Predicting Response to Selection

Breeding plans

can be understood using a few simple

principles of

Response to Selection

Predicting Genetic Merit of Progeny

Guiding principle of genetic improvement:

Parents with high additive genetic values (BV) tend to have progeny with high additive genetic values (==> High phenotypes)

Single progeny: $g_o = \frac{1}{2}g_s + \frac{1}{2}g_d + g_m$

$$E(g_o) = \tfrac{1}{2}g_s + \tfrac{1}{2}g_d$$

Progeny generation: $E(\bar{g}_o) = \frac{1}{2}\bar{g}_s^* + \frac{1}{2}\bar{g}_d^*$

* indicates the variable refers to selected individuals

Objectives:

developing 'simple' deterministic models to predict responses to selection for:

- **Evaluation and comparison** of breeding plans
- **Optimization** of breeding plans
- **Factors that affect responses to** breeding plans

$$E(\bar{g}_o) = \tfrac{1}{2}(\bar{g}_s^* - \bar{g}_s + \bar{g}_s) + \tfrac{1}{2}(\bar{g}_d^* - \bar{g}_d + \bar{g}_d)$$

$$= \tfrac{1}{2}(\bar{g}_s + \bar{g}_s^* - \bar{g}_s) + \tfrac{1}{2}(\bar{g}_d + \bar{g}_d^* - \bar{g}_d)$$

$$= \tfrac{1}{2}(\bar{g}_s + S_s) + \tfrac{1}{2}(\bar{g}_d + S_d)$$

S = genetic superiority of the selected parents $\quad S_s = \bar{g}_s^* - \bar{g}_s$

$$E(\bar{g}_o) = \underbrace{\tfrac{1}{2}(\bar{g}_s + \bar{g}_d)}_{\text{mean parental generation}} + \underbrace{\tfrac{1}{2}(S_s + S_d)}_{\text{average genetic superiority of selected parents}}$$

Response to selection

difference between
mean genetic value of **progeny**
and **parental generation**

$$E(R) = \bar{g}_o - \bar{g}_p \qquad \text{with } \bar{g}_p = ½(\bar{g}_s + \bar{g}_d)$$

$$E(R) = ½(\bar{g}_s + \bar{g}_d) + ½\,(S_s + S_d) - ½(\bar{g}_s + \bar{g}_d)$$

$$= ½(S_s + S_d)$$

expected response to selection =
average genetic superiority of selected parents

Linear regression theory

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

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

$$b_{yx} = \frac{\sigma_{xy}}{\sigma_x^2} = r_{xy}\frac{\sigma_y}{\sigma_x} \qquad r_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_y^2\sigma_x^2}}$$

Predict g_i given I_i : $\hat{g}_i = \bar{g} + b_{gI}\,(I_i - \bar{I}\,)$

Predict the mean genetic value of a group of selected animals:

$$\bar{g}^* = \bar{g} + b_{gI}\,(\bar{I}^* - \bar{I}\,)$$

Predict the genetic superiority of selected parents:

$$\hat{S} = \bar{g}^* - \bar{g} = b_{gI}\,(\bar{I}^* - \bar{I}\,)$$

Predicting Response per Generation

$\hat{g}$ not available for future generations or for alternative designs

=> need to predict genetic superiority of selected parents by other means

based on linear relationship between a selection criterion, **I**, and **g**

Define intensity of selection: $i = (\bar{I}^* - \bar{I}\,)/\sigma_I \quad \rightarrow \quad (\bar{I}^* - \bar{I}\,) = i\sigma_I$

$$\hat{S} = b_{g.I}\; i\;\sigma_I \qquad \text{with} \quad b_{g.I} = r_{g.I}\frac{\sigma_g}{\sigma_I}$$

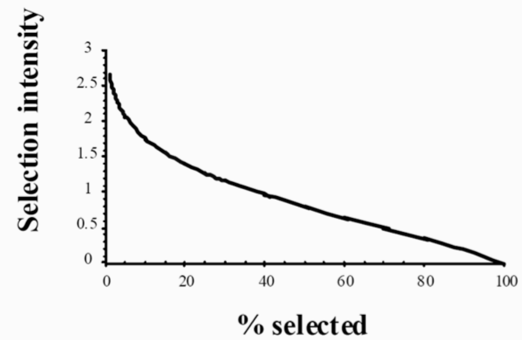
$$\hat{S} = r_{g.I}\frac{\sigma_g}{\sigma_I}\,(i\;\sigma_I) = i\;r_{gI}\;\sigma_g$$

$$\bar{g}_o = \tfrac{1}{2}(\bar{g}_s + S_s) + \tfrac{1}{2}(\bar{g}_d + S_d)$$

predict response per generation: $R = \tfrac{1}{2}(S_s + S_d)$

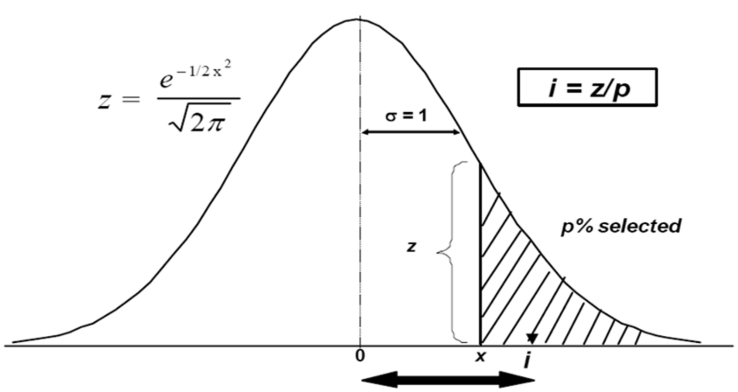
$$= \tfrac{1}{2}(i_s\, r_s\, \sigma_g + i_d\, r_d\, \sigma_g)$$

Effect of % Selected on Selection Intensity
Large population size

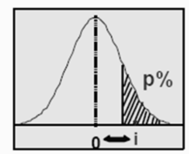


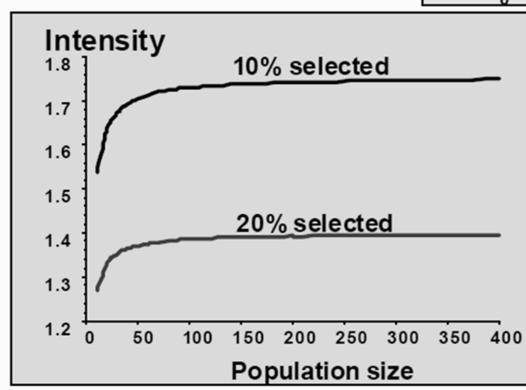
Selection Intensity with Truncation Selection

Derivation of selection intensity from the standard Normal distribution



Effect of Population Size
(Falconer and MacKay, 1996)





Adjustments to standard selection intensities:

- **for small population size: *i < z/p***
  - order statistics
  - approximation: $p^* = \dfrac{(s + {}^1/_2)}{n + \dfrac{s}{2n}}$

s = # selected n = population

and then $i^* = \dfrac{z^*}{p^*}$

Modeling Selection Across Multiple Age Groups

Age group	Relative number of candidates	Fraction selected from that group
1	w_{s1}	p_{s1}
2	w_{s2}	p_{s2}
3	w_{s3}	p_{s3}
	$\Sigma w_{si} = 1$	

Total proportion selected is $P_s = p_{s1}\, w_{s1} + p_{s2}\, w_{s2} + p_{s3}\, w_{s3}$

Adjustments to standard selection intensities:

- **For correlated EBV due to:**

1) genetic relationships between candidates of selection
2) use of the same information in calculating the EBV

e.g. n_{fs} full sib families with n_w individuals per family and selection on pedigree

$(\hat{g}_o = ½\hat{g}_s + ½\hat{g}_d)$ ➔ Correlation between EBV of full sibs = 1

➔ select n_c/n_w families out of n_{fs}
instead of: select n_c individuals out of $n_{fs}n_w$

} ➔ related to effect of pop. size

Genetic mean selected sires in group *i* : $\bar{g}^*_{si} = \bar{g}_{si} + S_{si}$

$S_{si} = i_{si}\, r_{si}\, \sigma_g$

$\bar{g}_{si}$ = genetic mean in age group i

r_{si} = accuracy of the selection criterion in group i

$$\bar{g}^*_s = \frac{1}{P_s}\{p_{s1}\, w_{s1}\, \bar{g}^*_{s1} + p_{s2}\, w_{s2}\, \bar{g}^*_{s2} + p_{s3}\, w_{s3}\, \bar{g}^*_{s3}\}$$

$$\bar{g}^*_d = \frac{1}{P_d}\Sigma p_{di}\, w_{di}\, (\bar{g}_{di} + S_{di})$$

Average genetic value of the progeny

$$E(\bar{g}_o) = ½\bar{g}^*_s + ½\bar{g}^*_d$$

Maximization in the selected group

What proportions should be selected from each group to maximize the average genetic value (given total fraction selected P_s)**?**

Assumption for each age group: $\mathrm{E(g_i) = I_i}$

- truncation selection across distributions of I
- find truncation point where selection across all distributions yields total proportion selected $\mathrm{P_s}$

Asymptotic Response per Unit Time

Generation interval:

L = average age of the parents when their progeny are born

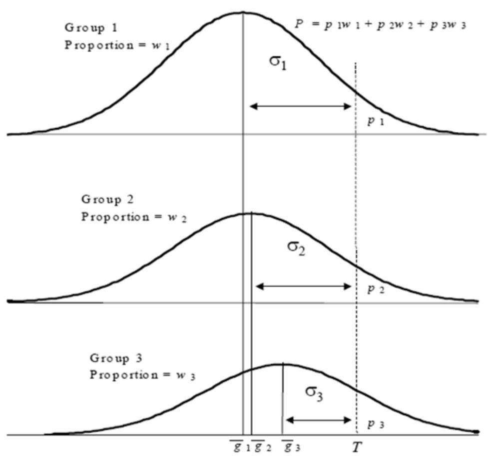
= average time between birth of the parents and birth of progeny

With equal selection in males and females:

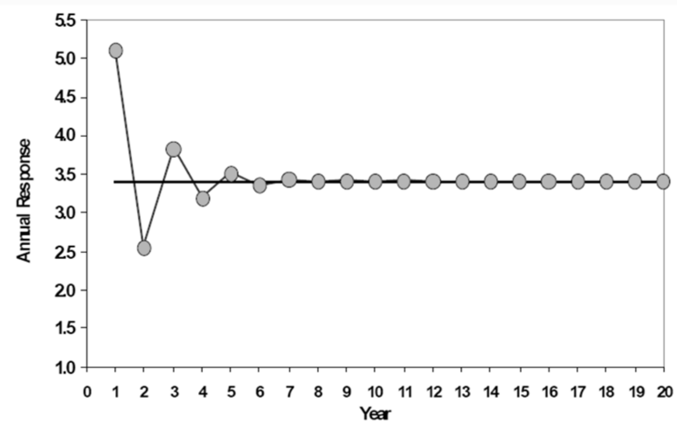
$$\text{Response per generation} = R = S = i r_{gI} \sigma_g$$

$$\text{Response per year} = R = \frac{i r_{g,I} \sigma_g}{L}$$

Schematic representation of truncation selection of a total proportion P across multiple overlapping distributions



Example of predicted annual versus asymptotic responses
Starting from an unselected population



Multiple Pathways of Selection
Unequal selection in males and females

$$R = \frac{S_s + S_d}{L_s + L_d}$$

= 'steady state' or 'asymptotic' response

When response constant at R/yr:

$\bar{g}_s = \bar{g}_o - L_s R$

$\bar{g}_d = \bar{g}_o - L_d R$

$\bar{g}_o = ½(\bar{g}_s + S_s) + ½(\bar{g}_d + S_d)$

$\bar{g}_o = ½(\bar{g}_o - L_s R + S_s) + ½(\bar{g}_o - L_d R + S_d)$

$= \bar{g}_o - ½R(L_s + L_d) + ½(S_s + S_d)$

Four basic pathways of genetic improvement

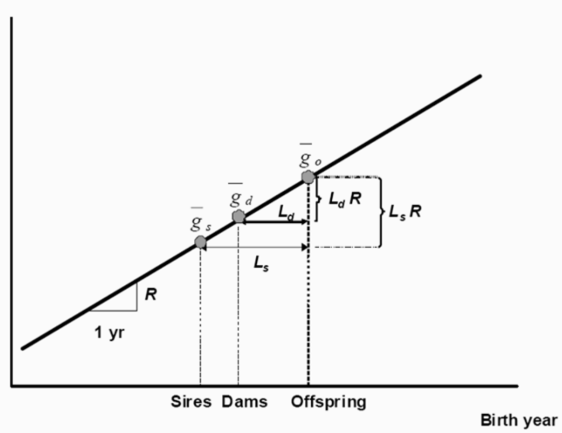
- Male parents of male progeny (sires of males, *sm)*
- Female parents of male progeny (dams of males, *dm)*
- Male parents of female progeny (sires of females, *sf)*
- Female parents of female prog. (dams of females, *df)*.

Asymptotic response/yr =

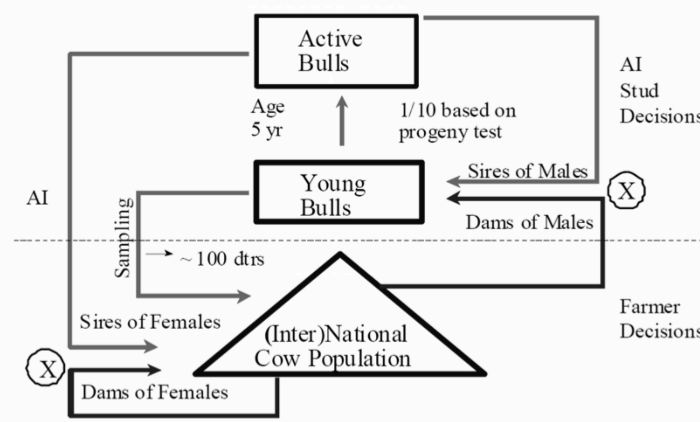
$$R = \frac{S_{sm} + S_{dm} + S_{sf} + S_{df}}{L_{sm} + L_{dm} + L_{sf} + L_{df}} = \frac{\sum_i S_i}{\sum_i L_i}$$

with $S_i = i_i r_i \sigma_g$

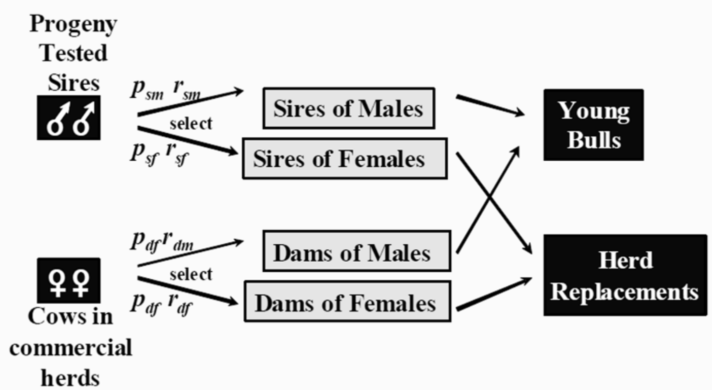
Asymptotic response to selection for breeding programs with overlapping generations



Dairy Cattle Progeny Testing Program



4 pathway program



Selection Across Age Groups

e.g. Selection of Bull Dams $\sigma_g = 550$ kg

Age Group	Age at Birth of Progeny	% of Bull Dams	% Selected	i	r	Genetic Superiority $ir\sigma_g$
Heifers	2 yr	50%	2.5%	2.34	.55	707.9
1st Lact.	3 yr	30%	1.5%	2.53	.68	946.2
2nd Lact.	4 yr	20%	1.5%	2.53	.72	1001.9

With selection across multiple age groups:

$$S_s = \frac{1}{P_s}\{p_{s1}\, w_{s1}\, S_{s1} + p_{s2}\, w_{s2}\, S_{s2} + p_{s3}\, w_{s3}\, S_{s3}\}$$

$$L_s = \frac{1}{P_s}\{p_{s1}\, w_{s1}\, L_{s1} + p_{s2}\, w_{s2}\, L_{s2} + p_{s3}\, w_{s3}\, L_{s3}\}$$

Selection Across Age Groups

- **Pooled Generation Interval**

  L_{dm} = 50% * 2
  + 30% * 3
  + 20% * 4

  = 2.7 yr

- **Pooled Genetic Superiority**

  S_{dm} = 50% * 707.9
  + 30% * 946.2
  + 20% * 1001.9

  = 838.2 kg

Accounting for the use of young bulls and pre-selection of young bulls

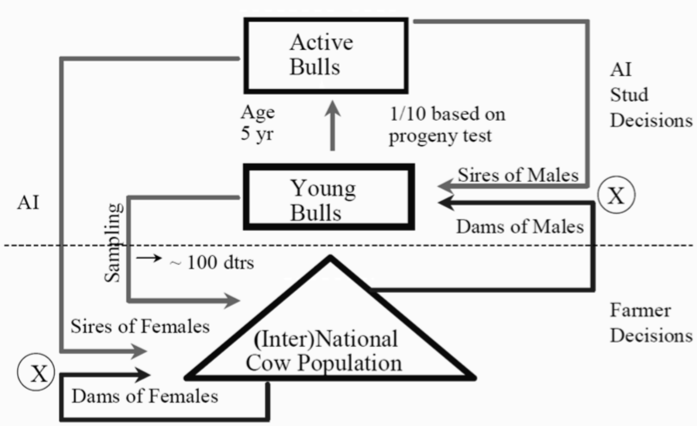
y = proportion of females produced from young bulls

Genetic superiority bull pathway =

weighted average of gen.superiorities of groups of bulls used

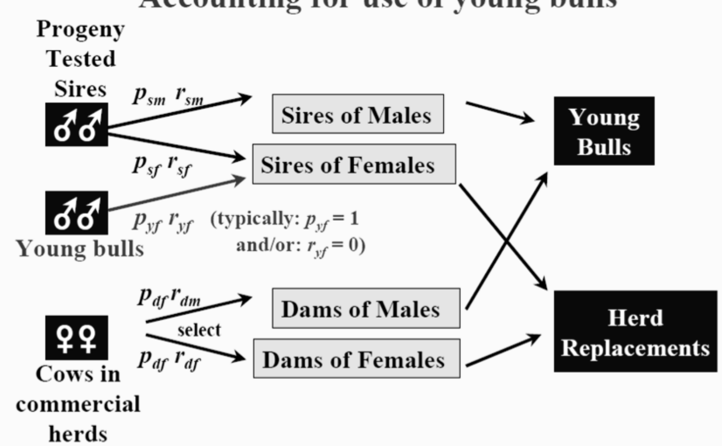
$$S_{sf} = y\, S_{yb,f} + (1\text{-}y)\,(S_{yb,f} + S_{pb,f})$$

Dairy Cattle Progeny Testing Program



Predicting Genetic Progress

Accounting for use of young bulls



Gen.superiority selected progeny-tested bulls is sum of genetic superiorities across stages = $(S_{yb,f} + S_{pb,f})$

Intensity and accuracy of selection and generation interval in a highly efficient hypothetical progeny-testing program for improving milk yield in dairy cattle with accounting for 20% use of young bulls to breed female replacements.

Pathway	Proportion Selected (p_i)	Intensity (i_i)	Accuracy (r_i)	Genetic Superiority ($S_i = i\, r_i \sigma_g$)		Generation Interval (yr) (L_i)	
Sires of males	2 %	2.42	0.90	$2.178\sigma_g$		6	
Sires of females - Young	100 %	0	0.50	0	$1.260\sigma_g$	2	6
Sires of females - Proven	10 %	1.75	0.90	1.575		7	
Dams of males	0.5 %	2.89	0.60	$1.734\sigma_g$		5	
Dams of females	90 %	0.19	0.60	$0.114\sigma_g$		6	
TOTAL				$\Sigma S = 5.268\sigma_g$		$\Sigma L = 23$	

$$R = \frac{5.268}{23}\sigma_g = 0.230\ \sigma_g \text{ per yr}$$

Generation interval bull pathway =

weighted average of gener.interval of groups of bulls used

$$L_{sf} = y\, L_{yb,f} + (1\text{-}y)\, L_{pb,f}$$

Predicting Correlated Response to Selection

- Select on trait 1
  - ↳ Response in trait 1
    - Direct response to selection
  - ↳ Response in trait 2
    - $r_{g_{1,2}} \neq 0$
    - Correlated response to selection
    - Change in trait due to selection on other trait

Indirect Selection

- Genetic improvement of a trait of economic importance through selection on EBV for a correlated trait, e.g.
  - Select on Somatic Cell Count to improve mastitis resistance.
  - Select on conformation traits to improve herd life.

Efficiency of Indirect vs Direct Selection

- Indirect selection: (correlated response in trait 2 to selection on trait 1)

$$\Delta G_{2.1} = r_{g_{1,2}} \frac{\sigma_{g_2}}{\sigma_{g_1}} (\Delta G_1)$$

$$\uparrow \quad \Delta G_1 = \frac{\mathrm{i} r_{\hat{A}_1} \sigma_{g_1}}{\Sigma L}$$

$$\text{Efficiency} = \frac{\Delta G_{2.1}}{\Delta G_2}$$

Indirect Selection (cont'd)

- Advocated over direct selection if:
  - Correlated trait is recorded and direct trait not.
  - Correlated trait is less expensive to measure.
  - Correlated trait is measured earlier in life ∴ L↓
  - Correlated trait has higher h^2.