STAT410/510 Supplement: Introduction to Random Effects

Objectives

- Introduction to Random Effects
- Fitting a model and related diagnostics in R
- Show results with plots & statistics
- Something for the future

What are Random Effects?

If the experiment were repeated	
Fixed effects:	Same levels would be used
Random effects:	Different levels would be used. "Nested" terms are often random effects.
Desired inference? The conclusions refer to	
Fixed effects:	The levels used.
Random effects:	A population from which the levels used are just a (random) sample.

Ex 1. Fixed or Random?

- Does one strain of barley grow faster than another? The strain vs. barley plants
- Do guppies swim faster at higher temperatures? Temperature vs. guppies
- Do starlings grow faster with different feeding schedule? The feeding schedule vs. starlings

The problem is a random effect produces an extra random term:

$$Y = \textit{fixed\_effects} + \textit{error}(\varepsilon)$$
$$Y = \textit{fixed\_effects} + \textit{random\_effects}(\delta) + \textit{error}(\varepsilon),$$
$$\text{where } \varepsilon \text{ has } \sigma^2_\varepsilon \text{ and } \textit{random\_effects} \text{ have } \sigma^2_\delta.$$

1. With random_effects, we now have more than one random term.
2. While the (pure) error is different for each datapoint, a random_effect is not.
3. We partition variance into one thats due to pure error (i.e., σ^2) and another due to the random_effects (i.e., σ^2_δ).

Let's Learn from an Example

To view and play with this sample dataset, go to `http://users.humboldt.edu/ygkim`, click "Sample Data," then click **Chapter 14** tab. It is in columns AJ1 through AN85 (i.e., 5 columns, 84 cases).

Ex 2. Three male students and three female students were randomly selected on a college campus. A researcher prepared one scenario, where each student is going to "ask for a favor." Subjects had to imagine asking a professor (or an authority) for a favor (polite condition) or asking a peer for a casual favor (informal condition). Another scenario was an "excusing for coming too late," which was similarly divided between polite and informal. In total, there were 7 such different scenarios. Thus, the variables are:

`subject` = F1, F2, F3, M3, M4, M7 ("random effect")
`gender` = F, M ("fixed effect")
`scenario` = 1, 2, 3, . . ., 7 ("random effect")
`attitude` = pol, inf ("fixed effect")
`frequency` = (in Hz.) Higher number means higher pitch

First, consider the following models:

```
model1 <- lm(frequency~gender)
model2 <- lm(frequency~attitude)
model3 <- lm(frequency~subject)
```

Can you write these models in SYMBOLS?
What's WRONG with these models?
Answer: We took multiple measurements per subject! That is, each subject gave multiple responses. This violates the independence assumption: multiple responses from the same subject cannot be regarded as independent from each other. These responses are inter-dependent rather than independent. Solution? Add random effect terms!

How to do it with R

```
data1 <- read.csv("U:\\STAT333\\politeness.csv")
dim(data1)
[1] 84  5
head(data1)
  subject gender scenario attitude frequency
1      F1      F        1      pol     213.3
2      F1      F        1      inf     204.5
3      F1      F        2      pol     285.1
4      F1      F        2      inf     259.7
5      F1      F        3      pol     203.9
6      F1      F        3      inf     286.9
library(ggplot2)
ggplot(data1,aes(reorder(scenario,frequency,median),frequency,fill=factor(gender)))+
    geom_boxplot()+xlab("scenario")+
    theme(legend.position="top",legend.direction="horizontal")+
    xlim("1","2","3","4","5","6","7")+
    scale_fill_manual(name = "This is demo", values = c("pink", "green"))
```

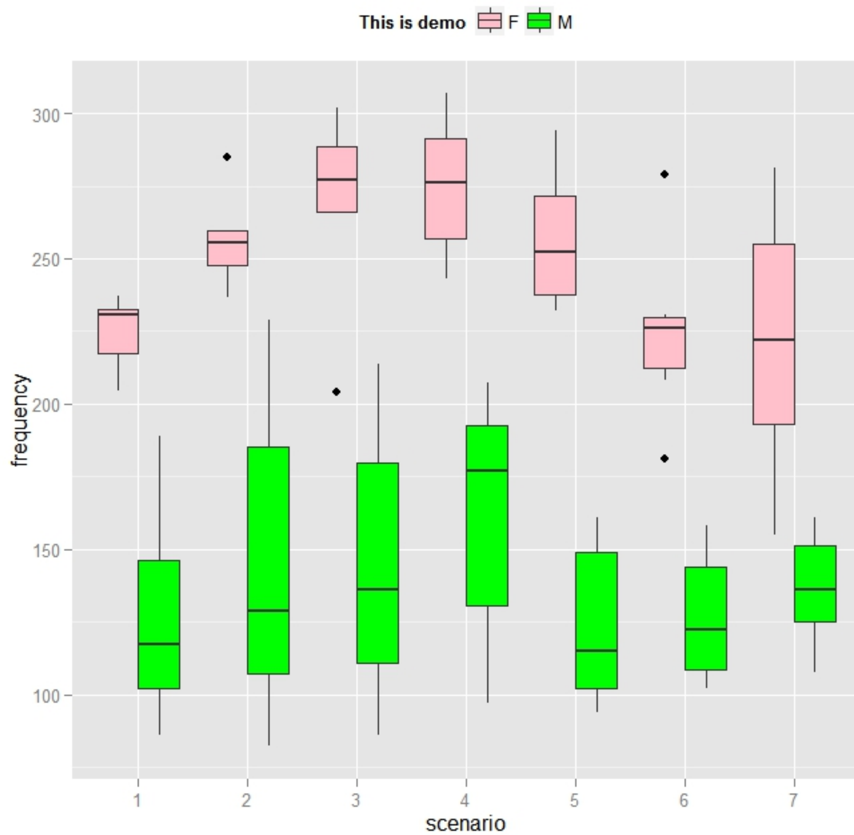

Figure 1: y vs. "scenario" per "gender"

The way we're going to deal with this non-independence is to include a random effect term in the model. This allows us to resolve this non-independence by assuming a different "baseline" frequency value for each subject. Notice that males have lower voices than females. Also, within the male and the female groups, you see lots of individual variation, with some people having relatively higher values for their sex and others having relatively lower values. We can model these individual differences by assuming different random intercepts for each subject. That is, each subject is assigned a different intercept value, and the mixed model estimates these intercepts for you.

By adding one or more random effect terms to the model, these random effects essentially breaks down original random variation into two parts: one due to "(pure) random" errors and another due to different "subjects."

```
library(lme4)
model4 <- lmer(frequency~attitude+gender+(1|subject))
```

Notation: The term `(1|subject)` is saying to "assume an intercept that's different for each subject" (i.e., subject is a random effect term) and "`1`" stands for the intercept here. You can think of this

formula as telling your model that it should expect that there's going to be multiple responses per subject, and these responses will depend on each subject's baseline level. This effectively resolves the non-independence that stems from having multiple responses by the same subject.

```
model5 <- lmer(frequency~attitude+gender+(1|subject)+(1|scenario))
```

So, on top of different intercepts for different subjects, we now also have different intercepts for different scenarios. We have now "resolved" those non-independencies (our model knows that there are multiple responses per subject and per scenario), and we accounted for by-subject and by-scenario variation in y (i.e., overall frequency measurements).

Note the efficiency and elegance of this model. In the past, people used to do a lot of averaging. For example, people would average over scenarios for a subjects-analysis (each data point comes from one subject, thus assuring independence), and then they would also average over subjects for an items-analysis (each data point comes from one scenario). There's much literature on the advantages and disadvantages of this approach. The upshot is: while traditional analyses that do averaging are in principle OK, mixed models give you much more flexibility and they take the full data into account. If you do a subjects-analysis (averaging over items), you're essentially disregarding by-item variation. Conversely, in the items-analysis (i.e., averaging over subjects), you're disregarding by-subject variation. Mixed models account for both sources of variation in a single model.

For demonstration, compare the following two models:

```
model_50 <- lmer(frequency~attitude+(1|subject)+(1|scenario))
model_51 <- lmer(frequency~attitude+gender+(1|subject)+(1|scenario))
```

```
> summary(model_50)
Linear mixed model fit by REML ['lmerMod']
Formula: frequency ~ attitude + (1 | subject) + (1 | scenario)

REML criterion at convergence: 802.5

Random effects:
 Groups   Name        Variance Std.Dev.
 scenario (Intercept)  227.8   15.09
 subject  (Intercept) 4043.4   63.59
 Residual              640.1   25.30
Number of obs: 84, groups:  scenario, 7; subject, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)  202.588     26.864   7.541
attitudepol  -20.188      5.521  -3.657

Correlation of Fixed Effects:
            (Intr)
attitudepol -0.103
```

See that scenario has much less variability than subject (227.8 vs. 4043.3). It means that there are more idiosyncratic differences among subjects than among scenarios. "Residual" stands for the (pure) variability thats not due to scenario or subject. This is the usual variance of the ε term, which is the "random" deviations from the predicted values that are not due to subjects or scenarios.

The fixed effects output is the same as usual "`lm`" printout. Here, the printout says from "informal" to "polite," you have to go down -20.188 (Hz). In other words: pitch is lower in polite speech than in informal speech, by about 20 (Hz), and that's significant (because of the t-value).

```
> summary(model_51)
Linear mixed model fit by REML ['lmerMod']
Formula: frequency ~ attitude + gender + (1 | subject) + (1 | scenario)

REML criterion at convergence: 784.4

Random effects:
 Groups   Name        Variance Std.Dev.
 scenario (Intercept) 227.8    15.09
 subject  (Intercept) 611.6    24.73
 Residual             640.1    25.30
Number of obs: 84, groups:  scenario, 7; subject, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)  257.080     16.102  15.965
attitudepol  -20.188      5.521  -3.657
genderM     -108.983     20.934  -5.206

Correlation of Fixed Effects:
            (Intr) atttdp
attitudepol -0.171
genderM     -0.650  0.000
```

Compared with `model_50` (i.e., the one without fixed effect "gender"), the variation that's associated with the random effect "subject" dropped considerably. This is because the variation that's due to gender was confounded with the variation that's due to subject. The model didn't know about males and females, and so its predictions were relatively more off, creating relatively larger residuals. Since we now have added the effect of gender, we have shifted a considerable amount of the variation that was previously in the random effects component (i.e., differences between males and females) to the "gender" component.

In the fixed effects output, we see that males and females differ by about 109 (Hz). And the intercept is now much higher (257.08 Hz), because it now represents the female category (for the "informal" attitude). The coefficient for the effect of attitude didn't change much.

TESTING for the STATISTICAL SIGNIFICANCE

1. To test for the "attitude" effect, for example, we use the likelihood ratio test between the following two models:

```
model_52 <- lmer(frequency~gender+(1|subject)+(1|scenario),REML=F)
model_53 <- lmer(frequency~attitude+gender+(1|subject)+(1|scenario),REML=F)
```

```
> anova(model_52,model_53)
Data:
Models:
model_52: frequency ~ gender + (1 | subject) + (1 | scenario)
model_53: frequency ~ attitude + gender + (1 | subject) + (1 | scenario)
         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
model_52  5 826.45 838.60 -408.22   816.45
model_53  6 816.04 830.63 -402.02   804.04 12.409      1  0.0004272 ***
```

   You would report this result in the following way:
   "... attitude significantly affected frequency (y) $\left(\chi^2_{df=1} = 12.409, p = 0.004272\right)$, in particular, "pol" (i.e., politeness) lowers frequency by about 20.19 Hz ± 5.52 ..."

2. To test for the "gender" effect, for example, use the following two models:

```
model_54 <- lmer(frequency~attitude+(1|subject)+(1|scenario),REML=F)
model_53 <- lmer(frequency~attitude+gender+(1|subject)+(1|scenario),REML=F)
anova(model_53,model_54)
```

3. To test for the "gender" by "attitude" interaction effecte, use the following two models:

```
model_55 <- lmer(frequency~attitude*gender+(1|subject)+(1|scenario),REML=F)
model_53 <- lmer(frequency~attitude+gender+(1|subject)+(1|scenario),REML=F)
anova(model_53,model_54)
```

RANDOM SLOPES VS. RANDOM INTERCEPTS

```
> coef(model5)
$scenario
  (Intercept) attitudepol   genderM
1    243.6407    -20.1881 -108.9833
2    263.9112    -20.1881 -108.9833
3    268.7796    -20.1881 -108.9833
4    278.0844    -20.1881 -108.9833
5    255.3155    -20.1881 -108.9833
6    243.6609    -20.1881 -108.9833
7    246.1661    -20.1881 -108.9833

$subject
   (Intercept) attitudepol   genderM
F1    243.1697    -20.1881 -108.9833
F2    267.5007    -20.1881 -108.9833
F3    260.5689    -20.1881 -108.9833
```

```
M3    285.9013    -20.1881 -108.9833
M4    261.5902    -20.1881 -108.9833
M7    223.7478    -20.1881 -108.9833

attr(,"class")
[1] "coef.mer"
```

See that each "scenario" and each "subject" is assigned a different intercept. That's what we would expect, given that we've told the model with "`(1|subject)`" and "`(1|scenario)`" to take by-subject and by-scenario variability into account. Note also that the coefficients for the fixed effects ("attitude" and "gender") are all the same for all subjects and items. That's because our model5 is a random intercept model. In this model, we account for baseline-differences in frequency, but we assume that whatever the effect of politeness is, it's going to be the same for all subjects and scenarios.

But is that a valid assumption? In fact, often times it's not – it is quite expected that some scenarios would elicit more or less politeness. That is, the effect of politeness might be different for different scenarios. Likewise, the effect of politeness might be different for different subjects. In that case, you would fit a random slope model, where subjects and scenarios are not only allowed to have differing intercepts, but are also allowed to have different slopes. This is how we would do this in R:

```
model6 <- lmer(frequency~attitude+gender+(1+attitude|subject)+(1+attitude|scenario),REML=FALSE)
```

Notation: The term "`(1+attitude|subject)`" means that there is differing baseline-levels of frequency (the intercept, represented by 1) as well as different responses to "attitude."

```
> summary(model6)
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: frequency ~ attitude + gender + (1 + attitude | subject) + (1 +
    attitude | scenario)

     AIC      BIC   logLik deviance df.resid
   823.8    848.1   -401.9    803.8       74

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 scenario (Intercept) 183.7779 13.5565
          attitudepol  35.7350  5.9779  0.26
 subject  (Intercept) 396.7728 19.9192
          attitudepol   0.9148  0.9565  1.00
 Residual             621.2325 24.9245
Number of obs: 84, groups:  scenario, 7; subject, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)  257.868     13.548  19.033
attitudepol  -20.188      5.903  -3.420
genderM     -110.561     17.479  -6.325

Correlation of Fixed Effects:
            (Intr) atttdp
attitudepol -0.108
```

```
genderM     -0.645  0.000

coef(model6)
$scenario
  (Intercept) attitudepol   genderM
1    245.3222   -20.97369 -110.5606
2    263.2699   -15.86582 -110.5606
3    269.1600   -20.98767 -110.5606
4    276.7772   -16.14591 -110.5606
5    256.0852   -19.74411 -110.5606
6    245.8987   -23.21350 -110.5606
7    248.5655   -24.38598 -110.5606

$subject
   (Intercept) attitudepol   genderM
F1    243.8906   -20.85926 -110.5606
F2    267.0154   -19.74889 -110.5606
F3    260.3888   -20.06707 -110.5606
M3    285.9216   -18.84107 -110.5606
M4    262.9769   -19.94280 -110.5606
M7    227.0170   -21.66948 -110.5606

attr(,"class")
[1] "coef.mer"
```

MODEL DIAGNOTICS: Available Plots

```
> plot(model6)
```

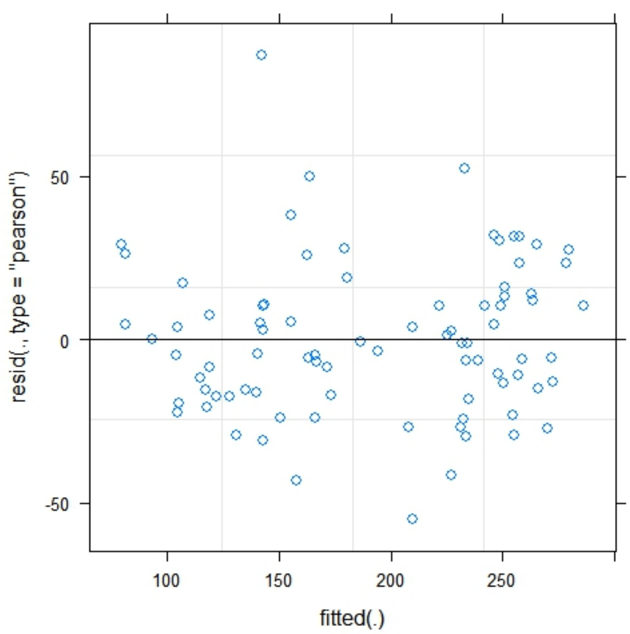

Figure 2: Residual plot

```
plot(model6,frequency~fitted(.))
```

```
plot(model6,frequency~fitted(.)|gender)
plot(model6,frequency~fitted(.)|subject)
qqmath(~resid(model6))

qqmath(~resid(model6)|gender)
qqmath(~resid(model6)|attitude)
qqmath(~resid(model6)|subject)
xyplot(resid(model6)~fitted(model6)|gender, layout=c(3,1))
xyplot(resid(model6)~fitted(model6)|gender)
xyplot(resid(model6)~fitted(model6)|attitude)
```

Some other plots.

```
ggplot(data1,aes(reorder(subject,frequency,median),frequency,fill=factor(attitude)))+
    geom_boxplot()+xlab("subject")+
    theme(legend.position="top",legend.direction="horizontal")+
    xlim("F1","F2","F3","M3","M4","M7")+
    scale_fill_manual(name = "This is demo", values = c("pink","green"))
```

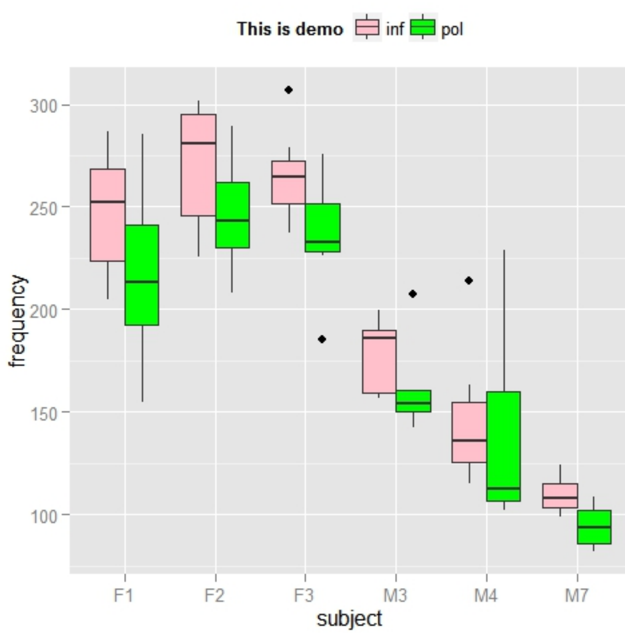

Figure 3: y vs. “subject” per “attitude”

```
ggplot(data1,aes(reorder(scenario,frequency,median),frequency,fill=factor(attitude)))+
    geom_boxplot()+xlab("scenario")+
    theme(legend.position="top",legend.direction="horizontal")+
    scale_fill_manual(name = "This is demo", values = c("pink","green"))
```

More Plots when you have converted it to “groupedData”

```
library(nlme)
```

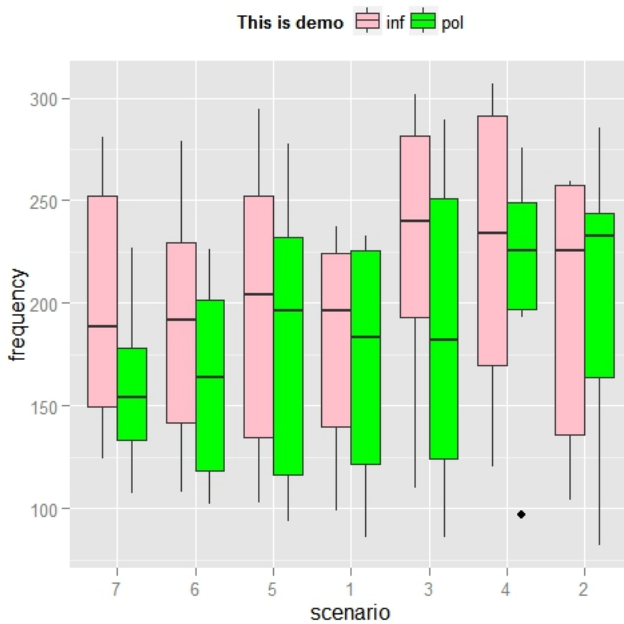

Figure 4: y vs. "scenario" per "attitude"

```
library(lattice)
data2 <- groupedData(frequency~scenario|gender/subject, outer=~attitude, data=data1)
xyplot(frequency~scenario|subject,pch=1:2,col=c(4,2),grid=T,data=data2)
```

```
model6 <- lmer(frequency~attitude+gender+(1+attitude|subject)+(1+attitude|scenario),REML=FALSE,data=data2)
plot(intervals(lmList(data2)))
```

To study more ...

Be sure to think about these sample models. Many are explained in your textbook.

```
model1 <- lmer(CACONTENTS~(1|PLANT)+(1|PLANT:LEAF))
model2 <- lmer(WTGAIN~SIRE+(1|SIRE:DAM), REML=F)
model3 <- lmer(root~fertilizer+(1|plant)+(week|plant))
model4 <- lmer(acuity~power+(1|subject)+(1|subject:eye))
model5 <- lmer(price~items+(1|items)+(store|items))
model6 <- lmer(SSS~test+diet+(1|subject))
model7 <- lmer(log(income)~year*sex+age+educ+(year|person))
model8 <- lmer(yield~irrigation*density*fertilizer+(1|block)+(1|block:irrigation)+
                            (1|block:irrigation:density))
model9 <- lmer(response~gender+(1|town)+(1|town:district)+
                            (1|town:district:street)+(1|town:district:street:house))
model10 <- lmer(Fat~1+(1|Lab)+(1|Lab:Technician)+(1|Lab:Technician:Sample))
model11 <- lmer(Glycogen~Treatment+(1|Treatment:Rat)+(1|Treatment:Rat:Liver))
```

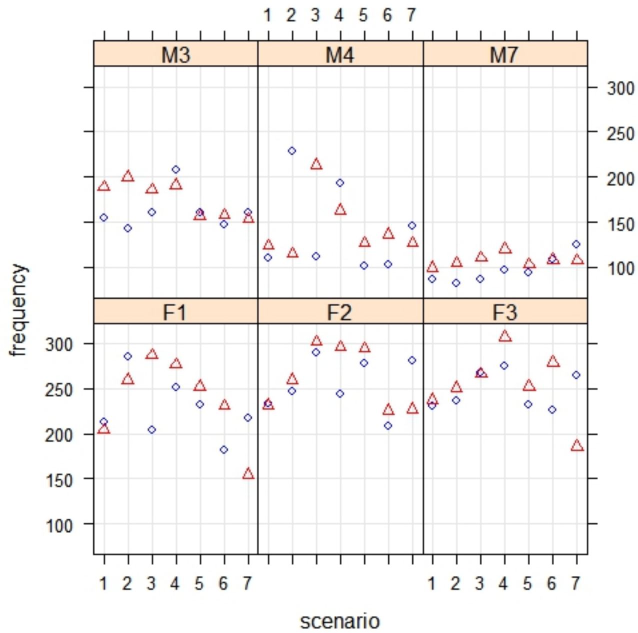

Figure 5: y vs. "scenario" for each "subject"

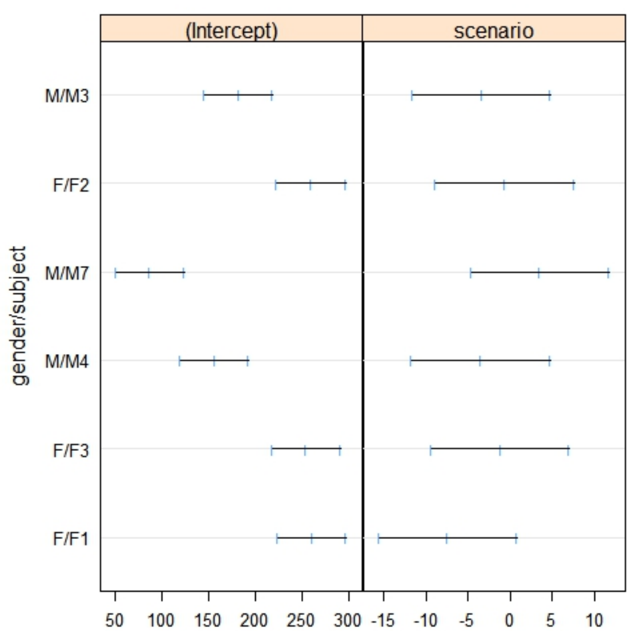

Figure 6: Latticeplot of coefficients