Latent process model for multivariate heterogeneous longitudinal data: application to cognitive aging

Cécile Proust-Lima & Hélène Jacqmin-Gadda

Department of Biostatistics, INSERM U897, University of Bordeaux 2

INSERM workshop 205 - june 2010 - Saint Raphael

Cognitive aging in the elderly

Dementia characterized by a progressive and continuous decline of cognitive functions

→ heterogeneous cognitive aging : normal/ pathological

Cognition : latent process defined in continuous time

→ interest in the evolution of this quantity

Psychometric tests : noisy measures of cognitive functions

→ collected in discrete times

→ usually one test as a reference marker of cognition

→ specific metrological properties (ceiling/floor effects,...)

Objective

Describe the different profiles of cognitive decline associated with dementia in the elderly

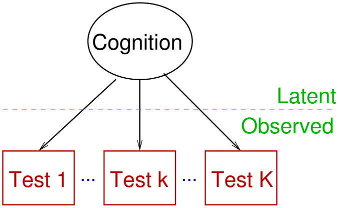
2 statistical problems addressed :

1. multiple markers of cognition (& different properties)
   $\longrightarrow$ nonlinear latent process model
2. heterogeneity of the declines & association with dementia
   $\longrightarrow$ joint latent class model

$\Rightarrow$ Joint latent class model for multivariate longitudinal data

Latent variable modelling (LVM)

Interest in a latent variable ("construct") measured by outcomes



ex1 :
cognition
measured by psychometric tests
ex2 :
arithmetic reasonning measured
by multi-item questionnaire

Principle :

- Structural equations : latent variable described according to covariates, time, etc
- Measurement model : link between the latent quantity and the outcomes

LVM in longitudinal settings

Latent process rather than latent variables defined at each time

→ linear mixed model for the latent process

Different types of outcomes

- Quantitative outcomes :

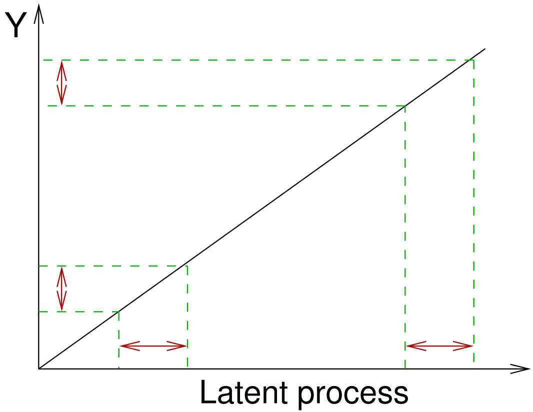
  standard : Gaussian (Roy, 2000)

  asymetric scales : non Gaussian (Proust, 2006)

- Ordinal outcomes :

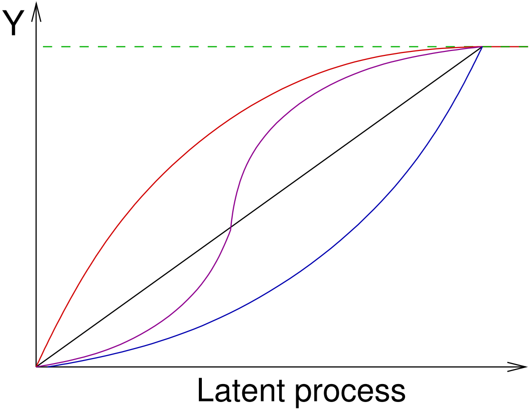
  threshold models (probit (Liu, 2006) ; proportional odds (Hambleton, 1991))

Symetric quantitative outcome



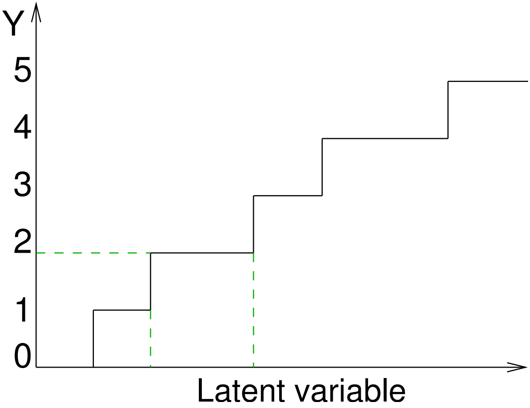
→ *Same sensitivity at each level*

Asymetric quantitative outcome



→ *Varying sensitivity depending on the level*

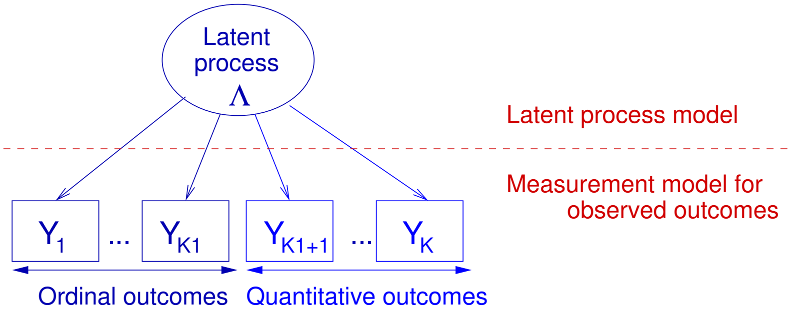
Ordinal outcome



→ *Range of latent process values for a given test value*

Structural model for the latent process

Notations : subject i, occasion j, outcome k



$$\Lambda_i(t) = \mathbf{X_{1i}}(t)^T \boldsymbol{\beta} + \mathbf{Z_i}(t)^T \mathbf{u_i}, t \geq 0$$

With $u_i \sim MVN(\mu, D)$ and identifiability constraints $u_{i0} \sim N(0,1)$

Measurement models for ordinal outcomes

Intermediate variable $\tilde{y}$ (with error, outcome-specific effects,...) :

$$\tilde{y}_{ijk} = \Lambda_i(t_{ijk}) + \mathbf{X_{2i}}(t)^T \gamma_k + \alpha_{ik} + \epsilon_{ijk}$$

with outcome-specific random intercept $\alpha_{ik} \sim N(0, \sigma_{\alpha_k})$

- Ordinal/binary outcome Y_k with C_k levels :

$$Y_{ijk} = c \quad \Leftrightarrow \quad \eta_{ck} \leq \tilde{y}_{ijk} < \eta_{(c+1)k} \text{ with } c \in \{0, C_k - 1\}$$

$\rightarrow$ constraints : $\eta_{0k} = -\infty$ and $\eta_{C_k k} = +\infty$

$\rightarrow$ Cumulative probit with Gaussian ϵ_{ijk}
and proportional odds model with logistic ϵ_{ijk}

Measurement models for quantitative outcomes

- Gaussian outcomes Y_k :

$$\frac{Y_{ijk} - \eta_{1k}}{\eta_{2k}} = \tilde{y}_{ijk}$$

- Non Gaussian quantitative outcomes Y_k :

$$H_k(y_{ijk}; \boldsymbol{\eta}) = \frac{h_k(y_{ijk}; \eta_{1k}; \eta_{2k}) - \eta_{3k}}{\eta_{4k}} = \tilde{y}_{ijk}$$

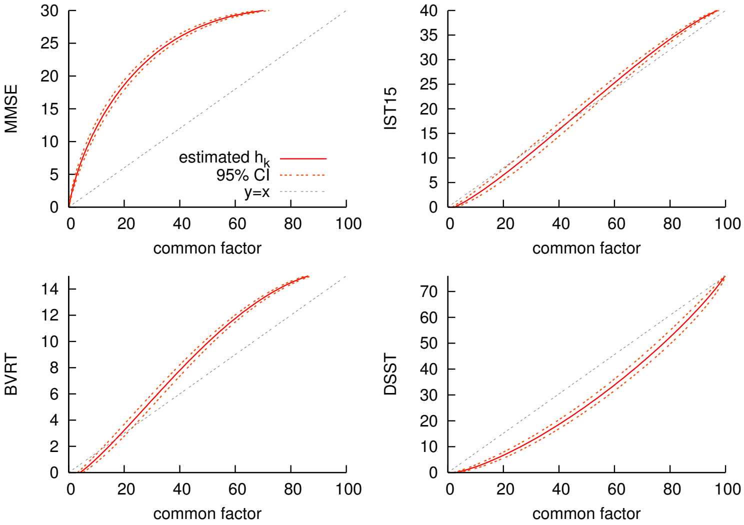
$\rightarrow h_k$ = CDF Beta (Proust, Bcs, 2006 ; Proust-Lima, CSDA 2009)

($h_k(.; 1, 1)$ = Identity $\Leftrightarrow$ special case for Gaussian outcomes)

$\rightarrow h_k$ = approximated by splines ...

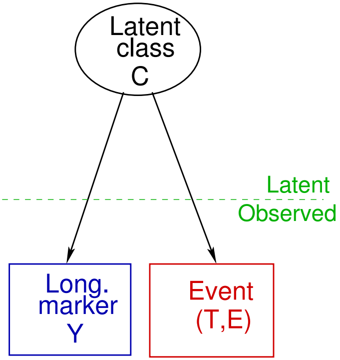
Estimated transformations for 4 psychometric tests

(Proust-Lima et al., AJE, 2007)



Joint latent class model (Lin et al., JASA, 2002)

With a single marker,



- Latent classes of subjects :

$\rightarrow$ latent class membership :

$$\pi_{ig} = P(c_i = g | X_{1i}) = \frac{e^{\xi_{0g} + X_{1i}{}^T \xi_{1g}}}{\sum_{l=1}^{G} e^{\xi_{0l} + X_{1i}{}^T \xi_{1l}}}$$

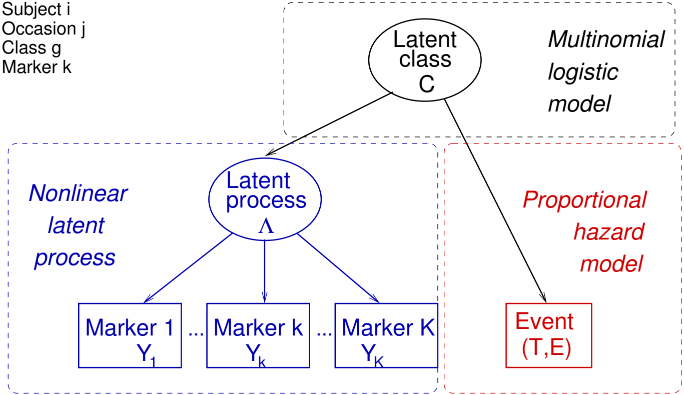
- Given class g,

$\rightarrow$ specific marker evolution

$\rightarrow$ specific risk of event

Extension to heterogeneous population : JLCM

Subject i
Occasion j
Class g
Marker k



$$\Lambda_i(t) \mid_{c_i=g} = Z_i(t)^T u_{ig} + X_{2i}(t)^T \beta_g$$ ← heterogeneous mixed model

← constraints : $u_{0i1} \sim N(0,1)$

$$Y_{ijk} \mid \Lambda_i(t_{ijk}, c_i = g),$$ ← marker-specific observation equation

Individual contribution to the likelihood

For a given number of classes G, the individual contribution is :

$$L_i(\boldsymbol{\theta}) = \sum_{g=1}^{G} \pi_{ig}(\boldsymbol{\theta}) \times f(\mathbf{y_i}|c_i = g; \boldsymbol{\theta}) \times \lambda(T_i|c_i = g; \boldsymbol{\theta})^{E_i} S(T_i|c_i = g; \boldsymbol{\theta})$$

with $f(\mathbf{y_i}|c_i = g; \boldsymbol{\theta})$:

- closed form for quantitative outcomes *(jacobian)*
- multivariate numerical integral over u_{ig} & α_{ik} for ordinal outcomes

Maximum likelihood estimators

- Log-likelihood $l(\psi) = \sum_{i=1}^{N} ln(L_i)$ maximised by a Marquardt algorithm

- Estimation achieved for a fixed number of latent classes G

& G selected using the Bayesian Information Criterion (BIC)

- Program in Fortran90/ R function in progress ...

Posterior classification

2 *posterior* class-membership probabilities :

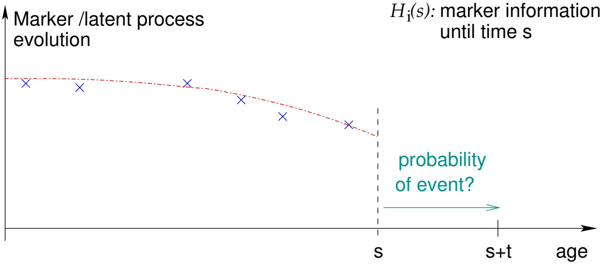
$$\hat{\pi}_{ig}^{y,T} = P(c_i = g \mid y_i, (T_i, E_i), x_i; \hat{\theta})$$

→ *used to assess the goodness-of-fit*

$$\hat{\pi}_{ig}^{y} = P(c_i = g \mid y_i, x_i; \hat{\theta}) = \frac{P(c_i = g \mid x_i; \hat{\theta}) f(y_i \mid c_i = g, x_i; \hat{\theta})}{\sum_{l=1}^{G} P(c_i = l \mid x_i; \hat{\theta}) f(y_i \mid c_i = l, x_i; \hat{\theta})}$$

→ *used for prognostic tools*

Prediction : prognostic /early detection tools



Predicted probability of event in (s,s+t) :

$$P(T_i \leq s+t \mid T_i > s, \mathcal{H}_i(s), X_i; \hat{\theta}) =$$

$$= \sum_{g=1}^{G} P(T_i \leq s+t \mid c_i = g, T_i > s, X_i; \hat{\theta}) \times \underbrace{P(c_i = g \mid \mathcal{H}_i(s), X_i, T_i > s; \hat{\theta})}_{\hat{\pi}_{ig}^{ys}}$$

Profiles of semantic memory decline associated with onset of Alzheimer's disease (AD) in the elderly

- Longitudinal outcomes : 2 measures of semantic memory

$\rightarrow$ 1 ordinal similarities test (WST- scale 0-10)

$\rightarrow$ 1 discrete quantitative fluency test (IST- scale 0-40)

- Time-to-event : age at onset of AD

$\rightarrow$ truncated data : entry in the cohort at age$>$65

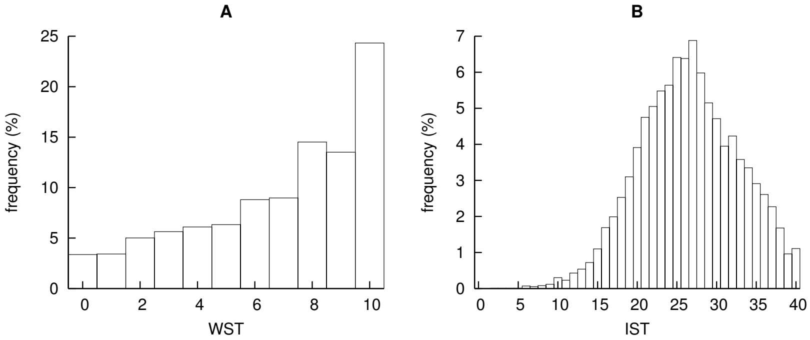
- Binary covariates : education, gender

- Subsample from a French cohort on aging (PAQUID) : N=2484

$\rightarrow$ followed-up during 14 years

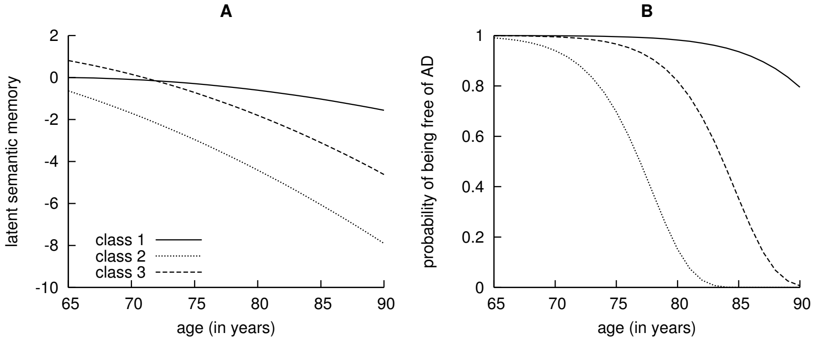
$\rightarrow$ 417 (16.8%) incident AD

Distribution of the tests



→ Median of 3 (IQR=[1,5]) repeated measures for IST
→ Median of 4 (IQR=[2,6]) repeated measures for WST

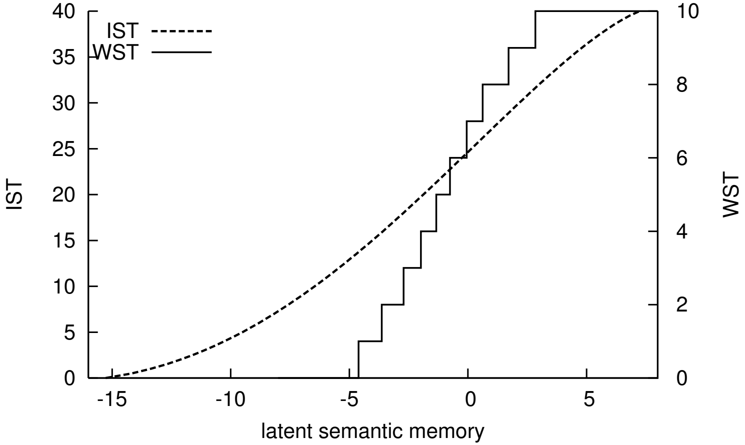
Predicted mean evolution of the latent process and probability of being free of dementia



Predicted mean evolution of the latent process in each class

Predicted probability of being free of AD in each class

Predicted transformations of the markers



Goodness-of-fit : class-specific marginal predictions 1

- For an ordinal outcome ($k = 1, ..., K_1$) :

$$\hat{y}_{ijk}|_{c_i=g} = E(y_{ijk}|\hat{\theta}; c_i = g) = \sum_{l=0}^{C_k-1} l \times P(\eta_{lk} \leq \tilde{y}_{ijk} < \eta_{(l+1)k}|\hat{\theta}; c_i = g)$$

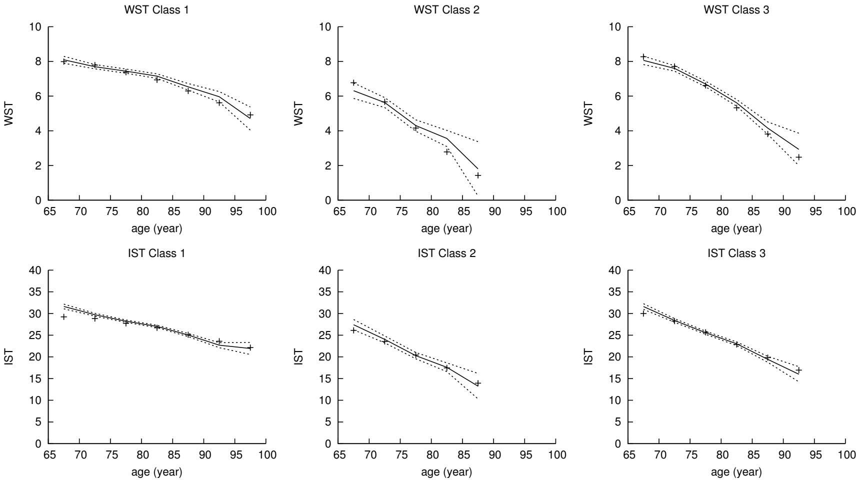
$$= C_k - 1 - \sum_{l=0}^{C_k-2} P(\tilde{y}_{ijk} < \eta_{(l+1)k}|\hat{\theta}; c_i = g)$$

- For a quantitative outcome ($k = K_1 + 1, ..., K$) :

$$\hat{y}_{ijk}|_{c_i=g} = E(H_k^{-1}(\tilde{y}_{ijk}; \hat{\eta}_k)|\hat{\theta}; c_i = g)$$

$\rightarrow$ numerical integration of $h_k^{-1}(\tilde{y}_{ijk}; \hat{\eta}_k)$ over the multivariate Gaussian distribution of $\tilde{y}_{ik}|_{c_i=g}$.

Class-specific marginal predictions in the test scale

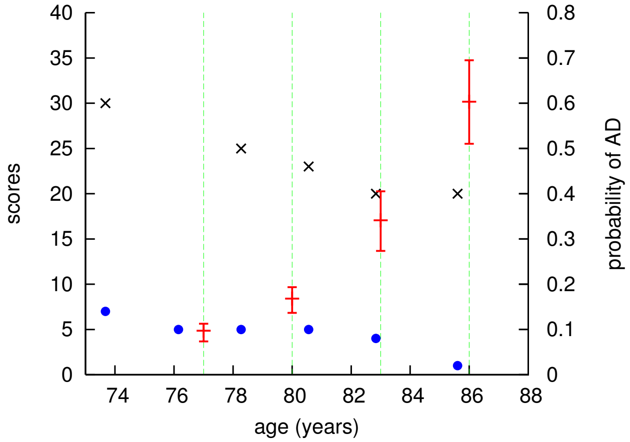


Goodness-of-fit : table of posterior classification

Final classif.	Number of subjects (%)	Mean of the class-membership probabilities in class : 1	2	3
1	2074 (83.5%)	**82.9**	2.3	14.8
2	142 (5.7%)	8.7	**78.3**	13.0
3	268 (10.8%)	18.5	7.7	**73.8**

→ unambiguous *posterior* classification

Dynamic predictive tool of AD



Probability of dementia in 5 years updated every 3 years

Diagnosed at

87 years old

x IST
● WST
+ Prediction with 95%CI

Concluding remarks

Advantages of the model :

- several markers *(latent process part)*
  - → avoids biases due to nonlinearity + ordinal scales
  - → increases the power of the analyses
- time-to-event *(joint model part)*
  - → avoids the selection biases
- latent class approach
  - → explicit interpretation of the association + heterogeneity

Possible applications :

- describe the natural history of a disease
- evaluate risk factors, treatments, ...
- develop tools for early detection/prognosis

References

- Hambleton R., Swaninathan H. & Rogers H. (1991). *Fundamentals of item response theory.* Newbury Park, CA : Sage.
- Lin H., Turnbull B.W. et al. (2002). Latent class models for joint analysis of longitudinal biomarker and event process data : application to longitudinal prostate-specific antigen readings and prostate cancer, *Journal of the American Statistical Association*, **97**,53-65
- Proust C., Jacqmin-Gadda H. et al. (2006). A nonlinear model with latent process for cognitive evolution using multivariate longitudinal data, *Biometrics*, **62**,1014-24
- Proust-Lima C., Amieva H. et al. (2007). Properties of 4 psychometric tests to measure cognitive changes in brain aging population based-studies, *American Journal of Epidemiology*, **165**, 344-50
- Proust-Lima C., Joly P. et al. (2008). Joint modelling of multivariate longitudinal outcomes & a time-to-event : a nonlinear latent class approach, *Computational Statistics & Data Analysis*, **53**, 1142-54
- Roy J., & Lin X. (2000). Latent variable models for longitudinal data with multiple continuous outcomes. *Biometrics*, **56**, 1047-1054.