JIMMA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

COLLEGE OF NATURAL SCIENCE

DEPARTMENT OF STATISTICS

MODELING EVOLUTION OF UPPER ARM CIRCUMFERENCE
OF INFANT'S USING LINEAR MIXED EFFECTS APPROACH

By: Tadele Akeba

Advisor: Wondwosen Kassahun(Phd scholar)

Co. Advisor: Zenebe fikrie   (Msc)

September, 2011

i

MODELING EVOLUTION OF UPPER ARM CIRCUMFERENCE
OF INFANT'S USING LINEAR MIXED EFFECTS APPROACH

By: Tadele Akeba

Advisor: Wondwosen Kassahun(Phd student)

Co. Advisor: Zenebe fikrie

A THESIS SUBMITTED TO SCHOOL OF GRADUATE STUDIES,
JIMMA UNIVERSITY IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN
BIOSTATISTICS

Declaration

I, the undersigned, declare that this evaluation thesis is my original work, has not been presented for a degree in this or any other university and that all sources of materials used for the thesis have been fully acknowledged.

_____	_____	_____
Name of investigator	Signature	Date

_____	_____	_____
Name of first advisor	Signature	Date

_____	_____	_____
Name of second advisor	Signature	Date

_____	_____	_____
Name of internal examiner	Signature	Date

_____	_____	_____
Name of external examiner	Signature	Date

_____	_____	_____
Department head	Signature	Date

TABLE OF CONTENTS

Content   pages

Chapter 4

Chapter 5

List of tables

List of Figure

List of abbreviations:

lme **-** linear mixed effects

LMEM- linear mixed effects model

ML- maximum likelihood

REML- restricted maximum likelihood

MME- mixed model equation

DIC- deviance information criteria

AC- arm circumference

AIC- Akakie's information criteria

BIC- Bayesian information criteria

pdBlocked - block-diagonal

pdCompSymm - compound-Symmetry structure

pdDiag - diagonal

pdIdent - multiple of an identity

pdSymm - general positive definite matrix

Acknowledgements

Abstract

Nowadays, quite a lot of methodology has been developed for the analysis of longitudinal studies, stemming from clinical trials, epidemiology, and other studies in humans. For example, hierarchical models are becoming ever more frequently. Such hierarchical models are standard in the analysis of longitudinal data, too to account for the correlation steaming from the repeated measures nature. This study will be dedicated to model models for longitudinal continuous, firmly rooted in hierarchical models such as the linear mixed model. One finds, coupled with methodological development, also the availability of standard software tools, including SAS, Stata, SPlus, R, etc. The Bayesian implementation of the models will also be explored using the freely available software WinBugs. The two approaches will then be applied on data set from the Jimma Infants longitudinal growth study.

The result demonstrated that the ML estimates of the random-effects standard deviations are smaller than the corresponding REML estimates which is different result from the Bayesian. The estimated within group residual standard deviations are identical. In general, the fixed-effects estimates obtained using ML, REML and Bayesian techniques are almost similar. The mean evolution of the upper arm circumference of infant for boys and girls is not different. For infants given supplementary food and without supplementary food their mean evolution is not different.

The linear mixed effect model estimate of the fixed effect obtained using likelihood and Bayesian techniques are almost similar but with different random effect standard deviations.

CHAPTER ONE

1. INTRODUCTION

1.1. Background of the study

Many longitudinal studies are designed to investigate changes over time in a characteristic which is measured repeatedly for each study participant. Multiple measurements are obtained from each individual under investigation at different times and possibly under changing experimental conditions, and there is considerable variation among individuals in the number and timing of observation (liard and Ware, 1982).

Statistics have often analyzed data of this form using some variant of a two stage model (harville, 1977). In this formulation the probability distribution for the multiple measurements have the same form for each individual, but the parameters of that distribution vary over individuals. The distribution of this parameters, or 'random effects', in the population constitute the second stage of the model

Such two-stage models have several desirable features. There is no requirement for balance in the data. They allow explicit modeling and analysis of between- and within-individual variation. Often the individual parameters have natural interpretation which is relevant to the goals of the study.

In this study infant considered to children between ages of one month to twelve month (under one year) on Jimma infants survival differential longitudinal growth study to establish risk factors affecting infant survival that contributes most to Childs early survival.

An infant or baby is the very young offspring of humans. A newborn is an infant who is within hours, days, or up to a few weeks from birth. In medical contexts, newborn or neonate (from Latin, neonatus, newborn) refers to an infant in the first 28 days of life (from birth up to 4 weeks after birth, less than a month old). The term "newborn" includes premature infants, post mature infants and full term newborns. The term infant is derived from the Latin word *infans*, meaning "unable to speak" or "speechless." It is typically applied to children between the ages of 1 month and 12 months; however, definitions vary between birth and three years of age. "Infant" is also a legal term referring to any child under the age of legal adulthood.

1.2 Statement of the problem

This study is on Jimma Infant Differential Longitudinal Growth to establish risk factors affecting infant growth that contribute most to the child's early survival by applying linear mixed effects modelto the upper AC of infants longitudinal continuous data. This longitudinal study is designed to investigate changes over time in a characteristic which is measured repeatedly for each study participant. Repeated measurements are obtained from each individual at different times to see if there is considerable variation among individuals in the number and timing of observation. And this study tries to answer,

- The basic Linear mixed effects model provides an adequate model for many different types of grouped data observed in particular , however there are many applications involving grouped data for which the within-group errors are hetroscedastic (i.e have unequal variances) or are correlated or are both hetroscedastic and correlated. And also, can we extend the basic Linear mixed effects model to allow for hetroscedastic, correlated within-group errors?
- Does the mean evolution of upper arm circumference of infant changes over time?
- Does the mean evolution of upper arm circumference of infants differ for gender as well as supplementary food behavior?
- Linear mixed effects model is based on thinking about individual behavior first. How is the change in the population represented?

1.3 Significance of the study

There are many factors that affect the growth of infants and this study designed to investigate factors that lead infants, to early survival by considering upper arm circumference of infants as a growth indicator for child's less than one year. The data are of repeated measurements, taken from child's starting from birth until the age of one year within two month interval. Hence this study opts to model arm circumference as a function some covariates like gender, age, etc. both the likelihood and Bayesian methodology are employed. It is expected that this study will be very useful in indicating the possible covariates for arm circumference. Further, it would helpful for further research in a difference or similar setting.

1.4. Objective of the study

1.4.1 General objective

The general objectives of this study was modeling evolution of upper arm circumference of infant's longitudinal study using linear mixed effects model to establish risk factors affecting infant early survival.

1.4.2 Specific objectives

- To present the linear mixed effects model for a longitudinal Gaussian data and to apply on the evolution of upper arm circumference of infant's.

- To explore and apply the Bayesian implementation of the Linear mixed effects models.

- To compare the Likelihood and Bayesian techniques as alternative approaches for a Gaussian longitudinal response.

CHAPTER TWO

2. LITRETURE REVIEW

There are many factors affecting the growth of infants. The Jimma Infant Survival Differential Longitudinal Growth Study is an Ethiopian study, set up to establish risk factors affecting infant survival and to investigate socio-economic, maternal, and infant-rearing factors that contribute most to the child's early survival. Several studies have been done on infant growth using arm circumference as growth indicator. A cross-sectional study was performed among term live birth newborns, from June 1997 to August 1999, at Hospital Maternidade Leonor Mendes de Barros, São Paulo, Brazil, a public maternity hospital within the healthcare system that serves a low-income population and is used as a reference center for high-risk pregnancies. The study group consisted of newborns from single pregnancies, with gestational ages of between 37 weeks and 41 weeks and 6 days, as estimated by Capurro's method. These newborns were examined by the main author within their first 48 hours of life. Only newborns whose mothers agreed to participate in the study were included. The study comprised 131 newborns: 66 males and 65 females. The average gestational age was 39 complete weeks, ranging from a minimum of 37 weeks to a maximum of 41 weeks. The normality test (Kolmogorov-Smirnov) showed that both variables studied followed the normal distribution. The Student t test, used in order to identify possible differences between sexes and the association of age with the dependent variable upper arm circumference, showed no significant differences for any of the evaluated parameters. And also there is direct association of upper arm circumference with age and other variables of the study.

Other studies which are done by WHO Child Growth Standards arm circumference-for-age and sex methods and development. There were a total of 10 770 arm circumference observations for boys and girls and The measurement of upper arm were followed starting from 3 to 60 months (de Onis et al., 2004b). The data of the longitudinal and cross-sectional samples were merged without any adjustments and a single model was fitted to generate one continuous set of curves constituting each sex-specific standard. The Box-Cox-power-exponential (BCPE) method (Rigby and Stasinopoulos, 2004), with curve smoothing by cubic splines was selected for constructing the WHO child growth curves. And the result shows that there is no significant difference between boys and girls. And also the arm circumference of both sexes increases with age.

Many longitudinal studies are designed to investigate changes over time in a characteristic which is measured repeatedly for each study participant. Multiple measurements are obtained from each individual, at different times and possibly under changing experimental conditions. Often, we cannot fully control the circumstances under which the measurements are taken, and there may be considerable variation among individuals in the number and timing of observations. The resulting unbalanced data sets are typically not amenable to analysis using a general multivariate model with unrestricted covariance structure (Nan M. liard; James H.Ware, 1982).

Random coefficient models, where we develop an overall statistical model by thinking first about individual trajectories in a "subject-specific" fashion, are a special case of a more general model framework based on the same perspective. This model framework, known popularly as the linear mixed effects model, is still based on thinking about individual behavior first, of course. However, the possibilities for how this is represented, and how the variation in the population is represented, are broadened. The result is a very flexible and rich set of models for characterizing repeated measurement data. One advantage of random coefficient models is that the model naturally represents Individual trajectories in a formal way, so that questions of interest about individual behavior may be considered. In modeling longitudinal continuous data the model emphasize characterizing the mean vectors, how the mean response change over with time and depends on other factors. And taking into account important source of variation (i.e. characterizing the nature of the random source deviations).

CHAPTER THREE
3. MATERIALS AND METHODS

3.1 Study Area

The study was conducted from children born in south western Ethiopia. And the children were from rural, urban and semi-urban areas. Of special interest in this study is focusing on infants in urban setting that means Jimma town.

3.2 Study Population

The study was conducted on children (infants) and they were visited every two months starting from birth until the age of one year (for twelve month).

3.3 Study Design

A longitudinal study design was used using longitudinal continuous data to study factors affecting infant's early survival. Inference was made using linear mixed models, we describe a unified approach to inference this models. Both likelihood (maximum likelihood (ML) and restricted maximum likelihood (REML)**)** and Bayesian method of estimation was discussed.

3.4 Description of the data

The Jimma Infant Survival Differential Longitudinal Growth Study is an Ethiopian study, set up to establish risk factors affecting infant survival and to investigate socio-economic, maternal, and infant-rearing factors that contribute most to the child's early survival. Children were examined for their first year growth characteristics. At baseline, there are a total of 7969 infants whereby 4317, 1494, and 2158 are from rural, urban, and semi-urban areas, respectively. The children were visited every two months starting from birth until the age of one year. Data were collected on demographic, behavioral and environmental factors on infants. Of special interest is the assessment of the upper arm circumference of the infants as an indicator for growth focusing on infants in urban settings i.e Jimma town.

3.5 Software to be used

In this study data were analyzed using R-software for the purpose of analyzing linear mixed models using different libraries in R packages and for Bayesian method of analysis we mainly used was open bug's software. And qq-plot of residual, histogram of residual, and the scatter plot of standardized residual versus predicted values to check whether the stated normality assumption is hold.

3.6 Variables in the study

The variables considered in this study are:

1. **Dependent Variable:**
   - AC (arm circumference)**:**

2. **Independent Variables:**
   - Age in month
   - Sex (coded as MALE=1, FEMALE=0)
   - Powdermil: whether the infant is given powder milk or no t(coded as Yes=1, No=0).
   - Suppfood: whether the infant is given supplementary food (coded as Yes=1, No=0).

3.7 Linear mixed effects model

3.7.1 The concept of random and fixed factors

Searle et al. (1995) states that variance component estimation originated from estimating the error variance in the analysis of variance by equating the error mean square to its expected value. This procedure was then extended to random effects model for balanced data and then for unbalanced data. The beginning of variance component has revolved around the one- way random effects classification model as:

$y_{ij} = \mu + \beta_i + e_{ij}$, where i = 1,…,a, the β_i's and e_{ij} ' s are assumed to be random variables and **μ** is the mean, var(β_i) = δ^2_β and var(e_{ij}) = δ^2_e with all covariance equal to zero. The variances associated with the random effects are called variance components.

The flow chart below is taken from Searle et al.(2006) it helps to decide on the appropriate model to be fitted(as an adequate representation of the data) using mixed model.

```
          _____
               ↓                        ↓
              No                       Yes
               ↓                        ↓
   Treat factors as fixed      treat factors as random
                                        ↓
        _____Where does interest lie?_____
               ↓                                      ↓
     It is only in the distribution of       in both the distribution and
         the   random effects                   the realized value of  the
               ↓                                     random effects
   Estimate the variance of the random                 ↓
            effects                          Estimate the variance of the
                                             random effects and calculate
                                             the best linear unbiased
                                             predictor of the realized value
                                             of the random effects.
```

Here we need to describe and define concepts of fixed and random effects that are applied in mixed effects model. A factor in a model is fixed if its levels are selected by a non random process or consists of entire population of all possible levels. A model is termed as fixed effect model if all the factors in the structure are fixed effects. A factor in a model is random if its levels consist of a random sample from a population of all possible levels. A model is termed as random effect model if all the factors in the structure are random. Therefore in modeling, if the level of independent variable that occurs in the study is considered to be the only level of interest, each level has a fixed effect on the response variable and the independent variable is known as a fixed effect factor. However if the levels are considered to be random sample from the population of possible levels (and/or the levels of the effects are assumed to be randomly selected from infinite population of possible levels), each level will have a random effect on the response variable and the variable is known as a random effect factor.

A model which contains both fixed and random effect factor with error terms is called mixed model or mixed effect model. And if the relationship among these factors, the error term and the response variable is linear we call such a model as linear mixed model. Hence, mixed effects model is a generalization of the standard linear model that enables the analysis of the data generated from several source of variation instead of just one.

3.8 The linear mixed model

The linear mixed effects model (LMM) introduced by Henderson (1975) is given by:

$$y_i = x_i \beta + z_i b_i + e_i \text{, where}$$

- x_i is (n_i x p) design matrix that characterize the systematic part of the response depending on covariate and time (fixed effects).
- β is (P x 1) vector of fixed effects parameters.
- z_i is (n_i x k) design matrix for random effects.
- b_i is (k x 1) vector of random effects , among unit variation.
- e_i is (n_i x 1) vector of within-unit variation.

Assumptions of the model

The theoretical assumptions about the covariance structure of the random vector **b** and e_i are as follows:

$$\mathrm{var}(e_i) = \delta^2_e \, \mathbf{I} = R_i$$

$$\mathbf{var(b)} = \mathbf{D} = \begin{pmatrix} \delta_1^{\,2} I_{q1} & \cdots & . \\ \vdots & \ddots & \vdots \\ . & \cdots & \delta_r^{\,2} I_{qr} \end{pmatrix}$$

we note that $\mathbf{var}(b_i) = \delta^2_i \mathbf{I_{qi}}$, with $q_1 + \ldots + q_r = n$, $\mathbf{cov}(b_i, b_j) = \mathbf{0}$, $i \neq j$, $\mathbf{cov(b,e)=0}$, $\mathbf{cov(x,e)=0}$. As a result, $\mathbf{var(y)} = zDz' + \delta^2_e I = zDz' + R = \varepsilon$. The assumptions for b_i and e_i are $b_i \sim N(0.D)$, $e_i \sim N(0. R_i)$.

3.9 Estimation

The model fitting consists of two parts, in estimating variance parameters, fixed effects, random effects. Estimation in Linear mixed effects model made by using restricted maximum likelihood (REML) and maximum likelihood (ML) and it is based on the maximum likelihood estimation approach which requires the assumption that the distribution of the response is normal.

3.9.1 Maximum likelihood method

The crucial requirement in estimating variance components of a set of data using maximum likelihood technique is the assumption of underlying probability distribution for the data. Maximum likelihood estimates are then, by definition, the parameter values for which the likelihood is maximized (Searle et al. (2006)). Gives the estimates of variance components,

$$\delta^{2(m+1)}{}_e = \frac{\left(y'\left(y - x\hat{\beta}^m - z\hat{b}^m\right)\right)}{N}$$

$$\delta^{2(m+1)}{}_i = \frac{\hat{b}^{t(m)}\hat{b}^{(m)} + \delta_i^{2(m)}tr(W_{ii}^{(m)})}{q_i} = \frac{\hat{b}^{t(m)}\hat{b}^{(m)}}{q_{i-tr(W_{ii})^{(m)}}}$$

where $W = (I + Z'R^{-1}ZD)^{-1} = W_{ij}$, $i,j = 1,2,..,r$ and D has q_i diagonal elements of δ^2_i. The superscript m is the number of iterations used to arrive at these estimates. tr represent trace operator.

3.9.1.1 Estimation of the parameters using maximum likelihood

After criterion for convergence is fulfilled the next step is estimation of the parameters for the mixed model. A key assumption in the analysis is that b and e is normally distributed with mean zero and variance D and R respectively.

$$E\binom{b}{e} = \binom{0}{0} \quad \text{and var}\binom{b}{e} = \begin{pmatrix} D & 0 \\ 0 & R \end{pmatrix}$$

Estimating D and R in LMM . estimation of parameters in the mixed model is more difficult. In many situations, the best approach is to use likelihood based methods, exploiting the assumption that b and e are normally distributed (Jennrich and schluchter, 1986). Using calculus, it is possible to reduce this maximization problem to one over only the parameters in D and R. the corresponding log likelihood functions are as follows:

$$\text{ML:l(D,R)} = \frac{-n}{2}\log(2\pi) - \frac{1}{2}\log|V| - \frac{1}{2}r'V^{-1}r$$

Where the vector $r = y - x(x'V^{-1}x)^{-1}x'V^{-1}y$ and $p = \text{rank}(x)$.

Mixed model actually minimizes -2 times these functions using a ridge-stabilized Newton- Raphson algorithm. The second derivative matrix (H) of the objective function evaluated at the optima is available upon completion and the asymptotic theory of maximum likelihood (Serfling ,1980) shows that $2H^{-1}$ is an asymptotic variance – covariance matrix of the estimated parameters of D and R. Mixed model profiles the residual variance out of log likelihood whenever it appears reasonable to do so. Therefore in LMM analysis, the ML provide estimates of D and R. which are denoted by $\hat{D}$ and $\hat{R}$, respectively.

3.9.2 Restricted maximum likelihood method (REML)

A major drawback of maximum likelihood estimation in Linear mixed effects model is that it ignores the loss in degrees of freedom due to fitting fixed effects. Fortunately, a modified maximum likelihood procedure, the so-called restricted maximum likelihood as described by Patterson and Thompson (1971), overcomes these problems by maximizing only part of the likelihood which is independent of fixed effects. REML is often interpreted as a technique that is based on linear combination of y. not forgetting that this linear combination do not contain any fixed effects. Not surprisingly these linear combinations of values not containing any fixed effects turn out to be equivalent to the residuals obtained after we fit the model of fixed effects.

Consider the set of values c'y where matrices c' of dimensions (r x n) can be chosen to satisfy c'y = c'xB+c'Zb such that no term in β is contained. That is from c'xβ = 0 follows that c'x=0.

$$P= V^{-1} - V^{-1}X (X'V^{-1}X)^{-1} X' V^{-1}$$

Where r= rank(x) and v = var(y)

Searle et al. (2006) give the following solutions for the estimate the variance components of the random factors form REML equations.

$$\delta^{2(m+1)} = \frac{\left(y'\left(y - x\hat{\beta}^m - z\hat{b}^m\right)\right)}{N - r}$$

$$\delta^{2(m+1)}_i = \frac{\hat{b}^{t(m)}\hat{b}^{(m)} + \delta_i^{2(m)} tr(T_{ii}^{(m)})}{q_i}$$

$$= \frac{\hat{b}^{t(m)}\hat{b}^{(m)}}{q_i - tr(T_{ii})^{(m)}}$$

where S= $R^{-1} - R^{-1}X(X'R^{-1}X)^{-1}X'R^{-1}$, T=$(I+Z'SZG)^{-1}$ =T_{ij} , i,j=1,2,....r and G has q_i diagonal elements of δ^2_i. the superscript m denotes the number of iteration needed to obtain the estimate.

3.9.2.1 Estimation of the parameters using Restricted maximum likelihood method

After criterion for convergence is fulfilled the next step is estimation of the parameters for the mixed model. A key assumption in the analysis is that b and e is normally distributed with mean zero and variance D and R respectively.

$$E\binom{b}{e} = \binom{0}{0} \quad \text{and var}\binom{b}{e} = \begin{pmatrix} D & 0 \\ 0 & R \end{pmatrix}$$

Estimating D and R in LMM. Estimation of parameters in the mixed model is more difficult. In many situations, the best approach is to use likelihood based methods, exploiting the assumption that b and e are normally distributed (Jennrich and schluchter, 1986). Using calculus, it is possible to reduce this maximization problem to one over only the parameters in D and R. the corresponding log likelihood functions are as follows:

$$\text{REML:} l_R(G,R) = -\frac{n-p}{2}\log(2\pi) - \frac{1}{2}\log|V| - \frac{1}{2}\log|x'V^{-1}x| - \frac{1}{2}r'V^{-1}r$$

Where the vector r = $y - x(x'V^{-1}x)^{-1}x'V^{-1}y$ and p = rank(x).

Mixed model actually minimizes -2 times these functions using a ridge-stabilized Newton- Raphson algorithm. The second derivative matrix (H) of the objective function evaluated at the optima is available upon completion and the asymptotic theory of maximum likelihood (Serfling ,1980) shows that $2H^{-1}$ is an asymptotic variance – covariance matrix of the estimated parameters of D and R. Mixed model profiles the residual variance out of log likelihood whenever it appears reasonable to do so. Therefore in LMM analysis, the REML provide estimates of D and R. which are denoted by $\widehat{D}$ and $\widehat{R}$, respectively.

3.10 Estimating β and predicting b in a LMM

Inferences about fixed effects have come to be called estimates, whereas those that concern random effects are known as predictions. Procedures for obtaining such estimators and predictors have been developed using a variety of approaches. The most widely used procedures are BLUE and BLUP, referring respectively to best linear unbiased estimator and to best linear unbiased predictor.

For mixed model ,the BLUE of β and the BLUP of b are offered a more compact method for jointly obtaining $\widetilde{\beta}$ and $\widetilde{b}$ in the form of (Henderson's, 1984)   mixed – model equations (MME),

$$\begin{bmatrix} x'R^{-1}x & x'R^{-1}z \\ x'R^{-1}x & z'R^{-1}z + D^{-1} \end{bmatrix} \begin{bmatrix} \widehat{\beta} \\ \widehat{b} \end{bmatrix} = \begin{bmatrix} x'R^{-1}y \\ z'R^{-1}y \end{bmatrix}$$

We used the qq-plot of residual, histogram of residual, and the scatter plot of standardized residual versus predicted values to check whether the stated normality assumption is hold (met).

3.11 Bayesian estimation approach

Inference will be made based on ML and empirical Bayes estimation methods. If Ө is q-vector of variances and covariance parameters found in R_i, i= 1,2,…,m and D, then estimation for β and Ө is based on ML from the marginal distributions of y an estimate to b can be obtained by use of an extended version of the Gauss-Markov theorem for random effects(Harville,1976). And it can be derived using Bayesian formulation of the model. Here we introduce a flat prior for the location parameter β and an estimate Ө from the marginal likelihood y after integrating out β and bi, i= 1,2,…,m. this approach was considered by Harville(1974,1976) and Dempster, Rubin and TSutakawa (1981). The empirical Bayes estimate of β and the b_i are the estimated means of the posterior distributions.

The Bayesian formulation is emphasized in this paper because it provides both a conceptual and computations unity. And we denote these estimates by $\tilde{\beta}(\hat{\ }_M)$ and $\tilde{b}_i(\Theta)$. When we consider the estimation of β and Ө simultaneously by maximizing there joint likelihood based on the marginal distribution of y. the ML estimates of $(\tilde{\beta}_M, \hat{\ }_M)$

Satisfy $\tilde{\beta}_M = \tilde{\beta}(\hat{\ }_M)$ an $\hat{b}_M = E(b/y, \tilde{\beta}_M, \hat{\ }_M)$ gives $\hat{b}_M = \tilde{b}(\hat{\ }_M)$ is the empirical Bayes estimate for b when Ө is estimated by ML, so it is estimated using
ML for β and empirical Bayes for b.

3.11.1 Bayesian model fitting

The Bayesian approach has a number of useful properties, for example it yields not only full posterior distributions for the parameters of interest but also full posterior (predictive) distributions for predicted values. It also provides a tractable method to fit more complex models - particularly of interest are those incorporating random effects that attempt to account for unobserved heterogeneity in the data set.

From a frequentist perspective, the unknown parameters θ are treated as fixed values that must be estimated from the data. In contrast the Bayesian approach instead treats the parameters as random variables that are generated from some probabilistic distribution.

A standard Bayesian model takes the form:

$$P(D/\theta) = \frac{P(D/\theta)P(\theta)}{\int P(D/\theta)P(\theta)}$$

That is the conditional posterior distribution for the parameters θ given the data D is equal to the likelihood (the distribution of D given θ) multiplied by a prior distribution for θ, up to some normalizing constant. Hence the unknown posterior distribution $p(\theta \mid D)$ is expressed in terms of a known likelihood $p(D \mid \theta)$ and a specified prior distribution $p(\theta)$. For simple models this can be calculated explicitly, however since the denominator involves integrating across the whole of the parameter space this becomes mathematically intractable when the number of parameters is large. Therefore a different fitting mechanism is required, the most widely used of which that of Markov chain Monte Carlo (MCMC) iterative is sampling. Monte Carlo integration involves sampling a large number of observations from a target distribution, and then using these samples to estimate various expected values. The law of large numbers ensures that the estimate can be made more accurate simply by increasing the sample size. Therefore if large numbers of samples can be obtained from the posterior distribution $p(\theta \mid D)$ then Monte Carlo integration offers a method to extract the required quantities of interest from these values. All that is required is a tractable method to sample from the posterior, and this can be done using a Markov chain. A Markov chain is a sequence of numbers where each number depends only on the previous value in the chain. It can be shown that under certain regularity conditions a Markov chain will converge to a so-called stationary distribution. If a Markov chain can be constructed such that its stationary distribution is identical to the posterior distribution of interest, then the required sample values can be obtained. MCMC combines these two techniques and has the advantage that it can produce estimates from the posterior distribution without requiring knowledge of the normalizing constant.

3.11.2 Gibbs sampling

The parameters θ do not have to be updated as a block, but can be updated separately if preferred, with corresponding changes to the proposal distributions. In this circumstance a special case of the Metropolis-Hastings algorithm occurs when knowledge of the full conditional distributions for individual parameters θ_i, $i = 1, \ldots, m$, given θ_{i-}, that is $p(\theta_i \mid x, \theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_m)$ are known. Hence the proposal distribution $q(\theta_{candi} \mid \theta_i, \theta_{i-}) = p(\theta_{candi} \mid \theta_{i-})$, and as a result the acceptance probability in is always equal to one. This technique is known as Gibbs sampling (Geman and Geman 1984, Gelfand and Smith 1990).

Combinations of Metropolis-Hastings and Gibbs sampling can be used if required, and the adaptive rejection sampling method proposed by Gilks and Wild (1992) means that as long as the conditional distributions of the parameters are log-concave, then Gibbs sampling can be used even if the distribution is complicated and is not specified explicitly. And These techniques are implemented in WinBUGS (Bayesian inference Using Gibbs Sampling).

3.12 Convergence analysis in Bayesian approach

The Gelman-Rubin statistic convergence analysis in Bayesian analysis is working to check convergence with three independent parallel chains. Before we summarize simulated parameters, we must ensure that the chains have converged and accurate.

History plot plots out a complete trace for the variable and can be used for diagnosing the convergence of parameter estimates in Bayesian analysis. The package gives the plot by making iteration number on the x-axis and parameter value on the y-axis for each parameter. For all parameters, the plots of the last iterations for two independently generated chains demonstrated well "chain mixture" an indication of convergence. The Time series plots (trace) show that the chains with two different colours overlap one over the other. Hence, we are reasonably confident that convergence has been achieved. Autocorrelation plot is the other recommended test for convergence of a Bayesian analysis. For all statistical parameters, the plots of the first lags of two independently generated chains demonstrated good "chain mixture" an indication of convergence. The plots show that the two independent chains were mixed or overlapped to each other and died out for higher lags and hence this is an evidence of convergence.

Density plot is another recommended technique for identifying non convergence. The plots indicates none of the coefficients have bimodal density, and hence the simulated parameter values were converged.

Deviance information criterion (DIC) can be used to assess model complexity and compare different models. It is important to note that DIC assumes the posterior mean to be a good estimate of the stochastic parameters. The deviance information criterion (DIC) is the generalization of the AIC for Bayesian model fitting using MCMC methods (Spiegel halter et al 2002). Dbar is the posterior mean of the deviance.

CHAPTER 4

4 Result and Discussion

4.1 Exploratory Data Analysis for evolution of upper arm circumference of infant

Exploratory analysis comprises techniques to visualize patterns in the data. Data analysis must begin by making displays that expose patterns relevant to the scientific question. Most longitudinal studies address the relationship of a response with explanatory variables, often including time. The following aspects of the data will be looked; individual profiles, the average evolution, the variance function, and the correlation structure. Data exploration is a very helpful tool in the selection of appropriate models. In the following sections exploratory analysis for the Jimma infants data sets considered.

Let us consider the jimma infant data set. In this data set, subjects are classified into two groups by the variable SEX, an indicator variable assuming the value MALE for boys and FEMALE for girls. Each subject has seven measures of AC, and the 8966 total records are grouped into 1480 groups by Id. We will try to see the mean profile for this data set.

4.2 Individual profile of upper arm circumference growth by sex

Figure 4.1 shows how the upper arm circumference of male and female infants evolves overtime. This is also something that we can observe from the individual profiles, which show an increasing pattern over age. There is an increase in upper arm circumference overtime for both males and females. Of course, at this point, it is not yet possible to decide on the significance of this difference. On the other hand, from the individual profiles, it seems that the variability is almost the same among the two groups.
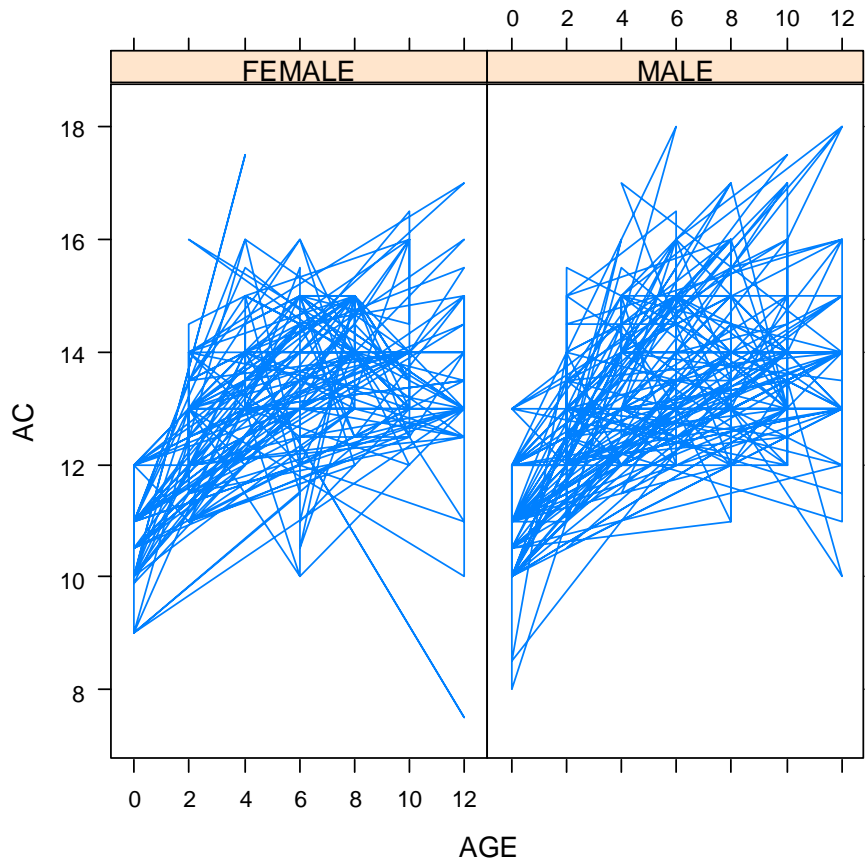
21

Figure 4.1: Individual profile plot of upper arm circumference

4.3 Mean profile of upper arm circumference growth by sex

Figure 4.2 shows the average mean profile plot of mean evolution of upper arm circumference by sex, males appear to have higher mean profile than females, though this is not quite clear from subject specific profile plots. Of course, at this point, it is not yet possible to decide on the significance of this difference.
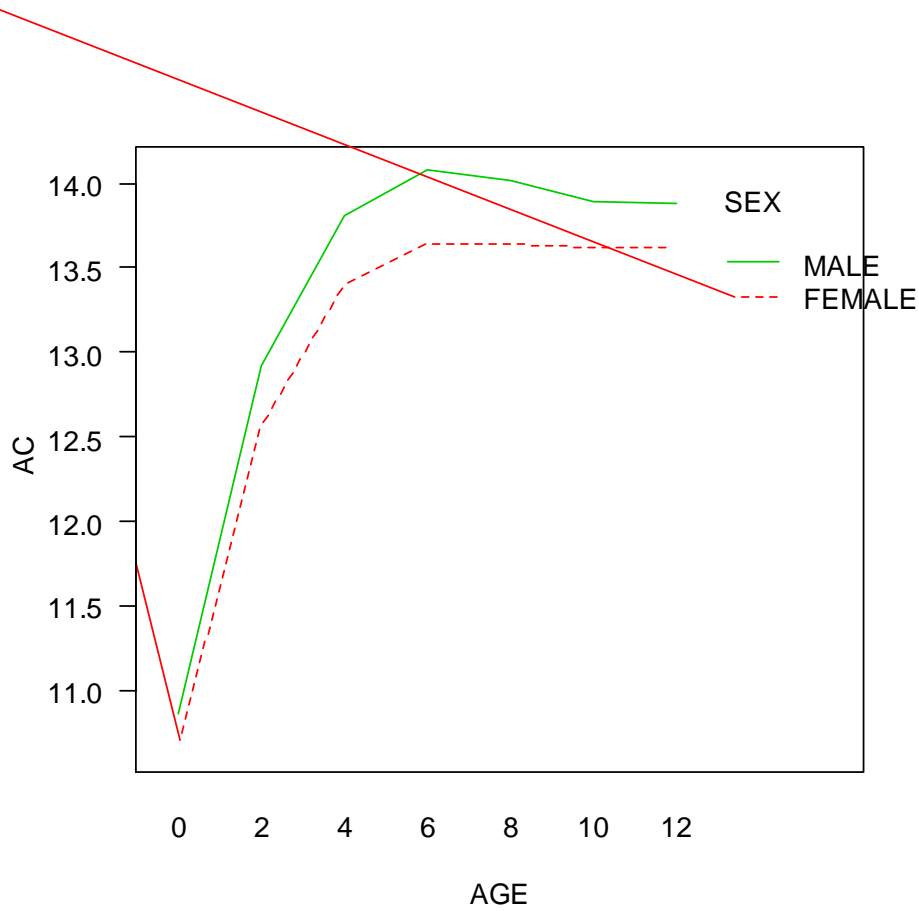
Figure 4.2: mean profile plot of upper arm circumference

From figure 4.2 **:**

  -It appears that there is a linear growth pattern up to age 6 and almost constant after
  age 6 for both sex.

  -The mean profile for males is higher than that of females

4.4 Exploring the Variability of the Observed Data

Having appropriate model, studying the evolution of the variance is very important step of the modeling approach. For the Jimma infant survival data, the observed variance displayed in the Figure below, shows an increase in variability overtime. Hence, a heterogeneous variance structure may be a good starting point.

Moreover, the variability for males and females seems to be more or less the same up to age 6 and different after age 6.
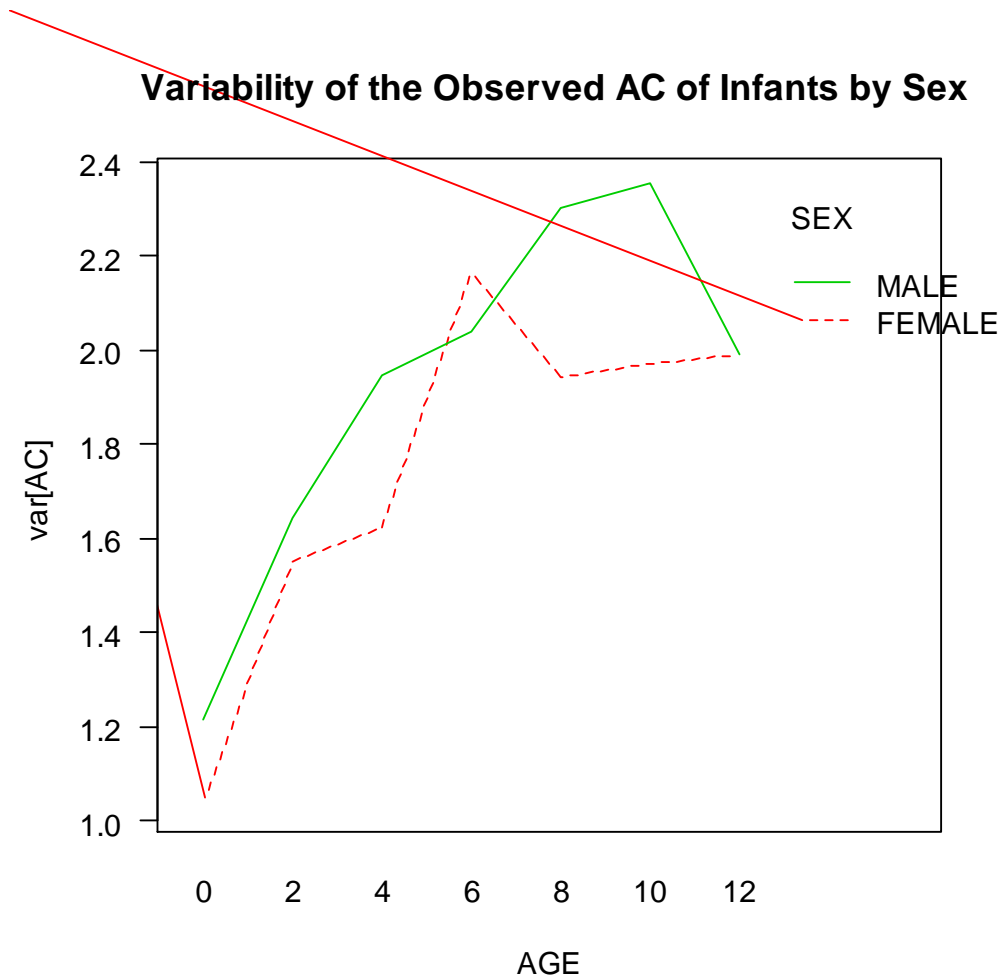

Figure 4.3: Variability of the Observed AC

4.5 Statistical analysis of Data Sets Using Mixed Models

4.5.1 Fitting Linear Mixed Effect Model for upper arm circumference of infant's

A model fitted to the evolution of upper arm circumference of infant using the covariate age(AGE), sex(SEX), supplementary food(SUPPFOOD), powder milk(POWDERMI) and their interaction with age(AGE) the model is,

$$AC_i = \beta_0 + \beta_1 AGE + \beta_2 * SEX + \beta_3 * SUPPFOOD + \beta_4 * POWDERMI + \beta_5 * AGE * SEX + \beta_6 * AGE * SUPPFOOD + \beta_7 * AGE * POWDERMI + b_{0i} + b_{1i} * AGE + \epsilon_i$$

Where β_i 's are the parameters and b_{oi}, b_{1i} are intercept and slope for the random effect and they are normaly distributed, ϵ_i is the within unit variation for i = 1,2,….

The random effects estimate using restricted maximum likelihood method.

Table 4.1 standard deviation of random effects estimate using restricted maximum likelihood method.

	StdDev	**Corr**
(Intercept)	0.8352031	(Intr)
AGE	0.1081101	-0.052
Residual	0.9670271	

The lme(linear mixed effect) function we are using in the above model is the restricted-maximum likelihood fit, which tends to produce more conservative estimates of the variance components. A maximum likelihood fit of Random effects estimate using maximum likelihood method can be obtained.

Table 4.2 standard deviation of random effects estimate using maximum likelihood method.

	StdDev	Corr
(Intercept)	0.8339264	(Intr)
AGE	0.1078235	-0.049
Residual	0.9661860	

As expected, the ML estimates of the random-effects standard deviations are smaller than the corresponding REML estimates. The estimated within group residual standard deviations are identical. In general, the fixed-effects estimates obtained using ML and REML are similar. Inferences regarding the fixed effects are essentially the same for the two estimation methods. By fitting linear model using linear model (lm) fit, we can compare it with lme fit as follows.

$AC_i = \beta_0 + \beta_1 AGE + \beta_2 * SEX + \beta_3 * SUPPFOOD + \beta_4 * POWDERMI + \beta_5 * AGE * SEX + \beta_6 * AGE * SUPPFOOD + \beta_7 * AGE * POWDERMI + \epsilon_i$

Table 4.3 comparison of lm with lme

	Model	*Df*	*AIC*	*BIC*	*logLik*	*Test*	*L.Ratio*	p-value
fit8	1	12	28636.29	28721.49	-14306.14			
fitlm8	2	9	31972.18	32036.08	-15977.09	1 vs 2	3341.887	<.0001

Where fit8 is lme and fitlm8 is linear regression model.

In this case, as evidenced by the low p-value for the likelihood ratio test, the linear mixed-effects model provides a much better description of the data than the linear regression model. This is because the model used in lm ignores the group structure of the data and incorrectly combines the between-group and the within-group variation in the residual standard error.

4.5.2 Patterned Variance Covariance Matrices for the Random Effects

The pdMat(positive definite matrix) classes are used to specify patterned variance covariance matrices for the random effects. The default class of positive definite matrix for the random effects is pdSymm, corresponding to a general symmetric positive definite matrix. The constructor for the pdDiag class is also called pdDiag. Because initial values for D can be derived internally in the lme function and it can be fitted.

$AC_i = \beta_0 + \beta_1 AGE + \beta_2 * SEX + \beta_3 * SUPPFOOD + \beta_{4*} POWDERMI + \beta_5 * AGE * SEX + \beta_6 * AGE * SUPPFOOD + \beta_7 * AGE * POWDERMI + b_{0i} + b_{1i} * AGE + \epsilon_i$

Where β_i 's are the parameters and b_{oi}, b_{1i} are intercept and slope for the random effect with diagonal structure.

When we fit with pdDiag class all estimates are similar to the previous object which was created previously with the default covariance structure (pdSymm). Another pdMat classes, that are used to specify patterned variance covariance matrices for the random effects is positive definite compound symmetry structure (pdCompSymm) and block diagonal (pdBlocked).

We can compare the four models one with general symmetric positive-definite (pdSymm) covariance matrix the other with the modified or simplified covariance structure(pdDiag) and the block diagonal(pdBlocked) covariance matrix using anova procedure as follows.

Table 4.4 comparison of the different patterned covariance structure

pdMat	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
Fit8	1	13	28633.51	28725.81	-14303.75			
pdDiag	2	12	28632.40	28717.60	-14304.20	1 vs 2	0.88919	0.3457
pdCompSymm	3	12	29287.58	29372.78	-14631.79			
pdBlocked	4	12	28632.40	28717.60	-14304.20			

since the fitted models are not nested we can compare them using AIC and BIC. So, the smaller AIC and BIC makes fit3diag preferable.

5.3 Variance Functions with linear mixed effect

The varFunc classes are used to specify within-group in the mixed effects model. Standard varFunc classes like varFixed (fixed variance), varIdent (different variances per stratum),varPower (power of covariate), varExp (exponential of covariate), varComb(combination of variance functions).

Figure 4.4 shows that the plot of the standardized residuals versus fitted values by gender for homoscedasticity variance and the variability in the infants' upper arm circumference measurements is almost the same for both genders but few outlying observations are clearly seen.

- Within each gender the variability is somewhat constant
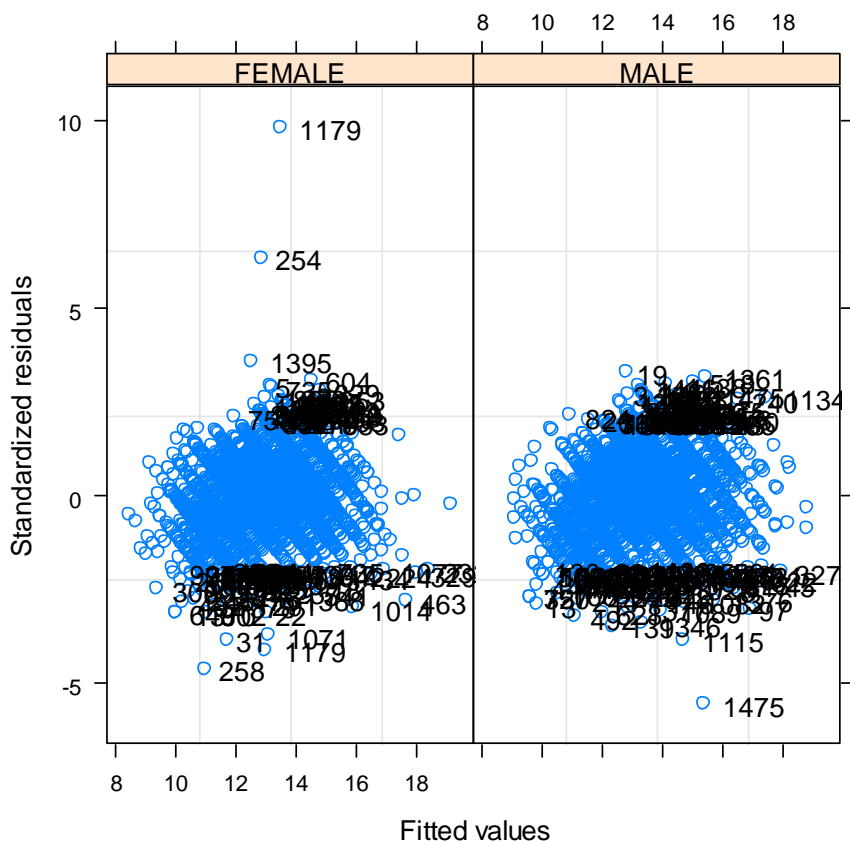- Few outlying observations are clearly seen.

Figure 4.4 plot of the standardized residuals versus fitted values

The above results suggest that a model that allows different variances by gender for the within-group error might represent the upper arm circumference growth of infant data set. The lme function allows the modeling of heteroscesdasticity of the within-error group via a weights argument. The varIdent variance function structure allows different variances for each level of a factor and can be used to fit the heteroscedastic model for the upper arm circumference evolution of infant.

The parameters for varIdent give the ratio of the stratum standard errors to the within-group standard error. To allow identifiably of the parameters, the within-group standard error is equal to the first stratum standard error. The standard error for the girls is about 0.965717/1 multiplied by 100 gives 96 percent of that for the boys. The remaining estimates are very similar to the ones in the homoscedastic fit.

We can assess the adequacy of the heteroscedastic fit by re-examining plots of the standardized residuals versus the fitted values by gender. Figure 4.5 gives us the Scatter plots of standardized residuals versus fitted values for the heteroscedastic fit of varIdent by gender.
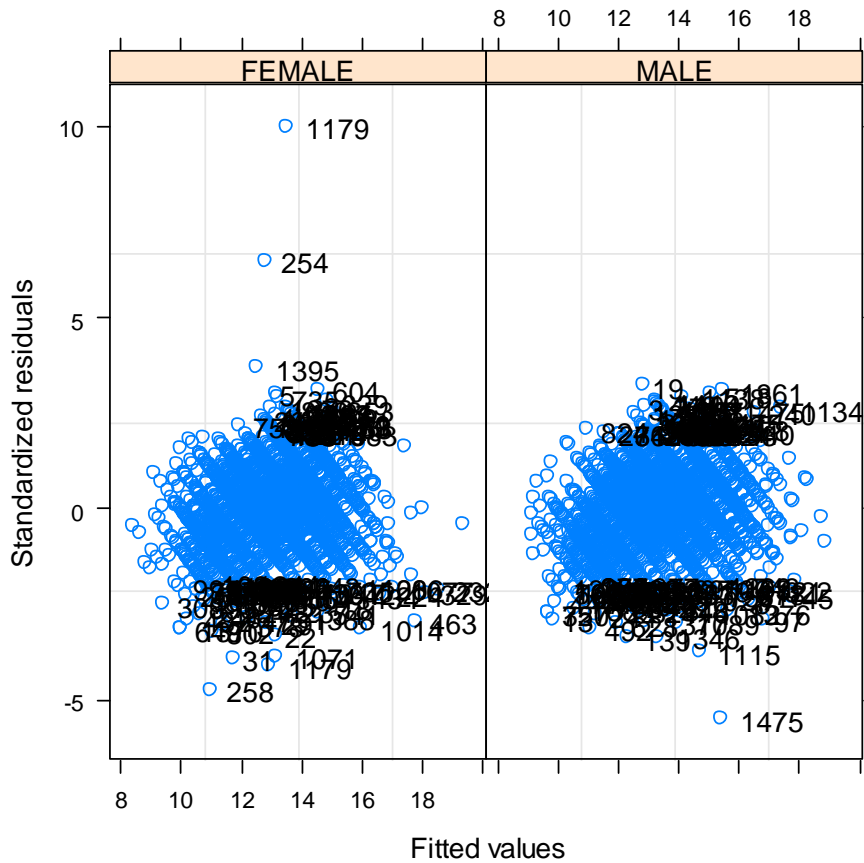
Figure 4.5 Plots of the standardized residuals versus the fitted values by gender.

- Like the previous plot the standardized residuals in each gender have about the same variability.
- We can still identify the outlying observations, corresponding to the previous Id.

A better way of seeing this by looking at a plot of the observed responses versus the within-group fitted values for hetroscedastic variance of model varIdent as shown in Figure 4.6.
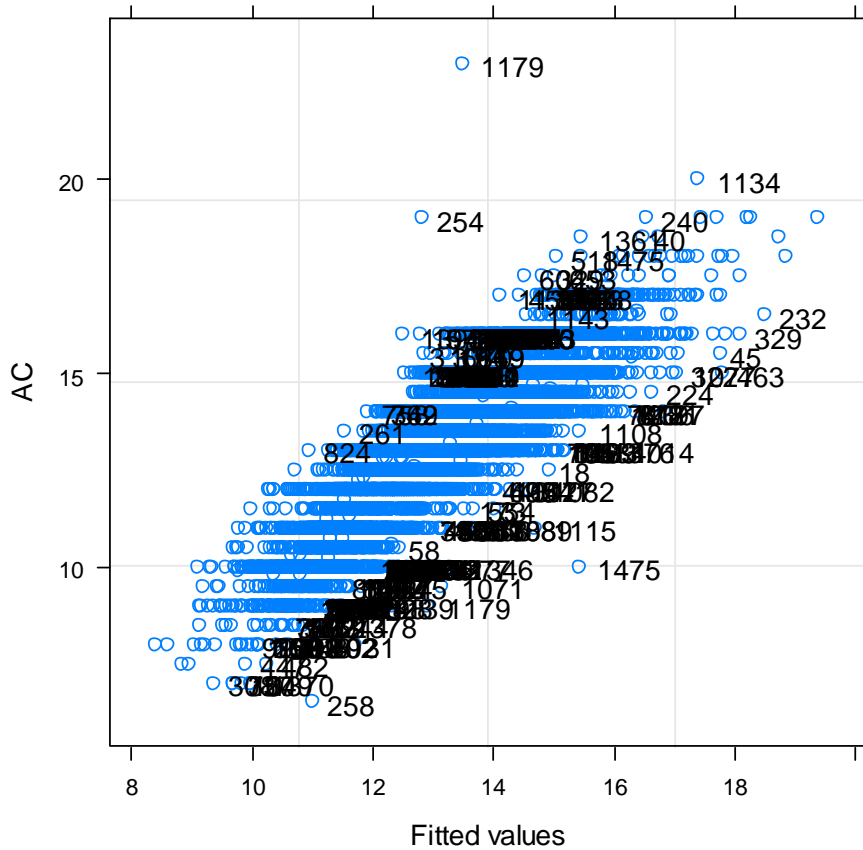
Figure 4.6 Plot of the observed responses versus the within-group fitted values

The varIdent fitted values are in close agreement with the observed infant upper arm circumference, except for the two extreme observations on Id 254 and 1179. The need for an heteroscedastic model growth data can be formally tested by comparison.

Table 4.5 comparison of varident variance function with the default variance function

Variance func	Model	df	AIC	BIC	logLik
default	1	12	28636.29	28721.49	-14306.14
varIdent	2	12	28633.30	28718.50	-14304.65

The small AIC of the anova test confirms that the heteroscedastic model explains the data significantly better than the homoscedastic model. The assumption of normality for the within-group errors can be assessed with the normal probability plot of the residuals, produced by the qqnorm method. Figure below, is a normal plots of the residuals corresponding to varIdent by gender.
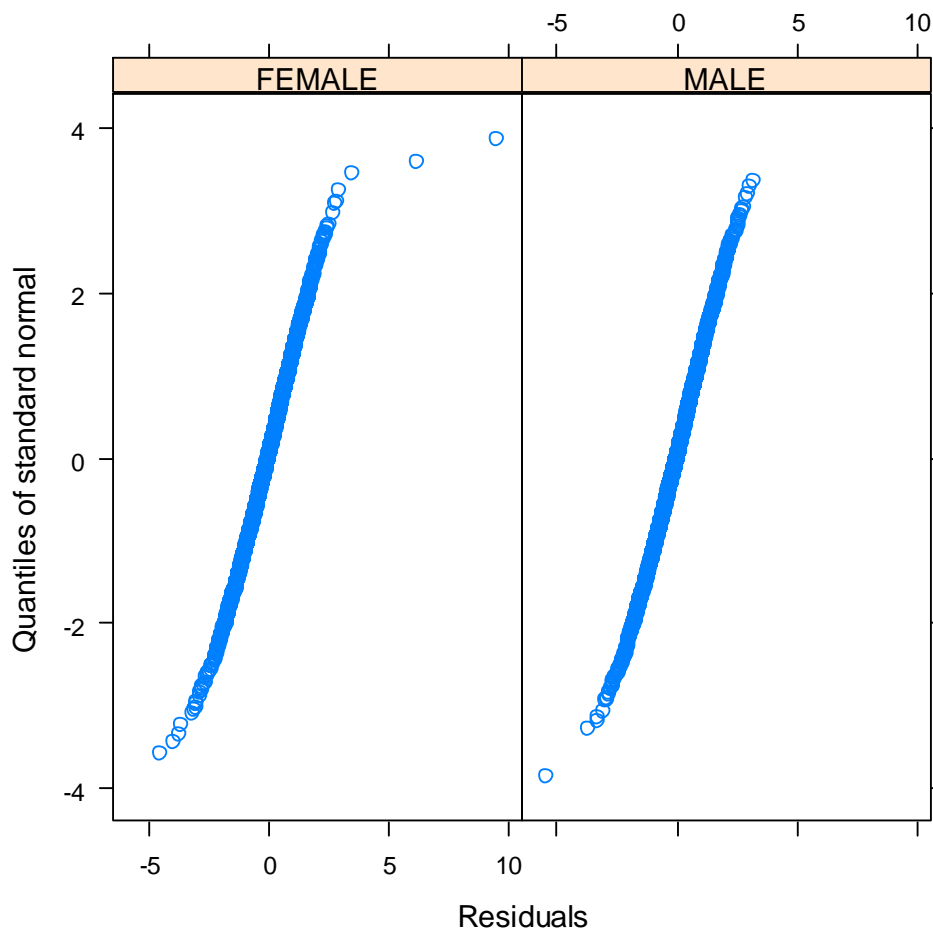

Figure 4.7 Normal Probability Plot of Residuals for the Model with Heteroscedastic Variance

Once again, we observe the two outlying points, but for the rest of the observations the normality assumption seems plausible.

4.5.3.1 Assessing Assumptions on the Random Effects

In this section, we describe diagnostic methods for assessing on the distribution of the random effects. As we have seen in the previous sections the random effects method is used to extract the estimated BLUPs of the random effects from lme objects. These are the primary quantities for assessing the distributional assumptions about the random effects.

Let us consider for which we have fitted the homoscedastic and heteroscedastic models. We first consider the homoscedastic fitted object and investigate the marginal normality of the corresponding random effects using the normal probability plots.

And figure 4.8 shows:

-The assumption of normality seems reasonable for both random effects, though there
     is some asymmetry in the distribution of the (AGE) random effects.

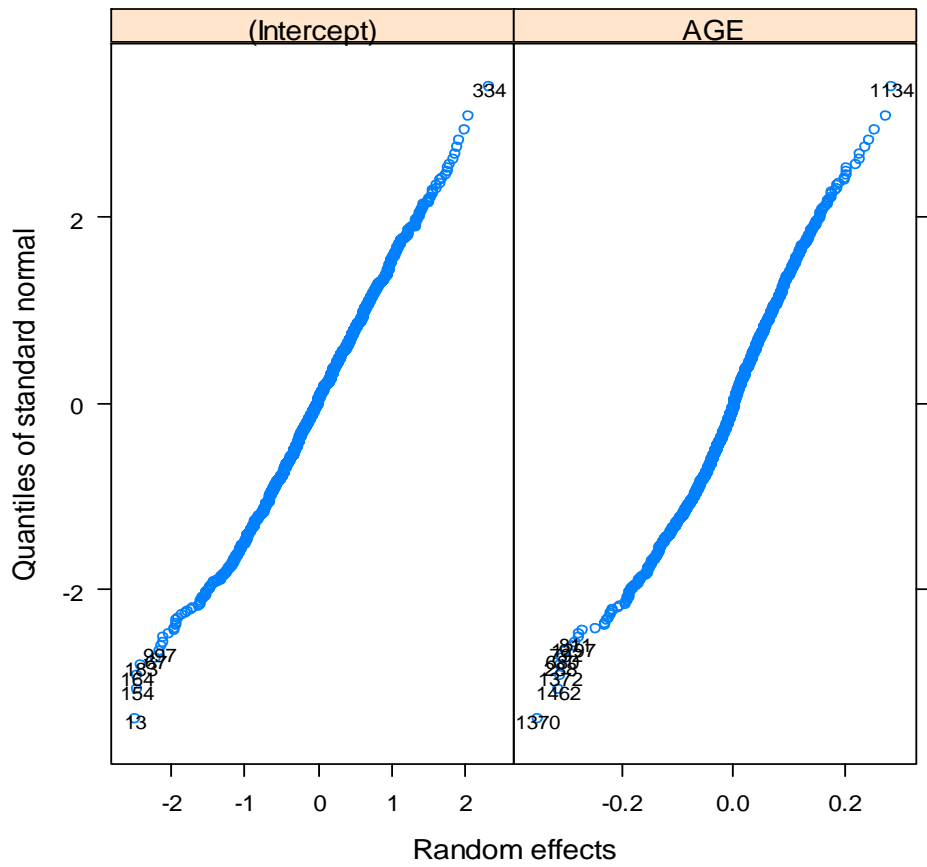- A few outliers appear to be present in both random effects.

Figure 4.8 Normal Probability Plot of Residuals for the Model with homoscedastic Variance

To investigate the homogeneity of the random-effects distribution for boys and girls, we use the pairs method to obtain scatter plots of the random-effects estimates by gender.
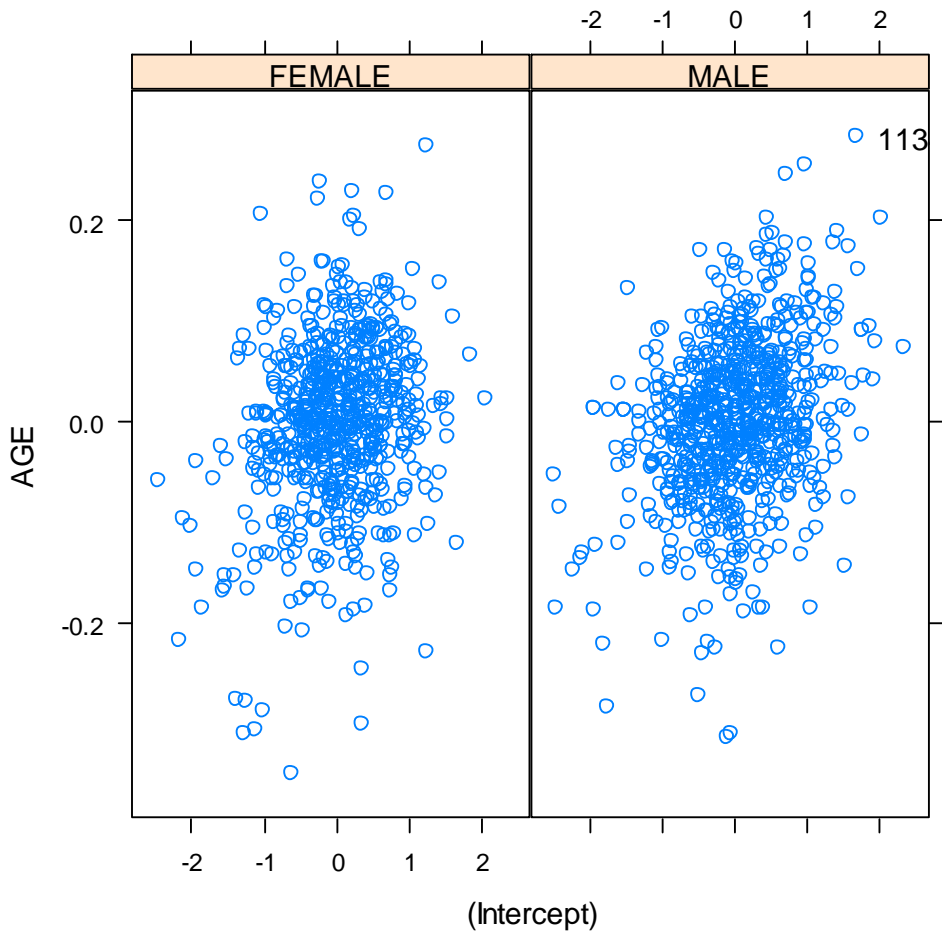
Figure 4.9 Scatter Plots of Random Intercept versus Slope for Homoscedastic Model

- This plot is the scatter plot of estimated random effects for the homoscedastic fit.
- Once again the observation from 1134 is identified as outlier.
- Except for the pair corresponding to Id 1134, the estimated random effects in
  the two groups seem to have similar distributions.

 Now let us consider the normal plot of estimated random effects for the heteroscedastic
varIdent lme fit.
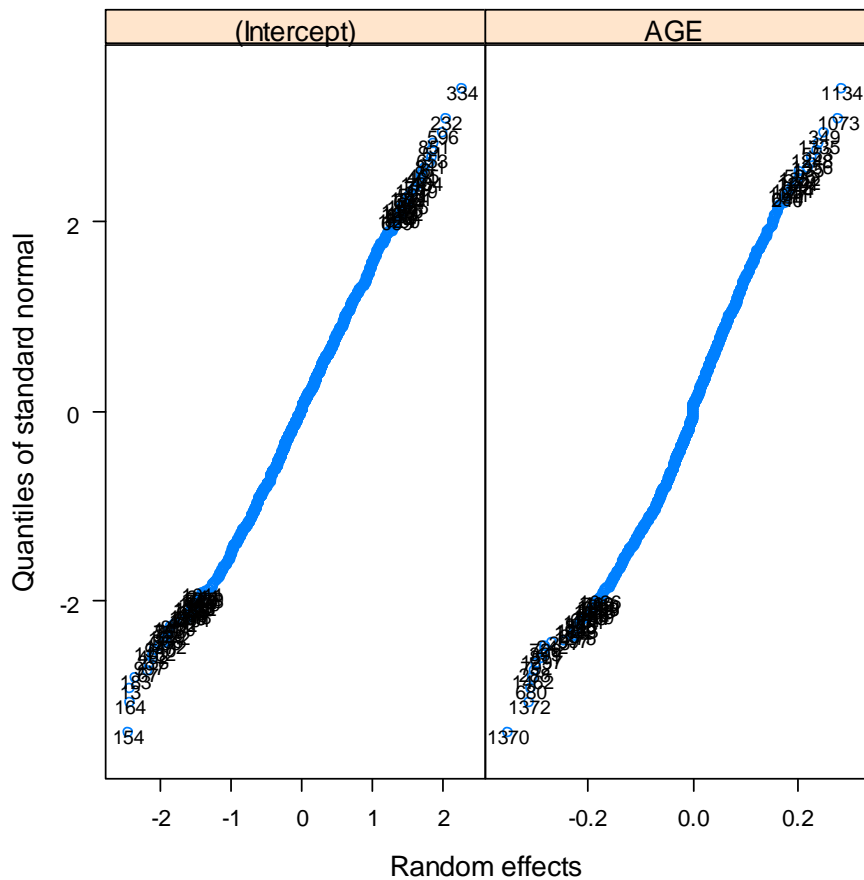
Figure 4.10 Normal Probability Plot of Residuals for a Model with Heteroscedastic Variance

- Figure 4.10 is a normal probability plot for the model with heteroscedastic random effects.
- The normal probability plots of the estimated random effects with heteroscedastic varIdent fit are almost similar to the corresponding plots for the model with homoscedastic fit.
- In mixed-effects estimation, there is a trade-off between the within group variability and the   between-group variability, when accounting for the overall variability in the data.

The homogeneity of the random-effects distribution for boys and girls can be investigated by using the pairs method. This gives us the scatter plots of the random-effects estimates by gender, for the heteroscedastic model varIdent are as shown in Figure 4.11.
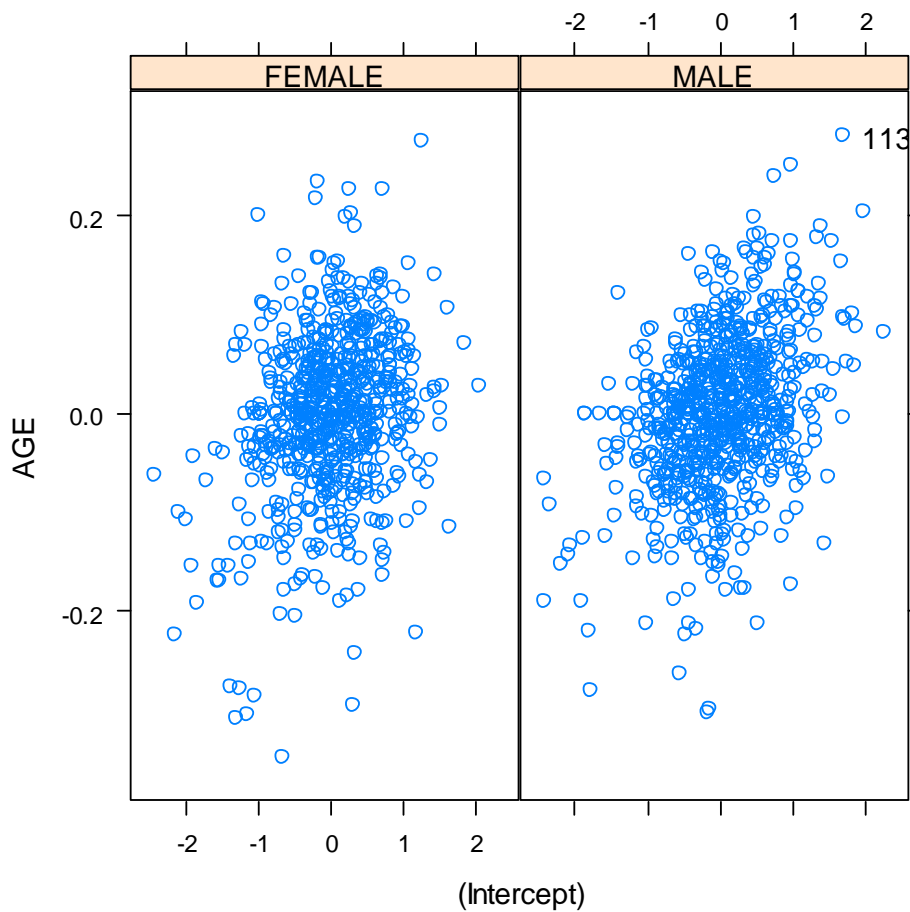

Figure 4.11 Scatter Plot of Random Intercept versus Slope for Heteroscedastic Model

The above plot does not suggest any departures from the assumption of homogeneity of the random effects distribution.

4.5.3.2 Different Variance Functions for Modeling Heteroscedasticity

Variance functions are used to model the variance structure of the within group errors using covariates. Here we will try to show how to apply the different variance functions in modeling heteroscedasity.

We begin by using the fitted homoscedastic linear mixed effects model. The primary tool for investigating within-group heteroscedasticity is the plots of residuals against the fitted values and other candidate variance covariates. Let us first see the plot of residuals with age by sex for homoscedastic fitted model.
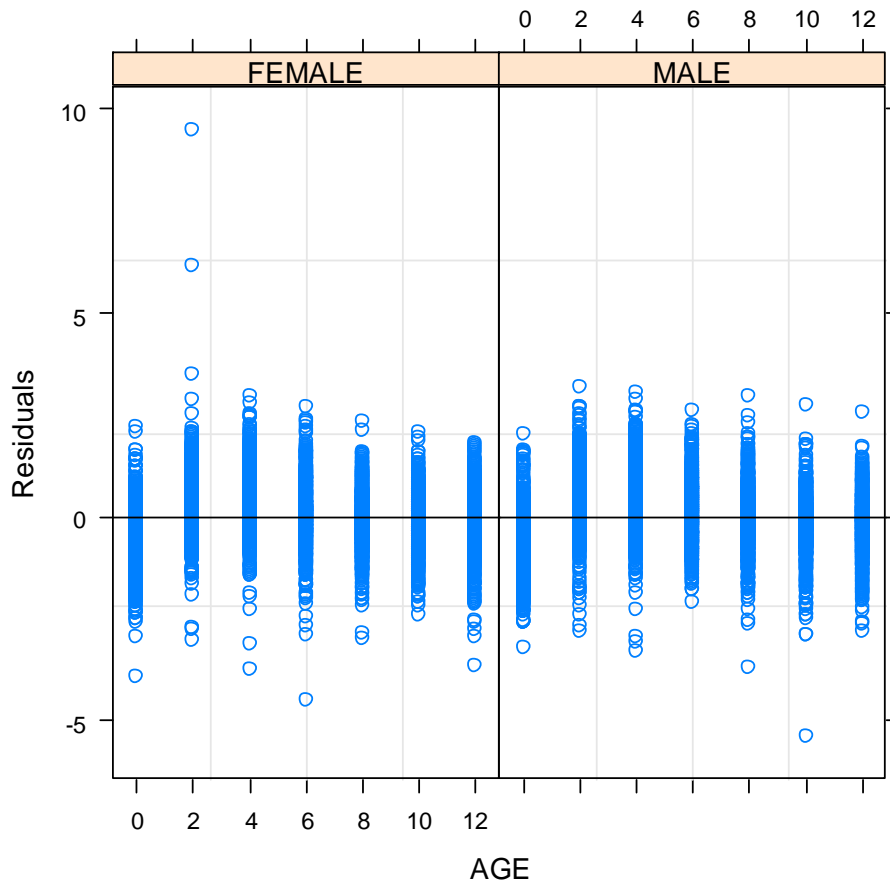

Figure 4.12   Plots of raw residuals versus AGE

From the plots of raw residuals AGE by SEX, shown in Figure 4.12, we can see the residual variability is some what higher for females than for males. However, the residuals variability remains constant with age classifying residuals by sex. The variance covariate used in this case is AGE. Let us see by fitting varPower variance function to compare with the homoscedastic model.

Table 4.6 Comparison of the homoscedastic with heteroscedasitic varPower models

Variance function	Model	df	AIC	BIC	logLik
defoult	1	12	28636.29	28721.49	-14306.14
varPower	2	12	28611.57	28696.78	-14293.79

The smallest AIC for hetroscedastic varPower model is preferable. We may wish to further investigate if a varExp and varComb variance structure. This is the combination of variance structures. With this variance structure, we can allow for both an increase in residual spread for larger AGE values as well as a different spread per SEX.

4.5.3.2.1 Comparison between models for the fitted variance function

Now let us see how we can compare all the models we have fitted above. In order to use the likelihood ratio test to compare models the models must be nested. Two models are called nested if one model can be obtained from the other model by setting specific parameters equal to zero. Therefore, in our case models are not nested and we use AIC to compare the models.

Table 4.7 comparison of the different variance function

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
fit3diag	1	11	28635.27	28713.38	-14306.64			
fm3Orth.lme	2	12	28633.30	28718.50	-14304.65	1 vs 2	3.976	0.0461
fitvp.lme	3	12	28611.57	28696.78	-14293.79			
fm6Orth.lme	4	13	28394.37	28486.67	-14184.18	3 vs 4	219.21	<.0001
fm7Orth.lme	5	14	28394.79	28494.20	-14183.40	4 vs 5	1.574	0.2097

Where fit3diag, fm3Orth.lme, fitvp.lme, fm6Orth.lme, and fm7Orth.lme are models for pdBlocked, varIdent, varPower, varExp and varComp variance functions.

The smallest AIC for hetroscedasticity model fm6Orth.lme is preferable. AIC is almost equal to model fm7Orth.lme but when we look their Bayesian information criteria(BIC) model fm6Orth.lme has smallest BIC. Therefore, model fm6Orth.lme is preferable.

4.5.4 Correlation Structure for Modeling Dependence

In the linear mixed effects model, $Yi = X_i\beta + Z_ib_i + \epsilon_i$, We have assumed that the within group error terms are independent and have homoscedastic variance. The assumption of homoscedastic variance was relaxed by using different types of variance functions. Now we shall discuss different approaches of handling the dependence of the within error terms in the model. Correlation structures are used to model dependence among observations. In the context of linear mixed-effects models the correlation structures are used to model dependence among the within-group errors. The nlme library provides a set of classes of correlation structures, the corStruct classes, which are used to specify within-group correlation models to model dependence of the within group errors in the above linear mixed effects model.

4.5.4.1 Different Correlation Structures with linear mixed effect

Correlation structures are specified as corStruct objects. Now we describe the use of correlation models in lme by considering the upper arm circumference growth of infant data set. While analyzing the upper arm circumference growth of infants data set we have selected one of the Heteroscedasitic variance model. We will begin from this model to search for a better correlation structure. The selected model was fm6Orth.lme.

Let us assume that our data are time series data and then study the autocorrelation function . The ACF method for the lme class obtains the empirical autocorrelation function from the residuals of an lme object. The autocorrelation function of the selected model fm6Orth.lme is indicated below.

Table 5.8 Autocorrelation function

s.no	lag	ACF
1	0	1.00000000
2	1	-0.02551666
3	2	-0.22255578
4	3	-0.26064938
5	4	-0.22406909
6	5	-0.04978312
7	6	0.15765067

Figure 4.13 shows The empirical autocorrelation functions are significantly different from 0 at lag one.

Figure 4.13   Plot of Auto Correlation
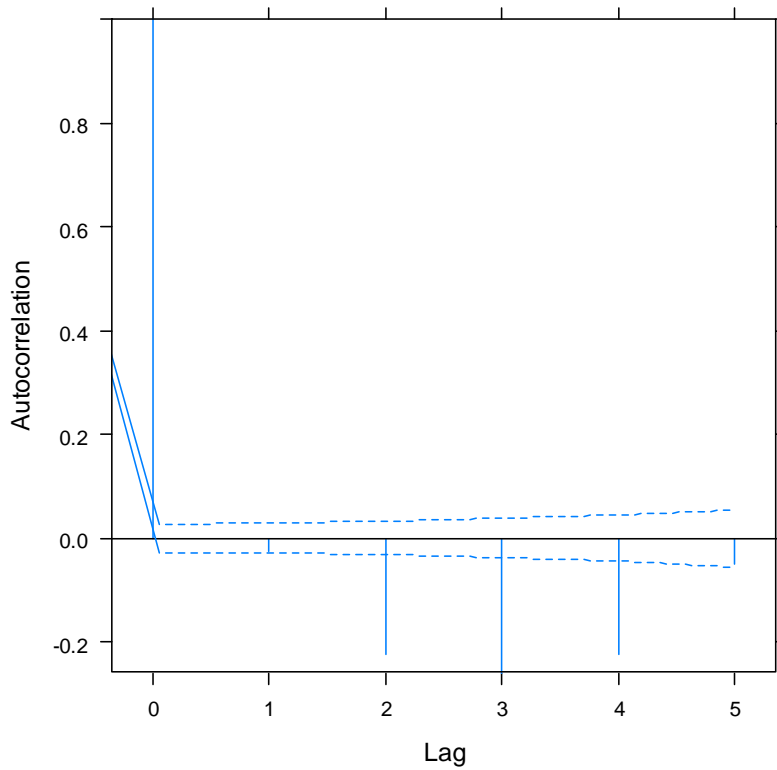
This suggests that an AR(1) model may be fitted and compared with model fm6Orth.lme. The new model that takes into account the dependence of the within error terms can be fitted with AR(1) correlation structure and using different correlation structures like moving average of order one, continuous-time AR(1), ARMA(1,1), exponential, compound symmetric, ratio, spherical, linear, and Gaussian.

4. 5.4.1.2 Comparison between models for the fitted correlation structure

The following output is model comparison for different correlation structures. This result shows that the model with corARMA (Autoregressive-moving average) correlation structure is preferable model for the infant data.

Table 4.9 Comparison of models for selection of different correlation structure.

	Model	df	AIC	BIC	logLik	Test	L.Ratio	P-value
fmcor.lme	1	14	28396.37	28495.77	-14184.2			
fm6Orth.lme	2	13	28394.37	28486.67	-14184.2	1 VS 2	0.0000	1.0000
fmcor.lme2	3	15	28025.73	28132.23	-13997.9	2 VS 3	372.6406	<.0001
fmcor.lme3	4	14	28194.17	28293.57	-14083.1	3 VS 4	170.4392	<.0001
fmcor.lme4	5	15	28188.74	28295.24	-14079.4	4 VS 5	7.4273	0.0064
fmcor.lme5	6	14	28194.17	28293.57	-14083.1	5 VS 6	7.4273	0.0064
fmcor.lme6	7	14	28736.58	28835.98	-14354.3			
fmcor.lme7	8	14	28194.17	28293.57	-14083.1			
fmcor.lme8	9	14	28212.24	28311.64	-14092.1			
fmcor.lme9	10	14	28212.24	28311.64	-14092.1			
fmcor.lme10	11	14	28211.69	28311.09	-14091.8			

Where fmcor.lme, fm6Orth.lme, fmcor.lme2, fmcor.lme3, fmcor.lme4, fmcor.lme5, fmcor.lme6, fmcor.lme7, fmcor.lme8, fmcor.lme9, and fmcor.lme10 are model for Autoregressive of order one, exponential variance function, Autoregressive moving average of order one, continuous-time AR(1), ARMA(1,1), exponential, compound symmetric, ratio, spherical, linear, and Gaussian respectively.

For infant data set we have tried several models with different correlation structure. We end up in selecting a model fmcor.lme2 with smallest AIC, BIC, using anova , with Autoregressive-moving average. To mention a few points about modeling in R first we deal with the variance structure by using different variance functions then we consider the correlation structure. The combined result will be equivalent to dealing with variance covariance structure.

The following table 4.10 is fixed effect estimate for the selected model with Autoregressive-moving average.

Table 4.10 fixed effect estimate of the parameter

	Value	Std.Error	DF	t-value	p-value
(Intercept)	11.41677	0.047146	7479	242.1603	0.0000
AGE	0.338638	0.007578	7479	44.68478	0.0000
SEXMALE	0.251061	0.06256	7479	4.01314	0.0001
SUPPFOOD	1.843013	0.069629	7479	26.46906	0.0000
POWDERMI	0.439541	0.27244	7479	1.61335	0.1067
AGE:SEXMALE	0.006951	0.008092	7479	0.85891	0.3904
AGE:SUPPFOOD	-0.31411	0.008809	7479	-35.6587	0.0000
AGE:POWDERMI	-0.06051	0.034399	7479	-1.75909	0.0786

The small p-values associated with SEXM and large p-value for AGE:SEXM in the summary output indicate that boys and girls haven't significant difference in upper arm circumference growth patterns. And also the parameters intercept, age, sex, supplementary food, and the interaction of age with supplementary food are found to be significant parameter. The other parameter powder milk, and the interaction of age with sex, the interaction of age with powder milk are not significant.

Table 4.11 is the summarized maximum likelihood and restricted maximum likelihood estimators of the model with random intercept and slope. The summary is made for the model that contains eight covariates and as we can see the standard errors of the two estimates are the same.

Table 4.11 summarized maximum likelihood and restricted maximum likelihood estimators

Effect	Parameter	MLE(s.e)	REML(s.e)
(Intercept)	β_0	11.416962(0.047)	11.416767(0.05)
AGE	β_1	0.338599(0.008)	0.338638 (0.008)
SEXMALE	β_2	0.251115 (0.063)	0.251061 (0.062)
SUPPFOOD	β_3	1.843030(0.069)	1.843013 (0.069)
POWDERMI	β_4	0.439355(0.272)	0.439541 (0.272)
AGE:SEXMALE	β_5	0.006944 (0.008)	0.006951 (0.008)
AGE:SUPPFOOD	β_6	-0.314096(0.009)	-0.314112 (0.009)
AGE:POWDERMI	β_7	-0.060463 (0.034)	-0.060512 (0.034)

Variance covariance matrix for Random effects of selected model.

Table 4.12 Variance covariance matrix for Random effects of selected model.

	(Intercept)	AGE
(Intercept)	0.495220	0.0142460
AGE	0.014246	0.0087795

Now let us consider the normal plot of estimated random effects for the heteroscedastic fmcor.lme2 lme fit using finally selected(preferable) model.
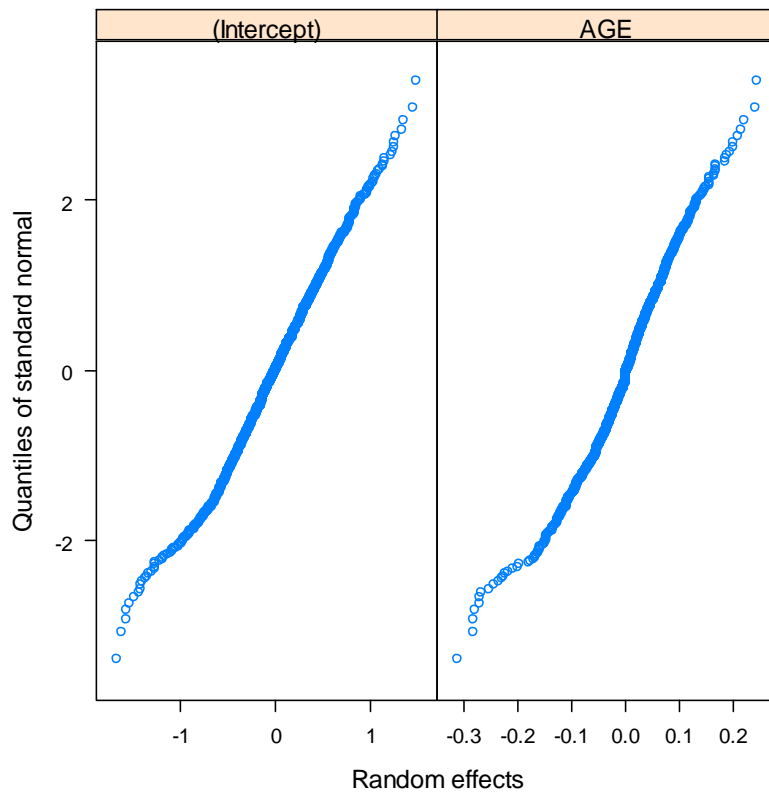
Figure 4.14 Normal plots of estimated random effects

- The assumption of normality seems reasonable for both random effects.
- A few outliers appear to be present in both random effects.

4.5.5 Inference for fixed effects using contrasts for Linear Mixed effects

Using the evolution of upper AC we shall briefly show how to apply contrasts to lme objects. Suppose our interest is to compare the mean evolution of upper arm circumference of infant between genders, mean evolution of upper arm circumference of infant changes over time and mean evolution of upper arm circumference of infant differ for those infant given supplementary food or not .

4.5.5.1 Mean evolution of upper arm circumference of infant over time and gender difference.

Table 4.13 fixed effect for mean evolution of upper arm circumference of infant

	Value	Std.Error	DF	t-value	P-value
(Intercept)	10.683090	0.04531637	7473	235.7446	0.0000
AGE 2	1.900051	0.04496911	7473	42.25235	0.0000
AGE 4	2.725439	0.05094171	7473	53.50113	0.0000
AGE 6	2.948209	0.05543581	7473	53.18239	0.0000
AGE 8	2.903339	0.05857493	7473	49.56624	0.0000
AGE 10	2.853430	0.06178234	7473	46.18520	0.0000
AGE 12	2.837058	0.06608990	7473	42.92726	0.0000
SEXMALE	0.179800	0.06205235	7473	2.89755	0.0038
AGE 2:SEXMALE	0.133679	0.06219779	7473	2.14926	0.0316
AGE 4:SEXMALE	0.191518	0.07016389	7473	2.72958	0.0064
AGE 6:SEXMALE	0.201832	0.07657086	7473	2.63589	0.0084
AGE 8:SEXMALE	0.153391	0.08066327	7473	1.90163	0.0573
AGE 10:SEXMALE	0.054971	0.08495498	7473	0.64706	0.5176
AGE 12:SEXMALE	0.041618	0.09048396	7473	0.45994	0.6456

From the above output we can see that the intercept = 10.683 is the mean evolution of upper arm circumference of infant for females at the reference age 0. The corresponding mean evolution of upper arm circumference of infant for males is given by 10.683+0.1798=10.851. The coefficients for age 2, age 4, age 6, age 8, age 10, and age

12 are 1.900, 2.7254, 2.9482, 2.9033, 2.8534, 2.8371 respectively. These values show the difference in mean evolution of upper arm circumference of infant between the reference age and the specified age level for females. For instance, the coefficient of age 2, which is 1.900, is the difference in mean evolution of upper arm circumference of infant between age 0 and age 2 for females.

The coefficients for the interaction terms refer to the difference between males and females at deferent age levels. From our above output we can produce the coefficients for males and females as follows.

Table 4.14 Estimate of the coefficients for male and female at different age levels

	Estimate For Female	Estimate for male
(Intercept)	10.683090	10.6830+0.180=10.863
AGE 2	1.900051	1.900051+0.134=2.034
AGE 4	2.725439	2.725439+0.192=2.917
AGE 6	2.948209	2.948209+0.202=3.150
AGE 8	2.903339	2.903339+0.155=3.058
AGE 10	2.853430	2.853430+0.055=2.908
AGE 12	2.837058	2.837058+0.042=2.879

Therefore small p-value for age 2, age 4, ….,age 12 indicates that there is significance difference in mean evolution of upper arm circumference of infants between the reference age 0 and age 2, age 4, ….,age 12. and for the interaction term AGE12:SEXMALE, AGE14:SEXMALE and AGE16:SEXMALE small p-value indicates that there is significance difference in mean evolution of upper arm circumference of infants between males and females at age 2, age 4, age 6.

The other interactions have large p-value which means that there is no significant difference in mean evolution of upper arm circumference of infants between males and females at age 8, 10, 12 respectively.

4.5.5.2 Mean evolution of upper arm circumference of infant differ for supplementary food behavior.

Table 4.15 fixed effect estimate of the mean evolution of upper arm circumference of infant differ for supplementary food behavior.

	Value	Std.Error	DF	t-value	P-value
(Intercept)	10.77773	0.032234	7473	334.3619	0.0000
AGE 2	1.968151	0.031089	7473	63.3069	0.0000
AGE 4	2.857342	0.036462	7473	78.3644	0.0000
AGE 6	3.109016	0.045026	7473	69.0491	0.0000
AGE 8	3.087466	0.058732	7473	52.5691	0.0000
AGE 10	3.002383	0.07666	7473	39.165	0.0000
AGE 12	3.048194	0.10691	7473	28.5118	0.0000
SUPPFOOD 1	-0.97829	1.007579	7473	-0.9709	0.3316
AGE12:SUPFOOD 1	0.92358	1.029172	7473	0.8974	0.3695
AGE14:SUPFOOD 1	0.747589	1.009926	7473	0.7402	0.4592
AGE16:SUPFOOD 1	0.856978	1.008695	7473	0.8496	0.3956
AGE18:SUPFOOD 1	0.841135	1.009098	7473	0.8336	0.4046
AGE110:SUPFOOD 1	0.848181	1.010103	7473	0.8397	0.4011
AGE112:SUPFOOD 1	0.784516	1.012876	7473	0.7745	0.4386

From the above output we can see that the intercept = 10.77796 is the mean evolution of upper arm circumference of infant for those Childs with out supplementary food at the reference age 0. The corresponding mean evolution of upper arm circumference of infants for those Childs with supplementary food is given by 10.778-0.979=9.799. The coefficients for those Childs with out supplementary food at age 2, age 4,age 6,age 8,age 10, and age 12 are 1.969, 2.8595, 3.1099, 3.0869, 2.9947, 3.0475 respectively. These values show the difference in mean evolution of upper arm circumference of infant between the reference age and the specified age level for those Childs with out supplementary food. For instance, the coefficient of age 2, which is 1.969 , is the

difference in mean evolution of upper arm circumference of infant between age 0 and age 2 for those Childs with out supplementary food.

The coefficients for the interaction terms refer to the difference between those Childs with out supplementary food and those Childs with supplementary food at deferent age levels.

From Table 4.15 we can produce the coefficients for those Childs with supplementary food and without.

Table 4.16 estimate of age level covariate with supplementary food and without

	Estimate for those Childs without Supplementary food	Estimate for those Childs with Supplementary food
Intercept	10.77773	10.77773+0.924=11.602
AGE12	1.968151	1.968151+0.748=2.716
AGE14	2.857342	3.604931
AGE16	3.109016	3.965994
AGE18	3.087466	3.928601
AGE110	3.002383	3.850564
AGE112	3.048194	3.83271

Therefore small p-value for age 2, age 4, ….,age 12 indicates that there is significance difference in mean evolution of upper arm circumference of infants between the reference age 0 and age 2, age 4, ….,age 12. and for the interaction term large p-value indicates that there is no significance difference in mean evolution of upper arm circumference of infants between infants those with supplementary food and with out supplementary food at each age levels.

4.**5.5.3 Gender difference for supplementary food behavior**

Table 4.17 Gender difference for supplementary food behavior

	Value	Std.Error	DF	t-value	p-value
(Intercept)	12.79165	0.044498	7483	287.4626	0.0000
SEXMALE	0.228847	0.061707	7483	3.70861	0.0002
SUPPFOOD11	0.202126	0.041131	7483	4.91424	0.0000
SEXMALE:SUPPFOOD11	0.051595	0.059302	7483	0.87004	0.3843

Large p-value for the interaction SEXM and SUPPFOOD11 indicates that there is is no significance difference in mean evolution of upper arm circumference of infants between genders.

If we see the plot of mean evolution of upper arm circumference of infants versus infants those with supplementary food and with out supplementary food there is no difference in mean evolution of upper arm circumference of infants between genders.
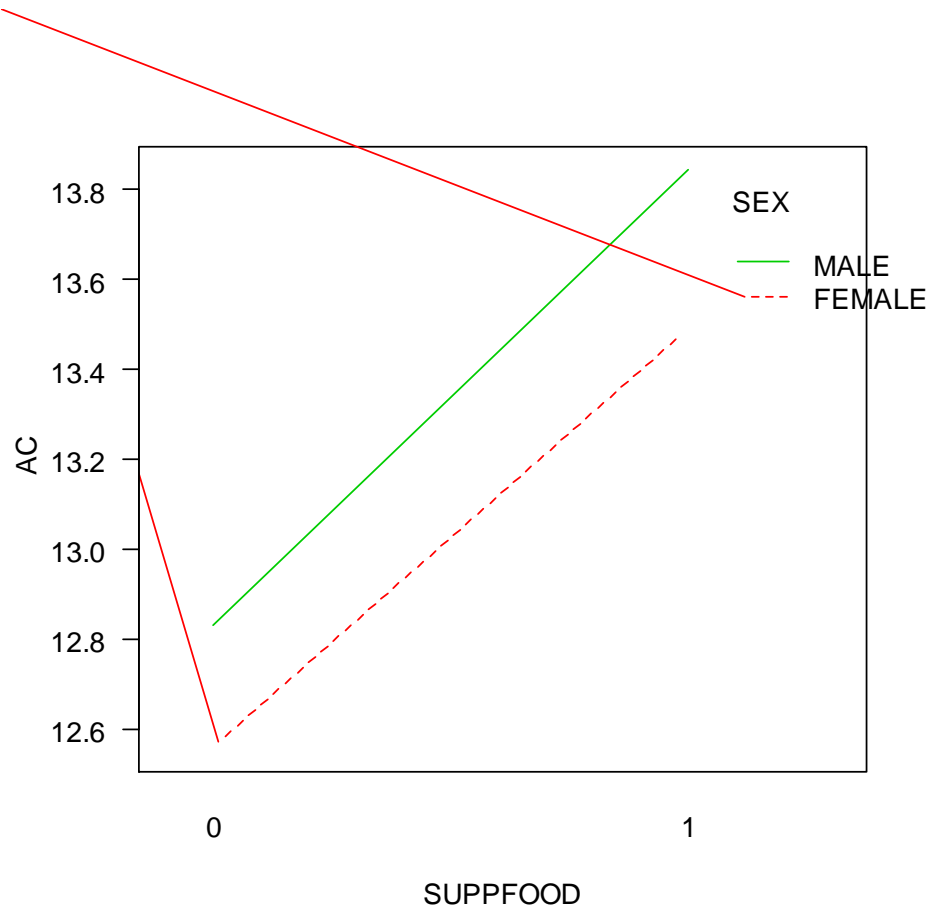

Figure 4.15 Plot of mean evolution of upper arm circumference for supplementary food behavior

4.5.6 Bayesian analysis of the Linear mixed effects models

In the analysis of Bayesian implementation of the Linear mixed effects models we run the Gibbs sampler defined with 60,000 iterations in two different chains, 10000 burn-in terms discarded, as to obtain 50000 samples from the full posterior distribution. Of which seven covariates which were significant in the likelihood approach were also statistically significant in the Bayesian implementation of the Linear mixed effect models. History plot plots out a complete trace for the variable and can be used for diagnosing the convergence of parameter estimates in Bayesian analysis. The package gives the plot by making iteration number on the x-axis and parameter value on the y-axis for each parameter. For all parameters, the plots of the last 40000 iterations for two independently generated chains demonstrated well "chain mixture" an indication of convergence. Figure below shows history plot for the intercept and other seven parameters. The Time series plots (trace) show that the chains with two different colours overlap one over the other. Hence, we are reasonably confident that convergence has been achieved.
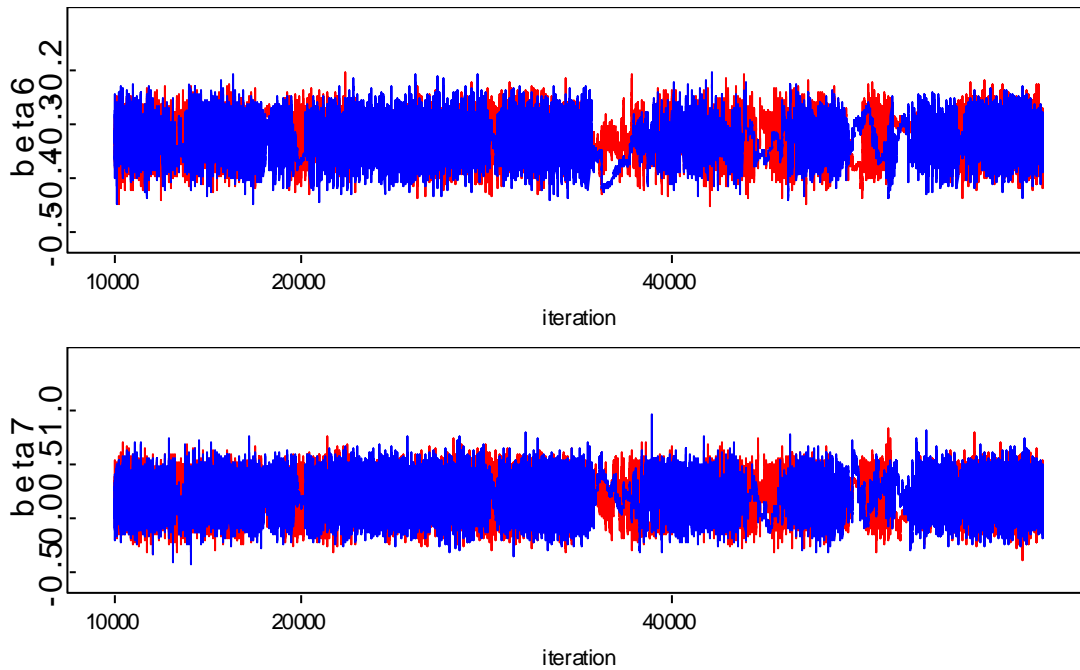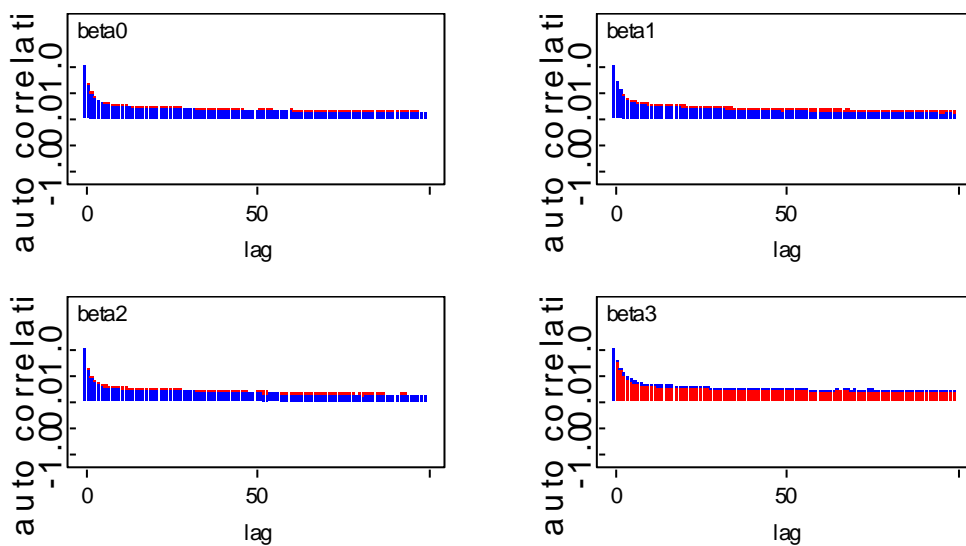
54
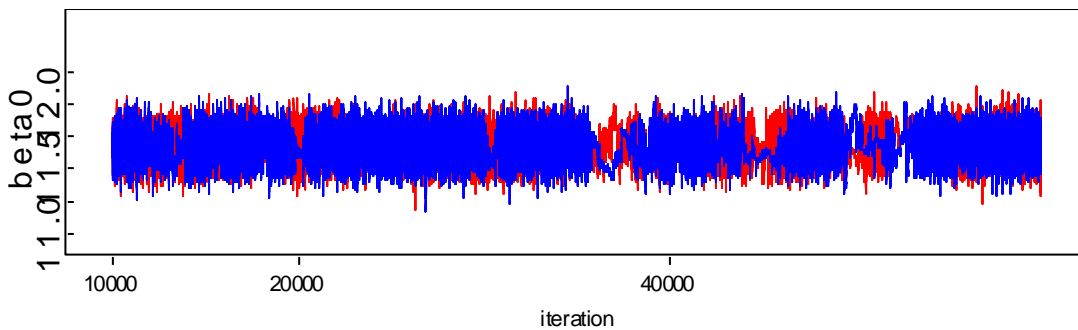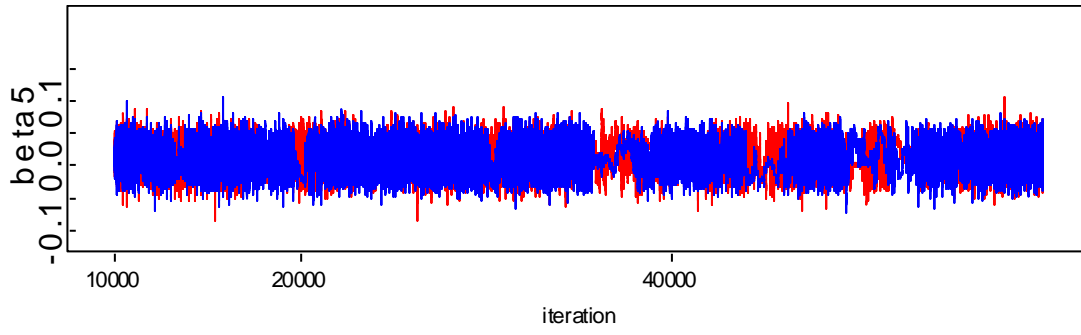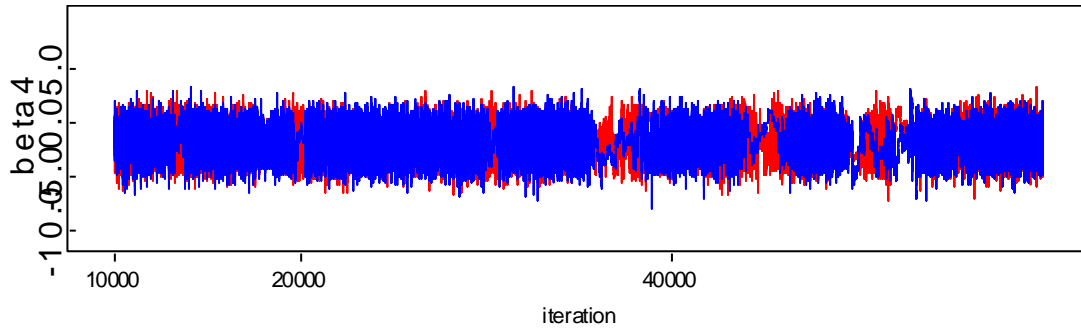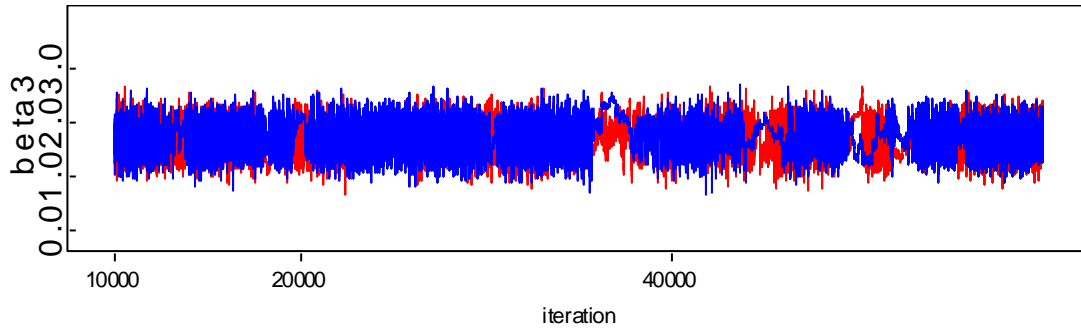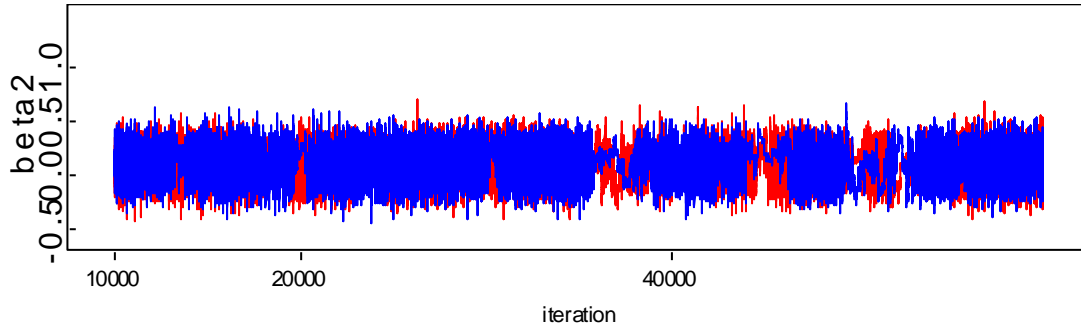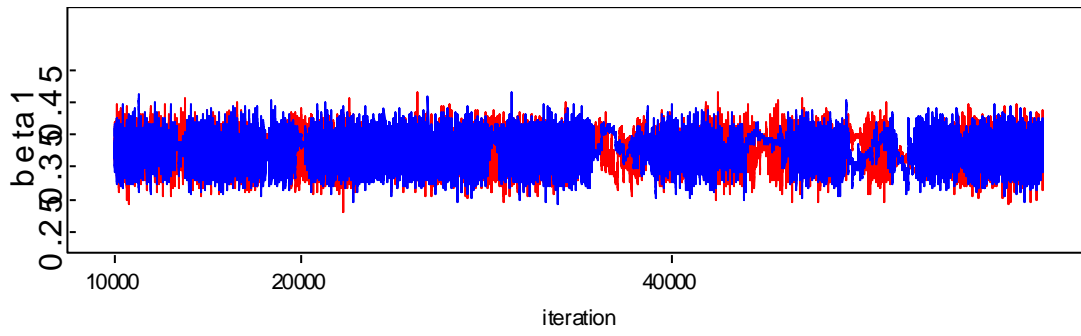
Figure 4.16 history plots for the seven parameter

Autocorrelation plot is the other recommended test for convergence of a Bayesian analysis. The plots show that the two independent chains were mixed or overlapped to each other and died out for higher lags and hence this is an evidence of convergence.
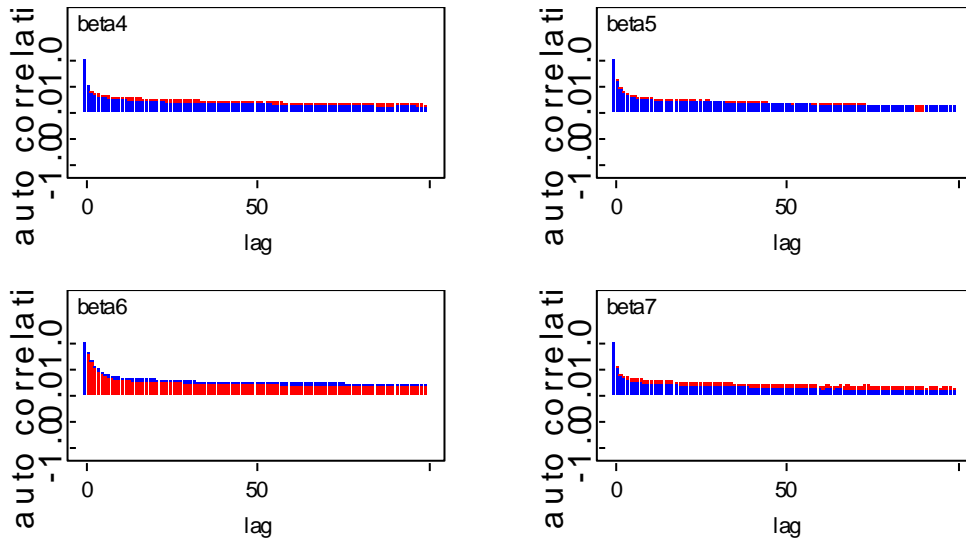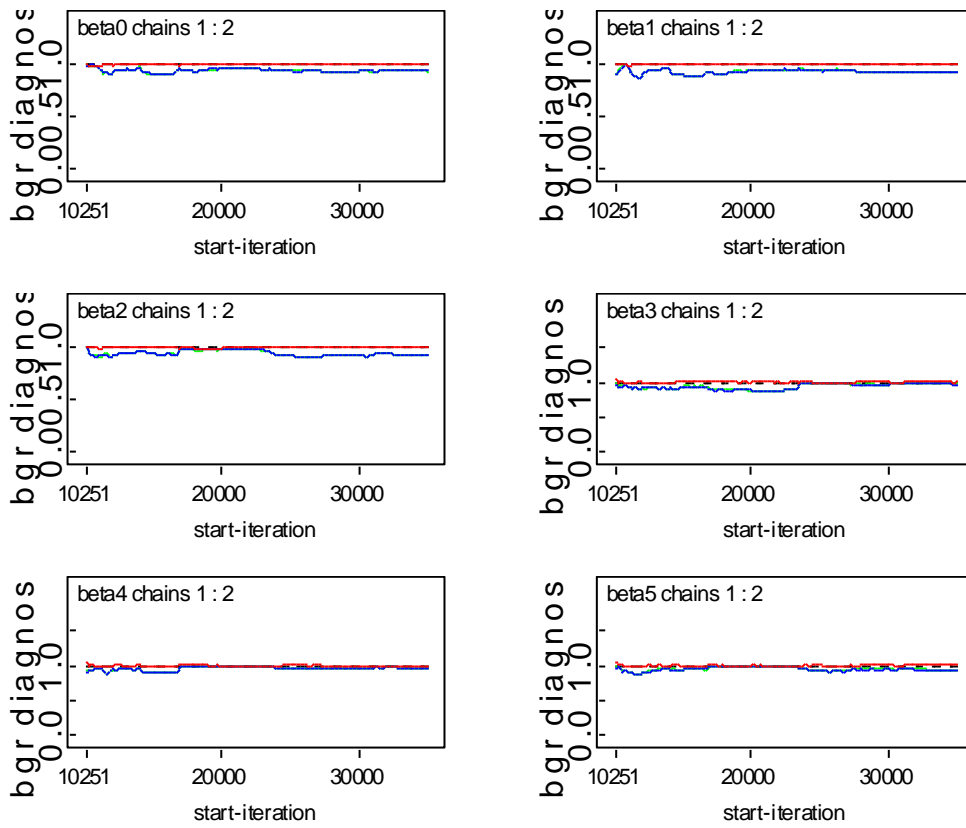
Figure 4.17 Autocorrelation plots for all seven parameter.

Gelman-Rubin statistic (GR) is also another way of checking convergence in Bayesian analysis.
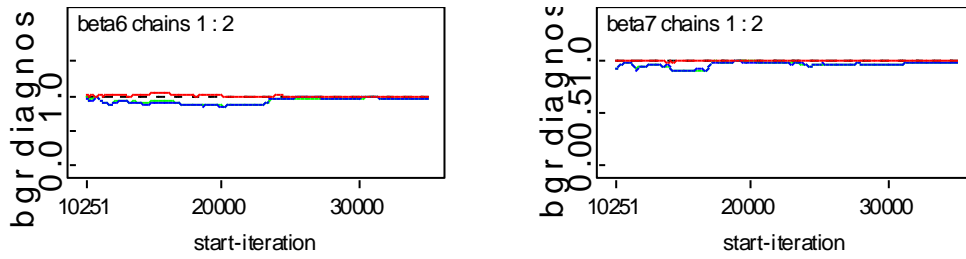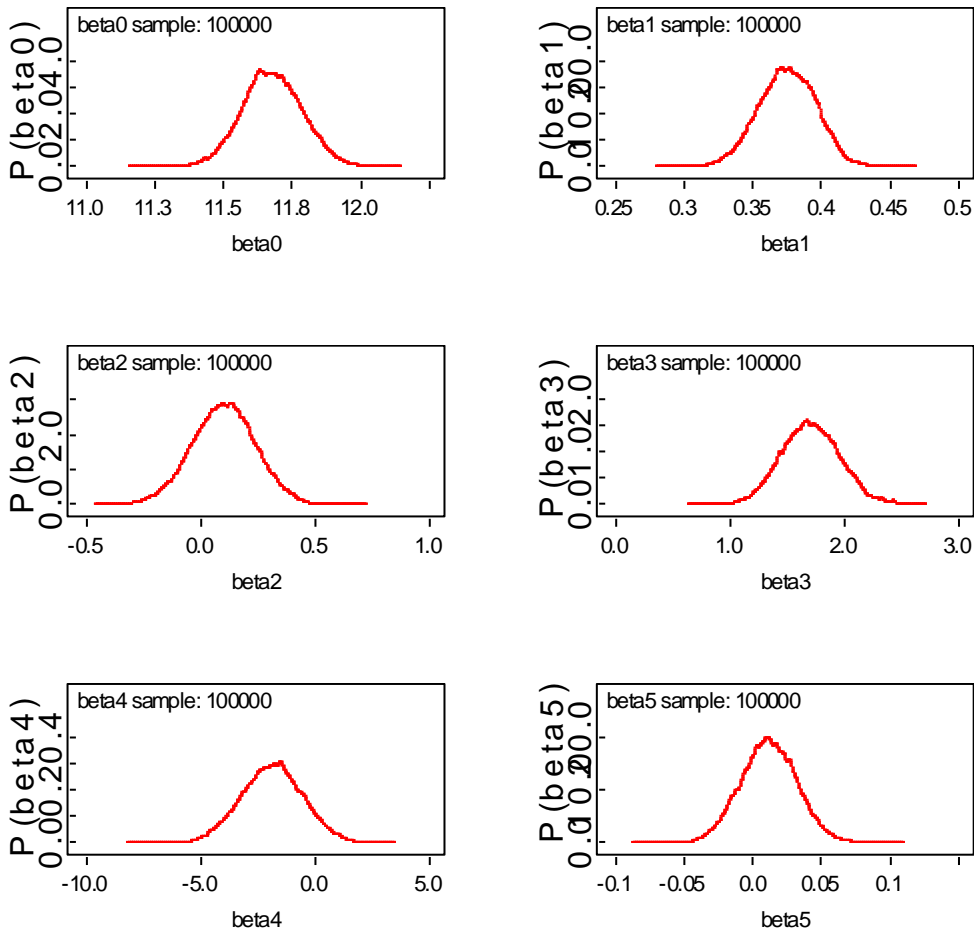
Figure 4.18 plot of Gelman-Rubin statistic for all parameter.

Density plot is another recommended technique for identifying non convergence.
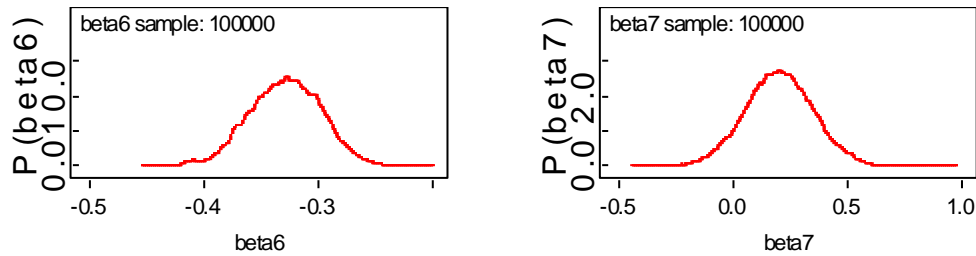
Figure 4.19 density plot for all parameter.

Deviance information criterion (DIC) can be used to assess model complexity and compare different models. It is important to note that DIC assumes the posterior mean to be a good estimate of the stochastic parameters. The deviance information criterion (DIC) is the generalization of the AIC for Bayesian model fitting using MCMC methods (Spiegel halter et al 2002). Dbar is the posterior mean of the deviance.

Table 4.18 Deviance information criterion (DIC)

	Dbar	*Dhat*	*DIC*	*pD*
AC	2431.0	81930.0	-77060.0	-79490.0
Total	2431.0	81930.0	-77060.0	-79490.0

One way to assess the accuracy of the posterior estimates is by calculating the Monte Carlo error for each parameter. To have accurate posterior estimates the simulation should be run until the Monte Carlo error for each parameter of interest is less than about 5% of the sample standard deviation. The output in Table below shows that all covariates have Monte Carlo errors much less than 0.05, and hence evidence for accuracy of posterior estimates in the Bayesian implementation of the Linear mixed effect models. Once convergence and accuracy of posterior estimates are maintained, summarizing the posterior statistics is possible. As a result the summary statistics of the posterior distributions are presented below.

Table: 4.19 Summary statistics of the posterior distributions of the model parameters

	Mean	standard deviation	MC-error	lower CI interval	median	upper CI interval
β_0	11.68	0.1082	0.002168	11.47	11.67	11.89**
β_1	0.3749	0.02072	4.195E-4	0.3338	0.3751	0.4138**
β_2	0.1991	0.1376	0.002734	0.1739	0.1994	0.3676**
β_3	1.704	0.2538	0.005756	1.221	1.699	2.208**
β_4	-1.895	1.331	0.02621	-4.497	-1.884	0.6962
β_5	0.01138	0.02026	3.973E-4	-0.02807	0.01132	0.05107
β_6	-0.3289	0.03131	7.158E-4	-0.3905	-0.3282	-0.2697**
β_7	0.2009	0.1472	0.002823	-0.08818	0.2006	0.4922
sigma	0.9968	0.5665	0.02594	0.03571	1.362	1.521
$\mathbf{b_{oi}}$	0.997	0.4234	0.04248	0.1394	1.132	1.484
$\mathbf{b_{1i}}$	0.04773	0.01239	0.00138	0.02666	0.04615	0.07345

The above table contains the estimated coefficients, the standard errors (sd), Monte Carlo (MC) errors, and credible interval. The significant results are indicated by (**) which indicate the importance of the predictor variables in the model since the 95% credible intervals do not contain zero. sigma, alpha.tau, and beta.tau are the residual standard deviation and the random effects standard deviation respectively and their results of model errors suggest that accounting for significant skewness, when the data exhibit skewness, provides a better model fit to the data and gives more accurate estimates to the parameters.

4.5.7 Comparison among the likelihood and Bayesian approach

The purpose of the present study is to compare the mean evolution of upper arm circumference of infant between genders, does mean evolution of upper arm circumference of infant changes over time and to compare mean evolution of upper arm circumference of infant differ for those infant given supplementary food or not using linear mixed effect model. Findings from the present study demonstrated that the ML estimates of the random-effects standard deviations are smaller than the corresponding REML estimates which is different result from the Bayesian. The estimated within group residual standard deviations are identical and also similar with result of the Bayesian techniques. In general, the fixed-effects estimates obtained using ML, REML and Bayesian techniques are similar. Inferences regarding the fixed effects are essentially the same for the two estimation methods.

Overall without looking each age level the mean evolution of the upper arm circumference of infant for boys and girls is not different and also the mean evolution of the upper arm circumference of infant increases as the infants age increases. For those infants given supplementary food and for infants without supplementary food their mean evolution of the upper arm circumference is not different and also there is no gender difference in mean evolution. But if we look the mean evolution at different age level there is significance difference in mean evolution of upper arm circumference of infants between males and females at age 2, age 4, age 6. There is no significant difference in mean evolution of upper arm circumference of infants between males and females at age 8, 10, 12 respectively.

In Bayesian techniques the coefficient estimate of the parameter is almost the same with the likelihood approach except the insignificant parameters and the interpretation of the fixed effect is the same.

CHAPTER 5

5 Conclusions

This paper is concerned in presenting the linear mixed effect model for a longitudinal Gaussian data using likelihood and Bayesian techniques. The main aim is to compare likelihood and Bayesian techniques as an alternative approach Gaussian longitudinal response. Within this frame work first exploratory analysis techniques is discussed to visualize patterns in the data. And then different variance function that are used to specify the within-group linear mixed effect model and different correlation structures in selection of an appropriate residual covariance structure, especially in the presence of random effect are used to select the best model. The result indicates the ML estimates of the random-effects standard deviations are smaller than the corresponding REML estimates. The estimated within group residual standard deviations are identical. In general, the fixed-effects estimates obtained using ML and REML are similar. Inferences regarding the fixed effects are essentially the same for the two estimation methods.

In Bayesian techniques the coefficient estimate of the parameter is almost the same with the likelihood approach except the insignificant parameters and the same residual standard deviation is obtained with different result of the random effect. The interpretation of the fixed effect is the same.

Numerous researches finding support the present finding that the mean evolution of the upper arm circumference of infant is not different for boys and girls and also the mean evolution of the upper arm circumference of infant increases as the infants' age increases. For those infants given supplementary food and for infants without supplementary food their mean evolution of the upper arm circumference is not different and also there is no gender difference in mean evolution.

REFERENCES

- Man M. Laird; James H. Ware (1982). Random – effects models for longitudinal data. Biometrics, Vol 38, No.4: 963-974.

- Verbeke, G. and Molenberghs, G. (2000). Linear mixed models for longitudinal data. Springer-verlag, New York.

- Henderson, C.R. (1975). Best linear unbiased estimation and prediction under a selection model. Biometrics, 31:423-447.

- Searle, S.R. (1995). An overview of variance components estimation. Metrica, 42: 215-230.

- Searle, S.R.,Casella, G., and McCulloch, C.E. (2006).Variance Components. John Wiley & Sons, Inc., New York.

- Self, S.G. and Liang, K.Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under non standard conditions. Journal of the American statistical association, 82: 605-610.

- Serfling, R.J. (1980). Approximation theorems of mathematical statistics. John Wiley & Sons, Inc., New York.

- Andrew Gelman, john B. Carlin, Hals. Stern and Donald B. Rubin. Bayesian data analysis. 2nd edition.

- Best N G, Cowles M K and Vines S K (1997) *CODA: Convergence diagnosis and output analysis software for Gibbs sampling output, Version 0.4*. MRC Biostatistics Unit, Cambridge:

- Brooks S P (1998) Markov chain Monte Carlo method and its application. *The Statistician*. **47**, 69-100.

- Brooks S P and Gelman A (1998) Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*. **7**, 434-455.

- Carlin B P and Louis T A (1996) *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman and Hall, London, UK.

- Congdon P (2001) *Bayesian Statistical Modelling*. John Wiley & Sons, Chichester, UK.

- Gilks W R, Richardson S and Spiegelhalter D J (Eds.) (1996) *Markov chain Monte Carlo in Practice*. Chapman and Hall, London, UK.

- Roberts G O (1996). Markov chain concepts related to sampling algorithms. In W R Gilks, S Richardson and D J Spiegelhalter (Eds.) *Markov chain Monte Carlo in Practice*. Chapman and Hall, London, UK.

- Spiegelhalter D J, Best N G, Carlin B P and van der Linde A (2002) Bayesian measures of model complexity and fit (with discussion). *J. Roy. Statist. Soc. B*. **64**, 583-640.

APPENDIX: WINBUGS CODE

In this Appendix we list the WinBUGS code used for the data analyses of Section 5.6

```
model
{
for( i in 1 : 8965) {
for( j in  Id[i] :
(Id[i+1]-1) )
      {
AC[j] ~ dnorm(mu[ j],tau.c)
mu[j]  <- beta0 +  beta1*AGE[j]+ beta2*SEX[j]+  beta3*SUPPFOOD[j]
+beta4*POWDERMI[j]+beta5*AGE[j]*SEX[j]+beta6*AGE[j]*SUPPFOOD[j]+beta7*A
GE[j]*POWDERMI[j]+alpha[i]+beta[i]*AGE[j]
}
alpha[i] ~ dnorm(0,alpha.tau)
beta[i] ~ dnorm(0,beta.tau)


                }
                tau.c~dgamma(0.001,0.001)
                alpha.tau ~ dgamma(0.001,0.001)
                sigma <- 1 / sqrt(tau.c)
                boi<-1/sqrt(alpha.tau)
                b1i<-1/sqrt(beta.tau)
                beta.tau ~ dgamma(0.001,0.001)
                beta0 ~ dnorm(0.0,1.0E-6)
                beta1 ~dnorm(0.0,1.0E-6)
                beta2 ~ dnorm(0.0,1.0E-6)
                beta3 ~ dnorm(0.0,1.0E-6)
                beta4 ~ dnorm(0.0,1.0E-6)
                beta5 ~ dnorm(0.0,1.0E-6)
                beta6 ~ dnorm(0.0,1.0E-6)
                beta7 ~ dnorm(0.0,1.0E-6)
```

```
        }
```

list(beta0=0, beta1=0, beta2=0, beta3=0, beta4=0, beta5=0, beta6=0, beta7=0, alpha.tau = 1,beta.tau=1,tau.c=1)

list(beta0=0.01, beta1=0.01, beta2=0.01, beta3=0.01, beta4=0.01, beta5=0.01, beta6=0.01, beta7=0.01,alpha.tau = 1,beta.tau=1,tau.c=1)