Use of Sandwich Variance Estimation in Generalized Linear Mixed Models: for Binary Repeated Mesures Data

A.A.Sunethra **M.R. Sooriyarachchi**

Department of Statistics, University of Colombo, Colombo 3, Sri Lanka

***Abstract*— Sandwich Variance Estimation (SVE) is a method of estimating variances of miss-specified models and has been popular for analyzing correlated/non-independent data to improve the variance estimation of models fitted for such data. This gained higher popularity when specialized models were not been developed for correlated data whereas with the development of statistical models for correlated data, the use of SVE in such models was at argument among the researches. Generalized Linear Mixed Models (GLMMs) are such models defined for correlated data. But, instances can be found in the literature where GLMMs have shown up model miss-specifications for correlated data. This brought forward the applicability of using SVE in GLMMs since SVE is a method of estimating variances of miss-specified models. Due to the dearth of literature on evaluating the impact of using SVE in GLMMS, this study was undertaken which used both simulated and actual data to evaluate the necessity of using SVE in GLMMs for analyzing Binary correlated data. Type I Error and power of the Type III F-test for fixed effects of the GLMMs fitted for both simulated and actual data showed up better results when GLMMs were fitted with SVE than fitted with the standard method of variance estimation. Further, simulation study demonstrated that at higher level of correlation present in the data, the necessity of using SVE in GLMMs becomes more desirable. Further, it was revealed that classical estimator of SVE perform poorly at small sample sizes (n<= 50) whereas small sample adjusted versions of the SVE showed up better performance at small sample sizes. Thus, this study highlighted that careful use of SVE in GLMMs can help on improving its functionality under model miss-specifications.**

Keywords: Generalized Linear Mixed Models (GLMM), Sandwich Variance Estimation, Binary Responses,Repeated Measures, Properties of the Test

I. INTRODUCTION

The presence of clusters or groups within a dataset is a frequently encountered scenario in data analysis. For example, the data from the members of the same family exhibit more similar characteristics than members from different families. Repeated measures data or longitudinal data pertains to another such situation where the data are clustered within the individuals that are being observed repeatedly over time. The collection of repeated measurements of an individual creates a cluster of observations that are similar to each other while the observations between two individuals are mostly not related to each other. In statistical terminology, such data are being termed as non-independent data or else correlated/clustered data. Thus, statistical techniques deployed for analyzing such data should not assume independence among the observations in the data. This becomes a challenging task since most of the commonly used statistical methods are being constrained on independent data.

Reference [1] pointed out 'statistical inference must control for clustering, as failure to do so can lead to under-estimated standard errors and consequent over-rejection using standard hypothesis tests'. Therefore, the methods that explicitly account for clustering outperforms when confronted with selecting an approach for analyzing clustered/correlated data. Among the handful of statistical methods that explicitly account for correlated data, this research is intended to explore the methods of robust standard errors which is often named as Sandwich Variance Estimation (SVE) within the class of Generalized Linear Mixed Models (GLMMs) for modeling correlated data.

The method of Sandwich Variance Estimation (SVE) was initially proposed by [2] where he discussed the use of maximum likelihood estimation under non-standard conditions and on how to improve the properties of the maximum likelihood estimates under model misspecifications.. SVE can be applied for correlated data in such a way that it adjusts the standard errors of the model parameters to suit the correlated structure in the data. This correction is generalized by [3] for independent heteroskedastic errors where in correlated data, SVEs are used for non-independent heteroskedastic data (errors) which is the case of this study. Basically, this method can improve only the standard errors of the model parameters, but not the parameter estimates of the models and hence improve the p-values of the model parameters which in turn improves model adequacy tests. Lately, new statistical models were developed which allows the presence of clustering in the model definition itself. Therefore, such models are intended to provide both parameter estimates and standard errors that are adjusted for clustering/correlation in accordance with the model fitted . The class of mixed models is one such modeling approach which facilitates modeling correlated data. Generalized Linear Mixed Models (GLMMs) ([4], [5]) are an extension of mixed models which has more flexibility than the general mixed models. Particularly, GLMMs can model correlated data with response variables being distributed in the exponential family.

4th Annual International Conference on
Operations Research and Statistics (ORS 2016)

ISSN 2251-1938
doi: 10.5176/2251-1938_ORS16.12

With the development of such specialized modeling approaches, the necessity and the popularity of sandwich variance estimation for variance estimation of such models became debatable. But few studies in the literature demonstrated intuitive results to exercise a comprehensive evaluation of this phenomenon of using SVE in GLMMs. Reference [6] have highlighted on miss-specifications that are probable with GLMMs mainly due to the errors/disparities that are prone with the definition of random effect used for representing correlation structure in data. Reference [7] have provided insight into usage of Sandwich Variance Estimation in Generalized Linear Mixed Models for analyzing two data sets where they made a comparison between SVE and standard method of variance estimation of GLMMs for the two datasets they considered. Reference [8] emphasized that miss specifications of GLMMs are mainly due to the random effect definition of the models which are difficult to observe and hence difficult to check the assumption of random effects. Reference [8] have used only the classical method of SVE and have claimed that Sandwich correction for the variance estimation have not been able to show up better performance under model miss-specification. They have wrapped up this scenario mentioning that a full discussion of this issue (sandwich correction in GLMMs) overtakes the objectives of their research where they suggest on alternative approaches for analyzing non-Gaussian correlated data such as use of nonparametric distribution for random effect distribution, finite mixture of normal and etc. Reference [7] have highlighted that SVE can be used indirectly as a diagnostic tool for assessing the misspecification of the random effects, but their findings were limited for two isolated data sets being analyzed with using SVE in GLMMs. Thus, this background was innate for a better evaluation of the intended phenomena which gave rise to this study with the objective of evaluating comprehensively the usage of Sandwich Variance Estimation in Generalized Linear Mixed Models for which both a large scale simulation study and an analysis of actual data was conducted.

The designing of the simulation study accommodated various correlation levels and various sample sizes (small, moderate, large) to enable a comprehensive examination of the intended scenario. Moreover, in addition to the classical form of the SVE, small sample adjusted version/s of the SVE were also considered which was not found in the literature where the use of SVE in GLMMs was looked at. Therefore, the findings of this study would provide a complete evaluation of the usage of SVE in GLMMs in terms of the correlation present in the data, in terms of sample size and in terms of the form of the SVE.

The impact resulted by using SVE (classical and small sample adjusted) in GLMMs was assessed by comparing Power and Type I error of the type III F-test for fixed effects in GLMMs.

II. SIMULATION STUDY DESIGN

The correlated data scenario simulated was of repeated measures data with Binary repeated measurements for each individual. The algorithm used was developed according to the algorithms proposed by [9] which was explained in detail by [10].

Three correlated binary variables were simulated for each individual, each taking the values 0 or 1. Let these 3 variables be denoted by Y_1, X_1 and X_2respectively, with distribution of variable Y_1 be given as $P_{Y1} = PR\ (Y_1=0)$. The joint distribution of $Y_1$and X_1 is denoted by $P_{Y1X1} = Pr(Y_1=0, X_1=0)$ which can be fully determined by the marginal and conditional distributions P_{Y1} and $P(X_1=0|\ Y_1=0)=P_{X1|Y1}$ using the theorem of conditional probabilities given by $P_{X1|Y1} = \frac{P_{Y1X1}}{P_{Y1}}$. The rest of the probabilities can be derived using Bayes theorem. As [9] had shown, the correlation between the two binary variables Y_1 and X_1 is given by $r_{Y1X1} = \frac{P_{Y1X1} - P_{Y1}P_{X1}}{\sqrt{P_{Y1}(1-P_{Y1})P_{X1}(1-P_{X1})}}$

.

The data were simulated for two distinct scenarios of having an equal probability of success (probability of getting zero) at the three periods and to have unequal probability of success at the three periods. The data simulated to have equal probability of success at each period is regarded as data under the null hypotheses where the effect of period is similar at each period and the data simulated for having unequal probabilities at each period is regarded as data under the alternative hypothesis of having different effect of period [10].

Rather than depending on a single set of correlations among the three periods, various correlation levels were implemented. To vary the level of correlation in the data, the conditional probabilities between the three periods were varied where the conditional probability of success in the second period given that a success occurred in the first period was varied among three values 0.7, 0.8 and 0.9. The probability of success in the third period given that the first two periods were successes was also varied similarly. These conditional probabilities (.7,.8 and .9) were selected so as to depict low, moderate and high positive correlation between the periods which is usually the case in repeated measures data. At each level of correlation, five different sample sizes (20, 50, 100, 250, 500) were simulated. Hence, the use of SVE in GLMMs was examined at varying levels of correlation in the data (that were representative of repeated measures data structures) and at varying sample sizes enabling a comprehensive examination of the performance of the GLMMs with SVE. The data were simulated both under the null and alternative hypothesis. Under the null hypotheses, the success probability of each of the three periods was taken to be 0.5 which impose the period to be affected similarly for the binary response at each period (i.e. no period effect to the response of interest). Under the alternative hypothesis, the success probabilities of the three periods were made to be different among the three time periods imposing an effect of the period to the response of interest. For each scenario, 1000 datasets were simulated.

Then the simulated data were analyzed using GLMMs with and without SVE. Both the classical SVE approach [3] and Mancl-DeRouen estimator of SVE [11] were examined. The SAS procedure; PROC GLIMMIX was used for fitting GLMMs where the significance of the fixed effects (i.e.

period) are tested using Type III F-tests. The hypotheses associated with the test are as follows.

H_0: Period effects are equal

H_1: At least one period effect is significantly different from the others

The data simulated under the null hypothesis are realizations of the case where the effect of the period is similar at each time period. Thus, the proportion of rejections of the null hypothesis for data simulated under the null hypothesis gives rise to the type I error of the test. In contrast, the data simulated under the alternative hypothesis are based on having different period effects at each time point. Therefore, the proportion of rejections of the null hypothesis for those data sets gives rise to power of the test. Finally, power and the type I error of the Type III F-tests were compared between the three types of GLMMs at each sample size, at each correlation level to inspect the impact of adopting SVE in GLMMs for the repeated measures scenario.

III. THEORY AND METHODOLOGY

Theory and methodology behind this study mainly consist of Sandwich Variance Estimation(SVE), Generalized Linear Mixed Modes (GLMMs) and simulation of Binary repeated measures data. The design of the simulation study was briefly explained in the above section above. Therefore, theory behind SVE and GLMMs is explained in this section.

A. Generalized Linear Mixed Models (GLMMs) [12]

The class of mixed models which consists Linear Mixed Models and Generalized Linear Mixed Models inherent the capability of modeling correlated/clustered data by specifying the linear predictor with an additional component that represents the clusters/groups in the data, whereas in non-mixed models the linear predictor comprises only with the explanatory variables that are regarded as constant or fixed effects. The specification of clusters/groups in the linear predictor is done by including the effect of a cluster as a random effect which is assumed to follow a particular distribution. Mostly, the distribution of the cluster/group effect is assumed to be normally distributed. GLMMs are an extension to the Linear Mixed Models which can accommodate Non-Gaussian data. The type of the GLMM considered in this study has a Binary response and the random effects (i.e clusters) were assumed to be Normally distributed with a mean zero and an unknown variance. The general form of a GLMM can be explained as follows.

$$Y = X\beta + Z\gamma + \varepsilon \qquad (1)$$

Where,

Y is a $N \times 1$column vector of the responses,

X is a $N \times p$ matrix for 'p' fixed effects,

β is a $p \times 1$ vector of fixed effect regression coefficients,

Z is a $N \times j$ design matrix for 'j' random effects,

γ is a $j \times 1$ vector of the random effects,

γ is assumed to follow normal distribution with $\gamma \sim Normal(0, G)$

ε is a $N \times 1$column vector of residuals assumed to be Normally distributed,

By default, GLMMs are fitted using the method of Residual Log Pseudo Likelihood (RSPL) for parameter estimation in PROC GLIMMIX (SAS/STAT(R) 9.2 User's Guide) for which the pseudo response model is considered as follows.

$$P = X\beta + Z\gamma + \varepsilon \qquad (2)$$

The resultant equations for fixed effects parameter estimates and for their variances are given as:

$$\hat{\beta} = (X' v(\hat{\theta})^{-1} X)^{-} X' v(\hat{\theta})^{-1} p \qquad (3)$$

$$V(\hat{\beta}) = (X' v(\hat{\theta})^{-1} X)^{-} \qquad (4)$$

Where $V(\theta)$ denotes the marginal variance in the linear mixed pseudo-model. The following sections are streamlined for deriving the equations for variances of fixed effect parameters of GLMMs when SVE is used as the method of variance estimation.

B. Sandwich Variance Estimation(SVE)

As per the corollary proposed by [2] which claimed that if the expected value of an estimating equation $E(\psi(x, \theta_0))$ has a nonsingular derivative A at θ_0, then for estimating function $T_n(x_1, x_2, \ldots, x_n)$ or simply T_n of θ_0, it can be shown that the asymptotic distribution of$\sqrt{n}(T_n - \theta_0) \sim Normal(0, V_s)$ where $V_s = A^{-1}BA^{-T}$ is the sandwich estimator. Note that B is the covariance matrix of the estimating equation and A is the Hessian matrix for the estimating equation. As explained in [13], though Huber's papers states that the above estimator in terms of maximum likelihood models, the required assumptions of the sandwich estimator allows the applying SVE on any type of estimating equations of the form T $(x, \theta) = 0$, that is for models where the parameter estimation is done by setting the estimating equation to zero, which doesn't require that the estimation equation to be a derivative of a log-likelihood.

When a Generalized Linear Mixed Models is fitted for analyzing a correlated data set, i.e data with random effects, the default method of estimation is the method of Residual Log Pseudo Likelihood (RSPL). The general form of the sandwich variance estimator for fixed effects estimated in GLMMs is:

$$V_s = c \times \hat{\Omega} \left(\sum_{i=1}^{m} A_i \hat{D}_i' \hat{\Sigma}_i^{-1} F_i' e_i e_i' F_i \hat{\Sigma}_i^{-1} \hat{D}_i A_i \right) \hat{\Omega} \qquad (5)$$

where $e_i = y_i - \hat{\mu}_i$, $\Omega = (D' \Sigma^{-1} D)^{-}$ m is the number of independent sampling units(i.e clusters)

But, for a GLMM with RSPL following substitutions are applicable

$$Y \to P, \quad \Sigma \to V(\theta), \quad D \to X, \quad \hat{\mu} \to X\hat{\beta}$$

For the classical sandwich variance estimator, above equation gets simplified to:

$$V_s = \hat{\Omega} \left(\sum_{i=1}^{m} \hat{D}_i' \hat{\Sigma}_i^{-1} e_i e_i' \hat{\Sigma}_i^{-1} \hat{D}_i \right) \hat{\Omega} \qquad (6)$$

Though the classical SVE was found to improve the error rates for large sample sizes it was found to perform poorly for small sample sizes. As [11] highlights, "robust estimator may be biased when the when the number of subjects is small since ordinary residuals are used for estimating …". Therefore, subsequently several other authors have made corrections to the classical SVE to deal with small(finite) sample sizes. Referece [14] have proposed such finite sample corrected SVE within linear regression models fitted with Least Squared estimation method where adjustment is done to eliminate bais of the SVE particluarly at small sample sizes. Subsequently, [11] proposed a similar modified version for SVE within Generalized Estimating Equations models fitted for clustered data. This laid the pathway for a small sample adjusted version applicable in GLMMS which can be obtained by setting $c = 1,\ A_i = I\ and\ F_i = (I - H_i^T)^{-1}$ [15].

IV. Simulation Results

A. *Simulations Under the null hypothesis*

The data simulated under the null hypothesis had an equal probability of success (p=0.5) at each period for each individual. The GLMM fitted consisted of a binary response variable while period was fitted as a factor with 3 levels (fixed effect). Three types of GLMMs i.e with the standard method of variance estimation, classical SVE and small sample SVE, were considered. Type III F-test, tests whether the 'period' effect is significantly different at three periods. As measurements of comparison the Type I Error and Power of the test were considered. Table 1 presents the Type I Errors across the various sample sizes at each level of correlation. The conditional probabilities of success among the three periods were varied from .7, .8 to .9 to vary the correlation between the three periods and resulted correlations are given in the second column of the Table 1. The last three columns presents the Type I Errors associated with the GLMMs fitted respectively with standard variance estimation, classical SVE and small sample adjusted SVE.

It is noteworthy that the Type I Errors of the GLMMs fitted with standard method of variance estimation were all below the 95% lower limit of the confidence interval for a 5% error rate (.036, 064) irrespective of the level of correlation and irrespective to the sample size which can be considered as an model miss-specification of the GLMM fitted for the data. That is, the Type I Errors for GLMM fitted with standard method of variance estimation resulted to be conservative under model miss specification. Moreover, at high correlation levels (p=0.9), the tendency to produce conservative Type I Errors is being increased highly as resulted Type I Errors were very small. In contrast, Type I Errors of GLMMs fitted with classical SVE were maintained within the 95% confidence interval of 5% error rate at sample sizes of 50 and above when the at low and moderate correlation levels (p=.7 and p=.8) while at high correlation level (p=.9) the classical SVE helped to maintain the error rate within the desired range at sample sizes of 100 or above. That is, the improvement to GLMMs that can be achieved by using classical SVE constraints with the level of correlation present in the data and with sample sizes where classical SVE could not provide improved results that the standard method of variance estimation at small samples and this inability get more viable at higher correlation levels. Therefore, adjustments for the classical SVE were considered and the sample form of SVE suggested by [11] was applied where last column of the Table 1 presents the Type I Error observed when GLMMs were fitted with small sample adjusted version of the SVE of [11]. Quite satisfactorily, small sample adjustment of [11] could maintained the Type I Errors within the 95% confidence interval of the 5% error rate at all the sample sizes when the level of correlation was low and moderate (p=.7 and p=.8) while with high correlation (p=.9) the resulted Type I Errors for the sample of size 20 was outside the 95% confidence interval. Therefore, the improvement of the Type I Errors resulted by small sample adjustment for SVE surpass the performance of standard GLMMs. Conservative Type I Errors obtained for standard GLMMs revealed the miss-specifications/errors viable in standard GLMMs for the data considered and the improvement achieved by applying Sandwich correction confirms that SVE can be used as a method of confronting model miss-specifications of GLMMs for analyzing clustered/correlated data.

Table 1: Type I Errors of the Type III F-test

Conditional Probability	Correlations among the periods	Sample size	Type I Error		
			Standard GLMM	GLMM with Classical SVE	GLMM with small sample SVE
0.7	r(y1,x1)=0.4	20	0.015	0.068	0.057
	r(x1,x2)= .34	50	0.022	0.062	0.057
	r(y1,x2)= .22	100	0.014	0.039	0.061
		250	0.022	0.041	0.054
		500	0.018	0.05	0.048
0.8	r(y1,x1)=0.6	20	0.01	0.088	0.055
	r(x1,x2)= .44	50	0.007	0.063	0.047
	r(y1,x2)= .52	100	0.009	0.056	0.059
		250	0.01	0.053	0.057
		500	0.012	0.05	0.035
0.9	r(y1,x1)=0.8	20	0.002	0.1	0.023
	r(x1,x2)= .43	50	0.002	0.08	0.05
	r(y1,x2)= .74	100	0.002	.061	0.055
		250	0.001	0.057	0.06
		500	0	0.059	0.055

B. Simulation U der the Alternative Hypothesis

The data simulated under the alternative hypothesis consists of realizations of the situations where the period effect is not similar at the three time periods. That is, there is an effect of the period on the response of interest. Thus, the proportion of rejections of the null hypothesis of the Type III F-test relates to the Power of the test. The Power of the test at each level of correlation across the varying sample sizes are presented in the Table 2.

When considered Table 2, the power of the test with standard method of variance estimation, can be seen to have reached the maximum level at large sample of sizes 250 and above with low and mild correlation levels in the data (p=.7 and p=.8) while at high correlation levels (p=.9) the power of the Type III F-test failed to achieve the maximum even at large sample sizes which indicates that standard method of variance estimation failed to confront model miss-specifications at small sample sizes and at higher levels of correlations. As per the Power results achieved by using classical SVE, a marginal improvement of the Power can be seen at small sizes whereas at large sample sizes classical SVE also gained the maximum power. The last column of the table presents the power attained by using small sample adjusted SVE into the GLMM which showed that small sample adjustment proposed by small sample adjustment has not been able to provide a significant improvement of the Power at small sample sizes though it improved the Type I Errors at small sample sizes. Reasoning for this will be elaborated at the Discussion section.

Table 2 : Power Results of the Type III F-test

Conditional Probability	Correlations among the periods	Sample size	Power		
			Standard GLMM	GLMM with Classical SVE	GLMM with small sample SVE
0.7	r(y1,x1)=0.35	20	0.13	0.27	0.211
	r(x1,x2)= .33	50	0.43	0.55	0.556
	r(y1,x2)= .19	100	0.84	0.91	0.867
		250	.99	1	1
		500	1	1	1
0.8	r(y1,x1)=0.54	20	0.061	0.24	0.134
	r(x1,x2)= .53	50	0.26	0.46	0.409
	r(y1,x2)= .35	100	0.59	0.74	0.725
		250	0.97	0.99	0.99
		500	1	1	1
0.9	r(y1,x1)=0.76	20	0.005	0.152	0.09
	r(x1,x2)= .72	50	0.034	0.266	0.275
	r(y1,x2) =.36	100	0.148	0.467	0.477
		250	0.556	0.874	0.874
		500	0.94	0.99	0.994

V. An Example based on Real Data

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression, "One of us (R. B. G.) thanks . . ." Instead, try "R. B. G. thanks". Put sponsor acknowledgments in the unnumbered footnote on the first page.

To envisage the impact of using SVE in GLMMs in real time data, a dataset from a CARDIA study conducted in US which evaluated the smoking status among young adults repeatedly for a period of 10 years was downloaded from JASA data archive (http://lib.stat.cmu.edu/jasadata/, retrieved on 8th July 2015).This data was from [16] where Generalized Estimating Equations methodology is being used by them to analyze the data. A complete analysis of the data is out of the scope of this study whereas the analysis will be restricted on to identifying the feasibility of using SVE in GLMMs for analyzing correlated Binary data. For simplicity, the only predictor variable used was the period or the visit on which the smoking status was evaluated and considered only three visits which were scheduled at years 0, 2 and 5 to ally with the simulated data scenario. The impact made by variance estimation methodology was gauged by the comparing the p-values of the type III F-test and by comparing the standard error estimates of the fixed effects of the GLMMs. The type III F-testes for testing whether there is a significant effect of the period to the smoking status of the individuals resulted p-values of 0.2247, 0.0313 and 0.0314 respectively for GLMMs fitted with standard method of variance estimation, classical SVE and small sample adjusted SVE. So, it should be noted that the resulted p-values resulted with SVE (both with classical and small sample adjusted) were significant and quite similar in quantity while in contrast the standard variance estimation produced insignificant p-value. This discrepancy itself can be considered as an evidence for highlighting the impact made by adopting SVE in GLMMs. A similar judgment is been proposed by [17] in linear regression frame work where they have recommended that "data analysts should correct for heteroscedasticity using Heteroscedasticity Consistent Standard Errors(HCCM) whenever there is reason to suspect heteroscedasticity" where HCCM is another term used to name SVE ,particularly in linear regression models. As the cause for classical SVE and small sample adjusted SVE to yield similar p-values, it can be noted that the data set consisted of 5078 young adults which is undoubtedly a large sample and hence an identical performance can be expected from classical SVE and small sample adjusted SVE.

Following table is presented with standard error estimates of the fixed effect parameters under each variance estimation method.

Table 3: Standard Errors Estimated

Parameter		Estimate	Standard errors		
			Standard GLMM	GLMM with Classical SVE	GLMM with Small Sample SVE
Intercept		1.22	0.053	0.045	0.045
Year	0	-0.10	0.059	0.039	0.039
Year	2	-0.05	0.061	0.037	0.039
Year	5	0	-	-	-

Thus, it can be clearly seen that standard error estimates of standard method of variance estimation had resulted standard errors that differ significantly to those of SVEs while the two types of SVEs resulted similar standard errors. The authors have claimed in the literature that SVE can be used as a diagnostic tool for assessing the miss-specifications of GLMMs in such a way that if the adoption of SVE resulted significantly different standard errors than those with model based (standard) standard errors, it can be regarded as a miss-specification of the mixed model to capture the correlation structure that actually exist in the data. In summary, the indication given by this example was an eye-opener to envisage how Sandwich correction can improve GLMMs while serving as a diagnosis tool for identifying model miss-specification which are usually being claimed as unobservable since the effect of the random effects are not measurable.

VI. Discussion

The objective of this study was formulated as a resultant of an in-depth literature review under taken on methods of analyzing correlated data and the study focused on evaluating the feasibility of using Sandwich Variance Estimation in Generalized Linear Mixed Models. The literature highlighted that method of SVE as a method of adjusting the standard errors of model parameter estimates to adjust for the correlation in the data while GLMM is a statistical modeling approach for fitting correlated data. This literature brought forward the feasibility of using SVE in GLMMs for which the related literature was very little. Thus, this study directed at examining the feasibility of using SVE in GLMMs for which a simulation study was identified to be well suitable. The repeated measures data scenario simulated had a Binary response variable with three repeated measurements for three time periods. A Binary GLMM was fitted, keeping all the options at the default settings while only the method of variance estimation was modified according to SVE. A significant difference of the functionality of GLMMs fitted for simulated data was examined with respect to the Type I Error of the type III F-test of the fixed effects while Power of the test did not show up significant improvement with the adoption of SVE. The simulation results demonstrated that Type I Errors of the standard GLMM were conservative irrespective of the level of correlation of the repeated measurements and irrespective of the sample sizes. It should be mentioned here that only three correlation levels were imposed to the data considering the nature of correlation that is viable in repeated measures data. The adoption of classical SVE maintained the Type I Errors within the desired confidence interval for a 5% error at larger sample sizes while the small sample adjusted SVE resulted the best performance of the GLMMs fitted since it maintained Type I Errors within the confidence interval even at small sample size at all the three levels of correlations imposed. Such an improved performance of GLMMs was not achieved with respect to the power of the Type III F-test since the results showed that only a marginal improvement in the power with GLMMs fitted with SVE (classical and small sample) than the standard GLMM. Both standard GLMMs and GLMMs with classical SVE reached maximum power at large sample sizes.

As reasons for the above behavior, it can be added that [17] have also shown that classical SVE provide incorrect inferences when the sample sizes is less than 250 when SVE is used in Linear Regression Models and they highlight that special versions of SVE work well even for samples as small as 25. Similarly the findings of this study also demonstrated a similar scenario where the use of classical SVE in Binary GLMMs performed poorly at samples of 20 and 50 whereas small sample adjusted SVE helped to correct this.

With respect to the power of the type II F-test, standard GLMM and GLMM with classical SVE attained maximum power at large sample sizes while only marginal improvement was achieved by using classical SVE at small sample sizes. A comparison of the power results attained with classical SVE and small sample adjusted SVE, it was observed classical SVE resulted in close better power results compared to that of small sample adjusted SVE. As the small sample adjusted SVE improves the error rates significantly over the classical approach particularly for small samples and there is only marginal differences in power between the two approaches, the small sample adjustment proposed by [11] can be recommended.

The analysis of the actual data set also revealed that the adoption of SVE in GLMMs impacted favorably on standard error estimates of the fixed effect parameters and indicated that SVE can be used even as an diagnostic tool for gauging model miss-specification which are usually not observable since random effect are not measurable as per the definition of GLMMs.

VII. Conclusions and Recommendations

This study evaluated the feasibility of using Sandwich Variance Estimation in Generalized Linear Mixed Models for Binary Repeated Measures Data. The results of the simulation study revealed that adoption of SVE in GLMMs would further improve its functionality. The type of the GLMM fitted in this study dealt with default options for parameter estimation (RSPL), integral approximations (Newton-Raphson with Ridging), Degrees of Freedom Method (Containment) and etc. Therefore, it can be recommended to examine the feasibility of SVE in GLMMs by adjusting above default settings appropriately and see whether the results of the simulations shows up a significant difference or not.

VIII. References

1. Cameron A. C and Miller D. L. (2011), Robust Inference with Clustered Data, in A. Ullah and D. E. Giles eds., *Handbook of Empirical Economics and Finance*, CRC Press, pp.1-28

2. Huber, P. J. (1967). The behavior of maximum likelihood estimates under non-standard conditions. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol.1, pp. 221–233). Berkeley, CA: University of California Press.

3. White, H. (1980), "Heteroskedasticity-Consistent Covariance Matrix Estimator and aDirect Test for Heteroskedasticity", *Econometrica*, 48 : 817-838.

4. Agresti, A. (2002). *Categorical data analysis*, 2nd edition. Hoboken, N.J. Wiley.

5. Molenberghs, G. and Verbeke, G. (2005). *Models for discrete longitudinal data*.Springer, New York.

6. Litière, S., Alonso, A. and Molenberghs, G. (2008), The impact of a misspecified random-effects distribution on the estimation and the performance of inferential procedures in generalized linear mixed models. *Statist. Med*., 27: 3125–3144. doi: 10.1002/sim.3157

7. Chavance M. and Escolano S (2012). Misspecification of the covariance structure in generalized linear mixed models, *Stat Methods Med Re, ,*doi: 10.1177/0962280212462859

8. Litière, S., Alonso, A., & Molenberghs, G. (2007). Type I and Type II Error Under Random-Effects Misspecification in Generalized Linear Mixed Models. *Biometrics*, *63*(4), 1038-1044.

9. Sebestian, K.,Dominik,T., Fredrich, L.(2011), Generating Correlated Ordinal Random Values, *Technical Report no 94*, Department of Statistics, University of Munich.

10. Gawarammana, M. B. M. B. K., &Sooriyarachchi, M. R. (2015). Comparison of Methods for Analyzing Binary Repeated Measures Data: A Simulation Based Study. *Communications in Statistics-Simulation and Computation*, (just-accepted), 00-00.

11. Mancl, L. A., &DeRouen, T. A. (2001). A covariance estimator for GEE with improved small-sample properties. *Biometrics*, 126-134.

12. Brown, H. and Prescott, R. (2014) Generalised linear mixed models, in Applied Mixed Models in Medicine, Third Edition, John Wiley & Sons, Ltd, Chichester, UK. doi: 10.1002/9781118778210.ch3

13. Hardin, J. W. (2003). The sandwich estimate of variance. *Advances in Econometrics*, *17*, 45-74.

14. MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of econometrics*, *29*(3), 305-325.

15. SAS Institute Inc. 2008. SAS/STAT® 9.2 User's Guide: The GLIMMIX Procedure (Book Excerpt). Cary, NC: SAS Institute Inc.

16. Preisser, J. S., Galecki, A. T., Lohman, K. K., & Wagenknecht, L. E. (2000). Analysis of smoking trends with incomplete longitudinal binary responses. *Journal of the American Statistical Association*, *95*(452), 1021-1031.

17. Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. The American Statistician, 54(3), 217-224.