Longitudinal Data Analysis

methods@manchester summer school

Day 2 | afternoon session

Thiago R. Oliveira

Lecturer in Quantitative Criminology, University of Manchester

30/06—04/07

Today

Growth curve models: a SEM approach

- ⇝ Introduction to Structural Equation Models
- ⇝ Measurement models and structural models
- ⇝ Repeated measurements as a multivariate response
- ⇝ Growth curve model as a factor model
- ⇝ Multilevel vs. SEM approaches

The structural equation modelling framework

The structural equation modelling framework

- ⇝ SEM = a general framework for specifying and estimating statistical models that include:
  - Measurement models (confirmatory factor analysis)
  - Structural (causal/path) models
- ⇝ Extends traditional regression by incorporating:
  - Latent (unobserved) variables
  - Measurement error
  - Simultaneous equations
- ⇝ SEM is especially powerful for longitudinal data

Brief overview of factor analysis

Factor analysis aims to explain the covariances among a set of observed variables using one or more latent variables (factors)

Factor analysis is most often applied to cross-sectional data where a latent construct is measured by a set of indicators, e.g.

Latent variable	Observed indicators
IQ ('intelligence')	Scores on set of test items
Area deprivation	Average income, unemployment rate, etc.
Trust in government	Responses to several survey questions
Internalising behaviour	Set of survey and test items

Latent variables are **unobserved constructs** that we assume to exist

We *infer* their distribution from a multivariate response

Components of an SEM

Measurement model:

$$y = \Lambda\eta + \epsilon$$

- ⇝ Links observed variables y to latent variables η
- ⇝ Factor loadings Λ
- ⇝ Measurement error ϵ

Structural model:

$$\eta = B\eta + \zeta$$

- ⇝ Describes relationships among latent variables
- ⇝ B: regression among η
- ⇝ ζ: structural disturbance

Latent growth curve analysis

Factor analysis of longitudinal data

In the application of factor analysis to repeated measures data:

- $\rightsquigarrow$ The observed variables are the responses over time $(y_{1i}, y_{2i}, \ldots, y_{Ti})$
- $\rightsquigarrow$ We can use this multivariate response of repeated measures to *infer* a latent construct
- $\rightsquigarrow$ Example: Depression score at T1, T2, T3 $\rightarrow$ 3 observed indicators
- $\rightsquigarrow$ Could they be manifestations of *underlying growth factors*?

$\Rightarrow$ This leads to the idea of a **latent growth curve model**

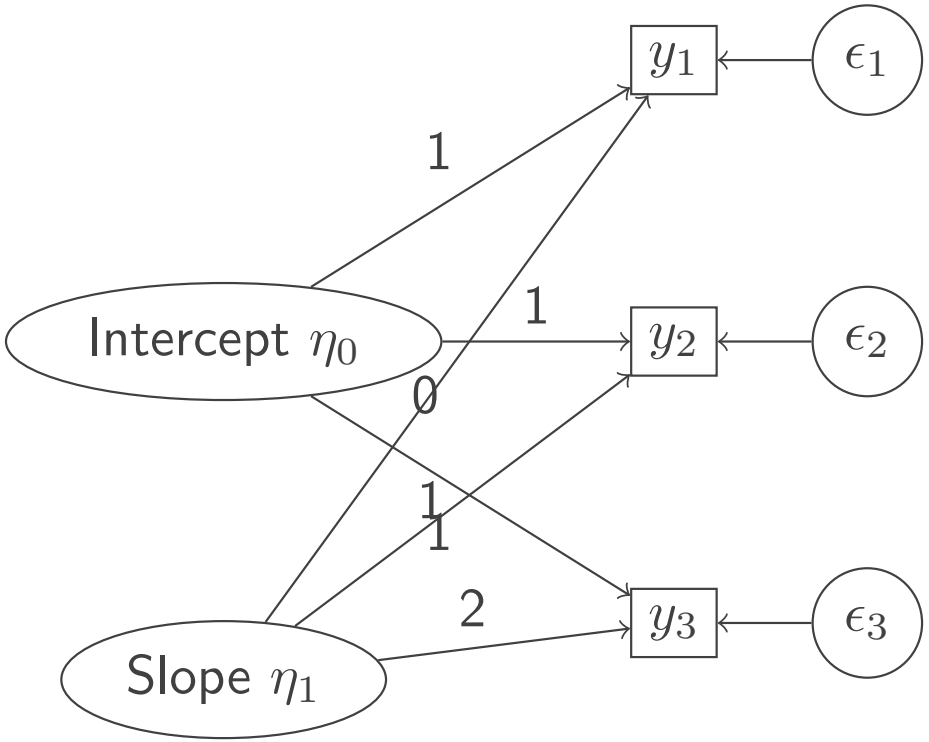
LGCM as a factor model

- ⇝ Let Y_t be repeated measurements over time ($t = 1, ..., T$)
- ⇝ Each Y_t is modeled as:

$$Y_t = \lambda_{0t}\eta_0 + \lambda_{1t}\eta_1 + \epsilon_t$$

- ⇝ η_0: latent intercept
  η_1: latent slope
- ⇝ Loadings λ_{0t}, λ_{1t} define the shape of the trajectory

Latent growth curve model



- ⇝ Latent intercept η_0: starting point
- ⇝ Latent slope η_1: rate of change
- ⇝ Time-specific residuals ϵ_t
- ⇝ Factor loadings reflect time (e.g., 0, 1, 2, ...)

Multilevel modelling vs SEM for growth curve analysis

Multilevel modelling vs SEM for growth curve analysis

A growth curve model can be framed as either a multilevel or a structural equation model

Multilevel

- ⇝ View data as two-level hierarchy
- ⇝ Data in long format
- ⇝ Time treated as time-varying predictor
- ⇝ Can specify highly flexible nonlinear functions of time
- ⇝ Easy to allow for between-individual variation in timing of measurements
- ⇝ More computationally efficient

SEM

- ⇝ View data as multivariate response
- ⇝ Data in wide format
- ⇝ Time specified through loadings on slope factor
- ⇝ Easy to incorporate latent constructs into the model when outcome is measured by multiple indicators
- ⇝ Easy to expand to bivariate latent growth curve models
- ⇝ Easy to extend to other models (group-based trajectory models, latent change score models, . . .)

Thank you!

thiago.oliveira@manchester.ac.uk

ThiagoROliveira.com

@oliveiratr.bsky.social