Exercises day 8

Basic Statistics for health researchers 2025

March 19, 2025

Warming up

Before starting the exercises, learn from the R-demo of Lecture 8 (available from the course webpage):

1. Read and run the code.
2. Check that the output matches the results presented on the slides.
3. Do not hesitate to add your own comments into the script.

Exercise A: Follicle data

For this exercise we will work with the "Human Follicle data" seen in the lecture. We will perform similar analyses of the follicle growth, but at earlier timepoints.

1. We first analyze the data of the follicle growth at day 6.
   (a) Make the necessary data management steps, similar to those made to study follicle growth at day 8 in the R-demo.
   (b) Fit an appropriate random effect model.
   (c) Report appropriate effect sizes and 95%-CI to compare the efficacy of the four plasma products on the follicle growth.
   (d) Are the results similar with those of the growth at day 8, seen in the lecture?
   (e) How large is the estimated intra-class correlation? How can this be interpreted?
2. Same question, but at days 2 and 4 instead of day 6. **Hint:** you can use more or less the exact same R code, after e.g., running this:

   ```
   whichDay <- 2 # just change to 2, 4, 6 or 8 as needed
   d$growth <- d[,paste0("Day",whichDay)]-d$Day0
   ```

   as well as the code below, to automatically extract the estimated values for ω_B and τ_W, to compute the intra-class correlation ρ, for the question of item (e),

```
ResSD <- as.data.frame(VarCorr(fitlmer))
omegaB <- ResSD[1,5]
tauW <- ResSD[2,5]
```

3. Are your conclusions consistent with that of Figure 3.B in Cristina's paper? `https://doi.org/10.1016/j.rbmo.2023.06.011`

Exercise B: knee surgery trial data

For this exercise we will work with the "Knee Surgery data" available from the course webpage. We will perform the analyis of an important outcome of the trial: the change in Oxford Knee Score at 24 months after surgery (the change is defined as the difference with the score before surgery). We will assume that, as often, adjustment for a few baseline covariates has been pre-specified in the statistical analysis plan (score before surgery, age, sex and study site). Covariates adjustment is often recommended as it usually leads to power gains, whithout increasing the risk of type-I error.

1. Load the data, visualize the first lines of the data and create a baseline table to summarize the distribution of study site, age, sex and baseline Oxford score, **per study arm**. Are their any noticeable differences? Why?

2. Produce a "spaghetti plot" for approximately 10 random patients of each arm. You can proceed as follows:

   (a) select 20 random rows of the dataset using, e.g., `d20 <- d[101:120,]`.

   (b) create a "long" format version `long20` of the dataset `d20`, using the `reshape()` function.

   (c) use the `xyplot()` function (after making the variable `arm` a factor variable).

3. Interpret the plot: Are the scores improving over time? Is there substantial patient to patient variability? Is there a clear, substantial, difference between the two arms (types of surgery)? Are your observations different if you randomly choose 20 other patients (e.g., `d20 <- d[1:20,]`).

4. Produce a relevant plot to describe the missing data pattern in the longitudinal data of the score.

   (a) Does it look like some patients have dropped out within 24 months (i.e., do no longer want to answer questionnaires)?

   (b) Were some patients not available to answer questionnaires at some early timepoints but available at later timepoints (i.e., intermittent missing data)?

   (c) For how many patients do we have the scores at all times?

   (d) For how many patients do we have the score at 24 months (i.e., the main outcome)?

5. We will assume that the data are either "Missing Completely At Random" (MCAR) or "Missing At Random" (MAR). What does it mean in our context?

6. We now perform the main analysis, using the following R code to fit a Mixed Model for Repeated Measurements (MMRM). First, we create the data set `long` in the "long" format. Second, we "center" the covariate `age` and Oxford score pre-surgey and choose 24 months as the reference level for the factor variable `time`. This is just to facilitate the interpretation of the default output of the software.

```
#-- make long format data from wide format data ----
long <- reshape(kneeSurgery,
                varying = c("Oxford.01", "Oxford.06","Oxford.12", "Oxford.24"),
                v.names = "Oxford",
                timevar = "time",
                times=c(1,6,12,24),
                direction = "long")
long <- long[order(long$id),] # reorder by subject ID
rownames(long) <- NULL # delete row names
head(long,n=5) # quick check
#--- data management steps ---
long$time <- factor(long$time)
long$time <- relevel(long$time,ref="24")
long$arm <- factor(long$arm)
long$change <- long$Oxford-long$Oxford.pre
long$age67 <- long$age-67
long$Oxford.pre22 <- long$Oxford.pre-22
#--- fit MMRM --------------------
lmmfit <- lmm(change~Oxford.pre22*time + site*time
              + sex*time + arm*time + age67*time,
              repetition = ~time|id,
              structure = "UN", data = long)
summary(lmmfit)
```

7. What is the main conclusion: do the results suggest that this trial brings sufficient evidence that one type of surgery is better than the other, for the mean Oxford score at 2 years?

8. Is there substantial evidence that one type of surgery is "much better" than the other? **Hint:** compare the estimated between-arm difference in mean of the change score to the patient to patient variability of the change score.

9. What is the interpretation of other parameters of the model? Especially:

   (a) Is the between-arm difference in mean change score similar at earlier timepoints? **Hint:** you can make checks using our "favorite trick": re-fit the model after appropriately changing the reference level of the appropriate covariate.

   (b) Does it seem that adjusting on sex and/or age was very important for the power of the trial?

10. A simpler analysis would have been to use a "complete case analysis" with an ANCOVA model. To do so, use the following R code:

```
dCCA <- d[!is.na(d$Oxford.24),]
dCCA$change <- dCCA$Oxford.24 - dCCA$Oxford.pre
fitANCOVA <- lm(change~Oxford.pre + site + sex + arm + age,data=dCCA)
```

(a) How many subjects are excluded from this analysis?

(b) How different or similar are the results of this analysis? Can you explain or hypothesize why?

11. By the way, an even simpler analysis would have been to use a "complete case analysis" with a t-test. What would have been the results and main conclusions? Can you remember why this is not the recommended approach?

12. Many researchers wonder whether they should define their primary outcome as the score at end of follow-up or as the change score at end of follow-up. The previous questions consider the later. Perform the MMRM analysis and t-test analysis using as outcome the score at end of follow-up, instead of the change from baseline. How different are the results? Try to explain or hypothesize why.