An Introduction to the Latent Class Model

28 – 29 April 2011

University of Texas – Dallas

A. L. McCutcheon, Ph.D.

Donald O. Clifton Chair of Survey Science
Director, Gallup Research Center
Chair, Graduate program in Survey Research and Methodology
Professor, Statistics and Sociology
University of Nebraska-Lincoln

Email: AmcCutcheon1@UNL.edu

www.unl.edu/ALM

Software

Glimmix (Michel Wedel)

gamma.post@gamma.rug.nl

Latent Gold (Jeroen Vermunt)

marketing@statisticalinnovations.com

Mplus (Bengt &Linda Muthén)

support @statmodel.com

LEM (Jeroen Vermunt)

Panmark (F. van de Pol)

fpol@compuserve.com

Levels of Measurement

- Continuous (metric, quantitative)
  - Ratio
    - True zero
    - Known metric
    - Ordered
    - Mutually exclusive categories
  - Interval
    - Known metric
    - Ordered
    - Mutually exclusive categories

Levels of Measurement (cont.)

- Categorical (discrete, qualitative)
  - Ordinal
    - Ordered
    - Mutually exclusive categories
  - Nominal
    - Mutually exclusive categories
- Latent class analysis focuses primarily on ***categorical*** data

Variable Types

- Observability
  - Manifest variables
    - Directly observed (e.g., frequency of church attendance)
    - Indicator variables
  - Latent variables
    - Indirectly observed (e.g., religiosity)
    - Frequently, variables of ***true*** interest

Variable Types (cont.)

- Causality
  - Independent (effects) variables
  - Dependent (response) variables
  - Intervening (mediating) variables
- Basic LCA's are non-causal in the traditional meaning of the term

Motivating Example

1. You are riding in a car driven by a close friend, and he hits a pedestrian. You know that he is going at least 35 miles an hour in a 20-mile-an-hour speed zone. There are no other witnesses. His lawyer says that if you testify under oath that the speed was only 20 miles an hour, it may save him from serious consequences. What right has your friend to expect you to protect him? (Universalistic: He has no right as a friend to expect me to testify to the lower figure.)

2. As a drama critic, your friend asks you to "go easy on a review" of a bad play in which all of his saving are invested.

3. As a physician, your friend asks you to "shade doubts" about a physical examination for an insurance policy.

4. As a member of the board of directors, does your friend have a right to expect that you "tip him off" about financially ruinous, though secret, company information.

Motivating Example

Samuel A. Stouffer and Jackson Toby (1951) "Role Conflict and Personality," *American Journal of Sociology* **56**: 395-406.

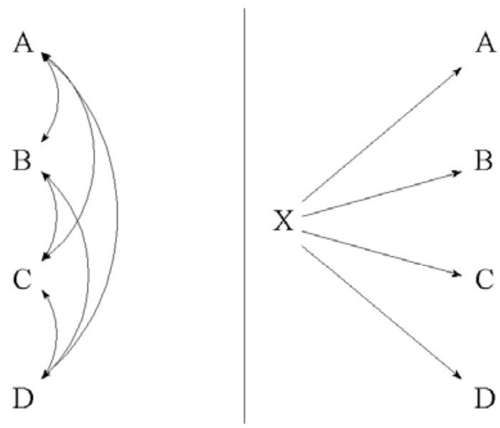
1	2	3	4		1	2	3	4	
+	+	+	+	20	-	+	+	+	38
+	+	+	-	2	-	+	+	-	7
+	+	-	+	6	-	+	-	+	25
+	+	-	-	1	-	+	-	-	6
+	-	+	+	9	-	-	+	+	24
+	-	+	-	2	-	-	+	-	6
+	-	-	+	4	-	-	-	+	23
+	-	-	-	1	-	-	-	-	42

Motivating Example

Table 1: Probability of Universalistic Response (-) and Relative Frequency for the Two-Class Model

Observed Variables	Respondent Type: Universalistic	Respondent Type: Particularistic
Auto Passenger	.993	.714
Drama Critic	.926	.354
Insurance Doctor	.940	.330
Board of Directors	.769	.132
Relative Class Frequency	.2792	.7203

Local Independence



Local Independence

	A = 1	A = 2
B = 1	95	55
B = 2	70	80

C = 1

	A = 1	A = 2
B = 1	80	20
B = 2	40	10

C = 2

	A = 1	A = 2
B = 1	15	35
B = 2	30	70

The I-by-J Table

- Assume we have 2 variables, A and B, and that A is causally related to B (i.e., $A \rightarrow B$)
- Further, let "i" index A and "j" index B, where i=1,…,I and j=I,…,J
- For example, if A_i is respondent's religious identification with, 1=Protestant, 2=Catholic, 3=Jewish, 4=None, 5=Other; (i.e., I=5), then A_2 represents the Catholics
- Let the cell count for the joint distribution of each of the IJ combinations of i and j be represented as f_{ij}

The I-by-J Table (cont.)

- The resulting I-by-J rectangular display of <u>cell</u> counts
  - contingency table
  - cross-classification table
  - "crosstabs"
- Frequencies/cell counts (f_{ij}) and probabilities (p_{ij})
  - $p_{ij} = f_{ij}/n$, where n is sample size (i.e., total number of observations [counts] recorded in the contingency table)
- Probability distributions (Agresti uses ✍$_{ij}$ as pop. parameters, we use p_{ij} and will note sample para.)
  - joint probability (distribution)
  - marginal probability (distribution)
  - conditional probability (distribution)

The I-by-J Table (cont.)

B

A		1	2	3	4	
	1	f_{11}	f_{12}	f_{13}	f_{14}	f_{1+}
	2	f_{21}	f_{22}	f_{23}	f_{24}	f_{2+}
	3	f_{31}	f_{32}	f_{33}	f_{34}	f_{3+}
	4	f_{41}	f_{42}	f_{43}	f_{44}	f_{4+}
	5	f_{51}	f_{52}	f_{53}	f_{54}	f_{5+}
		f_{+1}	f_{+2}	f_{+3}	f_{+4}	f_{++}

The I-by-J Table (cont.)

B

A		1	2	…	J	
	1	f_{11}	f_{12}	…	f_{1J}	f_{1+}
	2	f_{21}	f_{22}	…	f_{2J}	f_{2+}
	3	f_{31}	f_{32}	…	f_{3J}	f_{3+}
	…	…	…	…	..	…
	I	f_{I1}	f_{I2}	…	f_{IJ}	f_{I+}
		f_{+1}	f_{+2}	…	f_{+J}	f_{++}

Probability Distributions

- Joint Probability (distribution)

$$p_{ij} = f_{ij} / n, \quad \sum_{i,j} p_{ij} = 1$$

- Marginal Probability (distribution)

$$p_{i+} = \sum_{j} p_{ij} = \sum_{j} f_{ij} / n = f_{i+} / n, \quad \sum_{i} p_{i+} = 1$$

$$p_{+j} = \sum_{i} p_{ij} = \sum_{i} f_{ij} / n = f_{+j} / n, \quad \sum_{j} p_{+j} = 1$$

- Conditional Probability (distribution)

$$p_{j|i} = p_{ij} / p_{i+} = f_{ij} / f_{i+}, \quad \sum_{j} p_{j|i} = 1$$

Independence

- Two variables are statistically independent when

$$p_{ij} = p_{i+} p_{+j} \quad for\ i = 1, \cdots, I \quad and \quad j = 1, \cdots, J$$

- Thus, with independence

$$p_{j|i} = p_{ij} \big/ p_{i+} = (p_{i+} p_{+j}) \big/ p_{i+} = p_{+j}$$

Local Independence

B \ A	1	2
1	95	55
2	70	80

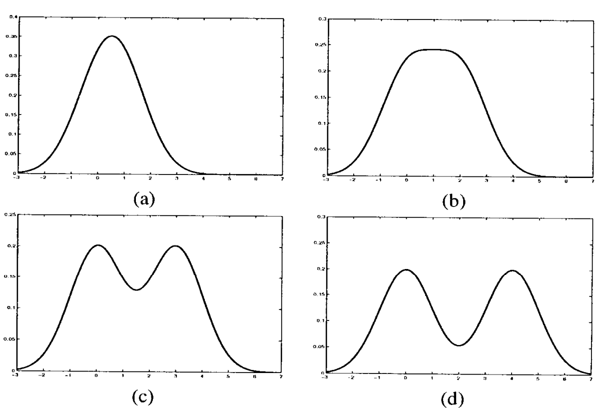
C = 1

B \ A	1	2
1	80	20
2	40	10

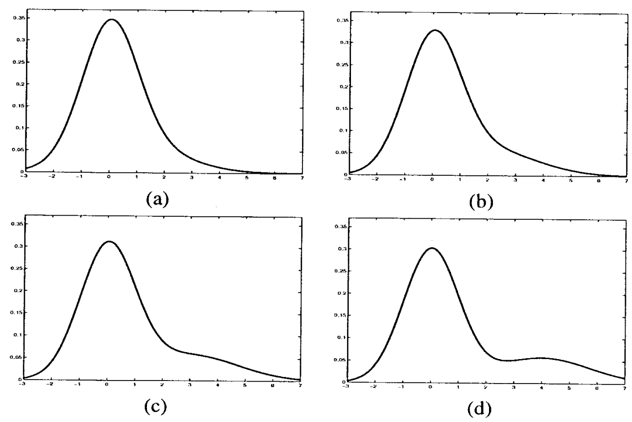
C = 2

B \ A	1	2
1	15	35
2	30	70

Mixtures



Mixtures (continued)



Introduction to the Latent Class Model 21

Finite Mixture Models

- **Discrete (finite) number of classes**

  $$f(y|z) = \Sigma_x \, \pi(x|z) \, \pi_m \, f(y_m|x,z)$$

- **$\pi(x|z)$ is the "mixing proportion"**

  $$\Sigma_x \, \pi(x|z) = 1.0$$

- **$f(y_m|x,z)$ is the "link function"**
  here a probability density

Introduction to the Latent Class Model 22

Basic Latent Class Model

- **Assume 4 dichotomous indicators: (A_i, B_j, C_k, D_l)**

$$\pi_{ijkl}^{ABCD} = \sum_{t=1}^{T} \pi_t^X \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X}$$

- **X_t is the latent variable**

$$\sum_{t=1}^{T} \pi_t^X = 1.0$$

Belief in God (Measurement Error): 1991 ISSP

Q.14 Please tick one box below to show which statement comes closest to expressing what you believe about God. (Please tick one box only)

----

1. ***I don't believe in God***
2. I don't know whether there is a God and I don't believe there is any way to find out
3. I don't believe in a personal God, but I do believe in a Higher Power of some kind
4. I find myself believing in God some of the time, but not at others
5. While I have doubts, I feel that I do believe in God
6. I know God really exists and I have no doubts about it
8. Can't choose, don't know

Q.15 How close do you feel to God most of the time? (Please tick one box only)

----

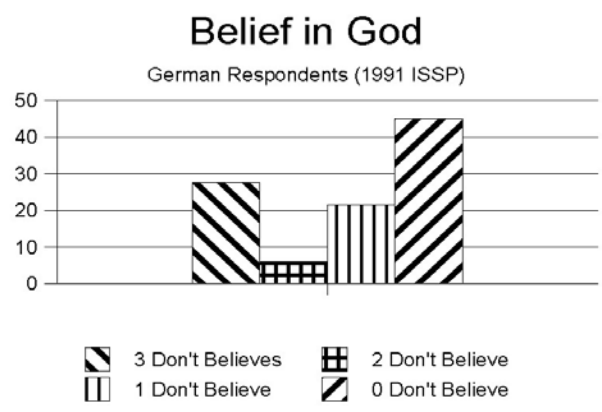
1. ***Don't believe in God***
2. Not close at all
3. Not very close
4. Somewhat close
5. Extremely close
8. Can't choose, don't know
9. NA, refused

Belief in God (Measurement Error): 1991 ISSP

Q.16 Which best describes your beliefs about God?
(Please tick one box only)
----

1. ***I don't believe in God now and I never have***
2. ***I don't believe in God now, but I used to***
3. I believe in God now, but I didn't used to
4. I believe in God now and I always have
8. Can't choose, don't know
9. NA, refused

Belief in God (Measurement Error):1991 ISSP



Belief in God (Measurement Error): 1991 ISSP

Q14	Q15	Q16	
+	+	+	1272
+	+	-	582
+	-	+	5
+	-	-	113
-	+	+	25
-	+	-	45
-	-	+	9
-	-	-	781

Motivating Example

1. You are riding in a car driven by a close friend, and he hits a pedestrian. You know that he is going at least 35 miles an hour in a 20-mile-an-hour speed zone. There are no other witnesses. His lawyer says that if you testify under oath that the speed was only 20 miles an hour, it may save him from serious consequences. What right has your friend to expect you to protect him? (Universalistic: He has no right as a friend to expect me to testify to the lower figure.)

2. As a drama critic, your friend asks you to “go easy on a review” of a bad play in which all of his saving are invested.

3. As a physician, your friend asks you to “shade doubts” about a physical examination for an insurance policy.

4. As a member of the board of directors, does your friend have a right to expect that you “tip him off” about financially ruinous, though secret, company information.

Motivating Example

16-fold response pattern (2x2x2x2=2^4)

1	2	3	4	1	2	3	4
+	+	+	+	-	+	+	+
+	+	+	-	-	+	+	-
+	+	-	+	-	+	-	+
+	+	-	-	-	+	-	-
+	-	+	+	-	-	+	+
+	-	+	-	-	-	+	-
+	-	-	+	-	-	-	+
+	-	-	-	-	-	-	-

Motivating Example

Samuel A. Stouffer and Jackson Toby (1951) "Role Conflict and Personality," *American Journal of Sociology* **56**: 395-406.

1	2	3	4		1	2	3	4	
+	+	+	+	20	-	+	+	+	38
+	+	+	-	2	-	+	+	-	7
+	+	-	+	6	-	+	-	+	25
+	+	-	-	1	-	+	-	-	6
+	-	+	+	9	-	-	+	+	24
+	-	+	-	2	-	-	+	-	6
+	-	-	+	4	-	-	-	+	23
+	-	-	-	1	-	-	-	-	42

Motivating Example

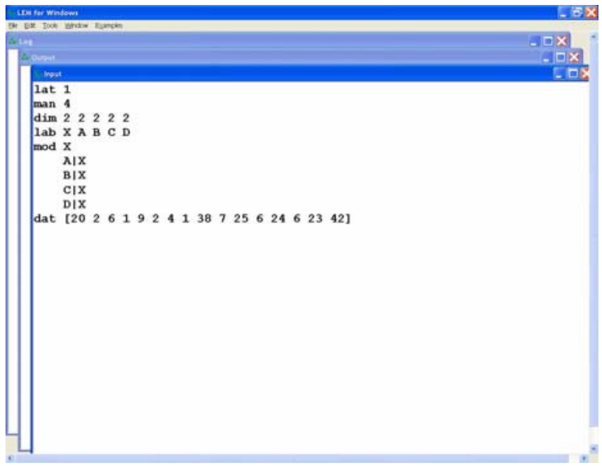
Table 1: Probability of Universalistic Response (-) and Relative Frequency for the Two-Class Model

Observed Variables	Respondent Type: Universalistic	Particularistic
Auto Passenger	.993	.714
Drama Critic	.926	.354
Insurance Doctor	.940	.330
Board of Directors	.769	.132
Relative Class Frequency	.2792	.7203

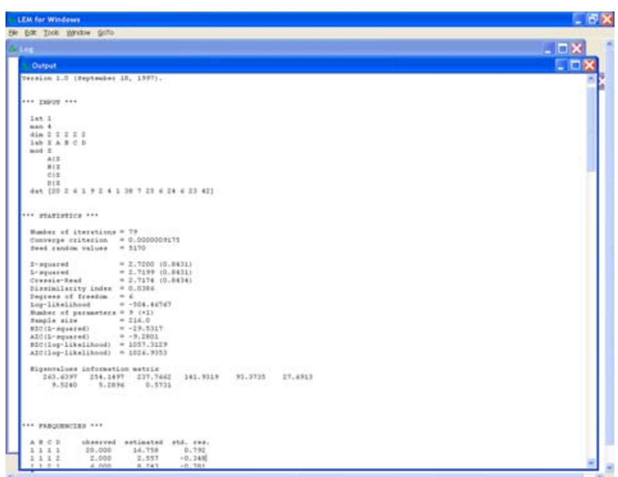
Exercise 1: Basic LEM Input

```
lat 1
man 4
dim 2 2 2 2 2
lab X A B C D
mod X
    A|X
    B|X
    C|X
    D|X
dat [20 2 6 1 9 2 4 1 38 7 25 6 24 6 23 42]
```

Exercise 1: LEM Basics



Example: LEM Basics (cont.)



```
*** STATISTICS ***

  Number of iterations = 77
  Converge criterion   = 0.0000009747
  Seed random values   = 4741

  X-squared            = 2.7200 (0.8431)
  L-squared            = 2.7199 (0.8431)
  Cressie-Read         = 2.7174 (0.8434)
  Dissimilarity index  = 0.0386
  Degrees of freedom   = 6
  Log-likelihood       = -504.46767
  Number of parameters = 9 (+1)
  Sample size          = 216.0
  BIC(L-squared)       = -29.5317
  AIC(L-squared)       = -9.2801
  BIC(log-likelihood)  = 1057.3129
  AIC(log-likelihood)  = 1026.9353

  Eigenvalues information matrix
     263.6368   254.1479   237.7647   141.9316    93.3723    27.6927
       9.5250     5.2903     0.5734
```


*** FREQUENCIES ***

A B C D	observed	estimated	std. res.
1 1 1 1	20.000	16.758	0.792
1 1 1 2	2.000	2.557	-0.348
1 1 2 1	6.000	8.243	-0.781
1 1 2 2	1.000	1.278	-0.246
1 2 1 1	9.000	9.186	-0.061
1 2 1 2	2.000	1.418	0.489
1 2 2 1	4.000	4.596	-0.278
1 2 2 2	1.000	0.965	0.036
2 1 1 1	38.000	41.804	-0.588
2 1 1 2	7.000	6.570	0.168
2 1 2 1	25.000	21.475	0.761
2 1 2 2	6.000	6.315	-0.125
2 2 1 1	24.000	23.643	0.073
2 2 1 2	6.000	6.064	-0.026
2 2 2 1	23.000	23.294	-0.061
2 2 2 2	42.000	41.834	0.026

$$\frac{f - \hat{f}}{\sqrt{\hat{f}}}$$

```
*** (CONDITIONAL) PROBABILITIES ***

* P(X) *

  1            0.2794  (0.0581)
  2            0.7206  (0.0581)

* P(A|X) *

  1 | 1        0.0068  (0.0253)
  2 | 1        0.9932  (0.0253)
  1 | 2        0.2865  (0.0404)
  2 | 2        0.7135  (0.0404)

* P(B|X) *

  1 | 1        0.0736  (0.0656)
  2 | 1        0.9264  (0.0656)
  1 | 2        0.6461  (0.0486)
  2 | 2        0.3539  (0.0486)

* P(C|X) *

  1 | 1        0.0604  (0.0660)
  2 | 1        0.9396  (0.0660)
  1 | 2        0.6705  (0.0497)
  2 | 2        0.3295  (0.0497)

* P(D|X) *

  1 | 1        0.2310  (0.0952)
  2 | 1        0.7690  (0.0952)
  1 | 2        0.8677  (0.0383)
  2 | 2        0.1323  (0.0383)
```

π_t^X

$\pi_{it}^{A|X}$ $(also\ \pi_{it}^{\bar{A}X})$

$\pi_{jt}^{B|X}$

$\pi_{kt}^{C|X}$

$\pi_{lt}^{D|X}$

Motivating Example

Table 1: Probability of Universalistic Response (-) and Relative Frequency for the Two-Class Model

Observed Variables	Respondent Type: Universalistic	Respondent Type: Particularistic
Auto Passenger	.993	.714
Drama Critic	.926	.354
Insurance Doctor	.940	.330
Board of Directors	.769	.132
Relative Class Frequency	.2792	.7203

How do you calculate the expected (f-hat)?

$$\pi_{ijklt}^{ABCDX} = \pi_t^X \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X}$$

And

$$\pi_{ijkl}^{ABCD} = \sum_{t=1}^{T} \pi_t^X \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X}$$

So, for example, for response pattern 2, 1, 2, 1, we have

$$\pi_{21211}^{ABCDX} = .2794 * .9932 * .0736 * .9396 * .2310$$
$$= .0044330$$

$$\pi_{21212}^{ABCDX} = .7206 * .7135 * .6461 * .3295 * .8677$$
$$= .0949758$$

And, thus

$$\pi_{2121}^{ABCD} = .0994088$$

$$\hat{f}_{2121} = \pi_{2121}^{ABCD} * N = .094088 * 216 = 21.4723$$

```
2 1 1 2       7.000      6.570      0.168
2 1 2 1      25.000     21.475      0.761
2 1 2 2       6.000      6.315     -0.125
```

Modal Probabilities:

$$\pi_{12121}^{X|ABCD} = \pi_{21211}^{ABCDX} / \pi_{2121}^{ABCD} = .004433/.0994088$$
$$= .0445936$$

$$\pi_{22121}^{X|ABCD} = \pi_{21212}^{ABCDX} / \pi_{2121}^{ABCD} = .0949758/.0994088$$
$$= .9554064$$

Exercise 1: Mplus Basics

```
TITLE:     this is an example of a LCM using the Stouffer and Toby data
           that has two latent classes and uses automatic start values with
           random starts
           MpLCA1
DATA:      FILE IS c:\workshops\utdallas05\lca1.dat;
VARIABLE:  NAMES ARE u1-u4;
           CLASSES = c (2);
           CATEGORICAL = u1-u4;
ANALYSIS:  TYPE = MIXTURE;
OUTPUT:    TECH1 TECH8;
```


Exercise 1: Mplus Basics (cont.)

LCA1.dat

```
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 1
1 1 1 2
1 1 1 2
1 1 2 1
1 1 2 1
...
```

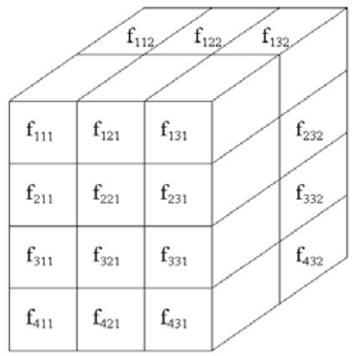
Unrestricted and Restricted Models

I-by-J Table

B

A	1	2	3	4	
1	f_{11}	f_{12}	f_{13}	f_{14}	f_{1+}
2	f_{21}	f_{22}	f_{23}	f_{24}	f_{2+}
3	f_{31}	f_{32}	f_{33}	f_{34}	f_{3+}
4	f_{41}	f_{42}	f_{43}	f_{44}	f_{4+}
5	f_{51}	f_{52}	f_{53}	f_{54}	f_{5+}
	f_{+1}	f_{+2}	f_{+3}	f_{+4}	f_{++}

Multi-way (IxJxK) Tables



Multi-way (IxJxK) Tables (cont.)

	C = 1				C = 2			
	B				B			
A	1	2	3		1	2	3	
1	f_{111}	f_{121}	f_{131}	f_{1+1}	f_{112}	f_{122}	f_{132}	f_{1+2}
2	f_{211}	f_{221}	f_{231}	f_{2+1}	f_{212}	f_{222}	f_{232}	f_{2+2}
3	f_{311}	f_{321}	f_{331}	f_{3+1}	f_{312}	f_{322}	f_{332}	f_{3+2}
4	f_{411}	f_{421}	f_{431}	f_{4+1}	f_{412}	f_{422}	f_{432}	f_{4+2}
	f_{+11}	f_{+21}	f_{+31}	f_{++1}	f_{+12}	f_{+22}	f_{+32}	f_{++2}

Multi-way (IxJxK) Tables (cont.)

A	B	C	count	A	B	C	count
1	1	1	f_{111}	3	1	1	f_{311}
1	1	2	f_{112}	3	1	2	f_{312}
1	2	1	f_{121}	3	2	1	f_{321}
1	2	2	f_{122}	3	2	2	f_{322}
1	3	1	f_{131}	3	3	1	f_{331}
1	3	2	f_{132}	3	3	2	f_{332}
2	1	1	f_{211}	4	1	1	f_{411}
2	1	2	f_{212}	4	1	2	f_{412}
2	2	1	f_{221}	4	2	1	f_{421}
2	2	2	f_{222}	4	2	2	f_{422}
2	3	1	f_{231}	4	3	1	f_{431}
2	3	2	f_{232}	4	3	2	f_{432}

Example

A A publicly available web site which allows customers to enter orders for products or services online, and for which your company has dedicated significant resources, and which represents a strategic commitment for your firm?

B An Extranet which allows selected businesses to place orders or check inventories via a web interface

C An Extranet which is used to interact via a web interface with outside suppliers for critical applications such as finance, procurement, or ERP

D A web site from which your company derives at least 25 percent of its overall revenue from web-based advertising, affiliate programs or subscriptions.

Multi-way Response Table as Array

A	B	C	D	count
1	1	1	1	2
1	1	1	2	12
1	1	2	1	1
1	1	2	2	11
1	2	1	1	0
1	2	1	2	8
1	2	2	1	3
1	2	2	2	42

A	B	C	D	count
2	1	1	1	0
2	1	1	2	3
2	1	2	1	0
2	1	2	2	9
2	2	1	1	1
2	2	1	2	17
2	2	2	1	5
2	2	2	2	147

Exercise 2: Basic LEM Input

```
* Unrestricted latent class model
* A = Web site allows customer orders
* B = Extranet allows selected business to
*     place orders, etc.
* C = Extranet which is used to interacct w/
*     outside suppliers
* D = Web site from which company derives at
*     least 25% revenue

lat 1
man 4
dim 2 2 2 2 2
lab X A B C D
mod X
    A|X
    B|X
    C|X
    D|X
dat [2 12 1 11 0 8 3 42 0 3 0 9 1 17 5 147 ]
see 79144
```

```
*** STATISTICS ***

  Number of iterations = 104
  Converge criterion   = 0.0000008765
  Seed random values   = 79144

  X-squared            = 2.5470 (0.8632)
  L-squared            = 3.5306 (0.7399)
  Cressie-Read         = 2.7190 (0.8432)
  Dissimilarity index  = 0.0174
  Degrees of freedom   = 6
  Log-likelihood       = -406.73333
  Number of parameters = 9 (+1)
  Sample size          = 261.0
  BIC(L-squared)       = -29.8566
  AIC(L-squared)       = -8.4694
  BIC(log-likelihood)  = 863.5473
  AIC(log-likelihood)  = 831.4667

  Eigenvalues information matrix
     292.3539   157.9235    93.4596    63.9076    58.5344    34.0635
      19.9269     6.7323     3.1609
```



```
*** FREQUENCIES ***

  A B C D     observed  estimated  std. res.
  1 1 1 1       2.000      1.444      0.463
  1 1 1 2      12.000     12.600     -0.169
  1 1 2 1       1.000      1.092     -0.088
  1 1 2 2      11.000     10.848      0.046
  1 2 1 1       0.000      0.477     -0.691
  1 2 1 2       8.000      7.374      0.230
  1 2 2 1       3.000      1.839      0.856
  1 2 2 2      42.000     43.326     -0.201
  2 1 1 1       0.000      0.225     -0.474
  2 1 1 2       3.000      2.538      0.290
  2 1 2 1       0.000      0.435     -0.660
  2 1 2 2       9.000      8.818      0.061
  2 2 1 1       1.000      0.719      0.331
  2 2 1 2      17.000     17.623     -0.148
  2 2 2 1       5.000      5.769     -0.320
  2 2 2 2     147.000    145.873      0.093
```

```
* (CONDITIONAL) PROBABILITIES ***

* P(X) *

  1              0.1216  (0.0432)
  2              0.8784  (0.0432)

* P(A|X) *

  1 | 1          0.8836  (0.1113)
  2 | 1          0.1164  (0.1113)
  1 | 2          0.2223  (0.0360)
  2 | 2          0.7777  (0.0360)

* P(B|X) *

  1 | 1          0.8355  (0.2017)
  2 | 1          0.1645  (0.2017)
  1 | 2          0.0501  (0.0227)
  2 | 2          0.9499  (0.0227)

* P(C|X) *

  1 | 1          0.5878  (0.1167)
  2 | 1          0.4122  (0.1167)
  1 | 2          0.1062  (0.0259)
  2 | 2          0.8938  (0.0259)

* P(D|X) *

  1 | 1          0.1041  (0.0607)
  2 | 1          0.8959  (0.0607)
  1 | 2          0.0379  (0.0131)
  2 | 2          0.9621  (0.0131)
```


Table 2: Probability of "Yes" Web Page has Functionality and Latent Class Probabilities for the Two-Class Model

Observed Variables	Web Page Type: Working Page	Poster Page
Take Customer Orders	.884	.222
Extranet for Business	.836	.050
Extranet for Suppliers	.588	.106
Revenue of 25% or more	.104	.038
Relative Class Frequency	.1216	.8784

Causal and non-causal associations

- -Asymmetric
  - -classic case: A → B
- -Symmetric
  - -alternative indicators of same concept: indicator variables
  - -parts of a common “system” or “complex”
  - -functional interdependence of elements
  - -effects of a common cause
  - -fortuitous

Latent and manifest measures

Why latent?
- -Measurement error
- -Unobserved heterogeneity
  - -Mixtures

Latent	Manifest: Categorical	Manifest: Continuous
Categorical	LCA	L Profile A
Continuous	LTA	Trad. FA

Probabilities

- Marginal e.g., P(a)
- Joint e.g., P(a,b)
- Conditional e.g., P(a|b)

Basics:

e.g. $P(a,b) = P(a|b)P(b)$
$P(a|b) = P(a,b)/P(b)$

Local (conditional) Independence

$$P(a,b,c,d|x)=P(a|x)P(b|x)P(c|x)P(d|x)\ P(x)$$

Independence:

-True independence: $P(a,b) = P(a)P(b)$

-Conditional or local Independence:

$$P(a,b,c) = P(c_k)P(a| c_k)P(b| c_k)$$

$$P(a,b) = \Sigma_k P(c_k)P(a| c_k)P(b| c_k)$$

$$P(a,b|c) = P(a| c_k)P(b| c_k)$$

	A = 1	A = 2
B = 1	95	55
B = 2	70	80

C = 1

	A = 1	A = 2
B = 1	80	20
B = 2	40	10

C = 2

	A = 1	A = 2
B = 1	15	35
B = 2	30	70

Marginal Table (top): $X^2 = 8.42$, 1 df
$P(a_i,b_j) = P(a_i) P(b_j)$

Partial Table (bottom):
$P(a_i,b_j|c_k) = P(a_i|c_k) P(b_j|c_k)$

Local or conditional independence

Symbols: P (observed, manifest) vs
π (unobserved, latent)

Conditional notation: e.g.

$$\pi_{it}^{A|X}$$

A_i where i = 1, … , I
B_j where j = 1, … , J
etc.

X_t where t = 1, … , T

The formal Latent Class Analysis (LCA) model
and other mixture (LCA) models

Bases
- -Measurement
- -Unobserved heterogeneity
  - -local independence (local homogeneity)

Parameterizations
- -Probabilistic
- -Loglinear

Formal Latent Class Model

$$\pi_{ijkl}^{ABCD} = \sum_{t=1}^{T} \pi_t^X \; \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X}$$

Where

$$\pi_t^X$$

is the latent class, or "mixing," probabilities, and

$$\pi_{it}^{A|X}, \quad \pi_{jt}^{B|X}, \quad \pi_{kt}^{C|X}, \; and \quad \pi_{lt}^{D|X}$$

are the conditional probabilities linking each of the indicator variables to the latent variable (X_t).

It follows that

$$\hat{F}_{ijkl} = N * \pi_{ijkl}^{ABCD}$$

Where

$$N = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{l=1}^{L} F_{ijkl}$$

When there are four indicator variables (A_i, B_j, C_k, D_l).

This can be generalized to any number of indicator variables. The current examples will assume four indicator variables.

Identifiability: the degree to which there is sufficient information in the sample observations to estimate the parameters in a proposed model.

Identifying Restrictions for LCMs:

Latent class probabilities

$$\sum_{t=1}^{T} \pi_t^X = 1.0$$

Thus, T-1 latent class probabilities must be estimated for the unrestricted LCM.

Conditional probabilities

$$\sum_{i=1}^{I} \pi_{it}^{A|X} = 1.0, \quad \sum_{j=1}^{J} \pi_{jt}^{B|X} = 1.0,$$

$$\sum_{k=1}^{K} \pi_{kt}^{C|X} = 1.0, \quad and \quad \sum_{l=1}^{L} \pi_{lt}^{D|X} = 1.0$$

Thus, for each of the T classes, I-1 parameters must be estimated for indicator variable A, J-1 for indicator variable B, K-l for indicator C, and L-1 for D.

So, for unrestricted LCMs, the number of necessary parameters that require estimation are:

$$(T-1)+T((I-1)+(J-1)+(K-1)+(L-1)=$$
$$T+T(I+J+K+L-4)-1=$$
$$T(I+J+K+L-3)-1$$

The information available in a contingency table of four variables is:

$$I*J*K*L-1$$

Thus, to have an overidentified model (i.e., one in which there is sufficient information to estimate a unique set of parameters), the unrestricted LCM must satisfy the following:

$$I*J*K*L-1 \quad > \quad T(I+J+K+L-3)-1$$

Note, however, that while this quick method of determining model identifiability works most of the time, there is one well-known instance in which it is known not to work (see Goodman's 1974 *Biometrika* article). Anecessary and sufficient condition for determining the local identifiability of an LCM involves determining the rank order of a matrix of partial derivatives of the nonredundant model parameters. Many programs, including LEM, will carry out the necessary calculations for determining the identifiability of a specific LCM.

df = (IJKL-1) – (T[I+J+K+L - 3] – 1)

The issue here is the maximum value permitted for T.

Example of Belief in God (3 dichotomous "I don't believe in God" indicator variables).

(2*2*2-1)>T(2+2+2 - 2) – 1 $\rightarrow$ 7 >T(4)-1

Example 2: Universalism

(2*2*2*2-1) > T(2+2+2+2-3)-1 $\rightarrow$ 15 > T(5) – 1

T < 3 (but…)

It may appear that simply adding yet another variable will solve many problems, since it gives us more degrees of freedom with which to test even more complex models.

Motivating Example

Table 1: Probability of Universalistic Response (-) and Relative Frequency for the Two-Class Model

Observed Variables	Respondent Type: Universalistic	Respondent Type: Particularistic
Auto Passenger	.993	.714
Drama Critic	.926	.354
Insurance Doctor	.940	.330
Board of Directors	.769	.132
Relative Class Frequency	.2792	.7203

Sparseness: many *sampling* zeros in observed dataset.

Latent Class Models (and loglinear models) are case intensive—they require relatively large sample sizes. All two-variable, three-variable and higher-order parameters are modeled, thus these models are information (i.e., sample size) intensive.

Sparseness leads to difficulties in model evaluation; specifically, to determining the number of degrees of freedom for the model test (Agresti, 1990*, Categorical Data Analysis*).

Resampling methods such as *jack-knifing* and *boot-strapping*, however, can be used to evaluate parameters (see Langeheine, Pannekoek and Van de Pol, 1996, SMR)

PANMARK provides *bootstrap* estimates of standard errors for LCM parameters estimated from sparse data tables.

Maximum likelihood estimation (mle) of latent class model parameters is through iterative estimation procedures: **EM** (expectation maximization; see McCutcheon 1987) or **Newton-Raphson**, or some variant of these two (see Vermunt, 1997*, Log-Linear Models for Event Histories*, esp. App. A, C, D)

Boundary Estimates

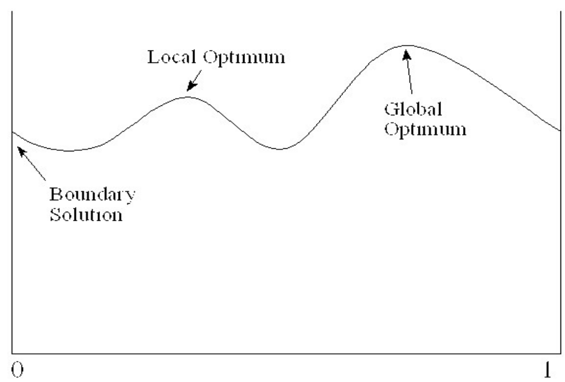


Figure 1: Maxima and Boundaries

75

Model Evaluation

Pearson Chi-Square

$$X^2 = \sum_{ijkl} \frac{(F_{ijkl} - \hat{F}_{ijkl})^2}{\hat{F}_{ijkl}}$$

Likelihood Ratio Chi-Square

$$G^2 = 2\sum_{ijkl} F_{ijkl} \ln\left(F_{ijkl} / \hat{F}_{ijkl} \right)$$

Where

$$\hat{F}_{ijkl} = N * \pi_{ijkl}^{ABCD}$$

Introduction to the Latent Class Model 76

Information criteria

Akaike Information Criteria (AIC)

$$AIC = G^2 - 2df$$

Bayesian Information Criteria (BIC)

$$BIC = G^2 - df * [\ln(N)]$$

Prefer lowest negative value for AIC and BIC.

Evaluation Criteria

```
X-squared            = 2.7200 (0.8431)
L-squared            = 2.7199 (0.8431)
Cressie-Read         = 2.7174 (0.8434)
Number of parameters = 9 (+1)
Sample size          = 216.0
BIC(L-squared)       = -29.5317
AIC(L-squared)       = -9.2801
```

Restricted Latent Class Models

Hypothesis Testing

- **-Equality restrictions**
- **-Deterministic restrictions**

- **-Conditional probabilities**
- **-Latent class probabilities**

Equality Restrictions on Conditional Probabilities

The *parallel indicators* hypothesis (e.g.)

$$\pi_{11}^{B|X} = \pi_{11}^{C|X} \quad and \quad \pi_{12}^{B|X} = \pi_{12}^{C|X}$$

Motivating Example

Table 1: Probability of Universalistic Response (-) and Relative Frequency for the Two-Class Model

Observed Variables	Respondent Type: Universalistic	Respondent Type: Particularistic
Auto Passenger	.993	.714
Drama Critic	.926	.354
Insurance Doctor	.940	.330
Board of Directors	.769	.132
Relative Class Frequency	.2792	.7203

Table 2: Latent Class Model Evaluation Criteria for Universalism Data

Model	X^2	G^2	AIC	BIC	DF
H_1: 2-Class Latent Class Model	2.72	2.72	-9.28	-29.53	6
H_2: H_1 + B & C Parallel Indicators	2.84	2.89	-13.11	-40.12	8
H_3: H_2 + D Equal Error Rate	3.60	3.65	-14.35	-44.73	9
H_4: H_3 + A as Perfect Indicator for Class 2	3.61	3.66	-16.34	-50.09	10

Conditional Likelihood Ratio Chi-Square tests

Equal Error Rates Hypothesis

$$\pi_{21}^{D|X} = \pi_{12}^{D|X}$$

Deterministic Restrictions on Conditional Probabilities

Perfect Indicator Hypothesis

$$\pi_{11}^{A|X} = 1.0$$

Table 3: Probability of Universalistic Response (-) and LC Probabilities for the Restricted Two-Class Model

Observed Variables	Respondent Type: Universalistic	Particularistic
Auto Passenger Friend	1.00*	.725
Drama Critic Friend	.954*	.364*
Insurance Doctor Friend	.954*	.364*
Board of Directors Friend	.852*	.148*
Latent Class Probabilities	.2426	.7574

LEM

```
lat 1
man 4
dim 2 2 2 2 2
lab X A B C D
mod X
    A|X
    B|X
    C|X eq1 B|X
    D|X eq2
des [1 0 0 1]
sta A|X [0 1 .5 .5]
dat [20 2 6 1 9 2 4 1 38 7 25 6 24 6 23 42]
```

Conditional Probability Restrictions

eq1 restricts current conditional probability estimate to equal that of some previously specified estimate

eq2 requires a design matrix and permits within class and across class equality restrictions

sta start values allows analyst to take control of start values. In current example, providing start values at the boundary fixes the values as “perfect indicator” hypothesis requires.

Mplus Specification for H_4

```
TITLE:          this is an example of a LCM using the Stouffer and Toby data which has
                two latent classes and the restrictions implied by Hypothesis 4
                MpLCA2
DATA:           FILE IS c:\workshops\utdallas05\lca1.dat;
VARIABLE:       NAMES ARE u1-u4;
                CLASSES = c (2);
                CATEGORICAL = u1-u4;
ANALYSIS:       TYPE = MIXTURE;
MODEL:
                %OVERALL%
                 %c#1%
                 [u1$1*-1];
                 [u2$1*-1] (1);
                 [u3$1*-1] (1);
                 [u4$1*-1] (p1);
                %c#2%
                 [u1$1@-15];
                 [u2$1*1] (2);
                 [u3$1*1] (2);
                 [u4$1*1] (p2);
MODEL CONSTRAINT:
                p2 = - p1;
OUTPUT:         TECH1 TECH8;
```


Model Selection and Evaluation

Some Basic Concepts

- Maximum likelihood principle
  - "In effect this principle says that when faced with several parameter values, any of which might be the true one for the population, the best 'bet' is that parameter value which would have made the sample actually obtained have the highest prior probability. When in doubt, place your bet on that parameter value which *would* have made the obtained result most likely." (Hays, 1981, pg. 182)
- Sampling distribution assumed
  - multinomial sampling
  - independent (product) multinomial sampling

Sampling Distributions

- Poisson sampling
  - counts (n_i) assumed as random variables
  - total sample size (n) is random variable
- Multinomial sampling
  - total sample size (n) is fixed
  - random sampling
  - n_i are conditioned on n; they can not exceed n
  - $\Sigma_i n_i = n$

$$P(n_i \mid n) = \left(\frac{n!}{\prod_i n_i!} \right) \prod_i \pi_i^{n_i}$$

Multinomial Distribution

- Most fundamental of all distributions
- Independent observations (frequencies)
- Multinomial distribution: with I possible outcomes

$$\Pr(n_1, n_2, \cdots, n_I) = \left(\frac{N!}{n_1! n_2! \cdots n_I!} \right) \prod_{i=1}^{I} \pi_i^{n_i}$$

- Binomial distribution: with 2 possible outcomes

$$\Pr(n_1, n_2) = \left(\frac{N!}{n_1!(N - n_1)!} \right) \pi_i^{n_i} (1 - \pi_1)^{n_2}$$

Sampling Distributions (cont.)

- Binomial distribution
  - I = 2
- Product (independent) multinomial sampling
  - When we fix n_i, such as in stratified sampling
  - Random sampling within strata

$$P(n_{j|i} \mid n_{i+}) = \left(\frac{n_{i+}!}{\prod_j n_{ij}!} \right) \prod_j \pi_{j|i}^{n_{ij}}$$

Probability and Likelihood

	p_1								
	0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
0	0.349	0.107	0.028	0.006	0.001	0.001			
1	0.387	0.268	0.121	0.041	0.011	0.002			
2	0.194	0.302	0.233	0.121	0.044	0.011	0.001		
3	0.057	0.201	0.267	0.215	0.117	0.042	0.009	0.001	
4	0.011	0.088	0.201	0.251	0.205	0.112	0.037	0.006	
5	0.002	0.026	0.103	0.201	0.246	0.201	0.103	0.026	0.002
6		0.006	0.037	0.112	0.205	0.251	0.201	0.088	0.011
7		0.001	0.009	0.042	0.117	0.215	0.267	0.201	0.057
8			0.001	0.011	0.044	0.121	0.233	302	0.194
9				0.002	0.011	0.041	0.121	0.268	0.387
10				0.001	0.001	0.006	0.028	0.107	0.349

Product Multinomial Sampling

- Prospective Stratification
  - Independent variables (n_i) are fixed
    - e.g., drawing random samples of 200 from each of 5 colleges
- Retrospective Stratification
  - Dependent variables (n_j) are fixed
    - e.g., drawing random samples of 200 teen mothers, 200 teenage women who not gotten pregnant and 200 who have had abortions
- Cross-sectional Studies
  - only the total sample size (n) is fixed
  - random sampling

Maximum likelihood estimation (mle) of latent class model parameters is through iterative estimation procedures: **EM** (expectation maximization; see McCutcheon 1987) or **Newton-Raphson**, or some variant of these two (see Vermunt, 1997*, Log-Linear Models for Event Histories*, esp. App. A, C, D)

Maximum Likelihood

- Given the observed data, the likelihood function is the probability of n_i for the sampling model, treated as an unknown set of parameters
- The factorials of the multinomial are constants, so the kernel of the probability function is

$$\prod_i \pi_i^{n_i}$$

- Maximize the log of the ML (monotonic function), log of the likelihood $L = \sum n_i \log \pi_i$

Maximum likelihood (cont.)

- Differentiate L with respect to π_i gives

$$\frac{\partial L}{\partial \pi_i} = \frac{n_i}{\pi_i} - \frac{n_I}{\pi_I} = 0$$

- The ML solution satisfies $\pi_i/\pi_I = n_i/n_I$

$$\sum \hat{\pi}_i = 1 = \frac{\hat{\pi}_I(\sum n_i)}{n_I} = \frac{\hat{\pi}_I n}{n_I}$$

- So, $\pi_I = n_I/n$ and $\pi_i = n_i/n = p_i$
- For contingency tables, the ML estimates of cell probabilities are the sample cell proportions

Boundary Estimates

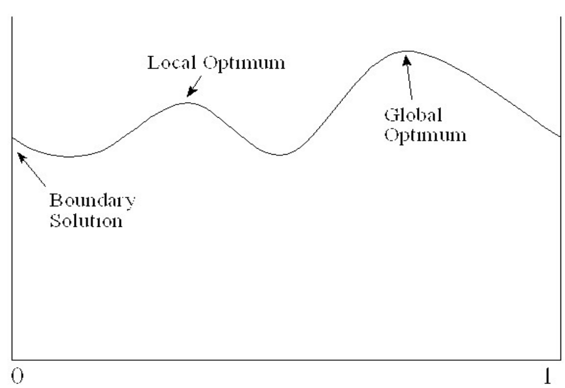


Figure 1: Maxima and Boundaries

Independence

- $E(p_{ij}) = p_{i+}p_{+j} = E(f_{ij}/n) = (f_{i+}f_{+j})/n = (n_{i+}n_{+j})/n$
- Pearson chi-squared statistics

$$X^2 = \sum_{i,j} \frac{(f_{ij} - \hat{f}_{ij})^2}{\hat{f}_{ij}}$$

- Likelihood ratio chi-squared statistic

$$G^2 = 2\sum_{i,j} f_{ij} \log(f_{ij} \big/ \hat{f}_{ij})$$

Hypothesis testing

- Degrees of freedom
  - Difference between knowns and unknowns
  - All models assume a set of known parameters
- Sparseness
  - When a “known” is not: sampling zeros
  - Structural zeros
- Chi-squared statistics
  - X^2 is pretty good approximation even with minimal (1.0) expected cell frequencies
  - G^2 often underestimates the probability of type I error and yields too high of value when expected cell sizes too small

Identifiability: the degree to which there is sufficient information in the sample observations to estimate the parameters in a proposed model.

Identifying Restrictions for LCMs:

Latent class probabilities

$$\sum_{t=1}^{T} \pi_t^X = 1.0$$

Thus, T-1 latent class probabilities must be estimated for the unrestricted LCM.

Degrees of freedom

B

A	1	2	3	4	
1	f_{11}	f_{12}	f_{13}	f_{14}	f_{1+}
2	f_{21}	f_{22}	f_{23}	f_{24}	f_{2+}
3	f_{31}	f_{32}	f_{33}	f_{34}	f_{3+}
4	f_{41}	f_{42}	f_{43}	f_{44}	f_{4+}
5	f_{51}	f_{52}	f_{53}	f_{54}	f_{5+}
	f_{+1}	f_{+2}	f_{+3}	f_{+4}	f_{++}

Degrees of Freedom (cont.)

- Knowns (data)
  - In the I x J contingency table: IJ
- Unknowns (parameters)
  - Sample size (n) is fixed, so we lose 1 df: (IJ-1)
  - Independence: I-1 row marginals and J-1 column marginals
  - Odds ratios (I-1)(J-1)
  - IJ = 1 + [(I-1) + (J-1)] + [(I-1)(J-1)]
- DF= (IJ-1) - estimated parameters
  - Independence model: df = (I-1)(J-1)
- Identifiability and Model Identification

Degrees of Freedom

- How many "knowns" do we have?
  - For the ABCD table, we have (IJKL)-1
- How many (unique) parameters are we estimating?
  - Recall the identifying restrictions (e.g., I-1, [I-1][J-1], etc.)
- DF = (IJKL-1) - # of estimated parameters

Conditional probabilities

$$\sum_{i=1}^{I} \pi_{it}^{A|X} = 1.0, \quad \sum_{j=1}^{J} \pi_{jt}^{B|X} = 1.0,$$

$$\sum_{k=1}^{K} \pi_{kt}^{C|X} = 1.0, \quad and \quad \sum_{l=1}^{L} \pi_{lt}^{D|X} = 1.0$$

Thus, for each of the T classes, I-1 parameters must be estimated for indicator variable A, J-1 for indicator variable B, K-l for indicator C, and L-1 for D.

So, for unrestricted LCMs, the number of necessary parameters that require estimation are:

$$(T-1) + T((I-1) + (J-1) + (K-1) + (L-1) =$$
$$T + T(I + J + K + L - 4) - 1 =$$
$$T(I + J + K + L - 3) - 1$$

The information available in a contingency table of four variables is:

$$I * J * K * L - 1$$

Thus, to have an overidentified model (i.e., one in which there is sufficient information to estimate a unique set of parameters), the unrestricted LCM must satisfy the following:

$$I * J * K * L - 1 \quad > \quad T(I + J + K + L - 3) - 1$$

Note, however, that while this quick method of determining model identifiability works most of the time, there is one well-known instance in which it is known not to work (see Goodman's 1974 *Biometrika* article). Anecessary and sufficient condition for determining the local identifiability of an LCM involves determining the rank order of a matrix of partial derivatives of the nonredundant model parameters. Many programs, including LEM, will carry out the necessary calculations for determining the identifiability of a specific LCM.

df = (IJKL-1) – (T[I+J+K+L - 3] – 1)

The issue here is the maximum value permitted for T.

Example of Belief in God (3 dichotomous "I don't believe in God" indicator variables).

(2*2*2-1)>T(2+2+2 - 2) – 1 $\rightarrow$ 7 >T(4)-1

Example 2: Universalism

(2*2*2*2-1) > T(2+2+2+2-3)-1 → 15 > T(5) – 1

T < 3 (but…)

It may appear that simply adding yet another variable will solve many problems, since it gives us more degrees of freedom with which to test even more complex models.

Sparseness: many *sampling* zeros in observed dataset.

Latent Class Models (and loglinear models) are case intensive—they require relatively large sample sizes. All two-variable, three-variable and higher-order parameters are modeled, thus these models are information (i.e., sample size) intensive.

Sparseness leads to difficulties in model evaluation; specifically, to determining the number of degrees of freedom for the model test (Agresti, 1990, *Categorical Data Analysis*).

Resampling methods such as *jack-knifing* and *boot-strapping*, however, can be used to evaluate parameters (see Langeheine, Pannekoek and Van de Pol, 1996, SMR)

PANMARK provides *bootstrap* estimates of standard errors for LCM parameters estimated from sparse data tables.

Model Evaluation

- Using the expected probabilities and n_i we can obtain an expected value for each cell in the table

$$\hat{f}_{ij} = \hat{\pi}_{ij} * n_i$$

- Goodness of fit statistic for model (M)

$$G^2(M) = 2\sum_i \sum_j f_{ij} \log\left(f_{ij} / \hat{f}_{ij} \right)$$

- Independence model (M_0) has df=(I-1)(J-1)

Calculating Expected Cell Counts

- Basic latent class model

$$\pi_{ijklt}^{ABCDX} = \pi_t^X \, \pi_{it}^{A|X} \, \pi_{jt}^{B|X} \, \pi_{kt}^{C|X} \, \pi_{lt}^{D|X}$$

- Expected cell proportions

$$\pi_{ijkl}^{ABCD} = \sum_{t=1}^{T} \pi_t^X \, \pi_{it}^{A|X} \, \pi_{jt}^{B|X} \, \pi_{kt}^{C|X} \, \pi_{lt}^{D|X}$$

- Expected cell counts

$$\hat{f}_{ijkl} = \pi_{ijkl}^{ABCD} * n$$

Model Evaluation (cont.)

- Conditional chi-square (G^2) testing:

$$\begin{aligned} G^2(M_2 \mid M_1) &= -2(L_2 - L_1) \\ &= -2(L_2 - L_s) - [-2(L_2 - L_s)] \\ &= G^2(M_2) - G^2(M_1) \end{aligned}$$

- L_2 must be "nested" within L_1 to use conditional chi-square testing

Model Diagnostics

- Standardized residuals

$$e_{ij} = \frac{f_{ij} - \hat{f}_{ij}}{\sqrt{\hat{f}_{ij}}}$$

- Diagnostics for GLMs

$$D = \frac{L_0 - L_M}{L_0}$$

$$D^* = \frac{G^2(M_0) - G^2(M)}{G^2(M_0)}$$

Evaluation Criteria

- Pearson goodness of fit Chi-square statistic
- Likelihood Ratio goodness of fit Chi-square statistic
  - Also, conditional L^2 statistic
- Information Criteria
  - Akaike Information Criteria (AIC)

$$AIC = G^2 - 2df$$

  - Bayesian Information Criteria (BIC)

$$BIC = G^2 - df * [\ln(n)]$$

Enter “Goodness of Fit”

- We need to evaluate how well the estimated expected cell counts from our model compare to the cell counts we’ve actually observed
- Goodness of fit statistics allow us to evaluate the nearness of our model *estimates* to the actually *observed* frequencies

Goodness of Fit Chi-Square Statistics

- Pearson chi-square

$$X_m^2 = \sum_i \sum_j \sum_k \frac{(f_{ijk} - \hat{f}_{ijk})^2}{\hat{f}_{ijk}}$$

- Likelihood ratio chi-square

$$G_m^2 = 2\sum_i \sum_j \sum_k f_{ijk} \log\left(f_{ijk} / \hat{f}_{ijk}\right)$$

Conditional Chi-Square Testing

- Assume we have a model M_1 that fits the observed data

$$G^2(M_2 \mid M_1) = -2(L_2 - L_1)$$
$$= -2(L_2 - L_s) - [-2(L_2 - L_s)]$$
$$= G^2(M_2) - G^2(M_1)$$

- This is distributed as chi-square with degrees of freedom

$$df\,(M_2 \mid M_1) = df\,(M_2) - df\,(M_1)$$

Conditional Chi-Square Test

- Conditional chi-square for $M_2|M_1$: 17.8104 – 1.0247 = 16.7857
- Degrees of freedom for $M_2|M_1$: 9 - 2 = 7
- Probability: p = .0188
- Decision: reject M_2

- Conditional chi-square for $M_3|M_1$: 7.9333 – 1.0247 = 6.9086
- Degrees of freedom for $M_3|M_1$: 7 - 2 = 5
- Probability: p = .2275
- Decision: accept M_3

Iterative ML Estimation

- Iterative Proportional Fitting (IPF)
  - Start with initial estimates
  - Apply scaling factor
  - Successively adjust expected values until convergence
- Newton-Raphson
  - Start with initial estimates
  - Approximate function in neighborhood of guess by a second-degree function
  - Successive approximations of second-degree functions until convergence

Motivating Example

Table 1: Probability of Universalistic Response (-) and Relative Frequency for the Two-Class Model

Observed Variables	Respondent Type: Universalistic	Particularistic
Auto Passenger	.993	.714
Drama Critic	.926	.354
Insurance Doctor	.940	.330
Board of Directors	.769	.132
Relative Class Frequency	.2792	.7203

Model Evaluation

Pearson Chi-Square

$$X^2 = \sum_{ijkl} \frac{(f_{ijkl} - \hat{f}_{ijkl})^2}{\hat{f}_{ijkl}}$$

Likelihood Ratio Chi-Square

$$G^2 = 2\sum_{ijkl} f_{ijkl} \ln\left(f_{ijkl} / \hat{f}_{ijkl} \right)$$

Where

$$\hat{f}_{ijkl} = n * \pi_{ijkl}^{ABCD}$$

Information criteria

Akaike Information Criteria (AIC)

$$AIC = G^2 - 2df$$

Bayesian Information Criteria (BIC)

$$BIC = G^2 - df * [\ln(N)]$$

Prefer lowest negative value for AIC and BIC.

Evaluation Criteria

```
X-squared            = 2.7200 (0.8431)
L-squared            = 2.7199 (0.8431)
Cressie-Read         = 2.7174 (0.8434)
Number of parameters = 9 (+1)
Sample size          = 216.0
BIC(L-squared)       = -29.5317
AIC(L-squared)       = -9.2801
```


Restricted Latent Class Models

Hypothesis Testing

- **-Equality restrictions**
- **-Deterministic restrictions**

- **-Conditional probabilities**
- **-Latent class probabilities**

Equality Restrictions on Conditional Probabilities

The *parallel indicators* hypothesis (e.g.)

$$\pi_{11}^{B|X} = \pi_{11}^{C|X} \quad and \quad \pi_{12}^{B|X} = \pi_{12}^{C|X}$$

Motivating Example

Table 1: Probability of Universalistic Response (-) and Relative Frequency for the Two-Class Model

Observed Variables	Respondent Type: Universalistic	Particularistic
Auto Passenger	.993	.714
Drama Critic	.926	.354
Insurance Doctor	.940	.330
Board of Directors	.769	.132
Relative Class Frequency	.2792	.7203

Table 2: Latent Class Model Evaluation Criteria for Universalism Data

Model	X^2	G^2	AIC	BIC	DF
H_1: 2-Class Latent Class Model	2.72	2.72	-9.28	-29.53	6
H_2: H_1 + B & C Parallel Indicators	2.84	2.89	-13.11	-40.12	8
H_3: H_2 + D Equal Error Rate	3.60	3.65	-14.35	-44.73	9
H_4: H_3 + A as Perfect Indicator for Class 2	3.61	3.66	-16.34	-50.09	10

Conditional Likelihood Ratio Chi-Square tests

Equal Error Rates Hypothesis

$$\pi_{21}^{D|X} = \pi_{12}^{D|X}$$

Deterministic Restrictions on Conditional Probabilities

Perfect Indicator Hypothesis

$$\pi_{11}^{A|X} = 1.0$$

Table 3: Probability of Universalistic Response (-) and
LC Probabilities for the Restricted Two-Class Model

Observed Variables	Respondent Type: Universalistic	Particularistic
Auto Passenger Friend	1.00*	.725
Drama Critic Friend	.954*	.364*
Insurance Doctor Friend	.954*	.364*
Board of Directors Friend	.852*	.148*
Latent Class Probabilities	.2426	.7574

LEM

```
 lat 1
man 4
dim 2 2 2 2 2
lab X A B C D
mod X
    A|X
    B|X
    C|X eq1 B|X
    D|X eq2
des [1 0 0 1]
sta A|X [0 1 .5 .5]
dat [20 2 6 1 9 2 4 1 38 7 25 6 24 6 23 42]
```

Conditional Probability Restrictions

eq1 restricts current conditional probability estimate to equal that of some previously specified estimate

eq2 requires a design matrix and permits within class and across class equality restrictions

sta start values allows analyst to take control of start values. In current example, providing start values at the boundary fixes the values as "perfect indicator" hypothesis requires.

Scale Analysis

Scaling Items with LCM's

- Assumes underlying latent continuum
  - Ability (e.g., mathematics, verbal)
  - Attitudes (e.g., racism, ethnocentrism)
- Dichotomous items of varying difficulty
- LCA can be restricted to introduce probabilistic model of item difficulty and person ability

Item Difficulty

3+4 ↓ | 4*7 ↓ | 2^2+3^2 ↓ | (a+1)(a-3) ↓

Low ———————————————— **High**

Mathematics Scale

Person's Ability

Low **High**

↑ A. McCutcheon ↑ A. Einstein

Mathematics Scale

Guttman Scaling Model

	A	B	C	D
0	-	-	-	-
1	+	-	-	-
2	+	+	-	-
3	+	+	+	-
4	+	+	+	+

Guttman

- Assume K dichotomous items
  - Right or wrong
  - Agree vs. disagree
- Permits 2^K possible response patterns
- Guttman model permits K+1 response patterns

Critics of Guttman Scales

- Deterministic measurement model
  - No possibility of measurement error
- Model evaluation criteria are “rules of thumb” (ad hoc evaluation criteria)
  - Coefficient of reproduceability

Guttman Scaling (Probability) Model

	A	B	C	D
0	0.0	0.0	0.0	0.0
1	1.0	0.0	0.0	0.0
2	1.0	1.0	0.0	0.0
3	1.0	1.0	1.0	0.0
4	1.0	1.0	1.0	1.0

Restricted LC Scaling Models

- Probabilistic measurement model
  - Explicitly models measurement error
- Probabilistic evaluation criteria
  - All evaluation criteria of usual latent class models are available
- Multiple models
  - Proctor model
  - Item-specific error model
  - Type-specific error model
  - Latent distance error model

Proctor Model

- A single measurement error term for all measures (conditional probabilities)
- Let the measurement error "a"
  - Replace the 0.0 response probabilities of the Guttman model with a
  - And, let 1-a replace the 1.0 response probabilities of the Guttman model
- Requires estimation of single conditional probability, and K latent class probabilities

Proctor Model

	A	B	C	D
0	a	a	a	a
1	1-a	a	a	a
2	1-a	1-a	a	a
3	1-a	1-a	1-a	a
4	1-a	1-a	1-a	1-a

Proctor Model (example)

	A	B	C	D
0	.08	.08	.08	.08
1	.92	.08	.08	.08
2	.92	.92	.08	.08
3	.92	.92	.92	.08
4	.92	.92	.92	.92

Item-specific Error Rate Model

- One measurement error term for each of the K measures (conditional probabilities)
- Let the measurement errors "a," "b," etc.
  - Replace the 0.0 response probabilities of the Guttman model with a (or b, or c, etc.)
  - And, let 1-a (1-b, or 1-c, etc.) replace the 1.0 response probabilities of the Guttman model
- Requires estimation of 2K conditional probability and latent class probabilities

Item-specific Error Rate Model

	A	B	C	D
0	a	b	c	d
1	1-a	b	c	d
2	1-a	1-b	c	d
3	1-a	1-b	1-c	d
4	1-a	1-b	1-c	1-d

Item-specific Error Rate Model (example)

	A	B	C	D
0	.08	.03	.12	.07
1	.92	.03	.12	.07
2	.92	.97	.12	.07
3	.92	.97	.88	.07
4	.92	.97	.88	.93

Type-specific Error Rate Model

- One measurement error (conditional probabilities) term for each type
- Let the measurement errors "a," "b," etc.
  - Replace the 0.0 response probabilities for each type in the Guttman model with a (or b, or c, etc.)
  - And, let 1-a (1-b, or 1-c, etc.) replace the 1.0 response probabilities for each type in the Guttman model
- Requires estimation of K+1 conditional probabilities and K latent class probabilities

Type-specific Error Rate Model

	A	B	C	D
0	a	a	a	a
1	1-b	b	b	b
2	1-c	1-c	c	c
3	1-d	1-d	1-d	d
4	1-e	1-e	1-e	1-e

Type-specific Error Rate Model (example)

	A	B	C	D
0	.08	.08	.08	.08
1	.97	.03	.03	.03
2	.82	.82	.18	.18
3	.91	.91	.91	.09
4	.99	.99	.99	.99

Latent Distance Error Rate Model

- One measurement error (conditional probabilities) term for each of the “anchor” items, 2 error terms for the other items
- Let the measurement errors “a,” “b,” etc.
  - Replace the 0.0 response probabilities in the Guttman model with a (or b, or c, etc.)
  - And, let 1-a (1-b, or 1-c, etc.) replace the 1.0 response probabilities of the Guttman model
- Requires estimation of 2K-2 conditional probabilities and K latent class probabilities

Latent Distance Error Rate Model

	A	B	C	D
0	a	b	d	f
1	1-a	b	d	f
2	1-a	1-c	d	f
3	1-a	1-c	1-e	f
4	1-a	1-c	1-e	1-f

Latent Distance Error Rate Model (example)

	A	B	C	D
0	.08	.03	.07	.02
1	.92	.03	.07	.02
2	.92	.82	.07	.02
3	.92	.82	.99	.02
4	.92	.82	.99	.98

Example Scale Analysis Set-ups

Data Taken from the Political Action Survey (1972) 3-nation panel

Protest Potential

A. Willing to sign a petition (yes/no)

B. Silling to participate in a lawful demonstration (yes/no)

C. Willing to participate in a "sit-in" (yes/no)

D. Willing to use personal violence (yes/no)

Germany								Netherlands							
94	163	11	460	1	2	1	107	45	120	5	210	1	15	4	288
0	4	0	18	0	2	0	15	0	1	0	7	1	1	0	39

Example Scale Analysis Set-ups

Data Taken from the Political Action Survey (1972) 3-nation panel

Protest Potential

A. Willing to sign a petition (yes/no)

B. Silling to participate in a lawful demonstration (yes/no)

C. Willing to participate in a "sit-in" (yes/no)

D. Willing to use personal violence (yes/no)

Germany								Netherlands							
94	163	11	460	1	2	1	107	45	120	5	210	1	15	4	288
0	4	0	18	0	2	0	15	0	1	0	7	1	1	0	39

Example Proctor Model Set-ups

```
*Proctor model: U.S. Data
lat 1
man 4
lab X A B C D
dim 5 2 2 2 2
mod X
    A|X eq2
    B|X eq2
    C|X eq2
    D|X eq2
des [0 1 1 0 1 0 1 0 1 0
     0 1 0 1 1 0 1 0 1 0
     0 1 0 1 0 1 1 0 1 0
     0 1 0 1 0 1 0 1 1 0]
dat [ 28 178 0 412 1 14 0 172 1 2 0 16 0 0 0 38]
```


Other Example Design Set-ups

```
*Item-specific error rate
des [0 1 1 0 1 0 1 0 1 0
     0 2 0 2 2 0 2 0 2 0
     0 3 0 3 0 3 3 0 3 0
     0 4 0 4 0 4 0 4 4 0]

*Type-specific error rate
des [0 1 2 0 3 0 4 0 5 0
     0 1 0 2 3 0 4 0 5 0
     0 1 0 2 0 3 4 0 5 0
     0 1 0 2 0 3 0 4 5 0]

*Latent distance model
des [0 1 1 0 1 0 1 0 1 0
     0 2 0 2 3 0 3 0 3 0
     0 4 0 4 0 4 5 0 5 0
     0 6 0 6 0 6 0 6 6 0]
```

Simultaneous Latent Class Models

Simultaneous LCMs

- Basic latent class model

$$\pi_{ijklt}^{ABCDX} = \pi_t^X \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X}$$

- Let G_s represent external (i.e., non-indicator) grouping variable
- Simultaneous latent class model

$$\pi_{ijklts}^{ABCDXG} = \pi_{ts}^{X|G} \pi_{its}^{A|XG} \pi_{jts}^{B|XG} \pi_{kts}^{C|XG} \pi_{lts}^{D|XG} \pi_s^G$$

Simultaneous Latent Class Models

Comparing LCMs in two or more populations

$$\pi_{ijkls}^{ABCDG} = \sum_{t=1}^{T} \pi_{ts}^{X|G} \ \pi_{its}^{A|XG} \pi_{jts}^{B|XG} \pi_{kts}^{C|XG} \pi_{lts}^{D|XG} \pi_{s}^{G}$$

The first order of business is to ascertain whether the latent variable characterized by the model is the same for all of the S groups (populations). That is, we test the hypotheses

$$\pi_{its}^{A|XG} = \pi_{it}^{A|X}, \quad \pi_{jts}^{B|XG} = \pi_{jt}^{B|X},$$

$$\pi_{kts}^{C|XG} = \pi_{kt}^{C|X}, \quad and \quad \pi_{lts}^{D|XG} = \pi_{lt}^{D|X}$$

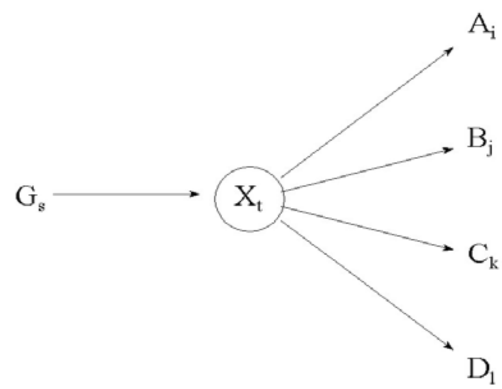
Simultaneous Latent Class Models

In which case, the latent variable is identically characterized by the indicator variables in each of the S groups (or populations) and the SLCM reduces to

$$\pi_{ijkls}^{ABCDG} = \sum_{t=1}^{T} \pi_{ts}^{X|G} \ \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X} \pi_{s}^{G}$$

This is equivalent to estimating a two-variable structural model with one “external” (i.e., non-indicator) variable and a latent variable.

Simultaneous Latent Class Models



Simultaneous Latent Class Models

Represents the model

$$\pi_{ijkls}^{ABCDG} = \sum_{t=1}^{T} \pi_{ts}^{X|G} \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X} \pi_{s}^{G}$$

Where the relationship between the grouping variable (G_s) and each of the indicator variables (e.g., A_i) is completely mediated through the latent variable (X_t).

The greater our ability to impose the restrictions of group independent conditional probabilities, the greater is our confidence that we have measured the same latent phenomenon in each of the groups.

SLCMs: Example 1

The Stouffer-Toby data set also contains a sample of students who were asked about the four scenarios, though the second group of respondents were asked if they had a right to ask their friend to act in a particularistic manner.

Table 1: Responses to Four Role Conflict Scenarios for Ego and Ego's Close Friend (Stouffer and Toby 1951)

Items	Ego Faces Dilemma		Ego's Friend Faces Dilemma	
A B C	Item D (+)	Item D (-)	Item D (+)	Item D (-)
+ + +	20	2	20	3
+ + -	6	1	4	3
+ - +	9	2	23	3
+ - -	4	1	4	2
- + +	38	7	25	6
- + -	25	6	15	6
- - +	24	6	29	5
- - -	23	42	31	37

Table 2: Simultaneous Latent Class Model Evaluation Criteria for Universalism Data

Model	X^2	G^2	AIC	BIC	DF
H_1: Unrestricted 2-Class / group Model	9.06	8.25	-15.75	-64.57	12
H_2: Structural Homogeneity	24.78	23.47	-16.53	-97.90	20
H_3: Complete Homogeneity	24.82	23.48	-18.52	-103.96	21

Table 3: Simultaneous Latent Class Parameters, Complete Homogeneity Two-Class Model: Stouffer-Toby Data

Observed Variables	Respondent Type: Universalistic	Respondent Type: Particularistic
Auto Passenger Friend	.990	.655
Drama Critic Friend	.892	.433
Insurance Doctor Friend	.979	.283
Board of Directors Friend	.681	.151
Latent Class Probabilities	.2917	.7083

LEM

```
* SLCM, Complete Homogeneity, Stouffer-
* Toby Data
lat 1
man 5
dim 2 2 2 2 2 2
lab X G A B C D
mod X|G eq2
    A|X
    B|X
    C|X
    D|X
des [1 1 0 0]
dat [20 2 6 1  9 2 4 1 38 7 25 6 24 6 23 42
    20 3 4 3 23 3 4 2 25 6 15 6 29 5 31 37]
```


Belief in God (Measurement Error): 1991 ISSP

Q.14 Please tick one box below to show which statement comes closest to expressing what you believe about God. (Please tick one box only)

----

1. ***I don't believe in God***
2. I don't know whether there is a God and I don't believe there is any way to find out
3. I don't believe in a personal God, but I do believe in a Higher Power of some kind
4. I find myself believing in God some of the time, but not at others
5. While I have doubts, I feel that I do believe in God
6. I know God really exists and I have no doubts about it
8. Can't choose, don't know

Q.15 How close do you feel to God most of the time? (Please tick one box only)

----

1. ***Don't believe in God***
2. Not close at all
3. Not very close
4. Somewhat close
5. Extremely close
8. Can't choose, don't know
9. NA, refused

Belief in God (Measurement Error): 1991 ISSP

Q.16 Which best describes your beliefs about God?
(Please tick one box only)
----

1. I don't believe in God now and I never have
2. I don't believe in God now, but I used to
3. I believe in God now, but I didn't used to
4. I believe in God now and I always have
8. Can't choose, don't know
9. NA, refused

Example 2: Belief in God (Germany, 1991 ISSP Data)

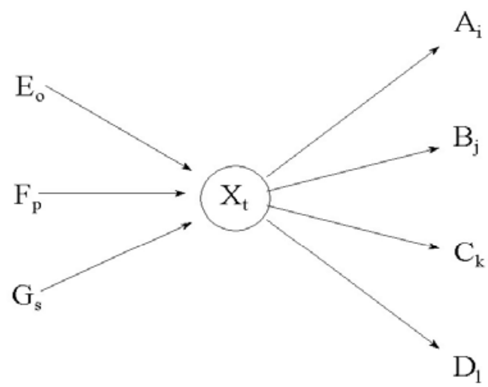
```
* 1991 ISSP Belief in God data: a=v33,
* b=v32, c=v31, L=lander; dichotomized with
* 1=don't, 2= believing
lat 1
man 4
dim 2 2 2 2 2
lab X L A B C
mod X|L
    A|XL eq2
    B|X
    C|XL eq2
des [0 0 0 0 1 0 1 0
     0 2 0 2 0 0 0 0]
sta A|XL [.5 .5 .5 .5 .9 .1 .9 .1]
dat [107 31  8 276 4 1 18 901
     674 82 37 306 5 4  7 371]
```

Table 7: Simultaneous Latent Class Parameters, Restricted Belief in God Data: ISSP 1991, West & East Germany

Observed Variables	West Believe	West Not	East Believe	East Not
Express	.767	.013*	.551	.013*
Close	.998*	.043*	.998*	.043*
Describe	.982*	.218	.982*	.110
LC Probs.	.8909	.1091	.4629	.5371


```
*1991 ISSP Belief in God data: a=v33,
* b=v32, c=v31, L=lander dichotomized
* with 1=don't believe, 2= believing
lat 1
man 4
dim 2 2 2 2 2
lab X L A B C
mod X|L
    A|XL {AX,AL}
    B|X
    C|XL {CX,CL}
dat [107 31  8 276 4 1 18 901
674 82 37 306 5 4  7 371]
```

Log-linear models with latent variables (latent logistic "mimic" model)



Probabilistic parameterization ("saturated model")

$$\pi_{ijklops}^{ABCDEFG} = \sum_{t=1}^{T} \pi_{tops}^{X|EFG} \pi_{itops}^{A|XEFG} \pi_{jtops}^{B|XEFG} \pi_{ktops}^{C|XEFG} \pi_{ltops}^{D|XEFG} \pi_{ops}^{EFG}$$

"Reduced" Model (above)

$$\pi_{ijklops}^{ABCDEFG} = \sum_{t=1}^{T} \pi_{to}^{X|E} \pi_{tp}^{X|F} \pi_{ts}^{X|G} \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X} \pi_{o}^{E} \pi_{p}^{F} \pi_{s}^{G}$$

Log-linear Parameterization

$$\ln(\hat{F}_{ijkltops}^{ABCDXEFG}) = \lambda + \lambda_o^E + \lambda_p^F + \lambda_s^G + \lambda_t^X + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{it}^{AX} + \lambda_{jt}^{BX} + \lambda_{kt}^{CX} + \lambda_{lt}^{DX} + \lambda_{to}^{XE} + \lambda_{tp}^{XF} + \lambda_{ts}^{XG} + \lambda_{io}^{AE} + \lambda_{jo}^{BE} + \lambda_{ko}^{CE} + \lambda_{lo}^{DE} + \cdots + \lambda_{ks}^{CG} + \lambda_{ls}^{DG} + \lambda_{ito}^{AXE} + \cdots + \lambda_{lts}^{DXG} + \lambda_{itop}^{AXEF} + \cdots + \lambda_{ltps}^{DXFG} + \lambda_{itops}^{AXEFG} + \cdots + \lambda_{ltops}^{DXEFG} + \lambda_{top}^{XEF} + \cdots + \lambda_{tops}^{XEFG} + \lambda_{op}^{EF} + \cdots + \lambda_{ops}^{EFG}$$

Reduced

$$\ln(\hat{F}_{ijkltops}^{ABCDXEFG}) = \lambda + \lambda_o^E + \lambda_p^F + \lambda_s^G + \lambda_t^X + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{it}^{AX} + \lambda_{jt}^{BX} + \lambda_{kt}^{CX} + \lambda_{lt}^{DX} + \lambda_{to}^{XE} + \lambda_{tp}^{XF} + \lambda_{ts}^{XG} + \lambda_{is}^{AG} + \lambda_{js}^{BG} + \lambda_{ks}^{CG} + \lambda_{ls}^{DG}$$

Logit

$$\ln\left(\hat{F}_{ijkl1ops}^{ABCDXEFG} \Big/ \hat{F}_{ijkl2ops}^{ABCDXEFG} \right) = \ln(\hat{F}_{ijkl1ops}^{ABCDXEFG}) - \ln(\hat{F}_{ijkl2ops}^{ABCDXEFG})$$

For our model, this yields (recalling the identifying restrictions)

$$\ln\left(\hat{F}^{ABCDXEFG}_{ijkl1ops} \Big/ \hat{F}^{ABCDXEFG}_{ijkl2ops} \right) = 2\lambda^{X}_{1} + 2\lambda^{AX}_{11} + 2\lambda^{BX}_{11} + 2\lambda^{CX}_{1} +$$

$$2\lambda^{CX}_{11} + 2\lambda^{DX}_{11} + 2\lambda^{EX}_{11} + 2\lambda^{FX}_{11} + 2\lambda^{GX}_{11}$$


```
* Belief in God Latent logistic example w/
* 3 indpendent variables; order f,e,d,c,b,a
* a=v31d, b=v32d, c=v33d, d=m/f, e=city,
* f=w/e; city: 1=50k+, 2=50k - 50k, 3=lt 5k
lat 1
man 6
dim 2 2 3 2 2 2 2
lab X F E D C B A
mod X|DEF {XD,XE,XF}
    A|X
    B|X
    C|XF {XC,FC}
dat [51  11   3 128   2   1   3 249  28   6
      2  96   1   0   8 313
     13   3   0  25   1   0   2 105   6   2
      3  23   0   0   2 106
      5   0   0  12   0   0   2  54   4   0
      0  18   0   0   1  74
    138  14   3  26   2   0   0  39 129  15
      4  50   0   1   2  64
    117  14   9  24   0   0   2  56  99  14
     11  40   1   1   2  85
    109  11   5  46   1   1   0  58  82  14
      5  61   1   1   1 111 ]
```

Simultaneous Latent Class Models: Model Selection

Simultaneous Latent Class Models

Comparing LCMs in two or more populations

$$\pi_{ijkls}^{ABCDG} = \sum_{t=1}^{T} \pi_{ts}^{X|G} \pi_{its}^{A|XG} \pi_{jts}^{B|XG} \pi_{kts}^{C|XG} \pi_{lts}^{D|XG} \pi_{s}^{G}$$

The first order of business is to ascertain whether the latent variable characterized by the model is the same for all of the S groups (populations). That is, we test the hypotheses

$$\pi_{its}^{A|XG} = \pi_{it}^{A|X}, \quad \pi_{jts}^{B|XG} = \pi_{jt}^{B|X},$$
$$\pi_{kts}^{C|XG} = \pi_{kt}^{C|X}, \quad and \quad \pi_{lts}^{D|XG} = \pi_{lt}^{D|X}$$

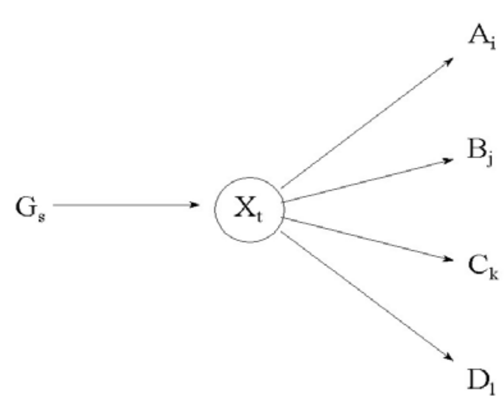
Simultaneous Latent Class Models

In which case, the latent variable is identically characterized by the indicator variables in each of the S groups (or populations) and the SLCM reduces to

$$\pi_{ijkls}^{ABCDG} = \sum_{t=1}^{T} \pi_{ts}^{X|G} \, \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X} \pi_{s}^{G}$$

This is equivalent to estimating a two-variable structural model with one “external” (i.e., non-indicator) variable and a latent variable.

Simultaneous Latent Class Models



Simultaneous Latent Class Models

Represents the model

$$\pi_{ijkls}^{ABCDG} = \sum_{t=1}^{T} \pi_{ts}^{X|G} \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X} \pi_{s}^{G}$$

Where the relationship between the grouping variable (G_s) and each of the indicator variables (e.g., A_i) is completely mediated through the latent variable (X_t).

The greater our ability to impose the restrictions of group independent conditional probabilities, the greater is our confidence that we have measured the same latent phenomenon in each of the groups.

SLCMs: Example 1

<u>1992-1994-1996 Czech Republic</u> (Economic Expectations data)

Do you generally prefer an economy:

1 **as socialist, which was in our country before 1989**

2 **as a social market with a high degree of state intervention**

3 **as a free market with minimal state intervention?"**

According to you, should the state administratively fix prices more?

Should the state provide a job for everyone who wants to work?

Do you think that the state should provide housing for every family which is not able to find it?

SLCMs: Example 1a

```
* Unrestricted latent class model
* Czech data: 1992-1994-1996 data
* A=Housing B=Prices C=Preferred form of economy D=Jobs
lat 1
man 5
dim 2 3 2 2 3 2
lab X T A B C D
mod T
    X|T
    A|XT
    B|XT
    C|XT
    D|XT
dat czecon92.dat
```


SLCMs: Example 1

czecon92.dat

```
 27   2  246 75  78 103   0   0  26  23  15  92
  4   2  48  22  32  50   1   0  21  37  24 164
 74   4 412  94  65 132   3   1  30  38  11  56
  4   0  76  39  18  41   0   0  34  30  14 114
159  10 437 116  62  96   3   1  37  23  11  38
  8   0  88  38  23  40   3   0  34  49  15 115
```

SLCMs: Example 1a (cont.)

```
*** STATISTICS ***
  Number of iterations = 72
  Converge criterion   = 0.0000008801
  Seed random values   = 3602
  X-squared            = 310.4235 (0.0000)
  L-squared            = 298.0800 (0.0000)
  Cressie-Read         = 302.8963 (0.0000)
  Dissimilarity index  = 0.0938
  Degrees of freedom   = 36
  Log-likelihood       = -13586.64568
  Number of parameters = 35 (+1)
  Sample size          = 3788.0
  BIC(L-squared)       = 1.4547
  AIC(L-squared)       = 226.0800
```


SLCMs: Example 1b

```
* Unrestricted latent class model
* Czech data: 1992-1994-1996 data
* A=Housing B=Prices C=Preferred form of economy D=Jobs
lat 1
man 5
dim 3 3 2 2 3 2
lab X T A B C D
mod T
    X|T
    A|XT
    B|XT
    C|XT
    D|XT
dat czecon92.dat
```

SLCMs: Example 1b (cont.)

```
*** STATISTICS ***
  Number of iterations = 2553
  Converge criterion   = 0.0000009983
  Seed random values   = 6253
  X-squared            = 102.3890 (0.0000)
  L-squared            = 104.0533 (0.0000)
  Cressie-Read         = 101.4031 (0.0000)
  Dissimilarity index  = 0.0398
  Degrees of freedom   = 18
  Log-likelihood       = -13489.63229
  Number of parameters = 53 (+1)
  Sample size          = 3788.0
  BIC(L-squared)       = -44.2594
  AIC(L-squared)       = 68.0533
```


SLCMs: Example 1c

```
* Unrestricted latent class model
* Czech data: 1992-1994-1996 data
* A=Housing B=Prices C=Preferred form of economy D=Jobs
lat 1
man 5
dim 4 3 2 2 3 2
lab X T A B C D
mod T
    X|T
    A|XT
    B|XT
    C|XT
    D|XT
dat czecon92.dat
```

SLCMs: Example 1c (cont.)

```
*** STATISTICS ***
  Number of iterations = 5000
  Converge criterion   = 0.0000027286
  Seed random values   = 5482
  X-squared            = 9.6596 (0.0000)
  L-squared            = 9.4632 (0.0000)
  Cressie-Read         = 9.0931 (0.0000)
  Dissimilarity index  = 0.0048
  Degrees of freedom   = 0
  Log-likelihood       = -13442.33725
  Number of parameters = 71 (+1)
  Sample size          = 3788.0
  BIC(L-squared)       = 9.4632
  AIC(L-squared)       = 9.4632
```


SLCMs: Example 1c (cont.)

```
Eigenvalues information matrix
   4573.9225  3982.8064  3552.2565  3189.1838  2715.6364  2461.4561
   2159.3769  2035.7557  1815.2758  1575.5695  1417.4125  1407.8151
   1285.0023  1183.5583  1181.6784  1154.7875  1122.9035  1020.1012
    989.4737   786.0133   770.8513   703.2509   701.9775   542.4750
    475.2142   441.5663   340.0557   257.4423   230.2436   226.0781
    216.4504   195.2566   172.4735   155.2111   117.3252   114.5920
     94.6768    91.2602    80.1883    69.8499    58.1719    41.6721
     40.3432    37.4084    31.7911    30.4830    25.2991    17.5388
     14.8002     8.3676     5.4431     4.4043     2.1683     2.0178
      1.3007     0.4437     0.0437     0.0001     0.0001     0.0000
     -0.0000    -0.0000    -0.0000    -0.0001    -0.0001    -0.0001
     -0.0002    -0.0003    -0.0003    -0.0004    -0.0022

WARNING: 14 (nearly) boundary or non-identified (log-linear) parameters
```

SLCMs: Example 1(cont.)

```
* Restricted latent class model
* Czech data: 1992-1994-1996 data
* A=Housing B=Prices C=Preferred form of economy D=Jobs
*
lat 1
man 5
dim 4 3 2 2 3 2
lab X T A B C D
mod T
    X|T
    A|XT
    B|XT
    C|XT
    D|XT
sta A|XT [0 1 0 1 0 1 .5 .5 .5 .5 .5 .5
          .6 .4 .6 .4 .6 .4 .4 .6 .4 .6 .4 .6 ]
sta B|XT [0 1 0 1 0 1 .5 .5 .5 .5 .5 .5
          .6 .4 .6 .4 .6 .4 .4 .6 .4 .6 .4 .6 ]
sta C|XT [0 .5 .5 0 .5 .5 0 .5 .5 .33 .33 .34 .33 .33 .34 .33 .33
          .34 0 .5 .5 0 .5 .5 0 .5 .5 0 .5 .5 0 .5 .5 0 .5 .5 ]
sta D|XT [0 1 0 1 0 1 .5 .5 .5 .5 .5 .5
          .6 .4 .6 .4 .6 .4 .4 .6 .4 .6 .4 .6 ]
dat czecon92.dat
```



```
*** STATISTICS ***
  Number of iterations = 2886
  Converge criterion   = 0.0000009992
  Seed random values   = 6220
  X-squared            = 38.9340 (0.0000)
  L-squared            = 28.9340 (0.0000)
  Cressie-Read         = 33.3781 (0.0000)
  Dissimilarity index  = 0.0133
  Degrees of freedom   = 0
  Log-likelihood       = -13452.07266
  Number of parameters = 71 (+1)
  Sample size          = 3788.0
  BIC(L-squared)       = 28.9340
  AIC(L-squared)       = 28.9340
```

SLCMs: Example 1(cont.)

```
 * Restricted latent class model
 * Czech data: 1992-1994-1996 data
 * A=Housing B=Prices C=Preferred form of economy D=Jobs
 *
 lat 1
 man 5
 dim 4 3 2 2 3 2
 lab X T A B C D
 mod T
     X|T
     A|XT eq2
     B|XT eq2
     C|XT eq2
     D|XT eq2
des [0 -1 0 -1  0 -1   0 0  0 0  0 0   0 0  0 0  0 0   0 0  0 0  0 0
     0 -1 0 -1  0 -1   0 0  0 0  0 0   0 0  0 0  0 0   0 0  0 0  0 0
    -1 0 0  -1 0 0  -1 0 0   0 0 0  0 0 0  0 0 0
    -1 0 0  -1 0 0  -1  0 0 -1 0 0 -1 0 0 -1 0 0
     0 -1 0 -1 0 -1 0  0 0  0 0  0 0   0 0  0 0  0 0   0 0  0 0  0 0]
sta A|XT [0 1 0 1 0 1 .5 .5 .5 .5 .5 .5
         .6 .4 .6 .4 .6 .4 .4 .6 .4 .6 .4 .6 ]
sta B|XT [0 1 0 1 0 1 .5 .5 .5 .5 .5 .5
         .6 .4 .6 .4 .6 .4 .4 .6 .4 .6 .4 .6 ]
sta C|XT [0 .5 .5 0 .5 .5 0 .5 .5 .33 .33 .34 .33 .33 .34 .33 .33
         .34 0 .5 .5 0 .5 .5 0 .5 .5 0 .5 .5 0 .5 .5 0 .5 .5 ]
sta D|XT [0 1 0 1 0 1 .5 .5 .5 .5 .5 .5
         .6 .4 .6 .4 .6 .4 .4 .6 .4 .6 .4 .6 ]
dat czecon92.dat
```



```
*** STATISTICS ***
  Number of iterations = 2929
  Converge criterion   = 0.0000009999
  Seed random values   = 3199
  X-squared            = 38.9330 (0.0029)
  L-squared            = 28.9340 (0.0492)
  Cressie-Read         = 33.3776 (0.0150)
  Dissimilarity index  = 0.0133
  Degrees of freedom   = 18
  Log-likelihood       = -13452.07267
  Number of parameters = 53 (+1)
  Sample size          = 3788.0
  BIC(L-squared)       = -119.3787
  AIC(L-squared)       = -7.0660

WARNING: no information is provided on identification of parameters
```

SLCMs: Example 1(cont.)

```
 * Restricted latent class model
 * Czech data: 1992-1994-1996 data
 * A=Housing B=Prices C=Preferred form of economy D=Jobs
 *
 lat 1
 man 5
 dim 4 3 2 2 3 2
 lab X T A B C D
 mod T
     X|T
     A|XT eq2
     B|XT eq2
     C|XT eq2
     D|XT eq2
des [0 -1 0 -1  0 -1 0   1 0  1 0  1  0   2 0 2 0 2 0   3 0  3 0  3 0
     0 -1 0 -1  0 -1 0   4 0  4 0  4  0   5 0 5 0 5 0   6 0  6 0  6 0
    -1 0 0 -1 0 0 -1 0 0   0 0 0  0 0 0  0 0 0
    -1 0 0 -1 0 0 -1 0 0  -1 0 0 -1 0 0 -1 0 0
     0 -1 0 -1  0 -1 0   7 0  7 0  7  0   8 0 8 0 8 0   9 0  9 0  9 0]
sta A|XT [0 1 0 1 0 1 .5 .5 .5 .5 .5 .5
          .6 .4 .6 .4 .6 .4 .4 .6 .4 .6 .4 .6 ]
sta B|XT [0 1 0 1 0 1 .5 .5 .5 .5 .5 .5
          .6 .4 .6 .4 .6 .4 .4 .6 .4 .6 .4 .6 ]
sta C|XT [0 .5 .5 0 .5 .5 0 .5 .5 .33 .33 .34 .33 .33 .34 .33 .33
          .34 0 .5 .5 0 .5 .5 0 .5 .5 0 .5 .5 0 .5 .5 0 .5 .5 ]
sta D|XT [0 1 0 1 0 1 .5 .5 .5 .5 .5 .5
          .6 .4 .6 .4 .6 .4 .4 .6 .4 .6 .4 .6 ]
dat czecon92.dat
```



```
*** STATISTICS ***

  Number of iterations = 299
  Converge criterion   = 0.0000009841
  Seed random values   = 132

  X-squared            = 79.3459 (0.0000)
  L-squared            = 62.2769 (0.0042)
  Cressie-Read         = 70.1651 (0.0006)
  Dissimilarity index  = 0.0312
  Degrees of freedom   = 36
  Log-likelihood       = -13468.74411
  Number of parameters = 35 (+1)
  Sample size          = 3788.0
  BIC(L-squared)       = -234.3485
  AIC(L-squared)       = -9.7231
```

Table 1: Conditional Response Probabilities for Modeled Trend Data: Czech Preferred Form of Economy

Indicator		Class 1	2	3	4
1992					
Economy					
	Socialist	.00*	.09	.00*	.00*
	Social Market	.14	.73	.50	.26
	Free Market	.86	.18	.50	.74
Housing		.00*	.92*	.33*	.25*
Prices		.00*	.97*	.57*	.66*
Jobs		.00*	.91*	.74*	.05*
1994					
Economy					
	Socialist	.00*	.14	.00*	.00*
	Social Market	.09	.76	.75	.33
	Free Market	.91	.10	.25	.67
Housing		.00*	.92*	.33*	.25*
Prices		.00*	.97*	.57*	.66*
Jobs		.00*	.91*	.74*	.05*
1996					
Economy					
	Socialist	.00*	.24	.00*	.00*
	Social Market	.22	.68	.72	.38
	Free Market	.78	.08	.28	.62
Housing		.00*	.92*	.33*	.25*
Prices		.00*	.97*	.57*	.66*
Jobs		.00*	.91*	.74*	.05*

Table 1: Latent Class Probabilities for Modeled Trend Data: Czech Preferred Form of Economy

Year	Class 1	2	3	4
1992	.145	.362	.178	.316
1994	.076	.485	.154	.285
1996	.087	.543	.162	.208

Belief in God (Measurement Error): 1991 ISSP

Q.14 Please tick one box below to show which statement comes closest to expressing what you believe about God. (Please tick one box only)
----

1. I don't believe in God
2. I don't know whether there is a God and I don't believe there is any way to find out
3. I don't believe in a personal God, but I do believe in a Higher Power of some kind
4. I find myself believing in God some of the time, but not at others
5. While I have doubts, I feel that I do believe in God
6. I know God really exists and I have no doubts about it
8. Can't choose, don't know

Q.15 How close do you feel to God most of the time? (Please tick one box only)
----

1. Don't believe in God
2. Not close at all
3. Not very close
4. Somewhat close
5. Extremely close
8. Can't choose, don't know
9. NA, refused

Belief in God (Measurement Error): 1991 ISSP

Q.16 Which best describes your beliefs about God? (Please tick one box only)
----

1. I don't believe in God now and I never have
2. I don't believe in God now, but I used to
3. I believe in God now, but I didn't used to
4. I believe in God now and I always have
8. Can't choose, don't know
9. NA, refused

Example 2: Belief in God (Germany, 1991 ISSP Data)

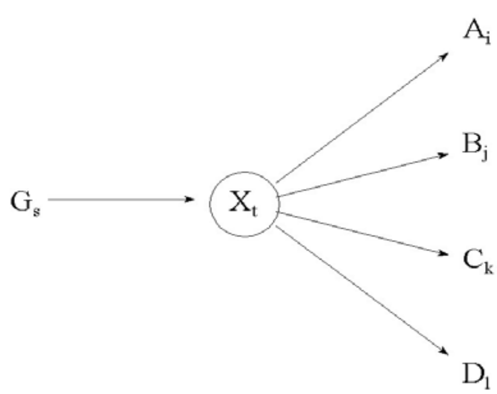
```
* 1991 ISSP Belief in God data: a=v33,
* b=v32, c=v31, L=lander; dichotomized with * 1=don't, 2= believing
lat 1
man 4
dim 2 2 2 2 2
lab X L A B C
mod X|L
    A|XL eq2
    B|X
    C|XL eq2
des [0 0 0 0 1 0 1 0
     0 2 0 2 0 0 0 0]
sta A|XL [.5 .5 .5 .5 .9 .1 .9 .1]
dat [107 31  8 276 4 1 18 901
     674 82 37 306 5 4  7 371]
```


Table 7: Simultaneous Latent Class Parameters, Restricted Belief in God Data: ISSP 1991, West & East Germany

Observed Variables	West Believe	West Not	East Believe	East Not
Express	.767	.013*	.551	.013*
Close	.998*	.043*	.998*	.043*
Describe	.982*	.218	.982*	.110
LC Probs.	.8909	.1091	.4629	.5371

Loglinear Structural Equation Modeling

Simultaneous Latent Class Models



Simultaneous Latent Class Models

Represents the model

$$\pi_{ijkls}^{ABCDG} = \sum_{t=1}^{T} \pi_{ts}^{X|G} \, \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X} \pi_{s}^{G}$$

Where the relationship between the grouping variable (G_s) and each of the indicator variables (e.g., A_i) is completely mediated through the latent variable (X_t).

The greater our ability to impose the restrictions of group independent conditional probabilities, the greater is our confidence that we have measured the same latent phenomenon in each of the groups.

Belief in God (Measurement Error): 1991 ISSP

Q.14 Please tick one box below to show which statement comes closest to expressing what you believe about God. (Please tick one box only)

----

1. ***I don't believe in God***
2. I don't know whether there is a God and I don't believe there is any way to find out
3. I don't believe in a personal God, but I do believe in a Higher Power of some kind
4. I find myself believing in God some of the time, but not at others
5. While I have doubts, I feel that I do believe in God
6. I know God really exists and I have no doubts about it
8. Can't choose, don't know

Q.15 How close do you feel to God most of the time? (Please tick one box only)

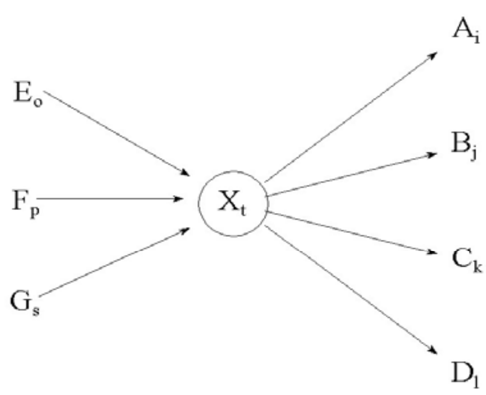
----

1. ***Don't believe in God***
2. Not close at all
3. Not very close
4. Somewhat close
5. Extremely close
8. Can't choose, don't know
9. NA, refused

Belief in God (Measurement Error): 1991 ISSP

Q.16 Which best describes your beliefs about God?
(Please tick one box only)
----
1. I don't believe in God now and I never have
2. I don't believe in God now, but I used to
3. I believe in God now, but I didn't used to
4. I believe in God now and I always have
8. Can't choose, don't know
9. NA, refused

Log-linear models with latent variables (latent logistic "mimic" model)



Odds and Odds Ratios

- **Odds** that an occurrence is of type i instead of type not i (or of type j instead of type not j)
  - e.g.,

$$\Omega_{1+} = p_{1+} / p_{2+} \quad or \quad \Omega_{1|1} = p_{11} / p_{12} = p_{1|1} / p_{2|1}$$

- **Odds ratios** compare odds
  - ratios of cross products
  - e.g.,

$$\theta_{11} = \frac{p_{11} / p_{12}}{p_{21} / p_{22}} = \frac{p_{11} p_{22}}{p_{12} p_{21}} = \frac{f_{11} f_{22}}{f_{12} f_{21}}$$

Odds and Odds Ratios (cont.)

- 2x2 table (Hagenaars and Heinen 1980)

Table 1: Age (A) and Religious Membership (B)

	B.Religious Membership		
A. Age	1. Member	2.Nonmember	Total
1. Young	143	193	336
2. Old	252	162	414
Total	395	355	750

Odds and Odds Ratios (cont.)

- Marginal odds: e.g., Ω_{1+} = 336/414 = .812
  - odds of subject being young relative to being old in this table
- Conditional odds: e.g. $\Omega_{1|1}$ = 143/193 = .741
  - Young people are (1/.741=) 1.35 times more likely to be nonmember than member

Odds and Odds Ratios (cont.)

- Odds ratio: θ_{11} = (143/193)/(252/162) = (143x162)/(252x193) = .476 (approx. .5)
  - the odds of a younger person being a member of a denomination is only half those of an older person
  - among members, the odds of being an older person is about twice that of being a younger person
- Odds ratios are independent of the marginal distributions of variables

Local Odds Ratios (e.g., θ_{11})

B

A	1	2	3	4	
1	f_{11}	f_{12}	f_{13}	f_{14}	f_{1+}
2	f_{21}	f_{22}	f_{23}	f_{24}	f_{2+}
3	f_{31}	f_{32}	f_{33}	f_{34}	f_{3+}
4	f_{41}	f_{42}	f_{43}	f_{44}	f_{4+}
5	f_{51}	f_{52}	f_{53}	f_{54}	f_{5+}
	f_{+1}	f_{+2}	f_{+3}	f_{+4}	f_{++}

Odds Ratios

B

A	1	2	3	4	
1	f_{11}	f_{12}	f_{13}	f_{14}	f_{1+}
2	f_{21}	f_{22}	f_{23}	f_{24}	f_{2+}
3	f_{31}	f_{32}	f_{33}	f_{34}	f_{3+}
4	f_{41}	f_{42}	f_{43}	f_{44}	f_{4+}
5	f_{51}	f_{52}	f_{53}	f_{54}	f_{5+}
	f_{+1}	f_{+2}	f_{+3}	f_{+4}	f_{++}

Geometric Averages

- Multiplicative models

$$\overline{X}_{geom} = \left(\prod_{i}^{n} X_i \right)^{1/n} = \sqrt[n]{X_1 X_2 \cdots X_n}$$

- Partial odds ratio

$$\theta_{11p} = \left(\prod_{k}^{K} \theta_{11k} \right)^{1/K} = \sqrt[K]{\theta_{111}\theta_{112}\cdots\theta_{11K}}$$

Multiplicative Model for Multi-way Tables

- A multiplicative model can be used to perfectly describe the cell counts of a multi-way contingency table

$$f_{ijk}^{ABC} = \eta \tau_i^A \tau_j^B \tau_k^C \tau_{ij}^{AB} \tau_{ik}^{AC} \tau_{jk}^{BC} \tau_{ijk}^{ABC}$$

- Where(for IJK):

$$\eta = \left(\prod_i \prod_j \prod_k f_{ijk} \right)^{1/IJK}$$

$$\tau_i^A = \frac{\left(\prod_j \prod_k f_{ijk} \right)^{1/JK}}{\eta}$$

Multiplicative Model for Multi-way Tables (cont.)

- And:

$$\tau_{ij}^{AB} = \frac{\left(\prod_k f_{ijk}\right)^{1/K}}{\eta \tau_i^A \tau_j^B}$$

$$\tau_{ijk}^{ABC} = \frac{f_{ijk}}{\eta \tau_i^A \tau_j^B \tau_k^C \tau_{ij}^{AB} \tau_{ik}^{AC} \tau_{jk}^{BC}}$$

- The other parameters are defined similarly
- This can be expanded to tables of any dimensionality

Multiplicative and Additive Models

- We have seen that a multi-dimensional table of cell counts can be perfectly represented by the multiplicative model:

$$f_{ijk}^{ABC} = \eta \tau_i^A \tau_j^B \tau_k^C \tau_{ij}^{AB} \tau_{ik}^{AC} \tau_{jk}^{BC} \tau_{ijk}^{ABC}$$

- By taking the natural log this equation, this model--known as the loglinear model, becomes a member of the family of general linear models (GLM) with the log link function:

$$\log(f_{ijk}^{ABC}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

Multiplicative and Additive Models (cont.)

- Where we have parallel definitions for the two parameterizations:
  - Let $G_{ijk} = \log(f_{ijk})$
  - Multiplicative — Additive

$$\eta = \left(\prod_i \prod_j \prod_k f_{ijk} \right)^{1/IJK} \qquad \mu = \frac{1}{IJK} \sum_I \sum_J \sum_K G_{ijk}$$

$$\tau_i^A = \frac{\left(\prod_j \prod_k f_{ijk} \right)^{1/JK}}{\eta} \qquad \lambda_i^A = \frac{1}{JK} \sum_J \sum_K G_{ijk} - \mu$$

Multiplicative and Additive Models (cont.)

- Multiplicative — Additive

$$\tau_{ij}^{AB} = \frac{\left(\prod_k f_{ijk} \right)^{1/K}}{\eta \tau_i^A \tau_j^B} \qquad \lambda_{ij}^{AB} = \frac{1}{K} \sum_k G_{ijk} - \mu - \lambda_i^A - \lambda_j^B$$

$$\tau_{ijk}^{ABC} = \frac{f_{ijk}}{\eta \tau_i^A \tau_j^B \tau_k^C \tau_{ij}^{AB} \tau_{ik}^{AC} \tau_{jk}^{BC}} \qquad \lambda_{ijk}^{ABC} = G_{ijk} - \mu - \lambda_i^A - \lambda_j^B - \lambda_k^C - \lambda_{ij}^{AB} - \lambda_{ik}^{AC} - \lambda_{jk}^{BC}$$

Multiplicative and Additive Models (cont.)

- Identifying constraints
  - We can estimate no more than IJK parameters for the IJK table, so this fixes the number of model parameters if our model is to be identified
- Multiplicative model

$$\prod_{i=1}^{I} \tau_i^{A} = 1.0$$

$$\prod_{i=1}^{I} \tau_{ij}^{AB} = \prod_{j=1}^{J} \tau_{ij}^{AB} = \prod_{i=1}^{I}\prod_{j=1}^{J} \tau_{ij}^{AB} = 1.0$$

Multiplicative and Additive Models (cont.)

- Multiplicative model identifying constraints (cont.)

$$\prod_{i=1}^{I} \tau_{ijk}^{ABC} = \prod_{j=1}^{J} \tau_{ijk}^{ABC} = \prod_{k=1}^{K} \tau_{ijk}^{ABC} = \prod_{i=1}^{I}\prod_{j=1}^{J} \tau_{ijk}^{ABC} = \prod_{i=1}^{I}\prod_{k=1}^{K} \tau_{ijk}^{ABC} =$$

$$\prod_{j=1}^{J}\prod_{k=1}^{K} \tau_{ijk}^{ABC} = \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K} \tau_{ijk}^{ABC} = 1.0$$

Multiplicative and Additive Models (cont.)

- This means, for instance:

$$\tau_1^A = \frac{1}{\tau_2^A} \qquad when\ I = 2$$

$$\tau_{11}^{AB} = \frac{1}{\tau_{12}^{AB}} = \frac{1}{\tau_{21}^{AB}} = \tau_{22}^{AB} \qquad when\ I = J = 2$$

$$\tau_{111}^{ABC} = \frac{1}{\tau_{112}^{ABC}} = \tau_{122}^{ABC} = \frac{1}{\tau_{222}^{ABC}} = \cdots \qquad when\ I = J = K = 2$$

Multiplicative and Additive Models (cont.)

- Identifying constraints on additive models:

$$\sum_{i=1}^{I} \lambda_i^A = 0$$

$$\sum_{i=1}^{I} \lambda_{ij}^{AB} = \sum_{j=1}^{J} \lambda_{ij}^{AB} = \sum_{i=1}^{I} \sum_{j=1}^{J} \lambda_{ij}^{AB} = 0$$

$$\sum_{i=1}^{I} \lambda_{ijk}^{ABC} = \sum_{j=1}^{J} \lambda_{ijk}^{ABC} = \sum_{k=1}^{K} \lambda_{ijk}^{ABC} = \sum_{i=1}^{I} \sum_{j=1}^{J} \lambda_{ijk}^{ABC} =$$

$$\sum_{i=1}^{I} \sum_{k=1}^{K} \lambda_{ijk}^{ABC} = \sum_{j=1}^{J} \sum_{k=1}^{K} \lambda_{ijk}^{ABC} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \lambda_{ijk}^{ABC} = 0$$

Multiplicative and Additive Models (cont.)

- Implications of these constraints on the loglinear model:

$$When \quad I = J = K = 2$$

$$\lambda_1^A = -\lambda_2^A$$

$$\lambda_{11}^{AB} = -\lambda_{12}^{AB} = -\lambda_{21}^{AB} = \lambda_{22}^{AB}$$

$$\lambda_{111}^{ABC} = -\lambda_{112}^{ABC} = -\lambda_{121}^{ABC} = -\lambda_{211}^{ABC} = \lambda_{221}^{ABC} = \lambda_{212}^{ABC} = \lambda_{122}^{ABC} = -\lambda_{222}^{ABC}$$

Interpreting Loglinear Models

- Logit models have a DV

$$\log(f_{1|ij} / f_{2|ij}) = \log(f_{ij1}) - \log(f_{ij2})$$

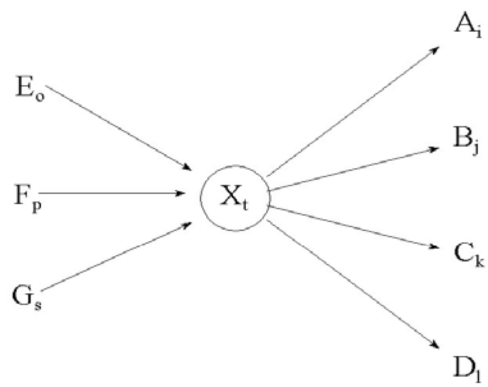
$$so,$$

$$\log(f_{1|11} / f_{2|11}) = (\mu + \lambda_1^A + \lambda_1^B + \lambda_1^C + \lambda_{11}^{AB} + \lambda_{11}^{AC} + \lambda_{11}^{BC} + \lambda_{111}^{ABC}) - (\mu + \lambda_1^A + \lambda_1^B + \lambda_2^C + \lambda_{11}^{AB} + \lambda_{12}^{AC} + \lambda_{12}^{BC} + \lambda_{112}^{ABC})$$

$$or,$$

$$\log(f_{1|11} / f_{2|11}) = (\lambda_1^C - \lambda_2^C) + (\lambda_{11}^{AC} - \lambda_{12}^{AC}) + (\lambda_{11}^{BC} - \lambda_{12}^{BC}) + (\lambda_{111}^{ABC} - \lambda_{112}^{ABC}) = 2\lambda_1^C + 2\lambda_{11}^{AC} + 2\lambda_{11}^{BC} + 2\lambda_{111}^{ABC}$$

Log-linear models with latent variables (latent logistic "mimic" model)



Probabilistic parameterization ("saturated model")

$$\pi_{ijklops}^{ABCDEFG} = \sum_{t=1}^{T} \pi_{tops}^{X|EFG} \pi_{itops}^{A|XEFG} \pi_{jtops}^{B|XEFG} \pi_{ktops}^{C|XEFG} \pi_{ltops}^{D|XEFG} \pi_{ops}^{EFG}$$

"Reduced" Model (above)

$$\pi_{ijklops}^{ABCDEFG} = \sum_{t=1}^{T} \pi_{to}^{X|E} \pi_{tp}^{X|F} \pi_{ts}^{X|G} \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X} \pi_{o}^{E} \pi_{p}^{F} \pi_{s}^{G}$$

Log-linear Parameterization

$$\ln(\hat{F}^{ABCDXEFG}_{ijkltops}) = \lambda + \lambda^{E}_{o} + \lambda^{F}_{p} + \lambda^{G}_{s} + \lambda^{X}_{t} + \lambda^{A}_{i} + \lambda^{B}_{j} + \lambda^{C}_{k} + \lambda^{D}_{l} +$$
$$\lambda^{AX}_{it} + \lambda^{BX}_{jt} + \lambda^{CX}_{kt} + \lambda^{DX}_{lt} + \lambda^{XE}_{to} + \lambda^{XF}_{tp} + \lambda^{XG}_{ts} + \lambda^{AE}_{io} + \lambda^{BE}_{jo} + \lambda^{CE}_{ko} + \lambda^{DE}_{lo} + \cdots$$
$$+ \lambda^{CG}_{ks} + \lambda^{DG}_{ls} + \lambda^{AXE}_{ito} + \cdots + \lambda^{DXG}_{lts} + \lambda^{AXEF}_{itop} + \cdots + \lambda^{DXFG}_{ltps} + \lambda^{AXEFG}_{itops} + \cdots$$
$$+ \lambda^{DXEFG}_{ltops} + \lambda^{XEF}_{top} + \cdots + \lambda^{XEFG}_{tops} + \lambda^{EF}_{op} + \cdots + \lambda^{EFG}_{ops}$$

Reduced

$$\ln(\hat{F}^{ABCDXEFG}_{ijkltops}) = \lambda + \lambda^{E}_{o} + \lambda^{F}_{p} + \lambda^{G}_{s} + \lambda^{X}_{t} + \lambda^{A}_{i} + \lambda^{B}_{j} + \lambda^{C}_{k} + \lambda^{D}_{l} +$$
$$\lambda^{AX}_{it} + \lambda^{BX}_{jt} + \lambda^{CX}_{kt} + \lambda^{DX}_{lt} + \lambda^{XE}_{to} + \lambda^{XF}_{tp} + \lambda^{XG}_{ts} + \lambda^{AG}_{is} + \lambda^{BG}_{js} + \lambda^{CG}_{ks} + \lambda^{DG}_{ls}$$

Logit

$$\ln\left(\hat{F}^{ABCDXEFG}_{ijkl1ops} \Big/ \hat{F}^{ABCDXEFG}_{ijkl2ops} \right) = \ln(\hat{F}^{ABCDXEFG}_{ijkl1ops}) - \ln(\hat{F}^{ABCDXEFG}_{ijkl2ops})$$

For our model, this yields (recalling the identifying restrictions)

$$\ln\left(\hat{F}^{ABCDXEFG}_{ijkl1ops} \Big/ \hat{F}^{ABCDXEFG}_{ijkl2ops} \right) = 2\lambda^{X}_{1} + 2\lambda^{AX}_{11} + 2\lambda^{BX}_{11} + 2\lambda^{CX}_{1} + 2\lambda^{CX}_{11} + 2\lambda^{DX}_{11} + 2\lambda^{EX}_{11} + 2\lambda^{FX}_{11} + 2\lambda^{GX}_{11}$$


```
* Belief in God Latent logistic example w/
* 3 indpendent variables; order f,e,d,c,b,a
* a=v31d, b=v32d, c=v33d, d=m/f, e=city,
* f=w/e; city: 1=50k+, 2=5-50k, 3=lt 5k
lat 1
man 6
dim 2 2 3 2 2 2 2
lab X F E D C B A
mod X|DEF {XD,XE,XF}
    A|X
    B|X
    C|XF {XC,FC}
dat [51  11   3 128   2   1   3 249  28   6
      2  96   1   0   8 313
     13   3   0  25   1   0   2 105   6   2
      3  23   0   0   2 106
      5   0   0  12   0   0   2  54   4   0
      0  18   0   0   1  74
    138  14   3  26   2   0   0  39 129  15
      4  50   0   1   2  64
    117  14   9  24   0   0   2  56  99  14
     11  40   1   1   2  85
    109  11   5  46   1   1   0  58  82  14
      5  61   1   1   1 111 ]
```

```
X-squared             = 81.1028 (0.2166)
  L-squared             = 85.3247 (0.1350)
  Cressie-Read          = 80.5868 (0.2284)
  Dissimilarity index   = 0.0519
  Degrees of freedom    = 72
  Log-likelihood        = -9959.26854
  Number of parameters  = 12 (+12)
  Sample size           = 2832.0
  BIC(L-squared)        = -486.9844
  AIC(L-squared)        = -58.6753
```



```
* P(A|X) *
  1 | 1        0.8834  (0.0111)
  1 | 2        0.0179  (0.0035)

* P(B|X) *
  1 | 1        0.9562  (0.0081)
  1 | 2        0.0026  (0.0019)

* P(C|XF) *
  1 | 1 1      0.9808  (0.0065)
  2 | 1 1      0.0192  (0.0065)
  1 | 1 2      0.9892  (0.0036)
  2 | 1 2      0.0108  (0.0036)
  1 | 2 1      0.2487  (0.0124)
  2 | 2 1      0.7513  (0.0124)
  1 | 2 2      0.3731  (0.0189)
  2 | 2 2      0.6269  (0.0189)
```

```
* P(X|FED) *
  1 | 1 1 1      0.1458  (0.0129)
  1 | 1 1 2      0.0811  (0.0084)
  1 | 1 2 1      0.1111  (0.0133)
  1 | 1 2 2      0.0607  (0.0083)
  1 | 1 3 1      0.0746  (0.0099)
  1 | 1 3 2      0.0400  (0.0059)
  1 | 2 1 1      0.7107  (0.0215)
  1 | 2 1 2      0.5596  (0.0239)
  1 | 2 2 1      0.6428  (0.0238)
  1 | 2 2 2      0.4821  (0.0248)
  1 | 2 3 1      0.5373  (0.0255)
  1 | 2 3 2      0.3752  (0.0234)
  2 | 1 1 1      0.8542  (0.0129)
  2 | 1 1 2      0.9189  (0.0084)
  2 | 1 2 1      0.8889  (0.0133)
  2 | 1 2 2      0.9393  (0.0083)
  2 | 1 3 1      0.9254  (0.0099)
  2 | 1 3 2      0.9600  (0.0059)
  2 | 2 1 1      0.2893  (0.0215)
  2 | 2 1 2      0.4404  (0.0239)
  2 | 2 2 1      0.3572  (0.0238)
  2 | 2 2 2      0.5179  (0.0248)
  2 | 2 3 1      0.4627  (0.0255)
  2 | 2 3 2      0.6248  (0.0234)
```

```
* Belief in God Latent logistic example w/
* a=v31d, b=v32d, c=v33d, d=m/f, e=city,
  * f=w/e; city: 1=50k+, 2=5-50k, 3=lt 5k
```


Models for Analyzing Categorically-Scored Panel Data

Two-wave Panel Data

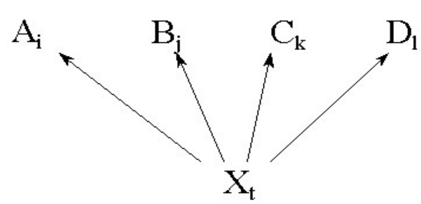
- LCMs can be used to analyze wide range of types of panel data
- With 2-wave panel data, multi-level variables
  - E.g., parent-child occupational status, party identification
  - Example from Jennings, van Deth, et al Political Action Panel Study (1991)

Example: U.S. Party Identification

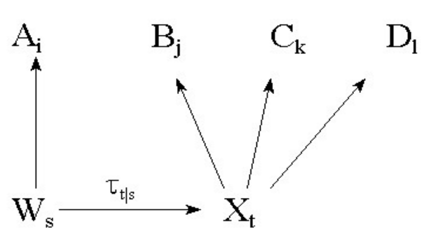
1981 / 1974	Democr.	Indep./ Democr.	Indep.	Indep./ Repub.	Repub.
Democr.	263	37	17	15	19
Ind./ Democr.	34	49	15	16	14
Ind./ Ind.	29	17	62	27	20
Ind./ Repub.	5	7	5	30	30
Repub.	10	4	8	24	157

Panel Data

- Multiple observations on same subjects
- Assume identical measure at four points in time: A_i, B_j, C_k, D_l
- The "no change" model
  - Usual LCM
- The "learning" model
  - Single trial learning, or variant
- Change models
  - Markov model variations



No Change Model



Socratic (learning) Model

No Change Model

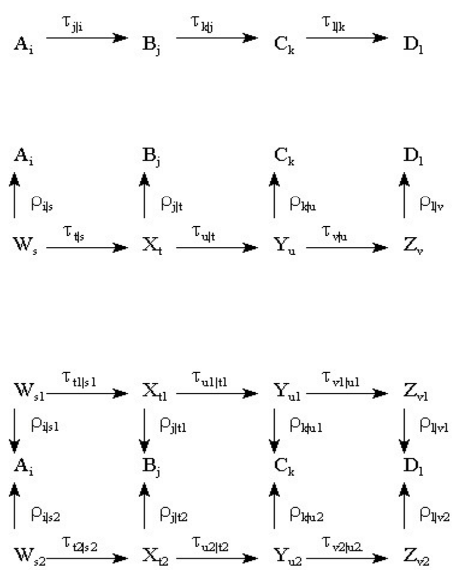
- Four measures of same phenomenon
  - E.g., brand purchase, vote intention
- Usual LCM tests that only measurement error is needed to explain obs. distrib.
- As will be shown later, this is equivalent to an identity matrix to the transitions

Learning Model

- Single-trial learning model: “Socratic”
  - Cumulative learning model: Markov

$$\pi_{ijkl}^{ABCD} = \sum_{s,t} \pi_{st}^{WX} \, \pi_{is}^{A|W} \, \tau_{t|s} \, \pi_{jt}^{B|X} \, \pi_{kt}^{C|X} \, \pi_{lt}^{D|X}$$

- Transition matrix $\tau_{t|s} = \begin{matrix} \tau_{1|1} & \tau_{2|1} \\ \tau_{1|2} & \tau_{2|2} \end{matrix}$



Markov Models

- Discrete time, discreet state
- Assumes no measurement error
- Estimate transition matrices (tau)
- E.g., $[.8 \;\; .2] * \begin{matrix} .9 & .1 \\ .2 & .8 \end{matrix} = [.76 \;\; .24]$
- “Stationary” models have only single tau

Markov Latent Class Models

- Discrete time, discrete state
- Measurement error is explicit aspect
- Each indicator, or set of indicators, measures latent state at specific time
- Measurement may change (rho)
- States may change over time
- Estimate transition matrices (tau)

Mixed Markov Latent Class Model

- MMLCMs:discrete time and state model
- Test that two (or more) latent Markov "chains" characterize obs. distribution
- Chain specific error structures
- Chain specific transition rates
- Estimation of so many parameters may result in identification problems
  - May want to impose some apriori restrictions

Mixed Markov Latent Class Models

- Let delta (r) represent the latent chains

$$\pi_{ijkl}^{ABCD} = \sum_{r=1}^{R}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L} \delta_r \, \rho_{i|sr} \, \tau_{tr|sr} \, \rho_{j|tr} \, \tau_{ur|tr} \, \rho_{k|ur} \, \tau_{vr|ur} \, \rho_{l|vr}$$

- Let rho represent the measurement of the latent states
- Let tau represent the latent transition matrices

British Election Study Panel Data

1993, 1995, 1996, 1997

Conservative, Labor, Other

296	2	10	3	2	0	7	1	1
2	0	0	2	5	1	0	0	1
12	2	4	0	1	0	7	1	13
3	0	0	0	1	0	0	0	0
0	0	1	4	345	11	1	10	6
0	0	0	1	9	2	0	6	9
15	0	4	1	1	0	3	0	2
0	2	0	2	31	6	0	4	8
9	1	4	0	5	8	14	14	155

Mixed Markov Latent Class Model: Mover-Stayer Model

```
{df 45,  indep.pars 35+0,  LR 36.228,  Chi-2 41.915,
AIC 4935.9,  BIC 5110.1}

chain proportions

   0.290276596   0.709723404


{initial probabilities of group 1,  }

{chain 1}   0.241333496   0.424339960   0.334326544
{chain 2}   0.414382039   0.371593932   0.214024029
```


Mixed Markov Latent Class Model: Mover-Stayer Model (cont.)

Time homogeneous response probabilities (restriction)

```
{response probs., group 1, , chain 1, indicator/wave 1 }
{class 1}   0.817388460   0.000005902   0.182605638
{class 2}   0.013193300   0.850936789   0.135869912
{class 3}   0.038748266   0.000000000   0.961251734


{response probs., group 1, , chain 2, indicator/wave 1 }
{class 1}   0.974524046   0.007899203   0.017576751
{class 2}   0.000000000   0.999954341   0.000045659
{class 3}   0.000000063   0.063325785   0.936674153
```

Mixed Markov Latent Class Model: Mover-Stayer Model (cont.)

```
{transition probs., group 1, , chain 1, indicators/waves  1, 2 }
{class 1}   0.458222021   0.113550651   0.428227327
{class 2}   0.000000000   1.000000000   0.000000000
{class 3}   0.142601921   0.328780884   0.528617194

{transition probs., group 1, , chain 1, indicators/waves  2, 3 }
{class 1}   0.912941878   0.087057889   0.000000233
{class 2}   0.000000000   0.964913867   0.035086133
{class 3}   0.243960957   0.000000520   0.756038523

{transition probs., group 1, , chain 1, indicators/waves  3, 4 }
{class 1}   0.820655669   0.066646082   0.112698249
{class 2}   0.057896931   0.937648068   0.004455001
{class 3}   0.291538432   0.075361473   0.633100094
```


Mixed Markov Latent Class Model: Mover-Stayer Model (cont.)

```
{transition probs., group 1, , chain 2, indicators/waves  1, 2 }
{class 1}   1.000000000   0.000000000   0.000000000
{class 2}   0.000000000   1.000000000   0.000000000
{class 3}   0.000000000   0.000000000   1.000000000

{transition probs., group 1, , chain 2, indicators/waves  2, 3 }
{class 1}   1.000000000   0.000000000   0.000000000
{class 2}   0.000000000   1.000000000   0.000000000
{class 3}   0.000000000   0.000000000   1.000000000

{transition probs., group 1, , chain 2, indicators/waves  3, 4 }
{class 1}   1.000000000   0.000000000   0.000000000
{class 2}   0.000000000   1.000000000   0.000000000
{class 3}   0.000000000   0.000000000   1.000000000
```