19

Analysis of Short-Term Selection Experiments: 2. Mixed-Model and Bayesian Approaches

Unnecessarily complex analysis should not be used as a foil to disguise lower quality datasets: estimates of genetic parameters are only as good as the data on which they are based — Kruuk (2004)

Draft version 24 June 2013

A least-squares (LS) analysis of a selection experiment distills the data down to the means and variances of the trait each generation. In some experiments, we have much more information, in the extreme measurements (or **records**) of all individuals throughout the course of the experiment and their complete pedigree. When such additional data is available, a LS analysis simply ignores it. A **mixed-model** (**MM**) **analysis** (LW Chapters 26, 27), on the other hand, fully considers the covariances between *all* observations. By virtue of using this additional information, a MM is potentially far more powerful. They are also more flexible, easily incorporating complicated fixed effects and highly unbalanced designs. Finally, breeders and evolutionary biologists (faced with complicated and shifting environments) are often more concerned with the genetic gain following selection, which may be different from the change in mean phenotype. LS estimates the later while MM, by estimating the mean breeding value in each generation, estimates the former.

There are two different procedures for a mixed-model analysis. The first is the **two-step approach**, wherein one first uses **REML** (LW Chapter 27) or some other method to estimate the appropriate variance components, and then uses these with **BLUP** to estimate breeding values (LW Chapter 26). With BLUPs for individual breeding values in hand, one can estimate the mean breeding value for any particular generation and hence (in theory) directly follow *genetic*, as opposed to *phenotypic* change, allowing for the separation of genetic versus environmental change, even when no control population is used. While straight-forward, the two-step approach not does account for the uncertainty in BLUP estimates that results from using estimates (as opposed to the true values) of the variances. In contrast, **Bayesian approaches** simultaneously estimate variance components and predict random effects (such as breeding values), returning (given a specified prior) a posterior distribution for random variables of interest given the data. Bayesian analysis provides an exact accounting of the uncertainty in the estimation of the unknown parameters, especially after the confounding effects of nuisance parameters are integrated out of the system.

Building around applications to the analysis of selection experiments (including applied breeding), we use this chapter to review some of the basic statistical machinery behind mixed models (which is used extensively both here and in Chapters 20 and 22) and to more formally introduce the Bayesian framework (briefly touched upon in Chapter 10, and more fully developed in Appendices 2 and 3). We start with a brief review of the theory of mixed models and then consider various applications of the **animal model** to the analysis of selection experiments. We then introduce the basics of Bayesian analysis, and conclude by re-analyzing many of the same mixed models under a Bayesian framework.

MIXED-MODEL vs. LEAST-SQUARES ANALYSIS

Figure 19.1 illustrates the result of a mixed-model analysis of selection response. Note

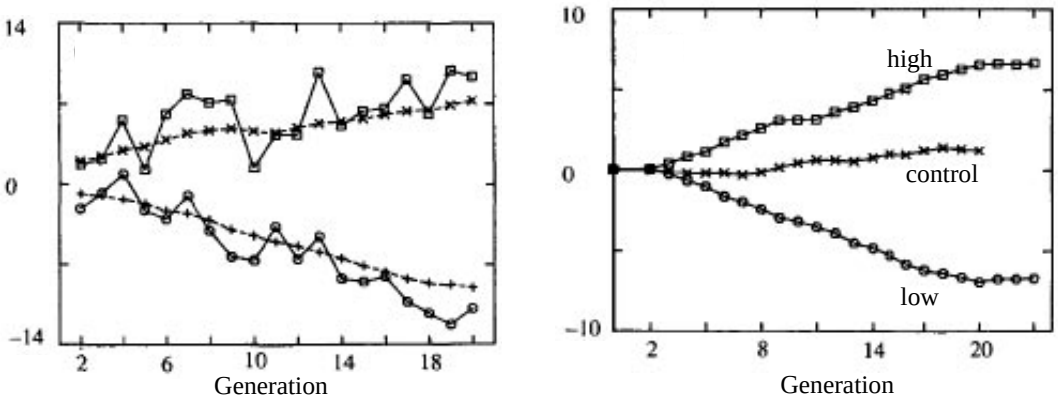


Figure 19.1. Results from high and low selection on 6-week weight in mice. **Left:** Observed (boxes and circles) and predicted (estimated mean breeding value plus estimated environmental value, crosses and pluses) mean phenotypic values in the up and down-selected lines, expressed as deviation from the control population. **Right:** Estimated mean breeding values for both selected populations and the control. See Examples 19.2 and 19.7 for more details on this experiment. After Meyer and Hill (1991).

that instead of measuring response from the *observed phenotypic means* (the LS approach), response is measured from the *estimated mean breeding values* obtained from BLUP. Further, instead of estimating a realized heritability, a mixed-model (REML) analysis estimates the additive genetic variance in the base population.

Mixed-models readily allow records to be adjusted for any number of **fixed effects**. For example, one might correct for trait differences between sexes, between individuals from different size litters, age effects, known environmental factors, etc. Such adjustments result in more accurate prediction of an individual's genetic value (and a more accurate estimate of the population's genetic response). While these are readily incorporated into a mixed-model analysis, it can be more problematic to adjust for fixed effects in a LS analysis. For example, a LS analysis cannot separate genetic from environmental trends when only a single line is considered. By contrast, under a properly-formatted MM analysis, one can separate phenotypic changes into genetic and environmental components without using a control population. This is possible since the covariance structure associated with the pedigree of all individuals in the experiment allows information to be borrowed across relatives from different generations.

Another advantage of mixed-model analysis is its *great flexibility* in handling any selection design within the same framework. For example, a MM analysis allows for overlapping generations (when a parent contributes offspring over several different years of selection), while LS analysis of response in overlapping generations can be difficult to formulate correctly. A properly-designed MM analysis can account for assortative mating, drift, and selection-induced gametic phase disequilibrium (provided we can assume the infinitesimal model). For all their power, a MM analysss has tradeoffs relative to a simpler LS analysis. First, a MM analysis requires far greater record keeping (e.g., following all individuals and their relatives) and is more computationally demanding. Second, a MM analysis can be rather model-sensitive, in particular the infinitesimal model assumption is critical. If selection-induced changes in allele frequencies are significant during the course of the experiment, the assumptions of a MM analysis can be violated. Finally, analysis of a mixed model *critically* depends on its correct covariance structure, so it this is incorrectly specified

the model is highly biased.

BLUP Selection

A previously introduced MM application is BLUP selection (Chapter 13), wherein a mixed model is used to find those individuals with the highest estimated breeding values, which are used as parents to form the next generation. This is the main route of selection used by animal and tree breeders, and to a growing extent by plant breeders working with outcrossing species. BLUP selection uses all of the information up to a given generation to choose those individuals who will be parents for the next cycle of selection. Conversely, a MM analysis of selection response is a *retrospective* analysis of the genetic gain of a population, wherein we start at some final time point and infer the trajectory of past genetic gain during the experiment/breeding program. Parents could have been chosen by simple mass selection, some selection or family index, or even by BLUP selection, but this is irrelevant to the MM analysis of the final genetic gain, which is solely based on the values of all individuals, and their relationships (pedigree), over the course of the experiment.

BASICS OF MIXED-MODEL ANALYSIS

While providing a brief overview of the analysis of mixed models here, we still encourage the reader to review LW Chapters 26 and 27 before proceeding. These chapters provide many additional worked examples to give a better feel for mixed models, as well as considering advanced topics in MM analysis in greater detail. We start with a quick review of some of the key theoretical results and then examine specific applications.

Mixed models are so named because they consider both fixed and random effects. Recall that fixed effects are unknown constants while random effects have values that are drawn from some underlying distribution (LW Chapters 8, 26). Hence, any particular value for a random effect represents just one possible realization from this underlying distribution, which is usually assumed to be a normal with mean zero and unknown (to be estimated) variance σ^2. A brief example will remind the reader of a powerful feature of random-effects models. Suppose we have t time points and include an environmental value E_i for each time point into our model. Treating these as fixed effects makes no assumption as to how the E_i from different generations are related to each other, but the cost is t degrees of freedom. Conversely, if we make a random-effects assumption that the E_is are draws from some underlying distribution, we use far fewer degrees of freedom. If we assume these are independent draws from an underlying normal, then we only need a single degree of freedom (the variance σ^2_E of this distribution). More generally, one could make additional assumptions about their distribution, for example, the E values for adjacent generations are autocorrelated by the same amount ρ, in which case the random model requires an additional degree of freedom.

Typically, statisticians speak of *estimating* fixed effects and *predicting* the realized values of random effects. Both LS and MM analyses estimate the fixed effects in a model, while MM analysis also predicts the values of the random effects by using the covariances between observations (after adjusting for fixed effects). Note that under a Bayesian analysis, every effect is assumed to be random, and this distinction between fixed vs. random effects becomes more subtle. The standard mixed model for a vector $\mathbf{y}$ of n observations is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{e}, \tag{19.1}$$

where $\boldsymbol{\beta}$ is a $q \times 1$ vector of fixed effects, $\mathbf{a}$ is a $p \times 1$ vector of random effects (in our case, the breeding values of the individuals in our experiment), $\mathbf{e}$ the $n \times 1$ vector of residuals (also random), $\mathbf{X}$ and $\mathbf{Z}$ are, respectively, the $n \times q$ **design matrix** and the $n \times p$ **incident matrix**

associated with the fixed and random effects. In the absence of the vector of random effects $\mathbf{a}$, Equation 19.1 reduces to a least-squares model, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$, with ordinary (unweighted) least squares (OLS) used to estimate $\boldsymbol{\beta}$ if the residuals are uncorrelated and homoscedastic, giving the covariance matrix for the vector of residuals as $\mathbf{Var}(\mathbf{e}) = \sigma_e^2\,\mathbf{I}$. More generally, if $\mathbf{Var}(\mathbf{e}) = \mathbf{V}$, where the only constraint is that $\mathbf{V}$ is symmetric and positive-definite, then generalized (weighted) least squares (GLS) is used (Equation 19.3a).

In order to solve Equation 19.1, we need to specify the covariance structure for the vectors of random effects. It is generally assumed that the residuals are uncorrelated and homoscedastic, $\mathbf{Var}(\mathbf{e}) = \sigma_e^2\,\mathbf{I}$. The covariance of $\mathbf{a}$ (in our case, the vector of breeding values) has a more complicated structure given by the pedigree, $\mathbf{Var}(\mathbf{a}) = \sigma_A^2\,\mathbf{A}$. Here $\mathbf{A}$ is a matrix of known constants (the **numerator relationship matrix**) whose elements are given by the pedigree structure. The resulting $n \times n$ covariance matrix $\mathbf{V}$ for the vector of observations $\mathbf{y}$ becomes

$$\mathbf{V} = \sigma_A^2\,\mathbf{Z}\mathbf{A}\mathbf{Z}^T + \sigma_e^2\,\mathbf{I}. \tag{19.2a}$$

The covariance matrix $\mathbf{V}$ is thus a function of the (usually unknown) variance components σ_A^2 and σ_e^2 and the matrices $\mathbf{Z}$, $\mathbf{A}$, and $\mathbf{I}$ of known constants. Since $\mathbf{a}$ is a vector of breeding values, we can alternatively express $\mathbf{V}$ as a function of the heritability (h^2) and phenotypic variance ($\sigma_z^2 = \sigma_A^2 + \sigma_e^2$) of the trait of interest (after removal of fixed effects),

$$\mathbf{V} = \sigma_z^2 \cdot \left(h^2\,\mathbf{Z}\mathbf{A}\mathbf{Z}^T + [1 - h^2]\,\mathbf{I}\right) \tag{19.2b}$$

as $\sigma_e^2 = \sigma_z^2 - \sigma_A^2 = \sigma_z^2(1 - h^2)$. Since our focus is generally on breeding values versus everything else, any dominance variance gets swept into the residual variance σ_e^2. Lumping any dominance into the residual potentially results in e_i values within families being correlated (due to full sibs sharing $\sigma_D^2/4$) , and we correct for this later.

Assuming $\mathbf{V}$ is known exactly, estimation of the vector $\boldsymbol{\beta}$ of fixed-effects follows from GLS (LW Chapter 8),

$$\widehat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y} \tag{19.3a}$$

Equation 19.3a is called the **best linear unbiased estimator** (**BLUE**) of the vector $\boldsymbol{\beta}$ of fixed effects. **Estimability** of the fixed effects can be an issue, as the structure of the data (indicated by the column rank of $\mathbf{X}$) may not allow for unique estimates of all fixed effects. In such cases, generalized inverses can be used to obtain unique estimates of certain linear combinations of the fixed effects (LW Appendix 2). If $\mathbf{X}$ has column rank $\ell \leq q$, then exactly ℓ combinations of the q fixed effects can be estimated, see LW Chapter 26 and LW Appendix 2 for further details. Finally, note that although the BLUEs are a function of $\mathbf{V}$ (and hence σ_A^2 and σ_e^2), applying Equation 19.2b shows that the phenotypic variance σ_z^2 in $\mathbf{V}$ cancels out in Equation 19.3a ($\mathbf{V}^{-1}$ scales as $1/\sigma_z^2$, while the inverse of $\mathbf{V}^{-1}$ scales as σ_z^2), leaving the BLUE estimate (for Equation 19.1) as a function of just the heritability.

Again assuming $\mathbf{V}$ is exactly known, the **best linear unbiased predictor** (**BLUP**) of the vector of random effects is given by

$$\widehat{\mathbf{a}} = \mathbf{Var}(\mathbf{a})\mathbf{Z}^T\mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right) = \sigma_A^2\mathbf{A}\mathbf{Z}^T\mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right) \tag{19.3b}$$

The BLUPs for breeding values are often called **PBVs** or **EBVs**, for **predicted** or **estimated breeding values**. Equation 19.3b is the regression of $\mathbf{a}$ on the vector of observations $\mathbf{y}$ adjusted for fixed effects ($\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}$) and then suitably scaled (by $\sigma_A^2\mathbf{A}\mathbf{Z}^T\mathbf{V}^{-1}$). Recall for a univariate regression predicting a from y that this scaling (the regression slope) is just $\sigma(a, y)/\sigma^2(y)$, while here the scaling is of the form of the covariance matrix $\sigma(\mathbf{a}, \mathbf{y}) = \sigma(\mathbf{a}, \mathbf{Z}\mathbf{a}) = \sigma_A^2\mathbf{A}\mathbf{Z}^T$ times the inverse of $\mathbf{V} = \sigma(\mathbf{y}, \mathbf{y})$. Even if the number

of random effects exceeds the number of actual observations (i.e., $p > n$), Equation 19.3b still provides unique estimates of each (provided $\mathbf{V}^{-1}$ exists). This occurs because the $p \times p$ covariance structure of the vector $\mathbf{a}$ is incorporated in the model. As with the BLUEs, BLUPs are just functions of h^2 as $\sigma_A^2 = h^2\sigma_z^2$, while $\mathbf{V}^{-1}$ scales as $1/\sigma_z^2$. An alternative expression is

$$\widehat{\mathbf{a}} = \left(\mathbf{Z}^T\mathbf{M}\mathbf{Z} + \lambda\mathbf{A}^{-1}\right)^{-1}\mathbf{Z}^T\mathbf{M}\mathbf{y}, \tag{19.3c}$$

where $\lambda = \sigma_e^2/\sigma_A^2 = (1-h^2)/h^2$ and

$$\mathbf{M} = \mathbf{I} - \mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T \tag{19.3d}$$

is the **absorption matrix** for the fixed effects (Kennedy and Trus 1993). As an aside,

$$\mathbf{H} = \mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T \tag{19.3e}$$

is referred to as the **Hat Matrix** (Hoaglin and Welsch 1978), because for OLS,

$$\widehat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y} = \mathbf{H}\mathbf{y} \tag{19.3f}$$

with $\mathbf{H}$ mapping the observed $\mathbf{y}$ values into those predicted $\widehat{\mathbf{y}}$ by the fixed-effects. Hence

$$\mathbf{M}\mathbf{y} = (\mathbf{I} - \mathbf{H})\mathbf{y} = \mathbf{y} - \widehat{\mathbf{y}} \tag{19.3g}$$

are the adjusted values for $\mathbf{y}$ once fixed effects are removed.

In practice, Equations 19.3a-c are often not used, as they require inversion of the potentially very large $(n \times n)$ matrix $\mathbf{V}$. Instead, $\widehat{\boldsymbol{\beta}}$ and $\widehat{\mathbf{a}}$ are obtained without computing an inverse by numerically solving (for example, by Gaussian elimination) **Henderson's mixed model equations**

$$\begin{pmatrix} \mathbf{X}^T\mathbf{X} & \mathbf{X}^T\mathbf{Z} \\ \mathbf{Z}^T\mathbf{X} & \mathbf{Z}^T\mathbf{Z} + \lambda\,\mathbf{A}^{-1} \end{pmatrix} \begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^T\mathbf{y} \\ \mathbf{Z}^T\mathbf{y} \end{pmatrix} \tag{19.4}$$

The careful reader might wonder why we worried so much about avoiding taking the inverse of $\mathbf{V}$, but that the mixed-model equations contain $\mathbf{A}^{-1}$, which as first blush looks just as complicated to invert as $\mathbf{V}$. However, $\mathbf{A}$ turns out to be very easy to invert (Henderson 1976, Quaas 1976). Indeed $\mathbf{A}^{-1}$ is often directly computed from the pedigree in place of $\mathbf{A}$. The variance-covariance matrices for $\widehat{\mathbf{a}}$ and $\widehat{\boldsymbol{\beta}}$ also follow from the mixed-model equations. Partition the inverse of the $(p+q) \times (p+q)$ matrix in Equation 19.4 as

$$\begin{pmatrix} \mathbf{X}^T\mathbf{X} & \mathbf{X}^T\mathbf{Z} \\ \mathbf{Z}^T\mathbf{X} & \mathbf{Z}^T\mathbf{Z} + \lambda\,\mathbf{A}^{-1} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}^T & \mathbf{C}_{22} \end{pmatrix} \tag{19.5a}$$

where $\mathbf{C}_{11}$, $\mathbf{C}_{12}$, and $\mathbf{C}_{22}$ are, respectively, $q \times q$, $q \times p$, and $p \times p$ submatrices. Using this notation, Henderson (1975) showed that the $q \times q$ covariance matrix for the BLUE vector of fixed effects $\boldsymbol{\beta}$ is given by

$$\mathbf{Var}(\widehat{\boldsymbol{\beta}}) = \sigma_e^2\,\mathbf{C}_{11}. \tag{19.5b}$$

The variance of the predicted breeding values $\widehat{\mathbf{a}}$ is a bit more subtle, as since $\mathbf{a}$ are random draws, the vector $\mathbf{a}$ of the actual breeding values has associated covariance structure $\sigma_A^2\mathbf{A}$. Hence, our concern are the **prediction error variances** (**PEVs**), the variances (and covariances) among the vector of prediction errors ($\widehat{\mathbf{a}} - \mathbf{a}$), given by the $p \times p$ matrix

$$\mathbf{Var}(\,\widehat{\mathbf{a}} - \mathbf{a}\,) = \sigma_e^2\,\mathbf{C}_{22}. \tag{19.5c}$$

Following Equation 19.3c, we can also express this as

$$\mathbf{Var}(\,\widehat{\mathbf{a}} - \mathbf{a}\,) = \sigma_e^2\left(\mathbf{Z}^T\mathbf{M}\mathbf{Z} + \lambda\mathbf{A}^{-1}\right)^{-1} \tag{19.5d}$$

The prediction error variances for *individual* breeding values are not a serious concern here, as we measure response by taking the average breeding values over all individuals within each generation. However, as we will see in Chapter 20, they can be important when we try to disentangle selection in natural populations. Finally, the covariances between estimated fixed effects and prediction errors is

$$\boldsymbol{\sigma}(\widehat{\boldsymbol{\beta}}, \widehat{\mathbf{a}} - \mathbf{a}\,) = \sigma_e^2\,\mathbf{C}_{12}. \tag{19.5d}$$

These covariance matrix expressions assume that the variance is known without error. When BLUP estimates are obtained using an *estimated* variance $\widehat{\sigma}_e^2$, this additional source of error is *not* accounted for by Equations 19.5b–d.

One concern is that the above BLUPs and BLUEs may be biased by selection. Assuming variances are known (the formal BLUP/BLUE framework), Henderson (1975) showed that these are unbiased if (i) selection decisions are based on linear combinations of data (such as truncation selection based on individual phenotypes or a linear index based on the phenotypes of an individual and its relatives) and (ii) that if selection is based on records that are adjusted for fixed effects, the model uses estimates of these fixed effects in an unbiased fashion when selection is absent. Hence, Henderson's conditions for BLUPs and BLUEs being unbiased by selection hold under many reasonable forms of artificial selection.

REML Estimation of Unknown Variance Components

The variance components (σ_A^2 and σ_e^2), or at a minimum the heritability $h^2 = \sigma_A^2/(\sigma_A^2 + \sigma_e^2)$, must be specified to obtain $\widehat{\boldsymbol{\beta}}$ and $\widehat{\mathbf{a}}$. These variances are generally unknown, but can be estimated using, for example, **restricted maximum likelihood** (**REML**). REML is closely related to BLUP, with (roughly speaking) REML estimates obtained from iterating and updating BLUP estimates until suitable convergence. REML maximizes that part of the likelihood function that is not influenced by fixed effects, with restricted likelihood referring to that part of the likelihood function unaffected by fixed effects (Patterson and Thompson 1971). Harville (1977) coined the term restricted ML, but Thompson (2008) notes that REML maximizes a *residual* likelihood, and hence prefers the term **residual maximum likelihood**. One advantage of REML estimates (over other variance estimation procedures) is that they are unbiased by the estimates of fixed effects (Patterson and Thompson 1971). We refer the reader to the extensive discussion of REML variance estimation in LW Chapter 27 for further details, and to Hofer (1998), Thompson and Mäntysaari (2004), Thompson et al. (2005), Thompson (2008), and Misztal (2008) for a review of more recent developments, including computational issues.

For the much of this chapter, we assume that BLUP variance components are first obtained by REML (relaxing this assumption when discussing Bayesian Mixed-model analysis). This two-stage approach of BLUP using estimated variance components (in place of

their true values) is called **empirical BLUP** or **REML/BLUP** (Sorensen and Kennedy 1986, Kennedy and Sorensen1988, Harville 1990). Kackar and Harville (1981) and Gianola et al. (1986, 1988) show that using the REML estimates does not result in biased values for BLUPs, but that the resulting predictors may not be "best" (there may be other linear predictors with smaller errors).

The covariance matrix for the REML estimates can be approximated by using the best quadratic fit of the restricted likelihood surface, centered at the REML estimates (Smith and Graser 1986, Graser et al. 1987). If $\boldsymbol{\sigma} = (\sigma_A^2, \sigma_e^2)^T$ is a vector of assumed variances, one computes the restricted likelihood $L(\boldsymbol{\sigma})$ for a grid of values close to the REML solution and then fits the best quadratic surface to the data,

$$L(\boldsymbol{\sigma}) = \mathbf{b}_0 + \boldsymbol{\sigma}^T \mathbf{b}_1 + \boldsymbol{\sigma}^T \mathbf{S} \boldsymbol{\sigma} \tag{19.6a}$$

where $\mathbf{b}_0$ and $\mathbf{b}_1$ are vectors and $\mathbf{S}$ a the symmetric matrix (Chapter 29 discusses fitting the best quadratic surfaces in the context of fitness surface estimation). The approximate covariance matrix for the vector of REMLs, $\widehat{\boldsymbol{\sigma}}$, is given by

$$\mathbf{Var}(\widehat{\boldsymbol{\sigma}}) \simeq (-2\mathbf{S})^{-1} \tag{19.6b}$$

The rationale for this approach is that, for large samples , the inverse of the matrix of second-order partial derivatives of the likelihood surface at the likelihood estimate approaches the covariance matrix of these estimates (LW Appendix 4). Equation 19.6a is a (second-order) multidimensional Taylor series (Appendix 6), with $2\mathbf{S}$ corresponding to the matrix of second-order partial derivatives (the **Hessian matrix**) of the likelihood function. Even though Equation 19.6 allows a (large-sample) estimate of the uncertainty in our estimate of the variance, this cannot easily be incorporated to provide a measure of how much additional uncertainty this introduces into BLUP estimates. As discussed later, Bayesian approaches are exact (given a specified prior) for any sample size and fully incorporate the uncertainty in the variance estimates into the uncertainty in the BLUP estimates.

REML Can (Often) Return Variance Estimates Unbiased by Selection

As with BLUPs and BLUEs (which assume variances are known), REML estimates of variance are often unbiased by selection. As we show below, under the infinitesimal model the relationship matrix $\mathbf{A}$ fully accounts for any change in σ_A^2 from disequilibrium and/or inbreeding/drift. As a result, if the base population consists of unselected and noninbred individuals in linkage equilibrium ($d(0) = 0$) and phenotypic data are available for all selected and unselected individuals, then under the infinitesimal model REML yields essentially unbiased estimates of the additive genetic variance in the base population (Henderson 1949, Henderson et al. 1959, Curnow 1961, Thompson 1973, Rothschild et al. 1979, Sorensen and Kennedy 1984b, Gianola and Fernando 1986, Gianola et al. 1988, Juga and Thompson 1989, Gianola et al. 1989, Fernando and Gianola 1990). Simulations by van der Werf and de Boer (1990) show that if the model includes the pedigree information for all individuals ($\mathbf{A}$) but is missing records (trait values) for some, then REML does not necessarily yield unbiased estimates. Simulations by Jeyaruban and Gibson (1996) showed that any such bias may vary with heritability, finding (for their model) that it increased with heritability.

When the base population consists of previously selected individuals, REML no longer provides protection from biased estimates of the additive genetic variance in the population prior to selection, even if the entire pedigree of individuals back to the base population is included (van der Werf 1990, van der Werf and de Boer 1990, van der Werf and Thompson 1992). This arises because $d \neq 0$ in the base population. While Equation 16.7 allows us to predict the dynamics of d_t, it requires the value of d_{t-1}. Without knowledge of the

actual base-population d value, $\mathbf{A}$ cannot fully account for the dynamics of d. Finally, if selection acts on a suite of unmeasured characters whose breeding values are correlated with characters included in the model, REML can generate biased estimates of the variances and covariances of the measured characters (Schaeffer and Song 1978).

ANIMAL-MODEL ANALYSIS OF SELECTION EXERIMENTS

The basic building block of mixed-model analysis of selection experiments is the **animal model** (Quaas and Pollak 1980), which estimates the breeding (or additive genetic) values of all individuals measured during the course of experiment. We examine its simplest version first, considering various elaborations in later sections (also see Chapter 22 and LW Chapter 26). While this model has its origin in the animal breeding literature, it has very widespread applicability. We trust that plant scientists will not be greatly offended, as the "animal" (or better yet, the "individual") model can be used to analyze plant selection experiments as well!

The Basic Animal Model

To apply the animal model to a selection experiment, first vectorize the observations from the entire experiment by letting y_{ij} denote the trait measurment on the jth individual from generation i, where $0 \le i \le t$ (generation 0 representing the unselected base population) and $1 \le j \le n_i$. Let the vector $\mathbf{y}$ denote the observations on all measured individuals from the entire experiment,

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_t \end{pmatrix}, \qquad \text{where} \qquad \mathbf{y}_i = \begin{pmatrix} \mathbf{y}_{i1} \\ \vdots \\ \mathbf{y}_{in_i} \end{pmatrix}$$

The vector $\mathbf{y}_i$ includes the values for all measured individuals from generation i, including those culled as well as those allowed to reproduce. The simplest animal model for these data is

$$y_{ij} = \mu + a_{ij} + e_{ij} \tag{19.7a}$$

where μ is an overall mean, a_{ij} the breeding value of the jth measured individual from generation i, and e_{ij} the deviation between breeding and phenotypic values. With exactly one record per individual, $\mathbf{Z} = \mathbf{I}$. In this simple model the only fixed effect is the mean, giving $\boldsymbol{\beta} = (\mu)$ and $\mathbf{X} = \mathbf{1}$ (a vector of ones), reducing Equation 19.1 to

$$\mathbf{y} = \mu \cdot \mathbf{1} + \mathbf{a} + \mathbf{e} \tag{19.7b}$$

where $\mathbf{a}$ is the vector of breeding values for all individuals measured during the course of the experiment, with $\mathbf{Var}(\mathbf{a}) = \sigma_A^2 \mathbf{A}$. The relationship matrix $\mathbf{A}$ is the key to mixed-model analysis, as it includes all the pedigree information. Because of this information, the animal model is easily extended to allow breeding values to be estimated for individuals without records, *provided* they have measured *relatives* in the analysis (see Example 19.5). The diagonal elements of $\mathbf{A}$ describe the amount of inbreeding, with $A_{ii} = (1 + f_i) = 2\Theta_{ii}$, while the off-diagonal elements given by $A_{ij} = 2\Theta_{ij}$ (twice the coefficient of coancestry, see LW Chapters 7, 26) describe the relatedness of individuals i and j. Recursive methods for obtaining the elements of $\mathbf{A}$ (and $\mathbf{A}^{-1}$) given a pedigree are discussed in LW Chapter 26. The simple animal model assumes that all genetic variance is additive, so that there is no (genetic)

covariance between residuals. In this case, it is generally assumed that $\mathbf{Var}(\mathbf{e}) = \sigma_e^2\,\mathbf{I}$, and the mixed-model equations (Equation 19.4) simplify to

$$\begin{pmatrix} n & \mathbf{1}^T \\ \mathbf{1} & \mathbf{I} + \lambda\,\mathbf{A}^{-1} \end{pmatrix} \begin{pmatrix} \widehat{\mu} \\ \widehat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} n\,\overline{y} \\ \mathbf{y} \end{pmatrix} \tag{19.7c}$$

where n is the total number of individuals in the experiment, $\lambda = \sigma_e^2/\sigma_A^2 = (1-h^2)/h^2$, $\widehat{\mathbf{a}}$ is the n-dimensional vector of the predicted breeding values of all measured individuals, and $\mathbf{1}$ an n-dimensional vector of ones. Likewise, the covariance matrices for the fixed effects $\mathbf{C}_{11}$, predictor errors for the BLUPs of breeding values $\mathbf{C}_{22}$, and the covariances $\mathbf{C}_{12}$ between these estimators are given by

$$\begin{pmatrix} n & \mathbf{1}^T \\ \mathbf{1} & \mathbf{I} + \lambda\,\mathbf{A}^{-1} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}^T & \mathbf{C}_{22} \end{pmatrix} \tag{19.7d}$$

Response is Measured by Change in Mean Breeding Values

Under a mixed-model analysis, response is measured by the change in the mean breeding value of a selected population over time. Mixed-models easily allow for overlapping generations by simply predicting breeding values at discrete time points (say every year) instead of each generation. Hence, in what follows, one can easily replace "generation" by "year" or some other time measure.

The estimated mean breeding value in generation k is simply given by the average of individual breeding value estimates for that generation,

$$\widehat{\overline{a}}_k = \frac{1}{n_k}\sum_{j=1}^{n_k} \widehat{a}_{kj} \tag{19.8a}$$

Total response at generation k is $\overline{a}_k - \overline{a}_0 = \overline{a}_k$, as the predicted mean breeding value from generation 0 (the unselected base population) is zero by construction ($\overline{a}_0 = 0$). In matrix notation, the vector $\overline{\mathbf{a}}$ of mean breeding values is estimated by

$$\widehat{\overline{\mathbf{a}}} = \begin{pmatrix} \widehat{\overline{a}}_0 \\ \vdots \\ \widehat{\overline{a}}_{t-1} \end{pmatrix} = \mathbf{K}^T\widehat{\mathbf{a}} \tag{19.8b}$$

where elements in the ith row of $\mathbf{K}$ are $1/n_j$ when the column corresponds to an individual from generation j, otherwise are zero (see Example 19.1). Thus, for t generations of data (corresponding to $t-1$ generations of selection, as the analysis includes the unselected base population, generation 0), $\mathbf{K}$ is $n \times t$, with $\mathbf{K}^T\mathbf{1}_n = \mathbf{1}_t$ (a $t \times 1$ vector of ones).

From Equation 19.5d, and recalling that $\mathbf{Var}(\mathbf{Bx}) = \mathbf{B}\,\mathbf{Var}(\mathbf{x})\mathbf{B}^T$ (LW Equation 8.21b), the sampling covariance matrix for the vector of estimated genotypic means becomes

$$\mathbf{Var}\left(\widehat{\overline{\mathbf{a}}}\right) = \sigma_e^2\,\mathbf{K}^T\mathbf{C}_{22}\mathbf{K} \tag{19.8c}$$

where the $n \times n$ matrix $\mathbf{C}_{22}$ is the solution to Equation 19.7d (under the simple animal model) or more generally by Equation 19.5c. Again, these expressions assume that the residual variance is known without error. Using an estimate for σ_e^2 adds an additional source of uncertainty.

Example 19.1. As a toy example of how one performs a mixed-model analysis of a selection experiment, consider the following very simple situation. From a base population of unrelated and non-inbred individuals, four (indexed by 1-4) are measured and have trait values (e.g., size) of 3, 6, 5, and 2, respectively. The two largest individuals are mated, and their resulting offspring (individuals 5-8) have values 4, 5, 6, 5. Now suppose we have either a REML-based estimate of the heritability (not advisable here given the very small sample size) or have previous knowledge of its value, so that our goal is to estimate the size of the genetic response. Assuming the only fixed effect is the mean, the resulting animal model is $\mathbf{y} = \mathbf{1}\boldsymbol{\beta} + \mathbf{a} + \mathbf{e}$, where

$$\mathbf{y} = \begin{pmatrix} 3 \\ 6 \\ 5 \\ 2 \\ 4 \\ 5 \\ 6 \\ 5 \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \end{pmatrix}, \quad \boldsymbol{\beta} = (\,\mu\,)$$

What is the relationship matrix $\mathbf{A}$? Since individuals 2 and 3 are the parents, and all offspring are full-sibs, all related individuals have values of $1/2$ as $2\theta_{ij} = 1/2$ for both parent-offspring and full-sibs. The resulting numerator relationship matrix becomes

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1/2 & 1/2 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 & 1/2 & 1/2 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 1 & 1/2 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 & 0 & 1/2 & 1 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 & 0 & 1/2 & 1/2 & 1 & 1/2 \\ 0 & 1/2 & 1/2 & 0 & 1/2 & 1/2 & 1/2 & 1 \end{pmatrix}$$

For example, individuals 2 and 5 are parent and offspring, so $A_{5,2} = A_{2,5} = 1/2$, as cov(P,O) = $\sigma_A^2/2$. Similarly, individuals 7 and 8 are full-sibs, so that $A_{7,8} = A_{8,7} = 1/2$. Note that the relationship matrix for the founders (base population members) is given by the 4×4 identity submatrix in the upper left of $\mathbf{A}$. This identity matrix implies that non-inbred (diagonal elements are one) and unrelated individuals (off-diagonal elements zero) formed the base population.

Turning to the covariance matrix of the residuals, we make the standard assumption that $\mathbf{Var}\,(\mathbf{e}) = \sigma_e^2\,\mathbf{I}$, i.e., all residuals are uncorrelated with common variance σ_e^2. However, if there is dominance, the residuals among full-sibs are inflated by $\sigma_D^2/4$. Likewise, if there are common family effects (e.g., maternal effects or other shared environmental effects) the residuals are inflated by σ_c^2, where c the is common family effect. We ignore these possible complications for now, which are easily accounted for by adding additional random effects to the model (see below).

Suppose (from REML or prior knowledge) that the heritability of the trait is $h^2 = 0.3$. Applying Equation 19.3a gives

$$\widehat{\mu} = \left(\mathbf{1}^T\mathbf{V}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}^T\mathbf{V}^{-1}\mathbf{y} = 4.22$$

where we have computed $\mathbf{V}$ using Equation 19.2b scaled to remove the phenotypic variance σ_z^2, e.g., $\mathbf{V} = 0.3\,\mathbf{A} + 0.7\,\mathbf{I}$. Substituting into Equation 19.3b gives the BLUPs for the genetic

values and resulting genetic means as

$$\widehat{\mathbf{a}} = \begin{pmatrix} -0.366 \\ 0.666 \\ 0.366 \\ -0.666 \\ 0.386 \\ 0.562 \\ 0.739 \\ 0.562 \end{pmatrix}. \quad \text{Here} \quad \mathbf{K} = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad \text{implying} \quad \mathbf{K}^T\widehat{\mathbf{a}} = \begin{pmatrix} 0 \\ 0.562 \end{pmatrix}$$

Note that (by construction) the mean breeding value in the base population is zero. Hence, the estimated response (for $h^2 = 0.3$) is 0.562. The estimated genetic gain (response) for different assumed heritabilities are found to be as follows:

h^2	Estimated response	h^2	Estimated response
0.0	0	0.6	0.940
0.1	0.211	0.7	1.026
0.2	0.398	0.8	1.083
0.4	0.707	0.9	1.095
0.5	0.833	1.0	1

The estimated gain increases with the assumed h^2 until it reaches a maximum of $\simeq 1.098$ for $h^2 = 0.86$, after which it decreases as the assumed h^2 increases. A Bayesian analysis removes this dependency of the estimated response on h^2 by computing a weighted average of response over all possible h^2 values (weighted by their posterior values), yielding in a marginal posterior distribution for the response that is independent of the heritability.

Turning to a more standard analysis, the selection differential is the mean of the selected parents minus the mean of all parents, $S = 5.5 - 4 = 1.5$. Likewise, the response is the mean of the offspring minus the mean in the previous generation, $5 - 4 = 1$, giving a realized heritability of $R/S = 1.0/1.5 = 0.67$. Using this value for h^2 in the MM analysis returns a genetic gain of one.

A mixed-model analysis of a selection experiment has a very different character from a LS analysis. In the latter, one estimates the realized heritability from a suitable regression of phenotypic means on selection differentials. Under a REML/BLUP analysis, one first estimates the heritability in the base population (using REML), and then uses (empirical) BLUP to estimate breeding values for all individuals. The genetic component of response in a given generation is given by the mean of the estimated breeding values within that generation. The information on the vector **a** of breeding values provided by the relationship matrix **A** allows us to separate genetic from environmental response even in the absence of a control population. Thompson and Atkins (1994) note that a fundamental difference between the two approaches in separating genetic from environmental change: a LS analysis typically uses *between-population* information (e.g., selection vs. control, up- vs. down-selected lines), while a REML/BLUP analysis uses *within-population* information (the connections provided via **A**).

Under a MM analysis, the correct estimate of heritability in a mixed-model analysis should be based on REML estimates of the base-population variance components, $\widehat{h}^2 = \widehat{\sigma}^2_A/(\widehat{\sigma}^2_A + \widehat{\sigma}^2_E)$. One must avoid the temptation to estimate heritabilities using estimated mean breeding values. For example, Blair and Pollak (1984) regressed the BLUP-estimated

mean breeding values on cumulative selection differentials to obtain a realized heritability estimate. The problem with this approach is that these mean estimates depend on the *assumed*, rather than the actual, heritability (Thompson 1986, Sorensen and Johannsson 1992, Ollivier 1999). In contrast, under a Bayesian analysis, the marginal posterior estimate of mean breeding value in any particular generation averages over all possible h^2 values, and is independent of the heritability (Example 19.9). A regression of these mean breeding values on generations can indeed return an unbiased estimate of realized heritability.

Fixed Effects Alter Heritabilities

Variance components in a mixed model are estimated *after* the variation introduced by fixed effects is removed, and this has an impact on the estimated heritability. Under the mixed model framework, $h^2 = \sigma_A^2/(\sigma_A^2 + \sigma_e^2)$, where σ_e^2 is the *residual* variance. This is potentially different from the more standard $h^2 = \sigma_A^2/\sigma_z^2$, which is a function of the phenotypic variance σ_z^2. In the absence of fixed effects that differ over individuals (such as sex- or age-specific means), $\sigma_z^2 = \sigma_A^2 + \sigma_e^2$, and the two definitions are equivalent. However, if such class-specific differences exist and are explicitly modeled in our analysis, then the variation they contribute to the overall phenotype is *removed*, and the resulting residual variance reduced, so that $\sigma_z^2 > \sigma_A^2 + \sigma_e^2$. As a result, mixed-model heritability estimates should be larger than for estimates ignoring fixed effects. Further, for different models using the same data, but different assumed fixed effects, the residual variance σ_e^2 (and thus the resulting heritability) can differ significantly (Wilson 2008). The implication is that comparison of heritabilities for different traits/populations estimated under a mixed-model framework can be somewhat problematic due to differences in the incorporated fixed effects. For example, if the trait is strongly influenced by sex, an analysis ignoring sex has a larger residual error. For a model with sex added as a fixed effect, this source of variation is removed, resulting in a smaller residual variance and a larger heritability.

Model Validation

Given the sensitivity of a mixed-model analysis to the validity of the assumptions (in particular, the infinitesimal model), some form of model validation is required to apply MM methods with confidence. One approach is to test the infinitesimal model prediction that estimates of the base population σ_A^2 should remain stable as additional generations of selection are considered. If the infinitesimal model holds, $\mathbf{A}$ completely accounts for changes in the additive variance in these later generations of selection (a point more fully developed shortly). On the other hand, if σ_A^2 is changing in ways not predictable from the infinitesimal model (e.g., significant allele frequency change), using data from additional generations of selection may result in dramatically different estimates of the base-population additive variance.

Example 19.2. One of the first REML/BLUP analyses of a selection experiment was by Meyer and Hill (1991), who examined the response to selection for adjusted food intake (AFI) in mice (Figure 19.1). AFI is defined as food intake between 4 and 6 weeks corrected for 4-week weight. Meyer and Hill had three replicates, each consisting of high, low, and control lines, for a total of almost 11,000 mice over the course of the experiment. Within-family selection (Chapter 21) on AFI was followed for 23 generations. Meyer and Hill included a number of fixed effects in their model, as well as adding a random effect to control for common litter (i.e., family) effects (see Example 19.7 for details).

As a check of the validity of the assumptions (in particular, the infinitesimal model), Meyer and Hill compared variance estimates based on data from generations 5–7 with estimates based

on generations 14–23. In both cases, the full pedigree structure was incorporated into the numerator relationship matrix (i.e., the relationships among individuals in generation 5 were used to generate the submatrix of $\mathbf{A}$ associated with this generation, and similarly for future generations). While incorporation of the complete pedigree information reduces the bias in estimates of the base-population additive variance, it does not completely reduce the bias if the records for all individuals from previous generations (back to the base population) are ignored (van der Werf and de Boer 1990). Even with this caveat in mind, Meyer and Hill observed a dramatic decline in the estimated additive variances (from 19.2 based on generations 5–7 to 2.5 based on generations 14–23). Under the infinitesimal model, both estimates should be for the base population variance. This large decrease suggests that the infinitesimal model may not be appropriate for this trait. It is interesting to note that this decrease occurred even as the total variance increased dramatically (from 23.88 to 33.93). This increase resulted mainly from an increase in the environmental variance (from 12.9 to 25.5), although there was also a slight increase in the litter-effects variance (from 4.78 to 5.96).

Several other REML/BLUP analyses of selection experiments in mice also found differences in estimates of base population additive variance when comparing data from early versus late generations. Beniwal et al. (1992a,b) observed decreases in the additive variance (in body weight, litter size, and lean mass), while Heath et al. (1995) observed an increase in the additive variance in body weight.

In contrast, Martinez et al. (2000) found no changes in variance estimates over 20 generations of selection for body composition (fat pad to body weight ratio) in mice. These authors examined REML estimates of the additive variance (and heritability) using various subsets of the full twenty-generation data. Consistent estimates of the additive variance were obtained using all records and the complete pedigree from generations 0-20, using records from generations 9–20 but with pedigree information from generation 0, and finally looking at the phenotypic data in blocks of three generations. They conclude that the selection response, while resulting in a roughly four-fold change in mean, was still well-fit by the infinitesimal model.

Separating Genetic and Environmental Trends

Observed improvement in a trait over time (such as milk yield) could be do entirely to improvement in the environment (better husbandry and nutrition), entirely from genetic changes (response from selected breeding), or (most likely) a combination of both. Thus, it is critical to partition an observed phenotypic change into genetic and environmental components. For example, Southwood and Kennedy (1991) showed that the improvement in several litter-size related traits in pigs over a ten year period in Quebec was entirely due to environmental, rather than genetic, changes.

In a least-squares analysis, any underlying environmental trend is assumed to be removed by contrasting selected and control populations (or contrasting populations selected in opposite directions). The rational is that the kth individual from population j in generation t can be written as

$$y_{tjk} = \mu + d_t + a_{tj} + e_{tjk} \tag{19.9}$$

where d_t is the environmental trend. If the common environmental value is the same in both selected and control populations, then the difference in phenotypic means in generation t is

$$\overline{y}_{s,t} - \overline{y}_{c,t} = (\,\overline{a}_{s,t} - \overline{a}_{c,t}\,) + (e_{s,t} - e_{c,t}) \tag{19.10}$$

The residuals e have expected value zero, hence the contrast provides an estimate of $\overline{a}_{s,t}$, *provided* there is no significant drift in the mean breeding value of the control population (so that $\overline{a}_{c,t} \simeq 0$). However, if genotype-environment interactions are present, the environment

value for generation t can differ between populations, in which case Equation 19.9 has an additional term $(d_{s,t} - d_{c,t})$. Hence, even when a control population is used, a least-squares analysis can still give biased results if there is significant drift in the mean of the control population ($|\,\overline{a}_{c,t}\,| \gg 0$) and/or significant $G \times E$.

Another case where the use of a control can be misleading was offered by Su et al. (1997), who examined response in body weight in chickens starting from a base population with a known previous history of selection on this trait. Because of prior selection, this base population showed a slippage of the mean back to the original (unselected) value. Using this population as a control would assume this slippage reflects a decay in the environment over time, resulting in an overestimation of the response. Because of this concern, Su et al. used a mixed model, instead of a control, to estimate the genetic response.

A mixed-model analysis estimates the mean *breeding value*, rather than the *phenotypic mean*, of the population. Hence a MM analysis allows for the separation of the genetic change from any environmental change (Henderson et al. 1959, Blair and Pollak 1984, Sorensen and Kennedy 1984a, Kennedy 1990). This occurs because $\mathbf{A}$ tracks the flow of genes through the population, allowing for estimates of breeding values independent of environmental effects. Of course, this is strictly dependent on the model assumptions holding, but if they do, a mixed-model analysis does not require a control population. This being said, Sorensen et al. (2003) showed that the inclusion of a control in a MM analysis generally improves the efficiency of estimates (smaller standard errors).

Common environment effects are incorporated into the basic animal model by simply adding a fixed effect d_t for common environmental effect in generation t,

$$y_{tj} = \mu + d_t + a_{tj} + e_{tj} \tag{19.11}$$

For T generations, there are T estimable fixed effects (μ and all but one d_i). Typically one could either constrain the d_i to sum to zero or arbitrarily set one equal to zero. We do this with d_1, so that the vector of fixed effects becomes

$$\boldsymbol{\beta} = (\mu, d_2, \cdots, d_T)^T$$

and the corresponding incident (or design) matrix $\mathbf{X}$ has ones in the columns (two through T) corresponding to the generation in which the individual was scored, viz.,

$$\mathbf{X} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 1 & 1 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 1 & 0 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \end{pmatrix}$$

An alternative approach is to treat the d_i as random effects, drawn independently (i.e., uncorrelated across generations) from, say, a normal distribution with mean zero and unknown variance σ_d^2. The assumption of no environment al correlation between adjacent generations can be delicate (as best), and treating the d_i as fixed removes these concerns (although at a cost of absorbing more degrees of freedom than a random-effects analysis).

Example 19.3. To examine the potential bias from $G \times E$ and drift in the control population, Blair and Pollak (1984) examined a seven-generation selection experiment on 14-month greasy

fleece weight in sheep. The model they assumed was that the *m*-th individual in generation *t* with fixed sex effect (male/female) *i*, fixed age of dam effect (mature/immature) *j*, and fixed rearing rank effect (single/twin) *k* had a phenotypic value of

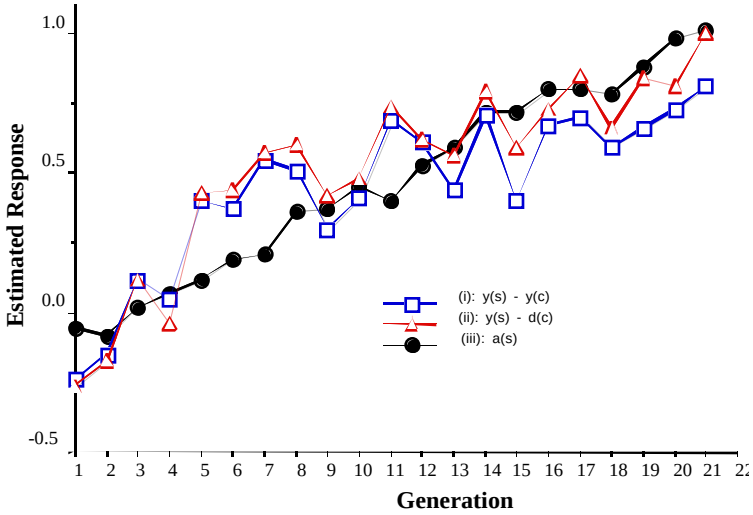
$$y_{tm} = sx_i + b_j + r_k + d_t + a_{tm} + e_{tm}$$

In matrix form, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Za} + \mathbf{e}$, where the vector $\boldsymbol{\beta}$ contains the fixed-effects for sex (*sx*), dam age (*b*), and rearing rank (*r*) in addition to the effects for years (*d*). Both the selected and control lines were subjected to a BLUP analysis using this model, and three different estimates of selection response were considered:

(i) $\widehat{\overline{y}}_{s,t} - \widehat{\overline{y}}_{c,t}$, the estimated phenotypic means following adjustment for the fixed-effects (sx, b, and r), obtained by $\widehat{\overline{y}}_{x,t} = \widehat{\overline{a}}_{x,t} + \widehat{d}_{x,t}$, where $x = c$ or s. This is an unbiased estimate of the response if there is no significant drift in the control population ($\overline{a}_{c,t} \simeq 0$) and no $G \times E$, so that $d_{c,t} = d_{s,t}$.

(ii) $\widehat{\overline{y}}_{s,t} - \widehat{d}_{c,t}$, the (fixed-effects adjusted) phenotypic mean of the selected population minus the common environmental effect, as estimated from the control population.

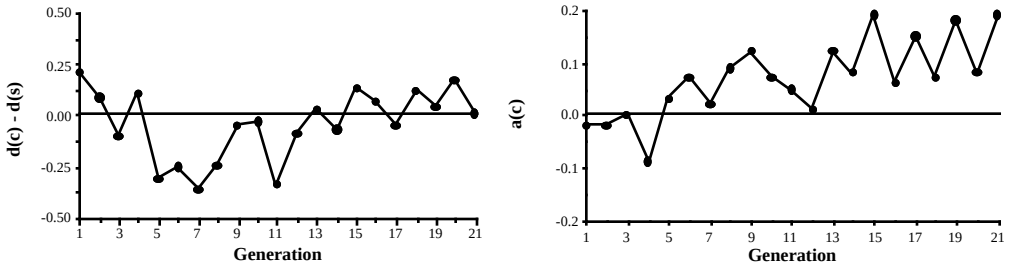
(iii) $\widehat{\overline{a}}_{s,t}$, the BLUP estimate of the mean breeding value in the selected population.



Estimate **(i)** mimics that used in a least-squares analysis, and Blair and Pollak show it is independent of the assumed heritability. Estimates **(ii)** and **(iii)** are highly dependent on the assumed (or estimated) heritabilties in the control and selected populations. As the above figure shows, all three estimates show a positive genetic trend (following a reversed response over the first few generations). The estimated response using only the predicted mean breeding value is smoother (as expected) than the other two estimates.

The potential biases in a least-squares analysis of these data using the contrast between the control and selected phenotypic means are seen in the two graphs below. The left graph plots the difference in the estimated common environmental effects between selected and control populations ($\widehat{d}_{ct} - \widehat{d}_{st}$). Ignoring the inherent variance in estimating the $\widehat{d}$'s, the difference (via a paired *t*-test) is not significantly different from zero. The right graph plots the predicted mean breeding value of the control population, which is assumed to be zero under the least-squares analysis. As the figure shows, there is a slight, but positive, trend. When the control mean is subtracted off, the net result is that the LS analysis slightly underestimates the true

response. Thus, there is no evidence of error being introduced by $G \times E$ differences between the control and selected lines, but error is introduced by the mean breeding value of the control population departing significantly from zero.



Validation That a Trend is Indeed Genetic

The estimated environmental (d) and additive-genetic (a) effects are highly-dependent on the estimated (or assumed) base-population heritability h^2. Hence, using BLUP to separate genetic from environmental values is highly dependent on the heritability used being close to its true value. Other departures from mixed-model assumptions (e.g., the infinitesimal model, and that BLUP and REML estimates are unaffected by selection) can also result an incorrect assignment of the relative importance of environmental versus genetic values. One must be cautious when relying on a BLUP analysis to separate genetic from environmental effects, and some sort of validation of the trend is critical.

We have already discussed validation of the general animal model, namely by looking at the consistency of the estimated additive variance using different subsets of the data. Similar validation that a trend is indeed genetic can be performed by again looking for consistency across the analysis. For example, if a control population is used, a BLUP analysis can estimate the amount of drift in the mean breeding value from its expected value of zero and estimates of the selection and control environmental effects can be compared (e.g., Example 19.3). If these are reasonably consistent, then a joint analysis (assuming the same environmental values in both populations) may yield more precise estimates of the generational environmental values. Likewise, if the estimates are significantly different, the possibility of either $G \times E$ and/or different environmental values in selected versus controlled lines needs to be seriously considered. Even when using a control population, there is still much to be gained by subjecting each to a mixed-model (e.g., BLUP) analysis.

Even in the absence of a control population, one can still attempt trend validation. For example, Boichard et al (1995) examined several different methods to attempt validation of genetic trends in dairy cattle, all of which involving comparing predicted trend values using different subsets of the data. The authors were interested in comparing the performance of a model (AM90) used for French Holsteins from 1990 to 1992 with a more recent model (AM93) used since 1993. For a variety of reasons, there was concern that the older AM90 model yielded biased estimates of any trend. One check was based on the fact that milk yield data is in the form of multiple records per individual so that a repeatability model is appropriate (Example 19.6). The authors compared estimates of BVs based only on the first lactations with estimates based on using all lacations. They found that the trend estimated from the first versus all lactations agree well under the AM93 model, but differ dramatically under the AM90 model. Two other measures of consistency (for example, looking at the stability of estimated breeding values of individuals as more information is added) also showed that the newer AM93 model seem relatively robust, while the older AM90 model seemed to produce biased estimates of the trend.

Replicate Lines

It is straightforward to jointly analyze multiple lines simultaneously. For n lines, write the total vector of observations as $\mathbf{y}^T = (\mathbf{y}_1^T, \mathbf{y}_2^T, \cdots, \mathbf{y}_n^T)$ where $\mathbf{y}_k$ is the vector of total observations from line k. If the generational environmental effects are assumed to be the same across lines, the model for the ith individual in generation t from line k is

$$y_{k,ti} = \mu + d_t + a_{k,ti} + e_{k,ti} \tag{19.12a}$$

Alternatively, if the environmental effects are potentially different in each population,

$$y_{k,ti} = \mu + d_{k,t} + a_{k,ti} + e_{k,ti} \tag{19.12b}$$

The power of combining multiple lines arises when environmental effects can be assumed to be the same across lines. In this case, the effective sample size for estimating each effect is increased, and (presumably) the resulting sampling variance decreased, improving the precision of the estimates. The assumed covariance matrix for the vector of joint breeding values can also take several forms. If the founding members for each line are drawn from the same base population (but otherwise unrelated), then the covariance matrix for the vector of breeding values $\mathbf{a}$ has block-diagonal form, with the ith block corresponding to $\sigma_A^2\mathbf{A}(i)$, the numerator relationship matrix for line i times the base-population additive variance. If the founding members of at least some different lines are related, then $\mathbf{A}$ is more complex, reflecting these relationships. Further modifications for joint analysis have been proposed by Visscher and Thompson (1990), and extended by Beniwal et al. (1992a,b) and Heath et al. (1995) by allowing the additive variance to change (discussed below). For example, one might assume that the additive variance remains constant for the first few generations of selection, after which it assumes a different value. This is a logical, but still ad-hoc, approach towards dealing with potential departures from the infinitesimal model.

Estimating the Additive Variance at Generation t

Even under the infinitesimal model, the additive variance changes over time due to the generation of disequilibrium (Chapter 16) and inbreeding (Chapters 3, 24), a point more fully unpacked in the next section. While REML provides an estimate of the base-population additive variance (which is unbiased provided the model assumptions hold), it does not immediately provide estimates of the actual additive variance in any particular generation of selection. The most straightforward approach is to use the parent-offspring regression for each generation of selection to estimate the additive genetic variance in the parents. With parents from generation t and offspring in generation $t+1$, the regression estimates the heritability of the parents, $h_A^2(t)$ (Robertson 1977). The drawback to this approach is the typically small sample size associated with each generation (resulting in large standard errors for each heritability/variance estimate). Ideally, one would like to be able to combine information across generations in such a way as to improve the variance estimates.

Sorensen and Kennedy (1984b) suggest one approach for combining information to estimate the variance in generation t, which is to use a mixed-model analysis treating generation t as the base population. In particular, one considers only the data from generation t onward (say to generation T) and the relationship matrix is adjusted to assume that generation t is the base population. The resulting covariance matrix for the breeding values becomes

$$\mathbf{Var}\begin{pmatrix} \mathbf{a}_t \\ \mathbf{a}_{t+1} \\ \vdots \\ \mathbf{a}_T \end{pmatrix} = \sigma_A^2(t)\begin{pmatrix} \mathbf{I} & \mathbf{A}_{t,t+1} & \cdots & \mathbf{A}_{t,T} \\ \mathbf{A}_{t,t+1} & \mathbf{A}_{t+1,t+1} & \cdots & \mathbf{A}_{t+1,T} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{A}_{t,T} & \mathbf{A}_{t+1,T} & \cdots & \mathbf{A}_{T,T} \end{pmatrix} \tag{19.13}$$

where $\mathbf{a}_k$ is the vector of breeding values in generation k. $\mathbf{A}_{jk}$ is the relationship matrix of associations between individuals in generations j and k. By taking $\mathbf{Var}(\mathbf{a}_t) = \mathbf{I}$, we are assuming that all individuals in generation t are unrelated and noninbred, as this is now our base population. All measured individuals from generation t (including those not leaving offspring) are included in the base population. While this approach seems logical, it is still somewhat ad-hoc and is not exact. Simulation studies by van der Werf and de Boer (1990) show that Sorensen and Kennedy's approach tends to overestimate the true variance. Another potential candidate would be to use the variance among the predicted breeding values within a generation, viz.,

$$\mathrm{Var}(A_t) = \frac{1}{n-1}\sum_{i=1}^{n_t}\left(\widehat{a}_{ti} - \widehat{\overline{a}}_t\right)^2 \tag{19.14}$$

While again this is a reasonable suggestion, there are complications. First, there is a level of uncertainty in that the breeding values are all estimated. Second, the assumed genetic variance used to obtain the BLUP estimates has a strong influence on the values of the estimated $\widehat{a}_{ti}$. Finally, the variance of the $\widehat{a}$ estimates $\rho^2\sigma_A^2$, where $\rho^2 < 1$ is the accuracy of the predicted breeding values (Equation 20.23c), so that Equation 19.14 is expected to underestimate the true variance. Sorensen et al. (2001) note that the use of Equation 19.14 in a *Bayesian* framework (wherein the uncertainty in variance estimates is naturally incorporated into the analysis) avoids many of these problems.

THE RELATIONSHIP MATRIX ACCOUNTS FOR DRIFT AND DISEQUILIBRIUM

There are three potential sources of change in the additive variance σ_A^2 during a selection experiment: allele-frequency change from selection, allele-frequency change from drift, and gametic-phase disequilibrium generated by selection (Chapters 5, 16). Under the infinitesimal model, selection does not change allele frequencies, leaving only drift and disequilibrium as potential agents of change. If the average effect for any allele is small, allele frequency changes are expected to be small for at least a few generations (Chapters 5, 25). In these settings, the relationship matrix $\mathbf{A}$ in a MM analysis fully accounts for the effects of gametic-phase disequilibrium as well as genetic drift (Sorensen and Kennedy 1983, 1984a), and hence for all significant changes in σ_A^2, *provided* the base population consists of unrelated, non-inbred individuals in linkage equilibrium. As a result, as long as selection-induced allele frequency change is negligible, the variance-covariance matrix of the vector of breeding values remains the product of the base-population additive genetic variance and the numerator relationship matrix, $\mathbf{Var}(\mathbf{a}) = \sigma_A^2\mathbf{A}$.

To both see this point, and to provide connections among these three potential agents of change, recall two important concepts from Chapter 16. First (Equation 16.2), the additive genetic variance can be decomposed as $\sigma_A^2 = \sigma_a^2 + d$, where d is the disequilibrium contribution and σ_a^2 the additive *genic* variance — the additive variance under linkage equilibrium and Hardy-Weinberg. Provided allele frequencies do not change, σ_a^2 is unchanged. For additive loci under drift, the expected genic variance at generation t is just $\sigma_a^2(t) = \sigma_a^2(0)(1-f_t)$, where f_t is the inbreeding coefficient (Equation 16.9c). Second (Equation 16.8a), the regression of the breeding value A_i of an individual on the breeding values A_{m_i} and A_{f_i} of its parents is

$$A_i = \frac{1}{2}A_{f_i} + \frac{1}{2}A_{m_i} + s_i, \tag{19.15a}$$

where the **segregation residual** s (also referred to as **Mendelian sampling**) results from segregation of alleles at heterozygous loci in the parents. The conditional independence

of the covariance relationships of the vector $\mathbf{a}$ from selection and drift (given $\mathbf{A}$) follows as a consequence of the behavior of this residual under the infinitesimal model. If the joint distribution of breeding values for parent and offspring is multivariate normal (as occurs under the infinitesimal model), this regression is linear and homoscedastic, so that s is independent of the parental breeding values. Its variance $\sigma^2_{s_i}$ for individual i is also independent of parental breeding values, but does depend on the average inbreeding $\overline{f}_i$ of the parents and the genic variance, with

$$\sigma^2_{s_i} = (1 - \overline{f}_i)\,\sigma^2_a/2 \tag{19.15b}$$

Provided the vector $\mathbf{s}$ of Mendelian sampling residuals remains multivariate normal, then $\mathbf{s} \sim \text{MVN}(\mathbf{0}, (\sigma^2_a/2)\,\mathbf{F})$. The matrix $\mathbf{F}$ is diagonal with ith element $(1 - \overline{f}_i)$, one minus the average inbreeding of the parents of i. If k and j are the parents of offspring i, then

$$F_{ii} = (1 - \overline{f}_i) = \left(1 - \frac{f_k + f_j}{2}\right) = \left(2 - \frac{A_{kk} + A_{jj}}{2}\right) \tag{19.16}$$

where $A_{kk} = (1 + f_k)$ denotes the kth diagonal element of the relationship matrix $\mathbf{A}$. The effects of drift on the additive variance enter through the inbreeding coefficients f. Thus, the distribution of the Mendelian sampling terms $\mathbf{s}$ is unaffected by the breeding values of the parents (and hence by selection and assortative mating), while the effects of drift on Mendelian sampling are fully accounted for by $\mathbf{F}$.

The final key concept is that when we have the complete pedigree of all individuals in the selection experiment, any breeding value can be expressed as a linear function of the base population breeding values and Mendelian sampling terms ($\mathbf{s}$) not affected by selection (Sorensen and Kennedy 1984a, Kennedy et al. 1988). This is an extension of the idea of ancestral regressions discussed in Chapter 15.

Example 19.4. Suppose individuals one through four are from the base population (and assumed unrelated), while individuals five and six are the offspring from one and two and three and four, respectively. The breeding values can be written as

$$A_5 = \frac{1}{2}\,A_1 + \frac{1}{2}\,A_2 + s_5, \qquad \text{and} \qquad A_6 = \frac{1}{2}\,A_3 + \frac{1}{2}\,A_4 + s_6$$

If individual seven is the offspring of these two parents, its breeding value is

$$\begin{aligned} A_7 &= \frac{1}{2}\,A_5 + \frac{1}{2}\,A_6 + s_7 = \frac{1}{2}\left(\frac{1}{2}\,A_1 + \frac{1}{2}\,A_2 + s_5\right) + \frac{1}{2}\left(\frac{1}{2}\,A_3 + \frac{1}{2}\,A_4 + s_6\right) + s_7 \\ &= \frac{1}{4}\,(A_1 + A_2 + A_3 + A_4) + \frac{1}{2}\,(s_5 + s_6) + s_7 \end{aligned}$$

Since the mean breeding value in the base population is zero (by construction), most of the response in a selection experiment comes not from base individuals with exceptional breeding values, but rather from Mendelian sampling (segregation) generating new variation.

This linear relationship between the vector $\mathbf{a}$ of all breeding values and the vectors $\mathbf{a}_b$ (breeding values in the base population) and $\mathbf{s}$ (Mendelian segregation value) can be formally expressed as follows. Defining $\mathbf{w}^T = (\mathbf{a}_b^T, \mathbf{s}^T)$,

$$\mathbf{a} = \mathbf{T}\mathbf{w} \tag{19.17a}$$

The matrix $\mathbf{T}$ describes the flow of genes from ancestors to relatives. Indexing the elements in $\mathbf{a}$ so that ancestors precede their relatives, $\mathbf{T}$ is lower-triangular with diagonal values of one, with the submatrix corresponding to the base population (and thus no known ancestors) being an identity matrix. The off-diagonal elements for their relatives in the pedigree are given by

$$T_{i,j} = \frac{1}{2}\left(T_{s,j} + T_{d,j}\right) \quad \text{for} \quad j < i \tag{19.17b}$$

where s and d correspond the index for the sire and dam of i.

From Equation 19.17a,

$$\mathbf{Var}(\mathbf{a}) = \mathbf{Var}(\mathbf{T}\mathbf{w}) = \mathbf{T}\,\mathbf{Var}(\mathbf{w})\mathbf{T}^2 \tag{19.18a}$$

Since we assumed the base population consists of unrelated individuals ($\mathbf{A}_b = \mathbf{I}$) in gametic-phase equilbrium (so that $\sigma_A^2 = \sigma_a^2$), $\mathbf{Var}(\mathbf{a}_b) = \sigma_a^2\,\mathbf{I}$, while $\mathbf{Var}(\mathbf{s}) = (\sigma_a^2/2)\,\mathbf{F}$. Under multivariate normality, $\mathbf{a}_b$ and $\mathbf{s}$ are independent, and hence uncorrelated, giving

$$\mathbf{Var}(\mathbf{w}) = \mathbf{Var}\begin{pmatrix}\mathbf{a}\\ \mathbf{s}\end{pmatrix} = \sigma_a^2\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{F}/2\end{pmatrix} = \sigma_a^2\,\mathbf{D} \tag{19.18b}$$

where $\mathbf{D}$ is a diagonal matrix with elements of one when both parents are unknown (base population) and $0.5(1-\overline{f})$ when both parents are known. (If due to missing data only a single parent is known, the diagonal element becomes $(1/4)(3-f)$, where f is the inbreeding in the known parent, Kennedy et al. 1988). Putting Equations 19.18a and b together,

$$\mathbf{Var}(\mathbf{a}) = \sigma_A^2\mathbf{A} = \mathbf{T}\,\mathbf{Var}(\mathbf{w})\mathbf{T}^T = \sigma_a^2\mathbf{T}\,\mathbf{D}\mathbf{T}^T$$

Assuming $\sigma_A^2 = \sigma_a^2$ (the base population additive variance equals its genic variance), then

$$\mathbf{A} = \mathbf{T}\mathbf{D}\mathbf{T}^T \tag{19.18c}$$

as obtained by Henderson (1976) and Thompson (1977). As an aside, this expression coupled with the simple form for $\mathbf{T}$ is what allowed Henderson (1976) to obtain a very rapid way of computing $\mathbf{A}^{-1}$. The critical feature of Equation 19.18c is that $\mathbf{D}$ is ***independent of selection and assortative mating*** (being Mendelian sampling residuals) and accounts for the reduction in additive variance from genetic drift.

If the infinitesimal model does not hold, then residual values may indeed vary with parental breeding values, in which case selection can certainly influence the distribution of residuals. Provided that the change in allele frequencies is small over the course of the experiment, this bias may not be too serious (e.g., Maki-Tanila and Kennedy 1986). Another key assumption from the infinitesimal model is that the distribution of residuals does not significantly deviate from normality. Chapter 24 examines this rather technical issue in some detail.

When the genic variance itself is changing due to selection, knowledge of $\mathbf{A}$ is no long sufficient to account for all changes in σ_A^2. In these cases, Beniwal et al. (1992a) suggest this might be somewhat accommodated by a modification of the covariance matrix. Consider the simple case of two time blocks with different (unknown) genic variances: the breeding values $\mathbf{a}_1$ in the first block have base genic variance $\sigma_a^2(1)$, while the vector of breeding values $\mathbf{a}_2$ from the second block start with genic variance $\sigma_a^2(2)$. Decompose the diagonal matrix $\mathbf{D}$ into two blocks,

$$\mathbf{D} = \begin{pmatrix}\mathbf{D}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{D}_2\end{pmatrix}$$

where $\mathbf{D}_i$ represents block i. The covariance matrix for the breeding values can be written as

$$\mathbf{Var}(\mathbf{a}) = \mathbf{Var}\begin{pmatrix}\mathbf{a}_1\\ \mathbf{a}_2\end{pmatrix} = \sigma_a^2(1)\mathbf{T}\begin{pmatrix}\mathbf{D}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{T}^T + \sigma_a^2(2)\mathbf{T}\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{D}_2\end{pmatrix}\mathbf{T}^T \qquad (19.19)$$

These variances can be estimated separately using REML, and they correspond to the genic variances at the start of each block, which may be reduced by drift within the block. Heath et al. (1995) generalize this approach to allow the additive variance to change each generation.

MODIFICATIONS OF THE BASIC ANIMAL MODEL

One strength of a mixed-model analysis is its flexibility. For example, one can (often) predict breeding values for individuals with no records, as the following example shows.

Example 19.5. Consider a pedigree with two (unrelated) parents, each of which has an offspring (via another unrelated and unmeasured parent) which is scored. Index the parents by 0 and 1, and the offspring by 2 (parent 0) and 3 (parent 1). Assuming a single fixed effect (the mean μ), the resulting mixed model becomes

$$\begin{pmatrix}y_1\\ y_2\\ y_3\end{pmatrix} = \begin{pmatrix}1\\ 1\\ 1\end{pmatrix}\mu + \begin{pmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}a_0\\ a_1\\ a_2\\ a_3\end{pmatrix} + \begin{pmatrix}e_1\\ e_2\\ e_3\end{pmatrix}$$

To complete the mixed-model, we need the relationship matix $\mathbf{A}$. Since all parents are unassumed to be unrelated, so are the two measured sibs, yielded the below $\mathbf{A}$ matrix on the left. If we assume 0 and 1 were crossed to a common (but unmeasured) parent, 2 and 3 are half-sibs, giving the $\mathbf{A}$ matrix on the right.

$$\mathbf{A} = \begin{pmatrix}1 & 0 & 1/2 & 0\\ 0 & 1 & 0 & 1/2\\ 1/2 & 0 & 1 & 0\\ 0 & 1/2 & 0 & 1\end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix}1 & 0 & 1/2 & 0\\ 0 & 1 & 0 & 1/2\\ 1/2 & 0 & 1 & 1/4\\ 0 & 1/2 & 1/4 & 1\end{pmatrix}$$

Thus, we can estimate the breeding value a_0 for the unmeasured parent. Provided unmeasured individuals have measured relatives, $\mathbf{A}$ allows us to estimate their breeding values.

Connections between relatives are often referred to as **links**, and the amount of links (or **connectiveness**) in a relationship matrix is one measure of its precision in estimating breeding values (i.e., the prediction error variance, see Kennedy and Trus 1993). The breeding value of an unmeasured individual with few measured relatives will have a much less precision than that for an individual with a large number of measured relatives. Despite this flexibility in predicting breeding values for individuals with missing records, it is important to stress that simple inclusion of all pedigree relationships appears to be *not* sufficient to yield unbiased BLUP/REML estimates when selection occurs (van der Werf 1990, van der Werf and de Boer 1990, van der Werf and Thompson 1992).

In many cases, it is prudent to modify the simple animal model by considering additional fixed and random effects. For example, genetic and environmental effects can be

separated without a control population by adding fixed effects to account for environmental trends. Likewise, it is often reasonable to include additional random effects, such as maternal/litter effects. Another modification of the basic model occurs when the phenotypic scores (records) of parents are unknown. Example 19.5 showed how we can estimate these as random effects, but likely with bias (when selection occurs). An alternative is to treat these breeding values of unmeasured parents as fixed, rather than random, effects. We deal with each of these modifications in turn.

Models with Additional Random Effects

We have assumed that residuals are uncorrelated and homoscedastic, giving their covariance matrix as $\mathbf{Var}(\mathbf{e}) = \sigma_e^2\,\mathbf{I}$. When additional random effects are present, but ignored by the model, they are subsumed into the residuals, potentially introducing correlations and heteroscedasticity. For example, if sibs share a common maternal environment, this introduces correlations between sibs beyond those accounted for by $\mathbf{A}$. If the model only includes $\mathbf{a}$ and $\mathbf{e}$, this additional covariance appears between the residuals of sibs, and the true covariance matrix for $\mathbf{e}$ is no longer the assumed diagonal, leading to biased estimates of the BLUEs and BLUPs. By suitably incorporating additional random effects, we can develop a new model where the residuals again have the simple covariance structure $\mathbf{Var}(\mathbf{e}) = \sigma_e^2\,\mathbf{I}$.

Suppose there is a second vector $\mathbf{u}$ of m random effects in addition to the vector $\mathbf{a}$ of p breeding values and vector of residuals $\mathbf{e}$. Equation 19.1 becomes

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{W}\mathbf{u} + \mathbf{e} \tag{19.20a}$$

where $\mathbf{X}$, $\mathbf{Z}$, and $\mathbf{W}$ are $n \times q$ (n observations and q fixed effects), $n \times p$ and $n \times m$ incident matrices. The covariance structure assumed is $\mathbf{Var}(\mathbf{a}) = \sigma_A^2\,\mathbf{A}$, $\mathbf{Var}(\mathbf{u}) = \sigma_u^2\,\mathbf{I}$, and $\mathbf{Var}(\mathbf{e}) = \sigma_e^2\,\mathbf{I}$, where $\mathbf{a}$, $\mathbf{u}$ and $\mathbf{e}$ are uncorrelated. This can be compactly written as

$$\mathbf{Var}\begin{pmatrix}\mathbf{a}\\ \mathbf{u}\\ \mathbf{e}\end{pmatrix} = \begin{pmatrix}\sigma_A^2\,\mathbf{A} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \sigma_u^2\,\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \sigma_e^2\,\mathbf{I}\end{pmatrix}, \tag{19.20b}$$

giving the covariance matrix for $\mathbf{y}$ as

$$\mathbf{Var}(\mathbf{y}) = \mathbf{V} = \mathbf{Z}\mathbf{A}\mathbf{Z}^T\sigma_A^2 + \mathbf{W}\mathbf{W}^T\sigma_u^2 + \mathbf{I}\,\sigma_e^2 \tag{19.20c}$$

If we had assumed the model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{e}$, the resulting covariance matrix for the residuals is $\mathbf{Var}(\mathbf{e}) = \mathbf{W}\mathbf{W}^T\sigma_u^2 + \mathbf{I}\,\sigma_e^2$, showing how the additional random effects alters the covariance structure. The resulting mixed-model equations become

$$\begin{pmatrix}\mathbf{X}^T\mathbf{X} & \mathbf{X}^T\mathbf{Z} & \mathbf{X}^T\mathbf{W}\\ \mathbf{Z}^T\mathbf{X} & \mathbf{Z}^T\mathbf{Z} + \lambda_A\mathbf{A}^{-1} & \mathbf{Z}^T\mathbf{W}\\ \mathbf{W}^T\mathbf{X} & \mathbf{W}^T\mathbf{Z} & \mathbf{W}^T\mathbf{W} + \lambda_u\mathbf{I}\end{pmatrix}\begin{pmatrix}\widehat{\boldsymbol{\beta}}\\ \widehat{\mathbf{a}}\\ \widehat{\mathbf{u}}\end{pmatrix} = \begin{pmatrix}\mathbf{X}^T\mathbf{y}\\ \mathbf{Z}^T\mathbf{y}\\ \mathbf{W}^T\mathbf{y}\end{pmatrix} \tag{19.21a}$$

where

$$\lambda_A = \frac{\sigma_e^2}{\sigma_A^2} \quad \text{and} \quad \lambda_u = \frac{\sigma_e^2}{\sigma_u^2} \tag{19.21b}$$

Additional vectors of random effects can be incorporated in a similar manner, see LW Chapters 26 and 27 for details. The mixed-model equations again form the basis for iterative REML estimates of the unknown variance components (σ_A^2, σ_u^2, and σ_e^2), as discussed in detail in LW Chapter 27.

Example 19.6. Often the same trait is measured multiple times in the same individual, for example, the sizes of different litters from a single female. When multiple records are present for at least some individuals, a **repeatability model** should be used (Chapter 13; LW Chapter 26). Repeated measures from the same individual have three components: a genetic value a_k, a common (permanent) environmental value p_k that is the same in each measurement, and the residual environmental value e varying between each measurement, giving the ith measurement of kth individual as $a_k + p_k + e_{ki}$. The **repeatability** of the trait is $r = (\sigma_A^2 + \sigma_p^2)/\sigma_z^2$, giving the variance of the residuals as $\sigma_e^2 = (1-r)\sigma_z^2$ and the variance of permanent environmental effects as $\sigma_p^2 = (r - h^2)\sigma_z^2$, where σ_z^2 is the trait variance after accounting for fixed effects. We remind the reader that "environmental" effects can also include non-additive genetic components, as these are not passed along to offspring.

The repeatability model was used by Estany et al. (1989) to examined the selection response for litter size in rabbits. Their model assumed two groups of fixed effects: the year-season (environmental) effects d_t which had 22 levels over the course this experiment (with d_0 set to zero) and the reproductive state l_i of the doe (three levels: l_1 for primiparious does, l_2 for lactating does, and l_3 for non-primiparious and non-lactating does). Since only two of these factors are estimable, l_1 was assigned a value zero. Their model had three random effects, a_k and p_k for the additive genetic and permanent environmental effect of the kth doe, and the residual e, giving the overall model as

$$y_{tk\ell i} = \mu + l_i + d_t + a_k + p_k + e_{tk\ell i}$$

where $y_{tk\ell i}$ denotes the litter size for the ℓth litter of doe k in reproductive state i in season-year t. In matrix form, the mixed-model becomes

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{Z}\mathbf{p} + \mathbf{e}$$

where $\mathbf{a}$ and $\mathbf{p}$ are $n \times 1$ vectors corresponding to the n does, $\mathbf{Var}(\mathbf{a}) = \sigma_A^2\,\mathbf{A}$, $\mathbf{Var}(\mathbf{p}) = \sigma_p^2\,\mathbf{I}$, and $\mathbf{Var}(\mathbf{e}) = \sigma_e^2\,\mathbf{I}$. $\mathbf{X}$ and $\mathbf{Z}$ are incidence matrices, and the vector of fixed effects is

$$\boldsymbol{\beta} = \begin{pmatrix} \mu \\ l_2 \\ l_3 \\ d_1 \\ d_2 \\ \vdots \\ d_{22} \end{pmatrix}$$

The mixed-model equations are given by Equation 19.21a with

$$\lambda_A = \frac{\sigma_e^2}{\sigma_A^2} = \frac{1-r}{h^2} \quad \text{and} \quad \lambda_u = \frac{\sigma_e^2}{\sigma_p^2} = \frac{1-r}{r-h^2}$$

The careful reader might notice that the two vectors of random effects, the breeding values $\mathbf{a}$ and permanent environment effects $\mathbf{p}$, enter the model as $\mathbf{Za}$ and $\mathbf{Zp}$, respectively. Why then do we simply not combine these, e.g., $\mathbf{Zu}$ where $\mathbf{u} = \mathbf{a} + \mathbf{p}$? The reason we cannot do this (and indeed the reason we can estimate $\mathbf{a}$ and $\mathbf{p}$ separately!) is that $\mathbf{a}$ and $\mathbf{p}$ have *different covariance structures*, $\sigma_A^2\,\mathbf{A}$ versus $\sigma_p^2\,\mathbf{I}$ — while estimates of $\mathbf{a}$ burrow information from relatives, estimates of $\mathbf{p}$ depend only on the focal individual.

Common Family and Material Effects

Random effects are frequently included to account for any common family environmental effect when sibs are present. For example, if two sibs (i and j) share a common environmental value c, then $\sigma(e_i, e_j) = \sigma_c^2$. Hence, there are off-diagonal elements in the covariance matrix of residuals and we no longer have the standard assumption $\mathbf{Var}(\mathbf{e}) = \sigma_e^2 \mathbf{I}$.

Example 19.7. Meyer and Hill (1991), examining the response to selection on adjusted food intake (AFI) in mice (Example 19.2), formulated a model incorporating shared family values c as random effects. In addition, their model accounts for fixed-effects due to generations (d, 22 levels), lines (ln, 3 levels), sex (sx, male/female), and litter size (lt, 7 levels for litters of size 6 to 12 individuals). Under their model, the observed value for AFI from the ith individual from generation t, line ℓ and full-sib family k is given by

$$y_{\ell,tki} = \mu + d_t + ln_\ell + sx_j + lt_m + a_{\ell,tki} + c_{\ell,tk} + e_{\ell,tki}$$

where this individual has sex j and experienced litter size m. In matrix form, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{W}\mathbf{c} + \mathbf{e}$. The vector of fixed effects $\boldsymbol{\beta}$ contains the d, ln, sx, and lt values, while the random effects are the vector of common family effects $\mathbf{c}$, the vector of additive genetic values $\mathbf{a}$, and the vector of residuals $\mathbf{e}$. The incident matrix $\mathbf{W}$ has as its *ik*-th element 1 if individual i is from family k, else the element is zero. Note that Meyer and Hill have two model variables to account for litter effects — a fixed effect lt common to all litters of the same size and a random effect $c_{\ell,tk}$ that varies between families but is the same for all individuals from a particular family. The resulting REML estimate for heritability was 0.15, while the fraction of the total variation (after removal of the fixed effects) accounted for by random family effects was estimated to be $c^2 = \sigma_c^2/\sigma_z^2 = 0.22$, where $\sigma_z^2 = \sigma_A^2 + \sigma_c^2 + \sigma_e^2$ is the trait variance following the removal of fixed effects. Hence, a larger fraction of the resemblance between sibs (the intraclass correlation, $t = h^2/2 + c^2$, see Chapter 21) is due to shared family environments (potentially including maternal effects), rather than due to shared genes. One caveat is that the model assumes no dominance variance. If present, sibs share dominance variance ($\sigma_D^2/4$), and under this model is incorporated into σ_c^2.

Sib correlations (beyond those accounted for by their correlations in breeding value for the focal trait) can arise for (at least) three reasons. The first is dominance, which we will address shortly. The second and third are common family effects, which can potentially have two components: shared family and/or maternal environmental effects and shared maternal effects with a genetic component (i.e., material performance itself has a genetic component, Chapters 15, 22). The distinction between the genetic versus environmental components of maternal effects (both of which may be included in a model) was foreshadowed in Example 19.6, in that they result in different covariance structures.

In Example 19.7, the common environment c was assumed to be uncorrelated across sibships, implying (assuming equal variances across families) that its covariance structure is $\mathbf{Var}(\mathbf{c}) = \sigma_c^2 \mathbf{I}$. Hence, there is no correlation across different sibships and thus no shared information. Now suppose that a mother has several sibships. In this case, we can treat the common family environment in two distinct ways. In the first, each particular litter has a unique common family effect that is uncorrrelated across litters from the same mother. The second is to consider a repeatability model (Example 19.6), wherein the contribution from the mother to a particular litter has two components. Considering litter i from mother k, the common family effect can be written as $c_{ki} + p_k$, where p_k is the permanent environmental effect of this mother (a common effect shared by all her litters), while c_{ki} is the unique

environment shared by all members of her ith litter. In this case, there are two additional random effects added to the model, and we also need to estimate the variances σ_p^2 and σ_c^2, responding to the contributions from p_k and c_{ki}, respectively. We assume each litter has two or more individuals, otherwise c_{ki} and p_k cannot be estimated seperately.

If we assume all of the maternal contribution is environmental, namely that a mother does not pass along any of her performance genetically to daughters, then the covariance structure for these two vectors of random effects would be $\mathbf{Var}(\mathbf{c}) = \sigma_c^2\,\mathbf{I}$ and $\mathbf{Var}(\mathbf{p}) = \sigma_p^2\,\mathbf{I}$. However, maternal performance could also have a genetic component, and thus female relatives have correlated maternal performances. In this case, we would add a third random effect ma (the breeding value of this maternal effect) to the model. This vector of breeding values for maternal performance has covariance matrix $\mathbf{Var}(\mathbf{ma}) = \sigma_A^2(ma)\mathbf{A}$. Distinguishing between maternal versus direct effects requires that there are paternal, as well as maternal, links in the pedigree (Clément et al. 2001, Kruuk 2004; Chapter 22).

A further complication is that if we treat maternal performance as having a heritable component, the focal trait is determined by direct and maternal contributions, and the breeding values for these two can be correlated. This now makes the problem a multiple trait one (LW Chapter 27). We discuss maternal effects in greater detail in Chapter 22, in the context of more general associate effects models.

From an operational standpoint, if maternal effects are suspected, at a minimum a common family effect should be included, and in the form of a repeatability model if the female has multiple litters (each of which results in several sibs). More generally, if there are many links between female relatives (with litters) in the data set, one should seriously consider a genetic maternal effects model. Failure to do so may result in contributions from shared genetic maternal performance being regarded as breeding values for the direct trait, giving a biased picture of the nature of selection response (Milner et al. 2000, Clément et al. 2001, Kruuk 2004).

Treating Certain Breeding Values as Fixed Effects

How should one proceed if the base-population has itself been under selection? If this is known or suspected to be the case, Graser et al. (1987) suggest that the base population breeding values be treated as fixed, rather than random, effects. The motivation for this suggestion is that if parents are selected, they are not a random sample from the base population. Since REML estimates are unbiased by fixed-effects, any bias in the variance of the initial sample is ignored by treating the original parental breeding values as fixed. Simulation studies, however, show that even if initial bias is reduced by treating the parents as fixed, selection on the resulting offspring (or future generations) introduces additional bias (van der Werf 1990). Despite this reservation, we briefly review the approach here as parents whose records are missing are also often treated as fixed, which requires some modifications of the mixed-model equations. Let $\mathbf{a_b}$ be the vector of breeding values for the base population and $\mathbf{a_r}$ breeding values of the remaining individuals that descent from the base population. Following Graser et al. (1987), we can express the dependence of $\mathbf{a_r}$ on the base population breeding values $\mathbf{a_b}$ as follows

$$\begin{pmatrix} \mathbf{a_b} \\ \mathbf{a_r} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{P}_1 & \mathbf{P}_2 \end{pmatrix} \begin{pmatrix} \mathbf{a_b} \\ \mathbf{a_r} \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{s} \end{pmatrix} \tag{19.22}$$

where $\mathbf{s}$ is the vector of segregational residuals (Equation 19.15) and $\mathbf{P}_1$ and $\mathbf{P}_2$ are matrices with values of $1/2$ in the parents' column in each row. Here, $\mathbf{a_r}$ is a random effect because it is a function of a fixed effect ($\mathbf{a_b}$) and a random effect ($\mathbf{s}$). The resulting mixed-model is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{a_b} + \mathbf{Z}_2\mathbf{a_r} + \mathbf{e} \tag{19.23}$$

Graser et al. show that the associated mixed-model equations are

$$\begin{pmatrix} \mathbf{X}^T\mathbf{X} & \mathbf{X}^T\mathbf{Z}_1 & \mathbf{X}^T\mathbf{Z}_2 \\ \mathbf{Z}_1^T\mathbf{X} & \mathbf{Z}_1^T\mathbf{Z}_1 + \lambda\mathbf{Q}^T\mathbf{G}^{-1}\mathbf{Q} & -\lambda\mathbf{Q}^T\mathbf{G}^{-1} \\ \mathbf{Z}_2^T\mathbf{X} & -\lambda\mathbf{G}^{-1}\mathbf{Q} & \mathbf{Z}_2^T\mathbf{Z}_2 + \lambda\mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\mathbf{a}}_\mathbf{b} \\ \widehat{\mathbf{a}}_\mathbf{r} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^T\mathbf{y} \\ \mathbf{Z}_1^T\mathbf{y} \\ \mathbf{Z}_2^T\mathbf{y} \end{pmatrix} \tag{19.24a}$$

where $\lambda = (2\sigma_e^2/\sigma_A^2)$,

$$\mathbf{Q} = (\mathbf{I} - \mathbf{P}_2)^{-1}\mathbf{P}_1 \quad \text{and} \quad \mathbf{G} = (\mathbf{I} - \mathbf{P}_2)^{-1}\mathbf{F}\left[(\mathbf{I} - \mathbf{P}_2)^{-1}\right]^T \tag{19.24b}$$

with the elements of the diagonal matrix $\mathbf{F}$ are given by Equation 19.16.

Dominance

Up to this point, we have been assuming all genetic variation is additive, requiring us to only consider the vector $\mathbf{a}$ of breeding values and its numerator relationship matrix $\mathbf{A}$. When nonadditive genetic variance is present, it creates additional genetic correlations between certain relatives beyond those accounted for by $\mathbf{A}$. The simplest setting is when dominance occurs, which inflates the covariance among (noninbred) full sibs by $\sigma_D^2/4$. As we have seen, sibs can also have their covariance inflated by common family effects, and separating the contribution of dominance from common family environment is nontrivial, requiring specific types of links in the pedigree (see below).

If the goal is simply to reduce the bias in predicted breeding values when dominance is present, an animal model with an additional random factor for common family effects (e.g., Example 19.7) will often be satisfactory. This model simply estimates the common sib variance σ_c^2, which may include contributions from both dominance and shared family environments.

The goal of estimating the dominance variance directly, and thus predicting dominance values, is considerably more difficult. Misztal (1997) found that roughly 20 times more data is required for domiance estimates to match the precision of their additive counterparts. This is because information on dominance only arises from relatives with a nonzero coefficient of fraternity (LW Equation 7.7), which requires that both parents for the two relatives are related. Even in these cases, the coefficient of fraternity is small (e.g., 0.25 for full sibs, 0.0625 for double first-cousins, and much less for more distant relatives). If the only such links in a pedigree are between full sibs, them common family environmental effects and dominance are fully intertwined.

Assuming no common family effects (a *major* assumption), we can attempt to estimate dominance as follows. Letting the vector $\mathbf{d}$ denote the dominance effects, the mixed model becomes

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{Z}\mathbf{d} + \mathbf{e} \tag{19.25a}$$

The overall genetic merit of an individual is estimated by $\widehat{\mathbf{g}} = \widehat{\mathbf{a}} + \widehat{\mathbf{d}}$. Turning to the covariance structure of this model, as before $\mathbf{Var}(\mathbf{a}) = \sigma_A^2\,\mathbf{A}$ and $\mathbf{Var}(\mathbf{e}) = \sigma_E^2\,\mathbf{I}$, while the covariance matrix for dominance effects is $\mathbf{Var}(\mathbf{d}) = \sigma_D^2\,\mathbf{D}$, giving

$$\mathbf{Var}(\mathbf{y}) = \mathbf{V} = \mathbf{Z}\mathbf{A}\mathbf{Z}^T\sigma_A^2 + \mathbf{Z}\mathbf{D}\mathbf{Z}^T\sigma_D^2 + \mathbf{I}\,\sigma_e^2 \tag{19.25b}$$

Equation 19.25b shows the error in the residual variance structure if one (incorrectly) assumes dominance is absent, where we would (incorrectly) use $\sigma_e^2\,\mathbf{I}$ instead of $\mathbf{Z}\mathbf{D}\mathbf{Z}^T\sigma_D^2 + \mathbf{I}\,\sigma_e^2$.

The elements of the **dominance genetic relationship matrix D** are obtained as follows. The covariance between dominance effects for (non-inbred) individuals *i* and *j* is the product of the dominance genetic variance and the coefficient of fraternity, $\sigma_D^2 \Delta_{ij}$. From LW Equation 7.7,

$$\Delta_{ij} = \Theta_{gk}\,\Theta_{hl} + \Theta_{gl}\,\Theta_{hk} \tag{19.26a}$$

where *i*'s parents are indexed by *g* and *h* and *j*'s by *k* and *l*, and (as above), Θ is the coefficient of coancestry. Recalling that the elements of the numerator relationship matrix **A** are $2\Theta_{ij}$, the off-diagonal elements of **D** can be computed from the elements of **A** by

$$D_{ij} = \frac{A_{gk}\,A_{hl} + A_{gl}\,A_{hk}}{4} \tag{19.26b}$$

whereas the diagonal elements are all $D_{ii} = 1$. Note that **D** is expected to be considerably more **sparse** (most off-diagonal elements zero) than **A**, and hence may not contribute information for most individuals. Ovaskainen et al. (2008) note that Equation 19.26a is an approximation, requiring that the four probabiltities (i.e., the Θ_{ij}) determining Δ_{ij} are independent. This is usually not a serious problem unless the pedigree is highly inbred. The resulting mixed-model equations for Equation 19.25a become

$$\begin{pmatrix} \mathbf{X}^T\mathbf{X} & \mathbf{X}^T\mathbf{Z} & \mathbf{X}^T\mathbf{Z} \\ \mathbf{Z}^T\mathbf{X} & \mathbf{Z}^T\mathbf{Z} + \lambda_A\mathbf{A}^{-1} & \mathbf{Z}^T\mathbf{Z} \\ \mathbf{Z}^T\mathbf{X} & \mathbf{Z}^T\mathbf{Z} & \mathbf{Z}^T\mathbf{Z} + \lambda_D\mathbf{D}^{-1} \end{pmatrix} \begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\mathbf{a}} \\ \widehat{\mathbf{d}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^T\mathbf{y} \\ \mathbf{Z}^T\mathbf{y} \\ \mathbf{Z}^T\mathbf{y} \end{pmatrix} \tag{19.27}$$

where $\lambda_A = \sigma_e^2/\sigma_A^2$, and $\lambda_D = \sigma_e^2/\sigma_D^2$. Hoeschele and Van Raden (1991) present a quick method for computing $\mathbf{D}^{-1}$ for a non-inbred population. In theory, epistatic terms can be included in the mixed model equations in a similar fashion, see LW Chapter 26 for details and LW Chapter 27 for modifications of the REML equations to estimate nonadditive variances. In practice, this is almost never done, as the effects are generally small and the precision of estimates is quite poor.

In addition to these concerns, inbreeding (which occurs in all selection experiments) introduces major complications. First, there may be inbreeding depression. In some situations, this can be dealt with by including level of inbreeding f as a covariate, for example by using a model such as

$$y_{ti} = \mu + I f_{ti} + a_{ti} + d_{ti} + e_{ti} \tag{19.28}$$

where f_{ti} is the inbreeding coefficient for the ith individual in generation t and I is the inbreeding depression under complete inbreeding (a fixed factor to be estimated). Since $A_{ii} = (1 + f_i)$, the value for f_{ti} immediately follows from the diagonal element of **A** corresponding to ti, $f_{ti} = A_{ti,ti} - 1$.

Recall (LW Chapter 10) that with only dominance, inbreeding depression is a linear function of the inbreeding f, but nonlinear when epistasis is present. Thus, if there is significant epistasis, Equation 19.28 may not properly correct for inbreeding depression, especially at high values of f (those approaching one). One potential solution is to use a control population with known levels of inbreeding to provide an independent estimation of I. In a LS analysis, inbreeding depression is removed by subtracting off the mean of such a control population. This assumes that the control and selected lines have the same level of inbreeding, when in fact the selected population is typically far more inbred (Chapters 3, 27). If the levels of inbreeding are very similar in both populations, the use of a control can account for nonlinear inbreeding depression. If f is expected to be small or modest (say

$f \leq 0.3$), then f^2 is 0.01 or less and the weighting on any epistatic term is quite small. In such cases, a simple linear model for inbreeding should be sufficient.

A second, and more subtle, complication is that the covariance between inbred relatives with dominance is no longer a function of just σ_A^2 and σ_D^2. As discussed in Chapter 12, these covariances now depend upon four other quadratic components (σ_{DI}^2, σ_{ADI}, $\imath^*$, $\imath^2 - \imath^*$), see Equations 12.14-12.15. While one could formulate a mixed model incorporating all six quadratic components (using the covariances given by Equation 12.14-12.16), the resulting model is extremely complex and numerically very demanding. A start in this direction has been developed by Smith and Mäki-Tanila (1990), who should be consulted for more details. A second approach is to use Equation 19.27, with a **D** matrix that approximates the elements under inbreeding (Smith and Mäki-Tanila 1990). One could also combine the use of a modified **D** with a covariate for inbreeding depression (Equation 19.28), but this is still a largely ad-hoc approach to a complex problem.

Some guidance is potentially offered by simulations by de Boer and van Arendonk (1992), who examined the consequences of ignoring a cofactor for inbreeding depression and the full covariance structure under inbreeding. When a standard dominance model not accounting for inbreeding was used, estimates of both additive and dominance effects were biased. However, when a simple cofactor for inbreeding depression was included, but the full covariance structure under inbreeding ignored, effects estimates were unbiased, at least up to the level of inbreeding used in the simulations ($f = 0.35$). Thus, with low to modest levels of inbreeding, simple inclusion of a fixed effect for inbreeding depression (Equation 19.25) and a random effect for common family effects (Example 19.7) appears to be a relatively robust way to handle dominance.

The situation with epistasis is even more complex than with simple dominance. The good news is that the weighting of the nonadditive variance component terms for the covariances between distant relatives is very small (e.g., LW Equation 7.12), so that even if nonadditive components are significant, their *actual* contribution to the covariance of most relatives, especially those separated by more than one generation, are very minor.

BAYESIAN MIXED-MODEL ANALYSIS

As mentioned throughout this chapter, a standard mixed-model analysis does not fully account for the uncertainty introduced into estimates of the breeding values by using estimates of the model variances components (instead of their true values). While there are large-sample approximations for the sample variance of a REML variance estimator, it is never fully clear what constitutes "large". Further, quantities of interest, such as the heritability, are often functions of the estimated quantities. The sample variance and sampling distribution of such functions are very complicated, and typically obtained by simulation. **Bayesian approaches** offer solutions to both these issues. While Bayesian statistics (as opposed to more standard, or **frequentist**, statistics) is often touted for its ability to incorporate prior information, we view their key utility as providing a complete description of the uncertainty of an estimate.

Frequentists assume the true value of a parameter is (typically) a constant, and the samples are variable. Statistics (such as confidence intervals) are computed by conceptually drawing an infinite number of samples, in which case (for example), the true value is in 95% of all such constructed confidence intervals. In contrast, a Bayesian assumes that the sample is fixed, while the parameter is random. Their interest is in how that data changes the probability distribution for possible locations of the parameter. Thus, the term "Bayesian mixed models" is (formally) inappropriate, as *all* terms in a Bayesian analysis are assumed to be random and hence never "mixed". However, we use this term to emphasize that much

of the basic foundations of a MM analysis of selection experiments remain unchanged (such as the model formulation). What does change is how we analyze such data.

Appendix 2 introduces some of the basic ideas in a Bayesian analysis (beyond our short introduction here). Computational issues are extremely important, and covered in Appendix 3. Indeed, the recent explosion in the application of Bayesian approaches largely follows from relatively new computational approaches (such as MCMC methods) that allow very complex distributions to be handled through straight-forward (but computationally intensive) procedures.

Introduction to Bayesian Statistics

While very deep (and very subtle) differences in philosophy separate hard-core Bayesians from hard-core frequentists (Efron 1986, Glymour 1981), our treatment here of Bayesian methods is motivated simply by their use as a powerful statistical tool. Their introduction into quantitative genetics can be largely credited to the influential paper of Gianola and Fernando (1986), which reviews Bayesian applications to animal breeding. Blasco (2001) provides a very nice overview of Bayesian vs. frequentist approaches in quantitative genetics and is highly recommended, while a very complete treatment of applications to quantitative genetics is provided by Sorensen and Gianola (2002).

The foundation of Bayesian Statistics is **Bayes' theorem** (Appendix 2; LW Chapter 13), which provides the relationship between $\Pr(x \mid y)$ and $\Pr(y \mid x)$, namely the flipped conditional probabilities. The continuous, vector-valued version of this theorem is

$$p(\boldsymbol{\Theta} \mid \mathbf{y}) = p(\mathbf{y} \mid \boldsymbol{\Theta}) \, \frac{p(\boldsymbol{\Theta})}{p(\mathbf{y})} = \frac{p(\mathbf{y} \mid \boldsymbol{\Theta}) \, p(\boldsymbol{\Theta})}{\int p(\mathbf{y}, \boldsymbol{\Theta}) \, d\boldsymbol{\Theta}} \tag{19.29a}$$

where $\boldsymbol{\Theta}^T = (\theta^{(1)}, \theta^{(2)}, \cdots, \theta^{(k)})$ is a vector of k random variables. Here $p(\boldsymbol{\Theta})$ is our prior belief (**prior** for short) about the distribution of the unknown values $\boldsymbol{\Theta}$, while $p(\mathbf{y} \mid \boldsymbol{\Theta})$ is just a standard likelihood function for the probability density of the observed vector of data $\mathbf{y}$ given that the unknown parameters have specified value $\boldsymbol{\Theta}$ (LW Appendix 4). The product of these two, normalized by $p(\mathbf{y})$ to form a proper probability distribution, is our posterior belief (**posterior**) for the distribution of the unknown parameters given both the data $\mathbf{y}$ and the prior information $p(\boldsymbol{\Theta})$. Since $p(\mathbf{y})$, the probablity of the data vector $\mathbf{y}$ is a constant, it is typically ignored, and the posterior is often written as

$$p(\boldsymbol{\Theta} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \boldsymbol{\Theta}) \, p(\boldsymbol{\Theta}) \tag{19.29b}$$

In words, the posterior is the product of the likelihood and the prior times a normalization constant to return a proper probability distribution (i.e., intergration to one).

Example 19.8. As an example of a Bayesian analysis, consider the simple case of n observations from a normal with unknown mean μ but known variance σ^2. The details for this analysis (and more realistic cases, such as both the mean and variance unknown) are developed in Appendix 2. Assuming the data $\mathbf{y} = (y_1, \cdots, y_n)^T$ are n independent draws from this distribution, the resulting likelihood function,which corresponds to $p(\mathbf{y} \mid \mu)$, is

$$p(\mathbf{y} \mid \mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\sum_{i=1}^{n} \frac{(y_i - \mu)^2}{2\sigma^2} \right)$$

Suppose we assume a Gaussian prior for the location of the mean, $\mu \sim \mathrm{N}(\mu_0, \sigma_0^2)$, so that

$$p(\mu) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{(\mu - \mu_0)^2}{2\sigma_0^2} \right)$$

The mean μ_0 and variance σ_0^2 are referred to as the prior **hyperparameters**, where μ_0 specifies a prior location for the mean and σ_0^2 specifies the uncertainty in this prior location. The larger this variance, the greater our uncertainty. In the limit (as $\sigma_o^2 \rightarrow \infty$) this corresponds to a $p(\mu) = c$, a constant, which is **uniform** or **flat** prior, where all values of μ are assumed to be equally likely. (Keyner 1921 called this is *principle of indifference* — all possible events are equally probable.) A little algebra (Appendix 2) yields

$$p(\mu \,|\, \mathbf{y}) \propto p(\mathbf{y} \,|\, \mu)\, p(\mu) = \exp\left(-\frac{(\mu - \mu_*)^2}{2\sigma_*^2}\right)$$

where the expressions for μ_* and σ_*^2 are given in Appendix 2 (Equation A2.22b). Thus, the posterior density function for μ is a normal with mean μ_* and variance σ_*^2,

$$\mu \,|\, (\mathbf{y}, \sigma^2) \sim \mathrm{N}\left(\mu_*, \sigma_*^2\right)$$

Notice that our Guassian prior **conjugated** with the likelihood function – the product of the prior and likelihood returned a distribution in the same family as the prior. The use of such **conjugate priors** (for a given likelihood) is a key tool in Bayesian analysis and is explored in detail in Appendix 2. For example, with normally-distributed data and an unknown variance, using a **scaled inverse chi-square** (χ^2) prior for the variance also conjugates the likelihood, with the posterior distribution for the variance following a scaled inverse χ^2 distribution (Appendix 2). Observe that while μ is a fixed (but unknown) value in a standard (frequentist) analysis, in a Bayesian analysis it is a *random variable*. There are no fixed effects in a Bayesian analysis, as everything is treated as a random variable (as opposed to being unknown, but fixed, values).

A Bayesian analysis returns a *distribution*, rather than a point estimate. A number of summary statistics can be reported for the posterior such as its mean, median (50% value) and mode (most frequent value, the MLE in a likelihood analysis). More generally, when nice analytic expressions (like ours above) are not available, one can simply plot the distribution via a histogram from values generated through MCMC methods (as discussed in Appendix 3).

What is the relative importance of the prior information $p(\mu)$ versus the actual data $\mathbf{y}$? Equation A2.22b gives the mean of the posterior distribution as

$$\mu_* = \mu_0 \frac{\sigma_*^2}{\sigma_0^2} + \overline{y} \frac{\sigma_*^2}{\sigma^2/n}$$

With a very diffuse prior on μ (i.e., $\sigma_0^2 \gg \sigma^2$), Equation A2.22b shows that $\sigma_*^2 \rightarrow \sigma^2/n$ and $\mu_* \rightarrow \overline{y}$, so that with very weak prior information the mean of the posterior distribution is close to the sample mean. Conversely, as we collect enough data (i.e., large n), $\sigma_*^2 \rightarrow \sigma^2/n$ and again $\mu_* \rightarrow \overline{y}$. Even with a very strong prior belief about the location of the mean, as our sample size becomes sufficiently large, the mean of the posterior approaches the sample mean.

The dependence of the posterior on the prior (which can easily be assessed by trying different priors) provides an indication of how much information on the unknown variable(s) is contained in the data. If the posterior is highly dependent on the prior, the data likely has little signal, while if the posterior is largely unaffected by the shape of the assumed prior, the data are highly informative. Such explorations of the effects of a prior under a careful Bayesian analysis offers some protection from incorrect conclusions based on weak likelihoods (the likelihood surface is rather flat). It is not uncommon in the same

analysis to find that the marginal posterior for some variables is very robust to choice of priors (and hence the data contains a strong signal generating a highly-peaked likelihood surface), while for other variables the marginal is highly dependent on the prior (and hence very little information on these variables is contained in the data).

Example 19.9. Sorensen et al. (1994) examined the effects of different priors for the variance components in a simulation study. The data were analyzed in two sets: the entire data set (ALL) and a partial subset (PART). The simulated heritability was 0.5. Both uniform and scaled inverse χ^2 priors for the additive and residual variance were used. The uniform spreads belief evenly over all possible values within a defined range, while the inverse χ^2 places more weight on specific values. The mean and variance of the marginal posterior distribution for h^2 in these four cases were as follows:

Data Set	Prior	$E[\,h^2 \,|\, y]$	$\sigma^2[\,h^2 \,|\, y]$
PART	Uniform	0.737	2.26×10^{-2}
PART	Inverse χ^2	0.501	0.29×10^{-2}
ALL	Uniform	0.550	1.63×10^{-2}
ALL	Inverse χ^2	0.529	0.24×10^{-2}

Note the disparity of the estimates under the two priors in the partial data case, and their agreement in the full data case. In the partial data case, the effect of the prior had a strong influence, indicating a weak signal (likelihood) for h^2 in this particular data set. With the full data set, the signal greatly increases, mitigating the effects of the prior. As expected, for both priors that the posterior variance decreased under the larger sample size (PART versus ALL). Also note that the posterior variance was smaller under the assumed inverse χ^2 priors. Thus, the choice of a prior not only influences the mean of the estimate, but also its variance as well. In this case, while the different priors gave the same mean heritability in the full data set, their variances differed by an order of magnitude.

This is an example of a **sensitivity analysis** using different priors to probe the stability of the posterior. With complex posteriors, one can observe broad stability for many of the variables (insensitivity to changes in the priors), but extreme dependence in the others. The use of different priors provides one means to explore the amount of signal along the different directions (variables) of the likelihood function.

Often, only a subset of the unknown variables are of concern the us, the rest being **nuisance variables** that we wish to remove (or at least ignore). Write the vector of unknown parameters as $\boldsymbol{\Theta}^T = (\boldsymbol{\Theta}_1^T, \boldsymbol{\Theta}_{nu}^T)$, where $\boldsymbol{\Theta}_{nu}$ is the column vector of nuisance variables. Integrating the full posterior over $\boldsymbol{\Theta}_{nu}$ gives the **marginal posterior distribution** for the variable of interest as

$$\begin{aligned} p(\boldsymbol{\Theta}_1 \,|\, \mathbf{y}) &= \int p(\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_{nu} \,|\, \mathbf{y})\, d\boldsymbol{\Theta}_{nu} \\ &= \int p(\boldsymbol{\Theta}_1 \,|\, \boldsymbol{\Theta}_{nu}, \mathbf{y})\, p(\boldsymbol{\Theta}_{nu} \,|\, \mathbf{y}) d\boldsymbol{\Theta}_{nu} \\ &= E_{\boldsymbol{\Theta}_{nu}} [\, p(\boldsymbol{\Theta}_1 \,|\, \boldsymbol{\Theta}_{nu}, \mathbf{y}) \,] \end{aligned} \tag{19.30}$$

The marginal posterior is very powerful, incorporating all of the uncertainly in the nusiance variables in terms of how this influences our level of certainty for the variables of

interest. This marginal probability calculation illustrates both the strength, and weakness, of a Bayesian analysis before the advent of MCMC approaches. The strength is that obtaining such a marginal is very powerful for inference, the weakness is that the integration to obtain this marginal can be horrendous (at best!). Fortunately, MCMC techniques allow one to easily simulate draw from this distribution (Appendix 3)

Example 19.10. In the context of analyzing a selection response experiment, the vector of breeding values $\mathbf{a}$ is of interest, while the q fixed effects ($\boldsymbol{\beta}$) and variances (σ_A^2, σ_e^2) are often regarded as nuisance parameters. In this case, Equation 19.30 gives the marginal distribution of the breeding values $\mathbf{a}$ given the data as

$$p(\mathbf{a} \mid \mathbf{y}) = \int p(\mathbf{a}, \boldsymbol{\beta}, \sigma_A^2, \sigma_e^2 \mid \mathbf{y})\, d\boldsymbol{\beta}\; d\sigma_A^2\; d\sigma_e^2$$

The integration is over the $q + 2$ dimensional space given by the q elements in $\boldsymbol{\beta}$ and the two variances. This conditioning removes any dependencies of estimates of the response on estimates of the variance components and fixed effects. Uncertainties introduced by estimating these nuisance parameters are automatically accommodated when considering the marginal distribution. While this multidimensional integral is complex, the Gibbs sampler (below) can be use to obtain draws from this marginal distribution.

With the marginal density $p(\mathbf{a} \mid \mathbf{y})$ in hand, one can obtain estimates of the response to selection $\mathbf{K}^T\mathbf{a}$ that are *independent* of the assumed (or estimated) additive variance σ_A^2. The error due to estimation of the additive variance from the data is directly incorporated when the marginal is computed, as we integrate over possible values of σ_A^2 and their support given the data. This independence of the estimate of response from the estimate of additive variance and the subsequent incorporation of the error in estimating σ_A^2 in the estimate of the response are two very compelling reasons for a Bayesian analysis of response.

This example hints at a key feature noted by Gianola and Fernando (1986). If the data on which selection was based are included in the analysis, then (by integrating over all nuisance parameters) the Bayesian approach accounts for any potential bias by selection (*provided* that the model assumptions, such as multivariate normality, hold).

Computing Posteriors and Marginals: MCMC and The Gibbs Sampler

Historically, the widespread implementation of Bayesian approaches was limited by the difficulty in obtaining marginal posterior distributions, which typically requires the integration of complex, high dimensional functions (e.g., Equations 19.29 and 19.30). **Markov Chain Monte Carlo** (**MCMC**) approaches (Appendix 3) provide a solution, offering a straightforward (although computationally demanding) procedure for generating random draws from very complex distributions.

Simulating random vectors directly drawn from some complex **target distribution** can be a very difficult task. The idea behind MCMC approaches is to successively draw samples from far simpler distributions in such a way that the distribution of the samples converges to the target distribution. MCMC approaches are so-named because one uses the previous sample value to randomly generate the next sample value, generating a **Markov chain** (Appendix 3). While there are a wide range of MCMC methods, two of the most commonly encountered in the quantitative-genetics literature are the **Metropolis-Hastings algorithm** (Metropolis and Ulam 1949, Metropolis et al. 1953, Hastings 1970), and the **Gibbs**

sampler (Geman and Geman 1984). Under Metropolis-Hastings, one simulates draws from a complex target distribution by first drawing a random variable from a specified (and simpler) distribution and then using a probability-based decision rule to decide whether to keep that realization or reject it (details in Appendix 3). The strength of Metropolis-Hastings is that is can be applied to a very wide range of problems, such as priors that do not conjugate with the likelihood (and hence do not have a simple form). Its weakness is that candidate values can end up being rejected with very high probability, making the sampler very inefficient (requiring very long runs to produce a reasonably-size trimmed sequence with low correlation between elements), especially when one is dealing with vectors of random variables. The Gibbs sampler is a special case of Metropolis-Hastings sampling wherein the random value is always accepted. The key to this sampler is that one only considers univariate conditional distributions — the distribution when all of the random variables but one are assigned fixed values. Typically, one uses conjugate priors to form a Gibbs sampler (see below). More generally, one can use also a **block implementation** of the sampler, sampling from multivariate distributions as well. For example, in a regression model one can sample the variance conditional on the elements of $\boldsymbol{\beta}$, and then sample all the elements of $\boldsymbol{\beta}$ at once, conditional on the variance.

To introduce the Gibbs sampler, consider a bivariate random vector (x, y), and suppose we wish to compute one or both marginals, $p(x)$ and $p(y)$. The idea behind the sampler is that it is far easier to consider a sequence of conditional distributions, $p(x \mid y)$ and $p(y \mid x)$, than it is to obtain the marginal by integration of the joint density $p(x, y)$, e.g., $p(x) = \int p(x, y)dy$. The sampler starts with some initial value y_0 for y and obtains x_0 by generating a random variable from the conditional distribution $p(x \mid y = y_0)$. The sampler then uses x_0 to generate a new value of y_1, drawing from the conditional distribution based on the value x_0, $p(y \mid x = x_0)$. The sampler proceeds as follows

$$x_i \sim p(x \mid y = y_{i-1}) \tag{19.31a}$$

$$y_i \sim p(y \mid x = x_i) \tag{19.31b}$$

Repeating this process k times, generates a **Gibbs sequence** of length k, where a subset of points (x_j, y_j) for $1 \leq j \leq m < k$ are taken as our simulated draws from the full joint distribution. To obtain the desired total of m sample points, one samples the chain (i) after a sufficient **burn-in** to remove the effects of the initial starting values and (ii) at set time points (say every n samples) following the burn-in (**trimming** the sequence). For example, Wang et al (1994b) in their Gibbs sampler for an animal model generated a total of 1,205,000 sample vectors. The first 5,000 were discarded (corresponding to the burn-in), and then every tenth subsequent iteration was saved (to reduce correlations between sample vectors) to yield a total sample of 120,000 vectors. The burn-in period, and sampling interval following the burn-in can be delicate, and careful analysis of the resulting sequence using conversion diagnostic tools is critical (Appendix 3).

When more than two variables are involved, the sampler is extended in the obvious fashion. For example, if there are four variables, (w, x, y, z), the sampler becomes

$$\begin{aligned}
w_i &\sim p(w \mid x = x_{i-1}, y = y_{i-1}, z = z_{i-1}) \\
x_i &\sim p(x \mid w = w_i, y = y_{i-1}, z = z_{i-1}) \\
y_i &\sim p(y \mid w = w_i, x = x_i, z = z_{i-1}) \\
z_i &\sim p(z \mid w = w_i, x = x_i, y = y_i)
\end{aligned}$$

Any feature of interest for the marginals can be computed from the m realizations of the Gibbs sequence. For example, the expectation of any function f of the random variable

x is approximated by

$$E[f(x)]_m = \frac{1}{m}\sum_{i=1}^{m} f(x_i) \tag{19.32a}$$

which is simply the average of the function evaluated over the points in the sampler. This is the **Monte-Carlo** (MC) **estimate** of $f(x)$, as $E[f(x)]_m \to E[f(x)]$ as $m \to \infty$. Likewise, the MC estimate for any function of n variables $(\theta^{(1)}, \cdots, \theta^{(n)})$ is given by

$$E[f(\theta^{(1)}, \cdots, \theta^{(n)})]_m = \frac{1}{m}\sum_{i=1}^{m} f(\theta_i^{(1)}, \cdots, \theta_i^{(n)}) \tag{19.32b}$$

Example 19.11. Suppose we are interested in the distribution of breeding values in a particular generation (measured by four individuals in the analysis), as well as in the base population heritability. A Gibbs sampler has been implemented and the realizations at three different iterations (say 100, 200, and 300) after a sufficiently burn-in period are as follows:

factor	Sample 100	Sample 200	Sample 300
$a(1)$	1.5	1.8	2.2
$a(2)$	2.1	3.4	1.4
$a(3)$	3.1	2.9	4.4
$a(4)$	3.3	4.3	3.6
σ_A^2	0.55	0.64	0.46
σ_e^2	1.10	0.98	1.20

Here $a(1)$ through $a(4)$ correspond to the values (realizations of the posterior distribution) of the four breeding values for our generation of interest in that particular iteration of the sampler, and σ_A^2 and σ_e^2 are similarly the realizations for the variances in that iteration. The values for the mean and variance for the breeding value of these four individuals

$$\overline{a}_i = \frac{1}{4}\sum_{j=1}^{4} a(j)_i \quad \text{and} \quad \text{Var}(a)_i = \frac{1}{4-1}\sum_{j=1}^{4} \left[\, a(j)_i - \overline{a}_i \,\right]^2,$$

and the base-population heritability

$$h_i^2 = \sigma_{A,i}^2/(\sigma_{A,i}^2 + \sigma_{e,i}^2)$$

for these three realizations are

	Sample 100	Sample 200	Sample 300
$\overline{a}$	2.5	3.1	2.9
Var(a)	0.72	1.09	1.83
h^2	0.33	0.40	0.28

Thus, the sampler has returned three values for each of the quantities of interest. Of course, a full sampler consist of thousands to hundreds of thousands of such realization, allowing us to empirically generate the full distribution of any for these functions. For example, the mean of the marginal posterior for the mean breeding value over these four individuals is simply the mean of $\overline{a}$ over the entire sample from the Gibbs sequence.

Bayesian Analysis of the Animal Model

The use of Bayesian approaches for the analysis of selection experiments was first suggested by Sorensen and Johansson (1992). Starting with the standard animal model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{e}, \quad \text{where} \quad \mathbf{y} \sim \text{MVN}(\mathbf{X}\boldsymbol{\beta}, \sigma_A^2\mathbf{Z}\mathbf{A}\mathbf{Z}^T + \sigma_e^2\mathbf{I}),$$

Wang et al. (1993, 1994a,b), Sorensen et al. (1994), and Jensen et al. (1994) developed Gibbs-samplers for this model and its extensions. As before, the conditional distribution of the data given the vectors of fixed effects $\boldsymbol{\beta}$, breeding values $\mathbf{a}$, and the environmental variance σ_e^2 is multivariate normal,

$$\mathbf{y} \mid \boldsymbol{\beta}, \mathbf{a}, \sigma_e^2 \sim \text{MVN}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a}, \mathbf{I}\sigma_e^2) \tag{19.33a}$$

The infinitesimal model is assumed, with the distribution of breeding values, given the relationship matrix $\mathbf{A}$ and additive genetic variance σ_A^2, being multivariate normal,

$$\mathbf{a} \mid \mathbf{A}, \sigma_A^2 \sim \text{MVN}(\mathbf{0}, \mathbf{A}\sigma_A^2) \tag{19.33b}$$

Sorensen et al. (1994) assumed a uniform prior for $\boldsymbol{\beta}$, a normal prior for $\mathbf{a}$, and both uniform and inverse χ^2 priors for the variances. For example, the joint prior for $\mathbf{a}$ and σ_A^2 is the product of $p(\mathbf{a} \mid \sigma_A^2, \mathbf{A}) \cdot p(\sigma_A^2)$, where the first distribution is a multivariate normal (Equation 19.33b) and the second either a uniform or an inverse χ^2 (Appendix 2). These choices are conjugate priors given the multivariate normal, resulting in analytic expressions for the $p + q + 2$ univariate conditional distributions for each factor in the model (p breeding values a_k, q fixed effects β_j, and the variances σ_A^2 and σ_e^2). Using these univariate conditionals, the Gibbs sampler can be constructed. The outline for the sampler is as follows:

1. Set initial values for $\mathbf{a}$, $\boldsymbol{\beta}$, σ_A^2, σ_e^2.

2. Using the current values of $\mathbf{a}$, $\boldsymbol{\beta}$, σ_A^2, σ_e^2 and the conditional distributions (see Sorensen et al. for exact expressions) :

(i) Update the fixed effects by sequentially drawing (for $j = 1, \cdots, q$) from the conditionals (which are univariate normals)

$$\beta_{j,i} \sim p(\beta_j \mid \beta_{1,i}, \cdot\cdot, \beta_{j-1,i}, \beta_{j+1,i-1}, \cdot\cdot, \beta_{q,i-1}, \mathbf{a}_{i-1}, \sigma_{A,i-1}^2, \sigma_{e,i-1}^2)$$

where $\beta_{j,i}$ is the value of β_j during the ith iteration of the sample. The current values of these parameters define the mean and variance for the normal, from which a random value is drawn. For factor j, we take the values for $\mathbf{a}$ and the variances from the last iteration $(i-1)$, the values of β_1 to β_{j-1} from the current iteration (i), and the values of β_{j+1} to β_q from the last iteration $(i-1)$. These values are inserted to give the parameters (here, the conditional mean and variance) for the univariate normal that corresponds to the conditional distribution for β_j and a random variable is drawn from this distribution to give $\beta_{j,i}$.

(ii) Update the breeding values by sequentially drawing (for $i = 1, \cdots, p$) from the conditionals (again univariate normals),

$$a_{j,i} \sim p(a_j \mid \boldsymbol{\beta}_i, a_{1,i}, \cdot\cdot, a_{j-1,i}, a_{j+1,i-1}, \cdot\cdot, a_{p,i-1}, \sigma_{A,i-1}^2, \sigma_{e,i-1}^2)$$

(iii) Update the additive variance, drawing from the conditional (a scaled inverse χ^2 distribution)

$$\sigma_{A,i}^2 \sim p(\sigma_A^2 \mid \boldsymbol{\beta}_i, \mathbf{a}_i, \sigma_{e,i-1}^2)$$

(iv) Update the error variance, drawing from the conditional (again, a scaled inverse χ^2 distribution)

$$\sigma^2_{e,i} \sim p(\sigma^2_e \,|\, \boldsymbol{\beta}_i, \mathbf{a}_i, \sigma^2_{A,i})$$

3. Using the updated values, repeat **(2)** until k samples are obtained, from which m are extracted (following the burn-in and trimming) for the Gibbs sampler chain.

This Bayesian analysis makes most of the standard animal-model assumptions, in particular that the infinitesimal model, and hence multivariate normality, holds. As with a MM analysis, a Bayesian approach is potentially biased by changes in allele frequencies. A Bayesian analysis has all the advantages of a MM analysis (over a LS analysis) and, in addition, the posterior marginals correctly give the distribution of any parameter of interest independent of the values assumed for other parameters. Any uncertainly introduced by estimating these additional parameters is fully captured by the marginal posterior. The Bayesian approach gives the correct distribution (assuming the model assumptions hold and the prior is reasonable) for the estimated response independent of the additive genetic variance. By contrast, a MM analysis is highly dependent on the assumed (or estimated) additive variance, and the standard error of a REML/BLUP estimate for the response (Equation 19.8c) does not account for the variance introduced by REML estimation of σ^2_A. One standard package for Bayesian analysis is Winbugs, and Damgaard (2007) shows how to apply this software to animal models.

Application: Estimating Response in Pig Litter Size Components

Blasco et al. (1998) used the method of Sorensen et al. (1994) to estimate the response to selection on ovulation rate and prenatal survival in French Large White pigs. Three lines were followed, one selecting on each trait and a control line. The relevant selection and control lines were jointly analyzed to estimate response. Ovulation rate was examined using the standard animal model,

$$\mathbf{y} \,|\boldsymbol{\beta}, \mathbf{a}, \sigma^2_e \sim N(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a}, \mathbf{I}\sigma^2_e), \quad \mathbf{a} \,|\, \mathbf{A}, \sigma^2_A \sim N(\mathbf{0}, \mathbf{A}\,\sigma^2_A)$$

Prenatal survival (as a function of the mother) was examined using the repeatability model, where c is the permanent environmental effect of a mother over multiple litters,

$$\mathbf{y} \,|\boldsymbol{\beta}, \mathbf{a}, \mathbf{c}, \sigma^2_e \sim N(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{W}\mathbf{c}, \mathbf{I}\,\sigma^2_e)$$

$$\mathbf{a} \,|\, \mathbf{A}, \sigma^2_A \sim N(\mathbf{0}, \mathbf{A}\,\sigma^2_A), \quad \mathbf{c} \,|\, \sigma^2_c \sim N(\mathbf{0}, \mathbf{I}\,\sigma^2_c)$$

Among the fixed effects in $\boldsymbol{\beta}$ are terms for the **parity** of the mother (1st parity = 1st litter, 2nd parity = second litter, and so on). The marginal posterior distribution for breeding values (and hence for the response via $\mathbf{K}^T\mathbf{a}$) was obtained by using the Gibbs sampler approach of Sorensen et al. (1994). For each trait, two independent chains of length 100,000 were computed, with the first 10,000 samples discarded (to remove burn-in effects) and sampling at every 30 iterations thereafter, generating a trimmed sampler of length 3,000. The authors obtained these burn-in and resampling values after several initial runs and using the diagnostics suggested by Raftery and Lewis (1992) for level of precision and Geyer (1992) for autocorrelation between samples. A uniform prior was taken for the fixed effects, while different priors used for the variances (discussed below).

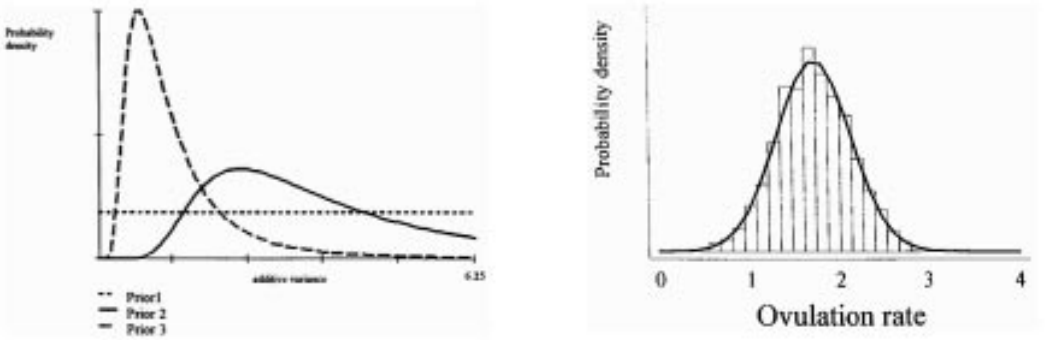


Figure 19.2. Analysis of ovulation rate at puberty in French Large White pigs. **Left:** Assumed priors for σ_A^2 (see text for details). **Right:** The Bayesian estimate of response is given by the marginal posterior density for the mean breeding value in ovulation rate in the last generation of selection, presented as a histogram. This distribution is approximately normal (solid curve). After Blasco et al. (1998).

Table 19.1 Estimated response to selection for ovulation rate at puberty and prenatal survival in French Large White pigs. Bayesian analysis with three different priors (Figure 19.2 for ovulation, Figure 19.3 for prenatal survival) were used to obtain Monte-Carlo estimates of the mean response and their associated standard deviations (the later incorporating the additional error from estimating the additive variance and other parameters). For comparison, least squares (LS) estimates ($\overline{z}_{i+1} - \overline{z}_i$) and mixed-model (REML/BLUP) estimates are also included. After Blasco et al. (1998).

Response in Ovulation Rate at Puberty

Method	Gen 1	Gen 2	Gen 3	Gen 4
Bayesian, Prior 1	0.30 ± 0.31	0.51 ± 0.35	1.03 ± 0.39	1.58 ± 0.43
Bayesian, Prior 2	0.31 ± 0.30	0.51 ± 0.34	1.05 ± 0.38	1.55 ± 0.42
Bayesian, Prior 3	0.31 ± 0.31	0.51 ± 0.35	1.01 ± 0.35	1.53 ± 0.38
LS	-0.09	0.35	1.98	1.87
REML/BLUP	0.27	0.45	1.00	1.54

Response in Prenatal Survival

Method	Gen 1	Gen 2	Gen 3	Gen 4
Bayesian, Prior 1	-0.53 ± 1.44	1.23 ± 1.61	2.83 ± 1.94	2.89 ± 2.12
Bayesian, Prior 2	-0.64 ± 1.70	1.50 ± 1.87	3.46 ± 2.05	3.49 ± 2.30
Bayesian, Prior 3	-0.46 ± 1.45	1.22 ± 1.60	2.84 ± 1.82	2.90 ± 2.01
LS	-5.71	2.11	4.13	-2.82
REML/BLUP	-0.54	1.49	3.26	3.42

Consider the results for ovulation rate at puberty first. Figure 19.2 shows the three priors assumed for the additive variance in this trait. The phenotypic variance of this trait is 6.25, setting an upper limit on σ_A^2. Prior one is a uniform distribution that weights all values in the parameter space equally. Priors two and three (scaled inverted χ^2 distributions) reflect additional information. Published heritabilities for this trait in pigs and rabbits range from 0.1 to 0.6, and prior two assumes a broad distribution around the approximate medium value ($\sigma_A^2 = 0.4 \cdot 6.25 = 2.5$). A study specifically in French Large Whites gave an estimate of $h^2 = 0.11 \pm 0.02$ and the tight distribution around this value is reflected in prior three. Using the approach for Sorensen et al. (1994), Blasco et al. obtain Monte-Carlo estimates of the (base population) heritability under these three priors of $h^2 = 0.39 \pm 0.07$, 0.39 ± 0.06, and 0.32 ± 0.06. Table 19.1 shows the estimated response during each of the

four generations of selection, comparing these with the least squares (differences between generation means) and mixed model (REML/BLUP) estimates. Note that the three different priors give very consistent estimates of response, implying that the data contain sufficient information to overpower most of the signal coming from the assumed prior. The Bayesian and MM analysis give very similar results, while the LS results give a slightly different estimates of response.

While the results for ovulation rate are consistent across the three priors and with the MM analysis, the results are more problematic for prenatal survival (Table 19.1). Figure 19.3 shows the assumed different priors. As with ovulation rate, prior one is uninformative, weighting all potential additive variances equally. Prior two (as with prior two for ovulation rate) assumes a broad distribution around the mean heritability ($h^2 \simeq 0.2$) for a number of studies, while prior three uses the estimate of $h^2 = 0.03 \pm 0.03$ found using French Large Whites. The three priors give Monte Carlo estimates of heritability (and its standard deviation) of $h^2 = 0.12 \pm 0.06$, 0.16 ± 0.04, and 0.11 ± 0.04. Likewise, these priors give Monte Carlo estimates of the repeatability of 0.23 ± 0.05, 0.23 ± 0.04, and 0.19 ± 0.04. As Table 19.1 shows, the standard deviations for the Monte Carlo estimates of mean response are very large, but that the three priors and the MM analysis give consistent results, while the LS results are quite different. Clearly, the information on prenatal survival in the experiment is not sufficiently large enough, as the posterior is significantly influenced by the prior.

Additional early examples of Bayesian analysis of selection experiments are Rodriguez et al. (1996) who looked at lean growth in pigs and Su et al. (1997) who examined body weight in chickens.

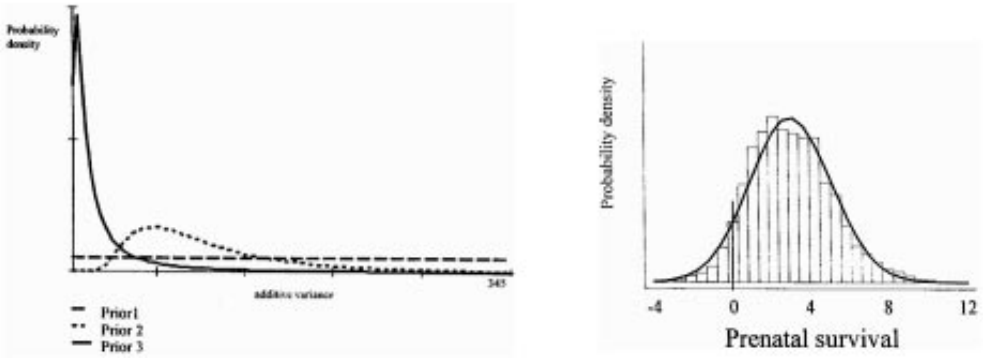


Figure 19.3. Analysis of French Large White pig prenatal survival. **Left:** Priors for the additive variance (see text for details). **Right:** Posterior distribution of mean breeding values (at generation four) in prenatal survival, which deviates from the best-fitting normal. After Blasco et al. (1998).

LS, MM, OR BAYES?

Just what analysis should an investigator use for a selection experiment? Obviously, in the absence of any pedigree information, a least-squares analysis is the only option, although this could also be placed in a Bayesian framework. With the pedigree in hand (either observed or inferred, see Chapter 20), a mixed-model analysis is much more powerful and is strongly preferred over LS, unless there is strong evidence that model assumptions are violated. If a mixed-model approach is appropriate and chosen, should the analysis be standard or Bayesian? As mentioned, the Bayesian approach does a much better job of treating uncertainty, but this comes at a higher computational cost, especially when one

does a proper analysis using several different priors to assess sensitivity. Perhaps the best advice is that offered by Blasco (2001):

> " The choice of one school or the other should be related to whether there are solutions in one school that the other does not offer, to how easily the problems are solved, and to how comfortable scientists feel with the way they convey their results. "

Blasco's last point is especially important: It is far more important for invesigators to use a method with which they are comfortable, in the sense of knowing its limitations and having some intution into the approach, than to simply use a method because it is new and trendy.

Generally speaking, simpler methods (such as OLS) tend to be more robust to model fragility than more complex approaches (e.g., mixed-models). While the latter can be considerably more powerful *when* model assumptions hold, they can also be significantly more biased when they fail. Best practice is to use several different approaches in the analysis of any dataset. If the results are consistent, one has additional confidence the model assumptions may be holding. If they yield rather different results, this is critical for the investigator to know, suggesting a much more careful examination of model assumptions may be in order.

Literature Cited

Beniwal, B. K., I. M. Hastings, R. Thompson, and W. G. Hill. 1992a. Estimation of changes in genetic parameters in selected lines of mice using REML with an animal model. 1. Lean mass. *Heredity* 69: 352–360. [19]

Beniwal, B. K., I. M. Hastings, R. Thompson, and W. G. Hill. 1992b. Estimation of changes in genetic parameters in selected lines of mice using REML with an animal model. 2. Body weight, body compostion and litter size. *Heredity* 69: 361–371. [19]

Blair, H. T., and E. J. Pollak. 1984. Estimation of genetic trend in a selected population with and without the use of a control population. *J. Animal Sci.* 58: 878–886. [19]

Blasco, A. 2001. The Bayesian controversy in animal breeding. *J. Anim. Sci.* 79: 2023–2046. [19]

Blasco, A., D. Sorensen, and J. P. Bidanel. 1998. Bayesian inference of genetic parameters and selection response for litter size components in pigs. *Genetics* 149: 301–306. [19]

Boichard, D., B. Bonaiti, A. Barbat, and S. Mattalia. 1995. Three methods to validate the estimation of genetic trend for dairy cattle. *J. Diary Sci.* 78: 431–437. [19]

Bulmer, M. G. 1971b. The effect of selection on genetic variability. *Amer. Nat.* 105: 201–211. [19]

Clément, V., B. Bibé, E. Verrier, J.-M. Elsen, E. Manfredi, J. Bouix, and E. Hanocq. 2001. Simulation analysis to test the influence of model adequacy and data structure on the estimation of genetic parameters for traits with direct and maternal effects. *Genet. Sel. Evol.* 33: 369–395. [19]

Curnow, R. N. 1961. The estimation of repeatability and heritability from records subject to culling. *Biometry* 17: 553–566. [19]

Damgaard, L. H. 2007. Technical note: How to use Winbugs to draw inferences in animal models. *J. ANim. Sci.* 85: 1363–1368. [19]

de Boer, I. J. M., and J. A. M. van Arendonk. 1992. Prediction of additive and dominance effects in selection or unselected populations with inbreeding. *Theor. Appl. Genet.* 84: 451–459. [19]

Efron, B. 1986. Why isn't everyone a Bayesian? *Am. Stat.* 40: 1-11. [19]

Estany, J., M. Baselga, A. Blasco, and J. Camacho. 1989. Mixed model methodology for the estimation of genetic response to selection in litter size of rabbits. *Live. Prod. Sc.* 21: 67–75. [19]

Fernando, R. L., and D. Gianola. 1990. Statistical inferences in populations undergoing selection or nonrandom mating. *In* D. Gianola and K. Hammond (eds.) *Statistical methods for genetic improvement of livestock*, pp. 437–453. Springer-Verlag, NY. [19]

Geman, S. and D. Geman. 1984. Stochastic relaxation, Gibbs distribution and Bayesian restoration of images. *IEE Transactions on Pattern Analysis and Machine Intelligence* 6: 721–741. [19]

Geyer, C. J. 1992. Practical Markov chain Monte Carlo (with discussion). *Stat. Sci.* 7: 473–511. [19]

Gianola, D., and R. L. Fernando. 1986. Bayesian methods in animal breeding theory. *J. Anim. Sci.* 63: 217–244. [19]

Gianola, D., R. L. Fernando, S. Im, and J. L. Foulley. 1989. Likelihood estimation of quantitative genetic parameters when selection occurs: models and problems. *Genome* 31: 768–777. [19]

Gianola, D., J. L. Foulley, and R. L. Fernando. 1986. Prediction of breeding values when variances are not know. *Genet. Sel. Evol.* 18: 485–498. [19]

Gianola, D., S. Im, and R. L. Fernando. 1988. Prediction of breeding values under Henderson's selection model: a revisitation. *J. Dairy Sci.* 71: 2790–2798. [19]

Glymour, C. 1981. Why I am not a Bayesian, *In* C. Glymour (ed.) *Theory and evidence*, pp. 63–93. University of Chicago Press, Chicago. [19]

Graser, H.-U., S. P. Smith, and B. Tier. 1987. A derivative-free approach for estimating variance components in animal models by restricted maximum likelihood. *J. Anim. Sci.* 64: 1362–1370. [19]

Harville, D. A. 1977. Maximum likelihood approaches to variance component estimation and to related problems. *J. Am. Stat. Assoc.* 72: 320–338. [19]

Harville, D. A. 1990. BLUP (Best Linear Unbiased Prediction) and beyond. *In* D. Gianola and K. Hammon (eds), *Advances in statistical methods for genetic improvement of livestock*, pp. 239–276. Springer-Verlag, NY. [19]

Hastings, W. K. 1970. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* 57: 97–109. [19]

Heath, S. C., G. Bulfield, R. Thompson, and P. D. Keightley. 1995. Rates of change of genetic parameters of body weight in selected mouse lines. *Genet. Res.* 66: 19–25. [19]

Henderson, C. R. 1949. Estimates of changes in herd environment. *J. Dairy Sci.* 32: 706. [19]

Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31: 423–447. [19]

Henderson, C. R. 1976. A simple method for the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32: 69–83. [19]

Henderson, C. R., O. Kempthorne, S. R. Searle, and C. M. von Krosigk. 1959. The estimation of environmental and genetic trends from records subject to culling. *Biometrics* 15: 192–218. [19]

Hoeschele, I, and P. M. Van Raden. 1991. Rapid inversion of dominance relationship matrices for noninbred populations by including sire by dam subclass effects. *J. Dairy Sci.* 74: 557–569. [19]

Hofer, A. 1998. Variance component estimation in animal breeding: a review. *J. Anim. Breed. Genet.* 115: 247–265. [19]

Hoaglin, D. C., and R. E. Welsch. 1978. The Hat Matrix in regression and ANOVA. *Amer. Stat.* 32: 17–22. [19]

Jensen, C. S. Wang, D. A. Sorensen, and D. Gianola. 1994. Marginal inferences of variance and covariance components for traits influenced by maternal and direct genetic effects using the Gibbs sampler. *Acta Agric.* 44: 193–201. [19]

Jeyaruban, M. G., and J. P. Gibson. 1996. Estimation of additive genetic variance in commercial layer poultry and simulated populations under selection. *Theor. Appl. Genet.* 92: 483–491. [19]

Juga, J., and R. Thompson. 1989. Estimation of variance components in populations selected over multiple generations. *Acta Agric. Scand.* 39: 79–89. [19]

Kackar, R. N., and D. A. Harville. 1981. Unbiasedness of two-stage estimation and prediction for mixed linear models. *Comm. Stat. Theor. Meth.* A10: 1249–1261. [19]

Kennedy, B. W. 1990. Use of mixed model methodology in analysis of designed experiments. *In* D. Gianola and K. Hammond (eds.), *Statistical methods for genetic improvement of livestock,* pp. 77–97. Springer-Verlag, NY. [19]

Kennedy, B. W., and D. A. Sorensen. 1988. Properties of mixed-model methods for prediction of genetic merit. *In* B. S. Weir, E. J. Eisen, M. M. Goodman, and G. Namkoong (eds.), *Proceedings of the second international conference on quantitative genetics,* pp. 91 – 103. Sinauer Assoc., Sunderland, MA. [19]

Kennedy, B. W., L. R. Schaeffer, and D. A. Sorensen. 1988. Genetic properties of animal models. *J. Dairy Sci.* 17: Supp. 2 17 – 26. [19]

Kennedy, B. W., and D. Trus. 1993. Considerations on genetic connectedness between management units under an animal model. *J. Anim. Sci.* 71: 2341–2352. [19]

Keynes, J. M. 1921. *A treatise on probability*. Macmillan, London. [19]

Kruuk, L. E. B. 2004. Estimating genetic parameters in natural populations using the 'animal model.' *Phil. Trans. R. Soc. Lond. B* 359: 873–890. [19]

Maki-Tanila, A., and B. W. Kennedy. 1986. Mixed model methodology under genetic models with a small number of additive and non-additive loci. *Proc. Third World Cong. on Genetics Applied to Livestock Production* XII: 443–448. Ames, Iowa. [19]

Martinez, V., L. Bünger, and W. G. Hill. 2000. Analysis of response to 20 generations of selection for body composition in mice: fit to the infinitesimal model assumptions. *Genet. Sel. Evol.* 32: 3–21. [19]

Metropolis, N., and S. Ulam. 1949. The Monte Carlo method. *J. Am. Statist. Assoc.* 44: 335–341. [19]

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A.Teller, and H. Teller. 1953. Equations of state calculations by fast computing machines. *Journal of Chemical Physics* 21: 1087–1091. [19]

Meyer, K., and W. G. Hill. 1991. Mixed model analysis of a selection experiment for food intake in mice. *Genet. Res. Camb.* 57: 71–81. [19]

Milner, J. M., J. M. Pemberton, S. Bortherstone, and S. D. Albon. 2000. Estimating variance components and heritabilities in the wild: a case study using the animal model approach. *J. Evol Biol.* 13: 804–813. [19]

Misztal, I. 1997. Estimation of variance components with large-scale dominance models. *J. Dairy Sci.* 80: 965–974. [19]

Misztal, I. 2008. Reliable computing in estimation of variance components. *J. Anim. Breed. Genet.* 125: 363–370. [19]

Ollivier, L. 1999. On the use of animal models in the analysis of selection experiments. *Genet. Sel. Evol.* 31: 135–148. [19]

Ovaskainen, O., J. M. Cano, and J. Merilä. 2008. A Bayesian framework for comparative quantitative genetics. *Proc. Roy. Soc. B* 275: 669–678. [19]

Patterson, H. D., and R. Thompson. 1971. Recovery of interblock information when block sizes are unequal. *Biometrika* 58: 545–554. [19]

Quaas, R. L. 1976. Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics* 32: 949–953. [19]

Quass, R. L., and E. J. Pollak. 1980. Mixed model methodology for farm and ranch beef cattle. *J. Anim. Sci.* 51: 1277–1287. [19]

Raftery, A. E., and S. Lewis. 1992 . How many iterations in the Gibbs sampler? *In* J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (eds.) *Bayesian statistics 4*, pp. 763–773. Oxford University Press. [19]

Robertson, A. 1977. The effect of selection on the estimation of genetic parameters. *Z. Tier Zuchtungsbiol.* 94: 131–135. [19]

Rodriguez, M. C., M. Toro and L. Silió. 1996. Selection on lean growth in a nucleus of Landrace pigs: an analysis using Gibbs sampling. *Anim. Sci.* 63: 243–253. [19]

Rothschild, M. F., C. R. Henderson, and R. L. Quaas. 1979. Effects of selection on variances and covariances of simulated first and second lactations. *J. Dairy Sci.* 62: 996–1002. [19]

Schaeffer, L. R. and H. Song. 1978. Selection bias and REML variance-covariance component estimation. *J. Dairy Sci.* 61: 91–92. [19]

Smith, S. P., and H.-U. Graser. 1986. Estimating variance components in a class of mixed models by restricted maximum likelihood. *J. Dairy Sci.* 69: 1156–1165. [19]

Smith, S. P., and A. Mäki-Tanila. 1990. Genotypic covariance matrices and their inverses for models allowing dominance and inbreeding. *Genet. Sel. Evol.* 22: 65-91. [19]

Sorensen, D., R. Fernando, and D. Gianola. 2001. Inferring the trajectory of genetic variance in the course of artificial selection. *Genet. Res. Camb.* 77: 83–94. [19]

Sorensen, D. and D. Gianola. 2002. *Likelihood, Bayesian and MCMC methods in quantitative genetics.* Springer. [19]

Sorensen, D. A., B. Guldbrandtsen, and J. Jensen. 2003. On the need for a control line in selection experiments. A likelihood analysis. *Genet. Sel. Evol.* 35: 3–20. [19]

Sorensen, D. A., and K. Johannsson. 1992. Estimation of direct and correlated response to selection using univariate animal models. *J. Anim. Sci.* 70: 2038–2044. [19]

Sorensen, D. A., and B. W. Kennedy. 1983. The use of the relationship matrix to account for genetic drift variances in the analysis of selection experiments. *Theor. Appl. Genet.* 66: 217–220. [19]

Sorensen, D. A., and B. W. Kennedy. 1984a. Estimation of response to selection using least-squares and mixed model methodology. *J. Anim. Sci.* 58: 1097–1106 [19]

Sorensen, D. A., and B. W. Kennedy. 1984b. Estimation of genetic variances from unselected and selected populations. *J. Anim. Sci.* 59: 1213–1223. [19]

Sorensen, D. A., and B. W. Kennedy. 1986. Analysis of selection experiments using mixed model methodology. *J. Anim. Sci.* 63: 245–258. [19]

Sorensen, D. A., C. S. Wang, J. Jensen, and D. Gianola. 1994. Bayesian analysis of genetic change due to selection using Gibbs sampling. *Genet. Sel. Evol.* 26: 333–360. [19]

Southwood, O. I., and B. W. Kennedy. 1991. Genetic and environmental trends for litter size in swine. *J. Anim. Sci.* 69: 3177–3182. [19]

Su, G., P. Sørensen, and D. Sorensen. 1997. Inferences about variance components and selection response for body weight in chickens. *Genet. Sel. Evol* 29: 413–425. [19]

Thompson, R. 1973. The estimation of variance and covariance components with an application when records are subject to culling. *Biometrics* 29: 527–550. [19]

Thompson, R. 1977. The estimation of heritability with unbalanced data. II. Data available on more than two generations. *Biometrics* 33: 496–504. [19]

Thompson, R. 1986. Estimation of realized heritability in a selected population using mixed model methods. *Genet. Sel. Evol.* 18: 475–484. [19]

Thompson, R. 2008. Estimation of quantitative genetic parameters. *Proc. R. Soc. B* 275: 679–686. [19]

Thompson, R., and K. D. Atkins. 1994. Sources of information for estimating heritability from selection experiments. *Genet. Res.* 63: 49–55. [19]

Thompson, R., S. Brotherstone, and I. M. White. 2005. Estimation of quantitative genetic parameters. *Phil. Trans. R. Soc. B* 360: 1469–1477. [19]

Thompson, R., and E. Mäntysaari. 2004. Prospects for statistical methods in animal breeding. *J. Ind. Soc. Agri. Stat.* 57: 15–25. [19]

van der Werf, J. H. J. 1990. A note on the use of conditional models to estimate additive genetic variance in selected populations. *In* W. G. Hill, R. Thompson, and J. A. Woolliams (eds.), *Proc. 4th World Congr. Genet. Appl. Livestock Prod.*, Vol. 13, pp. 476–479. Edinburgh. [19]

van der Werf, J. H. J., and I. J. M. de Boer. 1990. Estimation of additive genetic variance when base populations are selected. *J. Anim. Sci.* 68: 3124–3132. [19]

van der Werf, J. H. J., and R. Thompson. 1992. Variance decomposition in the estimation of genetic variance with selected data. *J. Anim. Sci.* 70: 2975–2985. [19]

Visscher, P., and R. Thompson. 1990. REML estimates of parameters for fat yield in pedigree herds in the U.K. using an individual animal model: male and female heritability estimates. *Proc. 4th World Congr. Genet. Appl. Livest. Prod. Edinburgh* 14: 484–487. [19]

Wang, C. S., J. J. Rutledge, and D. Gianola. 1993. Marginal inferences about variance components in a mixed linear model using Gibbs sampling. *Genet. Sel. Evol.* 21: 41–62. [19]

Wang, C. S., D. Gianola, D. A. Sorensen, J. Jensen, A. Christensen, and J. J. Rutledge. 1994a. Response to selection for litter size in Danish landrace pigs: a Bayesian analysis. *Theor. Appl. Genet.* 88: 229–230. [19]

Wang, C. S., J. J. Rutledge, and D. Gianola. 1994b. Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs. *Genet. Sel. Evol.* 26: 91–115. [19]

Wilson, A. 2008. Why h^2 does not always equal VA/VP. *J. Evol. Bio.* 21: 647–650. [19]