Role of Pathologies in the Long-Term Care process

Guillaume Biessy, PhD, SCOR Global Life SE

About the speaker

- **Guillaume Biessy**
- R&D actuary in SCOR LTCI international research center since 2013
- Member of the French Institute of Actuaries since 2013, PhD in applied mathematics from Paris-Saclay University defended in 2016.

- **SCOR**
- 4th largest reinsurance company
- Strongly supports research initiatives through a dedicated foundation, multiple research partnerships and the organization of actuarial prizes in several countries

Agenda

Context

Local likelihood

Study without pathologies

Study with pathologies

Take-away

Stakes for LTC insurance

Definition

LTC: **consolidated** and **permanent** state of incapacity to perform autonomously the **activities of daily living** (mobility, clothing, bathing and feeding).

- Phenomenon linked to ageing with multiple causes such as cancer and dementia
- Associated costs higher than 3,000 € / a month in France meaning people are forced to rely on private insurance products.

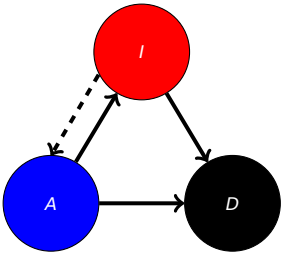
The private LTC market in France:

- Only exists since mid 1980's, Second market in the world
- **3.4 M insured** representing **600 M €** **of contributions** and **5 B €** **of reserve** in 2015

The illness-death model

Modelling based on a 3-state illness-death model.

Trajectories of insured lives represented through a continuous-time *càd-làg* process $(Z_u)_{u\geq 0}$ indexed by insured age u and taking its values in the set $E = \{A, I, D\}$ for autonomy, LTC (*impaired*) and death.



We work on the assumption of **no recovery from LTC**.

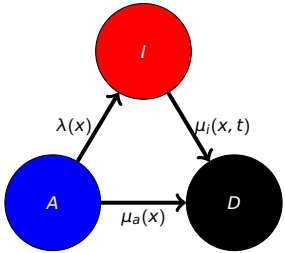
Transition intensities for the model

In a continuous-time framework, the model is defined through **transition intensities** (where x represents time spent in the autonomous state or current age for autonomous and t the time spent in LTC):

$$\mu_a(x) = \lim_{h \to 0} \frac{1}{h} \mathbb{P}\left(Z_{x+h} = D | Z_x = A\right),$$

$$\lambda(x) = \lim_{h \to 0} \frac{1}{h} \mathbb{P}\left(Z_{x+h} = I | Z_x = A\right),$$

$$\mu_i(x,t) = \lim_{h \to 0} \frac{1}{h} \mathbb{P}\left(Z_{x+t+h} = D | Z_{x^-} = A, Z_x = I, Z_{x+t} = I\right).$$



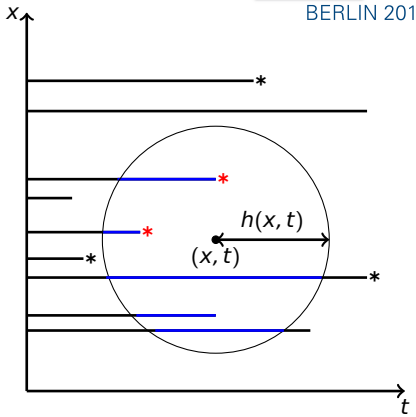
Data

Data comes from a gathering of French portfolios containing information about 200,000 insured and 25,000 claims.

Pathology	Number of claims made		Number of observed deaths	
	male	female	male	female
Dementia	3 088	8 460	2 405	5 602
Neurological diseases	1 628	2 657	1 338	1 898
Cancer	1 152	1 237	1 113	1 155
Other causes	1 561	3 722	1 276	2 469
Total	**7 429**	**16 076**	**6 132**	**11 124**

Local likelihood: a quick look

- **Inference / non-parametric smoothing** method
- **Aim:** infer **transition intensities** (functions of 1 or 2 variables)
- **Principle:** at **each point**, fit a **polynomial function** then compute its value at that point
- Coefficients of the polynomial functions estimated using **neighbouring observations**
- The further the observations are from the estimation point, the less they weigh
- Here, **continuous-time variant** applied directly to individual trajectories



Local likelihood: 1 dimension case

Let us assume trajectories of the form (x_i, y_i, c_i) where

- x_i is the **age at the beginning of the observation period**,
- y_i is the **age at the end of the observation period**,
- c_i is a **censoring index**: $c_i = 1$ if the event of interest was **not observed**.

For u close to x, $d \in \mathbb{N}$, there exists $a = (a_0, a_1, \ldots, a_d)^T$ such that

$$\begin{aligned}\log \mu(u) &= a_0 + a_1(u-x) + \cdots + a_d(u-x)^d + o((u-x)^d) \\ &= \langle a, A(u-x) \rangle + o((u-x)^d)\end{aligned}$$

where $A(u) = (1, u, \ldots, u^d)^T$.

Local likelihood: 1 dimension case

Let us begin by introducing the **local log-likelihood function**

$$\mathcal{L}_x(a) = \sum_{i=1}^{n} (1-c_i) W\left(\frac{y_i - x}{h(x)}\right) \langle a, A(y_i - x)\rangle - \int_{x_{min}}^{y_{max}} N(u) W\left(\frac{u-x}{h(x)}\right) e^{\langle a, A(u-x)\rangle} \, du$$

where

- $N(u) = \sum_{i=1}^{n} \mathbf{1}\{x_i < u \leq y_i\}$ is the **number of individuals under risk** at age u,
- $x_{min} = \min_{i \in \{1,\ldots,n\}} x_i$ and $y_{max} = \max_{i \in \{1,\ldots,n\}} y_i$,
- W is the **kernel**,
- h is the **bandwidth**.

Finally an estimator $\hat{\mu}(x)$ of $\mu(x)$ is given by $\hat{\mu}(x) = \exp(\hat{a}_0)$ where

$$(\hat{a}_0, \ldots, \hat{a}_d)^T = \underset{a}{\operatorname{argmax}} \, \mathcal{L}_x(a)$$

Local likelihood: meta-parameters

The local-likelihood relies on the following meta-parameters:

- The **degree** d of the adjusted polynomial function: higher degrees allow for more **flexibility** at the expense of higher **over-fitting risk**.
- The **kernel** W. In what follows we use **Epanechnikov** kernel

$$W(u) = \mathbf{1}\{|u| < 1\}\left(1 - u^2\right)^2.$$

- The **bandwidth** h. The bandwidth $h_\alpha(x)$ is here chosen so that α uncensored observations are no further than $h_\alpha(x)$ from x.

Local likelihood: model choice and residuals

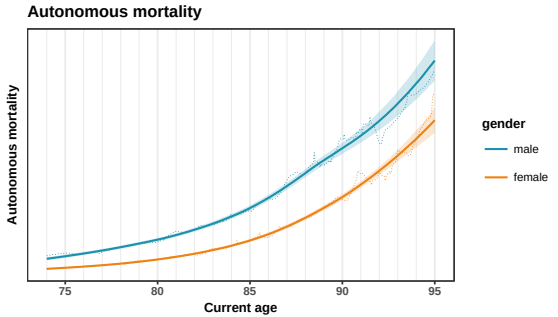
For each couple (d, α), we get a different estimator $\hat{\mu}_{d,\alpha}$. We then select one of them by using the **AIC** which balances the **accuracy** of the model with its **complexity**

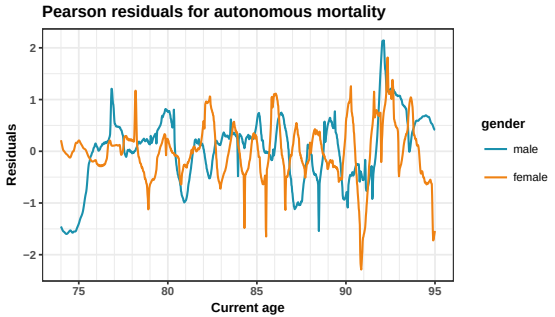
$$\text{AIC}(d, \alpha) = \underbrace{-2l(\hat{\mu}_{d,\alpha})}_{\text{log-likelihood}} + \underbrace{2\nu(d, \alpha)}_{\text{degrees of freedom}} .$$

Bias introduced by smoothing may be measured using *pseudo*-**Pearson residuals**. As **transition intensity** is not **directly observable**, the output of a local likelihood model with a very low amount of smoothing replaces the observation for this purpose

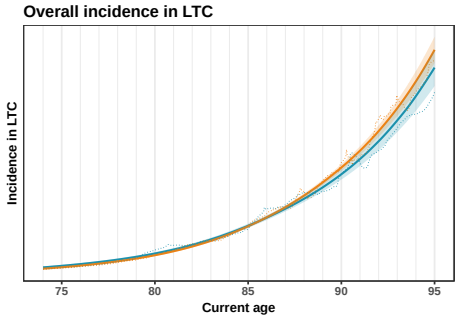
$$r_{d,\alpha}(x) = \frac{\hat{\mu}_{d,\alpha}(x) - \hat{\mu}_{d_0,\alpha_0}(x)}{\sqrt{\text{var}(\hat{\mu}_{d_0,\alpha_0}(x))}}.$$

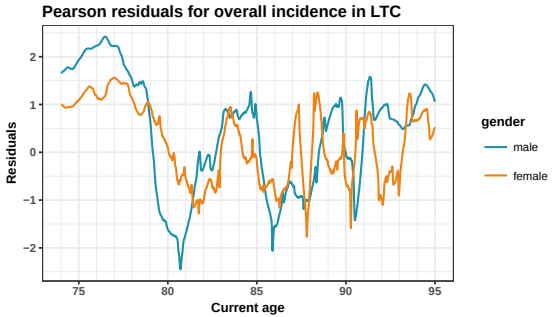
Autonomous mortality



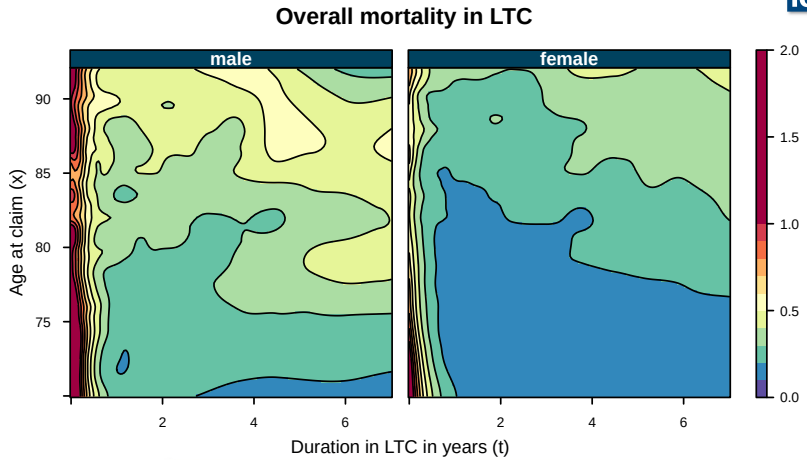


Incidence in LTC

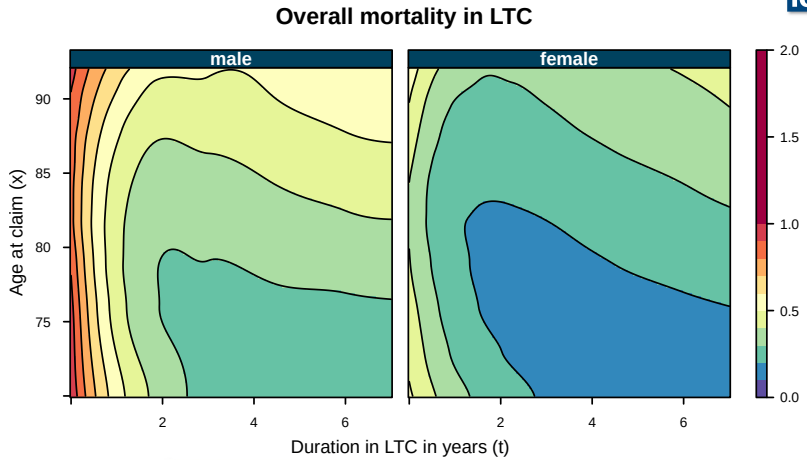




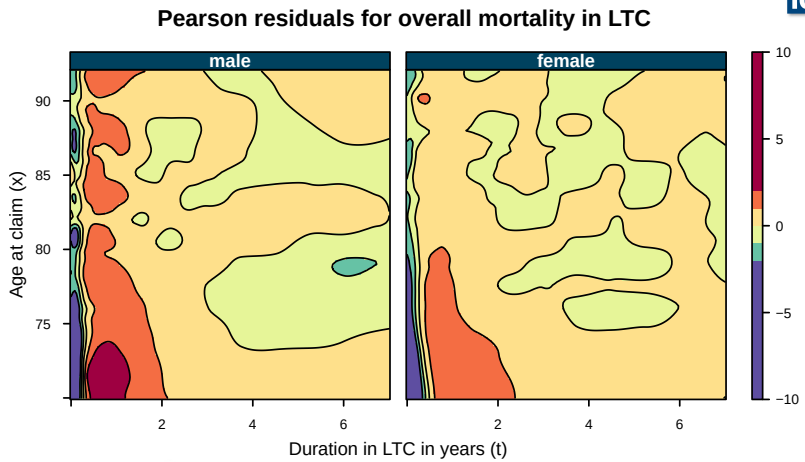
Overall mortality in LTC (minimal smooth)



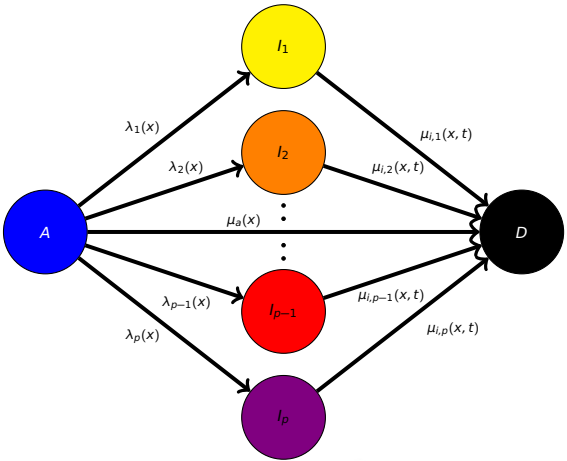
Overall mortality in LTC (optimal smooth)



Residuals for overall mortality in LTC



Multi-state model with pathologies



Lemma

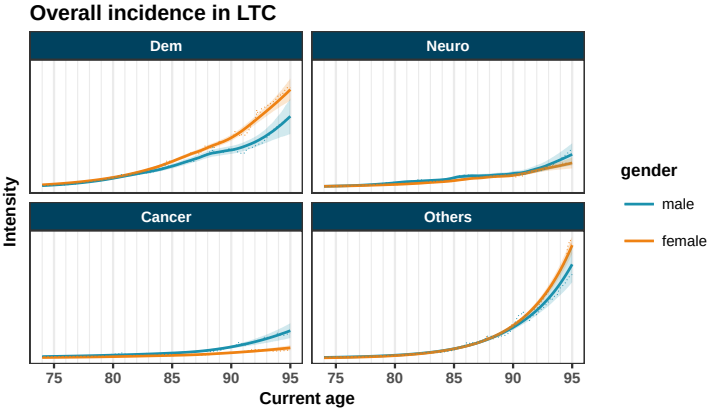
To ensure consistency we require, for $x, t \geq 0$

$$\lambda(x) = \sum_{k=1}^{p} \lambda_k(x)$$

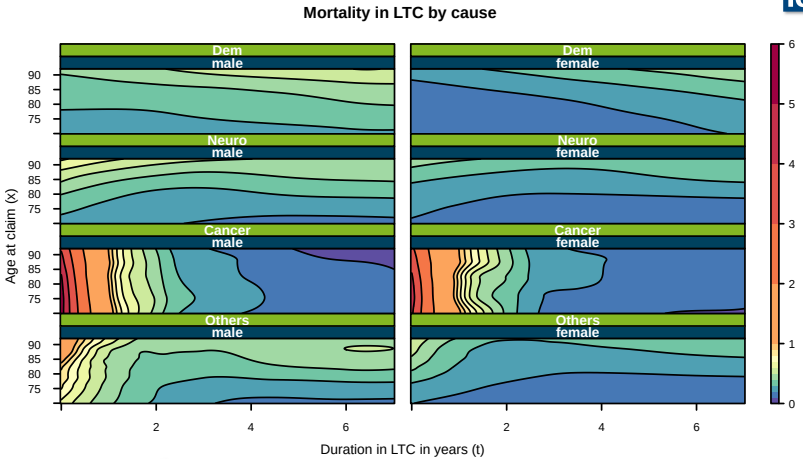
$$\mu_i(x,t) = \sum_{k=1}^{p} \eta_k(x,t)\mu_{i,k}(x,t)$$

$$\eta_k(x,t) = \frac{\lambda_k(x)\exp\left(-\int_0^t \mu_{i,k}(x,u)\mathrm{d}u\right)}{\sum_{j=1}^{p}\lambda_j(x)\exp\left(-\int_0^t \mu_{i,j}(x,u)\mathrm{d}u\right)}$$

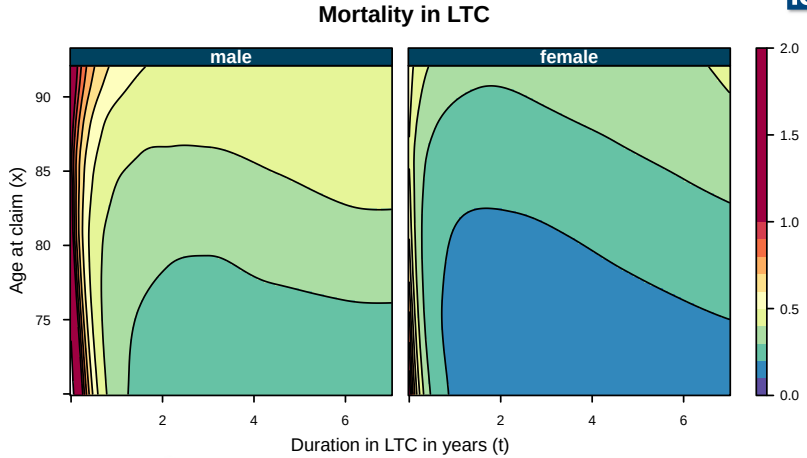
Overall incidence in LTC



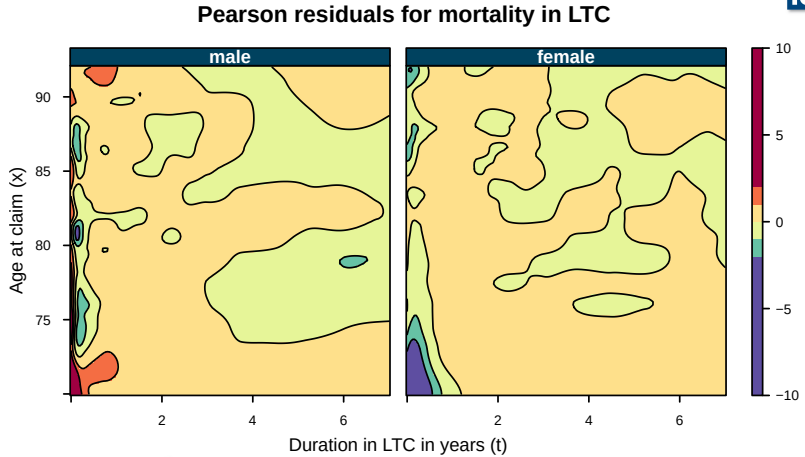
Mortality in LTC by cause



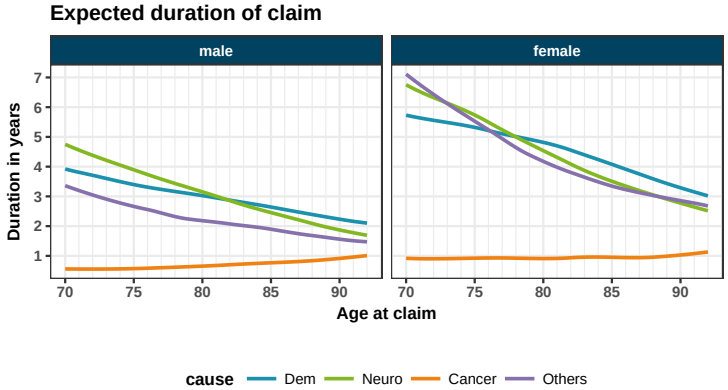
Second-step estimator of mortality in LTC



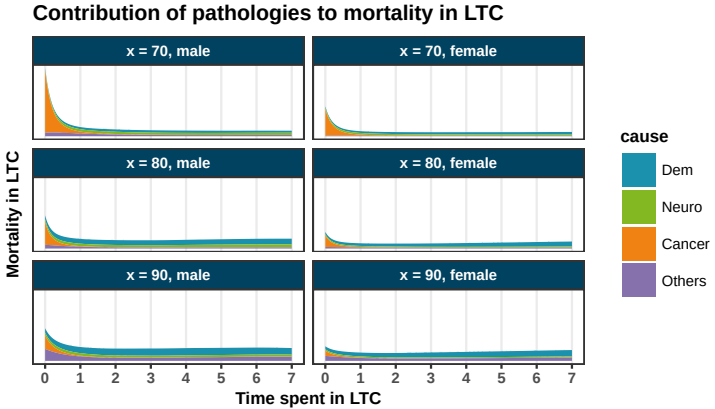
Residuals for second-step estimator of mortality in LTC



Life expectancy in LTC



Contribution of pathologies to mortality in LTC



Spin-off product: parametric model for mortality in LTC

Additive mixture model of 2 laws:

$$\mu_{i,1}(x,t) = \mu_0(x,t) + \Delta_1(x)$$
$$\lambda_1(x) = \theta(x)\lambda(x)$$

$$\mu_{i,2}(x,t) = \mu_0(x,t) + \Delta_2(x)$$
$$\lambda_2(x) = [1-\theta(x)]\lambda(x)$$

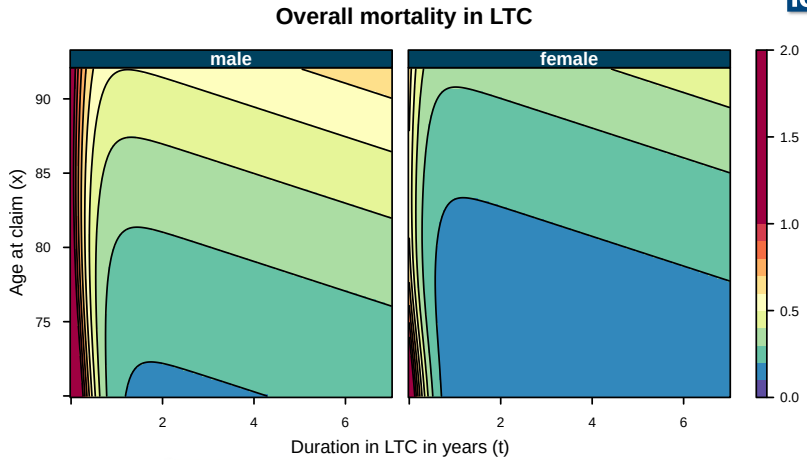
with:

- $\mu_0(x,t)$: common mortality term
- $\Delta_1(x)$: extra mortality for group 1 (cancer)
- $\Delta_2(x)$: extra mortality for group 2 (other causes)
- $\theta(x)$: initial weight of group 1 among new LTC entrants

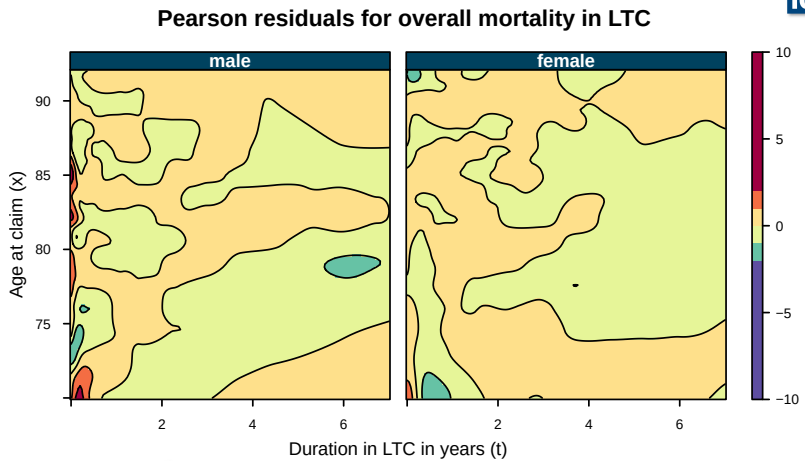
Resulting mortality in LTC for the whole population is then

$$\mu_i(x,t) = \mu_0(x,t) + \Delta_2(x) + \underbrace{\frac{\theta(x)\,[\Delta_1(x)-\Delta_2(x)]}{\theta(x)+[1-\theta(x)]\exp\left\{[\Delta_1(x)-\Delta_2(x)]\,t\right\}}}_{\substack{\to 0 \\ t\to\infty}}.$$

Overall mortality in LTC for parametric model



Residuals for parametric mortality in LTC



Conclusion

Chosen approach:

- Local-likelihood smoothing in a continuous-time framework
- Aggregated study followed by a study by group of pathologies

Main findings:

- Pathologies explain a large among of heterogeneity among trajectories
- Taking them into account greatly improve the model accuracy
- Using a mixture model extends the benefits even in case pathologies are not actually observed

Thank you very much for your attention!

Contact details:
Guillaume Biessy

address: 5 avenue Kleber
75016 Paris, France
phone: +33 (0)1 58 44 84 19
mail: gbiessy@scor.com
web: www.scor.com

References

- Tomas, J. and Planchet, F. (2013). Multidimensional smoothing by adaptive local kernel-weighted log-likelihood: Application to long-term care insurance. Insurance: Mathematics and Economics, 52(3), 573-589.
- Biessy, G. (2015). Long-term care insurance: A multi-state semi-Markov model to describe the dependency process in elderly people. Bulletin Français d'Actuariat, 15(29), 41-73
- Biessy, G. (2017). Continuous-time semi-markov inference of biometric laws associated with a long-term care insurance portfolio. ASTIN Bulletin: The Journal of the IAA, 47(2), 527-561.