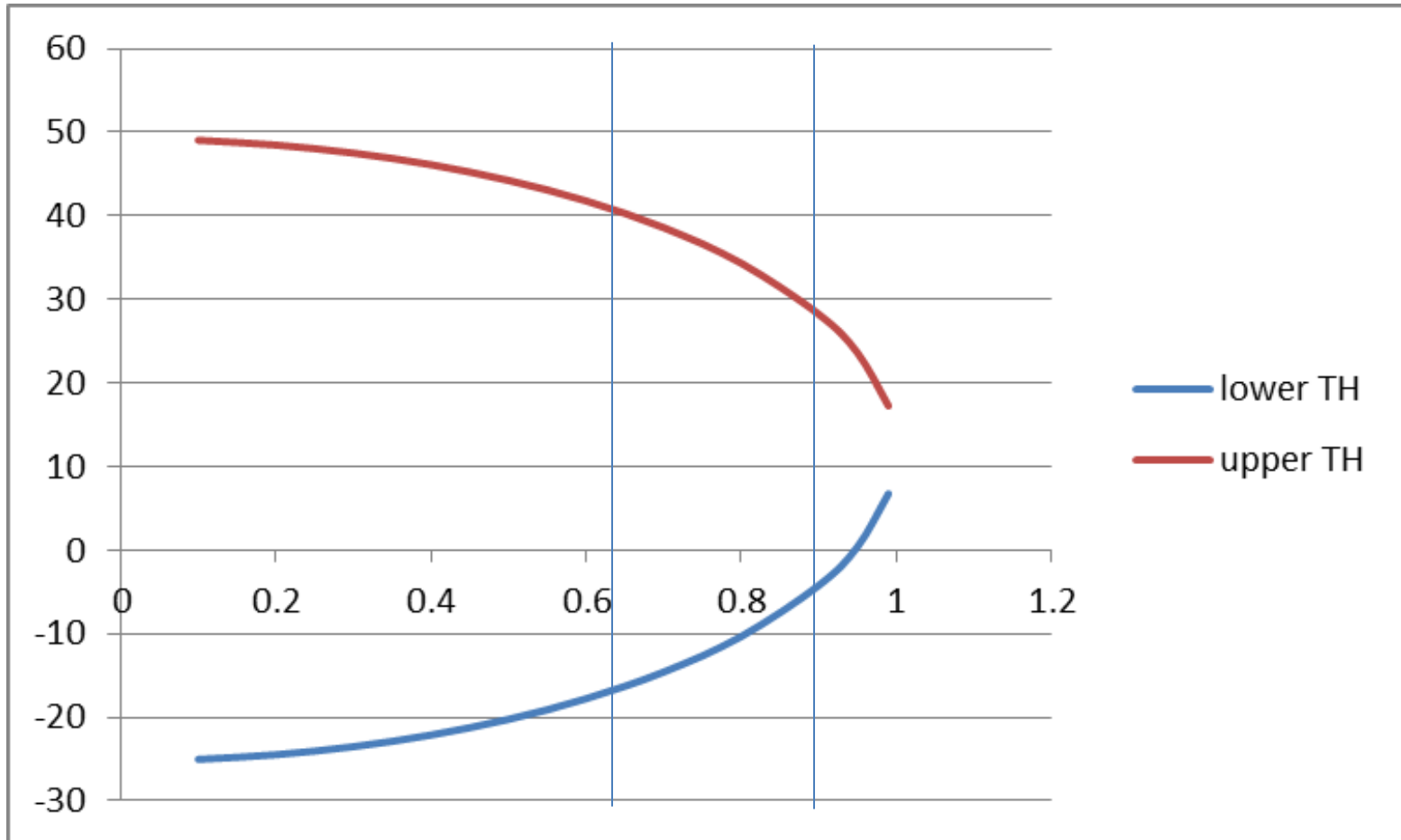
Own record: $r_{IA} = \sqrt{h^2} = 0.63$ Progeny: $r_{IA} = \sqrt{0.81} = 0.9$

Confidence interval of TBV|EBV

Selection on EBV vs inbreeding

X1 = Own Performance
X2 = Sire's Performance

		P		-1	G		b
h2	0.3						
varP	100	100	15		30	=	0.284
varA	30	15	100		15		0.107
varE	70						

Selection index weights for

heritability	X1 own perform.	X2 sire's perform.
0.1	0.098	0.045
0.3	0.284	0.107
0.5	0.467	0.133

Observe the weight on sire information, depending on h²

Lower h² means more emphasis on family information

Selection on EBV vs inbreeding

Simulation: *Belonsky and Kennedy, 1989*

Selection on INDividual performance vs selection on BLUP

Heritability	Increase in genetic merit after 10 years of selection			Inbreeding Increase	
	IND	BLUP	BLUP/IND	IND	BLUP
0.1	0.78	1.41	1.81	0.174	0.383
0.6	5.16	5.31	1.03	0.205	0.293

1. BLUP   more   response than phenotypic selection (IND)
   Difference is larger for smaller heritability.

2. BLUP   more inbreeding
   Especially for low heritabilities

   Co-selection of relatives

Selection on EBV vs inbreeding

Not only h² but also information available affects co-selection of relatives

h² =0.5

Information used	Nr.Records
nr of own records	0
nr. of records on dam	1
nr of records on sire	1
nr of full sib records	0
nr. of half sib records (excl. full sibs)	40
nr. of progeny	0

Index weight		value of variate:
-	-	-
0.250	dam	20%
0.065	sire	1%
-	-	-
0.741	HS	15%
-	-	-

Accuracy of EBV **0.585**

correlation EBV FS **1.000**
correlation EBV HS **0.633**

STEBVaccurcay.xls

h² =0.5

Information used	Nr.Records
nr of own records	1
nr. of records on dam	1
nr of records on sire	1
nr of full sib records	0
nr. of half sib records (excl. full sibs)	40
nr. of progeny	25

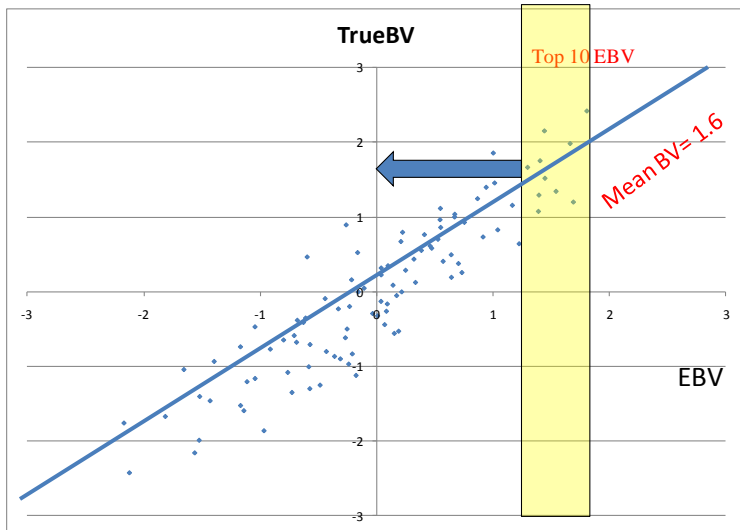| Index weight		value of variate:
0.164	own	2%
0.062	dam	0%
0.016	sire	0%
-	-	-
0.185	HS	0%
1.172	Prog	15%

Accuracy of EBV **0.914**

correlation EBV FS **0.516**
correlation EBV HS **0.455**

Note also that value of family info can be reduced due to Bulmer effect (reduced variance due to selection) – see later
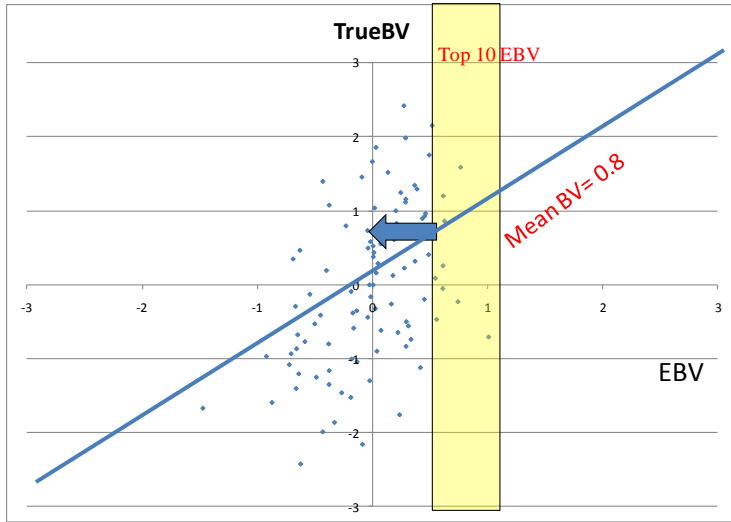
Predicted Response

i = selection intensity (standard normal)

Regression of A on EBV = 1
i.e. slope is the same for accurate and inaccurate EBVs, see left

select on EBV's:

Response = i * SD(EBV)

$R = i * r * \sigma_a$

Summary of this lecture

- Selection Index Theory can be used to work out weights and accuracy for a given set of information about an particular EBV

- Quantitative Genetic Theory and matrices (P, G) are used to work out such index weights (b) and accuracies

- In Genetic Evaluation we use BLUP where this all occurs 'automatically'

- Selection Index Theory still useful to predict what happens
  - Accuracy for a given amount of information
  - Accuracy if using genetic marker information
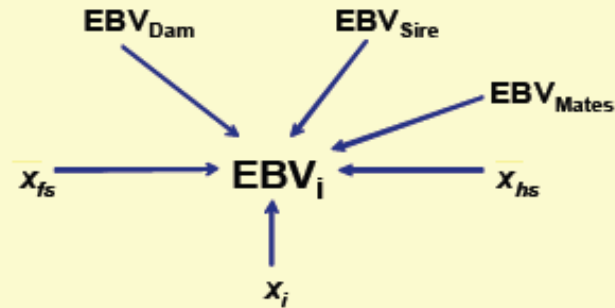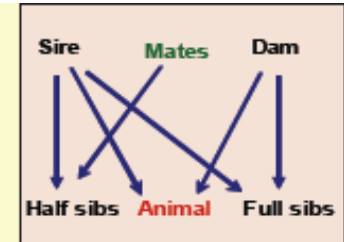 - Importance of own vs family information for given situations

Modeling selection on Animal Model BLUP

Animal model EBV
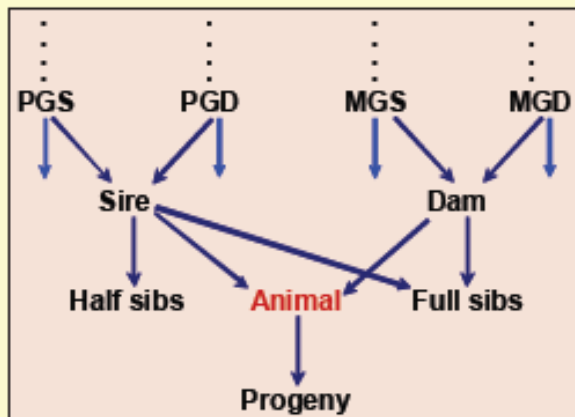
Try to fit 'all relatives'
in selection index
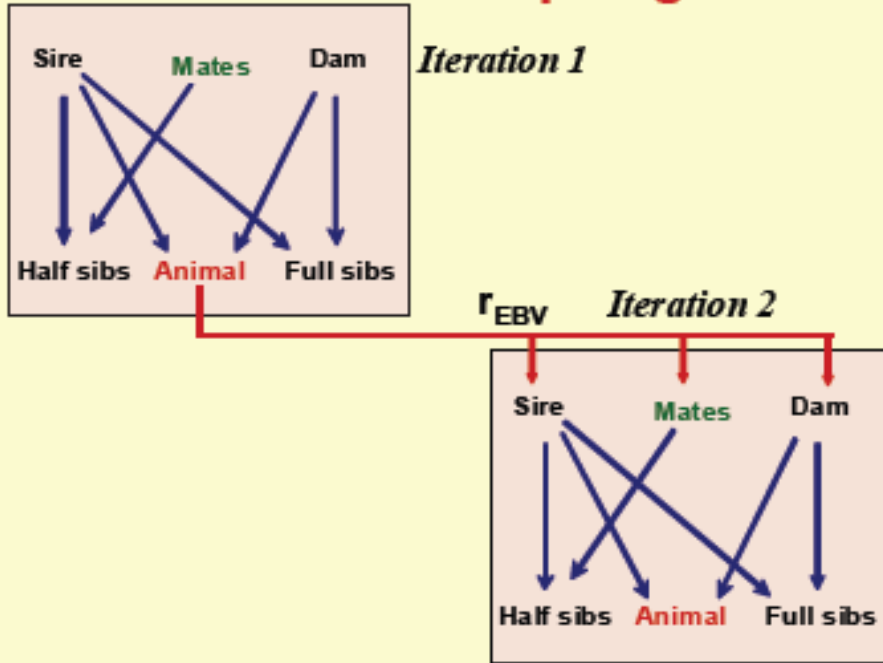

Pseudo-BLUP
Approximation of BLUP
EBV by Selection Index

Fit sire and dam EBV
mates of sire (HS-dams)
Half sibs
Full sibs
Own info

Iterate

Pseudo BLUP

Iteration builds pedigree info

Iteration 1

Sire — Mates — Dam

Half sibs — Animal — Full sibs

r_{EBV} *Iteration 2*

Sire — Mates — Dam

Half sibs — Animal — Full sibs

BLUP-EBV.xls
STEBVaccurcay.xls

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

$$EBV_I = b' \, [x_I, \overline{x_{fs}}, \overline{x_{hs}}, EBV_{Dam}, EBV_{Sire}, EBV_{Mates}]'$$

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

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