SPE-2017: postmortem

Next year. Martyn explores the possibility to organize the course in Lyon in June, 2018, before 21.6., and will report about the IARC response by the end of June 2017. Dates: Thursday 7 to Tuesday 12, or Thursday 14 to Tuesday 19. Sunday is day off, but Saturday is a whole course day.

Martyn and Damien have set up the repository SPE-R at GitHub. Everybody opens a github account and utilizes the repository.

Brief instructions:

1. Launch GitKraken
2. Pull
3. Go to the (sub)folder containing your own copy of the repository
4. Edit & save the source file
5. Stage file
6. Commit message: write a comment
7. Commit changes
8. Push

General points

It seems desirable to address tools in `tidyverse` in some way at some point, especially those for data processing: `dplyr` package, `data.table` class, ...

Create an optional exercise on advanced data handling in the first day for those already sufficiently familiar with R. – Janne & Martyn will prepare this.

Style of writing the code could be more consistent. Sweave actually would do it, if the pertinent parameter is not switched off.

Day 1

Lecture on history and ecology of R was OK.

Lecture on language and basic data is too detailed for newcomers at this stage of learning; impossible to digest and remember all of it. One should focus on the most essential items and leave less important ones aside or to be introduced in the exercise sheets. For instance, give some more time

in the lecture on function calls, encouraging the use of explicit naming of arguments. Emphasize this also in the exercises – Krista will develop the lecture along these lines.

Exercises 1.1 and 1.2: Only few participants reached 1.2, and many people could do 1.1 only to about 1.1.10, which is normal. In 1.2.1, provide the exact address from which the .zip containing the data sets can be downloaded to one's own computer and finally to the working directory. In the end of the exercise it would be nice to provide instructions on how to read data files from an url address. – Martyn.

Exercises 1.3 on tabulation and exercise 1.4 on graphics were generally OK.

Day 2

Lecture on Poisson & logistic regression. The presentation of diabetes example needs some editing. Analysis on the outbreak study was missing. The odds ratio OR associated with chocolate mousse cake was actually nearly 30, but the relative risk RR was only about 12, so in this instance OR cannot be interpreted to approximate RR. – Janne changes this after getting some numerical results from Esa.

Lecture on splines: Martyn revises the introduction explicating more clearly the aims of it.

Exercise 1.6: Technical problems in finding tilde '∼' for at least the Estonian and Spanish keyboards. Add to pre-course instructions: "Find out, how to get tilde from your local keyboard. Consult `https://en.wikipedia.org/wiki/Tilde`, especially the end of it". Same problems with the caret symbol! Instruct to do the optional subsection 1.6.7 only after you have done the other exercise of this session. ***Write this in bolded font***. – Esa.

Exercise 1.7: Modify some rows. In logical expressions change `T` and `F` into `TRUE` and `FALSE`! Introduce itemization of the tasks as in other practicals. – Janne

Exercise 1.8, item 11: Change at least `strata = gest4` – Esa

Exercise 1.9: Develop `plotPenSplines.R` to a more general function. See `Termplot()` and `termplot()`.

Day 3

Lecture on graphics: Maybe something more about `ggplot2` and less on `lattice` – Martyn modifies.

Exercise 1.10: Generally OK. Good to modify some tasks, like removing item 3 to be item 8, so that items 8 to 10 are presented as more advanced. – Martyn will do.

Day 4

Lecture on survival: In the post-mortem of 2016 it was said: "Should be shortened and somewhat reorganized.". Still there are 55 slides out of which only ~ 42 could be covered. It would be desirable to have time for the basics of relative survival, too, during the lecture. Include a brief comment on the principle of comparing O with E being applied also in SMR. Either speed up the oral presentation, or cut down the number of slides. Reduce the amount of mathematics but refer to Bendix's 5 page intro on concepts; see `http://bendixcarstensen.com/AdvCoh/papers/fundamentals.pdf`. Also, remove 21 to 23 on lung cancer risk up to high age, and 27 on `Lexis()`. Update slide 26 on R tools for competing risks; refer to `https://cran.r-project.org/web/views/Survival.html`. Perhaps `mstate` does not deserve special attention any more. Recall that `survfit()` in package `survival` computes AJ-estimates when argument `type="mstate"` is added in the call of `Surv()`, like is done in Exercise 1.11, item 13. – Janne does all of this.

Exercise 1.11: Nothing special to be done.

Lecture on representation of follow-up & SMR: When explaining SMR $= O/E$ make a comment of an analogy with relative survival. – Bendix.

Exercise 1.12: Basic calculation of SMR before spline modelling. Take out models based on unpenalized splines. Reduce the amount of typing by providing a decent function for repeated tasks, or a "housekeeping" script. – Bendix.

Day 5

Lecture on NCC & CC: Shorten the introductory part. Slide 16 is a bit crowded. – Esa.

Exercise 1.14 on NCC & CC: For many this was the first time when `merge()` was asked to use. Make a reference to the previous exercise where it was introduced. Change logicals "`T`" and "`F`" into `TRUE` and `FALSE`. Make sure that the right version will be included in `pracs.tex` and `pracs.pdf` in 2018. – Esa.

Lecture on causal inference: Some of the slides are too crowded. – Krista.

Exercise 1.13 on causal inference: Seems OK.

Day 6