

Stat 506 - Problem 4

Zhen Qin

Datasets and scripts are from <https://github.com/jensqin/Stat506> and [here](#).

1.

a.

First, transform data to proper type.

```
libname aux xport './NHANES/AUX_D.XPT';

libname demo xport './NHANES/DEMO_D.XPT';

libname lib './NHANES';

/* change the data type */
proc copy inlib=aux outlib=lib;

proc copy inlib=demo outlib=lib;

run;
```

Then merge the data.

```
libname lib './NHANES';

data auxd;
set lib.aux_d;

data demod;
set lib.demo_d;

/* merge the data */
data mer;
merge auxd(in=x) demod;
by seqn;
if x = 1;

data lib.mer;
set mer;

run;
```

b.

Select useful variables.

```

libname lib './NHANES';

data rdcdata;
set lib.mer;
keep seqn riagendr ridageyr auxu;;
drop auxu1k2r auxu1k2l;

data lib.rdcdata;
set rdcdata;

run;

```

Then reshape the data.

```

libname lib './NHANES';

data rdcdata;
set lib.rdcdata;

/* reshape data to long format */
proc transpose data=rdcdata out=rdclong prefix=auxu;
by seqn riagendr ridageyr;

/* remove missing value */
data rdclong;
set rdclong;
if nmiss(auxu1) then delete;

/* encode ear, freq and group */
/* ear = 1 when test the right ear */
/* group = 1 when age >= 25 */
data lib.rdclong;
set rdclong;
ear = 0;
freq = 1;
if prxmatch("/R/",_NAME_) then ear = 1;
if prxmatch("/2K/",_NAME_) then freq = 2;
if prxmatch("/3K/",_NAME_) then freq = 3;
if prxmatch("/4K/",_NAME_) then freq = 4;
if prxmatch("/6K/",_NAME_) then freq = 6;
if prxmatch("/8K/",_NAME_) then freq = 8;
if prxmatch("/500/",_NAME_) then freq = 0.5;
group = 0;
if ridageyr > 25 then group = 1;
rename auxu1=threshold;
keep seqn ridageyr riagendr auxu1 ear freq group;

run;

```

C.

Filter your data to contain only the 1000 Hz test for the right ear.

```

libname lib './';

/* Filter your data to contain only the 1000 Hz test for the right ear */
data test1kr;
set lib.rdc1long;
if freq > 1 then delete;
if freq < 1 then delete;
if ear < 1 then delete;
if threshold > 200 then delete;

data model1;
set test1kr;
gender = riagendr - 1;
inter_gg = group * gender;

/* fit the model with interaction */
proc reg data=model1;
model threshold = group gender inter_gg;

data model2;
set test1kr;
gender = riagendr - 1;
inter_rg=ridageyr*group;
inter_gr=ridageyr*gender;

/* fit the model to test age */
proc reg data=model2;
model threshold =ridageyr gender group inter_rg inter_gr;

data model3;
set test1kr;
inter_rg = ridageyr*group;

/* fit the model to test group */
proc reg data=model3;
model threshold = ridageyr group inter_rg;

run;

```

For the first model, the interaction is not significant because the p-value >0.05.

For the second model, after controlling for age group and gender, age is not important as a continuous variable because the p-value>0.05.

For the third model, 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.

The results are stored in reg1kr.lst.

d.

```

libname lib './';

data test;
set lib.rdcldong;
if threshold > 200 then delete;

data mmodel1;
set test1kr;
gender = riagendr - 1;
inter_gg = group * gender;

/* fit the mixed model with interaction */
proc reg data=mmodel1;
model threshold = group gender inter_gg;
random ear freq;

data mmodel2;
set test1kr;
gender = riagendr - 1;
inter_rg=ridageyr*group;
inter_gr=ridageyr*gender;

/* fit the mixed model to test age */
proc reg data=mmodel2;
model threshold =ridageyr gender group inter_rg inter_gr;
random ear freq;

data mmodel3;
set test1kr;
inter_rg = ridageyr*group;

/* fit the mixed model to test group */
proc reg data=mmodel3;
model threshold = ridageyr group inter_rg;
random ear freq;

run;

```

For model 1, in fixed effects, the interaction is significant.

For model 2, in fixed effects, RIAGENDR is significant.

For model 3, in fixed effects, the effect of age, as a continuous variable, is significantly different among the older and/or younger age groups.

The results are stored in regmx.lst.