# Misspecified random-effects distribution in generalized linear mixed models: perception and problems

Ariel Alonso[1], Saskia Litière[1] and Geert Molenberghs[1]

[1] Center for Statistics, Hasselt University, Belgium

Generalized linear mixed models have become a frequently used tool for the analysis of non-Gaussian longitudinal data. Estimation of the parameters and the inferential procedures are often based on maximum likelihood theory, which assumes that the underlying probability model is correctly specified. One of the assumptions at the core of generalized linear mixed model states that the random effects are normally distributed. Nevertheless, since random effects are unobserved, this assumption can rarely be verified accurately. There is a general perception between data analyst that misspecification of the random effects distribution only have a mild impact on the inferences based on the model.

In the present work we show that the maximum likelihood estimators are inconsistent in the presence of this misspecification. In general, the bias induced in the mean structure parameters is small, as far as the variability of the underlying random-effects distribution is small as well. However, the estimates of this variability are always severely biased. Given that the variance components are the only tool to study the variability of the true distribution, it can be difficult to assess whether problems in the estimation of the mean structure occur as well. Further, we illustrate that the power and the Type I error rate of the commonly used inferential procedures are also severely affected. The situation is aggravated if more than one random effect is included in the model. Finally, we propose to deal with possible misspecification by way of sensitivity analysis, considering several random-effects distributions.