Slide 1

GLLAMM models with discrete latent variables

Andrew Pickles, Sophia Rabe-Hesketh and Anders Skrondal

University of Manchester,
King's College London,
and Norwegian Institute for Public Health, Oslo.

Half day meeting on
Latent Class Analysis and Finite Mixture Models
Royal Statistical Society
Statistical Computing Section
London, UK

Slide 2

Outline

- General introduction to GLLAMM
- Exploratory and confirmatory latent class models
- Examples
  - Binary items on attitudes to abortion from a cross-sectional survey of complex design
  - Binary/ordinal diagnoses of autism from a longitudinal study
  - Ranked responses to items from a cross-sectional study of post-materialism

Slide 3

Generalized Linear Latent and Mixed Models: GLLAMM

Response model: linear predictor

(Illustrated for two level model)

Conditional on the latent variables, the response model is a generalized linear model with linear predictor

$$\nu_{ij} \;=\; \mathbf{x}'_{ij}\boldsymbol{\beta} + \sum_{m=1}^{M} \eta_{jm}\mathbf{z}'_{ij}\boldsymbol{\lambda},$$

- i indexes the units at level 1 (variables or items)
- j indexes the units at level 2 (units or subjects), with $i = 1, \cdots, n_j$
- $\boldsymbol{\beta}$ and $\boldsymbol{\lambda}$ are parameters,
- $\mathbf{x}_{ij}$ and $\mathbf{z}_{ij}$ are vectors of observed variables and known constants
- η_{jm} is the mth element of the latent variable vector $\boldsymbol{\eta}_j$.

Latent variables can be correlated, and can be dependent and/or independent variables

Slide 4

Response Model: The conditional distribution

- Conditional expectation:

  $$\mathrm{E}[y_{ij}|\mathbf{x}_{ij}, \mathbf{z}_{ij}, \boldsymbol{\eta}_j] \;=\; g^{-1}(\nu_{ij}),$$

  where $g(\cdot)$ is the link function
- The conditional distributions of the responses are from the exponential family
- The responses are conditionally independent given the latent variables
- The responses can be of mixed types (different links and distributions)

Slide 5

Response Model: Response types

- Continuous (normal, gamma)
- Binary (logit, probit or complementary log-log links)
- Ordinal
  - Cumulative models (logit, probit or complementary log-log links)
    including models for thresholds or scale parameter
  - Adjacent category odds model
  - Continuation ratio model
- Unordered categorical and rankings (multinomial logit)
- Counts (Poisson, binomial)
- Durations in continuous time
  - Proportional hazards model
  - Accelerated failure time model
- Durations in discrete time
  - Censored cumulative models
  - Continuation ratio model
  - Proportional hazards in continuous time
- Mixed responses

Slide 6

`gllamm` Platform, Command and Data Structure

- `gllamm` is implemented as a set of Stata procedures (MS Windows/NT, Mac and Unix/Linux)
- `gllamm` command for point and precision estimation
- `gllapred` command for prediction including empirical Bayes
- `gllasim` command for simulation given model, explanatory variables and parameter values
- "One response per record" data structure (MAR for missing responses under ML - no balance assumed)

Slide 7

Discrete Latent Variables

- Latent variable vector $\boldsymbol{\eta}_j$ for unit j with discrete values (or locations) $\boldsymbol{e}_c, c\!=\!1, \cdots, C$ in M dimensions.
- Individuals in the same latent class share the same value or location $\boldsymbol{e}_c$.
- Let π_{jc} denote the (prior) probability that unit j is in latent class c.
- This probability may depend on covariates $\mathbf{v}_j$ through a multinomial logit model

$$\pi_{jc} \;=\; \frac{\exp(\mathbf{v}_j'\boldsymbol{\alpha}_c)}{\sum_d \exp(\mathbf{v}_j'\boldsymbol{\alpha}_d)},$$

  where $\boldsymbol{\alpha}_c$ are parameters with $\boldsymbol{\alpha}_1\!=\!0$ for identification.

Slide 8

Model Likelihood

Joint distribution of responses $\boldsymbol{y}_j$ for unit j:

$$\Pr(\boldsymbol{y}_j|\mathbf{X}_j, \boldsymbol{Z}_j, \mathbf{v}_j; \boldsymbol{\theta}) \;=\;$$

$$\sum_{c=1}^{C} \Pr(\boldsymbol{\eta}_j\!=\!\boldsymbol{e}_c|\mathbf{v}_j; \boldsymbol{\alpha}) \prod_{i=1}^{n_j} f(y_{ij}|\mathbf{X}_j, \boldsymbol{Z}_j, \boldsymbol{\eta}_j\!=\!\boldsymbol{e}_c; \boldsymbol{\beta}, \boldsymbol{\lambda})$$

- $\mathbf{X}_j$ and $\boldsymbol{Z}_j$ are $n_j \times p$ and $n_j \times q$ design matrices for the vector of responses,
- $\boldsymbol{\theta}$ is the vector of all parameters
- $f(y_{ij}|\mathbf{X}_j, \boldsymbol{Z}_j, \boldsymbol{\eta}_j\!=\!\boldsymbol{e}_c; \boldsymbol{\beta}, \boldsymbol{\lambda})$ is the conditional probability (density) of the response given the observed and latent variables.

Slide 9

Conventional Exploratory Models

Conventional exploratory latent class model imposes no structure on the conditional response probabilities. The linear predictor has the form

$$\nu_{ijc} \;=\; \beta_i + e_{ic}.$$

The model can be written as

$$\nu_{ijc} \;=\; \boldsymbol{d}_i'\boldsymbol{\beta} + \sum_{m=1}^{I} e_{mc} d_{mi},$$

where $\boldsymbol{d}_i$ is a vector of length I with ith element equal to 1 and all other elements equal to 0, $\boldsymbol{d}_i = (d_{1i}, \cdots, d_{Ii})'$ where

$$d_{mi} \;=\; \begin{cases} 1 & \text{if } m = i \\ 0 & \text{if } m \neq i \end{cases}$$

Here e_{mc} is the cth location of the mth latent variable.

Slide 10

"Confirmatory" Model

- One-factor model would be specified by

$$\nu_{ijc} \;=\; \beta_i + \lambda_i e_c.$$

- In GLLAMM framework the model can be written as

$$\nu_{ijc} \;=\; \boldsymbol{d}_i'\boldsymbol{\beta} + e_c \boldsymbol{d}_i'\boldsymbol{\lambda}.$$

- If we increase the number of classes until the likelihood cannot be increased any further, the discrete distribtution can be viewed as a nonparametric maximum likelihood (NPML) estimator of a continuous distribution.

Slide 11

Example 1 : Attitudes to abortion data

- British Social Attitudes Survey 1983
- Respondents were asked whether or not abortion should be allowed by law if:

[wom] The woman decides on her own she does not wish to have the child

[cou] The couple agree that they do not wish to have the child

[mar] The woman is not married and does not wish to marry the man

[fin] The couple cannot afford any more children

[gen] There is a strong chance of a genetic defect in the baby

[ris] The woman's health is seriously endangered by the pregnancy

[rap] The woman became pregnant as a result of rape

- 720 respondents, 89% have complete data, a total of 7% of items are missing

Slide 12

Attitudes to abortion: Data structure

id	ab	wom	cou	mar	fin	gen	ris	rap	fem	wt2	pwt2	area83
1	1	1	0	0	0	0	0	0	0	1	.8281	102
1	1	0	1	0	0	0	0	0	0	1	.8281	102
1	1	0	0	1	0	0	0	0	0	1	.8281	102
1	1	0	0	0	1	0	0	0	0	1	.8281	102
1	1	0	0	0	0	1	0	0	0	1	.8281	102
1	1	0	0	0	0	0	1	0	0	1	.8281	102
1	1	0	0	0	0	0	0	1	0	1	.8281	102
2	0	1	0	0	0	0	0	0	0	1	.621075	102
2	1	0	1	0	0	0	0	0	0	1	.621075	102
2	1	0	0	1	0	0	0	0	0	1	.621075	102
2	1	0	0	0	1	0	0	0	0	1	.621075	102
2	1	0	0	0	0	1	0	0	0	1	.621075	102
2	1	0	0	0	0	0	1	0	0	1	.621075	102
2	1	0	0	0	0	0	0	1	0	1	.621075	102
3	1	1	0	0	0	0	0	0	0	1	.8281	102
3	1	0	1	0	0	0	0	0	0	1	.8281	102
...												

Slide 13

Attitudes to abortion: Model specification

- Model 1: Discrete one-factor two-class model

$$\text{logit}[\Pr(y_{ij}=1|\eta_j=e_c)] = \beta_i + \lambda_i e_c,$$

```
eq fac: wom cou mar fin gen ris rap
gllamm ab wom cou mar fin gen ris rap, nocons
    weight(wt) i(id) l(logit) f(binom)
    eqs(fac) ip(f) nip(2)
```

- Model 2: Class probabilities depend on sex (v_j=`[fem]`)

$$\pi_{j1} = \frac{\exp(\alpha_0+\alpha_1 v_j)}{1+\exp(\alpha_0+\alpha_1 v_j)},\ \pi_{j2} = 1-\pi_{j1}.$$

```
eq fem: fem
gllamm ab wom cou mar fin gen ris rap, nocons
     weight(wt) i(id) l(logit) f(binom)
     eqs(fac) peqs(fem) ip(f) nip(2)
```

- Model 3: Include a direct effect of gender on the second item `[cou]`.

$$\text{logit}[\Pr(y_{2j}=1|\eta_j=e_c, v_j)] = \beta_{02} + \beta_{12} v_j + \lambda_i e_c.$$

```
gen femcou = fem*cou
gllamm ab wom cou femcou mar fin gen ris rap, ...
```

Slide 14

Parameter estimates for models with two latent classes

	Model 1	Model 2
Intercepts:		
β_1 [wom]	-0.49 (0.12)	-0.46 (0.12)
β_2 [cou]	0.39 (0.24)	0.60 (0.28)
β_3 [mar]	-0.19 (0.15)	0.06 (0.17)
β_4 [fin]	0.22 (0.14)	0.43 (0.16)
β_5 [gen]	2.69 (0.26)	2.86 (0.29)
β_6 [ris]	3.48 (0.47)	3.66 (0.52)
β_7 [rap]	2.85 (0.22)	2.95 (0.24)
Factor loadings:		
λ_1 [wom]	1 (–)	1 (–)
λ_2 [cou]	1.62 (0.24)	1.64 (0.24)
λ_3 [mar]	1.33 (0.16)	1.32 (0.16)
λ_4 [fin]	1.16 (0.15)	1.15 (0.15)
λ_5 [gen]	0.94 (0.22)	0.93 (0.21)
λ_6 [ris]	1.05 (0.39)	1.04 (0.38)
λ_7 [rap]	0.61 (0.19)	0.60 (0.18)
Locations parameter:		
e_1	-1.28 (0.14)	-1.47 (0.16)
Probability parameters (class 1):		
α_0 [cons]	0.24 (0.12)	-0.01 (0.17)
α_1 [fem]	–	0.48 (0.17)
Log-likelihood:	-1967.89	-1963.82

Slide 15

Models Using Complex Survey Data

- British Attitudes Survey not a simple random sample
- Pseudolikelihoood estimation with inverse probability weights
- Robust standard errors (sandwich estimator) for cluster sampling with electoral ward as psu.

`gllamm` options `pweight` and `robust`:

```
gllamm ab wom cou mar fin gen ris rap, nocons
   weight(wt) i(id)l(logit) f(binom)
   eqs(fac) ip(f) nip(2) peqs(fem)
   pweight(pwt) robust cluster(area83)
```

Slide 16

Design based adjustment for clustering and weighting using PML and sandwich variance estimator

	no weights model-based se	weights robust se	weights robust se, cluster
Intercepts:			
β_1 [wom]	-0.46 (0.12)	-0.26 (0.15)	(0.16)
β_2 [cou]	0.60 (0.28)	0.82 (0.39)	(0.40)
β_3 [mar]	0.06 (0.17)	0.04 (0.21)	(0.24)
β_4 [fin]	0.43 (0.16)	0.35 (0.18)	(0.21)
β_5 [gen]	2.86 (0.29)	2.81 (0.31)	(0.30)
β_6 [ris]	3.66 (0.52)	3.72 (0.58)	(0.61)
β_7 [rap]	2.95 (0.24)	2.87 (0.31)	(0.32)
Factor loadings:			
λ_1 [wom]	1 (-)	1 (-)	
λ_2 [cou]	1.64 (0.24)	1.67 (0.29)	(0.30)
λ_3 [mar]	1.32 (0.16)	1.31 (0.18)	(0.21)
λ_4 [fin]	1.15 (0.15)	1.12 (0.18)	(0.19)
λ_5 [gen]	0.93 (0.21)	0.87 (0.24)	(0.27)
λ_6 [ris]	1.04 (0.38)	1.12 (0.46)	(0.45)
λ_7 [rap]	0.60 (0.18)	0.57 (0.26)	(0.25)
Location parameter:			
e_1	-1.47 (0.16)	-1.40 (0.21)	(0.20)
Probability parameters (class 1):			
α_0 [cons]	-0.01 (0.17)	0.07 (0.21)	(0.19)
α_1 [fem]	0.48 (0.17)	0.43 (0.18)	(0.19)

Slide 17

Attitudes to Abortion: Prediction using `gllapred`

```
* e1, e2, e3 are locations
gllapred mup, mu us(e)
gllapred mu, mu marg
```

Illustrated for model extended to 3 classes

	class 1	class 2	class 3		
Prior Probabilities					
male	6	47	47		
female	4	60	36		
Conditional Probabilities				Marginal	
				male	female
[wom]	0	18	78	45	38
[cou]	0	20	98	56	47
[mar]	0	16	91	51	42
[fin]	0	26	92	56	48
[gen]	7	90	98	89	89
[ris]	33	96	99	94	94
[rap]	53	93	97	93	93

Slide 18

Example 2: Early Diagnosis of Autism

- Approx 130 children suspected of autism diagnosed by
  - structured interview [`adi`]
  - structured observation [`adosg`]
  - clinical judgement [`clin`]
- Children diagnosed as autistic, pdd-nos or non-autistic spectrum
- Assessments made at age 2,3,5 and 9
- Comparatively sparse data often leads to boundary or near boundary solutions.

Slide 19

Two-class Model Fitted to Binary Age-2 Data (autism versus non-autism)

The simple two-class exploratory model for diagnosis measure i on child j

$$\text{logit}[\Pr(y_{ij} = 1|c] = \beta_i + e_{ic}.$$

is equivalent to the one-factor model

$$\text{logit}[\Pr(y_{ij} = 1|c] = \beta_i + \lambda_i e_c.$$

Intercepts	
β_1 interview	0.08 (0.24)
β_2 observation	6.68 (295)
β_3 clinician	-1.95 (0.61)
Conventional parameterisation:	
Locations (class 1)	
e_{11} interview	-1.07 (0.27)
e_{21} observation	-7.47 (295.)
e_{31} clinician	-1.98 (0.52)
Alternative parameterisation:	
Location (class 1)	
e_1	-1.07 (0.27)
Factor Loadings	
λ_2	6.41 (170.)
λ_3	1.84 (0.61)
Probability parameter (class 1)	
α_0	0.36 (0.24)

Slide 20

Avoiding Boundary Solutions Using a Simulated Data Prior

For example, could
(i) estimate model with sensitivity and specificity of interview and observation set equal,
(ii) use `gllasim` to simulate data from this model,
(iii) append simulated data to real data and re-estimate with simulated data given suitable low weight e.g. here 4 replicates each weighted 0.025

Intercepts	
β_1 interview	0.11 (0.23)
β_2 observation	5.73 (34.7)
β_3 clinician	-1.89 (0.52)
One-factor parameterisation:	
Location (class 1)	
e_1	-1.07 (0.26)
Factor Loadings	
λ_2	6.07 (31.0)
λ_3	1.78 (0.55)
Probability parameter (class 1)	
α_0	0.33 (0.25)

Slide 21

Two-factor Models: Latent Transition Model Age 2 to Age 5

- The model is defined by specifying classes for latent histories 00,01,10 and 11.
- We restrict the factor loadings to be the same at age 2 and age 5.

```
eq load2: adi_t2 adosg_t2 clin_t2
eq load5: adi_t5 adosg_t5 clin_t5
cons def 1 [id1_1l]adosg_t2 = [id1_2l]adosg_t5
cons def 2 [id1_1l]clin_t2 = [id1_2l]clin_t5
```

- We restrict the location of 0 at age 2 in the (00) class to be the same as that in the (01) class, and similarly for 1 age 2 in (10) and (11), 0 at age 5 in (01) and (11), and 1 at age 5 in (01) and (11).

```
cons def 3 [z2_1_1]adi_t2 = [z2_1_4]adi_t2
cons def 4 [z2_2_1]adi_t5 = [z2_2_2]adi_t5
cons def 5 [z2_1_2]adi_t2 = [z2_1_3]adi_t2
cons def 6 [z2_2_3]adi_t5 = [z2_2_4]adi_t5
```

- Fit the model with the constraints

```
gllamm aut, nocons i(id) nrf(2) eqs(load2 load5)
    ip(fn) nip(4) f(binom) constr(1 2 3 4 5 6)
```

Slide 22

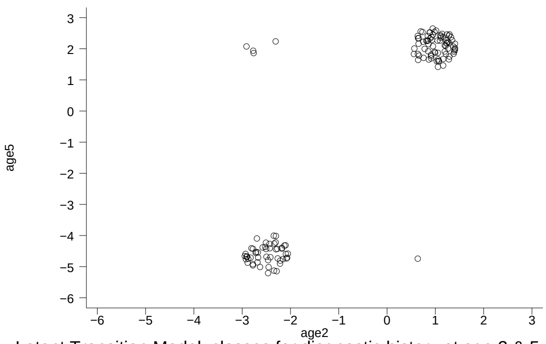
Autism Diagnosis: Child Classification to Latent Histories

- use

  ```
  gllapred p, p
  ```

  to calculate posterior class membership probabilities for each child, assigning each child to class with highest probability

Scatterplot of child class assignments (jittered)



Latent Transition Model: classes for diagnostic history at age 2 & 5

Slide 23

Random coefficient models: Latent Trajectory Models

- Discrete random intercept and slope of time
- Factor loadings for the three instruments
- Use the 3-category diagnosis as an ordinal response

$$\ln\left(\frac{\Pr(y_{ij}>s)}{\Pr(y_{ij}\leq s)}\right) = \beta_i + \lambda_i(e_{1c} + e_{2c}a_t) - \kappa_s,$$

$$\beta_1 = \kappa_1 = 0, \ \lambda_1 = 1, \ s = 1, 2.$$

```
gen adi_t = adi*t
gen adosg_t = adosg*t
gen clin_t = clin*t
eq int: adi adosg clin
eq slope: adi_t adosg_t clin_t
```

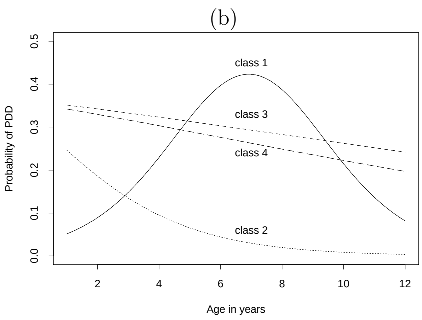
- Set factor loadings equal for intercept and slope and set $\kappa_1 = 0$

```
cons def 1 [id1_1l]adosg = [id1_2l]adosg_t
cons def 2 [id1_1l]clin = [id1_2l]clin_t
cons def 3 [_cut11]_cons=0
```

- Estimate model

```
gllamm dx adosg clin, i(id) f(binom) nrf(4)
    eqs(int slope) ip(fn) nip(4) constr(1 2 3)
```

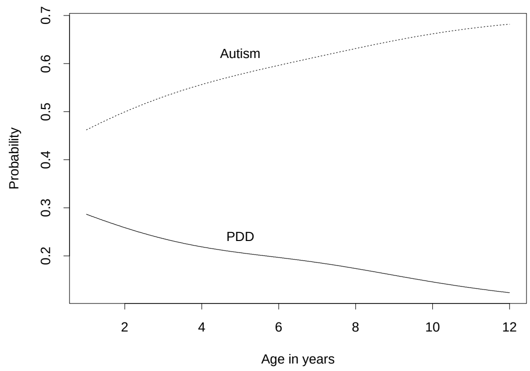
Slide 24



Conditional predicted probabilities of (a) autism and (b) pdd for the four latent classes

Slide 25

Marginal predicted probabilities of autism and pdd



Slide 26

Example 3: Post-materialism study

A latent class model for rankings

In 1974/1975, over 2000 German respondents ranked four political goals according to their desirability:

1. Maintain order in the nation (ORDER)
2. Give people more say in decisions of the government (SAY)
3. Fight rising prices (PRICES)
4. Protect freedom of speech (FREEDOM)

Croon describes a latent class analysis of rankings.

Slide 27

Model for Rankings

- Probability of a ranking of objects by subject j R_j can be written as

$$\Pr(R_j) \;=\; \frac{\exp(\nu_j^{r_j^1})}{\Sigma_{a=1}^{S}\exp(\nu_j^{r_j^a})} \times \frac{\exp(\nu_j^{r_j^2})}{\Sigma_{a=2}^{S}\exp(\nu_j^{r_j^a})} \times \cdots \times \frac{\exp(\nu_j^{r_j^S})}{\Sigma_{a=S-1}^{S}\exp(\nu_j^{r_j^a})}$$

  where r_j^a is the object given rank a.

- Each term represents the probability of choosing the object among the remaining objects.
- The latent class model is exploratory with

$$\nu_{jc}^s = e_c^s, \quad s = 1, 2, 3$$

$$\nu_{jc}^4 = 0$$

Slide 28

Rankings Model Parameter Estimates

	one class	two classes	three classes
class 1			
probability	1	0.79	0.45
locations			
e_1^1 [ORDER]	1.16 (0.04)	1.94 (0.09)	1.84 (0.15)
e_1^2 [SAY]	0.21 (0.04)	0.21 (0.05)	0.17 (0.09)
e_1^3 [PRICES]	1.28 (0.04)	1.87 (0.09)	2.96 (0.31)
class 2			
probability		0.21	0.23
locations			
e_2^1 [ORDER]		-0.87 (0.09)	-0.76 (0.26)
e_2^2 [SAY]		0.44 (0.12)	0.56 (0.12)
e_2^3 [PRICES]		-0.21 (0.16)	-0.09 (0.19)
class 3			
probability			0.32
locations			
e_3^1 [ORDER]			3.14 (0.40)
e_3^2 [SAY]			0.21 (0.10)
e_3^3 [PRICES]			1.18 (0.16)
log-likelihood	-6427.05	-6311.69	-6281.36

Slide 29

Data					Results for 3 class model				
						Posterior prob. (%)			
					pred.	class	1	2	3
Ranking				freq.	freq.	prior	0.45	0.23	0.32
1	2	3	4	137	126		10	10	**80**
1	2	4	3	29	46		1	31	**67**
1	3	2	4	309	315		36	3	**61**
1	3	4	2	255	257		37	2	**61**
1	4	2	3	52	40		2	26	**73**
1	4	3	2	93	93		11	6	**83**
2	1	3	4	48	50		20	**40**	**40**
2	1	4	3	23	29		2	**77**	22
2	3	1	4	61	57		**48**	46	6
2	3	4	1	55	61		7	**93**	0
2	4	1	3	33	32		1	**96**	3
2	4	3	1	59	61		2	**98**	0
3	1	2	4	330	339		**85**	3	12
3	1	4	2	294	281		**86**	2	12
3	2	1	4	117	109		**79**	18	4
3	2	4	1	69	56		25	**75**	0
3	4	1	2	70	81		**87**	9	4
3	4	2	1	34	41		32	**67**	0
4	1	2	3	21	18		3	**66**	32
4	1	3	2	30	30		27	21	**51**
4	2	1	3	29	25		2	**94**	4
4	2	3	1	52	47		3	**97**	0
4	3	1	2	35	33		**68**	23	9
4	3	2	1	27	33		13	**87**	0

Slide 30

`gllamm` continues to develop.

Documentation, examples and the stata code are freely available from

http://www.iop.kcl.ac.uk/IoP/
Departments/BioComp/programs/gllamm.html