

P8131_HW7

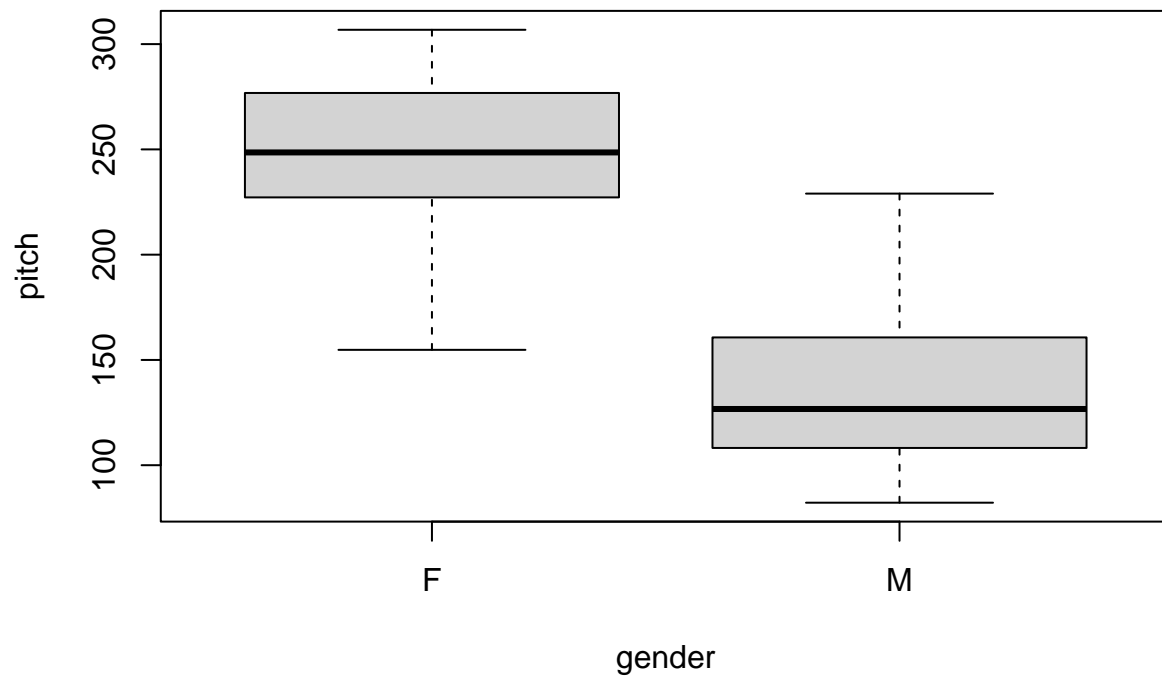
Zirui Zhang (zz3039)

2023-04-11

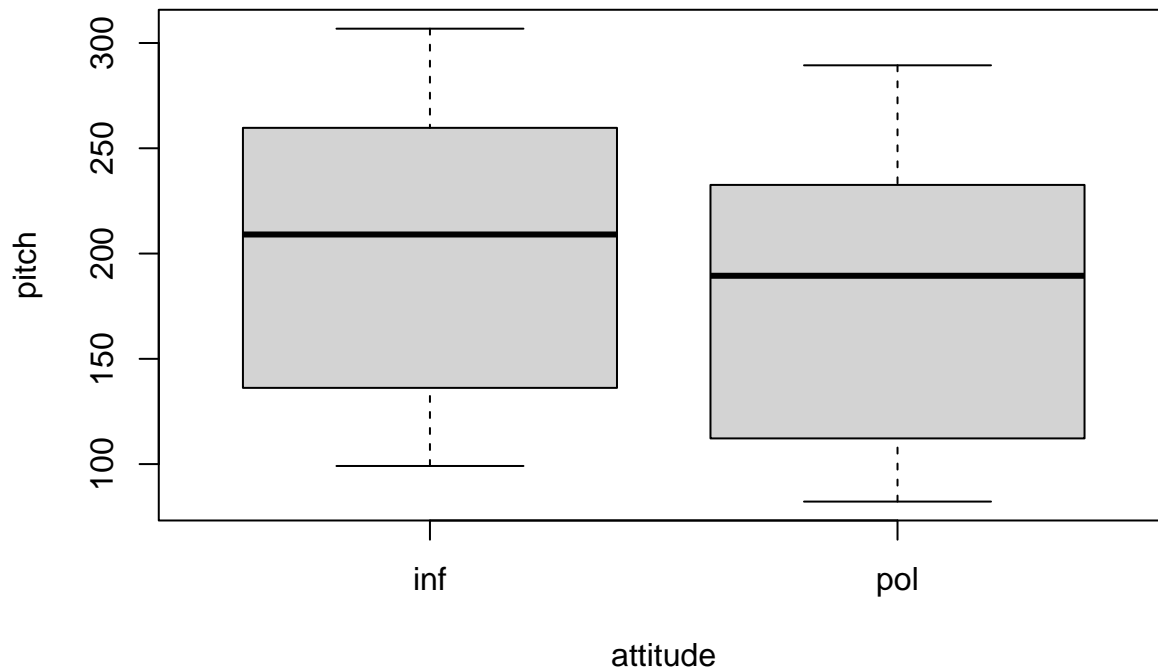
```
library(readr)
library(tidyverse)
library(dplyr)
library(nlme)
library(lme4)
```

Question (a)

```
# load data
pl = read.csv("./HW7-politeness_data.csv")
attach(pl)
# boxplot
boxplot(frequency~gender, xlab="gender", ylab="pitch")
```



```
boxplot(frequency~attitude, xlab="attitude", ylab="pitch")
```



Generally female tend to higher pitch than male, and inf attitude has slight higher pitch than pol attitude (ignoring scenarios).

Question (b)

```
# fit LMM with random intercept
lmm1 = lme(frequency ~ gender+attitude, random=~1 | subject, data=pl, method="REML" )
summary(lmm1)
```

```
## Linear mixed-effects model fit by REML
##   Data: pl
##       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
```

The covariance matrix for a subject Y_i is:

$$\text{cov}(Y_i) = \begin{bmatrix} \sigma^2 + \sigma_b^2 & \sigma_b^2 & \dots & \sigma_b^2 \\ \sigma_b^2 & \sigma^2 + \sigma_b^2 & \dots & \sigma_b^2 \\ \dots & \dots & \dots & \dots \\ \sigma_b^2 & \sigma_b^2 & \dots & \sigma^2 + \sigma_b^2 \end{bmatrix}$$

where

$$\sigma^2 = 847.7049, \sigma_b^2 = 598.1953$$

```
# covariance matrix for estimates of fixed effects
vcov(lmm1)
```

```
##          (Intercept)      genderM  attitudepol
## (Intercept)   229.67362 -2.195819e+02 -2.018345e+01
## genderM       -219.58189  4.391638e+02  7.288702e-15
## attitudepol   -20.18345  7.288702e-15  4.036690e+01
```

```
# BLUPs for subject-specific intercepts
random.effects(lmm1)
```

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

```
# residuals
pl$frequency-fitted(lmm1)
```

```
##      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"
```

Question (c)

```
lmm2.1 = lme(frequency ~ gender+attitude, random=~1|subject, data=pl, method="ML" )
lmm2.2 = lme(frequency ~ gender+attitude+gender*attitude, random=~1|subject, data=pl, method="ML" )
anova(lmm2.1, lmm2.2)
```

```
## Model df AIC BIC logLik Test L.Ratio p-value
## lmm2.1 1 5 825.6363 837.7904 -407.8182
## lmm2.2 2 6 826.2508 840.8357 -407.1254 1 vs 2 1.385523 0.2392
```

The P-value is greater than 0.05, thus we fail to reject the null and use the smaller model without the interaction term.

Question (d)

```
# fit LMM with random intercepts for subject and scenario
lmm3 = lmer(frequency ~ gender + attitude + (1|subject) + (1|scenario), data=pl)
summary(lmm3)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: frequency ~ gender + attitude + (1 | subject) + (1 | scenario)
## Data: pl
##
## 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

```

The covariance matrix for a subject Y_i is:

$$cov(Y_i) = \begin{bmatrix} \sigma^2 + \sigma_{b_1}^2 + \sigma_{b_2}^2 & \sigma_{b_1}^2 + \sigma_{b_2}^2 & \dots & \sigma_{b_1}^2 + \sigma_{b_2}^2 \\ \sigma_{b_1}^2 + \sigma_{b_2}^2 & \sigma^2 + \sigma_{b_1}^2 + \sigma_{b_2}^2 & \dots & \sigma_{b_1}^2 + \sigma_{b_2}^2 \\ \dots & \dots & \dots & \dots \\ \sigma_{b_1}^2 + \sigma_{b_2}^2 & \sigma_{b_1}^2 + \sigma_{b_2}^2 & \dots & \sigma^2 + \sigma_{b_1}^2 + \sigma_{b_2}^2 \end{bmatrix}$$

where

$$\sigma^2 = 637.78, \sigma_{b_1}^2 = 613.19, \sigma_{b_2}^2 = 224.5$$

Interpretation of the coefficient for the fixed effect term attitude: keeping scenario fixed, for a subject with pol attitude, the frequency decrease by -20.0024% compared to he or she with an inf attitude.