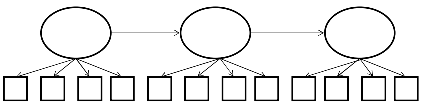
University of Ulster at Magee, Friday 15th June 2012

Latent Transition Analysis

Dr Oliver Perra
Institute of Child Care Research
Queen's University Belfast

email : o.perra@qub.ac.uk

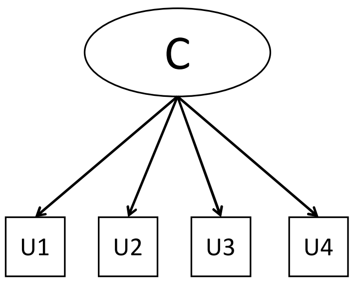


Overview of latent class and latent transition models

Latent Class Analysis

- Part of "mixture" models
  - Assumption: unobserved *heterogeneity* in the population
- Given a set of categorical indicators, individuals can be divided into subgroups (latent classes) based on an *unobserved* construct (e.g. Disordered v. Non-Disordered)
- Latent classes are *mutually exclusive* and *exhaustive*
- Individuals in each class are supposed to behave in the same manner (similar parameter values)
  - Intra-group homogeneity
  - Inter-group heterogeneity
- Latent classes describe the associations among the observed categorical variables

Latent Class Analysis



Latent Class Analysis

- Parameters of the model are:
  - Probability of being in each class (membership)
  - Probability of fulfilling each criterion (e.g. endorsing an item) given class membership
    - E.g. Probability of providing correct response to a test given membership in the “Mastery” latent class.
  - Furthermore, the model provides probability of being in each class for each individual (posterior probability)

Latent Class Analysis

- Categorical indicators : a b c d
- Latent class: x
- $P_{abcdx} = p_x * p_{a|x} * p_{b|x} * p_{c|x} * p_{c|x}$

Sum p_x = Sum $p_{a|x}$ = Sum $p_{b|x}$ = Sum $p_{c|x}$ =
= Sum $p_{c|x}$ = 1

Assumption of conditional independence

- Manifest variables are independent given latent class
  - Put it another way: the observed relationship between manifest variables (answers to questions, success in test items, etc.) <u>is attributable to a common factor</u>

If X is the latent variable with different classes, A and B are categorical outcomes:

$P_{abx} = P(a=1|x=1) * P(b=1|x=1) * P(x=1)$

with a=1 → pass in a ; b=1→ pass in b; x =1 → mastery

The probability any mastery respondent passes both tests (P of 111) is equal to the product of their estimated conditional probability of passing test a and estimated probability of passing test b

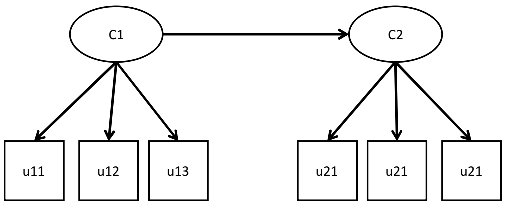
- Some variables are unlikely to be conditionally independent (e.g. related symptoms).

LC: Model Estimation

- Iterative maximum-likelihood estimation approaches
- Begin with a set of “start values” and proceed with re-estimation iterations until a criterion is met (usually convergence: each iteration in parameter estimation approaches some predesigned small change)
- Expectation-Maximization algorithm : robust with respect to initial start values
- Problems of local optima : convergence to local solutions

Latent Transition Analysis (LTA)

- Longitudinal extension of latent class models



LTA v. Growth models

- In growth models the focus is on average rate of change over time and the growth process is assumed to be continually occurring at the same rate
- In LTA, change can be discontinuous : movement through discrete categories or stages
  - "Qualitative growth": changes not restricted to quantitative growth
  - Different people may take different paths

Examples of LTA applications - I

- Stages of change for smoking cessation (Martin, Velicer & Fava, 1997)
  - 4 stages:
    - Pre-contemplation
    - Contemplation
    - Action
    - Maintenance
  - Movement was not always linear (forthsliders and backsliders; 2-stage progressions)
  - Probability of forthsliding> backsliding
  - Greater probability to move to adjacent stages than 2-stage progression

Examples of LTA applications - II

- LTA used to evaluate the stability of Typically Developing v. Reading Disability classification across grades 1 to 4 (Compton et al., 2008)
  - Results suggested a fair amount of stability
  - Results also suggested the importance of including a word reading fluency item in the model estimation, particularly after grade 1: inclusion of this indicator reduced "false negatives"

Examples of LTA applications - III

GRAHAM, COLLINS, WUGALTER, CHUNG, HANSEN

Table 1
Latent Transition Probability Estimates for Model 2 (Entire Sample)

	Time 2 latent status					
Time 1 latent status	1	2	3	4	5	6
1. No use	.704	.188	.013	.050	.038	.006
2. Alcohol only	.0	.783	.0	.157	.020	.039
3. Tobacco only	.0	.0	.622	.267	.000	.111
4. Alcohol + tobacco	.0	.0	.0	.793	.000	.207
5. Alcohol + tobacco + drunk	.0	.0	.0	.0	.628	.372
6. Advanced use	.0	.0	.0	.0	.043	.957

Note. Each value shown is the estimated probability of membership in the "Latent status" column at Time 2 conditional upon membership in the "Latent status" row at Time 1. A value of .0 indicates that the parameter was fixed at zero; .000 indicates the value was estimated at zero.

A model of substance use onset including both alcohol and tobacco use as possible starting points fit better than a model that included alcohol use as the only starting point.
Participants who had tried tobacco hut not alcohol in 7th grade seemed to be on an accelerated onset trajectory.

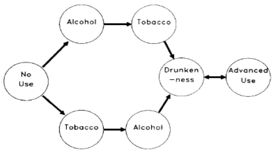


Figure 1. A model of substance use onset.

Latent Transition Analysis (LTA)

- Allows specification of number of stages in a model
- Transitions consistent with model, e.g. Cannabis lifetime use → no use (?)
- Estimate prevalence of class membership at first time of measurement
- Incidence of class transitions
- Probability of particular item responses conditional on stage membership

Example of LTA (Nylund, 2007)

- A longitudinal study of over 1,500 middle-school students in US
- Students completed 6-item Peer Victimization Scale in grade 6, 7 and 8 (e.g. being picked on, laughed at, hit and pushed around, etc.)
- Responses to items dichotomised

Note that is not necessary that items have the same number of response categories

Example of LTA (Nylund, 2007)

	Grade 6	Grade 7	Grade 8
Called bad names	37%	25%	20%
Talked about	33%	26%	23%
Picked on	28%	19%	14%
Hit and pushed	21%	15%	12%
Things taken/messed up	29%	19%	15%
Laughed at	30%	20%	18%

Proportion endorsed for 6 binary items by grade

3 classes in Grade 6

	Victimised (19%)	**Sometimes-victimised (29%)**	**Non-Victimised (52%)**
Called bad names	.85	.58	.08
Talked about	.74	.51	.07
Picked on	.81	.39	.03
Hit and pushed	.76	.17	.03
Things taken/messed up	.79	.31	.09
Laughed at	.86	.36	.06

Conditional item response probability (probability of endorsement) by latent class

3 classes in Grade 7

	Victimised (13%)	**Sometimes-victimised (20%)**	**Non-Victimised (67%)**
Called bad names	.76	.59	.05
Talked about	.69	.53	.09
Picked on	.82	.26	.03
Hit and pushed	.68	.12	.05
Things taken/messed up	.68	.29	.05
Laughed at	.75	.38	.03

Conditional item response probability (probability of endorsement) by latent class

Transition probabilities grade 6 to 7 (LTA model)

7th Grade			
	Victimised	*Sometimes-victm.*	*Non-victm.*
6th Grade			
Victimised	.42	.41	.17
Sometimes-victm.	.05	.48	.47
Non-victm.	.01	.10	.89

N of classes at each occasion

- Many LTA models will consider the same number of classes at each occasion
- However, there may be cases where the number of latent classes may be different across time:
  - e.g. : 2 classes of exposure to violence may be sufficient in early adolescence, but 5 classes may be necessary to describe heterogeneity of violence exposure in late adolescence (more *diversity* in phenomenon)
- The interpretation of each class is a function of its item response probabilities (see next)

LTA parameters

- Item response probabilities (some refer to these as rho, ρ)
  - Probability of endorsing a category of response at time t (e.g.: 1, 2,..., t) given latent status membership at time t
  - These allow to interpret latent statuses (e.g. Higher probability of endorsing victimisation items → victimised class)
  - One for each time-status-item combination
    - Constraints can be assumed and tested: E.g. identical across measurement occasions (measurement invariance)?

LTA Parameters (ctd.)

- Latent class prevalence at time t: probability of being in latent class a at time t
- Some (e.g. Collins) refer to these parameters as delta δ (with a subscript for class and time, e.g. δ_{at})
  - E.g. In Nylund's study, prevalence of "victimised" class in grade 6 was 19% , thus δ_{v6} = . 19)

LTA Parameters (ctd.)

- Transition probabilities: Probability of class *b* membership at time 2 given membership to class *a* at time 1

  E.g. Probability of being in "victimised" class in grade 7 given membership to "non-victimised" in grade 6 (= .01)
- Usually referred to as tau τ and underscript indicating class membership at time t given membership at time 1 , e.g.:
  - $\tau_{b|a}$
  - $\tau_{1|3}$

  The latter indicates probability of being in class 1 at time 2 given (|) membership in class 3 at time 1

LTA Parameters (ctd.)

- τ parameters arranged in a transition probability matrix like this:

Time 1		Time 2	
	Class 1	Class 2	Class 3
Class 1	$\tau_{1|1}$	$\tau_{2|1}$	$\tau_{3|1}$
Class 2	$\tau_{1|2}$	$\tau_{2|2}$	$\tau_{3|2}$
Class 3	$\tau_{1|3}$	$\tau_{2|3}$	$\tau_{3|3}$

LTA Parameters (ctd.)

- Restrictions and constraints can also be imposed on transition parameters:

Time 1		Time 2	
	Victimised	Sometimes victm.	Non-Victimis.
Victimised	$\tau_{1\|1}$	$\tau_{2\|1}$	$\tau_{3\|1}$
Sometimes-victm.	$\tau_{1\|2}$	$\tau_{2\|2}$	$\tau_{3\|2}$
Non-victimised	$\boldsymbol{\tau_{1\|3}}$	$\tau_{2\|3}$	$\tau_{3\|3}$

- E.g. $\tau_{1|3} = 0$ → fixing probability of transitioning from non-victimised to victimised to 0
- *Absorbing class*: one that has a zero probability of exiting : $\tau_{1|1} = 1$ → 100% probability of being victimised at time 2 if victimised at time 1

LTA Parameters (ctd.)

- Other restrictions and constraints can be imposed on transition parameters:
  - Transition probabilities to be the same across time points:

    E.g. :The probability of transitioning from victimised to non-victimised between grades 6 and 7 the same as between grades 7 and 8

$\tau_{n7|v6} = \tau_{n8|v7}$

Change process assumed *stationary*: individuals are transitioning between classes with the same probabilities across time points

Summary so far

- Latent Class Analysis: fundamentally a measurement model
- Latent Transition Analysis: measurement and structural model. Describes qualitative change across measurements points (2 or more)
- LTA parameters:
  - Conditional item response probabilities ρ (measurement model)
  - Prevalence of latent statuses at each time point δ
  - Transition probabilities between two time points τ

LTA Steps

- Step 1: Investigate measurement model alternatives for each time point (*separately* for each time point)
- Step 2: Test for measurement invariance across time
- Step 3: Explore specification of the latent transition model without covariates
  - Investigate transition probability specifications
- Step 4: Include covariates in LTA model
- Step 5: Include distal outcomes

Step 1

Investigate measurement model alternatives

Step 1: Investigate measurement model alternatives

- Decision does not involve only statistical indicators of fit to data, but also interpretability of results and aims of the study.
- "The choice of factor analysis or LCA is a matter of which model is most useful in practice. It cannot be determined statistically, because data that have been generated by an m-dimensional factor analysis model can be fit perfectly by a latent class model with m+1 classes"
  - Muthén & Muthén (2000). Integrating person-centred and variable-centred analyses. *Alcoholism: Clinical and Experimental Res*.
- If the aim is *diagnosis* or *categorisation*, then use LCA (avoids the use of arbitrary cut-points or ad-hoc rules)

Step 1: investigate measurement model

- 1.1 if LCA → determine number of classes at each time point
- 1.2 Test restrictions on item response parameters
- 1.3 Validate results including covariates

Determining n of classes

- The standard procedure is to test a series of LC models : from 2-class to n-class
- No accepted single indicator to decide on the appropriate number of classes:
  - Although log-likelihood value is provided in estimation, this cannot be used to compare models with different n classes (e.g. 2- vs. 3-class) via Likelihood Ratio Test (LRT)

Determining n of classes (ctd.)

- Consider χ^2 and likelihood ratio chi-square test G^2
- Use information criteria (the lower the value the better the fit)
  - AIC penalises by number of parameters → preference for "simpler" models
  - BIC penalises by number of parameters and sample size
  - *Mplus* provides the sample-size adjusted BIC

LC statistics and information criteria

- χ^2 = Sum [(observed f. – expected f.) ^{2} / exp. f.]
- G^2 = 2 sum [obs. f. * ln (obs. f. / exp. f.)]
- AIC = G^2 – 2 df
- BIC = G^2 – df * [ln(N)]
- Sample-size adjusted BIC : N * = (N + 2) / 24)

Practical

- Introduction to Mplus language
- Estimation of LC model using Mplus
- Imposing constraints on measurement parameters using Mplus

Intro to LCA in Mplus

- Mplus uses:
  - *input* files to instruct how to read separate data file, to specify type of analysis and model and to request information in *output* file and other functions (additional files, plots, etc.).
  - Results are reported in the *output* file
  - It can also provide (under request in input) files that can be used to create graphs
  - It can provide (under request) files with model parameters

```
TITLE: an example of LCA
DATA:  FILE = chap11.dat;
VARIABLE: NAMES are a b c d e male female ;
        MISSING = ALL (-9999);
        USEVARIABLES = a b c d ;
        CATEGORICAL = a b c d ;
        CLASSES = x(2);
```

One can use "= " or "is/are" . E.g. File is....;

; is used to separate arguments (NOT optional)

Names provides names for variables in dataset

Usevariables are the variables we will be using in the analyses

Missing values in dataset are indicated by -9999. If not provided, the program consider -9999 as a legitimate value for variable(s)

This command requests estimation of a latent categorical variable (x) with 2 classes (x1 and x2). The categorical outcomes (indicators) are to be regressed on the latent variable. x(3) would indicate a 3-class model.

The variables indicated are to be considered ordered categorical variables. The command is used only for outcome variables in model (ie do not indicate male as categorical because even if used as covariate in model, it is NOT an outcome). NOMINAL = ... would indicate variables with 2 or more categories but with no intrinsic order (e.g. political party preference)

Intro to Mplus (ctd.)

- Latent classes are indicated under the "Variable" command because they are effectively considered (unobserved) variables in the dataset.
- Unless specified otherwise (more about this later...) the outcome variables and the other variables in "usevariables" are regressed onto the latent categorical variable

Intro to Mplus (ctd.)

The other essential bit to conduct LCA:

```
ANALYSIS: TYPE = MIXTURE;
          STARTS = 100 10;
          STITERATIONS = 20;
```

TYPE: MIXTURE in the ANALYSIS command invokes a mixture model algorithm (necessary for "mixture" models such as LCA, LTA, LCGA, GMM, etc).
The default estimator for this type of analysis is Maximum Likelihood with robust standard errors (MLR in Mplus). [This can be changed with command ESTIMATOR =...]

By default, ML optimization in two stages: initial one with 10 random sets of starting values; 2 optimisations with highest likelihoods used as starting values in the final stage. This is what would happen if you do not provide the STARTS command in ANALYSIS.
In the example above, 100 random sets are used, with 10 values with highest likelihood used in the final stage. <u>Increase n starts is often necessary for the model to converge</u>.

The max number of iterations allowed in initial stage is 10 by default, but can be increased (in the example STITERATIONS = 20) for more thorough investigation of multiple solutions

Intro to Mplus (ctd.)

```
MODEL:
%OVERALL%
!this is the part of the model common for all
!classes
[x#1];

%x#1%
[a$1-d$1] (1-5);

%x#2%
[a$1-d$1] (6-10);
```

The other important part is the **MODEL: command.**
It is not necessary to specify a model if you are conducting a simple LCA, with no covariates and no restrictions on parameters (omit the MODEL command completely in this case).
%overall% describes the part of the model that is common to ALL latent classes (e.g. latent class affiliation is regressed on covariate x).
%x#1% is used to specify the part of the model that differs for class 1,
%x#2% specifies the part of the model specific to class 2
... And so on (if more than 2 classes)

Intro to Mplus (ctd).

- Mplus thinks of categorical variables (binary or with more categories) as continuous latent variables that are "cut" into different categories.
- The points in which to "cut" the underlying latent variable are called thresholds.
- If we take a binary variable:

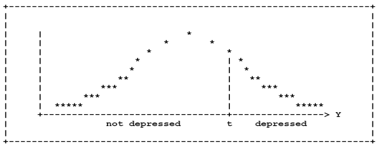


Figure 1 (draft). Latent continuous variable (depression severity, Y); and discretizing threshold (t).

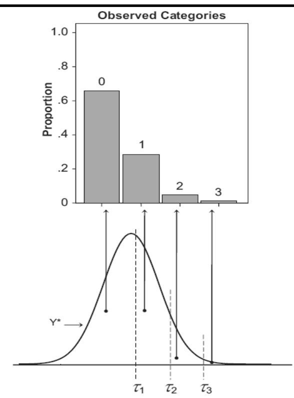


Figure 1. Pictorial representation of how thresholds (τs) in an underlying, continuous latent response variable (Y*) correspond to the categories in an ordinal variable.

From :Feldman, Masyn & Conger (2009). New approaches to studying problem behaviors. *Developmental Psychology, 45(3)*, 652-676.

Intro to Mplus (ctd.)

- We are considering a model with 4 binary indicators:
  - a b c d
- Categories of response are "No" (category 1) and "Yes" (category 2)
- Indicators have one threshold each [a$1 b$1 c$1 d$1]; the threshold represents the point in which the underlying distribution is cut to create the two response categories
- We want to fit a two-class model: x (latent class) → x#1 (latent class 1) x#2 (latent class 2)
  - In the same manner as for observed categorical variables, we need to estimate a threshold for x → [x#1] that cuts the distribution into two categories

Intro to Mplus(ctd).

- Number of thresholds = n of categories -1 (a binary variable needs only one cut to create two categories).
- Thresholds are indicated by the name of the variable followed by **$** and the progressive number: all within **square brackets**.
- A variable a with 3 categories (e.g. not yet, sometimes, often) would have 2 thresholds:
  - [a$1 ; a$2]
- The asterisk * is used to free a parameter. If followed by a number, it assigns a starting value to the thresholds;
- @ is used to *fix* the value of a thresholds to some pre-defined value (e.g. -15)

Thresholds are in a *logit* scale:

The LCA model with p observed binary items u, has a categorical latent variable C with K classes (C = k; k = 1, 2, ..., K). The marginal item probability for item $u_j = 1$ (j = 1, 2, ..., p) is given by:

$P(u_j = 1) = \text{sum } P(C=k) * P(u_j = 1 \mid C = k)$

where the conditional item probability in a given class is defined by :

$\mathbf{P(u_j = j \mid C = k) = 1 / [\,1 + \exp(- v_{jk})\,]}$

where the v_{jk} is the logit for each of the u_js for each of the latent classes, k

For example, if we want to constrain P(a=1|c=1) = .05, we fix logit threshold *v(jk)* to -2.95 → [a$1@-2.95] ;
A threshold = 0 will make P(a=1|c=1) = .50 ...and so on

Intro to Mplus (ctd.)

MODEL:
%OVERALL%
!this is the part of the model common for all
!classes

[x#1];

%x#1%
[a$1-d$1] (1-4);

%x#2%
[a$1-d$1] (5-8);

This means that the *threshold* for the latent categorical variable is being estimated: where do you cut the latent variable distribution to form two latent classes, as specified by CLASSES = x(2); estimates prob of being in x1 class

The parentheses after the indicators' thresholds assign a name (if a letter is used) or posit a constraint (if a number used) to each of these parameters.

If we wanted the thresholds of a, b, c and d to be the same for x1 and x2, we would have written:

%x#1%
[a$1-d$1] (1-4);
%x#2%
[a$1-d$1] (1-4);

By doing this, we are making the thresholds, therefore the item response probabilities, the same for x=1 and x=2

Constraints on measurement model: Parallel indicators

MODEL:
%OVERALL%
!this is the part of the model common for all
!classes

[x#1];

%x#1%

[a$1-d$1] (1);

%x#2%

[a$1-d$1] (2);

In this example, the thresholds for the latent class estimators (a to d: a, b, c, d) are equal to each other within each class, but not equal across classes → given membership in class 1, the probability of endorsing indicator *a* is the same as the probability of endorsing item *b*, and so on.
Referred as ***parallel indicators :*** have identical error rates with respect to each of the latent classes (if we consider one type of response within class as an error)

MODEL:
%OVERALL%
[x#1];

%x#1%
[a$1-b$1*-1] (1);
[c$1*-1] (p1);
[d$1*-1];

%x#2%
[a$1-b$1*1] (2);
[c$1*1] (p2);
[d$1*1];

MODEL CONSTRAINT:
p2 = - p1;

The * followed by a number assigns starting values to the thresholds, which helps specify the class meaning.
In the example, class 1 is the class with negative starting values for thresholds, hence the class with *higher* probability of endorsing items (category 2 = endorsement).
Thresholds for *c* are given names (p1, p2). A MODEL CONSTRAINT command defines a *linear constraint:* the threshold of *c* in class 1 is equal to the negative value of threshold of c in class 2.
This effectively means that the probability of NOT endorsing item *c* in class 1 (the endorsers) is the same as the probability of endorsing item *c* in class 2 (the non-endorsers):
Called ***equal error hypothesis:*** an indicator has the same error rate across the two classes (non endorsement of an item in the endorsers class = a response error)

Constraints on measurement model (ctd.)

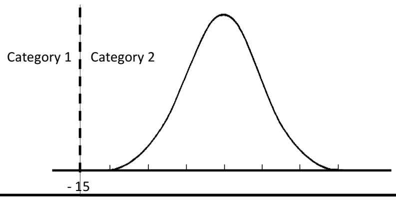
MODEL:
%OVERALL%
[x#1];

%x#1%
[a$1-b$1*-1] (1);
[c$1*-1] (p1);
[d$1@-15];

%x#2%
[a$1-b$1] (2);
[c$1*1] (p2);
[d$1*1];

MODEL CONSTRAINT:
p2 = - p1;

I added a statement to fix the thresholds of *d* in class 1 to (the logit value of) -15 (@ fixes the value of parameters). This means that individuals in class 1 have probability=1 of endorsing the item.
By placing the threshold at the lower limit of the underlying distribution, all scores will be above the “cut”, hence in category 2



Intro to Mplus (ctd.)

OUTPUT:
TECH1 TECH10;
PLOT:
SERIES = a(1) b(2) c(3) d(4);
TYPE = PLOT3;

Command OUTPUT allows you to choose options regarding information in the output. TECH1 for example will report arrays containing parameter specifications and starting values for all free parameters in the model (useful to check what the model is actually doing).
TECH10 reports univariate, bivariate and response pattern model fit information for the categorical dependent variables in the model.
The PLOT command creates graph files that can be useful for inspecting results. TYPE = PLOT3 provides plots with histograms, scatterplots, sample proportions and estimated probabilities (e.g. item

```
TITLE: 2-cl LCA unconstrained
DATA:  FILE = abcd.dat;
VARIABLE: NAMES are a b c d  male female ;
     MISSING = ALL (-9999);
     USEVARIABLES = a b c d;
     CATEGORICAL = a b c d ;
     CLASSES = x(2);
ANALYSIS: TYPE = MIXTURE;
      STARTS = 100 10;
      STITERATIONS = 20;
MODEL:
!the lines preceded by ! are not necessary
!%OVERALL%
![x#1];
!%x#1%
![a$1-d$1] (1-4);
!%x2%
![a$1-d$1] (5-8);
OUTPUT:
TECH1 TECH10;
PLOT:
SERIES = a(1) b(2) c(3) d(4);
TYPE = PLOT3;
```

```
TITLE: 2-cl LCA with measurement constraints
DATA:  FILE = abcd.dat;
VARIABLE: NAMES are a b c d  male female ;
            MISSING = ALL (-9999);
            USEVARIABLES = a b c d;
            CATEGORICAL = a b c d ;
            CLASSES = x(2);
ANALYSIS: TYPE = MIXTURE;
             STARTS = 100 10;
             STITERATIONS = 20;
MODEL:
%OVERALL%
[x#1];

%x#1%
[a$1-b$1*-1] (1);
[c$1*-1]  (p1);
[d$1@-15];

%x#2%
[a$1-b$1] (2);
[c$1*1] (p2);
[d$1*1];

MODEL CONSTRAINT:
p2 = - p1;
OUTPUT:
TECH1 TECH10;
PLOT:
SERIES = a(1) b(2) c(3) d(4);
TYPE = PLOT3;
```

What the output looks like:

```
Loglikelihood values at local maxima,
    seeds, and initial stage start
    numbers:

     -10148.718  987174           1689
     -10148.718  777300           2522
     -10148.718  406118           3827
     -10148.718  51296            3485
     -10148.718  997836           1208
     -10148.718  119680           4434
     -10148.718  338892           1432
     -10148.718  765744           4617
     -10148.718  636396           168
     -10148.718  189568           3651
     -10148.718  469158           1145
     -10148.718  90078            4008
     -10148.718  373592           4396
     -10148.718  73484            4058
     -10148.718  154192           3972
     -10148.718  203018           3813
     -10148.718  785278           1603
     -10148.718  235356           2878
     -10148.718  681680           3557
     -10148.718  92764            2064
```

A successfully converged model will have the best log likelihood values repeated at least twice. If the best (highest→closest to 0) value is not replicated in at least two final stage solutions, it is possible a local solution has been reached (the solution is not trustworthy)

Success   Not there yet

Success

```
Loglikelihood values at local maxima,
    seeds, and initial stage start
    numbers:

     -10148.718  987174         1689
     -10148.718  777300         2522
     -10148.718  406118         3827
     -10148.718  51296          3485
     -10148.718  997836         1208
     -10148.718  119680         4434
     -10148.718  338892         1432
     -10148.718  765744         4617
     -10148.718  636396         168
     -10148.718  189568         3651
     -10148.718  469158         1145
     -10148.718  90078          4008
     -10148.718  373592         4396
     -10148.718  73484          4058
     -10148.718  154192         3972
     -10148.718  203018         3813
     -10148.718  785278         1603
     -10148.718  235356         2878
     -10148.718  681680         3557
     -10148.718  92764          2064
```

The best solution replicated in all the final stages

Not there yet

```
Loglikelihood values at local maxima,
    seeds, and initial stage start
    numbers

     -10153.627  23688          4596
     -10153.678  150818         1050
     -10154.388  584226         4481
     -10155.122  735928         916
     -10155.373  309852         2802
     -10155.437  925994         1386
     -10155.482  370560         3292
     -10155.482  662718         460
     -10155.630  320864         2078
     -10155.833  873488         2965
     -10156.017  212934         568
     -10156.231  98352          3636
     -10156.339  12814          4104
     -10156.497  557806         4321
     -10156.644  134830         780
     -10156.741  80226          3041
     -10156.793  276392         2927
     -10156.819  304762         4712
     -10156.950  468300         4176
     -10157.011  83306          2432
```

A solution (-10155.482) is replicated 2 times, but is not the best solution. The best log-likelihood solution must be replicated for a trust-worthy solution

What if log likelihood not replicated?

If already increased STARTS (e.g. = 100 10) and STITERATIONS (e.g. =20) then:

- Increase the initial stage random sets of starting values further to 500 (e.g. STARTS = 500 10) or more.
- Take the seed value of the best loglikelihood values, then use the OPTSEED option in the ANALYSIS command indicating these seeds:

E.g. ANALYSIS: TYPE=mixture; OPTSEED=370560;

If estimates are replicated using different seeds of best log-likelihoods, we can trust we did not find local solutions

Note: problems in converging indicate the model is not well defined for the data: e.g. too many classes extracted

What does the output look like?

TESTS OF MODEL FIT

Loglikelihood

H0 Value	-2663.146
H0 Scaling Correction Factor for MLR	1.020

Information Criteria

Number of Free Parameters	9
Akaike (AIC)	5344.293
Bayesian (BIC)	5388.462
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	5359.878

What does the output look like?

Chi-Square Test of Model Fit for the Binary and Ordered Categorical (Ordinal) Outcomes

Pearson Chi-Square

Value	3.509
Degrees of Freedom	6
P-Value	0.7428

Likelihood Ratio Chi-Square

Value	3.496
Degrees of Freedom	6
P-Value	0.7445

What does the output look like?

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES
BASED ON THE ESTIMATED MODEL

Latent Classes		
1	524.25270	0.52425
2	475.74730	0.47575

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASS PATTERNS
BASED ON ESTIMATED POSTERIOR PROBABILITIES

Latent Classes		
1	524.25270	0.52425
2	475.74730	0.47575

CLASSIFICATION QUALITY
Entropy 0.467

ENTROPY serves as a measure of the precision of individual classification. It ranges from 0 (everybody has an equal posterior probability of membership in all classes) to 1 (each individual has posterior probability 1 of membership in a single class and probability 0 of membership in the remaining classes). **High entropy indicates clear class separation.**

Determining the number of classes

- Compare statistics and information criteria (BIC, AIC, sample-size adjusted BIC) → the lower, the better fit
- Likelihood Ratio Test (LRT) not applicable: but Mplus provides a Bootstrap LRT (OUTPUT: TECH14).
  - If CLASSES=x(3); the test provides p value of 3-class vs. 2-class fit. A significant value ($p < .05$) would indicate a significant improvement in fit with the inclusion of a third class.
- Mplus provides another similar test (Vu-Luong-Mendell-Rubin → TECH11)
- Consider Entropy (if the aim is finding homogenous clusters)
- Inspect bivariate and response patterns standardised residuals (TECH10): the model with more significant residuals (>|1.96|) has lower fit
- Interpretability of results

What does the output look like?

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Latent Class 1				
Thresholds				
A$1	-0.948	0.187	-5.056	0.000
B$1	-0.764	0.169	-4.529	0.000
C$1	-1.103	0.185	-5.957	0.000
D$1	-0.895	0.184	-4.860	0.000
Latent Class 2				
Thresholds				
A$1	1.272	0.250	5.093	0.000
B$1	0.953	0.174	5.492	0.000
C$1	0.901	0.205	4.397	0.000
D$1	1.023	0.191	5.372	0.000
Categorical Latent Variables				
Means				
X#1	0.097	0.241	0.402	0.688

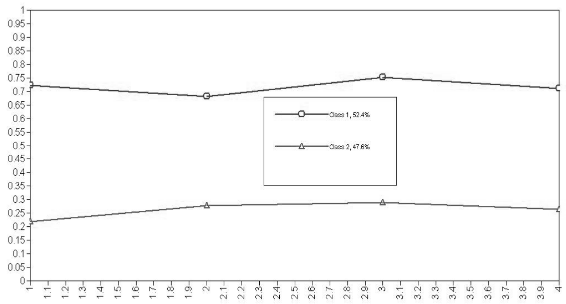
What does the output look like?

RESULTS IN PROBABILITY SCALE

Latent Class 1				
A				
Category 1	0.279	0.038	7.401	0.000
Category 2	0.721	0.038	19.097	0.000
B				
Category 1	0.318	0.037	8.697	0.000
Category 2	0.682	0.037	18.662	0.000
C				
Category 1	0.249	0.035	7.196	0.000
Category 2	0.751	0.035	21.674	0.000
D				
Category 1	0.290	0.038	7.645	0.000
Category 2	0.710	0.038	18.718	0.000
Latent Class 2				
A				
Category 1	0.781	0.043	18.290	0.000
Category 2	0.219	0.043	5.128	0.000
B				
Category 1	0.722	0.035	20.709	0.000
Category 2	0.278	0.035	7.986	0.000
C				
Category 1	0.711	0.042	16.895	0.000
Category 2	0.289	0.042	6.863	0.000
D				
Category 1	0.736	0.037	19.856	0.000
Category 2	0.264	0.037	7.136	0.000

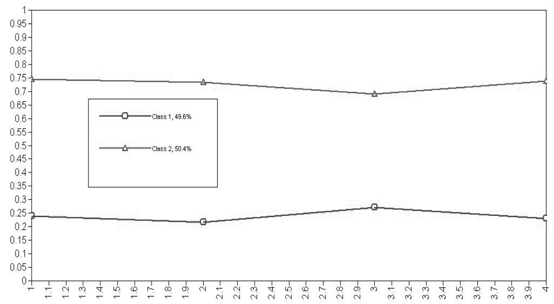
The conditional item response probabilities help attach meaning to each class (similarly to factor loadings in factor analysis).
In this case, class 1 includes individuals that have higher probability of endorsing the items (e.g. If items are symptoms, this class would be the “disorder” class)
Class 2 includes individuals with lower probabilities of endorsing items.
In this case, the profiles do not cross, but is possible to have classes where, for example, individuals in one class have higher probability of endorsing items a and b and individuals in another class endorse items c and d
Ordered vs. *Unordered* solutions

What does the output look like?



2-classes at time 1

What does the output look like?



2-classes at time 2

Step 1 : Validate results of LCA

- Test associations between latent classes (cross-sectional) and covariates: do they make sense?
  - E.g. Does the “victimised” latent class at each age point relate to known risk factors of this process (e.g. School safety)?
- It is also possible to investigate differential item functioning:
  - Two individuals in the same latent class have different item endorsement probabilities.

Note that the introduction of covariates (and distal outcomes) may change the model parameters, including class profiles and their respective size (more on this later)

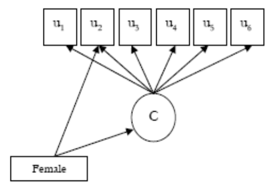


Figure 3.3. LCA model with gender as a covariate that has a direct effect on item (u_2) and on the latent class variable used to explore differential item functioning.

Validate results of LCA (ctd.)

- In Mplus, the regression of one variable on another one is expressed by "ON" in the MODEL command.
- To regress latent variable *class* on covariate gender (coded male) →
  - class ON male;
  - *Regression of class on male: the dependent (class) is regressed on the covariate (male)*

```
VARIABLE:
    NAMES are a1 b1 c1 d1
    male;
    USEVAR are a1-d1 male;
    CATEGORICAL are a1-d1;
    CLASSES are class(2);
MODEL:
%overall%
class on male;
```

Summary Step 1

– Assuming classification is the aim, determine the number of classes at each time point (consider information criteria, model residuals, interpretability of results, etc.)
– It is possible to test constraints on measurement model
– Test associations with covariates and DIF

Step 2: Investigate measurement invariance

Step 2: Investigate measurement invariance

- Assume we have settled for a measurement model at each time point (LCA), identified the number of classes and decided on other parameters constraints (e.g. parallel indicators)
- If the same number and type of classes across time, we can explore *measurement invariance:*
  - Equality of parameters of the measurement models, the conditional item response probabilities
- Measurement invariance assures that latent statuses can be interpreted in the same way across time

Types of measurement invariance

- Full invariance: conditional item probabilities are invariant across measurement occasions
  - Same number and type of classes occur at each time point
- Full measurement non-invariance: no constraints on measurement parameters across time
  - Even if the same n of classes, their profile and their meaning may be different
- Partial measurement invariance: equality of constraints for some measurement parameters across time

Assumptions tested before imposing relationships between latent variables

Measurement invariance

- Reduces the number of parameters estimated (as well as computation)
- Makes interpretation of parameters straightforward

- However, it may not be plausible, depending on the nature of latent classes, indicators, period spanned by measurement points

Mplus : cross-sectional LCA

- We assume 4 indicators (a b c d) measured at time 1 (a1 b1...d1) and at time 2 (a2...d2)
- We estimate two latent categorical variables, with two classes each (latent variables are *x* at time 1 and *y* at time 2).
- How can you make sure indicators a1 to d1 are regressed on x and a2 to d2 are regressed on y using Mplus?

Mplus: cross-sectional LCA (cd.)

```
VARIABLE:   NAMES ARE a1 b1 c1 d1
a2 b2 c2 d2
cova;
     usevar are a1-d1 a2-d2;
      CATEGORICAL = a1-d1  a2-d2  ;
        CLASSES = x (2) y(2) ;


  ANALYSIS:   TYPE = MIXTURE;
         STARTS = 100 10;
         STITERATIONS = 20;
```

Mplus: cross-sectional LCA (cd.)

```
MODEL:
      %OVERALL%

MODEL x:

%x#1%
 [a1$1-d1$1*-1] ;
 %x#2%
 [a1$1-d1$1*1] ;

MODEL y:

%y#1%
[a2$1-d2$1*-1] ;
%y#2%
[a2$1-d2$1*1] ;
```

No specification regarding the relationship between x and y as yet

When 2 or more latent variables are estimated, the part of the model specific to latent variable x and its categories is preceded by MODEL x:

Thresholds (therefore : response probabilities) are estimated for a1 b1 c1 and d1 within classes of latent variable x

Thresholds (therefore response probabilities) are estimated for a2 b2 c2 and d2 within classes of latent variable y .

No constraints on thresholds (freely estimated): conditional item response probabilities freely estimated (non-invariance)

Mplus: cross-sectional LCA (ctd.) Full measurement invariance

```
MODEL:
      %OVERALL%

MODEL x:

%x#1%
 [a1$1-d1$1*-1]  (1-4);
 %x#2%
 [a1$1-d1$1*1]  (5-8);

 MODEL y:

 %y#1%
 [a2$1-d2$1*-1]  (1-4);
 %y#2%
 [a2$1-d2$1*1]  (5-8);
```

Thresholds (therefore: response probabilities) are constrained to be the same for a1 in x1 and a2 in y1, or else: P(a1=1 | x=1) = P (a2 = 1 | y =1). The same is true for indicator b in class 1 of x and y, and so on.

Similar constraints are imposed for class 2 of x and y (5-8)

In this way, we specify a full-measurement invariance model

Mplus output: measurement non-invariance

x=1 and y=1

Latent Class Pattern 1 1

A1				
Category 1	**0.279**	0.038	7.401	0.000
Category 2	**0.721**	0.038	19.097	0.000
B1				
Category 1	0.318	0.037	8.697	0.000
Category 2	0.682	0.037	18.662	0.000
[...]				
A2				
Category 1	**0.254**	0.031	8.066	0.000
Category 2	**0.746**	0.031	23.734	0.000
B2				
Category 1	0.266	0.034	7.837	0.000
Category 2	0.734	0.034	21.577	0.000
[...]				

Prob of endorsing (category 2) item a1 if x=1 is 0.729; prob of endorsing item a2 if y=1 is 0.746

x=2 and y=2

Latent Class Pattern 2 2

A1				
Category 1	**0.781**	0.043	18.290	0.000
Category 2	**0.219**	0.043	5.128	0.000
B1				
Category 1	0.722	0.035	20.709	0.000
Category 2	0.278	0.035	7.986	0.000
[..]				
A2				
Category 1	**0.760**	0.032	23.693	0.000
Category 2	**0.240**	0.032	7.483	0.000
B2				
Category 1	0.783	0.030	26.134	0.000
Category 2	0.217	0.030	7.232	0.000
[..]				

Prob of endorsing (category 2) item a1 if x=2 Is 0.219; prob of endorsing item a2 if y=2 is 0.240

Mplus output: measurement invariance

x=1 and y=1

Latent Class Pattern 1 1

A1				
Category 1	0.269	0.025	10.820	0.000
Category 2	0.731	0.025	29.443	0.000
B1				
Category 1	0.293	0.025	11.729	0.000
Category 2	0.707	0.025	28.286	0.000
[..]				
A2				
Category 1	0.269	0.025	10.820	0.000
Category 2	0.731	0.025	29.443	0.000
B2				
Category 1	0.293	0.025	11.729	0.000
Category 2	0.707	0.025	28.286	0.000

x=2 and y=2

Latent Class Pattern 2 2

A1				
Category 1	0.771	0.026	29.694	0.000
Category 2	0.229	0.026	8.842	0.000
B1				
Category 1	0.755	0.023	33.225	0.000
Category 2	0.245	0.023	10.789	0.000
[..]				
A2				
Category 1	0.771	0.026	29.694	0.000
Category 2	0.229	0.026	8.842	0.000
B2				
Category 1	0.755	0.023	33.225	0.000
Category 2	0.245	0.023	10.789	0.000

Prob of endorsing (category 2) item a1 if x=1 is 0.731 and is the same probability of endorsing item a2 if y=1 (equality constraint imposed)

Test for measurement invariance
Run LRT test:

- *Non-invariance:*

TESTS OF MODEL FIT

Loglikelihood

H0 Value -5295.298
H0 Scaling Correction Factor 1.011
for MLR

Information Criteria

Number of Free Parameters 18
Akaike (AIC) 10626.597
Bayesian (BIC) 10714.936
Sample-Size Adjusted BIC 10657.767

- *Invariance:*

TESTS OF MODEL FIT

Loglikelihood

H0 Value -5300.601
H0 Scaling Correction Factor 1.012
for MLR

Information Criteria

Number of Free Parameters 10
Akaike (AIC) 10621.202
Bayesian (BIC) 10670.280
Sample-Size Adjusted BIC 10638.519

Test for measurement invariance
Run LRT test:

*LR = -2 * (L0-L1)*

L0 = log-likelihood null model (model with equality constraints)

L1= log-likelihood of unconstrained model

When using MLR estimator (as default in Mplus), LR needs to be adjusted using scaling factor

*LR = -2 *(L0-L1 / **cd**)*

cd = [(p0*c0)-(p1*c1)]/(p0-p1)

c0 = scaling factor null model

c1 = scaling factor alternative model

p0 = parameters in null model

p1 = parameters in alternative model

Test for measurement invariance
Run LRT test:

- *Non-invariance:*

TESTS OF MODEL FIT

Loglikelihood

H0 Value -5295.298
H0 Scaling Correction Factor 1.011
for MLR

Information Criteria

Number of Free Parameters 18

- *Invariance:*

TESTS OF MODEL FIT

Loglikelihood

H0 Value -5300.601
H0 Scaling Correction Factor 1.012
for MLR

Information Criteria

Number of Free Parameters 10

Cd=[(10*1.012)-(18*1.011)] / (10-18) = 1.0097
LR = -2 [-5300.601-(-5295.298)]/1.0097 = 10.50
Df = (p1-p0) = 18 – 10 = 8
Chi square (10.50, 8) = .23

The LRT indicates no significant worsening of fit if equality constraints imposed: assume measurement invariance

Partial Measurement Invariance

- Many different options, e.g.:
  - *Time-specific structure of one class*: in the example class 1 of x and y (time 1 and 2) is freely estimated across time, while equality constraints are imposed on class of x and y (this class is invariant)

```
MODEL x:
%x#1%
[a1$1-d1$1*-1] ;
%x#2%
[a1$1-d1$1*1] (5-8);
MODEL y:
%y#1%
[a2$1-d2$1*-1] ;
%y#2%
[a2$1-d2$1*1] (5-8);
```

Partial Measurement Invariance

- Many different options, e.g.:
  - *Differential item functioning with respect to time*: one item (or more) within a class is non-invariant across time (*a* in class 1 of *x* and *y*) , while the rest of the parameters are held invariant

```
MODEL x:
%x#1%
[a1$1*-1];
[b1$1-d1$1*-1] (2-4);
%x#2%
[a1$1-d1$1*1] (5-8);
MODEL y:
%y#1%
[a2$1*1];
[b2$1-d2$1*-1] (2-4);
%y#2%
[a2$1-d2$1*1] (5-8);
```

Explore transitions based on cross-sectional results

- Before imposing relationships between latent variables, it may be useful to inspect transitions between latent classes estimated cross-sectionally to get some preliminary idea of the type of movement in the sample across time.
- Use the modal class assignment (each individual assigned to the class with highest posterior probability).
- In Mplus: include "IDVAR=idnumber" in VARIABLE !*this tells Mplus to include an ID variable (idnumber) in the data file.*
- Command SAVEDATA writes a file:

SAVEDATA: file is modalclass_c2.dat;
SAVE = cprob ; *!cprob includes the modal class assignment !and the probability of being in each class for each individual in the !sample*

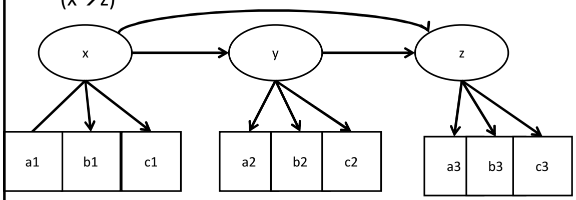
Summary Step 2

- Measurement invariance needs to be investigated before imposing a relationship between latent statuses at each time point
- Full measurement invariance facilitates estimation and interpretation, but may sometimes not be a plausible assumption
- If full measurement invariance not tenable, test partial measurement invariance (e.g. a time invariant "normative" class of non-victimised adolescents or non-violent children)

Step 3: Explore specification of the latent transition model without covariates

Step 3: Explore specification of the latent transition model without covariates

- LTA is an *autoregressive* model: one stage directly related to previously measured stage.
- First order effects (x→y) ; Second order effects (x→z)



Step 3: Explore LTA solution

- 3.1 Impose constraints on transition probabilities
- 3.2 First and second order effects
- 3.3 Stationary transitions (if 3 time measurements and no covariates)
- 3.4 Latent higher-order covariates (Mover-Stayer model)
- 3.5 Model fit

Step 3

- We have settled on class specifications and measurement characteristics of classes across time
- We can now impose auto-regressive relationships between latent variables across time
- In Mplus:

```
    CLASSES= x(2) y(2);
MODEL:
    %overall%
    y ON x
    MODEL x:
    %x#1%
    [a1$1-d1$1] (1-4);
    ....
```

This can also be written as :

```
%overall%
[x#1];
[y#1];   !estimates logit intercept
y#1 ON x#1 ;   !multinomial logistic
               !regression y on x
```

If y had 3 categories (hence: 2 thresholds)

```
%overall%
[x#1];  [y#1];  [y#2];
y#1 y#2 ON x#1;
```

Step 3.1: Restricting transition probabilities

Time 1		Time 2	
	Class1	Class2	Class3
Class1	$\tau_{1\|1}$	$\tau_{2\|1}$	$\tau_{3\|1}$
Class2	$\tau_{1\|2}$	$\tau_{2\|2}$	$\tau_{3\|2}$
Class3	$\tau_{1\|3}$	$\tau_{2\|3}$	$\tau_{3\|3}$

- Some tau parameters can be fixed
- This can help express a model of development (e.g. No backsliding)

Step 3.1: Restricting transition probabilities (ctd)

Time 1		Time 2	
	Class1	Class2	Class3
Class1	$\tau_{1\|1}$	$\tau_{2\|1}$	$\tau_{3\|1}$
Class2	**0**	$\tau_{2\|2}$	$\tau_{3\|2}$
Class3	**0**	**0**	$\tau_{3\|3}$

- A model of No backsliding among ordered classes:
  - If the classes represented degrees of ability (from 1=less able to 3=more able), the probability of transitioning from a more advanced level to a less advanced one is fixed to 0.

Step 3.1: Restricting transition probabilities (ctd).

Time 1		Time 2	
	Non problematic	Sometimes problematic	Often problematic
Non problematic	$\tau_{1\|1}$	$\tau_{2\|1}$	**0**
Sometimes problematic	$\tau_{1\|2}$	$\tau_{2\|2}$	$\tau_{3\|2}$
Often problematic	**0**	$\tau_{2\|3}$	$\tau_{3\|3}$

- In this example , we assume there are no transitions from a class at one extreme to a class at the other end (only transitions between adjacent stages allowed)

How to calculate transition probabilities

x \ y (3)	1	2	3
1	$a_1 + b_{11}$	$a_2 + b_{21}$	0
2	$a_1 + b_{12}$	$a_2 + b_{22}$	0
3	a_1	a_2	0

- The transition probabilities from x to y are given by unordered logistic regression expressions:

$P(y=c|x=1)= EXP(a_c + b_{c1}) / sum_1$
$P(y=c|x=2)= EXP(a_c + b_{c2}) / sum_2$
$P(y=c|x=3)= EXP(a_c + b_{c3}) / sum_3$

- sum_1 etc. represent the sum of the exponentiations across the classes of y in rows x (= 1, 2,3).
- The values in column y=3 are all 0 (a3=0; b31 =0 ; ...; b33=0) because the last class is the reference class.

How to calculate transition probabilities (ctd.)

x \ y (3)	1	2	3
1	$a_1 + b_{11}$	$a_2 + b_{21}$	0
2	$a_1 + b_{12}$	$a_2 + b_{22}$	0
3	a_1	a_2	0

- The parameters in the table are in Mplus :

a1: [y#1];
a2: [y#2];
b11: y#1 ON x#1; b12: y#1 ON x#2;
b21: y#2 ON x#1; b22: y#2 ON x#2;

How to calculate transition probabilities (ctd.)

y (3)

x	1	2	3
1	$a_1 + b_{11}$	$a_2 + b_{21}$	0
2	$a_1 + b_{12}$	$a_2 + b_{22}$	0
3	a_1	a_2	0

- Example:

a1 → [y#1] = -1.8;
a2 → [y#2] = 0.3;
b11 → y#1 ON x#1 = 2.6;
b12 → y#1 ON x#2 = 2.1;
b21 → y#2 ON x#1 = -1.3 ;
b22→ y#2 ON x#2 = -0.5;

P(y=1|x=1) = exp(a1+b11) / exp(a1+b11)+exp(a2+b21)+exp(0)

P(y=1|x=1) = exp(-1.8+ 2.6) / exp(-1.8+2.6)+exp(0.3+(-1.3))+1

P(y=2|x=1) = exp(a2+b21) / exp(a1+b11)+exp(a2+b21)+exp(0)

P(y=2|x=1) = exp(0.3+ (-1.3)) / exp(-1.8+2.6)+exp(0.3+(-1.3))+1

P(y=3|x=1) = exp(a3+b31) / exp(a1+b11)+exp(a2+b21)+exp(0)

P(y=3|x=1) = exp(0) / exp(-1.8+2.6)+exp(0.3+(-1.3))+1

How to calculate transition probabilities

y (3)

x	1	2	3
1	$a_1 + b_{11}$	$a_2 + b_{21}$	0
2	$a_1 + b_{12}$	$a_2 + b_{22}$	0
3	a_1	a_2	0

τ 1|3 = 0

If we want to fix P (y=3|x=1) to 0 we refer to the formula for its probability:
P(y=1|x=3) = exp(a1) / exp(a1)+exp(a2)+exp(0)
In this case, to ensure the result is 0, we make the value of exp(a1) very small by assigning to a1 a large negative number (for example, -15).
Since parameter a1 in Mplus is indicated by the logit intercept [y#1] :
MODEL:
%overall%"
[y#1@-15];
Thus, we should obtain that, whatever the other parameters, P(y=1|x=3) ≈ 0

How to calculate transition probabilities

y (3)

x	1	2	3
1	$a_1 + b_{11}$	$a_2 + b_{21}$	0
2	$a_1 + b_{12}$	$a_2 + b_{22}$	0
3	a_1	a_2	0

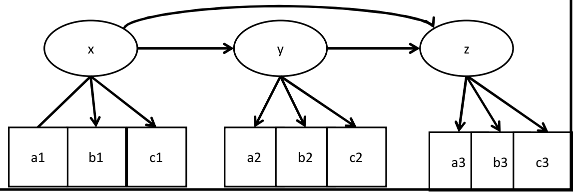
No backsliding

a1 → [y#1];
a2 → [y#2];
b12 → y#1 ON x#2;

We want:
P(y=2|x=3) = exp(a2) / exp(a1)+exp(a2)+exp(0) ≈ 0
P(y=1|x=3) = exp(a1) / exp(a1)+exp(a2)+exp(0) ≈ 0
P(y=1|x=2) = exp(a1+b12) / exp(a1+b12)+exp(a2+b22)+exp(0) ≈ 0
In this case, as well as ensuring a1 has a very small value (as before), we need to ensure a2 has a very small value, and also that the numerator of P(y=1|x=2) is small.
We can fix: [y#1@-10]; [y#2@-10]; y#1 ON x#2@-5;
In this case, the numerator of the three expressions above will be respectively : exp(-10) ; exp(-10) ; exp (-15)

Second order effects

- First order effects (x→y ; y→z): if no second order effects, non-adjacent latent variables are indirectly related
- Second order effects (x→z): lasting direct effects that being in category of x has on later class membership



Second order effects (ctd.)

VARIABLES:
...
Classes = x(2) y(2) z(2);

MODEL:
%overall%
y ON x; *!first order x→y*
z ON y; *!first order y→z*
z ON x ; *!2nd order x→z*

Can also be written:
y on x;
z ON x y;

Or:
[x#1];
[y#1];
[z#1];
y#1 ON x#1;
z#1 ON y#1 x#1;

Second order effects (ctd.)

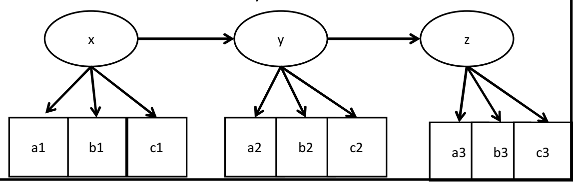
- Inspection of transition probabilities matrices estimated under different assumption (first- vs. second-order effects) help highlight impact of previous classification

1st ord.		**Grade 8**	
Grade 6	Victimised	Some. Victim.	Non Victim.
Victimised	**.27**	.37	.36
Some. Victim.	.06	**.29**	.65
Non Victim.	.02	.10	**.88**

2nd ord.		**Grade 8**	
Grade 6	Victimised	Some. Victim.	Non Victim.
Victimised	**.32**	.37	.31
Some. Victim.	.04	**.34**	.62
Non Victim.	.01	.06	**.93**

Stationary transitions

- Assume transitions across time points (> 2) are stationary: same probabilities to transition from a stage to another between time 1- time 2 and between time 2-time 1, and so on...
- However, if covariates are included, stationariety is no longer meaningful (it would bias estimation of covariates' coefficients)



Stationary transitions (ctd.)

MODEL:
%OVERALL%
[x#1];
[y#1](1);
[z#1] (1);

Logit intercepts of y and z constrained to be equal

y ON x (2);
z ON y (2);

Multinomial logistic regression of y on x (time 1-2) and z on y (time 2-3) constrained to be equal

Stationary transitions (ctd.): Output

LATENT TRANSITION PROBABILITIES BASED ON
THE ESTIMATED MODEL

X Classes (Rows) by Y Classes (Columns)

	1	2
1	0.339	0.661
2	0.863	0.137

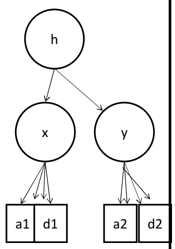
Y Classes (Rows) by Z Classes (Columns)

	1	2
1	0.339	0.661
2	0.863	0.137

Same transition probabilities: change happens at the same rate across time points

Higher-order latent variables

- It is possible to estimate a further latent variable to investigate unobserved heterogeneity in developmental process
- For example: a latent class of "movers" (individuals that transition between stages across measurement occasions) and one of "stayers" (individuals that remain in the same class across measurement occasions).
- E.g. If *x* and *y* are classes of depression, mover/stayer model help identify individuals chronically depressed



Movers/stayers model

- Allows more accurate estimation of transition probabilities if, indeed, there are individuals with zero probability of transitioning.
- Pre-requisite: same number of classes with same meaning (measurement invariance).

Movers: freely estimate the probability of transitioning across time points
Stayers: fix the probability of transitioning across time points to 0

STAYERS			
Time 1		Time 2	
	Class1	Class2	Class3
Class1	$\tau_{1\|1}$	0	0
Class2	0	$\tau_{2\|2}$	0
Class3	0	0	$\tau_{3\|3}$

How to calculate transition probabilities with covariates

	Time	2
Time 1	y1	y2
x1	a1+b11+g1(ms_i)	0
x2	a1+g1(ms_i)	0

- The Mover/Stayer latent variable (ms) is a (latent) covariate of the two latent variables x (time 1) and y (time 2)
- The latent variable ms has two categories.
- One category of ms (the last one) is the reference category
- The coefficient g describes the change in log odds for one category of ms as compared to the reference category

If ms = 1
P(y=1|x=1)= EXP(a1 + b11+g_1) / Exp(a1+b11+g1)+exp(0)

If ms=2 (ref. Cat.)→ g1(ms=0)=0
P(y=1|x=1)= EXP(a1 + b11) / Exp(a1+b11)+exp(0)

How to calculate transition probabilities with covariates (ctd.)

	Time	2
Time 1	y1	y2
x1	a1+b11+g1(msi)	0
x2	a1+g1(msi)	0

If ms = 1
P(y=1|x=1)= EXP(a1 + b11+g 1) / Exp(a1+b11+g1)+exp(0)

If ms=2 (ref. Cat.)→ g1(ms=0)=0
P(y=1|x=1)= EXP(a1 + b11) / Exp(a1+b11)+exp(0)

Assume ms=1 is the mover class and ms=2 the stayer

Fixing a1 → [y#1] = -15
ensures that in category 2 of ms (reference category) P(y=2|x=1) ≈ 0 (0 prob. of moving from 1 to 2)

If ms=2 → g1(ms)=0
P(y=2|x=1) = exp (a1) / exp(a1)+1
P(y=2|x=1) = exp (-15) / exp(-15)+1 ≈ 0

Mover / Stayer model in Mplus

```
VARIABLE:
   CLASSES = ms(2) x (2) y(2)

MODEL:
      %OVERALL%
x y ON ms;
[y#1@-15];

MODEL ms:
%ms#1%   !mover class
y#1 ON x#1;
%ms#2%   !stayer class
y#1 ON x#1@30;
```

- x y ON ms: Regresses x and y ON ms (mover/stayer)
- [y#1@-15]: Fixes prob of y=1|x=2 in ms2 to 0
- y#1 ON x#1: Freely estimates transitions in ms1
- y#1 ON x#1@30: Fixes P(y=1|x=1) = 1 in ms2

Mover / Stayer model in Mplus (ctd.)

```
MODEL ms.x:
%ms#1.x#1%
[a1$1-d1$1*-1] (1-4);
%ms#1.x#2%
[a1$1-d1$1*1] (5-8);
%ms#2.x#1%
[a1$1-d1$1] (1-4);
%ms#2.x#2%
[a1$1-d1$1] (5-8);

MODEL ms.y:
%ms#1.y#1%
[a2$1-d2$1] (1-4);
%ms#1.y#2%
[a2$1-d2$1] (5-8);
%ms#2.y#1%
[a2$1-d2$1] (1-4);
%ms#2.y#2%
[a2$1-d2$1] (5-8);
```

- MODEL ms.x: When a higher-order latent class is introduced
- (5-8) statements: This specifies measurement invariance: thresholds of x#2 the same as y#2. It is possible to specify different measurement constraints in ms1 and ms2, or for combinations of ms, x, y

Mover / Stayer model in Mplus : output

```
Categorical Latent Variables

 X#1     ON
   MS#1         -3.012    2.104    -1.432    0.152

 Y#1     ON
   MS#1         14.845    0.000   999.000   999.000

 Means
   MS#1          0.823    0.224     3.674    0.000
   X#1           2.384    2.078     1.147    0.251
   Y#1         -15.000    0.000   999.000   999.000

Latent Class Pattern 1 1 1

 Y#1     ON
   X#1          -3.579    0.000   999.000   999.000

Latent Class Pattern 2 1 1

 Y#1     ON
   X#1          30.000    0.000   999.000   999.000
```

- Means Y#1 -15.000: This had been fixed P(y=1|x=2)=0 in ms#2
- Pattern 1 1 1, Y#1 ON X#1: This had been freed in ms#1 (i.e. not equal across ms classes)
- Pattern 2 1 1, Y#1 ON X#1: This had been fixed P(y=1|x=1)=1 in ms#2

Mover / Stayer model in Mplus : output (ctd.)

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES BASED ON ESTIMATED POSTERIOR PROBABILITIES

ms=1 x=1 y=1
ms=1 x=1 y=2
...
ms=2 x=2 y=2

Latent Class Pattern		
1 1 1	5.64118	0.00564
1 1 2	236.08630	0.23609
1 2 1	208.96989	0.20897
1 2 2	244.08220	0.24408
2 1 1	279.44841	0.27945
2 1 2	0.00009	0.00000
2 2 1	0.00001	0.00000
2 2 2	25.77192	0.02577

Movers (ms=1)

Stayers (ms=2)

Model fit of LTA models

- The Chi-Square statistics (Pearson or Likelihood-ratio based) not recommended (distribution not well approximated when large number of sparse cells)
- Nested models (e.g. Stationary transitions vs. non-stationary) → compare with LRT (remember correction by scaling factor if MLR estimator)
- Consider residuals (less significant residuals→ better fit)

Important to build model step by step

Summary Step 3

- Impose autoregressive relationships (current status predicted by previous status)
- Consider and test constraints on transition probabilities
- If more than 2 time points, it is possible to consider stationary transitions (but not meaningful if covariates are included) and second-order effects
- It is possible to include higher-order latent covariates (e.g. Movers / Stayers model)

Step 4: Include covariates in the LTA model

Step 4: Include covariates in the LTA model

- Categorical, nominal and continuous covariates can be included as predictors of class membership and transition probabilities
- Covariates can be time-varying or time-invariant
- They can have time-varying or time-invariant effects (independently of their being time-varying or not)

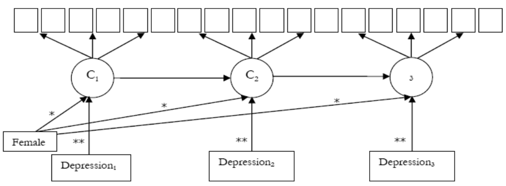


Figure 2.8. Model diagram of an LTA model with a time-invariant covariate (gender) and a time-varying covariate (depression).

Step 4: Categorical covariates

- If covariates are categorical (e.g. Gender) : **multiple-groups LTA**
  - It is possible to explore measurement invariance across groups: e.g. Do items map onto the latent variables in the same way for males and females?
  - Explore differences in latent class membership at start point
    - E.g. Does probability of being victimised in Grade 6 differ between males and females?
  - Explore differences in transition probabilities
    - E.g. Does probability of transitioning from victimised to non-victimised differ between males and females?

Investigating measurement invariance

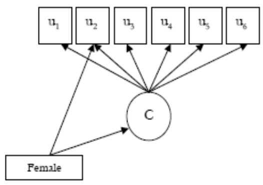


Figure 3.3. LCA model with gender as a covariate that has a direct effect on item (u_2) and on the latent class variable used to explore differential item functioning.

- This model can be extended to LTA models

Explore differences in latent class membership at start point

- E.g. Does probability of being victimised in Grade 6 differ between males and females?
  - VARIABLES: usevar are male a1 b1 c1 d1 a2 b2 c2 d2;

classes = x(2) y(2);
[...]
MODEL:
x on male;
y on x;
Model x:
%x#1%
[...]

Logistic regression of x on male (→ x on male;)

In this case, class membership at time 2 is predicted by class membership at time 1 (x) and NOT by gender : transition probabilities from x to y are the same for males and females (→ y on x;)

Explore differences in transition probabilities

- E.g. Does probability of transitioning from victimised to non-victimised differ between males and females?

VARIABLES: [...]
usevar are male a1 b1 c1 d1 a2 b2 c2 d2;
classes = x(2) y(2);
[...]
MODEL:
x on male;
y on x male;
Model x:
%x#1%
[...]

Logistic regression of x on male (→ x on male;)

In this case, class membership at time 2 is also predicted by gender while controlling for previous latent status : transition probabilities from x to y differ for males and females. (→ y on x male;)

Transition probabilities with categorical covariates

	Time	**2**	
Time 1	y1	y2	y3
x1	a1+b11+g1(male$_i$)	a2+b21+g2(male$_i$)	0
x2	a1+b12+g1(male$_i$)	a2+b22+g2(male$_i$)	0
x3	a1+g1(male$_i$)	a2+g2(male$_i$)	0

g1 and g2 are the logistic coefficients: change in the log odds of being in class y1 or y2 compared to class y3 (reference) for males (male=1) as compared to females (male=0).

In Mplus these parameters are:
g1 → y#1 ON male
g2 → y#2 ON male

Transition probabilities with categorical covariates

	Time	**2**	
Time 1	y1	y2	y3
x1	a1+b11+g1(male$_i$)	a2+b21+g2(male$_i$)	0
x2	a1+b12+g1(male$_i$)	a2+b22+g2(male$_i$)	0
x3	a1+g1(male$_i$)	a2+g2(male$_i$)	0

Transition probabilities for females (male=0) can be calculated considering that the g1 and g2 terms are equal to 0 (reference class). E.g. :
P(y=1|x=1) = exp(a1+b11) / [exp(a1+b11) + exp(a2+b21) + exp(0)]

Transition probabilities for males (male=1) are calculated adding g1 and g2 parameters. Eg:
P(y=1|x=1) = exp(a1+b11+g1) / [exp(a1+b11+g1) + exp(a2+b21+g2) + exp(0)]

Transition probabilities with categorical covariates

Table 3.17. Estimated transition probabilities presented by gender (males on the left, females on the right) based on model with only gender as a covariate

	Males Grade 7				Females Grade 7		
Grade 6	VI	SV	NV		VI	SV	NV
VI	0.42	0.42	0.16	VI	0.42	0.38	0.19
SV	0.05	0.51	0.44	SV	0.05	0.44	0.51
NV	0.01	0.11	0.88	NV	0.01	0.09	0.91

	Males Grade 8				Females Grade 8		
7th Grade	VI	SV	NV		VI	SV	NV
VI	0.51	0.39	0.10	VI	0.52	0.37	0.11
SV	0.07	0.51	0.42	SV	0.07	0.46	0.47
NV	0.02	0.07	0.91	NV	0.01	0.06	0.93

Note. VI class = victimized class, SV class = sometimes-victimized class, NV class = nonvictimized class.

One can thus obtain transition matrices for males and females

Multigroup LTA in Mplus

- Another possibility is to use the KNOWNCLASS option:

  – VARIABLES: usevar are a1 b1 c1 d1 a2 b2 c2 d2; *!NOTE: no male*
  classes = **cmale(2)** x(2) y(2);
  KNOWNCLASS = cmale (male = 0 male = 1);
  [...]
  MODEL:
  x on cmale;
  y on x cmale;
  Model x:
  %x#1%
  [...]

Defines a new class for which class membership is known (observed)

observed variable male is used to define known classes: first class is individuals with value 0 (females).

Multigroup LTA in Mplus

- Another possibility is to use the KNOWNCLASS option:

  – VARIABLES: usevar are a1 b1 c1 d1 a2 b2 c2 d2; *!NOTE: no male*
  classes = **cmale(2)** x(2) y(2);
  KNOWNCLASS = cmale (male = 0 male = 1);
  [...]
  MODEL:
  x on cmale;
  y on x cmale;
  Model x:
  %x#1%
  [...]

The known class is used as a predictor of latent class membership at time 1 and time 2

Multigroup LTA in Mplus

KNOWNCLASS allows another way to specify measurement invariance and parameters

Model cmale.**x**:	Model cmale.**y**:
%**cmale#1**.x#1%	%cmale#1.**y#1**%
[a1$1-d1$1] **(1-4)**;	[a2$1-d2$1] **(1-4)**;
%cmale#1.x#2%	%cmale#1.**y#2**%
[a1$1-d1$1] **(5-8)**;	[a2$1-d2$1] **(5-8);**
%**cmale#2**.x#1%	%cmale#2.y#1%
[a1$1-d1$1] **(9-12)**;	[a2$1-d2$1] **(9-12)**;
%cmale#2.x#2%	%cmale#2.y#2%
[a1$1-d1$1] **(12-16)**;	[a2$1-d2$1] **(12-16)**;

In this case, different thresholds (item response prob.) are estimated for females and males, but these are invariant at time 1 and 2 within groups

Estimation with covariates

- The inclusion of covariates changes estimation of LTA parameters, including class profiles, class size and transition probabilities (see formulae for calculating transition probabilities with covariates).
  - This is also the reason why stationary transition probabilities are not meaningful when covariates are included in the model: imposing these constraints would bias estimation of covariates coefficients
- If adding covariates changes the class structure substantially, this might point to the need to allow for measurement non-invariance (more investigation needed).

Estimation with covariates

http://bit.ly/Lr9Q6X

"classes seemed to change when adding xs as predictors of c:
I can think of 3 reasons:

1) more information is available when adding xs and therefore this solution is what one should trust.
2) Another is that the model may be misspecified when adding the xs because there may be some omitted direct effects from some xs to some ys/us (these can be included).
3) A third explanation is more subtle and has to do with individuals' misfit. There may be examples where for some individuals in the sample the ys/us "pull" the classes in a different direction than the xs. Note that both y/u and x information contribute to class formation. Consider the example where in a 2-class model a high x value has a positive influence on being in class 2, and being in class 2 gives a high probability for u=1 for most. *Individuals who have many u=1 outcomes but low x values are not fitting this model well*. If the x information dominates the u information then these individuals will be classified differently using only u versus using u and x."

Relating LCA results

http://www.statmodel.com/download/relatinglca.pdf

If one does not want to include covariates while estimating latent classes, there are different approaches where the latent class membership is regressed on covariates. E.g.:

- Consider the most likely class (modal class assignement based on posterior probabilities)
  - In this case, class membership is used as an observed variable ignoring the fact that individuals have different probabilities of being in one class
- Weight regression by each individual's posterior probability of being in a given class
- Clark & Muthén: including covariates while forming latent classes still performed the best

Summary Step 4

- Covariates can be time-varying or time invariant
- Interval or categorical covariates can be used to predict class affiliation at first measurement point and changes in transition probabilities
- Categorical covariates: multiple groups LTA
- Covariates may substantially change LTA parameters, including measurement parameters. This may warrant further investigation (e.g. DIF)

Step 5: Include distal outcomes

Step 5: Include distal outcomes

- Variables measured after the period considered by the model can be included as long-term outcomes related to the change process.
- Distal outcomes can be included in different ways. E.g.:
  – Can be related to a higher-order latent variable such as Mover-Stayer classification
  – Can be related to the latent status at the last time point of measurement

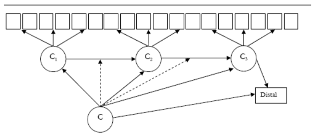


Figure 2.10. Model diagram of the LTA model with a higher-order latent class variable and a distal outcome.

Step 5: Include distal outcomes (ctd)

- Distal outcomes of different type (e.g. categorical or interval variables) can be included in LTA.
- In the case of interval variables, the variable means can be estimated for each class of the latent variable; these means can be compared to investigate significant differences.
- In the case of binary variables, proportions are estimated for each class of the latent variable.

Distal interval outcomes in Mplus

- The interval variable is *testscor*

```
VARIABLES are  male a1 b1 c1 d1 a2 b2 c2 d2 testscor ;
        usevar are a1-d2 testscor;
        categorical are a1-d2;
        classes are x(2) y(2);
[...]
```

testscor is in the USEVAR but not CATEGORICAL statement (therefore: interval variable)

```
MODEL:
%overall%
y ON x;
MODEL x:
%x#1%
[a1$1-d1$1] (1-4);
%x#2%
[a1$1-d1$1] (5-8);

MODEL y:
%y#1%
[a2$1-d2$1] (1-4);
[testscor] (p1);
%y#2%
[a2$1-d2$1] (1-4);
[testscor] (p2);
```

Estimates means of testscor in y1 and y2 : in MODEL command an interval variable name between brackets indicates the variable mean

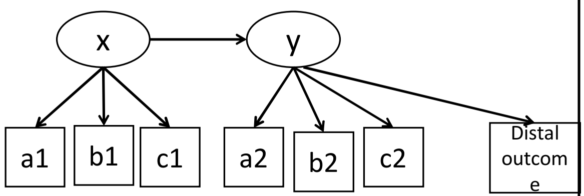
```
MODEL TEST:
p1 = p2;
```

This provides Wald test for H0 : p1 = p2

Binary distal variable in Mplus

- A binary distal outcome (or a categorical one) can be included in the same way that other categorical indicators are regressed on the latent variable
- According to Muthén, statistically the distal outcome is another latent class indicator (although one thinks of it in substantively different terms)
- The inclusion of a distal covariate may change some LTA parameters :
  - If this is the case, this warrants further investigation

Binary distal variable in Mplus (ctd).



Conditional independence is assumed between indicators and distal outcome given latent class membership

Binary distal variable in Mplus (ctd.)

- The outcome (binary) variable is *testbin*

```
VARIABLES : names are  male a1 b1 c1 d1 a2 b2 c2 d2 tes
        usevar are a1-d2 testbin;
        categorical are a1-d2 testbin;
        classes are x(2) y(2);
[…]

MODEL:
%overall%
y ON x;
MODEL x:
%x#1%
[a1$1-d1$1] (1-4);
%x#2%
[a1$1-d1$1] (5-8);

MODEL y:
%y#1%
[a2$1-d2$1] (1-4);
[testbin$1] ;
%y#2%
[a2$1-d2$1] (1-4);
[testbin$1] ;
```

testbin is in the USEVAR and indicated as CATEGORICAL outcome

Estimates thresholds of testbin in y1 and y2 (hence, proportions)

Binary distal variable in Mplus (ctd.)

- OUTPUT:

```
RESULTS IN PROBABILITY SCALE

Latent Class Pattern 1 1

A1
   Category 1      0.974    0.005   199.237    0.000
   Category 2      0.026    0.005     5.409    0.000
[…]
TESTBIN
   Category 1      0.395    0.012    33.182    0.000
   Category 2      0.605    0.012    50.832    0.000
Latent Class Pattern 1 2
[…]
TESTBIN
   Category 1      0.568    0.023    24.839    0.000
   Category 2      0.432    0.023    18.866    0.000
```

x = 1 ; y = 1
We specified estimation of testbin only in latent variable y, so will consider the different y classes

Proportion of “pass” scores (cat.2) in testbin is 60% in y1 and 43% in y2

Binary distal variable in Mplus (ctd.)

- OUTPUT:

LATENT CLASS ODDS RATIO RESULTS

Latent Class Pattern 1 1 Compared to Latent Class Pattern 1 2

TESTBIN
Category > 1 **2.017** 0.222 9.081 *0.000*

This is the odds ratio, its SE, Est/SE, *p value*

This means that compared to individuals in class 2 of y , individuals in class 1 of y are 2.017 times more likely to have a TESTBIN score greater than category 1 than they are to have a score in category 1. Since there are only 2 categories and category 2 is the "pass" score: compared to individuals in y2 individuals in y1 are 2 times more likely to have a pass score than they are to have a fail score.

Binary distal variable in Mplus (ctd.)

- If you want to treat the distal binary outcome as a different variable (not a latent class indicator) some options available:
  - create a binary latent variable measured by the binary indicator (your outcome) without error, then regress this variable on the latent class of interest (the predictor)
  - create a binary latent variable measured by the binary indicator with error (a LC measurement model of your outcome), then regress this on the predictor latent class
- *These approaches are not encouraged*

Create a binary latent variable measured by the binary indicator (your outcome) without error, then regress this variable on the latent class of interest (the predictor)

VARIABLES: names are male a1 b1 c1 d1 a2 b2 c2 d2 testbin;
usevar are a1-d2 **testbin**;
categorical are a1-d2 **testbin**;
classes are x(2) y(2) **outcome** (2);
[...]

A latent binary variable is created

MODEL:
%overall%
y ON x;
outcome ON y;
[...]
MODEL OUTCOME:
%outcome#1%
[**testbin$1@15**];
%outcome#2%
[**testbin$1@-15**];

outcome regressed on y
(y → outcome)

the latent variable is measured without error:
P(testbin=1|outcome=1) = 1
P(testbin=1|outcome=2) = 0

Create a binary latent variable measured by the binary indicator (your outcome) without error, then regress this variable on the latent class of interest (the predictor)

OUTPUT:
[...]
Y Classes (Rows) by OUTCOME Classes (Columns)

	1	2
1	0.395	0.605
2	0.568	0.432

Probability that individuals in y1 display category 2 of OUTCOME (pass score) is 60% ; only 43% for individuals in y2

[...]
OUTCOME# ON
Y#1 -0.702 0.110 -6.371 0.000

[...]

Logistic regression coefficient of y→OUTCOME . Converted to an **odds ratio:** exp(-0.702) = 0.496 (its inverse = 2.016) create a binary latent variable measured by the binary indicator with error (a LC measurement model of your outcome), then regress this on the predictor latent class

```
VARIABLES are  male a1 b1 c1 d1 a2 b2 c2 d2 testbin;
        usevar are a1-d2 testbin;
        categorical are a1-d2 testbin;
        classes are x(2) y(2) outcome (2);
[...]

MODEL:
%overall%
y ON x;
outcome ON y;
[...]
 MODEL OUTCOME:
   %outcome#1%
[testbin$1*1];
   %outcome#2%
[testbin$1*-1];
```

A latent binary variable is created

outcome regressed on y (y → outcome)

the latent variable is measured with error:

Create a binary latent variable measured by the binary indicator (your outcome) without error, then regress this variable on the latent class of interest (the predictor)

OUTPUT:
[...]
Y Classes (Rows) by OUTCOME Classes (Columns)

	1	2
1	0.524	0.476
2	0.824	0.176

Probability that individuals in y1 display category 2 of OUTCOME (pass score) is 48% ; only 18% for individuals in y2

Latent Class Pattern 1 1 1
[...]
TESTBIN

Category 1	0.670	0.055	12.213	0.000
Category 2	0.330	0.055	6.002	0.000

[...]
Latent Class Pattern 1 1 2
TESTBIN

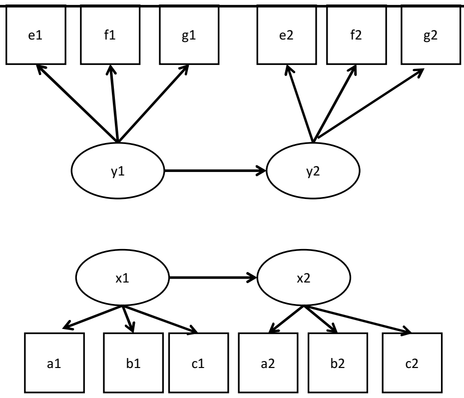
Category 1	0.091	0.041	2.212	0.027
Category 2	0.909	0.041	22.006	0.000

Individuals in class 1 of OUTCOME (fail) have 33% chance of passing test (substantial measurement error).
Individuals in class 2 (pass) of OUTCOME have 91% chance

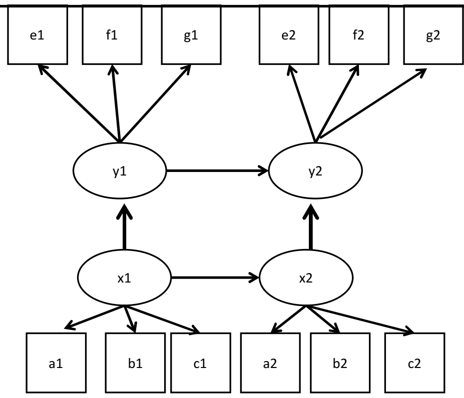
Further application

Further applications

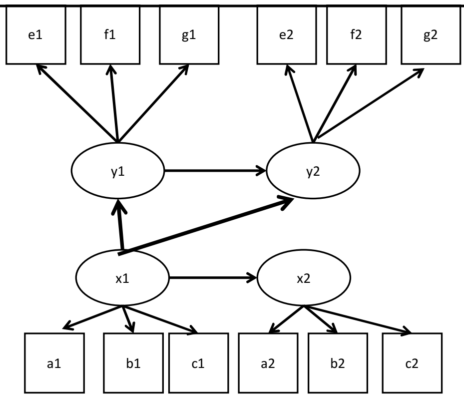
- Associative Latent Transition Analysis (ALTA):
  - Multiprocess model → examine change over time in two or more discrete developmental processes



independence



cross - sectional



baseline and lagged effects

References and resources

- A great resource to learn about stats in general:
  http://www.ats.ucla.edu/stat/
  Including examples from LCA textbooks:
  http://www.ats.ucla.edu/stat/mplus/examples/
- Mplus web page (visit the "*Mplus Web Notes*" and the "*Short Course Videos and Handouts*" pages for tutorials and examples)
  http://www.statmodel.com/

References and resources

- Nylund's dissertation on LTA (includes input files of some of the models tested):

http://www.statmodel.com/download/nylunddis.pdf

- Bray's dissertation on "advanced latent class modeling techniques" (also includes Mplus input files):

http://www.statmodel.com/download/Bray%20Dissertation%20%282007%29

References and resources

- Hagenaars, J.A & McCutcheon, A. (2002). Applied latent class analysis. Cambridge: Cambridge University Press.
- Langeheine, R. & van de Pol, F. (2002). Latent Markov chains. In Hagenaars, J.A. & McCutcheon, A.L. (eds.), Applied latent class analysis (pp. 304-341). Cambridge, UK: Cambridge University Press.
- Mooijaart, A. (1998). Log-linear and Markov modeling of categorical longitudinal data. In Bijleveld, C. C. J. H., & van der Kamp, T. (eds). Longitudinal data analysis: Designs, models, and methods. Newbury Park: Sage.

References and resources

- Chung, H., Park, Y., & Lanza, S.T. (2005). Latent transition analysis with covariates: pubertal timing and substance use behaviors in adolescent females. Statistics in Medicine, 24, 2895 - 2910.
- Collins, L.M. & Wugalter, S.E. (1992). Latent class models for stage sequential dynamic latent variables. Multivariate Behavioral Research, 27, 131-157.
- Collins, L.M., Graham, J.W., Rousculp, S.S., & Hansen, W.B. (1997). Heavy caffeine use and the beginning of the substance use onset process: An illustration of latent transition analysis. In K. Bryant, M. Windle, & S. West (Eds.), The science of prevention: Methodological advances from alcohol and substance use research. Washington DC: American Psychological Association. pp. 79-99.
- Kaplan, D. (2008). An overview of Markov chain methods for the study of stage-sequential developmental processes. Developmental Psychology, 44, 457-467.