

Marginal Correlation from an Extended Random-effects Model for Repeated and Overdispersed Counts

Clarice G.B. Demétrio¹, Tony Vangeneugden^{2,3}, Geert Molenberghs³ and Geert Verbeke⁴

¹ Departamento de Ciências Exatas, ESALQ, Universidade de São Paulo, Brasil

² Tibotec, Johnson & Johnson, Mechelen, Belgium

³ Center for Statistics, Hasselt University, Diepenbeek, Belgium

⁴ Biostatistical Centre, Katholieke Universiteit Leuven, Leuven, Belgium

Vangeneugden *et al* (2007, submitted) derived approximate correlation functions for longitudinal sequences of general data type, Gaussian and non-Gaussian, based on generalized linear mixed-effects models (GLMM). Their focus was on binary sequences, as well as on a combination of binary and Gaussian sequences. Here, we focus on the specific case of repeated count data, important in two respects. First, we employ the model proposed by Molenberghs, Verbeke, and Demétrio (2007, LIDA, 13, 513-531), which generalizes at the same time the Poisson-normal GLMM and conventional overdispersion models, in particular the negative-binomial model. The model flexibly accommodates data hierarchies, intra-sequence correlation, and overdispersion. Second, means, variances, and joint probabilities can be expressed in closed form, allowing for exact intra-sequence correlation expressions Vangeneugden *et al* (2008, submitted). Next to the general situation, some important special cases such as exchangeable clustered outcomes are considered, producing insightful expressions. The closed-form expressions are contrasted with the generic approximate expressions of Vangeneugden *et al* (2007, submitted). Data from an epileptic-seizures trial are analyzed and correlation functions derived. It is shown that the proposed extension strongly outperforms the classical GLMM.