HW 8

(1) $Y_{ij} = \beta_1 + \beta_2 t_{ij} + b_{1i} + b_{2i} t_{ij} + \epsilon_{ij}$

$i = 1 \ldots N \qquad j = 1 \ldots n_i$

$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$

Var(b_{1i}) Var-Cov Mat.

$$\begin{bmatrix} b_{1i} \\ b_{2i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{bmatrix} \right)$$

Cov(b_{1i}, b_{2i}) Var(b_{2i})

A) $Var(Y_{ij}) = ?$

$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X, Y)$

$$Var[\beta_1 + \beta_2 t_{ij} + b_{1i} + b_{2i} t_{ij} + \epsilon_{ij}]$$

$$= \cancel{Var[\beta_1]}^{0} + \cancel{Var[\beta_2 t_{ij}]}^{0} + Var[b_{1i} + b_{2i} t_{ij}] + Var[\epsilon_{ij}]$$

$$= Var[b_{1i} + b_{2i}] = Var[b_{1i}] + Var[b_{2i} t_{ij}] + 2Cov[b_{1i}, b_{2i} t_{ij}] + \sigma_\epsilon^2$$

$$= g_{11} + g_{22} t_{ij}^2 + 2 t_{ij} g_{12} + \sigma_\epsilon^2$$

B) Expression for $Cov[Y_{ij}, Y_{ik}]$

$$Cov(Y_{ik}, Y_{im}) = E\left[(Y_{ik} - E[Y_{ik}])(Y_{im} - E[Y_{im}])\right]$$

$$\rightarrow E\left[(b_{1i} + b_{2i}t_{ij} + \epsilon_{ij})(b_{1i} + b_{2i}t_{ik} + \epsilon_{ik})\right]$$

$$\rightarrow E\left[b_{1i}^2 + b_{1i}b_{2i}t_{ik} + \cancel{b_{1i}\epsilon_{ik}} + b_{1i}b_{2i}t_{ij} + b_{2i}^2 t_{ij}t_{ik} + \cancel{b_{2i}t_{ij}\epsilon_{ik}} + \cancel{b_{1i}\epsilon_{ij}} + \cancel{b_{2i}t_{ik}\epsilon_{ij}} + \cancel{\epsilon_{ij}\epsilon_{ik}}\right]$$ because $E[\epsilon] = \emptyset$

$E(XY) = Cov(X, Y) = g_{12}$

$E(X^2) = Var(X) + [E(X)]^2$

$$\rightarrow E\left[b_{1i}^2 + b_{1i}b_{2i}t_{ik} + b_{1i}b_{2i}t_{ij} + b_{2i}^2 t_{ij}t_{ik}\right]$$

$$\rightarrow E[b_{1i}^2] + t_{ik}\underbrace{E[b_{1i}b_{2i}]} + t_{ij}\underbrace{E[b_{1i}b_{2i}]} + t_{ij}t_{ik}E[b_{2i}^2]$$

$Cov(b_1, b_2) = E[b_1 b_2] - \cancel{E(b_1)E(b_2)}$ 0

$Cov(b_1, b_2) = E[b_1 b_2]$

$g_{12} = E[b_{1i} b_{2i}]$

$$\rightarrow E[b_{1i}^2] + t_{ik}g_{12} + t_{ij}g_{12} + t_{ij}t_{ik}E[b_{2i}^2]$$

b/c $E[b]=0 \rightarrow E[b_{1i}^2]=Var(b_{1i})$

$\rightarrow Var(b_{1i})+Var(b_{2i})+t_{ik}g_{12}+t_{ij}g_{12}$

$\rightarrow g_{11}+t_{ij}t_{ik}g_{22}+g_{12}[t_{ik}+t_{ij}]$

c)

(c) If

$$\begin{bmatrix} b_{1i} \\ b_{2i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \right)$$

what is the value of σ_e^2 if correlation between responses at time t=0 and time t=2 is equal to 0.5, i.e. $Corr(Y_{ij}, Y_{ik}) = 0.5$ for $t_{ij} = 0$ and $t_{ik} = 2$.

$$Corr(Y_{ij}, Y_{ik}) = \frac{Cov(Y_{ij}, Y_{ik})}{\sigma_{ij} \cdot \sigma_{ik}}$$

$$\rightarrow \frac{g_{11}+t_{ij}t_{ik}g_{22}+g_{12}[t_{ik}+t_{ij}]}{\sqrt{g_{11}+g_{22}t_{ij}^2+2t_{ij}g_{12}+\sigma_e^2} \cdot \sqrt{g_{11}+g_{22}t_{ik}^2+2t_{ik}g_{12}+\sigma_e^2}}$$

$t_{ij}=0$
$t_{ik}=2$
$Corr=.5$

$g_{11}=1$
$g_{22}=1$
$g_{12}=.5$

$$0.5 = \frac{g_{11}+0\cdot 2\cdot g_{22}+g_{12}[2+0]}{\sqrt{g_{11}+g_{22}0^2+2\cdot 0\cdot g_{12}+\sigma_e^2} \cdot \sqrt{g_{11}+g_{22}2^2+2\cdot 2\cdot g_{12}+\sigma_e^2}}$$

$$0.5 = \frac{g_{11}+2g_{12}}{\sqrt{g_{11}+\sigma_e^2} \cdot \sqrt{\ldots}}$$

$$\sqrt{g_{11}+\sigma^2_\epsilon} \cdot \sqrt{g_{11}+4g_{22}+4g_{12}+\sigma^2_\epsilon}$$

$$0.5 = \frac{1+2\cdot .5}{\sqrt{1+\sigma^2_\epsilon} \cdot \sqrt{1+4+2+\sigma^2_\epsilon}} \Rightarrow \frac{2}{\sqrt{1+\sigma^2_\epsilon} \cdot \sqrt{7+\sigma^2_\epsilon}}$$

$$0.5 = \frac{2}{\sqrt{7+8\sigma^2_\epsilon+(\sigma^2_\epsilon)^2}} \qquad \sigma^2_\epsilon = \phi$$

$$0.5 = \frac{2}{\sqrt{\phi^2+8\phi+7}}$$

$$\Rightarrow \sqrt{\phi^2+8\phi+7} = 4 \Rightarrow \phi^2+8\phi+7=16$$

$$\Rightarrow \phi^2+8\phi-9=0 \qquad \frac{-8\pm\sqrt{64-4\cdot 1\cdot(-9)}}{2}$$

$$\frac{-8\pm\sqrt{64+36}}{2} = \frac{-8\pm 10}{2}$$

$$= 1, -9$$

$\sigma^2_\epsilon = 1$

HW8

```
library(tidyverse)
```

```
-- Attaching packages --------------------------------------- tidyverse 1.3.1 --

v ggplot2 3.3.5     v purrr   0.3.4
v tibble  3.1.4     v dplyr   1.0.7
v tidyr   1.1.3     v stringr 1.4.0
v readr   2.0.1     v forcats 0.5.1

-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
```

```
library(lme4)
```

```
Loading required package: Matrix


Attaching package: 'Matrix'

The following objects are masked from 'package:tidyr':

    expand, pack, unpack
```

```
library(reshape2)
```

```

Attaching package: 'reshape2'

The following object is masked from 'package:tidyr':

    smiths
```

```
dat <- read.csv("table11_1_modified.csv")

dt.l <- melt(dat, id.vars=c("id", "group"), measure.vars=paste("week",1:8,sep=""),
             variable.name="week", value.name="fas")

dt.l$time <- recode(dt.l$week, "week1"=1, "week2"=2, "week3"=3, "week4"=4, "week5"=5, "week6"=6,

dt.l <- dt.l[order(dt.l$id),]
```

Question 2

Part A

Consider “Model 3” in the lecture note. Your collaborator Dr. Smith Lee, a clinician, email you to ask you explain the random intercept and random slope in the model. Write a response email to Dr. Lee. Use plain language as much as possible without numerical results in your reply.

```
fit1 <- lmer(fas ~ time + (1 + time| id), data=dt.l)
summary(fit1)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: fas ~ time + (1 + time | id)
   Data: dt.l

REML criterion at convergence: 1335

Scaled residuals: 
    Min      1Q  Median      3Q     Max 
-2.8482 -0.5401 -0.0040  0.4944  3.2309 

Random effects:
 Groups   Name        Variance Std.Dev. Corr 
 id       (Intercept) 190.225  13.792        
          time          8.273   2.876   -0.17
 Residual              27.638   5.257        
Number of obs: 192, groups:  id, 24

Fixed effects:
            Estimate Std. Error t value
(Intercept)   27.637      2.937   9.410
time           5.030      0.610   8.245

Correlation of Fixed Effects:
     (Intr)
time -0.225
```

In this model we are predicting the functional ability scores (FAS) of an individual. We allow for a random intercept, meaning that we expect individuals to have a random effect on the intercepts of the model, or the predicted FAS value for an individual at time = 0.

We also are allowing for a random slope, meaning that we are allowing for individuals to have different effects on the slope of the model, or, in other words, an individuals effect on FAS over time changes different between individuals.

Part B

Consider “Model 3” in the lecture note. You now want to study if the temporal effect differs by group. Consider GROUP to be a continuous variable. Assume random intercept and random slope for the time effect only in the model. Write down the linear mixed effects model. Fit the model to the provided dataset. Interpret the fixed effect for the interaction. In addition, what are the estimated variances of random effects? How do you interpret them?

```
dt.l$group <- unclass(as.factor(dt.l$group))
fit3 <- lmer(fas ~ time + group + time*group + (1+ time|id), data = dt.l)
summary(fit3)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: fas ~ time + group + time * group + (1 + time | id)
   Data: dt.l

REML criterion at convergence: 1322.4

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.8242 -0.5415  0.0040  0.5127  3.2420

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 id       (Intercept) 181.966  13.489
          time          6.475   2.545   -0.03
 Residual              27.638   5.257
Number of obs: 192, groups:  id, 24

Fixed effects:
            Estimate Std. Error t value
(Intercept)  17.7798     7.6137   2.335
time          8.5119     1.4423   5.901
group         4.9286     3.5244   1.398
time:group   -1.7411     0.6677  -2.608

Correlation of Fixed Effects:
           (Intr) time   group
time       -0.102
group      -0.926  0.095
time:group  0.095 -0.926 -0.102
```

$$Y_{ij} = \beta_0 + a_i + (\beta_1 + b_i)t_{ij} + \beta_2 * \text{Group} + \beta_3 t_{ij} * \text{Group} + e_{ij}$$

$$\text{For a given time, } t: \text{Group} = 1: \beta_0 + a_i + (\beta_1 + b_i + \beta_3)t + \beta_2$$

$$\text{Group} = 0: \beta_0 + a_i(\beta_1 + b_i)t \text{The difference between these two is: } \beta_3 t + \beta_2$$

$$\text{For a given time, } t+1: \text{Group} = 1: \beta_0 + a_i + (\beta_1 + b_i + \beta_3)(t+1) + \beta_2$$

$$\text{Group} = 0: \beta_0 + a_i(\beta_1 + b_i)(t+1) \text{The difference between these two is: } \beta_3(t+1) + \beta_2$$

The difference of the differences as shown above is $\neq 0$, so a 1 unit increase in time changes the effect of group on FAS, by a value of β_3.

The estimated variance of random effects for id is 181.966, and the variance for the random effect of time is 6.475. The variance is pretty large for for id, meaning our intercept varies a large amount between individuals. The variance for time or slope is pretty small, meaning the slope over time does not change particularly a lot between individuals.

Part C

Now, add a random slope for the group effect (so there are two random slopes) to the model in (b). Write down the updated model. Fit the model to the provided dataset. How do you explain the random slope for the group effect to Dr. Lee? Compare fixed effect coefficients in (b) and (c). Are they the same? Now, compared with results in (b), is the inference for the fixed effect of the interaction changed? Explain it. In addition, what are the estimated variances of random effects? Which main covariate effect does vary more across individuals, time or group?

$$Y_{ij} = \beta_0 + a_i + (\beta_1 + b_i)t_{ij} + (\beta_2 + c_i) * \text{Group} + \beta_3 t_{ij} * \text{Group} + e_{ij}$$

```
fit3c <- lmer(fas ~ time + group + time*group + (1+ time + group|id), data = dt.l)
summary(fit3c)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: fas ~ time + group + time * group + (1 + time + group | id)
   Data: dt.l

REML criterion at convergence: 1320.3

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.8209 -0.5266  0.0199  0.5041  3.2518

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 id       (Intercept) 309.242  17.585
          time          6.487   2.547   -0.70
          group        32.185   5.673   -0.68  1.00
 Residual              27.622   5.256
Number of obs: 192, groups:  id, 24

Fixed effects:
            Estimate Std. Error t value
(Intercept)  19.3877     7.7937   2.488
time          8.1286     1.4112   5.760
group         4.2669     3.5704   1.195
time:group   -1.5494     0.6507  -2.381

Correlation of Fixed Effects:
           (Intr) time   group
time       -0.363
```

```
group       -0.928  0.260
time:group   0.252 -0.922 -0.153
```

To explain the random slope for group effect to Dr. Lee, I would say that each group has its own effect on slope across time. Each individual has its own slope, but for a 1 unit increase in group and time, FAS increases by 4.27.

The fixed effects coefficients are somewhat similar, the first model intercept value is 17.78, and the second model is 19.39. The other estimates are pretty similar as well. They are not the same.

```r
library(car)
```

```
Loading required package: carData


Attaching package: 'car'

The following object is masked from 'package:dplyr':

    recode

The following object is masked from 'package:purrr':

    some
```

```r
Anova(fit3)
```

```
Analysis of Deviance Table (Type II Wald chisquare tests)

Response: fas
             Chisq Df Pr(>Chisq)    
time       85.1243  1  < 2.2e-16 ***
group       1.2932  1   0.255453    
time:group  6.7999  1   0.009116 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
Anova(fit3c)
```

```
Analysis of Deviance Table (Type II Wald chisquare tests)

Response: fas
             Chisq Df Pr(>Chisq)    
time       84.9760  1    < 2e-16 ***
group       0.7085  1    0.39994    
time:group  5.6695  1    0.01726 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value for both of the interaction terms are significant, thus the inference for the interaction term does not change.

The estimated variances of random effects for the intercept in the first model is 181.97, and 309.24, which are pretty different. For time, it is 6.475 and 6.487. The residual variance is nearly equal, at 27.638 and 27.622.

The main covariate effect of group varies more across individuals. This is because the variance is larger for group than it is for time