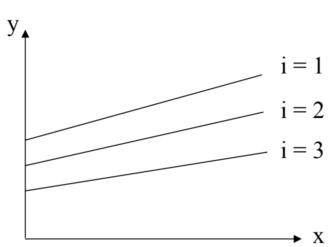
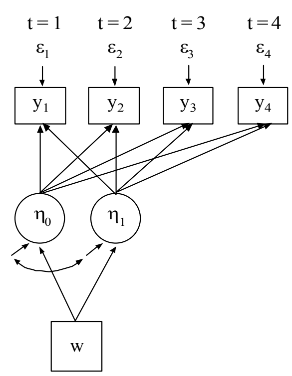
Growth Mixture Modeling

1

Individual Development Over Time





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

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

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

2

Mixtures and Latent Trajectory Classes

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)
- Subtypes: Subtypes of alcoholism (Cloninger, Zucker)

3

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

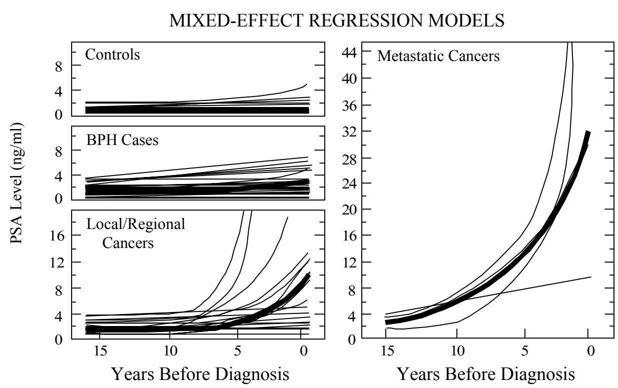
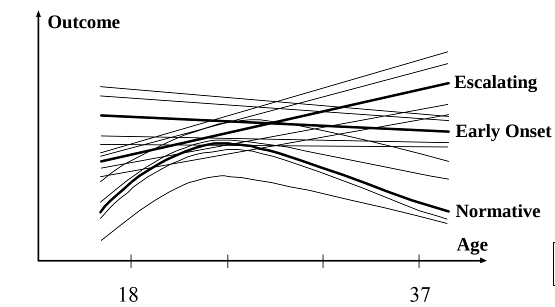


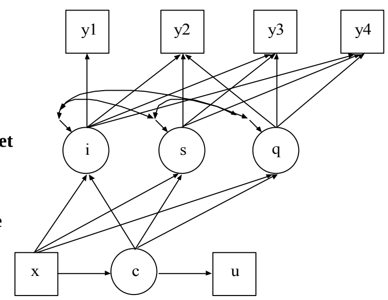
Figure 2. Longitudinal PSA curves estimated from the linear mixed-effects model for the group average (thick solid line) and for each individual in the study (thin solid lines)

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

4

Growth Mixture Modeling Of Developmental Pathways

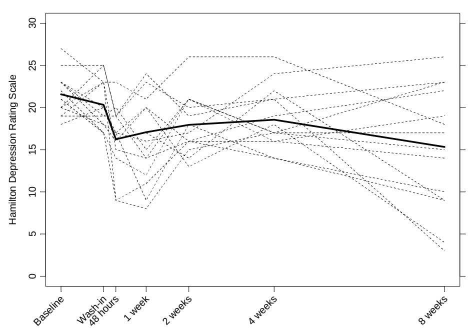




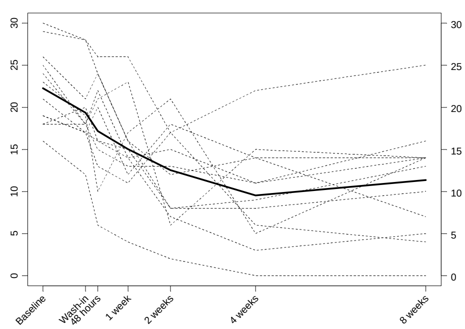
5

A Clinical Trial Of Depression Medication: Two-Class Growth Mixture Modeling

Placebo Non-Responders, 55%



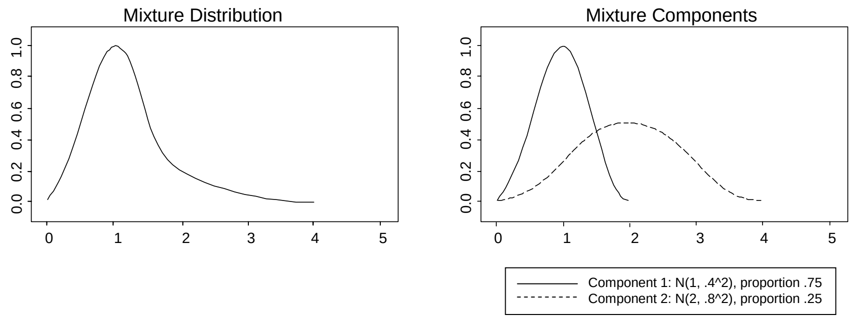
Placebo Responders, 45%



6

Mixture Distributions

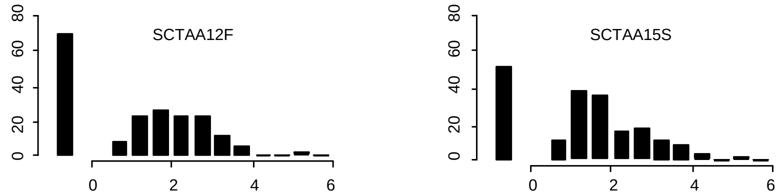
Non-normality for mixture, normality for mixture components



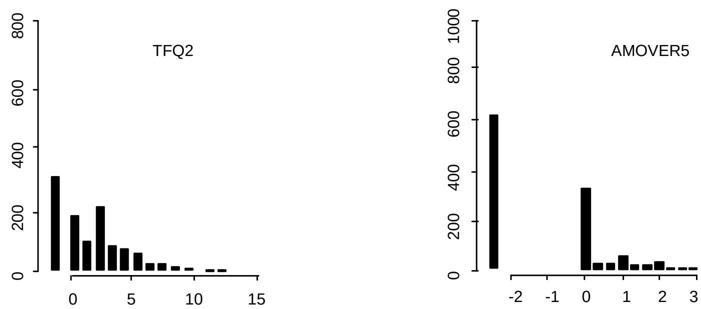
7

Two Types Of Distributions

(1) Normal mixture components



(2) Preponderance of zeroes



8

Growth Mixture Analysis

Generalization of conventional random effect growth modeling (multilevel modeling) to include qualitatively different developments (Muthén & Shedden, 1999 in Biometrics). Combination of conventional growth modeling and cluster analysis (finite mixture analysis).

- Setting
  - Longitudinal data
  - A single or multiple items measured repeatedly
  - Hypothesized trajectory classes (categorical latent variable)
  - Individual trajectory variation within classes

Growth Mixture Analysis (Continued)

- Aim
  - Estimate trajectory shapes
  - Estimate trajectory class probabilities
  - Estimate variation within class
  - Relate class probabilities to covariates
  - Relate within-class variation to covariates
  - Classify individuals into classes (posterior prob's)

Application: Mathematics achievement, grades 7 – 10 (LSAY) related to mother's education and home resources.
National sample, n = 846.

Strategies For Finding The Number Of Classes In Growth Mixture Modeling

- **Comparing models with different numbers of classes**
  - BIC – low BIC value corresponds to a high loglikelihood value and a parsimonious model
  - TECH11 – Lo-Mendell-Rubin likelihood ratio test (Biometrika, 2001)
  - TECH14 – bootstrapped LRT (Version 4)
- **Residuals and model tests**
  - TECH7 – class-specific comparisons of model-estimated means, variances, and covariances versus posterior probability-weighted sample statistics
  - TECH12 – class-mixed residuals for univariate skewness and kurtosis
  - TECH13 – multivariate skew and kurtosis model tests

11

Strategies For Finding The Number Of Classes In Growth Mixture Modeling (Continued)

- **Classification quality**
  - Posterior probabilities – classification table and entropy
  - Individual trajectory classification using pseudo classes (Bandeen-Roche et al., 1997; Muthén et al. in Biostatistics, 2002)
- **Interpretability and usefulness of the latent classes**
  - Trajectory shapes
  - Number of individuals in each class
  - Number of estimated parameters
  - Substantive theory
  - Auxiliary (external) variables – predictive validity

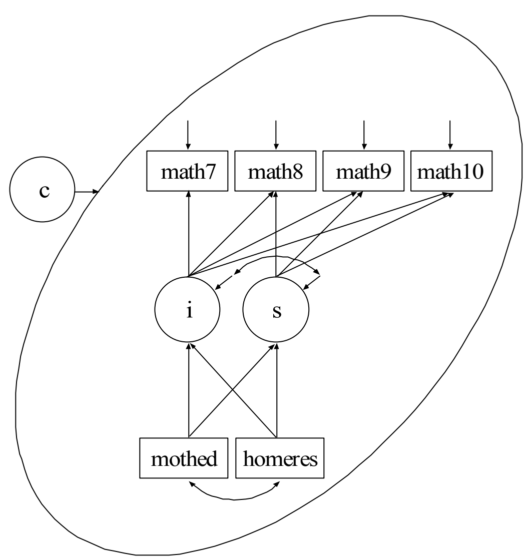
12

Strategies For Finding Starting Values In Growth Mixture Modeling (Continued)

- **Strategy 1**
  - Do a conventional growth analysis (one-class model)
  - Use estimated growth factor means and standard deviations as growth factor mean starting values in a multi-class model – mean plus and minus .5 standard deviation
- **Strategy 2**
  - Estimate a multi-class model with the variances and covariances of the growth factors fixed to zero (LCGA)
  - Use the estimated growth factor means as growth factor mean starting values for a model with growth factor variances and covariances free

Random starts makes it unnecessary to give starting values: starts = 50 5

13



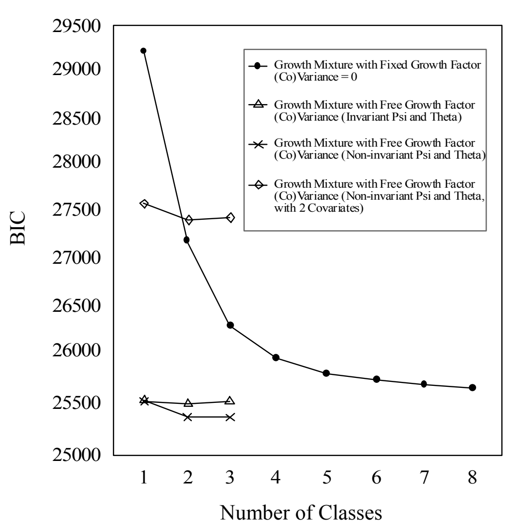
14

Deciding On The Number Of Classes For The LSAY Growth Mixture Model

n = 935

Number of Classes	1	2	3
Loglikelihood	-11,997.653	-11,864.826	-11,856.220
# parameters	15	29	36
BIC	24,098	23,928	23,959
AIC	24,025	23,788	23,784
Entropy	NA	.468	.474
TECH11 LRT p-value for k-1 classes	NA	.0000	.4041
Multivariate skew p-value	.00	.34	.26
Multivariate kurtosis p-value	.00	.10	.05

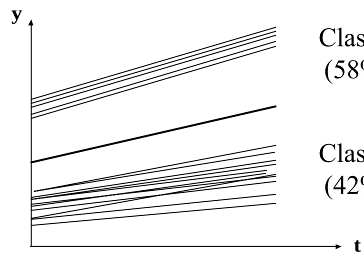
Model Fit by BIC: LSAY



LSAY: Estimated Two-Class Growth Mixture Model

Class 1
Intercept ON Mothed* Homeres
Slope ON Mothed Homeres*

Class 2
Intercept ON Mothed* Homeres*
Slope ON Mothed Homeres



Conventional Single-Class Analysis

Intercept ON Mothed* Homeres*
Slope ON Mothed Homeres*

17

Input For LSAY Two-Class Growth Mixture Model

```
TITLE:      2-class varying slopes on mothed and homeres
            varying Psi varying Theta

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

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

ANALYSIS:   TYPE = MIXTURE;
```

18

Input For LSAY Two-Class Growth Mixture Model (Continued)

```
MODEL:      %OVERALL%
            intercpt slope | math7@0 math8@1 math9 math10;
            intercpt slope ON mothed homeres;
            %c#2%
            intercpt slope ON mothed homeres;
            math7-math10 intercpt slope;
            slope WITH intercpt;

OUTPUT:     TECH8 TECH12 TECH13 RESIDUAL;
```

19

Output Excerpts LSAY Two-Class Growth Mixture Model

Tests Of Model Fit

```
Loglikelihood
             H0 Value                         -11864.825
Information Criteria
             Number of Free Parameters                29
             Akaike (AIC)                      23787.649
             Bayesian (BIC)                    23928.025
             Sample-Size Adjusted BIC          23835.923
                (n* = (n + 2) / 24)
             Entropy                               0.468
```

20

Output Excerpts LSAY Two-Class Growth Mixture Model (Continued)

Classification Information

FINAL CLASS COUNTS AND PROPORTIONS OF TOTAL SAMPLE SIZE

Class 1	392.19327	0.41946
Class 2	542.80673	0.58054

CLASSIFICATION OF INDIVIDUALS BASED ON THEIR MOST LIKELY CLASS MEMBERSHIP

Class Counts and Proportions

Class 1	342	0.36578
Class 2	593	0.63422

Average Class Probabilities by Class

	1	2
Class 1	0.853	0.147
Class 2	0.170	0.830

21

Output Excerpts LSAY Two-Class Growth Mixture Model (Continued)

TECHNICAL 12 OUTPUT

ESTIMATED MIXED MODEL AND RESIDUALS (OBSERVED – EXPECTED)

Observed Skewness

MATH7	MATH8	MATH9	MATH10	MOTHED	HOMERES
-0.184	-0.312	-0.349	-0.471	1.214	-0.087

Estimated Mixed Skewness

MATH7	MATH8	MATH9	MATH10	MOTHED	HOMERES
-0.160	-0.190	-0.348	-0.517	0.002	-0.012

22

Output Excerpts LSAY Two-Class Growth Mixture Model (Continued)

```
TECHNICAL 13 OUTPUT

   SKEW AND KURTOSIS TESTS OF MODEL FIT

   TWO-SIDED MULTIVARIATE SKEW TEST OF FIT
         Sample Value                          1.245
         Mean                                  0.999
         Standard Deviation                    0.275
         P-Value                              0.3400
   TWO-SIDED MULTIVARIATE KURTOSIS TEST OF FIT
         Sample Value                         29.612
         Mean                                 27.842
         Standard Deviation                    1.015
         P-Value                              0.1000
```

23

Output Excerpts LSAY Two-Class Growth Mixture Model (Continued)

Model Results

```
                         Estimates         S.E.     Est./S.E.
Class 1

 INTERCPT |
    MATH7                    1.000         .000          .000
    MATH8                    1.000         .000          .000
    MATH9                    1.000         .000          .000
    MATH10                   1.000         .000          .000

 SLOPE    |
    MATH7                     .000         .000          .000
    MATH8                    1.000         .000          .000
    MATH9                    2.422         .133        18.157
    MATH10                   3.580         .204        17.570

 INTERCPT ON
    MOTHED                   1.656         .626         2.645
    HOMERES                   .720         .377         1.911

 SLOPE    ON
    MOTHED                    .146         .154          .953
    HOMERES                   .228         .087         2.626
```

24

Output Excerpts LSAY Two-Class Growth Mixture Model (Continued)

	Estimates	S.E.	Est./S.E.
SLOPE WITH			
INTERCPT	.727	1.643	.443
Residual Variances			
MATH7	17.198	2.840	6.055
MATH8	15.257	2.077	7.347
MATH9	24.170	3.294	7.337
MATH10	49.112	10.037	4.893
INTERCPT	54.297	6.094	8.910
SLOPE	1.643	.627	2.620
Intercepts			
MATH7	.000	.000	.000
MATH8	.000	.000	.000
MATH9	.000	.000	.000
MATH10	.000	.000	.000
INTERCPT	42.733	1.648	25.922
SLOPE	.816	.366	2.228

25

Output Excerpts LSAY Two-Class Growth Mixture Model (Continued)

	Estimates	S.E.	Est./S.E.
Class 2			
INTERCPT |			
MATH7	1.000	.000	.000
MATH8	1.000	.000	.000
MATH9	1.000	.000	.000
MATH10	1.000	.000	.000
SLOPE |			
MATH7	.000	.000	.000
MATH8	1.000	.000	.000
MATH9	2.422	.133	18.157
MATH10	3.580	.204	17.570
INTERCPT ON			
MOTHED	2.085	.376	5.545
HOMERES	1.805	.285	6.334
SLOPE ON			
MOTHED	.028	.079	.358
HOMERES	.054	.058	.943

26

Output Excerpts LSAY Two-Class Growth Mixture Model (Continued)

	Estimates	S.E.	Est./S.E.
SLOPE WITH			
INTERCPT	.478	.575	.831
Residual Variances			
MATH7	12.248	1.587	7.720
MATH8	11.375	1.244	9.141
MATH9	8.014	1.420	5.642
MATH10	7.349	1.312	5.600
INTERCPT	32.431	4.651	6.973
SLOPE	.191	.264	.724
Intercepts			
MATH7	.000	.000	.000
MATH8	.000	.000	.000
MATH9	.000	.000	.000
MATH10	.000	.000	.000
INTERCPT	45.365	1.461	31.054
SLOPE	2.847	.316	9.015

LATENT CLASS REGRESSION MODEL PART

Means	Estimates	S.E.	Est./S.E.
C#1	-.325	.306	-1.063

27

Output Excerpts LSAY Two-Class Growth Mixture Model (Continued)

Residuals

ESTIMATED MODEL AND RESIDUALS (OBSERVED – ESTIMATED) FOR CLASS 1

Model Estimated Means

MATH7	MATH8	MATH9	MATH10	MOTHED	HOMERES
48.650	50.486	53.095	55.221	2.255	3.030

Residuals for Means

MATH7	MATH8	MATH9	MATH10	MOTHED	HOMERES
-0.113	0.072	0.236	-0.389	0.000	0.000

Model Estimated Covariances

	MATH7	MATH8	MATH9	MATH10	MOTHED
MATH7	76.058				
MATH8	60.370	78.945			
MATH9	62.515	68.408	100.958		
MATH10	64.263	72.253	83.615	141.979	
MOTHED	1.751	1.962	2.263	2.507	0.907
HOMERES	2.309	2.908	3.760	4.454	0.345

28

Output Excerpts LSAY Two-Class Growth Mixture Model (Continued)

	HOMERES
HOMERES	2.412

Residuals for Covariances

	MATH7	MATH8	MATH9	MATH10	MOTHED
MATH7	-0.153				
MATH8	-0.109	-0.143			
MATH9	0.413	0.572	0.451		
MATH10	-0.701	-0.614	-0.254	-0.338	
MOTHED	0.210	-0.252	-0.021	0.256	0.000
HOMERES	0.289	-0.254	-0.358	0.720	0.000

	HOMERES
HOMERES	0.000

29

Further Readings On Growth Mixture Modeling

Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In Collins, L.M. & Sayer, A. (Eds.), New methods for the analysis of change (pp. 291-322). Washington, D.C.: APA. (#82)

Muthén, B. (2001). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (eds.), New developments and techniques in structural equation modeling (pp. 1-33). Lawrence Erlbaum Associates. (#86)

Muthén, B. (2002). Beyond SEM: General latent variable modeling. Behaviormetrika, 29, 81-117. (#96)

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (ed.), Handbook of quantitative methodology for the social sciences (pp. 345-368). Newbury Park, CA: Sage Publications. (#100)

30

Further Readings On Growth Mixture Modeling (Continued)

Muthén, B. & Muthén, L. (2000). Integrating person-centered and variable-centered analysis: growth mixture modeling with latent trajectory classes. Alcoholism: Clinical and Experimental Research, 24, 882-891. (#85)

Muthén, B. & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. Biometrics, 55, 463-469. (#78)

General Growth Mixture Modeling

General Growth Mixture Modeling (GGMM)

GGMM goes beyond conventional random effect growth modeling by using latent trajectory classes which

- Allow for heterogeneity with respect to
  - Growth functions – different classes correspond to different growth shapes
  - Antecedents – different background variables have different importance for different classes
  - Consequences – class membership predicts later outcomes
- Allow for prediction of trajectory class membership
- Allow for confirmatory clustering
  - With respect to parameters – describing curve shapes
  - With respect to typical individuals – known classes
- Allow for classification of individuals
  - Early prediction of problematic development
- Allow for enhanced preventive intervention analysis
  - Different classes benefit differently and can receive different treatments

General Growth Mixture Modeling

Muthén (2001), Muthén & Muthén (2000), Muthén (2004)

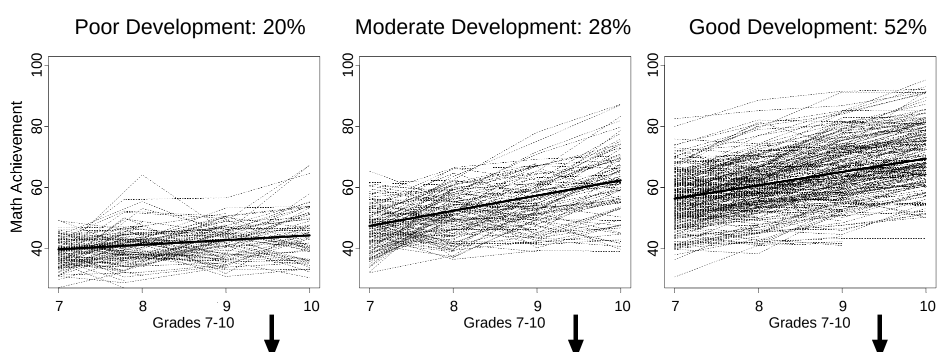
- Setting
  - Longitudinal data
  - A single or multiple items measured repeatedly
  - Hypothesized trajectory classes (categorical latent variable)
  - Individual trajectory variation within classes
  - Prediction of distal outcomes
- Aim
  - Estimate trajectory shapes
  - Estimate trajectory class probabilities
  - Estimate variation within class
  - Relate class probabilities to covariates
  - Relate within-class variation to covariates
  - Relate class membership to distal outcomes
  - Classify individuals into classes (posterior prob's)

General Growth Mixture Modeling (Continued)

- Applications
  - LSAY math achievement development and high school dropout
  - The development of heavy drinking ages 18-30 (NLSY) related to antecedents and consequences. National sample, n = 922

35

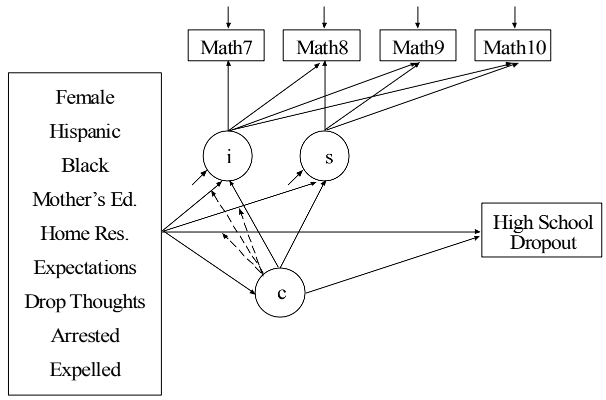
Mplus Graphics For LSAY Math Achievement Trajectory Classes



Dropout: 69% 8% 1%

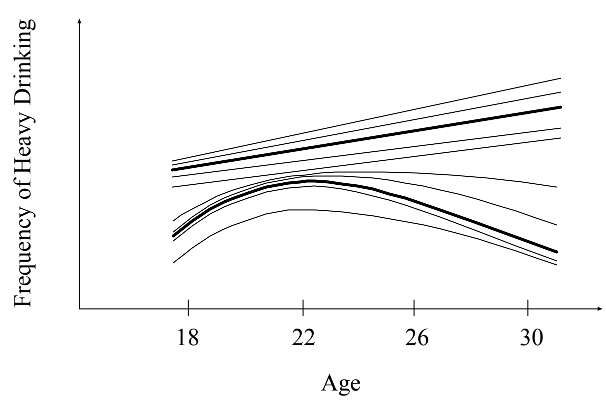
36

LSAY Math Achievement Trajectory Classes



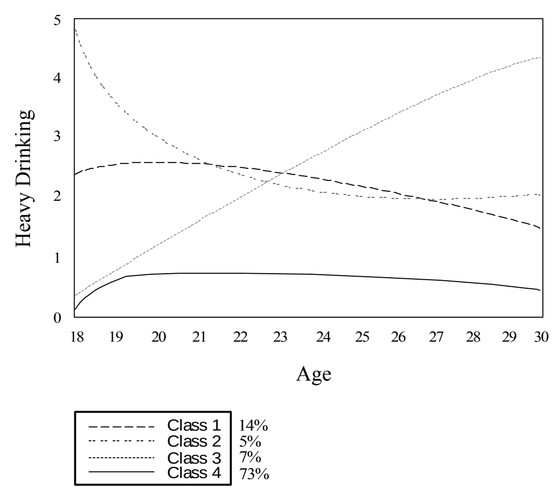
37

Example: NLSY Heavy Drinking Two Latent Trajectory Classes



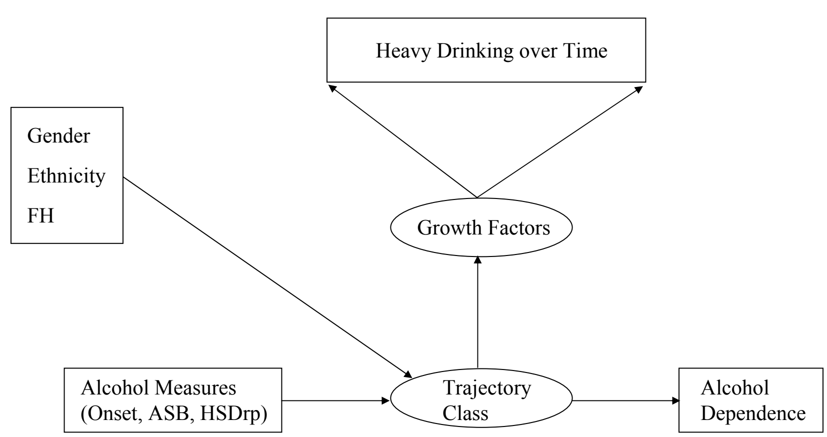
38

NLSY: Heavy Drinking, Cohort 64



39

NLSY: Antecedents And Consequences



40

Multinomial Logistic Regression Of c ON x

The multinomial logistic regression model expresses the probability that individual i falls in class k of the latent class variable c as a function of the covariate x,

$$P(c_j = k \mid x_i) = \frac{e^{\alpha_k + \gamma_k x_i}}{\sum_{s=1}^{K} e^{\alpha_s + \gamma_s x_i}} \,, \tag{90}$$

where $\alpha_K = 0$, $\gamma_K = 0$ so that $e^{\alpha_K + \gamma_K x_i} = 1$.

This implies that the log odds comparing class k to the last class K is

$$log[P(c_i = k \mid x_i)/P(c_i = K \mid x_i)] = \alpha_k + \gamma_k x_i \,. \tag{91}$$

41

Heavy Drinking And Alcohol Dependence NLSY Cohort64 (n=922)

HD Class on Covariates						
	HD Classes					
	1 (Down)		2 (High18)		3(Up)	
	Est.	t	Est.	t	Est.	t
Male	**1.21**	5.52	**1.25**	3.48	**1.45**	4.73
Black	**-0.89**	-3.43	**-3.14**	-2.86	-0.06	-0.17
Hisp	**-0.65**	-2.22	-0.35	-0.86	-0.01	-0.03
ES	**1.24**	4.79	**2.05**	5.72	0.71	1.78
FH1	0.03	0.09	-0.21	-0.41	-0.08	-0.16
FH23	0.04	0.15	0.25	0.56	0.08	0.23
FH123	-0.23	-0.58	**1.18**	2.59	**1.00**	2.60
HSDRP	0.57	1.98	0.32	0.76	**0.91**	2.93
Coll	-0.07	-0.31	**-1.31**	-2.85	**-1.08**	-2.59

42

Heavy Drinking And Alcohol Dependence NLSY Cohort64 (n=922) (Continued)

Alcohol Dependence as a Function of Heavy Drinking Class

	Probability	Odds Ratio
HD Class 1 (Down)	0.16	3.92
HD Class 2 (High 18)	0.26	7.06
HD Class 3 (Up)	0.60	30.00
HD Class 4 (Norm)	0.05	1.00

43

Input Excerpts NLSY Growth Mixture Model With Covariates And A Distal Outcome

```
TITLE:       NLSY for cohort 64 quadratic growth mixture model with
             covariates centered at 25: four-class model of heavy
             drinking with classes predicting dep94

VARIABLE:    CLASSES = c(4);
             CATEGORICAL IS dep94;

ANALYSIS:    TYPE = MIXTURE;

MODEL:       %OVERALL%
             c#1-c#3 ON male black hisp es fh1 fh23 fh123 hsdrp coll;
             i s1 s2 | hd82@-3.008 hd83@-2.197 hd84@-1.621 hd88@-.235
                       hd89@.000 hd94@.884;
```

44

Input Excerpts NLSY Growth Mixture Model With Covariates And A Distal Outcome (Continued)

```
!       log age scale: x_t = a*(ln(t-b) – ln(c-b));
!       where t is time, a and b are constants to fit the mean curve
!       (chosen as a = 2 and b = 16), and c is the centering age,
!       here set at 25.
        %c#1%                              ! Not needed
        [dep94$1*1 i*2 s1*-.5 s2*-.1];     ! Not needed
        %c#2%                              ! Not needed
        [dep94$1*0 i*1 s1*-.2 s2*-.3];     ! Not needed
        %c#3%                              ! Not needed
        [dep94$1*.6 i*3 s1*1.5 s2*.2];     ! Not needed
        %c#4%                              ! Not needed
        [dep94$1*2 i*.6 s1*-.2 s2*-.1];    ! Not needed
```


Output Excerpts NLSY Growth Mixture Model With Covariates And A Distal Outcome

```
C#1         ON
   MALE                    1.214         .220        5.515
   BLACK                   -.886         .258       -3.434
   HISP                    -.645         .290       -2.223
   ES                      1.240         .259        4.789
   FH1                      .026         .291         .088
   FH23                     .039         .261         .149
   FH123                   -.233         .399        -.583
   HSDRP                    .566         .286        1.976
   COLL                    -.071         .231        -.308
```

Output Excerpts NLSY Growth Mixture Model With Covariates And A Distal Outcome (Continued)

C#2 ON			
MALE	1.248	.359	3.481
BLACK	-3.138	1.097	-2.860
HISP	-.346	.401	-.864
ES	2.045	.357	5.722
FH1	-.211	.514	-.410
FH23	.247	.444	.556
FH123	1.178	.456	2.585
HSDRP	.323	.428	.756
COLL	-1.311	.460	-2.851
C#3 ON			
MALE	1.454	.308	4.727
BLACK	-.059	.344	-.171
HISP	-.011	.369	-.030
ES	.712	.399	1.784
FH1	-.079	.502	-.157
FH23	.084	.364	.232
FH123	1.004	.387	2.596
HSDRP	.913	.312	2.926
COLL	-1.075	.414	-2.594

47

Output Excerpts NLSY Growth Mixture Model With Covariates And A Distal Outcome (Continued)

Class 1			
Thresholds			
DEP94$1	1.631	0.248	6.574
Class 2			
Thresholds			
DEP94$1	1.041	0.338	3.077
Class 3			
Thresholds			
DEP94$1	-0.406	0.272	-1.493
Class 4			
Thresholds			
DEP94$1	2.987	0.208	14.392

48

Output Excerpts NLSY Growth Mixture Model With Covariates And A Distal Outcome (Continued)

Classification Information

```
FINAL CLASS COUNTS AND PROPORTIONS OF TOTAL SAMPLE SIZE
  Class 1         135.95653         0.14746
  Class 2          45.86689         0.04975
  Class 3          74.68767         0.08101
  Class 4         665.48891         0.72179
CLASSIFICATION OF INDIVIDUALS BASED ON THEIR MOST LIKELY CLASS
MEMBERSHIP
Class Counts and Proportions
  Class 1               134         0.14534
  Class 2                46         0.04989
  Class 3                72         0.07809
  Class 4               670         0.72668
Average Class Probabilities by Class
                     1         2         3          4
  Class 1         0.994     0.000     0.000      0.005
  Class 2         0.003     0.997     0.000      0.000
  Class 3         0.007     0.000     0.947      0.047
  Class 4         0.003     0.000     0.010      0.987
```


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.

Assessment of Reading Skills Development

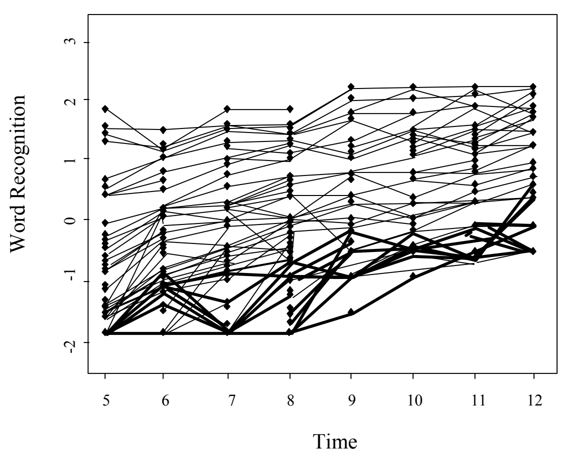
- Longitudinal multiple-cohort design involving approximately 1000 children with measurements taken four times a year from Kindergarten through grade two (October, December, February, April)
- Grade 1 – Grade 2: reading and spelling skills
- Precursor skills: phonemic awareness (Kindergarten, Grade 1,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

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

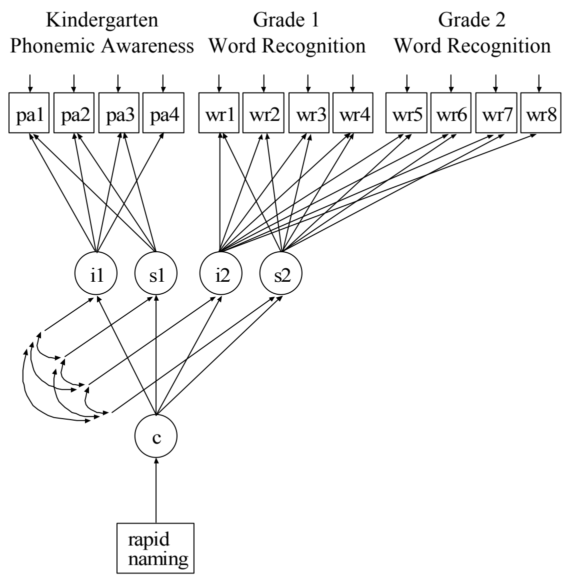
- Kindergarten children will differ in their growth and development in precursor skills
- The rate of development of the precursor skills will relate to the rate of development and the level of attainment of reading and spelling skills – and the individual growth rates in reading and spelling skills 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

51

Word Recognition Development In Grades 1 And 2

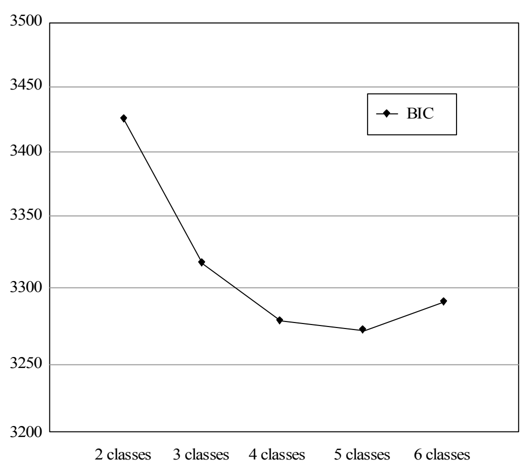


52



53

BIC Curve For Reading Skills Development



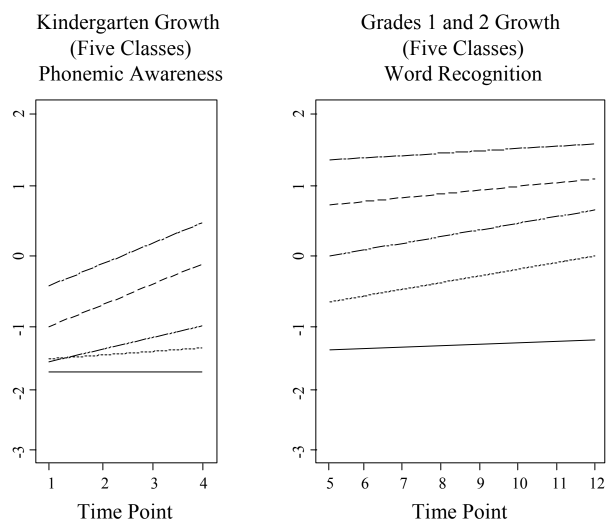
54

Input For Growth Mixture Model For Reading Skills Development

```
TITLE:     Growth mixture model for reading skills development

DATA:      FILE IS newran.dat;

VARIABLE:  NAMES ARE 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 MISSING;

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;
```

55

Five Classes Of Reading Skills Development



56

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)

57

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

		Full Model					
		1.00	2.00	3.00	4.00	5.00	Total
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

		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	145	100	49	411

58

Sensitivity And Specificity Of Early Classification (Continued)

		Full Model					
		1.00	2.00	3.00	4.00	5.00	Total
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	145	100	49	411

		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	145	100	49	411

		Full Model					
		1.00	2.00	3.00	4.00	5.00	Total
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	145	100	49	411

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	145	100	49	410

Growth Mixtures In Randomized Trials

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.

61

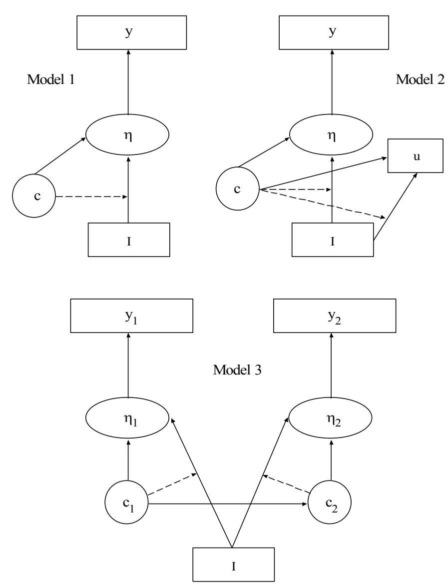


Figure 1. *Path Diagrams for Models 1 - 3*

62

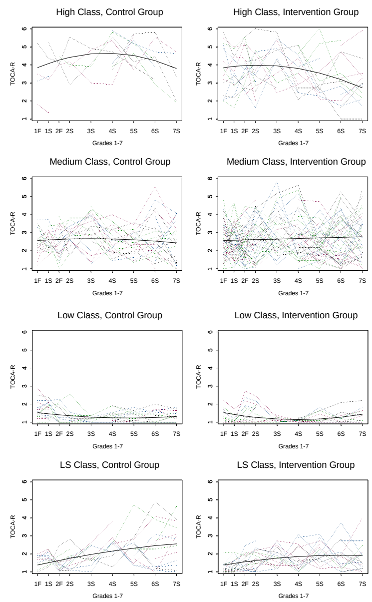
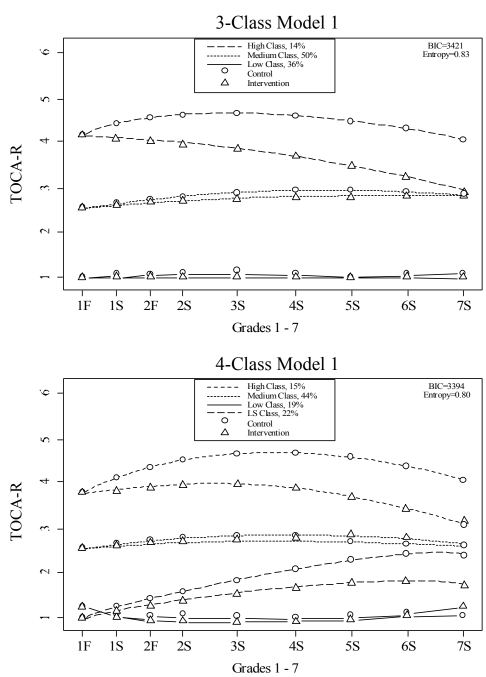


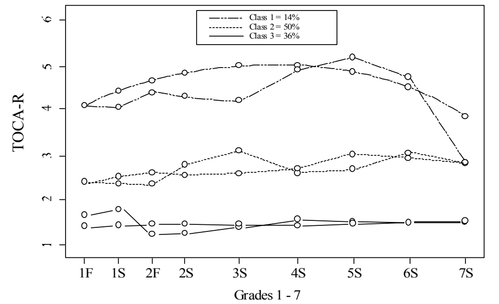
Figure 6. *Estimated Mean Growth Curves and Observed Trajectories for 4-Class model 1 by Class and Intervention Status*

63

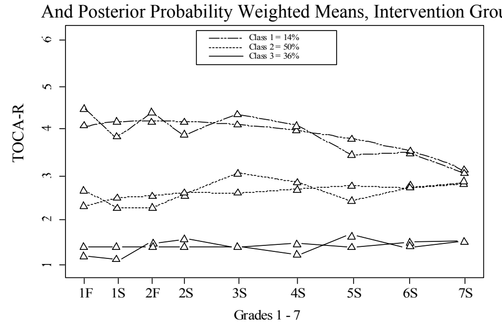


64

3-Class Model Estimated Mean Class Trajectories
And Posterior Probability Weighted Means, Control Group



3-Class Model Estimated Mean Class Trajectories
And Posterior Probability Weighted Means, Intervention Group



Input For Growth Mixtures In Randomized Trials

```
TITLE:     growth mixtures in randomized trials
DATA:      FILE IS toca.dat;
VARIABLE:  NAMES ARE sctaa11f sctaa11s sctaa12f sctaa12s sctaa13s
           sctaa14s sctaa15s sctaa16s sctaa17s intngrp;
           MISSING ARE ALL (999);
           USEVARIABLES ARE sctaa11f-sctaa17s tx;
           CLASSES = c(3);
DEFINE:    tx = (intngrp==4);
ANALYSIS:  TYPE = MIXTURE MISSING;
MODEL:     %OVERALL%
           ac bc qc | sctaa11f@0 sctaa11s@0.5 sctaa12f@1
           sctaa12s@1.5 sctaa13s@2.5 sctaa14s@3.5 sctaa15s@4.5
           sctaa16s@5.5 sctaa17s@6.5;
           qc@0;
           bc qc ON tx;
           sctaa11f WITH sctaa11s; sctaa12f WITH sctaa12s;
```

Input For Growth Mixtures In Randomized Trials (Continued)

```
%c#1%
[ac*3 bc qc]; bc qc ON tx;
%c#2%
[ac*2 bc qc]; bc qc ON tx;
%c#3%
[ac*1 bc qc]; bc qc ON tx;
ac sctaa11f-sctaa17s;
```

67

Further Readings On General Growth Mixture Modeling

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (ed.), Handbook of quantitative methodology for the social sciences (pp. 345-368). Newbury Park, CA: Sage Publications. (#100)

Muthén, B. & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. Biometrics, 55, 463-469. (#78)

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. (#87)

Muthén, B., Khoo, S.T., Francis, D. & Kim Boscardin, C. (2002). Analysis of reading skills development from Kindergarten through first grade: An application of growth mixture modeling to sequential processes. In S.R. Reise & N. Duan (eds), Multilevel modeling: Methodological advances, issues, and applications (pp. 71 – 89). Mahaw, NJ: Lawrence Erlbaum Associates. (#77)

68