**Homework 4: Mixed-Models course 2016**

**The homework consists of computing *p* values for a mixed effects models that you ran earlier and practice your write-up skills for how you computed these *p* values and the results you got; plus some reading/working through a tutorial (part 2 from what you did in week 1).**

**What you have to hand in:**

Hand in your R code and the output and figures you created with it: Copy/paste the code and the output/figures into a word document, print it, and bring it to the lab session on March 7 (=deadline!).

**(1) Compute *p* values for the cherry pit spitting model (using the *N*=30 data frame) you created in the last homework, which had trial, gender, and the interaction between trial and gender in the model (use the centered trial predictor throughout).**

**In principle, you want to get model-comparison based *p* values to investigate/answer the following research question (but see detailed instructions below for slight variations and how to get the *p* values):**

* Is there a significant trial effect?
* Is there a significant gender effect?
* Is there a significant interaction between gender and the trial effect?

**(a) Compute Type 2 tests to obtain *p* values for the interaction effect and the two main effects via LRTs using the anova command.**

Also include a write-up: Follow my example on the slides to report how you computed these *p* values and then report the results (besides the test statistic and *p* value, also include the regression coefficients and its standard error; feel free to include or not include the CIs).

**To determine p values, we computed Likelihood Ratio Tests using R's anova function to compare the model with the interaction effect between gender and trials to the model without the interaction effect between gender and trials (with models being identical otherwise). Females become better at spitting cherries as the trials advance and males get worse (Estimate = 1.12(.393), χ2 (1) = 7.64, p = .005). Moreover, we tested the main effects with restricted models and type 2 errors. We found that the trial (slope random effect) was not significant (Estimate = -0.07 (.43), χ2 (1) = .033, p = .85) and there was a significant difference in gender , where females spit further distance than men (Estimate = 4.03 (.99), χ2 (1) = 11.26, p < .001).**

Note: Compute the main effects even when the interaction effect is significant (which means that you have to suppress your inner Type 2 person). This means, you will do several model comparisons, not just one.

**(b) Do the same using the KRmodcomp function.**

Yes, also including the (adjusted) write-up.

**To determine p values, we used the KRmodcomp function of the package pbkrtest (Halekoh & Højsgaard, 2014), which performs conditional F tests with degrees of freedom based on Kenward-Roger approximation. Females become better at spitting cherries as the trials advance and males get worse (Estimate = 1.12(.393), F(1, 28) = 8.12, p = .008). Moreover, we tested the main effects with restricted models and type 2 errors. We found that the trial (slope random effect) was not significant (Estimate = -0.07 (.43), F(1, 29) = .032, p = .85) and there was a significant difference in gender , where females spit further distance than men (Estimate = 4.03 (.99), F(1, 28) = 15.302, p < .001).**

**(c) Do the same using the PBmodcomp function.**

Yes, you guessed it, also including the adjusted write-up.

Also: Make use of multiple cores!

**To determine p values, we used the PBmodcomp function of the package pbkrtest (Halekoh & Højsgaard, 2014), which performs model comparison of nested models using parametric bootstrap methods (i.e., bootstrapped Likelihood Ratio Tests).** **Females become better at spitting cherries as the trials advance and males get worse (Estimate = 1.12(.393), PBTest = 7.61, p = .01, Estimate = 1.12(.393), Gamma = 7.61, p = .01). Moreover, we tested the main effects with restricted models and type 2 errors. We found that the trial (slope random effect) was not significant (Estimate = -0.07 (.43), PBTest = .033, p = .85, Estimate = -0.07 (.43), Bartlett = .031, p = .85) and there was a significant difference in gender , where females spit further distance than men (Estimate = 4.03 (.99), PBTest= 11.19, p = .001, Estimate = 4.03 (.99), LRT= 11.19, p < .001).**

**(d) Try to use anova to do Type 3 tests of the two main effects and the interaction.**

Which step(s) go(es) wrong? I.e., which model(s) doesn't/don't do what you want them to do. Copy/paste the relevant summary output into your answer.

No need to include the usual write-up part that you had to do for (a) to (c).

**The “bug/feature” is present, this happens because the trial number categorical variable has more than two levels.**

Data: cherry

Models:

cherrysmallx11: distance ~ gender + numtrialc:gender + (1 + numtrialc | pid)

cherrylargex3: distance ~ numtrialc \* gender + (1 + numtrialc | pid)

Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)

cherrysmallx11 8 947.02 971.1 -465.51 931.02

cherrylargex3 8 947.02 971.1 -465.51 931.02 0 0 < 2.2e-16 \*\*\*

**(e) Compute Type 3 tests to obtain *p* values for the interaction effect and the two main effects via KR F tests using the Anova command.**

Here, include again the write-up part.

**To determine p values, we computed Type 3 conditional F tests with Kenward-Roger approximation for degrees of freedom as implemented in the Anova function of the package car (Fox & Weisberg, 2011), which in turn calls the function KRmodcomp of the package pbkrtest (Halekoh & Højsgaard, 2014).** **Females become better at spitting cherries as the trials advance and males get worse (Estimate = 1.12(.393), F(1, 28) = 8.12, p = .008). Moreover, we tested the main effects with restricted models and type 3 errors. We found that the trial (slope random effect) was not significant (Estimate = -0.07 (.39), F(1, 29) = .04, p = .84) and there was a significant difference in gender , where females spit further distance than men (Estimate = 3.62 (1), F(1, 28) = 12.94, p = .001).**

**(f) Compute Type 3 tests to obtain *p* values for the interaction effect and the two main effects via LRTs using the mixed command.**

Make use of multiple cores.

Include the write-up part.

**To determine p values, we computed Type 3 Likelihood Ratio Tests as implemented in the mixed function of the package afex (Singmann, Bolker, & Westfall, 2015). Females become better at spitting cherries as the trials advance and males get worse (Estimate = 1.12(.393), χ2 (7) = 11.4, p = .005). Moreover, we tested the main effects with restricted models and type 3 errors. We found that the trial (slope random effect) was not significant (Estimate = -0.07 (.39), χ2 (7) = .04, p = .83) and there was a significant difference in gender , where females spit further distance than men (Estimate = 3.62 (1), χ2 (7) = 7.64, p < .001).**

**(g) Compute Type 3 tests to obtain *p* values for the interaction effect and the two main effects via KR F tests using the mixed command.**

Make use of multiple cores.

Include the write-up part.

**To determine p values, we computed Type 3 conditional F tests with Kenward-Roger approximation for degrees of freedom as implemented in the mixed function of the package afex (Singmann, Bolker, & Westfall, 2015), which in turn calls the function KRmodcomp of the package pbkrtest (Halekoh & Højsgaard, 2014). We found that females become better at spitting cherries as the trials advance and males get worse (Estimate = 1.12(.393), F(1, 28) = 8.12, p = .008). Moreover, we tested the main effects with restricted models and type 3 errors. We found that the trial (slope random effect) was not significant (Estimate = -0.07 (.39), F(1, 28) = .04, p = .84) and there was a significant difference in gender , where females spit further distance than men (Estimate = 3.62 (1), F(1, 28) = 12.94, p = .001).**

**(h) Compute Type 3 tests to obtain *p* values for the interaction effect and the two main effects via PB tests using the mixed command.**

Make use of multiple cores for both the model-fitting and the bootstrapping part.

Include the write-up part.

**To determine p values, we computed Type 3 bootstrapped Likelihood Ratio Tests (using 1000 simulations) as implemented in the mixed function of the package afex (Singmann, Bolker, & Westfall, 2015), which in turn calls the function PBmodcomp of the package pbkrtest (Halekoh & Højsgaard, 2014). Females become better at spitting cherries as the trials advance and males get worse (Estimate = 1.12(.393), PBTest = 7.61, p = .008). Moreover, we tested the main effects with restricted models and type 2 errors. We found that the trial (slope random effect) was not significant (Estimate = -0.07 (.43), PBTest = .042, p = .83) and there was a significant difference in gender , where females spit further distance than men (Estimate = 4.03 (.99), PBTest= 11.4, p < .001)**

**(2) Read and work through the lme4 tutorial**

It is available here: http://www.bodowinter.com/tutorial/bw\_LME\_tutorial2.pdf

Full reference: Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. arXiv:1308.5499. [http://arxiv.org/pdf/1308.5499.pdf]

It's partly recap from what we discussed in the classes so far, and it's partly preparation for what we are going to discuss in the next classes (crossed random effects, i.e., random effects for participant and item).

In the tutorial, there are several things that the author tells you to compute yourself (create a data frame, run models, etc): Do all these things yourself in R; and as proof that you did it, you hand in this R script (here, you don't have to include the actual figures, but please do create them and make sure to include the code also for the figures in what you hand in).

**IN SCRIPT BELOW**

**install.packages('lme4')**

**library (lme4)**

**install.packages('car')**

**library (car)**

**install.packages('psych')**

**library (psych)**

**install.packages('afex')**

**library (afex)**

**install.packages('pbkrtest')**

**library (pbkrtest)**

**install.packages('"parallel')**

**library (parallel)**

**cherry <- read.csv("C:\\Users\\s4600479\\Desktop\\CherryPit\_large\_N30Gender.csv", sep = ",")**

**#contrast to zero sum, because...**

**contrasts(cherry$gender) = contr.sum(2)**

**#A)for this hw we need to scale trials**

**cherry$numtrialc <- scale(cherry$num\_trial, center = T, scale = F)**

**cherryg <- lmer(distance ~ numtrialc\*gender + (1 + numtrialc| pid), data = cherry)**

**summary(cherrylargex1)**

**summary(cherrylargex2)**

**summary(cherryg)**

**#Compute Type 2 tests to obtain p values for the interaction effect and the two main effects via LRTs using the anova command.**

**cherrylarge <- lmer(distance ~ numtrialc\*gender + (1 + numtrialc| pid), data = cherry)**

**cherrysmall <- lmer(distance ~ numtrialc + gender + (1 + numtrialc| pid), data = cherry)**

**#for main effects in type 2 errors**

**cherrylargex1 <- lmer(distance ~ numtrialc + gender + (1 + numtrialc| pid), data = cherry)**

**cherrysmallx1 <- lmer(distance ~ gender + (1 + numtrialc| pid), data = cherry)**

**#for main effects in type 2 errors**

**cherrylargex2 <- lmer(distance ~ numtrialc + gender + (1 + numtrialc| pid), data = cherry)**

**cherrysmallx2 <- lmer(distance ~ numtrialc + (1 + numtrialc| pid), data = cherry)**

**anova(cherrysmall, cherrylarge)**

**anova(cherrysmallx1, cherrylargex1)**

**anova(cherrysmallx2, cherrylargex2)**

**#Do the same using the KRmodcomp function.**

**cherryftest <- pbkrtest::KRmodcomp(cherrylarge, cherrysmall)**

**cherryftest**

**#main**

**cherryftestx1 <- pbkrtest::KRmodcomp(cherrylargex1, cherrysmallx1)**

**cherryftestx2 <- pbkrtest::KRmodcomp(cherrylargex2, cherrysmallx2)**

**cherryftestx1**

**cherryftestx2**

**#for PB test use multiple cores**

**#Multiple Cores with mixed()**

**# Create the cluster**

**n\_cores <- detectCores()**

**MyCluster <- makeCluster(rep("localhost", n\_cores - 1))**

**#Run your mixed() command (LRT, KR, PB)**

**cherrygcl <- mixed(distance ~ gender\*numtrialc + (1 + numtrialc | pid), type = 2, method = "LRT", data = cherry, cl = MyCluster)**

**#Each CPU runs a (sub)model! Can save a lot of time.**

**#Once you're done, stop the cluster**

**stopCluster(MyCluster)**

**#PBTEST**

**cherrypb <- pbkrtest::PBmodcomp(cherrylarge, cherrysmall)**

**summary(cherrypb)**

**#main**

**cherrypbx1 <- pbkrtest::PBmodcomp(cherrylargex1, cherrysmallx1)**

**summary(cherrypbx1)**

**cherrypbx2 <- pbkrtest::PBmodcomp(cherrylargex2, cherrysmallx2)**

**summary(cherrypbx2)**

**#Try to use anova to do Type 3 tests of the two main effects and the interaction**

**#for main effects in type 2 errors**

**#Compute Type 2 tests to obtain p values for the interaction effect and the two main effects via LRTs using the anova command.**

**cherrylargex3 <- lmer(distance ~ numtrialc\*gender + (1 + numtrialc| pid), data = cherry)**

**cherrysmallx3 <- lmer(distance ~ numtrialc + gender + (1 + numtrialc| pid), data = cherry)**

**#for main effects in type 2 errors**

**cherrysmallx11 <- lmer(distance ~ gender + numtrialc:gender + (1 + numtrialc| pid), data = cherry)**

**summary(cherrysmallx1)**

**cherrysmallx21 <- lmer(distance ~ numtrialc + numtrialc:gender + (1 + numtrialc| pid), data = cherry)**

**anova(cherrysmallx3, cherrylargex3)**

**anova(cherrysmallx11, cherrylargex3)**

**anova(cherrysmallx21, cherrylargex3)**

**#K-R F with type 3 and Anova fn**

**car::Anova(cherrylarge, type = 3, test = "F")**

**#another way**

**cherrykr <- mixed(distance ~ gender\*numtrialc + (1 + numtrialc |pid), data = cherry, type = 3, method = "KR")**

**# for output**

**anova(cherrykr)**

**#Type 3 tests to obtain p values for the interaction effect and the two main effects via LRTs**

**cherrylrt <- mixed(distance ~ gender\*numtrialc + (1 + numtrialc |pid), data = cherry, type = 3, method = "LRT")**

**# for output**

**anova(cherrylrt)**

**# Type 3 tests to obtain p values for the interaction effect and the two main effects via KR F tests**

**#another way**

**cherrykr <- mixed(distance ~ gender\*numtrialc + (1 + numtrialc |pid), data = cherry, type = 3, method = "KR")**

**# for output**

**anova(cherrykr)**

**#Type 3 tests to obtain p values for the interaction effect and the two main effects via PB tests**

**cherrypbt <- mixed(distance ~ gender\*numtrialc + (1 + numtrialc |pid), data = cherry, type = 3, method = "PB")**

**# for output**

**anova(cherrypbt)**

**############2####################**

**politenessdata<- read.csv("D:/mijn documenten/Nijmegen/mixedeffects/politeness\_data.csv", sep = ",")**

**view(politenessdata)**

**summary(politenessdata)**

**which(is.na(politenessdata$frequency))**

**boxplot(frequency ~ attitude\*gender, col=c("white","lightgray"),politenessdata)**

**lmer(frequency ~ attitude, data=politenessdata)**

**politeness.model = lmer(frequency ~ attitude + (1|subject) + (1|scenario), data=politenessdata)**

**politeness.model = lmer(frequency ~ attitude + gender + (1|subject) + (1|scenario), data=politenessdata)**

**summary(politeness.model)**

**politeness.null = lmer(frequency ~ gender + (1|subject) + (1|scenario), data=politenessdata, REML=FALSE)**

**politeness.model = lmer(frequency ~ attitude +**

**gender + (1|subject) + (1|scenario),**

**data=politenessdata, REML=FALSE)**

**anova(politeness.null,politeness.model)**

**politeness.int = lmer(frequency ~ attitude \***

**gender + (1|subject) + (1|scenario),**

**data=politenessdata, REML=FALSE)**

**anova(politeness.model, politeness.int)**

**coef(politeness.model)**

**politeness.model = lmer(frequency ~ attitude +**

**gender + (1+attitude|subject) +**

**(1+attitude|scenario),**

**data=politenessdata,**

**REML=FALSE)**

**coef(politeness.model)**

**politeness.null = lmer(frequency ~ gender +**

**(1+attitude|subject) + (1+attitude|scenario),**

**data=politenessdata, REML=FALSE)**

**anova(politeness.null,politeness.model)**

**all.res=numeric(nrow(politenessdata))**

**for(i in 1:nrow(politnessdata)){politeness.int = lmer(frequency ~ attitude \* gender + (1|subject) + (1|scenario), data=politenessdata, REML=FALSE), POP[-i,])**

**all.res[i]=fixef(politeness.int)[3]}**