Practical 5. Mixed effects models Hand-in: 13/11/2021

Resolve the following exercises in groups of two students. Write your solution in a Word, Latex or Markdown document and **generate a pdf file** with your solution. Upload the pdf file with your solution to the corresponding task at the Moodle environment of the course, no later than the hand-in date. Please take care to **write your names and surnames** on the first page of your report.

1. **Drug pharmacokinetics** (20p). The dataset `pharmacokinetics.dat` contains data regarding the concentration (`Conc`) of the drug cefamandole taken by volunteers (identifier `Subject`). Plasma concentrations of the drug were taken repeatedly over time (`Time`). Linear models and linear mixed models are considered to study the relationship between the concentration `Conc` and `Time`.

   (a) Read the datafile `pharmacokinetics.dat` into the R environment.

   (b) (1p) Does this data have a hierarchical structure? Do you have any a priori reasons to expect that plasma `cefamandole` measurements could not be independent? Transform the concentrations of `cefamandole` by taking their natural logarithm. Use the transformed concentration for all analyses of the rest of this practical.

   (c) Format the data for its use by the functions of the `nlme` package with the instruction:
   `X <- groupedData(Conc~Time|Subject,data=X)`

   (d) (1p) What is the total number of volunteers? How many measurements were taken on each volunteer?

   (e) (1p) Fit an ordinary linear regression model, with `Time` as the predictor and `Conc` as the response. Is there a significant relationship between these two variables? How much variance of `Conc` can be explained by `Time`?

   (f) (1p) Show the data adequately in a plot by adding the relationship obtained by the regression model.

   (g) (2p) Make the standard plots for the residuals of this regression (histogram, residuals versus fitted values, residuals versus order, normal probability plot) and indicate whether you believe if the standard regression assumptions hold or not.

   (h) (2p) Make boxplots of the residuals for each volunteer. Do you observe any problems?

   (i) (1p) Do separate regressions for each volunteer using the `lmList` instruction, and create all 95% confidence intervals for the intercepts and the slopes, using the `intervals` function. Display all intervals in a graph. Do you think intercepts and slopes vary significantly across volunteers? What model do the graphs suggest you?

   (j) (2p) Fit a random intercept model to the data with `lme`. Use the output to obtain an estimate of the intraclass correlation coefficient. Do you think observations are independent?

   (k) (2p) Compare the ordinary regression model with the random intercept model. Which model fits the data better? Give the value of the corresponding test statistic, its reference distribution and the p-value.

   (l) (2p) Fit a mixed model with random intercept and random slope for `time`. Does this model fit better than a random intercept only model? How many parameters has the best fitting model?

   (m) (1p) Investigate the residuals of the model of your choice by making plots. Comment on your results.

   (n) (2p) Investigate if there is evidence for a quadratic dependence of `Conc` on `Time`. Comment on your results.

   (o) (1p) Is there evidence for a random quadratic effect of `Time` on `Conc`? Investigate the issue by a hypothesis test and state your conclusion.

   (p) (1p) Make a scatterplot of `Conc` against `Time` and show the curves fitted to the data by the final model of your choice.