Practicals

Practical 1: Marginal Models Continuous

- We will use the PBC dataset; this is available as the object `pbc2` in the R workspace available on GitHub

- To load this workspace and make the data available execute the following steps:
  1. Open a new Rstudio session
  2. Create a new R script file (File → New File → R Script)
  3. Copy-paste and execute the following three lines

```
con <- url("https://raw.github.com/drizopoulos/Repeated_Measurements/master/Data.RData")
load(con)
close(con)
```

Practical 1: Marginal Models Continuous (cont'd)

- The data are available in the data frame `pbc2` – we will need the following variables
  * `id`: patient id number
  * `prothrombin`: prothrombin time in sec (the response variable of interest)
  * `year`: follow-up times in years
  * `drug`: the randomized treatment
  * `sex`: the gender of the patients
  * `age`: the age of the patients

Aim: To build an appropriate marginal model to investigate the relationships between the prothrombin time and the aforementioned variables

Practical 1: Marginal Models Continuous (cont'd)

- Q1: We will start by producing some descriptive plots for the prothrombin time, similar to those we have seen in Chapter 1, i.e.,
  - ▷ spaghetti plot per treatment group including the loess curve
  - ▷ spaghetti plot per sex including the loess curve

  (hint: see code for Section 1.1)

What observations can you make?

Practical 1: Marginal Models Continuous (cont'd)

- We will continue by starting our model building exercise
  Remember
  - ▷ we start with a full specification of the mean structure, and investigate the covariance structure
  - ▷ based on our chosen covariance structure we can make inferences for the mean structure

- Q2: Start by fitting a marginal model with independent error terms using `gls()` and the following specification of the mean structure (hint: see code for Section 2.4)
  - ▷ nonlinear time evolutions using natural cubic splines with 3 degrees of freedom
  - ▷ correct for `sex`, `drug` and `age`
  - ▷ interactions of the time effect with `sex` and `drug`

Practical 1: Marginal Models Continuous (cont'd)

- Q2:
  - ▷ interpret the results you obtained
  - ▷ should we simplify the model by excluding the non-significant terms?

- Q3: Continue with the same mean structure and try different covariance structures
  - ▷ first try different correlation structures, i.e., compound symmetry, continuous AR1, linear& Gaussian, and
  - ▷ then extend the above structures by assuming heteroscedastic errors, i.e., that the variance increases (or decreases) with time

  (hint: see code for Section 2.9)

Practical 1: Marginal Models Continuous (cont'd)

- Q4: Using appropriate tools (hypothesis tests, information criteria) decide which structure is the best
  - ▷ which models are nested to which models?

- For the remainder we will use the covariance structure you have chosen in Q4

- Q5: Check if we can drop **all** the interaction terms
  - ▷ with an F-test
  - ▷ with a Likelihood Ratio Test

  (hint: see code for Section 2.9)

Practical 1: Marginal Models Continuous (cont'd)

- Q6: Continue and check whether you can drop the nonlinear terms for the time effect
  - ▷ to do that fit a model that assumes a linear time trend, and
  - ▷ then do the likelihood ratio test to compare it to the model that includes the nonlinear terms

- Q7: Interpret the results of your final model
  - ▷ regression coefficients
  - ▷ covariance structure

Practical 1: Marginal Models Continuous (cont'd)

- Q8: Use an Effect Plot to depict the model with the following settings
  - ▷ `year`: in the range from 0 to 12 years of follow-up
  - ▷ `sex`: both males and females
  - ▷ `drug`: both treatment groups
  - ▷ `age`: fixed at 49 years old

(hint: see code for Section 2.4 – Effect Plot)

Practical 1: Marginal Models Continuous (cont'd)

- Q9: Check the assumptions of the model using scatterplots of the standardized & normalized residuals versus the fitted values,
  - ▷ overall
  - ▷ separately per sex
  - ▷ separately per treatment group

  (hint: see code for Section 2.11)

  What are your conclusions?

Practical 2: Mixed Models Continuous

- We will use the PBC dataset; this is available as the object `pbc2` in the R workspace available on GitHub

- To load this workspace and make the data available execute the following steps:
  1. Open a new Rstudio session
  2. Create a new R script file (File → New File → R Script)
  3. Copy-paste and execute the following three lines

```
con <- url("https://raw.github.com/drizopoulos/Repeated_Measurements/master/Data.RData")
load(con)
close(con)
```

Practical 2: Mixed Models Continuous (cont'd)

- The data are available in the data frame `pbc2` – we will need the following variables
  * `id`: patient id number
  * `prothrombin`: prothrombin time in sec (the response variable of interest)
  * `year`: follow-up times in years
  * `drug`: the randomized treatment
  * `sex`: the gender of the patients
  * `age`: the age of the patients

Aim: To build an appropriate linear mixed effects model to investigate the relationships between the prothrombin time and the aforementioned variables

Practical 2: Mixed Models Continuous (cont'd)

- Q1: Compute summary statistics for the number of repeated measurements per patient
  - ▷ do we have enough information to model potential nonlinearities in the subject-specific trajectories?

- Q2: Examine graphically for samples of patients

  (hint: see code for Section 1.1)
  - ▷ How do the individual longitudinal trajectories of the prothrombin time look like?
  - ▷ What observations can you make?

Practical 2: Mixed Models Continuous (cont'd)

- Q3: Start by fitting a linear mixed effects model using `lme()` with the following specification of the mean fixed and random effects

  (hint: see code for Section 3.2)

  ▷ *fixed effects:*

  * linear & quadratic time evolutions, nonlinear effect of age using natural cubic splines with 3 degrees of freedom
  * correct for `sex` and `drug`
  * interactions of time with `sex` and `drug`, and `age` with `sex` and `drug`

  ▷ *random effects:* random intercepts

Note: As in Practical 1, in the analysis requested above, and for the remainder of this practical exclude the prothrombin times that were above 18 sec.

Practical 2: Mixed Models Continuous (cont'd)

- Q4: Keeping the mean structure (i.e., the fixed effects as is), start elaborating the random-effects structure that captures the within subject correlation, i.e., consider
  - ▷ random intercepts & random slopes
  - ▷ random intercepts, linear & quadratic random slopes
  - ▷ random intercepts, linear, quadratic & cubic random slopes

  For each extra random effect that you add, perform the likelihood ratio test to see if it is required to add it
  - ▷ which are the null and alternative hypotheses for each of these tests?

Practical 2: Mixed Models Continuous (cont'd)

- Q5: Based on the model you selected Question 4, test whether you can drop all the *interaction terms* in order to simplify the model
  - ▷ first perform the omnibus test for all the interaction terms
  - ▷ if it is (highly) non-significant, you can drop them
  - ▷ if it is significant, find which group(s) are the significant ones

- Q6: In the same spirit as in Question 5, test whether you can drop all the *nonlinear terms* to simplify the model
  - ▷ first perform the omnibus test for all the nonlinear terms
  - ▷ if it is (highly) non-significant, you can drop them
  - ▷ if it is significant, find which group(s) are the significant ones

Practical 2: Mixed Models Continuous (cont'd)

- Q7: Interpret the results of your final model
  - ▷ regression coefficients
  - ▷ covariance structure

- Q8: Compare the marginal and subject-specific predictions from your final model, i.e.,
  - ▷ add in you data frame the marginal and subject-specific fitted values from the final model (remember to use the dataset that excludes the outliers)
  - ▷ select the following patients from the data set: 133, 36, 180, 11, 168, 116, 70, 58, 82, 104, 43, 21, 101, 210, 176, 157
  - ▷ create the plot that compares the predictions

  (hint: see code for Section 3.4)

Practical 2: Mixed Models Continuous (cont'd)

- Q9: Use an Effect Plot to depict the model with the following settings
  - ▷ `year`: in the range from 0 to 12 years of follow-up
  - ▷ `sex`: both males and females
  - ▷ `drug`: both treatment groups
  - ▷ `age`: the median age from the original data for the respective four groups of patients (i.e., the median age of male in placebo, females in placebo, males in active treatment & females in active treatment)

(hint: see code for Section 3.2 – Effect Plot)

Practical 2: Mixed Models Continuous (cont'd)

- Q10: Check the assumptions of the model using scatterplots of the standardized subject-specific & standardized marginals residuals versus the fitted values,
  - ▷ overall
  - ▷ separately per sex
  - ▷ separately per treatment group

  (hint: see code for Section 3.11)

  What are your conclusions?

Practical 3: Marginal Models Discrete

- We will use the PBC dataset; this is available as the object `pbc2` in the R workspace available on GitHub

- To load this workspace and make the data available execute the following steps:
  1. Open a new Rstudio session
  2. Create a new R script file (File → New File → R Script)
  3. Copy-paste and execute the following three lines

```
con <- url("https://raw.github.com/drizopoulos/Repeated_Measurements/master/Data.RData")
load(con)
close(con)
```

Practical 3: Marginal Models Discrete (cont'd)

- The data are available in the data frame `pbc2` – we will need the following variables
  * `id`: patient id number
  * `prothrombin`: prothrombin time in sec (the response variable of interest)
  * `year`: follow-up times in years
  * `drug`: the randomized treatment
  * `sex`: the gender of the patients
  * `age`: the age of the patients

Aim: To build an appropriate GEE model to investigate the relationships between a dichotomized version of the prothrombin time and the aforementioned variables

Practical 3: Marginal Models Discrete (cont'd)

- Q1:

- Q2:

Practical 3: Marginal Models Discrete (cont'd)

- Q3:

- Q4:

Practical 4: Mixed Models Discrete

- ***

Practical 4: Mixed Models Discrete (cont'd)

- ***