

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 30 April 2014

The assumption of normality (and therefore a homoscedastic linear parent-offspring regression) underlies most models of selection on quantitative traits, as the single-generation response can be predicted from the appropriate variance components (Chapters 6, 13, 16). This is in sharp contrast to response under the one- and two-locus models, where detailed knowledge of the underlying genotype frequencies and effects is required (Chapter 5). Under the **infinitesimal model**, which assumes a very large (effectively infinite) number of loci each with very small (effectively infinitesimal) effects, the genotypic distribution is normal. Hence, most models of short-term response are based on the implicit assumption that the infinitesimal model adequately describes the underlying genetics. BLUP and other mixed-model machinery (Chapters 19, 20, 22) explicitly assume breeding values are multivariate normal and hence also assume the infinitesimal. While the infinitesimal model is not taken as an exact description of biological reality, it represents one extreme of the possible assumptions about the underlying genetic architecture of a trait. When a large number of loci, each of small effect, underlie a character, it provides a satisfactory treatment of short-term response.

Selection (and drift) compromise predictions of response by changing allele frequencies and generating disequilibrium. Predicting changes in allele frequencies is especially problematic, requiring 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** (observable) predictors of response, based on easily-measured quantities, such as genetic variances. The breeder's (13.1 and 13.21) and Bulmer's (16.7b) equations are examples of macroscopic-based predictors of response. In contrast, the Price equation (6.8), which gives an exact prediction of response, has a composite transmission parameter $\sigma(w, \delta)$ that is, at best, very difficult to measure.

The goal of this chapter, which is rather technical in places, is to examine short-term 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, which not only changes the variance, but can also cause departures from normality. Finally, selection generates linkage-disequilibrium, which also alters genetic variances (Chapter 16) and causes departures from normality.

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 very short-term predictors of response based on macroscopic parameters with 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: DATA

Fisher's (1918) use of the infinitesimal model was part mathematical convenience and part biological approximation under the assumption that many genes of small effect underly a standard trait. Many geneticists, being trained in dissecting the action of single genes of modest to large effect, felt that this biological motivation was not justified. This objection was reinforced in the mid 1980's as QTL mapping experiments seemed to detect abundant major effect QTLs, each accounting for 10% or more of the phenotypic variation (LW Chapter 15). However, with the advent of more powerful genomic tools, trying to isolate the underlying sites behind these major peaks became increasingly frustrating. A major peak seen in the cross of two inbred lines often fractionated into several minor peaks upon finer mapping, and each of these peaks in turn fractionated as well as attention was turned to them. While genes of large effect were certainly found, more often what appeared to be a single major gene turned out to be a number of tightly-linked regions of much smaller effect (reviewed by Flint and Mackay 2009; Mackay et al. 2009).

This trend continued during the GWAS (genome-wide association study) phase of quantitative genetics starting in the early 2000's. Association mapping (LW Chapter 16) uses population-level disequilibrium and therefore allows for mapping on a kilobase (or smaller) scale, as opposed to the tens of megabase scale of QTLs. The broad conclusions (relevant to the infinitesimal model) from a large number of studies on human diseases with enormous sample sizes (in the tens of thousands or greater) were two fold (Visscher et al. 2012). First, the effects of detected sites tended to be very small. For example, over 600 variants associated with human height variation have been detected, most of which typically account for only a minuscule fraction of the additive variance (Allen et al. 2010). Second, the total additive variance accounted for by all detected sites was only a small fraction (around 10%) of the additive variance for the same trait estimated from relatives (typically twin studies in humans). The latter observation lead to concerns about "missing heritability" (Manolio et al. 2009), and a large number of papers attempting to account for this apparent paradox. In reality, this observation provides strong support for the infinitesimal model. In testing up to millions of SNPs for association in a GWAS, stringent thresholds are set to control for multiple comparisons (Appendix 4). As a result, power is greatly weakened, so that sites with small effects are unlikely have their test statistics exceed the threshold value for significance and are therefore not counted. Using mixed-model approaches that allowed all SNPs to be incorporated by shrinking the effects of most towards zero, Yang et al. (2010) could account for 45% of the additive variance in human height. Similar findings were seen for schizophrenia (Purcell et al. 2009). Example 24.1 illustrates how incomplete linkage disequilibrium can easily account for the remaining fraction of the "missing" heritability. Thus, we have come essentially full-circle in that current genomic data are most consistent with a very large number of loci, each of small effect, underlying traits. This is not to say that major alleles do not exist, but rather that they tend to be rare. Indeed, the effect detected in a GWAS study is the

additive variance of a site (LW Chapter 16), so finding that the vast majority of sites have low-effect variance does not imply they are due to alleles of small effects. If alleles of large effect tend to be rare, then some of these sites could indeed harbor large-effects alleles, which (as we detail below) has consequences in the prediction of selection response.

Example 24.1. A new QTL allele Q with additive effect a arises on an M background, where M and m are the two alleles at a SNP marker. The strongest association occurs when Q is completely restricted to the background on which it arose, so we assume this. Further, let q be the fraction of M haplotypes that carry Q , giving

Gamete	Frequency	Effect
MQ	pq	a
Mq	$p(1-q)$	0
mQ	$q(1-p)$	0
mq	$1-pq$	0

The additive variance associated with the casual site is

$$\sigma_A^2(Q) = 2a^2(pq)(1 - pq) \sim 2a^2pq$$

Conversely, the average effect of marker allele M is qa , giving the additive variation associated with this marker as

$$\sigma_A^2(M) = 2(qa)^2(p)(1 - p) = 2a^2p(1 - p)q^2$$

The resulting ratio of the actual to marker variances is $\sigma_A^2(M)/\sigma_A^2(Q) = q(1 - p)/(1 - qp)$, so that if Q is somewhat rare on M backgrounds, only a fraction of the actual variance is accounted for by the linked SNP. If, on average, roughly 50% of the M marker alleles carry Q , then the 45% of additive variance in human height accounted for by markers that Yang et al. (2010) observed will fully account for all of the additive variance. Further, this calculation is biased in favor of SNP variances by assuming complete disequilibrium. If some of the m marker alleles carried Q , the fraction accounted for by marker variance is even less.

THE INFINITESIMAL MODEL: THEORY

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, which serves as a starting point for the consequences on short-term response when they fail.

Allele Frequencies Do Not Change Under the Infinitesimal Model

Recall from Chapter 16 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 essentially 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 **inter-changeable**, with each locus having the same effects and frequencies (this is also called the **exchangeable model**). 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 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 remain 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 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 be seen from Equation 16.1. 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 model, but requires a rather delicate scaling of allelic effects so as to bound both the dominance variance and any inbreeding depression. 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 $0 : a_i + \delta_i : 2a_i$, where the frequency of the increasing allele is p_i (we use δ in place of our standard dominance notation of d to avoid any confusion with our other use of d for disequilibrium). The resulting dominance variance becomes

$$\sigma_D^2 = \sum_i^n (2p_i(1-p_i)\delta_i)^2$$

For n exchangeable loci,

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

For σ_D^2 to remain bounded as $n \rightarrow \infty$, d must be order $n^{-1/2}$, the same as we found for a . Thus, if both a and δ 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

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

Assuming n loci of equal effect gives

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

Note that if δ 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 δ 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. Under the exchangeable infinitesimal model, one cannot have both bounded dominance variance and inbreeding depression, a point first made by Robertson and Hill (1983). Of course, the flaw in this argument is our assumption of equal effects over all loci. In this case, all of the δ have the same sign. If we assume $E[\delta] = 0$, i.e., no directional dominance, then we can have 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 $0 < n E[\delta] < \infty$ (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 — implies that the distribution of breeding values is Gaussian under the infinitesimal model. This assumes that loci are unlinked and that there has been no previous selection (and hence no LD disequilibrium). 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 stops and random mating occurs, any 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. The issue remains as to how much of a departure from normality LD generate 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 homoscedasticity with the predictor error variance being a constant, independent of the parental values. This **Mendelian sampling variance**, which is half the genic variance in an infinitesimal model in a infinite population, is caused by segregation of heterozygous loci 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 and hence change the segregation variance.

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, while in the limit (number of loci going to infinity) the within-family random segregation typically approaches a normal, the formal infinitesimal model may 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 small), then convergence to a normal is by no means assured. Matthysee et al. (1979) show that a model with an infinite number of genes with gradually diminishing effects need not always converge to a normal, with the limiting variance of the offspring distribution depending on the parent's genetic values. 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, and hence changes in σ_a^2 , caused by a finite number of loci and/or genetic drift. This approximation breaks down over time, and hence is best regarded as an intermediate-term predictor of response. Next, we allow for linkage. Finally, we examine response when the genotypic distribution is no longer normal. None of these approaches fully account for allele frequency change, and 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 large, but finite. With a finite number of loci (and hence non-vanishing individual locus effects), some (albeit potentially very small) selection-induced allele frequency change occurs each generation. After a sufficient number of generations, the cumulative effect of these changes becomes 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 breeder's 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 (COA)**. These 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 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** (Chapter 2). 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 effectively 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 2). This model was concerned with allele frequencies, not their effects. 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. These continuum-of-alleles models were 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. While COA models are a very restrictive subset of all possible models that can lead to the infinitesimal, their advantage is that we can assume a *finite number of loci*, and hence partly accommodate allele frequency change.

COA models attempt to bridge short-term predictors (such as the breeder's and Bulmer equations) that rely on estimable quantities (σ_A^2 , h^2) with the long-term predictors of response (Chapters 25, 26) that are based on population-genetic models containing quantities that 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 16.7) 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 $\sigma_z^2 = (1 - \kappa) \sigma_z^2$ (Equation 16.10), then the equations for change in additive genic variance σ_a^2 and the genetic-phase

disequilibrium d when population size N_e is finite become

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

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

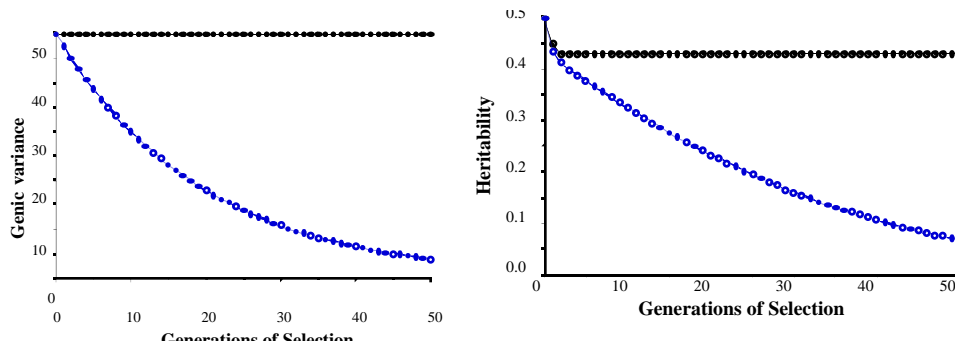
As before, $\sigma_A^2(t) = \sigma_a^2(t) + d(t)$ and $h^2(t) = \sigma_A^2(t)/\sigma_z^2(t)$, where $\sigma_z^2(t) = \sigma_A^2(t) + \sigma_e^2$ with $\sigma_e^2 = \sigma_z^2(0) - \sigma_A^2(0)$. The resulting response in the mean is given by the breeder's equation, $R(t) = h^2(t)S(t)$. If population size N_e is at least modest, the correction for drift effects on d is small. Drift effects on the genic variance, however, are quite substantial, removing all the initial genic variance after sufficient time (ignoring new mutation). Solving Equation 24.1a gives

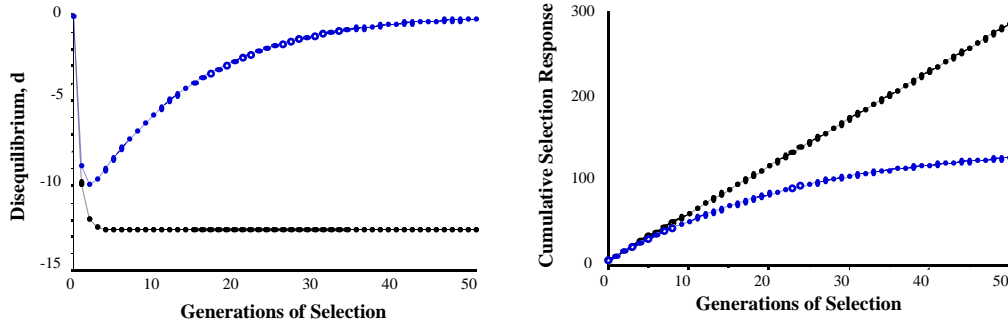
$$\sigma_a^2(t) = \left(1 - \frac{1}{2N_e}\right)^t \sigma_a^2(0) \simeq \sigma_a^2(0) \exp\left(-\frac{t}{2N_e}\right) \tag{24.1c}$$

This is just the standard loss of genetic variation under drift (Chapter 2). When dominance and/or epistasis is present, the additive variance can actually *increase* (for a while) under inbreeding (Chapter 11), so the assumption of only additive gene action is critical. With finite population size the response runs out of standing variation, as σ_a^2 is driven to zero by drift.

Example 24.2. To see the effects of drift on the infinitesimal model, reconsider Example 16.2 under finite population size. This example assumed truncation selection with the upper 20% saved (giving $\kappa = 0.787$ and $\bar{i} = 1.40$) 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 $\tilde{\sigma}_A^2 = 37.41$, and hence $\tilde{h}^2 = 0.43$ with an asymptotic value of response of $\tilde{R} = \tilde{i}\tilde{h}^2\tilde{\sigma}_z = 5.6$ 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 selection-induced 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 allele frequencies at these loci change (although potentially very slowly). 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), where for loci of equal effect and assuming $\sigma_{z^*}^2 = (1 - \kappa) \sigma_{z'}^2$, we have

$$\Delta \sigma_a^2(t) = - \left[\frac{\sigma_a^2(t)}{2N_e} + \left(1 - \frac{1}{N_e}\right) \frac{\kappa h^2(t) \sigma_A^2(t)}{2n} \right] \quad (24.2a)$$

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

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_a^2 and d . Note that, starting from an unselected base population, the only parameters needed to iterate the above equations are $\sigma_A^2(0)$, 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 Bulmer equation (16.7b),

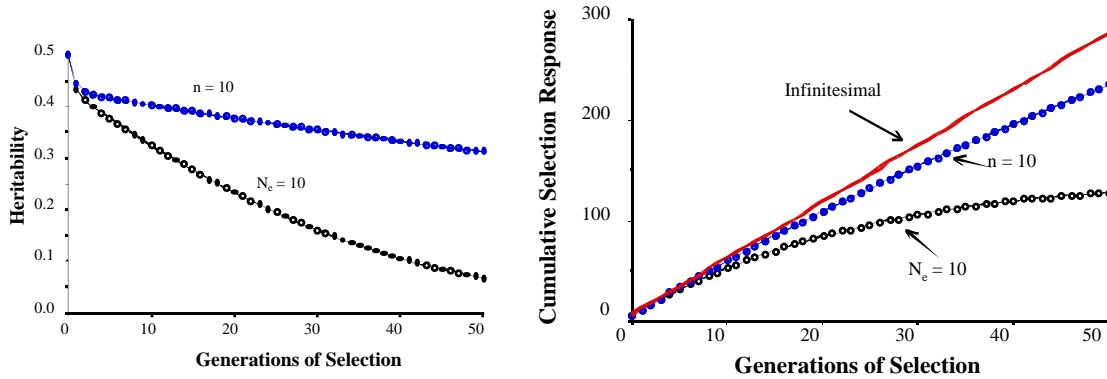
$$\Delta \sigma_a^2(t) = 0, \quad \Delta d(t) = - \frac{d(t) + \kappa h^2(t) \sigma_A^2(t)}{2}$$

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

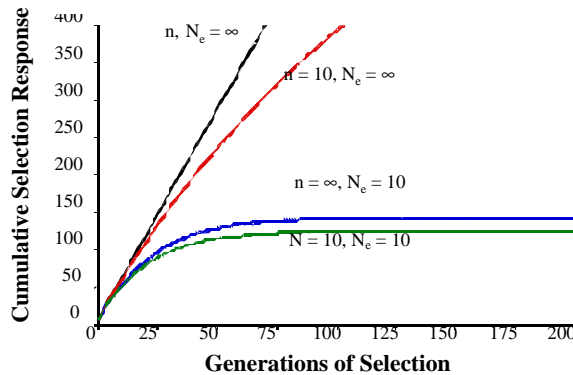
While finite n and/or N_e result in modifications of the simple Bulmer equation for the dynamics of d , these corrections are generally small. This is not the case for changes in σ_a^2 . With either finite population size and/or finite number of loci, the genic variance σ_a^2 decreases each generation, eventually going to zero (in the absence of mutation). The relative importance of drift versus a finite number of loci for changes in σ_a^2 can be compared using Equation 24.4a. With selection generating negative d (directional and/or stabilizing

selection), then $0 < \kappa < 1$ and $\sigma_A^2 = \sigma_a^2 + d < \sigma_a^2$, so that $\sigma_a^2(t) > \kappa h^2(t) \sigma_A^2(t)$. Thus, for comparable values of N_e and n , drift results in a greater per-generation reduction in σ_a^2 .

Example 24.3. Now let's consider what happens when the number of loci is finite. Assume the same model as in Example 24.2, but let $n = 10$ and $N_e = \infty$. We will contrast the behavior of this system with that in Example 24.2 ($N_e = 10, n = \infty$) and the infinitesimal model ($N_e = n = \infty$). As the figures below show, both h^2 and response decrease over time with a finite number of loci, and eventually a selection limit is reached as all initial variation is lost. However, these decreases is not nearly as dramatic as those in Example 24.2.



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, replacing the number of loci n in Equation 24.2 by n_e , the **effective number of loci**,

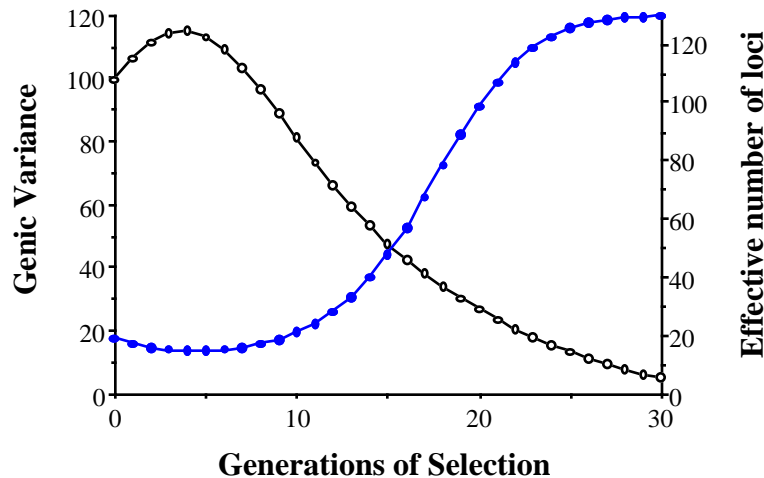
$$n_e = \frac{n}{1 + cv^2} \tag{24.3}$$

where cv is the coefficient of variation in the genic variance contributed by each locus,

$cv = \sigma(\sigma_{ai}^2) / E[\sigma_{ai}^2]$, where σ_{ai}^2 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 $n_e = n$, while $n_e \ll n$ when $cv \gg 1$.

It is important to 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 locus-specific 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 24.4. 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 $\sigma_a^2 = 100$ (half of which is from the major genes and half from the minor loci) and we assume a initial heritability of $h^2 = 0.5$ (by setting $\sigma_e^2 = 100$). 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 cv increases with the increases in genic variances at each major locus. 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,

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

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 $n_e \rightarrow \infty$, 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

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

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 16). 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

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

while the within-family (additive genetic) variance is

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

Response in Stabilizing Selection Experiments: Selection or Drift?

Gavrilets and Hastings (1994, 1995) noted that allele frequency changes can be quite slow under stabilizing selection. 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 16, 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 mentioned in Chapter 17, 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? The Gaussian assumption for each locus is the key, but one that is also incorrect at the start of selection, when COA approximation generally works well. If the distributions of genotypic values at each locus are initially *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. 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. Skewness measures departures from symmetry, while kurtosis measures if the tails of the distribution decline more rapidly, or more slowly, than a normal. 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 (LW Chapter 2).

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

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

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 (Example 24.6).

Lande (1983) and 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. If these rare alleles are captured when a population is sampled to form a laboratory stock for artificial selection, 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*. When loci are linked, the contribution d from gametic-phase disequilibrium decays by less than half each generation. This allows higher values to accrue, yielding larger value of $|d|$ at equilibrium. We examine the impact of linkage under two different settings. First, we consider the continuum-of-alleles approximation (requiring multivariate normality of the locus-specific distributions of effects). Second, we allow for departures from normality, which serves as a jumping-off point for our final section on treating non-Gaussian distributions.

An Approximate Treatment

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). Recall from Chapter 16 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 16.1b), $d(t) = 4 \sum_{j < i} C_{ij}(t)$, 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 $C_{ij}(t + 1)$. Recalling Equation 16.6, the change in $d(t)$ due to selection when genotypic and phenotypic values are normally distributed is

$$\frac{h^4(t)}{2} \delta(\sigma_z^2(t)) = 4 \sum_{j < i} \delta C_{ij}(t) \quad (24.7a)$$

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 (an exchangeable model). For n loci, there are $n(n - 1)/2$ unique pairs, giving the contribution from each pair as

$$4\delta C_{ij}(t) \simeq \frac{h^4(t) \delta(\sigma_z^2(t))}{n(n - 1)} \quad (24.7b)$$

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

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

This equation is approximate as the covariance *between* gametes $C_{i,j}$ (which is ignored here) 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 $r_{ij} \tilde{C}_{ij} = \tilde{\delta} C_{ij}$. Using Equation 24.7b gives the equilibrium covariance as

$$\tilde{C}_{ij} = \frac{\tilde{h}^4 \tilde{\delta}(\sigma_z^2)}{4n(n - 1) r_{ij}} \quad (24.7d)$$

thus

$$\tilde{d} = 4 \sum_{j < i} \tilde{C}_{ij}(t) = 4 \frac{\tilde{h}^4 \tilde{\delta}(\sigma_z^2)}{4n(n - 1)} \sum_{j < i} \frac{1}{r_{ij}} = \frac{1}{2} \tilde{h}^4 \tilde{\delta}(\sigma_z^2) \frac{1}{r_H} \quad (24.8a)$$

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

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

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

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

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

and the equilibrium heritability is given by Equation 16.13c using the above value of θ . As we saw in Chapter 16, with sufficiently strong disruptive selection ($\kappa < -r_H/2h^2(1-h^2)$), there is no real positive root for θ and the infinitesimal model predicts that the additive variance increases without limit (Bulmer 1976a). With a finite number of loci, this condition implies that selection creates almost complete disequilibrium, so that only a few of the possible gamete types are actually present, i.e., most gametes are either $aabbccdd \dots$ or $AABBCC \dots$ (Chapter 16). The general conclusion is that increasing the amount of linkage (e.g., decreasing r_H) increases the absolute value of \tilde{d} (Bulmer 1974, 1976a, 1980).

Example 24.5. As an example of the consequences of increased linkage, reconsider our analysis of the response under directional used in Examples 24.2 and 24.3. Here we assume infinite number of loci and infinite population size. Substituting into Equation 24.9 to obtain θ and recalling Equations 16.13a-c gives

r_H	θ	\tilde{d}	$\tilde{\sigma}_A^2$	\tilde{h}^2	\tilde{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

$r_H = 0.5$ corresponds to free recombination, while $r_H = 0.1$ 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 ($r_H = 0.1$), the response is reduced 33% relative to unlinked loci.

While the infinitesimal model predicts that \tilde{d} increases under disruptive selection as linkage tightens, Sorensen and Hill (1983) found exactly the opposite in their simulations, with \tilde{d} decreasing 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 \tilde{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 r_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, \tilde{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 Example 24.4 showed, response 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 recombination generating favorable gametes, which then increase in frequency. They also

found that linked systems are vulnerable to lower response due to hitchhiking fixing less favorable 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,

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

which is the covariance between the effect of an allele at the i th locus in the paternal (fa) gamete and an allele at the j th locus in the maternal (mo) 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),

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

Substituting for C^* gives

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

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 13.27b) that

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

$$\frac{\partial \ln \bar{w}}{\partial \mu_z} = \frac{S}{\sigma_z^2} \quad (24.12a)$$

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

(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

$$\delta C_{ij} = \frac{\partial \ln \bar{w}}{\partial \mu_z} \sum_k C_{ijk} + \frac{\partial \ln \bar{w}}{\partial \sigma_z^2} \sum_k \sum_l (C_{ijkl} - C_{ij} C_{kl}) + \dots \quad (24.13a)$$

$$\delta C_{i,j} = \frac{\partial \ln \bar{w}}{\partial \sigma_z^2} 2 \sum_k C_{ik} \sum_l C_{jl} + \dots \quad (24.13b)$$

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 ($C_{ijkl} = C_{ij} C_{kl} + C_{ik} C_{jl} + C_{il} C_{jk}$). In this case, $\delta C_{ij} = \delta C_{i,j}$, and combining Equations 24.12b and 24.13b gives

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

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

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

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

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

When $S^2 \ll |\delta(\sigma_z^2)|$ (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, 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 homoscedastic 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. This provides a simple solution to the vexing problem of modeling the transmission of a quantitative trait. The other option is Price's Equation (Equation 6.8), but its composite transmission parameter $\bar{\delta}_z$ is very difficult to estimate.

While changing allele frequencies and gametic-phase disequilibrium compromise prediction of response by changing the genetic variance, a more subtle, but no less important, issue is that they also compromise prediction by driving the 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 arbitrary distributions of genotypic values. While good progress has been made, we warn the reader that this can be a rather intimidating 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 translates 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. 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 both the breeder's and Bulmer equations are quite accurate for both directional and strong disruptive selection (Turelli and Barton 1994).

Describing the Genotypic Distribution: Moments

Under our assumption that genotypic and environmental values are additive and independent, $z = G + e$. When 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 if *phenotypes* are normally distributed, this tells us little about the distribution of *genotypes*. In theory, we can estimate this distribution by estimating the breeding values for a sample of individuals (LW Chapter 26), but this is generally impractical in most studies (but see Chapter 20). Further, methods used to estimate breeding values typically assume normality, and hence bias the distribution of estimated values towards a Gaussian.

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 $\mu_{3,z} = \mu_{3,G} + \mu_{3,e}$. In a constant environment, changes in any of the first three phenotypic moments exactly equals the change in the corresponding genotypic moments. Example 24.6 (below) derives the fourth phenotypic moment,

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

showing 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. When e is normal, $\mu_{3,e} = 0$ and $\mu_{4,e} = 3\sigma_e^4$, simplifying 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

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

where $X_{fa,i}$ ($X_{mo,i}$) is the value of the paternal (maternal) allele at the i th 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

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

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 16.2) and here we extend this partitioning to the third and higher genotypic moments. To describe the distribution of effects at locus i , let $\mu_{1,i} = E(X_i) = m_i$ denote the average value of an allele at locus i and define the k th moment for this locus by $\mu_{k,i} = E([X_i - m_i]^k)$ for $k \geq 2$. Summing over all n loci, define

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

$$M_2 = 2 \sum_i^n \mu_{2,i} \quad (24.18b)$$

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

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 as

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

While this may at first seem odd, recall for a normal that the fourth and second moments are related by $\mu_4 = 3\mu_2^2$ (LW Chapter 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, and these values provide a quantitative measure of the departure from a normal.

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_{ijjj} , 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 epistasis), so that the genotypic distribution G is the distribution of additive genetic (breeding) values (A). Expanding Equation 24.17 and taking expectations gives the familiar expressions for the mean and variance,

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

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

where M_2 corresponds to the genic variance σ_a^2 , and the double sum to the disequilibrium d (Equation 16.1). Similarly, the skew can be partitioned as

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

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),

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

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 $C_{ijkl} = C_{ij}C_{kl} + C_{ik}C_{jl} + C_{il}C_{jk}$.

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 from disequilibrium*. These 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_{ijkl}) 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 convergence to a Gaussian is that the central limit theorem assumes that the variables are independent (or at least only weakly correlated), while by introducing gametic-phase disequilibrium, selection generates dependence between loci.

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 $n \cdot n^{-3/2} = n^{-1/2}$ (also see Equation 24.6). 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 $n \cdot n^{-2} = n^{-1}$ (Equation 24.6). 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 it 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 its **cumulants**. The first use of cumulants in examining selection response appears to be O'Donald (1972) and Bulmer (1980). Sophisticated (and highly technical) treatments 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:

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

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

$$K_4 = \mu_4 - 3\mu_2^2, \quad K_5 = \mu_5 - 10\mu_2\mu_3 \quad (24.20)$$

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., $K_n(x+y) = K_n(x) + K_n(y)$. 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 24.6. Use cumulants to compute the fourth and fifth central moments of the phenotypic distribution. Here, $z = G + e$, so that the fourth moment is

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

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

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

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 24.7. If the underlying genes are additive across loci (no epistasis), 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. Following Turelli and Barton (1994),

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

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 between loci. This recovers Equation 24.19d by noting the $\mu_{4,G} = K_{4,G} + 3\sigma_A^4$ 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 convenient 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 by $(1/2)^{j-1}$ each generation, so that the distribution rapidly returns to a normal.

Finally, cumulants appear in series approximations of arbitrary 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:

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

where H_k denotes the **Chebyshev-Hermite polynomial** of order k , with

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

Equation 24.21 shows how the higher-order cumulants (K_3 and above) quantify departures from 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 was examined by Bulmer (1980) and Zeng (1987), while 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 14, 16), we choose the uppermost p percent of the population, giving a selection intensity of

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

where U is a unit normal random variable (Equation 14.2b). From Chapters 14 and 16, we already have the first two cumulants following selection as

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

while the next two cumulants are

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

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

Next, we translate these within-generation changes in the *phenotypic* distribution into the within-generation change in the distribution of *breeding values* and then examine how this changes (under random mating) during transmission to the next generation. Both these steps rely critically on assumptions of normality. If the distribution of breeding and phenotypic values is bivariate normal before selection (as might 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 (Chapter 16) for the mean breeding value and its variance following selection

$$\mu_A^* = \mu_z + h^2 \bar{v} \sigma_z \quad (24.23a)$$

and

$$\sigma_{A^*}^2 = \sigma_A^2 [1 - h^2 \bar{v} (\bar{v} - z_p)] \quad (24.23b)$$

When the joint distribution of breeding values and phenotypes is multivariate normal, Bulmer (1980) showed that all higher cumulants follow a very simple relationship,

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

Assuming unlinked loci, the cumulants for the distribution of breeding values in the next generation becomes

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

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

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

Notice that after a single generation of selection, $K_{3,A}$ and $K_{4,A}$ are non-zero, and hence 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 24.8. Suppose truncation selection occurs on a normally-distributed trait with initial mean $\mu_z = 0$ and initial variance $\sigma_z^2 = 100$. The upper 5% are saved, so that $\bar{v} = 2.063$ and $z_p = 1.645$ (Example 14.1). 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

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

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 16 (Equations 16.11a, 16.12) 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 perturbed away from a Gaussian. Further note that this is the most extreme case, as we have assumed $h^2 = 1$. 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 from 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 16.7b), 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 of 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 is generally 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 fully 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, recall Wright's formula for multiple alleles (Equation 5.11),

$$\Delta p_i = \sum_j G_{ij} \frac{\partial \ln \bar{w}}{\partial p_j} \quad (24.25a)$$

where $G_{ii} = p_i(1 - p_i)/2$ and $G_{ij} = -p_i p_j/2$ (for $i \neq j$). The assumption of linkage equilibrium is needed as Wright's formula fails when fitnesses at each locus are frequency-dependent, which can occur with constant fitnesses when linkage disequilibrium is present (Example 5.7). Now consider a function $f(p_1, p_2, \dots, p_{k-1})$ 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

$$\Delta f = \sum_i \frac{\partial f}{\partial p_i} \Delta p_i + \frac{1}{2} \sum_i \sum_j \frac{\partial^2 f}{\partial p_j \partial p_i} \Delta p_i \Delta p_j + \dots \quad (24.25b)$$

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

$$\frac{\partial \ln \bar{w}}{\partial p_j} = \frac{\partial \ln \bar{w}}{\partial f} \frac{\partial f}{\partial p_j}$$

gives (to first order)

$$\begin{aligned} \Delta f &\simeq \sum_i \frac{\partial f}{\partial p_i} \sum_j G_{ij} \frac{\partial \ln \bar{w}}{\partial p_j} \\ &= \frac{\partial \ln \bar{w}}{\partial f} \sum_i \sum_j \frac{\partial f}{\partial p_i} G_{ij} \frac{\partial f}{\partial p_j} \end{aligned} \quad (24.25c)$$

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

$$\Delta_{\mu_G} \simeq \mathbf{M} \nabla \ln \bar{w} \quad (24.26)$$

where

$$\Delta_{\mu_G} = \begin{bmatrix} \Delta \mu_{1,G} \\ \Delta \mu_{2,G} \\ \Delta \mu_{3,G} \\ \vdots \end{bmatrix}, \quad \nabla \ln \bar{w} = \begin{bmatrix} \frac{\partial \ln \bar{w}}{\partial \mu_{1,z}} \\ \frac{\partial \ln \bar{w}}{\partial \mu_{2,z}} \\ \frac{\partial \ln \bar{w}}{\partial \mu_{3,z}} \\ \vdots \end{bmatrix}$$

are (respectively) the vector of changes in the genotypic moments and of partial derivatives of long mean fitness with respect to each moment, and

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

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 $\partial \ln \bar{w} / \partial K_{i,z}$ (the partial derivative 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.

When 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

$$\Delta \mu_z \simeq \sigma_A^2 \frac{\partial \ln \bar{w}}{\partial \mu_z} + \mu_{3,G} \frac{\partial \ln \bar{w}}{\partial \mu_{2,z}} + k_4 \sigma_A^4 \frac{\partial \ln \bar{w}}{\partial \mu_{3,z}} \quad (24.27)$$

where $k_4 = (\mu_{4,G} - 3\sigma_A^4) / \sigma_A^4$ is the scaled coefficient of kurtosis. If the distribution of G is Gaussian, then $\mu_{3,z} = k_4 = 0$, and we recover the selection gradient version of the breeder's equation (Equation 13.27a). Under more general distributions, 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 ($\partial \ln \bar{w} / \partial \mu_{k,z}$). In order to proceed further, we have to make additional assumptions about the distribution of allelic effects at *individual loci*.

Example 24.9. 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

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

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 24.10. The expected response in the genotypic mean and variance becomes

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

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 breeder's equation and short-term changes in variance (from allele-frequency change) are expected to be small when n_e is modest. We remind the reader that this analysis ignores the effects of gametic-phase disequilibrium.

Since the locus-specific values $\mu_{2,i}$ change as allele frequencies change, predicting changes in variance over several generations even under these simplifying assumptions 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). Further, as allele frequencies change, so does n_e , and Example 24.4 showed just how unpredictable these changes 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). First, $\Delta\sigma_A^2 = \Delta\sigma_a^2$ as we ignore any disequilibrium. 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

$$\Delta\sigma_a^2 = -\frac{\kappa h^2 \sigma_A^2}{2n_e}$$

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

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

The allelic-moment approximation has an additional positive term relative to the COA approximation, predicting in a smaller change in σ_a^2 when $\kappa > 0$

Example 24.10. 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 casual 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),

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

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

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

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

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

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 acting directly 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 change at the individual loci or gametic-phase disequilibrium, is more important at producing departures from normality depends on whether there are alleles of modest to large effects. When these are present, the locus-specific selection coefficients (e.g., Equation 5.21) are sufficiently large that significant allele frequency

change can quickly occur, which has a far greater effect on departures from normal than do selection-generated disequilibrium. Conversely, when all alleles have small effects (selection is extremely weak on any single locus), over modest time scales departure from normality is largely due to selection generating third (and higher) level disequilibrium. For example, 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 change, while the contribution from third and fourth-order disequilibrium was quite small. When the number of loci is small, 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 considered one class of approximations for short-term response for non-Gaussian distributions of genotypic values, focusing solely on allele frequency changes. Here we consider the converse: a large enough number of loci such that allele frequency change (over our time span of interest) can be ignored, with the change in genotypic moments due to selection-generated 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

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

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

$$\Delta\mu = \sigma_A^2 L_1 + K_{G,3} L_2 + K_{G,4} L_3 + K_{G,5} L_4 \quad (24.29b)$$

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

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 cumulants of the distribution, driving 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 24.11. If phenotypes are normally-distributed, Equations 24.12a and b give

$$L_1 = \frac{S}{\sigma_z^2}, \quad L_2 = \frac{\delta(\sigma_z^2) + S^2}{2\sigma_z^4}, \quad L_i = 0 \quad \text{for } i \geq 3$$

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 breeder's 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 equation. 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 cumulants of order three and higher (and hence departures from normality), these are generally quite small (e.g., Example 24.8). As a result, the breeder's equation with the variance changes predicted from the Bulmer equation (16.7b) give quite accurate results for the predicted change in the mean and variance. Hence, the effects of disequilibrium in this case are essentially 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 still well predicted by the Bulmer equation.

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 presented), 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

$$\Delta\mu_z \simeq \sigma_a^2 \left(\frac{\partial \ln \bar{W}}{\partial \mu_z} \right) + \frac{\sigma_a^4}{r_{H_2}} \left(\frac{\partial \ln \bar{w}}{\partial \mu_z} \cdot \frac{\partial \ln \bar{w}}{\partial \sigma_z^2} \right) + \frac{3\sigma_a^6}{2r_{H_3}} \left(\frac{\partial \ln \bar{w}}{\partial \sigma_z^2} \cdot \frac{\partial \ln \bar{w}}{\partial \mu_{z,3}} \right) \quad (24.30)$$

where r_{H_2} and r_{H_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 approach equilibrium 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

$$\Delta\mu_z \simeq \frac{\sigma_a^2}{\tilde{\sigma}_z^2} \left(S + \sigma_a^2 \frac{\tilde{\delta}(\sigma_z^2) + S^2}{2\tilde{\sigma}_z^2} \left[\frac{S}{r_{H_2}} + \frac{3\sigma_a^2}{2r_{H_3}} \frac{\partial \ln \bar{w}}{\partial \mu_{3,z}} \right] \right) \quad (24.31)$$

where $\tilde{\sigma}_z^2$ is the equilibrium phenotypic variance and $\tilde{\delta}(\sigma_z^2)$ is the equilibrium within-generation change in phenotypic variance due to selection. This generalizes Bulmer's results, which corrected the breeder's equation for changes in the variance due to pair-wise disequilibrium. 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?

Predicting selection response is complicated. Even in the ideal setting where the breeder's equation exactly holds, drift and segregation generate a *variance* in response about the expected value (Chapter 18). Any particular *realization* of response will thus vary, and hence be less predictable. Second, even when the parent-offspring regression is assumed to be linear and homoscedatic, there are still a large number of confounding factors for even single-generation response (Table 13.2). Despite these concerns, short-term prediction of response is reasonable for many traits (Chapter 18). In contrast, prediction of long-term response is an unobtainable goal unless one essentially knows all of the very fine (microscopic) genetic details of a trait, *including* the distribution of allelic effects and frequencies.

As we have seen here and elsewhere (Chapters 5, 16), selection compromises response prediction through two avenues. First, when some of the underlying loci harbor alleles of modest to large effect, this generates selection coefficients that can result in significant allele frequency change over very short time scales (Equations 5.21, 5.3). Such changes alter the base-population heritability in ways not predictable from observable macroscopic features (such as the initial additive variance). Rather, their dynamics depend very fine details of the genetic architectures. A more subtle consequence of allele-frequency change is that it can drive a genotypic distribution away from normality by generating locus-specific skewness and kurtosis. This results in nonlinear and heteroscedatic parent-offspring regressions, and hence a failure of the breeder's equation, even when correctly updated values of h^2 are used.

The second consequence of selection, generation of linkage (or more correctly gametic-phase) disequilibrium, is often much more manageable. When the trait is controlled by a large number of loci, each of small effect, allele-frequency change is negligible over short time scales. However, selection-induced correlations (even among unlinked loci) change not only the genetic variance (Chapter 16), but can also generate skewness and kurtosis.

While the latter drive a genotypic distribution away from normality, with a large number of loci, usually the effect is modest and does not greatly compromise predictions of response. Further, as we saw in Chapter 16, the Bulmer equation accounts for changes in variances from disequilibrium using easily-observed parameters. Thus, allele frequency change is the more pernicious feature of selection, but (for short time scales) is restricted to traits whose underlying genetic architectures harbor one or more alleles of large effect.

Literature Cited

- Allen, H. Lango, K. Estrada, G. Lettre, S. I. Berndt, et al. 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 467: 832–838. [24]
- Barton, N. H. and M. Turelli. 1987. Adaptive landscapes, genetic distances and the evolution of quantitative 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. *Genet. Sel. Evol.* 26: 379–400 [24]
- Crow, J. F., and M. Kimura 1964. The theory of genetic loads. *Proced. of the XIth Inter. Congr. Genetics*, Vol. 2, 495–505 [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. Edin.* 52: 399–433. [24]
- Flint, J., and T. F. C. Mackay. 2009. Genetic architecture of quantitative traits in mice, flies, and humans. *Gen. Res.* 19: 723–733. [24]
- Gavrilets, S., and A. Hastings. 1994. Dynamics of genetic variability in two-locus models of stabilizing selection. *Genet.* 138: 519–532. [24]
- Gavrilets, S., and A. Hastings. 1995. 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. *Genet.* 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.

- Genet.* 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. *Evol.* 30: 314–334. [24]
- Lande, R. 1977. The influence of the mating system on the maintenance of genetic variability in polygenic characters. *Genet.* 86: 485–498. [24]
- Lande, R. 1983. The response to selection on major and minor mutations affecting a metrical trait. *Hered.* 50: 47–65. [24, 25]
- Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evol.* 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 populations with multiple alleles. II. Centripetal selection, mutation, and isoallelic variation. *Genet.* 66: 165–186. [14]
- Mackay, T. F. C., E. A. Stone, and J. F. Ayroles. 2009. The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* 10: 565–577. [24]
- Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorf, D. J. Hunter, M. I. McCarthy, E. M. Ramos, L. R. Cardon, A. Chakravarti, J. H. Cho, A. E. Guttmacher, A. Kong, L. Kruglyak, E. Mardis, C. N. Rotimi, M. Slatkin, D. Valle, A. S. Whittemore, M. Boehnke, A. G. Clark, E. E. Eichler, G. Gibson, J. L. Haines, T. F. C. Mackay, S. A. McCarrroll, and P. M. Visscher. 2009. Finding the missing heritability of complex diseases. *Nature* 461: 747–753. [24]
- Matthysse, S., K. Lange, and D. K. Wagener. 1979. Continuous variation caused 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]
- 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]
- Purcell, S. M., N. R. Wray, J. L. Stone, P. M. Visscher, et al. 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460: 748–752. [42]
- 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]
- Robertson, A., and W. G. Hill. 1983. Population and quantitative genetics of many linked loci in finite populations. *Proc. R. Soc. Lond. B* 219: 253–264. [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]
- Visscher, P. M., M. A. Brown, M. I. McCarthy, and J. Yang. 2012. Five years of GWAS discovery. *Amer. J. Hum. Genet.* 90: 7–24. [24]

- Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, G. W. , Montgomery, M. E. Goddard, and P. M. Visscher. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42, 565–569. [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 artificial selection in lines derived from natural populations. *Genet.* 169: 411–425. [24]