763 HW5

Ning Jiang

14 December 2015

```r
library(lme4)
library(RLRsim)
library(arm)
library(LMERConvenienceFunctions)
music = read.csv('ratings.csv')
```

1. (a).

```r
lm1a = lm(Classical ~ Instrument + Harmony + Voice, data = music)
summary(lm1a)
## Call:
## lm(formula = Classical ~ Instrument + Harmony + Voice, data = music)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8718 -1.7137 -0.0297  1.7576 11.4766
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         4.34016    0.12987  33.420  < 2e-16 ***
## Instrumentpiano     1.37359    0.11298  12.158  < 2e-16 ***
## Instrumentstring    3.13312    0.11230  27.899  < 2e-16 ***
## HarmonyI-V-IV      -0.03108    0.13008  -0.239 0.811168
## HarmonyI-V-VI       0.76909    0.13008   5.913 3.83e-09 ***
## HarmonyIV-I-V       0.05007    0.12997   0.385 0.700092
## Voicepar3rd        -0.41247    0.11271  -3.660 0.000258 ***
## Voicepar5th        -0.37058    0.11264  -3.290 0.001016 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.297 on 2485 degrees of freedom
##   (27 observations deleted due to missingness)
## Multiple R-squared:  0.255,  Adjusted R-squared:  0.2529
## F-statistic: 121.5 on 7 and 2485 DF,  p-value: < 2.2e-16

lm1.1 = lm(Classical ~ Harmony + Voice, data = music)
lm1.2 = lm(Classical ~ Instrument + Voice, data = music)
lm1.3 = lm(Classical ~ Instrument + Harmony, data = music)
anova(lm1a, lm1.1)

## Analysis of Variance Table
##
## Model 1: Classical ~ Instrument + Harmony + Voice
```

```
## Model 2: Classical ~ Harmony + Voice
##   Res.Df   RSS Df Sum of Sq      F    Pr(>F)    
## 1   2485 13108                                  
## 2   2487 17235 -2   -4127.6 391.26 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(lm1a, lm1.2)
```

```
## Analysis of Variance Table
## 
## Model 1: Classical ~ Instrument + Harmony + Voice
## Model 2: Classical ~ Instrument + Voice
##   Res.Df   RSS Df Sum of Sq      F    Pr(>F)    
## 1   2485 13108                                  
## 2   2488 13381 -3   -273.65 17.293 4.107e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(lm1a, lm1.3)
```

```
## Analysis of Variance Table
## 
## Model 1: Classical ~ Instrument + Harmony + Voice
## Model 2: Classical ~ Instrument + Harmony
##   Res.Df   RSS Df Sum of Sq      F    Pr(>F)    
## 1   2485 13108                                  
## 2   2487 13193 -2    -85.64 8.1181 0.0003061 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Individual test shows that only two indicators from the harmony variable are not statistically significant. Other variables are all highly statistically significant. Anova between lm1.1 and lm1 tests the importance of instruments on classical ratings. Overall, instrument significantly influences rating. The partial F-test is statistically significant, showing that at least one of the instrument predictors have coefficient that is significantly different from zero. Similarly, partial F-test between lm1 and lm1.2 tests the importance of harmony on classical ratings, and F-test between lm1 and lm1.3 tests the importance of voice on classical ratings. Both tests are statistically significant as well. So both harmony and voice are significant on ratings as well. But with the same degrees of freedom, test statistics for the voice is smaller than that of instrument. So instrument is more important in this sense, on classical ratings. Also, the coefficient in the regression out put shows that ratings change the greatest when we switch instruments, which corresponds to the first hypothesis of the researchers. I took a look at the diagnostic plots on the linear model, nothing interesting though.

(b).(i).

$$\text{for } i = 1,2,\ldots,2493, \text{and } j = 1,2,\ldots,70$$

$$y_i = \beta_{j[i]} + \beta_1 \text{Instrument}_i + \beta_2 \text{Harmony}_i + \beta_3 \text{Voice}_i + \epsilon_i$$

$$\beta_j = \alpha_0 + \eta_j$$

$$\epsilon_i \sim N(0, \sigma_\epsilon^2)$$

$$\eta_j \sim N(0, \tau_\eta^2)$$

(ii).

```
lmer1b2 = lmer(Classical ~ 1 + Instrument + Harmony + Voice + (1|Subjec
t), data = music, REML = F)
#first method
BIC(lmer1b2)

## [1] 10527.07

BIC(lm1a)

## [1] 11282.84

#second method
exactRLRT(lmer1b2)

## Using restricted likelihood evaluated at ML estimators.
## Refit with method="REML" for exact results.
##  simulated finite sample distribution of RLRT.
##  (p-value based on 10000 simulated values)
## data:
## RLRT = 763.37, p-value < 2.2e-16
```

So both methods show that we need a random intercept. The first method I used to see if the random intercept is needed is to compare the BIC score of the lmer and the linear model. By adding a random intercept, the BIC score dropped for about 700, showing that the model improved significantly.

The second method I used is to test whether the variance of the random effect is significantly different from zero using the RLRsim package. The test has a p-value is approximately zero, thus we reject the null that the variance is zero. This shows that there is significant amount of variation in the random effect, thus including a random intercept is helpful for capturing and explaining the extra amount of variation, thus is necessary for the case here.

(iii)

```
summary(lmer1b2)$coefficients

##                     Estimate Std. Error    t value
## (Intercept)       4.34374186 0.18808548 23.0945095
## Instrumentpiano   1.37704509 0.09304624 14.7995782
## Instrumentstring  3.13160679 0.09243416 33.8793218
## HarmonyI-V-IV    -0.03250823 0.10702814 -0.3037354
```

```
## HarmonyI-V-VI     0.77095828 0.10702477  7.2035500
## HarmonyIV-I-V     0.04989468 0.10693624  0.4665835
## Voicepar3rd      -0.41506472 0.09273406 -4.4758606
## Voicepar5th      -0.37438261 0.09267556 -4.0397125

summary(lm1a)$coefficients

##                     Estimate Std. Error    t value      Pr(>|t|)
## (Intercept)       4.34016117  0.1298670 33.4200529 1.447374e-202
## Instrumentpiano   1.37358790  0.1129806 12.1577271  4.450685e-33
## Instrumentstring  3.13312057  0.1123007 27.8993862 2.996472e-149
## HarmonyI-V-IV    -0.03108093  0.1300757 -0.2389450  8.111679e-01
## HarmonyI-V-VI     0.76908588  0.1300762  5.9125806  3.830504e-09
## HarmonyIV-I-V     0.05007038  0.1299715  0.3852413  7.000916e-01
## Voicepar3rd      -0.41246663  0.1127056 -3.6596803  2.577881e-04
## Voicepar5th      -0.37057793  0.1126373 -3.2900106  1.015828e-03

lmer1b3.1 = lmer(Classical ~ 1 + Harmony + Voice + (1|Subject), data =
music, REML = F)
lmer1b3.2 = lmer(Classical ~ 1 + Instrument + Voice + (1|Subject), data
 = music, REML = F)
lmer1b3.3 = lmer(Classical ~ 1 + Harmony + Instrument + (1|Subject), da
ta = music, REML = F)
anova(lmer1b2,lmer1b3.1)

## Data: music
## Models:
## lmer1b3.1: Classical ~ 1 + Harmony + Voice + (1 | Subject)
## lmer1b2: Classical ~ 1 + Instrument + Harmony + Voice + (1 | Subject)
##             Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer1b3.1  8 11408 11455 -5696.2    11392
## lmer1b2   10 10469 10527 -5224.4    10449 943.59      2  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

anova(lmer1b2,lmer1b3.2)

## Data: music
## Models:
## lmer1b3.2: Classical ~ 1 + Instrument + Voice + (1 | Subject)
## lmer1b2: Classical ~ 1 + Instrument + Harmony + Voice + (1 | Subject)
##             Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer1b3.2  7 10539 10580 -5262.4    10525
## lmer1b2   10 10469 10527 -5224.4    10449 75.931      3  2.288e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

anova(lmer1b2,lmer1b3.3)

## Data: music
## Models:
## lmer1b3.3: Classical ~ 1 + Harmony + Instrument + (1 | Subject)
```

```
## lmer1b2: Classical ~ 1 + Instrument + Harmony + Voice + (1 | Subject)
##           Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)

## lmer1b3.3  8 10489 10536 -5236.6    10473

## lmer1b2   10 10469 10527 -5224.4    10449 24.24      2   5.45e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Coefficients from the fixed effect of the full lmer model are basically the same when compared to the linear model, including random effect did not change the fixed effect much. Then, similar to part a, I fitted three separate lmer models, with the same random effect for intercept but the main effects are dropped one at a time, and did anova analysis between all these three models and the full model. All three anova analysis have p-values that are approximately zero. So the results are the same as in a, that all the main effect indicators have significant effect on the classical ratings. Results using RMEL and ML analysis are basically the same.

(c).(i).

```
lmer1c1 = lmer(Classical ~ 1 + Instrument + Harmony + Voice + (1|Subjec
t:Instrument) + (1|Subject:Harmony) + (1|Subject:Voice), data = music)
AIC(lmer1c1)

## [1] 10075.51

AIC(lmer1b2)

## [1] 10468.86

AIC(lm1a)

## [1] 11230.45
```

I compared the AIC score of the three random intercept effects model to the models in 1a and 1b. The AIC score is clearly the lowest for our model here in this part, we noticed significant drops in the AIC scores. So this model is the best compared with previous models.

(ii).

```
lmer1c1.1 = lmer(Classical ~ 1 + Harmony + Voice + (1|Subject:Instrumen
t) + (1|Subject:Harmony) + (1|Subject:Voice), data = music)
lmer1c1.2 = lmer(Classical ~ 1 + Instrument + Voice + (1|Subject:Instru
ment) + (1|Subject:Harmony) + (1|Subject:Voice), data = music)
lmer1c1.3 = lmer(Classical ~ 1 + Instrument + Harmony + (1|Subject:Inst
rument) + (1|Subject:Harmony) + (1|Subject:Voice), data = music)
anova(lmer1c1, lmer1c1.1)

## refitting model(s) with ML (instead of REML)
```

```
## Data: music
## Models:
## lmer1c1.1: Classical ~ 1 + Harmony + Voice + (1 | Subject:Instrument)
 +
## lmer1c1.1:     (1 | Subject:Harmony) + (1 | Subject:Voice)
## lmer1c1: Classical ~ 1 + Instrument + Harmony + Voice + (1 | Subject:
Instrument) +
## lmer1c1:     (1 | Subject:Harmony) + (1 | Subject:Voice)
##           Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer1c1.1 10 10160 10219 -5070.2    10140
## lmer1c1   12 10058 10127 -5016.8    10034 106.89      2  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(lmer1c1, lmer1c1.2)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: music
## Models:
## lmer1c1.2: Classical ~ 1 + Instrument + Voice + (1 | Subject:Instrum
ent) +
## lmer1c1.2:     (1 | Subject:Harmony) + (1 | Subject:Voice)
## lmer1c1: Classical ~ 1 + Instrument + Harmony + Voice + (1 | Subject:
Instrument) +
## lmer1c1:     (1 | Subject:Harmony) + (1 | Subject:Voice)
##           Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer1c1.2  9 10090 10143 -5036.3    10072
## lmer1c1   12 10058 10127 -5016.8    10034 39.013      3  1.724e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(lmer1c1, lmer1c1.3)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: music
## Models:
## lmer1c1.3: Classical ~ 1 + Instrument + Harmony + (1 | Subject:Instr
ument) +
## lmer1c1.3:     (1 | Subject:Harmony) + (1 | Subject:Voice)
## lmer1c1: Classical ~ 1 + Instrument + Harmony + Voice + (1 | Subject:
Instrument) +
## lmer1c1:     (1 | Subject:Harmony) + (1 | Subject:Voice)
##           Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer1c1.3 10 10081 10140 -5030.6    10061
## lmer1c1   12 10058 10127 -5016.8    10034 27.753      2  9.409e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As before, I still used the anova analysis, and fitted three extra models dropping one main effect at a time. All predictors are statistically significant. So that all the three main experimental factors (Instrument, Harmony & Voice) still have significant influence on classical ratings.

```r
summary(lmer1c1)$varcor

##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.66563
##  Subject:Voice      (Intercept) 0.16760
##  Subject:Instrument (Intercept) 1.48273
##  Residual                       1.56126

AIC(lmer1c1)

## [1] 10075.51

AIC(lm1a)

## [1] 11230.45

AIC(lmer1b2)

## [1] 10468.86
```

For the three estimated variance components, the variance for the Subject:Voice is the smallest, then is the variance for Subject:Harmony is bigger, variance for Subject:Instrument is the greatest. Since these random intercepts account for personal biases in rating music as classic, I guess this is telling us that there are less variation in bias for people to vote music as classic when they hear different voice leadings. More biases when they hear different harmonic motion, and the most biases when they hear different instruments in the music. However, they are all smaller than the residual variance. This is probably showing that we are having a pretty precise catch of the mean in the intercepts, so the variance is comparatively small. But they are still significant, and offers a better estimates than the model without three random effects.
(iii).

$$\text{for } i = 1,2,\ldots,2493, \text{and } j = 1,2,\ldots,70$$

$$y_i = \alpha^I_{j[i]kI[i]} + \alpha^H_{j[i]kH[i]} + \alpha^V_{j[i]kV[i]} + \beta_{Ii} + \beta_{Hi} + \beta_{Vi} + \epsilon_i$$

$$\alpha^I_{j[i]kI[i]} = \alpha^I_0 + \eta^I_{jk}$$

$$\alpha^H_{j[i]kH[i]} = \alpha^H_0 + \eta^H_{jk}$$

$$\alpha^V_{j[i]kV[i]} = \alpha^V_0 + \eta^V_{jk}$$

$$\epsilon_i \sim N(0, \sigma^2_\epsilon)$$

$$\eta^I_{jk} \sim N(0, \tau^2_{\eta I})$$

$$\eta^H_{jk} \sim N(0, \tau^2_{\eta H})$$

$$\eta_{jk}^{V} \sim N(0, \tau_{\eta V}^2)$$

2. The best model I obtained from question 1 is the 3 random intercept model. I will start from that model and select some terms.

(a).

```
music$CollegeMusic = as.factor(music$CollegeMusic)
music$APTheory = as.factor(music$APTheory)
music$first <- ifelse(is.na(music$X1stInstr), 0, 1)
music$second <- ifelse(is.na(music$X2ndInstr), 0, 1)

lmer2a = lmer(Classical ~ Harmony + Instrument + Voice + Selfdeclare +
OMSI + X16.minus.17 + ConsInstr + ConsNotes + PachListen + ClsListen +
KnowRob + KnowAxis + X1990s2000s + CollegeMusic + NoClass + APTheory +
Composing + PianoPlay + GuitarPlay + first + second + (1|Subject:Instru
ment) + (1|Subject:Harmony) + (1|Subject:Voice), data = music)
bfFixefLMER_F.fnc(lmer2a, method = c("AIC"))

## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Classical ~ Harmony + Instrument + Voice + ConsNotes + PachListen +
KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory + PianoPlay +(1 |
 Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
##    Data: music
## REML criterion at convergence: 6228.579
## Random effects:
##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.6698
##  Subject:Voice      (Intercept) 0.2204
##  Subject:Instrument (Intercept) 1.3059
##  Residual                       1.5747
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Voice, 129; Subject:Instrument, 129
## Fixed Effects:
##    (Intercept)      HarmonyI-V-IV     HarmonyI-V-VI      HarmonyIV-I-V
##       2.115307          -0.004452          0.850076           0.059833
##Instrumentpiano  Instrumentstring       Voicepar3rd        Voicepar5th
##       1.649108           3.588496         -0.403165          -0.299892
##      ConsNotes         PachListen           KnowRob           KnowAxis
##      -0.184582           0.199299          0.085961           0.080629
##    X1990s2000s            NoClass         APTheory1          PianoPlay
##       0.188702          -0.153935          0.631875           0.308238
```

(I am only showing the output for the last step of the selection.) I have recoded the whether you have taken music class in college, and AP music theory class as factor. I have also excluded the variables describing the difference between other variables, since I am worried that there will be multicollinearity in the regression. Then, since there are too many NAs in the proficiency of the 1st and 2nd instrument variable, I decided to combine these two variables (named as first and second). If the first musical instrument is missing, then I would regard the person as not playing any instrument, if not missing, then I would regard the person as playing a first instrument. Similarly for the second musical instrument variable. So if both the variables equals 0, then the person do not play any instrument. If the 1st equals 1 the 2nd variable is 0, then he only plays one instrument. If both are 1, then he plays 2 instruments.

Then, I included all variables left and used backward selection with AIC criteria on my lmer full model. The model selected by this method has the following variables: Harmony + Instrument + Voice + ConsNotes + PachListen + KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory + PianoPlay, and the other 3 random intercept terms.

```
lmer.selected = lmer(Classical ~ Harmony + Instrument + Voice + ConsNot
es + PachListen + KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory
 + PianoPlay + (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 |
Subject:Voice), data = music)

AIC(lmer.selected)

## [1] 6268.579

AIC(lmer2a)

## [1] 6298.681
```

AIC dropped about 30, showing that the selected model has a significant improvement over our full model with all predictors.

(b). Now it is time to chech whether each random effect still makes sense. I used the exactRLRT function to test whether the variance of the random effect is zero one at a time.

```
#for testing (1 | Subject:Instrument)
m1 = lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes + PachLi
sten + KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory + PianoPla
y + (1 | Subject:Instrument) , data = music)
m.null1 = lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes + P
achListen + KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory + Pia
noPlay+ (1| Subject:Harmony) + (1| Subject:Voice), data = music)


exactRLRT(m1, mA = lmer.selected, m0 = m.null1)
##  simulated finite sample distribution of RLRT.
##  (p-value based on 10000 simulated values)
## data:
## RLRT = 343.9, p-value < 2.2e-16

#for testing (1 | Subject:Harmony)
m2 = lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes + PachLi
sten + KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory + PianoPla
y + (1 | Subject:Harmony) , data = music)
```

```
m.null2 = lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes + P
achListen + KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory + Pia
noPlay+ (1 | Subject:Instrument) + (1| Subject:Voice), data = music)


exactRLRT(m2, mA = lmer.selected, m0 = m.null2)
##  simulated finite sample distribution of RLRT.
##  (p-value based on 10000 simulated values)
## data:
## RLRT = 63.955, p-value < 2.2e-16

#for testing (1 | Subject:Voice)
m3 = lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes + PachLi
sten + KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory + PianoPla
y + (1 | Subject:Voice) , data = music)
m.null3 = lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes + P
achListen + KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory + Pia
noPlay+ (1 | Subject:Instrument) + (1 | Subject:Harmony), data = music)


exactRLRT(m3, mA = lmer.selected, m0 = m.null3)
##  simulated finite sample distribution of RLRT.
##  (p-value based on 10000 simulated values)
## data:
## RLRT = 1.7802, p-value = 0.0836
```

Only the Subject:Voice interaction term is not statistically significant on a 5% level. So I am considering dropping this random effect. The final model based on previous part is the model with two random intercepts and all other covariates I have chosen.

```
lmer2c = lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes + Pa
chListen + KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory + Pian
oPlay+ (1 | Subject:Instrument) + (1 | Subject:Harmony), data = music)
BIC(lmer2c)

## [1] 6369.823

BIC(lmer.selected)

## [1] 6375.383
```

The AIC for the models before and after dropping the Subject:Voice is basically the same, but the BIC have dropped about 6, showing an improvement over the 3 random intercept model.

(c).

```
summary(lmer2c)$coefficients

##                       Estimate Std. Error     t value
## (Intercept)        2.115043506 0.98109163  2.15580628
## HarmonyI-V-IV     -0.004525948 0.18393853 -0.02460577
## HarmonyI-V-VI      0.850392080 0.18399150  4.62190959
```

```
## HarmonyIV-I-V     0.060228549 0.18387483  0.32755189
## Instrumentpiano   1.649136667 0.30032421  5.49118795
## Instrumentstring  3.588465707 0.30014726 11.95568360
## Voicepar3rd      -0.402679725 0.09899522 -4.06766828
## Voicepar5th      -0.299981838 0.09899535 -3.03026191
## ConsNotes        -0.184601541 0.07996101 -2.30864439
## PachListen        0.199297197 0.17584201  1.13338786
## KnowRob           0.085995660 0.08698230  0.98865695
## KnowAxis          0.080606693 0.07016950  1.14874260
## X1990s2000s       0.188716135 0.09091377  2.07577073
## NoClass          -0.153948461 0.10631890 -1.44798770
## APTheory1         0.631951900 0.36292056  1.74129538
## PianoPlay         0.308236066 0.08769287  3.51495017
```

For harmony, people are more willing to rate the music as classical when they hear HarmonyI-V-VI, the other three harmonies have basically the same ratings. For instruments, when String Quartet is identified, the rating is the highest. Then is the piano, and electric guitar is getting the lowest ratings on classical. For the voice leading, Contrary Motion is getting the highest score, then is the Parallel 5ths, Parallel 3rds is getting the lowest score.

Besides harmony, instrument, and voice, we have included 8 other variables into the model as fixed effects. Now let me take a look at each of them. Based on the regression output, it shows that the ratings for classical music actually decreases when people focus more on the notes while listening, and when they have taken more music classes. Ratings for classical music increases when people are more familiar with Pachelbel's Canon in D, have heard Rob Paravonian's Pachelbel Rant for more times, listen to more pop and rock from the 90's and 2000's, have taken AP Music Theory class, and plays plano. The higher they rated themselves in these questions, the higher they rated the music as classical.

3.

```
music$self[music$Selfdeclare<=2] = 0
music$self[music$Selfdeclare>2] = 1
```

I took a look at the distribution of the self declared musician variable. It seems that 2 is a cut off point for dichotomizing the data. Then, I will add all the interactions of this variable and use backward selection with AIC criteria to select model.

```
lmer3 = lmer(Classical ~ self + Harmony + Harmony:self + Instrument + I
nstrument:self + Voice + Voice:self + ConsNotes + ConsNotes:self + Pach
Listen + PachListen:self + KnowRob + KnowRob:self + KnowAxis + KnowAxis:
self + X1990s2000s + X1990s2000s:self + NoClass + NoClass:self + APTheo
ry + APTheory:self + PianoPlay + PianoPlay:self + (1 | Subject:Instrume
nt) + (1 | Subject:Harmony), data = music)
bfFixefLMER_F.fnc(lmer3, method = c("AIC"))
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Classical ~ self + Harmony + Instrument + Voice + ConsNotes
```

```
 + PachListen + KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory +
 PianoPlay + (1 | Subject:Instrument) + (1 | Subject:Harmony) + self:Ha
rmony + self:PianoPlay
##    Data: music
## REML criterion at convergence: 6205.618
## Random effects:
##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.5965
##  Subject:Instrument (Intercept) 1.2769
##  Residual                       1.5862
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Instrument, 129
## Fixed Effects:
##        (Intercept)                self       HarmonyI-V-IV
##           1.626729            0.081748            0.009662
##       HarmonyI-V-VI        HarmonyIV-I-V     Instrumentpiano
##           0.266757            0.004831            1.650690
##    Instrumentstring         Voicepar3rd         Voicepar5th
##           3.588724           -0.402700           -0.299994
##          ConsNotes          PachListen             KnowRob
##          -0.193661            0.277241            0.117490
##           KnowAxis         X1990s2000s             NoClass
##           0.028935            0.221350           -0.135895
##          APTheory1           PianoPlay  self:HarmonyI-V-IV
##           0.577703            0.761417           -0.029644
## self:HarmonyI-V-VI  self:HarmonyIV-I-V      self:PianoPlay
##           1.255995            0.118552           -0.605512
```

The final model chosen introduced the following interaction terms self:Harmony + self:PianoPlay. Since the harmony variable has 4 levels, there are actually four interaction terms that are included in the model.

```
lmer3.selected = lmer(Classical ~ self + Harmony + Instrument + Voice +
 ConsNotes + PachListen + KnowRob + KnowAxis + X1990s2000s + NoClass +
APTheory + PianoPlay + (1 | Subject:Instrument) + (1 | Subject:Harmony)
 + self:Harmony + self:PianoPlay, data = music)
AIC(lmer3)

## [1] 6264.794

AIC(lmer3.selected)

## [1] 6253.617

AIC(lmer2c)

## [1] 6268.359

summary(lmer3.selected)$coefficients

##                        Estimate Std. Error     t value
## (Intercept)         1.626728809 0.98891599  1.64496157
```

```
## self                 0.081747553 0.40956669  0.19959522
## HarmonyI-V-IV        0.009661836 0.23504390  0.04110652
## HarmonyI-V-VI        0.266756790 0.23518648  1.13423524
## HarmonyIV-I-V        0.004830918 0.23504390  0.02055326
## Instrumentpiano      1.650690252 0.29276108  5.63835281
## Instrumentstring     3.588723536 0.29258052 12.26576367
## Voicepar3rd         -0.402699786 0.09900453 -4.06748829
## Voicepar5th         -0.299993770 0.09900453 -3.03010130
## ConsNotes           -0.193661045 0.07811317 -2.47923669
## PachListen           0.277241340 0.17145912  1.61695299
## KnowRob              0.117490034 0.08480509  1.38541258
## KnowAxis             0.028934592 0.07021906  0.41206181
## X1990s2000s          0.221349586 0.09048509  2.44625475
## NoClass             -0.135895349 0.10437193 -1.30202968
## APTheory1            0.577702661 0.35122827  1.64480686
## PianoPlay            0.761417180 0.16301485  4.67084548
## self:HarmonyI-V-IV  -0.029643560 0.34503788 -0.08591393
## self:HarmonyI-V-VI   1.255995225 0.34513522  3.63913955
## self:HarmonyIV-I-V   0.118552484 0.34488331  0.34374665
## self:PianoPlay      -0.605511656 0.18902461 -3.20334834
```

AIC for the model with selected terms plus the two selected interactions is the lowest compared to model with no interaction and the model with all interactions. The interaction terms show that when participants categorize themselves as musicians, the relationship between harmony and ratings for how classic the music sounds, and the relationship between whether the participant plays piano and how classic the music sounds, are altered.

When they identify themselves as musicians, the pianoplay variable has a smaller effect on classical rating, and identifying harmony as I-V-IV would decrease the classical ratings compared to the baseline I-VI-V harmony. Identifying other harmonies increases classical ratings compared to the baseline.

```
summary(lmer3.selected)$varcor
```

```
##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.59647
##  Subject:Instrument (Intercept) 1.27694
##  Residual                       1.58617
```

Still, the estimated variance for the random intercepts are smaller than the residual variance, showing that the fixed effect in the model did a good job capturing the mean in the data, so that we have small variations.

4.(a).

```
lmer1c.pop = lmer(Popular ~ 1 + Instrument + Harmony + Voice + (1|Subje
ct:Instrument) + (1|Subject:Harmony) + (1|Subject:Voice), data = music)
lmer1c.pop1 = lmer(Popular ~ 1 + Harmony + Voice + (1|Subject:Instrumen
t) + (1|Subject:Harmony) + (1|Subject:Voice), data = music)
lmer1c.pop2 = lmer(Popular ~ 1 + Instrument + Voice + (1|Subject:Instru
```

```
ment) + (1|Subject:Harmony) + (1|Subject:Voice), data = music)
lmer1c.pop3 = lmer(Popular ~ 1 + Instrument + Harmony + (1|Subject:Inst
rument) + (1|Subject:Harmony) + (1|Subject:Voice), data = music)
anova(lmer1c.pop, lmer1c.pop1)

## refitting model(s) with ML (instead of REML)

## Data: music
## Models:
## lmer1c.pop1: Popular ~ 1 + Harmony + Voice + (1 | Subject:Instrument)
 + (1 | Subject:Harmony) + (1 | Subject:Voice)
## lmer1c.pop: Popular ~ 1 + Instrument + Harmony + Voice + (1 | Subjec
t:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
##             Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## lmer1c.pop1 10 10162 10220 -5070.9    10142
## lmer1c.pop  12 10079 10149 -5027.5    10055 86.87   2  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

anova(lmer1c.pop, lmer1c.pop2)

## refitting model(s) with ML (instead of REML)

## Data: music
## Models:
## lmer1c.pop2: Popular ~ 1 + Instrument + Voice + (1 | Subject:Instrum
ent) + (1 | Subject:Harmony) + (1 | Subject:Voice)
## lmer1c.pop: Popular ~ 1 + Instrument + Harmony + Voice + (1 | Subjec
t:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
##             Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer1c.pop2  9 10078 10130 -5030.0    10060
## lmer1c.pop  12 10079 10149 -5027.5    10055 5.1175      3     0.1634

anova(lmer1c.pop, lmer1c.pop3)

## refitting model(s) with ML (instead of REML)

## Data: music
## Models:
## lmer1c.pop3: Popular ~ 1 + Instrument + Harmony + (1 | Subject:Instr
ument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
## lmer1c.pop: Popular ~ 1 + Instrument + Harmony + Voice + (1 | Subjec
t:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice)
##             Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## lmer1c.pop3 10 10080 10138 -5030.0    10060
## lmer1c.pop  12 10079 10149 -5027.5    10055 5.0782      2    0.07894 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Following what I did in question 1 part c, I fitted lmer model with random intercept, and then used anova test to see if the individual factors are significant. The instrument variable is as important as usual, with a test statistics of approximately zero. However, on a 5% significance level, the harmony and voice variables are not statistically significant. This means that they do not have statistically significant influence ratings for how popular the music sounds.

(b).

```
lmer2.pop = lmer(Popular ~ Harmony + Instrument + Voice + Selfdeclare +
 OMSI + X16.minus.17 + ConsInstr + ConsNotes + PachListen + ClsListen +
 KnowRob + KnowAxis + X1990s2000s + CollegeMusic + NoClass + APTheory +
 Composing + PianoPlay + GuitarPlay + first + second + (1|Subject:Instr
ument) + (1|Subject:Harmony) + (1|Subject:Voice), data = music)
bfFixefLMER_F.fnc(lmer2.pop, method = c("AIC"))

## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Popular ~ Instrument + ConsNotes + PachListen + KnowRob + KnowAxis +
 X1990s2000s + NoClass + APTheory + (1 | Subject:Instrument) + (1 | Sub
ject:Harmony) + (1 | Subject:Voice)
##    Data: music
## REML criterion at convergence: 6342.393
## Random effects:
##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.6728
##  Subject:Voice      (Intercept) 0.2501
##  Subject:Instrument (Intercept) 1.3365
##  Residual                       1.6422
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Voice, 129; Subject:Instrument, 129
## Fixed Effects:
##    (Intercept)   Instrumentpiano  Instrumentstring         ConsNotes
##        7.49926          -1.14827          -3.02445           0.09936
##     PachListen           KnowRob          KnowAxis       X1990s2000s
##       -0.25424           0.07241           0.07219           0.01391
##        NoClass         APTheory1
##        0.09633          -0.03344
```

Then, I added all the predictors as before, and used backward selection with AIC criteria to select models. The final model chosen has the following predictors:Instrument + ConsNotes + PachListen + KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory, and the three random effects. Indeed, the selection dropped harmony and voice, which we have found out to be not important for popular rating previously.

Now let me see if random effect need to be fixed.

```
pop.selected = lmer(Popular ~ Instrument + ConsNotes + PachListen + Kno
wRob + KnowAxis + X1990s2000s + NoClass + APTheory + (1 | Subject:Instr
ument) +  (1 | Subject:Harmony) + (1 | Subject:Voice), data = music)
#for testing (1 | Subject:Instrument)
m1.pop = lmer(Popular ~ Instrument + ConsNotes + PachListen + KnowRob +
 KnowAxis + X1990s2000s + NoClass + APTheory + (1 | Subject:Instrument)
```

```
 , data = music)
mpop.null1 = lmer(Popular ~ Instrument + ConsNotes + PachListen + KnowR
ob + KnowAxis + X1990s2000s + NoClass + APTheory+ (1| Subject:Harmony)
+ (1| Subject:Voice), data = music)


exactRLRT(m1.pop, mA = pop.selected, m0 = mpop.null1)
##  simulated finite sample distribution of RLRT.
##  (p-value based on 10000 simulated values)
## data:
## RLRT = 248.16, p-value < 2.2e-16

#for testing (1 | Subject:Harmony)
m2.pop = lmer(Popular ~ Instrument + ConsNotes + PachListen + KnowRob +
 KnowAxis + X1990s2000s + NoClass + APTheory + (1 | Subject:Harmony) ,
data = music)
mpop.null2 = lmer(Popular ~ Instrument + ConsNotes + PachListen + KnowR
ob + KnowAxis + X1990s2000s + NoClass + APTheory+ (1 | Subject:Instrume
nt) + (1| Subject:Voice), data = music)


exactRLRT(m2.pop, mA = pop.selected, m0 = mpop.null2)
##  simulated finite sample distribution of RLRT.
##  (p-value based on 10000 simulated values)
## data:
## RLRT = 55.208, p-value < 2.2e-16

#for testing (1 | Subject:Voice)
m3.pop = lmer(Popular ~ Instrument + ConsNotes + PachListen + KnowRob +
 KnowAxis + X1990s2000s + NoClass + APTheory + (1 | Subject:Voice) , da
ta = music)
mpop.null3 = lmer(Popular ~ Instrument + ConsNotes + PachListen + KnowR
ob + KnowAxis + X1990s2000s + NoClass + APTheory+ (1 | Subject:Instrume
nt) + (1 | Subject:Harmony), data = music)


exactRLRT(m3.pop, mA = pop.selected, m0 = mpop.null3)
##  simulated finite sample distribution of RLRT.
##  (p-value based on 10000 simulated values)
## data:
## RLRT = 2.3345, p-value = 0.0585
```

The Subject:Voice random effect is statistically insignificant. So I decided to drop this random effect.

```
pop.2b = lmer(Popular ~ Instrument + ConsNotes + PachListen + KnowRob +
 KnowAxis + X1990s2000s + NoClass + APTheory + (1 | Subject:Instrument)
 +  (1 | Subject:Harmony), data = music)
BIC(pop.2b)

## [1] 6440.15
```

```r
BIC(pop.selected)
```

```
## [1] 6445.155
```

BIC dropped for 5 after we dropped the Subject:Voice random effect, showing an improvement of the model.

(c). Now let me check if we should include any any interactions.

```r
lmer3.pop = lmer(Popular ~ self + Instrument + Instrument:self + ConsNo
tes + ConsNotes:self + PachListen + PachListen:self + KnowRob + KnowRob:
self + KnowAxis + KnowAxis:self + X1990s2000s + X1990s2000s:self + NoCl
ass + NoClass:self + APTheory + APTheory:self + (1 | Subject:Instrument)
 + (1 | Subject:Harmony), data = music)
bfFixefLMER_F.fnc(lmer3.pop, method = c("AIC"))
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Popular ~ Instrument + ConsNotes + PachListen + KnowRob + KnowAxis +
X1990s2000s + NoClass + APTheory + (1 | Subject:Instrument) + (1 | Subj
ect:Harmony)
##    Data: music
## REML criterion at convergence: 6344.727
## Random effects:
##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.6706
##  Subject:Instrument (Intercept) 1.3555
##  Residual                       1.6557
## Number of obs: 1541, groups:
## Subject:Harmony, 172; Subject:Instrument, 129
## Fixed Effects:
##  (Intercept)   Instrumentpiano  Instrumentstring         ConsNotes
##      7.49928          -1.14833          -3.02454           0.09934
##   PachListen           KnowRob          KnowAxis       X1990s2000s
##     -0.25422           0.07243           0.07217           0.01393
##      NoClass         APTheory1
##      0.09631          -0.03341
```

No interactions were selected. It seems that identifying oneself as a musician does not have significant influence on the ratings for popular music.

5. Summary

Introduction

Researchers are interested in measuring the influence of instrument, harmonic motion, and voice leading on listeners' identification of music as classical or popular. A designed experiment was conducted on 70 listeners at the University of Pittsburgh, with 36 responses from each listener. The hypothesis researchers proposed are: instruments should have the largest influence on ratings; the harmonic progression, I-V-VI might be frequently rated as classical; and contrary motion would also be frequently rated as classical.

Methods

Ratings on classical and popular music were examined separately. Hierarchical models with random intercepts were employed to test above hypotheses. The models we have fitted are a bit different from the usual repeated measures model, where a random intercept was fitted for each participant to account for each individual level bias of ratings for classical music. Interaction terms are included in random intercepts in our model to account for potential bias that varies with the types of instrument, harmony, and voice leading. Thus we are drawing random samples from each subject and instrument, harmony, and voice leading combinations. By doing this, we captured extra piece of variance components in each individual predictor. Main effect predictors and interaction terms were chosen by backward selection methods with AIC criteria.

Results

Results are discussed on the final models only. Output of final model for classical ratings is in question 3. For classical music ratings, all proposed hypotheses have been proved. First, a 10-score-scale rating increase the most when the listeners distinguished there is change in instrument. When the instrument distinguished changes from electric guitar to piano, ratings would increase for 1.65, and the increase is expected to change as high as 3.59 when the instrument changes from piano to string quartet. Second, compared to the harmony I-VI-V, the greatest change in ratings occurs when listeners identifies the harmony I-V-VI, with an expected increase in score by 0.3. Third, for voice leading, ratings were highest when contrary motion was identified. Besides proofs for hypotheses, other factors were also proved to have significant influence on classical ratings. Classical rating decreases if the listener pays more attention on the notes, or has taken more music classes, increases if the listener knows more about Pachelbel's Canon, Rob Paravonian's Pachelbel Rant, Axis of Evil's Comedy bit on the 4 Pachelbel chords in popular music, listen to pop and rock from the 90's and 2000's, has taken AP Music classes, and plays piano. When they identify themselves as musicians, the more they play piano, the smaller effect it has on increasing classical rating, and identifying harmony as I-V-IV would decrease the classical ratings compared to the baseline I-VI-V harmony. Identifying other harmonies increases classical ratings compared to the baseline.

Output of final model for popular ratings is in question 4. The first hypothesis was still proved correct for popular ratings. Different from classical ratings, harmony and voice leading do not have significant influence on popular ratings. Only instrument has. Identifying instrument as piano decreases popular rating by 1.15, and by 3.02 for string quartet, both compared to electric guitar. Knowing more about Pachelbel's Canon, and have taken AP Music class decreases popular rating, whereas concentrating on notes, knowing more about Rob Paravonian's Pachelbel

Rant, and Axis of Evil's Comedy bit on the 4 Pachelbel chords in popular music, listen to pop and rock from the 90's and 2000's, taking less music classes increases the rating. Identifying oneself as a musician does not seem to have significant influence on popular ratings.