Assignment 4 STAT 315-463: Multivariable Statistical Methods and Applications

```
library(here)
library(lattice)
library(lme4)
library(ggplot2)

library(tidyverse)
```

```
# Read in data file
tern14 <- read.table("Terns2014.csv", header = TRUE, sep = ',', na.strings = "na")
```

```
# Only using simple Linear Regression
Model <- lm(Age ~ Wing * ID, data = tern14)
summary(Model)
```

```
##
## Call:
## lm(formula = Age ~ Wing * ID, data = tern14)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6826 -1.6299 -0.5009  1.5109  6.3868
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.6194915  0.4952646   5.289 2.06e-07 ***
## Wing         0.1530470  0.0050336  30.405  < 2e-16 ***
## ID           0.0101452  0.0159284   0.637    0.525
## Wing:ID     -0.0002035  0.0001494  -1.363    0.174
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.094 on 387 degrees of freedom
## Multiple R-squared:  0.8963, Adjusted R-squared:  0.8955
## F-statistic:  1115 on 3 and 387 DF,  p-value: < 2.2e-16
```

```
AIC(Model)
```

```
## [1] 1693.598
```

Model 1: Random intercepts

```
Model.1 <- lmer(Age ~ Wing + (1|ID), data = tern14)
AIC(Model.1)
```

```
## [1] 1143.745
```

In Model 1, we are looking at the model with the random intercepts, which assume that some terns have more and some have less wing length.

The AIC value of this mixed linear model indicates that it fits better than the model that ignores individual effects (The AIC score of random intercepts model is smaller than the simple linear regression model).

```
summary(Model.1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Age ~ Wing + (1 | ID)
##    Data: tern14
##
## REML criterion at convergence: 1135.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.4005 -0.4709 -0.0712  0.4367  5.6895
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 4.2721   2.0669
##  Residual             0.5619   0.7496
## Number of obs: 391, groups:  ID, 65
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 2.976157   0.283128   10.51
## Wing        0.147138   0.001146  128.34
##
## Correlation of Fixed Effects:
##      (Intr)
## Wing -0.395
```

Random Effects: By looking at the proportion of the random intercept variance compared to the total variance, we can see that around 88.4% of variability is explained by the random intercepts. The variance of residual shows that there are still some factors from the simple body measurements of terns which affect age and cannot be explained by the model.

Fixed Effects: The intercept in the summary above shows that the average tern's wing length when they are born is 2.97 mm and the slope suggests that in average, each increase in wing length by 1 mm takes 0.14 days.

```
model1_coef <- coef(Model.1)$ID
summary(model1_coef)
```

```
##   (Intercept)           Wing
##  Min.   :-0.4788   Min.   :0.1471
##  1st Qu.: 1.3941   1st Qu.:0.1471
##  Median : 2.6807   Median :0.1471
##  Mean   : 2.9762   Mean   :0.1471
##  3rd Qu.: 4.6634   3rd Qu.:0.1471
##  Max.   : 7.3462   Max.   :0.1471
```

Model 2: Random slopes

```
Model.2 <- lmer(Age ~ Wing + (0 + Wing|ID), data = tern14)
AIC(Model.2)
```

```
## [1] 1174.457
```

In this case, we are expecting all terns to start off at around the same wing length, but the effects of wing length on age differs across each tern. The AIC value suggests this model outperforms the simple linear regression model, which does not take the different growth rate into consideration.

```
summary(Model.2)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Age ~ Wing + (0 + Wing | ID)
##    Data: tern14
##
## REML criterion at convergence: 1166.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.0715 -0.4559 -0.0779  0.4010  4.2163
##
## Random effects:
##  Groups   Name Variance  Std.Dev.
##  ID       Wing 0.0005676 0.02382
##  Residual      0.5962503 0.77217
## Number of obs: 391, groups:  ID, 65
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 2.527025   0.122949   20.55
## Wing        0.153688   0.003315   46.35
##
## Correlation of Fixed Effects:
##      (Intr)
## Wing -0.405
```

Random Effects:

Fixed Effects:

```
model2_coef <- coef(Model.2)$ID
summary(model2_coef)
```

```
##   (Intercept)         Wing
##  Min.   :2.527   Min.   :0.1208
##  1st Qu.:2.527   1st Qu.:0.1370
##  Median :2.527   Median :0.1486
##  Mean   :2.527   Mean   :0.1537
##  3rd Qu.:2.527   3rd Qu.:0.1710
##  Max.   :2.527   Max.   :0.2388
```

Model 3: Random slopes and intercepts

```
Model.3 <- lmer(Age ~ Wing + (1 + Wing|ID), data = tern14)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.439389 (tol = 0.002, component 1)
```

```
AIC(Model.3)
```

```
## [1] 944.8728
```

From the AIC result above, we can see that this model with random slopes and random intercepts has the best fit (the smallest AIC value among all four models). It does make sense because each tern has different growth rate and their wing lengths also differ when they were just born.

```
summary(Model.3)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Age ~ Wing + (1 + Wing | ID)
##    Data: tern14
##
## REML criterion at convergence: 932.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.2764 -0.4639 -0.0748  0.4586  4.9026
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  ID       (Intercept) 3.4346907 1.85329
##           Wing        0.0003303 0.01817  -0.38
##  Residual             0.2256019 0.47498
## Number of obs: 391, groups:  ID, 65
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 2.697838   0.262709   10.27
## Wing        0.150908   0.002669   56.54
```

```
##
## Correlation of Fixed Effects:
##      (Intr)
## Wing -0.510
## optimizer (nloptwrap) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 0.439389 (tol = 0.002, component 1)
```

Random Effects:

Fixed Effects: The estimate values of intercept and wing is

All three mixed linear models suggest a negative correlation between slop and intercept. This means that a tern born with a longer wing tends to have a slower growth of wing length, whereas a tern's wing length will grow faster if it is born with a shorter wing length.

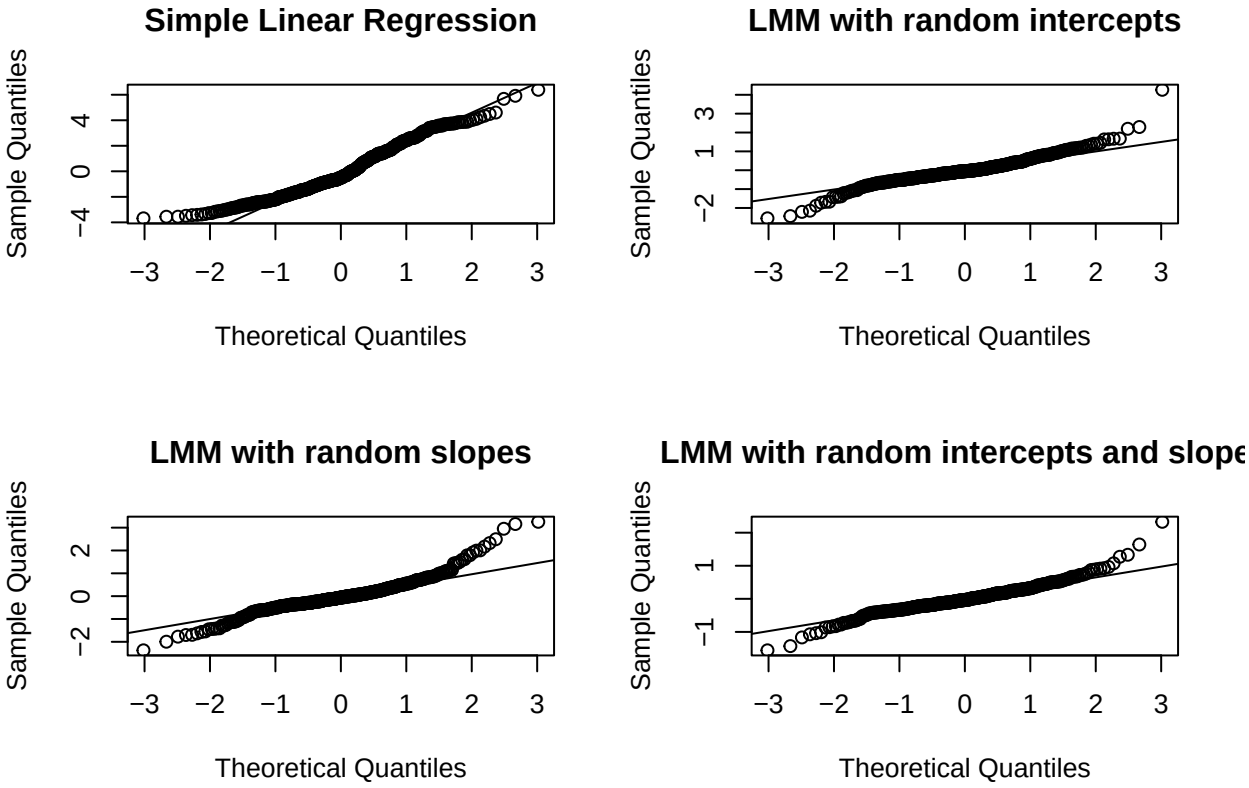
```
pardef <- par()
par(mfrow = c(2,2))

qqnorm(resid(Model), main = "Simple Linear Regression")
qqline(resid(Model))

qqnorm(resid(Model.1), main = "LMM with random intercepts")
qqline(resid(Model.1))

qqnorm(resid(Model.2), main = "LMM with random slopes")
qqline(resid(Model.2))

qqnorm(resid(Model.3), main = "LMM with random intercepts and slopes")
qqline(resid(Model.3))
```



If we look at the normal Q-Q plots from all these four models, we can notice that with the simple linear regression,