Continuous time semi-Markov inference of biometric laws associated with a Long-Term Care Insurance portfolio

Guillaume Biessy

SCOR Global Life, LaMME

Monday, September 7th, 2015

High stakes

Definition

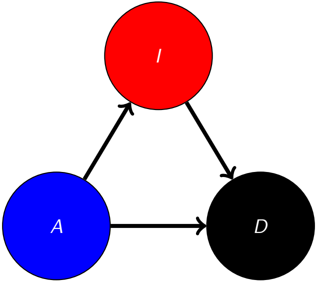
Dependency : permanent and consolidated state of inability to perform activities of daily living without someone else's help.

- Phenomenon linked to ageing
- Multiple causes : cancer, dementia, neurological and cardiovascular diseases ...
- Definition based on Activities of Daily Living (ADL) such as moving, clothing, bathing, feeding.
- Need to hire professional caregivers or join a nursing home.
- Associated costs up to 4 000 € a month in France.

The French long-term care insurance market :

- First products appeared in the 1980s.
- Second market in the world, behind the US market.
- 7,3 M insured lives, 665 M € premium paid in 2014.

Notation

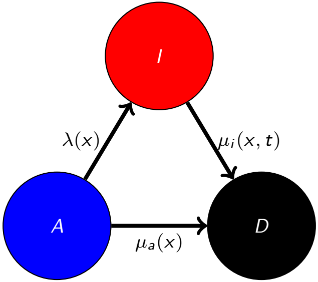


For $x_0 \geqslant 0$, let us consider a continuous-time process $(Z_x)_{x \geqslant x_0}$ with values in the 3-state set $E = \{A, I, D\}$ of autonomy, dependency, death. Let us further assume that Z is *cad-lag* and that $Z_{x_0} = A$ and denote

$$\begin{aligned} A_x =& P(Z_x = A), \\ I_x =& P(Z_x = I), \\ D_x =& P(Z_x = D). \end{aligned}$$

Hence $A_{x_0} = 1$ and for all $x \geqslant x_0$, $A_x + I_x + D_x = 1$.

Transition intensities



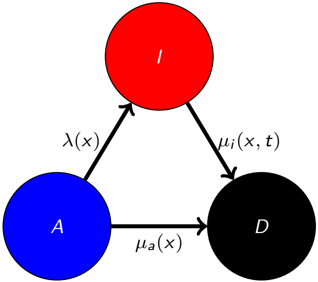
The model is then defined using the transition intensities :

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

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

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

Evolution equations

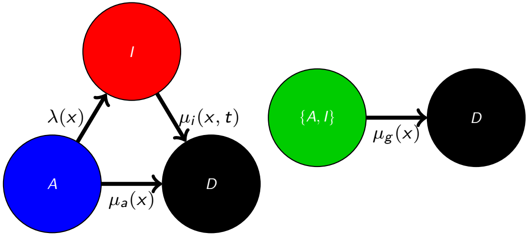


For $x \geqslant x_0$, $t \geqslant 0$, we have

$$A_x = \exp\left(-\int_{x_0}^{x} [\lambda(u) + \mu_a(u)]\, \mathrm{d}u\right),$$

$$I_x = \int_{x_0}^{x} \lambda(u) A_u \exp\left(-\int_{u}^{x} \mu_i(u, v-u)\mathrm{d}v\right)\mathrm{d}u.$$

Link with general mortality (1/2)



Intensity of mortality for the general population :

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

System of 3 differential equations :

$$\frac{\mathrm{d}}{\mathrm{d}x} A_x = -[\lambda(x) + \mu_a(x)] A_x, \tag{1}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} I_x = \lambda(x) A_x - \int_{x_0}^{x} \lambda(u) A_u \exp\left(-\int_u^x \mu_i(u, v-u)\mathrm{d}v\right) \mu_i(u, x-u)\mathrm{d}u, \tag{2}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} (A_x + I_x) = -\mu_g(x)(A_x + I_x). \tag{3}$$

Link with general mortality (2/2)

The following equation express the equivalence of forces of mortality

$$\mu_g(x)(A_x + I_x) = \mu_a(x)A_x + \int_{x_0}^{x} \lambda(u) A_u \exp\left(-\int_u^x \mu_i(u, v-u)\mathrm{d}v\right)\mu_i(u, x-u)\mathrm{d}u$$

By denoting

$$\Delta(x,t) = \mu_i(x,t) - \mu_a(x+t)$$

and after a few lines of calculation, we get

Mortality consistency equation

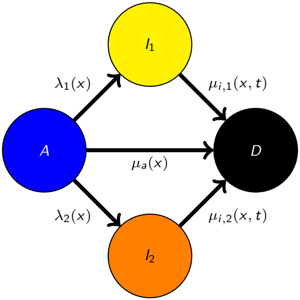
$$\mu_a(x) = \mu_g(x) - \frac{\int_{x_0}^{x} \lambda(u)\Delta(u, x-u)\exp\left(-\int_u^x \left[\Delta(u, v-u) - \lambda(v)\right]\mathrm{d}v\right)\mathrm{d}u}{1 + \int_{x_0}^{x} \lambda(u)\exp\left(-\int_u^x \left[\Delta(u, v-u) - \lambda(v)\right]\mathrm{d}v\right)\mathrm{d}u}.$$

Mixture model for dependent mortality (1/2)

How to model dependent mortality ?

- Insurer's experience shows very high death rates during the first few months spent in dependency.
- We believe this phenomenon can be explained by a strong heterogeneity among the dependent population, linked to causes of dependency.
- Causes like metastatic cancer, multiple stroke or severe cases of infarction are associated with very high death rates.
- Population structure therefore changes over time as individuals with higher mortality risks are among the first to die.
- Working assumption : dependent people can be separated in two populations, each with its own associated death rates.

Mixture model for dependent mortality (2/2)



Lemma

Consider the above dependency model with two states I_1 and I_2 and let us assume that for $x \geqslant x_0$, $t \geqslant 0$ and $k \in \{1,2\}$, $\mu_{i,k}(x,t) = \mu_a(x+t) + \Delta_k(x)$.
Then to ensure consistency with the 3-state model, the following relations must hold

$$\lambda(x) = \lambda_1(x) + \lambda_2(x)$$

$$\mu_i(x,t) = \mu_a(x+t) + \Delta_1(x) + \frac{\Delta_2(x) - \Delta_1(x)}{1 + \frac{\lambda_1(x)}{\lambda_2(x)} \exp([\Delta_2(x) - \Delta_1(x)]\, t)}$$

Data structure

We assume contributors/annuitant databases containing the following variables :

- DoB : date of birth of the individual,
- DoS : date of start. For contributors, it is the date of subscribing. For annuitants, the date of entry in dependency.
- DoE and CoE : Date of end and cause of end for the individuals. Code 1 for death, 2 for entry in dependency, 0 for right-censoring.

DoB	DoS	DoE	CoE
23/12/1941	11/10/1992	27/09/2006	2
14/06/1926	28/03/1997	31/12/2013	0
17/04/1937	28/03/1995	04/08/2003	1

TABLE 1 : Example of a database of contributors.

Estimation of parameters is then performed separately for men and women.

Presentation of the portfolio

- Dependency based on a 3ADL4 definition,
- High volume portfolio,
- Data for higher ages (up to 95) and high duration in dependency (up to 13 years).

Statistics	Contributors	Annuitants
Number of lines	160 669	17 632
Person-year exposure	1 325 758	43 010
Observation period	10 years	19 years
Censored trajectories	75,9 %	31,4 %
Percentage of women	65,3 %	65,9 %

Procedure for the estimation of biometric laws

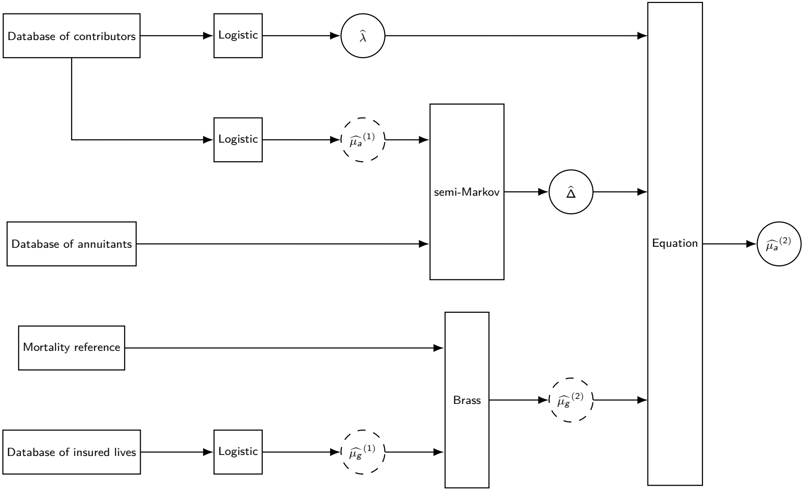


FIGURE 1 : Procedure for the estimation of biometric laws. Plain (resp. dashed) circle represents final (resp. intermediary) estimates of biometric laws.

Intensity of incidence in dependency

We consider a logistic form

$$\lambda(x) = \frac{\exp(a_\lambda x + b_\lambda)}{1 + \exp(a_\lambda x + c_\lambda)} + d_\lambda$$

with $a_\lambda > 0$, $b_\lambda, c_\lambda \in \mathbb{R}$ and $d_\lambda \geqslant 0$.

For an individual p defined by his/her age of entry in the portfolio $x_p \geqslant 0$, his/her age of exit $y_p > x_p$ and the associated exit cause $c_p \in \{0, 1, 2\}$ the partial log-likelihood has the following expression

$$\begin{aligned} l_p(\lambda) =& \delta^2_{c_p} \log(\lambda(y_p)) - \int_{x_p}^{y_p} \lambda(u) \mathrm{d}u \\ =& \delta^2_{c_p} \log\left(\frac{e^{a_\lambda y_p + b_\lambda}}{1 + e^{a_\lambda y_p + c_\lambda}} + d_\lambda\right) - \frac{e^{b_\lambda - c_\lambda}}{a_\lambda} \log\left(\frac{1 + e^{a_\lambda y_p + c_\lambda}}{1 + e^{a_\lambda x_p + c_\lambda}}\right) - d_\lambda (y_p - x_p), \end{aligned}$$

Intensity of general mortality (1/2)

Let F_g be the cumulative distribution function associated with the intensity of general mortality μ_g such that

$$F_g(x) = 1 - \exp\left(-\int_{x_0}^{x} \mu_g(u)\mathrm{d}u\right)$$

then we define the Cumulative Distribution Odds (CDO) by

$$\mathrm{CDO}_g(x) = \frac{F_g(x)}{1 - F_g(x)}.$$

We use Brass relational model with the assumption that the logarithm of the CDO associated with the mortality of observed and reference populations are parallel curves and we use the following estimator for the general mortality

$$\widehat{\mu_g}^{(2)}(x) = \frac{\widehat{\beta}\mu_g^{ref}(x)}{1 - (1 - \widehat{\beta})F_g^{ref}(x)}$$

where $\widehat{\beta}$ is the solution of the equation

$$\sum_x D_x = \sum_x \frac{\widehat{\beta}}{(1 - (1 - \widehat{\beta})F_g^{ref}(x))} q_g^{ref}(x) N_x$$

with D_x (resp. N_x) being the number of death observed (resp. the number of person at risk) between ages x and $x + 1$.

Intensity of general mortality (2/2)

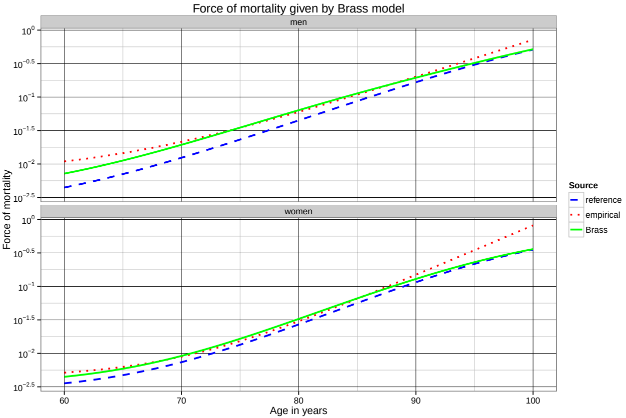


FIGURE 2 : Intensity of mortality estimated from data (dotted), from the mortality reference (dashed) and resulting from Brass model (plain). A translation of the y-scale has been applied to preserve confidentiality of results.

Intensity of dependent mortality (1/2)

We use the following model for the intensity of dependent mortality

$$\mu_i(x,t) = \mu_a(x+t) + \phi_1 + \exp(\phi_2 x + \phi_3) + \frac{\phi_4}{1+\exp(\phi_4 t + \phi_5 x + \phi_6)}.$$

which corresponds to

$$\begin{cases} \Delta_1(x) = & \phi_1 + \exp(\phi_2 x + \phi_3) \\ \Delta_2(x) = & \Delta_1(x) + \phi_4 \\ \lambda_1(x) = & \frac{\exp(\phi_5 x + \phi_6)}{1+\exp(\phi_5 x + \phi_6)} \lambda(x) \end{cases}$$

The associated partial log-likelihood for an individual p with an age of entry in dependency $x_p \geqslant 0$, an age of exit $y_p > x_p$ and the associated cause of exit $c_p \in \{0,1\}$ is

$$\begin{aligned} l_p(\mu_i) =& \delta^1_{c_p} \log(\mu_i(x_p, y_p - x_p)) - \int_{x_p}^{y_p} \mu_i(x_p, u - x_p) \mathrm{d}u \\ =& \delta^1_{c_p} \log\left(\mu_a(y_p) + \phi_1 + e^{\phi_2 x_p + \phi_3} + \frac{\phi_4}{1+\exp(\phi_4 [y_p - x_p] + \phi_5 x_p + \phi_6)}\right) - \\ & (\phi_1 + e^{\phi_2 x_p + \phi_3} + \phi_4)[y_p - x_p] + \log\left(\frac{1+\exp(\phi_5 x_p + \phi_6)}{1+\exp(\phi_4 [y_p - x_p] + \phi_5 x_p + \phi_6)}\right). \end{aligned}$$

Intensity of dependent mortality (2/2)

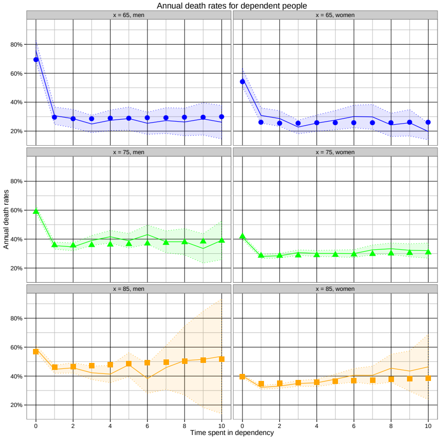


FIGURE 3 : Consecutive death rates for dependent people according to the model (points) with empirical rates (plain line) and 95 % confidence intervals (dashed lines). A translation of the y-scale has been applied to preserve confidentiality of results.

Intensity of autonomous mortality

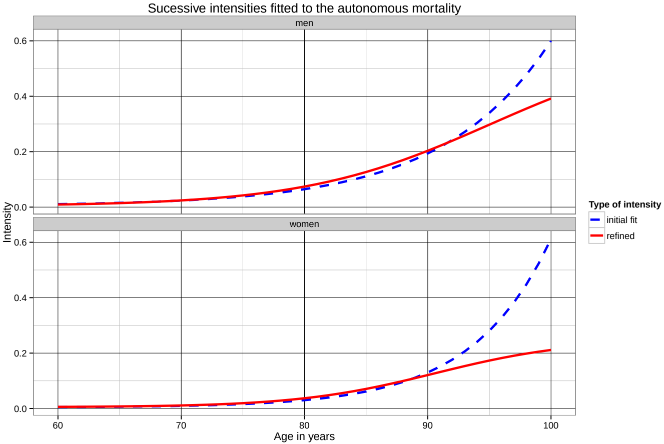


FIGURE 4 : Intensity of autonomous mortality : estimation *a priori* in red, computation using the mortality consistency equation in blue. A multiplicative factor has been applied to the y-scale to preserve confidentiality of results.

Summary of intensities associated with the model

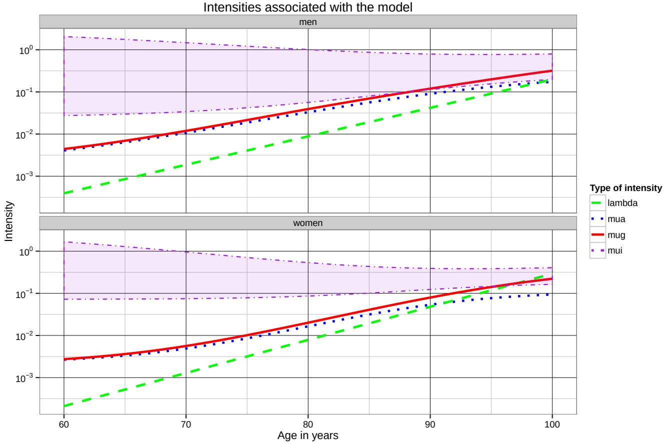


FIGURE 5 : Intensities of autonomous mortality in blue, incidence in dependency in green, general mortality in red and range for intensity of dependent mortality in purple. A translation of the y-scale has been applied to preserve confidentiality of results.

Link with annual rates (1/3)

For $x \geqslant x_0$ and $t \geqslant 0$

$$
\begin{aligned}
q_g(x) =& 1 - \exp\left(-\int_x^{x+1} \mu_g(u)\mathrm{d}u\right), \\
q_{aa}(x) =& \int_x^{x+1} \mu_a(u) \exp\left(-\int_x^u [\lambda(v) + \mu_a(v)\mathrm{d}v]\right)\mathrm{d}u, \\
i(x) =& \int_x^{x+1} \lambda(u) \exp\left(-\int_x^u [\lambda(v) + \mu_a(v)\mathrm{d}v]\right)\mathrm{d}u, \\
q_i(x,t) =& 1 - \exp\left(-\int_t^{t+1} \mu_i(x,u)\mathrm{d}u\right).
\end{aligned}
$$

Using those annual rates, we are able to compare the results with those of discrete time models. However, pricing and reserving should rely on continuous time formulas, as those are more accurate.

Link with annual rates (2/3)

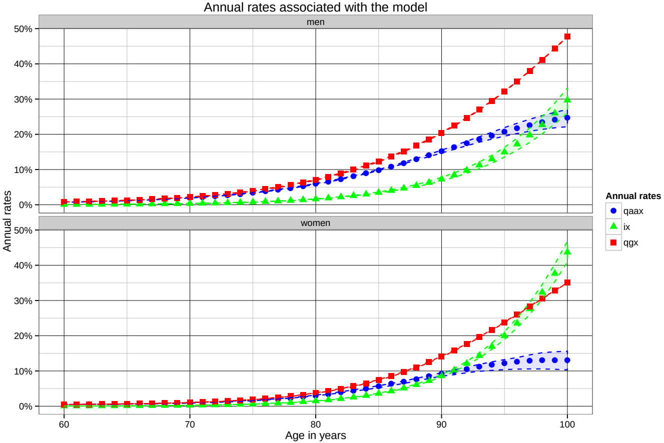


FIGURE 6 : Annual rates for autonomous mortality in blue, incidence in dependency in green, general mortality in red, with 95 % confidence intervals obtained by bootstrap. A multiplicative factor has been applied to the y-scale to preserve confidentiality of results.

Link with annual rates (3/3)

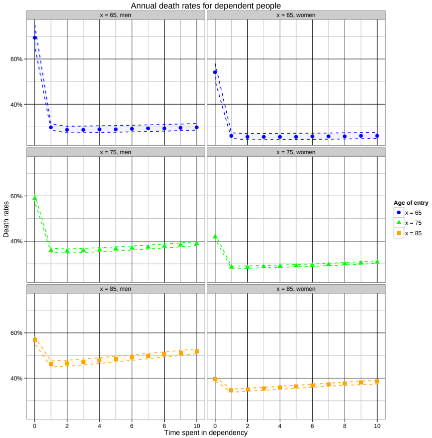


FIGURE 7 : Consecutive death rates for dependent people with respect to gender and age of entry, with 95 % confidence intervals obtained by bootstrap. A translation of the y-scale has been applied to preserve confidentiality of results.

Prevalence of dependency

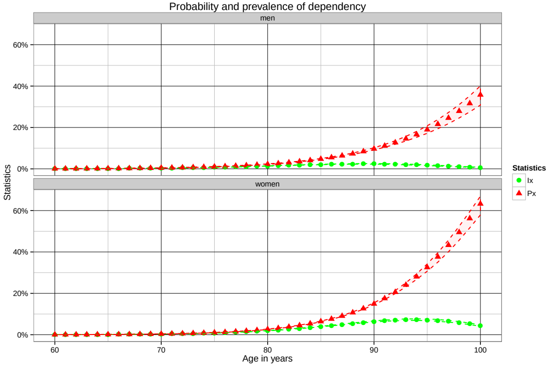


FIGURE 8 : Probability of being alive and dependent at age x and prevalence of dependency at age x, with 95 % confidence intervals obtained by bootstrap.

Pricing methodology (1/3)

We consider a product where the autonomous insured life pays a fixed amount of premium p/f_1, starting at subscribing and at the start of every period of duration $1/f_1$ year. Should he/she become dependent, he/she is entitled to an annuity R/f_2 at the end of every period of duration $1/f_2$.

- τ the continuous time actuarial rate used to compute discounted cash flows,
- f_1 (resp f_2) the number of payments of premium (resp. annuities) in one year, For most products in France, $f_1 = f_2 = 12$,
- ω the age of closure for mortality tables (we consider $\omega = 120$).

Pricing methodology (2/3)

We denote by

- $P(x_s, x)$ the expected value of insured liabilities for an autonomous insured life with age at subscribing x_s and current age x for an amount of premium of 1

$$P(x_s, x) = \frac{1}{f_1} \sum_{k \in \mathbb{N} \cap [f_1(x - x_s); f_1(\omega - x_s)[} \exp\left(- \int_x^{x_s + \frac{k}{f_1}} [\mu_a(u) + \lambda(u) + \tau] \, \mathrm{d}u \right),$$

- $C(R, x_i, t)$ the expected value of insurer liabilities for an insured life with age of entry in dependency x_i who survived t years in dependency and an annual amount of benefit R

$$C(R, x_i, t) = \frac{R}{f_2} \sum_{k \in \mathbb{N} \cap]f_2 t; f_2(\omega - x_i)]} \exp\left(- \int_t^{\frac{k}{f_2}} [\mu_i(x, u) + \tau] \, \mathrm{d}u \right),$$

- $B(R, x)$ the expected value of insurer liabilities for an autonomous insured life with current age x and an annual amount of benefit R

$$B(R, x) = \int_x^{\omega} \lambda(u) \exp\left(- \int_x^{u} [\mu_a(v) + \lambda(v) + \tau] \, \mathrm{d}v \right) C(R, u, 0) \mathrm{d}u$$

Pricing methodology (3/3)

- The stability premium $p^*(R, x_s)$. It is the value of premium that matches insurer and insured liabilities at the time of subscribing. For an age x_s at subscribing and an annual amount of benefit R we have

$$p^*(R, x_s) = \frac{B(R, x_s)}{P(x_s, x_s)}.$$

- The reserve for premium (RFP). This reserve is constituted for autonomous people. Its amount is equal to the expectancy of future discounted cash flows of benefit minus discounted cash flows of premium. For an insured of age at subscribing x_s, current age x, an annual amount of premium p and annual amount of benefit R, the associated amount of reserve is

$$\text{RFP}(p, R, x_s, x) = B(R, x) - p\ P(x_s, x).$$

- The reserve for claim (RFC). This reserve is constituted for dependent people. Its amount is equal to the expectancy of the future discounted cash flows of benefit. For an annual amount of benefit R, an age x_i at entry in dependency and a time t spent in dependency, the corresponding amount of reserve is

$$\text{RFC}(R, x_i, t) = C(R, x_i, t).$$

Resulting premium

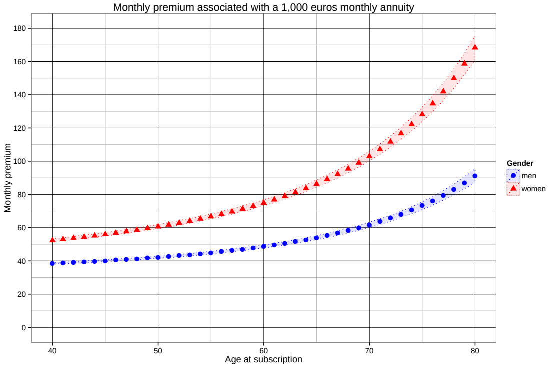


FIGURE 9 : Amount of monthly premium required according to the model, with 95 % confidence intervals obtained by bootstrap. A translation of the y-scale has been applied to preserve confidentiality of results.

Reserve for premium

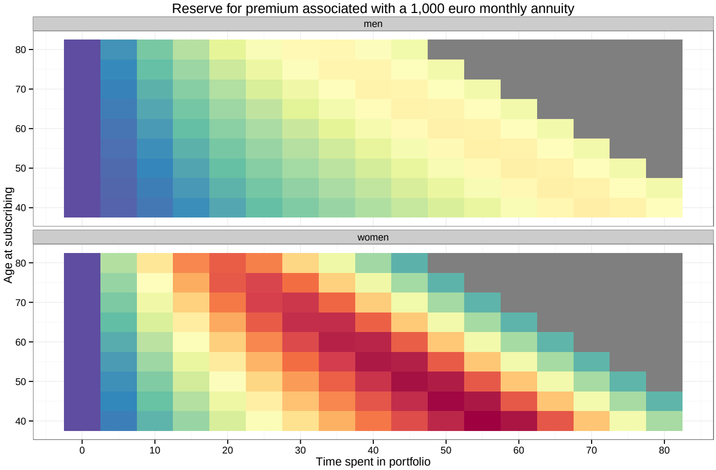


FIGURE 10 : Reserve for premium by age at subscribing and time spent in portfolio.

Reserve for claim

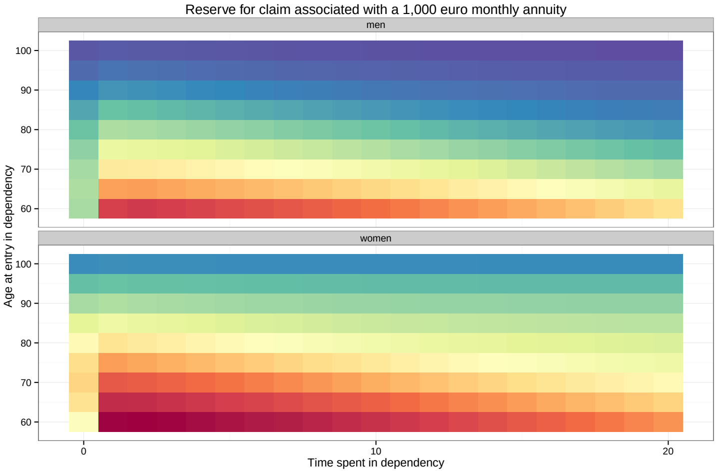


FIGURE 11 : Reserve for claim by age at entry in dependency and time spent in dependency.

Conclusion

Summary

- Continuous-time definition of the dependency process
- Model taking into account the general mortality
- Parametric expressions for transition intensities
- Estimation using the maximum likelihood method
- Application to data from an existing portfolio

Identified risks and solutions

- Robustness of estimation (bootstrap)
- Model risks (comparison with empirical rates, BIC)
- Drift for intensities (scenarios of evolution)

Potential improvements

- Study of pathologies causing dependency for a better model,
- Extension to consider multiple states of dependency,
- Model including long-term drifts.

Questions

Thank you for your attention !

Any questions ?