The Noise Data

Participants listened to brief political discussions under 5 levels of background noise. "Discrimination score" is a measure of how well they could tell what was being said. There are 5 lines of data per case. The variables are

- Subject 1dentification code
- Interest in topic (politics)
- Sex (0=Male, 1=Female)
- Age category (3 levels)
- Noise level
- Time (Order of noise level presentation)
- Discrimination score

```
> library(lme4) # Package must be downloaded and installed
Loading required package: Matrix
> loud =
read.table("http://www.utstat.toronto.edu/~brunner/data/legal/noise.data.txt")
> colnames(loud) = c("ident", "interest", "sex", "age", "noise", "time", "discrim")
> head(loud,n=10) # Notice the univariate data format
   ident interest sex age noise time discrim
1      1      2.5   1   2     1    4    50.7
2      1      2.5   1   2     2    1    27.4
3      1      2.5   1   2     3    3    39.1
4      1      2.5   1   2     4    2    37.5
5      1      2.5   1   2     5    5    35.4
6      2      1.9   1   2     1    3    40.3
7      2      1.9   1   2     2    1    30.1
8      2      1.9   1   2     3    5    38.9
9      2      1.9   1   2     4    2    31.9
10     2      1.9   1   2     5    4    31.6

> attach(loud)
> agefactor = factor(age); noisefactor=factor(noise); timefactr = factor(time)
>
> ########## Sex by Noise Level ##########
> table(sex,noise)/5 # There are 5 lines for each person
   noise
sex 1 2 3 4 5
  0 6 6 6 6 6
  1 6 6 6 6 6
> # Look at the means
> meanz = aggregate(discrim~sex+noise, FUN=mean)
> meantable = meanz[,3]; dim(meantable) = c(2,5)
> dimnames(meantable) = list(c("Male","Female"),1:5)
> # Add marginal means and round
> meantable = round(addmargins(meantable,FUN=mean),2); meantable

Margins computed over dimensions
in the following order:
1:
2:
           1     2     3     4     5  mean
Male   40.07 36.20 35.43 32.70 30.84 35.05
Female 39.57 37.46 35.18 36.07 32.05 36.06
mean   39.82 36.83 35.30 34.38 31.45 35.56
```

```
>
> # Tests
> noise1 = lmer(discrim ~ sex*noisefactor + (1 | ident))
> anova(noise1) # Matches the classical mixed model F statistics

Analysis of Variance Table
               Df  Sum Sq Mean Sq F value
sex             1   17.05   17.05  0.4192
noisefactor     4 2289.31  572.33 14.0718
sex:noisefactor 4  142.42   35.61  0.8754

> # But there are no p-values. As of December 2016, avoid the lmerTest package.
> # Large-sample likelihood ratio tests: Full vs reduced
> noint = update(noise1, . ~ . - sex:noisefactor) # Reduced model
> anova(noint,noise1) # Compare classical p = 0.4793

refitting model(s) with ML (instead of REML)

Data: NULL

Models:
noint: discrim ~ sex + noisefactor + (1 | ident)
noise1: discrim ~ sex * noisefactor + (1 | ident)

       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
noint   8 2063.2 2092.9 -1023.6   2047.2
noise1 12 2067.7 2112.1 -1021.8   2043.7 3.5954      4     0.4635

>
> # For a reduced model with interaction but missing a main effect, you
> # must make your own dummy variables. I will use effect coding.
> Sex = 2*sex-1
> n = length(discrim); Noise1 = Noise2 = Noise3 = Noise4 = numeric(n)
> Noise1[noise==1] = 1; Noise1[noise==5] = -1
> Noise2[noise==2] = 1; Noise2[noise==5] = -1
> Noise3[noise==3] = 1; Noise3[noise==5] = -1
> Noise4[noise==4] = 1; Noise4[noise==5] = -1
> SN1 = Sex*Noise1; SN2 = Sex*Noise2; SN3 = Sex*Noise3; SN4 = Sex*Noise4
>
> full = lmer(discrim ~ Sex+Noise1+Noise2+Noise3+Noise4+SN1+SN2+SN3+SN4
+                       + (1 | ident))
> red = lmer(discrim ~ Sex+Noise1+Noise2+Noise3+Noise4 + (1 | ident))
> anova(red,full) # Checks

refitting model(s) with ML (instead of REML)

Data: NULL

Models:
red: discrim ~ Sex + Noise1 + Noise2 + Noise3 + Noise4 + (1 | ident)
full: discrim ~ Sex + Noise1 + Noise2 + Noise3 + Noise4 + SN1 + SN2 +
full:     SN3 + SN4 + (1 | ident)

     Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
red   8 2063.2 2092.9 -1023.6   2047.2
full 12 2067.7 2112.1 -1021.8   2043.7 3.5954      4     0.4635
```

```
>
> # Main effect of sex
> nosex = update(full, . ~ . - Sex) # Reduced model
> anova(nosex,full) # Compare classical p = 0.5199

refitting model(s) with ML (instead of REML)

Data: NULL

Models:
nosex: discrim ~ Noise1 + Noise2 + Noise3 + Noise4 + SN1 + SN2 + SN3 +
nosex:     SN4 + (1 | ident)
full: discrim ~ Sex + Noise1 + Noise2 + Noise3 + Noise4 + SN1 + SN2 +
full:     SN3 + SN4 + (1 | ident)

      Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
nosex 11 2066.1 2106.8 -1022.0   2044.1
full  12 2067.7 2112.1 -1021.8   2043.7 0.432      1      0.511

> # Main effect of noise
> nonoise = update(full, . ~ . -Noise1-Noise2-Noise3-Noise4)
> anova(nonoise,full)

refitting model(s) with ML (instead of REML)

Data: NULL

Models:
nonoise: discrim ~ Sex + SN1 + SN2 + SN3 + SN4 + (1 | ident)
full: discrim ~ Sex + Noise1 + Noise2 + Noise3 + Noise4 + SN1 + SN2 +
full:     SN3 + SN4 + (1 | ident)

        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
nonoise  8 2111.8 2141.4 -1047.9   2095.8
full    12 2067.7 2112.1 -1021.8   2043.7 52.133      4  1.294e-10 ***
---
Signif. Codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
>
> # Sex by age by noise
> noise2 = lmer(discrim ~ sex*agefactor*noisefactor + (1 | ident)); anova(noise2)

Analysis of Variance Table
                          Df  Sum Sq Mean Sq F value
sex                        1   19.17   19.17  0.4730
agefactor                  2  433.82  216.91  5.3510
noisefactor                4 2289.31  572.33 14.1189
sex:agefactor              2   30.16   15.08  0.3720
sex:noisefactor            4  142.42   35.61  0.8784
agefactor:noisefactor      8  334.43   41.80  1.0313
sex:agefactor:noisefactor  8  345.66   43.21  1.0659

>
> # With a covariate
> noise3 = lmer(discrim ~ interest + sex*agefactor*noisefactor + (1 | ident))
> anova(noise3)

Analysis of Variance Table
                          Df  Sum Sq Mean Sq F value
interest                   1  260.00  260.00  6.4139
sex                        1   22.07   22.07  0.5445
agefactor                  2  584.43  292.22  7.2088
noisefactor                4 2289.31  572.33 14.1189
sex:agefactor              2    1.66    0.83  0.0205
sex:noisefactor            4  142.42   35.61  0.8784
agefactor:noisefactor      8  334.43   41.80  1.0313
sex:agefactor:noisefactor  8  345.66   43.21  1.0659
```