34-Comparison of

By Jaka Nugraha

COMPARISON OF LOGIT MODEL AND PROBIT MODEL ON MULTIVARIATE BINARY RESPONSE

JAKA NUGRAHA

Dept. of Statistics, Islamic University of Indonesia, Kampus Terpadu UII, Jl. Kaliurang Km.14, Yogyakarta, Indonesia. email : jk.nugraha@gmail.com

Abstract. On univariate binary response, Logit model is better interpretation compared to Probit model. Logit model and Probit model may be used to analyze same data sets for the same purpose but which model can perform better analysis on multivariate binary data is an interesting topics to be studied. In this study, a comparison of multivariate binary probit and logit models via a simulation study was performed under different correlations between dependent variables. We assume that each of n individual observed T response. Y_{it} is t^{nd} response on i^{nd} individual/subject and each response is binary. Each subject has covariate X_i (individual characteristic) and covariate Z_{ijt} (characteristic of alternative j). Individual response i can be represented by $\mathbf{Y}_i=(Y_{i1},\ldots,Y_{iT})$, Y_{it} is t^{nd} response on i^{nd} individual/subject and each response is binary. In order to simplify, we choose one of individual characteristics and alternative characteristics. We studied effects of correlations using data simulation. General Estimating Equations (GEE) was used to estimate the parameters in this study. Data were generated by using software R.2.8.1 as well as the estimation on the parameters. Based on the result, it can be concluded that estimator in the logit model is equivalent to 1.63 on the probit model. Estimator of the correlation base on Chaganty-Joe is more accurate compared to GEE base on Liang-Zeger.

Key words and Phrases : Random Utility Model, GEE, Simulated maximum likelihood estimator, Newton-Raphson, GHK.

1. Introduction

Investigators often encounter a situation in which plausible statistical models for observed data require an assumption of correlation between successive measurements on the same subjects (longitudinal data) or related subjects (clustered data) enrolled in clinical studies. Statistical models that fail to account for correlation between repeated measures are likely to produce invalid inferences since parameter estimates may not be consistent and standard error estimates may be wrong[1]. Statistical methods that appropriate for analyzing repeated measures include generalized estimating equations (GEE) and multi-level/mixed-linear models[2]. GEE involves specifying a marginal mean model relating the response to the covariates and a plausible correlation structure between responses at different time periods (or within each cluster). Estimated Parameter thus obtained are consistent irrespective of the underlying *true* correlation structure, but may be inefficient when the correlation structure is misspecified[2]. GEE parameter estimates are also sensitive to outliers[2,3]. Summary statistics derived from the likelihood ratio test can be used to check model adequacy in cross-sectional data analyses[1,4,5]. For mixed linear models, the process is often not straightforward due to the complexities involved[6]. Model selection is difficult in GEE due to lack of an absolute goodness-of-fit test to help in choosing the "best" model among several plausible models [4,5,7]. For repeated binary responses, Barnhart and Williamson[5] and Horton *et al.*[4] proposed ad-hoc goodness-of-fit statistics which are extensions of the Hosmer and Lemeshow method for cross-sectional logistic regression models[4,5,8].

Frequently, some dependent variables are observed in each individual. This observation results the multivariate data. Research of multivariate binary response models still gets a little attention, however the applications of multivariate binary response model are mostly extensive. GEE can be implemented on multivariate binary response. Variable Y_{it} with i=1,..n and t=1,..T in panel data (longitudinal) refer to the same variable. In multivariate binary response, Y_{it} refer to the different variable (T variables)[2]. Nugraha *et al.* [9,10] has tested properties of estimator of bivariate logistic regression using MLE and GEE. Both of MLE and GEE are consistent. Logistic model on multivariate binary data using GEE are more efficient compared to the univariate approximation. From simulation data, it was concluded that GEE was better than MLE, specifically GEE able to accommodate the correlations and the GEE's estimator was more precise than MLE.

Furthermore, Nugraha *et al*.[11,12] discussed estimating parameter of Probit model on multivariate binary response using simulated maximum likelihood estimator (SMLE) methods to estimate the parameter based on Geweke-Hajivassiliou-Keane (GHK) simulator. From computational side, simulation method applicable for Probit model is need to be developed to overcome the limitation of GHK method. For this limitation, Nugraha[13] have proposed mixed logit model. From simulation data, he conclude that mixed logit model is better than logit model.

Based on the fact that GEE was acceptable than MLE and is widely available in many statistical software applications, in this study, we compare the probit model and logit model on multivariate binary response using simulation data. In the R.2.8.1 program, the logit and probit models can be obtained by using library(*geepack*) and library(*mprobit*). The *geepack* is the GEE that is based on Liang-Zeger and the *mprobit* is the GEE that is based on Chaganty-Joe. Generating data and estimating paremater using R 2.8.1 software[14].

2. Utility Model

We assume that each of n individual is observed for T response in that Y_{it} is t^{nd} response on i^{nd} individual/subject and each response is binary. Response for i^{nd} individual can be represented by following statement:

$$Y_i = (Y_{i1},....,Y_{iT})$$

that is a vector of 1xT. $Y_{it} = 1$ if i[nd] subject and t[nd] response choose the first alternative and $Y_{it} = 0$ if the subject choose the second alternative. Each subject has covariate of X_i (individual characteristic) and covariate of Z_{ijt} (characteristic of alternative j=0,1). In order to simplify, one of individual characteristic and one of characteristic of alternative were chosen.

Utility of subject of i choose alternative of j on response t is

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt} \quad \text{for } t=1,2,\dots,T \text{ ; } i=1,2,\dots,n \text{ ; } j=0,1 \tag{1}$$

where

$$V_{ijt} = \alpha_{jt} + \beta_{jt} X_i + \gamma_t Z_{ijt}$$

U_{ijt} is utility that it is latent variable and V_{ijt} is named representative utility. In Random Utility Model (RUM), assumption that decision maker (subject) choosing alternative based on maximize utility, so equation (1) can represented in different of utility,

$$U_{it} = V_{it} + \varepsilon_{it} \tag{2}$$

where $V_{it} = (\alpha_{1t}-\alpha_{0t}) + (\beta_{1t}-\beta_{0t})X_i + \gamma_t(Z_{i1t} - Z_{i0t})$ and $\varepsilon_{it} = \varepsilon_{i1t} - \varepsilon_{i0t}$.

Association between Y_{it} and U_{it} is

$$y_{it} = 1 \Leftrightarrow U_{it} > 0 \Leftrightarrow -V_{it} < \varepsilon_{it} \text{ and } y_{it} = 0 \Leftrightarrow U_{it} < 0 \Leftrightarrow -V_{it} > \varepsilon_{it} .$$

Probability of subject i choose ($y_{i1} = 1,\dots, y_{iT} = 1$) is

$$= P(0 < U_{i1},\dots, 0 < U_{iT}) = P(-V_{i1} < \varepsilon_{i1},\dots, -V_{iT} < \varepsilon_{iT})$$

$$P(y_{i1} = 1,\dots, y_{iT} = 1) = \int_{\varepsilon i} I(-V_{it} < \varepsilon_{it}).f(\varepsilon_i)d\varepsilon_i \qquad \forall t \tag{3}$$

where $\varepsilon'_i = (\varepsilon_{i1},\dots, \varepsilon_{iT})$. The value of probability is multiple integral T and depend on parameters α, β, γ distribution ε.

Logit model derived by asumption that ε_{ijt} have extreme value distribution and independence each other. Density of extreme value (Gumbel) is

$$f(\varepsilon_{iit}) = \varepsilon^{-\varepsilon_{ijt}} \varepsilon^{-\varepsilon^{-\varepsilon_{ijt}}} \tag{4}$$

Marginal probability (for some t and i) is

$$P(y_{it} = 1) = \pi_{it} = \frac{\exp(V_{i1t})}{[\exp(V_{i0t}) + \exp(V_{i1t})]} \tag{5}$$

Probit model derived by assuming that vector $\varepsilon_i^{'}$ has a multivariate normal distribution with the mean of null and covarians of Σ. Density of ε_i is

$$f(\varepsilon_i) = \phi(\varepsilon_i) = \frac{1}{(2\pi)^{T/2} \mid \Sigma \mid^{1/2}} \exp[-\frac{1}{2}\varepsilon_i^{'}\Sigma^{-1}\varepsilon_i\ ] \tag{6}$$

Marginal probability (for some t and i) is

$$\pi_{it} = P(y_{it}=1|X_i,Z_i) = P(-V_{it} < \varepsilon_{it}) = \Phi(V_{it}) \quad (7)$$

where $\Phi(V_{it}) = \int_{-\infty}^{V_{it}} \frac{1}{(2\pi\sigma_t^2)^{1/2}} \exp[-\frac{1}{2\sigma_t^2}\varepsilon_{it}^2] d\varepsilon_{it}$

3. Overview of GEE

Marginal models are often fitted using the GEE methodology, whereby the relationship between the response and covariates is modeled separately from the correlation between repeated measurements on the same individual[2]. The correlation between successive measurements is modeled explicitly by assuming a "correlation structure" or "working correlation matrix". The assumption of a correlation structure facilitates the estimation of model parameters[2]. Examples of working correlation matrices include: exchangeable, auto-regressive of order 1 (AR(1)), unstructured, and independent correlation structures[2]. For binary data, correlation is often measured in terms of odds ratios[15]. Details of the correlation structure and response-covariate relationship are included in an expression known as the *quasi-likelihood function*[2], which is iteratively solved to obtain parameter estimates. Estimates obtained from the *quasi-likelihood function* are efficient when the true correlation matrix is closely approximated. In other words, the large-sample variance of the estimator reaches a Cramer-Rao type lower bound[3].

GEE for $\boldsymbol{\theta}$ can present in form

$$G(\boldsymbol{\theta}) = \sum_{i=1}^{n} W_i \Delta_i S_i^{-1} (Y'_i - \pi'_i) = 0 \quad (8)$$

$$W_i = diag\left(\begin{pmatrix} 1 \\ X_i \\ (Z_{i11} - Z_{i01}) \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ X_i \\ (Z_{i1T} - Z_{i0T}) \end{pmatrix}\right) \text{ and}$$

$$\Delta_i = diag\left(\pi_{i1}(1-\pi_{i1}) \quad \ldots \quad \pi_{iT}(1-\pi_{iT})\right)$$

$\mathbf{Y}_i = (Y_{i1},\ldots,Y_{iT})$; $\boldsymbol{\pi}_i = (\pi_{i1},\ldots,\pi_{iT})$; $\mathbf{S}_i = \mathbf{A}_i^{1/2}\mathbf{R}_i\mathbf{A}_i^{1/2}$;

$$\mathbf{A}_i^{1/2} = diag\left(\sqrt{Var(Y_{i1})} \quad \ldots \quad \sqrt{Var(Y_{iT})}\right)$$

$\mathbf{R}_i$ is working correlation matrix $\mathbf{Y}_i$ and $\mathbf{W}_i$ is an observation matrix.

To estimate R_i, Liang and Zeger[16] use vector of empirical corelation r_i with

$$r_{ist} = \frac{(Y_{is} - \pi_{is})(Y_{it} - \pi_{it})}{[\pi_{is}(1-\pi_{is})\pi_{it}(1-\pi_{it})]^{1/2}} \quad (9)$$

r_{ist} is usbias estimator for ρ_{ist} with i = 1,2,..n and s,r=1,2,…T.

In probit model, Chaganti and Joe[17] use

$$Kor(Y_{is},Y_{it})=\frac{\Phi(V_{is},V_{it};\rho_{st})-\Phi(V_{is})\Phi(V_{it}))}{[\Phi(V_{is})(1-\Phi(V_{is}))\Phi(V_{it})(1-\Phi(V_{it}))]^{1/2}} \tag{10}$$

to estimate R_i. If $\rho_{ist}=\rho_{st}$ for all i then

$$\hat{\rho}_{st}=\frac{1}{n}\sum_{i=1}^{n} r_{ist} \tag{11}$$

Equation (8) and (11) can be solved simultaneously for θ and ρ.

4. Generating Simulation Data

We will generate simulation data with T=3. Then, the equations of utility are

$$U_{i0t} = \alpha_{0t} + \beta_{0t}X_i + \gamma_t Z_{i0t} + \varepsilon_{i01} \text{ and } U_{i1t} = \alpha_{1t} + \beta_{1t}X_i + \gamma_t Z_{i1t} + \varepsilon_{i11} \tag{12}$$

for i=1,...,N; j=0,1 and t=1,...,3; ε_{ijt} ~Extreme Value Type I for logit model and ε_{ijt}~N(0,1) for probit model. Equation (12) can be presented in difference of utility $U_{it} = U_{i1t}-U_{i0t}$. On logit model, equations of utility difference are

$$U_{it} = \alpha_t + \beta_t X_i + \gamma_t Z_{it} + \varepsilon_{it} \tag{13}$$

where $Z_{it} = (Z_{i1t} - Z_{i0t})$; $\alpha_t = \alpha_{0t} - \alpha_{1t}$;$\beta_t = \beta_{0t} - \beta_{1t}$.

We generate data on α_1=-1, α_2=1, α_3=-1; β_1 = 0.5, β_1 = -0.5, β_t = 0.5, γ_1=0.3, γ_2=-0.3, γ_3=0.3 and some of correlations ρ =0;0.1;0.2;...; 0.9 using program R.2.8.1. Utility 1 (U_{i1}) was correlated with utility 2 (U_{i2}) and both utility is not correlated with utility 3 (U_{i3}). Data 1 are obtained from ε_{it}~extreme value and Data 2 are obtained from ε_{it}~N(0,1). For each of the data simulation, we estimate parameter using logit model and probit model. GEE-1 are estimator of logit model based on Liang-Zeger. GEE-2 are estimator of probit model based on Liang-Zeger. GEE-3 are estimator of probit model based on Chaganty-Joe.

Those data can be further analyzed by using the program *geepack* and *mprobit*, so the utility must be transformated in the form of:

$$U_i = \sum_{t=1}^{T}\left(\alpha_t D_{it} + \beta_t X_i D_{it} + \gamma_t (Z_{i1t} - Z_{i1t}) D_{it}\right) \tag{14}$$

where D_{ir} is *dummy* variable. D_{ir} =1 for r=t and D_{ir} =0 for r≠t, r=1,2,3. So

If t=1 then $D_{i1} = 1$ and $D_{i2} = D_{i3} = 0$. If t=2 then $D_{i2} = 1$ and $D_{i1} = D_{i3} = 0$. If t=3 then $D_{i3} = 1$ and $D_{i1} = D_{i2} = 0$

$$U_i = \alpha_t + \beta_t X_i + \gamma_t(Z_{i1t} - Z_{i0t})..$$

5. Main Result

Discrete Choice Model (DCM) was prepared based on the of the error distribution. So far there is no method to test the assumption of truth because the utilities(U_i) is also a latent variable can not be known by researchers in value.

5.1. Efek of ε variation to the Estimator

Logit model is constructed based on the assumption that variance ε_{it} is valued with $\pi^2/3$. The value of ε_{it} variance to the estimator is presented bellow. Based on the simulation data, it can be remarked that the value of variance ε_{it} give influence to the estimator. The bigger deviation of variace ε_{it} (from $\pi^2/3$) can affect to the resulted more bias estimator (bigger deviation). Suppose that Var(ε_{it}) = σ^2, where the utility model is

$$U_{it} = V_{it} + \varepsilon_{it}$$

Logit model use the assumption that value of error variance is $\pi^2/3$, so the utility model that will be estimated is

$$\frac{\pi}{\sigma\sqrt{3}}U_{it} = \frac{\pi}{\sigma\sqrt{3}}V_{it} + \frac{\pi}{\sigma\sqrt{3}}\varepsilon_{it}$$

$$\widetilde{U}_{it} = \widetilde{V}_{it} + \widetilde{\varepsilon}_{it} \text{ where } \frac{\pi}{\sigma\sqrt{3}}Var(\widetilde{\varepsilon}_{it}) = \frac{\pi^2}{3}$$

So, the estimator resulted will be deviated for $(1-\frac{\pi}{\sigma\sqrt{3}})\beta$.

5.2. EFFECT OF CORRELATION TO THE ESTIMATOR

In advance simulation, data were generated on n=1000 with 5 replications on each correlations of utility (0 to 0.9). From the simulation (Figure 1 to Figure 10), it can be concluded that :

a. Estimator GEE-1 (in the logit model) is equivalent to 1.63 GEE-2 on the probit model. From Data 2, On the Probit model using the assumption that error (ε_{it}) having normal standard distribution, the value of correlation between utilities give no effect to the estimator properties. Estimator in the logit model is equivalent to 1.64 on the probit model. This is caused by differences in the size of the variansi error on logit model ($\sigma^2=\pi^2/6\cong1.645$) and variansi error in probit models ($\sigma^2=1$) (Table 1, Table 2 and Table 3).
b. Estimator of the correlation using GEE-3 is the most accurate (small bias) compared to GEE-1 and GEE-2.. The estimator is not affected by the faulty of error distribution assumptions. On GEE-1 and GEE-2, the estimator of correlation tends to underestimate.
c. Estimator of the parameter α using GEE-3 is not accurate compared to GEE-1 and GEE-2. But, estimator of the parameter β and γ using the third method the results are relatively the same.
d. Utility 1 (U_{i1}) was correlated with utility 2 (U_{i2}) and both utility is not correlated with utility 3 (U_{i3}). Therefore the value of correlation was only give effect to the parameters whitin U_{i1} and U_{i2}. On both utilities, parameter estimation on the model have big deviation comparable to the correlation value. Therefore the value of correlation was only give effect to

the parameters whitin U_{i1} and U_{i2} on Logit and GEE-1. Estimator of the parameter β and γ using GEE-2 and GEE-3 are more accurate than Logit and GEE-1

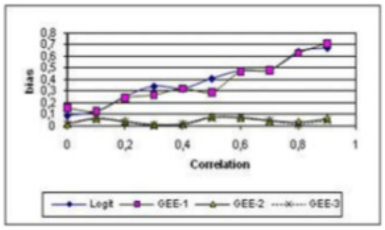

Figure 1. Bias of α_1

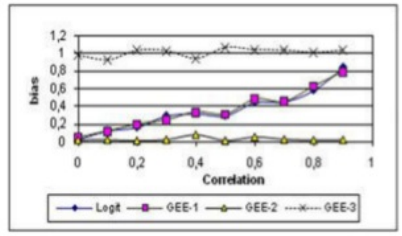

Figure 2. Bias of α_2

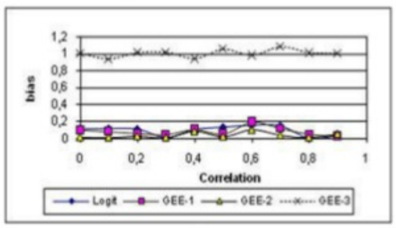

Figure 3. Bias of α_3

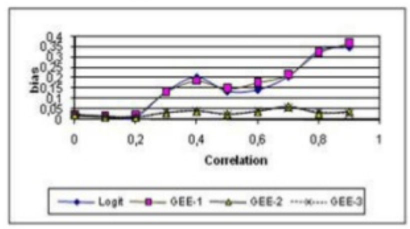

Figure 4. Bias of β_1

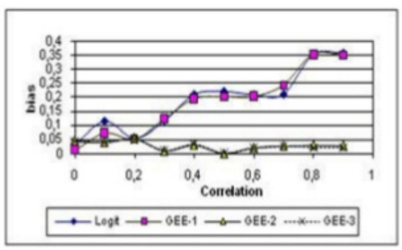

Figure 5. Bias of β_2

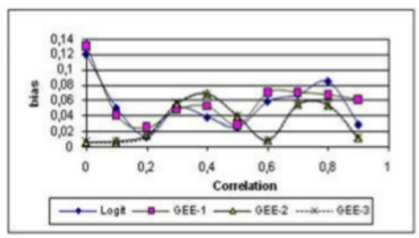

Figure 6. Bias of β_3

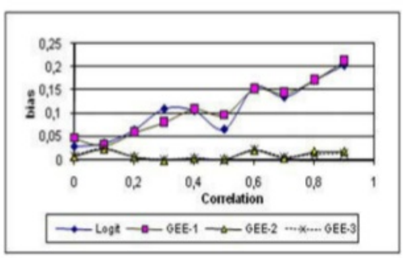

Figure 7. Bias of γ_1

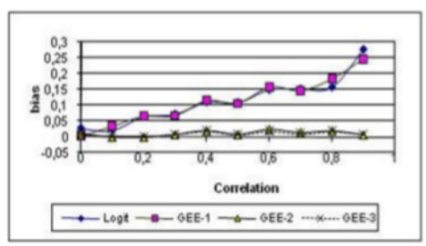

Figure 8. Bias of γ_2

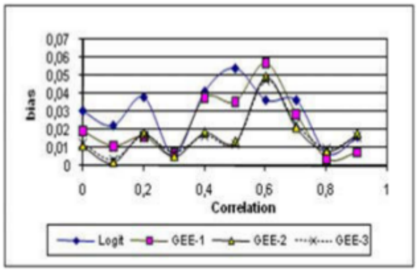

Figure 9. Bias of γ_3

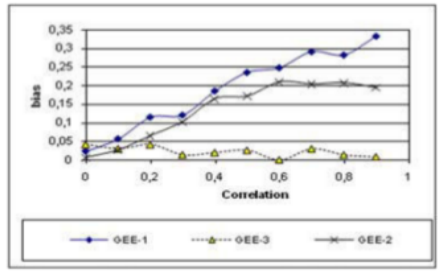

Figure 10. Bias of ρ

6. Concluding Remarks

Based on the results, it can be concluded that estimator in the logit model is equivalent to 1.63 on the probit model. Estimator of the correlation base on Chaganty-Joe is more accurate compared to GEE base on Liang-Zeger. So, we recommended to estimate correlation using GEE base on Chaganty-Joe and then using GEE base on Liang-Zeger to estimate coefisien regression.

References

[1] Dobson A, 2002, *An Introduction to Generalized Linear Models* Florida: Chapman & Hall/CRC.

[2] Diggle PJ, Heagerty P, Liang K, Zeger SL, 2002, A*nalysis of Longitudinal Data,* Second edition. Oxford: Oxford University Press.

[3] Qu, A., Lindsay, B. G. & Li, B., 2000, Improving generalised estimating equations using quadratic inference functions. Biometrika, 87, 823-836.

[4] Horton NJ, Bebchuk JD, Jones CL, Lipsitz SR, Catalano PJ, Zahner GE, Fitzmaurice GM., 1999, Goodness of fit for GEE: An example with mental health service utilization, *Stat Med*, 18(2):213-222.

[5] Barnhart H. X.,; Williamson J. M.., 1998, Goodness-of-fit tests for GEE modeling with binary responses. Biometrics 54(2):720–729.

[6] Schabenberger O., 2004, *Proceedings of the twenty-Ninth Annual SAS Users Group International Conference: May 9–12, 2004; Montreal* Cary, NC: SAS Institute Inc;189-29.

[7] Heagerty PJ, Zeger SL. 2000, Marginalized multilevel models and likelihood inference. Stat Sci. ;15: 1–26.

[8] Hosmer DW, Lemeshow S., 1980, Goodness-of-fit tests for the multiple logistic regression model. Communications in Statistics. A9:1043–1069.

[9] Nugraha J., Haryatmi S., Guritno S, 2009, Estimating Parameter of Logit Model on Multivariate Binary Response Using MLE and GEE. *Jurnal Ilmu Dasar*, Vol. 10. No. 1 , FMIPA Universitas Jember.

[10] Nugraha J., Guritno S., Haryatmi S. 2009, A Comparison of MLE and GEE on Modeling Binary Panel Response, ICoMS 3th IPB, 2008.

[11] Nugraha J., Guritno S., Haryatmi S., 2010, Probit Model on Multivariate Binary Respons Using SMLE, *Jurnal Ilmu Dasar*, FMIPA Univ. Jember,.

[12] Nugraha J., 2011, Simulation Study Of MLE on Multivariate Probit Models, Proceedings of The 6th SEAMS-UGM Conference 2011, FMIPA UGM.

[13] Nugraha J., 2011, Mixed Logit Model on Multivariate Binary Response Using Maximum Likelihood Estimator and Generalized Estimating Equations. Asian Journal of Mathematics and Statistics. Vol 4 (3) : 109-123.

[14] R Development Core Team, 2011, R : A Language and Enviroment Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

[15] Fitzmaurice G.M., Laird N.M., Ratnitzky, A.G.,1993, Regression Models for Discrete Longitudinal Responses. *Statistical Science* Vol. 8 No. 3. 284 – 309.

[16] Liang K.Y., and Zeger S.L.,1986, Longitudinal Data Analysis Using Generalised Linear Models, *Biometrika 73, 13-22.*

[17] Chaganty NR. and Joe H.,2004, Efficiency of Generalized Estimating Equations for Binary Responses, *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, Vol. 66, No. 4, pp. 851-860.

34-Comparison of

ORIGINALITY REPORT

10%

SIMILARITY INDEX

PRIMARY SOURCES

1	itswellington.massey.ac.nz Internet	91 words — 4%
2	www.scialert.net Internet	74 words — 3%
3	digitalcommons.wayne.edu Internet	47 words — 2%
4	www.jourlib.org Internet	44 words — 2%

EXCLUDE QUOTES	ON	EXCLUDE MATCHES	< 15 WORDS
EXCLUDE BIBLIOGRAPHY	ON		

Rektorat Universitas Islam Indonesia
Gedung GBPH Prabuningrat
Jl. Kaliurang Km 14,5 Yogyakarta 55584
T. (0274) 898444 (Hunting)
F. (0274) 898459
E. rektorat@uii.ac.id
W. www.uii.ac.id

Berita Acara Hasil Pengecekan Keaslian Karya Ilmiah
Atas Nama Jaka Nugraha, Dr., S.Si., M.Si. Untuk kenaikan Jabatan
Dari Lektor Kepala (400 AK) ke Guru Besar (850 AK)

Pada tanggal 23 Oktober 2020 telah dilakukan pengecekan *Originality* atau *Similarity* terhadap karya Ilmiah Dosen Tetap Universitas Islam Indonesia:

Nama : Jaka Nugraha, Dr., S.Si., M.Si.

NIP/NIDN/NIK : -/0524027101/956110102

Prodi : Statistika (S1)

Fakultas : Matematika dan Ilmu Pengetahuan Alam

NO	KARYA	REPORT ORIGINALITY	KETERANGAN
1.	Menghasilkan karya ilmiah berupa jurnal internasional terindeks bereputasi dengan impact factor pada Sustainable Chemistry and Pharmacy, Scopus & Scimagojr, SJR=0.62 pada bulan Juni 2018, dengan judul 'Preparation, characterization, and modelling activity of potassium flouride modified hydrotalcite for microwave assisted biodiesel conversion'. ISSN: 2352-5541. Volume 8; issue 43252; pages 63-70; jumlah halaman 8	14%	
2.	Menghasilkan karya ilmiah berupa jurnal internasional terindeks bereputasi dengan impact factor pada Pakistan Journal of Statistics and Operation Research, Scopus & Scimagojr, SJR=0.38 pada bulan September 2019, dengan judul 'Performance Analysis of Mixed Logit Models for Discrete Choice Models'. ISSN: 1816-2711, e-ISSN:2220 5818. Volume 15; issue 3; pages 563-575; jumlah halaman 13	9%	
3.	Menghasilkan karya ilmiah berupa jurnal internasional terindeks bereputasi dengan impact factor pada Journal of Engineering Science and Technology, Scopus & Scimagojr, SJR= 0.24 pada bulan Agustus 2020, dengan judul 'TOBIT MODELING AND CENTRAL COMPOSITE DESIGN FOR OPTIMIZATION OF SPECIFIC SURFACE AREA OF CLAY BY ACTIVATION'. ISSN: 1823-4690. Volume 15; issue 4; pages 2324-2338; jumlah halaman 15	9%	
4.	Menghasilkan karya ilmiah berupa jurnal internasional terindeks bereputasi dengan impact factor pada Applied Clay Science, Scopus & Scimagojr, SJR=1.07 pada bulan Agustus 2020, dengan judul 'Physicochemical characteristics and photocatalytic performance of Tin oxide/montmorillonite nanocomposites at various Sn/montmorillonite molar to mass ratios'. ISSN: 0169-1317. Volume 193; issue (2020) 105671; pages ; jumlah halaman 15	1%	

5.	Menghasilkan karya ilmiah berupa jurnal internasional terindeks pada Asian Journal of Mathematics and Statistics , ZBMATH-DOAJ pada bulan Juni 2011, dengan judul 'Mixed Logit Model on Multivariate Binary Response Using Maximum Likelihood Estimator and Generalized Estimating Equations'. ISSN: 1994-5418. Volume 4; issue 3; pages 113-127.; jumlah halaman 15	0%	
6.	Menghasilkan karya ilmiah berupa jurnal internasional terindeks pada Chemical Engineering Transactions dari International Journal of Chemical and Analytical Science, DOAJ pada bulan Juni 2013, dengan judul 'Evaluation Of Photodegradation Eficiency On Semiconductor Immobilized Clay Photocatalyst By Using Probit Model Approximation'. ISSN: 0976-1209. Volume 4; issue 2; pages 125-130; jumlah halaman 6	6%	
7.	Menghasilkan karya ilmiah berupa jurnal internasional terindeks pada The Turkish Online Journal of Design Art and Communicatio , Copernicus-DOAJ pada bulan September 2018, dengan judul 'Brown's Weighted Exponential Moving Average (B-Wema) with Levenberg-Marquardt Optimization to Forecasting Rate of Return'. ISSN: 2146-5193 . Volume Special Edition; issue ; pages 1744-1749 ; jumlah halaman 6	19%	
8.	Menghasilkan karya ilmiah berupa jurnal internasional terindeks pada International Journal of Engineering and Technology, ProQuest-DOAJ pada bulan September 2018, dengan judul 'Sentiment Analysis on Mobile Banking Application Using Naive Bayes Classifer and Association Methods'. ISSN: 2227-524X. Volume 7; issue 4.15 (Special Issue); pages 244-247; jumlah halaman 4	2%	
9.	Menghasilkan karya ilmiah berupa prosiding seminar internasional terindeks pada 4th International Conference on Process Engineering and Advanced Materials (ICPEAM 2016), Procedia Enginering, Scimagojr dan Scopus, Kuala Lumpur (Malaysia) pada tanggal 15Agustus2016 s.d. 16Agustus2016, dengan judul 'Modeling of Photocatalytic Activity of ZnO/AC By Using Linear Probability Model, Logit and Complementary Log Transformation'. ISBN: 1877-7058. Halaman 1112-1120, jumlah halaman: 9	25%	
10.	Menghasilkan karya ilmiah berupa prosiding seminar internasional terindeks pada International Conference on Chemistry, Chemical Process and Engineering (IC3PE) 2017, Scimagojr dan Scopus, Yogyakarta pada tanggal 15November2016 s.d. 16November2016, dengan judul 'Control of Wastewater Using Multivariate Control Chart'. ISBN: 0094243X, 15517616. Halaman 020126-1 - 020126-11, jumlah halaman: 11	2%	
11.	Menghasilkan karya ilmiah berupa prosiding seminar internasional terindeks pada 2nd International Conference on	0%	

	Chemistry, Chemical Process and Engineering (IC3PE) 2018, Scimagojr dan Scopus, Yogyakarta pada tanggal 14Agustus2018 s.d. 14Agustus2016, dengan judul 'Modelling on human immunodeficiency virus case using Poisson bivariate regression'. ISBN: 0094243X, 15517616. Halaman 020060-1 - 020060-8, jumlah halaman: 8		
12.	Menghasilkan karya ilmiah berupa prosiding seminar internasional terindeks pada 2nd International Conference on Chemistry, Chemical Process and Engineering (IC3PE) 2018, Scimagojr dan Scopus, Yogyakarta pada tanggal 14Agustus2018 s.d. 14Agustus2016, dengan judul 'Pattern analysis on staff of work accident handling using Chi-squared automatic interaction detection and log linear models'. ISBN: 0094243X, 15517616. Halaman 1112-1120, jumlah halaman: 9	0%	
13.	Menghasilkan karya ilmiah berupa prosiding seminar internasional terindeks pada 3rd International Conference on Learning Innovation and Quality Education (ICLIQE 2019), CPCI (part of Clarivate's Web of Science), Surakarta pada tanggal 07September2019 s.d. 07September2019, dengan judul 'The Influence of Religiosity Toward Universitas Islam Indonesia Students Nationalism Using Structural Equation Modeling and Loglinear Model'. ISBN: 2352-5398. Halaman 1192 -1204, jumlah halaman: 13	0%	
14.	Menghasilkan karya ilmiah berupa prosiding seminar internasional terindeks pada 3rd International Conference on Learning Innovation and Quality Education (ICLIQE 2019), CPCI (part of Clarivate's Web of Science), Surakarta pada tanggal 07September2019 s.d. 07September2019, dengan judul 'Designing Android-Based Library Application with Rough Set Method'. ISBN: 2352-5398. Halaman 965 -975, jumlah halaman: 11	8%	
15.	Menghasilkan karya ilmiah berupa prosiding seminar internasional terindeks pada 3rd International Seminar on Chemical Education, Scimagojr dan Scopus, Yogyakarta pada tanggal 17September2019 s.d. 17September2019, dengan judul 'Linear regression model and spatial autoregressive model for modeling high school dropout'. ISBN: 0094243X, 15517616. Halaman 020010-1 - 020010-5, jumlah halaman: 5	3%	
16.	Menghasilkan karya ilmiah berupa prosiding seminar internasional pada 6th SEAMS-UGM Conference 2011, Yogyakarta pada tanggal 12Juli2011 s.d. 15Juli2011, dengan judul 'SIMULATION STUDY OF MLE ON MULTIVARIATE PROBIT MODELS'. ISBN: 978-979-17979-3-1. Halaman 791 - 800, jumlah halaman: 10	9%	
17.	Menghasilkan karya ilmiah berupa prosiding seminar internasional pada 2nd IndoMS International Conference on Mathematics and Its Applications (IICMA 2013),Yogyakarta pada tanggal 06November2013 s.d. 07November2013, dengan judul 'Comparison Of Logit Model And Probit Model On Multivariate Binary Response'. ISBN: 978-602-96426-2-9 . Halaman 294-301, jumlah halaman: 8	10%	

18.	Menghasilkan karya ilmiah berupa prosiding seminar internasional pada International Conference On Research, Implementation, And Education Of Mathematics And Sciences (ICRIEMS) 2014, Yogyakarta pada tanggal 18Mei2014 s.d. 20Mei2014, dengan judul 'Random Effect Model And Generalized Estimating Equations For Binary Panel Response'. ISBN: 978-979-99314-8-1. Halaman M37-M44, jumlah halaman: 8	0%	
19.	Menghasilkan karya ilmiah berupa prosiding seminar internasional pada 3rd International Conference on Sustainable Built Environment (ICSBE) 2014, Yogyakarta pada tanggal 21Oktober2014 s.d. 22Oktober2014, dengan judul 'Energy Consumption and Thermal Comfort favored by the Occupants in the Air Conditioned House'. ISBN: 978-602-98397-4-6. Halaman 76-83, jumlah halaman: 8	0%	
20.	Menghasilkan karya ilmiah berupa prosiding seminar internasional pada The First International Conference on Statistical Methods in Engineering, Science, Economy and Education (SESEE-2015), Yogyakarta pada tanggal 19September2015 s.d. 20September2015, dengan judul 'The Ability The Chi Squares StatisticsTo Rejecting The Null Hypothesis on Contingency Tables 2x2'. ISBN: 978-602-99849-2-7. Halaman 66-71, jumlah halaman: 6	15%	
21.	Menghasilkan karya ilmiah berupa prosiding seminar internasional pada 4rd International Conference on Sustainable Built Environment (ICSBE) 2016, Yogyakarta pada tanggal 12Oktober2016 s.d. 14Oktober2016, dengan judul 'Analysis Of Community Capacity Indicators And Disaster Preparedness Using Structural Equation Modeling'. ISBN: 2541-223X. Halaman 602-610, jumlah halaman: 9	7%	
22.	Menghasilkan karya ilmiah berupa prosiding seminar internasional pada 2nd International Conference on Science, Mathematics, Environment, and Education (ICoSMEE), Surakarta pada tanggal 26Juli2019 s.d. 28Juli2019, dengan judul 'CLUSTERING AND DIVERSIFYING STUDENT SPREADS BASED ON ENVIRONMENTAL EFFECTS ASSISTED BY WEBGIS VISUALIZATION'. ISBN: 978-602-73159-6-9. Halaman 16-26, jumlah halaman: 11	2%	

Pengecekan di atas menggunakan alat Ithenticate dengan meniadakan (exclude) beberapa hal dengan ketentuan sebagai berikut:

1. Meniadakan (exclude) hasil cek kesamaan karya yang kurang dari 2 persen.
2. Meniadakan (exclude) hasil cek kesamaan karya yang disitasi oleh pihak lain.

3. Meniadakan (exclude) hasil cek kesamaan karya yang terindikasi plagiasi kepada karya ilmiah yang bersangkutan.
4. Meniadakan (exclude) hasil cek kesamaan karya yang menunjukan url atau laman karya ilmiah yang bersangkutan.
5. Meniadakan (exclude) hasil cek kesamaan karya yang diupload dalam bentuk yang berbeda (online pribadi) yang terdeteksi merupakan karya sendiri bukan merupakan laman publikasi Jurnal resmi hanya untuk kepentingan sharing (seperti https://www.researchgate.net facebook.com dll) sehingga bukan termasuk auto-plagisasi/self plagiarism.

Berdasarkan hasil pengecekan di atas, maka karya ilmiah tersebut di atas dapat diteruskan usulannya ke Lembaga Layanan Dikti Wilayah V.

Yogyakarta, 23 Oktober 2020
Rektor

UNIVERSITAS ISLAM INDONESIA
REKTOR
YOGYAKARTA

Prof. Fathul Wahid, S.T., M.Sc., Ph.D.