Sample Size and Measurement Occasion Planning for Latent Change Score Models through Monte Carlo Simulation

Zhiyong Zhang and Haiyan Liu
Department of Psychology, University of Notre Dame
Abstract

Latent change score models (LCSMs) proposed by McArdle (McArdle, 2000, 2009; McArdle & Nesselroade, 1994) offer a powerful tool for longitudinal data analysis. They are becoming increasingly popular in the social and behavioral research (e.g., Gerstorf et al., 2007; Ghisletta & Lindenberger, 2005; King et al., 2006; Raz et al., 2008). Although conducting both univariate and multivariate latent change score analysis is not a difficult task any more (e.g., Ghisletta & McArdle, 2012; Zhang et al., 2015), there is little discussion on the design issues such as sample size planning for LCSMs. To fill the gap, this study proposes a Monte Carlo based method to determine the required sample size and number of measurement occasions for both univariate and bivariate LCSMs. The method can obtain the power for testing each individual parameter of the models, especially the change rate and coupling parameters. The Monte Carlo procedure is implemented and provided in a free R package RAMpath (Zhang et al., 2015). Examples for sample size and measurement occasion planning for both univariate and bivariate LCSMs are provided.
Sample Size and Measurement Occasion Planning for Latent Change Score Models through Monte Carlo Simulation

Introduction

Longitudinal data collection and data analysis are becoming a norm for psychological research (e.g., Grimm et al., 2016; McArdle & Nesselroade, 2014). A longitudinal design often involves data collection on multiple variables from multiple participants at multiple times. Despite of the increased cost and complexity, there are many advantages to collect longitudinal data. For example, a longitudinal design naturally enables a researcher to study change and related phenomena. In addition, inter-individual differences in change can also be investigated.

Growth curve models are probably the most widely used technique for analyzing longitudinal data benefiting from the fact that a growth curve model can be fitted from a structural equation modeling (SEM) framework (e.g., McArdle, 1986; McArdle & Epstein, 1987; McArdle & Anderson, 1990; McArdle & Hamagami, 1992; McArdle, 1998; McArdle & Bell, 1998; McArdle & Nesselroade, 2014). With the increasing use of longitudinal design, it is not surprising that more and more complex models and methods have been developed. For example, in order to deal with missing data, full information maximum likelihood methods, multiple imputation, and Bayesian methods have been developed and used (e.g., Enders, 2011; Lu et al., 2013). To deal with non-normal data, robust methods have been proposed (e.g., Yuan & Zhang, 2012; Zhang, 2013; Zhang et al., 2013).

A more difficult issue in longitudinal research is to model the nonlinear trajectory of data. With more data collection, a linear growth curve model is often not sufficient. When moving to nonlinear models, issues such as computational difficulty can arise (e.g., Grimm et al., 2011; Wang & McArdle, 2008). Linearizing a nonlinear model provides an efficient way to deal with such difficulty. Although the method based on Taylor expansion is well known (e.g., Browne, 1993; Neale & McArdle, 2000), it is less known that the latent change score models (LCSMs) provide another way to model nonlinear trajectory that can be more efficient.

Proposed by McArdle and colleagues, LCSMs combine difference equations with growth
curves to study change in longitudinal studies (e.g., McArdle, 2000; McArdle & Hamagami, 2001; Hamagami & McArdle, 2007a; Hamagami et al., 2010). In such models, change is directly modeled, which is often the focus of a longitudinal study. As we will show shortly, the models in their nature accommodate to the nonlinear growth trajectories. In addition to the univariate LCSMs, bivariate LCSMs have also been proposed to model the inter-relationship between two growth processes (e.g., McArdle & Hamagami, 2001).

Fitting a LCSM in the SEM framework is easy to understand but can be tedious. It can be done in almost any SEM software. Recently, Ghisletta & McArdle (2012) showed how to estimate a univariate LCSM using difference R packages, including Lavaan (Rosseel, 2012), OpenMx (Boker et al., 2011) and sem (Fox, 2006). More recently, Zhang et al. (2015) automated the estimating procedure for the typical univariate and multivariate LCSMs through an R package RAMpath that is developed based on RAM notations (Boker et al., 2002; McArdle & Boker, 1990).

The importance of conducting statistical power analysis at the beginning of a study is universally accepted (e.g., Cohen, 1988; Hedges & Rhoads, 2010). Without adequate statistical power, the validity of statistical conclusions from all kinds of research is endangered (e.g., Cohen, 1988; Hedges & Rhoads, 2010; Myors & Wolach, 2014; Shadish et al., 2002). For example, without a carefully planned sample size, a study can easily fail to detect an existing effect by chance and, therefore, create problems for replication or cross-validation. Although there are studies on sample size planning and power calculation for growth curve analysis (e.g., Zhang & Wang, 2009), we are not aware of any discussion on such design issues for LCSMs.

To fill the gap, this study proposes a Monte Carlo based method to determine the required sample size and/or the number of measurement occasions for both univariate and bivariate LCSMs. The method can obtain the power for testing each individual parameter of the models such as the change rate and coupling parameters. We also implement the Monte Carlo procedure in a free R package RAMpath (Zhang et al., 2015).

In the rest of the chapter, we first present the univariate and bivariate LCSMs. Then, we
introduce the Monte Carlo based method for power analysis. After that, we show how to conduct power analysis for LCSMs through several examples using our developed software. We conclude the chapter with discussion and future directions.

A Univariate Latent Change Score Model

Let \( Y[t]_n \) denote the data from the \( n \)th \((n = 1, \ldots, N)\) participant at time \( t \) \((t = 1, \ldots, T)\) of a sample consisting of \( N \) participants measured for \( T \) times. The first part of a LCSM is a measurement error model where an observed score \( Y[t]_n \) is the sum of the latent true score \( y[t]_n \) and the measurement error/unicness score \( e_y[t]_n \):

\[
Y[t]_n = y[t]_n + e_y[t]_n.
\]

It is generally assumed that the error follows a normal distribution with mean 0 and variance \( \text{var} e_y \). The second part of the model builds the relationship between consecutive latent true scores so that the current score at time \( t \) is equal to the sum of the true score at the previous time \( t - 1 \) and the change, \( dy[t]_n \), from time \( t - 1 \) to time \( t \):

\[
y[t]_n = y[t - 1]_n + dy[t]_n.
\]

This effectively defines the change score as

\[
dy[t]_n = y[t]_n - y[t - 1]_n.
\]

Note that in the classic LCSM, the relationship between consecutive latent true scores is deterministic although it is not required to be so. The third part of the model concerns the modeling of the difference scores. One way is to model the difference score at time \( t \) as the sum of a linear constant effect \( y s \) and the proportional change from time \( t - 1 \) such that

\[
dy[t]_n = y s_n + \beta_y \times y[t - 1]_n,
\]

where \( \beta_y \) is a compound rate of change.
Given the three part of the model, we can model the observed score as

\[ Y[t]_n = y[t]_n + ey[t]_n \]

\[ = y[t - 1]_n + dy[t]_n + ey[t]_n \]

\[ = (1 + \beta_y)y[t - 1]_n + ys_n + ey[t]_n. \]

Successively expressing the above equation will lead to

\[ Y[t]_n = (1 + \beta_y)y[t - 1]_n + ys_n + ey[t]_n \]

\[ = (1 + \beta_y)(y[t - 2]_n + dy[t - 1]_n) + ys_n + ey[t]_n \]

\[ = (1 + \beta_y)^2y[t - 2]_n + (1 + \beta_y)ys_n + ys_n + ey[t]_n \]

\[ = (1 + \beta_y)^{t-1}y[1]_n + [1 + (1 + \beta_y) + \ldots + (1 + \beta_y)^{t-2}]ys_n + ey[t]_n \]

\[ = (1 + \beta_y)^{t-1}y0_n + [1 + (1 + \beta_y) + \ldots + (1 + \beta_y)^{t-2}]ys_n + ey[t]_n \]

where \(y0[n]\) is the initial latent score and note that the latent score at time \(t\) follows

\[ y[t]_n = (1 + \beta_y)^{t-1}y0_n + [1 + (1 + \beta_y) + \ldots + (1 + \beta_y)^{t-2}]ys_n. \]

Clearly, the observed and latent scores behave as a nonlinear function of time and therefore can capture the nonlinear trajectory except when \(\beta_y = 0\). To visually show this, we plot the latent scores with different values for \(\beta_y\) in Figure 1.

The initial latent score and the linear constant change can be correlated. In the model, they are assumed to have a bivariate normal distribution

\[
\begin{pmatrix}
  y0_n \\
  ys_n
\end{pmatrix}
\sim
MN
\begin{bmatrix}
  my0 & vary0ys \\
  mys & varys
\end{bmatrix}
\]

with \(MN\) denoting a multivariate, here bivariate, normal distribution. Therefore, the initial latent score follows a normal distribution with mean \(my0\) and variance \(vary0\) and the constant change also follows a normal distribution with mean \(mys\) and variance \(varys\). The covariance between them is \(vary0ys\) with the correlation expressed as

\[ \rho_{y0ys} = \frac{vary0ys}{\sqrt{vary0} \times varys}. \]
Using a path diagram, this model is portrayed in Figure 2. In the path diagram, squares represent observed variables, while circles represent latent variables. A single-headed arrow is for deterministic parameters such as regression coefficients, or factor loadings, while a double-headed arrow represents stochastic parameters such as variance and covariance. A triangle represents a constant. Any arrow originating from the triangle represents an intercept or mean of variables pointed by the arrow. We matched the notation in the formulas and in the path diagram. For simplification, we removed the brackets and the subscripts for the variables in the path diagram.

**A Bivariate Latent Change Score Model**

A bivariate LCSM is first a combination of two univariate LCSMs. Above and beyond that, it allows the two processes represented by the LCSMs to interact with each other. Let $Y[t]_n$ and $X[t]_n$ denote the observed data on two variables, respectively, from the $n$th ($n = 1, \ldots, N$) participant at time $t$ ($t = 1, \ldots, T$) of a sample consisting of $N$ participants measured for $T$ times. For the measurement error part of the model, we have

$$Y[t]_n = y[t]_n + ey[t]_n$$
$$X[t]_n = x[t]_n + ex[t]_n,$$

where $ey[t]_n$ follows a normal distribution with mean 0 and variance $varey$ and $ex[t]_n$ follows a normal distribution with mean 0 and variance $varex$. For the latent score from time $t - 1$ to time $t$, we have

$$y[t]_n = y[t - 1]_n + dy[t]_n$$
$$x[t]_n = x[t - 1]_n + dx[t]_n,$$

with $dy[t]_n$ and $dx[t]_n$ denoting the latent change score for the two variables respectively.

The innovative part of the bivariate LCSM is to allow the latent score of one variable to influence the change score of another variable. Specifically, we model the change scores as

$$dy[t]_n = ys_n + \beta_y \times y[t - 1]_n + \gamma_y x[t - 1]_n$$
$$dx[t]_n = xs_n + \beta_x \times x[t - 1]_n + \gamma_x y[t - 1]_n$$
where $\gamma_y$ and $\gamma_x$ are called coupling parameters. $\gamma_y$ represents the effect of $x$ on the change score of $y$ and $\gamma_x$ represents the effect of $y$ on the change score of $x$. We let $x_0$ be the initial latent score and $x_s$ be the constant change for $x$. A multivariate normal distribution is assumed for the initial latent scores and constant changes for the two variables such that

$$ \begin{pmatrix} y_{0n} \\ y_{sn} \\ x_{0n} \\ x_{sn} \end{pmatrix} \sim MN \begin{pmatrix} m_{y0} \\ m_{ys} \\ m_{x0} \\ m_{xs} \end{pmatrix}, \begin{pmatrix} \text{vary0} & \text{vary0ys} & \text{varx0y0} & \text{vary0xs} \\ \text{vary0ys} & \text{varys} & \text{varx0ys} & \text{varxs} \\ \text{varx0y0} & \text{varx0ys} & \text{varx0} & \text{varx0xs} \\ \text{vary0xs} & \text{varxs} & \text{varx0xs} & \text{varxs} \end{pmatrix} $$

Using a path diagram, a bivariate LCSM is portrayed in Figure 3.

**Statistical Power Analysis Based on Monte Carlo Simulation**

Statistical power analysis can be viewed as concerning a test whether one or a subset of parameters, denoted by $\gamma$, in a model with all the parameters $\theta$, is equal to 0 or known value $\gamma_0$. Therefore, the null and alternative hypotheses of interest are

$$ H_0 : \gamma = \gamma_0 \text{ vs. } H_1 : \gamma \neq \gamma_0. $$

Existing procedures for power evaluation are mostly based on the Wald test or the likelihood ratio test. The Wald test statistic is defined as

$$ T = (\hat{\gamma} - \gamma_0)' \hat{\Phi}(\hat{\gamma} - \gamma_0) $$

where $\hat{\gamma}$ is the parameter estimates in $\hat{\theta}$ corresponding to $\gamma$ and $\hat{\Phi}$ is the covariance matrix of $\hat{\gamma}$. The Wald test statistic can be compared to a critical value $C_\alpha$ under the null hypothesis. If $T > C_\alpha$, the null hypothesis is rejected. Under the null hypothesis and typical normality assumption, the Wald statistic asymptotically follows a chi-square distribution ($\chi^2_q$) with the degrees of freedom $q$, where $q$ is the length of $\gamma$. Then, the critical value at the significance level $\alpha$ is $C_\alpha = \chi^2_q(1 - \alpha)$. Note that when working with a single parameter, the Wald test is the square of a z test.
The likelihood ratio test works in the similar manner. In the likelihood ratio test, one first estimates the model under the alternative hypothesis to get the value of the likelihood function at $L_1$. Then, one can estimate the model under the null hypothesis by fixing the parameters in $\gamma$ to be $\gamma_0$ to get the value of the likelihood function at $L_0$. The likelihood ratio test statistic is defined as

$$T = -2 \ln \frac{L_0}{L_1}.$$  \hspace{1cm} (2)

The likelihood ratio test statistic is also compared to a critical value $C_\alpha$ to decide whether a null hypothesis should be rejected. If $T$ literally follows a chi-square distribution with degrees of freedom $q$, the critical value is $\chi^2_q(1 - \alpha)$. If $T > C_\alpha$, the null hypothesis is rejected.

By its definition, the statistical power is defined as

$$\pi = \Pr(\text{reject } H_0|H_1 \text{ is true})$$

$$= \Pr(T > C_\alpha|H_1 \text{ is true}),$$

where $T$ can be the Wald statistic or the likelihood ratio test statistic. For simple statistical analysis such as a t-test, one can obtain an analytical form for $\pi$ and therefore power and sample size planning can be conducted easily. However, for LCSMs, both the Wald and the likelihood ratio test statistics are complex functions of sample size and effect size. Therefore, the statistical power $\pi$ is also a complex function of these factors. Generally speaking, it is difficult or impossible to get a tractable form of $\pi$ so that the relationship between statistical power and sample size can be easily evaluated.

To deal with the difficulty in power analysis for LCSMs, we use a Monte Carlo simulation based method to approximate the power using the relative frequency to reject the null hypothesis given the alternative hypothesis is true. Specifically, the following procedure can be used.

1. Decide the significance level. Usually, the default 0.05 can be used. Based on that, get the critical value $C_\alpha$. If only one parameter is tested, the $C_\alpha$ based on the normal distribution is 1.96 for a z-test and 3.84 for the chi-squared distribution for the Wald test.

2. Specify a LCSM $M_1$ under $H_1$ with the hypothesized population parameter values ($\theta$).
3. Generate a set of data with the sample size \( N \) and number of measurement occasions \( T \) from the model using random number generation techniques.

4. Fit Models \( M_1 \) and \( M_0 \), the model by setting \( \gamma = \gamma_0 \), to the generated data and obtain the Wald statistic using Equation (1) and/or the likelihood ratio statistic using Equation (2).

5. If \( T > C_\alpha \), the null hypothesis \( H_0 \) is rejected.

6. Repeat Steps (2)–(5) for a total of \( R (R \geq 1000) \) times.

7. Suppose out of the \( R \) replications, the null hypothesis \( H_0 \) is rejected \( r \) times. Then the statistical power with the sample size \( n \) is estimated by \( \hat{\pi} = \frac{r}{R} \).

8. For sample size planning, if \( \hat{\pi} \) is smaller than the desired power, say 0.8, one can increase the sample size or the number of measurement occasions to repeat Steps 2 and 7 to recalculate the power. Otherwise, the sample size or the number of measurement occasions can be set to a smaller value.

The above Monte Carlo simulation based method for statistical power analysis has been widely used in the literature for mediation analysis and SEM (e.g., Muthén & Muthén, 2002; Thoemmes et al., 2010; Zhang & Wang, 2009; Zhang, 2014). This procedure is especially effective for advanced models. For example, Muthén & Muthén (2002) illustrated how to use Mplus to conduct statistical power analysis for structural equation models using such a procedure. Zhang & Wang (2009) focused on how to conduct statistical power analysis for growth curve models with and without missing data. Thoemmes et al. (2010) discussed how to apply the procedure in mediation analysis. Zhang (2014) extended Theommes et al. for the analysis of missing data and non-normal data.

For a typical power analysis for LCSM, a single parameter is often of interest. In this case, the power can be calculated using the above procedure based on the Z test. Furthermore, the power for each individual parameter in a LCSM model can be obtained all at once as we will show in our examples.
Software for Power Analysis for Latent Change Score Models

Although the idea of Monte Carlo simulation based power analysis is straightforward, it would still need the software to implement it to make it useful. Recent, Zhang et al. (2015) developed the R package RAMpath that can estimate both univariate and bivariate LCSMs. We expand RAMpath so that it can carry out power analysis for LCSMs. To further simplify power analysis for researchers who might not be familiar with R, we also develop an online software based on RAMpath.

R package

The R package RAMpath is now on CRAN and therefore it can be installed directly within R as a typical package. For example, to install it, use the R code

```
install.packages("RAMpath")
```

To use the package within R, use

```
library("RAMpath")
```

There are three functions in the package for power analysis:

- `powerLCM`
- `powerBLCM`
- `plot`

The function `powerLCM` is used to conduct power analysis for univariate LCSMs. The basic usage of the function is given below:

```
powerLCS(N = 100, T = 5, R = 1000, betay = 0, my0 = 0, mys = 0, varey = 1, vary0 = 1, varys = 1, vary0ys = 0, alpha = 0.05, ...)
```

In the function, N is the sample size and T is the number of measurement occasions. Both of them can be a single value or a vector. For example, using `N=c(100,200,500)` will calculate power for the three provided sample sizes. R is the number of Monte Carlo simulation used to estimate the power. A larger R will provide more accurate power estimation but also take longer computing time. As a rule of thumb, at least 1,000 should be used. alpha is the significance level for testing the hypothesis of the model parameters. The default value is 0.05.

To obtain power, the population parameter values have to be provided. Such values can be decided based on literature review, pilot study, expert opinions, etc. By default, all the mean,
intercept and covariance parameters are set at 0 and all the variance parameters are set at 1. Those values typically have to be changed in real power analysis. Note that the name of each parameter corresponds to that used in the path diagram in Figure 2. In addition to the basic input, for advanced users, other information can be provided to control the parameter and standard error estimation methods.

The output of the R function includes 4 main pieces of information for each parameter in the model. The first is the Monte Carlo estimate (\texttt{mc.est}). It is calculated as the mean of the \( R \) sets of parameter estimates from the simulated data. Note that the Monte Carlo estimates should be close to the population parameter values used in the model. The second is the Monte Carlo standard deviation (\texttt{mc.sd}), which is calculated as the standard deviation of the \( R \) sets of parameter estimates. The third is the Monte Carlo standard error (\texttt{mc.se}), which is obtained as the average of the \( R \) sets of standard error estimates of the parameter estimates. Lastly, \texttt{mc.power} is the statistical power for each parameter.

The function \texttt{powerBLCM} is used to conduct power analysis for bivariate LCSMs. The basic usage of the function is given below. It is the same as for the univariate LCSMs.

\begin{verbatim}
powerBLCM(N=100, T=5, R=1000, betay=0, my0=0, mys=0, varey=1,
          vary0=1, varys=1, vary0ys=0, betax=0, mx0=0, mxs=0, varex=1,
          varx0=1, varxs=1, varx0xs=0, varx0y0=0, varx0ys=0, vary0xs=0,
          varxsys=0, gammax=0, gammay=0, alpha=0.05, ...)
\end{verbatim}

The function \texttt{plot} is used to generate a power curve, which has the form \texttt{plot(x, parameter, ...)}. The first input of the function, \texttt{x}, is the output from either \texttt{powerLCM} or \texttt{powerBLCM}. In the input of the function for power analysis, either the sample size \( N \) or the number of occasions \( T \) should be a vector. The second input is the name of parameters to plot its power curve. Since there are multiple parameters in a LCSM, one can generate a plot for each model parameter. The name of a parameter should match the one in \texttt{powerLCM} or \texttt{powerBLCM}. This function will generate one or multiple line plots in which power is shown on the y-axis and sample size or the number of occasions is shown on the x-axis.
Online interface

In order to help researchers who are not familiar with R, we also provide a Web-based interface for power analysis for LCSMs. The URL for the univariate LCSMs is http://psychstat.org/lcsm and for the bivariate LCSMs is http://psychstat.org/blcsm.

The Web interface for the univariate LCSMs is shown in Figure 4. Since the interface is built on the R function shown earlier, it requires the same input information and gives the same output. For both sample size and number of occasions, multiple values can be provided in two ways to calculate power for each given value. We discuss this using the sample size as an example since the same method is used for the number of occasions. First, multiple sample sizes can be provided and separated by spaces. For example, inputting 100 150 200 will calculate power for the three sample sizes 100, 150 and 200. Second, a sequence of sample sizes can be generated using the method \( s:e:i \) with \( s \) denoting the starting sample size, \( e \) as the ending sample size, and \( i \) as the interval. Note that the values are separated by a colon “:”. For example, 100:150:10 will generate a sequence of sample sizes: 100 110 120 130 140 150.

The interface for the bivariate LCSMs is similar and is not provided here for the sake of space.

Examples

In this section, we show how to carry out our power analysis for both univariate and bivariate LCSMs through several examples.

Example 1. Type I error rate investigation for a univariate LCSM

Note that if the null hypothesis is true, the Monte Carlo procedure will yield the type I error rate. For example, suppose the parameter \( \beta_y = 0 \) in the population. Then the estimated power for it should be the same as the significance level, typically 0.05. For illustration, we set the population parameter values to those shown in the second column of Table 1. Therefore, if we
conduct a power analysis based on those parameter values, we will obtain the type I error rates for betay, my0, mys and vary0ys. If our Monte Carlo procedure performs well, we expect the type I error rates to be close to the alpha level used.

The R code for conducting the analysis is shown in Code 1. Note that the significance level is set at 0.05 and therefore, we expect the estimated values in the power column are close to 0.05.

Code 1: R input script for Example 1.

```r
powerLCS(N = 100, T = 5, R = 1000, betay = 0, my0 = 0, mys = 0, varey = 1, vary0 = 1, varys = 1, vary0ys = 0, alpha = 0.05)
```

The output of the R code is given in Code 2. First, the estimate for each parameter is very close to the true population parameter values as can be seen in the column named mc.est. This indicates the power calculation procedure runs well. Second, the Monte Carlo standard errors are close to the corresponding Monte Carlo standard deviations, another indicator that the power calculation is trustworthy. Third, as expected, the power for betay, my0, mys, and vary0ys is close to 0.05, the nominal type I error rate. Overall, this suggests that the Monte Carlo based method can provide well-controlled type I error rate.

Code 2: Type I error rate and power for parameters in Example 1.

```
pop.par  mc.est  mc.sd  mc.se  mc.power  N  T
betay    0  0.001  0.056  0.056  0.046  100  5
my0      0  0.001  0.129  0.126  0.056  100  5
mys      0  0.002  0.105  0.105  0.044  100  5
varey    1  0.994  0.083  0.081  1.000  100  5
vary0    1  0.990  0.236  0.230  1.000  100  5
vary0ys  0 -0.005  0.136  0.136  0.044  100  5
varys    1  1.006  0.227  0.227  1.000  100  5
```
Example 2. Power analysis for a univariate LCSM

To conduct a power analysis, one has to specify the population parameter values for the model. Zhang et al. (2015) included an example on using a univariate LCSM model to analyze the WISC data. In order to plan a future study with the sample size 100 and the number of measurement occasion 5, we use the estimates as our population parameter values. Column 3 in Table 1. shows the roundup parameter estimates being used in our example.

The R code for conducting the analysis is shown in Code 3 and the output of the R code is given in Code 4. From the output, we can see that the power to detect the parameter betay to be significant with a sample size 100 and a number of measurement occasions 5 is about 0.664. The power for another parameter, the constant change mys, is 0.274. Since oftentimes one hopes to get a power at least 0.8, a larger sample size is needed for this study. In addition, for studying power for different parameters, different sample sizes are often required.

Code 3: R input script for Example 2.

```r
powerLCS(N = 100, T = 5, R = 1000, betay = 0.1, my0 = 20, mys = 1.5, var = 9, vary0 = 2.5, varys = .05, vary0ys = 0, alpha = 0.05)
```

Code 4: Power for parameters in Example 2.

<table>
<thead>
<tr>
<th>pop.par</th>
<th>mc.est</th>
<th>mc.sd</th>
<th>mc.se</th>
<th>mc.power</th>
<th>N</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>betay</td>
<td>0.10</td>
<td>0.103</td>
<td>0.043</td>
<td>0.044</td>
<td>0.664</td>
<td>100</td>
</tr>
<tr>
<td>my0</td>
<td>20.00</td>
<td>19.999</td>
<td>0.324</td>
<td>0.319</td>
<td>1.000</td>
<td>100</td>
</tr>
<tr>
<td>mys</td>
<td>1.50</td>
<td>1.418</td>
<td>1.106</td>
<td>1.120</td>
<td>0.274</td>
<td>100</td>
</tr>
<tr>
<td>var</td>
<td>9.00</td>
<td>8.961</td>
<td>0.724</td>
<td>0.732</td>
<td>1.000</td>
<td>100</td>
</tr>
<tr>
<td>vary0</td>
<td>2.50</td>
<td>2.463</td>
<td>1.151</td>
<td>1.139</td>
<td>0.583</td>
<td>100</td>
</tr>
<tr>
<td>vary0ys</td>
<td>0.00</td>
<td>-0.004</td>
<td>0.408</td>
<td>0.403</td>
<td>0.048</td>
<td>100</td>
</tr>
<tr>
<td>varys</td>
<td>0.05</td>
<td>0.053</td>
<td>0.173</td>
<td>0.175</td>
<td>0.050</td>
<td>100</td>
</tr>
</tbody>
</table>
Example 3. Generate a power curve for different sample sizes for a univariate LCSM

Example 2 above showed that a larger sample size is needed in order to get sufficient power for parameters betay and mys. Although one can try a difference sample size greater than 100, for convenience, we can generate a power curve with multiple different sample sizes. For example, Figure 5 shows the power curves for the two parameters betay and mys with sample sizes ranging from 100 to 200 with an interval 10. From the plot, we can easily see that to get a power 0.8 for the parameter betay, a sample size about 150 is needed. On the other hand, a sample size larger than 200 is needed for the parameter mys to have a power 0.8, with the exact number undecided based on the plot.

The R code for generating the power curve is shown in Code 5. Note that in the plot function, we refer to a specific parameter directly using its name. In the input, `seq(100, 200, 10)` generate a sequence of sample sizes and in the output, power for each sample size is provided. Code 6 shows the output when the sample sizes are 100 and 200 only to save space.

Code 5: R input script for power curve in Example 3.

```r
res <- powerLCS(N = seq(100, 200, 10), T = 5, R = 1000, betay = 0.1, my0 = 20, mys = 1.5, varey = 9, vary0 = 2.5, varys = .05, vary0ys = 0, alpha = 0.05)
plot(res, 'betay')
plot(res, 'mys')
```

Code 6: Output for generating power curves in Example 3.

```
$`N=100-T5`

<table>
<thead>
<tr>
<th>pop.par</th>
<th>mc.est</th>
<th>mc.sd</th>
<th>mc.se</th>
<th>mc.power</th>
<th>N</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>betay</td>
<td>0.10</td>
<td>0.100</td>
<td>0.044</td>
<td>0.044</td>
<td>0.627</td>
<td>100</td>
</tr>
<tr>
<td>my0</td>
<td>20.00</td>
<td>20.002</td>
<td>0.331</td>
<td>0.319</td>
<td>1.000</td>
<td>100</td>
</tr>
<tr>
<td>mys</td>
<td>1.50</td>
<td>1.505</td>
<td>1.136</td>
<td>1.119</td>
<td>0.287</td>
<td>100</td>
</tr>
<tr>
<td>varey</td>
<td>9.00</td>
<td>8.970</td>
<td>0.744</td>
<td>0.732</td>
<td>1.000</td>
<td>100</td>
</tr>
</tbody>
</table>
```
Example 4. Generate a power curve for different number of occasions for a univariate LCSM

For LCSMs, power is not only related to the sample size but also the number of measurement occasions. With the increase of the number of occasion, one would expect the increase of power. For example, Figure 6 shows the power curves for the two parameters betay and mys with the number of occasions ranging from 4 to 10 with an interval 1 and with the fixed sample size 100. From the plot, we can easily see that the power increases along with the number of measurement occasions. For example, for the same sample size 100, the power is less than 0.2 with 4 occasions of data but increases to more than 0.8 with 7 occasions of data. The R code for generating the power curve is shown in Code 7.

Code 7: R input script for power curve with the number of occasions in Example 4.
res <- powerLCS(N = 100, T = 4:10, R = 1000, betay = 0.1, my0 = 20, mys = 1.5, varey = 9, vary0 = 2.5, varys = .05, vary0ys = 0, alpha = 0.05)

**Example 5. Power analysis for a bivariate LCSM**

Power analysis can be similarly conducted for bivariate LCSMs. As an example, we use the parameter estimates from a bivariate latent change score model in Zhang et al. (2015) with some modification as population parameter values (see Table 2).

The script in Code 8 shows the R code for power analysis for the bivariate LCSM with the sample size 100. From the output in Code 9, we can see that the parameter estimates are not very accurate. This is because the bivariate LCSM requires a much larger sample size to provide accurate parameter estimates. In this case, the statistical power obtained might not be accurate either.

**Code 8: R input script for power analysis for bivariate latent change score model in Example 5.**

gpowerBLCS(N=100, T=5, R=1000, betay=0.08, my0=20, mys=1.5, varey =9, vary0=3, varys=1, vary0ys=0, alpha=0.05, betax=0.2, mx0 =20, mxs=5, varex=9, varx0=3, varxs=1, varx0xs=0, varx0y0=1, varx0ys=0, vary0xs=0, varxsys=0, gammax=0, gammay=-.1)

**Code 9: Output for power analysis in Example 5.**

<table>
<thead>
<tr>
<th>pop.par</th>
<th>mc.est</th>
<th>mc.sd</th>
<th>mc.se</th>
<th>mc.power</th>
<th>N</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>betax</td>
<td>0.20</td>
<td>0.230</td>
<td>0.260</td>
<td>0.187</td>
<td>0.241</td>
<td>100</td>
</tr>
<tr>
<td>betay</td>
<td>0.08</td>
<td>0.164</td>
<td>0.572</td>
<td>0.435</td>
<td>0.081</td>
<td>100</td>
</tr>
<tr>
<td>gammax</td>
<td>0.00</td>
<td>-0.033</td>
<td>0.234</td>
<td>0.178</td>
<td>0.112</td>
<td>100</td>
</tr>
<tr>
<td>gammay</td>
<td>-0.10</td>
<td>-0.175</td>
<td>0.641</td>
<td>0.458</td>
<td>0.075</td>
<td>100</td>
</tr>
<tr>
<td>mx0</td>
<td>20.00</td>
<td>20.004</td>
<td>0.336</td>
<td>0.326</td>
<td>1.000</td>
<td>100</td>
</tr>
<tr>
<td>mxs</td>
<td>5.00</td>
<td>5.933</td>
<td>7.848</td>
<td>5.615</td>
<td>0.167</td>
<td>100</td>
</tr>
</tbody>
</table>
Increasing the sample size will lead to more accurate results as shown in Code 10 where the sample size is 500. In planning the sample size for LCSM models, one should pay attention to the parameter estimates to make sure they are accurate enough for power calculation. Specifically for the coupling parameters gammax and gammay, the power, type I error for gammax, is 0.057 and 0.271, respectively.

Code 10: Output for power analysis in Example 5 when the sample size is 500.

<table>
<thead>
<tr>
<th>pop.par</th>
<th>mc.est</th>
<th>mc.sd</th>
<th>mc.se</th>
<th>mc.power</th>
<th>N</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>betax</td>
<td>0.20</td>
<td>0.2009</td>
<td>0.031</td>
<td>0.031</td>
<td>1.000</td>
<td>500</td>
</tr>
<tr>
<td>betay</td>
<td>0.08</td>
<td>0.0830</td>
<td>0.070</td>
<td>0.068</td>
<td>0.199</td>
<td>500</td>
</tr>
<tr>
<td>gammmax</td>
<td>0.00</td>
<td>-0.0014</td>
<td>0.030</td>
<td>0.029</td>
<td>0.057</td>
<td>500</td>
</tr>
<tr>
<td>gammay</td>
<td>-0.10</td>
<td>-0.1022</td>
<td>0.072</td>
<td>0.073</td>
<td>0.271</td>
<td>500</td>
</tr>
<tr>
<td>mx0</td>
<td>20.00</td>
<td>19.9911</td>
<td>0.145</td>
<td>0.145</td>
<td>1.000</td>
<td>500</td>
</tr>
</tbody>
</table>
Discussion and Future Directions

To complement the research on LCSMs, in this chapter, we discuss how to plan the sample size and number of measurement occasions for both univariate and bivariate LCSMs. Specifically, we illustrate how to calculate power for each individual model parameter of interest. Since the analytical solution to power is untraceable, we used a Monte Carlo based method. We also provided an R package RAMpath and an online interface to carry out the power analysis procedure.

In calculating power, we need the information on the population parameter values. Each value can be viewed as the unstandardized effect size for the parameter of interest. We did not define the standardized effect size such as Cohen’s d (Cohen, 1988) for several reasons. First, given the complexity of LCSMs, it is difficult to define a standardized effect size. Second, in
general, it is easier to specify the unstandardized effect size because when conducting a literature review, one can simply adopt the parameter estimates directly from published results. Third, if a researcher is interested in standardized measures, he/she can use the standardized coefficients as the population parameter values in conducting power analysis. But note even in such condition, power from different studies might not be comparable.

The current study can be improved in many ways in the future. First, in this chapter, we have focused on the power analysis of a single model parameter as this is the most common situation. If a researcher wants to test multiple parameters simultaneously, a procedure based on the likelihood ratio test can be developed as for the growth curve analysis in Zhang & Wang (2009).

Second, in the current study, we have focused on the basic univariate and bivariate LCSMs. Since their invention, the basic univariate and bivariate LCSMs have been extended in many ways. For example, Hamagami & McArdle (2007b) expanded the traditional specifications of univariate and bivariate LCSMs to the parallel process change score model and the second-order LCSMs. Grimm et al. (2012) extended latent difference scores to allow for testing hypotheses where recent changes, as opposed to recent levels, are a primary predictor of subsequent changes. The Monte Carlo procedure used in this study can be flexibly extended to the more advanced models.

Third, the current study has assumed that the collected data will be complete. However, in practice, missing data are almost not avoidable in longitudinal studies. For example, Puma et al. (2009) found that student achievement outcomes are often missing for 10-20% in studies funded by the National Center for Education Evaluation and Regional Assistance. Missing data reduce power and without careful consideration, a well-planned study can become under-powered. Taking into consideration of missing data in power calculation requires the specification of missing data generating mechanism that can be used in the data generation step in our Monte Carlo method.

Third, in our Monte Carlo method, we have assumed that our data are normally distributed. However, practical data often deviate from a normal distribution. For example, Micceri (1989)
evaluated 440 large-sample achievement and psychometric measures distributions and found that all of them were nonnormal. More recently, Blanca et al. (2013) evaluated nonnormality using the skewness and kurtosis of 693 small samples and found that 94.5% of them violated the normality assumption. In addition, Cain et al. (2016) reviewed 254 multivariate distributions of data used in Psychological Science and the American Education Research Journal and found that 68% multivariate distributions deviated from normal distributions. Therefore, in the future, the influence of non-normal data should be considered when estimating power.

Finally, the Monte Carlo based method can be very computationally intensive because of the involvement of the Monte Carlo simulation. For example, it took about 10 minutes on a modern desktop to complete the power analysis in Example 4. At the same time, the Monte Carlo method can be easily parallelized to take advantage of modern hardware such as multi-core processors (e.g., Zhang, 2014). In the future, the R package RAMpath can be improved with the capacity of parallelization.
References


Table 1

*Population parameter values used in Examples 1-4*

<table>
<thead>
<tr>
<th></th>
<th>Example 1</th>
<th>Example 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>betay</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>my0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>mys</td>
<td>0</td>
<td>1.5</td>
</tr>
<tr>
<td>varey</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>vary0</td>
<td>1</td>
<td>2.5</td>
</tr>
<tr>
<td>varys</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>vary0ys</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 2

*Population parameter values used in Example 5*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_y$</td>
<td>0.08</td>
<td>$\beta_x$</td>
<td>0.2</td>
</tr>
<tr>
<td>$\gamma_x$</td>
<td>0</td>
<td>$\gamma_y$</td>
<td>$-0.1$</td>
</tr>
<tr>
<td>$m_y$</td>
<td>20</td>
<td>$m_x$</td>
<td>20</td>
</tr>
<tr>
<td>$M_y$</td>
<td>1.5</td>
<td>$m_x$</td>
<td>5</td>
</tr>
<tr>
<td>$\sigma^2_y$</td>
<td>9</td>
<td>$\sigma^2_x$</td>
<td>9</td>
</tr>
<tr>
<td>$\sigma_{y0}$</td>
<td>3</td>
<td>$\sigma_{x0}$</td>
<td>3</td>
</tr>
<tr>
<td>$\sigma_y$</td>
<td>0.05</td>
<td>$\sigma_x$</td>
<td>0.6</td>
</tr>
<tr>
<td>$\sigma_{y0}$</td>
<td>0</td>
<td>$\sigma_{x0}$</td>
<td>0</td>
</tr>
<tr>
<td>$\sigma_{y0}$</td>
<td>1</td>
<td>$\sigma_{x0}$</td>
<td>$\text{M}$</td>
</tr>
<tr>
<td>$\sigma_{y0}$</td>
<td>0</td>
<td>$\sigma_{x0}$</td>
<td>$\text{M}$</td>
</tr>
<tr>
<td>$\sigma_{y0}$</td>
<td>0</td>
<td>$\sigma_{x0}$</td>
<td>$\text{M}$</td>
</tr>
<tr>
<td>$\sigma_{y0}$</td>
<td>0</td>
<td>$\sigma_{x0}$</td>
<td>$\text{M}$</td>
</tr>
</tbody>
</table>
Figure 1. The trajectory plot of latent scores $y[t]$ from time 1 to time 5 with different $\beta_y$ values
Figure 2. Path diagram for a univariate latent change score models
Figure 3. The path diagram for a bivariate latent change score model.
## Univariate Latent Change Score Model

<table>
<thead>
<tr>
<th>Parameters (Help)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>100</td>
</tr>
<tr>
<td>Number of occasions</td>
<td>5</td>
</tr>
<tr>
<td>Number of replications</td>
<td>1000</td>
</tr>
<tr>
<td>betay</td>
<td>0</td>
</tr>
<tr>
<td>my0</td>
<td>0</td>
</tr>
<tr>
<td>mys</td>
<td>0</td>
</tr>
<tr>
<td>varyy</td>
<td>1</td>
</tr>
<tr>
<td>vary0</td>
<td>1</td>
</tr>
<tr>
<td>varies</td>
<td>1</td>
</tr>
<tr>
<td>vary0ys</td>
<td>0</td>
</tr>
<tr>
<td>Significance level</td>
<td>0.05</td>
</tr>
<tr>
<td>Power</td>
<td></td>
</tr>
<tr>
<td>Power curve</td>
<td>No power curve</td>
</tr>
<tr>
<td>Note</td>
<td>Univariate Latent Change Score</td>
</tr>
</tbody>
</table>

*Figure 4.* The online interface for power analysis for univariate latent change score models.
Figure 5. Power curve for betay and mys along with the sample size in the univariate latent change score model
Figure 6. Power curve for betay and mys along with the number of measurement occasions in the univariate latent change score model