The 6th
SEAMS - GMU 2011
International Conference on Mathematics and Its Applications

ISBN 978-979-17979-3-1

Proceedings of the 6th SEAMS-GMU International Conference on Mathematics and Its Applications

Mathematics and Its Applications in the Development of Sciences and Technology.

ISBN 978-979-17979-3-1

Proceedings of the 6th SEAMS-GMU International Conference on Mathematics and Its Applications

Yogyakarta - Indonesia, 12th - 15th July 2011

MATHEMATICS AND ITS APPLICATIONS IN THE DEVELOPMENT OF SCIENCES AND TECHNOLOGY

Department of Mathematics
Faculty of Mathematics & Natural Sciences
Universitas Gadjah Mada
Sekip Utara Yogyakarta - INDONESIA 55281
Phone : +62 - 274 - 552243 ; 7104933
Fax. : +62 - 274 555131

PROCEEDINGS OF THE 6TH
SOUTHEAST ASIAN MATHEMATICAL SOCIETY
GADJAH MADA UNIVERSITY
INTERNATIONAL CONFERENCE ON MATHEMATICS
AND ITS APPLICATIONS 2011

Yogyakarta, Indonesia, 12th – 15th July 2011

DEPARTMENT OF MATHEMATICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
UNIVERSITAS GADJAH MADA
YOGYAKARTA, INDONESIA
2012

Published by

Department of Mathematics
Faculty of Mathematics and Natural Sciences
Universitas Gadjah Mada
Sekip Utara, Yogyakarta, Indonesia
Telp. +62 (274) 7104933, 552243
Fax. +62 (274) 555131

PROCEEDINGS OF
THE 6TH SOUTHEAST ASIAN MATHEMATICAL SOCIETY-GADJAH MADA UNIVERSITY INTERNATIONAL CONFERENCE ON MATHEMATICS AND ITS APPLICATIONS 2011

ISBN 978-979-17979-3-1

PROCEEDINGS OF THE 6TH SOUTHEAST ASIAN MATHEMATICAL SOCIETY-GADJAH MADA UNIVERSITY INTERNATIONAL CONFERENCE ON MATHEMATICS AND ITS APPLICATIONS 2011

LIST OF REVIEWERS

Abdurakhman
Universitas Gadjah Mada, Indonesia
Achmad Muchlis
Institut Teknologi Bandung, Indonesia
Adhitya Ronnie Effendi
Universitas Gadjah Mada, Indonesia
Agus Buwono
Institut Pertanian Bogor, Indonesia
Agus Maman Abadi
Universitas Negeri Yogyakarta, Indonesia
Agus Yodi Gunawan
Institut Teknologi Bandung, Indonesia

Ari Suparwanto
Universitas Gadjah Mada, Indonesia

Asep. K. Supriatna
Universitas Padjadjaran, Indonesia
Atok Zulijanto
Universitas Gadjah Mada, Indonesia
Azhari SN
Universitas Gadjah Mada, Indonesia
Budi Nurani
Universitas Padjadjaran, Indonesia
Budi Santosa
Institut Teknologi Sepuluh Nopember, Indonesia
Budi Surodjo
Universitas Gadjah Mada, Indonesia
Cecilia Esti Nugraheni
Universitas Katolik Parahyangan, Indonesia
Ch. Rini Indrati
Universitas Gadjah Mada, Indonesia
Chan Basarrudin
Universitas Indonesia
Danardono
Universitas Gadjah Mada, Indonesia
Dedi Rosadi
Universitas Gadjah Mada, Indonesia

Insap Santosa
Universitas Gadjah Mada, Indonesia
Intan Muchtadi-Alamsyah
Institut Teknologi Bandung, Indonesia
Irawati
Institut Teknologi Bandung, Indonesia
Irwan Endrayanto A.
Universitas Gadjah Mada, Indonesia
Jailani
Universitas Negeri Yogyakarta, Indonesia
Janson Naiborhu
Institut Teknologi Bandung, Indonesia
Joko Lianto Buliali
Institut Teknologi Sepuluh Nopember, Indonesia
Khreshna Imaduddin Ahmad S.
Institut Teknologi Bandung, Indonesia
Kiki Ariyanti Sugeng
Universitas Indonesia
Lina Aryati
Universitas Gadjah Mada, Indonesia
M. Farchani Rosyid
Universitas Gadjah Mada, Indonesia
Mardiyana
Universitas Negeri Surakarta, Indonesia
MHD. Reza M. I. Pulungan
Universitas Gadjah Mada, Indonesia
Miswanto
Universitas Airlangga, Indonesia
Netty Hernawati
Universitas Lampung, Indonesia
Noor Akhmad Setiawan
Universitas Gadjah Mada, Indonesia
Nuning Nuraini
Institut Teknologi Bandung, Indonesia
Rieske Hadianti
Institut Teknologi Bandung, Indonesia

PREFACE

It is an honor and great pleasure for the Department of Mathematics – Universitas Gadjah Mada, Yogyakarta – INDONESIA, to be entrusted by the Southeast Asian Mathematical Society (SEAMS) to organize an international conference every four years. Appreciation goes to those who have developed and established this tradition of the successful series of conferences. The SEAMS - Gadjah Mada University (SEAMS-GMU) 2011 International Conference on Mathematics and Its Applications took place in the Faculty of Mathematics and Natural Sciences of Universitas Gadjah Mada on July 12th – 15th, 2011. The conference was the follow up of the successful series of events which have been held in 1989, 1995, 1999, 2003 and 2007.

The conference has achieved its main purposes of promoting the exchange of ideas and presentation of recent development, particularly in the areas of pure, applied, and computational mathematics which are represented in Southeast Asian Countries. The conference has also provided a forum of researchers, developers, and practitioners to exchange ideas and to discuss future direction of research. Moreover, it has enhanced collaboration between researchers from countries in the region and those from outside.

More than 250 participants from over the world attended the conference. They come from USA, Austria, The Netherlands, Australia, Russia, South Africa, Taiwan, Iran, Singapore, The Philippines, Thailand, Malaysia, India, Pakistan, Mongolia, Saudi Arabia, Nigeria, Mexico and Indonesia. During the four days conference, there were 16 plenary lectures and 217 contributed short communication papers. The plenary lectures were delivered by Halina France-Jackson (South Africa), Jawad Y. Abuihlail (Saudi Arabia), Andreas Rauber (Austria), Svetlana Borovkova (The Netherlands), Murk J. Bottema (Australia), Ang Keng Cheng (Singapore), Peter Filzmoser (Austria), Sergey Kryzhevich (Russia), Intan Muchtadi-Alamsyah (Indonesia), Reza Pulungan (Indonesia), Salmah (Indonesia), Yudi Soeharyadi (Indonesia), Subanar (Indonesia) Supama (Indonesia), Asep K. Supriatna (Indonesia) and Indah Emilia Wijayanti (Indonesia). Most of the contributed papers were delivered by mathematicians from Asia.

We would like to sincerely thank all plenary and invited speakers who warmly accepted our invitation to come to the Conference and the paper contributors for their overwhelming response to our call for short presentations. Moreover, we are very grateful for the financial assistance and support that we received from Universitas Gadjah Mada, the Faculty of Mathematics and Natural Sciences, the Department of Mathematics, the Southeast Asian Mathematical Society, and UNESCO.

We would like also to extend our appreciation and deepest gratitude to all invited speakers, all participants, and referees for the wonderful cooperation, the great coordination, and the fascinating efforts. Appreciation and special thanks are addressed to our colleagues and staffs who help in editing process. Finally, we acknowledge and express our thanks to all friends, colleagues, and staffs of the Department of Mathematics UGM for their help and support in the preparation during the conference.

The Editors
October, 2012

CONTENTS

Contributed Papers

Algebra

Analysis

Applied Mathematics

Mathematics Education

Statistics and Finance

Proceedings of "The 6th SEAMS-UGM Conference 2011"
Statistics, pp. 791 - 800.

SIMULATION STUDY OF MLE ON MULTIVARIATE PROBIT MODELS

JAKA NUGRAHA

Abstract. We have studied estimator properties of multivariate binary Probit Models using simulation study. Estimation of parameter is performed by GEE, MLE and SMLE method. Statistics software that was used in the calculation is R.2.8.1. Probit Model can be applied on binary multivariate response by using MLE and GEE estimation method. Based on the simulation data, MSLE estimator is inappropriate to multivariate Probit model. We recommend to combine GEE and MLE. GEE can be used to estimate parameter of regression. MLE can be used to estimate parameter correlations only.

*Keywords and Phrases :*Discrete Choice Model, MLE, GEE, simulation study.

1. INTRODUCTION

Discrete Choice Model (DCM) is a model constructed on the assumption that decision maker faced the choice among a group of alternatives based on their utilities. The alternatives or responses are nominal and one of them is having maximum utility. In this case, the decision maker can be a person, family, company or other unit of decision maker. DCM is correlated with two connected activities; determination of model and calculation of proportion for each choice. The model has been widely discussed are Logit Model and Probit Model. Methods of parameter estimation used are Maximum Likelihood Estimation (MLE) method, Moment method and Generalized Estimating Equation (GEE) method.

Some researchers have studied this similar estimating method on panel binary response. GEE estimator's have invariant properties, consistent and normally asymptotic [1, 2]. In binary panel Probit Model, MLE is the best compare to Solomon-Cox or Gibbs sampler [3, 4]. Probit Model needs multiple integral and it can be solved using Geweeke-Hajivassilou-Keane (GHK)[5,6] Frequently, some dependent variables are observed in each individual. Because the data include simultaneous measurements on many variables, this data is called multivariate data. However the applications of multivariate binary response model are most extensive. Researchon multivariate binary response models still gets a little attention.On binary response, MLE and GEE are consistent estimator [7] and the estimators of regression parameters are not influenced by the correlation [8]. On multivariate binary Logit Model, GEE is more efficient compared to univariate approximation but the estimator of correlation in GEE tends to be underestimate [9]. ProbitModel can be used in multivariate binary response by using some parameter estimation that can be used such as GEE, MLE and MSLE based on GHK simulation [10,11]

Based on the development of binary response model that also supported by computational field, we studied properties of estimator using simulated study in multivariate binary response data. Modeling of multivariate binary response utilized used Probit model and estimation of parameter is performed by GEE, MLE and SMLE method.

2. DCM ON THE 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 individual of i.

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

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

with

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

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

$$\boldsymbol{U_{it}} = \boldsymbol{U_{i1t}} - \boldsymbol{U_{i0t}} = \boldsymbol{V_{it}} + \boldsymbol{\varepsilon_{it}}$$

with

$$V_{it} = (V_{i1t} - V_{i0t}) = (\alpha_{1t}-\alpha_{0t}) + (\beta_{1t}-\beta_{0t})X_i + \gamma_t(Z_{i1t}-Z_{i0t}) = \alpha_t + \beta_t X_i + \gamma_t Z_{it}$$

and $\varepsilon_{it} = (\varepsilon_{i1t} - \varepsilon_{i0t})$. The probability of subject of i selecting ($y_{i1} = 1,...,y_{iT} = 1$) is

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

With $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1},...,\varepsilon_{iT})'$. Value of probability is calculated by multiple integral T depend on the parameter $\boldsymbol{\theta} = (\boldsymbol{\theta}_1',\boldsymbol{\theta}_2',...,\boldsymbol{\theta}_T')$ as well as distribution of $\boldsymbol{\varepsilon}$.

2.1 MLE on Multivariate Binary Probit Model. Vector $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1},...,\varepsilon_{iT})'$ is normally distributed with mean value of zero, covariance matrix $\boldsymbol{\Sigma}$ and each of ε_{it} is normally standard distributed.

$$\varepsilon_i^{'} \sim N(0, \Sigma) \text{ and } \varepsilon_{it} \sim N(0,1) \text{ for } t=1,...,T.$$

It is assume that $\boldsymbol{R}$ is correlation matrix of $\boldsymbol{\varepsilon}_i$, so $\boldsymbol{\Sigma} = \mathbf{Q}_i\mathbf{R}\mathbf{Q}_i$ with $\mathbf{Q}_i$ is diagonal matrix with its diagonal element ($2y_{it}$-1).

$$\Sigma = \begin{pmatrix} 1 & \sigma_{12} & \dots & \sigma_{1T} \\ \sigma_{12} & 1 & \dots & \sigma_{2T} \\ \dots & \dots & \dots & \dots \\ \sigma_{1T} & \sigma_{2T} & \dots & 1 \end{pmatrix} R = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1T} \\ \rho_{12} & 1 & \dots & \rho_{2T} \\ \dots & \dots & \dots & \dots \\ \rho_{1T} & \rho_{2T} & \dots & 1 \end{pmatrix}$$

$$Q_i = \begin{pmatrix} (2y_{i1}-1) & 0 & \dots & 0 \\ 0 & (2y_{i2}-1) & \dots & 0 \\ \dots & \dots & \dots & \dots. \\ 0 & 0 & \dots & (2y_{iT}-1) \end{pmatrix}$$

y_{it} = 0 at the respondent of i choosing first alternative and y_{it} = 1 if respondent of i choosing second alternative. So, $\sigma_{tt'} = (2y_{it}-1)(2y_{is}-1)\rho_{ts}$. To simplify the notation, take $V_{it} = \beta_t X_{it}$ (β_t is identified parameter and the parameters of α and γ are not included).

$$\Phi_T(w_i;0;\Sigma) = \int_{-\infty}^{w_{iT}} \dots \int_{-\infty}^{w_{il}} \dots \int_{-\infty}^{w_{i1}} \phi_T(\varepsilon_i;\theta;\Sigma)d\varepsilon_i = \int_{D(Y_i)} \phi_T(\varepsilon_i;\theta;\Sigma)d\varepsilon_i$$

With $w_{it} = (2y_{it}-1)x_{it}\beta_t$ and $D(\boldsymbol{Y}_i) = [-\infty, w_{i1}]\dots[-\infty, w_{il}]\dots[-\infty, w_{iT}]$. The $\phi_T(.)$is multivariate normal density of T response. Its log likelihood equation is

$$LL(\theta;\Sigma) = \sum_{i=1}^{n} \ln \Phi_T(w_i;0;\Sigma) \qquad (1)$$

or also can be represented as

$$LL(\theta;R) = \sum_{i=1}^{n} \ln \Phi_T(w_i;0;R)$$

Estimation of $\boldsymbol{\theta}$ and Σ(or $\boldsymbol{R}$) using MLE method can be derived from likelihood function of equation (1). Define

$$\Sigma^l = \left(\begin{array}{ccc|c} 1 & \dots & \sigma_{1T} & \sigma_{1l} \\ & \dots & & \dots \\ \sigma_{1T} & \dots & 1 & \sigma_{Tl} \\ \hline \sigma_{1l} & \dots & \sigma_{Tl} & 1 \end{array}\right) = \begin{pmatrix} \Sigma_{11}^l & \Sigma_{12}^l \\ \Sigma_{21}^l & 1 \end{pmatrix}$$

$$\Sigma^{kl} = \left(\begin{array}{ccc|cc} 1 & \ldots & \sigma_{1T} & \sigma_{1k} & \sigma_{1l} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \sigma_{1L} & \ldots & 1 & \sigma_{Tk} & \sigma_{Tl} \\ \hline \sigma_{1k} & .. & \sigma_{Tk} & 1 & \sigma_{kl} \\ \sigma_{1l} & \ldots & \sigma_{Tl} & \sigma_{kl} & 1 \end{array}\right) = \begin{pmatrix} \Sigma_{11}^{kl} & \Sigma_{12}^{kl} \\ \Sigma_{21}^{kl} & \Sigma_{22}^{kl} \end{pmatrix}$$

First derivative of log likelihood function (1) with respect to the parameter β is

$$\frac{\partial LL(\theta;\Sigma)}{\partial \beta_{1l}} = \sum_{i=1}^{n} \frac{\phi(w_{il};0;1)\Phi_{T-1}(w_{i,-l};M^l;S^l)(2y_{il}-1)x_{il}}{\Phi_T(w_i;0;\Sigma)} \tag{2}$$

with $M^l = \Sigma_{12}^l w_{il}$; $S^l = \Sigma_{11}^l - \Sigma_{12}^l \Sigma_{21}^l$ and $\mathbf{w}_{i,-l} = (w_{i1},..,w_{i(l-1)},w_{i(l+1)},...,w_{iT})$.

First derivative of log-likelihood function (1) with respect to parameter ρ is

$$\frac{\partial LL(\beta,R\,|\,x)}{\partial \rho_{kl}} = \sum_{i=1}^{N} \frac{\phi_2(w_{ik},w_{il};0,\Sigma_{22}^{kl})\Phi_{T-2}(w_{i,-kl},M_i^{kl};S^{kl})}{\Phi(w_i;\Sigma)}(2y_{ik}-1)(2y_{il}-1) \tag{3}$$

with $M_i^{kl} = \Sigma_{12}^{kl}(\Sigma_{22}^{kl})^{-1}(w_{il},w_{ik})'$; $S^{kl} = \Sigma_{11}^{kl} - \Sigma_{12}^{kl}(\Sigma_{22}^{kl})^{-1}\Sigma_{21}^{kl}$
and $\mathbf{w}_{i,-kl} = (w_1,...,w_{k-1},w_{k+1},...,w_{l-1},w_{l+1},...,w_T)$

Second derivative of log-likelihood functions (1) with respect to parameter β and ρ are

$$\frac{\partial^2 LL(\theta;\Sigma)}{\partial \beta_{1l}^2} = \sum_{i=1}^{n} \frac{\phi(w_{it};0;1)\Phi_{T-1}(w_{i,-l};M^l;\Sigma^l)[(2y_{il}-1)x_{il}]^2}{\Phi_T(w_i;0;\Sigma)}. \left(1-\phi(w_{it};0;1)\Phi_{T-1}(w_{i,-l};M^l;\Sigma^l)\right) \tag{4}$$

$$\frac{\partial^2 LL(\theta;\Sigma)}{\partial \beta_{1k}\partial \beta_{1l}} = \sum_{i=1}^{n} \frac{\phi(w_{ik};0;1)\phi(w_{il};0;1)(2y_{il}-1)(2y_{ik}-1)x_{ik}x_{il}}{\Phi_T(w_i;0;\Sigma)}. \left((\Phi_{T-2}(w_{i,-kl};M^{kl};\Sigma^{kl}) - 2\Phi_{T-1}(w_{i,-k};M^k;\Sigma^k)\Phi_{T-1}(w_{i,-l};M^l;\Sigma^l)\right) \tag{5}$$

with $M^{kl} = \Sigma_{12}^{kl} w_{ik}$; $\Sigma^{kl} = \Sigma_{11}^{kl} - \Sigma_{12}^{kl}\Sigma_{21}^{kl}$
and $\mathbf{w}_{i,-kl} = (w_{i1},.., w_{i(k-1)},w_{i(k+1)},w_{i(l-)},w_{i(l+1)}, ...,w_{iT})$

$$\frac{\partial^2 LL(\beta,R\,|\,x)}{\partial \rho_{kl}^2} = \sum_{i=1}^{n} \frac{\Phi_{T-2}(w_{i,-kl},M_i^{kl};S^{kl})}{[\Phi_T(w_i;\Sigma)]^2}(2y_{ik}-1)(2y_{il}-1) \left(\Phi_T(w_i;\Sigma)A1 - \left(\phi_2(w_{ik},w_{il};0,\Sigma_{22}^{kl})\right)^2 \Phi_{T-2}(\varepsilon_{i,-kl};M^{kl};S^{kl})(2y_{ik}-1)(2y_{il}-1)\right) \tag{6}$$

$$\text{with } A1 = \frac{\phi_2(w_{il}, w_{ik}; \rho_{kl})}{2\pi\sqrt{(1-\rho_{kl}^2)}}\left(-(w_{il}w_{ik}\pi)\sqrt{(1-\rho_{kl}^2)}\ln(1-\rho_{kl}^2)+4\pi\rho_{kl}\right)$$

$$\frac{\partial^2 LL(\beta,R\mid x)}{\partial\beta_l\partial\rho_{kl}} = \sum_{i=1}^{n}\frac{(2y_{ik}-1)(2y_{il}-1)\Phi_{T-2}(w_{i,-kl},M_i^{kl};S^{kl})\phi(w_{il};0;1)\phi(w_{ik};m^k;1)(2y_{il}-1)x_{il}}{\Phi(w_i;\Sigma)} -$$
$$\sum_{i=1}^{n}\frac{(2y_{ik}-1)(2y_{il}-1)^2 x_{il}\Phi_2(w_{ik},w_{il};0,\Sigma_{22}^{kl})\phi_{T-2}(w_{i,-kl},M_i^{kl};S^{kl})\phi(w_{il};0;1).\Phi_{T-1}(w_{i,-l};M^l;\Sigma^l)}{[\Phi(w_i;\Sigma)]^2} \quad (7)$$

with $m^k = \sigma_{kl}w_{il}$ and $w_{il} = (2y_{il}-1)x_{il}\beta_{1l}$.

MLE can be solved by Newton-Raphson iteration method based on first and second derivation of equation (2) to (7). Estimation by using MLE also needs multiple integral calculation T. Calculation value of $P(\boldsymbol{Y}_i)$ over Gaussian Square calculation can not be done for T more than four [12]. $P(\boldsymbol{Y}_i)$ can be solved only by simulation. GHK is found to be the most efficient simulation method and the unbiased estimator [13]. Therefore it is also called as maximum simulated likelihood estimator (MSLE) method. For getting MSLE in similar properties with MLE, high simulations are needed [14].

2.2 MSLE on Multivariate Binary ProbitModel. Probit Model is based on the assumption that vector $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1},...,\varepsilon_{iT})'$ on equation (5) has normal multivariate distribution with the mean null and the covariance matrix $\boldsymbol{\Sigma}$. Marginal probability (for t and i) is

$$\pi_{it} = P(y_{it}=1|X_i,Z_i) = P(-V_{it}<\varepsilon_{it}) = 1-\Phi(-V_{it}) \quad (8)$$

where

$$\Phi(-V_{it}) = \int_{-\infty}^{-V_{it}}\frac{1}{(2\pi\sigma_t^2)^{1/2}}\exp[-\frac{1}{2\sigma_t^2}\varepsilon_{it}^2]d\varepsilon_{it}$$
$$P(Y_{it}=y_{it}) = \pi_{it}^{y_{it}}(1-\pi_{it})^{1-y_{it}} \text{ for } y_{it}=0,1.$$

From the symmetric properties of normal distribution, the equation (8) can be expressed as

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

Marginal probability also can be represented by

$$P(Y_{it}=y_{it}) = \Phi[(2y_{it}-1)V_{it}]$$

and the combined probability is

$$P(Y_{i1}=y_{i1},...,Y_{iT}=y_{iT}) = P\left(\varepsilon_{i1}<(2y_{i1}-1)V_{i1},...,\varepsilon_{iT}<(2y_{iT}-1)V_{iT}\right) = \Phi(\boldsymbol{w}_i;\boldsymbol{0};\boldsymbol{\Sigma})$$

where $\boldsymbol{w}_i = ((2y_{i1}-1)V_{i1},..., (2y_{iT}-1)V_{iT})$ and Φ express multivariate normal density of T responses. The log-likelihood function is

$$LL(\boldsymbol{\theta};\boldsymbol{\Sigma}) = \sum_{i=1}^{n} \log \Phi(\boldsymbol{w}_i;\boldsymbol{0};\boldsymbol{\Sigma})$$

$\Phi(\boldsymbol{w}_i;\boldsymbol{0};\boldsymbol{\Sigma})$ is determined by GHK simulation using the factor of Cholesky$\boldsymbol{C}$. Because estimated parameter $\boldsymbol{\omega} = (\boldsymbol{\theta},\boldsymbol{c})$. $\boldsymbol{c}$ is elements of matrix $\boldsymbol{C}$.Utility equation become :

$$U_{it} = V_{it} + \sum_{l=1}^{t} c_{tl}\eta_{li} \text{ for t=1,...,T and } \eta_i \sim \mathbf{N(0,I)} \tag{9}$$

By using the algorithm of GHK simulation, the result is

$$\tilde{\pi}_{\mathrm{i}}^{(r)} = \prod_{t=1}^{T} \Phi_{it}^{(r)} = \prod_{t=1}^{T} \Phi\left(\frac{(2y_{it} - 1)V_{it} + \sum_{k=1}^{t-1} c_{tk}\eta_k^{(r)}}{c_{tt}} \right)$$

Therefore

$$\tilde{\pi}_i = \frac{1}{R}\sum_{r=1}^{R} \tilde{\pi}_i^{(r)}$$

Index of r states simulation r. The function of simulated log-likelihood is

$$\text{simlog} L(\boldsymbol{\omega}) = \sum_{i=1}^{n} \log\left(\frac{1}{R}\sum_{r=1}^{R} \tilde{\pi}^{(r)} \right) = \sum_{i=1}^{n} \log\left(\frac{1}{R}\sum_{r=1}^{R}\prod_{t=1}^{T} \Phi_{it}^{(r)} \right)$$

If it is known the utility model as represented in equation (9) that is fit to regularity condition and $\boldsymbol{\varepsilon}_i=(\varepsilon_{i1},..., \varepsilon_{iT})'$ is normal multivariate distributed with the mean of null and covariance matrix $\boldsymbol{\Sigma}$, so by using GHK simulation, MLE for the parameter of $\boldsymbol{\omega} = (\boldsymbol{\theta},\boldsymbol{\Sigma})$ is the solution of estimator equation:

$$\sum_{i=1}^{n} \frac{1}{\frac{1}{R}\sum_{r=1}^{R} \tilde{\pi}_i^{(r)}} \frac{1}{R}\sum_{r=1}^{R}\left(\left(\tilde{\pi}_i^{(r)}\right) \sum_{l=1}^{T} \frac{\phi_{li}^{(r)}}{\Phi_{li}^{r}} . \frac{a_{li}^{(r)}}{\partial \boldsymbol{\omega}} \right) = 0$$

where

$$\frac{\partial a_{li}^{(r)}}{\partial \boldsymbol{\theta}} = \begin{cases} \left[\sum_{h=1}^{l-1} \frac{c_{lh}}{c_{ll}} u_{hi}^{(r)} . \frac{\phi(a_{hi}^{(r)})}{\phi(\eta_{hi}^{(r)})} . \frac{\partial a_{hi}^{(r)}}{\partial \boldsymbol{\theta}} + \frac{(2y_{il} - 1)}{c_{ll}} \frac{\partial V_{il}}{\partial \boldsymbol{\theta}} \right], & l > 1 \\ \frac{(2y_{il} - 1)}{c_{11}} \frac{\partial V_{il}}{\partial \boldsymbol{\theta}}, & l = 1 \end{cases}$$

Index of r represent the simulation of r^{th}, r=1,..,R.

3. SIMULATION STUDY AND DISCUSSION

Multivariate analysis is generalization of univariate analysis. As data have low correlation, problem of multivariate can be solved by univariate analysis. However as correlation between the responses are strong, univariate approximation for multivariate case results the estimator that is under estimate. So we studied the effect of correlation level on estimating accuration based on simulation data generating at a fixed parameter value and some level of correlation.Statistics software that was used in this calculation is R.2.8.1 program, and three estimation methods used were :

a. GEE: estimation using GEE from Liang-Zeger
b. MLE: estimation using MLE method.
c. MSLE : estimation using MSLE method based on GHK simulation

The case of T=3 is taken. Utility model of subject i selecting the alternative j on the decision of t is

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

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

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

where i=1,...,n and t=1,2,3;j=0,1.ε_{ijt} ~N(0,1). Data were generated on the parameter value of α_t =-1,β_t = 0.5 and γ_t=0.3

Structure of correlation that will be examined is $r_{12}=\rho$ and $r_{13}=r_{32}= 0$. Utility on t=1 is correlated with the utility on t=2 with the values of correlation, ρ=0,0.2,...,0.9. The value of observation variable X_i and Z_{ijt} were taken from the Normal distribution,

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

Survey on the effect of correlation to the estimator was conducted on GEE, MLE and MSLE for n=1000. On each sample, the iteration for 50 times is performed. Results of estimating parameter are presented in Table 1, Table 2 and Table 3.

Table 1.Average of estimator using GEE.

θ	ρ_{21}					
	0	0.2	0.4	0.6	0.8	0.9
α_1=-1	-0.9852	-1.0363	-1.0125	-1.0685	-1.0284	-1.0600
α_2=-1	-1.0145	-1.0058	-1.0780	-1.0516	-1.0135	-1.0175
α_3=-1	-0.9875	-1.0250	-1.0804	-1.0981	-1.0085	-1.0481
β_1=0.5	0.4740	0.5127	0.4585	0.5391	0.4692	0.5387
β_2=0.5	0.5446	0.5530	0.4697	0.5212	0.4707	0.4716
β_3=0.5	0.4948	0.5147	0.5684	0.4924	0.5532	0.4889
γ_1=0.3	0.2915	0.3078	0.3047	0.3219	0.3180	0.3181
γ_2=0.3	0.3109	0.2996	0.3204	0.3239	0.3194	0.2930
γ_3=0.3	0.2890	0.3177	0.3183	0.3490	0.2918	0.3175
ρ_{21}	-0.0223	0.0861	0.2152	0.3526	0.5182	0.5672

Based on Table 1, GEE is good method to estimate regression parameter except correlation parameter (ρ_{21}). The estimatorsof α, β, and γ are close(small bias) to the true parameter at all of correlation value. So, the value of correlation among utility does not influence to GEE estimator but $\hat{\rho}_{21}$ is always underestimate.

Table 2. Average of estimator using MLE

θ	ρ_{21}					
	0	0.2	0.4	0.6	0.8	0.9
α_1=-1	-0.9829	-1.0331	-1.0108	-1.0710	-1.0159	-1.0455
α_2=-1	-0.0296	0.0274	-0.0648	0.0330	0.0006	0.0297
α_3=-1	-0.0025	0.0105	-0.0668	-0.0250	0.0092	-0.0010
β_1=0.5	0.4737	0.5115	0.4583	0.5415	0.4636	0.5287
β_2=0.5	0.5442	0.5515	0.4670	0.5190	0.4777	0.4775
β_3=0.5	0.4940	0.5135	0.5674	0.4919	0.5529	0.4885
γ_1=0.3	0.2900	0.3060	0.3038	0.3221	0.3127	0.3146
γ_2=0.3	0.3096	0.2990	0.3190	0.3185	0.3184	0.2914
γ_3=0.3	0.2878	0.3166	0.3167	0.3478	0.2907	0.3161
ρ_{21}	-0.0417	0.1584	0.3799	0.5996	0.8139	0.8915

Utility 1 (U_{i1}) is correlated with utility 2 (U_{i2}) and both utility are not correlated to utility 3 (U_{i3}). Therefore correlation value is only affecting the parameters within U_{i1} and U_{i2}. At both utility, bias of estimator is too high and proportional to the value of correlation. MLE of parameter α_2 and $\alpha_3$are not good because it produced the high bias (see Table 2). Estimator of ρ_{21} in MLE is more better than GEE.

Table 3. Average of estimator using MSLE

θ	ρ_{21}					
	0	0.2	0.4	0.6	0.8	0.9
α_1=-1	-0.3499	-0.6201	-0.2942	-0.1758	-0.3694	-0.1985
α_2=-1	0.3167	1.0063	0.5998	0.3530	0.5352	0.7078
α_3=-1	0.1514	-0.0859	0.0188	-0.0055	-0.1044	0.1018
β_1=0.5	-0.0966	-0.3595	-0.4914	-0.7192	-0.2877	-0.3125
β_2=0.5	0.3391	0.5567	0.3735	0.5089	0.4776	0.5845
β_3=0.5	-0.0084	-0.1065	0.1623	0.1867	-0.0379	0.0977
γ_1=0.3	-0.3336	-0.0899	-0.6505	-0.6186	-0.4069	-0.2815
γ_2=0.3	0.4304	0.8490	0.5980	0.6814	0.4885	0.4382
γ_3=0.3	-0.0789	-0.1702	0.2369	-0.2030	0.1326	0.1235
ρ_{21} =0	0.9876	0.9900	0.7483	0.4796	0.9476	0.9900

Result of parameter estimation by using GHK simulation produce very high bias (see Table 3). In other side, the value of estimator is influenced by initial estimator value. Problem encountered in ProbitModel is that its log-likelihood function are not globally concave. This causes the maximum global point difficult to define and the computation will be not efficient (time consuming) to reach convergence point.

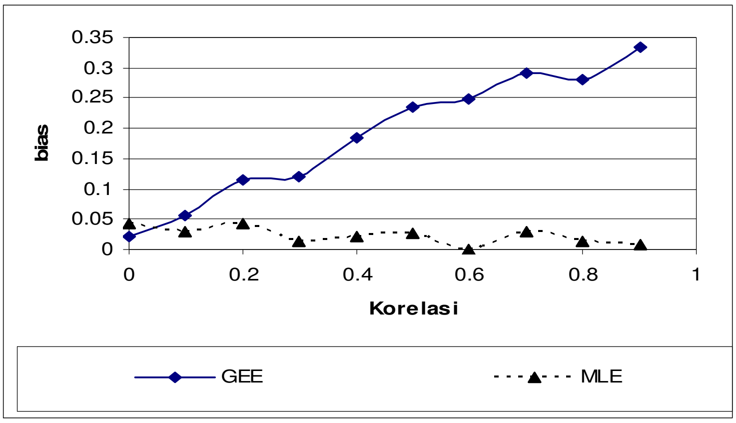

Figure 1. Comparisons of bias correlation

Comparisons of bias correlation have description on Figure 1. Estimator parameter using MLE is better (lower bias) compared to GEE. Estimator of correlation parameter in GEE has tendency to underestimate proportional to the magnitude of correlation.

4. CONCLUDING REMARKS

ProbitModel can be applied on binary multivariate response by using MLE and GEE estimation method. Based on simulation data,

1. GEE estimator for regression coefficients is not affected by the value of correlation between the responses.
2. MLE estimator for regression coefficients is affected by the value of correlation between the responses.
3. GEE estimator for correlation parameter tends to be underestimated. Whereas MLE method is more accurate to estimate the correlation parameter.
4. MSLE estimator not appropriate to multivariate ProbitModel

Open Problem

In this research, estimation of parameter used are MLE and GEE. It is very possible to use other estimation method such as Bayes methods. From computational side, simulation method applicable for Probit model is need to be developed to overcome the limitation of GHK method.

References

[1] LIANG, K.Y., AND ZEGER,S.L, Longitudinal Data Analysis Using Generalised Linear Models, *Biometrika*73, 13-22, 1986.

[2] PRENTICE, Correlated Binary Regression with Covariates Specific to Each Binary Observation. *Biometrics 44*, 1043-1048, 1988.

[3] HARRIS, M.N, MACQUARIE L.R AND SIOUCLIS AJ., Comparison of alternative Estimators for Binary Panel Probit Models, *Melbourne Institute Working Paper* no 3/00, 2000

[4] CONTOYANNIS P, ANDREW M. J, AND RICE N, Dynamics of Health in British Household: Simulation-Based Inference in Panel Probit Model, *Working Paper*, Department of Economics and Related Studies, University of York, 2001.

[5] HAJIVASSILIOU, V., D. MCFADDEN, AND RUUD P., Simulation of Multivariate Normal Rectangle Probabilities and Their derivatives: Theoretical and Computational Results, *Journal of Econometrics* 72, 85–134, 1996.

[6] GEWEKE J.F., KEANE M.P., ANDRUNKLE D.E., Statistical Inference in The Multinomial MultiperiodeProbit Model, *Journal of Econometrics* 80, 125-165, 1997.

[7] NUGRAHA J., GURITNO S., ANDHARYATMI S., Logistic Regression Model on Multivariate Binary Response Using Generalized Estimating Equation, National Seminar on Math and Education of Math conducted by UNY, Indonesia, 2006.

[8] NUGRAHA J, HARYATMI S. AND GURITNOS, A Comparison of MLE and GEE on Modeling Binary Panel Response, ICoMS3th IPB, 2008.

[9] NUGRAHA J., HARYATMI, ANDGURITNO, Logistic Regression Model on Multivariate Binary Response Using Generalized Estimating Equation, *Proceeding* of National Seminar on Mathematicsconducted by UNY, Indonesia FMIPA UNY, 2006.

[10] NUGRAHA J., GURITNO S.,AND HARYATMIS., Likelihood Function and its Derivatives of Probit Model on Multivariate Biner Response, *JurnalKalam*, Vol. 1 No. 2, Faculty of Science and Technology, Universiti Malaysia Terengganu, Malaysia, 2008

[11] NUGRAHA J., GURITNO S.,AND HARYATMIS., ProbitModel on Multivariate Binary ResponsUsing SMLE, *JurnalIlmuDasar*, FMIPA Univ. Jember, 2010.

[12] LECHNER M, LOLLIVIER S AND MAGNAGT, Parametric Binary Choice Models, *Discusion paper* no 2005-23, 2005.

[13] HAJIVASSILIOU V., MCFADDEN, AND RUUD P., Simulation of Multivariate Normal Rectangle Probabilities and Their derivatives: Theoretical and Computational Results, *Journal of Econometrics 72*, 85–134, 1996.

[14] TRAIN, KENNETH, *Discrete Choice Methods with Simulation*, UK Press, Cambridge, 2003.

JAKANUGRAHA
Dept. of Statistics, Islamic University of Indonesia, Kampus Terpadu UII, Jl. Kaliurang Km.14, Yogyakarta, Indonesia
e-mails: jnugraha@uii.ac.id or jk.nugraha@gmail.com