

The Generalized Estimating Equations for the Unknown Correlation Structure of the Data

Md. Soyebur Rahman¹

1

Received: 6 December 2015 Accepted: 3 January 2016 Published: 15 January 2016

Abstract

In many study the data are taken different period of time and the information about them is gathered relating to an event of concern at different time periods. The data are taken different time period are correlated. Regression analysis based on the Generalized Estimating Equation (GEE) is an increasing important method of such data. The Generalized Estimating Equation is an important and widely used approach in such analysis. Since the true correlation is unknown GEE offers to take a working correlation for analysis. In our study we consider four common correlation structure namely, independent, exchangeable, pair wise, autoregressive. In the study the data are taken from the Dhaka stock exchange (DSE) this data are highly correlated. At first we apply different methods of estimating parameter the we apply GEE for estimating the parameters. Finally we get the GEE gives better estimate than any other method.

Index terms— generalized estimating equation, GEE, OLS, GLM.

1 I. Introduction

The term Generalized Estimating Equations indicates that an estimating equation is not the result of a likelihood-based derivation, but that it is obtained by generalizing other estimating equation. Liang and Zeger (1986), Zeger and Liang (1986) and prentice (1988) developed a most recent method of estimating the parameters of the marginal model. They present a class of estimating equations that take the correlation into account to increase the efficiency. This class of estimating equations is called Generalized Estimating Equations (GEE) and the correlation matrix is called working correlation matrix. The name working is used in the sense that it is an approximate correlation matrix of Y 's (the response variable). The estimates of β obtained by GEE are consistent and in addition consistent variance estimates can be obtained under the weak assumption the weighted average of the estimated correlation matrices converge to a fixed matrix.

2 II. Data Analysis

The stock exchange data are highly correlated from one company to other company.

In our study, we have used stock exchange data because, we want to compare the different method of estimation. We have taken six company and their closing price and the general index in 2009 of DSE. In our calculation we have taken general index as an explanatory variable and the different company closing price as a independent variable. We want to estimate the following model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon$ where,

3 III. Summary and Conclusion

Selection of correlation structure is an important issue in Generalized Estimating Equation (GEE). We have earlier stated that there are four correlation structure namely, independent, exchangeable, autoregressive and pair wise. One needs to select under which correlation structure is unknown GEE works well or provides efficient estimate in several situation.

3 III. SUMMARY AND CONCLUSION

We have concentrated on standard error as a measure of accuracy for an estimator of the parameter. So in this case, we look only on the standard error under different correlation structure of different methods of estimation procedure.

In this study we consider stock exchange data, this data are highly correlated.

In the previous situation, we see that the standard error of the parameters of the GEE method is lowest than any other method. From this study we may conclude that from different method of estimation the GEE parameters are gives the efficient estimate and best approach.¹

¹© 2016 Global Journals Inc. (US)

Parameter Estimate		Using the GLM procedure The GLM Procedure		
		Standard Error		t Value
Intercept	-97.97753227	244.4865048		
x1	-27.79733308	7.7704859		
X2	-7.88169809	1.0654488		
X3	2.29819026	0.3979271		
X4	13.56079439	1.5965023		
X5	3.20755189	0.2794746		
X6	1.63484104	0.8332389	Using GEE procedure	The GENM
Analysis of Initial Parameter Estimates				
Standard Wald 95% Confidence Chi-				
Intercept	Calculation of the model (1) in the different methods as follows: Coefficients Standard Error t Sta			
X Vari-				
able 1				
X Vari-				
able 2				
X Vari-				
able 3				
Param-				
eter T				
Inter-				
cept				
x1 X2				
X3 X4				
X5 X6				
Scale				
X Vari-	13.603066	1.598677955	8.508947	2.16E-
able 4				15
X Vari-	3.2163438	0.279808985	11.49478	1.61E-
able 5				24
X Vari-	1.6285946	0.833022998	1.955042	0.051774
able 6				

[Note: y = DSE general index 2009 1 X = Aims first guaranteed mutual fund 2 X =ACI pharmaceuticals Limited 3 X =Glasco Smith-kline Bangladesh Limited 4 X =Beximco Pharmaceuticals Limited 5 X =Al-Arafah Islami Bank Limited 6 X =Bata Shoe Company (Bangladesh) Limited. G Author ? : Jahangirnagar University, Savar, Dhaka. e-mail: soyebur.rahman@gmail.com Author ? : Lecturer, Department of Statistics, Mawlana Bhashani Science and Technology University. Author ? : Assistant Professor, Department of Statistics, Mawlana Bhashani Science and Technology University.]

Figure 1:

-
- 49 [Biometrika] , *Biometrika* 73 p. .
- 50 [Mc Cullagh and Nelder ()] , P Mc Cullagh , J A Nelder . 1989. London: Chapman and Hall. (Generalized Linear
51 Models, 2 nd edition)
- 52 [Neter et al. ()] , J ; W Neter , M H Wasserman , Kutner . 1990.
- 53 [Hardin and Hilbe (ed.) ()] *Generalized Estimating Equations*, J W Hardin , J M Hilbe . U.S.A.: Chapman and
54 Hall/CRC (ed.) 2003.
- 55 [Liao ()] *Interpreting Probability Models Logit, probit, and other Generalized Linear Models*, T Liao , F . 1994.
56 New Delhi: Sage publication. (inc)
- 57 [Liang and Zeger ()] K-Y Liang , S L Zeger . *Longitudinal Data Analysis Using Generalized Linear Models*, 1986.
- 58 [Chang ()] 'Residual analysis of the Generalized Linear Models for Longitudinal Data'. Yue-Cune Chang .
59 *Statistics in medicine* 2000. 19 p. .