Modelling longitudinal cognitive test data with ceiling effects and left skewness

Denitsa Grigorova*, Dean Palejev **, Ralitza Gueorguieva ***

* Big Data for Smart Society Institute, Sofia University
125 Tsarigradsko Shosse, Bl. 2, 1113 Sofia, Bulgaria and
Faculty of Mathematics and Informatics, Sofia University
5 James Bourchier Blvd., 1164 Sofia, Bulgaria
denitsa.grigorova@gate-ai.eu, dgrigorova@fmi.uni-sofia.bg
** Institute of Mathematics and Informatics, Bulgarian Academy of Sciences
Acad. G. Bonchev St., Bl. 8, 1113 Sofia, Bulgaria
palejev@math.bas.bg
*** Department of Biostatistics, Yale School of Public Health,
60 College St, New Haven, CT 06520, U.S.A
ralitza.gueorguieva@yale.edu

Cognitive tests are among the markers for the development of cognitive diseases such as Alzheimer's disease. We model the scores from the Mini Mental State Examination (MMSE, discrete values in the range 0-30) over time on data from the Alzheimer's Disease Neuroimaging Initiative (ADNI, http://adni.loni.usc.edu/). The challenge of modelling such an outcome as MMSE is that the data are left-skewed with ceiling effect - the maximum possible score on the MMSE is 30 and this maximum is often achieved by healthy individuals (the higher the value of MMSE the better cognitive function of the individual). Different approaches for modeling MMSE have been considered in the statistical literature, such as linear mixed effects models on transformed data, mixture models based on latent class growth analysis and generalized additive models for location, scale and shape (GAMLSS). We find models such as binomial and beta-binomial from GAMLSS more appropriate for the MMSE score and we apply them to the ADNI data. We use random effects (parametric and non-parametric) in the models to account for correlations among repeated measures on the same individual. The estimation is based on the maximum likelihood approach. Using Bayesian Information Criterion, we select the best model that fits the data. Additionally, we propose a bootstrap approach for estimation of the covariance matrix of the estimates. Using statistical tests and inference we compare the cognitive function over time for individuals with cognitive impairment and normal controls in the ADNI data set. We also perform simulation studies with different sample sizes that evaluate the binomial and beta-binomial models in terms of bias and efficiency.