

This exercise sheet will familiarize you with generalized linear mixed models (GLMMs). The exercises refer to the content of the eighth and ninth lecture slides.

### Exercise 1: GLMM for binary data

The data set `madras` included in the package `drm` comprises monthly records (presence/absence) of the psychiatric symptom “thought disorder” of 86 patients over the first year after initial hospitalization for disease.

- (a) Load the data set and take a first look at the data.
- (b) Since we are interested in individual-specific effects for this data, a GLMM including fixed effects for `month` and `sex` as well as a random intercept per patient shall be estimated. Formulate the model and specify the distributional assumptions, the systematic component, as well as a suitable link function.
- (c) What differentiates the estimation of a GLMM from that of an LMM? What makes the estimation so complex?
- (d) Which three options for estimating GLMMs did you get to know? What is approximated in each case?
- (e) Estimate the model with each of these three options. Use the functions of the packages `MASS` and `lme4` which were presented in the lecture. Use 11 quadrature points for the adaptive Gaussian quadrature (AGQ). Compare the estimates of the fixed effects.
- (f) Why might PQL estimation be problematic for this model scenario?
- (g) Interpret the AGQ estimates of the fixed effects.
  - (i) Why can the coefficient estimates not be interpreted marginally (in general)?
  - (ii) Plot the individual probabilities  $P(Y_{ij} = 1|b_i)$  over time (`month`) by gender.
  - (iii) Add the probability curve over time for  $b_i = 0$  to both plots.
  - (iv) Now calculate the average probabilities  $P(Y_{ij} = 1)$  separately for men and women and add these to the plots as well. How can the resulting plots be interpreted?
- (h) You now want to perform an LR test to check whether there is an interaction between the sex of the patient and time.
  - (i) Why is an estimate based on PQL not suitable for this?
  - (ii) Use the AGQ estimation instead. What is your test decision?