

Mixed Logit-webprint

ORIGINALITY REPORT

0%

SIMILARITY INDEX

PRIMARY SOURCES

EXCLUDE QUOTES

ON

EXCLUDE MATCHES

< 2%

EXCLUDE
BIBLIOGRAPHY

ON

Mixed Logit-webprint

By Jaka Nugraha

WORD COUNT

3663

TIME SUBMITTED

10-AUG-2020 08:29AM

PAPER ID

62016328

INTRODUCTION

Discussion on binary response modeling was reported by [Sarkar and Midi \(2010\)](#), [Zadkarami \(2008\)](#), [Sarkar et al. \(2011\)](#) and the models usually adopted probit model and logit model ([Alpu and Fidan, 2004](#); [Damisa et al., 2007](#)). Generally for **panel data** observations towards same subjects and variables are conducted repeatedly. The model was constructed based on approach of fixed effect, random and dynamic models ([Lechner et al., 2005](#)). Simulation study of an efficient estimation of a mean true effect using a fixed effects model and a random effects model in order to establish appropriate confidence interval estimation have been reported ([Antoine et al., 2007](#)). Method for estimating parameter on **panel data** modeling used was Maximum Likelihood Estimating (MLE), moment method and Generalized Estimating Equation (GEE) ([Diggle et al., 1996](#)). The simplest model is independent model which the assumption taken based on the response on the same and inter subjects is independent. Based on this assumption, combined probability is result of multiplication of its marginal probability. [Liang and Zeger \(1986\)](#) suggest that logistic and probit analysis that are ignoring the correlation and produce consistence estimation parameter but still have close relation becomes inefficient. [Chaganty and Joe \(2004\)](#) have proposed the use of equations of the correlation estimator for probit models in GEE.

In many cases data are multivariate or correlated (e.g., due to repeated observations on a study subject or for subjects within centers) and it is appealing to have a model that maintains a marginal logistic regression interpretation for the individual outcomes. An alternative is to use a marginal analysis that avoids complete specification of the likelihood ([Lipsitz et al., 1990](#)). [Prentice \(1988\)](#) demonstrate modeling strategic to obtain a consistent and normal asymptotic regression coefficient estimator by using GEE. Double regression was not needed in this calculation.

Development of model by controlling unobserved and homogeneous individual characteristic towards replication was reported. However, for heterogeneous individual characteristic, problem would be found is that estimation parameter becomes bias ([Greene, 2005](#)). Double integral was needed in probit model. From several simulations, [Hajivassiliou et al. \(1996\)](#) reported that Geweke-Hajivasilou-Keane (GHK) was the best simulation. [Geweke et al. \(1997\)](#) also find double integral calculation by using Monte Carlo simulation known as unbiased and consistence GHK.

Eventually, several different independent variables are observed simultaneously. This observation produces multivariate data. [Nugraha \(2000\)](#) examined properties of parameter estimators on bivariate logistic regression by using MLE and GEE method. These method produce a consistence estimator. [Nugraha et al. \(2006\)](#) reported that for logistic model on multivariate binary, smaller variance estimator gained by GEE and compared to independent assumption approach. [Nugraha et al. \(2009a\)](#) evaluated the comparison of GEE and MLE method for multivariate binary logistic. It was concluded that GEE model has smaller bias. However, the estimator for correlation parameter is always underestimate. [Akanda et al. \(2005\)](#) proposed an alternative procedure for goodness of fit test based on GEE where, the correlation between two responses was considered.

[Nugraha et al. \(2008\)](#) presented that analytically, both first and second derivation of log-likelihood binary multivariate model can be calculated. Based on its first derivation and second derivation, estimation on the parameter can be executed by Newton-Raphson iteration procedure. [Nugraha et al. \(2009b\)](#) obtained the estimator of binary multivariate probit model by Simulated Maximum Likelihood Estimator (SMLE) based on GHK simulation.

During the last decade, practitioners and researchers are willing to move from the multinomial and nested logit models, that were the standard until only a few years ago, towards these more general models, specifically towards Mixed Logit ([Munizaga and Alvarez-Daziano, 2004](#)). In this present study, correlation effect on mixed logit model towards binary multivariate data is discussed. Estimation of parameter is conducted by MLE and GEE method and comparative study for simulation data is subjected. Both simulation data and estimation of parameter were performed utilizing R.2.10.0. program. (www.R-project.org).

MULTIVARIATE BINARY RESPONSE

It is assumed that Y_{it} is binary response, $Y_{it} = 1$ as the subject i at the response of t choosing the alternative 1 and $Y_{it} = 0$ if the subject of i at the response of t choosing the alternative of 2. Each individual has covariate X_i as individual characteristics i and covariate Z_{ijt} as characteristic of choice/alternative j at the individu of i .

Utility of subject i selecting the alternative of j on response t is:

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

with:

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

By assuming that decision maker select the alternative based on the maximum utility value, the model can be expressed in the form of difference utility:

$$U_{it} = U_{1it} - U_{0it} = V_{it} + \epsilon_{it} \tag{2}$$

with:

$$V_{it} = (V_{11t} - V_{10t}) = (\alpha_{1t} - \alpha_{0t}) + (\beta_{1t} - \beta_{0t}) X_i + \gamma_t (Z_{11t} - Z_{10t}) = \alpha_t + \beta_t X_i + \gamma_t Z_{it}$$

and $\epsilon_{it} = (\epsilon_{i1t}, \dots, \epsilon_{iTt})'$. The probability of subject of i selecting $(y_{i1} = 1, \dots, y_{iT} = 1)$ is:

$$P(y_{i1} = 1, \dots, y_{iT} = 1) = \int_{\epsilon} I(-V_{it} < \epsilon_{it}) \cdot f(\epsilon_i) d\epsilon_{i1} \dots d\epsilon_{iT} \forall_t \tag{3}$$

With $\epsilon_i = (\epsilon_{i1}, \dots, \epsilon_{iT})'$. Value of probability is calculated by multiple integral T depend on the parameter $\theta = (\theta_1', \theta_2', \dots, \theta_T')$ as well as distribution of ϵ (Train, 2003).

GEE ON MODEL LOGIT

GEE on multivariate binary response can be expressed in the following form :

$$G(\theta) = \sum_{i=1}^n W_i \Delta_i S_i^{-1} (Y_i - \pi_i) = 0 \tag{4}$$

W_i is called as observation matrix:

$$W_i = \text{diag} \left(\begin{pmatrix} 1 \\ X_i \\ (Z_{i1} - Z_{i0}) \end{pmatrix}, \dots, \begin{pmatrix} 1 \\ X_i \\ (Z_{iT} - Z_{i0}) \end{pmatrix} \right)$$

and $\Delta_i = \text{diag} (n_{i1} (1 - n_{i1}) \dots n_{iT} (1 - n_{iT}))$ S_i is matrix $T \times T$ and can be expressed as:

$$S_i = A_i^{-1} R_i A_i^{-1}$$

It is defined that $A_i^{-1} = \text{diag} (\sqrt{\text{Var}(Y_{i1})} \dots \sqrt{\text{Var}(Y_{iT})})$ and R_i is "working" matrix with the correlation ρ_{ij} in the size of $T \times T$.

Paired correlation coefficient between response of s and response of r , $\rho_{ist} = \text{corr}(Y_{is}, Y_{ir})$ for $i = 1, 2, \dots, n$ and $s, r = 1, 2, \dots, T$. In order to estimate R_i , empirical correlation, r_{ij} is defined as vector with the size of $T(T-1)/2$ with the elements:

$$r_{sr} = \frac{(Y_{is} - \pi_{is})(Y_{ir} - \pi_{ir})}{[\pi_{is}(1 - \pi_{is})\pi_{ir}(1 - \pi_{ir})]^{1/2}}$$

Empirical correlation r_{ist} that is contained parameter n_{is} and n_{it} is unbiased estimator for ρ_{ist} (Liang and Zeger, 1986). As it is assumed that $\rho_{ist} = \rho_{st}$ for all of i , the estimator is:

(5)

Equation 4 and 5 can be solved at the same time to get the parameter estimator θ and ρ .

Theorem 1: If $\hat{\theta}_G$ is GEE estimator that is solution of Eq. 4, so $\lim_{n \rightarrow \infty} \sqrt{n}(\hat{\theta}_G - \theta)$ is normal multivariate distributed:

$$\lim_{n \rightarrow \infty} \sqrt{n}(\hat{\theta}_G - \theta) \sim N(0; n \cdot \text{Var}(\hat{\theta}_G))$$

where,

$$\text{Var}(\hat{\theta}_G) = \lim_{n \rightarrow \infty} n \left(\sum_{i=1}^n \left(\frac{\partial \pi_i}{\partial \theta} \right) S_i^{-1} \left(\frac{\partial \pi_i}{\partial \theta} \right)' \right)^{-1} \left(\sum_{i=1}^n \left(\frac{\partial \pi_i}{\partial \theta} \right) \Sigma_i^{-1} [\text{Var}(Y_i)] S_i^{-1} \left(\frac{\partial \pi_i}{\partial \theta} \right)' \right) \left(\sum_{i=1}^n \left(\frac{\partial \pi_i}{\partial \theta} \right) S_i^{-1} \left(\frac{\partial \pi_i}{\partial \theta} \right)' \right)^{-1}$$

MIXED LOGIT MODEL

Mixed logit model is combination of extreme value distribution and normal distribution assumption. Based on equation of utility difference (2), the individual effect δ_{it} is added:

$$U_{it} = V_{it} + \delta_{it} + \epsilon_{it} \tag{6}$$

ϵ_{it} has extreme value type I distribution. δ_{it} is i th individual effect of t th response and is independent respect to ϵ_{it} . It is assumed that $\delta_i = (\delta_{i1}, \dots, \delta_{iT})$ has multivariate normal, $\delta_{it} \sim N(0, 1)$ and $\delta_i \sim N(0, \Sigma)$. Parameter that will be estimated on this mixed Logit model are the parameter of regression coefficient $(\alpha_t, \beta_t, \gamma_t)$ and Σ . Parameter Σ are estimated by the Cholesky factor, C .

Individual effect δ_i can be can be modified in the following form:

$$\delta_i = C\xi_i; \xi_i = (\xi_{i1}, \dots, \xi_{iT})' \tag{7}$$

where, C is matrix of Cholesky factor of Σ :

$$CC' = \Sigma \text{ and } \xi_i \sim N(0, 1)$$

I is identity matrix with the size of $T \times T$. Generally, for $t = 1, \dots, T$, the utility function can be constructed:

$$U_{it} = V_{it} + K_{it} + \epsilon_{it} \tag{8}$$

where,

$$K_{it} = \sum_{l=1}^i c_{tl} \xi_{il}, \xi_{il} \sim \text{NID}(0, 1)$$

ξ_{il} - extreme value type I, c_{tl} is an element of cholesky matrix. Based on Eq. 8 conditional logit model is derived with expression:

$$g_{it} = P(y_{it} = 1 | K_{it}) = \frac{\exp(V_{it} + \sum_{l=1}^i c_{tl} \xi_{il})}{[1 + \exp(V_{it} + \sum_{l=1}^i c_{tl} \xi_{il})]} \tag{9}$$

Marginal probability (for t and i) in the mixed logit model is:

$$\begin{aligned} P(y_{it} = 1) &= \pi_{it} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} P(y_{it} | c) f(\xi_{i1}, \dots, \xi_{iT}) d\xi_{i1} \dots d\xi_{iT} \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{\exp(V_{it} + \sum_{l=1}^i c_{tl} \xi_{il})}{[1 + \exp(V_{it} + \sum_{l=1}^i c_{tl} \xi_{il})]} \prod_{l=1}^i \varphi(\xi_{il}) d\xi_{i1} \dots d\xi_{iT} \end{aligned} \tag{10}$$

$\varphi(\xi_{it})$ is standard normal density.

The joint conditional probability is:

$$P(Y_{i1} = y_{i1}, \dots, Y_{iT} = y_{iT} | c, \xi_i) = \prod_{t=1}^T (g_{it})^{y_{it}} (1 - g_{it})^{1-y_{it}} \tag{11}$$

where, $c = (c_{11}, c_{12}, c_{13}, c_{22}, c_{23}, c_{33})$ and $\xi_i = (\xi_{i1}, \dots, \xi_{iT})$. The joint probability is:

$$P(Y_{i1} = y_{i1}, \dots, Y_{iT} = y_{iT}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left(\prod_{t=1}^T (g_{it})^{y_{it}} (1 - g_{it})^{1-y_{it}} \right) \varphi(\xi_i) d\xi_{i1} \dots d\xi_{iT} \tag{12}$$

$\varphi(\xi_i)$ is normal standard density.

Theorem 2: If it is known the utility model as represented in Eq. 8 that is fit to regularity condition, so MLE for parameter of $\theta_t = (\alpha_t, \beta_t, \gamma_t)'$ and Σ is the solution of estimator equation:

$$\sum_{i=1}^n \begin{pmatrix} 1 \\ X_i \\ Z_i \end{pmatrix} (y_{it} - \frac{Q_i}{Q}) = 0 \tag{a}$$

$$\frac{1}{Q} \sum_{i=1}^n \prod_{t=1}^T (g_{it})^{y_{it}} (1 - g_{it})^{1-y_{it}} (y_{it} - g_{it}) \xi_{it} \varphi(\xi_i) d\xi_{i1} \dots d\xi_{iT} = 0 \tag{b}$$

where, $l = 1, \dots, T$. with:

$$Q = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \prod_{i=1}^T ((g_{it})^{y_{it}} (1-g_{it})^{1-y_{it}}) \varphi(\xi_i) d\xi_{i1} \dots d\xi_{iT}$$

$$Q_1 = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \prod_{i=1}^T ((g_{it})^{y_{it}} (1-g_{it})^{1-y_{it}} g_{it} \varphi(\xi_i)) d\xi_{i1} \dots d\xi_{iT}$$

$\hat{\theta}_{MLE}$ and $\hat{\Sigma}_{MLE}$ can be obtained by using iteration process of methods BHHH ((Berndt, Hall, Hall, Hausman) or BFGS (Broyden, Fletcher, Goldfarb dan Shanno) (Chong and Zak, 1996).

SIMULATION DATA

Simulation data was simulated by taking case T = 3 and n = 1000. Utility equation is:

$$U_{i0t} = \alpha_{0t} + \beta_{0t} X_i + \gamma_t Z_{i0t} + \epsilon_{i0t} \text{ dan } U_{i1t} = \alpha_{1t} + \beta_{1t} X_i + \gamma_t Z_{i1t} + \epsilon_{i1t} \tag{13}$$

For i = 1, ..., n; t = 1, 2, 3 and j = 0, 1. Equation of utility (13) can be transformed in the difference of utility: $U_{it} = U_{i1t} - U_{i0t}$ and the corresponding model is:

$$U_{i1} = V_{i1} + \epsilon_{i1}; U_{i2} = V_{i2} + \epsilon_{i2}; U_{i3} = V_{i3} + \epsilon_{i3}$$

With $V_{it} = (V_{i1t} - V_{i0t}) = \alpha_t + \beta_t X_i + \gamma_t Z_{it}; Z_{it} = (Z_{i1t} - Z_{i0t}); \beta_t = \beta_{0t} - \beta_{1t}; \alpha_t = (\alpha_{0t} - \alpha_{1t})$. Data was generated on parameter value of $\alpha_t = -1, \beta_t = 0.5$ and $\gamma_t = 0.3$. Structure of correlation/covariances to be examined are $r_{12} = \rho$ and $r_{13} = r_{32} = 0$:

$$\text{Cor}(U_i) = \begin{pmatrix} 1 & \rho & 0 \\ \rho & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Utility for t = 1 is in correlation with the utility for t = 2 and the correlation order $\rho = 0, 0.1, 0.2, \dots, 0.9$. The value of observation variable X_i and Z_{ijt} taken from normal distribution is:

$$X_i \sim N(0, 1); Z_{i0t} \sim N(0, 1); Z_{i1t} \sim N(2, 1)$$

Generated data with the assumption of ϵ_{ijt} is extreme value distributed and $\delta_{it} \sim N(0, \sigma^2)$. Estimation towards parameter was conducted by utilizing GEE and MLE for mixed logit model.

In order to estimate parameter using MLE, double integration was required and for this purpose, Halton sequence can be implemented. Simulation program for Halton can be adopted from library randtoolbox. For optimizing log-likelihood function, optim function by available BFGS method in R.2.10 program was used (R Development Core Team, 2009).

RESULT OF SIMULATION

Figure 1-21 depicted the effect of correlation value to the each estimator bias. From results, it can be concluded that if assumption by using GEE does not appropriate, the utilities that are correlated each other will be bias. Utility 1 (U_{i1}) is in correlation with utility 2 (U_{i2}) and both are not correlated to utility 3 (U_{i3}). Therefore, value of correlation only influences on parameters within U_{i1} and U_{i2} . On both utilities, the estimator bias appreciable and comparable respect to the correlation values (Fig. 1-9).

- Estimation by using MLE is more accurate compared to as in using GEE. At bigger correlation between utilities, estimator on GEE will has higher deviation (Fig. 1-9) and this is caused misassumption. Model Logit model estimated by using GEE is based on the assumption that error has variance of $n^2/3$ while for data, the variance is $(\sigma^2 + n^2/3)$. As the assumption fulfilled, Mixed Logit model can estimate all parameter appropriately (Fig. 11-20)
- Correlation parameter estimator for MLE is better (smallest bias) compared to as or GEE (Fig. 20, 21) more over, Correlation parameter estimator for GEE tends to be underestimate (Fig. 10)
- In low correlation (less than 0.5), MLE is robust respect to assumption deviation and can estimate precisely. Deviation of assumption in GEE method produce in bias estimator

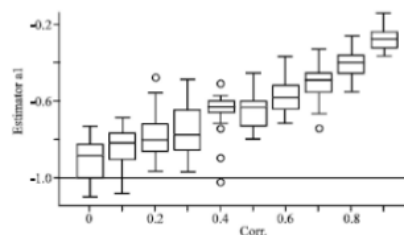


Fig. 1: Box plot α_1 ($t = 1$) in GEE

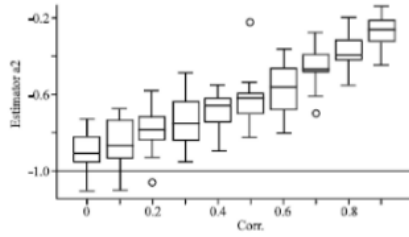


Fig. 2: Box plot α_2 ($t = 2$) in GEE

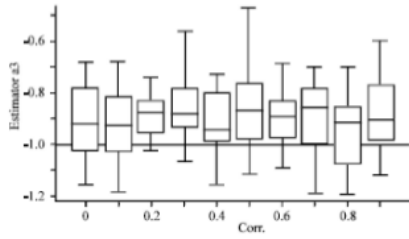


Fig. 3: Box plot α_3 ($t = 3$) in GEE

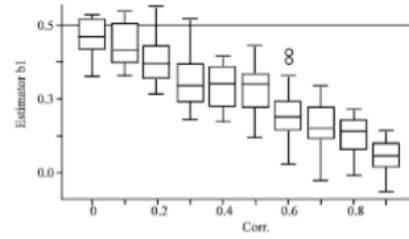


Fig. 4: Box plot β_1 ($t = 1$) in GEE

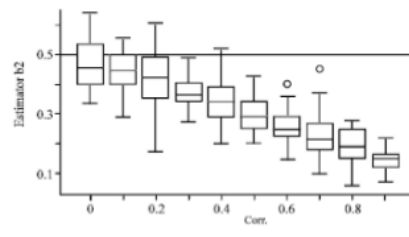


Fig. 5: Box plot β_2 ($t = 2$) in GEE

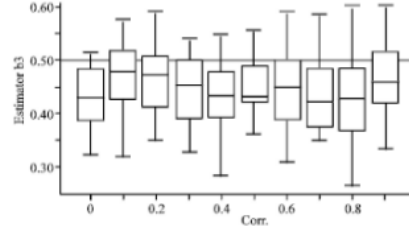


Fig. 6: Box plot β_3 ($t = 3$) in GEE

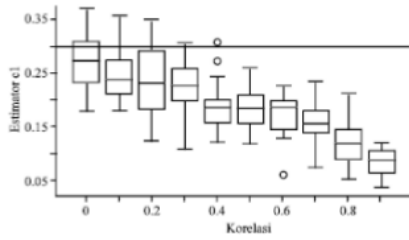


Fig. 7: Box plot γ_1 ($t = 1$) in GEE

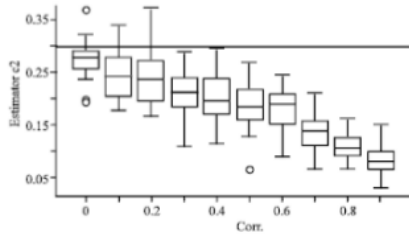


Fig. 8: Box plot γ_2 ($t = 2$) in GEE

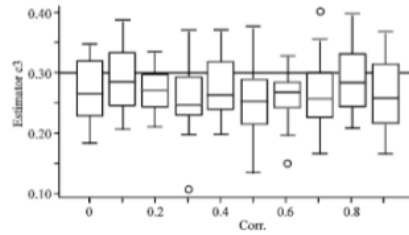


Fig. 9: Box plot γ_3 ($t = 3$) in GEE

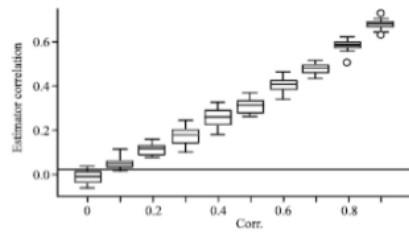


Fig. 10: Box plot ρ in GEE

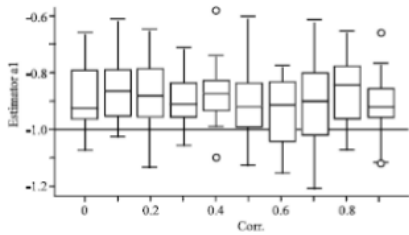


Fig. 11: Box plot α_1 ($t = 1$) in MLE

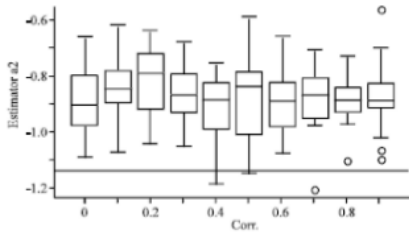


Fig. 12: Box plot a_2 ($t = 2$) in MLE

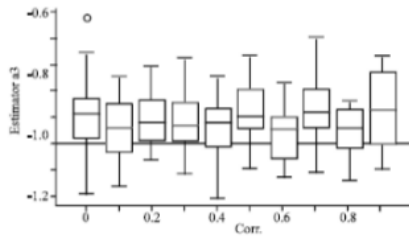


Fig. 13: Box plot a_3 ($t = 3$) in MLE

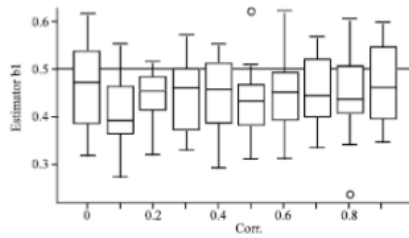


Fig. 14: Box plot β_1 ($t = 1$) in MLE

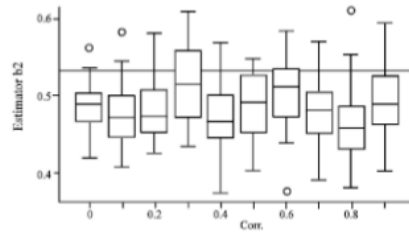


Fig. 15: Box plot β_2 ($t = 2$) in MLE

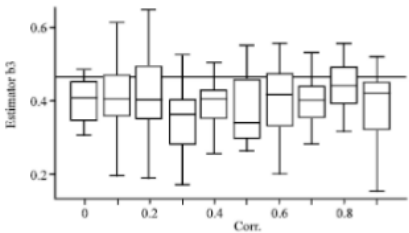


Fig. 16: Box plot β_3 ($t = 3$) in MLE

[Terjemahkan](#)

Alert

ABSTRACT

REFERENCES

PDF

XML

Research Article

Mixed Logit Model on Multivariate Binary Response using Maximum Likelihood Estimator and Generalized Estimating Equations[Jaka Nugraha](#)

ABSTRACT

This study presents discussion on the effects of correlation among response respect to estimator properties in mixed logit model on multivariate binary response. It is assumed that each respondent was observed for T response. Y_{it} is the t^{th} response for the i^{th} 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 that is represented by $Y_i = (Y_{i1}, \dots, Y_{iT})$, Y_{it} is t^{nd} response on i^{th} individual/subject and the response is binary. In order to simplify, one of individual characteristic was and alternative characteristics. We studied effects of correlations using data simulation. Methods of estimations used in this study are Generalized Estimating Equations (GEE) and Maximum Likelihood Estimator (MLE). We generate data and estimate parameters using software R.2.10. From simulation data, we conclude that MLE on mixed logit model is better than GEE. The higher correlation among utility, the higher deviation estimator to parameter.

Services

[Related Articles in ASCI](#)[Similar Articles in this Journal](#)[Search in Google Scholar](#)[View Citation](#)[Report Citation](#)

Science Alert

Like Page 68K likes

Science Alert
on Thursday

#Science

16
Comment
5

How to cite this article:

Jaka Nugraha , 2011. Mixed Logit Model on Multivariate Binary Response using Maximum Likelihood Estimator and Generalized Estimating Equations. *Asian Journal of Mathematics & Statistics*, 4: 113-127.

DOI: [10.3923/ajms.2011.113.127](https://doi.org/10.3923/ajms.2011.113.127)

URL: <https://scialert.net/abstract/?doi=ajms.2011.113.127>

Received: February 27, 2011; **Accepted:** April 14, 2011; **Published:** May 30, 2011

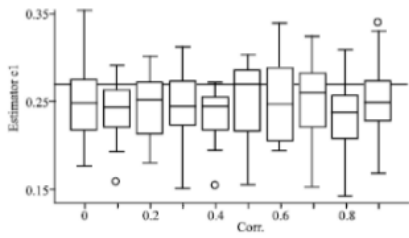


Fig. 17: Box plot γ_1 ($t = 1$) in MLE

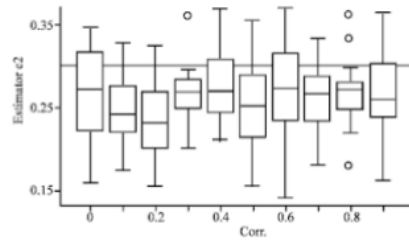


Fig. 18: Box plot γ_2 ($t = 2$) in MLE

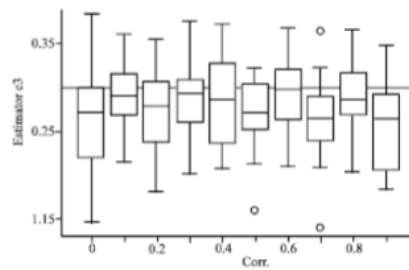


Fig. 19: Box plot γ_3 ($t = 3$) in MLE

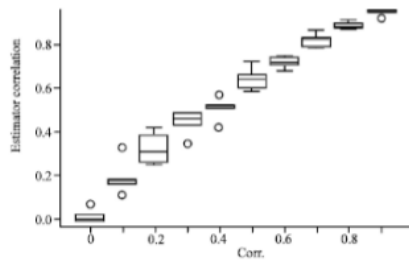


Fig. 20: Box plot ρ in MLE

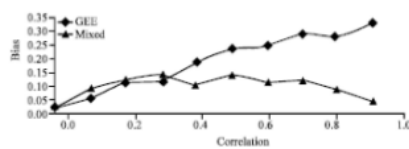


Fig. 21: Graph of a bias ρ in GEE and MLE

CASE STUDY

Data are obtained from library Ecdat in R.2.10.0 which are data Heating. Data are resulted from survey of room heater household in California. From 850 families, there are several variables observed. Y_1 is heater system utilized, that are gc (gas central), gr (gas room), ec (electric central) and er (electric room). X_{1j} is income of each family (income). X_{2j} is age of familyhed (Agehed). X_{3j} is amount of

room in house (rooms) where, $I = 1, \dots, 850$. Furthermore, multivariate approach with two responses was utilized. Y_{i1} is source of energy ($g = 1$) and electricity ($e = 0$). Y_{i2} is application type, central controled ($c = 1$) and room controled ($r = 0$). Characteristic variable for each response is defined in Table 1. From those data, tabulated data between energy source (Y_1) and application type (Y_2) (Table 2).

Based in Chi-squared Pearson's test, obtained value $\chi^2 = 95.8792$ ($p\text{-value} < 2.2e-16$), means that energy source choosen by respondents and application type are not independent.

If there is correlation among responses, parameter estimator will be bias. On this condition, Mixed Logit model is more accurate in estimating parameter compared with as in discrete choice model. For three approximations used; univariate, GEE, mixed logit model, instalation cost and operational cost are the significant variables influence to the Y variable. In contrast, effect of individual charactersitics of X_{1i} , X_{2i} and X_{3i} to the response is not significant.

Equation of marginal probability on independent model and GEE for energy source is as follow:

$$P(Y_{i1} = g) = \frac{\exp(\gamma_{1g}(Z_{1ig} - Z_{1ie}) + \gamma_{2g}(Z_{2ig} - Z_{2ie}))}{\exp(\gamma_{1g}(Z_{1ig} - Z_{1ie}) + \gamma_{2g}(Z_{2ig} - Z_{2ie})) + 1}, P(Y_{i1} = e) = 1 - P(Y_{i1} = g)$$

and the probability equation for application type is:

$$P(Y_{i2} = c) = \frac{\exp(\gamma_{1c}(Z_{1ic} - Z_{1ir}) + \gamma_{2c}(Z_{2ic} - Z_{2ir}))}{\exp(\gamma_{1c}(Z_{1ic} - Z_{1ir}) + \gamma_{2c}(Z_{2ic} - Z_{2ir})) + 1}, P(Y_{i2} = r) = 1 - P(Y_{i2} = c)$$

Estimator value for each parameter is presented in Table 3.

Table 1: Characteristic variabel of election

Response	Installation cost	Operational cost
Y_{i1}		
g	$Z_{1ig} = icS.1$	$Z_{2ig} = ocS.1$
e	$Z_{1ie} = isS.0$	$Z_{2ie} = ocS.0$
Y_{i2}		
c	$Z_{1ic} = icT.1$	$Z_{2ic} = ocT.1$
r	$Z_{1ir} = icT.0$	$Z_{2ir} = ocT.0$

Table 2: Crossed tabulation of energy source and application type

Type	Source	
	g	e
c	573	64
r	129	84

Table 3: Estimator for three models

Parameter	Independent	GEE	Mixed Logit
γ_{1g}	-0.003134464 (6.118290)	-0.002136034 (3.671803)	-0.0042592669 (5.357086)
γ_{1e}	-0.004997236 (182.300896)	-0.004964913 (196.656300)	-0.0034425002 (6.214808)
γ_{1r}	-0.004902223 (56.478941)	-0.004620066 (54.791350)	-0.0041680429 (39.067627)
γ_{1c}	0.006265215 (6.848393)	0.006951781 (9.265083)	-0.0006583833 (2.650140)
ρ	-	0.3131641 (47.84975)	0.6280863 (20.944647)

Values in blanket is statistic Wald

Equation of marginal probability for heater system on Logit model is:

$$P(Y_{i1} = g) = \int_{-\infty}^{\infty} I_{1g}(Z_g, Z_e | \xi_{i1}) \phi(\xi_{i1}) d\xi_{i1} \text{ dan } P(Y_{i1} = e) = 1 - P(Y_{i1} = g)$$

with:

$$I_{1g}(Z_g, Z_e | \xi_{i1}) = \frac{\exp(\gamma_{1g}(Z_{1ig} - Z_{1ie}) + \gamma_{2g}(Z_{2ig} - Z_{2ie}) + \sigma_{\xi_{i1}} \xi_{i1})}{\exp(\gamma_{1g}(Z_{1ig} - Z_{1ie}) + \gamma_{2g}(Z_{2ig} - Z_{2ie}) + \sigma_{\xi_{i1}} \xi_{i1}) + 1}$$

ξ_{i1} has normal distribution with density $\phi(\xi_{i1})$ and:

$$\sigma_{\xi_{i1}}^2 = \frac{\rho \pi^2 / 3}{1 - \rho}$$

Equation of marginal probability for the variable of application type is:

$$P(Y_{1i} = c) = \int_{-\infty}^{\infty} I_{\pi}(Z_i, Z_i | \xi_i) \varphi(\xi_i) d\xi_i \text{ dan } P(Y_{1i} = r) = 1 - P(Y_{1i} = c)$$

With:

$$I_{\pi}(Z_i, Z_i | \xi_i) = \frac{\exp(\gamma_{1T}(Z_{1i} - Z_{1i}) + \gamma_{2T}(Z_{2i} - Z_{2i}) + \sigma_{\xi_i})}{\exp(\gamma_{1T}(Z_{1i} - Z_{1i}) + \gamma_{2T}(Z_{2i} - Z_{2i}) + \sigma_{\xi_i}) + 1}$$

Equation of combined probability on mixed logit model is:

$$P(Y_{1i}, Y_{12}) = \int_{-\infty}^{\infty} (I_{11})^{y_{11}} (1 - I_{11})^{(1-y_{11})} (I_{12})^{y_{12}} (1 - I_{12})^{(1-y_{12})} \varphi(\xi_i) d\xi_i$$

If $Y_{11} = g$, so $y_{11} = 1$ and if $Y_{11} = e$ so $y_{11} = 0$. As $Y_{12} = c$, $y_{12} = 1$ and as $Y_{12} = r$ $y_{12} = 0$.

By using mixed logit model, correlation value between energy source and application source is 0.628 otherwise by using GEE correlation value is 0.313. From simulation result, it is noted that correlation estimator for GEE tend to be underestimate, therefore, estimator correlation on Mixed Logit model is more trusted.

CONCLUSION

Mixed Logit model is based on assumption that error component has extreme value distribution and parameter has normal distribution. Based on data simulation, if there is a violation to the assumption (specifically related to variance) GEE estimator will be bias. Estimator bias will be higher at higher variance value. Estimator bias in GEE can be reduced by MLE method. The correlation of parameter estimator in GEE tends to be underestimate but MLE can be used for estimating correlation between parameters accurately.

REFERENCES

Akanda, M.A.S., M. Khanam and M.A. Islam, 2005. Comparison of goodness-of-fit tests for GEE modeling with binary responses to diabetes mellitus. *J. Applied Sci.*, 5: 1214-1218.

[CrossRef](#) | [Direct Link](#) |

Alpu, O. and H. Fidan, 2004. Sequential probit model for infant mortality modelling in Turkey. *J. Applied Sci.*, 4: 590-595.

[CrossRef](#) | [Direct Link](#) |

Antoine, R., A. Sahai and P. Chami, 2007. Meta analysis of panel data generated by a set of randomized controlled trials. *J. Applied Sci.*, 7: 951-957.

[CrossRef](#) | [Direct Link](#) |

Chaganty, N.R. and H. Joe, 2004. Efficiency of generalized estimating equations for binary responses. *J. Royal Stat. Soc. Ser. B (Stat. Methodol.)*, 66: 851-860.

[Direct Link](#) |

Chong, E.K.P. and S.L. Zak, 1996. *An Introduction to Optimization*. John Wiley and Sons, Inc., New York.

Damisa, M.A., R. Samndi and M. Yohanna, 2007. Women participation in agricultural production: A probit analysis. *J. Applied Sci.*, 7: 412-416.

[CrossRef](#) | [Direct Link](#) |

Diggle, P., K.Y. Liang and S.L. Zeger, 1996. *Analysis of Longitudinal Data*. Oxford University Press, New York.

Geweke, J.F., M.P. Keane and D.E. Runkle, 1997. Statistical inference in the multinomial multiperiod probit model. *J. Econometrics*, 80: 125-165.

[CrossRef](#) |

Greene, W., 2005. *Econometric Analysis*. 5th Edn., Prentice Hall, Englewood Cliffs, New Jersey.

Hajivassiliou, V., D. McFadden and P. Ruud, 1996. Simulation of multivariate normal rectangle probabilities and their derivatives: Theoretical and computational results. *J. Econ.*, 72: 85-134.

[CrossRef](#) | [Direct Link](#) |

Lechner, M., S. Lollivier and T. Magnac, 2005. Parametric binary choice models. Discussion Paper No. 2005-23, University of St. Gallen.

<http://ideas.repec.org/p/usg/dp2005/2005-23.html>.

Liang, K.Y. and S.L. Zeger, 1986. Longitudinal data analysis using generalized linear models. *Biometrika*, 73: 13-22.

[Direct Link](#) |

Lipsitz, S.R., N.M. Laird and D.P. Harrington, 1990. Maximum likelihood regression methods for paired binary data. *Stat. Med.*, 9: 1517-1525.

[PubMed](#) | [Direct Link](#) |

Munizaga, M.A. and R. Alvarez-Daziano, 2004. Testing mixed logit and probit models by simulation. TRB 2005 Annual Meeting CD-ROM.

Nugraha, J., 2000. Estimating parameter on bivariate binary logistic regression. Proceeding of National Seminar, Faculty of Mathematic and Natural Science, UGM

Yogyakarta.

Nugraha, J., S. Guritno and S. Haryatmi, 2006. Logistic regression model for multivariate binary response using generalized estimating equation. Proceeding of National Seminar MIPA, UNY Yogyakarta.

Nugraha, J., S. Guritno and S. Haryatmi, 2008. Likelihood function and its derivatives of probit model on multivariate binary response. J. Kalam, 1: 28-35.

Nugraha, J., S. Guritno and S. Haryatmi, 2009. Comparison of logit model and probit model on multivariate binary response. International Conference on Natural and Material Sciences (NAMES 2009), Faculty of Mathematic and Natural Science -UNLAM Banjarmasin.

Nugraha, J., S. Guritno and S. Haryatmi, 2009. Estimating of parameter of logit model on multivariate binary response using MLE and GEE. J. Ilmu Dasar, 10: 82-92.

Prentice, R.L., 1988. Correlated binary regression with covariates specific to each binary observation. Biometrics, 44: 1033-1048.

[Direct Link](#) |

R Development Core Team, 2009. R: A Language and Environment Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Sarkar, S.K. and H. Midi, 2010. Importance of assessing the model adequacy of binary logistic regression. J. Applied Sci., 10: 479-486.

[CrossRef](#) | [Direct Link](#) |

Sarkar, S.K., H. Midi and S. Rana, 2011. Detection of outliers and influential observations in binary logistic regression: An empirical study. J. Applied Sci., 11: 26-35.

[CrossRef](#) | [Direct Link](#) |

Train, K.E., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, UK., ISBN-10: 0521017157.

Zadkarami, M.R., 2008. Bootstrapping: A nonparametric approach to identify the effect of sparsity of data in the binary regression models. J. Applied Sci., 8: 2991-2997.

[CrossRef](#) | [Direct Link](#) |



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.uui.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 Efficiency 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 Classifier 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 Engineering, 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 Statistics To 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 menunjukkan 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-plagiasi/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



The image shows a circular official stamp of Universitas Islam Indonesia (UII) Yogyakarta. The stamp contains the text "UNIVERSITAS ISLAM INDONESIA" at the top, "REKTOR" in the center, and "YOGYAKARTA" at the bottom. Overlaid on the stamp is a handwritten signature in blue ink that reads "Fathul Wahid".

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