Exercises Day 2

Your name

2026-01-28

Exercise 1

You performed a cognitive experiment with a pre- and post-test design. Participants completed one session before language training and one session after language training, with two types of training: Forced and Voluntary. The identical set switching task was used for pre- and post-test. Within each switching task reaction times from pure blocks are compared to switch/repeat blocks, with the expectation that participants react faster in pure blocks than non-pure blocks.

a) Prepare the data set for use in a LME. Create a new 2-level variable called BlockType that compares pure trials against non-pure trial.

```
load("CogTrainingData.RData")

Model_Data <-
  Cog_training_Data %>%
  mutate(BlockType = ifelse(TrialType == "pure", 0, 1))
```

b) Create a linear mixed effects model that tests whether there was a significant effect in the cognitive test performance for BlockType. Control for individual effects of Subject and Items.

```
Model_BlockType = lmer(RT ~ BlockType + (1 | Subject) + (1 | Item), data=Model_Data)

summary(Model_BlockType)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: RT ~ BlockType + (1 | Subject) + (1 | Item)
##    Data: Model_Data
##
## REML criterion at convergence: 423372
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5799 -0.6480 -0.1617  0.4501  5.5685
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept)  7040.8   83.91
##  Item     (Intercept)   120.7   10.98
##  Residual             29777.2  172.56
```

```
## Number of obs: 32191, groups:  Subject, 85; Item, 16
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  508.961      9.584    53.1
## BlockType    215.385      2.003   107.6
##
## Correlation of Fixed Effects:
##           (Intr)
## BlockType -0.076
```

c) Given that there were two sessions of the cognitive test, before and after training, and there were two types of training, create a new model that accounts for all three variables. Which fixed effects are significant?

```
Model_Large = lmer(RT ~ BlockType*Session*Group + (1 | Subject) + (1 | Item), data=Model_Data)

summary(Model_Large)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: RT ~ BlockType * Session * Group + (1 | Subject) + (1 | Item)
##    Data: Model_Data
##
## REML criterion at convergence: 422136.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.4268 -0.6389 -0.1570  0.4430  5.5338
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept)  7137.0   84.48
##  Item     (Intercept)   129.4   11.38
##  Residual             28683.1  169.36
## Number of obs: 32191, groups:  Subject, 85; Item, 16
##
## Fixed effects:
##                                  Estimate Std. Error t value
## (Intercept)                       574.888     14.056  40.899
## BlockType                         293.182      8.946  32.774
## Session                           -42.768      3.291 -12.997
## GroupVoluntary                      5.480     19.830   0.276
## BlockType:Session                 -51.987      5.516  -9.424
## BlockType:GroupVoluntary           -6.093     12.782  -0.477
## Session:GroupVoluntary             -5.600      4.744  -1.180
## BlockType:Session:GroupVoluntary   11.415      7.890   1.447
##
## Correlation of Fixed Effects:
##             (Intr) BlckTy Sessin GrpVln BlcT:S BlT:GV Sss:GV
## BlockType   -0.217
## Session     -0.353  0.554
## GroupVlntry -0.680  0.154  0.250
## BlckTyp:Sss  0.210 -0.952 -0.596 -0.149
```

```
## BlckTyp:GrV  0.152 -0.700 -0.387 -0.224  0.666
## Sssn:GrpVln  0.245 -0.384 -0.694 -0.361  0.413  0.560
## BlckTy:S:GV -0.147  0.665  0.417  0.217 -0.699 -0.951 -0.601
```

The significant effects are BlockType, Session and BlockType:Session as all three have a t-value greater than abs(2).

d) You also have information regarding the age for each participant. Create a model that includes this covariate as a fixed effect. Did the outcome of the model change compared to 1a? What might be possible reasons for this?

```
Model_Age = lmer(RT ~ BlockType*Session*Group + Age + (1 | Subject) + (1 | Item), data=Model_Data)

summary(Model_Age)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: RT ~ BlockType * Session * Group + Age + (1 | Subject) + (1 |
##     Item)
##    Data: Model_Data
##
## REML criterion at convergence: 422130.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.4269 -0.6389 -0.1568  0.4430  5.5326
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept)  7111.3   84.33
##  Item     (Intercept)   129.5   11.38
##  Residual             28683.1  169.36
## Number of obs: 32191, groups:  Subject, 85; Item, 16
##
## Fixed effects:
##                                 Estimate Std. Error t value
## (Intercept)                      487.473     78.184   6.235
## BlockType                        293.179      8.946  32.773
## Session                          -42.766      3.291 -12.997
## GroupVoluntary                     9.725     20.148   0.483
## Age                                3.695      3.251   1.137
## BlockType:Session                -51.986      5.516  -9.424
## BlockType:GroupVoluntary          -6.096     12.782  -0.477
## Session:GroupVoluntary            -5.601      4.744  -1.180
## BlockType:Session:GroupVoluntary  11.415      7.890   1.447
##
## Correlation of Fixed Effects:
##             (Intr) BlckTy Sessin GrpVln Age    BlcT:S BlT:GV Sss:GV
## BlockType   -0.039
## Session     -0.064  0.554
## GroupVlntry -0.302  0.151  0.246
## Age         -0.984  0.000  0.000  0.185
## BlckTyp:Sss  0.038 -0.952 -0.596 -0.147  0.000
## BlckTyp:GrV  0.027 -0.700 -0.387 -0.221  0.000  0.666
```

```
## Sssn:GrpVln  0.044 -0.384 -0.694 -0.355  0.000  0.413  0.560
## BlckTy:S:GV -0.026  0.665  0.417  0.213  0.000 -0.699 -0.951 -0.601
```

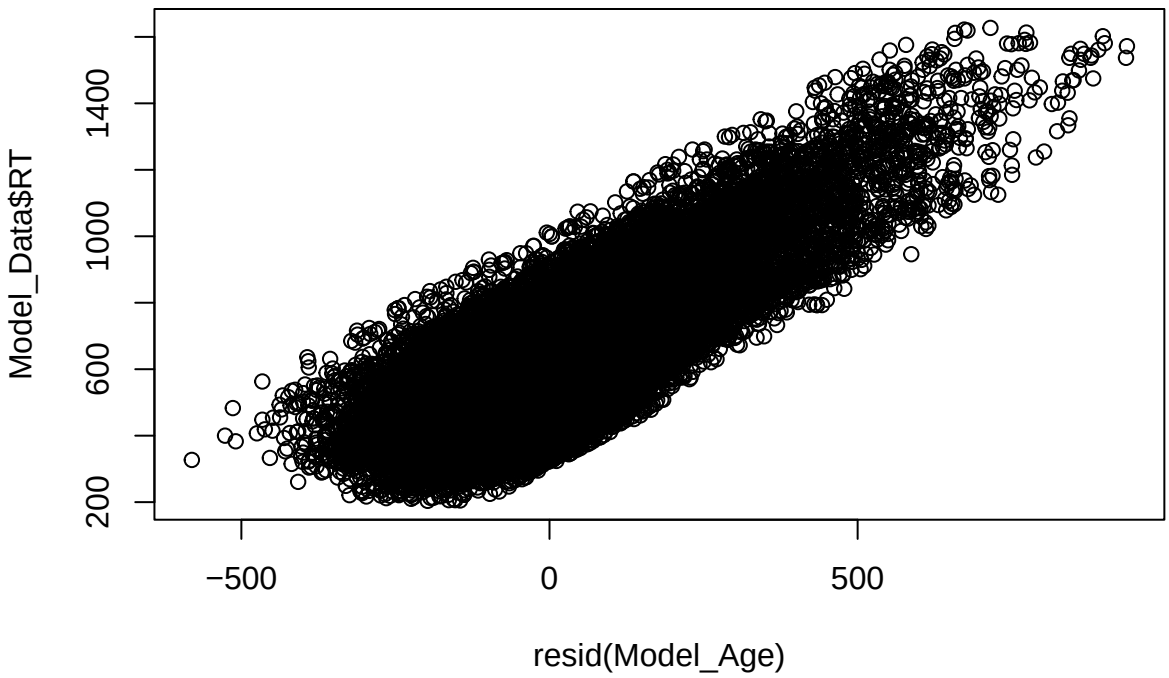
The fixed effects remained significant, however the t-values got slightly smaller, as a result of including an additional fixed effect. By accounting for more effects in the model, data variability is "distributed" between more variables, reducing the effects in single variables in comparison.

Exercise 2

Verify the assumptions of LMEs for the last model involving age as a covariate.

a) Is the dependent variable linear related to the residuals?

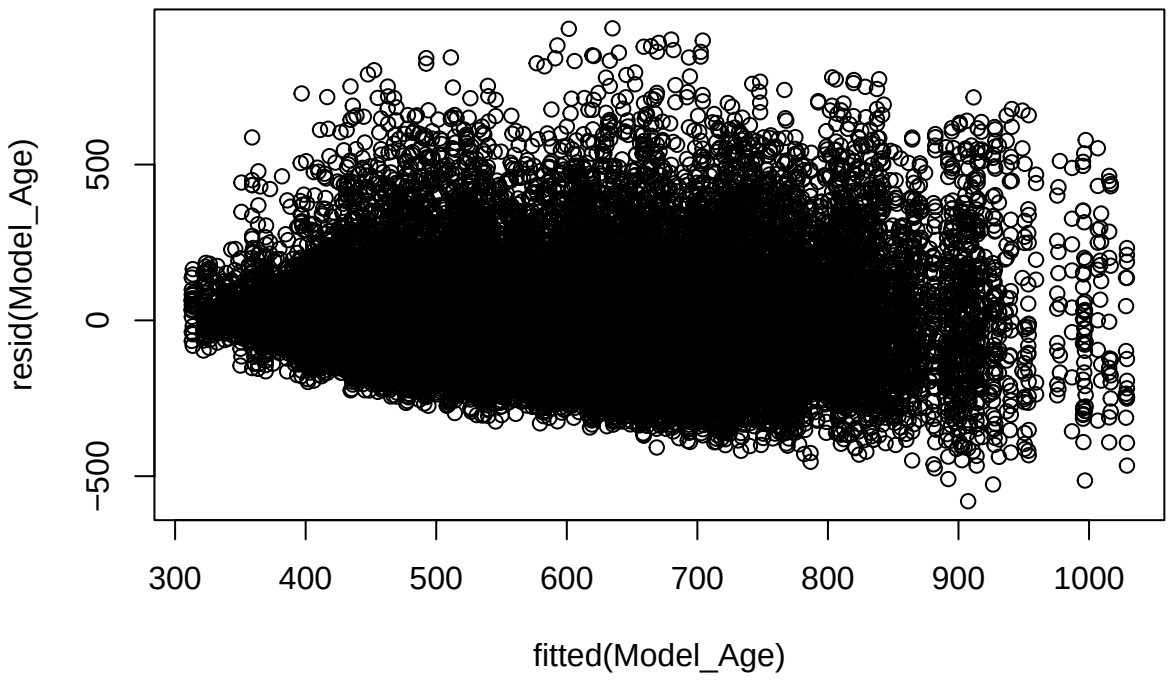
```
plot(resid(Model_Age), Model_Data$RT)
```



Yes, there is a linear relationship between residuals and RT

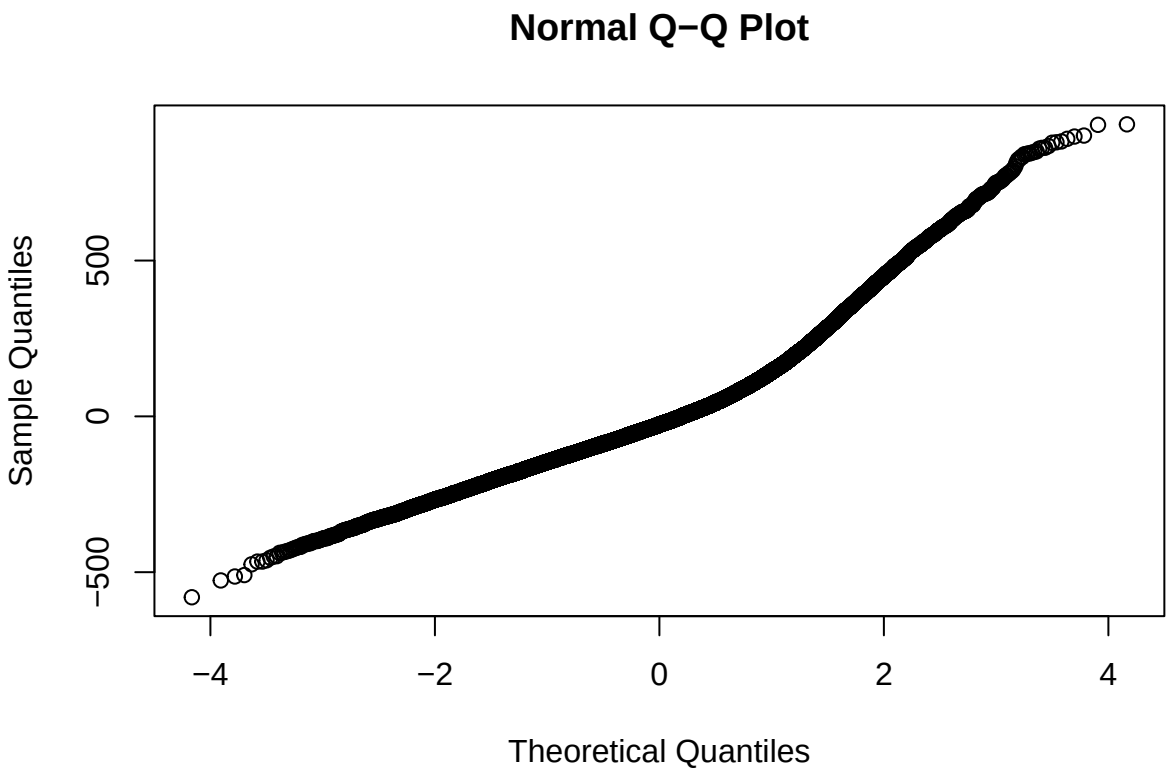
b) Test for constant variance in the residuals.

```
plot(fitted(Model_Age), resid(Model_Age))
```



c) Test for normal distributed residuals.

```
res_model <- residuals(Model_Age)
qqnorm(res_model)
```



d) Test for multicollinearity. Which variables are correlated with each other and what might be the reason?

```
vif(Model_Age)
```

```
##               BlockType                 Session                   Group
##                20.717069                3.022089                1.198370
##                     Age       BlockType:Session         BlockType:Group
##                1.041141                22.240478                20.725371
##           Session:Group BlockType:Session:Group
##                3.204804                22.235373
```

BlockType and its interactions show high correlation as BlockType is used within all of those interactions.