24

The Infinitesimal Model and Its Extensions

Normal theory is clearly the most powerful and problematic hypothesis in the present analysis. — Chevalet (1988)

What, me normal? — Turelli and Barton (1994)

Draft Version 18 July 2008

The joint assumptions of normality and linear parent-offspring regressions underlie most simple models of selection response, as in such cases the single-generation response can be predicted from knowledge of the appropriate variance components (Chapters 10, 13). This is in sharp contrast to response under the one- and two-locus models previously examined (Chapter 5) wherein prediction requires detailed knowledge of the underlying genotype frequencies and effects.

The **infinitesimal model**, which assumes a very large (effectively infinite) number of loci each with infinitesimal effect, satisfies both normality and linearity. Hence, most models of (short-term) selection response are based on the (either explicit or implicit) assumption that the infinitesimal model adequately describes the underlying dynamics. Of course, the infinitesimal model is not taken as an exact description of biological reality. It does, however, represent one extreme of assumptions about the underlying loci, allowing us to ignore the effects of allele frequency changes. When a large number of loci, each of small effect, underlie a character, the infinitesimal model often provides a very satisfactory treatment of short-term response.

Selection (and drift) compromise predictions of selection response by changing allele frequencies and generating disequilibrium. Predicting changes in allele frequencies are especially problematic as to do so requires very intimate knowledge of the underlying genetical details, such as the effects and frequencies of all alleles. As these are essentially unobservable for all but the most trivial cases (those rare one- or two-locus traits), these are often referred to as the **microscopic** (unobservable) parameters of the system. Ideally, we would like to have **macroscopic**-based (observable) predictors of response, based on easily-measured quantities, such as genetic variances. The breeders' equation (e.g., Equations 10.1 and 10.21), and the Bulmer equation (Equation 13.12) are examples of macroscopic-based predictors of response.

The goal of this chapter, which is rather technical in places, is to examine short-term selection response when the infinitesimal model fails. Basically there are four features we need to consider. The first two are closely connected: the joint assumptions that (i) the parent-offspring regression is linear and homoscedastic (the error variance is independent of the trait values of the parents) and (ii) the joint-distribution of breeding values between parent and offspring is multivariate normal. Since (i) is satisfied when (ii) occurs, this is the typical assumption made, and the strict infinitesimal model guarantees this, as we discuss below. The third feature is allele frequency change. While this does not occur when the number of loci and the population size are both infinite, either a finite number of loci and/or finite population size cause allele frequencies to change, and we need to account for these. Changes in allele frequencies can also cause departures from normality. Finally, selection generates linkage-disequilibrium, and this, too, can cause departures from normal-

ity. Further, allele frequencies changes and generation of disequilibrium change the genetic variances (and more generally other higher-order parameters of the genotypic distribution), further compromising predictions based on their pre-selection values.

We start by reviewing some of the basic properties of the infinitesimal model. We then introduce a class of finite locus models that have a connection with the infinitesimal model. These continuum-of-alleles models also make gaussian assumptions, namely that the distribution of allelic effects at *each locus* is normal. At the limit, these models recover the infinitesimal results. We then examine the effects of linkage on these models, and conclude by examining selection response when the distribution of allelic effects is no longer gaussian. The focus of this chapter is to start to bridge between very short term (few generations) predictors of response based on macroscopic parameters and long term response based on microscopic parameters. After sufficiently allele frequency change accrues, these bridging models break down and explicit population genetic models (Chapter 5) are required. We examine such models for long-term response in detail in Chapters 25 and 26.

THE INFINITESIMAL MODEL

Under the classic infinitesimal model, introduced by Fisher (1918), a character is determined by an infinite number of unlinked and nonepistatic loci, each with an infinitesimal effect. It is often assumed that each locus has two alleles and the effects and frequencies are the same (or very similar) across all loci, but we can (somewhat) relax these constraints. Here we examine some of the properties that result from these assumptions. This will serve as a starting point for relaxing these assumptions and examining their consequences for predicting short-term response.

Allele Frequencies Do Not Change Under the Infinitesimal Model

Recall from Chapter 13 that we can express the additive *genetic* variance σ_A^2 as the *genic* variance σ_a^2 plus the disequilibrium contribution d, $\sigma_A^2 = \sigma_a^2 + d$. Changes in either change the additive variance, and this partition decouples the effect of allele frequency change (changes in σ_a^2) from the effect of changes in linkage disequilibrium (d).

Under the infinitesimal model, *allele frequencies are unchanged by selection*, and thus σ_a^2 is assumed constant over time. Large changes in the mean occur by summing infinitesimal allele frequency changes at a large number of loci. To see this, consider a character determined by n completely additive diallelic loci. Further suppose that all loci are **interchangeable**, with each locus having the same effects and frequencies (this is also called the **exchangeable model**). In particular, assume each locus has two alleles, Q and q, with the genotypes QQ, Qq, and qq contributing $2a$, a, and 0 (respectively) to the genotypic value, so that allele Q has effect a. Further, assume the frequency of allele Q has the same value (p) at each locus. The resulting the mean is $2nap$ and the additive variance (ignoring the contribution from gametic-phase disequilibrium) is $\sigma_A^2 = \sigma_a^2 = 2na^2p(1-p)$. For σ_A^2 to remained bounded as the number of loci increase, a must be of order $n^{-1/2}$. The change in mean due to a single generation of selection is easily found to be $\Delta\mu = 2na\Delta p$. Assuming the frequency of Q changes by the same amount at each locus, $\Delta p = \Delta\mu/(2na)$. Since a is of order $n^{-1/2}$, Δp is of order $1/(n \cdot n^{-1/2}) = n^{-1/2}$, approaching zero as the number of loci becomes infinite. Thus the infinitesimal model allows for arbitrary changes in the mean with (essentially) no change in the allele frequencies at underlying loci. Biologically (i.e., with a finite number of loci), the infinitesimal model implies that large changes in the mean of a trait can occur with only small to modest changes in allele frequencies if all loci each make only a small contribution to the trait.

What effect does this amount of allele frequency change have on the variance? Letting

$p' = p + \Delta p$ denote the frequency after selection, the change in the additive genic variance is

$$\begin{aligned}\Delta\sigma_a^2 &= 2na^2p'(1-p') - 2na^2p(1-p)\\ &= 2na^2\Delta p(1-2p-\Delta p)\\ &\approx a\,(1-2p)\,\Delta\mu\end{aligned}$$

Since a is of order $n^{-1/2}$, the change in variance due to changes in allele frequencies is roughly $1/\sqrt{n}$ the change in mean. With a large number of loci, very large changes in the mean can occur without any significant change in the genic variance. The more loci of equal effect underlying a trait, the slower the change in σ_a^2 and hence the longer the response is predictable. In the limit of an infinite number of loci, there is no change in the genic variance ($\Delta\sigma_a^2 = 0$), while arbitrary changes in the mean can occur.

Disequilibrium Under the Infinitesimal Model

Allele frequency change (or lack thereof) is not the whole story for the infinitesimal model, as even with a constant genic variance, changes in d can significantly change σ_A^2. The reason for this can been seen from Equation 13.1. Changes in the covariances C_{ij} between loci i and j (for $i \neq j$) are roughly of order n^{-2} (Bulmer 1980, Turelli and Barton 1990). Since there are n^2 terms contributing to d, the total disequilibrium is of order one ($n^2 \cdot n^{-2}$) and does not necessarily approach zero as the number of loci becomes infinite. The same reasoning holds for changes in the higher-order moments, which are caused by higher-order associations between groups of loci. Indeed, for the k-th order moment there are n^k terms in the sum, each scaling as n^{-k} to potentially give a non-zero value in the limit (Turelli and Barton 1990).

Dominance

Dominance is certainly not excluded under an infinitesimal models, but there is a potential tradeoff in scaling effects so as to bound the dominance variance (on one hand) and bounding any inbreeding depression on the other. To see this, suppose we have n diallelic loci with no epistasis (the total genotypic value is simply the sum of the individual locus genotypic values), and let the genotypic values at locus i be $0 : a_i + d_i : 2a_i$ where the frequency of the increasing allele is p_i. The resulting dominance variance becomes

$$\sigma_D^2 = \sum_i^n \left(2p_i(1-p_i)d_i\right)^2$$

For n loci of equal effect,

$$\sigma_D^2 = 4np^2(1-p)^2d^2$$

For σ_D^2 to remain bounded as $n \to \infty$, d must be order $n^{-1/2}$, the same as we found for a. Thus, if both a and d scale as $1/\sqrt{n}$, the additive and dominance variances remained bounded as the number of locus goes to infinity.

Now consider the behavior of inbreeding depression, the difference between the mean trait value μ_f when population-level inbreeding is f versus that under random mating μ_0 (LW Chapter 10). Again, assuming no epistasis, from LW Equation 10.3, the inbreeding depression is given by

$$\mu_f - \mu_0 = -2f\sum_i^n p_i(1-p_i)d_i$$

Assuming n loci of equal effect gives

$$\mu_f - \mu_0 = -2nfp(1-p)d$$

Note that if d scales as $n^{-1/2}$, the amount of inbreeding depression scales as $n \cdot n^{-1/2} = n^{1/2}$ and hence goes to infinity. Conversely, if we scale d as order $1/n$, we have bounded inbreeding depression, but the dominance variance is now of order $n/n^2 = 1/n$ and hence is zero in the infinitesimal limit. Thus, under the exchangeable infinitesimal model, one cannot have both a bounded dominance variance and a bounded inbreeding depression. Of course, the flaw in this argument is our assumption of equal effects over all loci. In this case, all of the d have the same sign. If we assume $E[\,d\,] = 0$, i.e., no directional dominance, then we can have a bounded dominance variance, but no inbreeding depression. To have both a dominance variance and inbreeding depression in the infinitesimal limit requires a great deal of delicacy, in that individual effects have to be scaled so that $E[\,d\,] > 0$ (finite directional dominance).

Gaussian Features of the Infinitesimal

The **central limit theorem** from probability theory — sums of random variables typically converge to a normal, or Gaussian, distribution — gives the key feature that, under the infinitesimal model assumptions, the distribution of breeding values is Gaussian. This assumes that loci are unlinked and that there has been no previous selection. If the random variables being summed are sufficiently correlated, the central limit theorem fails and the distribution need not converge to a normal. This can happen when selection generates dependencies (gametic-phase disequilibrium, LD) among loci, driving the distribution away from a Gaussian. Under the infinitesimal model, there are no changes in allele frequencies, implying that once selection is stopped and random mating occurs, the departures from normality quickly decay. Indeed, Bulmer (1980) showed that the k-th order departure from normality (measured by cumulants, which we introduce shortly) decays by $(1/2)^{k-1}$ each generation, so that following t generations of random mating, the k-th order departure from normality is just $(1/2)^{t(k-1)}$, which quickly approaches zero. Thus, normality is rapidly restored by random mating once selection is stopped. The issue remains as to how much of a departure from normality LD generates under the infinitesimal model and whether this biases infinitesimal-model-based predictions of response. We return to this issue later in the chapter.

Another key feature of the infinitesimal model is that the distribution of breeding values A_o in the offspring, conditioned on the breeding values A_f, A_m of its parents, is normally distributed with mean $(a_f + a_m)/2$ and variance $\sigma_a^2/2$. Thus, we have homoscedasity with the predictor error variance being a constant, independent of the parental values. We have seen (Chapter 13) that this **Mendelian sampling variance**, which half the genic variance in an infinitesimal model in a infinite population, is cause by segregation in the parents (and hence is often called the **segregation variance**). Allele frequency change, as can occur with a finite number of loci and selection and/or drift under finite population size, can change the genic variance from one generation to the next, a point we examine shortly.

Not All Limits are Gaussian

The segregation variance is no longer a constant independent of parental genotypes when major genes are present. Indeed, this is one (albeit weak) test for the presence of a major gene (LW Chapter 8). Somewhat surprisingly, Dawson (1997) showed that, even under the infinitesimal limit, there are conditions where the segregation variance may vary over parents. Specifically, Dawson showed that while in the limit (number of loci going to infinity) the within-family random segregation typically approaches a normal, the formal infinitesimal model many not hold in that this variance can vary over families if the pair-wise LD is sufficiently large – in particular if a fraction of individuals are identical across a large fraction of loci. A large number of loci close to fixation can also violate the conditions for convergence to the strict (constant-variance) infinitesimal model. Dawson (1997) presents conditions on

the LD to ensure convergence to the infinitesimal model (i.e., constant family variance).

While the convergence to a normal distribution of breeding values occurs under the infinitesimal model (each locus has a vanishing small effect), simply having an infinite number of loci contributing to a trait is not sufficient for normality. If loci are sufficiently correlated, or if their effects are significantly different (for example, some remain at finite values while the rest get vanishingly smaller), then convergence to a normal is by no means assured. Matthyssee et al (1979) show that even with a model with an infinite number of genes with gradually diminishing effects need not always converge to a normal. Also see Lange (1978), who examines conditions under which a sum of effects over a large number of loci converges to a normal.

Modifications of the Infinitesimal Model

The rest of this chapter starts to move beyond the infinitesimal. First, by assuming a gaussian distribution of allelic effects *at each locus*, we can partly account for changes in allele frequencies caused by a finite number of loci and/or genetic drift. These approximation break down over time, and hence are best regarded as an intermediate-term predictor of response. Next, we allow for linkage. Finally, we examine what can be said when the distribution of allelic effects are no longer normal. None of these approaches fully accounts for allele frequency changes, and thus are best considered predictors for intermediate-term response. Prediction of long-term response requires explicit population-genetic models (Chapter 5). Chapters 25 – 27 examine the consequences for long-term response of these more explicit models.

GAUSSIAN CONTINUUM-OF-ALLELES MODELS

Simulation studies (e.g., Bulmer 1974, 1976a, Sorensen and Hill 1983, Mueller and James 1983, Chevalet 1988) have shown that the infinitesimal model gives a reasonably good fit to the change in variance over a few generations of selection when the number of loci is finite. However, with a finite number of loci, allele frequency changes occur and after a sufficient number of generations the cumulative effects of these changes become so large that they cannot be ignored. Likewise, if the population is finite, genetic drift also changes allele frequencies. Thus when either the number of loci n or the population size N is finite, we must incorporate changes in the *genic* variance σ_a^2 into our model.

Is there an intermediate step between the short-term predictions from the breeders' equation/infinitesimal model and the unpredictable long-term behavior when significant allele frequency changes have occurred? In many cases, the answer is yes, and is provided by approximations using **continuum-of-alleles models** (we will often simply refer to these as **COA models**). These models allow us to partly account for modest changes in allele frequencies due to selection (given a finite number of loci) and/or genetic drift (due to finite population size). The nice feature about these intermediate-term approximations for the selection response is that they are based entirely on macroscopic parameters, and thus there is some hope of estimating these.

Infinite Alleles and Continuum-of-alleles Models

The historical roots of the continuum-of-alleles model trace back to the classic paper of Kimura and Crow (1964), which introduced their **infinite alleles model**. Before this paper, most population-genetic models typically assumed two (or at most a few) alleles per locus. Kimura and Crow, in the first serious treatment of molecular evolution, noted that with an allele being represented by a long DNA sequence, each new mutation likely creates a new sequence, and hence an infinite number of alleles are possible. Kimura and

Crow's original paper simply dealt with how much variation (measured in terms of heterozygosity) could be maintained by the balance between drift and mutation (Chapter 6). Frequencies, rather than allelic effects, were the focus on the infinite alleles model. Crow and Kimura (1964) and Kimura (1965) quickly applied this notion of a very large number of alleles per locus to quantitative genetics by considering the distribution of allelic effects at each locus. Thus, **continuum-of-alleles** models were introduced, and further developed by Latter (1970), Lande (1975, 1977) and Felsenstein (1977) to model mutation-selection balance (Chapter 27). Kimura's (1965) original analysis found that if new mutations have small effects relative to the existing variation at the locus, then the distribution of effects (in an infinite population) converges to a normal. Thus, COA models make the assumption that the distribution of breeding values *at each locus* is gaussian (and jointly multivariate normal over a vector of loci), which can only be strictly correct if there are an infinite number of alleles at each locus, and hence an infinite population size. This assumption of *Gaussian distribution of effects at each locus* is much more restrictive than the assumption that the distribution of the *total* genotypic value is normal. While the distribution of *total* genotypic values is gaussian under continuum-of-alleles model, the central limit theorem allows the sum of non-normal distributions across loci to converge to a gaussian. Thus, COA models are a very restrictive subset of all possible models that can lead to the infinitesimal. The advantage of COA models is that we can assume a *finite number of loci*, and hence partly accommodate allele frequency change.

COA models attempt to approximate allele frequency changes, at least over some intermediate time scale. The breeders' equation holds for a single generation of response, and correction for Bulmer effect (changes in the genetic variances generated by disequilibrium, but not allele frequencies) allows us to extend the time scale for accurate short-term predictions. COA models attempt to bridge these simple short-term predictors which rely on estimable qualities (σ^2_A, h^2) with the long-term predictors of response (Chapters 25, 26) that are based on population-genetic models containing quantities are essentially unestimable. COA models attempt to capture the changes in variance not only for selection generating disequilibrium, but also (and more importantly) from changes in allele frequencies, while still using estimable quantities. Continuum-of-alleles approximations of the Bulmer equation for the change in variance (Equation 13.7) under a finite number of loci (n) were introduced by Lande (1975) and Felsenstein (1977, 1979), while Keightley and Hill (1987) allow for finite effective population size (N_e). We consider the effects of drift first.

Drift

Assuming that the phenotypic variance after selection has the form $\sigma^2_{z^*} = (1-\kappa)\,\sigma^2_z$ (Equation 13.10), then the equations for change in additive genic variance σ^2_a and the gametic-phase disequilibrium d when population size N_e is finite become

$$\Delta\,\sigma^2_a(t) = -\frac{\sigma^2_a(t)}{2N_e} \tag{24.1a}$$

$$\Delta\,d(t) = -\frac{1}{2}\left[\left(1+\frac{1}{N_e}\right)d(t) + \left(1-\frac{1}{N_e}\right)\kappa\,h^2(t)\,\sigma^2_A(t)\right] \tag{24.1b}$$

As before, $\sigma^2_A(t) = \sigma^2_a(t) + d(t)$ and $h^2(t) = \sigma^2_A(t)/\sigma^2_z(t)$, where $\sigma^2_z(t) = \sigma^2_A(t) + \sigma^2_e$ with $\sigma^2_e = \sigma^2_z(0) - \sigma^2_A(0)$. The resulting response in the mean is given by the breeders' equation, $R(t) = h^2(t)S(t)$.

If population size N_e is at least modest, corrections for its effects of drift on disequilibrium d are corresponding small. The effects of drift on the genic variance, however, are quite substantial, with drift removing all genic variance after sufficient time. Solving Equation

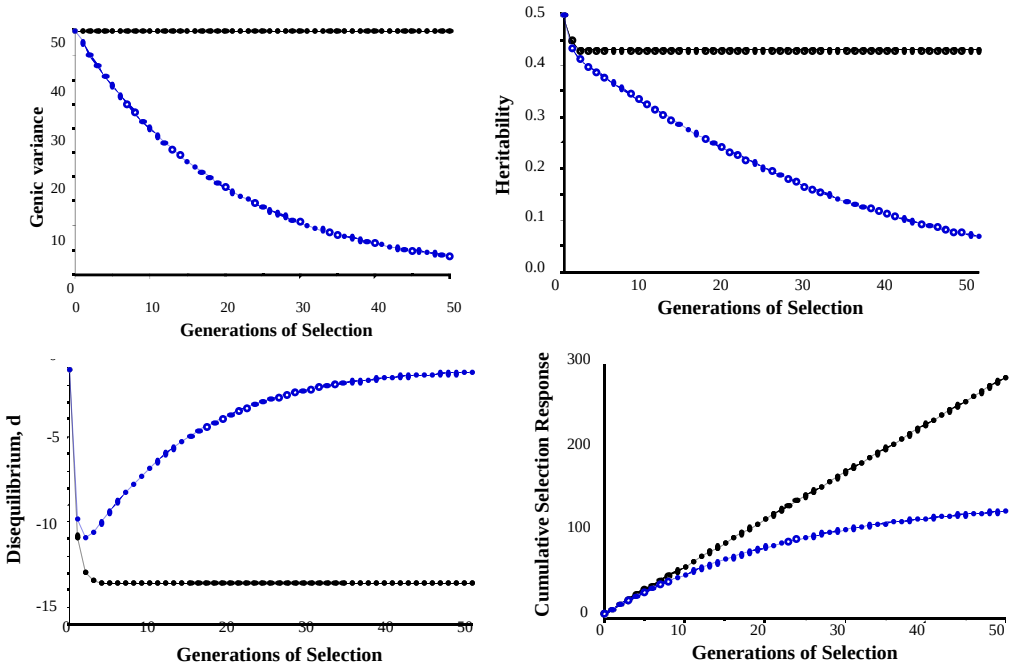
24.1a gives

$$\sigma_a^2(t) = \left(1 - \frac{1}{2N_e}\right)^t \sigma_a^2(0) \simeq \sigma_a^2(0)\, \exp\left(-\frac{t}{2N_e}\right) \tag{24.1c}$$

This is just the standard loss of genetic variation under drift. Recall (Chapter 9) that when dominance and/or epistasis is present, the additive variance can actually *increase* (for a while) under inbreeding, so the assumption of only additive gene action is critical. With finite population size the response runs out of variation, as σ_a^2 is driven to zero by drift.

Example 24.1. To see the effects of drift on the infinitesimal model, let's reconsider Example 13.2, but now allow finite population size. Recall this example assumed truncation selection with the upper 20% saved (giving $\kappa = 0.787$ and $\overline{z} = 1.40$) with $h^2(0) = 0.5$, and $\sigma_z^2(0) = 100$ with (initially) no disequilibrium. Under the infinitesimal model, the genic variance σ_a^2 remains unchanged at its original value of 50, while the additive variance decreases to its equilibrium value of $\widetilde{\sigma}_A^2 = 37.41$, and hence $\widetilde{h}^2 = 0.43$ with an asymptotic value of response of $\widetilde{R} = \overline{z}\widetilde{h}^2\widetilde{\sigma}_z = 5.6$ per generation.

Now suppose we have a finite population size with $N_e = 10$. Many artificial selection experiments have effective population sizes close to this value (Chapter 26). Iteration of Equation 24.1 gives the dynamics depicted below. The open circles correspond to the finite populations, the filled circles to the infinitesimal model in the absence of drift:



The key feature is that drift erodes away the genic variance, decreasing the heritability (and hence response) over time. The population (in the absence of mutation) will eventually run out of variation and reach a **selection limit** (Chapters 25, 26). Note the unusual behavior of the disequilibrium d, which (following an initial drop) decreases toward zero over time. This occurs because the genic variance is declining.

Drift and a Finite Number of Loci

Under the infinitesimal model, there is no change in allele frequencies and hence the genic variance remains unchanged. As we have seen, when the population size is finite, alleles are lost (and fixed) by drift, changing allelic frequencies and eventually reducing the genic variance to zero. A second route for allele frequency change is when the number of loci n is finite. In this case, there are non-zero selective effects on each locus and hence allele frequencies at these loci change. Again assuming that the distribution of genotypic values at *each* locus is Gaussian, continuum-of-alleles models can account for both finite N_e and n. The most general result is due to Chevalet (1988, 1994). Again assuming that $\sigma^2_{z^*} = (1-\kappa)\,\sigma^2_z$, then with loci of equal effect we have

$$\Delta\,\sigma^2_a(t) = -\left[\,\frac{\sigma^2_a(t)}{2N_e} + \left(1-\frac{1}{N_e}\right)\frac{\kappa\,h^2(t)\,\sigma^2_A(t)}{2n}\,\right] \tag{24.2a}$$

$$\Delta\,d(t) = -\frac{1}{2}\left[\left(1+\frac{1}{N_e}\right)d(t) + \left(1-\frac{1}{n}\right)\left(1-\frac{1}{N_e}\right)\kappa\,h^2(t)\,\sigma^2_A(t)\right] \tag{24.2b}$$

Provided we are willing to accept the assumption that the distribution of effects at each locus remains normally-distributed (a point we address later), we can simply iterate these expressions to obtain the current values of σ^2_a and d. Note that, starting from an unselected base population, the only parameters needed to iterate the above equations are $\sigma^2_A(0)$, h^2, n, and N_e, all of which are potentially estimable.

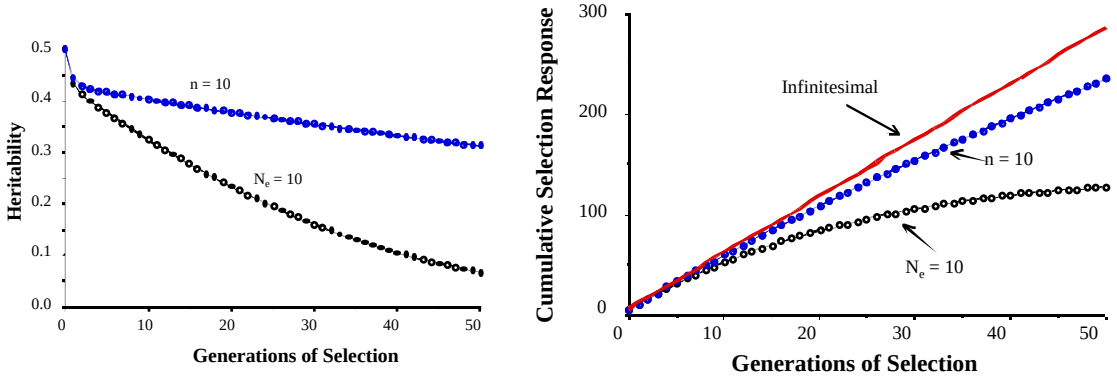
Equation 24.2 highlights the changes that occur when we assume a finite number of loci ($n < \infty$) and/or finite population size ($N_e < \infty$). When both are infinite, we recover the simple Bulmer equation (13.12),

$$\Delta\,\sigma^2_a(t) = 0, \qquad \Delta\,d(t) = -\,\frac{d(t) + \kappa\,h^2(t)\,\sigma^2_A(t)}{2}$$

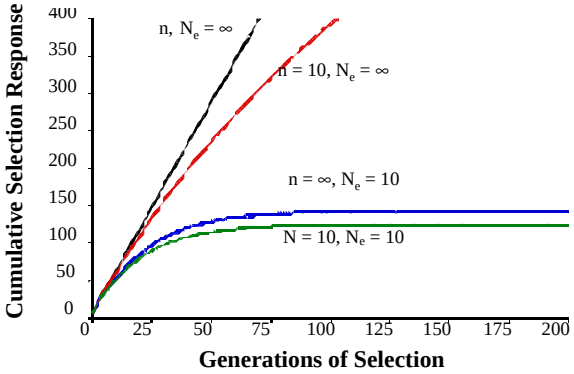
where the additive genic variance σ^2_a remains unchanged (as allele frequencies remain unchanged), while disequilibrium (nonzero d) is generated by selection, and decays to zero once selection stops.

While finite n and/or N_e result in modifications of the simple Bulmer equation for the dynamics of d, these corrections are generally small. This is not the case for changes in σ^2_a. With either finite population size and/or finite number of loci, the genic variance σ^2_a decreases each generation, eventually going to zero (in the absence of mutation). The relative importance of drift versus a finite number of loci on changes in σ^2_a can be compared using Equation 24.4a. With selection generating negative d (directional and/or stabilizing selection), then $d < 0$, $0 < \kappa < 1$, and $\sigma^2_A = \sigma^2_a + d < \sigma^2_a$ so that $\sigma^2_a(t) < \kappa\,h^2(t)\,\sigma^2_A(t)$. Thus, for comparable values of N_e and n, drift results in a greater per-generation reduction in σ^2_a.

Example 24.2. Now let's consider what happens when the number of loci is finite. We assume the same model as in Example 24.1, but now let $n = 10$ and $N_e = \infty$. We will contrast the behavior of this system with that in Example 24.1 ($N_e = 10, n = \infty$) and the infinitesimal model ($N_e = n = \infty$). As the figures below show, h^2, and response, both decrease over time with a finite number of loci, and eventually a selection limit is reached as all variation is lost. However, these decreases is not nearly as dramatic as the decreases we saw in Example 24.1.



It is of interest to consider the response when both n and N_e are finite. As the figure below shows, the cumulative response for a model with $N_e = n = 10$ is only very slightly less than a model with drift only ($N_e = 10$).



Effective Number of Loci, n_e

Chevalet (1994) further relaxed assumptions by allowing loci to differ in the amount of genetic variance they contribute. In this case, Chevalet showed that we replace the number of loci n in Equation 24.2 by n_e, the **effective number of loci**,

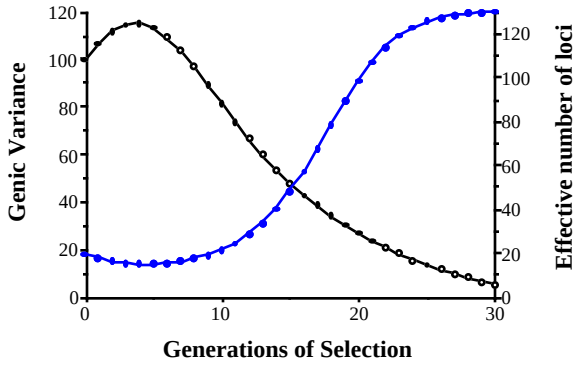
$$n_e = \frac{n}{1 + cv^2} \tag{24.3}$$

where cv is the coefficient of variation in the genic variance contributed by each locus, $cv = \sigma\left(\sigma^2_{ai}\right)/E[\sigma^2_{ai}]$, where σ^2_{ai} is the genic variance contributed by locus i. Note that this is closely related to the Castle-Wright estimator for number of segregating genes in a line cross F_2 population (LW Chapter 9). If all loci contribute the same variance, then $cv = 0$ and $n_e = n$.

Note that n_e changes over time, as allele frequency changes alter the genic variance contributed by any particular locus. Indeed, loci with the largest genetic variance should show the most initial response to selection, and hence the fastest depletion of genic variance. In such cases, one can move from a situation where the effective number of loci is quite small (a few of the loci have large effects, and hence a high cv) to a situation where n_e can be quite large (if the remaining loci all having roughly equal effects so that the cv is small). Hence,

n_e can increase over time, but we also correspondingly expect the total genic variance σ_a^2 for the remaining loci to decrease.

Example 24.3. Consider an additive model with both major and minor loci. There are five major loci, each with frequency $p = 0.25$ and effect $a = 5.16$, and 125 minor loci, each with $p = 0.5$ and $a = 0.89$. The resulting initial genic variance is $\sigma_a^2 = 100$ and we assume a initial heritability of $h^2 = 0.5$. Finally, we assume truncation selection with the uppermost 20% saved (further details for this model are given in Example 26.2). We ignore any effects of disequilibrium, focusing on how the genic variance (open circles) and the effective number of loci n_e (filled circles) change over time due simply to allele frequency changes.



While there are 130 loci in this system, initially the effective number is around 20, due to the large coefficient of variance in the locus-specific genic variances. As we start selection, the additive variance initially increases, as the major alleles increase their frequencies toward 0.5 (where they have maximal additive variance). Such an increase in variance is not predicted by COA models. Notice that the effective number of loci further decreases during this increase in variance, as the coefficient in variation for the locus-specific variance increases as the genic variances at each major locus increases. As these major loci become fixed, the total genic variance decreases, while the effective number of loci increases, reflecting a decrease in the coefficient of variation.

Dynamics: σ_a^2 and d Change on Different Time Scales

Chevalet (1988, 1994) and Gavrilets and Hastings (1994, 1995) noted that the dynamics of the genic variance and the disequilibrium occur on rather different time scales. The change in d is rather rapid, quickly approaching a **quasi-equilibrium** value,

$$d = -\kappa \left(1 - \frac{1}{n_e}\right) h^2 \sigma_A^2 \tag{24.4a}$$

This is not an equilibrium value, as changes in σ_a^2, which occur over much slower time scales, also change σ_A^2, albeit much more slowly. Note that as $n_e \to \infty$, we recover the equilibrium d value found by Bulmer (which is a true equilibrium as σ_a^2 does not change under the infinitesimal model). Thus the dynamics of COA models operate over two rather different time scales – for a given value of σ_a^2 there is a quick approach to the equilibrium value of d.

Over a much slower time scale, allele frequency changes change σ_a^2. Thus, at any particular time we (approximately) have

$$\sigma_A^2 = \sigma_a^2 - \kappa \left(1 - \frac{1}{n_e}\right) h^2 \sigma_A^2 \tag{24.4b}$$

Likewise, the distribution of additive genetic variance within the population can be decomposed into that held between families (the difference in family means) and that generated by segregation within each family (Chapter 13). When no disequilibrium is present, under random mating each component is $\sigma_A^2/2$. However, with selection the genetic variance among full sib families is

$$\sigma^2(FS) = \frac{\sigma_A^2}{2}\left(1 - \kappa h^2\right) \tag{24.5b}$$

while the within-family (additive genetic) variance is

$$\sigma_A^2(\text{within-family}) = \frac{\sigma_a^2}{2} = \frac{\sigma_A^2}{2}\left(1 - \left(1 - \frac{1}{n_e}\right)\kappa h^2\right) \tag{24.5b}$$

What forces dominate the dynamics of allele frequency change? Recall from Equation 24.1c that drift changes allele frequencies on a scale of t/N_e generations. Over this timescale, Gavrilets and Hastings (1995) showed that the dynamics of allele frequency changes from selection are largely governed by N_e/n. If n is sufficiently small relative to N_e, then selection is important, while is n is large relative to N_e then drift dominates (e.g., Example 24.2).

Response in Stabilizing Selection Experiments: Selection or Drift?

Gavrilets and Hastings (1994, 1995) noted that under stabilizing selection, the allele frequency changes can be quite slow. Simulations, as well as their analysis of two- and n- locus models, showed that while a rapid approach to the quasi-equilibrium d value (Equation 24.4a) occurs, the rate of change of allele frequencies can be very slow – on the order of a hundred (or more) of generations even with strong selection under a two locus model. As mentioned in Chapter 13, we expect an immediate decrease in the phenotypic variance due to selection generating negative disequilibrium, reducing σ_A^2. Gavrilets and Hasting suggest that, given the short time scales of most experiments, if further decreases in the variance occur from allele frequency changes, that these changes are more likely due to drift than selection. As we mentioned in Chapter 13, selection-driven allele frequency changes might also be involved, but through a different pathway than reduction in σ_A^2. If there is heritable variation in the environmental variance σ_E^2, then stabilizing selection can result in a reduction in σ_E^2, and hence a reduction in the phenotypic variance *independent* of any reduction due to reduction in the additive genetic variance.

How Robust is the Continuum-of-alleles Model?

If the trait is determined by a modest to large number of loci, all of roughly equal effect, and with alleles at intermediate frequencies, then COA models can perform reasonably, as least over intermediate time scales (Chevalet 1988, 1994). They generally tend to overestimate the cumulative response as generations increase, so that while the decrease in σ_a^2 from selection is partially captured (and hence an improvement over the infinitesimal model), after sufficient time the COA approximation break down.

What are possible causes for this? Of course, the Gaussian assumption for each locus is a key, but one that is also incorrect at the start of selection, when COA approximation generally works well. It appears that if the distributions of genotypic values at each locus are *close* to normal, then the COA approximation holds reasonable well as a predictor for selection

responses. As allele frequency changes drive the individual locus distributions further away from normality, the COA approximation breaks down. Generation of disequilibrium by selection also drive the total distribution of genotypic values away from normality.

Loci with alleles at extreme frequencies (near zero or one) can show large departures from normality. A theme that reappears through this chapter is a focus on the skewness and kurtosis of distributions. Both of these (with kurtosis appropriately defined) are zero under a normal and hence these third and fourth moments provide on measure of the departure of any particular distribution from normality. Skewness measured departures from symmetry, while kurtosis measures if the tails of the distribution fall of more rapidly, or more slowly, than a normal.

Suppose we have n exchangeable loci, each with frequency p of the favorable allele. Zeng (1987) showed that the resulting scaled (to unit variance) coefficients of skewness γ_3 and kurtosis γ_4 are

$$\gamma_3 = \frac{2p-1}{\sqrt{2np(1-p)}}, \qquad \text{and} \qquad \gamma_4 = \frac{1-2p(1-p)}{2np(1-p)} \tag{24.6}$$

Note that skewness is zero and kurtosis is minimized at intermediate allele frequencies ($p = 1/2$). As allele frequencies become more extreme, so does the skew and kurtosis. Rare alleles of large effect are especially problematic. Not only do these generate skewness and kurtosis, but as their frequencies increase, so does the genic variance. By contrast, under the COA model genic variance always declines.

Zhang and Hill (2005) note that natural selection tends to generate a correlation between allelic effect size and frequency, so that alleles of large effect may tend to be rare in natural populations, due to pleiotropic deleterious fitness effects often associated with major alleles. If this population is sampled to form a laboratory stock for artificial selection, then if the rare alleles are captured, an increase in additive variance is expected during selection. However, if the founding population is under strong drift for a few generations (such as due to the founding bottleneck), such rare alleles can be lost and the COA approximation may be a good predictor of short-term response. This theme of favorable rare alleles, and hence initially accelerated response to selection, will be revisited in Chapter 27, as it is central to certain predictions about response in a population under mutation-selection balance.

THE BULMER EFFECT UNDER LINKAGE

Both the infinitesimal and continuum-of-alleles models assumed loci are *unlinked*. Obviously with linkage d does not decay by half each generation, increasing its importance. To examine just how large an effect linkage creates, we examine two different models. The first is a continuum-of-alleles approximation (requiring multivariate normality of the locus-specific distributions of effects). The second is more technical, allowing for departures from normality. It also serves as a jumping-off point for our final section on treating non-gaussian distributions.

An Approximate Treatment

Recall from Chapter 13 that C_{ij} is the covariance between allelic effects at locus i and j, so that C_{ii} is the genic variance for locus i, while C_{ij} for $i \neq j$ measures the contribution from disequilibrium between i and j. Thus (Equation 13.1b), $d(t) = 4\sum_{j<i} C_{ij}(t)$, with the changes in the pairwise covariances describing the change in d. If r_{ij} is the recombination fraction between two loci, then $(1-r_{ij})\,C_{ij}(t)$ is the contribution passed on to $C_{ij}(t+1)$. Compared to unlinked loci, a larger fraction of the disequilibrium between linked loci persists

each generation. An approximate solution incorporating linkage was offered by Bulmer (1974, 1980), whose approach we follow (a more general solution by Turelli and Barton will be considered shortly). Recalling Equation 13.6, the change in $d(t)$ due to selection when genotypic and phenotypic values are normally distributed is

$$\frac{h^4(t)}{2}\,\delta(\sigma^2_{z(t)}) = 4\sum_{j<i}\delta C_{ij}(t) \tag{24.7a}$$

where δX denotes the *within-generation* change in the variable X. In order to approximate δC_{ij} (the new disequilibrium generated by selection), Bulmer assumed these changes are the same for each pair of loci (hence, he assumed an exchangeable model). Since for n loci there are $n(n-1)/2$ unique pairs, the contribution from each pair of loci to the left hand side of Equation 24.7a is

$$4\delta C_{ij}(t) \simeq \frac{h^4(t)\,\delta(\sigma^2_{z(t)})}{n(n-1)} \tag{24.7b}$$

The new disequilibrium equals the decay in the current disequilibrium plus the fresh disequilibrium generated by selection,

$$C_{ij}(t+1) = (1-r_{ij})\,C_{ij}(t) + \delta C_{ij}(t) \tag{24.7c}$$

This equation is approximate as the covariance *between* gametes $C_{i,j}$ enters as well as the *within*-gamete covariance C_{ij} (Equation 24.11b below gives a more exact treatment). Ignoring this for now, Equation 24.7c implies at equilibrium that $r_{ij}\,\widetilde{C}_{ij} = \widetilde{\delta} C_{ij}$. Using Equation 24.7b gives the equilibrium covariance as

$$\widetilde{C}_{ij} = \frac{\widetilde{h}^4\,\widetilde{\delta}(\sigma^2_z)}{4n\,(n-1)}\,\frac{1}{r_{ij}} \tag{24.7d}$$

thus

$$\widetilde{d} = 4\sum_{j<i}\widetilde{C}_{ij}(t) = 4\,\frac{\widetilde{h}^4\,\widetilde{\delta}(\sigma^2_z)}{4n\,(n-1)}\sum_{j<i}^{n}\frac{1}{r_{ij}} = \frac{1}{2}\,\widetilde{h}^4\,\widetilde{\delta}(\sigma^2_z)\,\frac{1}{H} \tag{24.8a}$$

where H is the harmonic mean of all pairwise recombination distances between loci,

$$H = \left(\frac{1}{n\,(n-1)/2}\sum_{j<i}^{n}\frac{1}{r_{ij}}\right)^{-1} \tag{24.8b}$$

The value of H varies with both the number of loci and the number of chromosomes, with H decreasing as the number of loci becomes increasing larger than the number of chromosomes. Using simulations of randomly distributed loci, Bulmer (1974) found that if the haploid chromosome number exceeds 10, H is likely no smaller than 0.4, while in *Drosophila melanogaster*, with its three main chromosomes and lack of recombination in males, H is around 0.1 if there are many loci. Even if only a few loci occur as tightly linked pairs, H can be lowered considerably from 0.5 since the harmonic mean disproportionately weights these very small values.

Assuming the phenotypic variance after selection is given by $\sigma^2_{z^*} = (1-\kappa)\,\sigma^2_z$, then the equilibrium additive genetic variance is given by $\widetilde{\sigma}^2_A = \sigma^2_z\,\theta$, where now

$$\theta = H\left(\frac{2h^2 - 1 + \sqrt{1 + 2h^2(1-h^2)\kappa/H}}{2H+\kappa}\right) \tag{24.9}$$

and the equilibrium heritability is given by Equation 13.13c using the above value of θ. As before, if disruptive selection is sufficiently strong ($\kappa < -H/2h^2(1-h^2)]$) there is no real positive root for θ and the infinitesimal model predicts that the additive variance increases without limit (Bulmer 1976a). (Recall that in this case with a finite number of loci, selection creates almost complete disequilibrium, so that only a few of the possible gamete types are actually present, i.e., most gametes are either $aabbccdd\cdots$ or $AABBCC\cdots$.) The general conclusion is that increasing the amount of linkage (e.g., decreasing H) increases the absolute value of $\widetilde{d}$ (Bulmer 1974, 1976a, 1980).

Example 24.4. As an example of the consequences of increased linkage, reconsider our analysis of the response under directional used in Examples 24.1 and 24.2. Here we assume infinite number of loci and infinite population size. Substituting into Equation 24.9 to obtain θ and recalling Equations 13.13a-c gives

H	θ	$\widetilde{d}$	$\widetilde{\sigma}_A^2$	$\widetilde{h}^2$	$\widetilde{R}$
0.5	0.37	−12.60	37.40	0.43	5.60
0.4	0.35	−14.50	35.50	0.42	5.37
0.3	0.33	−17.11	32.89	0.40	5.06
0.2	0.29	−20.97	29.03	0.37	4.57
0.1	0.23	−27.49	22.51	0.31	3.70

$H = 0.5$ corresponds to free recombination, while $H = 0.1$ might be expected in *Drosophila melanogaster*. As expected, decreasing the average amount of recombination between loci increases the effect of linkage disequilibrium, with more extreme d values, and hence smaller additive variances, heritabilities, and selection responses. For example, with strong linkage ($H = 0.1$), the response is reduced 33% relative to unlinked loci.

The infinitesimal model predicts that $\widetilde{d}$ increases under disruptive selection as linkage becomes tighter, while Sorensen and Hill (1983) found in their simulations that $\widetilde{d}$ *decreases* as linkage tightens. They reasoned that this discrepancy arises due to the interaction between a finite number of loci and the finite population sizes used in the simulations. To see this, consider complete linkage. In a finite population, the most extreme gamete observed (and hence the ultimate level of $\widetilde{d}$) is affected by sampling as selection can generate no gamete more extreme than those found in the initial sample. If the number of loci is small, the probability of sampling the most extreme possible gamete is high, but this probability decreases as the number of loci increases. Countering this, as recombination (measured by H) and/or the population size increases, the probability increases that recombination can regenerate the most extreme possible gametes before the relevant loci are fixed by drift and/or selection. When population size becomes large enough that drift effects are no longer important, $\widetilde{d}$ once again decreases with increasing linkage. Interactions of this sort between drift, selection and recombination are considered in some detail in Chapter 26.

As we saw before, realistic (i.e., few allele) models can result in an increase in the additive variance at some point during selection (typically fairly early) as favored alleles at low frequencies increase to immediate values. Hospital and Chevalet (1996) saw a similar phenomena in their simulation of linkage, namely that finite locus models can also show an increase in additive variation. The distinction is that this increase may come not rather quickly (as was the case for rare alleles), but rather many generations after selection was initiated, reflecting to recombination generating favorable gametes, which then increase in

frequency. They also found that linked systems are very vulnerable to lower response due to hitchhiking fixing deleterious alleles, an issue we return to in Chapter 26.

A More Careful Treatment

A more rigorous treatment of how selection changes the within-gamete covariances C_{ij} requires consideration of the **between-gamete covariance** $C_{i,j}$ as well as higher-order covariance terms that measure the amount of gametic-phase disequilibrium between groups of loci. A careful derivation highlights the importance of the normality assumptions we have liberally used above. The consequences of relaxing normality are considered in detail in the next section, while this section introduces some of the notation needed to examine non-Gaussian distributions of genotypic and phenotypic values.

We start by defining the between-gamete covariance,

$$C_{i,j} = \sigma\left(a_{fa}^{(i)}, a_{mo}^{(j)}\right) \tag{24.10}$$

which is the covariance between the effect of an allele at the ith locus in the paternal gamete and an allele at the jth locus in the maternal gamete. Under random mating, gametes unite at random and $C_{i,j} = 0$ at the start of each generation. Selection generates correlations *between* gametes in much the same way that it generates correlations among loci *within* gametes. For example, consider a particular chromosome containing multiple loci influencing a character under stabilizing selection. Initially, there is no correlation between the genetic values of the two copies of this chromosome in an offspring from randomly-mated parents. Stabilizing selection changes this initial distribution, favoring adults with an intermediate genotypic value. Thus surviving adults with a large genetic value on one chromosome are expected to have a small value on the other and vice-versa, generating negative $C_{i,j}$. Likewise, assortative mating generates positive $C_{i,j}$, while disassortative mating generates negative $C_{i,j}$.

We assume random mating, so that $C_{i,j}(t) = 0$ at the start of each generation. Letting C^* denote the covariance after selection, $C_{ij}^* = C_{ij} + \delta C_{ij}$ and $C_{i,j}^* = C_{i,j} + \delta C_{i,j} = \delta C_{i,j}$. Assuming recombination follows selection, with probability $1 - r_{ij}$ no recombination occurs between i and j and the within-gamete covariance is unchanged, while with probability r_{ij} recombination occurs and the new covariance depends on the covariance *between* gametes, giving the result of Lande (1975) and Bulmer (1980),

$$C_{ij}(t+1) = (1 - r_{ij})C_{ij}^*(t) + r_{ij}C_{i,j}^*(t) \tag{24.11a}$$

Substituting for C^* gives

$$\begin{aligned} C_{ij}(t+1) &= (1 - r_{ij})\left[\,\delta C_{ij}(t) + C_{ij}(t)\,\right] + r_{ij}\,\delta C_{i,j}(t) \\ &= (1 - r_{ij})C_{ij}(t) + \delta C_{ij}(t) - r_{ij}\left[\,\delta C_{ij}(t) - \delta C_{i,j}(t)\,\right] \end{aligned} \tag{24.11b}$$

Note that we recover Equation 24.7c only if $\delta C_{ij} = \delta C_{i,j}$ (selection changes the within-gamete and between-gamete covariances by the same amount). Turelli and Barton (1990) show that this occurs if there is either global gametic-phase equilibrium (all groups of loci are in gametic-phase equilibrium) or if the distribution of allelic effects over loci is multivariate normal. Thus, Equation 24.7c follows under COA assumptions. However, selection can drive a distribution away from normality, in which case 24.7c may no longer hold.

General expressions for δC_{ij} and $\delta C_{i,j}$ have been obtained by Turelli and Barton (1990) for the case of no dominance or epistasis. Their expressions involve (i) generalizations of measures of selection to higher moments of a distribution and (ii) generalizations of disequilibrium measures to groups of k loci. Starting with (i) first, recall (Equation 10.7a) that we defined the **directional selection gradient**, which measures how selection acts on the

phenotypic mean, as $\partial \ln \overline{w} / \partial \mu_z$. We can extend this notion to higher moments by considering $\partial \ln \overline{w} / \partial \mu_{k,z}$, where $\mu_{k,z} = E[\,(z-\mu_z)^k\,]$ is the kth central moment of the phenotypic distribution (for $k \geq 2$). If selection is primarily on the mean and variance, selection gradients for the skew and higher moments ($k \geq 3$) are generally negligible. For the case where phenotypes are normally distributed,

$$\frac{\partial \ln\, \overline{w}}{\partial \mu_z} = \frac{S}{\sigma_z^2} \tag{24.12a}$$

$$\frac{\partial \ln\, \overline{w}}{\partial \sigma_z^2} = \frac{\delta(\sigma_z^2) + S^2}{2\,\sigma_z^4} \tag{24.12b}$$

(Lande 1976, Lande and Arnold 1983). As will be shown in Chapters 28 and 29, when selection acts only on the mean, the within-generation change in the phenotype variance is $\delta(\sigma_z^2) = -S^2$, so that $\delta(\sigma_z^2) + S^2$ is the change in variance over that expected due to selection on the mean.

Using these extended selection gradients and ignoring selection acting on the skew and higher moments (by assuming that gradients for $k \geq 3$ are negligible), Turelli and Barton (1990) found that

$$\delta C_{ij} = \frac{\partial \ln\, \overline{w}}{\partial \mu_z} \sum_k C_{ijk} + \frac{\partial \ln\, \overline{w}}{\partial \sigma_z^2} \sum_k \sum_l (\, C_{ijkl} - C_{ij} C_{kl}\,) + \cdots \tag{24.13a}$$

$$\delta C_{i,j} = \frac{\partial \ln\, \overline{w}}{\partial \sigma_z^2}\, 2 \sum_k C_{ik} \sum_l C_{jl} + \cdots \tag{24.13b}$$

where C_{ijk} refers to the third-order covariance between the alleles at loci i, j, and k. If X_i is the additive value of a randomly chosen allele at locus i and $\mu_i = E(X_i)$ is the average value for this locus, then $C_{ijk} = E[\,(X_i - \mu_i)(X_j - \mu_j)(X_k - \mu_k)\,]$. Higher-order covariances are defined similarly. The covariances in Equation 24.13a measure the amount of third- (C_{ijk}) and fourth- (C_{ijkl}) order gametic-phase disequilibrium (the departures from random assortment for triplets and quadruplets of loci). If selection on the third (skew) or higher-order moments is significant, then Equation 24.13 includes covariance terms of order five and higher.

The key point about these equations is that *changes in covariances depend critically on very fine details of the genotypic distribution*, details that are essentially impossible to estimate empirically in realistic situations, and simplifying assumptions are required to proceed further. For example, if the distribution of genotypic values is multivariate normal (which, as previously mentioned, involves the rather strong assumption that allelic effects *at each locus* are normally distributed), Equation 24.13 simplifies greatly as $C_{ijk} = 0$ and C_{ijkl} can be expressed in terms of second-order covariances ($C_{ijlk} = C_{ij}\,C_{kl} + C_{ik}\,C_{jl} + C_{il}\,C_{jk}$). In this case, $\delta C_{ij} = \delta C_{i,j}$, and combining Equations 24.12b and 24.13b gives

$$\delta C_{ij} \simeq \frac{\delta(\sigma_z^2) + S^2}{\sigma_z^4}\, C_i\, C_j \tag{24.14a}$$

where $C_i = \sum_j C_{ij}$. Thus when allelic effects are multivariate normal (normal at each locus and multivariate normal for any subset of loci), the change in covariance is given by

$$\Delta C_{ij}(t+1) = \frac{\delta(\sigma_z^2) + S^2}{\sigma_z^4}\, C_i(t)\, C_j(t) - \, r_{ij} C_{ij}(t) \tag{24.14b}$$

a result due to Lande (1975, 1977). Since $2\sum_i C_i = 2\sum_{ij} C_{ij} = \sigma_A^2$, then assuming all the C_i are equivalent, it follows for n loci that $C_i = \sigma_A^2/(2n)$, and Equation 24.14a reduces to

$$\delta C_{ij} \simeq \frac{\sigma_A^4}{4n^2\sigma_z^4}\left(\delta(\sigma_z^2) + S^2\right) = \frac{h^4}{4n^2}\left(\delta(\sigma_z^2) + S^2\right) \tag{24.14c}$$

When $S^2 << | \delta(\sigma_z^2) |$ (selection is mainly on the variance), we recover Bulmer's approximation (Equation 24.7b) when the number of loci n is large.

RESPONSE UNDER NON-GAUSSIAN DISTRIBUTIONS

The assumption of normality has been pervasive in most previous discussions in this book. By assuming phenotypic and genotypic values are (and stay) normally distributed (or **Gaussian**), we can describe all changes in the phenotypic distribution from just the mean and variance. The assumption that genotypic values of offspring are described by a linear regression of the genotypic values of their parents, which is the basis for much of the theory of selection response, most easily follows if the joint distribution of parental and offspring values is multivariate normal.

By changing allele frequencies and/or by creating gametic-phase disequilibrium, selection can drive a genotypic distribution away from a Gaussian. An active area of research is to describe both how selection can alter a distribution and to extend selection theory to deal with arbitrary distributions of genotypic values. While good progress has been made, we warn the reader that this can be a rather intimating area of the literature. Our purpose here is to introduce some of the basic ideas and machinery used, as well as to summarize the major findings.

We start by considering how the distribution of effects at each of the individual loci translate into a distribution of genotypic values. In particular, we examine how within-locus moments translate into moments of the genotypic distribution. While moments are more intuitive measures of the shape of a distribution, it is the **cumulants** of the distribution that are more natural to work with when describing deviations from normality. With this basic machinery in hand, we consider two types of models — (i) a small to modest number of segregating loci underlie the trait and (ii) a very large number of loci of small effect underlie the trait. With a small number of loci, to an initial approximation, one can ignore effects of gametic-phase disequilibrium and instead focus on the changes in the higher genotypic moments caused by allele frequency changes at the underlying loci. The key results for such models is that even single-generation predictions require extensive information about the underlying genetics. In contrast, with a very large number of loci, the (short-term) effects of allele frequency changes can be essentially ignored, and changes from gametic-phase disequilibrium become critical. The nice (and somewhat surprising) result for this latter class of models is that the breeders' equation (and Bulmer's extension to changes in the additive variance) are quite accurate for both directional selection and for strong disruptive selection (Turelli and Barton, 1994).

Describing the Genotypic Distribution: Moments

We now proceed to the general theory of response under arbitarty distributions of genotypes. For starters, a few comments on the genotypic and phenotypic distributions are in order. Under our assumption that genotypic and environmental values are additive and independent, $z = G+e$. Thus, if environmental values e are normally distributed, phenotypes are normally distributed if and only if genotypic values are Gaussian. However, the converse is not true — an approximately normal distribution of phenotypes *does not* imply that genotypic values are Gaussian. While we can test to see if *phenotypes* are normally distributed, this tells us little about the distribution of *genotypes*. While in theory we can estimate the distribution of breeding values by computing the breeding values for a sample of individuals (LW Chapter 26), this is impractical in most studies of natural populations. Further, most of the methods typically used to estimate breeding values assume they are normally distributed, and hence bias the distribution of estimated values towards normality.

Since we assume no genotype-environment interactions, if the environment remains constant over time, changes in the phenotypic distribution are entirely due to changes in the genotypic distribution. The moments of a distribution provide a convenient measure to describe its shape, and hence changes in the moments provide descriptions of changes in the shape of the distribution. To see the connection between the moments of the phenotypic and genotypic distributions, note that the phenotypic mean, variance, and skew can be decomposed as $\mu_z = \mu_G$, $\sigma_z^2 = \sigma_G^2 + \sigma_e^2$, and $\mu_{3,z} = \mu_{3,G} + \mu_{3,e}$, hence changes in any of the first three phenotypic moments exactly equals the change in the corresponding genotypic moments. Example 24.5 derives the fourth phenotypic moment,

$$\mu_{4,z} = \mu_{4,G} + \mu_{4,e} + 6\,\sigma_G^2\,\sigma_e^2 \tag{24.15}$$

so that changes in the fourth moment of the phenotypic distribution can be due to either changes in the second (variance) and/or fourth moments of the genotypic distribution (derived in Example 24.5). When e is normal, $\mu_{3,e} = 0$ and $\mu_{4,e} = 3\sigma_e^4$, simplying these expressions.

How do the moments of G depend on the distribution of effects at individual loci? If n loci control the character, our assumption of complete additivity implies

$$G = \sum_i^n (X_{fa,i} + X_{mo,i}) \tag{24.16}$$

where $X_{fa,i}$ ($X_{mo,i}$) is the value of the paternal (maternal) allele at the ith locus. Assuming both sexes have the same distribution of allelic effects, the moments of G can be related to moments of the distribution of allelic effects at individual loci by expanding

$$\begin{aligned} \mu_{k,G} &= E\left([G - \mu_G]^k \right) \\ &= E\left(\left[\sum_i^n X_{p,i} + X_{m,i} - 2E(X_i) \right]^k \right) \qquad \text{for } k \geq 2 \end{aligned} \tag{24.17}$$

Finally, assume random mating so that $X_{fa,i}$ and $X_{mo,i}$ are independent at the start of each generation. Since we assume that the distribution of allelic effects is the same in both sexes, we drop the subscript referring to parental origin.

When considering a particular moment of G, it will be important to distinguish between contributions to that moment from individual loci (**within-locus moments**) and from gametic-phase disequilibrium (**between-locus contributions**). This distinction was used earlier with the additive genetic variance (Equation 13.2) and here we extend this partitioning to the third and higher genotypic moments. To describe the distribution of effects at locus i, let $\mu_{1,i} = E(X_i) = m_i$ be the average value of an allele at locus i and define the kth moment for this locus by $\mu_{k,i} = E([X_i - m_i]^k)$ for $k \geq 2$. Summing over all n loci, define

$$M_1 = 2\sum_i^n \mu_{1,i} \tag{24.18a}$$

$$M_2 = 2\sum_i^n \mu_{2,i} \tag{24.18b}$$

$$M_3 = 2\sum_i^n \mu_{3,i} \tag{24.18c}$$

as the contribution to the mean, variance, and skewness of the genotypic distribution due to the mean, variance, and skew at individual loci. Finally, define the within-locus kurtosis (LW Chapter 2) as

$$M_4 = 2\sum_i^n \left(\mu_{4,i} - 3\,\mu_{2,i}^2 \right) \tag{24.18d}$$

While this may at first seem odd, recall for a normal that the fourth and second moments are related by $\mu_4 = 3\mu_2^2$. Hence, if the distribution of allelic effects *at each locus* is normal, $M_4 = 0$ and likewise $M_3 = 0$ (as $\mu_{3,i} = 0$ as a normal does not display skew). On the other hand, nonzero values of M_3 and/or M_4 cause G to be non-Gaussian.

The between-locus contributions from gametic-phase disequilibrium are described by C_{ij}, C_{ijk} and C_{ijkl}, the covariances between groups of two, three, and four loci as defined previously. Note that with this notation $C_{ii} = \mu_{2,i}$, $C_{iii} = \mu_{3,i}$ and $C_{iiii} = \mu_{4,i}$, referring to the moments at locus i. If loci are independent (in gametic-phase equilibrium), then all other combinations involving four (or fewer) loci are zero expect C_{iijj}, which equals $C_{ii} \cdot C_{jj} = \mu_{2,i} \cdot \mu_{2,j}$.

Following Turelli and Barton (1990), we are now in position to decompose genotypic moments into within-locus effects (M_i) due to the moments at individual loci and between-locus effects due to covariances generated by gametic-phase disequilibrium. Remember that we are assuming the simplest case, complete additivity (no dominance or epistatis), so that the genotypic distribution G is the distribution of additive genetic values (A). Expanding Equation 24.17 and taking expectations gives the familiar expressions for the mean and variance,

$$\mu_G = 2\sum_i^n \mu_{1,i} = M_1 \tag{24.19a}$$

$$\sigma_G^2 = \sigma_A^2 = 2\sum_{i,j}^n C_{ij} = M_2 + 2\sum_i^n \sum_{j\neq i} C_{ij} \tag{24.19b}$$

Note that M_2 corresponds to the genic variance σ_a^2, while the double sum corresponds to the disequilibrium d. Similarly, the skew can be partitioned as

$$\mu_{3,G} = 2\sum_{i,j,k}^n C_{ijk} = M_3 + 2\sum_i^n \sum_{j,k\neq i} C_{ijk} \tag{24.19c}$$

All covariances in the second sums of Equations 24.19b and 19c are zero when all groups of two and three loci (respectively) are in gametic-phase equilibrium. Partitioning the kurtosis requires a little more care. After some simplification (Turelli and Barton 1990),

$$\mu_{4,G} = 3\,\sigma_A^4 + M_4 + 2\sum_i^n \sum_{j,k,l\neq i} (C_{ijlk} - C_{ij}\,C_{kl} - C_{ik}\,C_{jl} - C_{il}\,C_{jk}) \tag{24.19d}$$

Again, the covariance terms are zero when all groups of four loci are in gametic-phase equilibrium. Since $3\sigma_A^4$ is the value expected when genotypic values are Gaussian, the last two terms partition any kurtosis in G into the contribution from kurtosis at individual loci (M_4) and the contribution generated by gametic-phase disequilibrium between groups of four loci. If the distribution of allelic effects is multivariate normal, then $M_4 = 0$ and each term within the covariance sum is zero as $C_{ijlk} = C_{ij}\,C_{kl} + C_{ik}\,C_{jl} + C_{il}\,C_{jk}$.

Analogous to allele frequencies changing σ_a^2 and disequilibrium changing the covariances (and hence d), changes in M_3 and M_4 reflect *allele frequency change*, while changes

in the third- and fourth-order covariances reflect *changes in disequilibrium*. Thus, Equations 24.19a-d partition changes in the first four moments into those caused by allele frequency changes and those caused by changes in disequilibrium. When the number of loci is small, most changes are due to changes in allele frequencies, and hence the M_k, while when the number of loci is large (such that each has only a very small effect), most changes in the moments are due to changes in the covariances, and hence changes in the disequilibrium.

A few remarks on the implications of Equations 24.19a-d for deviations from normality are in order at this point. Higher-order moments can depart from their expectations under normality by the presence of skewness and/or kurtosis at the individual loci (generating non-zero M_3 and/or M_4), which can result from allele frequency changes. Alternatively, even if the within-locus moments are normal ($M_3 = M_4 = 0$), gametic-phase disequilibrium (nonzero C_{ijk} and/or $C_{ijk\ell}$) can introduce skewness and/or kurtosis. When the number of loci is small, skew and/or kurtosis at individual loci can be significant, giving nonzero M_3 and/or M_4, with the resulting genotypic distribution deviating from normality. The motivation for assuming that the distribution of G is normal is the central limit theorem, with sums of random variables (often) converging to a normal. Thus, as larger and larger numbers of loci underlie the character, the sum should approach a normal distribution as the contribution from each locus becomes smaller, reducing the effects of individual deviations from normality. The problem for strict converge to a Gaussian in our cases is that the central limit theorem assumes that the variables are independent (or at least only very weakly correlated). Selection, by introducing gametic-phase disequilibrium, generates just such a lack of independence.

To see these points, first consider the changes due to within-locus moments. If n is the number of loci, then as we have seen earlier the effects (a) of alleles at individual loci must scale as $1/\sqrt{n}$ in order for the genetic variance to remain bounded, hence C_{ii} terms scale as a^2 or n^{-1}. Summing over all n loci, M_2 is of order $n \cdot n^{-1} = 1$ and, as required, remains bounded as the number of loci increases. What happens to the skew and kurtosis as n increases? Since a is of order $n^{-1/2}$, $\mu_{3,i}$ is of order a^3 or $n^{-3/2}$, implying that M_3 is of order $n \cdot n^{-3/2} = n^{-1/2}$. Hence, as the number of loci becomes very large, the contribution from skew at individual loci becomes negligible. Likewise, $\mu_{4,i}$ is of order $n^{-4/2}$, implying M_4 is of order $n \cdot n^{-2} = n^{-1}$. As with skew, changes in kurtosis generated by within-locus (i.e., allele frequency) changes become negligible as the number of loci becomes (very) large.

The behavior of the between-locus contributions, however, is quite different (Turelli and Barton 1990). Under weak selection, Turelli and Barton show that C_{ijk} is proportional to $C_{ii}C_{jj}C_{kk}$ and is thus of order n^{-3}. However, there are $n(n-1)(n-2) \simeq n^3$ terms involving C_{ijk} in the covariance contribution to skew. Thus, this contribution is of order one and does not necessarily converge to zero even as the number of loci approaches infinity. The same argument holds for the kurtosis and higher moments (Turelli and Barton 1990). Thus, when the number of loci is very large, the distribution of genotypic values can depart from Gaussian due to selection generating third and higher-order covariances between loci, which in turn creates skew and kurtosis in the genotypic distribution. Even if the distribution of genotypes is originally Gaussian, selection generates these higher-order disequilibria, driving the distribution away from a Gaussian (Bulmer 1980; Zeng 1987; Turelli and Barton 1990, 1994).

Describing the Genotypic Distribution: Cumulants and Gram-Charlier Series

While most readers are familiar with moments, an alternate approach to describing the shape of a distribution, and in particular how it departs from a Gaussian, is to examine the **cumulants** of that distribution. The first use of cumulants in examining selection response appears to be O'Donald (1972) and Bulmer (1980). Sophisticated (and highly technical) treatments using cumulants have been developed by Bürger (1991, 1993) and Turelli and Barton (1994).

Our aim here is to both give the fearless reader sufficient background to access this literature and to show the connection between results using moments and those using cumulants.

Cumulants (the n-th of which we denote K_n) arise naturally in series approximations of probability distributions and are related to the central moments (μ_n). For example, the first five central moments can be expressed as functions of the cumulants as follows:

$$\mu_1 = K_1, \quad \mu_2 = K_2, \quad \mu_3 = K_3, \quad \mu_4 = K_4 + 3K_2^2, \quad \mu_5 = K_5 + 10\,K_2K_3$$

Hence, the first three cumulants are equal to the mean, variance, and skew, while the fourth and fifth cumulants are

$$K_4 = \mu_4 - 3\mu_2^2, \quad K_5 = \mu_5 - 10\,\mu_2\,\mu_3 \tag{24.20}$$

The major advantage of cumulants over moments is that they are *additive*, so that the n-th cumulant of a sum of random variables is just the sum of the cumulants for each, i.e., $K_n(x+y) = K_n(x) + K_x(y)$. This linearity property does not hold for higher order moments, which are highly nonlinear functions of the moments of the individual distributions.

The major disadvantage of using cumulants (in place of moments) is when dealing with recombination (Turelli and Barton 1994; Bürger 2000). In such cases, one works with cumulants to compute within-generation changes, converts these to moments for recombination, and then converts the recombinant products back into cumulants.

Example 24.5. Use cumulants to compute the fourth and fifth central moments of the phenotypic distribution. Here, $z = G + e$, so that the fourth moment is

$$\begin{aligned}
\mu_{4,z} &= K_{4,z} + 3K_{2,z}^2 \\
&= (\,K_{4,G} + K_{4,e}\,) + 3\,(K_{2,G} + K_{2,e})^2 \\
&= \left(\,\mu_{4,G} - 3\mu_{2,G}^2\,\right) + \left(\,\mu_{4,e} - 3\mu_{2,e}^2\,\right) + 3\,(\,\mu_{2,G} + \mu_{2,e}\,]^2 \\
&= \mu_{4,G} + \mu_{4,e} + 6\sigma_G^2\sigma_e^2
\end{aligned}$$

where the second and third steps follow from the additivity propery of cumulants ($K_{n,z} = K_{n,G} + K_{n,e}$) and Equation 24.20, respectively. Simplifying recovers Equation 24.15. Likewise,

$$\begin{aligned}
\mu_{5,z} &= K_{5,z} + 10K_{2,z}K_{3,z} \\
&= (\,K_{5,g} + K_{5,e}\,) + 10(K_{2,g} + K_{2,e})(K_{3,g} + K_{3,e}) \\
&= (\,\mu_{5,g} - 10\mu_{2,g}\mu_{3,g}\,) + (\,\mu_{5,e} - 10\mu_{2,e}\mu_{3,e}\,) \\
&\qquad + 10(\mu_{2,g} + \mu_{2,e})(\mu_{3,g} + \mu_{3,e}) \\
&= \mu_{5,g} + \mu_{5,e} + 10\,(\mu_{2,g}\mu_{3,e} + \mu_{2,e}\mu_{3,g})
\end{aligned}$$

These nonlinear expressions for the higher moments of a sum of variables is in sharp contrast to the expressions for cumulants, wherein $K_{n,z} = K_{n,g} + K_{n,e}$.

Example 24.6. If the underlying genes are additive, the n-th cumulant of the genotypic distribution is the sum of the appropriate cumulants for each of the underlying loci. To see the advantage of working with cumulants, consider the fourth cumulant of the genotypic distribution. Turelli and Barton (1994) show that this can be written as a sum of the fourth-order cumulants at the underlying loci,

$$K_{4,g} = \sum_{i,j,k,l} K_{ijkl} = \sum_i K_{iiii} + \sum_i \sum_{j,k,l \neq i} K_{ijkl}$$

the sum over K_{iiii} represents the within-locus contributions to the fourth cumulant, while the sum over the other indicies are the contributions to K_4 from fourth-order disequilibrium generation associations between loci. This recovers Equation 24.19d by noting the $\mu_{4,g} = M_4 + 3\sigma_A^4$ and that

$$K_{ijkl} = C_{ijkl} - C_{ij}\,C_{kl} - C_{ik}\,C_{jl} - C_{il}\,C_{jk} \quad \text{and} \quad M_4 = \sum_i K_{iiii}$$

A second advantage of working with cumulants is that the third and higher cumulants for a normal random variable are zero, so nonzero values for these higher cumulants provide a convient measure of departures from normality. Bulmer (1980) showed for the infinitesimal model (with unlinked loci) that, following the relaxation of selection, the j-th cumulant is decreased each generation by $(1/2)^{j-1}$, so that the distribution rapidly returns to a normal.

Finally, cumulants appear in series approximations of arbitary probability distributions. Consider a standardized random variable $y = (z - \mu)/\sigma$, which has mean zero and variance one. If the true density function for y is $\phi(y)$, we can approximate it as a unit normality density function $\varphi(y)$ plus correction terms. In particular, the **Gram-Charlier series** approximation (here to order five) is given by:

$$\phi(y) \simeq \varphi(y)\left[1 + \frac{K_3}{6}\,H_3(y) + \frac{K_4}{24}\,H_4(y) + \frac{K_5}{120}\,H_5(y)\right] \tag{24.21a}$$

where H_k denotes the **Chebsyshev-Hermite polynominal** of order k, with

$$\begin{aligned} H_3(x) &= x^3 - 3x \\ H_4(x) &= x^4 - 6x^2 + 3 \\ H_5(x) &= x^5 - 10x^3 + 15x \end{aligned} \tag{24.21b}$$

Equation 24.21 shows how the higher-order cumulants (K_3 and above) quantify departures for normality. If all of these are zero, the distribution is gaussian.

Bulmer (1980), Zeng (1987), and Turelli and Barton (1994) have used Gram-Charlier series to examine departures from normality under selection. Further properties of cumulants and Gram-Charlier (and other) series approximations are discussed in Johnson and Kotz (1970) and Kendall and Stuart (1977).

Application: Departure From Normality Under Truncation Selection

One application of this machinery is to compute the distribution of breeding values following a single generation of truncation selection, assuming that in the initial population the joint distribution of phenotypic and breeding values is multivariate normal. This example was considered by both Bulmer (1980) and Zeng (1987). Turelli and Barton (1994) present a very elegant (and elaborate) analysis for multiple generations. As before, we consider only additive models, so the distribution of genotypes is the distribution of breeding values.

First, let's just make the assumption of normality in phenotypic values, $z \sim N(\mu, \sigma_z^2)$, and compute the cumulants for the resulting distribution of phenotypic values after truncation selection. As before (Chapters 10, 13), we choose the uppermost p percent of the population, giving a selection intensity of

$$\overline{i} = \frac{\varphi(z_p)}{p}, \quad \text{where} \quad \Pr(U > z_p) = p$$

where U is a unit normal random variable (Equation 10.26a). From Chapters 10 and 13, we already have the first two cumulants following selection as

$$\mu^* = \mu + \overline{i}\,\sigma_z, \qquad \text{and} \qquad \sigma_z^{2\,*} = [1 - \overline{i}(\overline{i} - z_p)]\ \sigma_z^2$$

while the next two cumulants are

$$K_{3,z}^* = [(\overline{i} - z_p)(2\,\overline{i} - z_p) - 1]\,\overline{i}\sigma_z^3 \tag{24.22a}$$

$$K_{4,z}^* = \left[-6\,\overline{i}\,(\overline{i} - z_p)^2 + (3 - z_p^2)(\overline{i} - z_p) + \overline{i}\right]\ \overline{i}\sigma_z^4 \tag{24.22b}$$

We now need to translate these within-generation changes in the *phenotypic* distribution first to the within-generation change in the distribution of *breeding values* and then examine how this distribution of breeding values changes following random mating (assuming unlinked loci). Both these steps rely critically on assumptions of normality. If the distribution of breeding and phenotypic values is bivariate normal before selection (as would occur in the initial round of selection, but not necessarily in subsequent rounds), then the regression of breeding values on phenotypic values is linear,

$$A = \mu_z + h^2(z - \mu_z) + e$$

Rao et al (1968) show that truncation does not alter the regression, from which follows our standard results (Chapters 10, 13) for the mean breeding value and its variance following selection

$$\mu_A^* = \mu_z + h^2\overline{i}\sigma_z$$

and

$$\sigma_{A*}^2 = \sigma_z^2\left[1 - h^2\,\overline{i}\,(\,\overline{i} - z_p)\right]$$

All higher cumulants follow a very simple relationship between the breeding and phenotypic values,

$$K_{r,A}^* = (h^2)^r\,K_{r,z}^* \tag{24.23}$$

Following Bulmer (1980), if the regression of breeding value on phenotype is normal, then after selection the resulting cumulants for the distribution of breeding values become

$$K_{r,A}^* = \left(h^2\right)^r K_{r,z}^* \qquad \text{for} \quad r \geq 3 \tag{24.24a}$$

Likewise, (assuming unlinked loci), the cumulants for the distribution of breeding values in the next generation becomes

$$K_{r,A}(t+1) = \left(\frac{1}{2}\right)^{r-1} K_{r,A}^*(t) \tag{24.24b}$$

Hence, the cumulants for the distribution of breeding values at the start of the next generation are related to the cumulants of the post-selection phenotypic distribution by

$$K_{r,A}(t+1) = 2\left(\frac{h^2}{2}\right)^r K_{r,z}^*(t) \tag{24.24c}$$

Notice that after a single generation of selection, $K_{3,A}$ and $K_{4,A}$ are non-zero, and hence after one generation of selection the distribution of breeding values is no longer normal. At this point, the assumption of bivariate normality no longer holds and there is no longer

a simple relationship between $K^*_{r,A}$ and $K^*_{r,z}$. Thus, we can not simply iterate the above procedure over more than one generation of selection. See Turelli and Barton (1994) for a detailed analysis over multiple generations.

Example 24.7. Suppose truncation selection occurs on a normally-distributed trait with initial mean $\mu_z = 0$ and initial variance $\sigma_z^2 = 100$. Suppose the upper 5% is saved, so that $\overline{\imath} = 2.063$ and $z_p = 1.645$ (Example 10.10). Finally, to show an extreme case we first assume $h^2 = 1$, so that all variance is additive-genetic. Applying Equation 24.22a, the resulting third-order cumulant in the phenotypic distribution following selection is

$$\begin{aligned} K^*_{3,z} &= [(\overline{\imath} - z_p)(2\,\overline{\imath} - z_p) - 1]\,\overline{\imath}\sigma_z^3 \\ &= [(2.063 - 1.645)(2 \cdot 2.063 - 1.645) - 1] \cdot 2.063 \cdot 100^{3/2} \\ &= 76.45 \end{aligned}$$

Applying Equation 24.24c translates this into the third cumulant in the genotypic distribution in the next generation, as

$$K_{3,A}(t+1) = 2\left(\frac{h^2}{2}\right)^3 K^*_{3,z}(t) = 2\left(\frac{1}{2}\right)^3 76.45 = 19.11$$

Using the machinery from Chapter 13 gives the phenotypic variance in this generation as $\sigma_A^2 = \sigma_z^2 = 56.9$, Thus, the scaled skew becomes

$$\gamma_3 = \frac{K_3}{\sigma^3} = \frac{19.11}{56.9^{3/2}} = 0.045$$

A similar calculation gives $K_4 = 59.7$ and $\gamma_4 = 0.018$. The resulting (fourth-order) Gram-Charlier series approximation for the distribution $\phi(A)$ of breeding values in generation 1 is

$$\begin{aligned} \phi(A) &\simeq \varphi(A)\left[1 + \frac{0.045}{6}\,H_3(A) + \frac{0.018}{24}\,H_4(A)\right] \\ &= \varphi(A)\,[1 + 0.0075\,H_3(A) + 0.0075\,H_4(A)] \end{aligned}$$

where $\varphi(x)$ is the normal distribution and the $H_i(x)$ are defined by Equation 24.21b. The key point is that the resulting distribution is only very weakly perturbated away for a Gaussian. Further note that this is the most extreme case, as we have assume $h^2 = 1$. For a more typically heritability, say $h^2 = 0.3$, similar calculations give $\gamma_3 = 0.0039$ and $\gamma_4 = 0.0007$, and

$$\phi(A) \simeq \varphi(A)\,[1 + 0.00065\,H_3(A) + 0.00003\,H_4(A)]$$

so that the departure from a normal is very small indeed. Thus, under the infinitesimal model, the generation of linkage disequilibrium under truncation selection has very little impact on driving the distribution of breeding values away for a Gaussian. This point was initially made by Bulmer (1980). The much more extensive analysis by Turelli and Barton (1994) shows that the Bulmer equation (Equation 13.7b), even in the presence of strong truncation selection, can be used with little error. Turelli and Barton's analysis assumes a sufficiently number of loci so that changes (in both the genic variance and in cumulants or order three or higher) can be ignored. Thus, while the disequilibrium introduced by truncation selection can indeed drive a distribution of breeding values away from a strict gaussian, the error is assuming it remains gaussian in small.

Short-term Response Ignoring Linkage Disequilibrium

With the above machinery in hand, we are now ready to examine the response to selection under non-Gaussian genotypic distributions. We first consider the situation where a small to modest number of loci underlie the character so that most of the changes in the higher-order moments are due to changes in allele frequencies, rather than through generation of gametic-phase disequilibrium. Our treatment follows that of Barton and Turelli (1987).

If we are willing to assume additivity across loci and gametic-phase equilibrium, then genetic changes in the character can be completely described by the dynamics of allele frequency changes at each locus. The complete dynamics for a locus with k alleles is described by the $k-1$ allele frequency change equations. Alternatively, we could completely describe the dynamics by using equations based on any set of $k-1$ independent new variables that can be expressed as functions of allele frequencies (this is the standard multivariate transformation problem of vector calculus and requires that the determinant of the Jacobian transformation matrix is nonzero). Barton and Turelli show that one such set of new variables are the first $k-1$ moments of the allelic distribution. This is the motivation behind their approach, which focuses on *allelic moments* rather than *allelic frequencies*. If we ignore gametic-phase disequilibrium, then for n loci with k alleles each, we can completely describe the dynamics by using the first $n(k-1)$ moments of the genotypic distribution. This same approach can be used when linkage is considered, but the number of equations increases dramatically. While this approach of using a new set of variables is exact, it is also as fruitless to solve as the original set of allele-frequency change equations. The hope, however, is that by considering the first few moments we can gain considerable insight into the actual dynamics.

To briefly sketch the approach used by Barton and Turelli, we first express Wright's formula for multiple alleles (Equation 5.11) as

$$\Delta p_i = \sum_j G_{ij} \frac{\partial \ln \overline{w}}{\partial p_j} \tag{24.25a}$$

where $G_{ii} = p_i(1-p_i)/2$ and $G_{ij} = -p_i\, p_j/2$ (for $i \neq j$). Recall that Wright's formula holds as long as fitnesses are frequency-independent, which can be violated with constant fitness by linkage disequilibrium (Example 5.7). Now consider a function $f(p_1, p_2, \ldots\ p_{k-1})$ that depends on the allele frequencies at this locus, such as a particular moment of the allelic distribution. The change in f due to changes in allele frequencies can be approximated by a Taylor series to give

$$\Delta f = \sum_i \frac{\partial f}{\partial p_i}\, \Delta p_i + \frac{1}{2} \sum_i \sum_j \frac{\partial^2 f}{\partial p_j\, \partial p_i}\, \Delta p_i\, \Delta p_j + \cdots \tag{24.25b}$$

where we have ignored higher terms of Δp_i. Substituting for Δp_i using Equation 24.25a, and the chain rule of differentiation identity

$$\frac{\partial \ln \overline{w}}{\partial p_j} = \frac{\partial \ln \overline{w}}{\partial f}\, \frac{\partial f}{\partial p_j}$$

gives (to first order)

$$\begin{aligned} \Delta f &\simeq \sum_i \frac{\partial f}{\partial p_i} \sum_j g_{ij}\, \frac{\partial \ln \overline{w}}{\partial p_j} \\ &= \frac{\partial \ln \overline{w}}{\partial f} \sum_i \sum_j \frac{\partial f}{\partial p_i}\, g_{ij}\, \frac{\partial f}{\partial p_j} \end{aligned} \tag{24.25c}$$

This is a *weak-selection approximation*, as it assumes terms of order $\Delta p_i \, \Delta p_j$ and higher can be ignored (if drift is considered, these second-order terms must be included even if selection is weak, see Turelli 1988). Using this expression to consider changes in the allelic moments yields a set of equations where changes in a certain moment depend on higher order moments. After considerable algebra (see Barton and Turelli 1987 for details), the changes in genotypic moments (under the assumptions of complete additivity and gametic-phase equilibrium) can be expressed in matrix form as

$$\boldsymbol{\Delta}_{\mu_G} \simeq \mathbf{M} \, \nabla \ln \overline{w} \tag{24.26}$$

where

$$\boldsymbol{\Delta}_{\mu_G} = \begin{bmatrix} \Delta\mu_{1,G} \\ \Delta\mu_{2,G} \\ \Delta\mu_{3,G} \\ \vdots \end{bmatrix}, \qquad \nabla \ln \overline{w} = \begin{bmatrix} \dfrac{\partial \ln \overline{w}}{\partial \mu_{1,z}} \\ \dfrac{\partial \ln \overline{w}}{\partial \mu_{2,z}} \\ \dfrac{\partial \ln \overline{w}}{\partial \mu_{3,z}} \\ \vdots \end{bmatrix}$$

are (respectively) the vector of changes in the genotypic moments and of partial derviatives of long mean fitness with respect to each moment and the matrix $\mathbf{M}$, where

$$\mathbf{M} = 2 \sum_i \begin{bmatrix} \mu_{2,i} & \mu_{3,i} & (\mu_{4,i} - 3\mu_{2,i}^2) & \cdots \\ \mu_{3,i} & (\mu_{4,i} - \mu_{2,i}^2) & (\mu_{5,i} - 4\mu_{3,i}\,\mu_{2,i}) & \cdots \\ (\mu_{4,i} - 3\mu_{2,i}^2) & (\mu_{5,i} - 4\mu_{3,i}\,\mu_{2,i}) & (\mu_{6,i} - \mu_{3,i}^2 - 6\mu_{2,i}\,\mu_{4,i} + 9\mu_{2,i}^2) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

The elements of $\mathbf{M}$ corresponding to selection on the fourth and higher moments are more complicated than may be suggested by the simple dots in the matrix due to the nonadditive nature of higher moments. Expressions based on $\partial \ln \overline{w} \,/\, \partial K_{i,z}$ (the partial derviative of fitness with respect to the i-th cumulant of the phenotypic distribution) have a simpler form due to the additive nature of cumulants (Bürger 1991, 1993; Turelli and Barton 1994), but these still have the undesirable feature that the response of the i-th cumulant depends on cumulants of higher order.

For example, if selection on the first three phenotypic moments accounts for the majority of selection, Turelli and Barton found that the expected single-generation change in mean is

$$\Delta\mu_z \simeq \sigma_A^2 \, \frac{\partial \ln \overline{w}}{\partial \mu_z} + \mu_{3,G} \, \frac{\partial \ln \overline{w}}{\partial \mu_{2,z}} + k_4 \, \sigma_A^4 \, \frac{\partial \ln \overline{w}}{\partial \mu_{3,z}} \tag{24.27}$$

where $k_4 = (\mu_{4,G} - 3\sigma_A^4)/\sigma_A^4$ is the scaled coefficient of kurtosis, which is zero if G is Gaussian. If G is indeed Gaussian, $\mu_{3,z} = k_4 = 0$, and we recover the selection gradient version of the breeders' equation (Equation 10.23c). Thus predicting changes in even the simplest genotypic moment, the mean, requires a detailed knowledge of both higher order allelic moments ($\mu_{k,i}$) and the nature of selection on these higher order moments ($\partial \ln \overline{w} \,/\, \partial \mu_{k,z}$). In order to proceed further, we have to make additional assumptions about the distribution of allelic effects at *individual loci*.

Example 24.8. Assume the continuum-of-allele approximation, so that the distribution of allelic effects at each locus is normal. In this case all odd central moments at each locus are zero ($\mu_{2k+1} = 0$) and all even moments are related to the second moment by $\mu_{2k} = \mu_2^k\,(2k)!/(2^k k!)$ (Kendall and Stewart 1977). For example, $\mu_4 = 3\mu_2^2$ so that $\mu^4 - \mu_2^2 = 2\mu_2^2$. Assuming that most of selection is on the mean and variance, we can neglect the third and higher-order selection gradients. In this case, $\mathbf{M}$ becomes the 2×2 matrix

$$\mathbf{M} = \begin{pmatrix} 2\sum_{i=1}^{n} \mu_{2,i} & 0 \\ 0 & 4\sum_{i=1}^{n} \mu_{2,i}^2 \end{pmatrix} = \begin{pmatrix} \sigma_A^2 & 0 \\ 0 & \dfrac{\sigma_A^4}{n_e} \end{pmatrix}$$

where

$$n_e = \frac{\sigma_A^4}{4\sum_i \mu_{2,i}^2}$$

is equivalent to Chevalet's (1994) effective number of loci (Equation 24.3), see Example 24.9. The expected response in the genotypic mean and variance becomes

$$\begin{pmatrix} \Delta\mu \\ \Delta\sigma_A^2 \end{pmatrix} \simeq \begin{pmatrix} \sigma_A^2 & 0 \\ 0 & \sigma_A^4/n_e \end{pmatrix} \begin{pmatrix} \dfrac{\partial \ln \overline{w}}{\partial \mu_z} \\ \dfrac{\partial \ln \overline{w}}{\partial \sigma_z^2} \end{pmatrix} = \begin{pmatrix} \sigma_A^2 \dfrac{\partial \ln \overline{w}}{\partial \mu_z} \\ \dfrac{\sigma_A^4}{n_e} \dfrac{\partial \ln \overline{w}}{\partial \sigma_z^2} \end{pmatrix}$$

If the phenotypic distribution is *exactly* normal, since all moments can be expressed in terms of the mean and variance, only gradients measuring selection on the mean and variance appear and these equations are exact. Recalling Equation 24.12 gives

$$\Delta\mu \simeq h^2 S$$

and

$$\Delta\sigma_A^2 \simeq \frac{h^4}{2n_e}\left(\delta(\sigma_z^2) + S^2\right)$$

Thus the expected change in the mean follows the breeders' equation and short-term changes in variance are expected to be small if there are many loci of small effect (so that n_e is large). We remind the reader that this ignores the effects of gametic-phase disequilibrium.

Note that since $\mu_{2,i}$ changes as allele frequencies change, predicting changes in variance over several generations even under the extreme simplifying assumptions leading to these equations still requires a detailed knowledge about the distribution of allelic effects *at individual loci*. Thus, while short-term changes in the mean can be predicted without detailed knowledge of the underlying genetics (only σ_A^2 is required, which can be estimated from phenotypic resemblance between relatives), changes in variance cannot (unless an estimate of $\sum \mu_{2,i}^2$ can be obtained). Example 24.5 shows just how unpredictable changes in n_e over time can be.

Finally, let's try to connect these results for the change in the genic variance with those obtained under the continuum-of-alleles approximation (Equation 24.2a). Since we ignore any diseqiulibrium, $\Delta\sigma_A^2 = \Delta\sigma_a^2$. Ignoring drift, if the within-generation change in the phenotypic variance is $\delta(\sigma_z^2) = -\kappa\sigma_z^2$, then the COA approximation for the change in genic variance is

$$\Delta\sigma_a^2 = -\frac{\kappa\, h^2(t)\, \sigma_A^2(t)}{2n_e}$$

Since $\kappa h^4 \sigma_z^2 = \kappa h^2 \sigma_A^2$, by contrast, this allelic-moment approximation gives

$$\Delta\sigma_A^2 \simeq \frac{h^4}{2n_e}\left(\delta(\sigma_z^2) + S^2\right) = -\frac{\kappa h^2 \sigma_A^2}{2n_e} + \frac{h^4 S^2}{2n_e}$$

Thus, the allelic-moment approximation has an additional term relative to the COA approximation.

Example 24.9. Here we show that n_e, as defined in the previous example, is equivalent to Chevalet's (1994) n_e (Equation 24.3). This simply clears up a technical detail and can be skipped by the causal reader. Specifically, we need to show that

$$\frac{n}{1+cv^2} = \frac{\sigma_A^4}{4\sum_i \mu_{2,i}^2}$$

Since $\mu_{2,i}$ is the variance of allelic effects at locus i, the genic variance contributed by locus i (since there are two alleles) is $2\mu_{2,i}$. Thus,

$$1 + cv^2 = 1 + \frac{\sigma^2(2\mu_{2,i})}{E[\,2\mu_{2,i}\,]^2} = \frac{E[\,2\mu_{2,i}\,]^2 + \sigma^2(2\mu_{2,i})}{E[\,2\mu_{2,i}\,]^2}$$

Recalling that $\sigma^2(x) = E[\,x^2\,] - E[\,x\,]^2$, we have

$$\sigma^2(2\mu_{2,i}) = \frac{1}{n}\sum_i^n (2\mu_{2,i})^2 - E[2\mu_{2,i}]^2, \quad \text{hence} \quad E[2\mu_{2,i}]^2 + \sigma^2(2\mu_{2,i}) = \frac{1}{n}\sum_i^n (2\mu_{2,i})^2$$

Summing the genic variances at each locus gives the total genic variance (which is the additive variance as we are ignoring disequilibrium),

$$\sigma_A^2 = \sum_i^n 2\mu_{2,i} = n\,E[2\mu_{2,i}], \qquad \text{hence} \qquad E[2\mu_{2,i}]^2 = \frac{\sigma_A^4}{n^2}$$

Hence,

$$1 + cv^2 = \frac{E[\,2\mu_{2,i}\,]^2 + \sigma^2(2\mu_{2,i})}{E[\,2\mu_{2,i}\,]^2} = \frac{(1/n)\sum_i^n (2\mu_{2,i})^2}{\sigma_A^4/n^2} = \frac{4n\sum_i^n \mu_{2,i}^2}{\sigma_A^4}$$

Thus,

$$n_e = \frac{n}{1+cv^2} = \frac{n\sigma_A^4}{4n\sum_i^n \mu_{2,i}^2} = \frac{\sigma_A^4}{4\sum_i^n \mu_{2,i}^2}$$

If phenotypes are approximately normally-distributed (but allelic effects at individual loci are not necessarily Gaussian), the mean and variance terms of the selection gradient vector generally dominate. Considering only the first three genotypic moments, Equation 24.26 reduces to

$$\Delta\mu_G \simeq h^2 S + \left(\frac{\delta(\sigma_z^2) + S^2}{2\sigma_z^4}\right) M_3 \tag{24.28a}$$

$$\Delta\sigma_A^2 \simeq \frac{S}{\sigma_z^2} M_3 + \left(\frac{\delta(\sigma_z^2) + S^2}{\sigma_z^4}\right) \sum_i^n \left(\,\mu_{4,i} - \mu_{2,i}^2\,\right) \tag{24.28b}$$

$$\Delta\mu_{3,G} \simeq \frac{S}{\sigma_z^2} M_4 + \left(\frac{\delta(\sigma_z^2) + S^2}{\sigma_z^4}\right) \sum_i^n \left(\,\mu_{5,i} - 4\,\mu_{3,i}\,\mu_{2,i}\,\right) \tag{24.28c}$$

where M_3 and M_4 are as defined by Equations 24.18c and 18d. As discussed in Chapters 28 and 29, when selection acts only on the mean, $\delta(\sigma_z^2) = -S^2$, so that the first term in each of these three equations accounts for the effect of selection to change the mean and the second term accounts for the effect of selection on the variance. Note that we obtained Equation 24.28a previously by an alternative approach (Equation 5.44b). When the genotypic distribution is skewed ($M_3 \neq 0$), the single-generation change in the mean also depends on the nature of selection on the variance (O'Donald 1968, 1972; Bulmer 1980; Gillespie 1984; Barton and Turelli 1987; Mitchell-Olds and Shaw 1987). Further, even if skew is initially absent, Equation 24.28c shows that if the kurtosis of the genotypic distribution differs from that expected for a Gaussian ($M_4 \neq 0$), selection strictly on the mean generates skew. Thus, even ignoring the effects of gametic-phase disequilibrium, selection on the mean generates skew when the genotypic distribution displays kurtosis.

Which factor, allele frequency changes at the individual loci or the generation of gametic-phase disequilibrium, is more important at producing departures from normality depends on the number of underlying loci. Turelli and Barton (1990) simulated a character controlled by eight diallelic loci and found that most of the skew and kurtosis generated by selection was generated by allele frequency changes, while the contribution from third and fourth-order disequilibrium was quite small. Thus, when the number of loci is small, the relative contribution to the third and higher moments from linkage effects may be largely negligible, and the error by using Equation 24.26 (which assumes gametic-phase equilibrium) should be small. As the number of (equivalent) loci increases, within-locus effects make a smaller and smaller contribution, with departures from normality caused by disequilibrium eventually dominating as the number of loci becomes sufficiently large.

Short-term Response Ignoring Allele Frequency Change

The last section considers one class of approximations for short-term response without assuming gaussian distribution, namely focusing solely on allele frequency changes. Here we consider the converse, namely assuming a large enough number of loci such that allele frequency changes (over our time span of interest) can be ignored. Thus, all of the changes are due to selection generating disequilibrium. The difference from infinitesimal model is that we no longer make any Gaussian assumptions.

Turelli and Barton (1990, 1994) extended basic moment analysis (Equation 24.26) to allow for gametic-phase disequilibrium, by considering both within-locus moment changes due to allele frequency changes and between-locus contributions generated by disequilibrium. The 1994 paper is the more general of the two, with the analysis based on the cumulants of the distribution. While the mean, variance, and skew are equivalent to the first three cumulants, cumulants of order four and higher provide much more compact expressions than using moments, due to the additivity of cumulants versus the non-linear nature of higher order moments.

In parallel with their moments analysis, Turelli and Barton define the gradients of selection associated with the i-th cumulant of the phenotypic distribution $K_{z,i}$ by

$$L_i = \frac{\partial \ln(\overline{W})}{\partial K_{z,i}} \tag{24.29a}$$

L_1 and L_2 correspond to selection on the mean and variance, while L_i for $i \geq 3$ represents selection that drives the distribution away from normality (as cumulants of order three and higher are zero for a gaussian). Turelli and Barton present general expressions for the change in all the cumulants of the distribution. In particular, for a large number of loci, they show that if the majority of selection is on the first four cumulants of the distribution, the change

in the mean and variance is given by

$$\Delta\mu = \sigma_A^2\, L_1 + K_{g,3}\, L_2 + K_{g,4}\, L_3 + K_{g,5}\, L_4 \tag{24.29b}$$

$$\begin{aligned}\Delta\sigma_A^2 = {} & \frac{\sigma_a^2 - \sigma_A^2}{2} - \frac{\Delta\mu^2}{2} + \frac{K_{g,3}}{2} L_1 + \left(\sigma_A^4 + \frac{K_{g,4}}{2}\right) L_2 \\ & + \left(3\sigma_A^2 K_{g,3} + \frac{K_{g,5}}{2}\right) L_3 + \left(3K_{g,3}^2 + 4\sigma_A^2 K_{g,4} + \frac{K_{g,6}}{2}\right) L_4\end{aligned} \tag{24.29c}$$

where $K_{g,i}$ denotes the i-th cumulant of the genotypic distribution. Note for Equation 24.29b that if some cumulants of order three or higher are nonzero, selection to alter the higher-order culumants of the distribution (and hence drive the distribution away from a Gaussian) also results in a change in the mean. Further note the appearance of the *genic* variance σ_a^2 in Equation 24.29c. We are assuming (at least over our time scale) that allele frequency change can be ignored and hence this is a constant. All changes in the variance (and higher order moments/cumulants) are thus assumed to arise entirely from selection generating disequilibrium.

Example 24.10. If phenotypes are normally-distributed, since the first two cumulants equal the mean and the variance, recalling Equation 24.12a and b gives

$$L_1 = \frac{S}{\sigma_z^2}, \qquad L_2 = \frac{\delta(\sigma_z^2) + S^2}{2\,\sigma_z^4}$$

If the genotypes follow a normal distribution, then $K_{g,i} = 0$ for $i \geq 3$. In this case, Equation 24.29b reduces to

$$\Delta\mu = \sigma_A^2\, \frac{S}{\sigma_z^2} = h^2 S,$$

recovering the breeders' equation. Recalling that $\sigma_A^2 = \sigma_a^2 + d$, Equation 24.29c reduces to

$$\begin{aligned}\Delta\sigma_A^2 &= \frac{\sigma_a^2 - \sigma_A^2}{2} - \frac{(h^2 S)^2}{2} + \sigma_A^4 \left(\frac{\delta(\sigma_z^2) + S^2}{2\,\sigma_z^4}\right) \\ &= -\frac{d}{2} + \frac{h^4}{2}\,\delta(\sigma_z^2)\end{aligned}$$

and we recover Bulmer's formula for the short-term change in the variance (Equation 13.7b). Notice that there is no change in the genic variance, as a very large number of loci of small effect is assumed.

Turelli and Barton (1994) examined the effects of both strong truncation (directional) selection and strong disruptive selection on Gaussian (Infinitesimal and COA) models when the number of loci is large. For strong truncation selection, they found that while selection does indeed generate nonzero culumants of order three and higher (and hence departures from normality), these are generally quite small (e.g., Example 24.7). As a result, the breeder's equation with the variance changes predicted from Bulmers model give quite accurate results for the predicted change in the mean and variance. Hence, the effects of disequilibrium in this case are essentially all accounted for by considering only the second-order disequilibrium,

which is done in the basic Bulmer model. Barton and Turelli found that the distribution of genotypic values is highly non-normal under strong disruptive selection, with a significant fourth cumulant (kurtosis) being generated by significant fourth-order disequilibrium (generating correlations between groups of four loci). Surprisingly, even in this case the change in variance is well predicted by the basic Bulmer model (Equation 13.7b).

Effects of Linkage

As might be expected, when the above results are generalized to allow for linkage (as opposed to the expressions assuming unlinked loci that we present), they become rather complex (Turelli and Barton 1990 ,1994; Bürger 2000). However, when selection is weak, we can include linkage into an approximation for the asymptotic response for a generalized infinitesimal model that makes no assumptions about the distribution of genotypic values (Turelli and Barton 1990). In particular, higher-order genotypic moments can be expressed in terms of the initial additive variance in the absence of gametic-phase disequilibrium (the genic variance σ_a^2, which is assumed to be constant), giving the asymptotic response as approximately

$$\Delta\mu_z = \sigma_a^2 \left(\frac{\partial \ln \overline{W}}{\partial \mu_G} \right) + \frac{\sigma_a^4}{r_2} \left(\frac{\partial \ln \overline{w}}{\partial \mu_z} \cdot \frac{\partial \ln \overline{w}}{\partial \sigma_z^2} \right) + \frac{3\,\sigma_a^6}{2\,r_3} \left(\frac{\partial \ln \overline{w}}{\partial \sigma_z^2} \cdot \frac{\partial \ln \overline{w}}{\partial \mu_{z,3}} \right) \tag{24.30}$$

where r_2 and r_3 are the harmonic mean recombination rates (weighted by the allelic contributions at each locus) between pairs and triplets of loci. At equilibrium, the higher-order genotypic moments are constant as allele frequencies do not change, and with constant selection, covariances between loci are expected to approach constant values. Under these conditions, the expected change in mean following t generations of selection is just t times Equation 24.30.

Recalling Equations 24.7a and b, if phenotypes are approximately normally distributed, the asymptotic rate of response further reduces to

$$\Delta\mu_z = \frac{\sigma_a^2}{\widetilde{\sigma}_z^{\,2}} \left(S + \sigma_a^2 \frac{\widetilde{\delta}(\sigma_z^2) + S^2}{2\,\widetilde{\sigma}_z^{\,2}} \left[\frac{S}{r_2} + \frac{3\,\sigma_a^2}{2\,r_3} \frac{\partial \ln \overline{w}}{\partial \mu_{3,z}} \right] \right) \tag{24.31}$$

where $\widetilde{\sigma}_z^{\,2}$ is the equilibrium phenotypic variance and $\widetilde{\delta}(\sigma_z^2)$ is the equilibrium within-generation change in phenotypic variance due to selection. This generalizes Bulmer's results (and reduces to them when distributions are exactly Gaussian). Bulmer's approach corrects the breeders' equation for changes in the variance due to selection generating disequilibrium between pairs of loci. Equation 24.31 demonstrates that further corrections are required to account for the third (and higher-order) disequilibrium generated by selection.

SUMMARY: WHERE DOES ALL THIS MODELING LEAVE US?

While even short-term response can be notoriously difficult to predict, in part due to a large number of confounding factors for even single-generation response (e.g., Tables 10.1 and 14.7), prediction of response over many generations is an unobtainable goal unless one essentially knows all of the very fine (microscopic) genetic details of our trait, *including* the distribution of allelic effects and frequencies.

Selection generates unpredictability in response by changing allele frequencies and by generating disequilibrium. These changes, in turn, potentially generate unpredictability by (i) changing genetic variances in ways not predicable from simple knowledge of the base population variances (as occurs by allele frequency change, and (ii) by causing

the parent-offspring regression to be nonlinear. The later is typically caused by generating departures from bivariate normality for the parent-offspring regression. With a linear parent-offspring regression, changes in the mean follow from knowing the variance. If, however, this regression is not linear and homoscedastic (i.e., not gaussian), then higher order moments/cumulants are involved in predicting the response in the mean (Equations 24.27, 24.29b). For example, if skew is present, selection directly on the variance can result in a change in the mean (e.g, Equation 24.28a).

If we are willing to make simplifying assumptions, we can partly circumvent some of these issues. Under the strict infinitesimal model, allele frequencies do not change, and changes in variance are easily predicted from Bulmer's Equation (13.6), which follows changes in pair–wise disequilibrium generated by selection. *Provided* the parent-offspring regression remains normal, one can use the base population values of the variances to compute the chains in genetic variances out to arbitrary times and likewise use this to correct the heritability in the breeders' equation. The potential complication with this approach, *even* if we are willing to assume no (significant) allele frequency change (at least over our interval of interest) is that selection not only generates disequilibrium between pairs of loci, it also generates higher-order disequilibrium as well (e.g., Example 24.7). These are generated when there is either selection on the third (or higher) moments or cumulants, or if the distribution of breeding values has any nonzero cumulants above order two (as would occur if it is not normal). These changes by disequilibrium generally drive an initially gaussian distribution away from normality. However, in many cases (such as truncation selection), theses changes (while real) have a relatively minor impact on generating departures from normality.

At the other extreme are models where allele frequencies are changing fairly rapidly, typically because there are one (or more) alleles having significantly enough effects for selection on a trait to result in a reasonable rapid change in their frequencies. We examine these in detail in Chapter 25. The analysis of such models typically ignores any disequilibrium, although the quasi-equilibrium relationship (equation 24.4b) seen between the genetic variance and the disequilibrium d likely (approximately) holds in these models, allowing us to also include the effects of pairwise disequilibrium. However, changes in the third (and higher) moments of the genotypic distribution by disequilibrium are negligible relative to the changes in these moments generated by allele frequency change.

Gaussian continuum-of-alleles (COA) models attempt to capture *both* changes in allele frequencies and disequilibrium, but make the highly restrictive assumption that the distribution of allelic effects is gaussian *at each locus*. As long as this approximately holds, these models can track both changes in the genetic variance from selection and/or drift changing allele frequencies as well as the effects of both of these forces on pariwise disequilibrium (Equation 242a,b). However, these models spectacularly fail when major alleles are present. COA models always predict a *decrease* in the genic variance, but this can *increase* as rare favorable alleles increase in frequency. Thus occurs because , as these alleles increase, the gaussian distribution at their locus becomes more and more untenable.

So what do these models tell us about the behavior of real genetic systems? It is typically assumed that because a trait may have a large number of potentially underlying loci, that each has small effect. Indeed, the reality is more likely that there is a distribution of genic effects. Thus, a trait is (potentially) likely to have a few alleles of at least modest effects and a background constellation of loci with much smaller effects. The dynamics of the "major" alleles is largely driven by allele frequency changes, while the behavior of background polygenic effects is largely (at least on the short term) governed by disequilibrium. Such a system would have rapidly (and largely unpredictable) short-term response, while the intermediate response (following as the major alleles approach fixation) is closer to an infinitesimal model. Chapters 25-27 continues this discussion.

Literature Cited

Barton, N. H. and M. Turelli. 1987. Adaptive landscapes, genetic distances and the evolution of quantitiative characters. *Genet. Res.* 49: 157–173. [24]

Bulmer, M. G. 1974. Linkage disequilibrium and genetic variability. *Genet. Res.* 23: 281–289. [24]

Bulmer, M. G. 1976a. The effect of selection on genetic variability: a simulation study. *Genet. Res.* 28: 101–117. [24]

Bulmer, M. G. 1980. *The mathematical theory of quantitative genetics.* Oxford Univ. Press, NY. [24]

Bürger, R. 1991. Moments, cumulants, and polygenic dynamics. *J. Math. Biol.* 30: 199–213. [24]

Bürger, R. 1993. Predictions of the dynamics of a polygenic character under directional selection. *J. Theor. Biol.* 162: 487–513. [24]

Chevalet, C. 1988. Control of genetic drift in selected populations. *In* B. S. Weir, E. J. Eisen, M. M. Goodman, and G. Namkoong (eds.), *Proceedings of the second international conference on quantitative genetics,* pp. 379–394. Sinauer Assocs., Sunderland, MA. [24]

Chevalet, C. 1994. An approximate theory of selection assuming a finite number of quantitative trait loci. *Genetics, Selection, Evol.* 26: 379–400 [24]

Crow, J. F. and M. Kimura 1964. The theory of genetic loads. *Proced. of the XIth Inter. Congr. Genetics* [24]

Dawson, K. J. 1997. Linkage disequilibrium and the infinitesimal limit. *Theor. Pop. Biol.* 52: 127–154. [24]

Felsenstein, J. 1977. Multivariate normal genetic models with a finite number of loci. *In* E. Pollak, O. Kempthorne, and T. B. Bailey, Jr., (eds.), *Proceedings of the international conference on quantitative genetics*, pp. 227–246. Iowa State Univ. Press, Iowa. [24]

Felsenstein, J. 1979. Excursions along the interface between disruptive and stabilizing selection. *Genetics* 93: 773 – 795. [24]

Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Royal Soc. Edinburgh* 52: 399–433. [24]

Gavrilets, S., and A. Hastings. 1994. Dynamics of genetic variability in two-locus models of stabilizing selection. *Genetics* 138: 519–532. [24]

Gavrilets, S., and A. Hastings. 1994. Dynamics of polygenic variability under stabilizing selection, recombination, and drift. *Genet. Res. Camb.* 65: 63–74. [24]

Gillespie, J. H. 1984. Pleiotropic overdominance and the maintenance of genetic variation in polygenic traits. *Genetics* 107: 321–330. [24]

Hospital, F., and C. Chevalet. 1996. Interactions of selection, linkage and drift in the dynamics of polygenic characters. *Genet. Res. Camb.* 67: 77–87. [24]

Johnson, N. L., and S. Kotz. 1970. *Continuous univariate distributions – 1.* John Wiley & Sons, NY. [24]

Keightley, P. D. and W. G. Hill. 1987. Directional selection and variation in finite populations. *Genetics* 117: 573–582. [24]

Kendall, M., and A. Stuart. 1977. *The advanced theory of statistics. Vol. 1. Distribution theory*. 4th Ed. Macmillan, NY. [24]

Kimura, M. 1965. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Natl. Acad. Sci. USA* 54: 731–736. [24]

Kimura, M., and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49: 725-738. [24]

Lande, R. 1975. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.* 26: 221–235. [24]

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30: 314–334. [24]

Lande, R. 1977. The influence of the mating system on the maintenace of genetic variability in polygenic characters. *Genetics* 86: 485–498. [24]

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37: 1210–1226. [24]

Lange, K. 1978. Central limit theorems for pedigrees. *J. Math. Biology* 6: 59–66. [24]

Latter, B. D. H. 1970. Selection in finite popualtions with multiple alleles. II. Centripetal selection, mutation, and isoallelic variation. *Genetics* 66: 165–186. [14]

Matthysse, S., K. Lange, and D. K. Wagener. 1979. Continous variation casued by genes with graduated effects. *Proc. Natl. Acad. Sci. USA* 76: 2862–2865. [24]

Mitchell-Olds, T. and R. G. Shaw. 1987. Regression analysis of natural selection: statistical inference and biological interpretation. *Evolution* 41: 1149–1161. [24]

Mueller, J. P. and J. W. James. 1983. Effect on linkage disequilibrium of selection for a quantitative character with epistasis. *Theor. Appl. Genet.* 65: 25–30. [24]

Nagylaki, T. 1991. Error bounds for the fundamental and secondary theorems of natural selection. *Proc. Natl. Acad. Sci.* 88: 2402–2406. [24]

O'Donald, P. 1968. Measuring the intensity of selection. *Nature* 220: 197–198. [24]

O'Donald, P. 1972. Natural selection for a quantitative character over several generations. *Nature* 237: 113–114. [24]

Rao, B. R., M. L. Garg, and C. C. Li. 1968. Correlation between the sample variances in a singly truncated bivariate normal distribution. *Biometrika* 55: 433–436. [24]

Sorensen, D. A. and W. G. Hill 1983. Effects of disruptive selection on genetic variance. *Theor. Appl. Genet*. 65: 173–180. [24]

Turelli, M. 1988. Population genetic models for polygenic variation and evolution.*In* B. S. Weir, E. J. Eisen, M. M. Goodman, and G. Namkoong (eds.), *Proceedings of the second international conference on quantitative genetics,* pp. 601–618. Sinauer Assocs., Sunderland, MA. [24]

Turelli, M. and N. H. Barton. 1990. Dynamics of polygenic characters under selection. *Theor. Pop. Biol.* 38: 1–57. [24]

Turelli, M. and N. H. Barton. 1994. Genetic and statistical analyses of strong selection on polygenic traits: What, me normal? *Genetics* 138: 913–941. [24]

Zeng, Z.-B., 1987. Genotypic distribution at the limits of natural and artificial selection with mutation. *Theor. Pop. Biol.* 32: 90–113. [24]

Zhang, X.-S., and W. G. HIll. 2005. Predictions of patterns of response to artifical selection in lines derived from natural populations. *Genetics* 169: 411–425. [24]