hw7-Linear Mixed Effects model case study

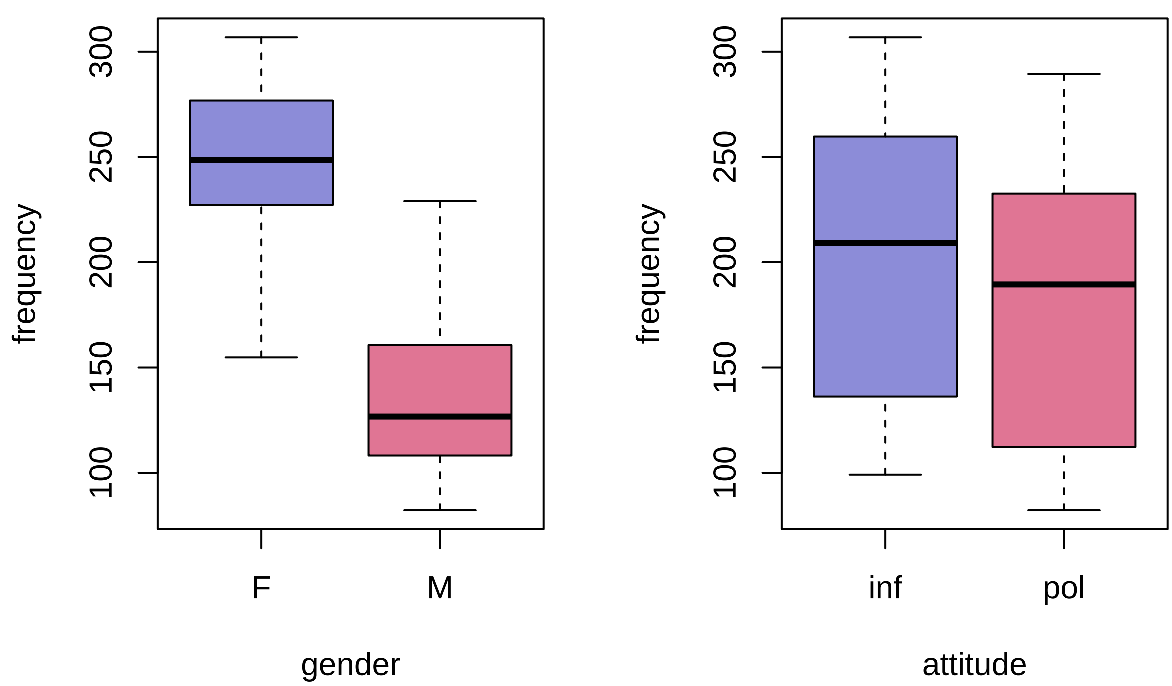
Cary Ni

```r
# load dataset and modify variable types
data_df = read_csv("./data/HW7-politeness_data.csv", show_col_types = FALSE)
data_df[, c(1:4)]= lapply(data_df[, c(1:4)], as.factor)
```

(a)

```r
# simple visualization of the categorical data
par(mfrow=c(1,2))
myColors = c(rgb(0.1,0.1,0.7,0.5) , rgb(0.8,0.1,0.3,0.6))

boxplot(frequency ~ gender, data = data_df, xlab = "gender", col = myColors)
boxplot(frequency ~ attitude, data = data_df, xlab = "attitude", col = myColors)
```



The boxplots show that female is associated with remarkably higher frequency than male while the informal register has a frequency that is only slightly higher than formal register.

(b)

```
model_1 = lme(frequency ~ gender + attitude, random = ~1 | subject, data = data_df)
summary(model_1)
```

```
## Linear mixed-effects model fit by REML
##   Data: data_df
##        AIC      BIC    logLik
##   806.0805 818.0527 -398.0402
##
## Random effects:
##  Formula: ~1 | subject
##         (Intercept) Residual
## StdDev:    24.45803 29.11537
##
## Fixed effects:  frequency ~ gender + attitude
##                  Value Std.Error DF   t-value p-value
## (Intercept)  256.98690 15.154986 77 16.957251  0.0000
## genderM     -108.79762 20.956235  4 -5.191659  0.0066
## attitudepol  -20.00238  6.353495 77 -3.148248  0.0023
##  Correlation:
##             (Intr) gendrM
## genderM     -0.691
## attitudepol -0.210  0.000
##
## Standardized Within-Group Residuals:
##        Min         Q1        Med         Q3        Max
## -2.3564422 -0.5658319 -0.2011979  0.4617895  3.2997610
##
## Number of Observations: 84
## Number of Groups: 6
```

Covariance Matrix for Yi

```
extract.lme.cov2(model_1, data_df)$V[[1]]
```

```
##            [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]
##  [1,] 1445.9002  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
##  [2,]  598.1953 1445.9002  598.1953  598.1953  598.1953  598.1953  598.1953
##  [3,]  598.1953  598.1953 1445.9002  598.1953  598.1953  598.1953  598.1953
##  [4,]  598.1953  598.1953  598.1953 1445.9002  598.1953  598.1953  598.1953
##  [5,]  598.1953  598.1953  598.1953  598.1953 1445.9002  598.1953  598.1953
##  [6,]  598.1953  598.1953  598.1953  598.1953  598.1953 1445.9002  598.1953
##  [7,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953 1445.9002
##  [8,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
##  [9,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
```

```
## [10,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
## [11,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
## [12,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
## [13,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
## [14,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
##           [,8]      [,9]     [,10]     [,11]     [,12]     [,13]     [,14]
##  [1,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
##  [2,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
##  [3,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
##  [4,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
##  [5,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
##  [6,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
##  [7,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
##  [8,] 1445.9002  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953
##  [9,]  598.1953 1445.9002  598.1953  598.1953  598.1953  598.1953  598.1953
## [10,]  598.1953  598.1953 1445.9002  598.1953  598.1953  598.1953  598.1953
## [11,]  598.1953  598.1953  598.1953 1445.9002  598.1953  598.1953  598.1953
## [12,]  598.1953  598.1953  598.1953  598.1953 1445.9002  598.1953  598.1953
## [13,]  598.1953  598.1953  598.1953  598.1953  598.1953 1445.9002  598.1953
## [14,]  598.1953  598.1953  598.1953  598.1953  598.1953  598.1953 1445.9002
```

Covariance matrix of fixed effects

```
vcov(model_1)
```

```
##             (Intercept)       genderM   attitudepol
## (Intercept)   229.67362 -2.195819e+02 -2.018345e+01
## genderM      -219.58189  4.391638e+02  6.451438e-15
## attitudepol   -20.18345  6.451438e-15  4.036690e+01
```

BLUP for intercepts

```
random.effects(model_1)
```

```
##    (Intercept)
## F1  -13.575831
## F2   10.170522
## F3    3.405309
## M3   27.960288
## M4    4.739325
## M7  -32.699613
```

Residuals

```
data_df$frequency-fitted(model_1)
```

```
##          F1          F1          F1          F1          F1          F1
## -10.1086926 -38.9110735  61.6913074  16.2889265 -19.5086926  43.4889265
##          F1          F1          F1          F1          F1          F1
##  27.3913074  33.3889265   8.4913074   8.9889265 -42.2086926 -12.7110735
##          F1          F1          F3          F3          F3          F3
## -26.9110735 -68.6086926 -10.6898326 -23.0922136  -3.5898326  -9.3922136
##          F3          F3          F3          F3          F3          F3
##  26.6101674   5.6077864  35.0101674  46.4077864  -7.7898326  -7.8922136
##          F3          F3          F3          F3          M4          M4
## -13.8898326  18.4077864   4.0077864 -54.8898326 -22.2262298 -29.3286108
##          M4          M4          M4          M4          M4          M4
##  96.0737702 -38.0286108 -20.7262298  60.6713892  60.4737702   9.9713892
##          M4          M4          M4          M4          M4          M4
## -31.1262298 -26.0286108 -22.9262298 -16.7286108  -6.9286108  -6.4262298
##          M7          M7          M7          M7          M7          M7
##  -9.3872916 -16.3896725 -13.2872916 -11.1896725  -9.5872916  -5.2896725
##          M7          M7          M7          M7          M7          M7
##   1.6127084   4.5103275  -1.7872916 -12.5896725  13.3127084  -7.2896725
##          M7          M7          F2          F2          F2          F2
##   8.9103275  12.1127084 -14.4550462 -35.8574271  -0.8550462  -7.4574271
##          F2          F2          F2          F2          F2          F2
##  42.2449538  34.6425729  -3.9550462  29.0425729  30.5449538  27.0425729
##          F2          F2          F2          F2          M3          M3
## -39.1550462 -41.2574271  13.8425729 -19.9550462  -2.3471929  12.6504261
##          M3          M3          M3          M3          M3          M3
## -13.7471929  23.5504261   4.0528071   9.9504261  51.3528071  14.7504261
##          M3          M3          M3          M3          M3          M3
##   4.5528071 -19.6495739  -9.4471929 -18.1495739 -15.0495739  -2.8471929
## attr(,"label")
## [1] "Fitted values"
```

```
model_1$residuals
```

```
##          fixed     subject
## 1  -23.6845238 -10.1086926
## 2  -52.4869048 -38.9110735
## 3   48.1154762  61.6913074
## 4    2.7130952  16.2889265
## 5  -33.0845238 -19.5086926
## 6   29.9130952  43.4889265
## 7   13.8154762  27.3913074
## 8   19.8130952  33.3889265
## 9   -5.0845238   8.4913074
## 10  -4.5869048   8.9889265
## 11 -55.7845238 -42.2086926
## 12 -26.2869048 -12.7110735
## 13 -40.4869048 -26.9110735
## 14 -82.1845238 -68.6086926
## 15  -7.2845238 -10.6898326
## 16 -19.6869048 -23.0922136
## 17  -0.1845238  -3.5898326
## 18  -5.9869048  -9.3922136
## 19  30.0154762  26.6101674
## 20   9.0130952   5.6077864
```

```
## 21  38.4154762  35.0101674
## 22  49.8130952  46.4077864
## 23  -4.3845238  -7.7898326
## 24  -4.4869048  -7.8922136
## 25 -10.4845238 -13.8898326
## 26  21.8130952  18.4077864
## 27   7.4130952   4.0077864
## 28 -51.4845238 -54.8898326
## 29 -17.4869048 -22.2262298
## 30 -24.5892857 -29.3286108
## 31 100.8130952  96.0737702
## 32 -33.2892857 -38.0286108
## 33 -15.9869048 -20.7262298
## 34  65.4107143  60.6713892
## 35  65.2130952  60.4737702
## 36  14.7107143   9.9713892
## 37 -26.3869048 -31.1262298
## 38 -21.2892857 -26.0286108
## 39 -18.1869048 -22.9262298
## 40 -11.9892857 -16.7286108
## 41  -2.1892857  -6.9286108
## 42  -1.6869048  -6.4262298
## 43 -42.0869048  -9.3872916
## 44 -49.0892857 -16.3896725
## 45 -45.9869048 -13.2872916
## 46 -43.8892857 -11.1896725
## 47 -42.2869048  -9.5872916
## 48 -37.9892857  -5.2896725
## 49 -31.0869048   1.6127084
## 50 -28.1892857   4.5103275
## 51 -34.4869048  -1.7872916
## 52 -45.2892857 -12.5896725
## 53 -19.3869048  13.3127084
## 54 -39.9892857  -7.2896725
## 55 -23.7892857   8.9103275
## 56 -20.5869048  12.1127084
## 57  -4.2845238 -14.4550462
## 58 -25.6869048 -35.8574271
## 59   9.3154762  -0.8550462
## 60   2.7130952  -7.4574271
## 61  52.4154762  42.2449538
## 62  44.8130952  34.6425729
## 63   6.2154762  -3.9550462
## 64  39.2130952  29.0425729
## 65  40.7154762  30.5449538
## 66  37.2130952  27.0425729
## 67 -28.9845238 -39.1550462
## 68 -31.0869048 -41.2574271
## 69  24.0130952  13.8425729
## 70  -9.7845238 -19.9550462
## 71  25.6130952  -2.3471929
## 72  40.6107143  12.6504261
## 73  14.2130952 -13.7471929
## 74  51.5107143  23.5504261
```

```
## 75  32.0130952   4.0528071
## 76  37.9107143   9.9504261
## 77  79.3130952  51.3528071
## 78  42.7107143  14.7504261
## 79  32.5130952   4.5528071
## 80   8.3107143 -19.6495739
## 81  18.5130952  -9.4471929
## 82   9.8107143 -18.1495739
## 83  12.9107143 -15.0495739
## 84  25.1130952  -2.8471929
## attr(,"std")
##  [1] 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537
##  [9] 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537
## [17] 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537
## [25] 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537
## [33] 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537
## [41] 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537
## [49] 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537
## [57] 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537
## [65] 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537
## [73] 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537 29.11537
## [81] 29.11537 29.11537 29.11537 29.11537
```

(c)

```r
# Rebuild the model_1 with ML
model_2 = lme(frequency ~ gender + attitude, random = ~1 | subject, data = data_df, method = "ML")
model_3 = lme(frequency ~ gender*attitude, random = ~1 | subject, data = data_df, method = "ML")
# Likelihood ratio test
anova(model_2, model_3)
```

```
##         Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## model_2     1  5 825.6363 837.7904 -407.8182
## model_3     2  6 826.2508 840.8357 -407.1254 1 vs 2 1.385523  0.2392
```

The likelihood ratio test gives a p value of 0.239, which fails to reject the null hypothesis. Therefore, we will take the small model and believe that the interaction term between gender and attitude is not significantly associated with pitch.

(d)

Mixed effects Model

$$Y_{ij} \ = \beta_0 + X^T\beta_1 + b_i + b_m + \epsilon_{ij}$$

$$b_i \sim N(0, \sigma_i^2) b_m \sim N(0, \sigma_m^2) \epsilon_{ij} \sim N(0, \sigma^2)$$

Note: b_i is the random intercept introduced by `subjects`, b_m is the random intercept introduced by `scenarios` (m scenarios in total). X^T is the fixed effects of `gender` and `attitude`.

```r
# Build the model
model_4 = lmer(frequency ~ gender + attitude + (1 | subject) + (1 | scenario), data = data_df)
summary(model_4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: frequency ~ gender + attitude + (1 | subject) + (1 | scenario)
##    Data: data_df
##
## REML criterion at convergence: 784.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.2690 -0.6331 -0.0878  0.5204  3.5326
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  scenario (Intercept) 224.5    14.98
##  subject  (Intercept) 613.2    24.76
##  Residual             637.8    25.25
## Number of obs: 84, groups:  scenario, 7; subject, 6
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  256.987     16.101  15.961
## genderM     -108.798     20.956  -5.192
## attitudepol  -20.002      5.511  -3.630
##
## Correlation of Fixed Effects:
##             (Intr) gendrM
## genderM     -0.651
## attitudepol -0.171  0.000
```

Covariance Matrix for Y_i

```r
# get variance of random effects
VarCorr(model_4)
```

```
##  Groups   Name        Std.Dev.
##  scenario (Intercept) 14.983
##  subject  (Intercept) 24.763
##  Residual             25.254
```

```r
# build the covariance matrix for Yi
knitr::include_graphics("./data/corr matrix.png")
```

$$
\hat{var}(y_i) = \begin{pmatrix}
1476 & \dots & 613 & 838 & \dots & 613 \\
\vdots & 1476 & \vdots & \vdots & 838 & \vdots \\
613 & \dots & 1476 & 613 & \dots & 838 \\
838 & \dots & 613 & 1476 & \dots & 613 \\
\vdots & 838 & \vdots & \vdots & 1476 & \vdots \\
613 & \dots & 838 & 613 & \dots & 1476
\end{pmatrix}
$$

Coefficient of attitude

```
fixed.effects(model_4)
```

```
## (Intercept)     genderM attitudepol
##   256.98690  -108.79762   -20.00238
```

The coefficient of `attitude` indicates that talking in formal register is expected to have a 20 Hz decrease in frequency compared to informal register while holding gender fixed.