Les limites du modèle linéaire mixte dans l'étude longitudinale des échelles ordinales et quantitatives bornées

Cécile Proust-Lima

Département de Biostatistique, INSERM U897, Université de Bordeaux 2

en collaboration avec
Hélène Jacqmin-Gadda *(Département de Biostatistique)*
& Hélène Amieva *(Département de Neuropsychologie)*

EPICLIN - 5 mai 2011 - Marseille

Analyse des données répétées d'un marqueur

Modèle linéaire mixte (MLM) utilisé couramment pour :

- décrire le changement au cours du temps d'un marqueur longitudinal
- évaluer l'effet de variables explicatives sur la trajectoire du marqueur

Hypothèses du MLM :

- la variable d'intérêt (marqueur longitudinal) est continue
- les erreurs et les effets aléatoires sont Gaussiens
- les variables (explicatives et fonctions du temps) sont associées à un changement constant estimé du marqueur

Problème avec les échelles de mesure

Exemples d'échelles de mesure :

- somme d'items mesurant la qualité de vie
- échelle mesurant le niveau de dépendance
- test psychométrique mesurant le niveau cognitif

Echelles de mesure : marqueurs quantitatifs discrets bornés ou ordinaux

- mesures avec erreur du processus sous-jacent d'intérêt (processus biologique/psychologique)
- effets plafonds et planchers
- sensibilité variable au changement = curvilinéarité

Objectif

Evaluer l'impact de la non prise en compte

des propriétés métrologiques (curvilinéarité)

des échelles de mesure dans l'évaluation

des déterminants de leurs trajectoires ?

Modèle linéaire mixte classique (MLM)

Soit $Y_i(t)$, marqueur mesuré au temps t ($t \in \mathbb{R}$) pour le sujet i ($i = 1, ..., N$) :

$$Y_i(t) = \beta_0 + \beta_1 t + \beta_2 X_i + \beta_3 X_i \times t + u_{0i} + u_{1i} t + \epsilon_i(t)$$

avec X_i, variable explicative pour le sujet i et β_2 & β_3 ses effets sur Y

u_{0i} & u_{1i}, intercept et pente aléatoires Gaussiens centrés de matrice de variance covariance B

$\epsilon_i(t)$, erreur de mesure Gaussienne centrée de variance σ_ϵ^2

Modèle linéaire mixte classique (MLM)

Soit $Y_i(t)$, marqueur mesuré au temps t ($t \in \mathbb{R}$) pour le sujet i ($i = 1, ..., N$) :

$$Y_i(t) = \beta_0 + \underbrace{\beta_1 t + \beta_2 X_i + \beta_3 X_i \times t + u_{0i} + u_{1i} t} + \epsilon_i(t)$$

$$Y_i(t) = \beta_0 + \sigma_0 \times \Lambda_i(t) + \epsilon_i(t)$$

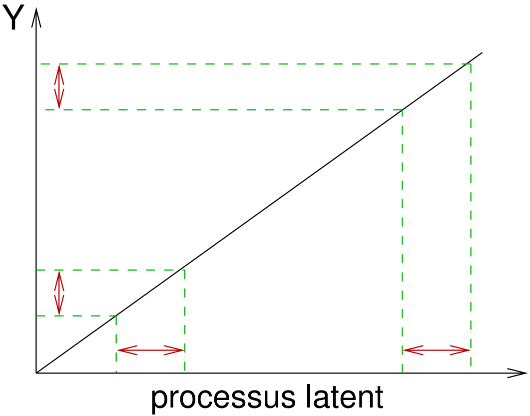
avec X_i, variable explicative pour le sujet i et β_2 & β_3 ses effets sur Y

u_{0i} & u_{1i}, intercept et pente aléatoires Gaussiens centrés de matrice de variance covariance B

$\epsilon_i(t)$, erreur de mesure Gaussienne centrée de variance σ_ϵ^2

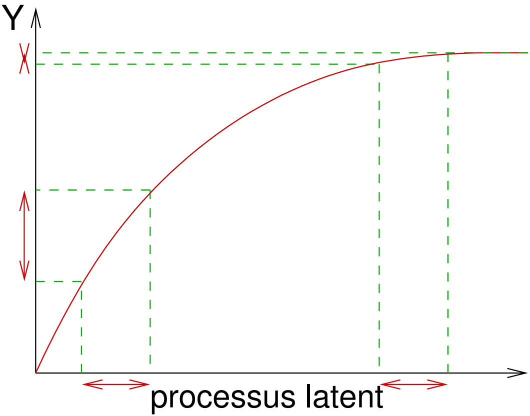
$\Lambda_i(t)$ processus latent au temps t ou quantité sous-jacente d'intérêt

Transformations linéaires



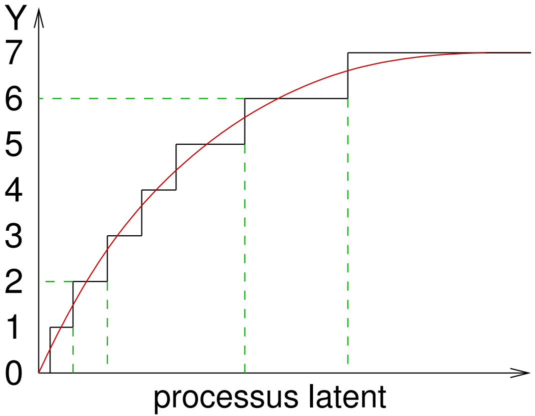
→ *Même sensibilité dans tout le range de Y*

Transformations nonlinéaires



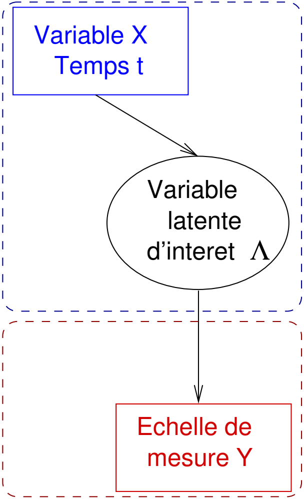
→ *Sensibilité variable + effets plafond/plancher*

Transformation en escalier pour marqueur ordinal



$\longrightarrow$ *Intervalle de valeurs de* Λ *pour une valeur de* Y *donnée*

Principe général du modèle mixte à processus latent



$\leftarrow$ Modèle linéaire mixte standard

$$\Lambda_i(t) = \beta_1 t + \beta_2 X_i + \beta_3 X_i \times t + u_{0i} + u_{1i} t$$

$\leftarrow$ Equation d'observation

$$H(Y_i(t);\eta) = \Lambda_i(t) + \epsilon_i(t)$$

avec $H(\ ;\eta) =$ linéaire, FdR Beta,
I-splines, seuils, etc

Evaluation des facteurs de risque de déclin cognitif

4 tests psychométriques :

- Mini Mental State Examination (MMSE) [0-30]
- Isaacs Set Test (IST) [0-40]
- Benton Visual Retention Test (BVRT) [0-15]
- Sous-score de calcul du MMSE (CALC) [0-5]

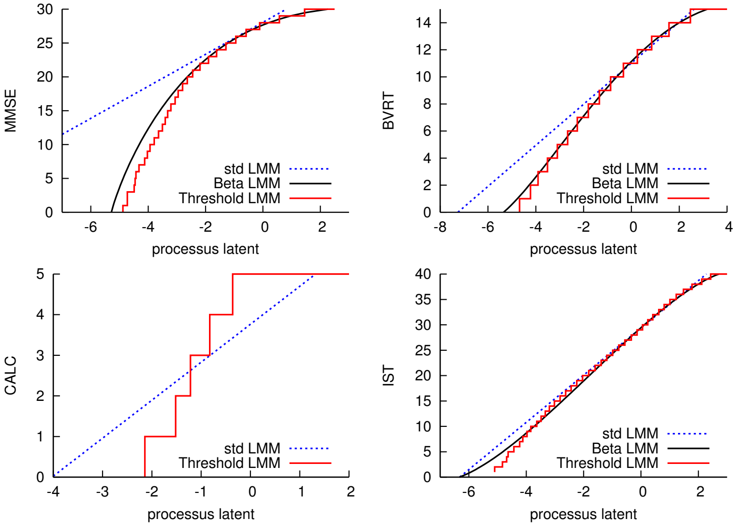
1 variable explicative : certificat d'études primaires (CEP - 1/0)

Plusieurs modèles mixtes :

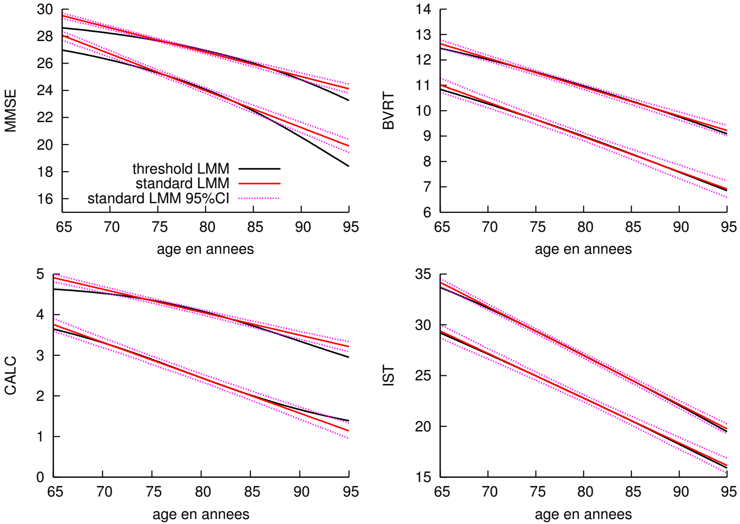
- Modèle linéaire mixte standard (linéaire/standard)
- Modèle mixte avec transformation continue estimée (FdR Beta)
- Modèle mixte avec transformation à seuils (ordinal/threshold)

$\rightarrow$ *Trajectoire individuelle linéaire en fonction de l'âge et effet du CEP*

Transformations nonlinéaires estimées



Trajectoires prédites selon certificat d'études primaires



Effet du CEP sur le déclin cognitif

Test psych.	Transformation dans le MLM	Effet du CEP $\hat{\beta}$	Effet du CEP p	Effet du CEP $\times$ t $\hat{\beta}$	Effet du CEP $\times$ t p
MMSE	linéaire	0.618	(<0.0001)	0.389	(<0.0001)
	FdR Beta	1.026	(<0.0001)	0.114	(0.034)
	ordinal	1.000	(<0.0001)	0.092	(0.067)
BVRT	linéaire	1.067	(<0.0001)	0.151	(0.029)
	FdR Beta	1.033	(<0.0001)	0.03	(0.618)
	ordinal	1.04	(<0.0001)	0.02	(0.621)
IST	linéaire	1.031	(<0.0001)	-0.085	(0.0836)
	FdR Beta	1.038	(<0.0001)	-0.131	(0.0050)
	ordinal	1.017	(<0.0001)	-0.148	(0.0008)
CALC	linéaire	1.206	(<0.0001)	0.321	(<0.0001)
	ordinal	1.09	(<0.0001)	0.037	(0.453)

Simulations : Erreur de 1$^{\text{ère}}$ espèce α dans le MLM

8 scenarios avec 4 distributions de tests & 2 effets de variables explicatives

- Modèle à seuils choisi comme “vrai” modèle
- 500 répliques d’échantillons de 500 sujets

Résultats pour le MLM standard (valeur nominale : $\alpha = 5\%$) :

- distribution proche du MMSE : $\hat{\alpha}$ jusqu’à 93.2%
- distribution proche du BVRT : $\hat{\alpha}$ jusqu’à 21.4%
- distribution proche du CALC : $\hat{\alpha}$ jusqu’à 45.6%
- distribution proche du IST : $\hat{\alpha}$ jusqu’à 8.0%

Résultats pour le MLM avec FdR Beta : tous les $\hat{\alpha} \approx$ 5%

Conclusions

Curvilinéarité : propriété intrinsèque des échelles de mesure

$\Leftrightarrow$ sensibitilité variable au changement

- indépendante des relations avec variables/populations/etc
- potentiellement valable pour toutes les échelles

Ne pas en tenir compte dans le MLM implique :

- de potentielles fausses associations
- une confusion entre les effets sur l'intercept et la pente

$\longrightarrow$ mais difficile à diagnostiquer à partir du fit du MLM

Utilisation des modèles mixtes nonlinéaires recommandée :

- fonction **`lcmm`** du package R **`lcmm`**
- plus de détails : Proust-Lima et al., AJE - *révision mineure*

Transformations du MMSE dans PAQUID