Latent Variable Modeling Using Mplus: Day 2

Bengt Muthén & Linda Muthén

Mplus
www.statmodel.com

October, 2012

Table of Contents I

Table of Contents II

Table of Contents III

Table of Contents IV

Mplus Integrates A Multitude Of Analysis Types Using The Unifying Theme Of Latent Variables

- Exploratory factor analysis
- Structural equation modeling
- Item response theory analysis
- Growth modeling
- Latent class analysis
- Latent transition analysis (Hidden Markov modeling)
- Growth mixture modeling
- Survival analysis
- Missing data modeling
- Multilevel analysis
- Complex survey data analysis
- Bayesian analysis
- Causal inference

Mplus Integrates A Multitude Of Analysis Types Using The Unifying Theme Of Latent Variables

- Exploratory factor analysis
- Structural equation modeling

Mplus Integrates A Multitude Of Analysis Types Using The Unifying Theme Of Latent Variables

- Structural equation modeling

- Bayesian analysis

Mplus Integrates A Multitude Of Analysis Types Using The Unifying Theme Of Latent Variables

- Latent class analysis

- Causal inference

Mplus Integrates A Multitude Of Analysis Types Using The Unifying Theme Of Latent Variables

- Growth mixture modeling
- Survival analysis
- Missing data modeling

- Latent class analysis

Mplus Integrates A Multitude Of Analysis Types Using The Unifying Theme Of Latent Variables

- Survival analysis

- Latent class analysis

Used to capture heterogeneity when individuals come from different unobserved subpopulations

Application Areas

- Cross-sectional data
  - Medical and psychiatric diagnosis such as Alzheimer's disease, schizophrenia, depression, alcoholism
  - Market segmentation
  - Mastery in educational development
- Longitudinal data
  - Multiple disease processes such as prostate-specific antigen development
  - Developmental pathways such as adolescent-limited versus life-course persistent antisocial behavior

Analysis Methods

- Regression mixture models - Modeling of counts, randomized interventions with non-compliance
- Latent class analysis with and without covariates
- Latent transition analysis
- Latent class growth analysis
- Growth mixture modeling
- Survival mixture modeling

A Mixture Example: Biomarker For Alzheimer's Disease

Meyer et al. (2010). Diagnosis-independent Alzheimer disease biomarker signature in cognitively normal elderly people. Arch Neurol.

- Alzheimer's Disease (AD) Neuroimaging Initiative
- Adults aged 55 to 90 years
- 3 groups based on cognitive tests and clinical ratings:
  - Mild AD
  - Mild cognitive impairment
  - Cognitively normal
- Measure of cerebrospinal fluid-derived β-amyloid protein

Cerebrospinal Fluid Protein Distributions: A 2-Class Mixture

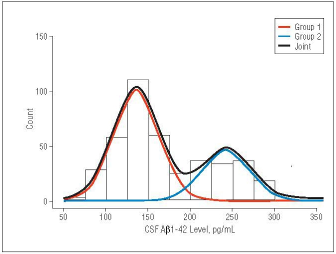


Figure 1. Mixture model classification for cerebrospinal fluid–derived β-amyloid protein 1-42 (CSF Aβ1-42). Results are presented as a histogram of observed counts overlaid with the 2 mixture distributions and the joined distribution based on the mixture proportion.

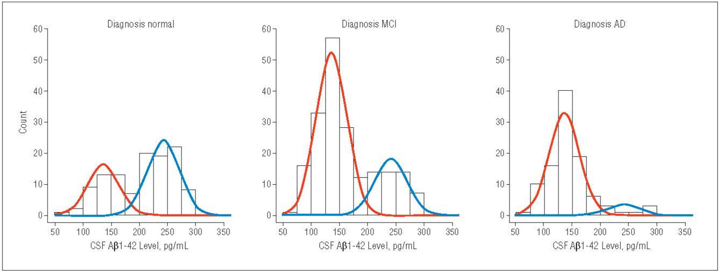
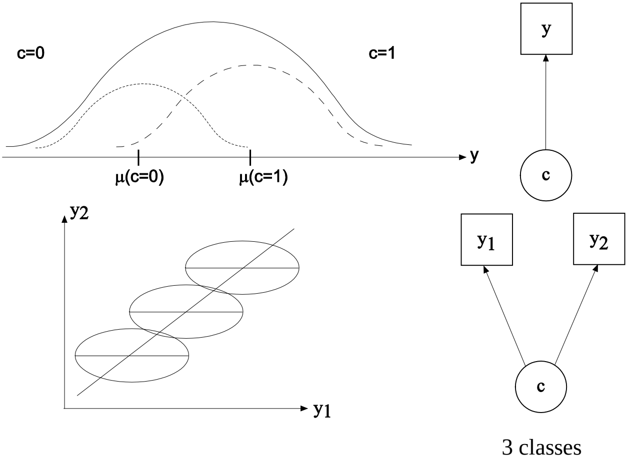
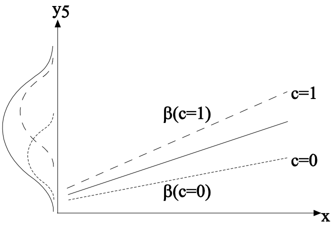
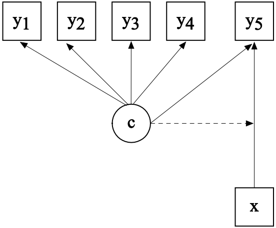


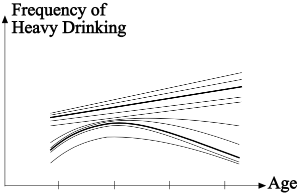
Figure 2. Cerebrospinal fluid–derived β-amyloid protein 1-42 (CSF Aβ1-42) mixture model applied to the clinically diagnosed subject groups. AD indicates Alzheimer disease; MCI, mild cognitive impairment.

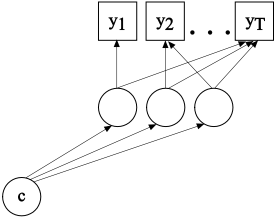

c=0
c=1
y
μ(c=0)
μ(c=1)
y
c
y2
y1
y1
y2
c
3 classes

Mixture Modeling Prototypes, Continued







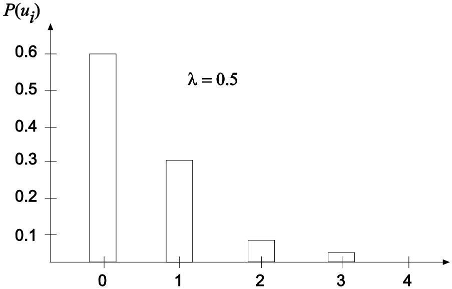


2. Regression With A Count Dependent Variable

- Poisson
- Zero-inflated Poisson (ZIP)
- Negative binomial
- Mixture ZIP
- ZI negative binomial
- Mixture negative binomial

A Poisson distribution for a count variable u_i has

$$P(u_i = r) = \frac{\lambda_i^r e^{-\lambda_i}}{r!}, \text{ where } u_i = 0, 1, 2, \ldots$$



λ is the rate at which a rare event occurs (rate = mean count)

Regression equation for the log rate:

$$^{e}log\ \lambda_i = ln\ \lambda_i = \beta_0 + \beta_1\ x_i$$

Zero-Inflated Poisson (ZIP) Regression

A Poisson variable has variance = mean = λ, but count data often have variance $>$ mean due to preponderance of zeros. Zero-inflated Poisson modeling avoids this restriction.

Alcohol abuse example: How many times in the last month did you drink 5 or more drinks at one occasion?

- Two classes of subjects: Drinkers and Non-drinkers
- A zero observation may be obtained because the subject is a non-drinker or because he/she is a drinker but did not drink 5 or more drinks at one occasion during the last month
- "Mixture at zero"

ZIP is a model with two latent classes:

- $\pi = P$ (being in the zero class where only $u = 0$ is seen)
- $1 - \pi = P$ (not being in the zero class with u following a Poisson distribution)

Zero-Inflated Poisson (ZIP) Regression, Continued

- A mixture at zero (π is the probability of being in the zero class):
    - $P(u=0) = \pi + (1-\pi)e^{-\lambda}$, where $e^{-\lambda}$ = Poisson for 0 count
- ZIP mean count: $\lambda(1-\pi)$
- ZIP variance: $\lambda(1-\pi)(1+\lambda \times \pi)$
- The ZIP model implies two regressions:

$$
\begin{aligned}
logit(\pi_i) &= \gamma_0 + \gamma_1\, x_i, \\
ln\, \lambda_i &= \beta_0 + \beta_1\, x_i
\end{aligned}
$$

2.2 Negative Binomial Regression

Unobserved heterogeneity e_i is added to the Poisson model

$$ln\ \lambda_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \text{ where } exp(\varepsilon) \sim \Gamma$$

Poisson assumes

$$E(u_i|x_i) = \lambda_i$$
$$V(u_i|x_i) = \lambda_i$$

Negative binomial assumes

$$E(u_i|x_i) = \lambda_i$$
$$V(u_i|x_i) = \lambda_i(1 + \lambda_i \alpha)$$

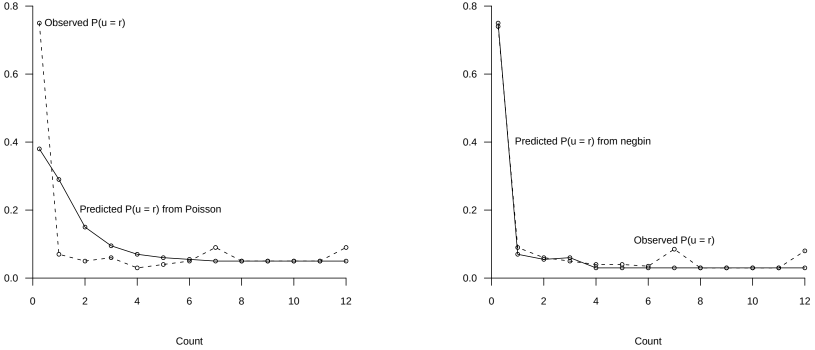
NB with $\alpha = 0$ gives Poisson. When the dispersion parameter $\alpha > 0$, the NB model gives substantially higher probability for low counts and somewhat higher probability for high counts than Poisson.

Further variations are zero-inflated NB and zero-truncated NB (hurdle model or two-part model).

Allowing any number of latent classes, not only a mixture at zero:

$$
\begin{aligned}
logit\ (\pi_i) &= \gamma_0 + \gamma_1\, x_i, \\
ln\ \lambda_{i|C=c_i} &= \beta_{0c} + \beta_{1c}\, x_i.
\end{aligned}
$$

An equivalent generalization of zero-inflated negative binomial is possible.



Dependent variable: Number of affairs reported in the last year

Covariates: Having kids, marital happiness, religiosity, years married

Source: Hilbe (2011). Negative Binomial Regression. Second Edition. Cambridge.

Model Alternatives For Counts Of Marital Affairs ($n = 601$)

Model	Log-Likelihood	# of Parameters	BIC
Poisson	-1,399.913	13	2883
Negative Binomial	-724.240	14	1538
Zero-inflated Poisson	-783.002	14	1656
Zero-inflated negative binomial	-718.064	15	1532
2-class Poisson mixture	-728.001	15	1552
2-class negative binomial mixture	-718.064	16	1539
2-class zero-inflated Poisson	**-700.718**	**16**	**1504**

Model Alternatives For Counts Of Marital Affairs (continued)

Model	Log-Likelihood	# of Parameters	BIC
2-class zero-inflated negative binomial	-700.718	17	1510
2-class negative binomial hurdle	-726.039	15	1548
Poisson with normal residual	-735.953	14	1561

Input For Two-Class ZIP Regression

```
TITLE:      Hilbe page 112 example
DATA:       FILE = affairs1.dat;
VARIABLE:   NAMES = ID
            male age yrsmarr kids relig educ occup ratemarr naffairs affair
            vryhap hapavg avgmarr unhap vryrel smerel slghtrel notrel;
            USEVAR = naffairs kids vryhap hapavg avgmarr vryrel smerel
            slghtrel notrel yrsmarr3 yrsmarr4 yrsmarr5 yrsmarr6;
            COUNT = naffairs(pi);
            CLASSES = c(2);
DEFINE:     IF (yrsmarr==4) THEN yrsmarr3=1 ELSE yrsmarr3=0;
            IF (yrsmarr==7) THEN yrsmarr4=1 ELSE yrsmarr4=0;
            IF (yrsmarr==10) THEN yrsmarr5=1 ELSE yrsmarr5=0;
            IF (yrsmarr==15) THEN yrsmarr6=1 ELSE yrsmarr6=0;
ANALYSIS:   TYPE = MIXTURE;
            ESTIMATOR = ML;
            PROCESSORS = 8;
```

Input For Two-Class ZIP Regression (Continued)

```
MODEL:          %OVERALL%
                naffairs ON kids-yrsmarr6 (p1-p12);
                ! It is also possible to model the logit of the probability
                ! of being in the zero class:
                ! naffairs#1 ON kids-yrsmarr6;
                ! Also possible: c ON kids-yrsmarr6;
MODEL CONSTRAINT:
                NEW(e1-e12);
                DO(1,12) e# = exp(p#);
OUTPUT:         TECH1;
```

3. Estimating Treatment Effects in Randomized Trials with Non-Compliance

- Angrist, Imbens, Rubin (1996). Identification of causal effects using instrumental variables. **Journal of the American Statistical Association**
- Little & Yau (1998). Statistical techniques for analyzing data from prevention trials: treatment of no-shows using Rubins causal model. **Psychological Methods**
- Jo (2002). Estimation of intervention effects with noncompliance: Alternative model specifications. **Journal of Educational and Behavioral Statistics**
- Jo, Asparouhov & Muthén (2008). Intention-to-treat analysis in cluster randomized trials with noncompliance. **Statistics in Medicine**

Potential outcomes, principal stratification, latent classes (mixtures)

Randomized Trials With Non-Compliance

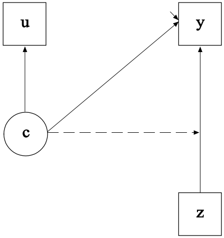
- Tx group (compliance status observed)
  - Compliers
  - Noncompliers
- Control group (compliance status unobserved)
  - Compliers
  - Noncompliers

Compliers and Noncompliers are typically not randomly equivalent subgroups.

Four approaches to estimating treatment effects:

1. Tx versus Control (Intent-To-Treat; ITT)
2. Tx Compliers versus Control (Per Protocol)
3. Tx Compliers versus Tx NonCompliers + Control (As-Treated)
4. Mixture analysis (Complier Average Causal Effect; CACE):
   - Tx Compliers versus Control Compliers
   - Tx NonCompliers versus Control NonCompliers

Causal Effect For Compliers (CACE) Via Mixture Modeling



- z is a 0/1 dummy variable indicating treatment assignment
- c is a latent class variable (Complier and Non-Complier)
- u is a categorical variable with categories Show and No-Show.
  - u is missing for the control group
  - u is identical to c for the treatment group (c observed for Tx)

JOBS Example

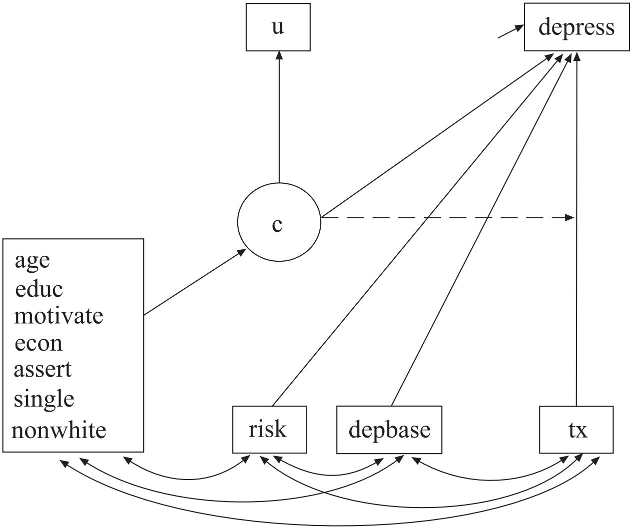
The JOBS data are from a Michigan University Prevention Research Center study of interventions aimed at preventing poor mental health of unemployed workers and promoting high quality of reemployment. The intervention consisted of five half-day training seminars that focused on problem solving, decision making group processes, and learning and practicing job search skills. The control group received a booklet briefly describing job search methods and tips. Respondents were recruited from the Michigan Employment Security Commission. After a series of screening procedures, 1801 were randomly assigned to treatment and control conditions. Of the 1249 in the treatment group, only 54% participated in the treatment.

The variables collected in the study include depression scores and outcome measures related to reemployment. Background variables include demographic and psychosocial variables.

JOBS Example, Continued

Data for the analysis include the outcome variable of depression and the background variables of treatment status, age, education, marital status, SES, ethnicity, a risk score for depression, a pre-intervention depression score, a measure of motivation to participate, and a measure of assertiveness. A subset of 502 individuals classified as having high-risk of depression were analyzed.

The analysis replicates that of Little & Yau (1998).


u
depress
c
age
educ
motivate
econ
assert
single
nonwhite
risk
depbase
tx

Input For JOBS Example

```
TITLE:      Complier Average Causal Effect (CACE) estimation
            in a randomized trial.
            Data from the JOBS II intervention trial, courtesy of
            Richard Price and Amiram Vinokur, University of Michigan.
            The analysis below replicates that of:
            Little, R.J. & Yau, L.H.Y. (1998). Statistical techniques
            for analyzing data from prevention trials:
            Treatment of no-shows using Rubin's causal model.
            Psychological Methods, 3, 147-159.
DATA:       FILE = jobs with u.dat;
VARIABLE:   NAMES = depress risk Tx depbase age motivate educ
            assert single econ nonwhite u;
            CATEGORICAL = u; ! u=1 show, u=0 no-show
            MISSING = ALL(999);
            CLASSES = c(2);
ANALYSIS:   TYPE = MIXTURE;
            STARTS = 100 20;
            PROCESSORS = 8;
```

Input For JOBS Example, Continued

```
MODEL:     %OVERALL%
           depress ON Tx risk depbase;
           c ON age educ motivate econ assert single nonwhite;
           %c#1%
           ! c#1 is the complier class (shows)
           [u$1@-15]; ! P(u = 1) = 1
           ! [depress]; different across class as the default
           %c#2%
           ! c#2 is the noncomplier class (no-shows)
           [u$1@15]; ! P(u = 1) = 0
           ! [depress]; different across class as the default
           depress ON Tx@0;
OUTPUT:    TECH1 TECH8;
```

Output For JOBS Example, Continued

Final Class Counts And Proportions For the Latent Classes Based On The Estimated Model

Latent Classes		
1	271.93480	0.54170
2	230.06520	0.45830
Classification Quality		
Entropy	0.727	

Average Latent Class Probabilities For Most Likely Latent Class Membership (Row) By Latent Class (Column)

Latent Classes		
	1	2
1	0.900	0.100
2	0.097	0.903

Output For JOBS Example, Continued

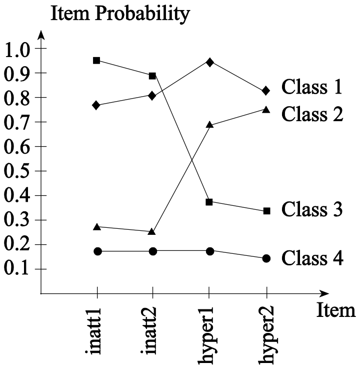
	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
Latent Class 1				
depress ON				
tx	-0.310	0.130	-2.378	0.017
risk	0.912	0.247	3.685	0.000
depbase	-1.463	0.181	-8.077	0.000
Intercepts				
depress	1.812	0.299	6.068	0.000
Thresholds				
u$1	-15.000	0.000	999.000	999.000
Residual Variances				
depress	0.506	0.037	13.742	0.000

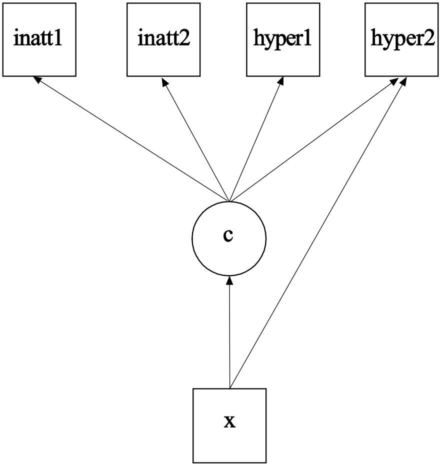
Output For JOBS Example, Continued

	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
Latent Class 2				
depress ON				
tx	0.000	0.000	999.000	999.000
risk	0.912	0.247	3.685	0.000
depbase	-1.463	0.181	-8.077	0.000
Intercepts				
depress	1.633	0.273	5.977	0.000
Thresholds				
u$1	15.000	0.000	999.000	999.000
Residual Variances				
depress	0.506	0.037	13.742	0.000

Output For JOBS Example, Continued

	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
Categorical Latent Variables				
c#1 ON				
age	0.079	0.015	5.184	0.000
educ	0.300	0.068	4.390	0.000
motivate	0.667	0.157	4.243	0.000
econ	-0.159	0.152	-1.045	0.296
assert	-0.376	0.143	-2.631	0.009
single	0.540	0.283	1.908	0.056
nonwhite	-0.499	0.317	-1.571	0.116
Intercepts				
c#1	-8.740	1.590	-5.498	0.000


Item Probability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Class 1
Class 2
Class 3
Class 4
inatt1
inatt2
hyper1
hyper2
Item


inatt1
inatt2
hyper1
hyper2
c
x

Latent Class Analysis
Alcohol Dependence Criteria, NLSY 1989 (n = 8313)

	Latent Classes				
	Two-class solution[1]		Three-class solution[2]		
	I	II	I	II	III
Prevalence	0.78	0.22	0.75	0.21	0.03
DSM-III-R-Criterion	Conditional Probability of Fulfilling a Criterion				
Withdrawal	0.00	0.14	0.00	0.07	0.49
Tolerance	0.01	0.45	0.01	0.35	0.81
Larger	0.15	0.96	0.12	0.94	0.99
Cut down	0.00	0.14	0.01	0.05	0.60
Time spent	0.00	0.19	0.00	0.09	0.65
Major role-Hazard	0.03	0.83	0.02	0.73	0.96
Give up	0.00	0.10	0.00	0.03	0.43
Relief	0.00	0.08	0.00	0.02	0.40
Continue	0.00	0.24	0.02	0.11	0.83

[1] Likelihood ratio chi-square fit = 1779 with 492 degrees of freedom

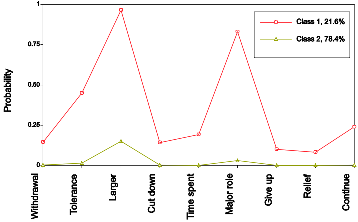
[2] Likelihood ratio chi-square fit = 448 with 482 degrees of freedom

Latent Class Membership By Number Of DSM-III-R Alcohol Dependence Criteria Met (n=8313)

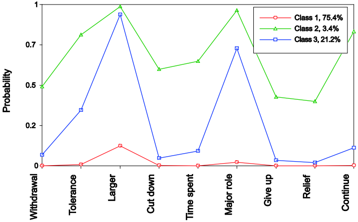
Number of Criteria Met	%	Two-class solution I	Two-class solution II	Three-class solution I	Three-class solution II	Three-class solution III
0	64.2	5335	0	5335	0	0
1	14.0	1161	1	1161	1	0
2	10.2	0	845	0	845	0
3	5.6	0	469	0	469	0
4	2.6	0	213	0	211	2
5	1.4	0	116	0	19	97
6	0.8	0	68	0	0	68
7	0.5	0	42	0	0	42
8	0.5	0	39	0	0	39
9	0.3	0	24	0	0	24
%	100.0	78.1	21.9	78.1	18.6	3.3

LCA Item Profiles For NLSY Alcohol Criteria

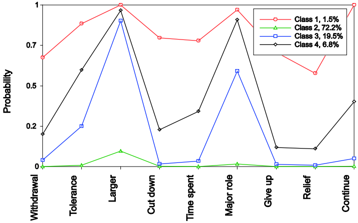
2-class LCA Item Profiles



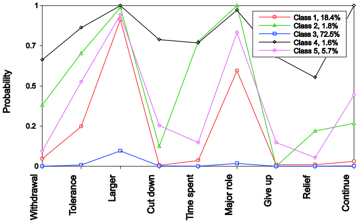
3-class LCA Item Profiles



4-class LCA Item Profiles



5-class LCA Item Profiles



Input For NLSY Alcohol LCA

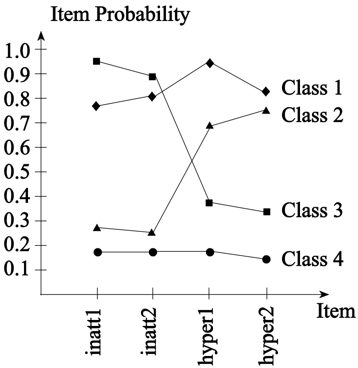
```
TITLE:      Alcohol LCA M & M (1993)
DATA:       FILE = bengt03_spread.dat;
VARIABLE:   NAMES = u1-u9;
            CATEGORICAL = u1-u9;
            CLASSES = c(3);
ANALYSIS:   TYPE = MIXTURE;
PLOT:       TYPE = PLOT3;
            SERIES = u1-u9(*);
```

Modeling With A Combination Of Continuous And Categorical Latent Variables

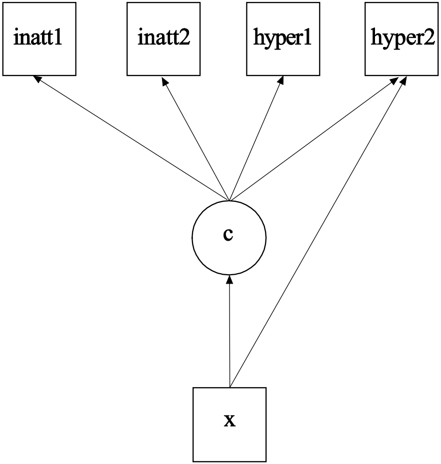
- Factor mixture analysis
  - Generalized factor analysis
  - Generalized latent class analysis

Muthén, B. (2008). Latent variable hybrids: Overview of old and new models. In Hancock, G. R., & Samuelsen, K. M. (Eds.), Advances in latent variable mixture models, pp. 1-24. Charlotte, NC: Information Age Publishing, Inc.

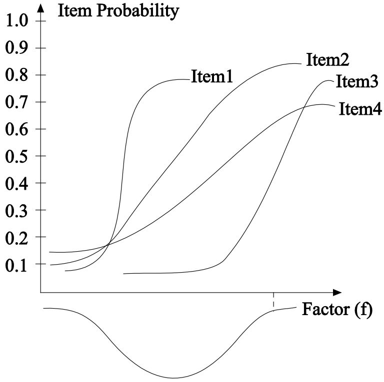
a. Item Profiles



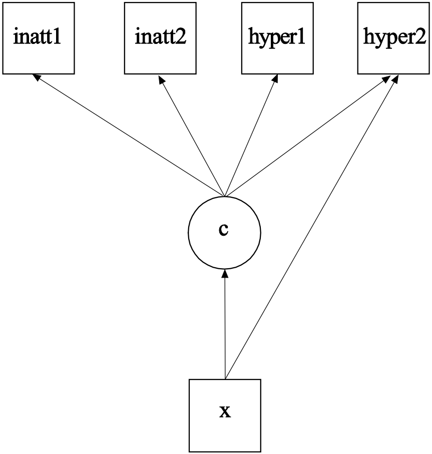
b. Model Diagram



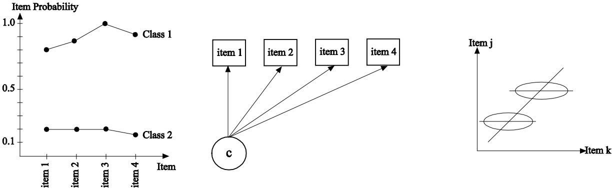
a. Item Response Curves



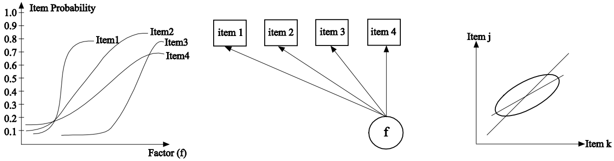
b. Model Diagram



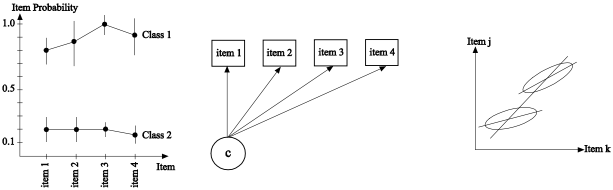
Latent Class Analysis



Factor Analysis (IRT)



Factor Mixture Analysis



Latent Class, Factor, And Factor Mixture Analysis Alcohol Dependence Criteria, NLSY 1989 (n = 8313)

	Latent Classes				
	Two-class solution[1]		Three-class solution[2]		
	I	II	I	II	III
Prevalence	0.78	0.22	0.75	0.21	0.03
DSM-III-R criterion	conditional probability of fulfilling a criterion				
Withdrawal	0.00	0.14	0.00	0.07	0.49
Tolerance	0.01	0.45	0.01	0.35	0.81
Larger	0.15	0.96	0.12	0.94	0.99
Cut down	0.00	0.14	0.01	0.05	0.60
Time spent	0.00	0.19	0.00	0.09	0.65
Major role-hazard	0.03	0.83	0.02	0.73	0.96
Give up	0.00	0.10	0.00	0.03	0.43
Relief	0.00	0.08	0.00	0.02	0.40
Continue	0.00	0.24	0.02	0.11	0.83

[1]Likelihood ratio chi-square fit = 1779 with 492 degrees of freedom

[2]Likelihood ratio chi-square fit = 448 with 482 degrees of freedom

LCA, FA, And FMA For NLSY 1989

- LCA, 3 classes: logL = -14,139, 29 parameters, BIC = 28,539
- FA, 2 factors: logL = -14,083, 26 parameters, BIC = 28,401
- FMA 2 classes, 1 factor, loadings invariant:

  logL = -14,054, 29 parameters, BIC = 28,370

  Models can be compared with respect to fit to the data using TECH10:
- Standardized bivariate residuals
- Standardized residuals for most frequent response patterns

Estimated Frequencies And Standardized Residuals

Obs. Freq.	LCA 3c		FA 2f		FMA 1f, 2c	
	Est. Freq.	Res.	Est. Freq.	Res.	Est. Freq.	Res.
5335	5332	-0.07	5307	-0.64	5331	-0.08
941	945	0.12	985	1.48	946	0.18
601	551	**-2.22**	596	-0.22	606	0.21
217	284	**4.04**	211	-0.42	228	0.75
155	111	**-4.16**	118	**-3.48**	134	1.87
149	151	0.15	168	1.45	147	0.17
65	68	0.41	46	**-2.79**	53	1.60
49	52	0.42	84	**3.80**	59	1.27
48	54	0.81	44	-0.61	46	0.32
47	40	-1.09	45	-0.37	45	0.33

Bolded entries are significant at the 5% level.

Input For FMA Of 9 Alcohol Items In The NLSY 1989

```
TITLE:      Alcohol LCA M & M (1995)
DATA:       FILE = bengt05_spread.dat;
VARIABLE:   NAMES = u1-u9;
            CATEGORICAL = u1-u9;
            CLASSES = c(2);
ANALYSIS:   TYPE = MIXTURE;
            ALGORITHM = INTEGRATION;
            STARTS = 200 10; STITER = 20;
            ADAPTIVE = OFF;
            PROCESSORS = 8;
```

Input For FMA Of 9 Alcohol Items In The NLSY 1989 (Continued)

```
MODEL:      %OVERALL%
            f BY u1-u9;
            f*1; [f@0];
            %c#1%
            [u1$1-u9$1];
            f*1;
            %c#2%
            [u1$1-u9$1];
            f*1;
OUTPUT:     TECH1 TECH8 TECH10;
PLOT:       TYPE = PLOT3;
            SERIES = u1-u9(*);
```

Factor (IRT) Mixture Example: The Latent Structure Of ADHD

- UCLA clinical sample of 425 males ages 5-18, all with ADHD diagnosis
- Subjects assessed by clinicians:
  1) direct interview with child ($>$ 7 years),
  2) interview with mother about child
- KSADS: Nine inattentiveness items, nine hyperactivity items; dichotomously scored
- Families with at least 1 ADHD affected child
- Parent data, candidate gene data on sib pairs
- What types of ADHD does a treatment population show?

The Latent Structure Of ADHD (Continued)

Inattentiveness items:	Hyperactivity items:
'Difficulty sustaining attn on tasks/play'	'Difficulty remaining seated'
'Easily distracted'	'Fidgets'
'Makes a lot of careless mistakes'	'Runs or climbs excessively'
'Doesn't listen'	'Difficulty playing quietly'
'Difficulty following instructions'	'Blurts out answers'
'Difficulty organizing tasks'	'Difficulty waiting turn'
'Dislikes/avoids tasks'	'Interrupts or intrudes'
'Loses things'	'Talks excessively'
'Forgetful in daily activities'	'Driven by motor'

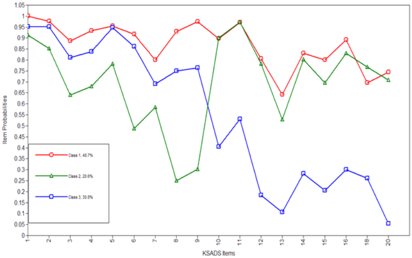
The Latent Structure Of ADHD: Model Results

Model	Likelihood	# Parameters	BIC	BLRT p value k-1 classes
LCA - 2c	-3650	37	7523	0.00
LCA - 3c	**-3545**	**56**	**7430**	**0.00**
LCA - 4c	-3499	75	7452	0.00
LCA - 5c	-3464	94	7496	0.00
LCA - 6c	**-3431**	**113**	**7547**	**0.00**
LCA - 7c	-3413	132	7625	0.27

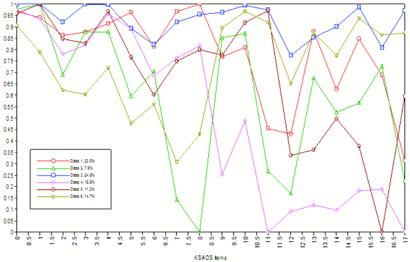
LCA-3c is best by BIC and LCA-6c is best by BLRT

Three-Class And Six-Class LCA Item Profiles

LCA- 3c



LCA -6c



The Latent Structure Of ADHD: Model Results

Model	Likelihood	# Parameters	BIC	BLRT p value k-1 classes
LCA - 2c	-3650	37	7523	0.00
LCA - 3c	**-3545**	**56**	**7430**	**0.00**
LCA - 4c	-3499	75	7452	0.00
LCA - 5c	-3464	94	7496	0.00
LCA - 6c	**-3431**	**113**	**7547**	**0.00**
LCA - 7c	-3413	132	7625	0.27
EFA - 2f	**-3505**	**53**	**7331**	

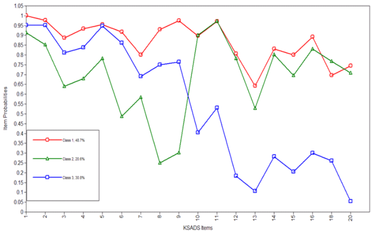
The EFA model is better than LCA-3c, but no classification of individuals is obtained

The Latent Structure Of ADHD: Model Results

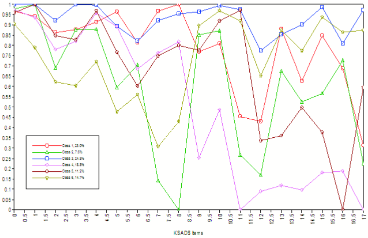
Model	Likelihood	# Parameters	BIC	BLRT p value k-1 classes
LCA - 2c	-3650	37	7523	0.00
LCA - 3c	**-3545**	**56**	**7430**	**0.00**
LCA - 4c	-3499	75	7452	0.00
LCA - 5c	-3464	94	7496	0.00
LCA - 6c	**-3431**	**113**	**7547**	**0.00**
LCA - 7c	-3413	132	7625	0.27
EFA - 2f	**-3505**	**53**	**7331**	
FMA - 2c, 2f	-3461	59	7280	
FMA - 2c, 2f Class-varying Factor loadings	**-3432**	**75**	**7318**	**χ^2-diff (16)=58 p $<$ 0.01**

Item Profiles For Three-Class LCA, Six-Class LCA And Two-Class, Two-Factor FMA

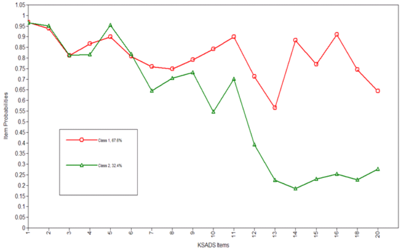
LCA- 3c

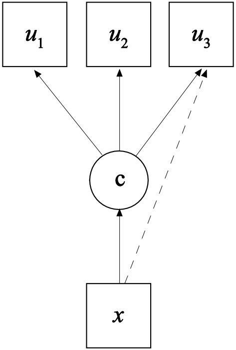


LCA- 6c



FMA- 2c, 2f




u_1
u_2
u_3
c
x

The Antisocial Behavior (ASB) data were taken from the National Longitudinal Survey of Youth (NLSY) that is sponsored by the Bureau of Labor Statistics. These data are made available to the public by Ohio State University. The data were obtained as a multistage probability sample with oversampling of blacks, Hispanics, and economically disadvantaged non-blacks and non Hispanics.

Data for the analysis include 17 antisocial behavior items that were collected in 1980 when respondents were between the ages of 16 and 23 and the background variables of age, gender, and ethnicity. The ASB items assessed the frequency of various behaviors during the past year. A sample of 7,326 respondents has complete data on the antisocial behavior items and the background variables of age, gender, and ethnicity.

Antisocial Behavior (ASB) Data (Continued)

Following is a list of the 17 items:

Property offense:	Person offense:	Drug offense:
Damaged property	Fighting	Use marijuana
Shoplifting	Use of force	Use other drugs
Stole $< \$50$	Seriously threaten	Sold marijuana
Stole $> \$50$	Intent to injure	Sold hard drugs
"Con" someone	Gambling operation	
Take auto		
Broken into building		
Held stolen goods		

The items were dichotomized 0/1 with 0 representing never in the last year.

Are there different groups of people with different ASB profiles?

Input For LCA Of 9 Antisocial Behavior (ASB) Items With Covariates

```
TITLE:      LCA of 9 ASB items with three covariates
DATA:       FILE = asb.dat;
            FORMAT = 34x 51f2;
VARIABLE:   NAMES = property fight shoplift lt50 gt50 force threat injure
            pot drug soldpot solddrug con auto bldg goods gambling dsm1-
            dsm22 male black hisp single divorce dropout college onset f1
            f2 f3 age94;
            USEVARIABLES = property fight shoplift lt50 threat pot drug
            con goods age94 male black;
            CLASSES = c(4);
            CATEGORICAL = property-goods;
ANALYSIS:   TYPE = MIXTURE;
```

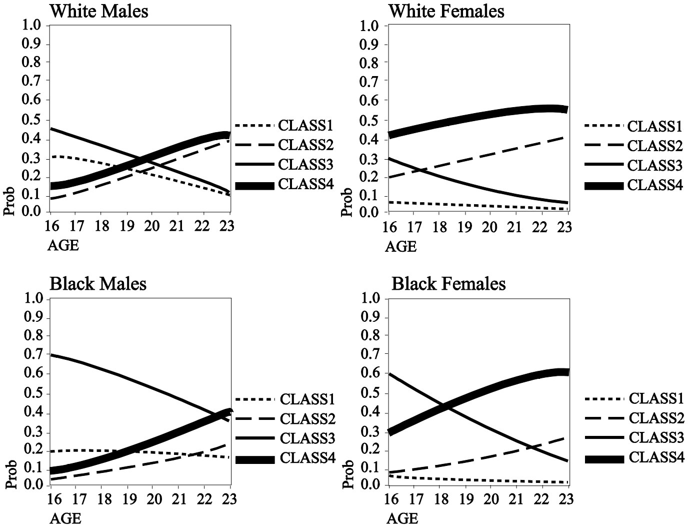
Input For LCA Of 9 Antisocial Behavior (ASB) Items With Covariates (continued)

```
MODEL:     %OVERALL%
           c#1-c#3 ON age94 male black;
           %c#1% !Not needed – High class
           [property$1-goods$1*0]; !Not needed
           %c#2% !Not needed – Drug class (pot, drugs)
           [property$1-goods$1*1]; !Not needed
           %c#3% !Not needed – Person class (fight, threaten)
           [property$1-goods$1*2]; !Not needed
           %c#4% !Not needed – Low class
           [property$1-goods$1*3]; !Not needed
OUTPUT:    TECH1 TECH8;
```

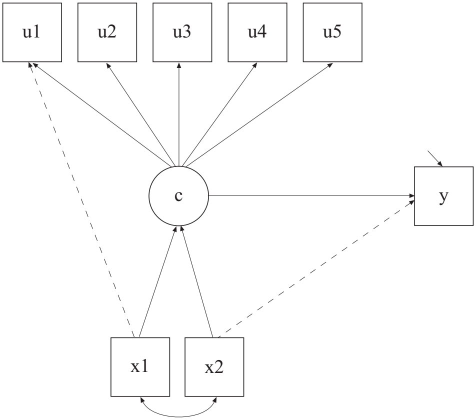
Output: Antisocial Behavior (ASB) Items With Covariates

Parameter	Estimate	S.E.	Est./S.E.
c#1 ON			
age94	-0.285	0.028	-10.045
male	2.578	0.151	17.086
black	0.158	0.139	1.141
c#2 ON			
age94	0.069	0.022	3.182
male	0.187	0.110	1.702
black	-0.606	0.139	-4.357
c#3 ON			
age94	-0.317	0.028	-11.311
male	1.459	0.101	14.431
black	0.999	0.117	8.513

ASB Classes Regressed On Age, Male, Black In The NLSY (n=7326)



1-step analysis versus 3-step analysis (analyze-classify-analyze)



Critique of 1-Step: Vermunt (2010)

*However, the one-step approach has certain disadvantages. **The first** is that it may sometimes be impractical, especially when the number of potential covariates is large, as will typically be the case in a more exploratory study. Each time that a covariate is added or removed not only the prediction model but also the measurement model needs to be reestimated. **A second** disadvantage is that it introduces additional model building problems, such as whether one should decide about the number of classes in a model with or without covariates. **Third**, the simultaneous approach does not fit with the logic of most applied researchers, who view introducing covariates as a step that comes after the classification model has been built. **Fourth**, it assumes that the classification model is built in the same stage of a study as the model used to predict the class membership, which is not necessarily the case.*

1-Step vs 3-Step: An Example

Substantive question: Should the latent classes be defined by the indicators alone or also by covariates and distals?

Example: Study of genotypes influencing phenotypes.

Phenotypes may be observed indicators of mental illness such as DSM criteria. The interest is in finding latent classes of subjects and then trying to see if certain genotype variables influence class membership.

Possible objection to 1-step: If the genotypes are part of deciding the latent classes, the assessment of the strength of relationship is compromised.

3-step: Determine the latent classes based on only phenotype information. Then classify subjects. Then relate the classification to the genotypes.

Problem Of 3-Step Approach Based On Most Likely Class Ignoring The Measurement Error

- Step 1: Do LCA on the latent class indicators
- Step 2: Classify subjects into most likely class
- Step 3: Regress the nominal most likely class variable on covariates

Problem: The Step 3 regression ignores the misclassification of the nominal observed variable being different from the latent class variable. This causes biased Step 3 estimates and SEs. The biases increase when entropy (classification quality: 0-1) decreases.

Auxiliary Variables In Mixture Modeling: The Correct 3-Step Approach

- Prior to Mplus Version 7: Pseudo-class (PC) approach. Estimate LCA model, impute C, regress imputed C on X
- New improved method in Mplus Version 7: 3-step approach
  1. Estimate the LCA model
  2. Create a nominal most likely class variable N
  3. Use a mixture model for N, C and X, where N is a C indicator with measurement error rates prefixed at the uncertainty rate of N estimated in the step 1 LCA analysis
- Mplus Web Note 15. Asparouhov and Muthén (2012). Auxiliary Variables in Mixture Modeling: A 3-Step Approach Using Mplus
- Vermunt (2010). Latent Class Modeling with Covariates: Two Improved Three-Step Approaches. Political Analysis, 18, 450-469

Auxiliary Variables In Mixture Modeling: Latent Class Predictor Example

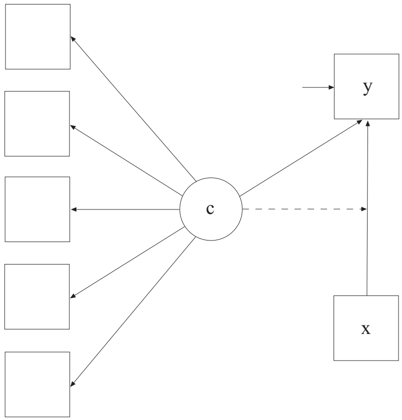
```
VARIABLE:   NAMES = u1-u5 x;
            CATEGORICAL = u1-u5;
            CLASSES = c(2);
            AUXILIARY = x(R3STEP);
DATA:       FILE = 1.dat;
ANALYSIS:   TYPE = MIXTURE;
MODEL:      !no model is needed, LCA is default
```

Auxiliary Variables In Mixture Modeling: Latent Class Predictor Example

TESTS OF CATEGORICAL LATENT VARIABLE MULTINOMIAL LOGISTIC REGRESSIONS
THE 3-STEP PROCEDURE

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
C#1 ON				
X	0.488	0.190	2.569	0.010

- The latent class variable can be identified by any mixture model, not just LCA, for example Growth Mixture Models
- Multiple auxiliary variables can be analyzed at the same time
- Auxiliary variables can be included in a Montecarlo setup
- The 3-step procedure can be setup manually for other types of models, different from the distal outcome model and the latent class regression. For example, distal outcomes regressed on the latent class variable and another predictor


y
c
x

3-Step Mixture Modeling For Special Models, Continued

- How can we estimate a mixture regression model independently of the LCA model that defines C

$$Y = \alpha_c + \beta_c X + \varepsilon$$

- We simulate data with $\alpha_1 = 0$, $\alpha_2 = 1$, $\beta_1 = 0.5$, $\beta_2 = -0.5$
- Step 1: Estimate the LCA model (without the auxiliary model) with the following option
  SAVEDATA: FILE=1.dat; SAVE=CPROB;
- The above option creates the most likely class variable N
- Step 2: Compute the error rate for N. In the LCA output find

```
Average Latent Class Probabilities for Most Likely Latent Class Membership
by Latent Class (Column)

              1          2

     1    0.835      0.165
     2    0.105      0.895
```

- Compute the nominal variable N parameters

$$\log(0.835/0.165) = 1.621486$$

$$\log(0.105/0.895) = -2.14286$$

- Step 3: estimate the model where N is a latent class indicator with the above fixed parameters and include the class specific Y on X model
- When the class separation in the LCA is pretty good then N is almost a perfect C indicator

```
VARIABLE:   NAMES = u1-u5 y x p1 p2 n;
            NOMINAL = n;
            CLASSES = c(2);
            USEVARIABLES = y x n;
MODEL:
            %OVERALL%
            y ON x;
            %c#1%
            [n#1@1.621486];
            y ON x;
            %c#2%
            [n#1@-2.14286];
            y ON x;
```

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Latent Class 1				
Y ON				
X	0.546	0.054	10.185	0.000
Means				
N#1	1.621	0.000	999.000	999.000
Intercepts				
Y	0.101	0.052	1.961	0.050
Latent Class 2				
Y ON				
X	-0.483	0.051	-9.496	0.000
Means				
N#1	-2.143	0.000	999.000	999.000
Intercepts				
Y	1.037	0.056	18.498	0.000

7.2 Rules of Thumb: 1-Step Versus 3-Step Versus Most Likely Class Ignoring The Measurement Error

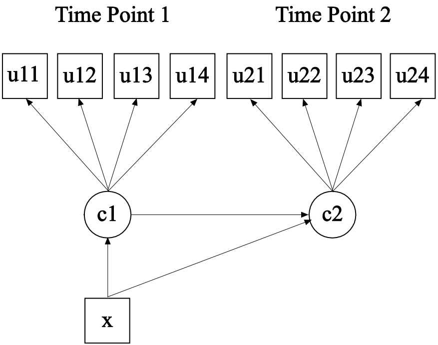
Consider a latent class model with covariates that is correctly specified. Choosing among 1-step, 3-step, and 3-step using Most likely class and ignoring misclassification error, the best approach to use depends to a large extent on the entropy (classification quality: 0-1):

- Entropy < 0.6: Use 1-step. 3-step and Most likely class don't work well
- $0.6 <$ Entropy < 0.8: 1-step and 3-step work well, but not Most likely class
- Entropy > 0.8: All three approaches work well

3-step needs large sample size ($n > 500$?); see Mplus Web Note 15.

Transition Probabilities

		c2: 1	c2: 2
c1	1	0.8	0.2
	2	0.4	0.6



Steps In Latent Transition Analysis

- Step 1: Study measurement model alternatives for each time point
- Step 2: Explore transitions based on cross-sectional results
  - Cross-tabs based on most likely class membership
- Step 3: Explore specification of the latent transition model without covariates
  - Testing for measurement invariance across time
- Step 4: Include covariates in the LTA model
- Step 5: Include distal outcomes and advanced modeling extensions

  Source: Nylund (2007)

8.1.1 LTA Example 1: Stage-Sequential Development in Reading Using ECLS-K Data

Kaplan (2008). An overview of Markov chain methods for the study of stage-sequential developmental processes. Developmental Psychology, 44, 457-467.

- Early Childhood Longitudinal Study - Kindergarten cohort
- Four time points: Kindergarten Fall, Spring and Grade 1 Fall, Spring; $n = 3,575$
- Five dichotomous proficiency scores: Letter recognition, beginning sounds, ending letter sounds, sight words, words in context
- Binary poverty index
- LCA suggests 3 classes: Low alphabet knowledge (LAK), early word reading (EWR), and early reading comprehension (ERC)

LTA Example 1: ECLS-K, Continued

- Three latent classes:
  - Class 1: Low alphabet knowledge (LAK)
  - Class 2: Early word reading (EWR)
  - Class 3: Early reading comprehension (ERC)

The ECLS-K LTA model has the special feature of specifying no decline in knowledge as zero transition probabilities. For example, transition from Kindergarten Fall to Spring:

LATENT TRANSITION PROBABILITIES
BASED ON THE ESTIMATED MODEL

c1 classes (rows) by c2 classes (columns)

	1	2	3
1	0.329	0.655	0.017
2	0.000	0.646	0.354
3	0.000	0.000	1.000

LTA Example 1: ECLS-K. Transition Tables for the Binary Covariate Poverty

		Poverty = 0 (cp=1) c2			Poverty = 1 (cp=2) c2		
		1	2	3	1	2	3
	1	0.252	0.732	0.017	0.545	0.442	0.013
c1	2	0.000	0.647	0.353	0.000	0.620	0.380
	3	0.000	0.000	1.000	0.000	0.000	1.000

8.1.2 LTA Example 2: Mover-Stayer LTA Modeling of Peer Victimization During Middle School

Nylund (2007) Doctoral dissertation: Latent Transition Analysis: Modeling Extensions and an Application to Peer Victimization

- Student's self-reported peer victimization in Grade 6, 7, and 8
- Low SES, ethnically diverse public middle schools in the Los Angeles area (11% Caucasian, 17% Black, 48 % Latino, 12% Asian)
- $n = 2045$
- 6 binary items: Picked on, laughed at, called bad names, hit and pushed around, gossiped about, things taken or messed up (Neary & Joseph, 1994 Peer Victimization Scale)

LTA Example 2: Mover-Stayer Model

- Class 1: Victimized (G6-G8: 19%, 10%, 8%)
- Class 2: Sometimes victimized (G6-G8: 34%, 27%, 21%)
- Class 3: Non-victimized (G6-G8: 47%, 63%, 71%)

Movers (60%)							
	c2 (Grade 7)				c3 (Grade 8)		
	0.29	0.45	0.26		0.23	0.59	0.18
c1	0.06	0.44	0.51	c2	0.04	0.47	0.49
(Grade 6)	0.04	0.46	0.55	(Grade 7)	0.06	0.17	0.77

Stayers (40%)							
	c2 (Grade 7)				c3 (Grade 8)		
	1	0	0		1	0	0
c1	0	1	0	c2	0	1	0
(Grade 6)	0	0	1	(Grade 7)	0	0	1

8.2 Latent Transition Analysis: Review of Logit Parameterization

Consider the logit parameterization for CLASSES = c1(3) c2(3):

		c2		
		1	2	3
c1	1	a1 + b11	a2 + b21	0
	2	a1 + b12	a2 + b22	0
	3	a1	a2	0

where each row shows the logit coefficients for a multinomial logistic regression of c2 on c1 with the last c2 class as reference class.

Zero lower-triangular probabilities are obtained by fixing the a1, a2, and b12 parameters at the logit value -15. The parameters b11, b21, and b22 are estimated.

LTA Example 1: ECLS-K, Mplus Input

```
TITLE:       LTA of Kindergarten Fall and Spring (3 x 3)
DATA:        FILE = dp.analytic.dat;
             FORMAT = f1.0, 20f2.0;
VARIABLE:    NAMES = pov letrec1 begin1 ending1 sight1 wic1
             letrec2 begin2 ending2 sight2 wic2
             letrec3 begin3 ending3 sight3 wic3
             letrec4 begin4 ending4 sight4 wic4;
             USEVARIABLES = letrec1 begin1 ending1 sight1 wic1
             letrec2 begin2 ending2 sight2 wic2;
             ! letrec3 begin3 ending3 sight3 wic3
             ! letrec4 begin4 ending4 sight4 wic4;
             CATEGORICAL = letrec1 begin1 ending1 sight1 wic1
             letrec2 begin2 ending2 sight2 wic2;
             ! letrec3 begin3 ending3 sight3 wic3
             ! letrec4 begin4 ending4 sight4 wic4;
             CLASSES = c1(3) c2(3);
             MISSING = .;
```

LTA Example 1: ECLS-K, Mplus Input, Continued

```
ANALYSIS:   TYPE = MIXTURE;
            STARTS = 400 80;
            PROCESSORS = 8;
MODEL:      %OVERALL%
            ! fix lower triangular transition probabilities = 0:
            [c2#1@-15 c2#2@-15]; ! fix a1 = a2 = -15
            c2#1 ON c1#2@-15; ! fix b12 = -15
            c2#1 ON c1#1*15; ! b11: start at 15 to make total logit start=0
            c2#2 ON c1#1-c1#2*15; ! b21, b22
```

LTA Example 1: ECLS-K, Mplus Input, Continued

```
MODEL c1:     %c1#1%
              [letrec1$1-wic1$1] (1-5) ;
              %c1#2%
              [letrec1$1-wic1$1] (6-10);
              %c1#3%
              [letrec1$1-wic1$1] (11-15);
MODEL c2:     %c2#1%
              [letrec2$1-wic2$1] (1-5);
              %c2#2%
              [letrec2$1-wic2$1] (6-10);
              %c2#3%
              [letrec2$1-wic2$1] (11-15);
OUTPUT:       TECH1 TECH15;
PLOT:         TYPE = PLOT3;
              SERIES = letrec1-wic1(*) | letrec2-wic2(*);
```

Output: Latent Transition Table

LATENT TRANSITION PROBABILITIES BASED ON THE ESTIMATED MODEL

c1 classes (rows) by c2 classes (columns)

	1	2	3
1	0.329	0.655	0.017
2	0.000	0.646	0.354
3	0.000	0.000	1.000

- TECH15 output with conditional class probabilities useful for studying transition probabilities with an observed binary covariate such as treatment/control or a latent class covariate
- LTA transition probability calculator for continuous covariates
- Probability parameterization to simplify input for Mover-Stayer LTA and other models with restrictions on the transition probabilities

8.3.1 TECH15 For LTA Example 1, ECLS-K: Mplus Input, Adding Poverty As Knownclass

```
               CLASSES = cp(2) c1(3) c2(3);
               KNOWNCLASS = cp(pov=0 pov=1);
ANALYSIS:      TYPE = MIXTURE;
               STARTS = 400 80; PROCESSORS = 8;
MODEL:         %OVERALL%
               c1 ON cp;
               [c2#1@-15 c2#2@-15];
               c2#1 ON c1#2@-15;
MODEL cp:      %cp#1%
               c2#1 ON c1#1*15;
               c2#2 ON c1#1-c1#2*15;
               %cp#2%
               c2#1 ON c1#1*15;
               c2#2 ON c1#1-c1#2*15;
MODEL c1:      %c1#1% etc as before
```

$P(CP=1)=0.808$
$P(CP=2)=0.192$

$P(C1=1|CP=1)=0.617$
$P(C1=2|CP=1)=0.351$
$P(C1=3|CP=1)=0.032$

$P(C1=1|CP=2)=0.872$
$P(C1=2|CP=2)=0.123$
$P(C1=3|CP=2)=0.005$

$P(C2=1|CP=1,C1=1)=0.252$
$P(C2=2|CP=1,C1=1)=0.732$
$P(C2=3|CP=1,C1=1)=0.017$

$P(C2=1|CP=1,C1=2)=0.000$
$P(C2=2|CP=1,C1=2)=0.647$
$P(C2=3|CP=1,C1=2)=0.353$

TECHNICAL 15 Output, Continued

$$P(C2=1|CP=1,C1=3)=0.000$$
$$P(C2=2|CP=1,C1=3)=0.000$$
$$P(C2=3|CP=1,C1=3)=1.000$$

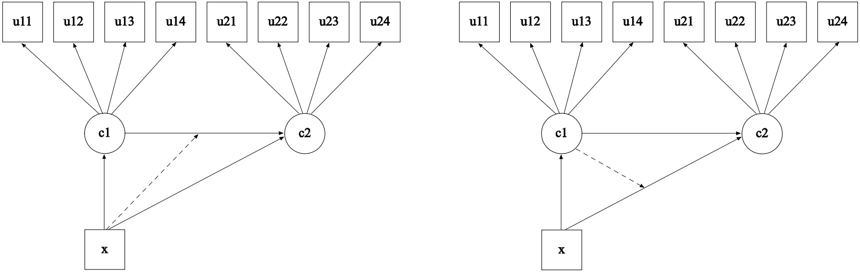
$$P(C2=1|CP=2,C1=1)=0.545$$
$$P(C2=2|CP=2,C1=1)=0.442$$
$$P(C2=3|CP=2,C1=1)=0.013$$

$$P(C2=1|CP=2,C1=2)=0.000$$
$$P(C2=2|CP=2,C1=2)=0.620$$
$$P(C2=3|CP=2,C1=2)=0.380$$

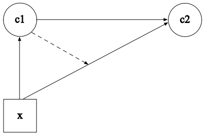
$$P(C2=1|CP=2,C1=3)=0.000$$
$$P(C2=2|CP=2,C1=3)=0.000$$
$$P(C2=3|CP=2,C1=3)=1.000$$

8.3.2 Latent Transition Probabilities Influenced By A Continuous Covariate

- Muthén & Asparouhov (2011). LTA in Mplus: Transition probabilities influenced by covariates. Mplus Web Notes: No. 13. July 27, 2011. www.statmodel.com
- New feature in Version 7: The LTA calculator


u11
u12
u13
u14
u21
u22
u23
u24
c1
c2
x
u11
u12
u13
u14
u21
u22
u23
u24
c1
c2
x

Review of Logit Parameterization with Covariates: Parameterization 2



		c2		
		1	2	3
c1	1	a1 + b11 + g11 x	a2 + b21 + g21 x	0
	2	a1 + b12 + g12 x	a2 + b22 + g22 x	0
	3	a1 + g13 x	a2 + g23 x	0

```
MODEL:      %OVERALL%
            c1 ON x;
            c2 ON c1;

MODEL c1:   %c1#1%
            c2#1 ON x (g11);
            c2#2 ON x (g21);

            %c1#2%
            c2#1 ON x (g12);
            c2#2 ON x (g22);

            %c1#3%
            c2#1 ON x (g13);
            c2#2 ON x (g23);
```

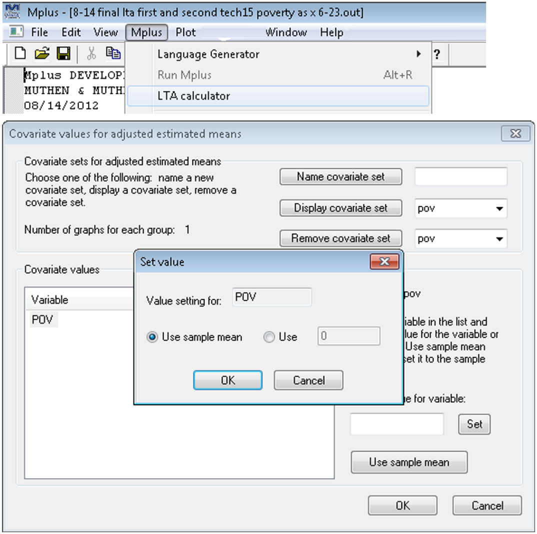
LTA Example 1: ECLS-K, Adding Poverty as Covariate

```
              USEVARIABLES = letrec1 begin1 ending1 sight1 wic1
              letrec2 begin2 ending2 sight2 wic2 pov;
              ! letrec3 begin3 ending3 sight3 wic3
              ! letrec4 begin4 ending4 sight4 wic4;
              CATEGORICAL = letrec1 begin1 ending1 sight1 wic1
              letrec2 begin2 ending2 sight2 wic2;
              ! letrec3 begin3 ending3 sight3 wic3
              ! letrec4 begin4 ending4 sight4 wic4;
              CLASSES = c1(3) c2(3);
              MISSING = .;
ANALYSIS:     TYPE = MIXTURE;
              STARTS = 400 80;
              PROCESSORS = 8;
```

LTA Example 1: ECLS-K, Adding Poverty as Covariate, Continued

```
MODEL:      %OVERALL%
            c1 ON pov;
            ! do c2 ON pov in c1-specific model part to get interaction
            [c2#1@-15 c2#2@-15]; ! to give zero probability of declining
            c2#1 ON c1#2@-15; ! to give zero probability of declining
            c2#1 ON c1#1*15;
            c2#2 ON c1#1-c1#2*15;

MODEL c1:   %c1#1%
            c2 ON pov; ! (g11) and (g21)
            %c1#2%
            c2#1 ON pov@-15; ! to give zero probability of declining (g12)
            c2#2 ON pov; ! (g22)

            ! %c1#3% not mentioned due to g13=0, g23=0 by default
```

LTA Calculator Applied to Poverty



LTA Calculator Applied to Poverty, Continued

Estimated conditional probabilities for the latent class variables:

Condition(s):	POV = 1.000000
	P(C1=1)=0.872
	P(C1=2)=0.123
	P(C1=3)=0.005
	P(C2=1|C1=1)=0.545
	P(C2=2|C1=1)=0.442
	P(C2=3|C1=1)=0.013
	P(C2=1|C1=2)=0.000
	P(C2=2|C1=2)=0.620
	P(C2=3|C1=2)=0.380
	P(C2=1|C1=3)=0.000
	P(C2=2|C1=3)=0.000
	P(C2=3|C1=3)=1.000

8.3.3 Probability Parameterization

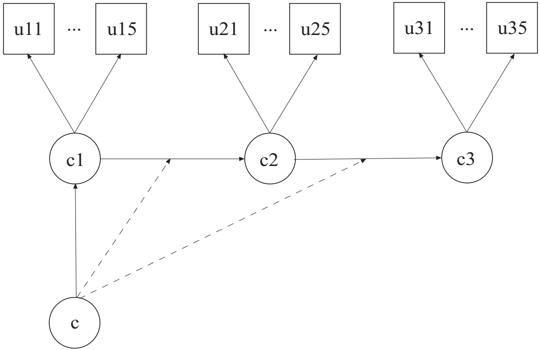
- New feature in Mplus Version 7
- LTA models that do not have continuous x's can be more conveniently specified using
  PARAMETERIZATION=PROBABILITY
  to reflect hypotheses expressed in terms of probabilities
- Useful for Mover-Stayer LTA models

Probability parameterization for CLASSES = c1(3) c2(3):

		c2		
		1	2	3
c1	1	p11	p12	0
	2	p21	p22	0
	3	p31	p32	0

where the probabilities in each row add to 1 and the last c2 class is not mentioned. The p parameters are referred to using ON. The latent class variable c1 which is the predictor has probability parameters [c1#1 c1#2], whereas "intercept" parameters are not included for c2.

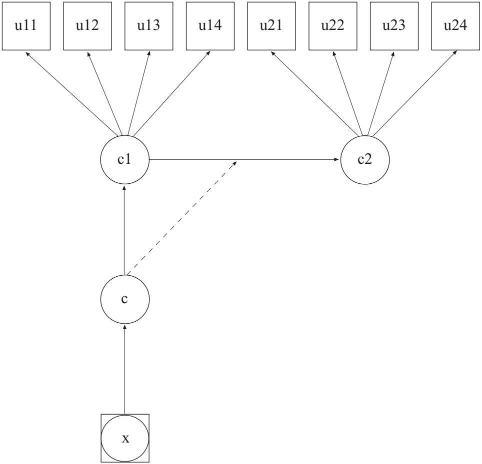
A transition probability can be conveniently fixed at 1 or 0 by using the p parameters.


u11
...
u15
u21
...
u25
u31
...
u35
c1
c2
c3
c

Mover-Stayer LTA in Probability Parameterization

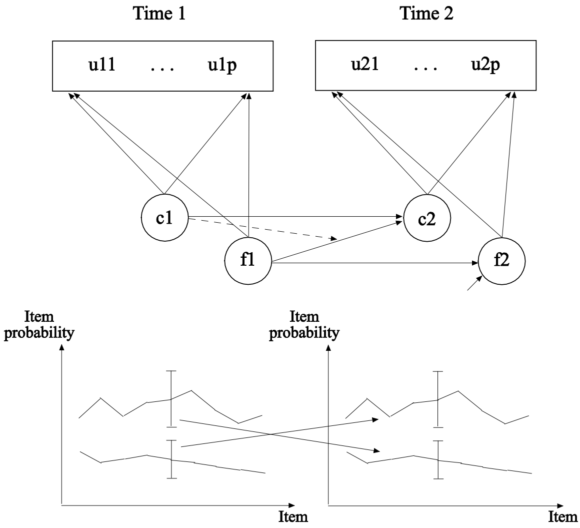
```
ANALYSIS:    PARAMETERIZATION = PROBABILITY;
MODEL:       %OVERALL% ! Relating c1 to c:
             c1 ON c;
MODEL c:     %c#1% ! Mover class
             c2 ON c1;
             c3 ON c2;
             %c#2% ! Stayer class
             c2#1 ON c1#1@1; c2#2 ON c1#1@0;
             c2#1 ON c1#2@0; c2#2 ON c1#2@1;
             c2#1 ON c1#3@0; c2#2 ON c1#3@0;

             c3#1 ON c2#1@1; c3#2 ON c2#1@0;
             c3#1 ON c2#2@0; c3#2 ON c2#2@1;
             c3#1 ON c2#3@0; c3#2 ON c2#3@0;
             !measurement part as before
```

Mover-Stayer LTA in Probability Parameterization: Predicting Mover-Stayer Class Membership From A Nominal Covariate



```
VARIABLE:       CLASSES = cg(5) c (2) c1(3) c2(3) c3(3);
                KNOWNCLASS = cg(eth=0 eth=1 eth=2 eth=3 eth=4);
ANALYSIS:       TYPE = MIXTURE COMPLEX;
                STARTS = 400 100;
                PROCESS = 8;
                PARAMETERIZATION = PROBABILITY;
MODEL:          %OVERALL%
                c ON cg#1-cg#5 (b1-b5);
                c1 ON c;
MODEL c:        etc
MODEL CONSTRAINT:
                NEW(logor2-logor5);
                ! log of ratio of odds of being Mover vs Stayer for the groups
                logor2 = log((b2/(1-b2))/(b1/(1-b1))); ! eth=1 (cg=2) vs 0 (cg=1)
                logor3 = log((b3/(1-b3))/(b1/(1-b1))); ! eth=2 (cg=3) vs 0
                logor4 = log((b4/(1-b4))/(b1/(1-b1))); ! eth=3 (cg=4) vs 0
                logor5 = log((b5/(1-b5))/(b1/(1-b1))); ! eth=4 (cg=5) vs 0
```

8.4 Latent Transition Analysis Extensions: Factor Mixture Latent Transition Analysis Muthén (2006)



Factor Mixture Latent Transition Analysis: Aggressive-Disruptive Behavior In The Classroom

- 1,137 first-grade students in Baltimore public schools
- 9 items: Stubborn, Break rules, Break things, Yells at others, Takes others property, Fights, Lies, Teases classmates, Talks back to adults
- Skewed, 6-category items; dichotomized (almost never vs other)
- Two time points: Fall and Spring of Grade 1
- For each time point, a 2-class, 1-factor FMA was found best fitting

Factor Mixture Latent Transition Analysis: Aggressive-Disruptive Behavior In The Classroom (Continued)

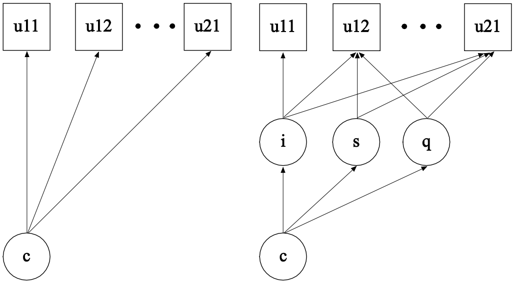
Model	Loglikelihood	# parameters	BIC
conventional LTA	-8,649	21	17,445
LTA			
FMA LTA factors related across time	-8,102	40	16,306

Factor Mixture Latent Transition Analysis: Aggressive-Disruptive Behavior In The Classroom (Continued)

Estimated latent transition probabilities, fall to spring

conventional LTA		
	low	high
low	0.93	0.07
high	0.17	0.83
FMA-LTA		
	low	high
low	0.94	0.06
high	**0.41**	**0.59**

9. Latent Class Growth Analysis: LCA vs LCGA



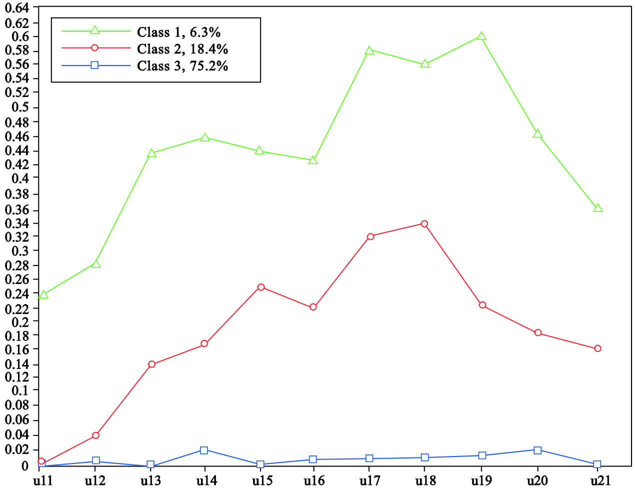
Example: Number of parameters for 11 u's and 3 classes:

	LCA	LCGA
binary u	35	11
3-categ. u	68	12

9.1 Single Process Latent Class Growth Analysis: Cambridge Delinquency Data

- 411 boys in a working class section of London
  (n = 403 due to 8 boys who died)
- Ages 10 to 32 (ages 11 - 21 used here)
- Outcome is number of convictions in the last 2 years, modeled as an ordered polytomous variable scored 0 for 0 convictions, 1 for one conviction, and 2 for more than one conviction

  Sources: Farrington & West (1990); Nagin & Land (1993); Roeder, Lynch & Nagin (1999); Muthén (2004)

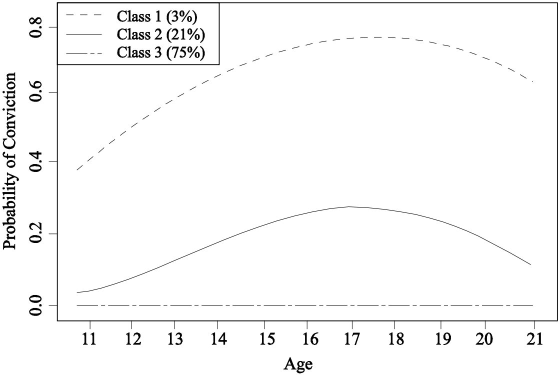


LogL = -1,032 (68 parameters), BIC = 2,472

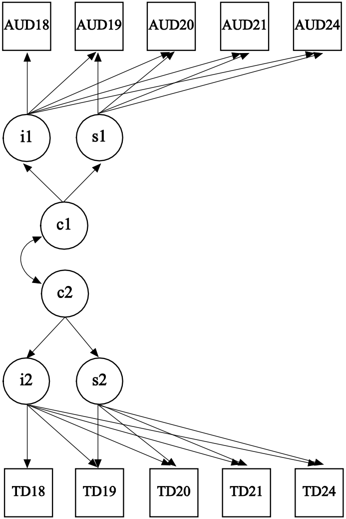
Input LCGA On Cambridge Data

```
TITLE:      LCGA ordered polytomous variables for conviction at each age11-21
            dep. variable 0, 1 , 2 (0, 1, or more convictions)
DATA:       FILE = naginordered.dat;
VARIABLE:   NAMES = u11 u12 u13 u14 u15 u16 u17 u18 u19 u20 u21 c1 c2 c3
            c4;
            USEVAR = u11-u21;
            CATEGORICAL = u11-u21;
            CLASSES = c(3);
ANALYSIS:   TYPE = MIXTURE;
MODEL:      %OVERALL%
            i s q | u11@-.6 u12@-.5 u13@-.4 u14@-.3 u15@-.2 u16@-.1 u17@0
            u18@.1 u19@.2 u20@.3 u21@.4;
OUTPUT:     TECH1 TECH8;
PLOT:       SERIES = u11-u21(s);
            TYPE = PLOT3;
```

- 3-class LCGA
  - LogL = -1,072
  - 12 parameters
  - BIC = 2,215
- 3-class LCA
  - LogL = -1,032
  - (68 parameters)
  - BIC = 2,472

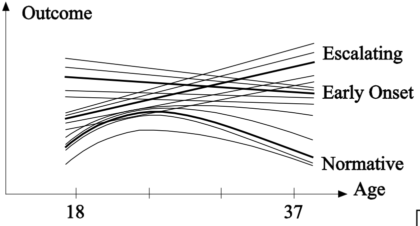


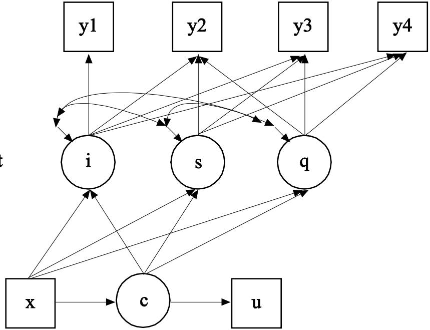
Co-Occurrence Of Alcohol And Tobacco Use Disorder



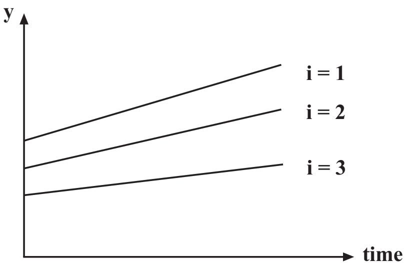
Class-Probability Estimates

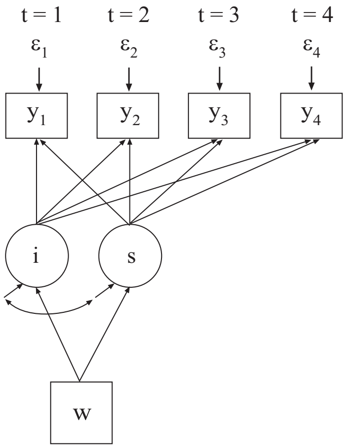
		TD			
		low	up	chronic	
AUD	low	.61	.08	.001	.69
	down	.15	.07	.03	.25
	crhonic	.04	.02	.003	.06
		.80	.17	.03	1.000


Outcome
Escalating
Early Onset
Normative
Age
18
37


y1
y2
y3
y4
i
s
q
x
c
u

Are All Individuals From The Same Population?





$$(1) \quad y_{ti} = i_i + s_i \; time_{ti} + \varepsilon_{ti}$$

$$(2a) \quad i_i = \alpha_0 + \gamma_0 \; w_i + \zeta_{0i}$$

$$(2b) \quad s_i = \alpha_1 + \gamma_1 \; w_i + \zeta_{1i}$$

Modeling motivated by substantive theories of:

- Multiple Disease Processes: Prostate cancer (Pearson et al.)
- Multiple Pathways of Development: Adolescent-limited versus life-course persistent antisocial behavior (Moffitt), crime curves (Nagin), alcohol development (Zucker, Schulenberg)
- Different response to medication such as placebo response to antidepressants (Muthén & Brown, 2009, Statistics in Medicine; Muthén et al., 2011, APPA book)

Example: Mixed-Effects Regression Models For Studying The Natural History Of Prostate Disease

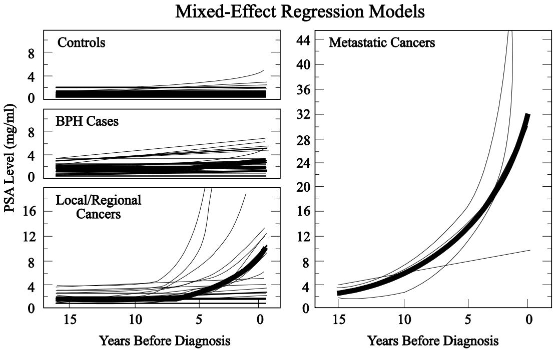
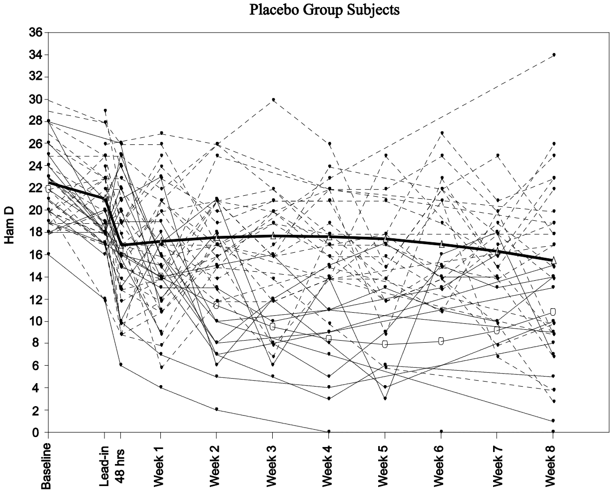


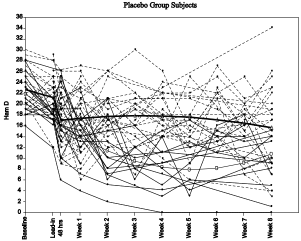
Figure 2. Longitudinal PSA curves estimated from the linear mixed-effects model for the group average (thick solid line) and for each individual in the study (thin solid lines)

Source: Pearson, Morrell, Landis & Carter (1994), Statistics in Medicine

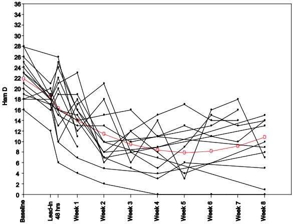
A Clinical Trial Of Antidepressants Growth Mixture Modeling With Placebo Response (Muthén et al, 2011)



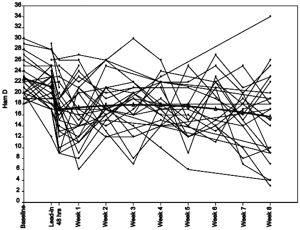
A Clinical Trial Of Antidepressants Growth Mixture Modeling With Placebo Response (Muthén et al, 2011)



Responder Class

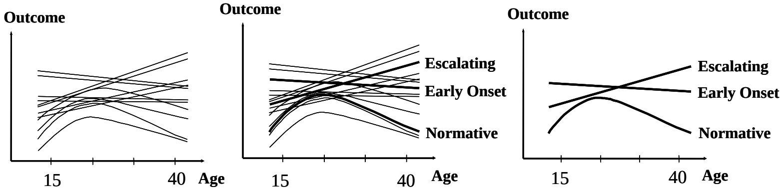


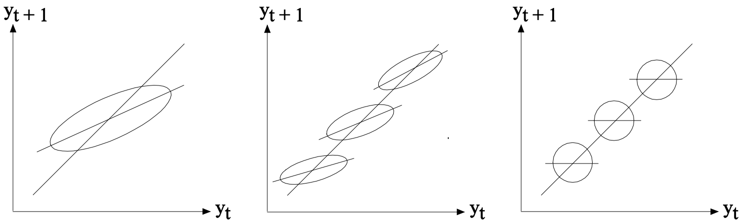
Non-Responder Class



Growth Modeling Paradigms

HLM (Raudenbush) GMM (Muthén) LCGA (Nagin)





- Replace rhetoric with statistics – let likelihood decide
- HLM and LCGA are special cases of GMM

10.1 Philadelphia Crime Data: ZIP Growth Mixture Modeling

- 13,160 males ages 4 - 26 born in 1958 (Moffitt, 1993; Nagin & Land, 1993)
- Annual counts of police contacts
- Individuals with more than 10 counts in any given year deleted (n=13,126)
- Data combined into two-year intervals

Zero-Inflated Poisson (ZIP) Growth Mixture Modeling Of Counts

$$u_{ti} = \begin{cases} 0 & \text{with probability } \pi_{ti} \\ \text{Poisson}(\lambda_{ti}) & \text{with probability } 1 - \pi_{ti} \end{cases} \tag{1}$$

$$\ln \lambda_{ti|C_i=c} = \eta_{0i} + \eta_{1i}\alpha_{ti} + \eta_{2i}\alpha_{ti}^2 \tag{2}$$

$$\eta_{0i|C_i=c} = \alpha_{0c} + \zeta_{0i} \tag{3}$$

$$\eta_{1i|C_i=c} = \alpha_{1c} + \zeta_{1i} \tag{4}$$

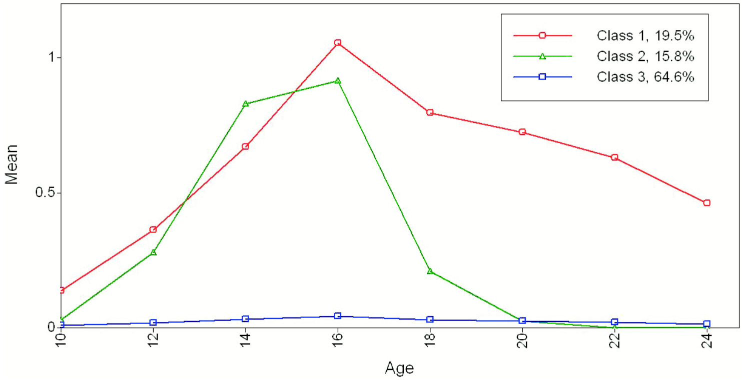
$$\eta_{2i|C_i=c} = \alpha_{2c} + \zeta_{2i} \tag{5}$$

In Mplus, $\pi_{ti} = P(u\#_{ti} = 1)$, where $u\#$ is a binary latent inflation variable and $u\# = 1$ indicates that the individual is unable to assume any value except 0.

ZIP Modeling Of Philadelphia Crime: Log-Likelihood And BIC Comparisons For GMM And LCGA

Model	Log-Likelihood	# Parameters	BIC	# Significant Residuals
1-class GMM	-40,606	17	81,373	5
2-class GMM	-40,422	21	81,044	4
3-class GMM	-40,283	25	80,803	1
4-class GMM	-40,237	29	80,748	0
4-class LCGA	-40,643	23	81,503	4
5-class LCGA	-40,483	27	81,222	3
6-class LCGA	-40,410	31	81,114	3
7-class LCGA	-40,335	35	81,003	2
8-class LCGA	-40,263	39	80,896	1

Three-Class ZIP GMM For Philadelphia Crime

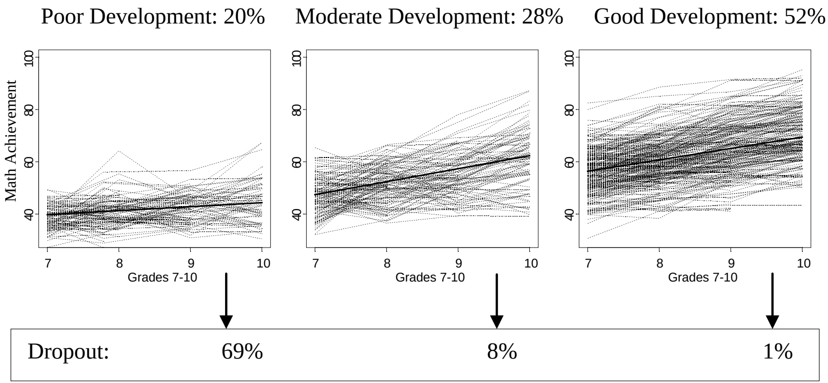


Input Excerpts Three-Class ZIP GMM For Philadelphia Crime

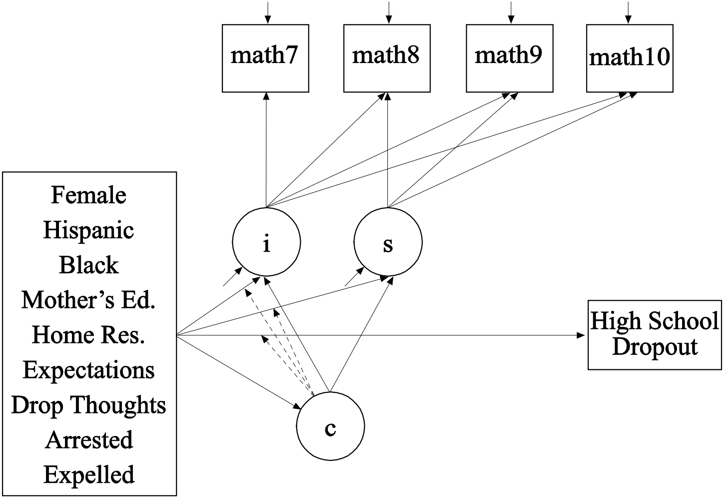
```
VARIABLE:   USEVAR = y10 y12 y14 y16 y18 y20 y22 y24;
            !y10 = ages 10-11, y12 = ages y12-13, etc
            IDVAR = cohortid;
            USEOBS = y10 LE 10 AND y12 LE 10 AND y14 LE 10 AND
            y16 LE 10 AND y18 LE 10 AND y20 LE 10 AND y22 LE 10
            AND y24 LE 10;
            COUNT = y10-y24(i);
            CLASSES = c(3);
ANALYSIS:   TYPE = MIXTURE;
            ALGORITHM = INTEGRATION;
            PROCESS = 8;
            INTEGRATION = 10;
```

Input Excerpts Three-Class ZIP GMM For Philadelphia Crime (Continued)

```
              STARTS = 50 5;
              INTERACTIVE = control.dat;
MODEL:        %OVERALL%
              i s q | y10@0 y12@.1 y14@.2 y16@.3 y18@.4 y20@.5 y22@.6
              y24@.7;
OUTPUT:       TECH1 TECH10;
PLOT:         TYPE = PLOT3;
              SERIES = y10-y24(s);
```


Poor Development: 20%
Moderate Development: 28%
Good Development: 52%
Math Achievement
40
60
80
100
7
8
9
10
Grades 7-10
Dropout:
69%
8%
1%

LSAY Math Achievement Trajectory Classes



```
VARIABLE:   USEVARIABLES = female mothed homeres math7 math8
            math9 math10 expel arrest hisp black hsdrop expect droptht7;
            CATEGORICAL = hsdrop;
            CLASSES = c(3);
ANALYSIS:   TYPE = MIXTURE; STARTS = 0;
MODEL:      %OVERALL%
            i s | math7@0 math8@1 math9@2 math10@3;
            i s ON mothed homeres expect droptht7 expel arrest
            female hisp black;
            c ON mothed homeres expect droptht7 expel arrest
            female hisp black;
            hsdrop ON mothed homeres expect droptht7 expel arrest
            female hisp black;
            %c#1%
            [i*36]; [s*0]; [hsdrop$1*-1]; ! to get the low class first
OUTPUT:     SAMPSTAT STANDARDIZED TECH1 TECH8;
PLOT:       TYPE = PLOT3;
```

Output Excerpts For LSAY Trajectory Classes

	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
Categorical Latent Variables				
c#1 ON				
mothed	-0.251	0.122	-2.055	0.040
homeres	-0.240	0.114	-2.111	0.035
expect	-0.512	0.183	-2.805	0.005
droptht7	1.267	0.659	1.922	0.055
expel	1.903	0.526	3.616	0.000
arrest	0.385	0.485	0.795	0.427
female	-0.635	0.339	-1.870	0.061
hisp	0.977	0.555	1.760	0.078
black	25.391	0.380	66.780	0.000

	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
hsdrop ON				
mothed	-0.402	0.098	-4.123	0.000
homeres	-0.113	0.044	-2.547	0.011
expect	-0.281	0.055	-5.067	0.000
droptht7	0.403	0.240	1.678	0.093
expel	1.232	0.195	6.306	0.000
arrest	0.322	0.255	1.262	0.207
female	0.554	0.150	3.708	0.000
hisp	0.219	0.222	0.986	0.324
black	-0.003	0.203	-0.014	0.989

Output Excerpts For LSAY Trajectory Classes, Continued

	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
Thresholds Class 1				
hsdrop$1	-0.488	0.305	-1.599	0.110
Thresholds Class 2				
hsdrop$1	0.495	0.374	1.325	0.185
Thresholds Class 3				
hsdrop$1	1.637	0.578	2.834	0.005

$P(\text{hsdrop}=1 \mid x=0) = 1/(1+\exp(\text{threshold}))$.

10.3 General Growth Mixture Modeling With Sequential Processes

- New setting:
  - Sequential, linked processes
- New aims:
  - Using an earlier process to predict a later process
  - Early prediction of failing class

Application: General growth mixture modeling of first- and second-grade reading skills and their Kindergarten precursors; prediction of reading failure (Muthén, Khoo, Francis, Boscardin, 1999). Suburban sample, n = 410.

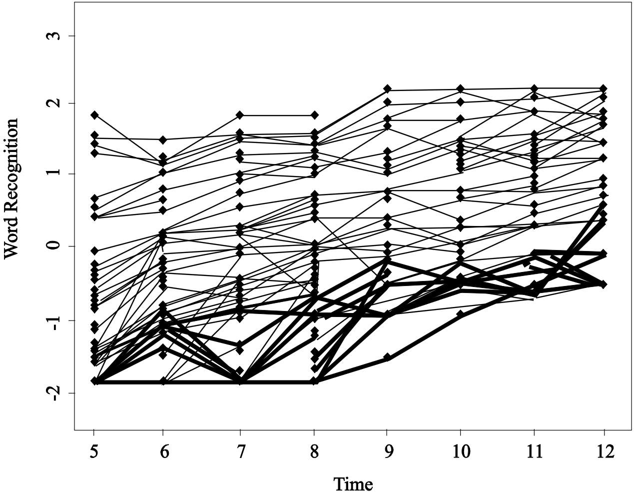
10.3.1 Assessment Of Reading Skills Development: Early Classification

- Longitudinal multiple-cohort design involving approximately 1000 children with measurements taken four times a year from Kindergarten through grade two (October, December, February, April)
- Grade 1 - Grade 2: reading and spelling skills
- Precursor skills: phonemic awareness (Kindergarten, Grade 1, Grade 2), letters/names/sounds (Kindergarten only), rapid naming
- Standardized reading comprehension tests at the end of Grade 1 and Grade 2 (May).

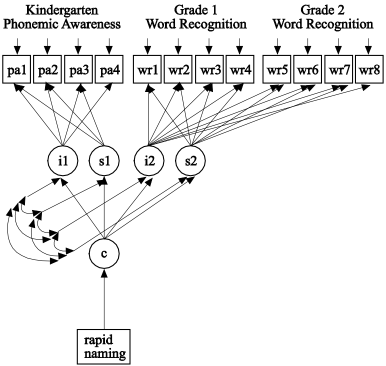
Assessment Of Reading Skills Development (Continued)

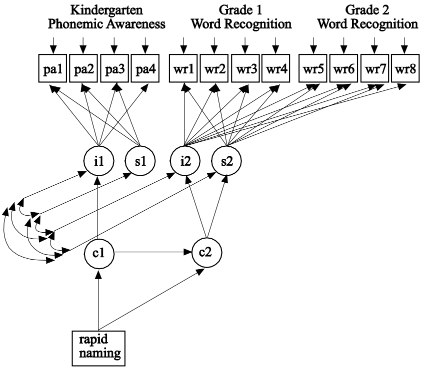
Three research hypotheses (EARS study; Francis, 1996):

- Kindergarten children will differ in their growth and development in precursor skills
- The rate of development of the precursor skills will relate to the rate of development and the level of attainment of reading and spelling skills - and the individual growth rates in reading and spelling skills will predict performance on standardized tests of reading and spelling
- The use of growth rates for skills and precursors will allow for earlier identification of children at risk for poor academic outcomes and lead to more stable predictions regarding future academic performance


Word Recognition
Time
-2
-1
0
1
2
3
5
6
7
8
9
10
11
12

Word Recognition Development In Grades 1 And 2

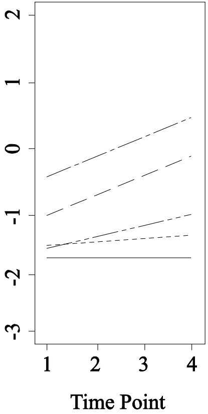




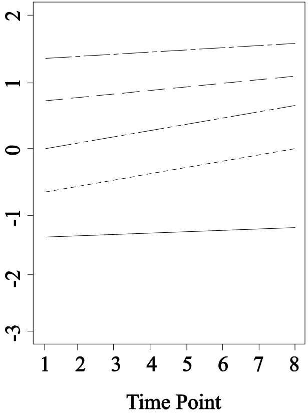
Input For Growth Mixture Model For Reading Skills Development

```
TITLE:      Growth mixture model for reading skills development
DATA:       FILE = newran.dat;
VARIABLE:   NAMES = gender eth wc pa1-pa4 wr1-wr8 l1-l4 s1 r1 s2 r2
            rnaming1 rnaming2 rnaming3 rnaming4;
            USEVAR = pa1-wr8 rnaming4;
            MISSING ARE ALL (999);
            CLASSES = c(5);
ANALYSIS:   TYPE = MIXTURE;
MODEL:      %OVERALL%
            i1 s1 | pa1@-3 pa2@-2 pa3@-1 pa4@0;
            i2 s2 | wr1@-7 wr2@-6 wr3@-5 wr4@-4 wr5@-3 wr6@-2
            wr7@-1 wr8@0;
            c#1-c#4 ON rnaming4;
OUTPUT:     TECH8;
```

Kindergarten Growth
(Five Classes)
Phonemic Awareness



Grades 1 and 2 Growth
(Five Classes)
Word Recognition



How Early Can A Good Classification Be Made?

Focus on Class 1, the failing class.

1. Estimate full growth mixture model for Kindergarten, Grade 1, and Grade 2 outcomes
2. Use the estimated full model to classify students into classes based on the posterior probabilities for each class, where a student is classified into the class with the largest posterior probability.
3. Classify students using early information by holding parameters fixed at the estimates from the full model of Step 1 and classifying individuals using Kindergarten information only, adding Grade 1 outcomes, adding Grade 2 outcomes
4. Study quality of early classification by cross-tabulating individuals classified as in Steps 2 and 3 (sensitivity and specificity)

Sensitivity And Specificity Of Early Classification

		Full Model					
		1.00	2.00	3.00	4.00	5.00	Total
K	1.00	28	7	3			38
Only	2.00	10	29	16			55
	3.00	8	33	100	25		166
	4.00		1	24	63	17	105
	5.00		1	1	12	32	46
Total		46	71	144	100	49	410
K + 1	1.00	28	7	3			38
Only	2.00	15	44	24			83
	3.00	3	20	112	20		155
	4.00			5	79	4	88
	5.00				1	45	46
Total		46	71	144	100	49	410

Sensitivity And Specificity Of Early Classification (Continued)

		Full Model					
		1.00	2.00	3.00	4.00	5.00	Total
K + 2	1.00	28	8				36
Only	2.00	16	54	22			92
	3.00	2	9	119	7		137
	4.00			4	91	4	99
	5.00				2	45	47
Total		46	71	144	100	49	410
K + 3	1.00	37	12				49
Only	2.00	9	53	8			70
	3.00		6	136	4		146
	4.00			1	95	1	97
	5.00				1	48	49
Total		46	71	144	100	49	410

		Full Model					
		1.00	2.00	3.00	4.00	5.00	Total
K + 4	1.00	45	11				56
Only	2.00	1	57	3			61
	3.00		3	141	2		146
	4.00			1	97		98
	5.00				1	49	50
Total		46	71	144	100	49	410
K + 5	1.00	45	3				48
Only	2.00	1	66				67
	3.00		2	145	1		148
	4.00				98		98
	5.00				1	49	50
Total		46	71	144	100	49	410

Sensitivity And Specificity Of Early Classification (Continued)

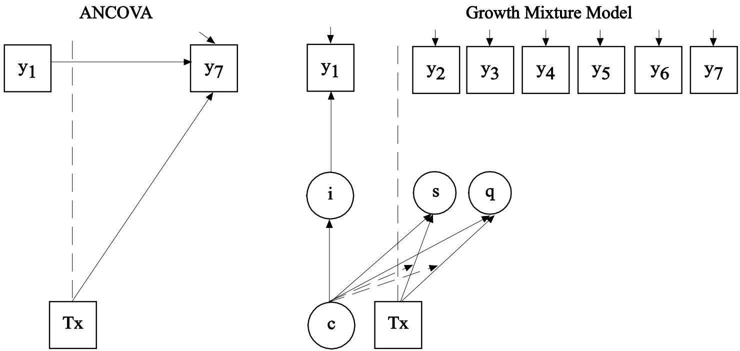
		Full Model					
		1.00	2.00	3.00	4.00	5.00	Total
K + 6	1.00	46					46
Only	2.00		69				69
	3.00		1	145	1		147
	4.00				98		98
	5.00				1	49	50
Total		46	70	144	100	49	410

Different treatment effects in different trajectory classes

Muthén, B., Brown, C.H., Masyn, K., Jo, B., Khoo, S.T., Yang, C.C., Wang, C.P. Kellam, S., Carlin, J., & Liao, J. (2002). General growth mixture modeling for randomized preventive interventions. Biostatistics, 3, 459-475.

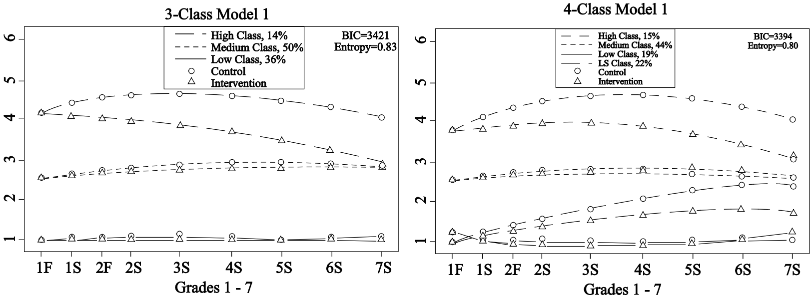
Muthén & Brown (2009), Statistics in Medicine: A sizable portion of responders in antidepressant trials may be placebo responders

See also Muthén & Curran, 1997 for monotonic treatment effects



GMM: treatment changes trajectory shape.

Modeling Treatment Effects

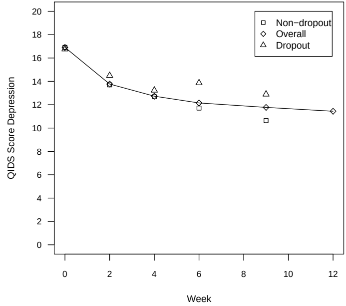


11. Data Not Missing At Random: Non-Ignorable Dropout In Longitudinal Studies

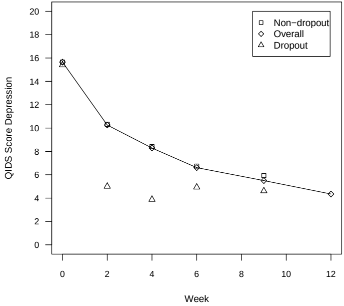
- Missing completely at random (MCAR)
- Missing at random (MAR)
- Not missing at random (NMAR)
    - Selection modeling
    - Pattern-mixture modeling
    - General latent variable modeling

11.1 Longitudinal Data From An Antidepressant Trial (STAR*D) $n = 4041$

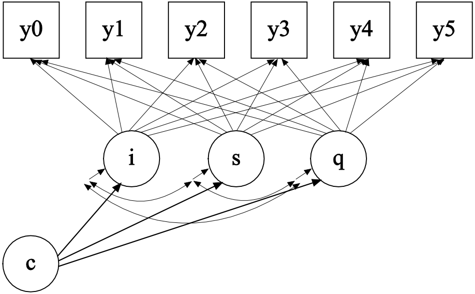
Subjects treated with citalopram (Level 1). No placebo group
Sample means of the QIDS depression score at each visit:



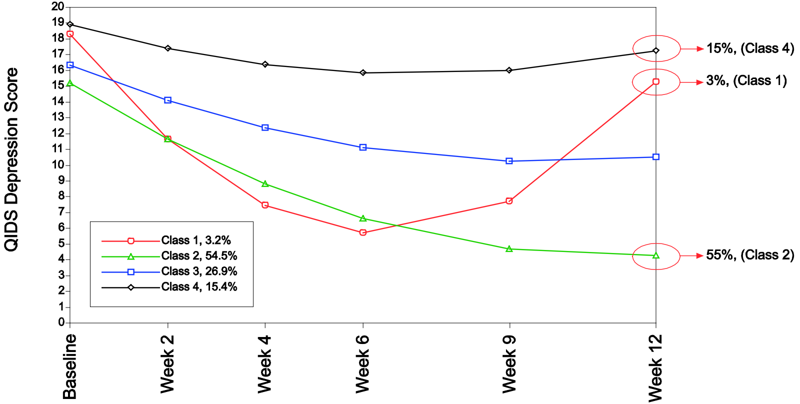
Next level



Follow-up


y0
y1
y2
y3
y4
y5
i
s
q
c

4-Class Growth Mixture Model



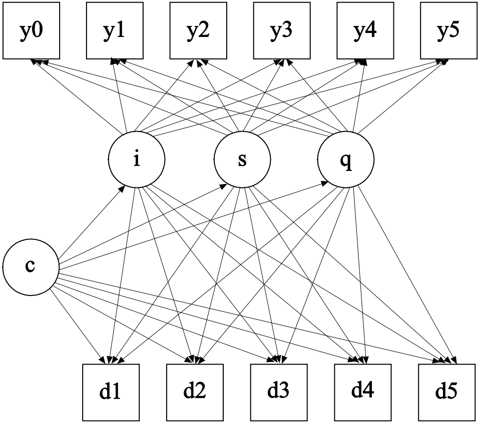
Not Missing At Random (NMAR): Non-Ignorable Dropout Modeling

NMAR: Missingness influenced by latent variables

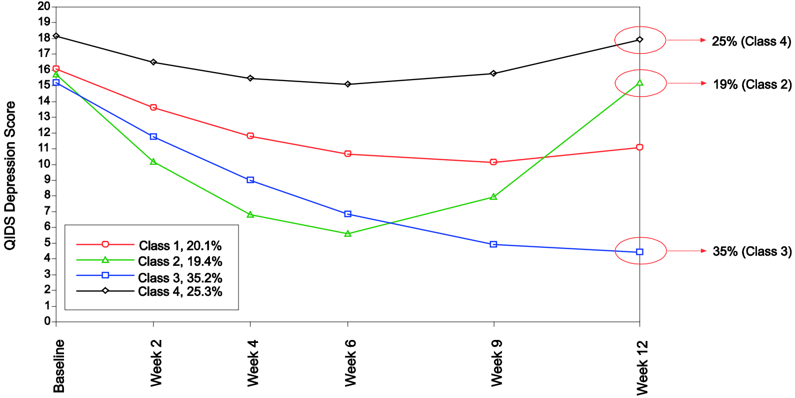
- Data to be modeled are not only outcomes but also missing data indicators
- Two general approaches:
  - Selection modeling: Growth features influence dropout occasion
  - Pattern-mixture modeling: Dropout occasion influences growth parameters

Muthén, Asparouhov, Hunter & Leuchter (2011). Growth modeling with non-ignorable dropout: Alternative analyses of the STAR*D antidepressant trial. **Psychological Methods**.

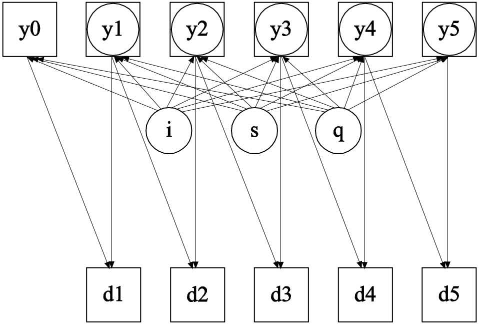
Beunckens Mixture Model (Mixture Wu-Carroll Model): Adding Dropout Information (Survival Indicators)



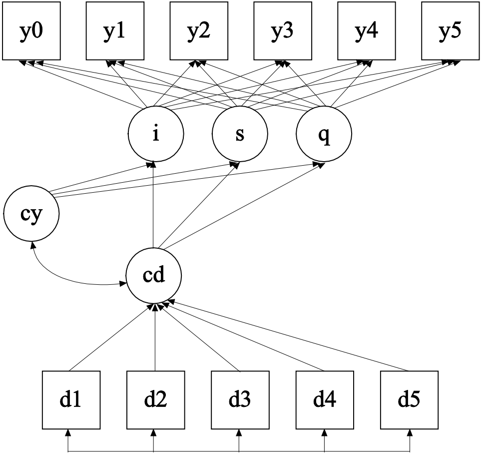
4-class Beunckens Selection Mixture Model



Diggle-Kenward NMAR Selection Model



Muthén-Roy Pattern-Mixture Model (d's are dropout dummies)



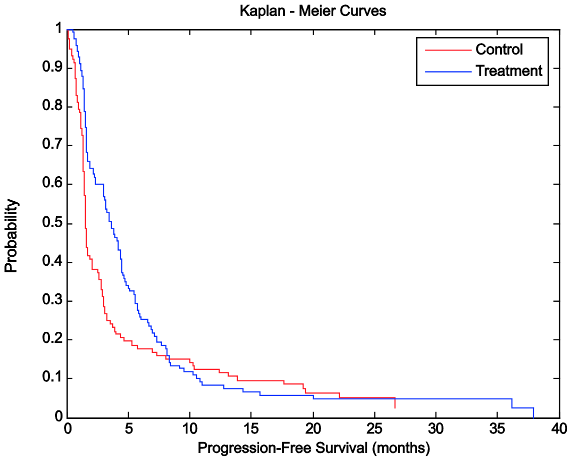
Comparing Trajectory Class Percentages Across Models

The NMAR approach of adding dropout information gives a less favorable conclusion regarding drug response than the standard assumption of MAR.

Model	Response class	Temporary response class	Non-response class
MAR	55 %	3 %	15 %
NMAR models:			
Beuncken	35 %	19 %	25%
Muthén-Roy	32 %	15 %	14 %

- Muthén & Masyn (2005). Discrete-time survival mixture analysis. **Journal of Educational and Behavioral Statistics**
- Larsen (2004). Joint analysis of time-to-event and multiple binary indicators of latent classes. **Biometrics**
- Larsen (2005). The Cox proportional hazards model with a continuous latent variable measured by multiple binary indicators. **Biometrics**
- Asparouhov, Masyn, & Muthén (2006). Continuous time survival in latent variable models. ASA section on Biometrics, 180-187
- Muthén, Asparouhov, Boye, Hackshaw & Naegeli (2009). Applications of continuous-time survival in latent variable models for the analysis of oncology randomized clinical trial data using Mplus. Technical Report

12.1 Cancer Survival Trial of Second-Line Treatment of Mesothelioma

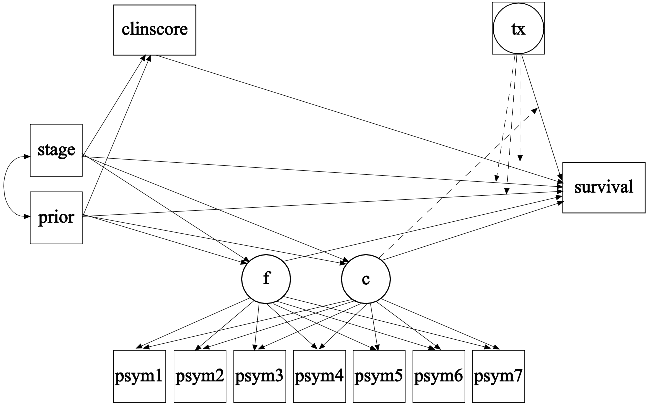


Patient-Reported Lung Cancer Symptom Scale (LCSS)

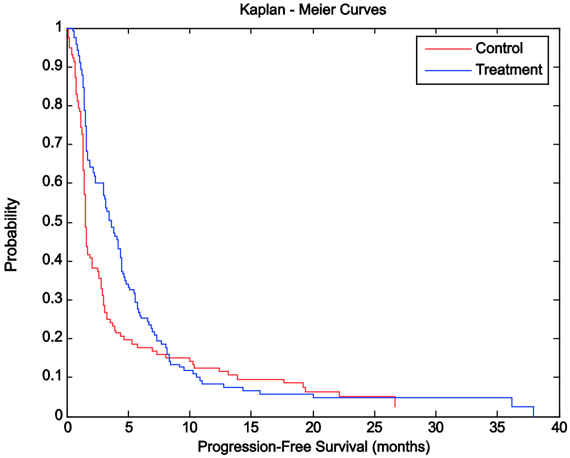
Directions: Please place a mark along each line where it would best describe the symptoms of your lung illness DURING THE PAST DAY (during the past 24 hours)

1. How is your appetite?	
As good as it could be	As bad as it could be
2. How much fatigue do you have?	
None	As much as it could be
3. How much coughing do you have?	
None	As much as it could be
4. How much shortness of breath do you have?	
None	As much as it could be
5. How much blood do you see in your sputum?	
None	As much as it could be
6. How much pain do you have?	
None	As much as it could be
7. How bad are your symptoms from your lung illness?	
I have none	As much as it could be
8. How much has your illness affected your ability to carry out normal activities?	
Not at all	So much that I can do nothing for myself
9. How would you rate the quality of your life today?	
Very high	Very low

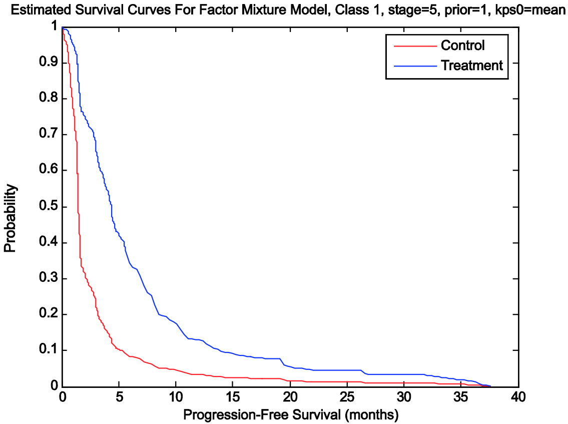
Predicting Survival From Visit 0 Using a Factor Mixture Model For LCSS Items



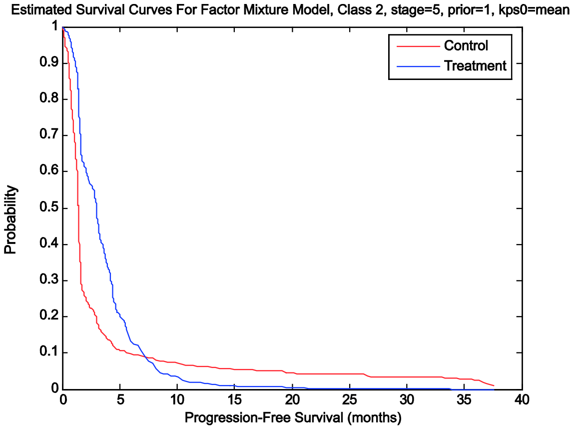
Survival Curves Showing Overall Treatment Effect



Survival Curves For Low-Symptom Class



Survival Curves For High-Symptom Class



References

For references, see handouts for Topics 1 - 9 at
`http://www.statmodel.com/course_materials.shtml`

For handouts and videos of Version 7 training, see
`http://mplus.fss.uu.nl/2012/09/12/the-workshop-new-features-of-mplus-v7/`

For papers using special Mplus features, see
`http://www.statmodel.com/papers.shtml`