0.1 Introduction to Mixed Models

All models are characterized by the mean α and the error terms. In addition to these terms, any model described so far will have either random effects terms or fixed effects terms and accordingly are referred to as random or fixed models. Models that have both fixed effects terms and random effects terms are known as 'mixed effects models'. Once the theory underlying fixed and random effects models has been fully understood, the progression to understanding mixed models is very simple.

Elaborating on the original mice litter example, the six litters by each mouse were fed according to three different dietary treatments (Searle, 1997). Therefore a fixed effect ϕ_j has been added to the model, which is now formulated as follows;

$$y_{ij} = \mu + \delta_i + \phi_j + \gamma_{ij} + \epsilon_{ijk} \tag{1}$$

As before, an interaction effect γ_{ij} must also be added to the model. In cases where the interaction term describes the combined effect of fixed and random components, it should be treated as random effect. The variance of the above model is composed of the σ^2_δ, σ^2_γ and σ^2_ϵ .

It may be shown that the interaction factors make no contribution to the outcome, i.e γ_{ij} is consistently calculated as zero. Considering the skin tumour example, a person's age would bear no relation to their gender and hence there would be plausible interaction between the two factors. Indeed , in keeping with the 'Law of Parsimony', factors should be specified such that each would convey separate information. However, interaction terms are extant when the model specifies repeated observations, as there is necessarily a relationship between observations from the same subject. Importantly, interaction effects, being random effects, are attended by variance component terms and therefore also contribute to the overall variance of the model.

Searle (1997) gives a mixed effects model formulation for the Grubbs artillery study.

y_{ij} is the muzzle velocity of the ith shell, as measured by the jth chronometer.

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij} \tag{2}$$

In this formulation α_i is the random effect of round i, and the fixed effect component β_j is the bias in chronometer j. (Also, no interaction term is used).

0.2 Matrix Formulation

There are matrix (i.e multivariate) formulations of both fixed effects models and random effects models. Brown and Prescott (1999) remarks that the matrix notation makes the underlying theory of mixed effects models much easier to work with. The fixed effects models can be specified as follows;

$$\mathbf{Y} = \mathbf{Xb} + \mathbf{e} \tag{3}$$

$\mathbf{Y}$ is the vector of n observations, with dimension $n \times 1$. $\mathbf{b}$ is a vector of fixed p effects, and has dimension $p \times 1$. It is composed of coefficients, with the first element being the population mean. For the skin tumour example, with the three specified fixed effects, $p = 4$. $\mathbf{X}$ is known as the design 'matrix', model matrix for fixed effects, and comprises 0s or 1s, depending on whether the relevant fixed effects have any effect on the observation is question. $\mathbf{X}$ has dimension $n \times p$. $\mathbf{e}$ is the vector of residuals with dimension $n \times 1$.

The random effects models can be specified similarly. $\mathbf{Z}$ is known as the 'model matrix for random effects', and also comprises 0s or 1s. It has dimension $n \times q$. $\mathbf{u}$ is a vector of random q effects, and has dimension $q \times 1$.

$$\mathbf{Y} = \mathbf{Zu} + \mathbf{e} \tag{4}$$

Again, once the component fixed effects and random effects components are considered, progression to a mixed model formulation is a simple step. Further to Laird and Ware (1982), it is conventional to formulate a mixed effects model in matrix form as follows:

$$\mathbf{Y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e} \tag{5}$$

($E(\mathbf{u}) = 0$, $E(\mathbf{e}) = 0$ and $E(\mathbf{y}) = \mathbf{Xb}$)

0.2.1 Stating the LME Model

The general linear mixed model is

$$Y = X\beta + Zu + \varepsilon$$

where Y is a $(n \times 1)$ vector of observed data, X is an $(n \times p)$ fixed-effects design or regressor matrix of rank k, Z is a $(n \times g)$ random-effects design or regressor matrix, u is a $(g \times 1)$ vector of random effects, and ε is an $(n \times 1)$ vector of model errors (also random effects). The distributional assumptions made by the MIXED procedure are as follows: is normal with mean 0 and variance G; ε is normal with mean 0 and variance R; the random components u and ε are independent. Parameters of this model are the fixed-effects and all unknowns in the variance matrices G and R. The unknown variance elements are referred to as the covariance parameters and collected in the vector *theta*.

The concept of critiquing the model-data agreement applies in mixed models in the same way as in linear fixed-effects models. In fact, because of the more complex model structure, you can argue that model and data diagnostics are even more important. For example, you are not only concerned with capturing the important variables in the model. You are also concerned with distributing them correctly between the fixed

and random components of the model. The mixed model structure presents unique and interesting challenges that prompt us to reexamine the traditional ideas of influence and residual analysis. This paper presents the extension of traditional tools and statistical measures for influence and residual analysis to the linear mixed model and demonstrates their implementation in the MIXED procedure (experimental features in SAS 9.1). The remainder of this paper is organized as follows. The Background section briefly discusses some mixed model estimation theory and the challenges to model diagnosis that result from it.

0.3 Statement of the LME model

A linear mixed effects model is a linear mdoel that combined fixed and random effect terms formulated by Laird and Ware (1982) as follows;

$$Y_i = X_i\beta + Z_i b_i + \epsilon_i$$

- Y_i is the $n \times 1$ response vector
- X_i is the $n \times p$ Model matrix for fixed effects
- β is the $p \times 1$ vector of fixed effects coefficients
- Z_i is the $n \times q$ Model matrix for random effects
- b_i is the $q \times 1$ vector of random effects coefficients, sometimes denoted as u_i
- ϵ is the $n \times 1$ vector of observation errors

The linear mixed effects model is given by

$$Y = X\beta + Zu + \epsilon \tag{6}$$

0.4 Standard Deviation of Differences

In computing limits of agreement, it is first necessary to have an estimate for the standard deviations of the differences. When the agreement of two methods is analyzed using LME models, a clear method of how to compute the standard deviation is required. As the estimate for inter-method bias and the quantile would be the same for both methodologies, the focus hereon is solely on the variance of differences.

The standard deviation of the differences of methods x and y is computed using values from the overall VC matrix.

$$\mathrm{Var}(x - y) = \mathrm{Var}(x) + \mathrm{Var}(y) - 2\mathrm{Cov}(x, y)$$

0.5 Restricted Likelihood Estimation

Restricted maximum likelihood (REML) is an alternative methods of computing parameter estimated, developed by Paterson and Thompson (1971) and Harville (1977) to provide unbiased estimates of variance and covariance parameters. The REML approach does not base estimates on a maximum likelihood fit of all the information, but instead uses a likelihood function derived from a data set, transformed to remove the irrelevant influences (Dodge, 2003).

REML obtains estimates of the fixed effects using non-likelihoodlike methods, such as ordinary least squares or generalized least squares, and then using these estimates it maximizes the likelihood of the residuals (subtracting off the fixed effects) to obtain estimates of the variance parameters. In most software packages REML is the default algorithm used to compute coefficients for the predictor variables. REML estimation reduces the bias in the variance component, and also handles high correlations more effectively, and is less sensitive to outliers than ML.

The variance components in the LME model may be estimated by ML or REML. Maximum Likelihood estimates do not take into account the estimation of fixed effects and so are biased downwards. REML estimates accounts for the presence of these nuisance parameters by maximising the linearly independent error contrasts to obtain more unbiased estimates.

McCullough and Searle (2001) describes two important outcomes of using REML. Firstly variance components can be estimated without being affected by fixed effects. Secondly in estimating variance components with REML, degrees of freedom for the fixed effects can be taken into account implicitly, whereas with ML they are not.

Restricted maximum likelihood is often preferred to maximum likelihood because REML estimation reduces the bias in the variance component by taking into account the loss of degrees of freedom that results from estimating the fixed effects in $\boldsymbol{\beta}$. Restricted maximum likelihood also handles high correlations more effectively, and is less sensitive to outliers than maximum likelihood. The problem with REML for model building is that the likelihoods obtained for different fixed effects are not comparable. Hence it is not valid to compare models with different fixed effects using a likelihood ratio test or AIC when REML is used to estimate the model. Therefore models derived using ML must be used instead.

0.6 nlme - Variance functions

Variance functions are applied to LME models through the `'weights'` argument. *R* supports several variance functions.

'`varIdent`' cosntructs a model with different variances per stratum.

0.6.1 Diagnostic plots

[Pinheiro Bates Page 391] Diagnostic plots for identifying within-group heteroscedascity and assessing the adequacy of a variance function can also be used with 'nlme' objects.

0.7 Likelihood and estimation

Likelihood is the hypothetical probability that an event that has already occurred would yield a specific outcome. Likelihood differs from probability in that probability refers to future occurrences, while likelihood refers to past known outcomes.

The likelihood function ($L(\theta)$)is a fundamental concept in statistical inference. It indicates how likely a particular population is to produce an observed sample. The set of values that maximize the likelihood function are considered to be optimal, and are used as the estimates of the parameters. For computational ease, it is common to use the logarithm of the likelihood function, known simply as the log-likelihood ($\ell(\theta)$).

0.8 Linear Mixed effects Models

A linear mixed effects (LME) model is a statistical model containing both fixed effects and random effects (random effects are also known as variance components). LME models are a generalization of the classical linear model, which contain fixed effects only. When the levels of factors are considered to be sampled from a population, and each level is not of particular interest, they are considered random quantities with associated variances. The effects of the levels, as described, are known as random effects. Random effects are represented by unobservable normally distributed random variables. Conversely fixed effects are considered non-random and the levels of each factor are of specific interest.

Fisher (1918) introduced variance components models for use in genetical studies. Whereas an estimate for variance must take an non-negative value, an individual variance component, i.e. a component of the overall variance, may be negative.

The methodology has developed since, including contributions from Tippett (1931), who extend the use of variance components into linear models, and Eisenhart (1947), who introduced the 'mixed model' terminology and formally distinguished between mixed and random effects models. Henderson (1950) devised a methodology for deriving estimates for both the fixed effects and the random effects, using a set of equations that would become known as 'mixed model equations' or 'Henderson's equations'. LME methodology is further enhanced by Henderson's later works (Henderson, 1953; Henderson et al., 1959, 1963, 1973, 1984). The key features of Henderson's work provide the basis for the estimation techniques.

Hartley and Rao (1967) demonstrated that unique estimates of the variance components could be obtained using maximum likelihood methods. However these estimates are known to be biased 'downwards' (i.e. underestimated) , because of the assump-

tion that the fixed estimates are known, rather than being estimated from the data. Patterson and Thompson (1971) produced an alternative set of estimates, known as the restricted maximum likelihood (REML) estimates, that do not require the fixed effects to be known. Thusly there is a distinction the REML estimates and the original estimates, now commonly referred to as ML estimates.

Laird and Ware (1982) provides a form of notation for notation for LME models that has since become the standard form, or the basis for more complex formulations. Due to computation complexity, linear mixed effects models have not seen widespread use until many well known statistical software applications began facilitating them. SAS Institute added PROC MIXED to its software suite in 1992 (Singer, 1998). Pinheiro and Bates (1994) described how to compute LME models in the `S-plus` environment.

Using Laird-Ware form, the LME model is commonly described in matrix form,

$$y = X\beta + Zb + \epsilon \tag{7}$$

where y is a vector of N observable random variables, β is a vector of p fixed effects, X and Z are $N \times p$ and $N \times q$ known matrices, and b and ϵ are vectors of q and N, respectively, random effects such that $\mathrm{E}(b) = 0,\ \mathrm{E}(\epsilon) = 0$ and where D and Σ are positive definite matrices parameterized by an unknown variance component parameter vector θ. The variance-covariance matrix for the vector of observations y is given by $V = ZDZ' + \Sigma$. This implies $y \sim (X\beta, V) = (X\beta, ZDZ' + \Sigma)$. It is worth noting that V is an $n \times n$ matrix, as the dimensionality becomes relevant later on. The notation provided here is generic, and will be adapted to accord with complex formulations that will be encountered in due course.

0.8.1 Estimation

Estimation of LME models involve two complementary estimation issues'; estimating the vectors of the fixed and random effects estimates $\hat{\beta}$ and $\hat{b}$ and estimating the variance covariance matrices D and Σ. Inference about fixed effects have become known as 'estimates', while inferences about random effects have become known as 'predictions'. The most common approach to obtain estimators are Best Linear Unbiased Estimator (BLUE) and Best Linear Unbiased Predictor (BLUP). For an LME model given by (7), the BLUE of $\hat{\beta}$ is given by

$$\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y,$$

whereas the BLUP of $\hat{b}$ is given by

$$\hat{b} = DZ'V^{-1}(y - X\hat{\beta}).$$

Estimation of the fixed parameters

The vector y has marginal density $y \sim \mathrm{N}(X\beta, V)$, where $V = \Sigma + ZDZ'$ is specified through the variance component parameters θ. The log-likelihood of the fixed parameters (β, θ) is

$$\ell(\beta, \theta|y) = -\frac{1}{2}\log|V| - \frac{1}{2}(y - X\beta)'V^{-1}(y - X\beta), \tag{8}$$

and for fixed θ the estimate $\hat{\beta}$ of β is obtained as the solution of

$$(X'V^{-1}X)\beta = X'V^{-1}y. \tag{9}$$

Substituting $\hat{\beta}$ from (9) into $\ell(\beta, \theta|y)$ from (8) returns the *profile* log-likelihood

$$\begin{aligned} \ell_P(\theta \mid y) &= \ell(\hat{\beta}, \theta \mid y) \\ &= -\frac{1}{2}\log|V| - \frac{1}{2}(y - X\hat{\beta})'V^{-1}(y - X\hat{\beta}) \end{aligned}$$

of the variance parameter θ. Estimates of the parameters θ specifying V can be found by maximizing $\ell_P(\theta \mid y)$ over θ. These are the ML estimates.

For REML estimation the *restricted* log-likelihood is defined as

$$\ell_R(\theta \mid y) = \ell_P(\theta \mid y) - \frac{1}{2}\log|X'VX|.$$

The REML approach does not base estimates on a maximum likelihood fit of all the information, but instead uses a likelihood function derived from a data set, transformed to remove the irrelevant influences (Dodge, 2003). Restricted maximum likelihood is often preferred to maximum likelihood because REML estimation reduces the bias in the variance component by taking into account the loss of degrees of freedom that results from estimating the fixed effects in $\boldsymbol{\beta}$. Restricted maximum likelihood also handles high correlations more effectively, and is less sensitive to outliers than maximum likelihood. The problem with REML for model building is that the likelihoods obtained for different fixed effects are not comparable. Hence it is not valid to compare models with different fixed effects using a likelihood ratio test or AIC when REML is used to estimate the model. Therefore models derived using ML must be used instead.

Estimation of the random effects

The established approach for estimating the random effects is to use the best linear predictor of b from y, which for a given β equals $DZ'V^{-1}(y-X\beta)$. In practice β is replaced by an estimator such as $\hat{\beta}$ from (9) so that $\hat{b} = DZ'V^{-1}(y - X\hat{\beta})$. Pre-multiplying by the appropriate matrices it is straightforward to show that these estimates $\hat{\beta}$ and $\hat{b}$ satisfy the equations in (**??**).

Algorithms for likelihood function optimization

Iterative numerical techniques are used to optimize the log-likelihood function and estimate the covariance parameters θ. The procedure is subject to the constraint that R and D are both positive definite. The most common iterative algorithms for optimizing the likelihood function are the Newton-Raphson method, which is the preferred method, the expectation maximization (EM) algorithm and the Fisher scoring methods.

The EM algorithm, introduced by Dempster et al. (1977), is an iterative technique for maximizing complicated likelihood functions. The algorithm alternates between performing an expectation (E) step and the maximization (M) step. The 'E' step computes the expectation of the log-likelihood evaluated using the current estimate for the variables. In the 'M' step, parameters that maximize the expected log-likelihood, found on the previous 'E' step, are computed. These parameter estimates are then used to determine the distribution of the variables in the next 'E' step. The algorithm alternatives between these two steps until convergence is reached.

The main drawback of the EM algorithm is its slow rate of convergence. Consequently the EM algorithm is rarely used entirely in LME estimation, instead providing an initial set of values that can be passed to other optimization techniques.

The Newton Raphson (NR) method is the most common, and recommended technique for ML and REML estimation. The NR algorithm minimizes an objective function defines as -2 times the log likelihood for the covariance parameters θ. At every iteration the NR algorithm requires the calculation of a vector of partial derivatives, known as the gradient, and the second derivative matrix with respect to the covariance parameters. This is known as the observed Hessian matrix. Due to the Hessian matrix, the NR algorithm is more time-consuming, but convergence is reached with fewer

iterations compared to the EM algorithm. The Fisher scoring algorithm is an variant of the NR algorithm that is more numerically stable and likely to converge, but not recommended to obtain final estimates.

0.8.2 Formulation of the response vector

Information of individual i is recorded in a response vector $\boldsymbol{y}_i$. The response vector is constructed by stacking the response of the 2 responses at the first instance, then the 2 responses at the second instance, and so on. Therefore the response vector is a $2n_i \times 1$ column vector. The covariance matrix of $\boldsymbol{y_i}$ is a $2n_i \times 2n_i$ positive definite matrix $\boldsymbol{\Omega}_i$.

Consider the case where three measurements are taken by both methods A and B, $\boldsymbol{y}_i$ is a 6×1 random vector describing the ith subject.

$$\boldsymbol{y}_i = (y_i^{A1}, y_i^{B1}, y_i^{A2}, y_i^{B2}, y_i^{A3}, y_i^{B3})\boldsymbol{\prime}$$

The response vector $\boldsymbol{y_i}$ can be formulated as an LME model according to Laird-Ware form.

$$\boldsymbol{y_i} = \boldsymbol{X_i\beta} + \boldsymbol{Z_i b_i} + \boldsymbol{\epsilon_i}$$

$$\boldsymbol{b_i} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{D})$$

$$\boldsymbol{\epsilon_i} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R_i})$$

Information on the fixed effects are contained in a three dimensional vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)\boldsymbol{\prime}$. For computational purposes β_2 is conventionally set to zero. Consequently $\boldsymbol{\beta}$ is the solutions of the means of the two methods, i.e. $E(\boldsymbol{y}_i) = \boldsymbol{X}_i\boldsymbol{\beta}$. The variance covariance matrix $\boldsymbol{D}$ is a general 2×2 matrix, while $\boldsymbol{R}_i$ is a $2n_i \times 2n_i$ matrix.

0.8.3 Decomposition of the response covariance matrix

The variance covariance structure can be re-expressed in the following form,

$$\text{Cov}(\text{y}_i) = \boldsymbol{\Omega_i} = \boldsymbol{Z}_i \boldsymbol{D} \boldsymbol{Z}_i' + \boldsymbol{R_i}.$$

$\boldsymbol{R_i}$ can be shown to be the Kronecker product of a correlation matrix $\boldsymbol{V}$ and $\boldsymbol{\Lambda}$. The correlation matrix $\boldsymbol{V}$ of the repeated measures on a given response variable is assumed to be the same for all response variables. Both Hamlett et al. (2004) and Lam et al. (1999) use the identity matrix, with dimensions $n_i \times n_i$ as the formulation for $\boldsymbol{V}$. Roy (2009) remarks that, with repeated measures, the response for each subject is correlated for each variable, and that such correlation must be taken into account in order to produce a valid inference on correlation estimates. **?** proposes various correlation structures may be assumed for repeated measure correlations, such as the compound symmetry and autoregressive structures, as alternative to the identity matrix.

However, for the purposes of method comparison studies, the necessary estimates are currently only determinable when the identity matrix is specified, and the results in Roy (2009) indicate its use.

For the response vector described, Hamlett et al. (2004) presents a detailed covariance matrix. A brief summary shall be presented here only. The overall variance matrix is a 6×6 matrix composed of two types of 2×2 blocks. Each block represents one separate time of measurement.

$$\boldsymbol{\Omega}_i = \begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{D} & \boldsymbol{D} \\ \boldsymbol{D} & \boldsymbol{\Sigma} & \boldsymbol{D} \\ \boldsymbol{D} & \boldsymbol{D} & \boldsymbol{\Sigma} \end{pmatrix}$$

The diagonal blocks are Σ, as described previously. The 2×2 block diagonal matrix in $\boldsymbol{\Omega}$ gives $\boldsymbol{\Sigma}$. $\boldsymbol{\Sigma}$ is the sum of the between-subject variability $\boldsymbol{D}$ and the within subject

variability $\boldsymbol{\Lambda}$.

$\boldsymbol{\Omega_i}$ can be expressed as

$$\boldsymbol{\Omega_i} = \boldsymbol{Z}_i \boldsymbol{D} \boldsymbol{Z}_i' + (\boldsymbol{I}_{\boldsymbol{n_i}} \otimes \boldsymbol{\Lambda}).$$

The notation $\dim_{n_i}$ means an $n_i \times n_i$ diagonal block.

0.8.4 Correlation terms

Hamlett et al. (2004) demonstrated how the between-subject and within subject variabilities can be expressed in terms of correlation terms.

$$\boldsymbol{D} = \begin{pmatrix} \sigma_A^2 \rho_A & \sigma_A \sigma_b \rho_{AB} \delta \\ \sigma_A \sigma_b \rho_{AB} \delta & \sigma_B^2 \rho_B \end{pmatrix}$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} \sigma_A^2 (1 - \rho_A) & \sigma_{AB} (1 - \delta) \\ \sigma_{AB} (1 - \delta) & \sigma_B^2 (1 - \rho_B) \end{pmatrix}.$$

ρ_A describe the correlations of measurements made by the method A at different times. Similarly ρ_B describe the correlation of measurements made by the method B at different times. Correlations among repeated measures within the same method are known as intra-class correlation coefficients. ρ_{AB} describes the correlation of measurements taken at the same same time by both methods. The coefficient δ is added for when the measurements are taken at different times, and is a constant of less than 1 for linked replicates. This is based on the assumption that linked replicates measurements taken at the same time would have greater correlation than those taken at different times. For unlinked replicates δ is simply 1. Hamlett et al. (2004) provides a useful graphical depiction of the role of each correlation coefficients.

0.8.5 For Expository Purposes

For expository purposes consider the case where each item provides three replicates by each method. Then in matrix notation the model has the structure

$$\boldsymbol{y}_i = \begin{pmatrix} y_{1i1} \\ y_{2i1} \\ y_{1i2} \\ y_{2i2} \\ y_{1i3} \\ y_{2i3} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} b_{1i} \\ b_{2i} \end{pmatrix} + \begin{pmatrix} \epsilon_{1i1} \\ \epsilon_{2i1} \\ \epsilon_{1i2} \\ \epsilon_{2i2} \\ \epsilon_{1i3} \\ \epsilon_{2i3} \end{pmatrix}.$$

The between item variance covariance $\boldsymbol{G}$ is as before, while the within item variance covariance is given as

$$\boldsymbol{G} = \begin{pmatrix} g_1^2 & g_{12} \\ g_{12} & g_2^2 \end{pmatrix}$$

$$\boldsymbol{R}_i = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & 0 & 0 & 0 & 0 \\ \sigma_{12} & \sigma_2^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_1^2 & \sigma_{12} & 0 & 0 \\ 0 & 0 & \sigma_{12} & \sigma_2^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_1^2 & \sigma_{12} \\ 0 & 0 & 0 & 0 & \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

Assumptions made on the structures of $\boldsymbol{G}$ and $\boldsymbol{R}_i$ will be discussed in due course.

0.9 Linear mixed effects models

These models are used when there are both fixed and random effects that need to be incorporated into a model.

Fixed effects usually correspond to experimental treatments for which one has data for the entire population of samples corresponding to that treatment.

Random effects,on the other hand, are assigned in the case where we have measurements on a group of samples, and those samples are taken from some larger sample pool, and are presumed to be representative.

As such, linear mixed effects models treat the error for fixed effects differently than the error for random effects.

0.10 Extended LME model

The extended single level LME model relaxes the independence assumption, allowing heteroscedastic and correlated within group errors.

$$\epsilon_i = \mathcal{N}(0, \sigma^2 \Lambda_i) \tag{10}$$

Λ_i are positive definite matrices. σ^2 is factored out of the matrix for computational reasons.

Bibliography

Brown, H. and R. Prescott (1999). *Applied Mixed Models In Medicine*. John Wiley and Sons.

Dempster, A., N. Laird, and D. Rubin (1977). Maximum likelihood from incomplete data via the em algorithme. *ournal of the Royal Statistical Society. Series B 39*(1), 1–38.

Dodge, Y. (2003). *The Oxford Dictionary of Statistical Terms*. Oxford University Press.

Eisenhart, C. (1947). The assumptions underlying the analysis of variance. *Biometrics 3*(1), 1–21.

Fisher, R. (1918). The correlation between relatives on the supposition of mendelian inheritance. *Transactions of the Royal Society of Edinburgh 2*, 399–433.

Hamlett, A., L. Ryan, and R. Wolfinger (2004). On the use of PROC MIXED to estimate correlation in the presence of repeated measures. *Proceedings of the Statistics and Data Analysis Section, SAS Users Group International 198-229*, 1–7.

Hartley, H. and J. Rao (1967). Maximum-likelihood estimation for the mixed analysis of variance model. *Biometrika 54*(1/2), 93–108.

Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of American Statistical Association 72*(358), 320–338.

Henderson, C. (1953). Estimation of variance and covariance components. *Biometrics 9*(2), 226–252.

Henderson, C., O. Kempthorne, S. Searle, and C. von Krosigk (1959). The estimation of environmental and genetic trends from records subject to culling. *Biometrics 15*, 192–218.

Henderson, C., O. Kempthorne, S. Searle, and C. von Krosigk (1963). The estimation of environmental and genetic trends from records subject to culling. *Biometrics 15*, 192–218.

Henderson, C., O. Kempthorne, S. Searle, and C. von Krosigk (1973). The estimation of environmental and genetic trends from records subject to culling. *Biometrics 15*, 192–218.

Henderson, C., O. Kempthorne, S. Searle, and C. von Krosigk (1984). The estimation of environmental and genetic trends from records subject to culling. *Biometrics 15*, 192–218.

Henderson, C. R. (1950). Estimation of genetic parameters (abstract). *Annals of Mathematical Statistics 21*, 309–310.

Laird, N. M. and J. H. Ware (1982). Random-effects models for longitudinal data. *Biometrics 38*(4), 963–974.

Lam, M., K. Webb, and D. O'Donnell (1999). Correlation between two variables in

repeated measurements. *American Statistical Association, Proceedings of the Biometric Session*, 213–218.

McCullough, C. and S. Searle (2001). *Generalized , Linear and Mixed Models*. Wiley Interscience.

Paterson, H. and R. Thompson (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika 58*(3), 545–554.

Patterson, H. and R. Thompson (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika 58*(3), 545–554.

Pinheiro, J. and D. Bates (1994). *Mixed Effects Models in S and S plus* (2nd ed.). Reading, Massachusetts: Springer.

Roy, A. (2009). An application of the linear mixed effects model to ass the agreement between two methods with replicated observations. *Journal of Biopharmaceutical Statistics 19*, 150–173.

Searle, S. (1997). *Linear Models*. Wiley classics Library.

Singer, J. D. (1998). Using sas proc mixed to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics 24*(4), 323–355.

Tippett, L. (1931). *The Methods of Statistics* (1st ed.). London: Williams and Norgate.