Mplus Short Courses
Day 2

Growth Modeling With Latent Variables Using Mplus

Linda K. Muthén
Bengt Muthén

www.statmodel.com

1

Table Of Contents

2

Mplus Background

- Inefficient dissemination of statistical methods:
  - Many good methods contributions from biostatistics, psychometrics, etc are underutilized in practice
- Fragmented presentation of methods:
  - Technical descriptions in many different journals
  - Many different pieces of limited software
- Mplus: Integration of methods in one framework
  - Easy to use: Simple, non-technical language, graphics
  - Powerful: General modeling capabilities
- Mplus versions
  - V1: November 1998
  - V2: February 2001
  - V3: March 2004
  - V4: February 2006
- Mplus team: Linda & Bengt Muthén, Thuy Nguyen, Tihomir Asparouhov, Michelle Conn

3

Statistical Analysis With Latent Variables A General Modeling Framework

Statistical Concepts Captured By Latent Variables

Continuous Latent Variables

- Measurement errors
- Factors
- Random effects
- Frailties, liabilities
- Variance components
- Missing data

Categorical Latent Variables

- Latent classes
- Clusters
- Finite mixtures
- Missing data

4

Statistical Analysis With Latent Variables A General Modeling Framework (Continued)

Models That Use Latent Variables

Continuous Latent Variables

- Factor analysis models
- Structural equation models
- Growth curve models
- Multilevel models

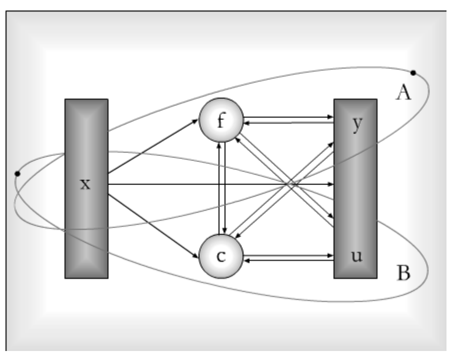
Categorical Latent Variables

- Latent class models
- Mixture models
- Discrete-time survival models
- Missing data models

Mplus integrates the statistical concepts captured by latent variables into a general modeling framework that includes not only all of the models listed above but also combinations and extensions of these models.

5

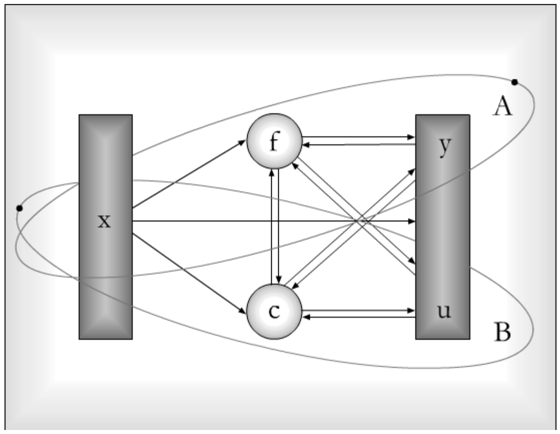
General Latent Variable Modeling Framework



- Observed variables
  - x background variables (no model structure)
  - y continuous and censored outcome variables
  - u categorical (dichotomous, ordinal, nominal) and count outcome variables
- Latent variables
  - f continuous variables
    - – interactions among f's
  - c categorical variables
    - – multiple c's

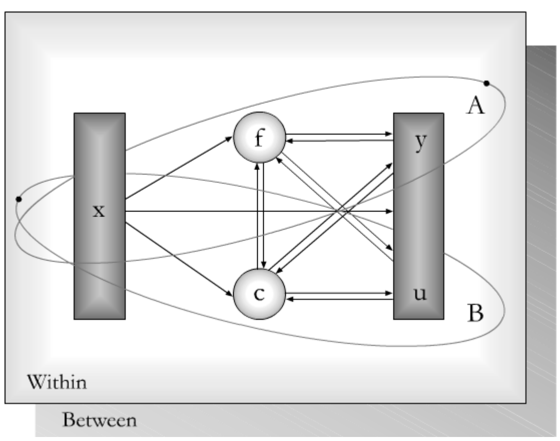
6

General Latent Variable Modeling Framework



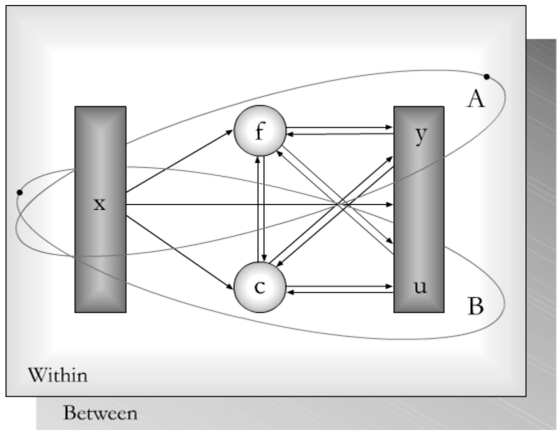
7

General Latent Variable Modeling Framework



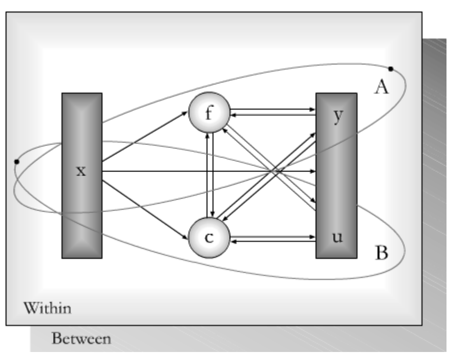
8

General Latent Variable Modeling Framework



9

General Latent Variable Modeling Framework



- Observed variables
  - x background variables (no model structure)
  - y continuous and censored outcome variables
  - u categorical (dichotomous, ordinal, nominal) and count outcome variables
- Latent variables
  - f continuous variables
    - – interactions among f's
  - c categorical variables
    - – multiple c's

10

Overview

Single-Level Analysis

	Cross-Sectional	Longitudinal
Continuous Observed And Latent Variables	***Day 1*** Regression Analysis Path Analysis Exploratory Factor Analysis Confirmatory Factor Analysis Structural Equation Modeling	***Day 2*** Growth Analysis
Adding Categorical Observed And Latent Variables	***Day 3*** Regression Analysis Path Analysis Exploratory Factor Analysis Confirmatory Factor Analysis Structural Equation Modeling Latent Class Analysis Factor Mixture Analysis Structural Equation Mixture Modeling	***Day 4*** Latent Transition Analysis Latent Class Growth Analysis Growth Analysis Growth Mixture Modeling Discrete-Time Survival Mixture Analysis Missing Data Analysis

11

Overview (Continued)

Multilevel Analysis

	Cross-Sectional	Longitudinal
Continuous Observed And Latent Variables	***Day 5*** Regression Analysis Path Analysis Exploratory Factor Analysis Confirmatory Factor Analysis Structural Equation Modeling	***Day 5*** Growth Analysis
Adding Categorical Observed And Latent Variables	***Day 5*** Latent Class Analysis Factor Mixture Analysis	***Day 5*** Growth Mixture Modeling

12

Typical Examples Of Growth Modeling

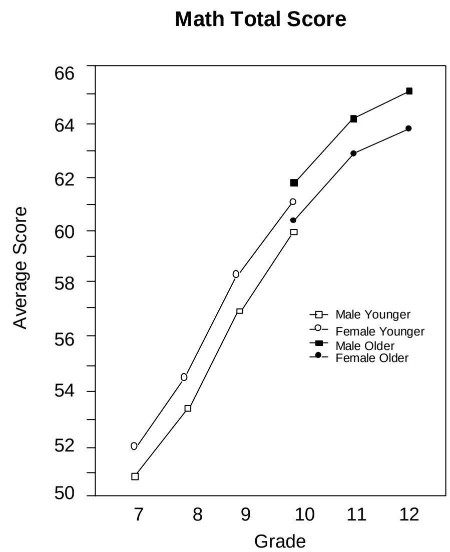
13

LSAY Data

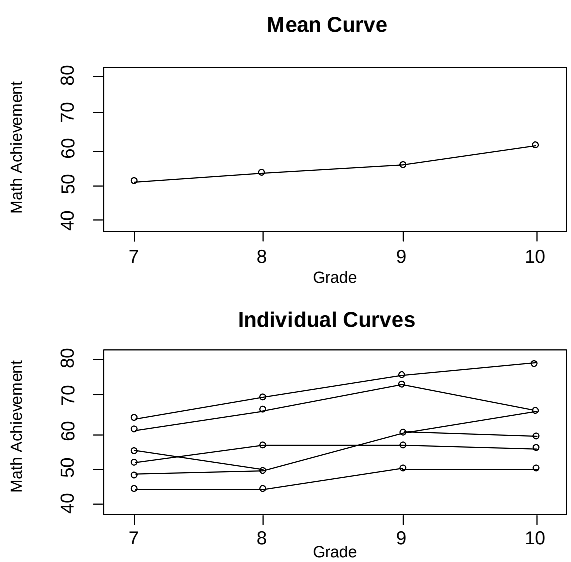
Longitudinal Study of American Youth (LSAY)

- Two cohorts measured each year beginning in 1987
  - Cohort 1 - Grades 10, 11, and 12
  - Cohort 2 - Grades 7, 8, 9, and 10
- Each cohort contains approximately 60 schools with approximately 60 students per school
- Variables - math and science achievement items, math and science attitude measures, and background variables from parents, teachers, and school principals
- Approximately 60 items per test with partial item overlap across grades - adaptive tests

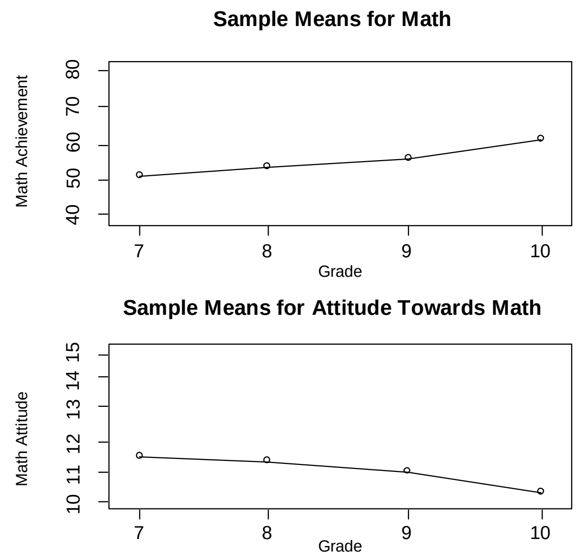
14


Math Total Score
Average Score
Grade
66
64
62
60
58
56
54
52
50
7
8
9
10
11
12
Male Younger
Female Younger
Male Older
Female Older

15


Mean Curve
Math Achievement
Grade
40
50
60
70
80
7
8
9
10
Individual Curves
Math Achievement
Grade
40
50
60
70
80
7
8
9
10

16



17

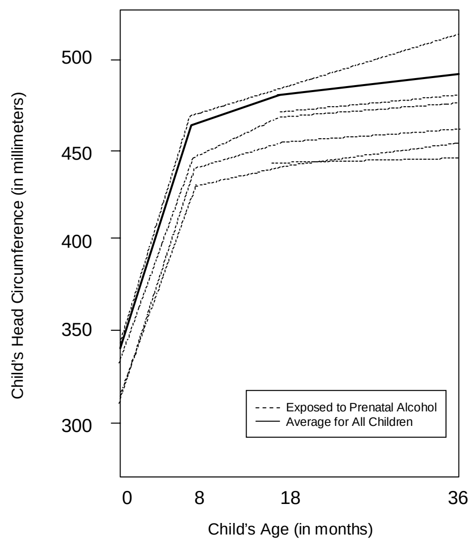
Maternal Health Project Data

Maternal Health Project (MHP)

- Mothers who drank at least three drinks a week during their first trimester plus a random sample of mothers who used alcohol less often
- Mothers measured at fourth month and seventh month of pregnancy, at delivery, and at 8, 18, and 36 months postpartum
- Offspring measured at 0, 8, 18 and 36 months
- Variables for mothers - demographic, lifestyle, current environment, medical history, maternal psychological status, alcohol use, tobacco use, marijuana use, other illicit drug use
- Variables for offspring - head circumference, height, weight, gestational age, gender, and ethnicity

18

MHP: Offspring Head Circumference



19

Basic Modeling Ideas

20

Longitudinal Data: Three Approaches

Three modeling approaches for the regression of outcome on time (n is sample size, T is number of timepoints):

- **Use all *n* x *T* data points to do a single regression analysis**: Gives an intercept and a slope estimate for all individuals - does not account for individual differences or lack of independence of observations
- **Use each individual's *T* data points to do *n* regression analyses:** Gives an intercept and a slope estimate for each individual. Accounts for individual differences, but does not account for similarities among individuals
- **Use all *n* x *T* data points to do a single random effect regression analysis:** Gives an intercept and a slope estimate for each individual. Accounts for similarities among individuals by stipulating that all individuals' random effects come from a single, common population and models the non-independence of observations as show on the next page

Individual Development Over Time

(1) $y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}$

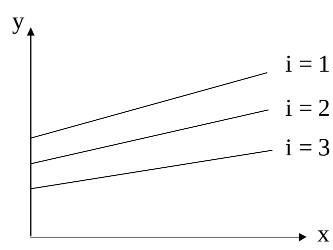
(2a) $\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}$

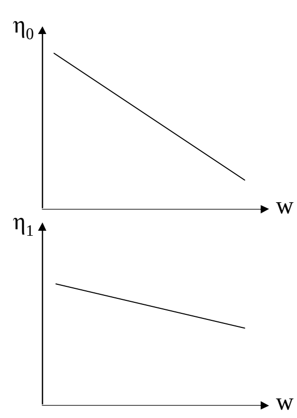
(2b) $\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}$

t = timepoint   i = individual   w = time-invariant covariate

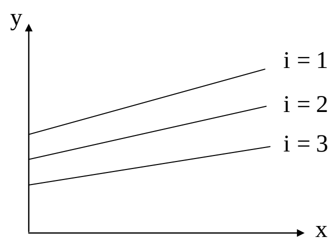
y = outcome   x = time score

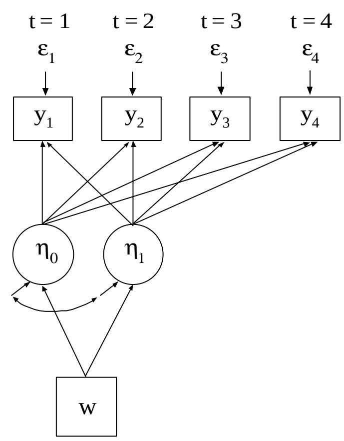
η_0 = intercept η_1 = slope





Individual Development Over Time





(1) $y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}$

(2a) $\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}$

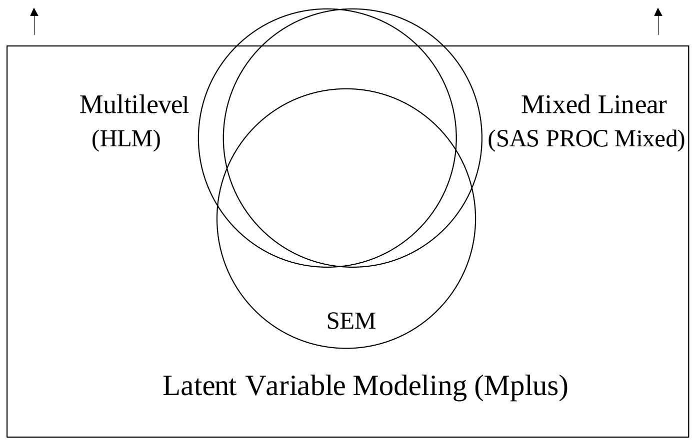
(2b) $\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}$

23

Growth Modeling Frameworks

24

Growth Modeling Frameworks/Software



25

Comparison Summary Of Multilevel, Mixed Linear, And SEM Growth Models

- Multilevel and mixed linear models are the same
- SEM differs from the multilevel and mixed linear models in two ways
  - Treatment of time scores
    - Time scores are data for multilevel and mixed linear models -- individuals can have different times of measurement
    - Time scores are parameters for SEM growth models -- time scores can be estimated
  - Treatment of time-varying covariates
    - Time-varying covariates have random effect coefficients for multilevel and mixed linear models -- coefficients vary over individuals
    - Time-varying covariates have fixed effect coefficients for SEM growth models -- coefficients vary over time

26

Random Effects: Multilevel And Mixed Linear Modeling

Individual i (i = 1, 2, …, n) observed at time point t (t = 1, 2, … T).

Multilevel model with two levels (e.g. Raudenbush & Bryk, 2002, HLM).

- Level 1: $y_{ti} = \eta_{0i} + \eta_{1i}\, x_{ti} + \kappa_i\, w_{ti} + \varepsilon_{ti}$ (39)
- Level 2: $\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}$ (40)
  $\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}$ (41)
  $\kappa_i = \alpha + \gamma\, w_i + \zeta_i$ (42)

27

Random Effects: Multilevel And Mixed Linear Modeling (Continued)

Mixed linear model:

$$y_{ti} = \textit{fixed part} + \textit{random part} \quad (43)$$
$$= \alpha_0 + \gamma_0\, w_i + (\alpha_1 + \gamma_1\, w_i)\, x_{ti} + (\alpha + \gamma\, w_i)\, w_{ti} \quad (44)$$
$$+ \zeta_{0i} + \zeta_{1i}\, x_{ti} + \zeta_i\, w_{ti} + \varepsilon_{ti}\,. \quad (45)$$

E.g. "*time* x w_i" refers to γ_1 (e.g. Rao, 1958; Laird & Ware, 1982; Jennrich & Sluchter, 1986; Lindstrom & Bates, 1988; BMDP5V; Goldstein, 1995, MLn; SAS PROC MIXED - Littell et al. 1996 and Singer, 1999).

28

Random Effects: SEM And Multilevel Modeling

SEM (Tucker, 1958; Meredith & Tisak, 1990; McArdle & Epstein 1987; SEM software):

Measurement part:

$$y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \kappa_t\, w_{ti} + \varepsilon_{ti}\,. \qquad (46)$$

Compare with level 1 of multilevel:

$$y_{ti} = \eta_{0i} + \eta_{1i}\, x_{ti} + \kappa_i\, w_{ti} + \varepsilon_{ti}\,. \qquad (47)$$

Multilevel approach:

- x_{ti} as data: Flexible individually-varying times of observation
- Slopes for time-varying covariates vary over individuals

Random Effects: SEM And Multilevel Modeling (Continued)

SEM approach:

- x_t as parameters: Flexible growth function form
- Slopes for time-varying covariates vary over time points

Structural part (same as level 2, except for κ_t):

$$\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}\,, \qquad (48)$$

$$\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}\,, \qquad (49)$$

κ_t not involved (parameter).

Random Effects: Mixed Linear Modeling And SEM

Mixed linear model in matrix form:

$$\boldsymbol{y}_i = (y_{1i}, y_{2i}, \ldots, y_{Ti})' \qquad (51)$$

$$= \boldsymbol{X}_i \boldsymbol{\alpha} + \boldsymbol{Z}_i \boldsymbol{b}_i + \boldsymbol{e}_i . \qquad (52)$$

Here, $\boldsymbol{X}, \boldsymbol{Z}$ are design matrices with known values, $\boldsymbol{\alpha}$ contains fixed effects, and $\boldsymbol{b}$ contains random effects. Compare with (43) - (45).

31

Random Effects: Mixed Linear Modeling And SEM (Continued)

SEM in matrix form:

$$y_i = \nu + \Lambda\, \eta_i + K\, x_i + \varepsilon_i , \qquad (53)$$

$$\eta_i = \alpha + B\, \eta_i + \Gamma\, x_i + \zeta_\mathrm{i} . \qquad (54)$$

$$y_i = \textit{fixed part} + \textit{random part}$$

$$= \mathrm{v} + \Lambda\, (I - B)^{-1}\, \alpha + \Lambda\, (I - B)^{-1}\, \Gamma\, x_i + K\, x_i$$

$$+ \Lambda\, (I - B)^{-1}\, \zeta_\mathrm{i} + \varepsilon_i .$$

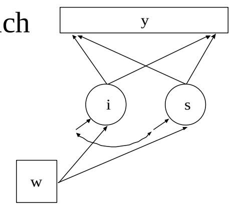
Assume $x_{ti} = x_t$, $\kappa_i = \kappa_t$ in (39). Then (39) is handled by (53) and (40) – (41) are handled by (54), putting x_t in Λ and w_{ti}, w_i in x_i.

Need for Λ_i, K_i, B_i, Γ_i.

32

Growth Modeling Approached In Two Ways: Data Arranged As Wide Versus Long

- Wide: Multivariate, Single-Level Approach

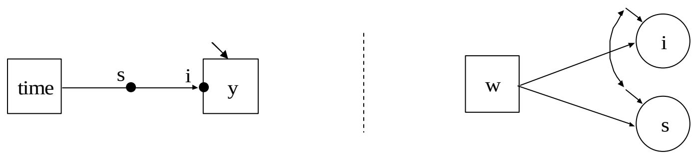


$y_{ti} = i_i + s_i \times \text{time}_{ti} + \varepsilon_{ti}$

i_i regressed on w_i
s_i regressed on w_i

- Long: Univariate, 2-Level Approach (CLUSTER = id)

Within   Between



The intercept *i* is called *y* in Mplus

33

Multilevel Modeling In A Latent Variable Framework

Integrating multilevel and SEM analyses (Asparouhov & Muthén, 2002).

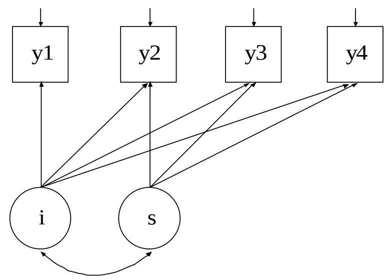
Flexible combination of random effects and other latent variables:

- Multilevel models with random effects (intercepts, slopes)
  - Individually-varying times of observation read as data
  - Random slopes for time-varying covariates
- SEM with factors on individual and cluster levels
- Models combining random effects and factors, e.g.
  - Cluster-level latent variable predictors with multiple indicators
  - Individual-level latent variable predictors with multiple indicators
- Special applications
  - Random coefficient regression (no clustering; heteroscedasticity)
  - Interactions between continuous latent variables and observed variables

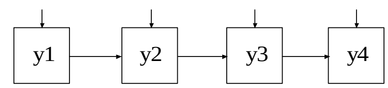
34

Alternative Models For Longitudinal Data

Growth Curve Model



Auto-Regressive Model



Hybrid Models

Curran & Bollen (2001)
McArdle & Hamagami (2001)

35

Advantages Of Growth Modeling In A Latent Variable Framework

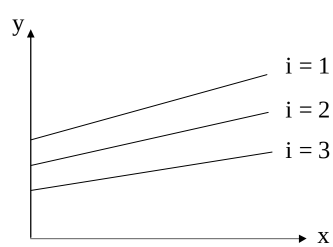
- Flexible curve shape
- Individually-varying times of observation
- Regressions among random effects
- Multiple processes
- Modeling of zeroes
- Multiple populations
- Multiple indicators
- Embedded growth models
- Categorical latent variables: growth mixtures

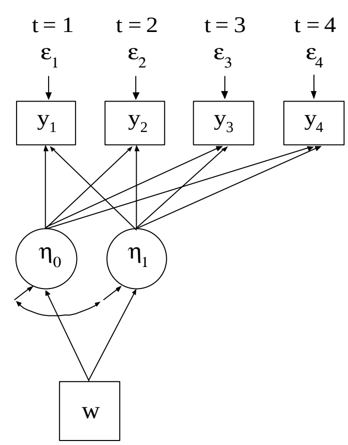
36

The Latent Variable Growth Model In Practice

37

Individual Development Over Time





(1) $y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}$

(2a) $\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}$

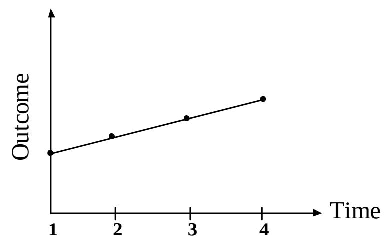
(2b) $\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}$

38

Specifying Time Scores For Linear Growth Models

Linear Growth Model

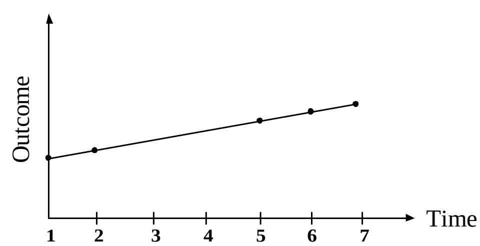
- Need two latent variables to describe a linear growth model: Intercept and slope



- Equidistant time scores for slope:

Equidistant time scores	0	1	2	3
for slope:	0	.1	.2	.3

Specifying Time Scores For Linear Growth Models (Continued)



- Nonequidistant time scores for slope:

Nonequidistant time scores	0	1	4	5	6
for slope:	0	.1	.4	.5	.6

Interpretation Of The Linear Growth Factors

Model:

$$y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}\,, \qquad (17)$$

where in the example t = 1, 2, 3, 4 and x_t = 0, 1, 2, 3:

$$y_{1i} = \eta_{0i} + \eta_{1i}\, 0 + \varepsilon_{1i}\,, \qquad (18)$$
$$\eta_{0i} = y_{1i} - \varepsilon_{1i}, \qquad (19)$$
$$y_{2i} = \eta_{0i} + \eta_{1i}\, 1 + \varepsilon_{2i}\,, \qquad (20)$$
$$y_{3i} = \eta_{0i} + \eta_{1i}\, 2 + \varepsilon_{3i}\,, \qquad (21)$$
$$y_{4i} = \eta_{0i} + \eta_{1i}\, 3 + \varepsilon_{4i}\,. \qquad (22)$$

41

Interpretation Of The Linear Growth Factors (Continued)

Interpretation of the intercept growth factor
η_{0i} (initial status, level):
Systematic part of the variation in the outcome variable at the time point where the time score is zero.

- Unit factor loadings

Interpretation of the slope growth factor
η_{1i} (growth rate, trend):
Systematic part of the increase in the outcome variable for a time score increase of one unit.

- Time scores determined by the growth curve shape

42

Interpreting Growth Model Parameters

- Intercept Growth Factor Parameters
  - Mean
    - Average of the outcome over individuals at the timepoint with the time score of zero;
    - When the first time score is zero, it is the intercept of the average growth curve, also called initial status
  - Variance
    - Variance of the outcome over individuals at the timepoint with the time score of zero, excluding the residual variance

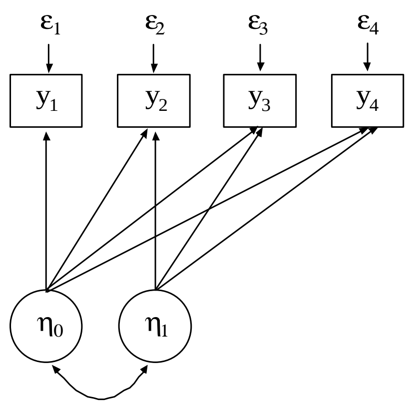
43

Interpreting Growth Model Parameters (Continued)

- Linear Slope Growth Factor Parameters
  - Mean – average growth rate over individuals
  - Variance – variance of the growth rate over individuals
  - Covariance with Intercept – relationship between individual intercept and slope values
- Outcome Parameters
  - Intercepts – not estimated in the growth model – fixed at zero to represent measurement invariance
  - Residual Variances – time-specific and measurement error variation
  - Residual Covariances – relationships between time-specific and measurement error sources of variation across time

44

Latent Growth Model Parameters And Sources Of Model Misfit



45

Latent Growth Model Parameters For Four Time Points

Linear growth over four time points, no covariates.

Free parameters in the H_1 unrestricted model:

- 4 means and 10 variances-covariances

Free parameters in the H_0 growth model:

(9 parameters, 5 d.f.):

- Means of intercept and slope growth factors
- Variances of intercept and slope growth factors
- Covariance of intercept and slope growth factors
- Residual variances for outcomes

Fixed parameters in the H_0 growth model:

- Intercepts of outcomes at zero
- Loadings for intercept growth factor at one
- Loadings for slope growth factor at time scores
- Residual covariances for outcomes at zero

46

Latent Growth Model Sources Of Misfit

Sources of misfit:
- Time scores for slope growth factor
- Residual covariances for outcomes
- Outcome variable intercepts
- Loadings for intercept growth factor

Model modifications:
- Recommended
  - Time scores for slope growth factor
  - Residual covariances for outcomes
- Not recommended
  - Outcome variable intercepts
  - Loadings for intercept growth factor

47

Latent Growth Model Parameters For Three Time Points

Linear growth over three time points, no covariates.

Free parameters in the H_1 unrestricted model:
- 3 means and 6 variances-covariances

Free parameters in the H_0 growth model
(8 parameters, 1 d.f.)
- Means of intercept and slope growth factors
- Variances of intercept and slope growth factors
- Covariance of intercept and slope growth factors
- Residual variances for outcomes

Fixed parameters in the H_0 growth model:
- Intercepts of outcomes at zero
- Loadings for intercept growth factor at one
- Loadings for slope growth factor at time scores
- Residual covariances for outcomes at zero

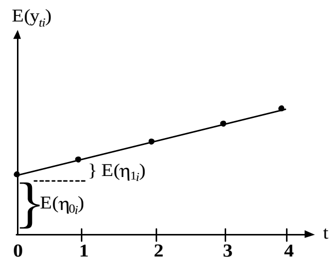
48

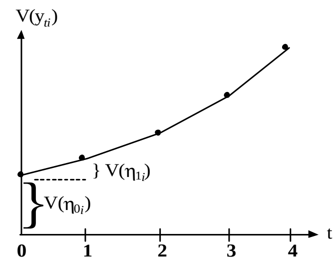
Growth Model Means And Variances

$$y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}\,,$$
$$x_t = 0, 1, \ldots, T-1.$$

Expectation (mean; E) and variance (V):

$$E\,(y_{ti}) = E\,(\eta_{0i}\,) + E\,(\eta_{1i})\, x_t\,,$$
$$V\,(y_{ti}) = V\,(\eta_{0i}\,) + V\,(\eta_{1i})\, x_t^2$$
$$+\; 2x_t\, Cov\,(\eta_{0i}\,, \eta_{1i}) + \mathrm{V}\,(\varepsilon_{ti})$$





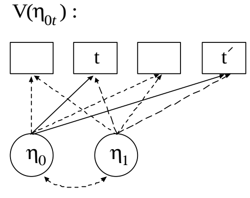
$\mathrm{V}(\varepsilon_{ti})$ constant over t
$\mathrm{Cov}(\eta_0\,, \eta_1) = 0$

Growth Model Covariances

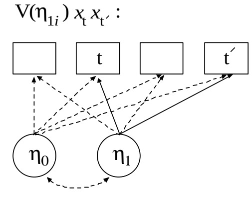
$$y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}\,,$$
$$x_t = 0, 1, \ldots, T-1.$$

$$Cov(y_{ti}\,, y_{t'i}\,) = V(\eta_{0i}) + V(\eta_{1i})\, x_t\, x_{t'}$$
$$+\; Cov(\eta_{0i}\,, \eta_{1i})\,(x_t + x_{t'})$$
$$+\; Cov(\varepsilon_{ti}\,, \varepsilon_{t'i}\,).$$

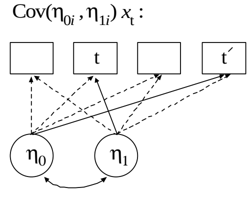
$\mathrm{V}(\eta_{0t})$:



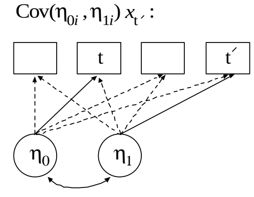
$\mathrm{V}(\eta_{1i})\, x_t\, x_{t'}$:



$\mathrm{Cov}(\eta_{0i}\,, \eta_{1i})\, x_t$:



$\mathrm{Cov}(\eta_{0i}\,, \eta_{1i})\, x_{t'}$:



Growth Model Estimation, Testing, And Model Modification

- Estimation: Model parameters
  - Maximum-likelihood (ML) estimation under normality
  - ML and non-normality robust s.e.'s
  - Quasi-ML (MUML): clustered data (multilevel)
  - WLS: categorical outcomes
  - ML-EM: missing data, mixtures
- Model Testing
  - Likelihood-ratio chi-square testing; robust chi square
  - Root mean square of approximation (RMSEA): Close fit (≤ .05)
- Model Modification
  - Expected drop in chi-square, EPC
- Estimation: Individual growth factor values (factor scores)
  - Regression method – Bayes modal – Empirical Bayes
  - Factor determinacy

51

Alternative Growth Model Parameterizations

Parameterization 1 – for continuous outcomes

$$y_{ti} = \mathbf{0} + \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}\,, \qquad (32)$$

$$\eta_{0i} = \boldsymbol{\alpha}_0 + \zeta_{0i}\,, \qquad (33)$$

$$\eta_{1i} = \alpha_1 + \zeta_{1i}\,. \qquad (34)$$

Parameterization 2 – for categorical outcomes and multiple indicators

$$y_{ti} = \boldsymbol{v} + \eta_{0i} + \eta_{1i}\, x_t + \varepsilon_{ti}\,, \qquad (35)$$

$$\eta_{0i} = \mathbf{0} + \zeta_{0i}\,, \qquad (36)$$

$$\eta_{1i} = \alpha_1 + \zeta_{1i}\,. \qquad (37)$$

52

Alternative Growth Model Parameterizations

Parameterization 1 – for continuous outcomes

- Outcome variable intercepts fixed at zero
- Growth factor means free to be estimated

```
MODEL:    i BY y1-y4@1;
          s BY y1@0 y2@1 y3@2 y4@3;
          [y1-y4@0 i s];
```

Parameterization 2 – for categorical outcomes and multiple indicators

- Outcome variable intercepts constrained to be equal
- Intercept growth factor mean fixed at zero

```
MODEL:    i BY y1-y4@1;
          s BY y1@0 y2@1 y3@2 y4@3;
          [y1-y4] (1);
          [i@0 s];
```

53

Simple Examples Of Growth Modeling

54

Steps In Growth Modeling

- Preliminary descriptive studies of the data: means, variances, correlations, univariate and bivariate distributions, outliers, etc.
- Determine the shape of the growth curve from theory and/or data
    - Individual plots
    - Mean plot
- Consider change in variance across time
- Fit model without covariates using fixed time scores
- Modify model as needed
- Add covariates

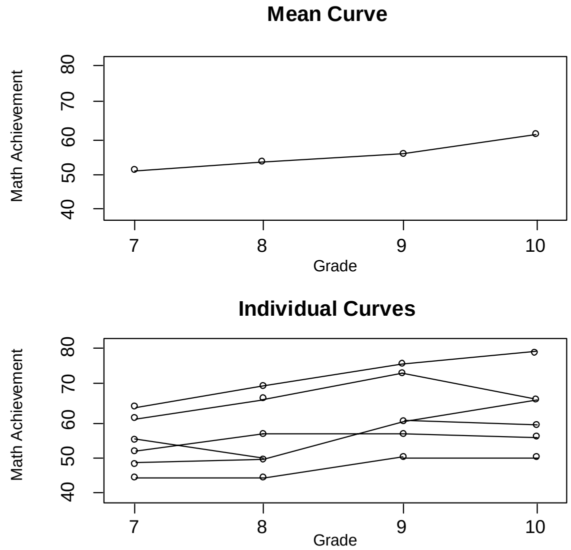
55

LSAY Data

The data come from the Longitudinal Study of American Youth (LSAY). Two cohorts were measured at four time points beginning in 1987. Cohort 1 was measured in Grades 10, 11, and 12. Cohort 2 was measured in Grades 7, 8, 9, and 10. Each cohort contains approximately 60 schools with approximately 60 students per school. The variables measured include math and science achievement items, math and science attitude measures, and background information from parents, teachers, and school principals. There are approximately 60 items per test with partial item overlap across grades – adaptive tests.

Data for the analysis include the younger females. The variables include math achievement from Grades 7, 8, 9, and 10 and the background variables of mother's education and home resources.

56



57

Input For LSAY TYPE=BASIC Analysis

```
TITLE:      LSAY For Younger Females With Listwise Deletion
            TYPE=BASIC Analysis
DATA:       FILE IS lsay.dat;
            FORMAT IS 3F8.0 F8.4 8F8.2 3F8.0;
VARIABLE:   NAMES ARE cohort id school weight math7 math8 math9
            math10 att7 att8 att9 att10 gender mothed homeres;
            USEOBS = (gender EQ 1 AND cohort EQ 2);
            MISSING = ALL (999);
            USEVAR = math7-math10;
ANALYSIS:   TYPE = BASIC;
PLOT:       TYPE = PLOT1;
```

58

Sample Statistics For LSAY Data

n = 984

Means

	MATH7	MATH8	MATH9	MATH10
	52.750	55.411	59.128	61.796

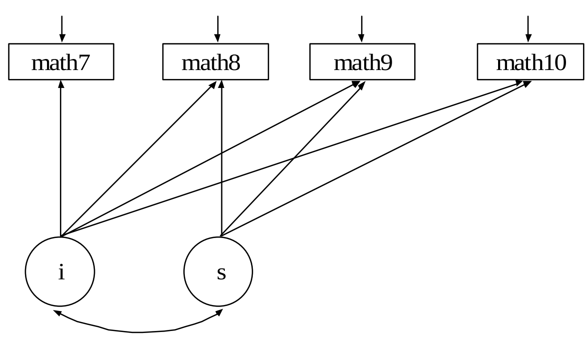
Covariances

	MATH7	MATH8	MATH9	MATH10
MATH7	81.107			
MATH8	67.663	82.829		
MATH9	73.150	76.513	100.986	
MATH10	77.952	82.668	95.158	131.326

Correlations

	MATH7	MATH8	MATH9	MATH10
MATH7	1.000			
MATH8	0.826	1.000		
MATH9	0.808	0.837	1.000	
MATH10	0.755	0.793	0.826	1.000

59



60

Input For LSAY Linear Growth Model Without Covariates

```
TITLE:     LSAY For Younger Females With Listwise Deletion
           Linear Growth Model Without Covariates
DATA:      FILE IS lsay.dat;
           FORMAT IS 3F8.0 F8.4 8F8.2 3F8.0;
VARIABLE:  NAMES ARE cohort id school weight math7 math8 math9
           math10 att7 att8 att9 att10 gender mothed homeres;
           USEOBS = (gender EQ 1 AND cohort EQ 2);
           MISSING = ALL (999);
           USEVAR = math7-math10;
ANALYSIS:  TYPE = MEANSTRUCTURE;
MODEL:     i BY math7-math10@1;
           s BY math7@0 math8@1 math9@2 math10@3;
           [math7-math10@0];
           [i s];
OUTPUT:    SAMPSTAT STANDARDIZED MODINDICES (3.84);
```

Alternative language:

```
MODEL: i s | math7@0 math8@1 math9@2 math10@3;
```

61

Output Excerpts LSAY Linear Growth Model Without Covariates

Tests Of Model Fit

```
Chi-Square Test of Model Fit
          Value                         22.664
          Degrees of Freedom                 5
          P-Value                       0.0004

CFI/TLI
          CFI                           0.995
          TLI                           0.994

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                      0.060
          90 Percent C.I.               0.036   0.086
          Probability RMSEA <= .05      0.223

SRMR (Standardized Root Mean Square Residual)
          Value                         0.025
```

62

Output Excerpts LSAY Linear Growth Model Without Covariates (Continued)

Modification Indices

	M.I.	E.P.C.	Std.E.P.C.	StdYX E.P.C.
S BY MATH7	6.793	0.185	0.254	0.029
S BY MATH8	14.694	-0.169	-0.233	-0.025
S BY MATH9	9.766	0.155	0.213	0.021

63

Output Excerpts LSAY Linear Growth Model Without Covariates (Continued)

Model Results

	Estimates	S.E.	Est./S.E.	Std	StdYX
I BY					
MATH7	1.000	.000	.000	8.029	.906
MATH8	1.000	.000	.000	8.029	.861
MATH9	1.000	.000	.000	8.029	.800
MATH10	1.000	.000	.000	8.029	.708
S BY					
MATH7	.000	.000	.000	.000	.000
MATH8	1.000	.000	.000	1.377	.148
MATH9	2.000	.000	.000	2.753	.274
MATH10	3.000	.000	.000	4.130	.364

64

Output Excerpts LSAY Linear Growth Model Without Covariates (Continued)

	Estimates	S.E.	Est./S.E.	Std	StdYX
Means					
I	52.623	.275	191.076	6.554	6.554
S	3.105	.075	41.210	2.255	2.255
Intercepts					
MATH7	.000	.000	.000	.000	.000
MATH8	.000	.000	.000	.000	.000
MATH9	.000	.000	.000	.000	.000
MATH10	.000	.000	.000	.000	.000

65

Output Excerpts LSAY Linear Growth Model Without Covariates (Continued)

	Estimates	S.E.	Est./S.E.	Std	StdYX
I WITH					
S	3.491	.730	4.780	.316	.316
Residual Variances					
MATH7	14.105	1.253	11.259	14.105	.180
MATH8	13.525	.866	15.610	13.525	.156
MATH9	14.726	.989	14.897	14.726	.146
MATH10	25.989	1.870	13.898	25.989	.202
Variances					
I	64.469	3.428	18.809	1.000	1.000
S	1.895	.322	5.894	1.000	1.000

R-Square

Observed Variable	R-Square
MATH7	0.820
MATH8	0.844
MATH9	0.854
MATH10	0.798

66

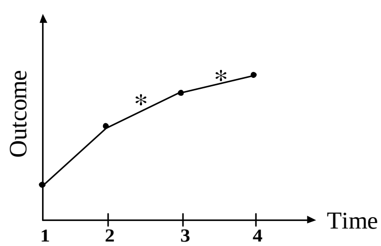
Growth Model With Free Time Scores

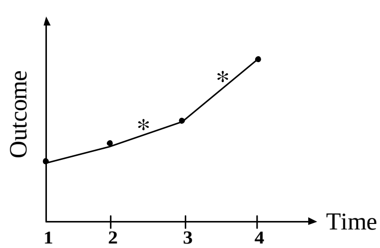
67

Specifying Time Scores For Non-Linear Growth Models With Estimated Time Scores

Non-linear growth models with estimated time scores

- Need two latent variables to describe a non-linear growth model: Intercept and slope





Time scores: 0 1 Estimated Estimated

68

Interpretation Of Slope Growth Factor Mean For Non-Linear Models

- The slope growth factor mean is the change in the outcome variable for a one unit change in the time score
- In non-linear growth models, the time scores should be chosen so that a one unit change occurs between timepoints of substantive interest.
  - An example of 4 timepoints representing grades 7, 8, 9, and 10
    - Time scores of 0 1 * * – slope factor mean refers to change between grades 7 and 8
    - Time scores of 0 * * 1 – slope factor mean refers to change between grades 7 and 10

Growth Model With Free Time Scores

- Identification of the model – for a model with two growth factors, at least one time score must be fixed to a non-zero value (usually one) in addition to the time score that is fixed at zero (centering point)
- Interpretation—cannot interpret the mean of the slope growth factor as a constant rate of change over all timepoints, but as the rate of change for a time score change of one.
- Approach—fix the time score following the centering point at one
- Choice of time score starting values if needed
  - Means 52.75 55.41 59.13 61.80
  - Differences 2.66 3.72 2.67
  - Time scores 0 1 >2 >2+1

Input Excerpts For LSAY Linear Growth Model With Free Time Scores Without Covariates

```
MODEL:     i s | math7@0 math8@1 math9 math10;

OUTPUT:    RESIDUAL;
```

Alternative language:

```
MODEL:     i BY math7-math10@1;
           s BY math7@0 math8@1 math9 math10;
           [math7-math10@0];
           [i s];
```

71

Output Excerpts LSAY Growth Model With Free Time Scores Without Covariates

n = 984

Tests Of Model Fit

```
Chi-Square Test of Model Fit
             Value                               4.222
             Degrees of Freedom                      3
             P-Value                            0.2373
CFI/TLI
             CFI                                 1.000
             TLI                                 0.999
RMSEA (Root Mean Square Error Of Approximation)
             Estimate                            0.020
             90 Percent C.I.                     0.000  0.061
             Probability RMSEA <= .05            0.864
SRMR (Standardized Root Mean Square Residual)
             Value                               0.015
```

72

Output Excerpts LSAY Growth Model With Free Time Scores Without Covariates (Continued)

Selected Estimates

	Estimates	S.E.	Est./S.E.	Std	StdYX
I |					
MATH7	1.000	.000	.000	8.029	.903
MATH8	1.000	.000	.000	8.029	.870
MATH9	1.000	.000	.000	8.029	.797
MATH10	1.000	.000	.000	8.029	.708
S |					
MATH7	.000	.000	.000	.000	.000
MATH8	1.000	.000	.000	1.134	.123
MATH9	**2.452**	**.133**	**18.442**	**2.780**	**.276**
MATH10	**3.497**	**.199**	**17.540**	**3.966**	**.350**

73

Output Excerpts LSAY Growth Model With Free Time Scores Without Covariates (Continued)

	Estimates	S.E.	Est./S.E.	Std	StdYX
S WITH					
I	3.110	.600	5.186	.342	.342
Variances					
I	64.470	3.394	18.994	1.000	1.000
S	**1.286**	**.265**	**4.853**	**1.000**	**1.000**
Means					
I	52.785	.283	186.605	6.574	6.574
S	**2.586**	**.167**	**15.486**	**2.280**	**2.280**

74

Output Excerpts LSAY Growth Model With Free Time Scores Without Covariates (Continued)

Residuals

Model Estimated Means/Intercepts/Thresholds

MATH7	MATH8	MATH9	MATH10
52.785	55.370	59.123	61.827

Residuals for Means/Intercepts/Thresholds

MATH7	MATH8	MATH9	MATH10
-.035	.041	.004	-.031

75

Output Excerpts LSAY Growth Model With Free Time Scores Without Covariates (Continued)

Model Estimated Covariances/Correlations/Residual Correlations

	MATH7	MATH8	MATH9	MATH10
MATH7	79.025			
MATH8	67.580	85.180		
MATH9	72.094	78.356	101.588	
MATH10	75.346	82.952	93.994	128.477

Residuals for Covariances/Correlations/Residual Correlations

	MATH7	MATH8	MATH9	MATH10
MATH7	1.999			
MATH8	.014	-2.436		
MATH9	.981	-1.921	-.705	
MATH10	2.527	-.368	1.067	2.715

76

Covariates In The Growth Model

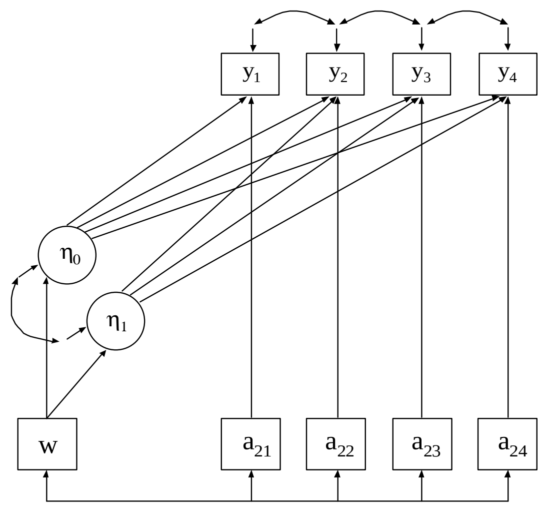
77

Covariates In The Growth Model

- Types of covariates
  - Time-invariant covariates—vary across individuals not time, explain the variation in the growth factors
  - Time-varying covariates—vary across individuals and time, explain the variation in the outcomes beyond the growth factors

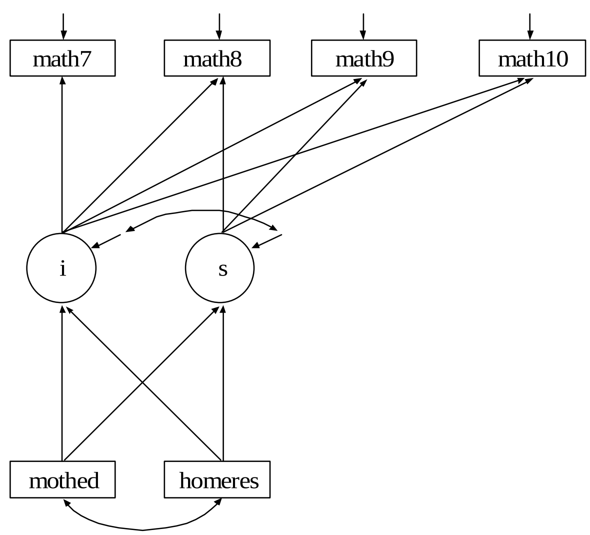
78

Time-Invariant And Time-Varying Covariates



79

LSAY Growth Model With Time-Invariant Covariates



80

Input Excerpts For LSAY Linear Growth Model With Free Time Scores And Covariates

```
VARIABLE:  NAMES ARE cohort id school weight math7 math8 math9
           math10 att7 att8 att9 att10 gender mothed homeres;
           USEOBS = (gender EQ 1 AND cohort EQ 2);
           MISSING = ALL (999);
           USEVAR = math7-math10 mothed homeres;
ANALYSIS:  !ESTIMATOR = MLM;
MODEL:     i s | math7@0 math8@1 math9 math10;
           i s ON mothed homeres;
```

Alternative language:

```
MODEL:     i BY math7-math10@1;
           s BY math7@0 math8@1 math9 math10;
           [math7-math10@0];
           [i s];
           i s ON mothed homeres;
```


Output Excerpts LSAY Growth Model With Free Time Scores And Covariates

n = 935

Tests Of Model Fit for ML

```
Chi-Square Test of Model Fit
             Value                              15.845
             Degrees of Freedom                      7
             P-Value                            0.0265
CFI/TLI
             CFI                                 0.998
             TLI                                 0.995
RMSEA (Root Mean Square Error Of Approximation)
             Estimate                            0.037
             90 Percent C.I.                     0.012  0.061
             Probability RMSEA <= .05            0.794
SRMR (Standardized Root Mean Square Residual)
             Value                               0.015
```

Output Excerpts LSAY Growth Model With Free Time Scores And Covariates (Continued)

Tests Of Model Fit for MLM

```
Chi-Square Test of Model Fit
             Value                                8.554*
             Degrees of Freedom                       7
             P-Value                             0.2862
             Scaling Correction Factor            1.852
                 for MLM
CFI/TLI
             CFI                                  0.999
             TLI                                  0.999
RMSEA (Root Mean Square Error Of Approximation)
             Estimate                             0.015
SRMR (Standardized Root Mean Square Residual)
             Value                                0.015
WRMR (Weighted Root Mean Square Residual)
             Value                                0.567
```

83

Output Excerpts LSAY Growth Model With Free Time Scores And Covariates (Continued)

Selected Estimates For ML

	Estimates	S.E.	Est./S.E.	Std	StdYX
I ON					
MOTHED	**2.054**	**.281**	**7.322**	**.257**	**.247**
HOMERES	**1.376**	**.182**	**7.546**	**.172**	**.255**
S ON					
MOTHED	.103	.068	1.524	.094	.090
HOMERES	**.149**	**.045**	**3.334**	**.136**	**.201**
I WITH					
S	2.604	.559	4.658	.297	.297
Residual Variances					
I	53.931	2.995	18.008	.842	.842
S	1.134	.253	4.488	.942	.942
Intercepts					
I	43.877	.790	55.531	5.484	5.484
S	1.859	.221	8.398	1.695	1.695

84

Output Excerpts LSAY Growth Model With Free Time Scores And Covariates (Continued)

R-Square

```
Observed
Variable      R-Square

MATH7         0.813
MATH8         0.849
MATH9         0.861
MATH10        0.796

Latent
Variable      R-Square

I             .158
S             .058
```

85

Model Estimated Average And Individual Growth Curves With Covariates

Model:

$$y_{ti} = \eta_{0i} + \eta_{1i} x_t + \varepsilon_{ti} , \qquad (23)$$

$$\eta_{0i} = \alpha_0 + \gamma_0 w_i + \zeta_{0i} , \qquad (24)$$

$$\eta_{1i} = \alpha_1 + \gamma_1 w_i + \zeta_{1i} , \qquad (25)$$

Estimated growth factor means:

$$\hat{E}(\eta_{0i}) = \hat{\alpha}_0 + \hat{\gamma}_0 \overline{w} , \qquad (26)$$

$$\hat{E}(\eta_{1i}) = \hat{\alpha}_1 + \hat{\gamma}_1 \overline{w} . \qquad (27)$$

Estimated outcome means:

$$\hat{E}(y_{ti}) = \hat{E}(\eta_{0i}) + \hat{E}(\eta_{1i}) x_t . \qquad (28)$$

Estimated outcomes for individual *i* :

$$\hat{y}_{ti} = \hat{\eta}_{0i} + \hat{\eta}_{1i} x_t \qquad (29)$$

where $\hat{\eta}_{0i}$ and $\hat{\eta}_{1i}$ are estimated factor scores. $\hat{y}_{ti}$ can be used for prediction purposes.

86

Model Estimated Means With Covariates

Model estimated means are available using the TECH4 and RESIDUAL options of the OUTPUT command.

Estimated Intercept Mean = Estimated Intercept +
Estimated Slope (Mothed)*
Sample Mean (Mothed) +
Estimated Slope (Homeres)*
Sample Mean (Homeres)

43.88 + 2.05*2.31 + 1.38*3.11 = 52.9

Estimated Slope Mean = Estimated Intercept +
Estimated Slope (Mothed)*
Sample Mean (Mothed) +
Estimated Slope (Homeres)*
Sample Mean (Homeres)

1.86 + .10*2.31 + .15*3.11 = 2.56

87

Model Estimated Means With Covariates (Continued)

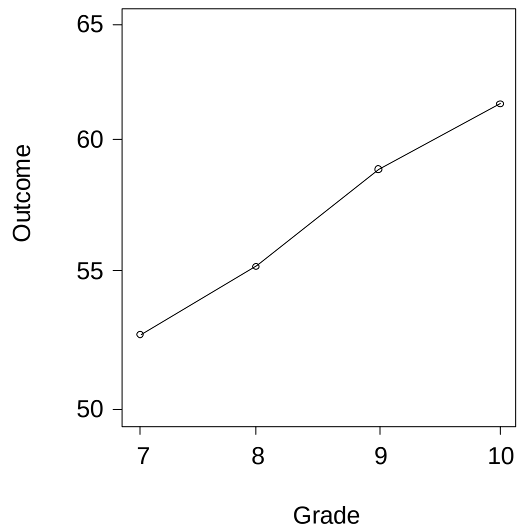
Estimated Outcome Mean at Timepoint t =

Estimated Intercept Mean +
Estimated Slope Mean * (Time Score at Timepoint t)

Estimated Outcome Mean at Timepoint 1 =
52.9 + 2.56 * (0) = **52.9**
Estimated Outcome Mean at Timepoint 2 =
52.9 + 2.56 * (1.00) = **55.46**
Estimated Outcome Mean at Timepoint 3 =
52.9 + 2.56 * (2.45) = **59.17**
Estimated Outcome Mean at Timepoint 4 =
52.9 + 2.56 * (3.50) = **61.86**

88

Estimated LSAY Curve



89

Centering

90

Centering

- Centering determines the interpretation of the intercept growth factor
- The centering point is the timepoint at which the time score is zero
- A model can be estimated for different centering points depending on which interpretation is of interest
- Models with different centering points give the same model fit because they are reparameterizations of the model
- Changing the centering point in a linear growth model with four timepoints

Timepoints	1	2	3	4	Centering at
Time scores	0	1	2	3	Timepoint 1
	-1	0	1	2	Timepoint 2
	-2	-1	0	1	Timepoint 3
	-3	-2	-1	0	Timepoint 4

91

Input Excerpts For LSAY Growth Model With Free Time Scores And Covariates Centered At Grade 10

```
MODEL:     i s | math7*-3 math8*-2 math9@-1 math10@0;
           i s ON mothed homeres;
```

Alternative language:

```
MODEL:     i BY math7-math10@1;
           s BY math7*-3 math8*-2 math9@-1 math10@0;
           [math7-math10@0];
           [i s];
           i s ON mothed homeres;
```

92

Output Excerpts LSAY Growth Model With Free Time Scores And Covariates Centered At Grade 10

n = 935

Tests of Model Fit

```
CHI-SQUARE TEST OF MODEL FIT

         Value                          15.845
         Degrees of Freedom                  7
         P-Value                         .0265

RMSEA (ROOT MEAN SQUARE ERROR OF APPROXIMATION)

         Estimate                        .037
         90 Percent C.I.                 .012    .061
         Probability RMSEA <= .05        .794
```

93

Output Excerpts LSAY Growth Model With Free Time Scores And Covariates Centered At Grade 10 (Continued)

Selected Estimates

	Estimates	S.E.	Est./S.E.	Std	StdYX
I ON					
MOTHED	2.418	0.353	6.851	0.238	0.229
HOMERES	1.903	0.229	8.294	0.187	0.277
S ON					
MOTHED	0.111	0.073	1.521	0.094	0.090
HOMERES	0.161	0.049	3.311	0.136	0.201

94

Further Readings On Introductory Growth Modeling

Bijleveld, C. C. J. H., & van der Kamp, T. (1998). Longitudinal data analysis: Designs, models, and methods. Newbury Park: Sage.

Bollen, K.A. & Curran, P.J. (2006). Latent curve models. A structural equation perspective. New York: Wiley.

Muthén, B. & Khoo, S.T. (1998). Longitudinal studies of achievement growth using latent variable modeling. Learning and Individual Differences, Special issue: latent growth curve analysis, 10, 73-101. (#80)

Muthén, B. & Muthén, L. (2000). The development of heavy drinking and alcohol-related problems from ages 18 to 37 in a U.S. national sample. Journal of Studies on Alcohol, 61, 290-300. (#83)

Raudenbush, S.W. & Bryk, A.S. (2002). Hierarchical linear models: Applications and data analysis methods. Second edition. Newbury Park, CA: Sage Publications.

Singer, J.D. & Willett, J.B. (2003). Applied longitudinal data analysis. Modeling change and event occurrence. New York, NY: Oxford University Press.

Snijders, T. & Bosker, R. (1999). Multilevel analysis. An introduction to basic and advanced multilevel modeling. Thousand Oakes, CA: Sage Publications.

95

Further Practical Issues

96

Five Ways To Model Non-Linear Growth

- Estimated time scores
- Quadratic (cubic) growth model
- Fixed non-linear time scores
- Piece-wise growth modeling
- Time-varying covariates

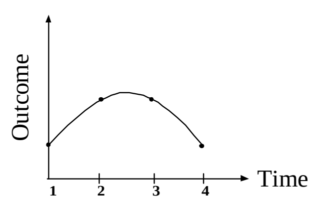
97

Specifying Time Scores For Quadratic Growth Models

Quadratic growth model

$$y_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \eta_{2i}\; x_t^2 + \varepsilon_{ti}$$

- Need three latent variables to describe a quadratic growth model: Intercept, linear slope, quadratic slope



- Linear slope time scores: 0 1 2 3
  0 .1 .2 .3
- Quadratic slope time scores: 0 1 4 9
  0 .01 .04 .09

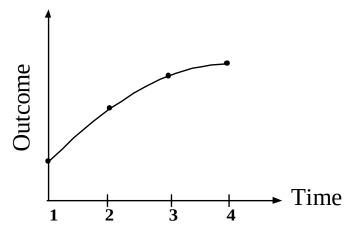
98

Specifying Time Scores For Non-Linear Growth Models With Fixed Time Scores

Non-Linear Growth Models with Fixed Time scores

- Need two latent variables to describe a non-linear growth model: Intercept and slope

Growth model with a logarithmic growth curve--ln(t)

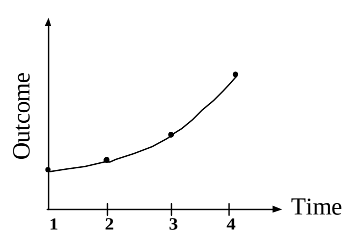


Time scores: 0 0.69 1.10 1.39

99

Specifying Time Scores For Non-Linear Growth Models With Fixed Time Scores (Continued)

Growth model with an exponential growth curve–
exp(t-1) - 1



Time scores: 0 1.72 6.39 19.09

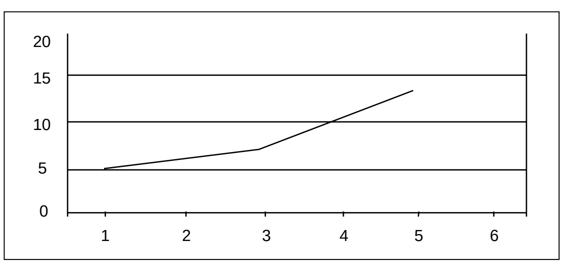
100

Piecewise Growth Modeling

101

Piecewise Growth Modeling

- Can be used to represent different phases of development
- Can be used to capture non-linear growth
- Each piece has its own growth factor(s)
- Each piece can have its own coefficients for covariates

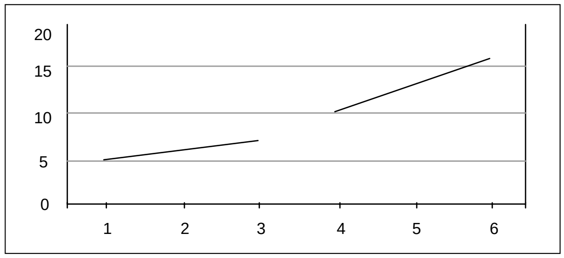


One intercept growth factor, two slope growth factors

0	1	2	2	2	2	Time scores piece 1
0	0	0	1	2	3	Time scores piece 2

102

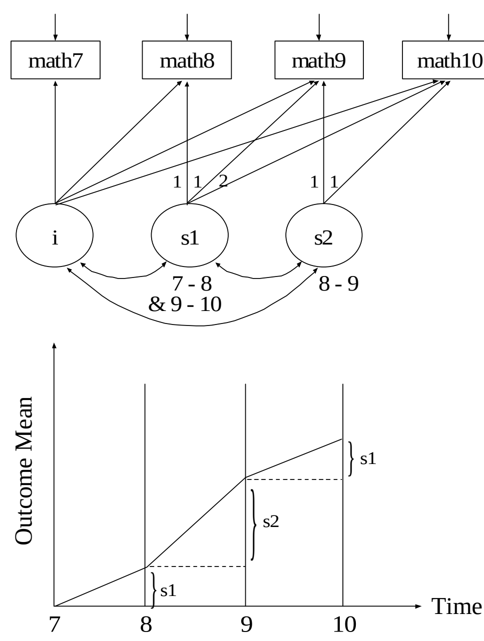
Piecewise Growth Modeling (Continued)



Two intercept growth factors, two slope growth factors

0 1 2 Time scores piece 1

0 1 2 Time scores piece 2



Input For LSAY Piecewise Growth Model With Covariates

```
MODEL:     i s1 | math7@0 math8@1 math9@1 math10@2;
           i s2 | math7@0 math8@0 math9@1 math10@1;
           i s1 s2 ON mothed homeres;


Alternative language:


MODEL:     i BY math7-math10@1;
           s1 BY math7@0 math8@1 math9@1 math10@2;
           s2 BY math7@0 math8@0 math9@1 math10@1;
           [math7-math10@0];
           [i s1 s2];
           i s1 s2 ON mothed homeres;
```


Output Excerpts LSAY Piecewise Growth Model With Covariates

n = 935

Tests of Model Fit

```
CHI-SQUARE TEST OF MODEL FIT

          Value                        11.721
          Degrees of Freedom                3
          P-Value                       .0083

RMSEA (ROOT MEAN SQUARE ERROR OF APPROXIMATION)

          Estimate                       .056
          90 Percent C.I.                .025   .091
          Probability RMSEA <= .05       .331
```

Output Excerpts LSAY Piecewise Growth Model With Covariates (Continued)

Selected Estimates

		Estimates	S.E.	Est./S.E.	Std	StdYX
I	ON					
	MOTHED	2.127	.284	7.488	.266	.256
	HOMERES	1.389	.185	7.524	.174	.257
S1	ON					
	MOTHED	-.126	.147	-.858	-.113	-.109
	HOMERES	.091	.096	.950	.081	.120
S2	ON					
	MOTHED	**.436**	**.191**	**2.285**	**.185**	**.178**
	HOMERES	**.289**	**.124**	**2.329**	**.123**	**.181**

Growth Model With Individually-Varying Times Of Observation And Random Slopes For Time-Varying Covariates

Growth Modeling In Multilevel Terms

Time point *t*, individual *i* (two-level modeling, no clustering):

y_{ti} : repeated measures of the outcome, e.g. math achievement
a_{1ti} : time-related variable; e.g. grade 7-10
a_{2ti} : time-varying covariate, e.g. math course taking
x_i : time-invariant covariate, e.g. grade 7 expectations

Two-level analysis with individually-varying times of observation and random slopes for time-varying covariates:

$$\textit{Level } 1: y_{ti} = \pi_{0i} + \pi_{1i}\, a_{1ti} + \pi_{2ti}\, a_{2ti} + e_{ti}, \qquad (55)$$

$$\textit{Level } 2: \begin{cases} \pi_{0i} = ß_{00} + ß_{01}\, x_i + r_{0i}, \\ \pi_{1i} = ß_{10} + ß_{11}\, x_i + r_{1i}, \\ \pi_{2i} = ß_{20} + ß_{21}\, x_i + r_{2i}. \end{cases} \qquad (56)$$

109

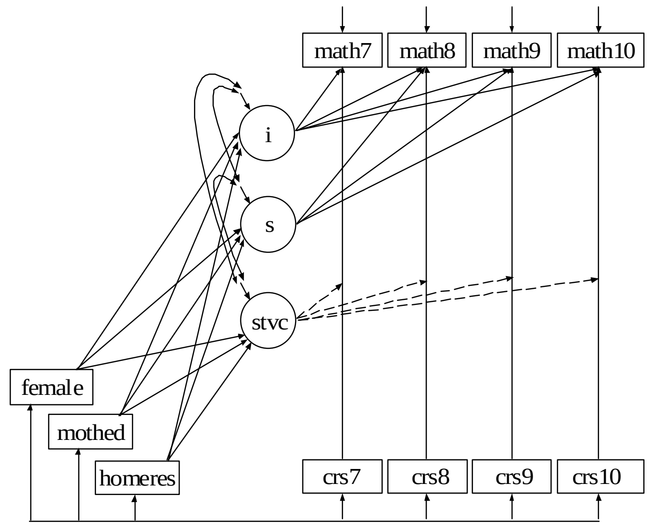
Growth Modeling In Multilevel Terms (Continued)

Time scores a_{1ti} read in as data (not loading parameters).

- π_{2ti} possible with time-varying random slope variances
- Flexible correlation structure for $V(e) = \Theta\ (T \text{ x } T)$
- Regressions among random coefficients possible, e.g.

$$\pi_{1i} = ß_{10} + \gamma_1\, \pi_{0i} + ß_{11}\, x_i + r_{1i}, \qquad (57)$$
$$\pi_{2i} = ß_{20} + \gamma_2\, \pi_{0i} + ß_{21}\, x_i + r_{2i}. \qquad (58)$$

110



111

Input For Growth Model With Individually Varying Times Of Observation

```
TITLE:     Growth model with individually varying times of
           observation and random slopes
DATA:      FILE IS lsaynew.dat;
           FORMAT IS 3F8.0 F8.4 8F8.2 3F8.0;
VARIABLE:  NAMES ARE math7 math8 math9 math10 crs7 crs8 crs9
           crs10 female mothed homeres a7-a10;

           ! crs7-crs10 = highest math course taken during each
           ! grade (0=no course, 1=low, basic, 2=average, 3=high.
           ! 4=pre-algebra, 5=algebra I, 6=geometry,
           ! 7=algebra II, 8=pre-calc, 9=calculus)

           MISSING ARE ALL (9999);
           CENTER = GRANDMEAN (crs7-crs10 mothed homeres);
           TSCORES = a7-a10;
```

112

Input For Growth Model With Individually Varying Times Of Observation (Continued)

```
DEFINE:     math7 = math7/10;
            math8 = math8/10;
            math9 = math9/10;
            math10 = math10/10;
ANALYSIS:   TYPE = RANDOM MISSING;
            ESTIMATOR = ML;
            MCONVERGENCE = .001;
MODEL:      i s | math7-math10 AT a7-a10;
            stvc | math7 ON crs7;
            stvc | math8 ON crs8;
            stvc | math9 ON crs9;
            stvc | math10 ON crs10;
            i ON female mothed homeres;
            s ON female mothed homeres;
            stvc ON female mothed homeres;
            i WITH s;
            stvc WITH i;
            stvc WITH s;
OUTPUT:     TECH8;
```

113

Output Excerpts For Growth Model With Individually Varying Times Of Observation And Random Slopes For Time-Varying Covariates

n = 2271

Tests of Model Fit

```
Loglikelihood

          H0 Value                      -8199.311

Information Criteria

          Number of Free Parameters            22
          Akaike (AIC)                  16442.623
          Bayesian (BIC)                16568.638
          Sample-Size Adjusted BIC      16498.740
             (n* = (n + 2) / 24)
```

114

Output Excerpts For Growth Model With Individually Varying Times Of Observation And Random Slopes For Time-Varying Covariates (Continued)

Model Results

	Estimates	S.E.	Est./S.E.
I ON			
FEMALE	0.187	0.036	5.247
MOTHED	0.187	0.018	10.231
HOMERES	0.159	0.011	14.194
S ON			
FEMALE	-0.025	0.012	-2.017
MOTHED	0.015	0.006	2.429
HOMERES	0.019	0.004	4.835
STVC ON			
FEMALE	-0.008	0.013	-0.590
MOTHED	0.003	0.007	0.429
HOMERES	0.009	0.004	2.167
I WITH			
S	0.038	0.006	6.445
STVC WITH			
I	0.011	0.005	2.087
S	0.004	0.002	2.033

115

Output Excerpts For Growth Model With Individually Varying Times Of Observation And Random Slopes For Time-Varying Covariates (Continued)

Intercepts			
MATH7	0.000	0.000	0.000
MATH8	0.000	0.000	0.000
MATH9	0.000	0.000	0.000
MATH10	0.000	0.000	0.000
I	4.992	0.025	198.456
S	0.417	0.009	47.275
STVC	0.113	0.010	11.416
Residual Variances			
MATH7	0.185	0.011	16.464
MATH8	0.178	0.008	22.232
MATH9	0.156	0.008	18.497
MATH10	0.169	0.014	12.500
I	0.570	0.023	25.087
S	0.036	0.003	12.064
STVC	0.012	0.002	5.055

116

Random Slopes

- In single-level modeling random slopes $ß_i$ describe variation across individuals i,

$$y_i = \alpha_i + ß_i\, x_i + \varepsilon_i\,, \tag{100}$$
$$\alpha_i = \alpha + \zeta_{0i}\,, \tag{101}$$
$$ß_i = ß + \zeta_{1i}\,, \tag{102}$$

Resulting in heteroscedastic residual variances

$$V\,(y_i \mid x_i) = V\,(\,ß_i\,)\;\; x_i^2 \;\; + \theta \tag{103}$$

- In two-level modeling random slopes $ß_j$ describe variation across clusters j

$$y_{ij} = a_j + ß_j\; x_{ij} + \varepsilon_{ij}\,, \tag{104}$$
$$a_j = a + \zeta_{0j}\,, \tag{105}$$
$$ß_j = ß + \zeta_{1j}\,. \tag{106}$$

A small variance for a random slope typically leads to slow convergence of the ML-EM iterations. This suggests respecifying the slope as fixed.

Mplus allows random slopes for predictors that are

- Observed covariates
- Observed dependent variables
- Continuous latent variables

117

Computational Issues For Growth Models

- Decreasing variances of the observed variables over time may make the modeling more difficult
- Scale of observed variables – keep on a similar scale
- Convergence – often related to starting values or the type of model being estimated
  - Program stops because maximum number of iterations has been reached
    - If no negative residual variances, either increase the number of iterations or use the preliminary parameter estimates as starting values
    - If there are large negative residual variances, try better starting values
  - Program stops before the maximum number of iterations has been reached
    - Check if variables are on a similar scale
    - Try new starting values
- Starting values – the most important parameters to give starting values to are residual variances and the intercept growth factor mean
- Convergence for models using the | symbol
  - Non-convergence may be caused by zero random slope variances which indicates that the slopes should be fixed rather than random

118

Advanced Growth Models

119

Advantages Of Growth Modeling In A Latent Variable Framework

- Flexible curve shape
- Individually-varying times of observation
- **Regressions among random effects**
- Multiple processes
- Modeling of zeroes
- Multiple populations
- Multiple indicators
- Embedded growth models
- Categorical latent variables: growth mixtures

120

Regressions Among Random Effects

121

Regressions Among Random Effects

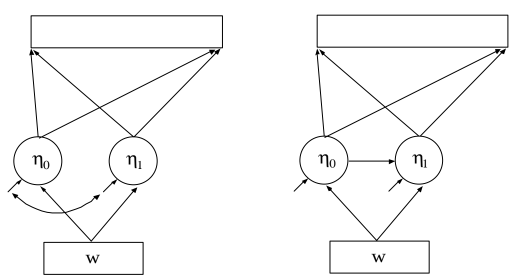
Standard multilevel model (where $x_t = 0, 1, \ldots, T$):

$$\text{Level 1: } y_{ti} = \eta_{0i} + \eta_{1i} x_t + \varepsilon_{ti}, \quad (1)$$
$$\text{Level 2a: } \eta_{0i} = \alpha_0 + \gamma_0 w_i + \zeta_{0i}, \quad (2)$$
$$\text{Level 2b: } \eta_{1i} = \alpha_1 + \gamma_1 w_i + \zeta_{1i}. \quad (3)$$

A useful type of model extension is to replace (3) by the regression equation

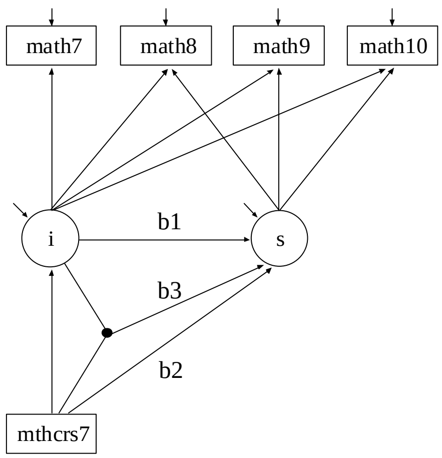
$$\eta_{1i} = \alpha + \beta \eta_{0i} + \gamma w_i + \zeta_i. \quad (4)$$

Example: Blood Pressure (Bloomqvist, 1977)



122

Growth Model With An Interaction



Input For A Growth Model With An Interaction Between A Latent And An Observed Variable

```
TITLE:     growth model with an interaction between a latent and an
           observed variable
DATA:      FILE IS lsay.dat;
VARIABLE:  NAMES ARE math7 math8 math9 math10 mthcrs7;
           MISSING ARE ALL (9999);
           CENTERING = GRANDMEAN (mthcrs7);
DEFINE:    math7 = math7/10;
           math8 = math8/10;
           math9 = math9/10;
           math10 = math10/10;
ANALYSIS:  TYPE=RANDOM MISSING;
MODEL:     i s | math7@0 math8@1 math9@2 math10@3;
           [math7-math10] (1);    !growth language defaults
           [i@0 s];               !overridden

           inter | i XWITH mthcrs7;
           s ON i mthcrs7 inter;
           i ON mthcrs7;
OUTPUT:    SAMPSTAT STANDARDIZED TECH1 TECH8;
```

Output Excerpts Growth Model With An Interaction Between A Latent And An Observed Variable

Tests Of Model Fit

```
Loglikelihood

          H0 Value                            -10068.944
Information Criteria

          Number of Free Parameters                   12
          Akaike (AIC)                         20161.887
          Bayesian (BIC)                       20234.365
          Sample-Size Adjusted BIC             20196.236
            (n* = (n + 2) / 24)
```

125

Output Excerpts Growth Model With An Interaction Between A Latent And An Observed Variable (Continued)

Model Results

```
                  Estimates      S.E.  Est./S.E.
 I        |
    MATH7            1.000      0.000     0.000
    MATH8            1.000      0.000     0.000
    MATH9            1.000      0.000     0.000
    MATH10           1.000      0.000     0.000
 S        |
    MATH7            0.000      0.000     0.000
    MATH8            1.000      0.000     0.000
    MATH9            2.000      0.000     0.000
    MATH10           3.000      0.000     0.000
```

126

Output Excerpts Growth Model With An Interaction Between A Latent And An Observed Variable (Continued)

	Estimates	S.E.	Est./S.E.
S ON			
I	0.087	0.012	7.023
INTER	-0.047	0.006	-7.301
S ON			
MTHCRS7	0.045	0.013	3.555
I ON			
MTHCRS7	0.632	0.016	40.412

127

Output Excerpts Growth Model With An Interaction Between A Latent And An Observed Variable (Continued)

	Estimates	S.E.	Est./S.E.
Intercepts			
MATH7	5.019	0.015	341.587
MATH8	5.019	0.015	341.587
MATH9	5.019	0.015	341.587
MATH10	5.019	0.015	341.587
I	0.000	0.000	0.000
S	0.417	0.007	57.749
Residual Variances			
MATH7	0.184	0.011	16.117
MATH8	0.178	0.009	20.109
MATH9	0.164	0.009	18.369
MATH10	0.173	0.015	11.509
I	0.528	0.018	28.935
S	0.037	0.004	10.027

128

Interpreting The Effect Of The Interaction Between Initial Status Of Growth In Math Achievement And Course Taking In Grade 6

- Model equation for slope s
  s = a + b1*i + b2*mthcrs7 + b3*i*mthcrs7 + e

  or, using a moderator function (Klein & Moosbrugger, 2000) where i moderates the influence of mthcrs7 on s
  s = a + b1*i + (b2 + b3*i)*mthcrs7 + e

- Estimated model

  Unstandardized
  s = 0.417 + 0.087*i + (0.045 – 0.047*i)*mthcrs7

  Standardized with respect to i and mthcrs7
  s = 0.42 + 0.08 * i + (0.04-0.04*i)*mthcrs7

129

Interpreting The Effect Of The Interaction Between Initial Status Of Growth In Math Achievement And Course Taking In Grade 6 (Continued)

- Interpretation of the standardized solution
  At the mean of i, which is zero, the slope increases 0.04 for 1 SD increase in mthcrs7

  At 1 SD below the mean of i, which is zero, the slope increases 0.08 for 1 SD increase in mthcrs7

  At 1 SD above the mean of i, which is zero, the slope does not increase as a function of mthcrs7

130

Growth Modeling With Parallel Processes

131

Advantages Of Growth Modeling In A Latent Variable Framework

- Flexible curve shape
- Individually-varying times of observation
- Regressions among random effects
- **Multiple processes**
- Modeling of zeroes
- Multiple populations
- Multiple indicators
- Embedded growth models
- Categorical latent variables: growth mixtures

132

Multiple Processes

- Parallel processes
- Sequential processes

133

Growth Modeling With Parallel Processes

- Estimate a growth model for each process separately
  - Determine the shape of the growth curve
  - Fit model without covariates
  - Modify the model
- Joint analysis of both processes
- Add covariates

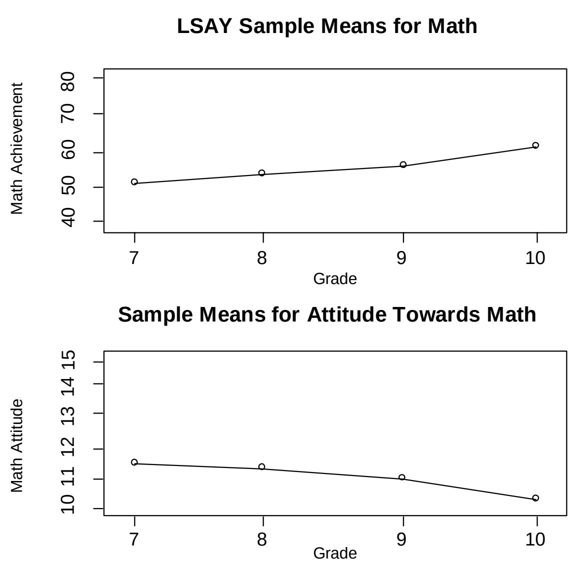
134

LSAY Data

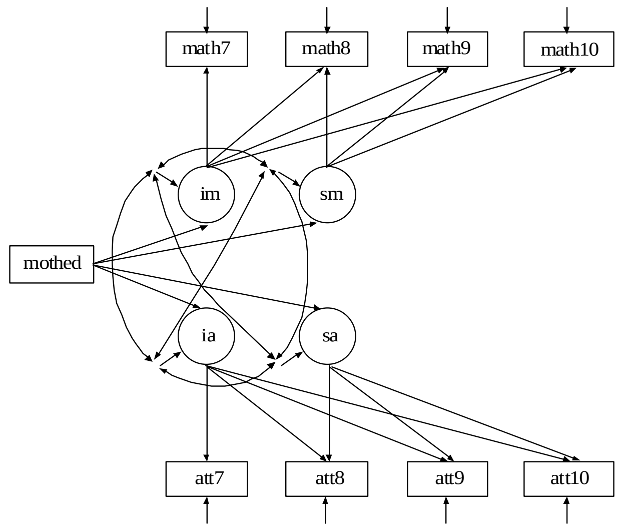
The data come from the Longitudinal Study of American Youth (LSAY). Two cohorts were measured at four time points beginning in 1987. Cohort 1 was measured in Grades 10, 11, and 12. Cohort 2 was measured in Grades 7, 8, 9, and 10. Each cohort contains approximately 60 schools with approximately 60 students per school. The variables measured include math and science achievement items, math and science attitude measures, and background information from parents, teachers, and school principals. There are approximately 60 items per test with partial item overlap across grades—adaptive tests.

Data for the analysis include the younger females. The variables include math achievement and math attitudes from Grades 7, 8, 9, and 10 and mother's education.

135



136



137

Input For LSAY Parallel Process Growth Model

```
TITLE:     LSAY For Younger Females With Listwise Deletion
           Parallel Process Growth Model-Math Achievement and
           Math Attitudes

DATA:      FILE IS lsay.dat;
           FORMAT IS 3f8 f8.4 8f8.2 3f8 2f8.2;

VARIABLE:  NAMES ARE cohort id school weight math7 math8 math9
           math10 att7 att8 att9 att10 gender mothed homeres
           ses3 sesq3;
           USEOBS = (gender EQ 1 AND cohort EQ 2);
           MISSING = ALL (999);
           USEVAR = math7-math10 att7-att10 mothed;

ANALYSIS:  TYPE = MEANSTRUCTURE;
```

138

Input For LSAY Parallel Process Growth Model

```
MODEL:     im sm | math7@0 math8@1 math9 math10;
           ia sa | att7@0 att8@1 att9@2 att10@3;
           im-sa ON mothed;

OUTPUT:    MODINDICES STANDARDIZED;
```

Alternative language:

```
           im BY math7-math10@1;
           sm BY math7@0 math8@1 math9 math10;

           ia BY att7-att10@1;
           sa BY att7@0 att8@1 att9@2 att10@3;

           [math7-math10@0 att7-att10@0];
           [im sm ia sa];

           im-sa ON mothed;
```

139

Output Excerpts LSAY Parallel Process Growth Model

n = 910

Tests of Model Fit

```
Chi-Square Test of Model Fit

          Value                         43.161
          Degrees of Freedom                24
          P-Value                        .0095

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                        .030
          90 Percent C.I.                 .015   .044
          Probability RMSEA <= .05        .992
```

140

Output Excerpts LSAY Parallel Process Growth Model (Continued)

		Estimates	S.E.	Est./S.E.	Std	StdYX
IM	**ON**					
	MOTHED	**2.462**	**.280**	**8.798**	**.311**	**.303**
SM	**ON**					
	MOTHED	**.145**	**.066**	**2.195**	**.132**	**.129**
IA	**ON**					
	MOTHED	**.053**	**.086**	**.614**	**.025**	**.024**
SA	**ON**					
	MOTHED	**.012**	**.035**	**.346**	**.017**	**.017**

141

Output Excerpts LSAY Parallel Process Growth Model (Continued)

		Estimates	S.E.	Est./S.E.	Std	StdYX
SM	WITH					
	IM	3.032	.580	5.224	.350	.350
IA	WITH					
	IM	4.733	.702	6.738	.282	.282
	SM	.544	.164	3.312	.235	.235
SA	WITH					
	IM	-.276	.279	-.987	-.049	-.049
	SM	.130	.066	1.976	.168	.168
	IA	-.567	.115	-4.913	-.378	-.378

142

Modeling With Zeroes

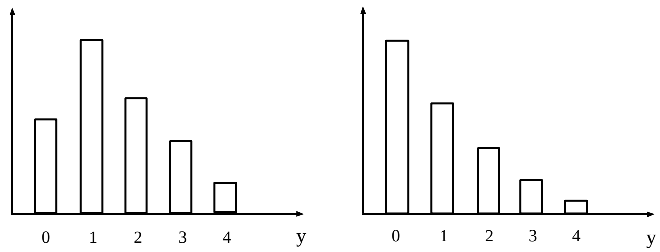
143

Advantages Of Growth Modeling In A Latent Variable Framework

- Flexible curve shape
- Individually-varying times of observation
- Regressions among random effects
- Multiple processes
- **Modeling of zeroes**
- Multiple populations
- Multiple indicators
- Embedded growth models
- Categorical latent variables: growth mixtures

144

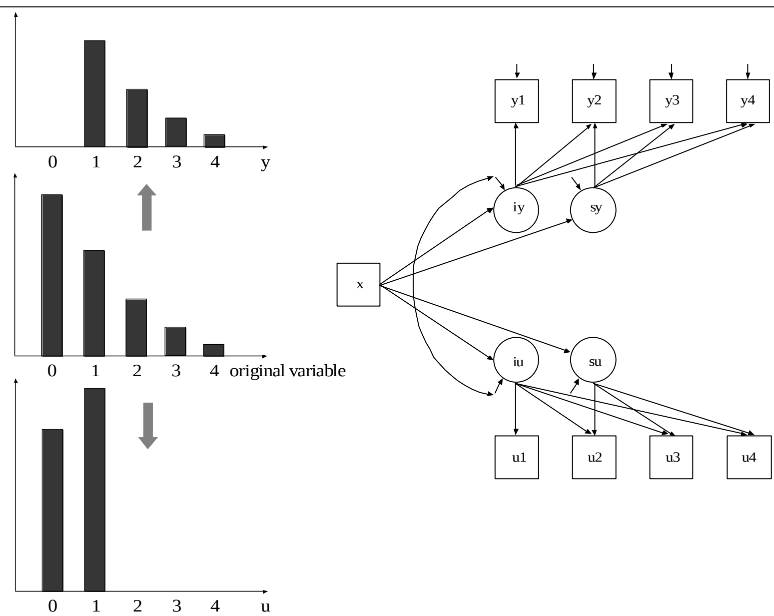
Modeling With A Preponderance Of Zeros



- Outcomes: non-normal continuous – count – categorical
- Censored-normal modeling
- Two-part (semicontinuous modeling): Duan et al. (1983), Olsen & Schafer (2001)
- Mixture models, e.g. zero-inflated (mixture) Poisson (Roeder et al., 1999), censored-inflated, mover-stayer latent transition models, growth mixture models
- Onset (survival) followed by growth: Albert & Shih (2003)

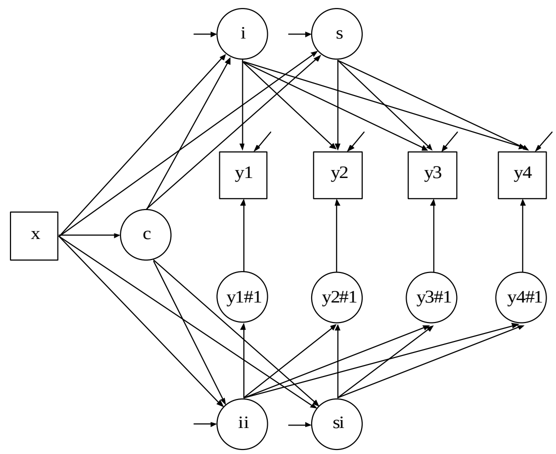
145

Two-Part (Semicontinuous) Growth Modeling

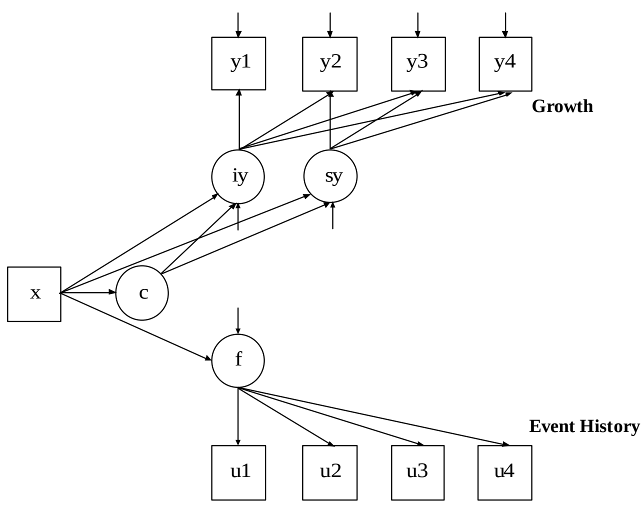


146

Inflated Growth Modeling (Two Classes At Each Time Point)

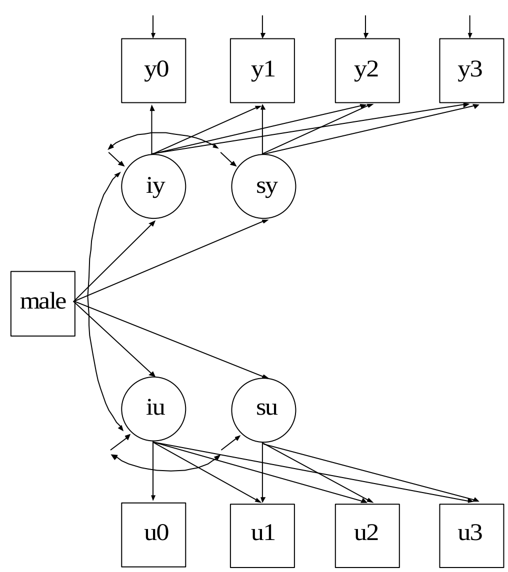


Onset (Survival) Followed By Growth



Two-Part Growth Modeling

149



150

Input For Step 1 Of A Two-Part Growth Model

```
TITLE:     step 1 of a two-part growth model
              Amover u    y
               >0    1   >0
                0    0   999
               999  999  999
DATA:      FILE = amp.dat;
VARIABLE:  NAMES ARE caseid
           amover0 ovrdrnk0 illdrnk0 vrydrn0
           amover1 ovrdrnk1 illdrnk1 vrydrn1
           amover2 ovrdrnk2 illdrnk2 vrydrn2
           amover3 ovrdrnk3 illdrnk3 vrydrn3
           amover4 ovrdrnk4 illdrnk4 vrydrn4
           amover5 ovrdrnk5 illdrnk5 vrydrn5
           amover6 ovrdrnk6 illdrnk6 vrydrn6
           tfq0-tfq6 v2 sex race livewith
           agedrnk0-agedrnk6 grades0-grades6;
           USEV = amover0 amover1 amover2 amover3
           sex race u0-u3 y0-y3;
         ! MISSING = ALL (999);
```

151

Input For Step 1 Of A Two-Part Growth Model (Continued)

```
DEFINE:       u0 = 1;                                  !binary part of variable
              IF(amover0 eq 0) THEN u0 = 0;
              IF(amover0 eq 999) THEN u0 = 999;
              y0 = amover0;                            !continuous part of variable
              IF (amover0 eq 0) THEN y0 = 999;
              u1 = 1;
              IF(amover1 eq 0) THEN u1 = 0;
              IF(amover1 eq 999) THEN u1 = 999:
              y1 = amover1;
              IF(amover1 eq 0) THEN y1 = 999;
              u2 = 1;
              IF(amover2 eq 0) THEN u2 = 0;
              IF(amover2 eq 999) THEN u2 = 999;
              y2 = amover2;
              IF(amover2 eq 0) THEN y2 = 999;
              u3 = 1;
              IF(amover3 eq 0) THEN u3 = 0;
              IF(amover3 eq 999) THEN u3 = 999;
              y3 = amover3;
              IF(amover3 eq 0) THEN y3 = 999;
ANALYSIS:     TYPE = BASIC;
SAVEDATA:     FILE = ampyu.dat;
```

152

Output Excerpts Step 1 Of A Two-Part Growth Model

SAVEDATA Information

```
Order and format of variables
    AMOVER0  F10.3
    AMOVER1  F10.3
    AMOVER2  F10.3
    AMOVER3  F10.3
    SEX      F10.3
    RACE     F10.3
    U0       F10.3
    U1       F10.3
    U2       F10.3
    U3       F10.3
    Y0       F10.3
    Y1       F10.3
    Y2       F10.3
    Y3       F10.3
Save file
     ampyu.dat
Save file format
     14F10.3
Save file record length    1000
```

153

Input For Step 2 Of A Two-Part Growth Model

```
TITLE:      two-part growth model with linear growth for both
            parts
DATA:       FILE = ampyau.dat;
VARIABLE:   NAMES = amover0-amover3 sex race u0-u3 y0-y3;
            USEV = u0-u3 y0-y3 male;
            USEOBS = u0 NE 999;
            MISSING = ALL (999);
            CATEGORICAL = u0-u3;
DEFINE:     male = 2-sex;
```

154

Input For Step 2 Of A Two-Part Growth Model (Continued)

```
ANALYSIS:   TYPE = MISSING;
            ESTIMATOR = ML;
            ALGORITHM = INTEGRATION;
            COVERAGE = .09;
MODEL:      iu su | u0@0 u1@0.5 u2@1.5 u3@2.5;
            iy sy | y0@0 y1@0.5 y2@1.5 y3@2.5;
            iu-sy ON male;
            ! estimate the residual covariances
            ! iu with su, iy with sy, and iu with iy
            iu WITH sy@0;
            su WITH iy-sy@0;
OUTPUT:     PATTERNS SAMPSTAT STANDARDIZED TECH1 TECH4 TECH8;
PLOT:       TYPE = PLOT3;
            SERIES = u0-u3(su) | y0-y3(sy);
```

155

Output Excerpts Step 2 Of A Two-Part Growth Model

Tests of Model Fit

```
Loglikelihood

          H0 Value                    -3277.101

Information Criteria

          Number of Free parameters          19
          Akaike (AIC)                 6592.202
          Bayesian (BIC)               6689.444
          Sample-Size Adjusted BIC     6629.092
             (n* = (n + 2) / 24)
```

156

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

Model Results

		Estimates	S.E.	Est./S.E.	Std	StdYX
IU |						
	U0	1.000	0.000	0.000	2.839	0.843
	U1	1.000	0.000	0.000	2.839	0.882
	U2	1.000	0.000	0.000	2.839	0.926
	U3	1.000	0.000	0.000	2.839	0.905
SU |						
	U0	0.000	0.000	0.000	0.000	0.000
	U1	0.500	0.000	0.000	0.416	0.129
	U2	1.500	0.000	0.000	1.249	0.407
	U3	2.500	0.000	0.000	2.082	0.664

157

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

IY |						
	Y0	1.000	0.000	0.000	0.534	0.787
	Y1	1.000	0.000	0.000	0.534	0.738
	Y2	1.000	0.000	0.000	0.534	0.740
	Y3	1.000	0.000	0.000	0.534	0.644
SY |						
	Y0	0.000	0.000	0.000	0.000	0.000
	Y1	0.500	0.000	0.000	0.117	0.162
	Y2	1.500	0.000	0.000	0.351	0.487
	Y3	2.500	0.000	0.000	0.586	0.707

158

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

```
 IU      ON
    MALE           0.569   0.234      2.433    0.200    0.100
 SU      ON
    MALE          -0.181   0.119     -1.518   -0.218   -0.109
 IY      ON
    MALE           0.149   0.061      2.456    0.279    0.139
 SY      ON
    MALE          -0.068   0.038     -1.790   -0.290   -0.145

 IU      WITH
    SU            -1.144   0.326     -3.509   -0.484   -0.484
    IY             1.193   0.134      8.897    0.788    0.788
    SY             0.000   0.000      0.000    0.000    0.000
 IY      WITH
    SY            -0.039   0.019     -2.109   -0.316   -0.316
 SU      WITH
    IY             0.000   0.000      0.000    0.000    0.000
    SY             0.000   0.000      0.000    0.000    0.000
```

159

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

```
 Intercepts
    Y0             0.000   0.000    0.000    0.000     0.000
    Y1             0.000   0.000    0.000    0.000     0.000
    Y2             0.000   0.000    0.000    0.000     0.000
    Y3             0.000   0.000    0.000    0.000     0.000
    IU             0.000   0.000    0.000    0.000     0.000
    SU             0.855   0.098    8.716    1.027     1.027
    IY             0.232   0.059    3.901    0.435     0.435
    SY             0.240   0.031    7.830    1.025     1.025

 Thresholds
    U0$1           2.655   0.206   12.877
    U1$1           2.655   0.206   12.877
    U2$1           2.655   0.206   12.877
    U3$1           2.655   0.206   12.877
```

160

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

Residual Variances

Y0	0.175	0.032	5.470	0.175	0.380
Y1	0.266	0.029	9.159	0.266	0.509
Y2	0.238	0.027	8.810	0.238	0.457
Y3	0.269	0.054	5.014	0.269	0.392
IU	7.982	1.086	7.351	0.990	0.990
SU	0.685	0.202	3.400	0.988	0.988
IY	0.279	0.040	7.019	0.981	0.981
SY	0.054	0.017	3.224	0.979	0.979

161

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

Observed Variable	R-Square
U0	0.710
U1	0.682
U2	0.650
U3	0.666
Y0	0.620
Y1	0.491
Y2	0.543
Y3	0.608

Latent Variable	R-Square
IU	0.010
SU	0.012
IY	0.019
SY	0.021

162

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

Technical 4 Output

ESTIMATED MEANS FOR THE LATENT VARIABLES

	IU	SU	IY	SY	MALE
1	0.305	0.758	0.312	0.204	0.536

ESTIMATED COVARIANCE MATRIX FOR THE LATENT VARIABLES

	IU	SU	IY	SY	MALE
IU	8.062				
SU	-1.170	0.694			
IY	1.214	-0.007	0.285		
SY	-0.010	0.003	-0.042	0.055	
MALE	0.142	-0.045	0.037	-0.017	0.249

Output Excerpts Step 2 Of A Two-Part Growth Model (Continued)

ESTIMATED CORRELATION MATRIX FOR THE LATENT VARIABLES

	IU	SU	IY	SY	MALE
IU	1.000				
SU	-0.495	1.000			
IY	0.801	-0.015	1.000		
SY	-0.014	0.016	-0.336	1.000	
MALE	0.100	-0.109	0.139	-0.145	1.000

Multiple Populations

165

Advantages Of Growth Modeling In A Latent Variable Framework

- Flexible curve shape
- Individually-varying times of observation
- Regressions among random effects
- Multiple processes
- Modeling of zeroes
- **Multiple populations**
- Multiple indicators
- Embedded growth models
- Categorical latent variables: growth mixtures

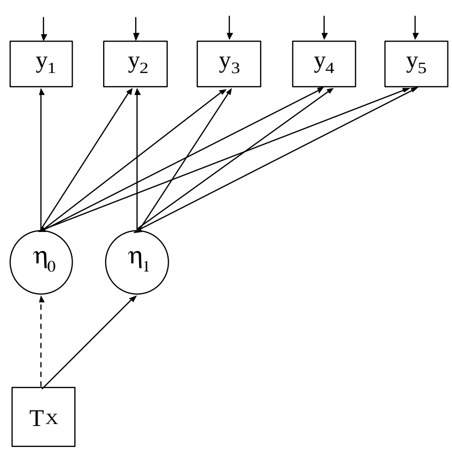
166

Multiple Population Growth Modeling

- Group as a dummy variable
- Multiple-group analysis
- Multiple-group analysis of randomized interventions

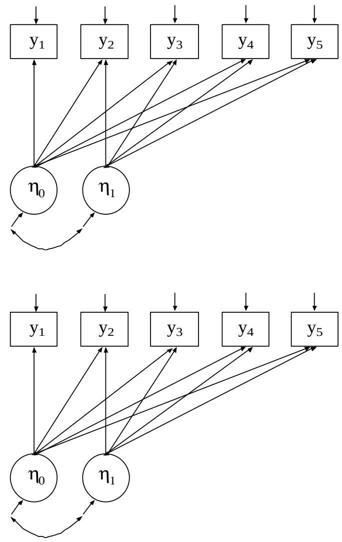
167

Group Dummy Variable As A Covariate



168

Two-Group Model



Multiple Population Growth Modeling Specifications

Let y_{git} denote the outcome for population (group) g, individual i, and timepoint t,

$$\textit{Level 1}: y_{gti} = \eta_{g0i} + \eta_{g1i}\, x_t + \varepsilon_{gti}, \quad (65)$$
$$\textit{Level 2a}: \eta_{g0i} = \alpha_{g0} + \gamma_{g0}\, w_{gi} + \zeta_{g0i}, \quad (66)$$
$$\textit{Level 2b}: \eta_{g1i} = \alpha_{g1} + \gamma_{g1}\, w_{gi} + \zeta_{g1i}, \quad (67)$$

Measurement invariance (level-1 equation): time-invariant intercept 0 and slopes 1, x_t

Structural differences (level-2): α_g , γ_g , $V(\zeta_g)$

Alternative parameterization:

$$\textit{Level 1}: y_{gti} = \nu + \eta_{g0i} + \eta_{g1i}\, x_t + \varepsilon_{gti}, \quad (68)$$

with α_{10} fixed at zero in level 2a.

Analysis steps:

1. Separate growth analysis for each group
2. Joint analysis of all groups, free structural parameters
3. Join analysis of all groups, tests of structural parameter invariance

NLSY: Multiple Cohort Structure

Birth Year Cohort	Age[a]																			
	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
57								82	83	84	85	86	87	88	89	90	91	92	93	94
58							82	83	84	85	86	87	88	89	90	91	92	93	94	
59						82	83	84	85	86	87	88	89	90	91	92	93	94		
60					82	83	84	85	86	87	88	89	90	91	92	93	94			
61				82	83	84	85	86	87	88	89	90	91	92	93	94				
62			82	83	84	85	86	87	88	89	90	91	92	93	94					
63		82	83	84	85	86	87	88	89	90	91	92	93	94						
64	82	83	84	85	86	87	88	89	90	91	92	93	94							

[a] Non-shaded areas represent years in which alcohol measures were obtained

171

Multiple Group Modeling Of Multiple Cohorts

- Data – two cohorts born in 1961 and 1962 measured on the frequency of heavy drinking in the years 1983, 1984, 1988, and 1989
- Development of heavy drinking across chronological age, not year of measurement, is of interest

Cohort/Year	1983	1984	1988	1989
1961 (older)	**22**	23	**2**7	28
1962 (younger)	21	**22**	26	**27**

Cohort/Age	21	22	23	24	25	26	27	28
1961 (older)		**83**	**84**				**88**	**89**
1962 (younger)	**83**	**84**				**88**	**89**	

172

Multiple Group Modeling Of Multiple Cohorts (Continued)

- Time scores calculated for age, not year of measurement

Age	**21**	**22**	**23**	24	25	**26**	**27**	**28**
Time score	**0**	**1**	**2**	3	4	**5**	**6**	**7**

Cohort 1961 time scores 1 2 6 7
Cohort 1962 time scores 0 1 5 6

- Can test the degree of measurement and structural invariance
  - Test of full invariance
    - Growth factor means, variances, and covariances held equal across cohorts
    - Residual variances of shared ages held equal across cohorts

173

Input For Multiple Group Modeling Of Multiple Cohorts

```
TITLE:        Multiple Group Modeling Of Multiple Cohorts
DATA:         FILE IS cohort.dat;
VARIABLE:     NAMES ARE cohort hd83 hd84 hd88 hd89;
              MISSING ARE *;
              USEV = hd83 hd84 hd88 hd89;
              GROUPING IS cohort (61 = older 62 = younger);
MODEL:        i s | hd83@0 hd84@1 hd88@5 hd89@6;
              [i] (1);
              [s] (2);
              i (3);
              s (4);
              i WITH s (5);
```

174

Input For Multiple Group Modeling Of Multiple Cohorts (Continued)

```
MODEL older:
              i s | hd83@1 hd84@2 hd88@6 hd89@7;
              hd83 (6);
              hd88 (7);
MODEL younger:
              hd84 (6);
              hd89 (7);

OUTPUT:       STANDARDIZED;
```

175

Output Excerpts Multiple Group Modeling Of Multiple Cohorts

Tests Of Model Fit

```
Chi-Square Test of Model Fit
          Value                       68.096
          Degrees of Freedom              17
          P-Value                      .0000

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                      .047
          90 Percent C.I.               .036   .059
```

176

Output Excerpts Multiple Group Modeling Of Multiple Cohorts (Continued)

Model Results

	Estimates	S.E.	Est./S.E.	Std	StdYX
Group OLDER					
I WITH					
S	-.111	.010	-11.390	-.537	-.537
Residual Variances					
HD83	1.141	.046	24.996	1.141	.445
HD84	1.062	.057	18.489	1.062	.453
HD88	1.028	.041	25.326	1.028	.455
HD89	.753	.053	14.107	.753	.358
Variances					
I	1.618	.068	23.651	1.000	1.000
S	.026	.002	13.372	1.000	1.000
Means					
I	1.054	.030	35.393	.828	.828
S	-.032	.005	-6.611	-.200	-.200

177

Output Excerpts Multiple Group Modeling Of Multiple Cohorts (Continued)

GROUP YOUNGER					
Residual Variances					
HD83	1.049	.066	15.916	1.049	.393
HD84	1.141	.046	24.996	1.141	.445
HD88	1.126	.056	19.924	1.126	.491
HD89	1.028	.041	25.326	1.028	.455

178

Preventive Interventions Randomized Trials

Prevention Science Methodology Group (PSMG)

Developmental Epidemiological Framework:

- Determining the levels and variation in risk and protective factors as well as developmental paths within a defined population in the absence of intervention
- Directing interventions at these risk and protective factors in an effort to change the developmental trajectories in a defined population
- Evaluating variation in intervention impact across risk levels and contexts on proximal and distal outcomes, thereby empirically testing the developmental model

179

Aggressive Classroom Behavior: The GBG Intervention

Muthén & Curran (1997, Psychological Methods)

The Johns Hopkins Prevention Center carried out a school-based preventive intervention randomized trial in Baltimore public schools starting in grade 1. One of the interventions tested was the Good Behavior Game intervention, a classroom based behavior management strategy promoting good behavior. It was designed specifically to reduce aggressive behavior of first graders and was aimed at longer term impact on aggression through middle school.

One first grade classroom in a school was randomly assigned to receive the Good Behavior Game intervention and another matched classroom in the school was treated as control. After an initial assessment in fall of first grade, the intervention was administered during the first two grades.

180

Aggressive Classroom Behavior: The GBG Intervention (Continued)

The outcome variable of interest was teacher ratings (TOCA-R) of each child's aggressive behavior (breaks rules, harms property, fights, etc.) in the classroom through grades 1 – 6. Eight teacher ratings were made from fall and spring for the first two grades and every spring in grades 3 – 6.

The most important scientific question was whether the Good Behavior Game reduces the slope of the aggression trajectory across time. It was also of interest to know whether the intervention varies in impact for children who started out as high aggressive versus low aggressive.

Analyses in Muthén-Curran (1997) were based on data for 75 boys in the GBG group who stayed in the intervention condition for two years and 111 boys in the control group.

181

The GBG Aggression Example: Analysis Results

Muthén & Curran (1997):

- Step 1: Control group analysis
- Step 2: Treatment group analysis
- Step 3: Two-group analysis w/out interactions
- Step 4: Two-group analysis with interactions
- Step 5: Sensitivity analysis of final model
- Step 6: Power analysis

182

MUTHÉN AND CURRAN

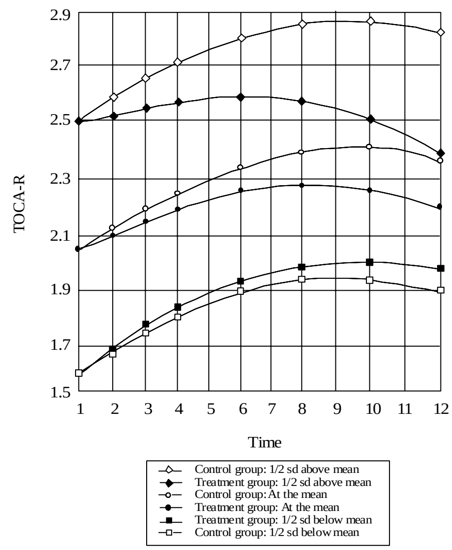
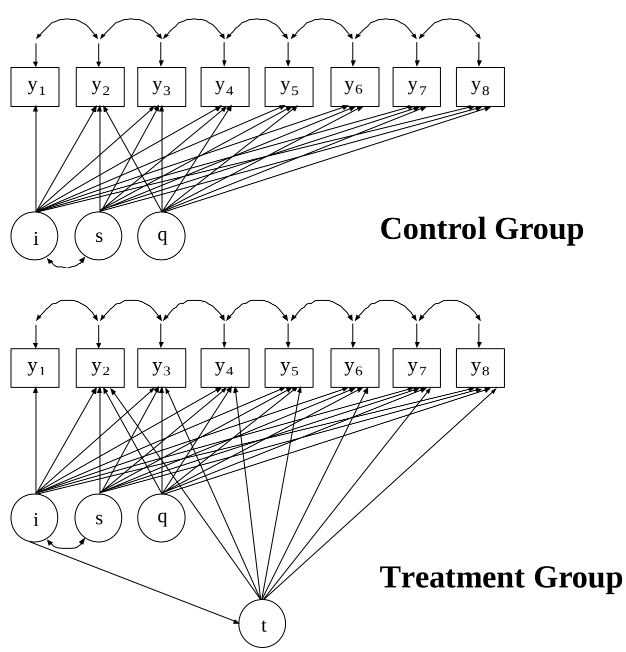


Figure 15. Model implied growth trajectories of Teacher Observation of Classroom Behavior—Revised (TOCA-R) scores as a function of initial status. Each timepoint represents one 6-month interval.

183



184

Input Excerpts For Aggressive Behavior Intervention Using A Multiple Group Growth Model With A Regression Among Random Effects

```
TITLE:          Aggressive behavior intervention growth model
                n = 111 for control group
                n = 75 for tx group
MODEL:          i s q | y1@0 y2@1 y3@2 y4@3 y5@5 y6@7 y7@9 y8@11;
                i t | y1@0 y2@1 y3@2 y4@3 y5@5 y6@7 y7@9 y8@11;
                [y1-y8] (1);   !alternative growth model
                [i@0];         !parameterization
                i (2);
                s (3);
                i WITH s (4);
                [s] (5);
                [q] (6);
                t@0 q@0;
                q WITH i@0 s@0 t@0; y1-y7 PWITH y2-y8;
                t ON i;
```

185

Input Excerpts For Aggressive Behavior Intervention Using A Multiple Group Growth Model With A Regression Among Random Effects (Continued)

```
MODEL control:
                [s] (5);
                [q] (6);
                t ON i@0;
                [t@0];
```

186

Output Excerpts Aggressive Behavior Intervention Using A Multiple Group Growth Model With A Regression Among Random Effects

Tests Of Model Fit

```
Chi-Square Test of Model Fit

         Value                            64.553
         Degrees of Freedom                   50
         P-Value                           .0809

RMSEA (Root Mean Square Error Of Approximation)

         Estimate                           .056
         90 Percent C.I.                    .000   .092
```

187

Output Excerpts Aggressive Behavior Intervention Using A Multiple Group Growth Model With A Regression Among Random Effects (Continued)

Group Control

Observed Variable	R-Square
Y1	.644
Y2	.642
Y3	.663
Y4	.615
Y5	.637
Y6	.703
Y7	.812
Y8	.818

Group Tx

Observed Variable	R-Square
Y1	.600
Y2	.623
Y3	.568
Y4	.464
Y5	.425
Y6	.399
Y7	.703
Y8	.527

188

Output Excerpts Aggressive Behavior Intervention Using A Multiple Group Growth Model With A Regression Among Random Effects (Continued)

Group Control

	Estimates	S.E.	Est./S.E.	Std	StdYX
T ON					
I	**.000**	**.000**	**.000**	**999.000**	**999.000**
Residual Variances					
Y1	.444	.088	5.056	.444	.356
Y2	.449	.079	5.714	.449	.358
Y3	.414	.069	6.026	.414	.337
Y4	.522	.080	6.551	.522	.385
Y5	.512	.079	6.469	.512	.363
Y6	.422	.074	5.677	.422	.297
Y7	.264	.083	3.186	.264	.188
Y8	.291	.094	3.097	.291	.182
T	.000	.000	.000	999.000	999.000
Variances					
I	.803	.109	7.330	1.000	1.000
S	.004	.001	3.869	1.000	1.000
Q	.000	.000	.000	999.000	999.000

189

Output Excerpts Aggressive Behavior Intervention Using A Multiple Group Growth Model With A Regression Among Random Effects (Continued)

	Estimates	S.E.	Est./S.E.	Std	StdYX
Means					
I	.000	.000	.000	.000	.000
S	.086	.021	4.035	1.285	1.285
Q	-.005	.002	-3.005	999.000	999.000
Intercepts					
Y1	2.041	.078	26.020	2.041	1.828
Y2	2.041	.078	26.020	2.041	1.823
Y3	2.041	.078	26.020	2.041	1.841
Y4	2.041	.078	26.020	2.041	1.753
Y5	2.041	.078	26.020	2.041	1.718
Y6	2.041	.078	26.020	2.041	1.711
Y7	2.041	.078	26.020	2.041	1.724
Y8	2.041	.078	26.020	2.041	1.612
T	.000	.000	.000	999.000	999.000

190

Output Excerpts Aggressive Behavior Intervention Using A Multiple Group Growth Model With A Regression Among Random Effects (Continued)

Group Tx

	Estimates	S.E.	Est./S.E.	Std	StdYX
T ON					
I	**-.052**	**.015**	**-3.347**	**-1.000**	**-1.000**
Residual Variances					
Y1	.535	.141	3.801	.535	.400
Y2	.439	.122	3.595	.439	.377
Y3	.501	.108	4.653	.501	.432
Y4	.701	.132	5.332	.701	.536
Y5	.736	.133	5.545	.736	.575
Y6	.805	.152	5.288	.805	.601
Y7	.245	.104	2.364	.245	.297
Y8	.609	.182	3.351	.609	.473
T	.000	.000	.000	.000	.000
Variances					
I	.803	.109	7.330	1.000	1.000
S	.004	.001	3.869	1.000	1.000
Q	.000	.000	.000	999.000	999.000

Output Excerpts Aggressive Behavior Intervention Using A Multiple Group Growth Model With A Regression Among Random Effects (Continued)

	Estimates	S.E.	Est./S.E.	Std	StdYX
Means					
I	.000	.000	.000	.000	.000
S	.086	.021	4.035	1.285	1.285
Q	-.005	.002	-3.005	999.000	999.000
Intercepts					
Y1	2.041	.078	26.020	2.041	1.764
Y2	2.041	.078	26.020	2.041	1.893
Y3	2.041	.078	26.020	2.041	1.895
Y4	2.041	.078	26.020	2.041	1.785
Y5	2.041	.078	26.020	2.041	1.805
Y6	2.041	.078	26.020	2.041	1.764
Y7	2.041	.078	26.020	2.041	2.248
Y8	2.041	.078	26.020	2.041	1.799
T	-.016	.013	-1.225	-.341	-.341

Growth Modeling With Multiple Indicators

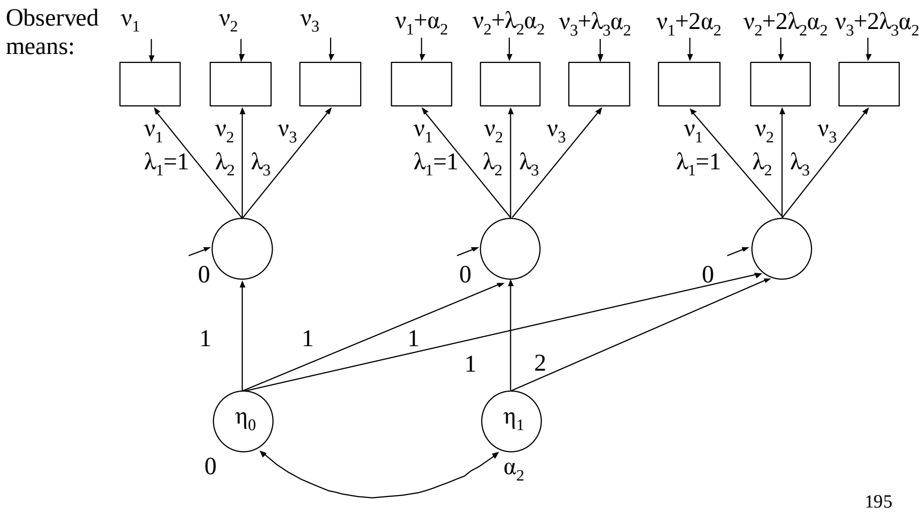
193

Advantages Of Growth Modeling In A Latent Variable Framework

- Flexible curve shape
- Individually-varying times of observation
- Regressions among random effects
- Multiple processes
- Modeling of zeroes
- Multiple populations
- **Multiple indicators**
- Embedded growth models
- Categorical latent variables: growth mixtures

194

Growth Of Latent Variable Construct Measured By Multiple Indicators



195

Multiple Indicator Growth Modeling Specifications

Let y_{jti} denote the outcome for individual i, indicator j, and timepoint t, and let η_{ti} denote a latent variable construct,

Level 1a (measurement part):

$$y_{jti} = \nu_{jt} + \lambda_{jt}\, \eta_{ti} + \varepsilon_{jti}, \qquad (44)$$

Level 1b : $\eta_{ti} = \eta_{0i} + \eta_{1i}\; x_t + \zeta_{ti}$, (45)

Level 2a : $\eta_{0i} = \alpha_0 + \gamma_0\, w_i + \zeta_{0i}$, (46)

Level 2b : $\eta_{1i} = \alpha_1 + \gamma_1\, w_i + \zeta_{1i}$, (47)

Measurement invariance: time-invariant indicator intercepts and slopes:

$$\nu_{j1} = \nu_{j2} = \ldots \nu_{jT} = \nu_j, \qquad (48)$$

$$\lambda_{j1} = \lambda_{j2} = \ldots \lambda_{jT} = \lambda_j, \qquad (49)$$

where $\lambda_1 = 1$, $\alpha_0 = 0$. $V(\varepsilon_{jti})$ and $V(\zeta_{ti})$ may vary over time. Structural differences: $E(\eta_{ti})$ and $V(\eta_{ti})$ vary over time.

196

Steps In Growth Modeling With Multiple Indicators

- Exploratory factor analysis of indicators for each timepoint
- Determine the shape of the growth curve for each indicator and the sum of the indicators
- Fit a growth model for each indicator—must be the same
- Confirmatory factor analysis of all timepoints together
  - Covariance structure analysis without measurement parameter invariance
  - Covariance structure analysis with invariant loadings
  - Mean and covariance structure analysis with invariant measurement intercepts and loadings
- Growth model with measurement invariance across timepoints

197

Advantages Of Using Multiple Indicators Instead Of An Average

- Estimation of unequal weights
- Partial measurement invariance—changes across time in individual item functioning
- No confounding of time-specific variance and measurement error variance
- Smaller standard errors for growth factor parameters (more power)

198

Classroom Aggression Data (TOCA)

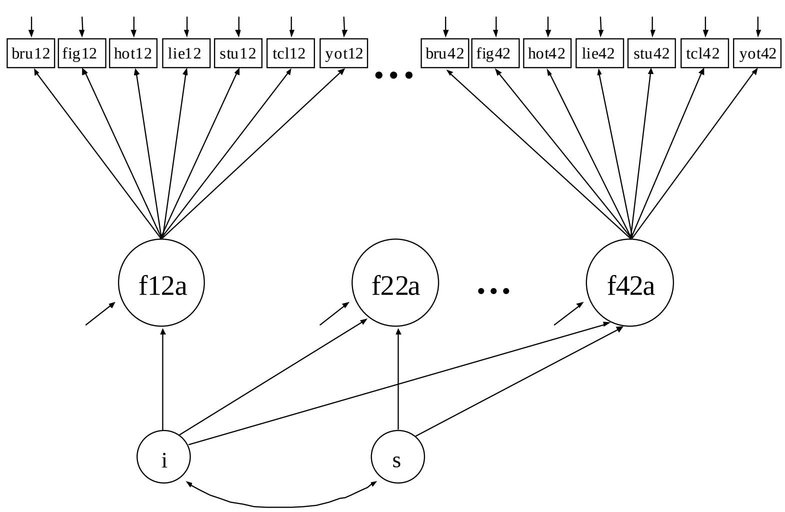
The classroom aggression data are from an intervention study in Baltimore public schools carried out by the Johns Hopkins Prevention Research Center. Subjects were randomized into treatment and control conditions. The TOCA-R instrument was used to measure 10 aggression items at multiple timpoints. The TOCA-R is a teacher rating of student behavior in the classroom. The items are rated on a six-point scale from almost never to almost always.

Data for this analysis include the 342 boys in the control group. Four time points are examined: Spring Grade 1, Spring Grade 2, Spring Grade 3, and Spring Grade 4.

Seven aggression items are used in the analysis:

- Break rules
- Fights
- Harms others
- Lies
- Stubborn
- Teasing classmates
- Yells at others

199



200

Input Excerpts For TOCA Data Multiple Indicator CFA With No Measurement Invariance

```
TITLE:      Multiple indicator CFA with no measurement invariance
.
.
.
MODEL:      f12a BY bru12
                    fig12
                    hot12
                    lie12
                    stu12
                    tcl12
                    yot12;

            f22a BY bru22
                    fig22
                    hot22
                    lie22
                    stu22
                    tcl22
                    yot22;
```

201

Input Excerpts For TOCA Data Multiple Indicator CFA With No Measurement Invariance (Continued)

```
MODEL:      f32a BY bru32
                    fig32
                    hot32
                    lie32
                    stu32
                    tcl32
                    yot32;

            f42a BY bru42
                    fig42
                    hot42
                    lie42
                    stu42
                    tcl42
                    yot42;
```

202

Input Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading Invariance

```
TITLE:      Multiple indicator CFA with factor loading invariance
.
.
.
MODEL:      f12a BY bru12
                    fig12 (1)
                    hot12 (2)
                    lie12 (3)
                    stu12 (4)
                    tcl12 (5)
                    yot12 (6);

            f22a BY bru22
                    fig22 (1)
                    hot22 (2)
                    lie22 (3)
                    stu22 (4)
                    tcl22 (5)
                    yot22 (6);
```

203

Input Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading Invariance (Continued)

```
MODEL:      f32a BY bru32
                    fig32 (1)
                    hot32 (2)
                    lie32 (3)
                    stu32 (4)
                    tcl32 (5)
                    yot32 (6);

            f42a BY bru42
                    fig42 (1)
                    hot42 (2)
                    lie42 (3)
                    stu42 (4)
                    tcl42 (5)
                    yot42 (6);
```

204

Input Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading And Intercept Invariance

```
TITLE:       Multiple indicator CFA with factor loading and intercept
             incariance
.
.
.
MODEL:       f12a BY bru12
                     fig12 (1)
                     hot12 (2)
                     lie12 (3)
                     stu12 (4)
                     tcl12 (5)
                     yot12 (6);
             f22a BY bru22
                     fig22 (1)
                     hot22 (2)
                     lie22 (3)
                     stu22 (4)
                     tcl22 (5)
                     yot22 (6);
```

205

Input Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading And Intercept Invariance (Continued)

```
MODEL:       f32a BY bru32
                     fig32 (1)
                     hot32 (2)
                     lie32 (3)
                     stu32 (4)
                     tcl32 (5)
                     yot32 (6);

             f42a BY bru42
                     fig42 (1)
                     hot42 (2)
                     lie42 (3)
                     stu42 (4)
                     tcl42 (5)
                     yot42 (6);
```

206

Input Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading And Intercept Invariance (Continued)

```
[bru12 bru22 bru32 bru42] (7);
[fig12 fig22 fig32 fig42] (8);
[hot12 hot22 hot32 hot42] (9);
[lie12 lie22 lie32 lie42] (10);
[stu12 stu22 stu32 stu42] (11);
[tcl12 tcl22 tcl32 tcl42] (12);
[yot12 yot22 yot32 yot42] (13);

[f12a@0 f22a f32a f42a];
```

207

Input Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading Invariance And Partial Intercept Invariance

```
TITLE:      Multiple indicator CFA with factor loading and partial
            intercept invariance

MODEL:      f12a BY bru12
                    fig12 (1)
                    hot12 (2)
                    lie12 (3)
                    stu12 (4)
                    tcl12 (5)
                    yot12 (6);
            f22a BY bru22
                    fig22 (1)
                    hot22 (2)
                    lie22 (3)
                    stu22 (4)
                    tcl22 (5)
                    yot22 (6);
```

208

Input Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading Invariance And Partial Intercept Invariance (Continued)

```
f32a BY bru32
        fig32 (1)
        hot32 (2)
        lie32 (3)
        stu32 (4)
        tcl32 (5)
        yot32 (6);
f42a BY bru42
        fig42 (1)
        hot42 (2)
        lie42 (3)
        stu42 (4)
        tcl42 (5)
        yot42 (6);
```

209

Input Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading Invariance And Partial Intercept Invariance (Continued)

```
[bru12 bru22 bru32 bru42] (7);
[fig12 fig22 fig32 fig42] (8);
[hot12 hot22 hot32      ] (9);
[lie12 lie22 lie32 lie42] (10);
[stu12 stu22            ] (11);
[tcl12 tcl22 tcl32      ] (12);
[yot12 yot22 yot32 yot42] (13);

[f12a@0 f22a f32a f42a];
```

210

Summary of Analysis Results For TOCA Measurement Invariance Models

Model	Chi-Square (d.f.)	Difference (d.f. diff.)
Measurement non-invariance	567.08 (344)	
Factor loading invariance	581.29 (362)	14.21 (18)
Factor loading and intercept invariance	654.59 (380)	73.30* (18)
Factor loading and partial intercept invariance	606.97 (376)	25.68* (14)
Factor loading and partial intercept invariance with a linear growth structure	614.74 (381)	7.77 (5)

211

Summary of Analysis Results For TOCA Measurement Invariance Models (Continued)

Explanation of Chi-Square Differences

Factor loading invariance	(18)	6 factor loadings instead of 24
Factor loading and intercept invariance	(18)	7 intercepts plus 3 factor means instead of 28 intercepts
Factor loading and partial intercept invariance	(14)	4 additional intercepts
Factor loading and partial intercept invariance with a linear growth structure	(5)	1 growth factor mean instead of 3 factor means 2 growth factor variances, 1 growth factor covariance, 4 factor residual variances instead of 10 factor variances/covariances

212

Input Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading And Intercept Invariance With A Linear Growth Structure

```
MODEL:      f12a BY bru12
                    fig12 (1)
                    hot12 (2)
                    lie12 (3)
                    stu12 (4)
                    tcl12 (5)
                    yot12 (6);
            f22a BY bru22
                    fig22 (1)
                    hot22 (2)
                    lie22 (3)
                    stu22 (4)
                    tcl22 (5)
                    yot22 (6);
```

213

Input Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading And Intercept Invariance With A Linear Growth Structure (Continued)

```
MODEL:      f32a BY bru32
                    fig32 (1)
                    hot32 (2)
                    lie32 (3)
                    stu32 (4)
                    tcl32 (5)
                    yot32 (6);
            f42a BY bru42
                    fig42 (1)
                    hot42 (2)
                    lie42 (3)
                    stu42 (4)
                    tcl42 (5)
                    yot42 (6);
```

214

Input Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading And Intercept Invariance With A Linear Growth Structure (Continued)

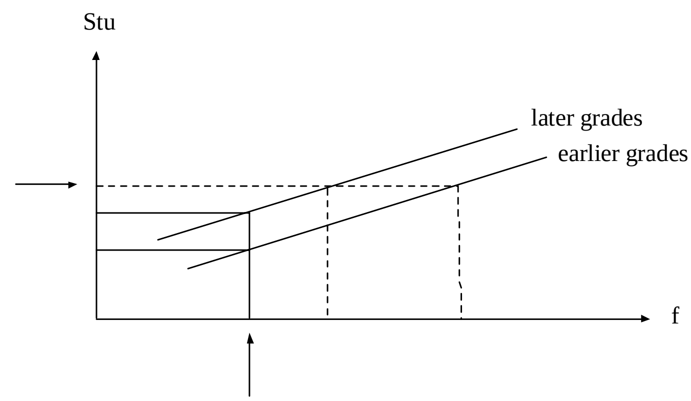
```
[bru12 bru22 bru32 bru42] (7);
[fig12 fig22 fig32 fig42] (8);
[hot12 hot22 hot32      ] (9);
[lie12 lie22 lie32 lie42] (10);
[stu12 stu22            ] (11);
[tcl12 tcl22 tcl32      ] (12);
[yot12 yot22 yot32 yot42] (13);

i s | f12a@0 f22a@1 f32a@2 f42a@3;



Alternative language:

i BY f12a-f42a@1;
s BY f12a@0 f22a@1 f32a@2 f42a@3;
[f12a-f42a@0 i@0 s];
```

215

Output Excerpts For TOCA Data Multiple Indicator CFA With Factor Loading And Intercept Invariance With A Linear Growth Structure

	Estimates	S.E.	Est./S.E.	Std	StdYX
F12A |					
BRU12	1.000	.000	.000	.190	.786
FIG12	1.097	.039	28.425	.208	.868
HOT12	.986	.037	26.586	.187	.811
LIE12	.967	.041	23.769	.184	.742
STU12	.880	.041	21.393	.167	.667
TCL12	1.034	.039	26.206	.196	.786
YOT12	.932	.039	23.647	.177	.709
Intercepts					
STU12	.331	.013	25.408	.331	1.324
STU22	.331	.013	25.408	.331	1.231
STU32	.417	.017	24.345	.417	1.592
STU42	.390	.017	23.265	.390	1.496

216

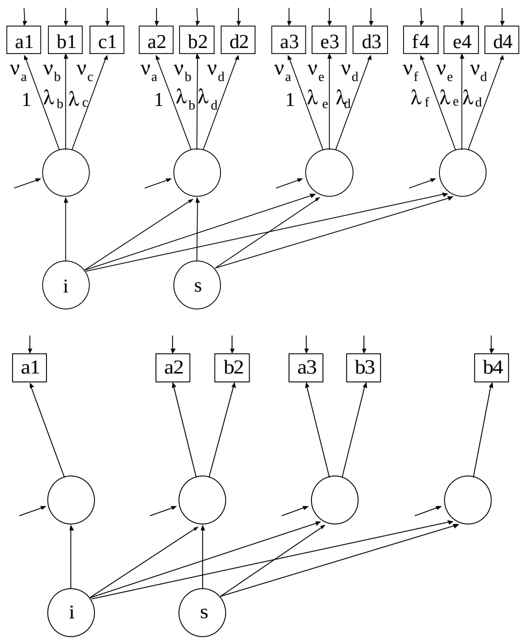


217

Degrees Of Invariance Across Time

- Case 1
  - Same items
  - All items invariant
  - Same construct
- Case 2
  - Same items
  - Some items non-invariant
  - Same construct
- Case 3
  - Different items
  - Some items invariant
  - Same construct
- Case 4
  - Different items
  - Some items invariant
  - Different construct

218



219

Embedded Growth Models

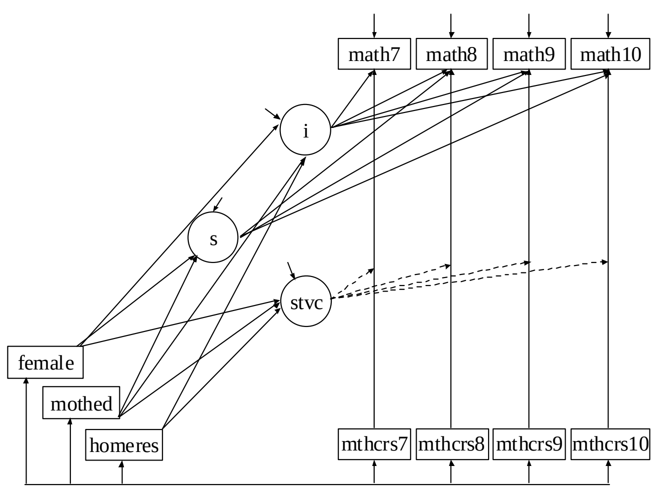
220

Advantages Of Growth Modeling In A Latent Variable Framework

- Flexible curve shape
- Individually-varying times of observation
- Regressions among random effects
- Multiple processes
- Modeling of zeroes
- Multiple populations
- Multiple indicators
- **Embedded growth models**
- Categorical latent variables: growth mixtures

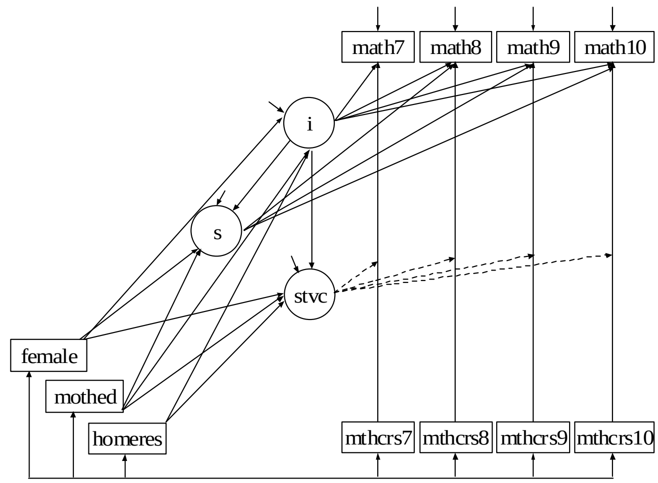
221

Growth Modeling With Time-Varying Covariates



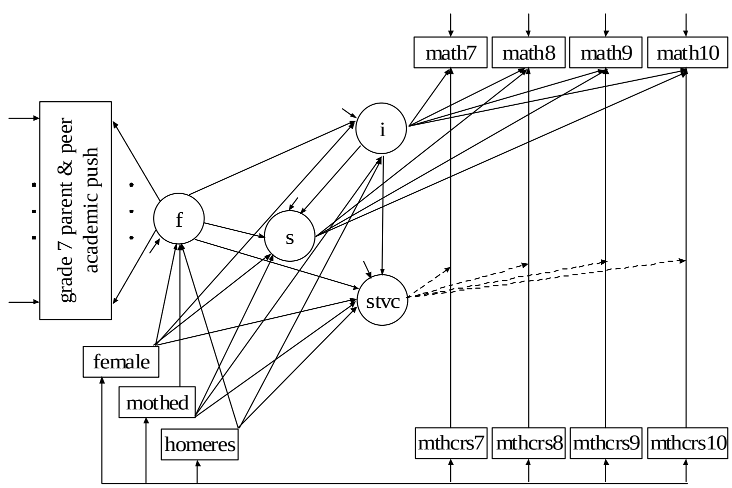
222

A Generalized Growth Model



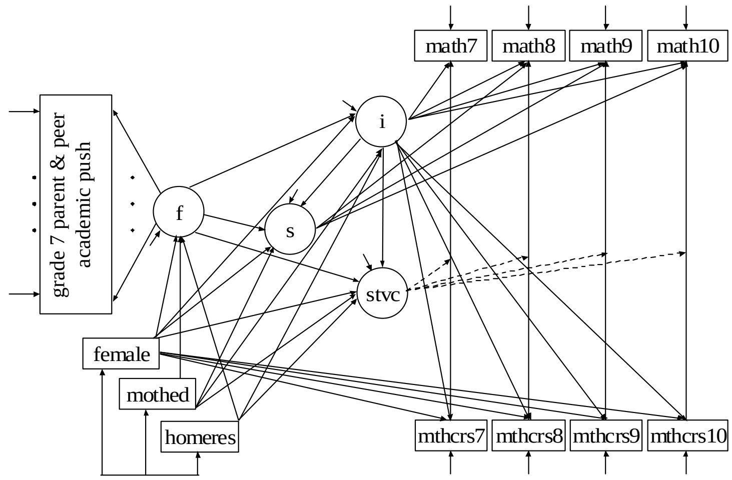
223

A Generalized Growth Model



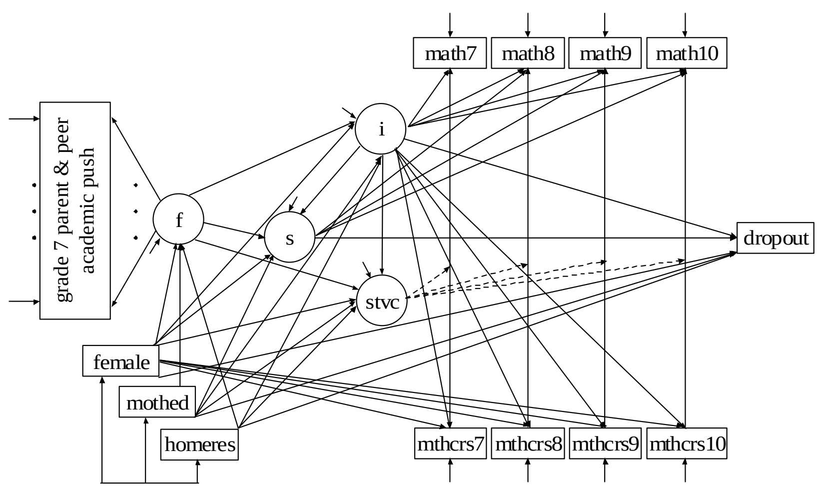
224

A Generalized Growth Model



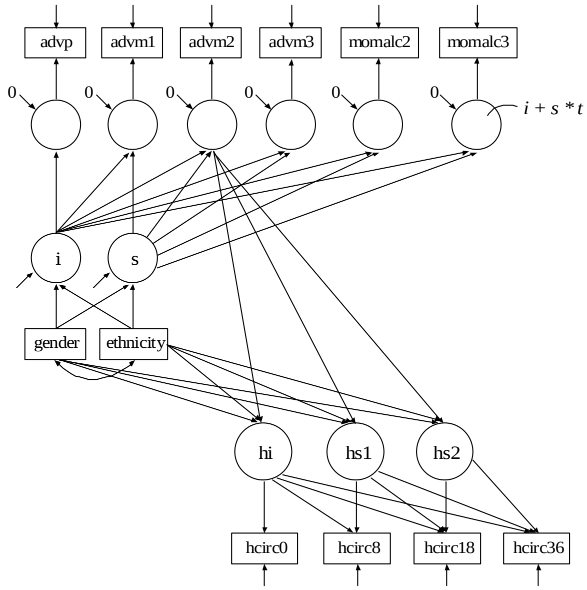
225

A Generalized Growth Model



226

Two Linked Processes



227

Input Excerpts For Two Linked Processes With Measurement Error In The Covariates

```
TITLE:     Embedded growth model with measurement error in the
           covariates and sequential processes
           advp: mother’s drinking before pregnancy
           advm1-advm3: drinking in first trimester
           momalc2-momalc3: drinking in 2nd and 3rd trimesters
           hcirc0-hcirc36; head circumference
MODEL:     fadvp    BY  advp;       fadvp@0;
           fadvm1   BY  advm1;     fadvm1@0;
           fadvm2   BY  advm2;     fadvm2@0;
           fadvm3   BY  advm3;     fadvm3@0;
           fmomalc2 BY  momalc2;  fmomalc2@0;
           fmomalc3 BY  momalc3;  fmomalc3@0;
           i BY fadvp-fmomalc3@1;
           s BY fadvp@0  fadvm1@1  fadvm2*2  fadvm3*3
                  fmomalc2-fmomalc3*5 (1);
           [advp-momalc3@0 fadvp-fmomalc3@0 i s];
```

228

Input Excerpts For Two Linked Processes With Measurement Error In The Covariates (Continued)

```
advp WITH advm1; advm1 WITH advm2; advm3 WITH advm2;

i s ON gender eth; s WITH i;


hi   BY  hcirc0-hcirc36@1;
hs1  BY  hcirc0@0 hcirc8@1.196 hcirc36@1.196 hcirc36@1.196;
hs2  BY  hcirc0@0 hcirc8@0 hcirc18@1 hcirc36*2;

[hcirc0-hcirc36@0 hi*34 hs1 hs2];

hs1 WITH hs2@0; hi WITH hs2@0; hi WITH hs1@0;
hi  WITH i@0; hi   WITH s@0; hs1  WITH i@0;
hi1 WITH s@0; hs2  WITH i@0; hs2  WITH s@0;

hi-hs2 ON gender eth fadvm2;
```

229

Power For Growth Models

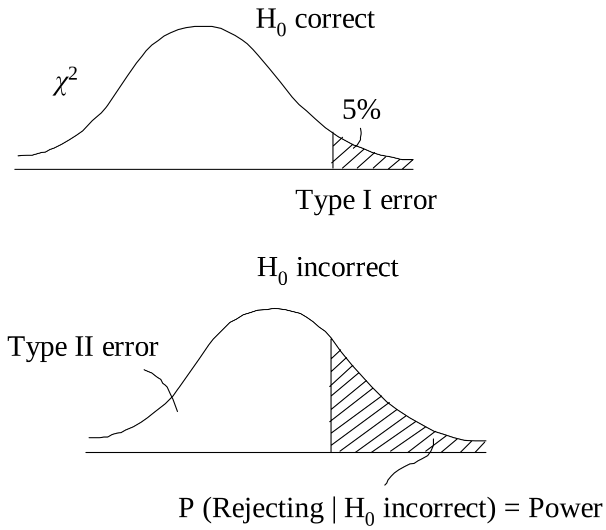
230

Designing Future Studies: Power

- Computing power for growth models using Satorra-Saris (Muthén & Curran, 1997; examples)
- Computing power using Monte Carlo studies (Muthén & Muthén, 2002)
- Power calculation web site – PSMG
- Multilevel power (Miyazaki & Raudenbush, 2000; Moerbeek, Breukelen & Berger, 2000; Raudenbush, 1997; Raudenbush & Liu, 2000)
- School-based studies (Brown & Liao, 1999: Principles for designing randomized preventive trials)
- Multiple- (sequential-) cohort power
- Designs for follow-up (Brown, Indurkhia, & Kellam, 2000)

231

Designing Future Studies: Power



232

Power Estimation For Growth Models Using Satorra & Saris (1985)

- Step 1: Create mean vector and covariance matrix for hypothesized parameter values
- Step 2: Analyze as if sample statistics and check that parameter values are recovered
- Step 3: Analyze as if sample statistics, misspecifying the model by fixing treatment effect(s) at zero
- Step 4: Use printed x^2 as an appropriate noncentrality parameter and computer power.

Muthén & Curran (1997): Artificial and real data situations.

233

Input For Step 1 Of Power Calculation

```
TITLE:     Power calculation for a growth model
           Step 1: Computing the population means and
           covariance matrix
DATA:      FILE IS artific.dat;
           TYPE IS MEANS COVARIANCE;
           NOBSERVATIONS = 500;
VARIABLE:  NAMES ARE y1-y4;
MODEL:     i s | y1@0 y2@1 y3@2 y4@3;
           i@.5;
           s@.1;
           i WITH s@0;
           y1-y4@.5;
OUTPUT:    STANDARDIZED RESIUDAL;
```

234

Data For Step 1
Of Power Calculation (Continued)

```
0  0  0  0
1
0  1
0  0  1
0  0  0  1
```

235

Input For Step 2
Of Power Calculation

```
TITLE:     Power calculation for a growth model
           Step 2: Analyzing the population means and
           covariance matrix to check that parameters are
           recovered
DATA:      FILE IS pop.dat;
           TYPE IS MEANS COVARIANCE;
           NOBSERVATIONS = 500;
VARIABLE:  NAMES ARE y1-y4;
MODEL:     i s | y1@0 y2@1 y3@2 y4@3;
OUTPUT:    STANDARDIZED RESIUDAL;
```

236

Data For Step 2 Of Power Calculation (Continued)

Data From Step 1 Residual Output

```
0   .2   .4   .6
1
.5   1.1
.5   .7   1.4
.5   .8   1.1   1.9
```

237

Input For Step 3 Of Power Calculation

```
TITLE:     Power calculation for a growth model
           Step 3: Analyzing the population means and
           covariance matrix with a misspecified model
DATA:      FILE IS pop.dat;
           TYPE IS MEANS COVARIANCE;
           NOBSERVATIONS = 50;
VARIABLE:  NAMES ARE y1-y4;
MODEL:     i s | y1@0 y2@1 y3@2 y4@3;
OUTPUT:    STANDARDIZED RESIUDAL;
```

238

Step 4 Of Power Calculation

Output Excerpt From Step 3

```
Chi-Square Test of Model Fit

          Value                            9.286
          Degrees of Freedom                   6
          P-Value                          .1580
```

Power Algorithm in SAS

```
DATA POWER;
DF=1; CRIT=3.841459;
LAMBDA=9.286;
Power=(1 – (PROBCHI(CRIT, DF, LAMBDA)));
RUN;
```

239

Step 4 Of Power Calculation (Continued)

Results From Power Algorithm

SAMPLE SIZE	POWER
44	0.80
50	0.85
100	0.98
200	0.99

Note: **Non-centrality parameter = printed chi-square value from Step 3 = 2*sample size*F**

240

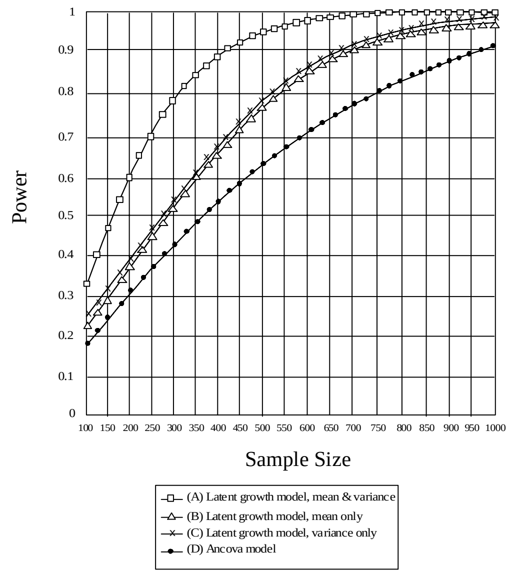


Figure 6. Power to detect a main effect of *ES* = .20 assessed at Time 5.

241

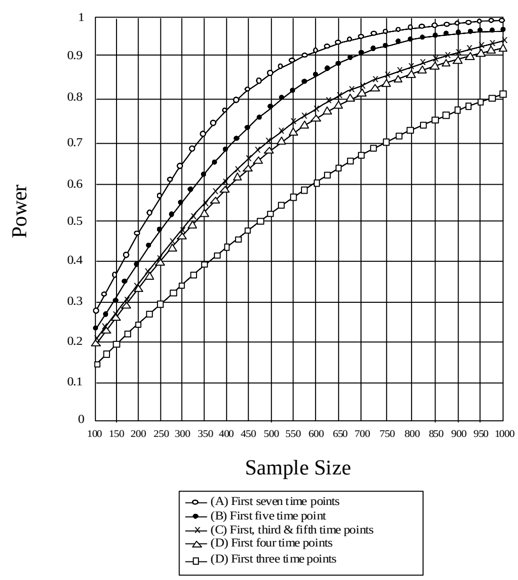


Figure 7. Power to detect a main effect of *ES* = .20 assessed at Time 5 varying as a function of total number of measurement occasions.

242

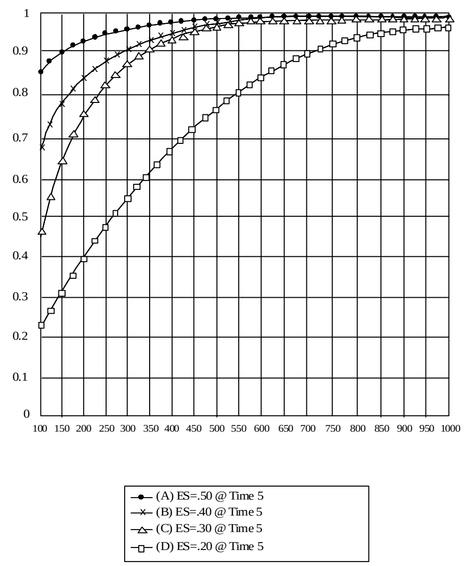


Figure 8. Power to detect various effect sizes assessed at Time 5 based on the first five measurement occasions

243

Power Estimation For Growth Models Using Monte Carlo Studies

Muthén & Muthén (2002)

244

Input Power Estimation For Growth Models Using Monte Carlo Studies

```
TITLE:        This is an example of a Monte Carlo
              simulation study for a linear growth model
              for a continuous outcome with missing data
              where attrition is predicted by time-
              invariant covariates (MAR)
MONTECARLO:   NAMES ARE y1-y4 x1 x2;
              NOBSERVATIONS = 500;
              NREPS = 500;
              SEED = 4533;
              CUTPOINTS = x2(1);
              MISSING = y1-y4;
```

245

Input Power Estimation For Growth Models Using Monte Carlo Studies (Continued)

```
MODEL POPULATION:    x1-x2@1;
                     [x1-x2@0];
                     i s | y1@0 y2@1 y3@2 y4@3;
                     [i*1 s*2];
                     i*1; s*.2; i WITH s*.1;
                     y1-y4*.5;
                     i ON x1*1 x2*.5;
                     s ON x1*.4 x2*.25;
MODEL MISSING:       [y1-y4@-1];
                     y1 ON x1*.4 x2*.2;
                     y2 ON x1*.8 x2*.4;
                     y3 ON x1*1.6 x2*.8;
                     y4 ON x1*3.2 x2*1.6;
```

246

Input Power Estimation For Growth Models Using Monte Carlo Studies (Continued)

```
ANALYSIS:     TYPE = MISSING H1;
MODEL:        i s | y1@0 y2@1 y3@2 y4@3;
              [i*1 s*2];
              i*1; s*.2; i WITH s*.1;
              y1-y4*.5;
              i ON x1*1 x2*.5;
              s ON x1*.4 x2*.25;
OUTPUT:       TECH9;
```

247

Output Excerpts Power Estimation For Growth Models Using Monte Carlo Studies

Model Results

```
                          ESTIMATES              S.E.     M. S. E.  95%   %Sig
           Population   Average  Std. Dev.  Average              Cover Coeff

 I      ON
  X1           1.000   1.0032     0.0598   0.0579    0.0036    0.936  1.000
  X2           0.500   0.5076     0.1554   0.1570    0.0241    0.952  0.908

 S      ON
  X1           0.400   0.3980     0.0366   0.0349    0.0013    0.936  1.000
  X2           0.250   0.2469     0.0865   0.0877    0.0075    0.938  0.830
```

248

Cohort-Sequential Designs and Power

Considerations:

- Model identification
- Number of timepoints needed substantively
- Number of years of the study
- Number of cohorts: More gives longer timespan but greater risk of cohort differences
- Number of measurements per individual
- Number of individuals per cohort
- Number of individuals per age

Tentative conclusion:

Power most influenced by total timespan, not the number of measures per cohort

249

References

(To request a Muthén paper, please email bmuthen@ucla.edu.)

Analysis With Longitudinal Data

Introductory

Bollen, K.A. & Curran, P.J. (2006). Latent curve models. A structural equation perspective. New York: Wiley.

Collins, L.M. & Sayer, A. (Eds.) (2001). New methods for the analysis of change. Washington, D.C.: American Psychological Association.

Curran, P.J. & Bollen, K.A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In Collins, L.M. & Sayer, A.G. (Eds.) New methods for the analysis of change (pp. 105-136). Washington, DC: American Psychological Association.

Duncan, T.E., Duncan, S.C., Strycker, L.A., Li, F., & Alpert, A. (1999). An introduction to latent variable growth curve modeling: concepts, issues, and applications. Mahwah, NJ: Lawrence Erlbaum Associates.

Goldstein, H. (1995). Multilevel statistical models. Second edition. London: Edward Arnold.

Jennrich, R.I. & Schluchter, M.D. (1986). Unbalanced repeated-measures models with structured covariance matrices. Biometrics, 42, 805-820.

250

References (Continued)

Laird, N.M., & Ware, J.H. (1982). Random-effects models for longitudinal data. Biometrics, 38, 963-974.

Lindstrom, M.J. & Bates, D.M. (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. Journal of the American Statistical Association, 83, 1014-1022.

Littell, R., Milliken, G.A., Stroup, W.W., & Wolfinger, R.D. (1996). SAS system for mixed models. Cary NC: SAS Institute.

McArdle, J.J. & Epstein, D. (1987). Latent growth curves within developmental structural equation models. Child Development, 58, 110-133.

McArdle, J.J. & Hamagami, F. (2001). Latent differences score structural models for linear dynamic analyses with incomplete longitudinal data. In Collins, L.M. & Sayer, A. G. (Eds.), New methods for the analysis of change (pp. 137-175). Washington, D.C.: American Psychological Association.

Meredith, W. & Tisak, J. (1990). Latent curve analysis Psychometrika, 55, 107-122.

251

References (Continued)

Muthén, B. (1991). Analysis of longitudinal data using latent variable models with varying parameters. In L. Collins & J. Horn (Eds.), Best methods for the analysis of change. Recent advances, unanswered questions, future directions (pp. 1-17). Washington D.C.: American Psychological Association.

Muthén, B. (2000). Methodological issues in random coefficient growth modeling using a latent variable framework: Applications to the development of heavy drinking. In Multivariate applications in substance use research, J. Rose, L. Chassin, C. Presson & J. Sherman (eds.), Hillsdale, N.J.: Erlbaum, pp. 113-140.

Muthén, B. & Khoo, S.T. (1998). Longitudinal studies of achievement growth using latent variable modeling. Learning and Individual Differences. Special issue: latent growth curve analysis, 10, 73-101.

Muthén, B. & Muthén, L. (2000). The development of heavy drinking and alcohol-related problems from ages 18 to 37 in a U.S. national sample. Journal of Studies on Alcohol, 61, 290-300.

Muthén, B. & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. Biometrics, 55, 463-469.

252

References (Continued)

Rao, C.R. (1958). Some statistical models for comparison of growth curves. Biometrics, 14, 1-17.

Raudenbush, S.W. & Bryk, A.S. (2002). Hierarchical linear models: Applications and data analysis methods. Second edition. Newbury Park, CA: Sage Publications.

Singer, J.D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics, 23, 323-355.

Singer, J.D. & Willett, J.B. (2003). Applied longitudinal data analysis. Modeling change and event occurrence. New York, NY: Oxford University Press.

Snijders, T. & Bosker, R. (1999). Multilevel analysis: An introduction to basic and advanced multilevel modeling. Thousand Oakes, CA: Sage Publications.

Tucker, L.R. (1958). Determination of parameters of a functional relation by factor analysis. Psychometrika, 23, 19-23.

References (Continued)

Advanced

Albert, P.S. & Shih, J.H. (2003). Modeling tumor growth with random onset. Biometrics, 59, 897-906.

Brown, C.H. & Liao, J. (1999). Principles for designing randomized preventive trials in mental health: An emerging development epidemiologic perspective. American Journal of Community Psychology, special issue on prevention science, 27, 673-709.

Brown, C.H., Indurkhya, A. & Kellam, S.K. (2000). Power calculations for data missing by design: applications to a follow-up study of lead exposure and attention. Journal of the American Statistical Association, 95, 383-395.

Collins, L.M. & Sayer, A. (Eds.), New methods for the analysis of change. Washington, D.C.: American Psychological Association.

Duan, N., Manning, W.G., Morris, C.N. & Newhouse, J.P. (1983). A comparison of alternative models for the demand for medical care. Journal of Business and Economic Statistics, 1, 115-126.

References (Continued)

Ferrer, E. & McArdle, J.J. (2004). Alternative structural models for multivariate longitudinal data analysis. Structural Equation Modeling, 10, 493-524.

Klein, A. & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. Psychometrika, 65, 457-474.

Khoo, S.T. & Muthén, B. (2000). Longitudinal data on families: Growth modeling alternatives. In Multivariate applications in substance use research, J. Rose, L. Chassin, C. Presson & J. Sherman (eds), Hillsdale, N.J.: Erlbaum, pp. 43-78.

Miyazaki, Y. & Raudenbush, S.W. (2000). A test for linkage of multiple cohorts from an accelerated longitudinal design. Psychological Methods, 5, 44-63.

Moerbeek, M., Breukelen, G.J.P. & Berger, M.P.F. (2000). Design issues for experiments in multilevel populations. Journal of Educational and Behavioral Statistics, 25, 271-284.

Muthén, B. (1996). Growth modeling with binary responses. In A.V. Eye & C. Clogg (eds), Categorical variables in developmental research: methods of analysis (pp. 37-54). San Diego, CA: Academic Press.

255

References (Continued)

Muthén, B. (1997). Latent variable modeling with longitudinal and multilevel data. In A. Raftery (ed), Sociological Methodology (pp. 453-380). Boston: Blackwell Publishers.

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (ed.), Handbook of quantitative methodology for the social sciences (pp. 345-368). Newbury Park, CA: Sage Publications.

Muthén, B. & Curran, P. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. Psychological Methods, 2, 371-402.

Muthén, B. & Muthén, L. (2000). The development of heavy drinking and alcohol-related problems from ages 18 to 37 in a U.S. national sample. Journal of Studies on Alcohol, 61, 290-300.

Muthén, L.K. and Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. Structural Equation Modeling, 4, 599-620.

256

References (Continued)

Olsen, M.K. & Schafer, J.L. (2001). A two-part random effects model for semicontinuous longitudinal data. Journal of the American Statistical Association, 96, 730-745.

Raudenbush, S.W. (1997). Statistical analysis and optimal design for cluster randomized trials. Psychological Methods, 2, 173-185.

Raudenbush, S.W. & Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. Psychological Methods, 5, 199-213.

Roeder, K., Lynch, K.G., & Nagin, D.S. (1999). Modeling uncertainty in latent class membership: A case study in criminology. Journal of the American Statistical Association, 94, 766-776.

Satorra, A. & Saris, W. (1985). Power of the likelihood ratio test in covariance structure analysis. Psychometrika, 51, 83-90.