ProblemSet4_2

Name: Zhen Qin, Uniqname: qinzhen, UMID: 48800866, Dept: Statistics

Datasets are from https://github.com/jensqin/Stat506 (https://github.com/jensqin/Stat506) .

c.

```
setwd("E:/UM academy/stat 506/ProblemSet4")
library(sas7bdat)
library(lme4)

# load the data
rdclong = read.sas7bdat('rdclong.sas7bdat')

# Filter your data to contain only the 1000 Hz test for the right ear
test1kr = rdclong[rdclong$ear==1&rdclong$freq==1&rdclong$threshold<200,]
test = rdclong[rdclong$threshold<200,]
```

```
# fit the model with interaction
fit_interaction=lm(threshold ~group*RIAGENDR, data=test1kr)
summary(fit_interaction)
```

```
##
## Call:
## lm(formula = threshold ~ group * RIAGENDR, data = test1kr)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.896  -5.480  -0.480   4.668  89.520
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      5.6280     0.8443   6.666 3.17e-11 ***
## group           21.7885     1.6164  13.480  < 2e-16 ***
## RIAGENDR        -0.1480     0.5333  -0.278    0.781
## group:RIAGENDR   1.3878     1.0379   1.337    0.181
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.95 on 2729 degrees of freedom
## Multiple R-squared:  0.4369, Adjusted R-squared:  0.4363
## F-statistic: 705.9 on 3 and 2729 DF,  p-value: < 2.2e-16
```

The interaction is not significant because the p-value is 0.181>0.05.

```
# fit the model to test age
fit_age=lm(threshold~RIDAGEYR+group+RIAGENDR+RIDAGEYR:group+RIDAGEYR:RIAGENDR,data=test1kr)
summary(fit_age)
```

```
##
## Call:
## lm(formula = threshold ~ RIDAGEYR + group + RIAGENDR + RIDAGEYR:group +
##     RIDAGEYR:RIAGENDR, data = test1kr)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -34.263  -5.597  -0.538   4.566  89.004
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         3.37173    2.02428   1.666   0.0959 .
## RIDAGEYR            0.14196    0.11503   1.234   0.2173
## group             -67.72894    6.77964  -9.990  < 2e-16 ***
## RIAGENDR           -0.27486    0.67439  -0.408   0.6836
## RIDAGEYR:group      1.05344    0.14060   7.492 9.09e-14 ***
## RIDAGEYR:RIAGENDR   0.01063    0.01600   0.665   0.5063
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.52 on 2727 degrees of freedom
## Multiple R-squared:  0.477,  Adjusted R-squared:  0.476
## F-statistic: 497.4 on 5 and 2727 DF,  p-value: < 2.2e-16
```

After controlling for age group and gender, age is not important as a continuous variable because the p-value is 0.217>0.05.

```
# fit the model to test group
fit_group=lm(threshold~RIDAGEYR+group+RIDAGEYR:group,data=test1kr)
summary(fit_group)
```

```
##
## Call:
## lm(formula = threshold ~ RIDAGEYR + group + RIDAGEYR:group, data = test1kr)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -34.554  -5.643  -0.486   4.514  89.041
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.9641     1.7633   1.681   0.0929 .
## RIDAGEYR         0.1576     0.1126   1.400   0.1617
## group          -67.9324     6.7705 -10.034  < 2e-16 ***
## RIDAGEYR:group   1.0561     0.1405   7.516 7.61e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.52 on 2729 degrees of freedom
## Multiple R-squared:  0.4769, Adjusted R-squared:  0.4763
## F-statistic: 829.3 on 3 and 2729 DF,  p-value: < 2.2e-16
```

The effect of age, as a continuous variable, is significantly different among the older and/or younger age groups because the p-value is less than 0.05.

d.

```
# fit the mixed model with interaction
mx_interaction=lmer(threshold ~group*RIAGENDR+(ear|freq),data = test)
summary(mx_interaction)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: threshold ~ group * RIAGENDR + (ear | freq)
##    Data: test
##
## REML criterion at convergence: 314088.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0395 -0.5673 -0.0310  0.4464  7.3432
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  freq     (Intercept)  34.0129  5.8321
##           ear           0.6943  0.8332  -0.94
##  Residual             219.1067 14.8023
## Number of obs: 38170, groups:  freq, 7
##
## Fixed effects:
##                Estimate Std. Error t value
## (Intercept)      6.7540     1.1611    5.82
## group           50.7721     0.5378   94.41
## RIAGENDR        -0.8448     0.1765   -4.79
## group:RIAGENDR  -7.2817     0.3449  -21.11
##
## Correlation of Fixed Effects:
##             (Intr) group  RIAGEN
## group       -0.125
## RIAGENDR    -0.227  0.493
## gr:RIAGENDR  0.116 -0.947 -0.512
```

In fixed effects, the absolute value of t value of the interaction is greater than 2, so it is significant.

```
# fit the mixed model to test age
mx_age=lmer(threshold~RIDAGEYR+group+RIAGENDR+RIDAGEYR:group+RIDAGEYR:RIAGENDR+(ear|freq),dat
a=test)
summary(mx_age)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## threshold ~ RIDAGEYR + group + RIAGENDR + RIDAGEYR:group + RIDAGEYR:RIAGENDR +
##     (ear | freq)
##    Data: test
##
## REML criterion at convergence: 312346.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.2724 -0.5542 -0.0293  0.4636  7.4690
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  freq     (Intercept)  34.0358  5.8340
##           ear           0.6933  0.8326  -0.93
##  Residual             209.2355 14.4650
## Number of obs: 38170, groups:  freq, 7
##
## Fixed effects:
##                     Estimate Std. Error t value
## (Intercept)         2.010511   1.326354   1.516
## RIDAGEYR            0.309199   0.038591   8.012
## group             -51.441914   2.286115 -22.502
## RIAGENDR            1.059910   0.226346   4.683
## RIDAGEYR:group      1.078443   0.047265  22.817
## RIDAGEYR:RIAGENDR  -0.124462   0.005391 -23.087
##
## Correlation of Fixed Effects:
##             (Intr) RIDAGEYR group  RIAGEN RIDAGEYR:g
## RIDAGEYR    -0.469
## group       -0.112  0.243
## RIAGENDR    -0.250  0.147   -0.015
## RIDAGEYR:gr  0.349 -0.777   -0.784  0.016
## RIDAGEYR:RI  0.190 -0.203    0.038 -0.756 -0.026
```

In fixed effects, the absolute value of t value of RIAGENDR is greater than 2, so it is significant.

```
# fit the mixed model to test group
mx_group=lmer(threshold~RIDAGEYR+group+RIDAGEYR:group+(ear|freq),data=test)
summary(mx_group)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: threshold ~ RIDAGEYR + group + RIDAGEYR:group + (ear | freq)
##    Data: test
##
## REML criterion at convergence: 313238
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0182 -0.5659 -0.0426  0.4497  7.4264
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  freq     (Intercept)  33.9349  5.8254
##           ear           0.6926  0.8322  -0.93
##  Residual             214.2326 14.6367
## Number of obs: 38170, groups:  freq, 7
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)      3.64073    1.28460   2.834
## RIDAGEYR         0.12056    0.03823   3.153
## group          -48.51150    2.31108 -20.991
## RIDAGEYR:group   1.04568    0.04781  21.872
##
## Correlation of Fixed Effects:
##             (Intr) RIDAGEYR group
## RIDAGEYR    -0.461
## group       -0.121  0.256
## RIDAGEYR:gr  0.369 -0.800   -0.784
```

In fixed effects, the t-value is greater than 2 so the effect of age, as a continuous variable, is significantly different among the older and/or younger age groups.