Slide 1

Generalized linear latent and mixed modeling (GLLAMM)

Sophia Rabe-Hesketh
Institute of Psychiatry, London

joint work with

Anders Skrondal
Norwegian Institute of Public Health, Oslo

and

Andrew Pickles
The University of Manchester

UCLA, March 2003

Slide 2

Outline

- Structure of GLLAMM models and `gllamm` syntax
  - Response model: Generalized linear model conditional on latent variables
    - Linear predictor: latent variables as factors or random coefficients
    - Links and distributions
  - Structural model:
    - Regressions of latent variables on observed variables
    - Regressions of latent variables on other latent variables
  - Distribution of the latent variables (disturbances)
    - Multivariate normal
    - Discrete
- Application: Cluster randomized study of sex education in Norwegian schools

Slide 3

Linear Predictor in GLLAMM

$$\nu = \mathbf{x}'\boldsymbol{\beta} + \sum_{l=2}^{L}\sum_{m=1}^{M_l} \eta_m^{(l)} \mathbf{z}_m^{(l)\prime} \boldsymbol{\lambda}_m^{(l)} \quad \text{for identification, } \lambda_{m1}^{(l)} = 1$$

- Fixed part: $\mathbf{x}'\boldsymbol{\beta}$ as usual
- Random part:
  - $\eta_m^{(l)}$ is mth latent variable at level l, $m = 1, \cdots, M_l$, $l = 2, \cdots, L$
  - $\eta_m^{(l)}$ can be a **factor** or a **random coefficient**
  - $\mathbf{z}_m^{(l)}$ are variables and $\boldsymbol{\lambda}_m^{(l)}$ are parameters
  - Unless regressions for the latent variables are specified, latent variables at different levels are independent whereas latent variables at the same level may be dependent.

Slide 4

Random coefficient models in GLLAMM

- One covariate multiplies each latent variable,

$$\eta_m^{(l)} z_m^{(l)} \quad (\lambda_m^{(l)} = 1)$$

- e.g. Latent growth curve model for individuals j (level 2) observed at times t_{ij}, $i = 1, \cdots, n_j$ (level 1)

$$\nu_{ij} = \beta_1 + \beta_2 t_{ij} + \eta_{1j}^{(2)} + \eta_{2j}^{(2)} t_{ij}$$

β_1, β_2 : mean intercept and slope

$\eta_{1j}^{(2)}$, $\eta_{2j}^{(2)}$: random deviations of the subject-specific intercepts and slopes from their means

  - The model can also be defined as

$$\begin{aligned} \nu_{ij} &= b_{1j} + b_{2j} t_{ij} \\ b_{1j} &= \beta_1 + \eta_{1j}^{(2)} \\ b_{2j} &= \beta_2 + \eta_{2j}^{(2)} \end{aligned}$$

Factor models in GLLAMM

- A linear combination of dummy variables for the items multiplies each latent variable,

$$\eta_m^{(l)} \mathbf{z}_m^{(l)\prime} \boldsymbol{\lambda}_m^{(l)}, \ \lambda_{m1}^{(l)} = 1$$

- e.g. One-factor model for items i, $i = 1, \cdots, I$ (level 1) and subjects j (level 2)

$$\begin{aligned} \nu_{ij} &= \mathbf{d}_i'\boldsymbol{\beta} + \eta_j^{(2)}\mathbf{d}_i'\boldsymbol{\lambda} \\ &= d_{1i}\beta_1 + \cdots + d_{Ii}\beta_I + \eta_j^{(2)}(d_{1i} + d_{2i}\lambda_2 \cdots + d_{Ii}\lambda_I) \\ &= \beta_i + \eta_j^{(2)}\lambda_i \end{aligned}$$

Slide 5

where

$$d_{pi} = \begin{cases} 1 & \text{if } p = i \\ 0 & \text{otherwise} \end{cases}, \quad \mathbf{d}_i' = (d_{1i}, \cdots, d_{Ii})$$

β_i: intercept for item i

$\eta_j^{(2)}$: common factor

λ_i: factor loading for item i, $\lambda_1 = 1$

unit j	item i	d_{1i}	d_{2i}	$\cdots$	d_{Ii}	y_{ij}
1	1	1	0	$\cdots$	0	y_{11}
1	2	0	1	$\cdots$	0	y_{21}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
1	I	0	0	$\cdots$	1	y_{I1}

Syntax for linear predictor in `gllamm`

`gllamm` [*varlist*] [`if` *exp*] [`in` *range*] , `i`(*varlist*)
[`nrf`(*numlist*) `eqs`(*eqnames*) `noconstant`
`offset`(*varname*) `constraints`(*numlist*) $\cdots$

Slide 6

`i`(*varlist*) $L-1$ variables identifying the hierarchical, nested clusters, from level 2 to L, e.g., `i(pupil class school)`.

`nrf`(*numlist*) $L-1$ numbers specifying the numbers of latent variables M_l at each level.

`eqs`(*eqnames*) $M = \sum M_l$ equations for the $\mathbf{z}_m^{(l)\prime}\boldsymbol{\lambda}_m^{(l)}$ multiplying each latent variable. No constant is assumed unless explicitly included in the equation definition.

`noconstant` no constant in the fixed part $\mathbf{x}'\boldsymbol{\beta}$.

`offset`(*varname*) variable in fixed part with regression coefficient set to 1.

`constraints`(*numlist*) list of linear parameter constraints defined using the `constraint define` command.

Slide 7

Syntax examples: linear predictor

- Two-level growth curve model (occasions in subjects)

```
gen cons=1
eq int: cons
eq slope: time
gllamm y time, i(subject) nrf(2) eqs(int slope) ...
```

- Three-level growth curve model (occasions in subjects in centers)

```
gllamm y time, i(subject center) nrf(2 2) /*
   */ eqs(int slope int slope) ...
```

- One-factor model

```
tab items, gen(d)  /* create dummy variables */
eq fact: d1-d5
gllamm y d1-d5, i(subject) nrf(1) eqs(fact) nocons ...
```

- Two-factor model (independent clusters)

```
eq fact1: d1-d5
eq fact2: d6-d10
gllamm y d1-d10, i(subject) nrf(2) /*
   */ eqs(fact1 fact2) nocons ...
```

Slide 8

Links and families in GLLAMM

- The conditional expectation of the response is 'linked' to the linear predictor

$$g(E[y|\mathbf{x}, \boldsymbol{\eta}, \mathbf{z}]) = \nu$$

- The conditional distribution of the response is from the exponential family:

Links
identity
reciprocal
logarithm
logit
probit
scaled probit
compl. log-log

Families
Gaussian
gamma
Poisson
binomial

Ordinal responses
ordinal logit
ordinal probit
ordinal compl. log-log
scaled ord. probit

Nominal & Rankings
multinomial logit

- Heteroscedasticity: Standard deviation or scale parameter σ can be modelled as $\log \sigma = \mathbf{z}^{(1)\prime}\boldsymbol{\alpha}$

Slide 9

Options for links and families

[··· family(*families*) fv(*varname*) link(*links*) lv(*varname*) s(*eqname*) ···]

`family(`*families*`)` family (or families) to be used.

`fv(`*varname*`)` variable whose values indicate which family applies to which observation.

`link(`*links*`)` and `lv(`*varname*`)` analogous to `family(`*families*`)` and `fv(`*varname*`)`.

`s(`*eqname*`)` equation for log standard deviation or scale parameter.

Slide 10

Syntax examples: One-factor models

- Dichotomous responses: binomial probit

```
tab item, gen(d)
eq fact: d1-d5
gllamm d1-d5, i(subject) nrf(1) eqs(fact) nocons /*
  */ link(probit) family(binom)
```

- Continuous responses: normal with item-specific unique factor variances

```
eq het: d1-d5
gllamm d1-d5, i(subject) nrf(1) eqs(fact) nocons /*
  */ link(ident) family(gauss) s(het)
```

- Mixed responses (items 1,2 normal and 3,4,5 binomial probit)

```
/* key = 1 for items 1,2; key=2 for items 3,4,5 */
gen key = (d1+d2) + 2(d3+d4+d5)
/* different variances for items 1 and 2 */
eq het: d1 d2
eq fact: d1-d5
gllamm y d1-d5, i(subject) nrf(1) eqs(fact) nocons /*
  */ link(ident probit) family(gauss binom) /*
  */ lv(key) fv(key) s(het)
```

Slide 11

Structural model in GLLAMM

- Regressions of latent variables on other latent and explanatory variables at the same or higher levels

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\mathbf{w} + \boldsymbol{\zeta}$$

- $\boldsymbol{\eta} = (\eta_1^{(2)}, \eta_2^{(2)}, \cdots, \eta_{M_2}^{(2)}, \cdots, \eta_1^{(l)}, \cdots, \eta_{M_l}^{(l)}, \cdots, \eta_{M_L}^{(L)})'$
  - $M = \sum_l M_l$ latent variables:
    - factors
    - random coefficients
- $\mathbf{B}$ is an upper diagonal $M \times M$ matrix of regression coefficients
- $\boldsymbol{\Gamma}$ is an $M \times p$ matrix of regression coefficients
- $\mathbf{w}$ is a p dimensional vector of explanatory variables
- $\boldsymbol{\zeta}$ is an M dimensional vector of errors/disturbances (same level as corresponding elements in $\boldsymbol{\eta}$).

Slide 12

Options for the structural model

[··· `geqs`(*eqnames*) `bmatrix`(*matname*)
`constraints`(*numlist*) ···]

`geqs`(*eqnames*) equations for regressions of latent variables on explanatory variables. The second character of each equation name indicates which latent variables is regressed on the predictors.

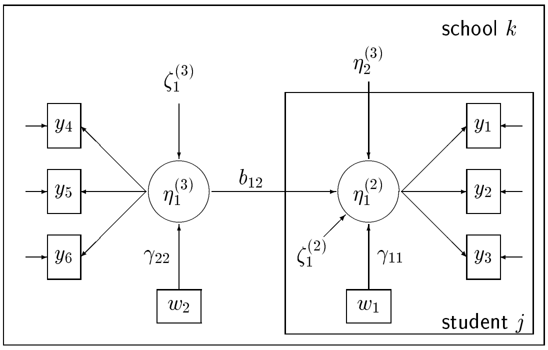
`bmatrix`(*matrix*) $M \times M$ matrix of 1s and 0s. Elements equal to 0 indicate that the corresponding element in $\mathbf{B}$ is 0; elements equal to 1 that the corresponding element in $\mathbf{B}$ should be estimated.

`constraint`(*numlist*) list of linear parameter constraints defined using the `constraint define` command.

Slide 13

Specifying a multilevel structural equation model

$$\begin{bmatrix} \eta^{(2)}_{1jk} \\ \eta^{(3)}_{1k} \\ \eta^{(3)}_{2k} \end{bmatrix} = \begin{bmatrix} 0 & b_{12} & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta^{(2)}_{1jk} \\ \eta^{(3)}_{1k} \\ \eta^{(3)}_{2k} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & 0 \\ 0 & \gamma_{22} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w_{1jk} \\ w_{2k} \end{bmatrix} + \begin{bmatrix} \zeta^{(2)}_{1jk} \\ \zeta^{(3)}_{1k} \\ \zeta^{(3)}_{2k} \end{bmatrix}$$



Slide 14

```
/* Equations for response model */
eq fact1: d1-d3
eq fact2: d4-d6
gen zero = 0
eq zero: zero

/* B-matrix */
matrix B=(0, 1, 1\ 0, 0, 0\ 0, 0, 0)
constr def 1 [b1_3]_cons = 1

/* Equations for regressions of latent variables
on observed variables */
eq f1: w1
eq f2: w2

gllamm y d1-d6, i(student school) nrf(1 2)     /*
  */ eqs(fact1 fact2 zero) bmat(B) geqs(f1 f2) /*
  */ constr(1) nip(8 4 4) nocons nocor adapt ...
```

Slide 15

Distribution of the latent variables

- Specify distribution of $\boldsymbol{\zeta}$. If there is no structural model, $\boldsymbol{\eta} = \boldsymbol{\zeta}$
- Disturbances at different levels are independent
- Multivariate normal $\boldsymbol{\zeta}^{(l)} = (\zeta_1^{(l)}, \cdots, \zeta_{M_l}^{(l)})'$ with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_l$
  - Program estimates Cholesky decomposition $\mathbf{Q}_l$ of $\boldsymbol{\Sigma}_l$, $\boldsymbol{\Sigma}_l = \mathbf{Q}_l\mathbf{Q}_l'$
  - The integral over the distribution of $\boldsymbol{\zeta}^{(l)}$ is approximated by integrating over independent standard normal $\mathbf{v}$ with $\boldsymbol{\zeta} = \mathbf{Q}_l\mathbf{v}$ using product quadrature
- Discrete $\boldsymbol{\zeta}^{(l)} = \mathbf{e}_c^{(l)}$ with probability π_c, where $\mathbf{e}_c^{(l)}$ are points in M_l dimensions, $c = 1, \cdots, C_l$
  - Interpretation as latent classes
  - Nonparametric maximum likelihood (NPML)

Slide 16

Options for the distribution of the disturbances

[··· `ip`(*string*) `nocorrel` `nip`(*numlist*) `adapt` ···]

`ip`(*string*) if string is `g`, the disturbances have a multivariate normal distribution and if string is `f`, the mass-points are freely estimated. The default is `g`.

`nocorrel` sets all correlations to zero if latent variables are multivariate normal.

`nip`(*numlist*) numbers of quadrature points or locations of the latent variables. For quadrature, a number is given for each latent variable (total M); for discrete latent variables, a number is given for each level (total $L-1$). A single number means that all values are the same.

`adapt` adaptive quadrature will be used (if `ip(g)` is specified) instead of ordinary Gauss-Hermite quadrature.

Slide 17

Syntax for prediction using gllapred

gllapred *varname* [if *exp*] [in *range*] [, u
fac linpred mu marginal us(*varname*)
outcome(#) above(#) ···]

u posterior means and standard deviations of disturbances $\boldsymbol{\zeta}$ returned in *varname*m1, *varname*s1, *varname*m2, etc.

fac posterior means and standard deviations of latent variables $\boldsymbol{\eta}$ returned in *varname*m1, *varname*s1, *varname*m2, etc.

linpred linear predictor (with posterior means of latent variables) returned in *varname*.

mu mean response returned in *varname*. Without further options, mean w.r.t. posterior distribution.

marginal together with mu, causes marginal or population average mean to be returned (mean w.r.t. prior distribution).

us(*varname*) together with mu, causes conditional mean to be returned, conditional on latent variables being equal to the values in *varname*1, *varname*2, etc.

outcome(#) with mlogit link, causes mu option to return probability that the response equals #.

above(#) with ordinal links, causes mu option to return probability that the response exceeds #.

Slide 18

Syntax for simulation using gllasim

gllasim *varname* [if *exp*] [in *range*] [, u fac
us(*varname*) from(*matrix*) ···]

By default, responses are simulated for the model just estimated and returned in *varname*.

u disturbances $\boldsymbol{\zeta}$ are simulated and returned in *varname*p1, *varname*p2, etc.

fac latent variables $\boldsymbol{\eta}$ are simulated and returned in *varname*p1, *varname*p2, etc.

us(*varname*) response variables are simulated for latent variables equal to *varname*1, *varname*2, etc.

from(*matrix*) causes simulations to be based on the model just estimated in gllamm but with parameter values in *matrix*.

Slide 19

Estimation and prediction in `gllamm`

- To obtain the likelihood of GLLAMM's, the latent variables must be integrated out
  - Sequentially integrate over latent variables, starting with the lowest level using a recursive algorithm
  - Use Gauss-Hermite quadrature to replace integrals by sums
  - Scale and translate quadrature locations to match the peak of the integrand using **adaptive quadrature**
- Maximum likelihood estimates obtained using Newton-Raphson
- Empirical Bayes (EB) predictions of latent variables and EB standard errors obtained using adaptive quadrature

Slide 20

Application

- Cluster randomized study of sex education in Norway
- Schools were randomized to receive sex education or not
- Assessments pre randomization, 6 months and 18 months post randomization
- Three ordinal outcomes (5-point scale) measuring 'contraceptive self-efficacy':
  "If my partner and I were about to have intercourse without either of us having mentioned contraception ...
  - I would have no problems saying that I have no contraception"
  - I would have no problems asking my partner whether he/she has contraception"
  - it would be easy for me to produce a condom (if I brought one)"
- 46 schools and 1183 pupils contributed to the analysis

Slide 21

Measurement and structural models

- **Measurement model** for item i at occasion t for student j in school k:

$$y^*_{itjk} = \delta_i + \lambda_i \eta^{(2)}_{tjk} + \epsilon_{itjk}, \quad \delta_1 = 0,\ \lambda_1 = 1$$

where

$$y_{itjk} = \begin{cases} 1 & \text{if} \quad y^*_i \le \kappa_1 \\ 2 & \text{if} \quad \kappa_1 < y^*_i \le \kappa_2 \\ 3 & \text{if} \quad \kappa_2 < y^*_i \le \kappa_3 \\ 4 & \text{if} \quad \kappa_3 < y^*_i \le \kappa_4 \\ 5 & \text{if} \quad \kappa_4 < y^*_i, \end{cases}$$

This is a constrained version of a 'graded response model' (Samejima, 1969).

- **Structural model** with covariates [Time] x_{1t}, [Treat] x_{2k} and [Treat]×[Time] x_{3tk}:

$$\eta^{(2)}_{tjk} = \beta_1 x_{1t} + \beta_2 x_{2k} + \beta_3 x_{3tk} + \eta^{(3)}_{jk} + \eta^{(4)}_k + \zeta^{(2)}_{tjk}$$

This is a three-level random intercept model for 'contraceptive self-efficacy'.

Slide 22

Missingness model

- The probability that student j fails to complete the questionnaire on occasion t ($d_{tjk} = 1$) is modeled as a logistic regression:

$$d_{tjk} = \begin{cases} 0 & \text{if } d^*_{tjk} \le 0 \\ 1 & \text{if } d^*_{tjk} > 0 \end{cases}$$

The latent response (propensity not to complete the questionnaire) depends on the contemporaneous self efficacy as well as self efficacy on previous occasions:

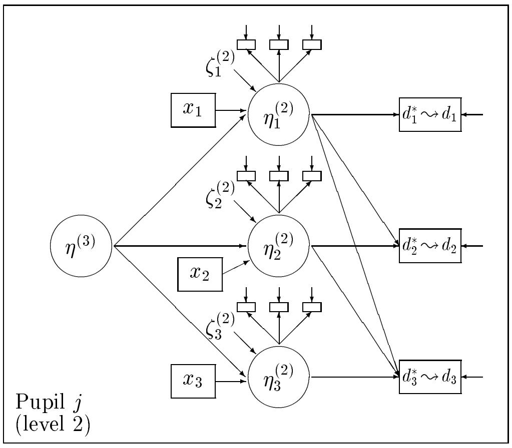
$$\begin{aligned} d^*_{1jk} &= \gamma + \alpha_0 \eta^{(2)}_{1jk} + \epsilon_{1jk} \\ d^*_{2jk} &= \gamma + \alpha_0 \eta^{(2)}_{2jk} + \alpha_1 \eta^{(2)}_{1jk} + \epsilon_{2jk} \\ d^*_{3jk} &= \gamma + \alpha_0 \eta^{(2)}_{3jk} + \alpha_1 \eta^{(2)}_{2jk} + \alpha_2 \eta^{(2)}_{1jk} + \epsilon_{3jk} \end{aligned}$$

This is analogous to the models by Hausman & Wise (1979) and Diggle & Kenward (1994), but missingness depends on latent variables $\boldsymbol{\eta}$ instead of the observable responses $\mathbf{y}$.

Slide 23

Path diagram

- The school variance was estimated as nearly zero; therefore school effects were omitted



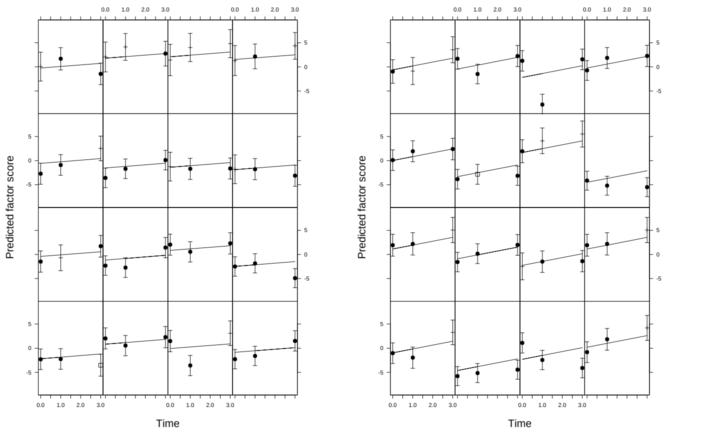
Slide 24

	Model 1		Model 2		Model 3		Model 4	
	Est	(SE)	Est	(SE)	Est	(SE)	Est	(SE)
Missingness model								
γ	-1.99	(0.26)	-2.10	(0.29)	-0.85	(0.04)	-1.09	(0.06)
α_0 [lag 0]	0.77	(0.09)	0.64	(0.09)	–		–	
α_1 [lag 1]	-0.10	(0.04)	–		–		-0.07	(0.03)
α_2 [lag 2]	-0.20	(0.05)	–		–		-0.25	(0.04)
Structural model								
β_1 [Time]	0.32	(0.10)	0.45	(0.09)	-0.06	(0.09)	-0.39	(0.09)
β_2 [Treat]	-0.91	(0.26)	-0.25	(0.23)	-0.28	(0.24)	-1.46	(0.21)
β_3 [Treat]×[Time]	**0.48**	**(0.11)**	**0.34**	**(0.10)**	**0.20**	**(0.11)**	**0.56**	**(0.11)**
$\text{var}(\zeta_{tjk}^{(2)})$	6.74	(0.60)	7.29	(0.68)	4.57	(0.41)	5.04	(0.45)
$\text{var}(\eta_{jk}^{(3)})$	5.30	(0.57)	4.29	(0.98)	3.72	(0.43)	3.51	(0.39)
Measurement model			Not shown					
log-likelihood	-8624		-8632		-8680		-8658	

Slide 25

Factor scores and standard errors

Control group Intervention group



Lines: $\widehat{\beta}_1 x_t + \widehat{\beta}_2 x_k + \widehat{\beta}_3 x_{tk} + \widetilde{\eta}^{(3)}_{jk}$, Points: $\widetilde{\eta}^{(2)}_{tjk} = \widehat{\beta}_1 x_t + \widehat{\beta}_2 x_k + \widehat{\beta}_3 x_{tk} + \widetilde{\eta}^{(3)}_{jk} + \widetilde{\zeta}^{(2)}_{tjk}$.

• all three items, □ one or two questions, − no questions.

Slide 26

Alternative interpretations of missingness/dropout modeling

- Bias correction in structural model (overly optimistic?)
- Sensitivity analysis

Slide 27

Some links and references

- gllamm and a manual can be downloaded from www.iop.kcl.ac.uk/iop/departments/biocomp/programs/gllamm.html

- Rabe-Hesketh, S., Skrondal, A. & Pickles, A. (2003). Generalized multilevel structural equation modelling. *Psychometrika*, in press.
- Skrondal, A. & Rabe-Hesketh, S. (2003). Mulitlevel logistic regression for polytomous data and rankings *Psychometrika*, in press.
- Rabe-Hesketh, S., Skrondal, A. & Pickles, A. (2002). Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal, 2*, 1–21.
- Rabe-Hesketh, S., Skrondal, A. & Pickles, A. (2003). Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *Submitted for publication.*
- Skrondal, A. & Rabe-Hesketh, S. (2003). *Generalized latent variable modeling: Multilevel, longitudinal and structural equation models.* Boca Raton, FL: Chapman & Hall/ CRC, to appear