763 HW5

Dai Chen

1 a 9/9
b 9/9
c 9/9

2 a 9/9
b 9/9
c 9/9

3 9/9

4 a 9/9
b 9/9
c 9/9

5 10/10

Total 100/100

Problem 1: The three main experimental factors

(a)

Fit a linear model on the three main experimental factors.

```
lm1 <- lm(Classical ~ Instrument + Harmony + Voice, data = rat.c)
summary(lm1)
```

```
Call:
lm(formula = Classical ~ Instrument + Harmony + Voice, data = rat.c)

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
```

Comparing to guitar, holding all other variables constant, piano increases the rating of about 1.37 unit; string increases the rating of about 3.13 units. Comparing to I-VI-V, holding all other variables constant, Harmony I-V-IV decreases the rating of about 0.0031 unit; I-V-VI increases the rating of about 0.769 unit; IV-I-V increases the rating of about 0.05 unit. Comparing to contrary, parallel 3rd decreases the rating of about 0.412 unit; parallel 5th decreases the rating of about 0.371 unit.

To determine if Instrument, Harmony or Leading Voices are significant, we fit models without each of them and compare with the previous model.

```
lm1.1 <- lm(Classical ~ Harmony + Voice, data = rat.c)
lm1.2 <- lm(Classical ~ Instrument + Voice, data = rat.c)
lm1.3 <- lm(Classical ~ Instrument + Harmony, data = rat.c)
anova(lm1, 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(lm1, 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(lm1, 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 ***
```

9

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the analysis of variance table, we see that Instrument, Harmony and Voice are all important factors. By partial F-test, the results are significant between the full model and the reduced models, as p-values for these three < 0.05.

(b)

i.

Hierarchical linear model:

$$y_i = \beta_0 + \beta_{I[i]} + \beta_{H[i]} + \beta_{V[i]} + \eta_{j[i]} + \epsilon_i$$
$$\epsilon_i \overset{iid}{\sim} N(0, \sigma_\epsilon^2)$$
$$\eta_j = \alpha_0 + \lambda_j$$

9

$$\lambda_j \overset{iid}{\sim} N(0, \tau_\eta^2)$$
$$for\ i = 1, 2, ..., 2493, j = 1, 2, ..., 70$$

ii.

Fit repeated measures model for each of the 70 participants.

```
library(arm)
lmer1 <- lmer(Classical ~ Instrument + Harmony + Voice + (1|Subject),
              data = rat.c, REML=FALSE)
c(AIC(lmer1), AIC(lm1))
```

```
[1] 10468.86 11230.45
```

```
c(BIC(lmer1), BIC(lm1))
```

```
[1] 10527.07 11282.84
```

Method 1: According to the rule of thumb, the decrease in both AIC and BIC for the random intercept model are greater than 3, thus the model with random intercept is significantly better.

Method 2: From the anova we see that the lm model does not fit as well as the lmer model, as the p-value is smaller than 0.05. The random intercept term is significant.

```
library(RLRsim)
exactRLRT(lmer1)
```

```


    simulated finite sample distribution of RLRT.

    (p-value based on 10000 simulated values)

data:
RLRT = 763.37, p-value < 2.2e-16
```

The p-value of the Restricted Likelihood Ratio Tests is less than 0.05, which means that the result is significant between the model with the random intercept and the model without random intercept.

The two methods agree with each other, that we should keep the random intercept for a better fit.

iii.

```
display(lmer1)
```

```
lmer(formula = Classical ~ Instrument + Harmony + Voice + (1 |
    Subject), data = rat.c, REML = FALSE)
                 coef.est coef.se
(Intercept)       4.34     0.19
Instrumentpiano   1.38     0.09
Instrumentstring  3.13     0.09
HarmonyI-V-IV    -0.03     0.11
HarmonyI-V-VI     0.77     0.11
HarmonyIV-I-V     0.05     0.11
```

```
Voicepar3rd      -0.42     0.09
Voicepar5th      -0.37     0.09

Error terms:
 Groups   Name        Std.Dev.
 Subject  (Intercept) 1.29
 Residual             1.89
---
number of obs: 2493, groups: Subject, 70
AIC = 10468.9, DIC = 10448.9
deviance = 10448.9
```

We see that comparing to guitar, holding all other variables constant, piano increases the rating of about 1.38 unit; string increases the rating of about 3.13 units. Comparing to I-VI-V, holding all other variables constant, Harmony I-V-IV decreases the rating of about 0.03 unit; I-V-VI increases the rating of about 0.77 unit; IV-I-V increases the rating of about 0.05 unit. Comparing to contrary, parallel 3rd decreases the rating
9 of about 0.42 unit; parallel 5th decreases the rating of about 0.37 unit. Comparing to the lm model, the lmer model's fixed effect is not much different from the coefficient estimation.

Now use ANOVA to check if any of the main design effects is not needed for the repeated measurement method regression.

```r
lmer1.1 <- lmer(Classical ~ Harmony + Voice + (1|Subject),
                data = rat.c, REML=F)
lmer1.2 <- lmer(Classical ~ Instrument + Voice + (1|Subject),
                data = rat.c, REML=F)
lmer1.3 <- lmer(Classical ~ Instrument + Harmony + (1|Subject),
                data = rat.c, REML=F)
anova(lmer1, lmer1.1)
```

```
Data: rat.c
Models:
lmer1.1: Classical ~ Harmony + Voice + (1 | Subject)
lmer1: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer1.1  8 11408 11455 -5696.2    11392
lmer1   10 10469 10527 -5224.4    10449 943.59      2  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(lmer1, lmer1.2)
```

```
Data: rat.c
Models:
lmer1.2: Classical ~ Instrument + Voice + (1 | Subject)
lmer1: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer1.2  7 10539 10580 -5262.4    10525
lmer1   10 10469 10527 -5224.4    10449 75.931      3  2.288e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(lmer1, lmer1.3)
```

```
Data: rat.c
Models:
lmer1.3: Classical ~ Instrument + Harmony + (1 | Subject)
lmer1: Classical ~ Instrument + Harmony + Voice + (1 | Subject)
        Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmer1.3  8 10489 10536 -5236.6    10473
lmer1   10 10469 10527 -5224.4    10449 24.24      2   5.45e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the analysis of variance table, we see that Instrument, Harmony and Voice are all important factors. By partial F-test, the results are significant between the full model and the reduced models, as p-values for these three < 0.05.

I checked the diagnostics plots and there are nothing particular suspicious.

(c)

i.

Fit the model with all three random effect.

```r
lmer3 <- lmer(Classical ~ Instrument + Harmony + Voice + (1|Subject:Instrument ) +
                (1|Subject:Harmony) +  (1|Subject:Voice), data=rat.c, REML = F)
```

9

Comparing this model with models in 1a and 1b by AIC.

```r
AIC(lmer3, lm1)
```

```
      df      AIC
lmer3 12 10057.53
lm1    9 11230.45
```

```r
AIC(lmer3, lmer1)
```

```
      df      AIC
lmer3 12 10057.53
lmer1 10 10468.86
```

```r
BIC(lmer3, lm1)
```

```
      df      BIC
lmer3 12 10127.38
lm1    9 11282.84
```

```r
BIC(lmer3, lmer1)
```

```
      df      BIC
lmer3 12 10127.38
lmer1 10 10527.07
```

We see that the model in 1(c) is better than the model in 1(b) and 1(a), because the AIC and BIC is the smallest and the changes are greater than 3, thus the model with the three random effects fits better.

ii.

Similar to 1a and 1b, we compare the fit of models with each of the design factors and a model without each of the design factors.

```r
# Instrument
lmer3.1 <- lmer(Classical ~  Harmony + Voice +
                  (1|Subject:Instrument ) + (1|Subject:Harmony) +
                  (1|Subject:Voice), data=rat.c, REML = F)
# Harmony
lmer3.2 <- lmer(Classical ~ Instrument + Voice +
                  (1|Subject:Instrument ) + (1|Subject:Harmony) +
                  (1|Subject:Voice), data=rat.c, REML = F)
# Voice
lmer3.3 <- lmer(Classical ~ Instrument + Harmony +
                  (1|Subject:Instrument ) + (1|Subject:Harmony) +
                  (1|Subject:Voice), data=rat.c, REML = F)
```

```r
anova(lmer3, lmer3.1)
```

```
Data: rat.c
Models:
lmer3.1: Classical ~ Harmony + Voice + (1 | Subject:Instrument) + (1 |
lmer3.1:     Subject:Harmony) + (1 | Subject:Voice)
lmer3: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer3:     (1 | Subject:Harmony) + (1 | Subject:Voice)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer3.1 10 10160 10219 -5070.2    10140
lmer3   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(lmer3, lmer3.2)
```

```
Data: rat.c
Models:
lmer3.2: Classical ~ Instrument + Voice + (1 | Subject:Instrument) + (1 |
lmer3.2:     Subject:Harmony) + (1 | Subject:Voice)
lmer3: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer3:     (1 | Subject:Harmony) + (1 | Subject:Voice)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer3.2  9 10090 10143 -5036.3    10072
lmer3   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(lmer3, lmer3.3)
```

```
Data: rat.c
Models:
lmer3.3: Classical ~ Instrument + Harmony + (1 | Subject:Instrument) +
lmer3.3:     (1 | Subject:Harmony) + (1 | Subject:Voice)
lmer3: Classical ~ Instrument + Harmony + Voice + (1 | Subject:Instrument) +
lmer3:     (1 | Subject:Harmony) + (1 | Subject:Voice)
        Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer3.3 10 10081 10140 -5030.6    10061
lmer3   12 10058 10127 -5016.8    10034 27.753      2  9.409e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the analysis of variance table, we see that Instrument, Harmony and Voice are all important factors. By partial F-test, the results are significant between the full model and the reduced models, as p-values for these three < 0.05.

Comment on the size of variance component:

```r
summary(lmer3)$varcor
```

```
 Groups             Name        Std.Dev.
 Subject:Harmony    (Intercept) 0.65792
 Subject:Voice      (Intercept) 0.15726
 Subject:Instrument (Intercept) 1.47285
 Residual                       1.56116
```

We see that the error term of the Subject:Instrument has the biggest variance, the error term of Subject: Voice has the smallest variance.

The variance of the residual is a lot larger than the Subject:Harmony and the Subject:Voice, and is a little larger than the Subject:Instrument.

It shows that the fixed effect coefficients are capturing the mean pretty well. Even though the variance components are statistically significant, the magnitude is comparatively small. The difference between each individual pretty small compared to the rating each individual give to different musical excerpts.

iii.

Hierarchical

$$y_i = \alpha^I_{j[i]I[i]} + \alpha^H_{j[i]j:H[i]} + \alpha^V_{j[i]j:V[i]} + \beta_0 + \beta_{I[i]} + \beta_{H[i]} + \beta_{V[i]} + \epsilon_i$$
$$\alpha^I_{j[i]I[i]} = \alpha^I_0 + \eta^I_{jI}$$
$$\alpha^H_{j[i]H[i]} = \alpha^H_0 + \eta^H_{jH}$$
$$\alpha^V_{j[i]V[i]} = \alpha^V_0 + \eta^V_{jV}$$

great!

$$\epsilon_i \sim N(0, \sigma^2_\epsilon)$$
$$\eta^I_{jI} \sim N(0, \tau^2_{\eta I})$$
$$\eta^H_{jH} \sim N(0, \tau^2_{\eta H})$$
$$\eta^V_{jV} \sim N(0, \tau^2_{\eta V})$$
$$for\ i = 1, 2, \ldots, 2493, and\ j = 1, 2, \ldots, 70$$

Problem 2: Individual covariates

(a)

word missing?

The best model from the previous question is the Because there are many NAs in "how proficient are you at your first musical instrument" and "how proficient are you in your second musical instrument", we decide to code NA as 0, and code people who play instrument as 1 for the first and second instrument. We also removed the "first 12" variable which is not needed for this assignment and the "X1990s2000s.minus.1960s1970s" and "Instr.minus.Notes" are not included because of collinearity.

```
rat$CollegeMusic <- as.factor(rat$CollegeMusic)
rat$APTheory <- as.factor(rat$APTheory)
rat$X1stInstr <- ifelse(is.na(rat$X1stInstr), 0,1)
rat$X2ndInstr <- ifelse(is.na(rat$X2ndInstr), 0,1)
```

Now we use backward selection to select the covariates.

```
library(LMERConvenienceFunctions)
lmer2a <- lmer(Classical ~ Harmony + Instrument + Voice + Selfdeclare +
                 OMSI + X16.minus.17 + ConsInstr + ConsNotes + PachListen +
                 ClsListen + KnowRob + KnowAxis + X1990s2000s + CollegeMusic +
                 NoClass + APTheory + Composing + PianoPlay + GuitarPlay +
                 X1stInstr + X2ndInstr + (1|Subject:Instrument) +
                 (1|Subject:Harmony) + (1|Subject:Voice), data = rat)
# bfFixefLMER_F.fnc(lmer2a, method = c("AIC"))
lmer2a.select <- lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes + PachListen +
    KnowRob + KnowAxis + X1990s2000s + NoClass + APTheory + PianoPlay +
    (1 | Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice), data=rat)
AIC(lmer2a, lmer2a.select)
```

9

```
              df      AIC
lmer2a        30 6298.681
lmer2a.select 20 6268.579
```

We see that the difference between the full model and the selected model is significant according to rule of thumb, because the AIC dropped about 30.

(b)

```
### mA
### m0
### m.test

# Test for Instrument
m.1 <- lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes +
              PachListen + KnowRob + KnowAxis + X1990s2000s + NoClass +
              APTheory + PianoPlay + (1 | Subject:Instrument) , data = rat)
m.0.1 <- lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes +
                PachListen + KnowRob + KnowAxis + X1990s2000s + NoClass +
                APTheory + PianoPlay + (1| Subject:Harmony) + (1| Subject:Voice),
              data = rat)
exactRLRT(m.1, mA = lmer2a.select , m0 = m.0.1)
```

```
    simulated finite sample distribution of RLRT.

    (p-value based on 10000 simulated values)

data:
RLRT = 343.9, p-value < 2.2e-16
```

```r
# Test for Harmony
m.2 <- lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes +
              PachListen + KnowRob + KnowAxis + X1990s2000s + NoClass +
              APTheory + PianoPlay + (1| Subject:Harmony) , data = rat)
m.0.2 <- lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes +
                PachListen + KnowRob + KnowAxis + X1990s2000s + NoClass +
                APTheory + PianoPlay+ (1| Subject:Instrument) + (1| Subject:Voice),
              data = rat)
exactRLRT(m.2, mA = lmer2a.select , m0 = m.0.2)
```

```
    simulated finite sample distribution of RLRT.

    (p-value based on 10000 simulated values)

data:
RLRT = 63.955, p-value < 2.2e-16
```

```r
# Test for Voice
m.3 <- lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes +
              PachListen + KnowRob + KnowAxis + X1990s2000s + NoClass +
              APTheory + PianoPlay + (1| Subject:Voice) , data = rat)
m.0.3 <- lmer(Classical ~ Harmony + Instrument + Voice + ConsNotes +
                PachListen + KnowRob + KnowAxis + X1990s2000s + NoClass +
                APTheory + PianoPlay+ (1|Subject:Instrument) + (1| Subject:Harmony),
              data = rat)
exactRLRT(m.3, mA = lmer2a.select , m0 = m.0.3)
```

```
    simulated finite sample distribution of RLRT.

    (p-value based on 10000 simulated values)

data:
RLRT = 1.7802, p-value = 0.0813
```

According to the Restricted Likelihood Ratio Tests, we found that the p-value for the random effect of **Subject:Voice became not statistically significant** (p-value=0.08). Testing the other random effect, the other random effects should be kept as the p-value is smaller than 0.05.

```r
AIC(lmer2a.select, m.0.3)
```

```
              df      AIC
lmer2a.select 20 6268.579
m.0.3         19 6268.359
```

9

```r
BIC(lmer2a.select, m.0.3)
```

```
              df      BIC
lmer2a.select 20 6375.383
m.0.3         19 6369.823
```

The AIC and BIC both favors the model without the Subject:Voice random effect. According to the rule of thumb, the differences between these two models are significant because the BIC decreases 5.56 if we take the random effect of the Voice:Subject interaction out.

(c)

```r
summary(m.0.3)$coef
```

```
                    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
```

9 Comparing to guitar, holding all other variables constant, piano increases the rating of about 1.65 unit; string increases the rating of about 3.59 units. Comparing to I-VI-V, holding all other variables constant, Harmony I-V-IV decreases the rating of about 0.0045 unit; I-V-VI increases the rating of about 0.85 unit; IV-I-V increases the rating of about 0.060 unit. Comparing to contrary, parallel 3rd decreases the rating of about 0.403 unit; parallel 5th decreases the rating of about 0.30 unit.

With one unit of increase in the “how much did you concentrate on the notes while listening”, holding all other variables constant, the rating on Classical decreases about 0.18. With one unit of increase in “have you heard rob paravonian’s Pachelbel rant”, the rating increases about 0.085, holding other variables constant. With one unit of increase in “have you heard Axis of Evil’s comedy bit”, the rating increases about 0.081; With one unit of increase in “how much do you listen to pop and rock from the 90s and 20s”, the rating increases about 0.189, holding other variables constant; With one more music classes taken, the rating decreases about 0.15; Comparing to people who did not take AP music theory class in high school, people who took rated classical 0.63 higher. With one unit of increase in “do you play piano”, the rating increases about 0.31.

Problem 3: Musicians vs. Non-musicians

First dichotomize the Self-declare (“are you a musician”) variable so that about half of the participants are categorized as self-declared musicians. From the summary and boxplot we choose 2 to be the cutoff,i.e. if the

great explanation

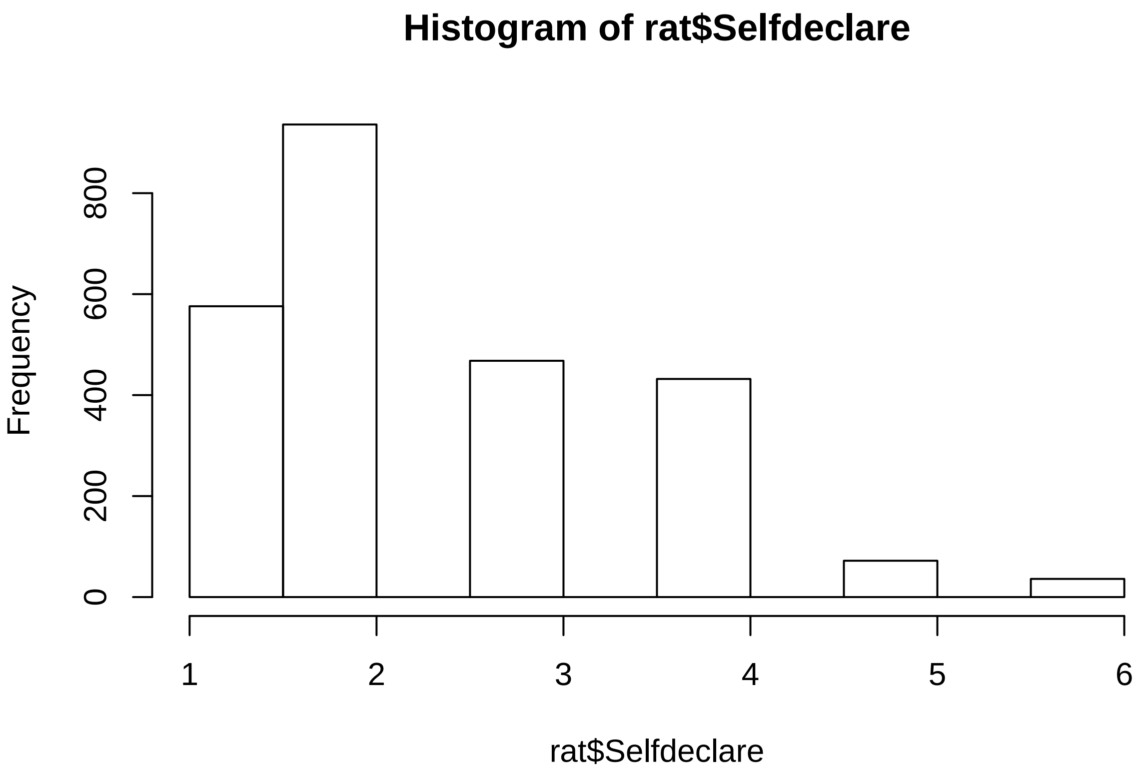
participant chose 1 or 2 on the "Are you are musician?", then this person is identified as "non-musician" and is coded as 0; otherwise, the person is identified as "musician" and is coded as 1. There are 1008 self-identified musicians and 1512 non-musician according to this standard.

There aren't that many musicians in the data set, so you must mean something else...

```
summary(rat$Selfdeclare)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  1.000   2.000   2.000   2.443   3.000   6.000 
```

```
hist(rat$Selfdeclare)
```




```
rat$Musician <- ifelse((rat$Selfdeclare==1|rat$Selfdeclare==2), 0, 1)
c(sum(rat$Musician), sum(rat$Musician==0))
```

```
[1] 1008 1512
```

In order to examine the relationship between dichotomized musician variable and other predictors in the model, we add interaction terms of Musician and other covariates to the previous model. Then we perform backward selection to this lmer model to see if any interaction term is in the model.

```
m.new.3 <- lmer(Classical ~ Harmony*Musician + Instrument*Musician +
                  Voice*Musician + ConsNotes*Musician + PachListen*Musician +
                  KnowRob*Musician + KnowAxis*Musician + X1990s2000s*Musician +
                  NoClass*Musician + APTheory*Musician + PianoPlay*Musician +
                  (1|Subject:Instrument) + (1| Subject:Harmony), data = rat)

# bfFixefLMER_F.fnc(m.new.3, method = c("AIC","BIC"))
```

```r
# The selected model is
m.new.select <- lmer(Classical ~ Harmony + Musician + Instrument +
                       Voice + ConsNotes +  PachListen + KnowRob +
                       KnowAxis + X1990s2000s + NoClass +
                       APTheory + PianoPlay + (1 | Subject:Instrument) +
                       (1 |Subject:Harmony) +  Harmony:Musician +
                       Musician:PianoPlay, data=rat)

AIC(m.new.select, m.new.3)
```

```
             df      AIC
m.new.select 24 6253.617
m.new.3      35 6264.794
```

```r
BIC(m.new.select, m.new.3)
```

```
             df      BIC
m.new.select 24 6381.782
m.new.3      35 6451.700
```

```r
m.new.select.piano <- lmer(Classical ~ Harmony + Musician + Instrument +
                       Voice + ConsNotes +  PachListen + KnowRob +
                       KnowAxis + X1990s2000s + NoClass +
                       APTheory + PianoPlay + (1 | Subject:Instrument) +
                       (1 |Subject:Harmony) +  Harmony:Musician, data=rat)

m.new.select.harmony <- lmer(Classical ~ Harmony + Musician + Instrument +
                       Voice + ConsNotes +  PachListen + KnowRob +
                       KnowAxis + X1990s2000s + NoClass +
                       APTheory + PianoPlay + (1 | Subject:Instrument) +
                       (1 |Subject:Harmony) + Musician:PianoPlay, data=rat)

# anova(m.new.select, m.new.select.piano)
# anova(m.new.select, m.new.select.harmony)
```

We see that the interaction for Harmony:Musician and Musician:PianoPlay is left in the model. These are the two most significant interaction that we want to consider. Comparing the AIC and BIC we see that both AIC and BIC favors our selected model, which is

9

```r
m.new.select <- lmer(Classical ~ Harmony + Musician + Instrument +
                       Voice + ConsNotes +  PachListen + KnowRob +
                       KnowAxis + X1990s2000s + NoClass +
                       APTheory + PianoPlay + (1 | Subject:Instrument) +
                       (1 |Subject:Harmony) +  Harmony:Musician +
                       Musician:PianoPlay, data=rat)
```

Comparing models without these two interaction terms, and the partial F-test have p-values less than 0.01, showing that the two interaction terms are very statistically significant.

The coefficient for the musician vs piano play is -0.606, which means that being a musician weakens the effect of piano play on Classical rating. The coefficient for HarmonyI-V-IV:Musician is -0.03, HarmonyI-V-VI:Musician is 1.26, HarmonyIV-I-V:Musician is 0.12. It means that musician rating Harmony I-V-VI higher on the classical rating. (Which is is the beginning progression for Pachelbel's Canon!)

Problem 4: Classical vs. Popular

(a)

i. examine only the main effect

Fit a linear model on the three main experimental factors. To determine if Instrument, Harmony or Leading Voices are significant, we fit models without each of them and compare with the previous model.

```
lm.pop <- lm(Popular ~ Instrument + Harmony + Voice, data = rat)
```

```
Analysis of Variance Table

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

Analysis of Variance Table

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

Analysis of Variance Table

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

From the analysis of variance table, we see that only Instrument is significant for the rating of Popular. Compared to guitar, piano decreases the rating of popular by -0.95; string decreases the rating of popular by 2.6, holding all other variables constant.

ii. repeated measurement method

Fit repeated measures model for each of the 70 participants. Comparing the AIC and BIC We see that the lmer model is significantly better. The restricted likelihood ratio test also shows that the difference is significant between the model with and without random intercept of (1|Subject). This means that the difference between subject is significant. However notice that the variance component is small compared to the residual variance.

```
lmer.pop.1 <- lmer(Popular ~ Instrument + Harmony + Voice + (1|Subject),
              data = rat, REML=FALSE)
```

```
           df      AIC
lmer.pop.1 10 10430.30
lm.pop.1    7 11657.31

           df      BIC
lmer.pop.1 10 10488.51
lm.pop.1    7 11698.06


    simulated finite sample distribution of RLRT.

    (p-value based on 10000 simulated values)

data:
RLRT = 714.74, p-value < 2.2e-16
```

Test the three main effect on this model, Instrument and Harmony are significant. Leading Voice is not significant.

```
Data: rat
Models:
lmer1.1.pop: Popular ~ Harmony + Voice + (1 | Subject)
lmer.pop.1: Popular ~ Instrument + Harmony + Voice + (1 | Subject)
            Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer1.1.pop  8 11138 11184 -5560.8    11122
lmer.pop.1  10 10430 10488 -5205.1    10410 711.31      2  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Data: rat
Models:
lmer1.2.pop: Popular ~ Instrument + Voice + (1 | Subject)
lmer.pop.1: Popular ~ Instrument + Harmony + Voice + (1 | Subject)
            Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
lmer1.2.pop  7 10433 10474 -5209.7    10419
lmer.pop.1  10 10430 10488 -5205.1    10410 9.0032      3    0.02925 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Data: rat
Models:
lmer1.3.pop: Popular ~ Instrument + Harmony + (1 | Subject)
lmer.pop.1: Popular ~ Instrument + Harmony + Voice + (1 | Subject)
            Df   AIC   BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
lmer1.3.pop  8 10431 10477 -5207.4    10415
lmer.pop.1  10 10430 10488 -5205.1    10410 4.429      2     0.1092
```

iii. the interactions as random effects

```r
lmer3.pop <- lmer(Popular ~ Instrument + Harmony + Voice +
                    (1|Subject:Instrument ) + (1|Subject:Harmony) +
                    (1|Subject:Voice), data=rat, REML = F)
```

Comparing this model with models in i and ii by AIC. The model includes interactions as random effects is the best.

```r
# the best model in i
lm.pop.ins <- lm(Popular ~ Instrument, data = rat)
# the best model in ii
lmer.pop.ins <- lmer(Popular ~ Instrument + Harmony + (1|Subject),
                     data = rat, REML=F)

AIC(lmer3.pop, lm.pop.ins, lmer.pop.ins)
```

```
             df      AIC
lmer3.pop    12 10078.97
lm.pop.ins    4 11142.27
lmer.pop.ins  8 10430.73
```

```r
BIC(lmer3.pop, lm.pop.ins, lmer.pop.ins)
```

```
             df      BIC
lmer3.pop    12 10148.82
lm.pop.ins    4 11165.55
lmer.pop.ins  8 10477.30
```

In conclusion, the model we chose is the model with all three main effcts and three interactions of design factors interact with the subjects.

```r
summary(lmer3.pop)$coef
```

9

```
                    Estimate Std. Error     t value
(Intercept)       6.57990842 0.20571275  31.9859053
Instrumentpiano  -0.94902658 0.25001902  -3.7958175
Instrumentstring -2.60591003 0.24971525 -10.4355262
HarmonyI-V-IV    -0.02557523 0.13958034  -0.1832295
HarmonyI-V-VI    -0.27150724 0.13956500  -1.9453820
HarmonyIV-I-V    -0.18544407 0.13950653  -1.3292859
Voicepar3rd       0.16390541 0.08263730   1.9834314
Voicepar5th       0.16206576 0.08257373   1.9626793
```

Comparing to guitar, holding all other variables constant, piano decreases the rating of about 0.94 unit; string decreases the rating of about 2.61 units. Comparing to I-VI-V, holding all other variables constant, Harmony I-V-IV decreases the rating of about 0.0256 unit; I-V-VI decreases the rating of about 0.27 unit; IV-I-V decreases the rating of about 0.185 unit. Comparing to contrary, parallel 3rd increases the rating of about 0.164 unit; parallel 5th increases the rating of about 0.16 unit.

(b)

Select the covariates for popular and perform stepwise selection. The selected model is

Check to see if there should be changes in the random effects. According to the Restricted Likelihood Ratio Tests, only the Subject:Instrument random effect was significant.

Therefore our final model for (b) is:

```
m.pop.final <- lmer(Popular ~ Instrument + Harmony + Voice + PachListen +
                              X1990s2000s + Composing + GuitarPlay +
                       (1 | Subject:Instrument),data=rat)

summary(m.pop.final)$coef
```

Comparing to guitar, holding all other variables constant, piano decreases the rating of about 0.98 unit; string decreases the rating of about 2.71 units. Comparing to I-VI-V, holding all other variables constant,
9 Harmony I-V-IV decreases the rating of about 0.018 unit; I-V-VI decreases the rating of about 0.28 unit; IV-I-V decreases the rating of about 0.21 unit. Comparing to contrary, parallel 3rd increases the rating of about 0.198 unit; parallel 5th increases the rating of about 0.176 unit.

With one unit of increase in "have you heard rob paravonian's Pachelbel rant", the rating decreases about 0.117, holding other variables constant. With one unit of increase in "how much do you listen to pop and rock from the 90s and 20s", the rating increases about 0.084, holding other variables constant; With one unit of increase in "Have you done any music composing", the rating increases about 0.20.

(c)

Similar to poblem 3, we add interaction terms of Musician and other covariates to the previous model.

The selected model is:

```
m.new.3.pop.selected <- lmer(Popular ~ Instrument + Musician +
                               Harmony + PachListen + X1990s2000s +
                               Composing + (1 | Subject:Instrument) +
                               Musician:Harmony, data=rat)
summary(m.new.3.pop.selected)$coef
```

```
                         Estimate Std. Error    t value
(Intercept)            6.64192800 0.58100263 11.4318381
Instrumentpiano       -0.97768681 0.27379213 -3.5709091
Instrumentstring      -2.71531182 0.27351262 -9.9275557
Musician               0.26539367 0.30611413  0.8669762
HarmonyI-V-IV         -0.09356725 0.12802400 -0.7308571
HarmonyI-V-VI         -0.03574347 0.12812541 -0.2789725
HarmonyIV-I-V         -0.22807018 0.12802400 -1.7814642
PachListen            -0.11732101 0.10106747 -1.1608187
X1990s2000s            0.08975983 0.07198046  1.2470028
Composing              0.17261736 0.10247912  1.6844148
Musician:HarmonyI-V-IV  0.19381763 0.20421738  0.9490751
Musician:HarmonyI-V-VI -0.63323693 0.20428098 -3.0998330
Musician:HarmonyIV-I-V  0.03437648 0.20405839  0.1684639
```

```r
summary(m.new.3.pop.selected)$varcor
```

```
 Groups             Name        Std.Dev.
 Subject:Instrument (Intercept) 1.4570
 Residual                       1.6741
```

We see that the interaction for Harmony:Musician is left in the model. This is the most significant interaction that we want to consider.

The coefficient for HarmonyI-V-IV:Musician is 0.19, HarmonyI-V-VI:Musician is -0.63, HarmonyIV-I-V:Musician is 0.03. It means that musician rating Harmony I-V-VI lower on the classical rating. (Which is is the beginning progression for Pachelbel's Canon!)

See Next Page for the One Page Summary

Problem 5: Brief Writeup

Overview

Dr.Jimenez and a student Vincent Rossi collected data in a designed experiment to investigate the influence of instrument, harmonic motion and voice leading on listener's identification of musical as "classical" or "popular".

There are 70 subjects participated in this study. They were presented with 36 musical stimuli and were asked to rate the music on two different scales, one for how classical the music sound (1-10) and one for how popular the music sound (1-10). The three experimental design factors are Instrument (3 levels), Harmonic Motion (4 levels) and Voice Leading (3 levels). The combination gives $3 \times 3 \times 4 = 36$ different music pieces.

Statistical Methods and Results

We first looked at the data and performed some exploratory data analysis, examined the structure of the dataset before fitting the model. Linear models and linear mixed effect models were built to investigate whether the design factors (Instrument, Harmony & Voice) affect Classical and Popular ratings. In the linear mixed effect models, a "repeated measures" model was fitted, i.e. a random intercept model is fit for each of the 70 participants on the 36 ratings for both Classical and Popular. Restricted likelihood ratio test was simulated to test if the random effects of interest were statistically significant. Stepwise selection of the lmer model was performed to choose the most significant covariates based on Akaike information criterion.

It was shown in 1(a) and 4.1 that for Classical ratings, all three factors significantly affect the rating. For Popular ratings, referring to 4(a), Instrument is the only main effect which has a significant effect on the
10 musical excerpts' popular rating. Guitar played music is 0.96 unit higher than piano and 2.6 unit higher than music performed by String. As for Leading Voice, from 1(c), we see that contrary motion is more likely to be rated as Classical music.

For Harmony, in the Classical music ratings, the interaction between Harmony I-V-VI and Musician is 1.26, meaning that Harmony I-V-VI is rated higher in Classical for musicians compared to non-musicians, holding other variables constant. On the contrary, in the Popular music ratings, the interaction between Harmony I-V-VI and Musician is -0.63, meaning that if Harmony I-V-VI is rated lower compared to non-musicians, holding other variables constant.

For Classical ratings, the random effect of Subject interact with Leading Voice is not statistically significant. For Popular ratings, the random effect of Subject interact with Leading Voice and the random effect of Subject interact with Harmony are not statistically significant. For both Classical and popular ratings, the random effects of Subject interact with Harmony or Voice are relatively small compared to the residual variance. Therefore, our fixed effects of Harmony and Voice is capturing the mean pretty well, and we can leave these two random slopes out of the model for practical reasons.

Interestingly, the more music classes the participants took the less likely to rate music as classical; however if the participant took AP music theory, he or she is more likely to rate the music as classical. The more familiar the participant is with the Pachelbel's Canon, the higher the rating he or she will give the Classical ratings, and lower he or she will give the popular ratings.

Discussion

Two types of random intercept were fitted (a random intercept for each subject; random intercept for each of the design factors interact with each subject). There might be other types of interaction that we could have investigated. The relationship between AP theory class, college music classes and the number of total music classes taken in relation to the ratings can be further analyzed.