Linear and generalized linear mixed-effects model in

Matteo Lisi
Laboratoire Psychologie de la Perception
Universitè Paris Descartes

Linear model

- In basic linear models, the dependent variable is modeled as a weighted linear combination of n independent variables (predictors) with an additive error ε

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \ldots + \beta_n X_{n,i} + \varepsilon_i$$

$$\varepsilon \sim N\left(0, \sigma^2\right)$$

Linear model

- In basic linear models, the dependent variable is modeled as a weighted linear combination of n independent variables (predictors) with an additive error ε

$$Y = \boldsymbol{X}\beta + \varepsilon$$

$$\varepsilon \sim N\left(0, \sigma^2\right)$$

Matrix notation:

$$Y = \boldsymbol{X}\beta + \varepsilon$$

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_p \end{pmatrix} = \begin{pmatrix} 1 & X_{11} & \cdots & X_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{p1} & \cdots & X_{pn} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_p \end{pmatrix}$$

Matrix multiplication:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} a\beta_1 + b\beta_2 \\ c\beta_1 + d\beta_2 \end{pmatrix}$$

Linear mixed-effects models

- Basic linear models can be considered as *fixed-effects* only, i.e. the independent variables are not random (experimental manipulations)

- The only random source of variation is the residual error $\varepsilon \sim N(0, \sigma^2)$

- However, in most cases some of the independent variables represent random samples from a larger population on which we would like to draw conclusions (e.g., individual participants)

- If we want to generalize from the sample to the population, these variables must be treated as *random effects*

- A model containing both *fixed*- and *random-effects* is called a *mixed-effects* model

- *Mixed-effects* models are used primarily to describe the relationship between a dependent variable and some independent variables that are grouped according to one or more classification factors

- A typical example is **repeated measures data** where observations are grouped according to the subject: in this case common random effects are associated with observations made on the same subject (i.e., sharing the same level of the classification factor)

- Random effects are treated as random variations around a population mean

- The dependent variable is taken conditionally on the random effects, and modeled as a sum of a fixed effect term $\boldsymbol{X}$ and a random effect term $\boldsymbol{Z}$

$$\left(Y \mid b\right) = \boldsymbol{X}\beta + \boldsymbol{Z}b + \varepsilon$$

$$\varepsilon \sim N\left(0, \sigma^2\right)$$

$$b \sim N\left(0, \Sigma\right)$$

$$\left(y|b\right)=\beta_0+\beta_1 x+b_s+\varepsilon$$

simple random additive
term: *random intercept*

$$\left(y|b\right)=\beta_0+\beta_1 x+b_{0,s}+b_{1,s}x+\varepsilon$$

random slope

Pinheiro, J. C., & Bates, D. M. (2000).
Mixed-Effects Models in S and S-PLUS.
New York: Springer-Verlag.
doi:10.1007/b98882

library **`nlme`** → linear & nonlinear mixed-effects model

library **`lme4`** → linear,nonlinear & generalized linear mixed-effects model

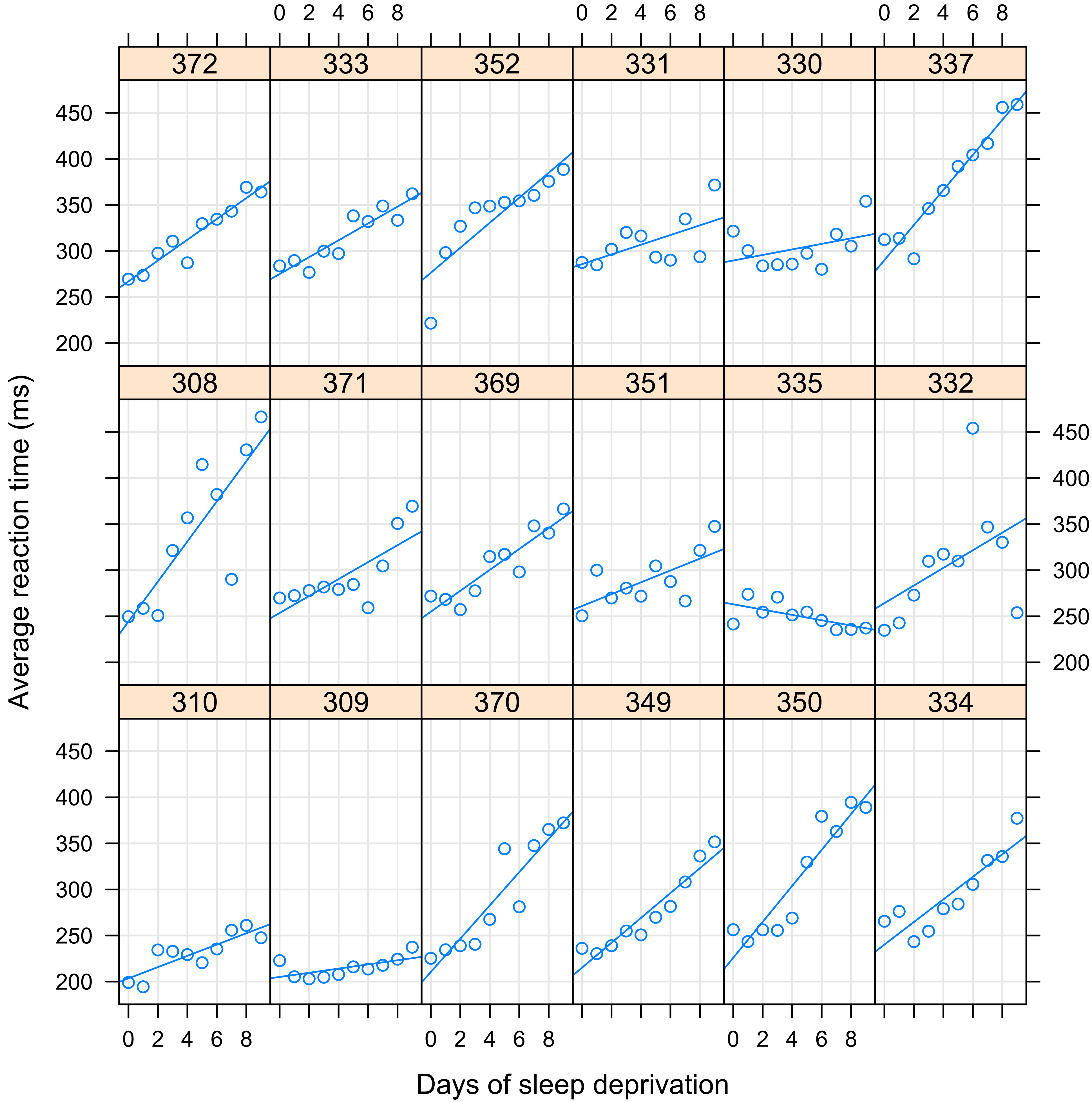
see also http://lme4.r-forge.r-project.org/

Example 1

The **sleepstudy** dataset, included in the package **lme4**, contains data from a study of the effects of sleep deprivation on reaction times of long-distance truck drivers

It includes only data from a group of 18 subjects who were restricted to 3 hours of sleep per night for 10 days

```
> library(lme4)
> str(sleepstudy)
'data.frame':	180 obs. of  3 variables:
 $ Reaction: num  250 259 251 321 357 ...
 $ Days    : num  0 1 2 3 4 5 6 7 8 9 ...
 $ Subject : Factor w/ 18 levels "308","309","310",..: 1 1 1 1 1 1 1 1 1 
1 ...
```



372
333
352
331
330
337
308
371
369
351
335
332
310
309
370
349
350
334
Average reaction time (ms)
Days of sleep deprivation
200
250
300
350
400
450
0
2
4
6
8

- Two fixed-effects parameters (intercept and slope)
- Two random-effects for each subjects (individual random variations in intercept and slope)

```
> sleep.m <- lmer(Reaction ~ Days + (Days|Subject), data = sleepstudy)
```

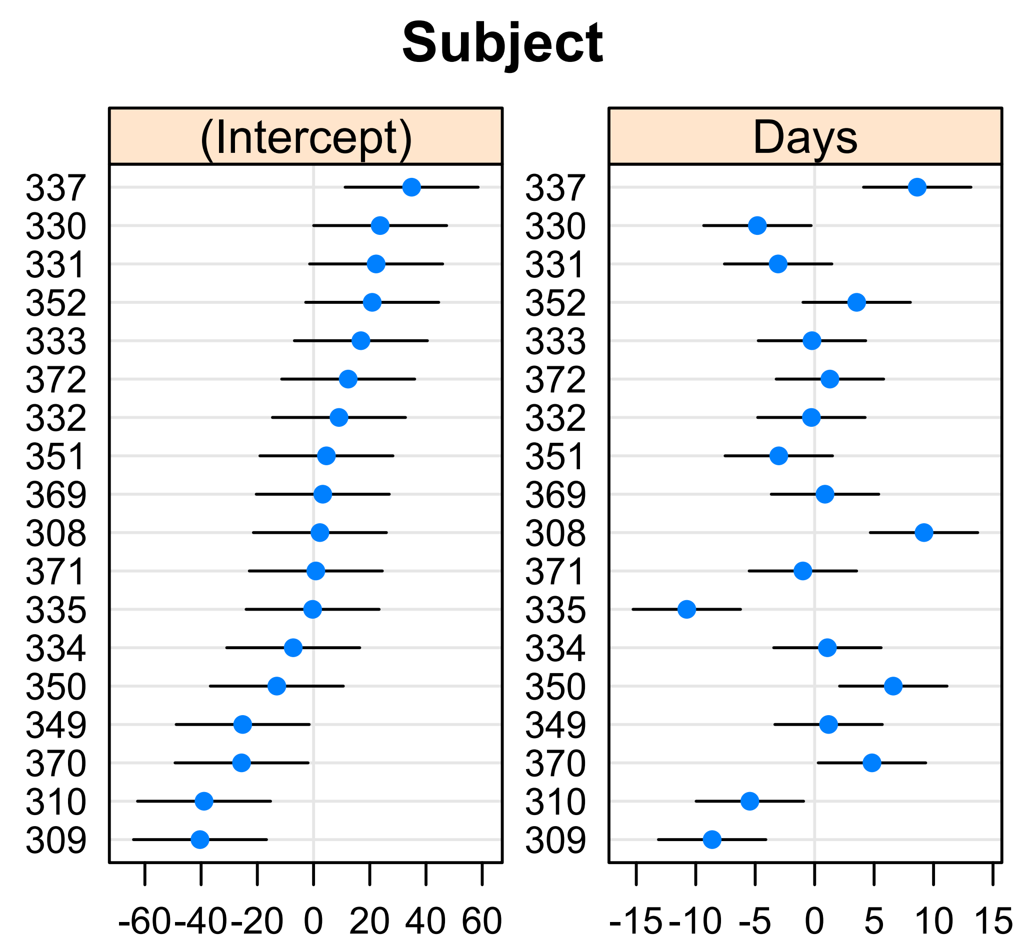
the coefficient for the intercept (1) is implicit; an equivalent formulation is:

```
Reaction ~ 1 + Days + (1 + Days|Subject)
```

grouping factor on the right of the “|”

```
> summary(sleep1.m)
Linear mixed model fit by REML ['lmerMod']
Formula: Reaction ~ Days + (Days | Subject)
   Data: sleepstudy

REML criterion at convergence: 1743.6

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.9536 -0.4634  0.0231  0.4634  5.1793

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Subject  (Intercept) 612.09   24.740
          Days         35.07    5.922   0.07
 Residual             654.94   25.592
Number of obs: 180, groups:  Subject, 18

Fixed effects:
            Estimate Std. Error t value
(Intercept)  251.405      6.825   36.84
Days          10.467      1.546    6.77

Correlation of Fixed Effects:
     (Intr)
Days -0.138
```

The distribution of the random effects allows for correlation of the random effects on the same subject



Typical initial RT is 251 ms. The reaction time increases about 10 ms for each day of sleep deprivation

- A model with uncorrelated random effect

$$b \sim N(0, \Sigma) \qquad \Sigma = \sigma^2 I$$

```
> sleep2.m <- lmer(Reaction ~ Days + (1|Subject) + (0 + Days|Subject),
data = sleepstudy)
```

- The two models (with correlated vs. uncorrelated random effects) can be compared with a likelihood ratio test

$$LRT = -2\ln\left(L_0 / L_1 \right) \qquad LRT \sim \chi^2\left(df_1 - df_0\right)$$

```
> anova(sleep1.m, sleep2.m)
refitting model(s) with ML (instead of REML)
Data: sleepstudy
Models:
sleep2.m: Reaction ~ Days + (1 | Subject) + (0 + Days | Subject)
sleep1.m: Reaction ~ 1 + Days + (1 + Days | Subject)
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
sleep2.m  5 1762.0 1778.0 -876.00   1752.0
sleep1.m  6 1763.9 1783.1 -875.97   1751.9 0.0639      1     0.8004
```

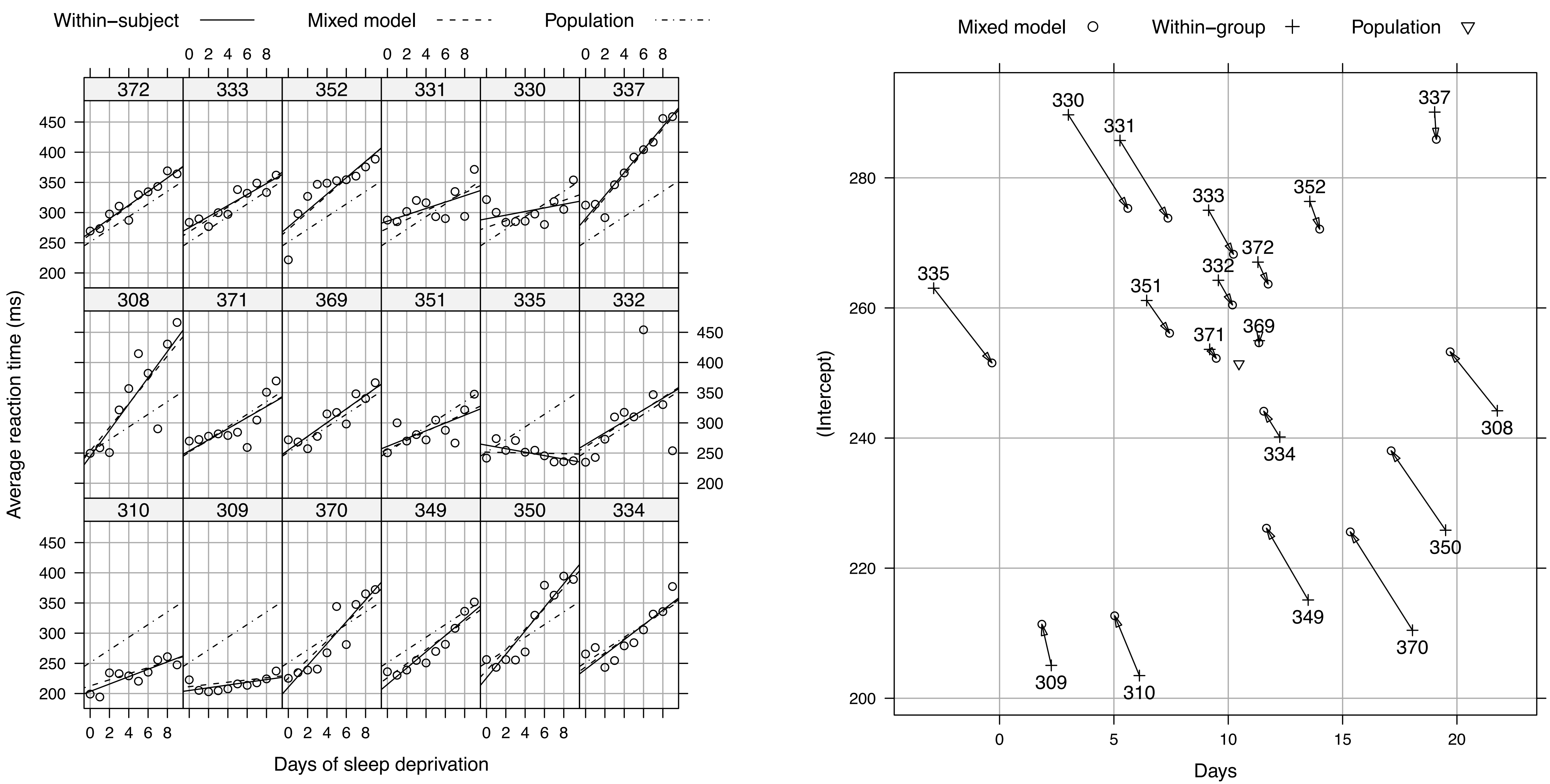
- A likelihood ratio test can also be used to obtain a *p* value for the fixed-effects parameters

```
> anova(sleep2.m, update(sleep2.m, .~. - Days))
refitting model(s) with ML (instead of REML)
Data: sleepstudy
Models:
update(sleep2.m, . ~ . - Days): Reaction ~ (1 | Subject) + (0 + Days | Subject)
sleep2.m: Reaction ~ Days + (1 | Subject) + (0 + Days | Subject)
                               Df    AIC    BIC logLik deviance Chisq Chi Df Pr(>Chisq)
update(sleep2.m, . ~ . - Days)  4 1783.6 1796.4 -887.8   1775.6
sleep2.m                        5 1762.0 1778.0 -876.0   1752.0  23.6      1  1.186e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

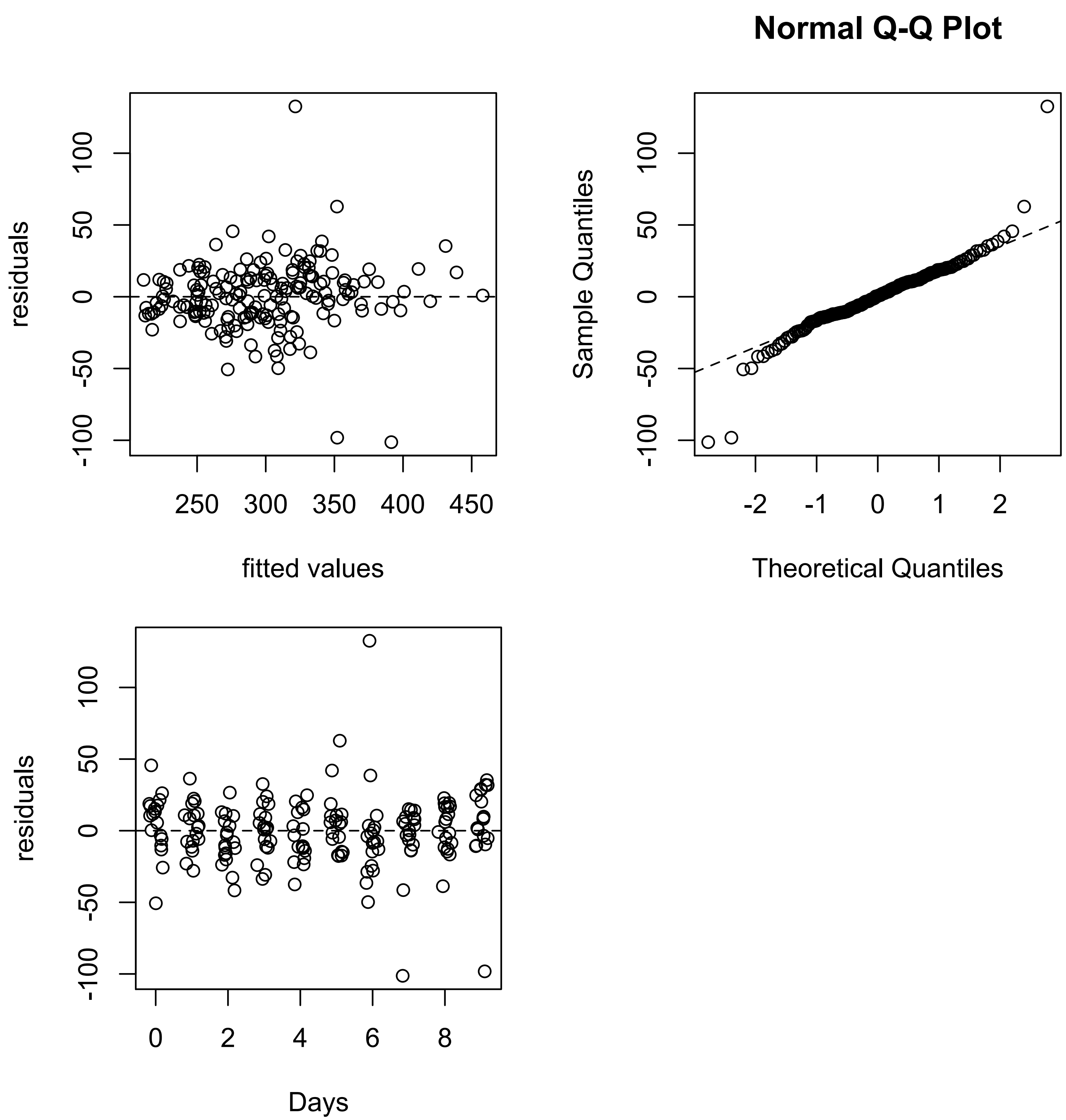
- …but a better way to test fixed effects is bootstrap

```
> confint.merMod(sleep2.m, nsim=1000)
Computing profile confidence intervals ...
                 2.5 %     97.5 %
.sig01       15.258647  37.786472
.sig02        3.964074   8.769159
.sigma       22.880555  28.787598
(Intercept) 237.572148 265.238062
Days          7.334067  13.600505
```

Comparison with within-subjects estimates



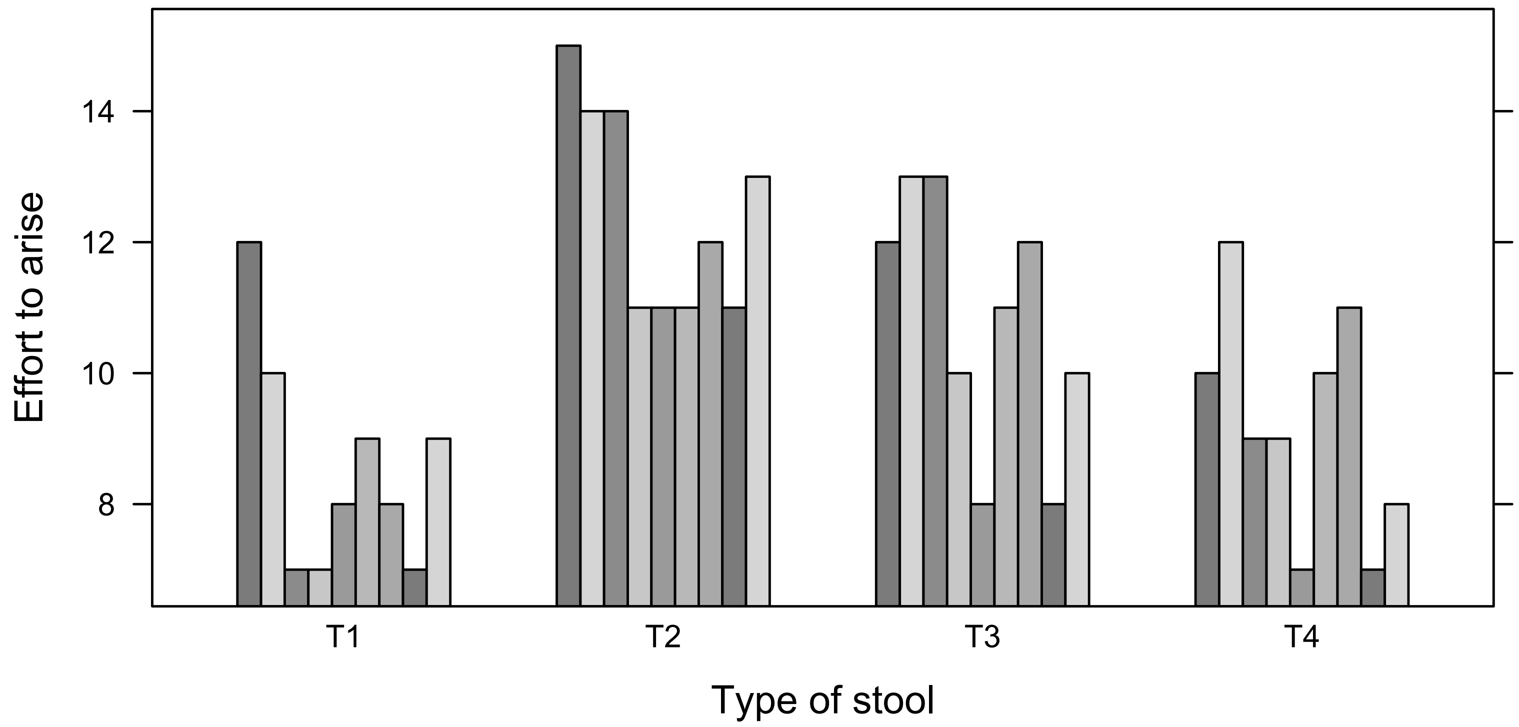
- **Diagnostics:** as for other linear models, it is important to check if the residuals have constant variance, are independent and normally distributed



Example 2

The **ergoStool** dataset, included in the package **MEMSS**, contains data from an ergonomic study in which 9 subjects evaluated the difficulty to arise of 4 types of stool

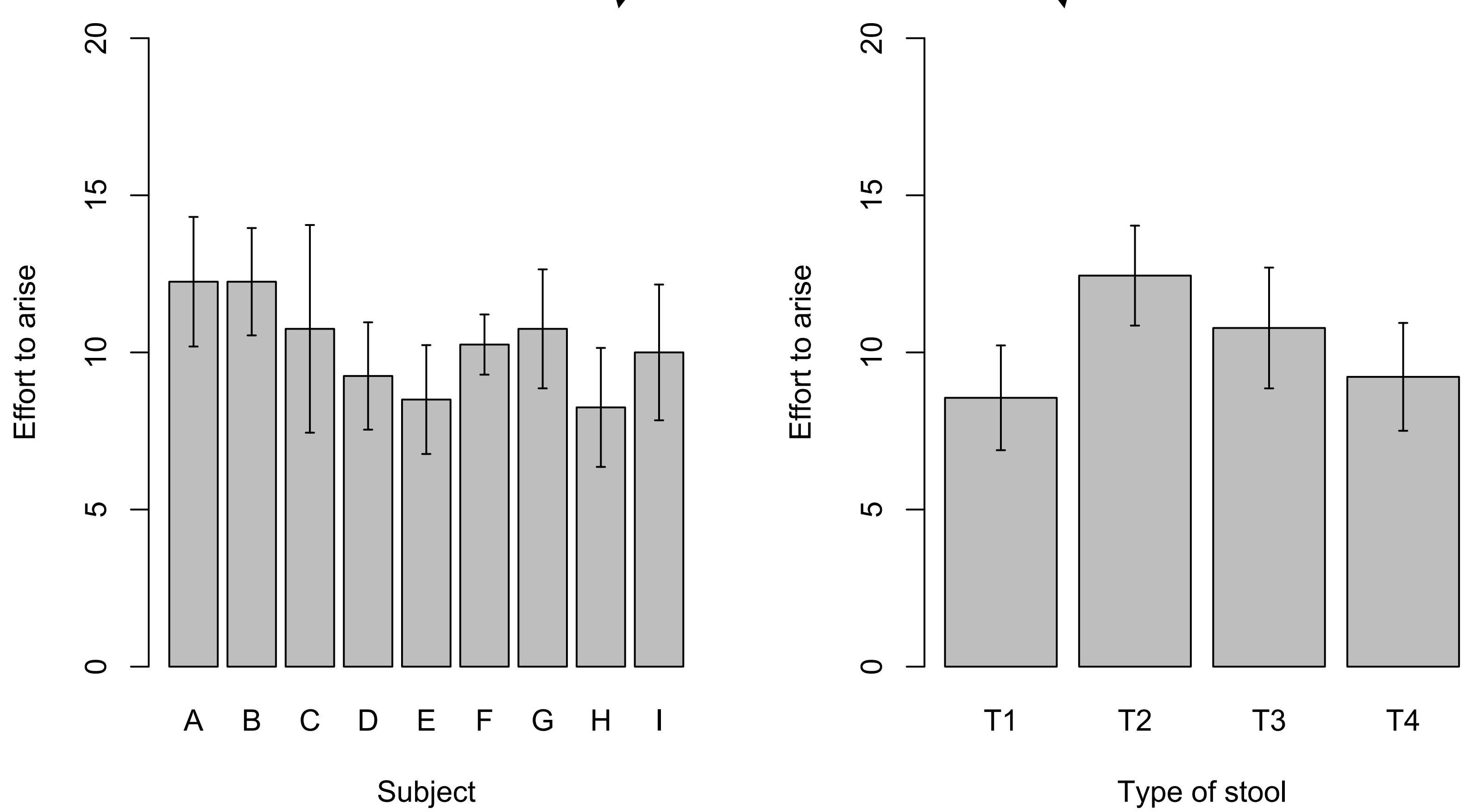
```
> data(ergoStool,package="MEMSS")
> str(ergoStool)
'data.frame': 36 obs. of  3 variables:
 $ effort : num  12 15 12 10 10 14 13 12 7 14 ...
 $ Type   : Factor w/ 4 levels "T1","T2","T3",..: 1 2 3 4 1 2 3 4 1 2 ...
 $ Subject: Factor w/ 9 levels "A","B","C","D",..: 1 1 1 1 2 2 2 2 3 3 ...
```



Repeated measures-design

grouping factor (individual observational units) *random-effects* (between subjects)

experimental manipulation: *fixed-effects* (within subjects)



```
> data(ergoStool,package="MEMSS")
> stool1.m <- lmer(effort ~ Type + (1|Subject), ergoStool)
> summary(stool1.m)
Linear mixed model fit by REML ['lmerMod']
Formula: effort ~ Type + (1 | Subject)
   Data: ergoStool

REML criterion at convergence: 121.1

Scaled residuals:
     Min       1Q   Median       3Q      Max
-1.80200 -0.64317  0.05783  0.70100  1.63142

Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 1.775    1.332
 Residual             1.211    1.100
Number of obs: 36, groups:  Subject, 9

Fixed effects:
            Estimate Std. Error t value
(Intercept)   8.5556     0.5760  14.853
TypeT2        3.8889     0.5187   7.498
TypeT3        2.2222     0.5187   4.284
TypeT4        0.6667     0.5187   1.285

Correlation of Fixed Effects:
       (Intr) TypeT2 TypeT3
TypeT2 -0.450
TypeT3 -0.450  0.500
TypeT4 -0.450  0.500  0.500
```

Stool types T2, T3, and T4 are tested against T1 (so the coefficients represent the difference from T1)

The intercept indicates the mean value for stool T1

```
> contrasts(ergoStool$Type)
   T2 T3 T4
T1  0  0  0
T2  1  0  0
T3  0  1  0
T4  0  0  1
```

Visualize the contrast matrix for factor Type

```
> # use model parameters to test contrasts of interests
> # you can also adjust the confidence level of the interval
> # to correct for multiple comparisons
> confint(stool1.m, parm=4:6, level = 1 - 0.05/6)
Computing profile confidence intervals ...
          0.417 % 99.583 %
TypeT2  2.5109497 5.266828
TypeT3  0.8442830 3.600161
TypeT4 -0.7112726 2.044606
```

T2 and T3 are significantly different from T1

```
> stool2.m <- lmer(effort ~ Type + (1|Subject), within(ergoStool, Type <-
relevel(Type, ref = "T2")))
> confint(stool2.m, parm=5:6, level = 1 - 0.05/6) # T3, T4 vs T
Computing profile confidence intervals ...
         0.417 %   99.583 %
TypeT3 -3.044606 -0.2887275
TypeT4 -4.600161 -1.8442831
```

T3 and T4 are significantly different from T2

Generalized linear model

- Generalization of linear models in which the linear predictor is related to the response variable by a *link function*

$$\phi^{-1}\left[P\left(Y=1\right)\right]=\boldsymbol{X}\beta$$

- In the example the link function is the inverse of the cumulative distribution function of the standard-normal distribution (*probit* model)

$$\phi\left(z\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-w^2/2}\,dw$$

- The coefficients of the GLM (the linear predictor part) can be directly translated into the parameter of the probability function (i.e., the psychometric function)

$$\boldsymbol{X}\beta = \beta_0 + \beta_1 X$$

$$\mu = -\beta_0 / \beta_1 \longrightarrow$$ *location parameter* (or *PSE*, threshold,…)

$$\sigma = 1 / \beta_1 \longrightarrow$$ *scale parameter*

- Normally psychometric functions like ϕ are fitted separately for each participants and conditions, and the individual estimates of parameters of interest are used as input for group analysis

- Therefore, **group analysis does not take into account the subject-specific standard error, or the number of repetitions or trials**.

- Inferences from this two-level analysis (individual and group) apply, strictly speaking, only to the sample studied and not to the general population

Generalized linear mixed-effects models

- Generalized linear models can be extended to include *random* variation both in the *location* (criterion) and *scale* (sensitivity) parameter of the psychometric functions

$$\phi^{-1}\left[P\left(Y=1 \mid b\right)\right]=\boldsymbol{X}\beta+\boldsymbol{Z}b$$

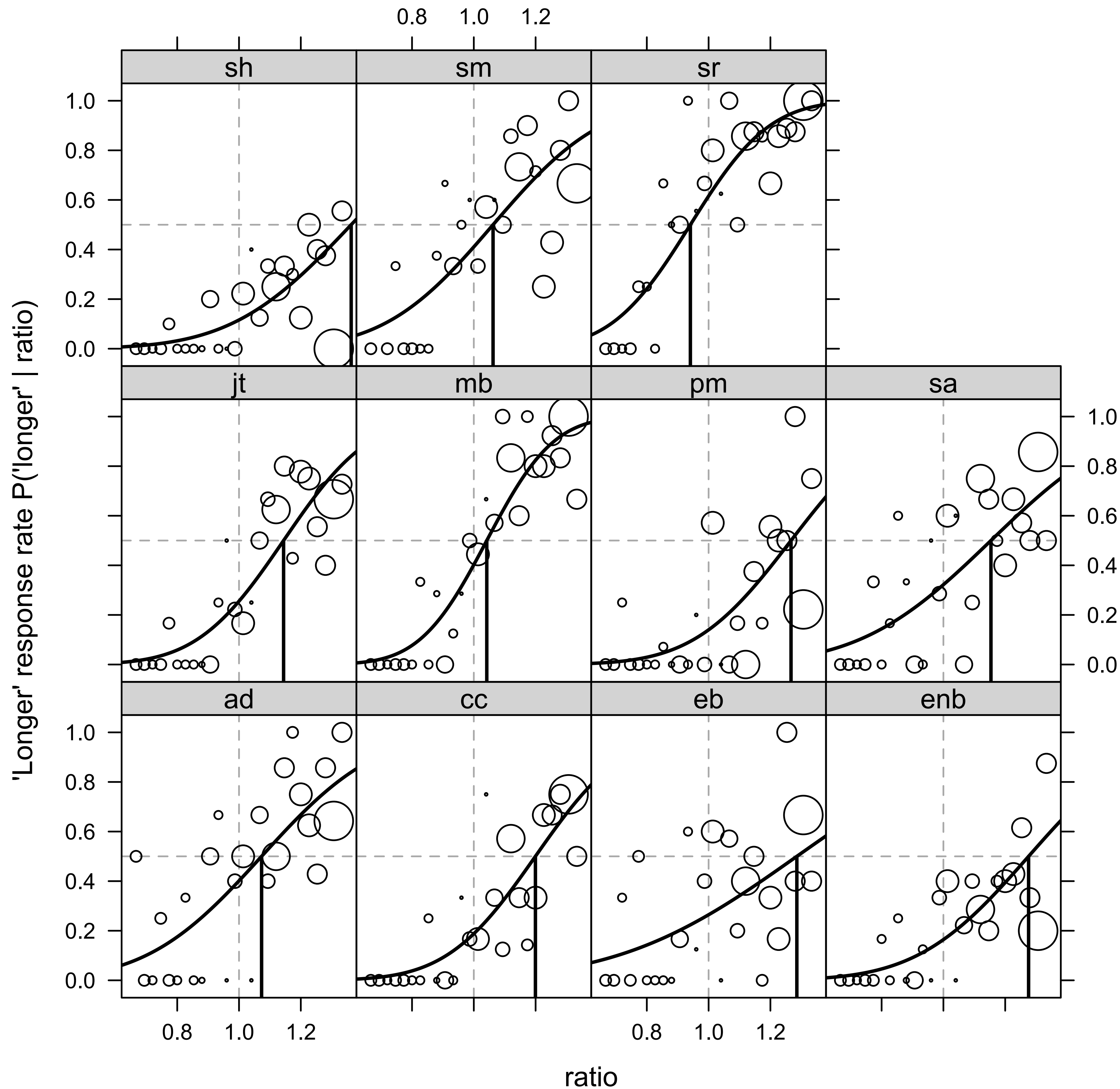
$$b \sim N\left(0,\Sigma\right)$$

Example 1

The dataset in “blinkStudy.txt” contains data from a study of the effects of voluntary eye blinks on the perceived durations of visual stimuli.

```
> bridge <- read.table("blinkStudy.txt",header=T,sep="\t")
> str(bridge)
'data.frame': 1615 obs. of  3 variables:
 $ SUBJ: Factor w/ 11 levels "ad","cc","eb",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ DUR : int  290 490 380 410 460 330 450 440 420 450 ...
 $ RESP: int  0 1 0 1 1 0 1 1 1 1 ...
```

Participants were asked to judge the duration of a visual stimulus (uniformly distributed between 250 and 500ms) with reference to the average duration. They were asked to blink during the stimulus presentation. How does blinking affect the estimated duration?



sh
sm
sr
jt
mb
pm
sa
ad
cc
eb
enb
0.8
1.0
1.2
0.0
0.2
0.4
0.6
0.8
1.0
'Longer' response rate P('longer' | ratio)
ratio

The model will have both *random* location and scale parameter: in the case of a cumulative gaussian, it means there will be individual variation both in the mean (μ) and standard deviation (σ)

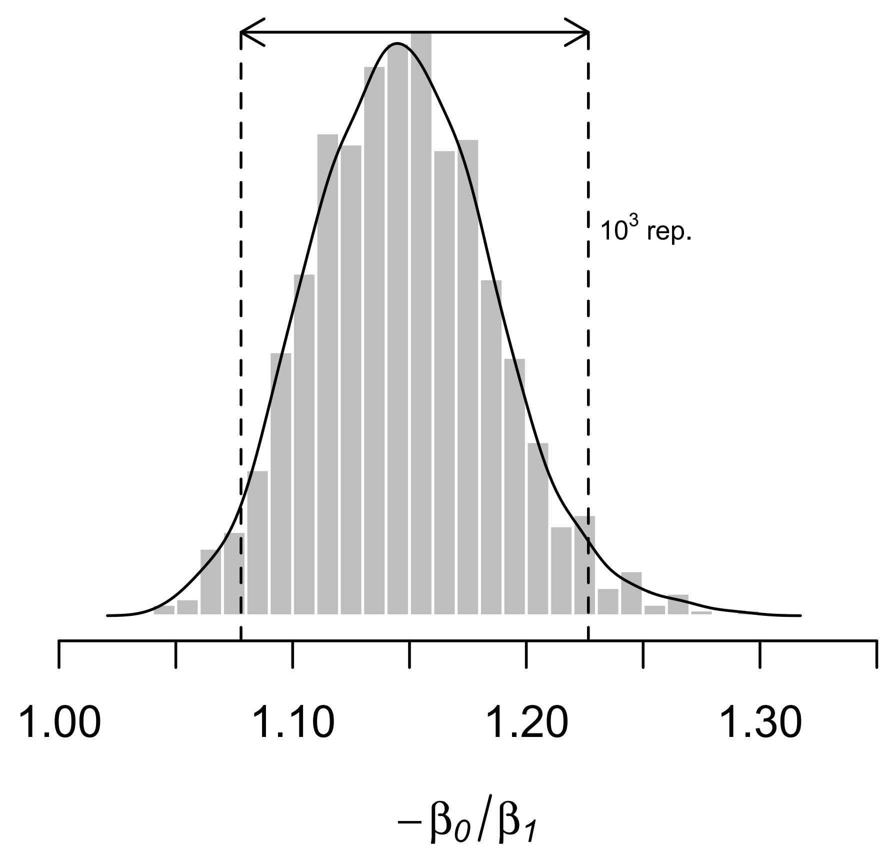
```
> bridge.m <- glmer(RESP ~ ratio + (ratio|SUBJ), data=bridge, family =
binomial(link=probit))

> summary(bridge.m)

(...)

Random effects:
 Groups Name        Variance Std.Dev. Corr
 SUBJ   (Intercept) 0.3225   0.5679
        ratio       0.4920   0.7014   -0.83
Number of obs: 1615, groups:  SUBJ, 11

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -4.2619     0.2912  -14.63   <2e-16 ***
ratio         3.7167     0.3026   12.28   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(...)
```

The average PSE can be computed from the parameters

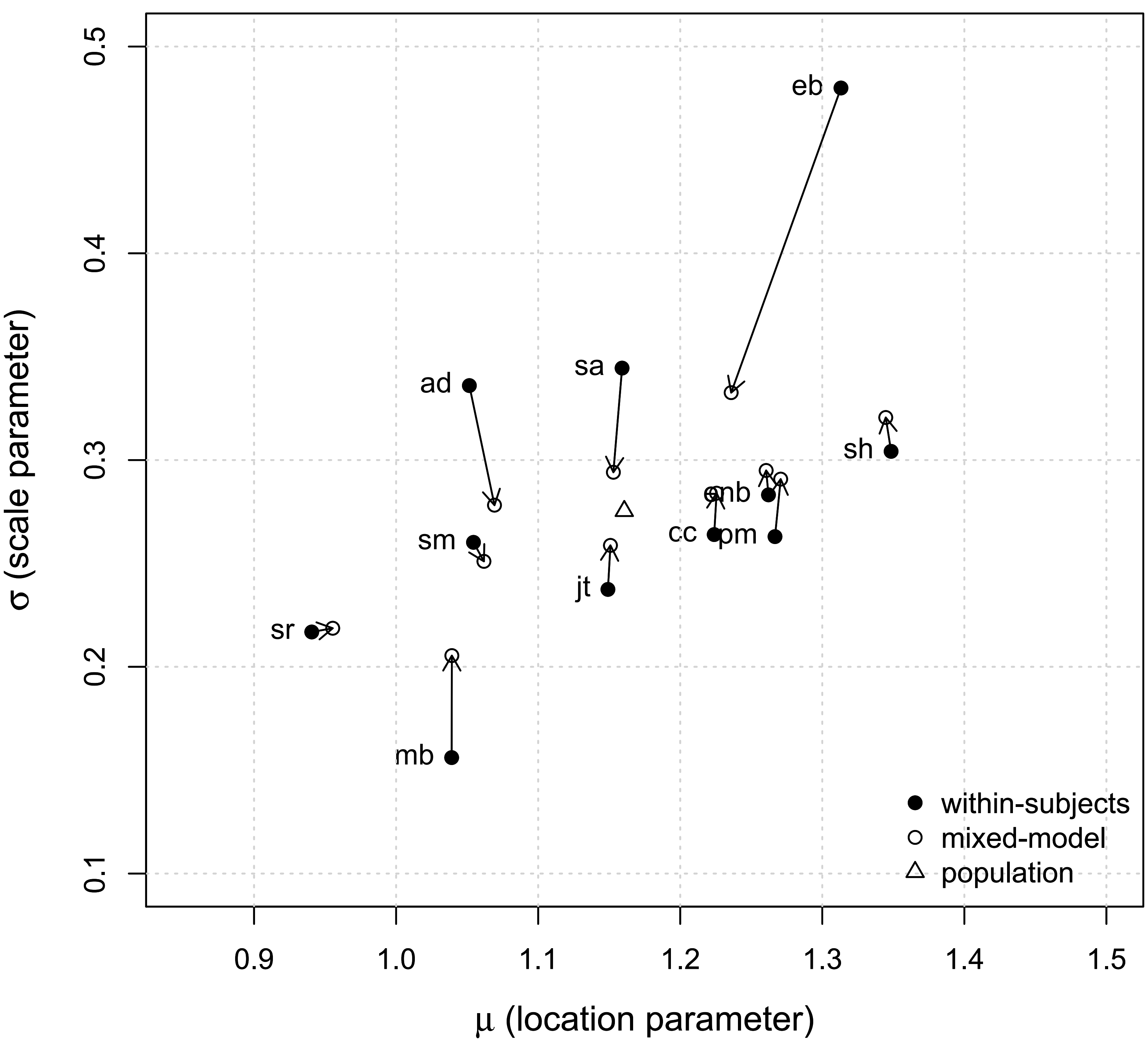
```
> fixef(bridge.m) # linear predictor parameters
(Intercept)       ratio
  -4.261918    3.716717

> PSE <- unname(-fixef(bridge.m)[1]/fixef(bridge.m)[2])
> PSE
[1] 1.146689
```

Is it significantly greater than 1?

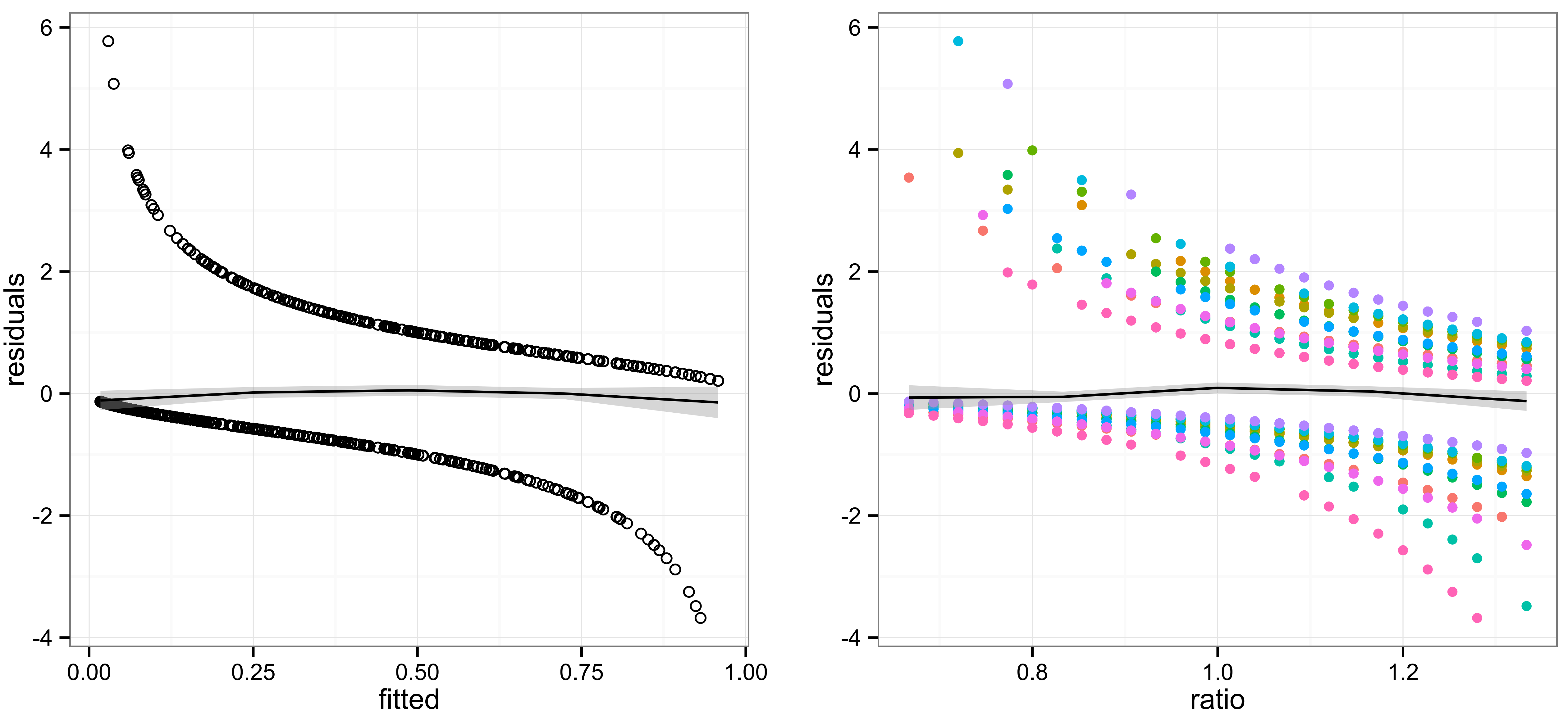
We can compute a bootstrapped 95% CI; it will include both the variability due to the binomial variable, as well as the variability between subjects.



Comparison with within-subjects estimates



Diagnostic plots



In the package `MPDiR` by Kenneth Knoblauch and Laurence T. Maloney (http://cran.r-project.org/web/packages/MPDiR/index.html) there are additional link functions, that allows for example to adjust the lower asymptote of the function in order to fit data from *n*AFC task, where the lower asymptote is at 1/*n*.

References & useful resources:

- Bates, D. M. ("in progress") ***lme4: Mixed-effects modeling with R***. freely available at http://lme4.r-forge.r-project.org

- Knoblauch, K., & Maloney, L. T. (2012). ***Modeling psychophysical data in R***

- Pinheiro, J. C., & Bates, D. M. (2000). ***Mixed-Effects Models in S and S-PLUS***.

- Moscatelli, A., Mezzetti, M., & Lacquaniti, F. (2012). ***Modeling psychophysical data at the population-level: The generalized linear mixed model.*** Journal of Vision, 12(11)(26)

- Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2010). ***Experimental Effects and Individual Differences in Linear Mixed Models: Estimating the Relationship between Spatial, Object, and Attraction Effects in Visual Attention***. Frontiers in Psychology, 1, 238.

- http://www.r-bloggers.com
  http://glmm.wikidot.com/faq
  http://stats.stackexchange.com
  http://stackoverflow.com/questions/tagged/r