GENERALIZED QUASILIKELIHOOD METHOD FOR MISCLASSIFIED LONGITUDINAL BINARY DATA

YI TAO

Generalized Quasilikelihood Method for Misclassified Longitudinal Binary Data

by

© Yi Tao

B. Science (Memorial University)

A practicum submitted to the

School of Graduate Studies

in partial fulfilment of the

requirements for the degree of

Master of Applied Statistics

Department of Mathematics and Statistics

Memorial University

June 30, 2010

St. John's Newfoundland and Labrador

Abstract

In this practicum we develop the generalized quasi-likelihood approach to analyzing longitudinal binary data with misclassification in response. We utilize the method of Monahan and Stefanski (1992) to approximate the expectation of an unknown function involved in the calculation of the means and covariances, which are further used to develop GQL estimating functions. The results of an intensive simulation study show that the proposed method works very well in all the preselected settings. The efficiency gain as compared to the naive method is remarkable. The method is robust in the sense that the performance varies just slightly when model parameters change in the simulation.

Keywords: Logit link; longitudinal binary response; GQL; Misclassification.

Acknowledgements

I would like to thank my supervisors, Zhaozhi Fan and Hong Wang , for their guidance, supervision and encouragement during my Master program, which was supported by the Department of Mathematics and Statistics, School of Graduate Studies, and who's research grants, at Memorial University.

Many thanks are due to Professors J. Concepcion Loredo-Osti, Gary Sneddon, Brajendra Sutradhar, Veeresh Gadag, who either taught me courses or gave many valuable helps for my studying and living.

My gratitude goes to all follow students: Yunqi Ji, Min Chen, Fang Fang, Qiong Li, Bingrui Sun, for their friendship. Especially Yunqi Ji, with whom I had many inspiring discussions.

Special thanks go to my family: my parents. Their support has provided me the strength to keep myself going.

Contents

List of tables

Chapter 1

Introduction

Longitudinal data are those which are collected from the same individuals over time. In biological or medical sciences, longitudinal data are accrued when the behavior of individuals over time is of interest. There is a considerable demand for adequate methods for the evaluation of data of this type in applications, especially when the underlying studies create binary responses.

The misclassification problem in studies with binary response is found to be of considerable interest among researchers recently. For example, Carroll, Spiegelman, Lan, Bailey, and Abbott (1984) deliberated measurement error problems in the Framingham Heart study; Burr (1988) investigated misclassification in quantal bioassay, among others. In practice, the misclassification of the observations happens frequently due to technical or fiscal reasons. However, we usually need to assume that the observations are error-free during the data analysis. Ignoring the existing misclassification can cause biased estimation of the model parameters, see Neuhaus (1999) and the references therein for estimation bias of covariates effects caused by ignoring measurement errors.

The literature on misclassification and measurement errors is abundant. Rosychuk (1999) studied the estimating bias of the covariate effects when ignoring errors in responses. Magder and Hughes (1997) proposed an EM approach to the inference of model effects. The method of Carroll, Maca, and Ruppert (1999) handles complex models with a simple measurement error structure. Neuhaus (1999) proposed a computationally more efficient approximation to accommodate measurement error when estimating the model effects. Roy et al. (2009) proposed a model-based approach to the case of misclassification. Based on approximated joint probabilities, a minimum Pearson's chi-square type approach was developed to deal with semi-Markov data by Rosychuk and Thompson (2001). Gustafson (2003), McGlothlin et al. (2008), and Rosychuk et al. (2009) also gave some different approaches to correct the bias of the estimates. Moreover, the simulation extrapolation (SIMEX) method is also becoming more and more popular in misclassification study (Cook and Stefanski, 1994; Carroll, Kchenhoff, Mwalili, and Lesaffre, 2006).

For correlated such as longitudinal binary response data, the misclassification problem is even harder to handle, mainly due to the complex nature of the likelihood when complex correlation structure is involved. Ji and Fan (2010) investigated the GQL and MLE approaches for a kind of dynamical binary response model, taking misclassification into account. The simulation results therein show remarkable efficiency gain by appropriately modeling the misclassification.

In this practicum, we develop a new approach that combines the method of Monahan and Stefanski (1992) (see also Roy, Baneriee, and Maiti, 2005) and the generalized quasi-likelihood approach (Sutradhar, 2003). An approximate generalized quasi-likelihood (GQL) method is proposed to correct the estimation bias of the regression parameters in the presence of classification error in binary responses.

In the logistic mixed model, the likelihood function is very hard to calculate. To avoid the complexity of the likelihood function we apply the GQL method, which is proved to be almost as efficient as MLE for binary data (Farrell and Sutradhar, 2007). Furthermore, the conditional inference leads to a loss of efficiency of statistical inference about regression parameters. The unconditional expectation is also very hard to calculate. The integration over the random effects is difficult for the logistic case, especially for multivariate random effects. We focus on unconditional generalized quasi-likelihood inference that involves unconditional moments of up to order 2. The integrations are approximated by using the method of Monahan and Stefanski (1992). By doing so we can avoid any extra distributional assumptions on the correlated binary responses, as well as the mechanism of the misclassification. The method is hence widely applicable.

Through out this practicum, we use T to denote the longitudinal binary response without error, and Y the observed corresponding binary response with error. The probabilities of misclassification denoted by ε_0 and ε_1 are defined as:

$$\varepsilon_0 = P(Y = 1|T = 0) \quad (1.1)$$

$$\varepsilon_1 = P(Y = 0|T = 1), \quad (1.2)$$

where 1 denotes a positive while 0 denotes a negative outcome. As an example in medical studies, a negative result means that a person is disease free, and a positive result means that the person has the disease. Furthermore, ε_0 denotes the probability of misdiagnosing healthy person, and ε_1 expresses probability of misdiagnosing a non-healthy person. The terms of sensitivity and specificity which are widely used in medical practice, are defined as below:

$$sensitivity = 1 - type\ II\ error = 1 - \varepsilon_1 \quad (1.3)$$

$$specificity = 1 - type\ I\ error = 1 - \varepsilon_0. \quad (1.4)$$

The remaining parts of this practicum are organized as follows.

The misclassification model and longitudinal model are introduced in Chapter 2. In Chapter 3, we summarize the logistic regression with Logit link and Probit link. We introduce the covariance matrix and corrected generalized quasi-likelihood (C-GQL) method in Chapter 4. The work here extends and unifies previous work by generalized quasi-likelihood (GQL)(Sutradhar, 2003) and Logit link approaches (Roy, Banerjee, and Maiti, 2005). The emphasis of this practicum is the bias correction of model effects when estimating functions are constructed based on quasi-likelihood.

The calculation and approximation are provided in Chapter 4. Simulation studies in Chapters 5 and 6 is conducted to investigate the performance of the proposed method when the sample size is small. We conclude this practicum with a few remarks in Chapter 7.

Chapter 2

The Models

In this practicum, we use T_{ij} to denote the true binary response without error, and Y_{ij} to denote the corresponding observed binary response with misclassification for the ith individual and jth time, where $i = 1, ..., I$ and $j = 1, ..., J$. The covariates, denoted by x_{ij}, may be time dependent, and is of dimension $p \times 1$.

Usually, the methods to analyze longitudinal binary data are based on the assumption that observed binary response Y_{ij} has no observation error. But in practice Y_{ij} is often prone to classification error. In this paper we define the relationship between Y_{ij} and T_{ij} in term of ε_0 and ε_1 given in (2.1).

The rates of misclassification are usually unknown in practice. But one can frequently obtain a small validation sample and hence get reasonably good estimates of them. Or some similar studies are available and an estimate of the unknown rate can be obtained from there.

We use η to denote the mean of T, the binary response without error, and μ the mean of Y, the binary response with error. Let γ_i be the random effect of each ith individual, which follows $N(0, \sigma^2_{\gamma_i})$. Moreover, η^c and μ^c denote the conditional mean of T and Y, given the random effects, respectively. The conditional expectation of T_{ij} given the random effect for subject i is

$$\eta^c_{ij} = P(T_{ij} = 1|\gamma_i),$$

and the unconditional expectation of T_{ij} is

$$\eta_{ij} = P(T_{ij} = 1) = E(\eta^c_{ij}).$$

The conditional expectation of Y_{ij} given the random effect for subject i is

$$\mu^c_{ij} = P(Y_{ij} = 1|\gamma_i).$$

Similarly the unconditional expectation of Y_{ij} is

$$\mu_{ij} = P(Y_{ij} = 1) = E(\mu^c_{ij}).$$

We describe the relationship between the observed response Y_{ij} and the inherent response T_{ij} by the misclassification model (Ji and Fan, 2010):

$$Y_{ij} = (1 - \varepsilon_1) * T_{ij} + \varepsilon_0 * (1 - T_{ij}), \tag{2.1}$$

where T_{ij} is a Bernoulli random variable. Further, the operation $*$ is called the binomial thinning operation. It is defined as $\rho * T = \sum_{k=1}^{T} b_k(\rho)$ with $\sum_{k=1}^{0} b_k(\rho) = 0$, and $\{b_k(\rho)\}$ is a sequence of independently and identically distributed binary random variables with $P(b_k(\rho) = 1) = \rho$ (McKenzie, 1988).

The logistic random effect longitudinal model is defined as

$$\begin{aligned} \eta_{ij}^c &= P(T_{ij} = 1 \mid \gamma_i) \\ &= \frac{exp(x'_{ij}\beta + \gamma_i)}{1 + exp(x'_{ij}\beta + \gamma_i)} = g(x'_{ij}\beta + \gamma_i). \end{aligned} \tag{2.2}$$

In this practicum, the parameters of scientific interest include both the regression coefficients β and the random effect γ_i.

Based on the misclassification model (2.1), we can find the unconditional expectation of Y:

$$\begin{aligned} \mu_{ij} &= P(Y_{ij} = 1) \\ &= E(Y_{ij}) \\ &= E((1 - \varepsilon_1) * T_{ij} + \varepsilon_0 * (1 - T_{ij})) \\ &= (1 - \varepsilon_1)E(T_{ij}) + \varepsilon_0(1 - E(T_{ij})) \\ &= E(T_{ij}) - \varepsilon_1 E(T_{ij}) + \varepsilon_0 - \varepsilon_0 E(T_{ij}) \\ &= \varepsilon_0 + (1 - \varepsilon_0 - \varepsilon_1)E(T_{ij}) \\ &= \varepsilon_0 + (1 - \varepsilon_0 - \varepsilon_1)\eta_{ij}, \end{aligned} \tag{2.3}$$

similarly we can present the conditional expectation of Y through the following equations:

$$\begin{aligned}\mu^c_{ij} &= P(Y_{ij}=1|\gamma_i)\\ &= \varepsilon_0 + (1-\varepsilon_0-\varepsilon_1)\eta^c_{ij}. \end{aligned} \tag{2.4}$$

As far as η_{ij} is concerned:

$$\begin{aligned}\eta_{ij} &= E(T_{ij})\\ &= E(\eta^c_{ij})\\ &= E(g(x'_{ij}\beta+\gamma_i))\\ &= \int_{-\infty}^{+\infty} g(x'_{ij}\beta+\gamma_i)f(\gamma_i)d\gamma_i. \end{aligned} \tag{2.5}$$

Therefore, by the equations (2.4) and (2.5), the unconditional expectation μ_{ij} is:

$$\begin{aligned}\mu_{ij} &= P(Y_{ij}=1)\\ &= E(\mu^c_{ji})\\ &= E(\varepsilon_0 + (1-\varepsilon_0-\varepsilon_1)\eta^c_{ij})\\ &= \int_{-\infty}^{+\infty}(\varepsilon_0+(1-\varepsilon_0-\varepsilon_1)g(x'_{ij}\beta+\gamma_i))f(\gamma_i)d\gamma_i\\ &= \varepsilon_0+(1-\varepsilon_0-\varepsilon_1)\int_{-\infty}^{+\infty} g(x'_{ij}\beta+\gamma_i)f(\gamma_i)d\gamma_i, \end{aligned} \tag{2.6}$$

where $\int_{-\infty}^{+\infty} g(x'_{ij}\beta+\gamma_i))f(\gamma_i)d\gamma_i = E(\eta^c_{ij}) = \eta_{ij} = E(T_{ij})$.

In (2.5), the link function $g(.)$ is nonlinear, which makes it difficult to calculate the integral over the density of γ_i. Roy (2005) discussed for the Probit link and logit link functions to deal with the approximation of the integral, that we will present in detail in Chapter 3.

Chapter 3

Link Functions

Roy, Banerjee, and Maiti(2005) applied a method of approximation for probit link and logit link functions to deal with the misclassified binary data based on a mixed model. In this section we introduce our approaches.

Let $g(.)$ be the logit link and Φ be the distribution function of a standard normal variable. Suppose G is a function of g, then we have:

$$\begin{aligned} g(x'_{ij}\beta+\gamma_i) = \tfrac{exp(x'_{ij}\beta+\gamma_i)}{1+exp(x'_{ij}\beta+\gamma_i)} & \quad function\ A \\ \overset{one\ to\ one}{\longleftrightarrow} \quad x'_{ij}\beta+\gamma_i & \quad function\ B \\ \overset{one\ to\ one}{\longleftrightarrow} \quad \Phi(x'_{ij}\beta+\gamma_i) & \overset{one\ to\ one}{\longleftrightarrow} G(\Phi). \end{aligned}$$

Function B can be rewritten as function A therefore,

$$g(x'_{ij}\beta+\gamma_i) = G(\Phi(x'_{ij}\beta+\gamma_i)) = \frac{exp(x'_{ij}\beta+\gamma_i)}{1+exp(x'_{ij}\beta+\gamma_i)}. \tag{3.1}$$

3.1 Probit Link Function

Let Φ be the cumulative distribution function and ϕ the probability density function of $N(0,1)$. We calculate $\int_{-\infty}^{+\infty}\Phi(x'_{ij}\beta+\gamma_i)f(\gamma_i)d\gamma_i$, where $\gamma_i \sim N(0,\sigma^2_{\gamma_i})$ and $f(\gamma_i)=\frac{1}{\sigma_{\gamma_i}\sqrt{2\pi}}exp(-\frac{\gamma_i^2}{2\sigma^2_{\gamma_i}})$:

$$\begin{aligned}
&\int_{-\infty}^{+\infty}\Phi(x'_{ij}\beta+\gamma_i)f(\gamma_i)d\gamma_i \\
=& \int_{-\infty}^{+\infty}\Phi(x'_{ij}\beta+\gamma_i)\frac{1}{\sigma_{\gamma_i}\sqrt{2\pi}}exp(-\frac{\gamma_i^2}{2\sigma^2_{\gamma_i}})d\gamma_i \\
=& \int_{-\infty}^{+\infty}\int_{-\infty}^{x'_{ij}\beta+\gamma_i}\frac{1}{\sqrt{2\pi}}exp(-\frac{z^2}{2})dz\frac{1}{\sigma_{\gamma_i}\sqrt{2\pi}}exp(-\frac{\gamma_i^2}{2\sigma^2_{\gamma_i}})d\gamma_i. \qquad (3.2)
\end{aligned}$$

Equation (3.2) can be viewed as an integral over a bivariate normal distribution, thus we can write it as a probability based on a bivariate normal distribution

$$\begin{aligned}
&= P(-\infty<\gamma_i<+\infty\ ,\ z<x'_{ij}\beta+\gamma_i) \\
&= P(z<x'_{ij}\beta+\gamma_i) \\
&= P(z-\gamma_i<x'_{ij}\beta).
\end{aligned}$$

Where $z\sim N(0,1)$ and $\gamma_i\sim N(0,\sigma^2_{\gamma_i})$, hence $z-\gamma_i\sim N(0-0,\sqrt{1+\sigma^2_{\gamma_i}})$ and

$$\begin{aligned}
&P(z-\gamma_i<x'_{ij}\beta) \\
=& \Phi(\frac{x'_{ij}\beta-0}{\sqrt{1+\sigma^2_{\gamma_i}}}) \\
=& \Phi(\frac{x'_{ij}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) \\
=& \Phi(x'_{ij}\frac{\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}).
\end{aligned}$$

Thus we have the equation:

$$\int_{-\infty}^{+\infty} \Phi(x'_{ij}\beta + \gamma_i) f(\gamma_i) d\gamma_i = \Phi(x'_{ij} \frac{\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}). \tag{3.3}$$

3.2 Logit Link Function

We denote $\Phi = \Phi(x'_{ij}\beta + \gamma_i)$ and $\Phi_0 = \Phi(x'_{ij}\beta)$. Therefore, by Taylor expansion we have:

$$G(\Phi) \approx G(\Phi_0) + G'(\Phi)|_{\Phi=\Phi_0}(\Phi - \Phi_0). \tag{3.4}$$

When the logit link is used in the generalized linear mixed model, we have

$$\begin{aligned}
& E(T_{ij}) = \eta_{ij} = E(\eta_{ij}^c) \\
= & \int_{-\infty}^{+\infty} g(x'_{ij}\beta + \gamma_i) f(\gamma_i) d\gamma_i \\
= & \int_{-\infty}^{+\infty} G(\Phi(x'_{ij}\beta + \gamma_i)) f(\gamma_i) d\gamma_i \\
\approx & \int_{-\infty}^{+\infty} (G(\Phi_0) + G'(\Phi)|_{\Phi=\Phi_0}(\Phi - \Phi_0)) f(\gamma_i) d\gamma_i \\
= & G(\Phi_0) + G'(\Phi)|_{\Phi=\Phi_0} \int_{-\infty}^{+\infty} \Phi f(\gamma_i) d\gamma_i - G'(\Phi)|_{\Phi=\Phi_0}\Phi_0 \\
= & G(\Phi_0) + G'(\Phi)|_{\Phi=\Phi_0}\Phi - G'(\Phi)|_{\Phi=\Phi_0}\Phi_0 \\
= & G(\Phi_0) + G'(\Phi)|_{\Phi=\Phi_0}(\Phi - \Phi_0) \\
\approx & G(\Phi) \\
= & G(\Phi(\gamma', W(z))) \\
= & \Phi(\frac{x'_{ij}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) \\
\approx & \frac{exp(\frac{x'_{ij}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})}{1 + exp(\frac{x'_{ij}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})}. \qquad (3.5)
\end{aligned}$$

The last approximation is due to Monahan and Stefanski(1992), where $k^2 = 1.7$.

In principle, the expectation of Y can be well approximated by:

$$\begin{aligned}
\mu_{ij} &= \varepsilon_0 + (1 - \varepsilon_0 - \varepsilon_1) \int_{-\infty}^{+\infty} g(x'_{ij}\beta + \gamma_i) f(\gamma_i) d\gamma_i \\
&\approx \varepsilon_0 + (1 - \varepsilon_0 - \varepsilon_1) \frac{exp(\frac{x'_{ij}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})}{1 + exp(\frac{x'_{ij}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})}. \qquad (3.6)
\end{aligned}$$

In the next chapter we will use the Generalized Quasi-likelihood (GQL) method to estimate the regression coefficients. The method needs μ_{ij} and the second ordered moments to formulate the estimation equations.

Chapter 4

Estimation Method

To correct the bias of the estimates of regression coefficients β due to ignoring the classification error, we apply a corrected generalized quasi-likelihood method(CGQL). We assume misclassification rates ε_0 and ε_1 are available during the estimation procedure. The challenging part of this chapter is the unconditional generalized quasi-likelihood inference which involves unconditional moments of up to second order.

The aim of this practicum is to compare the corrected generalized quasi-likelihood (CGQL) estimates with the estimates from the naive generalized quasi-likelihood (NGQL) approach. Whereas in the case of NGQL, the estimates of regression coefficients β are based on the GQL method which ignores ε_0 and ε_1.

4.1 Generalized Quasi-likelihood Method

Let $y_i = (y_{i1}, \ldots, y_{im})'$ denote the longitudinal observations of the binary response with error. Let $\mu_i = (\mu_{i1}, \ldots, \mu_{im})'$ be the vector of the unconditional expectations of y_i. Let Σ_i denote the $m \times m$ covariance matrix of Y_i. Sutradhar (2003) proposed a Generalized quasi-likelihood method to estimate the regression coefficients β by solving the estimating equations:

$$\sum_{i=1}^{n} \frac{\partial \mu_i'}{\partial \beta} \Sigma_i^{-1}(y_i - \mu_i) = 0, \tag{4.1}$$

where $\frac{\partial \mu_i'}{\partial \beta}$ is the $(p+1) \times m$ first derivative matrix of unconditional expectation of the longitudinal observations. Σ_i is the covariance of the binary response Y_i (see Section 4.3). Moreover, the ijth element for the first order derivatives of μ_{ij} is given by:

$$\begin{aligned}\frac{\partial \mu_{ij}}{\partial \beta} &\approx \frac{\partial}{\partial \beta}\left(\varepsilon_0 + (1 - \varepsilon_0 - \varepsilon_1) \frac{exp(\frac{x_{ij}'\beta}{\sqrt{1+\sigma_{\gamma_i}^2/k^2}})}{1 + exp(\frac{x_{ij}'\beta}{\sqrt{1+\sigma_{\gamma_i}^2/k^2}})} \right) \\ &\approx (1 - \varepsilon_0 - \varepsilon_1) \frac{exp(\frac{x_{ij}'\beta}{\sqrt{1+\sigma_{\gamma_i}^2/k^2}}) \frac{x_{ij}'\beta}{\sqrt{1+\sigma_{\gamma_i}^2/k^2}}}{(1 + exp(\frac{x_{ij}'\beta}{\sqrt{1+\sigma_{\gamma_i}^2/k^2}}))^2}.\end{aligned}$$

Newton-Raphson method has been applied to solve the estimating equation:

$$\hat{\beta}_i = \hat{\beta}_{i-1} + COV^{-1} FUN|_{\beta=\hat{\beta}_{i-1}}, \tag{4.2}$$

where FUN denotes the estimating function, and COV denotes the variance-covariance matrix. In practice, $cov(\hat{\beta}_{CGQL})$ can be estimated by

$$COV^{-1} = (\sum_{i=1}^{n} \frac{\partial \mu_i'}{\partial \beta} \Sigma_i^{-1} \frac{\partial \mu_i}{\partial \beta})^{-1}|_{\beta=\hat{\beta}_{CGQL}}, \tag{4.3}$$

where CGQL is the notation of our method that we will describe in detail in the following subsections.

4.2 Construction of the second order moments

In Section 4.1 we have constructed the estimating equations. The covariance of the binary response Y_i has its (u,v)th element as

$$cov(Y_{iu}, Y_{iv}) = E(Y_{iu}Y_{iv}) - \mu_{iu}\mu_{iv}. \quad (4.4)$$

We estimate the regression coefficients using the NGQL method by solving the estimating equation as following

$$\sum_{i=1}^{n} \frac{\partial \eta_i'}{\partial \beta} \Sigma_1^{-1}(y_i - \eta_i) = 0. \quad (4.5)$$

How to use Σ_1 in (4.5)? $E(T_{iu}T_{iv})$ need be calculated. Where $\Sigma_1 = cov(T_{iu}, T_{iv}) = E(T_{iu}T_{iv}) - \eta_{iu}\eta_{iv}$ and $\varepsilon_0 + (1-\varepsilon_0-\varepsilon_1)E(T_{iu}T_{iv}) = E(Y_{iu}Y_{iv})$. Whereas, the estimating equation of the CGQL method is

$$\sum_{i=1}^{n} \frac{\partial \mu_i'}{\partial \beta} \Sigma_2^{-1}(y_i - \mu_i) = 0, \quad (4.6)$$

where $\Sigma_2 = cov(Y_{iu}, Y_{iv}) = (1-\varepsilon_0-\varepsilon_1)^2 cov(T_{iu}, T_{iv})$ (Neuhaus, 2002). The second order moment $E(Y_{iu}Y_{iv})$ is calculated below.

The issue of approximating the second order moment $E(Y_{iu}Y_{iv})$ is considered by Monahan and Stefanski(1992). In Chapter 3 we have,

$$\begin{aligned} \eta_{ij} = E(\eta_{ij}^c) &= \int_{-\infty}^{+\infty} g(x_{ij}' + \gamma_i) f(\gamma_i) d\gamma_i \\ &\approx \frac{exp(\frac{x_{ij}'\beta}{\sqrt{1+\sigma_{\gamma_i}^2/k^2}})}{1 + exp(\frac{x_{ij}'\beta}{\sqrt{1+\sigma_{\gamma_i}^2/k^2}})}. \end{aligned} \quad (4.7)$$

Let x_{iu} denotes the covariates of the ith individual at uth time, and x_{iv} denote the covariates of the ith individual at vth time. From (3.6) we have the unconditional expectations μ_{iu} and μ_{iv} as:

$$\begin{aligned}
\mu_{iu} &= \varepsilon_0 + (1-\varepsilon_0-\varepsilon_1)\eta_{iu} \\
&\approx \varepsilon_0 + (1-\varepsilon_0-\varepsilon_1)\frac{exp(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})}{1+exp(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})} \qquad (4.8) \\
\mu_{iv} &= \varepsilon_0 + (1-\varepsilon_0-\varepsilon_1)\eta_{iv} \\
&\approx \varepsilon_0 + (1-\varepsilon_0-\varepsilon_1)\frac{exp(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})}{1+exp(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})}. \qquad (4.9)
\end{aligned}$$

The first step of finding the covariance of Y_{iu} and Y_{iv} is to calculate the second order moment:

$$
\begin{aligned}
& E(Y_{iu}Y_{iv}) \\
= \; & E(E(Y_{iu}Y_{iv}|\gamma_i)) \\
= \; & \int_{-\infty}^{+\infty} \mu_{iu}^c \mu_{iv}^c f(\gamma_i) d\gamma_i \\
= \; & \int_{-\infty}^{+\infty} (\varepsilon_0 + (1-\varepsilon_0-\varepsilon_1)\frac{exp(x_{iv}'\beta+\gamma_i)}{1+exp(x_{iv}'\beta+\gamma_i)}) \\
& (\varepsilon_0 + (1-\varepsilon_0-\varepsilon_1)\frac{exp(x_{iu}'\beta+\gamma_i)}{1+exp(x_{iu}'\beta+\gamma_i)}) f(\gamma_i) d\gamma_i \\
= \; & \int_{-\infty}^{+\infty} (\; \varepsilon_0^2 f(\gamma_i) \\
& + \varepsilon_0(1-\varepsilon_0-\varepsilon_1) f(\gamma_i) \frac{exp(x_{iu}'\beta+\gamma_i)}{1+exp(x_{iu}'\beta+\gamma_i)} \\
& + \varepsilon_0(1-\varepsilon_0-\varepsilon_1) f(\gamma_i) \frac{exp(x_{iv}'\beta+\gamma_i)}{1+exp(x_{iv}'\beta+\gamma_i)} \\
& + \; (1-\varepsilon_0-\varepsilon_1)^2 f(\gamma_i) \frac{exp(x_{iv}'\beta+\gamma_i)}{1+exp(x_{iv}'\beta+\gamma_i)} \\
& \frac{exp(x_{iu}'\beta+\gamma_i)}{1+exp(x_{iu}'\beta+\gamma_i)} \;) d\gamma_i \\
= \; & \int_{-\infty}^{+\infty} \varepsilon_0^2 f(\gamma_i) d\gamma_i \qquad \longrightarrow (1) \\
& + \int_{-\infty}^{+\infty} \varepsilon_0(1-\varepsilon_0-\varepsilon_1) f(\gamma_i) \frac{exp(x_{iu}'\beta+\gamma_i)}{1+exp(x_{iu}'\beta+\gamma_i)} d\gamma_i \qquad \longrightarrow (2) \\
& + \int_{-\infty}^{+\infty} \varepsilon_0(1-\varepsilon_0-\varepsilon_1) f(\gamma_i) \frac{exp(x_{iv}'\beta+\gamma_i)}{1+exp(x_{iv}'\beta+\gamma_i)} d\gamma_i \qquad \longrightarrow (3) \\
& + \int_{-\infty}^{+\infty} (1-\varepsilon_0-\varepsilon_1)^2 f(\gamma_i) \frac{exp(x_{iv}'\beta+\gamma_i)}{1+exp(x_{iv}'\beta+\gamma_i)} \\
& \qquad \frac{exp(x_{iu}'\beta+\gamma_i)}{1+exp(x_{iu}'\beta+\gamma_i)} d\gamma_i \qquad \longrightarrow (4).
\end{aligned}
$$

4.3 Computation of the covariance matrix

We now show how to compute the covariance matrix. First of all, we calculate the second order moment by parts.

Part 1:

$$\begin{aligned} & \int_{-\infty}^{+\infty} \varepsilon_0^2 f(\gamma_i) d\gamma_i \\ = & \ \varepsilon_0^2 \int_{-\infty}^{+\infty} f(\gamma_i) d\gamma_i \\ = & \ \varepsilon_0^2 = CM_1. \end{aligned} \tag{4.10}$$

Part 2:

$$\begin{aligned} & \int_{-\infty}^{+\infty} \varepsilon_0(1-\varepsilon_0-\varepsilon_1) f(\gamma_i) \frac{exp(x'_{iu}\beta+\gamma_i)}{1+exp(x'_{iu}\beta+\gamma_i)} d\gamma_i \\ = & \ \varepsilon_0(1-\varepsilon_0-\varepsilon_1) \int_{-\infty}^{+\infty} f(\gamma_i) \frac{exp(x'_{iu}\beta+\gamma_i)}{1+exp(x'_{iu}\beta+\gamma_i)} d\gamma_i \\ \approx & \ \varepsilon_0(1-\varepsilon_0-\varepsilon_1) \frac{exp(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})}{1+exp(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})} = CM_2. \end{aligned} \tag{4.11}$$

Part 3:

$$\begin{aligned} & \int_{-\infty}^{+\infty} \varepsilon_0(1-\varepsilon_0-\varepsilon_1) f(\gamma_i) \frac{exp(x'_{iv}\beta+\gamma_i)}{1+exp(x'_{iv}\beta+\gamma_i)} d\gamma_i \\ \approx & \ \varepsilon_0(1-\varepsilon_0-\varepsilon_1) \frac{exp(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})}{1+exp(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})} = CM_3. \end{aligned} \tag{4.12}$$

Part 4:

Similar to chapter 3, let $g(.)$ be the logit link and Φ be the distribution function of a standard normal variable. Suppose G is a function of g, and let $g(x'_{iu}\beta + \gamma_i) = \frac{exp(x'_{iu}\beta+\gamma_i)}{1+exp(x'_{iu}\beta+\gamma_i)}$ and $g(x'_{iv}\beta + \gamma_i) = \frac{exp(x'_{iv}\beta+\gamma_i)}{1+exp(x'_{iv}\beta+\gamma_i)}$ then we have:

$$\begin{aligned} g(x'_{iu}\beta + \gamma_i) = \frac{exp(x'_{iu}\beta + \gamma_i)}{1 + exp(x'_{iu}\beta + \gamma_i)} &\overset{one\ to\ one}{\longleftrightarrow} x'_{iu}\beta + \gamma_i \\ &\overset{one\ to\ one}{\longleftrightarrow} \Phi(x'_{iu}\beta + \gamma_i) \overset{one\ to\ one}{\longleftrightarrow} G(\Phi) \end{aligned}$$

and

$$\begin{aligned} g(x'_{iv}\beta + \gamma_i) = \frac{exp(x'_{iv}\beta + \gamma_i)}{1 + exp(x'_{iv}\beta + \gamma_i)} &\overset{one\ to\ one}{\longleftrightarrow} x'_{iv}\beta + \gamma_i \\ &\overset{one\ to\ one}{\longleftrightarrow} \Phi(x'_{iv}\beta + \gamma_i) \overset{one\ to\ one}{\longleftrightarrow} G(\Phi), \end{aligned}$$

thus

$$g(x'_{iu}\beta + \gamma_i) = G(\Phi(x'_{iu}\beta + \gamma_i)) = \frac{exp(x'_{iu}\beta + \gamma_i)}{1 + exp(x'_{iu}\beta + \gamma_i)} \tag{4.13}$$

$$g(x'_{iv}\beta + \gamma_i) = G(\Phi(x'_{iv}\beta + \gamma_i)) = \frac{exp(x'_{iv}\beta + \gamma_i)}{1 + exp(x'_{iv}\beta + \gamma_i)}. \tag{4.14}$$

Therefore

$$\begin{aligned} &\int_{-\infty}^{+\infty} (1 - \varepsilon_0 - \varepsilon_1)^2 f(\gamma_i) \frac{exp(x'_{iv}\beta + \gamma_i)}{1 + exp(x'_{iv}\beta + \gamma_i)} \frac{exp(x'_{iu}\beta + \gamma_i)}{1 + exp(x'_{iu}\beta + \gamma_i)} d\gamma_i \\ = \; &(1 - \varepsilon_0 - \varepsilon_1)^2 \int_{-\infty}^{+\infty} f(\gamma_i) \frac{exp(x'_{iv}\beta + \gamma_i)}{1 + exp(x'_{iv}\beta + \gamma_i)} \frac{exp(x'_{iu}\beta + \gamma_i)}{1 + exp(x'_{iu}\beta + \gamma_i)} d\gamma_i \\ = \; &(1 - \varepsilon_0 - \varepsilon_1)^2 \int_{-\infty}^{+\infty} f(\gamma_i) G(\Phi_1) G(\Phi_2) d\gamma_i. \end{aligned}$$

The integral in the above equation has no complete analytic form for the logit link. Its approximation can be obtained below.

Let Φ be the cumulative distribution function of a standard normal distribution. We denote $\Phi_1 = \Phi(x'_{iu}\beta + \gamma_i)$, $\Phi_{10} = \Phi(x'_{iu}\beta)$, $\Phi_2 = \Phi(x'_{iv}\beta + \gamma_i)$, and $\Phi_{20} = \Phi(x'_{iv}\beta)$:

$$\begin{aligned}
& \int_{-\infty}^{+\infty} f(\gamma_i) G(\Phi_1) G(\Phi_2) d\gamma_i \\
\approx & \int_{-\infty}^{+\infty} f(\gamma_i)(G(\Phi_{10}) + G'(\Phi_{10})(\Phi_1 - \Phi_{10})) \\
& \qquad (G(\Phi_{20}) + G'(\Phi_{20})(\Phi_2 - \Phi_{20})) d\gamma_i.
\end{aligned}$$

Where

$$\begin{aligned}
& \int_{-\infty}^{+\infty} f(\gamma_i)(G(\Phi_{10}) + G'(\Phi_{10})(\Phi_1 - \Phi_{10})) \\
& \qquad (G(\Phi_{20}) + G'(\Phi_{20})(\Phi_2 - \Phi_{20})) d\gamma_i \\
\approx & \int_{-\infty}^{+\infty} f(\gamma_i)(G(\Phi_{10})G(\Phi_{20}) + G(\Phi_{10})G'(\Phi_{20})\Phi_2 - G(\Phi_{10})G'(\Phi_{20})\Phi_{20} \\
& + G'(\Phi_{10})\Phi_1 G(\Phi_{20}) + G'(\Phi_{10})\Phi_1 G'(\Phi_{20})\Phi_2 - G'(\Phi_{10})\Phi_1 G'(\Phi_{20})\Phi_{20} \\
& - G'(\Phi_{10})\Phi_{10} G(\Phi_{20}) - G'(\Phi_{10})\Phi_{10} G'(\Phi_{20})\Phi_2 + G'(\Phi_{10})\Phi_{10} G'(\Phi_{20})\Phi_{20}) d\gamma_i \\
= & \; G(\Phi_{10})G(\Phi_{20}) \int_{-\infty}^{+\infty} f(\gamma_i) d\gamma_i + G(\Phi_{10})G'(\Phi_{20}) \int_{-\infty}^{+\infty} \Phi_2 f(\gamma_i) d\gamma_i \\
& - G(\Phi_{10})G'(\Phi_{20})\Phi_{20} \int_{-\infty}^{+\infty} f(\gamma_i) d\gamma_i \\
& + G'(\Phi_{10})G(\Phi_{20}) \int_{-\infty}^{+\infty} \Phi_1 f(\gamma_i) d\gamma_i + G'(\Phi_{10})G'(\Phi_{20}) \int_{-\infty}^{+\infty} \Phi_1 \Phi_2 f(\gamma_i) d\gamma_i \\
& - G'(\Phi_{10})G'(\Phi_{20})\Phi_{20} \int_{-\infty}^{+\infty} \Phi_1 f(\gamma_i) d\gamma_i \\
& - G'(\Phi_{10})\Phi_{10} G(\Phi_{20}) \int_{-\infty}^{+\infty} f(\gamma_i) d\gamma_i - G'(\Phi_{10})\Phi_{10} G'(\Phi_{20}) \int_{-\infty}^{+\infty} \Phi_2 f(\gamma_i) d\gamma_i \\
& + G'(\Phi_{10})\Phi_{10} G'(\Phi_{20})\Phi_{20} \int_{-\infty}^{+\infty} f(\gamma_i) d\gamma_i
\end{aligned}$$

By definition and (3.3) we know:

$$\begin{aligned}
\int_{-\infty}^{+\infty} f(\gamma_i)d\gamma_i &= 1 \\
\int_{-\infty}^{+\infty} \Phi_1 f(\gamma_i)d\gamma_i &= \Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma_{\gamma_i}^2}}) \\
\int_{-\infty}^{+\infty} \Phi_2 f(\gamma_i)d\gamma_i &= \Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma_{\gamma_i}^2}}).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\int_{-\infty}^{+\infty} f(\gamma_i)(G(\Phi_{10}) + G'(\Phi_{10})(\Phi_1 - \Phi_{10})) \\
&\qquad (G(\Phi_{20}) + G'(\Phi_{20})(\Phi_2 - \Phi_{20}))d\gamma_i \\
\approx\; & G(\Phi_{10})G(\Phi_{20}) + G(\Phi_{10})G'(\Phi_{20})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma_{\gamma_i}^2}}) \\
& -G(\Phi_{10})G'(\Phi_{20})\Phi_{20} \\
& +G'(\Phi_{10})G(\Phi_{20})\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma_{\gamma_i}^2}}) + G'(\Phi_{10})G'(\Phi_{20})\int_{-\infty}^{+\infty} \Phi_1\Phi_2 f(\gamma_i)d\gamma_i \\
& -G'(\Phi_{10})G'(\Phi_{20})\Phi_{20}\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma_{\gamma_i}^2}}) \\
& -G'(\Phi_{10})\Phi_{10}G(\Phi_{20}) - G'(\Phi_{10})\Phi_{10}G'(\Phi_{20})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma_{\gamma_i}^2}}) \\
& +G'(\Phi_{10})\Phi_{10}G'(\Phi_{20})\Phi_{20}
\end{aligned}$$

Then we add and subtract $G'(\Phi_{10})G'(\Phi_{20})\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})$:

$$
\begin{aligned}
&\int_{-\infty}^{+\infty} f(\gamma_i)(G(\Phi_{10}) + G'(\Phi_{10})(\Phi_1 - \Phi_{10})) \\
&\qquad (G(\Phi_{20}) + G'(\Phi_{20})(\Phi_2 - \Phi_{20}))d\gamma_i \\
\approx\ & G(\Phi_{10})G(\Phi_{20}) + G(\Phi_{10})G'(\Phi_{20})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) \\
& -G(\Phi_{10})G'(\Phi_{20})\Phi_{20} \\
& +G'(\Phi_{10})G(\Phi_{20})\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) + G'(\Phi_{10})G'(\Phi_{20})\int_{-\infty}^{+\infty}\Phi_1\Phi_2 f(\gamma_i)d\gamma_i \\
& -G'(\Phi_{10})G'(\Phi_{20})\Phi_{20}\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) \\
& -G'(\Phi_{10})\Phi_{10}G(\Phi_{20}) - G'(\Phi_{10})\Phi_{10}G'(\Phi_{20})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) \\
& +G'(\Phi_{10})\Phi_{10}G'(\Phi_{20})\Phi_{20} \\
& +G'(\Phi_{10})G'(\Phi_{20})\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) \\
& -G'(\Phi_{10})G'(\Phi_{20})\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) \\
\approx\ & (G(\Phi_{10}) + G'(\Phi_{10})\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) - G'(\Phi_{10})\Phi_{10}) \\
& (G(\Phi_{20}) + G'(\Phi_{20})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) - G'(\Phi_{20})\Phi_{20}) \\
& +G'(\Phi_{10})G'(\Phi_{20})(\int_{-\infty}^{+\infty}\Phi_1\Phi_2 f(\gamma_i)d\gamma_i - \Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})).
\end{aligned}
$$

Thus

$$
\begin{aligned}
&\int_{-\infty}^{+\infty} f(\gamma_i)G(\Phi_1)G(\Phi_2)d\gamma_i \\
\approx\ & (G(\Phi_{10}) + G'(\Phi_{10})\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) - G'(\Phi_{10})\Phi_{10}) \\
& (G(\Phi_{20}) + G'(\Phi_{20})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) - G'(\Phi_{20})\Phi_{20}) \\
& + G'(\Phi_{10})G'(\Phi_{20})(\int_{-\infty}^{+\infty} \Phi_1\Phi_2 f(\gamma_i)d\gamma_i - \Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}))
\end{aligned}
$$

Consider Taylor's Theorem we have:

$$
G(\Phi_{10}) + G'(\Phi_{10})(\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) - \Phi_{10}) = G(\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})) \tag{4.15}
$$

$$
G(\Phi_{20}) + G'(\Phi_{20})(\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) - \Phi_{20}) = G(\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})). \tag{4.16}
$$

From (4.15) and (4.16), we obtain:

$$
\begin{aligned}
&\int_{-\infty}^{+\infty} f(\gamma_i)G(\Phi_1)G(\Phi_2)d\gamma_i \\
\approx\ & (G(\Phi_{10}) + G'(\Phi_{10})\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) - G'(\Phi_{10})\Phi_{10}) \\
& (G(\Phi_{20}) + G'(\Phi_{20})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}) - G'(\Phi_{20})\Phi_{20}) \\
& + G'(\Phi_{10})G'(\Phi_{20})(\int_{-\infty}^{+\infty} \Phi_1\Phi_2 f(\gamma_i)d\gamma_i - \Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})) \\
\approx\ & G(\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}))G(\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})) \\
& + G'(\Phi_{10})G'(\Phi_{20})(\int_{-\infty}^{+\infty} \Phi_1\Phi_2 f(\gamma_i)d\gamma_i - \Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}))
\end{aligned}
$$

Use a similar method to chapter 3 to deal with functions Φ_1 and Φ_2, we have

$$\begin{aligned}
& \Phi_1(\frac{x_{iu}'\beta}{\sqrt{1+\sigma_{\gamma_i}^2}}) \\
= & \int_{-\infty}^{+\infty} \Phi_1(x_{iu}'\beta+\gamma_i)f(\gamma_i)d\gamma_i \\
= & \int_{-\infty}^{+\infty}\int_{-\infty}^{x_{iu}'\beta+\gamma_i} \frac{1}{\sqrt{2\pi}}exp(-\frac{z_1^2}{2})dz_1\frac{1}{\sigma_{\gamma_i}\sqrt{2\pi}}exp(-\frac{\gamma_i^2}{2\sigma_{\gamma_i}^2})d\gamma_i \\
= & P(-\infty<\gamma_i<+\infty\ ,\ z_1<x_{iu}'\beta+\gamma_i) \\
= & P(z_1-\gamma_i<x_{iu}'\beta),
\end{aligned}$$

and

$$\begin{aligned}
& \Phi_2(\frac{x_{iv}'\beta}{\sqrt{1+\sigma_{\gamma_i}^2}}) \\
= & \int_{-\infty}^{+\infty} \Phi_2(x_{iv}'\beta+\gamma_i)f(\gamma_i)d\gamma_i \\
= & \int_{-\infty}^{+\infty}\int_{-\infty}^{x_{iv}'\beta+\gamma_i} \frac{1}{\sqrt{2\pi}}exp(-\frac{z_2^2}{2})dz_2\frac{1}{\sigma_{\gamma_i}\sqrt{2\pi}}exp(-\frac{\gamma_i^2}{2\sigma_{\gamma_i}^2})d\gamma_i \\
= & P(-\infty<\gamma_i<+\infty\ ,\ z_2<x_{iv}'\beta+\gamma_i) \\
= & P(z_2-\gamma_i<x_{iv}'\beta).
\end{aligned}$$

Moreover,

$$
\begin{aligned}
& \int_{-\infty}^{+\infty} \Phi_1 \Phi_2 f(\gamma_i) d\gamma_i \\
= & \int_{-\infty}^{+\infty} \Phi_1 \Phi_2 \frac{1}{\sigma_{\gamma_i}\sqrt{2\pi}} exp(-\frac{\gamma_i^2}{2\sigma_{\gamma_i}^2}) d\gamma_i \\
= & \int_{-\infty}^{+\infty} \int_{-\infty}^{x_{iu}'\beta+\gamma_i} \frac{1}{\sqrt{2\pi}} exp(-\frac{z_1^2}{2}) dz_1 \int_{-\infty}^{x_{iv}'\beta+\gamma_i} \frac{1}{\sqrt{2\pi}} exp(-\frac{z_2^2}{2}) dz_2 \\
& \frac{1}{\sigma_{\gamma_i}\sqrt{2\pi}} exp(-\frac{\gamma_i^2}{2\sigma_{\gamma_i}^2}) d\gamma_i \\
= & P(-\infty < \gamma_i < +\infty,\ z_1 < x_{iu}'\beta + \gamma_i,\ z_2 < x_{iv}'\beta + \gamma_i) \\
= & P(z_1 - \gamma_i < x_{iu}'\beta,\ z_2 - \gamma_i < x_{iv}'\beta).
\end{aligned}
$$

Thus

$$
\begin{aligned}
& \int_{-\infty}^{+\infty} f(\gamma_i) G(\Phi_1) G(\Phi_2) d\gamma_i \\
\approx & G(\Phi_1(\frac{x_{iu}'\beta}{\sqrt{1+\sigma_{\gamma_i}^2}})) G(\Phi_2(\frac{x_{iv}'\beta}{\sqrt{1+\sigma_{\gamma_i}^2}})) \\
& + G'(\Phi_{10}) G'(\Phi_{20}) (P(z_1 - \gamma_i < x_{iu}'\beta,\ z_2 - \gamma_i < x_{iv}'\beta) \\
& - P(z_1 - \gamma_i < x_{iu}'\beta) P(z_2 - \gamma_i < x_{iv}'\beta))
\end{aligned}
\tag{4.17}
$$

Now $G'(\Phi_{20})$ need be calculated. Let $u = x_{iu}'\beta$, we find

$$
\begin{aligned}
G(\Phi_{10}) &= G(\Phi(x_{iu}'\beta)) = \frac{exp(x_{iu}'\beta)}{1+exp(x_{iu}'\beta)} \\
&= G(\Phi(u)) = \frac{exp(u)}{1+exp(u)}.
\end{aligned}
$$

Let $\alpha = \Phi(u)$ and $u = \Phi^{-1}(\alpha)$, we have

$$G(\alpha) = \frac{exp(\Phi^{-1}(\alpha))}{1 + exp(\Phi^{-1}(\alpha))} = \frac{exp(u)}{1 + exp(u)}.$$

In chapter 3 we have stated that Φ is the cumulative distribution function and ϕ is the probability density function of $N(0, 1)$. Then we obtain

$$\begin{aligned}
\frac{\partial G(\alpha)}{\partial \alpha} &= \frac{\partial \frac{exp(u)}{1+exp(u)}}{\partial u} \frac{\partial u}{\partial \alpha} \\
&= \frac{\partial \frac{exp(u)}{1+exp(u)}}{\partial u} \frac{\partial \Phi^{-1}(\alpha)}{\partial \alpha} \\
&= (\frac{exp(u)}{1 + exp(u)} - \frac{exp(u)exp(u)}{(1 + exp(u))^2}) \frac{\partial \Phi^{-1}(\alpha)}{\partial \alpha} \\
&= (\frac{exp(u)}{1 + exp(u)} - \frac{exp(u)exp(u)}{(1 + exp(u))^2}) \frac{1}{\frac{\partial \Phi(u)}{\partial u}} \\
& \quad \frac{\partial \Phi(u)}{\partial u} = \phi(u) = \frac{1}{\sqrt{2\pi}} exp(\frac{-u^2}{2}) \\
& \quad \frac{\partial \Phi^{-1}(\alpha)}{\partial \alpha} = \frac{1}{\frac{\partial \Phi(u)}{\partial u}} = \frac{1}{\frac{1}{\sqrt{2\pi}} exp(\frac{-u^2}{2})} \\
&= (\frac{exp(u)}{1 + exp(u)} - \frac{exp(u)exp(u)}{(1 + exp(u))^2}) \frac{1}{\frac{1}{\sqrt{2\pi}} exp(\frac{-u^2}{2})}.
\end{aligned}$$

This yields

$$\begin{aligned}
G'(\Phi_{10}) &= (\frac{exp(x'_{iu}\beta)}{1 + exp(x'_{iu}\beta)} - \frac{exp(x'_{iu}\beta)exp(x'_{iu}\beta)}{(1 + exp(x'_{iu}\beta))^2}) \frac{1}{\frac{1}{\sqrt{2\pi}} exp(\frac{-(x'_{iu}\beta)^2}{2})} \\
&= \frac{exp(x'_{iu}\beta)}{(1 + exp(x'_{iu}\beta))^2} (\frac{1}{\sqrt{2\pi}} exp(\frac{-(x'_{iu}\beta)^2}{2}))^{-1}.
\end{aligned}$$

Using the method from above for $G'(\Phi_{20})$yields

$$G'(\Phi_{20}) = \frac{exp(x'_{iv}\beta)}{(1 + exp(x'_{iv}\beta))^2} (\frac{1}{\sqrt{2\pi}} exp(\frac{-(x'_{iv}\beta)^2}{2}))^{-1}.$$

Hence

$$
\begin{aligned}
&\int_{-\infty}^{+\infty} f(\gamma_i)G(\Phi_1)G(\Phi_2)d\gamma_i \\
\approx\ & G(\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}}))G(\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}}})) \\
&+\frac{exp(x'_{iu}\beta)}{(1+exp(x'_{iu}\beta))^2}(\frac{1}{\sqrt{2\pi}}exp(\frac{-(x'_{iu}\beta)^2}{2}))^{-1} \\
&\frac{exp(x'_{iv}\beta)}{(1+exp(x'_{iv}\beta))^2}(\frac{1}{\sqrt{2\pi}}exp(\frac{-(x'_{iv}\beta)^2}{2}))^{-1} \\
&(P(z_1-\gamma_i<x'_{iu}\beta,\ z_2-\gamma_i<x'_{iv}\beta) \\
&-P(z_1-\gamma_i<x'_{iu}\beta)P(z_2-\gamma_i<x'_{iv}\beta)).
\end{aligned}
$$

Using the approximation exploited by Monahan and Stefanski(1992), this leads to

$$
\begin{aligned}
&\int_{-\infty}^{+\infty}(1-\varepsilon_0-\varepsilon_1)^2 f(\gamma_i)\frac{exp(x'_{iv}\beta+\gamma_i)}{1+exp(x'_{iv}\beta+\gamma_i)}\frac{exp(x'_{iu}\beta+\gamma_i)}{1+exp(x'_{iu}\beta+\gamma_i)}d\gamma_i \\
\approx\ & (1-\varepsilon_0-\varepsilon_1)^2 G(\Phi_1(\frac{x'_{iu}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}}))G(\Phi_2(\frac{x'_{iv}\beta}{\sqrt{1+\sigma^2_{\gamma_i}/k^2}})) \\
&+\frac{exp(x'_{iu}\beta)}{(1+exp(x'_{iu}\beta))^2}(\frac{1}{\sqrt{2\pi}}exp(\frac{-(x'_{iu}\beta)^2}{2}))^{-1} \\
&\frac{exp(x'_{iv}\beta)}{(1+exp(x'_{iv}\beta))^2}(\frac{1}{\sqrt{2\pi}}exp(\frac{-(x'_{iv}\beta)^2}{2}))^{-1} \\
&(P(z_1-\gamma_i<x'_{iu}\beta,\ z_2-\gamma_i<x'_{iv}\beta)-P(z_1-\gamma_i<x'_{iu}\beta)P(z_2-\gamma_i<x'_{iv}\beta)) \\
=\ & CM_4. \qquad (4.18)
\end{aligned}
$$

Summarizing part one to part four, we have the approximation of the uvth element of the covariance matrix:

$$
\begin{aligned}
& cov(Y_{iu}, Y_{iv}) \\
= \ & E(Y_{iu}Y_{iv}) - \mu_{iu}\mu_{iv} \\
\approx \ & (CM_1 + CM_2 + CM_3 + CM_4) - \mu_{iu}\mu_{iv}.
\end{aligned}
\tag{4.19}
$$

However, the diagonal element of the covariance matrix can be presented in different ways. We denote x_{it} as the covariates of the ith individual at the tth time. The tth diagonal of the covariance matrix can then be written as:

$$
\begin{aligned}
& var(Y_{it}) \\
= \ & E(Y_{it}^2) - \mu_{it}^2 \\
= \ & E(Y_{it}^2) - (E(Y_{it}))^2 \\
= \ & E(Y_{it}) - (E(Y_{it}))^2 \\
= \ & E(Y_{it})(1 - E(Y_{it})) \\
= \ & \mu_{it}(1 - \mu_{it}) \\
\approx \ & \varepsilon_0 + (1 - \varepsilon_0 - \varepsilon_1)\frac{exp(\frac{x'_{it}\beta}{\sqrt{1+\sigma_{\gamma_i}^2/k^2}})}{1 + exp(\frac{x'_{it}\beta}{\sqrt{1+\sigma_{\gamma_i}^2/k^2}})} \\
& -(\varepsilon_0 + (1 - \varepsilon_0 - \varepsilon_1)\frac{exp(\frac{x'_{it}\beta}{\sqrt{1+\sigma_{\gamma_i}^2/k^2}})}{1 + exp(\frac{x'_{it}\beta}{\sqrt{1+\sigma_{\gamma_i}^2/k^2}})})^2
\end{aligned}
\tag{4.20}
$$

Chapter 5

Simulation Study

In this chapter, we investigate the performance of our newly proposed CGQL method through simulation studies. the comparison between CGQL and NGQL is made at a variety of model settings, which reflect reasonably well many practical situations. We first provide the designs for our simulation, and then we discuss the simulation results. A brief discussion with conclude this chapter.

5.1 Designs

We performed $k = 500$ simulations each time with the sample size $I = 100, 500$, and 1000 respectively under the assumption that the response misclassification probabilities are known. Each independent individual has $m = 4$ repeated observations. The true parameter values are: the regression coefficients $\beta = (0.6, 0.7, 0.0)$ and $(0.4, 0.6, 0.2)$ for each set respectively; the misclassification rates for each simulation will be increase by degrees as $\alpha = (0.05, 0.05), (0.05, 0.10), (0.10, 0.05), (0.10, 0.10)$, (0.10, 0.20), (0.20, 0.10), and (0.20, 0.20); the variance of the random effect v_r will be 0.04 and 0.25 for each estimation of β. Finally, we generated 500 simulated data sets under the longitudinal model and the misclassification model that we have described in Chapter 2. The setup of the parameter values is presented as below:

$$\alpha \begin{cases} \beta = (0.6, 0.7, 0.0) \begin{cases} \sigma^2 = 0.05 \begin{cases} I = 100 \\ I = 500 \\ I = 1000 \end{cases} \\ \sigma^2 = 0.25 \begin{cases} I = 100 \\ I = 500 \\ I = 1000 \end{cases} \end{cases} \\ \beta = (0.4, 0.6, 0.2) \begin{cases} \sigma^2 = 0.05 \begin{cases} I = 100 \\ I = 500 \\ I = 1000 \end{cases} \\ \sigma^2 = 0.25 \begin{cases} I = 100 \\ I = 500 \\ I = 1000 \end{cases} \end{cases} \end{cases}$$

The probabilities of misclassification are denoted by ε_0 and ε_1 previously. In simulation studies the misclassification rates $\alpha = (\alpha_1, \alpha_2)$ where α_1 is type I error and α_2 is type II error. α_1 is the probability of observing $Y = 0$ when $T = 1$, and α_2 is the probability of observing $Y = 1$ when $T = 0$.

We use x_{mit} to denote the time-dependent covariate for the ith individual and the tth time, where $m = 1, 2, 3$. x_{1it} follows binomial distribution, x_{2it} follows normal distribution, and x_{3it} follows poisson distribution, to mimick the gender, body mass index, and the number of cigarette smoked for a specific individual. The covariates change over time and subject. It can be generated as follows

$$x_{1it} \sim \begin{cases} binomial(1,\ 0.5 - 0.1 * t) & for\ i = 1, ..., I/2 \\ binomial(1,\ 0.6 - 0.1 * t) & for\ i = I/2 + 1, ..., I, \end{cases}$$

$$x_{2it} \sim \begin{cases} normal(0.1 * (t-1),\ 1.5) & for\ i = 1, ..., I/2 \\ normal(0.1 * t,\ 1.5) & for\ i = I/2 + 1, ..., I, \end{cases}$$

and

$$x_{3it} \sim \begin{cases} poisson(4) & for\ i = 1, ..., I/2;\ t = 1 \\ poisson(18) & for\ i = 1, ..., I/2;\ t = 2 \\ poisson(22) & for\ i = 1, ..., I/2;\ t = 3 \\ poisson(30) & for\ i = 1, ..., I/2;\ t = 4 \\ \\ poisson(5) & for\ i = I/2 + 1, ..., I;\ t = 1 \\ poisson(20) & for\ i = I/2 + 1, ..., I;\ t = 2 \\ poisson(25) & for\ i = I/2 + 1, ..., I;\ t = 3 \\ poisson(35) & for\ i = I/2 + 1, ..., I;\ t = 4. \end{cases}$$

The values of T_{ij}'s are generated from model (2.2) with covariates x_1, x_2, x_3, and corresponding true values of the regression coefficients β. The observed data Y_{ij}'s are generated following (2.1).

5.2 Results

In this section, we examine the performance of CGQL and NGQL approaches in the estimation of β. Simulation studies were conducted for regression coefficient $\beta = (0.6, 0.7, 0.0)$ and $(0.4, 0.6, 0.2)$, misclassification rate $\alpha = (0.05, 0.05), ..., (0.20, 0.20)$, and variance of random effect $v_r = 0.04, 0.25$. sample sizes were chosen to be 100, 500, and 1000, separately. For each of the two estimation approaches we calculate the bias of the estimated β (Bias), simulated standard errors (SSE), estimated standard error(ESE), and coverage probabilities of 90% confidence interval (CPR). The simulation results are reported in the following tables.

Tables 5.1-5.3: $\beta = (0.6, 0.7, 0.0)$ and $vr = 0.04$.

Tables 5.4-5.6: $\beta = (0.6, 0.7, 0.0)$ and $vr = 0.25$.

Tables 5.7-5.9: $\beta = (0.4, 0.6, 0.2)$ and $vr = 0.04$.

Tables 5.10-5.12: $\beta = (0.4, 0.6, 0.2)$ and $vr = 0.25$.

Table 5.1: Estimation when $n = 100$, $\beta = (0.6, 0.7, 0.0)$, and the variance of the random effect $v_r = 0.04$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta 1$	$\beta 2$	$\beta 3$	$\beta 1$	$\beta 2$	$\beta 3$
(0.05,0.05)	Bias	0.0055	0.0080	-0.0041	-0.1104	-0.0854	-0.0002
	SSE	0.2498	0.1127	0.0497	0.2313	0.0951	0.0424
	ESE	0.2698	0.1127	0.0504	0.2305	0.0901	0.0438
	CPR	0.9200	0.9060	0.9240	0.8660	0.7160	0.9240
(0.05,0.10)	Bias	-0.0285	0.0271	-0.0015	-0.1717	-0.1352	-0.0399
	SSE	0.2958	0.1189	0.0539	0.2158	0.0767	0.0439
	ESE	0.3073	0.1156	0.0560	0.2325	0.0790	0.0447
	CPR	0.9140	0.9000	0.9160	0.8240	0.4560	0.7800
(0.10,0.05)	Bias	0.0004	0.0193	0.0104	-0.0977	-0.1287	0.0452
	SSE	0.2958	0.1204	0.0553	0.2334	0.0889	0.0450
	ESE	0.2921	0.1181	0.0558	0.2341	0.0830	0.0446
	CPR	0.9080	0.9040	0.9140	0.8620	0.5280	0.7160
(0.10,0.10)	Bias	0.0291	0.0112	0.0114	-0.1459	-0.1860	0.0063
	SSE	0.3398	0.1251	0.0571	0.2395	0.0818	0.0444
	ESE	0.3313	0.1285	0.0586	0.2353	0.0797	0.0433
	CPR	0.9100	0.9140	0.9060	0.8240	0.2500	0.8640
(0.10,0.20)	Bias	0.0521	0.0160	0.0005	-0.2810	-0.2732	-0.0693
	SSE	0.3846	0.1632	0.0658	0.2234	0.0812	0.0405
	ESE	0.3779	0.1582	0.0673	0.2215	0.0778	0.0430
	CPR	0.8920	0.9020	0.9000	0.6240	0.0500	0.5140
(0.20,0.10)	Bias	0.0048	0.0223	-0.0034	-0.1773	-0.2437	0.0671
	SSE	0.3916	0.1532	0.0657	0.2287	0.0805	0.0435
	ESE	0.3782	0.1500	0.0674	0.2362	0.0796	0.0429
	CPR	0.8900	0.9080	0.9220	0.8260	0.1000	0.5380
(0.20,0.20)	Bias	0.6513	0.7327	-0.0186	0.3393	0.3612	-0.0054
	SSE	0.4686	0.1981	0.0884	0.2194	0.0727	0.0418
	ESE	0.4407	0.1904	0.0827	0.2175	0.0736	0.0429
	CPR	0.8860	0.9200	0.8900	0.6480	0.0020	0.9080

Table 5.2: Estimation when $n = 500$, $\beta = (0.6, 0.7, 0.0)$, and the variance of the random effect $v_r = 0.04$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta1$	$\beta2$	$\beta3$	$\beta1$	$\beta2$	$\beta3$
(0.05,0.05)	Bias	0.0151	0.0069	0.0011	-0.0783	-0.0955	0.0016
	SSE	0.1178	0.0483	0.0221	0.0994	0.0387	0.0180
	ESE	0.1211	0.0490	0.0229	0.1027	0.0385	0.0198
	CPR	0.9120	0.9020	0.9200	0.8240	0.2160	0.9220
(0.05,0.10)	Bias	-0.0034	0.0010	0.0036	-0.1720	-0.1457	-0.0316
	SSE	0.1283	0.0510	0.0234	0.1033	0.0372	0.0185
	ESE	0.1322	0.0531	0.0242	0.1020	0.0372	0.0198
	CPR	0.9080	0.9010	0.9080	0.4860	0.0180	0.4900
(0.10,0.05)	Bias	0.0129	0.0065	-0.0016	-0.0794	-0.1313	0.0350
	SSE	0.1288	0.0510	0.0236	0.0998	0.0386	0.0190
	ESE	0.1335	0.0522	0.0242	0.1074	0.0380	0.0196
	CPR	0.9180	0.9140	0.9120	0.8460	0.0500	0.4520
(0.10,0.10)	Bias	0.0180	0.0096	-0.0045	-0.1507	-0.1800	-0.0026
	SSE	0.1536	0.0562	0.0259	0.0997	0.0357	0.0199
	ESE	0.1438	0.0571	0.0260	0.1033	0.0363	0.0194
	CPR	0.8920	0.8840	0.8880	0.5700	0.0020	0.8920
(0.10,0.20)	Bias	0.0178	0.0150	-0.0022	-0.2836	-0.2690	-0.0697
	SSE	0.1702	0.0720	0.0317	0.0961	0.0357	0.0188
	ESE	0.1638	0.0678	0.0304	0.0973	0.0340	0.0194
	CPR	0.8820	0.8840	0.8900	0.0980	0.0000	0.0240
(0.20,0.10)	Bias	0.0085	0.0035	0.0019	-0.1670	-0.2534	0.0732
	SSE	0.1618	0.0663	0.0287	0.1002	0.0362	0.0197
	ESE	0.1600	0.0666	0.0299	0.1025	0.0354	0.0193
	CPR	0.8900	0.9080	0.9060	0.4920	0.0000	0.0200
(0.20,0.20)	Bias	0.0039	-0.0017	0.0016	-0.2799	-0.3378	0.0012
	SSE	0.1971	0.0823	0.0369	0.1019	0.0341	0.0184
	ESE	0.1951	0.0811	0.0355	0.0993	0.0337	0.0190
	CPR	0.9140	0.8920	0.9060	0.1180	0.0000	0.9080

Table 5.3: Estimation when $n = 1000$, $\beta = (0.6, 0.7, 0.0)$, and the variance of the random effect $v_r = 0.04$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta1$	$\beta2$	$\beta3$	$\beta1$	$\beta2$	$\beta3$
(0.05,0.05)	Bias	0.0110	0.0008	0.0029	-0.0904	-0.0995	0.0046
	SSE	0.0872	0.0327	0.0162	0.0717	0.0268	0.0135
	ESE	0.0869	0.0343	0.0161	0.0738	0.0271	0.0139
	CPR	0.9040	0.9220	0.8880	0.6480	0.0280	0.8840
(0.05,0.10)	Bias	0.0057	-0.0006	-0.0020	-0.1492	-0.1489	-0.0392
	SSE	0.0891	0.0364	0.0165	0.0720	0.0259	0.0137
	ESE	0.0935	0.0367	0.0173	0.0720	0.0258	0.0139
	CPR	0.9180	0.9080	0.9080	0.3320	0.0000	0.1080
(0.10,0.05)	Bias	0.0211	0.0053	-0.0037	-0.0752	-0.1339	0.0338
	SSE	0.0951	0.0373	0.0166	0.0740	0.0254	0.0134
	ESE	0.0933	0.0372	0.0172	0.0751	0.0269	0.0139
	CPR	0.8980	0.9000	0.9020	0.7460	0.0000	0.2020
(0.10,0.10)	Bias	-0.0092	0.0013	-0.0002	-0.1680	-0.1879	0.0006
	SSE	0.1029	0.0390	0.0181	0.0709	0.0280	0.0139
	ESE	0.1003	0.0406	0.0184	0.0727	0.0258	0.0137
	CPR	0.8760	0.9200	0.9060	0.2540	0.0000	0.8920
(0.10,0.20)	Bias	-0.0027	0.0018	0.0037	-0.2930	-0.2704	-0.0661
	SSE	0.1147	0.0449	0.0215	0.0711	0.0255	0.0138
	ESE	0.1149	0.0473	0.0213	0.0690	0.0242	0.0137
	CPR	0.9060	0.9200	0.8880	0.0040	0.0000	0.0000
(0.20,0.10)	Bias	0.0001	0.0014	0.0010	-0.1650	-0.2518	0.0708
	SSE	0.1080	0.0474	0.0206	0.0694	0.0249	0.0135
	ESE	0.1119	0.0469	0.0212	0.0716	0.0250	0.0137
	CPR	0.9040	0.8920	0.9100	0.2640	0.0000	0.0000
(0.20,0.20)	Bias	0.0131	0.0024	-0.0019	-0.2876	-0.3376	-0.0009
	SSE	0.1356	0.0576	0.0259	0.0683	0.0243	0.0128
	ESE	0.1360	0.0569	0.0251	0.0692	0.0236	0.0134
	CPR	0.9020	0.8940	0.8880	0.0080	0.0000	0.9140

Table 5.4: Estimation when $n = 100$, $\beta = (0.6, 0.7, 0.0)$, and the variance of the random effect $v_r = 0.25$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta1$	$\beta2$	$\beta3$	$\beta1$	$\beta2$	$\beta3$
(0.05,0.05)	Bias	-0.0277	0.0252	-0.0009	-0.0591	-0.0620	0.0255
	SSE	0.2907	0.1187	0.0559	0.2657	0.1003	0.0451
	ESE	0.2960	0.1131	0.0567	0.2477	0.0949	0.0494
	CPR	0.8960	0.9080	0.9080	0.8700	0.7980	0.8840
(0.05,0.10)	Bias	0.0804	0.0444	-0.0082	-0.0977	-0.1100	-0.0454
	SSE	0.3412	0.1235	0.0591	0.2493	0.0895	0.0434
	ESE	0.3242	0.1219	0.0600	0.2468	0.0848	0.0486
	CPR	0.8820	0.8920	0.9040	0.8780	0.5960	0.7940
(0.10,0.05)	Bias	-0.0664	0.0296	0.0171	-0.1270	-0.1109	0.0594
	SSE	0.2930	0.1202	0.0586	0.2422	0.0892	0.0460
	ESE	0.2928	0.1260	0.0599	0.2494	0.0891	0.0498
	CPR	0.8880	0.9080	0.8820	0.8740	0.6240	0.6820
(0.10,0.10)	Bias	0.0431	0.0513	-0.0173	-0.1333	-0.1657	-0.0112
	SSE	0.3649	0.1442	0.0601	0.2361	0.0844	0.0446
	ESE	0.3570	0.1377	0.0621	0.2549	0.0862	0.0472
	CPR	0.8980	0.8940	0.9080	0.8740	0.4080	0.9020
(0.10,0.20)	Bias	-0.0436	0.0377	0.0146	-0.3272	-0.2507	-0.0671
	SSE	0.3940	0.1516	0.0738	0.2242	0.0859	0.0449
	ESE	0.3941	0.1595	0.0734	0.2364	0.0813	0.0480
	CPR	0.9020	0.9320	0.9120	0.5920	0.1020	0.6080
(0.20,0.10)	Bias	0.0659	0.0268	0.0101	-0.1064	-0.2419	0.0777
	SSE	0.4117	0.1573	0.0631	0.2522	0.0802	0.0425
	ESE	0.3980	0.1563	0.0699	0.2491	0.0810	0.0464
	CPR	0.9080	0.9180	0.9220	0.8540	0.1020	0.4920
(0.20,0.20)	Bias	0.1642	0.0585	-0.0062	-0.2274	-0.3323	-0.0010
	SSE	0.4651	0.2017	0.0870	0.2322	0.0785	0.0439
	ESE	0.4736	0.1952	0.0880	0.2306	0.0750	0.0470
	CPR	0.9120	0.9140	0.9160	0.7120	0.0120	0.9220

Table 5.5: Estimation when $n = 500$, $\beta = (0.6, 0.7, 0.0)$, and the variance of the random effect $v_r = 0.25$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta1$	$\beta2$	$\beta3$	$\beta1$	$\beta2$	$\beta3$
(0.05,0.05)	Bias	0.0174	0.0192	-0.0070	-0.0625	-0.0795	-0.0046
	SSE	0.1251	0.0502	0.0231	0.1103	0.0421	0.0209
	ESE	0.1285	0.0515	0.0250	0.1098	0.0412	0.0220
	CPR	0.9080	0.8920	0.9160	0.8540	0.3920	0.9160
(0.05,0.10)	Bias	0.0098	0.0027	0.0067	-0.1666	-0.1283	-0.0284
	SSE	0.1398	0.0549	0.0240	0.1048	0.0391	0.0206
	ESE	0.1402	0.0555	0.0264	0.1073	0.0392	0.0216
	CPR	0.8920	0.8890	0.9100	0.5240	0.0460	0.6240
(0.10,0.05)	Bias	-0.0243	0.0233	-0.0066	-0.1144	-0.1106	0.0399
	SSE	0.1419	0.0551	0.0256	0.1040	0.0389	0.0193
	ESE	0.1397	0.0544	0.0260	0.1113	0.0394	0.0216
	CPR	0.8900	0.8740	0.8960	0.7380	0.1320	0.3900
(0.10,0.10)	Bias	0.0240	0.0185	-0.0033	-0.1438	-0.1733	-0.0036
	SSE	0.1500	0.0562	0.0284	0.1111	0.0400	0.0191
	ESE	0.1486	0.0584	0.0281	0.1078	0.0372	0.0214
	CPR	0.9120	0.9100	0.9000	0.6260	0.0060	0.9360
(0.10,0.20)	Bias	0.0248	0.0152	-0.0171	-0.2574	-0.2598	-0.0823
	SSE	0.1722	0.0685	0.0315	0.1064	0.0365	0.0209
	ESE	0.1711	0.0701	0.0322	0.1041	0.0367	0.0214
	CPR	0.9040	0.9120	0.8540	0.2100	0.0000	0.0100
(0.20,0.10)	Bias	0.0371	0.0260	0.0104	-0.1505	-0.2401	0.0835
	SSE	0.1640	0.0659	0.0310	0.1098	0.0373	0.0196
	ESE	0.1704	0.0688	0.0319	0.1091	0.0369	0.0211
	CPR	0.9060	0.9000	0.8800	0.5940	0.0000	0.0080
(0.20,0.20)	Bias	0.0279	-0.0001	0.0134	-0.2717	-0.3316	0.0065
	SSE	0.2056	0.0806	0.0374	0.1037	0.0337	0.0212
	ESE	0.2049	0.0825	0.0373	0.1049	0.0355	0.0207
	CPR	0.9100	0.9180	0.8860	0.1660	0.0000	0.8620

Table 5.6: Estimation when $n = 1000$, $\beta = (0.6, 0.7, 0.0)$, and the variance of the random effect $v_r = 0.25$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta 1$	$\beta 2$	$\beta 3$	$\beta 1$	$\beta 2$	$\beta 3$
(0.05,0.05)	Bias	0.0116	0.0121	0.0073	-0.0754	-0.0887	0.0078
	SSE	0.0915	0.0336	0.0160	0.0788	0.0276	0.0140
	ESE	0.0919	0.0354	0.0175	0.0784	0.0282	0.0153
	CPR	0.9120	0.9080	0.9140	0.7400	0.0700	0.8920
(0.05,0.10)	Bias	0.0019	0.0164	0.0005	-0.1402	-0.1393	-0.0332
	SSE	0.0954	0.0391	0.0171	0.0769	0.0276	0.0137
	ESE	0.0968	0.0386	0.0185	0.0762	0.0277	0.0152
	CPR	0.8980	0.8700	0.9240	0.4140	0.0000	0.2700
(0.10,0.05)	Bias	0.0321	0.0091	0.0007	-0.0586	-0.1287	0.0339
	SSE	0.0954	0.0377	0.0179	0.0793	0.0263	0.0146
	ESE	0.0972	0.0382	0.0186	0.0793	0.0278	0.0152
	CPR	0.8920	0.9120	0.9140	0.8040	0.0020	0.2660
(0.10,0.10)	Bias	0.0180	0.0021	0.0024	-0.1435	-0.1809	0.0028
	SSE	0.1044	0.0403	0.0187	0.0768	0.0258	0.0136
	ESE	0.1042	0.0412	0.0198	0.0763	0.0269	0.0151
	CPR	0.8980	0.8940	0.9060	0.3940	0.0000	0.9260
(0.10,0.20)	Bias	0.0349	0.0227	-0.0053	-0.2615	-0.2590	-0.0748
	SSE	0.1205	0.0478	0.0228	0.0725	0.0267	0.0142
	ESE	0.1236	0.0494	0.0228	0.0745	0.0255	0.0151
	CPR	0.9060	0.8940	0.9000	0.0280	0.0000	0.0000
(0.20,0.10)	Bias	0.0343	0.0177	0.0042	-0.1465	-0.2432	0.0759
	SSE	0.1189	0.0487	0.0217	0.0742	0.0257	0.0142
	ESE	0.1183	0.0486	0.0226	0.0761	0.0261	0.0150
	CPR	0.8820	0.8900	0.9140	0.3780	0.0000	0.0000
(0.20,0.20)	Bias	0.0390	0.0216	-0.0059	-0.2742	-0.3275	-0.0032
	SSE	0.1374	0.0596	0.0258	0.0731	0.0249	0.0139
	ESE	0.1435	0.0596	0.0264	0.0736	0.0250	0.0146
	CPR	0.9040	0.8900	0.9040	0.0120	0.0000	0.9080

Table 5.7: Estimation when $n = 100$, $\beta = (0.4, 0.6, 0.2)$, and the variance of the random effect $v_r = 0.04$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta 1$	$\beta 2$	$\beta 3$	$\beta 1$	$\beta 2$	$\beta 3$
(0.05,0.05)	Bias	0.0178	0.0083	0.0120	-0.0180	-0.0690	-0.0153
	SSE	0.2671	0.1073	0.0517	0.2334	0.0862	0.0446
	ESE	0.2751	0.1072	0.0541	0.2335	0.0849	0.0462
	CPR	0.9180	0.9160	0.9100	0.9020	0.7600	0.8780
(0.05,0.10)	Bias	0.0138	0.0085	-0.0035	-0.1026	-0.1329	-0.0820
	SSE	0.2933	0.1167	0.0574	0.2292	0.0821	0.0449
	ESE	0.2985	0.1156	0.0572	0.2274	0.0814	0.0438
	CPR	0.9220	0.8920	0.9140	0.8500	0.4980	0.3940
(0.10,0.05)	Bias	0.0315	0.0205	0.0005	-0.0144	-0.1026	-0.0009
	SSE	0.3211	0.1154	0.0554	0.2466	0.0876	0.0471
	ESE	0.3028	0.1156	0.0552	0.2443	0.0859	0.0449
	CPR	0.8720	0.8980	0.9160	0.8880	0.6540	0.9080
(0.10,0.10)	Bias	0.0456	0.0117	-0.0097	-0.1044	-0.1657	-0.0576
	SSE	0.3324	0.1383	0.0580	0.2414	0.0819	0.0430
	ESE	0.3391	0.1259	0.0603	0.2376	0.0783	0.0435
	CPR	0.9180	0.8680	0.9140	0.8720	0.3480	0.6440
(0.10,0.20)	Bias	0.0060	0.0212	0.0168	-0.2082	-0.2477	-0.1379
	SSE	0.3981	0.1574	0.0713	0.2152	0.0765	0.0422
	ESE	0.3929	0.1514	0.0740	0.2200	0.0754	0.0425
	CPR	0.9020	0.9040	0.9220	0.7660	0.0700	0.0720
(0.20,0.10)	Bias	0.0442	0.0044	-0.0013	-0.0566	-0.2146	-0.0003
	SSE	0.3766	0.1432	0.0727	0.2461	0.0786	0.0436
	ESE	0.3889	0.1439	0.0686	0.2474	0.0779	0.0442
	CPR	0.9200	0.9060	0.8740	0.8900	0.1520	0.9020
(0.20,0.20)	Bias	0.0327	0.0442	0.0111	-0.1753	-0.3045	-0.0960
	SSE	0.4894	0.2030	0.0824	0.2278	0.0797	0.0410
	ESE	0.4846	0.1975	0.0867	0.2313	0.0773	0.0434
	CPR	0.9260	0.9240	0.9280	0.7980	0.0200	0.2640

Table 5.8: Estimation when $n = 500$, $\beta = (0.4, 0.6, 0.2)$, and the variance of the random effect $v_r = 0.04$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta1$	$\beta2$	$\beta3$	$\beta1$	$\beta2$	$\beta3$
(0.05,0.05)	Bias	0.0319	0.0027	-0.0022	-0.0245	-0.0841	-0.0282
	SSE	0.1207	0.0476	0.0228	0.1072	0.0386	0.0193
	ESE	0.1266	0.0470	0.0234	0.1074	0.0375	0.0199
	CPR	0.9100	0.9000	0.9140	0.8880	0.2740	0.6000
(0.05,0.10)	Bias	0.0187	0.0046	0.0019	-0.1061	-0.1384	-0.0773
	SSE	0.1369	0.0500	0.0259	0.1046	0.0347	0.0190
	ESE	0.1356	0.0518	0.0257	0.1024	0.0363	0.0194
	CPR	0.9060	0.9100	0.9040	0.7200	0.0200	0.0060
(0.10,0.05)	Bias	0.0080	0.0017	0.0003	-0.0353	-0.1089	-0.0030
	SSE	0.1303	0.0499	0.0237	0.1045	0.0353	0.0186
	ESE	0.1330	0.0501	0.0245	0.1077	0.0375	0.0200
	CPR	0.9040	0.9140	0.9200	0.8960	0.0960	0.9120
(0.10,0.10)	Bias	0.0017	0.0078	0.0059	-0.1016	-0.1657	-0.0474
	SSE	0.1414	0.0534	0.0270	0.1068	0.0343	0.0205
	ESE	0.1454	0.0556	0.0268	0.1040	0.0356	0.0194
	CPR	0.9160	0.9200	0.8960	0.7380	0.0000	0.2420
(0.10,0.20)	Bias	-0.0156	-0.0025	0.0004	-0.2236	-0.2534	-0.1463
	SSE	0.1612	0.0678	0.0308	0.0924	0.0316	0.0180
	ESE	0.1654	0.0648	0.0321	0.0965	0.0327	0.0189
	CPR	0.9120	0.8840	0.9080	0.2500	0.0000	0.0000
(0.20,0.10)	Bias	0.0022	0.0080	-0.0037	-0.0662	-0.2160	-0.0071
	SSE	0.1643	0.0630	0.0301	0.1081	0.0356	0.0192
	ESE	0.1661	0.0624	0.0299	0.1063	0.0351	0.0195
	CPR	0.9040	0.9000	0.8860	0.8260	0.0000	0.8680
(0.20,0.20)	Bias	0.0199	0.0073	-0.0010	-0.2045	-0.2942	-0.0912
	SSE	0.2059	0.0763	0.0371	0.0983	0.0366	0.0191
	ESE	0.1974	0.0804	0.0372	0.0982	0.0346	0.0192
	CPR	0.8880	0.9140	0.9080	0.3300	0.0000	0.0000

Table 5.9: Estimation when $n = 1000$, $\beta = (0.4, 0.6, 0.2)$, and the variance of the random effect $v_r = 0.04$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta1$	$\beta2$	$\beta3$	$\beta1$	$\beta2$	$\beta3$
(0.05,0.05)	Bias	0.0012	0.0022	-0.0014	-0.0654	-0.0868	-0.0278
	SSE	0.0880	0.0317	0.0167	0.0692	0.0286	0.0131
	ESE	0.0882	0.0335	0.0166	0.0747	0.0267	0.0142
	CPR	0.9080	0.9140	0.9000	0.7960	0.0760	0.3760
(0.05,0.10)	Bias	0.0016	0.0021	0.0017	-0.1214	-0.1417	-0.0753
	SSE	0.0952	0.0348	0.0179	0.0706	0.0257	0.0140
	ESE	0.0957	0.0362	0.0182	0.0722	0.0250	0.0139
	CPR	0.9000	0.9180	0.8940	0.4920	0.0000	0.0000
(0.10,0.05)	Bias	0.0028	-0.0006	0.0046	-0.0424	-0.1084	-0.0038
	SSE	0.0929	0.0361	0.0172	0.0769	0.0264	0.0131
	ESE	0.0916	0.0354	0.0176	0.0766	0.0262	0.0142
	CPR	0.8920	0.8940	0.8860	0.8220	0.0060	0.9120
(0.10,0.10)	Bias	0.0168	0.0039	-0.0014	-0.1007	-0.1659	-0.0524
	SSE	0.1004	0.0397	0.0190	0.0714	0.0255	0.0135
	ESE	0.1010	0.0391	0.0189	0.0726	0.0253	0.0137
	CPR	0.8960	0.8820	0.8900	0.6040	0.0000	0.0100
(0.10,0.20)	Bias	0.0020	0.0032	0.0003	-0.2189	-0.2492	-0.1413
	SSE	0.1111	0.0448	0.0225	0.0683	0.0238	0.0130
	ESE	0.1174	0.0462	0.0225	0.0686	0.0235	0.0134
	CPR	0.9100	0.9000	0.8980	0.0520	0.0000	0.0000
(0.20,0.10)	Bias	-0.0047	0.0013	0.0053	-0.0963	-0.2135	0.0047
	SSE	0.1177	0.0460	0.0223	0.0731	0.0260	0.0135
	ESE	0.1146	0.0444	0.0216	0.0740	0.0250	0.0140
	CPR	0.8900	0.8780	0.8700	0.6140	0.0000	0.8800
(0.20,0.20)	Bias	0.0086	-0.0004	0.0002	-0.1961	-0.3029	-0.0945
	SSE	0.1383	0.0538	0.0265	0.0721	0.0218	0.0132
	ESE	0.1401	0.0550	0.0262	0.0698	0.0229	0.0134
	CPR	0.9100	0.9080	0.9120	0.1360	0.0000	0.0000

Table 5.10: Estimation when $n = 100$, $\beta = (0.4, 0.6, 0.2)$, and the variance of the random effect $v_r = 0.25$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta1$	$\beta2$	$\beta3$	$\beta1$	$\beta2$	$\beta3$
(0.05,0.05)	Bias	0.0759	0.0437	0.0020	-0.0078	-0.0566	-0.0253
	SSE	0.3216	0.1079	0.0569	0.2683	0.0898	0.0426
	ESE	0.3172	0.1136	0.0565	0.2674	0.0891	0.0484
	CPR	0.8980	0.9060	0.9100	0.9000	0.8080	0.8980
(0.05,0.10)	Bias	-0.0491	0.0149	0.0072	-0.1187	-0.1254	-0.0782
	SSE	0.3285	0.1157	0.0601	0.2440	0.0859	0.0444
	ESE	0.3303	0.1190	0.0627	0.2515	0.0827	0.0484
	CPR	0.9080	0.9100	0.9220	0.8660	0.4960	0.5040
(0.10,0.05)	Bias	0.0545	0.0168	-0.0064	-0.0052	-0.0997	-0.0056
	SSE	0.3146	0.1112	0.0594	0.2523	0.0833	0.0485
	ESE	0.3051	0.1151	0.0613	0.2474	0.0866	0.0507
	CPR	0.8940	0.9120	0.9080	0.8980	0.6800	0.9080
(0.10,0.10)	Bias	-0.0087	0.0119	0.0076	-0.1150	-0.1603	-0.0410
	SSE	0.3306	0.1247	0.0629	0.2468	0.0864	0.0461
	ESE	0.3441	0.1305	0.0646	0.2533	0.0841	0.0478
	CPR	0.9180	0.9220	0.9060	0.8860	0.3860	0.7760
(0.10,0.20)	Bias	0.0344	0.0154	-0.0012	-0.2256	-0.2453	-0.1371
	SSE	0.4094	0.1556	0.0702	0.2360	0.0820	0.0444
	ESE	0.4100	0.1537	0.0746	0.2328	0.0778	0.0459
	CPR	0.9220	0.9160	0.9200	0.7280	0.0840	0.1000
(0.20,0.10)	Bias	0.0043	0.0493	-0.0022	-0.0961	-0.1921	0.0096
	SSE	0.4034	0.1550	0.0727	0.2510	0.0901	0.0461
	ESE	0.3988	0.1582	0.0740	0.2533	0.0875	0.0495
	CPR	0.9040	0.9200	0.9260	0.8600	0.3100	0.9220
(0.20,0.20)	Bias	-0.0501	0.0290	0.0087	-0.2567	-0.2942	-0.0936
	SSE	0.4691	0.1863	0.0894	0.2294	0.0754	0.0448
	ESE	0.4770	0.1815	0.0914	0.2358	0.0743	0.0475
	CPR	0.9120	0.9240	0.9060	0.7320	0.0220	0.3460

Table 5.11: Estimation when $n = 500$, $\beta = (0.4, 0.6, 0.2)$, and the variance of the random effect $v_r = 0.25$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta 1$	$\beta 2$	$\beta 3$	$\beta 1$	$\beta 2$	$\beta 3$
(0.05,0.05)	Bias	0.0028	0.0226	-0.0104	-0.0437	-0.0657	-0.0379
	SSE	0.1219	0.0482	0.0247	0.1032	0.0375	0.0204
	ESE	0.1277	0.0489	0.0255	0.1092	0.0391	0.0221
	CPR	0.9200	0.8740	0.8860	0.8860	0.4960	0.4840
(0.05,0.10)	Bias	0.0187	0.0055	-0.0077	-0.0612	-0.1266	-0.0758
	SSE	0.1352	0.0536	0.0262	0.1017	0.0373	0.0202
	ESE	0.1405	0.0530	0.0279	0.1090	0.0382	0.0214
	CPR	0.9140	0.8900	0.9100	0.8760	0.0480	0.0240
(0.10,0.05)	Bias	0.0334	0.0168	0.0074	-0.0329	-0.0937	0.0090
	SSE	0.1411	0.0507	0.0255	0.1162	0.0399	0.0214
	ESE	0.1379	0.0521	0.0269	0.1118	0.0392	0.0223
	CPR	0.8960	0.9020	0.9000	0.8800	0.2680	0.8820
(0.10,0.10)	Bias	-0.0348	0.0112	0.0147	-0.1323	-0.1536	-0.0380
	SSE	0.1436	0.0570	0.0268	0.1051	0.0383	0.0208
	ESE	0.1478	0.0573	0.0294	0.1074	0.0374	0.0218
	CPR	0.9020	0.9000	0.8840	0.6560	0.0140	0.4200
(0.10,0.20)	Bias	0.0226	0.0063	0.0009	-0.1910	-0.2329	-0.1460
	SSE	0.1730	0.0697	0.0347	0.1038	0.0359	0.0193
	ESE	0.1789	0.0668	0.0339	0.1061	0.0352	0.0208
	CPR	0.9160	0.8940	0.8760	0.4560	0.0000	0.0000
(0.20,0.10)	Bias	-0.0346	0.0045	0.0057	-0.1016	-0.2056	0.0078
	SSE	0.1723	0.0641	0.0319	0.1069	0.0377	0.0206
	ESE	0.1681	0.0648	0.0325	0.1093	0.0369	0.0219
	CPR	0.9060	0.9020	0.9160	0.7720	0.0020	0.8920
(0.20,0.20)	Bias	0.0045	0.0330	0.0043	-0.2052	-0.3043	-0.0799
	SSE	0.2153	0.0794	0.0393	0.0488	0.0176	0.0093
	ESE	0.2087	0.0832	0.0394	0.1046	0.0348	0.0210
	CPR	0.8800	0.8980	0.9080	0.2820	0.0000	0.0000

Table 5.12: Estimation when $n = 1000$, $\beta = (0.4, 0.6, 0.2)$, and the variance of the random effect $v_r = 0.25$ based on 500 simulations under the CGQL method and the NGQL method. Where α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		CGQL			NGQL		
α	Quantity	$\beta 1$	$\beta 2$	$\beta 3$	$\beta 1$	$\beta 2$	$\beta 3$
(0.05,0.05)	Bias	-0.0102	0.0161	0.0079	-0.0661	-0.0777	-0.0203
	SSE	0.0914	0.0326	0.0165	0.0786	0.0278	0.0146
	ESE	0.0920	0.0352	0.0180	0.0781	0.0281	0.0155
	CPR	0.8840	0.8840	0.9100	0.7720	0.1500	0.6300
(0.05,0.10)	Bias	0.0220	0.0130	0.0071	-0.1017	-0.1312	-0.0726
	SSE	0.0992	0.0360	0.0176	0.0775	0.0277	0.0146
	ESE	0.0999	0.0380	0.0194	0.0759	0.0267	0.0151
	CPR	0.8980	0.9100	0.9060	0.6100	0.0000	0.0000
(0.10,0.05)	Bias	-0.0188	0.0104	0.0013	-0.0600	-0.0935	0.0017
	SSE	0.0941	0.0381	0.0183	0.0820	0.0286	0.0146
	ESE	0.0965	0.0370	0.0187	0.0788	0.0279	0.0155
	CPR	0.9040	0.8920	0.9240	0.7760	0.0560	0.9400
(0.10,0.10)	Bias	0.0388	0.0056	0.0040	-0.1180	-0.1576	-0.0523
	SSE	0.1042	0.0419	0.0191	0.0764	0.0264	0.0147
	ESE	0.1081	0.0406	0.0204	0.0769	0.0263	0.0152
	CPR	0.9000	0.8880	0.9260	0.5440	0.0000	0.0380
(0.10,0.20)	Bias	0.0003	0.0144	0.0029	-0.2171	-0.2382	-0.1432
	SSE	0.1221	0.0464	0.0228	0.0722	0.0248	0.0140
	ESE	0.1257	0.0476	0.0240	0.0739	0.0245	0.0147
	CPR	0.9080	0.9080	0.9140	0.1040	0.0000	0.0000
(0.20,0.10)	Bias	-0.0114	0.0137	0.0055	-0.1031	-0.2025	0.0084
	SSE	0.1216	0.0458	0.0216	0.0759	0.0270	0.0140
	ESE	0.1207	0.0460	0.0228	0.0786	0.0260	0.0153
	CPR	0.8820	0.8880	0.9140	0.6420	0.0000	0.8860
(0.20,0.20)	Bias	0.0193	0.0101	0.0082	-0.1911	-0.2920	-0.0901
	SSE	0.1483	0.0563	0.0267	0.0727	0.0231	0.0144
	ESE	0.1465	0.0560	0.0279	0.0742	0.0242	0.0147
	CPR	0.8920	0.8960	0.9040	0.1640	0.0000	0.0000

5.3 Comparison

From the simulation results in Section 5.2, we observe that the CGQL method makes a remarkable reduction in bias as compared to the NGQL method. In the attached Tables 5.1-5.12, the performance of the CGQL method and the NGQL method is compared with the concern of the bias of $\hat{\beta}$ (Bias), simulated standard error (SSE), estimated standard error (ESE), and coverage probability (CPR)of the 90% confidence interval. These simulated Bias, SSE, ESE, and CPR are reported in Tables 5.1-5.6 for the case when $\beta = (0.6, 0.7, 0.0)$, and Tables 5.7-5.12 for $\beta = (0.4, 0.6, 0.2)$. In this subsection, we discuss the simulation results from three different perspectives: bias, standard errors, and coverage probabilities.

The NGQL estimates are severely biased. Most of relative biases ($\frac{|\hat{\beta}-\beta|}{\beta}$) are greater than 5%. For example, in Table 5.9 when $\alpha = (0.05, 0.05)$ and $(0.2, 0.2)$ the relative biases are (16.4%, 14.5%, 13.9%) and (49.0%, 50.5%, 47.3%) for the CGQL and NGQL methods, respectively. This is understandable: the higher the misclassification rate, the bigger the bias. Also, the NGQL tends to always underestimate the regression parameter.

To rectify the attenuation effect of the NGQL method, we developed the CGQL method. The biases of the estimated β for the CGQL method are remarkably smaller compared to those of the NGQL method. The efficiency gain increases when the sample size I increases. The simulation in Tables 5.1-5.12 show that all of relative biases of the CGQL estimates are smaller than 5%. That makes it reasonable to claim the unbiasedness of the CGQL estimators. For instance, in Table 9 when $\alpha = (0.05, 0.05)$ and $(0.2, 0.2)$ the relative biases are $(0.3\%, 0.4\%, 0.7\%)$ and $(2.2\%, 0.1\%, 0.1\%)$ for the CGQL method, respectively. The improvement of the CGQL method over the NGQL is obvious.

From Tables 5.1-5.12 we can see that the estimated standard errors based on equation 4.3 are almost unbiased, in the sense that the estimated standard error and simulated standard error are very close for the CGQL approach. As an example, $SSE = (0.1004, 0.0397, 0.0190)$ and $ESE = (0.1010, 0.0391, 0.0189)$ when $\alpha = (0.10, 0.10)$, in Table 9. Furthermore for the CGQL method, the standard error of the estimated regression coefficients $\hat{\beta}$ increases as the misclassification rate α increases.

From the simulation results we can also observe that the random effects in the model are well accounted. The biases and standard errors of the estimates of the model parameters do not vary significantly with the change of the variance of the random effects. For example, when the variance of the random effects changes from $v_r = 0.04$ to $v_r = 0.25$, the estimate of the regression coefficients and their standard errors do not change much.

Moreover, when $\alpha_1 < \alpha_2$ the values of the ESE's tend to be larger as compared to the ones when $\alpha_1 > \alpha_2$. Since

$$specificity \; = 1 - type\ I\ error = 1 - \alpha_1,$$

and

$$sensitivity \ = 1 - type\ II\ error = 1 - \alpha_2,$$

the sensitivity has more significant influence on the estimation of the model parameters than specificity does. For the case of decreasing sensitivity the estimation bias does not change much in our simulations.

The coverage probability is computed as the proportion of situations that the 90% confidence interval includes the true β. It is clear from Tables 5.1-5.12 that the coverage probabilities for the CGQL method are much closer to the nominal level of 90% than that of the NGQL method. Since the NGQL method creates biased estimates of the model parameters, the corresponding confidence interval is already meaningless. We included it in the tables to indicate that the CGQL method is more efficient as compared to the NGQL method.

Chapter 6

Sensitivity Analysis

In this chapter, we investigate the performance of the CGQL method when the estimate of standard deviation of the random effect(SD), usually unknown, is slightly biased. We take the SD equal to 0.2 and 0.5 for each setting of the model parameters in the simulation. We use different values of the SD in simulation to see if the CGQL method is robust. For the sake of simplicity, we still use SD to denote the standard deviation in the tables, although its values are changed from case to case in the simulation. The simulation results are given in Tables 5.1-5.12.

Table 6.1: Estimation when $n = 500$, $\beta = (0.4, 0.6, 0.2)$, and standard deviation of the random effect $SD = 0.20$ based on 500 simulations under the CGQL method. Where SD: standard deviation of the random effect; α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		$\alpha = (0.05, 0.10)$			$\alpha = (0.10, 0.15)$		
SD	Quantity	$\beta 1$	$\beta 2$	$\beta 3$	$\beta 1$	$\beta 2$	$\beta 3$
0.15	Bias	0.0115	-0.0020	-0.0022	-0.0203	0.0014	-0.0001
	SSE	0.1416	0.0526	0.0263	0.1482	0.0578	0.0288
	ESE	0.1368	0.0514	0.0251	0.1527	0.0601	0.0292
	CPR	0.8840	0.9020	0.8860	0.9140	0.9140	0.9060
0.20	Bias	0.0187	0.0039	0.0019	-0.0021	0.0037	-0.0006
	SSE	0.1369	0.0500	0.0259	0.1540	0.0595	0.0301
	ESE	0.1356	0.0504	0.0257	0.1524	0.0607	0.0295
	CPR	0.9060	0.9020	0.9040	0.9040	0.9100	0.8960
0.25	Bias	-0.0146	0.0118	0.0018	0.0167	0.0067	0.0006
	SSE	0.1347	0.0513	0.0265	0.1617	0.0609	0.0287
	ESE	0.1351	0.0523	0.0258	0.1566	0.0609	0.0299
	CPR	0.9040	0.9140	0.8900	0.8960	0.9040	0.9000

Table 6.2: Estimation when $n = 500$, $\beta = (0.4, 0.6, 0.2)$, and standard deviation of the random effect $SD = 0.5$ based on 500 simulations under the CGQL method. Where SD: standard deviation of the random effect; α: misclassification rate; Bias: the bias of the estimated β; SSE: simulated standard errors; ESE: estimated standard error; CPR: coverage probabilities of 90% confidence interval.

		$\alpha = (0.05, 0.10)$			$\alpha = (0.10, 0.15)$		
SD	Quantity	$\beta 1$	$\beta 2$	$\beta 3$	$\beta 1$	$\beta 2$	$\beta 3$
0.4	Bias	-0.0227	0.0049	-0.0058	-0.0354	-0.0008	0.0046
	SSE	0.1349	0.0537	0.0263	0.1535	0.0636	0.0304
	ESE	0.1395	0.0526	0.0261	0.1549	0.0619	0.0304
	CPR	0.9100	0.8940	0.9000	0.8960	0.9100	0.9060
0.5	Bias	0.0113	0.0099	0.0043	-0.0000	-0.0024	0.0092
	SSE	0.1409	0.0535	0.0264	0.1605	0.0620	0.0289
	ESE	0.1406	0.0537	0.0273	0.1658	0.0629	0.0306
	CPR	0.9060	0.9100	0.9070	0.9060	0.9040	0.8980
0.6	Bias	0.0127	0.0343	0.0138	-0.0470	0.0562	-0.0046
	SSE	0.1509	0.0551	0.0265	0.1679	0.0670	0.0290
	ESE	0.1514	0.0550	0.0287	0.1677	0.0652	0.0325
	CPR	0.9060	0.8420	0.8900	0.8760	0.7860	0.9300

Tables 6.1 and 6.2 show the comparison of the CGQL method by using different values of SD. From the results we can see that the CGQL estimates are robust to different values of SD. For instance, in Table 6.1, the GQL misclassification probabilities are set to be $\alpha = (0.05, 0.10)$ and $\alpha = (0.10, 0.15)$ for $SD = 0.20$. When the value of SD is taken as 0.15, 0.20, or 0.25, the differences between Bias, SSE, ESE, and CPR of the estimate of other model parameters are not significant. In Table 6.2 where $SD = 0.5$, the coverage probabilities are less satisfactory when the SD is overestimated. The situation gets worse when the misclassification rates are also high. For slightly underestimated SD, the CGQL works well.

Chapter 7

Discussion

The likelihood function is very hard to calculate for the logistic mixed model thus the MLE is difficult to develop. As in Sutradhar (2003), we use generalized quasi-likelihood method to conduct statistical inference of the model parameters. When misclassification in binary responses exists, the estimates of the regression parameters are attenuated.

In this practicum, we develop an approach that successfully corrects the estimation bias. We focus on unconditional generalized quasi-likelihood inference that involves unconditional moments of up to the second order. Following Monahan and Stefanski (1992), we formulate the generalized linear mixed model with Logit link and Probit link in Chapter 3, and in Chapter 4 we also utilize the method to approximate the expectation of an unknown function involved in the calculation of the expectations and covariances (see also Roy, Banerjee, and Maiti, 2005).

This practicum develops a corrected generalized quasi-likelihood method to analyze longitudinal binary data with misclassification in response. In addition, the work here assumed the misclassification probability is known. In practice, they can usually be estimated from a small validation sample or other similar studies. The simulation results in Section 5.2 show that the CGQL approach has very good performance. In Section 5.3 we also discussed the remarkable efficiency gain as compared to the naive method. Chapter 6 shows the robustness of the CGQL method when model parameters change slightly in the simulation. This practicum should be useful in biological or medical sciences to analyze similar binary data, which is to be taken up in a future communication.

Appendix

The program is running by R.

1. Data Generation:

```
n<-100
t<-4
p<-3
k<-500
parameter<-matrix(0,p,k)
Pr<-matrix(0,p,k)
ste<-matrix(0,p,k)
step<-matrix(0,p,k)
x1<-matrix(0,n,t)
x2<-matrix(0,n,t)
x3<-matrix(0,n,t)
x1[1:(n*0.5),1]<-rbinom(n*0.5,1,0.4)
x1[1:(n*0.5),2]<-rbinom(n*0.5,1,0.3)
```

```
x1[1:(n*0.5),3]<-rbinom(n*0.5,1,0.2)
x1[1:(n*0.5),4]<-rbinom(n*0.5,1,0.1)
x1[(n*0.5+1):n,1]<-rbinom(n*0.5,1,0.5)
x1[(n*0.5+1):n,2]<-rbinom(n*0.5,1,0.4)
x1[(n*0.5+1):n,3]<-rbinom(n*0.5,1,0.3)
x1[(n*0.5+1):n,4]<-rbinom(n*0.5,1,0.2)
x2[1:(n*0.5),1]<-rnorm(n*0.5,0.0,1.5)
x2[1:(n*0.5),2]<-rnorm(n*0.5,0.1,1.5)
x2[1:(n*0.5),3]<-rnorm(n*0.5,0.2,1.5)
x2[1:(n*0.5),4]<-rnorm(n*0.5,0.3,1.5)
x2[(n*0.5+1):n,1]<-rnorm(n*0.5,0.1,1.5)
x2[(n*0.5+1):n,2]<-rnorm(n*0.5,0.2,1.5)
x2[(n*0.5+1):n,3]<-rnorm(n*0.5,0.3,1.5)
x2[(n*0.5+1):n,4]<-rnorm(n*0.5,0.4,1.5)
x3[1:(n*0.5),1]<-rpois(n*0.5,4)
x3[1:(n*0.5),2]<-rpois(n*0.5,18)
x3[1:(n*0.5),3]<-rpois(n*0.5,22)
x3[1:(n*0.5),4]<-rpois(n*0.5,30)
x3[(n*0.5+1):n,1]<-rpois(n*0.5,5)
x3[(n*0.5+1):n,2]<-rpois(n*0.5,20)
x3[(n*0.5+1):n,3]<-rpois(n*0.5,25)
x3[(n*0.5+1):n,4]<-rpois(n*0.5,35)
x<-array(0,dim=c(p,n,t))
```

```
x[1,,]<-t(x1)
x[2,,]<-t(x2)
x[3,,]<-t(log(x3+0.1))
beta<-c(0.5,0.7,0.3)
truebeta<-c(0.4,0.6,0.2)
alpha<-c(0.05, 0.05)
ks<-1.7
sigma<-rnorm(n, 0, 0.3)
vr<-0.04
gamma<-rnorm(n,0,sqrt(vr))
muY<-matrix(0,n,t)
muT<-matrix(0,n,t)
Y<-array(0,dim=c(n,t))
T<-array(0,dim=c(n,t))
error<-10^-4
d<-c(1,1,1)
library(mnormt)
library(MASS)
q<-qnorm(0.95,0,1)
indalpha1<-function(x,Y){if (x-q*Y<=truebeta[1] & x+q*Y>=truebeta[1]) 1 else 0}
indalpha2<-function(x,Y){if (x-q*Y<=truebeta[2] & x+q*Y>=truebeta[2]) 1 else 0}
indbeta<-function(x,Y){if (x-q*Y<=truebeta[3] & x+q*Y>=truebeta[3]) 1 else 0}
covY<-array(0, dim=c(t,t,n))
```

```
devmuY<-array(0,dim=c(p,t,n))
sigma<-matrix(0,n,p)
fun<-c(0,0,0)
std<-c(0,0,0)
CI<-c(0,0,0)
cov1<-array(0, dim=c(p,p,n))
Tcov<-matrix(0,p,p)
Thecov<-matrix(0,p,p)
muTf<-function(beta,gamma,x){
exp(t(x[,i,])%*%truebeta+gamma[i])/(1+exp(t(x[,i,])%*%truebeta+gamma[i]))}
RESULT<-array(0,dim=c(3,4,k))
for (k in 1:500){
print("simulation:")
print(k)
for (i in 1:n){
muT[i,]<-muTf(truebeta,gamma,x)
}
for (d in 1:t)
for (i in 1:n){
T[i,d]<-rbinom(1,1,muT[i,d])
Y[i,d]<-(T[i,d]==1)*rbinom(1,1,1-alpha[2])+(T[i,d]==0)*rbinom(1,1,alpha[1])
}
RESULT[,,k]<-GQL(Y,x,beta)
```

```
parameter[,k]<-RESULT[,1,k]
Pr[,k]<-RESULT[,2,k]
ste[,k]<-RESULT[,3,k]
step[,k]<-RESULT[1,4,k]
}
SM<-c(0,0,0)
SSE<-c(0,0,0)
ESE<-c(0,0,0)
CPR<-c(0,0,0)
for (i in 1:3){
SM[i]<-mean(parameter[i,])
ESE[i]<-mean(ste[i,])
SSE[i]<-sd(parameter[i,])
CPR[i]<-sum(Pr[i,])/k
}
table1<-matrix(0,4,3)
table1[1,]<-SM-truebeta
table1[2,]<-SSE
table1[3,]<-ESE
table1[4,]<-CPR
print("Mathed: GQL")
print("used: Y")
print(k)
```

```
print(n)
print(alpha)
print(truebeta)
print(vr)
table1
```

2. Method:

```
GQL<-function(Y, x, beta){
r<-0
while( (d[1]>error || d[2]>error || d[3]>error) & all(is.finite(beta)) & r<20){
r=r+1
beta0<-beta
alpha0<-alpha
for (i in 1:n){ muY[i,]<-alpha0[1]+(1-alpha0[1]-alpha0[2])*exp(t(x[,i,])%*%beta0/s
}
for (h in 1:n){
for (i in 1:t){
for (j in 1:t){
if (i==j){
j=i
s<-exp(x[,h,i]%*%beta0/sqrt(1+vr/ks))
g<-(s/(1+s))
bm<-alpha0[1]+(1-alpha0[1]-alpha0[2])*g
covY[i,i,h]<-bm- bm^2
}
else{
ss1<-exp(x[,h,i]%*%beta0/sqrt(1+vr/ks))
gg1<-(ss1/(1+ss1))
```

```
bm1<-alpha0[1]+(1-alpha0[1]-alpha0[2])*gg1
ss2<-exp(x[,h,j]%*%beta0/sqrt(1+vr/ks))
gg2<-(ss2/(1+ss2))
bm2<-alpha0[1]+(1-alpha0[1]-alpha0[2])*gg2
Fs<-alpha0[1]^2
s1<-exp(x[,h,i]%*%beta0/sqrt(1+vr/ks))
g1<-(s1/(1+s1))
Sc<-alpha0[1]*(1-alpha0[1]-alpha0[2])*g1
s2<-exp(x[,h,j]%*%beta0/sqrt(1+vr/ks))
g2<-(s2/(1+s2))
Th<-alpha0[1]*(1-alpha0[1]-alpha0[2])*g2
m1<- exp(x[,h,i]%*%beta0)/(1+exp(x[,h,i]%*%beta0))^2 * 1/(1/sqrt(2*pi)*exp(-(x[,h,
m2<- exp(x[,h,j]%*%beta0)/(1+exp(x[,h,j]%*%beta0))^2 * 1/(1/sqrt(2*pi)*exp(-(x[,h,
R<-vr/(1+vr)
vc<-matrix(c(1+vr,vr,vr,1+vr),2,2)
Fr<-(1-alpha0[1]-alpha0[2])^2*(g1*g2+m1*m2*(pmnorm(c(x[,h,i]%*%beta0, x[,h,j]%*%be
[,h,i]%*%beta0, 0, sqrt(1+vr))*pnorm(x[,h,j]%*%beta0, 0, sqrt(1+vr))))
covY[i,j,h]<-Fs+Sc+Th+Fr - bm1*bm2
}
}
}
}
for (i in 1:n)
```

```
for (t in 1:4){
devmuY[,t,i]<-
(1-alpha0[1]-alpha0[2])*(
(exp(t(x[,i,t])%*%beta0/sqrt(1+vr/ks)) / ((1+exp(t(x[,i,t])%*%beta0/sqrt(1+vr/ks))
)
}
for (i in 1:n){
sigma[i,]<-devmuY[,,i] %*% ginv(covY[,,i]) %*% t(t(Y[i,]-muY[i,]))
}
for (p in 1:3){
fun[p]<-sum(sigma[,p])
}
print("fun")
print(fun)
for (i in 1:n){
cov1[,,i]<-devmuY[,,i] %*% ginv(covY[,,i]) %*% t(devmuY[,,i])
}
for (i in 1:p)
for (j in 1:p){
Tcov[i,j]<-sum(cov1[i,j,])
}
Thecov<-ginv(Tcov)
print(r)
```

```
print("cov")
print(Thecov)
beta<-beta0 + ginv(Tcov)%*%fun
print("beta")
print(beta)
d<-abs(beta-beta0)
}
std<-(diag(Thecov))^0.5
CI[1]<-indalpha1(beta[1], std[1])
CI[2]<-indalpha2(beta[2], std[2])
CI[3]<-indbeta(beta[3], std[3])
result<-matrix(0,3,4)
result[,1]<-beta
result[,2]<-CI
result[,3]<-c(std[1], std[2], std[3])
result[1,4]<-r
print("result")
print(result)
}
```

Bibliography

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *J. Amer. Statist. Assoc.* 88, 9-25.

Carroll, R., Maca, J., and Ruppert, D. (1999). Nonparametric regression in the presence of measurement error. *Biometrika* 86, 541-554.

Greenland, S. (1988). Variance estimation of epidemiologic effect estimates under misclassification. *Statistics In Medicine* 7, 745-757.

Imbens, G. W. (2002). Generalized method of moments and empirical likelihood. *Journal of Business and Economic Statistics* 20, 493-506.

Ji, Y. and Fan, Z. (2010). Analysis of longitudinal categorical and count data subject to measurement error. *Doctorial Thesis*, Memorial University of Newfoundland, St.John's, Newfoundland, Canada.

Jiang, J. and Zhang, W. (2001). Robust estimation in generalized linear mixed models. *Biometrika* 88, 753-765.

Kuha, J., Skinner, C., and Palmgren, J. (2001). Misclassification error. *In Encyclopedia of Biostatistics*, P.Armitage and T. Colton (eds), 2651-2621. Chichester:

Wiley.

Küchenhoff, H., Mwalili, S. M. and Lesaffre, E. (2006). A general method for dealing with misclassification in regression: the misclassification SIMEX. *Biometrics* 62, 85-96.

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

McKenzie, E. (1988). Some ARMA models for dependent sequences of Poisson counts. *Applied Probability* 20, 822-835.

Monahan, H., Stefanski, LA. (1992). Normal scale mixture approximations to $F^{*}(z)$ and computation of the logisticnormal integral. *In Hand book of the Logistic Distribution*, Balakrishnan N (eds), 529-540. Marcel Dekker: New York.

Morrissey, JM., Spiegelman, D. (1999). Matrix methods for estimating odds ratios with misclassified exposure data: extensions and comparisons. *Biometrics* 55, 338-344.

Neuhaus, J. M. (1999). Bias and efficiency loss due to misclassified responses in binary regression. *Biometrika* 86, 4: 843-855.

Neuhaus, J. M. (2001). Assessing change with longitudinal and clustered binary data. *Annual Review of Public Health* 22, 115-128.

Neuhaus, J. M. (2002). Analysis of clustered and longitudinal binary data subject to response misclassification. *Biometrics* 58, 3:675-683.

Roy, S., Banerjee, T., and Maiti, T. (2005). Measurement error model for misclassified binary responses. *Statistics In Medicine* 24, 269-283.

Sutradhar, B. C. (2003). An overview on regression models for discrete longitudinal responses. *Statistical Science* 18, 377-393.

Sutradhar, B. C. (2008). Generalized method of moments versus generalized quasi-likelihood inferences in binary panel data models. *The Indian Journal of Statistics* 70-B, 34-62.

Sutradhar, B. C. and Farrell, P. J. (2007). On optimal lag 1 dependence estimating for binary models with application to asthma data. *Sankhya: The Indian Journal of Statistics* 69, 448-467.

Wedderburn, R. (1974). Quasi-likelihhod functions, generalized linear models, and the Gauss-Newton method. *Biometrika* 61, 439-447.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica* 50, 1-25.