Long-Term Care and Longevity

Yang Lu [1] & Christian Gouriéroux [2]

[1] *SCOR and CREST, Email: yang.lu@normalesup.org*
[2] *CREST and University of Toronto, 15 Boulevard Gabriel Peri Malakoff 92245 France, Email: Christian.Gourieroux@ensae.fr*

Résumé. L'accroissement de la durée de vie est accompagné de l'augmentation du nombre de seniors en dépendance-incapacité. Compte tenu du coût élevé de la dépendance, il est important d'analyser l'évolution jointe de ces deux phénomènes, i.e. comment augmentera l'espérance de vie sans incapacité par rapport à l'espérance de vie? Ce papier répond à cette question quand les données de la dépendance sont manquantes ou non utilisables, mais seulement celles de mortalité aggrégée sont disponibles.

Nous présentons un modèle joint sur la dépendance et la mortalité, et expliquons pourquoi les paramètres sont identifiables avec uniquement les données aggrégées de mortalité, i.e. sans distinguer les personnes avec ou sans dépendance. Les modèles sont estimés sur les données de la population masculine française. D'abord un modèle déterministe est considéré, qui sera ensuite généralisé en un modèle avec facteur dynamique latent. La prévision, en fonction de l'année de naissance, de l'évolution de la probabilité d'entrer en dépendance durant sa vie, ainsi que de l'espérance de vie avec/sans incapacité est également fournie. Méthodologiquement ceci est une généralisation dynamique des modèles d'hétérogénéïté latente tels que les mélanges finis ou le frailty statique.

Mots-clés. Longevité, dépendance, risque semi-concurrents, hétérogénéité non observable, facteur dynamique non observable, identification.

Abstract. The increase of the expected lifetime, that is the longevity phenomenon, is accompanied by an increase of the number of seniors with a severe disability. Because of the significant costs of long-term care facilities, it is important to analyze the time spent in long-term care, as well as the probability of entering into this state during its lifetime, and how they evolve with longevity. Our paper considers such questions, when lifetime data are available, but long-term care data are either unavailable, or too aggregated, or unreliable, as it is usually the case.

We specify a joint structural model of long-term care and mortality, and explain why parameters of such models are identifiable from only the lifetime data. The methodology is applied to the mortality data of French males, first with a deterministic trend and then with a dynamic factor process. Various prediction formulas are then provided and illustrated using the same data.

Keywords. Longevity, Long-Term Care (LTC), Semi-Competing Risks, Unobserved Heterogeneity, Dynamic Frailty, Identification.

1 Introduction

Questions raised in the abstract are quite hard to answer, partly because of the competing risks nature of the two risks, that are, the risk of entering into long-term care (LTC) and the risk of dying directly. Besides, since the mortality of disabled people is higher than that of the autonomous people, when only the lifetime is observed, the health state at a given age[1] is an unobserved heterogeneity whose distribution changes over time and will cause a **spurious duration dependence**. This effect should be identified in order to study the true duration dependence, that is, the age dependence of the mortality evolution, and how this dynamics changes between different cohorts, that is, the longevity phenomenon. Therefore, it is essential to analyze the joint behavior of the two risks for the modeling and prediction of either the LTC, or the longevity, even when only lifetime is observed.

We introduce in this paper a joint model of LTC and mortality, based on the intensity of entry and on the mortality intensities. The model disentangles the mortality intensities according to the time spent in LTC. Moreover we assume that these intensities depend on unobservable dynamic factors (or dynamic frailties) with nonstationary features, able to capture the longevity phenomenon and its potential impact on both the mortality and the long-term care.

Such a joint model would be simple to estimate if individual data on both mortality and LTC were available. However there does not exist a universal definition of LTC, since the criteria of losing autonomy are quite vague and may differ both by country and insurance company. Another problem is that most datasets are cross-sectional, either by nature or because the observation period is too short to deliver longitudinal information.

Our paper will develop a methodology to estimate this joint model when the mortality data are the only available ones. The possibility to identify the characteristics of LTC from the mortality data is due to the jumps in mortality intensity arising when entering LTC. Our model allows us to predict jointly the future evolution of the LTC entry probabilities and the mortality rates.

The original paper can be downloaded in the SSRN website : `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2347735`.

2 Structural versus reduced form approach

We consider a situation where an individual can either experience first a **non terminal event** and then fail, or he/she can fail directly. In both situations the failure is called the **terminal event**. In the second case, the terminal event censors the non terminal event. The corresponding model is called the semi-competing risks model in the literature. In our framework, the non terminal event is the potential entering into LTC and the terminal

[1] Either autonomous, or in LTC, the second group can be further decomposed by the age of entry.

event is the death. The migration from the autonomous state to the LTC is assumed irreversible.

2.1 Structural approach

We define the folllowing latent variables:

- X_1 the potential time of the non terminal event,
- X_2 the (potential) time of death for an individual which has not experienced the non terminal event,
- X_3 the residual lifetime up to the death once the individual experienced the non terminal event.

Some of these variables are really latent even for an econometrician with the maximal available information. Indeed an individual dying before the potential time of the non terminal event will never experience spell X_1 or X_3. At most the observations include the indicator variable Z defined by: $Z = \mathbb{1}_{X_1 \leq X_2}$, that is, whether or not the individual experiences the non terminal event before the death, and the duration variable(s):

$$\begin{cases} Y_1 &= X_1 Z, \\ Y_2 &= (X_1 + X_3)Z + X_2(1-Z). \end{cases} \tag{1}$$

In regime 1, we observe the time Y_1^* up to the entry into LTC and the lifetime Y_2. In regime 0, we observe the lifetime only. The first equation corresponds to a standard Tobit model and is completed by an equation providing the observed lifetime depending on the regime. Under the assumption that $(X_1, X_3) \perp\!\!\!\perp X_2$, this structural approach can be written in an equivalent way as a multistate model, or the reduced form approach [see details in the paper].

2.1.1 Distribution of the (potentially) observable variables

Let us now derive the joint distribution of the potentially observable variables (Y_1, Y_2). The couple (Y_1, Y_2) has a bi-dimensional continuous component on domain $\mathcal{D}_1 = \{(y_1, y_2) : y_1 < y_2\}$, and a one-dimensional continuous component on $\mathcal{D}_0 = \{(y_1, y_2) : y_1 = 0, y_2 > 0\}$. The joint distribution of (Y_1, Y_2) admits a density with respect to the dominating measure $\lambda_{\mathcal{D}_1} + \lambda_{\mathcal{D}_0}$, where $\lambda_{\mathcal{D}}$ denotes the Lebesgue measure on domain $\mathcal{D}$. This density is:

$$f(y_1, y_2) = \lambda_1(y_1)\lambda_{2|1}(y_2 - y_1|y_1)e^{-\Lambda_1(y_1)-\Lambda_2(y_1)-\Lambda_{2|1}(y_2-y_1|y_1)}, \text{ on domain } \mathcal{D}_1 = \{(y_1, y_2) : y_1 < y_2\}, \tag{2}$$

and

$$f(0, y_2) = \lambda_2(y_2)e^{-\Lambda_1(y_2)-\Lambda_2(y_2)}, \text{ on domain } \mathcal{D}_0 = \{(y_1, y_2) : y_1 = 0, y_2 > 0\}. \tag{3}$$

We deduce the marginal survival function and the p.d.f. of the lifetime Y_2, which is later on the only really observable duration variable:

Proposition 1. *The survival function of the lifetime Y_2 is:*

$$\begin{aligned} S_2(y_2) &= \mathbb{P}(Y_2 > y_2) \\ &= \int_0^{y_2} \lambda_1(t) e^{-\Lambda_1(t)-\Lambda_2(t)-\Lambda_{2|1}(y_2-t|t)} dt + e^{-\Lambda_1(y_2)-\Lambda_2(y_2)}, \end{aligned} \tag{4}$$

and its p.d.f. is:

$$f_2(y_2) = \int_0^{y_2} \lambda_1(t)\lambda_{2|1}(y_2-t|t) e^{-\Lambda_1(t)-\Lambda_2(t)-\Lambda_{2|1}(y_2-t|t)} dt + \lambda_2(y_2) e^{-\Lambda_1(y_2)-\Lambda_2(y_2)}. \tag{5}$$

2.2 Identification in a model with constant intensities

Let us explain why it is possible to identify the parameters related to the intensity functions λ_1, λ_2, and $\lambda_{2|1}$ when none of them, but only the lifetime variable Y_2 is observed. For illustration purpose, we look at the special case when the intensities $\lambda_1, \lambda_2, \lambda_{2|1}$ are constant. We can prove that in both cases, the distribution of lifetime Y_2 is a mixture of two lifetime distributions. Then we have the following Proposition:

Proposition 2. *i) If $\lambda_1 + \lambda_2 - \lambda_{2|1} \neq 0$ and $\lambda_2 \neq \lambda_{2|1}$, the mixture representation has two distinct components and three parameters $\lambda_1, \lambda_2, \lambda_{2|1}$ can be identified from the distribution of lifetime Y_2.*

ii) If $\lambda_2 = \lambda_{2|1}$, the non terminal event has no effect on the mortality intensity. We get $S_2(y_2) = e^{-\lambda_{2|1} y}$. The parameter $\lambda_2 = \lambda_{2|1}$ is identifiable, but not the parameter λ_1.

iii) If $\lambda_1 + \lambda_2 - \lambda_{2|1} = 0$, the three parameters $\lambda_1, \lambda_2, \lambda_{2|1}$ can all be identified.

3 Statistical inference

3.1 Specification of a model without frailty

Let us first consider the basic model introduced in Section 3.1. We denote by $i, i = 1, ..., n$, the individuals and assume that the latent variables $X_{1,i}, X_{2,i}, X_{3,i}$, $i = 1, ..., n$ are independent, with a joint distribution, which depends on the generation only. We denote by $\lambda_1(x_1|t_0)$, $\lambda_2(x_2|t_0)$, $\lambda_{2|1}(x_3|x_1, t_0)$ the intensities for the individuals with the same birth date t_0. Then the individual lifetimes $Y_{2,i}, i = 1, ..., n$ are also independent with a distribution depending on t_0 only. The associated p.d.f. and survival functions are denoted $f_2(y_2; t_0)$ and $S_2(y_2; t_0)$, respectively. Taking into account the right censoring of the lifetimes, the log-likelihood function is:

$$\log l(Y_2, \theta) = \sum_{t_0} \Big\{ \sum_{i \in \eta^u_{t_0}} \log f_2(y_{2,i}, t_0, \theta) + \sum_{i \in \eta^c_{t_0}} \log S_2(y_{2,i}, t_0, \theta) \Big\}, \tag{6}$$

where $\eta^u_{t_0}$ (respectively $\eta^c_{t_0}$) is the set of uncensored (resp. censored) individuals in generation t_0, $y_{2,i}$ denotes either the observed failure time if the individual is not censored, or the censoring time, otherwise, and θ denotes the parameter.

Let us now specify the intensities by generation in a parametric way. We first disentangle the effects of the duration and of the current date in the intensities. More precisely we assume:

$$\begin{cases} \lambda_1(x_1|\underline{F}, t_0) & = a_1(x_1) + b_1(x_1)F_{t_0+x_1}, \\ \lambda_2(x_2|\underline{F}, t_0) & = a_2(x_2) + b_2(x_2)F_{t_0+x_2}, \\ \lambda_{2|1}(x_3|\underline{F}, x_1, t_0) & = a_3(x_3|x_1) + b_3(x_3|x_1)F_{t_0+x_1+x_3}, \end{cases} \tag{7}$$

where (F_t) is a deterministic function of time, such as:

$$F_t = \exp(-mt), \tag{8}$$

where $m > 0$. The factor is deterministic and known up to the value of parameter m.

The discussion of the parametric specification of the baseline hazard functions $a_1(\cdot), a_2(\cdot)$,$a_3(\cdot|\cdot)$, $b_1(\cdot), b_2(\cdot)$,$b_3(\cdot|\cdot)$ is provided in the original paper but omitted here. In next section's application, besides this general semi-Markov case, we will also consider the following special (Markov) case, where the intensity of X_3 given X_1 depends only on the current age x_1+x_3:

$$\begin{cases} a_3(x_3|x_1) & = a_3(x_3 + x_1), \\ \tilde{b_3}(x_3|x_1) & = \tilde{b_3}(x_3 + x_1). \end{cases} \tag{9}$$

3.2 Specification of a model with dynamic frailty

The above deterministic model is easily extended to introduce a stochastic dynamic frailty. The aim of introducing a stochastic factor is to quantify the uncertainty of both the model fit and the future evolution, which should be taken into account when it comes to pricing insurance contracts written on the LTC risk, and/or the longevity risk.

Instead of using a deterministic trend [equation (8)] for the log-factor $\log F_t$, we assume now a special Cox, Ingeroll, Ross (CIR) dynamics without the mean reversion coefficient:

$$dF_t = -mF_t dt + \sigma\sqrt{F_t}dW_t, \tag{10}$$

where $\sigma > 0$, $m > 0$, and W is a standard Brownian motion. The initial condition is $F_{\min t_0} = 1$, where $\min t_0$ is the birth date of the first cohort.

4 Application

The methodology is applied to a set of observations from the Human Mortality Database (HMD). We consider the French male population of males who survive up to age 50.

We estimated both the model with deterministic time factor and the model with dynamic frailty. For each model, once the parameters estimated, we can compare the model-implied values of the intensity function of Y_2 at each integer age, to the historical values of the data to look at the goodness of fit of the model in terms of the observed intensity. We can also plot the evolution of the implied prevalence for different cohorts, that is, the proportion of the people in LTC for each given age. The results of the estimations are provided in the full paper.

For the model with dynamic frailty, we can also infer the path of the unobserved frailty process F. This is useful for several reasons. First, after filtering out the unobserved frailty process, we can check the specification of its dynamics (CIR process), as well as the goodness of fit of the model in terms of observable mortality rates. Second, its values can be used for the predictions of the future mortality and of the LTC transition probability which depend on the frailty process.

This nonlinear filtering is based on simulations of the factor path after substituting the estimated parameters to their true values. Because of the strong nonlinearity of our mode, the simulation requires MCMC methods and is explained in details in the paper.

5 Prediction of individual risk

Now we can infer for each individual the value of the unobserved variables given the observed ones. We consider below an individual of cohort t_0 at calendar date $t_0 + y_2$ and assume that the available information set is $X_1 > y, X_2 > y$, that is, the individual is still autonomous. A special case is when $y = 50$, any individual enrolled in the study at this age is in good health at the beginning, and we are interested in the prediction of Y_1 and Y_2. First, let us compute the probability that a person will enter the LTC during his or her lifetime, given survival in good health up to age y. For each cohort, this probability is given by:

$$\mathbb{P}(Y_1 > 0 | X_1 > y, X_2 > y, t_0) = \frac{\int_y^\infty \lambda_1(x, t_0) e^{-\Lambda_1(x,t_0)-\Lambda_2(x,t_0)} dx}{e^{-\Lambda_1(y,t_0)-\Lambda_2(y,t_0)}}, \tag{11}$$

This probability is also called the cumulative incidence (at age $Y_2 = \infty$). Other interesting quantities include the residual life expectancy with (potential) LTC.

$$e_1(y) = \mathbb{E}[Y_2 - y | X_1 > y, X_2 > y],$$

as well as the residual life expectancy without LTC (or Healthy Life Years) defined by:

$$e_2(y) = \mathbb{E}[\min(X_1, X_2) - y | X_1 > y, X_2 > y].$$

Bibliographie

[1] Fine, J. P., Jiang, H., Chappell, R. (2001). On semi-competing risks data. Biometrika, 88(4), 907-919.
[2] Duffie, D., Eckner, A., Horel, G., Saita, L. (2009). Frailty correlated default. The Journal of Finance, 64(5), 2089-2123.
[3] Imai, K., Soneji, S. (2007). On the estimation of disability-free life expectancy: Sullivan's method and its extension. Journal of the American Statistical Association, 102(480), 1199-1211.
[4] Xu, J., Kalbfleisch, D., Tai, B. (2010). Statistical analysis of illness death processes and semicompeting risks data. Biometrics, 66(3), 716-725.
[5] Elbers, C., Ridder, G. (1982). True and spurious duration dependence: The identifiability of the proportional hazard model. The Review of Economic Studies, 49(3), 403-409.
[6] Duffie, D., Filipovic, D., Schachermayer, W. (2002). Affine processes and application in finance. Annals of Applied Probability.
[7] Andersen, P. K., Keiding, N. (2012). Interpretability and importance of functionals in competing risks and multistate models. Statistics in medicine, 31(11-12), 1074-1088.