

Basic Principles of Response to Selection

Jack Dekkers

Response to selection = (mean BV of progeny) – (mean BV parental generation)

Predicting Mean BV of Progeny Generation

Guiding principle of genetic improvement: parents with high BV tend to have progeny with high BV

Single progeny: $g_o = \frac{1}{2}g_s + \frac{1}{2}g_d + g_m$ $g_m =$ Mendelian sampling terms
 $E(g_o) = \frac{1}{2}g_s + \frac{1}{2}g_d$ because $E(g_m) = 0$

Progeny generation: $E(\bar{g}_o) = \frac{1}{2}\bar{g}_s^* + \frac{1}{2}\bar{g}_d^*$ Average BV of progeny
* refers to selected individuals

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

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

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

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

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


 mean parental
generation


 average genetic superiority
of selected parents

Response = $R = \Delta G =$ (mean BV progeny generation) – (mean BV parental generation)

$$E(R) = \bar{g}_o - \bar{g}_p \quad \text{with } \bar{g}_p = \frac{1}{2}(\bar{g}_s + \bar{g}_d)$$

$$E(R) = \frac{1}{2}(\bar{g}_s + \bar{g}_d) + \frac{1}{2}(S_s + S_d) - \frac{1}{2}(\bar{g}_s + \bar{g}_d)$$

$$= \frac{1}{2}(S_s + S_d)$$

→ expected response to selection = average genetic superiority of selected parents

For current generation, genetic superiority of selected parents can be estimated based on their EBV:

If EBV are unbiased: $E(g) = \hat{g}$

Then – for an individual: $E(g_o) = \frac{1}{2}\hat{g}_s + \frac{1}{2}\hat{g}_d$

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

$$= \frac{1}{2}(\bar{\hat{g}}_s + \hat{S}_s) + \frac{1}{2}(\bar{\hat{g}}_d + \hat{S}_d)$$

$\hat{S} =$ estimated genetic superiority of the selected parents $\hat{S} = \bar{\hat{g}}^* - \bar{\hat{g}}$

→ Expected response = $E(R) = \frac{1}{2}(\hat{S}_s + \hat{S}_d)$

Predicting Response per Generation into the Future

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

→ need to predict genetic superiority of selected parents by other means

- do this based on the linear relationship between the selection criterion, I , and g

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

Predict the mean BV of a group of selected animals: $\hat{\bar{g}}^* = \bar{g} + b_{gI} (\bar{I}^* - \bar{I})$

Predict the genetic superiority of selected parents: $\hat{S} = \hat{\bar{g}}^* - \bar{g} = b_{gI} (\bar{I}^* - \bar{I})$

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

How many SD (for the selection criterion) is the selected group better than all candidates

Using $(\bar{I}^* - \bar{I}) = i \sigma_I$ $\hat{S} = b_{gI} i \sigma_I$

From regression: $b_{g,I} = r_{g,I} \frac{\sigma_g}{\sigma_I}$ $\hat{S} = r_{g,I} \frac{\sigma_g}{\sigma_I} (i \sigma_I) = i r_{gI} \sigma_g$

r_{gI} = accuracy of selection

$$\hat{S} = i r_{gI} \sigma_g$$

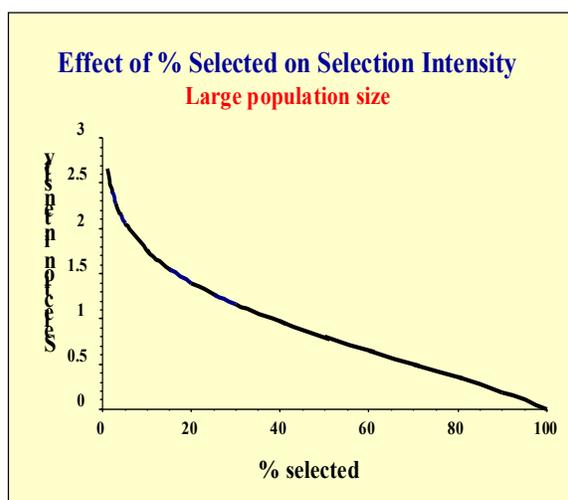
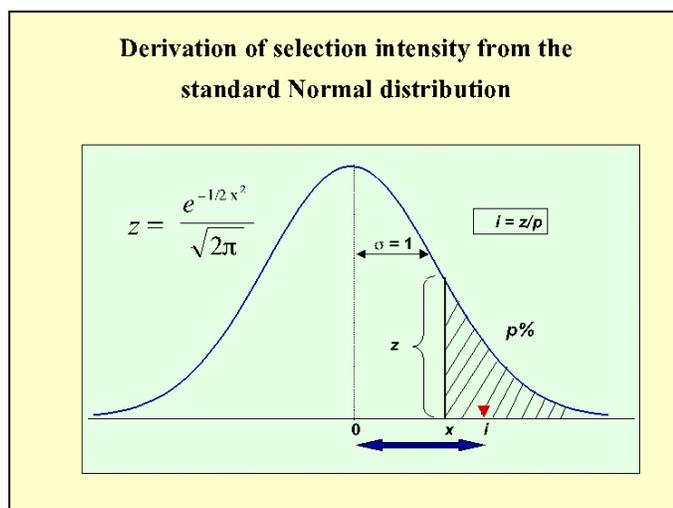
is general equation to predict genetic superiority of selected parents
applies whenever selection criterion (I) is linearly related to the BV (g)

Use $S = i r_{gI} \sigma_g$ to model the genetic mean of the next generation recursively:

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

or predict response per generation: $R = \frac{1}{2}(S_s + S_d) = \frac{1}{2}(i_s r_s \sigma_g + i_d r_d \sigma_g)$

Selection Intensity with Truncation Selection

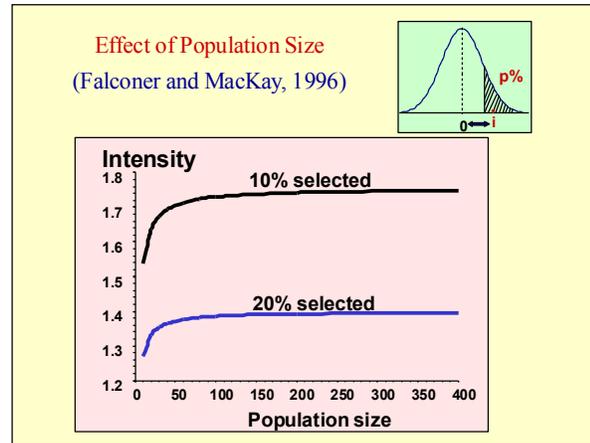


Adjustments to standard (large unrelated population) selection intensities:

- **small populations:** $i < z/p$
 - order statistics
 - approximation:
$$p = \frac{(s + 1/2)}{n + \frac{s}{2n}}$$

$$s = \# \text{ selected} \quad n = \text{population size}$$

$$i = \frac{z^*}{p^*}$$
 use p^* rather than p to derive intensity



- **Correct intensities for correlated EBV** among selection candidates 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

$$\left(\rightarrow \hat{g}_o = \frac{1}{2} \hat{g}_s + \frac{1}{2} \hat{g}_d \right) \quad \text{Correlation between EBV of full sibs} = 1$$

\rightarrow select n_c/n_w families out of n_{fs} } \rightarrow related to effect of pop. size
 instead of: select n_c individuals out of $n_{fs}n_w$

Approximations to account for effect of correlated EBV on intensity:

Rawlings (1976)
$$i^* = \sqrt{1 - t_{av}} i$$

t_{av} = average correlation between the selection criterion across all possible pairs of selection candidates

Unrelated full sib families
$$t_{av} = t_{fs} \frac{n_w - 1}{n_w n_{fs} - 1}$$

t_{fs} = correlation between selection criterion of full sibs

Full sib families nested within half sib families (Meuwissen 1991):
 n_{hs} sires mated with n_{fs} dams each producing n_w offspring

- Correlation between half sibs = t_{hs}
- Correlation between full sibs = t_{fs}
- Correlation between non-relatives = 0

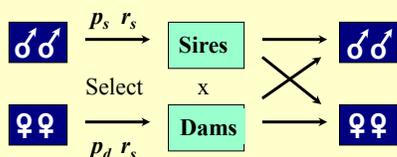
$$t_{av} = \frac{t_{fs}(n_w - 1) + t_{hs}n_w(n_{fs} - 1)}{n_w n_{fs} n_{hs} - 1}$$

Correlations between EBV of relatives based on selection index theory :

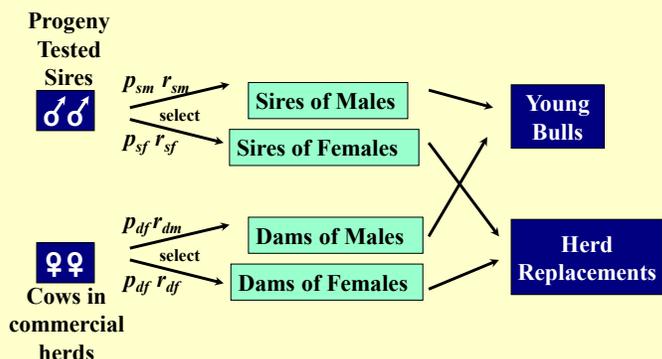
$$t_{i,j} = \text{corr}(I_i, I_j) = \text{corr}(\mathbf{b}'\mathbf{x}_i, \mathbf{b}'\mathbf{x}_j) = \frac{\mathbf{b}'\text{cov}(\mathbf{x}_i, \mathbf{x}_j)\mathbf{b}}{\mathbf{b}'\mathbf{Pb}} = \frac{\mathbf{b}'\mathbf{Rb}}{\mathbf{b}'\mathbf{Pb}}$$

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

2 Pathway Program



4 pathway program



→→ [Genetic_gain.xls](#)

Four pathways of genetic improvement

(Rendel and Robertson, 1950):

- sires of males sm
- dams of males dm
- sires of females sf
- dams of females df

$$\text{Asymptotic response/yr} = R_{yr} = \frac{S_{sm} + S_{dm} + S_{sf} + S_{df}}{L_{sm} + L_{dm} + L_{sf} + L_{df}}$$

$$R_{yr} = \frac{\sum_i S_i}{\sum_i L_i} \quad S_i = i r_i \sigma_g$$

Dairy Cattle Progeny Testing Program

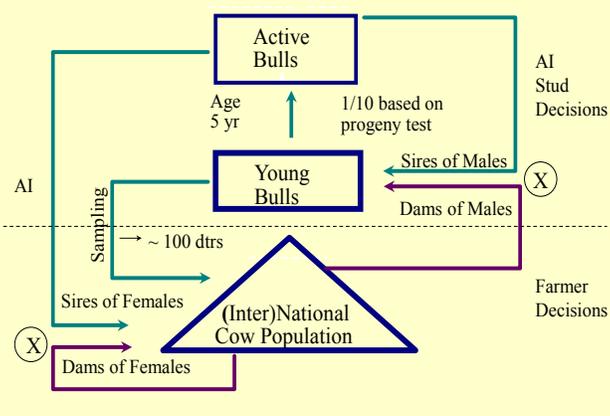


Table 1 Intensity and accuracy of selection and generation interval in a highly efficient hypothetical progeny-testing program for improving milk yield in dairy cattle.

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	10 %	1.75	0.90	$1.575 \sigma_g$	7
Dams of males	0.5 %	2.89	0.60	$1.743 \sigma_g$	5
Dams of females	90 %	0.19	0.60	$0.114 \sigma_g$	6
TOTAL				$\Sigma S = 5.601 \sigma_g$	$\Sigma L = 24$

$$R_{yr} = \frac{(2.42 \times 0.9 + 1.75 \times 0.9 + 2.89 \times 0.6 + 0.19 \times 0.6)}{6 + 7 + 5 + 6} \sigma_g = 0.233 \sigma_g \text{ per yr}$$

Selection Across Multiple Age Groups

E.g. three age groups:

Age group	Relative number of candidates	Generation interval	Mean within group	Accuracy of selection within group	Fraction selected within group	Genetic superiority within group	Genetic mean selected individuals
1	w_1	L_1	\bar{g}_1	r_1	p_1	$S_1 = i_1 r_1 \sigma_g$	$\bar{g}_1^* = \bar{g}_1 + S_1$
2	w_2	L_2	\bar{g}_2	r_2	p_2	$S_2 = i_2 r_2 \sigma_g$	$\bar{g}_2^* = \bar{g}_2 + S_2$
3	w_3	L_3	\bar{g}_3	r_3	p_3	$S_3 = i_3 r_3 \sigma_g$	$\bar{g}_3^* = \bar{g}_3 + S_3$
$\sum w_i = 1$							

Total proportion selected is:

$$P = p_1 w_1 + p_2 w_2 + p_3 w_3$$

Genetic mean of selected individuals across groups is:

$$\bar{g}^* = \frac{1}{P} \{ p_1 w_1 \bar{g}_1^* + p_2 w_2 \bar{g}_2^* + p_3 w_3 \bar{g}_3^* \}$$

$$\bar{g}_i^* = \bar{g}_i + S_i = i_i r_i \sigma_g$$

Applied to separate selection of males and females:

Genetic mean of selected sires:

$$\bar{g}_s^* = \frac{1}{P_s} \sum p_{si} w_{si} (\bar{g}_{si} + S_{si})$$

Genetic mean of selected dams:

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

Genetic mean of progeny:

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

For asymptotic gain/yr: Compute average within-group genetic superiority and generation interval across age classes:

For sires: $S_s = \frac{1}{P_s} \sum p_{si} w_{si} S_{si}$

$$L_s = \frac{1}{P_s} \sum p_{si} w_{si} L_{si}$$

For dams $S_d = \frac{1}{P_d} \sum p_{di} w_{di} S_{di}$

$$L_d = \frac{1}{P_d} \sum p_{di} w_{di} L_{di}$$

$$R_{yr} = \frac{S_s + S_d}{L_s + L_d}$$

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 $i r \sigma_g$
Heifers	2 yr	50%	2.5%	2.34	.55	707.9
1 st Lact.	3 yr	30%	1.5%	2.53	.68	946.2
2 nd Lact.	4 yr	20%	1.5%	2.53	.72	1001.9

Selection Across Age Groups

• Pooled Generation Interval

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

$$= 2.7 \text{ yr}$$

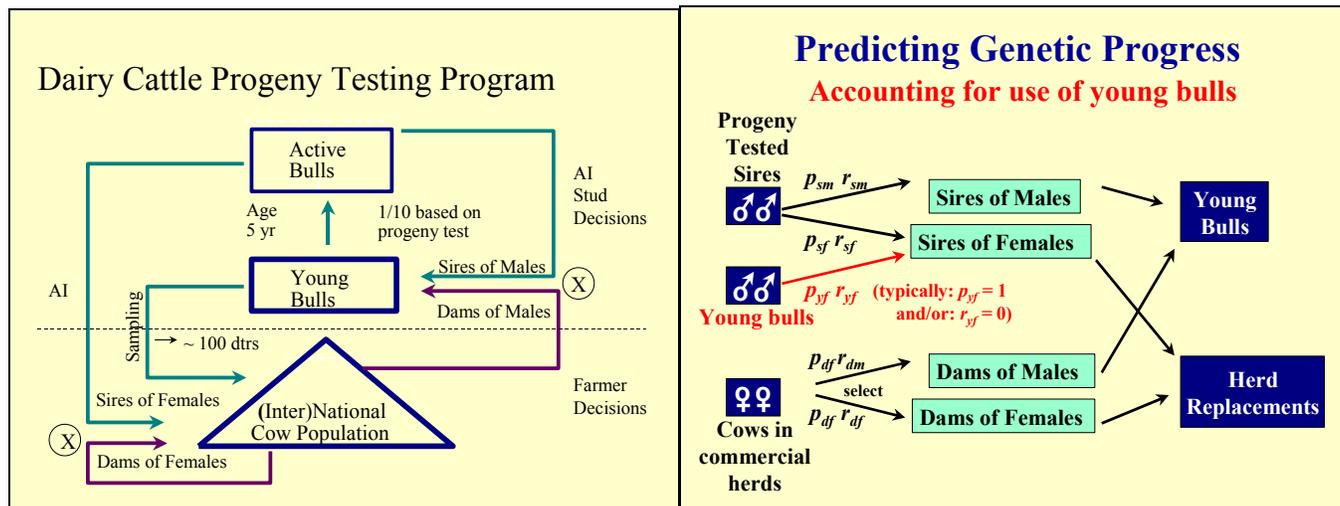
• Pooled Genetic Superiority

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

$$= 838.2 \text{ kg}$$

Accounting for the use of young bulls

y = proportion of females produced from young bulls



Genetic superiority sires of females

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

Without MAS/GS

$$S_{yb,f} = 0 \text{ because } p_{yb,f} = 1 \Rightarrow i_{yb,f} = 0 \text{ or } r_{yf} = 0$$

Generation interval of sires of females

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

Table 2. 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 σ_g		6	
Sires of females - Proven	100 %	0	0.50	0	1.260 σ_g	2	6
	10 %	1.75	0.90	1.575		7	
Dams of males	0.5 %	2.89	0.60	1.734 σ_g		5	
Dams of females	90 %	0.19	0.60	0.114 σ_g		6	
TOTAL				$\Sigma S = 5.268 \sigma_g$		$\Sigma L = 23$	

Now response per year becomes: $R = \frac{5.268}{23} \sigma_g = 0.230 \sigma_g$ per yr

→→ Genetic_gain.xls

Optimizing selection across (age) groups

Optimize proportions to select from each group to maximize the average genetic value of the selected group (for total fraction selected P).

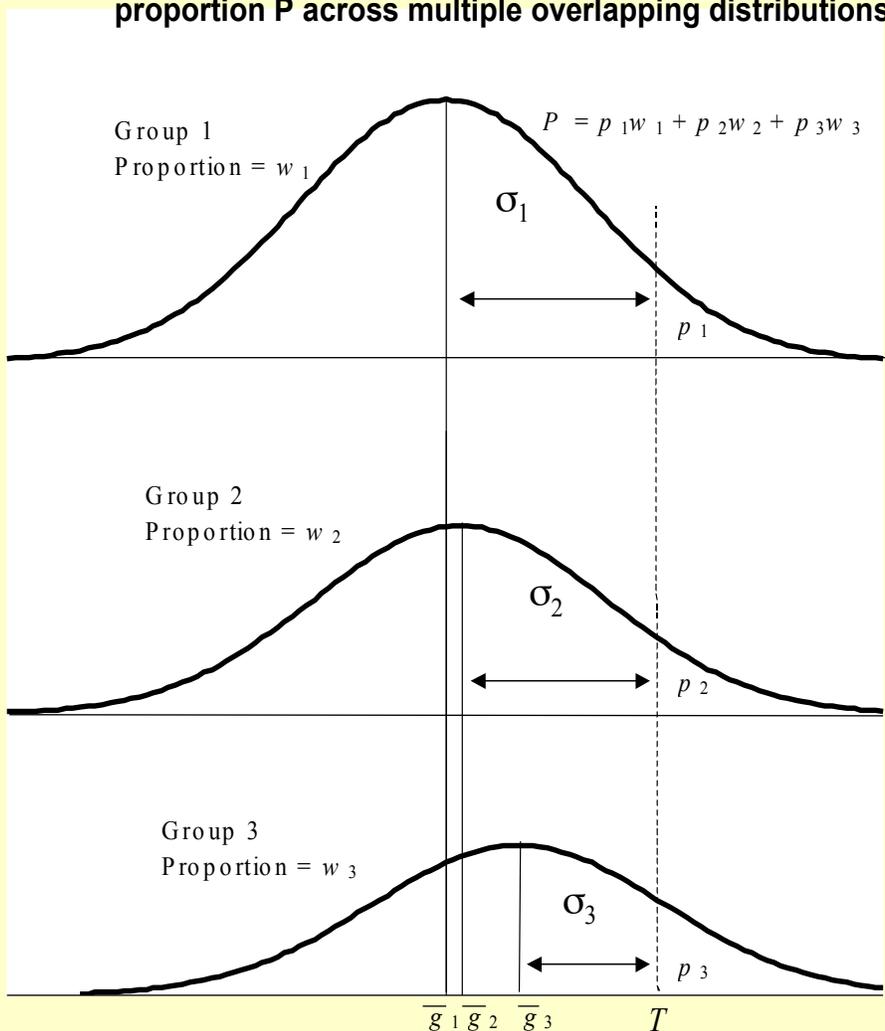
Assume that selection criterion is unbiased for each age group: $E(g_i) = I_i$

→ truncation selection across distributions of I

→ find the truncation point where selection across all distributions yields a total proportion selected of P

- Various algorithms can be used
- Bisection
 - gradient algorithms
 - Genetic algorithms

Figure 3.2 Schematic representation of truncation selection of a total proportion P across multiple overlapping distributions



→→ [Multrunc.xls](#)
→→ [truncselDE.xls](#)

Balancing Accuracy vs Generation Interval

Selection on EBV Across Age Groups

