Department of Animal Sciences

Quantitative Genetics and Animal Breeding

Lecture notes for ANSC 241

Prof. Dr. A. K. Kahi

May 2013

Egerton University

ANIMAL BREEDING IN A NUTSHELL:

WHERE TO GO?	**OBJECTIVES** -What types of animals to breed for -Economic value of commercial traits
HOW TO GET THERE?	**QUANTITATIVE GENETICS** **SELECTION THEORY** **CROSSING THEORY** -Which animals to breed from Mate allocation
GETTING THERE	**IMPLEMENTATION** -Education -Industry structures -Business structures

Overview of Breeding Objectives

You should develop a feel for objectives in the Animal Production Lectures. The aim is to improve traits of commercial importance. We usually aim at more than one trait by converting each to dollar value per head and making a total score. The objective is to find animals with the best ***breeding values*** for this overall score. An example of how to construct such an overall score (or an overall index) is shown with units:

10 x Fleece Weight + -2 x Fibre Diameter + 1.25 x Body Weight				
KSh/kg. kg	+	KSh/µ.µ.	+	KSh/kg. kg

Note that the overall units are Kenya Shillings.

The economic value of improving fleece weight on sheep by 1kg is at KSh 10 (per animal) and economic value of one unit fibre diameter has a cost of KSh 2.

However, breeding objectives often involve pitfalls. For example. There is widespread effort to increase growth rate, and yet this tends to make animals larger at all stages of growth larger, faster growing animals eat more food.

GENERATION INTERVAL EXAMPLE: Fleece weight in sheep.

Consider a 1000 ewe flock with an age structure typified by the numbers in the table below. Notice that, fairly typically, we keep rams 2 years, keep ewes 6 years, and drop first progeny at 2 years. Mating ratio is 1 ram to 50 ewes, and there is some mortality.

The question is, what is the predicted response to selection per year

Assume. Heritability 0.3
Standard Deviation 0.4kg
Weaning rate 0.8

Age at drop of progeny	**2**	**3**	**4**	**5**	**6**	**7**	**TOTAL**
No. of rams:	**12**	**8**					**20**
No. of ewes:	**250**	**200**	**180**	**150**	**120**	**100**	**1000**

Calculation of predicted response per year to selection

$L_m = \frac{12x2+8x3}{12+8}$ = 2.4 years = average age of rams 'dropping' progeny

$L_f = \frac{250x2+200x3+180x4+150x5+120x6+100x7}{250+200+180+150+120+100}$ = 3.99 years

1000 ewes give 0.8 x 1000 = progeny, 400 male and 400 female

p_m = 12 young rams selected out of 400 available = 12/400 = 0.03

From tables, p_m =0.03 gives i_m = 2.268

p_f = 250/400 = 0.625 giving i_f = 0.607

$R_{yr} = \frac{i_m + i_f}{L_m + L_f} h^2 \sigma_p = \frac{2.268+0.607}{2.4+3.99}$ 0.3 x 0.4 = 0.054 kg

R_{year} = 0.054 kg increase in fleece weight predicted per year

We have now tools to compare alternative selection programs!

An alternative to the previous program is to cull older ewes and retain more young ones

Lower female selection intensity	decreases R_{yrear}
Shorter generation interval	increases R_{year}

USES OF REPEATABILITY

1 As $h^2 = V_A/V_p$ and r is as above, an estimate of repeatability can act as an upper estimate of repeatability. As a heritability estimate requires much more data, this id a useful feature

2. Knowledge of repeatability tells the value of taking repeated measures, as shown next.

VALUE OF TAKING REPAETED MEASURES - The variance of the mean of a number of measures on each animal has a lower phenotypic variance, as the influence of unwanted temporary environmental effects are diluted:

$V_p = V_G + V_{Eg} + V_{Es}$ Vp with measurement per animal.

$V_{p(n)} = V_G + V_{Eg} + V_{Es}/n$ Variance of mean of n measurements of a trait on each animal

$V_{P(n)} = \left(r + \frac{1-r}{n} \right) V_P$ Showing the proportional reduction in V_P

$\sigma_{P(n)} = \sqrt{\left(r + \frac{1-r}{n} \right)} \sigma_p$ Showing the proportional reduction in σ_p

SELECTION STRATEGY. Take a number of measures on each animal, reducing the influence of unwanted temporary effects, to give a better prediction of each animal's 'underlying phenotype'. Select animals on the mean of n measures.

Effect on selection response

Response with one measure = $R = ih^2\sigma_p = i\frac{V_A}{V_P}\sigma_P = i\frac{\sigma_A}{\sigma_P}\sigma_A$

Response with n measures = $R_n = i\frac{\sigma_A}{\sigma_{P_n}}\sigma_A = \frac{ih^2\sigma_P}{\sqrt{\left(r + \frac{1-r}{n} \right)}}$

Which gives $R_n = R\frac{\sigma_P}{\sigma_{P_n}}$

So the proportional decrease in phenotypic standard deviation tells us the proportional increase in response.

REPEATABILITY EXAMPLES

Fertility in cattle r =0.2 Mean =0.8 σ_p=0.4 V_p=0.16

$V_{p(2)} = V_p = [0.2 + 0.8/2]\ 0.16 = 0.096$

$$V_{P(2)} = \left(r + \frac{1-r}{n}\right)V_P = (0.2 + 0.8/2) = 0.096$$

$$R_n = R\frac{\sigma_P}{\sigma_{P_n}} = R\sqrt{(0.16/0.096)} = 1.29R$$ 29% more response

Body weight in cattle:

r = 0.95 Mean =300Kg σ_p=30kg V_p=900

$$V_{P(2)} = \left(r + \frac{1-r}{2}\right)V_P = (0.95 + 0.05/2) = 877.5$$

$$R_n = R\frac{\sigma_P}{\sigma_{P_n}} = R\sqrt{(900/877.5)} = 1.01R$$ 1% more response.

MESSAGE : If a trait is highly repeatable, then taking extra measures of it adds little information. However, lowly repeatable traits yield useful new information with each new measurement.

GENOTYPE X ENVIRONMENT INTERACTION

The mean performance of any one breed depends on the environment which it is farmed	This is an effect of environment
The mean performance of some breed is higher than that of others	This is an effect of genotype.
The ranking of breeds on performance can depend on environment .	This is Genotype x environment interaction.

→ as V_E decreases and both **heritability** $h^2 = \frac{V_A}{V_A + V_E}$ **and response increase.**

How do we account for identifiable environmental effects?

Fixed effects e.g.: Birth type, herd, management group.

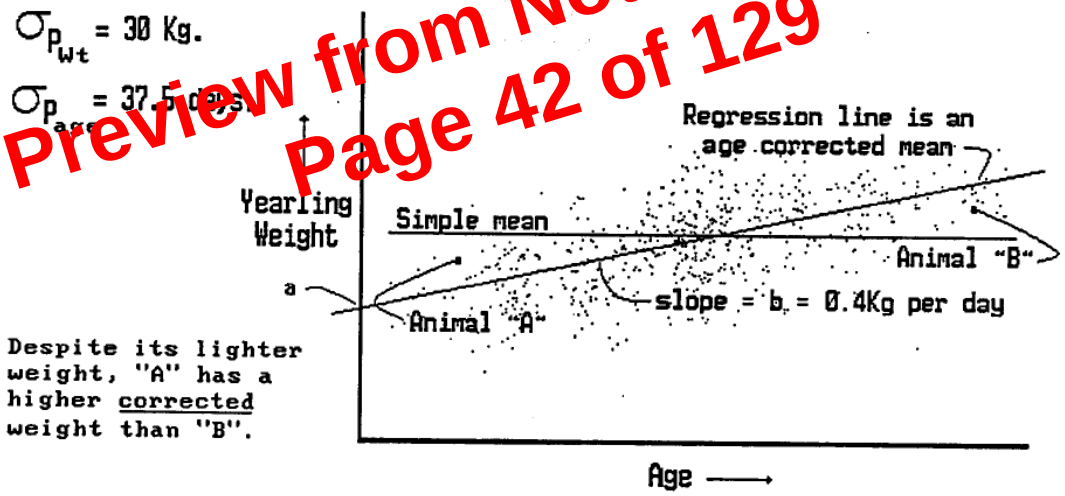
Example:

Birth Type	Mean Weaning Weight
Single	**25**
Twin	**23**
Triple	**20**

Strategy: Express all phenotypes as deviations from their group means.
A 25kg. Twin is given a value of 25-23 =+2kg.
but a 25kg Triple is given a value of 25-20 = +5kKg

Continuous effects e.g.: age of individual, lactation length.

Example: 'Yearling' weight in cattle vs. age at weight:



Strategy: express all phenotypes as deviations from the "age-corrected mean":

Example:

The 'yearling' of Alice (Animal A) is 280kg at 11 months of age,
The 'yearling weight' of Bessy (Animal B) is 295kg at 13 months of age.

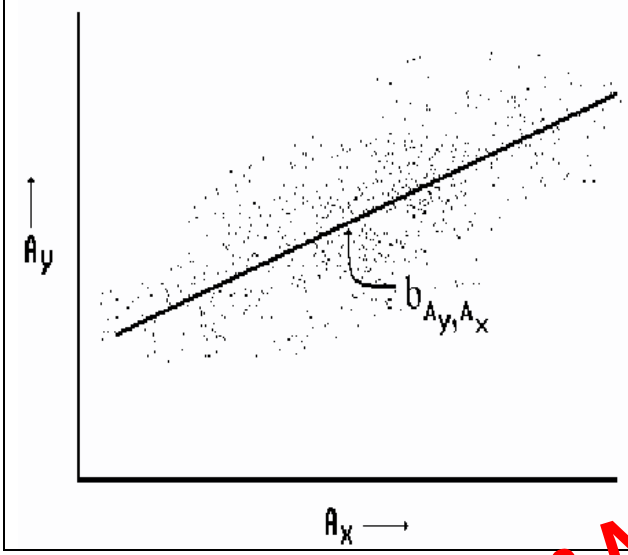
Is Bessy better than Alice?

Correlated response

When selection is for a given trait x, we can also expect an effect on trait y if y is correlated to x. The response in a trait correlated to a trait under direct selection, is termed *correlated response*.

Correlated Response (CR_y) - Response in trait y due to selection on trait x

If we knew animals' breeding values, we could draw the relationship between breeding values for two traits, say trait x = growth and y = fat depth.



From the simple scatterplot of animals' values, we could predict (ˆ) breeding value for fat depth (trait y) from breeding value for growth (trait x):

$$\hat{A}_y = b_A A_X$$

Now if we selected on growth, then animals' breeding values for growth (A_x) would increase over time, and we would have a predicted response per generation:

$$R_x = i_x.h^2_x.\sigma_{Px}$$

The predicted increase in breeding values for fat depth (trait y) is then the correlated response. As the breeding value of y can be predicted from the BV of x, so can the Response for y (average of BV in selected group) be predicted from the response for x.

Hence, just as $\hat{A}_y = b_A A_X$ then $CR_y = b_A R_x$.

Now we can express this in terms we have already handled:

Correlated Response =

$$CR_y = b_A R_x = r_A \frac{\sigma_{Ay}}{\sigma_{Ax}} R_x \qquad \left[\text{as } b_A = \frac{\text{Cov}(A_x,A_y)}{\sigma^2_{A_x}} \text{ and } r_A = \frac{\text{Cov}(A_x,A_y)}{\sigma_{A_x}\sigma_{A_y}}\right]$$

$$CR_y = r_A \frac{\sigma_{A_y}}{\sigma_{A_x}} i_x h^2_x \sigma_{P_x} \qquad \left[\text{as } CR_y = b_A R\right]$$

$$CR_y = r_A \frac{h_y \sigma_{P_y}}{h_x \sigma_{P_x}} i_x h^2_x \sigma_{P_x} \qquad \left[\text{as } \sigma_A = h\sigma_P\right]$$

We now have a basis to describe what proportion of GENES and GENOTYPES relatives share in common- to what extent they are genetically similar. Here are two more examples:

EXAMPLE- if **X** and **Y** are **half sibs**, consider one **locus**

Mum	**Dad**	**Mum**
$\mathbf{M_1M_2}$	$\mathbf{D_1D_2}$	$\mathbf{M_3M_4}$

X	**Y**
$\mathbf{D_1M_1}$	$\mathbf{D_1M_3}$
$\mathbf{D_1M_2}$	$\mathbf{D_1M_4}$
$\mathbf{D_2M_1}$	$\mathbf{D_2M_3}$
$\mathbf{D_2M_2}$	$\mathbf{D_2M_4}$

Pick a GENE. Chances of **2nd** choice of animal having same **gene** is 0 (if a maternal gene had been picked) or ½ (if a paternal gene had been picked) averaging r = ¼.

Pick a GENOTYPE, chances of 2nd choice of animal having same **genotype** is **u = 0**

EXAMPLE- if X and Y are **parent** and **offspring**.

Parent (X)	**Mum**
X_1X_2	M_1M_2

Y
$\mathbf{X_1M_1}$
$\mathbf{X_1M_2}$
$\mathbf{X_2M_1}$
$\mathbf{X_2M_2}$

Pick a GENE in X, chances of 2nd choice of animal (Y) having same **gene is r = ½.**

Pick a GENE in Y, chances of 2nd choice of animal (X) having same **gene is 0** (if a maternal gene had been picked) or 1 (if paternal gene had been picked) averaging r =½.

Pick a GENOTYPE, chances of 2nd choice of animal having same **genotype** is **u = 0**.

To illustrate this for another example, consider 9 animal to be allocated to three groups. The animals have observed values of 1, 2 or 3 units - three animals of each value. The table shows three ways of allocating animals to groups:

Type of allocation to groups.

	'Dead fair'			Random			Full Similarity		
Individual values	**1**	**1**	**1**	**2**	**2**	**1**	**1**	**2**	**3**
	2	**2**	**2**	**3**	**1**	**3**	**1**	**2**	**3**
	3	**3**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
Group Averages	**2**	**2**	**2**	**2.0**	**1.7**	**2.3**	**1**	**2**	**3**

We can use knowledge of similarity (covariance) between group members to

- predict performance of relatives: we expect an animal to be good because its full sib was good
- determine the importance of the effects that cause similarity: common genetic effects to family members is the most important one for us, we want to know what extent difference that we observe are heritable.

Genetic Covariance

We can use genetic relationship among relatives to predict how much they 'look alike' phenotypically. For a quantitative trait, this is measured by *covariance*.

As variance is an indication of the extend that unrelated individuals can be different from each other, covariance is a measure to what extend two individuals would have the same value. Suppose individual y has an extreme value, a covariance would tell us what extend another individual should also have a similar extreme value

Note that covariance is used to indicate similarity between individuals (for the trait). Earlier, we used covariance to indicate similarity between traits (for the same individual). There is a parallel here: genetic correlations indicate to what extend the genes for two different trait are common. A genetic relationship indicates to what extend two individuals have genes in common.

The genetic covariance between two individuals is equal to their genetic relationship times the genetic variance

additive genetic covariance between individuals x and y: $Cov(A_x,A_y)=a_{xy}\ V_A$

dominance covariance between individuals x and y: $Cov(D_x,D_y)=d_{xy}\ V_D$

genetic covariance between individuals x and y: $Cov(A_x,A_y)+Cov(D_x,D_y)$

$$3 \text{ x } \left(\sigma_B^2 + \frac{\sigma_W^2}{n} \right) = \sigma_W^2 + 3\sigma_B^2$$

We can estimate the contribution of the variance of temporary effects within groups by taking all deviations within groups (we estimate the group mean and take the deviation of each record from each group mean). These deviations are called *residual effects* and if we square all these within group deviations, we obtain the *residual sums of squares*. If the residual sums of squares are divided by the number of residuals that we can compare (this is the *degree of freedom* for the residual) than we obtain an estimate of the residual variance: σ^2_W for this example:

Analysis of the variance example data set

Effect	Degr. of Free	Sums of squares	Mean Squares	Expected mean squares
Mean	1	8640		
Group effect (Between groups)	4	192	48	$\sigma^2_w + 3\sigma^2_B$
Residual (Within groups)	10	18	1.8	σ^2_w
Total	15	8850		

Here is how these figures are calculated….

1) sums of squares due to means: $15 \times 24^2 = 8640$

2) sums of squares due to differences:
$3 \times (21^2 + 25^2 + 29^2 + 19^2 + 26^2) = 8832$
corrected for mean:
$3 \times ((21-24)^2 + (25-24)^2 + (29-24)^2 + (19-24)^2 + (26-24)^2) = 92$
or directly: 8832 - 8640 = 92

3) total sums of squares:
$21^2 + 25^2 + \ldots$ <all individual weighings squared> $\ldots + 26 = 8850$

4) residual sums of squares:
total SS-SS groups = 8850-8832 = 18
notice that also: $(-1)^2 + 0 + (-2)^2 + \ldots (2)^2 = 18$

<u>The estimated variance components for example data set 1:</u>

Between groups	$\sigma^2_B = 15.4$
Within groups	$\sigma^2_w = 1.8$
Total variance is	$\sigma_B^2 + \sigma_w^2 = 17.2$

Repeatability = intra-class correlation = 15.4/17.2 = 0.895

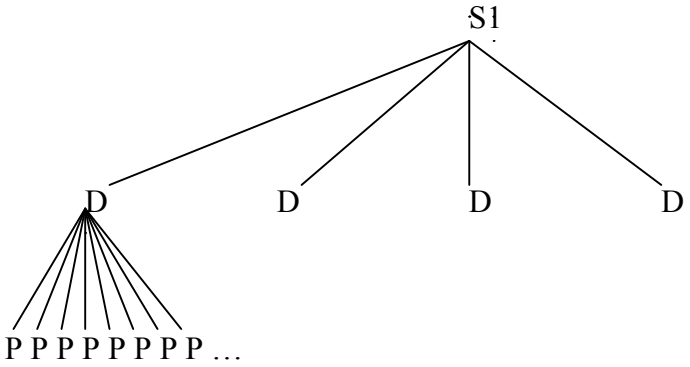
Variance of the group means $\sigma^2_B + \frac{\sigma^2_W}{n} = 15.4 + 1.8/3 = 16$

The heritability is estimated as $\hat{h}^2 = \frac{\hat{V}_A}{\hat{V}_P} = \frac{4\hat{\sigma}_S^2}{\hat{\sigma}_S^2 + \hat{\sigma}_W^2}$

or. h^2= 4 x intraclass correlation between half sibs…

Analysis of full-sib families

A common structure of the data is that we have observation on full sib families, where each sire is mated to more dams, and each dam has more than one offspring. Hence, we have full sib families within half sib families and we can make groups of sire and dam groups, but dams are different for each sires. (This is called a Nested or Hierarchical design: dams are nested within sires).



In the analysis we don't have two, but three variance components: the variance between sires, the variance between dams (within sires) and the variance within dams (within full sib families). This example gives the following table:

SOURCE	EXPECTED MEAN SQUARE
Sires	$\sigma_w^2 + 8\sigma_d^2 + 32\sigma_s^2$
Dams	$\sigma_w^2 + 8\sigma_d^2$
Progeny	σ_w^2

Note: 8 progeny per dam
32 progeny per sire

And the expected value of the variance components is:

Variance due to	Component	Expectation/ interpretation
Sires	σ_S^2	$\frac{1}{4}V_A$
Dams within sire	σ_d^2	$\frac{1}{4}V_A + \frac{1}{4}V_D + V_{Ec}$
Progeny within dam	σ_W^2	$\frac{1}{2}V_A + .75V_D + V_{EW}$
Total	σp^2	$V_A + V_D + V_{Ec} + V_{EW}$
Sires + dams	$\sigma_S^2 + \sigma_d^2$	$\frac{1}{2}V_A + \frac{1}{4}V_D + V_{Ec}$

VA = Additive genetic variance
VD = Dominance variance.
VEc = Common environmental variance for full sibs
VEw = Environmental variance specific for each individual

The intraclass correlation between full sibs is the between group (full sib family)

$$t_{FS} = \frac{\left(\hat{\sigma}_s^2 + \hat{\sigma}_d^2\right)}{\hat{\sigma}_p^2} = \frac{\frac{1}{2}\hat{V}_A + \frac{1}{4}\hat{V}_D + \hat{V}_{Ec}}{\hat{V}_P} \geq \frac{1}{2}h^2 \qquad \text{So } \hat{h}^2 \leq 2t_{FS}$$

Since full sibs have more in common that just genetic effects, their intra-class correlation will overestimate heritability. Only the half-sib correlation can give an unbiased estimate of heritability, since that contain genetic effects only.

Assumptions in such ANOVA estimates of heritability:

1.Randomly chosen sires:

Since the estimate is based on the variance among sires. The variance among a selected group of sires will be smaller. An estimate of heritability based on progeny of selected sire will be biased downward.

2.Randomly allocated dams

3. Equal environment for each progeny group

Estimation of Heritability - by Regression

1. Regression of offspring on **one** parent

What is the covariance between the performance of a sire, and the performance of its offspring?

We expect there will be some covariance, because a sire and its offspring are genetically related. Of all the variation we observe between performances of sire (i.e. the phenotypic variance) we expect the sire only to transfer its genetic effects to its offspring. The random environmental effects of the sire and the offspring are assumed unrelated. Since the sire has only half of its genes in common with its offspring. Therefore, the theoretical expectation between performances of sires, and performances of their offspring is expected to be equal to half the additive genetic variance. The regression of the performances of offspring on performances of their parents is therefore.

Regression of offspring on parent: $b_{OP} = \frac{\text{Cov}(\text{parent,offspring})}{\text{Var(parent)}} = \frac{\frac{1}{2}V_A}{V_P} = \frac{1}{2}h^2$

Therefore, if we calculate the regression of offspring on parents, we know that, based on our quantitative genetic model, this regression should be equal to $\frac{1}{2}h^2$. We can use this knowledge to estimate heritability based on data. We can also use it to predict differences between offspring of two parents. If the parents differ an amount of 40 (say in mature weight) we expect their offspring to differ an amount to $\frac{1}{2}h^2 * 40$.

The regression of performance of offspring on performance of parent is equal to $\frac{1}{2}h^2$

10 USE OF INFORMATION FROM RELATIVES

Introduction

If we want to select the best from individuals of a population in order to achieve genetic improvement, it is important that we are able to rank these individuals based on expected genetic merit. Such an *estimation of breeding value* can be based on individuals' own phenotype, as we saw in chapter 4. However, generally we have more information available.

We expect a bull to be genetically good also if it is an offspring from a very good sire or dam (remember the expected value of progeny was equal to the average breeding value of the parents). In addition, if a bull has good performing sibs, or very good offspring, we tend to give more credit to its breeding value. In general, we can use information from genetically related animals to estimate breeding values. It is not instead of own phenotype, but additional information.

The problem is then to determine how important a good own phenotype is in relation to good phenotypes of related animals. If a trait can be measured on one sex only, we can not use own information at all. Again, we can rely on sibs, but a common source of information about an animal's EBV is also to test a number of its progeny.

In this chapter we will analyse when relatives' information can be important, and how it can be used in estimating breeding value. We will particularly consider information from sibs and information from progeny.

Principle of estimation of breeding values

We would like to rank and select animals on their true breeding values (TBV' or A) but we don't have this perfect knowledge we can not see genes and breeding values. Instead, we must use observed phenotypes to get estimated breeding values (EBV's or A), and accept a slower rate of selection response.

ESTIMATED BREEDING VALUES-EBV's are estimating TRUE BREEDING VALUES using phenotypic information

The most obvious piece of phenotypic information we can use to estimate an animal's breeding value is the animal's own phenotype. We saw in chapter 4:

$\hat{A}_i = h^2 P$ Note that $P = (P_i - \bar{P})$, a deviation.

The principle of breeding value estimation is based on regression. We want to know how much better a breeding value is when we observe a certain phenotypic difference. If we regress the breeding values on phenotypic observation, the slope of the regression line tells us how much difference we have in breeding values per unit of difference in phenotype. This slope is equal to the heritability. Using quantitative genetic theory:

$$b_{xy} = \frac{Cov(x,y)}{var(y)} \text{ which is now equal to } \frac{Cov(P,A)}{Var(P)} = \frac{Var(A)}{Var(P)} = h^2$$

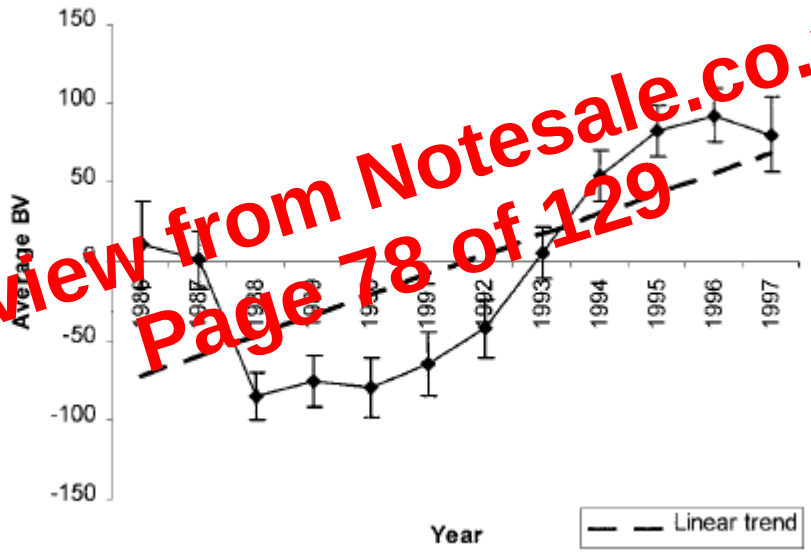
recalling that cov(P,A) = cov(A + E,A) = cov(A,A) = var(A).

But by how much?

In Flock 1, the reference sire's progeny are worse than Flock 2 sire's progeny by 0.5 kg. Assuming many progeny, the reference sire's breeding value inferiority must be twice this, because of the diluting effect of ewe mates of equal merit. So the reference sire is 1kg genetically inferior to Flock 1 sires, and by a similar argument, he must be 1kg genetically superior to Flock 2 sires. Thus, if the flock sires are representative of their flocks (or if they are equally selected) then Flock 1 is 2kg genetically superior to Flock 2. Given the observed average merit of the flocks, the flock 1 environmental effect must be 2.5kg below that of Flock 2.

BLUP can both calculate and use this information automatically whenever there are such genetic linkages available – i.e. whenever relatives are spread across different groups.

3. BLUP gives genetic trends. The approach used in the last example could be used to test the genetic differences between animals born in different years, instead of different flocks. This ability to compare the EBV's of animals born and measured in different years means that a year mean EBV's can be calculated and genetic trends reported. Here is an example from Ojango and Pollot (2001) for the Kenyan Holsteins.



4. BLUP can handle unbalanced designs easily. A selection index using sib information faces the problem that each candidate does not have the same number of sibs (n):

$$\text{Index} = b_f P_f + b_w P_w \quad \ldots \text{from lecture on using sib information}$$

The weight for the family information (b_f) depends, besides on heritability and the type of family, also on the number of individuals in the family.

One solution is to construct an index for each number of sibs involved – but if progeny information is available the same problem exists … BLUP handles this imbalance automatically by constructing a custom selection index for each animal. BLUP only needs to report the EBV's ($\hat{A}$'s) and not the index weights (b's).

of a single additive genetic effect, two different models can be distinguished: **sire model and animal model**. In a sire model only the additive genetic effects of the sires of each individual are considered. With such a model breeding values for only sires can be estimated. With an animal model, breeding values for all individuals can be estimated.

The model to describe the observations contains fixed effects for the environmental factors and random effects for the genetic effects. The model contains both fixed and random effects and is, therefore, referred to as a mixed model. In matrix notation, a mixed model can be expressed as follows:

$$y = X\,b + Z\,g + e$$

where, **y** vector with n observations (n*1),
b vector with f fixed effects (f*1),
g vector with s random effects (s*1),
e vector with error terms (n*1),
X incidence-matrix indicating for each observation the **fixed** effects by which it is influenced (n*f),
Z incidence-matrix indicating for each observation the **random** effects by which it is influenced (n*s).

In brackets the size of the matrices and vectors is given where:
n= number of observations
f= number of fixed effect classes (sum over all factors)
s= number of animals for which breeding value has to be estimate.

Write down the set of so-called least squares equations corresponding to the model
In setting up the least squares equations, all effects in the model (**b** and **g**) are treated as fixed effects. In matrix notation the least squares equations are:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z \end{bmatrix} \begin{bmatrix} b \\ g \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

where **X**, **Z**, **b**, **g** and **y** are as defined above.

Let us look at this expression in more detail. The matrix **X** relates each observation to fixed effect classes. In case we have only one fixed effect with several classes, the matrix **X'X** contains the number of observations in each class at the diagonal and zero elsewhere. With more than one fixed effect the diagonal element still contains the number of observations but the off-diagonal elements are no longer all equal to zero. They represent how observations for one class of a fixed effect are distributed over the classes for the other fixed effects. In other words, **X'X** contains information on number of observations for each fixed effect class. **Z'Z** is a diagonal matrix which contains information on the number of observations for each class of **g**. The matrix **X'Z** (and its transpose **Z'X**) contains the number of observations for all combinations of classes of **b** and **g**.

The size of the matrices can be derived from the size of **X** and **Z** which are (n*f) and (n*s), where the first number represents the number of rows and the second the number of columns. Let us look at the size of **X'Z**: this is the product of a matrix (transpose of **X**) with f rows and n columns (f*n) with a (n*s) matrix. The result is a (f*n)x(n*s)=(f*s) matrix. (Note: matrices can only be multiplied when number of columns in first matrix is equal to number of rows in second matrix). The vector **X'y** contains the sum of the observations in each class in **b** and **Z'y** contains the sum of observations for each class in **g**.

Transform the least squares equations into mixed model equation

In the least squares equations we have not used information on the heritability of the trait or on genetic relationships between animals. To include this information, the part of the left-hand sides of the equations that relates to the genetic effects has to be modified. This modification, which involves adding an additional matrix, depends on:
a) the model (animal or sire);
b) whether or not relationships between animals have to be taken into consideration;
c) heritability of the trait.

For now we look at a sire model without relationships between sires. Later on the other options will be described. The mixed model equations for a sire model without relationships between sires is:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z+\alpha I \end{bmatrix}\begin{bmatrix} b \\ g \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

where **I** is the identity matrix and α is $\sigma_e^2/\sigma_s^2=(1-¼h^2)/(¼h^2)$. In words: from the least squares equations we can get the mixed model equations by adding α to the diagonal elements for the sires. The term α represents the variance ratio of the error in the model and the genetic effect in the model. The genetic effect in the sire model is the sire's transmitting ability (s_i) which is equal to half the additive genetic effect of the sire (a_i). The half results from the fact that the sire only contributes 50% of the genes of the animals on which we have the observations. The variance of the sire effects is: $\sigma_s^2 =var(s_i)=var(½a_i)=¼var(a_i)=¼\sigma_a^2=¼h^2\sigma_p^2$ Consequently, the variance of the error term in the model (σ_e^2) is equal to: $\sigma_p^2-\sigma_s^2=\sigma_p^2 - ¼\sigma_a^2=(1-¼h^2)\sigma_p^2$. By adding α**I** to the **Z'Z** and solving the equation we get BLUP estimates for sires. In principle the effect of this modification is to change the estimate of a sire from a mean of daughter performances (adjusted for fixed effects) to a 'regressed' mean as is also the case with a selection index. In other words, adding α to the appropriate diagonal coefficients has an effect analogous to multiplying the mean of daughter performances (adjusted for fixed effects) by the appropriate selection index weighting factor.

4) Obtain estimates
Estimates for all effects can be obtained by solving the mixed-model equations:

$$\begin{bmatrix} \hat{b} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z+\alpha I \end{bmatrix}^{-1}\begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

In this case of a sire model contains estimates for the genetic effect transmitted by the sire to its offspring, which is also referred to as estimated transmitting ability (ETA). The estimated transmitting ability is half the breeding value of the sire. With an animal model, breeding values (and not transmitting abilities) are estimated for all animals (sire, dams and offspring).
Note: When there is more than one fixed effect in the model, restrictions have to be used to avoid that the matrix is singular, in which case the generalized inverse rather than inverse of the left hand sides should be used.

Example:
We have observation on milk production of cows measured on two different herds. All productions were observed in the same year and all cows had the same age. The cows were progeny of 3 different sires. The objective is to estimate the breeding value of these sires using a sire model. The following data are available:

Deviations from Hardy Weinberg in an inbred population (for single locus –2 allele model)

Genotype	HW-equilibrium (Non-inbred)	Inbred population
A_1A_1	P^2	$P^2 + pqF$
$A_1 A_2$	2pq	2pq(1-F)
$A_2 A_2$	q^2	$q^2 + pqF$
Allele Frequency A_1	p	p

An individual with an inbreeding coefficient F has therefore F % less heterozygosity.
In the following paragraph we see that this has negative consequences, for example in case of genetic defects.

- Inbreeding is temporarily. It is a configuration of genotype frequencies that typically has more homozygotes (of either kind). However, as soon as different inbred strains cross, the inbreeding is completely disappeared.
- Of course, if we had no 'other' lines, an inbred populations might fix its genes due to drift (or due to selection if selected), thereby loosing its genetic variation. In that case, inbreeding is not so temporarily.

Consequences of Inbreeding

Why is inbreeding bad?

1) Increased frequency of affected individuals due to genetic defects'
Inbreeding increases the frequency of homozygotes. This is a disadvantage, since many mutations that occur have a negative effect, but luckily they are usually recessive (otherwise they might not have survived).The effect of deleterious recessive alleles comes only to expression in homozygotes (carrying two copies of the recessive allele).

This is applicable to genetic defects, which are usually due to recessive alleles in small frequencies. If the frequency of the recessive allele is q, than in a non-inbred population, the probability of being an affected individual is q2. An inbred individual would have a probability of q2+pqF.

Let q be equal to 1 %. We have then

	Probability of being affected
Normal individual:	1 in 10,000
Inbred individual (F=0.125)	13.4 in 10,000

Hence a large increase!

2) Inbreeding depression
The effect of increased frequencies of individuals that are homozygous for negative recessive effects translates for quantitative traits, regulated by possibly many genes, into inbreeding depression. Increased homozygosity means most traits are depressed by between 2% and 7% per 10% increase in F.

Since we observe the phenomena only for alleles that are recessive, we should observe inbreeding depression only for traits that show dominance. Those are typically traits that relate to *fitness and reproduction*.

Inbreeding depression is a 'mirror image' of heterosis, the first is due to a shortage of heterozygotes, the second due to an excess of heterozygotes. Heterosis if more distant line or

$$\Delta F = \frac{1}{2N_e}$$

> ***Please note:*** *this prediction of inbreeding only holds for the case of no selection and random mating. In reality, there is selection going on in breeding programs and the inbreeding rate can be a multiple of the one predicted with the no-selection assumption.*

EXAMPLE

What is F after 20 years' breeding on the following structure?

AGE	2	3	4	5	6
RAMS	1	1			
EWES	20	20	20	20	20

L_m =2.5 L =3.25 years
L_f =4

N_m =1 x 3.25 = 3.25 males entering the flock per generation.
N_f =20 x 3.25 = 65 females entering the flock per generation.

$$N_e = \frac{4N_mN_f}{N_m + N_f} = \frac{845}{68.25} = 12.38$$

$F_{20years} = F_{20\backslash 3.25gens} = 1- [1-1/(2x12.38)]^{20\backslash 3.25}$

$F_{20years}$ = 0.224 which is probably unacceptable – but you need to look at response to selection together with this when looking for the best overall strategy. Also, a flock is unlikely to remain fully closed for 20 years.

Alternative strategy: Use only 4 new rams each year:

AGE:	2	3	4	5	6
RAMS	4				
EWES	20	20	20	20	20

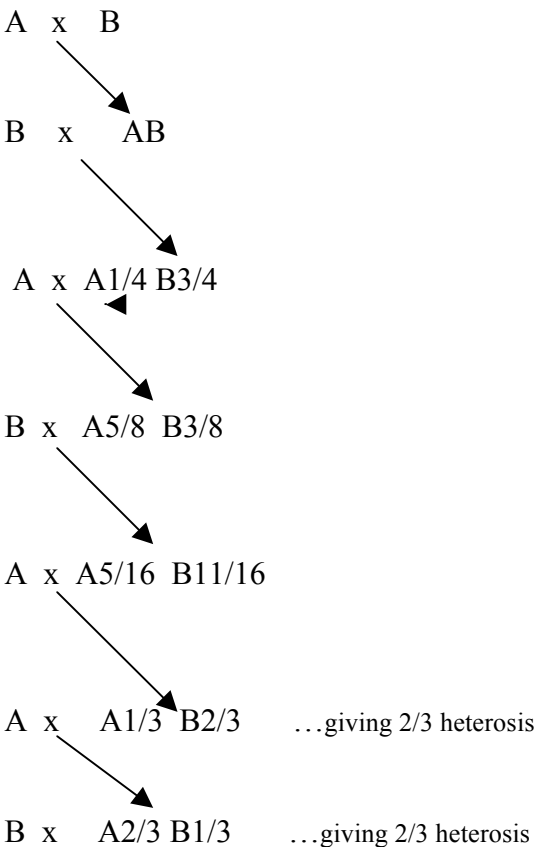
L_m =2 L =3 years
L_f =4

Calculation
L_m =2, L_f =4, L = 3
$N_{m.}$ = 4 x 3 = 12, N_f = 20 x 3 = 60, N_e = 40
F_{20yrs} = 0.0804 which is probably acceptable.

NOTE:

$$\text{As } F_t = 1 - \left[1 - \frac{1}{2N_e}\right]^t$$

F_t can only approach unity. For the previous example, F = 0.5 is reached in about 166 years.



Heterosis expressed in established rotations with n breeds contributing equally is

$$\frac{2^n - 2}{2^n - 1}$$

From this it can be shown that rotational crosses express more heterosis than synthetics which use the number of breeds

Calculating predicted merit in crossbreeding systems

EXAMPLE: Yearling weight in cattle, three parental breeds: 1, 2 and 3. Hypothetical values of the following parameters:

Direct additive effects A_{d1}, A_{d2} and A_{d3}. The expression of this is equal to the proportion of genes from each breed.

Maternal additive effects A_{m1}, A_{m2} ***and*** A_{m3}. These relate to dam genotype, and are probably caused by milk production and rearing ability. Note that these effects add to zero- they describe the relative maternal performance of each pure breed.

Direct dominance effect D_d. this is the effect of heterosis in crossbred individuals, when fully expressed as in an F_1 cross.

Maternal dominance effect D_m. this is the effect of heterosis due to crossbreeding in the dam, when fully expressed as in an F_1 dam.

3. If lower tiers buy from the average rams (and no ewes) from the tiers above, they will lag behind the tier above by 2 generations (about 7 years in sheep) of selection response (Bichard, 1971).

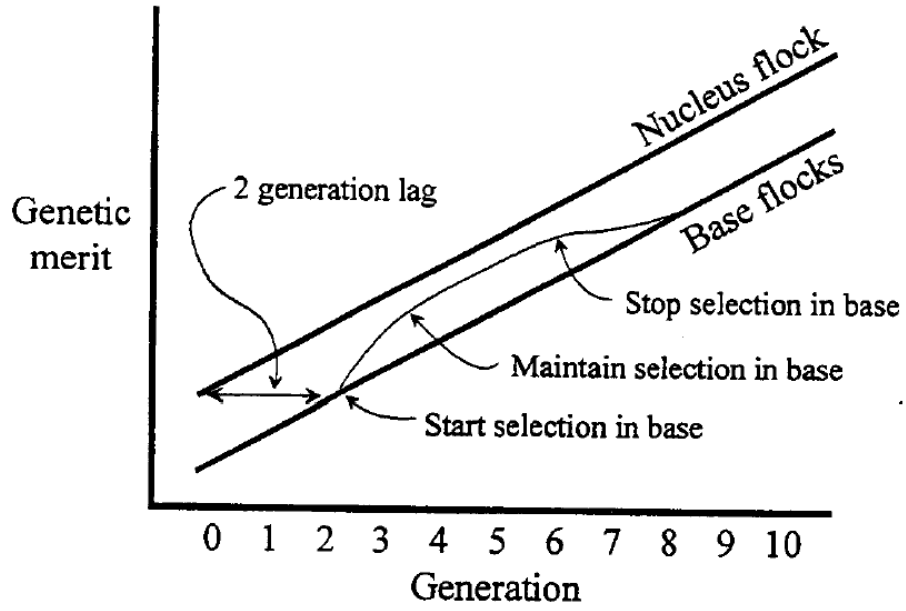


Figure 2. Selection response in a 2-tier closed nucleus scheme. The base lags about 2 generations behind the nucleus. Any selection effort in the base needs to be maintained just to keep a non-increasing advantage. Opening the nucleus will give more sustained returns from selection in the base.

Open nucleus breeding schemes

Stock in the base tier(s) can have higher EBV's than nucleus stock that would have otherwise been selected. This is most true for animals of low fecundity, such as ewes.

These high- merit base ewes can be migrated up to be bred in the nucleus, giving an open nucleus scheme. This pushes the nucleus to progress more quickly and this benefits the whole scheme as the base will move as fast as the nucleus after things have settled down. Overall response in open 2-tier schemes is 10 – 15 % faster than in closed schemes when optimal design is applied: about 10% of the population in the nucleus and about 50% of nucleus mated ewes born in the base (James, 1977).

Different measurement strategies in nucleus and base

A major use of nucleus schemes is to avoid or reduce measurement costs in lower tiers. Increased (or decreased) accuracy can be got by measuring more (or less) traits as selection criteria for index or using more (or less) information from, relatives. As $r_{\hat{A}A}$ increases, $\sigma_{\hat{A}}$ increases ($\sigma_{\hat{A}}=r_{\hat{A}A}\,\sigma_A$), and the distribution of EBV widen as shown in figure in the next page.

The classic approach to calculating economic weightings is economically rationale – it takes no account of genetic parameters. This makes sense in that the value of making a unit change in a given trait should not be influenced by how difficult it is to generate this change. These difficulties can be handled appropriately at the genetic evaluation phase. In this setting, breeding objectives should reflect the costs and returns involved in a production system, and should not consider costs and gains generated in a breeding programme.

A simple example

A very simple breeding objective is presented here. The reader is directed to Ponzoni (1988) for a comprehensive worked example. The key tactical objective of a selection programme is to choose animals of high breeding value to be used as parents. An animal's breeding value (A) is the value of its genes to its progeny. The breeding objective is simply a multi-trait breeding value, with each trait weighted by a relative economic weight, for example:

Breeding objective:	6 x fleece weight		+	-1 x Fibre diameter
Units = KSh	KSh /Kg.	Kg	+	KSh/ μ . μ

In order to combine different traits into such a single soccer they have to be converted to a common scale. This is generally dollars or some other monetary unit. The economic weights in this simple example are taken to have been calculated from market prices of KSh 6 per kilogram of wool and - KSh1 per micron for an average fleece. The units of expression are thus dollars [per unit (kg or u)] per head shown. Note that these weights involved no consideration of genetic parameters.

Units of expression

All economic weightings in a breeding objective should have the same basis for units of expression, such as 'KSh per head shorn' as used above. Choice of this basis can have an important influence on the consequences of using the breeding objective.

A simple basis for unit of expression such as dollars per head shown can be used for situations in which all traits are directly related to economic costs or returns, and thus excludes reproductive traits, whose effect is at least partly manifested through progeny.

A less simple basis is 'dollars per breeding ewe per year', which accommodates both production and reproduction traits. In all cases, each trait should use this same basis, such that an objective might be for example:

KSh 7.20 per kg per breeding ewe per year x clean fleece weight
+ KSh -120 per micron per breeding ewe per year x Fibre diameter
+ KSh 7.20 per lamb weaned per breeding ewe per year x Number of lambs weaned

This means that in a flock of 150 breeding ewes, a marginal increase of 1kg in clean fleece weight would increase profit by KSh7.20 x 150 = KSh1,080 per year. This accommodates wool shorn form all classes of stock, through the way in which the economic weight is calculated. The economic weight for an increase of one lamb weaned is more difficult to calculate, due to expression via progeny, but in this case it is the same as for clean fleece weight. Delays in returns due to expression in progeny can be accommodated by considering the pattern of flow of genes through the population, and discounting future returns to give current values (McClintock and Cunningham, 1974).