Exercises for course day 4

Statistical analysis of repeated measurements and clustered data

Exercise 1

The data `bloodpressureL` (long format) contains results from a crossover trial in which 12 male volunteers received three different formulations of a drug for lowering blood pressure:

Treatment A: 50 mg tablet

Treatment B: 100 mg tablet

Treatment C: Sustained-release formulation capsule

The volunteers were randomized to sequences ABC, BCA, and CAB, with a one week wash-out between the treatments. The outcome was the duration of the drug, in hours.

1. Load the data and make a suitable spaghettiplot to view it. What trends do you see?
2. Compute summary statistics for each combination of treatment, period, and sequence and use these to visualize potential trends in the sample means:
   (a) Plot treatment on the x-axis and connect points for each sequence.
   (b) Plot period on the x-axis and connect points for each sequence.
   (c) Plot treatment on the x-axis and connect points for each period.
   (d) Plot period on the x-axis and connect point for each treatment.

   What impression do you get of the effects of treatment and period?
3. Make a state-of-the-arts linear mixed model analysis to compare the duration of the drug between the three treatments while adjusting for potential effects of period.
   (a) Estimate the mean differences in duration between the treatments. Don't forget the 95% confidence intervals. What can you conclude?
   (b) Are there any substantial effects of period? Should you be concerned about these?
   (c) Take a look at the estimated residual correlations. Do they confirm that the crossover design is advantageous compared to an independent samples design?

(d) Use residualplots to check whether the durations are approximately normally distributed. Does it matter if they are?

4. Is it important to adjust for the period effect? What happens if we don't?

5. Assuming that the study is a pilot study to investigate whether the wash-out is sufficient, make an analysis to estimate potential carryover effects. You can do this e.g. by going through the following steps:

   (a) Make a new variable `pre` that describes the treatment in the previous period. Assuming that your data is named `bp`, you can use the following code:

   ```
   bp$pre <- rep("N", dim(bp)[1])
   bp$pre[bp$sequence=="ABC" & bp$period=="2"] <- "A"
   bp$pre[bp$sequence=="ABC" & bp$period=="3"] <- "B"
   bp$pre[bp$sequence=="BCA" & bp$period=="2"] <- "B"
   bp$pre[bp$sequence=="BCA" & bp$period=="3"] <- "C"
   bp$pre[bp$sequence=="CAB" & bp$period=="2"] <- "C"
   bp$pre[bp$sequence=="CAB" & bp$period=="3"] <- "A"
   bp$pre <- factor(bp$pre)
   ```

   Set the reference category to "`N`" correspondig to no previous treatment.

   (b) Estimate the effects of the previous treatment in the linear mixed model including `treatment` and `pre` as fixed effects (leave out `period` for the moment). What can you conclude?

   (c) **Difficult:** Try to run the analysis with `treatment`, `pre`, and `period` as fixed effects. Can you figure out why this analysis is problematic?