Pricing and Risk Analysis of a Long-Term Care Insurance Contract in a non-Markov Multi-State Model

Quentin Guibert

Univ Lyon, Université Claude Bernard Lyon 1, ISFA, Laboratoire SAF EA2429, F-69366, Lyon, France
Prim'Act, Paris, France
Email: q.guibert@hotmail.fr

3rd European Actuarial Journal Conference
5-8 September 2016, Lyon, France

Joint work with F. Planchet (ISFA and Prim'Act)

Outline

Motivation

- Multi-state models are the most natural tools for pricing and reserving LTC guarantees.
- In general, a Markov model is used. This may be too strong.
- Inference methodology based on crude intensities → GLM Poisson models that depend on age or duration time (e.g. Haberman and Pitacco, 1998; Pritchard, 2006; Levantesi and Menzietti, 2012; Fong *et al.*, 2015) .
- Lack of (detailed) national data. Difficulty to have precise and accurate insurance data. No covariate.
- Need for realistic (best estimate) tables for LTC insurance contracts.

Motivation

- Insurers observe longitudinal data with censoring and truncation $\rightarrow$ interest of non-parametric techniques for estimating transition probabilities.
- Can be estimated efficiently by direct approaches (without calculating transition intensities) for particular acyclic multi-state models, even when the Markov assumption is not satisfy (Meira-Machado *et al.*, 2006; de Uña-Álvarez and Meira-Machado, 2015; Guibert and Planchet, 2016). $\rightarrow$ We use directly quantities of interest.
- Lack of regression methodology in this context as traditional methods are based on intensities or use the Markov assumption.
- Several direct regression models on the survival function, the cumulative incidence function or the state occupation probability based on pseudo-values (Andersen *et al.*, 2003; Klein and Andersen, 2005; Murawska and Rizopoulos, 2015). Helms *et al.* (2005) use these techniques for transition probabilities with a Markov model.
- Recent extentions to dynamic prediction (Nicolaie *et al.*, 2013; Grand and Putter, 2016).

Aims

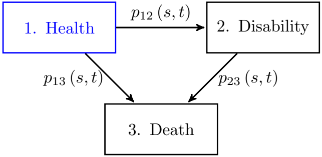
- Extend existing dynamic regression methods using directly the transition probabilities of a non-Markov illness-death model.
  $\rightarrow$ Flexible models and robust against misspecification.
- Measure the effect of observed covariates with LTC insurance data.
- Illustration for pricing and reserving.

Long-Term Care Insurance Model

- Consider an illness-death model X with only one heavy dependency state.
- Transition probabilities $h \to j$ between s and t with individual covariates Z

$$p_{hj}(s,t;Z(s)) = \mathbb{P}(X(t)=j \mid X(s)=h, Z(s)).$$

- A very simple annuity guarantee paid in dependency.

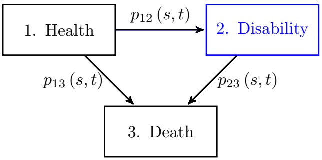


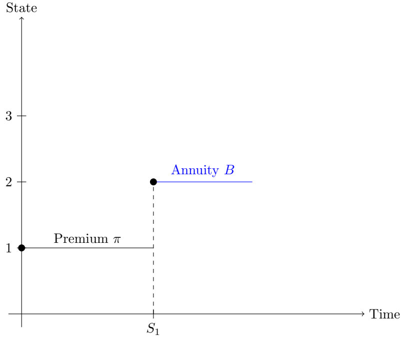
Long-Term Care Insurance Model

- Consider an illness-death model X with only one heavy dependency state.
- Transition probabilities $h \to j$ between s and t with individual covariates Z

$$p_{hj}(s, t; Z(s)) = \mathbb{P}(X(t) = j \mid X(s) = h, Z(s)) .$$

- A very simple annuity guarantee paid in dependency.



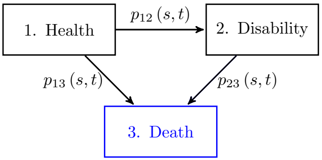


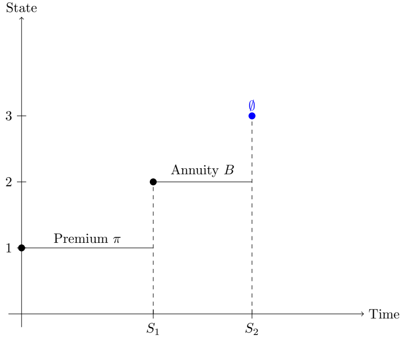
Long-Term Care Insurance Model

- Consider an illness-death model X with only one heavy dependency state.
- Transition probabilities $h \to j$ between s and t with individual covariates Z

$$p_{hj}(s, t; Z(s)) = \mathbb{P}(X(t) = j \mid X(s) = h, Z(s)).$$

- A very simple annuity guarantee paid in dependency.





Long-Term Care Insurance Model

- Actuarial calculation in discrete time at the valuation time 0 for a s-year old insured.
- Actuarial value of premium payments

$$a_1\left(s;Z\left(s\right)\right)=\sum_{\tau=0}^{\omega-s-1} v^{\tau} p_{11}\left(s,s+\tau;Z\left(s\right)\right)\pi.$$

- Actuarial values of annuity payments in state $h=1,2$

$$A_h\left(s;Z\left(s\right)\right)=\sum_{\tau=0}^{\omega-s-1} v^{\tau} p_{h2}\left(s,s+\tau;Z\left(s\right)\right)B.$$

- Calculating these probabilities is not simple without the Markov assumption. For state 1, we suppose that it's verified (no previous state).
- More sophisticated payments (e.g. depending on duration time, cause of entry) require estimating more complex transition probabilities (Guibert and Planchet, 2016).

Non-Parametric Transition Probabilities

- Introduce S, the lifetime in healthy state, and T, the overall lifetime.
- Let C and L be a right-censoring and a left-truncation variables. Truncations only occur in state 1.
- Suppose $(C,L) \perp\!\!\!\perp (S,T,Z)$ and $C \perp\!\!\!\perp L$. If $L \leq S$, we observe

$$\begin{cases} \widetilde{S} = \min(S,C) \text{ and } \gamma = \mathbb{1}_{\{S \leq C\}}, \\ \widetilde{T} = \min(T,C) \text{ and } \delta = \mathbb{1}_{\{T \leq C\}}. \end{cases}$$

Consistent estimators of transition probabilities

By selecting individuals in $\mathcal{L}_s^1 = \left\{ i : \widetilde{S}_i > s \right\}$, we have

$$\widehat{p}_{11}(s,t) = 1 - \widehat{H}_s^1(t), \ \widehat{p}_{12}(s,t) = \widehat{H}_s^1(t) - \widehat{F}_s^1(t),$$

and with individuals in $\mathcal{L}_s^2 = \left\{ i : \widetilde{S}_i \leq s < \widetilde{T}_i \right\}$

$$\widehat{p}_{22}(s,t) = 1 - \widehat{F}_s^2(t),$$

with $\widehat{H}_s^1(t)$ and $\widehat{F}_s^1(t)$ be the Kaplan-Meier c.d.f of S and T on $\mathcal{L}_s^1$ (Tsai *et al.*, 1987), and $\widehat{F}_s^2(t)$ that of T calculated on $\mathcal{L}_s^2$.

Dynamic Pseudo-Values

- Define a set of landmark time points $\mathcal{S} = \{s_1, \ldots, s_K\}$ and, for each s_k, a grid of points $\mathcal{T}_{s_k}$.
- Let be a (multivariate) parameter $\theta = \mathbb{E}\left[f\left(S, T\right)\right]$ for some function f. For varying landmark time points s and for several prediction points t, our aim is to construct a regression GLM model of the form

$$\theta_i = \mathbb{E}\left[f\left(S_i, T_i\right) \mid Z_i\right] = g^{-1}\left(\beta^\top Z_i\right),$$

with g, an invertible link function (cloglog-link function in the application).

- For each landmark s, define the jackknife pseudo-values for individual i at each time $t \in \mathcal{T}_s \rightarrow$ responses in the regression model.

$h \rightarrow j$	f_{st}^{hj}	Pseudo-values	Selection
$p_{11}\left(s, t\right)$	$\mathbb{1}_{\{S>t\}}$	$\widehat{\theta}_{ist}^{11} = n_s^1 \widehat{p}_{11}\left(s, t\right) - \left(n_s^1 - 1\right) \widehat{p}_{11}^{(-i)}\left(s, t\right)$	$\mathcal{L}_s^1$
$p_{12}\left(s, t\right)$	$\mathbb{1}_{\{S\leq t\}} - \mathbb{1}_{\{T\leq t\}}$	$\widehat{\theta}_{ist}^{12} = n_s^1 \widehat{p}_{12}\left(s, t\right) - \left(n_s^1 - 1\right) \widehat{p}_{12}^{(-i)}\left(s, t\right)$	$\mathcal{L}_s^1$
$p_{22}\left(s, t\right)$	$\mathbb{1}_{\{T>t\}}$	$\widehat{\theta}_{ist}^{22} = n_s^2 \widehat{p}_{22}\left(s, t\right) - \left(n_s^2 - 1\right) \widehat{p}_{22}^{(-i)}\left(s, t\right)$	$\mathcal{L}_s^2$

- Stacked models for each transition $h \rightarrow j$: hj index is suppressed to simplify.

Regression Approach 1: Fixed Landmark

- For each landmark s, consider $f_s = (f_{st}, t \in \mathcal{T}_s)$, $\theta_{is} = (\theta_{ist}, t \in \mathcal{T}_s)$ and a link

$$\theta_{ist} = g^{-1}\left(\beta(s)^\top Z_{it}(s)\right).$$

- Each $\beta(s)$ is estimated using the GEE approach (Liang and Zeger, 1986), i.e. by resolving the score equation

$$\mathcal{U}(\beta(s)) = \sum_i \left(\frac{\partial}{\partial \beta(s)} \theta_{is}\right) M_i^{-1} \left(\widehat{\theta}_{is} - \theta_{is}\right) = 0$$

- A working covariance matrix M_i must be fixed and can be selected using the QIC criteria (Pan, 2001).
- Variance of $\widehat{\beta}(s)$ is obtained by using a sandwich estimator.
- Convergence and asymptotic normality of $\widehat{\beta}(s)$ is not formally demonstrated here, but is undoubtedly satisfy (Graw *et al.*, 2009). However, performances should be investigated.

Regression Approach 2: Supermodels

- If lankmark times are close enough, the effect of covariates should vary smoothly over s. We consider simultaneously all landmarks, which can now interact with covariables.
- Consider $f = (f_{st}, s \in \mathcal{S}, t \in \mathcal{T}_s)$, $\theta_i = (\theta_{ist}, s \in \mathcal{S}, t \in \mathcal{T}_s)$ and a link

$$\theta_{ist} = g^{-1}\left(\beta(s)^\top Z_{it}(s)\right).$$

- Each component l of the vector $\beta(s)$ can be written as a polynomial function

$$\beta_l(s) = \beta_l{}^\top q_l(s).$$

- The vector β, such that $\beta(s) = Q(s)\beta$, is estimated using the GEE approach as for the regression Approach 1.
- To have consistent results, we follow Nicolaie *et al.* (2013) and choose an independence working correlation matrix.

Insurance Data

- Individual insurance data with left-truncation (only in healthy state) and right-censoring.
- $\simeq$ 15,900 (63% of women) contracts observed on a 13-years period and almost 47% are censored in healthy state $\rightarrow$ only an extract of the database to have acceptable computation times.
- 4 groups of pathologies. Deaths in the healthy state are grouped with cancellations (Guibert and Planchet, 2014).
- Mortality varies a lot with pathologies causing entry into dependency. Thus, we consider a model for each pathology.

Pathology	n	$\%$ of women	Age of entry S				
			Mean	Median	Std err.	Q1	Q3
Neurological pathologies	457	58.2%	79.3	79.5	5.5	76.0	83.3
Various pathologies	426	69.2%	82.1	82.6	5.3	79.1	85.6
Terminal cancers	285	45.3%	78.4	78.8	6.2	74.0	82.9
Dementia	1038	70.6%	81.1	81.4	5.1	78.2	84.8

Estimated Transition Probabilities

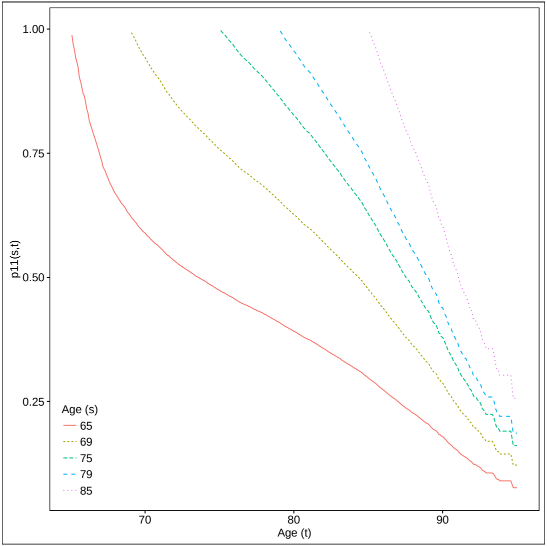


Figure: $\widehat{p}_{11}(s,t)$

Estimated Transition Probabilities

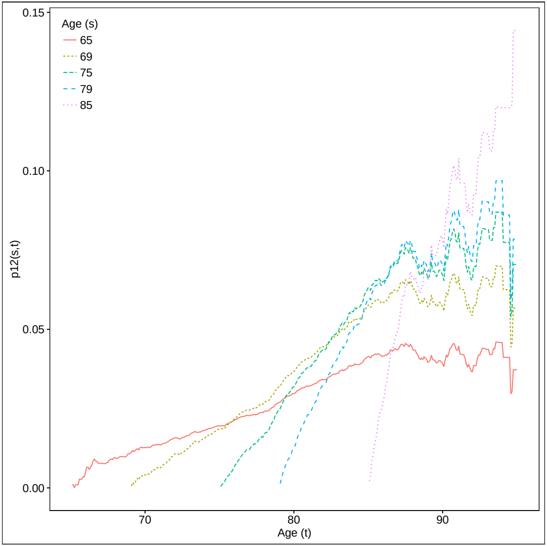


Figure: $\widehat{p}_{12}\left(s,t\right)$

Estimated Transition Probabilities

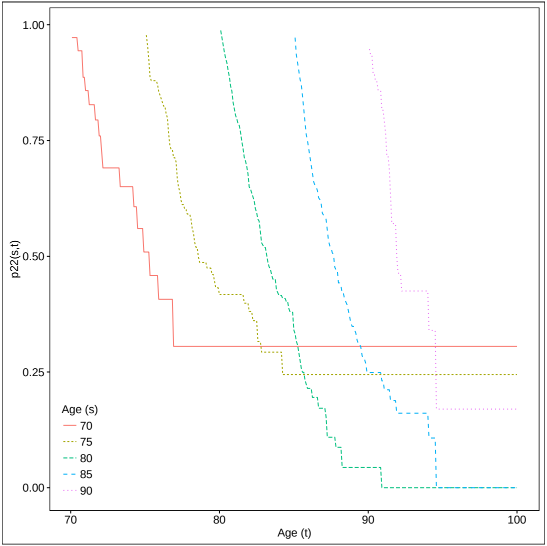


Figure: $\widehat{p}_{22}\,(s,t)$

Approach 1 vs. Approach 2 for $p_{12}(s,t)$

Model specification

- $\mathcal{S} = \{65, 67, \ldots, 85\}$ and, for each s, $\mathcal{T}_s = \{s+1, s+2, \ldots, 90\}$.
- Approach 1: $g(\theta_{ist}) = \beta_0(s) + \beta_1(s)\,t + \beta_2(s)\,t^2 + \beta_3(s)\,Sex$.
- Approach 2: $g(\theta_{ist}) = \beta_0(s) + \beta_1(s)\,t + \beta_2(s)\,t^2 + \beta_3(s)\,Sex + \beta_4 s^2 t \mathbb{1}_{\{t-s\leq 3\}} + \beta_5 s^2 t \mathbb{1}_{\{t-s>3\}}$, where $\beta_l(s) = \beta_{l0} + \beta_{l1}s$, $l = 0, \ldots, 3$.

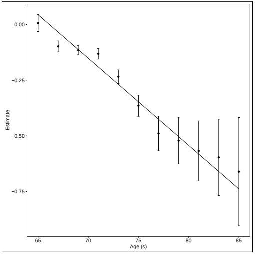


Figure: $\widehat{p}_{12}(s,t)$ - Coefficients of men with associated 95% confidence intervals. Points : Approach 1. Solid line : Approach 2.

Approach 1 vs. Approach 2 for $p_{12}(s,t)$

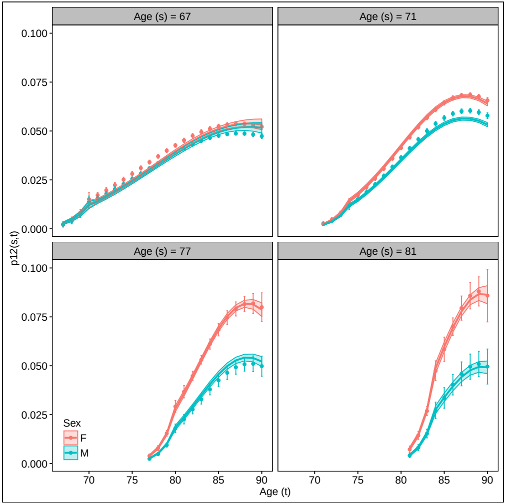


Figure: $\widehat{p}_{12}(s,t)$ - Predicted transition probabilities for men and women with associated 95% prediction intervals. Points : Approach 1. Solid line : Approach 2.

Approach 2 for $p_{22}(s,t)$

- A lot of time points should be consider to satisfy actuarial needs.
- Results obtained with Approach 1 are not robust. We consider only Approach 2.
- The other causes are in the study phase.

Model specification

- $\mathcal{S} = \{70, 70.25, \ldots, 90\}$. For each s, $\mathcal{T}_s = \left\{s + \frac{1}{12}, s + \frac{2}{12}, \ldots, s+5\right\}$.
- Neurological pathologies:

$$g(\theta_{ist}) =$$
$$\left(\beta_0(s) + \beta_1(s)t + \beta_2(s)t^2 + \beta_3(s)Sex + \beta_4(s)S + \beta_5(s)S^2\right)\mathbb{1}_{\{t-s\leq 1\}}$$
$$+\left(\beta_6(s) + \beta_7(s)t + \beta_8(s)t^2 + \beta_9(s)Sex + \beta_{10}(s)S + \beta_{11}(s)S^2\right)\mathbb{1}_{\{t-s>1\}}$$
$$+\beta_{12}Sex\cdot t$$

where $\beta_l(s) = \beta_{l0} + \beta_{l1}s$, $l = 0, \ldots, 11$.

Fitted $p_{22}(79, t)$ with Approach 2

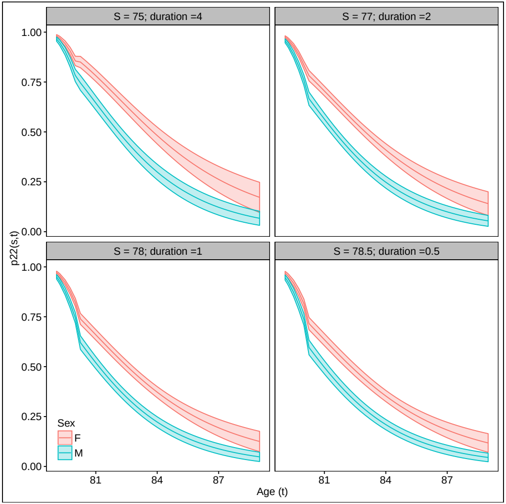


Figure: Neurological pathologies - Predicted transition probabilities for men and women with associated 95% prediction intervals.

Actuarial Application - Premiums Calculation

Principle of equivalence for a s-year old insured

$$a_1(s; Z(s)) = A_1(s; Z(s))$$

- Annual payments.
- $B = 1$ Eur, $v = (1 + 0.01)^{-1}$, Age limit$= 90$.

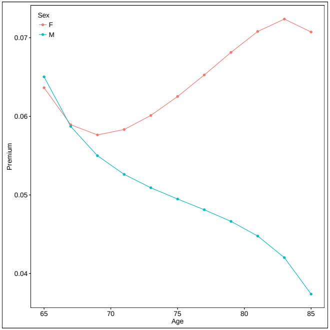


Figure: Comparison of Premiums.

Conclusion

- New dynamic regression approach based on pseudo-values for the transition probabilities of a non-Markov illness-death model.
- Two alternatives: fixed landmark models and landmark supermodels in a similar way that Nicolaie *et al.* (2013):
  - flexible methodologies,
  - the effects of covariates are directly estimated on pseudo-values $\rightarrow$ more interpretable.
- Tables construction by sex where the cause of entry in dependency are distinguished.
- Pseudo-values calculation needs important computation times. Estimation with a GEE approach can be done easily using the `geepack` `R`-package (Halekoh and Højsgaard, 2006).
- New researches are need to check models' performances.
- Fixed landmark models are tricky to use with $p_{22}(s,t)$ due to missing data. $\rightarrow$ Our dataset will be completed soon by adding 5 years of observations.
- The effect of working correlation matrices and performances of our methods should be tested with simulated data.

References I

Andersen, P. K., Klein, J. P., and Rosthøj, S. (2003). Generalised linear models for correlated pseudo-observations with applications to multi-state models. *Biometrika*, **90**(1), 15–27.

de Uña-Álvarez, J. and Meira-Machado, L. (2015). Nonparametric estimation of transition probabilities in the non-Markov illness-death model: A comparative study. *Biometrics*.

Fong, J. H., Shao, A. W., and Sherris, M. (2015). Multistate Actuarial Models of Functional Disability. *North American Actuarial Journal*, **19**(1), 41–59.

Grand, M. K. and Putter, H. (2016). Regression models for expected length of stay. *Statistics in Medicine*, **35**(7), 1178–1192.

Graw, F., Gerds, T. A., and Schumacher, M. (2009). On pseudo-values for regression analysis in competing risks models. *Lifetime Data Analysis*, **15**(2), 241–255.

Guibert, Q. and Planchet, F. (2014). Construction de lois d'expérience en présence d'évènements concurrents – Application à l'estimation des lois d'incidence d'un contrat dépendance. *Bulletin Français d'Actuariat*, **13**(27), 5–28.

Guibert, Q. and Planchet, F. (2016). Non-Parametric Inference of Transition Probabilities Based on Aalen-Johansen Integral Estimators for Acyclic Multi-State Models: Application to LTC Insurance. Preprint, Institut de Science Financière et d'Assurances, Lyon.

Haberman, S. and Pitacco, E. (1998). *Actuarial Models for Disability Insurance*. Chapman and Hall/CRC, 1 edition.

Halekoh, U. and Højsgaard, S. (2006). The R Package geepack for Generalized Estimating Equations. *Journal of Statistical Software*, **15**(2). 00551.

Helms, F., Czado, C., and Gschlößl, S. (2005). Calculation of LTC Premiums Based on Direct Estimates of Transition Probabilities. *ASTIN Bulletin: The Journal of the International Actuarial Association*, **35**(02), 455–469.

Klein, J. P. and Andersen, P. K. (2005). Regression Modeling of Competing Risks Data Based on Pseudovalues of the Cumulative Incidence Function. *Biometrics*, **61**(1), 223–229.

Levantesi, S. and Menzietti, M. (2012). Managing longevity and disability risks in life annuities with long term care. *Insurance: Mathematics and Economics*, **50**(3), 391–401.

Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, **73**(1), 13–22.

References II

Meira-Machado, L., de Uña-Álvarez, J., and Cadarso-Suárez, C. (2006). Nonparametric estimation of transition probabilities in a non-Markov illness–death model. *Lifetime Data Analysis*, **12**(3), 325–344.

Murawska, M. and Rizopoulos, D. (2015). Simple analysis of non-Markov models: A case study on heart transplant data. *Statistical Modelling*, **15**(1), 51–69.

Nicolaie, M. A., van Houwelingen, J. C., de Witte, T. M., and Putter, H. (2013). Dynamic Pseudo-Observations: A Robust Approach to Dynamic Prediction in Competing Risks. *Biometrics*, **69**(4), 1043–1052.

Pan, W. (2001). Akaike's information criterion in generalized estimating equations. *Biometrics*, **57**(1), 120–125.

Pritchard, D. J. (2006). Modeling disability in long-term care insurance. *North American Actuarial Journal*, **10**(4), 48–75.

Tsai, W.-Y., Jewell, N. P., and Wang, M.-C. (1987). A note on the product-limit estimator under right censoring and left truncation. *Biometrika*, **74**(4), 883–886.