A Simulation Study of the Properties of the F-test for Type III Fixed Effects in Binary Generalized Linear Mixed Models (GLMMs)

A. A. Sunethra and M. R. Sooriyarachchi

Department of Statistics

The presence of clusters within data is often encountered in medicine, biology and social sciences. With such data, the observations in the data cannot be regarded as independent since the observations within a cluster might have similar patterns than observations across clusters. The method of sandwich variance estimation (SVE) is commonly used with correlated data to adjust the standard errors for correlation. This study evaluates the usefulness of SVE in generalized linear mixed models (GLMMs) which are specialized to model correlated data. The properties of the test for examining a repeated measures scenario with three periods is evaluated using a simulation study for the case of binary responses. A known method for simulating correlated binary data was modified and the correlation of first and second periods was set to 0.4 and 0.35, for first and third periods to 0.22 and 0.19 and for the second and third periods to 0.34 and 0.33 under the null and alternative hypotheses respectively to approximately represent an autoregressive pattern which is often observed in this type of repeated measurements. As properties of comparison, Type I error and the power were compared among the two GLMMs with and without SVE. The simulation consisted of 1000 replicates of sample sizes 20, 50, 100, 250 and 500. While the Type I errors of the GLMMs without SVE were conventional, the type I errors of the GLMMs with SVE were within the 95% probability interval for a 5% error rate for all the sample sizes except 20. For sample size 20, the error rate given by SVE is inflated. This indicates that except for very small sample size (20), the use of SVE in GLMMS gives superior results than not using SVE. In par with the power of the test, both the GLMMs with and without SVE reached 100% power in large sample sizes (n=250,500). But, the GLMM with SVE had a higher power than that of the GLMM without SVE for small sizes (n=20, 50, 100). Therefore, the results of this research suggested that the performance of GLMMs for binary correlated data can be enhanced further by using SVE.