1 a 9/9
b 7/9
c 7/9

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

3 6/9

4 a 9/9
b 5/9
c 9/9

5 10/10

Total 79/100

36-763 - Homework 5

Brandon Sherman

December 18, 2015

```r
set.seed(8675309)
ratings <- read.csv("data/ratings.csv")[, -1]
ratings <- select(ratings, -first12)
```

```r
head(filter(ratings, is.na(Classical), is.na(Popular)))
```

```
##   Subject Harmony Instrument    Voice Selfdeclare OMSI X16.minus.17
## 1      24  I-IV-V     guitar   par3rd           1   55            9
## 2      24  I-IV-V      piano   par5th           1   55            9
## 3      24  I-V-IV     guitar contrary           1   55            9
## 4      24  I-V-IV     guitar   par5th           1   55            9
## 5      24  I-V-IV      piano contrary           1   55            9
## 6      24  I-V-IV      piano   par5th           1   55            9
##   ConsInstr ConsNotes Instr.minus.Notes PachListen ClsListen KnowRob
## 1         0         0                 0         NA        NA       0
## 2         0         0                 0         NA        NA       0
## 3         0         0                 0         NA        NA       0
## 4         0         0                 0         NA        NA       0
## 5         0         0                 0         NA        NA       0
## 6         0         0                 0         NA        NA       0
##   KnowAxis X1990s2000s X1990s2000s.minus.1960s1970s CollegeMusic NoClass
## 1        0          NA                           NA           NA      NA
## 2        0          NA                           NA           NA      NA
## 3        0          NA                           NA           NA      NA
## 4        0          NA                           NA           NA      NA
## 5        0          NA                           NA           NA      NA
## 6        0          NA                           NA           NA      NA
##   APTheory Composing PianoPlay GuitarPlay X1stInstr X2ndInstr Classical
## 1       NA         2         0          0        NA        NA        NA
## 2       NA         2         0          0        NA        NA        NA
## 3       NA         2         0          0        NA        NA        NA
## 4       NA         2         0          0        NA        NA        NA
## 5       NA         2         0          0        NA        NA        NA
## 6       NA         2         0          0        NA        NA        NA
##   Popular
## 1      NA
## 2      NA
## 3      NA
## 4      NA
## 5      NA
## 6      NA
```

```r
tail(filter(ratings, is.na(Classical), is.na(Popular)))
```

```
##    Subject Harmony Instrument    Voice Selfdeclare OMSI X16.minus.17
## 22      73  I-V-IV      piano contrary           3  233           -1
## 23      73  I-V-IV      piano   par3rd           3  233           -1
## 24      73  I-V-VI      piano   par3rd           3  233           -1
## 25      73  I-V-VI      piano   par5th           3  233           -1
## 26      73  IV-I-V      piano contrary           3  233           -1
## 27      73  IV-I-V      piano   par3rd           3  233           -1
##    ConsInstr ConsNotes Instr.minus.Notes PachListen ClsListen KnowRob
## 22         5         5                 0          5         3      NA
## 23         5         5                 0          5         3      NA
## 24         5         5                 0          5         3      NA
## 25         5         5                 0          5         3      NA
## 26         5         5                 0          5         3      NA
## 27         5         5                 0          5         3      NA
##    KnowAxis X1990s2000s X1990s2000s.minus.1960s1970s CollegeMusic NoClass
## 22        0           5                            3            1       0
## 23        0           5                            3            1       0
## 24        0           5                            3            1       0
## 25        0           5                            3            1       0
## 26        0           5                            3            1       0
## 27        0           5                            3            1       0
##    APTheory Composing PianoPlay GuitarPlay X1stInstr X2ndInstr Classical
## 22        1         1         0          0         4         1        NA
## 23        1         1         0          0         4         1        NA
## 24        1         1         0          0         4         1        NA
## 25        1         1         0          0         4         1        NA
## 26        1         1         0          0         4         1        NA
## 27        1         1         0          0         4         1        NA
##    Popular
## 22      NA
## 23      NA
## 24      NA
## 25      NA
## 26      NA
## 27      NA
```

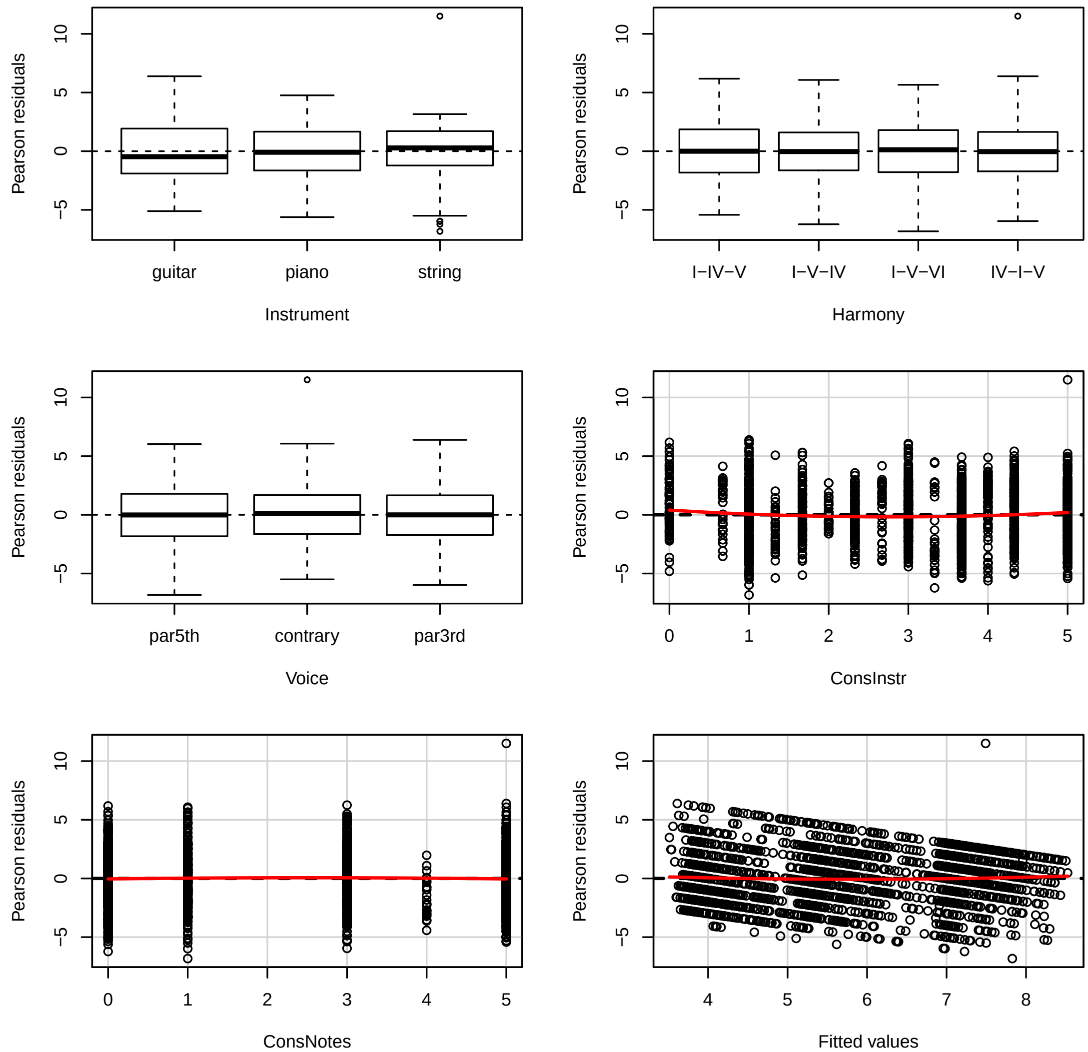
```r
ratings <- filter(ratings, !is.na(Classical), !is.na(Popular))

# Set baseline of Voice as "parallel 5th" because we want a coefficient for "contrary"
ratings$Voice <- factor(ratings$Voice, levels = c("par5th", "contrary", "par3rd"))
```

It seems like a few subjects filled out the survey, then did not complete the part of the survey of interest on classical and popular music. Many of them barely filled out the survey, so we cannot claim that individuals are missing at random. Therefore we decided to exclude any individuals with both classical and popular ratings missing.

We also reset the baseline for voice as “parallel fifths” because they were considered off-limits by the classical community until the early 1900s and so that we get a coefficient for “contrary”, and we keep the baseline for harmony as I-IV-V because as a common blues progression, it makes sense as a baseline to compare to classical.

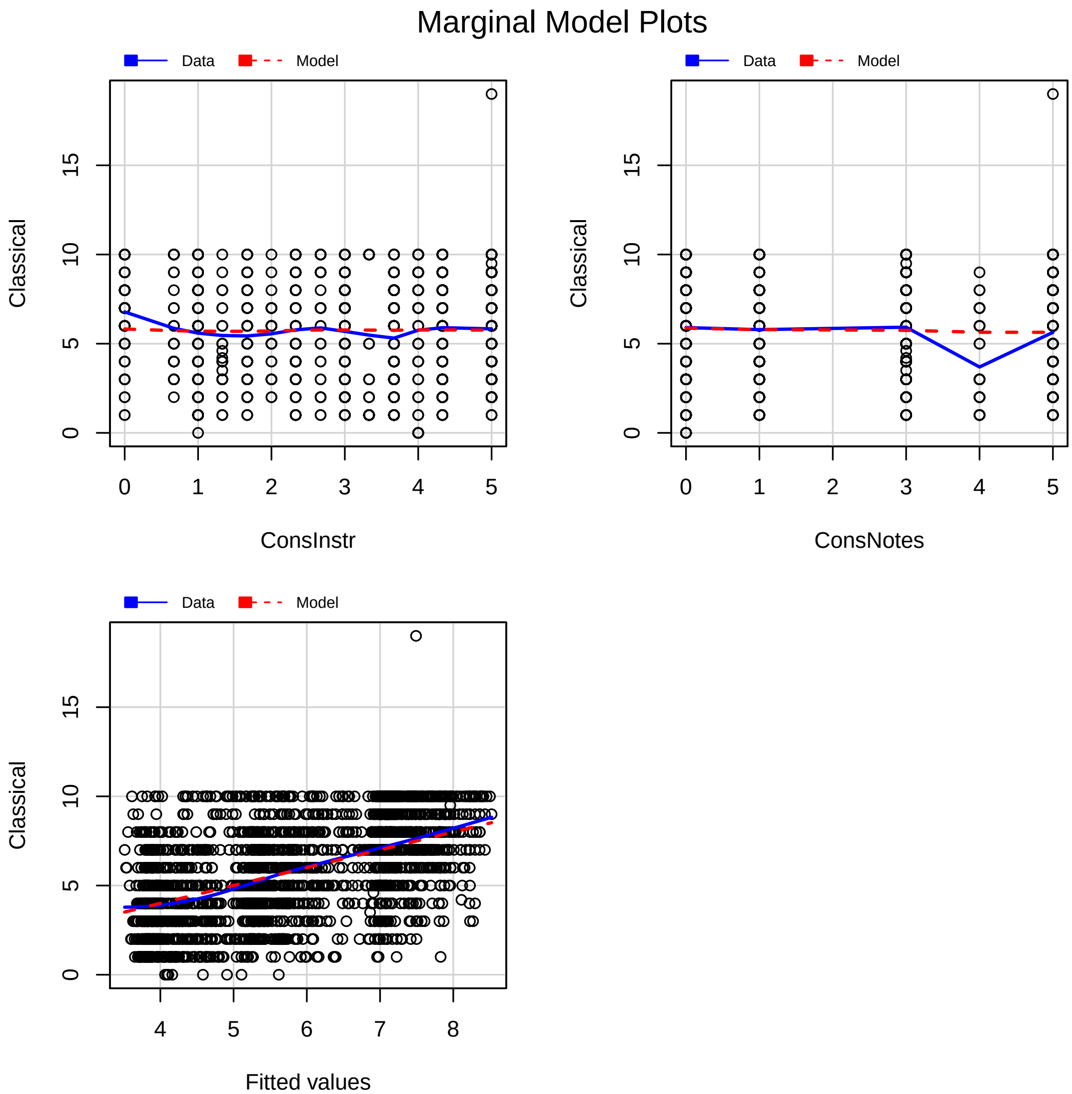
1) a)

```r
library("car")

lm.main.three <- lm(Classical ~ Instrument + Harmony + Voice +
                    ConsInstr + ConsNotes,
                    data = ratings)
lm.none <- lm(Classical ~ ConsInstr + ConsNotes, data = ratings)
anova(lm.none, lm.main.three)

## Analysis of Variance Table
##
## Model 1: Classical ~ ConsInstr + ConsNotes
## Model 2: Classical ~ Instrument + Harmony + Voice + ConsInstr + ConsNotes
##   Res.Df   RSS Df Sum of Sq      F    Pr(>F)
## 1   2130 15355
## 2   2123 11371  7    3984.3 106.27 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(lm.main.three)

##
## Call:
## lm(formula = Classical ~ Instrument + Harmony + Voice + ConsInstr +
##     ConsNotes, data = ratings)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8272 -1.7373  0.0015  1.7266 11.5087
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3.83848    0.17514  21.917  < 2e-16 ***
## Instrumentpiano    1.45451    0.12314  11.811  < 2e-16 ***
## Instrumentstring   3.18958    0.12228  26.085  < 2e-16 ***
## HarmonyI-V-IV     -0.02747    0.14170  -0.194  0.84631
## HarmonyI-V-VI      0.80099    0.14170   5.653 1.79e-08 ***
## HarmonyIV-I-V      0.07151    0.14157   0.505  0.61354
## Voicecontrary      0.40078    0.12270   3.266  0.00111 **
## Voicepar3rd       -0.01804    0.12275  -0.147  0.88320
## ConsInstr          0.06774    0.04039   1.677  0.09368 .
## ConsNotes         -0.06954    0.03117  -2.231  0.02581 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.314 on 2123 degrees of freedom
##   (360 observations deleted due to missingness)
## Multiple R-squared:  0.2608,Adjusted R-squared:  0.2577
## F-statistic: 83.22 on 9 and 2123 DF,  p-value: < 2.2e-16

residualPlots(lm.main.three)
```




```
##                Test stat Pr(>|t|)
## Instrument           NA       NA
## Harmony              NA       NA
## Voice                NA       NA
## ConsInstr         3.155    0.002
## ConsNotes        -0.888    0.375
## Tukey test        1.758    0.079
```

```
marginalModelPlots(lm.main.three)

## Warning in mmps(...):  Interactions and/or factors skipped
```




```r
durbinWatsonTest(lm.main.three)

##  lag Autocorrelation D-W Statistic p-value
##    1       0.4272912      1.145292       0
##  Alternative hypothesis: rho != 0
```

A likelihood ratio test shows that compared to a model with no covariates, a model with instrument, harmony, voice, and the two attention variables, we see that the model with the covariates of interest is a good fit ($F = 106.3$, $p < 0.0001$). We decided to use the attention variables as controlling factors because it seems likely that if someone was not paying attention, then that person may be picking a rating somewhat at random. In the full model with all three covariates, we see
9 that a I-V-VI harmony is statistically significantly more likely to be rated as positive than I-VI-V (our baseline) and that voice and instrument are both highly statistically significant predictors.

When we look at the model with the main three covariates, we obtain a decent fit. Residual plots appear evenly distributed and marginal model plots have roughly concordant smoothers for the data and the model (with the caveat that our predictors aren't truly continuous).

However a Durbin-Watson test shows clear evidence of autocorrelation, with $p < 0.0001$. Therefore our standard errors are invalid and so are the p-values for our estimated coefficients.

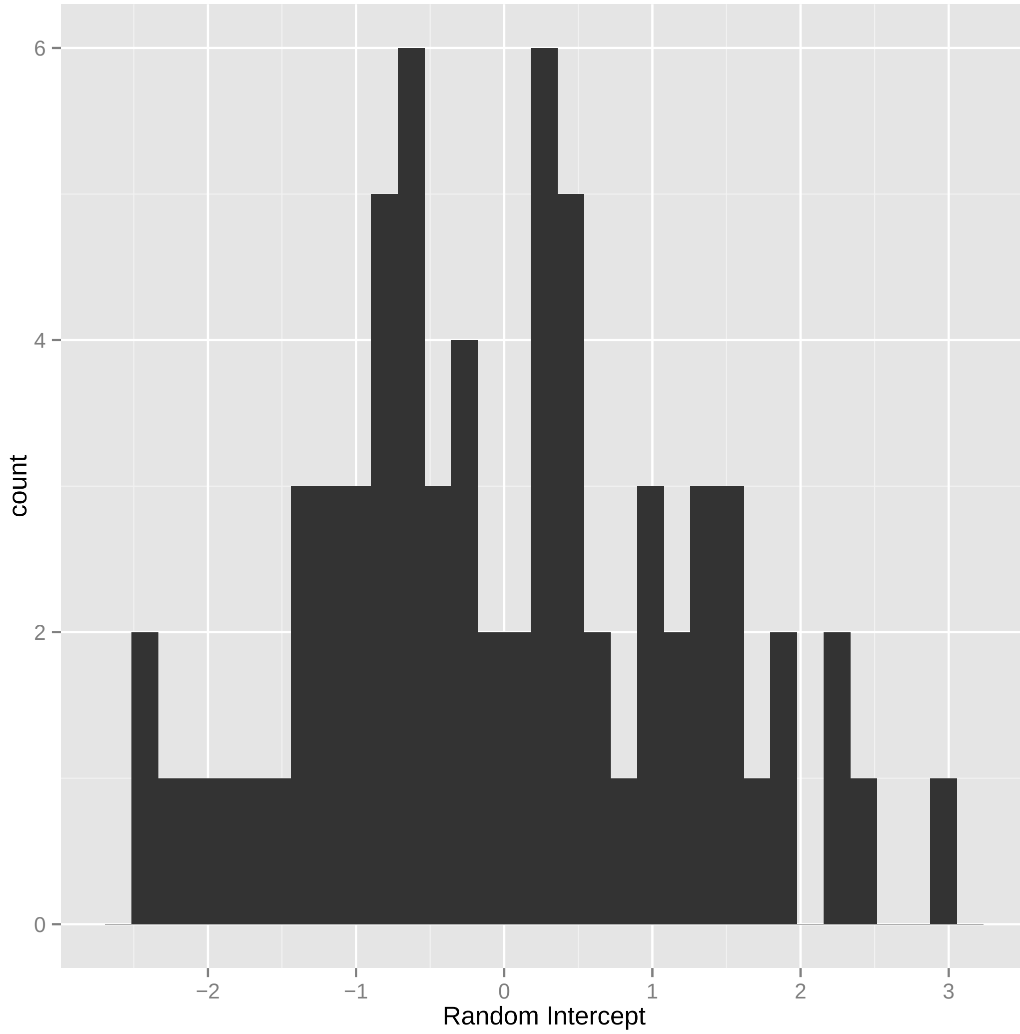
b) i)

$$\text{Level 1: } y_i \overset{iid}{\sim} N(\mu_{j[i]}, \sigma_y^2)$$
$$\text{Level 2: } \mu_j \sim N(\alpha_j, \sigma_\alpha^2).$$

how would this be modified for the experimental factors that we know have to be in the model?

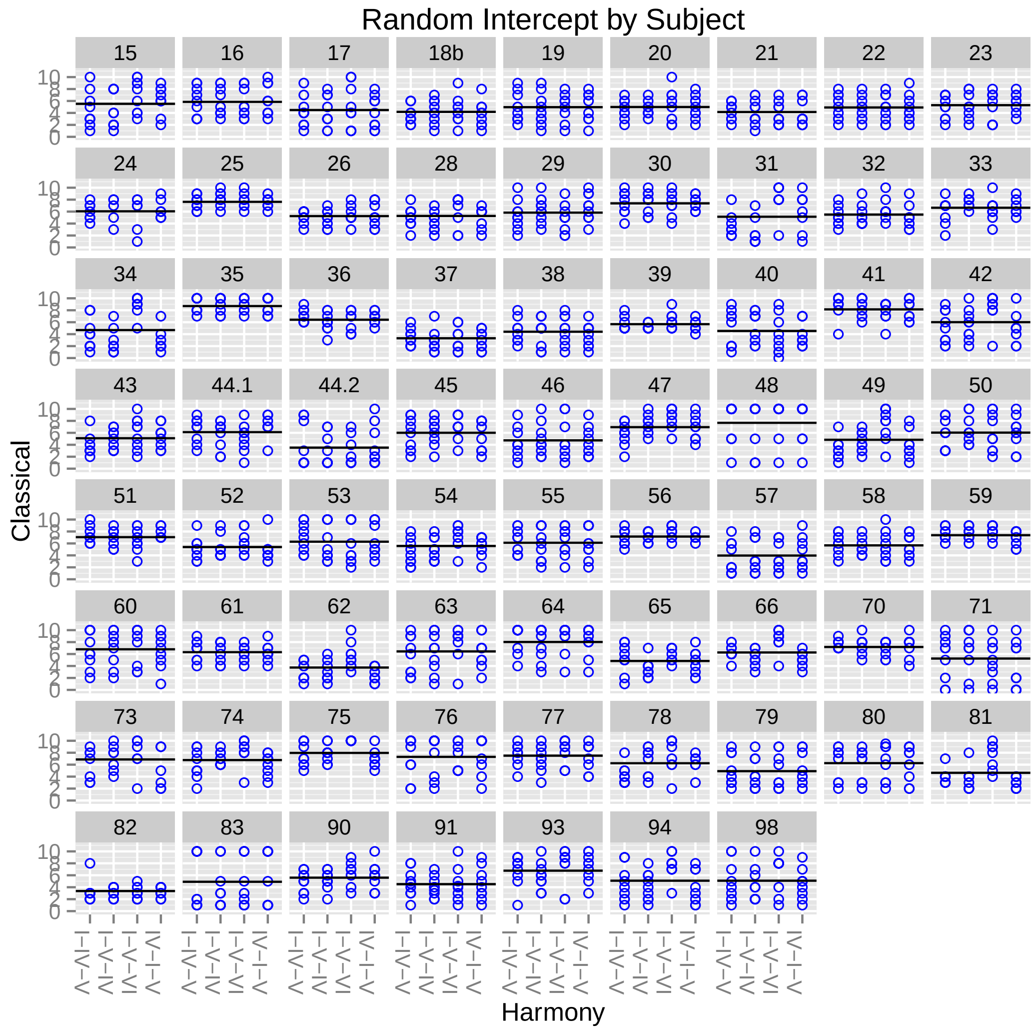
Note that I'm assuming that this is for a model without considering anything but subject.

ii)
```
library("lme4")
library("ggplot2")
library("arm")

# Method 1:  Fit an intercept-only model with JUST the random intercept,
# then look at the magnitudes of the random effect coefficients.

intercept.only <- lmer(Classical ~ (1 | Subject), data = ratings,
                       REML = FALSE)
display(intercept.only)

## lmer(formula = Classical ~ (1 | Subject), data = ratings, REML = FALSE)
## coef.est  coef.se
##     5.79     0.16
##
## Error terms:
##  Groups   Name        Std.Dev.
##  Subject  (Intercept) 1.28
##  Residual             2.33
## ---
## number of obs: 2493, groups: Subject, 70
## AIC = 11466.2, DIC = 11460.2
## deviance = 11460.2

head(ranef(intercept.only)[[1]])

##     (Intercept)
## 15  -0.26293514
## 16   0.06762002
## 17  -1.30545527
## 18b -1.61058311
## 19  -0.82233619
## 20  -0.79690887

qplot(ranef(intercept.only)[[1]][[1]], data = ranef(intercept.only)[[1]],
      geom = "histogram") + xlab("Random Intercept")

## stat_bin:  binwidth defaulted to range/30.  Use 'binwidth = x' to adjust this.
```




```
# Method 2:  Fit a bunch of random intercept models and visualize their fits
split.ratings <- split(filter(ratings), ratings$Subject)
a0.2bii <- fixef(intercept.only)[[1]] + ranef(intercept.only)$Subject[[1]]


p.2bii <- ggplot(data = ratings, aes(x = Harmony, y = Classical)) +
    geom_point(pch = 1, color = "blue") +
    facet_wrap(~ Subject) +
    ggtitle("Intercept-Only Model by Subject") +
    theme(axis.text.x = element_text(angle = -90)) + # Turn x-axis labels 90 deg.
    ggtitle("Random Intercept by Subject") +
    scale_y_continuous(limits = c(0, 11), breaks = seq(0, 10, by=2))

for (i in 1:length(split.ratings)) {
    p.2bii <- p.2bii + geom_abline(data = split.ratings[[i]],
                                   slope = 0,
                                   intercept = a0.2bii[i],
                                   color = "black")
}
```

```
plot(p.2bii)
## Warning:  Removed 1 rows containing missing values (geom_point).
```



7

I would also like to see some more formal tests or comparisons, using model fit indices or tests that we mentioned in class.

It appears that using a random intercept is quite useful.
The first method I did was to look at the raw values and the residual standard deviation and make a histogram of the random intercepts for all subjects. The distribution looks fairly normal, implying that individual intercepts vary from the mean along a roughly normal distribution.
When I plotted the fixed intercept added to each random intercept for each subject on a plot for that individual subject, the calculated intercepts can vary by quite a bit. For example, Subject 82 has a low intercept at around 4, while Subject 35 has a higher intercept at around 8. Since the scale for Classical ratings goes from 0 to 10, different intercepts can lead to vastly different base intercepts for each subject. Therefore we conclude that a random intercept model is necessary.

iii)
```
full.model.ri <- lmer(Classical ~ Instrument + Harmony + Voice + (1 | Subject),
                      data = ratings, REML = FALSE)
display(full.model.ri)

## lmer(formula = Classical ~ Instrument + Harmony + Voice + (1 |
##     Subject), data = ratings, REML = FALSE)
```

```
##                  coef.est coef.se
## (Intercept)       3.97     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
## Voicecontrary     0.37     0.09
## Voicepar3rd      -0.04     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
```

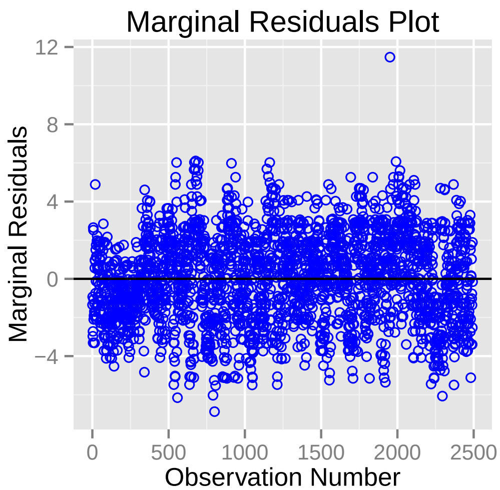
We notice that the I-V-VI harmony is still the largest type of harmony in magnitude in predicting how classical subjects perceive a piece of music (with I-VI-V as the baseline). We also see that compared to parallel 5ths, contrary voicings lead to a much larger perception in whether or not music is classical, and that compared to electric guitar, both piano and string quartets were perceived as vastly more classical. The coefficient standard errors for these estimates are quite small compared to their coefficients. But how well does our model fit? Let's look at some residual plots:

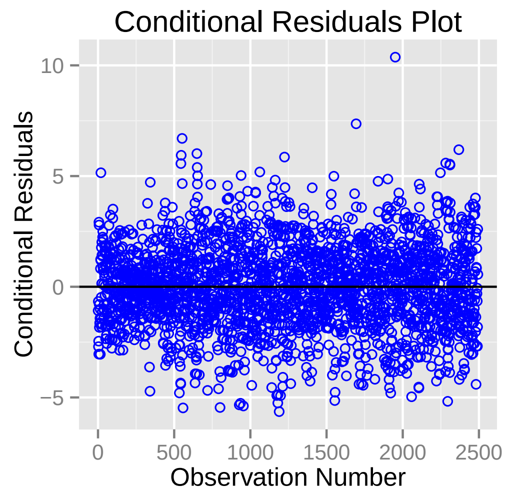
```
source("residual-functions.R") # Load Brian's residual functions

make.lmer.resid.df <- function(model) {
    # Returns a data frame with all three types of residuals
    # and all three types of predicted values, along with
    # subjects and observation numbers

    data.frame(      marg.resids = r.marg(model),
                     cond.resids = r.cond(model),
                     reff.resids = r.reff(model),
                     marg.yhats = yhat.marg(model),
                     cond.yhats = yhat.cond(model),
                     reff.yhats = yhat.reff(model),
                     Subject = model@frame$Subject,
                     ObsNum = seq(1, length(model@frame$Subject))
               )
}
resid.analysis.2biii <- make.lmer.resid.df(full.model.ri)
head(resid.analysis.2biii)
```

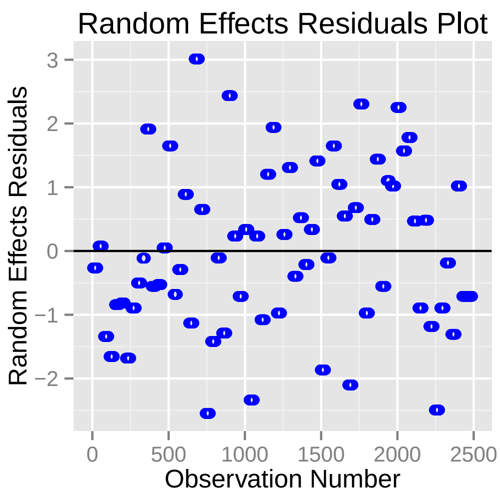
```
##   marg.resids cond.resids reff.resids marg.yhats cond.yhats reff.yhats
## 1  -1.3437419  -1.0788458   -0.264896   4.343742   4.078846   3.264896
## 2  -0.9286771  -0.6637811   -0.264896   3.928677   3.663781   3.264896
## 3  -2.9693592  -2.7044632   -0.264896   3.969359   3.704463   1.264896
## 4  -2.7207869  -2.4558909   -0.264896   5.720787   5.455891   3.264896
## 5  -3.3057222  -3.0408262   -0.264896   5.305722   5.040826   2.264896
## 6   2.6535957   2.9184917   -0.264896   5.346404   5.081508   8.264896
##   Subject ObsNum
```

```
## 1      15      1
## 2      15      2
## 3      15      3
## 4      15      4
## 5      15      5
## 6      15      6
```

```r
library("gridExtra")

plot.lmer.residuals <- function(lmer.model) {

    residual.df <- make.lmer.resid.df(lmer.model)
    residual.plots <- list(0, 0, 0)

    # Plot marginal residuals against observation number
    residual.plots[[1]] <-
    ggplot(data = residual.df, aes(x = ObsNum, y = marg.resids)) +
        geom_point(pch = 1, color = "blue") +
            geom_abline(slope = 0, intercept = 0) +
                ggtitle("Marginal Residuals Plot") +
                    xlab("Observation Number") +
                        ylab("Marginal Residuals")

    # Plot conditional residuals against observation number
    residual.plots[[2]] <-
    ggplot(data = residual.df, aes(x = ObsNum, y = cond.resids)) +
        geom_point(pch = 1, color = "blue") +
            geom_abline(slope = 0, intercept = 0) +
                ggtitle("Conditional Residuals Plot") +
                    xlab("Observation Number") +
                        ylab("Conditional Residuals")

    # Plot random effect residuals against observation number
    residual.plots[[3]] <-
    ggplot(data = residual.df, aes(x = ObsNum, y = reff.resids)) +
        geom_point(pch = 1, color = "blue") +
            geom_abline(slope = 0, intercept = 0) +
                ggtitle("Random Effects Residuals Plot") +
                    xlab("Observation Number") +
                        ylab("Random Effects Residuals")

    # Three side-by-side plots
    grid.arrange(residual.plots[[1]], residual.plots[[2]], residual.plots[[3]],
                 nrow = 2)
}

plot.lmer.residuals(full.model.ri)
```







The marginal residuals look fine. They have roughly mean 0, implying that the fixed effects fit well. The conditional residuals look roughly homoskedastic and have mean 0, although there appears to be a bit of grouping structure, since some observations close to each other have more high residuals than low residuals. And although the random effects residuals look a bit odd, there do not appear to be any clear outliers.

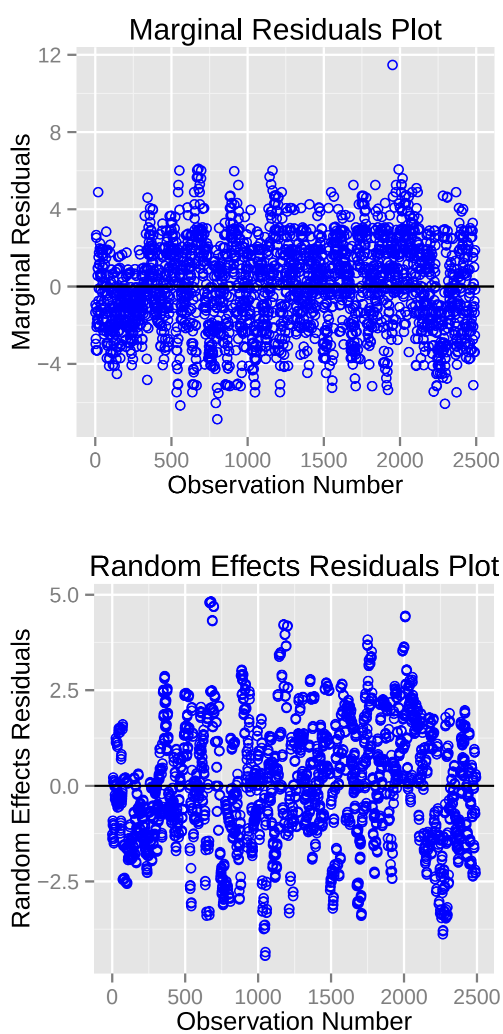
c) i)

```
library("lme4")
library("arm")

interaction.model.ri <- lmer(Classical ~ Instrument + Harmony + Voice +
                             (1 | Subject:Instrument) + (1 | Subject:Harmony) +
                             (1 | Subject:Voice),
                      data = ratings, REML = FALSE)
AIC(interaction.model.ri)
## [1] 10057.53
AIC(lm.main.three)
## [1] 9644.821
AIC(intercept.only)
## [1] 11466.25
```

```
AIC(full.model.ri)
## [1] 10468.86
```

The model with the three random interaction effects has an AIC of 10058, while the model with just the random intercept has an AIC of 10469, and the model with just the intercept has an AIC of 11466. We therefore determine that the model with the random interaction intercepts is the best fit of the three.
Note that while the model with just the main three factors and no grouping by subject has an AIC of 9645, which is by far the lowest, the issue with autocorrelation is important enough to determine that the model is not a good fit.
Just to be sure that the random interaction effects model is a good fit, let's check our residual plots:

```
plot.lmer.residuals(interaction.model.ri)
```



The marginal residuals plot looks fine, with approximate mean 0 and only one large outlier. The conditional residual plot is mostly homoskedastic, but there is a bit of increased variance above observation 2000. Most of this increased variance appears to be due to only a few points though. Finally, the random effects residuals, while oddly shaped, do not appear to have any large outliers except for a small cluster of points at around observation 600. However since it is only about 5 points, the model fit seems fine.

ii)
```r
display(interaction.model.ri)
```

```
## lmer(formula = Classical ~ Instrument + Harmony + Voice + (1 |
##     Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice),
##     data = ratings, REML = FALSE)
##                    coef.est coef.se
## (Intercept)         3.97     0.21
## Instrumentpiano     1.36     0.26
## Instrumentstring    3.13     0.26
## HarmonyI-V-IV      -0.03     0.14
## HarmonyI-V-VI       0.77     0.14
## HarmonyIV-I-V       0.06     0.14
## Voicecontrary       0.37     0.08
## Voicepar3rd        -0.04     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:Instrumen
## AIC = 10057.5, DIC = 10033.5
## deviance = 10033.5
```

We see that as in our earlier models, a piano and a string quartet are considered far more classical than electric guitar, with piano considered about 1.4 points higher than guitar and string quartet considered about 3.1 points higher than guitar. While a I-V-IV harmony does not seem to be that different from I-IV-V in predicting "classicalness", I-V-IV (the "Pachelbel progression") is associated with an increase in classicalness of about 3/4 of a point compared to I-IV-V. Contrary voice is fairly associated with classicalness, since it is, on average, about 0.37 points more "classical" than parallel 5ths. None of the fixed standard errors are unusually large.

We'll discuss standard deviations because that is what the `display()` function gives us, and because they are easier to interpret than variances, which have squared units. The estimated variance components for most of the random effect terms appear quite fine. The standard deviation for harmony is 0.66, meaning that subjects' perception of harmony vary by almost 3/4 from the mean perception across each combination. This is a variation under 10% of the total scale, which is a remarkably consistent amount of personal bias between subjects.

The standard deviation for subjects' perception of voice has a standard deviation of about 0.16, which is even smaller, implying that subjects view voice even more consistently than they do harmony.

Finally, instrument had the largest standard deviation. Subjects varied heavily by how classical they considered piano and string quartets relative to electric guitar, but perhaps this is due to some underlying tendency to rate things as "extreme". It is possible that a moderately-minded subject views piano as quite classical, but might conservatively rate it a 7. Meanwhile, someone with a more polarizing mindset may view piano as incredibly classical and rate it a 10.

The estimated residual standard deviation is 1.56, which is larger than the other standard deviations. Since it covers almost 1/4 of the scale, we probably need more covariates to explain why people rate music as classical.

iii)

you don't need to reproduce the fixed effects at level 2, but you do need to have multiple random effects for the variation among subjects in reaction to each level of each design factor

Level 1:

$y_i \sim N(\mu_{j[i]}, \sigma_y^2)$

$\mu_{j[i]} = \alpha_{j[i]} + \beta_{piano} piano_i + \beta_{str} str_i + \beta_{H1} H1_i + \beta_{H2} H2_i + \beta_{H3} H3_i + \beta_{VC} X_{i,vc} + \beta_{VP3} VP3$

Level 2:

7

$\alpha_{j[i]} \sim N(\psi_j, \sigma_\alpha^2)$

$\psi_j = \gamma_0 + \gamma_{piano} piano_i s_j + \gamma_{str} str_i s_j + \gamma_{H1} H1_i s_j + \gamma_{H2} H2_i s_j + \gamma_{H3} H3_i s_j + \gamma_{VC} X_{i,vc} + \gamma_{VP3} VP3$

where s_j is the jth subject and all variables above are indicators.

2) a) A footnote on p. 245 of Gelman & Hill gives a few definitions of fixed effects, including "constant across individuals", "interesting in themselves (independently of the population)", and "the sample exhausts the population." Under these definitions, it seems like harmony, instrument, and voice are fixed effects, since they are the only factors that are constant across individuals and exhaust the whole population of measured effects. Additionally, I would classify "CollegeMusic" and "APTheory" as fixed effects because in both situations, someone either did something or did not.

part of the confusion is "what is an individual", which is somewhat different in G&H examples, vs this data set

Far trickier are the 0 to 5 scale-based variables. Since the scale is intended to be all-inclusive and there are no other levels that can be modeled, we can conclude that they are the only levels, so they should be fixed. However they are *not* constant across individuals since different individuals clearly have different scale values. We will try treating them as fixed since we have data on almost all levels of interest. In that case, I would add the following fixed variables to the model containing instrument, harmony, and voice. Note that I would not include the instrument proficiency covariates because many individuals did not play an instrument.

- Selfdeclare
- ConsInstr
- ConsNotes
- PachListen
- ClsListen
- KnowRob
- KnowAxis
- X1990s2000s
- CollegeMusic
- APTheory
- Composing
- PianoPlay
- GuitarPlay

4

I don't see that you've tried any variable selection here

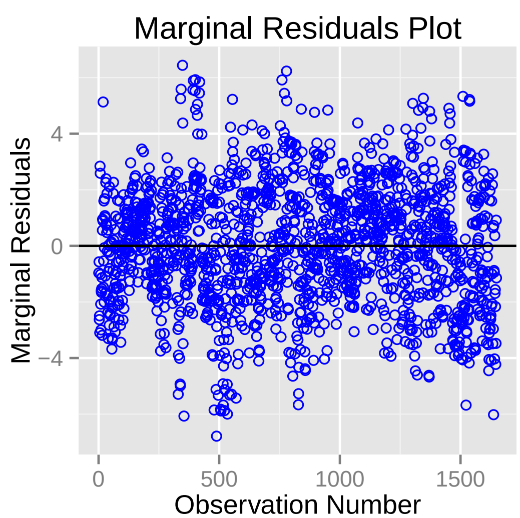
Note that I'm treating all of these scale variables as continuous. There is no reason to believe that the difference between a 1 and a 2 differs from between a 4 and a 5, and having 4 indicator variables for each scale would lead to a harder to interpret model.

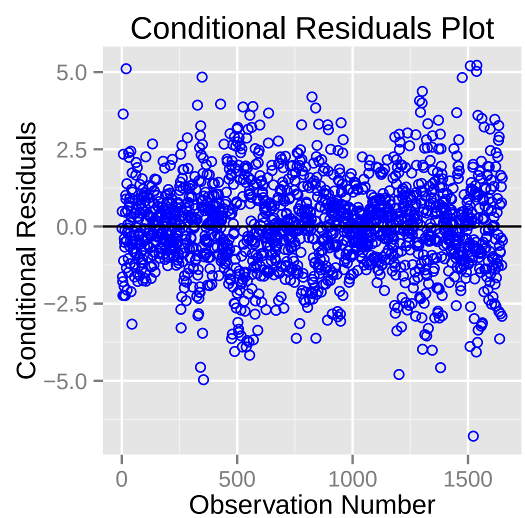
```
lmer.2a <- lmer(Classical ~ Instrument + Harmony + Voice + Selfdeclare + ConsInstr + ConsNotes
                PachListen + ClsListen + KnowRob + KnowAxis + X1990s2000s + CollegeMusic + APTh
                Composing + PianoPlay + GuitarPlay +
                            (1 | Subject:Instrument) + (1 | Subject:Harmony) +
                            (1 | Subject:Voice),
                      data = ratings, REML = FALSE)
# We'll display the actual model in part c)
```

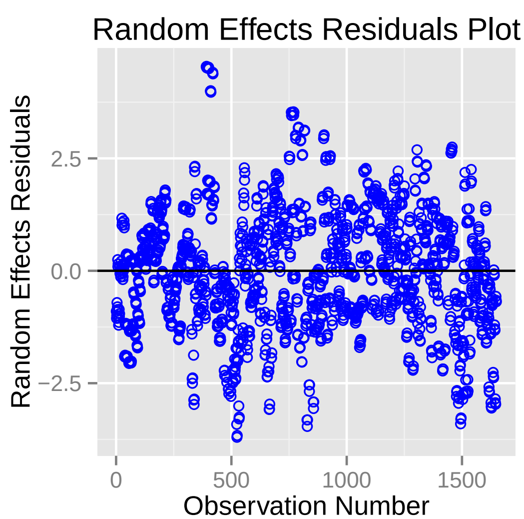
b)
```
VarCorr(interaction.model.ri)

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

VarCorr(lmer.2a)

## Groups             Name        Std.Dev.
## Subject:Harmony    (Intercept) 0.64055
## Subject:Voice      (Intercept) 0.18474
## Subject:Instrument (Intercept) 1.24839
## Residual                       1.57409

plot.lmer.residuals(lmer.2a)
```

4

no formal rechecking of the randm effects?







Adding more fixed effects does not appear to impact the random effects much. The standard deviation of the random effect for subject and harmony in our new model is about 0.01 lower than in our original model, and for subject and voice, it is slightly higher by about 0.03. The residual standard deviation is 1.57, which is about 0.01 larger than our original model. The original model has an intraclass correlation coefficient of

The only notable decrease is in the standard deviation of subject on instrument, which decreased by almost 1/4 of a point. This implies that the scale covariates, which deal with individuals' prior musical knowledge, reduce the variability in how subjects perceive the classicalness of each instrument.

The random effects residuals also look virtually identical to how they did in our original model.

c)
```
display(lmer.2a)

## lmer(formula = Classical ~ Instrument + Harmony + Voice + Selfdeclare +
##     ConsInstr + ConsNotes + PachListen + ClsListen + KnowRob +
##     KnowAxis + X1990s2000s + CollegeMusic + APTheory + Composing +
##     PianoPlay + GuitarPlay + (1 | Subject:Instrument) + (1 |
##     Subject:Harmony) + (1 | Subject:Voice), data = ratings, REML = FALSE)
##                  coef.est coef.se
## (Intercept)       3.11     1.10
## Instrumentpiano   1.53     0.28
## Instrumentstring  3.45     0.28
## HarmonyI-V-IV     0.03     0.17
## HarmonyI-V-VI     0.91     0.17
## HarmonyIV-I-V     0.11     0.17
## Voicecontrary     0.33     0.10
## Voicepar3rd      -0.06     0.10
## Selfdeclare      -0.35     0.17
## ConsInstr         0.08     0.12
## ConsNotes        -0.24     0.10
## PachListen        0.11     0.17
## ClsListen         0.25     0.09
## KnowRob           0.10     0.09
## KnowAxis          0.04     0.07
## X1990s2000s       0.16     0.09
## CollegeMusic     -0.76     0.34
## APTheory          0.36     0.39
## Composing         0.27     0.13
## PianoPlay         0.32     0.09
## GuitarPlay       -0.12     0.13
##
## Error terms:
##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.64
##  Subject:Voice      (Intercept) 0.18
##  Subject:Instrument (Intercept) 1.25
##  Residual                       1.57
## ---
## number of obs: 1649, groups: Subject:Harmony, 184; Subject:Voice, 138; Subject:Instrument, 1
## AIC = 6669, DIC = 6619
## deviance = 6619.0
```

We see the following effects for the variables we added to the model:

nice summary

- Selfdeclare: Increasing identification as a musician leads to *decreasing* rating of music as classical. This is interesting. Perhaps people who don't consider themselves musicians tend to have a broader definition of "classical". However the standard error appears a bit large.

9

- ConsInstr: The amount a person concentrates on an instrument has barely any effect on "classicalness", and the standard error is larger than the coefficient.
- ConsNotes: The amount a person concentrates on notes has a decreasing effect on classicalness, which seems unusual. Maybe people see classical music as complex, and more memorable notes seem too simple to them.
- PachListen: Appears to have a very small effect and has a larger standard error than coefficient.
- ClsListen: Has an increasing effect on classical rating. This could be because people who listen to classical music are more likely to notice classical patterns in other music.
- KnowRob: Also does not appear to have much of an impact due to having a standard error almost the size of the coefficient.
- KnowAxis: Bigger standard error than coefficient, so we will not interpret it.
- X1990s2000s: Appears to have a slight positive effect, but not much of an overall effect. Maybe people who listen to a lot of pop and rock have a more stereotypical view of what classical music is, but the effect size does not seem particularly big.
- CollegeMusic: People who took music classes in college are LESS likely to identify music as classical than people who did not. This is quite interesting, since I would expect that people with college music classes would be more easily able to pick classical elements out of non-classical songs. Maybe college music classes aren't classically-focused?
- APTheory: People who took AP Music Theory are more likely to rate music as classical than people who did not. It is interesting that this effect has the opposite sign as CollegeMusic did, but note that the standard error is bigger than the coefficient, so this may have the opposite sign in another experiment.
- Composing: The more an individual composes, the more likely that person is to rate music as classical. This makes sense, since they probably have more knowledge of classical elements than people who compose less.
- PianoPlay: The more someone plays piano, the more likely that person is to rate music as classical. This makes sense for the same reason as for composing.
- GuitarPlay: The more someone plays guitar, the less likely that person is to rate music as classical. However the effect size seems quite low and may not be real because the coefficient's standard error is bigger than it.

3)
```
table(ratings$Selfdeclare)

##
##   1   2   3   4   5   6
## 564 935 460 426  72  36

564 / 2493

## [1] 0.2262335

(564 + 935) / 2493

## [1] 0.6012836
```

Dichotimizing into "musician" or "Non-musician" is quite difficult, as 1499 out of 2493 individuals with recorded ratings, which is 60.1% of the dataset, rated themselves a 1 or a 2. This is the closest we can get to 50%, so we will dichotimize into `isMusician` by declaring 1s and 2s as non-musicians (coded as 0), and 3+ as musicians (coded as 1).

```r
ratings$isMusician <- ifelse(ratings$Selfdeclare <= 2, 0, 1)
sum(ratings$isMusician) / 2493
```

```
## [1] 0.3987164
```

Now about 40% of the dataset consists of musicians.

What should we consider interactions with? We will definitely consider interaction with how much someone plays piano and how much someone plays guitar, as we suspect that pianists have more experience with classical music and guitarists with rock and pop. We'll also look at interactions with how much classical music someone listens to for obvious reasons, and how much someone composes. We won't take any predictors out of the last model because even if they are not large or have large standard errors, their interactions may be interesting.

```r
lmer.3 <- lmer(Classical ~ Instrument + Harmony + Voice + isMusician + ConsInstr + ConsNotes +
                PachListen + ClsListen + KnowRob + KnowAxis + X1990s2000s + CollegeMusic + APTheory
                Composing + PianoPlay + GuitarPlay + isMusician:PianoPlay + isMusician:GuitarPlay +
                isMusician:ClsListen + isMusician:Composing +
                             (1 | Subject:Instrument) + (1 | Subject:Harmony) +
                             (1 | Subject:Voice),
                      data = ratings, REML = FALSE)
display(lmer.3)
```

```
## lmer(formula = Classical ~ Instrument + Harmony + Voice + isMusician +
##     ConsInstr + ConsNotes + PachListen + ClsListen + KnowRob +
##     KnowAxis + X1990s2000s + CollegeMusic + APTheory + Composing +
##     PianoPlay + GuitarPlay + isMusician:PianoPlay + isMusician:GuitarPlay +
##     isMusician:ClsListen + isMusician:Composing + (1 | Subject:Instrument) +
##     (1 | Subject:Harmony) + (1 | Subject:Voice), data = ratings,
##     REML = FALSE)
##                       coef.est coef.se
## (Intercept)            1.64     1.04
## Instrumentpiano        1.53     0.26
## Instrumentstring       3.45     0.26
## HarmonyI-V-IV          0.03     0.17
## HarmonyI-V-VI          0.91     0.17
## HarmonyIV-I-V          0.11     0.17
## Voicecontrary          0.34     0.10
## Voicepar3rd           -0.06     0.10
## isMusician             0.63     0.51
## ConsInstr              0.02     0.11
## ConsNotes             -0.31     0.09
## PachListen             0.25     0.17
## ClsListen              0.13     0.12
## KnowRob                0.07     0.10
## KnowAxis               0.01     0.07
## X1990s2000s            0.23     0.10
## CollegeMusic          -0.74     0.37
## APTheory               0.45     0.45
## Composing              0.49     0.26
## PianoPlay              0.71     0.24
## GuitarPlay             0.12     0.56
## isMusician:PianoPlay  -0.68     0.26
## isMusician:GuitarPlay -0.33     0.59
```

again, no variable selection?

6

```
## isMusician:ClsListen   0.26     0.24
## isMusician:Composing  -0.47     0.34
##
## Error terms:
##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.63
##  Subject:Voice      (Intercept) 0.18
##  Subject:Instrument (Intercept) 1.18
##  Residual                       1.58
## ---
## number of obs: 1649, groups: Subject:Harmony, 184; Subject:Voice, 138; Subject:Instrument, 138
## AIC = 6661.5, DIC = 6603.5
## deviance = 6603.5
```

```
sum(ratings$isMusician & ratings$PianoPlay) / 2493
```

```
## [1] 0.2719615
```

```
# P(piano | musician)
sum(ratings$isMusician & ratings$PianoPlay) / sum(ratings$isMusician)
```

```
## [1] 0.6820926
```

```
# P(musician | piano)
sum(ratings$isMusician & ratings$PianoPlay) / sum(ifelse(ratings$PianoPlay == 0, 0, 1))
```

```
## [1] 0.6538091
```

Now we have some interesting effects. Identification as a musician now has a positive coefficient, implying that people who identify as musicians are more likely to rate music as classical than non-musicians.

When we look at the interaction between whether or not someone is a musician and they play piano or guitar, both coefficients are negative (except for guitar, where the interaction term still has a very large standard error). Piano on its own is associated with a large increase, but self-identified musicians that play piano are associated with a decrease in rating that almost cancels out the main effect.

When we look at the dataset, we notice something VERY unusual: only 65% of individuals who play piano (when we dichotimize those who play piano to be about 50% of the dataset) identify as musicians, even though we would expect people who play an instrument to consider themselves musicians. There therefore seems to be a big disparity in how identification as a musician impacts how piano players view classical music. Those who identify as musicians have a lower increase per point of piano experience than those who do not identify as musicians.

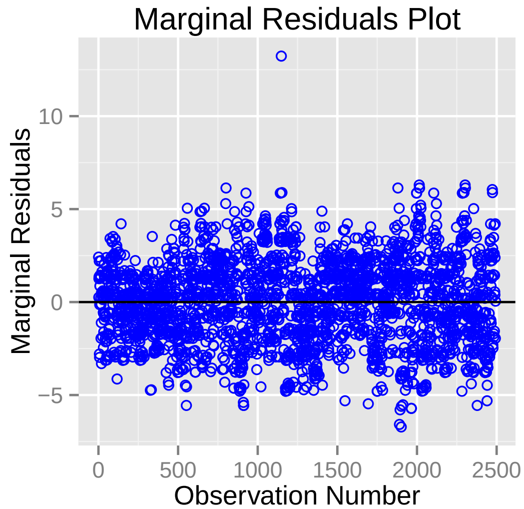
There is a similar discrepancy with the interaction term for self-identified musicians who compose music. The interaction effect almost cancels out the main effect for musicians, implying that while nonmusicians who compose tend to rate music an additional point higher for every point of composition, the effect practically disappears for musicians who compose.

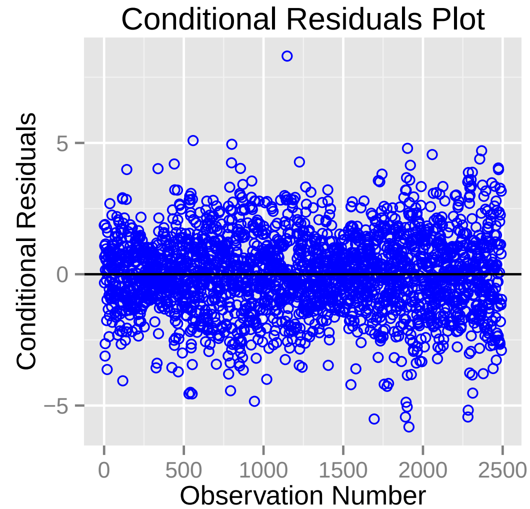
With regard to the other effects, listening to classical music and identifying as a musician is associated with an additional 0.26 point increase for every additional point of listening to classical music. This is larger than the main effect for listening to classical music, implying that people who listen to a lot of classical music and identify as musicians view music as more classical than people who do not play classical music.

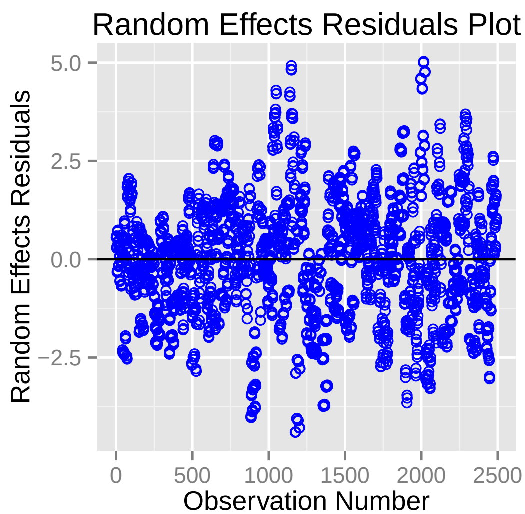
Although it wasn't asked for, I'll also note that this model has a lower AIC than the model without so many added fixed effects. The model with just the covariates of interest has an AIC of 6669, and this one has an AIC of 6662.

4) a)

```
lmer.4a <- lmer(Popular ~ Instrument + Harmony + Voice +
                             (1 | Subject:Instrument) + (1 | Subject:Harmony) +
                             (1 | Subject:Voice),
                      data = ratings, REML = FALSE)
display(lmer.4a)

## lmer(formula = Popular ~ Instrument + Harmony + Voice + (1 |
##     Subject:Instrument) + (1 | Subject:Harmony) + (1 | Subject:Voice),
##     data = ratings, REML = FALSE)
##                  coef.est coef.se
## (Intercept)       6.74     0.21
## Instrumentpiano  -0.95     0.25
## Instrumentstring -2.61     0.25
## HarmonyI-V-IV    -0.03     0.14
## HarmonyI-V-VI    -0.27     0.14
## HarmonyIV-I-V    -0.19     0.14
## Voicecontrary    -0.16     0.08
## Voicepar3rd       0.00     0.08
##
## Error terms:
##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.63
##  Subject:Voice      (Intercept) 0.17
##  Subject:Instrument (Intercept) 1.40
##  Residual                       1.58
## ---
## number of obs: 2493, groups: Subject:Harmony, 280; Subject:Voice, 210; Subject:Instrument, 2
## AIC = 10079, DIC = 10055
## deviance = 10055.0

plot.lmer.residuals(lmer.4a)
```







The fixed effects look smaller here than for the classical model, although they are in the opposite direction. Piano and string instruments are associated with a lower pop rating, both with similar magnitudes to classical rating. Presence of piano is associated with a roughly one-point decrease in popular compared to electic guitar, and string quartet with a 2.6-point decrease compared to
9 electric guitar. I-V-IV harmony is still not an important predictor (and has literally the same coefficient and standard deviation), and the other harmony coefficients are very small. Although I-V-VI has a fairly large effect on identification as classical, it does not have that much of an impact on identification as popular. Voice has an even lower impact on identification as popular, with contrary voice being less than fifth of a point less popular than 5th.

Residual plots for the hierarchical model seem to show that the model is a good fit. The marginal residuals appear to have mean 0, and the conditional residuals seem homoskedastic. The random effects residuals also look fine, with no outliers.

b)
```
lmer.4b <- lmer(Popular ~ Instrument + Harmony + Voice + Selfdeclare + ConsInstr + ConsNotes +
                PachListen + ClsListen + KnowRob + KnowAxis + X1990s2000s + CollegeMusic + APTI
                Composing + PianoPlay + GuitarPlay +
                           (1 | Subject:Instrument) + (1 | Subject:Harmony) +
                           (1 | Subject:Voice),
```

```
                      data = ratings, REML = FALSE)
display(lmer.4b)
```

```
## lmer(formula = Popular ~ Instrument + Harmony + Voice + Selfdeclare +
##     ConsInstr + ConsNotes + PachListen + ClsListen + KnowRob +
##     KnowAxis + X1990s2000s + CollegeMusic + APTheory + Composing +
##     PianoPlay + GuitarPlay + (1 | Subject:Instrument) + (1 |
##     Subject:Harmony) + (1 | Subject:Voice), data = ratings, REML = FALSE)
##                  coef.est coef.se
## (Intercept)       7.82     1.10
## Instrumentpiano  -1.13     0.28
## Instrumentstring -3.00     0.28
## HarmonyI-V-IV     0.02     0.18
## HarmonyI-V-VI    -0.31     0.18
## HarmonyIV-I-V    -0.26     0.18
## Voicecontrary    -0.24     0.11
## Voicepar3rd      -0.04     0.11
## Selfdeclare      -0.03     0.17
## ConsInstr        -0.01     0.12
## ConsNotes         0.13     0.10
## PachListen       -0.21     0.17
## ClsListen        -0.08     0.09
## KnowRob           0.04     0.09
## KnowAxis          0.06     0.07
## X1990s2000s      -0.04     0.09
## CollegeMusic      0.15     0.34
## APTheory         -0.06     0.39
## Composing         0.12     0.13
## PianoPlay        -0.04     0.09
## GuitarPlay        0.06     0.13
##
## Error terms:
##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.66
##  Subject:Voice      (Intercept) 0.19
##  Subject:Instrument (Intercept) 1.25
##  Residual                       1.63
## ---
## number of obs: 1649, groups: Subject:Harmony, 184; Subject:Voice, 138; Subject:Instrument, 1
## AIC = 6779.2, DIC = 6729.2
## deviance = 6729.2
```

The fixed effects in modeling "Popular" are, interesting enough, almost all completely useless. The largest additional coefficient we measured in magnitude has a value of -0.21, and it's whether or not someone knows about Pachelbel's Canon. It also has a large standard deviation, and many of the other coefficients have a standard deviation larger than the coefficient itself. The three main covariates are the best we have.

Meanwhile, the standard deviations of the random effects are almost identical to those in the classical model. The standard deviation of Subject:Harmony is 0.66 (compared to 0.64), the standard deviation of Subject:Voice is 0.19 (compared to 0.18), and the standard deviation of Subject:Instrument is 1.25 (which is identical to 1.25 in the Classical model). The residual standard deviation is 1.63 (compared to 1.57).

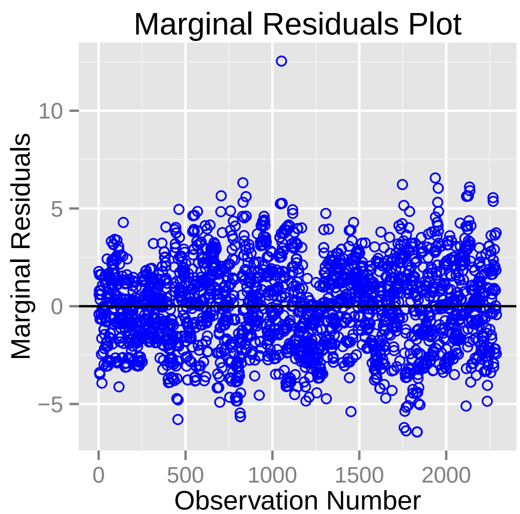
5 Because none of the added seem to help, I will continue onto Part c) without them. However because I hypothesize that covariates related to pop music probably have an interaction effect,

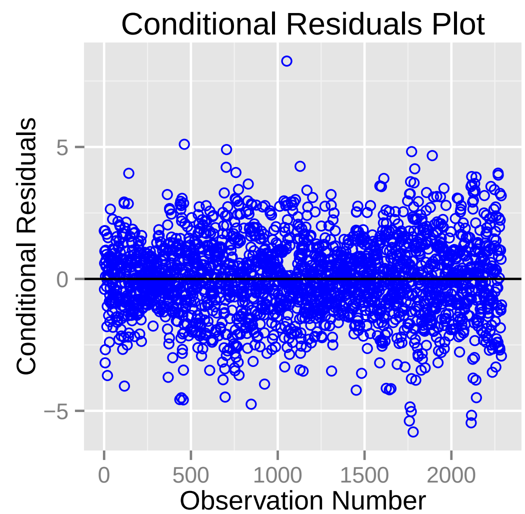
I'll include an extra interaction effect for X1990s2000s.

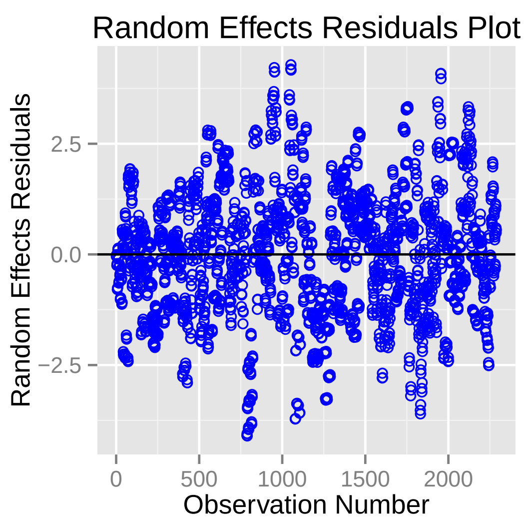
c)
```
lmer.4c <- lmer(Popular ~ Instrument + Harmony + Voice + isMusician + ClsListen +
                          X1990s2000s + Composing + PianoPlay + GuitarPlay +
                isMusician:PianoPlay + isMusician:GuitarPlay + isMusician:ClsListen +
                isMusician:Composing + isMusician:X1990s2000s +
                            (1 | Subject:Instrument) + (1 | Subject:Harmony) +
                            (1 | Subject:Voice),
                      data = ratings, REML = FALSE)
display(lmer.4c)
```

```
## lmer(formula = Popular ~ Instrument + Harmony + Voice + isMusician +
##     ClsListen + X1990s2000s + Composing + PianoPlay + GuitarPlay +
##     isMusician:PianoPlay + isMusician:GuitarPlay + isMusician:ClsListen +
##     isMusician:Composing + isMusician:X1990s2000s + (1 | Subject:Instrument) +
##     (1 | Subject:Harmony) + (1 | Subject:Voice), data = ratings,
##     REML = FALSE)
##                        coef.est coef.se
## (Intercept)             5.77     0.57
## Instrumentpiano        -1.02     0.25
## Instrumentstring       -2.75     0.25
## HarmonyI-V-IV          -0.02     0.15
## HarmonyI-V-VI          -0.29     0.15
## HarmonyIV-I-V          -0.22     0.15
## Voicecontrary          -0.19     0.09
## Voicepar3rd             0.00     0.09
## isMusician              0.96     0.87
## ClsListen               0.06     0.11
## X1990s2000s             0.21     0.11
## Composing              -0.38     0.19
## PianoPlay               0.03     0.12
## GuitarPlay             -0.24     0.43
## isMusician:PianoPlay    0.02     0.15
## isMusician:GuitarPlay   0.24     0.45
## isMusician:ClsListen   -0.15     0.18
## isMusician:Composing    0.62     0.25
## isMusician:X1990s2000s -0.20     0.15
##
## Error terms:
##  Groups             Name        Std.Dev.
##  Subject:Harmony    (Intercept) 0.64
##  Subject:Voice      (Intercept) 0.18
##  Subject:Instrument (Intercept) 1.32
##  Residual                       1.57
## ---
## number of obs: 2289, groups: Subject:Harmony, 256; Subject:Voice, 192; Subject:Instrument, 
## AIC = 9247.8, DIC = 9201.8
## deviance = 9201.8
```

```
plot.lmer.residuals(lmer.4c)
```







This is interesting. When we exclude many of the other covariates, the scale rating of how much 1990s and 2000s pop subjects listen to becomes slightly positive, leading to a quarter of a point increase in "poppiness" per scale increase. Whether or not someone is a musician appears to have a large effect, with an increase of almost a full point in poppiness compared to non-musicians,
9 however the associated coefficient has a standard error almost equal to itself, so perhaps the effect is not as large as it seems. Composing is, interestingly enough, negative for self-identified non-musicians and positive for musicians. Maybe this is because the study has more musicians who play pop music than classical music, so they have a better ear for pop.

None of the other interaction terms, including musician combined with how much pop someone listens to, appears that important, since their standard errors are either larger than or equal to the coefficients in magnitude. But whether or not someone is a musician appears to be an important predictor only when we look at it in tandem with other covariates.

5) We began our analysis by excluding all individuals missing final ratings. We then built a model to predict classical rating and determined based on a low AIC that a model treating the interaction between subject and each of instrument, harmony, and voice was the best fit.

   In this model, we found that instrument is easily the most influential of the three factors. Piano is associated with an increase of about 1.4 points higher than electric guitar and string quartet is associated with an increase of about 3.1 points higher than guitar. With regard to harmony, we found that I-V-VI, is associated with an increase in classical rating of about 3/4 of a point, but the other harmonies are not associated with notable differences. Contrary voice was associated with an increase of 1/3 of a point, while none of the other voicings were associated with notable effects.

   The interaction model's variance components were quite large. Subject on harmony had a standard deviation of 0.66, while subject on instrument had a standard deviation of 1.47. Subject on voice had a particularly small standard deviation of 0.16. These standard deviations imply that subjects vary in their conceptions of harmony and voice by a reasonable amount, but a residual variance of 1.56 implies that more factors are needed than just the basic repeated measures model.

   We examined numerous other fixed covariates on subject ratings, none of which impacted the variance of the random effects. Despite suspecting that individuals who had heard of Pachelbel's Canon and the associated bits on it, none of these factors associated an individual's rating, even though the I-V-VI progression was perceived as classical. We found an increasing relationship in classical rating with listening to classical music, composing, and playing piano. Interestingly enough, we found a *decreasing* relationship with classical rating for people who took music classes in college. Perhaps college music classes expose students to other kinds of music with similar characteristics to classical music, resulting in students associating "classicality" with these other types of music. We also found a slight negative association between identification as a musician and classical rating.

   We also found that how strongly someone identifies as a musician completely changes some of the relationships observed. Before performing this analysis on the same covariates, we discretized our
10 musician scale into "musician" and "non-musician" by declaring those who rated themselves a 1 or a 2 "non-musician", and those who rated themselves a 3 or above a "musician". We found some interesting effects, including that discretizing musician results in a positive coefficient. However the interaction effects for musicians who play piano and musicians who play guitar are both negative. Similarly, individuals who play classical music and identify as musicians view music as more classical than those who do not.

   Playing the piano for nonmusicians is associated with an increase in 0.71 in classical rating, however playing piano for musicians is associated with an increase of just 0.03 points, virtually removing the effect. This seemed quite unusual, so we tabulated self-identified musicians and self-identified piano players and determined that when we dichotimize piano such that about 50% of individuals are piano players and 50% are not, only 65% of self-identified piano players identify as musicians. We suspect this is because "musician" is also a profession, and some people do not consider themselves musicians because they are not professional musicians.

   There were far fewer factors involved in predicting high popular ratings than high classical ratings. When we built a random interaction model with the same covariates to predict popular rating, we found that virtually no covariates were predictive of it except for instrument. Piano is associated with a decrease in 1.1 point compared to electric guitar, and string quartet is associated with a decrease in 3.0 points compared to electric guitar. Variance components were virtually identical to those in the classical model. When we built the model involving musicians, we found that musicians who compose have lower pop ratings than musicians who do not compose. We suspect this is because the study has more musicians who play pop music than musicians who play classical music.