

Lecture 8

Short-term Response Selection

Bruce Walsh. Aug 2003. Nordic Summer Course

Quantitative geneticists usually distinguish between the short- and long-term response to selection. If we are only trying to predict a few generations of selection response, knowledge of the base population genetic variances (and in particular the heritability) is usually sufficient to make a satisfactory prediction. However, as selection proceeds and allele frequencies change significantly, the initial genetic variances essentially lack any predictive power for the long-term response. The focus in this lecture is on just what predictions we can make for short-term response and some of the complications that can arise.

Changes in the Mean: the Breeder's Equation and its Extensions

It is critical to distinguish between the **within-** and **between-generation** changes induced by selection. The within-generation change is the difference in a population before and after an episode of selection, while the between-generation change (the **response to selection**) is the difference between the population distribution before selection and the distribution of the trait in the next generation (measured at the suitable stage). The response to selection depends not only on the strength of within-generation selection, but also on the fraction of offspring trait value that can be predicted from parental value. If the latter is zero, no matter how strong the within-generation selection is, there will be no response to selection.

The Selection Differential S and Response R

The within-generation change in the mean due to selection is

$$S = \mu_* - \mu \quad (8.1)$$

where μ is the population mean before selection and μ_* the mean of the parents that reproduce (the population mean after selection). S is called the **selection differential**.

The between-generation change, (the response to selection) R , is the change in means between the population before selection and the population in the next generation,

$$R = \mu_o - \mu \quad (8.2)$$

where μ_o is the character mean in the offspring (measured at the same stage as in their parents).

Another useful way to think about the response is in terms of breeding values, as the average deviation of offspring from the population mean is just the mean breeding value of their parents. Hence, the response can also be simply thought of as the net change in breeding value.

The Selection Intensity i

Much akin to the covariance being a poor indicator of the strength of an association, the selection differential S is not particularly informative when trying to compare the strength of selection on different traits and/or in different populations. A much more useful measure is the **selection intensity** i ,

$$i = \frac{S}{\sigma_z} \quad (8.3)$$

which is the differential expressed as in fractions of phenotypic standard deviations.

The Breeders' Equation: Translating S into R

The parent-offspring regression allows us to translate the within-generation change S into the between-generation change R . Recall (Lecture 1) that the predicted value \hat{y} given we know x is

$$\hat{y} = \mu_y + b_{y|x}(x - \mu_x)$$

Here we are trying to predict the offspring value y_O given $x = (P_f + P_m)/2$, the midparent value. Hence, $b_{y|x} = b_{O|MP} = h^2$ is the slope of the midparent-offspring regression, while $\mu_y = \mu_x = \mu$, the mean trait value in the population, giving

$$y_O = \mu + h^2 \left(\frac{P_f + P_m}{2} - \mu \right) \quad (8.4)$$

This regression holds for each midparent-offspring pair. Averaging over all parents, the average difference between the mean (μ_*) of selected parents and the (before selection) population mean is

$$E[(P_f + P_m)/2 - \mu] = \mu_* - \mu = S$$

Likewise, the average value over all the offspring of these selected parents is $E[y_O] = \mu_O$. Thus, averaging over all the midparents gives

$$\mu_O = \mu + h^2 S$$

since $R = \mu_O - \mu$, this gives

$$R = h^2 S \quad (8.5)$$

This relationship is often called the **breeders' equation**, and shows that the heritability of a character is the link between the within-generation change S and the between-generation response R . If $h^2 \simeq 0$, then $R \simeq 0$ no matter how strong the amount of selection is applied.

In some situations, males and females are subjected to different amounts of selection. Recall from lecture 5 (Equation 5.30a) that the regression of offspring value on the value of its sire and dam can also be written as

$$\mu_o = \mu_z + \frac{h^2}{2} (z_s - \mu_s) + \frac{h^2}{2} (z_d - \mu_d) + e$$

giving the expected response as

$$R = \frac{h^2}{2} S_s + \frac{h^2}{2} S_d \quad (8.6)$$

In this case, the breeders equation holds with the selection differential is simply the average differential of both sexes,

$$S = \frac{S_s + S_d}{2}$$

There are several equivalent expressions for the Breeders' Equation. First,

$$R = \frac{\sigma_A^2}{\sigma_z^2} S = \sigma_A \frac{\sigma_A}{\sigma_z} \frac{S}{\sigma_z} = \sigma_A h i \quad (8.7)$$

Alternatively,

$$R = h^2 S \frac{\sigma_z}{\sigma_z} = h^2 \sigma_z i \quad (8.8)$$

While the breeders' equation holds for a single generation of selection from an unselected base population, its validity in predicting response over several generations depends on:

- The reliability of the h^2 estimate
- Absence of environmental change between generations
- The absence of genetic change between the generation in which h^2 was estimated and the generation in which selection is applied.

The later point is critical, as strictly speaking, the prediction equation is true for one generation only, since selection changes gene frequencies and thus h^2 (through changes in the genetic variances). In practice, the breeders' equation is generally valid over several generations.

The Generalized Breeders' Equation: Accuracy

We can extend the breeders' equation to apply to much more general selection schemes beyond simply choosing an individual solely on the basis of its phenotype. To obtain this extension, first note that

$$h^2\sigma_z = \left(\frac{\sigma_A^2}{\sigma_z^2}\right)\sigma_z = \left(\frac{\sigma_A}{\sigma_z}\right)\sigma_A = h\sigma_A$$

Hence, we can rewrite Equation 8.8 as

$$R = ih\sigma_A \quad (8.9a)$$

Recall that h is simply the correlation between an individual's breeding (A) and phenotypic (P) values, $h = \rho_{PA}$. This correlation quantifies the ability to predict the breeding value of an individual from some measure (here that individual's phenotype) and is called the **accuracy** of the selection scheme used to chose parents. We can thus express the breeders equation in terms of the accuracy of selection as

$$R = i\rho_{PA}\sigma_A \quad (8.9b)$$

Hence, the breeders equation can be consider as the following product:

Response = (Intensity)*(Accuracy in Predicting Breeding Value)*(Usable Variance)

More generally, if we use some measure u for predicting the breeding value of an individual, then the breeders' equation can be expressed in terms of the accuracy ρ_{uA} of that measure in predicting breeding value,

$$R = i\rho_{uA}\sigma_A \quad (8.10)$$

Example 8.1. Progeny testing, using the mean of a parent's offspring to predict the parent's breeding value, is an alternative predictor of an individual's breeding value. In this case, the correlation between the mean of n offspring and the breeding value of the parent is

$$\rho_{uA} = \sqrt{\frac{n}{n+a}}, \quad \text{where } a = \frac{4-h^2}{h^2}$$

From Equation 8.10, the response to selection under progeny testing is

$$R = i\sigma_A\sqrt{\frac{n}{n+a}} = i\sigma_A\sqrt{\frac{h^2n}{4+h^2(n-1)}}$$

Note that for very large n that the accuracy approaches one. Progeny testing gives a larger response than simple selection on the phenotypes of the parents (**mass selection**) when

$$\sqrt{\frac{n}{4+h^2(n-1)}} > 1, \quad \text{or } n > \frac{4-h^2}{1-h^2}$$

In particular, $n > 4, 5,$ and $7,$ for $h^2 = 0.1, 0.25,$ and $0.5.$ Also note that the ratio of response for progeny testing (R_{pt}) to mass selection (R_{ms}) is just

$$\frac{R_{pt}}{R_{ms}} = \frac{1}{h}\sqrt{\frac{h^2n}{4+h^2(n-1)}} = \sqrt{\frac{n}{4+h^2(n-1)}}$$

The Generalized Breeders' Equation: Generation Intervals

So far, we have been assuming non-overlapping generations — all parents only reproduce in one generation interval. Of course, in most settings with domesticated animals, they live multiple years and can have progeny over different years. In such cases, the response should be expressed in terms of response per year. To express the breeders' equation in terms of response per year, we first need to compute the **generation intervals** L_x (the average age of parents when progeny are born) for both sexes.

Example 8.2. Compute L_s and L_d for the following age structure:

		Age of Birth of Progeny				
Sires		year 2	year 3	year 4	year 5	Total
Number		60	30	0	0	90
Dams		year 2	year 3	year 4	year 5	Total
Number		400	600	100	40	1140

$$L_s = \frac{2 \cdot 60 + 3 \cdot 30}{60 + 30} = 2.33, \quad L_d = \frac{2 \cdot 400 + 3 \cdot 600 + 4 \cdot 100 + 5 \cdot 40}{400 + 600 + 100 + 40} = 2.81$$

Incorporating the generation intervals, the yearly rate of response can be expressed as

$$R_y = \left(\frac{i_s + i_d}{L_s + L_d} \right) h^2 \sigma_p = \left(\frac{i_s + i_d}{L_s + L_d} \right) h \sigma_A \quad (8.11a)$$

Thus, one way to increase response is to reduce the generation intervals, for example by having younger parents. The problem is that there is a tradeoff between generation interval and selection intensity. In species that are reproductively-limited (few offspring per dam), using younger dams means that a higher fraction of the dams must be chosen to replace the population (i.e., to keep the same number of animals in a herd). As a consequence, the selection intensity on these parents (which increases as fewer parents are chosen) is reduced.

More generally, we can combine both selection accuracy and generation interval to give a more general version of the breeders' equation:

$$R_y = \left(\frac{i_s + i_d}{L_s + L_d} \right) \rho_{uA} \sigma_A \quad (8.11b)$$

Expressed this way, there are three components of response that the breeder has some control over: selection intensity, generation interval, and selection accuracy (not much can be done with increasing σ_A^2). Response is increased by decreasing L and increasing ρ and i . We have already discussed tradeoffs between L and i , and there are similar tradeoffs between L and ρ . Clearly, the longer we wait to allow a parent to reproduce, the more accurate we can predict their breeding value, as information from other relatives and from progeny-testing accumulates over time. However, these increases in ρ also result in increases in L . The optimal selection program must balance all of these competing interests.

Equation 8.11b also highlights the importance to animal breeding of advances in reproductive technology such as artificial insemination (AI) and multiple ovulation embryo transplant (MOET) schemes. The more offspring a parent can produce, the stronger a selection intensity we can apply (and still keep a required fixed number of animals in our herd). Hence, AI has resulted in the potential for far greater selection intensities (and unfortunately far more inbreeding) than would be possible under natural insemination. Likewise, MOET schemes to increase the number of offspring from females potentially allow for increases in the selection intensity on dams as well as decreases in the generation interval.

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 8.1). 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.

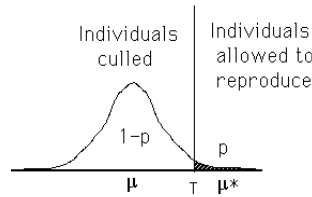


Figure 8.1. Under truncation selection, the uppermost (or lowermost) fraction p of a population is selected to reproduce. Alternatively, one could set a threshold level in advance, above (below) which individuals are allowed to reproduce. To predict response given either p or T , we need to know the mean of the selected tail (μ^*), 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 μ and variance σ^2 , this conditional mean is given in Lecture 1 (Equation 1.12a), which gives the expected selection differential as

$$S = \varphi \left(\frac{T - \mu}{\sigma} \right) \frac{\sigma}{p} \tag{8.12a}$$

where p is the fraction saved and $\varphi(x) = (2\pi)^{-1/2} e^{-x^2/2}$ is the unit normal density function evaluated at x . Hence, the (within-generation) mean after selection is just

$$\mu_* = \mu + S = \mu + \varphi \left(\frac{T - \mu}{\sigma} \right) \frac{\sigma}{p} \tag{8.12b}$$

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 8.12a, we must first find the threshold value T_p satisfying $\Pr(z > T_p) = p$. Notice that T in Equation 8.12a 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 **probit transformation** of p , by

$$\Pr(U \leq z_{[p]}) = p \tag{8.13a}$$

Hence

$$\Pr(U > z_{[1-p]}) = p \quad (8.13b)$$

It immediately follows that $z_{[1-p]} = (T_p - \mu)/\sigma$, and Equation 8.12a gives the expected selection intensity as

$$i = \frac{S}{\sigma} = \frac{\varphi(z_{[1-p]})}{p} \quad (8.14a)$$

One can obtain $z_{[1-p]}$ from normal distribution tables. Alternatively, a number of approximations have been suggested for Equation 8.14a. Assuming normality, Smith (1969) suggests

$$i \simeq 0.8 + 0.41 \ln\left(\frac{1}{p} - 1\right) \quad (8.14b)$$

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

$$i \simeq \frac{2.97425 - 3.38197 p^{0.2} - 1.9319 p^{0.4} + 2.3097 p^{0.6}}{0.51953 + 0.88768 p^{0.2} - 2.38388 p^{0.4} + p^{0.6}} \quad (8.14c)$$

Example 8.3. 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,

$$i = \frac{\varphi(1.645)}{0.05} = \frac{0.103}{0.05} \simeq 2.06$$

Applying Equation 8.8 gives the expected response to this amount of selection as $R = h^2 \sigma 2.06$. Smith's approximation gives the selection intensity as

$$i \simeq 0.8 + 0.41 \ln\left(\frac{1}{0.05} - 1\right) \simeq 2.01$$

which is quite reasonable. Saxton's approximation gives $i \simeq 2.06$.

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 $z_{1,M} \geq z_{2,M} \dots \geq z_{M,M}$ 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,

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

where $z'_{k,M} = (z_{k,M} - \mu)/\sigma$ are the **standardized order statistics**. Figure 8.2 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 8.14a overestimating the actual selection intensity, although the difference is small unless N is very small.

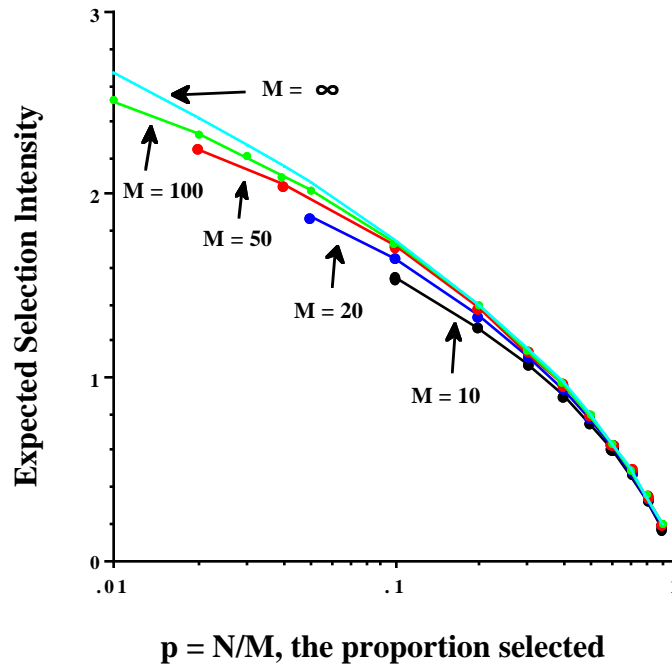


Figure 8.2. The expected selection intensity $E(i)$ under truncation selection with normally-distributed phenotypes, as a function of the total number of individuals measured M and the fraction of these saved $p = N/M$, N being the number of these adults allowed to reproduce. The curve $M = \infty$ is given by using Equation 8.14a, 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. Note that Equation 8.14a 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., $\text{Pr}(y \geq y_p) = p$), Burrows' approximation is

$$E(i_{(M,N)}) \simeq \mu_{y_p} - \frac{(M - N)p}{2N(M + 1)\phi(y_p)} \quad (8.15a)$$

where

$$\mu_{y_p} = E(y | y \geq y_p) = \frac{1}{p} \int_{y_p}^{\infty} x \phi(x) dx$$

is the truncated mean, which can be obtained by numerical integration. Since the second term of Equation 8.15a is positive, if M is finite the expected truncated mean overestimates the expected standardized selection differential.

For a normal distribution,

$$E(i_{(M,N)}) \simeq i - \left[\frac{M - N}{2N(M + 1)} \right] \frac{1}{i} = i - \left[\frac{1 - p}{2p(M + 1)} \right] \frac{1}{i} \quad (8.15b)$$

where i is given by Equation 8.14a. Bulmer (1980) suggests an alternative finite-sample approximation for $E(i_{(M,N)})$ when phenotypes are normally distributed, using Equation 8.14a with p replaced by

$$\tilde{p} = \frac{N + 1/2}{M + N/(2M)} \quad (8.16)$$

Example 8.4. 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 8.3, the expected selection intensity is $i \simeq 2.06$. 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 $E(i_{(20,1)}) \simeq 1.87$. There is considerable spread about this expected value, as the standard deviation of this order statistic is 0.525. How well do the approximations of $E(i_{(20,1)})$ perform? Burrow's approximation gives

$$E(i_{(20,1)}) \simeq 2.06 - \frac{(20 - 1)}{2(20 + 1)2.06} = 2.06 - 0.22 = 1.84$$

while Bulmer's approximation uses

$$\tilde{p} = \frac{1 + 1/2}{20 + 1/40} \simeq 0.075$$

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

$$E(i_{(20,1)}) \simeq 0.1415/0.075 \simeq 1.89$$

A final correction for finite population size was noted by Rawlings (1976) and (especially) Hill (1976, 1977). If families are sampled, such that n individuals are chosen per family, then the selection intensity is further reduced because the correlations between family members. In particular, if a total of M individuals are sampled, with n individuals per family then Burrow's correction (Equation 8.15b) is modified to become

$$i - \left[\frac{1 - p}{2p(M + 1)(1 - \tau + \tau/n)} \right] \frac{1}{i} \quad (8.17)$$

where τ is the intra-class correlation of family members.

Selection on Threshold Traits

Discrete characters can often be modeled by assuming an underlying continuous character z (usually referred to as the the **liability**) that maps to the observed discrete character states (Figure 8.3). Consider the simplest case where the character is either present ($z \geq T$) or absent ($z < T$). Let μ_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 μ_t from the observed frequencies of character states.

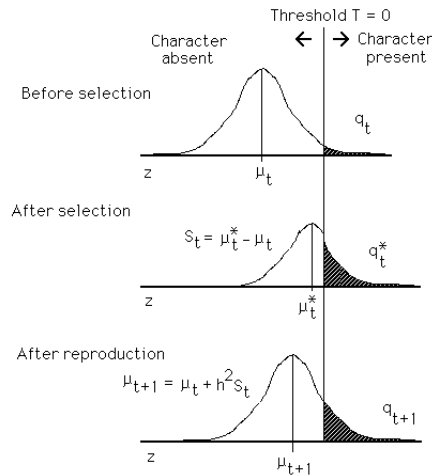


Figure 8.3. Selection on a character with a single threshold T . z is the value on the underlying scale of liability. Assuming z is a well-behaved quantitative character, $\mu_{t+1} = \mu_t + S_t h^2$, where $S_t = \mu_t^* - \mu_t$. Using the probit transform (Equation 8.13) we can translate from q_t , the frequency of individuals displaying the character, to μ_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 hence (Equation 8.13b)

$$\mu_t = -z_{[1-q_t]} \quad (8.18)$$

where $z_{[p]}$, the probit transformation of p (Equation 8.13). For example, if 5% of a large population displays the trait, $\mu = -z_{[0.95]}$. From normal probability tables, $\Pr(U < 1.645) = 0.95$, hence $z_{[0.95]} = 1.645$ and $\mu = -1.645$. For small samples, estimation of μ requires the use of order statistics.

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

$$\begin{aligned} q_{t+1} &= \Pr(U \geq -\mu_{t+1}) \\ &= \Pr(U \geq -\mu_t - h^2 S_t) \\ &= \Pr(U \geq z_{[1-q_t]} - h^2 S_t) \end{aligned} \quad (8.19)$$

It remains to obtain $S_t = \mu_t^* - \mu_t$, where μ_t^* 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,

$$\mu_t^* = (1 - p_t) E(z|z < 0; \mu_t) + p_t E(z|z \geq 0; \mu_t) \quad (8.20)$$

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 8.20 gives

$$S_t = \mu_t^* - \mu_t = \frac{\varphi(\mu_t)}{q_t} \frac{p_t - q_t}{1 - q_t} \quad (8.21)$$

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 8.21 reduces to $S_t = \varphi(\mu_t)/q_t$.

Example 8.5. 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. Equation 8.21 gives the selection differential as

$$S_0 = \frac{\varphi(-1.645)}{0.05} \simeq \frac{0.106}{0.05} \simeq 2.064$$

giving

$$\mu_1 = \mu_0 + h^2 S = -1.645 + 0.25 \cdot 2.062 = -1.129$$

and hence (Equation 8.18)

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

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 8.4.

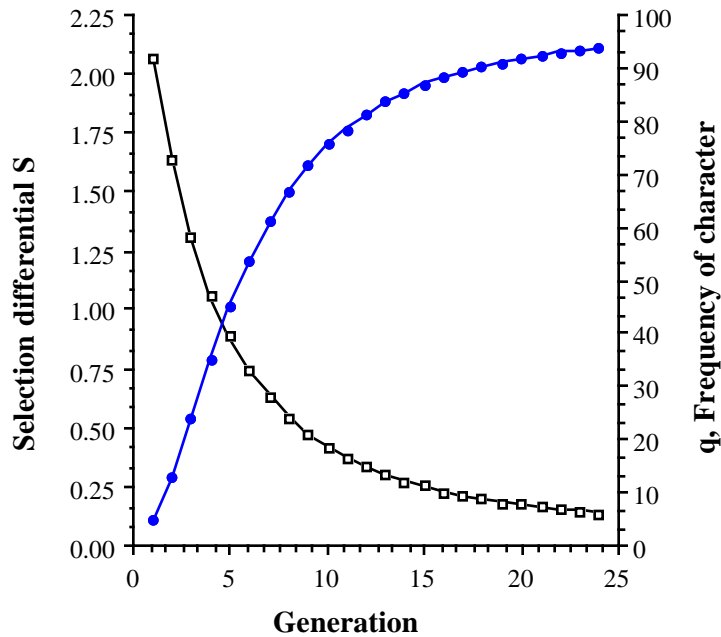


Figure 8.4. Response to selection on a threshold character. The initial frequency of individuals displaying the trait is 5%. Selection occurs by only allowing adults displaying the trait to reproduce ($p_t = 1$). Solid circles denote q_t , open squares denote S_t . The liability is assumed to have $h^2 = 0.25$.

Permanent Versus Transient Response

The slope of the parent-offspring regression can be inflated by epistasis, genotype \times 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,

$$b_{op} = \frac{h^2}{2} + \frac{1}{\sigma_z^2} \left(\frac{\sigma_{AA}^2}{4} + \frac{\sigma_{AAA}^2}{8} + \frac{\sigma_{AAAA}^2}{16} + \dots + \sigma(E_p, E_o) \right) \quad (8.22a)$$

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

$$R = h^2 S + \frac{S}{\sigma_z^2} \left(\frac{\sigma_{AA}^2}{2} + \frac{\sigma_{AAA}^2}{4} + \frac{\sigma_{AAAA}^2}{8} + \dots + \sigma(E_{fa}, E_o) + \sigma(E_{mo}, E_o) \right) \quad (8.22b)$$

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 lecture 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 8.22 is $h^2 S$. One exception is when significant inbreeding occurs. In this case, σ_{AA}^2 and other non-additive variance components can contribute to permanent response.

Response with Epistasis

The response when additive \times additive epistatic variance is present was examined by Griffing (1960a,b) for the infinitesimal model. 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

$$R = S \left(h^2 + \frac{\sigma_{AA}^2}{2\sigma_z^2} \right) \quad (8.23)$$

One might expect that $R(n)$, the cumulative response after n generations of selection, is simply n times the result given by Equation 8.23. 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 τ generations of no selection is

$$S \left(h^2 + (1 - c)^\tau \frac{\sigma_{AA}^2}{2\sigma_z^2} \right) \quad (8.24)$$

which converges to $h^2 S$. Equation 8.24 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 τ generations only $(1 - c)^\tau$ of the favorable two-locus combinations selected at $\tau = 0$ remain unaltered by recombination.

Summing Equation 8.24 over t gives the cumulative response after t generations with constant selection differential S as

$$R(t) = t h^2 S + R_{AA}(t) \quad (8.25)$$

where $R_{AA}(t)$ denotes the cumulative additive \times additive epistatic contribution. With t generations of selection followed by τ generations of no selection, the cumulative response is

$$t h^2 S + (1 - c)^\tau R_{AA}(t) \quad (8.26)$$

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

The presence of epistasis can result in a curvilinear selection response if σ_{AA}^2/σ_z^2 is sufficiently large. However, 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 $h^2 S$ 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 σ_{AA}^2 and higher order epistatic variances.

Maternal Effects: Response Under Falconer's Model

Maternal effects are another potential complication for the breeders' equation. They can result in apparent **reversed responses** wherein the mean embarrassingly changes in the opposite direction of that predicted from the breeders' equation. They can also result in other unusual dynamics such as time lags.

The simplest model of maternal effects (motivated by the inheritance of litter size in mice) is that of Falconer (1965). Falconer assumes the maternal contribution is a linear function of the maternal phenotype z_{mo} , so that $M = mz_{mo}$ and the phenotypic decomposition becomes

$$z = G + m z_{mo} + e \quad (8.27)$$

Conceivably, M could be a nonlinear function of z_{mo} , but linearity is assumed for tractability. We refer to Equation 8.27 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. Negative estimates of m have been reported, which as we will shortly see can (in some cases) lead to a reversed response. For example, 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.

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

$$E(z_o | A_{mo}, A_{fa}, z_{mo}) = \frac{A_{mo}}{2} + \frac{A_{fa}}{2} + m z_{mo} \quad (8.28a)$$

where A_{mo} and A_{fa} are the maternal and paternal breeding values. Averaging over the selected parents, the mean in generation $t + 1$ is

$$\mu_z(t+1) = \frac{A_{fa}^*(t) + A_{mo}^*(t)}{2} + m \mu_{mo}^*(t) \quad (8.28b)$$

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

$$\begin{aligned} E_s(A_{mo}) &= E_s \left(\mu_A(t) + b_{Az} [z_{mo} - \mu_z(t)] + e \right) \\ &= \mu_A(t) + b_{Az} S_{mo}(t) \end{aligned} \quad (8.29)$$

where $E_s(\cdot)$ 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 . The resulting regression slope (at equilibrium) is

$$b_{Az} = h^2 \frac{2}{2-m} \quad (8.30)$$

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

$$\Delta\mu_z(1) = S_{mo}(1) \left(\frac{h^2}{2-m} + m \right) + S_{fa}(1) \frac{h^2}{2-m} \quad (8.31)$$

An interesting consequence of Equation 8.28 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

$$m < 1 - \sqrt{1 + 2h^2} \quad (8.32a)$$

If selection is only occurring on females, this condition is

$$m < 1 - \sqrt{1 + h^2} \quad (8.32b)$$

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 8.32b implies that a reversed response in the first generation is expected (as $1 - \sqrt{1 + 0.11} \approx -0.05 > m$).

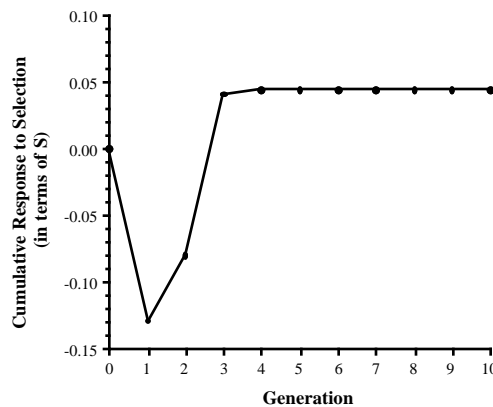


Figure 8.5. Prediction from the model, using Falconer's estimated values of $h^2 = 0.11$ and $m = -0.13$. The predicted change in population mean following a single generation of selection on females with $S_{mo} > 0$ is plotted. There is a reversed response in the first generation, even though the net 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 $S_{mo} h^2 / [(1-m)(2-m)] = 0.044S_{mo}$.

Gene Frequency Changes Under Selection

Next, we turn to population-genetic considerations of the expected allele frequency changes at loci underlying a quantitative trait.

How quickly does selection change the frequency of alleles at loci contributing to a trait under selection? We start by reviewing a few results from population genetics. Consider a diallelic locus, with alleles A_1 and A_2 , whose genotypes have the following relative fitnesses:

Genotype	A_1A_1	A_1A_2	A_2A_2
Fitness	1	$1 + s$	$1 + 2s$

This is an example of **additive fitness**. With these fitnesses, for every offspring left by an individual with an A_1A_1 genotype, $1 + 2s$ offspring are left (on average) by individuals with an A_2A_2 genotype. If q represents the frequency of allele A_2 before selection, then the change in the frequency of q after selection is given by

$$\Delta q = \frac{sq(1 - q)}{1 + 2sq} \simeq sq(1 - q) \quad \text{when } |2sq| \ll 1 \quad (8.33)$$

Thus, under these fitnesses, the change in the frequency of the favorable allele is proportional to s . In finite populations, genetic drift can overpower the effects of selection. In particular, when

$$4N_e |s| \ll 1$$

the fate of an allele is largely determined by gene drift, rather than selection. In such cases, favorable alleles can easily be lost by drift.

Now consider a locus contributing to a character z under selection. Suppose the genotypes at this locus make the following contribution to the character:

Genotype	A_1A_1	A_1A_2	A_2A_2
Contribution	0	a	$2a$

For a trait with phenotypic variation σ_z^2 under selection intensity i , this induces additive fitnesses on these genotypes, with

$$s \simeq \frac{a}{\sigma_z} i \quad (8.34)$$

Hence, the change in allele frequency depends on both the strength of selection i and the relative contribution a/σ_z of the character to the overall trait value. As expected, loci with larger contributions are under stronger selection than loci with minor contributions and hence have faster allele frequency changes. Further note that if

$$4N_e |s| = \frac{4N_e |a i|}{\sigma_z} \ll 1 \quad (8.35)$$

then the effect of selection on this locus is weaker than the effects of drift. Thus, many favorable QTL alleles can be lost by drift if either their effects (a/σ_z), the strength of selection on the character (i), or the effective population size (N_e) are sufficiently small.

More generally, if the locus shows dominance towards the character, the fitnesses become

Genotype	A_1A_1	A_1A_2	A_2A_2
Contribution	0	$a(1 + k)$	$2a$
Induced fitness	1	$1 + s(1 + h)$	$1 + 2s$

where for the induced fitnesses $s = ai/\sigma_z$ (as above) and $h = k$.

Short-term Changes in the Variance

Selection has two routes by which to change the genetic variances, and hence the heritability and selection response. First, it can change the frequencies at individual alleles. When the contribution to a trait from any locus is very small, these selection-induced changes in allele frequencies over a few generations are also very small. However, selection also creates correlations between alleles at different loci (**linkage disequilibrium**), and this can result in an immediate change in the variance.

Consider the within-generation change in the variance, $\delta\sigma_z^2 = \sigma_{z^*}^2 - \sigma_z^2$. Using regression arguments similar to those leading to the breeders' equation, the expected response in the variance to a single generation of selection is

$$d = \sigma_O^2 - \sigma_P^2 = \frac{h^4}{2} \delta\sigma_z^2 \quad (8.36)$$

where σ_O^2 is the variance in the offspring and σ_P^2 the variance in the unselected population. This is the variance response analog to the response in mean (the breeders' equation), with $h^4/2$ replacing h^2 and $\delta\sigma_z^2$ replacing S . In many situations (such as truncation selection), we can write

$$\sigma_{z^*}^2 = (1 - k)\sigma_z^2$$

so that the result of selection is a proportional change in the variance.

It turns out that all the change in the variance is due to a change in the additive genetic variance, so that if V_a denotes the additive variance before selection, then after one generation of selection

$$V_A(1) = V_a + d, \quad V_P(1) = V_A(1) + V_D + V_E = V_P + d \quad (8.37)$$

where V_P is the phenotypic variance in the base (pre-selection) population. The heritability thus becomes

$$h^2(1) = \frac{V_A(1)}{V_P(1)} = \frac{V_a + d}{V_P + d}$$

Truncation selection reduces the variance ($\delta\sigma_z < 0$), which results in reduced additive genetic variance and heritability in the next generation, slowing response. This reduction in variance due to selection creating linkage disequilibrium is referred to as the **Bulmer effect**, after Michael Bulmer's pioneering work on this subject in the 1970's.

One subtle feature of changes in the variance is that recombination breaks down the selection-induced correlations, so that in the absence of selection, $d(t+1) = d(t)/2$. Hence, one must iterate to obtain the value of the variance in generation t . Starting with an unselected base population, $d(0) = 0$, we obtain the value for $d(t+1)$ by iterating

$$\begin{aligned} d(t+1) &= \frac{d(t)}{2} + \frac{h^4(t)}{2} \delta\sigma_{z(t)} \\ &= \frac{d(t)}{2} - k \frac{h^4(t)}{2} \sigma_z^2(t) \end{aligned} \quad (8.38)$$

The first term ($d/2$) is the decay in linkage disequilibrium from recombination while the second term is the amount of new disequilibrium created by selection. Note for above that

$$\sigma_z^2(t) = \sigma_z^2(0) + d(t), \quad \text{and} \quad h^2(t) = \frac{V_A(t)}{V_P(t)} = \frac{V_a + d(t)}{V_P + d(t)} \quad (8.39)$$

While all this looks rather complicated at first glance, it's really a very straight forward series of substitutions. The net result for directional selection is that most of the reduction in variance occurs over the first few generations, which rapidly approaches an equilibrium value (the equilibrium reduction in the additive variance). However, under disruptive selection (selection to increase the variance, for example by selecting both the largest and smallest parents), the variance may continue to increase substantially over many generations before settling on its equilibrium value.

Artificial Selection Problems

1. Taking the selection differential as the difference between the means of selected parents and the mean before selection makes the assumption that each selected parent contributes equally to the next generation. Biases introduced by differential fertility can be removed by using **effective selection differentials**, S_e ,

$$S_e = \frac{1}{n_p} \sum_{i=1}^{n_p} \left(\frac{n_i}{\bar{n}} \right) (z_i - \mu_z) = \left(\frac{1}{n_p} \sum_{i=1}^{n_p} \left(\frac{n_i}{\bar{n}} \right) z_i \right) - \mu_z$$

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, \bar{n} the average number of offspring for selected parents, and μ_z is the mean before selection. If all selected parents have the same number of offspring ($n_i = \bar{n}$ 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**.

Suppose 5 parents are selected, with the following trait values and offspring number:

Parent	phenotypic value	number of offspring
1	45	1
2	40	2
3	35	3
4	33	5
5	32	5

If the mean before selection is 30, compute the S and S_e . If $h^2 = 0.3$, what is the expected response that would be estimated under the two differentials?

2. Consider a population not currently under selection, with $\sigma_z^2 = 100$ and $h^2 = 0.5$ and $d(0) = 0$ (no disequilibrium). Consider two types of selection (i) stabilizing where $\sigma_{z^*}^2 = 0.5\sigma_z^2$ (i.e., $k = 1/2$) and (ii) disruptive selection $\sigma_{z^*}^2 = 1.5\sigma_z^2$ ($k = -1/2$). For both types of selection compute $d(1)$ and $d(2)$, $\sigma_A^2(1)$ and $\sigma_A^2(2)$, $\sigma_z^2(1)$ and $\sigma_z^2(2)$, and $h^2(1)$ and $h^2(2)$.

Solutions to Chapter 8 Problems

1. Here $\mu_* = 37$, giving $S = 7$, while $\bar{n} = 3.2$ and

i	z_i	n_i	n_i/\bar{n}	$z_i \cdot n_i/\bar{n}$
1	45	1	0.3125	14.06
2	40	2	0.6250	25.00
3	35	3	0.9375	32.81
4	33	5	1.563	51.56
5	32	5	1.563	50.0

$$\frac{1}{n_p} \sum_{i=1}^{n_p} \left(\frac{n_i}{\bar{n}} \right) z_i = 34.69$$

Giving $S_e = 4.69$. Assuming $h^2 = 0.3$, using the uncorrected S gives a response of $R = 0.3 \cdot 7 = 2.1$, while the true expected response if $R = 0.3 \cdot 4.69 = 1.4$

2. Here $\sigma_a^2 = h^2 \sigma_z^2 = 50$, and $d(0) = 0$

$$d(1) = d(0) - k(h^4/2)\sigma_z^2(0) = \begin{cases} 0 - 0.5 * 0.125 * 100 = -6.25 & \text{for stabilizing, } k = 0.5 \\ 0 + 0.5 * 0.125 * 100 = 6.25 & \text{for disruptive, } k = -0.5 \end{cases}$$

$$\sigma_A^2(1) = \sigma_a^2 + d(1) = \begin{cases} 43.75 & \text{for stabilizing} \\ 56.25 & \text{for disruptive} \end{cases}, \quad \sigma_z^2(1) = \sigma_z^2 + d(1) = \begin{cases} 93.75 & \text{for stabilizing} \\ 106.25 & \text{for disruptive} \end{cases}$$

$$h^2(1) = \sigma_A^2(1)/\sigma_z^2(1) = \begin{cases} 0.467 & \text{for stabilizing} \\ 0.529 & \text{for disruptive} \end{cases}$$

$$d(2) = d(1)/2 - k(h^4(1)/2)\sigma_z^2(1) = \begin{cases} -6.25/2 - 0.5(0.46^2/2) * 93.75 = -8.02 & \text{for stabilizing} \\ 6.25/2 + 0.5(0.53^2/2) * 106.25 = 10.59 & \text{for disruptive} \end{cases}$$

$$\sigma_A^2(2) = \sigma_a^2 + d(2) = \begin{cases} 41.77 & \text{for stabilizing} \\ 60.57 & \text{for disruptive} \end{cases}, \quad \sigma_z^2(2) = \sigma_z^2 + d(2) = \begin{cases} 91.77 & \text{for stabilizing} \\ 110.6 & \text{for disruptive} \end{cases}$$

$$h^2(2) = \sigma_A^2(2)/\sigma_z^2(2) = \begin{cases} 0.455 & \text{for stabilizing} \\ 0.548 & \text{for disruptive} \end{cases}$$