



HAL
open science

Tables de mortalité best estimate: une approche de crédibilité pour les portefeuilles de petite taille

Yahia Salhi

► **To cite this version:**

Yahia Salhi. Tables de mortalité best estimate: une approche de crédibilité pour les portefeuilles de petite taille. *l'actuariat*, 2019, 31. hal-01997734

HAL Id: hal-01997734

<https://hal.science/hal-01997734>

Submitted on 4 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



RISQUE BIOMÉTRIQUE



Yahia SALHI

Maître de conférences à l'ISFA, université Lyon 1, et chercheur associé à la chaire Data Analytics & Models for Insurance (DAMI)

Tables de **mortalité** best estimate : une approche de crédibilité pour les **portefeuilles** de petite taille

L'approche permet de réviser, au fur et à mesure que de nouvelles observations arrivent, l'ajustement de Makeham ou encore d'un modèle de lissage semi-paramétrique tout en prenant en compte la bonne représentation par âge des effectifs.

Solvabilité II et IFRS 17 concourent à renforcer la nécessité pour les assureurs de disposer des meilleures hypothèses actuarielles relatives à leurs portefeuilles, eu égard à l'information disponible. Que cette information soit externe ou propre à leur expérience. S'agissant du risque de mortalité/longévité, une approche naturelle pour répondre à cette problématique consiste en l'utilisation des données disponibles au niveau de chaque portefeuille et de mener une démarche endogène de construction d'une table de mortalité spécifique. Par exemple, au moyen d'estimateurs non paramétriques tels que Kaplan-Meier ou Nelson-Aalen. Néanmoins un certain nombre d'acteurs se heurtent à des problématiques techniques liées à la taille de leurs portefeuilles. Ainsi, les taux de mortalité estimés présentent souvent des irrégularités et souffrent d'une forte instabilité due à la variance significative des estimateurs. Dans de tels contextes, il est délicat de construire des tables de mortalité au moyen de la seule expérience de chaque type de produit, de chaque portefeuille. D'autant plus que la révision dans le temps de celles-ci risque d'induire des impacts significatifs dans l'évolution des provisions techniques. Hardy et Panjer (1998) ont proposé un modèle de crédibilité (cf. Bühlmann & Gisler 2005) permettant une mise à jour des prévisions des décès agrégés pour chaque portefeuille, compte tenu de l'information émanant de l'ensemble des portefeuilles. Cependant, cette dernière ne tient pas compte de la distribution des effectifs par âge, ce qui est un point crucial compte tenu de la fréquente hétérogénéité du coût des prestations en cas de décès en fonction de l'âge (par exemple, en assurance des prêts ou encore pour des contrats d'épargne en UC avec garantie plancher).

Cette constatation a motivé les travaux récents de Salhi, Thérond et Tomas (2016) et Salhi et Thérond (2018), qui proposent une approche également basée sur la théorie de crédibilité mais qui prend en compte la structure par âge des populations (individuellement de taille restreinte).

Contourner **la forte incertitude**

Il s'agit de considérer l'information provenant de chaque portefeuille et de mesurer la confiance que l'on peut lui accorder. À titre d'illustration, dans Salhi, Thérond et Thomas (2016), un modèle paramétrique (Makeham) est ajusté à chaque portefeuille, intégrant la bonne représentation de la mortalité aux différents âges. Pour rappel, le modèle de Makeham (1887) propose pour chaque portefeuille i une forme en « S » pour la force de mortalité sous-jacente :

$$\mu_i(x) = A_i + B_i \times (C_i)^x,$$

où A_i , B_i et C_i sont des paramètres fixes à estimer à partir des observations. Pour les portefeuilles de petite taille, les estimateurs de ces paramètres présentent une forte variance. Pour contourner cette problématique, l'approche proposée requiert l'estimation des paramètres A_i et C_i sur une population de référence¹ et de caractériser chaque portefeuille par le seul paramètre B_i censé capter l'hétérogénéité de la population sous-jacente. En d'autres termes, le modèle considéré est le suivant :

$$\mu_i(x) = A_r + B_i \times (C_r)^x.$$

À chaque période t , le modèle ci-dessus est ajusté à chaque population et un nouveau paramètre $B_i(t)$ est estimé. Nous suivrons donc l'évolution dans le temps du ratio $X_i(t) = B_i(t)/B_r(t)$ représentant le niveau relatif de la mortalité du portefeuille i par rapport à la référence. En tenant compte de l'information jusqu'en T , l'objectif est donc de

50

1 - Par exemple, la population nationale du pays d'où sont issus les assurés ou la population agrégée de l'ensemble des portefeuilles.

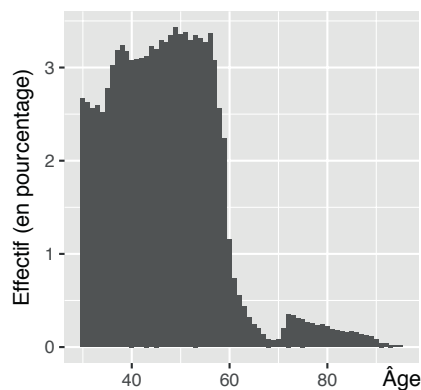
proposer la meilleure prédiction de l'évolution de ce ratio pour la période suivante, c'est-à-dire en $T+1$. Et d'en déduire, pour chaque portefeuille, la table best estimate qui en résulte. Dans cette optique, chaque portefeuille est caractérisé par un profil de risque Θ_i non observé et la prédiction best estimate peut s'interpréter comme l'espérance du ratio $X_i(T+1)$ conditionnellement à la connaissance du profil de risque Θ_i , c'est-à-dire la variable aléatoire $E[X_i(T+1)|\Theta_i]$. Cela invoque naturellement l'utilisation de la théorie de crédibilité afin de proposer une estimation adéquate dudit ratio tout en introduisant certaines hypothèses supplémentaires. En effet, les différents portefeuilles sont supposés être indépendants conditionnellement à la connaissance de leur profil de risque (hétérogénéité), et l'évolution de la moyenne du niveau relatif $X_i(t)$ d'un portefeuille donné est supposée être stable dans le temps. De plus, la variance de $X_i(t)$, compte tenu de l'hétérogénéité, est supposée évoluer en raison inverse de la taille de l'effectif de la population sous-jacente par rapport aux autres. Le poids de la population i , en présence d'autres portefeuilles, jouera un rôle déterminant pour mesurer la confiance que l'on accorde aux observations qu'elle engendre. En effet, la prédiction best estimate du niveau relatif, sous les conditions citées supra, prend la forme suivante : $E[X_i(T+1)|Q_i] = \alpha_i \bar{X}_i + (1-\alpha_i)$, où \bar{X}_i est le niveau moyen pondéré du ratio sur les périodes précédentes et α_i est un facteur mesurant la crédibilité accordée aux données du portefeuille i et évolue en fonction du poids du portefeuille concerné. Ainsi, la méthodologie préconise d'emprunter une information supplémentaire des autres portefeuilles avec une qualité mais aussi une quantité de données suffisamment importantes.

À titre d'exemple, le facteur de crédibilité est nul lorsque le poids du portefeuille l'est aussi. Dans ce cas précis, la prédiction pour la période $T+1$ du ratio est égale à 1. En d'autres termes, la table de référence est la meilleure estimation pour la mortalité de cette population. Pour résumer, cette méthode répond à une double problématique opérationnelle : la construction de lois d'expérience pour des portefeuilles de taille restreinte (pour lesquels une construction endogène n'est pas envisageable) et un processus de mise à jour automatique de ces hypothèses au fur et à mesure que de nouvelles observations sont disponibles (par exemple, de manière annuelle).

Tester la prédiction

L'application de cette méthodologie est illustrée sur des données collectées par le groupe de travail « mortalité » de l'IA. Il s'agit de 14 portefeuilles de tailles différentes allant

de 7 500 à presque deux millions d'assurés observés entre 2007 et 2011. Le portefeuille présenté est composé de 93 165 assurés répartis selon la figure ci-dessous.



Nous avons utilisé les données disponibles jusqu'en 2010 pour calibrer les différents modèles et proposons un ajustement en 2011. Cette prédiction est comparée à la réalisation des décès en 2011 pour mieux apprécier le pouvoir prédictif du modèle proposé. Ci-dessous, nous avons représenté les données de 2011 (en cercles noirs) ainsi que les prédictions faites par le modèle de Makeham ajusté (en vert) et celle proposée par le modèle Hardy-Panjer (en rose). Nous observons que le modèle de Makeham propose un meilleur ajustement et s'adapte mieux à la structure par âge de la mortalité. Des critères quantitatifs de proximité entre les observations et la prédiction confirment également que l'ajustement par crédibilité du modèle de Makeham intègre la bonne représentation aux différents âges.

Prédiction en 2011

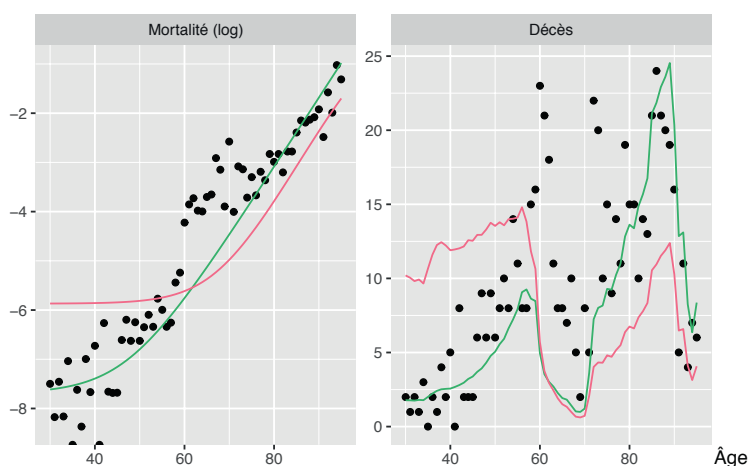


Figure 1 Prédiction en 2011 par l'ajustement du modèle de Makeham (ligne verte) et le modèle de Hardy-Panjer (ligne rose) et comparaison avec les données brutes (cercles noirs) de décès (à droite) et de taux de mortalité (à gauche).