Cheat Sheet: linear mixed-effects models

Measurement and Evaluation of HCC Systems

Scenario

Use linear mixed-effects models if you want to test the effect of several variables variables `varX1`, `varX2`, … on a continuous outcome variable `varY`, where the *Y* and some of the *X*s are repeated measurements on the same unit (e.g. multiple conditions or time points per participant, multiple people per group, etc.). This cheat sheet assumes that `data` is in the long format, with a user `id` to tie together the repeated measures. (if it is not, use the `melt` function in the `reshape2` package to fix that).

Power analysis

- The power analysis for a linear mixed-effects model is beyond the scope of this course.

Plotting

- For plotting the effects of within-subjects manipulations, see the Plotting sections of the dependent *t* test and repeated and mixed ANOVA cheat sheets.
- For plotting linear effects, see the Plotting section of the regression cheat sheet.

Pre-testing assumptions

- For between-subjects factors, refer to Pre-testing assumptions in the ANOVA and/or factorial ANOVA cheat sheets.
- For linear effects, refer to Pre-testing assumptions in the regression cheat sheet.
- Note that unlike for repeated and mixed ANOVAs, sphericity is not assumed for linear mixed-effects models.

(optional) Preparing dummies and/or contrasts

- If one or more of your *X*s are nominal variables, you need to create dummy variables or contrasts for them.
- For simple dummies, refer to the regression cheat sheet.
- For contrasts, refer to the ANOVA cheat sheet.

Running the test

- Start with the simplest possible baseline: the mean:
  `baseline <- gls(varY ~ 1, data=data, method="ML")`
- Then add the random intercept for id:
  `random <- lme(varY ~ 1, random = ~1|id, data=data, method="ML")`
- Test whether the random intercept is needed:
  `anova(baseline, random)`
  If this ANOVA is significant, you indeed need the random intercept. The χ^2-value and its *p*-value represent the significance of the random intercept.
- Introduce your first *X* variable:
  `model1 <- update(random, .~. + varX1)`
- (optional) Introduce another *X* variable (add more *X*s one by one if needed):
  `model2 <- update(model1, .~. + varX2)`
- (optional) Introduce interaction effects:
  `model3 <- update(model2, .~. + varX1:varX2)`
- Compare all models:
  `anova(random, model1, model2, model3)`
- (optional) Add a random slope for a within-subjects variable (interpretation: is the effect of `varX1` different per `id`?):
  `slope1 <- update(model3, random = ~varX1|id)`
- Compare with the previous model to see if this random slope improves the model:
  `anova(model3, slope1)`
  The χ^2-value and its *p*-value represent the significance of the random slope.
- (optional) Introduce a specific error covariance matrix; e.g. random slopes can be combined with an AR(1) covariance matrix, which makes within-subjects data points that are closer together in time correlate more strongly than data points that are further apart:
  `ARmodel <- update(slope1, correlation = corAR1(0, form= ~varX1|id))`
- Get the summary and confidence intervals of your final model:
  ```
  summary(ARmodel)
  intervals(ARmodel)
  ```
- For the correct interpretation of the coefficients, refer to the regression cheat sheet.
- The only additional thing is the random intercept (which has an *SD*), and optionally the random slope (which also has an *SD*) and the correlation between the two. All these measures have a confidence interval as well.

Post-testing assumptions and inspecting outliers

- For post-testing assumptions and inspecting outliers, refer to the regression cheat sheet.

Reporting

- For random intercepts and slopes: The relationship between [varY] and [VarX1 and other within-subjects variables] showed significant variance in intercept across participants, *SD* = x.xx (95% CI: x.xx, x.xx), $\chi^2(1)$ = x.xx, p = .xxx. In addition, the slopes varied across participants, *SD* = x.xx (x.xx, x.xx), $\chi^2(2)$ = 38.87, p = .xxx, the slopes and intercepts were significantly correlated, cor = .xx (.xx, .xx).
- For reporting further results, refer to the regression cheat sheet. Note that there is no model R^2, and model comparisons are χ^2 tests.