Our examination of selection of quantitative characters, and its consequences, comprises the remainder of this book. We start by discussing the simplest models — changes in the mean of a single character following a single generation of selection. As the reader will see over the next several chapters, even these simple models are not without complications. Over subsequent chapters we build on these models, making them more complex and biologically realistic.

Selection changes the distribution of both phenotypic and genotypic values from one generation to the next. While it is convenient to assume that only the means of these distributions change, this is generally not the case. Rigorous prediction of the expected change in the mean over several generations requires predictions of how higher moments (such as the variance and skew) of the genotypic distribution change as well (Chapters 9, 10). Unfortunately, this requires considerable knowledge about the underlying distributions of allelic effects and frequencies, both of which are almost always unknown. Despite this gloomy initial assessment, the change in mean over a few generations of selection can often be accurately projected using estimates of genetic variances from an unselected base population. This chapter considers the basis for much of this success: models that assume a linear parent-offspring regression and short-term stability of the genetic variance (i.e., over short time scales, selection-induced changes in variance are assumed to be negligible). Over longer time scales, changes in variances can be substantial and are extremely difficult to predict. Thus, while the results presented here hold for a single generation, they are expected to become increasingly less accurate as the number of generations of selection increases.

Besides the constancy of variances and linearity of parent-offspring regression, the third major assumption required here is that there is no selection acting on characters correlated to our character of interest. Breeders have the luxury of
artificial selection and hence fairly precise control over the nature of selection, while evolutionary biologists face the additional problem of estimating the form of selection acting on the character(s) under consideration. However, even in the most controlled breeding designs, natural selection can also be acting on the character of interest (either directly, or through correlations on other characters under selection). Accounting for selection on multiple characters is discussed in Chapters 17-20.

There is a huge literature on different selection schemes that exploit special features of specific organisms (such as artificial insemination or complex crosses of selfed plants). Our focus in this chapter is individual (or mass) selection, wherein individuals are chosen solely on their phenotypic value alone (i.e., information from relatives, other characters, etc is ignored). Family selection, where individuals are chosen based on either their family mean or their relative ranking within a family is discussed in Chapter 6. Using such information from relatives can improve our ability to accurately predict breeding values and hence increase response relative to individual selection. The general theory for selection using information from relatives relies on index selection and BLUP (LW Chapter 26), both of which are discussed in Chapter 20. Additional complications (such as threshold traits, overlapping generations, inbreeding, and group selection) are examined in Chapter 6. The reader is also referred to Turner and Young (1969), Pirchner (1983), Ollivier (1988), Weller (1994), and Cameron (1997) for specialized applications in animal breeding and to Namkoong (1979), Hallauer and Miranda (1981), Wricke and Weber (1986), and Mayo (1987) for specialized applications in plant breeding.

**SINGLE-GENERATION RESPONSE: THE BREEDERS’ EQUATION**

The simplest selection model is the classic breeders’ equation

\[ R = h^2 S \]  

which relates response to heritability times the selection differential is perhaps the most well know expression in quantitative genetics. The actual origin of the breeders’ equation is somewhat unclear, but it was clearly suggested (in multivariate form) in the early writings of Pearson and popularized by Lush (1937). The simplicity of this equation is compelling, relating the change in mean across a generation (the response \( R \)) to the product of the within generation change in the mean (the **directional selection differential** \( S \)) with a measure of how character value is passed across generations (the slope \( h^2 \) of the midparent-offspring regression). As shown in Chapter 1, the simplest derivation of the breeders equation follows by assuming the parent-offspring regression is linear. Even when the parent-offspring regression is not assumed to be linear, the breeders’ equation still approximately holds (Equations 9.31b and 9.35b).
While the selection different $S$ is a convenient and simple measure of the selection, it does not really tell us much about the strength of selection. Consider selection acting on the same character in two different populations. In one, the largest five percent of measured individuals are allowed to reproduce while in the second the largest 25 percent reproduce. Clearly selection is more intense in the first population. However, if the characters are normally distributed, Equation 4.14 gives the selection differentials for these two populations as $S_1 = 2.06 \sigma_1$ and $S_2 = 1.27 \sigma_2$, where $\sigma_k^2$ is the character variance in population $k$. Provided the second population is sufficiently more variable than the first, it can have the larger selection differential.

A better measure is to use the standardized directional selection differential (or selection intensity)

$$\tau = \frac{S}{\sigma}$$  \hspace{1cm} (4.2)

The selection intensity accounts for differences in the phenotypic variances, much akin to the correlation being a better measure of the strength of association than is the covariance (LW Chapter 3). Substituting $\tau \sigma$ for $S$ gives the selection intensity version of the breeders’ equation,

$$R = h^2 \tau \sigma_z = \tau h \sigma_A$$  \hspace{1cm} (4.3)

Under appropriate linearity assumptions, the breeders’ equation generalizes to much more general settings, predicting the mean of character $y$ measured in one group when selection occurs on character $x$ measured in another group. Assuming the regression of $y$ on $x$ is linear, standard regression theory (LW Equation 3.14) gives

$$E[y - \mu_y | x] = \frac{\sigma_{x,y}}{\sigma_x^2} (x - \mu_x)$$

giving the expected change in $y$ from selection on $x$ as

$$R_y = \mu_y^* - \mu_y = \frac{\sigma_{x,y}}{\sigma_x^2} (\mu_x^* - \mu_x) = \frac{\sigma_{x,y}}{\sigma_x^2} S_x$$  \hspace{1cm} (4.4a)

where $\mu^*$ denotes the mean following selection. The selection intensity version is

$$R_y = \frac{\sigma_{x,y}}{\sigma_x} \tau_x = \tau_x \sigma_y \rho(x, y)$$  \hspace{1cm} (4.4b)

where $\rho(x, y)$ is the correlation between $x$ and $y$. Expressing this in terms of the relative response, the change in $y$ in phenotypic standard deviations of $y$, gives

$$\frac{R_y}{\sigma_y} = \tau_x \rho(x, y)$$  \hspace{1cm} (4.4c)

Equation 4.4c allows one to compare the relative efficiencies of different selection schemes. Fixing the selection intensity $\tau$, the maximal expected response in $y$
occurs when we select on the measure \( x \) that has the largest correlation with \( y \). The relative responses of two different schemes (\( x_1 \) versus \( x_2 \)) is given by \( \rho(x_1, y)/\rho(x_2, y) \).

Suppose \( y \) is the breeding value \( A \) (LW Chapter 4) of an individual so that \( R_y \) is the change in breeding value from selection on \( x \), giving \( R_y = \tau \sigma_A \rho(x, A) \). If \( x \) is the phenotypic value \( z \) of the character whose breeding value is of interest, the correlation between \( x \) and \( y \) here is the correlation between an individual’s phenotype \( z \) and its breeding value \( A \),

\[
\rho(z, A) = \sigma(z, A)/\sigma_A = \sigma_A/\sigma_z = h
\]

recovering Equation 4.3. In evaluating other breeding schemes \( h \) is replaced by the appropriate correlation between the breeding value and the measure \( x \) being selected on. This correlation is referred to as the **accuracy** in predicting breeding value from knowledge of \( x \). For example, we could take \( x \) to be the family mean of the character for the individual being considered, in which case the correlation between an individual’s breeding value and its family mean determines the response to selection under this scheme (Chapter 6). Since \( \tau \) and \( \sigma_A \) are constant, the largest response occurs by taking the measure \( x \) that gives the largest correlation with breeding value. This idea forms the foundation of index selection (Chapters 19, 20) wherein \( x \) is a linear combination of several characters \( x = \sum a_i z_i \).

---

**Example 1.** As an example of applying Equation 4.4a, consider selection on clones or other pure lines. With clones, parents pass on their entire genome to their offspring. The phenotypic value \( z \) of an offspring from a parent with genotypic value \( G \) can be written as \( z = G + E \), so that parent-offspring covariance (in the absence of any genotype \( \times \) environment covariance and/or interactions) equals the total genetic variance \( \sigma_G^2 \). The resulting parent-offspring regression is linear with slope \( b_{op} = \sigma_G^2/\sigma_z^2 = H^2 \), the **broad-sense heritability** (LW Chapter 19), giving

\[
R = H^2 S
\]

Since \( H^2 \geq h^2 \), the single-generation response to selection is larger for clones than for a sexual population with the same variance components. When selection continues for several generations, using clones is expected to be less efficient than using a sexual population. Selection among clones very rapidly removes genetic variation from the population without any mechanism (other than mutation) to generate new variation. With selection among sexual individuals, recombination generates an almost endless supply of new variation if a large number of segregating loci underly the trait. For this reason, selection schemes involving clones often randomly mate lines every few generations to allow for recombination, generating variation required for continued response.
Prelude to the Multivariate Breeders’ Equation

Three other alternate parameterizations of the breeders’ equation are popular in the evolutionary literature, and we introduce them here to motivate their multivariate versions (discussed in Chapters 15, 17). Expressing the heritability in terms of additive genetic and phenotypic variance,

$$ R = \sigma_A^2 \sigma_z^{-2} S $$

(4.5a)

While this decomposition seems rather trivial, it suggests that multivariate version of the breeders’ equation (under appropriate linearity assumptions) is given by

$$ R = \mathbf{GP}^{-1} \mathbf{s} $$

(4.5b)

where $\mathbf{R}$ is the vector of responses (changes in the mean of each character), $\mathbf{s}$ the vector of selection differentials for each character, and $\mathbf{G}$ and $\mathbf{P}$ are the additive genetic and phenotypic covariance matrices (LW Chapter 21). We will examine this equation, and its consequences, in Chapters 15 and 17.

A second parametrization follows from the Robertson-Price identity, $S = \sigma(z, w)$, which states that the selection differential equals the covariance between relative fitness $w$ (fitnesses normalized so that mean fitness is one) and character value $z$ (LW Equation 3.21). Since the slope of the least-squares linear regression of relative fitness on phenotypic value is $\beta = \sigma(z, w)/\sigma_z^2$, it follows that $S = \sigma_z^2 \beta$. The regression slope $\beta$ is called the directional selection gradient, for reasons that will become apparent shortly. Substituting $S = \sigma_z^2 \beta$ into Equation 4.1 gives

$$ R = \sigma_A^2 \beta $$

(4.6a)

The multiple linear regression of relative fitness on a vector $\mathbf{z}$ of characters has slope $\beta = \mathbf{P}^{-1} \mathbf{s}$ (LW Equation 8.10c). Substituting into Equation 4.5b gives

$$ R = \mathbf{G} \beta $$

(4.6b)

We examine Equation 4.6b, and its consequences, in detail in Chapter 17.

These alternative expressions point out that we can distinguish between phenotypic selection, the change in a phenotypic distribution within a generation (measured by $S$ or $\beta$), and the evolutionary response to selection, the transmission of these within-generation changes to the next generation ($R$). Lande and Arnold (1983) and Arnold and Wade (1984a,b), following Fisher (1930, 1958) and Haldane (1954), have stressed the utility of this approach. Attempts to measure selection by comparing phenotypic distributions across generations are confounded by inheritance, as $R$ depends on $\beta$ through $\sigma_A^2$ (or $\mathbf{G}$ in the multivariate case). For example, if a character has no additive genetic variation ($\sigma_A^2 = 0$), there is no permanent change in the mean across generations regardless of the selection intensity, and a cross-generation comparison would conclude that selection was
not operating. Chapters 14 and 16 examine in detail methods for estimating the nature of phenotypic selection in natural populations.

Finally, the breeders’ equation can be expressed in a form analogous to Wright’s formula (Equation 9.4) from population genetics. As is discussed in Chapter 9, Wright’s formula states that the change in an allele frequency \( p_i \) at a locus is proportional to \( \partial W / \partial p_i \), so that allele frequencies change so as to increase population fitness \( W \). Lande (1976) showed there is an analogous expression relating the change in mean of a quantitative trait to the change in mean population fitness. If \( w(z) \) is the expected fitness of an individual with character value \( z \), phenotypes are normally distributed, and fitness is frequency-independent (individual fitnesses are not a function of the mean of the character), then Lande showed that the directional selection gradient satisfies \( \beta = \partial \ln w / \partial \mu \). Hence we can express the breeders equation as

\[
R = \sigma_A^2 \left( \frac{\partial \ln w}{\partial \mu} \right) \tag{4.7a}
\]

The multivariate version is the gradient of mean fitness (Chapter 15) with respect to the vector of character means,

\[
\beta = \frac{\partial \ln w}{\partial \mu}
\]

giving

\[
R = G \frac{\partial \ln w}{\partial \mu} \tag{4.7b}
\]

The gradient terms represents the changes in character means that produce the maximal change in population fitness, while the actual response involves the product of this vector with the genetic covariance matrix \( G \). The resulting response vector is generally not parallel to \( \beta \) and hence the genetic covariance structure causes the character means to change in a direction that does not result in the optimal change in population fitness. We examine these issues in detail in Chapters 15 and 17.

The Importance of Linearity

A variety of factors can result in a nonlinear parent-offspring regression (Chapter 10, LW Chapter 17). In such cases, the mean of the selected parents (and hence the selection differential \( S \)) is not sufficient to predict the offspring mean. As Figure 4.1 shows, two selected parental populations with the same mean, but different variances, can have different expected responses when the parent-offspring regression is nonlinear. Even if phenotypes are normally distributed and the character is completely determined by additive loci (no dominance or epistasis), if the underlying distribution of genotypic values shows skew, selection on the variance also results in a change in the mean (see Equation 9.35b). A sufficient condition
for linearity is that the joint distribution of breeding and phenotypic values is bivariate normal (Chapter 15, LW Chapter 8).

![Figure 4.1](image)

**Figure 4.1.** The importance of linearity in the parent-offspring regression. If this regression is nonlinear, different subsets of the population with the same mean can have different offspring means. Suppose equal number of parents with values $z_a$ and $z_b$ are chosen. Denoting the expected value of offspring from parents with value $z$ by $E[z_o | z]$, the offspring mean is given by $(E[z_o | z_a] + E[z_o | z_n]) / 2$. Conversely, choosing parents all with value $(z_a + z_b)/2$, gives the same parental mean and hence the same $S$, but the expected offspring mean is now $E[z_o | (z_a + z_b)/2]$, which, as shown, can deviate considerably from $(E[z_o | z_a] + E[z_o | z_n]) / 2$.

Selection generally causes the distribution of genotypic values to depart from normality, creating at least slight departures from linear parent-offspring regressions. Response under such non-normal distributions can be very complicated, depending on underlying genetic factors that do not easily translate into standard (and measurable) variance components (see Chapters 9 and 10).

**Correcting for Fertility Differences: Effective Selection Differentials**

In artificial selection experiments, $S$ is usually estimated as the difference between the mean of the selected adults and the sample mean of the population before selection. Selection need not stop at this stage. For example, strong artificial selection to increase a character might be countered by natural selection due to a decrease in the fertility of individuals with extreme character values. Biases introduced by such differential fertility can be removed by randomly choosing the same number of offspring from each selected parent, ensuring equal fertility.

Alternatively, biases introduced by differential fertility can be removed by using **effective selection differentials, $S_e$**, $S_e = \frac{1}{n_p} \sum_{i=1}^{n_p} \left( \frac{n_i}{n} \right) \left( z_i - \mu_z \right)$ (4.8)

where $z_i$ and $n_i$ are the phenotypic value and total number of offspring of the $i$th
parent, \( n_p \), the number of parents selected to reproduce, \( \pi \) the average number of offspring for selected parents, and \( \mu_z \) is the mean before selection. If all selected parents have the same number of offspring (\( n_i = \pi \) for all \( i \)), then \( S_e \) reduces to \( S \). However, if there is variation in \( n_i \) among selected parents, \( S_e \) can be considerably different from \( S \). This corrected differential is occasionally referred to as the

**realized selection differential.**

The derivation of Equation 4.8 follows directly from applying the Robertson–Price identity (LW Equation 3.21). If a total of \( n \) parents are examined, \( n_p \) of which are selected as parents, then

\[
S = \sigma(z, w) = E[w|z] - E[z]E[w] = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{W_i}{\bar{W}} \right) z_i - \mu_z \frac{1}{n}
\]

where the fitness of individual \( i \) is \( W_i = n_i \) (with \( n_i = 0 \) for individuals not chosen as parents). The mean fitness becomes

\[
\bar{W} = \frac{1}{n} \sum_{i=1}^{n} n_i = \frac{\pi n_p}{n}, \quad \text{where} \quad \pi = \sum_{i=1}^{n} n_i / n_p
\]

so that \( \pi \) is the mean number of offspring left by adults selected to reproduce. Hence

\[
\frac{W_i}{\bar{W}} = \frac{n_i n}{\pi n_p} \quad \text{giving} \quad \sigma(z, w) = \sum_{i=1}^{n} z_i \frac{n_i}{\pi n_p} - \mu_z
\]

Rearranging recovers Equation 4.8. Chapter 14 examines individual fitness and fitness calculations such as this in detail.

**MORE GENERAL PARENT-OFFSPRING REGRESSIONS**

The regression coefficients for parent-offspring regressions can vary with the sex of both parents and offspring. In such cases, the breeders’ equation can be extended by incorporating these sex-dependent regression coefficients. To see how this is done, denote the phenotypic values of the father and mother by \( z_{fa} \) and \( z_{mo} \) and an offspring by \( z_o \) (if its sex is unimportant) or by \( z_{so} \) and \( z_{da} \) for sons and daughters (respectively) when sex is important. Let \( E(z_o | z_{fa}, z_{mo}) \) be the expected phenotypic value of an offspring whose parents have phenotypic values \( z_{mo} \) and \( z_{fa} \). The importance of this conditional expectation (the **biparental regression**) is that the expected character value in the next generation (assuming no fertility differences) is the average of \( E(z_o | z_{fa}, z_{mo}) \) over all selected parents. Taking expectations is particularly straightforward when the biparental regression is linear, i.e.,

\[
E(z_o | z_{fa}, z_{mo}) = \mu_o + b_{o,fa} (z_{fa} - \mu_{fa}) + b_{o,mo} (z_{mo} - \mu_{mo}) \quad (4.9)
\]
where $\mu_{fa}$ and $\mu_{mo}$ are the mean character values of males and females before selection and $\mu_o$ the mean for the offspring sex being considered. Denoting the expectation taken over all selected parents by $E_s$, the expected offspring mean after selection is

$$E_s\left[E(z_o | z_{fa}, z_{mo})\right] = \mu_o + b_{o,fa} S_{fa} + b_{o,mo} S_{mo} \quad (4.10)$$

where $S_{fa}$ and $S_{mo}$ are the directional selection differentials on fathers and mothers (respectively). With fertility differences, expectations are computed by appropriately weighting each $(z_{fa}, z_{mo})$ pair (see Equation 4.8).

Equations 4.9 and 4.10 allow for the possibility of differences between sexes in regression coefficients, in which case separate equations for sons and daughters are required. For example, the expected change in the mean character value of daughters, $R_{da}$, equals the expected mean of daughters of selected parents minus the mean of females before selection. Applying Equation 4.10,

$$E_s \left[E(z_{da} | z_{fa}, z_{mo})\right] = \mu_{mo} + b_{da,fa} S_{fa} + b_{da,mo} S_{mo}$$

implying

$$R_{da} = b_{da,fa} S_{fa} + b_{da,mo} S_{mo} \quad (4.11a)$$

where $b_{da,fa}$ is the regression coefficient of daughters on their fathers and $b_{da,mo}$ the mother-daughter regression coefficient. Likewise, for sons

$$R_{so} = b_{so,fa} S_{fa} + b_{so,mo} S_{mo} \quad (4.11b)$$

**Example 2.** Coyne and Beecham (1987) estimated the following parent-offspring regression coefficients for abdominal bristle number in laboratory populations of *Drosophila melanogaster*:

- **Mother-son**: $b_{so,mo} = 0.39 \pm 0.08$
- **Mother-daughter**: $b_{da,mo} = 0.32 \pm 0.08$
- **Father-son**: $b_{so,fa} = 0.13 \pm 0.10$
- **Father-daughter**: $b_{da,fa} = 0.40 \pm 0.08$

Note that the father-son regression has a significantly smaller slope than the three other parent-offspring sex combinations. Assume that these estimated values are indeed the true values and that different amounts of selection are applied to fathers and mothers. Suppose the mean increase in bristle number in selected fathers is 2, while selected mothers show a mean decrease of 1 bristle. What is the expected change in mean bristle number in the male and female offspring using these estimated regression coefficients, assuming all parent-offspring regressions
are linear? Here $S_{mo} = -1$ and $S_{fa} = 2$, and from Equation 4.11a, the expected change in bristle number in females is

$$R_{da} = b_{da,mo}S_{mo} + b_{da,fa}S_{fa} = 0.32(-1) + 0.40(2) = 0.48$$

Likewise, from Equation 4.11b, the expected change in males is

$$R_{so} = b_{so,fa}S_{fa} + b_{so,mo}S_{mo} = 0.13(2) + 0.39(-1) = -0.13$$

Even though we selected for increased bristle number in males and decreased number in females, the expected response is the exact opposite: a decrease in males and an increase in females.

Equation 4.11 is very general, requiring only that the biparental regression is linear, in which case $b_{o,fa}$ and $b_{o,mo}$ are partial regression coefficients and can be computed from covariances between relatives. Again, linearity is ensured if the joint distribution of both parents and their offspring is multivariate normal. If there is no correlation between the phenotypes of parents, the partial regression coefficients are standard univariate regression coefficients (LW Chapter 8), so that LW Equation 3.14b gives

$$b_{o,fa} = \frac{\sigma(z_o,z_{fa})}{\sigma^2(z_{fa})} \quad \text{and} \quad b_{o,mo} = \frac{\sigma(z_o,z_{mo})}{\sigma^2(z_{mo})}$$

If mating is random, and genotype × environmental interactions, shared environmental effects, epistasis, and sex-specific effects can all be neglected, the regression slope (for each parent-offspring combination) is $b_{o,p} = h^2/2$ (LW Chapters 7, 17). Defining the total selection differential as the average of both parental values, $S = (S_{fa} + S_{mo})/2$, recovers the breeders’ equation

$$R = \frac{h^2}{2} S_{fa} + \frac{h^2}{2} S_{mo} = h^2 S \quad (4.13)$$

Equation 4.13 shows how differential selection on parents is incorporated into the breeders’ equation. For example, consider selection on dioecious plants. If plants to form the next generation are chosen after pollination, fathers (pollen donors) are chosen at random with respect to the character under selection and $S_{fa} = 0$, giving $R = (h^2/2)S_{mo}$. Conversely, if parents are selected before pollination with equal amounts of selection ($S$) on both sexes, $R = h^2 S$.

**TRUNCATION SELECTION**

Truncation selection is by far the commonest form of artificial selection in plant and animal breeding and in laboratory experiments. Under truncation selection,
only the largest (or smallest) individuals are allowed to reproduce (Figure 4.2). Truncation selection is usually described by either the percent \( p \) of the population saved or the threshold phenotypic value \( T \) below (above) which individuals are culled. The investigator usually sets these in advance of the actual selection. Hence, while \( S \) is trivially computed after the parents are chosen, we would like to predict the expected selection differential given either \( T \) or \( p \). Specifically, given either \( T \) or \( p \), what is the expected mean of the selected parents? We initially assume a large number of individuals are saved. We then turn to complications introduced by finite population size.

**Figure 4.2.** Under truncation selection, the uppermost (lowermost) fraction \( p \) of a population is selected to reproduce. Alternatively, one could set a threshold level in advance, above (below) which individuals are allowed to reproduce. To predict response given either \( p \) or \( T \), we need to know the mean of the selected tail \( (\mu^*) \), from which we can compute \( S = \mu^* - \mu \) and then apply the breeders’ equation.

### Selection Intensities and Differentials Under Truncation Selection

Given the threshold cutoff \( T \), the expected mean of the selected adults is given by the conditional mean, \( E(z | z \geq T) \). Generally it is assumed that phenotypes are normally distributed, and we use this assumption throughout. With initial mean \( \mu \) and variance \( \sigma^2 \), this conditional mean is given by LW Equation 2.14, which gives the expected selection differential as

\[
S = \varphi \left( \frac{T - \mu}{\sigma} \right) \frac{\sigma}{p} \quad (4.14)
\]

where \( p \) is the fraction saved and \( \varphi(x) = \left(\frac{2\pi}{\sigma^2}\right)^{-1/2} e^{-x^2/2} \) is the unit normal density function evaluated at \( x \) (LW Equation 2.11).

Generally, it is the fraction to be saved \( p \) (rather than \( T \)) that is preset by the investigator (for example, if we save the uppermost 5 percent, what is the expected \( S \)?). Given \( p \), to apply Equation 4.14, we must first find the threshold value \( T_p \) satisfying \( Pr(z > T_p) = p \). Notice that \( T \) in Equation 4.14 enters only as \( (T - \mu)/\sigma \), which transforms \( T_p \) to a scale with mean zero and unit variance.
Hence,

\[
\Pr \left( \frac{z - \mu}{\sigma} > \frac{T_p - \mu}{\sigma} \right) = \Pr \left( U > \frac{T_p - \mu}{\sigma} \right) = p
\]

where \( U \sim N(0, 1) \) denotes a unit normal random variable. Define \( z[p] \), the \textbf{probit transformation} of \( p \) (LW Chapter 11), by

\[
\Pr \left( U < z[p] \right) = p
\]

Hence

\[
\Pr \left( U \geq z[1-p] \right) = p
\]

It immediately follows that \( z[1-p] = (T_p - \mu)/\sigma \), and Equation 4.14 gives the expected selection intensity as

\[
\bar{t} = \frac{S}{\sigma} = \frac{\varphi(z[1-p])}{p}
\]

(LW Figure 2.4 plots \( \bar{t} \) as a function of \( p \) (labeled in this graph as \( (\mu_s - \mu)/\sigma \) versus \( \Phi_T \)). One can obtain \( z[1-p] \) from normal distribution tables (LW Chapter 11). Alternatively, a number of approximations have been suggested for Equation 4.16a. Assuming normality, Smith (1969) has suggested

\[
\bar{t} \approx 0.8 + 0.41 \ln \left( \frac{1}{p} - 1 \right)
\]

Simmonds (1977) found that this approximation is generally quite good for \( 0.004 \leq p \leq 0.75 \), and offered alternative approximations for \( p \) values outside this range. The most precise approximation is Saxton’s (1988), with

\[
\bar{t} \approx \frac{2.97425 - 3.35197 p^{0.2} - 1.9319 p^{0.4} + 2.3097 p^{0.6}}{0.51953 + 0.88768 p^{0.2} - 2.38388 p^{0.4} + p^{0.6}}
\]

\[ \text{Example 3.} \] Consider selection on a normally distributed character in which the upper 5% of the population is saved (\( p = 0.05 \)). From unit normal tables (e.g., Table 11.1 of LW), \( z[1-0.05] = 1.645 \) as \( \Pr[U \geq 1.645] = 0.05 \). Hence,

\[
\bar{t} = \frac{\varphi(1.645)}{0.05} = \frac{0.103}{0.05} \approx 2.06
\]

Applying Equation 4.3 gives the expected response to this amount of selection as \( R = h^2 \sigma 2.06 \). Smith’s approximation gives the selection intensity as

\[
\bar{t} \approx 0.8 + 0.41 \ln \left( \frac{1}{0.5} - 1 \right) \approx 2.01
\]
which is quite reasonable. Saxton’s approximation gives $\tau \simeq 2.06$.

Correcting the Selection Intensity for Finite Samples

If the number of individuals saved is small, the preceding formulae overestimate the selection differential because of sampling effects (Burrows 1972). To see this, assume $M$ adults are sampled at random from the population and the largest $N$ of these are used to form the new generation, giving $p = N/M$. The expected selection coefficient is computed from the distribution of order statistics. Rank the $M$ observed phenotypes as $z_{1,M} \geq z_{2,M} \ldots \geq z_{M,M}$ where $z_{k,M}$ is referred to as the $k$th order statistic when $M$ observations are sampled. The expected selection intensity is given by the expected mean of the $N$ selected parents, which is the average of the first $N$ order statistics,

$$E(\tau) = \frac{1}{\sigma} \left( \frac{1}{N} \sum_{k=1}^{N} E(z_{k,M}) - \mu \right) = \frac{1}{N} \sum_{k=1}^{N} E(z'_{k,M})$$

where $z'_{k,M} = (z_{k,M} - \mu)/\sigma$ are the standardized order statistics. Properties of order statistics have been worked out for many distributions (Kendall and Stuart 1977, David 1981, Sarhan and Greenberg 1962, Harter 1970a,b). For the unit normal, see Harter (1961) for expected values and Sarhan and Greenberg (1962) for variances and covariances. Figure 4.3 plots exact values for the expected selection intensity for small values of $N$ when phenotypes are normally distributed. Note that finite population size results in Equation 4.16a overestimating the actual selection intensity, although the difference is small unless $N$ is very small.
p = N/M, the proportion selected

Figure 4.3. The expected selection intensity $E(\tau)$ under truncation selection with normally-distributed phenotypes, as a function of the total number of individuals measured $M$ and the fraction of these saved $p = N/M$, $N$ being the number of these adults allowed to reproduce. The curve $M = \infty$ is given by using Equation 4.16a, which is exact if $N$ and $M$ are infinite. The values on the curves for $M = 10, 20, 50, and 100$ were obtained from the average of the expected values of the $N$ largest unit normal order statistics (Harter 1961). Note that Equation 4.16a is generally a good approximation, unless $N$ is very small.

Burrows (1972) developed a finite-sample approximation for the expected selection intensity for any reasonably well-behaved continuous distribution. Using the standardized variable $y = (z - \mu)/\sigma$ simplifies matters considerably. Letting $\phi(y)$ be the probability density function of the phenotypic distribution, and $y_p$ the truncation point (e.g., $Pr(y \geq y_p) = p$), Burrows’ approximation is

$$E(\tau_{(M,N)}) \simeq \mu_{y_p} - \frac{(M - N) p}{2N(M + 1) \phi(y_p)} \quad (4.17a)$$
where

\[ \mu_{y_p} = E(y|y \geq y_p) = \frac{1}{p} \int_{y_p}^{\infty} x \varphi(x) \, dx \]

is the truncated mean, which can be obtained by numerical integration. Since the second term of Equation 4.17a is positive, if \( M \) is finite the expected truncated mean overestimates the expected standardized selection differential.

For a normal distribution, LW Equation 2.14 implies \( \mu_{y_p} = \varphi(y_p)/p \) giving

\[ E(\tau_{(M,N)}) \simeq \tau - \left[ \frac{1}{N} \right] \frac{1}{\tau} \]

\[ = \tau - \left[ \frac{1 - p}{2p(M + 1)} \right] \frac{1}{\tau} \]

(4.17b)

(4.17c)

where \( \tau \) is given by Equation 4.16a. Bulmer (1980) suggests an alternative finite-sample approximation for \( E(\tau_{(M,N)}) \) when phenotypes are normally distributed, using Equation 4.16a with \( p \) replaced by

\[ \bar{p} = \frac{N + 1/2}{M + N/(2M)} \]

(4.17d)

**Example 4.** Consider the expected selection intensity on males when the upper 5% of the sampled males are used to form the next generation and phenotypes are normally distributed. If the number of males sampled is very large, then from Example 3, the expected selection intensity is \( \tau \simeq 2.06 \). Suppose, however, that only 20 males are sampled so that only the largest male is allowed to reproduce in order to give \( p = 0.05 \). The expected value for this male is the expected value of the largest order statistic for a sample of size 20. For the unit normal, the expected value of the largest order statistic in a sample of 20 is \( \simeq 1.87 \) (Harter 1961) and hence \( E(\tau_{(20,1)}) \simeq 1.87 \). There is considerable spread about this expected value, as the standard deviation of this order statistic is 0.525 (Sarhan and Greenberg 1962). How well do the approximations of \( E(\tau_{(20,1)}) \) perform? Burrow’s approximation gives

\[ E(\tau_{(20,1)}) \simeq 2.06 - \frac{(20 - 1)}{2(20 + 1) \cdot 2.06} = 2.06 - 0.22 = 1.84 \]

while Bulmer’s approximation uses

\[ \bar{p} = \frac{1 + 1/2}{20 + 1/40} \simeq 0.075 \]

which gives \( z_{0.075} \simeq 1.44 \). Since \( \varphi(1.44) = 0.1415 \),

\[ E(\tau_{(20,1)}) \simeq 0.1415/0.075 \simeq 1.89 \]
SELECTION ON THRESHOLD CHARACTERS

As discussed in LW Chapters 11 and 25, discrete characters can often be modeled by assuming an underlying continuous character $z$ (usually referred to as the liability) that maps to the observed discrete character states (Figure 4.4). Consider the simplest case where the character is either present ($z \geq T$) or absent ($z < T$). Let $\mu_t$ be the mean liability and $q_t$ the frequency of individuals displaying the character in generation $t$. If liability is well enough behaved to satisfy the assumptions of the breeders’ equation (e.g., a linear biparental-offspring regression, no epistasis, genotype-environmental interactions or correlations), then $\mu_{t+1} = \mu_t + h^2 S_t$. We index the selection differential $S_t$ by generation, as the amount of selection on threshold characters changes each generation. The problem is to estimate $\mu_t$ from the observed frequencies of character states. Our analysis will be restricted to a single threshold, but extension to multiple thresholds is straightforward (see Lande 1978).

![Figure 4.4.](image)

Figure 4.4. Selection on a character with a single threshold $T$. $z$ is the value on the underlying scale of liability. Assuming $z$ is a well-behaved quantitative character, $\mu_{t+1} = \mu_t + S_t h^2$, where $S_t = \mu^* - \mu_t$. Using the probit transform (Equation 4.15) we can translate from $q_t$, the frequency of individuals displaying the character, to $\mu_t$, the mean of $z$.

If the values on the underlying scale are normally distributed, we can choose a scale that sets the threshold value at $T = 0$ and assigns $z$ a variance of one. Since
$z - \mu_t$ is a unit normal, $\Pr(z \geq 0) = \Pr(z - \mu_t \geq -\mu_t) = \Pr(U \geq -\mu_t) = q_t$ and from Equation 6.16

$$\mu_t = -z_{[1-q_t]}$$ (6.17)

where $z_p$, the probit transformation of $p$ is given by Equation 4.15. For example, if 5% of a large population displays the trait, $\mu = -z_{[0.05]}$. From normal probability tables, $\Pr(U < 1.645) = 0.95$, hence $z_{[0.05]} = 1.645$ and $\mu = -1.645$. For small samples, estimation of $\mu$ requires the use of order statistics.

The response to selection, as measured by the change in the frequency of the character, is

$$q_{t+1} = \Pr(U \geq -\mu_{t+1})$$
$$= \Pr(U \geq -\mu_t - h^2 S_t)$$
$$= \Pr(U \geq z_{[1-q_t]} - h^2 S_t)$$ (6.18)

It remains to obtain $S_t = \mu^*_t - \mu_t$, where $\mu^*$ is the mean value of $z$ in the selected parents in generation $t$. The selected population may consist entirely of adults displaying the character. However, more individuals than this may be required to keep the population at constant size, especially if $q_t$ and the number of sampled individuals are small. In this case, the selected adults consist of two populations: those displaying the trait (hence $z \geq 0$) and those not ($z < 0$). Letting $p_t$ be the fraction of selected adults displaying the character,

$$\mu^*_t = (1 - p_t) E(z|z < 0; \mu_t) + p_t E(z|z \geq 0; \mu_t)$$ (6.19)

Applying LW Equation 2.14 and noting that the unit normal density function satisfies $\varphi(x) = \varphi(-x)$, gives

$$E(z|z \geq 0; \mu_t) = \mu_t + \frac{\varphi(\mu_t)}{q_t}$$
$$E(z|z < 0; \mu_t) = \mu_t - \frac{\varphi(\mu_t)}{1 - q_t}$$

Substituting into Equation 6.19 gives

$$S_t = \mu^*_t - \mu_t = \frac{\varphi(\mu_t)}{q_t} \frac{p_t - q_t}{1 - q_t}$$ (6.20)

As expected, if $p_t > q_t$, $S_t > 0$. $S_t$ depends critically on $q_t$ and is very unlikely to remain constant over several generations of selection. Maximal selection occurs if only individuals displaying the trait are saved ($p_t = 1$), in which case Equation 6.20 reduces to $S_t = \varphi(\mu_t)/q_t$. 
Example 5. Consider a threshold character whose liability has heritability $h^2 = 0.25$. What is the expected response to selection if the initial frequency of individuals displaying the character is 5% and selection is practiced by selecting only adults displaying the character? As was calculated earlier, $q_0 = 0.05$ implies $\mu_0 = -1.645$. In each generation, only individuals displaying the trait are saved. Thus

$$S_0 = \varphi(-1.645)/0.05 \approx 0.106/0.05 \approx 2.064$$

giving

$$\mu_1 = \mu_0 + 0.25, \quad S_0 = -1.645 + 0.25 \times 2.062 = -1.129$$

and hence

$$q_1 = \Pr(U \geq 1.129) = 0.129$$

After one generation of selection, the character frequency is expected to increase to 12.9%. Changes in $q$ and $S$ after further iterations are plotted in Figure 4.5.

**Figure 4.5.** Response to selection on a threshold character. The initial frequency of individuals displaying the trait is 5%. Selection occurs by only allowing adults displaying the trait to reproduce ($p_t = 1$). Solid circles denote $q_t$, open squares denote $S_t$. The liability is assumed to have $h^2 = 0.25$.

Why did we not estimate $\mu_t^*$ using $z_{[1-q_t]}$? The reason is that the distribution of $z$ values in selected parents is a weighted average of two truncated normal
density functions. This distribution is not normal. However, we assume that normality is restored in the next generation due to segregation plus the addition of the environmental value. We examine the validity of this assumption in Chapter 10.

PERMANENT VERSUS TRANSIENT RESPONSE

As was discussed in LW Chapters 7 and 17, the slope of the parent-offspring regression can be inflated by epistasis, genotype × environment interactions and correlations, and environmental effects shared by parents and their offspring. Hence, even if the parent-offspring regression is linear, the slope can deviate significantly from $h^2/2$, altering the response from that predicted by the breeders’ equation. For example, with epistasis and correlation between parental and offspring environmental values,

$$b_{op} = \frac{h^2}{2} + \frac{1}{\sigma^2_\varepsilon} \left( \frac{\sigma^2_{AA}}{4} + \frac{\sigma^2_{AAA}}{8} + \frac{\sigma^2_{AAAA}}{16} + s + \sigma(E_p, E_o) \right)$$

Assuming a linear biparental regression, the response to a single generation of selection becomes

$$R = h^2 S + \frac{S}{\sigma^2_\varepsilon} \left( \frac{\sigma^2_{AA}}{2} + \frac{\sigma^2_{AAA}}{4} + \frac{\sigma^2_{AAAA}}{8} + s + \sigma(E_{fa}, E_o) + \sigma(E_{mo}, E_o) \right) \quad (4.21)$$

which can deviate significantly from $h^2 S$. Why then do we pay so much attention to $h^2$?

The reason is that we are interested in the permanent response to selection. Recall that one of our main assumptions throughout this chapter is that changes in allele frequencies are negligible, hence any changes in genetic variances are due to gametic-phase disequilibrium rather than changes in allele frequencies. Under this assumption, epistasis and/or shared environmental factors inflate the transient response to selection, but once selection stops, their contribution to response decays away. Changes in genotypic frequencies attributable to gametic-phase disequilibrium and environmental correlations are due to nonrandom associations built up by selection. Recombination and randomization of environmental effects causes these correlations to decay. Conversely, changes in allele frequencies are permanent. Once selection is stopped, the new allele frequencies are stable (assuming that our time scale for observing a population is such that drift and mutation have negligible effects). Hence, as will be shown shortly, the permanent response under the conditions leading to Equation 4.21 is $h^2 S$. One exception, discussed in Chapter 6, is when significant inbreeding occurs. In this case, $\sigma^2_{AA}$ and other non-additive variance components ($\sigma^2_{DI}, \sigma_{ADI}$) introduced in Chapter 3 can contribute to permanent response.
Response with Epistasis

The response when additive × additive epistatic variance is present was examined by Griffing (1960a,b) for the infinitesimal model (Chapter 10). Under the assumption that phenotypes are normally distributed and that the effects at any particular locus are small relative to the total phenotypic variation, the response to one generation of selection is

\[ R = S \left( h^2 + \frac{\sigma_{AA}^2}{2\sigma_z^2} \right) \] (4.22)

One might expect that \( R(n) \), the cumulative response after \( n \) generations of selection, is simply \( n \) times the result given by Equation 4.22. However, any increased response due to epistasis is only temporary, reflecting gametic-phase disequilibrium generated by selection. As disequilibrium decays under recombination, so does the component of response due to epistasis. This occurs because the contribution from epistasis is due to favorable combinations of alleles at different loci, specifically those alleles that interact epistatically to change the character in the direction favored by selection. Recombination breaks down these combinations, removing the epistatic contribution. Griffing showed that for two linked loci (separated by recombination fraction \( c \)), the response when a generation of selection is followed by \( \tau \) generations of no selection is

\[ S \left( h^2 + (1 - c)^\tau \frac{\sigma_{AA}^2}{2\sigma_z^2} \right) \] (4.23)

which converges to \( h^2 S \). Equation 4.23 follows by noting that the probability a gamete containing specific alleles from both loci remains intact following one generation of recombination is \( 1 - c \). Thus, after \( \tau \) generations only \( (1 - c)^\tau \) of the favorable two-locus combinations selected at \( \tau = 0 \) remain unaltered by recombination.

Summing Equation 4.23 over \( t \) gives the cumulative response after \( t \) generations with constant selection differential \( S \) as

\[ R(t) = t h^2 S + R_{AA}(t) \] (4.24a)

where

\[ R_{AA}(t) = S \frac{\sigma_{AA}^2}{2\sigma_z^2} \left( \sum_{i=1}^{t} (1 - c)^{i-1} \right) = S \left( \frac{1 - (1 - c)^t}{1 - c} \right) \left( \frac{\sigma_{AA}^2}{2\sigma_z^2} \right) \] (4.24b)

denotes the cumulative additive × additive epistatic contribution. The last equality follows using the partial sum of a geometric series:

\[ \sum_{i=0}^{n} x^i = \frac{1 - x^{n+1}}{1 - x} \] (4.25)
If loci are unlinked \((c = 1/2)\), \(R_{AA}(t)\) rapidly converges to \(S \sigma_{AA}^2 / \sigma_z^2\), while if loci are completely linked \((c = 0)\), \(R_{AA}(t) = t S \sigma_{AA}^2 / (2 \sigma_z^2)\). Hence, with tight linkage, the transient contribution from epistasis can be considerable. With \(t\) generations of selection followed by \(\tau\) generations of no selection, the cumulative response is

\[
t h^2 S + (1 - c)^\tau R_{AA}(t)
\]

which converges to \(R = t h^2 S\) for large \(\tau\) (i.e., many generations after selection is stopped), which is the value predicted from the breeders’ equation.

**Figure 4.6.** The permanent and transient response to selection assuming pairwise epistasis in a diploid, with \(h^2 = 1/4\) and \(\sigma_{AA}^2 / \sigma_z^2 = 1/2\). **Top:** The cumulative response (in units of \(S\)) assuming a constant amount of selection for various values of \(c\). Note that even with this large amount of epistasis (\(\sigma_{AA}^2\) accounting for half the total variance), it is difficult to distinguish the curvilinear response with epistasis from a linear response. Considerable epistasis can be present and still not be detected from the observed cumulative response. **Bottom:** The decay of the response to a single generation of selection due to the decay of the contribution from epistasis. Provided \(c > 0\), the cumulative response eventually decays to \(h^2 S = S/4\), the expectation under no epistasis.
The presence of epistasis can result in a curvilinear selection response if $\sigma_{AA}^2/\sigma_z^2$ is sufficiently large significant. However, as Figure 4.6 shows, it is usually difficult to distinguish this from a linear response. Further, much of the curvilinearity occurs in the first few generations, as with a constant selection differential, the increment to response decreases each generation, eventually converging to $h^2S$ and hence a linear response from that point on.

Once selection is relaxed, the total response decays back to that predicted from the breeders’ equation. Interestingly, this situation mimics the effects of natural selection countering artificial selection, which also results in a decay of the cumulative response once artificial selection stops. Thus, in order to predict the permanent response correctly we must know $h^2$. If only the parent-offspring slope is estimated, this can overestimate the final amount of response due to the inclusion of $\sigma_{AA}^2$ and higher order epistatic variances.

Griffing’s analysis is restricted to two loci, and hence limited to only pair-wise (additive $\times$ additive) epistasis. Equation 4.21 gives the single-generation response for arbitrary levels of additive epistasis, provided the biparental offspring regression is linear. Again assuming the infinitesimal model (and unlinked loci), Bulmer (1980) found the response due to a single generation of selection decays in the next generation to

$$R = S \left( h^2 + \frac{1}{4} \frac{\sigma_{AA}^2}{\sigma_z^2} + \frac{1}{16} \frac{\sigma_{AAA}^2}{\sigma_z^2} + \frac{1}{64} \frac{\sigma_{AAAA}^2}{\sigma_z^2} + \ldots \right)$$

which again rapidly converges to $R = h^2S$ after several generations without selection. For $n$-locus additive epistasis (e.g., $\sigma_{AA...A}^2$, where there are $n$ A’s), the per generation decay rate for unlinked loci is $(1/2)^{n-1}$, the probability that a parental gamete containing specific alleles at $n$ unlinked loci is passed on to an offspring. The probability that such a gamete remains unchanged after $t$ generations is $2^{-t(n-1)}$, which rapidly converges to zero. A final caveat is that these results apply to infinite populations. As we will see in Chapter 12, with finite populations some of the additive epistatic contribution can be permanent.

**Selection on Autotetraploids**

Polyploidy is very common in plants and can introduce complications in predicting the response to selection. For example, the dynamics of selection response for autotetraploids with dominance is very similar to diploids with epistasis. From LW Equation 7.22 and LW Table 7.5, the autotetraploid parent-offspring covariance when dominance (but no epistasis) is present is

$$\sigma(z_p, z_o) = \frac{\sigma_A^2}{2} + \frac{\sigma_D^2}{6}$$

This also assumes no shared environmental effects, genotype $\times$ environment interactions or correlations.
The inflation in the parent-offspring covariance is due to dominance interactions between the two alleles per locus that each autotetraploid parent passes on to its offspring. Thus, like epistasis in diploids, favorable combinations of alleles can be passed down from parent to offspring in polyploids. With equal amounts of selection on both sexes (e.g., selection occurs before pollination), the resulting response (assuming linearity of the parent-offspring regression) is

\[ R = S \left( h^2 + \frac{\sigma_D^2}{3\sigma^2} \right) \]  

(4.27)

If selection occurs after pollination, \( S \) is replaced throughout the rest of our discussion by \( S/2 \). Gallais (1975) extended Griffing’s (1960a) method (and hence assumed phenotypes are normally distributed with each gene having a very small effect on the character) to obtain the response after \( t \) generations of selection with constant differential \( S \) as

\[ R(t) = th^2S + R_D(t) \]  

(4.28a)

where

\[ R_D(t) = S \frac{3}{2} \left[ 1 - \left(1/3\right)^t \right] \frac{\sigma_D^2}{3\sigma^2} \]  

(4.28b)

which converges to \( S \left( \sigma_D^2/2\sigma^2 \right) \). Segregation reduces the departure from tetraploid Hardy-Weinberg proportions generated by the selection of favorable combinations of allelic pairs, reducing their contribution to response. The response for \( t \) generations of selection followed by \( \tau \) generations of no selection is

\[ th^2S + \left(1/3\right)^\tau R_D(t) \]

which again converges to \( th^2S \). In LW Chapter 4, it was shown that in an autotetraploid the difference in the frequency of pairs of alleles from Hardy-Weinberg expectation decays by 1/3 each generation in the absence of double reduction \((c = 0)\), as would occur for a locus completely linked to the centromere. More generally, if \( c \) is the per-generation probability of a double reduction, the decay rate of 1/3 is replaced in the above equations by \( (1-c)/3 \). Wricke and Weber (1986) discuss additional topics on autotetraploid selection, while single-locus models have been examined by R. Hill (1971). Swanson et al. (1974) found that if some double reductions occur \((c > 0)\), the additive variance is slightly inflated over the value expected with no double reductions \((c = 0)\), permanently increasing selection response. This results from the slight excess of homozygotes at equilibrium over the frequency expected from Hardy-Weinberg expectations (see LW Chapter 4).

Response due to Environmental Correlation

As Equation 4.21 indicates, shared parent-offspring environmental effects (e.g., \( \sigma(E_p, E_o) \neq 0 \)) can influence response. We show here that this contribution is
also transient. Consider a character whose variation is entirely environmental, in which case the phenotypic value can be decomposed as

\[ z = \mu + E = \mu + c_{fa} + c_{mo} + e \]

where \( \mu \) is the mean value of the character when environmental effects are randomly distributed, and the environmental value \( E \) has been decomposed into the maternal and paternal contributions to the offspring due to shared environmental effects \( (c_{mo} \text{ and } c_{fa}) \) and a residual due to special environmental effects \( (e) \). In order to predict the shared environmental contribution from a parent, we assume the simplest model, that a fraction \( b \) of the total environmental value of a parent is passed onto its offspring. This model serves as a useful introduction to some of the dynamics that can occur with maternal effects (which are examined in the next section). Thus, the expected contribution from a father to his offspring is

\[ e_{fa} = bE_{fa} = b(z_{fa} - \mu) \]

where \( E_{fa} \) is the father’s total environmental value. To simplify matters further, we assume that this regression coefficient is independent of the sexes of the parent and offspring, although this can easily be relaxed. If we assume that parents and offspring have the same phenotypic variance, then

\[ b = \rho/2 \]

where \( \rho \) is the midparent-offspring correlation. Note that \( \rho \) can be negative. For example, suppose parents and offspring compete for a limited amount of common resource. Larger parents may gather a disproportionate share of resources, resulting in smaller offspring.

Provided \( E_{fa} \) and \( E_{mo} \) are independent, the expected value of an offspring from parents with phenotypic values \( z_{fa} \) and \( z_{mo} \) is

\[ E(z_o | z_{mo}, z_{fa}) = \mu + \frac{\rho}{2} (z_{fa} - \mu) + \frac{\rho}{2} (z_{mo} - \mu) \quad (4.29) \]

Denoting the mean of adults selected in generation \( t \) by \( \mu_t^* \), the mean at generation \( t + 1 \) is given by

\[ \mu_{t+1} = \mu + \rho (\mu_t - \mu) \]

where \( (\mu_t^* - \mu) \) is the environmental deviation in selected parents at generation \( t \), \( \rho \) of which is passed on to their offspring. Rewriting the mean after selection as \( \mu_t^* = \mu_t + S_t \),

\[ \mu_{t+1} = \mu + \rho (\mu_t + S_t - \mu) \quad (4.30) \]

The change in mean in generation \( t \), \( \Delta \mu_t = \mu_{t+1} - \mu_t \), is

\[ \Delta \mu_t = [\mu + \rho (\mu_t + S_t - \mu) - [\mu + \rho (\mu_{t-1} + S_{t-1} - \mu)]
\]

\[ = \rho [(\mu_t - \mu_{t-1}) + (S_t - S_{t-1})]
\]

\[ = \rho [\Delta \mu_{t-1} + (S_t - S_{t-1})] \quad (4.31a) \]

Suppose that constant selection (with differential \( S \)) is applied starting at generation 1. Here, \( \Delta \mu_0 = 0 \), \( S_0 = 0 \), and \( S_t = S \) for \( t \geq 1 \). Equation 4.31a gives \( \Delta \mu_1 = \rho S \). Further iterations yield

\[ \Delta \mu_t = \rho^t S \quad (4.31b) \]
which decreases each generation, approaching zero for large \( t \). Hence, even under continued selection, the response to selection eventually stops. The reason for this decline in the per generation rate of response can be seen from Equation 4.31b. Changes in the character mean due to previous generations of selection decay, countering selection in the current generation. Only \( \rho \) of the change from generation \( t - 1 \) is passed on, and, in general, only \( \rho^k \) of the response from generation \( t - k \) persists by generation \( t \). Summing over Equation 4.31b, the total response to selection after \( t \) generations is

\[
R(t) = \mu_{t+1} - \mu_0 = \sum_{i=1}^{t} \Delta \mu_i = S \sum_{i=1}^{t} \rho^i \tag{4.31c}
\]

Recalling the partial sum of a geometric series (Equation 4.25), this reduces to

\[
R(t) = S \frac{\rho}{1 - \rho} (1 - \rho^t) \tag{4.32}
\]

Thus, no matter how long selection is applied, the mean can never change by more than \( S\rho/(1 - \rho) \). Further, none of this response is permanent. Suppose selection is stopped after \( t \) generations, giving \( S_t = S \), \( S_{t+\tau} = 0 \) for \( \tau \geq 1 \). Substituting into Equation 4.31a and using Equation 4.31b, the expected change in generation \( t + 1 \) is

\[
\Delta \mu_{t+1} = \rho^\tau (\Delta \mu_t - S) = -S \rho^\tau (1 - \rho^t) \tag{4.33}
\]

By generation \( t + \tau \) the cumulative response is

\[
R(t + \tau) = R(t) - S (1 - \rho^\tau) \sum_{i=1}^{\tau} \rho^i = \rho^\tau R(t) \tag{4.34}
\]

which converges to zero, with the rate of decay being set by the amount of environmental correlation, \( \rho \) (Figure 4.7). Hence, while there can be some selection response when the resemblance between relatives is entirely environmental, any response is transient, decaying away once selection stops.
Figure 4.7. Response when resemblance between relatives is due entirely to correlation between environmental values in parents and offspring. Selection with constant differential $S$ starts at $t = 0$ and continues until generation 10 (indicated by the arrow), at which point selection is stopped. Note the interesting dynamics that occur if environmental values are negatively correlated. The response to selection is reversed with respect to the selection differential. In this case, selection for increased character value results in a decreased mean value, with the total response eventually converging to $-S/3$ (for $\rho = -0.5$). Once selection is relaxed there is an initial positive response (generation 11), although response quickly decays to zero.

MATERNAL EFFECTS

Any influence that the mother’s phenotype has on the phenotype of her offspring (beyond being a predictor of breeding value) is considered a maternal effect (LW Chapter 23). Body size, amount of care invested in offspring, and endosperm production are examples of maternal performance characters that potentially influence a variety of characters in the progeny. Paternal effects are also possible, especially in situations where the father plays some role in caring for the offspring. While paternal effects are not considered here, they can be treated in exactly the same fashion as maternal effects.

Assuming a maternal effect, the phenotype of a character can be decomposed as $z = G + M + e$, the sum of a genotypic value $G$ plus a maternal component $M$ plus an environmental deviation $e$ (this assumes no interactions between $G$, $M$, and $e$). There are a number of subtle (and important) features with this apparently simple model. First, how many characters are we really considering? The maternal effect may be regarded as either a direct function of the character being considered, or it may be due to a character (or suite of characters) correlated with the one being considered. For example, weight may be correlated with the amount of lactation, in which case maternal weight can influence offspring weight indirectly because larger females tend to have increased lactation. In the latter case, if there is a genetic basis to the maternal effect, selection on the direct character (e.g., body weight) is expected to give a correlated response in the maternal performance characters (e.g., lactation), and vice versa (Kirkpatrick and Lande 1989). We return to these points when we discuss evolution of correlated characters in Chapters 17-20. For now attention is restricted to single-character models.

Response under Falconer’s Model

The simplest model of maternal effects (motivated by the inheritance of litter size in mice) is that of Falconer (1965), which deals with a single maternal character: the only character that has a maternal influence on $z$ is $z$ itself (reviewed in LW Chapter 23). Falconer assumes the maternal contribution is a linear function of the
maternal phenotype \( z_{mo} \) so that \( M = m z_{mo} \) and the phenotypic decomposition becomes

\[
z = G + m z_{mo} + e \tag{4.35}
\]

Conceivably, \( M \) could be a nonlinear function of \( z_{mo} \), but linearity is assumed for tractability. We refer to Equation 4.35 as the **dilution model**, as the effect of the maternal phenotype is diluted over several generations. The parameter \( m \) can be regarded as the partial regression coefficient (holding genotypic value constant) of offspring phenotype on maternal phenotype and can be estimated as the difference between the mother- and father-offspring regression slopes (LW Equation 23.13). Negative estimates of \( m \) have been reported. Falconer (1965) estimated \( m = -0.15 \) for litter size in mice and Janssen et al. (1988) estimated \( m \) values of \(-0.58\) and \(-0.40\) for age of maturity in two replicate lines of springtails. Maternal effects can result in unusual dynamics such as time lags. This is not surprising considering the dynamics of selection response when variation in the character is entirely environmental (Figure 4.7).

Assume that the joint distribution of phenotypes and breeding values in parents and offspring is multivariate normal. Further assuming no epistasis, the expected phenotypic value of an offspring whose mother has phenotypic value \( z_{mo} \) is

\[
E(z_o \mid A_{mo}, A_{fa}, z_{mo}) = \frac{A_{mo}}{2} + \frac{A_{fa}}{2} + m z_{mo} \tag{4.36a}
\]

where \( A_{mo} \) and \( A_{fa} \) are the maternal and paternal breeding values (see example 7 in LW Chapter 8). Averaging over the selected parents, the mean in generation \( t + 1 \) is

\[
\mu_z(t + 1) = \frac{A_{fa}^*(t) + A_{mo}^*(t)}{2} + m \mu_{mo}^*(t) \tag{4.36b}
\]

where \( A_{fa}^*(t) \) and \( A_{mo}^*(t) \) are the mean breeding values of the selected parents and \( \mu_{mo}^*(t) \) the mean phenotypic value of selected mothers in generation \( t \). Using the regression of breeding value on phenotype,

\[
A = \mu_A + b_{Az} (z - \mu_z) + e
\]

allows us to predict the breeding value \( A \) of an individual from its phenotypic value \( z \). Thus we can rewrite \( A_{mo}^*(t) \) as

\[
E_s(A_{mo}) = E_s \left( \mu_A(t) + b_{Az} [z_{mo} - \mu_z(t)] + e \right)
= \mu_A(t) + b_{Az} S_{mo}(t) \tag{4.37}
\]

where \( E_s(\ ) \) denotes the expected value over the selected parents. A similar expression holds for \( A_{fa}^*(t) \). In the absence of maternal effects, \( b_{Az} = h^2 \). However, the dilution model generates a covariance between \( M \) and \( A \), specifically
\[ \sigma_{A,M} = m \sigma_A^2 / (2 - m), \] which in turn alters the covariance between \( z \) and \( A \) (Falconer 1965, Kirkpatrick and Lande 1989, LW Chapter 23). The resulting regression slope (at equilibrium) is

\[ b_{Az} = h^2 \frac{2}{2 - m} \] (4.38)

(Falconer 1965, Van der Steen 1985, Kirkpatrick and Lande 1989). If there is a negative maternal effect \( (m < 0) \), \( b_{Az} < h^2 \), reducing the correlation between breeding value and phenotype. Conversely, \( m > 0 \) increases the correlation between breeding value and phenotype above \( h^2 \). Applying Equations 4.36–4.38 and using \( \mu_m^*(t) = \mu_z^*(t) + S_{mo}(t) \), gives

\[ \mu_z(t + 1) = \mu_A(t) + \frac{h^2}{2 - m} \left( S_{mo}(t) + S_{fa}(t) \right) + m \left( \mu_z(t) + S_{mo}(t) \right) \] (4.39)

(Van der Steen 1985, Kirkpatrick and Lande 1989). The change in population mean over one generation, \( \Delta \mu_z(t) \), is thus

\[
\Delta \mu_z(t) = \mu_z(t + 1) - \mu_z(t) = \\
\left[ \mu_A(t) + \frac{h^2}{2 - m} \left( S_{mo}(t) + S_{fa}(t) \right) + m \left( \mu_z(t) + S_{mo}(t) \right) \right] \\
- \left[ \mu_A(t - 1) + \frac{h^2}{2 - m} \left( S_{mo}(t - 1) + S_{fa}(t - 1) \right) + m \left( \mu_z(t - 1) + S_{mo}(t - 1) \right) \right] \\
= \frac{h^2}{2 - m} \left( S_{mo}(t) + S_{fa}(t) \right) + m S_{mo}(t) + m \left( \Delta \mu_z(t - 1) - S_{mo}(t - 1) \right) \] (4.40)

The last simplification follows from the regression of breeding value on phenotype, with

\[ \mu_A(t) = \mu_A(t - 1) + \frac{h^2}{2 - m} \left( S_{mo}(t - 1) + S_{fa}(t - 1) \right) \]

Equation 4.40 can be interpreted as follows: the first two terms are the change in character value resulting from selection in generation \( t \) due to genetic \( (h^2 / [2 - m]) \) and maternal \( (m) \) contributions. The final term, which can also be expressed as \( m [ \mu_z(t) - \mu_z^*(t - 1) ] \), represents the decay in the maternal contribution from the previous generation.

Starting with an unselected base population, the response to a single generation of selection is

\[ \Delta \mu_z(1) = S_{mo}(1) \left( \frac{h^2}{2 - m} + m \right) + S_{fa}(1) \frac{h^2}{2 - m} \] (4.41)
An interesting consequence of Equation 4.41 is that if \( m < 0 \), there is some possibility of a **reversed response**, where \( \Delta \mu_z \) has opposite sign of \( S \). If \( S_{fa} = S_{mo} = S \), a reversed response is expected if

\[
m < 1 - \sqrt{1 + 2h^2}
\]  

(4.42a)

If selection is only occurring on females, this condition is

\[
m < 1 - \sqrt{1 + h^2}
\]  

(4.42b)

An example of an apparent maternally-induced reversed response was seen by Falconer (1960, 1965) in his selection experiments on litter size in mice. This character shows a negative maternal effect, with \( m \) and \( h^2 \) estimated to be \(-0.13\) and \(0.11\), respectively. Since selection for litter size occurs only in females, Equation 4.42b implies that a reversed response in the first generation is expected (as \( 1 - \sqrt{1 + 0.11} \approx -0.05 > m \)). As Figure 4.8 shows, a reversed response was indeed observed.
Figure 4.8. Top: Falconer’s (1960, 1965) experiments on selection response for litter size in mice. The dashed line is the response to selection for small litters, the thick line selection for large litters, and the thin line the control. Note the reversed response in the first generation in both the up- and down-selected lines. Bottom: Prediction from the model, using Falconer’s estimated values of \( h^2 = 0.11 \) and \( m = -0.13 \). The predicted change in population mean following a single generation of selection on females with \( S_{mo} > 0 \) is plotted. There is a reversed response in the first generation, even though the net genetic change is to increase the character. By generation 3, the nongenetic change in phenotypic mean has largely decayed away, revealing the net genetic change of 

\[
S_{mo} h^2 / [(1 - m)(2 - m)] = 0.044S_{mo}.
\]

An observed reversed response is misleading because the permanent response is expected to have the same sign as \( S \), while the initial observed response also includes a transient component that (in this case) is of opposite sign and of larger magnitude than the permanent response component. It may take several generations for this transient component to decay and reveal the actual genetic changes (Figure 4.8), which from Equation 4.48 is (for a single generation of selection) 

\[
S h^2 [2 / (2 - m)(1 - m)].
\]

The possibility of reversed response hints at some of the complicated dynamics that can appear when maternal effects are present. To examine these dynamics in more detail, consider the dilution model with constant directional selection occurring equally on both sexes. i.e., \( S_{fa}(t) = S_{mo}(t) = S \) for \( t \geq 1 \). Iteration of Equation 4.40 gives

\[
\Delta \mu_z(t) = S \left[ \frac{2h^2}{(1 - m)(2 - m)} \left( 1 - m^t \right) + m^t \right] \quad (4.43a)
\]

which converges (for \( |m| < 1 \)) to

\[
\Delta \mu_z = S \frac{2h^2}{(1 - m)(2 - m)} \quad (4.43b)
\]

Thus after a sufficient number of generations, the per generation change is constant. If \( |m| \) is near zero, the per generation response rapidly converges to the asymptotic value, while if \( |m| \) is near one, the rate of convergence is considerably slower. Summing over the single-generation changes (Equation 4.53a) the cumulative response to \( t \) generations of selection is

\[
R(t) = \frac{S}{1 - m} \left[ t \frac{2h^2}{2 - m} + m (1 - m^t) \left( 1 - \frac{2h^2}{(1 - m)(2 - m)} \right) \right] \quad (4.44a)
\]

which converges (for \( |m| < 1 \)) to

\[
\frac{S}{1 - m} \left[ \frac{2h^2}{2 - m} \left( t - \frac{m}{1 - m} \right) + m \right] \quad (4.44b)
\]
How much of this response is permanent? Suppose selection ends at generation $t$, and denote by $\tau$ the number of generations since selection was stopped. Iterating Equation 4.40 with $S(t) = S, S(t + \tau) = 0$ for $\tau \geq 1$ yields

$$\Delta \mu_z (t + \tau) = m^\tau (\Delta \mu_z (t) - S)$$

(4.45)

where $\Delta \mu_z (t)$ is given by Equation 4.43a. Thus, response continues even after the cessation of selection. Summing Equation 4.45 over $\tau$ yields the cumulative response following the last generation of selection,

$$R^*(\tau) = \frac{m(1 - m^\tau)}{1 - m} (\Delta \mu_z (t) - S)$$

(4.46)

which converges as $\tau \to \infty$ to

$$R^* = S m \left( \frac{2h^2}{(1 - m)(2 - m)} - 1 \right)$$

(4.47)

Summing Equations 4.44b and 4.47, the permanent response to $t$ generations of selection is

$$R(t) + R^* = t h^2 S \frac{2}{(1 - m)(2 - m)}$$

(4.48)

If $R^*$ is opposite in sign to $S$, there is some erosion of the cumulative response upon relaxation of selection (we have already seen a special case of this with reversed response). For $|m| < 1$, erosion in response occurs if

$$0 < m < \frac{3 - \sqrt{1 + 8h^2}}{2}$$

(4.49a)

On the other hand, if maternal effects are either negative ($m < 0$) or sufficiently large

$$m > \frac{3 - \sqrt{1 + 8h^2}}{2}$$

(4.49b)

the response continues for a few generations following the relaxation of selection. Figure 4.9 plots some sample trajectories.
Figure 4.9. Examples of the predicted selection response with maternal effects under Falconer’s dilution model. Selection starts at generation zero, with $S_{fa} = S_{mo} = S$ until generation 10 (arrow), at which point selection stops. We assume $h^2 = 0.35$, with the different curves corresponding to different maternal effect values, $m$. **Top:** Positive maternal effects ($m > 0$). For this value of $h^2$, Equations 4.49b gives the critical $m$ value as 0.52, so that for $m = 0.75$ response continues after selection is relaxed, while response decays for $m = 0.5$ and 0.25. **Bottom:** Negative maternal effects. The dynamics here are considerably more interesting. For this $h^2$ value, Equation 4.49a implies that response continues once selection stops for all values of $m < 0$. Compare with Figure 4.7.

In summary, the presence of maternal effects introduces several complications. First, predicting the response to selection requires not only of the inheritance parameters ($m, h^2$) and current selection differential, but also requires knowledge
of previous selection \(\Delta \mu_z(t-1), S_{mo}(t-1)\). Second, after selection is stopped, the mean is likely to continue to change due to lag effects (e.g., Figure 4.9). If \(m < 0\), the response will continue, while if \(m > 0\) the response can either continue or decay. This clearly causes problems if we are trying to estimate the nature of selection acting on a character by comparing changes in means between generations. For example, an observed cross-generation decrease in a character could be due to four very different causes: (i) \(S < 0\), (ii) \(S > 0\) and a reversed response due to maternal effects, (iii) no selection in the observed generation but a previous history of \(S > 0\), with the decrease in mean due to a positive maternal effect (reflecting a decay in response), or (iv) no current selection but a previous history of \(S < 0\), with the decrease in mean due to a negative (or sufficiently large positive) maternal effect (reflecting a continuation of response).

Other Models of Maternal Effects
Willham (1963, 1972), expanding on the early results of Dickerson (1947), obtained covariances between relatives under a more general model that assumes the maternal effect can be further decomposed as \(M = G_M + e_M\). \(G_M\) is the contribution to \(z\) resulting from the mother’s genotypic value for the maternal performance character, while \(e_M\) is the contribution resulting from the environmental value of the maternal performance character (reviewed in LW Chapter 23). Hence, even though from the offspring’s standpoint, \(M\) is treated as an environmental effect, it can have both a genetic and environmental basis in the mother. As Kirkpatrick and Lande (1989) have pointed out, this model is really a two-character model: the character directly being followed \(z\) (e.g., body weight) and some other character \(z_M\) (e.g., lactation) that has a maternal effect on \(z\). Generalizations of the Dickerson-Willham model have been proposed by Riska et al. (1985) and Kirkpatrick and Lande (1989), but we will defer further discussion of these multiple character models until we discuss the evolution of correlated characters in Chapters 17-20. Mueller and James (1985) examine selection on maternal characters with overlapping generations.

ANCESTRAL REGRESSIONS

A general approach for examining which components of the response are transient is to consider the expected value of an offspring as a function of all its direct relatives that have been under selection. If this ancestral regression is linear (as would occur if the joint distribution of the phenotypic values of all relatives is multivariate normal), response can be described by specifying the regression coefficients by an obvious extension of the biparental regression to now include all selected relatives back to the original unselected base population. For example, if selection starts in generation 0, the response in the first generation is \(R(1) = 2 \beta_{1,0} S_0\), where \(\beta_{1,0}\) is the regression of offspring at generation one on a parent from gen-
eration zero (this assumes both parents have the same regression coefficients and selection differentials, an assumption that will be relaxed shortly). Likewise, the total response after two generations, \( R(2) = 4 \beta_{2,0} S_0 + 2 \beta_{2,1} S_1 \), depends on the nature of selection on the four grandparents and both parents. Note that this formulation allows the parent-offspring regression to change through time (e.g., \( \beta_{2,1} \) need not equal \( \beta_{1,0} \)), as might happen with inbreeding. Similarly, the response following three generations of selection depends upon the nature of selection on that individual’s eight great-grand parents, four grandparents and two parents,

\[
R(3) = 8 \beta_{3,0} S_0 + 4 \beta_{3,1} S_1 + 2 \beta_{3,2} S_2
\]

Proceeding in this fashion gives the response for generation \( T \) as

\[
R(T) = \sum_{t=0}^{T-1} \beta_{T,t} \left( 2^{T-t} S_t \right)
\]

where \( \beta_{T,t} \) is the regression coefficient for the phenotype of an individual in generation \( T \) on one of its relatives in generation \( t < T \). With pure selfing each individual has only a single relative in each previous generation, giving the ancestral regression as

\[
R(T) = \sum_{t=0}^{T-1} \beta_{T,t} S_t
\]

Recall from standard regression theory (LW Chapter 8) that the vector of partial regression coefficients \( \beta = V^{-1} \sigma \), where \( \sigma \) is a vector of covariances between the individuals in generation \( T \) with all relatives in previous generations and \( V \) is the phenotypic covariance matrix for the entire collection of individuals. The key here is that the regression coefficients are entirely determined by the covariances between relatives. If we have independence so that the partial regression coefficients reduce to univariate regression coefficients (i.e., \( \beta_i = \sigma(y,x_i) / \sigma^2 x_i \)), then we have

\[
R(T) = \sum_{t=0}^{T-1} \sigma_{G}(T,t) / \sigma^2(z_t) \left( 2^{T-t} S_t \right)
\]

where \( \sigma_{G}(T,t) = \sigma(z_T,z_t) \) is the cross-generation covariance, the phenotypic covariance between an individual in generation \( t \) and its descendant in generation \( T > t \). With selection under pure selfing, each individual has a single ancestor and the \( 2^{T-t} \) term in Equation 4.50c is absent.

If different relatives in the same generation experience different amounts of selection, with \( S_{k,i} \) being the selection differential on relative \( i \) in generation \( k \), then

\[
R(T) = \sum_{t=0}^{T-1} \beta_{T,t} \left( \sum_{i=1}^{n(t,T)} S_{t,i} \right)
\]
where $n(t, T)$ is the number of relatives in generation $t$ that contribute to response in generation $T$. Note for the case of pure selfing $n(t, T) = 1$. Finally, we can also allow for different regression coefficients on each relative to completely generalize this approach,

$$R(T) = \sum_{t=0}^{T-1} \left( \sum_{i=1}^{n(t, T)} \beta_{T,t,i} S_{t,i} \right)$$  (4.52)

where $\beta_{T,t,i}$ is the regression coefficient of the phenotype on an individual in generation $T$ on its $i$-th relative in generation $t$.

To apply ancestral regression for predicting response, we require that the regression remains linear and that selection-induced changes in the variance and covariances are negligible. Thus, while we allow changes in $\beta_{T,t}$ due to the particular genetic system being considered (e.g., selfing wherein the additive genetic variance decreases by a predictable amount each generation in the absence of selection) we assume that selection does not confound these changes. Bulmer (1980) shows that under the infinitesimal model the joint distribution of an offspring and all its direct ancestors is multivariate normal and hence the ancestral regression is linear. Since selection does not change allele frequencies under the infinitesimal model (Chapter 10), this might suggest that the regression coefficients $\beta_{T,t}$ are unaffected by selection. The problem, however, is that selection generates gametic-phase disequilibrium that can significantly alter the genotypic moments (Chapter 10). For now, we assume that these changes (over short time scales) are small enough to be neglected.

As an application of ancestral regressions, consider additive by additive epistasis. In this case, Cockerham (1984b) found that for two linked loci, the cross-generation covariance is

$$\sigma_G(\tau + t, \tau) = \sigma_A^2(\tau) + \sigma_{AA}^2(\tau) \left( \frac{1 - c}{2} \right)^t$$

giving

$$2^t \sigma_G(\tau + t, \tau) = \sigma_A^2(\tau) + (1 - c)^t \frac{\sigma_{AA}^2(\tau)}{2}$$

If the genetic variances remain constant, then applying Equation 4.50a we recover Equation 4.26.

The behavior of the regression coefficients over time thus informs us about the permanency of response. Note from Equation 4.50a that unless $2^t \beta_{\tau+t,\tau}$ remains constant as $t$ increases, the contribution to cumulative response from selection on adults in generation $t$ changes over time. For example, when loci are strictly additive (no dominance or epistasis), $\sigma_G(\tau + t, \tau) = 2^{-t} \sigma_A^2(\tau)$ and thus $2^t \beta_{\tau+t,\tau} = h^2_\tau$, the standard result from the breeders’ equation. Note that unless $2^t \sigma_G(\tau + t, \tau)$ remains constant, any response contributed decays. Hence any term of $\sigma_G(\tau + t, \tau)$ that decreases by more than $1/2$ each generation contributes only to the
transient response. An exception is with pure selfing where the total contribution in generation \( t \) from an ancestor \( \tau \) previous generations is \( \sigma_G(\tau + t, \tau) \), so that any components that decline as \( \tau \) increases will contribute to the transient response.

**MODIFYING THE BREEDERS’ EQUATION FOR NATURAL POPULATIONS**

One of the most serious limitations to applying the breeders’ equation in natural populations is that selection can occur on (unmeasured) characters that are correlated with the particular character under consideration. Further, genotype-environment correlations can be a concern, as (for example) larger individuals may be able to obtain the best environments. Just how these complications can bias the breeders’ equation was examined by van Tienderen and de Jong (1994). They assume complete additivity (no dominance or epistasis), multivariate normality, and linear parent-offspring regressions. As is discussed in Chapter 9, under these conditions a more general expression than the breeders’ equation for the response to selection is given by Robertson’s Secondary Theorem of Natural Selection, which states that response equals the covariance between relative fitness \( w \) and breeding value \( A \), 

\[
R = \sigma(w, A) .
\]

van Tienderen and de Jong use a path analysis argument (LW Appendix 2) to explore the relationship between response \( R \) and the selection differential \( S \) when complications such as selection on correlated characters and genotype-environment correlations exist.

To present their analysis, we first decompose the phenotype \( z \) as

\[
z = A + E + E_s
\]

where \( A \) is the additive genetic value, \( E \) the general environmental value (for example, the average value for a particular macrohabitat) and \( E_s \) the special environmental value unique to each individual (LW Chapter 6). By construction, \( E_s \) is independent of other variables (so that the total environmental variance is \( \sigma_E^2 + \sigma_{E_s}^2 \)), but \( A \) and \( E \) may be correlated. Consider the path diagrams (LW Appendix 2) in Figure 4.10, which shows possible paths of how the environment value \( E \), the genotypic value \( A \), and the phenotypic value \( z \) can influence fitness. The standard breeders’ equation assumes \( E \) and \( A \) influence fitness only through phenotypic value \( z \). van Tienderen and de Jong examines the more general situation when \( E \) and \( A \) can influence fitness independent of (or in addition to) their effects on \( z \), such as can occur if the character is phenotypically or genetically correlated with under characters under selection. If fitness is entirely set by the phenotypic value of the character of interest, there should be no expected differences in fitness of individuals with the same phenotypic value \( z \) but different underlying genetic \( A \) or environmental values \( E \). However, if other correlated characters are under selection, then individuals with the same \( z \) value can have different fitnesses as correlations between the \( A \) and / or \( E \) values with the genetic
and/or environmental values at these other characters with have influences on fitnesses (also see Chapters 15-17).

To quantify the effects from these different paths to influencing fitness, van Tienderen and de Jong consider the multiple regression of relative fitness \( w \) as a function of \( z \), \( A \), and \( E \), viz.,

\[
w = \alpha + \beta_z z + \beta_A A + \beta_E E + \epsilon
\]

Assumptions under the standard breeders' equation

\[ \text{Es} \rightarrow z \rightarrow w \]

Possible relations between component values (\( A, E, z \)) and fitness

\[ E_s \rightarrow z \rightarrow w \]

**Figure 4.10.** The pathways by which the components of a character, here its phenotype \( z \), additive genetic value \( A \), common environmental effect \( E \), and special environmental effect \( E_s \), influence fitness \( w \). **Top.** The breeders' equation assumes that only the phenotype (\( z \)) of a character influences fitness. This may not be an unreasonable starting assumption for artificial selection wherein the breeder directly chooses individuals on the basis of phenotypes and randomized environments with respect to phenotypes. **Bottom.** Other pathways by which the components of a character can influence fitness. In addition to possible effects from the phenotype, either (or both) of the additive genetic or environmental value can influence fitness independent of their influence on phenotype. For example, an environmental value can influence both the character of interest and independently influence fitness. This creates an association between character value and fitness so that we might observe a character mean being changed within a generation without that character itself being under any selection.

The partial regression coefficients \( \beta \) represent the expected change in fitness holding the other variables constant (LW Chapter 8). For example, \( \beta_z \) is the effect of phenotype on fitness, holding the other variables (\( A \) and \( E \)) constant. From mul-
multiple regression theory (LW Chapter 8), the partial regression coefficients satisfy

$$
\begin{pmatrix}
\sigma(w, z) \\
\sigma(w, A) \\
\sigma(w, E)
\end{pmatrix}
= 
\begin{pmatrix}
S \\
R \\
\sigma(w, E)
\end{pmatrix}
= 
\begin{pmatrix}
\sigma_A^2 \\
\sigma^2_A \\
\sigma^2_A
\end{pmatrix}
\begin{pmatrix}
\sigma(z, A) \\
\sigma(z, E) \\
\sigma(A, E)
\end{pmatrix}
\begin{pmatrix}
\beta_z \\
\beta_A \\
\beta_E
\end{pmatrix}
\begin{pmatrix}
\sigma(w, z) \\
\sigma(w, A) \\
\sigma(w, E)
\end{pmatrix}
(4.53)
$$

Recall that $S = \sigma(w, z)$ follows from the Robertson-Price identity and $R = \sigma(w, A)$ from Robertson’s Secondary Theorem. The first vector contains the covariances between relative fitness and the predictor variables ($z$, $A$, and $E$), while the matrix is the variance-covariance matrix for these predictor variables. Note that

$$
\sigma(z, A) = \sigma(A + E + E_s, A) = \sigma_A^2 + \sigma(A, E)
$$

Likewise, $\sigma(z, E) = \sigma_E^2 + \sigma(E, A)$. Using these identities and taking the first two rows of the above matrix equation gives

$$
S = \sigma_A^2 \beta_z + (\sigma_A^2 + \sigma(E, A)) \beta_A + (\sigma_E^2 + \sigma(E, A)) \beta_E
= \sigma_A^2 \beta_z + \sigma_A^2 \beta_A + \sigma_E^2 \beta_E + \sigma(E, A) (\beta_A + \beta_E)
(4.54a)
$$

and

$$
R = (\sigma_A^2 + \sigma(E, A)) \beta_z + \sigma_A^2 \beta_A + \sigma(E, A) \beta_E
= \sigma_A^2 \beta_z + \sigma_A^2 \beta_A + \sigma(E, A) (\beta_z + \beta_E)
(4.54b)
$$

If there are no genotype-environment correlations [$\sigma(E, A) = 0$],

$$
R = \sigma_A^2 (\beta_z + \beta_A)
(4.55a)
$$

and

$$
S = \sigma_A^2 \beta_z + \sigma_A^2 \beta_A + \sigma_E^2 \beta_E
(4.55b)
$$

Multiplying both sides of Equation 4.55b by $h^2$ and rearranging gives

$$
\sigma_A^2 \beta_z = h^2 \sigma_A^2 \beta_z = h^2 \left( S - \left[ \sigma_A^2 \beta_A + \sigma_E^2 \beta_E \right] \right)
$$

Substituting into Equation 4.55a gives

$$
R = h^2 S + \sigma_A^2 (1 - h^2) \beta_A - h^2 \sigma_E^2 \beta_E
(4.55c)
$$

Hence, selection entirely on additive genetic values ($\beta_A \neq 0$) inflates response over the breeders’ equation, while selection on entirely environmental values ($\beta_E \neq 0$) decreases response relative to the breeders’ equation. Following this same approach gives the general solution [when $\sigma(E, A) \neq 0$] of

$$
R = h^2 S + \sigma_A^2 (1 - h^2) \beta_A - h^2 \sigma_E^2 \beta_E + \sigma(E, A) (\beta_z - h^2 \beta_A + (1 - h^2) \beta_E)
(4.56)
$$
If selection acts only on the phenotype of the character being considered, then \( \beta_A = \beta_E = 0 \) and Equation 4.54a reduces to \( S = \sigma_z^2 \beta_z \), implying \( \beta_z = S/\sigma_z^2 \). Substituting into Equation 4.54b gives the response as

\[
R = \beta_z \left[ \sigma_z^2 + \sigma(E, A) \right] = \left( h^2 + \frac{\sigma(E, A)}{\sigma_z^2} \right) S
\]  

(4.57a)

which (as expected) reduces to the breeders’ equation when there is no genotype-environment correlation. Under artificial selection, it is generally assumed (to a first approximation) that individual fitness is entirely based on the phenotype of the character of interest, specifically those phenotypes chosen by the breeder. In this case, the partial regression coefficients of fitness on genotype and environmental values are zero (again, as first approximation), as phenotype entirely determines fitness. In natural populations, we do not have this luxury and another possibility is that there is no natural selection of the character of interest (its phenotype, by itself has no effect on fitness), but rather selection occurs on characters genetically correlated with the one we are following. If these characters under selection are only connected to the character we are following through genetic value (i.e., no environmental correlation between characters), then \( \beta_A \neq 0 \) while \( \beta_z = \beta_E = 0 \). In this case response becomes

\[
R = \beta_A \sigma_A^2 = S \frac{\sigma_A^2}{\sigma_A^2 + \sigma(E, A)}
\]  

(4.57b)

which reduces to \( R = S \) in the absence of genotype-environment correlations. Another possibility is that the only correlation between characters under selection and our character is through shared environmental effects, giving \( \beta_E \neq 0 \) while \( \beta_A = \beta_E = 0 \), in which case the response becomes

\[
R = \beta_E \sigma(E, A) = S \frac{\sigma(E, A)}{\sigma_E^2 + \sigma(E, A)}
\]  

(4.57c)

which equals zero unless a genotype-environment correlation exists.

**SUMMARY: LIMITATIONS OF THE BREEDERS’ EQUATION**

As many of the previous sections have illustrated, there are number of situations that can compromise the breeders’ equation, even if we are interested only in the single-generation response to selection and can assume a linear parent-offspring regression. One particularly important (and usually unstated) assumption is that we start from an *unselected* base population. If the base population itself has been under selection, decay of transient response components from previous selection compromises the predicted single-generation response. Another troublesome assumption is that our model has accounted for all the selection on the character of
interest. This is especially tricky as selection on any character correlated with the
one of interest can introduce significant bias. The problem of selection of multiple
characters is discussed in Chapters 17–20, but often there is no easy solution, or
even any indication of a problem before the experiment begins. Thus, even in the
best of situations (linearity and no selection-induced changes in allele and gamete
frequencies), there are pitfalls in predicting even a single generation of response
from the slope of the parent-offspring regression. Things get worse if the parent-
offspring regression is nonlinear, where the single-generation change in the mean
can depend on higher order moments of the genotypic distribution, and hence is
not predictable from simple variance components (e.g., Equation 9.35b).

Table 4.1 reviews some of the various factors that can compromise the breed-
ers’ equation, and the chapters in which these complications are examined in
detail. Provided one can assume linearity of the regressions of relatives, many of
these complications can be accounted for. The importance of linearity is that if the
regression of an individual on all its direct relatives selected in previous gener-
ations (back to the original unselected base population) remains linear, response
is entirely determined by the covariances between a current individual and these
previous relatives (Equation 4.50a).

Even if we have corrected for all of the potential complications listed in Table
4.1, the breeders’ equation is expected to be an increasingly poor predictor as
selection proceeds. Even a single generation of selection can significantly change
the underlying variance components, which in turn changes the regression co-
efficients. Further, selection can introduce nonlinearities into an initially linear
regression by transforming the starting distribution away from normality (Chap-
ter 10). In the absence of major genes, allele frequencies changes over the first
few generations of selection are expected to be rather small, but genotype frequen-
cies can change dramatically due to selection generating linkage disequilibrium
(Chapters 7, 10). Directional selection generates negative disequilibrium, decreas-
ing heritability and hence reducing response. This reduction can be significant if
heritability is high. Likewise, selection on the variance itself (through disruptive
or stabilizing selection) also creates disequilibrium which changes the genetic
variance. Chapter 7 examines such short-term changes in disequilibrium. As se-
lection continues over several generations, allele frequencies themselves start to
change, even if all loci have very small effects (Chapter 11). Drift and mutation also
become increasingly important and these complications are examined in Chapters
12, 13.

<table>
<thead>
<tr>
<th>Table 4.1</th>
<th>Summary of various factors that complicate prediction of short-term selection response in the phenotypic mean, even assuming all regressions are linear and that we are considering a single generation of selection from an unselected base population.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>G × E interactions</strong> (Chapters 17, 20)</td>
<td>Possibility of nonlinear parent-offspring regressions.</td>
</tr>
<tr>
<td></td>
<td>Possibility of reversed response.</td>
</tr>
</tbody>
</table>
Age-structure (Chapter 6)  Several generations are required to propagate genetic change uniformly through the population.

Epistasis (Chapter 4)  Component of response due to epistasis is transient. Parent-offspring covar. overestimates permanent response.

Correlated environmental effects (Chapter 4)  Contribution from parent-offspring correlation decays away after selection relaxed.

Maternal effects (Chapter 4)  Complicated lags in response — mean changes unpredictably after selection is relaxed. Possibility of reversed response.

Inbreeding (Chapter 6)  Response depends on additional variance components that are difficult to estimate ($\sigma_{DI}^2$, $\sigma_{ADI}$, etc). Response has permanent and transient components.

Drift (Chapter 8)  Generates variance in the short-term response.

Gametic-phase disequilibrium (Chapter 7)  Changes additive genetic variance. Directional selection generates negative gametic-phase disequilibrium, reducing $h^2$ and slowing response.

Assortative Mating (Chapter 7)  Generates gametic-phase disequilibrium which either enhances (positive correlation between mates) or retards (negative correlation between mates) response.

Selection on Correlated Characters (Chapters 17-20)  Response completely unpredicatable unless selection on correlated characters are accounted for.

Environmental Change (Chapter 5)  A significant change in the environment can obscure the true amount of genetic change.