

# Contents

<b>1</b>	<b>Model Diagnostics</b>	<b>3</b>
1.1	Model Validation . . . . .	3
1.2	Residual diagnostics . . . . .	4
1.3	Standardization and Studentization . . . . .	7
1.3.1	Pearson Residual . . . . .	8
1.4	Distinction From Linear Models (Schabenberger page 3) . . . . .	9
1.5	Taxonomy of LME Residuals . . . . .	11
1.6	Residuals in LME Models : Marginal, Conditional and Scaled . . . . .	12
1.6.1	Internally and Externally Studentized Residuals . . . . .	12
<b>2</b>	<b>Residuals diagnostics in LME Models</b>	<b>14</b>
2.0.2	Marginal Residuals . . . . .	15
2.0.3	Conditional Residuals . . . . .	16
2.1	Residual diagnostics . . . . .	16
2.2	Conditional and Marginal Residuals . . . . .	16
2.3	Pearson and Deviance Residuals . . . . .	17
<b>3</b>	<b>Galecki</b>	<b>18</b>
3.1	Model Diagnostics for Roy's Models . . . . .	18
3.2	Case Deletion Diagnostics for the Variance Ratios . . . . .	19
3.2.1	Likelihood Distances . . . . .	21
3.3	Likelihood Distance . . . . .	23

<b>4</b>	<b>Cook's Distance</b>	<b>24</b>
4.1	Cook's Distance for LMEs . . . . .	24
4.1.1	Cook's Distance . . . . .	27
4.2	Cook's Distance for LMEs . . . . .	28
4.2.1	Cook's Distance . . . . .	28
4.2.2	Cooks's Distance . . . . .	28

# Chapter 1

## Model Diagnostics

### 1.1 Model Validation

Model validation is possibly the most important step in the model building sequence. It is also one of the most overlooked. Often the validation of a model seems to consist of nothing more than quoting the  $R^2$  statistic from the fit (which measures the fraction of the total variability in the response that is accounted for by the model). Unfortunately, a high  $R^2$  value does not guarantee that the model fits the data well. Use of a model that does not fit the data well will not provide any meaningful insights to the underlying research questions.

In classical linear models, model diagnostics techniques determine whether or not the distributional assumptions are satisfied, and to assess the influence of unusual observations, and have become a required part of any statistical analysis. Well established methods are commonly available in statistical packages and standard textbooks on applied regression. However it has been noted by several papers that model diagnostics do not often accompany LME model analyses. Schabenberger (2004) discusses the state of LME diagnostics tools, providing a useful summary of established measures.

Schabenberger (2004) describes the examination of model-data agreement as com-

prising several elements; residual analysis, goodness of fit, collinearity diagnostics and influence analysis.

## 1.2 Residual diagnostics

Residual analysis is a widely used model validation technique. For classical linear models, residual diagnostics are typically implemented as a plot of the observed residuals and the predicted values. A visual inspection for the presence of trends inform the analyst on the validity of distributional assumptions, and to detect outliers and influential observations. The rationale is that, if the model is properly fitted to the model, then the residuals would approximate the random errors that one should expect; if the residuals behave randomly, with no noticeable patterns or trends. If some sort of non-random trend is evident in the model, then the model can be considered to be poorly fitted. For classical linear models, residual diagnostics are typically implemented as a plot of the observed residuals and the predicted values. A visual inspection for the presence of trends inform the analyst on the validity of distributional assumptions, and to detect outliers and influential observations. Statistical software environments, such as the R programming language, provides a suite of tests and graphical procedures for appraising a fitted linear model, with several of these procedures analysing the model residuals. However, for LME models the matter of residual is more complex, both from a theoretical point of view and from the practical matter of implementing a comprehensive analysis using statistical software. As the LME model can be tailored to the needs of the particular research question, the rationale behind the model appraisal must follow accordingly.

For an LME model, the *raw* residuals at level  $i$  are obtained by subtracting the fitted levels at that level from the response vector. If we require a *Pearson* type residual, we further divide this by the estimated within-group standard error. The Pearson residual is the raw residual divided by the square root of the within-group standard error for both methods. The fitted values at level  $i$  are obtained by adding together

the population fitted values (based only on the fixed effects estimates) and the estimated contributions of the random effects to the fitted values at grouping levels less or equal to  $i$ . In addition to Raw and Pearson residuals, there are normalized residuals : standardized residuals pre-multiplied by the inverse square-root factor of the estimated error correlation matrix. For the forthcoming examples, we will use the raw residuals.

## Preliminary Residual Analysis

The underlying assumptions for LME models are similar to those of classical linear models. There are two key techniques: a residual plot and the normal probability plot. Using the nlme package it is possible to create plots specific to each method. This is useful in determine which methods ‘disagree’ with the rest. Analysis of the residuals would determine if the methods of measurement disagree systematically, or whether or not erroneous measurements associated with a subset of the cases are the cause of disagreement. Erroneous measurements are incorrect measurements that indicate disagreement between methods that would otherwise be in agreement. Once the residuals are computed, they can be used to make an assessment about the model fit. For the LME model described in Chapter 2, we can plot the residuals against the fitted values, to assess the assumption of constant variance. If the points in a residual plot are randomly dispersed around the horizontal axis, a linear regression model is appropriate for the data; otherwise, a non-linear model is more appropriate.

LME models assume that the residuals of the model are normally distributed. The residuals can be divided according to groups according to the method of measurement. In the following examples, we separately assess normality the  $J$  method residuals (the first 255 residuals) and  $S$  method residuals (the remaining 255). Importantly the residuals from the  $J$  method are normally distributed, but there is non-normality of the residuals according to the  $S$  method.

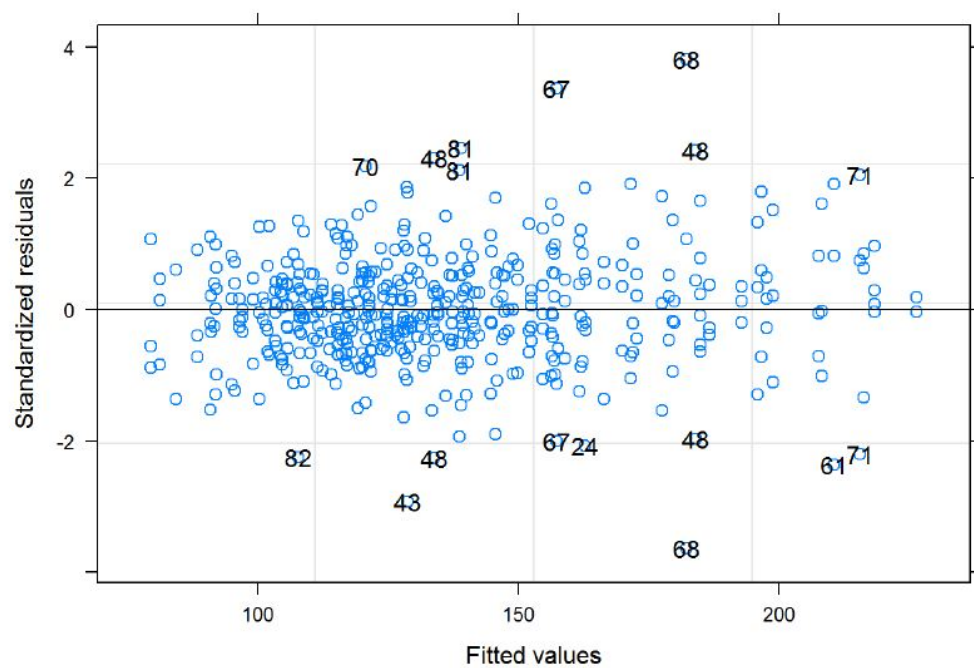


Figure 1.2.1:

```
> shapiro.test(resid(JS.roy1)[1:255])
```

Shapiro-Wilk normality test

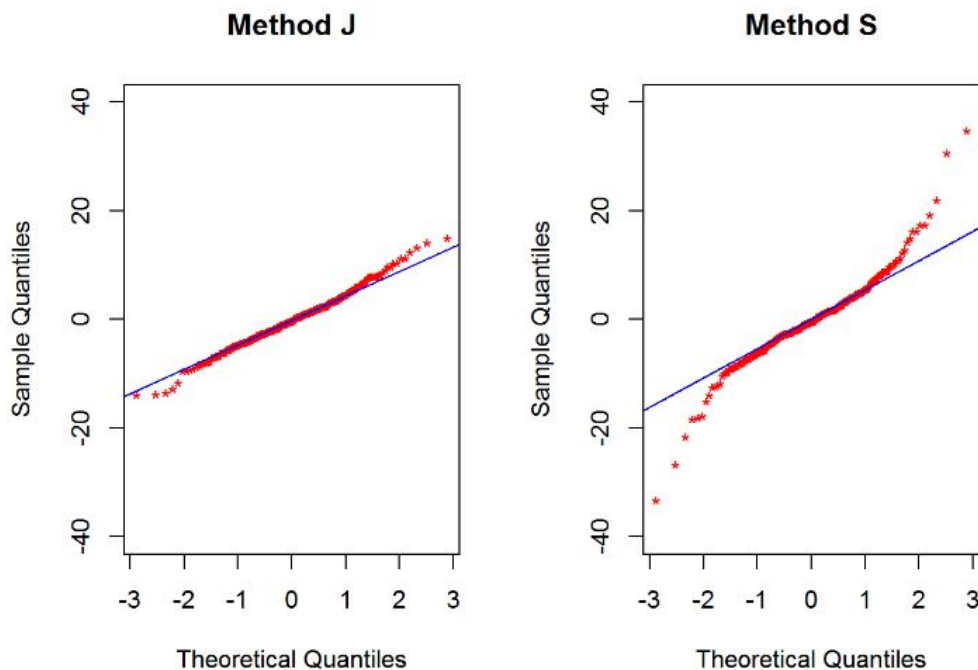
data: resid(JS.roy1)[1:255]

W = 0.9931, p-value = 0.2852

```
> shapiro.test(resid(JS.roy1)[256:510])
```

Shapiro-Wilk normality test

```
data: resid(JS.roy1)[256:510]
W = 0.9395, p-value = 9.503e-09
```



Normal probability plots can be rendered for each level of the random effects. LME models assume that not only the within-cluster residuals are normally distributed, but that each level of the random effects are as well.

### 1.3 Standardization and Studentization

To alleviate the problem caused by inconstant variance, the residuals can be scaled (i.e. divided) by their standard deviations. This results in a ‘standardized residual’. A random variable is said to be standardized if the difference from its mean is scaled by its standard deviation. The residuals have mean zero but their variance is unknown, it depends on the true values of  $\theta$ . Standardization is not possible in practice. Because

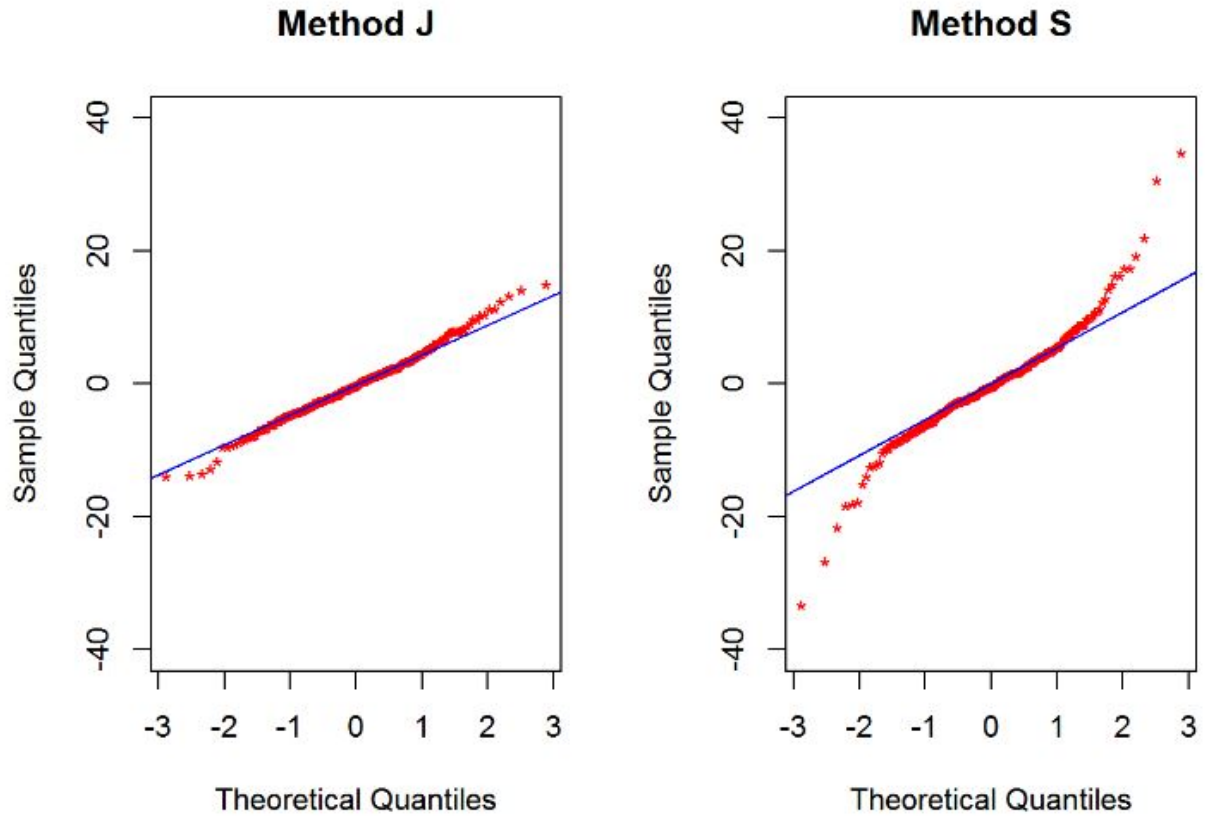


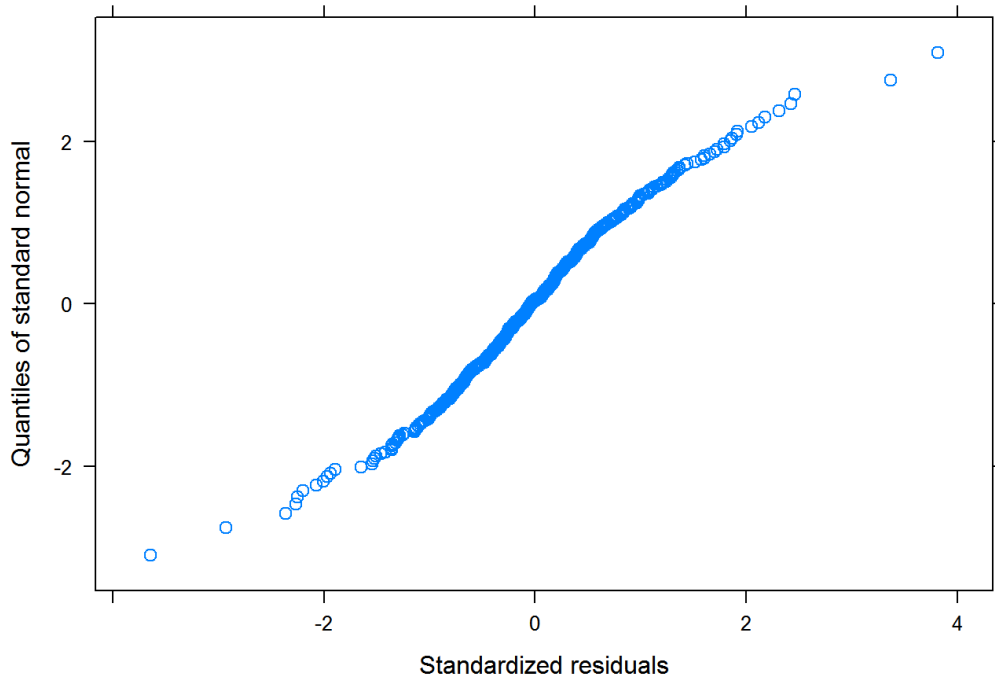
Figure 1.2.2:

true standard deviations are frequently unknown, one can instead divide a residual by the estimated standard deviation to obtain the ‘studentized residual’. If that estimate is independent of the  $i$ –th observation, the process is termed ‘external studentization’. This is usually accomplished by excluding the  $i$ –th observation when computing the estimate of its standard error. If the observation contributes to the standard error computation, the residual is said to be internally studentized. Externally studentized residual require iterative influence analysis or a profiled residuals variance (**CITE**).

The computation of internally studentized residuals relies on the diagonal entries of  $\mathbf{V}(\hat{\theta}) - \mathbf{Q}(\hat{\theta})$ , where  $\mathbf{Q}(\hat{\theta})$  is computed as

$$\mathbf{Q}(\hat{\theta}) = \mathbf{X}(\mathbf{X}'\mathbf{Q}(\hat{\theta})^{-1}\mathbf{X})\mathbf{X}^{-1}$$





### 1.3.1 Pearson Residual

Another possible scaled residual is the ‘Pearson residual’, whereby a residual is divided by the standard deviation of the dependent variable. The Pearson residual can be used when the variability of  $\hat{\beta}$  is disregarded in the underlying assumptions.

In this chapter, we will look at residual analysis and diagnostic tool for LME models, and discuss how they can be applied to the Method Comparison Problem. In classical linear models, a residual is the difference between an observed value and its estimated or predicted value. For LME models, the topic is more complicated, and still a matter of active research.

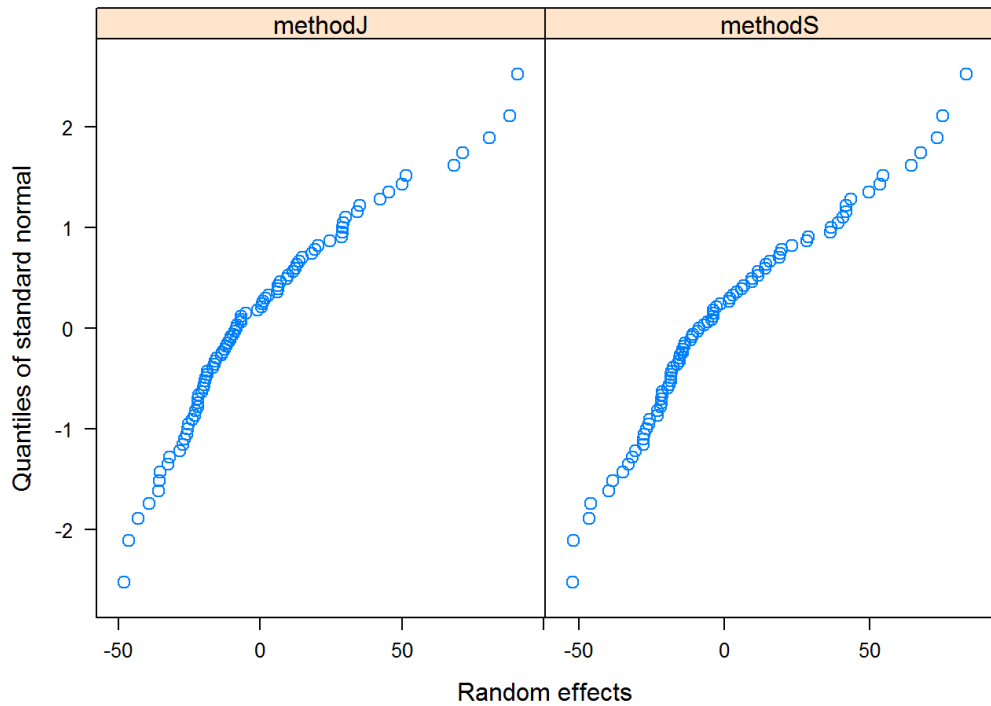


Figure 1.2.3:

## 1.4 Distinction From Linear Models (Schabenberger page 3)

The differences between perturbation and residual analysis in the linear model and the linear mixed model are connected to the important facts that  $b$

and  $b$  depend on the estimates of the covariance parameters, that  $b$

has the form of an (estimated) generalized least squares (GLS) estimator, and that  $b$  is a random vector. In a mixed model, you can consider the data in a conditional and an unconditional sense. If you imagine a particular realization of the random effects, then you are considering the conditional distribution  $Y$ — If you are interested in quantities averaged over all possible values of the random effects, then you are interested in  $Y$ ; this is called the marginal formulation. In a clinical trial, for example, you may be interested in drug efficacy for a particular patient. If random effects vary by patient,

that is a conditional problem. If you are interested in the drug efficacy in the population of all patients, you are using a marginal formulation. Correspondingly, there will be conditional and marginal residuals, for example. The estimates of the fixed effects

depend on the estimates of the covariance parameters. If you are interested in determining the influence of an observation on the analysis, you must determine whether this is influence on the fixed effects for a given value of the covariance parameters, influence on the covariance parameters, or influence on both. Mixed models are often used to analyze repeated measures and longitudinal data. The natural experimental or sampling unit in those studies is the entity that is repeatedly observed, rather than each individual repeated observation. For example, you may be analyzing monthly purchase records by customer. An influential data point is then not necessarily a single purchase. You are probably more interested in determining the influential customer. This requires that you can measure the influence of sets of observations on the analysis, not just influence of individual observations. The application of well-known concepts in model-data diagnostics to the mixed model can produce results that are at first counter-intuitive, since our understanding is steeped in the ordinary least squares (OLS) framework. As a consequence, we need to revisit these important concepts, ask whether they are portable to the mixed model, and gain new appreciation for their changed properties. An important example is the ostensibly simple concept of leverage.

## 1.5 Taxonomy of LME Residuals

Standard residual and influence diagnostics for linear models can be extended to linear mixed models. The dependence of fixed-effects solutions on the covariance parameter estimates has important ramifications in perturbation analysis. To gauge the full impact of a set of observations on the analysis, covariance parameters need to be updated, which requires refitting of the model.

The conditional (subject-specific) and marginal (population-averaged) formulations in the linear mixed model enable you to consider conditional residuals that use the

estimated BLUPs of the random effects, and marginal residuals which are deviations from the overall mean. Residuals using the BLUPs are useful to diagnose whether the random effects components in the model are specified correctly, marginal residuals are useful to diagnose the fixed-effects components.

Pinheiro and Bates (1994) describes three types of residual that describe the variabilities present in LME models

1. marginal residuals,  $\hat{\xi}$ , which predict marginal errors,
2. conditional residuals,  $\hat{\epsilon}$ , which predict conditional errors,
3. the BLUP,  $\mathbf{Z}\hat{\mathbf{b}}$ , that predicts random effects.

Each type of residual is useful to evaluates some assumption of the model.

According to Hilton & Minton (1995), a residual is considered pure for a specific type fo error if it depends only on the fixed components and on the error that it is supposed to predict. Residuals that depend on other types of error are known as ‘confounded errors’.

## 1.6 Residuals in LME Models : Marginal, Conditional and Scaled

### 1.6.1 Internally and Externally Studentized Residuals

The computation of internally studentized residuals relies on the diagonal values of  $\mathbf{V}(\hat{\boldsymbol{\theta}}) - \mathbf{Q}(\hat{\boldsymbol{\theta}})$  Externally studentized residuals require iterative influence analysis or a profiled residual variance.

Cook’s Distance

$$\boldsymbol{\delta}_{(U)} = \hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(U)}$$

A DFFIT measures the change in predicted values due to the removal of data points. (Belsey, Kuh and Welsch (1980))

$D(\beta) = \delta'_{(U)} \delta_{(U)} / \text{rank}(\mathbf{X})$  Cook's D can be calibrated according to a chi-square distribution with degrees of freedom equal to the rank of  $\mathbf{X}$  ?.

## Residuals

The raw residuals  $r_{mi}$  and  $r_{ci}$  are usually not well suited for these purposes.

- Conditional Residuals  $r_{ci}$
- Marginal Residuals  $r_{mi}$

## Chapter 2

# Residuals diagnostics in LME Models

Residuals are used to examine model assumptions and to detect outliers and potentially influential data point. In LME models, there are two types of residuals, marginal residuals and conditional residuals. A marginal residual is the difference between the observed data and the estimated marginal mean. A conditional residual is the difference between the observed data and the predicted value of the observation. In a model without random effects, both sets of residuals coincide. Schabenberger (2004) provides a useful summary. In a model without random effects, both sets of residuals coincide.

The definitions of both marginal residuals ( $r_m$ ) and conditional residuals ( $r_c$ ) follow from the definitions of marginal and conditional means in the LME model  $E[\mathbf{Y}] = \mathbf{X}\boldsymbol{\beta}$  and  $E[\mathbf{Y}|\mathbf{u}] = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$ , respectively.

A marginal residual is the difference between the observed data and the estimated (marginal) mean,

$$r_{mi} = y_i - x_0' \hat{b} = x_i^T \hat{\beta}$$

$$y - X\beta = Z\eta + \epsilon$$

A conditional residual is the difference between an observed value  $y_i$  and the con-

ditional predicted value  $\hat{y}_i$ ,

$$r_{ci} = y_i - x_i' \hat{b} - z_i' \hat{\gamma}$$

$$y - X\beta - Z\eta = \epsilon$$

Conditional residuals include contributions from both fixed and random effects, whereas marginal residuals include contribution from only fixed effects. Marginal residuals should have mean of zero, but may show grouping structure. Also they may not be homoscedastic. Marginal residuals are good for checking fixed effects. Plots of the elements of the marginal residual vector versus the explanatory variables in  $X$  can be used to check the linearity of  $\mathbf{y}$  in a similar manner to the residual plots used in linear models. Conditional residuals should have mean of zero with no grouping structure. They should be homoscedastic. Conditional residuals are useful for checking normality of outliers.

In linear mixed effects models, diagnostic techniques may consider ‘conditional’ residuals. A conditional residual is the difference between an observed value  $y_i$  and the conditional predicted value  $\hat{y}_i$ .

$$\epsilon_{\hat{y}_i} = y_i - \hat{y}_i = y_i - (X_i \hat{\beta} + Z_i \hat{b}_i)$$

However, using conditional residuals for diagnostics presents difficulties, as they tend to be correlated and their variances may be different for different subgroups, which can lead to erroneous conclusions.

## 2.0.2 Marginal Residuals

The marginal residuals are defined according to

$$\hat{\xi} = y - X\hat{\beta} = M^{-1}Qy.$$

### 2.0.3 Conditional Residuals

A conditional probability is the difference between the observed value and the predicted value of the dependent variable.

$$\hat{\epsilon}_i = y_i - X_i\hat{\beta} + Z_i\hat{b}_i$$

In general conditional residuals are not well suited for verifying model assumptions and detecting outliers. Even if the true model residuals are uncorrelated and have equal variance, conditional variances will tend to be correlated and their variances may be different for different subgroups of individuals (West et al., 2007).

## 2.1 Residual diagnostics

For classical linear models, residual diagnostics are typically implemented as a plot of the observed residuals and the predicted values. A visual inspection for the presence of trends inform the analyst on the validity of distributional assumptions, and to detect outliers and influential observations.

## 2.2 Conditional and Marginal Residuals

Conditional residuals include contributions from both fixed and random effects, whereas marginal residuals include contribution from only fixed effects.

Suppose the linear mixed-effects model lme has an  $n \times p$  fixed-effects design matrix  $\mathbf{X}$  and an  $n \times q$  random-effects design matrix  $\mathbf{Z}$ .

Also, suppose the p-by-1 estimated fixed-effects vector is  $\hat{\beta}$ , and the q-by-1 estimated best linear unbiased predictor (BLUP) vector of random effects is  $\hat{b}$ . The fitted conditional response is

$$\hat{y}_{Cond} = X\hat{\beta} + Z\hat{b}$$

and the fitted marginal response is



$$\hat{y}_{Mar} = X\hat{\beta}$$

residuals can return three types of residuals:

- raw,
- Pearson, and
- standardized.

For any type, you can compute the conditional or the marginal residuals. For example, the conditional raw residual is

$$r_{Cond} = y - X\hat{\beta} - Z\hat{b}$$

and the marginal raw residual is

$$r_{Mar} = y - X\hat{\beta}$$

## 2.3 Pearson and Deviance Residuals

The **deviance residual** is the measure of deviance contributed from each observation and is given by

$$r_{Di} = \text{sign}(r_i)\sqrt{d_i}$$

where  $d_i$  is the individual deviance contribution. The deviance residuals can be used to check the model fit at each observation for generalized linear models.

The standardized and studentized deviance residuals are

$$r_{Dsi} = \frac{r_{Di}}{\sqrt{\hat{\phi}(1 - h_i)}}$$

$$r_{Dti} = \frac{r_{Di}}{\sqrt{\hat{\phi}_{(i)}(1 - h_i)}}$$

# Chapter 3

## Galecki

### 3.1 Model Diagnostics for Roy's Models

Further to previous work, this section revisits case-deletion and residual diagnostics, and explores how approaches devised by Galecki & Burzykowski (2013) can be used to appraise Roy's model. These authors specifically look at Cook's Distances and Likelihood Distances. For the Roy Model, Cook's Distances may also be generated using the *predictmeans*

As the model is structurally different from the models discussed in the earlier sections, Residual analysis will be briefly revisited.

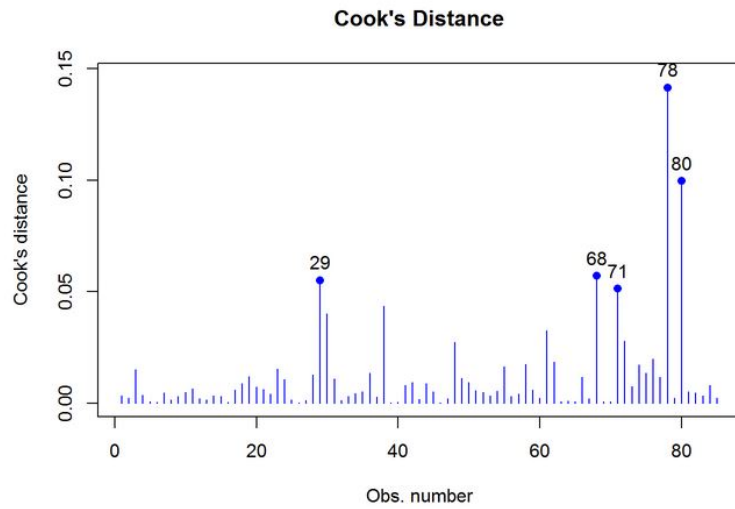


Figure 3.1.1:

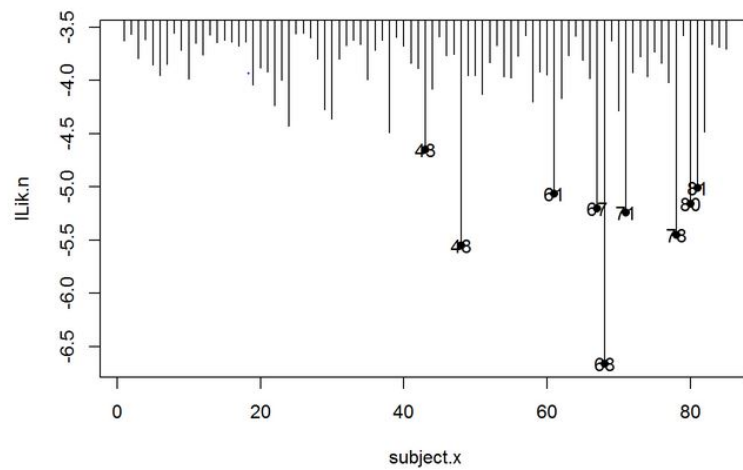


Figure 3.1.2:

## 3.2 Case Deletion Diagnostics for the Variance Ratios

Schabenberger advises on the use of deletion diagnostics for variance components of an LME model. Taking the core principals of his methods, and applying them to the Method Comparison problem, case deletion diagnostics are used on the variance

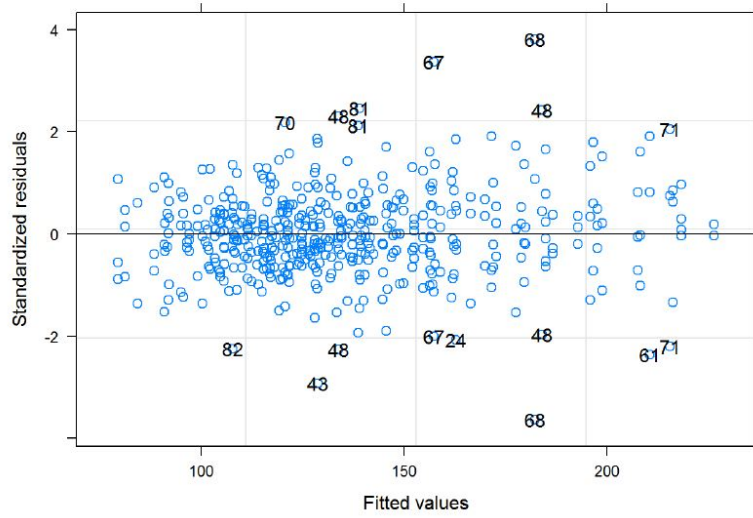


Figure 3.1.3:

components of the Roy model., specifically the ratio of between subject variances and the within subject covariances respectively.

$$\text{BSVR} = \frac{\sigma_2^2}{\sigma_2^2} \quad \text{WSVR} = \frac{d_2^2}{d_2^2}$$

These variance ratios are re-computed for each case removed, and may be analysed separately or jointly for outliers.

The Grubbs' Test for Outliers is a commonly used technique for assessing outlier in a univariate data set. As there may be several outliers (i.e. influential cases) present, the Grubbs test is not useful. However, as identification is all that is required, a simple outlier detection technique, such as using to Tukey's specification for boxplots (i.e. greater than  $Q_3 + 1.5IQR$  or less than  $Q_1 - 1.5IQR$ ), will suffice. Ranking the absolute values of the standardized scores can also be used to identify influential cases, even if the data is not normally distributed.

Bivariate Analyses may be applied jointly to the both sets of data sets, e.g. Mahalanobis distances. The Mahalanobis distance, while not an intuitive measure in the context of the data, can be used to rank highly influential cases.

## Leave-One-Out Diagnostics with lmeU

Galecki et al discuss the matter of LME influence diagnostics in their book, although not into great detail.

The command `lmeU` fits a model with a particular subject removed. The identifier of the subject to be removed is passed as the only argument

A plot of the per-observation diagnostics individual subject log-likelihood contributions can be rendered.

### Likelihood Displacement

#### 3.2.1 Likelihood Distances

The likelihood distance is a global summary measure that expresses the joint influence of the subsets of observations,  $U$ , on all parameters in  $\phi$  that were subject to updating. For classical linear models, the implementation of influence analysis is straightforward. ? points out the likelihood distance gives the amount by which the log-likelihood of the model fitted from the full data changes if one were to estimate the model from a reduced-data estimates. Importantly  $LD(\psi_{(U)})$  is not the log-likelihood obtained by fitting the model to the reduced data set. It is obtained by evaluating the likelihood function based on the full data set (containing all  $n$  observations) at the reduced-data estimates.

$$LD((U)) = 2[l(\hat{\phi}) - l_{\hat{\phi}_{\omega}}]$$

$$RLD((U)) = 2[l_R(\hat{\phi}) - l_R(\hat{\phi})_{\omega}]$$

An overall influence statistic measures the change in the objective function being minimized. For example, in OLS regression, the residual sums of squares serves that purpose. In linear mixed models fit by maximum likelihood (ML) or restricted maximum likelihood (REML), an overall influence measure is the likelihood distance (?). In LME models, fitted by either ML or REML, an important overall influence measure

is the likelihood distance (?). The procedure requires the calculation of the full data estimates  $\hat{\psi}$  and estimates based on the reduced data set  $\hat{\psi}_{(U)}$ . The likelihood distance is given by determining

$$LD_{(U)} = 2\{l(\hat{\psi}) - l(\hat{\psi}_{(U)})\} \quad (3.1)$$

$$RLD_{(U)} = 2\{l_R(\hat{\psi}) - l_R(\hat{\psi}_{(U)})\} \quad (3.2)$$

For classical linear models, the implementation of influence analysis is straightforward. However, for LME models, the problem is more complex. Update formulas for the fixed effects are available only when the covariance parameters are assumed to be known. A measure of total influence requires updates of all model parameters. This can only be achieved in general is by omitting observations or cases, then refitting the model. This is a very simplistic approach, and computationally expensive.

West et al. (2007) examines a group of methods that examine various aspects of influence diagnostics for LME models. For overall influence, the most common approaches are the *likelihood distance* and the *restricted likelihood distance*.

### The logLik Function

`logLik.lme` returns the log-likelihood value of the linear mixed-effects model represented by object evaluated at the estimated coefficients. It is also possible to determine the restricted log-likelihood, if relevant, using this function. For the Blood Data Example, the loglikelihood of the JS.roy1 model can be computed as follows.

```
> logLik(JS.roy1)
'log Lik.' -2030.736 (df=8)
```

### 3.3 Likelihood Distance

The likelihood distance is a global, summary measure, expressing the joint influence of the observations in the set  $U$  on all parameters in  $\phi$  that were subject to updating.

The likelihood distance gives the amount by which the log-likelihood of the full data changes if one were to evaluate it at the reduced-data estimates. The important point is that  $l(\psi_U)$  is not the log-likelihood obtained by fitting the model to the reduced data set.

It is obtained by evaluating the likelihood function based on the full data set (containing all  $n$  observations) at the reduced-data estimates.

The likelihood distance is a global, summary measure, expressing the joint influence of the observations in the set  $U$  on all parameters in  $\psi$  that were subject to updating.

# Chapter 4

## Cook's Distance

### 4.1 Cook's Distance for LMEs

Diagnostic methods for fixed effects are generally analogues of methods used in classical linear models. Diagnostic methods for variance components are based on ‘one-step’ methods. Cook (1986) gives a completely general method for assessing the influence of local departures from assumptions in statistical models. Cook (1977) greatly expanded the study of residuals and influence measures. Cook’s key observation was the effects of deleting each observation in turn could be computed without undue additional computational expense. Consequently deletion diagnostics have become an integral part of assessing linear models.

Cook’s Distance ( $D_i$ ) is a well known diagnostic technique used in classical linear models, that functions as an overall measure of the combined impact of the  $i$ th case on all estimated regression coefficients. It uses the same structure for measuring the combined impact of the differences in the estimated regression coefficients when the  $i$ th case is deleted.  $D_{(i)}$  can be calculated without fitting a new regression coefficient each time an observation is deleted. Importantly,  $D_{(i)}$  can be calculated without fitting a new regression coefficient each time an observation is deleted.

Cook’s  $D$  statistics (i.e. colloquially Cook’s Distance) is a measure of the influence



of observations in subset  $U$  on a vector of parameter estimates (Cook, 1977).

$$\delta_{(U)} = \hat{\beta} - \hat{\beta}_{(U)}$$

If  $V$  is known, Cook's  $D$  can be calibrated according to a chi-square distribution with degrees of freedom equal to the rank of  $\mathbf{X}$  (?). (**State  $V$  and  $X$** )

Cook's distance can be used in several ways: to indicate data points that are particularly worth checking for validity; to indicate regions of the design space where it would be good to be able to obtain more data points. Large values for Cook's distance indicate observations for special attention.

Use of threshold value for Cook's Distance is discouraged (**cite: JohnFox**). However, informal heuristics do exist for OLS models; Observations for which Cook's distance is higher than 1 are to be considered as influential. Alternatively there is an informal threshold of  $4/N$  or  $4/(Nk+1)$ , where  $N$  is the number of observations and  $k$  the number of explanatory variables.

**cite: JohnFox** advises the use of diagnostic plotting and to examine in closer details the points with "*values of  $D$  that are substantially larger than the rest*", and that thresholds should just be used to enhance graphical displays.

The effect on the precision of estimates is separate from the effect on the point estimates. Data points that have a small Cook's distance, for example, can still greatly affect hypothesis tests and confidence intervals, if their influence on the precision of the estimates is large.

? would later adapt the Cook's Distance measure for the analysis of LME models. For LME models, two formulations exist; a Cook's distance that examines the change in fixed parameter estimates, and another that examines the change in random effects parameter estimates. The outcome of either Cook's distance is a scaled change in either  $\beta$  or  $\theta$ .

For LME models, Cook's distance can be extended to model influence diagnostics by defining.

$$CD_{\beta i} = \frac{(\hat{\beta} - \hat{\beta}_{[i]})^T (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X}) (\hat{\beta} - \hat{\beta}_{[i]})}{p}$$

It is also desirable to measure the influence of the case deletions on the covariance matrix of  $\hat{\beta}$ .

Schabenberger (2004) discusses a taxonomy of Cook's distance when applied to LME models.

- For variance components  $\gamma$ :  $CD(\gamma)_i$ ,
- For fixed effect parameters  $\beta$ :  $CD(\beta)_i$ ,
- For random effect parameters  $\mathbf{u}$ :  $CD(u)_i$ ,
- For linear functions of  $\beta$ :  $CD(\psi)_i$

It is also desirable to measure the influence of the case deletions on the covariance matrix of  $\hat{\beta}$ .

For fixed effects parameter estimates in LME models, the Cook's distance can be extended to measure influence on these fixed effects.

$$CD_i(\beta) = \frac{(c_{ii} - r_{ii}) \times t_i^2}{r_{ii} \times p}$$

For fixed effects parameter estimates in LME models, the Cook's distance can be extended to measure influence on these fixed effects.

$$CD_i(\beta) = \frac{(c_{ii} - r_{ii}) \times t_i^2}{r_{ii} \times p}$$

For random effect estimates, the Cook's distance is

$$CD_i(b) = g'_{(i)} (I_r + \text{var}(\hat{b}) D)^{-2} \text{var}(\hat{b}) g_{(i)}.$$

Large values for Cook's distance indicate observations for special attention.

A large value for  $CD(u)_i$  indicates that the  $i$ -th observation is influential in predicting random effects. For linear functions,  $CD(\psi)_i$  does not have to be calculated unless  $CD(\beta)_i$  is large.

### 4.1.1 Cook's Distance

- For variance components  $\gamma$ :  $CD(\gamma)_i$ ,
- For fixed effect parameters  $\beta$ :  $CD(\beta)_i$ ,
- For random effect parameters  $\mathbf{u}$ :  $CD(u)_i$ ,
- For linear functions of  $\hat{\beta}$ :  $CD(\psi)_i$

Diagnostic tool for variance components

$$C_{\theta i} = ((\hat{\theta})_{[i]} - \hat{\theta})^T \text{cov}(\hat{\theta})^{-1} ((\hat{\theta})_{[i]} - \hat{\theta})$$

**Random Effects** A large value for  $CD(u)_i$  indicates that the  $i$ -th observation is influential in predicting random effects.

**linear functions**  $CD(\psi)_i$  does not have to be calculated unless  $CD(\beta)_i$  is large.

#### Random Effects

A large value for  $CD(u)_i$  indicates that the  $i$ -th observation is influential in predicting random effects.

#### linear functions

$CD(\psi)_i$  does not have to be calculated unless  $CD(\beta)_i$  is large.

*Cook (1977)* greatly expanded the study of residuals and influence measures. Cook's key observation was the effects of deleting each observation in turn could be computed without undue additional computational expense. Consequently deletion diagnostics have become an integral part of assessing linear models.

Cook (1986) gave a completely general method for assessing influence of local departures from assumptions in statistical models.

## 4.2 Cook's Distance for LMEs

### 4.2.1 Cook's Distance

For linear mixed effects models, Cook's distance can be extended to model influence diagnostics by defining.

$$C_{\beta_i} = \frac{(\hat{\beta} - \hat{\beta}_{[i]})^T (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X}) (\hat{\beta} - \hat{\beta}_{[i]})}{p}$$

It is also desirable to measure the influence of the case deletions on the covariance matrix of  $\hat{\beta}$ .

### 4.2.2 Cook's Distance

Cook's  $D$  statistics (i.e. colloquially Cook's Distance) is a measure of the influence of observations in subset  $U$  on a vector of parameter estimates (Cook, 1977).

$$\delta_{(U)} = \hat{\beta} - \hat{\beta}_{(U)}$$

If  $V$  is known, Cook's  $D$  can be calibrated according to a chi-square distribution with degrees of freedom equal to the rank of  $\mathbf{X}$  (?).

For LME models, Cook's distance can be extended to model influence diagnostics by defining.

It is also desirable to measure the influence of the case deletions on the covariance matrix of  $\hat{\beta}$ .

# Bibliography

- Cook, R. (1977). Detection of influential observations in linear regression. *Technometrics* 19, 15–18.
- Cook, R. (1986). Assessment of local influence. *Journal of the Royal Statistical Society. Series B (Methodological)* 48(2), 133–169.
- Pinheiro, J. and D. Bates (1994). *Mixed Effects Models in S and S plus* (2nd ed.). Reading, Massachusetts: Springer.
- Schabenberger, O. (2004). Mixed model influence diagnostics. 18929.
- West, B., K. Welch, and A. Galecki (2007). *Linear Mixed Models: a Practical Guide Using Statistical Software*. Chapman and Hall CRC.