Cheat Sheet Linear Mixed Models

Install and load the NPL package: https://github.com/EricMenetre/NPL
Install and load the lme4, lmerTest, emmeans, optimx, cAIC4 and MuMIn packages. **Do not forget to cite the packages in the publications !**

Step 1: investigate the normality of the different variables

```
exp_plots_LMM(data, data$DV, data$ID, data$class)
```

Perform other plots if needed

Step 2: build an empty model (m0) with the minimum of random variables

```
M0 <- lmer(DV ~ 1 + (1|subject), data = data, REML = FALSE)
```

Step 3: build several models adding always one fixed effect at a time

```
M1 <- lmer(DV ~ IV1 + (1|subject), data = data, REML = FALSE)
M2 <- lmer(DV ~ IV1 + IV2 + (1|subject), data = data, REML = FALSE)
Mn <- lmer(DV ~ IV1 + IV2 + IVn + (1|subject), data = data, REML = FALSE)
```

Do not add the interactions at this step.
REML should be set to FALSE, see step 5

Convergence of the model

Constraint due to the presence of the random effects on the fitting of the curve. If the model does not converge, change the optimizer (function to find the best fit of the regression line)

```
M0 <- lmer(DV ~ 1 + (1|subject), data = data, REML = FALSE,
control = lmerControl(optimizer = "optimx", calc.derivs = FALSE,
optCtrl = list(method = "nlminb", starttests = FALSE, kkt = FALSE)))
```

Step 4: for each model, check the postulates and estimate the marginal and conditional R^2

```
summary(Mn)
LMM_check(Mn)
ICC_ranef(Mn)
r.squaredGLMM(Mn)
```

All the plots should show normal distributions
Note the effect size and the ICC for each model, including the m0

Step 5: add the other eventual random effects

```
Mn <- lmer(DV ~ IV1 + IV2 + IVn (1|subject)+ (1|item), data = data, REML = TRUE)
cAIC(Mn) #See step 6
```

Do not add the interactions at this step.
To compare models based on their fixed effects, we need to set REML = FALSE. When comparing the relevance of random effects, REML must be TRUE

Step 6: add the interactions to the model

```
Mn <- lmer(DV ~ IV1 + IV2 + IVn + IV1:IV2 + (1|subject)+ (1|item), data = data, REML = FALSE)
```

REML must be set as FALSE since we compare fixed effects again

Useful external ressources

- D. Bates lme4 article
- A beginner's guide to mixed models
- How to write random effects in (g)lmer models
- Original article regarding the optimx package
- Post-hocs with emmeans

Step 6: compare the models to find the best one

- Systamatically compare the models based on (information available in the summary):
  - the difference in deviance The AIC; BIC and cAIC (only the cAIC is available in the REML models) see step 5
  - The two R^2
- Select the model with the **lowest** AIC; BIC or cAIC and the **highest** R^2

Step 7: get the main effects and the post-hocs

```
anova(model_opt)
summary(model_opt)
cAIC(Mn)
emmeans(model_opt, list(pairwise ~ IV | IV2), adjust = 'tukey')
```

Step 8: report

Information to report in the publications:

- (F(df) = X.XX; SE = X.XX; p-val = X.XXX) in the text
- As a table with all the information given by the anova and/or the summary/emmeans output

Dictionary:

- ***Fixed effects***: same as the independant variables in the ANOVA.
- ***Random effects***: either random intercept or random slope → part of the variance explained by the fact that a certain subject belongs to a certain group, or that a certain RT belongs to a certain subject.
- ***Marginal R^2***: part of variance explained by the fixed effects.
- ***Conditional R^2***: part of variance explained by the fixed and the random effects.
- ***REML***: adjustment method, either maximum likelihood (ML) or REML for restricted maximum likelihood.
- ***ICC***: interclass correlation → variance inside each random variable.