Thailand Statistician
January 2012; 10(1) : 1-14
http://statassoc.or.th
Contributed paper

# A Comparison of Parameter Estimation Approaches for Generalized Linear Mixed Models with Correlated Ordinal Responses

**Chayanat Phonork and Veeranun Pongsapukdee***

Department of Statistics, Faculty of Science, Silpakorn University, SnamChandra Palace Campus, Nakhon Pathom 73000, Thailand.

* Author for correspondence; e-mail: veeranun@su.ac.th

Received: 25 April 2011
Accepted: 25 September 2011

**Abstract**

Parameter estimation approaches under random effects such as in Generalized Linear Mixed Models (GLMMs) for categorical responses are generally complicated, especially; those models with correlated ordinal response categories. Then, research and development in these areas are still currently interested. The objective of this research is to compare the parameter estimation approaches for GLMMs under cumulative logit models that based on both the fixed and random effects and that based on only fixed effects models (Generalized Linear Models (GLMs)) including marginal models by considering the bias, the mean square error, and the coverage from 1,000 datasets in each condition of sample size (number of clusters and cluster size) and the intra-cluster correlation. The GLMs are investigated for based comparative study approach using maximum likelihood or Fisher scoring. In marginal models, two approaches, Generalized Estimating Equation (GEE) and Marginal Quasi-Likelihood (MQL) are performed. In random effects models, we use mixed effects such that GLMMS models which corresponding the two approaches: Penalized Quasi-Likelihood (PQL) and Adaptive Gaussian Quadrature (AGQ) are investigated. These models are distinguished in the form of cumulative logit models. The results from every condition of sample sizes and intra-cluster correlations, strongly indicate that for the marginal model, the better approach is MQL compared to the GEE, as the intra-cluster correlation is medium and large. By contrast, they give approximately the same results as that of the ML when the intra-cluster correlation is small. For the random effects model, the AGQ is better than the PQL. In conclusion, the AGQ is the best when it is compared with all the others. Moreover, as the number of cluster and the cluster size are increased, the accuracy of parameter estimates is much improved. However, the increasing of the intra-cluster correlation affects the GEE and the MQL estimates. But, such impact is only little for the AGQ and the PQL estimates.

______________________________

**Keywords:** Mixed-effects modeling, Cumulative logit models; Marginal model; Generalized Estimating Equations (GEEs); Marginal Quasi-Likelihood (MQL); Generalized Linear Mixed Models (GLMMs); Adaptive Gaussian Quadrature (AGQ); Penalized Quasi-Likelihood (PQL); Correlated ordinal response categories.

## 1. Introduction

Many observations of the response variable for each subject are possibly collected and measured repeatedly at several times or under various different conditions. The data are examined and combined with the models of interest under a process called fitting a statistical model. The statistical analyses of trying to understand the data involves looking for patterns in the data. Parameter estimation approaches under random effects models such as Generalized Linear Mixed Models (GLMMs) for categorical responses are generally complicated, especially; those models with correlated ordinal response categories. In this article, we empirically investigate firstly the modeling the ordinal response variable, primarily through the use of multinomial generalized linear models, or shortly multinomial GLMs [1]. Model parameters that describe a factor's effects in ordinary linear models and GLMs are called fixed effects. They apply to all categories of interest. GLMs appeared on the statistical scene in the path breaking article of Nelder and Wedderburn [2], McCullagh and Nelder [3]. These authors generalize the classical linear models based on the normal distribution to involve two aspects: a variety of distributions from continuous to discrete or categorical, exponential family distribution models, and allowing transformations of mean, through the link functions of mean, linking the systematic part of models to the mean of one of the distributions. GLMs are now a mature and well known data-analytic methodology, including the most popular cumulative logit models, known as proportional odds models [4, 5]. For fixed effects the maximum likelihood approach or Fisher's scoring (ML), Generalized Estimating Equation (GEE, [6, 7]), and Marginal Quasi-Likelihood (MQL, Goldstein [8, 9]) are performed under GLMs and Marginal models. In marginal models, two approaches, Generalized Estimating Equation (GEE) and Marginal Quasi-Likelihood (MQL) are performed.

On the other hand, secondly, we also investigate the mixed modeling, GLMMs. Random effects usually apply to a sample, for example, the model treats observations from a given clinics as a cluster, and it has a random effect in each clinic. In mixed effects models, the two corresponding approaches, Penalized Quasi-Likelihood (PQL, [10]) and Adaptive Gaussian Quadrature (AGQ, [11, 12]) are carried out under GLMMs models. Thus, not only using GLMs, by contrast, GLMMs are further extended that permit random effects and fixed effects in the linear predictors. Following up to Searle's classic linear models [13] and variance components by Searle et al. [14] the modern perspective on GLMMs can be useful to provide a unified and accessible treatment of the newest statistical methods for analyzing correlated, non-normally distributed data [15]. However, building and extending GLMMs to analyzing repeated categorical response data are still not straight forward statistical methods. Several offered statistical estimation approaches must be selected efficiently and correctly but there are only few previous comparisons among these works. Therefore, in this paper, five approaches for model parameter estimation that consist of ML, GEE, MQL for GLMs and the others, PQL and AGQ for GLMMs [16] are investigated and compared for their efficacies considering the bias, the mean square error, and the coverage from 1,000 datasets in each sample size (number of clusters and cluster size) and the intra-cluster correlations. All simulation works are generated and processed using the authors' rewritten macro run with SAS® version 9.1.

## 2. Models and Parameter estimation approaches

### 2.1 Models for fixed effects

Let N be the number of independent clusters or subjects, $n_i$ the number of observations for cluster *i*, *i = 1,…, N* and let $Y_{ij}$ denote the response *j, j = 1,…, n*$_i$, whose values consist of k level, *k = 1, 2,…, c.* Furthermore, let $\boldsymbol{x}_{ij} = (x_{1ij}, \ldots, x_{pij})'$ denote a column vector of p-covariates associated with the response j, in cluster i. We assume a cumulative model in the form:

$$\text{logit}[P(Y_{ij} \le k \mid \boldsymbol{x}_{ij})] = \eta_{ijk} = \alpha_k + \boldsymbol{x}_{ij}'\boldsymbol{\beta}, \quad k = 1, \ldots, c-1. \qquad \ldots\ldots\ldots (1)$$

Where $\alpha_k$ denote the cutpoint for the response level *k*, $\alpha_1 < \alpha_2 < \ldots < \alpha_{c-1}$ and $\boldsymbol{\beta}$ denote the fixed effects of the explanatory variables. Thus,

$$P(Y_{ij} \le k) = 1/[1+\exp(-\eta_{ijk})],$$

and $\pi_k(\boldsymbol{x}_{ij}) = P(Y_{ij} = k \mid \boldsymbol{x}_{ij}) = P(Y_{ij} \le k \mid \boldsymbol{x}_{ij}) - P(Y_{ij} \le k-1 \mid \boldsymbol{x}_{ij})$.

The maximum likelihood function for the ordinal response for the ML approach is

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\prod_{k=1}^{c} \pi_k(\boldsymbol{x}_{ij})^{y_{ijk}}.$$

Where $y_{ijk}$ = 1 if $Y_{ij}$ = *k* and $y_{ijk}$ = 0 if $Y_{ij}$ $\neq$ *k*. The log-likelihood function is

$$l(\boldsymbol{\beta}) = \log L(\boldsymbol{\beta}) = \sum_{i=1}^{N}\sum_{j=1}^{n_i}\sum_{k=1}^{c} y_{ijk}\log \pi_k(\boldsymbol{x}_{ij}).$$

The estimates of $\boldsymbol{\beta}$ (including $\alpha_k$) are the results from Fisher scoring [17] then have the form

$$u(\boldsymbol{\beta}) = \frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{N} \boldsymbol{D}_i' \boldsymbol{V}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{\pi}_i) = \boldsymbol{0}.$$

Where, $D_i$ denotes the matrix of derivatives for the element i, $\frac{\partial \boldsymbol{\pi}_i}{\partial \boldsymbol{\beta}} = \frac{\partial \boldsymbol{\pi}_i}{\partial \boldsymbol{\eta}_i}\frac{\partial \boldsymbol{\eta}_i}{\partial \boldsymbol{\beta}} = \frac{\partial \boldsymbol{\pi}_i}{\partial \boldsymbol{\eta}_i}\boldsymbol{x}_i$ and $\boldsymbol{V}_i = [diag\{var(Y_{ij})\}]$, *i = 1,..., N*.

**2.2 Models for random effects**

Let $\boldsymbol{u}_i$ denote a column vector of random effect values for cluster i. This is common to all observations in the cluster. Let $\boldsymbol{z}_{ij}$ denote a column vector of their explanatory variables. The linear predictor for a GLMM has the form (2)

$$\text{logit}[P(Y_{ij} \le k \mid \boldsymbol{x}_{ij}, \boldsymbol{u}_i)] = \alpha_k + \boldsymbol{x}_{ij}'\boldsymbol{\beta} + \boldsymbol{z}_{ij}'\boldsymbol{u}_i, \quad k = 1,\ldots,c-1. \quad \ldots\ldots. \ (2)$$

The random effect $\boldsymbol{u}_i$ is assumed to have a multivariate normal distribution N($\boldsymbol{0}$, $\boldsymbol{\Sigma}_u$). The covariance matrix $\boldsymbol{\Sigma}_u$ depends on unknown variance components and possibly also correlation parameters. The special case with univariate random effect and $\boldsymbol{z}'_{ij} = 1$. Then with $\boldsymbol{u}_i$ replaced by $u_i$, where $\{u_i\}$ are $N(0, \sigma_u^2)$, the GLMM has the form

$$\text{logit}[P(Y_{ij} \le k \mid \boldsymbol{x}_{ij}, u_i)] = \alpha_k + \boldsymbol{x}'_{ij}\boldsymbol{\beta} + u_i, \quad k = 1, \dots, c-1.$$

This has the form of ordinary GLMs with unobserved values $\{u_i\}$ of a particular covariate. The random effects also provide a mechanism for explanatory over dispersion in basic models not having those effects [18]. Conditional on $\boldsymbol{u}_i$, the conditional likelihood function for cluster i have the form

$$f(Y_i \mid \boldsymbol{u}_i; \boldsymbol{\beta}) = \prod_{j=1}^{n_i} \prod_{k=1}^{c} \pi_k^u(\boldsymbol{x}_{ij})^{y_{ijk}}.$$

Where, $\pi_k^u(\boldsymbol{x}_{ij}) = P(Y_{ij} = k \mid \boldsymbol{x}_{ij}, \boldsymbol{u}_i) = P(Y_{ij} \le k \mid \boldsymbol{x}_{ij}, \boldsymbol{u}_i) - P(Y_{ij} \le k-1 \mid \boldsymbol{x}_{ij}, \boldsymbol{u}_i)$.

Without conditional on $\boldsymbol{u}_i$, the conditional likelihood function for cluster i (marginal) have the form

$$L_i(\boldsymbol{\beta}, \boldsymbol{\Sigma}_u) = f(Y_i; \boldsymbol{\beta}, \boldsymbol{\Sigma}_u) = \int f(Y_i \mid \boldsymbol{u}_i; \boldsymbol{\beta}) \times f(\boldsymbol{u}_i; \boldsymbol{\Sigma}_u) d\boldsymbol{u}_i.$$

Then its log likelihood function is $l(\boldsymbol{\beta}, \boldsymbol{\Sigma}_u) = \log L = \sum_{i=1}^{N} \log L_i$. The log likelihood function is evaluated numerically and maximized as a function of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}_u$ or Monte Carlo in combination with Newton-Raphson or Fisher scoring may be used iteratively [17].

### 2.2.1 Adaptive Gaussian Quadrature Approach

Adaptive Gaussian Quadrature (AGQ) approach was proposed by Pinheiro & Bates [11]. The approximation is a finite weighted sum that evaluates the function at certain points in the form

$$\int_{-\infty}^{\infty} f(\boldsymbol{u}) exp(-\boldsymbol{u}^2) d\boldsymbol{u} \approx \sum_{t=1}^{q} w_t f(d_t).$$

Where, $w_t$ denote weights, $d_t$ denote quadrature points that are tabulated. The approximation improves as t, the number of quadrature points (t) increases.

### 2.2.2 Penalized Quasi-Likelihood Approach

Penalized quasi-likelihood (PQL) was proposed by Green [10]. His integral has closed form. This type of integral approximation is called Laplace approximation that evaluates the function at certain points around $\bar{\boldsymbol{u}} \approx E(\boldsymbol{u} \mid \boldsymbol{Y})$ in the form

$$q(\boldsymbol{\beta}, \boldsymbol{Y}) - (1/2)\bar{\boldsymbol{u}}'\boldsymbol{\Sigma}_u^{-1}\bar{\boldsymbol{u}},$$

where, $q(\boldsymbol{\beta}, \boldsymbol{Y})$ is similar to the quasi-log-likelihood of GLMs. Then, Monte Carlo method in combination with Newton-Raphson or Fisher scoring may be used iteratively. However, in case when the response is binary PQL tends to produce estimates with negative bias [19].

## 2.3 Marginal models

Conditional on the random effects, observations are assumed to follow a GLM. Then, the following parameter estimation approaches may be applied.

### 2.3.1 Generalized Estimating Equation Approach

Generalized Estimating Equation (GEE) was originally proposed by Liang and Zeger [6] for modeling univariate marginal distribution, such as the binomial and Poisson and Lipsitz et al. [20] outlined a GEE approach for cumulative logit models with repeated ordinal responses. GEE is an alternative to ML estimation but does not completely specify the joint distribution, then it does not have a likelihood function. The quasi-likelihood parameter estimates are the solutions of quasi-score equations

$$u(\boldsymbol{\beta}) = \frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{N} \boldsymbol{D}_i' \boldsymbol{V}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{\pi}_i) = \boldsymbol{0},$$

where, $\pi_i$ denote the probability associated with $\boldsymbol{Y}_i$, $D_i$ denotes the matrix of derivatives for the element i, $\frac{\partial \boldsymbol{\pi}_i}{\partial \boldsymbol{\beta}}$ and $\boldsymbol{V}_i = \boldsymbol{A}_i^{1/2}\boldsymbol{R}_i(\rho)\boldsymbol{A}_i^{1/2}$ is working covariance matrix for $\boldsymbol{Y}_i$, $\boldsymbol{A}_i$ denotes working correlation matrix. Let $\boldsymbol{R}_i(\rho)$ denotes the working correlation matrix $n_i \times n_i$ for $Y_i$ from which it follows independence ($\rho_{jj'} = 0$), exchangeable ($\rho_{jj'} = \rho$) or autocorrelation $\rho_{jj'} = \rho^{|j-j'|}$. The consistency also follows from general results for unbiased estimating functions [21].

### 2.3.2 Marginal quasi-likelihood Approach

Marginal Quasi-Likelihood (MQL) was proposed by Goldstein [8]. The model for marginal distribution has the form $E(\boldsymbol{y}_i)=\boldsymbol{\mu}_i=g^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}^*)$ or $\boldsymbol{y}_i=\boldsymbol{\mu}_i^u+\boldsymbol{\varepsilon}_i$, where, $\boldsymbol{u}\sim N(\boldsymbol{0},\boldsymbol{\Sigma}_u)$ and $var(\boldsymbol{\varepsilon}_i)=a_i\upsilon(\boldsymbol{\mu}_i^u)$, $a_i$ is a constant and $\upsilon(\boldsymbol{\mu}_i^u)$ denotes the function of the conditional mean. The likelihood parameter estimates are the solutions for GLMs which follow the equation $u(\boldsymbol{\beta})=\boldsymbol{D}'[\mathrm{var}(\boldsymbol{y})]^{-1}(\boldsymbol{y}-\boldsymbol{\mu})$ in the form

$$\boldsymbol{x}'(\boldsymbol{D}\boldsymbol{V}_0\boldsymbol{D}+\boldsymbol{z}\boldsymbol{\Sigma}_u\boldsymbol{z})^{-1}\boldsymbol{D}(\boldsymbol{y}-\boldsymbol{\mu})=\boldsymbol{0},$$

where, $var(\boldsymbol{y})=\boldsymbol{V}_0+\boldsymbol{D}^{-1}\boldsymbol{z}\boldsymbol{\Sigma}_u\boldsymbol{z}\boldsymbol{D}^{-1}$, $D$ and $V$ denote the diagonal matrices whose element are $\dfrac{\partial\,\boldsymbol{\mu}_i}{\partial\,\boldsymbol{\beta}}$ and $a_i\upsilon(\boldsymbol{\mu}_i^u)$, respectively.

## 3. Simulation and Estimation

Each condition of sample size or number of clusters ($N$ = 50, 100, 250) and the cluster size ($n_i$ = 5, 7, 10), the intra-cluster correlation (ICC = 0.05, 0.20, 0.40) which are corresponding to the variance of the random effects ($\sigma_u^2$ = 0.17, 0.82, 2.19) since $ICC=\sigma_u^2/\left(\sigma_u^2+\sigma_\varepsilon^2\right)$ where $\sigma_\varepsilon^2=\pi^2/3$ the logit models in (1) – (2) and parameter values that are used similarly ($\alpha_1$ = -1, $\alpha_2$ = 1, $\beta_1$ = 0.5, $\beta_2$ = 0.5) as that of Spiess & Hamerle [22] are performed. 1,000 datasets are simulated and processed for each parameter estimation and each model condition as the following stages in Figure 1.

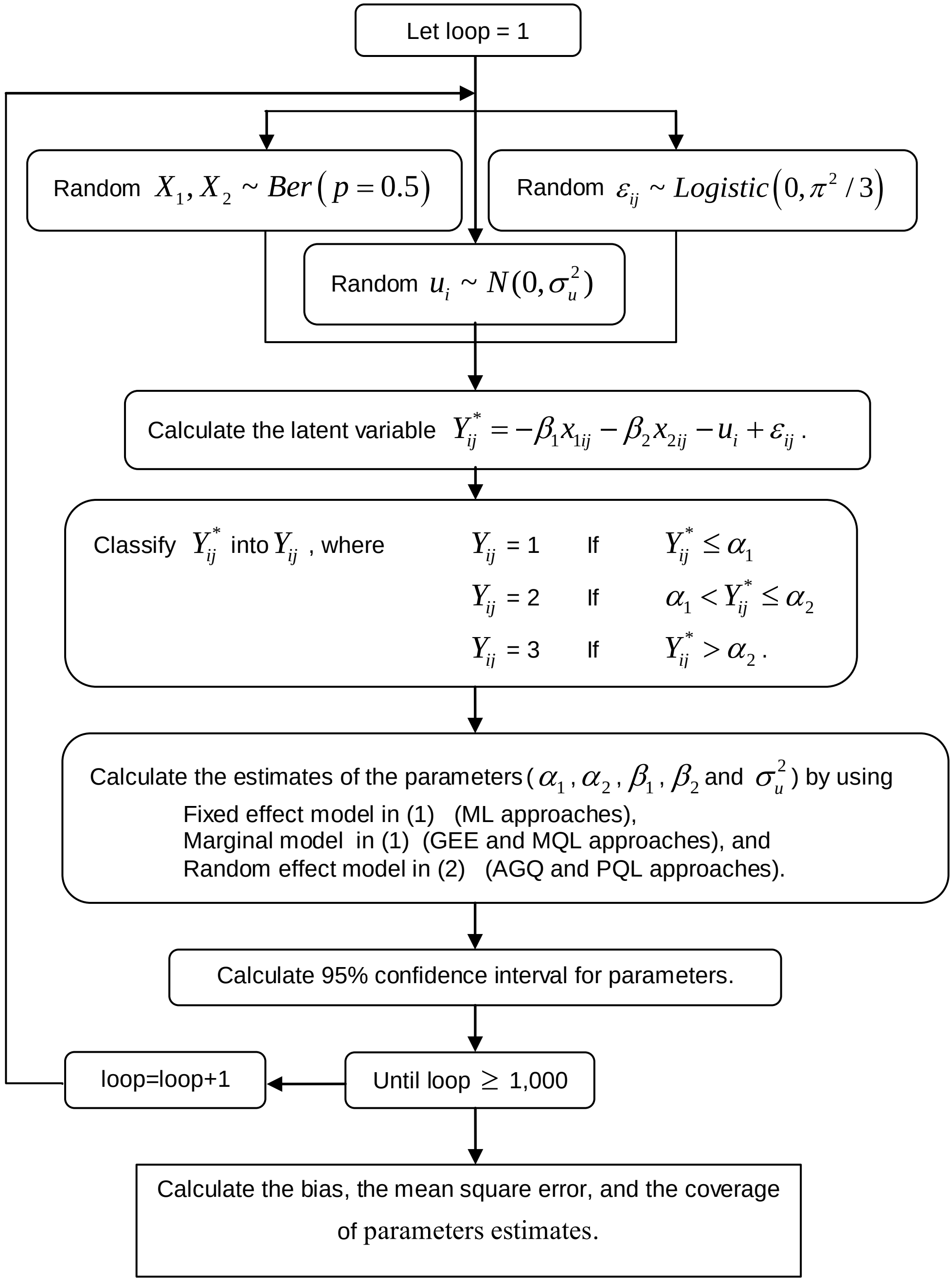


**Figure 1. Simulation and Estimation.**

## 4. Research results

The parameter estimation approaches for generalized linear mixed models with correlated ordinal responses are compared for the model parameters; $\alpha_1, \alpha_2, \beta_1, \beta_2$ and $\sigma_u^2$. The results of which shown in Tables 1-3 is for the parameter $\beta_1$ when ICC = 0.05, 0.20, 0.40; $N$ = 50, 100, 250; $n_i$ = 5, 7, 10, respectively. The rest of parameter estimations are also obtained very similarly results to those for the parameter $\beta_1$.

Table 1. Bias, MSE, and Coverage (%) for $\beta_1$ under estimation approaches when ICC = 0.05.

| N | n | Bias | | | | | MSE | | | | | Coverage(%) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ML | GEE | MQL | AGQ | PQL | ML | GEE | MQL | AGQ | PQL | ML | GEE | MQL | AGQ | PQL |
| | 5 | -.049 | -.049 | -.049 | -.033 | -.041 | .058 | .058 | .057 | .051 | .059 | 94.6 | 94.1 | 95.3 | 96.5 | 95.4 |
| 50 | 7 | -.041 | -.041 | -.039 | -.026 | -.031 | .043 | .043 | .043 | .041 | .044 | 94.7 | 93.7 | 95.2 | 96.1 | 94.6 |
| | 10 | -.048 | -.048 | -.050 | -.033 | -.041 | .031 | .031 | .030 | .029 | .030 | 93.9 | 92.9 | 93.9 | 95.2 | 93.9 |
| | 5 | -.055 | -.055 | -.054 | -.041 | -.048 | .030 | .030 | .030 | .028 | .030 | 94.4 | 94.1 | 95.3 | 95.8 | 94.9 |
| 100 | 7 | -.043 | -.043 | -.044 | -.028 | -.034 | .022 | .022 | .022 | .022 | .022 | 93.3 | 93.3 | 93.9 | 94.3 | 93.8 |
| | 10 | -.052 | -.052 | -.052 | -.038 | -.045 | .017 | .017 | .017 | .016 | .017 | 93.1 | 92.6 | 93.8 | 94.6 | 94.1 |
| | 5 | -.056 | -.056 | -.055 | -.042 | -.045 | .014 | .014 | .014 | .013 | .014 | 92.0 | 91.6 | 92.3 | 93.5 | 91.9 |
| 250 | 7 | -.049 | -.049 | -.048 | -.035 | -.043 | .010 | .010 | .010 | .009 | .010 | 92.3 | 92.1 | 92.9 | 94.0 | 92.8 |
| | 10 | -.053 | -.053 | -.057 | -.039 | -.048 | .008 | .008 | .009 | .007 | .008 | 89.8 | 89.8 | 91.0 | 92.4 | 91.7 |

Table 2. Bias, MSE, and Coverage (%) for $\beta_1$ under estimation approaches when ICC= 0.20.

| N | n | Bias | | | | | MSE | | | | | Coverage(%) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ML | GEE | MQL | AGQ | PQL | ML | GEE | MQL | AGQ | PQL | ML | GEE | MQL | AGQ | PQL |
| | 5 | -.086 | -.086 | -.085 | -.031 | -.054 | .063 | .063 | .059 | .068 | .062 | 93.6 | 92.0 | 94.9 | 95.2 | 95.0 |
| 50 | 7 | -.081 | -.081 | -.080 | -.025 | -.044 | .045 | .045 | .042 | .047 | .044 | 93.6 | 92.8 | 95.3 | 96.1 | 95.6 |
| | 10 | -.088 | -.088 | -.088 | -.034 | -.051 | .035 | .035 | .032 | .031 | .031 | 91.3 | 90.6 | 93.4 | 96.1 | 95.1 |
| | 5 | -.095 | -.095 | -.096 | -.041 | -.065 | .038 | .038 | .036 | .035 | .034 | 91.8 | 91.3 | 94.0 | 95.4 | 94.8 |
| 100 | 7 | -.084 | -.084 | -.083 | -.027 | -.048 | .027 | .027 | .025 | .024 | .024 | 90.2 | 90.4 | 93.2 | 95.5 | 95.1 |
| | 10 | -.094 | -.094 | -.094 | -.038 | -.055 | .023 | .023 | .021 | .017 | .017 | 87.0 | 87.3 | 90.5 | 94.5 | 93.4 |
| | 5 | -.094 | -.094 | -.102 | -.039 | -.064 | .021 | .021 | .021 | .016 | .017 | 85.4 | 84.6 | 86.8 | 92.9 | 91.3 |
| 250 | 7 | -.089 | -.089 | -.089 | -.033 | -.056 | .015 | .015 | .015 | .010 | .012 | 83.8 | 83.1 | 86.1 | 93.1 | 91.0 |
| | 10 | -.093 | -.093 | -.092 | -.038 | -.056 | .014 | .014 | .013 | .008 | .008 | 76.1 | 76.4 | 80.0 | 94.2 | 92.9 |

Table 3. Bias, MSE, and Coverage (%) for $\beta_1$ under estimation approaches when ICC= 0.40.

| N | n | Bias | | | | | MSE | | | | | Coverage(%) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ML | GEE | MQL | AGQ | PQL | ML | GEE | MQL | AGQ | PQL | ML | GEE | MQL | AGQ | PQL |
| | 5 | -.144 | -.144 | -.143 | -.028 | -.069 | .077 | .077 | .066 | .079 | .069 | 90.6 | 89.5 | 95.4 | 95.1 | 95.5 |
| 50 | 7 | -.136 | -.135 | -.132 | -.015 | -.047 | .058 | .058 | .047 | .054 | .049 | 89.9 | 89.0 | 94.7 | 96.2 | 96.3 |
| | 10 | -.141 | -.141 | -.141 | -.025 | -.049 | .049 | .049 | .041 | .036 | .034 | 85.6 | 84.2 | 91.6 | 96.6 | 95.8 |
| | 5 | -.152 | -.152 | -.151 | -.037 | -.078 | .052 | .052 | .046 | .040 | .038 | 84.4 | 83.8 | 90.2 | 95.5 | 95.1 |
| 100 | 7 | -.139 | -.139 | -.138 | -.019 | -.052 | .039 | .039 | .035 | .027 | .026 | 83.3 | 82.5 | 88.2 | 95.5 | 94.5 |
| | 10 | -.150 | -.150 | -.149 | -.034 | -.058 | .037 | .037 | .033 | .019 | .019 | 74.7 | 74.9 | 81.9 | 94.4 | 94.2 |
| | 5 | -.151 | -.151 | -.150 | -.033 | -.077 | .034 | .034 | .031 | .016 | .018 | 68.9 | 68.9 | 75.8 | 94.4 | 91.2 |
| 250 | 7 | -.148 | -.148 | -.147 | -.029 | -.062 | .029 | .029 | .027 | .011 | .013 | 63.3 | 63.2 | 68.9 | 94.2 | 91.2 |
| | 10 | -.150 | -.150 | -.150 | -.033 | -.057 | .028 | .028 | .027 | .008 | .010 | 48.9 | 49.3 | 50.5 | 93.8 | 90.7 |

Tables 1-3 show that for the fixed effects models, the MQL has bias, MSE, and Coverage (%) that are closed to those of the GEE as well as the ML when the intra-cluster correlation is small. But it performs better than the GEE and the ML when the intra-cluster correlation is medium and large. For the random effects models the AGQ is better than the PQL.

Considering on all the five approaches' MSE when the cluster size is increased, it is found that the AGQ is better than the PQL and both of them give smaller MSE than that from the other approaches' (Figures 2-3).

Furthermore, considering on all the five approaches' MSE when the intra-cluster correlation is increased, this condition does affect the GEE and the MQL estimates. But, such impact is only little for the AGQ and the PQL estimates (Figure 4).

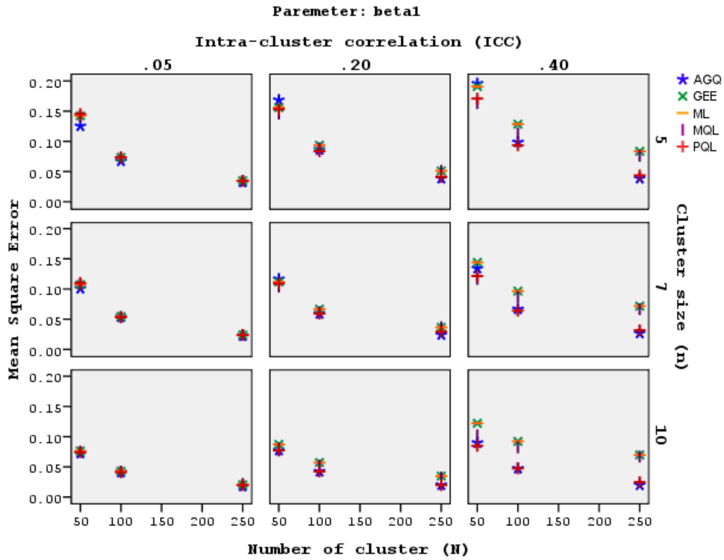


Figure 2. MSE and number of cluster (N) classified by intra-cluster correlation (ICC) and cluster size ( $n$ ).

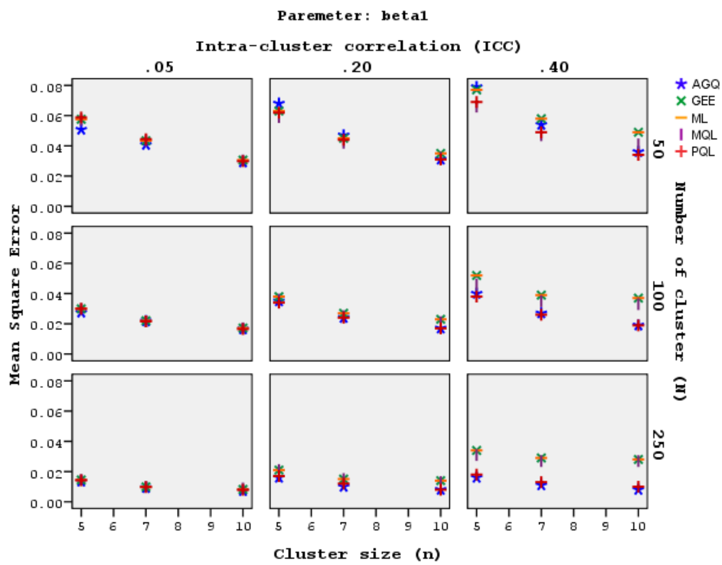


Figure 3. MSE and cluster size ( $n$ ) classified by Number of clusters (N) and Intra-cluster correlation (ICC)

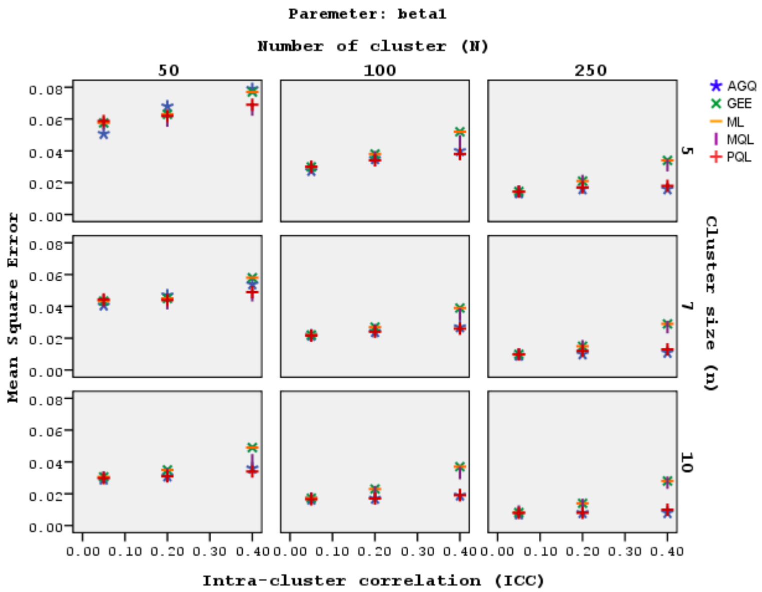


Figure 4. MSE and intra-cluster correlation (ICC) classified by Cluster sizes ( $n$ ) and Number of clusters (N).

## 5. Conclusion and Discussion

In general conclusion, the AGQ is the best when it is compared with all the other four approaches. Furthermore, as the number of cluster and the cluster size are increased, the accuracy of parameter estimates is much improved. However, the increasing of the intra-cluster correlation affects the GEE and the MQL estimates. This probably means that it might get harder and harder to classify the right categories when there are so many units in the same cluster or when the intra-cluster correlation occurs. But, such impact is only little for the AGQ and the PQL estimates. Thus, the AGQ approach trend to be chosen and would be the recommended approach for the generalized linear mixed models with correlated ordinal response categories.

## Acknowledgement

This research was partially supported by the Department of Statistics and Graduate School, Silpakorn University, Thailand. We also would like to thank the editors and the referees for their very helpful comments.

## References

[1] McFadden, D., Conditional logit analysis of qualitative choice behavior in Frontiers in Econometrics, ed. by P. Zarembka. Academic Press, New York, 1974; 105-142.

[2] Nelder, J.A. and Wedderburn, R.W.M., Generalized Linear Models. Journal Royal Statistics, 1972; A135: 370-384.

[3] McCullagh, P. and Nelder, J.A., Generalized Linear Models. Chapman & Hall. London, 1983. 2nd ed., 1989.

[4] Walker, S. H. and Duncan, D. B., Estimation of probability of an event as a function of several independent variables. Biometrika, 1967; 54: 167-179.

[5] McCullagh, P., Regression models for ordinal data (with discussion) Journal Royal Statistics, 1980; B42: 109–142.

[6] Liang, K. Y. and Zeger, S. L., Longitudianl data analysis using generalized linear models. Biometrika, 1986; 13-22.

[7] Zeger, S. L., Liang, K. Y., and Albert, P. S., Models for longitudinal data: A generalized estimating equation approach. Biometrics, 1988; 44: 1049-1060.

[8] Goldstein, H., Nonlinear multilevel models, with an application to discrete response data. Biometrika, 1991; 78: 45-51.

[9] Goldstein, H., Multilevel Statistical Models, 2nd ed., Edward Arnold. London, 1995.

[10] Green, P. J., Penalized likelihood for general semi-parametric regression models. Int. Statist Rev, 1987; 55: 245-60.

[11] Pinheiro, J. C. and Bates, D. M., Approximations to the log-likelihood function in the nonlinear mixed-effects models. Journal of Computational and Graphical Statistics, 1995; 4: 12-35.

[12] Liu, Q. and Pierce, D. A., A note on Guass-Hermite quadrature. Biometrika, 1994; 81: 624-629.

[13] Searle, S. R. Linear Models. Classic Edition. John Wiley & Sons. New York. (Reprinted from 1971), 1997.

[14] Searle, S. R., Casella, G. and McCulloch, C.E., Variance Components. John Wiley & Sons. New York, 1992.

[15] McCulloch, C.E. and Searle, S., Generalized, Linear, and Mixed Models. John Wiley & Sons. New York, 2001.

[16] Kuss, O. and McLerran, D., A note on the estimation of the multinomial logistic model with correlated response in SAS. Computer Method and Programs in Medicine, 2007; 262-269.

[17] Agresti, A., Categorical Data Analysis. 2nd ed., John Wiley & Sons. New York, 2002.

[18] Breslow N. E. and Clayton, D. G., Approximate inference in generalized linear mixed model. Journal of the American Statistical Association, 1993; 88 No 421: 9-25.

[19] Breslow, N. E. and Lin, X. Bias correlation in generalized linear mixed models with a single component of dispersion. Biometrika, 1995; 82: 81-91.

[20] Lipsitz, S.R., KIM, K. and Zhao, L., Analysis of repeated categorical data using generalized estimating equations. Statist. Medic, 1994; 13: 1149-1163.

[21] Liang, K. Y. and Zeger, S. L., Inference based on estimating functions in the presence of nuisance parameters. Statist. Sci. 1995; 10:158-173.

[22] Spiess, M. and Hamerle, A, A comparison of different methods for the estimation of regression models with correlated binary responses. Computational Statistics and Data Analysis, 2000; 33: 439-455.