Structural Equation Modeling

P.12 - Growth Curve Analysis

01.12.2021

Lab Description

In this assignment you are going to learn how to estimate a *Latent Growth Curve* (LGC) model in `R` using the `lavaan`. In the previous lab you learn how you can use hierarchical models (i.e., multi-level models for nested data). In this lab, we are are still going to work with nested data (i.e., observations nested under participants), but this time we are going to approach the problem from the perspective of *Structural Equation Models* (SEM) For this practical you will need the following packages: `lavaan` and `semPlot`. You can install and load these packages using the following code:

```
# Install packages.
install.packages(c("lavaan", "semPlot"))

# Load the packages.
library(lavaan)
library(semPlot)
```

Exercise 1

In this exercise, you are going to estimate a LGC model similar as the one depicted in *Figure 1*, but with six waves instead of four. Specifically, you are going to investigate changes in body weight over 12 years (i.e., six waves of data separated by two years each) using the health and aging data set `health.dat` with $N = 5335$. Body weight was operationalized as the *Body Mass Index* (BMI), which is a ratio of weight to square of height (i.e., kg/m^2). The data set `health.dat` is available on Canvas in the module corresponding to this lab. You can use the code below to load the data and set the variables names.

Set the working directory to the location where your data file has been downloaded and load the data.

```
# For example.
setwd("/Users/mihai/Downloads")

# Load data.
data <- read.table("health.dat")

# Inspect the data.
```
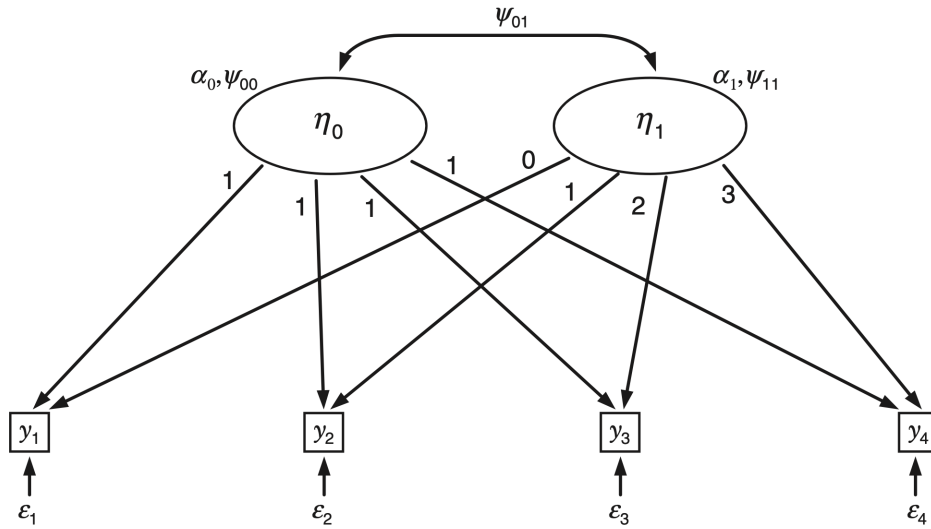
Figure 1: Example of a LGT model reproduced from Newsom (2015, p. 174).

```
View(data)
```

Set the variable names.

```
# Variable names.
variable_names_ex_1 = c(
    "age", "srh1", "srh2", "srh3", "srh4", "srh5", "srh6", "bmi1",
    "bmi2", "bmi3", "bmi4", "bmi5", "bmi6", "cesdna1", "cesdpa1", "cesdso1",
    "cesdna2", "cesdpa2", "cesdso2", "cesdna3", "cesdpa3", "cesdso3",
    "cesdna4", "cesdpa4", "cesdso4", "cesdna5", "cesdpa5", "cesdso5",
    "cesdna6", "cesdpa6", "cesdso6", "diab1", "diab2", "diab3 ", "diab4", "diab5", "diab6"
)

# Set the names.
names(data) <- variable_names_ex_1
```

List the variables.

```
# List variables.
str(data)
```

a. Estimate a LGC model in which you set each intercept factor loading equal to 1 and the slope factor loadings equal to 0, 1, 2, 3, 4, and 5. Do not include correlated measurement residuals in this model. Evaluate the fit of this model, and interpret the mean of the latent intercept and mean of the latent slope.

b. Re-estimate the model from point *(a)*, but now add auto-correlations among adjacent time points of the measurement residuals (e.g., ε_1 with ε_2, ε_2 with ε_3, and so on). Evaluate the fit of this model and test its fit against the more restricted model estimated at point *(a)*.

c. In the model estimated at point *(b)*, inspect the estimate for the covariance between the latent intercept

and the latent slope. Provide a substantive interpretation for this value.

d. Estimate a model that assumes homogeneity of variance of the measurement residuals. In this model, remove the correlated measurement residuals so you can test this model against the unconstrained model that was estimated at point *(a)*. What can you conclude from the comparison of both models, and from the model that included the auto-correlations between error terms?

Exercise 2

In this exercise you are going to keep using the `health.dat` and extend the model from *Exercise 1* with a time-invariant covariate, namely the age of the participants.

a. Compute a new variable `age_c` which is the grand mean centered age of the participants.

- *Tip.* You can obtain the sample mean using `mean(data$age)`.

b. Estimate the conditional *LGC* model of *BMI* by regressing the latent intercept and the latent slope on `age_c`.

c. Evaluate the fit of this model, and the effect of `age_c` on the latent intercept and latent slope.

d. Compute a new binary age variable called `age_b` for implementing the *MIMIC* approach discussed during *Lecture 8*, using the following `R` code:

```
data$age_b <- ifelse(data$age <= 65, 0, 1)
```

e. Estimate the *MIMIC* model by regressing the intercept and the slope factors on the binary age variable. Interpret the effects of the binary age variable.

f. Test the same hypothesis of age differences in trajectories, using the multi-group approach and comparing those under 65 to those aged 65 and older.

References

Newsom, J. T. (2015). *Longitudinal structural equation modeling: A comprehensive introduction.* Routledge, Taylor and Francis Group.