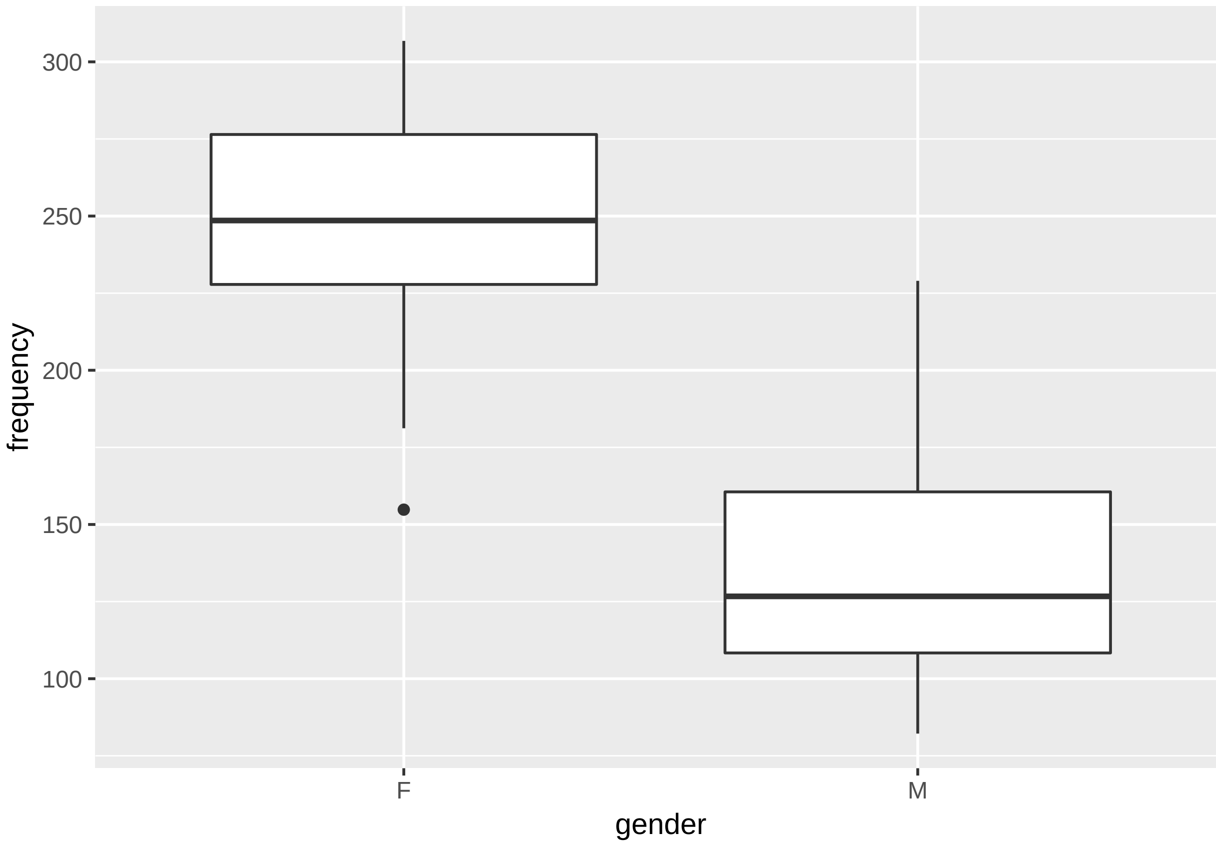
BM_HW7

Siyan Chen

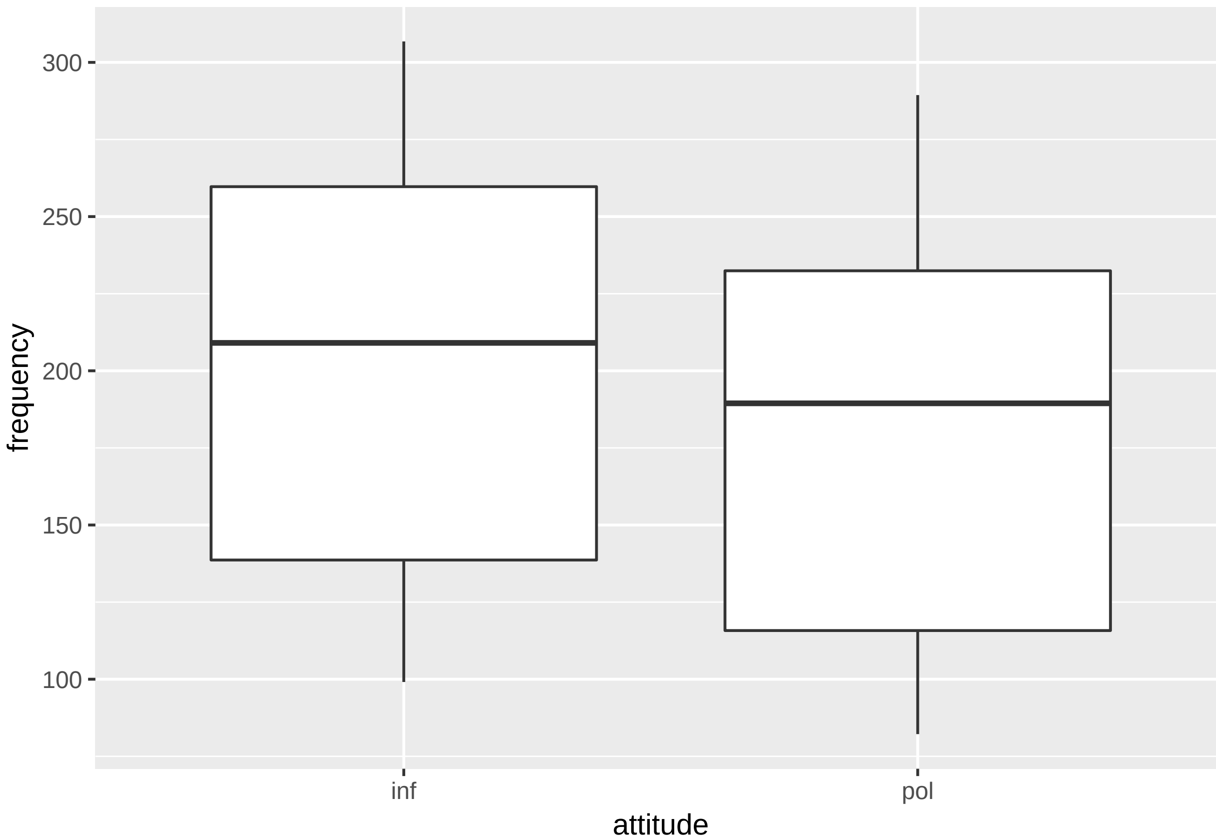
4/15/2019

1.

```r
ggplot(df) + geom_boxplot(aes(x = gender, y = frequency))
```




```r
ggplot(df) + geom_boxplot(aes(x = attitude, y = frequency))
```



Based on the plot, female have higher frequency of pitcher compared to male. Inf attitude tends to have higher frequency of pitcher compared to plr attitude.

2.

```r
# fit a random intercept model
lmm1 = lme(frequency ~ gender + attitude, random = ~1 | subject, data = df, method = 'REML')
summary(lmm1)
```

```
## Linear mixed-effects model fit by REML
##  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
```

```r
VarCorr(lmm1)
```

```
## subject = pdLogChol(1)
##             Variance StdDev
## (Intercept) 598.1953 24.45803
## Residual    847.7049 29.11537
```

covariance matrix for a subject Yi

$\sigma_b = 24.45803$
$\sigma = 29.11537$

$$cov(Y_i) = \begin{bmatrix} 1445.9 & 598.2 & \dots & 598.2 \\ 598.2 & 1445.9 & \dots & 598.2 \\ \vdots & \vdots & \ddots & \vdots \\ 598.2 & 598.2 & \dots & 1445.9 \end{bmatrix}_{14\times 14}$$

covariance matrix for the REML estimates of fixed effects

```r
vcov(lmm1)
```

```
##             (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
```

BLUPs

```r
random.effects(lmm1)
```

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

Residuals

```r
lmm1$residuals[, 2]
```

```
##           1           2           3           4           5           6 
## -10.1086926 -38.9110735  61.6913074  16.2889265 -19.5086926  43.4889265 
##           7           8           9          10          11          12 
##  27.3913074  33.3889265   8.4913074   8.9889265 -42.2086926 -12.7110735 
##          13          14          15          16          17          18 
## -26.9110735 -68.6086926 -10.6898326 -23.0922136  -3.5898326  -9.3922136 
##          19          20          21          22          23          24 
##  26.6101674   5.6077864  35.0101674  46.4077864  -7.7898326  -7.8922136 
##          25          26          27          28          29          30 
## -13.8898326  18.4077864   4.0077864 -54.8898326 -22.2262298 -29.3286108 
##          31          32          33          34          35          36 
##  96.0737702 -38.0286108 -20.7262298  60.6713892  60.4737702   9.9713892 
##          37          38          39          40          41          42 
## -31.1262298 -26.0286108 -22.9262298 -16.7286108  -6.9286108  -6.4262298 
##          43          44          45          46          47          48 
##  -9.3872916 -16.3896725 -13.2872916 -11.1896725  -9.5872916  -5.2896725 
##          49          50          51          52          53          54 
##   1.6127084   4.5103275  -1.7872916 -12.5896725  13.3127084  -7.2896725 
##          55          56          57          58          59          60 
##   8.9103275  12.1127084 -14.4550462 -35.8574271  -0.8550462  -7.4574271 
##          61          62          63          64          65          66 
##  42.2449538  34.6425729  -3.9550462  29.0425729  30.5449538  27.0425729 
##          67          68          69          70          71          72 
## -39.1550462 -41.2574271  13.8425729 -19.9550462  -2.3471929  12.6504261 
##          73          74          75          76          77          78 
## -13.7471929  23.5504261   4.0528071   9.9504261  51.3528071  14.7504261 
##          79          80          81          82          83          84 
##   4.5528071 -19.6495739  -9.4471929 -18.1495739 -15.0495739  -2.8471929
```

3

```r
lmm2 = lme(frequency ~ gender * attitude, random = ~ 1|subject, data = df, method = 'REML')
summary(lmm2)
```

```
## Linear mixed-effects model fit by REML
##  Data: df 
##        AIC     BIC    logLik
##   799.8018 814.094 -393.9009
## 
## Random effects:
##  Formula: ~1 | subject
##         (Intercept) Residual
## StdDev:    24.46382 29.04716
## 
## Fixed effects: frequency ~ gender * attitude 
##                         Value Std.Error DF   t-value p-value
## (Intercept)         260.68571 15.481307 76 16.838740  0.0000
## genderM            -116.19524 21.893875  4 -5.307203  0.0061
```

```
## attitudepol          -27.40000  8.964149 76 -3.056620  0.0031
## genderM:attitudepol   14.79524 12.677221 76  1.167073  0.2468
##  Correlation:
##                     (Intr) gendrM atttdp
## genderM             -0.707
## attitudepol         -0.290  0.205
## genderM:attitudepol  0.205 -0.290 -0.707
##
## Standardized Within-Group Residuals:
##        Min         Q1        Med         Q3        Max
## -2.2344163 -0.5454437 -0.1646159  0.4697182  3.1800944
##
## Number of Observations: 84
## Number of Groups: 6
```

```r
# compare two  models

lmm_1 = lme(frequency ~ gender + attitude, random = ~1 | subject, data = df, method = 'ML')
lmm_2 = lme(frequency ~ gender * attitude, random = ~ 1|subject, data = df, method = 'ML')

anova(lmm_2, lmm_1)
```

```
##       Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## lmm_2     1  6 826.2508 840.8357 -407.1254
## lmm_1     2  5 825.6363 837.7904 -407.8182 1 vs 2 1.385523  0.2392
```

p-value is 0.2392 which is greater than the significant level. We cannot reject the null hypothesis that the small model fit data well. Therefore, we have not evidence to show that the interaction term is signifcantly associted with pitch.

4

```r
lmm3 = lme(frequency ~ gender + attitude, random = ~ 1 + attitude|subject, data = df, method = 'REML')
summary(lmm3)
```

```
## Linear mixed-effects model fit by REML
##  Data: df
##        AIC      BIC    logLik
##   810.0805 826.8416 -398.0402
##
## Random effects:
##  Formula: ~1 + attitude | subject
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev       Corr
## (Intercept) 24.458032213 (Intr)
## attitudepol  0.003285569 0
## Residual    29.115372269
##
## Fixed effects: frequency ~ gender + attitude
##                  Value Std.Error DF   t-value p-value
## (Intercept)  256.98691 15.154987 77 16.957250  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  0.4617896  3.2997610
##
## Number of Observations: 84
## Number of Groups: 6
```

```
VarCorr(lmm3)
```

```
## subject = pdLogChol(1 + attitude)
##             Variance     StdDev       Corr
## (Intercept) 5.981953e+02 24.458032213 (Intr)
## attitudepol 1.079496e-05  0.003285569 0
## Residual    8.477049e+02 29.115372269
```

model with random intercept and slope

$Y_ij = \beta_1 + \beta_2 * I\{gender = male\} + \beta_3 * I\{attitide = pol\} + b_{1i} + b_{2i} * I\{attitide = pol\} + \epsilon_{ij}$
$b_{i\mathrm{N(0,G)}}\ {}_{\epsilon_i} N(O, \sigma^2 I)$

Attitude of inf:

$var(Y_{ij} = \sigma_{b1i}^2 + \sigma^2\ cov(Y_{ij} = cov(b_{1i} + \epsilon_{ij}, b_{1i} + \epsilon_{ik}) = \sigma_{b1i}^2$

$$cov(Y_i) = \begin{bmatrix} \sigma_{b1i}^2 + \sigma^2 & \sigma_{b1i}^2 & \dots & \sigma_{b1i}^2 \\ \sigma_{b1i}^2 & \sigma_{b1i}^2 + \sigma^2 & \dots & \sigma_{b1i}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{b1i}^2 & \sigma_{b1i}^2 & \dots & \sigma_{b1i}^2 + \sigma^2 \end{bmatrix}_{14\times 14}$$

Attitude of pol

$$cov(Y_i) = \begin{bmatrix} \sigma_{b2i}^2 + \sigma_{b1i}^2 + \sigma^2 & \sigma_{b2i}^2 + \sigma_{b1i}^2 & \dots & \sigma_{b2i}^2 + \sigma_{b1i}^2 \\ \sigma_{b2i}^2 + \sigma_{b1i}^2 & \sigma_{b2i} + \sigma_{b1i}^2 + \sigma^2 & \dots & \sigma_{b2i}^2 + \sigma_{b1i}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{b2i}^2 + \sigma_{b1i}^2 & \sigma_{b2i}^2 + \sigma_{b1i}^2 & \dots & \sigma_{b2i} + \sigma_{b1i}^2 + \sigma^2 \end{bmatrix}_{14\times 14}$$

$cov(Y_{ij} - in, Y_{ij} - pol) = \sigma_{b1i}^2 + \sigma_{b2i}^2$ approximate to σ_{b1i}

Therefore, the coveriance matrix for subject Yi is

$$cov(Y_i) = \begin{bmatrix} 1445.9 & 598.2 & \dots & 598.2 \\ 598.2 & 1445.9 & \dots & 598.2 \\ \vdots & \vdots & \ddots & \vdots \\ 598.2 & 598.2 & \dots & 1445.9 \end{bmatrix}_{14\times 14}$$

Which is approximate to compound symmetry.

```
fixed.effects(lmm3)
```

```
## (Intercept)     genderM attitudepol
##   256.98691  -108.79762   -20.00238
```

```r
random.effects(lmm3)
```

```
##    (Intercept)   attitudepol
## F1  -13.575831 -8.408891e-07
## F2   10.170522  1.499413e-07
## F3    3.405308 -2.981919e-07
## M3   27.960288  1.009764e-06
## M4    4.739325  7.794162e-07
## M7  -32.699612 -8.000404e-07
```

Fixed effect of first female is 257.0 and random effect is -0.0000008408891. BLUP of first female is -13.575831