Joint Tweedie Mixed Models for Longitudinal Data of Mixed Types

by

Jiaxiu Li

Bsc. (Management)- Shanghai Institute of Technology, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

In the Graduate Academic Unit of Mathematics and Statistics

Supervisor(s):	Guohua Yan, Ph.D, Statistics
	Renjun Ma, Ph.D, Statistics
Examining Board:	James Watmough,Ph.D, Statistics, Chair
	M. Tariqul Hasan, Ph.D, Statistics
	Fan-Rui Meng, Ph.D, Forestry & Environmental Management

This thesis is accepted

Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

September, 2017

Abstract

Recently, there has been tremendous interest in developing approaches for joint analysis of longitudinal data. This has gained emphasis particularly in medical studies, where different types of outcomes are frequently collected over time on each of many subjects. For example, both the spike burst rate and the duration are observed longitudinally in the myoelectric activity study in ponies. Univariate mixed-effects models are usually used for a single outcome since they can analyze variability of different sources. However, difficulties in analyzing the correlation between the responses arose because the two responses of the same pony are not independent. As the relationships between these two outcomes are of particular interest, these two outcomes should be analyzed jointly instead of separately.

In this thesis, we propose to model data of mixed types jointly by incorporating both subject-specific and time-specific random effects through the Generalized Linear Mixed Models (GLMM) based on the class of Tweedie distributions of different index parameters. This class of models can handle a variety of data types, including continuous, discrete and mixed data. An optimal equation to predict the random effects has been obtained based on the best linear unbiased predictor of the random effects. This approach also has the advantage of computational efficiency. An important feature of the proposed model is that two types of correlations are taken into account: correlations between measurements on different response variables and correlations between measurements on the same response variable within subject or cluster. Joint modelling facilitates analysis of correlation between observations on the same subject or within the cluster. What is more, the analysis results do not rely on any distributional assumption of random effects.

The proposed method is illustrated with the analysis of the myoelectric activity in ponies dataset which consists of two responses. Both univariate model and joint model were conducted. Through comparing the results, we conclude that analyzing the responses jointly instead of separately can not only assess the association of responses but also improve the estimation efficiency. Simulation study is also conducted to verify the proposed model.

Dedication

This report is dedicated to my parents who give me supports with their unconditional love forever.

Acknowledgements

Foremost, I have a tremendous amount of gratitude towards my supervisor Dr. Guohua Yan for his constant instruction, persistent help and great patience. Without his advice and help, this thesis would not have been completed.

I would like to express my sincere gratitude to my another supervisor Dr. Renjun Ma. He has played a very important role of my life since I started my graduate school studies, and his excellent guidance, constructive suggestions and occasional pushing have made all the difference when things might have been harder than expected.

I am also grateful to Dr. Tariqul Hasan and Dr. Jeffery Picka for their encouragement and kindness.

I am indebted to the University of New Brunswick and the Department of Mathematics and Statistics for providing such an excellent academic program and friendly environment. Many thanks to faculty members and staff.

Finally, I would like to thank my brother and my friend Qing Yu for their love and support.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In longitudinal studies, mixed-effects models are commonly used as they are able to account for both within- and between-subject sources of variability. A wide range of applications of univariate mixed-effects models for a single response variable that may be continuous or categorical have been investigated by many researchers (Throp, 2009). However, in medical studies, different types of outcomes are frequently collected over time on each of many subjects. For example, both the spike burst rate and the duration are observed longitudinally in the myoelectric activity study in ponies. Therefore, difficulties have arisen in analyzing bivariate responses because of correlations usually present between the two responses on the same subject. In recent years, extensions have been made in Fitzmaurice et al.(2008) for jointly modelling multiple response variables or mixtures of discrete and continuous outcomes measures longitudinally.

Models for clustered bivariate discrete and continuous responses are introduced by Catalano and Ryan (1992) and Fitzmaurice and Laird (1995), but their approaches mainly focus on the marginal modelling instead of joint, random effects models (Gueorguieva RV and Agresti A, 2001). Latent variables models were used to analyze the mixtures of discrete and continuous outcomes by Sammel et al.(1997). Rochon (1996) has applied the generalized estimating equations to fit the extended marginal models for bivariate repeated measures of discrete or continuous outcomes. Unfortunately, such a method risks attenuating the estimates of the regression parameters in cases in which the focus is on subject-specific inference. Regan and Catalano (1999) also considered a maximum likelihood estimation for all marginal and correlation parameters using a joint model equation. But their approach is only suitable when there is an exchangeable correlation structure between the continuous and binary responses. However, in many cases, these methods are of limited use, especially when the random effects is of primary interest.

When multiple variables need to be analyzed, the joint models are preferred since it not only considers the correlation between measurements on different variables for each subject but also takes into account the correlation between measurements at different time points of the same subject (S.Fieuws et al., 2007). The goal of joint modelling is to provide a framework in which relationships between and within subjects can be investigated and the accuracy of parameter estimation can be improved.

In this thesis, we propose a Generalized Linear Mixed Model (GLMM) framework for longitudinal responses of mixed types. The framework is based on the class of Tweedie exponential dispersion distributions (Jørgensen 1987) which includes as special cases the Poisson, normal gamma, inverse Gaussian, compound Poisson, gamma and so on (Ma, 1999); for this very reason, the proposed model framework is able to handle various data types, including both continuous and discrete data. Multilevel random effects are introduced in the proposed model to account for correlation structure of the longitudinal data.

The novelty of our approach lies in modelling data of mixed types jointly by incorporating both subject-specific and time-specific random effects into Tweedie models through different index. Moreover, this approach incorporates two types of correlations: correlations between measurements on different response variables and correlations between measurements on the same response variable within a subject, which can reflect the hierarchical structure clearly in the context of longitudinal data.

An optimal equation to predict the random effects has been obtained based on the best linear unbiased predictor of the random effects. This approach leads to improved computational efficiency since all the formulas are explicit in each iterative step.

One important feature of the model is that the subject-specific random effects is shared between the two response variables; this helps to study the heterogenity between different responses and allows easy extensions to multiple responses.

We demonstrate the model using a motivating example. The purpose of the example is to compare the effects of the nine drugs and a placebo on the patterns of myoelectric activity in the intestines of the ponies.([1]Lester et al., 1998a; 1998b;

1998c). The researcher chose six ponies, attached electrodes to different areas of the intestines of each pony and administered each treatment to each pony twice. Then, the spike burst rate and the duration were collected at equally spaced time intervals. The goal is to assess the effects of these drugs on spike burst rate and duration.

The remaining part of this thesis is organized as follows:

Chapter 2 presents an introductory overview of statistical models for longitudinal studies. Two general approaches for analyzing longitudinal data are discussed. The GLMM and Tweedie exponential dispersion models are briefly introduced. In addition, the basic idea of the joint model is described.

In chapter 3, the proposed model is presented hierarchically through a few assumptions. The moment structures of the responses and the moment structures of the random effects are derived and also presented in matrix forms. The orthodox best linear unbiased predictors of the random effects, as well as the iterative equations for estimating regression and random effects parameters are also developed.

To demonstrate the proposed model, we analyze the myoelectric activity in ponies dataset in chapter 4. Analysis results along with relevant graphs and figures are presented and discussed. Separate analysis of the two response variables are also included in order to compare the results. In addition, we conducted a brief simulation study to assess the behaviour of our estimating algorithm.

We conclude the thesis in Chapter 5 with a brief discussion of our proposed model and point to a few aspects for further research.

Chapter 2

Literature Review

2.1 Introduction of longitudinal study

2.1.1 Longitudinal study

Longitudinal studies are commonly used in many areas of research, such as medical, public health and psychology. For example, if a subject is involved in an experiment, then the subject is assessed several times during a period of time, the response variable and the explanatory variables are measured repeatedly.

Let Y_{it} represent the t^{th} time of the i^{th} subject; x_{itk} denote the k^{th} covariate for the it^{th} measurement, and $i = 1, \cdots m$, $t = 1, \cdots t_m$ and $k = 1, \cdots p$. Therefore, a longitudinal data set can be expressed in Table 2.1.

For longitudinal data, the defining characteristic is that the same measurements are taken repeatedly over a period of time. The most important feature of

Table 2.1: Layout of a longitudinal data set

ID	Repeated numbers	y_{it}	x_{itk}
1	1	y_{11}	$x_{111} \cdots x_{11p}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	t_1	y_{1t_1}	$x_{1t_1 1} \cdots x_{1t_1 p}$
2	1	y_{21}	$x_{211} \cdots x_{21p}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
2	t_2	y_{2t_2}	$x_{2t_2 1} \cdots x_{2t_2 p}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	1	y_{m1}	$x_{m11} \cdots x_{m1p}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	t_m	y_{mt_m}	$x_{mt_m 1} \cdots x_{mt_m p}$

longitudinal data is that it is able to study the change. Therefore, the primary goal of most longitudinal studies is to capture the changes within the subject and characterize the factors that affected the change (Fitzmaurice et al., (2004)).

The benefit of the longitudinal study is that researchers are able to detect the changes or developments of the subject beyond a specific time period. This is also the critical distinction between cross-sectional and longitudinal studies. Since the defining feature of cross-sectional studies is that they allow comparisons among different population groups at a single point in time.

Generally speaking, it is quicker to work on cross-sectional studies than longitudinal studies. However, longitudinal studies can provide more information about the correlation among the repeated measurements than cross-sectional studies. That is why researchers might start with cross-sectional studies first to investigate the links between certain variables. Then they would launch the longitudinal study to analyze the correlation within or between subjects.

Both methods are useful approaches to research. Our purpose in this thesis will focus on studying longitudinal data.

2.1.2 General approaches

Researchers have spent a great deal of time deriving methods for analyzing longitudinal data. Several general approaches have been described to model longitudinal data by Diggle *et al.*(2002). In Fitzmaurice et al.(2008), various methods are introduced for the analysis of such data. Two methods are introduced below.

One approach for modelling longitudinal data is to consider the marginal mean. The term marginal here means that the model for the mean response depends only on the covariates of interest not on random effects at each occasion. A marginal model can be applied to use the framework of generalized linear models(GLM) (Nedler and Wedderburn, 1972). The model can be devided into three parts. Firstly, the mean of each response, $E(Y_{ij}|X_{ij}) = \mu_{ij}$, is assumed to be independent across the subjects but correlated with each subject. The correlation between μ_{ij} and the covariates X_{ij} follows the equation, $g(\mu_{ij}) = X_{ij}'\beta$, where $g(\cdot)$ is a known function. Secondly, the conditional variance of each response Y_{ij}, given the covariates X_{ij} is specified as $Var(Y_{ij}|X_{ij}) = \phi v(\mu_{ij})$, where $v(\mu_{ij})$ is a known variance function of μ_{ij} and ϕ is a scale parameter which may need to be estimated. Thirdly, the conditional within-subject relationship among the vectors of repeated responses, given the covariates, is assumed to be a function of association parameters. The advantage of this approach is that it models the mean and covariance separately, which ensures that the regression parameters of the marginal model have interpretations that do not depend on the assumptions made about the within-subject associations. Although the marginal model can handle the missing data, there are only a limited number of variance-covariance structures to be used for unbalanced data. What is more, if the correlation structure is not specified correctly, the parameter estimators are inefficient. Moreover, the marginal model is unable to differentiate within-subject variability and between-subject variability. Liang and Zeger(1986) improved this model by estimating consistent estimators or regression parameters, and their variances, under relatively mild conditions, which are called Generalized Estimating Equations(GEE). (Throp, 2009)

Another modelling approach is about a mixed-effects or random-effects model (Laird and Ware, 1982). The concept of this approach is that the mean response depends both on the covariates of the interest and on a vector of random effects. Since some subsets of the regression parameters are different from one subject to another, therefore the natural heterogeneity in the population should also be explained. Each subject is assumed to have its own individual changing trajectory across time, and a subset of the regression parameters are treated as random.The mixed-effects model contains both fixed effects and random effects. Here, the fixed effects are usually assumed to be shared by all subjects and the random effects are the part that are unique to a particular subject. Through involving the random effects in the model, a mixed-effects model can make a distinction between within-subject and between-subject. By fitting a mixed-effects model, it is able to describe the general change of the mean response over time through the estimation of the parameters in addition to the particular change trajectory of each subject over time. What's more, the mixed-effects model is often used since it can handle missing data and unbalanced data.

Two basic methods for analyzing longitudinal studies have been discussed above. More approaches can be found in Fitzmaurice (2008) and other researchers' articles.

2.1.3 Covariance structure

One of the challenges for analyzing longitudinal data is that they are correlated since measurements on the same subject are taken repeatedly over time, thus appropriate methods should take the correlation into consideration. In addition, a noticeable feature for the longitudinal data is that for most datasets, they are positively correlated.

Every subject share a common variance matrix and correlation matrix to characterize the average dependence among repeated observations even though the covariance matrix may vary from subject to subject.

Let $\mathrm{C}ov\{Y_i, Y_i\}$ denote the covariance structure between the i^{th} subject. Let $\mathrm{Corr}(Y_{it}, Y_{is})$ represent the correlation between times t and s of i^{th} subject.

The covariance structures that characterize the correlated data within subjects are defines as

$$Cov\{Y_i, Y_i\} = E\{((Y_i) - E(Y_i))((Y_i) - E(Y_i))^T\} \tag{2.1}$$

and its matrix can be defined as follows:

$$Cov(Y_i) = \begin{bmatrix} Var(Y_{i1}) & Cov(Y_{i1}, Y_{i2}) & Cov(Y_{i1}, Y_{i3}) & \dots & Cov(Y_{i1}, Y_{in}) \\ Cov(Y_{i2}, Y_{i1}) & Var(Y_{i2}) & Cov(Y_{i2}, Y_{i3}) & \dots & Cov(Y_{i2}, Y_{in}) \\ \dots & \dots & \dots & \dots & \dots \\ Cov(Y_{in}, Y_{i1}) & Cov(Y_{in}, Y_{i2}) & Cov(Y_{i1}, Y_{i3}) & \dots & Var(Y_{in}) \end{bmatrix} .(2.2)$$

We can also define the correlation matrix Corr(Y_i) as:

$$Corr(Y_i) = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \rho_{23} & \dots & \rho_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{n1} & \rho_{n2} & \rho_{n3} & \dots & 1 \end{bmatrix} . \tag{2.3}$$

where Corr$(Y_{is}, Y_{it}) = \rho_{st} = \rho_{ts} =$ Corr(Y_{it}, Y_{is}) since this matrix is also symmetric.

Many of the covariance structures for the time series data can also be applicable to the longitudinal data. The following are two types of covariance structures and the corresponding correlation matrix. (Guo, 2011)

The first strategy is referred to as the Toeplitz. This method assumes that any pair of responses that are equally separated in time have the same correlation which can be defines as Corr$(Y_{is}, Y_{it}) = \rho_k$ for all s and t; in addition, the variance σ^2 is constant. This structure may be written as:

$$Cov(Y_i) = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \dots & 1 \end{bmatrix}. \tag{2.4}$$

The second approach is the autoregressive model which can also be called as the first-order autoregressive covariance, which is the special case of the Toeplitz covariance.(Fitzmaurice et al., 2004) This method assumes that the correlation among subjects "tails off", which can be defined as Corr$(Y_{is}, Y_{it}) = \rho^{|t-s|}$ for all s and t; in addition, the variance σ^2 is constant. This structure may be written as:

$$Cov(Y_i) = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{bmatrix}. \tag{2.5}$$

Note that this method is used in the proposed model.

2.2 Generalized linear mixed models

The Generalized Linear Mixed Model (GLMM) are often used in longitudinal data analysis. They are considered as an extension of two modelling methods, linear mixed models and generalized linear models. Linear mixed models present a useful extension of linear regression models that they include both fixed effects and random effects since observations are dependent due to kinds of correlation. Random effects can be treated as the reflection of natural heterogeneity for the reason that many factors have not been measured. However, the generalized linear models not only cater to the need of continuous response variables, but meet the discrete response variables.

In this thesis, the GLMM is based on the class of Tweedie exponential dispersion model distributions. In Section 2.3, the general idea of GLMM is presented and Tweedie exponential dispersion models will discussed in Section 2.3.

Let $Y_{it} = (y_{i1}, y_{i2}, \ldots, y_{iT})^T$ be the response for the t^{th} observation within individual or cluster $i, i = 1, 2, \ldots, n$ and $t = 1, 2, \ldots, T$. Let X_{it} and Z_{it} be the vectors of the covariates with fixed effects and random effects, respectively. In addition, we assume the response Y_{it} is conditionally independent, given the fixed and random effects, and follows a distribution in the exponential family.

$$g(\mu_{it}) = X_{it}^T\beta + Z_{it}^T b_i \tag{2.6}$$

where $\mu_{it} = E[Y_{it}|X_i, Z_i, b_i]$ is the conditional mean of Y_{it}, β is a vector of regression coefficients for the fixed effects. The random effects b_i is assumed to follow a certain distribution. Here, $g(\cdot)$ is known as the link function.

2.3 Tweedie exponential dispersion models

Probability distributions are widely used since they are helpful for understanding the behaviour in nature. In practice, usually two types of data are recorded; one is continuous and the other is discrete. Therefore, probability density function or probability mass function could be used. But in nature, some behaviour can take both continuous as well as discrete values, such as rainfall observations. For studying such kind of data, Jørgensen (1997) proposed the Tweedie family of distributions. The Tweedie exponential dispersion distributions is consist of a class of distributions which can be used to deal different responses types (Lall, 2014).

The following definition helps to get the general idea of Tweedie exponential dispersion models. Let $V()$ be the variance function which describes the mean-variance association of the distribution when the dispersion is constant. Let ϕ be a dispersion parameter. Then, if Y follows an exponential dispersion model distribution with mean μ, the variance of Y can be written as

$$V(Y) = \phi V(\mu) \tag{2.7}$$

For the Tweedie distribution, denoted by $Tw_p(\mu, \sigma^2)$ with $E(Y) = \mu$ and $Var(Y) = \phi\mu^p$, where p is a parameter that controls the variance of Tweedie distributions. More details can be found in Jørgensen (1987a).

In order to simplify the derivation of the estimation function for the regression parameter, following Ma and Jørgensen (2007), the probability density function of Tweedie exponential dispersion models can be transformed via the following functions:

$$f_p(y;\mu,\sigma^2) = \begin{cases} c_p(y;\sigma^2)\exp\left\{\frac{1}{\sigma^2}\left(\frac{y\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p}\right)\right\} & \text{if } p \neq 1,2, \\ c_2(y;\sigma^2)\exp\left[-\frac{1}{\sigma^2}\left\{\frac{y}{\mu} + \log(\mu)\right\}\right] & \text{if } p = 2, \\ c_1(y)\exp\left[y\log(\mu) - \mu\right] & \text{if } p = 1, \end{cases}$$

where the explicits expression for $c_p(y;\sigma^2)$ are given by Jørgensen (1987, 1997).

Some popular cases of Tweedie distributions include the following: For $p = 0$, the Tweedie distribution becomes the normal distribution; for $p = 1$, the Tweedie distribution becomes the Poisson distribution, for $p = 2$, the Tweedie distribution becomes the Gamma distribution and for $p = 3$, the Tweedie distribution becomes the inverse Gaussian distribution.

For $1 < p < 2$, the Tweedie distribution correspond to the Compound Possion distribution which accounts for continuous data with exact zeroes.

For $0 < p < 1$, the Tweedie distribution is not defined.

2.4 Joint models of longitudinal data

2.4.1 Introduction

In longitudinal studies, it is common that multiple responses are measured repeatedly over time. Perhaps the most common method is to model each longitudinal outcome independently. However, this does not give full consideriction to the multivariate nature of the data for the following reasons. First, there is likely to be a correlation within or between subjects. Second, the variability for each response seems to be different. The primary goal of joint modelling is to provide a framework to answer questions of scientific interest pertaining to relationships among and between the multiple outcomes and formalize the factors. In order for there to be valid inference, the joint models must be obligated to characterize the correlations within and between each outcome. (Throp, 2009)

Broadly, in longitudinal studies, there are two basic methods. The first approach is to apply a conditioning argument that allows the joint distribution to be factorized into a marginal component and a conditional component, where the conditioning can be done either on the discrete outcome or on the continuous outcome. Another approach is to formulate a joint model for the outcomes directly.(Fitzmaurice et al., (2004))

In the following sections, more details will be described for the joint modelling approach.

2.4.2 Approaches to joint model

Consider modelling two outcomes; let Y_1 and Y_2 denote two outcomes measured on a subject. Here, we restrict our attention to two outcomes, noting that extension to more than two outcomes is straightforward. There has been a great deal of research on joint models to analyze longitudinal data. Several methods will be briefly described below, while more approaches can be found in texts by Fitzmaurice et al., (2008).

2.4.2.1 Conditional models

One strategy is referred to as the conditional model. The idea here is to factorize the density into a product of a marginal and a conditional density in order to avoid specifying a joint distribution for (Y_1, Y_2) directly. That is,

$$f(y_1, y_2) = f(y_1|y_2)f(y_2) = f(y_2|y_1)f(y_1) \tag{2.8}$$

This reduces the problem to the specification of models for each of the outcomes separately, with a marginally specified model for one outcome and a conditionally specified model for the other. More attention should be given to the choice of which outcome is modelled marginally and which outcome is modelled conditionally when both outcomes are considered as longitudinal data, since different choices may lead to very different, even completely opposite, results and conclusions.

Nevertheless, the limitation of this approach is that the marginal inferences can not be attained directly. For example, in Equation (2.8), although the inferences about the marginal evolution of Y_2 can be derived directly, the inference about the characteristics of the marginal distribution of Y_1 which is also needed cannot be derived directly.

In addition, when Y_1 and Y_2 are highly correlated and thought to be manifestations of a common underlying treatment effect, conditioning on one outcome will attenuate the treatment effect.

When the case has two outcomes, only two factorizations are needed to be calculated. However, this methodology is not the preferred choice when there are more than two outcomes since more factorizations must be taken into account.

2.4.2.2 Shared-parameter models

Another strategy is the shared-parameter model. As described earlier, random effects can be used to account for the associations in the longitudinal studies. The same concept can be applied to generate the additional association for the multivariate longitudinal studies. Shared-parameter models can be used.

Define b as a vector of random effects, let Y_1 and Y_2 remain the specific model and assume both outcomes are independent, conditionally on b. Then, the joint density for (Y_1, Y_2) is

$$f(y_1, y_2) = \int f(y_1, y_2|b)f(b)db = \int f(y_1|b)f(y_2|b)f(b)db \tag{2.9}$$

where $f(b)$ denotes the marginal density of the random effects. Here, the random effects b which is a "shared-parameter" causes a correlation between $Y_1$and Y_2 since they both depend on b. That Y_1 and Y_2 are conditionally independent can be treated as a common set of underlying characteristics of the subject that governs the processes of both responses.

One advantage of this approach is that Y_1 and Y_2 can be two different types of data. For example, Y_1 could be the discrete response and Y_2 could be the continuous response, but they share the same random effects. Another advantage is that each parameter has the same interpretation both in the joint model and "univariate" model. In addition, extensions to more than two responses are straightforward and do not increase the computational intensity.

The main disadvantage of this model is that it implies a very strong assumption about the association between the outcomes. Specifically, the product of the correlation between measurements from the first outcome and measurements from the second outcome must be the same even the correlation between pairs of measurements are from different outcomes(Fitzmaurice et al., (2008)). That is

$$Corr\{Y_1(s), Y_2(s)\} \leq \sqrt{corr\{Y_1(s), Y_1(t)\}}\sqrt{corr\{Y_2(s), Y_2(t)\}} \tag{2.10}$$

As such, this approach is limited to representing the association structure of the multiple longitudinal outcomes. This provides the motivation for the more flexible correlation patterns, even if the model is more complex.

Chapter 3

Joint Model of Longitudinal Data

3.1 Model Specification

Let Y_{ijt} represent the jth ($j = 1, 2, \ldots, J$) longitudinal response recorded at the tth ($t = 1, 2, \ldots, T$) time point of the ith ($i = 1, 2, \ldots, m$) subject. Let $Y = (Y_{111}, \ldots, Y_{11T}, Y_{121}, \ldots, Y_{12T}, \ldots, Y_{1J1}, \ldots, Y_{1JT}, \ldots, Y_{mJ1}, \ldots, Y_{mJT})'$ denote the vector of the responses. Let $Y_{ij} = (Y_{ij1}, Y_{ij2}, \ldots, Y_{ijT})'$, $Y_i = ({Y_{i1}}', {Y_{i2}}', \ldots, {Y_{iJ}}')'$ and $Y = ({Y_1}', {Y_2}', \ldots, {Y_m}')'$. We consider the subject-specific random effects U_i for the ith subject and time-specific random effect V_{ijt} for the response from the jth longitudinal response at the tth time point of the ith subject. Let $W = (U', V')'$ denote the vector of the random effects where $U = (U_1, \ldots, U_i, \ldots, U_m)'$ and $V = (V'_{11}, V'_{12} \ldots, V'_{iJ}, V'_{21}, V'_{22}, \ldots, V'_{2J}, \ldots, V'_{m1}, \ldots, V'_{mJ})'$ with $V_i = (V_{ij1}, \ldots, V_{ijt}, \ldots, V_{iJT})'$ and $V_{ij} = (V_{ij1}, V_{ij2}, \ldots, V_{ijT})'$ respectively. The model is characterized by the following three assumptions.

3.1.1 Assumption 1

Subject-specific random effects $U_1, \ldots, U_i, \ldots, U_m$ are positive, independently and identically distributed with mean 1 and variance σ^2. That is

$$E(U_i) = 1 \quad \text{and} \quad \text{Var}(U_i) = \sigma^2.$$

3.1.2 Assumption 2

Given the subject-specific random effects U, the moment structure of the positive time-specific random effects $V_{111}, \ldots, V_{ijt}, \ldots, V_{mJT}$ can be expressed as

$$E(V_{ijt}|U) = U_i \qquad \text{and} \qquad \text{Cov}(V_{ijt}, V_{i'j't'} \mid U) \quad = \quad \begin{cases} \tau_j^2 \rho_{j(t,t')} U_i & \text{if } i = i' \quad \text{and} \quad \text{j} = j' \\ 0 & \text{if } i \neq i' \quad \text{or} \quad \text{j} \neq j' \end{cases},$$

with $\rho_{j(t,t')} = 1$ for $t = t'$ and $j = j'$. This correlation structure of the random effect is flexible as it can accommodate various correlation structures including exchangeable, AR(1) and unstructured correlation structures.

3.1.3 Assumption 3

Given the random effects W, the components of Y are conditionally independent, and the conditional distribution of Y_{ijt}, given W, depends on V_{ijt} only, which is

$$Y_{ijt} \mid W \quad \sim \quad \text{Tw}_{p_j}(\mu_{ijt} V_{ijt}, \epsilon_j^2 {V_{ijt}}^{1-p_j}) \tag{3.1}$$

where $\mu_{ijt} = \exp(x'_{ijt}\beta_j)$ with vector of covariates x_{ijt} and regression parameter vector β_j.

In (3.1), Tw_{p_j} is the Tweedie exponential family with index parameter p. This Tweedie family can also be called as the power-variance family with $E(Y_{ijt} \mid W) = \mu_{ijt}V_{ijt}$ and $Var(Y_{ijt} \mid W) = \epsilon_j^2 {V_{ijt}}^{1-p_j} \mu_{ijt}^{p_j} V_{ijt}^{p_j} = \epsilon_j^2 \mu_{ijt}^{p_j} V_{ijt}$.

3.2 Moment structure

The moment structure of the model is based on the method of conditioning on random effects through some algebraic calculations. The unconditional mean and covariance are presented to facilitate the parameter estimation.

The unconditional expectation of the response Y_{ijt} can be expressed as:

$$E(Y_{ijt}) = EE(Y_{ijt} \mid W) = \mu_{ijt}E(U_i) = \mu_{ijt} \tag{3.2}$$

The unconditional variance of the responses Y_{ijt} has the form:

$$\begin{aligned}
\text{Var}(Y_{ijt}) &= E\{\text{Var}(Y_{ijt} \mid W)\} + \text{Var}\{E(Y_{ijt} \mid W)\} \\
&= {\epsilon_j}^2 \mu_{ijt}^{p_j} E(V_{ijt}) + \mu_{ijt}^2 \text{Var}(V_{ijt}) \\
&= {\epsilon_j}^2 \mu_{ijt}^{p_j} E(V_{ijt}) + \mu_{ijt}^2 \{\text{Var}(E(V_{ijt} \mid U)) + E(\text{Var}(V_{ijt} \mid U))\} \\
&= {\epsilon_j}^2 \mu_{ijt}^{p_j} + \mu_{ijt}^2(\sigma^2 + \tau_j^2) \qquad (3.3)
\end{aligned}$$

Since all the derivations of these moment structures are similiar, the covariance between the response Y_{ijt} and $Y_{i't'j'}$ is:

$$\text{Cov}(Y_{ijt}, Y_{i'j't'}) = E\{\text{Cov}(Y_{ijt}, Y_{i'j't'} \mid W)\} + \text{Cov}\{E(Y_{ijt} \mid W), E(Y_{i'j't'} \mid W)\} \tag{3.4}$$

1). $if\ i = i',\ j = j',\ t = t'\ then:$

$$\begin{aligned}\text{Cov}(Y_{ijt}, Y_{ijt}) &= Var(Y_{ijt}) \\ &= {\epsilon_j}^2 \mu_{ijt}^{p_j} + \mu_{ijt}^2(\sigma^2 + \tau_j^2)\end{aligned} \tag{3.5}$$

2). $if\ i = i',\ j = j',\ t \neq t'\ then:$

$$\begin{aligned}\text{Cov}(Y_{ijt}, Y_{ijt'}) &= E\{\text{Cov}(Y_{ijt}, Y_{ijt'} \mid W)\} + \text{Cov}\{E(Y_{ijt} \mid W), E(Y_{ijt'} \mid W)\} \\ &= \mu_{ijt}\mu_{ijt'}(\sigma^2 + \tau_j^2 \rho_{j(t,t')})\end{aligned} \tag{3.6}$$

3). $if\ i = i',\ j \neq j'\ then:$

$$\begin{aligned}\text{Cov}(Y_{ijt}, Y_{ij't'}) &= E\{\text{Cov}(Y_{ijt}, Y_{ij't'} \mid W)\} + \text{Cov}\{E(Y_{ijt} \mid W), E(Y_{ij't'} \mid W)\} \\ &= 0 + Cov(\mu_{ijt}V_{ijt}, \mu_{ij't'}V_{ij't'}) \\ &= \mu_{ijt}\mu_{ij't'}Cov(V_{ijt}, V_{ij't'}) \\ &= \mu_{ijt}\mu_{ij't'}\sigma^2\end{aligned} \tag{3.7}$$

4). *Otherwise*

$$\mathrm{Cov}(Y_{ijt}, Y_{i'j't'}) = 0 \tag{3.8}$$

Therefore, the unconditional covariance structure of the response is obtained as follows:

$$\mathrm{Cov}(Y_{ijt}, Y_{i'j't'}) = \begin{cases} {\epsilon_j}^2 \mu_{ijt}^{p_j} + \mu_{ijt}^2 \left(\sigma^2 + \tau_j^2\right) & \text{if } i = i', j = j', t = t' \\ \mu_{ijt}\mu_{ijt'}(\sigma^2 + \tau_j^2 \rho_{j(t,t')}) & \text{if } i = i', j = j', t \neq t' \\ \mu_{ijt}\mu_{ij't'}\sigma^2 & \text{if } i = i', j \neq j' \\ 0 & \text{otherwise} \end{cases} \tag{3.9}$$

where $\mu_i = ({\mu_{i1}}', {\mu_{i2}}', \ldots, {\mu_{iJ}}')'$ and $\mu_{ij} = (\mu_{ij1}, \mu_{ij2}, \ldots, \mu_{ijT})'$.

The estimating equations for the regression parameters and random effects parameters can be constructed by using the approach of the best linear unbiased predictors.

3.3 The best linear unbiased predictors of random effects

Let the inverse of the covariance matrix of Y be denoted as $\mathrm{Cov}^{-1}(Y)$. The best linear unbiased predictors of random effects (BLUP) of subject-specific random effects

$U_1, \ldots, U_i, \ldots, U_m$ given Y can be predicted as follows:

$$\hat{U} = E(U) + \text{Cov}(U, Y)\text{Cov}^{-1}(Y)\{Y - E(Y)\} \tag{3.10}$$

where the vector $E(U) = (1, \ldots, 1, \ldots, 1)'$ is of dimension $m \times 1$.

Also, diag(μ_{ij}) and $diag(\tau_j^2 R_j)$ are defined as:

$$diag(\mu_{ij}) = \begin{bmatrix} \mu_{ij1} & 0 & \ldots & 0 \\ 0 & \mu_{ij2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \mu_{ijT} \end{bmatrix}, \tag{3.11}$$

$$\text{diag}(\tau_j^2 R_j) = \begin{bmatrix} \tau_1^2 R_1 & 0 & \ldots & 0 \\ 0 & \tau_2^2 R_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \tau_j^2 R_j \end{bmatrix}, \tag{3.12}$$

The covaraince structure between U and Y obtains as:

$$\text{Cov}(U, Y) = \begin{bmatrix} \mu_1'\sigma^2 & 0 & 0 & \ldots & 0 \\ 0 & \mu_2'\sigma^2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \mu_m'\sigma^2 \end{bmatrix}, \tag{3.13}$$

which can be simplified as:

$$\text{Cov}(U, Y) \quad = \quad \sigma^2 diag(\mu_1', \mu_2', \ldots, \mu_m') \tag{3.14}$$

The covaraince structure of Y obtain the matrix as:

$$\text{Cov}(Y) = \begin{bmatrix} Cov(Y_1) & 0 & 0 & \ldots & 0 \\ 0 & Cov(Y_2) & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & Cov(Y_m) \end{bmatrix}, \tag{3.15}$$

where $Cov(Y_i)$ can be expressed as:

$$\text{Cov}(Y_i) = \begin{bmatrix} Var(Y_{i1}) & Cov(Y_{i1}, Y_{i2}) & \ldots & Cov(Y_{i1}, Y_{iJ}) \\ Cov(Y_{i2}, Y_{i1}) & Var(Y_{i2}) & \ldots & Cov(Y_{i2}, Y_{iJ}) \\ \ldots & \ldots & \ldots & \ldots \\ Cov(Y_{iJ}, Y_{i1}) & Cov(Y_{iJ}, Y_{i2}) & \ldots & Var(Y_{iJ}) \end{bmatrix}. \tag{3.16}$$

$Cov(Y_{ij}, Y_{ij})$ obtains the matrix as:

$$\text{Cov}(Y_{ij}, Y_{ij}) = \begin{bmatrix} Var(Y_{ij1}) & Cov(Y_{ij1}, Y_{ij2}) & \ldots & Cov(Y_{ij1}, Y_{ijT}) \\ Cov(Y_{ij2}, Y_{ij2}) & Var(Y_{ij2}) & \ldots & Cov(Y_{ij2}, Y_{ijT}) \\ \ldots & \ldots & \ldots & \ldots \\ Cov(Y_{ijT}, Y_{ij1}) & Cov(Y_{ijT}, Y_{ij2}) & \ldots & Var(Y_{ijT}) \end{bmatrix}, \tag{3.17}$$

and this matrix can be simplified as:

$$\text{cov}(Y_{ij}, Y_{ij}) = \epsilon_j{}^2 diag(\mu_{ij}^{p_j}) + \sigma^2 \mu_{ij} \mu_{ij}{}' + diag(\mu_{ij})(\tau_j^2 R_j) diag(\mu_{ij}) \quad (3.18)$$

$cov(Y_{ij}, Y_{ij'})$ obtains the matrix:

$$\text{cov}(Y_{ij}, Y_{ij'}) = \begin{bmatrix} Cov(Y_{ij1}, Y_{ij'1}) & Cov(Y_{ij1}, Y_{ij'2}) & \ldots & Cov(Y_{ij1}, Y_{ij'T}) \\ Cov(Y_{ij2}, Y_{ij'1}) & Cov(Y_{ij2}, Y_{ij'2}) & \ldots & Cov(Y_{ij2}, Y_{ij'T}) \\ \ldots & \ldots & \ldots & \ldots \\ Cov(Y_{ijT}, Y_{ij'1}) & Cov(Y_{ijT}, Y_{ij'2}) & \ldots & Cov(Y_{ijT}, Y_{ij'T}) \end{bmatrix}, (3.19)$$

which can be simplified as:

$$\text{cov}(Y_{ij}, Y_{ij'}) = \sigma^2 \mu_{ij} \mu_{ij}{}' \quad (3.20)$$

Similiarly, we can derive the orthodox BLUP of time-specific random effects V given the response Y as follows:

$$\begin{aligned} \hat{V} &= E(V) + \text{Cov}(V, Y)\text{Cov}^{-1}(Y)\{Y - E(Y)\} \\ &= E(V) + \text{Cov}(V)\mathbf{D}'\text{Cov}^{-1}(Y)\{Y - E(Y)\} \quad (3.21) \end{aligned}$$

where $E(V)$ is a vector of length mJT which can be defined as $(1, \ldots, 1, \ldots, 1)'$ and $Cov(V)$ is a $mJT \times mJT$ matrix with elements $Cov(V_{ijt}, V_{i'j't'})$ which can be expressed as follows:

$$\begin{aligned}\mathrm{Cov}(V_{ijt}, V_{i'j't'}) &= E\{\mathrm{Cov}(V_{ijt}, V_{i'j't'} \mid \boldsymbol{U})\} + \mathrm{Cov}\{E(V_{ijt} \mid \boldsymbol{U}), E(V_{i'j't'} \mid \boldsymbol{U})\} \\ &= E\{\mathrm{Cov}(V_{ijt}, V_{i'j't'} \mid \boldsymbol{U})\} + Cov(U_i, U_{i'}) \qquad (3.22)\end{aligned}$$

$1^\circ \; if \;\; i = i', \; j = j', \; t = t' \; then:$

$$\begin{aligned}\mathrm{Cov}(V_{ijt}, V_{ijt}) &= E\{\mathrm{Cov}(V_{ijt}, V_{i'j't'} \mid \boldsymbol{U})\} + \sigma^2 \\ &= E({\tau_j}^2 \rho_{j(t,t')} U_i) + \sigma^2 \\ &= {\tau_j}^2 \times 1 \times E(U_i) + \sigma^2 \\ &= {\tau_j}^2 + \sigma^2 \qquad (3.23)\end{aligned}$$

$2^\circ \; if \;\; i = i', \; j = j', \; t \neq t' \; then:$

$$\begin{aligned}\mathrm{Cov}(V_{ijt}, V_{ijt'}) &= E\{\mathrm{Cov}(V_{ijt}, V_{ijt'} \mid \boldsymbol{U})\} + \sigma^2 \\ &= E({\tau_j}^2 \rho_{j(t,t')} U_i) + \sigma^2 \\ &= {\tau_j}^2 \rho_{j(t,t')} + \sigma^2 \qquad (3.24)\end{aligned}$$

$3^\circ \; If \;\; i = i', \; j \neq j' \; then:$

$$\begin{aligned}\mathrm{Cov}(V_{ijt}, V_{ij't'}) &= E\{\mathrm{Cov}(V_{ijt}, V_{ij't'} \mid \boldsymbol{U})\} + \sigma^2 \\ &= \sigma^2 \qquad (3.25)\end{aligned}$$

4° *Otherwise*

$$\text{Cov}(V_{ijt}, V_{i'j't'}) \;=\; 0 \tag{3.26}$$

which has the general form:

$$\text{Cov}(V_{ijt}, V_{i'j't'}) \;=\; \begin{cases} {\tau_j}^2 + \sigma^2 & \text{if } i = i', j = j', t = t' \\ {\tau_j}^2 \rho_{j(t,t')} + \sigma^2 & \text{if } i = i', j = j', t \neq t' \\ \sigma^2 & \text{if } i = i', j \neq j' \\ 0 & \text{otherwise} \end{cases} . \tag{3.27}$$

The matrix of D in (3.21) is an $mJT \times mJT$ diagonal matrix of the vector $(\mu_{111}, \ldots, \mu_{ijt}, \ldots, \mu_{mJT})'$.

The estimating equations for the regression parameters can be constructed by using the linear predictors of V and Y.

The covaraince structure of V and Y obtains:

$$\text{Cov}(V, Y) = \begin{bmatrix} Cov(V_1, Y_1) & 0 & 0 & \ldots & 0 \\ 0 & Cov(V_2, Y_2) & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & Cov(V_m, Y_m) \end{bmatrix}, \tag{3.28}$$

In (3.21), D has already been defined as a diagonal matrix and $Cov(V, Y) =$

$C0v(V)D'$, so the matrix of $Cov(V_i, V_i)$ can be expressed as:

$$\mathrm{Cov}(V_i, V_i) = \begin{bmatrix} Cov(V_{i1}, V_{i1}) & Cov(V_{i1}, V_{i2}) & \dots & Cov(V_{i1}, V_{iJ}) \\ Cov(V_{i2}, V_{i1}) & Cov(V_{i2}, V_{i2}) & \dots & Cov(V_{i2}, V_{iJ}) \\ \dots & \dots & \dots & \dots \\ Cov(V_{iJ}, V_{i1}) & Cov(V_{iJ}, V_{i2}) & \dots & Cov(V_{iJ}, V_{iJ}) \end{bmatrix}, \quad (3.29)$$

According to the correlation between V_i and V_i we have derived, the matrix also equals:

$$\mathrm{Cov}(V_i, V_i) = \begin{bmatrix} Cov(V_{i1}, V_{i1}) & Cov(V_{i1}, V_{i2}) & \dots & Cov(V_{i1}, V_{iJ}) \\ Cov(V_{i2}, V_{i1}) & Cov(V_{i2}, V_{i2}) & \dots & Cov(V_{i2}, V_{iJ}) \\ \dots & \dots & \dots & \dots \\ Cov(V_{iJ}, V_{i1}) & Cov(V_{iJ}, V_{i2}) & \dots & Cov(V_{iJ}, V_{iJ}) \end{bmatrix}, \quad (3.30)$$

where the matrix of $Cov(V_{ij}, V_{ij})$ can be expressed as:

$$\mathrm{Cov}(V_{ij}, V_{ij}) = \begin{bmatrix} Cov(V_{ij1}, V_{ij1}) & Cov(V_{ij1}, V_{ij2}) & \dots & Cov(V_{ij1}, V_{ijT}) \\ Cov(V_{ij2}, V_{ij1}) & Cov(V_{ij2}, V_{ij2}) & \dots & Cov(V_{ij2}, V_{ijT}) \\ \dots & \dots & \dots & \dots \\ Cov(V_{ijT}, Y_{ij1}) & Cov(V_{ijT}, V_{ij2}) & \dots & Cov(V_{ijT}, V_{ijT}) \end{bmatrix}, \quad (3.31)$$

which can be simplified as:

$$\begin{aligned}\text{Cov}(V,Y) &= \left\{\mathbf{1}_m \otimes (diag(\tau_1^2 R_1, \tau_2^2 R_2, \ldots, \tau_J^2 R_J) + \sigma^2 \mathbf{1}_{JT}\mathbf{1}'_{JT})\right\}. \\ &\qquad \times \left\{diag(\mu_{ij1}, \mu_{ij2}, \ldots, \mu_{ijT})\right\}, \end{aligned} \tag{3.32}$$

3.4 Estimation of Parameters

3.4.1 Estimation of regression parameters

First, we consider the estimation for the regression parameters in the case of known random effect parameters. To do that, following Ma and Jørgensen (2007), we differentiate the partially observed 'joint' log-likelihood of the Tweedie mixed model for the data and random effects with respect to $\boldsymbol{\beta}$ which yields the partially observed 'joint' score function. By replacing the random effects with their BLUP predictors, we obtain an unbiased estimating function for the regression parameters $\boldsymbol{\beta}$, which can be expressed as:

$$\psi(\boldsymbol{\beta}) = \sum_{i=1}^{m}\sum_{j=1}^{J}\sum_{t=1}^{T} \boldsymbol{X}'_{ijt} \frac{\mu_{ijt}^{1-p_j}(\boldsymbol{\beta})}{\epsilon_j^2}\left[y_{ijt} - \hat{V}_{ijt}(\boldsymbol{\beta})\mu_{ijt}(\boldsymbol{\beta})\right] \tag{3.33}$$

where the ith component corresponds to the ith individual. Without loss of generality, we assume that the matrix of the covariates is full rank.

As indicated by Ma and Jørgensen (2007), under mild regularity conditions, the solution of the estimating equation $\psi(\boldsymbol{\beta}) = 0$ is consistent and asymptotically normal with asymptotic mean $\boldsymbol{\beta}$ and asymptotic variance given by the inverse of the sensitivity matrix $\mathrm{S}(\boldsymbol{\beta}) = \mathrm{E}_{\boldsymbol{\beta}}\{\partial\psi(\boldsymbol{\beta})/\partial\boldsymbol{\beta}\}$ as the subjects are assumed to be independent. Thus, this estimating equation $\psi(\boldsymbol{\beta}) = 0$ is optimal in the sense that it attains the minimum asymptotic covariance for the estimator of $\boldsymbol{\beta}$ within the class of all linear functions of Y (Ma and Jørgensen, 2007). Furthermore, as in Ma et al.(2009), this estimation equation $\psi(\boldsymbol{\beta}) = 0$ can be solved iteratively using the scoring algorithm, where the value of $\boldsymbol{\beta}$ is updated as follows:

$$\boldsymbol{\beta}^* = \boldsymbol{\beta} - \mathrm{S}^{-1}(\boldsymbol{\beta})\psi(\boldsymbol{\beta}),$$

with the explicit expression of the sensitivity matrix given by

$$\mathrm{S}(\boldsymbol{\beta}) = \boldsymbol{X}'\mathrm{diag}\left\{E(\mathbf{Y})\right\}\mathrm{cov}^{-1}(\mathbf{Y})\mathrm{diag}\left\{E(\mathbf{Y})\right\}\boldsymbol{X} \tag{3.34}$$

where

$$E(\mathbf{Y}) = \{E(Y_{111}), \ldots, E(Y_{mJT})\}'$$

and

$$\psi(\boldsymbol{\beta}) = \boldsymbol{X}^{\boldsymbol{T}} diag\boldsymbol{E}(\boldsymbol{Y})cov^{-1}(\mathbf{Y})\{\mathbf{Y} - E(\mathbf{Y})\}$$

.

3.4.2 Estimation of random effects parameters

To estimate the regression parameter β, in the previous subsection we assumed the random effect parameters are known. In this section, we present a moment approach to estimate the unknown random effect parameters σ^2, τ_j^2 and ϵ_j^2. To do that we assume that the correlation structure of the random effects are known. In next subsection we present the estimation of the correlation parameters under various underlying correlation structures. To estimate subject-specific random effect variation σ^2 we use the BLUPs of the random effect $\hat{U}$. After some algebraic calculation, the iterative equation for estimating σ^2 can be expressed as

$$\hat{\sigma}_r^2 = \frac{1}{m} \sum_{i=1}^{m} \left\{ (\hat{U}_i - 1)^2 + \hat{\sigma}_{r-1}^2 - \hat{\sigma}_{r-1}^4 \; \mu_i{}' \text{var}(\, Y_i)^{-1} \; \mu_i \right\}, \tag{3.35}$$

where $\hat{\sigma}_{r-1}^2$ is the estimate from the previous iteration.

Similarly, the iterative equation for estimating τ_j^2 and ϵ_j^2 can be expressed as

$$\begin{aligned} \hat{\tau}_{j,r}^2 \;\; = \;\; & \frac{1}{mT} \sum_{i=1}^{m} \sum_{t=1}^{T} \Big\{ (\hat{V}_{ijt} - \hat{U}_i)^2 + \hat{\tau}_{j,r-1}^2 - [\sigma^4 \; \mu_{ij}{}' \text{var}(\, Y_{ij})^{-1} \; \mu_{ij} + \text{cov}(V_{ijt}, \; Y_{ij}) \\ & \text{var}(\, Y_{ij})^{-1} \text{cov}(\, Y_{ij}, V_{ijt}) - 2\text{cov}(U_i, \; Y_{ij}) \text{var}(\, Y_{ij})^{-1} \text{cov}(\, Y_{ij}, V_{ijt})] \Big\}, \end{aligned} \tag{3.36}$$

and

$$\hat{\epsilon}^2_{j,r} = \frac{1}{mT}\sum_{i=1}^{m}\sum_{t=1}^{T}\frac{1}{\mu^p_{ijt}}\Big\{(y_{ijt}-\mu_{ijt}\hat{V}_{ijt})^2+\mu^2_{ijt}(\sigma^2+\tau^2_j)$$
$$-\mu^2_{ijt}\text{cov}(V_{ijt},\ Y_{ij})\text{var}(\ Y_{ij})^{-1}\text{cov}(\ Y_{ij},V_{ijt})\Big\}, \qquad (3.37)$$

respectively. The explicit forms for the estimators of σ^2, τ^2_j and ϵ^2_j are similar to those presented in Ma and Jørgensen (2007). In next subsection we present the estimation of the correlation parameters $\rho_{j(t,t')}$.

3.4.3 Estimation of correlation parameter under flexible correlation structures

In this section, we present a moment approach to estimate $\rho_{j,(t,t')}$ under flexible correlation structure which can be expressed as

$$\Re_j = \left[\rho_{j,(t,t')}\right]_{T\times T} = \begin{bmatrix} 1 & \rho_{j,(1,2)} & \rho_{j,(1,3)} & \cdots & \rho_{j,(1,T-1)} \\ \rho_{j,(2,1)} & 1 & \rho_{j,(2,3)} & \cdots & \rho_{j,(2,T-1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho_{j,(T-1,1)} & \rho_{j,(T-1,2)} & \rho_{j,(T-1,3)} & \cdots & 1 \end{bmatrix}. \qquad (3.38)$$

For unstructured correlation structure $\rho_{(t,t')}$ is

$$\begin{aligned}\rho_{j,(t,t')} &= \frac{\text{cov}\left\{(V_{ijt}-U_i)(V_{ijt'}-U_i)\right\}}{\left[\left\{\text{var}(V_{ijt}-U_i)\right\}\left\{\text{var}(V_{ijt'}-U_i)\right\}\right]^{1/2}} \\ &= \frac{\text{cov}\left\{(\hat{V}_{ijt}-\hat{U}_i)(\hat{V}_{ijt'}-\hat{U}_i)+b_{j,i}(t,t')\right\}}{\left[\left\{\text{var}(\hat{V}_{ijt}-\hat{U}_i)+b_{j,i}(t,t)\right\}\left\{\text{var}(\hat{V}_{ijt'}-\hat{U}_i)+b_{j,i}(t',t')\right\}\right]^{1/2}},\end{aligned} \tag{3.39}$$

where $b_{j,i}(t,t')$ is the correction term which can be simplified as

$$\begin{aligned}b_{j,i}(t,t') &= \rho_{j,(t,t')}\tau_j^2 - \Big\{\text{cov}(\hat{V}_{ijt},\hat{V}_{ijt'}) + \sigma^2\ \mu_{ij}{}'\text{var}(\ Y_{ij})^{-1}\left[\sigma^2\ \mu_{ij}\right] \\ &\quad -\text{cov}(\ Y_{ij},V_{ijt}) - \text{cov}(\ Y_{ij},V_{ijt'})]\Big\}.\end{aligned}$$

In (3.41), $\rho_{j,(t,t')}$ can be estimated using adjusted Pearson estimator as

$$\hat{\rho}_{j,(t,t')} = \frac{\sum_{i=1}^{m}\left\{(\hat{V}_{ijt}-\hat{U}_i)(\hat{V}_{ijt'}-\hat{U}_i)+b_{j,i}(t,t')\right\}}{\left[\left\{\sum_{i=1}^{m}(\hat{V}_{ijt}-\hat{U}_i)^2+b_{j,i}(t,t)\right\}\left\{\sum_{i=1}^{m}(\hat{V}_{ijt'}-\hat{U}_i)^2+b_{j,i}(t',t')\right\}\right]^{1/2}}, \tag{3.40}$$

For unstructured correlation structure $\rho_{(t,t')}$ is

$$\begin{aligned}
\rho_{j,(t,t')} &= \frac{\operatorname{cov}\{(V_{ijt} - U_i)(V_{ijt'} - U_i)\}}{[\{\operatorname{var}(V_{ijt} - U_i)\}\{\operatorname{var}(V_{ijt'} - U_i)\}]^{1/2}} \\
&= \frac{\operatorname{cov}\left\{(\hat{V}_{ijt} - \hat{U}_i)(\hat{V}_{ijt'} - \hat{U}_i) + b_{j,i}(t,t')\right\}}{\left[\left\{\operatorname{var}(\hat{V}_{ijt} - \hat{U}_i) + b_{j,i}(t,t)\right\}\left\{\operatorname{var}(\hat{V}_{ijt'} - \hat{U}_i) + b_{j,i}(t',t')\right\}\right]^{1/2}},
\end{aligned} \tag{3.41}$$

Many covariance pattern models are available as options. In this thesis, the autoregressive of order 1 (AR(1)) structure is used.

For the autoregressive of order 1 (AR(1)) structure, we can set $\rho_{j(t,t')} = \rho_j^{|t-t'|}$ for any $t \neq t'$. Under the AR(1) structure, the correlation matrix has the form

$$\mathfrak{R}_j = \begin{bmatrix} 1 & \rho_j & \rho_j^2 & \dots & \rho_j^{T-1} \\ \rho_j & 1 & \rho_j & \dots & \rho_j^{T-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_j^{T-1} & \rho_j^{T-2} & \rho_j^{T-2} & \dots & 1 \end{bmatrix}. \tag{3.42}$$

To estimate ρ_j under AR(1) structure, it would be sufficient to estimate lag 1 ($\rho_j^1 = \rho_j$) correlation only which can be obtained from (3.41) as:

$$\begin{aligned}
&\hat{\rho}_j \\
&= \frac{\sum_{i=1}^{m}\sum_{t=1}^{T-1}\left\{(\hat{V}_{ijt} - \hat{U}_i)(\hat{V}_{ij(t+1)} - \hat{U}_i) + b_{j,i}(t,t+1)\right\}}{\left[\left\{\sum_{i=1}^{m}\sum_{t=1}^{T-1}(\hat{V}_{ijt} - \hat{U}_i)^2 + b_{j,i}(t,t)\right\}\left\{\sum_{i=1}^{m}\sum_{t=1}^{T-1}(\hat{V}_{ij(t+1)} - \hat{U}_i)^2 + b_{j,i}(t+1,t+1)\right\}\right]^{1/2}}.
\end{aligned} \tag{3.43}$$

where $b_{j,i}(t,t')$ is the correction term which can be simplified as

$$\begin{aligned} b_{j,i}(t,t') &= \rho_{j,(t,t')}\tau_j^2 - \Big\{\mathrm{cov}(\hat{V}_{ijt}, \hat{V}_{ijt'}) + \sigma^2\ \mu_{ij}{}'\mathrm{var}(\ Y_{ij})^{-1}\big[\sigma^2\ \mu_{ij} \\ &\quad -\mathrm{cov}(\ Y_{ij}, V_{ijt}) - \mathrm{cov}(\ Y_{ij}, V_{ijt'})\big]\Big\}\,. \end{aligned}$$

.

Chapter 4

Data Analysis

The proposed model for analyzing the longitudinal data with different types of responses based on conventional Tweedie models is demonstrated by a motivating example in section 4.1.

The example, which is called myoelectric activity study in ponies, is to compare the effects of the nine drugs and a placebo on the patterns of myoelectric activity in the intestines of the ponies (Lester et al., 1998a; 1998b; 1998c). The researchers attached electrodes to different areas of the intestines of the six ponies. Repeated responses of the spike burst rate and duration were collected at equally spaced time intervals around the time of each drug administration. The purpose of the study is to evaluate the joint effects on both spike burst rate and duration.

In section 4.2 and 4.3, we focus on the analysis of the myoelectric activity study in ponies using the proposed model and the results from the study are also discussed. A univariate mixed effects model will also be conducted to compare the difference of the two methods. The simulation process will be presented in section 4.4.

4.1 Myoelectric activity study in ponies

The Myoelectric activity in ponies dataset was selected because it contains two different types of responses. Also, it is considered as a longitudinal dataset since the the two responses were recorded repeatedly over time. Our attention will be restricted to the placebo and one of the 9 drugs. Every pony had three treatments in random order. Two types of outcomes were recorded; the first response is the spike burst rate which is considered as a count variable reflecting the number of contractions exceeding a certain threshold in 15 minutes; the second response is the duration which is considered as a continuous variable reflecting average duration of the contractions in each 15-minute interval. Only 6 ponies were used in the experiment. For computational feasibility, every pony had 12 time points of each drug administration. The main objective of the study is to evaluate the effectiveness of the drug.

For the response of the spike burst rate, we use Poisson distribution, conditional on random effects $W = (U', V')'$, which correspond to $p = 1$. Let Y_{i1t} be the t^{th} spike burst rate on the i^{th} pony; therefore the random effects $U_i, \cdots, U_m$ are the pony-specific random effects and and $V_{i1t}, \cdots, V_{m1t}$ are time-specific random effects for the spike burst rate respectively. The formulation is as follows:

$$Y_{i1t} \mid W \quad \sim \quad \mathrm{Tw}_{p_1}(\mu_{i1t} V_{i1t}, \epsilon_1^2 {V_{i1t}}^{1-p_1})$$

where $p_1 = 1$, $i = 1, \cdots, 6$ and $t = 1, \cdots, 36$.

For the response of the duration, we use Gamma distribution, conditional on random effects $W = (U', V')'$, which correspond to $p = 2$. Let Y_{i2t} be the t^{th} duration measurement on the i^{th} pony; therefore the random effects $U_i, \cdots, U_m$ are the pony-specific random effects and and $V_{i2t}, \cdots, V_{m2t}$ are time-specific random effects for the duration measurement respectively. The formulation is as follows:

$$Y_{i2t} \mid W \quad \sim \quad \mathrm{Tw}_{p_2}(\mu_{i2t} V_{i2t}, \epsilon_2^2 {V_{i2t}}^{1-p_2})$$

where $p_2 = 2$, $i = 1, \cdots, 6$ and $t = 1, \cdots, 36$.

The pony-specific random effects of U_i and time-specific random effects of V_{ijt} as specified in assumptions 1 and 2 respectively are discussed in Chapter 3.

We consider drug; bdur, which denotes the baseline duration; and bcount, which denotes the baseline score measurements as covariates. For the purpose of analysis, the covariate drug was coded as 1 if the active drug is administered and 0 if the placebo is administered. The collection of regression parameters is denoted as $\beta_j = (\beta_{j0}, \beta_{j1}, \beta_{j2})$, where β_{j0} corresponds to the intercept.

The models are fitted using the R software. To obtain the convergence, 1000 iterations have been performed.

4.2 Univariate model analysis

Ma et al.(2015) considered a similar model which also use a generalized linear mixed models based on the Tweedie distributions. This models produces the same assumptions and steps but analyze the responses separately. The results for the separate fits of the two response variables are listed in Table 1.

Table 4.1: Parameter estimates for the pony data based on the univariate mixed effects model

	Spike Burst Rate			Duration		
Parameter	Estimate	SE	P-value	Estimate	SE	P-value
Intercept	3.0414	0.2973	0.0001	0.0732	0.0501	0.2038
Drug effect	0.0335	0.0843	0.7075	**0.1774**	**0.0466**	**0.0125**
Baseline	**0.3108**	**0.0656**	**0.0052**	**0.3804**	**0.0869**	**0.0072**
σ^2	0.0304			0.0030		
τ^2	0.1672			0.0298		
ϵ^2	1			0.0076		
ρ	0.3859			0.7670		

In Table 4.1, the estimates of the regression and random effects parameters for both responses are listed. Examining Table 4.1, a noticeable feature could be found that the drug effects have a significant effect for the response of the duration, but not for the response of the spike burst rate and the baseline covariate were significant to both responses.

4.3 Joint model analysis

In the analyses of this dataset, the first strategy was employed by studying the correlation between the two reponses.

The sample correlation between two vectors of responses Y_1, Y_2 can be expressed as follows:

$$\text{Corr}(Y_1, Y_2) = \frac{\sum_{i=1}^{n}(Y_{i1} - \bar{Y_1})(Y_{i2} - \bar{Y_2})}{\sqrt{\sum_{i=1}^{n}(Y_{i1} - \bar{Y_1})^2 \sum_{i=1}^{n}(Y_{i2} - \bar{Y_2})^2}} \tag{4.1}$$

So, $\text{Corr}(Y_1, Y_2) = 0.3097$ can be calculated from this equation, which indicates that there is a weak relationship between the two responses and that the relationship between the responses is positive represents that if the spike burst rate increases, the duration increases as well.

Similarly, the correlation between two responses of each pony can be attained and the results are shown in Table 4.2.

Table 4.2: Sample correlation between two responses of each pony

Pony	Corr.
1	0.1438
2	0.2263
3	0.4979
4	-0.5735
5	0.1316
6	0.7709

The result from Table 4.2 which is the correlation coefficient that measures the strength and direction between two responses, indicates that for most ponies, there is a positive relationship between two responses. But for the 4^{th} pony, the coefficient of correlation is negative which indicates that for the two responses, an increase in spike burst rate is associated with a decrease in the duration or the opposite trend.

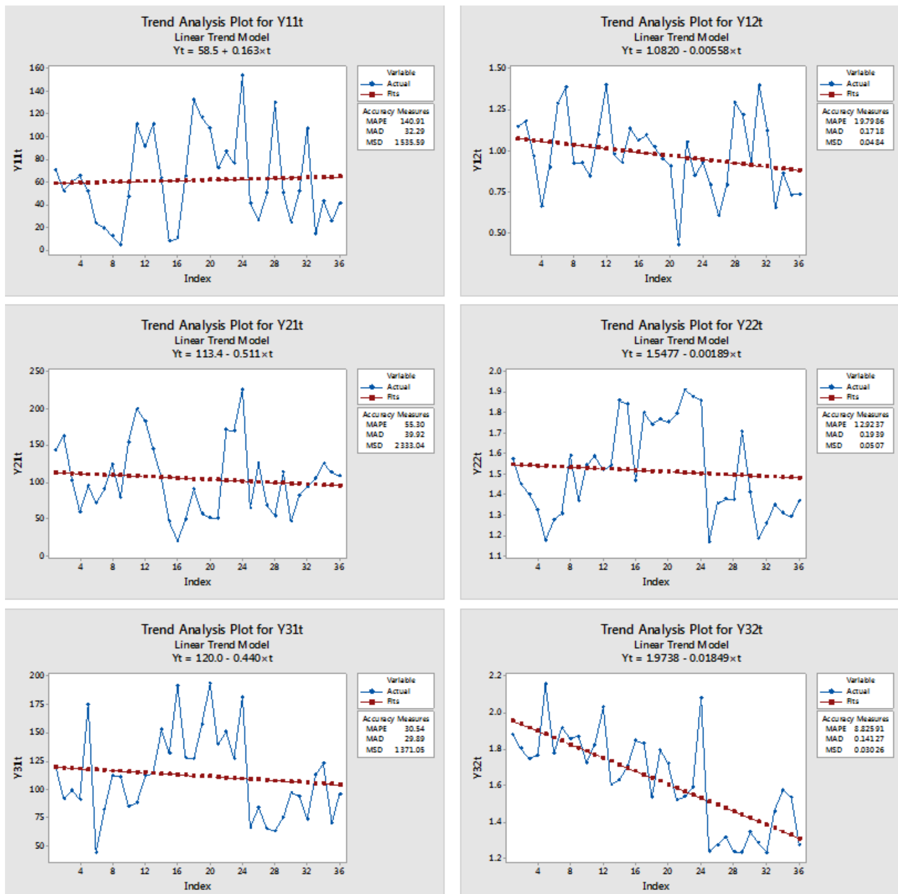

Figure 4.1: Time plot for pony 1-3 of the two response with linear trend

Figure 4.1 and Figure 4.2 explains how the two responses grow as the time changes. Linear trend was used to fit the data. The plots show that for the 1^{st}, 5^{th} and 6^{th} pony, they have the same trend as they have an increasing trend of the spike burst rate and a decreasing trend of the duration while the 4^{th} pony has the

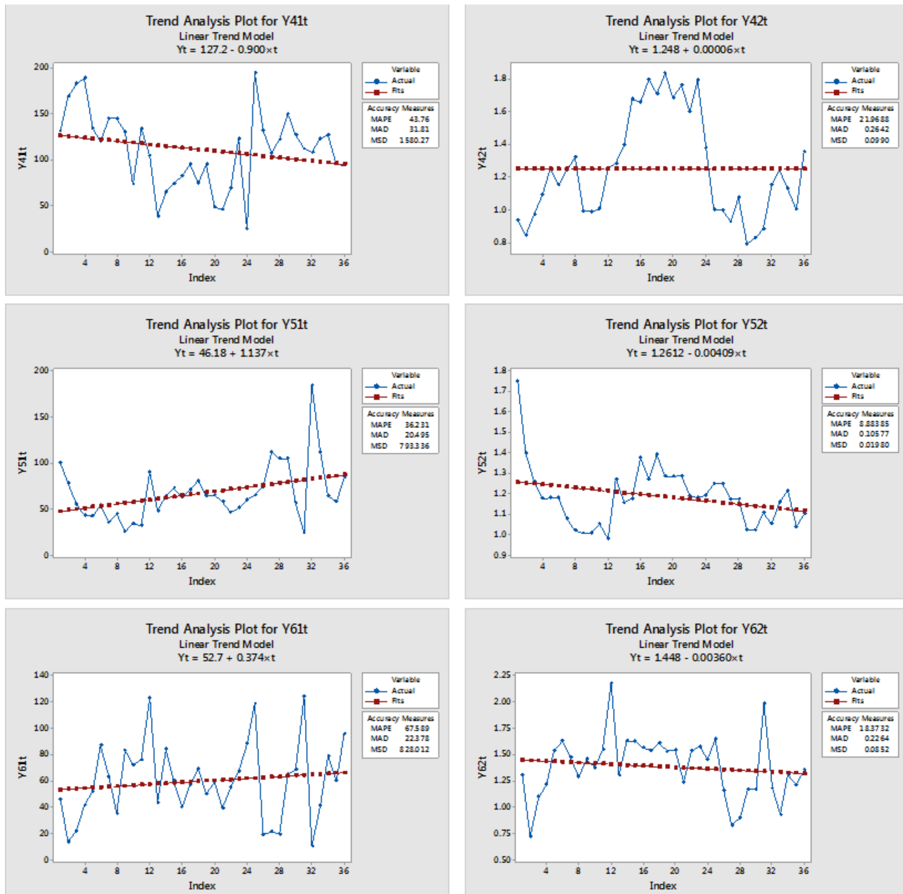

Figure 4.2: Time plot for pony 4-6 of the two response with linear trend

opposite trend. In addition, the 2^{nd} and 3^{rd} pony both have a deceasing trend of the two responses.

The proposed model was fit allowing for analyzing the two responses jointly. The results for the joint analysis are summarized in Table 4.3.

The results are presented in Table 4.3, which reports the estimates and standard errors for all parameters, and the covariates in bold fonts illustrate that they have

Table 4.3: Parameter estimates for the pony data based on the joint Tweedie mixed model

	Spike Burst Rate			Duration		
Parameter	Estimate	SE	P-value	Estimate	SE	P-value
Intercept	2.8655	0.3015	0.0000	0.1014	0.0565	0.0727
Drug effect	0.0486	0.0889	0.5846	**0.1589**	**0.0414**	**0.0001**
Baseline	**0.3488**	**0.0677**	**0.0000**	**0.3194**	**0.0829**	**0.0001**
σ^2	0.0108			0.0108		
τ^2	0.1787			0.0332		
ϵ^2	1.0000			0.0001		
ρ	0.4317			0.5723		

significant effects to the response. The results indicate that significant differences were observed between the active drug and placebo for the duration response. But for spike burst rate response, the drug effects are insignificant. The covariates of the baseline appear to have a significant effect on the duration response but not on the spike burst rate.

The estimates of dispersion and correlation parameters σ^2, τ_1^2, ϵ_1^2 and ρ_1 for the spike burst rate are 0.0108, 0.1787, 1.0000 and 0.4317, respectively. The estimates of dispersion and correlation parameters σ^2, τ_2^2, ϵ_2^2 and ρ_2 for the duration are 0.0108, 0.0332, 0.0001 and 0.5723, respectively. The estimates of the variance parameters σ^2, τ^2 and ϵ^2 indicate that there is additional variation in the responses beyond what can be characterized by the random effect. The values of σ^2 and τ^2 indicate the variation of each subject and of the responses at different measurement times, respectively. The value of ρ indicates the correlation among the repeated measurements of each response. What is more, the first-order autoregressive model is fitted to this dataset.

Figure 4.3 is the scatter plot of the predicted subject-specific random effects of each pony. The plot shows that there is no pony which deviates so much from the other pony so that we need to pay more attention on that pony.

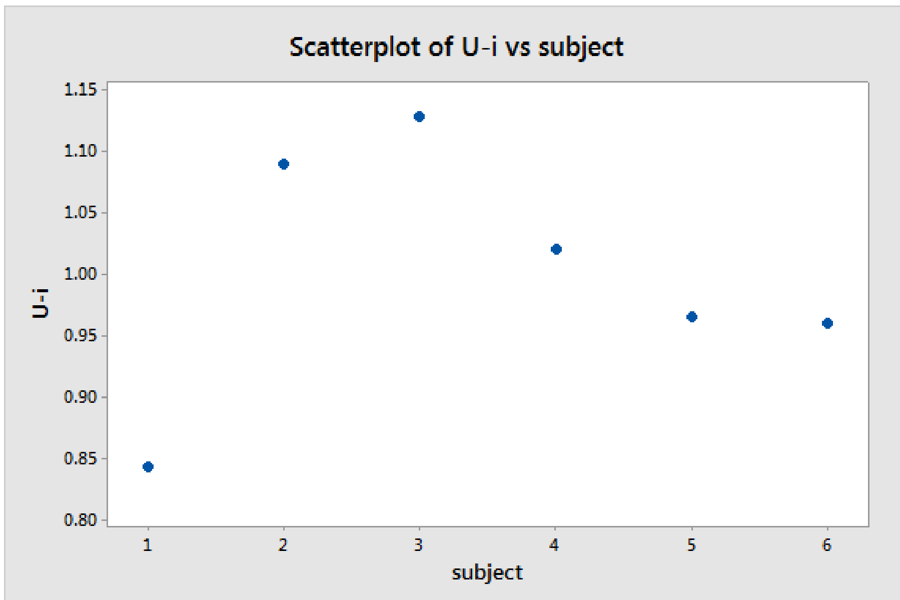

Figure 4.3: Predicted subject-specific random effects of each pony

Therefore, the following part will focus on analyzing the time-specific random effects for both responses.

A similar idea that may be investigated is how the association between the two random effects evolves over time. By definition, the sample correlation between the two sequences of predicted random effects is given by

$$\mathrm{Corr}(\hat{V}_1, \hat{V}_2) = \frac{\sum_{i=1}^{n}(\hat{V}_{i1} - \bar{\hat{V}}_1)(\hat{V}_{i2} - \bar{\hat{V}}_2)}{\sqrt{\sum_{i=1}^{n}(\hat{V}_{i1} - \bar{\hat{V}}_1)^2 \sum_{i=1}^{n}(\hat{V}_{i2} - \bar{\hat{V}}_2)^2}} \tag{4.2}$$

Hence, the sample correlation between the two sequences of predicted random effects over time can be calculated and is 0.4590 which is a increase compared to the correlation between two responses, indicating that the association of the random effects is stronger when we do not consider any other covariates.

The sample correlation between the two sequences of predicted random effects of each pony can also be determined by using equation (4.2). The results are presented in Table 4.4. Compared with the value of correlation between two responses of each pony, the correlation between two random effects for the 1^{st}, 5^{th} and 6^{th} pony increased and the correlation between two random effects for the rest ponies decreased.

Table 4.4: Sample correlation between the two sequences of predicted random effects

Pony	Corr.
1	0.2406
2	0.1898
3	0.3548
4	-0.2121
5	0.4359
6	0.8822

To illustrate the latent characteristic of each pony visually, plots can be conducted in Minitab. Figure 4.4 depicts the latent trend of the two sequences of predicted random effects with linear trend for pony 1- 3. Figure 4.5 depicts the latent trend of the two sequences of predicted random effects with linear trend for the rest ponies.

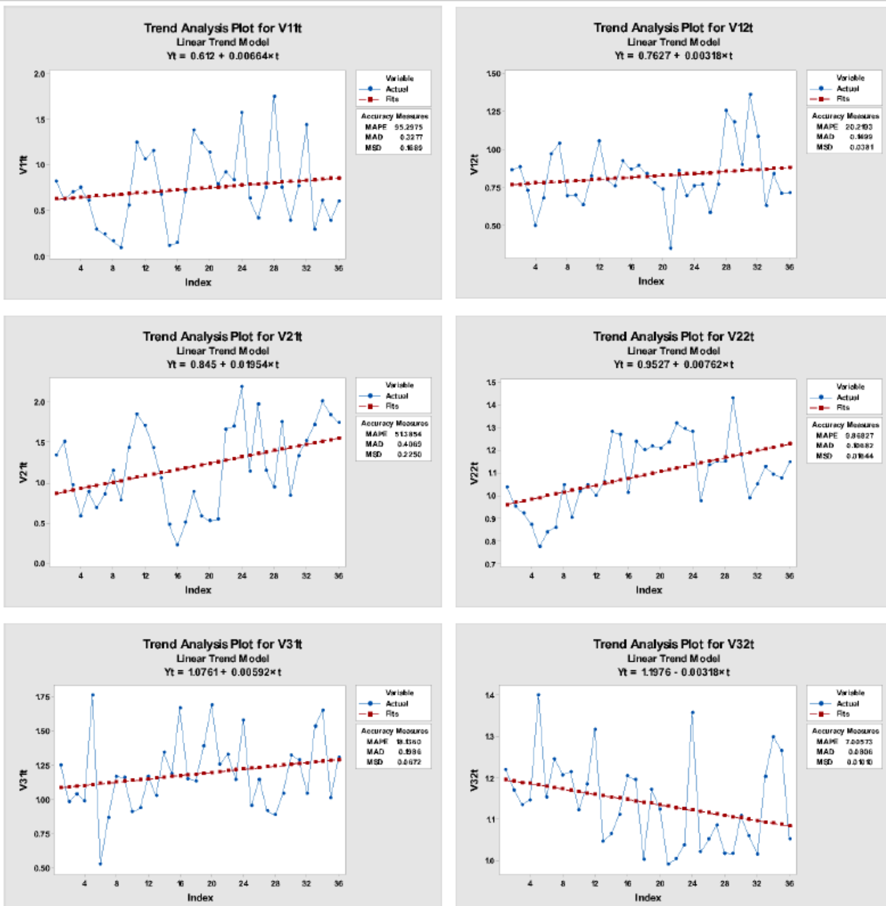

Figure 4.4: Latent trend for pony 1-3 of the two sequences of predicted random effects with linear trend

Through by summarizing the two figures, we can conclude that all ponies have a increasing trend on both random effects except the 3^{rd} and 4^{th} pony have the

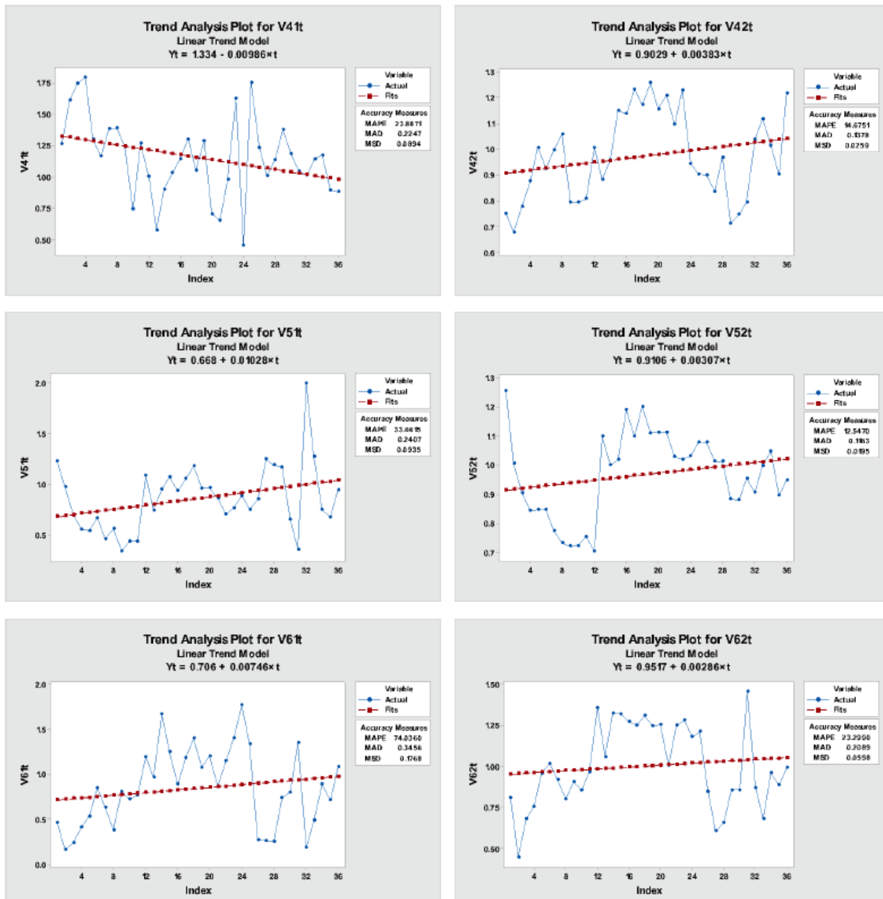

Figure 4.5: Latent trend for pony 4-6 of the two sequences of predicted random effects with linear trend

different trend to different random effects.

These plots can give us a general idea of the latent characteristic of each pony. Therefore, it appears that most ponies have an increasing latent trend to both random effects, which is indicative of a positive correlation. This positive correlation indicates that a pony which has an increasing trend in one response will also have an increasing trend in the other response or conversely. The 3^{rd} and 4^{th} pony have a different trend when compared with both responses which indicate the correlation between two responses is negative. The negative correlation indicates that a pony has an increasing trend to one response but has a decreasing trend to the other response or vice versa.

Now that both approaches for analyzing the responses separately or jointly have been used to fit the data. The regression and random effects parameter estimates from both methods are similar. However, comparing two models, significant difference were found between the active drug and the placebo in the joint model and univariate model. Thus two methods give the same conclusions on the same objective, which is to test the efficacy of the drug.

What is more, we can tell also that the drug effect factor is not the only aspect to analyze, whereas the correlations between measurements on different response variables and correlations between measurements on the same response variable within subject or cluster should be taken into account as they also play the important roles.

Therefore, it may be better to use two random effects rather than a common random effect since it can give us more information.

4.4 Simulation

In this section, we evaluate the performance of the proposed model framework and estimation methods through the simulation studies. To do that, we carried out simulation runs for the joint model with random effects. In section 4.4.1, we will first not consider the random effects parameter of the correlation ρ. In section 4.4.2, all random effects parameters will be taken into account.

4.4.1 Results over 100 simulations without correlation parameter ρ

To investigate the behaviour of the joint modelling methods under realistic conditions, we first do the simulations without considering of the correlation parameter ρ. Over 100 iterations of simulations will be processed by using the proposed model. The generation procedure is described as follow:

- Step 1:

Generate 1000 variates $U_1, \cdots, U_{1000}$, following Gamma distribution with mean 1 and variance σ^2;

- Step 2:

Generate 5 covariates $V_{ij1}, \cdots, V_{ij5}$ for each V_i, following Gamma distribution with mean u_i and variance $\tau_j^2 u_i$ for each u_i, where $i = 1, \cdots, 1000$ and $j = 1, 2$;

- Step 3:

Generate $Y_{ij1}, \cdots, Y_{ij5}$ for each u_i and v_j, following Gamma and Poisson distribution respectively by using

$$Tw_{p_j}(\mu_{ijt} V_{ijt}, \epsilon_j^2 {V_{ijt}}^{1-p_j}) \tag{4.3}$$

The conditional mean and variance of the Tweedie distribution is given as

$$E(Y_{ijt}|W) = \mu_{ijt}V_{ijt} \quad \text{and} \quad Var(Y_{ijt}|W) = \epsilon_j^2 \mu_{ijt}^q {V_{ijt}}^{1-p_j} \tag{4.4}$$

where $\mu_{ijt} = exp(x'_{ijt}\beta_j)$

The variance for both responses is assumed to be 0.5. The regression and dispersion parameters for the first response were fixed; that is, $\beta_1 = (0, 0.3, -0.2)$, $\tau_1^2 = 0.1$ and $\epsilon_1^2 = 0.5$ respectively. Also, the regression and dispersion parameters for the second response were fixed; that is, $\beta_2 = (0.9, 0.1, 0.3)$, $\tau_2^2 = 0.2$ and $\epsilon_1^2 = 1$ respectively. The 100 data sets are obtained by repeating this procedure.

The average of the regression and dispersion estimates over 500 simulations are displayed in Table 4.5.

Table 4.5: Summary statistics for the Simulation without considering the random effects parameter ρ_j

Parameter	True Value	Estimated Value	Bias	Simulated SE
β_{01}	0	-0.0016	0.0002	0.0266
β_{11}	0.3	0.3008	0.0004	0.0114
β_{21}	-0.2	-0.2002	0.0001	0.0118
β_{02}	0.9	0.8982	0.0001	0.0267
β_{12}	0.1	0.1005	0.0003	0.0113
β_{22}	0.3	0.2998	0.0009	0.0120
σ^2	0.5	0.4908	0.0692	0.0299
τ_1^2	0.1	0.3544	0.0736	0.0158
ϵ_1^2	0.5	0.2904	0.0719	0.0098
ρ_1	0.0000	0.0000	0.0000	0.0000
τ_2^2	0.2	0.2986	0.1177	0.0156
ϵ_2^2	1	1	0.0000	0.0000
ρ_2	0.0000	0.0000	0.0000	0.0000

The regression parameters are almost the same as the true value and the random effects parameters are reasonably estimated through our model.

4.4.2 Results over 100 simulations with all random effects parameters

In this section, we will do the simulations that consider of the correlation parameter. Over 500 iterations of simulations will be processed by using the joint model. The generation procedure is described as follow:

- Step 1:

Generate 1000 variates $U_1, \cdots, U_{1000}$, following Gamma distribution with mean 1 and variance σ^2;

- Step 2:

Generate 5 covariates $V_{ij1}, \cdots, V_{ij5}$ for each V_i, following multivariate log-normal distribution with mean U_i and variance $\tau_j^2 \rho_j(t, t')U_i$ for each U_i, where $i = 1, \cdots, 1000$ and $j = 1, 2$;

- Step 3:

Generate $Y_{ij1}, \cdots, Y_{ij5}$ for each u_i and v_j, following Gamma and Poisson distribution respectively by using

$$Tw_{p_j}(\mu_{ijt}V_{ijt}, \epsilon_j^2 {V_{ijt}}^{1-p_j}) \tag{4.5}$$

The conditional mean and variance of the Tweedie distribution is given as

$$E(Y_{ijt}|W) = \mu_{ijt}V_{ijt} \quad \text{and} \quad Var(Y_{ijt}|W) = \epsilon_j^2 \mu_{ijt}^q {V_{ijt}}^{1-p_j} \tag{4.6}$$

where $\mu_{ijt} = exp(x'_{ijt}\beta_j)$

The variance for both responses is assumed to be 0.5. The regression and dispersion parameters for the first response were fixed; that is, $\beta_1 = (0, 0.3, -0.2)$, $\tau_1^2 = 0.5$, $\rho_1 = 0.5$ and $\epsilon_1^2 = 0.5$ respectively. Also, the regression and dispersion parameters for the second response were fixed; that is, $\beta_2 = (0.9, 0.1, 0.3)$, $\tau_2^2 = 0.2$, $\rho_2 = 0.5$ and $\epsilon_1^2 = 1$ respectively. The 500 data sets are obtained by repeating this procedure.

The average of the regression and dispersion estimates over 500 simulations are displayed in Table 4.6.

Table 4.6: Summary statistics for the Simulation with all random effects parameters

Parameter	True Value	Estimated Value	Bias	Simulated SE
β_{01}	0	-0.0002	0.0002	0.0289
β_{11}	0.3	0.2996	0.0004	0.0145
β_{21}	-0.2	-0.1991	0.0001	0.0136
β_{02}	0.9	0.8999	0.0001	0.0277
β_{12}	0.1	0.1003	0.0003	0.0112
β_{22}	0.3	0.2991	0.0009	0.0111
σ^2	0.5	0.4308	0.0692	0.0311
τ_1^2	0.5	0.5736	0.0736	0.0694
ϵ_1^2	0.5	0.5719	0.0719	0.0631
ρ_1	0.5	0.6971	0.1971	0.0649
τ_2^2	0.2	0.3177	0.1177	0.0290
ϵ_2^2	1	1	0.0000	0.0000
ρ_2	0.5	0.7396	0.2396	0.0465

The regression parameters are almost the same as the true value and the random effects parameters are reasonably estimated through our model. On the other hand, the random effects parameter σ^2 is underestimated and the other random effects parameters are overestimated.

Chapter 5

Discussion

5.1 Conclusion

In longitudinal studies, different types of outcomes are frequently collected over time on each of many subjects. Plenty of different methods have been developed according to the background and the objective of the study.

The proposed joint Tweedie mixed models can be used for analyzing the situations with mixed types of longitudinal data. Because the shared subject-specific random effects can incorporate heterogenity between different responses and the distinct time-specific random effects can describe the correlations between measurements on the same response variable within a subject, the proposed models are able to investigate the correlation between different responses of the same subject. Moreover, the models are proposed through the generalized linear mixed models based on the class of Tweedie distributions of different index parameters. Comparing with the univariate model, the new model incorporates additional information that increase estimation and prediction accuracy.

The prediction of the random effects can be obtained from the best linear unbiased predictor of the random effects which has been proved to be a good method for prediction of random effects (BLUP) by Ma and Jørgensen (2007). Because the BLUP depends only on the first two moments of random effects, the proposed model has the advantage of computational efficiency and robustness against misspecification of random-effects distributions (Ha and Lee, 2005).

Both the distributional shape and the intra-dependence of clustered data are considered in Tweedie mixed models. It makes the resulting variance component decomposition of the structure of the covariance matrix of the response easier to interpret as well as simpler to fit the model compared with the conditional mode approach (Ma et al., 2007). The expressions of the estimating equations derived in Chapter 3 are explicit in matrix forms.

Our BLUP provides a common method for estimating regression parameters and random effects parameters for all Tweedie mixed models. In fact, the orthodox BLUP approach specifically includes random effects and combines the "subject-specific" and the "population-averaged" inferences within a common model (Lee and Nelder, 1996). The asymptotic justifications of the conditional mode approaches which were proposed by Draper (1996) mainly rely on the approximate normality for random effects or the "the right transformed normality for random effects on the right scale", but in BLUP approach, the asymptotic justifications do not depend on any distributional assumption of random effects.

5.2 Further study

There are still many areas we could discuss in more detail.

In this thesis, we did not address some important research questions such as how to check the model. It is challenging in random-effects modelling approaches. Also, we did not discuss the significance of random effects parameters σ^2, τ^2, ϵ^2 and ρ as their standard errors are not estimated. However, it would be of interest to develop suitable tests for their significance. The bootstrap method may be used since it could help us to construct such confidence intervals (Lele 1991).

Observations of each subject are serially correlated in longitudinal data since subjects are measured over a period of time. In our model, the correlation parameter ρ is assumed under the AR(1) structure. However, it is of interest to investigate other structures in the model.

The correlation ρ between two measurement times for the same object can be expressed as follows:

$$\Re_j = \left[\rho_{(t,t')}\right]_{T\times T} = \begin{bmatrix} 1 & \rho_{(1,2)} & \rho_{(1,3)} & \cdots & \rho_{(1,T-1)} \\ \rho_{(2,1)} & 1 & \rho_{(2,3)} & \cdots & \rho_{(2,T-1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho_{(T-1,1)} & \rho_{(T-1,2)} & \rho_{(T-1,3)} & \cdots & 1 \end{bmatrix} \tag{5.1}$$

Many correlation structures are available as options, such as compound symmetry and Toeplitz. Compound symmetry is of interest since it has a simpler form than the AR(1); it assumes that $\text{Corr}(Y_{i,t}, Y_{i,t+s}) = \rho$, its correlation structure can be obtained as follows:

$$\Re_j = \left[\rho_{(t,t')}\right]_{T\times T} = \begin{bmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \dots & \dots & \dots & \dots & \dots \\ \rho & \rho & \rho & \dots & 1 \end{bmatrix} \quad (5.2)$$

For the Toeplitz structure, it assumes assumes that $\text{Corr}(Y_{i,t}, Y_{i,t+s}) = \rho_t$ for all s and t, which the correlation matrix can be expressed as:

$$Cov(Y_i) = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \dots & 1 \end{bmatrix}. \quad (5.3)$$

In future work, we may add these correlation structures into the proposed model.

Model selection is an important part in data analysis. A common issue in our model is how to judge the performance of the regression and random effects parameters. In our analysis, we just considered the generalized linear mixed effects models only, however, there is no model-selection criteria available in GLMM. Therefore, a generalized estimating equation (GEE) could be used since we could use the well-known Akaike Information Criterion (AIC) to select a better model through comparing the values of AIC. To be more specific, we set different values to random random effects parameters. For example:

Firstly, suppose we set $\sigma^2 = 0$ and calculate the associated AIC value. Then we estimate σ^2 from data and calculate another AIC value. If the simpler model with $\sigma^2 = 0$ has a smaller AIC value, we may think that it is a better model according to its properties. Thus, the subject-specific random effects can be expressed as follows:

$$E(U_i) = 1 \quad \text{and} \quad \text{Var}(U_i) = \sigma^2 = 0.$$

It can be concluded that there is no correlation between different responses. Hence, the responses can be analyzed separately or jointly.

According to these steps, it can be shown that our aim of using GEE method is to find a simpler model which is also adequate for the data. In addition, this may give a more convincing interpretation and more precise parameter estimates.

Thus, further work on the comparison of GEE model and GLMM would be of interest.

Appendix A

R Code for Pony Data

```
# process data & load in data
ponys <- read.csv("/Users/JiaxiuLi/Desktop/ponys.csv",head=TRUE)

#transform the responses,i=6,t=36,j=2
Iidx<-rep(1:6, each=36)
Tidx<-rep(1:36, 6)
Jidx<-rep(rep(1:2,each=36),6)
y1<-ponys$count  #possion,power=1; the spike burst rate
y2<-ponys$duram  #gamma,power=2; the duration

#transform the predictor variable
X1<-cbind(ponys$drug,ponys$logbcount)
X1<-cbind(1, X1)
```

```
X2<-cbind(ponys$drug,ponys$logbduram)
X2<-cbind(1, X2)

nx<-3

Iidx<-rep(1:6, each=36); Tidx<-rep(1:36, 6)

# The Main Function
blup<-function(Y,X1,X2,Iidx,Tidx,q1,q2,
silent=TRUE,sigma=NULL,estsigma=TRUE,
tau1=NULL,eps1=1,rho1=0.5,esttau1=FALSE,estrho1=TRUE,esteps1=FALSE,
tau2=NULL,eps2=0.03,rho2=0.5,esttau2=TRUE,estrho2=TRUE,esteps2=TRUE,
tol=1e-4, niter=100,intercept=TRUE){

require(MASS)

#denfine I,T
n <- length(Y); I<-length(unique(Iidx)); T<-length(unique(Tidx))
J<-length(unique(Jidx))
if (!silent) cat("I=", I, "T=", T, "J=", J, "n=", n, "\n")
```

```
# A function of rho to correlation matrix R.
rho2R<-function(rho,T){
return(outer(1:T,1:T,function(irow,icol){
rho^(abs(irow-icol))}))
}

# A function to evaluate a'Ab.
quad<-function(a, A, b){
sum(a*apply(t(A)*b,2,sum))}

# Initialize beta, mu.
require(tweedie)
#judge q1
if (q1==1){
ft1<-glm(round(y1)~.-1, data=as.data.frame(X1),
family=poisson(link=log))} else{
if (q1>1 & q1<2){
ft1<-glm(y1~.-1, data=as.data.frame(X1),
family=tweedie(var.power=q1,link.power=0))
} else{
ft1<-glm(I(y1+0.001)~.-1, data=as.data.frame(X1),
family=Gamma(link=log))}
```

```
  }
    beta1<-ft1$coef
    mu1<-ft1$fitted
    r1<-predict(ft1)
if (!silent) cat("beta from glm=", beta1, "\n")

#judge q2
if (q2==1){
  ft2<-glm(round(y2)~.-1, data=as.data.frame(X2),
  family=poisson(link=log))} else{
 if (q2>1 & q2<2){
  ft2<-glm(y2~.-1, data=as.data.frame(X2),
  family=tweedie(var.power=q2,link.power=0))
 } else{
  ft2<-glm(I(y2+0.001)~.-1, data=as.data.frame(X2),
  family=Gamma(link=log))}
}
    beta2<-ft2$coef
    mu2<-ft2$fitted
    r2<-predict(ft2)
if (!silent) cat("beta from glm=", beta2, "\n")
```

```
# Initial values of parameters and random effects
u<-rep(NA, I)  #; v<-rep(NA, I*T)
v1<-rep(NA, I*T)
v2<-rep(NA, I*T)
for (i in 1:I){
u[i]<-mean(y1[Iidx==i])/mean(y1)
for (t in 1:T){
v1[(i-1)*T+t]<-mean(y1[Iidx==i & Tidx==t])/mean(y1)
v2[(i-1)*T+t]<-mean(y2[Iidx==i & Tidx==t])/mean(y2)
}
}
thres<-1e-4; u<-thres+u; v1<-v1; v2<-v2+thres

##############################################################
#initial values of parameters: mu(beta), sigma2, tau2,rho, eps2.
#initial for the sigma,  sigma=NULL
if (is.null(sigma)){
  sigma2<-mean((u-1)^2)
} else{
  sigma2<-sigma
}
```

```
#initial for the 1st response
if (is.null(tau1)){
  tau01<-mean((v1-rep(u, each=T))^2)
} else{tau01<-tau1}
  rho01<-rho1
if (!esttau1) estrho1<-FALSE
eps01<-eps1
if (q1==1) {eps01<-1; esteps1<-FALSE}

#initial for the 2nd response
if (is.null(tau2)){
  tau02<-mean((v2-rep(u, each=T))^2)
} else{tau02<-tau2}
  rho02<-rho2
if (!esttau2) estrho2<-FALSE
eps02<-eps2
if (q2==1) {eps02<-1; esteps2<-FALSE}

############################################################
#updating...
iter<-1
if (!silent) cat("iter=0", "beta1=", beta1, "beta2=", beta2,
```

```
"sigma2=", sigma2, "tau01=", tau01, "rho01=", rho01,
"eps01=", eps01,"tau02=", tau02,
"rho02=", rho02, "eps02=", eps02, "\n")
continue<-TRUE
container<-NULL
container<-rbind(container,c(beta1, sigma2, tau01, rho01, eps01,
      beta2, sigma2, tau02, rho02, eps02))
while (continue){
if (!silent){
cat("iter=", iter, "sigma2=", sigma2,
"tau01=", tau01, "rho01=", rho01, "eps01=", eps01,
"tau02=", tau02, "rho02=", rho02, "eps02=", eps02, "\n")
}

# Current beta, sigma2,tau01, rho01, eps01, tau02, rho02, eps02.
#-------------------- -----------------
beta1.prev<-beta1
beta2.prev<-beta2
sigma2.prev<-sigma2
tau01.prev<-tau01
rho01.prev<-rho01
eps01.prev<-eps01
```

```
tau02.prev<-tau02
rho02.prev<-rho02
eps02.prev<-eps02

##############################################################
# Update random effects u and v, and beta, mu.
psibeta<-rep(0, ncol(X1)*2)
sbeta<-matrix(0, ncol(X1)*2, 2*ncol(X1))
for (i in 1:I){
  mui1<-mu1[Iidx==i];mui2<-mu2[Iidx==i];mui<-c(mui1,mui2)
  yi1<-y1[Iidx==i]; yi2<-y2[Iidx==i];yi<-c(yi1,yi2)
  X1i<-X1[Iidx==i,]
  X2i<-X2[Iidx==i,]
  XXi<-rbind(cbind(X1i,matrix(0,T,nx)),cbind(matrix(0,T,nx),X2i))

  cov_v<-rbind(cbind(tau01*rho2R(rho01,T),matrix(0,T,T)),
               cbind(matrix(0,T,T),tau02*rho2R(rho02,T)))+sigma2

  var_y_c1<-eps01*diag(mui1^q1)+outer(mui1, mui1)
                 *(tau01*rho2R(rho01,T)+sigma2)
  var_y_c2<-eps02*diag(mui2^q2)+outer(mui2, mui2)
                 *(tau02*rho2R(rho02,T)+sigma2)
```

```
var_y_b1<-sigma2*outer(mui1, mui2)
var_y_b2<-t(var_y_b1)

#covariance of y
var_y<-rbind(cbind(var_y_c1,var_y_b1),cbind(var_y_b2,var_y_c2))
var_y_inv<-ginv(var_y)

#covariance of v and y
var_v_y<- cov_v %*% diag(mui)   #checked

# U pdate random effects.
vi<-1+c(var_v_y %*% var_y_inv %*% matrix(yi-mui,2*T,1))
v1[((i-1)*T+1): (i*T)]<-vi[1:T]
v2[((i-1)*T+1): (i*T)]<-vi[(T+1):(T*2)]
u[i]<-1+quad(sigma2*mui, var_y_inv, yi-mui)

# Calculate psibeta and sbeta.
psibeta<-psibeta+c(t(XXi)%*%diag(mui)%*%var_y_inv
               %*%matrix(yi-mui,2*T,1))
sbeta<-sbeta-t(XXi)%*%diag(mui)%*%var_y_inv
               %*%diag(mui)%*%(XXi)
```

```
}

# Update beta and mu.
beta<-c(beta1,beta2)
beta<-beta-apply(ginv(sbeta)*psibeta, 2, sum)
beta1<-beta[1:nx]
beta2<-beta[(nx+1):(nx*2)]
mu1<-exp(apply(t(X1)*beta1, 2, sum))
mu2<-exp(apply(t(X2)*beta2, 2, sum))

################################################################
#Update sigma2, tau2, rho and eps2
# 1. sigma2%%%%
sigma2.updater<-function(){
  ci<-rep(0, I)
  for (i in 1:I){
   mui1<-mu1[Iidx==i];mui2<-mu2[Iidx==i];mui<-c(mui1,mui2)
   yi1<-y1[Iidx==i]; yi2<-y2[Iidx==i];yi<-c(yi1,yi2)
   X1i<-X1[Iidx==i,]
    X2i<-X2[Iidx==i,]
    XXi<-rbind(cbind(X1i,matrix(0,T,4)),cbind(matrix(0,T,4),X2i))
```

```
    cov_v<-rbind(cbind(tau01*rho2R(rho01,T),matrix(0,T,T)),
              cbind(matrix(0,T,T),tau02*rho2R(rho02,T)))+sigma2

    var_y_c1<-eps01*diag(mui1^q1)+outer(mui1, mui1)
                   *(tau01*rho2R(rho01,T)+sigma2)
    var_y_c2<-eps02*diag(mui2^q2)+outer(mui2, mui2)
                   *(tau02*rho2R(rho02,T)+sigma2)

     var_y_b1<-sigma2*outer(mui1, mui2)
     var_y_b2<-t(var_y_b1)

     #covariance of y
     var_y<-rbind(cbind(var_y_c1,var_y_b1),cbind(var_y_b2,var_y_c2))
     var_y_inv<-ginv(var_y)

     ci[i]<-sigma2-sigma2^2*quad(mui, var_y_inv, mui)
   }
   sigma2<-mean((u-1)^2)+mean(ci)
   return(sigma2)
 }
 if (estsigma){
   sigma2<-sigma2.updater()
```

```
} else{
  sigma2<-sigma
}

# 2/(1). tau01%%%%%
tau01.updater<-function(){
b<-matrix(0, I, T)
R1<-rho2R(rho01, T)
for (i in 1:I){
mui1<-mu1[Iidx==i];mui2<-mu2[Iidx==i];mui<-c(mui1,mui2)
yi1<-y1[Iidx==i]; yi2<-y2[Iidx==i];yi<-c(yi1,yi2)

#covariance of v_i1 (the 1st response)
cov_v<-sigma2+tau01*R1

#covariance of y_i1 (the 1st response)
var_y<-eps01*diag(mui1^q1)+outer(mui1, mui1)*cov_v
var_y_inv<-ginv(var_y)

#covariance of v and y
cov_v_y<- cov_v %*% diag(mui1)
#convariance of u and y
```

```
cov_u_y<-sigma2*(mui1)

for (t in 1:T){
b[i,t]<-tau01-(sigma2^2*quad(mui1, var_y_inv, mui1)+quad(cov_v_y[t,],
           var_y_inv, cov_v_y[t,]-2*cov_u_y))
}
}
  tau01<-mean((v1-rep(u, each=T))^2)+mean(c(b))
  if (!silent) cat("iter=", iter,
  "tau01=", tau01, "meanb=", mean(c(b)), "\n")
  return(tau01)
}
if (esttau1){
  tau01<-tau01.updater()
} else{
  tau01<-tau1
}

# 2/(2). tau02%%%%
tau02.updater<-function(){
b<-matrix(0, I, T)
R2<-rho2R(rho02, T)
```

```
for (i in 1:I){
mui1<-mu1[Iidx==i];mui2<-mu2[Iidx==i];mui<-c(mui1,mui2)
yi1<-y1[Iidx==i]; yi2<-y2[Iidx==i];yi<-c(yi1,yi2)

#covariance of v_i1 (the 1st response)
cov_v<-sigma2+tau02*R2

#covariance of y_i1 (the 1st response)
var_y<-eps02*diag(mui2^q2)+outer(mui2, mui2)*cov_v
var_y_inv<-ginv(var_y)

#covariance of v and y
cov_v_y<- cov_v %*% diag(mui2)
#convariance of u and y
cov_u_y<-sigma2*(mui2)

for (t in 1:T){
b[i,t]<-tau02-(sigma2^2*quad(mui2, var_y_inv, mui2)+quad(cov_v_y[t,],
          var_y_inv, cov_v_y[t,]-2*cov_u_y))
}
}
tau02<-mean((v2-rep(u, each=T))^2)+mean(c(b))
```

```
if (!silent) cat("iter=", iter,
"tau02=", tau02, "meanb=", mean(c(b)), "\n")
return(tau02)
}
if (esttau2){
  tau02<-tau02.updater()
} else{
  tau02<-tau2
}

# 3/(1). eps01%%%%%
eps01.updater<-function(){
R1<-rho2R(rho01, T)
b<- mu1^2*(sigma2+tau01)
for (i in 1:I){
mui1<-mu1[Iidx==i];mui2<-mu2[Iidx==i];mui<-c(mui1,mui2)
yi1<-y1[Iidx==i]; yi2<-y2[Iidx==i];yi<-c(yi1,yi2)

#covariance of v
cov_v<-tau01*R1+sigma2

#covariance of y
```

```
var_y<-eps01*diag(mui1^q1)+outer(mui1, mui1)*cov_v
var_y_inv<-ginv(var_y)

#covariance of v and y
cov_v_y<- cov_v %*% diag(mui1)

b[Iidx==i]<-b[Iidx==i]-(mui1)^2
    *diag(cov_v_y %*% var_y_inv %*% t(cov_v_y))
    }
    eps01<-mean(((y1-mu1*v1)^2+c(b))/(mu1^q1))
    return(eps01)
}
if (esteps1) {eps01<-eps01.updater()}

# 3/(2). eps02%%%%%
eps02.updater<-function(){
R2<-rho2R(rho02, T)
b<- mu2^2*(sigma2+tau02)
for (i in 1:I){
mui1<-mu1[Iidx==i];mui2<-mu2[Iidx==i];mui<-c(mui1,mui2)
yi1<-y1[Iidx==i]; yi2<-y2[Iidx==i];yi<-c(yi1,yi2)
```

```
#covariance of v
cov_v<-tau02*R2+sigma2

#covariance of y
var_y<-eps02*diag(mui2^q2)+outer(mui2, mui2)*cov_v
var_y_inv<-ginv(var_y)

#covariance of v and y
cov_v_y<- cov_v %*% diag(mui2)

b[Iidx==i]<-b[Iidx==i]-(mui2)^2
*diag(cov_v_y %*% var_y_inv %*% t(cov_v_y))
}
eps02<-mean(((y2-mu2*v2)^2+c(b))/(mui2^q2))
return(eps02)
}
if (esteps2) {eps02<-eps02.updater()}

# 4/(1). rho01%%%%
# Bias-corrected formula from random effects.
rho01.updater.RE<-function(){
e<-rep(0, I*(T-1))->s->e1->e2->s1->s2
```

```
R1<-rho2R(rho01, T)
for (i in 1:I){
mui1<-mu1[Iidx==i];mui2<-mu2[Iidx==i];mui<-c(mui1,mui2)
yi1<-y1[Iidx==i]; yi2<-y2[Iidx==i];yi<-c(yi1,yi2)

#covariance of v
cov_v<-tau01*R1+sigma2

#covariance of y
var_y<-eps01*diag(mui1^q1)+outer(mui1, mui1)*cov_v
var_y_inv<-ginv(var_y)

#covariance of v and y
cov_v_y<- cov_v %*% diag(mui1)

cov_v_hat<-cov_v_y%*%var_y_inv%*%t(cov_v_y)
vi<-v1[((i-1)*T+1):(i*T)]

for (t in 1:(T-1)){
e[(i-1)*(T-1)+t]<-rho01*tau01-(cov_v_hat[t,t+1]+sigma2
*quad(mui1, var_y_inv, sigma2*mui1-cov_v_y[t,]-cov_v_y[t+1,]))
e1[(i-1)*(T-1)+t]<-tau01-(cov_v_hat[t,t]
```

```
+sigma2*quad(mui1, var_y_inv, sigma2*mui1-2*cov_v_y[t,]))
e2[(i-1)*(T-1)+t]<-tau01-(cov_v_hat[t+1,t+1]
+sigma2*quad(mui1,var_y_inv,sigma2*mui1-2*cov_v_y[t+1,]))

s[(i-1)*(T-1)+t]<-(vi[t]-u[i])*(vi[t+1]-u[i])
s1[(i-1)*(T-1)+t]<-(vi[t]-u[i])^2
s2[(i-1)*(T-1)+t]<-(vi[t+1]-u[i])^2
}
}
rho01<-sum(s+e)/sqrt(sum(s1+e1)*sum(s2+e2))
if (rho01>.99999|rho01<0.00001)
{cat("rho exceeds limits!"); rho01<-runif(1,0.2, 0.8)}
return(rho01)
}
if (estrho1){
  rho01<-rho01.updater.RE()
} else{
  rho01<-rho1
}

# 4/(2). rho02%%%%
rho02.updater.RE<-function(){
```

```
e<-rep(0, I*(T-1))->s->e1->e2->s1->s2
R2<-rho2R(rho02, T)
for (i in 1:I){
mui1<-mu1[Iidx==i];mui2<-mu2[Iidx==i];mui<-c(mui1,mui2)
yi1<-y1[Iidx==i]; yi2<-y2[Iidx==i];yi<-c(yi1,yi2)

#covariance of v
cov_v<-tau02*R2+sigma2

#covariance of y
var_y<-eps02*diag(mui2^q2)+outer(mui2, mui2)*cov_v
var_y_inv<-ginv(var_y)

#covariance of v and y
cov_v_y<- cov_v %*% diag(mui2)

cov_v_hat<-cov_v_y%*%var_y_inv%*%t(cov_v_y)
vi<-v2[((i-1)*T+1):(i*T)]

for (t in 1:(T-1)){
e[(i-1)*(T-1)+t]<-rho02*tau02-(cov_v_hat[t,t+1]+sigma2*
quad(mui2, var_y_inv, sigma2*mui2-cov_v_y[t,]-cov_v_y[t+1,]))
```

```
e1[(i-1)*(T-1)+t]<-tau02-(cov_v_hat[t,t]+sigma2*
quad(mui2, var_y_inv, sigma2*mui2-2*cov_v_y[t,]))
e2[(i-1)*(T-1)+t]<-tau02-(cov_v_hat[t+1,t+1]+sigma2*
quad(mui2, var_y_inv, sigma2*mui2-2*cov_v_y[t+1,]))

s[(i-1)*(T-1)+t]<-(vi[t]-u[i])*(vi[t+1]-u[i])
s1[(i-1)*(T-1)+t]<-(vi[t]-u[i])^2
s2[(i-1)*(T-1)+t]<-(vi[t+1]-u[i])^2
}
}
rho02<-sum(s+e)/sqrt(sum(s1+e1)*sum(s2+e2))
if (rho02>.99999|rho02<0.00001)
{cat("rho exceeds limits!"); rho02<-runif(1,0.2, 0.8)}
return(rho02)
}
if (estrho2){
  rho02<-rho02.updater.RE()
} else{
  rho02<-rho2
}

############################################################
```

```
difference1<-sum(abs(beta1-beta1.prev))+
             abs(sigma2-sigma2.prev)+
             abs(tau01-tau01.prev)+
             sum(abs(rho01-rho01.prev))+
             abs(eps01-eps01.prev)
difference2<-sum(abs(beta2-beta2.prev))+
             abs(sigma2-sigma2.prev)+
             abs(tau02-tau02.prev)+
             sum(abs(rho02-rho02.prev))+
             abs(eps02-eps02.prev)
difference<-difference1+difference2
iter<-iter+1
if (!silent) cat("iter=", iter, "continue=", continue,
               "difference1=", difference1,"\n")
if (!silent) cat("iter=", iter, "continue=", continue,
               "difference2=", difference2,"\n")
if (!silent) cat("beta1=", beta1, "sigma2=", sigma2,
    "tau01=", tau01, "rho01=", rho01, "eps01=", eps01, "\n")
if (!silent) cat("beta2=", beta2, "sigma2=", sigma2,
    "tau02=", tau02, "rho02=", rho02, "eps02=", eps02, "\n")

continue<-(difference>tol) & (iter<=niter)
```

```
container<-rbind(container, c(beta1, sigma2, tau01, rho01, eps01,
                   beta2, sigma2, tau02, rho02, eps02))
}

##############################################################
# Calculate the covariance matrix of beta (estimated)
if (!silent) cat("Eigenvalues of sbeta=",
eigen(-sbeta)$values, "\n")
invS<-ginv(- sbeta)
return(list(beta01=as.numeric(summary(ft1)$coef[,1]),
            SE_beta01=as.numeric(summary(ft1)$coef[,2]),
            beta02=as.numeric(summary(ft2)$coef[,1]),
            SE_beta02=as.numeric(summary(ft2)$coef[,2]),
            beta=as.numeric(beta), SE_beta=sqrt(diag(invS)),
            u=u, v1=v1, v2=v2,r1=r1,r2=r2,
            sigma2=sigma2, tau01=tau01, rho01=rho01, eps01=eps01,
                           tau02=tau02, rho02=rho02, eps02=eps02,

            container=container))
}
# End of the Main Function.
```

```
#################################################################
tmp<-blup(Y,X1,X2,Iidx,Tidx,q1=1,q2=2,
silent=TRUE,sigma=NULL,estsigma=TRUE,
tau1=NULL,eps1=1,rho1=0.5,esttau1=TRUE,estrho1=TRUE,esteps1=FALSE,
tau2=NULL,eps2=0.03,rho2=0.5,esttau2=TRUE,estrho2=TRUE,esteps2=TRUE,
tol=1e-4, niter=1000,intercept=TRUE)
```

Appendix B

R Code for Simulation

B.1 Simulation with correlation

```
q1<-2;sigma2=0.5;tau1=0.5;eps1=0.5;rho1=0.5
q2<-1;tau2=0.2;eps2=1;rho2=0.5

#randomize the value of beta
set.seed(100)
beta1<-round(rnorm(3, 0, 2), 1); beta1[1]<-0
beta2<-round(rnorm(3, 0, 1), 1)

I<-1000;T<-5;J<-2
Iidx<-rep(1:I, each=T); Tidx<-rep(1:T, I)
```

```
x1<-matrix(rnorm(2*I*T),I*T,2)
X1<-cbind(1,x1)
emu1<-c(exp(X1%*%matrix(beta1, 3, 1)))
x2<-matrix(rnorm(2*I*T),I*T,2)
X2<-cbind(1,x2)
emu2<-c(exp(X2 %*% matrix(beta2, 3, 1)))
source("/Users/JiaxiuLi/Desktop/code.R")

set.seed(100)
container<-NULL

for (kk in 1:10){
#level 1:
u<-rgamma(I,1/sigma2,1/sigma2)

# level 2:
v1<-rep(NA, I*T)
v2<-rep(NA, I*T)

mvrlognorm<-function(n, Omega, v=NULL){
  # v=E(X); Omega=V(X).
```

```
  p<-nrow(Omega)
  if (is.null(v)) v<-rep(1, p)
  Sigma<-matrix(NA, p, p)
  for (i in 1:p){
    for (j in 1:p){
      Sigma[i,j]<-log(1+Omega[i,j]/(v[i]*v[j]))
    }
  }
  mu<- (-0.5)*diag(Sigma)+log(v)
  require(MASS)
  Y<-mvrnorm(n, mu,Sigma)
  X<-exp(Y)
  return(X)
}
# A function of rho to correlation matrix R.
rho2R<-function(rho,T){
return(outer(1:T,1:T,function(irow,icol){
rho^(abs(irow-icol))}))
}

for (i in 1:I){
  Omega1<-u[i]*tau1*rho2R(rho1, T)
```

```
  Omega2<-u[i]*tau2*rho2R(rho2, T)
  mn<-rep(u[i], T)
  v1[((i-1)*T+1):(i*T)]<-mvrlognorm(1, Omega1, mn)
  v2[((i-1)*T+1):(i*T)]<-mvrlognorm(1, Omega2, mn)
}

#level 3:
#shape=V_ijt/eps
#rate=1/(eps*mu)

y1<-matrix(0,I,T)
y2<-matrix(0,I,T)
for(i in 1:I){
  emui1<-emu1[Iidx==i];emui2<-emu2[Iidx==i]
  vi1<-v1[Iidx==i]; vi2<-v2[Iidx==i]
  for (t in 1:T){
    y1[i,t]<-rgamma(1,shape=vi1[t]/eps1,rate=1/(eps1*emui1[t]))
    y2[i,t]<-rpois(1,vi2[t]*emui2[t])
  }
}

y1<-c(t(y1))
```

```
y2<-c(t(y2))

tmp<-blup(y1, y2, X1, X2,Iidx,Tidx,q1=2,q2=1,
silent=TRUE,sigma=NULL,estsigma=TRUE,
tau1=NULL,eps1=0.5,rho1=0.5, esttau1=TRUE,estrho1=TRUE,esteps1=TRUE,
tau2=NULL,eps2=1,rho2=0.5,esttau2=TRUE,estrho2=TRUE,esteps2=FALSE,
tol=1e-4, niter=200,intercept=TRUE)

container<-rbind(container, tmp$container[nrow(tmp$container), ])
}
container
apply(container,2,mean)
```

B.2 Simulation without correlation

```
q1<-2;sigma2=0.5;tau1=0.1;eps1=0.5;rho1=0.5
q2<-1;tau2=0.2;eps2=1;rho2=0.5

#randomize the value of beta
set.seed(100)
beta1<-round(rnorm(3, 0, 2), 1); beta1[1]<-0
beta2<-round(rnorm(3, 0, 1), 1)
```

```
I<-1000;T<-5;J<-2
Iidx<-rep(1:I, each=T); Tidx<-rep(1:T, I)

x1<-matrix(rnorm(2*I*T),I*T,2)
X1<-cbind(1,x1)
emu1<-c(exp(X1%*%matrix(beta1, 3, 1)))
x2<-matrix(rnorm(2*I*T),I*T,2)
X2<-cbind(1,x2)
emu2<-c(exp(X2 %*% matrix(beta2, 3, 1)))
source("/Users/JiaxiuLi/Desktop/code.R")

container<-NULL

for (kk in 1:3){
#level 1:
u<-rgamma(I,1/sigma2,1/sigma2)

# level 2: 1st mehod:
v1<-rep(NA, I*T)
v2<-rep(NA, I*T)
for (i in 1:I){
```

```
  v1[((i-1)*T+1):(i*T)]<-rgamma(T,u[i]/tau1,1/tau1)
  v2[((i-1)*T+1):(i*T)]<-rgamma(T,u[i]/tau2,1/tau2)
}

#level 3:
#shape=V_ijt/eps
#rate=1/(eps*mu)

y1<-matrix(0,I,T)
y2<-matrix(0,I,T)
for(i in 1:I){
  emui1<-emu1[Iidx==i];emui2<-emu2[Iidx==i]
  vi1<-v1[Iidx==i]; vi2<-v2[Iidx==i]
  for (t in 1:T){
    y1[i,t]<-rgamma(1,shape=vi1[t]/eps1,rate=1/(eps1*emui1[t]))
    y2[i,t]<-rpois(1,vi2[t]*emui2[t])
  }
}

y1<-c(t(y1))
y2<-c(t(y2))
```

```
tmp<-blup(y1, y2, X1, X2,Iidx,Tidx,q1=2,q2=1,
silent=TRUE,sigma=NULL,estsigma=TRUE,
tau1=NULL,eps1=0.5,rho1=0, esttau1=TRUE,estrho1=FALSE,esteps1=TRUE,
tau2=NULL,eps2=1,rho2=0,esttau2=TRUE,estrho2=FALSE,esteps2=FALSE,
tol=1e-4, niter=10,intercept=TRUE)

container<-rbind(container, tmp$container[nrow(tmp$container), ])
}
container
apply(container,2,mean)
```

Bibliography

[1] Throp, J. (2009). *Joint mixed-effects models for longitudinal data analysis: an application for the metabolic syndrome*, Virginia Commonwealth University, Master thesis.

[2] Catalano PJ and Ryan LM (1992). Bivariate latent variable models for clustered discrete and continuous outcomes. *Journal of the American Statistical Association*, 87, pp651-658.

[3] Fitzmaurice GM and Laird NM (1995). Regression models for a bivariate discrete and continuous outcome with clustering. *Journal of the American Statistical Association*, 90, pp845-852.

[4] Gueorguieva RV and Agresti A, (2001). A correlated probit model for joint modelling of clustered binary and continuous responses. *Journal of the American Statistical Association*, 96, pp1102-1112.

[5] Fitzmaurice GM and Laird NM (1995). Latent variable models for mixed discrete and continuous outcomes. *Journal of the Royal Statistical Society, Series*

B, 59, pp667-678.

[6] Rochon, J (1996). Analyzing bivariate repeated measures for discrete and continuous outcome variables. *Biometrics*, 52, pp740-750.

[7] Regan MM, Catalano, PJ (1999). Likelihood models for clustered binary and continuous outcomes: application to development toxicology. *Biometrics*, 55, pp760-768.

[8] S. Fieuws, G. Verbeke, and G. Molenberghs (2007). Random-effects models for multivariate repeated measures. *Stat. Methods Med. Res.* 16, pp387-397.

[9] Jrgensen, B. (1987). Exponential dispersion models (with discussion). *J. Roy. Statist. Soc. Ser. B* 49, pp127-162.

[10] Ma, R. (1999). *An orthodox blup approach to generalized linear mixed models*, University of British Columbia, PhD thesis.

[11] Garret M. Fitzmaurice, Nan M. Laird and James H. Ware. (2011). *Applied Longitudinal Analysis (*2^{nd} *ed.).* John Wiley & Sons: New Jersey.

[12] Diggle, P.J., Heagerty, P.J., Liang, K.Y., and Zeger, S.L. (2002). *Analysis of Longitudinal Data (*2^{nd} *ed.).* Oxford Science Publications: Oxford.

[13] Fitzmaurice, G., Davidian, M., Verbeke, G. and Molenberghs, G. (2009). *Longitudinal Data Analysis*. Chapman & Hall/CRC Taylor Francis Group: Boca Raton.

[14] Nedler, J.A. and Wedderburn, R.W.M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society, Series A*, 135, pp370-384.

[15] Liang K.Y. and Zeger, S.L. (1986). Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*, 73, pp13-22.

[16] Laird, N. M. and Ware, J. H. (1982). Random-Effects Models for Longitudinal Data. *Biometrics*, 38, pp963-974.

[17] Guo, X. (2011). *Longitudinal Data Analysis in Social Science Data*, University of Alberta, Master thesis.

[18] Lall, S. (2014). *A Study of Tweedie Family of Distribution for Rainfall Modelling*, Indian Agricultural Research Institute, Master thesis.

[19] Jrgensen, B. (1997). *The Theory of Dispersion Models*. Chapman and Hall: London.

[20] Jrgensen, B. and Souza, M.C.P. (1994). Fitting Tweedie's compound Poisson model to insurance claims data. *Scand. Actuarial J.* 1, pp69-93.

[21] Jrgensen, B. (1987). Exponential dispersion models (with discussion). *J. Roy. Statist. Soc. Ser. B* 49, pp127-162.

[22] Ma, R. and Jrgensen, B. (2007). *Nested generalized linear models: an orthodox best linear unbiased predictor approach,* Journal of Royal Statistical Society, B69, pp 625-641.

[23] Ma, R., Hasan, M.T., and Sneddon, G. (2009). Modelling heterogeneity in clustered count data with extra zeros using compound Poisson random effect. *Statistics in medicine*, 28, pp 2356-2369.

[24] Lester GD, Merritt AM, Neuwirth L, VietroWidenhouse T et al. (1998a). Myoelectric activity of the ileum, cecum, and right ventral colon, and cecal emptying of radiolabeled markers in clinically normal ponies. *American Journal of Veterinary Medicine*, 59, pp 313-319.

[25] Lester GD, Merritt AM, Neuwirth L, VietroWidenhouse T et al. (1998a). Effect of α_2-adrenergic, cholinergic, and nonsteroidal anti-inflammatory drugs on myoelectric activity of ileum, cecum, and right ventral colon, and on cecal emptying of radiolabeled markers in clinically normal ponies. *American Journal of Veterinary Medicine*, 59, pp 320-327.

[26] Lester GD, Merritt AM, Neuwirth L, VietroWidenhouse T et al. (1998a). Effect of erythromycin lactobionate on myoelectric activity of ileum, cecum, and right ventral colon, and cecal emptying of radiolabeled markers in clinically normal ponies. *American Journal of Veterinary Medicine*, 59, pp 328-335.

[27] Ha, I.D. and Lee, Y. (2005). Comparison of hierarchical likelihood versus orthodox best linear unbiased predictor approaches for frailty models. *Biometrika*, 92, pp717-723.

[28] Ha, I.D. and Lee, Y. (2005).Hierarchical generalized linear models (with discussion). *J. R. Statist. Soc.* B58, pp619-678.

[29] Draper,D. (1996). Discussion on ?Hierarchical generalized linear models? (byY. Lee and J. A. Nelder). *J. R. Statist. Soc.* B58, pp662-663.

[30] Lele, S. (1991). Resampling using estimating equations. *Theory of Estimating Equations*. (ed. Godambe, V.P.) Oxford: Clarendon Press.

Vita

Candidate's full name: Jiaxiu Li

Universities Attended:

Master of Science, August 2017, University of New Brunswick,
Fredericton, NB, Canada

Bachelor of Management, June 2014, Shanghai Institute of Technology
Shanghai, China

Conference Presentations:

Li, Jiaxiu, Yan, Guohua and Ma, Renjun (2017). Joint Tweedie mixed models for longitudinal data of mixed types
The 44th Annual Meeting of the Statistical Society of Canada.