

# 6A

## Least-Squares Analysis of Short-Term Selection Experiments

*Draft Version 26 May1999*

This chapter introduces some basic statistical machinery for evaluating selection experiments and then assesses the success (or failure) of the breeders' equation in predicting short-term response. Our focus is on least-squares (LS) analysis. Under the LS framework, response is measured from differences in phenotypic means and heritability is estimated from the (LS) regression of response on selection differential. Chapter 7 examines the more sophisticated mixed-model (MM) analysis, which requires knowledge of both the values of all individuals in the selection experiment (as opposed to just their generational means) and the complete pedigree of these individuals. For many readers, a LS analysis will be sufficient for their needs, so we refer the bulk of Chapter 7 only to the advanced (and dedicated) reader.

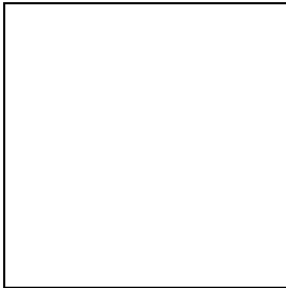
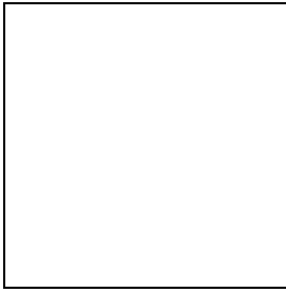
We start this chapter by considering the expected variance in response, then discuss simple estimating heritability from selection response, using the results for expected variances to obtain standard errors for these realized heritability estimates. We then turn to considerations of just how well the breeders' equation predicts short-term selection response. The statistical machinery developed in the first part of the chapter allows us to assess the quantitative fit of the breeders' equation, which often fails. This is not surprisingly given the large number of assumptions underlying the breeders' equation (e.g., Table 4.1). More interesting, even the gross qualitative predictions of the breeders' equations can fail. For example, a population selected to increase in a character may actually show a decrease. Likewise, replicate populations selected by equal amounts in opposite directions may show significantly greater response to selection in one direction. We discuss possible causes for these gross departures from the breeders' equation. While most selection experiments involve laboratory/domesicated populations in controlled environments, a few controlled experiments have looked at selection in natural populations and we end our discussion on the goodness of fit of the breeders' equation by reviewing these. We conclude by returning to a few statistical issues on experimental design, discussing the use of control populations and

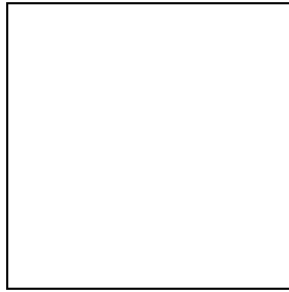
issues on the optimal design of (LS) selection experiments.

The literature on selection experiments is truly massive, and our goal here is to introduce the important concepts, rather than to exhaustively review all experiments. Recent reviews of selection experiments include Wilson (1977), Wright (1977), Roberston (1980), Mather (1983), Hill (1984), Sheridan (1988), Hill and Mackay (1989), Eisen (1989), Falconer (1992), and Hill and Caballero (1992). We again remind the reader that our focus here is short-term response — that expected over the first few generations of selection in the absence of significant change in allele frequencies. Chapters 9-12 examine long-term response.

### VARIANCE IN SHORT-TERM RESPONSE

Even if the assumptions underlying the breeders' equation hold exactly,  $h^2S$  is the predicted *mean* value, and the response to selection in replicate lines can vary considerably around this expected value. Figure 6.1 displays the results from three different experiments wherein replicate lines were subjected to identical selection. Since artificial selection usually involves choosing a small number of parents to form the next generation, a major source of this (between-line) variation is genetic drift. Drift also changes the within-line variance, but our assumption here is that the time scale of the experiment is such that we can ignore these changes. Chapter 12 examines the consequences of drift for long-term response in some detail.





**Figure 6.1.** Examples of the variance in response to selection among replicate lines. **Top:** Postweaning weight gain in male and female mice (Hanrahan et al. 1973). Each replicate consists of a single family, propagated by selecting the largest pair within each family. **Middle:** Abdominal bristle number in *Drosophila melanogaster* (Frankham et al. 1968a). Here, 50 pairs of parents were scored and largest 10 of each sex were used to form the next generation. **Bottom:** Pupal weight at 21 days in *Tribolium castaneum* (Gall 1971). Here, 50 males and 75 females were scored, with the largest 5 males and 20 females used to form the next generation.

Since our focus is short-term response, we assume that selection-induced changes in allele frequencies are sufficiently small to be ignored over the course of the experiment. In particular, we assume that selection has only a minor role in changing the genetic variance over the course of the experiment. This is the standard infinitesimal model assumption used throughout Chapters 4 and 5 (these assumptions are relaxed in Chapters 9-12). One complication that we ignore for now (addressing it later in the chapter) is that, under the infinitesimal model, selection decreases the additive genetic variance in a predictable manner due to gametic-phase disequilibrium (Chapter 5).

#### Expected Variance in Response Generated by Drift

The expected amount of variance between the means of replicate lines generated by drift and selection has been examined by Prout (1962a) and in detail by Hill (1971, 1972d,e, 1974c, 1977b, 1980, 1986). We start by considering the between-line variance generated by drift alone. As discussed in Chapter 3, there are two sources of between-line variance: variance in the phenotypic mean of the  $N$  individuals chosen (since under a pure drift model these individuals are chosen independent of their phenotype) and variance in breeding values given this phenotypic mean. For example, the population mean may be zero, but the mean phenotypic value in a randomly-drawn sample might be 2.5, while the mean breeding value for individuals in this sample might be 3.75.

With selection, the  $N$  individuals are not chosen at random with respect to their phenotypes, decreasing the first source of variance and hence decreasing the sampling variance relative to pure drift. Selection also alters the within- and

between-line variances relative to drift — generating negative gametic-phase disequilibrium which reduces the additive variance within a line (Chapters 5, 10) and further reducing the within-line additive variance by decreasing the effective population size below that expected from drift alone (Chapter 12). Countering these reductions in the within-line variance, selection can increase the between-line variance relative to drift by changing allele frequencies more rapidly than expected from drift alone. It is these opposing (and potentially offsetting) changes in variance that lead to suggestions that many of these effects cancel out, leaving the pure drift variance as an adequate approximation (Robertson 1977b; Hill 1977b, 1980, 1986; Nicholas 1980). Simulations (Robertson 1977b) and experimental results (Falconer 1973, López-Fanjul 1982) suggest this is not an unreasonable approximation. Hence, we use the between-line variance in means under pure drift (developed in Chapter 3) to approximate the variance in response about its expected value.

To examine the various sources of variation contributing to the variance in response, write the population mean in generation  $t$  as

$$\bar{z}_t = \mu + g_t + d_t + e_t \quad (6.1)$$

where  $g_t$  is the mean breeding value and  $d_t$  the mean environmental deviation in generation  $t$ . [In the notation of Chapter 3,  $\mu + g_t = \mu_g(t)$ ; since  $\mu$  is constant, note that  $\sigma_{\mu_g}^2(t) = \sigma_g^2(t)$ .] Under this model, a series of lines initiated simultaneously have expected value

$$E(\bar{z}_t) = \mu + E(g_t) + d_t \quad (6.2)$$

The variance of  $\bar{z}_t$  around its expected value  $E(g_t)$  in this particular generation is  $\sigma_g^2(t) + \sigma_e^2(t)$ . However, we usually don't know the particular value of  $d_t$ , which is assumed to be uncorrelated between generations ( $\sigma(d_t, d_{t'}) = 0$ ), with mean 0 and variance  $\sigma_d^2$ . Thus, the expected value of  $\bar{z}_t$  measured at a random time (e.g., the value of  $d$  is chosen at random) is  $\mu + E(g_t)$  and the variance of  $\bar{z}_t$  about this mean value is inflated by  $\sigma_d^2$  to give

$$\sigma_{\bar{z}}^2(t) = \sigma_g^2(t) + \sigma_e^2(t) + \sigma_d^2 \quad (6.3)$$

**Control populations** can be used to remove  $\sigma_d^2$ . Briefly, if one also has an unselected (control) population raised in the same environment, then the difference between the selected and control means ( $\bar{z}_{s,t} - \bar{z}_{c,t}$ ) removes the common environmental effect  $d_t$ . Complications with using controls can arise if there are genotype  $\times$  environment interactions and/or significant drift in the control population (such that  $\bar{z}_{c,t}$  deviates significantly from its expected value of zero). We examine these issues in more detail at the end of the chapter.

If  $M_0$  individuals are initially sampled to form each line, then from Equation 3.9b,

$$\sigma_g^2(t) = \left( \frac{1}{M_0} + 2f_t \right) h^2 \sigma_z^2 \quad (6.4)$$

where  $f_t$  is the amount of inbreeding at generation  $t$ . The  $M_0$  term accounts for variation in mean breeding value between lines in the founding generation while the  $f_t$  term accounts for variation generated by subsequent drift.

The variance of  $e_t$  is more involved (Equation 3.9b), with its exact form depending on the distribution of family sizes. Hill (1971, 1980) shows that it is bounded by

$$\frac{\sigma_z^2 - \sigma_A^2/2}{M_t} \leq \sigma_e^2(t) \leq \frac{\sigma_z^2}{M_t} \quad (6.5a)$$

To be conservative, we will use the upper bound

$$\sigma_e^2(t) = \frac{\sigma_z^2}{M_t} \quad (6.5b)$$

in the rest of our discussions. Equation 6.2 thus becomes

$$\sigma_{\bar{z}}^2(t) = \left( \frac{1}{M_0} + 2f_t \right) h^2 \sigma_z^2 + \sigma_d^2 + \sigma_z^2/M_t \quad (6.6)$$

Since drift variance accumulates each generation (via  $f_t$  increasing each generation) while the other terms do not, drift is expected to dominate, usually after a few generations. If population size remains constant,

$$2f_t = 2 \left[ 1 - \left( 1 - \frac{1}{2N_e} \right)^t \right] \simeq t/N_e \quad \text{for } t/N_e \ll 1 \quad (6.7)$$

If different numbers of males ( $N_m$ ) and females ( $N_f$ ) are sampled and/or  $N$  varies over time,

$$2f_t \simeq \sum_{k=0}^{t-1} \left[ \frac{1}{4N_m(k)} + \frac{1}{4N_f(k)} \right] \quad t > 0 \quad (6.8)$$

Equation 6.6 describes the divergence *between* lines due to drift. Drift also introduces a positive correlation between the means at different generations *within* a line. If the errors in estimating the mean breeding values in generations  $t$  and  $t'$  ( $g_t$  and  $g_{t'}$ ) from the phenotypic means ( $\bar{z}_t$  and  $\bar{z}_{t'}$ ) are uncorrelated and if between-generation environmental effects are uncorrelated, then from Equation 3.10b,

$$\sigma(g_t, g_{t'}) = \sigma(\bar{z}_t, \bar{z}_{t'}) = \left( \frac{1}{M_0} + 2f_t \right) h^2 \sigma_z^2 \quad \text{for } t < t' \quad (6.9)$$

The assumption that  $\sigma_A^2$  and  $\sigma_z^2$  remain constant is a major one and can be violated in at least three ways. First, changes in the underlying allele frequencies can change the variance. If major alleles are segregating, large changes in the variance can occur within a few generations (especially when selection is acting, see Figure 6.5). Major alleles initially at low frequencies can result in a substantial increase

in the between-line variance over that predicted by Equation 3.9b, as these alleles are lost in some lines and increase in frequency in others (e.g., Table 5.2). The net result being that different lines from the same base population can start with very different values of additive variance, increasing the variance in response (James 1970). Second, directional selection generates negative gametic-phase disequilibrium, reducing the additive genetic variance within a line (Chapter 5). Third (as mentioned above), inbreeding due to finite population size reduces additive genetic variance within a line by fixing alleles. In the absence of mutational input the expected additive genetic variance within a line in generation  $t$  is (Equation 3.2)

$$\sigma_A^2(t) = \left(1 - \frac{1}{2N_e}\right)^t \sigma_A^2(0) \simeq \left(1 - \frac{t}{2N_e}\right) \sigma_A^2(0)$$

where  $\sigma_A^2(0)$  is the variance in the base population. If dominance and/or epistasis is present, the within-line variance can actually increase under drift (Chapter 3), further inflating the between-line variance. What little theory has been developed for expected divergence when dominance and/or epistasis is present was discussed in Chapter 3, and we restrict considerations to those situations where additive variance accounts for all but a negligible fraction of the total genetic variation. Provided  $t/2N_e \ll 1$ , the error introduced in Equations 6.6 and 6.9 by ignoring the reduction in the within-line variance due to inbreeding is small. A final complication that we have ignored is that drift generates between-line variance in  $\sigma_A^2$  itself (Chapter 3), further inflating the between-line variance in means.

Despite all these potential complications, those few experimental tests of the pure-drift approximation have found it to be fairly reasonably (Falconer 1973, López-Fanjul 1982). Clearly, this is an area for further experimental research. The mixed-model (REML/BLUP) methods introduced in Chapter 7 account for many of the areas of concern raised above. However, they require a complete pedigree of all individuals involved in the selection experiment and also assume that the infinitesimal model hold.

## REALIZED HERITABILITIES

The breeders' equation immediately suggests that heritability can be estimated as the ratio of observed response to observed selection differential,

$$\hat{h}_r^2 = \frac{R}{S} \quad (6.10)$$

Estimates of heritability based on the response to selection are referred to as **realized heritabilities** (Falconer 1954), and we denote these estimates by  $\hat{h}_r^2$ . While one can use this approach to estimate  $h^2$ , any complication in predicting response

using the breeders' equation (Table 4.1) will usually make  $\widehat{h}_r^2$  a biased estimator of  $h^2$ . Turning this point around, however, suggests that one test for the success of the breeders' equation is to compare how close realized heritabilities are to heritabilities estimated from resemblance between relatives in the unselected base population. If the breeders' equation generally provides an accurate model of selection response, we expect these two different estimates to be similar (i.e., within sampling error).

### Estimators for Several Generations of Selection

While the estimator given by Equation 6.10 is unambiguous for a single generation of selection, two different estimates have been proposed when several generations of selection are considered. Both are based on the **cumulative selection response**  $R_C(t)$  and **cumulative selection differential**  $S_C(t)$ , which are defined by

$$S_C(t) = \sum_{i=1}^t S_i \quad (6.11a)$$

and

$$R_C(t) = \sum_{i=1}^t R_i \quad (6.11b)$$

where  $S_i$  and  $R_i$  are the (directional) selection differential and single-generation response for generation  $i$  ( $\bar{z}_i^* - \bar{z}_i$  and  $\bar{z}_{i+1} - \bar{z}_i$ , respectively). Perhaps the most common multigeneration estimator of realized heritability is to use the slope of the unweighted (ordinary) least-squares (OLS) regression of cumulative response on cumulative selection differential,

$$R_C(t) = b_C S_C(t) + e_t \quad \text{for } t = 1, \dots, T \quad (6.12)$$

with  $\widehat{h}_r^2 = b_C$  for mass selection (Falconer 1954). Modifications of the approach can be used for family selection (Chapter 8) and other designs, such as divergent selection (discussed at the end of this chapter). Since the expected response is zero if there is no selection, the regression line is constrained to pass through the origin and hence lacks an intercept term. Recalling LW Equation 8.33a, the OLS estimator for the slope is  $(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$ . Since the design matrix  $\mathbf{X}$  for the regression given by Equation 6.12 is just the vector of cumulative selection differentials  $\mathbf{S}$  and  $\mathbf{y}$  is the vector of cumulative responses  $\mathbf{R}$ , it follows that the OLS estimate of the slope is given by

$$\widehat{b}_C(\text{OLS}) = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \mathbf{R} = \frac{\sum_{i=1}^T S_C(i) \cdot R_C(i)}{\sum_{i=1}^T S_C^2(i)} \quad (6.13)$$

We will refer to this as the **OLS regression estimator** of the realized heritability.

An alternate estimator for multigenerational data is to simply consider the ratio of total selection response to total selection differential,

$$\hat{b}_T = \frac{R_C(T)}{S_C(T)} \quad (6.14)$$

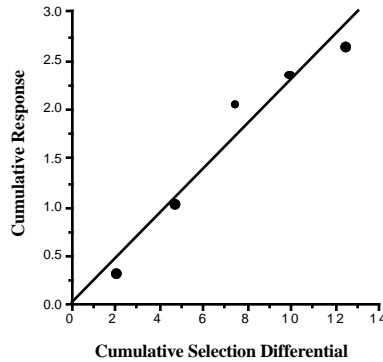
with  $\hat{h}_r^2 = \hat{b}_T$  for mass selection. We refer to this as the **ratio estimator** of the realized heritability.

**Example 1.** Consider the following data from Mackay (1985), who performed a divergent selection experiment on abdominal bristle number in replicate lines of *Drosophila melanogaster*. Fifty males and fifty females were measured in each line, with ten of each sex selected to form the next generation. Her data for the High (up-selected) line from replicate pair 2 for the first five generations of selection are

t	$\bar{z}$	$\bar{z}^*$	$S(t)$	$R(t)$	$S_C(t)$	$R_C(t)$
1	18.02	20.10	20.10–18.02 = 2.08	18.34–18.02 = 0.32	2.08	0.32
2	18.34	21.00	21.00–18.34 = 2.66	19.05–18.34 = 0.71	4.74	1.03
3	19.05	21.75	21.75–19.05 = 2.70	20.07–19.05 = 1.02	8.44	2.05
4	20.07	22.55	22.55–20.07 = 2.48	20.36–20.07 = 0.29	9.92	2.34
5	20.36	22.95	22.95–20.36 = 2.59	20.65–20.36 = 0.29	12.51	2.63
6	20.65					

The ratio estimate of the realized heritability (Equation 6.14) is

$$\hat{h}_r^2 = \frac{2.63}{12.51} = 0.2102$$



The regression, forced through the origin, of cumulative response ( $R_C$ ) on cumulative selection differential ( $S_C$ ) is plotted above. Equation 6.13 gives the OLS regression estimator of the realized heritability as

$$\hat{h}_r^2 = \hat{b}_C(OLS) = \frac{\sum_{i=1}^5 S_C(i) \cdot R_C(i)}{\sum_{i=1}^5 S_C^2(i)} = \frac{78.96}{350.45} = 0.2245$$

---

The theory for estimating realized heritabilities (via a LS analysis) in populations with overlapping generations is less well developed. Approaches assuming an asymptotic selection response are flawed in that many generations are required to reach a stable genetic structure starting from an unselected base population (**Examples x and y from Chapter 8**). For nonasymptotic response, see Hill (1974b) and Johnson (1977b) for the relevant theory and Atkins and Thompson (1986) for an example with Scottish Blackface sheep. Mixed-model approaches (Chapter 7) easily accommodate overlapping generations.

#### Weighted Least-Squares Estimates of Realized Heritability

Ordinary least-squares regression assumes that the residuals are homoscedastic and uncorrelated, so that the covariance matrix for the vector of residuals  $\mathbf{e}$  is  $\text{Var}(\mathbf{e}) = \sigma_e^2 \cdot \mathbf{I}$  (LW Chapter 8). However, genetic drift causes the covariance structure of the regression given by Equation 6.12 to depart significantly from this simple form. In particular, the residual variance increases with time (Equation 6.6) and residuals from different generations are correlated (Equation 6.9). Thus,  $\text{Var}(\mathbf{e}) = \mathbf{V}$  and generalized least-squares (GLS) must be used to account for this covariance structure. From LW Equation 8.34, the GLS estimator of the regression slope is given by

$$\hat{b}_C(\text{GLS}) = (\mathbf{S}^T \mathbf{V}^{-1} \mathbf{S})^{-1} \mathbf{S}^T \mathbf{V}^{-1} \mathbf{R} \quad (6.15)$$

where  $\mathbf{V}$  is the variance-covariance matrix associated with selection response,

$$V_{ij} = \sigma_e(i, j) = \sigma [R_C(i), R_C(j)]$$

The elements of  $\mathbf{V}$  can be obtained from the pure-drift approximation, with the variances  $V_{ii}$  given by Equation 6.6 and covariances  $V_{ij}$  given by Equation 6.9. Even though OLS assumes an incorrect residual structure, it still provides an unbiased estimate of  $b_C$ . However, OLS significantly underestimates the standard error of the OLS estimator, and it is this reason that GLS estimators are greatly preferred and should be used when ever possible (Hill 1971; 1972d,e).

While Richardson et al. (1968) and Irgang et al. (1985) suggested using weighted least-squares to account for differences in the variance of residuals that occur when different number of adults are measured in different generations,

the full power of GLS regressions incorporating the drift-generated covariance structure was developed by Hill (1971, 1972d,e, 1974c, 1977b, 1980, 1986).

**Example 2.** To compute the GLS regression using the data from Example 1, we need to variance-covariance matrix of the residuals, which we obtain by using the pure-drift approximation. From Example 1,  $M = 100$ ,  $N = 20$ , while the estimated phenotypic variance is  $\text{Var}(z) = 3.293$  (Mackay, personal communication). Assuming that both initial sampling and between-generation environmental effects can be ignored (i.e.,  $M_0 \gg 1$  and  $\sigma_d^2 \simeq 0$ ), then from Equations 6.6 and 6.7, the pure-drift approximation gives the variance associated with the response in generation  $i$  as:

$$\sigma^2 [R_C(i)] = \left( \frac{i}{N} \right) h^2 \sigma_z^2 + \frac{\sigma_z^2}{M} = i \cdot h^2 \cdot 0.1647 + 0.03292$$

Similarly, the pure-drift approximation (Equations 6.9 and 6.7) for the covariance between generations is

$$\sigma [R_C(i), R_C(j)] = \left( \frac{i}{N} \right) h^2 \sigma_z^2 = i \cdot h^2 \cdot 0.1647 \quad \text{for } i < j$$

The resulting covariance matrix becomes

$$\mathbf{V} = 0.1647 \cdot \begin{pmatrix} h^2 + 0.2 & h^2 & h^2 & h^2 & h^2 \\ h^2 & 2h^2 + 0.2 & 2h^2 & 2h^2 & 2h^2 \\ h^2 & 2h^2 & 3h^2 + 0.2 & 3h^2 & 3h^2 \\ h^2 & 2h^2 & 3h^2 & 4h^2 + 0.2 & 4h^2 \\ h^2 & 2h^2 & 3h^2 & 4h^2 & 5h^2 + 0.2 \end{pmatrix}$$

Since the matrix  $\mathbf{V}$  is a function of the unknown heritability, estimation is an iterative process, starting with some initial estimate of  $h^2$ , updating  $\mathbf{V}$  in subsequent iterations with the current estimate until convergence. For a starting value, we will use the ratio estimate  $h^2 = 0.21$ . Applying Equation 6.15 gives a first estimate as

$$= \hat{b}_C(GLS)^{(1)} = (\mathbf{S}^T \mathbf{V}^{-1} \mathbf{S})^{-1} \mathbf{S}^T \mathbf{V}^{-1} \mathbf{R} = 0.222197$$

Substituting this new value into  $\mathbf{V}$  gives upon a second iteration  $\hat{b}_C(GLS)^{(2)} = 0.222135$ , which remains unchanged upon subsequent iteration.

## Standard Errors for Realized Heritability Estimates

The final piece of statistical machinery necessary for assessing the success of the breeders' equation is computation of the standard errors for realized heritability estimates. Consider first the realized heritability estimated from the unweighted regression (Equation 6.13). Recalling LW Equation 8.33b for the variance for an OLS estimator,

$$\begin{aligned}\text{Var}[\hat{b}_C(\text{OLS})] &= \sigma_e^2 (\mathbf{X}^T \mathbf{X})^{-1} = \sigma_e^2 (\mathbf{S}^T \mathbf{S})^{-1} \\ &= \sigma_e^2 / \sum_{i=1}^T S_C^2(i)\end{aligned}\quad (6.16a)$$

The residual variance  $\sigma_e^2$  can be estimated from the residual sums of squares divided by the degrees of freedom (see LW Chapter 8). Here

$$\hat{\sigma}_e^2 = \frac{1}{T-1} \sum_{i=1}^T \hat{e}_i^2 = \frac{1}{T-1} \sum_{i=1}^T \left( R_C(i) - \hat{h}_r^2 S_C(i) \right)^2 \quad (6.16b)$$

As mentioned, because the OLS estimator assumes residuals are uncorrelated and have equal variances (both of which are incorrect), it significantly underestimates the correct variance (see Example 3). The GLS regression estimator (Equation 6.15) avoids these problems by properly accounting for the variance structure. From standard GLS theory (LW Equation 8.35),

$$\text{Var}[\hat{b}_C(\text{GLS})] = (\mathbf{S}^T \mathbf{V}^{-1} \mathbf{S})^{-1} \quad (6.17)$$

As above, the pure drift approximation (Equation 6.6 and 6.9) is generally used to obtain the elements of  $\mathbf{V}$ , with  $\hat{h}_r^2$  used in place of  $h^2$ .

Finally, consider the variance for the estimator  $b_T$ , the ratio of total response to total selection (Equation 6.14). Since  $\text{Var}(y/c) = \text{Var}(y)/c^2$  for a constant  $c$ , it immediately follows that

$$\text{Var}(\hat{b}_T) = \frac{\text{Var}[R_C(T)]}{S_C^2(T)} \simeq \frac{(T/N) \hat{h}_r^2 \sigma_z^2 + \sigma_z^2/M}{S_C^2(T)} \quad (6.18)$$

To obtain the variance in response in Generation  $T$ , we again used the pure-drift approximation (Equation 6.6), assuming that initial sampling can be ignored ( $M_0 \gg 1$  and no significant between-generation environmental variance ( $\sigma_d^2 = 0$ )).

Hill (1972d,e) noted that  $\hat{b}_C(\text{GLS})$  is generally a slightly better estimator than  $\hat{b}_T$  when  $h^2$  is small, while  $\hat{b}_T$  is slightly better estimator when  $h^2$  and/or the number of generations is large. Hill suggests computing the standard errors for both methods and using the estimator with the smaller SE.

**Example 3.** Using the data from Examples 1 and 2, we compare the standard errors associated with the three different realized heritability estimates (total response, weighted and unweighted regressions). Consider the unweighted regression estimator  $\hat{b}_C(\text{OLS})$  first. The residual sums of squares is

$$\sum_{i=1}^T \left( R_C(i) - \hat{h}_r^2 S_C(i) \right)^2 = 0.091$$

giving an estimated residual variance of  $\hat{\sigma}_e^2 = 0.091/4 = 0.0228$ . Equation 6.16a gives

$$\text{Var} \left[ \hat{b}_C(\text{OLS}) \right] = \sigma_e^2 / \sum_{i=1}^T S_C^2(i) = \frac{0.0228}{350.45} = 0.0000649$$

Taking the square root gives the standard error as 0.0081. Turning to the estimate  $\hat{b}_T$  based on the total response to total selection, Equation 6.18 (using the pure-drift approximation) gives

$$\text{Var}(\hat{b}_T) = \frac{(5/20) \cdot 0.21 \cdot 3.292 + 0.03292}{12.51^2} = 0.00132$$

for a standard error of 0.0363. Finally, substituting the GLS estimate of  $\hat{h}_r^2 = 0.222135$  in  $\mathbf{V}$  (Example 2), Equation 6.17 gives variance of this estimate as

$$(\mathbf{S}^T \mathbf{V}^{-1} \mathbf{S})^{-1} = \frac{1}{790.4} = 0.00126554$$

for a standard error of 0.0356.

In summary, the three approaches give extremely similar estimates,

Unweighted least-squares regression, $\hat{b}_C(\text{OLS})$	$\hat{h}_r^2 = 0.2245 \pm 0.0081$
Total response/total differential, $\hat{b}_T$	$\hat{h}_r^2 = 0.2102 \pm 0.0363$
Weighted least-squares regression, $\hat{b}_C(\text{GLS})$	$\hat{h}_r^2 = 0.2221 \pm 0.0356$

Note that the difference in the standard error between  $\hat{b}_C(\text{GLS})$  and  $\hat{b}_T$  is very small, with  $\hat{b}_T$  considerably more straightforward to compute. Also note that the unweighted regression badly underestimates the standard error.

---

## Empirical vs. Predicted Standard Errors

While the standard error in the response can be directly estimated from the between-line variance in a series of replicate lines subjected to identical selection, this approach is generally not cost-effective. As a result, the standard errors reported for most experiments used the pure-drift approximation discussed previously. How closely do the empirical (observed) standard errors match those predicted from the drift approximation? Unfortunately, very few experiments have addressed this issue.

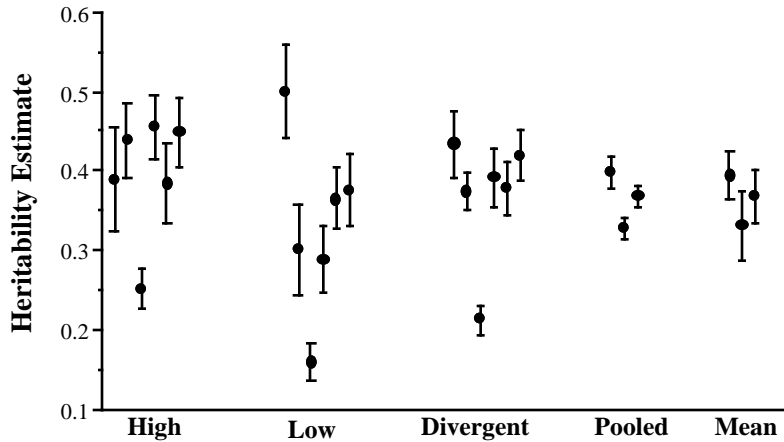
Perhaps the most extensive study is that of Falconer (1973), who performed two way selection for 6-week weight in mice. A total of six replicate sets of lines were used. Each set consisted of a line selected for larger size, a line selected for smaller size, and an unselected control, for a total of 18 lines in the entire experiment. Realized heritabilities were estimated (by OLS regression) using the three different contrasts available within each replicate set: large versus control (high lines), small versus control (low lines), and large versus small (divergent lines).

As shown in Table 6.1, the average realized heritability estimates under each of the three different contrasts are very similar — 0.395, 0.331, and 0.369 for high, low, and divergent lines (respectively). Likewise, the three pooled estimates, obtained by OLS regression of the mean response (over the six replicates) on the mean selection differential, give very similar estimates. However, the standard errors for the mean estimates (based on variation between the replicate lines) were 1.6, 3.3, and 2.3 times as large as the estimated standard error using OLS regression of the pooled data.

**Table 6.1** OLS Realized heritability estimates (Equation 6.13) for Falconer’s selection experiments on bodyweight in mice. Standard errors were computed using the OLS solution given by Equation 6.16. The **pooled** estimate is obtained by using the means for a given contrast (high, low, or divergence) over all six replicate lines for the data points in the regression and then using the OLS estimate and standard error. The **mean** estimate is the average of the realized heritability estimates over all six replications and the associated standard error is the variation about this mean seen in the replicates.

Replicate	High Lines		Low lines		Divergent lines	
	$\hat{h}_r^2$	SE	$\hat{h}_r^2$	SE	$\hat{h}_r^2$	SE
1	0.390	0.066	0.501	0.058	0.434	0.042
2	0.438	0.047	0.301	0.058	0.375	0.024
3	0.251	0.025	0.159	0.024	0.212	0.020
4	0.457	0.041	0.288	0.041	0.392	0.037
5	0.385	0.051	0.365	0.039	0.379	0.033
6	0.448	0.043	0.376	0.046	0.420	0.032
Pooled	0.398	0.020	0.328	0.014	0.369	0.014
Mean	0.395	0.031	0.331	0.046	0.369	0.033

Figure 6.2 displays the range in estimates and standard errors. The SEs for the pooled estimates (computed from OLS) are much less than the true SEs computed from the observed variation between replicates. Note, however, that the OLS-computed standard errors for any given replicate are not substantially different from the observed standard errors (compare the replicates with the results for the mean estimates). One caveat is that, as mentioned previously, standard errors obtain by OLS regression are expected to underestimate the true SEs.



**Figure 6.2.** Realized heritability estimates (and their associated standard errors) for Falconer's (1973) selection experiment in mice. See Table 6.1 for more details.

A second experiment (López-Fanjul and Domínguez 1982, summarized by López-Fanjul 1982) followed twenty replicate *Drosophila* lines selected for sternopleural bristle number. They found that the empirical standard errors were less than OLS-estimated SEs, which in turn were less than GLS-estimated SEs. This is contrary to the expectation that the OLS-estimated SEs is smallest, with the empirical and GLS-estimated SEs being expected to be roughly equal. The authors suggest that this may be at least partly due to a scale effect (LW Chapter 11), as the phenotypic variance greatly decreased during selection.

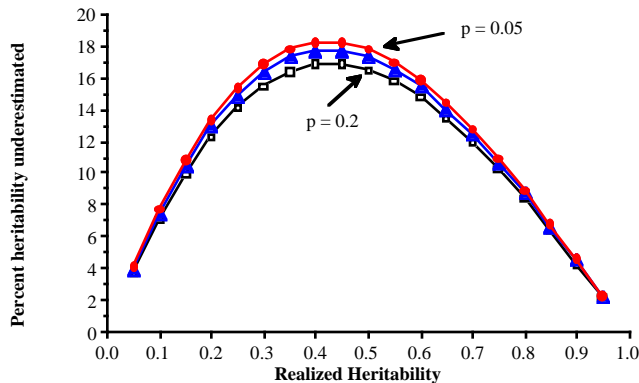
A final experiment (Bohn et al. 1983) examined the variance in response to selection on pupa weight in *Trioblium castaneum*. The authors found a poor fit when using Hill's (1971) variance formulae that attempt to correct for the effects of selection. The authors examined three replicates from each of six different base populations. Interestingly, a strong correlation was observed between departures in the predicted variance and departures from normality in the base population. The variance among replicate lines drawn from base populations with increasing (absolute) amounts of kurtosis showed larger departures from the predicted value.

### Infinitesimal-model Corrections for Disequilibrium

Our final comment on estimation is that selection is expected to decrease the additive variance (and hence the heritability) due to the generation of gametic-phase disequilibrium (Chapters 5, 10). Hence, the realized heritability is depressed relative to the base population heritability (Figure 6.2). Under the infinitesimal model, the majority of this decrease occurs over the first few generations, after which the heritability is essentially at its equilibrium value,  $\tilde{h}^2$ , where (Equations 5.13a and 5.13c),

$$\tilde{h}^2 = \frac{\theta}{1 + \theta - h^2}, \quad \text{with} \quad \theta = \frac{2h^2 - 1 + \sqrt{1 + 4h^2(1 - h^2)(1 - \kappa)}}{2(2 - \kappa)} \quad (6.19)$$

Here  $h^2$  is the (unselected) base-population heritability and  $\kappa$  is a measure of the reduction in phenotypic variance due to selection. In particular,  $\kappa = 1 - \bar{v}(\bar{v} - z_{[1-p]})$  for truncation selection saving a fraction  $p$  of the population (Table 5.1). Thus, one can treat the realized heritability as an estimate of the equilibrium heritability ( $\hat{h}_r^2 \simeq \tilde{h}^2$ ) and numerically solve for the base-population heritability ( $h^2$ ). This correction was first applied to realized heritability estimates by Atkins and Thompson (1986). Figure 6.2 shows the relative percentage by which the base population heritability is underestimated by the uncorrected realized heritability.



**Figure 6.2.** The relative percentage by which the base population heritability is underestimated by the realized heritability due to reduction in the additive variance by gametic-phase disequilibrium. The three curves correspond to different levels of truncation selection, with 5 (upper curve), 10 (middle curve), and 20 (lower curve) percent of the population saved. Values were obtained by numerically solving Equation 6.19 and assuming the infinitesimal model (no significant selection-induced changes in allele frequencies).

**Example 4.** What are the disequilibrium-corrected estimates for the realized heritability values obtained in Examples 1 and 2? Recall that for this experiment,  $p = 20/100 = 0.2$ . From Example 2 in Chapter 5, this value of  $p$  gives  $\kappa = 0.213$ . Substituting this value into Equation 16.19 and numerically solving for each of the three previous estimates gives infinitesimal model corrected realized heritabilities of:

Estimator	Estimate of $h_r^2$	
	Uncorrected	Corrected
Unweighted least-squares regression, $\hat{b}_C(\text{OLS})$	0.2245	0.2541
Total response/total differential, $\hat{b}_T$	0.2102	0.2368
Weighted least-squares regression, $\hat{b}_C(\text{GLS})$	0.2221	0.2512

Assuming that the infinitesimal model is a reasonable assumption over the term of this experiment, all three methods underestimated the base population heritability by about 13 percent (see Figure 6.2).

## EXPERIMENTAL EVALUATION OF THE BREEDERS' EQUATION

Although the breeders' equation requires a number of assumptions (Table 4.1) and strictly speaking holds only for a single generation of selection from an unselected base population, the general claim (e.g., Falconer 1981) is that it is usually satisfactory over a few (5-10) generations. A slightly more refined statement is that the time must not exceed  $N_e/2$  generations (Hill 1977a), as drift significantly changes the genetic variance within a line after this time. After a sufficient number of generations, drift and selection change the genetic variances significantly from their base-population values and the breeders' equation (using the initial heritability value) fails.

What do the actual data say about the adequacy of the breeders' equation? Starting with the first formal comparisons by Reeve and Robertson (1953) and Clayton et al. (1957), a number of authors have compared their observed short-term responses with those predicted from the breeders' equation. As summarized below, the results are mixed. One complication is that most analyses simply used ordinary (unweighted) least squares, resulting in significantly underestimated standard errors. Likewise, almost all realized heritability estimates have not corrected for the expected decline due to gametic-phase disequilibrium. Only a few selection experiments have used the more powerful mixed-model (REML/BLUP) design in place of least-squares designs, and we discuss these in Chapter 7.

### Sheridan's Analysis

One of the most extensive reviews of the fit of the breeders' equation is by Sheridan (1988), who examined 198 experiments involving laboratory and domesticated animals, comparing realized heritabilities with estimates of heritability based on resemblances between relatives. Sheridan first considered those experiments with an extensive past history of selection (indeed, the base populations appeared to be at a selective plateau). In these populations the response is very poorly predicted by the breeders' equation, with all 11 experiments (6 in *Drosophila*, 2 in *Tribolium*, 3 in mice) showing greater than 50 percent disagreement between the realized and estimated (i.e., base-population) heritabilities.

Table 6.2 shows the fit for the remaining 187 experiments whose base populations were apparently not at a selective plateau. As the data show, the fit was rather poor in many experiments — almost half have a disagreement of at least 30 percent, and one in three exceeds 50 percent. Besides the biological reasons listed in Table 4.1, there are also design issues that could account for the apparent poor fit. First, none of the experiments reviewed by Sheridan corrected for the expected decline in the realized heritability due to gametic-phase disequilibrium. Second, small absolute disagreements can translate into large relative percentages for those traits with low heritabilities (Hill and Caballero 1992). For example, if  $|\hat{h}_r^2 - \hat{h}^2| = 0.04$ , this is a 20 percent relative disagreement if  $\hat{h}_r^2 = 0.2$  but an 80 percent disagreement if  $\hat{h}_r^2 = 0.05$ . Finally, some fraction of the disagreement could simply be due to variance in response and thus the differences are not significant.

**Table 6.2.** Comparison between estimates of realized heritability ( $\hat{h}_r^2$ ) and heritability estimates based on resemblances between relatives ( $\hat{h}^2$ ). Within each group, the table gives the distribution of the percent absolute disagreement ( $|\hat{h}^2 - \hat{h}_r^2|/\hat{h}_r^2$ ) between the two estimates, while  $n$  is the number of experiments considered for each species group. For example, 8% of the 60 *Drosophila* experiments had a percent absolute disagreement between estimates of 30 to 50 percent. After Sheridan (1988).

Species	Percent absolute disagreement (relative to $\hat{h}_r^2$ )					n
	0–10%	10–20%	20–30%	30–50%	>50%	
<i>Drosophila</i>	48%	12%	5%	8%	27%	60
<i>Tribolium</i>	31%	8%	12%	38%	12%	26
Mice and Rats	23%	0	13%	23%	41%	39
Poultry and Quail	20%	5%	17%	5%	54%	41
Swine and Sheep	20%	0%	10%	20%	50%	10
<b>Summary over all groups</b>						
Laboratory Species	37%	7%	9%	19%	28%	125

Commerical Species	19%	5%	18%	11%	47%	62
All Species	31%	6%	12%	17%	34%	187

To address the issue of significance, Sheridan considered the disagreement to be significant if it exceeded two standard errors. Assuming the two estimates are uncorrelated, the variance for their difference is the sum of the variance for each estimate, giving a standard error of

$$SE(\hat{h}^2 - \hat{h}_r^2) = \sqrt{(SE[\hat{h}^2])^2 + (SE[\hat{h}_r^2])^2} \quad (6.20)$$

Taking those 131 experiments that have both standard errors (for  $\hat{h}_r^2$  and  $\hat{h}^2$ ), Table 6.3 shows that 25 percent had realized heritabilities significantly different from  $\hat{h}^2$  by this criteria. One problem with this approach is that most of the reported standard errors used by Sheridan were based on unweighted least-squares (OLS), which underestimates the true standard error, giving confidence intervals that are too narrow.

**Table 6.3.** Tests of significance between estimated and realized heritabilities. Differences of more than two standard errors (computed from Equation 6.20) are regarded as significant. Only those experiments with estimated standard errors for both heritability estimates are included, and  $n$  denotes the number of such experiments. After Sheridan (1988).

Species	Significant Differences	NS Differences	Total
<i>Drosophila</i>	14 (23%)	47 (77%)	61
<i>Tribolium</i>	7 (27%)	19 (73%)	26
Mice and Rats	6 (18%)	28 (82%)	34
Poultry and Quail	5 (45%)	6 (55%)	11
Swine and Sheep	8 (53%)	7 (47%)	15
<b>Summary over all groups</b>			
Laboratory Species	27 (21%)	104 (79%)	131
Commerical Species	11 (37%)	19 (63%)	30
All Species	38 (25%)	113 (75%)	151

Finally, Sheridan looked at the goodness-of-fit as a function of the duration of the experiment (Table 6.4). Surprisingly, longer experiments tended to have a better fit. While this is contrary to expectations, this could also be design artifact. First, longer experiments tend to have smaller standard errors, as the SE scales

as the inverse of the total selection differential. Second, in many cases longer experiments may employ larger population sizes than experiments of shorter duration, reducing the effect of drift.

**Table 6.4.** Agreement between realized and estimated base-population heritability as a function of the duration of the experiment. After Sheridan (1988).

Generations	Percent absolute disagreement (relative to $\hat{h}_r^2$ )			n
	0–10%	10–30%	> 30 %	
1 - 5	18%	27%	55%	44
6 - 10	24%	17%	59%	98
10 - 15	52%	10%	38%	90

**Realized Heritabilities and Selection Intensity**

In addition to quantitative differences in the predictions of the breeders’ equation discussed above, there are also reported cases of major qualitative departures as well. For example, the breeders’ equation predicts that while the response should increase with selection intensity, the ratio of response to selection differential,  $R/S$ , should be constant. Some studies have reported a dependence of realized heritability on the selection intensity, although a survey of selection experiments finds no consistent patterns between the two (Table 6.5).

There are several reasons why some dependence may arise between realized heritability and selection intensity. First, increasing selection intensity increases gametic-phase disequilibrium, reducing  $\sigma_A^2$  and hence  $\hat{h}_r^2$ . As Figure 6.2 shows, the prediction is that (uncorrected) realized heritabilities should decrease with increasing selection intensity. Again, essentially none of the reported selection experiments correct for this expected reduction. Second, with increasing selection, allele frequencies are expected to change more rapidly. Whether this results in an increased or decreased response depends on the initial distribution of allele frequencies and effects (see the discussion on genetic asymmetries below). Finally,  $N_e$  decreases as selection intensity increases (Chapter 12), increasing the amount of inbreeding and (generally) reducing the additive variance (Chapter 3). Thus, lines experiencing different amounts of selection have different amounts of inbreeding, even if the census sizes are identical.

**Table 6.5.** Summary of experiments examining the effects of selection intensity on estimated realized heritability,  $\hat{h}_r^2$ .

Clayton et al. 1957 abdominal bristles in <i>D. melanogaster</i>	$\hat{h}_r^2$ decreases with increasing selection intensity.
Frankham et al 1968a	Agreement between base population estimate

abdominal bristles in <i>D. melanogaster</i>	of $h^2$ and $\widehat{h}_r^2$ best at highest selection intensity, becoming worse as selection intensity decreases.
Hanrahan et al. 1973 postweaning weight gain in mice	No consistent effect of selection intensity on $\widehat{h}_r^2$ .
Meyer and Enfield 1975 pupa weight in <i>Tribolium castaneum</i>	$\widehat{h}_r^2$ decreases with selection intensity in down-selected lines; no effect in up-selected lines.
Silvela et al. 1989 kernal oil content in maize	No effect of selection intensity on $\widehat{h}_r^2$ .

---

### Inbreeding and Short-term Response

When inbreeding occurs, the short-term response is generally found to be less than that predicted from the breeders' equation using variance components estimated from the base population (Table 6.6). This is expected, as inbreeding (generally) decreases the additive genetic variance within a line, reducing response. The rough rule of thumb (Hill 1977a) is that the effects of drift on reducing the within-line variation can be ignored provided that the selection experiment lasts less than  $N_e/2$  generations. Exceptions can occur if nonadditive variance is present, in which case inbreeding may actually result in an *increase* in the within-line additive variance (Chapter 3), increasing response. The consequences for drift on long-term experiments (those greatly exceeding  $N_e/2$  generations) are examined in detail in Chapter 12.

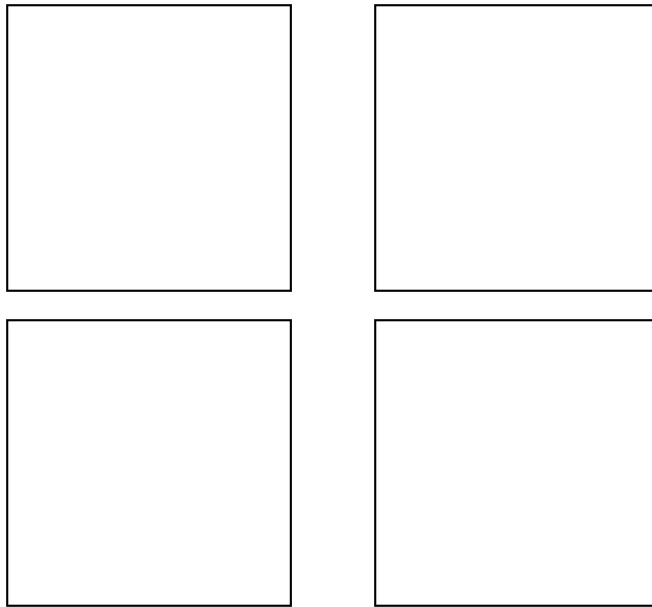
**Table 6.6.** Results of experiments examining the effects of finite population size and inbreeding on short-term response.

---

Tantawy and Reeve 1956 Wing length in <i>D. melanogaster</i>	Short-term response for outbreds > double first cousins > sib mating.
Chung and Chapman 1958 Gonadotrophic hormone level in rats	Outbred and crossbred lines had larger short-term responses than inbred lines.
Lewis and Warwick 1953 Body size in mice	Outbred line had a larger short-term response than an inbred line.
Frankham et al. 1968a Abdominal bristles in <i>D. melanogaster</i>	Weak trend for larger population sizes to have a greater short-term response
Hanrahan et al. 1973 Postweaning weight gain in mice	Short-term response increased with increasing population size
Silvela et al. 1989 Kernal oil content in maize	Short-term response increased with increasing population size

### Asymmetric Selection Response

A common design is to perform a **divergent selection experiment**, wherein replicate lines are selected in opposite directions. Many such experiments (e.g., Figure 6.3) show different amounts of response in the up versus down direction, a phenomenon referred to as an **asymmetric selection response** (Falconer 1954). This is in sharp contrast with the expectation from the breeders' equation, which predicts that the absolute magnitude of response should depend only the absolute value of  $S$ .



**Figure 6.3.** Examples of asymmetric response to selection. **A:** Abdominal bristles in male *Drosophila melanogaster* (Sheldon 1963). **B:** Rearing activity in rats (Sanders 1981). **C:** Twelve-week body weight in chickens (Maloney et al. 1963). **D:** Weight at 6 weeks in mice (Falconer 1954). The horizontal axis is in generations in **A** and **B** and is the cumulative selection differential in **C** and **D**.

There are a variety of possible explanations for asymmetric responses (Table 6.7). It may simply be an artifact of the experimental design and/or analysis. In particular, the prediction of equal positive and negative slopes holds only for plots of cumulative response versus cumulative selection differentials ( $R_C(t)$  vs.  $S_C(t)$ ). Asymmetry in response based on differences in slope of cumulative

response versus generations of selection ( $R_C(t)$  vs.  $t$ ) can thus be very misleading as the different lines may have experienced different amounts of selection.

**Table 6.7.** Possible explanations for asymmetric response (including reversed response).

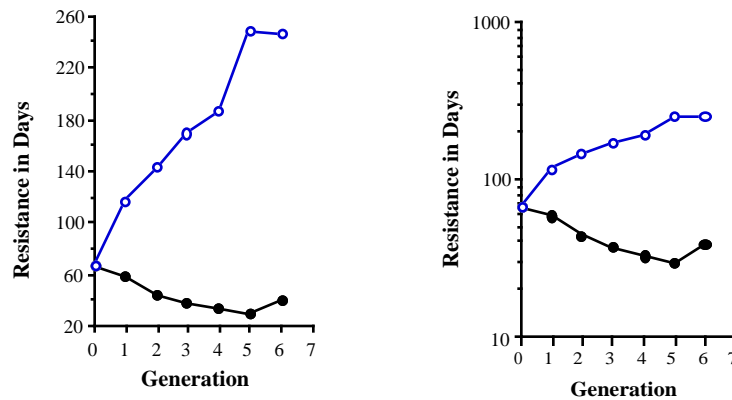
---

<b>Design Defects</b>	<ul style="list-style-type: none"> <li>Drift</li> <li>Scale effects</li> <li>Different effective selection differentials</li> <li>Undetected environmental trends</li> <li>Transient effects from previous selection in the base population</li> <li>Undetected selection on correlated characters</li> </ul>
<b>Nonlinear parent-offspring regressions</b>	<ul style="list-style-type: none"> <li>Major genes with dominance</li> <li>Genotype-environment interactions</li> <li>Departures from normality</li> </ul>
<b>Other sources</b>	<ul style="list-style-type: none"> <li>Genetic asymmetries</li> <li>Inbreeding depression</li> <li>Maternal effects</li> </ul>

---

Even if the amount of artificial selection is the same in both directions, there may be major differences in natural selection — e.g., up-selected lines may experience lower fertility and thus have a lower effective selection differential. Using effective selection differentials (Equation 4.8) allows for some correction, but the investigator often lacks the data (e.g., fertilities for each parent) necessarily to compute them.

Even though lines may have quite different values of  $\widehat{h}_r^2$ , there is still the issue of whether these differences are significant. As we have seen, genetic drift can generate considerable variation between replicate lines and it is important to distinguish between a real difference in response versus the expected variation in two realizations of a process with the same (absolute) expected value. Directional trends in environmental change can also produce asymmetry, accentuating the response in the direction of the trend and retarding response in the opposite direction. Finally, differences in response could simply be scale effects (Figure 6.4, see LW Chapter 11). For example, if the genetic variance increases with the mean, heritability can increase with the mean, giving a faster response in the upwardly selected lines.



**Figure 6.4.** An example of a scale effect generating an asymmetric selection response. **Left:** Selection for resistance to dental caries in albino rats (*Rattus norvegicus*), response measured as the expected number of days to develop caries on a standard diet. **Right:** Same data on a log scale. Based on Falconer's (1954) analysis of data from Hunt et al. (1944).

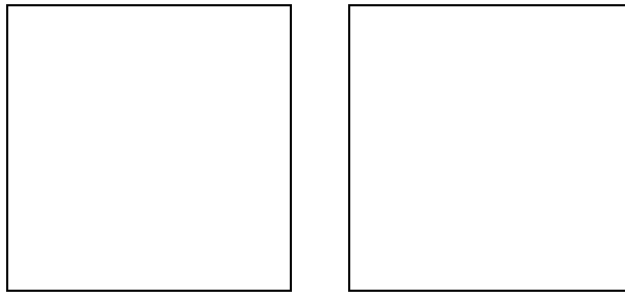
Although spurious asymmetric responses can result from defects in design and analysis, true asymmetric responses can be generated by a variety of genetic situations. For example, if the parent-offspring regression is nonlinear,  $S$  is not sufficient to predict response (Chapters 9, 10) and it is not surprising that asymmetric responses can be generated in such cases. Failure to detect departures from linearity in the base population does not rule out nonlinearity as an explanation for an observed asymmetric response. The range of variation in the base population may not be sufficient to detect departures at the extreme ends of the initial range of phenotypes. As the selected lines diverge, differences in the tails of the initial phenotypic distribution can become quite important.

Characters displaying inbreeding depression (LW Chapter 10) show asymmetric selection response, accentuating the response in one direction and retarding it in the other. A simple test for inbreeding as an explanation of asymmetric response is to see whether the mean of an unselected control population changes in the direction of greater response when inbred. The effects of inbreeding depression can be corrected for by inbreeding a control population to the same level as the divergent selection lines, using the contrast between selected and control lines to estimate response. However, this is not necessarily as straightforward as it appears, as selection generally increases the amount of inbreeding within a line by decreasing the effective population size (Chapter 12). Thus, simply keeping the control line at the same size as the selected lines underestimates the amount of inbreeding, especially when selection is intense.

When major alleles are segregating, asymmetries in response arise as a con-

sequence of allele frequencies changing in different directions in the selected populations, as an increase in an allele from its initial frequency results in a different additive variance from that produced by an equal decrease in allele frequency. Falconer (1954) refers to this feature as **genetic asymmetry**. This is most easily seen by considering LW Figure 4.6, which shows additive genetic variation as a function of allele frequency for a single diallelic locus. If alleles are completely additive,  $\sigma_A^2$  as a function of allele frequency is symmetric about  $p = 1/2$ . Suppose that the frequency of an allele that increases the character value is initially below  $1/2$ . In upwardly-selected lines, this allele increases in frequency (ignoring drift), resulting in an increase in  $\sigma_A^2$  and  $h^2$  as the allele frequency increases to  $1/2$ . Conversely, in downwardly-selected lines,  $\sigma_A^2$  always decreases. If dominance is present,  $\sigma_A^2$  is no longer a symmetric function of allelic frequencies (LW Figure 4.8) and asymmetric changes in the contribution to  $\sigma_A^2$  from a single locus are almost always expected.

Frankham and Nurthen (1981) give an interesting example wherein a base population was constructed with the major recessive allele  $sm^{lab}$  (which greatly reduces abdominal bristle number in *Drosophila melanogaster*) initially at low frequency. As shown in Figure 6.5, in most down-selected lines, this allele increased in frequency, resulting in a large increase in  $h^2$  as the  $sm^{lab}$  allele reaches intermediate frequencies. Heritability subsequently returns to the base population value as this allele becomes fixed.



**Figure 6.5.** **Left:** Asymmetric selection response for abdominal bristle number in *Drosophila melanogaster* from a base population containing a major allele ( $sm^{lab}$ ) initially at low frequency. L1-L3 are three replicate low lines, H1-H3 three replicate high lines and C the control line. Data were log-transformed to remove scale effects. **Right:** Changes in heritabilities in the control and lines L1-L3. Heritability was estimated from phenotypic correlations between bristle numbers of adjacent segments of the same fly (see Frankham and Nurthen 1981 for details). Note the large increases in heritability for lines L2 and L3, the lines that show an asymmetric selection response (relative to H1- H3), while the heritability is roughly constant in L1 which does not show an asymmetric response. The increase in  $h^2$  reflects an increase in the major allele  $sm^{lab}$  due to selection. This allele increased rapidly

in frequency after generation 5 and was essentially fixed by generation 10. This is reflected by a rapid increase in  $h^2$  after generation 5, with  $h^2$  returning to normal (the level in the unselected controls) as the allele becomes fixed. After Frankham and Nurthen (1981).

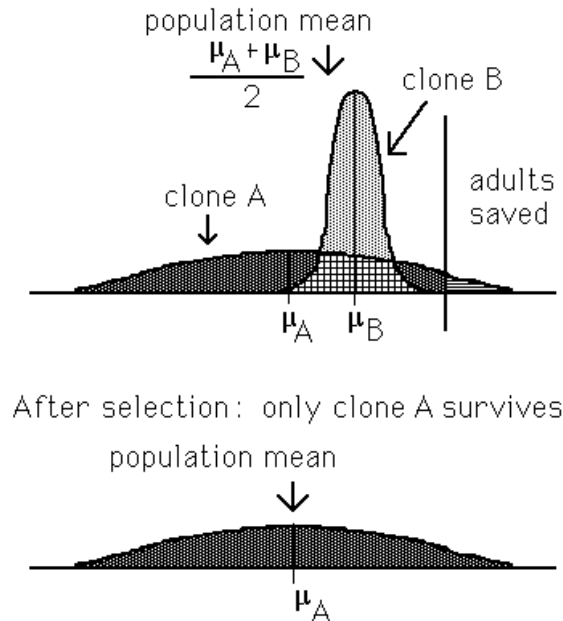
While we have focused on the effects of a single major gene, the effects of unequal allele frequencies and dominance apply to any QTL. Alleles of small effect are expected to have much slower changes in allele frequencies (and thus expected to have smaller effects on asymmetry) over the time scales of most short-term experiments. Hence, if many loci of small effect underlie a character, the effects of genetic asymmetry on selection response are expected to be slight unless the number of generations is large. Further, at least some cancellation is expected between the asymmetries in  $\sigma_A^2$  generated by allele frequency changes at different loci, analogous to directional dominance being required to generate significant inbreeding depression (LW Chapter 10).

Components of reproductive fitness, such as fecundity and development time, are expected to have asymmetric allele frequencies as natural selection increases the frequency of alleles increasing fitness. Such characters are also expected to have substantial nonadditive genetic variance (Chapter 9). These conditions suggest that asymmetric responses in reproduction are expected, and further predict that response should be larger in lines selected for a decrease in reproductive fitness. This was seen by Frankham (1990), who found a larger response in the direction of reduced reproductive fitness in 24 of 30 experiments reviewed. Frankham suggested that the presence of rare recessives decreasing reproductive fitness was the most likely cause for this trend. As Figure 11.2 shows, very marked asymmetric responses are expected in such situations.

### Reversed Response

The most extreme departure from the breeders' equation is a **reversed response**, a response in the opposite direction of selection. Negative maternal effects can result in such a response (Figure 4.8). Likewise, they can arise from sufficiently large genotype-environment interactions (Haldane 1931). Haldane's intuition was based on selection on two asexual clones, the first with a higher mean but smaller environmental variance than the second. As shown in Figure 6.6, if the most extreme individuals are selected, there is an excess of the clone with the smaller mean but higher variance, resulting in a decrease in mean value in the next generation. Wright (1969) expanded Haldane's model to a single diallelic locus. Gimelfarb (1986) gives a particularly interesting analysis when there is a multiplicative genotype  $\times$  environment interaction, showing for this model that while *phenotypes* are subjected to directional selection, the nature of the interaction between genotype and environment is such that *genotypic* values are actually under either stabilizing or disruptive selection. The interesting feature of both Haldane's and

Gimelfarb's models is that reversed response is most probable when selection is very intense.



**Figure 6.6.** Haldane's (1931) example of a reversed response generated by genotype  $\times$  environment interaction. The population consists of equal numbers of two asexual clones *A* and *B*. The mean phenotypic value of *A* ( $\mu_A$ ) is less than the mean phenotypic value of *B* ( $\mu_B$ ). However, clone *A* has a larger environmental variance than *B*. Strong truncation selection culls out all *B* clones in the population, leaving only *A*. The new mean in the generation following selection  $\mu_A$  is less than the mean before selection  $(\mu_A + \mu_B)/2$ . In this case, selection for *increased* character value *z* resulted in a *lowering* of the population mean.

## RESPONSE IN NATURAL POPULATIONS

For several reasons, prediction in natural populations is more difficult than prediction under artificial selection in highly controlled environments. One obvious problem is associating parents and their offspring. Even if the mother is known, in species with multiple paternity the offspring must be assessed as to whether they are full or half sibs. Polymorphic molecular markers can be used to determine paternity (methods are reviewed by Lynch 1988c, Queller and Goodman 1989, Avise 1994, Weir 1996), but this can be quite resource consuming for a large study.

**Example 5.** An interesting example of applying artificial selection in a natural setting is Flux and Flux (1982), who examined response on clutch size in starlings (*Sturnus vulgaris*) in New Zealand. The adults nest almost exclusively in nest boxes, allowing for identification of the mothers and careful monitoring of their offspring. Increased clutch size was selected for by removing all eggs from clutches below a specified brood size (artificial truncation selection on clutch size). Lumping results from the entire study, the 516 clutches from the offspring of selected females had an average size of  $5.60 \pm 0.04$ , while the average size of 2050 clutches from offspring of unselected females was  $5.48 \pm 0.02$ , giving an estimated response of  $R = 0.12 \pm 0.04$ . The mean clutch size of selected female parents was 5.20 versus 5.48 for unselected (control) females, giving  $S_f = 0.72$ . Since the expected response in daughters is  $(h^2/2) S_f$ , the realized heritability is

$$\hat{h}_r^2 = 2 \cdot 0.12 / 0.72 = 0.33$$

This is in good agreement with the estimated heritability based on mother-daughter regressions of  $\hat{h}^2 = 0.34 \pm 0.08$ .

A serious complication in natural populations is that the investigator has very little (if any) control over environmental factors. In contrast to artificial selection experiments where individuals are usually randomized by the investigator across environments, genotypes in natural populations may nonrandomly assort over environments. For example, if resources are limited, individuals with a genetic disposition for larger body size may subsequently gather a larger share of resources than individuals with a genetic disposition for smaller size. This generates a correlation between environmental and genetic effects for body size.

A final complication is that the selection differential on a character must be inferred, rather than being under the control of the investigator. Recall from LW Chapter 8 (also see Chapters 14 and 16) that the observed amount of selection on character  $i$  depends on the amount of direct selection  $\beta_i$  on that character plus indirect effects due to direct selection  $\beta_j$  on all characters phenotypically correlated with the character of interest,

$$S_i = \sigma^2(z_i) \beta_i + \sum_{j \neq i} \sigma(z_i, z_j) \beta_j$$

Hence, a nonzero selection differential can result even if there is no direct selection on the character of interest ( $\beta_i = 0$ ). This again is in sharp contrast to artificial selection wherein the amount of direct selection on a character is largely controlled (and hence is known) by the investigator.

---

**Example 6.** In contrast to Example 5, not all studies show a good agreement between predicted and observed response. An apparent example of a reversed response in a natural population was seen by Larsson et al. (1998), who examined selection on body size in barnacle geese (*Branta leucopsis*). Analysis of more than 2,000 banded individuals showed that larger females produce larger clutches, had earlier hatching dates, and had more (and heavier) offspring compared with smaller females. Males showed no relationship between size and these fitness traits. Despite the fact that size is highly heritable in this population, and there appeared to be significant selection to increase body size, size has significantly *decreased* (by 0.5 and 0.6 SD in females and males, respectively) over a 13-year period in the colonies studied. Since individuals were banded, the authors could largely rule out immigration from other populations (perhaps under different selection regimes). Although selection acting on a character genetically correlated to body size could counterbalance any direct selection on body size (Chapter 17), no obvious candidates of such selection could be suggested by the authors.

---

### Observed Responses in Natural Bird Populations

Some of the best studies of the response to selection in natural populations come from birds (see Cooke and Buckley 1987 for reviews, also see Hailman 1986). In certain settings (such as isolated islands), the entire population can be banded and all nests located, allowing for an accurate measurement of individual fitness (Chapter 14) as well as mother-offspring associations.

---

**Example 7.** Some of the most classic studies of selection response in natural populations are the work on various species of Darwin's finches in the Galápagos Islands (Grant 1986). We consider two particular studies on Darwin's ground finch (*Geospiza fortis*) here.

Boag (1983) had reason to suspect that strong directional selection occurred on body size (Boag and Grant 1981). All birds on the small island of Isla Daphne Major were banded and followed. Midparents measured in 1978 were 0.44 standard deviations (SD) larger than the mean of adults measured in 1976. The adult offspring of the 1978 parents were 0.31 SD larger than the mean for 1976. Comparisons using the mean of a previous generation must be done with extreme care, due to the possibility of large between-generation environmental differences. Further complications concern selection on characters correlated with size and problems with overlapping generations. Assuming the complications introduced by these effects are negligible, the realized heritability estimate is  $\widehat{h}_r^2 = 0.31/0.44 = 0.70$ , consistent with the estimate from sib analysis ( $\widehat{h}^2 = 0.78 \pm 0.19$ ).

Grant and Grant (1995) report on two periods of selection for *G. fortis*. A drought in 1976–1997 resulted in strong selection for large birds with deep beaks. A change in the dominant food supply during a subsequent drought from 1984–1986 resulted in populations being selected in the opposite direction for beak traits. Six characters were followed during both droughts. As shown below, with three exceptions, the response in these traits was well predicted. Grant and Grant suggest that the main reasons for the discrepancies are environmental effects on growth and adult size and possible selection on offspring before they were measured.

Character	1976-1977		1984-1986	
	Predicted	Observed	Predicted	Observed
Weight	17.39 ± 0.22	17.52 ± 0.25	16.82 ± 0.13	15.48 ± 0.08*
Wing length	69.98 ± 0.39	69.65 ± 0.35	67.93 ± 0.17	67.21 ± 0.11***
Tarsus length	19.45 ± 0.09	19.32 ± 0.14	19.02 ± 0.04	19.02 ± 0.04
Bill length	11.14 ± 0.10	11.06 ± 0.11	10.86 ± 0.05	10.96 ± 0.03
Bill depth	9.83 ± 0.12	9.94 ± 0.09	9.51 ± 0.06	9.32 ± 0.03**
Bill width	8.96 ± 0.08	8.97 ± 0.08	8.77 ± 0.04	8.70 ± 0.03

where \* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$

Examples 5 and 7 are rather unusual in that they show excellent agreement between predicted and observed response. Many other studies in natural bird populations show apparent directional selection on a character coupled with significant heritability, but no response to selection (or worse, a reverses response, e.g., Example 6). These include several studies on clutch size (reviewed in Price and Liou 1989, Cooke et al. 1990), tarsus length (a body size measure) in the collared flycatcher *Ficedula albicollis* (Alatalo et al. 1989), and breeding date (Price et al. 1988). A variety of explanations have been proposed to account for this lack of response. For example, Price et al. (1988) suggest that the apparent directional selection on breeding date is just that, apparent. They argue that an individual's nutritional level, which they assume to be entirely environmental, directly influences fitness and also allows individuals to breed earlier. This results in an association between fitness and breeding date and hence a selection differential on breeding date as from the Price-Robertson identity  $S = \sigma(w, \text{breeding date})$ . This association is, however, entirely due to both sharing a common environmental factor so that there is no response to decrease breeding date. Although there sufficient heritable variation for a response, the additive genetic value for breeding date is not under directional selection and hence no response is observed. Similar explanations (again based on environmental factors influencing both an individual's fitness and the character of interest) have been offered to account for lack of response in clutch size (Price and Liou 1989) and, as the next example shows, tarsus length.

**Example 8.** Alatalo et al. (1989) examined tarsus length in the collared flycatcher (*Ficedula albicollis*) on a population using nest boxes residing in the southern part of the island of Gotland in the Baltic Sea. Measurements of lifetime fitness on this isolated population are possible since most surviving offspring return to breed in the area they were reared as offspring. In addition to tarsus length, fledging weight was also measured and a Pearson-Lande-Arnold regression (Chapter 16, LW Chapter 8) performed to compute the amounts of direct selection (the estimated selection gradients  $\beta$ ) on both characters.

Year	Observed $S$	Estimated selection gradients, $\beta$	
	on tarsus length	tarsus length	fledging weight
1981	0.19**	0.01	0.25*
1983	0.08	-0.01	0.21*
1984	0.20**	0.12	0.33***
1985	0.02	-0.06	0.27***
pooled	0.12**	0.03	0.27***

where \* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$

Although there is a significant selection differential on tarsus length in two of the years (and in the pooled data), there is no significant direct selection on tarsus length itself ( $\beta$  not significantly different from zero). Rather, direct selection is on fledging weight. While there is a significant phenotypic correlation between tarsus length and fledging weight ( $r = 0.32, p < 0.001$ ), it appears to be entirely due to within-individual correlations of environmental effects as there is no correlation between offspring weight and parental tarsus length ( $r = -0.01, p > 0.1$ ). Hence, the observed selection on tarsus length is a consequence of selection on fledging weight, which has no genetic correlation with tarsus length and hence no response in tarsus length is expected (Chapter 17).

## CONTROL POPULATIONS AND EXPERIMENTAL DESIGNS

One complication with estimating realized heritabilities is distinguishing between genetic versus environmental trends. For example, Newman et al. (1973) found that 60 percent of the increase in yearling weight in a selected line of Shorthorn cattle was due to environmental, rather than genetic, improvement. Using a large unselected **control population** reared under the same environmental conditions as the selected line(s) allows for correction of between-generation environmental change. Hill reviews both different approaches (1972b) and some of the features seen in control populations from a number of different species (1972c). As we show shortly, the use of control populations is not a fool-proof approach for removing

environmental trends, as the control and selected lines can develop different genotype-environment interactions. Likewise, extensive genetic drift in the control population can also result in biases. When full pedigree data are available, the powerful mixed-model analysis (discussed below) can also be used to remove genetic and environmental trends, even in the absence of control populations (although they are certainly preferred).

### Basic Theory of Control Populations

Assuming no genotype-environment interaction, the true mean of a line in generation  $t$  can be decomposed as  $\mu + g_t + d_t$ , where  $\mu$  is the base population mean,  $g_t$  the change in mean breeding value due to selection and drift, and  $d_t$  the change in mean due to environmental change. Under this model, observed means  $\bar{z}$  for a selected (s) and control (c) population reared in a common environment can be decomposed as

$$\bar{z}_{s,t} = \mu + g_{s,t} + d_t + e_{s,t} \quad (6.21a)$$

$$\bar{z}_{c,t} = \mu + g_{c,t} + d_t + e_{c,t} \quad (6.21b)$$

where  $e_t$  is the error in estimating the mean breeding value ( $\mu + g_t$ ) from the observed mean corrected for the change in the environment ( $\bar{z}_t - d_t$ ) and has expected value 0.

Assuming that the breeders' equation holds, the expected total response at generation  $t$  is  $E[R_C(t)] = E(g_{s,t}) = h^2 S_C(t)$ . Under drift alone, there is no expected directional change in the mean breeding value of the control population,  $E(g_{c,t}) = 0$ . Assuming no genotype-environment interactions (i.e.,  $d_{s,t} = d_{c,t}$ ), the contrast between the selected and control populations has expected value

$$E(\bar{z}_{s,t} - \bar{z}_{c,t}) = E(g_{s,t}) - E(g_{c,0}) = h^2 S_C(t) \quad (6.22a)$$

If the control population is small, then  $g_{c,t}$  can drift significantly away from zero, resulting in an over (or under) estimation of the true selection response. Hence, if the goal is simply to remove an environmental trend, control populations should be kept as large as possible so that  $g_{c,t}$  is close to zero. However, control populations are also used to attempt to correct for any effects of inbreeding depression on the trait (Chapter 6). In these cases, significant drift has likely occurred, and the use of a control population introduces additional uncertainty in the estimate of the response (although this may be more than compensated for the accounting for inbreeding depression and/or environmental trends).

If genotype-environmental interactions are present, then

$$E(\bar{z}_{s,t} - \bar{z}_{c,t}) = h^2 S_C(t) + (d_{s,t} - d_{c,t}) \quad (6.22b)$$

resulting in  $\bar{z}_{s,t} - \bar{z}_{c,t}$  being a potentially biased estimator of  $h^2 S_C(t)$ . If the environmental trends are positively correlated between populations, then the use

of a control will still improve the estimate of the genetic trend. However, if the environmental values are negatively correlated between populations, the use of a control can lead to a more inaccurate estimate than simply using a selected population without a control. Mixed-model analysis may be able to provide some insights (as they estimate the  $d_{x,t}S$ ), but they require extensive information (full pedigrees) and that the model assumptions hold. Alternatively, Muir (1986a,b) has suggested that an analysis of covariance approach (along the lines of LW Equation 10.12) be used when extensive  $G \times E$  is expected.

If a control population is used, then the responses and differentials are estimated by

$$R_t = (\bar{z}_{s,t} - \bar{z}_{c,t}) - (\bar{z}_{s,t-1} - \bar{z}_{c,t-1}) \quad (6.23a)$$

$$R_C(t) = \bar{z}_{s,t} - \bar{z}_{c,t} \quad (6.23b)$$

$$S_t = \bar{z}_{s,t-1}^* - \bar{z}_{s,t-1} \quad (6.23c)$$

where  $\bar{z}^*$  denotes the mean of the selected individuals. No correction is necessary for the selection differential  $S_t$ , as we have assumed the environment stays constant within a generation. If selection differs between sexes,

$$S_t = \frac{S_t(m) + S_t(f)}{2}$$

where  $S_t(m)$  and  $S_t(f)$  are, respectively, the observed differentials on males and females.

### Divergent Selection Designs

A related approach to comparing a selected and a control line is the **divergent** (or **bidirectional**) selection design, wherein one compares lines selected in opposite directions (typically denoted by the up and down, or high and low, lines). Again assuming no significant genotype  $\times$  environment interactions between lines, the basic statistical model for this design is

$$\bar{z}_{u,t} = \mu + g_{u,t} + d_t + e_{u,t} \quad (6.24a)$$

$$\bar{z}_{d,t} = \mu + g_{d,t} + d_t + e_{d,t} \quad (6.24b)$$

where  $u$  and  $d$  refer to the upwardly- and downwardly-selected lines. With this design, the responses and differentials are estimated by

$$R_t = (\bar{z}_{u,t} - \bar{z}_{u,t-1}) - (\bar{z}_{d,t} - \bar{z}_{d,t-1}) \quad (6.25a)$$

$$R_C(t) = \bar{z}_{u,t} - \bar{z}_{d,t} \quad (6.25b)$$

$$S_t = (\bar{z}_{u,t-1}^* - \bar{z}_{u,t-1}) - (\bar{z}_{d,t-1}^* - \bar{z}_{d,t-1}) \quad (6.25c)$$

Again, the expected response (using Equations 6.25a and 6.25c for the response and selection differential) is just  $R = h^2S$ . More generally,  $S_x$  denotes the

selection differential in line  $x$ , the the total (cumulative) divergence between lines is just

$$R_C(t) = h^2 \sum_{i=1}^t (S_{u,i} - S_{d,i}) = h^2 (S_{C,u} - S_{C,d}) \quad (6.25d)$$

In addition to previous concerns about genotype-environment interactions, asymmetric response to selection (Chapter 6) also complicates the interpretation of results with divergent selection and can result in a biased estimate of the realized heritability.

### Variance in Response

Chapter 6 developed the pure-drift approximation for the variance (Equation 6.6) and within-line covariance across generations (Equation 6.9) for the design of a single selected lines. Here we present the corresponding expressions for the selection + control and divergence selection designs (assuming no genotype-environment interactions between lines). For unidirectional selection plus a control population,

$$\begin{aligned} R_C(t) &= \bar{z}_{s,t} - \bar{z}_{c,t} = (\mu + g_{s,t} + d_t + e_{s,t}) - (\mu + g_{c,t} + d_t + e_{c,t}) \\ &= g_{s,t} - g_{c,t} + e_{s,t} - e_{c,t} \end{aligned} \quad (6.26)$$

Similarly, for divergent selection,

$$R_C(t) = \bar{z}_{su,t} - \bar{z}_{sd,t} = g_{su,t} - g_{sd,t} + e_{su,t} - e_{sd,t} \quad (6.27)$$

Since each term in Equations 6.26 and 6.27 is independent, applying Equations 6.4 and 6.5b gives the pure-drift approximations for the variance in response as

$$\begin{aligned} \sigma^2 [R_C(t)] &= (2f_t + B_0) 2f_t h^2 \sigma_z^2 + B_t \sigma_z^2 \\ &\simeq (tA + B_0) h^2 \sigma_z^2 + B_t \sigma_z^2 \end{aligned} \quad (6.28a)$$

and the covariance between generations in the same line

$$\begin{aligned} \sigma [R_C(t), R_C(t')] &= (2f_t + B_0) h^2 \sigma_z^2 \\ &\simeq (tA + B_0) \sigma_z^2 h^2 \quad \text{for } t < t' \end{aligned} \quad (6.28b)$$

where the coefficients  $A$  and  $B_t$  are given in Table 6.8. Recall (Equation 6.3) that for unidirectional selection without a control, the variance in response has an additional term,  $\sigma_d^2$  accounting for the between-generation environmental variation

**Table 6.8.** Coefficients for the pure-drift variances and covariances in response (Equations 6.28a and 6.28b).  $M_x$  individuals are sampled, of which  $N_x$  are allowed to reproduce. The subscripts  $s$  and  $c$  refer to the selected and control populations,  $u$  and  $d$  to the up- and down-selected lines, respectively.

---

**Selection in a single direction without a control line.** Equation 6.28a has an extra term,  $\sigma_d^2$ , accounting for the between-generation' variation in environmental effects.

$$f_t = f_{s,t}, \quad A = \frac{1}{N_s}, \quad B_t = \frac{1}{M_{s,t}} \quad \text{for } t \geq 0$$

**Selection in a single direction with a control line**

$$f_t = f_{s,t} + f_{c,t}, \quad A = \frac{1}{N_s} + \frac{1}{N_c}, \quad B_t = \frac{1}{M_{s,t}} + \frac{1}{M_{c,t}} \quad \text{for } t \geq 0$$

**Divergent Selection Without a Control Line**

$$f_t = f_{u,t} + f_{d,t}, \quad A = \frac{1}{N_u} + \frac{1}{N_d}, \quad B_t = \frac{1}{M_{u,t}} + \frac{1}{M_{d,t}} \quad \text{for } t \geq 0$$

### Control Populations and Variance in Response

When does using a control population reduce the variance in response? Assuming  $M = M_s = M_c$  and  $N = N_s = N_c$ , subtracting the expected variance in response using a unidirectionally-selected population adjusted using a control from the response estimated without a control gives

$$\left( \frac{t}{N} + \frac{1}{M_0} \right) h^2 \sigma_z^2 + \frac{1}{M} \sigma_z^2 - \sigma_d^2 \quad (6.29a)$$

Assuming the between-line drift variance dominates (terms involving  $M$  are ignored), the condition for the variance in response with a control to be larger than the response without one is approximately

$$\frac{t\sigma_z^2 h^2}{N} > \sigma_d^2 \quad (6.29b)$$

Hence, regardless of the value of  $\sigma_d^2$ , if sufficient generations are used, the optimal design (in terms of giving the smallest expected variance in response) is not to use a control. However, this approach runs the risk of an undetected directional environmental trend compromising the estimated heritability. Further, as the number of generations becomes large, the key assumption of short-term response (essentially no changes in the genetic and phenotypic variances) becomes untenable.

## OPTIMAL EXPERIMENTAL DESIGNS

As Equation 6.29b illustrates, when to use different designs is not entirely clear-cut. What in general can we say? The coefficient of variation in response,

$$\text{CV} [R_C(t)] = \frac{\sigma^2 [R_C(t)]}{E[R_C(t)]}$$

is especially useful in comparing efficiencies of different designs, as it is independent of  $\sigma_z^2$  and further provides an appropriate measure of comparing efficiencies when the expected response differs between designs. Table 6.9 gives expressions for the CV under some simplifying assumptions. Note that the coefficient of variation is a function of  $tN$ , the total number of adults selected during the course of the experiment (provided drift variance dominates error variance). A short experiment with many selected adults per generation thus gives the same expected CV as a long experiment with few adults per generation (provided the total numbers are the same). However, if the error variance is nontrivial relative to the drift variance (as would be expected if  $h^2$  small), increasing the duration of the experiment results in some improvement in precision (Hill 1980).

**Table 6.9.** Coefficients of variation for various designs, assuming the pure-drift approximation and further that  $\sigma^2 [R_C(t)] \simeq t A h^2 \sigma_z^2$ . This assumes that the selection experiment is sufficiently long that the between-line drift dominates (i.e.,  $t A \gg B_0$  and  $t A h^2 \gg B_t$ ). For all designs, we assume that the absolute selection intensity on all selected lines is  $\bar{i}$ .

---

**Selection in a single direction with a control line**

$$E [R_C(t)] = t h^2 \bar{i} \sigma_z, \quad \text{CV} [R_C(t)] \simeq \frac{1}{h\bar{i}} \sqrt{\frac{2}{Nt}}$$

**Selection in a Single Direction Without a Control Line** (Assuming  $\sigma_d^2 = 0$ )

$$E [R_C(t)] = t h^2 \bar{i} \sigma_z, \quad \text{CV} [R_C(t)] \simeq \frac{1}{h\bar{i}} \sqrt{\frac{1}{Nt}}$$

**Divergent Selection Without a Control Line**

$$E [R_C(t)] = 2t h^2 \bar{i} \sigma_z, \quad \text{CV} [R_C(t)] \simeq \frac{1}{h\bar{i}} \sqrt{\frac{1}{2Nt}}$$


---

As an example of using CV, consider unidirectional selection without a control population versus divergent selection. Which is more efficient if the same

total number of adults are selected (e.g.,  $N$  under unidirectional,  $N_d = N_u = N/2$  under divergent selection)? If there is no between-generation environmental variance both designs are equally efficient, while divergent selection is more efficient if  $\sigma_d^2 > 0$ .

---

**Example 9.** Suppose we plan to select the upper 5% of a population for a normally distributed character with  $h^2 = 0.25$ . What value of  $Nt$  is needed for the expected CV of response to be no greater than 0.01 if no control population is used? From Example 3 of Chapter 4,  $E(\bar{z}) = 2.06$  if the population is large, and slightly less in small populations (for simplicity assume the large population value). Further assuming that drift variance dominates error variances (including  $\sigma_d^2$ ), applying the expressions from Table 6.9 gives

$$0.01 = \frac{1}{0.5 \times 2.06} \times \sqrt{\frac{1}{Nt}}$$

Solving, gives  $Nt \simeq 9426$ . Hence, during the entire course of the experiment using a total of at least 9426 selected parents gives an approximate expected CV less than 1%. If the desired CV is 0.05 or 0.10,  $Nt \simeq 377$  and  $Nt \simeq 94$ , respectively.

---

### Nicholas' Criterion

An alternative criterion for choosing  $Nt$  was suggested by Nicholas (1980). Often the investigator is interested in insuring that at least a certain response will occur with a preset probability. To a reasonable approximation, the expected mean value in any given replicate line after  $t$  generations of selection is normally distributed, with mean  $E[R_C(t)]$  and variance  $\sigma^2[R_C(t)]$ . Consider the probability that the observed response is at least  $\beta$  of the expected response,

$$\begin{aligned} \Pr(R_C(t) > \beta E[R_C(t)]) &= \Pr\left(\frac{[R_C(t) - E[R_C(t)]]}{\sigma[R_C(t)]} > \frac{(\beta - 1)E[R_C(t)]}{\sigma[R_C(t)]}\right) \\ &= \Pr\left(U > \frac{\beta - 1}{\text{CV}[R_C(t)]}\right) \end{aligned} \quad (6.30)$$

where  $U$  is a unit normal random variable. Note that the probability that the observed response exceeds the expected response ( $\beta = 1$ ) is one half (as  $\Pr[U > 0] = 1/2$ ).

---

**Example 10.** Again suppose that  $\bar{i} = 2.06$ ,  $h^2 = 0.25$ , and the design is unidirectional selection without a control population. What value of  $Nt$  is required in order for a 95% probability that the observed response is at least 90% of its expected response? Here,  $\beta = 0.9$  and (from normal tables)  $\Pr[U > -1.65] = 0.95$ . Hence,

$$\frac{\beta - 1}{\text{CV}[R_C(t)]} = \frac{-0.1}{\text{CV}[R_C(t)]} = -1.65$$

Rearranging gives

$$\text{CV}[R_C(t)] = \frac{1}{0.5 \times 2.06} \sqrt{\frac{1}{Nt}} = \frac{0.1}{1.65}$$

implying that  $Nt \simeq 257$ .

### Replicate Lines

There is no loss of efficiency when replicate lines are used (Hill 1980). To see this, let  $\bar{z}_{i,t}$  for  $1 \leq i \leq r$  be the sample mean for replicate population  $i$  at time  $t$ . The overall mean is

$$\bar{z}_t = \frac{1}{r} \sum_{i=1}^r \bar{z}_{i,t} \quad (6.31a)$$

Taking variances and assuming each line is independent,

$$\sigma_{\bar{z}}^2(t) = \frac{1}{r^2} \sum_{i=1}^r \sigma_{\bar{z}_i}^2(t) = \frac{1}{r} \sigma_{\bar{z}_1}^2(t) \quad (6.31b)$$

If the number sampled and number used as parents within a replicate are  $M^*$  and  $N^*$ , with  $N^* = N/r$  and  $M^* = M/r$ , then it is easily seen from Table 6.9 that the variance of a replicate line is just  $r$  times the variance of a population with  $N$  and  $M$ . Hence, variance in the sample mean from  $r$  replicate lines with  $N^*$  and  $M^*$  is the same as the variance with a single line with  $N$  and  $M$ , provided that the number of individuals within each replicate line is sufficiently large to avoid significant inbreeding. Richardson et al. (1968), Irgang et al. (1985), and Muir (1986a,b) develop regression approaches correcting for between-generation environmental changes when replicate lines are used.