Performance of Model Selection Statistics in Growth Mixture Modeling of Homogeneous Data

A Dissertation Presented

by

Ruixue Wang

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philsophy

in

Applied Mathematics and Statistics

Stony Brook University

December 2011

Stony Brook University

The Graduate School

Ruixue Wang

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Stephen J. Finch – Dissertation Advisor
Professor, Ph.D., Department of Applied Math and Statistics

Nancy Mendell - Chairperson of Defense
Professor, Ph.D., Department of Applied Math and Statistics

Song Wu – Third Committee Member
Assistant Professor, Ph.D., Department of Applied Math and Statistics

Derek Gordon – Outside Member
Associate Professor, Ph.D., Department of Genetics
Rutgers, the State University of New Jersey

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Performance of Model Selection Statistics in Growth Mixture Modeling of Homogeneous Data

by

Ruixue Wang

Doctor of Philosophy

in

Applied Mathematics and Statistics

Stony Brook University

2011

Growth mixture modeling (GMM) is used to detect the existence of two or more trajectory patterns among participants in a longitudinal study.

One crucial issue is the determination of the number of longitudinal trajectory patterns. I study the properties of three statistics used to identify the number of components in a sample of data. These are the Bayesian information criterion (BIC), Lo-Mendell-Rubin test (LMRT), and bootstrap likelihood ratio test (BLRT). I estimate the probability that each of these statistics identifies that there is a single component for homogeneous data using the M-plus and SAS PROC TRAJ statistical packages.

I use four distributions for the longitudinal outcome measures: the censored normal distribution, the gamma distribution, the zero-inflated Poisson distribution and the Bernoulli distribution. I considered these factors: trajectory pattern, intra-class

correlation, time measurements, random effects and sample size. For the censored normal distribution, the BIC and LMRT (set at the 0.01 significance level) have the highest fraction of replicates identified as homogeneous. These rates for LMRT are 0.92 or better at significance level 0.01 and 0.98 or better for the BIC. The identification rates of these two statistics are not significantly affected by the intra-class correlation in the trajectory, the trajectory pattern, the number of time measurements, and the sample size. A similar pattern was observed for the gamma distribution using the M-plus statistical package. The identification rate of the LMRT is better than that of the BLRT at both the 0.01 and 0.05 significance levels.

For the ZIP and Bernoulli distribution, PROC TRAJ computations have a higher correct identification rate than those from M-plus. Larger sample size is associated with an increase in the probability that two or more components will be identified for ZIP distributed data following a linear trend and with random effects. The same pattern holds for Bernoulli data. Overall, the BIC statistic has the highest correct identification rate. These rates are on the order of 95% for homogeneous data following either a censored normal or gamma distribution.

TABLE OF CONTENTS

LIST OF TABLES

List of Figures

Chapter 1

Introduction

In a longitudinal study, repeated observations of the same variable at the individual level are made over a period of time. Longitudinal studies are commonly used in social science and biomedical research (White et al., 2000, Chang et al., 2009, Hong &Ho, 2005, Fan, 2001, Duncan et al., 1996). For instance, criminologists track delinquent behavior over time of a sample of individuals and relate it to variables such as income, life style, and health (Eklund, Kerr & Stattin, 2010). Psychologists investigate children's social development and adjustment using longitudinal studies (Crick, Ostrov & Werner, 2006). Doctors use longitudinal studies to identify the progression patterns of diseases and then assess treatment approaches (Gomez et al, 2007).

One longitudinal model is the hierarchical linear model (Raudenbush & Bryk, 2002). The formal specification of this model is given in page 5. The concept is that there is one homogeneous pattern described by coefficients that are random variables specific to each individual. Additionally, growth mixture modeling (GMM) is used to detect the existence of two or more longitudinal patterns among the participants (Muthén & Asparaouhov, 2006). GMM can be treated as a multilevel modeling technique to explore heterogeneity in a population (Kerner & Muthén, 2009).

One crucial issue is the determination of the number of longitudinal patterns present in a sample of data. The Bayesian information criterion (BIC) and likelihood ratio test (LRT) are used to select the number of components. The determination of the number of components and other modeling results in mixture modeling are sensitive to heteroscedasticity in the underlying stochastic process.

My research question is: Given homogeneous data, what are the statistical properties of model selection procedures for the number of components selected? How sensitive are model selection procedures to heteroscedasticity.

I study the probability that the Bayesian information criterion (BIC), entropy measure, Lo-Mendell –Rubin likelihood ratio test (LMRT) and bootstrap likelihood ratio test (BLRT) identify that there is a single component for homogeneous data using the M-plus and SAS statistical packages.

Chapter 2 Literature Review

2.1 Growth Mixture Modeling (GMM)

Conventional growth modeling models longitudinal data as coming from a homogeneous population. The latent curve model (LCM) can be written as

$$y_i = \Lambda\eta_i + \varepsilon_i$$

where y_i is a $p \times 1$vector of repeated measures for individual *i,* and *p* is the number of data measurements across time. The matrix Λ is a $p \times q$ factor-loading matrix whose coefficients are randomly distributed; η_i is a $q \times 1$vector of random coefficients, e.g., intercept and slope, and ε_i is a $p \times 1$vector of residuals. Usually, it is assumed that the random coefficients and residuals are normally distributed; that is, y_i is normally distributed with probability density function

$$f(y_i) = \phi(y_i, \Lambda\alpha, \Lambda\Psi\Lambda' + \Theta)$$

where α is the mean of the random coefficients, Ψ is the covariance matrix of random coefficients, Θ is the covariance matrix of the residuals, and $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$, the pdf of the standard normal distribution.

The assumption of homogeneity is incorrect when individuals belong to different subpopulations. B.O. Nagin (1999) advanced a semiparmeteric group-based trajectory model. Muthén (2001) proposed an extension of latent curve modeling methodology that used growth mixture models. Muthén’s growth mixture modeling uses a finite-mixture random effects model to represent unobserved heterogeneity. Growth mixture modeling partitions the population into an unknown number of latent classes. That is, the GMM model is that the population is composed of *K* latent classes, each characterized by its own LCM.

The probability density function for the GMM is a mixture of normally distributed classes

$$f(y_i) = \sum_{k=1}^{K} \pi_k \Phi_k (y_i, \mathrm{A}_k \alpha_k, \mathrm{A}_k \Psi_k \mathrm{A}_k' + \Theta_k)$$

GMM estimates mean growth curves for each class and models individual variation by the estimation of growth factor variances for each class. GMMs are applied to longitudinal data that are heterogeneous and contain a finite number of latent classes. Each class has its own mean trajectory. Figure 1 presents a GMM with three repeated measures. This model has two quantitative latent parameters, the intercept (*i*) and the slope (*s*). The categorical latent variable that models class membership is *c*. The variables Y_1, Y_2, Y_3 are the repeated outcome measures.

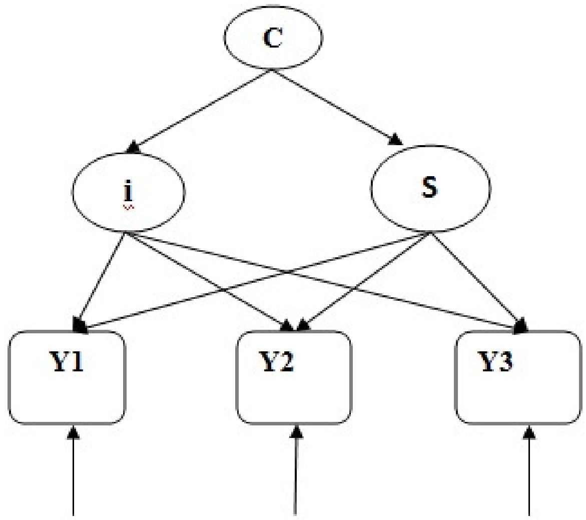

Figure 1 General growth mixture model with three continuous outcomes

The formal statement of a simple growth mixture model describes two levels.

Level 1 models individual change with the model:

$$y_{it} = \pi_{0i} + \pi_{1i} t + \varepsilon_{it}.$$

Level 2 models variation between persons with the model:

$$\pi_{0i} = \beta_{00} + C_{0i} X_i + u_{0i} \quad , \quad \pi_{1i} = \beta_{10} + C_{1i} X_i + u_{1i},$$

where

y_{it} is the outcome measure for subject *i*(*i*=1,…*n*) at time *t*.

The random variable ε_{it} models the within-participant random effects, which are independently and identically normally distributed as $N(0, \sigma^2)$.

The random coefficient π_{0i} is the intercept coefficient for participant *i*, and the random coefficient π_{1i} is the linear rate of change for participant *i*.

The random variables u_{0i} and u_{1i} are individual effects with specified probability distribution in the random effects model. They are not used in a fixed effects model. Finally,

X_i is the group indicator.

2.2. Probability distributions of the outcome measures:

I consider four distributions of the outcome measures in this research: the censored normal (CNORM), the gamma distribution, the zero-inflated Poisson (ZIP) distribution, and the Bernoulli distribution. The censored normal (CNORM) distribution is useful for psychometric scale data (Nagin&Tremblay, 1999). Given that subject *i* belongs to group *k*, the likelihood of observing the trajectory for this subject is

$$p(Y_i = y_i \mid C_i = k, W_i = w_i) = \prod_{y_{ij=Max}} (1-\phi(\frac{Max-\mu_{ijk}}{\sigma})) \prod_{y_{ijk=Min}} (1-\phi(\frac{Min-\mu_{ijk}}{\sigma})) \prod_{Min<y_{ij}<Max} \frac{1}{\sigma}(\frac{y_{ij}-\mu_{ijk}}{\sigma})$$

where

$$\mu_{ijk} = \beta_{0k} + t_{ij}\beta_{1k} + t_{ij}^2\beta_{2k} + ... + w_{ij}\delta_k \;,$$

$W_i = (w_{i1}, ...w_{iT})$ is a time-dependent covariate, and t_{ij} denotes subject *i*'s time measurement at period *j* (Jones et al., 2001).

The gamma distribution is a two-parameter family of continuous distributions. The two parameters are the scale parameter and the shape parameter. Its probability function can be expressed as

$$f(x;k,\theta) = x^{k-1}\frac{e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} \qquad x \geq 0 \quad and\ k,\theta > 0$$

Specifically, if $X \sim \Gamma(k = \frac{\upsilon}{2}, \theta = 2)$, then *X* has a chi-square distribution with υ degrees of freedom.

The zero-inflated Poisson (ZIP) distribution is used for integer valued outcomes when the count outcome is equal to zero more often than one would expect from a Poisson distribution. The ZIP distribution can be treated as a mixture model with two components, a zero component and the Poisson component. A zero count occurs with probability one in the zero component. The probability in the nonzero component follows a Poisson distribution, which includes a zero count as well. A probability is estimated for each outcome to be either in the zero component or the Poisson component.

Formally, the probability mass function of the ZIP is:

$$f(x;\rho,\lambda) \sim \begin{cases} \rho + (1-\rho)e^{-\lambda} & if\ x = 0 \\ (1-\rho)P(x;\lambda) & if\ x > 0 \end{cases}$$

Here ρ is the probability of the zero component, and λ is expected value for the Poisson component.

Given that there are *k* classes, each following a ZIP distribution, the probability of observing y_i given membership in class *k* is

$$p(Y_i = y_i \mid C_i = k, W_i = w_i) = \prod_{y_{ij=0}} (\rho_{ijk} + (1-\rho_{ijk})e^{-\pi_{ijk}}) \prod_{y_{ij}>0} (1-\rho_{ijk}) \frac{\exp(-\lambda_{ijk})\lambda_{ijk}^{y_{ij}}}{y_{ij}!}$$

Additionally, the model's parameters can be linked with polynomial functions such as the following

$$\log(\lambda_{ijk}) = \beta_{0k} + t_{ij}\beta_{1k} + t_{ij}^2\beta_{2k} + ... + w_{ij}\delta_k \quad ,$$

and $\log(\frac{\rho_{ijk}}{1-\rho_{ijk}}) = \alpha_{0k} + t_{ij}\alpha_{1k} + t_{ij}^2\alpha_{2k} + ...$

Finally, the Bernoulli distribution is used for the conditional distribution of dichotomous data, given class membership. The probability for Bernoulli model is

$$p(Y_i = y_i \mid C_i = k, W_i = w_i) = \prod_{y_{ij=0}} (1-p_{ijk}) \prod_{y_{ij}} p_{ijk}$$

with

$$p_{it} = \frac{\exp(\beta_{0k} + t_{ij}\beta_{1k} + t_{ij}^2\beta_{2k} + ... + w_{ij}\delta_k)}{1+\exp(\beta_{0k} + t_{ij}\beta_{1k} + t_{ij}^2\beta_{2k} + ... + w_{ij}\delta_k)}$$

where *i* represents an individual and *t* represents time.

2.3 The Bayesian Information Criterion (BIC; Schwartz, 1978)

In a longitudinal study, determining the number of trajectory classes is a difficult problem, with the determination of optimal statistic unresolved. There are a number of criteria in use, such as the Bayesian information criteria (BIC), the likelihood ratio test and entropy.

Yang (2006) reported that the adjusted BIC (Sclove, 1987) is superior to other information criterion statistics in LCA models. Magidson and Vermunt (2004) gave several examples in which BIC was a good measure to select the number of classes. The formula for the BIC

$BIC = -2 \ln L + s \ln (n)$

where

n =the number of observations; that is, the sample size.

s=the number of free parameters to be estimated.

L=the maximized value of the likelihood function of the estimated model.

BIC value differences of 10 or more are considered as evidence favoring one model over another (Raftery, 1995).

2.4 M-plus statistical program including the Vuong-Lo-Mendell-Rubin test (LMRT) and bootstrapped parametric likelihood ratio test (BLRT)

M-plus is a program to perform growth mixture modeling, with Monte Carlo procedures for power estimation. M-plus provides the entropy measure, the Vuong-Lo-Mendell-Rubin test, and the bootstrapped parametric likelihood ratio test in addition to BIC.

Entropy is not a measure of model fit. A value of entropy near 1 indicates clear delineation of classes (Celeux&Soromenho, 1996). An entropy value of 0.8 or higher for a model suggests that the model can clearly identify the trajectory class that an individual follows. A model with a lower value of entropy may still produce good parameter estimates. The entropy of a model with g classes extracted from n individuals is given by

$$E(g) = -\sum_{i=1}^{g}\sum_{j=1}^{n}\hat{\tau}_{ij}\ln\hat{\tau}_{ij}$$

where $\hat{\tau}_{ij}$ is the posterior probability of membership in class *i*, for subject *j*, $j = 1, \ldots, n$.

The Lo, Mendell, and Rubin likelihood ratio test (Lo, Mendell, & Rubin, 2001) can be used to compare latent class models. The test compares the improvement in fit due to increasing the number of classes by one (i.e., comparing *k* and *k*-1 class models) and provides a p-value that can be used to determine if there is a statistically significant improvement due to adding one more class. That is, the null and alternative hypotheses are:

$H_0 : k-1$ trajectory classes VS $H_1 : k$ trajectory classes

A p-value less than 0.05 suggests rejecting the model with one less class in favor of the model with the larger number of classes.

The likelihood ratio test is given by

$LR = -2[\log L(\hat{\theta}_r) - \log L(\hat{\theta}_u)]$ (Nylund, Asparouhv&Muthén, 2007)

where $\hat{\theta}_r$ is the maximum likelihood estimator for the restricted model and $\hat{\theta}_u$ is the maximum likelihood for the model with fewer restrictions.

In the adjusted Lo, Mendell, and Rubin likelihood ratio test, the k-1 class model is obtained by deleting the first class in the estimated k class model.

The parametric bootstrapped likelihood ratio test (McLachlan &

Peel, 2000) compares the *k* class model to the *k-1* class model. A p-value smaller than the significance level suggests that the model with *k*-1 classes is rejected in favor of the *k* class model. In M-plus, the bootstrap method estimates models for both the *k* class model and the *k-1*class model. Then M-plus generates several data sets using bootstrap draws according to the estimated parameters from the *k*-1 class model. These data are used to generated loglikelihood values and then calculate the test statistic for each bootstrap draw. From the initial analysis, the test statistic is compared to the distribution of test statistics obtained from the bootstrap draws to get a p-value. Then M-plus uses a minimum number of bootstrap draws to give an approximate p-value. The default number of bootstrap draws ranges from 2 to 100 (Muthén & Muthén, 1998-2010).

2.5 PROC TRAJ

PROC TRAJ, a SAS procedure, fits a discrete mixture model to longitudinal data (Jones et al. 2001). The model groups data to different classes with different parameter values. In contrast to traditional regression which models only one regression function within the population, PROC TRAJ is a specialized mixture model for multiple regression functions within the population. PROC TRAJ does not provide any individual level information. It focuses on class membership and identifying distinct classes. It assumes that every subject in a trajectory class follows the same trajectory.

Model selection in PROC TRAJ uses the Bayesian Information Criterion (BIC) values to select the number of classes in the model. The BIC values given in the output are negative. Typically, researchers select the model with the largest BIC value, often with constraints on minimum group size. The posterior group membership probabilities can be used to explore differences in covariates between classes.

Chapter 3 Methods

My simulation study is designed to examine the performance of the Bayesian Information Criterion (BIC) and related statistics when sampling from homogeneous trajectories. In my simulation study, I use a factorial design on four different families of longitudinal dependent variables. These are the normal distribution, the gamma distribution, the zero-inflated Poisson (ZIP) distribution, and the Bernoulli distribution.

3.1 Factorial experiments

3.1.1 Normal distribution

There are 4 factors in this part of the simulation study: Intra-class correlations, trajectory patterns, sample size and number of time measurements.

Factor 1: Intra-class correlation (ICC)

In the GMM model for homogeneous data,

level 1 models individual change with the model

Level 1: $y_{it} = \pi_{0i} + \pi_{1i} t + \varepsilon_{it}$.

Level 2 models variation between persons according to the model

Level 2: $\pi_{0i} = \beta_{00} + u_{0i} \qquad \pi_{1i} = \beta_{10} + u_{1i}$,

where

y_{it} : Outcome for subject *i*(*i*=1,…n) at time point *t*.

ε_{it}: Within-participant random effects, which are independently and identically distributed as $N(0,\sigma^2)$.

π_{0i} is the intercept outcome for participant *i*, and π_{1i} is the linear rate of change for participant *i* .

The random variables u_{0i} and u_{1i} model individual effects. They are not in the fixed effects model. In a random effects model, they have specified probability distributions.

Here I specify the intra-class correlation= $\frac{\sigma_0^2 + \sigma_1^2 t^2}{\sigma_0^2 + \sigma_1^2 t^2 + \sigma_R^2}$

where *t* is fixed at *t*=1/2.

Two sets of variances were chosen to insure that the two intra-class correlation values are around ¼, ¾ and variances and standard deviations are integers.

Table 1 Value of Intra-class correlation at t=1/2 for Factor 1

Intra-class correlation	σ_0^2	σ_1^2	σ_R^2
Smaller intra-calss correlation value=0.265	9	16	36
larger intra-calss correlation value=0.765	9	16	4

Factor 2: Trajectory Pattern: (linear or quadratic)

I fit the simulated data to either a linear or quadratic trajectory pattern.

Factor 3: Factor 3: Number of time measurements with low setting at 4 points and high at 6 points. With

4 time measurements, $t = 0, \frac{1}{3}, \frac{2}{3}, 1$

With 6 time measurements, $t = 0, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1$

Factor 4: Sample Size: low setting at 400 trajectories and high setting at 800 trajectories.

The general model I use to generate normal distribution data is given in section 3.1 and Table 1. I use y_{it} as the dependent variable.

3.1.2 Gamma distribution

The parameters of the gamma distribution are (*k*/2, 2), where k is the mean of the random variable. For each time point and individual, the mean $k_{it} = y_{it}$, where y_{it} is given in the preceding section.

I generated 100 replicates for each setting of the four factors; that is, setting of of intra-class correlation, distribution of error term (Normal and Gamma), sample size (400 and 800), and number of time measurements (4 and 6). Then they are fit to linear or quadratic trajectory models.

3.1.2 Zero-inflated Poisson distribution (ZIP)

Factor 1: Fraction in the zero component

Here I set the fraction of the zero component to be either constant or varying. When the fraction of the component is constant, the smaller fraction setting is 1/8, and the larger is ½.

When the fraction in the zero component is not constant, the logit of the fraction changes linearly over time as following

$$\log(\frac{\rho_{it}}{1-\rho_{it}}) = \alpha_{0i} + \alpha_{1i}t$$

Factor 2: Trend of the log value of means

I studied two trends: linear and s-shaped.

The model for a linear trend is: $\ln(\lambda_{it}) = \pi_{0i} + \pi_{1i}t$

The model for an s-shaped trend is $\ln(\lambda_{it}) = \dfrac{\exp \pi_{it}}{(1+\exp \pi_{it})^2}$

Factor 3: Fixed effect or random effect

When the effects are fixed, I treat the parameters in factor 2 as constant at the individual level.

When the effects are random, the parameters in level 1 have a level 2 models given by

Level 2: $\pi_{0i} = \beta_{00} + u_{0i}$

$$\pi_{1i} = \beta_{10} + u_{1i}$$

Here, π_{0i} and π_{1i} are latent random variables that vary across individuals and follow a normal distribution.

Factor 4: The sample size

The low setting of sample size is 400, and the high setting is 800.

Factor 5: Trajectory Pattern (linear or quadratic)

I fit the simulated data to either a linear or quadratic trajectory pattern.

I generate 100 replicates for every setting of the four factors and fit each replicate to both a linear and quadratic trajectory. There are in total 48 scenarios.

3.1.3 Bernoulli Distribution

The Bernoulli distribution is used to model dichotomous data, given group membership. The parameter for the dichotomous model is the mean of the Bernoulli variable p_{it}, where

$p_{it} = \frac{\exp(f(i,t))}{1+\exp(f(i,t))}$. Here *i* represents the individual and t the time period.

Factor 1: Trend of *f (i, t)*

I study two trends: linear and s-shaped.

The model for the linear trend is: $f(i,t) = \pi_{0i} + \pi_{1i} t$

The model for an s-shaped trend is $f(i,t) = \frac{\exp \pi_{it}}{(1 + \exp \pi_{it})^2}$

Factor 2: Fixed effects and Random effects

When the effects are fixed, I treat the parameters in factor 2 as constant at the individual level.

When the effects are random, parameters in level 1 have a level 2 model given by

Level 2: $\pi_{0i} = \beta_{00} + u_{0i}$

$$\pi_{1i} = \beta_{10} + u_{1i}$$

Here, π_{0i} and π_{1i} are latent random variables that vary across individuals and follow a normal distribution.

Factor 3: Sample size: 400 and 800.

Factor 4: Trajectory Pattern (linear or quadratic)

I fit the simulated data to either a linear or quadratic trajectory pattern.

I generated 100 replicates for each combination of the 3 factors and fit them to both a linear and quadratic trajectory. There are in total 16 scenarios.

3.2 Data analysis

The performances of three primary criteria-BIC, LMR and BLRT- are compared under each factor. Entropy is used as a supplemental measure. I focus on the rate of identifying one trajectory class since my simulated data are homogeneous.

I ran models with 1 to 4 trajectory classes and record results with non-negative estimated variances of latent variables. When only the1-trajectory class results had no negative variances, I treated the results as indicating a 1-trajectory class model. In the event a simulation scenario had one or more replicates in which all trajectory class models had one or more negative estimated variances, I ran additional replications until I had 100 replicates with at least the 1-trajectory class model having non-negative variances.

Since a BIC value difference of 10 or more is considered as evidence favoring one model over another (Raftery, 1995), I report two sets of BIC rates. The first rate is the fraction of replicates in which the 1-trajectory class model had the best BIC value. The second rate is the fraction of replicates in which the 1-trajectory class was best or within 10 units of the best trajectory class result.

M-plus reports p-values of the Vuong-Lo-Mendell-Rubin likelihood ratio test and the bootstrap likelihood ratio test. I use these p-values sequentially. Specifically, if the p-value for 2 components compared to 1 is greater than the level of significance, I choose 1 component. If this p-value is less than the level of significance, I choose a model with more than one component. . I use two significance levels: 0.05 and 0.01.

Chapter 4 Results

4.1 Number of random starting points (RSPS)

In M-plus, the default number of sets of random starting values is 10. Hipp and Bauer (2006) reported that at least 50 to 100 sets of starting values should be used. I ran a study using 5 replicates and 100, 400, 800, and 1600 RSPs in each replicate to identify the smallest number of RSPs that had maximum value of the likelihood function within 0.1 of the maximum of the likelihood function from 1,600 RSPs. I studied 8 scenarios in three factors: two intra-class correlations (V1=0.265,V2-0.765), two trajectory patterns(Linear(L),Quadratic(Q)), and two types of error terms(Normal(N),Non-normal(G)). I called these 8 scenarios: LNV1, QNV1, LNV2, QNV2, LGV1, QGV1, LGV2, and QGV2 respectively.

The values of the likelihood for the 5 replicates in each of the 8 scenarios are all identical except replicate 5 in scenario LNV2 (linear trajectory pattern, normal distribution, and larger intra-class correlation setting). In replicate 5 of this scenario,, the likelihood using 100 RSPs for 4-classes is 0.977 higher than the maximum using 400, 800, or 1600 RSPs. Therefore, I decided to use 100 RSPs in my simulation study.

4.2 Computational failure rate for normal and gamma distribution

In the event that a replicate has one or more negative variance estimates in the one trajectory class model, the replicate is scored as a computational failure.

I then generated additional replicates until there were 100 replicates without computational failure. I list the total numbers and proportions of extra replicates under each scenario for the normal distribution in Table 2 and for the gamma distribution in Table 3.

Tables 4 and 5 contain logistic regression results when the dependent variable is the computational failure rate and the independent variables are the experimental factors.

Table 2 Computational failure rate statistics for normal distribution data (100 replicates without computational failure)

Factors				Total repliates	Additional replicates	Failure rates
Sample size	Number of time measurements	Trajectory pattern	Intra-class correlation			
400	4	Linear	0.265	102	2	0.02
400	4	Linear	0.765	100	0	0
400	4	Quadratic	0.265	264	164	0.62
400	4	Quadratic	0.765	238	138	0.58
400	6	Linear	0.265	105	5	0.05
400	6	Linear	0.765	100	0	0
400	6	Quadratic	0.265	252	152	0.60
400	6	Quadratic	0.765	193	93	0.48
800	4	Linear	0.265	104	4	0.04
800	4	Linear	0.765	100	0	0
800	4	Quadratic	0.265	258	158	0.61
800	4	Quadratic	0.765	227	127	0.56
800	6	Linear	0.265	100	0	0
800	6	Linear	0.765	100	0	0
800	6	Quadratic	0.265	219	119	0.54
800	6	Quadratic	0.765	163	63	0.39

Note: A replicate has a computation failure if there is at least one negative variance estimate for a one trajectory class model.

Table 3 Computational failure rate statistics for gamma distribution data (100 replicates without computational failure)

Factor				Total replicates	Additional replicates	failure rates
Sample size	Number of time measurements	Trajectory pattern	Intra-class correlation			
400	4	Linear	0.265	151	51	0.34
400	4	Linear	0.765	128	28	0.22
400	4	Quadratic	0.265	294	194	0.66
400	4	Quadratic	0.765	244	144	0.59
400	6	Linear	0.265	124	24	0.19
400	6	Linear	0.765	105	5	0.05
400	6	Quadratic	0.265	247	147	0.6
400	6	Quadratic	0.765	246	146	0.59
800	4	Linear	0.265	125	25	0.20
800	4	Linear	0.765	110	10	0.09
800	4	Quadratic	0.265	253	153	0.60
800	4	Quadratic	0.765	262	162	0.62
800	6	Linear	0.265	108	8	0.07
800	6	Linear	0.765	104	4	0.04
800	6	Quadratic	0.265	259	159	0.61
800	6	Quadratic	0.765	221	121	0.55

Note: A replicate has a computation failure if there is at least one negative variance estimate for a one trajectory class model.

From Tables 2 and 3, the trajectory pattern factor has a significant effect on the computational failure rate. The computational failure rate when fitting to a quadratic pattern is generally higher than the computational failure rate fitting to a linear pattern. The computational failure rate of linear trajectory pattern of normal distributions is close to 0, much lower than the rate when fitting to gamma distribution data which is less than 50%. The computational failure rates fitting to a quadratic pattern are larger than 50% for both distributions. For these cases, I than ran more than 200 replicates to get 100 replicates without computational failure. Computational failure rates for gamma data are generally higher than those for normal data.

Table 4 contains the results of using logistic regression to model the probability of computational failure when analyzing normal data. The trajectory pattern is the most significant factor (Wald chi-square 216.96, p<0.0001), with odds ratio of failure 92 for a quadratic pattern compared to a linear pattern (Appendix Table 23). The intra-class correlation factor (Wald chi-square 17.86, p<0.0001) and the number of time points (Wald chi-square 12.88, p=0.0003) are the next most significant factors. The lower intra-class correlation setting, which is 0.265, is associated with a higher computational failure rate. Four time points has a lower computational failure rate than 6 time points. P-value and odds ratio of the sample size factor show that changing the sample size has no significant effect on the computational failure rate.

Table 4 Analysis of factors contributing normal data computational failure rates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept		1	-0.26	0.10	6.42	0.011
Intra-class correlation	***0.265***	***1***	***0.40***	***0.10***	***17.86***	***<.0001***
Trajectory Pattern	***Linear***	***1***	***-4.53***	***0.31***	***216.96***	***<.0001***
Time measurement	4	1	0.34	0.10	12.88	0.0003
Sample size	400	1	0.18	0.09	3.71	0.054

Note: A replicate has a computation failure if there is at least one negative variance estimate for a one trajectory class model.

Table 5 contains the logistic results for the gamma distributed data. The trajectory pattern is still the most significant factor (Wald chi-square 444.73, p<0.0001). The number of time measurements (Wald chi-square 13.91, p=0.0002) and the intra-class correlation factor (Wald chi-square 10.99, p=0.0009) are the next most significant factors. The sample size factor also has a significant effect on failure rate (Wald chi-square 6.66, p=0.0099). Data with a linear trajectory pattern, having larger intra-class correlations, more time measurements, and larger sample size will have a lower computational failure rate for data from a gamma distribution.

The odds ratio of failure rates for a quadratic pattern compared to a linear pattern is 8.21, the largest odds ratio among the four factors. The odds ratios for the other three factors, while significant, are relatively close to 1. See Appendix Table 24.

Table 5 Analysis of factors contributing gamma data computational failure rates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept		1	0.03	0.08	0.11	0.745
Intra-class correlations	0.265	1	0.27	0.08	10.99	0.0009
Trajectory Pattern	***Linear***	***1***	***-2.11***	***0.1***	***444.73***	***<.0001***
Time measurements	4	1	0.30	0.08	13.91	0.0002
Sample size	400	1	0.21	0.08	6.66	0.01

Note: A replicate has a computation failure if there is at least one negative variance estimate for a one trajectory class model.

I summarize the computational failure rates for the trajectory pattern factor in Table 6, which contains summary statistics about the computation failure rates averaging over sample size, intra-class correlation coefficient, and numbers of time points. The differences of rates between linear and quadratic patterns are substantial. When fitting normal data with a linear trajectory pattern, the failure rates are no more than 0.05 with average 0.01, while they are larger than 0.39 with average 0.55 for quadratic pattern. A similar pattern occurs for gamma data.

Table 6 Summary statistics of computational failure rate for settings of the trajectory pattern factor

Normal	Trajectory pattern	Mean	Std dev	Minimum	Maximum
	Linear	0.01	0.02	0	0.05
	Quadratic	0.55	0.08	0.39	0.63
Gamma	Trajectory pattern	Mean	Std dev	Minimum	Maximum
	Linear	0.15	0.11	0.04	0.34
	Quadratic	0.60	0.03	0.55	0.66

Note: A replicate has a computation failure if there is at least one negative variance estimate for a one trajectory class model.

4.3 Results of Censored Normal Data

Table 7 contains the rates of identifying one trajectory class for normally distributed data for 100 replicates with no computational failure. The first BIC identification rate, called the absolute BIC rule rate, is the fraction of replicates in which the one trajectory class model had smaller BIC value than the BIC for two, three, and four classes. The second BIC identification rate, called the significant BIC rule rate, is the fraction of replicates in which the BIC statistic for one trajectory class was largest or within 10 units of the largest BIC. The two rates indicate that the BIC value is effective at identifying the correct number of trajectories for homogeneous data for normally distributed data. The BIC identification rate is 98% or higher for each setting of the four factors.

Table 7 BIC Rate of identifying one trajectory class for Normal Data

Factors (100 replicates for each scenario)				Absolute BIC rule	Significant BIC rule
Sample size	Number of time measurements	Trajectory pattern	Intra-class correlation		
400	4	Linear	0.265	0.99	0.99
400	4	Linear	0.765	1	1
400	4	Quadratic	0.265	1	1
400	4	Quadratic	0.765	1	1
400	6	Linear	0.265	1	1
400	6	Linear	0.765	1	1
400	6	Quadratic	0.265	1	1
400	6	Quadratic	0.765	1	1
800	4	Linear	0.265	1	1
800	4	Linear	0.765	0.98	1
800	4	Quadratic	0.265	1	1
800	4	Quadratic	0.765	1	1
800	6	Linear	0.265	1	1
800	6	Linear	0.765	1	1
800	6	Quadratic	0.265	1	1
800	6	Quadratic	0.765	1	1

Table 8 reports the rates of selecting a one component model using the LMRT (Lo-Mendell-Rubin likelihood ratio test) rule and the BLRT (Bootstrap likelihood ratio test) rule specified in Chapter 3 at levels of significance 0.05 or 0.01. I also report the 95% confidence intervals for the rates when the significance level is 0.05 and the 99% confidence intervals when the significance level is 0.01.

Table 8 LMRT and BLRT rates of identifying one trajectory class for normal data with significance levels 0.05 & 0.01 (100 replicates for each setting)

Factors				**Significance level:0.05**			
Sample size	Time measure -ments	Trajectory pattern	Intra-class correlation	LMRT	Confidence intervals	BLRT	Confidence intervals
400	4	Linear	0.265	0.94	[0.93 0.95]	0.87	[0.85 0.89]
400	4	Linear	0.765	0.78	[0.75 0.81]	0.79	[0.76 0.82]
400	4	Quadratic	0.265	0.89	[0.87 0.91]	0.96	[0.95 0.97]
400	4	Quadratic	0.765	0.9	[0.88 0.92]	0.99	[0.99 0.99]
400	6	Linear	0.265	0.88	[0.86 0.90]	0.93	[0.92 0.94]
400	6	Linear	0.765	0.83	[0.80 0.86]	0.78	[0.75 0.81]
400	6	Quadratic	0.265	0.9	[0.88 0.92]	0.92	[0.91 0.93]
400	6	Quadratic	0.765	0.93	[0.92 0.94]	0.94	[0.93 0.95]
800	4	Linear	0.265	0.94	[0.93 0.95]	0.97	[0.96 0.98]
800	4	Linear	0.765	0.87	[0.85 0.89]	0.93	[0.92 0.94]
800	4	Quadratic	0.265	0.88	[0.86 0.90]	0.99	[0.99 0.99]
800	4	Quadratic	0.765	0.83	[0.80 0.86]	0.96	[0.95 0.97]
800	6	Linear	0.265	0.91	[0.89 0.93]	0.93	[0.92 0.94]
800	6	Linear	0.765	0.87	[0.85 0.89]	0.82	[0.79 0.85]
800	6	Quadratic	0.265	0.93	[0.92 0.94]	0.98	[0.98 0.98]
800	6	Quadratic	0.765	0.93	[0.92 0.94]	0.96	[0.95 0.97]
Factors				**Significance level:0.01**			
Sample size	Time measure -ments	Trajectory pattern	Intra-class correlation	LMRT	Confidence intervals	BLRT	Confidence intervals
400	4	Linear	0.265	0.98	[0.97 0.99]	0.91	[0.89 0.93]
400	4	Linear	0.765	0.95	[0.94 0.96]	0.87	[0.84 0.90]
400	4	Quadratic	0.265	0.93	[0.91 0.95]	0.97	[0.96 0.98]
400	4	Quadratic	0.765	0.96	[0.95 0.97]	0.99	[0.99 0.99]
400	6	Linear	0.265	0.92	[0.90 0.94]	0.93	[0.91 0.95]
400	6	Linear	0.765	0.93	[0.91 0.95]	0.78	[0.73 0.83]
400	6	Quadratic	0.265	0.95	[0.94 0.96]	0.92	[0.90 0.94]
400	6	Quadratic	0.765	0.96	[0.95 0.97]	0.94	[0.93 0.95]
800	4	Linear	0.265	1	[1.00 1.00]	0.97	[0.96 0.98]
800	4	Linear	0.765	0.94	[0.93 0.95]	0.94	[0.93 0.95]
800	4	Quadratic	0.265	0.95	[0.94 0.96]	0.99	[0.99 0.99]
800	4	Quadratic	0.765	0.95	[0.94 0.96]	0.96	[0.95 0.97]
800	6	Linear	0.265	0.97	[0.96 0.98]	0.93	[0.91 0.95]
800	6	Linear	0.765	0.97	[0.96 0.98]	0.9	[0.88 0.92]
800	6	Quadratic	0.265	0.96	[0.95 0.97]	0.98	[0.97 0.99]
800	6	Quadratic	0.765	0.96	[0.95 0.97]	0.99	[0.99 0.99]

LMRT: Lo-Mendell-Rubin likelihood ratio test   BLRT: bootstrap likelihood ratio test

At significance level 0.05, the target correct identification rate is 0.95. At significance level 0.01, the target correct identification rate is 0.99. The average one component selection rate for LMRT was 0.89 at the 0.05 level and 0.96 at 0.01 level. The average one component selection rate for BLRT was 0.92 at the 0.05 level and 0.94 at the 0.01 level. Both correct identification rates are smaller than the target rates, possibly reflecting the multiple testing implicit in choosing among three hypotheses; namely 1 component vs. 2, 2 components vs. 3, and 3 components vs. 4.

The probabilities of identifying one trajectory class with significance levels 0.05 and 0.01 using LMRT or BLRT are modeled in Tables 9 and 10. In Table 9, for the LMRT using significance level 0.05 for normally distributed data, the setting of the random component is the only significant factor, with the 0.265 setting having the higher probability of selecting one component. The odds ratio for intra-class correlations using significance level 0.05 is 1.52, with a confidence interval that excludes 1. The confidence intervals for the odds ratio of the other factors include 1(Appendix Table 26).

No factor was significant for significance level 0.01 with the LMRT on normal data.

Table 9 Analysis of factors determining correct identification rate of a single trajectory of normal data based on LMRT with significance level 0.05 and 0.01

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
			significance level 0.05			
Intercept		1	2.16	0.18	142.81	<.0001
Intra-class correlation	***0.265***	***1***	***0.42***	***0.16***	***6.80***	***0.009***
Trajectory Pattern	Linear	1	-0.22	0.16	1.82	0.177
Time measurements	4	1	-0.19	0.16	1.42	0.233
Sample size	400	1	-0.14	0.16	0.77	0.382
			significance level 0.01			
Intercept		1	3.08	0.27	127.95	<.0001
Intra-class correlation	0.265	1	0.12	0.24	0.23	0.629
Trajectory Pattern	Linear	1	0.12	0.24	0.23	0.629
Time measurements	4	1	0.12	0.24	0.23	0.629
Sample size	400	1	-0.35	0.25	2.08	0.15

LMRT: Lo-Mendell-Rubin likelihood ratio test

Table 10 shows the logistic results for the BLRT identification rate using the 0.05 and 0.01 significance levels.

The identification rate using the BLRT at both significance levels is associated with the trajectory pattern (Wald chi-square with 0.05 level=35.87, $p<0.0001$). The identification rate with the quadratic trajectory pattern is higher than the rate for the linear trajectory pattern. The sample size and intra-class correlation are the next most significant effects for the BLRT at 0.05 significance level. Larger sample size results in higher correct identification rate at both the 0.05 and 0.01 significance levels. The intra-class correlation factor is not significant for the BLRT at the 0.01 level of significance. At the 0.05 level, the lower intra-class correlation coefficient is associated with a higher correct identification rate.

Table 10 Analysis of factors determining correct identification rate of a single trajectory of normal data based on BLRT with significance level 0.05 and 0.01

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Significance level=0.05						
Intercept		1	3.16	0.25	163.14	<.0001
Intra-class correlation	0.265	1	0.69	0.20	12.36	0.0004
Trajectory Pattern	***Linear***	***1***	***-1.30***	***0.22***	***35.87***	***<.0001***
Time measurements	4	1	0.36	0.19	3.52	0.061
Sample size	400	1	-0.65	0.20	11.14	0.0008
Significance level=0.01						
Intercept		1	3.41	0.27	155.50	<.0001
Intra-class correlation	0.265	1	0.50	0.21	5.58	0.018
Trajectory Pattern	***Linear***	***1***	***-1.17***	***0.23***	***25.01***	***<.0001***
Time measurements	4	1	0.50	0.21	5.58	0.018
Sample size	400	1	-0.77	0.22	12.55	0.0004

BLRT: bootstrap likelihood ratio test

Compared to the linear pattern, the odds ratio for the quadratic pattern is 3.66 for the BLRT using the 0.05 significance level and 3.23 using the 0.01 significance level. The odds ratios for trajectory patterns are the largest among the four factors. The odds ratio for the intra-class correlations decreases from 1.99 to 1.65 while odds ratio for sample size increase from 1.92 to 2.17 with lower significance level.

LMRT and BLRT rates of identifying one trajectory class for censored normal data increase when the significance level decreases, as documented in Table 8.

As shown in Table 11, the identification rate using the LMRT is different from the BLRT rate using McNemar's Test.

In Table 11, with significance level 0.05, there are five BLRT rates that are significantly higher than the LMRT rates, especially with the quadratic trajectory pattern. There is one scenario (0.265 intra-class correlation, linear trajectory, 4 time measurements, and sample size 400) in which LMRT has a higher identification rate. As expected, decreasing significance level improves the identification rates of both the LMRT and the BLRT.

For significance level 0.01, four LMRT identification rates appear to be higher than rates using the BLRT, with no scenario having the BLRT rate significantly higher than the LMRT rate. That is, the identification rate of the LMRT seems to be better than the rate of the BLRT with significance level 0.01.

Table 11 McNemar's Test comparing LMRT and BLRT correct identification rates with significance level 0.05&0.01

Factors				Significance level:0.05	
Sample size	Time measurements	Trajectory pattern	Intra-class correlation	Difference LMRT-BLRT	p-value
400	**4**	**Linear**	**0.265**	***0.07***	***0.0156***
400	4	Linear	0.765	-0.01	1
400	**4**	**Quadratic**	**0.265**	***-0.07***	***0.0156***
400	**4**	**Quadratic**	**0.765**	***-0.09***	***0.0039***
400	6	Linear	0.265	-0.05	0.0625
400	6	Linear	0.765	0.05	0.0625
400	6	Quadratic	0.265	-0.02	0.5
400	6	Quadratic	0.765	-0.01	1
800	4	Linear	0.265	-0.03	0.25
800	**4**	**Linear**	**0.765**	***-0.06***	***0.0313***
800	**4**	**Quadratic**	**0.265**	***-0.11***	***0.001***
800	**4**	**Quadratic**	**0.765**	***-0.13***	***0.0002***
800	6	Linear	0.265	-0.02	0.5
800	6	Linear	0.765	0.05	0.0625
800	6	Quadratic	0.265	-0.05	0.0625
800	6	Quadratic	0.765	-0.03	0.25
Factors				**Significance level:0.01**	
Sample size	Time measurements	Trajectory pattern	Intra-class correlation	Difference LMRT-BLRT	p-value
400	**4**	**Linear**	**0.265**	***0.07***	***0.0156***
400	**4**	**Linear**	**0.765**	***0.08***	***0.0078***
400	4	Quadratic	0.265	-0.04	0.125
400	4	Quadratic	0.765	-0.03	0.25
400	6	Linear	0.265	-0.01	1
400	**6**	**Linear**	**0.765**	***0.15***	***0.0001***
400	6	Quadratic	0.265	0.03	0.25
400	6	Quadratic	0.765	0.02	0.5
800	4	Linear	0.265	0.03	0.25
800	4	Linear	0.765	*	*
800	4	Quadratic	0.265	-0.04	0.125
800	4	Quadratic	0.765	-0.01	1
800	6	Linear	0.265	0.04	0.125
800	**6**	**Linear**	**0.765**	***0.07***	***0.0156***
800	6	Quadratic	0.265	-0.02	0.5
800	6	Quadratic	0.765	-0.03	0.25

*: cannot perform test

LMRT: Lo-Mendell-Rubin likelihood ratio test BLRT: bootstrap likelihood ratio test

4.4 Results of Gamma Distribution data

Table 12 reports the BIC identification rate for data following the gamma distribution using the format of Table 7. The absolute BIC correct identification rates are lower than the significant BIC correct identification rates. The absolute BIC rule has only 3 identification rates smaller than 0.95. While changing the distribution of data from normal to gamma reduces the identification rate for both BIC rules, the BIC identification rates are high. The BIC identification rates are significantly affected by the intra-class correlation factor and time measurement factor (Appendix Table 28).

Table 12 BIC Rate of identifying one trajectory class for Gamma Data

Factors				Absolute BIC rule	Significant BIC rule
Sample size	Time measurements	Trajectory pattern	Intra-class correlation		
400	4	Linear	0.265	0.99	1
400	4	Linear	0.765	0.98	1
400	4	Quadratic	0.265	0.93	0.99
400	4	Quadratic	0.765	0.87	0.99
400	6	Linear	0.265	1	1
400	6	Linear	0.765	0.98	1
400	6	Quadratic	0.265	0.99	1
400	6	Quadratic	0.765	0.96	1
800	4	Linear	0.265	1	1
800	4	Linear	0.765	0.96	0.98
800	4	Quadratic	0.265	1	1
800	4	Quadratic	0.765	0.82	0.9
800	6	Linear	0.265	0.98	1
800	6	Linear	0.765	0.98	1
800	6	Quadratic	0.265	0.99	1
800	6	Quadratic	0.765	0.96	0.99

Table 13 contains the identification rates for the LMRT and BLRT using both the 0.05 and 0.01 levels of significance, following the format of Table 8.

The effect of changing significance levels improves LMRT performance as expected. There are two rates that are larger than 0.95 at significance level 0.05 and six rates larger than or equal to 0.95 at significance level 0.01.

The average one component selection rate for LMRT is 0.90 at the 0.05 level and 0.93 at the 0.01 level. The average one component selection rate for BLRT is 0.86 at the 0.05 level and 0.88 at the 0.01 level.

Table 13 LMRT and BLRT rates of identifying one trajectory class for gamma data with significance levels 0.05 & 0.01(100 replicates for each setting)

Factors				**Significance level:0.05**			
Sample size	Time measure -ments	Trajectory pattern	Intra-class correlation	LMRT	Confidence intervals	BLRT	Confidence intervals
400	4	Linear	0.265	0.93	[0.92 0.94]	0.93	[0.92 0.94]
400	4	Linear	0.765	0.91	[0.89 0.93]	0.91	[0.89 0.93]
400	4	Quadratic	0.265	0.84	[0.81 0.87]	0.78	[0.75 0.81]
400	4	Quadratic	0.765	0.82	[0.79 0.85]	0.72	[0.68 0.76]
400	6	Linear	0.265	0.88	[0.86 0.90]	0.9	[0.88 0.92]
400	6	Linear	0.765	0.88	[0.86 0.90]	0.9	[0.88 0.92]
400	6	Quadratic	0.265	0.97	[0.96 0.98]	0.96	[0.95 0.97]
400	6	Quadratic	0.765	0.94	[0.93 0.95]	0.83	[0.80 0.86]
800	4	Linear	0.265	0.91	[0.89 0.93]	0.88	[0.86 0.90]
800	4	Linear	0.765	0.89	[0.87 0.91]	0.89	[0.87 0.91]
800	4	Quadratic	0.265	0.9	[0.88 0.92]	0.87	[0.85 0.89]
800	4	Quadratic	0.765	0.8	[0.77 0.83]	0.73	[0.69 0.77]
800	6	Linear	0.265	0.89	[0.87 0.91]	0.9	[0.88 0.92]
800	6	Linear	0.765	0.87	[0.85 0.89]	0.79	[0.76 0.82]
800	6	Quadratic	0.265	0.96	[0.95 0.97]	0.89	[0.87 0.91]
800	6	Quadratic	0.765	0.94	[0.93 0.95]	0.87	[0.85 0.89]
Factors				**Significance level:0.01**			
Sample size	Time measure -ments	Trajectory pattern	Intra-class correlation	LMRT	Confidence intervals	BLRT	Confidence intervals
400	4	Linear	0.265	0.94	[0.93 0.95]	0.95	[0.94 0.96]
400	4	Linear	0.765	0.93	[0.91 0.95]	0.95	[0.94 0.96]
400	4	Quadratic	0.265	0.87	[0.84 0.90]	0.85	[0.82 0.88]
400	4	Quadratic	0.765	0.85	[0.82 0.88]	0.72	[0.67 0.77]
400	6	Linear	0.265	0.95	[0.94 0.96]	0.91	[0.89 0.93]
400	6	Linear	0.765	0.95	[0.94 0.96]	0.92	[0.90 0.94]
400	6	Quadratic	0.265	0.97	[0.96 0.98]	0.96	[0.95 0.97]
400	6	Quadratic	0.765	0.94	[0.93 0.95]	0.83	[0.79 0.87]
800	4	Linear	0.265	1	[1.00 1.00]	0.96	[0.95 0.97]
800	4	Linear	0.765	0.94	[0.93 0.95]	0.9	[0.88 0.92]
800	4	Quadratic	0.265	0.95	[0.94 0.96]	0.89	[0.86 0.92]
800	4	Quadratic	0.765	0.85	[0.82 0.88]	0.74	[0.75 0.83]
800	6	Linear	0.265	0.95	[0.94 0.96]	0.9	[0.88 0.92]
800	6	Linear	0.765	0.92	[0.90 0.94]	0.79	[0.75 0.83]
800	6	Quadratic	0.265	0.97	[0.96 0.98]	0.89	[0.86 0.92]
800	6	Quadratic	0.765	0.94	[0.93 0.95]	0.87	[0.84 0.90]

Table 14 contains the logistic results for the LMRT correct identification rate for gamma data.

Only the time measurement factor has a significant effect on the correct identification rate of LMRT set at the 0.05 level of significance (Wald Chi-square=7.21, p=0.0073). More time measurements increase the correct identification rate. For the LMRT set at significance level 0.01, the intra-class correlation factor is the most significant. Smaller intra-class correlation increases the identification rate of the LMRT (Wald Chi-square=7.69, p=0.0056). Trajectory pattern and sample size are not significant factors when the LMRT is set at either the 0.05 or 0.01 level.

Table 14 Analysis of factors determining correct identification rate of a single trajectory of gamma data based on LMRT with significance level 0.05 and 0.01

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
		Significance level=0.05				
Intercept		1	2.25	0.19	144.40	<.0001
Intra-class correlation	0.265	1	0.31	0.17	3.53	0.060
Trajectory Pattern	Linear	1	-0.01	0.16	0.0	0.935
Time measurements	***4***	***1***	***-0.45***	***0.17***	***7.21***	***0.007***
Sample size	400	1	0.01	0.16	0.01	0.935
		Significance level=0.01				
Intercept		1	2.58	0.23	130.76	<.0001
Intra-class correlation	***0.265***	***1***	***0.57***	***0.21***	***7.69***	***0.006***
Trajectory Pattern	Linear	1	0.49	0.21	5.69	0.017
Time measurements	4	1	-0.53	0.21	6.65	0.01
Sample size	400	1	-0.24	0.20	1.44	0.23

The odds ratio for each of the four factors is further from 1 when the LMRT level of significance is 0.01, compared to results when the LMRT level is 0.05.

For example, the odds ratio of number of time points is 1.56 with LMRT significance level set to 0.05 compared to 1.70 for significance level 0.01. Similarly, the odds ratio of intra-class correlations set at 0.265 changes from 1.36 for significance level 0.05 to 1.77 at significance level 0.01.

Table 15 contains the logistic regression results for the correct identification rate of the BLRT (Bootstrap likelihood ratio test) with significance level set to either 0.05 or 0.01.

For the BLRT at either significance level 0.05 or 0.01, the intra-class correlations (Wald Chi-square=11.40, p=0.0007 at 0.05 level, Wald Chi-square=19.82, p <0.0001 at 0.01 level) and the trajectory pattern factor (Wald Chi-square=10.47, p=0.0012 at 0.05 level, Wald Chi-square=16.13, p<0.0001 at 0.01 level) are strongly associated with the correct identification rate when analyzing data with a gamma distribution. The odds ratio of intra-class correlations 0.265 versus 0.765 is 1.65 at the 0.05 level and 2.03 at 0.01 level. (Appendix Table 30) The odds ratio of the linear versus the quadratic pattern changes from 1.61 at 0.05 level to 1.89 at 0.01 level.

For data following a gamma distribution, smaller random component and linear trajectory pattern are associated with higher BLRT correct identification rate. The number of time measurements and the sample size are not significant.

Table 15 Analysis of factors determining correct identification rate of a single trajectory of gamma data based on BLRT with significance level 0.05 and 0.01

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Significance level=0.05						
Intercept		1	1.49	0.16	91.97	<.0001
Intra-class correlations	***0.265***	***1***	***0.50***	***0.15***	***11.40***	***0.0007***
Trajectory Pattern	***Linear***	***1***	***0.48***	***0.15***	***10.47***	***0.0012***
Time measurements	4	1	-0.35	0.15	5.68	0.017
Sample size	400	1	0.12	0.15	0.64	0.425
Significance level=0.01						
Intercept		1	1.36	0.16	73.23	<.0001
Intra-class correlations	***0.265***	***1***	***0.71***	***0.16***	***19.82***	***<.0001***
Trajectory Pattern	***Linear***	***1***	***0.64***	***0.16***	***16.13***	***<.0001***
Time measurements	4	1	-0.13	0.15	0.72	0.397
Sample size	400	1	0.18	0.15	1.33	0.248

For data with the gamma distribution, the correct identification rate of the LMRT performance is often higher than the correct identification rate of BLRT at both significance levels according to McNemar's Test as shown in Table 16.

Seven LMRT identification rates are significantly larger than the BLRT identification rate using the setting of significance level to either 0.05 or 0.01. For either significance level, there is no scenario in which the rate for the BLRT is significantly higher than the rate for LMRT.

Table 16 McNemar's Test comparing LMRT and BLRT identifying rates with significance level 0.05 and 0.01

Factors				Significance level:0.05	
Sample size	Time measure-ments	Trajectory pattern	Intra-class correlation	Difference LMRT-BLRT	p-value
400	4	Linear	0.265	0	*
400	4	Linear	0.765	0	*
400	4	Quadratic	0.265	***0.06***	***0.0313***
400	4	Quadratic	0.765	***0.1***	***0.002***
400	6	Linear	0.265	-0.02	0.5
400	6	Linear	0.765	-0.02	0.5
400	6	Quadratic	0.265	0.01	1
400	6	Quadratic	0.765	***0.11***	***0.001***
800	4	Linear	0.265	0.03	0.25
800	4	Linear	0.765	0	*
800	4	Quadratic	0.265	0.03	0.25
800	4	Quadratic	0.765	***0.07***	***0.0156***
800	6	Linear	0.265	-0.01	1
800	6	Linear	0.765	***0.08***	***0.0078***
800	6	Quadratic	0.265	***0.07***	***0.0156***
800	6	Quadratic	0.765	***0.07***	***0.0156***
Factors				**Significance level:0.01**	
Sample size	Time measure-ments	Trajectory pattern	Intra-class correlation	Difference LMRT-BLRT	p-value
400	4	Linear	0.265	-0.01	1
400	4	Linear	0.765	-0.02	0.5
400	4	Quadratic	0.265	0.02	0.5
400	4	Quadratic	0.765	***0.13***	***0.0002***
400	6	Linear	0.265	0.04	0.125
400	6	Linear	0.765	0.03	0.25
400	6	Quadratic	0.265	0.01	1
400	6	Quadratic	0.765	***0.11***	***0.001***
800	4	Linear	0.265	0.04	0.125
800	4	Linear	0.765	0.04	0.125
800	4	Quadratic	0.265	***0.06***	***0.0313***
800	4	Quadratic	0.765	***0.11***	***0.001***
800	6	Linear	0.265	0.05	0.0625
800	6	Linear	0.765	***0.13***	***0.0002***
800	6	Quadratic	0.265	***0.08***	***0.0078***
800	6	Quadratic	0.765	***0.07***	***0.0156***

*: cannot perform test

LMRT: Lo-Mendell-Rubin likelihood ratio test BLRT: bootstrap likelihood ratio test

4.5 Zero inflated Poisson (ZIP) data

I studied three selection criteria: BIC, LMRT, and BLRT. I chose the optimal number of class using the significant BIC rule; that is, the difference in BIC scores must be at least 10 to choose more than one group. The significance level for the LMRT and for the BLRT was set to 0.01.

There were 5 factors with 48 scenarios in my factorial design for ZIP data. I ran 100 replicates for each scenario. The rates of identifying one trajectory class given that the ZIP data are homogeneous were zero for the three criteria. I computed the average of the number of components selected for each scenario using the three criteria For example, suppose that a scenario had the result that the LMRT procedure chose a 2-trajectory model twenty times and a 3-trajectory model eighty times. Then, the average number of components selected for this scenario is 2.8. (2*0.2+3*0.8=2.8). I listed these averages using the BIC, LMRT and BLRT for each scenario in Appendix Table 31.

I compared the performance of BIC, LMRT and BLRT using a randomized block analysis. That is, the 48 scenarios were blocks and the BIC, LMRT, BLRT criteria were treatments.

The analysis of variance documented that the average of the number of components selected was significantly different among the different scenarios and that the number of components selected by the three criteria are different.

Tukey's multiple comparisons test indicated that the average number of components selected by BIC, LMRT and BLRT were significantly different from each other at the 0.0001 level of significance.

The BIC criterion had the lowest number of components selected on average. BIC chose two classes in 50% of the 48 scenarios and three classes in the other 50% of scenarios so that the mean number of classes selected was 2.5. The number of classes selected by the LMRT was 3 for 36 out of 48 scenarios. The mean of LMRT number of classes is three with minimum two and maximum four classes. BLRT chose 4-class results in most of the scenarios and had mean number of selected classes 3.5.

Table 17 Analysis of ZIP average number of selected components using BIC, LMRT and BLRT with M-plus: means comparisons

Least Squares Means		**Adjustment for Multiple Comparisons: Tukey**			
Treatments	Averages	LSMean number (i)	Std dev	min	max
BIC	2.5	1	0.505	2	3
LMRT	2.96	2	0.505	2	4
BLRT	3.5	3	0.546	2	4

Least Squares Means for effect treatment		**Pr > |t| for H0: LSMean(i)=LSMean(j)**	
i/j	BIC	LMRT	BLRT
BIC		<.0001	<.0001
LMRT	<.0001		<.0001

LMRT: Lo-Mendell-Rubin likelihood ratio test BLRT: bootstrap likelihood ratio test

Since the rates of identifying one Trajectory class in M-plus were zero for each scenarios, I ran the same replicates using PROC TRAJ. PROC TRAJ assumes that every subject in a trajectory group follows the same trajectory. The PROC TRAJ rates of identifying one trajectory class were always 100% except for the scenario of linear trend data with random effect and 1/8 proportion of extra zeroes or changing proportion of extra zeros.

Table 18 contains the scenarios for which the identification rates were not equal to one using the significant BIC rule. The effect of sample size was substantial. The correct identification rate for linear trend data with random effect and sample size of 400 were more than 0.8 while is the rate was less than 0.5 with sample size 800.

Table 18 Partial table of correct identification rates for ZIP data using PROC TRAJ for scenarios with correct identification rates less than 1

scenarios				Fit to	Significant BIC rule
1/8	Linear trend	Random effect	400	linear	0.85
1/8	Linear trend	Random effect	400	Quadratic	0.91
ρ_{it}	Linear trend	Random effect	400	linear	0.82
ρ_{it}	Linear trend	Random effect	400	Quadratic	0.88
1/8	Linear trend	Random effect	800	linear	0.18
1/8	Linear trend	Random effect	800	Quadratic	0.3
ρ_{it}	Linear trend	Random effect	800	linear	0.22
ρ_{it}	Linear trend	Random effect	800	Quadratic	0.3

Table 19 contains the results of the logistic regression analysis for ZIP data with BIC criterion using PROC TRAJ. The significant factors were S-shaped trend as opposed to linear trend (Wald Chi-Square=152.74, p<0.0001, odds ratio=190.96), sample size 400 compared to 800 (Wald Chi-Square=236.29, p<0.0001, odds ratio=16.18 (Appendix Table 32), ½ fraction of extra zeros compared to 1/8 and varying proportions (Wald Chi-Square=109.18, p<0.0001, odds ratio=95.54), and fixed effect (Wald Chi-Square=49.16, p<0.0001, odds ratio>999.999). That is, the fraction of replicates identified as having one trajectory group was higher for data with S-shaped trend than for data with a linear trend, was higher with fixed effect than random effect data, higher with sample size 400 than sample size 800, and higher with proportion of extra zeroes set to ½ compared to other proportions.

Table 19 Analysis of factors determining correct identification rate of one trajectory of ZIP data based on significant BIC rate

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept		1	4.05	0.43	90.4	<.0001
Proportion of extra zeros	***1/2***	***1***	***4.56***	***0.44***	***109.18***	***<.0001***
Proportion of extra zeros	1/8	1	0.03	0.17	0.03	0.863
Trend	***linear***	***1***	***-5.25***	***0.43***	***152.74***	***<0.0001***
Effects	***Fixed***	***1***	***7.06***	***1.01***	***49.16***	***<0.0001***
Sample size	***400***	***1***	***2.78***	***0.18***	***236.29***	***<0.0001***
Trajectory pattern	Quadratic	1	0.43	0.17	6.31	0.012

4.6 Bernoulli Distribution

I used the same method to compute and analyze the number of components selected using Bernoulli data. There were 100 replicates for each scenario. The number of scenarios was 16.

. The correct identification rates for Bernoulli data using BIC, LMRT and BLRT were always zero using the M-plus statistic package (Appendix Table 33). Table 20 contains the results from the complete block analysis. The average number selected was significantly different for the three criteria.

Table 20 Analysis of Bernoulli average number of selected components using BIC, LMRT and BLRT with M-plus: means comparisons

Least Squares Means		**Adjustment for Multiple Comparisons: Tukey**			
Treatments	Averages	LSMean number (i)	Std dev	min	max
BIC	2.63	1	0.719	2	4
LMRT	3.19	2	0.403	3	4
BLRT	3.31	3	0.602	2	4
Least Squares Means for effect treatment Pr > |t| for H0: LSMean(i)=LSMean(j)					
i/j	BIC	LMRT		BLRT	
BIC		<.0001		0.0007	
LMRT	<.0001			0.6285	
BLRT	0.0007	0.6285			

LMRT: Lo-Mendell-Rubin likelihood ratio test BLRT: bootstrap likelihood ratio test

Tukey's multiple comparison procedure indicated, at the 0.001 significance level; the average numbers of selected components using BIC was significantly less than the number selected using LMRT and BLRT. The numbers selected by LMRT and BLRT were not significantly different.

Due to the low correct identification rate, I again ran the same replicates using PROC TRAJ. The results showed that BIC is a good indicator for Bernoulli data except for linear trend data with random effect and 800 sample size. The optimal choice was always 1 trajectory class except for the two scenarios listed in Table 22. None of the factors appeared to have significant effect on BIC correct identification rate. (Appendix Table 34)

Table 21 Partial table of correct identification rates for Bernoulli using PROC TRAJ for scenarios with correct identification rates less than 1

Scenario			Fit to	BIC	BIC (Dif>10)
Linear trend	Random effect	800	linear	0.05	0.05
Linear trend	Random effect	800	Quadratic	0	0

Chapter 5 Conclusions

My simulation study examined the performance of the BIC, LMRT and BLRT statistics for identifying one trajectory class when the data were homogenous.

Computational failure rate

With the M-plus program, there was a high probability of computational failure for both normal and gamma distributed data. The level of intra-class correlation, trajectory pattern and time measurements were significant factors related to the computational failure rate.

A decreased computational failure rate was associated with data having higher intra-class correlations, linear trajectory pattern, and more time measurements. As expected, a larger sample size was associated with decreasing computational failure rate. Compared to the normal distribution, gamma distribution data had a larger computational failure rate in every scenario. When the estimated variances are negative and not significant, the M-plus creators suggest fixing them to zero (Muthén L.K. M-plus Discussion, 2007). Otherwise, the original model should be replaced.

Comparison of the BIC and likelihood ratio test

Overall, the BIC statistic had the highest correct identification rate. These rates are of the order of 95% for homogeneous data following either a censored normal or gamma distribution. For normal and gamma distribution data, the BIC rate for correctly identifying one trajectory class is greater than the rate using the LMRT or the BLRT as calculated in M-plus. The four factors considered appeared to have no significant effect on the BIC performance for normally distributed data. I randomly chose 10 replicates without computational failure and analyzed them using PROC TRAJ. The BIC identification rates results were identical to the M-plus results.

None of the three criteria correctly identified that there is only one trajectory class for ZIP and Bernoulli data. Table 18 showed that the BIC was a better indicator than the LMRT and the BLRT in the sense that it selected models with fewer trajectory components. The BIC identified 2-classes in twenty four of the scenarios considered and 3-classes as optimal results for the remaining twenty four scenarios of ZIP data. The LMRT chose a 3-class model as the most optimal result in 36 (out of 48) scenarios, and the BLRT chose a 4-class model as the most optimal result in 25 (out of 48) scenarios. As a special case of ZIP, I ran 10 replicates of Poisson data (the probability of the zero component equals zero). The results were similar to other ZIP replicates. The correct identification rates for Poisson data were zero using M-plus. PROC TRAJ provided more accurate BIC results for ZIP.

For Bernoulli data, the average numbers of trajectory classes selected for 16 scenarios using the three criteria equaled 3 after rounding. Using the BIC I observed significantly fewer components compared to the LMRT and the BLRT. The LMRT and the BLRT chose the same number of trajectory components on average. PROC TRAJ produced better BIC results. The correct identification rate is 1 except for linear Bernoulli data with random effect.

Factor effects and comparison of LMRT and BLRT

In general, the LMRT was less affected by the four factors considered than the BLRT for normal and gamma data. The LMRT tended to choose a smaller number of selected components than the BLRT for ZIP data. Consequently, I conclude that the performance of the LMRT is somewhat better than the BLRT.

Table 22 Summary of significant factor of LMRT and BLRT for normal and gamma data

	Normal		Gamma	
	0.05	0.01	0.05	0.01
LMRT	Intra-class correlations	None	Number of time measurements	Intra-class correlations
BLRT	Trajectory pattern	Trajectory pattern	Intra-class correlations	Intra-class correlations
	Intra-class correlations Sample size	Sample size	Trajectory pattern	Trajectory pattern

The setting of the intra-class correlations factor was the only significant factor on LMRT with significance level 0.05 for normal data. At significance level 0.01, none of the four factors was significant for the correct identification rate of LMRT for normal data. For gamma data, time measurements were the only significant factor for the correct identification rate of the LMRT at significance level 0.05. The intra-class correlation factor was also significant for the correct identification rate of the LMRT at significance level 0.01. Lower intra-class correlations and higher time measurements increased the correct identification rate of LMRT.

The trajectory pattern factor was the significant factor for the BLRT correct identifying rate at each significance level for both distributions of data.

For normally distributed data, trajectory pattern, intra-class correlations and sample size were significant factors for BLRT correct identifying rate at significance level 0.05. The intra-class correlation factor was not significantly associated with the correct identification rate of the BLRT with significance level setting 0.01. Quadratic pattern, lower intra-class correlation and larger sample size were associated with higher correct identification rate of the BLRT with significance level set at both 0.05 and 0.01.

For gamma distribution data, intra-class correlations and trajectory patterns were significant factors. Lower intra-class correlations and linear trajectory pattern increased BLRT correct identification rate.

According to McNemar's test, the correct identification rate of the LMRT is higher than rate of the BLRT for normal distribution data at significance level 0.01 and gamma distribution data at both significance levels. For ZIP and Bernoulli data, none of the tests considered correctly identified that there was one trajectory class. The average number of the selected number of components for the LMRT was lower than that of the BLRT for both ZIP and Bernoulli distributions with lower standard deviation. The BLRT usually chose 4 trajectory components the most frequently for the ZIP distributed data.

In conclusion, it is better to use either the BIC or LMRT with significance setting at 0.01 as the model selection criterion. For normal and gamma distribution, a higher level of intra-class correlations was associated with a decreased failure rate but also a decreased correct identification rate. A linear trajectory pattern was associated with lowered the failure rate and lowered correct identification rate for data with normal distribution errors. None of the four factors effects were consistent across failure rate, normal distribution and gamma distribution data.

I suggest PROC TRAJ instead of M-plus for ZIP and Bernoulli data. For the ZIP and Bernoulli distribution, PROC TRAJ computations had a higher correct identification rate than those from M-plus. Larger sample size was associated with an increase in the probability that two or more components will be identified for ZIP distributed data following a linear trend and with random effects. The same pattern held for Bernoulli data.

Future Work

My research studied the null hypothesis of various model selection criteria using four types of distributions. I observed that there were computational failure rates for normal and gamma data as evidenced by the negative estimated variances. As suggested in M-plus Online Discussion, if the variances are negative and not significant, they can be set to zero. However, what can we do if the estimated variances are significant and negative?

My proposed future research would investigate whether application of the Box-Cox transformation will help to improve the model selection precision. I ran a pilot study using normally distributed data with four time measurements and 800 sample size. The lamda values I used were 0.86 and -1. For my randomly selected 10 effective replicates, I used the maximum likelihood estimation presented in Box&Cox (1964) and identified the lamda value of 0.86. The identification rates were identical to the original data. However, the -1 value of lamda changed some of the 10 replicates results to ineffective ones with negative estimated variances.

In my future study, I will also study the property of model selection criteria for that alternative hypothesis that there are two or more components with unequal probabilities? It would also be interest to see if allowing for the mixture of binomials makes the detection of components more difficult.

My future study will also examine ordinal data besides the ZIP and Bernoulli data with growth mixture models. The correct identification rates for ZIP and Bernoulli data were zero using M-plus but 1 for most scenarios using PROC TRAJ.

In conclusion, the BIC and LMRT are effective for censored normal and gamma data. The correct identification rates for ZIP and Bernoulli data using M-plus are 0. I suggest PROC TRAJ instead of M-plus for ZIP and Bernoulli data. More research will be done in the future.

Reference

1. Bauer D. J. & Curran P. J. (2003). Distributional Assumptions of Growth Mixture Models: Implications for Overextraction of Latent Trajectory Classes. Psychological Methods. Vol.8, No.3,338-363

2. Bauer D. J. & Curran P. J. (2004) .The Integration of Continuous Discrete Latent Variable Models: Potential Problems and Promising Opportunities. Psychological Methods. Vol.9, No.1,3-29

3. Bauer D. J. (2011). Observations on the Use of Growth Mixture Models in Psychological Research. Multivariate Behavioral Research.

4. Bollen K.&Brand J. (2008). Fixed and Random Effects in Panel Data Using Structural Equations Models. California Center for Population Research On-Line Working Paper Series.

5. Box G. E. P. &Cox D. R. (1964). An analysis of Transformations. Journal of the Royal Statistical Society. Series B (Methodological). Vol. 26 No.2, 211-252

6. Celeux G. &Soromenho G. (1996). An Entropy Criterion for Assessing the Number of Clusters in a Mixture Model. Journal of Classification 13:195-212

7. Chang S.,Choi S., Li K., Fleur R., Huang C., Shen T., Ahn K., Gordon D., Kim W., Wu R., Mendell N. & Finch S. (2009). Growth mixture modeling as an exploratory analysis tool in longitudinal quantitative trait loci analysis. BMC PROCeeding 3 Suppl 7:S112

8. Connell A. M. & Frye A. A. (2006). Growth Mixture Modeling in Developmental Psychology: Overview and Demonstration of Heterogeneity in Developmental TRAJectories of Adolescent Antisocial Behavior. Infant and Child Development. Dev.15:609-621.

9. Colder C. R., Mehta P., Balanda K., Campbell R. T., Mayhew K.P., Stanton W.R., Pentz M.A., & Flay B.R. (2001). Identifying TRAJectories of Adolescent Smoking: An Application of Latent Growth Mixture Modeling. Health Psychology. Vol.20, No.2, 127-135.

10. Crick N. & Ostrov J. & Werner N. (2006). A Longitudinal Study of Relational Aggression, Physical Aggression, and Children's Social- Psychological Adjustment. Journal of Abnormal Child Psychology. Vol.34, No.2, 131-142

11. Duncan T., Duncan S.&Hops H. (1996). The Role of Parents and Older Siblings in Predicting Adolescent Substance Use: Modeling Development via Structural Equation Latent Growth Methodology. Journal of Family Psychology. Vol.10, No. 2,158-172

12. Duncan T. & Duncan S. (2004). An Introduction to Latent Growth Curve Modeling. Behavior Therapy 35, 333-363

13. Eklund J. M., Margaret K. & Stattin H. (2010). Romantic relationships and delinquent behavior in adolescence: the moderating role of delinquency propensity. Journal of Adolescence 33, 377-386

14. Fan X. (2001). Parental Involvement and Students' Academic Achievement: A Growth Modeling Analysis. The Journal of Experimental Education. 70(1).27-61

15. Fan X. (2003). Power of Latent Growth Modeling for Detecting Group Differences in Linear Growth Trajectory Parameters. STRUCTURAL EQUATION MODELING,10(3),380-400

16. Gomez S. &O`Malley C. & Stroup A. & Shema S. & Satariano W. (2007). Longitudinal, population-based study of racial/ethnic differences in colorectal cancer survival: impact of neighborhood socioeconomic status, treatment and comorbidity. BMC Cancer.

17. Hertzog C., Oertzen T., Ghisletta P. &Lindenberger U.(2006).On the Power of Multivariate Latent Growth Curve Models to Detect Correlated Change. Psychological Methods,Vol.11,No.3,244-252

18. Hertzog, C., Oertzen T.,Ghisletta,P.&Lindenberger,U.(2008).Evaluating the Power of Latent Growth Curve Models to Detect Individual Differences in Change. Structural Equation Modeling,15:541-563

19. Hipp J. &Bauer D. (2006). Local Solutions in the Estimation of Growth Mixture Models. Psychological Methods, Vol. 11.No.1,36-53

20. Hong S. & Ho H. (2005). Direct and Indirect Longitudinal Effects of Parental Involvement on Student Achievement: Second-Oder Latent Growth Modeling across Ethnic Groups. Journal of Educational Psychology. Vol. 97, No. 1, 32-42

21. Jones B. & Nagin D. &Roeder K. (2001). A SAS PROCedure Based on Mixture Models for Estimating Developmental Trajectories. Sociological Methods& Research, Vol.29 No.3 374-393

22. Jones B. & Nagin D. (2007). Advances in Group-Based Trajectory Modeling and an SAS Procedure for Estimating Them. Sociological Methods& Research Vol.35 No.4

23. Jung T. & Wickrama K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. Social and Personality Psychology Compass. 2/1. 302-317

24. Kreuter F. & Muthén B. (2007) Analyzing criminal trajectory profiles: Bridging multilevel and group-based approaches using growth mixture modeling. Quantitative Criminology

25. Kerner B. & Muthén B. (2009) Growth mixture modeling in families of the Framingham heart study. BMC PROCeedings, 2009, 3(suppl 7):S114

26. Lo Y., Mendell N. & Rubin D. (2001). Testing the number of components in a normal mixture. Biometrika 88(3). 767-778

27. Lubke G. & Muthén B. (2005). Investigating Population Heterogeneity with Factor Mixture Models. Psychological Methods. Vol.10, No. 1,22-39

28. Lubke G. & Muthén B. (2007). Performance of Factor Mixture Models as a Function of Model Size, Covariate Effects, and Class-Specific Parameters. Structural Equation Modeling,14(1),26-47

29. Lubke G. & Neale M. (2006). Distinguishing Between Latent Classes and Continuous Factors: Resolution by Maximum Likelihood? Multivariate Behavioral Research, 41(4), 499-532.

30. Magidson J. &Vermunt J. (2004). Latent class models. In D. Kaplan (Ed.), Handbook of quantitative methodology ofr the social sciences (pp. 175-198). Newbury Park, CA: Sage

31. Mclachlan G. &Peel D. (2000). Finite mixture models. Wiley Series in Probability and Statistics

32. Mike W. & Cheung L. (2008). A Model for Integrating Fixed-, Random-, and Mixed-Effects Meta Analyses Into Structural Equation Modeling. Psychological Methods. Vol. 13.No.3,182-202

33. Morgan B.J.T. & Ridout M. S. (2008). A new mixture model for capture heterogeneity. Appl. Statist. 57, part 4,433-466

34. Muthén B. (1984). A general structural equation model with Bernoulli, ordered categorical and continuous latent variable indicators. Psychometrika. Vol.49, No. 1,115-132

35. Muthén, B. O. (1989). Latent Variable Modeling in Heterogeneous Populations. Psychometrika. VOl.54, No.4,557-585

36. Muthén B. O. & Curran P. (1997).General Longitudinal Modeling of Individual Differences in Experimental Designs: A latent Variable Framework for Analysis and Power Estimation. Psychological Methods,Vol 2.No.4,371-402

37. Muthén B. O. (2001). Latent Variable Mixture Modeling. In G.A. Marcoulides & R.E. Schumacker (Eds.), New developments and techniques in structural equation modeling. Pp.1-33. Mahwah, NJ: Erlbaum.

38. Muthén L. K. &Muthén B. O. (1998-2010). M-plus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

39. Muthén L. K. &Muthén B. O. (2000).Integrating Person-Centered and Variable-Centered Analyses: Growth Mixture Modeling With Latent Trajectory Classes. Alcoholism: Clinical and Experimental Research,Vol,24,No.6

40. Muthén L. K. &Muthén B. O. (2002).How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power. Structural Equation Modeling,9(4),599-620

41. Muthén B. O., Brown C., Masyn K., Booil Jo, Siek-Toon K., Chin-Chien Y., Chenping,W., Kellam G., John B. &Jason L.(2002).General Growth Mixture Modeling for Randomized Preventive Interventions.Biostatistics,3,4,459-475

42. Muthe´n, B. O. (2003). Statistical and Substantive Checking in Growth Mixture Modeling: Comment on Bauer and Curran (2003). Psychological Methods,Vol 8.No.4,369-377

43. Nagin D. S. (1999). Analyzing Developmental Trajectories: A Semiparametric, Group-Based Approach. Psychological Methods. Vol.4, No. 2,139-157

44. Nagin D. & Tremblay R. (1999). Trajectories of Boy's Physical Aggression, Opposition, and Hyperactivity on the Path to Physically Violent and Nonviolent Juvenile Delinquency. Child Development. Vol.70 No.5, 1181-1196

45. Ning Y. & Finch S. J. (2000). The null distribution of the likelihood ratio test for a mixture of two normals after a restricted box-cox transformation. Communications in Statistics-Simulation and Computation, 29:2,449-461

46. Ning Y. & Finch S. J. (2004). The Likelihood Ratio Test with the Box-Cox Transformation for the Normal Mixture Problem: Power and Sample Size Study. Communications in Statistics-Simulation and Computation, 33:3,553-565

47. Nylund K, Asparouhov T. & Muthén B.(2007).Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. STRUCTURAL EQUATION MODELING,14(4),535-569

48. Raftery A. (1995). Bayesian Model Selection in Social Research. Sociological Methodology, Vol.25, 111-163

49. Raudenbush S. &Bryk A. (2002). Hierarchical Linear Models: Applications and Data Analysis Methods. Sage Publications, Inc.

50. Raudenbush S.& Liu X. F. (2001).Effects of Study Duration, Frequency of Observation, and Sample Size on Power in Studies of Group Differences in Polynomial Change. Psychological Methods, Vol.6, No.4, 387-401

51. Rescorla L., Mirak J., & Singh L. (2000). Vocabulary growth in late talkers: lexical development from 2; 0 to 3; 0. Child Lang.27, 293-311

52. Rindskopf D. (2003). Mixture or homogeneous? Comment on Bauer and Curran (2003). Psychological Methods, Vol.8, No.3, 364-368

53. Schwarz G. (1978). Estimating the Dimension of a Model. Annals of Statistics. Vol. 6, No. 2, 461-464

54. Sclove L. (1987). Application of model-selection criteria to some problems in multivariate analysis. Psychometrika, 52,333-343

55. Singer J. D. (1998). Using SAS PROC Mixed to fit multilevel models, hierarchical models and individual growth models. Journal of Educational and Behavioral Statistics. Vol. 24. No.4.323-355

56. Terry E. Duncan, Susan C. Duncan, Lisa A. Strycker. (2006).An introduction to latent variable growth curve modeling: concepts, issues and Applications

57. Tofighi D. & Enders C. K. (2007). Identifying the correct number of classes in a growth mixture. Advances in Latent Variable Mixture Models, pages 317-341

58. Tony J. & Wickrama K.A.S. (2008).An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling. Social and Personality Psychology,302-317

59. Whites H., Johnson V. & Buyske S. (2000). Parental modeling and parenting behavior effects on offspring alcohol and cigarette use a growth curve analysis. Journal of Substance Abuse 12. 287-310

60. Yang C. (2006). Evaluating latent class analyses in qualitative phenotype identification. Computational Statistics & Data Analysis, 50, 1090-1104

61. Yang C. C., Muthe´n, B. O. & Yang C. C. (1999). Finite mixture multivariate generalized linear models using Gibbs sampling and E-M algorithms. PROC. Natl. Sci. Counc. ROC (A). Vol. 23, No. 6,695-702

62. Zhiyong Z. & Lijuan W. (2009).Statistical Power Analysis for Growth Curve Models Using SAS. Behavior Research Methods,41(4),1083-1094

Appendix

Table 23 Odds Ratio Estimates for factors of normal data computational failure rates

Effect		Point estimate	95% Wald Confidence Limits	
Intra-class correlations	0.265 vs0.765	1.49	1.24	1.80
Trajectory Pattern	Quadratic vs Linear	92.80	80.78	169.56
Time measurements	4 vs 6	1.41	1.17	1.69
Sample size	400 vs 800	1.20	1.00	1.44

Table 24 Odds Ratio Estimates for factors of gamma data computational failure rates

Effect		Point estimate	95% Wald Confidence Limits	
Intra-class correlations	0.265 vs0.765	1.31	1.12	1.54
Trajectory Pattern	Quadratic vs Linear	8.21	6.75	9.98
Time measurements	4 vs 6	1.36	1.16	1.59
Sample size	400 vs 800	1.23	1.05	1.45

Table 25 Analysis of factors determining correct identification rate of a single trajectory of normal data based on BIC with minimum BIC rule

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept		1	26.53	210.40	0.02	0.90
Intra-class correlations	0.265	1	0.70	1.23	0.32	0.57
Trajectory Pattern	Linear	1	-11.11	148.80	0.01	0.94
Time measurements	4	1	-11.11	148.80	0.01	0.94
Sample size	400	1	0.70	1.23	323.00	0.57

Table 26 Odds Ratio Estimates for correct identification rate of normal data using LMRT with significance level 0.05 and 0.01

Effect		Point estimate	95% Wald Confidence Limits	
significance level 0.05				
Intra-class correlations	0.265 vs0.765	1.52	1.11	2.09
Trajectory Pattern	Quadratic vs Linear	1.24	0.91	1.70
Time measurements	6 vs 4	1.21	0.89	1.65
Sample size	800 vs 400	1.15	0.84	1.57
significance level 0.01				
Intra-class correlations	0.265 vs0.765	1.12	0.70	1.81
Trajectory Pattern	Linear vs Quadrati	1.12	0.70	1.81
Time measurements	4 vs 6	1.12	0.70	1.81
Sample size	800 vs 400	1.42	1.12	1.71

Table 27 Odds Ratio Estimates for correct identification rate of normal data using BLRT with significance level 0.05 and 0.01

Effect		Point estimate	95% Wald Confidence Limits	
Significance level=0.05				
Intra-class correlations	0.265 vs0.765	1.99	1.36	2.92
Trajectory Pattern	***Quadratic vs Linear***	***3.66***	***2.39***	***5.60***
Time measurements	4 vs 6	1.43	0.98	2.08
Sample size	800 vs 400	1.92	0.30	2.66
Significance level=0.01				
Intra-class correlations	0.265 vs0.765	1.65	1.09	2.50
Trajectory Pattern	***Quadratic vs Linear***	***3.23***	***2.04***	***5.11***
Time measurements	4 vs 6	1.65	1.09	2.50
Sample size	800 vs 400	2.17	1.41	3.33

Table 28 Analysis of factors determining correct identification rate of a single trajectory of gamma data based on BIC with absolute BIC rates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept		1	4.93	1.02	23.64	<0.0001
Intra-class correlations	0.265	1	2.71	1.04	6.78	0.009
Trajectory Pattern	Linear	1	1.94	0.77	6.39	0.012
Time measurement	4	1	-2.71	1.04	6.78	0.009
Sample size	400	1	1.94	0.77	6.39	0.012

Table 29 Odds Ratio Estimates for correct identification rate of gamma data using LMRT with significance level 0.05 and 0.01

Effect		Point estimate	95% Wald Confidence Limits	
Significance level=0.05				
Intra-class correlations	0.265 vs0.765	1.36	0.99	1.89
Trajectory Pattern	Quadratic vs Linear	1.01	0.74	1.40
Time measurements	6 vs 4	1.56	1.13	2.17
Sample size	400 vs 800	1.01	0.74	1.40
Significance level=0.01				
Intra-class correlations	0.265 vs0.765	1.77	1.18	2.66
Trajectory Pattern	Linear vs Quadratic	1.63	1.09	2.44
Time measurements	6 vs4	1.70	1.14	2.55
Sample size	800 vs 400	1.27	0.86	1.89

Table 30 Odds Ratio Estimates for correct identification rate of gamma data using BLRT with significance level 0.05 and 0.01

Effect		Point estimate	95% Wald Confidence Limits	
Significance level=0.05				
Intra-class correlations	0.265 vs0.765	1.65	1.23	2.20
Trajectory Pattern	Linear vs Quadratic	1.61	1.21	2.15
Time measurements	6 vs 4	1.42	1.06	1.89
Sample size	400 vs 800	1.12	0.85	1.49
Significance level=0.01				
Intra-class correlations	0.265 vs0.765	2.03	1.49	2.78
Trajectory Pattern	Linear vs Quadratic	1.89	1.39	2.58
Time measurements	4 vs 6	1.14	0.84	1.54
Sample size	400 vs 800	1.20	0.88	1.62

Table 31 Average selected number of classes of BIC, LMRT and BLRT for ZIP model

scenarios				Fit to	BIC	LMRT	BLRT
Fraction	Trend	Effects	Sample size				
1/8	Linear trend	Fixed	400	linear	2	3	3
1/8	Linear trend	Fixed	400	Quadratic	2	3	3
1/8	Linear trend	Random	400	linear	2	2	4
1/8	Linear trend	Random	400	Quadratic	2	2	3
1/8	s-shaped trend	Fixed	400	linear	2	3	3
1/8	s-shaped trend	Fixed	400	Quadratic	2	3	3
1/8	s-shaped trend	Random	400	linear	2	3	3
1/8	s-shaped trend	Random	400	Quadratic	2	3	3
1/2	Linear trend	Fixed	400	linear	2	3	4
1/2	Linear trend	Fixed	400	Quadratic	2	3	3
1/2	Linear trend	Random	400	linear	2	2	3
1/2	Linear trend	Random	400	Quadratic	2	2	4
1/2	s-shaped trend	Fixed	400	linear	2	3	3
1/2	s-shaped trend	Fixed	400	Quadratic	2	3	3
1/2	s-shaped trend	Random	400	linear	2	3	3
1/2	s-shaped trend	Random	400	Quadratic	2	3	3
ρ_{it}	Linear trend	Fixed	400	linear	2	3	3
ρ_{it}	Linear trend	Fixed	400	Quadratic	2	3	3
ρ_{it}	Linear trend	Random	400	linear	2	2	3
ρ_{it}	Linear trend	Random	400	Quadratic	2	2	4
ρ_{it}	s-shaped trend	Fixed	400	linear	2	3	3

ρ_{it}	s-shaped trend	Fixed	400	Quadratic	2	3	2
ρ_{it}	s-shaped trend	Random	400	linear	2	3	3
ρ_{it}	s-shaped trend	Random	400	Quadratic	2	2	3
1/8	Linear trend	Fixed	800	linear	3	3	4
1/8	Linear trend	Fixed	800	Quadratic	3	4	4
1/8	Linear trend	Random	800	linear	3	3	4
1/8	Linear trend	Random	800	Quadratic	3	3	4
1/8	s-shaped trend	Fixed	800	linear	3	4	4
1/8	s-shaped trend	Fixed	800	Quadratic	3	3	4
1/8	s-shaped trend	Random	800	linear	3	4	4
1/8	s-shaped trend	Random	800	Quadratic	3	3	4
1/2	Linear trend	Fixed	800	linear	3	3	3
1/2	Linear trend	Fixed	800	Quadratic	3	3	4
1/2	Linear trend	Random	800	linear	3	3	4
1/2	Linear trend	Random	800	Quadratic	3	3	3
1/2	s-shaped trend	Fixed	800	linear	3	3	4
1/2	s-shaped trend	Fixed	800	Quadratic	3	3	4
1/2	s-shaped trend	Random	800	linear	3	3	3
1/2	s-shaped trend	Random	800	Quadratic	3	3	4
ρ_{it}	Linear trend	Fixed	800	linear	3	4	4
ρ_{it}	Linear trend	Fixed	800	Quadratic	3	3	4
ρ_{it}	Linear trend	Random	800	linear	3	3	4
ρ_{it}	Linear trend	Random	800	Quadratic	3	3	4
ρ_{it}	s-shaped trend	Fixed	800	linear	3	4	4

ρ_{it}	s-shaped trend	Fixed	800	Quadratic	3	3	4
ρ_{it}	s-shaped trend	Random	800	linear	3	3	4
ρ_{it}	s-shaped trend	Random	800	Quadratic	3	3	4

Table32 Odds Ratio Estimates for correct identification rate of ZIP data using BIC

Effect		Point estimate	95% Wald Confidence Limits	
Proportion of extra zeros	½ vs rho	95.54	40.62	224.72
Proportion of extra zeros	1/8 vs rho	1.03	0.74	1.44
Trend	linear	190.96	83.03	439.22
Effects	Fixed	>999.999	161.79	>999.999
Sample size	400	16.18	11.34	23.07
Trajectory pattern	Quadratic vs linear	1.53	1.10	2.14

Table 33 Average selected number of classes of BIC, LMRT and BLRT for bernoulli model

Scenario			Fit to	BIC	LMRT	BLRT
Linear trend	Fixed effect	400	linear	2	3	3
Linear trend	Fixed effect	400	Quadratic	2	3	3
Linear trend	Random effect	400	linear	2	3	3
Linear trend	Random effect	400	Quadratic	2	3	3
S-shaped trend	Fixed effect	400	linear	2	3	3
S-shaped trend	Fixed effect	400	Quadratic	2	3	3
S-shaped trend	Random effect	400	linear	2	3	3
S-shaped trend	Random effect	400	Quadratic	2	3	2
Linear trend	Fixed effect	800	linear	3	4	4
Linear trend	Fixed effect	800	Quadratic	3	3	4
Linear trend	Random effect	800	linear	4	4	4
Linear trend	Random effect	800	Quadratic	4	3	4
S-shaped trend	Fixed effect	800	linear	3	4	4
S-shaped trend	Fixed effect	800	Quadratic	3	3	3
S-shaped trend	Random effect	800	linear	3	3	4
S-shaped trend	Random effect	800	Quadratic	3	3	3

Table 34 Analysis of factors determining correct identification rate of a single trajectory of gamma data based on BIC with significant BIC rate

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept		1	26.24	217.80	0.02	0.90
Trend	linear	1	-29.19	217.80	0.02	0.89
Effects	Fixed	1	-29.19	217.80	0.02	0.89
Sample size	400	1	-29.19	217.80	0.02	0.89
Trajectory pattern	Quadratic	1	-11.23	119.60	0.01	0.93