Linear Mixed Models

Andreas Futschik

Dept. of Statistics

07/2016

Contents . . .

Linear Model: Assumptions

$$Y = X\beta + \epsilon$$

- normally distributed responses
- homogeneous variances
- linear Relationship
- independent observations

Linear Model: Assumptions

$$Y = X\beta + \epsilon$$

- normally distributed responses
- homogeneous variances
- linear Relationship
- independent observations

Linear Model: Assumptions

$$Y = X\beta + \epsilon$$

- normally distributed responses
- homogeneous variances
- linear Relationship
- independent observations

Linear Model: Assumptions

$$Y = X\beta + \epsilon$$

- normally distributed responses
- homogeneous variances
- linear Relationship
- independent observations

What if assumptions are violated?

$$Y = X\beta + \epsilon$$

- normally distributed responses ⇒ generalized linear models
- homogeneous variances ⇒ transformation of y, or weighted LS
- linear relationship ⇒ additional predictors, splines, nonlinear regression
- independent observations ⇒ next slide
- more than one assumption might be violated simultaneously

What if assumptions are violated?

$$Y = X\beta + \epsilon$$

- normally distributed responses ⇒ generalized linear models
- homogeneous variances ⇒ transformation of y, or weighted LS
- linear relationship ⇒ additional predictors, splines, nonlinear regression
- independent observations ⇒ next slide
- more than one assumption might be violated simultaneously

What if assumptions are violated?

$$Y = X\beta + \epsilon$$

- normally distributed responses ⇒ generalized linear models
- homogeneous variances ⇒ transformation of y, or weighted LS
- linear relationship ⇒ additional predictors, splines, nonlinear regression
- independent observations ⇒ next slide
- more than one assumption might be violated simultaneously

What if assumptions are violated?

$$Y = X\beta + \epsilon$$

- normally distributed responses ⇒ generalized linear models
- homogeneous variances ⇒ transformation of y, or weighted LS
- linear relationship ⇒ additional predictors, splines, nonlinear regression
- independent observations ⇒ next slide
- more than one assumption might be violated simultaneously

What if assumptions are violated?

$$Y = X\beta + \epsilon$$

- normally distributed responses $\Rightarrow$ generalized linear models
- homogeneous variances $\Rightarrow$ transformation of y, or weighted LS
- linear relationship $\Rightarrow$ additional predictors, splines, nonlinear regression
- independent observations $\Rightarrow$ next slide
- more than one assumption might be violated simultaneously

What if observations are dependent?

- Approach 1: Mixed Models
- Approach 2: Generalized Least Squares (GEE, GLS)

What if observations are dependent?

- Approach 1: Mixed Models
- Approach 2: Generalized Least Squares (GEE, GLS)

Some reasons for dependent observations

- Repeated measurements (e.g. on same subject, or experimental evolution)
- Population substructure/genetic relatedness
- Spatial dependence

Some reasons for dependent observations

- Repeated measurements (e.g. on same subject, or experimental evolution)
- Population substructure/genetic relatedness
- Spatial dependence

Some reasons for dependent observations

- Repeated measurements (e.g. on same subject, or experimental evolution)
- Population substructure/genetic relatedness
- Spatial dependence

Linear mixed model

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \gamma_1 Z_1 + \cdots + \gamma_q Z_q + \epsilon$$

- $\beta_0, \ldots, \beta_p$ . . . unknown parameters
- $\gamma_1, \ldots, \gamma_q$ . . . unobserved random variables (random effects)
- ϵ . . . random errors (i.i.d.)

Alternative Phrasing:

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + f, \quad f = \gamma_1 Z_1 + \cdots + \gamma_q Z_q + \epsilon$$

structured/dependent errors

LMM–data structure

We observe (or know) Y, X, Z:

- $Y = X\beta + Z\gamma + \epsilon$
- $X \dots n \times p$ design matrix
- $\beta \dots p \times 1$ parameter vector
- $Z \dots n \times q$ design matrix
- $\gamma \sim N(0, \sigma^2 D)$ $q \times 1$ random effects
- $D \dots q \times q$ covariance matrix of γ
- $\epsilon \sim N_n(0, \sigma^2 I)$ independent of γ.

Example: GWAS I

Correcting for population structure and kinship using the linear mixed model: theory and extensions.
G.E. Hoffman, PLoS One. 2013

- GLS: $y = \mu + x_j\beta + \alpha + \epsilon$
- $\alpha \sim N(0, K\sigma_a^2)$
- LMM: $y = \mu + x_j\beta + R\gamma + \epsilon$
- $\gamma \sim N(0, \sigma_a^2)$
- R chosen such that $R\gamma \sim N(0, K\sigma_a^2)$
- $K = XX^t$, $n \times n$ genetic similarity matrix from (centered & scaled) genotype matrix $X \ldots n \times p$ for n individuals and p markers

Example: GWAS II

Genome-wide efficient mixed-model analysis for association studies

Zhou X., Stephens M., Nat. Genetics 2012

- LMM: $y = W\alpha + x\beta + Zu + \epsilon$
- $y \ldots$, quantitative trait
- $W \ldots$ matrix of fixed covariate (age, gender, environment, etc.), $x \ldots$ vector of marker genotypes
- $Z \ldots n \times m$ indicator matrix for m strains,
- $u \sim N_m(0, \lambda K)$, with K known $m \times m$ relatedness (kinship) matrix

Example: Repeated measurements in population genetics

Genomic Evidence of Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in Drosophila
A. O. Bergland, E. L. Behrman, et al. Plos Genetics, 2014.

- LM: $y_i = season + \epsilon_i$
- GLMM: $y_i = season + (1|population) + \epsilon_i$
- GEE: $y_i = season + population + \epsilon_i$, $\epsilon_i \sim$ AR(1|population)

Example

Influence of sleep deprivation on reaction time

- 18 volunteers: sleep only 3 hours for 9 consecutive nights
- Reaction time (ms) measured at start (day 0) and at each of 9 consecutive days

```
  Reaction Days Subject
1 249.5600    0     308
2 258.7047    1     308
3 250.8006    2     308
4 321.4398    3     308
5 356.8519    4     308
6 414.6901    5     308
```

Models for the example
Random Intercept Model

$$Y = \beta_0 + \beta_1 X_1 + \sum_{j=1}^{18} \gamma_j Z_j + \epsilon$$

- $X_1 \ldots \text{day} \in 0, \ldots, 9$
- $\beta_0 \ldots$ mean reaction time at beginning
- $Z_j = \begin{cases} 1 & person = j \\ 0 & else \end{cases}$
- $D = \tau^2 I$ (diagonal)

Models for the example

Random Slope Model

$$Y = \beta_0 + \beta_1 X_1 + \sum_{j=1}^{18} \gamma_j^{(1)} Z_j^{(1)} + \sum_{j=1}^{18} \gamma_j^{(2)} Z_j^{(2)} + \epsilon$$

- X_1, β, $Z_j^{(1)}$ as before
- $Z_j^{(2)} = \begin{cases} X_{1,j}(Tag) & person = j \\ 0 & else \end{cases}$
- $D = diag(\tau_1^2 I, \tau_2^2 I)$ (diagonal)
- (further models later!)

General model for repeated measurements

$$\begin{pmatrix} Y_{(1)} \\ Y_{(2)} \\ \vdots \\ Y_{(n)} \end{pmatrix} = \begin{pmatrix} X_{(1)} \\ X_{(2)} \\ \vdots \\ X_{(n)} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} Z_{(1)} & & & \\ & Z_{(2)} & & \\ & & \ddots & \\ & & & Z_{(n)} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

- $Y_{(i)}, \epsilon_{(i)} \ldots n_i \times 1$, $\epsilon_i \sim N(0, \sigma^2 I_{n_i})$
- $X_{(i)} \ldots n_i \times p$, $Z_{(i)} \ldots n_i \times q_i$
- $\gamma_{(i)}$ independent (from each other and from ϵ_i), $\gamma_{(i)} \sim N_{n_i}(0, \sigma^2 D_i)$

Repeated Measurements versus Longitudinal Analysis

- Repeated Measurements: Multipe measurements within a group
- Longitudinal analysis: group = individual (repeated measurements on individual)
- Hierarchical model: combination & groups and longitudinal analysis: e.g. several populations and from each population some individuals followed over time
- time series: one ore more measurements taken at typically many time points

Parameter Estimation
Maximum Likelihood

- Recall: $Y = X\beta + Z\gamma + \epsilon$
- $E(Y) = X\beta$, $\mathrm{Var}(Y) = \mathrm{Var}(Z\gamma) + \mathrm{Var}(\epsilon) = \sigma^2(ZDZ^t + I)$
- so: $Y \sim N(X\beta, \sigma^2 V)$ mit $V = I + ZDZ^t$
- ML-Schätzer obtained by maximizing

$$\frac{1}{(2\pi)^{n/2}|\sigma^2 V|^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(y - X\beta)^t V^{-1}(y - X\beta)\right]$$

in β, σ^2, V.

Restricted Maximum Likelihood (REML)

- Let K denote a matrix satisfying $KX = 0$. Then also $E(KY) = K(X\beta) = 0$.
- Choose for this purpose $KY = Y - X\hat{\beta}_{OLS}$, i.e. $K = I - X(X^tX)^{-1}X^t$ then OLS resduals orthogonal on X.
- We have: $KY \sim N_k(0, \sigma^2 K(ZDZ^t + I)K^t)$
- maximize in σ^2, D
- β is then estimated via GLS $(y - X\beta)^t\hat{V}^{-1}(y - X\beta) \to$ min, with $\hat{V}$ computed using REML:

$$\hat{\beta} = (X^t\hat{V}^{-1}X)^{-1}X^t\hat{V}^{-1}Y.$$

Restricted Maximum Likelihood (REML)

- Let K denote a matrix satisfying $KX = 0$. Then also $E(KY) = K(X\beta) = 0$.
- Choose for this purpose $KY = Y - X\hat{\beta}_{OLS}$, i.e. $K = I - X(X^tX)^{-1}X^t$ then OLS resduals orthogonal on X.
- We have: $KY \sim N_k(0, \sigma^2 K(ZDZ^t + I)K^t)$
- maximize in σ^2, D
- β is then estimated via GLS $(y - X\beta)^t\hat{V}^{-1}(y - X\beta) \to$ min, with $\hat{V}$ computed using REML:

$$\hat{\beta} = (X^t\hat{V}^{-1}X)^{-1}X^t\hat{V}^{-1}Y.$$

Restricted Maximum Likelihood (REML)

- Let K denote a matrix satisfying $KX = 0$. Then also $E(KY) = K(X\beta) = 0$.
- Choose for this purpose $KY = Y - X\hat{\beta}_{OLS}$, i.e. $K = I - X(X^tX)^{-1}X^t$ then OLS resduals orthogonal on X.
- We have: $KY \sim N_k(0, \sigma^2 K(ZDZ^t + I)K^t)$
- maximize in σ^2, D
- β is then estimated via GLS $(y - X\beta)^t\hat{V}^{-1}(y - X\beta) \to$ min, with $\hat{V}$ computed using REML:

$$\hat{\beta} = (X^t\hat{V}^{-1}X)^{-1}X^t\hat{V}^{-1}Y.$$

Restricted Maximum Likelihood (REML)

- Let K denote a matrix satisfying $KX = 0$. Then also $E(KY) = K(X\beta) = 0$.
- Choose for this purpose $KY = Y - X\hat{\beta}_{OLS}$, i.e. $K = I - X(X^tX)^{-1}X^t$ then OLS resduals orthogonal on X.
- We have: $KY \sim N_k(0, \sigma^2 K(ZDZ^t + I)K^t)$
- maximize in σ^2, D
- β is then estimated via GLS $(y - X\beta)^t\hat{V}^{-1}(y - X\beta) \to$ min, with $\hat{V}$ computed using REML:

$$\hat{\beta} = (X^t\hat{V}^{-1}X)^{-1}X^t\hat{V}^{-1}Y.$$

Random Intercept

```
Linear mixed model fit by REML
Formula: Reaction ~ Days + (1 | Subject)
   Data: sleepstudy
  AIC  BIC logLik deviance REMLdev
 1794 1807 -893.2     1794    1786
Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 1378.18  37.124
 Residual              960.46  30.991
Number of obs: 180, groups: Subject, 18

Fixed effects:
            Estimate Std. Error t value
(Intercept) 251.4051     9.7459   25.80
Days         10.4673     0.8042   13.02
```

Random Intercept via ML

```
Linear mixed model fit by maximum likelihood
Formula: Reaction ~ Days + (1 | Subject)
   Data: sleepstudy
  AIC   BIC logLik deviance REMLdev
 1802  1815   -897     1794    1786
Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 1296.87  36.012
 Residual              954.53  30.895
Number of obs: 180, groups: Subject, 18

Fixed effects:
            Estimate Std. Error t value
(Intercept) 251.4051     9.5053   26.45
Days         10.4673     0.8017   13.06
```

REML versus ML

- REML leads to more precise estimates, especially for σ^2, D!
- For hypotheses testing ML-estimate needs to be used!

Random Slope

```
Linear mixed model fit by REML
Formula: Reaction ~ Days + (1 | Subject) + (0 + Days | Subject)
   Data: sleepstudy
  AIC  BIC logLik deviance REMLdev
 1754 1770 -871.8     1752    1744
Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 627.568  25.0513
 Subject  Days         35.858   5.9882
 Residual             653.584  25.5653
Number of obs: 180, groups: Subject, 18

Fixed effects:
            Estimate Std. Error t value
(Intercept)  251.405      6.885   36.51
Days          10.467      1.559    6.71
```

Random Slope: with Days2 = Days/10 as random effect

```
Linear mixed model fit by REML
Formula: Reaction ~ Days + (1 | Subject) + (0 + Days2 | Subject)
   Data: sleepstudy
  AIC   BIC logLik deviance REMLdev
 1754 1770 −871.8     1752    1744
Random effects:
 Groups    Name        Variance Std.Dev.
 Subject  (Intercept)  627.57  25.051
 Subject  Days2       3585.82  59.882
 Residual              653.58  25.565
Number of obs: 180, groups: Subject, 18

Fixed effects:
            Estimate Std. Error t value
(Intercept)  251.405      6.885   36.51
Days          10.467      1.559    6.71
```

Reconstrukting the Random Effects

- $E(\gamma|Y) = DZ^t(I + ZDZ^t)^{-1}(Y - X\beta)$
- Plug in estimates for D and β into equation

```
> head(ranef(fm2))——random slope Modell
—Subject
      (Intercept)          Days
308     1.5138200    9.3241219
309   —40.3749105   —8.5997562
310   —39.1816682   —5.3881596
330    24.5182907   —4.9689806
331    22.9140346   —3.1941494
```

Predicting Y

- Reconstructing the systematic part of observations:

$$E(X\beta + Z\gamma|Y) = X\beta + ZE(\gamma|Y) = (I - W)X\beta + WY$$

- here $W = ZDZ^t(I + ZDZ^t)^{-1}$
- replace β and D with estimates

Prediction–model with repeated measurements

- What value do we expect for Y from future observations for a given group/subject?
- Approach from before for a given group:

$$E(X_i\beta + Z_i\gamma | Y_i) = (I - W_i)X_i\beta + W_iY_i$$

- $W_i = Z_iD_iZ_i^t(I + Z_iD_iZ_i^t)^{-1}$
- *“borrowing strength”:* Weighted mean of estimators given a specific group and using all groups:
- Weights depend on ratio “Var-within group” zu “Var between groups”.

Predictions in Example

Reaction times–predictions with/without random effects

```
     Reaction   fitted   fixest Reaction   fitted   fixest
1    249.5600 252.9189 251.4051 222.7339 211.0302 251.4051
2    258.7047 272.7103 261.8724 205.2658 212.8977 261.8724
3    250.8006 292.5017 272.3397 202.9778 214.7653 272.3397
4    321.4398 312.2931 282.8070 204.7070 216.6328 282.8070
5    356.8519 332.0846 293.2742 207.7161 218.5003 293.2742
6    414.6901 351.8760 303.7415 215.9618 220.3678 303.7415
7    382.2038 371.6674 314.2088 213.6303 222.2354 314.2088
8    290.1486 391.4588 324.6761 217.7272 224.1029 324.6761
9    430.5853 411.2502 335.1434 224.2957 225.9704 335.1434
10   466.3535 431.0416 345.6107 237.3142 227.8380 345.6107
mean 342.1338 341.9803 298.5079 215.2330 219.4341 298.5079
```

Simulations: Advantage using Shrinkage

- 10 individuals with 3 repeated measurements each
- $Y_{ij} = U_i + \epsilon_{i,j}$, $1 \leq i \leq 10$, $1 \leq j \leq 3$, $U_i \sim N(0, 10^2)$, $\epsilon_{i,j} \sim N(0, 20^2)$, i.i.d.
- Want to reconstruct U_i.
- Method 1: Estimate U_i as fixed effects
  $res < -lm(y \sim x1, data = data)$, x_1...dummy for individual
- Method 2: Estimate U_i as random effects
  $reslme < -lmer(y \sim (1|x1), data = data, REML = F)$
- $MSE(\hat{U}_i^{(fixed)}) = 1099$, $MSE(\hat{U}_i^{(random)}) = 562$

Simulation Results

```
            Ui    lme.est  fixed.est
1    5.0523638   5.333149  13.901627
2  -18.4665423 -10.387320 -27.076057
3    0.8945209  -3.557302  -9.272624
4   -5.0107337  -1.438628  -3.749993
5   25.2917984   8.473786  22.088152
6   -1.0618720  -2.709170  -7.061846
7    3.7328746  11.333587  29.542639
8  -10.8108774  -6.034696 -15.730311
9  -12.5460652  -4.103138 -10.695426
10  -2.0376523   3.089734   8.053839
```

Simulation code

```
indSD  <-  10  # sd between individuals
errSD  <-  20  # sd variance (within individuals)

x1 <- as.factor(rep(1:10,each= 3))  # individual indicator

Ui <- rnorm(10,0,sd=indSD); z1 <- rep(Ui,each=3)
y<-z1+rnorm(10*3,0,sd=errSD)

data<-data.frame(y,x1,z1)

options(contrasts=c("contr.sum","contr.sum"))
res<- lm(y~x1,data=data)

library('lme4')
reslme<- lmer(y~( 1 | x1),data=data,REML=F)

fixed.est<-c(coef(res)[-1],-sum(coef(res)[-1]))
lme.est  <- ranef(reslme)[[1]]

sum((fixed.est-Ui)^2)
sum((lme.est-Ui)^2)
```

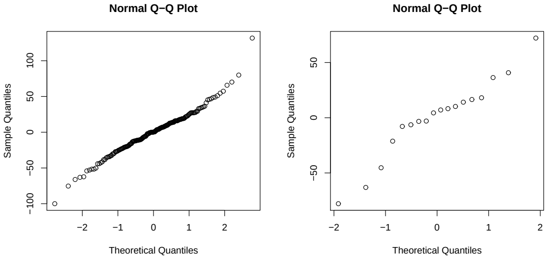
Diagnostics/visualisation

One can check...

- normal distribution of random effects and residuals
- homogeneous variances of residuals
- independence between residuals and random effects
- linearity assumption

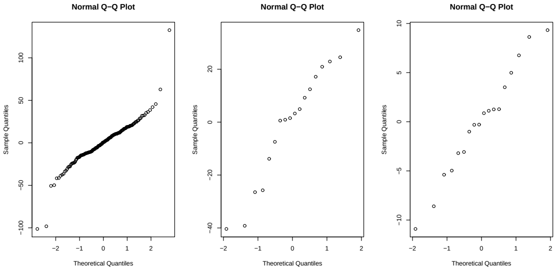
Normal distribution of residuals (le) and random effects (ri)

QQ-plot: Random intercept model



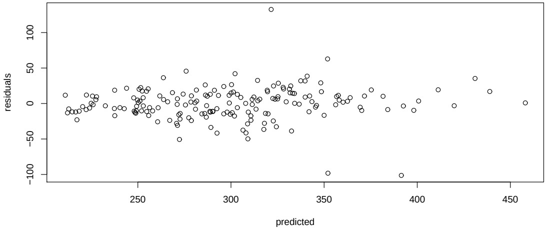
Normal distribution of residuals (le), rand. intercepts (ce), rand. slopes (ri)

QQ-plot: Random slope model



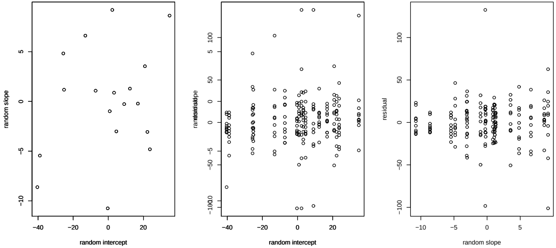
Checking homogeneous variances of the residuals

Random slope model



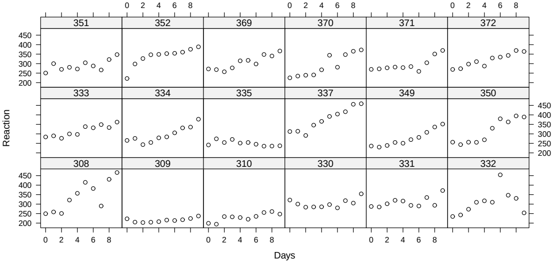
Independence Random effects/Residuals

Random slope model



scatterplot separately by individual

xy-plot



R-Code

```
# random intercept model, normal distribution of residuals
qqnorm(residuals(fm1))
qqnorm(ranef(fm1)[[1]][,1])  # and random effects

# independence random effects/residuals
plot(ranefs$Subject[,1],ranefs$Subject[,2],xlab="random intercept",ylab="random slope")
lines(lsfit(ranefs$Subject[,1],ranefs$Subject[,2]))
plot(rep(ranefs$Subject[,1],each=10),residuals(fm2),,xlab="random intercept",ylab="residual")
plot(rep(ranefs$Subject[,2],each=10),residuals(fm2),,xlab="random slope",ylab="residual")

# Variance homogeneity erros
plot(sleepstudy$Reaction-residuals(fm2),residuals(fm2),xlab="predicted",ylab="residuals")

# Individual plots
library(lattice)
xyplot(Reaction~Days|Subject,data=sleepstudy)
```

Mixed Models for Block Design

Penicillin manufactured in 4 ways (treat) using 5 blends of ingredients. The response is "yield". (blend=Block)

```
> library(faraway)
> data(penicillin)
> head(penicillin)
  treat  blend yield
1     A Blend1    89
2     B Blend1    88
3     C Blend1    97
4     D Blend1    94
5     A Blend2    84
6     B Blend2    77
```

Mixed models for Block Design

Model can be estimated either via ANOVA oder using mixed models.

$$Y_{i,j} = \alpha_i + \beta_j + \epsilon_{i,j}$$

- ANOVA: β_j is (nuisance) parameter–not of primary interest
- Mixed Model: β_j ist normally distributed random effect $N(0, \sigma_\beta^2)$.

Mixed Models for Block Design
ANOVA

```
Call:
lm(formula = yield ~ treat + blend, data = penicillin)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  86.0000     0.9704  88.624  < 2e-16 ***
treat1       -2.0000     1.6808  -1.190  0.25708
treat2       -1.0000     1.6808  -0.595  0.56292
treat3        3.0000     1.6808   1.785  0.09956 .
blend1        6.0000     1.9408   3.092  0.00934 **
blend2       -3.0000     1.9408  -1.546  0.14812
blend3       -1.0000     1.9408  -0.515  0.61573
blend4        2.0000     1.9408   1.031  0.32310
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.34 on 12 degrees of freedom
Multiple R-squared: 0.5964, Adjusted R-squared: 0.361
F-statistic: 2.534 on 7 and 12 DF,  p-value: 0.07535
```

Mixed Models for Block Design

```
> res <− lmer(yield~treat+(1|blend),data=penicillin)
> summary(res)
Linear mixed model fit by REML
Formula: yield ~ treat + (1 | blend)
   Data: penicillin
   AIC   BIC logLik deviance REMLdev
 118.6 124.6  −53.3    117.3   106.6
Random effects:
 Groups   Name        Variance Std.Dev.
 blend    (Intercept) 11.792   3.4339
 Residual             18.833   4.3397
Number of obs: 20, groups: blend, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)   86.000      1.817   47.34
treat1        −2.000      1.681   −1.19
treat2        −1.000      1.681   −0.59
treat3         3.000      1.681    1.78
```

Mixed models for Block Design

ANOVA

```
> anova(lmod)
Analysis of Variance Table

Response: yield
          Df Sum Sq Mean Sq F value   Pr(>F)
treat      3     70  23.333  1.2389 0.33866
blend      4    264  66.000  3.5044 0.04075 *
Residuals 12    226  18.833
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
```

Hypotheses tests within mixed models

- Classical t-Tests for parameters: Distribution under null hypothesis: t-distribution only works approximatively, classical formula for degrees of freedom does not work.
- Likelihood-ratio test: χ^2–distribution is usually not very accurate as H_0 distribution.
- χ^2–distribution in particular problematic when testing random effects: under H_0 $Var = 0$, d.h. we are at the boundary of the parameter space. *(Alternative: Bootstrap)*
- When applying LR-Tests: Fit model using option ML instead of REML!

Likelihood-Ratio test for mixed model

```
> res0 <- lmer(yield~treat+(1|blend),data=penicillin,REML=FALSE)
> res1 <- lmer(yield~(1|blend),data=penicillin,REML=FALSE)
>
>
> anova(res0,res1)
Data: penicillin
Models:
res1: yield ~ (1 | blend)
res0: yield ~ treat + (1 | blend)
     Df    AIC    BIC  logLik  Chisq Chi Df Pr(>Chisq)
res1  3 127.33 130.31 -60.662
res0  6 129.28 135.25 -58.639 4.0474      3     0.2564
```

Likelihood-Ratio test for mixed model

Critical values Werte using bootstrap

```
res1  <- lmer(yield~treat+(1|blend),data=penicillin,REML=FALSE)
res0  <- lmer(yield~(1|blend),data=penicillin,REML=FALSE)
tstat <- 2*(logLik(res1)-logLik(res0))

> lrstat <- rep(0,1000)
> for (i in 1:1000){
+ penicillin$sy<-unlist(simulate(res0))
+ sim1  <- lmer(sy~treat+(1|blend),data=penicillin,REML=FALSE)
+ sim0  <- lmer(sy~(1|blend),data=penicillin,REML=FALSE)
+ lrstat[i] <- 2*(logLik(sim1)-logLik(sim0))
+ }
> mean(lrstat<tstat)

   0.635
```

Linear Mixed Models II

Andreas Futschik

IFAS

07/2016

Generalized mixed models

- Exponential family:

$$f(y|\theta,\phi) = \exp\left[\frac{y\theta - b(\theta)}{a(\phi)} - c(y,\phi)\right]$$

- $\mathrm{E}(Y) = \mu = b'(\theta)$
- Variance function (models relationship expected value–variance: $\mathrm{Var}(Y) = b''(\theta)a(\phi)$
- Link function g: $\mathrm{E}(Y) = \mu = g^{-1}(X\beta)$, or

$$g(\mathrm{E}(Y)) = X\beta$$

- Estimate β using maximum likelihood:

$$\max_{\beta,\phi} f(y, b'^{-1}(g^{-1}[X\beta]),\ \phi)$$

Generalized linear model: Examples

	$b(\theta)$	$g(\mu)$
Binomial	$-n\log(1+\exp(\theta))$	$\log(\mu/1-\mu)$
Poisson	$\exp(\theta)$	$\log(\mu)$
Gamma		
Negative Binomial		
...		

Example: Logistic regression

- Logit-Modell (Y... 0/1 (Bernoulli)):

$$P_\beta(Y=1)=\frac{\exp(\beta_0+\sum_{i=1}^k \beta_i X_i)}{1+\exp(\beta_0+\sum_{i=1}^k \beta_i X_i)}$$

- With one explanatory variable S-shaped function:

$$P_\beta(Y=1)=\frac{\exp(\beta_0+\beta_1 X)}{1+\exp(\beta_0+\beta_1 X)}$$

- Alternative: Probit Modell

$$P_\beta(Y=1)=\Phi(\beta_0+\beta_1 X)$$

Logit-Transformation

- Write: $\pi = P(Y = 1)$
- We have:

$$\log(\frac{\pi}{1-\pi}) = \beta_0 + \sum_{i=1}^{k} \beta_i X_i$$

- $\log(\frac{\pi}{1-\pi})$ is called log-odds or logit.
- parameter estimation using maximum likelihood (iterative numerical algorithm)

Interpretation of coefficients

- Consequence of increase in X_j on odds:

$$X_j \to X_j + 1 \Longrightarrow \frac{\pi}{1-\pi} \to \frac{\pi}{1-\pi} e^{\beta_j}$$

- Odds ratio for $X_j + 1$ versus X_j: e^{β_j}

Generalized mixed model

- Like in the classical GLM (Exponential family):

$$f(y|\theta,\phi) = \exp\left[\frac{y\theta - b(\theta)}{a(\phi)} - c(y,\phi)\right]$$

- Link-Funktion g: $\mathrm{E}(Y) = \mu = \mathrm{E}g^{-1}(X\beta + Z\gamma)$, or

$$\mathrm{E}(g(Y)) = X\beta + Z\gamma$$

- Estimate β using maximum likelihood:

$$\max_{\beta,\phi} f(y, b'^{-1}(g^{-1}[X\beta + Z\gamma]),\ \phi)$$

Generalized mixed model

- Like in the classical GLM (Exponential family):
$$f(y|\theta,\phi) = \exp\left[\frac{y\theta - b(\theta)}{a(\phi)} - c(y,\phi)\right]$$
- Link function g: $\mathrm{E}(Y|\gamma) = \mu = g^{-1}(X\beta + Z\gamma)$, or
$$\mathrm{E}(Y) = \mathrm{E}g^{-1}(X\beta + Z\gamma)$$
- Be careful: For g non-linear, i.g. no closed form formula for $\mathrm{E}(Y)$, especially
$$E(Y) \neq g^{-1}(X\beta)!$$
- Exception: probit link, logistic regression: $E(Y) = \Phi(\tau X\beta)$, $\tau^2 = 1 + ZDZ^t$.
- Estimate β, γ using maximum likelihood:
$$\max_{\beta,\gamma,\phi} f(y, b'^{-1}(g^{-1}[X\beta + Z\gamma]), \phi)$$

Example: Stable standing

- Balance/stable standing of test persons investigated
- Response: stable standing of test persons
- Explanatory variables: Type of underground (normal/unstable), vision: eyes closed/open/blinded?, Age, gender, etc...
- 12 measurements per person, two per combination underground*vision

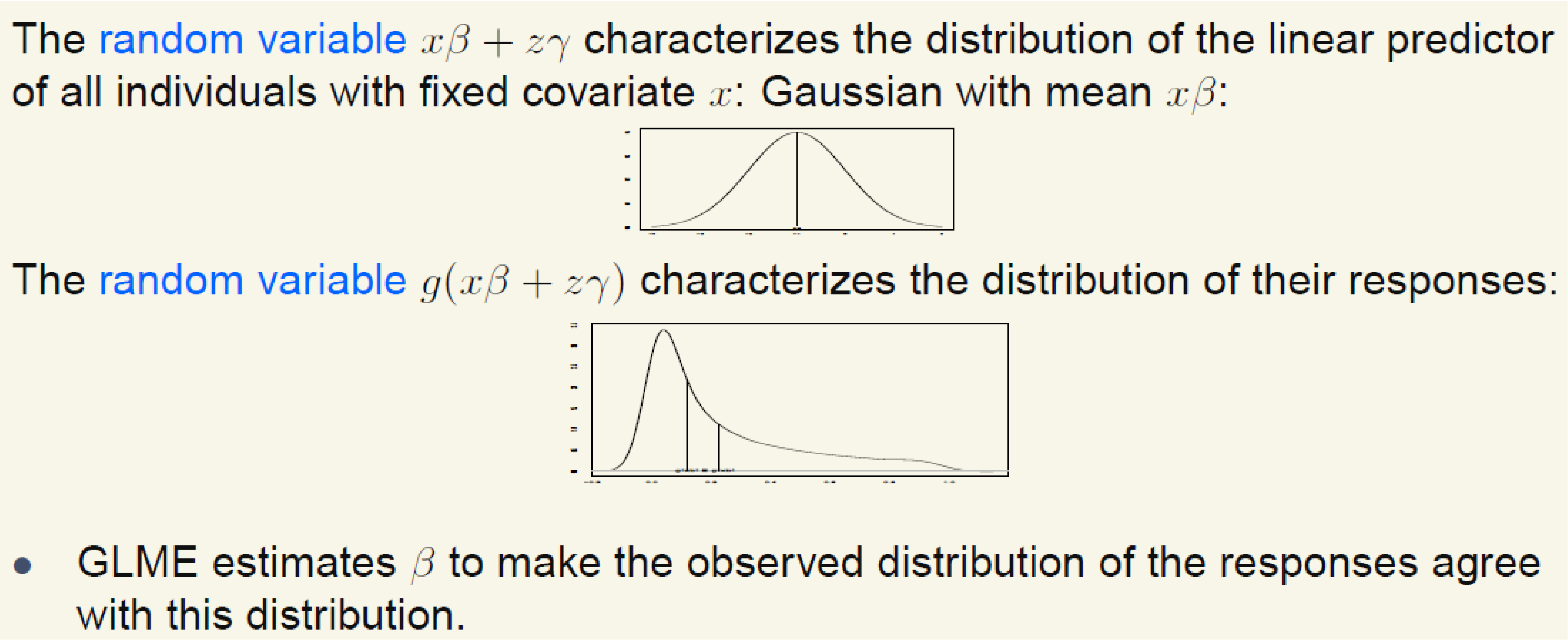
Results

```
> library(MASS)
> gg <— glmmPQL(stable ~ Sex+Age+Height+Weight+Surface+Vision,

  random=~1 | Subject, family=binomial,data=ctsib)
> summary(gg)

Random effects:
Formula: "1 | Subject
        (Intercept) Residual
StdDev:      3.0608  0.59062

Variance function:
Structure: fixed weights
Formula: ~invwt
Fixed effects: stable ~ Sex + Age + Height + Weight + Surface + Vision

              Value Std.Error  DF t—value p—value

(Intercept) 15.5716   13.4985 437  1.1536  0.2493
Sexmale      3.3554    1.7526  35  1.9145  0.0638
Age         —0.0066    0.0820  35 —0.0810  0.9359
Height      —0.1908    0.0920  35 —2.0736  0.0455
Weight       0.0695    0.0629  35  1.1051  0.2766
Surfacenorm  7.7241    0.5736 437 13.4666  0.0000
Visiondome   0.7265    0.3259 437  2.2289  0.0263
Visionopen   6.4853    0.5440 437 11.9220  0.0000
```

Supplementary Material: GLMM

(adopted from van der Vaart)

Interpretation GLMM



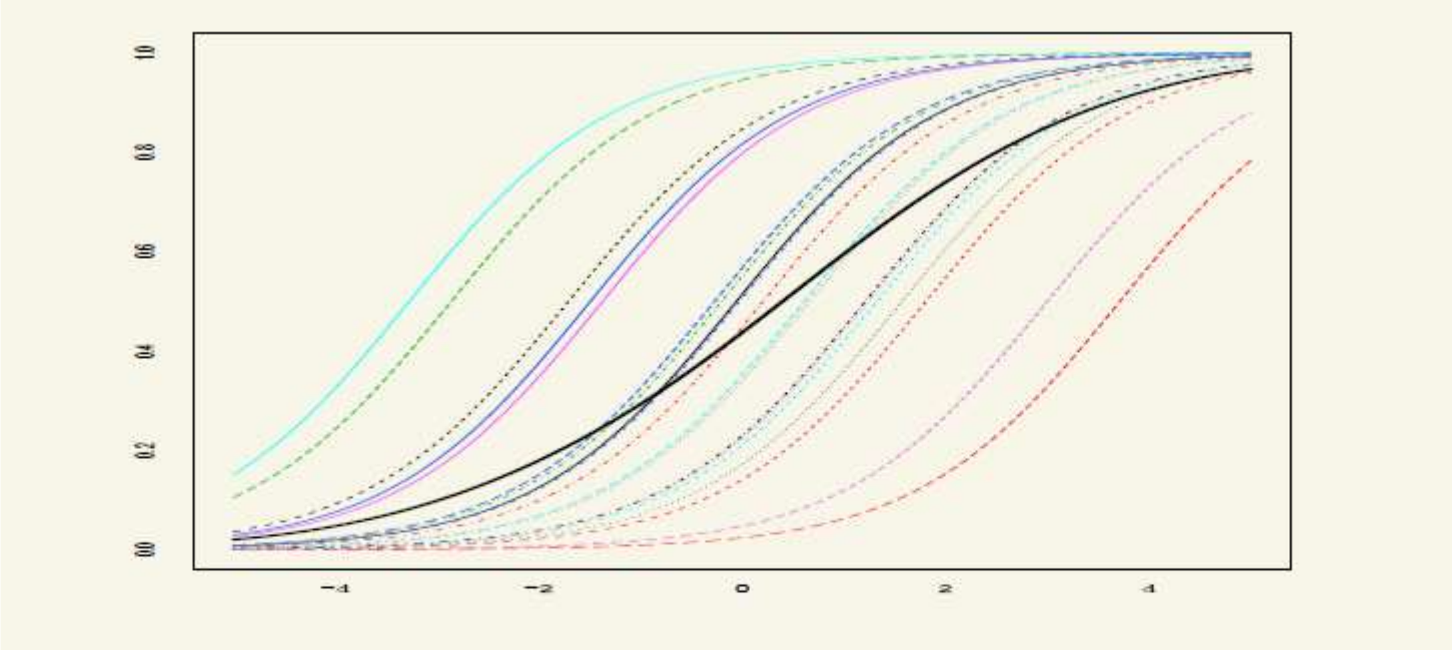
The random variable $x\beta + z\gamma$ characterizes the distribution of the linear predictor of all individuals with fixed covariate x: Gaussian with mean $x\beta$:

The random variable $g(x\beta + z\gamma)$ characterizes the distribution of their responses:

- GLME estimates β to make the observed distribution of the responses agree with this distribution.

- GLM/GEE estimates β so that $g^{-1}(X\beta)$ is close to the mean of this distribution.
- For monotone Link-functions & GLMM=GLME: $g^{-1}(X\beta)=\text{median}(g^{-1}(X\beta+Z\gamma))$

Mean Response



- The logistic function $x\beta \rightarrow g^{-1}(x\beta + z\gamma)$ gives the probability of response 1 as a function of $x\beta$, for a given random effect $z\gamma$.
- The random effects $z\gamma$ shift this curve left or right.
- The resulting population of curves has mean curve $x\beta \rightarrow E[g^{-1}(x\beta + z\gamma)]$.
- Because the mean curve is flatter, the GLM/GEE estimates are smaller than the GLME estimates for β.