

Random Effects

- We need to understand (at least qualitatively) what are the likely sources of random variation
- One possible source is Random Effects, when units are sampled at random from a population and various aspects of their behavior may show stochastic variation between units
- We introduce Linear Random Effects model where
 - the response is assumed to be a linear function of explanatory variables with regression coefficients that vary from one individual to the next
 - variability reflects natural heterogeneity due to unmeasured factors

Example: Children birth weight and growth rate.

- A random effects model is a reasonable description if the set of coefficients from a population of children can be thought of as a sample from a distribution
- Given the actual coefficient for a children, the linear Random Effects model assumes that repeated observations for that person are independent
- Correlation arises because we cannot observe the underlying growth curve, that is the regression coefficient, but we have only imperfect measurements of weight on each infant
- So the model takes the form

$$E(Y_{ij}|U_i) = (\beta_0 + U_i) + \beta_1(\text{time})_{ij}$$

- Typically, a parametric model such as Gaussian with mean=0 and unknown variance ν^2 is used for U_i .

Linear Mixed Models

- The Usual Linear Model

$$y = X\beta + e,$$

where

- $y = (y_1, \dots, y_n)'$ is an $n \times 1$ vector of independent observations
- β is a $p \times 1$ vector of unknown parameters
- X an $n \times p$ design (model) matrix
- $e = (e_1, \dots, e_n)'$ is an $n \times 1$ vector of independent errors

- The linear mixed model (general)

$$Y_i = X_i\beta + Z_ib_i + e_i,$$

where

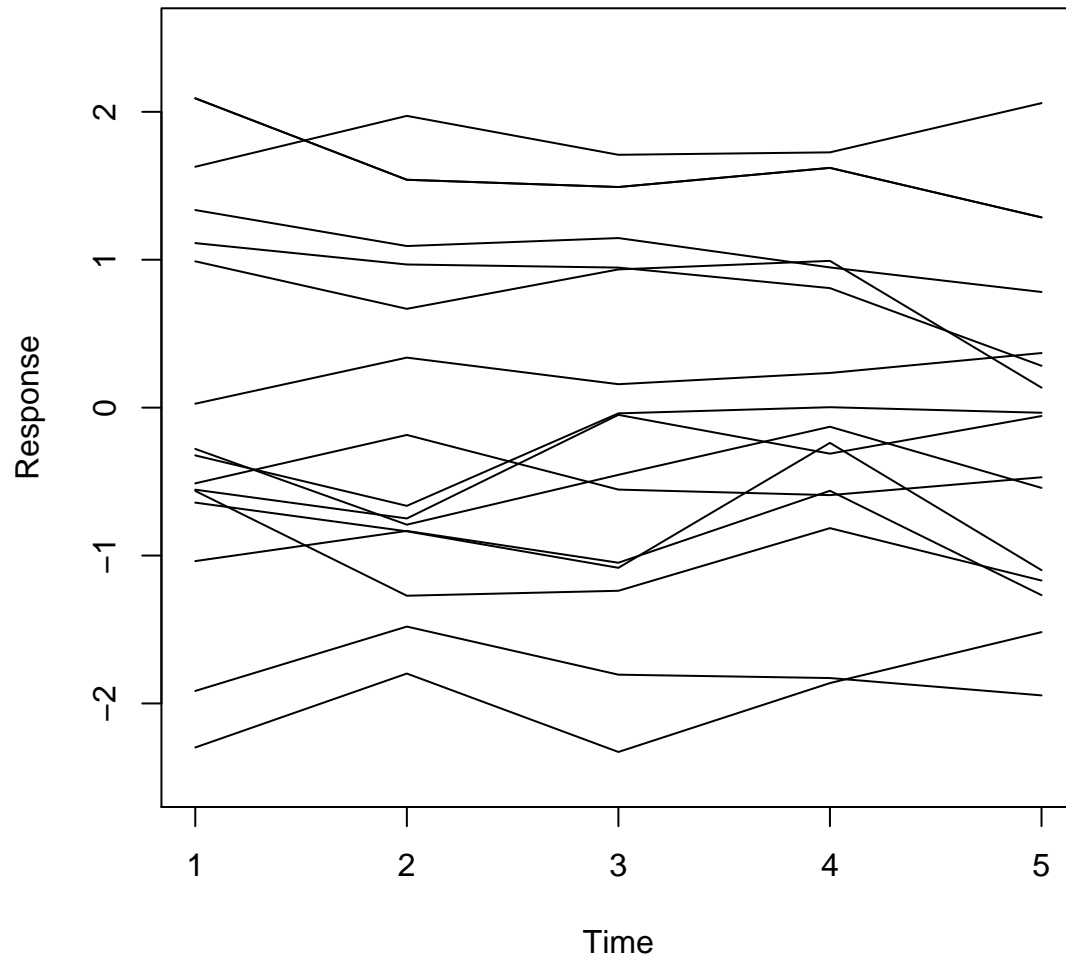
- Y_i , β and e as before with
 - * $E(e_i) = 0_n$
 - * $Var(e_i) = W$
- Matrix Z is a given $n \times q$ matrix (the columns of Z is a subset of the columns of X)
- b_i is an unobservable random vector of dimensions $q \times 1$, following (theoretically) any multivariate distribution with the following assumptions
 - * $E(b_i) = 0_q$
 - * $Var(b_i) = B$
- In practice b_i follow a multivariate normal distribution.
- In addition, vectors b_i and e_i are assumed uncorrelated.
- $E(Y_i) = X_i\beta$
- $Var(Y_i) = Var(X\beta + Zb + e) = ZBZ' + W$.

Random Intercept Model

Consider the model

$$\begin{aligned} Y_{ij} &= X'_{ij}\beta + b_i + e_{ij} \\ &= (\beta_1 + b_i) + X_{ij2}\beta_2 + \dots + X_{ijp}\beta_p + e_{ij} \end{aligned}$$

- Each subject's profile appears flat (across occasions) - [or parallel]
- Observations Y_{ij} vary around a different value for each subject. These values are the intercepts of the line each subject's responses vary around, where b_i represents the deviations of subject's i intercept from the population one (β_1).
- The set of intercepts are a sample from the population of intercepts.
- This implies that there is *between-subject variability* (equivalent to *within-subject correlation*)



- Furthermore, the variance of Y_{ij} takes the form

$$\begin{aligned} \text{Var}(Y_{ij}) &= \text{Var}(X'_{ij}\beta + b_i + e_{ij}) \\ &= \text{Var}(b_i) + \text{Var}(e_{ij}) \\ &= \sigma_b^2 + \sigma^2 \end{aligned}$$

and the covariance between any pair of observations of the same subject

$$\begin{aligned} \text{Cov}(Y_{ij}, Y_{ik}) &= \text{Cov}(X'_{ij}\beta + b_i + e_{ij}, X'_{ik}\beta + b_i + e_{ik}) \\ &= \text{Cov}(b_i, b_i) \\ &= \sigma_b^2. \end{aligned}$$

The covariance matrix then becomes

$$\text{Cov}(Y_i) = \begin{pmatrix} \sigma_b^2 + \sigma^2 & \sigma_b^2 & \sigma_b^2 & \cdots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 + \sigma^2 & \sigma_b^2 & \cdots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 + \sigma^2 & \cdots & \sigma_b^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \cdots & \sigma_b^2 + \sigma^2 \end{pmatrix},$$

and the correlation between two observations becomes

$$\rho = \text{Corr}(Y_{ij}, Y_{ik}) = \frac{\sigma_b^2}{\sigma_b^2 + \sigma^2}.$$

- The presence of random effect induce correlation among repeated measurements. This is also known as *intra-class correlation*.

Note: In statistics, the intraclass correlation is a descriptive statistic that can be used when quantitative measurements are made on units that are organized into groups. It describes how strongly units in the same group resemble each other. While it is viewed as a type of correlation, unlike most other correlation measures it operates on data structured as groups, rather than data structured as paired observations.

- The model

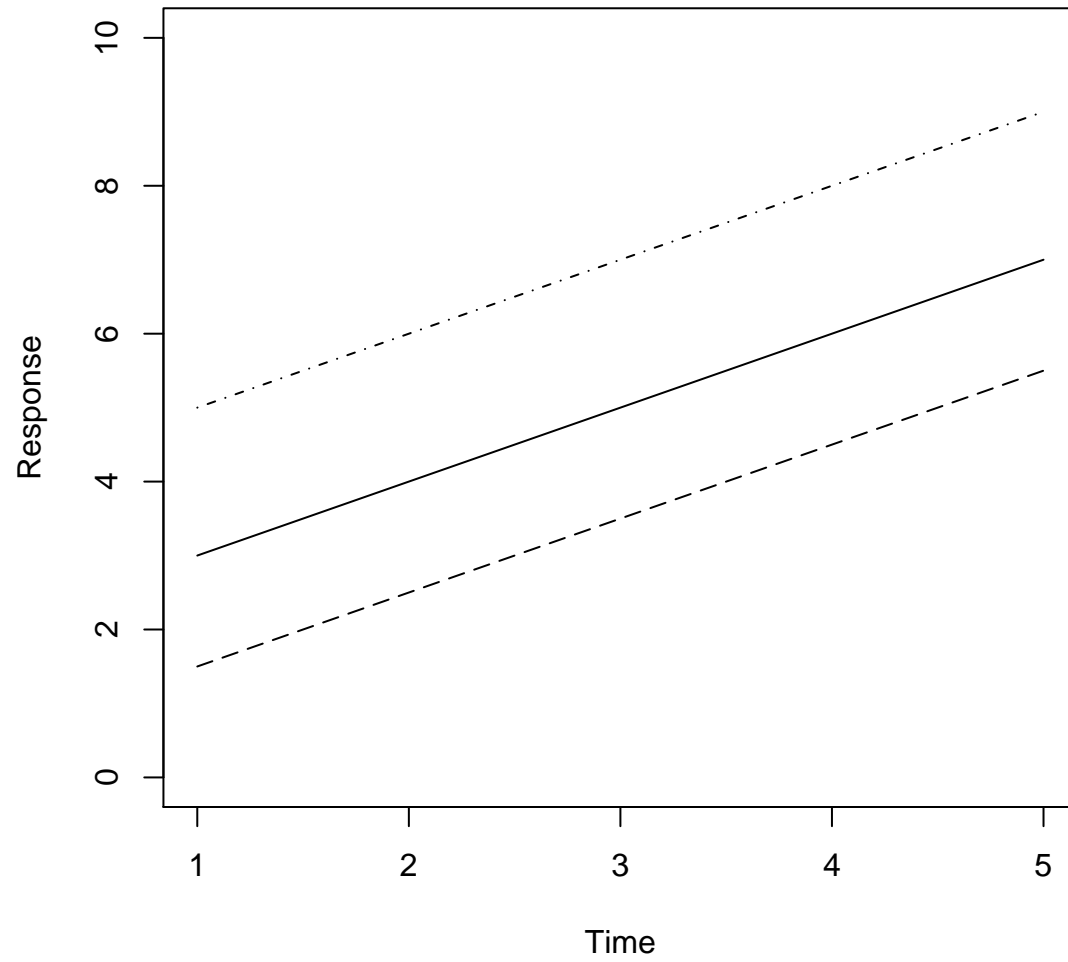
$$E(Y_{ij}|b_i) = X'_{ij}\beta + b_i$$

is referred to as the *conditional* or *subject specific* mean model

- The model

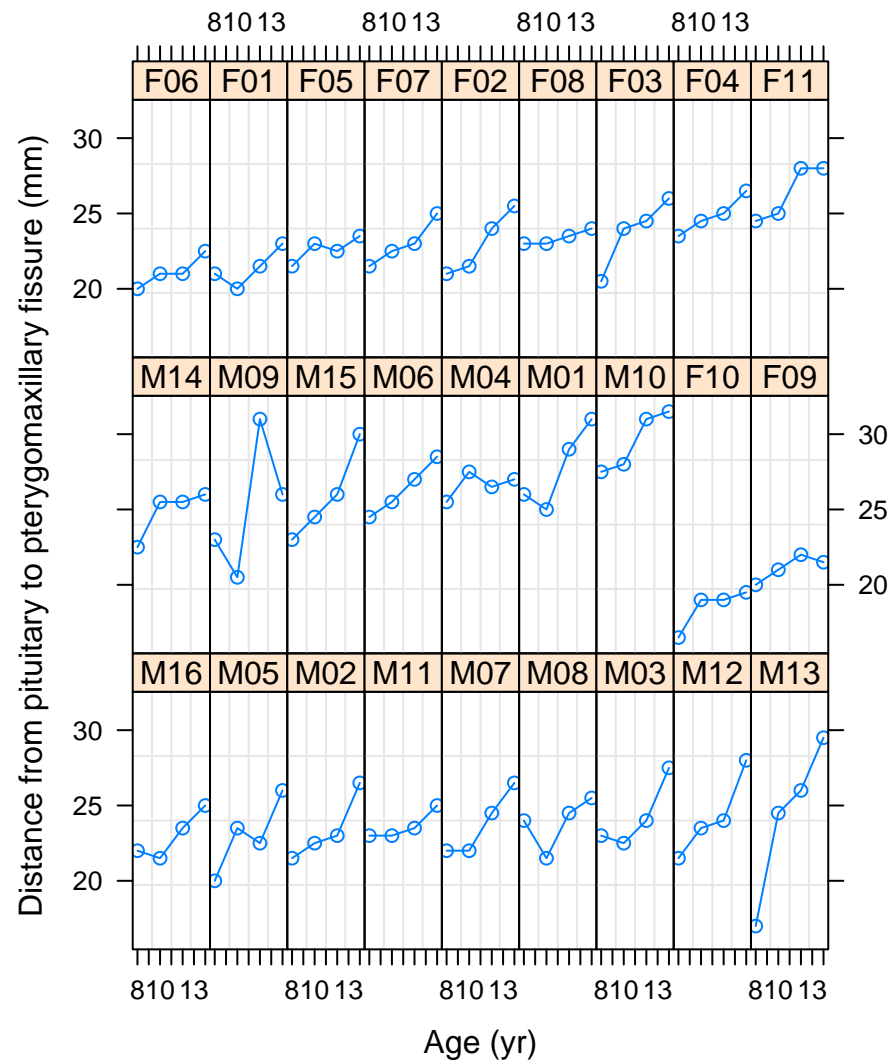
$$E(Y_{ij}) = X'_{ij}\beta$$

is referred to as the *marginal* or *population averaged* mean model



Example: Orthodont Data [included in nlme package]

- A set of measurements of the distance from the pituitary gland to the pterygomaxillary fissure taken every 2 years.
- Measurements taken from 8 till 14 years of age.
- We have 27 children: 16 males - 11 females
- Data collected from x-rays.



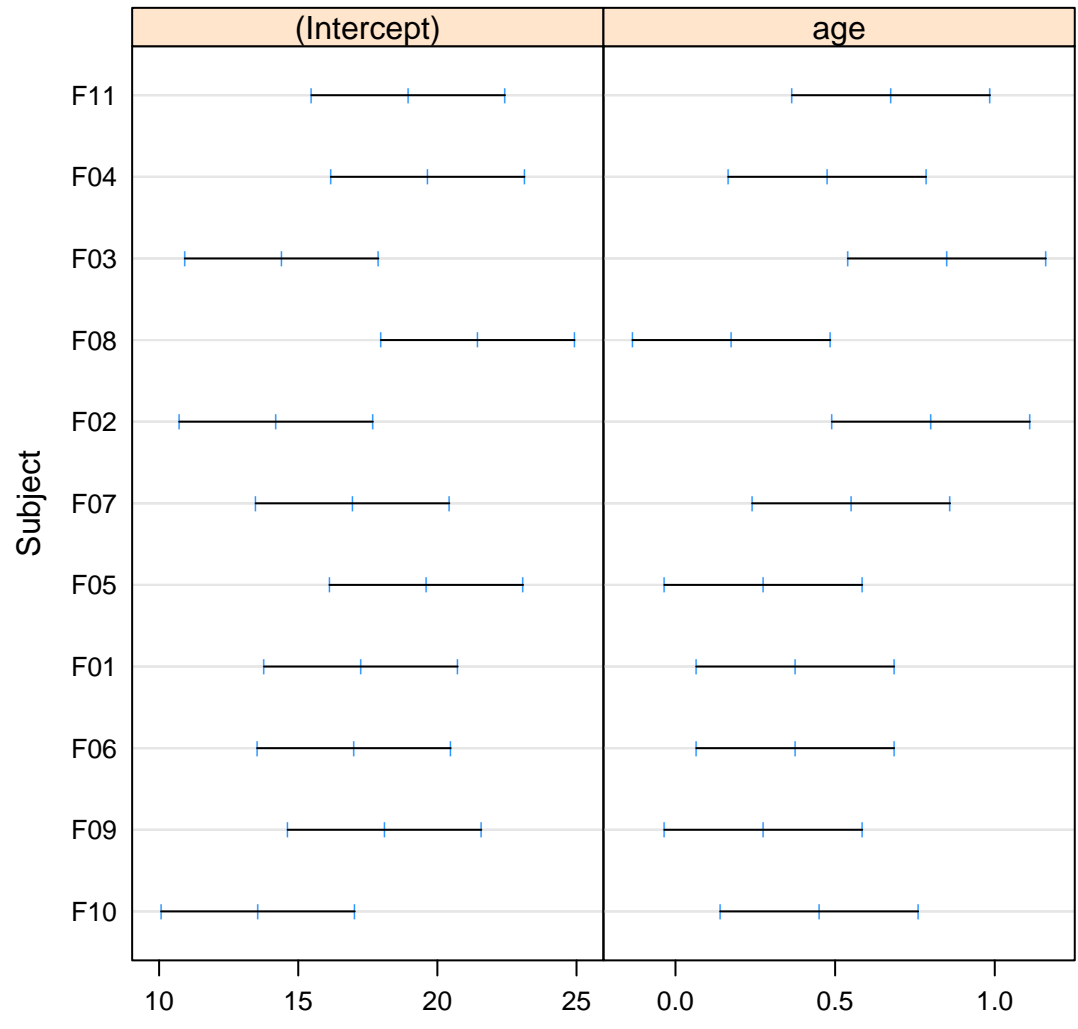
```
> levels(Orthodont$Sex)
[1] "Male" "Female"
> OrthoFem=Orthodont[Orthodont$Sex=="Female",]
> lmF=lmList(distance ~ age, data=OrthoFem)
> coef(lmF)
      (Intercept)      age
F10           13.55  0.450
F09           18.10  0.275
F06           17.00  0.375
F01           17.25  0.375
F05           19.60  0.275
F07           16.95  0.550
F02           14.20  0.800
F08           21.45  0.175
F03           14.40  0.850
F04           19.65  0.475
F11           18.95  0.675
```

```
> intervals(lmF)
, , (Intercept)
```

	lower	est.	upper
F10	10.07138	13.55	17.02862
F09	14.62138	18.10	21.57862
F06	13.52138	17.00	20.47862
F01	13.77138	17.25	20.72862
F05	16.12138	19.60	23.07862
F07	13.47138	16.95	20.42862
F02	10.72138	14.20	17.67862
F08	17.97138	21.45	24.92862
F03	10.92138	14.40	17.87862
F04	16.17138	19.65	23.12862
F11	15.47138	18.95	22.42862

```
, , age
```

	lower	est.	upper
F10	0.14009962	0.450	0.7599004
F09	-0.03490038	0.275	0.5849004
F06	0.06509962	0.375	0.6849004
F01	0.06509962	0.375	0.6849004
F05	-0.03490038	0.275	0.5849004
F07	0.24009962	0.550	0.8599004
F02	0.49009962	0.800	1.1099004
F08	-0.13490038	0.175	0.4849004
F03	0.54009962	0.850	1.1599004
F04	0.16509962	0.475	0.7849004
F11	0.36509962	0.675	0.9849004

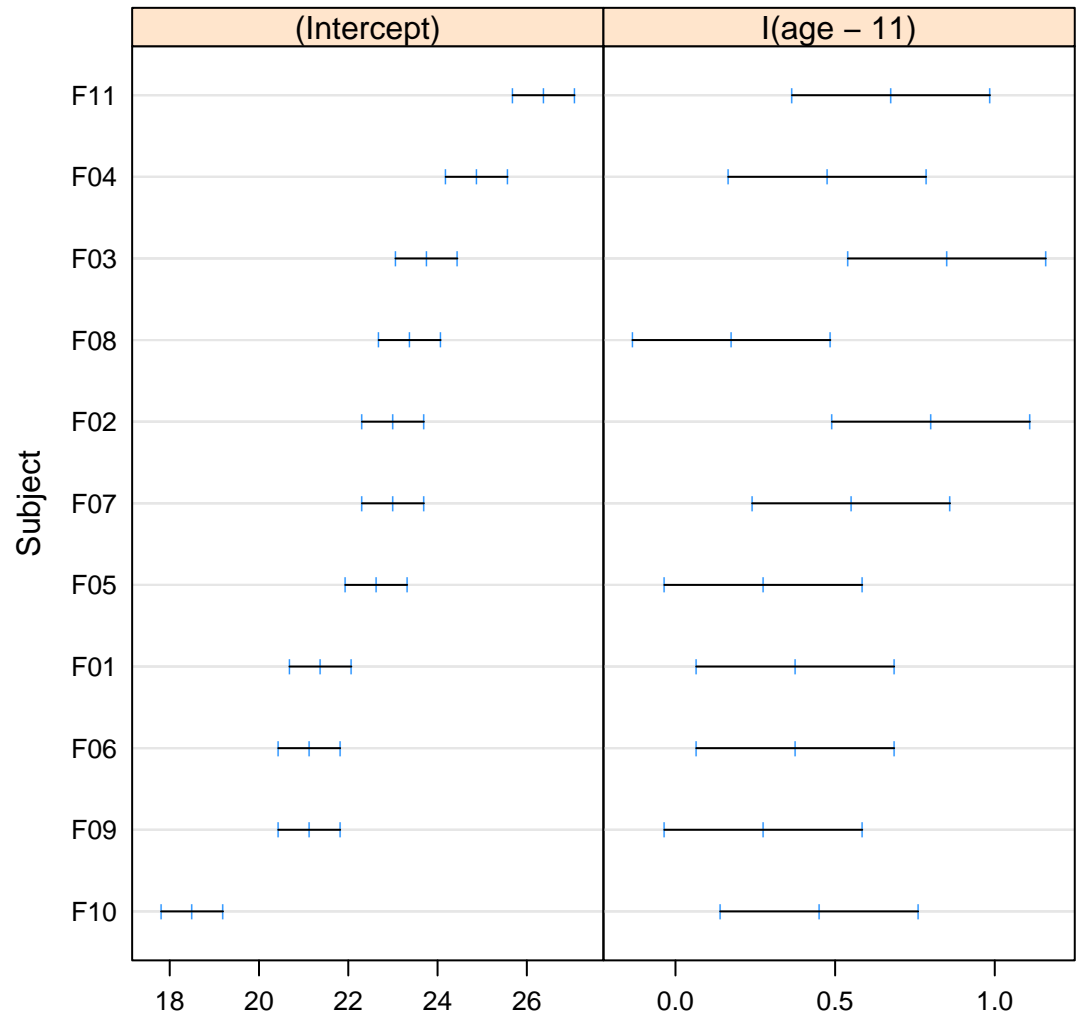


```
> lmF2=update(lmF,distance~I(age-11))
> intervals(lmF2)
, , (Intercept)
```

	lower	est.	upper
F10	17.80704	18.500	19.19296
F09	20.43204	21.125	21.81796
F06	20.43204	21.125	21.81796
F01	20.68204	21.375	22.06796
F05	21.93204	22.625	23.31796
F07	22.30704	23.000	23.69296
F02	22.30704	23.000	23.69296
F08	22.68204	23.375	24.06796
F03	23.05704	23.750	24.44296
F04	24.18204	24.875	25.56796
F11	25.68204	26.375	27.06796

```
, , I(age - 11)
```

	lower	est.	upper
F10	0.14009962	0.450	0.7599004
F09	-0.03490038	0.275	0.5849004
F06	0.06509962	0.375	0.6849004
F01	0.06509962	0.375	0.6849004
F05	-0.03490038	0.275	0.5849004
F07	0.24009962	0.550	0.8599004
F02	0.49009962	0.800	1.1099004
F08	-0.13490038	0.175	0.4849004
F03	0.54009962	0.850	1.1599004
F04	0.16509962	0.475	0.7849004
F11	0.36509962	0.675	0.9849004



```
> lmeF=lme(distance~age,data=OrthoFem,random=~1)# Using REML
> summary(lmeF)
Linear mixed-effects model fit by REML
Data: OrthoFem
      AIC      BIC    logLik
149.2183 156.169 -70.60916

Random effects:
Formula: ~1 | Subject
      (Intercept) Residual
StdDev:      2.06847 0.7800331

Fixed effects: distance ~ age
              Value Std.Error DF   t-value p-value
(Intercept) 17.372727 0.8587419 32 20.230440    0
age          0.479545 0.0525898 32  9.118598    0
Correlation:
(Intr)
age -0.674

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-2.2736479 -0.7090164  0.1728237  0.4122128  1.6325181

Number of Observations: 44
Number of Groups: 11
```

```
> lmeF0=lme(distance~I(age-11),data=OrthoFem,random=~1)
> summary(lmeF0)
Linear mixed-effects model fit by REML
Data: OrthoFem
      AIC      BIC    logLik
149.2183 156.169 -70.60916

Random effects:
Formula: ~1 | Subject
      (Intercept) Residual
StdDev:      2.06847 0.7800331

Fixed effects: distance ~ I(age - 11)
              Value Std.Error DF t-value p-value
(Intercept) 22.647727 0.6346568 32 35.6850    0
I(age - 11)  0.479545 0.0525898 32  9.1186    0
Correlation:
      (Intr)
I(age - 11) 0

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-2.2736479 -0.7090164  0.1728237  0.4122128  1.6325181

Number of Observations: 44
Number of Groups: 11
```

Random Intercept and Slope Model

Consider the model

$$Y_{ij} = (\beta_1 + b_{1i}) + (\beta_2 + b_{2i})t_{ij} + e_{ij}.$$

- Each subject varies with respect
 - (i) baseline level when $t_{i1} = 0$ and
 - (ii) rate of change of response over time.
- In this particular case we have $q = p = 2$ and

$$X_i = Z_i = \begin{pmatrix} 1 & t_{i1} \\ 1 & t_{i2} \\ \vdots & \vdots \\ 1 & t_{in_i} \end{pmatrix}.$$

- Additionally, consider the variance

$$\begin{aligned}
 \text{Var}(Y_{ij}) &= \text{Var}(X'_{ij}\beta + Z'_{ij}b_i + e_{ij}) \\
 &= \text{Var}(Z'_{ij}b_i + e_{ij}) \\
 &= \text{Var}(b_{1i} + b_{2i}t_{ij} + e_{ij}) \\
 &= \text{Var}(b_{1i}) + 2t_{ij}\text{Cov}(b_{1i}, b_{2i}) + t_{ij}^2\text{Var}(b_{2i}) + \text{Var}(e_{ij}).
 \end{aligned}$$

and the covariance among the repeated observations of the same subject becomes

$$\text{Cov}(Y_{ij}, Y_{ik}) = \text{Var}(b_{1i}) + (t_{ij} + t_{ik})\text{Cov}(b_{1i}, b_{2i}) + t_{ij}t_{ik}\text{Var}(b_{2i}).$$

- Hence, the covariance matrix can be expressed as a function of time.

Covariance Structure

In the linear mixed model

$$Y_i = X_i\beta + Z_ib_i + e_i,$$

the matrix $W_i = Cov(e_i)$ introduces the covariance between the repeated observations when focusing on the conditional mean response profile of a specific individual. In other words, it is the covariance of the i^{th} individual's deviations from the response profile

$$E(Y_i|b_i) = X_i\beta + Z_ib_i.$$

- The usual assumption is $W = \sigma^2 I_n$. This is referred as the *conditional independence* assumption.
- The conditional covariance becomes

$$Cov(Y_i|b_i) = Cov(e_i) = W_i$$

- The marginal then takes the form

$$Cov(Y_i) = Z_i B Z_i' + W_i$$

- The $Cov(Y_i)$ allows for *between-subject* (B) and *within-subject* (W_i) sources of variation.
- Due to the fact that $Cov(Y_i)$ is a function of times of measurements (when time is in Z_i), in principle each subject may have its own measurement times.
- The comparison of random effects models for the covariance is based on the likelihood ratio test (REML). A test of two nested models, one with q and another one with $q + 1$ correlated random effects lead to a chi-square test on $q + 1$ df (1 for variance and q covariances). However, caution is needed when the null hypothesis is on the boundary of the parameter space.

Some Characteristics

- There is no need of balanced data.
- The covariances are functions of time. As a result, if time is included in Z_i , each patient can have his own sequence of measurement times. This property makes these models suitable for the analysis of *real life* longitudinal data.
- The number of covariance parameters that need to be estimated remains unchanged regardless of the number of measurements.
- The random effects covariance structure allows the variances and covariances to change (increase or decrease) as a function of measurement times, without introducing restrictive structures as the covariance pattern models do.

Prediction

- In the analysis of longitudinal data the interest in fixed effects β is obvious. The interpretation of the parameters is clear and associated with the mean response over time and changes in covariates.
- In many cases, however, subject-specific trajectories are of interest.
- Under the linear mixed-effects model patient specific response trajectories can be predicted/estimated.
- This is possible by obtaining predictions of the subject-specific effects b_i (random effects),
or

$$X_i\beta + Z_ib_i.$$

- Generally, the issue of predicting a random variable and as a result the patient specific response trajectory is that of predicting its conditional mean given the available data.
- There are two pieces of information that contribute in the estimation/prediction of b_i .

- The first is the statement that

$$b_i \sim N(0, B)$$

(the prior of b_i).

- The second is the likelihood of the data Y_i , which say that

$$Y_i | b_i \sim N(X_i \beta + Z_i b_i, W_i)$$

- We combine information by multiplying the two densities (joint) and ...after some maths... we get

$$E(b_i|Y_i) = BZ_i'\Sigma_i^{-1}(Y_i - X_i\hat{\beta}),$$

where $\Sigma_i = Cov(Y_i) = Z_iBZ_i' + W_i$. This is known as the BLUP.

- The predictor of b_i depends on B . Hence, when this is replaced by its REML estimator, we have

$$\hat{b}_i = \hat{B}Z_i'\hat{\Sigma}_i^{-1}(Y_i - X_i\hat{\beta}),$$

also known as the empirical BLUP (or empirical Bayes estimate).

- Given \hat{b}_i we obtain

$$\hat{Y}_i = X_i\hat{\beta} + Z_i\hat{b}_i.$$

- As a result we have

$$\begin{aligned}
\hat{Y}_i &= X_i\hat{\beta} + Z_i\hat{b}_i \\
&= X_i\hat{\beta} + Z_i\hat{B}Z_i'\hat{\Sigma}_i^{-1}(Y_i - X_i\hat{\beta}) \\
&= (I_{n_i} - Z_i\hat{B}Z_i'\hat{\Sigma}_i^{-1})X_i\hat{\beta} + Z_i\hat{B}Z_i'\hat{\Sigma}_i^{-1}Y_i \\
&= (\hat{W}_i\hat{\Sigma}_i^{-1})X_i\hat{\beta} + (I_{n_i} - \hat{W}_i\hat{\Sigma}_i^{-1})Y_i
\end{aligned}$$

where

$$\hat{\Sigma}_i\hat{\Sigma}_i^{-1} = I_{n_i} = (Z_i\hat{B}Z_i' + \hat{W}_i)\hat{\Sigma}_i^{-1} = Z_i\hat{B}Z_i'\hat{\Sigma}_i^{-1} + \hat{W}_i\hat{\Sigma}_i^{-1}.$$

This expression shows that \hat{Y}_i is a weighted mean of $X_i\hat{\beta}$, the population-averaged mean response profile and Y_i the i^{th} patient's observed response profile.

- As a result the predicted response profile is pulled (shrinks) towards the population-averaged mean response profile.

- The amount of shrinkage depends on W_i and Σ_i .
- If W_i is "large" then the within-subject variability is greater than the between subject variability and hence more weight is given on the population averaged mean response profile $X_i\hat{\beta}$.
- The opposite holds when W_i is "small".

Example: Orthodont (**cont.**)

```
>lmeOrth1=lme(distance ~ I(age-11),data=Orthodont,random=~1)
```

```
>lmeOrth1ml=update(lmeOrth1,method='ML')
```

```
>lmeOrth2=lme(distance ~ I(age-11),data=Orthodont)
```

```
>lmeOrth2ml=update(lmeOrth2,method='ML')
```

```
>lmeOrth3=update(lmeOrth2,fixed=distance ~ Sex*I(age-11))
```

```
> summary(lmeOrth1)
Linear mixed-effects model fit by REML
Data: Orthodont
      AIC      BIC    logLik
455.0025 465.6563 -223.5013

Random effects:
Formula: ~1 | Subject
      (Intercept) Residual
StdDev:    2.114724 1.431592

Fixed effects: distance ~ I(age - 11)
              Value Std.Error DF  t-value p-value
(Intercept) 24.023148 0.4296605 80 55.91193    0
I(age - 11)  0.660185 0.0616059 80 10.71626    0
Correlation:
      (Intr)
I(age - 11) 0

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-3.66453932 -0.53507984 -0.01289591  0.48742859  3.72178465

Number of Observations: 108
Number of Groups: 27
```

```
> OrthRElml=random.effects(lmeOrthlml)
> OrthRElml
      (Intercept)
M16  -0.9152788
M05  -0.9152788
M02  -0.5798146
M11  -0.3561719
M07  -0.2443505
M08  -0.1325291
M03   0.2029351
M12   0.2029351
M13   0.2029351
M14   0.7620421
M09   0.9856849
M15   1.6566133
M06   2.1038989
M04   2.3275416
M01   3.3339342
M10   4.8994337
F10  -4.9408491
F09  -2.5925998
F06  -2.5925998
F01  -2.3689570
F05  -1.2507430
F07  -0.9152788
F02  -0.9152788
F08  -0.5798146
F03  -0.2443505
F04   0.7620421
F11   2.1038989
```



```
> coef(lmeOrth1)#subject specific coefficients (random intercept only)
      (Intercept) I(age - 11)
M16      23.10517      0.6601852
M05      23.10517      0.6601852
M02      23.44163      0.6601852
M11      23.66593      0.6601852
M07      23.77808      0.6601852
M08      23.89023      0.6601852
M03      24.22668      0.6601852
M12      24.22668      0.6601852
M13      24.22668      0.6601852
M14      24.78744      0.6601852
M09      25.01174      0.6601852
M15      25.68464      0.6601852
M06      26.13325      0.6601852
M04      26.35755      0.6601852
M01      27.36691      0.6601852
M10      28.93702      0.6601852
F10      19.06774      0.6601852
F09      21.42291      0.6601852
F06      21.42291      0.6601852
F01      21.64721      0.6601852
F05      22.76872      0.6601852
F07      23.10517      0.6601852
F02      23.10517      0.6601852
F08      23.44163      0.6601852
F03      23.77808      0.6601852
F04      24.78744      0.6601852
F11      26.13325      0.6601852
```

```
> summary(lmeOrthlml)
Linear mixed-effects model fit by maximum likelihood
Data: Orthodont
      AIC      BIC    logLik
451.3895 462.1181 -221.6948

Random effects:
Formula: ~1 | Subject
      (Intercept) Residual
StdDev:    2.072142 1.422728

Fixed effects: distance ~ I(age - 11)
              Value Std.Error DF  t-value p-value
(Intercept) 24.023148 0.4255878 80 56.44699    0
I(age - 11)  0.660185 0.0617993 80 10.68272    0
Correlation:
      (Intr)
I(age - 11) 0

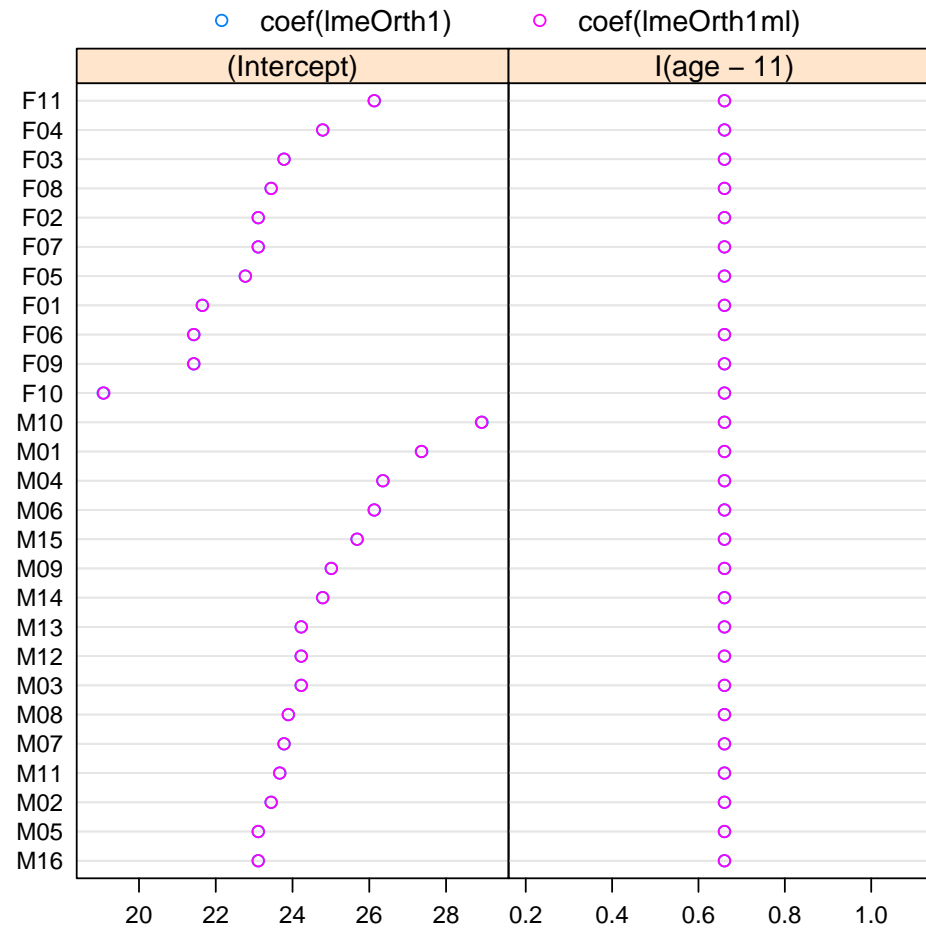
Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-3.68695130 -0.53862941 -0.01232442  0.49100161  3.74701483

Number of Observations: 108
Number of Groups: 27
```

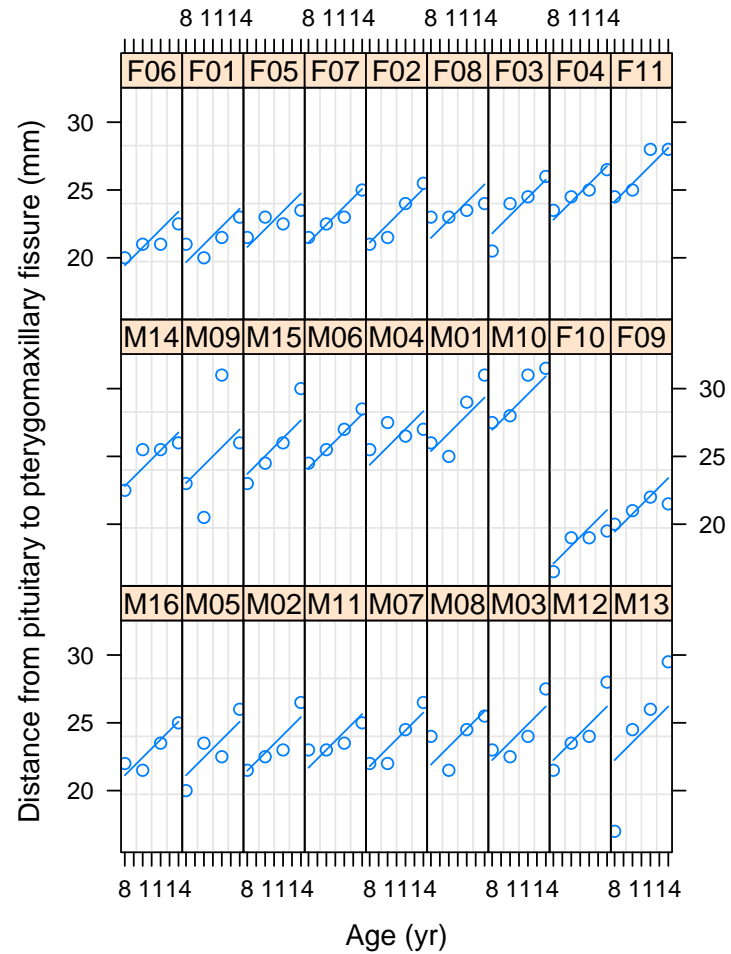
```
> OrthRElml=random.effects(lmeOrthlml)
> OrthRElml
      (Intercept)
M16  -0.9152788
M05  -0.9152788
M02  -0.5798146
M11  -0.3561719
M07  -0.2443505
M08  -0.1325291
M03   0.2029351
M12   0.2029351
M13   0.2029351
M14   0.7620421
M09   0.9856849
M15   1.6566133
M06   2.1038989
M04   2.3275416
M01   3.3339342
M10   4.8994337
F10  -4.9408491
F09  -2.5925998
F06  -2.5925998
F01  -2.3689570
F05  -1.2507430
F07  -0.9152788
F02  -0.9152788
F08  -0.5798146
F03  -0.2443505
F04   0.7620421
F11   2.1038989
```

```
> coef(lmeOrthlml)#subject specific coefficients (random intercept only)
      (Intercept) I(age - 11)
M16      23.10787      0.6601852
M05      23.10787      0.6601852
M02      23.44333      0.6601852
M11      23.66698      0.6601852
M07      23.77880      0.6601852
M08      23.89062      0.6601852
M03      24.22608      0.6601852
M12      24.22608      0.6601852
M13      24.22608      0.6601852
M14      24.78519      0.6601852
M09      25.00883      0.6601852
M15      25.67976      0.6601852
M06      26.12705      0.6601852
M04      26.35069      0.6601852
M01      27.35708      0.6601852
M10      28.92258      0.6601852
F10      19.08230      0.6601852
F09      21.43055      0.6601852
F06      21.43055      0.6601852
F01      21.65419      0.6601852
F05      22.77241      0.6601852
F07      23.10787      0.6601852
F02      23.10787      0.6601852
F08      23.44333      0.6601852
F03      23.77880      0.6601852
F04      24.78519      0.6601852
F11      26.12705      0.6601852
```

```
>plot(compareFits(coef(lmeOrth1),coef(lmeOrth1ml)))
```



```
>plot(augPred(lmeOrth1), aspect="xy", grid=T)
```



```
> summary(lmeOrth2)
Linear mixed-effects model fit by REML
Data: Orthodont
      AIC      BIC    logLik
454.6367 470.6173 -221.3183

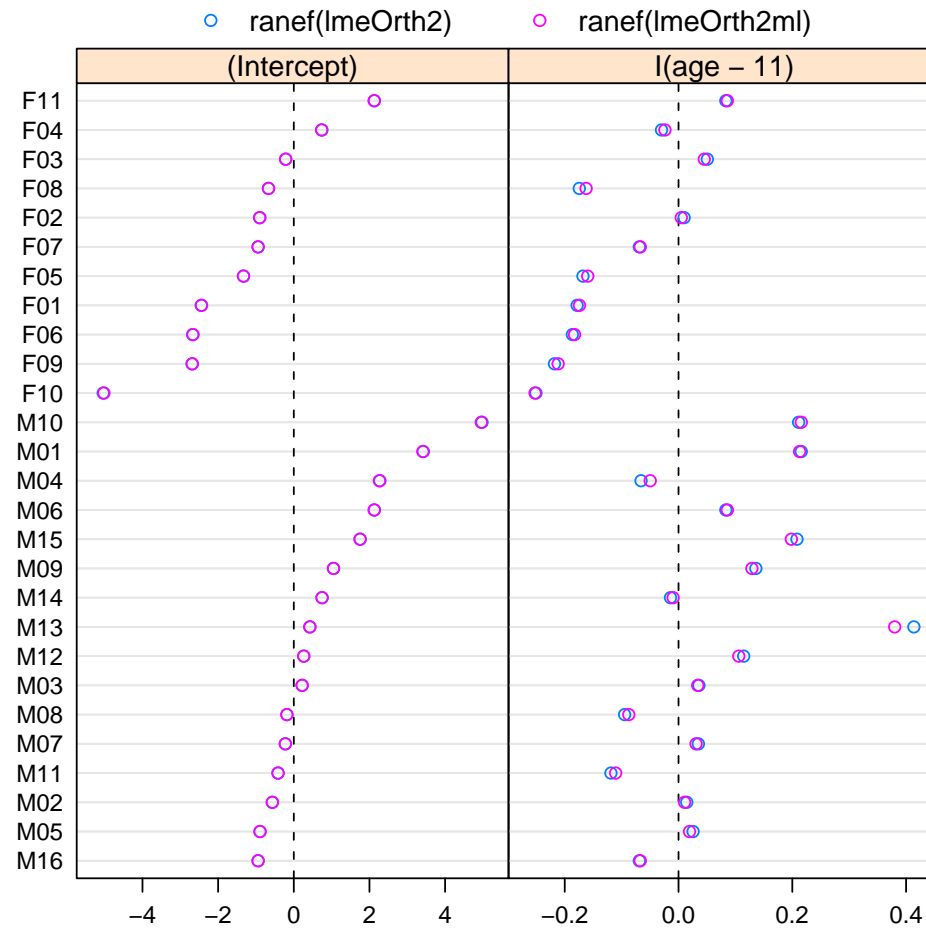
Random effects:
Formula: ~I(age - 11) | Subject
Structure: General positive-definite
      StdDev   Corr
(Intercept) 2.1343327 (Intr)
I(age - 11) 0.2264275 0.503
Residual    1.3100394

Fixed effects: distance ~ I(age - 11)
      Value Std.Error DF  t-value p-value
(Intercept) 24.023148 0.4296608 80 55.91189    0
I(age - 11)  0.660185 0.0712532 80  9.26534    0
Correlation:
      (Intr)
I(age - 11) 0.294

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-3.223106405 -0.493761198  0.007316808  0.472151143  3.916034231

Number of Observations: 108
Number of Groups: 27
```

```
>plot(compareFits(ranef(lmeOrth2), ranef(lmeOrth2ml)), mark=c(0,0))
```




```

> summary(lmeOrth3)
Linear mixed-effects model fit by REML
Data: Orthodont
      AIC      BIC    logLik
458.9891 498.655 -214.4945

Random effects:
Formula: ~Sex + I(age - 11) + Sex:I(age - 11) | Subject
Structure: General positive-definite
              StdDev   Corr
(Intercept)   1.7178454 (Intr) SexFml I(-11)
SexFemale     1.6956351 -0.307
I(age - 11)   0.2937695 -0.009 -0.146
SexFemale:I(age - 11) 0.3160597 0.168 0.290 -0.964
Residual      1.2551778

Fixed effects: distance ~ Sex + I(age - 11) + Sex:I(age - 11)
              Value Std.Error DF  t-value p-value
(Intercept)  24.968750 0.4572240 79 54.60945 0.0000
SexFemale    -2.321023 0.7823126 25 -2.96687 0.0065
I(age - 11)   0.784375 0.1015733 79 7.72226 0.0000
SexFemale:I(age - 11) -0.304830 0.1346293 79 -2.26421 0.0263
Correlation:
              (Intr) SexFml I(-11)
SexFemale    -0.584
I(age - 11)  -0.006 0.004
SexFemale:I(age - 11) 0.005 0.144 -0.754

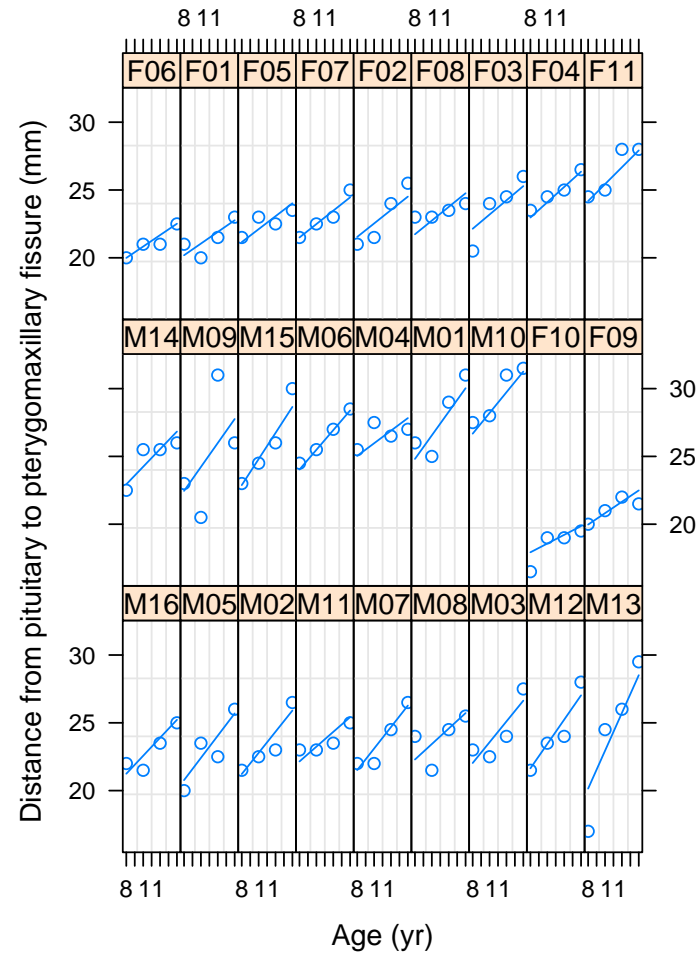
Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-2.96534486 -0.38609670 0.03647795 0.43142668 3.99155835

Number of Observations: 108
Number of Groups: 27

```

```
> OrthRE3=random.effects(lmeOrth3)
> OrthRE3
      (Intercept)  SexFemale  I(age - 11) SexFemale:I(age - 11)
M16 -1.73612668   0.63199885 -0.121203414   0.0748642681
M05 -1.73713471   0.49796730  0.035630448   -0.0876368146
M02 -1.40604191   0.43103963 -0.003830025   -0.0370896958
M11 -1.18396932   0.56512991 -0.239248823    0.2132764937
M07 -1.07528511   0.31943477  0.008987456   -0.0407096045
M08 -0.96357680   0.47583428 -0.213277852    0.1928075453
M03 -0.63399603   0.20785928 -0.017487532   -0.0003969599
M12 -0.63483606   0.09616632  0.113207353   -0.1358145288
M13 -0.63802816  -0.32826691  0.609847916   -0.6504012907
M14 -0.08183867   0.14099033 -0.135532941    0.1380152657
M09  0.13720981  -0.12701403  0.099549847   -0.0991217929
M15  0.79838740  -0.39490093  0.177462762   -0.1605286380
M06  1.24102052  -0.32776769 -0.058124041    0.0964521169
M04  1.46326110  -0.17133882 -0.319681817    0.3739018202
M01  2.45317943  -0.81889368  0.084716304   -0.0161270990
M10  3.99777519  -1.19823860 -0.021015641    0.1385089141
F10 -1.91258504  -1.84210386  0.071770763   -0.2293495874
F09 -0.72087067  -0.69430276  0.027050737   -0.0864435068
F06 -0.71120815  -0.68499782  0.026688309   -0.0852850411
F01 -0.59610113  -0.57413261  0.022368854   -0.0714818606
F05 -0.03022851  -0.02911148  0.001134008   -0.0036244236
F07  0.16900395  0.16277491 -0.006341852    0.0202661280
F02  0.19316023  0.18603726 -0.007247922    0.0231622923
F08  0.30543005  0.29417922 -0.011461928    0.0366266523
F03  0.54331257  0.52328536 -0.020387500    0.0651510668
F04  1.02505976  0.98728531 -0.038465941    0.1229211327
F11  1.73502694  1.67108646 -0.065107527    0.2080571474
```

```
>plot(augPred(lmeOrth3), aspect="xy", grid=T)
```

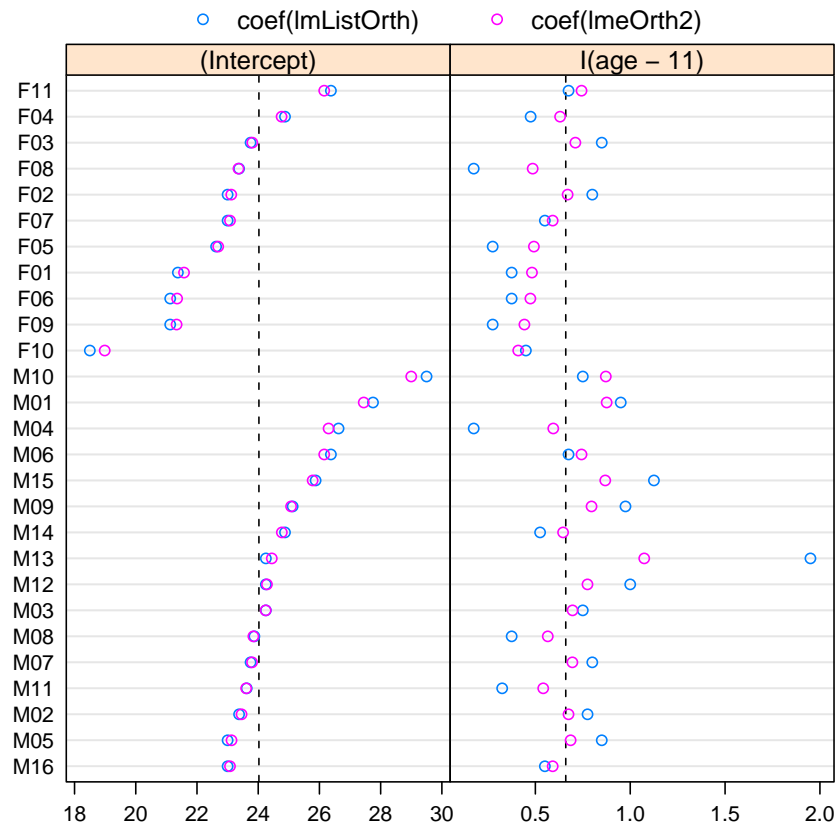


```
> newOrth=data.frame(Subject=rep(c("M11","F03"),c(3,3)),Sex=rep(c("Male","Female"),c(3,3)),age
=rep(16:18,2) )
> newOrth
  Subject    Sex age
1     M11   Male  16
2     M11   Male  17
3     M11   Male  18
4     F03 Female  16
5     F03 Female  17
6     F03 Female  18
> predict(lmeOrth3,newdata=newOrth,level=0:1)
  Subject predict.fixed predict.Subject
1     M11      28.89063      26.51041
2     M11      29.67500      27.05554
3     M11      30.45938      27.60066
4     F03      25.04545      26.33587
5     F03      25.52500      26.86018
6     F03      26.00455      27.38449
```

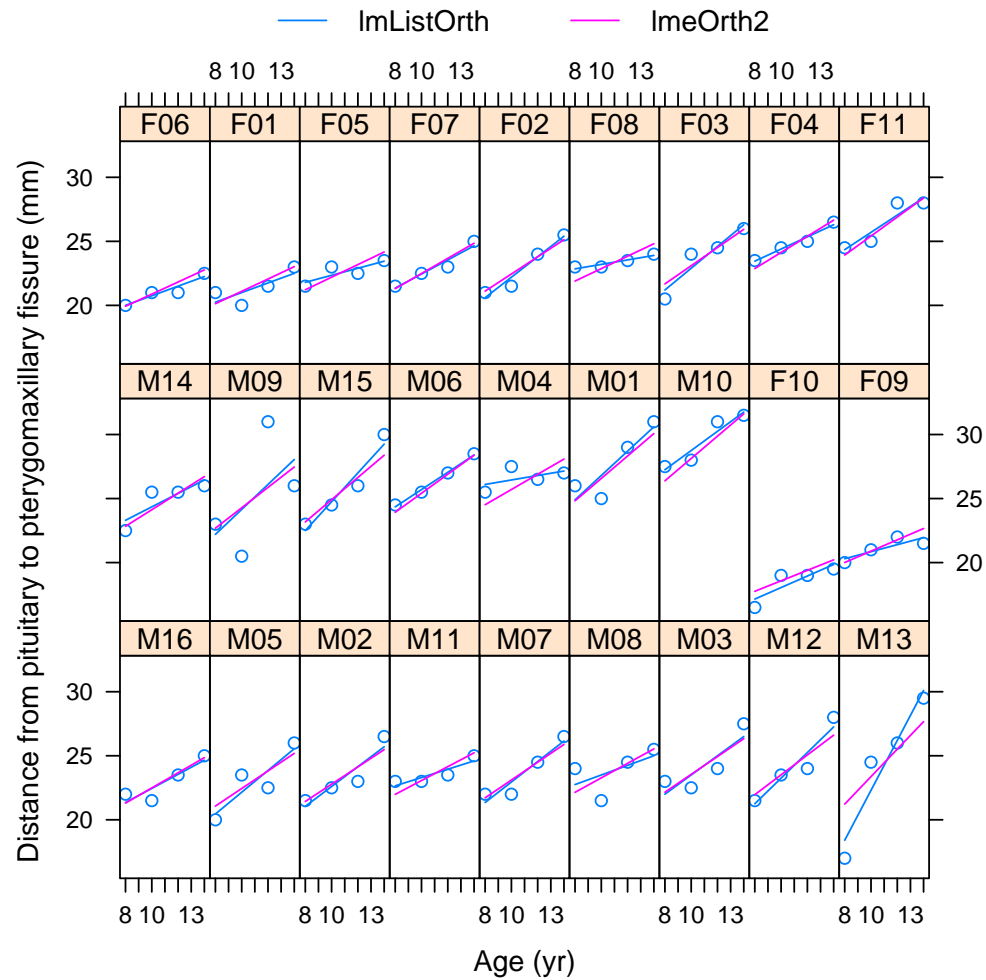
```

>lmListOrth=lmList(distance I(age-11), data=Orthodont)
>compFOrth=compareFits(coef(lmListOrth),coef(lmeOrth2))

```



```
>plot(comparePred(lmListOrth, lmeOrth2, length.out=2), layout=c(9,3))
```



Examining a Fitted Model

There are two basic assumptions that need to be assessed

1. the within-group errors are assumed independent and identically normally distributed with mean zero and variance σ^2 (since $W_i = \sigma^2 I$), and they are independent of the random effects
2. the random effects are normally distributed with mean zero and covariance matrix B (not depending on the group) and are independent for different groups.

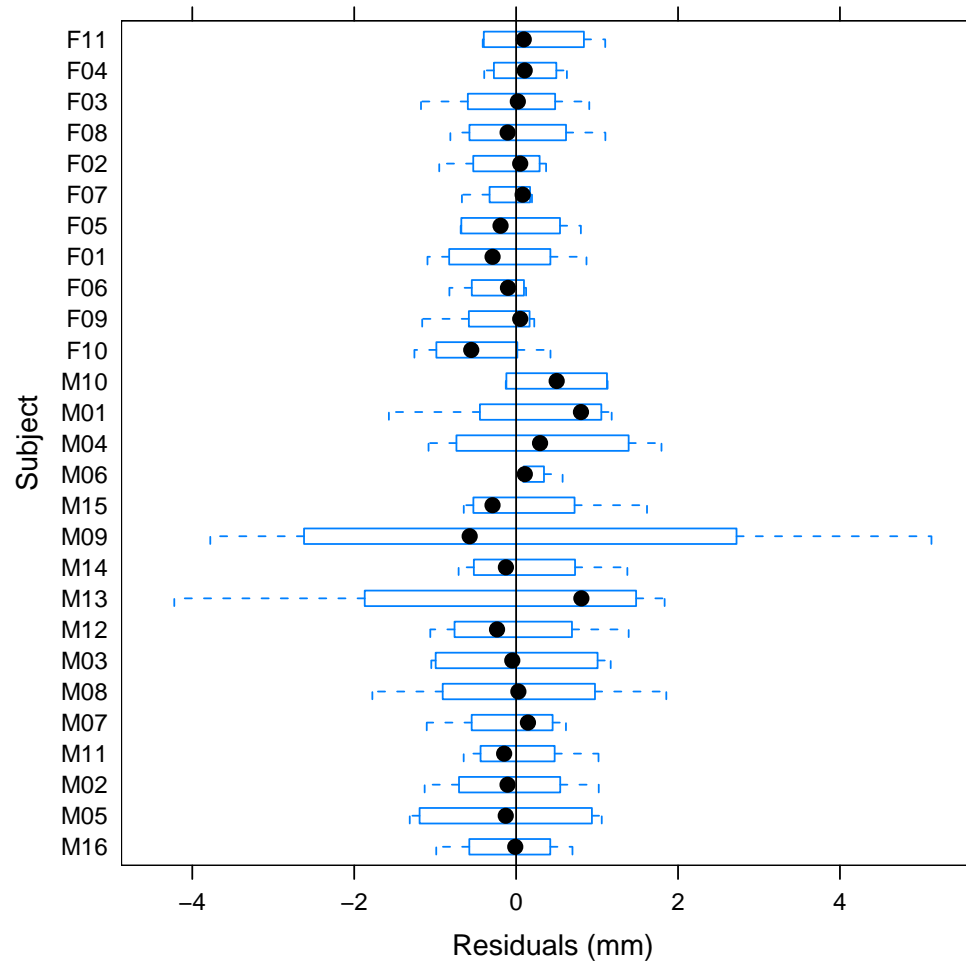
Assessing assumptions on the within-group error

- The primary quantities used to assess the adequacy of the first assumption are the *within-group residuals*, defined as the difference between the observed and the within-group fitted value.
- The `plot` method of `lme` class is the primary tool for obtaining diagnostics for the first assumption.

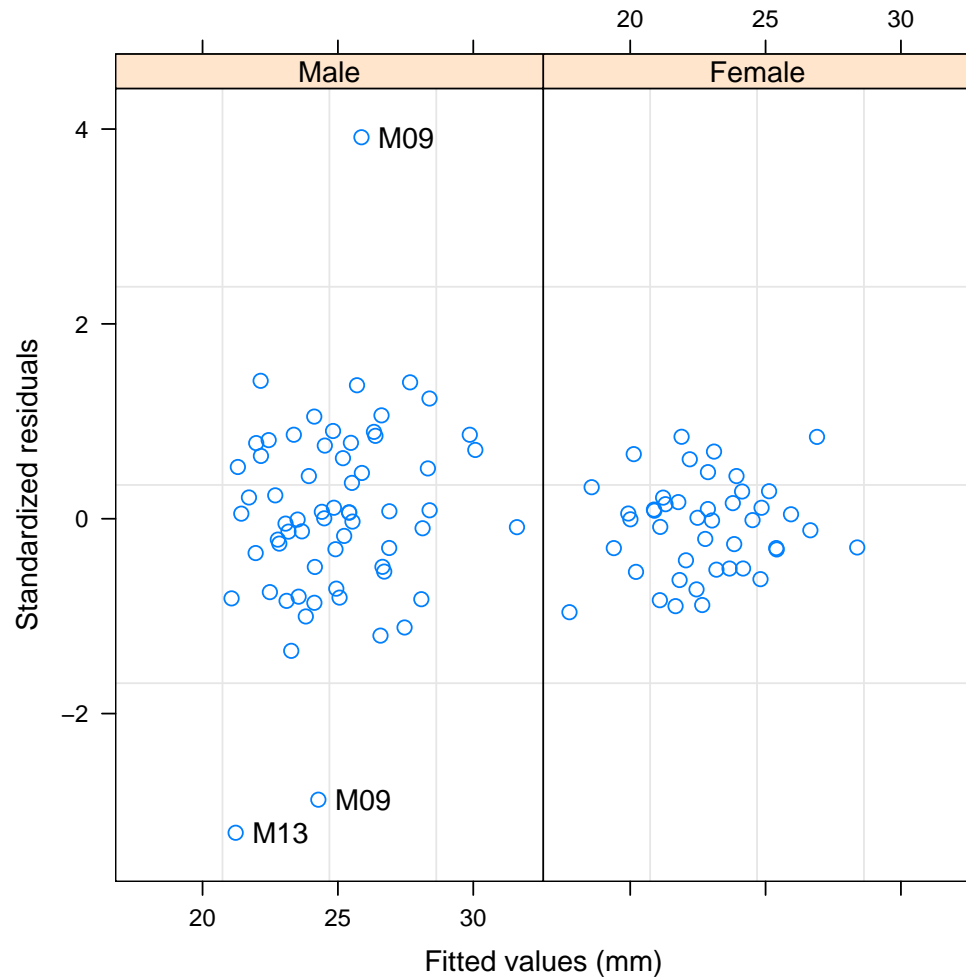
Example: Orthodont (cont.)

- Initially we consider the box plot of the residuals, by group.
- We add a vertical line at zero so we can assess whether
 - the residuals are centered at zero
 - have constant variance across groups
 - are independent of the group level

```
>plot(lmeOrth2,Subject ~ resid(.),abline=0)
```



```
>plot(lme0rth2,resid(.,type='p') ~ fitted(.)|Sex,id=0.05,adj=-0.3)
```



```

> lmeOrth5=lme(distance~I(age-11),data=Orthodont,weights=varIdent(form=~1|Sex))
> summary(lmeOrth5)
Linear mixed-effects model fit by REML
Data: Orthodont
      AIC      BIC    logLik
435.6466 454.2907 -210.8233

Random effects:
Formula: ~I(age - 11) | Subject
Structure: General positive-definite
      StdDev   Corr
(Intercept) 2.1590091 (Intr)
I(age - 11) 0.1980627 0.617
Residual    1.6452598

Variance function:
Structure: Different standard deviations per stratum
Formula: ~1 | Sex
Parameter estimates:
      Male   Female
1.0000000 0.4040981
Fixed effects: distance ~ I(age - 11)
      Value Std.Error DF  t-value p-value
(Intercept) 23.97377 0.4341697 80 55.21752      0
I(age - 11) 0.60686 0.0594260 80 10.21203      0
Correlation:
      (Intr)
I(age - 11) 0.391

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-3.02779067 -0.48052007 0.04214476 0.51813201 3.18632228

Number of Observations: 108
Number of Groups: 27

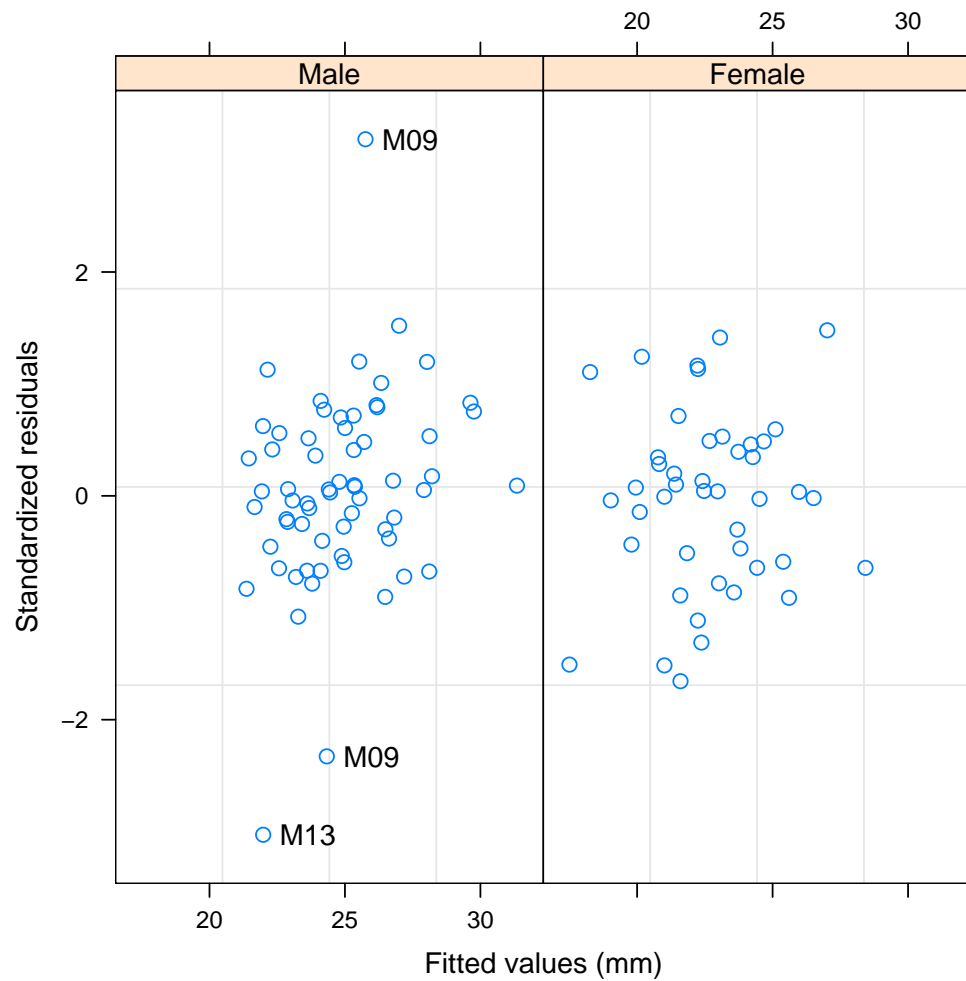
```

R Console

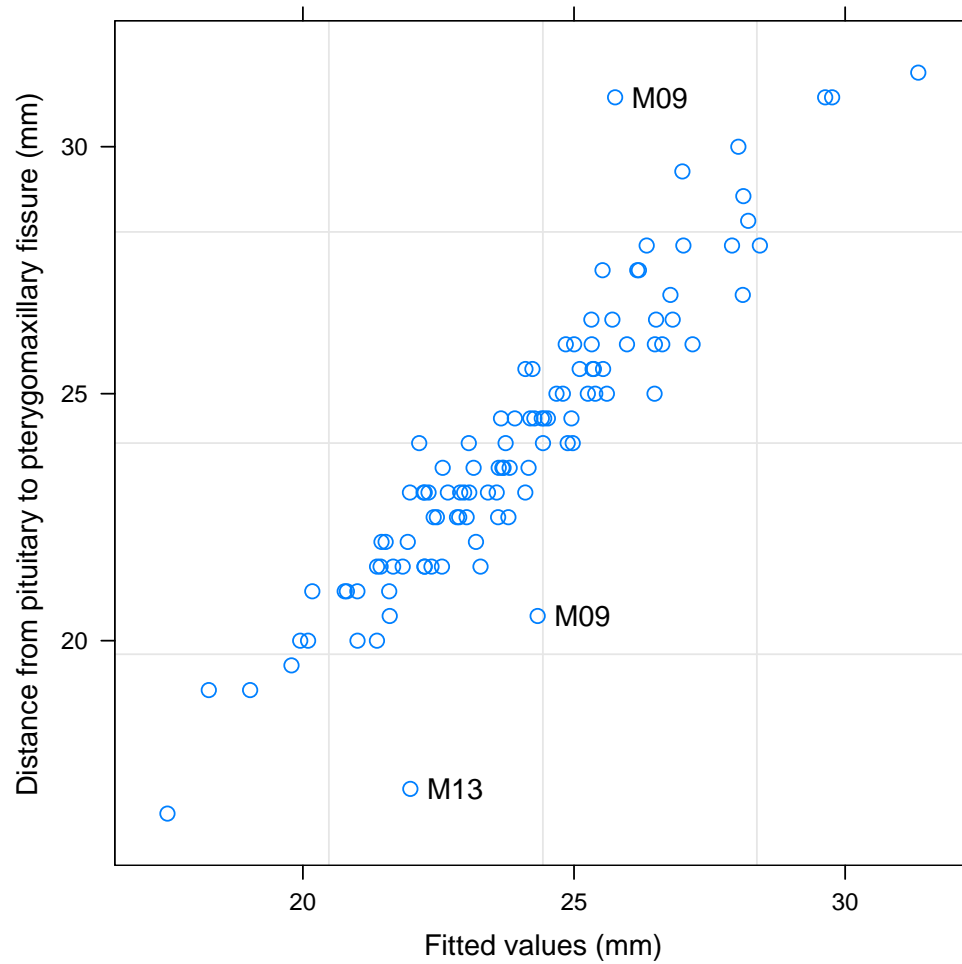
Page 1

```
> anova(lmeOrth2,lmeOrth5)
      Model df      AIC      BIC    logLik  Test  L.Ratio p-value
lmeOrth2    1  6 454.6367 470.6173 -221.3183
lmeOrth5    2  7 435.6466 454.2907 -210.8233 1 vs 2 20.99004 <.0001
```

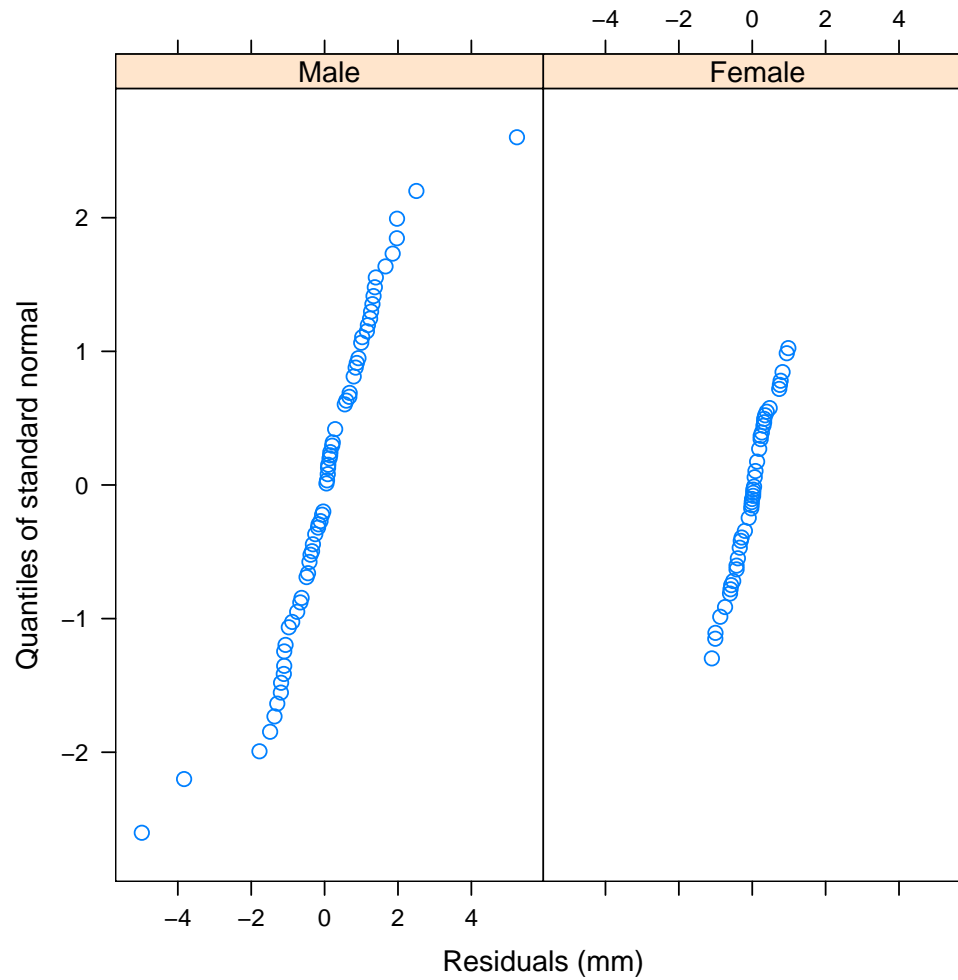
```
>plot(lme0rth5,resid(.,type='p') ~ fitted(.)|Sex,id=0.05,adj=-0.3)
```



```
>plot(lme0rth5,distance ~ fitted(.),id=0.05,adj=-0.3)
```



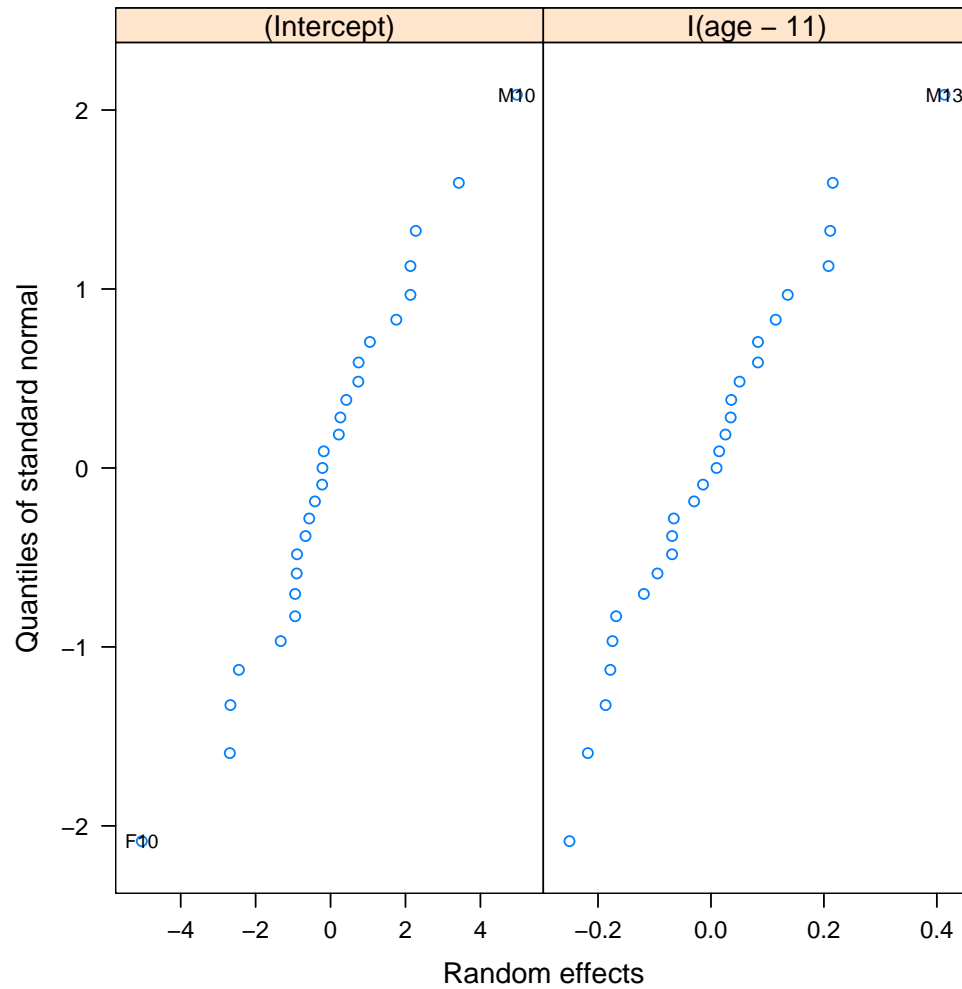
```
>qqnorm(lmeOrth5, ~ resid(.)|Sex)
```



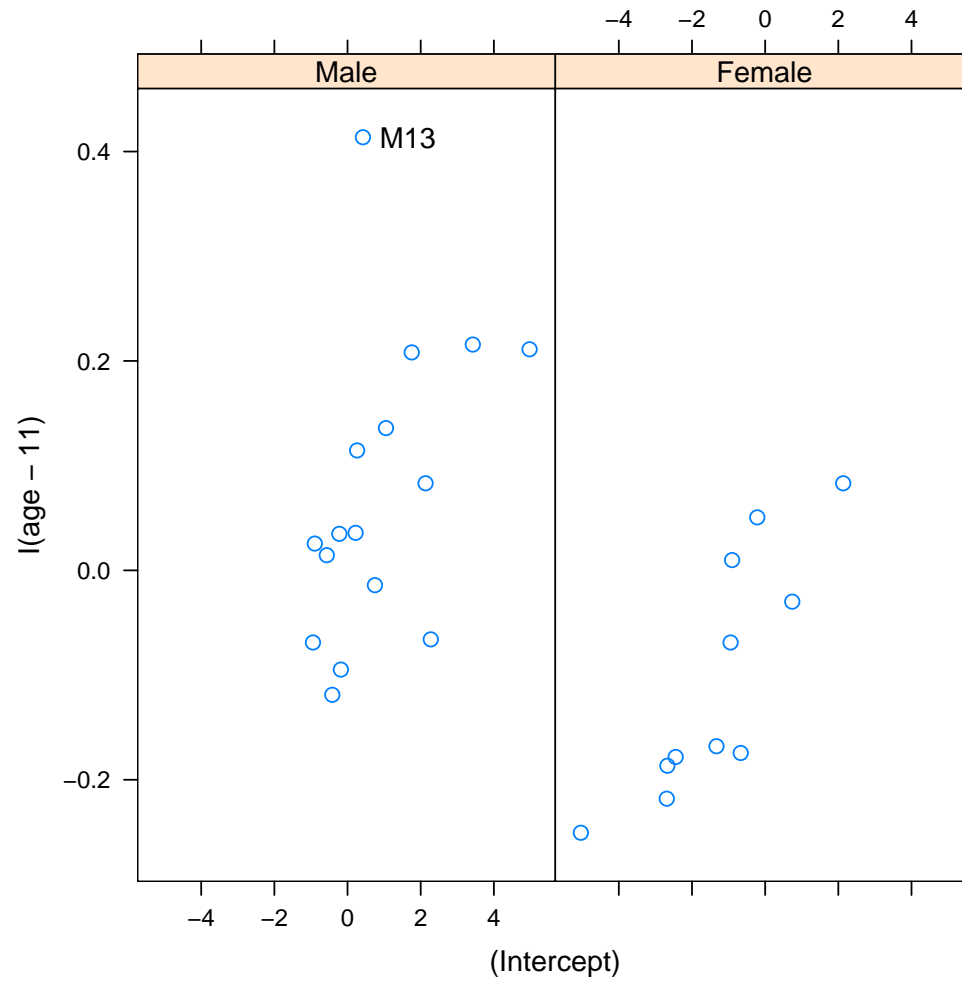
Assessing assumptions on the random effects

- The `ranef` method is used to obtain the estimated BLUP of the random effects for `lme` objects.
- Two types of diagnostic plots will be used to assess the second assumption
 - `qqnorm`: normal plot
 - `pairs`: scatter plot

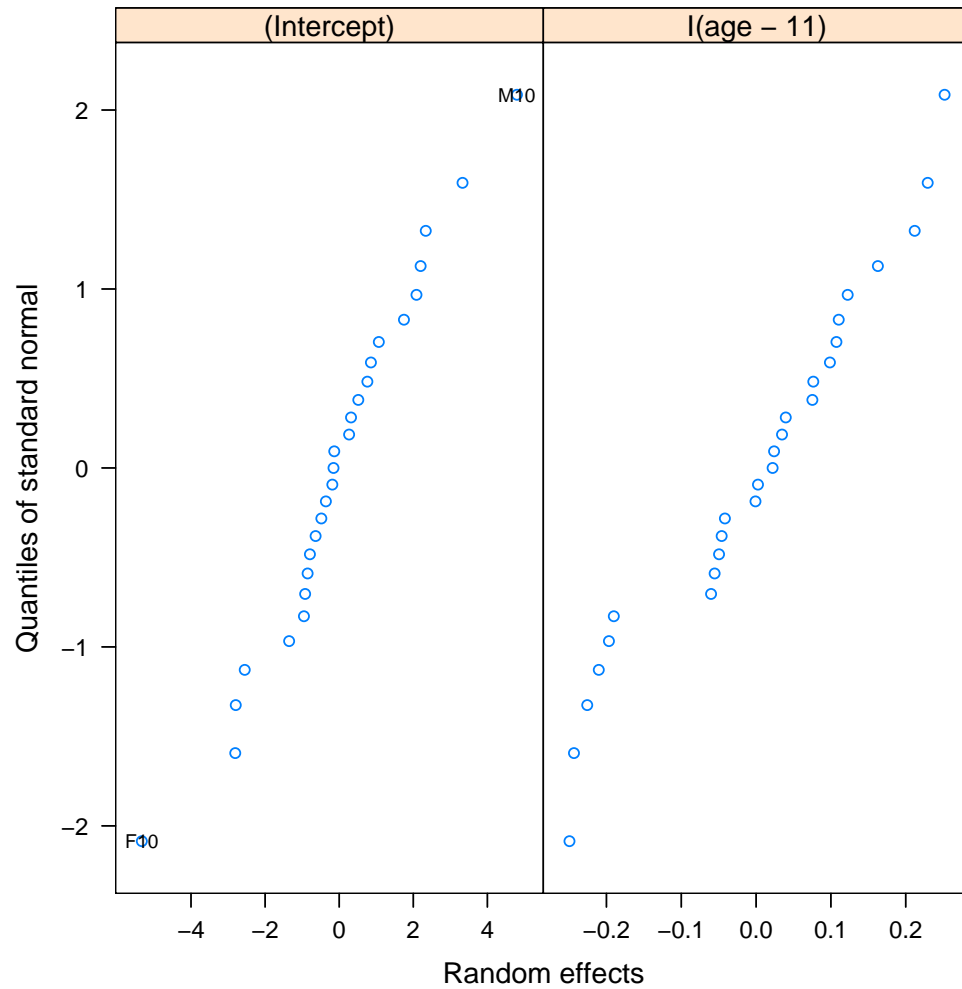
```
>qqnorm(lme0rth2, ~ ranef(.),id=0.10,cex=0.7)
```



```
>pairs(lmeOrth2, ~ranef(.)|Sex, id= ~ Subject=='M13', adj=-0.3)
```



```
>qqnorm(lme0rth5, ~ ranef(.),id=0.10,cex=0.7)
```



Generalized Linear Mixed Effects Models

- GLMs can be extended, with the inclusion of random parameters, to allow variation between subjects
- Random effects follow multivariate normal distribution
- Conditional on random effects, responses are independent following a distribution that belongs to the *exponential family*.

Model Specification:

- The distribution of Y_{ij} , conditional to random effects, belongs to the exponential family of distributions.

- It's variance is

$$\text{Var}(Y_i) = \phi v(E[Y_{ij}|b_i])$$

- Given b_i , Y_{ij} are independent from one another

- In matrix notation, the linear predictor can be written

$$\eta_{ij} = X'_{ij}\beta + Z'_{ij}b_i,$$

and for some known link function $g()$

$$g(E[Y_{ij}|b_i]) = \eta_{ij} = X'_{ij}\beta + Z'_{ij}b_i,$$

- Random effects, in theory, can follow any multivariate distribution. In practice, they follow multivariate normal with mean equal zero and a covariance matrix G .

GLMM for Continuous Response:

- Responses Y_{ij} are independent, conditional on b_i , and normally distributed

- Variance has the form

$$\text{Var}(Y_{ij}|b_i) = \sigma^2,$$

where $\phi = \sigma^2$ and $v(\mu) = 1$.

- The linear predictor is

$$\eta_{ij} = X'_{ij}\beta + Z'_{ij}b_i,$$

where $X'_{ij} = Z'_{ij} = (1, t_{ij})$ (illustration). Then

$$\begin{aligned} E(Y_{ij}|b_i) &= \eta_{ij} = X'_{ij}\beta + Z'_{ij}b_i \\ &= (\beta_1 + b_{1i}) + (\beta_2 + b_{2i})t_{ij}. \end{aligned}$$

- Although the link is the identity function, more options are available

- Random effects have a bi-variate Normal with covariance matrix $G_{2 \times 2}$

GLMM for Binary Response:

- Responses Y_{ij} are independent, conditional on b_i , Bernoulli variables
- Variance has the form

$$\text{Var}(Y_{ij}|b_i) = E(Y_{ij}|b_i)(1 - E(Y_{ij}|b_i)).$$

This means that $\phi = 1$.

- The linear predictor is given by

$$\begin{aligned}\eta_{ij} &= X'_{ij}\beta + Z'_{ij}b_i \\ &= X'_{ij}\beta + b_i,\end{aligned}$$

where $Z'_{ij} = 1$ for all i, j (illustration). Then

$$\log \left[\frac{P(Y_{ij} = 1|b_i)}{P(Y_{ij} = 0|b_i)} \right] = \eta_{ij} = X'_{ij}\beta + b_i$$

- $b_i \sim N(0, \sigma^2)$.
- This is a random intercept model, equivalent to the compound symmetry model.

GLMM for Counts:

- Responses Y_{ij} are independent, conditional on b_i , following Poisson distribution

- Variance has the form

$$\text{Var}(Y_{ij}|b_i) = E(Y_{ij}|b_i).$$

This means that $\phi = 1$.

- The linear predictor is given by

$$\eta_{ij} = X'_{ij}\beta + Z'_{ij}b_i,$$

where $Z'_{ij} = (1, t_{ij})$ for all i, j (illustration). Then

$$\log E(Y_{ij}|b_i) = \eta_{ij} = X'_{ij}\beta + Z'_{ij}b_i$$

- Random effects follow bivariate normal with zero mean and 2x2 covariance matrix

Parameter Interpretation

- Parameters in the linear predictor are now interpreted in terms of conditional probabilities, given subject (random) effects
- Regression parameters β in GLMM have different interpretation than in marginal models
- In GLMM, β represent subject-specific interpretation
- Specifically, β represent the impact of covariates on changes in an individual's transformed mean response

- Consider the example with the logistic regression model

$$\log \left[\frac{P(Y_{ij} = 1 | b_i)}{P(Y_{ij} = 0 | b_i)} \right] = X'_{ij} \beta + b_i,$$

where $b_i \sim N(0, g_{11})$. Furthermore, consider covariate X_{ijk} takes some value x , leading to the log-odds

$$\begin{aligned} \log \left[\frac{P(Y_{ij} = 1 | b_i, X_{ij1}, X_{ij2}, \dots, X_{ijk} = x, \dots, X_{ijp})}{P(Y_{ij} = 0 | b_i, X_{ij1}, X_{ij2}, \dots, X_{ijk} = x, \dots, X_{ijp})} \right] \\ = \beta_1 X_{ij1} + \beta_2 X_{ij2} + \dots + \beta_k x + \dots + \beta_p X_{ijp} + b_i. \end{aligned}$$

Additionally, if $X_{ijk} = x + 1$, then the log-odds takes the form

$$\begin{aligned} \log \left[\frac{P(Y_{ij} = 1 | b_i, X_{ij1}, X_{ij2}, \dots, X_{ijk} = x + 1, \dots, X_{ijp})}{P(Y_{ij} = 0 | b_i, X_{ij1}, X_{ij2}, \dots, X_{ijk} = x + 1, \dots, X_{ijp})} \right] \\ = \beta_1 X_{ij1} + \beta_2 X_{ij2} + \dots + \beta_k (x + 1) + \dots + \beta_p X_{ijp} + b_i, \end{aligned}$$

and hence β_k measures the changes in the log-odds resulted from a unit change in covariate X_{ijk} while the remaining ones were held fixed. In terms of interpretation:

– If the covariate X_{ijk} varies within individual (subject-specific, time-varying) then

$$\log \left[\frac{P(Y_{ij'} = 1 | b_i, X_{ij'1}, X_{ij'2}, \dots, X_{ij'k} = x + 1, \dots, X_{ij'p})}{P(Y_{ij'} = 0 | b_i, X_{ij'1}, X_{ij'2}, \dots, X_{ij'k} = x + 1, \dots, X_{ij'p})} \right] \\ - \log \left[\frac{P(Y_{ij} = 1 | b_i, X_{ij1}, X_{ij2}, \dots, X_{ijk} = x, \dots, X_{ijp})}{P(Y_{ij} = 0 | b_i, X_{ij1}, X_{ij2}, \dots, X_{ijk} = x, \dots, X_{ijp})} \right] = \beta_k,$$

where the interpretation is quite straight forward since all other covariates as well as random effects are the same and hence removed. Hence,

$$\log \left[\frac{P(Y_{ij'}=1|b_i,\dots)/P(Y_{ij'}=0|b_i,\dots)}{P(Y_{ij}=1|b_i,\dots)/P(Y_{ij}=0|b_i,\dots)} \right] = \log OR = \beta_k \Rightarrow \\ OR = \exp(\beta_k)$$

is the within subject OR.

- If the covariate X_{ijk} is time invariant (between-subject), like treatment group, interpretation becomes complicated. Hence

$$\begin{aligned}
 & \log \left[\frac{P(Y_{ij} = 1 | b_i, X_{ij1}, X_{ij2}, \dots, X_{ijk} = 1, \dots, X_{ijp})}{P(Y_{ij} = 0 | b_i, X_{ij1}, X_{ij2}, \dots, X_{ijk} = 1, \dots, X_{ijp})} \right] \\
 - & \log \left[\frac{P(Y_{i'j} = 1 | b_{i'}, X_{i'j1}, X_{i'j2}, \dots, X_{i'jk} = 0, \dots, X_{i'jp})}{P(Y_{i'j} = 0 | b_{i'}, X_{i'j1}, X_{i'j2}, \dots, X_{i'jk} = 0, \dots, X_{i'jp})} \right] = \beta_k + (b_i - b_{i'}),
 \end{aligned}$$

and as a result the change in log-odds is confounded by $b_i - b_{i'}$. It is misleading to give to this change a subject-specific interpretation. It is seen as a model based extrapolation (no data available) and could be sensitive to various assumptions concerning the random effects.

Estimation and Inference

- The distribution of the random effects as well as the distribution of the responses are known
- As a result, the joint distribution of random effects and responses is fully specified

$$f(Y_i, b_i) = f(Y_i|b_i)f(b_i),$$

where

$$f(Y_i|b_i) = f(Y_{i1}|b_i) f(Y_{i2}|b_i) \dots f(Y_{in_i}|b_i)$$

under the conditional independence assumption.

- Then, the likelihood function takes the form

$$L(\beta, \phi, G) = \prod_{i=1}^N \int f(Y_i|b_i) f(b_i) db_i,$$

where the random effects are integrated out of the likelihood, obtaining in that way a marginal likelihood averaged over the b_i .

- There is now way the likelihood can be written in a closed form
- As a result, numerical integration techniques are required

Prediction of b_i

- Given the MLE of β , ϕ and G , b_i can be predicted as

$$\hat{b}_i = E(b_i | Y_i; \hat{\beta}, \hat{\phi}, \hat{G})$$

- This is the empirical Bayes or BLUP used before
- Numerical integration techniques are also required

The `lmer` function (R: `lme4` package)

lmer(lme4)

R Documentation

Fit (Generalized) Linear Mixed-Effects Models

Description

Fit a linear or generalized linear mixed-effects model with nested or crossed grouping factors for the random effects.

Usage

```
lmer(formula, data, family, method, control, start,
      subset, weights, na.action, offset, contrasts,
      model, ...)
lmer2(formula, data, family, method, control, start,
       subset, weights, na.action, offset, contrasts,
       model, ...)
```

Arguments

<code>formula</code>	a two-sided linear formula object describing the fixed-effects part of the model, with the response on the left of a <code>~</code> operator and the terms, separated by <code>+</code> operators, on the right. The vertical bar character <code> </code> separates an expression for a model matrix and a grouping factor.
<code>data</code>	an optional data frame containing the variables named in <code>formula</code> . By default the variables are taken from the environment from which <code>lmer</code> is called.
<code>family</code>	a GLM family, see glm . If <code>family</code> is missing then a linear mixed model is fit; otherwise a generalized linear mixed model is fit.
<code>method</code>	a character string. For a linear mixed model the default is <code>"REML"</code> indicating that the model should be fit by maximizing the restricted log-likelihood. The alternative is <code>"ML"</code> indicating that the log-likelihood should be maximized. (This method is sometimes called "full" maximum likelihood.) For a generalized linear mixed model the criterion is always the log-likelihood but this criterion does not have a closed form expression and must be approximated. The default approximation is <code>"NPS"</code> or penalized quasi-likelihood. Alternatives are <code>"E"</code> , <code>"L"</code> , <code>"A"</code> or <code>"ASO"</code> .

Example: Respiratory Data

```
resp_lmer1 =lmer(status ~ centre + treatment + sex + baseline + age +  
(1|subject), data = resp, family = "binomial")
```

```

> summary(resp_lmer1)
Generalized linear mixed model fit using Laplace
Formula: status ~ centre + treatment + sex + baseline + age + (1 | subject)
Data: resp
Family: binomial(logit link)
AIC   BIC logLik deviance
443 471.7 -214.5    429
Random effects:
Groups Name          Variance Std.Dev.
subject (Intercept) 3.8402   1.9596
number of obs: 444, groups: subject, 111

Estimated scale (compare to 1 ) 0.7770601

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.64382    0.75668  -2.172  0.0298 *
centre2       1.04635    0.53075   1.971  0.0487 *
treatmenttreatment 2.16087    0.51652   4.183 2.87e-05 ***
sexmale       0.20740    0.65969   0.314  0.7532
baselinegood  3.07037    0.52499   5.848 4.96e-09 ***
age          -0.02549    0.01994  -1.278  0.2012
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
              (Intr) centr2 trtmnt sexmal bslnsd
centre2      -0.054
trtmnttrtmn -0.407  0.018
sexmale      -0.008 -0.151  0.222
baselinegod -0.347 -0.236  0.206  0.101
age          -0.753 -0.226 -0.015 -0.255  0.069

```



```
resp_lmer2 = lmer(status ~ centre + treatment + sex + baseline + age  
+ (age|subject), data = resp, family = "binomial")
```

```

> summary(resp_lmer2)
Generalized linear mixed model fit using Laplace
Formula: status ~ centre + treatment + sex + baseline + age + (age | subject)
Data: resp
Family: binomial(logit link)
AIC   BIC logLik deviance
445.8 482.7 -213.9   427.8
Random effects:
Groups Name          Variance Std.Dev. Corr
subject (Intercept) 1.964799 1.401713
age                0.001584 0.039799 0.003
number of obs: 444, groups: subject, 111

Estimated scale (compare to 1 ) 0.7859826

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.29487    0.72534  -1.785   0.0742 .
centre2      0.99755    0.50953   1.958   0.0503 .
treatmenttreatment 2.01372    0.50179   4.013 5.99e-05 ***
sexmale      0.24017    0.68883   0.349   0.7273
baselinegood 2.97704    0.51023   5.835 5.39e-09 ***
age          -0.03354    0.02107  -1.592   0.1114
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
              (Intr) centr2 trtmnt sexmal bslnsd
centre2      -0.084
trtmnttrtmn -0.396  0.013
sexmale      0.053 -0.130  0.215
baselinegod -0.337 -0.226  0.217  0.076
age          -0.753 -0.173 -0.038 -0.316  0.042

```