A Test and Confidence Set for Comparing the Location of Quadratic Growth Curves

by

Wanchunzi Yu

A Dissertation Presented in Partial Fulfillment
of the Requirement for the Degree
Doctor of Philosophy

Approved June 2015 by the
Graduate Supervisory Committee:

Mark Reiser, Chair
Jarrett Barber
Ming-Hung Kao
Robert D. St Louis
Jeffrey Wilson

ARIZONA STATE UNIVERSITY

August 2015

ABSTRACT

Quadratic growth curves of 2^{nd} degree polynomial are widely used in longitudinal studies. For a 2^{nd} degree polynomial, the vertex represents the location of the curve in the XY plane. For a quadratic growth curve, we propose an approximate confidence region as well as the confidence interval for x and y-coordinates of the vertex using two methods, the gradient method and the delta method. Under some models, an indirect test on the location of the curve can be based on the intercept and slope parameters, but in other models, a direct test on the vertex is required. We present a quadratic-form statistic for a test of the null hypothesis that there is no shift in the location of the vertex in a linear mixed model. The statistic has an asymptotic chi-squared distribution. For 2^{nd} degree polynomials of two independent samples, we present an approximate confidence region for the difference of vertices of two quadratic growth curves using the modified gradient method and delta method. Another chi-square test statistic is derived for a direct test on the vertex and is compared to an F test statistic for the indirect test. Power functions are derived for both the indirect F test and the direct chi-square test. We calculate the theoretical power and present a simulation study to investigate the power of the tests. We also present a simulation study to assess the influence of sample size, measurement occasions and nature of the random effects. The test statistics will be applied to the Tell Efficacy longitudinal study, in which sound identification scores and language protocol scores for children are modeled as quadratic growth curves for two independent groups, TELL and control curriculum. The interpretation of shift in the location of the vertices is also presented.

I dedicate my dissertation to Puyu, and my beloved family.

ACKNOWLEDGEMENTS

To begin with, I would like to truly thank my fantastic advisor, Dr. Mark Reiser. He is not only statistically significant, but also practically significant for me. He has been a constant source and guide of knowledge for the last five years, and without his support and encouragement, this dissertation would not have been possible. Thank you.

Gratitude to my committee, past and present: Dr. Jeffery Wilson, Dr. Jason Kao, Dr. St Louis, Dr. Jennifer Broatch, Dr. Randy Eubank, Dr. Jarrett Barber, and Dr. William Verdini for their valuable time and feedback, as well as their mentorship and professional development. I would also like to thank the graduate coordinator, Debbie Olson, for her help in many administrative tasks and her patience for all my questions about graduation. Thanks to my graduate statistics club colleagues, your endless help made my graduate study more memorable. Very specially, I would like to express my deep appreciation for my three close friends who were right there with me: Wenjun Ke, Jun Zhang and Sarah Burke.

More personally, many thanks to the most important persons in my life who have helped me along the path to a Ph.D. I would like to thank my beloved parents, Defu Yu and Jing Wan, for their tremendous love and support. Lastly, and most importantly, I wish to thank my husband, Puyu San, who is my constant companion, confidant and true friend. He has given me a great deal of help and concern throughout the PhD program.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

Table Page

Table Page

Table Page

LIST OF FIGURES

Chapter 1

INTRODUCTION

Many longitudinal studies are designed to investigate a characteristic of an individual, where the characteristic is measured repeatedly over the occasions for each study subject. Often the observations are considerably correlated across time points. A multivariate model with a general unrestricted covariance structure may be applied to analyze the correlated data, whereas the growth curve model is usually considered. The analysis of growth curves focuses on the explanation of within-individual variation by the aging process or natural development. In some longitudinal studies, the relation between the time measurement and response cannot be adequately described by a linear trend model. Adding a square term of the fixed effect time t to the model gives a quadratic growth curve model, which often describes the true unknown model better. The coefficient parameters of the fixed effect are necessary to determine the growth curve. The vertex of a quadratic curve provides the location of such a curve, which is interesting. By all means reasonable, it is important to derive the confidence region of the parabola's vertex as well as the confidence interval of x-coordinate and y-coordinate.

For two independent groups, such as control and treatment, the confidence region as well as the confidence interval for the difference of vertices of two quadratic growth curves are useful. Both the x-coordinate and y-coordinate of the vertex are given by a non-linear combination of the model fixed regression coefficients, not simply only one of them. However, common statistical computer packages usually display statistical inferences for the fixed regression coefficient, but not for any of their functions.

For a one-sample study, the test of the null hypothesis of no shift in the location can be performed indirectly with an F test on the model parameters. The location of the vertex is a function of the model parameters, and a statistic for a direct test on the location of the vertex is also presented. Power calculations are proposed to investigate the performance of the indirect F test and the direct test. For a two-sample study, the null hypothesis of no difference in the location of the vertices can sometimes be conducted with the indirect F test, but sometimes only the direct test is available. Power calculations for comparing the F test and direct test are performed for the two-sample study.

Two models, linear mixed effects model and growth curve model, and three methods, the gradient method, the delta method and mean response method are reviewed in Chapter 2. For a one-sample quadratic growth curve, test statistics for confidence region and confidence interval of the vertex are derived. To show the validity of the statistics, simulations using different models, parameters and sample sizes as well as power analysis for testing non-vertex points are performed in Chapter 3. For two independent samples, modified test statistics for confidence region and confidence interval of the difference of vertices are obtained and simulations for testing the updated statistics are conducted in Chapter 4, as well as the comparison of the theoretical power and simulated power for both the direct and indirect test. In Chapter 5, an application of analysis, TELL Efficacy Study, using the derived statistics is presented. The conclusion and discussion for future research are presented in Chapter 6.

Chapter 2

LITERATURE REVIEW

2.1 Longitudinal Study

The defining feature of a longitudinal study is that it involves repeated observations of the same variable over long periods of time, thereby allowing the direct study of change over time. The fundamental objective of a longitudinal analysis is the assessment of within-individual changes in the response and the explanation of systematic differences among individuals in their changes. Rao (1965) described a two conceptually distinct stages for longitudinal analysis of within-individual changes. Stage 1, within-individual change in the response is characterized in terms of some appropriate summary of the changes in the repeated measurements on each individual during the occasions. Stage 2, the relationship between theses estimates of within-individual changes and the inter-individual differences in selected covariates (Bijleveld *et al.*, 1999). The goal of a longitudinal analysis is to determine whether the individuals have higher or lower values on selected covariates provided that certain individuals change more or less than others. It may also be interesting to make predictions about how specific subjects change over time in some longitudinal studies. In the second situation, the prediction of longitudinal studies are more reliable since they borrow the information form all individuals to give a better prediction of the within-individual change over time for a specific subject.

A distinctive feature of longitudinal data is that the data are clustered. Clustering is grouping a set of observations in such a way that observations in the same cluster are more similar to each other than to those in other clusters. In longitudinal studies

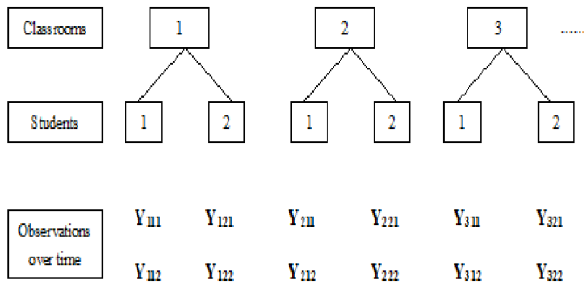


Figure 2.1: Nested Structure for Students within Classrooms

clusters are composed of the repeated measurements obtained from a single individual at different occasions. Observations within a cluster will typically exhibit positive correlation, and this correlation must be accounted for in the analysis. Longitudinal data also have a temporal order; the first time point within a cluster necessarily comes before the second time point, and so on. The ordering of the repeated measures has important implications for analysis. When all the individuals are measured at a commons set of occasions and there are no missing values, the longitudinal data are balanced and complete. However, it is more common that the longitudinal data are unbalanced and/or incomplete. As a consequence, to be of real practical use, methods for the analysis of longitudinal data must be able to handle data that are unbalanced over time and possibly incomplete (Fitzmaurice *et al.*, 2004). For instance a two-stage nested design, it is often used in analyzing processes to identify the main sources of variability. In a two-stage nested design, the levels of one factor are nested under the levels of the other factor, such as students nested within classrooms or patients nested within physicians. Figure 2.1 shows the concept of a nested model intuitively.

2.2 Linear Mixed Model and Growth Curve

A mixed model is a statistical model containing mixed effects, where the mixed effects consists of both fixed effects and random effects. They are appropriate in settings where repeated measurements are provided on the same individual, or where measurements are made on clusters of related individuals. Random effects models were first introduced by Fisher (1919) to study the correlations of trait values between relatives. Afterwards, the best linear unbiased estimates (BLUE) for fixed effects and best linear unbiased predictions (BLUP) for random effects were provided by Henderson *et al.* (1959). Subsequently, mixed modeling has become a major area of statistical research, including work in many fields, such as computing maximum likelihood estimates, missing data in mixed effects models, non-linear mixed effect models, and Bayesian estimation of mixed effects models. Mixed models are applied in many disciplines where multiple correlated measurements are made on each individual of interest.

Mixed models are based on explicit identification of individual and population characteristics; most mixed models in the literature can be described either as growth models or as repeated-measures models. Growth-curve analyses emphasize the explanation of within-subject variation by the nature developmental or aging process (Ware, 1985). These analyses often compare growth characteristics for different populations, emphasizing the contribution of experimental conditions to between-subject variability (Laird and Ware, 1982).

A linear mixed model for longitudinal data can be expressed in matrix notation,

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\alpha}_i + \boldsymbol{\epsilon}_i, \tag{2.1}$$

where,

$\boldsymbol{y}_i$ is a known vector of observations for subject i, $\boldsymbol{Y}' = [\boldsymbol{y}_1', \cdots, \boldsymbol{y}_N']$,

$\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ are known model matrices of regressors for subject i relating the observations $\boldsymbol{y}_i$ to $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}_i$, $\boldsymbol{X}' = [\boldsymbol{X}_1', \cdots, \boldsymbol{X}_N']$,

$\boldsymbol{\beta}$ is an unknown vector of fixed effects parameters,

$\boldsymbol{\alpha}_i$ is an unknown vector of random effects with mean $E(\boldsymbol{\alpha}_i) = \mathbf{0}$ and covariance $\text{Cov}(\boldsymbol{\alpha}_i) = \boldsymbol{G}$; the covariance matrix $\boldsymbol{G}$ is usually identical for all the subjects,

$\boldsymbol{\epsilon}_i$ is an unknown vector of random error terms with mean $E(\boldsymbol{\epsilon}_i) = \mathbf{0}$ and covariance $\text{Cov}(\boldsymbol{\epsilon}_i) = \boldsymbol{R}_i$; the set of unknown parameters in $\boldsymbol{R}_i$ do not depend on the subject i, only the dimension of $\boldsymbol{R}_i$ depends on the subject i,

$\boldsymbol{\alpha}_i$ and $\boldsymbol{\epsilon}_i$ are independent, $\text{Cov}(\boldsymbol{\alpha}_i, \boldsymbol{\epsilon}_i) = \mathbf{0}$, that is,

$$E\begin{pmatrix} \boldsymbol{\alpha}_i \\ \boldsymbol{\epsilon}_i \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \qquad \text{Cov}\begin{pmatrix} \boldsymbol{\alpha}_i \\ \boldsymbol{\epsilon}_i \end{pmatrix} = \begin{pmatrix} \boldsymbol{G} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R}_i \end{pmatrix}.$$

Assuming the random effect $\boldsymbol{\alpha}_i$ is known, the conditional distribution of $\boldsymbol{y}_i$ given $\boldsymbol{\alpha}_i$ is multivariate normal,

$$\boldsymbol{y}_i | \boldsymbol{\alpha}_i \sim \boldsymbol{N}_n(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\alpha}, \boldsymbol{R}_i).$$

Further, $\boldsymbol{\alpha}_i$ is assumed to be normally distributed with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{G}$. Then the marginal density function of the random vector $\boldsymbol{y}_i$ is given by (Verbeke and Molenberghs, 2009),

$$f(\boldsymbol{y}_i) = \int f(\boldsymbol{y}_i|\boldsymbol{\alpha}_i) f(\boldsymbol{\alpha}_i) d\boldsymbol{\alpha}_i$$

which is multivariate normally distributed with the dimension of time measurements n, i.e. the marginal model of $\boldsymbol{y}_i$ is,

$$\boldsymbol{y}_i \sim \boldsymbol{N}_n(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i).$$

When all the covariance parameters are known, the maximum likelihood (ML) function of $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\alpha}_i)'$ is (Verbeke and Molenberghs, 2009),

$$\begin{aligned} &L_{ML}(\boldsymbol{\theta}) \\ &= \prod_{i=1}^{N} \left\{ (-2\pi)^{-n/2} \left|\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i\right|^{-1/2} \times exp\left(\sum_{i=1}^{n} (\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}_{ML})'(\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i)^{-1}(\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}_{ML}) \right) \right\}, \end{aligned}$$

where N is the sample size. The ML estimator for fixed regression coefficients and their variance are (Laird and Ware, 1982),

$$\hat{\boldsymbol{\beta}}_{ML} = \left(\sum_{i=1}^{N} \boldsymbol{X}_i'(\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i)^{-1}\boldsymbol{X}_i\right)^{-1} \left(\sum_{i=1}^{N} \boldsymbol{X}_i'(\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i)^{-1}\boldsymbol{y}_i\right) \tag{2.2}$$

$$\Sigma_{\hat{\boldsymbol{\beta}}_{ML}} = \left(\sum_{i=1}^{N} \boldsymbol{X}_i'(\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i)^{-1}\boldsymbol{X}_i\right)^{-1}. \tag{2.3}$$

Denote $\boldsymbol{\zeta}$ as the vector of variance and covariance parameters found in $\boldsymbol{R}_i$ and $\boldsymbol{G}$. The restricted maximum likelihood (REML) function of $\boldsymbol{\zeta}$ is (Verbeke and Molenberghs, 2009),

$$\begin{aligned} L_{REML}(\boldsymbol{\zeta}) =& (2\pi)^{-(n-k)/2} \left|\sum_{i=1}^{N} \boldsymbol{X}_i'\boldsymbol{X}_i\right|^{1/2} \\ &\times \left|\sum_{i=1}^{N} \boldsymbol{X}_i'(\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i)^{-1}\boldsymbol{X}_i\right|^{-1/2} \prod_{i=1}^{N} |\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i|^{-1/2} \\ &\times exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}_{ML})'(\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i)^{-1}(\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}_{ML})\right\}. \end{aligned}$$

The $\boldsymbol{\zeta}$ is a function of a set of error contrasts $\boldsymbol{U} = \boldsymbol{A}'\boldsymbol{Y}$ where $\boldsymbol{A}$ is any $(n \times (n-k))$ full-rank matrix with columns orthogonal to the columns of the $\boldsymbol{X}$ matrix (Verbeke and Molenberghs, 2009). Then for each individual i, the REML estimator through an empirical bayesian algorithm for the random effect and its variance are (Laird and Ware, 1982),

$$\hat{\boldsymbol{\alpha}}_{i(REML)} = \boldsymbol{G}\boldsymbol{Z}_i'(\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i)^{-1}(\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}_{ML}) \tag{2.4}$$

$$\Sigma_{\hat{\boldsymbol{\alpha}}_{i(REML)}} = \boldsymbol{G}\boldsymbol{Z}_{\boldsymbol{i}}' \left\{(\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i)^{-1} - (\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i)^{-1}\boldsymbol{X}_i\hat{\Sigma}_{\hat{\boldsymbol{\beta}}_{ML}}\boldsymbol{X}_i'(\boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i)^{-1}\right\} \boldsymbol{Z}_i\boldsymbol{G},$$

assuming that the necessary matrix inverses exist when it is implied. For the case of less than full rank, we could work out the relevant formulas using generalized inverses. When the covariance matrices are unknown, the literature on the estimation of variance components is extensive. By default, the estimation method for the covariance parameters used in SAS is REML through expectation and maximization (EM) algorithm (Laird and Ware, 1982).

The $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\alpha}_i)'$ is the parameter vector and let $\boldsymbol{y}_i$ be a function of $\boldsymbol{\theta}$, say $\boldsymbol{y}_i = f(\boldsymbol{\theta})$, $\boldsymbol{y}_i \in \mathbb{R}^s$. If $\boldsymbol{\theta}^*$ denotes the true parameter vector, under $H_0 : \boldsymbol{\theta}^* \in \Theta_0 \subset R^s$, then the regularity conditions for the delta method, used frequently to obtain the asymptotic distributions, and maximum likelihood estimators (MLE) for the linear mixed model can be stated as follows,

1. $\boldsymbol{y}_i, i = 1, ...N$ are independently and identically distributed (i.i.d.) with probability density function $f(\boldsymbol{y}_i; \boldsymbol{\theta})$.

2. The probability distribution is identifiable. That is, the probability distribution are distinct for different parameters $\boldsymbol{\theta}$; if $\boldsymbol{\theta} \neq \boldsymbol{\theta}'$, then $f(\boldsymbol{y}_i; \boldsymbol{\theta}) \neq f(\boldsymbol{y}_i; \boldsymbol{\theta}')$.

3. The parameter space Θ_0 is compact and there exists a $\boldsymbol{\theta}^* \in Interior(\Theta_0)$ such that $E_{\boldsymbol{\theta}^*}\ log\ f(\boldsymbol{y}_i; \boldsymbol{\theta})$ exists and $\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta} \in \Theta}{argmax}\ E_{\boldsymbol{\theta}^*}\ log\ f(\boldsymbol{y}_i; \boldsymbol{\theta})$.

4. The probability density function is positive, i.e. $f(\boldsymbol{y}_i; \boldsymbol{\theta}) > 0$ and is three times continuously differentiable in $\boldsymbol{\theta}$ in some neighborhood of $\boldsymbol{\theta}^*$.

5. The integration and differential operators are interchangeable.

6. The Jacobian matrix $\frac{\partial \boldsymbol{f}(\boldsymbol{\theta}^*)}{\partial \boldsymbol{\theta}}$ is of full rank.

7. The mapping $\boldsymbol{f} : \Theta_0 \mapsto \boldsymbol{y}_i$ is continuous at every point $\boldsymbol{\theta} \in \Theta_0$.

8. The fisher information matrix $I(\boldsymbol{\theta}^*) = E_{\boldsymbol{\theta}^*}\left(\frac{\partial^2 log\ f(\mathbf{y}_i;\theta^*)}{\partial\theta\partial\theta'}\right)$ exists and is nonsingular.

9. The first and second derivative of the log-likelihood function $log\ f(\boldsymbol{y}_i; \boldsymbol{\theta})$ are defined and the boundary is $\left|\frac{\partial^3 log\ f(\boldsymbol{y}_i;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^3}\right| \leqslant M(\boldsymbol{y}_i)$ with $E(M(\boldsymbol{y}_i)) < \infty$.

2.3 Modeling the Covariance

Approaches for appropriately modeling the covariance or time dependence among the repeated measures obtained on the same individuals must be considered because of the correlated feature of longitudinal data. When an appropriate structure for the covariance has been achieved, correct standard errors are obtained and inferences about the regression parameters can be made (Fitzmaurice *et al.*, 2004). In this dissertation, three covariance structures are considered; they are compound symmetry (CS), autoregressive one (AR(1)) and unstructured (UN).

One of the earliest covariance pattern models considered historically for the analysis of repeated measures data was compound symmetry. The compound symmetry covariance structure has a randomization justification in certain repeated measures designs. It assumes the correlation between any pair of measurements is the same regardless of the time interval between the measurements, which is a quite strong assumption. Therefore, with a compound symmetry covariance structure, the variance is constant across occasions, say σ_e^2 $(0 < \sigma_e^2 < \infty)$, and the correlation of any two responses at different occasions, j and j', for the same individual i is $\mathrm{Corr}(y_{i,j}, y_{i,j'}) = \rho$ for $|\rho| \leqslant 1$,

$$\mathrm{Cov}(\boldsymbol{y}_i) = \Sigma_{\boldsymbol{y}_i} = \sigma_e^2 \begin{pmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \rho \\ \rho & \rho & \rho & \cdots & 1 \end{pmatrix},$$

where y_{ij} denotes the response variable for the i^{th} individual at j^{th} occasion and $\boldsymbol{y}_i$ denotes the response vector of individual i at all occasions as before.

The autoregressive covariance structure is a parsimonious as compound symmetry, since it also has only two parameters, regardless of the number of time points. In the autoregressive model, it is assumed that the variance is constant across occasions, say σ_e^2, and the correlation of any two responses at different occasions for the same individual i is $\mathrm{Corr}(y_{i,j}, y_{i,j'}) = \rho^{|j'-j|}$ for all j and j', and ρ,

$$\mathrm{Cov}(\boldsymbol{y}_i) = \Sigma_{\boldsymbol{y}_i} = \sigma_e^2 \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{pmatrix}.$$

When the number of time points is relatively small and all individuals are measured at the same set of time points, it maybe reasonable to consider the unstructured covariance structure which allows the covariance matrix to be arbitrary, with all of its elements unconstrained. Let the covariance matrix be symmetric and positive-definite is the only formal requirement. The main advantage of an unstructured covariance is that no assumptions are made about the structure of variances and covariances. One potential drawback is that the number of covariance parameters to be estimated grows rapidly with the number of measurement occasions (Fitzmaurice *et al.*, 2004). For each individual i with n measurement occasions, the unstructured covariance matrix has $\frac{n\times(n+1)}{2}$ parameters, the n variances at each occasion and the $\frac{n\times(n-1)}{2}$ pairwise covariances,

$$\mathrm{Cov}(\boldsymbol{y}_i) = \Sigma_{\boldsymbol{y}_i} \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix},$$

where $\mathrm{Var}(y_{ij}) = \sigma_j^2$ and $\mathrm{Cov}(y_{i,j}, y_{i,j'}) = \sigma_{jj'}$; the correlation is $\mathrm{Corr}(y_{i,j}, y_{i,j'}) = \frac{\sigma_{jj'}}{\sigma_j \sigma_{j'}}$.

Given different models, the correlation ρ between two responses with respect to the same individual at different occasions may be different. Given the correlation ρ, different covariance structures can be expressed for specific models. To make a selection among models with different types of covariance structure, information criteria can be applied. To compare non-nested models, Akaike Information Criterion (AIC), Akaike (1974), is one of the earliest proposed information criteria. According to the AIC, given a set of competing models for the covariance, the model should be selected with minimum

$$AIC = -2(\hat{l} - c),$$

where $\hat{l}$ is the maximized REML log-likelihood and c is the number of covariance parameters. The preferred model is the one with the minimum AIC value, given a set of candidate models. Hence AIC not only rewards goodness of fit, but also includes a penalty that is an increasing function of the number of estimated parameters. However, Woodroofe (1982) showed that AIC is not theoretically consistent; consequently, the correct model will not be selected when sample size (N) approaches infinity. AICc, is AIC with a finite population correction, is produced as

$$AICc = AIC + \frac{2c(c+1)}{N-c-1},$$

and AICc is similar to AIC with a greater penalty for extra parameters. Burnham and Anderson (2002) strongly recommend using AICc, rather than AIC, if N is small or c is large even though AICc is also not consistent (Eslamian, 2014). Since AICc converges to AIC as N gets large, AICc generally should be employed. Another valuable criterion is Bayesian Information Criterion (BIC). Schwarz *et al.* (1978) proposed that when choosing competing models for the covariance structure, the model should be selected with minimum

$$BIC = -2(\hat{l} - c\ log\sqrt{N}).$$

Bozdogan (1987) proved BIC is consistent when the sample size approaches infinity.

2.4 Vertex of Quadratic Curve

The vertex of a quadratic growth curve provides the location of the curve, which is interesting to be investigated. In the geometry of curves, a vertex is defined as a point where the first derivative of curvature is zero (Agoston, 2005). This is typically a local maximum or minimum of curvature in the optimization field (Gibson, 2001).

A quadratic function, in mathematics, is a polynomial function of the form

$$f(x) = ax^2 + bx + c, \qquad a \neq 0.$$

where a, b and c denote coefficients for quadratic term, linear term and intercept respectively. The graph of a quadratic function is a parabola whose axis of symmetry is parallel to the $y-$axis. The expression $ax^2 + bx + c$ in the definition of a quadratic function is a polynomial of degree 2 or second order, or a 2^{nd} degree polynomial, because the highest exponent of x is the second degree.

The vertex of a quadratic curve is also called the turning point since it is the location when the curve turns. By the method of completing the square, the standard form of a quadratic function can be expressed as

$$f(x) = a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a}, \qquad a \neq 0,$$

thus the vertex of the curve in the vector form is

$$\left(-\frac{b}{2a}, -\frac{b^2 - 4ac}{4a}\right). \tag{2.5}$$

If $a < 0$, the vertex is the maximum point; otherwise, if $a > 0$, the vertex is the minimum point. The vertex point can be also be obtained by finding the roots of the first derivative using calculus:

$$f'(x) = 2ax + b,$$

with the corresponding function value,

$$f(x) = a\left(-\frac{b}{2a}\right)^2 + b\left(-\frac{b}{2a}\right) + c = -\frac{b^2 - 4ac}{4a},$$

therefore again the vertex point can be expressed as $\left(-\frac{b}{2a}, -\frac{b^2-4ac}{4a}\right)$. The vertical line $x = -\frac{b}{2a}$ that passes through the vertex is also the axis of symmetry of the quadratic curve.

2.5 Delta Method

The delta method is a method for deriving an approximate probability distribution for a function of an asymptotically normal statistical estimator from knowledge of the limiting variance of that estimator. More broadly, the delta method is known as a generalization of the Central Limit Theorem using Taylor series approximations for mean and variance. Using a Taylor series expansion if a function $g(Y)$ has derivatives of order r, that is, $g^r(Y) = \frac{d^r}{dy^r}g(Y)$ exists, then for any constant a, Casella and Berger (2002) displayed the Taylor polynomial of order r about a as,

$$T_r(Y) = \sum_{i=0}^{r} \frac{g^{(i)}(a)}{i!}(Y-a)^i.$$

The major of Taylor theorem is that the remainder from the approximation, $g(Y) - T_r(Y)$ always tends to zero faster than the highest-order explicit term, namely,

$$\lim_{x\to a} \frac{g(Y) - T_r(Y)}{(Y-a)} = 0.$$

Hence we can drop the higher-order terms to give the first-order approximation,

$$g(Y) \approx g(a) + g'(a)(Y-a) \ .$$

Let $a = \mu$, the mean of random variable Y, a Taylor series expansion of $g(Y)$ about μ gives the approximation,

$$g(Y) = g(\mu) + g'(\mu)(Y - \mu) \ .$$

Taking the variance of both sides yields,

$$\text{Var}(g(Y)) \approx (g'(\mu))^2 \text{Var}(Y) \ .$$

For the univariate delta method (Casella and Berger, 2002), the function $g(Y)$ is a real-valued continuous function of Y. Let Y_N be a sequence of random variables that satisfies $\sqrt{N}(Y_N - \mu) \xrightarrow{D} N(0, \sigma^2)$, $0 < \sigma^2 < \infty$. For a given function g and a specific value of μ, suppose that $g'(\mu)$ exists and is not 0, then

$$\sqrt{N}(g(Y_N) - g(\mu)) \xrightarrow{D} N(0, \sigma^2(g'(\mu))^2).$$

For the multivariate delta method (Casella and Berger, 2002), define the random vector $\boldsymbol{Y} = (Y_1, ..., Y_p)$ with mean $\boldsymbol{\mu} = (\mu_1, ..., \mu_p)$ and covariances $\text{Cov}(Y_i, Y_j) = \sigma_{ij}$. An $i.i.d$ random sample of size N from the population of $\boldsymbol{Y}$ can be observed and denote these observations as $\boldsymbol{Y}^{(1)}, ..., \boldsymbol{Y}^{(N)}$. Furthermore, let the sample means for each element of the vector $\boldsymbol{Y}$ be $\bar{Y}_i = \frac{\sum_{k=1}^{N} Y_i^{(k)}}{N}$, $i = 1, ..., p$ and $\bar{\boldsymbol{Y}} = (\bar{Y}_1, ..., \bar{Y}_p)$ be the vector of sample means. Consequently, we consider the multivariate function $g : \mathbb{R} \mapsto \mathbb{R}$ with $g(\boldsymbol{Y}) = g(Y_1, ..., Y_p)$, and use a Taylor series expansion to write (Rencher and Schaalje, 2008)

$$g(\bar{Y}_1, ..., \bar{Y}_p) \approx g(\mu_1, ...\mu_p) + \sum_{i=1}^{p} g_i'(\mu_i)(\bar{Y}_i - \mu_i).$$

In vector notation (Papanicolaou, 2009),

$$g(\bar{\boldsymbol{Y}}) \approx g(\boldsymbol{\mu}) + \nabla' g(\boldsymbol{\mu})(\bar{\boldsymbol{Y}} - \boldsymbol{\mu}), \tag{2.6}$$

with the notation $\nabla' g(\boldsymbol{\mu}) = (\nabla' g(\boldsymbol{Y}))|_{\boldsymbol{Y}=\boldsymbol{\mu}}$. The multivariate delta method in vector form is, let $\boldsymbol{Y}^{(1)}, ..., \boldsymbol{Y}^{(N)}$ be a random sample with $E(\boldsymbol{Y}^{(k)}) = \boldsymbol{\mu}$ and covariance matrix $E(\boldsymbol{Y}^{(k)} - \boldsymbol{\mu})(\boldsymbol{Y}^{(k)} - \boldsymbol{\mu})' = \boldsymbol{\Sigma}$. For a given function g with continuous first partial derivatives and a specific value of $\boldsymbol{\mu}$ for which $\boldsymbol{\tau}^2 = \nabla' g(\boldsymbol{\mu})\boldsymbol{\Sigma}\nabla g(\boldsymbol{\mu}) > \mathbf{0}$,

$$\sqrt{N}(g(\bar{\boldsymbol{Y}}) - g(\boldsymbol{\mu})) \xrightarrow{D} N_p(\mathbf{0}, \boldsymbol{\tau}^2). \tag{2.7}$$

2.6 Interval Estimation for Mean Response

When the x-coordinate of the vertex of a quadratic growth curve is obtained, the y-coordinate of the vertex can be estimated as the mean response of the X value. Denote X_h the level of X for which we would like to estimate the mean response. Then X_h may be a value which occurred in the sample, or it may be some other value within the scope of the data. The mean response is denoted by $E\{Y_h\}$ at $X = X_h$. If repeated samples were selected, each holding the levels of the variable $X = X_h$, the sampling distribution of $\hat{Y}_h$ with regard to the different values of $\hat{Y}_h$ that would be obtained by calculating $\hat{Y}_h$ for each sample. For the normal error fixed effects model with i.i.d. observation, the sampling distribution of $\hat{Y}_h$ is normal, with mean $E\{\hat{Y}_h\} = E\{Y_h\}$ and variance $\mathrm{Var}\{\hat{Y}_h\} = \sigma^2 \left[\frac{1}{N} + \frac{(X_h - \bar{X})^2}{\Sigma(X_i - \bar{X})^2}\right]$ (Kutner *et al.*, 2005). When the mean square error (MSE) is substituted for σ^2, the estimated variance of $\hat{Y}_h$, $s^2\{\hat{Y}_h\}$, is

$$s^2\{\hat{Y}_h\} = MSE\left[\frac{1}{N} + \frac{(X_h - \bar{X})^2}{\Sigma(X_i - \bar{X})^2}\right] .$$

Then the positive square root of $s^2\{\hat{Y}_h\}$ is $s\{\hat{Y}_h\}$, the estimated standard deviation of $\hat{Y}_h$. Hence,

$$\frac{\hat{Y}_h - E\{Y_h\}}{s\{\hat{Y}_h\}} \quad \text{is distributed as t distribution with } (N-p) \text{ degrees of freedom} \tag{2.8}$$

where p is the number of fixed regression coefficients. A confidence interval for $E\{Y_h\}$ is constructed using the t distribution. The $(1-\alpha)\%$ confidence limits are,

$$\hat{Y_h} \pm t(1-\alpha/2; N-p)s\{\hat{Y_h}\},$$

where α is the type I error rate (Kutner *et al.*, 2005). For a quadratic model, if the vertex $\boldsymbol{V'} = (V_x, V_y)$ exists and the value of x-coordinate V_x is known, we could estimate the value of y-coordinate $\hat{V_y}$ by substituting V_x in the regression model and the standard deviation $s\{\hat{V_y}\} = MSE\left[\frac{1}{N} + \frac{(V_x - \bar{X})^2}{\Sigma(X_i - \bar{X})^2}\right]$. Using distribution (2.8), the $(1-\alpha)\%$ confidence limits of $\hat{V_y}$ are,

$$\hat{V_y} \pm t(1-\alpha/2; N-p)s\{\hat{V_y}\}\ .$$

We consider the confidence interval for a mixed linear model will be given in Chapter 3.

2.7 Confidence Set for X-Coordinate With a Given Gradient

Bachmaier (2009) proposed an exact confidence set for the $x-$coordinate for fixed effects quadratic model with a given gradient; the model is given by

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i\ , \tag{2.9}$$

where y_i denotes the response variable, β_0, β_1 and β_2 are fixed regression coefficients, N is the number of observations, and the errors, ϵ_i, are assumed to be independent and normally distributed random variables with an expected value 0 and a common unknown variance $\sigma^2 > 0$, i.e. $\epsilon_i \sim N(0,\ \sigma^2)$. The function $E(y_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$ is a parabola with respect to x_i. The x-coordinate where this function has given gradient m leads to

$$x_{\text{gigrad}} = \frac{m - \beta_1}{2\beta_2} \quad \text{if} \quad \beta_2 \neq 0\ . \tag{2.10}$$

A point estimate of (2.10) is $\hat{x}_{\text{gigrad}} = (m - b_1)/(2b_2)$, where b_1 and b_2 are the least squares estimates of β_1 and β_2.

An exact $(1-\alpha)$-confidence set for x_{gigrad} is obtained as a solution of function (2.11), where x_0 is any point in the confidence set and $t^2_{N-3,1-\alpha/2}$ denotes the squared t-quantile with $N-3$ degrees of freedom, (Bachmaier, 2009):

$$
\begin{aligned}
& x_0 \in C(x_{\text{gigrad}}) \\
\Leftrightarrow\ & \frac{(b_1 - m + 2x_0 b_2)^2}{\hat{\text{Var}}(b_1) + 4x_0 \hat{\text{Cov}}(b_1, b_2) + 4x_0^2 \hat{\text{Var}}(b_2)} \leqslant t^2_{N-3,1-\alpha/2} \\
\Leftrightarrow\ & (b_1 - m + 2x_0 b_2)^2 \leqslant [\hat{\text{Var}}(b_1) + 4x_0 \hat{\text{Cov}}(b_1, b_2) + 4x_0^2 \hat{\text{Var}}(b_2)] \cdot t^2_{N-3,1-\alpha/2} \\
\Leftrightarrow\ & A \cdot x_0^2 + B \cdot x_0^2 + C \leqslant 0 \ .
\end{aligned} \tag{2.11}
$$

$$
\begin{aligned}
\text{where, } A &= b_2^2 - \hat{\text{Var}}(b_2) \cdot t^2_{N-3,1-\alpha/2} \\
B &= (b_1 - m)b_2 - \hat{\text{Cov}}(b_1, b_2) \cdot t^2_{N-3,1-\alpha/2} \\
C &= \frac{1}{4}((b_1 - m)^2 - \hat{\text{Var}}(b_1) \cdot t^2_{N-3,1-\alpha/2}) \ .
\end{aligned}
$$

Only if the denominator in expression (2.11) is positive, the medium equivalent sign can be applied; this is satisfied if the mean square error is positive. The mean square error equals to zero only occurs with probability 0, therefore an optional confidence interval for this case is chosen without violating the coverage probability of the confidence interval. To solve the inequality, if $A \neq 0$, then $A \cdot x_0^2 + B \cdot x_0^2 + C$ is a quadratic function; it has two solutions if the discriminant $\text{Dis} = B^2 - 4AC$ is positive. With regard to the numerical stability concerning small values of $4AC$, we compute either zero in two different ways (Bachmaier, 2009):

$$
x_{01} = \begin{cases} \frac{-2C}{B - \sqrt{B^2 - 4AC}} & \text{when } B < 0, \\ \frac{-B - \sqrt{B^2 - 4AC}}{2A} & \text{when } B \geqslant 0. \end{cases}
\qquad
x_{02} = \begin{cases} \frac{-B + \sqrt{B^2 - 4AC}}{2A} & \text{when } B \leqslant 0, \\ \frac{-2C}{B + \sqrt{B^2 - 4AC}} & \text{when } B > 0. \end{cases}
$$

Therefore when $A > 0$ and $\text{Dis} > 0$, this leads to a two-sided confidence interval $[x_{01}, x_{02}]$. When $A < 0$ and $\text{Dis} > 0$, the exact confidence interval goes to

$(-\infty, x_{02}] \bigcup [x_{01}, +\infty)$. The confidence interval for a mixed linear model will be given in Chapter 3.

2.8 Confidence Set for the Difference of the X-Coordinates of Two Vertices

Bachmaier (2010) published the confidence set for the difference of the $x-$coordinate of two quadratic regression models. For two independent samples, such as control and treatment groups, the quadratic regression model given in Bachmaier's paper is,

$$y_i = \beta_0^{(\text{mid})} + \beta_0^{(\text{eff})} I_i + \beta_1^{(\text{mid})} x_i + \beta_1^{(\text{eff})} I_i x_i + \beta_2^{(\text{mid})} x_i^2 + \beta_2^{(\text{eff})} I_i x_i^2 + \epsilon_i \qquad (2.12)$$

where

$$I_i = \begin{cases} -1 & \text{if } y_i \text{ comes from control group } C, \\ +1 & \text{if } y_i \text{ comes from treatment group } T. \end{cases}$$

is a dummy variable to indicate the group, and other parameters are defined as explained in model (2.9). From model (2.12), the distinct models for control and treatment groups are,

$$y_i = \beta_0^{(\text{C})} + \beta_1^{(\text{C})} x_i + \beta_2^{(\text{C})} x_i^2 + \epsilon_i \qquad \text{for group } C,$$

$$y_i = \beta_0^{(\text{T})} + \beta_1^{(\text{T})} x_i + \beta_2^{(\text{T})} x_i^2 + \epsilon_i \qquad \text{for group } T,$$

with the coefficients,

$$\beta_k^{(\text{C})} = \beta_k^{(\text{mid})} - \beta_k^{(\text{eff})} \qquad \text{for } k = 0, 1, 2,$$

$$\beta_k^{(\text{T})} = \beta_k^{(\text{mid})} + \beta_k^{(\text{eff})} \qquad \text{for } k = 0, 1, 2.$$

The difference of $x-$coordinates of vertices from two groups can be tested for model (2.12). In order to obtain an exact F-test or t-test for testing the vertices and the corresponding confidence set within the framework of general linear models, linearity of the difference of x-coordinate of the vertices with respect to regression coefficients is required. Therefore one more assumption is needed, which is equal quadratic coefficients in both groups, i.e.,

$$\beta_2^{(\mathrm{C})} = \beta_2^{(\mathrm{T})}, \text{or equivalent, } \beta_2^{(\mathrm{eff})} = 0.$$

Denote $\boldsymbol{V}^{(\mathrm{C})'} = (V_x^{(\mathrm{C})}, V_y^{(\mathrm{C})})$ and $\boldsymbol{V}^{(\mathrm{T})'} = (V_x^{(\mathrm{T})}, V_y^{(\mathrm{T})})$ as the vectors of vertices for the two groups respectively. With the additional assumption, the point estimates of the x-coordinate are,

$$\hat{V_x}^{(\mathrm{C})} = \frac{-b_1^{(\mathrm{C})}}{2b_2^{(\mathrm{C})}} = \frac{-(b_1^{(\mathrm{mid})} - b_1^{(\mathrm{eff})})}{2b_2^{(\mathrm{mid})}}, \qquad \hat{V_x}^{(\mathrm{T})} = \frac{-b_1^{(\mathrm{T})}}{2b_2^{(\mathrm{T})}} = \frac{-(b_1^{(\mathrm{mid})} + b_1^{(\mathrm{eff})})}{2b_2^{(\mathrm{mid})}}.$$

where $b_1^{(\mathrm{mid})}$, $b_1^{(\mathrm{eff})}$, and $b_2^{(\mathrm{mid})}$ are the least squares estimates of fixed regression coefficients. The difference of x-coordinate of vertices, $V_x^{(\mathrm{diff})} = V_x^{(\mathrm{T})} - V_x^{(\mathrm{C})} = 0$ is equivalent to $\beta_1^{(\mathrm{eff})} + \beta_2 V_x^{(\mathrm{diff})} = 0$, which is a linear combination of β's. Then $\beta_1^{(\mathrm{eff})} + \beta_2 V_x^{(\mathrm{diff})}$ distributes as a t distribution with $N-5$ degrees of freedom, where N is the sample size and the lost 5 degrees of freedom is with regard to the number of regression coefficients β's, the confidence interval for $V_x^{(\mathrm{diff})}$ is (Bachmaier, 2010):

$$\begin{aligned}
& V_x^{(\mathrm{diff})} \in C(V_x^{(\mathrm{T})} - V_x^{(\mathrm{C})}) \\
\Leftrightarrow & \frac{\left(b_1^{(\mathrm{eff})} + b_2 V_x^{(\mathrm{diff})}\right)^2}{\hat{\mathrm{Var}}(b_1^{(\mathrm{eff})}) + 2V_x^{(\mathrm{diff})}\hat{\mathrm{Cov}}(b_1^{(\mathrm{eff})}, b_2) + \left[V_x^{(\mathrm{diff})}\right]^2 \hat{\mathrm{Var}}(b_2)} \leqslant t^2_{N-5,1-\alpha/2} \\
\Leftrightarrow & \left(b_1^{(\mathrm{eff})} + b_2 V_x^{(\mathrm{diff})}\right)^2 \leqslant \\
& \left(\hat{\mathrm{Var}}(b_1^{(\mathrm{eff})}) + 2V_x^{(\mathrm{diff})}\hat{\mathrm{Cov}}(b_1^{(\mathrm{eff})}, b_2) + \left[V_x^{(\mathrm{diff})}\right]^2 \hat{\mathrm{Var}}(b_2)\right) \cdot t^2_{N-5,1-\alpha/2} \\
\Leftrightarrow & A \cdot \left[V_x^{(\mathrm{diff})}\right]^2 + B \cdot V_x^{(\mathrm{diff})} + C \leqslant 0, \\
& \quad \mathit{where},\ A = b_2^2 - \hat{\mathrm{Var}}(b_2) \cdot t^2_{N-5,1-\alpha/2} \\
& \quad B = 2b_1^{(\mathrm{eff})} b_2 - 2\hat{\mathrm{Cov}}(b_1^{(\mathrm{eff})}, b_2) \cdot t^2_{N-5,1-\alpha/2} \\
& \quad C = \left[b_1^{(\mathrm{eff})}\right]^2 - \hat{\mathrm{Var}}(b_1^{(\mathrm{eff})}) \cdot t^2_{N-5,1-\alpha/2}\ .
\end{aligned} \tag{2.13}$$

The inequality in (2.13) reveals the isotonicity of the confidence interval with regard to the confidence level. The terms in the brackets of the right-hand side give a variance, which cannot be negative, hence the right side is monotone increasing with the confidence level. To solve this inequality (2.13), if $A \neq 0$, then $A \cdot \left[V_x^{(\text{diff})}\right]^2 + B \cdot V_x^{(\text{diff})} + C$ is a quadratic function. It has two roots if the discriminant $\text{Dis} = B^2 - 4AC$ is positive. With respect to the numerical stability concerning small values of $4AC$, we compute either root in two different ways (Bachmaier, 2010):

$$x_{01} = \begin{cases} \frac{-2C}{B-\sqrt{B^2-4AC}} & \text{when } B < 0, \\ \frac{-B-\sqrt{B^2-4AC}}{2A} & \text{when } B \geqslant 0. \end{cases} \qquad x_{02} = \begin{cases} \frac{-B+\sqrt{B^2-4AC}}{2A} & \text{when } B \leqslant 0, \\ \frac{-2C}{B+\sqrt{B^2-4AC}} & \text{when } B > 0. \end{cases}$$

Therefore when $A > 0$ and $\text{Dis} > 0$, this leads to a two-sided confidence interval $[x_{01}, x_{02}]$. When $A < 0$ and $\text{Dis} > 0$, the exact confidence interval goes to $(-\infty, x_{02}] \bigcup [x_{01}, +\infty)$. The confidence set for a mixed linear model will be given in Chapter 4.

Chapter 3

A TEST AND CONFIDENCE SET FOR THE LOCATION OF A QUADRATIC GROWTH CURVE

3.1 Two Quadratic Growth Curve Models

In this dissertation, growth curves with random parameters are studied. Since the responses for each individual are measured repeatedly over time, the models may be polynomial growth curves. Two specific quadratic models for the growth curves from model (2.1) are explored, one is a mixed model with second-order polynomial and random intercept, named the random intercept model; the other is a mixed model with second-order polynomial and both random intercept and random slope, named the random slope model. They are defined as follows:

Second-order mixed model with random intercept (random intercept model),

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + \alpha_{0i} + \epsilon_{ij} \qquad i = 1, ..., N \quad j = 1, ..., n_i \tag{3.1}$$

where,

N is the number of individuals, n_i is the number of occasions for the i^{th} individual,

β_0, β_1 and β_2 are fixed regression coefficients, assuming $\beta_2 \neq 0$

α_{0i} is random effect of the i^{th} individual, $\alpha_{0i} \sim N(0, \sigma^2_{\alpha_0})$,

ϵ_{ij} is the random error term of the i^{th} individual at the j^{th} occasion, $\epsilon_{ij} \sim N(0, \sigma^2_e)$, α_{0i} and ϵ_{ij} are independent, i.e. $\text{Cov}(\alpha_{0i}, \epsilon_{ij}) = 0$ for all i,

y_{ij} is the response at j^{th} occasion of i^{th} individual, and t_{ij} is a time measurement.

In matrix notation, model (3.1) can be written as,

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\alpha}_i + \boldsymbol{\epsilon}_i$$

where,

$\boldsymbol{X}_i$ is the model matrix of regressors for individual i, and $\boldsymbol{X}_i = \begin{pmatrix} 1 & t_{i1} & t_{i1}^2 \\ 1 & t_{i2} & t_{i2}^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{i,n_i} & t_{i,n_i}^2 \end{pmatrix}$,

$\boldsymbol{Z}_i$ is matrix known model matrix, and $\boldsymbol{Z}_i' = (1, 1, \cdots, 1)$,

$\boldsymbol{\beta}$ is an unknown vector of fixed effects, and $\boldsymbol{\beta'} = (\beta_0, \beta_1, \beta_2)$,

$\boldsymbol{\alpha}_i$ is an unknown vector of random effect, $\boldsymbol{\alpha}_i = \alpha_{0i}$ and $\text{Cov}(\alpha_{0i}) = \boldsymbol{G}_{(1\times 1)} = \sigma_{\alpha_0}^2$, $0 < \sigma_{\alpha_0}^2 < \infty$,

$\boldsymbol{\epsilon}_i$ is an unknown vector of random errors for individual i with mean $E(\boldsymbol{\epsilon}_i) = \mathbf{0}$ and covariance $\text{Cov}(\boldsymbol{\epsilon}_i) = \boldsymbol{R}_i$, and $\boldsymbol{R}_{i(n_i\times n_i)} = \sigma_e^2\boldsymbol{I}_{(n_i\times n_i)}$, $0 < \sigma_e^2 < \infty$, $\boldsymbol{\alpha}_i$ and $\boldsymbol{\epsilon}_i$ are independent,

$\boldsymbol{y}_i$ is a known vector of observations for individual i, with mean $E(\boldsymbol{y}_i) = \boldsymbol{X}_i\boldsymbol{\beta}$ and covariance $\Sigma_{\boldsymbol{y}_i} = \boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i$, because of the normality derivation of marginal distribution of $\boldsymbol{y}_i$ in Section 2.2.

To derive the covariance structure for the random intercept model (3.1), the variance for each response is,

$$\text{Var}(y_{ij}) = \text{Var}(\boldsymbol{X}_{ij}\boldsymbol{\beta} + \alpha_{0i} + \epsilon_{ij}) = \sigma_{\alpha_0}^2 + \sigma_e^2.$$

Similarly, the marginal covariance and correlation between any pair of responses, $y_{i,j}$ and $y_{i,j'}$, are,

$$\text{Cov}(y_{i,j}, y_{i,j'}) = \text{Cov}(\boldsymbol{X}_{ij}\boldsymbol{\beta} + \alpha_{0i} + \epsilon_{ij},\ \boldsymbol{X}_{ij'}\boldsymbol{\beta} + \alpha_{1i} + \epsilon_{ij'}) = \sigma_{\alpha_0}^2,$$

and

$$\rho = Corr(y_{i,j}, y_{i,j'}) = \frac{\sigma_{\alpha_0}^2}{\sigma_{\alpha_0}^2 + \sigma_e^2}.$$

Therefore the marginal covariance matrix of the repeated measurements has the following compound symmetry pattern,

$$\Sigma_{\boldsymbol{y_i}} = \begin{pmatrix} \sigma^2_{\alpha_0}+\sigma^2_e & \sigma^2_{\alpha_0} & \sigma^2_{\alpha_0} & \cdots & \sigma^2_{\alpha_0} \\ \sigma^2_{\alpha_0} & \sigma^2_{\alpha_0}+\sigma^2_e & \sigma^2_{\alpha_0} & \cdots & \sigma^2_{\alpha_0} \\ \sigma^2_{\alpha_0} & \sigma^2_{\alpha_0} & \sigma^2_{\alpha_0}+\sigma^2_e & \cdots & \sigma^2_{\alpha_0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma^2_{\alpha_0} & \sigma^2_{\alpha_0} & \sigma^2_{\alpha_0} & \cdots & \sigma^2_{\alpha_0}+\sigma^2_e \end{pmatrix} = \sigma^2_e \boldsymbol{I} + \sigma^2_{\alpha_0}\boldsymbol{J}.$$

Second-order mixed model with random intercept and random slope (random slope model),

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + \alpha_{0i} + \alpha_{1i} t_{ij} + \epsilon_{ij} \qquad i = 1, ..., N \quad j = 1, ..., n_i \tag{3.2}$$

where,

α_{0i} and α_{1i} are random effects of individual i, $\alpha_{0i} \sim N(0, \sigma^2_{\alpha_0})$, $\alpha_{1i} \sim N(0, \sigma^2_{\alpha_1})$ and $\mathrm{Cov}(\alpha_{0i}, \alpha_{1i}) = \sigma_{\alpha_0\alpha_1}$,

ϵ_{ij}, β_0, β_1, β_2, n_i, N, y_{ij} and t_{ij} are defined the same as in model (3.1),

α_{0i}, and α_{1i} are independent of ϵ_{ij}, i.e. $\mathrm{Cov}(\alpha_{0i}, \epsilon_{ij}) = 0$ and $\mathrm{Cov}(\alpha_{1i}, \epsilon_{ij}) = 0$ for all i.

In matrix notation, model (3.2) can be written as,

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\alpha}_i + \boldsymbol{\epsilon}_i$$

where,

$\boldsymbol{X}_i$ is model matrix of regressors for individual i, and $\boldsymbol{X}_i = \begin{pmatrix} 1 & t_{i1} & t_{i1}^2 \\ 1 & t_{i2} & t_{i2}^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{i,n_i} & t_{i,n_i}^2 \end{pmatrix}$,

$\boldsymbol{Z}_i$ is matrix known model matrix, and $\boldsymbol{Z}_i' = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ t_{i1} & t_{i2} & \cdots & t_{i,n_i} \end{pmatrix}$,

$\boldsymbol{\beta}$ is an unknown vector of fixed effects, and $\boldsymbol{\beta}' = (\beta_0, \beta_1, \beta_2)$,

$\boldsymbol{\alpha}_i$ is an unknown vector of random effects, $\boldsymbol{\alpha}_i' = (\alpha_{0i}, \alpha_{1i})$, and $\text{Cov}(\boldsymbol{\alpha}_i) = \boldsymbol{G}_{(2\times 2)} = \begin{pmatrix} \sigma^2_{\alpha_0} & \sigma_{\alpha_0\alpha_1} \\ \sigma_{\alpha_0\alpha_1} & \sigma^2_{\alpha_1} \end{pmatrix}$,

$\boldsymbol{\epsilon}_i$ is an unknown vector of random errors for individual i with mean $E(\boldsymbol{\epsilon}_i) = \mathbf{0}$ and covariance $\text{Cov}(\boldsymbol{\epsilon}_i) = \boldsymbol{R}_i$, and $\boldsymbol{R}_{i(n_i\times n_i)} = \sigma_e^2 \boldsymbol{I}_{(n_i\times n_i)}$, $\boldsymbol{\alpha}_i$ and $\boldsymbol{\epsilon}_i$ are independent,

$\boldsymbol{y}_i$ is a known vector of observations for individual i, with mean $E(\boldsymbol{y}_i) = \boldsymbol{X}_i\boldsymbol{\beta}$ and covariance $\Sigma_{\boldsymbol{y}_i} = \boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i$.

To derive the covariance structure for the random slope model (3.2), the variance of each response is

$$\text{Var}(y_{ij}) = \text{Var}(\boldsymbol{X}_{ij}\boldsymbol{\beta} + \boldsymbol{Z}_{ij}\boldsymbol{\alpha_i} + \epsilon_{ij}) = g_{11} + 2t_{ij}g_{12} + t_{ij}^2 g_{22} + \sigma_e^2 \ , \tag{3.3}$$

where $\text{Var}(\epsilon_{ij}) = \sigma_e^2$, $g_{11} = \sigma^2_{\alpha_0}$, $g_{22} = \sigma^2_{\alpha_1}$ and $g_{12} = \sigma_{\alpha_0\alpha_1}$; g_{11} and g_{22} are the diagonal elements of $\boldsymbol{G}$, and g_{12} is the off diagonal element of $\boldsymbol{G}$. Similarly, the marginal covariance and correlation between any pair of responses, $y_{i,j}$ and $y_{i,k}$, are

$$\text{Cov}(y_{i,j}, y_{i,k}) = \text{Cov}(\boldsymbol{X}_{ij}\boldsymbol{\beta}+\boldsymbol{Z}_{ij}\boldsymbol{\alpha_i}+\epsilon_{ij},\ \boldsymbol{X}_{ik}\boldsymbol{\beta}+\boldsymbol{Z}_{ik}\boldsymbol{\alpha_i}+\epsilon_{ik}) = g_{11}+(t_{ij}+t_{ik})g_{12}+t_{ij}t_{ik}g_{22}, \tag{3.4}$$

and

$$\rho = Corr(Y_{i,j}, Y_{i,k}) = \frac{g_{11} + (t_{ij} + t_{ik})g_{12} + t_{ij}t_{ik}g_{22}}{\sqrt{g_{11} + 2t_{ij}g_{12} + t_{ij}^2 g_{22} + \sigma^2}\sqrt{g_{11} + 2t_{ik}g_{12} + t_{ik}^2 g_{22} + \sigma^2}}$$

which is close to the unstructured covariance pattern.

For the random intercept model (3.1) and the random slope model (3.2), denote $\boldsymbol{b}' = (b_0, b_1, b_2)$ as the maximum likelihood estimator (MLE) , defined in equation (2.2), of fixed regression coefficients $\boldsymbol{\beta}' = (\beta_0, \beta_1, \beta_2)$. As proved in Section 2.2, under some situations such as all the covariance parameters of random effects are known, the distribution of $\boldsymbol{b}$ is exactly normal. More generally, such as the covariance parameters of random effects are unknown, $\boldsymbol{b}$ is approximately normally distributed in large samples with mean $\boldsymbol{\beta}$ and covariance $\Sigma_{\boldsymbol{b}}$, defined in equation (2.3),

$$\Sigma_{\boldsymbol{b}} = \begin{pmatrix} \sigma_{b_0}^2 & \sigma_{b_0b_1} & \sigma_{b_0b_2} \\ \sigma_{b_0b_1} & \sigma_{b_1}^2 & \sigma_{b_1b_2} \\ \sigma_{b_0b_2} & \sigma_{b_1b_2} & \sigma_{b_2}^2 \end{pmatrix} = \left(\sum_i \boldsymbol{X}_i' \Sigma_{\boldsymbol{y}_i}^{-1} \boldsymbol{X}_i \right)^{-1} , \tag{3.5}$$

The corresponding estimated covariance of $\Sigma_{\boldsymbol{b}}$ is,

$$\hat{\Sigma}_{\boldsymbol{b}} = \begin{pmatrix} \hat{\sigma}_{b_0}^2 & \hat{\sigma}_{b_0b_1} & \hat{\sigma}_{b_0b_2} \\ \hat{\sigma}_{b_0b_1} & \hat{\sigma}_{b_1}^2 & \hat{\sigma}_{b_1b_2} \\ \hat{\sigma}_{b_0b_2} & \hat{\sigma}_{b_1b_2} & \hat{\sigma}_{b_2}^2 \end{pmatrix} = \left(\sum_i \boldsymbol{X}_i' \hat{\Sigma}_{\boldsymbol{y}_i}^{-1} \boldsymbol{X}_i \right)^{-1} . \tag{3.6}$$

Denote $\Omega_{\boldsymbol{b}} = \frac{1}{N} \left(\sum_i \boldsymbol{X}_i' \Sigma_{\boldsymbol{y}_i}^{-1} \boldsymbol{X}_i \right)^{-1}$, then

$$\sqrt{N}(\boldsymbol{b} - \boldsymbol{\beta}) \xrightarrow{L} N_3(\boldsymbol{0}, \Omega_{\boldsymbol{b}}). \tag{3.7}$$

3.2 Methods for Confidence Intervals and Region

Let $\boldsymbol{V'} = (V_x, V_y)$ be the vertex of a quadratic growth curve; the vertex can be expressed as a non-linear function of $\boldsymbol{\beta}$, as shown in formula (2.4),

$$V_x(\beta_1, \beta_2) = -\frac{1}{2}\beta_1\beta_2^{-1} , \qquad V_y(\beta_0, \beta_1, \beta_2) = \beta_0 - \frac{1}{4}\beta_1^2\beta_2^{-1} . \tag{3.8}$$

Denote $\hat{\boldsymbol{V}}' = (\hat{V}_x, \hat{V}_y)$ as an estimate of the vertex $\boldsymbol{V}$, $\hat{\boldsymbol{V}}' = (\hat{V}_x, \hat{V}_y)$ can be expressed using the estimator of $\boldsymbol{\beta}$,

$$\hat{V}_x(b_1, b_2) = -\frac{1}{2}b_1b_2^{-1} , \qquad \hat{V}_y(b_0, b_1, b_2) = b_0 - \frac{1}{4}b_1^2b_2^{-1} .$$

In order to obtain the confidence set of the vertex through the fixed regression coefficients $\boldsymbol{\beta}$'s, the first-order partial derivative of $\boldsymbol{V}$ with respect to $\boldsymbol{\beta}$ is required. For the vertex $\boldsymbol{V}$,

$$\frac{\partial \boldsymbol{V}}{\partial \boldsymbol{\beta}} = \boldsymbol{D} = \begin{pmatrix} \frac{\partial V_x}{\partial \beta_0} & \frac{\partial V_x}{\partial \beta_1} & \frac{\partial V_x}{\partial \beta_2} \\ \frac{\partial V_y}{\partial \beta_0} & \frac{\partial V_y}{\partial \beta_1} & \frac{\partial V_y}{\partial \beta_2} \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{2}\beta_2^{-1} & \frac{1}{2}\beta_1\beta_2^{-2} \\ 1 & -\frac{1}{2}\beta_1\beta_2^{-1} & \frac{1}{4}\beta_1^2\beta_2^{-2} \end{pmatrix} . \tag{3.9}$$

Similarly, for the estimated vertex $\hat{\boldsymbol{V}}$, the first-order partial derivative evaluated at $\boldsymbol{\beta} = \boldsymbol{b}$ is,

$$\left.\frac{\partial \boldsymbol{V}}{\partial \boldsymbol{\beta}}\right|_{\boldsymbol{\beta}=\boldsymbol{b}} = \hat{\boldsymbol{D}} = \begin{pmatrix} 0 & -\frac{1}{2}b_2^{-1} & \frac{1}{2}b_1 b_2^{-2} \\ 1 & -\frac{1}{2}b_1 b_2^{-1} & \frac{1}{4}b_1^2 b_2^{-2} \end{pmatrix} .$$

Methods for confidence set of the vertex through $\boldsymbol{\beta}$ will be presented in Section 3.2.1.

3.2.1 Delta Method for Confidence Intervals of Coordinates X and Y

Statistical software routinely compute estimates and inference of the fixed regression coefficients $\boldsymbol{\beta}$, but not usually for non-linear functions of $\boldsymbol{\beta}$. The vertex of the quadratic growth curve given equation (3.8) is obviously a non-linear function of $\boldsymbol{\beta}$. The mean and covariance matrix of the estimated vertex can be obtained through $\boldsymbol{b}$ by large sample theory. The multivariate delta method (2.6) applies by satisfying all the conditions. As sample size N tends to infinity, $\hat{\boldsymbol{V}}(\boldsymbol{b})$ converges to $\boldsymbol{V}$ in probability and the asymptotic covariance is,

$$\mathrm{ACov}(\hat{\boldsymbol{V}}) = \Sigma_{\hat{\boldsymbol{V}}} = \boldsymbol{D}\Sigma_{\boldsymbol{b}}\boldsymbol{D}' = \begin{pmatrix} \sigma^2_{\hat{V}_x} & \sigma_{\hat{V}_x\hat{V}_y} \\ \sigma_{\hat{V}_x\hat{V}_y} & \sigma^2_{\hat{V}_y} \end{pmatrix} ,$$

where $\boldsymbol{D}$ and $\Sigma_{\boldsymbol{b}}$ come from equations (3.9) and (3.5). The estimated asymptotic covariance of estimated vertex, $\mathrm{A\hat{C}ov}(\hat{\boldsymbol{V}})$, is obtained at $\boldsymbol{\beta} = \boldsymbol{b}$,

$$\mathrm{A\hat{C}ov}(\hat{\boldsymbol{V}}) = \hat{\Sigma}_{\hat{\boldsymbol{V}}} = \hat{\boldsymbol{D}}\hat{\Sigma}_{\boldsymbol{b}}\hat{\boldsymbol{D}}' = \begin{pmatrix} \hat{\sigma}^2_{\hat{V}_x} & \hat{\sigma}_{\hat{V}_x\hat{V}_y} \\ \hat{\sigma}_{\hat{V}_x\hat{V}_y} & \hat{\sigma}^2_{\hat{V}_y} \end{pmatrix} , \qquad (3.10)$$

where,

$$\begin{aligned} \hat{\sigma}^2_{\hat{V}_x} &= (0,\ -\frac{1}{2}b_2^{-1},\ \frac{1}{2}b_1 b_2^{-2}) \cdot \hat{\Sigma}_{\boldsymbol{b}} \cdot (0,\ -\frac{1}{2}b_2^{-1},\ \frac{1}{2}b_1 b_2^{-2})' , \\ \hat{\sigma}^2_{\hat{V}_y} &= (1,\ -\frac{1}{2}b_1 b_2^{-1},\ \frac{1}{4}b_1^2 b_2^{-2}) \cdot \hat{\Sigma}_{\boldsymbol{b}} \cdot (1,\ -\frac{1}{2}b_1 b_2^{-1},\ \frac{1}{4}b_1^2 b_2^{-2})' . \end{aligned} \qquad (3.11)$$

Since the estimated regression coefficients $\boldsymbol{b}$ are approximately normally distributed (3.7) and the estimated vertex $\hat{\boldsymbol{V}}$ is a function of $\boldsymbol{b}$, by the delta method, $\hat{\boldsymbol{V}}$ is approximately multivariate normal with mean $\boldsymbol{V}(\boldsymbol{\beta})$ and covariance $\Sigma_{\hat{\boldsymbol{V}}}$ when the sample size N is large. Define $\Omega_{\hat{\boldsymbol{V}}} = \boldsymbol{D}\Omega_{\boldsymbol{b}}\boldsymbol{D}'$, then $\Sigma_{\hat{\boldsymbol{V}}} = N\Omega_{\hat{\boldsymbol{V}}}$ and

$$\sqrt{N}(\hat{\boldsymbol{V}} - \boldsymbol{V}) \xrightarrow{L} N_2(\boldsymbol{0}, \Omega_{\hat{\boldsymbol{V}}}). \tag{3.12}$$

Based on linear model theory, each variable of a multivariate normal vector is normal, therefore $\hat{V}_x$ is approximately normally distributed with mean V_x and variance $\sigma^2_{\hat{V}_x}$, i.e. $\hat{V}_x \overset{a}{\sim} N(V_x, \sigma^2_{\hat{V}_x})$, where symbol $\overset{a}{\sim}$ is defined as asyptotically distributed. Similarly, $\hat{V}_y$ is approximately normally distributed with mean V_y and variance $\sigma^2_{\hat{V}_y}$, i.e. $\hat{V}_y \overset{a}{\sim} N(V_y, \sigma^2_{\hat{V}_y})$. Hence, the approximate $(1-\alpha)\%$ confidence interval of $\hat{V}_x$ is,

$$(\hat{V}_x - Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_x}, \hat{V}_x + Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_x}).$$

The approximate $(1-\alpha)\%$ confidence interval of $\hat{V}_y$ is,

$$(\hat{V}_y - Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_y}, \hat{V}_y + Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_y}).$$

where α is the type I error rate, $\hat{\sigma}_{\hat{V}_x}$ and $\hat{\sigma}_{\hat{V}_y}$ are given in equation (3.11).

3.2.2 Gradient Method for Confidence Interval of Coordinate X

Bachmaier (2009) presented the confidence set for the x-coordinate where a quadratic regression model has a given gradient, as reviewed in Section 2.7; Inequality (2.11) can be used to compute a confidence interval for x-coordinate with a given gradient. For the vertex of a quadratic growth curve, the specific gradient equals zero, i.e. $m = 0$. If the covariance parameters are unknown, we have shown that the estimated vertex, $\hat{\boldsymbol{V}}$, is approximate normal as the sample size tends to infinity. Hence, to form an approximate asymptotic confidence interval of the x-coordinate of the vertex of a quadratic growth curve, the normal approximation is applied. The adjusted method is,

$$x_0 \in C(V_x)$$

$$
\begin{aligned}
&\Leftrightarrow \frac{(b_1+2x_0b_2)^2}{\hat{\sigma}^2_{b_1}+4x_0\hat{\sigma}_{b_1b_2}+4x_0^2\hat{\sigma}^2_{b_2}} \leqslant Z^2_{1-\alpha/2} \\
&\Leftrightarrow (b_1+2x_0b_2)^2 \leqslant [\hat{\sigma}^2_{b_1}+4x_0\hat{\sigma}_{b_1b_2}+4x_0^2\hat{\sigma}^2_{b_2}]\cdot Z^2_{1-\alpha/2} \\
&\Leftrightarrow A\cdot x_0^2+B\cdot x_0^2+C \leqslant 0\ .
\end{aligned}
\tag{3.13}
$$

$$
\begin{aligned}
where,\ A &= b_2^2-\hat{\sigma}^2_{b_2}\cdot Z^2_{1-\alpha/2} \\
B &= b_1b_2-\hat{\sigma}_{b_1b_2}\cdot Z^2_{1-\alpha/2} \\
C &= \frac{1}{4}(b_1^2-\hat{\sigma}^2_{b_1}\cdot Z^2_{1-\alpha/2})\ .
\end{aligned}
$$

To solve the inequality, if $A \neq 0$, then $A\cdot x_0^2+B\cdot x_0^2+C$ in inequality (3.13) is a parabola. It has two roots if the discriminant $D = B^2-4AC$ is positive. With regard to the numerical stability concerning small values of $4AC$, root is computed in either two different ways:

$$
x_{01} = \begin{cases} \frac{-2C}{B-\sqrt{B^2-4AC}} & \text{when } B<0 \\ \frac{-B-\sqrt{B^2-4AC}}{2A} & \text{when } B \geqslant 0 \end{cases}, \qquad x_{02} = \begin{cases} \frac{-B+\sqrt{B^2-4AC}}{2A} & \text{when } B \leqslant 0 \\ \frac{-2C}{B+\sqrt{B^2-4AC}} & \text{when } B>0 \end{cases}
\tag{3.14}
$$

Hence when $A>0$ and $D>0$, this leads to a two-sided confidence interval $[x_{01}, x_{02}]$. When $A<0$ and $D>0$, the confidence interval goes to $(-\infty, x_{02}]\bigcup[x_{01}, +\infty)$. In this dissertation, only the first situation is considered. An approximate $(1-\alpha)\%$ confidence interval for coordinate x of the estimated vertex, $\hat{V}_x$, is $[x_{01}, x_{02}]$ given in equation (3.14).

3.2.3 Mean Response Method for Confidence Interval of Coordinate Y

If the $x-$coordinate of vertex V_x is given and substituted into the regression model

$$
\hat{y_{ij}} = b_0 + b_1x_{ij} + b_2x_{ij}^2,
$$

Then $V_y = \boldsymbol{C'b}$, where $\boldsymbol{C} = (1,\ V_x,\ V_x^2)$, where $\hat{V}_y$ is treated as a mean response of y at $x = V_x$. The variance of $y-$coordinate of the vertex is $\sigma^2_{\hat{V}_y} = \boldsymbol{C'}\Sigma_{\boldsymbol{b}}\boldsymbol{C}$, and the estimated variance is $\hat{\sigma}^2_{\hat{V}_y} = \boldsymbol{C}'\hat{\Sigma}_{\boldsymbol{b}}\boldsymbol{C}$. Then,

$$\frac{\hat{V}_y - V_y}{\hat{\sigma}^2_{\hat{V}_y}} \sim N(0,1) \ .$$

Therefore the $(1-\alpha)\%$ confidence interval for the $y-$coordinate of vertex is

$$(\hat{V}_y - Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_y},\ \hat{V}_y + Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_y}).$$

If the $x-$coordinate of the vertex $\hat{V}_x$ is estimated, the $y-$coordinate of vertex $\hat{V}_y$ can be calculated $\hat{V}_y = b_0 + b_1 \cdot \hat{V}_x + b_2 \cdot \hat{V}_x^{\ 2}$. Using the equations $\hat{V}_x = -\frac{1}{2}b_1 b_2^{-1}$ and $\hat{V}_x^2 = \frac{1}{4}b_1^2 b_2^{-2}$,

$$\hat{\sigma}^2_{\hat{V}_y} = (1,\ -\frac{1}{2}b_1 b_2^{-1},\ \frac{1}{4}b_1^2 b_2^{-2}) \cdot \hat{\Sigma}_{\boldsymbol{b}} \cdot (1,\ -\frac{1}{2}b_1 b_2^{-1},\ \frac{1}{4}b_1^2 b_2^{-2})' \ ,$$

it is same as the estimated variance of $\hat{V}_y$ from the delta method. Therefore, in this condition, the mean response method becomes the delta method.

3.2.4 Confidence Region for Vertex

As defined in equation (3.8), the vertex of a quadratic growth curve is a two-dimensional vector; the two elements are the x-coordinate and the y-coordinate and they are related. In order to find the confidence region for the vertex, the large sample distribution of a quadratic form can be applied. Consider the chi-square distribution with k degrees of freedom, defined as the distribution of a sum of the squares of k independent standard normal random variables. It was proven in (3.12) that the $\hat{\boldsymbol{V}}$ has an approximate bivariate normal distribution, hence $\left(\hat{\boldsymbol{V}} - \boldsymbol{V}\right)\Sigma_{\hat{\boldsymbol{V}}}^{-1/2} \overset{a}{\sim} N_2(\boldsymbol{0}, \boldsymbol{I}_{(2\times 2)})$. Let $\boldsymbol{z} = \left(\hat{\boldsymbol{V}} - \boldsymbol{V}\right)\Sigma_{\hat{\boldsymbol{V}}}^{-1/2}$, by definition $\boldsymbol{z}'\boldsymbol{z}$ is $\chi^2_{(2)}$ (Rencher and Nchaalh, 2007). In the quadratic form notation,

$$\begin{pmatrix} \hat{V}_x - V_x \\ \hat{V}_y - V_y \end{pmatrix}' \Sigma_{\hat{\boldsymbol{V}}}^{-1} \begin{pmatrix} \hat{V}_x - V_x \\ \hat{V}_y - V_y \end{pmatrix} \sim \chi^2_{(2)} \ .$$

Because $\hat{\Sigma}_{\hat{\boldsymbol{V}}}$ is a consistent statistic for $\Sigma_{\hat{\boldsymbol{V}}}$, by large sample theory an approximate chi-square distribution with 2 degrees of freedom is obtained,

$$\begin{pmatrix} \hat{V}_x - V_x \\ \hat{V}_y - V_y \end{pmatrix}' \hat{\Sigma}_{\hat{\boldsymbol{V}}}^{-1} \begin{pmatrix} \hat{V}_x - V_x \\ \hat{V}_y - V_y \end{pmatrix} \overset{a}{\sim} \chi^2_{(2)} \ .$$

Therefore the approximate $(1-\alpha)\%$ confidence region of the vertex is

$$\begin{pmatrix} \hat{V}_x - V_x \\ \hat{V}_y - V_y \end{pmatrix}' \hat{\Sigma}_{\hat{\boldsymbol{V}}}^{-1} \begin{pmatrix} \hat{V}_x - V_x \\ \hat{V}_y - V_y \end{pmatrix} \leqslant \chi^2_{1-\alpha,2} \ , \tag{3.15}$$

where α is the type I error rate and $\chi^2_{1-\alpha,2}$ is the critical value. The confidence region for the vertex is the area covered by an ellipse, since (3.15) with equality is an elliptic equation.

3.3 Power Analysis

Power analysis plays an important role to reject the null hypothesis if it specifies a vertex point that is actually not the true vertex point for quadratic growth curve. Consider the hypotheses,

$$H_0 : \boldsymbol{V} = \boldsymbol{V}_0 \quad \text{v.s.} \quad H_a : \boldsymbol{V} = \boldsymbol{V}_a \tag{3.16}$$

where $\boldsymbol{V}_a$ is the true vertex and $\boldsymbol{V}_0$ is the hypothesized vertex point under the null hypothesis. The power function of a statistical test is the probability that the test statistic falls in the rejection region R (Kenward and Roger, 1997). The approximation (3.15) can be used to obtain a direct method to test the hypothesis (3.16). The power function of the direct chi-square test will be presented in Section 3.3.3.

An indirect method to test the hypotheses (3.16) would use an F statistic with respect to $\boldsymbol{\beta}$'s, since the x and y-coordinates of the vertex (3.8) are nonlinear functions of $\boldsymbol{\beta}$'s. Transform the hypotheses (3.16) to the hypotheses with regard to $\boldsymbol{\beta}$'s; the new hypotheses are stated as follows,

$$H_0 : \begin{pmatrix} V_x \\ V_y \end{pmatrix} = \begin{pmatrix} -\frac{1}{2}\beta_{0,1}\beta_{0,2}^{-1} \\ \beta_0 - \frac{1}{4}\beta_{0,1}^2\beta_{0,2}^{-1} \end{pmatrix} \quad vs \quad \begin{pmatrix} V_x \\ V_y \end{pmatrix} \neq \begin{pmatrix} -\frac{1}{2}\beta_{0,1}\beta_{0,2}^{-1} \\ \beta_0 - \frac{1}{4}\beta_{0,1}^2\beta_{0,2}^{-1} \end{pmatrix}$$

where V_x and V_y are the coordinates of $\boldsymbol{V}$. Alternatively, the null hypothesis may be simply stated as,

$$H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0, \tag{3.17}$$

where $\boldsymbol{\beta}_0' = (\beta_{0,0}, \beta_{0,1}, \beta_{0,2})$ and $V_{0x} = -\frac{1}{2}\beta_{0,1}\beta_{0,2}^{-1}$ and $V_{0y} = \beta_{0,0} - \frac{1}{4}\beta_{0,1}^2\beta_{0,2}^{-1}$. Power functions of the indirect F test will be presented in Section 3.3.1 and Section 3.3.2 for the random intercept model (3.1) and the random slope model (3.2).

The two null hypotheses (3.16) and (3.17) are not necessarily equivalent. For the $x-$coordinate of the vertex, $V_x = -\frac{1}{2}\beta_1\beta_2^{-1}$, if β_2 is shifted by amount Δ, V_x can remain unchanged by changing β_1 with certain amount Δ , i.e. the change of β_2 can be offset by the change of β_1. Similarly, for the $y-$coordinate of the vertex, $V_y = \beta_0 - \frac{1}{4}\beta_1^2\beta_2^{-1}$, if the ratio $\beta_1^2\beta_2^{-1}$ is shifted amount Δ, V_y can remain the same by shifting the same amount Δ for β_0, i.e. the change of ratio $\beta_1^2\beta_2^{-1}$ can be offset by the change of β_0. The explanation is also shown in Figure 3.1; different quadratic functions with different coefficients $\boldsymbol{\beta}$'s share the same vertex $\boldsymbol{V}$. In conclusion, "do not reject $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$" implies "do not reject $H_0 : \boldsymbol{V} = \boldsymbol{V}_0$", while "reject $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$" does not necessarily imply "reject $H_0 : \boldsymbol{V} = \boldsymbol{V}_0$".

3.3.1 Power Function of F Test for Random Intercept Model

To derive the power function for testing the hypothesis (3.17) with respect to $\boldsymbol{\beta}$ for the random intercept model (3.1), a randomized block design with random block can be presented since it is applicable to model the longitudinal data. Repeated measurements on a single sample from a population can be represented by a randomized

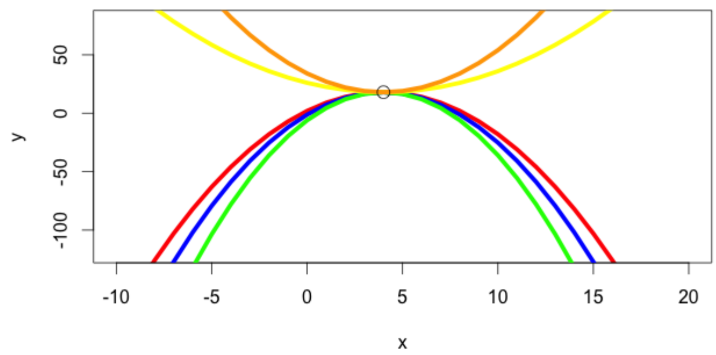


Figure 3.1: Different Quadratic Functions with Same Vertex

block model,

$$y_{ij} = \mu_{..} + \alpha_{0i} + \tau_j + \epsilon_{ij} \tag{3.18}$$

where,

y_{ij} is the response at jth occasion for ith subject with $E(y_{ij}) = \mu_{..} + \tau_j$,

$\mu_{..}$ is a constant for grand mean of all the observations,

α_{0i} is the random effect, and $\alpha 0_i$ are independent $N(0, \sigma^2_{\alpha_0})$,

τ_j is the fixed effect, and τ_j's are constants subject to the restriction $\Sigma\tau_j = 0$,

ϵ_{ij} are independent $N(0, \sigma^2_e)$, and independent of the α_{0i},

$i = 1, 2, ..., N$; $j = 1, 2, ...n_i$. N is sample size, and n_i is number of occasions assuming to be same for all the subjects as n.

Testing hypothesis (3.17) for random intercept model (3.1) is equivalent to testing a potential quadratic trend for the randomized block model (3.18). The null hypothesis of no potential trend for model (3.18) can be stated as $H_0 : \boldsymbol{\tau} = \mathbf{0}$. Under the assumption of the compound symmetry covariance structure, $\Sigma_{\boldsymbol{y}_i} = \sigma^2_e \cdot \boldsymbol{I}_{n\times n} + \sigma^2_{\alpha_0} \cdot \boldsymbol{J}_{n\times n}$, the test statistic for $H_0 : \boldsymbol{\tau} = 0$ is an F statistic based on sum of squares error and sum of squares treatment (occasion), where

$$SS(Occasion) = N \cdot \sum_j (\bar{y}_{.j} - \bar{y}_{..})^2 \qquad SS(Error) = \sum_i \sum_j (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 \ .$$

The F statistic is exact and uniformly most powerful (UMP); a UMP test is a hypothesis test which has the greatest power among all possible tests of a given Type I error rate α (Casella and Berger, 2002). Sum of squares occasion can be partitioned into sum of squares for polynomial trend using Gram-Schmidt orthonormalization or the Cholesky factorization of $\boldsymbol{X}_i'\boldsymbol{X}_i$, where $\boldsymbol{X}_i$ is the model matrix for subject i. The Cholesky factor produces orthonormalization, however it is less numerically stable. A specific example of Cholesky factorization is illustrated to test quadratic trend for the randomized block model. Consider the number of occasion, $n = 3$, then the design matrix $\boldsymbol{X}_i$ for the randomized block model is,

$$\boldsymbol{X}_i = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} .$$

The equivalent 2^{nd} order random intercept model (3.1) is,

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + \alpha_{0i} + \epsilon_{ij}$$

where t_{ij} is the time measurement, β_0, β_1 and β_2 are parameters of regression coefficients, α_{0i} is the random effect and normally distributed, $\alpha_{0i} \sim N(0, \sigma^2_{\alpha_0})$, y_{ij} and ϵ_{ij} are same as denoted in model (3.18).

The design matrix $\boldsymbol{X}_i$ for the random intercept model and matrix $\boldsymbol{K} = \boldsymbol{X}_i'\boldsymbol{X}_i$ are,

$$\boldsymbol{X}_i = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}, \qquad \boldsymbol{K} = \boldsymbol{X}_i'\boldsymbol{X}_i = \begin{pmatrix} 3 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix} .$$

The Cholesky factor $\boldsymbol{S}$ of matrix $\boldsymbol{K}$ based on the regression model is lower triangular with positive diagonal elements such that $\boldsymbol{K} = \boldsymbol{S}\boldsymbol{S}'$. In this example,

$$\boldsymbol{S} = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 6/\sqrt{3} & \sqrt{2} & 0 \\ 14/\sqrt{3} & 8/\sqrt{2} & \sqrt{2}/\sqrt{3} \end{pmatrix} .$$

In general, the design matrix $\boldsymbol{X}_i$ for randomized block design and the random intercept model with up to $(q-1)^{\text{th}}$ order polynomial are,

$$\boldsymbol{X}_i = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{pmatrix}, \qquad \boldsymbol{X}_i = \begin{pmatrix} 1 & 1 & 1^2 & \cdots & 1^{q-1} \\ 1 & 2 & 2^2 & \cdots & 2^{q-1} \\ 1 & 3 & 3^2 & \cdots & 3^{q-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & n & n^2 & \cdots & n^{q-1} \end{pmatrix} .$$

The assumption of equally spaced intervals for time measurements has to be satisfied, otherwise, the columns for the design matrix should be the actual values of occasions. The Cholesky factorization algorithm is usable to obtain the Cholesky factor of matrix $\boldsymbol{K} = \boldsymbol{X}_i'\boldsymbol{X}_i$ from the random intercept model (Johnson *et al.*, 1992);

1. Partition matrices $\boldsymbol{K} = \boldsymbol{S}\boldsymbol{S}'$ as,

$$\boldsymbol{H} = \begin{pmatrix} a_{11} & \boldsymbol{K}_{21}' \\ \boldsymbol{K}_{21} & \boldsymbol{K}_{22} \end{pmatrix} = \begin{pmatrix} s_{11} & 0 \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{pmatrix} \begin{pmatrix} s_{11} & \boldsymbol{S}_{21}' \\ 0 & \boldsymbol{S}_{22}' \end{pmatrix}$$
$$= \begin{pmatrix} s_{11}^2 & s_{11}S_{21}' \\ s_{11}S_{21} & S_{21}S_{21}' + S_{22}S_{22}' \end{pmatrix} .$$

2. Determine s_{11} and $\boldsymbol{S}_{21}$:

$$s_{11} = \sqrt{a_{11}}, \qquad \boldsymbol{S}_{21} = \frac{1}{s_{11}}\boldsymbol{K}_{21} .$$

3. Compute $\boldsymbol{S}_{22}$ from

$$\boldsymbol{k}_{22} - \boldsymbol{S}_{21}\boldsymbol{S}_{21}' = \boldsymbol{S}_{22}\boldsymbol{S}_{22}'$$

this is a Cholesky factorization of order $n-1$.

The null hypothesis $H_0 : \boldsymbol{\beta}_{q\times 1} = \boldsymbol{\beta_0}$, testing a potential $(q-1)^{\text{th}}$ order polynomial trend, is a component of the null hypothesis $H_0 : \boldsymbol{\tau} = 0$, testing all polynomial trends; then sum of squares for $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta_0}$ can be obtained from $H_0 : \boldsymbol{L\tau} = 0$ by reparametrization, where $\boldsymbol{L}$ contains coefficients for orthogonal polynomial contrasts. Denote l_m as the mth row for $\boldsymbol{L}$, the sum of squares for each contrast is,

$$SS(\text{Contrast}_k) = \frac{N \cdot (\sum_j l_{mj}\bar{y}_{.j})^2}{\sum_j l_{mj}^2} = \left(\boldsymbol{L}\hat{\boldsymbol{\beta}}\right)' \left(\boldsymbol{L}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{L}'\right)^{-1} \left(\boldsymbol{L}\hat{\boldsymbol{\beta}}\right) .$$

$SS(\text{Contrast}_k)/\sigma^2 \sim \chi^2(n, \lambda)$, where $\lambda = \left(\boldsymbol{L}\hat{\boldsymbol{\beta}}\right)' \left(\boldsymbol{L}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{L}'\right)^{-1} \left(\boldsymbol{L}\hat{\boldsymbol{\beta}}\right)/(2\sigma^2)$ Then the test is based on $F = \frac{MS(\text{Contrast}_k)}{MS(\text{Error})}$; it is an exact test (Khuri *et al.*, 2011).

For the random intercept model, the generalized F statistic for testing $H_0 : \boldsymbol{\beta}_{q\times 1} = \boldsymbol{\beta_0}$ is,

$$F = \frac{(\boldsymbol{b} - \boldsymbol{\beta}_0)' \left(\left(\sum_i \boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1}\right)^{-1} (\boldsymbol{b} - \boldsymbol{\beta}_0)}{q}, \tag{3.19}$$

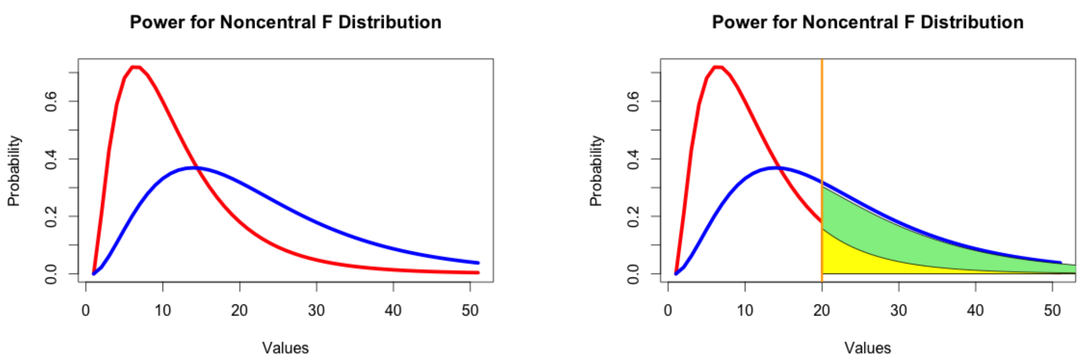
where the numerator degrees of freedom is $\text{ndf}_1 = q$ and the denominator degrees of freedom is $\text{ddf}_1 = N \cdot (n-1) - (q-1)$; it is an approximate test. The non-centrality parameter, λ_1, is

$$\lambda_1 = (\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \left(\left(\sum_i \boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1}\right)^{-1} (\boldsymbol{\beta} - \boldsymbol{\beta}_0).$$

Under $H_0, \lambda_1 = 0$; on the other hand, given H_a is true, $\lambda_1 > 0$. Therefore the power function is

$$\text{Power} \approx \text{Prob}\{F(\text{ndf}_1,\ \text{ddf}_2,\ \lambda_1) > F_{1-\alpha,\ \text{ndf}_1,\ \text{ddf}_1}\}$$

where λ_1 is the value of the non-centrality parameter and $F_{1-\alpha}$ is the critical value of the central F at the designated α level. Comparison of the central F and non-central F distribution is displayed in Figure 3.2. In Figure 3.2(a), the probability density function (pdf) for the central F distribution is the red curve and pdf for the non-central F distribution is the blue curve. In Figure 3.2(b), given the critical value, vertical orange line, the yellow area represents the rejection region under central F distribution and the sum of yellow and green area is power for the test under non-central F distribution.



(a) PDF of Central and Non-Central F Distribution

(b) Power for F Test

Figure 3.2: Power for Non-Central F Distribution

3.3.2 Power Function of F Test for Random Slope Model

For random slope model (3.2), the variance of each response and covariance between any two responses of same subject are given in equation (3.3) and (3.4). The test of $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta_0}$ using a F-type statistic (3.19) is an approximate test in that the denominator degrees of freedom ddf_{1a} is not exact. The power function for the approximated F test is

$$\text{Power} \approx \text{Prob}\{F(\text{ndf}_1,\ \text{ddf}_{1a},\ \lambda_1) > F_{1-\alpha,\ \text{ndf}_1,\ \text{ddf}_{1a}}\}$$

where $F_{1-\alpha}$ is the critical value of the central F distribution with the approximate denominator degrees of freedom. Two main methods for computing denominator degrees of freedom for longitudinal studies, Satterthwaite and Kenward-Roger, are briefly illustrated. As defined, $\boldsymbol{\theta}$ is the vector of unknown parameters in $\Sigma_{\boldsymbol{y}_i}$ which includes all fixed regression coefficients $\boldsymbol{\beta}$ and variance components, and suppose $\boldsymbol{C} = \left(\sum_i \boldsymbol{X}_i' \Sigma_{\boldsymbol{y}_i} \boldsymbol{X}_i\right)^{-}$, where $\boldsymbol{M}^{-}$ denotes a generalized inverse of matrix $\boldsymbol{M}$. Let $\hat{\boldsymbol{C}}$ and $\hat{\boldsymbol{\theta}}$ be the corresponding estimates. For an estimable contrast matrix $\boldsymbol{L}_{h\times p}$ with the rank of $\boldsymbol{L}\hat{\boldsymbol{C}}\boldsymbol{L}'$, $q > 1$. The Satterthwaite denominator degrees of freedom for the F statistic are computed by first performing the spectral decomposition $\boldsymbol{L}\hat{\boldsymbol{C}}\boldsymbol{L}' = \boldsymbol{P}'\boldsymbol{D}\boldsymbol{P}$, where $\boldsymbol{P}$ is an orthogonal matrix of eigenvectors and $\boldsymbol{D}$ is a diagonal matrix of eigenvalues, both of dimension $q \times q$. Define l_m to be the mth row of $\boldsymbol{L}$, and let

$$v_m = \frac{2(D_m)^2}{\boldsymbol{g}_m' \boldsymbol{M} \boldsymbol{g}_m}$$

where D_m is the mth diagonal element of $\boldsymbol{D}$, $\boldsymbol{g}_m$ is the gradient of $l_m \boldsymbol{C} l_m$ with respect to $\boldsymbol{\theta}$, evaluated at $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{M}$ is the asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}$ obtained from the second derivative matrix of the likelihood equations. Then let

$$E = \sum_{m=1}^{q} \frac{v_m}{v_m - 2}\ I(v_m > 2)$$

where the indicator function eliminates terms for which $v_m \leqslant 2$. The degrees of freedom for F are then computed as

$$v = \frac{2E}{E - q}$$

provided $E > q$; otherwise v is set to zero.

Kenward and Roger (1997) derived another approximation for denominator degrees of freedom,

$$v = 4 + \frac{h + 2}{h \times u - 1}$$

where $u = \frac{\mathrm{Var}(F)}{2\mathrm{E}(F)^2}$, $\mathrm{E}(F)$ and $\mathrm{Var}(F)$ are the mean and variance for test statistic $F(\mathrm{ndf}_1,\ \mathrm{ddf}_{1a},\ \lambda_1)$. The observed information matrix of the covariance parameter estimates is used in the calculations. For covariance structures that have nonzero second derivatives with regard to the covariance parameters, the Kenward-Roger covariance matrix adjustment includes a second order term.

3.3.3 Power Function for Chi-Square Test

The non-central chi-square distribution can be applied to compute power for the hypotheses (3.16), $H_0 : \boldsymbol{V} = \boldsymbol{V}_0$ vs $H_a : \boldsymbol{V} = \boldsymbol{V}_a$, since $\boldsymbol{V}$ has an asymptotic multivariate normal distribution as proven in (3.12). Using Theorem 5.5 in textbook Rencher and Schaalje (2008), $\hat{\boldsymbol{V}}'\Sigma_{\hat{\boldsymbol{V}}}^{-1}\hat{\boldsymbol{V}}$ distributes as a non-central chi-square with 2 degrees of freedom with the non-centrality parameter

$$\begin{aligned}
\lambda_2 &= (\boldsymbol{V} - \boldsymbol{V}_0)'\Sigma_{\hat{\boldsymbol{V}}}^{-1}(\boldsymbol{V} - \boldsymbol{V}_0) \\
&= \begin{pmatrix} V_x - V_{0x} \\ V_y - V_{0y} \end{pmatrix}' \Sigma_{\hat{\boldsymbol{V}}}^{-1} \begin{pmatrix} V_x - V_{0x} \\ V_y - V_{0y} \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{2}\beta_1\beta_2^{-1} - V_{0x} \\ \beta_0 - \frac{1}{4}\beta_1^2\beta_2^{-1} - V_{0y} \end{pmatrix}' \Sigma_{\hat{\boldsymbol{V}}}^{-1} \begin{pmatrix} -\frac{1}{2}\beta_1\beta_2^{-1} - V_{0x} \\ \beta_0 - \frac{1}{4}\beta_1^2\beta_2^{-1} - V_{0y} \end{pmatrix} .
\end{aligned}$$

Namely, $\hat{\boldsymbol{V}}'\Sigma_{\hat{\boldsymbol{V}}}^{-1}\hat{\boldsymbol{V}} \sim \chi^2_{2,\lambda_2}$. Under the null hypothesis, the non-centrality parameter $\lambda_2 = 0$, an approximate distribution with 2 degrees of freedom follows:

$$\begin{pmatrix} \hat{V}_x - V_{0x} \\ \hat{V}_y - V_{0y} \end{pmatrix}' \hat{\Sigma}_{\hat{\boldsymbol{V}}}^{-1} \begin{pmatrix} \hat{V}_x - V_{0x} \\ \hat{V}_y - V_{0y} \end{pmatrix} \overset{a}{\sim} \chi^2_{(2),\lambda_2} .$$

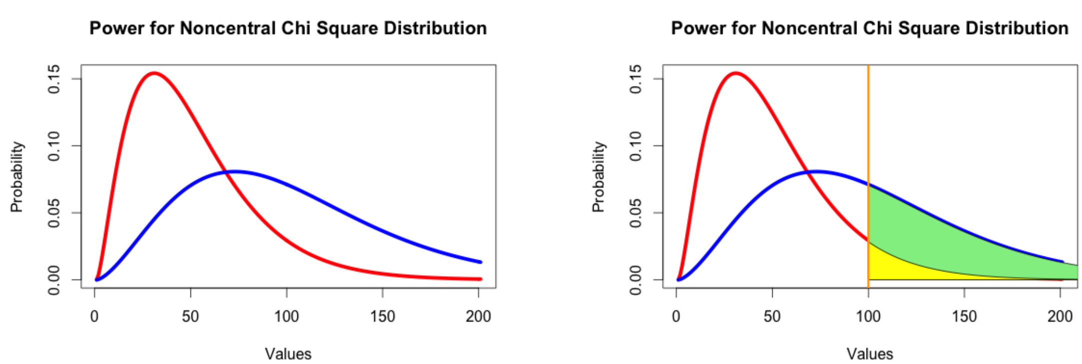
where estimated covariance $\hat{\Sigma}_{\hat{\boldsymbol{V}}}$ is the consistent statistic for $\Sigma_{\hat{\boldsymbol{V}}}$. Therefore the decision rule is reject the null hypothesis if

$$\begin{pmatrix} \hat{V}_x - V_{0x} \\ \hat{V}_y - V_{0y} \end{pmatrix}' \hat{\Sigma}_{\hat{\boldsymbol{V}}}^{-1} \begin{pmatrix} \hat{V}_x - V_{0x} \\ \hat{V}_y - V_{0y} \end{pmatrix} > \chi^2_{1-\alpha,2} , \tag{3.20}$$

otherwise do not reject the null hypothesis, where $\chi^2_{1-\alpha,2}$ is the critical value given test size level α. Comparing the confidence region for the vertex (3.15) and the rejection region in (3.20), both are obtained through the approximate chi-square distribution with 2 degrees of freedom in a quadratic form; the only difference is the reversed inequality sign. The relationship between the confidence region and power analysis will be shown in Section 3.4. The power function for the test is

$$\text{Power} \approx \text{Prob}\left\{\chi^2(2,\ \lambda_2) > \chi^2_{1-\alpha,2}\right\}.$$

Comparison of central chi-square and non-central chi-square distributions is displayed in Figure 3.3. In Figure 3.3 (a), the pdf for the central chi-square distribution is the red curve and the pdf for the non-central chi-square distribution is the blue curve. In Figure 3.3 (b), given the critical value, the vertical orange line, the yellow area represents the rejection region under central chi-square distribution and the sum of yellow and green area is power for the test under the non-central chi-square distribution.



(a) PDF of Central and Non-Central Chi-Square Distribution

(b) Power for Chi-Square Test

Figure 3.3: Power for Non-Central Chi-Square Distribution

3.3.4 Leverage Value for the Vertex

The hat matrix $\boldsymbol{H}$, plays an important role in diagnostics for linear regression analysis. The hat matrix, sometimes also called the influence matrix and projection matrix, maps the vector of observed values to the vector of predicted values. It describes the influence each observed value has on each predicted value, $\partial\hat{\boldsymbol{y}}/\partial\boldsymbol{y}$, where $\boldsymbol{y}$ is the known vector of response of all the observations and $\hat{\boldsymbol{y}}$ is the estimated vector of response. Suppose that a linear model is solved using ordinary least squares (OLS), the estimator for the regression coefficient is $\hat{\boldsymbol{\beta}} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'y}$ and the predicted value is $\hat{\boldsymbol{y}} = \boldsymbol{X}(\boldsymbol{X'X})^{-1}\boldsymbol{X'y}$, where $\boldsymbol{X}$ is the model matrix. Therefore the hat matrix for OLS is $\boldsymbol{H}_{\mathrm{OLS}} = \boldsymbol{X}(\boldsymbol{X'X})^{-1}\boldsymbol{X'}$. The leverage h_{ii}, the diagonal element of the hat matrix $\boldsymbol{H}$, is commonly used to diagnose influential observations for linear regression; it identifies observations whose distance from the center of the data causes them to have a potentially large effect on the fitted values. In OLS regression, the leverage values are always between 0 and 1; the trace of the hat matrix H_{OLS} is q, then the average value of leverage is q/N, where where q is the number of regression coefficients $\boldsymbol{\beta'}s$ and N is sample size. The larger the leverage, the more likely that the observed value is an outlier with respect to x direction (Kutner *et al.*, 2005).

Leverage for subject i, $\partial\hat{\boldsymbol{y}}_{\boldsymbol{i}}/\partial\boldsymbol{y}'_i$, is also an important diagnostic tool for longitudinal studies. Unlike the value of leverage for OLS, the leverage for generalized least square (GLS) is a matrix quantity (Gruttola *et al.*, 1987). As illustrated in equation (2.2) and (2.3) for mixed model, the estimate for the regression coefficient and the predicted value using the MLE method are $\hat{\boldsymbol{\beta}} = \left(\sum\limits_{i=1}^{N} \boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1}\left(\sum\limits_{i=1}^{N} \boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{y}_i\right)$ and $\Sigma_{\hat{\boldsymbol{\beta}}} = \left(\sum\limits_{i=1}^{N} \boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1}$, then the estimated response for subject i is, $\hat{\boldsymbol{y}}_i = \boldsymbol{X}_i\hat{\boldsymbol{\beta}} = \boldsymbol{X}_i\left(\sum\limits_{i=1}^{N} \boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1}\left(\sum\limits_{i=1}^{N} \boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{y}_i\right)$. Therefore the leverage matrix for subject i

is $\boldsymbol{H}_i^* = \partial\hat{\boldsymbol{y}}_{\boldsymbol{i}}/\partial\boldsymbol{y}_i' = \boldsymbol{X}_i\left(\sum_{i=1}^{N}\boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1}\boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}$ given the covariance matrix held fixed. The leverage at j^{th} occasion for i^{th} subject is the partial derivative of the fitted value, $\partial\hat{y}_{ij}/\partial y_{ij}$; it may be positive or negative. If the correlation between observations is high, it is particularly likely to be a negative leverage; otherwise, if the sample correlation is 0, the leverage are always positive.

For the quadratic growth curve, since $V_x(\beta_1,\beta_2) = -\frac{1}{2}\beta_1\beta_2^{-1}$, the $x-$coordinate of the vertex maybe far away from the scope of the studied occassion values if β_2 is small in magnitude compared to β_1. Moreover, $\frac{\partial V_y}{\partial\beta_1} = -\frac{1}{2}\beta_1\beta_2^{-1}$, so both $\sigma^2_{\hat{V}_x}$ and $\sigma^2_{\hat{V}_y}$ may become large if $-\frac{1}{2}\beta_1\beta_2^{-1}$ is large, and the power of the test $H_0 : \boldsymbol{V} = \boldsymbol{V}_0$ will decrease as $V_x = -\frac{1}{2}\beta_1\beta_2^{-1}$ becomes large. It is proposed to use the leverage value of the $x-$coordinate of $\boldsymbol{V}$ as a measure of the distance from the scope of the studied occasion values to the vertex. Power for the chi-square test of $H_0 : \boldsymbol{V} = \boldsymbol{V}_0$ is related to the leverage. The model matrix $\boldsymbol{X}_i$ with $x-$coordinate of the vertex for the i^{th} subject of random intercept model (3.1) and random slope model (3.2) is,

$$\boldsymbol{X}_{i,((n_i+1)\times 3)} = \begin{pmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{n_i} & t_{n_i}^2 \\ 1 & V_x & V_x^2 \end{pmatrix}.$$

Then the leverage $\boldsymbol{H}_i^*$ for subject i is, $\boldsymbol{H}^*_{i,((n_i+1)\times(n_i+1))} = \boldsymbol{X}_i\left(\sum_{i=1}^{N}\boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1}\boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}$; the leverage on the vertex point $\boldsymbol{V}$ is,

$$H^*_{i,(n_i+1,n_i+1)} = (1\ V_x\ V_x^2)\left(\sum_{i=1}^{N}\boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1}(1\ V_x\ V_x^2)'(\Sigma_{\boldsymbol{y}_i}^{-1})_{(n_i+1,n_i+1)} \qquad (3.21)$$

As deduced in Section 3.3.3, the non-centrality parameter for chi-square distribution for computing power is,

$$\lambda_2 = \Delta\boldsymbol{V}'\Sigma_{\hat{\boldsymbol{V}}}^{-1}\Delta\boldsymbol{V} = \Delta\boldsymbol{V}'\left(\boldsymbol{D}\left(\sum_{i=1}^{N}\boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1}\boldsymbol{D}'\right)^{-1}\Delta\boldsymbol{V}, \tag{3.22}$$

where $\boldsymbol{D}$ is the derivative of $\boldsymbol{V}$ with regard to the regression coefficients $\boldsymbol{\beta}$ defined in equation (3.9), and $\Delta\boldsymbol{V} = \boldsymbol{V} - \boldsymbol{V}_0$, the difference between true vertex $\boldsymbol{V}$ and hypothesized vertex $\boldsymbol{V}_0$. Keeping the difference $\Delta\boldsymbol{V}$ fixed, the influence of the leverage on the vertex point (3.21) can be seen in the non-centrality parameter (3.22). When the vertex is farther outside the scope of the occasions, the x-coordinate of the vertex point $V_x = -\frac{1}{2}\beta_1\beta_2^{-1}$, an element in $\boldsymbol{D}$ matrix, becomes large, which results in smaller λ_2 and greater $H^*_{i,(n_i+1,n_i+1)}$. Therefore as V_x is further outside the scope of occasions, the lower the non-centrality parameter, the greater the leverage on the vertex point, the lower the power.

3.4 Studies of Coverage and Power

To test the validity of test statistics presented in this chapter, Monte Carlo simulation studies were performed for two growth curves, random intercept model (3.1) and random slope model (3.2) For each model, we construct the confidence intervals for the x and y coordinates of the vertex using two different methods; the gradient method and the delta method for x-coordinate and the delta method and mean response method for y-coordinate respectively. As proved that the delta method and the mean response method are identical, the simulation results are only presented once. Confidence region and power analysis are also provided. Since some methods in this dissertation assume large sample, the validity of these methods are examined for small sample size. Thus different sample sizes are selected for the simulation studies across different type I error rates. Six occasions are chosen for the growth curve, that is, each individual is measure at six different time points, assuming no missing data. In order to investigate the influence of the leverage of the vertex, different models with vertex within and outside the scope of occasion are examined.

3.4.1 Random Intercept Model, X-coordinate of Vertex Within Scope of Occasions

For random intercept model (3.1), 1000 data sets are generated with regression coefficients $\boldsymbol{\beta}' = (\beta_0, \beta_1, \beta_2) = (2, 8, -1)$ and variances of random effect $\sigma^2_{\alpha_0} = 1, \sigma^2_e = 0.5$ for three different sample sizes 100, 50 and 25. These parameters are chosen for easy explanation. The true model is,

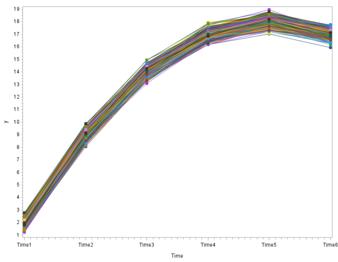
$$y_{ij} = 2 + 8t_{ij} - t^2_{ij} + \alpha_{0i} + \epsilon_{ij}, \qquad i = 1, 2, ..., N \quad j = 1, 2, ..., 6.$$

The true vertex of this quadratic growth curve is $\boldsymbol{V}' = (4, 18)$, where $V_x = 4$ is within the scope of occasions, $[0, 5]$. Profile plots and smoothed profile plots of 1000 data sets are shown in Figure 3.4 for sample size 20, 50 and 100; the figure indicates the quadratic trend of the growth curve. The larger the sample size, the narrower the width of all the curves.

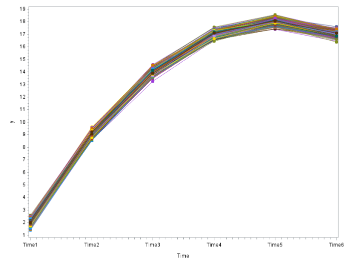
The chi-square QQ plots of the response variable $\boldsymbol{y}$ are also provided in Figure 3.5 for sample size 20, 50 and 100. In the figure, when the sample size is relatively large, 50 or 100, the points are around the straight line, which reveals the variables are multivariate normal. However, when the sample size is small, 20, the chi-square QQ plot reflects the minor violation of the normality assumption, since the points are not around or on a straight line.

Simulation Results for Confidence Intervals and Confidence Region

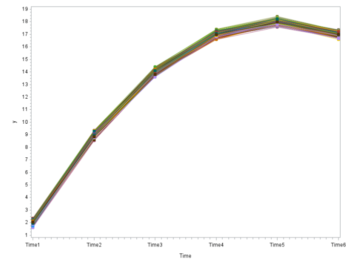
The results of the simulation for confidence intervals of x-coordinate are shown in Table 3.1, where symbol D and G represent delta method and gradient method respectively. The table includes the empirical coverage p as well as the lower bound,



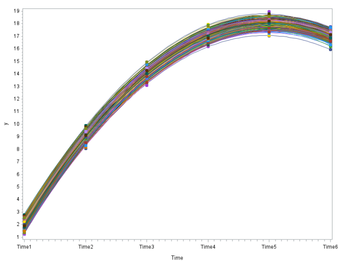
(a) Profile Plot for N = 20



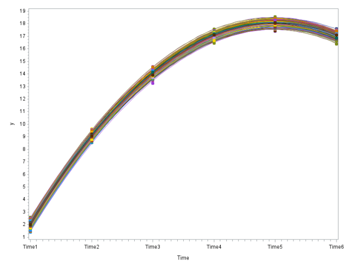
(b) Profile Plot for N = 50



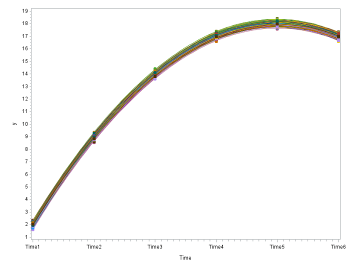
(c) Profile Plot for N = 100



(d) Smoothed Plot for N = 20

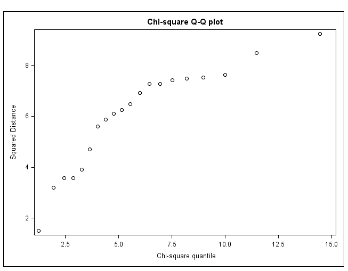


(e) Smoothed Plot for N = 50

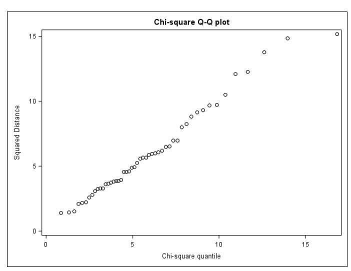


(f) Smoothed Plot for N = 100

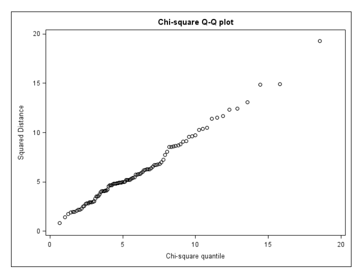
Figure 3.4: Profile and Smoothed Profile Plots for Random Intercept Model



(a) $N = 20$



(b) $N = 50$



(c) $N = 100$

Figure 3.5: Chi-square QQ Plot for Random Intercept Model

upper bound and the width of the empirical coverage, where the empirical coverage is the count that the computed confidence interval contains the true V_x divided by 1000, the number of data sets, and the lower and upper bounds are computed using Wald-type confidence interval (Brown *et al.*, 2001). For each data set, if the computed confidence interval contains the true value, it is coded as 1, otherwise 0. Using the count divided by the total number of data sets 1000, we obtain the empirical coverage $p = \frac{count}{1000}$. Because of only two possible outcomes, 0 or 1, the count has a binomial distribution of 1000 independent experiments, each of which yields success with probability $1 - \alpha$. The mean and variance of this binomial distribution are $1000 \cdot p$ and $1000 \cdot p \cdot (1 - p)$. Hence the standard deviation of coverage is $\sqrt{\frac{(\alpha)(1-\alpha)}{1000}}$ and the approximate bounds using normal approximation of the true coverage are $p \pm Z_{1-\alpha/2}\sqrt{\frac{(\alpha)(1-\alpha)}{1000}}$. The width is the mean of the difference of the upper and lower limits of interval for 1000 data sets. From Table 3.1, only one of the 18 conditions has the nominal coverage outside of the bounds; it is sample size 50 with α level 0.05 for the delta method. The width shows that the difference between the delta method and gradient method is small, since all widths are less than 0.3. All coverages are slightly low which may be due to relatively small sample sizes. The conclusion is drawn that both methods are applicable to obtain the confidence interval of the estimated x-coordinate of the vertex for different sample sizes tested.

The results of simulation for confidence intervals of y-coordinate are shown in Table 3.2. The two methods, the delta method and the mean response method, are identical, hence only one result is shown in the table. Similar to Table 3.1, Table 3.2 includes the empirical coverage, the count that the computed confidence interval contains the true V_y divided by the number of data sets 1000, as well as lower bound, upper bound and width for the empirical coverage. From the table, one of the 9 conditions have nominal coverage outside the bound, sample size 20 and α level 0.1. However, for the other four conditions, sample size 20 and α level 0.01 and 0.05, the method gives reasonable results. We conclude that the method is valid for computing the confidence interval of $y-$coordinate of the vertex.

The confidence region (3.15) has an elliptic shape, as discussed in Section 3.2.4. Figure 3.6 displays the true quadratic curve, $y = 2 + 8x + x^2$, and the ellipse for the confidence region. For the graph, the value of chi-square, 599, 100 times 5.99

Table 3.1: Confidence Intervals for X-Coordinate of the Vertex

α	Sample Size	Empirical Coverage D	lower bound D	upper bound D	width D	Empirical Coverage G	lower bound G	upper bound G	width G
0.01	100	0.985	0.9751	0.9949	0.099	0.984	0.97378	0.99422	0.099
0.01	50	0.984	0.97378	0.99422	0.141	0.987	0.97778	0.99622	0.141
0.01	20	0.984	0.97378	0.99422	0.223	0.983	0.97247	0.99353	0.224
0.05	100	0.940	0.92528	0.95472	0.076	0.937	0.92194	0.95206	0.076
✗ 0.05	50	**0.932**	**0.9164**	**0.9476**	**0.107**	0.935	0.91972	0.95028	0.107
0.05	20	0.944	0.92975	0.95825	0.170	0.945	0.93087	0.95913	0.170
0.1	100	0.887	0.87053	0.90347	0.064	0.890	0.87372	0.90628	0.064
0.1	50	0.888	0.87159	0.90441	0.090	0.886	0.86947	0.90253	0.090
0.1	20	0.888	0.87159	0.90441	0.142	0.890	0.87372	0.90628	0.143

* Random intercept model, when x-coordinate of vertex is within occasions for one sample

* D represents the delta method, and G represents the gradient method

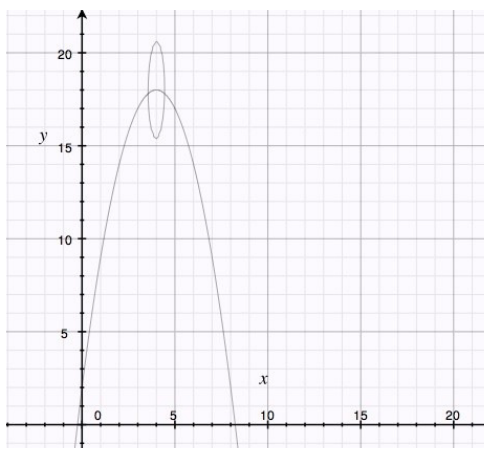


Figure 3.6: Confidence Region of Vertex

(the critical value of chi-square distribution with two degrees of freedom when α level equals 0.05) is chosen to make the ellipse clearer to see. If the chosen critical value is small, the ellipse in the figure reduces to a dot.

The results of the simulation for the confidence region based on delta method of the vertex are shown in Table 3.3. The table includes the empirical coverage, the count of confidence regions containing the true vertex divided by the number of data

Table 3.2: Confidence Intervals for Y-Coordinate of the Vertex

α	Sample Size	Empirical Coverage	lower bound	upper bound	width
0.01	100	0.990	0.98190	0.99810	0.554
0.01	50	0.981	0.96988	0.99212	0.778
0.01	20	0.984	0.97378	0.99422	1.223
0.05	100	0.942	0.92751	0.95649	0.422
0.05	50	0.945	0.93087	0.95913	0.592
0.05	20	0.941	0.92640	0.95560	0.931
0.1	100	0.896	0.88012	0.91188	0.354
0.1	50	0.899	0.88333	0.91467	0.497
✗ 0.1	20	**0.88**	**0.86310**	**0.89690**	**0.781**

* Random intercept model, when x-coordinate of vertex is within occasions for one sample

sets 1000, as well as lower bound and upper bound for the empirical coverage. From Table 3.3, only one of the 9 conditions has the nominal coverage outside the bounds; it is sample size 100 and α level 0.05. Although the chi-square distribution with two degrees of freedom applied to compute the confidence region is approximate, we conclude that the method is practicable for the confidence region for different sample sizes tested. A before, empirical coverage is slightly low for all conditions.

Results for Power

To compare the simulated power and the theoretical power, calculated using power functions derived in Section 3.3, hypothesized points as shown in Table 3.4 are tested; the results are also displayed in Table 3.4. The hypothesized points are selected due to the medium effect size. As proved, there is an unique solution for the vertex if $\boldsymbol{\beta}$ is fixed, however there are infinite many solutions for $\boldsymbol{\beta}$ if the vertex is given. Therefore

Table 3.3: Confidence Region of the Vertex

Type I Error	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.988	0.97913	0.99687
0.01	50	0.985	0.97510	0.99490
0.01	20	0.98	0.96860	0.99140
✗ 0.05	100	**0.933**	**0.91750**	**0.94850**
0.05	50	0.936	0.92083	0.95117
0.05	20	0.941	0.92640	0.95560
0.1	100	0.889	0.87266	0.90534
0.1	50	0.884	0.86734	0.90066
0.1	20	0.886	0.86947	0.90253

* Random intercept model, when x-coordinate of vertex is within occasions for one sample

the fixed regression parameters $\boldsymbol{\beta}$ are chosen initially, then the vertex is computed using the selected $\boldsymbol{\beta}$ to obtain the unique solution. In the table, the theoretical power for chi-square and F test are displayed as well as the simulated power with the lower and upper bounds for chi-square and F test respectively. The result shows that the confidence intervals of simulated power for chi-square test contain all the chi-square theoretical power; only 2 of 15 confidence intervals of simulated power for F test not include the F theoretical power. Further more, 13 of 15 theoretical power for F test are no less than the theoretical power for chi-square test. Roughly, we conclude that the power for F test is greater than that for chi-square test.

Table 3.4: Comparison of Theoretical Power and Simulated Power*

N	β_0	β_1	β_2	V_{0x}	V_{0y}	SPC	LBC	UBC	ChiSqP	SPF	LBF	UBF	FP
N = 100	2.05	8.05	-1.02	3.946	17.933	0.729	0.701	0.757	0.721	0.824	0.800	0.848	0.828
	1.98	8.02	-1.02	3.931	17.745	0.974	0.964	0.984	0.972	0.994	0.989	0.999	0.996
	1.95	8.05	-1.03	3.908	17.679	1	1	1	1	1	1	1	1
	2.02	7.95	-0.98	4.06	18.143	0.801	0.776	0.826	0.813	0.839	0.816	0.862	0.847
	2.1	7.9	-1	3.95	17.7025	0.923	0.906	0.940	0.925	0.999	0.997	1	1
N = 50	2.05	8.05	-1.02	3.946	17.933	0.397	0.367	0.427	0.418	0.487	0.456	0.518	0.501
	1.98	8.02	-1.02	3.931	17.745	0.768	0.742	0.794	0.762	0.885	0.865	0.905	0.886
	1.95	8.05	-1.03	3.908	17.679	0.947	0.971	0.949	0.986	0.992	0.986	0.998	0.62
	2.02	7.95	-0.98	4.06	18.143	0.505	0.474	0.536	0.501	0.524	0.493	0.555	0.522
	2.1	7.9	-1	3.95	17.7025	0.656	0.627	0.685	0.650	✗ **0.974**	**0.964**	**0.984**	**0.961**
N = 20	2.05	8.05	-1.02	3.946	17.933	0.197	0.172	0.222	0.189	0.214	0.188	0.239	0.213
	1.98	8.02	-1.02	3.931	17.745	0.380	0.350	0.410	0.372	✗ **0.504**	**0.473**	**0.535**	**0.464**
	1.95	8.05	-1.03	3.908	17.679	0.62	0.590	0.650	0.595	0.716	0.688	0.744	0.696
	2.02	7.95	-0.98	4.06	18.143	0.265	0.238	0.292	0.224	0.233	0.207	0.259	0.222
	2.1	7.9	-1	3.95	17.7025	0.317	0.288	0.346	0.300	0.619	0.589	0.649	0.598

* Random intercept model with x-coordinate of vertex within occasions

* SPC = simulated power for chi-square test, LBC = lower bound of simulated power for chi-square test, UBC = upper bound of simulated power for chi-square test, ChiSqP = theoretical chi square power

* SPF = simulated power for F test, LBF = lower bound of simulated power for F test, UBF = upper bound of simulated power for F test, FP = theoretical F power

The simulation result for power of the chi-square test are also presented in Table 3.5. The hypothesized vertex value is chosen directly depending on the distance of 0.05 and 0.1 for x-coordinate and y-coordinate between the tested point and the true vertex. All the pairwise combinations of these points are tested. The results include the empirical power, the count of samples which reject the null hypothesis divided by 1000, the total number of data sets, as well as lower bound and upper bound for the interval around the empirical power.

From the table, when we keep V_{0x} at the true value, the change of V_{0y} does not affect the power much. However, when we keep V_{0y} at the true value, the change of V_{0x} affects the power much more. The result shows that the x-coordinate is more sensitive than y-coordinate. It can be explained on the width of confidence interval; the widths of y-coordinate confidence interval are commonly larger than x-coordinate confidence interval, which means that the variation of y-coordinate is larger than x-coordinate. The reason is that the number of time points we choose for t is only six, i.e. the domain of t is $\{0, 1, 2, 3, 4, 5\}$, however the range of y-coordinate is much wider. At last, if the true vertex is tested, and the empirical power is nearly equal to the size of the test, 0.05. For a non-vertex point, there is a positive relationship between power and sample size, i.e. the larger the sample size, the greater the power.

Simulation results reveal that all the methods and statistics perform reasonably for the random intercept model when the x-coordinate of the true vertex is within the domain of the data.

3.4.2 Random Slope Model, X-coordinate of Vertex Within Scope of Occasions

For random intercept model (3.2), 1000 data sets are generated with the fixed coefficient parameters $\boldsymbol{\beta}' = (\beta_0, \beta_1, \beta_2) = (2, 8, -1)$ and variances of random effect $\sigma^2_{\alpha_0} = 1, \sigma^2_{\alpha_1} = 0.5, \sigma_{\alpha_0,\alpha_1} = 0, \sigma^2_e = 0.5$ for three sample sizes 100, 50 and 20. The

Table 3.5: Power Analysis for Chi-square Test ($\alpha = 0.05$)

			$N=100$			$N=50$			$N=20$	
V_{0x}	V_{0y}	SPower	LB	UB	SPower	LB	UB	SPower	LB	UB
3.9	18.05	0.999	0.99704	1.00096	0.93	0.91419	0.94581	0.56	0.52923	0.59007
3.9	17.95	0.999	0.99704	1.00096	0.936	0.92083	0.95117	0.57	0.53931	0.60069
3.9	18.1	0.999	0.99704	1.00096	0.932	0.91640	0.94760	0.569	.53831	0.59969
3.9	17.9	0.999	0.99704	1.00096	0.982	0.91198	0.94402	0.558	0.52722	0.58878
3.95	18.05	0.671	0.64188	0.70012	0.346	0.31652	0.37548	0.169	0.14577	0.19223
3.95	17.95	0.657	0.62758	0.68642	0.339	0.30966	0.37548	0.172	0.14861	0.19539
3.95	18.1	0.712	0.68393	0.74007	0.389	0.35878	0.41922	0.183	0.15903	0.20697
3.95	17.9	0.694	0.66544	0.72256	0.372	0.34204	0.40196	0.812	0.15809	0.20591
4.05	18.05	0.644	0.61432	0.67368	0.37	0.34008	0.39992	0.198	0.17330	0.22270
4.05	17.95	0.653	0.62350	0.68250	0.392	0.36174	0.42226	0.217	0.19145	0.24255
4.05	18.1	0.673	0.64392	0.70208	0.382	0.35189	0.41211	0.205	0.17998	0.23002
4.05	17.9	0.7	0.6716	0.7284	0.417	0.38644	0.44756	0.215	0.18954	0.24046
4.1	18.05	0.994	0.98921	0.99879	0.894	0.87492	0.91308	0.53	0.49907	0.56093
4.1	17.95	0.995	0.99063	0.99937	0.9	0.88141	0.91859	0.534	0.50308	0.56492
4.1	18.1	0.994	0.98921	0.99879	0.895	0.87600	0.91400	0.54	0.50911	0.57089
4.1	17.9	0.995	0.99063	0.99937	0.909	0.89117	0.92683	0.549	0.51816	0.57984
4	18	0.067	0.05150	0.08250	0.064	0.04883	0.07917	0.059	0.04440	0.07360

* SPower = Simulated Power, LU = Lower Bound, UB = Upper Bound

* random intercept model, when x-coordinate of vertex is within occasions for one sample

true model is,

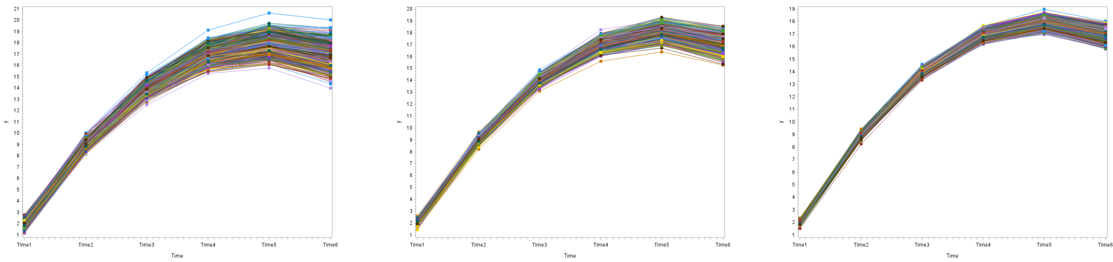
$$y_{ij} = 2 + 8t_{ij} - t_{ij}^2 + \alpha_{0i} + \alpha_{1i}t_{ij} + \epsilon_{ij}, \qquad i = 1, 2, ..., N \quad j = 1, 2, ..., 6.$$

Then,

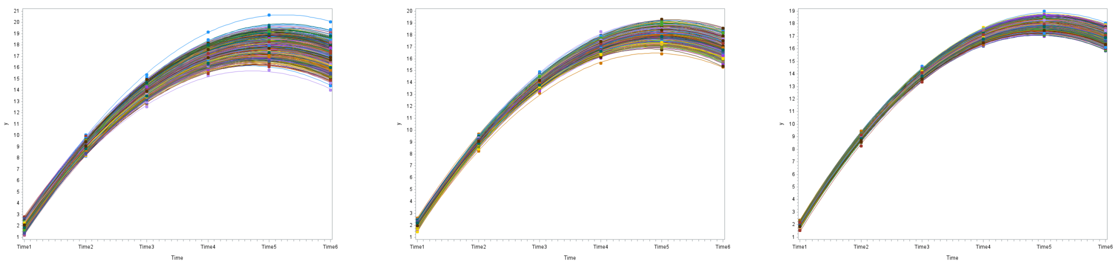
$$E\{y_{ij}\} = 2 + 8t_{ij} - t_{ij}^2, \qquad i = 1, 2, ..., N \quad j = 1, 2, ..., 6.$$

The true vertex of the quadratic growth curve is $\boldsymbol{V}' = (4, 18)$, where $V_x = 4$ is within the scope of occasions $[0, 5]$. Profile plots of 1000 data sets are displayed in Figure 3.7 for sample size 20, 50 and 100 as well as the smoothed profile plots. The figure suggests the quadratic trend intuitively; the width of the graph becomes more narrow as the sample size increases. Comparing profile plots for the random slope model 3.7 to random intercept model 3.4, the former curves are wider even for larger sample size.

The chi-square QQ plots of the response variable $\boldsymbol{y}$ for sample sizes 20, 50, and 100 are displayed in Figure 3.8. Similar findings are drawn from the figure: if the sample size is relatively large, the normality assumption is satisfied; if the sample size is small, the normality assumption is violated slightly.

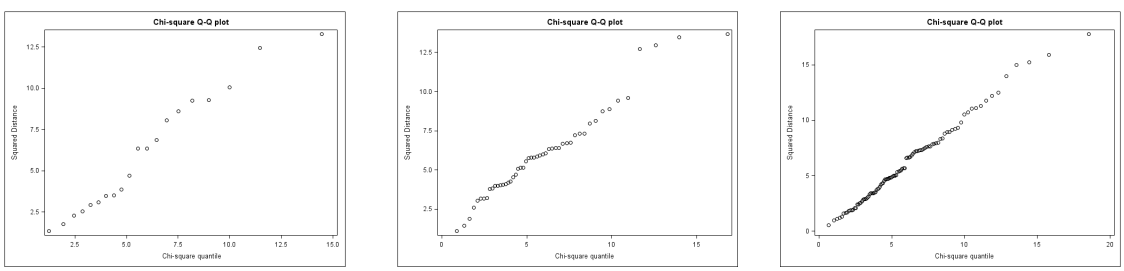


(a) Profile Plot for N = 20 (b) Profile Plot for N = 50 (c) Profile Plot for N = 100



(d) Smoothed Plot for N = 20 (e) Smoothed Plot for N = 50 (f) Smoothed Plot for N = 100

Figure 3.7: Profile and Smoothed Profile Plots for Random Slope



(a) $N = 20$ (b) $N = 50$ (c) $N = 100$

Figure 3.8: Chi-square QQ Plot for Random Slope Model

Simulation Results for Confidence Interval and Confidence Region

The results of simulation for confidence intervals of the x-coordinate with two covariance structures, UN and CS, are displayed in Table 3.6; UN is the more appropriate covariance structure. The results include the empirical coverage as well as lower bound, upper bound and width for the empirical coverage. In Table 3.6 (a), four of the 18 conditions have nominal coverage outside the bounds; they are sample size 100 and α level 0.1, both methods, and sample size 20 and α level 0.1, both methods. The widths indicate that the variation of the two methods is small. In Table 3.6 (b), the unreasonably high empirical coverage of all the tests result from the inappropriate covariance structure. To sum up, we conclude that both the methods are appropriate for the confidence interval of x-coordinate of the estimated vertex for different sample sizes tested if the covariance structure is correctly specified.

Comparing the two different covariance structures, the coverage for the model with covariance structure CS is extremely high. The reason is that the number of estimated covariance parameters for compound symmetry structure is only 2; it does not match the random slope model as provided in Section 3.1. That is, the strong assumption that the correlation between any pair of measurements is the same regardless of the time interval between the measurements is not satisfied for the random intercept model. Simulation results also illustrate that the covariance structure UN is more reasonable than CS.

The results of simulation for confidence intervals of the y-coordinate of the estimated vertex with covariance structure UN and CS are shown in Table 3.7. Only one result is shown, since the mean response method and the delta method are equivalent. The tables include the empirical coverage as well as lower bound, upper bound and width for the empirical coverage. In Table 3.7 (a), two of the 9 conditions have nominal coverage that is not within the bounds; they are sample size 100 and α level 0.1, and sample size 20 and α level 0.05. In Table 3.7 (b), all the conditions have the nominal coverage within the bounds. In conclusion, the method performs well for the confidence interval of y-coordinate of vertex for different sample sizes tested.

Comparing the two different covariance structures as for x-coordinate, the nominal coverage obtained from covariance structure CS is always large. As stated, there are

Table 3.6: Confidence Intervals for X-Coordinate

a) Covariance Structure Unstructured

α	Sample Size	Empirical Coverage D	lower bound D	upper bound D	width D	Empirical Coverage G	lower bound G	upper bound G	width G
0.01	100	0.984	0.97378	0.99422	0.207	0.985	0.97510	0.99490	0.207
0.01	50	0.983	0.97247	0.99353	0.291	0.982	0.97117	0.99283	0.291
0.01	20	0.986	0.97643	0.99557	0.459	0.985	0.97510	0.99490	0.460
0.05	100	0.945	0.93087	0.95913	0.157	0.944	0.92975	0.95825	0.157
0.05	50	0.94	0.92528	0.95472	0.221	0.939	0.92417	0.95383	0.221
0.05	20	0.938	0.92305	0.95295	0.349	0.937	0.92194	0.95206	0.350
✗ 0.1	100	**0.879**	**0.86204**	**0.89596**	**0.132**	**0.876**	**0.85886**	**0.89314**	**0.132**
0.1	50	0.89	0.87372	0.90628	0.186	0.889	0.87266	0.90534	0.186
✗ 0.1	20	**0.882**	**0.86522**	**0.89878**	**0.293**	**0.882**	**0.86522**	**0.89878**	**0.294**

b) Covariance Structure Compound Symmetry

α	Sample Size	Empirical Coverage D	lower bound D	upper bound D	width D	Empirical Coverage G	lower bound G	upper bound G	width G
0.01	100	0.992	0.98475	0.99925	0.234	0.993	0.98621	0.99979	0.234
0.01	50	0.991	0.98331	0.99869	0.329	0.99	0.98190	0.99810	0.329
0.01	20	0.99	0.98190	0.99810	0.522	0.99	0.98190	0.99810	0.524
0.05	100	0.991	0.98515	0.99685	0.240	0.99	0.98383	0.99617	0.240
0.05	50	0.968	0.95709	0.97891	0.250	0.966	0.95477	0.97723	0.250
0.05	20	0.962	0.95015	0.97385	0.340	0.964	0.95245	0.97555	0.340
0.1	100	0.93	0.91673	0.94327	0.150	0.929	0.91564	0.94236	0.150
0.1	50	0.926	0.91238	0.93962	0.210	0.923	0.90913	0.93687	0.210
0.1	20	0.93	0.91673	0.94327	0.334	0.93	0.91673	0.94327	0.334

* Random slope model, when x-coordinate of vertex is within occasions for one sample
* D represents the delta method, and G represents the gradient method

only two parameters for CS which is not appropriate for the random slope model. The same conclusion is drawn that covariance structure UN is more applicable than CS.

The simulation results for confidence region of the vertex with covariance structure UN and CS are displayed in Table 3.8. The tables include the empirical coverage as well as lower bound and upper bound for the empirical coverage. From Table 3.8 (a), only one of the 9 conditions had nominal coverage outside the bounds; it is sample size 50 and α level 0.1. In Table 3.8 (b), all 9 conditions result in good empirical coverage. We conclude that the approximate chi-square distribution employed for confidence region is valid for different sample sizes tested.

Results for Power

We investigate power analysis with covariance structure UN only, since UN has been shown to be more appropriate than CS using the simulation results of confidence

Table 3.7: Confidence Intervals for Y-Coordinate

a) Covariance Structure Unstructured

α	Sample Size	Empirical Coverage	lower bound	upper bound	width
0.01	100	0.989	0.98051	0.99749	1.550
0.01	50	0.991	0.98331	0.99869	2.178
0.01	20	0.984	0.97378	0.99422	3.430
0.05	100	0.941	0.92640	0.95560	1.180
0.05	50	0.938	0.92305	0.95295	1.658
✗ 0.05	20	**0.933**	**0.91750**	**0.94850**	**2.610**
✗ 0.1	100	**0.88**	**0.86310**	**0.89690**	**0.990**
0.1	50	0.892	0.87585	0.90815	1.392
0.1	20	0.889	0.87266	0.90534	2.191

b) Covariance Structure Compound Symmetry

α	Sample Size	Empirical Coverage	lower bound	upper bound	width
0.01	100	0.994	0.98771	1.00029	1.756
0.01	50	0.998	0.99436	1.00164	2.465
0.01	20	0.992	0.98475	0.99925	3.917
0.05	100	0.983	0.97499	0.99101	2.035
0.05	50	0.965	0.95361	0.97639	1.876
0.05	20	0.963	0.95130	0.97470	2.981
0.1	100	0.926	0.91238	0.93962	1.122
0.1	50	0.927	0.91347	0.94053	1.574
0.1	20	0.931	0.91782	0.94418	2.502

* Random slope model, when x-coordinate of vertex is within occasions for one sample

interval and confidence region. Values of $\boldsymbol{\beta}$ and $\boldsymbol{V}$ under H_0 are given in Table 3.9; the results for comparing the theoretical power and simulated power for both chi-square test and F test are also displayed in the table. As illustrated, the fixed regression parameters $\boldsymbol{\beta}$ are selected first, then the vertex is computed using the $\boldsymbol{\beta}$ to obtain the solution. From the table, the theoretical power for the chi-square and F tests are displayed as well as the simulated power with the lower and upper bounds for chi-square and F test respectively. Only 2 of 15 confidence intervals of simulated power for the chi-square test do not contain the chi-square theoretical power; Seven of 15 confidence intervals of simulated power for F test do not include the F theoretical power. Comparing Table 3.4 and 3.9, both theoretical and simulated power for chi-square test and F test decreases when the random slope term is added in the model.

Table 3.8: Confidence Region of the Vertex

a) Covariance Structure Unstructured

α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.99	0.98190	0.99810
0.01	50	0.986	0.97643	0.99557
0.01	20	0.984	0.97378	0.99422
0.05	100	0.936	0.92083	0.95117
0.05	50	0.936	0.92083	0.95117
0.05	20	0.94	0.92528	0.95472
0.1	100	0.886	0.86947	0.90253
✗ 0.1	50	**0.875**	**0.85780**	**0.89220**
0.1	20	0.885	0.86840	0.90160

b) Covariance Structure Compound Symmetry

α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.992	0.98475	0.99925
0.01	50	0.99	0.98190	0.99810
0.01	20	0.989	0.98051	0.99749
0.05	100	0.951	0.93762	0.96438
0.05	50	0.949	0.93536	0.96264
0.05	20	0.954	0.94102	0.96698
0.1	100	0.903	0.88760	0.91840
0.1	50	0.894	0.87799	0.91001
0.1	20	0.913	0.89834	0.92766

* Random slope model, when x-coordinate of vertex is within occasions for one sample

Table 3.9: Comparison of Theoretical Power and Simulated Power

N	β_0	β_1	β_2	V_{0x}	V_{0y}	SPC	LBC	UBC	ChiSqP	SPF	LBF	UBF	FP
N = 100	2.05	8.05	-1.02	3.946	17.933	0.436	0.405	0.467	0.440	0.389	0.359	0.419	0.390
	1.98	8.02	-1.02	3.931	17.745	0.406	0.376	0.436	0.408	0.400	0.370	0.430	0.380
	1.95	8.05	-1.03	3.908	17.679	0.697	0.669	0.725	0.683	0.700	0.672	0.728	0.691
	2.02	7.95	-0.98	4.06	18.143	0.369	0.339	0.399	0.360	0.316	0.287	0.345	0.321
	2.1	7.9	-1	3.95	17.7025	0.207	0.182	0.232	0.183	✗ **0.250**	**0.233**	**0.277**	**0.230**
N = 50	2.05	8.05	-1.02	3.946	17.933	0.251	0.224	0.278	0.239	✗ **0.234**	**0.208**	**0.260**	**0.206**
	1.98	8.02	-1.02	3.931	17.745	0.229	0.203	0.255	0.221	0.221	0.195	0.247	0.200
	1.95	8.05	-1.03	3.908	17.679	0.382	0.352	0.412	0.390	0.407	0.377	0.437	0.381
	2.02	7.95	-0.98	4.06	18.143	✗ **0.226**	**0.200**	**0.252**	**0.197**	✗ **0.205**	**0.180**	**0.230**	**0.173**
	2.1	7.9	-1	3.95	17.7025	0.123	0.103	0.143	0.113	✗ **0.163**	**0.140**	**0.186**	**0.132**
N = 20	2.05	8.05	-1.02	3.946	17.933	0.112	0.092	0.132	0.120	✗ **0.127**	**0.106**	**0.148**	**0.104**
	1.98	8.02	-1.02	3.931	17.745	0.110	0.091	0.129	0.114	0.119	0.099	0.139	0.103
	1.95	8.05	-1.03	3.908	17.679	0.168	0.145	0.191	0.177	0.168	0.145	0.191	0.167
	2.02	7.95	-0.98	4.06	18.143	✗ **0.145**	**0.123**	**0.167**	**0.105**	✗ **0.121**	**0.101**	**0.141**	**0.094**
	2.1	7.9	-1	3.95	17.7025	0.072	0.056	0.088	0.074	✗ **0.1**	**0.08**	**0.119**	**0.080**

* Random slope model with x-coordinate of vertex within occasions

* SPC = simulated power for chi-square test, LBC = lower bound of simulated power for chi-square test, UBC = upper bound of simulated power for chi-square test, ChiSqP = theoretical chi square power

* SPF = simulated power for F test, LBF = lower bound of simulated power for F test, UBF = upper bound of simulated power for F test, FP = theoretical F power

The results of power for chi-square test by directly choosing the hypothesized vertex points are displayed in Table 3.10. The points to be tested are chosen based on the distance of 0.05 and 0.1 between the hypothesized point and true vertex, for both the x and y direction. The table includes the power as well as lower bound and upper bound for the interval around the empirical power.

From the table, when V_{0x} is kept equal to the true value, a change of V_{0y} does not affect the power dramatically. However, when we keep V_{0y} equal to the true value, a change of V_{0x} extremely affects the results. It indicates that the x-coordinate is more sensitive than the y-coordinate; the reason is similar as shown for random intercept model. Checking the width from the confidence interval tables, the width of y-coordinate confidence interval is commonly larger than that of x-coordinate, which means that the variation of y-coordinate is larger than x-coordinate. It is because the number of occasions for the x-coordinate is only six, however the range of y-coordinate is much broader than x-coordinate. Finally, the true vertex point is tested, the result shows the empirical power is nearly equal to the size of the test. One more conclusion is that the relationship between sample size and power is positive.

Table 3.10: Power Analysis for Chi-square Test ($\alpha = 0.05$, UN)

			$N = 100$			$N = 50$			$N = 20$	
V_{0x}	V_{0y}	SPower	LB	UB	SPower	LB	UB	SPower	LB	UB
3.9	18.05	0.994	0.98921	0.99879	0.854	0.83211	0.87589	0.439	0.40824	0.46976
3.9	18.1	0.997	0.99361	1.00039	0.894	0.87492	0.91308	0.484	0.45303	0.51497
3.9	17.9	0.964	0.95245	0.97555	0.711	0.68290	0.73910	0.32	0.29109	0.34891
3.9	17.95	0.977	0.96771	0.98629	0.763	0.73664	0.78936	0.352	0.32240	0.38160
3.95	18.05	0.594	0.56356	0.62444	0.341	0.31162	0.37038	0.153	0.13069	0.17531
3.95	17.95	0.397	0.36667	0.42733	0.241	0.21449	0.26751	0.104	0.08508	0.12292
3.95	18.1	0.689	0.66031	0.71769	0.396	0.36569	0.42631	0.178	0.15429	0.20171
3.95	17.9	0.326	0.29695	0.35505	0.192	0.16759	0.21641	0.094	0.075912	0.11209
4.05	18.05	0.442	0.41122	0.47278	0.247	0.22027	0.27373	0.149	0.12693	0.17107
4.05	18.1	0.348	0.31848	0.37752	0.207	0.18189	0.23211	0.137	0.11569	0.15831
4.05	17.95	0.604	0.57369	0.63431	0.351	0.32142	0.38058	0.183	0.15903	0.20697
4.05	17.9	0.69	0.66133	0.71867	0.413	0.38248	0.44352	0.215	0.18954	0.24046
4.1	18.05	0.952	0.93875	0.96525	0.752	0.72523	0.77877	0.408	0.37754	0.43846
4.1	18.1	0.935	0.91972	0.95028	0.71	0.68188	0.73812	0.378	0.34795	0.40805
4.1	17.95	0.984	0.97622	0.99178	0.827	0.80356	0.85044	0.476	0.44505	0.50695
4.1	17.9	0.993	0.98783	0.99817	0.86	0.83849	0.88151	0.509	0.47801	0.53999
4	18	0.064	0.04883	0.07917	0.064	0.04883	0.07917	0.06	0.04528	0.07472

* SPower = Simulated Power, LU = Lower Bound, UB = Upper Bound
* Random slope model, when x-coordinate of vertex is within occasions for one sample

In conclusion, simulation results illustrate that the methods and statistics perform reasonably for the random slope model when the x-coordinate of vertex is within the scope of occasions.

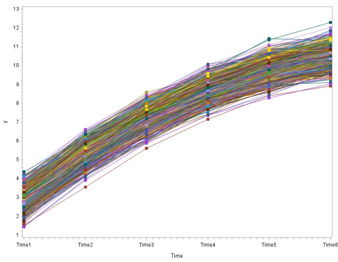
3.4.3 Random Intercept Model, X-coordinate of Vertex Outside Scope of Occasions

In the this section, a vertex with the x-coordinate outside the scope of occasions is examined for the random intercept model (3.1). One thousand data sets are generated with regression coefficients $\boldsymbol{\beta}' = (\beta_0, \beta_1, \beta_2) = (3, 2.5, -0.2)$, variances for random effect and error $\sigma^2_{\alpha_0} = 1, \sigma^2_e = 0.5$ for three sample sizes 50, 25 and 10. The true model is,

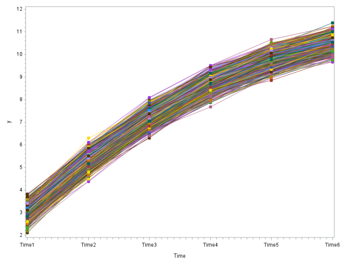
$$y_{ij} = 3 + 2.5t_{ij} - 0.2t_{ij}^2 + \alpha_{0i} + \epsilon_{ij}, \qquad i = 1, 2, ..., N \quad j = 1, 2, ..., 6.$$

The true vertex of this quadratic growth curve is $\boldsymbol{V}' = (6.25, 10.8125)$, where $V_x = 6.25$ is outside the scope of occasions, $[0, 5]$. Profile plots for 1000 data sets are shown in Figure 3.9 for sample size 10, 25 and 50 and smoothed profile plots as well. The width of graph tends to be narrower as the sample size increases. From these plots, intuitively vertices of quadratic curves are outside the scope of the occasions, since the curves always increase along the whole range of occasions and never turn.

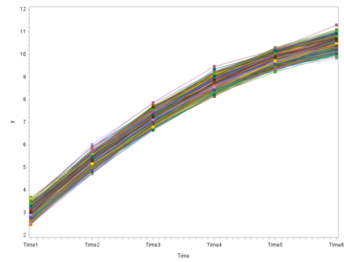
The chi-square QQ plots of the response variable $\boldsymbol{y}$ for sample size 10, 25, and 50 are displayed in Figure 3.10. The plots for sample size 25 and 50 support the assumption of normality, however the plot for sample size 10 violates the normality assumption severally. We would expect the simulation result for sample size 10 may not be good enough due to the small sample size, the following simulation results confirm it.



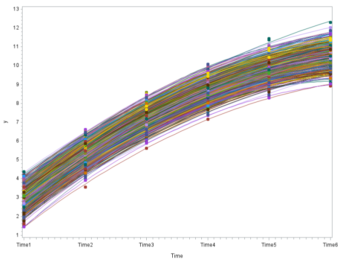
(a) Profile Plot for N = 10



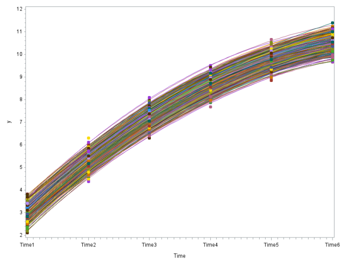
(b) Profile Plot for N = 25



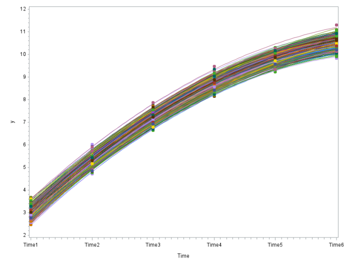
(c) Profile Plot for N = 50



(d) Smoothed Plot for N = 10

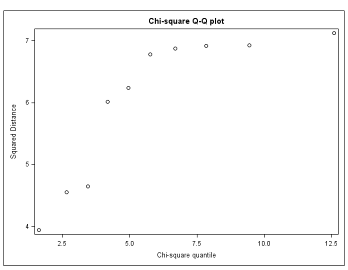


(e) Smoothed Plot for N = 25

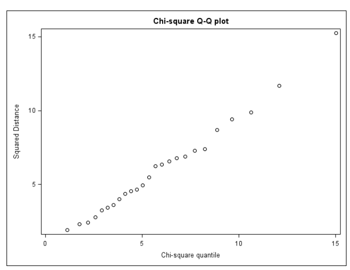


(f) Smoothed Plot for N = 50

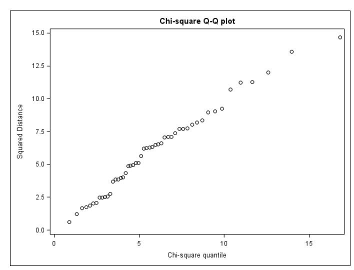
Figure 3.9: Profile and Smoothed Plots for Mixed Model with Random Intercept



(a) $N = 10$



(b) $N = 25$



(c) $N = 50$

Figure 3.10: Chi-square QQ Plot for Random Intercept Model

Simulation Results for Confidence Interval and Confidence Region

The results of the simulation for confidence intervals of x-coordinate are displayed in Table 3.11. The table contains the empirical coverage p as well as the lower bound, upper bound and the width of the empirical coverage. Only one of the 18 conditions has the bounds that do not include the nominal coverage; it is sample size 10 and

α level 0.01. The width becomes smaller as sample sizes are larger, however the widths of two methods are similar which indicates the similarity of these methods. In conclusion, both methods are applicable to derive the confidence interval of the estimated x-coordinate of vertex for different sample sizes tested.

Table 3.11: Confidence Intervals for X-Coordinate of the Vertex

α	Sample Size	Empirical Coverage D	lower bound D	upper bound D	width D	Empirical Coverage G	lower bound G	upper bound G	width G
0.01	50	0.991	0.98331	0.99869	1.634	0.991	0.98331	0.99869	1.712
0.01	25	0.986	0.97642	0.99557	2.386	0.989	0.98051	0.99749	2.649
✗ 0.01	10	**0.97**	**0.95611**	**0.98389**	**4.033**	0.982	0.97117	0.99283	6.336
0.05	50	0.954	0.94102	0.96698	1.244	0.948	0.93424	0.96176	1.278
0.05	25	0.95	0.93649	0.96351	1.816	0.956	0.94329	0.96871	1.926
0.05	10	0.943	0.92863	0.95737	3.070	0.937	0.92194	0.95206	4.013
0.1	50	0.912	0.89726	0.92674	1.044	0.908	0.89297	0.92303	1.064
0.1	25	0.913	0.89834	0.92766	1.524	0.898	0.88226	0.91374	1.588
0.1	10	0.901	0.88546	0.91654	2.577	0.902	0.88654	0.91747	2.998

* Random intercept model, when x-coordinate of vertex is outside occasions for one sample
* D represents the delta method, and G represents the gradient method

Table 3.12 shows the simulation results for confidence intervals of the y-coordinate. Similarly as in Table 3.11, Table 3.12 includes the empirical coverage, the count for the computed confidence interval contains the true V_y divided by total 1000 data sets, as well as lower bound, upper bound and width for the empirical coverage. From the table, three of the 9 conditions do not contain the nominal coverage within the bounds; they are sample size 50 with α level 0.05 and 0.1 and sample size 25 with α level 0.1. For the simulated width, it decreases as sample size increases as usual.

Table 3.12: Confidence Intervals for Y-Coordinate of the Vertex

α	Sample Size	Empirical Coverage	lower bound	upper bound	width
0.01	50	0.987	0.97778	0.99622	1.410
0.01	25	0.993	0.98621	0.99979	2.048
0.01	10	0.981	0.96988	0.99212	3.438
✗ 0.05	50	**0.963**	**0.95130**	**0.97470**	**1.074**
0.05	25	0.958	0.94557	0.97043	1.559
0.05	10	0.946	0.93199	0.96001	2.617
✗ 0.1	50	**0.927**	**0.91347**	**0.94053**	**0.901**
✗ 0.1	25	**0.921**	**0.90697**	**0.93503**	**1.309**
0.1	10	0.912	0.89726	0.92674	2.196

* Random intercept model, when x-coordinate of vertex is outside occasions for one sample

The results of the simulation for the confidence region of the vertex using the delta method are displayed in Table 3.13. The table includes the empirical coverage, as well as lower bound and upper bound for the empirical coverage. From Table 3.13, three of the 9 conditions have the nominal coverage outside the bounds; They are sample size 10 with α level 0.01, 0.05 and 0.1. Due to the approximate chi-square distribution we applied in the method, sample size 10 is too low to satisfy the large sample assumption. The conclusion is drawn that the chi-square distribution with two degrees of freedom is practical to compute the confidence region, however small sample size must be paid special attention to.

Table 3.13: Confidence Region of the Vertex

α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	50	0.99	0.98190	0.99810
0.01	25	0.98	0.96860	0.99140
✗ 0.01	10	0.953	0.93577	0.97023
0.05	50	0.943	0.92863	0.95737
0.05	25	0.943	0.92863	0.95737
✗ 0.05	10	0.899	0.88032	0.91768
0.1	50	0.901	0.88546	0.91654
0.1	25	0.89	0.87372	0.90628
✗ 0.1	10	0.842	0.82303	0.86097

* Random intercept model, when x-coordinate of vertex is outside occasions for one sample

Results for Power

The hypothesized points as shown in Table 3.14 are tested in order to compare the theoretical power and simulated power for both chi-square test and F test; the results are also given in Table 3.14. Again, the fixed regression parameters $\boldsymbol{\beta}$ are selected first, then the vertex is computed using $\boldsymbol{\beta}$ to obtain the solution. From the table, the theoretical power for chi-square and F tests are presented as well as the simulated power with the lower and upper bounds for chi-square and F tests. Seven of 15 confidence intervals for simulated power for the chi-square test not contain the chi-square theoretical power, especially for sample size 10. The result may be due to the violation of normality assumption as shown by the chi-square QQ plot. Only two of 15 confidence intervals for simulated power of the F test not include the F theoretical power. Furthermore, the theoretical power for F test is greater than the theoretical power for chi-square test under all conditions.

Table 3.14: Comparison of Theoretical Power and Simulated Power

N	β_0	β_1	β_2	V_{0x}	V_{0y}	SPC	LBC	UBC	ChiSqP	SPF	LBF	UBF	FP
N = 50	2.8	2.8	-0.25	5.6	10.64	✗ **0.550**	**0.519**	**0.579**	**0.855**	0.905	0.887	0.923	0.890
	3.2	2.8	-0.22	6.364	12.110	1	1	1	1	1	1	1	1
	2.5	3.0	-0.25	6	11.5	0.988	0.981	0.995	0.990	1	1	1	1
	2.5	2.6	-0.18	7.222	11.889	✗ **0.929**	**0.913**	**0.945**	**0.956**	1	1	1	1
	3.1	2.3	-0.18	6.389	10.447	0.610	0.580	0.640	0.594	✗ **0.987**	**0.980**	**0.994**	**0.976**
N = 25	2.8	2.8	-0.25	5.6	10.64	✗ **0.236**	**0.210**	**0.262**	**0.318**	0.600	0.570	0.630	0.574
	3.2	2.8	-0.22	6.364	12.110	0.992	0.986	0.998	0.991	1	1	1	1
	2.5	3.0	-0.25	6	11.5	0.821	0.797	0.845	0.841	1	1	1	1
	2.5	2.6	-0.18	7.222	11.889	0.693	0.664	0.722	0.717	1	1	1	1
	3.1	2.3	-0.18	6.389	10.447	✗ **0.359**	**0.329**	**0.389**	**0.328**	0.758	0.731	0.785	0.757
N = 10	2.8	2.8	-0.25	5.6	10.64	✗ **0.119**	**0.099**	**0.139**	**0.149**	0.255	0.228	0.282	0.240
	3.2	2.8	-0.22	6.364	12.110	0.749	0.722	0.776	0.757	0.933	0.918	0.949	0.943
	2.5	3.0	-0.25	6	11.5	0.458	0.427	0.489	0.439	0.984	0.976	0.992	0.986
	2.5	2.6	-0.18	7.222	11.889	✗ **0.462**	**0.431**	**0.493**	**0.340**	0.865	0.844	0.886	0.876
	3.1	2.3	-0.18	6.389	10.447	✗ **0.207**	**0.182**	**0.232**	**0.153**	✗ **0.385**	**0.355**	**0.415**	**0.340**

* Random intercept model with x-coordinate of vertex outside occasions

* SPC = simulated power for chi-square test, LBC = lower bound of simulated power for chi-square test, UBC = upper bound of simulated power for chi-square test, ChiSqP = theoretical chi square power

* SPF = simulated power for F test, LBF = lower bound of simulated power for F test, UBF = upper bound of simulated power for F test, FP = theoretical F power

The results of simulation for power are shown in Table 3.15. The points under null hypothesis are chosen based on the difference of 0.5 for both x-coordinate and y-coordinate between the non-vertex tested point and the true vertex. All the pairwise combinations of these points are tested. Even though all the tested x-coordinates are out the scope of occasions, these tested points are assumed to be reasonable. The results include the power, the count, which is the number of rejecting null hypothesis when the tested point is non-vertex point, divided by the total 1000 data sets, and the lower bound and upper bound for the interval around the empirical power. From the table, when we keep V_{0x} fixed, a change of V_{0y} affects the power dramatically. Similarly, when we keep V_{0y} fixed, a change of V_{0x} also affects the power much. At last, the true vertex is tested, and the empirical power is nearly equal to the size of the test. For a non-vertex point, there is a strongly positive relationship between power and sample size, i.e. power increases remarkably as sample size increases.

Table 3.15: Power Analysis for Chi-square Test ($\alpha = 0.05$)

		$N=50$			$N=25$			$N=10$		
V_{0x}	V_{0y}	SPower	LB	UB	SPower	LB	UB	SPower	LB	UB
5.25	9.8125	0.987	0.97998	0.99402	0.703	0.67468	0.73132	0.19	0.16568	0.21432
5.25	10.3125	0.954	0.94102	0.96698	0.539	0.50810	0.56990	0.127	0.10636	0.14764
5.25	11.3125	1	1	1	0.999	0.99704	1.00096	0.743	0.71592	0.77008
5.25	11.8125	1	1	1	1	1	1	0.955	0.94215	0.96785
5.75	9.8125	0.966	0.95477	0.97723	0.692	0.66339	0.72061	0.321	0.29206	0.34994
5.75	10.3125	0.327	0.29792	0.35608	0.124	0.10357	0.14443	0.077	0.060476	0.093524
5.75	11.3125	0.986	0.97872	0.99328	0.827	0.80356	0.85044	0.449	0.41817	0.47983
5.75	11.8125	1	1	1	0.998	0.99523	1.00077	0.789	0.76371	0.81429
6.75	9.8125	1	1	1	0.995	0.99063	0.99937	0.811	0.78673	0.83527
6.75	10.3125	0.993	0/98783	0.99817	0.826	0.80250	0.84950	0.502	0.47101	0.53299
6.75	11.3125	0.39	0.35977	0.42023	0.291	0.26285	0.31915	0.221	0.19528	0.24672
6.75	11.8125	0.955	0.94215	0.96785	0.725	0.69732	0.75268	0.452	0.42115	0.48285
7.25	9.8125	1	1	1	0.995	0.99063	0.99937	0.961	0.94900	0.97300
7.25	10.3125	1	1	1	0.987	0.97998	0.99402	0.788	0.76267	0.81333
7.25	11.3125	0.793	0.76789	0.81811	0.535	0.50409	0.56591	0.372	0.34204	0.40196
7.25	11.8125	0.89	0.87061	0.90939	0.632	0.60211	0.66189	0.441	0.41023	0.47177
6.25	10.8125	0.057	0.04263	0.07137	0.057	0.04263	0.07137	0.101	0.08232	0.11968

* SPower = Simulated Power, LU = Lower Bound, UB = Upper Bound
* Random intercept model, when x-coordinate of vertex is outside occasions for one sample

Simulation results indicate that all the methods and statistics performs reasonably for the random intercept model when the x-coordinate of vertex is outside the occasions.

3.4.4 Random Slope Model, X-coordinate of Vertex Outside Scope of Occasions

For the random slope model (3.2), 1000 data sets are generated with the fixed coefficient parameters $\boldsymbol{\beta}' = (\beta_0, \beta_1, \beta_2) = (3, 2.5, -0.2)$, and covariances for random effects and error $\sigma^2_{\alpha_0} = 1.5, \sigma^2_{\alpha_1} = 1, \sigma_{\alpha_0,\alpha_1} = 0, \sigma^2_e = 0.5$ for sample size 100, 50 and 20. The true model is,

$$y_{ij} = 3 + 2.5t_{ij} - 0.2t_{ij}^2 + \alpha_{0i} + \alpha_{1i}t_{ij} + \epsilon_{ij}, \qquad i = 1, 2, ..., N \quad j = 1, 2, ..., 6.$$

Then,

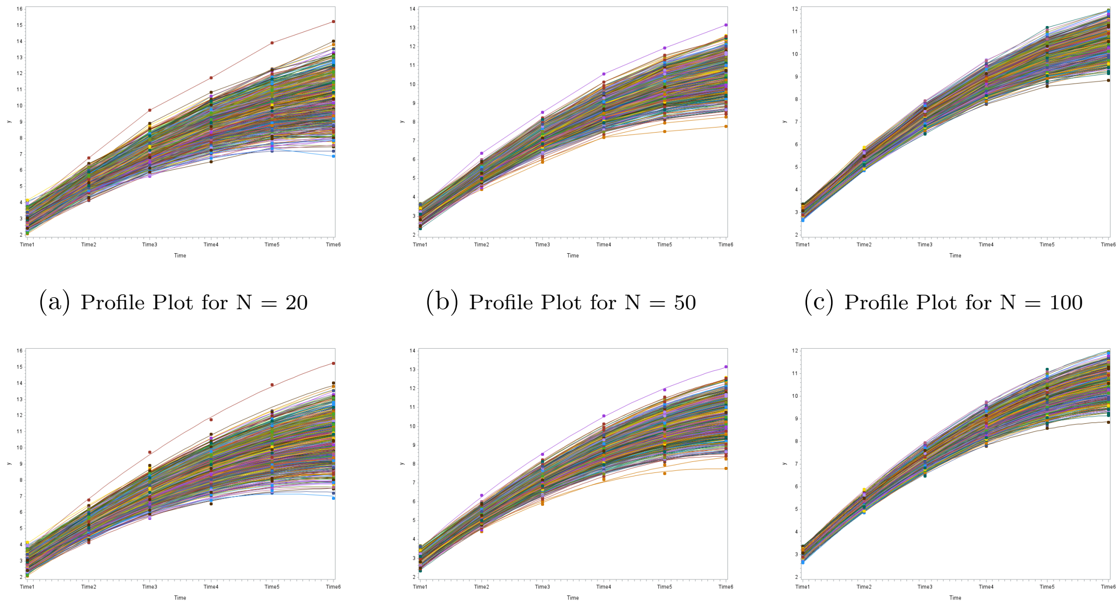
$$E\{y_{ij}\} = 3 + 2.5t_{ij} - 0.2t_{ij}^2, \qquad i = 1, 2, ..., N \quad j = 1, 2, ..., 6.$$

The true vertex of the quadratic growth curve is $\boldsymbol{V}' = (6.25, 10.8125)$, where the x-coordinate of vertex $V_x = 6.25$ is outside the scope of occasions $[0, 5]$. Profile plots and the smoothed profile plots of 1000 data sets are displayed in Figure 3.11 for sample size 20, 50 and 100. As before, the larger the sample size, the narrower the width. From the plots, vertices of quadratic growth curves are out the scope of occasions. The profile plots for random slope model are much wider than those for random intercept model.

The chi-square QQ plots of the response variable $\boldsymbol{y}$ for sample size 20, 50, and 100 are displayed in Figure 3.12. As previous, we conclude that if the sample size is large enough, the assumption of normality is satisfied. However, due to the random linear term in the true model, all three plots seem to reveal a heavy tail.

Simulation Results for Confidence Interval and Confidence Region

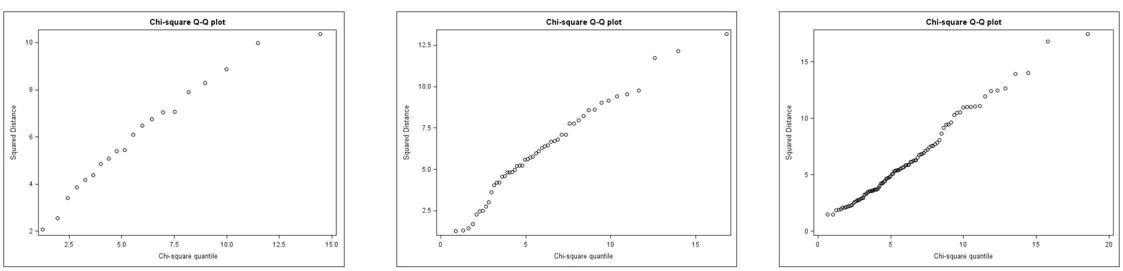
The results of simulation for confidence intervals of x-coordinate with covariance structure UN are displayed in Table 3.16. The results include the empirical coverage as well as lower bound, upper bound and width for the empirical coverage. In Table



(a) Profile Plot for N = 20 (b) Profile Plot for N = 50 (c) Profile Plot for N = 100

(d) Smoothed Plot for N = 20 (e) Smoothed Plot for N = 50 (f) Smoothed Plot for N = 100

Figure 3.11: Profile Plots for Mixed Model with Random Intercept and Slope



(a) $N = 20$ (b) $N = 50$ (c) $N = 100$

Figure 3.12: Chi-square QQ Plot for Random Slope Model

3.16, none of the 18 conditions has the nominal coverage outside the bounds. Width for each method becomes wider as the sample size decreases. Simulation results are reasonable which indicates the validity of both methods.

Table 3.17 displays the simulation results for confidence intervals of the y-coordinate of the estimated vertex with covariance structure UN. In the table, the empirical coverage is shown as well as lower bound, upper bound and width for the empirical

Table 3.16: Confidence Intervals for X-Coordinate

α	Sample Size	Empirical Coverage D	lower bound D	upper bound D	width D	Empirical Coverage G	lower bound G	upper bound G	width G
0.01	100	0.99	0.98190	0.99810	1.728	0.989	0.98051	0.99749	1.757
0.01	50	0.99	0.98190	0.99810	2.461	0.992	0.98475	0.99925	2.546
0.01	20	0.988	0.97913	0.99687	3.965	0.991	0.98331	0.99869	4.390
0.05	100	0.957	0.94443	0.96957	1.315	0.955	0.94215	0.96785	1.328
0.05	50	0.95	0.93649	0.96351	1.873	0.96	0.94785	0.97215	1.910
0.05	20	0.952	0.93875	0.96525	3.018	0.956	0.94329	0.96871	3.193
0.1	100	0.907	0.89189	0.92211	1.104	0.9	0.88439	0.91561	1.111
0.1	50	0.903	0.88760	0.91840	1.572	0.892	0.87585	0.90815	1.594
0.1	20	0.9	0.88439	0.91561	2.533	0.902	0.88653	0.91747	2.634

* Random slope model, when x-coordinate of vertex is outside occasions for one sample
* D represents the delta method, and G represents the gradient method

coverage. None of the 9 conditions has the nominal coverage outside the bound. In conclusion, the method is appropriate given the relatively large sample size.

Table 3.17: Confidence Intervals for Y-Coordinate

α	Sample Size	Empirical Coverage	lower bound	upper bound	width
0.01	100	0.994	0.98771	1.00029	3.371
0.01	50	0.980	0.96860	0.99140	4.782
0.01	20	0.985	0.97510	0.99490	7.539
0.05	100	0.958	0.94557	0.97043	2.566
0.05	50	0.944	0.92975	0.95825	3.64
0.05	20	0.941	0.92640	0.95560	5.738
0.1	100	0.906	0.89082	0.92118	2.154
0.1	50	0.891	0.87479	0.90721	3.055
0.1	20	0.902	0.88653	0.91747	4.816

* Random slope model, when x-coordinate of vertex is outside occasions for one sample

The simulation results for confidence region of the vertex with covariance structure UN are shown in Table 3.18. The table includes the empirical coverage as well as lower bound and upper bound for the empirical coverage. From Table 3.18, none of the 9 conditions had the nominal coverage outside the bounds. Similarly the appropriate results obtained for α level suggest the properness of the approximate chi-squared distribution.

Table 3.18: Confidence Region of the Vertex

α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.987	0.97778	0.99622
0.01	50	0.985	0.97510	0.99490
0.01	20	0.986	0.97643	0.99557
0.5	100	0.947	0.93311	0.96089
0.5	50	0.944	0.92975	0.95825
0.5	20	0.936	0.92083	0.95117
0.1	100	0.903	0.88760	0.91840
0.1	50	0.900	0.88439	0.91561
0.1	20	0.888	0.87159	0.90441

* Random slope model, when x-coordinate of vertex is within occasions for one sample

Results for Power

We investigate power of both direct chi-square test and indirect F test with covariance structure UN. Values for $\boldsymbol{\beta}$ and $\boldsymbol{V}$ under H_0 are given in Table 3.19. Results for comparing the theoretical power and simulated power for chi-square and F test are also displayed in Table 3.19. As illustrated, initially the fixed regression parameters $\boldsymbol{\beta}$ are selected, then the vertex is computed using the $\boldsymbol{\beta}$ to obtain the solution. In the table, the theoretical power for chi-square and F tests are displayed as well as the simulated power with the lower and upper bounds for chi-square and F tests. From the table, six of 15 confidence intervals of simulated power for chi-square test contain the chi-square theoretical power; it may because that the true vertex is far outside the scope of model. Four of 15 confidence intervals of simulated power for F test do not include the F theoretical power. Further more, most of theoretical powers for F test are no greater than the theoretical power for chi-square test but not always.

Table 3.19: Comparison of Theoretical Power and Simulated Power

N	β_0	β_1	β_2	V_{0x}	V_{0y}	SPC	LBC	UBC	ChiSqP	SPF	LBF	UBF	FP
N = 100	2.8	2.8	-0.25	5.6	10.64	✗ **0.897**	**0.878**	**0.916**	**0.860**	✗ **0.952**	**0.939**	**0.965**	**0.966**
	3.2	2.8	-0.22	6.364	12.110	0.826	0.803	0.850	0.849	0.812	0.788	0.836	0.813
	2.5	3.0	-0.25	6	11.5	0.866	0.845	0.887	0.865	0.998	0.995	1	0.998
	2.5	2.6	-0.18	7.222	11.889	✗ **0.812**	**0.788**	**0.836**	**0.859**	✗ **0.996**	**0.992**	**1**	**0.991**
	3.1	2.3	-0.18	6.389	10.447	0.386	0.356	0.416	0.353	0.353	0.323	0.383	0.355
N = 50	2.8	2.8	-0.25	5.6	10.64	0.570	0.539	0.601	0.554	0.747	0.720	0.774	0.727
	3.2	2.8	-0.22	6.364	12.110	✗ **0.504**	**0.473**	**0.535**	**0.541**	0.487	0.456	0.518	0.485
	2.5	3.0	-0.25	6	11.5	0.545	0.514	0.576	0.560	0.914	0.900	0.931	0.915
	2.5	2.6	-0.18	7.222	11.889	0.559	0.528	0.590	0.553	0.845	0.823	0.867	0.837
	3.1	2.3	-0.18	6.389	10.447	✗ **0.222**	**0.196**	**0.248**	**0.194**	0.207	0.182	0.232	0.189
N = 20	2.8	2.8	-0.25	5.6	10.64	✗ **0.133**	**0.112**	**0.154**	**0.250**	0.351	0.321	0.381	0.326
	3.2	2.8	-0.22	6.364	12.110	0.233	0.207	0.259	0.242	✗ **0.247**	**0.220**	**0.274**	**0.205**
	2.5	3.0	-0.25	6	11.5	✗ **0.19**	**0.166**	**0.214**	**0.252**	0.530	0.500	0.561	0.502
	2.5	2.6	-0.18	7.222	11.889	✗ **0.354**	**0.324**	**0.384**	**0.249**	0.439	0.408	0.470	0.410
	3.1	2.3	-0.18	6.389	10.447	✗ **0.143**	**0.121**	**0.165**	**0.104**	✗ **0.125**	**0.105**	**0.146**	**0.099**

* Random slope model with x-coordinate of vertex outside occasions

* SPC = simulated power for chi-square test, LBC = lower bound of simulated power for chi-square test, UBC = upper bound of simulated power for chi-square test, ChiSqP = theoretical chi square power

* SPF = simulated power for F test, LBF = lower bound of simulated power for F test, UBF = upper bound of simulated power for F test, FP = theoretical F power

Power is smaller in Table 3.19 than in Table 3.14, which reveals that adding the random slope term in the model leads to power reduction.

The results of power for the chi-square test by directly selecting the hypothesized vertex points are displayed in Table 3.20. The points to be tested are chosen based on the difference of 0.5 for both x and y-coordinates between the point under the null hypothesis and true vertex. We examine all the pairwise combinations of these points. The table includes the power as well as lower bound and upper bound for the interval around the empirical power. From the table, when V_{0x} is kept fixed, a change of V_{0y} influences the power noticeably. Similarly, if V_{0y} is fixed, a change of V_{0x} extremely affects the results. Finally, when the true vertex point is tested, the result shows the empirical power is nearly equal to the size of the test. The relationship between sample size and power is strongly positive.

Table 3.20: Power Analysis for Chi-square Test ($\alpha = 0.05$, UN)

		$N = 100$			$N = 50$			$N = 20$		
V_{0x}	V_{0y}	SPower	LB	UB	SPower	LB	UB	SPower	LB	UB
5.25	9.8125	0.964	0.95245	0.97555	0.673	0.64392	0.70208	0.137	0.11569	0.15831
5.25	10.3125	0.999	0.99704	1.00096	0.926	0.90978	0.94222	0.341	0.31162	0.37038
5.25	11.3125	1	1	1	1	1	1	0.918	0.90099	0.93501
5.25	11.8125	1	1	1	1	1	1	0.984	0.97622	0.99178
5.75	9.8125	0.236	0.20968	0.26232	0.109	0.08968	0.95383	0.047	0.03388	0.06012
5.75	10.3125	0.297	0.26868	0.32532	0.118	0.09801	0.77684	0.049	0.03562	0.06238
5.75	11.3125	0.981	0.97254	0.98946	0.802	0.77730	0.82670	0.313	0.28426	0.34174
5.75	11.8125	1	1	1	0.974	0.96414	0.98386	0.64	0.61025	0.66975
6.75	9.8125	0.999	0.99704	1.00096	0.939	0.92417	0.95383	0.611	0.58078	0.64122
6.75	10.3125	0.965	0.95361	0.97639	0.75	0.72316	0.77684	0.444	0.41320	0.47480
6.75	11.3125	0.395	0.36470	0.42530	0.246	0.21931	0.27269	0.18	0.15619	0.20381
6.75	11.8125	0.336	0.30672	0.36528	0.205	0.17998	0.23002	0.157	0.13445	0.17955
7.25	9.8125	1	1	1	0.997	0.99361	1.00039	0.884	0.86415	0.90385
7.25	10.3125	1	1	1	0.984	0.97622	0.99178	0.763	0.73664	0.78936
7.25	11.3125	0.965	0.95361	0.97639	0.78	0.75432	0.80568	0.485	0.45402	0.51598
7.25	11.8125	0.859	0.83743	0.88057	0.599	0.56862	0.62938	0.386	0.35583	0.41617
6.25	10.8125	0.053	0.03911	0.06689	0.056	0.04175	0.07025	0.064	0.04883	0.07917

* SPower = Simulated Power, LU = Lower Bound, UB = Upper Bound
* Random slope model, when x-coordinate of vertex is outside occasions for one sample

In conclusion, simulation results illustrate that the methods and statistics perform reasonably for the random slope model when the x-coordinate of the vertex is outside the occasions.

3.4.5 Theoretical Power Analysis for Chi-Square Test

As illustrated in Section 3.3.4, the larger the leverage on the vertex point, the lower the power. Tables 3.21, 3.22 and 3.23 present additional results on the effect of leverage. These tables contain the leverage, non-centrality parameter and power with different x-coordinates of the vertex outside the scope of occasions for sample size 100, 50 and 20 respectively. The variance of the random linear term for quadratic growth curve, $\sigma^2_{\alpha_1}$, changes from 0 to 0.5 in the table; the random slope model with variance for the random linear term equal to zero reduces to the random intercept model. From the tables, if the difference $\Delta \boldsymbol{V}$ of the true vertex and the hypothesized vertex is unchanged, when the $x-$coordinate of vertex is further from the scope of occasions, the leverage on vertex point increases; simultaneously, the non-centrality parameter of the chi-square test decreases and the power decreases. When the variance of the random linear term becomes higher, the power tends to be lower.

For the random intercept model (3.1), we also explore the relationship between the intraclass correlation coefficient (ICC) $\rho = \frac{\sigma^2_{\alpha_0}}{\sigma^2_{\alpha_0}+\sigma^2_e}$ and the theoretical power of the chi-square test. The results are displayed in Table 3.24 and 3.25 for variance of random effect $\sigma^2_{\alpha_0} = 1$, with the $x-$coordinate of vertex within the scope of occasions and the with $x-$coordinate of the vertex outside the range. From the tables, if the difference $\Delta \boldsymbol{V}$ between the true vertex and hypothesized vertex remains the same, increasing the intraclass correlation coefficient results in a larger non-centrality parameter and power. When the variance of error $\sigma^2_e = 0.5$ is unchanged, Table 3.26 and 3.27 show the simulation result for the x-value of the vertex within the scope of time points and for the x-value of the vertex outside the occasions. The results show that keeping $\Delta \boldsymbol{V}$ the same, increasing the intraclass correlation coefficient leads to decreasing of the non-centrality parameter and power.

Table 3.21: Leverage and Theoretical Power for Chi-Square Test ($N = 100$)

		$\sigma^2_{\alpha_1} = 0$			$\sigma^2_{\alpha_1} = 0.25$			$\sigma^2_{\alpha_1} = 0.5$		
V_x	V_y	Leverage	λ_2	Power	Leverage	λ_2	Power	Leverage	λ_2	Power
5.0000	9.2500	0.0071	19.8117	0.98442	0.00438	5.16708	0.51768	0.00439	3.68243	0.38608
5.2083	9.5104	0.0096	19.4672	0.98287	0.00536	4.87036	0.49264	0.00536	3.39589	0.35918
5.4348	9.7935	0.0132	19.1172	0.98114	0.00671	4.57919	0.46739	0.00667	3.12384	0.33335
5.6818	10.1023	0.0185	18.7571	0.97919	0.00853	4.29353	0.44204	0.00845	2.86617	0.30870
5.9524	10.4405	0.0262	18.3821	0.97696	0.01100	4.01325	0.41664	0.01084	2.62254	0.28530
6.2500	10.8125	0.0372	17.9869	0.97437	0.01431	3.73819	0.39127	0.01405	2.39244	0.26316
6.5789	11.2237	0.0534	17.5652	0.97131	0.01876	3.46819	0.36600	0.01834	2.17527	0.24231
6.9444	11.6806	0.0770	17.1103	0.96762	0.02471	3.20311	0.34090	0.02405	1.97034	0.22272
7.3529	12.1912	0.1119	16.6140	0.96310	0.03267	2.94291	0.31606	0.03166	1.77695	0.20435
7.8125	12.7656	0.1641	16.0673	0.95744	0.04332	2.68761	0.29155	0.04181	1.59442	0.18716
8.3333	13.4167	0.2430	15.4599	0.95021	0.05761	2.43737	0.26748	0.05536	1.42211	0.17111
8.9286	14.1607	0.3644	14.7802	0.94081	0.07683	2.19247	0.24396	0.07352	1.25947	0.15614
9.6154	15.0192	0.5550	14.0158	0.92830	0.10282	1.95340	0.22110	0.09799	1.10605	0.14221
10.4167	16.0208	0.8614	13.1537	0.91138	0.13827	1.72081	0.19904	0.13125	0.96149	0.12929
11.3636	17.2045	1.3688	12.1817	0.88807	0.18712	1.49558	0.17793	0.17696	0.82559	0.11733
12.5000	18.6250	2.2390	11.0902	0.85562	0.25548	1.27882	0.15791	0.24073	0.69828	0.10632
13.8889	20.3611	3.7975	9.8752	0.81027	0.35307	1.07187	0.13914	0.33158	0.57966	0.09625
15.6250	22.5313	6.7444	8.5429	0.74749	0.49623	0.87637	0.12178	0.46459	0.46997	0.08710
17.8571	25.3214	12.7130	7.1147	0.66303	0.71402	0.69419	0.10597	0.66667	0.36964	0.07888
20.8333	29.0417	25.9401	5.6324	0.55550	1.06228	0.52748	0.09187	0.98946	0.27926	0.07161

* x-coordinate of vertex outside the occasions

* Scope of occasions [0, 5]

* Parameters $\boldsymbol{\beta}' = (3, 2.5, -0.25)$, $\sigma^2_{\alpha_0} = 1$, $\sigma^2_e = 0.5$, and $\sigma_{\alpha_0\alpha_1} = 0$

3.5 Discussion

This chapter describes several methods for obtaining confidence intervals and a confidence region for the vertex of a quadratic growth curve model, including the x-coordinate of the vertex within and outside the time domain. Initially, the delta method and gradient method were performed for the confidence interval of x-coordinate of the vertex, while delta method and mean response method are equivalent for the y-coordinate. The approximate chi-square distribution with two degrees of freedom was derived for the confidence region analysis and power analysis. Furthermore, in the power and simulation studies, two models, random intercept model and random slope model, were considered. For each model, three different sample

Table 3.22: Leverage and Theoretical Power for Chi-Square Test ($N = 50$)

		$\sigma^2_{\alpha_1} = 0$			$\sigma^2_{\alpha_1} = 0.25$			$\sigma^2_{\alpha_1} = 0.5$		
V_x	V_y	Leverage	λ_2	Power	Leverage	λ_2	Power	Leverage	λ_2	Power
5.0000	9.2500	0.0142	9.90583	0.81154	0.00877	2.58354	0.28154	0.00879	1.84122	0.21044
5.2083	9.5104	0.0192	9.73362	0.80428	0.01073	2.43518	0.26727	0.01071	1.69794	0.19689
5.4348	9.7935	0.0265	9.55858	0.79666	0.01341	2.28960	0.25328	0.01334	1.56192	0.18412
5.6818	10.1023	0.0370	9.37853	0.78856	0.01706	2.14677	0.23958	0.01690	1.43309	0.17212
5.9524	10.4405	0.0523	9.19105	0.77985	0.02199	2.00663	0.22618	0.02169	1.31127	0.16088
6.2500	10.8125	0.0745	8.99343	0.77036	0.02862	1.86910	0.21308	0.02811	1.19622	0.15037
6.5789	11.2237	0.1068	8.78262	0.75987	0.03751	1.73409	0.20030	0.03667	1.08763	0.14055
6.9444	11.6806	0.1541	8.55514	0.74813	0.04942	1.60156	0.18783	0.04810	0.98517	0.13139
7.3529	12.1912	0.2239	8.30699	0.73481	0.06534	1.47146	0.17568	0.06332	0.88847	0.12284
7.8125	12.7656	0.3282	8.03365	0.71950	0.08664	1.34381	0.16387	0.08361	0.79721	0.11486
8.3333	13.4167	0.4860	7.72993	0.70169	0.11521	1.21868	0.15242	0.11071	0.71106	0.10742
8.9286	14.1607	0.7288	7.39010	0.68077	0.15365	1.09624	0.14133	0.14703	0.62974	0.10048
9.6154	15.0192	1.1100	7.00789	0.65596	0.20564	0.97670	0.13064	0.19599	0.55302	0.09401
10.4167	16.0208	1.7229	6.57684	0.62636	0.27653	0.86041	0.12038	0.26250	0.48074	0.08799
11.3636	17.2045	2.7376	6.09084	0.59091	0.37424	0.74779	0.11058	0.35391	0.41279	0.08240
12.5000	18.6250	4.4779	5.54509	0.54855	0.51096	0.63941	0.10130	0.48146	0.34914	0.07722
13.8889	20.3611	7.5951	4.93762	0.49838	0.70614	0.53594	0.09258	0.66315	0.28983	0.07245
15.6250	22.5313	13.4888	4.27147	0.44006	0.99245	0.43819	0.08448	0.92918	0.23499	0.06809
17.8571	25.3214	25.4259	3.55734	0.37438	1.42804	0.34710	0.07705	1.33333	0.18482	0.06415
20.8333	29.0417	51.8801	2.81620	0.30391	2.12455	0.26374	0.07037	1.97893	0.13963	0.06063

* x-coordinate of vertex outside the occasions

* Scope of occasions [0, 5]

* Parameters $\boldsymbol{\beta}' = (3, 2.5, -0.25)$, $\sigma^2_{\alpha_0} = 1$, $\sigma^2_e = 0.5$, and $\sigma_{\alpha_0\alpha_1} = 0$

sizes were selected in order to examine the influence of sample size for all the methods. Three different Type I error rates were chosen as well for the purpose of making the methods more convincing. Given the simulation results, a conclusion can be drawn that all methods described in this study for the confidence region of the vertex of growth curves of 2$^{\text{nd}}$ degree polynomial are applicable for different sample sizes, different Type I error rates and different models. For the power analysis, non-vertex points were tested to show the power of the tests as well as the relationship between confidence region and power. The theoretical power does not conform well to empirical power for the direct chi-square test when the $x-$coordinate of the vertex is far outside the scope of occasions.

Table 3.23: Leverage and Theoretical Power for Chi-Square Test ($N = 20$)

		$\sigma^2_{\alpha_1} = 0$			$\sigma^2_{\alpha_1} = 0.25$			$\sigma^2_{\alpha_1} = 0.5$		
V_x	V_y	Leverage	λ_2	Power	Leverage	λ_2	Power	Leverage	λ_2	Power
5.0000	9.2500	0.036	3.96233	0.41197	0.02191	1.03342	0.13569	0.02197	0.73649	0.10961
5.2083	9.5104	0.048	3.89345	0.40564	0.02682	0.97407	0.13040	0.02679	0.67918	0.10469
5.4348	9.7935	0.066	3.82343	0.39918	0.03354	0.91584	0.12525	0.03336	0.62477	0.10006
5.6818	10.1023	0.093	3.75141	0.39250	0.04266	0.85871	0.12023	0.04226	0.57323	0.09571
5.9524	10.4405	0.131	3.67642	0.38552	0.05498	0.80265	0.11533	0.05422	0.52451	0.09163
6.2500	10.8125	0.186	3.59737	0.37813	0.07155	0.74764	0.11057	0.07026	0.47849	0.08780
6.5789	11.2237	0.267	3.51305	0.37022	0.09378	0.69364	0.10593	0.09169	0.43505	0.08422
6.9444	11.6806	0.385	3.42205	0.36165	0.12354	0.64062	0.10140	0.12025	0.39407	0.08087
7.3529	12.1912	0.560	3.32280	0.35227	0.16334	0.58858	0.09700	0.15831	0.35539	0.07772
7.8125	12.7656	0.820	3.21346	0.34189	0.21661	0.53752	0.09271	0.20904	0.31888	0.07478
8.3333	13.4167	1.215	3.09197	0.33031	0.28803	0.48747	0.08855	0.27678	0.28442	0.07202
8.9286	14.1607	1.822	2.95604	0.31732	0.38413	0.43849	0.08450	0.36759	0.25189	0.06943
9.6154	15.0192	2.775	2.80315	0.30266	0.51411	0.39068	0.08059	0.48997	0.22121	0.06701
10.4167	16.0208	4.307	2.63074	0.28608	0.69133	0.34416	0.07682	0.65626	0.19230	0.06474
11.3636	17.2045	6.844	2.43634	0.26738	0.93561	0.29912	0.07319	0.88478	0.16512	0.06261
12.5000	18.6250	11.195	2.21804	0.24641	1.27740	0.25576	0.06974	1.20366	0.13966	0.06064
13.8889	20.3611	18.988	1.97505	0.22316	1.76536	0.21437	0.06647	1.65789	0.11593	0.05881
15.6250	22.5313	33.722	1.70859	0.19789	2.48113	0.17527	0.06340	2.32295	0.09399	0.05712
17.8571	25.3214	63.565	1.42293	0.17118	3.57009	0.13884	0.06057	3.33333	0.07393	0.05559
20.8333	29.0417	129.700	1.12648	0.14405	5.31138	0.10550	0.05800	4.94732	0.05585	0.05421

* x-coordinate of vertex outside the occasions

* Scope of occasions [0, 5]

* Parameters $\boldsymbol{\beta}' = (3, 2.5, -0.25)$, $\sigma^2_{\alpha_0} = 1$, $\sigma^2_e = 0.5$, and $\sigma_{\alpha_0\alpha_1} = 0$

In the next chapter, we will investigate confidence intervals and test statistics for two independent samples, such as treatment and control groups. A test for difference in the location of the vertices will be developed. Power functions for testing the difference between two vertices from two groups will be derived.

Table 3.24: Intraclass Correlation and Theoretical Power for Chi-Square Test

	$N = 100$		$N = 50$		$N = 20$	
ρ	λ_2	Power	λ_2	Power	λ_2	Power
0.05	3.7118	0.38882	1.8559	0.21183	0.7424	0.11011
0.10	6.7838	0.64079	3.3919	0.3919	1.3568	0.16507
0.15	9.3882	0.78900	4.6941	0.47743	1.8776	0.21389
0.20	11.6450	0.87302	5.8225	0.57041	2.3290	0.25707
0.25	13.6417	0.92136	6.8209	0.64333	2.7283	0.29547
0.30	15.4451	0.95002	7.7225	0.70125	3.0890	0.33003
0.35	17.1088	0.96761	8.5544	0.74809	3.4218	0.36162
0.40	18.6790	0.97875	9.3395	0.78677	3.7358	0.39105
0.45	20.1987	0.98601	10.0994	0.81944	4.0397	0.41906
0.50	21.7119	0.99086	10.8560	0.84768	4.3424	0.44641
0.55	23.2690	0.99415	11.6345	0.87271	4.6538	0.47392
0.60	24.9341	0.99640	12.4670	0.89542	4.9868	0.50255
0.65	26.7977	0.99794	13.3989	0.91652	5.3595	0.53355
0.70	29.0019	0.99894	14.5010	0.93649	5.8004	0.56869
0.75	31.7936	0.99956	15.8968	0.95551	6.3587	0.61072
0.80	35.6581	0.99987	17.8291	0.97326	7.1316	0.66414
0.85	41.7184	0.99998	20.8592	0.98837	8.3437	0.73681
0.90	53.3322	1.00000	26.6661	0.99785	10.6664	0.84099
0.95	87.2698	1.00000	43.6349	0.99999	17.4540	0.97045

* Random intercept model with x-coordinate of vertex within the occasions

* Parameters $\boldsymbol{\beta}' = (2, 8, -1)$, $\boldsymbol{V}' = (4, 18)$, $\Delta\boldsymbol{V}' = (0.05, 0.5)$ and $\sigma^2_{\alpha_0} = 1$

Table 3.25: Intraclass Correlation and Theoretical Power for Chi-Square Test

ρ	$N = 100$ λ_2	Power	$N = 50$ λ_2	Power	$N = 20$ λ_2	Power
0.05	3.0613	0.32738	1.5306	0.18120	0.61225	0.09900
0.10	5.6328	0.55553	2.8164	0.30393	1.12656	0.14406
0.15	7.8239	0.70729	3.9119	0.40734	1.56478	0.18438
0.20	9.7136	0.80342	4.8568	0.49147	1.94272	0.22008
0.25	11.3607	0.86435	5.6803	0.55929	2.27214	0.25160
0.30	12.8097	0.90368	6.4048	0.61406	2.56193	0.27947
0.35	14.0950	0.92970	7.0475	0.65860	2.81899	0.30418
0.40	15.2437	0.94738	7.6218	0.69515	3.04874	0.32618
0.45	16.2775	0.95970	8.1388	0.72546	3.25551	0.34588
0.50	17.2143	0.96850	8.6071	0.75085	3.44286	0.36361
0.55	18.0687	0.97493	9.0343	0.77235	3.61374	0.37966
0.60	18.8535	0.97973	9.4268	0.79076	3.77070	0.39429
0.65	19.5803	0.98339	9.7902	0.80669	3.91607	0.40772
0.70	20.2605	0.98625	10.1303	0.82067	4.05210	0.42019
0.75	20.0969	0.98852	10.4534	0.83316	4.18137	0.43193
0.80	21.5374	0.99039	10.7687	0.84463	4.30747	0.44329
0.85	22.1856	0.99201	11.0928	0.85571	4.43711	0.45485
0.90	22.9416	0.99357	11.4708	0.86776	4.58833	0.46820
0.95	24.2391	0.99559	12.1196	0.88642	4.84783	0.49071

* Random intercept model with x-coordinate of vertex outside the occasions

* Parameters $\boldsymbol{\beta}' = (3, 2.5, -0.25)$, $\boldsymbol{V}' = (5, 9.25)$, $\boldsymbol{\Delta V}' = (0.05, 0.5)$ and $\sigma^2_{\alpha_0} = 1$

Table 3.26: Intraclass Correlation and Theoretical Power for Chi-Square Test

	$N = 100$		$N = 50$		$N = 20$	
ρ	λ_2	Power	λ_2	Power	λ_2	Power
0.05	70.5245	1.00000	35.2623	0.99985	14.1049	0.92988
0.10	61.0543	1.00000	30.5272	0.99934	12.2109	0.88885
0.15	53.1999	1.00000	26.5999	0.99781	10.6400	0.84003
0.20	46.5801	1.00000	23.2901	0.99419	9.3160	0.78569
0.25	40.9252	0.99998	20.4626	0.98700	8.1850	0.72806
0.30	36.0386	0.99988	18.0193	0.97460	7.2077	0.66910
0.35	31.7735	0.99955	15.8867	0.95540	6.3547	0.61042
0.40	28.0185	0.99857	14.0093	0.92819	5.6037	0.55322
0.45	24.6873	0.99613	12.3437	0.89230	4.9375	0.49836
0.50	21.7119	0.99086	10.8560	0.84768	4.3424	0.44641
0.55	19.0383	0.98073	9.5192	0.79491	3.8077	0.39772
0.60	16.6227	0.96318	8.3114	0.73505	3.3245	0.35243
0.65	14.4295	0.93534	7.2148	0.66956	2.8859	0.31060
0.70	12.4294	0.89448	6.2147	0.60015	2.4859	0.27215
0.75	10.5979	0.83850	5.2989	0.52859	2.1196	0.23697
0.80	8.9145	0.76648	4.4573	0.45664	1.7829	0.20491
0.85	7.3621	0.67900	3.6810	0.38595	1.4724	0.17577
0.90	5.9258	0.57838	2.9629	0.31797	1.1852	0.14937
0.95	4.5931	0.46862	2.2966	0.25395	0.9186	0.12550

* Random intercept model with x-coordinate of vertex within the occasions

* Parameters $\boldsymbol{\beta}' = (2, 8, -1)$, $\boldsymbol{V}' = (4, 18)$, $\boldsymbol{\Delta V}' = (0.05, 0.5)$ and $\sigma_e^2 = 0.5$

Table 3.27: Intraclass Correlation and Theoretical Power for Chi-Square Test

	$N = 100$		$N = 50$		$N = 20$	
ρ	λ_2	Power	λ_2	Power	λ_2	Power
0.05	58.1640	1.00000	29.0820	0.99897	11.6328	0.87266
0.10	50.6952	1.00000	25.3476	0.99682	10.1390	0.82102
0.15	44.3353	0.99999	22.1677	0.99197	8.8671	0.76412
0.20	38.8544	0.99995	19.4272	0.98268	7.7709	0.70414
0.25	34.0821	0.99978	17.0410	0.96702	6.8164	0.64303
0.30	29.8892	0.99920	14.9446	0.94322	5.9778	0.58235
0.35	26.1763	0.99751	13.0882	0.90995	5.2353	0.52334
0.40	22.8655	0.99343	11.4328	0.86659	4.5731	0.46686
0.45	19.8948	0.98478	9.9474	0.81326	3.9790	0.41350
0.50	17.2143	0.96850	8.6071	0.75085	3.4429	0.36361
0.55	14.7835	0.94085	7.3917	0.68088	2.9567	0.31738
0.60	12.5690	0.89794	6.2845	0.60529	2.5138	0.27484
0.65	10.5433	0.83650	5.2716	0.52634	2.1087	0.23593
0.70	8.6831	0.75479	4.3415	0.44634	1.7366	0.20053
0.75	6.9690	0.65336	3.4845	0.36753	1.3938	0.16849
0.80	5.3843	0.53557	2.6922	0.29199	1.0769	0.13959
0.85	3.9151	0.40763	1.9576	0.22150	0.7830	0.11363
0.90	2.5491	0.27823	1.2745	0.15752	0.5098	0.09040
0.95	1.2757	0.15763	0.6379	0.10117	0.2551	0.06969

* Random intercept model with x-coordinate of vertex outside the occasions

* Parameters $\boldsymbol{\beta}' = (3, 2.5, -0.25)$, $\boldsymbol{V}' = (5, 9.25)$, $\boldsymbol{\Delta V}' = (0.05, 0.5)$ and $\sigma_e^2 = 0.5$

Chapter 4

A TEST AND CONFIDENCE SET FOR THE DIFFERENCE OF LOCATION OF TWO QUADRATIC GROWTH CURVES

4.1 Models and Methods for Confidence Set

The confidence region of the vertex for one growth curve has been discussed in Chapter 3. In this chapter, we investigate the confidence region for the difference of vertices for growth curves from two independent samples, such as the control and treatment groups. Similar to the one sample case, two growth curve models are explored; one is the second-order random intercept model, and the other is the second-order random slope model. They are defined as follows,

Second-order mixed model with random intercept (random intercept model),

$$y_{ij} = \beta_0^{(\mathrm{mid})} + \beta_0^{(\mathrm{eff})} I_i + \beta_1^{(\mathrm{mid})} t_{ij} + \beta_1^{(\mathrm{eff})} I_i t_{ij} + \beta_2^{(\mathrm{mid})} t_{ij}^2 + \beta_2^{(\mathrm{eff})} I_i t_{ij}^2 + \alpha_{0i} + \epsilon_{ij} \tag{4.1}$$

where

$$I_i = \begin{cases} -1 & \text{if } y_{ij} \text{ comes from control group C,} \\ +1 & \text{if } y_{ij} \text{ comes from treatment group T.} \end{cases}$$

is a dummy variable to indicate the group,

$i = 1, ..., N,\ j = 1, ..., n_i,\ N = N_1 + N_2$ is the total number of individuals, N_1 and N_2 are sample sizes for treatment group and control group, n_i is the number of time measurements for subject i,

β's are fixed regression coefficients,

α_{0i} is a random effect, $\alpha_{0i} \sim N(0, \sigma^2_{\alpha_0})$, $0 < \sigma^2_{\alpha_0} < \infty$, assuming the variance for individual across groups are same, i.e. homogeneous variances,

ϵ_{ij} is the random error term for the i^{th} individual at the j^{th} occasion, $\epsilon_{ij} \sim N(0, \sigma_e^2)$, $0 < \sigma_e^2 < \infty$,

α_{0i} and ϵ_{ij} are independent, $\text{Cov}(\alpha_{0i}, \epsilon_{ij}) = 0$ for all i,

y_{ij} denotes response at j^{th} occasion for the i^{th} individual, and t_{ij} is the time measurement.

From model (4.1), the distinct models for the control and the treatment groups are

$$y_{ij} = \beta_0^{(\text{C})} + \beta_1^{(\text{C})} t_{ij} + \beta_2^{(\text{C})} t_{ij}^2 + \alpha_{0i} + \epsilon_{ij} \qquad \text{for group C,}$$

$$y_{ij} = \beta_0^{(\text{T})} + \beta_1^{(\text{T})} t_{ij} + \beta_2^{(\text{T})} t_{ij}^2 + \alpha_{0i} + \epsilon_{ij} \qquad \text{for group T,}$$

where

$$\beta_k^{(\text{C})} = \beta_k^{(\text{mid})} - \beta_k^{(\text{eff})} \qquad \text{for } k = 0, 1, 2,$$

$$\beta_k^{(\text{T})} = \beta_k^{(\text{mid})} + \beta_k^{(\text{eff})} \qquad \text{for } k = 0, 1, 2.$$

In matrix notation,

$$\boldsymbol{y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{Z}_i \boldsymbol{\alpha}_i + \boldsymbol{\epsilon}_i,$$

where

$\boldsymbol{X}_i$ is the model matrix of regressors for individual i, and

$$\boldsymbol{X}_i^{(\text{C})} = \begin{pmatrix} 1 & 0 & t_{i1} & 0 & t_{i1}^2 & 0 \\ 1 & 0 & t_{i2} & 0 & t_{i2}^2 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & t_{i,n_i} & 0 & t_{i,n_i}^2 & 0 \end{pmatrix}, \ \boldsymbol{X}_i^{(\text{T})} = \begin{pmatrix} 0 & 1 & 0 & t_{i1} & 0 & t_{i1}^2 \\ 0 & 1 & 0 & t_{i2} & 0 & t_{i2}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & t_{i,n_i} & 0 & t_{i,n_i}^2 \end{pmatrix},$$

$$\boldsymbol{X}'_{(N \cdot n_i) \times 6} = \left(\left(\boldsymbol{X}_1^{(\text{C})}\right)', \cdots, \left(\boldsymbol{X}_{N_1}^{(\text{C})}\right)', \left(\boldsymbol{X}_1^{(\text{T})}\right)', \cdots, \left(\boldsymbol{X}_{N_2}^{(\text{T})}\right)' \right)$$

$\boldsymbol{Z}_i$ is a known model matrix, and $\boldsymbol{Z}'_i = (1, 1, \cdots, 1)$,

$\boldsymbol{\beta}$ is an unknown vector of fixed effects, and $\boldsymbol{\beta}' = (\beta_0^{(\text{C})}, \beta_0^{(\text{T})}, \beta_1^{(\text{C})}, \beta_1^{(\text{T})}, \beta_2^{(\text{C})}, \beta_2^{(\text{T})})$,

$\boldsymbol{\alpha}_i$ is an unknown vector of random effects, $\boldsymbol{\alpha}_i = \alpha_{0i}$ and $\text{Cov}(\alpha_{0i}) = \boldsymbol{G}_{(1\times 1)} = \sigma_{\alpha_0}^2$,

$\boldsymbol{\epsilon}_i$ is an unknown vector of random errors for individual i with mean $E(\boldsymbol{\epsilon}_i) = \mathbf{0}$ and covariance $\mathrm{Cov}(\boldsymbol{\epsilon}_i) = \boldsymbol{R}_i$, and $\boldsymbol{R}_{i(n_i \times n_i)} = \sigma_e^2 \boldsymbol{I}_{(n_i \times n_i)}$, $\boldsymbol{\alpha}_i$ and $\boldsymbol{\epsilon}_i$ are independent,

$\boldsymbol{y}_i$ is a known vector of observations for individual i, with mean $E(\boldsymbol{y}_i) = \boldsymbol{X}_i \boldsymbol{\beta}$ and covariance $\Sigma_{\boldsymbol{y}_i} = \boldsymbol{Z}_i \boldsymbol{G} \boldsymbol{Z}_i' + \boldsymbol{R}_i$.

Second-order mixed model with random intercept and random slope (random slope model),

$$y_{ij} = \beta_0^{(\mathrm{mid})} + \beta_0^{(\mathrm{eff})} I_i + \beta_1^{(\mathrm{mid})} t_{ij} + \beta_1^{(\mathrm{eff})} I_i t_{ij} + \beta_2^{(\mathrm{mid})} t_{ij}^2 + \beta_2^{(\mathrm{eff})} I_i t_{ij}^2 + \alpha_{0i} + \alpha_{1i} t_{ij} + \epsilon_{ij}, \quad (4.2)$$

where

$$I_i = \begin{cases} -1 & \text{if } y_{ij} \text{ comes from control group C,} \\ +1 & \text{if } y_{ij} \text{ comes from treatment group T,} \end{cases}$$

is a dummy variable to indicate the group.

α_{0i} and α_{1i} are random effects, $\alpha_{0i} \sim N(0, \sigma_{\alpha_0}^2)$, $\alpha_{1i} \sim N(0, \sigma_{\alpha_1}^2)$, $0 < \sigma_{\alpha_0}^2 < \infty$, assuming the variances for individual are homogeneous.

ϵ_{ij}, β_0's, n, N, y_{ij} and t_{ij} are defined the same as in model (4.1),

α_{0i}, α_{1i} are independent of ϵ_{ij}, $\mathrm{Cov}(\alpha_{0i}, \epsilon_{ij}) = 0$, and $\mathrm{Cov}(\alpha_{1i}, \epsilon_{ij}) = 0$.

From model (4.2), the distinct models for control and treatment group are,

$$y_{ij} = \beta_0^{(\mathrm{C})} + \beta_1^{(\mathrm{C})} t_{ij} + \beta_2^{(\mathrm{C})} t_{ij}^2 + \alpha_{0i} + \alpha_{1i} t_{ij} + \epsilon_{ij} \qquad \text{for group C,}$$

$$y_{ij} = \beta_0^{(\mathrm{T})} + \beta_1^{(\mathrm{T})} t_{ij} + \beta_2^{(\mathrm{T})} t_{ij}^2 + \alpha_{0i} + \alpha_{1i} t_{ij} + \epsilon_{ij} \qquad \text{for group T,}$$

where

$$\beta_k^{(\mathrm{C})} = \beta_k^{(\mathrm{mid})} - \beta_k^{(\mathrm{eff})} \qquad \text{for } k = 0, 1, 2,$$

$$\beta_k^{(\mathrm{T})} = \beta_k^{(\mathrm{mid})} + \beta_k^{(\mathrm{eff})} \qquad \text{for } k = 0, 1, 2.$$

In matrix notation, random slope model (4.2) is,

$$\boldsymbol{y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{Z}_i \boldsymbol{\alpha}_i + \boldsymbol{\epsilon}_i$$

where

$\boldsymbol{X}_i$ is model matrix of regressors for individual i, and

$$\boldsymbol{X}_i^{(\mathrm{C})} = \begin{pmatrix} 1 & 0 & t_{i1} & 0 & t_{i1}^2 & 0 \\ 1 & 0 & t_{i2} & 0 & t_{i2}^2 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & t_{i,n_i} & 0 & t_{i,n_i}^2 & 0 \end{pmatrix}, \boldsymbol{X}_i^{(\mathrm{T})} = \begin{pmatrix} 0 & 1 & 0 & t_{i1} & 0 & t_{i1}^2 \\ 0 & 1 & 0 & t_{i2} & 0 & t_{i2}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & t_{i,n_i} & 0 & t_{i,n_i}^2 \end{pmatrix},$$

$$\boldsymbol{X}'_{(N\cdot n_i)\times 6} = \left(\left(\boldsymbol{X}_1^{(\mathrm{C})}\right)', \cdots, \left(\boldsymbol{X}_{N_1}^{(\mathrm{C})}\right)', \left(\boldsymbol{X}_1^{(\mathrm{T})}\right)', \cdots, \left(\boldsymbol{X}_{N_2}^{(\mathrm{T})}\right)'\right)$$

$\boldsymbol{Z}_i$ is a known model matrix, and $\boldsymbol{Z}_i' = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ t_{i1} & t_{i2} & \cdots & t_{in_i} \end{pmatrix}$,

$\boldsymbol{\beta}$ is an unknown vector of fixed effects, and $\boldsymbol{\beta'} = (\beta_0^{(\mathrm{C})}, \beta_0^{(\mathrm{T})}, \beta_1^{(\mathrm{C})}, \beta_1^{(\mathrm{T})}, \beta_2^{(\mathrm{C})}, \beta_2^{(\mathrm{T})})$,

$\boldsymbol{\alpha}_i$ is an unknown vector of random effect, $\boldsymbol{\alpha}_i' = (\alpha_{0i}, \alpha_{1i})$ and $\mathrm{Cov}(\boldsymbol{\alpha}_i) = \boldsymbol{G}_{(2\times 2)} = \begin{pmatrix} \sigma_{\alpha_0}^2 & \sigma_{\alpha_0\alpha_1} \\ \sigma_{\alpha_0\alpha_1} & \sigma_{\alpha_1} \end{pmatrix}$,

$\boldsymbol{\epsilon}_i$ is an unknown vector of random errors for individual i with mean $E(\boldsymbol{\epsilon}_i) = \mathbf{0}$ and covariance $\mathrm{Cov}(\boldsymbol{\epsilon}_i) = \boldsymbol{R}_i$, and $\boldsymbol{R}_{i(n_i\times n_i)} = \sigma_e^2 \boldsymbol{I}_{(n_i\times n_i)}$, $\boldsymbol{\alpha}_i$ and $\boldsymbol{\epsilon}_i$ are independent,

$\boldsymbol{y}_i$ is a known vector of observations for individual i, with mean $E(\boldsymbol{y}_i) = \boldsymbol{X}_i\boldsymbol{\beta}$ and covariance matrix $\Sigma_{\boldsymbol{y}_i} = \boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i$.

4.1.1 Covariance Matrix for the Difference of Vertices

The second-order random intercept model for control and treatment groups is defined in expression (4.1); the second-order random slope model for control and treatment groups is presented in expression (4.2). For models (4.1) and (4.2), denote $\boldsymbol{b'} = (b_0^{(\mathrm{C})}, b_0^{(\mathrm{T})}, b_1^{(\mathrm{C})}, b_1^{(\mathrm{T})}, b_2^{(\mathrm{C})}, b_2^{(\mathrm{T})})$ as the maximum likelihood estimator (MLE) of the regression coefficients $\boldsymbol{\beta'} = (\beta_0^{(\mathrm{C})}, \beta_0^{(\mathrm{T})}, \beta_1^{(\mathrm{T})}, \beta_1^{(\mathrm{T})}, \beta_2^{(\mathrm{C})}, \beta_2^{(\mathrm{C})})$. Provided that the covariance parameters of random effects are known, $\boldsymbol{b}$ is normally distributed with mean $\boldsymbol{\beta}$ and covariance $\Sigma_{\boldsymbol{b}}$,i.e. $\sqrt{N}(\boldsymbol{b}-\boldsymbol{\beta}) \to N_6(\mathbf{0}, \Omega_{\boldsymbol{b}})$, where $\Omega_{\boldsymbol{b}} = \frac{1}{N}\Sigma_{\boldsymbol{b}}$ and

$$\Sigma_{\boldsymbol{b}} = \begin{pmatrix} \sigma^2_{b_0^{(\mathrm{C})}} & 0 & \sigma_{b_0^{(\mathrm{C})}b_1^{(\mathrm{C})}} & 0 & \sigma_{b_0^{(\mathrm{C})}b_2^{(\mathrm{C})}} & 0 \\ 0 & \sigma^2_{b_0^{(\mathrm{T})}} & 0 & \sigma_{b_0^{(\mathrm{T})}b_1^{(\mathrm{T})}} & 0 & \sigma_{b_0^{(\mathrm{T})}b_2^{(\mathrm{T})}} \\ \sigma_{b_0^{(\mathrm{C})}b_1^{(\mathrm{C})}} & 0 & \sigma^2_{b_1^{(\mathrm{C})}} & 0 & \sigma_{b_1^{(\mathrm{C})}b_2^{(\mathrm{C})}} & 0 \\ 0 & \sigma_{b_0^{(\mathrm{T})}b_1^{(\mathrm{T})}} & 0 & \sigma^2_{b_1^{(\mathrm{T})}} & 0 & \sigma_{b_1^{(\mathrm{T})}b_2^{(\mathrm{T})}} \\ \sigma_{b_0^{(\mathrm{C})}b_2^{(\mathrm{C})}} & 0 & \sigma_{b_1^{(\mathrm{C})}b_2^{(\mathrm{C})}} & 0 & \sigma^2_{b_2^{(\mathrm{C})}} & 0 \\ 0 & \sigma_{b_0^{(\mathrm{T})}b_2^{(\mathrm{T})}} & 0 & \sigma_{b_1^{(\mathrm{T})}b_2^{(\mathrm{T})}} & 0 & \sigma^2_{b_2^{(\mathrm{T})}} \end{pmatrix} = \left(\sum_{i=1}^N \boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1}.$$

If the covariance parameters are unknown, $\boldsymbol{b}$ is approximately normal. The estimated covariance matrix $\hat{\Sigma}_{\boldsymbol{b}}$ is,

$$\hat{\Sigma}_{\boldsymbol{b}} = \begin{pmatrix} \hat{\sigma}^2_{b_0^{(\mathrm{C})}} & 0 & \hat{\sigma}_{b_0^{(\mathrm{C})}b_1^{(\mathrm{C})}} & 0 & \hat{\sigma}_{b_0^{(\mathrm{C})}b_2^{(\mathrm{C})}} & 0 \\ 0 & \hat{\sigma}^2_{b_0^{(\mathrm{T})}} & 0 & \hat{\sigma}_{b_0^{(\mathrm{T})}b_1^{(\mathrm{T})}} & 0 & \hat{\sigma}_{b_0^{(\mathrm{T})}b_2^{(\mathrm{T})}} \\ \hat{\sigma}_{b_0^{(\mathrm{C})}b_1^{(\mathrm{C})}} & 0 & \hat{\sigma}^2_{b_1^{(\mathrm{C})}} & 0 & \hat{\sigma}_{b_1^{(\mathrm{C})}b_2^{(\mathrm{C})}} & 0 \\ 0 & \hat{\sigma}_{b_0^{(\mathrm{T})}b_1^{(\mathrm{T})}} & 0 & \hat{\sigma}^2_{b_1^{(\mathrm{T})}} & 0 & \hat{\sigma}_{b_1^{(\mathrm{T})}b_2^{(\mathrm{T})}} \\ \hat{\sigma}_{b_0^{(\mathrm{C})}b_2^{(\mathrm{C})}} & 0 & \hat{\sigma}_{b_1^{(\mathrm{C})}b_2^{(\mathrm{C})}} & 0 & \hat{\sigma}^2_{b_2^{(\mathrm{C})}} & 0 \\ 0 & \hat{\sigma}_{b_0^{(\mathrm{T})}b_2^{(\mathrm{T})}} & 0 & \hat{\sigma}_{b_1^{(\mathrm{T})}b_2^{(\mathrm{T})}} & 0 & \hat{\sigma}^2_{b_2^{(\mathrm{T})}} \end{pmatrix} = \left(\sum_{i=1}^N \boldsymbol{X}_i'\hat{\Sigma}_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1}.$$

The distinct estimated covariance matrices for control and treatment groups are

$$\hat{\Sigma}_{\boldsymbol{b}^{(\mathrm{T})}} = \begin{pmatrix} \hat{\sigma}^2_{b_0^{(\mathrm{T})}} & \hat{\sigma}_{b_0^{(\mathrm{T})}b_1^{(\mathrm{T})}} & \hat{\sigma}_{b_0^{(\mathrm{T})}b_2^{(\mathrm{T})}} \\ \hat{\sigma}_{b_0^{(\mathrm{T})}b_1^{(\mathrm{T})}} & \hat{\sigma}^2_{b_1^{(\mathrm{T})}} & \hat{\sigma}_{b_1^{(\mathrm{T})}b_2^{(\mathrm{T})}} \\ \hat{\sigma}_{b_0^{(\mathrm{T})}b_2^{(\mathrm{T})}} & \hat{\sigma}_{b_1^{(\mathrm{T})}b_2^{(\mathrm{T})}} & \hat{\sigma}^2_{b_2^{(\mathrm{T})}} \end{pmatrix}, \quad \hat{\Sigma}_{\boldsymbol{b}^{(\mathrm{C})}} = \begin{pmatrix} \hat{\sigma}^2_{b_0^{(\mathrm{C})}} & \hat{\sigma}_{b_0^{(\mathrm{C})}b_1^{(\mathrm{C})}} & \hat{\sigma}_{b_0^{(\mathrm{C})}b_2^{(\mathrm{C})}} \\ \hat{\sigma}_{b_0^{(\mathrm{C})}b_1^{(\mathrm{T})}} & \hat{\sigma}^2_{b_1^{(\mathrm{C})}} & \hat{\sigma}_{b_1^{(\mathrm{C})}b_2^{(\mathrm{C})}} \\ \hat{\sigma}_{b_0^{(\mathrm{C})}b_2^{(\mathrm{T})}} & \hat{\sigma}_{b_1^{(\mathrm{C})}b_2^{(\mathrm{C})}} & \hat{\sigma}^2_{b_2^{(\mathrm{C})}} \end{pmatrix}.$$

Denote $\boldsymbol{V}^{(\mathrm{C})'} = (V_x^{(\mathrm{C})}, V_y^{(\mathrm{C})})$ and $\boldsymbol{V}^{(\mathrm{T})'} = (V_x^{(\mathrm{T})}, V_y^{(\mathrm{T})})$ as the vertices of the control and treatment groups respectively, then $\boldsymbol{V}^{(\mathrm{C})}, \boldsymbol{V}^{(\mathrm{T})}$ and their estimates $\hat{\boldsymbol{V}}^{(\mathrm{C})}, \hat{\boldsymbol{V}}^{(\mathrm{T})}$ are given by,

$$V_x^{(\mathrm{C})} = \frac{-\beta_1^{(\mathrm{C})}}{2\beta_2^{(\mathrm{C})}}, \qquad V_y^{(\mathrm{C})} = \beta_0^{(\mathrm{C})} - \frac{[\beta_1^{(\mathrm{C})}]^2}{4\beta_2^{(\mathrm{C})}}, \qquad V_x^{(\mathrm{T})} = \frac{-\beta_1^{(\mathrm{T})}}{2\beta_2^{(\mathrm{T})}}, \qquad V_y^{(\mathrm{T})} = \beta_0^{(\mathrm{T})} - \frac{[\beta_1^{(\mathrm{T})}]^2}{4\beta_2^{(\mathrm{T})}}.$$

$$\hat{V}_x^{(\mathrm{C})} = \frac{-b_1^{(\mathrm{C})}}{2b_2^{(\mathrm{C})}}, \qquad \hat{V}_y^{(\mathrm{C})} = b_0^{(\mathrm{C})} - \frac{[b_1^{(\mathrm{C})}]^2}{4b_2^{(\mathrm{C})}}, \qquad \hat{V}_x^{(\mathrm{T})} = \frac{-b_1^{(\mathrm{T})}}{2b_2^{(\mathrm{T})}}, \qquad \hat{V}_y^{(\mathrm{T})} = b_0^{(\mathrm{T})} - \frac{[b_1^{(\mathrm{T})}]^2}{4b_2^{(\mathrm{T})}}.$$

For the treatment group, the first-order partial derivative of $\hat{\boldsymbol{V}}^{(\mathrm{T})}$ with respect to $\boldsymbol{\beta}^{(\mathrm{T})} = \left(\beta_0^{(\mathrm{T})}, \beta_1^{(\mathrm{T})}, \beta_2^{(\mathrm{T})}\right)'$ evaluated at $\boldsymbol{\beta}^{(\mathrm{T})} = \boldsymbol{b}^{(\mathrm{T})} = \left(b_0^{(\mathrm{T})}, b_1^{(\mathrm{T})}, b_2^{(\mathrm{T})}\right)'$ is,

$$\frac{\partial \boldsymbol{V}^{(\mathrm{T})}}{\partial \boldsymbol{\beta}^{(\mathrm{T})}}\Big|_{\boldsymbol{\beta}^{(\mathrm{T})}=\boldsymbol{b}^{(\mathrm{T})}} = \hat{\boldsymbol{D}}^{(\mathrm{T})} = \begin{pmatrix} \frac{\partial V_x^{(\mathrm{T})}}{\partial \beta_0^{(\mathrm{T})}} & \frac{\partial V_x^{(\mathrm{T})}}{\partial \beta_1^{(\mathrm{T})}} & \frac{\partial V_x^{(\mathrm{T})}}{\partial \beta_2^{(\mathrm{T})}} \\ \frac{\partial V_y^{(\mathrm{T})}}{\partial \beta_0^{(\mathrm{T})}} & \frac{\partial V_y^{(\mathrm{T})}}{\partial \beta_1^{(\mathrm{T})}} & \frac{\partial V_y^{(\mathrm{T})}}{\partial \beta_2^{(\mathrm{T})}} \end{pmatrix}\Big|_{\boldsymbol{\beta}^{(\mathrm{T})}=\boldsymbol{b}^{(\mathrm{T})}} = \begin{pmatrix} 0 & -\frac{1}{2}[b_2^{(\mathrm{T})}]^{-1} & \frac{1}{2}b_1^{(\mathrm{T})}[b_2^{(\mathrm{T})}]^{-2} \\ 1 & -\frac{1}{2}b_1^{(\mathrm{T})}[b_2^{(\mathrm{T})}]^{-1} & \frac{1}{4}b_1^{(\mathrm{T})2}[b_2^{(\mathrm{T})}]^{-2} \end{pmatrix}.$$

For the control group, the first-order partial derivative of $\hat{\boldsymbol{V}}^{(\mathrm{C})}$ with regard to $\boldsymbol{\beta}^{(\mathrm{C})} = \left(\beta_0^{(\mathrm{C})}, \beta_1^{(\mathrm{C})}, \beta_2^{(\mathrm{C})}\right)'$ evaluated at $\boldsymbol{\beta}^{(\mathrm{C})} = \boldsymbol{b}^{(\mathrm{C})} = \left(b_0^{(\mathrm{C})}, b_1^{(\mathrm{C})}, b_2^{(\mathrm{C})}\right)'$ is,

$$\frac{\partial \boldsymbol{V}^{(\mathrm{C})}}{\partial \boldsymbol{\beta}^{(\mathrm{C})}}\Big|_{\boldsymbol{\beta}^{(\mathrm{C})}=\boldsymbol{b}^{(\mathrm{C})}} = \hat{\boldsymbol{D}}^{(\mathrm{C})} = \begin{pmatrix} \frac{\partial V_x^{(\mathrm{C})}}{\partial \beta_0^{(\mathrm{C})}} & \frac{\partial V_x^{(\mathrm{C})}}{\partial \beta_1^{(\mathrm{C})}} & \frac{\partial V_x^{(\mathrm{C})}}{\partial \beta_2^{(\mathrm{C})}} \\ \frac{\partial V_y^{(\mathrm{C})}}{\partial \beta_0^{(\mathrm{C})}} & \frac{\partial V_y^{(\mathrm{C})}}{\partial \beta_1^{(\mathrm{C})}} & \frac{\partial V_y^{(\mathrm{C})}}{\partial \beta_2^{(\mathrm{C})}} \end{pmatrix}\Big|_{\boldsymbol{\beta}^{(\mathrm{C})}=\boldsymbol{b}^{(\mathrm{C})}} = \begin{pmatrix} 0 & -\frac{1}{2}[b_2^{(\mathrm{C})}]^{-1} & \frac{1}{2}b_1^{(\mathrm{C})}[b_2^{(\mathrm{C})}]^{-2} \\ 1 & -\frac{1}{2}b_1^{(\mathrm{C})}[b_2^{(\mathrm{C})}]^{-1} & \frac{1}{4}b_1^{(\mathrm{C})2}[b_2^{(\mathrm{C})}]^{-2} \end{pmatrix}.$$

4.1.2 Delta Method for the Difference of X- Coordinates

For the one sample case in Section 3.3.1, when the sample size tends to infinity, by the multivariate delta method (2.6), $\hat{\boldsymbol{V}}^{(\mathrm{T})}$, the estimate of $\boldsymbol{V}^{(\mathrm{T})}$ for the treatment group, is approximately multivariate normal with mean $\boldsymbol{V}^{(\mathrm{T})}$ and covariance $\Sigma_{\hat{\boldsymbol{V}}^{(\mathrm{T})}}$, i.e., $\sqrt{N_1}(\hat{\boldsymbol{V}}^{(\mathrm{T})} - \boldsymbol{V}^{(\mathrm{T})}) \xrightarrow{L} MVN_2(0, \Omega_{\hat{\boldsymbol{V}}^{(\mathrm{T})}})$, where $\Omega_{\hat{\boldsymbol{V}}^{(\mathrm{T})}} = \frac{1}{N_1}\Sigma_{\hat{\boldsymbol{V}}^{(\mathrm{T})}}$. Using the estimated covariance $\hat{\Sigma}_{\hat{\boldsymbol{V}}^{(\mathrm{T})}}$,

$$\hat{\Sigma}_{\hat{\boldsymbol{V}}^{(\mathrm{T})}} = \boldsymbol{D}^{(\mathrm{T})}\hat{\Sigma}_{\boldsymbol{b}^{(\mathrm{T})}}\boldsymbol{D}^{(\mathrm{T})'} = \begin{pmatrix} \hat{\sigma}^2_{\hat{V}_x^{(\mathrm{T})}} & \hat{\sigma}_{\hat{V}_x^{(\mathrm{T})}\hat{V}_y^{(\mathrm{T})}} \\ \hat{\sigma}_{\hat{V}_x^{(\mathrm{T})}\hat{V}_y^{(\mathrm{T})}} & \hat{\sigma}^2_{\hat{V}_y^{(\mathrm{T})}} \end{pmatrix}$$

$\hat{V}_x^{(\mathrm{T})}$ is approximately normally distributed with mean $V_x^{(\mathrm{T})}$ and variance $\sigma^2_{\hat{V}_x^{(\mathrm{T})}}$, i.e. $\hat{V}_x^{(\mathrm{T})} \overset{a}{\sim} N\left(V_x^{(\mathrm{T})}, \sigma^2_{\hat{V}_x^{(\mathrm{T})}}\right)$. $\hat{V}_y^{(\mathrm{T})}$ is approximately normally distributed with mean $V_y^{(\mathrm{T})}$ and variance $\sigma^2_{\hat{V}_y^{(\mathrm{T})}}$, i.e. $\hat{V}_y^{(\mathrm{T})} \overset{a}{\sim} N\left(V_y^{(\mathrm{T})}, \sigma^2_{\hat{V}_y^{(\mathrm{T})}}\right)$. Similarly, the estimated vertex for the control group $\sqrt{N_2}(\hat{\boldsymbol{V}}^{(\mathrm{C})} - \boldsymbol{V}^{(\mathrm{C})}) \xrightarrow{L} MVN_2(0, \Omega_{\hat{\boldsymbol{V}}^{(\mathrm{C})}})$ and $\hat{V}_x^{(\mathrm{C})} \overset{a}{\sim} N\left(V_x^{(\mathrm{C})}, \sigma^2_{\hat{V}_x^{(\mathrm{C})}}\right)$, $\hat{V}_y^{(\mathrm{C})} \overset{a}{\sim} N\left(V_y^{(\mathrm{C})}, \sigma^2_{\hat{V}_y^{(\mathrm{C})}}\right)$, where $\Omega_{\hat{\boldsymbol{V}}^{(\mathrm{C})}} = \frac{1}{N_2}\Sigma_{\hat{\boldsymbol{V}}^{(\mathrm{C})}}$ and

$$\hat{\Sigma}_{\hat{\boldsymbol{V}}^{(\mathrm{C})}} = \boldsymbol{D}^{(\mathrm{C})}\hat{\Sigma}_{\boldsymbol{b}^{(\mathrm{C})}}\boldsymbol{D}^{(\mathrm{C})'} = \begin{pmatrix} \hat{\sigma}^2_{\hat{V}_x^{(\mathrm{C})}} & \hat{\sigma}_{\hat{V}_x^{(\mathrm{C})}\hat{V}_y^{(\mathrm{C})}} \\ \hat{\sigma}_{\hat{V}_x^{(\mathrm{C})}\hat{V}_y^{(\mathrm{C})}} & \hat{\sigma}^2_{\hat{V}_y^{(\mathrm{C})}} \end{pmatrix}$$

The summation of two independent normal distribution (Casella and Berger 2002), is normal with the summation of mean and variance. Define the difference between the two vertices of the control and treatment group, $\boldsymbol{V}^{(\text{diff})'} = \boldsymbol{V}^{(\text{T})'} - \boldsymbol{V}^{(\text{C})'} = (V_x^{(\text{diff})}, V_y^{(\text{diff})})$. Suppose that the control group and the treatment groups are independent, the covariance of $\boldsymbol{V}^{(\text{diff})}$ is $\Sigma_{\boldsymbol{V}^{(\text{diff})}} = \Sigma_{\boldsymbol{V}^{(\text{C})}} + \Sigma_{\boldsymbol{V}^{(\text{T})}}$. The distribution for the difference of $x-$coordinates, $V_x^{(\text{diff})} = V_x^{(\text{T})} - V_x^{(\text{C})}$, and the difference of $y-$coordinates, $V_y^{(\text{diff})} = V_y^{(\text{T})} - V_y^{(\text{C})}$, are approximate normal. Namely, $\hat{V}_x^{(\text{diff})} \overset{a}{\sim} N\left((V_x^{(\text{T})} - V_x^{(\text{C})}), (\sigma^2_{\hat{V}_x^{(\text{T})}} + \sigma^2_{\hat{V}_x^{(\text{C})}})\right)$ and $\hat{V}_y^{(\text{diff})} \overset{a}{\sim} N\left((V_y^{(\text{T})} - V_y^{(\text{C})}), (\sigma^2_{\hat{V}_y^{(\text{T})}} + \sigma^2_{\hat{V}_y^{(\text{C})}})\right)$. Therefore, the approximate $(1-\alpha)\%$ confidence interval of $\hat{V}_x^{(\text{diff})}$ is

$$(\hat{V}_x^{(\text{diff})} - Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_x^{(\text{diff})}},\ \hat{V}_x^{(\text{diff})} + Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_x^{(\text{diff})}}).$$

Similarly, the approximate $(1-\alpha)\%$ confidence interval of $\hat{V}_y^{(\text{diff})}$ is

$$(\hat{V}_y^{(\text{diff})} - Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_y^{(\text{diff})}},\ \hat{V}_y^{(\text{diff})} + Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_y^{(\text{diff})}}).$$

4.1.3 Gradient Method for the Difference of X- Coordinates with Common Quadratic Term

When assuming the quadratic terms of two growth curves are the same, $\beta_2^{(\text{C})} = \beta_2^{(\text{T})} = \beta_2$, for mixed models (4.1) and (4.2), the $x-$coordinates of vertices for control and treatment groups become,

$$V_x^{(\text{C})} = \frac{-\beta_1^{(\text{C})}}{2\beta_2^{(\text{C})}} = \frac{-(\beta_1^{(\text{mid})} - \beta_1^{(\text{eff})})}{2\beta_2}, \qquad V_x^{(\text{T})} = \frac{-\beta_1^{(\text{T})}}{2\beta_2^{(\text{T})}} = \frac{-(\beta_1^{(\text{mid})} + \beta_1^{(\text{eff})})}{2\beta_2},$$

$$\hat{V}_x^{(\text{C})} = \frac{-b_1^{(\text{C})}}{2b_2^{(\text{C})}} = \frac{-(b_1^{(\text{mid})} - b_1^{(\text{eff})})}{2b_2}, \qquad \hat{V}_x^{(\text{T})} = \frac{-b_1^{(\text{T})}}{2b_2^{(\text{T})}} = \frac{-(b_1^{(\text{mid})} + b_1^{(\text{eff})})}{2b_2}.$$

For large samples the distribution for the estimator of the difference of the two vertices, $\hat{\boldsymbol{V}}^{(\text{diff})}$, is approximately bivariate normal. Since the estimator vector $\boldsymbol{b}$ is approximately multivariate normal, using a large sample approximation,

$$V_x^{(\text{diff})} \in C(V_x^{(\text{T})} - V_x^{(\text{C})})$$

$$\begin{aligned}
&\Leftrightarrow \frac{(b_1^{(\text{eff})} + b_2 V_x^{(\text{diff})})^2}{\hat{\text{Var}}(b_1^{(\text{eff})}) + 2V_x^{(\text{diff})}\hat{\text{Cov}}(b_1^{(\text{eff})}, b_2) + [V_x^{(\text{diff})}]^2\hat{\text{Var}}(b_2)} \leqslant Z^2_{1-\alpha/2} \\
&\Leftrightarrow (b_1^{(\text{eff})} + b_2 V_x^{(\text{diff})})^2 \leqslant \\
&\left(\hat{\text{Var}}(b_1^{(\text{eff})}) + 2V_x^{(\text{diff})}\hat{\text{Cov}}(b_1^{(\text{eff})}, b_2) + [V_x^{(\text{diff})}]^2\hat{\text{Var}}(b_2)\right) \cdot Z^2_{1-\alpha/2} \\
&\Leftrightarrow A \cdot [V_x^{(\text{diff})}]^2 + B \cdot V_x^{(\text{diff})} + C \leqslant 0,
\end{aligned} \tag{4.3}$$

$$\begin{aligned}
\text{where, } A &= b_2^2 - \hat{\text{Var}}(b_2) \cdot Z^2_{1-\alpha/2} \\
B &= 2b_1^{(\text{eff})} b_2 - 2\hat{\text{Cov}}(b_1^{(\text{eff})}, b_2) \cdot Z^2_{1-\alpha/2} \\
C &= [b_1^{(\text{eff})}]^2 - \hat{\text{Var}}(b_1^{(\text{eff})}) \cdot Z^2_{1-\alpha/2} \ .
\end{aligned}$$

To solve the inequality (4.3), if $A \neq 0$, then $A \cdot x_0^2 + B \cdot x_0 + C$ is a parabola. It has two solutions if the discriminant $D = B^2 - 4AC$ is positive. With regard to the numerical stability concerning small values of $4AC$, we compute either root in two different ways:

$$x_{01} = \begin{cases} \frac{-2C}{B - \sqrt{B^2 - 4AC}} & \text{when } B < 0, \\ \frac{-B - \sqrt{B^2 - 4AC}}{2A} & \text{when } B \geqslant 0. \end{cases} \qquad x_{02} = \begin{cases} \frac{-B + \sqrt{B^2 - 4AC}}{2A} & \text{when } B \leqslant 0, \\ \frac{-2C}{B + \sqrt{B^2 - 4AC}} & \text{when } B > 0. \end{cases}$$

Therefore when $A > 0$ and $D > 0$, this leads to a two-sided confidence interval $[x_{01}, x_{02}]$. When $A < 0$ and $D > 0$, the confidence interval goes to $(-\infty, x_{02}] \bigcup [x_{01}, +\infty)$. In this dissertation, only the first situation is applied; then the confidence interval for the difference of x-coordinates for vertices $\hat{V}_x^{(\text{diff})}$ is $[x_{01}, x_{02}]$.

4.1.4 Mean Response Method for the Difference of Y-Coordinates

The mean response given a set of values of regressors for the OLS model is reviewed in Section 2.6. The variance of $\hat{V}_y^{(\text{diff})}$ is not the same when $V_x^{(\text{T})}$ and $V_x^{(\text{C})}$ are known and when they are estimated, where $\hat{V}_y^{(\text{diff})}$ is treated as a difference of the mean responses $\hat{V}_y^{(\text{C})}$ and $\hat{V}_y^{(\text{T})}$. When the $x-$coordinates of two vertices for the control and treatment groups $\hat{V}_x^{(\text{C})}$ and $\hat{V}_x^{(\text{T})}$ are given, the difference of $y-$coordinate of vertex

$\hat{V}_y^{(\mathrm{diff})}$ can be calculated as,

$$\hat{V}_y^{(\mathrm{diff})} = \hat{V}_y^{(\mathrm{T})} - \hat{V}_y^{(\mathrm{C})} = \left(b_0^{(\mathrm{T})} + b_1^{(\mathrm{T})} \cdot V_x^{(\mathrm{T})} + b_2^{(\mathrm{T})} \cdot V_x^{(\mathrm{T})2}\right) - \left(b_0^{(\mathrm{C})} + b_1^{(\mathrm{C})} \cdot V_x^{(\mathrm{C})} + b_2^{(\mathrm{C})} \cdot V_x^{(\mathrm{C})2}\right) = \boldsymbol{C}'\boldsymbol{b},$$

where $\boldsymbol{C}' = \left(1,\ V_x^{(\mathrm{T})},\ [V_x^{(\mathrm{T})}]^2,\ -1,\ -V_x^{(\mathrm{C})},\ -[V_x^{(\mathrm{C})}]^2\right)$. Then $\sigma^2_{\hat{V}_y^{(\mathrm{diff})}} = \boldsymbol{C}'\Sigma_{\hat{\boldsymbol{b}}}\boldsymbol{C}$, and $\hat{\sigma}^2_{\hat{V}_y^{(\mathrm{diff})}} = \boldsymbol{C}'\hat{\Sigma}_{\hat{\boldsymbol{b}}}\boldsymbol{C}$, and the difference of $y-$coordinate of vertex $\hat{V}_y^{(\mathrm{diff})}$ distributes approximately normally,

$$\frac{\hat{V}_y^{(\mathrm{diff})} - V_y^{(\mathrm{diff})}}{\hat{\sigma}^2_{\hat{V}_y^{(\mathrm{diff})}}} \overset{a}{\sim} N(0,1)\ ,$$

Therefore the $(1-\alpha)\%$ confidence interval of $V_y^{(\mathrm{diff})}$ is

$$(\hat{V}_y^{(\mathrm{diff})} - Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_y^{(\mathrm{diff})}},\ \hat{V}_y^{(\mathrm{diff})} + Z_{1-\alpha/2}\hat{\sigma}_{\hat{V}_y^{(\mathrm{diff})}}).$$

When the $\hat{V}_x^{(\mathrm{C})}$ and $\hat{V}_x^{(\mathrm{T})}$ are estimated, then they are random. For the treatment group, $\hat{V}_y^{((\mathrm{T}))} = b_0^{(\mathrm{T})} + b_1^{(\mathrm{T})}\hat{V}_x^{(\mathrm{T})} + b_2^{(\mathrm{T})}(\hat{V}_x^{(\mathrm{T})})^2$. Substitute $\hat{V}_x^{(\mathrm{T})} = -\frac{b_1^{(\mathrm{T})}}{2b_2^{(\mathrm{T})}}$, then use the delta method, $\hat{\sigma}^2_{\hat{V}_y^{(\mathrm{T})}}$ can be computed. Similarly, $\hat{\sigma}^2_{\hat{V}_y^{(\mathrm{C})}}$ can be obtained using the delta method by substituting $\hat{V}_x^{(\mathrm{C})} = -\frac{b_1^{(\mathrm{C})}}{2b_2^{(\mathrm{C})}}$. Therefore the estimated variance of $V_y^{(\mathrm{diff})}$ for the mean response method, $\hat{\sigma}^2_{\hat{V}_y^{(\mathrm{diff})}} = \hat{\sigma}^2_{\hat{V}_y^{(\mathrm{T})}} + \hat{\sigma}^2_{\hat{V}_y^{(\mathrm{C})}}$, is equivalent to the estimated variance from the delta method if the x-coordinate of the vertices are estimated. Hence the conclusion is drawn that the two methods provide identical results.

4.1.5 Confidence Region for the Difference of Vertices

In order to compute a confidence region for the difference of vertices, the approximate chi-square distribution for a quadratic form is employed. The chi-square distribution with p degrees of freedom is the distribution of a sum of the squares of p independent standard normal random variables. As proved, the estimated difference of the vertices follows an approximate bivariate normal distribution,

$$\hat{\boldsymbol{V}}^{(\mathrm{diff})} \overset{a}{\sim} N_2\left(\boldsymbol{V}^{(\mathrm{diff})},\ \Sigma_{\hat{\boldsymbol{V}}^{(\mathrm{diff})}}\right)$$

where $\Sigma_{\hat{\boldsymbol{V}}^{(\text{diff})}} = \Sigma_{\hat{\boldsymbol{V}}^{(\text{T})}} + \Sigma_{\hat{\boldsymbol{V}}^{(\text{C})}}$. For the bivariate standard normal distribution in vector form, the sum of the squares of two independent standard normal variables is chi-square distribution with two degrees of freedom,

$$\begin{pmatrix} \hat{V}_x^{(\text{diff})} - V_x^{(\text{diff})} \\ \hat{V}_y^{(\text{diff})} - V_y^{(\text{diff})} \end{pmatrix}' \Sigma^{-1}_{\hat{\boldsymbol{V}}^{(\text{diff})}} \begin{pmatrix} \hat{V}_x^{(\text{diff})} - V_x^{(\text{diff})} \\ \hat{V}_y^{(\text{diff})} - V_y^{(\text{diff})} \end{pmatrix} \sim \chi^2_{(2)} \ .$$

As $\hat{\Sigma}_{\hat{\boldsymbol{V}}^{(\text{diff})}} = \hat{\Sigma}_{\hat{\boldsymbol{V}}^{(\text{T})}} + \hat{\Sigma}_{\hat{\boldsymbol{V}}^{(\text{C})}}$ is consistent for $\Sigma_{\hat{\boldsymbol{V}}^{(\text{diff})}}$, an approximate chi-square distribution with two degrees of freedom follows,

$$\begin{pmatrix} \hat{V}_x^{(\text{diff})} - V_x^{(\text{diff})} \\ \hat{V}_y^{(\text{diff})} - V_y^{(\text{diff})} \end{pmatrix}' \hat{\Sigma}^{-1}_{\hat{\boldsymbol{V}}^{(\text{diff})}} \begin{pmatrix} \hat{V}_x^{(\text{diff})} - V_x^{(\text{diff})} \\ \hat{V}_y^{(\text{diff})} - V_y^{(\text{diff})} \end{pmatrix} \overset{a}{\sim} \chi^2_{(2)} \ . \tag{4.4}$$

Therefore the approximate $(1-\alpha)\%$ confidence region for the difference of the vertices for two groups is

$$\begin{pmatrix} \hat{V}_x^{(\text{diff})} - V_x^{(\text{diff})} \\ \hat{V}_y^{(\text{diff})} - V_y^{(\text{diff})} \end{pmatrix}' \hat{\Sigma}^{-1}_{\hat{\boldsymbol{V}}^{(\text{diff})}} \begin{pmatrix} \hat{V}_x^{(\text{diff})} - V_x^{(\text{diff})} \\ \hat{V}_y^{(\text{diff})} - V_y^{(\text{diff})} \end{pmatrix} \leqslant \chi^2_{1-\alpha,2} \ , \tag{4.5}$$

The confidence region is the area of an ellipse since it is an elliptic equation with equality sign in (4.5).

4.2 Analysis of Simulation Results for Confidence Interval and Confidence Region

4.2.1 Simulation Results: Two Quadratic Growth Curves With Common Quadratic Term

In this section, the quadratic terms of two growth curves are assumed to be identical. For the two-sample random intercept model (4.1), two combinations of datasets are generated; they are mixed model with x-coordinate of the vertex within and outside the scope of occasions; each condition contains both the control and treatment group. For the case x-coordinate of vertex within occasions, 1000 data sets are generated with the regression parameters $\boldsymbol{\beta}' = (\beta_0^{(\mathrm{C})}, \beta_0^{(\mathrm{T})}, \beta_1^{(\mathrm{C})}, \beta_1^{(\mathrm{T})}, \beta_2) = (2, 2, 8, 8.1, -1)$ and variances for random effect and error $\sigma^2_{\alpha_0} = 1$, $\sigma^2_e = 0.5$ for sample size 20 and 100. Then the true model for the control group is

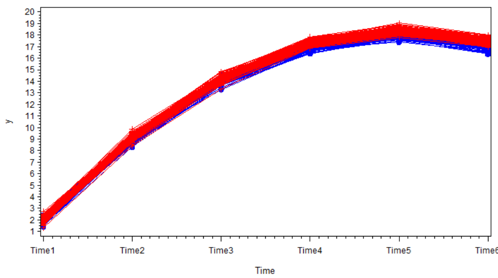
$$y_{ij} = 2 + 8t_{ij} - t_{ij}^2 + \alpha_{0i} + \epsilon_{ij},$$

and the true model for the treatment group is

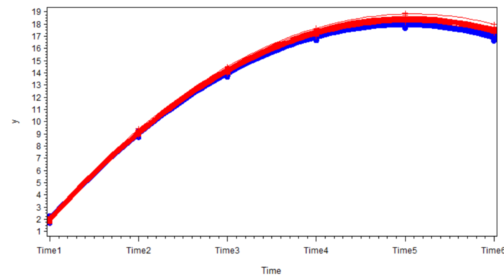
$$y_{ij} = 2 + 8.1t_{ij} - t_{ij}^2 + \alpha_{0i} + \epsilon_{ij}.$$

The true vertex of control group is $\boldsymbol{V}' = (4, 18)$ and the vertex of treatment groups is $\boldsymbol{V}' = (4.05, 18.4025)$, and $\boldsymbol{V}^{(\mathrm{diff})'} = (0.05, 0.4025)$. The time points are $t_{ij} = 0,\ 1,\ 2,\ 3,\ 4,\ 5$, then the x-coordinates for both vertices, $V_x^{(\mathrm{C})} = 4$ and $V_x^{(\mathrm{T})} = 4.05$ are within the scope of occasions, [0,5]. The profile plots and smoothed profile plots are shown in Figure 4.1. For a better display, only 100 datasets are randomly selected from each group; red represents treatment and blue is for control group. The quadratic trend is intuitively suggested from the figure. The red curves are above the blue curves which indicates the y-value of vertex for treatment group is higher than that for control group.

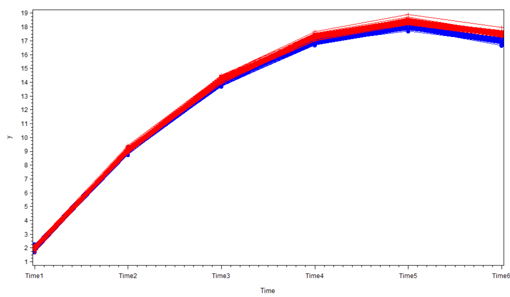
The results of simulations for the confidence interval of difference of x-coordinates are shown in Table 4.1. In this table, symbol D represents the delta method and symbol G represents the gradient method. The results include the empirical coverage as well as lower bound and upper bound for the empirical coverage, where the lower and upper bounds are computed using Wald-type confidence interval. From the columns of the empirical coverage, only 2 conditions had the nominal coverage outside the bounds; they are α level 0.1, sample size 20, for both methods. Bias results from the small sample size; for sample size 100, both methods perform reasonably for the



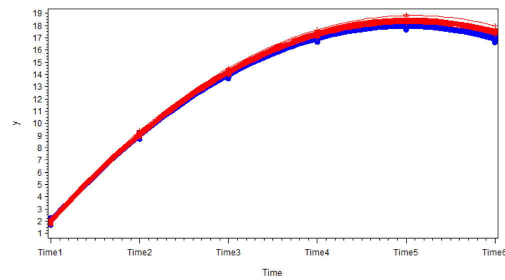
(a) Profile Plot for N = 20



(b) Smoothed Plot for N = 20



(c) Profile Plot for N = 100



(d) Smoothed Plot for N = 100

Figure 4.1: Profile and Smoothed Plots for Random Intercept Model

different α levels. The conclusion is that both the delta method and the gradient method are applicable for the confidence interval of the difference of x-coordinates.

The results of the simulation for confidence intervals for the difference of y-coordinates are displayed in Table 4.2. The table contains the empirical coverage as well as lower bound and upper bound for the empirical coverage. From the table, all 18 conditions had the nominal coverage within the bounds. Therefore, the delta method is appropriate to compare the difference of y-coordinates.

Table 4.3 shows the simulation results of the confidence region for the difference of vertices. The table includes the empirical coverage as well as lower bound and upper bound for the empirical coverage. From the table, one condition has the nominal coverage outside the bounds; it is sample size 20 with α level 0.1. The approximate

Table 4.1: Confidence Interval for Difference of X-Coordinates

α	Sample Size	Empirical Coverage D	lower bound	upper bound	Empirical Coverage G	lower bound	upper bound
0.01	100	0.993	0.98621	0.99979	0.994	0.98771	1.00029
0.01	20	0.991	0.98331	0.99869	0.991	0.98331	0.99869
0.05	100	0.955	0.94215	0.96785	0.955	0.94215	0.96785
0.05	20	0.959	0.94671	0.97129	0.961	0.94900	0.97300
0.1	100	0.91	0.89511	0.92489	0.91	0.89511	0.92489
✗ 0.1	20	**0.931**	**0.91782**	**0.94418**	**0.931**	**0.91782**	**0.94418**

* Random intercept model, when x-coordinates of vertices are within occasions for two samples
* D represents the delta method and G represents the gradient method

Table 4.2: Confidence Interval for Difference of Y-Coordinates

α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.987	0.97778	0.99622
0.01	20	0.986	0.97643	0.99557
0.05	100	0.94	0.92528	0.95472
0.05	20	0.954	0.94102	0.96698
0.1	100	0.893	0.87692	0.90908
0.1	20	0.906	0.89082	0.92118

* Random intercept model, when x-coordinates of vertices are within occasions for two samples

chi-square distribution with two degrees of freedom applied to obtain the confidence region for the difference of vertices seems practicable.

Table 4.3: Confidence Region of Difference of Two Vertices

α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.992	0.98475	0.99925
0.01	20	0.989	0.98051	0.99749
0.05	100	0.956	0.94329	0.96871
0.05	20	0.95	0.93649	0.96351
0.1	100	0.897	0.88119	0.91281
✗ 0.1	20	**0.916**	**0.90157**	**0.93043**

* Random intercept model, when x-coordinates of vertices are within occasions for two samples

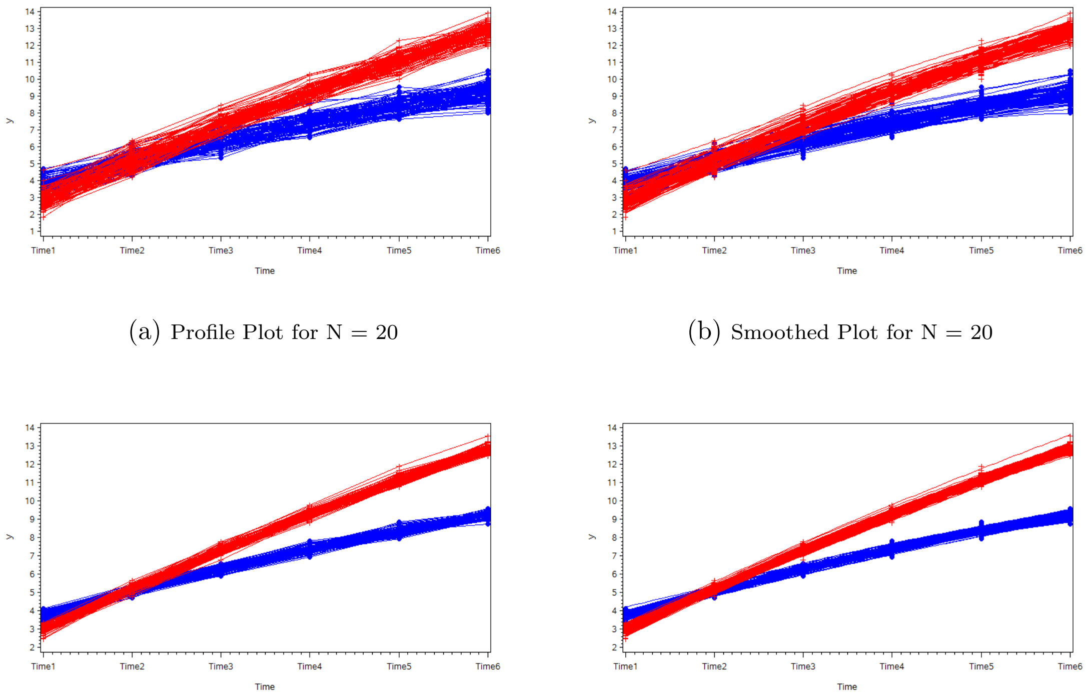
For the random intercept model with x-coordinate of vertex outside the scope of occasions, 1000 data sets are generated with the fixed regression coefficients $\boldsymbol{\beta}' = (\beta_0^{(\mathrm{C})}, \beta_0^{(\mathrm{T})}, \beta_1^{(\mathrm{C})}, \beta_1^{(\mathrm{T})}, \beta_2) = (3.73, 2.98, 1.41, 2.29, -0.062)$, and variances $\sigma^2_{\alpha 0} = 1.44, \sigma^2_e = 2$ with sample size 20 and 100. The time points are $t_{ij} = 0,\ 1,\ 2,\ 3,\ 4,\ 5$. The true models for control and treatment groups are,

$$y_{ij} = 3.73 + 1.41t_{ij} - 0.062t_{ij}^2 + \alpha_{0i} + \epsilon_{ij},$$

$$y_{ij} = 2.98 + 2.29t_{ij} - 0.062t_{ij}^2 + \alpha_{0i} + \epsilon_{ij}.$$

With the vertices for the control and treatment groups, $\boldsymbol{V'} = (11.37, 11.75)$ and $\boldsymbol{V'} = (18.47, 24.13)$, the difference is $\boldsymbol{V}^{(\text{diff})'} = (7.1, 12.38)$. Obviously, the x-coordinates for both vertices, $V_{xc} = 11.37$ and $V_{xt} = 18.47$ are far outside the scope of occasions, [0, 5]. The profile plots and smoothed profile plots are shown in Figure 4.2. For a clearer display, only 100 datasets are randomly chosen from each group; red represents treatment and blue is for control group. The quadratic trend is not intuitive in the figure; actually the vertices are far away from the highest time point in the figure. The red curves are higher than the blue curves which indicates the y-coordinate of vertex for treatment group is greater than that for control group.

The results of simulations for the confidence interval of the difference of x and y coordinates are shown in Table 4.4 and Table 4.5. In Table 4.4, symbol D is for delta method and symbol G represents gradient method. The empirical coverage as well as lower bound and upper bound for the empirical coverage is displayed in the table. For sample size 100, three conditions have the nominal coverage outside the bounds. While for sample size 20, both methods seem inappropriate, especially the gradient method. The smoothed profile plot is flat which shows a small quadratic term. The reason for lower empirical coverage is that the variance of $\hat{\beta}_2$ is too large compared to the square of the β_2 estimate. Table 4.5 contains the empirical coverage as well as the lower bound and upper bound for the empirical coverage. For sample size 100, one case has the nominal coverage within the bounds; it is α level 0.1. However, for sample size 20, none of the three conditions had the nominal coverage within the bounds due to the flatness of the curves. We conclude that delta method and



(a) Profile Plot for N = 20

(b) Smoothed Plot for N = 20

(c) Profile Plot for N = 100

(d) Smoothed Plot for N = 100

Figure 4.2: Profile and Smoothed Plots for Random Intercept Model

gradient method are suitable to compute the difference of x-coordinates for quadratic growth curves with large sample size. If the curve is nearly flat, the delta method is applicable to investigate the confidence interval of the difference of y-coordinates for large sample size.

Table 4.6 shows the simulation results of the confidence region for the difference of vertices. The table includes the empirical coverage as well as lower bound and upper bound for the empirical coverage. Once more, the approximate chi-square distribution with two degrees of freedom performs better for the confidence region of the difference of vertices with larger sample size than that with smaller sample size.

For random slope model (4.2), two cases are generated; they are mixed model with x-coordinate of vertex within and outside the scope of time points. Initially, 1000

Table 4.4: Confidence Interval for Difference of X-Coordinates

α	Sample Size	Empirical Coverage D	lower bound	upper bound	Empirical Coverage G	lower bound	upper bound
✗ 0.01	100	**0.963**	**0.94763**	**0.97837**	**0.878**	**0.85135**	**0.90465**
✗ 0.01	20	**0.913**	**0.89005**	**0.93593**	**0.195**	**0.16274**	**0.22726**
✗ 0.05	100	**0.931**	**0.91529**	**0.94671**	0.942	0.92751	0.95649
✗ 0.05	20	**0.871**	**0.85022**	**0.89178**	**0.368**	**0.33811**	**0.39789**
0.1	100	0.901	0.88546	0.91654	0.893	0.87692	0.90908
✗ 0.1	20	**0.850**	**0.83143**	**0.86857**	**0.46**	**0.43407**	**0.48593**

* Random intercept model, when x-coordinates of vertices are outside occasions for two samples
* D represents the delta method and G represents the gradient method

Table 4.5: Confidence Interval for Difference of Y-Coordinates

α	Sample Size	Empirical Coverage	lower bound	upper bound
✗ 0.01	100	**0.969**	**0.95489**	**0.98311**
✗ 0.01	20	**0.920**	**0.89791**	**0.94209**
✗ 0.05	100	**0.921**	**0.90428**	**0.93772**
✗ 0.05	20	**0.873**	**0.85236**	**0.89364**
0.1	100	0.904	0.88868	0.91932
✗ 0.1	20	**0.856**	**0.83774**	**0.87426**

* Random intercept model, when x-coordinates of vertices are outside occasions for two samples

data sets are generated for random slope model with x-coordinate of vertex within occasions. The parameters are regression coefficients $\boldsymbol{\beta}' = (\beta_0^{(\mathrm{C})}, \beta_0^{(\mathrm{T})}, \beta_1^{(\mathrm{C})}, \beta_1^{(\mathrm{T})}, \beta_2) = (2, 2, 8, 8.1, -1)$ and covariance coefficients $\sigma^2_{\alpha_0} = 1$, $\sigma^2_{\alpha_1} = 0.5$ and $\sigma^2_e = 0.5$. The true models for control and treatment group are

$$y_{ij} = 2 + 8t_{ij} - t_{ij}^2 + \alpha_{0i} + \alpha_{1i}t_{ij} + \epsilon_{ij},$$

Table 4.6: Confidence Region of Difference of Two Vertices

α	Sample Size	Empirical Coverage	lower bound	upper bound
✗ 0.01	100	**0.97**	**0.95611**	**0.98389**
✗ 0.01	20	**0.929**	**0.90809**	**0.94991**
✗ 0.05	100	**0.928**	**0.91198**	**0.94402**
✗ 0.05	20	**0.876**	**0.85557**	**0.89643**
0.1	100	0.890	0.87372	0.90628
✗ 0.1	20	**0.837**	**0.81779**	**0.85621**

* Random intercept model, when x-coordinates of vertices are outside occasions for two samples

and

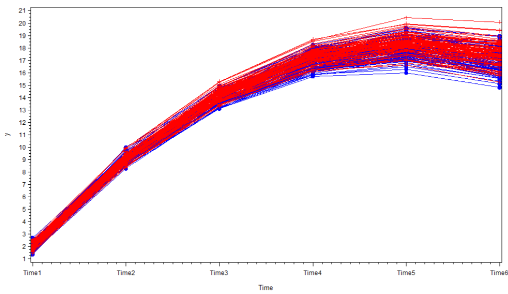
$$y_{ij} = 2 + 8.1t_{ij} - t_{ij}^2 + \alpha_{0i} + \alpha_{1i}t_{ij} + \epsilon_{ij}.$$

With the vertex of the control group $\boldsymbol{V}' = (4, 18)$, and the vertex of the treatment group $\boldsymbol{V}' = (4.05, 18.4025)$, the difference is $\boldsymbol{V}^{(\text{diff})'} = (0.05, 0.4025)$. The time points are $t_{ij} = 0,\ 1,\ 2,\ 3,\ 4,\ 5$, then the x-coordinates for both vertices, $V_x^{(\text{C})} = 4$ and $V_x^{(\text{T})} = 4.05$ are within the scope of occasions, [0,5]. The profile plots and smoothed profile plots are shown in Figure 4.3. For a better display, only 100 datasets are randomly selected from each group; red represents treatment and blue is for control group. The quadratic trend is intuitively suggested from the figure. The red curves are roughly above the blue curves which indicates the y-coordinate of vertex for treatment group maybe larger than that for control group.

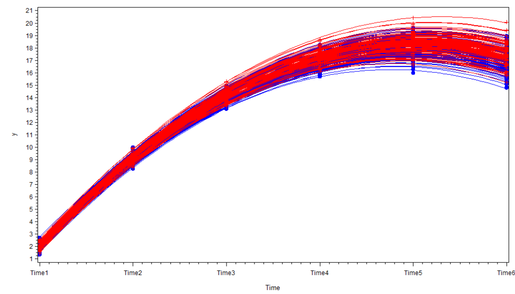
The results of the simulation for confidence intervals of difference of x-values of vertices are shown in Table 4.7. In this table, symbol D represents delta method and symbol G represents gradient method. The results include the empirical coverage as well as lower bound and upper bound for the empirical coverage. From the columns of the empirical coverage, none of the 12 conditions had the nominal coverage outside the bounds. The simulation results under the same method for different sample sizes are close; both the delta method and the gradient method are applicable for the confidence interval of the difference for x-coordinates.

The results of simulations for confidence interval for the difference of y-coordinates of vertices are displayed in Table 4.8. The table contains the empirical coverage as well as lower bound and upper bound for the empirical coverage. From the table, all 18 conditions had the nominal coverage within the bounds. Therefore, the delta method is appropriate to compare the difference of y-values.

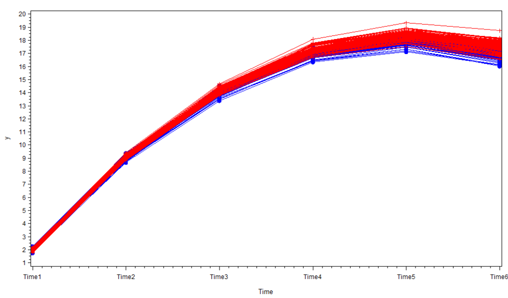
Table 4.9 shows the simulation results of the confidence region for the difference of vertices. The table includes the empirical coverage as well as the lower bound and



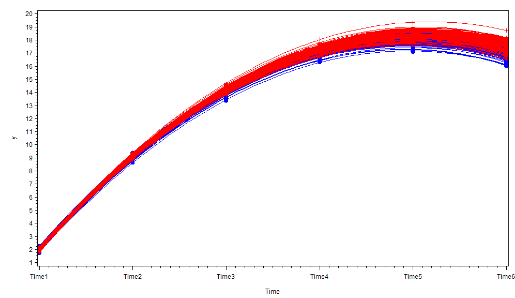
(a) Profile Plot for N = 20



(b) Smoothed Plot for N = 20



(c) Profile Plot for N = 100



(d) Smoothed Plot for N = 100

Figure 4.3: Profile and Smoothed Plots for Random Slope Model

upper bound for the empirical coverage. From the table, only one condition has the nominal coverage outside the bounds; it is sample size 20 with α level 0.1. Hence, the approximate chi-square distribution with two degrees of freedom applied to obtain the confidence region for the difference of vertices is useful.

For the random slope model with x-coordinate of the vertex outside the scope of occasions, 1000 data sets are generated with the regression parameters $\boldsymbol{\beta}' = (\beta_0^{(\mathrm{C})}, \beta_0^{(\mathrm{T})}, \beta_1^{(\mathrm{C})}, \beta_1^{(\mathrm{T})}, \beta_2) = (3.73, 2.98, 1.41, 2.29, -0.062)$ and variances $\sigma^2_{\alpha_0} = 1.44$, $\sigma^2_{\alpha_1} = 0.5$, and $\sigma^2_e = 2$ with sample size 20 and 100. The time measurements are $t_{ij} = 0,\ 1,\ 2,\ 3,\ 4,\ 5$. The true distinct model for control group is

$$y_{ij} = 3.73 + 1.41t_{ij} - 0.062t_{ij}^2 + \alpha_{0i} + \alpha_{1i}t_{ij} + \epsilon_{ij},$$

Table 4.7: Confidence Interval for Difference of X-Coordinates

α	Sample Size	Empirical Coverage D	lower bound	upper bound	Empirical Coverage G	lower bound	upper bound
0.01	100	0.982	0.97117	0.99283	0.982	0.97117	0.99283
0.01	20	0.984	0.97378	0.99422	0.984	0.97378	0.99422
0.05	100	0.944	0.92975	0.95825	0.944	0.92975	0.95825
0.05	20	0.944	0.92975	0.95825	0.944	0.92975	0.95825
0.1	100	0.896	0.88012	0.91188	0.896	0.88012	0.91188
0.1	20	0.895	0.87905	0.91095	0.894	0.87799	0.91001

* Random intercept model, when x-coordinates of vertices are within occasions for two samples
* D represents the delta method and G represents the gradient method

Table 4.8: Confidence Interval for Difference of Y-Coordinates

α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.986	0.97643	0.99557
0.01	20	0.985	0.97510	0.99490
0.05	100	0.949	0.93536	0.96264
0.05	20	0.944	0.92975	0.95825
0.1	100	0.905	0.88975	0.92025
0.1	20	0.893	0.87692	0.90908

* Random slope model, when x-coordinates of vertices are within occasions for two samples

The true model for treatment group is

$$y_{ij} = 2.98 + 2.29t_{ij} - 0.062t_{ij}^2 + \alpha_{0i} + \alpha_{1i}t_{ij} + \epsilon_{ij}.$$

With the vertices for the control and treatment groups, $\boldsymbol{V}' = (11.37, 11.75)$ and $\boldsymbol{V}' = (18.47, 24.13)$, the difference is $\boldsymbol{V}^{(\text{diff})'} = (7.1, 12.38)$. Obviously, the x-coordinates for both vertices, $V_x^{(\text{C})} = 11.37$ and $V_x^{(\text{T})} = 18.47$ are far outside the scope of occasions, [0, 5]. The profile plots and smoothed profile plots are shown in Figure 4.4. For a clearer display, only 100 datasets are randomly chosen from each group; red represents treatment and blue is for control group. Quadratic trend is not intuitive in the figure, actually the vertices are far away from the largest occasion in the figure. The red curves are higher than the blue curves which indicates the y-value of vertex for treatment group is greater than that for control group.

Table 4.9: Confidence Region of Difference of Two Vertices

α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.987	0.97778	0.99622
0.01	20	0.988	0.97913	0.99687
0.05	100	0.946	0.93199	0.96001
0.05	20	0.936	0.92083	0.95117
0.1	100	0.898	0.88226	0.91374
✗ 0.1	20	**0.872**	**0.85462**	**0.88938**

* Random slope model, when x-coordinates of vertices are within occasions for two samples

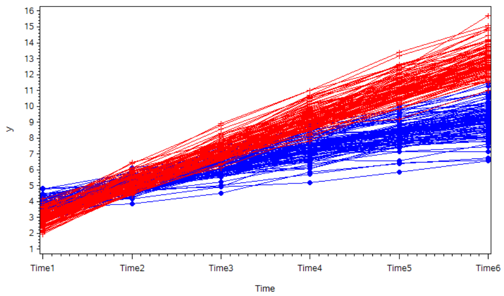
Table 4.10 show the results of simulations for confidence interval of the difference of x-coordinates of vertices. In this table, symbol D represents delta method and symbol G represents gradient method. The empirical coverage as well as lower bound and upper bound for the empirical coverage is displayed in the table. For sample size 100, only one condition has the nominal coverage within the bounds; it is α level 0.1 for the delta method. However, the empirical coverage of other 5 conditions are close to the nominal coverage. For sample size 20, both methods seems inappropriate, especially the gradient method, due to a small quadratic term. Therefore only the delta method is suitable to compute the difference of x-coordinates for quadratic growth curves with large sample size.

Table 4.10: Confidence Interval for Difference of X-Coordinates

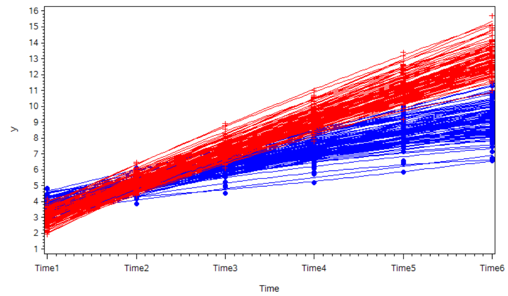
α	Sample Size	Empirical Coverage D	lower bound	upper bound	Empirical Coverage G	lower bound	upper bound
✗ 0.01	100	**0.963**	**0.94763**	**0.97837**	**0.885**	**0.85902**	**0.91098**
✗ 0.01	20	**0.932**	**0.91150**	**0.95250**	**0.185**	**0.15338**	**0.21662**
✗ 0.05	100	**0.928**	**0.91198**	**0.94402**	**0.929**	**0.91308**	**0.94492**
✗ 0.05	20	**0.895**	**0.87600**	**0.91400**	**0.366**	**0.33614**	**0.39586**
✗ 0.1	100	0.896	0.88012	0.91188	0.893	0.87692	0.90908
✗ 0.1	20	**0.867**	**0.84934**	**0.88466**	**0.472**	**0.44603**	**0.49797**

* Random slope model, when x-coordinates of vertices are outside occasions for two samples
* D represents the delta method and G represents the gradient method

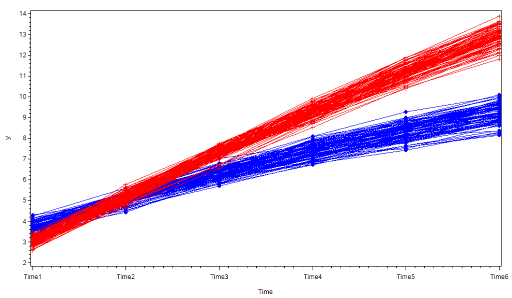
The simulation results for confidence intervals for the difference of y-coordinates are displayed in Table 4.11. The results contains the empirical coverage as well as



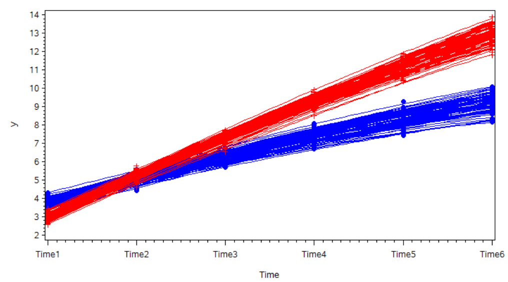
(a) Profile Plot for N = 20



(b) Smoothed Plot for N = 20



(c) Profile Plot for N = 100



(d) Smoothed Plot for N = 100

Figure 4.4: Profile and Smoothed Plots for Random Slope Model

lower bound and upper bound for the empirical coverage. For sample size 100, only one condition has the nominal coverage inside the bounds. However, the empirical coverage of other conditions are slightly lower than the nominal coverage. For sample size 20, none of condition has the nominal coverage within the bounds due to the flatness of the curves. To sum the conclusion, the method is applicable to compute the confidence interval of difference of y-values of vertices for large sample size.

Table 4.12 shows the simulation results of the confidence region for the difference of vertices. The table includes the empirical coverage as well as lower bound and upper bound for the empirical coverage. From the table, the approximate chi-square distribution with two degrees of freedom performs better for the confidence region

Table 4.11: Confidence Interval for Difference of Y-Coordinates

α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	**0.968**	**0.95367**	**0.98233**
✗ 0.01	20	**0.948**	**0.92992**	**0.96608**
✗ 0.05	100	**0.934**	**0.91861**	**0.94939**
✗ 0.05	20	**0.909**	**0.89117**	**0.92683**
✗ 0.1	100	0.91	0.89511	0.92489
✗ 0.1	20	**0.879**	**0.86204**	**0.89596**

* Random slope model, when x-coordinates of vertices are outside occasions for two samples

with larger sample size than that with smaller sample size, but the method does not perform well if the location is far outside the scope of occasions which happens when the curve is nearly flat.

Table 4.12: Confidence Region of Difference of Two Vertices

α	Sample Size	Empirical Coverage	lower bound	upper bound
✗ 0.01	100	**0.957**	**0.94048**	**0.97352**
✗ 0.01	20	**0.922**	**0.90016**	**0.94384**
✗ 0.05	100	**0.924**	**0.90758**	**0.94024**
✗ 0.05	20	**0.861**	**0.83956**	**0.88244**
✗ 0.1	100	**0.877**	**0.85991**	**0.89409**
✗ 0.1	20	**0.825**	**0.80523**	**0.84477**

* Random slope model, when x-coordinates of vertices are outside occasions for two samples

4.2.2 Simulation Results: Two Quadratic Growth Curves With Heterogeneous Quadratic Terms

Quadratic growth curves without parameter restrictions for the control and treatment groups will be investigated in this section. As well as the quadratic growth curves with distinct quadratic term, two situations are discussed: random intercept models with x-values of vertex within and outside the scope of time measurements.

For x-coordinate within the scope of the time measurements, 1000 data sets are generated with the regression parameters $\boldsymbol{\beta}' = (\beta_0^{(C)}, \beta_0^{(T)}, \beta_1^{(C)}, \beta_1^{(T)}, \beta_2^{(C)}, \beta_2^{(T)}) =$

$(2, 2.1, 8, 8.1, -1, -0.9)$ and variances parameters $\sigma^2_{\alpha_0} = 1$ and $\sigma^2_e = 0.5$, with sample size 20 and 100. The true model for the control group is

$$y_{ij} = 2 + 8t_{ij} - t^2_{ij} + \alpha_{0i} + \epsilon_{ij}.$$

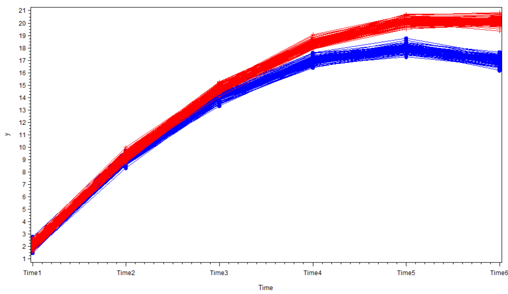
And the true model for the treatment group is

$$y_{ij} = 2.1 + 8.1t_{ij} - 0.9t^2_{ij} + \alpha_{0i} + \epsilon_{ij}.$$

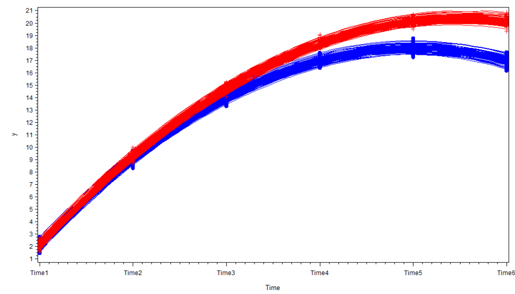
The true vertices for control and treatment group are $\boldsymbol{V}' = (4, 18)$ and $\boldsymbol{V}' = (4.5, 20.325)$; the difference between them is $\boldsymbol{V}^{(\text{diff})'} = (0.5, 2.325)$. The time points are $t_{ij} = 0,\ 1,\ 2,\ 3,\ 4,\ 5$, obviously, the x-coordinates for both vertices, $V_x^{(\text{C})} = 4$ and $V_x^{(\text{T})} = 4.5$ are within the scope of occasions, $[0,\ 5]$. The profile plots and smoothed profile plots are displayed in Figure 4.5. For a clearer view, only 100 datasets are randomly selected from each group; color red represents treatment and blue is for control group; quadratic trend is intuitive in the figure. The blue curves are below the red curves which indicates that the y-coordinate of vertex for control group is smaller than that for the treatment group.

The results of simulations for the confidence interval of difference of x and y-coordinates of vertices are shown in Table 4.13. The empirical coverage as well as lower bound and upper bound for the empirical coverage are displayed in the tables. For the x-coordinate, only the delta method is examined since the gradient method is not applicable when the quadratic terms of two groups are different; all six conditions had the nominal coverage within the bounds. For the y-coordinate, one condition has the nominal coverage outside the bounds; it is α level 0.1 for sample size 100. We conclude that the the delta method is suitable for computing the difference of x and y-coordinates.

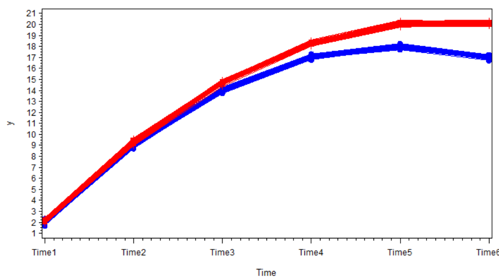
Table 4.14 displays the simulation results of the confidence region for the difference of two vertices. The table includes the empirical coverage as well as lower bound and



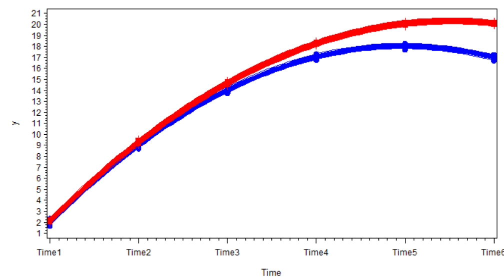
(a) Profile Plot for N = 20



(b) Smoothed Plot for N = 20



(c) Profile Plot for N = 100



(d) Smoothed Plot for N = 100

Figure 4.5: Profile and Smoothed Plots for Random Intercept Model

upper bound for the empirical coverage. The approximate chi-square distribution with two degrees of freedom is valid for the confidence region of the difference of vertices as shown in the table for different α level and sample size, since all the bounds of empirical coverage contain the nominal value.

For $x-$coordinate outside the scope of the time measurements, 1000 data sets are generated with the fixed regression parameters $\boldsymbol{\beta}' = (\beta_0^{(\mathrm{C})}, \beta_0^{(\mathrm{T})}, \beta_1^{(\mathrm{C})}, \beta_1^{(\mathrm{T})}, \beta_2^{(\mathrm{C})}, \beta_2^{(\mathrm{T})}) = (4, 3, 1.5, 2.5, -0.15, -0.2)$ and variances $\sigma^2_{\alpha_0} = 1.5$ and $\sigma^2_e = 0.5$, with sample size 20 and 100.

The true distinct model for the control group is

$$y_{ij} = 4 + 1.5t_{ij} - 0.15t_{ij}^2 + \alpha_{0i} + \epsilon_{ij},$$

Table 4.13: Confidence Interval for Difference of X and Y-Coordinates

		x-coordinate			y-coordinate		
α	Sample Size	Empirical Coverage	lower bound	upper bound	Empirical Coverage	lower bound	upper bound
0.01	100	0.991	0.98331	0.99869	0.991	0.98331	0.99869
0.01	20	0.987	0.9778	0.99622	0.994	0.98771	1.00000
0.05	100	0.947	0.93311	0.96089	0.955	0.94215	0.96785
0.05	20	0.95	0.93649	0.96351	0.953	0.93988	0.96612
✗ 0.1	100	0.899	0.88333	0.91467	**0.915**	**0.90049**	**0.92951**
0.1	20	0.898	0.88226	0.91374	0.9	0.88439	0.91561

* Random intercept model, when x-coordinates of vertices are within occasions for two samples

Table 4.14: Confidence Region of Difference of Two Vertices

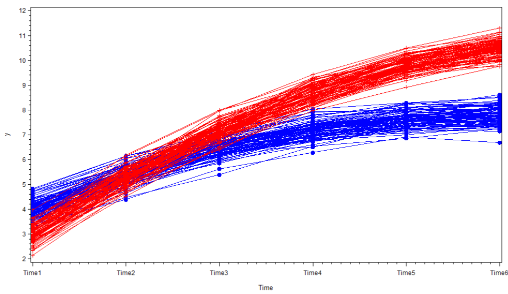
α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.995	0.98926	1.00074
0.01	20	0.992	0.98475	0.99925
0.05	100	0.953	0.93988	0.96612
0.05	20	0.947	0.93311	0.96089
0.1	100	0.901	0.88546	0.91654
0.1	20	0.893	0.87692	0.90908

* Mixed model with random intercept, when x-coordinates of vertices are within occasions for two samples

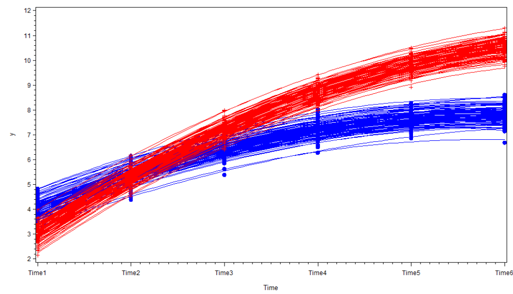
And the true model for the treatment group is

$$y_{ij} = 3 + 2.5t_{ij} - 0.2t_{ij}^2 + \alpha_{0i} + \epsilon_{ij}.$$

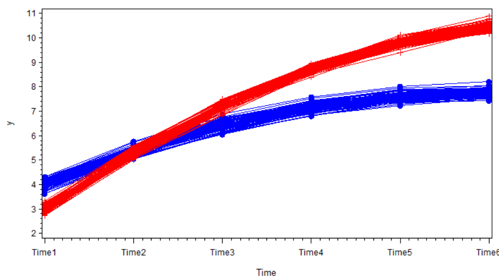
The true vertices for control and treatment group are $\boldsymbol{V}' = (5, 7.75)$ and $\boldsymbol{V}' = (6.25, 10.8125)$; the difference between them is $\boldsymbol{V}^{(\text{diff})'} = (1.25, 3.0625)$. The time points are $t_{ij} = 0,\ 1,\ 2,\ 3,\ 4,\ 5$, then the x-coordinate of the vertex for control group, $V_{xc} = 5$, is the largest time measurement while the x-coordinate of vertex for treatment group, $V_{xt} = 6.25$, is slightly outside the scope of occasion, $[0, 5]$. The profile plots and smoothed profile plots are displayed in Figure 4.6. For a better view, only 100 datasets are randomly selected out of 1000 datasets from each group; color red is for treatment while color blue represents control group; quadratic trend is intuitive in the figure. The blue curves are below the red curves which indicates that the y-coordinate of vertex for control group is much less than the treatment group.



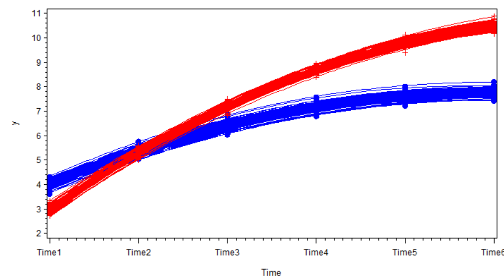
(a) Profile Plot for N = 20



(b) Smoothed Plot for N = 20



(c) Profile Plot for N = 100



(d) Smoothed Plot for N = 100

Figure 4.6: Profile and Smoothed Plots for Random Intercept Model

The results of simulations for the confidence interval of the difference of x and y-coordinates of vertices are displayed in Table 4.15. The empirical coverage as well as lower bound and upper bound for the empirical coverage are displayed in the table for the delta method. For the x-coodinate, five conditions have the nominal coverage higher than the upper bound, the only exception is α level 0.01 for sample size 100. For the y-coordinate, only one condition have the nominal coverage outside the bounds; it is α level 0.01 for sample size 100. We conclude that the delta method is applicable for the confidence interval of the difference of y-coordinates.

Table 4.16 shows the simulation results of the confidence region for the difference of two vertices from two groups. The table includes the empirical coverage as well as lower bound and upper bound for the empirical coverage. Five conditions have

Table 4.15: Confidence Interval for Difference of X and Y-Coordinates

		x-coordinate			y-coordinate		
α	Sample Size	Empirical Coverage	lower bound	upper bound	Empirical Coverage	lower bound	upper bound
✗ 0.01	100	0.995	0.98926	1.00074	**0.996**	**0.99086**	**1.00114**
✗ 0.01	20	**1**	**1**	**1**	0.991	0.98331	0.99869
✗ 0.05	100	**0.972**	**0.96177**	**0.98223**	0.952	0.93875	0.96525
✗ 0.05	20	**0.982**	**0.97376**	**0.99024**	0.949	0.93536	0.96246
✗ 0.1	100	**0.922**	**0.90805**	**0.93595**	0.908	0.89297	0.92303
✗ 0.1	20	**0.933**	**0.91999**	**0.94601**	0.902	0.88653	0.91747

* Random intercept model, when x-coordinates of vertices are outside occasions for two samples

the nominal coverage within the bounds in the table. Therefore the approximate chi-square distribution with two degrees of freedom is valid for the confidence region of the difference of vertices given that the growth curves have a fairly steep rate of change.

Table 4.16: Confidence Region of Difference of Two Vertices

α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.993	0.98621	0.99979
0.01	20	0.991	0.98331	0.99869
0.05	100	0.96	0.94785	0.97215
0.05	20	0.956	0.94329	0.96871
✗ 0.1	100	**0.917**	**0.90265**	**0.93135**
0.1	20	0.909	0.89404	0.92396

* Random intercept model, when x-coordinates of vertices are outside occasions for two samples

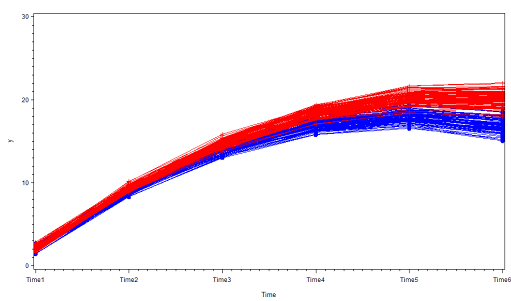
For the random slope model (4.2), two situations are discussed; they are x-values of vertices within and outside the scope of time measurements. For $x-$coordinate within the scope of the time measurements, 1000 data sets are generated with the fixed regression parameters $\boldsymbol{\beta}' = (\beta_0^{(C)}, \beta_0^{(T)}, \beta_1^{(C)}, \beta_1^{(T)}, \beta_2^{(C)}, \beta_2^{(T)}) = (2, 2.1, 8, 8.1, -1, -0.9)$ and variances $\sigma^2_{\alpha_0} = 1$, $\sigma^2_{\alpha_1} = 0.5$ and $\sigma^2_e = 0.5$, with sample size 20 and 100. The true distinct models for the control and the treatment groups are

$$y_{ij} = 2 + 8t_{ij} - t_{ij}^2 + \alpha_{0i} + \alpha_{1i}t_{ij} + \epsilon_{ij},$$

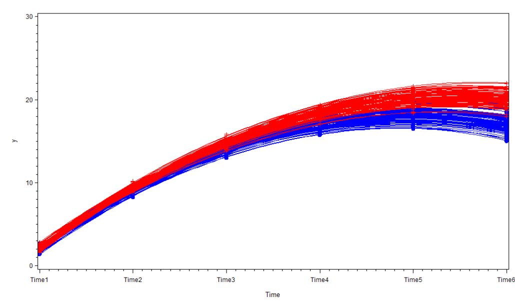
and

$$y_{ij} = 2.1 + 8.1t_{ij} - 0.9t_{ij}^2 + \alpha_{0i} + \alpha_{1i}t_{ij} + \epsilon_{ij}.$$

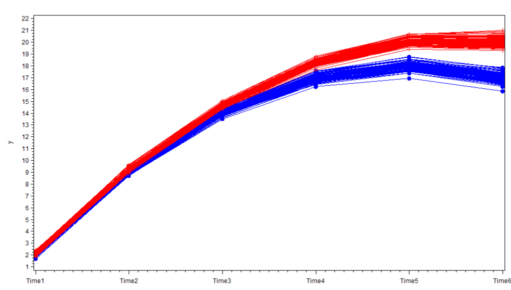
The true vertices for control and treatment groups are $\boldsymbol{V}' = (4, 18)$ and $\boldsymbol{V}' = (4.5, 20.325)$; the difference between them is $\boldsymbol{V}^{(\text{diff})'} = (0.5, 2.325)$. The time points are $t_{ij} = 0,\ 1,\ 2,\ 3,\ 4,\ 5$, then the x-coordinates for both vertices, $V_x^{(\text{C})} = 4$ and $V_x^{(\text{T})} = 4.5$ are within the scope of occasions, $[0, 5]$. The profile plots and smoothed profile plots are displayed in Figure 4.7. For a clearer view, only 100 datasets are randomly selected from each group; color red represents treatment and blue is for control group; quadratic trend is intuitive in the figure. The blue curves are below the red curves which indicates that the y-coordinate of vertex for control group is smaller than that for treatment group.



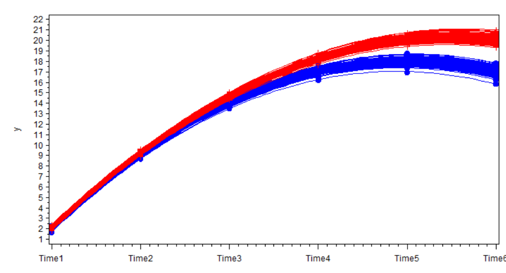
(a) Profile Plot for N = 20



(b) Smoothed Plot for N = 20



(c) Profile Plot for N = 100



(d) Smoothed Plot for N = 100

Figure 4.7: Profile and Smoothed Plots for Random Slope Model

The results of simulations for confidence interval of the difference of x and y-coordinates of vertices are shown in Table 4.17. The empirical coverage as well as lower bound and upper bound for the empirical coverage are displayed in the table. For the $x-$coordinate of the vertex, the delta method is examined. All six conditions have the nominal coverage within the bounds in the table. For y-coordinate of the vertex, none of 12 conditions have the nominal coverage outside the bounds. Therefore the delta method is suitable for computing confidence intervals for the difference of x and y-coordinates of vertices when the vertices are inside the scope of occasions.

Table 4.17: Confidence Interval for Difference of X and Y-Coordinates

		x-coordinate			y-coordinate		
α	Sample Size	Empirical Coverage	lower bound	upper bound	Empirical Coverage	lower bound	upper bound
0.01	100	0.989	0.98051	0.99749	0.989	0.98051	0.99749
0.01	20	0.991	0.98331	0.99869	0.983	0.97247	0.99353
0.05	100	0.954	0.94102	0.96698	0.943	0.92863	0.95737
0.05	20	0.957	0.94443	0.96957	0.94	0.92528	0.95471
0.1	100	0.901	0.88546	0.91654	0.894	0.87799	0.91001
0.1	20	0.903	0.88760	0.91840	0.893	0.87692	0.90908

* Random slope model, when x-coordinates of vertices are within occasions for two samples

Table 4.18 displays the simulation results of the confidence region for the difference of two vertices. The table includes the empirical coverage as well as lower and upper bounds for the empirical coverage. Since the bounds of empirical coverage contain the nominal value for all conditions, the approximate chi-square distribution with two degrees of freedom is valid for the confidence region of the difference of vertices as shown in the table.

For $x-$coordinate outside the scope of occasions, 1000 data sets are generated with the fixed regression parameters $\beta_0^{(C)}, \beta_0^{(T)}, \beta_1^{(C)}, \beta_1^{(T)}, \beta_2^{(C)}\beta_2^{(T)}$ equal to 4, 3, 1.5, 2.5, -0.15, -0.2, and variances $\sigma^2_{\alpha_0} = 1.5$, $\sigma^2_{\alpha_1} = 0.5$ and $\sigma^2_e = 0.5$, with sample size 20 and 100. The distinct models for control and treatment groups are

$$y_{ij} = 4 + 1.5t_{ij} - 0.15t_{ij}^2 + \alpha_{0i} + \alpha_{1i}t_{ij} + \epsilon_{ij},$$

and

$$y_{ij} = 3 + 2.5t_{ij} - 0.2t_{ij}^2 + \alpha_{0i} + \alpha_{1i}t_{ij} + \epsilon_{ij}.$$

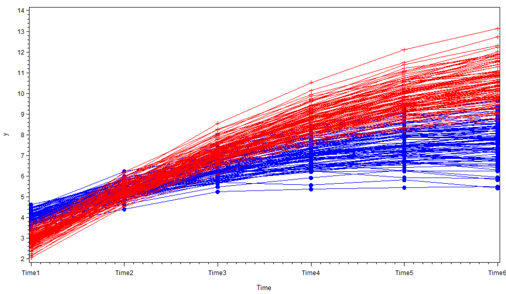
$\boldsymbol{V}' = (5, 7.75)$ and $\boldsymbol{V}' = (6.25, 10.8125)$ are the vertices for control and treatment group; the difference between them is $\boldsymbol{V}^{(\text{diff})'} = (1.25, 3.0625)$. The time points are $t_{ij} = 0,\ 1,\ 2,\ 3,\ 4,\ 5$, then the x-coordinate of the vertex for control group, $V_x^{(\text{C})} = 5$, is the largest time measurement while the x-coordinate of vertex for treatment group, $V_x^{(\text{T})} = 6.25$, is slightly outside the scope of occasion, $[0, 5]$. The profile plots and smoothed profile plots are displayed in Figure 4.8. For a better view, only 100 datasets are randomly selected out of 1000 datasets from each group; color red is for treatment while color blue represents control group; quadratic trend is intuitive in the figure. Most of the blue curves are below the red curves which indicates that the y-coordinate of vertex for control group is much less than that for treatment group.

Table 4.18: Confidence Region of Difference of Two Vertices

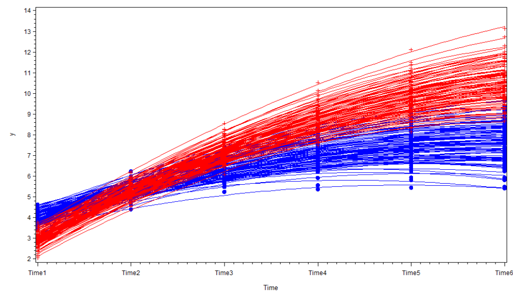
α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.988	0.97913	0.99687
0.01	20	0.989	0.98051	0.99749
0.05	100	0.948	0.93424	0.96176
0.05	20	0.938	0.92305	0.95295
0.1	100	0.902	0.88653	0.91747
0.1	20	0.897	0.88119	0.91281

* Random slope model, when x-coordinates of vertices are within occasions for two samples

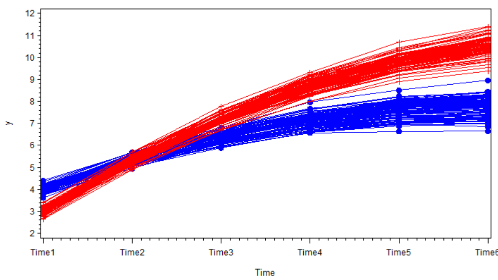
The results of simulations for confidence intervals of the difference of x and y-coordinates of vertices are displayed in Table 4.19. The empirical coverage as well as lower bound and upper bound for the empirical coverage are displayed in the table only for the delta method. For the x-coordinate of the vertex, all three conditions for sample size 20 have nominal coverage higher than the upper bound, while the delta method is appropriate for sample size 100. For the y-coordinate of the vertex, none of 12 conditions have the nominal coverage outside the bounds. Hence the delta method is applicable for the confidence interval of the difference of y-coordinates.



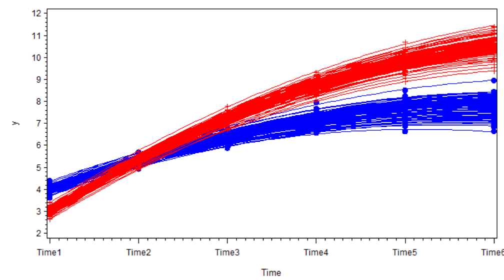
(a) Profile Plot for N = 20



(b) Smoothed Plot for N = 20



(c) Profile Plot for N = 100



(d) Smoothed Plot for N = 100

Figure 4.8: Profile and Smoothed Plots for Random Slope Model

Table 4.20 shows the simulation results of the confidence region for the difference of two vertices from two groups. The table includes the empirical coverage as well as lower bound and upper bound for the empirical coverage. Five conditions have the nominal coverage within the bounds in the table; therefore the approximate chi-square distribution with two degrees of freedom is valid for the confidence region of the difference of vertices if the vertex is not far from the scope of occasions. In section 4.2.1, when the $x-$coordinate of the vertex was outside the scope of occasions, coverage for confidence intervals and confidence region was not large enough compared to the α level. In this section, the $x-$coordinate of the vertex is only slightly outside the scope of occasions, and coverage for confidence intervals and confidence region is acceptable. Therefore we conclude that when the curve is almost flat, the methods for confidence interval and confidence region are less reliable than the intuitive quadratic curve.

Table 4.19: Confidence Interval for Difference of X and Y-Coordinates

		x-coordinate			y-coordinate		
α	Sample Size	Empirical Coverage	lower bound	upper bound	Empirical Coverage	lower bound	upper bound
0.01	100	0.988	0.97913	0.99687	0.998	0.97913	0.99687
✗ 0.01	20	**0.999**	**0.99643**	**1.00157**	0.992	0.98475	0.99925
0.05	100	0.952	0.93875	0.96525	0.949	0.93536	0.96264
✗ 0.05	20	**0.966**	**0.95477**	**0.97723**	0.951	0.93762	0.96438
0.1	100	0.909	0.89404	0.92396	0.897	0.88119	0.91281
✗ 0.1	20	**0.93**	**0.91673**	**0.94327**	0.894	0.87799	0.91001

* Random slope model, when x-coordinates of vertices are outside occasions for two samples

Table 4.20: Confidence Region of Difference of Two Vertices

α	Sample Size	Empirical Coverage	lower bound	upper bound
0.01	100	0.988	0.97913	0.99687
0.01	20	0.995	0.98926	1.00074
0.05	100	0.952	0.93875	0.96525
0.05	20	0.957	0.94443	0.96957
0.1	100	0.906	0.89082	0.92118
✗ 0.1	20	**0.918**	**0.90373**	**0.93227**

* Random slope model, when x-coordinates of vertices are outside occasions for two samples

4.3 Power Analysis

Power analysis plays an important role to reject the null hypothesis of identical vertex for two groups given that the vertices of two groups are actually different. The power function is interesting to be developed for testing the difference of two vertices. Consider the null hypothesis,

$$H_0 : \boldsymbol{V}^{(\mathrm{C})} = \boldsymbol{V}^{(\mathrm{T})} \text{ v.s. } H_a : \boldsymbol{V}^{(\mathrm{C})} \neq \boldsymbol{V}^{(\mathrm{T})} \tag{4.6}$$

where $\boldsymbol{V}^{(\mathrm{C})}$ and $\boldsymbol{V}^{(\mathrm{T})}$ are distinct vertices of control and treatment groups. Since the vertices are nonlinear functions of $\boldsymbol{\beta}$, the null hypothesis can also be expressed as

$$H_0: \begin{pmatrix} \frac{-\beta_1^{(\mathrm{C})}}{2\beta_2^{(\mathrm{C})}} \\ \beta_0^{(\mathrm{C})} - \frac{[\beta_1^{(\mathrm{C})}]^2}{4\beta_2^{(\mathrm{C})}} \end{pmatrix} = \begin{pmatrix} \frac{-\beta_1^{(\mathrm{T})}}{2\beta_2^{(\mathrm{T})}} \\ \beta_0^{(\mathrm{T})} - \frac{[\beta_1^{(\mathrm{T})}]^2}{4\beta_2^{(\mathrm{T})}} \end{pmatrix}.$$

Under some conditions,

$$H_0: \boldsymbol{\beta}^{(\mathrm{C})} = \boldsymbol{\beta}^{(\mathrm{T})} \text{ v.s. } H_a: \boldsymbol{\beta}^{(\mathrm{C})} \neq \boldsymbol{\beta}^{(\mathrm{T})} \tag{4.7}$$

is equivalent. Therefore the difference of two vertices may be tested either indirectly by an F test with respect to regression coefficients $\boldsymbol{\beta}$'s or directly by a chi-square test with regard to vertices $\boldsymbol{V}$'s.

Comparing the hypotheses (4.6) and (4.7), provided that the quadratic terms of two populations are equal, $\beta_2^{(\mathrm{C})} = \beta_2^{(\mathrm{T})} = \beta_2$, the null hypothesis (4.7) becomes

$$H_0: \begin{pmatrix} \frac{-\beta_1^{(\mathrm{C})}}{2\beta_2} \\ \beta_0^{(\mathrm{C})} - \frac{[\beta_1^{(\mathrm{C})}]^2}{4\beta_2} \end{pmatrix} = \begin{pmatrix} \frac{-\beta_1^{(\mathrm{T})}}{2\beta_2} \\ \beta_0^{(\mathrm{T})} - \frac{[\beta_1^{(\mathrm{T})}]^2}{4\beta_2} \end{pmatrix}.$$

For the $x-$coordinate of the vertex, if $\beta_1^{(\mathrm{C})} = \beta_1^{(\mathrm{T})}$, then $V_x^{(\mathrm{C})} = V_x^{(\mathrm{T})}$ and vice versa. Similarly, for the $y-$coordinate of the vertex, if the $\beta_1^{(\mathrm{C})} = \beta_1^{(\mathrm{T})}$ and $\beta_0^{(\mathrm{C})} = \beta_0^{(\mathrm{T})}$ then $V_y^{(\mathrm{C})} = V_y^{(\mathrm{T})}$. Therefore the two null hypotheses $H_0: \boldsymbol{\beta}^{(\mathrm{C})} = \boldsymbol{\beta}^{(\mathrm{T})}$ and $H_0: \boldsymbol{V}^{(\mathrm{C})} = \boldsymbol{V}^{(\mathrm{T})}$ are necessarily equivalent. More specifically, comparing a chi-square statistic χ_p^2 with p degrees of freedom, and a F statistic $F_{p,\ q}$ with numerator degrees of freedom p and denominator degrees of freedom q, when q tends to infinity, $\chi_p^2 \to p \cdot F_{p,\ q}$ (Casella and Berger, 2002). On the other hand, if the quadratic terms of two samples are different, $\beta_2^{(\mathrm{C})} \neq \beta_2^{(\mathrm{T})}$, the two null hypothesis $H_0: \boldsymbol{\beta}^{(\mathrm{C})} = \boldsymbol{\beta}^{(\mathrm{T})}$ and $H_0: \boldsymbol{V}^{(\mathrm{C})} = \boldsymbol{V}^{(\mathrm{T})}$ are not necessarily equivalent. Since for the $x-$coordinate of vertex, the ratio $\frac{\beta_1^{(\mathrm{C})}}{\beta_2^{(\mathrm{C})}} = \frac{\beta_1^{(\mathrm{T})}}{\beta_2^{(\mathrm{T})}}$ leads to $V_x^{(\mathrm{C})} = V_x^{(\mathrm{T})}$, i.e. even $\beta_1^{(\mathrm{C})} \neq \beta_1^{(\mathrm{T})}$ and $\beta_2^{(\mathrm{C})} \neq \beta_2^{(\mathrm{T})}$ may result in $V_x^{(\mathrm{C})} = V_x^{(\mathrm{T})}$. Similarly, for the $y-$coordinate of the vertex, the difference of the ratios $\frac{[\beta_1^{(C)}]^2}{\beta_2}$ and $\frac{[\beta_1^{(T)}]^2}{\beta_2}$ can be offset by the difference of $\beta_0^{(C)}$ and $\beta_0^{(T)}$. Namely, even $\beta_0^{(C)} \neq \beta_0^{(T)}$, $\beta_1^{(C)} \neq \beta_1^{(T)}$ and $\beta_2^{(C)} \neq \beta_2^{(T)}$ may not preclude $V_y^{(C)} = V_y^{(T)}$.

4.3.1 Power Function of F Test for Growth Curves with Common Quadratic Term

Repeated measurements on two independent samples, control and treatment, can be presented by a split plot design model,

$$y_{ijk} = \mu_{...} + \alpha_{0i(k)} + \tau_j + \gamma_k + (\tau\gamma)_{jk} + \epsilon_{ijk}, \tag{4.8}$$

where

y_{ijk} is the response at j^{th} occasion for i^{th} subject from group k,

$\mu_{...}$ is a constant for grand mean of all the observations,

$\alpha_{0i(k)}$ is the random effect for subject i nested within group k, and $\alpha_{0i(k)} \sim N(0, \sigma^2_{\alpha_0})$,

τ_j is the fixed time effect and τ_j's are constants subject to the restriction $\sum \tau_j = 0$,

γ_k is the fixed group effect and γ_k's are constants subject to the restriction $\sum \gamma_k = 0$,

$\epsilon_{ijk} \sim N(0, \sigma_e^2)$, and independent of the $\alpha_{0i(k)}$,

$i = 1, 2, ..., N;\ N = N_1 + N_2;\ j = 1, 2, ...n_i;$ and $k = 1, 2.$ N is the total sample size, N_1 and N_2 are sample sizes for control and treatment groups and n_i is the number of occasions assuming to be same for all the subjects as n.

The corresponding 2nd order random intercept model with compound symmetry covariance structure with respect to model (4.8) is model (4.1), Given the common quadratic term for control and treatment groups, $\beta_2^{(C)} = \beta_2^{(T)} = \beta_2$, the equivalent null hypothesis to test $H_0 : \boldsymbol{\beta}^{(C)} = \boldsymbol{\beta}^{(T)}$ with regard to the F test is $H_0 : \boldsymbol{C}_1\boldsymbol{\beta} = \mathbf{0}$, where

$$\boldsymbol{C}_1 = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \end{pmatrix}, \qquad \boldsymbol{\beta} = \begin{pmatrix} \beta_0^{(\mathrm{C})} \\ \beta_1^{(\mathrm{C})} \\ \beta_0^{(\mathrm{T})} \\ \beta_1^{(\mathrm{T})} \\ \beta_2 \end{pmatrix}$$

The F test statistic is,

$$F = \frac{(\boldsymbol{C}_1\hat{\boldsymbol{\beta}})' \left[\boldsymbol{C}_1 \left(\sum_i \boldsymbol{X}_i' \hat{\Sigma}_{\boldsymbol{y}_i}^{-1} \boldsymbol{X}_i\right)^{-1} \boldsymbol{C}_1'\right]^{-1} (\boldsymbol{C}_1\hat{\boldsymbol{\beta}})}{\mathrm{rank}(\boldsymbol{C}_1)}, \tag{4.9}$$

with the non-centrality parameter

$$\lambda_3 = (\boldsymbol{C}_1\boldsymbol{\beta})' \left[\boldsymbol{C}_1 \left(\sum_i \boldsymbol{X}_i' \Sigma_{\boldsymbol{y}_i}^{-1} \boldsymbol{X}_i\right)^{-1} \boldsymbol{C}_1'\right]^{-1} (\boldsymbol{C}_1\boldsymbol{\beta}),$$

where $\Sigma_{\boldsymbol{y}_i} = \sigma_e^2 \cdot \boldsymbol{I}_{n\times n} + \sigma_{\alpha_0}^2 \cdot \boldsymbol{J}_{n\times n}$ and $\boldsymbol{X}_i$ is the model matrix for control group and treatment group,

$$\boldsymbol{X}_i^{(\mathrm{C})} = \begin{pmatrix} 1 & t_{i1} & 0 & 0 & t_{i1}^2 \\ 1 & t_{i2} & 0 & 0 & t_{i2}^2 \\ 1 & t_{i3} & 0 & 0 & t_{i3}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_{in} & 0 & 0 & t_{in}^2 \end{pmatrix}, \qquad \boldsymbol{X}_i^{(\mathrm{T})} = \begin{pmatrix} 0 & 0 & 1 & t_{i1} & t_{i1}^2 \\ 0 & 0 & 1 & t_{i2} & t_{i2}^2 \\ 0 & 0 & 1 & t_{i3} & t_{i3}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & t_{in} & t_{in}^2 \end{pmatrix}.$$

The numerator degrees of freedom is $\mathrm{ndf}_2 = \mathrm{rank}(\boldsymbol{C}_1)$, and the between-within denominator degrees of freedom $\mathrm{ddf}_2 = N \cdot (n-1) - \mathrm{rank}(\boldsymbol{C}_1)$, and the power function is,

$$\mathrm{Power} \approx \mathrm{Prob}\left\{F(\mathrm{ndf}_2,\ \mathrm{ddf}_2,\ \lambda_3) > F_{1-\alpha,\ \mathrm{ndf}_2,\ \mathrm{ddf}_2}\right\},$$

where $F_{1-\alpha}$ is the critical value for the central F distribution with Type I error rate α.

For the 2nd order random slope model (4.2), the test of $H_0 : \boldsymbol{\beta}^{(\mathrm{C})} = \boldsymbol{\beta}^{(\mathrm{T})}$ using an F-type statistic (4.9) is approximate since the denominator degrees of freedom ddf_{2a} are not known. As reviewed in Section 3.3.2, the commonly used methods to compute the denominator degrees of freedom are Satterthwaite and Kenward-Roger. The power function for the approximate F test is

$$\text{Power} \approx \text{Prob}\left\{F(\mathrm{ndf}_2,\ \mathrm{ddf}_{2a},\ \lambda_3) > F_{1-\alpha,\ \mathrm{ndf}_2,\ \mathrm{ddf}_{2a}}\right\},$$

where $F_{1-\alpha}$ is the critical value of the central F distribution with the approximate denominator degrees of freedom.

4.3.2 Power Function of F Test for Growth Curves with Heterogeneity of the Quadratic Term

Assume the quadratic terms of two growth curves are not identical, $\beta_2^{(\mathrm{C})} \neq \beta_2^{(\mathrm{T})}$, for the 2nd order random intercept model (4.1), the equivalent null hypothesis to test $H_0 : \boldsymbol{\beta}^{(\mathrm{C})} = \boldsymbol{\beta}^{(\mathrm{T})}$ is $H_0 : \boldsymbol{C}_2\boldsymbol{\beta} = \boldsymbol{0}$ where

$$\boldsymbol{C}_2 = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{pmatrix}, \qquad \boldsymbol{\beta} = \begin{pmatrix} \beta_0^{(\mathrm{C})} \\ \beta_1^{(\mathrm{C})} \\ \beta_2^{(\mathrm{C})} \\ \beta_0^{(\mathrm{T})} \\ \beta_1^{(\mathrm{T})} \\ \beta_2^{(\mathrm{T})} \end{pmatrix}$$

The F test statistic and the corresponding non-centrality parameter is,

$$F = \frac{(\boldsymbol{C}_2\hat{\boldsymbol{\beta}})' \left[\boldsymbol{C}_2 \left(\sum_i \boldsymbol{X}_i'\hat{\Sigma}_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1} \boldsymbol{C}'\right]^{-1} (\boldsymbol{C}_2\hat{\boldsymbol{\beta}})}{\mathrm{rank}(\boldsymbol{C}_2)}, \tag{4.10}$$

$$\lambda_4 = (\boldsymbol{C}_2\boldsymbol{\beta})' \left[\boldsymbol{C}_2 \left(\sum_i \boldsymbol{X}_i'\Sigma_{\boldsymbol{y}_i}^{-1}\boldsymbol{X}_i\right)^{-1} \boldsymbol{C}_2'\right]^{-1} (\boldsymbol{C}_2\boldsymbol{\beta}),$$

where $\Sigma_{\boldsymbol{y}_i} = \sigma_e^2 \cdot \boldsymbol{I}_{n\times n} + \sigma_{\alpha_0}^2 \cdot \boldsymbol{J}_{n\times n}$ and $\boldsymbol{X}_i$ is the model matrix for control group or treatment group,

$$\boldsymbol{X}_i^{(\mathrm{C})} = \begin{pmatrix} 1 & t_{i1} & t_{i1}^2 & 0 & 0 & 0 \\ 1 & t_{i2} & t_{i2}^2 & 0 & 0 & 0 \\ 1 & t_{i3} & t_{i3}^2 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_{in} & t_{in}^2 & 0 & 0 & 0 \end{pmatrix}, \qquad \boldsymbol{X}_i^{(\mathrm{T})} = \begin{pmatrix} 0 & 0 & 0 & 1 & t_{i1} & t_{i1}^2 \\ 0 & 0 & 0 & 1 & t_{i2} & t_{i2}^2 \\ 0 & 0 & 0 & 1 & t_{i3} & t_{i3}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & t_{in} & t_{in}^2 \end{pmatrix}.$$

The numerator degrees of freedom is $\mathrm{ndf}_3 = \mathrm{rank}(\boldsymbol{C}_2)$, and the denominator degrees of freedom $\mathrm{ddf}_3 = N \cdot (n-1) - \mathrm{rank}(\boldsymbol{C}_2)$ for between-within method (Schluchter and Elashoff, 1990), and the power function is,

$$\mathrm{Power} \approx \mathrm{Prob}\left\{F(\mathrm{ndf}_3,\ \mathrm{ddf}_3,\ \lambda_4) > F_{1-\alpha,\ \mathrm{ndf}_3,\ \mathrm{ddf}_3}\right\},$$

where $F_{1-\alpha}$ is the critical value for the central F distribution with Type I error rate α.

For the 2$^{\text{nd}}$ order random slope model (4.2), the F distribution for test statistic (4.10) becomes approximate in that the denominator degrees of freedom is not exact. The approximate power function is,

$$\mathrm{Power} \approx \mathrm{Prob}\left\{F(\mathrm{ndf}_3,\ \mathrm{ddf}_{3a},\ \lambda_4) > F_{1-\alpha},\ \mathrm{ndf}_3,\ \mathrm{ddf}_{3a}\right\},$$

where ddf_{3a} is the approximate denominator degrees of freedom that can be calculated by the Satterthwaite or Kenward-Roger method.

4.3.3 Power Function for Chi-Square Test

The non-central chi-square distribution is used to compute power for the null hypothesis for a direct test $H_0 : \boldsymbol{V}^{(\mathrm{C})} = \boldsymbol{V}^{(\mathrm{T})}$. As proved in Section 4.1.1, $\hat{\boldsymbol{V}}^{(\mathrm{diff})} \overset{a}{\sim}$

$N_2\left(\boldsymbol{V}^{(\text{diff})}, \Sigma_{\hat{\boldsymbol{V}}^{(\text{diff})}}\right)$, then $\hat{\boldsymbol{V}}^{(\text{diff})\prime}\Sigma^{-1}_{\hat{\boldsymbol{V}}^{(\text{diff})}}\hat{\boldsymbol{V}}^{(\text{diff})}$ distributes approximately as a non-central chi-square with 2 degrees of freedom with the non-centrality parameter

$$\begin{aligned}
\lambda_5 &= \boldsymbol{V}^{(\text{diff})\prime}\Sigma^{-1}_{\hat{\boldsymbol{V}}^{(\text{diff})}}\boldsymbol{V}^{(\text{diff})} \\
&= \begin{pmatrix} \boldsymbol{V}_x^{(\text{T})} - \boldsymbol{V}_x^{(\text{C})} \\ \boldsymbol{V}_y^{(\text{T})} - \boldsymbol{V}_y^{(\text{C})} \end{pmatrix}' \Sigma^{-1}_{\hat{\boldsymbol{V}}^{(\text{diff})}} \begin{pmatrix} \boldsymbol{V}_x^{(\text{T})} - \boldsymbol{V}_x^{(\text{C})} \\ \boldsymbol{V}_y^{(\text{T})} - \boldsymbol{V}_y^{(\text{C})} \end{pmatrix} \\
&= \begin{pmatrix} \frac{-\beta_1^{(\text{T})}}{2\beta_2^{(\text{T})}} - \frac{-\beta_1^{(\text{C})}}{2\beta_2^{(\text{C})}} \\ \beta_0^{(\text{T})} - \frac{\beta_1^{(\text{T})2}}{4\beta_2^{(\text{T})}} - \beta_0^{(\text{C})} + \frac{\beta_1^{(\text{C})2}}{4\beta_2^{(\text{C})}} \end{pmatrix}' \Sigma^{-1}_{\hat{\boldsymbol{V}}^{(\text{diff})}} \begin{pmatrix} \frac{-\beta_1^{(\text{T})}}{2\beta_2^{(\text{T})}} - \frac{-\beta_1^{(\text{C})}}{2\beta_2^{(\text{C})}} \\ \beta_0^{(\text{T})} - \frac{\beta_1^{(\text{T})2}}{4\beta_2^{(\text{T})}} - \beta_0^{(\text{C})} + \frac{\beta_1^{(\text{C})2}}{4\beta_2^{(\text{C})}} \end{pmatrix}.
\end{aligned}$$

That is, $\hat{\boldsymbol{V}}^{(\text{diff})\prime}\Sigma^{-1}_{\hat{\boldsymbol{V}}^{(\text{diff})}}\hat{\boldsymbol{V}}^{(\text{diff})} \overset{a}{\sim} \chi^2_{2,\lambda_5}$. Under the null hypothesis, the non-centrality parameter $\lambda_5 = 0$. The approximate power function is,

$$\text{Power} \approx \text{Prob}\left\{\chi^2(2,\lambda_5) > \chi^2_{1-\alpha,2}\right\},$$

where $\chi^2_{1-\alpha,2}$ is the critical value given test size level α. Using $\hat{\Sigma}_{\hat{\boldsymbol{V}}^{(\text{diff})}}$, the consistent statistic for $\Sigma_{\hat{\boldsymbol{V}}^{(\text{diff})}}$, the decision rule is, reject the null hypothesis if

$$\begin{pmatrix} \boldsymbol{V}_x^{(\text{T})} - \boldsymbol{V}_x^{(\text{C})} \\ \boldsymbol{V}_y^{(\text{T})} - \boldsymbol{V}_y^{(\text{C})} \end{pmatrix}' \hat{\Sigma}^{-1}_{\hat{\boldsymbol{V}}^{(\text{diff})}} \begin{pmatrix} \boldsymbol{V}_x^{(\text{T})} - \boldsymbol{V}_x^{(\text{C})} \\ \boldsymbol{V}_y^{(\text{T})} - \boldsymbol{V}_y^{(\text{C})} \end{pmatrix} > \chi^2_{1-\alpha,2},$$

otherwise do not reject the null hypothesis.

4.3.4 Power Results for Growth Curves with Common Quadratic Term

In this section, we investigate the indirect F test for $H_0 : \boldsymbol{\beta}^{(\text{C})} = \boldsymbol{\beta}^{(\text{T})}$ and the direct chi-square test for $H_0 : \boldsymbol{V}^C = \boldsymbol{V}^{(\text{T})}$, assuming $\beta_2^{(\text{C})} = \beta_2^{(\text{T})}$. For the random intercept model (4.1) and parameter sets I, II, and III as shown in Table B.1, twelve combination of datasets are considered with different regression coefficients, variances of random effect, sample sizes, but the same time points. The six time points are $t_{ij} = 0, 1, 2, 3, 4, 5$; and sample sizes are selected to be 20 and 50. Two variance parameters chosen for the random effect are 10 and 80 with apparent difference between them. The vertices for parameter sets I, II, and III are outside the scope of occasions; while the vertices for parameter set IV is within the scope of occasions. The fixed regression coefficients for control and treatment groups, and vertex, are listed in Table B.1.

Table 4.21: Parameters for Power Analysis

		β_0	β_1	β_2	Vertex	Within
Parameter I	Control	6.05	3.0	-0.2	(7.5, 17.3)	No
	Treatment	4.5	3.3	-0.2	(8.25, 18.1125)	No
Parameter II	Control	5	1.7	-0.1	(8.5, 12.225)	No
	Treatment	4	1.9	-0.1	(9.5, 13.025)	No
Parameter III	Control	10	1.15	-0.06	(9.5833, 15.51)	No
	Treatment	9.5	1.45	-0.06	(12.0833, 18.26)	No
Parameter IV	Control	2	8	-1	(4, 18)	Yes
	Treatment	2	8.1	-1	(4.05, 18.4025)	Yes

The simulated power and confidence intervals as well as the theoretical power are displayed in Table 4.22 and 4.23. In the table, parameter sets (a) have the variances $\sigma^2_{\alpha_0} = 10$, $\sigma^2_e = 5$, and parameter sets (b) have $\sigma^2_{\alpha_0} = 80$, $\sigma^2_e = 5$. For parameter sets I, II, and III, with the smaller random effect variance $\sigma^2_{\alpha_0} = 10$, the F test has higher power than the chi-square test for every combination. When the variance of the random effect is larger, $\sigma^2_{\alpha_0} = 80$, it is more obvious that the F test has higher power than the chi-square test for every combination; and the power for both the F and the chi-square tests increases. Then, the increase of the variance $\sigma^2_{\alpha_0}$ would result in a decrease of power for both F and chi-square test. Parameter set IV is for a random intercept model with x-value of the vertex within the scope of the model. In this condition, the results show that there is a small difference between the theoretical power of the chi-square test and the F test even for small sample size. However, for parameter sets I, II, and III, with vertices outside the scope of the occasions, all the asymptotic F power are greater than the power of the chi-square test. As the vertices move further away from parameter set I to parameter set III, the power for both the F test and the chi-square test become lower. Hence the further the vertices are away from the scope of the occasions, the F and chi-square power becomes smaller; and it affects the chi-square power more. The theoretical power of the F test is always between the lower and upper bounds of the simulated power, for the vertex both within and outside the scope of occasions. As the sample size increases, the power will increase as a consequence. However, the theoretical power of chi-square test is between the lower and upper bounds of the simulated power only when the vertex is within the scope of the model. Even worse, when the vertex is further outside the occasions, the simulated power of the chi-square test decreases dramatically; and the difference between the simulated power and the theoretical power of chi-square test is very large. Therefore, when the vertex is far away from the scope of occasions, the use of chi-square test should be given more attention. For all the conditions, increasing sample size will lead to an increase in power. Table 4.22 and 4.23 provide little useful information to compare the denominator degrees of freedom for F test, since the simulated model is random intercept model which has an exact denominator degrees of freedom; the three different degrees of freedom methods, between-within, Satterthwaite and Kenward-Roger, provide similar power.

The random slope models (4.2) are generated using the fixed regression parameters listed in Table B.1 with variances $\sigma_e^2 = 5$, $\sigma_{\alpha_0}^2 = 10$, and $\sigma_{\alpha_1}^2 = 5$. The results are displayed in Table 4.24. Compared to Table 4.22 and 4.23, in all the conditions, the theoretical power and simulated power decrease simutaneously. Hence, adding a random slope term in the model results in a decrease of power for both the F and chi-square tests. Other findings are similar; the theoretical power of F test is higher than the chi-square test when the vertex is outside the scope of occasions, while the theoretical power of chi-square test is competitive with the F test otherwise. The theoretical power of F test is within the lower and upper bounds of the simulated power in all conditions. However, the lower and upper bounds of the simulated power of chi-square test only contain the theoretical power when the vertex is within the scope of occasions.

Table 4.22: Power for Random Intercept Model with Common Quadratic Term

Parameters	Sample Size	Method	Simulated Power	Lower Bound	Upper Bound	Theoretical Power
I (a)	$N = 20$	BWF	0.374	0.344	0.404	0.386
		KRF	0.366	0.336	0.396	0.379
		SATF	0.369	0.339	0.399	0.376
		Chisq	0.208	0.183	0.233	0.329
	$N = 50$	BWF	0.771	0.745	0.797	0.785
		KRF	0.763	0.737	0.789	0.781
		SATF	0.766	0.739	0.792	0.767
		Chisq	0.708	0.680	0.736	0.700
I (b)	$N = 20$	BWF	0.339	0.310	0.368	0.339
		KRF	0.327	0.298	0.356	0.333
		SATF	0.326	0.297	0.355	0.330
		Chisq	0.165	0.142	0.188	0.280
	$N = 50$	BWF	0.725	0.697	0.753	0.718
		KRF	0.718	0.690	0.746	0.715
		SATF	0.718	0.690	0.746	0.714
		Chisq	0.635	0.605	0.665	0.616
II (a)	$N = 20$	BWF	0.212	0.187	0.237	0.188
		KRF	0.204	0.179	0.229	0.185
		SATF	0.207	0.182	0.232	0.184
		Chisq	0.061	0.046	0.076	0.149
	$N = 50$	BWF	0.431	0.400	0.462	0.420
		KRF	0.428	0.397	0.459	0.417
		SATF	0.429	0.398	0.460	0.406
		Chisq	0.163	0.140	0.186	0.317
II (b)	$N = 20$	BWF	0.169	0.146	0.192	0.170
		KRF	0.160	0.137	0.183	0.167
		SATF	0.160	0.137	0.183	0.166
		Chisq	0.042	0.030	0.054	0.131
	$N = 50$	BWF	0.374	0.344	0.404	0.374
		KRF	0.370	0.109	0.151	0.372
		SATF	0.369	0.339	0.399	0.370
		Chisq	0.130	0.109	0.151	0.270

Table 4.23: Power for Random Intercept Model with Common Quadratic Term

Parameters	Sample Size	Method	Simulated Power	Lower Bound	Upper Bound	Theoretical Power
III (a)	$N = 20$	BWF	0.345	0.315	0.275	0.337
		KRF	0.333	0.304	0.362	0.331
		SATF	0.335	0.306	0.364	0.328
		Chisq	0.025	0.015	0.035	0.120
	$N = 50$	BWF	0.720	0.692	0.748	0.716
		KRF	0.714	0.686	0.742	0.712
		SATF	0.716	0.688	0.744	0.698
		Chisq	0.110	0.091	0.129	0.239
III (b)	$N = 20$	BWF	0.359	0.329	0.389	0.332
		KRF	0.346	0.316	0.376	0.326
		SATF	0.345	0.315	0.375	0.323
		Chisq	0.043	0.030	0.056	0.116
	$N = 50$	BWF	0.716	0.688	0.744	0.708
		KRF	0.711	0.683	0.739	0.705
		SATF	0.711	0.683	0.739	0.703
		Chisq	0.089	0.071	0.107	0.228
IV (a)	$N = 20$	BWF	0.080	0.063	0.097	0.081
		KRF	0.079	0.058	0.090	0.081
		SATF	0.077	0.060	0.094	0.081
		Chisq	0.079	0.062	0.096	0.082
	$N = 50$	BWF	0.141	0.119	0.163	0.134
		KRF	0.135	0.114	0.156	0.133
		SATF	0.137	0.116	0.158	0.133
		Chisq	0.142	0.120	0.164	0.134
IV (b)	$N = 20$	BWF	0.082	0.065	0.099	0.077
		KRF	0.074	0.058	0.090	0.077
		SATF	0.074	0.058	0.090	0.077
		Chisq	0.083	0.066	0.100	0.078
	$N = 50$	BWF	0.129	0.108	0.150	0.123
		KRF	0.129	0.108	0.150	0.122
		SATF	0.128	0.107	0.149	0.122
		Chisq	0.131	0.110	0.152	0.123

Table 4.24: Power for Random Slope Model with Common Quadratic Term

Parameters	Sample Size	Method	Simulated Power	Lower Bound	Upper Bound	Theoretical Power
I (a)	$N = 20$	BWF	0.227	0.200	0.253	0.222
		KRF	0.186	0.161	0.210	0.210
		SATF	0.190	0.166	0.214	0.210
		Chisq	0.091	0.073	0.109	0.208
	$N = 50$	BWF	0.482	0.451	0.513	0.499
		KRF	0.470	0.439	0.501	0.489
		SATF	0.476	0.445	0.507	0.489
		Chisq	0.375	0.345	0.405	0.463
II (a)	$N = 20$	BWF	0.166	0.143	0.189	0.117
		KRF	0.140	0.118	0.162	0.112
		SATF	0.148	0.126	0.170	0.112
		Chisq	0.028	0.018	0.038	0.107
	$N = 50$	BWF	0.247	0.220	0.274	0.232
		KRF	0.247	0.220	0.247	0.227
		SATF	0.239	0.213	0.265	0.227
		Chisq	0.075	0.059	0.091	0.204
III (a)	$N = 20$	BWF	0.074	0.058	0.090	0.076
		KRF	0.061	0.046	0.076	0.074
		SATF	0.064	0.049	0.079	0.074
		Chisq	0.004	0.001	0.008	0.065
	$N = 50$	BWF	0.124	0.104	0.144	0.119
		KRF	0.111	0.091	0.131	0.117
		SATF	0.116	0.096	0.136	0.117
		Chisq	0.012	0.005	0.019	0.090
IV (a)	$N = 20$	BWF	0.066	0.051	0.081	0.051
		KRF	0.057	0.043	0.072	0.051
		SATF	0.060	0.045	0.075	0.051
		Chisq	0.061	0.046	0.076	0.051
	$N = 50$	BWF	0.058	0.043	0.073	0.053
		KRF	0.054	0.040	0.068	0.053
		SATF	0.054	0.040	0.068	0.053
		Chisq	0.058	0.043	0.073	0.054

Chapter 5

APPLICATION

We apply the statistics derived in Chapter 3 and 4 for vertices on a study of growth of language and early literacy skills in preschoolers who have developmental speech and language impairment. The confidence intervals and confidence region for the difference of vertices from the control and treatment groups are conducted. The direct F test and the indirect chi-square test are also provided.

5.1 Description of Study: Tell Language Efficacy for Preschoolers with Developmental Speech and Language Impairment

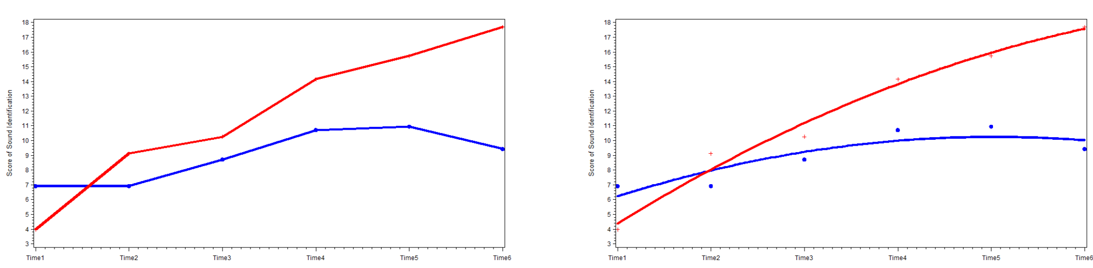
U.S. Department of Education data for the Individuals with Disabilities Education Act (IDEA) reported that 13% of four-year olds and five-year olds are receiving special education services in preschool and that 82% of these children show developmental speech and language impairment (DSLI) as a primary diagnosis (Wilcox *et al.*, 2011). Young children with DSLI often fail to develop crucial pre-literacy skills, which will place those children at high risk for later reading failure and literacy difficulties. Some researchers have pointed out that preschoolers with DSLI demonstrate persistently depressed academic achievement, lower rates of post-secondary school attendance and greater grade retention than their normally-developing peers. Due to these potential risks, it is necessary to address children's early literacy skills and oral language during the preschool years and to increase their ability to benefit from reading and writing instruction in elementary school. Some intervention studies have been performed on children with DSLI, including targeting code-related early literacy skills, inferential language skills and oral language or curriculum supplements based on shared reading. Studies on evaluating an effectiveness of an early childhood intervention with respect to improving early literacy and oral language skills for young children with DSLI are also performed by researchers. One of these studies is examining the efficacy of "Teaching Early Literacy and Language"(TELL) curriculum in promoting the early literacy and oral language growth trajectories of preschoolers with DSLI. The variables in the TELL curriculum include a series of instructions, scripted teaching activities, materials for implementation of oral language and early literacy activities, and professional development for teachers. They targeted one specific skill (e.g., vocabulary, identification of beginning sounds in a word) or small set of skills (e.g., inferential language, print concepts, letter sounds and identification) over a relatively short period of time (e.g., weeks). The TELL curriculum has shown positive results in oral language and early literacy activities in an earlier small randomized controlled trial. Researchers compare those trajectories of children who were enrolled in the TELL curriculum with those who were randomly assigned to control classes (Wilcox *et al.*, 2011).

5.2 Comparison of Vertices for TELL Curriculum and Control Group for Letter Sound Identification Score

First of all, we focused on one specific item from TELL curriculum, Curriculum Based Measurement (CBM) Letter Sound Identification (SoundID) in year 2011. Fifty-seven children with DSLI nested under teacher are randomly assigned to offer the TELL curriculum or accept those with business as usual (BAU). The efficacy variable, SoundID test score, was obtained by six follow-up time measurements (1, 2.25, 3.5, 5.25, 6.5, 7.75 months). The mean and standard deviation of SoundID scores for both children with DSLI from TELL and BAU curriculum are displayed in Table 5.1. On average, compared to the children received BAU, children who were enrolled in the TELL curriculum have higher SoundID scores starting at the second time point. The profile plot and smoothed profile plot for children with DSLI receiving TELL curriculum and BAU are shown in Figure 5.1, which indicate the quadratic curve for the trend; red curve represents the TELL curriculum and blue curve is for BAU curriculum.

Table 5.1: Sound Identification Score by Group (TELL vs. Control): Mean, Standard Deviations at Each Occasion

Variables	TELL ($N_1 = 32$)	Control ($N_2 = 25$)
	Mean (SD)	Mean (SD)
SoundID (T1) Scores	3.970 (4.730)	6.917 (9.180)
SoundID (T2) Scores	9.120 (7.310)	6.920 (8.524)
SoundID (T3) Scores	10.260 (8.080)	8.720 (9.095)
SoundID (T4) Scores	14.148 (8.023)	10.714 (9.670)
SoundID (T5) Scores	15.741 (7.744)	10.955 (8.666)
SoundID (T6) Scores	17.692 (8.480)	9.429 (8.818)



(a) Profile Plot for TELL Efficacy Example

(b) Smoothed Plot for TELL Efficacy Example

Figure 5.1: Profile and Smoothed Plots for TELL Efficacy Example

Table 5.2: Model Selection for Children Received TELL Curriculum

Information Criteria	Random Intercept Model	Random Slope Model
AIC	982.1	✓ **951.2**
AICc	982.3	✓ **951.4**
BIC	983.8	✓ **953.4**

In order to apply the methods for one sample case derived in Chapter 3, the single group was analyzed initially. The model for children enrolled in the TELL curriculum was conducted; two models are compared, they are random intercept model

$$y_{ijk} = \beta_0 + \beta_1 t_{ijk} + \beta_2 t_{ijk}^2 + \beta_{c1} x_{1ij} + \gamma_{0j} + \alpha_{0i(j)} + \epsilon_{ijk},$$

and random slope model

$$y_{ijk} = \beta_0 + \beta_1 t_{ijk} + \beta_2 t_{ijk}^2 + \beta_{c1} x_{1ij} + \gamma_{0j} + \alpha_{0i(j)} + \alpha_{1i(j)} t_{ijk} + \epsilon_{ijk}.$$

Where,

y_{ijk} is the sound identification score at the k^{th} time point for child i nested under teacher j; x_{1ij} is a covariate of mother's education for child i nested under teacher j; γ_{0j} is the random intercept effect of j^{th} teacher, $\gamma_{0j} \sim N(0,\ \sigma^2_{\gamma_0})$; $\alpha_{0i(j)}$ and $\alpha_{1i(j)}$ are the random intercept and random slope effects of i^{th} children nested under j^{th} teacher, $\alpha_{0i(j)} \sim N(0,\ \sigma^2_{\alpha_0})$ and $\alpha_{1i(j)} \sim N(0,\ \sigma^2_{\alpha_1})$; ϵ_{ijk} is the random error term, $\epsilon_{ijk} \sim N(0,\ \sigma^2_e)$. Based on the three criteria AIC, AICc and BIC displayed in Table 5.2, the random slope model is selected because of the smaller values of the criteria. Substituting the estimates of fixed regression parameters, the estimated model is,

$$\hat{y}_{ijk} = 1.803 + 2.923 \cdot t_{ijk} - 0.115 \cdot t_{ijk}^2 + 1.445 \cdot x_{1ij}$$

with the estimates of variance components, $\sigma^2_{\gamma_0} = 5.675$, $\sigma^2_{\alpha_0} = 18.016$, $\sigma^2_{\alpha_1} = 0.833$, and $\sigma^2_e = 9.214$.

Table 5.3: Confidence Intervals of Vertex for Children with DSLI Who Received TELL Curriculum

Method	Vertex	Lower Limit	Upper Limit
Delta for x-coordinate	12.745	4.731	20.759
Gradient for x-coordinate	12.475	8.503	127.986
Delta for y-coordinate	20.430	11.842	29.019

The proposed methods for the confidence interval of vertex are applied, the estimated vertex, lower and upper limits are displayed in Table 5.3. The delta method confidence interval of x-coordinate is $(4.731,\ 20.759)$, while the gradient method obtains the confidence interval $(8.503,\ 127.986)$ which is too wide to be useful. The delta method confidence interval of y-coordinate is $(11.842,\ 29.019)$. The confidence region for the vertex $\boldsymbol{V}' = (V_x,\ V_y)$ is the area under the ellipse,

$$\begin{pmatrix} 12.745 - V_x \\ 20.430 - V_y \end{pmatrix}' \begin{pmatrix} 0.309 & -0.259 \\ -0.259 & 0.269 \end{pmatrix} \begin{pmatrix} 12.745 - V_x \\ 20.430 - V_y \end{pmatrix} \leqslant 5.991 \ .$$

Similar model analysis are also applied for children with DSLI enrolled in BAU. The estimated model is

$$\hat{y}_{ijk} = 2.631 + 2.011 \cdot t_{ijk} - 0.135 \cdot t_{ijk}^2 + 1.735 \cdot x_{1ij}$$

with the estimates of variance components, $\sigma^2_{\gamma_0} = 0$, $\sigma^2_{\alpha_0} = 60.841$, $\sigma^2_{\alpha_1} = 0.332$, and $\sigma^2_e = 4.587$. The random intercept effect of teacher can be removed from the model, since it is too small to be significant. The results for model selection and confidence interval are displayed in Table 5.4 and Table 5.5. For x-coordinate of the estimated vertex, confidence interval from the delta method is $(5.109,\ 9.839)$ while that from gradient method is $(5.979,\ 15.281)$. For the y-coordinate of the estimated vertex, confidence interval from the delta method is (6.520, 13.771). The confidence region for the vertex $\boldsymbol{V'} = (V_x,\ V_y)$ is

$$\begin{pmatrix} 7.474 - V_x \\ 10.145 - V_y \end{pmatrix}' \begin{pmatrix} 0.758 & -0.151 \\ -0.151 & 0.322 \end{pmatrix} \begin{pmatrix} 7.474 - V_x \\ 10.145 - V_y \end{pmatrix} \leqslant 5.991,$$

it is the area within an ellipse.

Table 5.4: Model Selection for Children Received BAU

Information Criteria	Random Intercept Model	Random Slope Model
AIC	692.8	✓ **682.4**
AICc	692.9	✓ **682.6**
BIC	693.4	✓ **683.4**

Table 5.5: Confidence Intervals of Vertex for Children with DSLI Who Received BAU

Method	Vertex	Lower Limit	Upper Limit
Delta for x-coordinate	7.474	5.109	9.839
Gradient for x-coordinate	7.474	5.979	15.281
Delta for y-coordinate	10.145	6.520	13.771

For the letter sound identification of children with DSLI who received the TELL curriculum, the estimated vertex is 20.43 letters at time 50 weeks, while for children with DSLI enrolled in BAU, the estimated vertex is 10.145 letters at 30 weeks. The TELL curriculum treatment produced a shift up to 10.285 letters and a shift to the right of 20 weeks. The vertex of TELL group is outside the scope of the occasions, and results can be interpreted that children from BAU class have reached a plateau at 30 weeks but that children enrolled in the TELL curriculum would continue to increase proficiency after week 30. In order to test the difference of locations between the two groups, methods for confidence set for difference of vertices derived in Chapter 4 are applied. The joint random slope model for the TELL and control group is

$$\begin{aligned} y_{ijkl} =& \beta_0^{(\text{mid})} + \beta_0^{(\text{eff})} \cdot I_l + \beta_1^{(\text{mid})} \cdot t_{ijkl} + \beta_0^{(\text{eff})} \cdot I_l \cdot t_{ijkl} + \beta_2^{(\text{mid})} \cdot t_{ijkl}^2 + \beta_2^{(\text{eff})} \cdot I_l \cdot t_{ijkl}^2 \\ &+ \beta_{c1} \cdot x_{1ijl} + \gamma_{0j(l)} + \alpha_{0i(jl)} + \alpha_{1i(jl)} t_{ijkl} + \epsilon_{ijkl}, \end{aligned}$$

where,

$$I_l = \begin{cases} 1 & \text{if } y_{ijkl} \text{ comes from the control group,} \\ 0 & \text{if } y_{ijkl} \text{ comes from the TELL group.} \end{cases}$$

y_{ijkl} is the sound identification score at the k^{th} time point for child i under teacher j and curriculum l; x_{1ijl} is a covariate of mother's education for child i under teacher j and curriculum l; $\gamma_{0j(l)}$ is the random effect of j^{th} teacher nested under curriculum, $\gamma_{0j(l)} \sim N(0,\ \sigma^2_{\gamma_0})$; $\alpha_{0i(jl)}$ and $\alpha_{1i(jl)}$ are the random intercept and random slope effects of i^{th} children nested under j^{th} teacher and l^{th} curriculum, $\alpha_{0i(jl)} \sim N(0,\ \sigma^2_{\alpha_0})$ and $\alpha_{1i(jl)} \sim N(0,\ \sigma^2_{\alpha_1})$; ϵ_{ijkl} is the random error term, $\epsilon_{ijkl} \sim N(0,\ \sigma^2_e)$. The fitted regression model is

$$\hat{y}_{ijkl} = 1.766 + 1.109 \cdot I_l + 2.914 \cdot t_{ijkl} - 0.996 \cdot I_l \cdot t_{ijkl} - 0.113 \cdot t_{ijk}^2 - 0.011 \cdot I_l \cdot t_{ijk}^2 + 1.573 \cdot x_{1ijl},$$

with the estimates of variance components, $\sigma^2_{\gamma_0} = 0$, $\sigma^2_{\alpha_0} = 38.209$, $\sigma^2_{\alpha_1} = 0.666$, and $\sigma^2_e = 7.136$. The results of the confidence interval for difference of the vertices from the TELL and control group are displayed in Table 5.6. For the difference of x-coordinates of the vertices, the gradient method is not applicable since the quadratic term for control children and TELL children are not equal, which is against the assumption. The analysis illustrates that the time for children reach the plateau is not significantly different while the sound identification of letters is significantly different, which indicates the advantages of the TELL curriculum. The confidence region for the difference of the vertices from the TELL and BAU groups is,

$$\begin{pmatrix} 5.271 - V_x^{(\text{diff})} \\ 10.285 - V_y^{(\text{diff})} \end{pmatrix}' \begin{pmatrix} 0.174 & -0.129 \\ -0.129 & 0.140 \end{pmatrix} \begin{pmatrix} 5.271 - V_x^{(\text{diff})} \\ 10.285 - V_y^{(\text{diff})} \end{pmatrix} \leqslant 5.991$$

Table 5.6: Confidence Intervals for difference of Vertices for Control and TELL Children

Method	Difference of Vertices	Lower Limit	Upper Limit
Delta for x-coordinates	5.271	-3.084	13.626
Delta for y-coordinates	10.285	0.963	19.607

To compare the vertices from the TELL and the BAU groups for sound identification score, hypothesis testing is performed for a direct chi-square test $H_0 : \boldsymbol{V}^{(\text{C})} = \boldsymbol{V}^{(\text{T})}$, and a indirect F test $H_0 : \boldsymbol{\beta}^{(\text{C})} = \boldsymbol{\beta}^{(\text{T})}$. The test statistic of the chi-square test is $\chi_2^2 = 6.482$ with 2 degrees of freedom; and the p-value of the test is 0.039. At the significance level $\alpha = 0.05$, we reject the null hypothesis that the vertices from control and TELL group are identical, since the p-value is less than α. The test statistic of the F test is shown in Table 5.7 with different denominator degrees of freedom methods. All three p-values are less than the significance level $\alpha = 0.05$, therefore we reject the null hypothesis that the fixed regression coefficients of TELL and control group are equivalent. The chi-square and F tests conclude the identical result.

5.3 Comparison of Vertices for TELL Curriculum and Control Group for Language Protocol Score

Furthermore, we compared another specific item from TELL curriculum, Curriculum Based Measurement (CBM) Language protocol (LPT) in year 2011, 2012, and 2013. There are 162 children with DSLI nested under teacher who are randomly

Table 5.7: F Test for the Difference of Vertices for Control and TELL Children

DDFM	Test Statistic	NDF	DDF	P-value
Between-Within	5.38	3	261	0.0013
Kenward-Roger	5.31	3	112	0.0019
Satterthwaite	5.38	3	112	0.0017

assigned to offer the TELL curriculum or accept those with BAU. The mean and standard deviation of LPT scores for both children with DSLI from the TELL and BAU groups are displayed in Table 5.8; the six follow-up time measurements are 1, 2.25, 3.5, 5.25, 6.5 and 7.75 months. On average, compared to children who received BAU, children who enrolled in the TELL curriculum have higher LPT scores from the beginning.

Table 5.8: Language Protocol Score by Group (TELL Curriculum vs. Control): Mean, Standard Deviations at Each Occasion

Variables	TELL ($N_1 = 82$)	Control ($N_2 = 80$)
	Mean (SD)	Mean (SD)
LPT (T1) Scores	2.263 (2.029)	1.413 (1.821)
LPT (T2) Scores	3.338 (2.016)	2.700 (1.958)
LPT (T3) Scores	3.413 (1.904)	2.45 (1.813)
LPT (T4) Scores	3.886 (1.556)	2.845 (1.600)
LPT (T5) Scores	5.400 (2.216)	3.214 (2.028)
LPT (T6) Scores	4.732 (2.348)	2.956 (1.807)

The profile plot and smoothed profile plot for children with DSLI enrolled in the TELL curriculum and BAU are shown in Figure 5.2, which indicate the quadratic curve for the trend; red curve represents the TELL curriculum and blue curve is for BAU curriculum. The red curve is greater than the blue curve at all time measurements. The joint fitted regression model for the TELL and control group is,

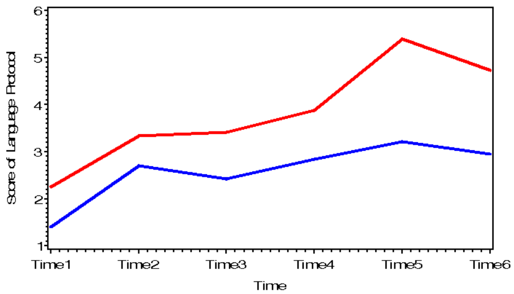
$$\hat{y}_{ijkl} = 1.693 - 0.469 \cdot I_l + 0.651 \cdot t_{ijkl} - 0.073 \cdot I_l \cdot t_{ijkl} - 0.029 \cdot t_{ijk}^2 - 0.017 \cdot I_l \cdot t_{ijk}^2 + 0.055 \cdot x_{1ijl},$$

with the estimates of variance components, $\sigma_{\gamma_0}^2 = 0.582$, $\sigma_{\alpha_0}^2 = 0.999$, $\sigma_{\alpha_1}^2 = 0.008$, and $\sigma_e^2 = 2.076$. Where y_{ijkl} is the language protocol score at k^{th} time point of child i under teacher j and curriculum l, and x_{1ijl} is the mother's education of child i under teacher j and curriculum l.

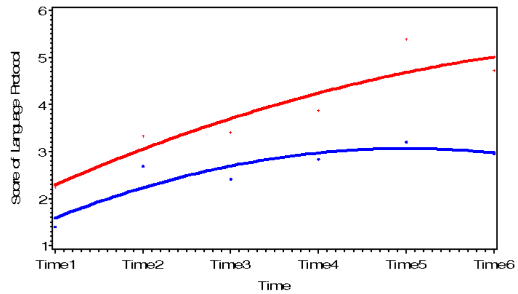
The vertices for TELL and BAU groups are $(11.352,\ 5.382)$ and $(6.368,\ 3.069)$, and the difference is $(4.985,\ 2.314)$. The confidence interval for difference of two vertices for LPT are displayed in Table 5.9. For the difference of x-coordinates of vertices, the gradient method is not applicable since the quadratic term for control children and TELL children are not equal, which violates the assumption. The results illustrate that the time for children reach the plateau is not significantly different while the CBM language protocol scores are significantly different, which indicates the advantages of the TELL curriculum. The confidence region, an ellipse, of difference of the two vertices are shown in (5.1).

$$\begin{pmatrix} 4.985 - V_x^{(\text{diff})} \\ 2.314 - V_y^{(\text{diff})} \end{pmatrix}' \begin{pmatrix} 0.322 & -1.459 \\ -1.459 & 8.304 \end{pmatrix} \begin{pmatrix} 4.985 - V_x^{(\text{diff})} \\ 2.314 - V_y^{(\text{diff})} \end{pmatrix} \leqslant 5.991 \tag{5.1}$$

To compare the vertices from the TELL and the BAU groups for language protocol score, hypothesis testing is performed for a direct chi-square test $H_0 : \boldsymbol{V}^{(\text{C})} = \boldsymbol{V}^{(\text{T})}$, and a indirect F test $H_0 : \boldsymbol{\beta}^{(\text{C})} = \boldsymbol{\beta}^{(\text{T})}$. The test statistic of the chi-square test is $\chi_2^2 = 19.016$ with 2 degrees of freedom; and the p-value of the test is 0+. At the significance level $\alpha = 0.05$, there is a strong evidence to reject the null hypothesis, therefore the



(a) Profile Plot for TELL Efficacy Example



(b) Smoothed Plot for TELL Efficacy Example

Figure 5.2: Profile and Smoothed Plots for TELL Efficacy Example

Table 5.9: Confidence Intervals for difference of Vertices for Control and TELL Children for LPT

Method	Difference of Vertices	Lower Limit	Upper Limit
Delta for x-coordinates	4.985	-2.835	12.804
Delta for y-coordinates	2.314	0.766	3.861

difference of the vertices from control and TELL curriculum are significantly different. The test statistic of the F test is shown in Table 5.10 with different denominator degrees of freedom methods. All p-values are far less than the significance level $\alpha = 0.05$, therefore there is a strong evidence to reject the null hypothesis that the difference between the fixed regression coefficients is not significant. In this example, the chi-square test and F test provide the identical conclusion.

To conclude, the methods of confidence set for vertex of one sample and the methods of confidence set for the difference of vertices of the two independent samples can be applied to analyze the TELL Efficacy example. The F test statistic and chi-square test statistic can also be computed to test the difference of the vertices of TELL and control group.

Table 5.10: F Test for the Difference of Vertices for Control and TELL Children

DDFM	Test Statistic	NDF	DDF	P-value
Between-Within	12.90	3	773	0+
Kenward-Roger	12.77	3	224	0+
Satterthwaite	12.90	3	230	0+

Chapter 6

DISCUSSION

6.1 Conclusion

Initially methods for the confidence interval and confidence region for the vertex of one quadratic growth curve model were discussed in this dissertation. The delta method and the gradient method were developed for the confidence interval of the x-coordinate of the vertex, while delta method was developed for the y-coordinate. An approximate chi-square distribution with two degrees of freedom was derived for the confidence region analysis. Power functions for direct chi-square test and indirect F test on the location of the vertex were also derived. In the simulation studies, random intercept model and random slope model were investigated. For each model, different sample sizes were chosen in order to examine the influence of sample size for all methods. Three different Type I error rates were selected as well for the purpose of making the results more convincing. For the power analysis, both theoretical and simulated power were computed. Depending on all the simulation results, a conclusion is drawn that all methods described in this study for confidence region of the location of quadratic growth curves of 2^{nd} degree polynomial are applicable for different sample sizes, different Type I error rates and different models when the location of the vertex is inside the scope of occasions.

Furthermore, methods for a confidence interval, confidence region, and hypothesis test for the difference of vertices from two independent groups with quadratic growth curves were discussed in this dissertation. The delta method and the gradient method were developed for confidence interval of the difference of x-coordinates for the vertices, while the delta method and the mean response method were developed for the difference of y-coordinates. The delta method was used for the confidence region when the curve is almost flat, and the vertex is far outside the scope of occasions, the methods for confidence intervals and confidence region may have low reliability.

Power functions were also obtained for the test of difference of vertices. Different power functions for chi-square and F test are applicable for quadratic growth curves. Simulation studies were conducted for two independent quadratic growth curves as well to verify the validity of the methods. The conclusions are that when the vertices are within the scope of occasions, both the F test and the chi-square test are valid to test the equality of the vertices of two groups. When the vertex is outside the scope of the model, the use of chi-square test should be given more attention, since the simulated power was much smaller than the theoretical power. For the random intercept model, the larger the variance of random intercept, $\sigma^2_{\alpha_0}$, the lower the power for both F and chi-square tests. Increasing the sample size will always help to increase the power of both tests. For the random slope model, adding a random slope variance, $\sigma^2_{\alpha_0}$, the power of both tests will decrease as a consequence. When the fixed quadratic term, β_2, is close to zero, the vertex of the quadratic growth curve will be further away outside the occasions which will lead to reduce of power for both the F and the chi-square tests.

Finally, using the TELL Efficacy Study, the provided methods for investigating the vertex were demonstrated. We conclude that the delta method is appropriate for both x and y-coordinates of vertex regardless with one or two samples when the vertex is within the scope of occasions. The gradient method is useful for the x-coordinate of the vertex for one samples and two sample with common quadratic term.

6.2 Future Research

An interesting topic for further research can be dealing with vertices of quadratic growth curves under heterogeneity in the random effects population. The linear mixed model (2.1) $\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\alpha}_i + \boldsymbol{\epsilon}_i$ is the homogeneity model, since this model can be seen as a hierarchical Bayes model where $\boldsymbol{\alpha}_i|\boldsymbol{\mu} \sim N(\boldsymbol{\mu}, \boldsymbol{G})$, $\boldsymbol{\mu}$ equals zero with probability 1. Normality for the random effects and error structure are assumptions to the model. Butler and Louis (1992) showed that this assumption has little effect on the fixed effects estimates. However, sometimes the impact of the normality assumption on the estimates of the random effects is severe and very difficulty to examine. For longitudinal data where the systematic part has been misspecified due to the omission of a categorical variable, it might result in a mixture in the distribution of the random effects which violates the normality assumption. Each category for the categorical variable could be a latent subgroup. For instance, studies on the evolution of the blood pressure of patients treated with an antihypertensive drug often report "responders" and "non-responders" to medication, since responders and non responders could be two groups with different normal distributions i.e. the types, "responders" and "non-responders" could be the latent subgroups. In the educational application, there may be responser and non-responder to the TELL curriculum.

Assume that the random effects are sampled from a mixture of g normal distributions with means $\boldsymbol{\mu}_k$ and covariance matrix $\boldsymbol{G}^*$, the linear mixed model can be extended to accommodate clustered $\boldsymbol{\alpha}_i$. A cluster is presented by each component of the mixture and it contains a proportion p_k from the population, $\sum_{k=1}^{g} p_k = 1$. The additional constraint is $E(\boldsymbol{\alpha}_i) = \sum_{k=1}^{g} p_k\boldsymbol{\mu}_k = 0$ to assure that $E(\boldsymbol{y}_i) = \boldsymbol{X}_i\boldsymbol{\beta}$. Then the marginal distribution of the response variable $\boldsymbol{y}_i$ is

$$\boldsymbol{y}_i \sim \sum_{k=1}^{g} p_k \; N(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\mu}_k, \boldsymbol{Z}_i\boldsymbol{G}^*\boldsymbol{Z}_i' + \boldsymbol{R}_i). \tag{6.1}$$

Opposite to the homogeneity model, model (6.1) is called the heterogeneity model by Verbeke and Lesaffre (1996) because it can be seen as a hierarchical Bayes model where $\boldsymbol{\alpha}_i|\boldsymbol{\mu} \sim N(\boldsymbol{\mu}, G^*)$ and the distribution of $\boldsymbol{\mu}$ is assumed to be discrete with probabilities p_k at support points $\boldsymbol{\mu}_k, k = 1, ...g$. To calculate the maximum likelihood estimates for all parameters in model (6.1), the EM algorithm can be applied. Let $\boldsymbol{\pi}$ denote the vector of component probabilities, $\boldsymbol{\pi'} = (p_1, \cdots, p_g)$, let $\boldsymbol{\theta}$ be the vector containing the remaining parameters and assume that $\Psi' = (\boldsymbol{\pi}', \boldsymbol{\theta})'$. Denote p_{ik} the posterior probability for the ith individual to belong to the kth component of the mixture as,

$$p_{ik} = p_{ik}(\Psi) = \frac{p_k f_k(\boldsymbol{y}_i|\boldsymbol{\theta})}{\sum\limits_{m=1}^{g} p_m f_m(\boldsymbol{y}_i|\boldsymbol{\theta})}$$

where $p_m f_m(\boldsymbol{y}_i|\boldsymbol{\theta})$ is the density function of a multivariate normal distribution with mean $\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\mu}_m$ and covariance matrix $\boldsymbol{V}_i$. The empirical Bayes estimate for the random effect $\boldsymbol{\alpha_i}$ is then

$$\hat{\boldsymbol{\alpha}}_i = E(\boldsymbol{\alpha}_i|\boldsymbol{y}_i, \Psi) = \boldsymbol{G}^*\boldsymbol{Z}_i'(\boldsymbol{Z}_i\boldsymbol{G}^*\boldsymbol{Z}_i' + \boldsymbol{R}_i)^{-1}(\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}) + \boldsymbol{A}_i \sum_{k=1}^{g} p_{ik}(\Psi)\boldsymbol{\mu}_k, \tag{6.2}$$

where $\boldsymbol{A}_i = \boldsymbol{I}_i - \boldsymbol{G}^*\boldsymbol{Z}_i'(\boldsymbol{Z}_i\boldsymbol{G}^*\boldsymbol{Z}_i' + \boldsymbol{R}_i)\boldsymbol{Z}_i$. For (6.2), the first component has the same expression as the estimate for $\boldsymbol{\alpha}_i$ obtained in (2.4) under the normality assumption. However, the overall covariance matrix for the random effect $\boldsymbol{G}$ has been replaced by the within-cluster covariance matrix $\boldsymbol{G}^*$. The second component of (6.2) can be treated as a correction term toward the component means, proportional to the posterior probability of belonging to each of the components. In the case of univariate random effects, A_i can be seen as an increasing function of $\sigma_e^2(\boldsymbol{Z}_i\boldsymbol{Z}_i')^{-1}$ satisfying $0 < A_i < 1$. Therefore the correction term will receive more weight in those cases for which the random effect is poorly estimated under the homogeneity model. The choice of the appropriate number g of mixture components is the one of the main issues in deciding on the correct random effects distributions. The test of heterogeneity for a mixed model can be performed using a likelihood ratio test and goodness of fit test. In this topic, two quadratic growth curves with heterogeneity in the random effects population will be studied. They are stated as follows.

Second-order mixed model with random intercept for latent subgroup k,

$$y_{ijk} = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + p_k \mu_{0ik} + \epsilon_{ijk} \tag{6.3}$$

where,

$i = 1, 2, ..., N$, $j = 1, 2, ..., n_i$

n_i is the number of occasions for subject i, N is the number of individuals, and N_k is the number of observations in subgroup k.

β_0, β_1 and β_2 are regression coefficients of fixed effect,

μ_{0ik} is random effect for each individual i in latent group k, $\mu_{0ik} \sim N(\mu_{0k}, \sigma^2_{\mu_0})$ i.e. $\boldsymbol{G}^* = \sigma^2_{\mu_0}$,

ϵ_{ijk} is the random error term at the j^{th} occasion for the i^{th} individual in latent group k, $\epsilon_{ijk} \sim N(0, \sigma^2_e)$ i.e. $\boldsymbol{R}_i = \sigma^2_e \boldsymbol{I}_{n_i \times n_i}$,

μ_{0ik} and ϵ_{ijk} are independent,

p_k is the proportion for latent group k from the population,

y_{ijk} denotes the response variable for i^{th} individual at j^{th} occasion in k^{th} latent group, $E(y_{ijk}) = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + p_k \mu_{0ik}$.

Second-order mixed model with random intercept and random slope for latent subgroup k,

$$y_{ijk} = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + p_k(\mu_{0ik} + \mu_{1ik} t_{ij}) + \epsilon_{ijk} \tag{6.4}$$

where,

μ_{0ik} and μ_{1ik} are independent random effects for individual i in group k, $\mu_{0ik} \sim N(\mu_{0k}, \sigma^2_{\mu_0})$, $\mu_{1ik} \sim N(\mu_{1k}, \sigma^2_{\mu_1})$ and $\mathrm{Cov}(\mu_{0ik}, \mu_{1ik}) = \sigma_{\mu_0\mu_1}$, i.e. $\boldsymbol{G}^* = \begin{pmatrix} \sigma^2_{\mu_0} & \sigma_{\mu_0\mu_1} \\ \sigma_{\mu_0\mu_1} & \sigma^2_{\mu_1} \end{pmatrix}$,

$\epsilon_{ijk}, \beta_0, \beta_1, \beta_2, n_i$, N and N_k are defined same as in model (6.3),

μ_{0ik}, μ_{1ik} and ϵ_{ijk} are independent, i.e. $\mathrm{Cov}(\mu_{0ik}, \epsilon_{ijk}) = 0$ and $\mathrm{Cov}(\mu_{1ik}, \epsilon_{ijk}) = 0$.

y_{ijk} denotes the response variable at j^{th} occasion for the i^{th} individual in the k^{th} latent group, $E(y_{ijk}) = \beta_0 + \beta_1 t_{ij} + \beta_2 t^2_{ij} + p_k(\mu_{0ik} + \mu_{1ik} t_{ij})$.

Denote $\boldsymbol{b'} = (b_0, b_1, b_2)$ the maximum likelihood estimator (MLE) of fixed regression coefficients $\boldsymbol{\beta'} = (\beta_0, \beta_1, \beta_2)$ and let $\boldsymbol{V'_k} = (V_x, V_y)$ be the vertex of quadratic growth curve for latent subgroup k and $\hat{\boldsymbol{V}}'_k = (\hat{V_{xk}}, \hat{V_{yk}})$ be the estimated vertex. For a random intercept model (6.3), the vertex and its estimate for subgroup k are,

$$V_{xk}(\beta_1, \beta_2) = -\frac{1}{2}\beta_1\beta_2^{-1} \ , \qquad V_{yk}(\beta_0, \beta_1, \beta_2, p_k, \mu_{0ik}) = (\beta_0 + p_k\mu_{0k}) - \frac{1}{4}\beta_1^2\beta_2^{-1} \ ,$$

$$\hat{V_{xk}}(b_1, b_2) = -\frac{1}{2}b_1 b_2^{-1} \ , \qquad \hat{V_{yk}}(b_0, b_1, b_2, p_k, \mu_{0k}) = (b_0 + p_k\mu_{0k}) - \frac{1}{4}b_1^2 b_2^{-1} \ .$$

For a random slope model (6.4), the vertex and its estimate for subgroup k are,

$$V_{xk}(\beta_1, \beta_2, p_k, \mu_{1ik}) = -\frac{1}{2}(\beta_1 + p_k\mu_{1ik})\beta_2^{-1} \ ,$$

$$V_{yk}(\beta_0, \beta_1, \beta_2, p_k, \mu_{0ik}, \mu_{1ik}) = (\beta_0 + p_k\mu_{0k}) - \frac{1}{4}(\beta_1 + p_k\mu_{1ik})^2\beta_2^{-1} \ ,$$

$$\hat{V_{xk}}(b_1, b_2, p_k, \mu_{1k}) = -\frac{1}{2}(b_1 + p_k\mu_{1k})b_2^{-1} \ ,$$

$$\hat{V_{yk}}(b_0, b_1, b_2, p_k, \mu_{0k}, \mu_{1k}) = (b_0 + p_k\mu_{0k}) - \frac{1}{4}(b_1 + p_k\mu_{1k})^2 b_2^{-1} \ .$$

In future work, for a quadratic growth curve with heterogeneity in the random effects population, the methods for confidence interval for $x-$coordinate and $y-$coordinate for the vertex as well as the confidence region for the vertex are interesting to be conducted. Furthermore, the confidence intervals and confidence region for the difference of the vertices from different latent populations can also be tested.

Finally, power function and power analysis for testing the difference of locations of the latent populations can be performed.

REFERENCES

Agoston, M. K., *Computer graphics and geometric modeling*, vol. 1 (Springer, 2005).

Akaike, H., "A new look at the statistical model identification", Automatic Control, IEEE Transactions on **19**, 6, 716–723 (1974).

Bachmaier, M., "A confidence set for that x-coordinate where a quadratic regression model has a given gradient", Stat Papers **50** (2009).

Bachmaier, M., "Test and confidence set for the difference of the x-coordinates of the vertices of two quadratic regression models", Stat Papers **51** (2010).

Bijleveld, C. C., L. J. T. v. d. Kamp, A. Mooijaart, W. v. d. Kloot, R. v. d. Leeden and E. V. D. Burg, *Longitudinal Data Analysis: Designs, Models and Methods* (SAGE Publications, 1999).

Bozdogan, H., "Model selection and akaike's information criterion (aic): The general theory and its analytical extensions", Psychometrika **52**, 3, 345–370 (1987).

Brown, L. D., T. T. Cai and A. DasGupta, "Interval estimation for a binomial proportion", Statistical science pp. 101–117 (2001).

Burnham, K. P. and D. R. Anderson, *Model selection and multimodel inference: a practical information-theoretic approach* (Springer Science & Business Media, 2002).

Butler, S. M. and T. A. Louis, "Random effects models with nonparametric priors", Statistics in Medicine **11**, 2981–2000 (1992).

Casella, G. and R. L. Berger, *Statistical Inference* (Pacific Grove, CA: Duxbury Thomson Learning, 2002), second" edn.

Castelloe, J. M. and R. G. O'Brien, "Power and sample size determination for linear models", Statistics, Data Analysis, and Data Mining **240**, 26 (2000).

Davidian, M. and D. M. Giltinan, *Nonlinear Models for Repeated Measurement Data* (London: Chapman & Hall, 1995).

Eslamian, S., *Handbook of Engineering Hydrology: Environmental hydrology and water management* (CRC Press, 2014).

Fan, X., A. Felsóvályi, S. A. Sivo and S. C. Keenan, *SAS for Monte Carlo Studies A Guide for Quantitative Researchers* (SAS Institute Inc., Cary, NC, USA, 2002).

Fisher, R. A., "Xv.?the correlation between relatives on the supposition of mendelian inheritance.", Transactions of the royal society of Edinburgh **52**, 02, 399–433 (1919).

Fitzmaurice, G. M., N. M. Laird and J. H. Ware, *Applied Longitudinal Analysis* (A John Wiley & Sons, INC. Publication, 2004).

Gibson, C. G., *Elementary geometry of differentiable curves: An Undergraduate Introduction* (Cambridge University Press, 2001).

Gruttola, V. D., J. H. Ware and T. A. Louis, "Influence analysis of generalized least squares estimators", Journal of the American Statistical Association **82**, 399 (1987).

Henderson, C. R., O. Kempthorne, S. R. Searle and C. Von Krosigk, "The estimation of environmental and genetic trends from records subject to culling", Biometrics **15**, 2, 192–218 (1959).

Johnson, R. A., D. W. Wichern *et al.*, *Applied multivariate statistical analysis*, vol. 4 (Prentice hall Englewood Cliffs, NJ, 1992).

Kenward, M. G. and J. H. Roger, "Small sample inference for fixed effects from restricted maximum likelihood", International Biometric Society **53**, 3, 983–997 (1997).

Khuri, A. I., T. Mathew and B. K. Sinha, *Statistical tests for mixed linear models*, vol. 906 (John Wiley & Sons, 2011).

Kshirsagar, A. M. and W. B. Smith, *Growth Curves* (New York: Marcel Dekker, 1995).

Kutner, M. H., W. Li, C. J. Nachtsheim and J. Neter, *Applied Linear Statistical Models* (Boston: Montreal McGraw-Hill, 2005), fifth edn.

Laird, N. M. and J. H. Ware, "Random-effects models for longitudinal data", Biometrics **38** (1982).

Littell, R. C., G. A. Milliken, W. W. Stroup, R. D. Wolfinger and O. Schabenberger, *SAS for Mixed Models* (SAS Institute Inc., Cary, NC, USA, 2006), second edn.

Muller, K. E., L. M. LaVange, S. L. Ramey and C. T. Ramey, "Power calculations for general linear multivariate models including repeated measures applications", Journal of American Statistical Association **87**, 420, 1209–1226 (1992).

Pan, J. and K. Fang, *Growth Curve Models and Statistical Diagnostics* (New York: Springer, 2002).

Papanicolaou, A., "Taylor approximation and the delta method", (2009).

Rao, C. R., "The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves", Biometrika pp. 447–458 (1965).

Rencher, A. C. and G. B. Schaalje, *Linear Models in Statistic* (Hoboken, N.J.: Wiley-Interscience, 2008), second edn.

Schabenberger, O. and F. Pierce, *Contemporary Statistical Models for the Plant and Soil Sciences* (CRC Press, 2002).

Schluchter, M. D. and J. T. Elashoff, “Small-sample adjustments to tests with unbalanced repeated measures assuming several covariance structures”, Journal of Statistical Computation and Simulation **37**, 1-2, 69–87 (1990).

Schwarz, G. *et al.*, “Estimating the dimension of a model”, The annals of statistics **6**, 2, 461–464 (1978).

Slaughter, S. J. and L. D. Delwiche, “SAS macro programming for beginners”, SUGI Tutorials **29** (2003).

Verbeke, G. and E. Lesaffre, “A linear mixed-effects model with heterogeneity in the random-effects population”, Journal of the American Statistical Association **91**, 433, 217–221 (1996).

Verbeke, G. and G. Molenberghs, *Linear mixed models for longitudinal data* (Springer Science & Business Media, 2009).

Ware, J. H., “Linear models for the analysis of longitudinal studies”, The American Statistician **39**, 2, 95–101 (1985).

Wilcox, M., S. I. Gray, A. B. Guimond and A. E. Lafferty, “Efficacy of the tell language and literacy curriculum for preschoolers with developmental speech and/or language impairment”, Early Childhood Research Quarterly pp. 278–294 (2011).

Woodroofe, M., “On model selection and the arc sine laws”, The Annals of Statistics pp. 1182–1194 (1982).

APPENDIX A

MODEL SELECTION USING CRITERIA

Table A.1: Selection of Covariance Structure for Random Intercept Model

Information Criteria	$N = 20$			$N = 50$			$N = 100$		
	CS	AR(1)	UN	CS	AR(1)	UN	CS	AR(1)	UN
AIC	966	6	28	989	0	11	994	0	6
AICc	991	7	2	997	0	3	996	0	4
BIC	993	7	0	1000	0	0	1000	0	0

* Mixed model with random intercept, when x-coordinate of vertex is within occasions for one sample

Table A.2: Selection of Covariance Structure for Random Slope Model

Information Criteria	$N = 20$			$N = 50$			$N = 100$		
	CS	AR(1)	UN	CS	AR(1)	UN	CS	AR(1)	UN
AIC	256	260	484	143	144	713	36	27	937
AICc	265	237	462	148	148	704	37	27	936
BIC	314	329	357	232	240	528	118	108	774

* Mixed model with random intercept and random slope, when x-coordinate of vertex is within occasions for one sample

Table A.3: Selection of Covariance Structure for Random Intercept Model

Information Criteria	$N = 10$			$N = 25$			$N = 50$		
	CS	AR(1)	UN	CS	AR(1)	UN	CS	AR(1)	UN
AIC	865	32	103	982	4	14	991	0	9
AICc	964	35	1	996	4	0	997	0	3
BIC	927	32	41	996	4	0	1000	0	0

* Mixed model with random intercept, when x-coordinate of vertex is outside occasions for one sample

Table A.4: Selection of Covariance Structure for Random Slope Model

Information Criteria	$N = 20$			$N = 50$			$N = 100$		
	CS	AR(1)	UN	CS	AR(1)	UN	CS	AR(1)	UN
AIC	0	2	998	0	0	1000	0	0	1000
AICc	0	11	989	0	0	1000	0	0	1000
BIC	0	42	958	0	0	1000	0	0	1000

* Mixed model with random intercept and random slope, when x-coordinate of vertex is outside occasions for one sample

Table A.5: Selection of Covariance Structure for Random Slope Model

Information Criteria	$N = 20$			$N = 100$		
	CS	AR(1)	UN	CS	AR(1)	UN
AIC	982	0	18	993	0	7
AICc	997	0	3	995	0	5
BIC	1000	0	0	1000	0	0

* Mixed model with random intercept, when x-coordinates of vertices are within occasions for two samples

* Same quadratic term $\beta_2^{(C)} = \beta_2^{(T)}$

Table A.6: Selection of Covariance Structure for Random Intercept Model

Information Criteria	$N = 20$			$N = 100$		
	CS	AR(1)	UN	CS	AR(1)	UN
AIC	984	3	13	995	0	5
AICc	992	3	5	995	0	5
BIC	997	3	0	1000	0	0

* Mixed model with random intercept, when x-coordinates of vertices are outside occasions for two samples

* Same quadratic term $\beta_2^{(C)} = \beta_2^{(T)}$

Table A.7: Selection of Covariance Structure for Random Slope Model

Information Criteria	$N = 20$			$N = 100$		
	CS	AR(1)	UN	CS	AR(1)	UN
AIC	153	155	691	0	1	999
AICc	157	160	682	0	1	999
BIC	255	252	492	16	15	969

* Mixed model with random intercept and random slope, when x-coordinates of vertices are within occasions for two samples

* Same quadratic term $\beta_2^{(C)} = \beta_2^{(T)}$

Table A.8: Selection of Covariance Structure for Random Slope Model

Information Criteria	$N = 20$			$N = 100$		
	CS	AR(1)	UN	CS	AR(1)	UN
AIC	222	247	531	6	10	984
AICc	228	254	518	6	11	983
BIC	310	354	336	58	48	893

* Mixed model with random intercept and random slope, when x-coordinates of vertices are outside occasions for two samples

* Same quadratic term $\beta_2^{(C)} = \beta_2^{(T)}$

Table A.9: Selection of Covariance Structure for Random Intercept Model

Information Criteria	$N = 20$			$N = 100$		
	CS	AR(1)	UN	CS	AR(1)	UN
AIC	988	0	12	993	0	7
AICc	999	0	1	993	0	7
BIC	1000	0	0	1000	0	0

* Mixed model with random intercept, when x-coordinates of vertices are within occasions for two samples

* Different quadratic terms $\beta_2^{(C)} \neq \beta_2^{(T)}$

Table A.10: Selection of Covariance Structure for Random Intercept Model

Information Criteria	$N = 20$			$N = 100$		
	CS	AR(1)	UN	CS	AR(1)	UN
AIC	986	0	14	992	0	8
AICc	996	0	4	992	0	8
BIC	1000	0	0	1000	0	0

* Mixed model with random intercept, when x-coordinates of vertices are outside occasions for two samples

* Different quadratic terms $\beta_2^{(C)} \neq \beta_2^{(T)}$

Table A.11: Selection of Covariance Structure for Random Slope Model

Information Criteria	$N = 20$			$N = 100$		
	CS	AR(1)	UN	CS	AR(1)	UN
AIC	0	1	999	0	0	1000
AICc	0	1	999	0	0	1000
BIC	0	39	961	0	0	1000

* Mixed model with random intercept and random slope, when x-coordinates of vertices are within occasions for two samples

* Different quadratic terms $\beta_2^{(C)} \neq \beta_2^{(T)}$

Table A.12: Selection of Covariance Structure for Random Slope Model

Information Criteria	N =20			N = 100		
	CS	AR(1)	UN	CS	AR(1)	UN
AIC	0	0	1000	0	0	1000
AICc	0	1	999	0	0	1000
BIC	0	70	930	0	0	1000

* Mixed model with random intercept and random slope, when x-coordinates of vertices are outside occasions for two samples
* Different quadratic terms $\beta_2^{(C)} \neq \beta_2^{(T)}$

APPENDIX B

MORE RESULTS OF POWER ANALYSIS

Table B.1: Parameters for Power Analysis

		β_0	β_1	β_2	Vertex
Parameter I	Control	6.05	3.0	-0.2	(7.5, 17.3)
	Treatment	4.5	3.3	-0.2	(8.25, 18.1125)
Parameter II	Control	5	1.7	-0.1	(8.5, 12.225)
	Treatment	4	1.9	-0.1	(9.5, 13.025)
Parameter III	Control	10	1.15	-0.06	(9.5833, 15.51)
	Treatment	9.5	1.45	-0.06	(12.0833, 18.26)

Table B.2: Power Analysis for Random Intercept Model with Same Quadratic Term

Parameters	Sample Size	Method	Simulated Power	Lower Bound	Upper Bound	Theoretical Power
I (a)	$N = 20$	BWF	0.374	0.344	0.404	0.386
		KRF	0.366	0.336	0.396	0.379
		SATF	0.369	0.339	0.399	0.383
		Chisq	0.208	0.183	0.233	0.329
	$N = 50$	BWF	0.771	0.745	0.797	0.785
		KRF	0.763	0.737	0.789	0.781
		SATF	0.766	0.739	0.792	0.783
		Chisq	0.708	0.680	0.736	0.700
I (b)	$N = 20$	BWF	0.200	0.175	0.225	0.190
		KRF	0.190	0.166	0.214	0.187
		SATF	0.190	0.166	0.214	0.192
		Chisq	0.073	0.057	0.089	0.162
	$N = 50$	BWF	0.440	0.409	0.471	0.425
		KRF	0.436	0.405	0.467	0.422
		SATF	0.435	0.404	0.466	0.424
		Chisq	0.240	0.213	0.267	0.351
II (a)	$N = 20$	BWF	0.212	0.187	0.237	0.188
		KRF	0.204	0.179	0.229	0.185
		SATF	0.207	0.182	0.232	0.187
		Chisq	0.061	0.046	0.076	0.149
	$N = 50$	BWF	0.431	0.400	0.462	0.420
		KRF	0.428	0.397	0.459	0.417
		SATF	0.429	0.398	0.460	0.419
		Chisq	0.163	0.140	0.186	0.317
II (b)	$N = 20$	BWF	0.117	0.097	0.137	0.108
		KRF	0.114	0.094	0.134	0.107
		SATF	0.114	0.094	0.134	0.108
		Chisq	0.030	0.019	0.041	0.090
	$N = 50$	BWF	0.217	0.191	0.243	0.208
		KRF	0.215	0.189	0.241	0.207
		SATF	0.215	0.189	0.241	0.207
		Chisq	0.059	0.044	0.074	0.157
III (a)	$N = 20$	BWF	0.345	0.315	0.275	0.337
		KRF	0.333	0.304	0.362	0.331
		SATF	0.335	0.306	0.364	0.334
		Chisq	0.025	0.015	0.035	0.120
	$N = 50$	BWF	0.720	0.692	0.748	0.716
		KRF	0.714	0.686	0.742	0.712
		SATF	0.716	0.688	0.744	0.714
		Chisq	0.110	0.091	0.129	0.239
III (b)	$N = 20$	BWF	0.198	0.173	0.223	0.184
		KRF	0.186	0.162	0.210	0.181
		SATF	0.186	0.162	0.210	0.183
		Chisq	0.032	0.021	0.043	0.082
	$N = 50$	BWF	0.408	0.377	0.439	0.409
		KRF	0.401	0.371	0.431	0.406
		SATF	0.401	0.371	0.431	0.408
		Chisq	0.032	0.021	0.043	0.135
IV	$N = 20$	BWF	0.412	0.381	0.443	0.417
		KRF	0.397	0.367	0.427	0.409
		SATF	0.400	0.370	0.430	0.413
		Chisq	0.417	0.386	0.448	0.422
	$N = 100$	BWF	0.987	0.980	0.994	0.986
		KRF	0.987	0.980	0.994	0.986
		SATF	0.987	0.980	0.994	0.986
		Chisq	0.987	0.980	0.994	0.987

* (a): $\sigma_e^2 = 5$, $\sigma_{\alpha_0}^2 = 10$. * (b): $\sigma_e^2 = 10$, $\sigma_{\alpha_0}^2 = 80$.

* Parameters IV: $\boldsymbol{\beta}^{(C)'} = (2, 8, -1)$, $\boldsymbol{\beta}^{(T)'} = (2, 8.1, -1)$, $t_{ij} = (0, 1, 2, 3, 4, 5)$, $\sigma_e^2 = 0.5$, $\sigma_{\alpha_0}^2 = 1$.