# Longitudinal Data Analysis

## Prof. Dr. Holger Brandt

### December 11, 2023

## Assignment 8: Bayesian Growth Curve Mixture Models (GCMM)

This assignment uses the autism data from assignment 5. In this analyses, we will try to identify subgroups with different patterns.

For all models, specify a model with intercept and linear slope factor. Use `jags` for these analyses.

1. Specify a GCMM with two latent classes. Class #2 is different from class #1 in the average slope parameter. For this class #2, we assume a higher average values in the slope factor (i.e., these are the children with a strong increase in social skills). This is achieved by using a censored normal distribution (with $N(0, 0.01)^+$ in precision notation).

   (a) *Estimate the model with 20,000 iterations and a burn-in of 1. Investigate the Rhat statistics (Gelman plot) to identify how many burn-in iterations you need. [1 pt.]*

   (b) *From the same model, investigate the trace plots. Which parameters need longer to converge. Try to provide a reason for this behavior. [2 pts.]*

   (c) *Rerun the analysis with the selected burn-in rate. Report on the results (parameter estimates, Rhat, effective sample size). [2 pts.]*

2. Adapt the model in two steps: (a) with a class-specific covariance matrix for the factors; and (b) in addition class-specific residual variances. Use the same number of iterations and burn-in as above.

   (a) *Compare the model fit via the DIC provided by these models to the original model and a simple latent growth curve model without class. Which model do you prefer? [1 pt.]*

   (b) *Shortly describe the parameter estimates from the models and how they differ from each other. Are there any problems with convergence? [3 pts.]*

3. Extract the class memberships from model 2(a) (i.e. with class-specific covariance matrix in addition to the class-specific slope mean).

   (a) *Illustrate the development trajectories (spagetthi plots) separate for each class and describe the results. [1 pt.]*

   (b) *Compare the class membership with the observed grouping variable (low, medium, high language skills at age 2). Describe the findings. [1 pt.]*

   * *(Extrapoint): Would it make sense to include a class-specific mean in addition to the model investigated? Why or why not [+1 pt.]*