

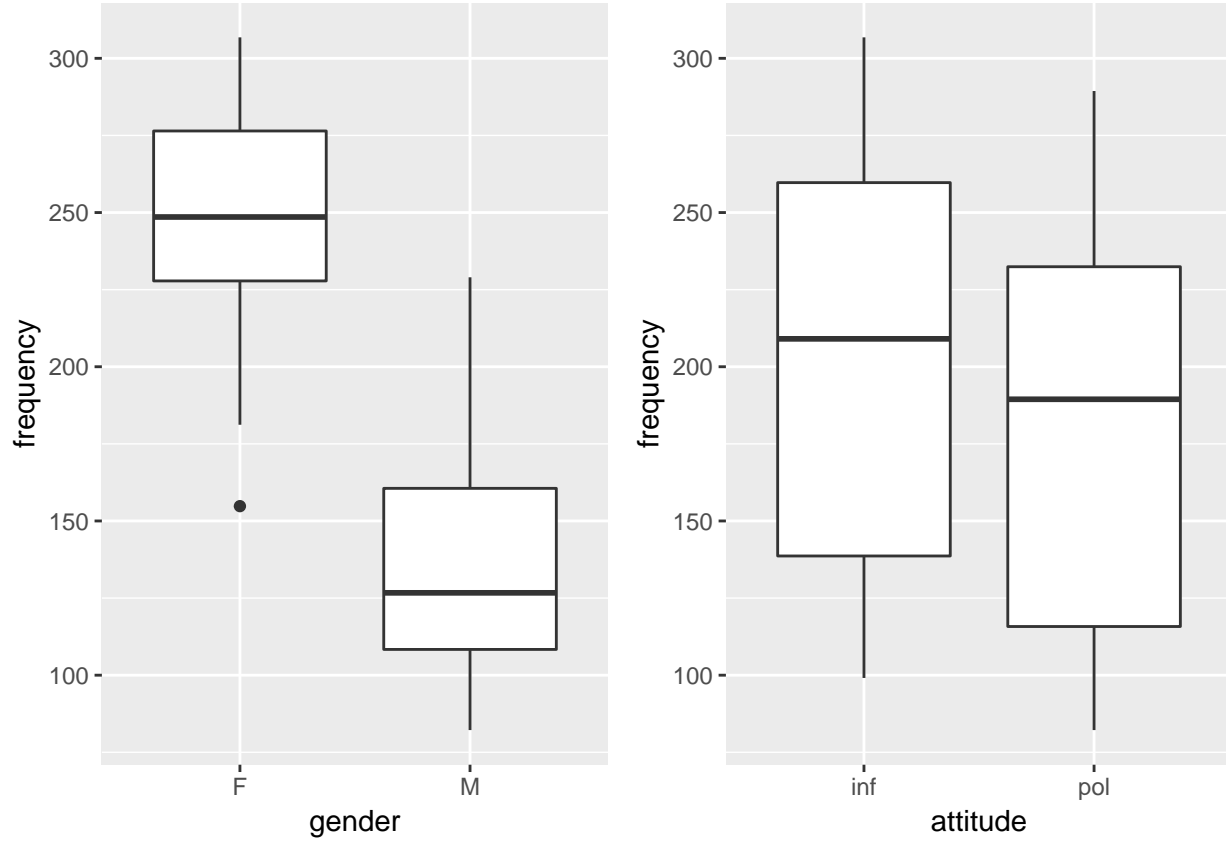
# HW7\_\_answer

Guojing Wu

4/15/2019

## Problem

### 1) Exploratory analysis



The boxplot shows that there are differences of pitch between male and female and between formal and informal

### 2) Fixed effect model with random intercept

The covariance matrix for a subject  $Y_i$  is

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

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

And the covariance matrix for REML is:

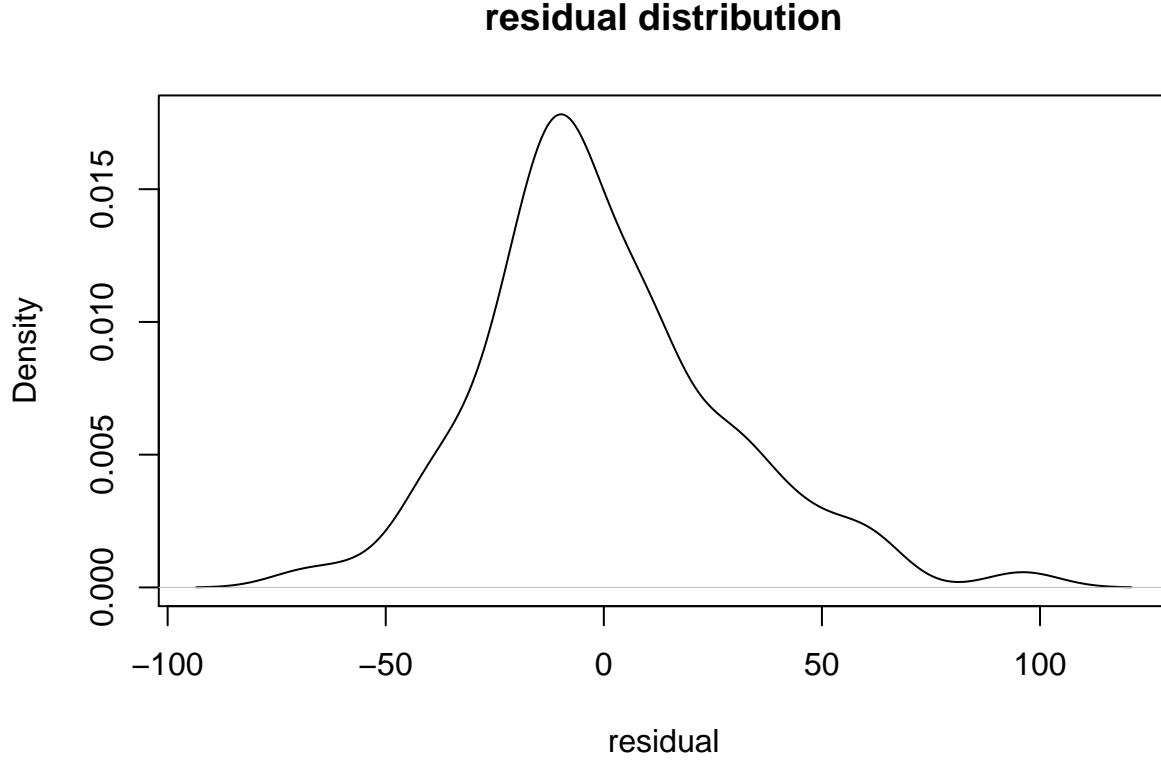
	(Intercept)	genderM	attitudepol
(Intercept)	229.67362	-219.5819	-20.18345
genderM	-219.58189	439.1638	0.00000
attitudepol	-20.18345	0.0000	40.36690

The BLUPs for subject-specific intercept

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

The residuals are:

	F1	F2	F3	M3	M4	M7
	-10.108693	-10.689833	-22.226230	-9.387292	-14.4550462	-2.347193
	-38.911074	-23.092214	-29.328611	-16.389672	-35.8574271	12.650426
	61.691307	-3.589833	96.073770	-13.287292	-0.8550462	-13.747193
	16.288926	-9.392214	-38.028611	-11.189673	-7.4574271	23.550426
	-19.508693	26.610167	-20.726230	-9.587292	42.2449538	4.052807
	43.488926	5.607786	60.671389	-5.289672	34.6425729	9.950426
	27.391307	35.010167	60.473770	1.612708	-3.9550462	51.352807
	33.388926	46.407786	9.971389	4.510327	29.0425729	14.750426
	8.491307	-7.789833	-31.126230	-1.787292	30.5449538	4.552807
	8.988926	-7.892214	-26.028611	-12.589673	27.0425729	-19.649574
	-42.208693	-13.889833	-22.926230	13.312708	-39.1550462	-9.447193
	-12.711073	18.407786	-16.728611	-7.289672	-41.2574271	-18.149574
	-26.911074	4.007786	-6.928611	8.910327	13.8425729	-15.049574
	-68.608693	-54.889833	-6.426230	12.112708	-19.9550462	-2.847193



### 3) Fixed effects model with intercepts for different subjects.

For likelihood ratio test, we use ML instead of REML, and then doing ANOVA. The pvalue = 0.2392, so in this case, we fail to reject the null hypothesis and state that the interaction term is not significantly associated with pitch.

### 4) Fixed effect model with random intercept and random slope

Consider the model form is

$$\begin{aligned}
 Y_{i,j} &= \beta_1 + \beta_2 \cdot \text{gender}_{i,j} + \beta_3 \cdot \text{attitude}_{i,j} + b_{1,i} + b_{2,i} \text{attitude}_{i,j} + \epsilon_{i,j} \\
 &= (\beta_1 + b_{1,i}) + \beta_2 \cdot \text{gender}_{i,j} + (\beta_3 + b_{2,i}) \cdot \text{attitude}_{i,j} + \epsilon_{i,j}
 \end{aligned}$$

where  $b_{1,i} \sim N(0, g_{11})$ ,  $b_{2,i} \sim N(0, g_{22})$ ,  $\text{cov}(b_{1,i}, b_{2,i}) = g_{12}$ ,  $\epsilon_{i,j} \sim N(0, \sigma^2)$ , and  $(b_{1,i}, b_{2,i})^T$  and  $\epsilon$  are assumed to be independent.

To compute the covariance structure, the  $g_{11}, g_{12}, g_{22}, \sigma$  can be obtained from below:

	Variance	StdDev	Corr
(Intercept)	5.981953e+02	24.458032213	(Intr)
attitudepol	1.079496e-05	0.003285569	0
Residual	8.477049e+02	29.115372269	

The fixed effect is:

	x
(Intercept)	256.98691

	x
genderM	-108.79762
attitudepol	-20.00238

The random effect is:

	(Intercept)	attitudepol
F1	-13.575831	-8e-07
F2	10.170522	1e-07
F3	3.405308	-3e-07
M3	27.960288	1e-06
M4	4.739325	8e-07
M7	-32.699612	-8e-07

The BLUP for the first female subject in scenario 1 with polite attitude is 223.4086921.

## Code

```
knitr::opts_chunk$set(echo = FALSE, message = FALSE, warning = FALSE, comment = "")
library(tidyverse)
library(grid)
library(gridExtra)
library(nlme)
data.poli <- read.csv("HW7-politeness_data.csv", sep = ",")
plot.gender <- data.poli %>%
  ggplot(aes(x = gender, y = frequency)) +
  geom_boxplot()

plot.attl <- data.poli %>%
  ggplot(aes(x = attitude, y = frequency)) +
  geom_boxplot()

grid.arrange(plot.gender, plot.attl, ncol = 2)
LMM1 <- lme(frequency ~ gender + attitude, data = data.poli, random = ~1 | subject, method = 'REML')
vcov(LMM1) %>% knitr::kable()
random.effects(LMM1) %>% knitr::kable()
# residual
resid.mat <- matrix(as.numeric(data.poli$frequency - fitted(LMM1)), nrow = 14, ncol = 6)
colnames(resid.mat) = rownames(random.effects(LMM1))
resid.mat %>% knitr::kable()

plot(density(as.numeric(resid.mat)), xlab = "residual", main = "residual distribution")
# interaction term
LMM.1 <- lme(frequency ~ gender + attitude, data = data.poli, random = ~1 | subject, method = 'ML')
LMM.2 <- lme(frequency ~ gender * attitude, data = data.poli, random = ~1 | subject, method = 'ML')
res.inter <- anova(LMM.1, LMM.2)
LMM2 <- lme(frequency ~ gender + attitude, random = ~ 1 + attitude | subject, data = data.poli)
VarCorr(LMM2) %>% knitr::kable()
fixed.effects(LMM2) %>% knitr::kable()
random.effects(LMM2) %>% knitr::kable()
```