H2020 Marie Skłodowska-Curie innovative training network

Workshop:
Missing Data & Longitudinal Models in M*plus*

November 10, 2017

Prof. Wim Beyers
Ghent University, Belgium
wim.beyers@ugent.be

Who are you?

- What experience have you with Structural Equation Modeling (SEM) so far?
  - 12% no experience; 36% beginners; 40% occasional users; 12% experienced users
- Have you used M*plus* before?
  - 13% no; 87% yes
- Have you used M*plus* before to analyze longitudinal data?
  - 48% no; 52% yes
- What is your main learning objective today?
  - 21% curious about SEM and M*plus*; 74% learning SEM and M*plus* for longitudinal data; 10% my advisor requires me; 36% complex data; 10% latent variable stuff

With special thanks to...

Luc Goossens

Tod Little

Patrick Curran

Karl Jöreskog   Bengt & Linda Muthén
(no pics available)

Change is inevitable. Change is constant.

(Benjamin Disraeli)

Change is the nursery of music, joy, life, and eternity.

(John Donne)

When you're finished changing, you're finished.

(Benjamin Franklin)

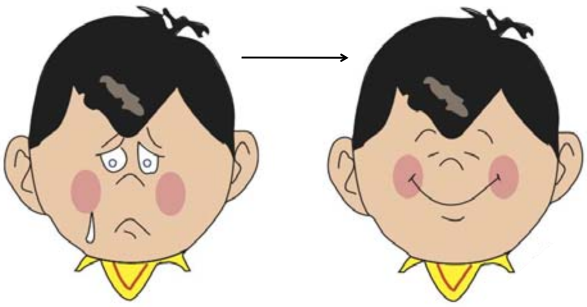
We are restless because of incessant change, but we would be frightened if change were stopped.

(Lyman Bryson)

Change is a measure of time.

(Edwin Way Teale)

sound



DEVELOPMENT

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

- General measurement and design issues
  - Time and intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Absolute change: the difference score
  - Absolute change+: repeated measures ANOVA
  - Relative change: autoregression
  - Relative change+: cross-lagged models
- Newer and better methods
  - With 2+ waves: LCM
  - With 3+ waves: LGCM + extension
  - With 3+ waves: LCGA
  - With 4+ waves: LGMM
  - With 3/4+ waves: LTA (mover-stayer)

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

The example DATA

- *N* = 405 adolescents + mothers
- From three cohorts
  - 1: M_{age} = 13 at Time 1
  - 2: M_{age} = 15 at Time 2
  - 3 :M_{age} = 17 at Time 3
- Measures:

Time 1:		
- **Gender**: 203 boys (1) and 202 girls (0) (A-report)		
- **Support** from mother (A-report)	1-6	3.00 (0.80)
- **Structure** by mother (A-report)	1-6	3.29 (0.96)
- **Shaming** by mother (A-report)	1-6	2.49 (0.82)
Time 1-2-3-4 (yearly measument):		
- **Antisocial behavior** (M-report)	0-10	1.66 - 1.83 - 2.03 - 2.06
- School **GPA on PE-class** (A-report)	0-10	2.52 - 4.08 - 5.00 - 5.77

- Missing data (coded 9999):
  - dropout and nonresponse from T3 onwards!
  - from 7% (ANTI-3) to 34% (GPA-4)
  - 11% overall

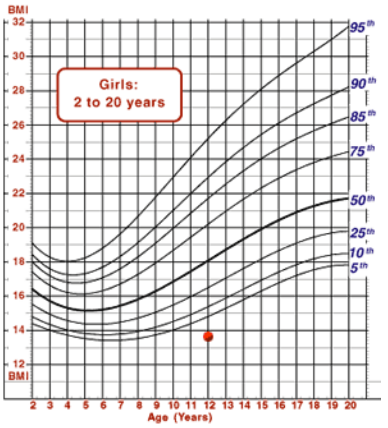
SEMDATA.SAV
SEMDATA.DAT
SEMDATA.XLS

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

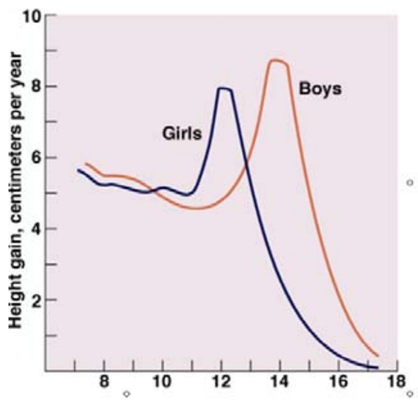
- **When measuring CHANGE, how can we define TIME?**
  - **Age** in years, months, days.
  - **Experiential time**: Amount of time something is experienced
    - Years of schooling (grade), length of relationship, amount of practice
    - Calibrate on beginning of event, measure time experienced
  - **Episodic time**: Time of onset of a life event
    - Age, toilet trained, driver license, puberty, birth of child, retirement
    - Early onset, on-time, late onset: used to classify or calibrate
    - Time since onset or time from normative or expected occurance.
- **What measurement Intervals should we take?**
  - **How fast** is the developmental process?
  - **Intervals** must be equal to or less than expected processes of change (e.g., schooling studies at half-year intervals)
  - **If too short**: too sensitive to measurement error
  - **If too long**: insensitive to change and variability in change

A great example article:
Reuter & Conger (1998). doi: 10.1037/0012-1649.34.6.1470

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !



- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !



Missing data

- What is the problem?
- Types of missing data?
- Preventing
- Curing
  - Bad methods
  - Questionable methods
  - Good methods

DROPOUT/MISSING VALUES

- Aim of statistic analyses =
  - based on sample data, draw conclusions about a population
  - estimate the population parameters as good as possible, based on the sample data
- What if we have incomplete data?
  - Can we still estimate correctly the population parameters?
  - Can we still draw correct conclusions about the population?

→ Missing values!
- occur in about very empirical study
- particularly in longitudinal research (dropout)!

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

TYPES of MISSING VALUES

- **Missing Completely at Random (MCAR)**
  - No relationship of missingness with non-observed (missing) data, and no relationship with observed data (= completely a-selective dropout)
- **Missing at Random (MAR) !**
  - No relationship of missingness with non-observed (missing) data, but possibly (and preferably) a relationship with observed data
- **Missing Not at Random (MNAR)**
  - A relationship of missingness with non-observed (missing) data (and possibly also with observed data

gpa4	gpa4_MCAR	test_MCAR	gpa4_MAR	test_MAR	gpa4_MNAR	test_MNAR
4,50	4,50	0	-999999	1	-999999,0	1
4,60	-999999,0	1	4,60	0	-999999,0	1
6,20	6,20	0	-999999	1	6,20	0
4,00	-999999,0	1	4,00	0	-999999,0	1
7,50	-999999,0	1	7,50	0	7,50	0
6,90	6,90	0	6,90	0	6,90	0
6,10	6,10	0	6,10	0	-999999,0	1
4,20	-999999,0	1	-999999	1	-999999,0	1
6,30	6,30	0	6,30	0	6,30	0
7,20	7,20	0	7,20	0	7,20	0
5,80	5,80	0	5,80	0	-999999,0	1
5,80	-999999,0	1	5,80	0	5,80	0

		gpa4	gpa1	support
test_ MCAR	Pearson Correlation	-,038	,017	-,077
	Sig. (2-tailed)	,576	,798	,252
	N	221	221	221

		support	gpa1	gpa4
test_ MAR	Pearson Correlation	**,266**	-,079	,009
	Sig. (2-tailed)	,000	,241	,889
	N	221	221	221

		support	gpa1	gpa4
test_ MNAR	Pearson Correlation	-,067	**-,188**	**-,399**
	Sig. (2-tailed)	,320	,005	,000
	N	221	221	221

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

DROPOUT

	Y (wave 2)			
X (wave1)	Compleet	MCAR	MAR	MNAR
130	101	101	**101**	101
145	155			
136	140	140	**140**	
146	134	134		134
111	129		**111**	129
134	124		**124**	124
153	112			112
137	122	122	**122**	122
118	118	118	**118**	118

- How to test for?
  - Not really possible. But…
  - MVA (SPSS) ➔
- References:
  - Little & Rubin (2002). *Statistical analysis with missing data.* Wiley
  - Schafer & Graham (2002): doi: 10.1037/1082-989X.7.2.147

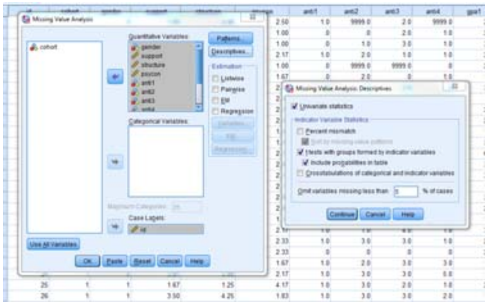


*Table with **Separate Variance t Test** in output, with in rows all variables with (+5%) missings, and in colums all variables in dataset. Cell contain a t-value (+ p) indicating whether or not missingness in the row variable is correlated significantly with the values of the column variable, and therefore is selective. Check patterns of significant t-values. If not clear pattern, MAR is very likely!*

*When selecting EM in the 'Estimation' -menu, **Little's MCAR test** is provided (= summary of t-tests above). If not significant: MCAR! If X^2/df (normed X^2) < 2: MAR.*

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

PREVENTING DROPOUT & MNAR!

- ***Dillman (1978)***
  - ✓ *Intensive follow-up and tracking of subjects*
  - ✓ *Repeated invitations to participation, reminders*
  - ✓ *Repeated sending of the measurements*
  - ✓ *Do everything to prevent large dropout!*
- ***Planned missingness***
  - ✓ *Do not measure all variables in all participants at all times.*
- ***Cohort-sequential design!!***
  - ✓ *Let new persons come in at each wave of the study,*
  - ✓ *This way you create different patterns of missingness, not only dropout!*

= Different ways to increase the chances of MAR or MCAR!

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

CURING!

- Purpose is NOT to fill in empty cells in the data!
- Purpose IS to estimate the population parameters as good as possible, using a sample with missing data!

Which methods can help us in this challenging task?

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

BAD ways to deal with missing data

- **List-wise Deletion**
  - Variances biased, means biased
  - *Acceptable only if power is not an issue and the incomplete data is MCAR*
- **Pair-wise Deletion**
  - *N* varies for each correlation
  - Variances biased, means biased
  - Sometimes estimation problems!
  - *Acceptable only if power is not an issue and the incomplete data is MCAR*
- **Sample-wise Mean Substitution**
  - For long time very popular method!
  - Variances reduced, correlations biased
  - *Never acceptable!*
- **Subject-wise Mean Substitution**
  - Depends on homogeneity of the items used
  - *Acceptable only if set of items is homogeneous and only few missings!*

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

QUESTIONABLE ways to deal with missing data

(single imputation methods)

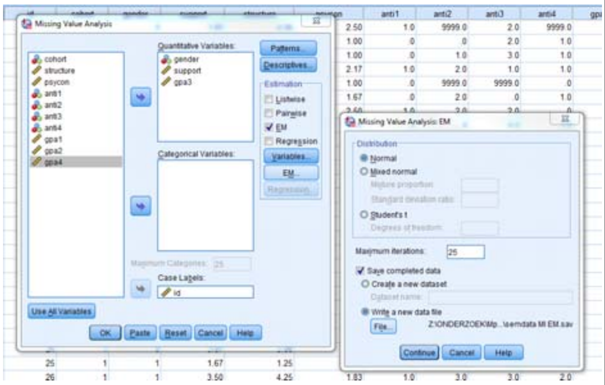
- **Regression Imputation**
  - All subjects with same values on IVs get the same estimated value on the DV.
  - Variances reduced
  - *Assumes MCAR*
- **Stochastic Regression Imputation**
  - Same as above but a random error component is added to reduce the loss in variance
  - *Still assumes MCAR*

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

GOOD ways to deal with missing data

- **But only if enough variables related to missingness are included in analysis (MAR), or missingness is MCAR**
- **EM Imputation**
  - Imputes the missing data values in an iterative way, starting with the E step
  - The E(stimation)-step is a stochastic regression-based imputation for each variable.
  - The M(aximization)-step is to calculate a complete covariance matrix based on the estimated values.
  - The E-step is repeated for each variable but the regression is now on the covariance matrix estimated in the previous M-step.
  - The EM-steps are repeated until the imputed estimates don't differ from one iteration to the other

EM in SPSS (onder assumptie MCAR)



EM in SPSS (onder assumptie MCAR)

Univariate Statistics

	N	Mean	Std. Deviation	Missing Count	Missing Percent	No. of Extremes[a] Low	No. of Extremes[a] High
gender	405	.50	.501	0	.0	0	0
support	405	3.0074	.80219	0	.0	22	7
gpa3	275	5.005	1.1610	130	32.1	0	2

a. Number of cases outside the range (Q1 - 1.5*IQR, Q3 + 1.5*IQR).

EM Means[a]

gender	support	gpa3
.50	3.0074	4.997

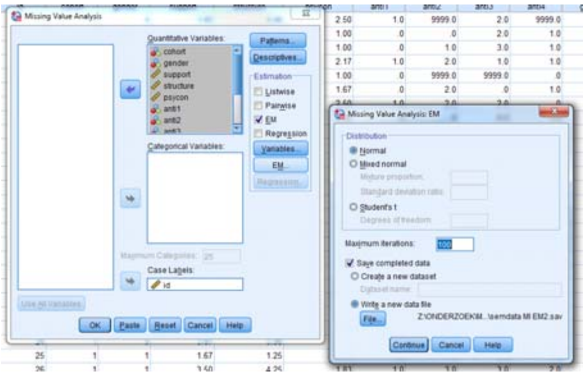
a. Little's MCAR test: Chi-Square = 2.512, DF = 2, Sig. = .285

EM Covariances[a]

	gender	support	gpa3
gender	.251		
support	-.026	.64351	
gpa3	-.028	.15530	1.3492

a. Little's MCAR test: Chi-Square = 2.512, DF = 2, Sig. = .285

EM in SPSS (onder assumptie MAR)



EM in SPSS (onder assumptie MAR)

Univariate Statistics

	N	Mean	Std. Deviation	Missing Count	Missing Percent	No. of Extremes[a] Low	No. of Extremes[a] High
cohort	405	2.00	.818	0	.0	0	0
gender	405	.50	.501	0	.0	0	0
support	405	3.0074	.80219	0	.0	22	7
structure	405	3.2928	.96177	0	.0	34	1
psycon	405	2.4859	.82896	0	.0	0	27
anti1	405	1.662	1.6581	0	.0	0	1
anti2	297	1.828	1.8979	108	26.7	0	4
anti3	374	2.027	2.0370	31	7.7	0	8
anti4	294	2.061	2.1520	111	27.4	0	6
gpa1	405	2.523	.9249	0	.0	0	8
gpa2	375	4.076	1.0836	30	7.4	0	3
gpa3	275	5.005	1.1610	130	32.1	0	2
gpa4	270	5.774	1.2480	135	33.3	0	0

a. Number of cases outside the range (Q1 - 1.5*IQR, Q3 + 1.5*IQR).

EM Means[a]

cohort	gender	support	structure	psycon	anti1	anti2	anti3	anti4	gpa1	gpa2	gpa3	gpa4
2.00	.50	3.0074	3.2928	2.4859	1.662	1.890	1.988	2.188	2.523	4.068	5.023	5.801

a. Little's MCAR test: Chi-Square = 330.094, DF = 266, Sig. = .005

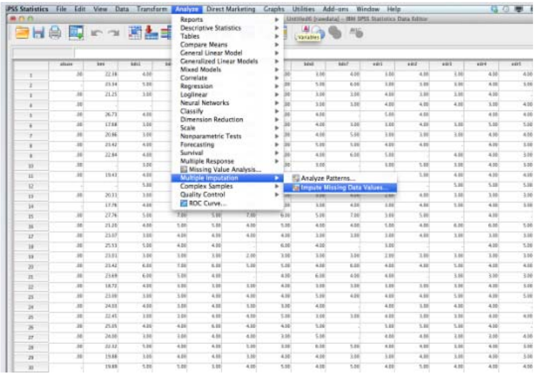
Ook beschikbaar in Lisrel (Prelis), SAS (PROC MI, met OUTEM= optie), e.a.

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

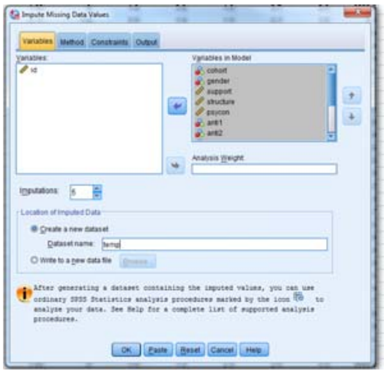
GOOD ways to deal with missing data

- **But only if enough variables related to missingness are included in analysis (MAR), or missingness is MCAR**
- **Multiple Imputation**
  - Estimate *N* (e.g., 5) datasets using the EM algorithm
  - Each dataset is based on a kind of resampling of the original sample (equivalent to a random selection of a different sample from the population)
  - Possible way 1:
    - Run the analyses *N* times
    - Summarize the results of these *N* analyses using the formulas of Rubin (1987)
  - Possible way 2:
    - Collapse the *N* samples to one dataset and do the analyses.

MI in SPSS

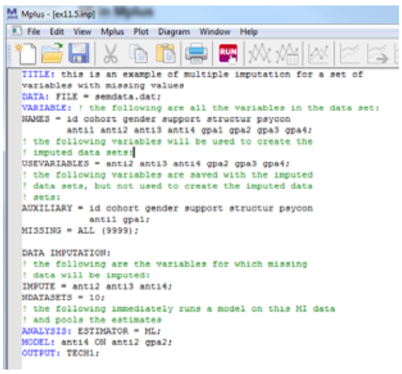


MI in SPSS



Further info at www.appliedmissingdata.com/spss-multiple-imputation.pdf

MI in Mplus



Further info at https://www.statmodel.com/download/Imputations7.pdf

- General measurement and design issues
  – Time & Intervals
  – Differential growth
  – Missing values
- Classic methods and disadvantages
  – Difference score
  – Repeated ANOVA
  – Autoregression
  – Cross-lagged models
- New(er) methods
  – LCM
  – LGCM
  – LCGA
  – LGMM
  – LTA
- Excercise !

GOOD ways to deal with missing data

- **But only if enough variables related to missingness are included in analysis (MAR), or missingness is MCAR, but even in cases of MNAR!**
- **Full Information Maximum Likelihood (FIML)**
  – Sufficient statistics (means, covariances) are estimated with the Expectation Maximization (EM) algorithm
  – Those estimates then serve as the start values for the Maximum Likelihood model estimation
  – Does not impute the missing values.
  – Can only be used when testing a SEM-model.
  – Available in Lisrel, AMOS, M*plus*, EQS, etc.

EXAMPLES comes with LGC and other models.

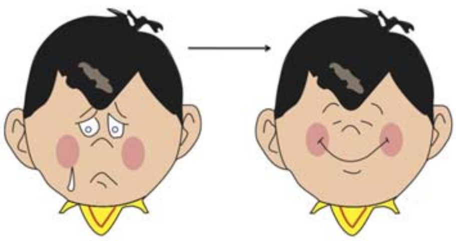
- General measurement and design issues
  – Time & Intervals
  – Differential growth
  – Missing values
- Classic methods and disadvantages
  – Difference score
  – Repeated ANOVA
  – Autoregression
  – Cross-lagged models
- New(er) methods
  – LCM
  – LGCM
  – LCGA
  – LGMM
  – LTA
- Excercise !

Missing values: Conclusions

- Missing values are part of every empirical study.
- Neglecting the missing data (listwise deletion) is a wrong approach.
- Different good methods are available to handle data that are MAR or MCAR, and give us correct population parameters!
- Even methods are available in case data are MNAR!

How to analyze change?

- Classic methods & disadvantages
- New(er) & better methods, using SEM



- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

$CHANGE_{14} = ANTI_4 – ANTI_1$

- Though many problems with it, still popular (e.g., intervention, pretest-posttest, or clinical studies)
- Most cited problem: Unreliability of the difference score
  - when the measures comprising the difference are only modestly reliable and positively correlated →
  - which is typically the case in longitudinal research!
- And therefore also: lack of valididity

+ Change is measured without taking level into account!

difference score.inp

```
DEFINE: change14 = anti4-anti1;
MODEL: change14 ON support;
```

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate      S.E.  Est./S.E.    P-Value

 CHANGE14 ON
    SUPPORT           -0.038     0.183     -0.207      0.836
```

r_2		0.60	0.70	0.80	0.90	1.00
				$r_1 = 0.60$		
r_{12}	0.60	0.00	0.13	0.25	0.38	0.50
	0.70	–0.33	–0.17	0.00	0.17	0.33
	0.80	–1.00	–0.75	–0.50	–0.25	0.00
	0.90	–	–	–	–	–1.00
	0.95	–	–	–	–	–
				$r_1 = 0.70$		
	0.60	0.13	0.25	0.38	0.50	0.63
	0.70	–0.17	0.00	0.17	0.33	0.50
	0.80	–0.75	–0.50	–0.25	0.00	0.25
	0.90	–	–	–	–1.00	–0.50
	0.95	–	–	–	–	–
				$r_1 = 0.80$		
	0.60	0.25	0.38	0.50	0.63	0.75
	0.70	0.00	0.17	0.33	0.50	0.67
	0.80	–0.50	–0.25	0.00	0.25	0.50
	0.90	–	–	–	–1.00	–0.50
	0.95	–	–	–	–	–
				$r_1 = 0.90$		
	0.60	0.38	0.50	0.63	0.75	0.88
	0.70	0.17	0.33	0.50	0.67	0.83
	0.80	–0.25	0.00	0.25	0.50	0.75
	0.90	–	–	–1.00	0.00	0.50
	0.95	–	–	–	–1.00	0.00
				$r_1 = 1.00$		
	0.60	0.50	0.63	0.75	0.88	1.00
	0.70	0.33	0.50	0.67	0.83	1.00
	0.80	0.00	0.25	0.50	0.75	1.00
	0.90	–	–1.00	0.00	0.50	1.00
	0.95	–	–	–1.00	0.00	1.00

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

$CHANGE_{14} = ANTI_4 – ANTI_1$

- Though many problems with it, still popular (e.g., intervention, pretest-posttest, or clinical studies)
- Most cited problem: Unreliability of the difference score
  - when the measures comprising the difference are only modestly reliable and positively correlated →
  - which is typically the case in longitudinal research!
- And therefore also: lack of valididity

+ Change is measured without taking level into account!

difference score.inp

```
DEFINE: change14 = anti4-anti1;
MODEL: change14 ON support;
```

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate      S.E.  Est./S.E.    P-Value

 CHANGE14 ON
    SUPPORT           -0.038     0.183     -0.207      0.836
```

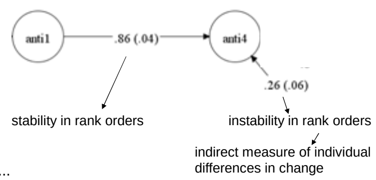
- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

CHANGE = effect of TIME in a repeated ANOVA

- SPSS output: R ANOVA.spv
- So, a good method
  - To describe and test an overall mean change function, and test for the form of it (linear, quadratic, cubic, etc.)
  - To test for the efffect of covariates on the change function (e.g., support): Time x Support interactions!
  - To test for the effect of between-subject factors on the change fuction (e.g., gender): Time x Gender interactions!
- But
  - Only tests mean change over time in the whole sample and not deviations from that mean change
  - And... group statistics (e.g., mean) represent everyone, and no one!
  - Equal intervals between measurements are necessary!
  - Change is an outcome of the repeated measures (time) and cannot be used as a predictor of outcomes.

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

- With observed variables: autoregression.inp
- With latent variables! autoregression latent.inp

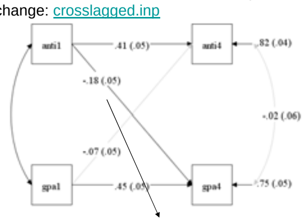


- But...
  - implicit assumption of decreasing correlations across time!
  - indifferent to the functional form of change!
  - only RELATIVE change in terms of rank order!
  - Interpretation?
    - Does ANTI1 predict (.86) an increase in ANTI4?
    - Increase only relative to others in the sample!
    - Even if everyone decreases!
  - No trajectories of individual change over time!

Extensions: mediation: autoregressionb.inp
moderation: autoregression-multigroup.inp

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

- A multivariate extension of the autoregressive model of change: crosslagged.inp



Antisocial behavior is 'causing' change in GPA!

- Interesting for examining direction of effects, by comparing cross-lagged coefficients (easily done in SEM)
- But
  - Same problem with interpretation of change as in AR model!
  - Also shares all other problems of the AR model!
  - \+ Statistical drawbacks:
    If X is (much) more stable than Y, then Y will have a stronger effect on X than vice versa!
    And, the less reliable X is, the less Y can explain it.

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

So,

What do we want exactly when we try to assess change?

RELIABLE estimates of change!
Estimates of ABSOLUTE change!
INDIVIDUAL estimates of change!

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 2 WAVES of data

Latent Change Models (LCM)
(McArdle & Nesselroade, 1994; Hertzog & Nesselroade, 2003)

- Change in **latent** variables, using (CFA)
- First step = Longitudinal invariant factor model (multiple indicators), resulting in reliable scores for Anti-1 and 4. ➔
- Second step = Restructuring Anti-1 and Anti-4 in latent Level and Change factors, using a very simple equation:

Anti-1 = 1xAnti-Level
Anti-4 = 1xAnti-Level + 1xAnti-Change

- As a consequence:

Anti-Change = Anti-4 – Anti-1
Anti-Level = Anti-1

a reliable difference score !

- lcm_ANTI.inp

Observed variables?

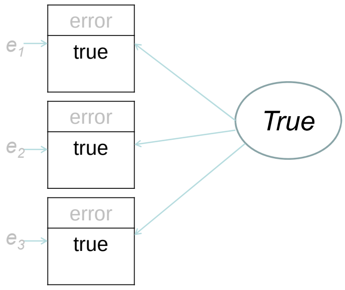
The problem of measurement error
DATA = MODEL + ERROR

- True variance: correlated
- Error variance: not correlated

➔Total covariance: underestimated!

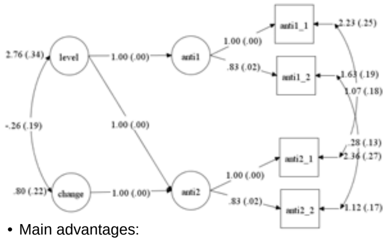
Latent variables!

Solution for measurement error!



SEM = analysis with latent variables!

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !



- Main advantages:
  - Only two waves of data needed!
  - Can be extended to multiple succesive change factors!
  - Reliable estimates of change!
  - Change is assessed as a latent factor!
    - With a mean: Mean change in the total sample!
    - With a variance: Individual differences in change!
    - That can be predicted and used as a predictor!
    - Latent factor scores can be estimated!
- But
  - Only linear change fuction!

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)

(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

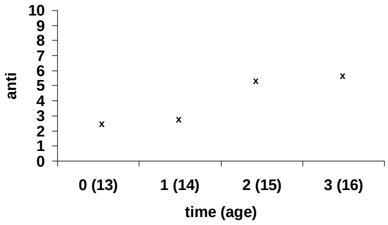
- Questions:
  - Does an individual characteristic (e.g., antisocial behavior) change over time?
  - Which trajectory is followed?
  - Interindividual differences?
- Step 1: Within-Person
  - Equation for every subject in the sample:
    anti = **intercept** + (**slope** x Time) + error (regression)
  - Growth can be non-linear too!
    anti = **intercept** + (**slope** x Time) + (**curve** x Time^2) + error
  - Assumption: Indivuals share the shape of the change function (e.g., linear), but can differ in the amount or rate of change (individual growth parameters: intercept, slope, etc.)
- Step 2: Between-Person
  - Means (fixed) & variances (random) of intercepts, slopes
  - Predictors of change (conditional growth models).

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)

(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

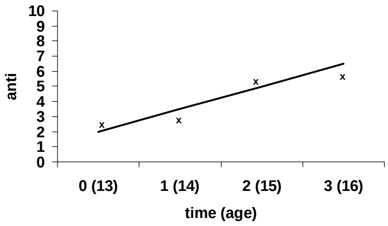


- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)

(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)



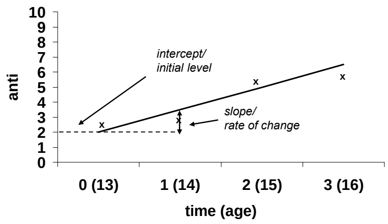
- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)

(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

anti = 2 + (1.5 x time) + error

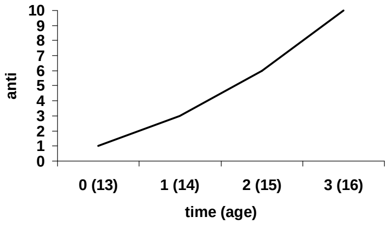


- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

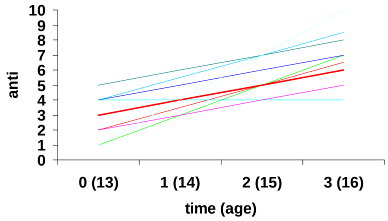
anti = 1 + (1.5 x time) + (.5 x time²) + error



- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

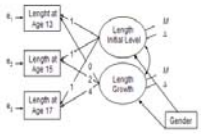
With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)



- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

Two Approaches to GCM

Multilevel	SEM: LGCM
- regression approach	- factor approach
- fixed & random effects	- means & variances
- flexible with missing values & individual-varying intervals	- not that flexible with individual-varying intervals
- less flexible in prediction	- very flexible in prediction
$length = b_{00} + b_{01}xgender + error_0 + (b_{10} + b_{11}xgender + error_1)*time + error$	
- little attention for model fit	- overload of fit indices

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

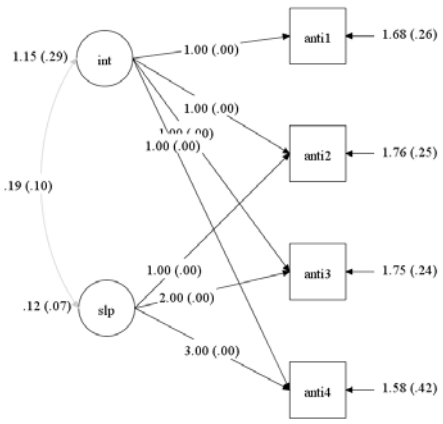
With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

PARAMETERS in the model

- Mean intercept / fixed effect intercept
  = Mean initial level of all individuals
- Variance intercept / random effect intercept
  = Interindividual differences in initial level
- Mean slope / fixed effect slope
  = Mean rate of growth across individuals
- Variance slope / random effect slope
  = Interindividual differences in rate of change

lgcm_anti.inp (with missing data, and FIML)



- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

PARAMETERS in the model

- Mean intercept / fixed effect intercept
  = Mean initial level of all individuals
- Variance intercept / random effect intercept
  = Interindividual differences in initial level
- Mean slope / fixed effect slope
  = Mean rate of growth across individuals
- Variance slope / random effect slope
  = Interindividual differences in rate of change

lgcm_anti.inp (with missing data, and FIML)

lgcm_anti + predictors.inp (gender and support as predictors)
lgcm_anti + predictors + interaction.inp (support X level anti)
lgcm anti - piecewise.inp (piecewise model with 2 slopes)

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

- Can be extended to a multvariate LGCM!
  lgcm_anti + gpa.inp
  - Correlated intercepts = cross-sectional association
  - Correlated intercept & slope = level of IV is predicting rates of change in DV!
  - Correlated slopes = common underlying growth in two constructs = change associated with change (causality?)
- But
  - Assumption: same shape of the growth function for all subjects; interindividual differences in change are modeled as deviations from that overall mean.

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

Extension: Cohort sequential design!

= mixing cross-sectional & longitudinal design
- remember: CS-design ➔ MAR
- Many longitudinal studies have multiple cohorts.
- Example: PhD of 4 years, but wanting to measure antisocial behavior from age 13 to 20 (8 years)!

	Time			
Cohort	**T1**	**T2**	**T3**	**T4**
1 (1985)	13	14	15	16
2 (1983)	15	16	17	18
3 (1981)	17	18	19	20

➔ Linking adjacent segments of change from different cohorts to estimate a common growth curve

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

Extension: Cohort sequential design!

Method 1: Multigroup modeling

	Time			
Cohort	**T1**	**T2**	**T3**	**T4**
1 (1985)	13	14	15	16
2 (1983)	15	16	17	18
3 (1981)	17	18	19	20

- Growth curve for each cohort, but:
  - Fix mean and variance of intercept and slope equal across cohorts
  - Fix int-slp correlation equal across cohorts
  - Adjust slope factor loadings of anti in line with cohort or birthyear
    - Cohort 1 (1985): 0 1 2 3
    - Cohort 2 (1983): 2 3 4 5
    - Cohort 3 (1981): 4 5 7 8

cohort-sequential growth model anti.inp (linear)
cohort-sequential growth model antib.inp (curvilinear)

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

Extension: Cohort sequential design!

Method 2: Data rearrangement (DATA COHORT)

	Age							
Cohort	13	14	15	16	17	18	19	20
1 (1985)	x	x	x	x	MCAR	MCAR	MCAR	MCAR
2 (1983)	MCAR	MCAR	x	x	x	x	MCAR	MCAR
3 (1981)	MCAR	MCAR	MCAR	MCAR	x	x	x	x

- Estimate 1 intercept and 1 slope, using all available data
  - Works out the time score based on birth and measurement year
  - Idea = rearrange our cohort & time data to age data
  - Then a growth curve for the complete age span!
  - Only works with continuous variables!

cohort-sequential growth model anti_2.inp (linear)
cohort-sequential growth model anti_2b.inp (curvilinear)
!? cohort-sequential growth model anti_2_piecewise.inp

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3/4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)
(Muthén, 2001, 2004)

and a special case of it:
Latent Class Growth Analysis (LCGA)
(Nagin, 1999, 2001)

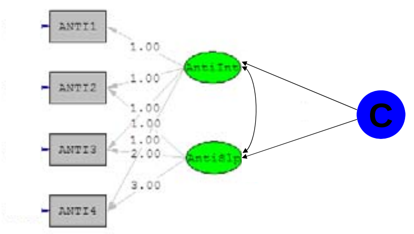
- Cfr. LGCM: heterogenity in the sample is captured by variation around a mean growth function, i.e., it is assumed that all individuals are drawn from the same population.
- LGMM
  - relaxes this single population assumption to allow for parameter differences across unobserved subpopulations
  - by using a combination of continuous and categorical latent variables (mixed)
  - by introducing a latent class variabele C, a trajectory class variable, representing *k* unobserved subpopulations in the sample (note that in LGCM *k* = 1)
  - by estimating a separate growth model for each of the latent classes
  - and estimating latent class conditional probabilities (membership)

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3/4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)
(Muthén, 2001, 2004)

and a special case of it:
Latent Class Growth Analysis (LCGA)
(Nagin, 1999, 2001)



- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3/4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)
Latent Class Growth Analysis (LCGA)

- **LGMM**: Class differences
  - Typically between mean intercepts between classes
  - Mean slopes in the classes
  - Variance in intercept and slope
  - Shape of the growth function!
  - Influence of covariates!
- **LGMM**: Extensions
  - Including covariates of change
  - Including outcomes that are predicted from growth
- **LCGA** (Nagin)
  - Very similar, only no variances within classes are estimated (therefore a semi-parametric approach)
  - Individuals within classes are treated as homogeneous

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3/4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)
Latent Class Growth Analysis (LCGA)

- Estimation: Using the EM algorithm
  - Estimation of each individual's probability of membership in each class (conditional probabilities)
  - Measures of fit and classification quality:
    - BIC
      - Small values correspond to a good model with a large likelihood and not too many parameters
      - Look at the big drops in BIC from one solution to another!
      - Sensitive to the number of classes!
      - Less sensitive to differences in growth shape between classes
    - LMR-LR test
      - Test of a solution with k-1 classes against a solution with k classes (e.g. 2 vs 1): low p-value indicates that solution k-1 should be rejected in favor of the solution with k classes.
    - Entropy
      - Measure of classification quality based on the individual class probabilities.
      - High values (closer to 1) indicate good classification.
- Examples with 2 classes:
  - lcga_anti2.inp (Nagin approach)
  - lgmm_anti2.inp (with equal variances across classes)
  - lgmm_anti2free.inp (with free variances across classes)

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)
Latent Class Growth Analysis (LCGA)

Antisocial data: fit statistics

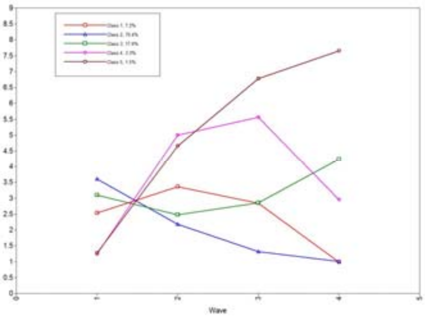
	LCGA			LGMM		
k	**BIC**	**entropy**	***p*LMRT**	**BIC**	**entropy**	***p*LMRT**
2l	4707.35	.97	.000	**4653.78**	**.95**	**.002**
3l	**4681.69**	**.80**	**.030**	4635.85	.79	.332
4l	4685.03	.81	.248	4625.88	.83	.154
2c	4533.67	.97	.000	**4506.76**	**.99**	**.000**
3c	4497.46	.82	.045	4487.26	.83	.411
4c	4476.20	.82	.014	4475.70	.85	.230
5c	**4462.39**	**.83**	**.006**			
6c	4460.83	.83	.426			

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)
Latent Class Growth Analysis (LCGA)

Antisocial data: Best solution: 5 classes, LCGA, curvilinear

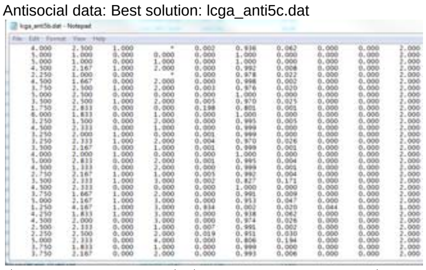


- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)

Antisocial data: Best solution: lcga_anti5c.dat



Observed scores

Individual estimated class probabilities

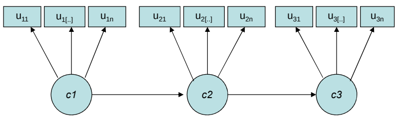
Classification

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3/4+ WAVES of data

Latent Transition Analysis (LTA)
(Kaplan, 2008; Meeus et al., 2010)

- A longitudinal extension of LCA or LPA
  - first, uses class-specific parameters (the continuous/ categorical observed multidimensional scores) as measurement parameters (~ LCA or cluster analysis)
    - invariant across time!
  - and, uses class probabilities as structural parameters to estimate the number of participants in each of the classes
  - then using latent transition probabilities to calculate patterns of stability and change over time in the movement or transition between classes (~ mover-stayer model)



➔ An ideal model for testing typological or person-oriented developmental theories!

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3/4+ WAVES of data

Latent Transition Analysis (LTA)
(Kaplan, 2008; Meeus et al., 2010)

- Example: data on parenting across three waves
  - 659 adolescents, measured 3 times: 13, 15 and 17 years
  - psychological control, support/warmth, firm control (continuous)
  - scores standardized (*z*) to ease interpretation
  - hypotheses: at least 4 latent classes (parenting styles) an gradually better parenting from 13 (peak of puberty years; reactive parenting) to 17 years
  - estimated using Robust Maximum Likelihood estimation
  - models with different # latent classes fitted and compared
  - decision based on BIC, entropy, and interpretability

lta parenting4.inp

k	BIC	entropy
3	15308.59	.93
4	15016.72	.92
5	14982.35	.90
6	15324.43	.92

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadv…
  - Differen…
  - Repeat…
  - Autore…
  - Cross-l… models
- New(er) m…
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3/4+ WAVES of data

Latent Transition Analysis (LTA)
(Kaplan, 2008; Meeus et al., 2010)

- Selected output *k* = 5 solution
  - Measurement part + latent class probabilities

	Psychological Control	Support	Firm Control	13	15	17
Neglecting	-0.38	-0.11	-0.54	49%	29%	30%
Permissive	-0.67	1.21	0.64	24%	18%	19%
Rejecting	0.34	-1.17	-1.19	10%	8%	9%
Controlling	1.00	-0.62	-0.07	15%	16%	16%
Democratic	-1.66	1.36	1.23	1%	28%	26%

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic met… and disadva…
  - Differenc…
  - Repeate…
  - Autoregr…
  - Cross-la… models
- New(er) me…
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

With 3/4+ WAVES of data

Latent Transition Analysis (LTA)
(Kaplan, 2008; Meeus et al., 2010)

- Selected output *k* = 5 solution
  - Transition probabilities

	Transition Probabilities into latent classes at T+1				
	Neglecting	Permissive	Rejecting	Controlling	Democratic
Neglecting 1	**.574**	.020	.000	.006	.400
Permissive 1	.000	**.728**	.000	.008	.264
Rejecting 1	.090	.000	**.813**	.000	.097
Controlling 1	.000	.000	.000	**.974**	.026
Democratic 1	.000	.000	.000	.272	**.728**
Neglecting 2	**.971**	.003	.000	.000	.026
Permissive 2	.000	**.894**	.000	.004	.102
Rejecting 2	.027	.000	**.953**	.000	.021
Controlling 2	.007	.000	.006	**.870**	.117
Democratic 2	.033	.094	.026	.086	**.761**

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

To conclude

- Since two decades, we have interesting new methods to analyze change and development!
- M*plus* provides a powerful tool to analyze change and development, and is constantly improving!
- And, as said before…

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
  - LTA
- Excercise !

1. Check the direction of effects (using a cross-lagged model) between GPA and antisocial behavior, using data of Times 1 and 3. Check wether results are the same for boys and girls.
   DATA are SEMDATA.DAT; software = Mplus 8.0
2. Estimate a LGCM of GPA, using the FIML approach for missing data. Check models with linear and curvilinear change. Interprete the parameters that are found.
   DATA are SEMDATA.DAT; software = Mplus 8.0
3. Find the optimal LCGA/LGMM solution of the GPA data. Explain why this solution was chosen and interprete the different classes.
   DATA are SEMDATA.DAT; software = Mplus 7.3
4. Evaluate the effect of mother support on development of GPA, using a conditional growth model and FIML
   DATA are SEMDATA.DAT; software = Mplus 8.0
5. Test the multivariate LGCM of ANTI & GPA, using the FIML approach for missing data. Search for the best fitting model. Interprete the all the estimated parameters in this model.

- *Setup and do the analyses using Mplus!*
- *Ask for help while doing the analyses!*
- *Present the results to the audience, using a single or two slides and explain the effects in words!*

Informative websites

- www.statmodel.com: thé M*plus* site!
- http://davidakenny.net/cm/causalm.htm: great SEM page!
- https://stats.idre.ucla.edu/mplus/: online examples and videos on M*plus*
- http://users.ugent.be/~wbeyers/workshop/index.html: the website for this workshop!

Literature

An introduction to SEM

- Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: The Guilford Press.
- Hox, J. J., & Bechger, T. M. (1998). An introduction to structural equation modeling. *Family Science Review, 11*, 354-373.

Defining time and choosing intervals

- Rueter, M. A., & Conger, R. D. (1998). Reciprocal influences between parenting and adolescent problem-solving behavior. *Developmental Psychology, 34,* 1470-1482.

Missing data and how to handle them

- Little, T. D., Lindenberger, U., & Maier, H. (2000). Selectivity and generalizability in longitudinal research: On the effects of continuers and dropouts. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 187-200). Mahwah, NJ: Erlbaum.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147-177.

Literature

Latent Change Model (LCM)

- McArdle, J. J., & Nesselroade, J. R. (1994). Structuring data to study development and change. In S. H. Cohen & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological innovations* (pp. 223-268). Hillsdale, NJ: Erlbaum.
- Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: An overview of methodological issues. *Psychology and Aging, 18,* 639-657.

Latent Growth Curve modeling (LGCM)

- Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116,* 363-381.
- 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* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavior Therapy, 35,* 333.
- McArdle, J. J., & Nesselroade, J. R. (2002). Growth curve analysis in contemporary psychological research. In J. Schinka & W. Velicer (Eds.), *Comprehensive handbook of psychology (Vol. 2): Research methods in psychology* (pp. 447-480). New York: Wiley.

Literature

Latent Class Growth Analysis (LCGA)

- Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric group-based approach. *Psychological Methods, 4,* 139-157.
- Nagin, D. S. (2001). Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods, 6,* 18-34.

Latent Growth Mixture Modeling (LGMM)

- Muthen, B., & Muthen, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research, 24,* 882-891.
- Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 291-322). Washington, DC: American Psychological Association.
- 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.

Literature

Longitudinal measurement invariance

- Bishop, J., Geiser, C., & Cole, D. A. (2014). Modeling latent growth with multiple indicators: A comparison of three approaches. *Psychological Methods*. doi: 10.1037/met0000018
- Coertjens, L., Donche, V., De Maeyer, S., Vanthournout, G., & Van Petegem, P. (2012). Longitudinal measurement invariance of learning strategy scales: Are we using the same ruler at each wave? *Journal of Psychoeducational Assessment, 30*, 577-587. doi: 10.1177/0734282912438844
- Wu, A., Liu, Y., Gadermann, A. M., & Zumbo, B. D. (2010). Multiple-indicator multilevel growth model: A solution to multiple methodological challenges in longitudinal studies. *Social Indicators Research, 97*, 123-142. doi: 10.1007/s11205-009-9496-8
- Metha, P. D., Neale, M. C., & Flay, B. R. (2004). Squeezing interval change from ordinal panel data: Latent growth curves with ordinal outcomes. *Psychological Methods, 9*, 301-333. doi: 10.1037/1082-989X.9.3.301
- 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* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Literature

Latent Transition Analysis

- Collins, L. M., & Lanza, S. T. (2009). *Latent class and latent transition analysis: With applications in the social, behavioral and health sciences.* Hoboken, NJ: John Wiley & Sons.
- Graham, J. W., Collins, L. M., Wugalter, S. E., Chung, N. K., & Hansen, W. B. (1991). Modeling transitions in latent stage-sequential processes: A substance use prevention example. *Journal of Consulting and Clinical Psychology, 59,* 48-57.
- Kaplan, D. (2008). An overview of Markov chain methods for the study of stage-sequential developmental processes. *Developmental Psychology, 44,* 457–467.
- Meeus, W., van de Schoot, R., Keijsers, L., Schwartz, S., & Branje, S. (2010). On the progression and stability of adolescent identity formation: A five-wave longitudinal study in early-to-middle and middle-to-late adolescence. *Child Development, 81,* 1565-1581.
- Velicer, W. F., Martin, R. A., & Collins, L. M. (1996). Latent transition analysis for longitudinal data. *Addiction, 91,* 197-210.

Further literature

- **Model fit**
  - Hu, L.-T., & Bentler, P. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural Equation Modeling. Concepts, Issues, and Applications* (pp. 76-99). London: Sage.
  - Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.
  - Hu, L.-t., & Bentler, P.-M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological-Methods, 3*, 424-453.
  - Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organizational Research Methods, 3, 4–70.

- **MLR (and related Chi-square test)**
  - Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*, 507-514.
  - Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399-419). Thousand Oaks, CA: Sage Publications.

- **Reporting about SEM**
  - Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 158-176). Thousand Oaks, CA: Sage Publications.
  - McDonald, R. P., & Ringo Ho, M.-H. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*, 64.
  - Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics. Fourth edition*. Needham Heigths, MA: Allyn & Bacon.