

```

# Mixed Model
#
library(lattice)
library(lme4)
int = read.table(file.path(pfadu, "dbwort.df.txt"))
stimm2 = read.table(file.path(pfadu, "stimm2.df.txt"))
phrasen = read.table(file.path(pfadu, "phrasen.df.txt"))

# Inwiefern wird die Intensität (db) vom Vokal (V) beeinflusst?
# Vokal ist in diesem Fall der 'fixed Faktor' (= der zu prüfende Faktor).
bwplot(db ~ V, data = int)

# Dies soll berechnet werden, in dem die Variabilität
# wegen Versuchspersonen- und Wortunterschiede
# ausgeklammert wird:

# Variabilität im Bezug zum FF wegen Vpn:
bwplot(db ~ V|Vpn, data = int)

# Variabilität im Bezug zum FF wegen Wort:
bwplot(db ~ V|Wort, data = int)

# In der lmer() Funktion wird diese Variabilität
# durch (1+FF|RF) ausgeklammert: FF ist der Fixed Faktor,
# RF der Random Faktor, daher:
int.lmer = lmer(db ~ V + (1+V|Vpn) + (1+V|Wort), data = int)

# anova() gibt eine Ahnung, ob der FF signifikant sein wird
# (vor allem wenn der F-Ratio > 8 ist)
anova(int.lmer)

# Um zu testen, ob db signifikant von V beeinflusst wird,
# das Modell noch einmal laufen lassen, ohne FF (ohne V):
ohneff = update(int.lmer, ~ . -V)
# entspricht:
ohneff = lmer(db ~ (1+V|Vpn) + (1+V|Wort), data = int)

# Die Signifikanz wird durch ein Likelihood-Ratio-Test berechnet,
# indem die Modelle (a) ohne und (b) mit dem FF verglichen werden:
anova(ohneff, int.lmer)
# db wurde signifikant von V beeinflusst ( $\chi^2[1] = 10.7$ ,  $p < 0.01$ ).

#####
# Die Bedeutung von (1+FF|RF) und (1|RF)
#####

# (1+FF|RF) bedeutet: die Variabilität vom RF wird
# ausgeklammert, unter der Annahme, dass
# RF und FF interagieren.
#
# Interaktion zwischen RF und FF:
# der Einfluss von /a/ und /i/ auf db ist nicht derselbe
# für die verschiedenen Vpn.
# Interaktion zwischen RF (Vpn) und FF?
bwplot(db ~ V|Vpn, data = int)

# Interaktion zwischen RF (Wort) und FF?
bwplot(db ~ V|Wort, data = int)

# Man kann prüfen, ob RF und FF interagieren, wenn mit (1|RF)

```

```

# ein Modell ohne Interaktion neu berechnet wird.
# z.B. RF-Variabilität für Vpn mit FF ausklammern,
# ohne dass RF und FF interagieren:

int.lmer2 = lmer(db ~ V + (1|Vpn) + (1+V|Wort), data = int)
# oder
int.lmer2 = update(int.lmer, ~ . -(1+V|Vpn) + (1|Vpn))
# Und die Modelle mit und ohne Interaktion vergleichen:
anova(int.lmer, int.lmer2)

# Wie ist es mit Wort? Interaktion mit FF?
int.lmer3 = update(int.lmer, ~ . -(1+V|Wort) + (1|Wort))
anova(int.lmer, int.lmer3)

# Daher müsste der Unterschied zwischen 1|W und 1+W|V auf
# die Ergebnisse kaum einen Einfluss haben:
#
ohneff3 = update(int.lmer3, ~ . -V)
anova(int.lmer3, ohneff3)

# Vergleichen mit:
int.lmer = lmer(db ~ V + (1+V|Vpn) + (1+V|Wort), data = int)
ohneff = update(int.lmer, ~ . -V)
anova(int.lmer, ohneff)

#####
# Die Wahl zwischen (1+FF|RF) und (1|RF)
#####

# Im Prinzip soll immer (1+FF|RF) anstatt (1|RF) im
# Modell deklariert wird - vorausgesetzt, dass
# FF ein 'within' Faktor ist im Bezug
# zu RF, wie hier (beide Stufen von FF werden
# pro Vpn belegt):
with(int, table(Vpn, V))

# und hier (beide Stufen von FF werden
# pro Wort belegt):
with(int, table(Wort, V))

# (1+FF|RF) geht auch wenn es sich um einen 'within' Faktor
# handelt, und einige Werte fehlen, wie hier:
head(stimm2)
with(stimm2, table(Vpn, K))
stimm2.lmer = lmer(vot ~ K + (1+K|Vpn), data = stimm2)

# Wenn jedoch der FF 'between' ist im Bezug zum RF, dann
# funktioniert oft (1+FF|RF) nicht; oder das Ergebnis ist genau
# das gleiche wie (1|RF). z.B.
#
# Alter ist offensichtlich between
with(stimm2, table(Vpn, Alter))
# Daher wird dies wahrscheinlich nicht funktionieren:
alter.lmer = lmer(vot ~ Alter + (1+Alter|Vpn), data = stimm2)
# Warning message:
# In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
# Model failed to converge: degenerate Hessian with 1 negative eigenvalues
#
# Das kann nicht funktionieren, weil Alter und Vpn nicht
# interagieren können:

```

```
bwplot(vot ~ Alter|Vpn, data = stimm2)
```

```
# Daher:
```

```
alter.lmer = lmer(vot ~ Alter + (1|Vpn), data = stimm2)
```

```
#####
```

```
# Fragen
```

```
#####
```

```
##### Frage 1
```

```
# F1 wurde in drei verschiedenen Wörtern produziert von 3 Versuchspersonen
```

```
# in phrasenfinaler und phrasenmedialer Position erhoben.
```

```
# Inwiefern wird F1 (a) von der Position (b) vom Geschlecht beeinflusst.
```

```
bwplot(F1 ~ Pos, data = phrasen)
```

```
# (a)
```

```
# Wort und Versuchspersonen sind RF.
```

```
# Pos ist 'within' im Bezug zu Versuchsperson und Wort
```

```
with(phrasen, table(Vpn, Pos))
```

```
with(phrasen, table(W, Pos))
```

```
# Daher:
```

```
pos.lmer = lmer(F1 ~ Pos + (1+Pos|W) + (1+Pos|Vpn), data = phrasen)
```

```
# Modell ohne FF
```

```
pos.ohne = update(pos.lmer, ~ . -Pos)
```

```
# Statistischer Test
```

```
anova(pos.lmer, pos.ohne)
```

```
# Position hat keinen signifikanten Einfluss auf F1.
```

```
# (b)
```

```
bwplot(F1 ~ G, data = phrasen)
```

```
# G ist 'between' im Bezug zu Vpn. aber 'within' im Bezug zu Wort
```

```
with(phrasen, table(Vpn, G))
```

```
with(phrasen, table(W, G))
```

```
# daher
```

```
g.lmer = lmer(F1 ~ G + (1|Vpn) + (1+G|W) , data = phrasen)
```

```
# ohne FF
```

```
g.ohne = update(g.lmer, ~ . -G)
```

```
# Statistischer Test
```

```
anova(g.lmer, g.ohne)
```

```
# Geschlecht hat einen signifikanten Einfluss
```

```
# auf F1 ( $\chi^2[1] = 4.4, p < 0.05$ ).
```

```
#####
```

```
##### Frage 2
```

```
# Diese Daten:
```

```
preasp = read.table(file.path(pfadu, "preasp.txt"))
```

```
head(preasp)
```

```
bwplot(vdur ~ ptonic, data = preasp)
```

```
# within
```

```
with(preasp, table(sp, ptonic))
```

```
# between
```

```
with(preasp, table(word, ptonic))
```

```
ptonic.lmer = lmer(vdur ~ ptonic + (1|word) + (1+ptonic|sp), data = preasp)
```

```
ptonic.ohne = lmer(vdur ~ (1|word) + (1+ptonic|sp), data = preasp)
```

```
ptonic.ohne = update(ptonic.lmer, ~ . -ptonic)
```

```
anova(ptonic.lmer, ptonic.ohne)
```

```
# vdur wird signifikant von der Satzbetonung beeinflusst ( $X^2[1] = 18.6, < 0.001$ )
```

```

bwplot(vdur ~ Pre, data = preasp)
with(preasp, table(sp, Pre))
with(preasp, table(word, Pre))
pre.lmer = lmer(vdur ~ Pre + (1+Pre|word) + (1+Pre|spk), data = preasp)
pre.ohne = update(pre.lmer, ~ . -Pre)
anova(pre.lmer, pre.ohne)
bwplot(vdur ~ Pre | spk, data = preasp)
bwplot(vdur ~ Pre | word, data = preasp)
# zeigen Dauerwerte (vdur) von Vokalen (vdur), die in verschiedenen
# Wörtern (word) und von verschiedenen Versuchspersonen (spk) produziert
# wurden.
# Inwiefern wird die Vokaldauer
# (a) von der Satzbetonung (ptonic: Y = betont, N = unbetont)
# (b) von der Präaspiration (Faktor Pre) beeinflusst?

##### Frage 3
# Diese Daten:
asp = read.table(file.path(pfadu, "asp.txt"))
# zeigen Dauerwerte von der Aspiration, die in verschiedenen
# Wörtern und von verschiedenen Versuchspersonen produziert wurden.
# Inwiefern wird die Aspirationsdauer von der Artikulationstelle beeinflusst?
bwplot(d ~ Kons, data = asp)
kons.lmer = lmer(d ~ Kons + (1 + Kons | Vpn) + (1|Wort), data = asp)
ohne = update(kons.lmer, ~ . -Kons)
ohne2 = lmer(d ~ (1 + Kons | Vpn) + (1|Wort), data = asp)
anova(kons.lmer, ohne)

##### Frage 4
# In diesen Daten:
laerm = read.table(file.path(pfadu, "laerm.txt"))
# wurden Reaktionszeiten von verschiedenen Versuchspersonen
# auf verschiedene Wörter gemessen. Inwiefern wird
# die Reaktionszeit (a) vom Lärm (Faktor Laerm) und (b)
# von der Altersgruppe der Versuchsperson beeinflusst?
laerm.lmer = lmer(rt ~ Laerm + (1 + Laerm|Vpn) + (1+Laerm|Wort), data = laerm)
anova(laerm.lmer)
ohne = update(laerm.lmer, ~ . -Laerm)
anova(laerm.lmer, ohne)
# Der Einfluss von Lärm auf die Reaktionszeit war nicht ganz signifikant (p =
# 0.05074).
bwplot(rt ~ Alter, data = laerm)
alter.lmer = lmer(rt ~ Alter + (1|Vpn) + (1 + Alter|Wort), data = laerm)
anova(alter.lmer)
# Alter hatte keinen signifikanten Einfluss auf die Reaktionszeit.

##### Frage 5
# In diesen Daten:
stefan = read.table(file.path(pfadu, "stefan.txt"))
# wurde F2 (f2mid) in verschiedenen Wörtern (Word) produziert
# von verschiedenen Versuchspersonen (Vpn) gemessen. Inwiefern wird
# F2 (a) vom Tempo (b) vom Vokal (Faktor (V)) beeinflusst?
tempo.lmer = lmer(f2mid ~ Tempo + (1 + Tempo|Vpn) + (1+Tempo|Word), data =
# stefan)
anova(tempo.lmer)
with(stefan, table(Vpn, V))
with(stefan, table(Word, V))
bwplot(f2mid ~ V, data = stefan)
v.lmer = lmer(f2mid ~ V + (1 + V | Vpn) + (1|Word), data = stefan)

```

```

anova(v.lmer)
ohne = update(v.lmer, ~ . -V)
anova(v.lmer, ohne)
# Vokal hatten einen signifikanten Einfluss auf F2 (X^2[1] = 12.7, p < 0.001).
v.lmer2 = lmer(f2mid ~ V + (1 | Vpn) + (1|Word), data = stefan)
v.lmer2 = update(v.lmer, ~ . -(1 + V | Vpn) + (1|Vpn))
anova(v.lmer2, v.lmer)

```

```

v.lmer3 = update(v.lmer, ~ . -(1|Word))
anova(v.lmer, v.lmer3)

```

```
##### Frage 6
```

```

sc = read.table(file.path(pfadu, "schule.txt"))
# F1 von Kindern aus 107 Schulen wurde gemessen. Die Eltern
# der Kinder stammten aus zwei Regionen (Region, J/N).
# Inwiefern wird F1 von der Region beeinflusst?
reg.lmer = lmer(F1 ~ Region + (1+Region|Schule) + (1|Gender), data = sc)
ohne = update(reg.lmer, ~ . -Region)
anova(reg.lmer, ohne)
# F1 wird von der Region beeinflusst (X^2[1] = 22.4, p < 0.001).

```

```
##### Frage 7
```

```

# In diesen Daten
lex = read.table(file.path(pfadu, "lex.txt"))
# wurden Reaktionszeiten von verschiedenen Sprechern (Subj) auf verschiedene
# Wörter (Word) gemessen. Die Sprecher waren entweder Muttersprache
# englisch oder nicht (Faktor Lang). Zusätzlich gibt es eine
# Verschlüsselung für die Wortlänge (Length): je kleiner die Zahl, umso
# kürzer das Wort. Inwiefern werden die Reaktionszeiten von (a) der
# Sprache und (b) der Wortlänge beeinflusst?
with(lex, table(Lang, Subj))
with(lex, table(Lang, Word))
bwplot(rt ~ Lang, data = lex)
bwplot(rt ~ Lang|Word, data = lex)
sprache.lmer = lmer(rt ~ Lang + (1|Subj) + (1+Lang|Word), data = lex)
sprache.ohne = update(sprache.lmer, ~ . -Lang)
anova(sprache.lmer, sprache.ohne)
# Sprache hat einen signifikanten Einfluss auf die Reaktionszeiten X^2[1] =
# 6.0, p < 0.05.
xyplot(rt ~ Length, data = lex, type = c("p", "smooth"))
xyplot(rt ~ Length|Subj, data = lex, type = c("p", "smooth"))
len.lex = lmer(rt ~ Length + (1|Word) + (1+Length|Subj), data = lex)
len.ohne = update(len.lex, ~ . -Length)
anova(len.lex, len.ohne)

```

```
##### Frage 8
```

```

# In diesen Daten
read.table(file.path(pfadu, "x24.txt"))
# wurden für 3 Vokale (V) Entfernung zum Vokalmittelpunkt (ent) für
# verschiedene Sprecher (Vpn), die verschiedene Wörter (Word) produziert
# haben. Inwiefern werden die Entfernungen vom Vokal beeinflusst?

```

```
##### Frage 9
```

```

# Diese Studie von Bodo Winter
pl = read.csv("http://www.bodowinter.com/tutorial/politeness_data.csv")

```

```
# siehe auch: http://arxiv.org/pdf/1308.5499.pdf  
# befasst sich mit der Beziehung zwischen  $f_0$  (frequency) und Höflichkeit  
  (attitude). Höflichkeit hat zwei Stufen: pol (höflich) und inf (normal).  
  Dir  $f_0$ -Daten sind von verschiedenen Versuchspersonen (subject) und in  
  verschiedenen Sprechsituationen (scenario) erhoben worden. Inwiefern  
  wird  $f_0$  von der Höflichkeit beeinflusst.
```