Homework 5

Alex Reinhart

December 10, 2013

Problem 1.

(a) A simple fixed-effect ANOVA analysis indicates that all three variables are statistically significant predictors of classical rating:

```
> anova(lm(Classical ~ Harmony + Instrument + Voice - 1, data=ratings))

Analysis of Variance Table

Response: Classical
             Df Sum Sq Mean Sq   F value    Pr(>F)
Harmony       4  83657 20914.3 3965.0604 < 2.2e-16 ***
Instrument    2   4127  2063.6  391.2243 < 2.2e-16 ***
Voice         2     86    42.8    8.1181 0.0003061 ***
Residuals  2485  13107     5.3
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The F test results indicate that the model improvement when each variable is added is greater than we would expect by chance. According to the coefficient estimates, `I-V-VI` harmony results in the highest Classical scores, and stringed instruments beat pianos and guitars. `Par3rd` and `Par5th` vocals are "less classical" than `contrary`.

(b) Our model is

$$y_i = \beta_{j[i]} + a_{h[i]} + b_{t[i]} + c_{v[i]} + \epsilon \tag{1}$$

$$\beta_j \sim N(0, \tau^2) \tag{2}$$

$$\epsilon \sim N(0, \sigma^2), \tag{3}$$

where y_i is the ith classical rating, $j[i]$ is the rater responsible for the rating, and a, b and c are coefficients corresponding to each fixed effect. The fixed effects are factor variables, so the coefficients are indexed by the particular categories each observation falls in. The vectors h, v, and t indicate the particular harmony, vocal or instrument category each observation falls into.

Adding the random effects term decreases AIC by about 762, suggesting that the random effects term is very useful. Alternately, we can use the exact likelihood ratio test from `RLRsim` to test the presence of the random effect:

```
> lm.a <- lmer(Classical ~ Harmony + Instrument + Voice + (1 | Subject) - 1,
+              data=ratings, REML=FALSE)
> lm.0 <- lm(Classical ~ Harmony + Instrument + Voice - 1, data=ratings)
> exactLRT(lm.a, lm.0)

        simulated finite sample distribution of LRT. (p-value based on 10000
        simulated values)

data:
LRT = 763.588, p-value < 2.2e-16
```

This tiny p value suggests the random effect is useful. (Both models had 2,493 non-`NA` observations.)

The fitted random effects model is

```
> display(lm.a)

lmer(formula = Classical ~ Harmony + Instrument + Voice + (1 |
    Subject) - 1, data = ratings, REML = FALSE)
                coef.est coef.se
HarmonyI-IV-V    4.34     0.19
HarmonyI-V-IV    4.31     0.19
HarmonyI-V-VI    5.11     0.19
HarmonyIV-I-V    4.39     0.19
Instrumentpiano  1.38     0.09
Instrumentstring 3.13     0.09
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
```

The coefficient estimates are almost identical to the ordinary linear model, but with slightly larger standard errors for harmony and smaller standard errors for instrument and voice. My main conclusions described above are unchanged.

(c) Including all three random intercepts decreases AIC by 411 and BIC by about 400, suggesting that the model is substantially improved. (Adding these random effects did not add or remove any observations from the dataset, so the AIC and BIC are comparable.)

```
> display(lmer(Classical ~ Harmony + Instrument + Voice + (1 | Subject:Harmony) +
+              (1 | Subject:Instrument) + (1 | Subject:Voice) - 1, data=ratings,
+              REML=FALSE))

lmer(formula = Classical ~ Harmony + Instrument + Voice + (1 |
    Subject:Harmony) + (1 | Subject:Instrument) + (1 | Subject:Voice) -
    1, data = ratings, REML = FALSE)
                coef.est coef.se
HarmonyI-IV-V    4.34     0.21
HarmonyI-V-IV    4.31     0.21
HarmonyI-V-VI    5.11     0.21
HarmonyIV-I-V    4.40     0.21
Instrumentpiano  1.36     0.26
Instrumentstring 3.13     0.26
Voicepar3rd     -0.41     0.08
Voicepar5th     -0.37     0.08

Error terms:
 Groups             Name        Std.Dev.
 Subject:Harmony    (Intercept) 0.66
 Subject:Voice      (Intercept) 0.16
 Subject:Instrument (Intercept) 1.47
 Residual                       1.56
---
number of obs: 2493, groups: Subject:Harmony, 280; Subject:Voice, 210; Subject:Instrumer
AIC = 10057.5, DIC = 10033.5
deviance = 10033.5
```

The coefficient estimates are still almost identical to the simple linear model, with larger standard errors for harmony and instrument and smaller SEs for voice.

Apparently, choice of instrument is responsible for the widest inter-subject variance in random instruments, followed by harmony and then voice. Apparently there are widely differing opinions on how classical different instruments are. All of the variance components are smaller than the estimated residual variance, so while we have "soaked up" some residual variance by adding the random effects, we have not removed all of it.

Written mathematically, we have the model

$$y_i = \alpha_{j[i],h[i]} + \beta_{j[i],t[i]} + \gamma_{j[i],v[i]} + a_{h[i]} + b_{t[i]} + c_{v[i]} + \epsilon \tag{4}$$
$$\alpha_{j,k} \sim N(0, \tau_1^2) \tag{5}$$
$$\beta_{j,k} \sim N(0, \tau_2^2) \tag{6}$$
$$\gamma_{j,k} \sim N(0, \tau_3^2) \tag{7}$$
$$\epsilon \sim N(0, \sigma^2), \tag{8}$$

where y_i is the ith classical rating, $j[i]$ is the rater responsible for the rating, and a, b and c are coefficients corresponding to each fixed effect. The fixed effects are factor variables, so the coefficients are indexed by the particular categories each observation falls in. The random effects are indexed by subject and by harmony (h), voice (v) and instrument (t), so unique pairs of subject and instrument type (for example) have a single random β draw.

Problem 2.

(a) To do this, I ran backwards model selection starting with the full model with all covariates. (This means I started with observations removed due to `NA`s; no observations were added back into the model when variables were removed.) Variables were kept in the model if they caused a decrease in BIC of greater than 5. To do this, I used the `bfFixefLMER_F.fnc` method from the LMERConvenienceFunctions package, though I first had to fix a bug where it could not correctly read BIC from `lmer` output.

The model selection process removed the voice variable, which is an experimental factor, so I reintroduced it into the model manually after selection.

```
> lmer.full <- lmer(Classical ~ Harmony + Instrument + Voice +
+                   (1 | Subject:Harmony) + (1 | Subject:Instrument) +
+                   (1 | Subject:Voice) - 1 + CollegeMusic + APTheory +
+                   NoClass + KnowRob + KnowAxis + PachListen +
+                   ClsListen + OMSI + ConsInstr + ConsNotes + Composing +
+                   PianoPlay + GuitarPlay, data=ratings, REML=FALSE)
> lmer.sel <- bfFixefLMER_F.fnc(lmer.full, method="BIC", threshold=5,
+                               log.file=FALSE, reset.REML.TRUE=FALSE)
> # reintroduce voice to lmer.sel@call
> lmer.sel <- lmer(Classical ~ Harmony + Instrument +
+                  (1 | Subject:Harmony) + (1 | Subject:Instrument) +
+                  (1 | Subject:Voice) + APTheory + NoClass + KnowRob +
+                  KnowAxis + PachListen + ConsNotes + Voice - 1,
+                  data = ratings, REML = FALSE)

> display(lmer.sel)
```

```
lmer(formula = Classical ~ Harmony + Instrument + (1 | Subject:Harmony) +
    (1 | Subject:Instrument) + (1 | Subject:Voice) + APTheory +
    NoClass + KnowRob + KnowAxis + PachListen + ConsNotes + Voice -
    1, data = ratings, REML = FALSE)
                coef.est coef.se
HarmonyI-IV-V    2.50     0.91
HarmonyI-V-IV    2.49     0.91
HarmonyI-V-VI    3.39     0.91
HarmonyIV-I-V    2.55     0.91
Instrumentpiano  1.61     0.30
Instrumentstring 3.53     0.30
APTheoryYes      1.02     0.36
NoClass         -0.07     0.11
KnowRob         -0.03     0.08
KnowAxis         0.14     0.07
PachListen       0.25     0.18
ConsNotes       -0.05     0.08
Voicepar3rd     -0.40     0.11
Voicepar5th     -0.29     0.11

Error terms:
 Groups             Name        Std.Dev.
 Subject:Harmony    (Intercept) 0.67
 Subject:Voice      (Intercept) 0.20
 Subject:Instrument (Intercept) 1.35
 Residual                       1.58
---
number of obs: 1577, groups: Subject:Harmony, 176; Subject:Voice, 132; Subject:Instrumer
AIC = 6396.3, DIC = 6360.3
deviance = 6360.3
```

(b) I compared the BIC of the new model to the no-random-effects model and the one with a single random intercept per participant:

```
> BIC(lm(Classical ~ Harmony + Instrument + APTheory +
+        NoClass + KnowRob + KnowAxis + PachListen + ConsNotes +
+        Voice - 1, data = ratings))

[1] 7092.075

> BIC(lmer(Classical ~ Harmony + Instrument + (1 | Subject) + APTheory +
+          NoClass + KnowRob + KnowAxis + PachListen + ConsNotes +
+          Voice - 1, data = ratings, REML = FALSE))

[1] 6780.466
```

```
> BIC(lmer.sel)

[1] 6492.82
```

It's clear that allowing interactions in random effects improves the model, with a BIC drop of nearly 290. (All models had 1,577 observations to test on.)

(c) Based on the model coefficients displayed in (a) above, my interpretation of harmony, voice and instrument remains unchanged. Listeners with experience in an AP Music Theory course were predisposed to rating songs as classical, and those who were more familiar with Pachelbel's Canon or the Axis of Evil's comedy piece on it were also predisposed to classical ratings. Having taken more music classes, and being more familiar with Rob Paravonian's rant, was weakly associated with lower classical ratings, along with concentrating more on the notes of the music.

Problem 3. I split `Selfdeclare` ratings so that responses 1 and 2 are rated as not musicians and 3-6 are rated as musicians. I then constructed the random effects model with interactions between musicianship and all other predictors, and subjected this model to backwards selection using BIC as the criterion. (With and without interactions, the model had 1,577 non-`NA` observations.) Only interactions were considered for removal, and all of them were removed. They apparently did not substantially benefit the model.

Several interactions did appear to be significant based on t statistic alone: musicianship and harmony, for example, and musicianship and `ConsNotes`. Being a musician who concentrated more on the notes was associated with a higher classical rating, and being a musician who heard a I-V-VI harmony was also correlated with higher ratings.

Problem 4.

(a) Examining popular ratings, I determined via BIC that including all three interaction random effects was beneficial, as it was before with classical ratings. In backwards selection, the same fixed effects were chosen as beneficial to the model. Hence I will interpret the selected model.

no voice fixed eff??

```
> lmer.pop <- lmer(Popular ~ Harmony + Instrument +
+                  (1 | Subject:Harmony) + (1 | Subject:Instrument) +
+                  (1 | Subject:Voice) + APTheory + NoClass + KnowRob +
+                  KnowAxis + PachListen + ConsNotes + Voice - 1,
+                  data = ratings, REML = FALSE)
> display(lmer.pop)

lmer(formula = Popular ~ Harmony + Instrument + (1 | Subject:Harmony) +
    (1 | Subject:Instrument) + (1 | Subject:Voice) + APTheory +
    NoClass + KnowRob + KnowAxis + PachListen + ConsNotes + Voice -
    1, data = ratings, REML = FALSE)
                coef.est coef.se
```

```
HarmonyI-IV-V     7.67     0.88
HarmonyI-V-IV     7.70     0.88
HarmonyI-V-VI     7.38     0.88
HarmonyIV-I-V     7.44     0.88
Instrumentpiano  -1.13     0.29
Instrumentstring -3.00     0.29
APTheoryYes      -0.07     0.35
NoClass           0.08     0.10
KnowRob           0.13     0.08
KnowAxis          0.04     0.07
PachListen       -0.26     0.17
ConsNotes         0.07     0.07
Voicepar3rd       0.18     0.11
Voicepar5th       0.21     0.11

Error terms:
 Groups              Name        Std.Dev.
 Subject:Harmony     (Intercept) 0.66
 Subject:Voice       (Intercept) 0.22
 Subject:Instrument (Intercept) 1.28
 Residual                        1.64
---
number of obs: 1577, groups: Subject:Harmony, 176; Subject:Voice, 132; Subject:Instrumer
AIC = 6493.3, DIC = 6457.3
deviance = 6457.3
```

All kinds of harmony appear to have similar effects on popular ratings (within a standard error); stringed instruments are the least pop-y, followed by pianos and then guitars. `Par3rd` and `Par5th` vocals are more Popular than `contrary`.

(b) Taking an AP Music Theory course was very weakly associated with lower Popular ratings. Taking more music classes, being familiar with Rob Paravonian's rant, knowing the Axis of Evil's comedy piece, and concentrating on the notes of the music resulted in slightly higher Popular ratings. Familiarity with Pachelbel's Canon correlated with slightly lower Popular ratings.

(c) I performed the same procedure as in problem 3. As before, all interactions with musicianship were removed by backwards selection. The main effect for musicianship showed no significant difference between both states; the coefficient estimates were within one standard error of each other.

Problem 5. In summary, classical ratings are dependent upon the harmony and instrument played in the piece, with a lesser dependence on the type of vocals. Using a hierarchical model with random intercepts for each experimental subject (and separate intercepts depending on the harmony, instrument, and voice in the rated piece), I estimated that stringed instruments sound the most classical, followed by pianos and then guitars. Harmony produced a less obvious distinction, with the coefficients not clearly significantly different; nonetheless, `I-V-VI` harmony was estimated to have the highest classical ratings, and `I-V-IV` harmony the lowest. Vocals coded as `par3rd` and `par5th` produced lower classical ratings.

Other fixed effects were selected and entered into the model based on backwards selection using BIC. Listeners with experience in an AP Music Theory course were predisposed to rating songs as classical, and those who were more familiar with Pachelbel's Canon or the Axis of Evil's comedy piece on it were also predisposed to classical ratings. Having taken more music classes, and being more familiar with Rob Paravonian's rant, was weakly associated with lower classical ratings, along with concentrating more on the notes of the music.

For Popular ratings, a similar hierarchical model was fit with an identical random effects structure. All kinds of harmony appeared to have similar effects on popular ratings (within a standard error); stringed instruments were the least Popular, followed by pianos and then guitars. `Par3rd`- and `Par5th`-coded vocals are more Popular than contrary.

Taking an AP Music Theory course was very weakly associated with lower Popular ratings. Taking more music classes, being familiar with Rob Paravonian's rant, knowing the Axis of Evil's comedy piece, and concentrating on the notes of the music resulted in slightly higher Popular ratings. Familiarity with Pachelbel's Canon correlated with slightly lower Popular ratings.

In both models, the piece's instrument was responsible for the most between-subject random intercept variance, followed by the piece's harmony.

To test whether the musicianship (or lack thereof) of the reader had any interactions with other fixed effects influencing Popular or Classical ratings, I fit models with interaction terms and performed backwards selection to remove non-influential terms. This removed all interaction terms, suggesting that musicianship is not associated with any significant interactions for either Popular or Classical ratings.

4: 18/20
5: 20/20

38