Modelling of Multivariate Binary Responses

H.A.C.S. Hapuarachchi[1*], S. Samita[2], and W.N.N.K. Perera[3]

[1]*Department of Sports Sciences and Physical Education, Faculty of Applied Sciences, Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka*
[2]*Department of Crop Science, Faculty of Agriculture, University of Peradeniya, Peradeniya, Sri Lanka*
[3]*Department of Statistics, Faculty of Applied Sciences, University of Sri Jayewardenepura, Sri Lanka*
[*]*sampathhac@appsc.sab.ac.lk*

In different fields of study, multivariate binary data are often encountered especially when several different characteristics or attributes are measured on the same unit or from the same person. However, bivariate or multivariate responses observed from the same individual or a unit is possibly expected to be correlated. This study aimed to assess the influence on the estimates of marginal parameters of the joint binary response model when the data is non-clustered and random effect is assumed bridge and normal distributions. Additionally, evaluating the performance of the joint response model compared to the marginal models. Generalized linear mixed models (GLMMs) are used to model correlated binary responses jointly by adding the random effect to the linear predictor. The GLMMs typically adopt a logit or probit link functions embedded with random effects into the linear predictor for correlated binary data. Therefore, the random intercept logistic regression model is used to estimate the model parameters since its ease of interpretation. The 2011 Bangladesh Demographic and Health Survey data were used to study the practical application of the proposed method. A simulation study was employed to illustrate the impact on the joint response model when the use of the bridge and normal distributions for the random effect and to examine the performance of the joint response model compared to the marginal models. R software version 4.0.2 as well as the SAS University Edition®, and 'NLMIXED" procedure were utilized to simulate and to estimate the parameters of the joint response model and marginal models. The simulation study implied that joint modeling of the correlated binary responses provide a better gain in efficiency in the parameter estimates compared to separate models. It is more noticeable for smaller sample sizes. Moreover, this study results revealed that, even after the random effects follow the normal distribution, assuming bridge distribution for the random effect in the joint response model leads to slightly more accurate results. Assuming a bridge or normal distribution for the random effect to accommodate the correlation between two binary responses provided somewhat closer parameter estimates although the standard errors are slightly different.

Keywords: *GLMMs, Joint modeling, Bridge distribution, Correlated binary responses, Random intercept logistic regression*