Hierarchical Linear Models Homework 5

Cici Wang

December 10, 2013

1.

(a)

```
> setwd("C:/Users/Chencheng Wang/Dropbox/CMU")
> library(ggplot2)
> library(arm)
> library(lme4)
> library(RLRsim)
> rating=read.csv("ratings.csv",header=T)
> attach(rating)
> lm1=lm(Classical~Harmony+Voice+Instrument,data=rating)
> summary(lm1)

Call:
lm(formula = Classical ~ Harmony + Voice + Instrument, data = rating)

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 ***
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 **
Instrumentpiano   1.37359    0.11298  12.158  < 2e-16 ***
Instrumentstring  3.13312    0.11230  27.899  < 2e-16 ***
---
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

> lm2=lm(Classical~Voice+Instrument,data=rating)
> lm3=lm(Classical~Harmony+Instrument,data=rating)
> lm4=lm(Classical~Voice+Harmony,data=rating)
> anova(lm1,lm2)
```

```
Analysis of Variance Table

Model 1: Classical ~ Harmony + Voice + Instrument
Model 2: Classical ~ Voice + Instrument
  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

> anova(lm1,lm3)

Analysis of Variance Table

Model 1: Classical ~ Harmony + Voice + Instrument
Model 2: Classical ~ Harmony + Instrument
  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

> anova(lm1,lm4)

Analysis of Variance Table

Model 1: Classical ~ Harmony + Voice + Instrument
Model 2: Classical ~ Voice + Harmony
  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
```

From the linear model test, we can see that for experimental factor Harmony, only the second kind HarmonyI-V-VI is significant, it has small p-value less than 0.01. And we found that every kinds of experimental factor Voice are significant and also for Instrument. From the anova test, we find that each P-value is smaller than 0.01, which means that experimental factors Harmony, voice and Instrument are all significant.

(b)

(i)

model:

$$Classical_i = \beta_1 * Instrument_i + \beta_2 * Voice_i + \beta_3 * Harmony_i + \alpha_{j[i]} + \epsilon_i, \epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

$$\alpha_j = \gamma_0 + \phi_j,\ \phi_j \overset{iid}{\sim} N(0, \tau_0^2)$$

(ii)

Method 1:

```
> repeat1=lmer(Classical~Instrument+Voice+Harmony+(1|Subject),data=rating)
> model1=list(lm1,repeat1)
> sapply(model1,AIC)
```

```
[1] 11230.45 10491.51
```

```
> sapply(model1,BIC)
```

```
[1] 11282.84 10549.73
```

If we use AIC and BIC to test whether we need the random intercept, we can find that the model included random intercept has a smaller AIC value and a smaller BIC value. So in this aspect, random intercept is needed.

Method 2:

```
> exactRLRT(repeat1)

        simulated finite sample distribution of RLRT.

        (p-value based on 10000 simulated values)

data:
RLRT = 763.3759, p-value < 2.2e-16
```

Here we have strong rejection of $H_0 : \tau^2 = 0$ since p-value $<$ 2.2e-16. So we keep the random effect.

(iii)

```
> repeat2=lmer(Classical~Instrument+Voice+(1|Subject),data=rating)
> repeat3=lmer(Classical~Instrument+Harmony+(1|Subject),data=rating)
> repeat4=lmer(Classical~Voice+Harmony+(1|Subject),data=rating)
> model2=list(repeat1,repeat2,repeat3,repeat4)
> sapply(model2,AIC)
```

```
[1] 10491.51 10552.74 10505.58 11423.04
```

```
> sapply(model2,BIC)
```

```
[1] 10549.73 10593.49 10552.15 11469.60
```

We fitted all models again and compared their AIC and BIC scores. We found that the model has all three main experimental factors with random intercept in it has the smallest AIC and the smallest BIC. So this model should be the best model so far.

(c)

(i)

```
> lmer1=lmer(Classical~Instrument+Voice+Harmony+(1|Subject:Instrument)+(1|Subject:Voice)+(1|Subject:Harm
> model3=list(lm1,repeat1,lmer1)
> sapply(model3,AIC)
```

```
[1] 11230.45 10491.51 10075.51
```

```
> sapply(model3,BIC)
```

```
[1] 11282.84 10549.73 10145.37
```

From the result of AIC values, we can find that model in part 1a has the largest AIC and BIC value, model with three new random effects has the smallest AIC and BIC value.Thus, This model is better than both models in part 1a and 1b.

(ii)

```
> lmer2=lmer(Classical~Instrument+Voice+(1|Subject:Instrument)+(1|Subject:Voice)+(1|Subject:Harmony),da
> lmer3=lmer(Classical~Instrument+Harmony+(1|Subject:Instrument)+(1|Subject:Voice)+(1|Subject:Harmony),
> lmer4=lmer(Classical~Voice+Harmony+(1|Subject:Instrument)+(1|Subject:Voice)+(1|Subject:Harmony),data=
> model4=list(lmer1,lmer2,lmer3,lmer4)
> sapply(model4,AIC)

[1] 10075.51 10101.74 10092.66 10176.17

> sapply(model4,BIC)

[1] 10145.37 10154.13 10150.87 10234.38

> lmer5=lmer(Classical~Voice+Harmony+(1|Subject:Voice)+(1|Subject:Harmony),data=rating)
> lmer6=lmer(Classical~Voice+Harmony+(1|Subject:Instrument)+(1|Subject:Harmony),data=rating)
> lmer7=lmer(Classical~Voice+Harmony+(1|Subject:Instrument)+(1|Subject:Voice),data=rating)
> model4.1=list(lmer1,lmer5,lmer6,lmer7)
> sapply(model4.1,AIC)

[1] 10075.51 11617.96 10174.65 10269.04

> sapply(model4.1,BIC)

[1] 10145.37 11670.35 10227.04 10321.43

>
```

When we do the re-examine, our strategy is taking one main factor out each time, and fit the model again. After comparing main effects, we took out one random intercept each time, and compare them. When we compared all models together, we could find out that the most complicated model with all three experiment factors and their random effects has the smallest AIC value and the smallest BIC value. So this model should be the best model we have so far.

```
> display(lmer1)

lmer(formula = Classical ~ Instrument + Voice + Harmony + (1 |
    Subject:Instrument) + (1 | Subject:Voice) + (1 | Subject:Harmony),
    data = rating)
                 coef.est coef.se
(Intercept)       4.34     0.21
Instrumentpiano   1.36     0.26
Instrumentstring  3.13     0.26
Voicepar3rd      -0.41     0.08
Voicepar5th      -0.37     0.08
HarmonyI-V-IV    -0.03     0.14
HarmonyI-V-VI     0.77     0.14
HarmonyIV-I-V     0.06     0.14

Error terms:
 Groups             Name        Std.Dev.
 Subject:Harmony    (Intercept) 0.67
 Subject:Voice      (Intercept) 0.17
 Subject:Instrument (Intercept) 1.48
 Residual                       1.56
---
number of obs: 2493, groups: Subject:Harmony, 280; Subject:Voice, 210; Subject:Instrument, 210
AIC = 10075.5, DIC = 10015.5
deviance = 10033.5
```

We have three groups in our model.The variance of this group equals to $0.67^2 = 0.4489$. The second group is Subject:Voice, which has a variance equals to $0.17^2 = 0.0289$. And the third group is Subject:Instrument, the variance is $1.48^2 = 2.1904$. The residual variance for this model is $1.56^2 = 2.4336$.All three groups have smaller variance than the residual variance.

(iii)

$$Classical_i = \alpha_1 * Instrument_i + \alpha_2 * Voice_i + \alpha_3 * Harmony_i + \beta_{1j[i]} + \beta_{2j[i]} + \beta_{3j[i]} + \epsilon_i, \epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

,

$$\beta_{1j} = \zeta_{Subject:Instrument} + \psi_j,\ \psi_j \overset{iid}{\sim} N(0, \tau_1^2)$$

$$\beta_{2j} = \gamma_{Subject:Voice} + \phi_j,\ \phi_j \overset{iid}{\sim} N(0, \tau_2^2)$$

$$\beta_{3j} = \delta_{Subject:Harmony} + \omega_j,\ \omega_j \overset{iid}{\sim} N(0, \tau_3^2)$$

2.

(a)

```
> newrating=subset(rating,X16.minus.17!='NA' & ConsNotes!='NA'
+ & ClsListen!='NA' & PianoPlay!='NA' & GuitarPlay!='NA'
+ & CollegeMusic!='NA' & Selfdeclare!='NA' )

The following objects are masked from rating:

    APTheory, Classical, ClsListen, CollegeMusic, Composing, ConsInstr,
    ConsNotes, first12, GuitarPlay, Harmony, Instr.minus.Notes,
    Instrument, KnowAxis, KnowRob, NoClass, OMSI, PachListen,
    PianoPlay, Popular, Selfdeclare, Subject, Voice, X, X16.minus.17,
    X1990s2000s, X1990s2000s.minus.1960s1970s, X1stInstr, X2ndInstr

> lmer0=lmer(Classical~Instrument+Voice+Harmony+(1|Subject:Instrument)+(1|Subject:Voice)+(1|Subject:Harm
> add1=update(lmer0,.~.+X16.minus.17,data=newrating)
> anova(lmer0,add1)

Data: newrating
Models:
lmer0: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
lmer0:     (1 | Subject:Voice) + (1 | Subject:Harmony)
add1: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add1:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17
      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer0 12 8396.3 8463.9 -4186.1   8372.3
add1  13 8393.9 8467.2 -4184.0   8367.9 4.3258      1    0.03754 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> add2=update(add1,.~.+ConsNotes1,data=newrating)
> anova(add1,add2)

Data: newrating
Models:
add1: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add1:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17
add2: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
```

```
add2:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17 +
add2:     ConsNotes1
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
add1 13 8393.9 8467.2 -4184.0   8367.9
add2 14 8395.4 8474.3 -4183.7   8367.4 0.5526      1     0.4572

> add3=update(add1,.~.+ClsListen1,data=newrating)
> anova(add1,add3)

Data: newrating
Models:
add1: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add1:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17
add3: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add3:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17 +
add3:     ClsListen1
     Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
add1 13 8393.9 8467.2 -4184.0   8367.9
add3 14 8388.2 8467.1 -4180.1   8360.2 7.709      1   0.005495 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> add4=update(add3,.~.+PianoPlay1,data=newrating)
> anova(add3,add4)

Data: newrating
Models:
add3: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add3:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17 +
add3:     ClsListen1
add4: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add4:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17 +
add4:     ClsListen1 + PianoPlay1
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
add3 14 8388.2 8467.1 -4180.1   8360.2
add4 15 8389.8 8474.4 -4179.9   8359.8 0.3716      1     0.5421

> add5=update(add3,.~.+GuitarPlay1,data=newrating)
> anova(add3,add5)

Data: newrating
Models:
add3: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add3:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17 +
add3:     ClsListen1
add5: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add5:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17 +
add5:     ClsListen1 + GuitarPlay1
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
add3 14 8388.2 8467.1 -4180.1   8360.2
add5 15 8389.1 8473.6 -4179.5   8359.1 1.1503      1     0.2835

> add6=update(add3,.~.+CollegeMusic1,data=newrating)
> anova(add3,add6)
```

```
Data: newrating
Models:
add3: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add3:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17 +
add3:     ClsListen1
add6: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add6:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17 +
add6:     ClsListen1 + CollegeMusic1
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
add3 14 8388.2 8467.1 -4180.1   8360.2
add6 15 8390.2 8474.8 -4180.1   8360.2 0.0012      1     0.9723

> add7=update(add3,.~.+musician,data=newrating)
> anova(add3,add7)

Data: newrating
Models:
add3: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add3:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17 +
add3:     ClsListen1
add7: Classical ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
add7:     (1 | Subject:Voice) + (1 | Subject:Harmony) + X16.minus.17 +
add7:     ClsListen1 + musician
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
add3 14 8388.2 8467.1 -4180.1   8360.2
add7 15 8382.5 8467.1 -4176.3   8352.5 7.7154      1   0.005475 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since there are some variables with NA values, so we restricted our data set to a smaller data set that has no missing values.Then, We added seven different possible new variables into our best model. They are X16.minus.17, ConsNotes, ClsListen, PianoPlay, GuitarPlay,CollegeMusic and Selfdeclare. X16.minus.17 is the auxiliary measure of listener's ability to distinguish classical and popular music, so I think include this variable should help me to discuss. And since this variable is from -4 to 9, so I treated this variable as a continues variable. Moreover, I also included ConsNotes, ClsListen, PianoPlay, GuitarPlay,CollegeMusic and Selfdeclare. I think these variables are all reasonable for us to include them to discuss.Concentration of notes taking, the level of listening to classical music, do they think they are musicians, do they play piano or not and do they play guitar or not all could be influences of affecting the result of classical rating. And I used boxplots to see their data distribution, I decided to categorized them into two levels,high or low.

When doing the anova test, if we got a small p-value, we will say that this variable is significant for the model. So we will include that variable and continue to examine the next one. Finally, our model has new variables X16.minus.17, ClsListen,and musician.

(b)

```
> random0=lm(Classical~Instrument+Voice+Harmony+X16.minus.17+ClsListen+musician,data=newrating)
> random=lmer(Classical~Instrument+Voice+Harmony+X16.minus.17+ClsListen+musician+(1|Subject:Instrument)
> modelr=list(random0,random)
> sapply(modelr,BIC)

[1] 9333.255 8545.463

> random.1=lmer(Classical~Instrument+Voice+Harmony+X16.minus.17+ClsListen+musician+(1|Subject:Instrumen
> anova(random,random.1)
```

```
Data: newrating
Models:
random: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
random:     musician + (1 | Subject:Instrument)
random.1: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
random.1:     musician + (1 | Subject:Instrument) + (1 | Subject:Voice)
         Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
random   13 8445.2 8518.5 -4209.6   8419.2
random.1 14 8447.2 8526.1 -4209.6   8419.2     0      1          1
```

```
> random.2=lmer(Classical~Instrument+Voice+Harmony+X16.minus.17+ClsListen+musician+(1|Subject:Instrumen
> anova(random,random.2)
```

```
Data: newrating
Models:
random: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
random:     musician + (1 | Subject:Instrument)
random.2: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
random.2:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony)
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
random   13 8445.2 8518.5 -4209.6   8419.2
random.2 14 8381.7 8460.6 -4176.8   8353.7 65.539      1  5.698e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> random1=update(random.2,.~.+(1|Subject:X16.minus.17))
> anova(random.2,random1)
```

```
Data: newrating
Models:
random.2: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
random.2:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony)
random1: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
random1:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
random1:     (1 | Subject:X16.minus.17)
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
random.2 14 8381.7 8460.6 -4176.8   8353.7
random1  15 8362.9 8447.5 -4166.5   8332.9 20.742      1  5.255e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> random2=update(random1,.~.+(1|Subject:ClsListen1))
> anova(random1,random2)
```

```
Data: newrating
Models:
random1: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
random1:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
random1:     (1 | Subject:X16.minus.17)
random2: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
random2:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
random2:     (1 | Subject:X16.minus.17) + (1 | Subject:ClsListen1)
        Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
random1 15 8362.9 8447.5 -4166.5   8332.9
random2 16 8364.9 8455.1 -4166.5   8332.9     0      1          1
```

```
> random3=update(random1,.~.+(1|Subject:musician))
> anova(random1,random3)

Data: newrating
Models:
random1: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
random1:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
random1:     (1 | Subject:X16.minus.17)
random3: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
random3:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
random3:     (1 | Subject:X16.minus.17) + (1 | Subject:musician)
        Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
random1 15 8362.9 8447.5 -4166.5   8332.9
random3 16 8364.9 8455.1 -4166.5   8332.9     0      1          1

>
>
```

We firstly went back and checked three random effects for Harmony,Voice and Instrument.Since in problem 1 we used the whole dataset to do our analysis, and in this problem we deleted some columns that have NAs. So we fitted the best model from problem 1 again using our new dataset. In the anova test, we can see that adding random intercepts for Harmony and Instrument are significant. Adding random effect for Voice is not significant.

We added three new random effects into the model each of a time. After testing by anova, we found out that adding new random effects for ClsListen and GuitarPlay in model are not significant since they all have a really large p-value. And adding new random effects for X16.minus.17 is significant. So we should also include random effects for X16.minus.17 into our model.Also, we found AIC, BIC values for the new model. It shows that this model has smaller AIC and BIC values than the original model. So this model should be better.

(c)

```
> summary(random1)

Linear mixed model fit by REML ['lmerMod']
Formula: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +      musician + (1 | Subj
   Data: newrating

REML criterion at convergence: 8356.635

Random effects:
 Groups               Name        Variance Std.Dev.
 Subject:Harmony      (Intercept) 0.3518   0.5931
 Subject:Instrument   (Intercept) 1.2996   1.1400
 Subject:X16.minus.17 (Intercept) 1.0206   1.0102
 Residual                         2.5134   1.5854
Number of obs: 2073, groups: Subject:Harmony, 232; Subject:Instrument, 174; Subject:X16.minus.17, 58

Fixed effects:
                 Estimate Std. Error t value
(Intercept)       4.53783    0.48176   9.419
Instrumentpiano   1.46594    0.22848   6.416
Instrumentstring  3.21743    0.22813  14.104
Voicepar3rd      -0.38949    0.08534  -4.564
Voicepar5th      -0.37507    0.08531  -4.396
```

```
HarmonyI-V-IV    -0.01659    0.14776  -0.112
HarmonyI-V-VI     0.84146    0.14780   5.693
HarmonyIV-I-V     0.07918    0.14772   0.536
X16.minus.17     -0.07833    0.05670  -1.381
ClsListen         0.29445    0.11379   2.588
musician         -0.94547    0.48579  -1.946

Correlation of Fixed Effects:
            (Intr) Instrmntp Instrmnts Vcpr3r Vcpr5t HI-V-I HI-V-V HIV-I-
Instrumntpn -0.237
Instrmntstr -0.237  0.499
Voicepar3rd -0.089  0.000     0.000
Voicepar5th -0.088  0.000     0.000     0.500
HrmnyI-V-IV -0.153  0.000     0.000     0.000 -0.001
HrmnyI-V-VI -0.153  0.000     0.000     0.001 -0.001  0.500
HrmnyIV-I-V -0.153  0.000     0.000     0.001 -0.001  0.500  0.500
X16.mins.17 -0.229  0.000     0.000     0.000  0.000  0.000  0.000  0.000
ClsListen   -0.254  0.000     0.000     0.000  0.000  0.000  0.000  0.000
musician    -0.689  0.000     0.000     0.000  0.000  0.000  0.000  0.000
            X16..1 ClsLst
Instrumntpn
Instrmntstr
Voicepar3rd
Voicepar5th
HrmnyI-V-IV
HrmnyI-V-VI
HrmnyIV-I-V
X16.mins.17
ClsListen    0.085
musician    -0.007 -0.318
```

In the final model, we can interpret the effect of each variable individually. When we have one more unit of piano instrument, the rating of classical will increase by 1.4659. When we have one more unit of string instrument, the rating of classical will increase by 3.2174. When we have one more unit of 3rd par voice, the rating of classical will decrease by 0.3895. When we have one more unit of 5th par voice, the rating of classical will decrease by 0.3751. And if the Harmony level 1 increase one unit, the rating of classical will be decreased by 0.0166.If the Harmony level 2 increase one unit, the rating of classical will be increased by 0.8415.If the Harmony level 3 increase one unit, the rating of classical will be increased by 0.0792.

If the measure of listener's ability to distinguish classical or popular music increase one unit, the rating of classical music will be reduced by 0.0783. If each level of listening to classical music increase one unit, then the rating of classical would be increased by 0.2944. If the number of musicians increase one unit, the rating of classical will be decreased by 0.9454.

3.

```
> music1=update(random1,.~.+Harmony:musician)
> anova(random1,music1)

Data: newrating
Models:
random1: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
random1:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
random1:     (1 | Subject:X16.minus.17)
music1: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
music1:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
```

```
music1:     (1 | Subject:X16.minus.17) + Harmony:musician
        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
random1 15 8362.9 8447.5 -4166.5   8332.9
music1  18 8358.7 8460.2 -4161.3   8322.7 10.251      3    0.01655 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> music2=update(music1,.~.+Voice:musician)
> anova(music1,music2)

Data: newrating
Models:
music1: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
music1:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
music1:     (1 | Subject:X16.minus.17) + Harmony:musician
music2: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
music2:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
music2:     (1 | Subject:X16.minus.17) + Harmony:musician + Voice:musician
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
music1 18 8358.7 8460.2 -4161.3   8322.7
music2 20 8360.8 8473.6 -4160.4   8320.8 1.8408      2     0.3984
```

```
> music3=update(music1,.~.+Instrument:musician)
> anova(music1,music3)

Data: newrating
Models:
music1: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
music1:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
music1:     (1 | Subject:X16.minus.17) + Harmony:musician
music3: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
music3:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
music3:     (1 | Subject:X16.minus.17) + Harmony:musician + Instrument:musician
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
music1 18 8358.7 8460.2 -4161.3   8322.7
music3 20 8358.9 8471.7 -4159.5   8318.9 3.7441      2     0.1538
```

```
> music4=update(music1,.~.+X16.minus.17:musician)
> anova(music1,music4)

Data: newrating
Models:
music1: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
music1:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
music1:     (1 | Subject:X16.minus.17) + Harmony:musician
music4: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
music4:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
music4:     (1 | Subject:X16.minus.17) + Harmony:musician + X16.minus.17:musician
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
music1 18 8358.7 8460.2 -4161.3   8322.7
music4 19 8360.1 8467.2 -4161.1   8322.1 0.5723      1     0.4494
```

```
> music5=update(music1,.~.+ClsListen1:musician)
> anova(music1,music5)
```

```
Data: newrating
Models:
music1: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
music1:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
music1:     (1 | Subject:X16.minus.17) + Harmony:musician
music5: Classical ~ Instrument + Voice + Harmony + X16.minus.17 + ClsListen +
music5:     musician + (1 | Subject:Instrument) + (1 | Subject:Harmony) +
music5:     (1 | Subject:X16.minus.17) + Harmony:musician + musician:ClsListen1
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
music1 18 8358.7 8460.2 -4161.3   8322.7
music5 19 8358.6 8465.7 -4160.3   8320.6 2.0872      1     0.1485

>
```

After adding interactions into the model, in results of anova we found that the interaction between Harmony and musician variables are significant since it has a small p-value equals to 0.01655. And the interactions all have non-significant results. So we say that there is an interaction between musician variable and predictor Harmony.

4.

(a)

```
> lmm1=lm(Popular~Harmony+Voice+Instrument,data=rating)
> summary(lmm1)

Call:
lm(formula = Popular ~ Harmony + Voice + Instrument, data = rating)

Residuals:
    Min      1Q  Median      3Q     Max
-6.7218 -1.7026  0.2008  1.4691 13.2248

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       6.58263    0.12761  51.583   <2e-16 ***
HarmonyI-V-IV    -0.02405    0.12782  -0.188   0.8508
HarmonyI-V-VI    -0.26829    0.12782  -2.099   0.0359 *
HarmonyIV-I-V    -0.18564    0.12772  -1.454   0.1462
Voicepar3rd       0.16859    0.11075   1.522   0.1281
Voicepar5th       0.16326    0.11068   1.475   0.1403
Instrumentpiano  -0.95200    0.11102  -8.575   <2e-16 ***
Instrumentstring -2.61173    0.11035 -23.667   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.257 on 2485 degrees of freedom
  (27 observations deleted due to missingness)
Multiple R-squared:  0.1901,	Adjusted R-squared:  0.1878
F-statistic: 83.32 on 7 and 2485 DF,  p-value: < 2.2e-16

> lmm2=lm(Popular~Voice+Instrument,data=rating)
> lmm3=lm(Popular~Harmony+Instrument,data=rating)
> lmm4=lm(Popular~Voice+Harmony,data=rating)
> anova(lmm1,lmm2)
```

```
Analysis of Variance Table

Model 1: Popular ~ Harmony + Voice + Instrument
Model 2: Popular ~ Voice + Instrument
  Res.Df   RSS Df Sum of Sq      F Pr(>F)
1   2485 12656
2   2488 12688 -3   -31.092 2.0349 0.1069
```

```
> anova(lmm1,lmm3)
```

```
Analysis of Variance Table

Model 1: Popular ~ Harmony + Voice + Instrument
Model 2: Popular ~ Harmony + Instrument
  Res.Df   RSS Df Sum of Sq      F Pr(>F)
1   2485 12656
2   2487 12672 -2   -15.263 1.4984 0.2237
```

```
> anova(lmm1,lmm4)
```

```
Analysis of Variance Table

Model 1: Popular ~ Harmony + Voice + Instrument
Model 2: Popular ~ Voice + Harmony
  Res.Df   RSS Df Sum of Sq      F    Pr(>F)
1   2485 12656
2   2487 15580 -2   -2923.9 287.05 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the linear model test we could find that the second level of Harmony is significant, and those two kinds of instrument are also significant in the model. So we used anova test to see whether those variables are significant or not. In the anova test, we found that only the model without instrument is significant. Therefore, the final model should only include instrument as a predictor variable.

```
> pop0=lm(Popular~Instrument,data=rating)
> pop1=lmer(Popular~Instrument+Voice+Harmony+(1|Subject),data=rating)
> model5=list(pop0,pop1)
> sapply(model5,AIC)

[1] 11142.27 10453.12

> sapply(model5,BIC)

[1] 11165.55 10511.34

> pop1.2=lmer(Popular~Instrument+Voice+(1|Subject),data=rating)
> pop1.3=lmer(Popular~Instrument+Harmony+(1|Subject),data=rating)
> pop1.4=lmer(Popular~Instrument+(1|Subject),data=rating)
> model6=list(pop1,pop1.2,pop1.3,pop1.4)
> sapply(model6,AIC)

[1] 10453.12 10447.49 10447.40 10441.77

> sapply(model6,BIC)

[1] 10511.34 10488.24 10493.97 10470.87
```

```
> pop2=lmer(Popular~Instrument+Voice+Harmony+(1|Subject:Instrument)+(1|Subject:Voice)+(1|Subject:Harmony
> model7=list(pop1.4,pop2)
> sapply(model7,AIC)

[1] 10441.77 10097.24

> sapply(model7,BIC)

[1] 10470.87 10167.09

> pop2.1=lmer(Popular~Instrument+Harmony+(1|Subject:Instrument)+(1|Subject:Voice)+(1|Subject:Harmony),da
> pop2.2=lmer(Popular~Instrument+Voice+(1|Subject:Instrument)+(1|Subject:Voice)+(1|Subject:Harmony),data
> pop2.3=lmer(Popular~Instrument+(1|Subject:Instrument)+(1|Subject:Voice)+(1|Subject:Harmony),data=ratin
> model8=list(pop2,pop2.1,pop2.2,pop2.3)
> sapply(model8,AIC)

[1] 10097.24 10091.75 10089.39 10083.91

> sapply(model8,BIC)

[1] 10167.09 10149.96 10141.78 10124.66

> pop3.1=lmer(Popular~Instrument+(1|Subject:Voice)+(1|Subject:Harmony),data=rating)
> pop3.2=lmer(Popular~Instrument+(1|Subject:Instrument)+(1|Subject:Harmony),data=rating)
> pop3.3=lmer(Popular~Instrument+(1|Subject:Instrument)+(1|Subject:Voice),data=rating)
> model9=list(pop2.3,pop3.1,pop3.2,pop3.3)
> sapply(model9,AIC)

[1] 10083.91 10579.86 10083.69 10173.57

> sapply(model9,BIC)

[1] 10124.66 10614.79 10118.61 10208.50

>
```

We repeated what did in problem 1. First we fitted the model using the conventional linear model. We found that only variable Instrument is significant in our model, so we omitted the other two variables. Then we included the random intercept for each participant into the model. When we compare them, we can find that the model has only Instrument and also with random intercept is better. So we used this model to continue discussion. Then We added three new random effects in the model. After checking AIC and BIC values for each model we fit, for instance, deleting one main effect or deleting one random intercept, we could found that the model with main effect Instrument with combination random effect of person/Instrument, person/Harmony has the smallest AIC and BIC values. So we should use that to be our best model.

voice and harmony have to be in the model because they are design variables...

(b)

```
> popnew=lmer(Popular~Instrument+Voice+Harmony+(1|Subject:Instrument)+(1|Subject:Harmony),data=newrating
> pop4=update(popnew,.~.+X16.minus.17,data=newrating)
> anova(popnew,pop4)

Data: newrating
Models:
popnew: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
popnew:     (1 | Subject:Harmony)
pop4: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop4:     (1 | Subject:Harmony) + X16.minus.17
```

```
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
popnew 11 8458.9 8520.9 -4218.5   8436.9
pop4   12 8456.8 8524.5 -4216.4   8432.8 4.0663      1    0.04375 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> pop5=update(pop4,.~.+ConsNotes1,data=newrating)
> anova(pop4,pop5)

Data: newrating
Models:
pop4: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop4:     (1 | Subject:Harmony) + X16.minus.17
pop5: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop5:     (1 | Subject:Harmony) + X16.minus.17 + ConsNotes1
     Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
pop4 12 8456.8 8524.5 -4216.4   8432.8
pop5 13 8458.5 8531.8 -4216.2   8432.5 0.353      1     0.5524

> pop6=update(pop4,.~.+ClsListen1,data=newrating)
> anova(pop4,pop6)

Data: newrating
Models:
pop4: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop4:     (1 | Subject:Harmony) + X16.minus.17
pop6: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop6:     (1 | Subject:Harmony) + X16.minus.17 + ClsListen1
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
pop4 12 8456.8 8524.5 -4216.4   8432.8
pop6 13 8458.6 8531.9 -4216.3   8432.6 0.2595      1     0.6105

> pop7=update(pop4,.~.+PianoPlay1,data=newrating)
> anova(pop4,pop7)

Data: newrating
Models:
pop4: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop4:     (1 | Subject:Harmony) + X16.minus.17
pop7: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop7:     (1 | Subject:Harmony) + X16.minus.17 + PianoPlay1
     Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
pop4 12 8456.8 8524.5 -4216.4   8432.8
pop7 13 8458.6 8531.9 -4216.3   8432.6   0.2      1     0.6548

> pop8=update(pop4,.~.+GuitarPlay1,data=newrating)
> anova(pop4,pop8)

Data: newrating
Models:
pop4: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop4:     (1 | Subject:Harmony) + X16.minus.17
pop8: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop8:     (1 | Subject:Harmony) + X16.minus.17 + GuitarPlay1
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
pop4 12 8456.8 8524.5 -4216.4   8432.8
```

```
pop8 13 8456.1 8529.4 -4215.1   8430.1 2.7219      1    0.09898 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> pop9=update(pop8,.~.+CollegeMusic1,data=newrating)
> anova(pop8,pop9)

Data: newrating
Models:
pop8: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop8:     (1 | Subject:Harmony) + X16.minus.17 + GuitarPlay1
pop9: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop9:     (1 | Subject:Harmony) + X16.minus.17 + GuitarPlay1 + CollegeMusic1
     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
pop8 13 8456.1 8529.4 -4215.1   8430.1
pop9 14 8457.6 8536.5 -4214.8   8429.6 0.5526      1     0.4572

> pop10=update(pop8,.~.+musician,data=newrating)
> anova(pop8,pop10)

Data: newrating
Models:
pop8: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop8:     (1 | Subject:Harmony) + X16.minus.17 + GuitarPlay1
pop10: Popular ~ Instrument + Voice + Harmony + (1 | Subject:Instrument) +
pop10:     (1 | Subject:Harmony) + X16.minus.17 + GuitarPlay1 + musician
      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
pop8  13 8456.1 8529.4 -4215.1   8430.1
pop10 14 8453.4 8532.3 -4212.7   8425.4 4.7012      1    0.03014 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

>
```

We added seven different possible new variables into our best model. They are same variables we added for classical rating.After testing by anova, we found out that adding new random effects for X16.minus.17, GuitarPlay and musician are significant. So we should include them into our model.

```
> popmodel=lm(Popular~Instrument+Harmony+Voice+X16.minus.17 + GuitarPlay1 + musician,data=newrating)
> popmodel0=lmer(Popular~Instrument+Harmony+Voice+X16.minus.17 + GuitarPlay1 + musician+(1|Subject:Harmo
> model10=list(popmodel,popmodel0)
> sapply(model10,AIC)

[1] 9252.051 8866.068

> sapply(model10,BIC)

[1] 9319.692 8939.345

> poprandom1=update(popmodel0,.~.+(1|Subject:Instrument))
> anova(popmodel0,poprandom1)

Data: newrating
Models:
popmodel0: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
popmodel0:     musician + (1 | Subject:Harmony)
poprandom1: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
```

```
poprandom1:     musician + (1 | Subject:Harmony) + (1 | Subject:Instrument)
           Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
popmodel0  13 8841.7 8914.9 -4407.8   8815.7
poprandom1 14 8453.4 8532.3 -4212.7   8425.4 390.24      1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> poprandom2=update(poprandom1,.~.+(1|Subject:X16.minus.17))
> anova(poprandom1,poprandom2)

Data: newrating
Models:
poprandom1: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
poprandom1:     musician + (1 | Subject:Harmony) + (1 | Subject:Instrument)
poprandom2: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
poprandom2:     musician + (1 | Subject:Harmony) + (1 | Subject:Instrument) +
poprandom2:     (1 | Subject:X16.minus.17)
           Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
poprandom1 14 8453.4 8532.3 -4212.7   8425.4
poprandom2 15 8427.9 8512.5 -4199.0   8397.9 27.481      1  1.587e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> poprandom3=update(poprandom2,.~.+(1|Subject:GuitarPlay1))
> anova(poprandom2,poprandom3)

Data: newrating
Models:
poprandom2: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
poprandom2:     musician + (1 | Subject:Harmony) + (1 | Subject:Instrument) +
poprandom2:     (1 | Subject:X16.minus.17)
poprandom3: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
poprandom3:     musician + (1 | Subject:Harmony) + (1 | Subject:Instrument) +
poprandom3:     (1 | Subject:X16.minus.17) + (1 | Subject:GuitarPlay1)
           Df    AIC    BIC logLik deviance Chisq Chi Df Pr(>Chisq)
poprandom2 15 8427.9 8512.5  -4199   8397.9
poprandom3 16 8429.9 8520.1  -4199   8397.9     0      1          1

> poprandom4=update(poprandom2,.~.+(1|Subject:musician))
> anova(poprandom2,poprandom4)

Data: newrating
Models:
poprandom2: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
poprandom2:     musician + (1 | Subject:Harmony) + (1 | Subject:Instrument) +
poprandom2:     (1 | Subject:X16.minus.17)
poprandom4: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
poprandom4:     musician + (1 | Subject:Harmony) + (1 | Subject:Instrument) +
poprandom4:     (1 | Subject:X16.minus.17) + (1 | Subject:musician)
           Df    AIC    BIC logLik deviance Chisq Chi Df Pr(>Chisq)
poprandom2 15 8427.9 8512.5  -4199   8397.9
poprandom4 16 8429.9 8520.1  -4199   8397.9     0      1          1

>
>
```

Firstly, we went back and checked random effects we got before. We found that both random effects person/Instrument,person/Harmony were significant.Then we added three new random effects into the model each of a time. After testing by anova, we found out that adding new random effects for X16.minus.17 in model was significant since it had a small p-value. And other two random effects were not significant. So our final model should include the random effect for Harmony, Instrument and X16.minus.17.

LRT not appropriate for testing random effects

```
> summary(poprandom4)

Linear mixed model fit by REML ['lmerMod']
Formula: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +      musician + (1 | Subj
   Data: newrating

REML criterion at convergence: 8418.859

Random effects:
 Groups               Name        Variance  Std.Dev.
 Subject:Harmony      (Intercept) 3.357e-01 0.579429
 Subject:Instrument   (Intercept) 1.078e+00 1.038271
 Subject:musician     (Intercept) 1.097e+00 1.047198
 Subject:X16.minus.17 (Intercept) 5.277e-05 0.007264
 Residual                         2.637e+00 1.623958
Number of obs: 2073, groups: Subject:Harmony, 232; Subject:Instrument, 174; Subject:musician, 58; Subjec

Fixed effects:
                 Estimate Std. Error t value
(Intercept)       6.44679    0.70240   9.178
Instrumentpiano  -1.00753    0.21195  -4.754
Instrumentstring -2.66689    0.21156 -12.606
HarmonyI-V-IV    -0.04464    0.14751  -0.303
HarmonyI-V-VI    -0.30941    0.14755  -2.097
HarmonyIV-I-V    -0.22608    0.14747  -1.533
Voicepar3rd       0.15112    0.08741   1.729
Voicepar5th       0.15208    0.08739   1.740
X16.minus.17      0.07243    0.05804   1.248
GuitarPlay1Low   -0.44258    0.50438  -0.877
musician          0.71719    0.46845   1.531

Correlation of Fixed Effects:
            (Intr) Instrmntp Instrmnts HI-V-I HI-V-V HIV-I- Vcpr3r Vcpr5t
Instrumntpn -0.150
Instrmntstr -0.151  0.499
HrmnyI-V-IV -0.105  0.000     0.000
HrmnyI-V-VI -0.105  0.000     0.000     0.500
HrmnyIV-I-V -0.105  0.000     0.000     0.500  0.500
Voicepar3rd -0.062  0.000     0.000     0.000  0.001  0.001
Voicepar5th -0.062  0.000     0.000    -0.001 -0.001 -0.001  0.501
X16.mins.17 -0.309 -0.001     0.000     0.000  0.000  0.000  0.000  0.000
GuitrPly1Lw -0.751 -0.001     0.000     0.000  0.000  0.000  0.000  0.000
musician    -0.683  0.000     0.000     0.000  0.000  0.000  0.000  0.000
            X16..1 GtrP1L
Instrumntpn
Instrmntstr
HrmnyI-V-IV
HrmnyI-V-VI
HrmnyIV-I-V
```

```
Voicepar3rd
Voicepar5th
X16.mins.17
GuitrPly1Lw  0.226
musician     0.061  0.181
```

In the final model, we can interpret the effect of each variable individually. When we have one more unit of piano instrument, the rating of popular will reduced by 1.0075 comparing to the comparison group. When we have one more unit of string instrument, the rating of popular will reduced by 2.6671 comparing to the comparison group.When we have one more unit of 3rd par voice, the rating of popular will increase by 0.1414 comparing to the comparison group.If the level of HarmonyI-V-IV increase one unit, the rating of popular will be decreased by 0.0446 comparing to the comparison group.If the level of HarmonyI-V-VI increase one unit, the rating of popular will be decreased by 0.3094 comparing to the comparison group.If the level of HarmonyIV-I-V increase one unit, the rating of popular will be decreased by 0.3094comparing to the comparison group.If the level of HarmonyIV-I-V increase one unit, the rating of popular will be decreased by 0.2260 comparing to the comparison group.

If the measure of listener's ability to distinguish classical or popular music increase one unit, the rating of popular music will be reduced by 0.0724. If the number of rarely playing guitar increase one unit, the rating of popular music will be reduced by 0.4427. And if the number of musicians increase one unit, the rating of popular will be decreased by 0.7174.

(c)

```
> popmusic1=update(poprandom4,.~.+Instrument:musician)
> anova(poprandom4,popmusic1)

Data: newrating
Models:
poprandom4: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
poprandom4:     musician + (1 | Subject:Harmony) + (1 | Subject:Instrument) +
poprandom4:     (1 | Subject:X16.minus.17) + (1 | Subject:musician)
popmusic1: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
popmusic1:     musician + (1 | Subject:Harmony) + (1 | Subject:Instrument) +
popmusic1:     (1 | Subject:X16.minus.17) + (1 | Subject:musician) + Instrument:musician
           Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
poprandom4 16 8429.9 8520.1 -4199.0   8397.9
popmusic1  18 8430.6 8532.0 -4197.3   8394.6 3.3711      2     0.1853
```

I don't see much motivation for the new random effects here.

Again, p-values and simple LRTs are not approporiate for random effects

```
> popmusic2=update(poprandom4,.~.+X16.minus.17:musician)
> anova(poprandom4,popmusic2)

Data: newrating
Models:
poprandom4: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
poprandom4:     musician + (1 | Subject:Harmony) + (1 | Subject:Instrument) +
poprandom4:     (1 | Subject:X16.minus.17) + (1 | Subject:musician)
popmusic2: Popular ~ Instrument + Harmony + Voice + X16.minus.17 + GuitarPlay1 +
popmusic2:     musician + (1 | Subject:Harmony) + (1 | Subject:Instrument) +
popmusic2:     (1 | Subject:X16.minus.17) + (1 | Subject:musician) + X16.minus.17:musician
           Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
poprandom4 16 8429.9 8520.1 -4199.0   8397.9
popmusic2  17 8431.9 8527.7 -4198.9   8397.9 0.0666      1     0.7964

>
```

When we were using anova test to test the significance of adding interactions for musician variable and other predictors, we can see that the result was all interactions were not significant because they are have large p-values. Thus, we don't need to include any interactions into our model.

5.

This is a study about the influence of instrument, harmonic motion, and voice leading on listeners' identification of music as "classical" or "popular".In my statistical analysis, I found that the influence of instrument, harmonic motion and voice leading have different influence on different kind of music. For classical ~~music,~~ ratings as I showed in my answer to part 1a, all three main experimental factors are significant, which means that we should consider all their effects for classical rating. And if include random intercept for each participant, still all three main effects are significant as shown in part 1b. After that, when I considered one step further, I thought there might be some personal biases, such as person A is more inclined to rate everything as classical, and person B is more inclined to rate everything as popular. So on this aspect, I included three new random effects, which are combinations of person/instrument,person/voice and person/harmony. In the result of anova test, all three new random effects were taken into account. So three main experimental factors and their corresponding random effects are all counted as factors that could influence the rating of Classical music. For popular ~~music~~, as I did in part 4a,when I was using the conventional linear model, I ratings found that only the variable instrument is significant. And after adding random effects into my model, I found that the random effect for person/Voice is not significant. So for popular music, only Instrument as a main effect has influence on the rating.But when I was trying to add new variables into the model, I still included all three main experiment factors in my model to continue analyzing.

For this repeated measures model, I would not say that this is a standard repeated measure since we included other variance components. For Classical, besides those three main random effects, I included one more variance component which is the combination of person/X16.minus.17. For Popular music, I included one more variance components based on old random effects, it's the combination of person/X16.minus.17. Before adding new variance components, I always go back and check those three main random effects first.For classical music, I included all three main random effects in part 1c, and after adding new fixed effects, I still need to include all of them. On the other hand, for popular music, I included random effects for Harmony and Instrument before and also after adding new fixed effects. Since the anova result for adding new variance components is significant, so I would say that adding new variance components is valuable.

In the model for Classical and the model for Popular music, I studied same individual covariates other than those three main experimental factors. The reason is that I chose variables that can influence both classical and popular music.After getting best models for classical and popular, it is reasonable that two models resulted in adding different new fixed effects since variables can have influence on one kind of music but not the other one. It is noticeable that those common new variables in different model have same level of influence on the rating. For instance, both model have variables X16.minus.17 and musician. These two variables both have positive influences on classical rating and popular rating. So this can prove that my choice of using these variables for two different kinds of music analysis is correct, they do have influences on both type of music. Moreover, for popular music, we have one more new covaraite, which is guitarplay, and this covaraite is not in the model for classical music. So we can conclude that some variables have influence on one of the music kind, not both. And this is very reasonable, according to my result from part 4c, we could say that people who rarely play guitar tend to listen to popular music more nowadays.

4: 14/20
5: 18/20

32/40