Inference on random changepoint models: application to pre-dementia cognitive decline

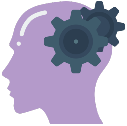
Corentin Segalas[1], Hélène Jacqmin-Gadda[2]

[1]Medical Statistics Department, London School of Hygiene & Tropical Medicine

[2]Biostatistics Team, Bordeaux Population Health, Inserm U1219

Journées *Statistiques & Santé* 2020, October 2020

Context: pre-dementia cognitive decline



- Very long and progressive pre-diagnosis phase
- Heterogeneous and non-linear decline trajectories
- Subject-specific acceleration of cognitive decline

Context: pre-dementia cognitive decline trajectories

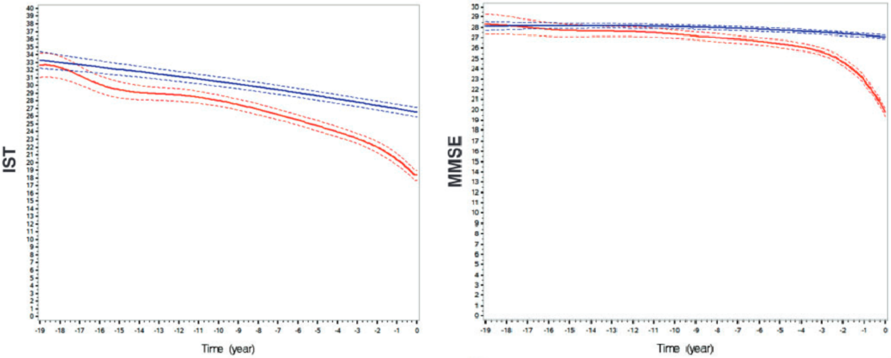


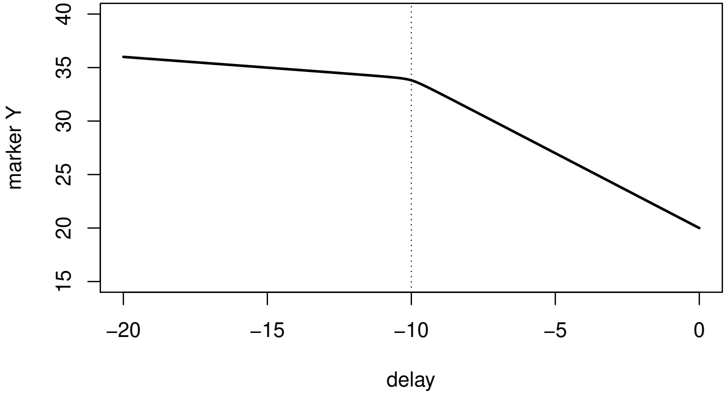
Figure: Estimated cognitive trajectories for cases (red) and matched controls (blue) for high educational subjects from French cohort PAQUID (Amieva et al., 2014).

Project 1

Objective: Testing the existence of a random changepoint in a mixed model

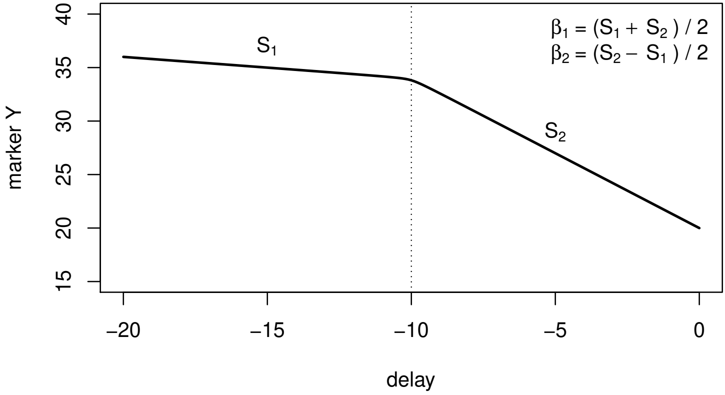
Segalas C, Amieva H, Jacqmin-Gadda H. A hypothesis testing procedure for random changepoint mixed models. *Statistics in Medicine*, 2019.
https://doi.org/10.1002/sim.8195

The random changepoint mixed model



The random changepoint mixed model

$$Y(t_{ij}) = \beta_{0i} + \beta_{1i} t_{ij} + \beta_2 \sqrt{(t_{ij} - \tau_i)^2 + \gamma} + \varepsilon_{ij}$$



The random changepoint mixed model

$$Y(t_{ij}) = \beta_{0i} + \beta_{1i} t_{ij} + \beta_2 \sqrt{(t_{ij} - \tau_i)^2 + \gamma} + \varepsilon_{ij}$$

- $\beta_{ki} = \beta_k + b_{ki}$ for $k = 0, 1$ with $b_i = (b_{0i}, b_{1i}) \sim \mathcal{N}(0, B)$
- $\tau_i = \mu_\tau + \sigma_\tau \tilde{\tau}_i$ with $\tilde{\tau}_i \sim \mathcal{N}(0, 1)$ and $\tilde{\tau}_i \perp b_i$
- $\sqrt{. + \gamma}$ a smooth transition function
- $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma)$ residual error $\perp$ of the random effects

At this stage β_2 is assumed non random

A score test approach

$$Y(t_{ij}) = \beta_{0i} + \beta_{1i} t_{ij} + \beta_2 \sqrt{(t_{ij} - \tau_i)^2 + \gamma} + \varepsilon_{ij}$$

- Objective: H_0: $\beta_2 = 0$ vs. H_1: $\beta_2 \neq 0$

A score test approach

$$Y(t_{ij}) = \beta_{0i} + \beta_{1i} t_{ij} + \beta_2 \sqrt{(t_{ij} - \tau_i)^2 + \gamma} + \varepsilon_{ij}$$

- Objective: H_0: $\beta_2 = 0$ vs. H_1: $\beta_2 \neq 0$
- Nuisance parameters: $\underbrace{\beta_0, \beta_1, \sigma, \sigma_0, \sigma_1, \sigma_{01}}_{\theta}, \mu_\tau, \sigma_\tau$
- Classic score test statistics depends upon $\hat{\mu}_{\tau 0}, \hat{\sigma}_{\tau 0}$

$$S_n(0; \hat{\mu}_{\tau 0}, \hat{\sigma}_{\tau 0}, \hat{\theta}_0) = \frac{U_n(0; \hat{\mu}_{\tau 0}, \hat{\sigma}_{\tau 0}, \hat{\theta}_0)^2}{Var(U_n(0; \hat{\mu}_{\tau 0}, \hat{\sigma}_{\tau 0}, \hat{\theta}_0))}$$

with

$$U_n(0, \mu_\tau, \sigma_\tau, \theta) = \left. \frac{\partial \ell_n(Y; \beta_2, \mu_\tau, \sigma_\tau, \theta)}{\partial \beta_2} \right|_{\beta_2 = 0} \text{ and } U_n = \sum_{i=1}^{n} u_i$$

The supremum score test (Hansen, 1996)

- Test statistic:

$$T_n = \sup_{(\mu_\tau, \sigma_\tau)} S_n(0; \mu_\tau, \sigma_\tau, \hat{\theta}_0)$$

with $\hat{\theta}_0$ MLE of identifiable nuisance parameters under H_0

The supremum score test (Hansen, 1996)

- Test statistic:

$$T_n = \sup_{(\mu_\tau, \sigma_\tau)} S_n(0; \mu_\tau, \sigma_\tau, \hat{\theta}_0)$$

with $\hat{\theta}_0$ MLE of identifiable nuisance parameters under H_0

- Empirical distribution of T_n under H_0: perturbation algorithm (van der Vaart et al., 1996). For $k = 1, \dots, K$, we generate n r.v. $\xi_i^{(k)} \sim \mathcal{N}(0,1)$ and compute

$$T_n^{(k)} = \sup_{(\mu_\tau, \sigma_\tau)} \frac{\left(\sum_{i=1}^n u_i(0; \mu_\tau, \sigma_\tau, \hat{\theta}_0)\xi_{\mathbf{i}}^{(\mathbf{k})}\right)^2}{\sum_{i=1}^n u_i(0; \mu_\tau, \sigma_\tau, \hat{\theta}_0)^2}$$

- Empirical $p-$value $p_K = \frac{1}{K}\sum_{k=1}^K \mathbf{1}_{T_n^{(k)} > T_n^{(obs)}}$

Additional tests for heterogeneity

Heterogeneity in β_2?

- Is β_2 subject specific (i.e. random)?

  $$H_0\colon\ B = \begin{pmatrix} \sigma_0^2 & \sigma_{01} & 0 \\ \sigma_{01} & \sigma_1^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ vs. } H_1\colon\ B \text{ unstructured}$$

  $\Rightarrow$ corrected test for variance components (Stram and Lee, 1994)

- Does β_2 depend on covariate?
  $\Rightarrow$ Wald test

Additional tests for heterogeneity

Heterogeneity in β_2?

- Is β_2 subject specific (i.e. random)?

$$H_0\colon\ B = \begin{pmatrix} \sigma_0^2 & \sigma_{01} & 0 \\ \sigma_{01} & \sigma_1^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ vs. } H_1\colon\ B \text{ unstructured}$$

$\Rightarrow$ corrected test for variance components (Stram and Lee, 1994)

- Does β_2 depend on covariate?
  $\Rightarrow$ Wald test

Heterogeneity in τ_i?

- Does τ_i depend on covariate?
  $\Rightarrow$ Wald test

Application: the PAQUID cohort

- cohort of 3777 elderly subjects ($\geq$ 65yo) from the French departments of Gironde and Dordogne, 25 years follow-up
- 901 incident cases of dementia between year 1 and 25
- Isaac 15s score (verbal fluency)
- Stratified analysis on the educational level

Application: results

	obs. statistic test	$p-$value
High education	143.7	<0.001
Low education	56.9	<0.001

Table: Score test results with $K = 500$

$\Rightarrow$ We clearly reject H_0: $\beta_2 = 0$ for both group

Application: results

	obs. statistic test	$p-$value
High education	143.7	<0.001
Low education	56.9	<0.001

Table: Score test results with $K = 500$

$\Rightarrow$ We clearly reject H_0: $\beta_2 = 0$ for both group

$$\beta_{2i} = \beta_2 + \alpha_{2i} \text{ with } \alpha_i = (\alpha_{0i}, \alpha_{1i}, \alpha_{2i}) \sim \mathcal{N}(0, B)$$

$$(H_0) : \mathrm{B} = \begin{pmatrix} \sigma_0^2 & \sigma_{01} & 0 \\ \sigma_{01} & \sigma_1^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ vs. } (H_1) : B \text{ unstructured}$$

$\Rightarrow$ We reject H_0: $\sigma_2 = 0$ for both group ($p < 0.001$)

Discussion

- Valid test with good power
- `testRCPM` function in `rcpm` package
- Assumption of a fixed β_2 (test with random β_{2i} robust)
- Relaxing the assumption of a Gaussian distribution for $\tilde{\tau}_i$

Project 2

Objective: Compare mean CP date between markers

Segalas C, Helmer C, Jacqmin-Gadda H. A curvilinear bivariate random changepoint model to assess temporal order of markers. *Statistical Methods in Medical Research*, 2020.
https://doi.org/10.1177/0962280219898719

The bivariate random changepoint mixed model

$$Y^\ell(t_{ij}^\ell) = \beta_{0i}^\ell + \beta_{1i}^\ell(t_{ij}^\ell - \tau_i^\ell) + \beta_{2i}^\ell\sqrt{(t_{ij}^\ell - \tau_i^\ell)^2 + \gamma} + \varepsilon_{ij}^\ell \quad \ell = 1, 2$$

- $\beta_{ki}^\ell = \beta_k^\ell + b_{ki}^\ell$ with $b_i^\ell = (b_{0i}^\ell, b_{1i}^\ell, b_{2i}^\ell) \sim \mathcal{N}(0, B^\ell)$
- $\tau_i^\ell = \mu_\tau^\ell + \sigma_\tau^\ell \tilde{\tau}_i^\ell$ with $\tilde{\tau}_i^\ell \sim \mathcal{N}(0,1)$ and $\tilde{\tau}_i^\ell \perp b_i$
- $\sqrt{. + \gamma}$ a smooth transition function
- $\varepsilon_{ij}^\ell \sim \mathcal{N}(0, \sigma^\ell)$ residual error $\perp$ of the random effects

$+$ $corr(b_i^1, b_i^2) = B^{12}$ and $corr(\tilde{\tau}_i^1, \tilde{\tau}_i^2) = \rho_\tau^{12} \Rightarrow$ bivariate model

Curvilinearity

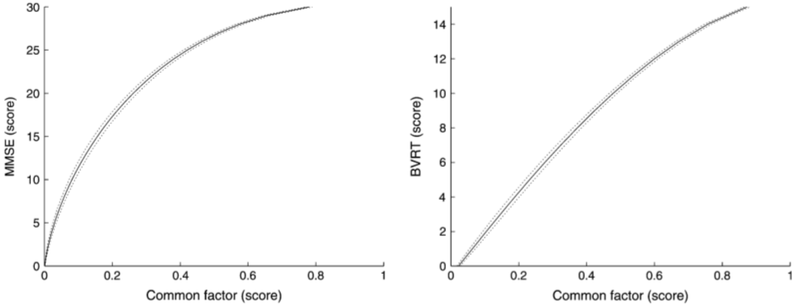


Figure: Estimated link function between crude score and the underlying latent process (Proust Lima et., 2006)

Curvilinearity

I-spline transformation of both crude markers Y^ℓ:

$$\tilde{Y}_{ij}^\ell = g^\ell(Y_{ij}^\ell, \eta^\ell) = \eta_0^\ell + \sum_{k=1}^{5} \eta_k^{\ell 2} I_k^\ell(Y_{ij}^\ell) \quad \ell = 1, 2$$

- I-splines of degree 2 with 2 internal knots at the quantiles
- $\tilde{Y} = (\tilde{Y}^1, \tilde{Y}^2)$ follows bivariate random changepoint model
- Identifiability constraints on the model: $\beta_0^\ell = 0$ and $\sigma_\epsilon^\ell = 1$

Inference

- Log-likelihood $\tilde{\tau}_i = (\tilde{\tau}_i{}^1, \tilde{\tau}_i{}^2)$:

$$\ell(\theta) = \sum_{i=1}^{n} \log \int f(\tilde{Y}_i|\tilde{\tau}_i) f(\tilde{\tau}_i) \mathrm{d}\tilde{\tau}_i + n \log |J_g^1||J_g^2|$$

where $\tilde{Y}_i|\tilde{\tau}_i$ is a multivariate Gaussian.

- Optimization: Levenberg-Marquardt algorithm and pseudo adaptive Gaussian quadrature
- Test: H_0: $\mu_\tau^1 - \mu_\tau^2 = 0$ vs. H_1: $\mu_\tau^1 - \mu_\tau^2 \neq 0$: a Wald test

Application: the Three City (3C) cohort

- cohort of 2104 elderly subjects ($\geq$ 65yo)
- 401 incident cases from Bordeaux center
- Grober and Bushke (GB) immediate vs. free recall

Application: the Three City (3C) cohort

- cohort of 2104 elderly subjects ($\geq$ 65yo)
- 401 incident cases from Bordeaux center
- Grober and Bushke (GB) immediate vs. free recall

Table: Results of the preliminary tests on the 3C sample.

	$\beta_2 = 0$ vs. $\beta_2 \neq 0$	$\sigma_2 = 0$ vs. $\sigma_2 \neq 0$
GB immediate recall	< 0.001	< 0.001
GB free recall	< 0.001	< 0.001

Application: results

Table: Results of the bivariate estimation on the 3C sample.

	GB immediate recall		GB free recall		Wald test	
	$\hat{\beta}$	$\widehat{se}(\hat{\beta})$	$\hat{\beta}$	$\widehat{se}(\hat{\beta})$	stat.	$p-$value
β_1	-0.286	0.023	-0.262	0.037	0.589	0.443
β_2	-0.230	0.022	-0.229	0.029	0.024	0.877
μ_τ	-3.177	0.347	-5.820	0.579	3.937	0.047

se: standard error

$\Rightarrow$ difference between GB immediate and free recall

Application: marginal estimation

$$E(\tilde{Y}^{\ell}(t), \hat{\theta}^{\ell}) = \int E(\tilde{Y}^{\ell}(t)|\tau_i^{\ell}, \hat{\theta}^{\ell}) f(\tau_i^{\ell}|\hat{\theta}^{\ell}) \mathrm{d}\tau_i^{\ell}$$

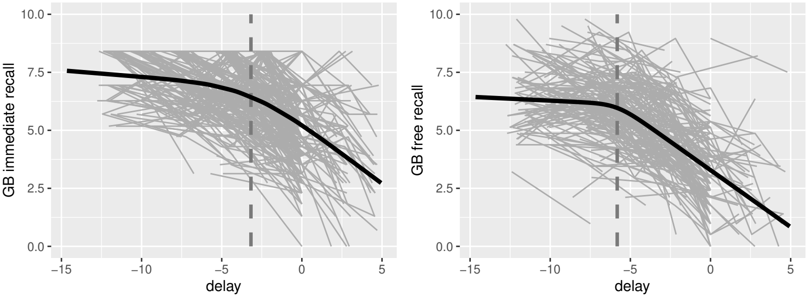


Figure: All individual GB immediate and free recall trajectories on the transformed scale compared to the estimated marginal trajectory $E(\tilde{Y}^{\ell}(t))$

Application: fit of the model

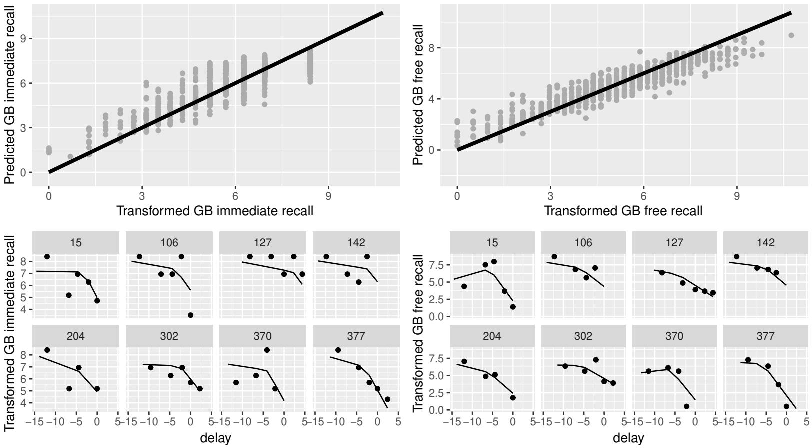


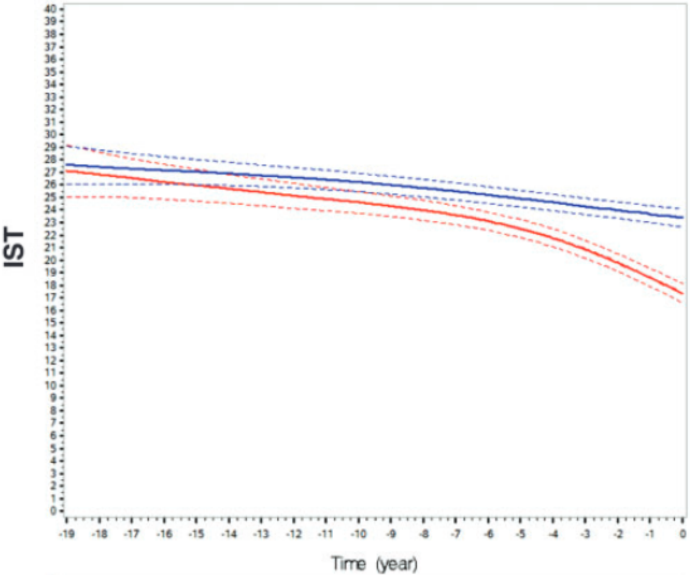
Figure: Upper panes: true transformed observation vs. predicted observations; Lower panes: individual observations (dots) vs. their predicted trajectories (solid line).

Discussion

- Valid estimation procedure and valid test
- `bircpme` function in `rcpm` package
- Identification of a late acceleration of cognitive decline
  $\Rightarrow$ modelling cases and controls together?

Perspectives & discussion

Time of differenciation versus late accelerated decline



A semi-latent class random changepoint model

$$Y(t_{ij}) = \beta_{0i} + \beta_{1i} t_{ij} + c_i \beta_{2i} f(t_{ij} - \tau_i, \eta) + \varepsilon_{ij}$$

A semi-latent class random changepoint model

$$Y(t_{ij}) = \beta_{0i} + \beta_{1i} t_{ij} + c_i \beta_{2i} f(t_{ij} - \tau_i, \eta) + \varepsilon_{ij}$$

- with a class membership model

$$\pi_i = \mathbb{P}(c_i = 1 | X_i, \delta_i) = \left(\frac{\exp(\eta^\top X_i)}{1 + \exp(\eta^\top X_i)} \right)^{1-\delta_i}$$

- δ_i case indicator (1 for cases, 0 for controls)

⇒ all cases have a changepoint
⇒ some controls have a changepoint

Discussion

- Selection bias: a joint model approach
  - the longitudinal marker $Y(t_{ij}) = \tilde{Y}(t_{ij}) + \varepsilon_{ij}$
  - the time to dementia: $\lambda(t_{ij}) = \lambda_0(t_{ij}) \exp(\nu^\top Z_i + \gamma \tilde{Y}(t_{ij}))$

$\Rightarrow$ possible to test for the existence of the random CP

- The timescale issue: age or delay?

- Random changepoint model vs. flexible nonlinear model

Thank you for your attention!

1. Amieva H. et al. "Compensatory mechanisms in higher-educated subjects with Alzheimer's disease: a study of 20 years of cognitive decline." Brain: A Journal of Neurology (2014)
2. Hansen B.E. "Inference When a Nuisance Parameter Is Not Identified under the Null Hypothesis." Econometrica (1996)
3. Stram D.O. and Lee J.W. "Variance Components Testing in the Longitudinal Mixed Effects Model." Biometrics (1994)
4. Marquardt D.W. "An Algorithm for Least-Squares Estimation of Nonlinear Parameters." Journal of the Society for Industrial and Applied Mathematics (1963)
5. Grober E. and Buschke H. "Genuine Memory Deficits in Dementia." Developmental Neuropsychology (1987)
6. van der Vaart A.W. and Wellner J.A. "Weak Convergence and Empirical Processes." Chapter 2.9, Springer Series in Statistics
7. Proust-Lima C. et al. "Sensitivity of Four Psychometric Tests to Measure Cognitive Changes in Brain Aging-Population-based Studies. American Journal of Epidemiology (2006)

https://github.com/crsgls