

DÉPENDANCE DES PERSONNES ÂGÉES : UNE APPROCHE MULTI-ÉTATS BASÉE SUR LA NOTION DE PROCESSUS SEMI-MARKOVIENT

Guillaume Biessy

*SCOR Global Life SE - 5 avenue Kléber, 75795 Paris Cedex 16, France.
Laboratoire de Mathématiques et Modélisation d'Évry - Université d'Évry Val
d'Essonne, UMR CNRS 8071, USC INRA, Évry, France.
gbiessy@scor.com*

Résumé. La tarification des produits d'assurance dépendance se base aujourd'hui sur des modèles simples, où la dépendance est considérée comme un état unique et homogène. En raison du vieillissement de la population et des progrès rapides de la médecine, il est primordial d'acquiescer une vision plus claire de ce risque. Nous pensons que cet objectif peut être atteint en prenant en compte plusieurs niveaux de dépendance. Un processus multi-états est dit semi-markovien lorsque les probabilités de transition du processus dépendent à la fois de l'état actuel et du temps passé dans cet état. De tels processus s'avèrent plus flexibles que les processus markoviens simples, et ont fait l'objet de nombreuses publications dans le domaine de l'épidémiologie. Cependant, leur application à l'assurance dépendance est restée principalement théorique, en raison notamment du manque de données accessibles aux assureurs.

Cette communication a pour but de présenter la démarche de construction d'un modèle semi-markovien considérant 2 niveaux de dépendance. Ce travail s'appuie sur des données recueillies dans le cadre de l'Allocation Personnalisée d'Autonomie (APA). Tout d'abord, nous introduisons les paramètres intervenant dans la modélisation des transitions entre les états. Nous procédons ensuite à l'estimation de ces paramètres par la méthode du maximum de vraisemblance, en tenant compte des spécificités liées aux données APA. Enfin, nous proposons une application du modèle à la tarification d'un produit d'assurance dépendance fictif, à l'aide d'une méthode de type Monte Carlo.

Mots-clés. Assurance Dépendance, processus semi-markovien, APA, censure à droite, loi de Weibull, risques semi-proportionnels, modèle de mélange, maximum de vraisemblance, Monte Carlo

Abstract. The pricing of today's long-term care insurance products relies on simple models where dependency is considered as a single homogeneous state. Because of aging population and rapid evolution in the field of medicine, it becomes paramount to get a clearer picture of the underlying risk. We believe it may be achieved by taking into account several levels of dependency. A semi-Markov process is a multi-state process whose transition probabilities not only depend on the current state but also on the time spent in this state. This process has proven more flexible than the simple Markov process, and is core to numerous publications in the field of epidemiology. However its use in relation with long-term care insurance has remained mostly theoretical, mainly because of the lack of data available to insurers.

The present communication aims at introducing the construction process of a semi-Markov model with 2 levels of dependency. This work is based on data from the French long-term care public aid: the "Allocation Personnalisée d'Autonomie" (APA). Firstly, we introduce the parameters used to model transitions between states. We then proceed to the calibration of those parameters, using a likelihood maximization method, while taking into account the peculiarities of the APA data set. Finally, we apply this model to the pricing of a fictive long-term care insurance product, using a Monte Carlo method.

Keywords. Long-Term Care Insurance, semi-Markov process, APA, right censoring, Weibull law, semi-proportional hazard, mixture model, maximum likelihood, Monte Carlo.

Cette communication vise à présenter des travaux de recherche en cours portant sur la modélisation du processus de dépendance chez les personnes âgées. Ces travaux sont menés par l'auteur dans le cadre d'une thèse CIFRE liant SCOR Global Life, branche d'assurance vie du premier réassureur français, et le Laboratoire de Mathématiques et Modélisation d'Evry, qui a démarré en novembre 2013. La thèse est supervisée par Catherine Matias, directrice de recherche CNRS au Laboratoire Probabilités et Modèles aléatoires de l'UPMC pour le volet académique, et Vincent Lepez, responsable de la tarification au sein de SCOR Global Life, ayant lui même soutenu une thèse CIFRE, pour le volet opérationnel. L'auteur possède quant à lui un diplôme d'ingénieur de Télécom Bretagne, et le titre d'actuaire de l'Euro-Institut d'Actuariat (EURIA), obtenus en 2013.

Le phénomène de dépendance est au cœur de nombreux produits d'assurance, et la connaissance du risque sous-jacent est un enjeu important. En sa qualité de réassureur, SCOR a joué un rôle de pionnier dans la connaissance de ce risque dès la fin des années 1980 et il s'agit toujours aujourd'hui d'un axe stratégique majeur pour le groupe, qui réassure près de 2 des 6 millions de contrats Dépendance en France, et qui possède depuis 2001 un centre de Recherche & Développement entièrement consacré à l'étude du risque.

Aujourd'hui, les modèles utilisés par les assureurs considèrent la dépendance comme un état unique. Nous pensons qu'une modélisation multi-états semi-markovienne est nécessaire afin d'améliorer la connaissance du risque. Une telle modélisation a déjà été utilisée avec succès dans le domaine de l'épidémiologie, on pourra se référer par exemple à Mathieu (2006), et son application au phénomène de dépendance semble donc assez

naturelle. Cependant, les produits d'assurance dépendance sont encore très récents et les données disponibles demeurent limitées, ce qui explique pourquoi, en dépit de nombreuses publications théoriques, voire Janssen (2007) ou Christiansen (2012), les applications à des données réelles d'assurance sont aujourd'hui très rares.

Notre communication présentera la démarche de construction d'un modèle multi-états semi-markovien utilisant des données de l'aide publique française pour les personnes âgées dépendantes, l'Allocation Personnalisée d'Autonomie (APA). Une première étude de ces données a été réalisée par Lepez (2006). Ces données contiennent les évaluations du niveau de dépendance de plus de 30,000 bénéficiaires résidant dans 4 départements français sur les années 2002 à 2005 inclus. Elles sont censurées à droite et contiennent des valeurs manquantes. Les hypothèses majeures de notre travail comprennent le fait de pouvoir se débarrasser de l'effet stock constaté sur les données, lié au lancement de l'APA en 2002, à l'aide de troncature à gauche de l'échantillon, et de pouvoir assimiler les dates d'évaluations aux dates de transitions entre états de dépendance, en passant ainsi d'un problème de censure par intervalle à un problème de censure à droite, dont le traitement est beaucoup plus abordable. Nous réalisons l'ajustement d'un modèle semi-markovien à 2 états de dépendance utilisant des lois de durée de type Weibull, et intégrant des risques proportionnels, on pourra se référer à Cox (1984), afin de modéliser l'impact du sexe et de l'âge d'entrée en dépendance. Nous introduisons également une variable latente dans le modèle, pour tenir compte de l'hétérogénéité causée par les pathologies, responsables de la dépendance mais qui ne sont pas observées dans l'échantillon. L'estimation des paramètres du modèle se base sur la méthode du maximum de vraisemblance, adaptée pour tenir compte des spécificités des données. Une méthode de tarification, reposant sur des simulations, est finalement développée et appliquée pour la tarification d'un produit d'assurance dépendance fictif.

Bibliographie

- [1] Christiansen, M. C. (2012), *Multistate models in health insurance*, Advances in Statistical Analysis, 96(2), p155-186.
- [2] Cox, D. R. et Oakes, D. (1984), *Analysis of survival data*, CRC Press.
- [3] Janssen, J. et Manca, R. (2007) *Semi-Markov Risk Models for Finance, Insurance and Reliability*, Springer-Verlag New York Inc.
- [4] Lepez, V. (2006), *Trajectoires en Dépendance des personnes âgées : modélisation, estimation et application en assurance vie*, Mémoire d'Actuariat, Centre d'Etudes Actuarielles.
- [5] Mathieu, E. (2006), *Modélisations multi-états markoviennes et semi-markoviennes. Applications à l'état de santé des patients atteints par le virus du SIDA*, thèse de doctorat, Université de Montpellier.