LIGNES DIRECTRICES MORTALITE

Version approuvée

après consultation des membres de l'Institut des Actuaires

insérée en tant que recommandation dans les règles professionnelles de l'Institut par le Conseil d'Administration du 20 juin 2006

SOMMAIRE

Chapitre 1 : Validation des fichiers initiaux

La justesse d'une table d'expérience est avant tout conditionnée par la pertinence et la qualité des données utilisées pour sa construction. A ce titre, la validation des fichiers initiaux[1] est une étape incontournable du travail de l'actuaire certificateur et il est essentiel que celui-ci puisse accéder à l'ensemble des données utilisées, notamment au détail par police d'assurance.

Lors de la vérification des données initiales, l'actuaire certificateur peut se trouver confronté à une telle variété de situations qu'il serait difficile d'en donner une description exhaustive. Pour cette raison, nous avons identifié des questions qui se posent de façon récurrente d'un travail de certification à l'autre et les abordons ici afin de guider les actuaires certificateurs lorsqu'ils s'y trouveront confrontés.

I. La période d'observation

La période d'observation est l'intervalle de temps retenu pour l'étude de mortalité. Il est important de vérifier comment elle a été choisie.

Les déclarations tardives de sinistres

Il est évidemment souhaitable que la période d'observation inclue les années d'expérience les plus récentes possible. Cependant, pour tenir compte des déclarations tardives de décès et ne pas sous-estimer le risque de mortalité, il faut veiller à ce qu'un certain décalage ait été conservé entre la date de fin d'observation et la date d'extraction des données. L'idéal serait de disposer de données qui permettent d'estimer l'écart entre la date de survenance et la date de déclaration du décès de façon à valider le décalage utilisé.

La durée d'observation doit être un multiple de douze mois

Il est en général préférable que la période d'observation porte sur des intervalles d'une longueur multiple de douze mois car le risque de mortalité fluctue naturellement tout au long de l'année. Si cela n'a pas été possible, des ajustements peuvent s'avérer nécessaires pour corriger les estimations de ce phénomène de fluctuation.

La longueur de la période d'observation

En raison des fluctuations conjoncturelles de la mortalité, l'actuaire certificateur a intérêt à exiger, dans la mesure où cela est possible, que la période d'observation comporte plusieurs années.

Le choix d'une période de plusieurs années présente un deuxième avantage : cela permet d'augmenter le volume d'observations, ce qui est souvent utile pour améliorer la qualité des estimations. Il faut toutefois éviter de tomber dans l'excès inverse et ne pas accepter une période trop longue sans quelques précautions préalables.

Une période d'observation de trois à cinq ans paraît un choix raisonnable, une table construite avec une période différente devra faire l'objet d'une attention particulière. Dans tous les cas, il conviendra de vérifier qu'il n'y a pas d'évolution sensible de la mortalité sur la période retenue.

[1] Par « fichiers initiaux », nous désignons l'ensemble des fichiers de données, tels qu'ils ont été extraits du système de gestion, qui ont servi à construire la table de mortalité.

Pour une période d'observation relativement longue, il devient également nécessaire, dans les analyses de mortalité en montants[2], de contrôler l'évolution du taux d'intérêt technique et de l'inflation et, si nécessaire, de corriger les données pour éliminer l'impact de ces évolutions sur les estimations des probabilités de décès.

II. La représentativité des données

Le fichier initial doit être représentatif du portefeuille pour lequel la table de mortalité sera utilisée.

II.1 Le risque de biais dans l'extraction

Afin de s'assurer qu'il n'y a pas eu de biais dans la sélection des données retenues pour l'étude, l'actuaire certificateur a besoin de contrôler la méthode d'extraction employée. A cet effet, il est nécessaire qu'il sache comment le système de gestion informatique fonctionne (mode de stockage des données, durée de l'historique conservé...) et comment les enregistrements ont été sélectionnés.

Nous tenons à attirer l'attention des actuaires certificateurs sur les extractions filtrées qui, si elles sont mal utilisées, peuvent entacher la représentativité des données et produire des fichiers initiaux biaisés.

Exemple :
Sur la période d'observation qu'il a choisie, le constructeur de la table a conservé uniquement les contrats où il y a eu un sinistre ainsi que ceux qui sont encore en vigueur à la fin de la période d'observation.

- ⇨ Cette extraction est biaisée : les contrats qui ont pris fin au cours de la période d'observation sans survenance d'un sinistre n'ont pas été pris en compte. Ce biais conduit à surestimer le risque de décès.

Il nous semble également important d'attirer l'attention des actuaires certificateurs sur le cas des portefeuilles réassurés. Le niveau de mortalité sera en général différent selon que l'on affecte à chaque contrat le montant total assuré ou le montant net de réassurance. Dans ce dernier cas, un changement du programme de réassurance peut rendre la table de mortalité d'expérience caduque.

II.2 Utilisation de techniques d'échantillonnage

Nous recommandons à l'actuaire certificateur de vérifier que toutes les données disponibles ont bien été utilisées dans la construction de la table. Dans le cas contraire, il devra s'assurer de la pertinence de la méthode d'échantillonnage retenue.

Pour les fichiers volumineux, signalons cependant la technique statistique qui consiste à construire la table de mortalité sur une partie du fichier puis à la tester sur le reste des données. Cela revient, au final, à valider la table sur l'ensemble des données disponibles.

[2] Sous ce terme, nous désignons les études où chaque observation est pondérée par le montant assuré.

III. La connaissance de l'assuré : un élément essentiel

Certaines variables doivent naturellement être renseignées pour chaque enregistrement car, sans cette information, l'estimation même des probabilités de décès ne peut être réalisée. Il s'agit de :

- la date de naissance de l'assuré
- la date de début d'assurance
- le cas échéant, la date de fin d'assurance et la cause de sortie (décès, chute ou expiration naturelle du contrat)

Le fichier initial ne peut cependant pas se limiter à ces champs. Il faut également disposer de variables qui permettront de s'assurer de la cohérence des données et ainsi que de vérifier leur homogénéité.

Le numéro identifiant ou « numéro d'assuré » que l'assureur attribue souvent à chacun de ses clients est, lui aussi, une variable fondamentale. Il permet de faire le lien entre les différents contrats de l'assuré (communément appelés « doublons ») et s'avère extrêmement utile dans le cadre d'une étude de mortalité. Grâce à lui, on peut effectuer les regroupements nécessaires pour ne comptabiliser chaque assuré qu'une seule fois et lui affecter le capital assuré total. Cependant, cette variable n'est, en général, pas complètement fiable ; citons, par exemple, le cas d'un client qui achèterait de l'assurance dans des réseaux de distribution différents ou encore la fusion de systèmes de gestion qui pourrait engendrer une perte d'information.

Quel est l'intérêt d'éliminer les « doublons » ? La multi-comptabilisation d'un assuré ne devrait pas avoir une influence trop importante sur l'estimation des niveaux de mortalité. En revanche, elle conduirait à sous-estimer les intervalles de confiance. Elle pourrait également biaiser les estimations si le risque était segmenté en fonction du montant assuré.

IV. La cohérence des données

Une synthèse du travail de nettoyage des données ainsi qu'une analyse descriptive des données finalement conservées sont normalement remis à l'actuaire certificateur. Le rapport sur le nettoyage des données indique la proportion de polices éliminées et précise pour quelles raisons elles n'ont pas été conservées. L'analyse descriptive permet, quant à elle, de vérifier que la distribution des variables ne présente pas d'anomalie et de repérer les spécificités du portefeuille. Au-delà de ces documents d'étude, l'actuaire certificateur peut effectuer ses propres contrôles ou demander des analyses complémentaires. A cet égard toutes les variables peuvent être utiles. Nous soulignons ici l'importance des dates, de l'âge et du montant assuré.

De nombreuses dates-clé sont souvent stockées dans le système de gestion des assureurs : date de souscription de la police, date de prise d'effet de la couverture, date de décès, date d'enregistrement du décès, date de paiement de la garantie... Le contrôle de la bonne chronologie de ces dates donne un aperçu de la qualité du fichier. Il est également intéressant de vérifier si l'âge à la souscription et l'âge maximum de couverture correspondent bien aux limites fixées par l'assureur dans les conditions de souscription ou s'il s'agit d'une véritable dérogation. L'objectif de ces contrôles n'est pas d'éliminer les contrats souscrits en dérogation aux règles générales, mais d'identifier d'éventuelles erreurs de saisies afin qu'elles soient corrigées.

Dans un fichier initial de bonne qualité, les dates devraient contenir tant le jour et le mois que l'année. Si la connaissance des dates se limite aux années, l'actuaire certificateur est en droit de soupçonner soit des faiblesses dans les structures de gestion, soit un effort insuffisant dans la conduite des travaux d'analyse de la mortalité. La connaissance du système de gestion lui permettra d'affiner son jugement et d'en tirer ses conclusions en termes d'incertitude des estimations ou de besoin d'approfondissement des travaux.

D'autres variables peuvent permettre à l'actuaire certificateur de contrôler la cohérence des données. Il a, par exemple, intérêt à vérifier si le montant du capital garanti en cas de décès est bien compris dans les limites fixées par l'assureur et, dans le cas où il ne s'agirait pas d'une dérogation, de s'assurer que les erreurs de saisie ont été corrigées. Il peut aussi juger utile de croiser des données comptables avec celles issues du fichier d'étude afin de déterminer si le chiffre d'affaire annuel et les montants de sinistres payés sont cohérents (cela implique de disposer, pour chaque police, du montant de la cotisation, de sa périodicité et, éventuellement, des taux de revalorisation).

V. Les variables explicatives de la mortalité

Certaines variables permettent de déterminer l'assiette d'application de la table de mortalité. Même si elles n'interviennent pas directement dans le calcul des taux de mortalité et ne jouent pas de rôle majeur dans la vérification de la cohérence du fichier initial, elles sont tout aussi essentielles que les variables citées précédemment. Elles servent à tester l'homogénéité des données et, si une hétérogénéité est observée, elles sont utilisées pour mettre en place un suivi de la structure du portefeuille qui permettra de vérifier, dans le futur, que la table reste adaptée au risque couvert. Un des rôles majeurs de l'actuaire certificateur est, en effet, de contrôler l'assiette d'application de la table de mortalité et, une fois la certification établie, de surveiller son adéquation au portefeuille de l'assureur.

Si l'assureur a déjà constitué des segments tarifaires construits sur des différences de mortalité, l'actuaire certificateur a naturellement besoin de connaître la classification de chaque police. A titre d'exemple, citons la segmentation assez usuelle entre les assurés qui ont passé sans encombre la sélection (assurés à « risque normal ») et ceux qui se sont vus appliquer une surprime (assurés à « risque aggravé »). N'oublions pas non plus la segmentation, de plus en plus fréquente, entre fumeurs et non-fumeurs. Dans son travail de suivi, l'actuaire certificateur devra s'assurer qu'il n'est pas intervenu de changement majeur dans les méthodes de sélection et de segmentation utilisées par l'assureur.

Les variables qui peuvent influer sur le risque de décès sont nombreuses. Les critères les plus éprouvés sont certainement le sexe, le tabagisme ou non-tabagisme, la classification en risque normal ou aggravé, le montant assuré et le niveau de sélection médicale (contrats relevant du même processus de sélection). Il est également souvent très pertinent d'analyser le risque en fonction de la durée écoulée depuis l'adhésion pour mesurer l'impact de la sélection médicale ou de l'antisélection éventuelle. D'autres variables peuvent aussi s'avérer pertinentes, telles que, par exemple, la catégorie socioprofessionnelle de l'assuré, son lieu de résidence ou, encore, le mode de vente du produit d'assurance.

Chapitre 2 : Estimation brute des taux annuels de décès

I. Introduction et notations

I.1 Cadre de l'analyse

Pour estimer le taux annuel de décès à l'âge entier x, traditionnellement noté q_x, la première étape consiste à utiliser uniquement les individus observés vivants sur une partie ou sur la totalité de leur x+1ème année. Il n'est pas tenu compte des observations relatives aux âges voisins, qu'elles proviennent des mêmes individus ou d'individus différents. On parle alors d'estimation ***brute*** du taux q_x, par opposition aux estimations dites *lissées* qui, elles, prennent en compte un ensemble d'informations relatives à une plage d'âges.

Dans la suite, nous appelons Pop_x l'ensemble des individus utiles à l'estimation du taux q_x. L'indice i est employé pour désigner un individu de Pop_x (i = 1, 2, …N_x), où N_x représente la taille de la population Pop_x.

La première hypothèse, commune à l'ensemble des méthodes présentées ici, consiste à supposer que, sur l'intervalle d'âge [x,x+1], tous les décès éventuels des individus sont des événements indépendants et que les N_x individus de Pop_x ont la même probabilité de décéder dans leur x+1ème année (rappelons que, par définition de Pop_x, tous sont vivants à l'âge x).

Le problème soulevé dans la procédure d'estimation de la probabilité q_x tient au fait que les individus ne sont pas tous observables sur la totalité de la plage d'âge [x, x+1]. Le modèle binomial classique ne peut plus s'appliquer. De très nombreuses méthodes relevant pour la plupart d'hypothèses différentes sur la loi de répartition des décès sur la plage [x,x+1] et (ou) d'approximations construites sur l'ordre de grandeur du taux q_x ont été développées.

I.2 Données incomplètes

Un individu de Pop_x peut ne pas être pas être observable durant la totalité de sa x+1^{ième} année principalement pour deux raisons :

- soit il a fêté son x^{ième} anniversaire avant le début de la période d'observation et il rentre donc sous observation entre les âges x et x+1. On parle alors d'observation *tronquée à gauche* .
- soit son x+1^{ième} anniversaire (qu'il soit décédé ou vivant) est postérieur à la date de fin d'observation. Dans le cas où, à la fin de l'expérience (c'est-à-dire a posteriori), on constate que l'individu est vivant on parle d'observation *censurée à droite à âge prévisible a priori.*

Dans ces deux situations les âges au début et à la fin d'observation sont a priori connus et non aléatoires. Seul le décès (ou la survie) est aléatoire.

A ces deux raisons s'ajoute un troisième motif d'observation incomplète : l'individu, pour une cause non calendaire (résiliation du contrat, perte d'information, etc…), sort du champ d'observation vivant à un âge imprévisible (inférieur à l'âge x+1 bien sur). On parle alors de *censure à droite à âge imprévisible a priori*.

Il convient d'ores et déjà de porter une attention particulière à ces observations. En effet la cause de cette sortie vivant du champ d'observation est très souvent porteuse d'information sur la probabilité de décès de l'individu (par exemple une résiliation sur un contrat temporaire décès peut être un indice d'une meilleure santé de l'assuré). Comme nous le verrons par la suite, cette modification de la probabilité de décès invalide tous les estimateurs proposés dans la littérature.

Il conviendra, lors d'une mission de certification d'identifier ces situations, de contrôler (autant que faire se peut) que les causes de sorties n'influent pas sur la probabilité de décès et d'évaluer le poids de ces observations par rapport à l'effectif total N_x.

I.3 Notations

Rappelons que tout individu atteignant l'âge x+1 est considéré comme sortant de l'observation.

Notons $x + a_i$ l' âge de début d'observation de l'individu *i* sur la plage $[x, x+1[$.

Notons $x+b_i$ l'âge qu'aurait l'individu *i* en fin d'observation si il ne décédait pas et si il ne sortait pas de manière imprévisible de l'étude.

$0 \le a_i \le b_i \le 1$.

Notons $x+T_i$ la variable aléatoire « âge de l'individu *i* à sa sortie de l'observation ».

Quand l'individu sort de l'observation à un âge inférieur strictement à b_i on associe à cet évènement la cause de sortie, variable à deux modalités : « décès » et « autre ».

Enfin D_i représente la variable de Bernoulli indicatrice du décès de l'individu *i*

- $D_i = 0$ si la variable aléatoire T_i prend la valeur b_i ou si T_i prend une valeur inférieure strictement à b_i et la cause de sortie n'est pas le décès.
- $D_i = 1$ dans les autres cas.

$D_x = \sum_{i=1..N_x} D_i$ représente le nombre de décès dans la population Pop_x.

Le tableau ci-dessous récapitule la plupart des situations possibles :

Valeur de a_i	Valeur de b_i	réalisation t_i de T_i	Cause de sortie	Di	bilan
0	1	1	#	0	Observation complète et non décès
0	1	<1	décès	1	Observation complète et décès
> 0	1	1	#	0	Observation tronquée à gauche et non décès
> 0	1	<1	décès	1	Observation tronquée à gauche et décès
> 0	< 1	$t_i=b_i$	#	0	Observation tronquée à gauche et censurée à droite
0	< 1	$t_i=b_i$	#	0	Observation censurée à droite
0	< 1	$t_i<b_i$	décès	1	Observation potentiellement partielle et décès
0	1	<1	autre	0	Observation censurée à droite à âge imprévisible
0	< 1	$t_i<b_i$	autre	0	Observation censurée à droite à âge imprévisible

: signifie que l'assuré a été observé sur toute la période d'observation

II. Méthodes d'estimations

II.1 Généralités

Si toutes les observations étaient complètes et si aucune sortie inopinée n'était possible l'estimation de q_x relèverait du modèle binomial classique. Dans ce modèle $\hat{q}_x = \frac{D_x}{N_x}$ est à la fois l'estimateur du maximum de vraisemblance et celui de la méthode des moments.

Les très nombreuses extensions de cet estimateur aux données incomplètes peuvent se classer en deux familles selon qu'il s'agit d'une extension n'utilisant comme information relative au décès que la seule indicatrice de décès (on peut parler alors ***d'approche de type Bernoulli*)** ou d'une extension ***prenant aussi en compte l'âge[3] à la survenance de l'éventuel décès.***
A l'intérieur de ces deux familles d'estimateurs on peut encore identifier ceux qui relèvent de la ***méthode du maximum de vraisemblance***, ceux qui relèvent de la méthode des moments (labellisé dans ce contexte de ***méthode d'estimation cohérente***), voire ceux qui peuvent revendiquer au même titre que l'estimateur binomial les deux méthodes.
Pour les estimateurs de type Bernoulli la présence de données incomplètes conduit, quelque soit la méthode utilisée, y compris la méthode des moments qui dans le cas classique est non paramétrique, à introduire une hypothèse de répartition des décès sur la plage d'âges [x,x+1] donc un paramétrage de la loi que l'on pourrait qualifier de « local ».
Pour les estimateurs qui intègrent la connaissance exacte des âges au décès la méthode des moments retrouve sa vertu non paramétrique (***estimateur de Kaplan Meier***). On ne sera pas surpris que la méthode du maximum de vraisemblance pour être opérationnelle dans sa phase d'optimisation demande elle aussi une hypothèse de répartition des décès sur la plage d'âges [x,x+1].

Le plus souvent les estimateurs obtenus n'ont pas d'expression explicite mais sont évalués par des techniques de résolution numérique au demeurant fort simples. Pour pallier cet inconvénient et obtenir des formules explicites des approximations sont proposées, construites sur le fait que le taux estimé q_x est « petit ». Etant explicites ces approximations sont traditionnellement les plus connues voire les plus utilisées. Sans les remettre en cause, il convient de bien contrôler leur domaine de validation qui peut être contesté notamment :

- dans leur application aux âges élevés où le taux de mortalité devient plus important. L'obtention numérique sans approximation de l'estimateur est de ce point de vue incontournable.
- dans le cumul, sur tous les âges des erreurs d'approximation qui sont toujours de même signe et qui sur des applications à des produits de rente peuvent finir par « peser » sur les résultats.

Par contre ces approximations sont particulièrement utiles pour le calcul des variances (et donc des intervalle de confiance) des estimateurs. Le plus souvent ce calcul n'est d'ailleurs possible qu'à partir des solutions explicites et par ailleurs l'imprécision due aux approximations est nettement moins pénalisante.

Lors d'une mission de certification il pourra être utile de bien identifier de quelle famille l'estimateur proposé s'inspire, si il s'agit d'une formule exacte ou approchée, voire quand les données sont disponibles d'élaborer des estimateurs « concurrents » et de confronter les résultats.

Les méthodes présentées ci-dessous ne doivent pas être vues comme un ensemble de « figures imposées ». Bien au contraire elles doivent permettre au certificateur de garder un esprit ouvert devant telle ou telle variante suggérée par les conditions d'expérience. Les exemples d'estimateurs présentés ci-après ne prétendent pas à l'exhaustivité.

[3] nous faisons ici référence à l'âge exact par opposition à l'âge arrondi à un entier

II.2 Estimateurs de type Bernoulli

II.2.a Méthode du maximum de vraisemblance

Si l'observation *i* ne relève pas d'une censure à droite à âge imprévisible, sa vraisemblance est égale à :

$$\begin{cases} {}_{b_i-a_i}q_{x+a_i} \text{ en cas de décès} \\ 1-{}_{b_i-a_i}q_{x+a_i} = {}_{b_i-a_i}p_{x+a_i} \text{ en cas de survie à l'age } x+b_i \end{cases}$$

Si l'observation *i* est censurée à droite à l'âge imprévisible $x+t_i$, de très nombreuses modélisations probabilistes relatives à la loi du couple (âge à la sortie, cause de sortie) ont été développées. Nous nous limitons ici à proposer un raisonnement d'inférence où l'âge $x+t_i$ est supposé connu (mais la censure à droite, c'est-à-dire la sortie « vivant », reste elle aléatoire). La vraisemblance d'une telle observation est alors égale à la probabilité d'être vivant à l'âge $x+t_i$ (soit ${}_{t_i-a_i}p_{x+a_i}$) et ce cas n'a plus lieu d'être distingué des censures à droite à âge prévisible. ***Cependant l'écriture de cette vraisemblance suppose que la cause de censure n'influe pas sur la mortalité de l'individu. Cette hypothèse peut, comme nous l'avons signalé en préambule, être assez souvent contestée.***

La vraisemblance sur l'ensemble de la population Pop_x s'obtient par le produit des vraisemblances de chaque observation et la log-vraisemblance est alors égale à :

$$\sum_{i=1..N_x} D_i \times ln({}_{b_i-a_i}q_{x+a_i}) + (1-D_i) \times ln({}_{t_i-a_i}p_{x+a_i}) \quad (1)$$

Pour estimer le taux q_x, il est indispensable de choisir un modèle permettant d'exprimer, pour tout couple (a,b) tel que 0<a<b<1, les probabilités ${}_{b-a}q_{x+a}$ et ${}_{b-a}p_{x+a}$ en fonction de la probabilité q_x. Chaque modèle retenu conduit à une expression différente de l'estimateur maximisant la vraisemblance ou, de manière équivalente, de la log-vraisemblance. Cette expression peut être explicite (rarement), avoir une expression approximative explicite ou être obtenue par des méthodes de résolution numérique en général très simples.

II.2.b Méthode des moments. Notion d'estimateur cohérent.

Dans le modèle binomial l'estimateur $\hat{q}_x$ vérifie : $\hat{q}_x.N_x = D_x$. Ainsi le nombre de décès espéré $\hat{q}_x.N_x$ par l'estimateur $\hat{q}_x$ est égal au nombre de décès D_x effectivement observés. L'extension au cas des données incomplètes est, avec cet éclairage, naturelle. Supposons (comme pour la méthode du maximum de vraisemblance) qu'un modèle de répartition des décès sur l'intervalle $[x,x+1]$ ait été choisi. On peut donc exprimer, pour tout couple (a,b), 0<a<b<1, la probabilité ${}_{b-a}q_{x+a}$ comme fonction de q_x notée : ${}_{b-a}q_{x+a} = {}_{b-a}q_{x+a}(q_x)$

$\hat{q}_x$ **est dit estimateur cohérent si** le nombre de décès espéré par l'estimateur $\hat{q}_x$, soit $\sum_{i=1..N_x} {}_{b_i-a_i}q_{x+a_i}(\hat{q}_x)$, est égal à la somme :

- du nombre de décès observés a posteriori, D_x,
  et
- du nombre de décès espéré par l'estimateur et inobservés à cause du phénomène de censure droite, $\sum_{i=1..N_x, t_i<b_i; \text{sortie vivant}} {}_{b_i-t_i}q_{x+t_i}(\hat{q}_x)$.

En remarquant que, si t_i est égal à b_i, la probabilité ${}_{b_i-t_i}q_{x+t_i}(\hat{q}_x)$ est évidemment nulle, la restriction $t_i<b_i$ peut être omise.

Finalement $\hat{q}_x$ cohérent vérifie :

$$\sum_{i=1..N_x} {}_{b_i-a_i}q_{x+a_i}(\hat{q}_x) = D_x + \sum_{i=1..N_x,} (1-D_i).{}_{b_i-t_i}q_{x+t_i}(\hat{q}_x) \quad (2)$$

Remarque :

Certains auteurs propose une variante de la relation de cohérence en raisonnant pour chacun des individus de Pop_x sur l'intervalle $[x+a_i,x+1]$ au lieu de l'intervalle $[x+a_i,x+b_i]$

La relation de cohérence s'écrit alors :

$$\sum_{i=1..N_x} {}_{1-a_i}q_{x+a_i}(\hat{q}_x) = D_x + \sum_{i=1..N_x} (1-D_i).{}_{1-t_i}q_{x+t_i}(\hat{q}_x) \quad (3)$$

II.3 Estimateurs prenant en compte l'âge au décès

II.3.a Méthode du maximum de vraisemblance.

La vraisemblance des observations où les individus sont vivants à la fin de l'observation est inchangée. Par contre la vraisemblance des individus décédés n'est plus une vraisemblance de type Bernoulli mais la densité de la loi de la variable aléatoire $x+T_i$, « âge au décès », conditionnée par l'évènement $T_i > a_i$ qui peut aussi s'écrire : ${}_{t_i-a_i}p_{x+a_i} \times \mu(x+t_i)$ où $\mu(x+t_i)$ représente le taux de hasard de la loi de l'âge au décès.

Ici aussi une hypothèse sur la loi de répartition des décès sur la plage $[x,x+1]$ est indispensable. Elle permet de relier le taux q_x aux vraisemblances ${}_{t_i-a_i}p_{x+a_i}$ et ${}_{t_i-a_i}p_{x+a_i} \times \mu(x+t_i)$.

La log-vraisemblance s'écrit alors :

$$\sum_{i=1...N_x} D_i \times \ln\left(\mu(x+t_i)\right) + \sum_{i=1...N_x} \ln\left({}_{t_i-a_i}p_{x+a_i}\right) \quad (4)$$

II.3.b Méthode des moments. Estimateur de Kaplan Meier

Dans le cas classique de données complètes, la loi de répartition de la variable Min (T-x,1) est estimée par la distribution empirique (à l'origine de la méthode des moments) qui prend uniquement les valeurs s_i-x et 1. La probabilité de la valeur 1 est alors la fréquence des survivants $\frac{N_x - D_x}{N_x}$, celle des autres points est égale à $1/N_x$. L'estimation de q_x (égale à $\frac{D_x}{N_x}$) reste alors inchangée et cette approche ne rajoute rien au modèle binomial.

En présence de données incomplètes, l'extension naturelle de la fonction de survie empirique permet de proposer une estimation du taux q_x sans faire aucune hypothèse sur la loi de répartition des décès sur $[x,x+1]$. Nous décrivons ci-dessous l'algorithme conduisant à la fonction de survie de Kaplan Meier (on pourra trouver en annexe une bibliographie relative à cet estimateur).

(i) Fonction d'exposition au risque :

On appelle fonction d'exposition au risque la fonction notée *expo*(*t*) et définie par :

$$\text{expo}(t) = \sum_{i=1...N_x} 1_{a_i \le t} - \sum_{i=1...N_x} 1_{s_i \le t}$$

Cette fonction compte le nombre d'individus soumis au risque de décès à l'âge x+t et sous observation.

Un individu entrant ou sortant par censure en t est conventionnellement considéré comme exposé au risque de décès. Ce choix est sans importance pour toutes les valeurs de t qui

n'enregistrent pas de décès. Par contre, l'expression de *expo*(*t*) aux âges de décès étant cruciale, cette subjectivité peut être gênante dans la mesure ou une autre convention ne donne pas la même expression pour la fonction de survie Kaplan Meier. C'est la raison pour laquelle il est recommandé de choisir l'unité de temps la plus fine possible afin de réduire au minimum les situations ou entrées, censures et décès se produisent simultanément.

(ii) La fonction de survie (notée S_{KM}) de Kaplan-Meier relève alors de l'algorithme récurrent décrit ci-dessous :

a- $S_{KM}(0)=1$

b- Sur tout intervalle (*a,b*) tel que, entre *x*+*a* et *x*+*b*, on n'observe pas d'âge au décès, la fonction S_{KM} reste constante.

c- Pour tout âge de décès *x*+*t* :

$S_{KM}(t+) = S_{KM}(t-) \times \left[1 - \frac{n(t)}{expo(t)}\right]$ où *n(t)* représente le nombre de décès observés à l'âge *x*+*t* (en général n(t) est égal à 1).

La justification intuitive de cet algorithme vient du fait que toute fonction de survie S vérifie, pour tout couple (*y,y'*) *y*<*y'* : $S(y') = S(y) \times (1 - {}_{y'-y}q_y)$ et que la probabilité de décès dans un intervalle (*t-,t+*) est estimée classiquement par la fréquence de décès observés soit *n(t)/expo(t).*

L'estimateur de q_x s'obtient à partir de la fonction de survie de Kaplan Meier par la relation : $\hat{q}_{KM}(t) = 1 - S_{KM}(t)$.

La fonction de survie Kaplan Meier présente de nombreux avantages :

- Pas d'hypothèse a priori sur la répartition des décès sur [x,x+1].
- Facilité de programmation
- Possibilité d'obtention d'intervalles de confiance.
- Possibilité, si besoin est, d'estimer le taux ${}_{b-a}q_{x+a}$ par $1 - \frac{S_{KM}(b)}{S_{KM}(a)}$
- Possibilité si les capacités informatiques le permettent, de traiter simultanément l'ensemble de la population et d'obtenir ainsi les estimations des q_x pour tous les âges *x* sans avoir à isoler pour chaque âge *x* la population Pop_x.

L'obtention de la fonction de survie Kaplan Meier doit être une exigence de la part du certificateur dès lors que les données permettent son calcul. En même temps, celui-ci est tenu de vérifier que l'unité de temps choisie a été la plus fine possible et, en cas d'abondance d'ex aequo, de vérifier que les résultats restent très peu sensibles aux choix des conventions permettant le calcul de la fonction d'exposition au risque.

III. Exemples d'estimateurs construits sur un modèle de répartition des décès sur la plage [*x,x*+1]

III.1 Expression de la probabilité ${}_{b-a}q_{x+a}$ en fonction de q_x

Pour exprimer les probabilités ${}_{b-a}q_{x+a}$ et ${}_{b-a}p_{x+a}$ il est indispensable de préciser la loi de répartition de l'âge au décès sur la plage d'âges [x,x+1] sachant qu'il y a eu décès sur cet intervalle. Cette loi conditionnelle peut se déduire d'hypothèses faites sur la loi non conditionnelle de l'âge au décès (par exemple : modèle de taux de hasard constant sur [*x,x*+1], modèle de Balducci) mais l'hypothèse peut aussi se faire directement sur la loi conditionnelle elle-même (par exemple : loi uniforme sur [*x,x*+1]).
Dans tous les cas les formules ci-dessous peuvent être très utiles :

$$ {}_{b-a}q_{x+a} = \frac{F_C(b) - F_C(a)}{1 - q_x F_C(a)} q_x \text{ et } {}_{b-a}p_{x+a} = \frac{1 - q_x F_C(b)}{1 - q_x F_C(a)} $$

où F_c désigne la fonction de répartition conditionnelle de la différence entre l'âge au décès et l'âge x sachant que le décès intervient entre les âges x et x+1.

A partir de ces relations, on peut construire une multitude d'estimateurs, qui varieront en fonction de l'hypothèse faite sur la loi de répartition des décès, de la méthode utilisée (maximum de vraisemblance ou méthode des moments) et de l'approximation utilisée si le taux q_x est considéré comme petit. Nous n'avons pas l'intention d'en proposer ici un catalogue exhaustif.

Nous présentons ci-après, pour chaque hypothèse de répartition des décès, un tableau synthétique donnant :

1. la définition mathématique de l'hypothèse
2. la densité de la loi de l'âge au décès sur [*x*,*x*+1]
3. le taux de hasard de la loi de l'âge au décès sur [*x*,*x*+1]
4. la loi conditionnelle de *S*-*x* sachant $S \in [x, x+1]$
5. l'expression de ${}_{b-a}q_{x+a}$
6. son approximation
7. l'équation conduisant à l'estimateur maximum de vraisemblance de type Bernoulli obtenue par dérivation de la log-vraisemblance
8. et sa solution approchée
9. l'équation conduisant à l'estimateur cohérent (2)
10. et sa solution approchée
11. l'équation conduisant à l'estimateur cohérent (3)
12. et sa solution approchée
13. l'équation conduisant à l'estimateur maximum de vraisemblance utilisant les âges au décès
14. et sa solution approchée

III.2 Hypothèse de répartition uniforme :

Définition	*densité de loi de S constante sur [x,x+1]*	#
Densité de loi de S sur [x,x+1]	$\prod_{i=0..x-1} p_i \times q_x$	#
Taux de hasard	$\frac{q_x}{1-tq_x}$ *en x+t*	*croissant sur tout intervalle [x,x+1]. L'individu vieillit*
Loi conditionnelle de S-x	*loi uniforme sur [0,1]*	*ne dépend plus de q_x*
$_{b-a}q_{x+a}$	$\frac{(b-a)}{1-aq_x}q_x$	#
approximation de $_{b-a}q_{x+a}$	$(b-a)q_x$	*relation classique mais qui sous évalue la probabilité.*
équation MV de type Bernoulli	$\frac{D_x}{q} = \sum_{i=1..N_x}\left[(1-D_i)\frac{t_i}{1-t_iq} - \frac{a_i}{1-a_iq}\right]$	*résolution numérique*
solution approchée	$\hat{q}_x = \frac{D_x}{\sum_{i=1...N_x}(1-D_i)(t_i-a_i) - \sum_{i=1...N_x} D_i a_i}$	*supérieure à la solution exacte*
équation de cohérence (2)	$\frac{D_x}{q} = \sum_{i=1..N_x}\left[\frac{b_i-a_i}{1-a_iq} - \frac{(1-D_i)(b_i-t_i)}{1-t_iq}\right]$	*résolution numérique*
solution approchée	$\hat{q}_x = \frac{D_x}{\sum_{i=1...N_x}(b_i-a_i) - \sum_{i=1...N_x}(1-D_i)(b_i-t_i)}$	*supérieure à la solution exacte*
équation de cohérence (3)	$\frac{D_x}{q} = \sum_{i=1..N_x}\left[\frac{1-a_i}{1-a_iq} - \frac{(1-D_i)(1-t_i)}{1-t_iq}\right]$	*résolution numérique*
solution approchée	$\hat{q}_x = \frac{D_x}{\sum_{i=1..N_x}(1-D_i)(t_i-a_i) + \sum_{i=1..N_x} D_i(1-a_i)}$	*supérieure à la solution exacte*
équation MV avec âges au décès	$\frac{D_x}{q} = \sum_{i=1..N_x}\left[(1-D_i)\frac{t_i}{1-t_iq} - \frac{a_i}{1-a_iq}\right]$	*coincide avec MV Bernoulli*
Solution approchée	$\hat{q}_x = \frac{D_x}{\sum_{i=1...N_x}(1-D_i)(t_i-a_i) - \sum_{i=1...N_x} D_i a_i}$	*coincide avec MV Bernoulli*

III.3 Hypothèse de taux de hasard constant :

Définition	*taux de hasard de S constant sur [x,x+1]*	#
densité de loi de S sur [x,x+1]	$-ln(p_x)S(x)e^{+ln(p_x)t}$ *en x+t*	#
taux de hasard	$\lambda = -ln(p_x)$ *sur [x,x+1]*	*L'individu ni ne vieillit ni ne rajeunit sur [x,x+1]*
loi conditionnelle de S-x sachant $S \in [x, x+1]$	$S_C(t) = \frac{p_x^{\ t} - p_x}{q_x} = \frac{e^{-\lambda t} - e^{-\lambda}}{1 - e^{-\lambda}}$	*dépend de q_x et donc de λ. Préférer λ pour les résolutions*
${}_{b-a}q_{x+a}$	$1 - p_x^{\ b-a} = 1 - e^{-\lambda(b-a)}$	#
approximation de ${}_{b-a}q_{x+a}$	$(b-a)\lambda \approx (b-a)q_x$	*commune aux trois modèles*
équation MV de type Bernoulli	$\sum_{i=1..N_x} \frac{D_i(b_i - a_i)}{e^{\lambda(b_i - a_i)} - 1} = \sum_{i=1..N_x} (1 - D_i)(t_i - a_i)$	*résolution numérique*
solution approchée	$\hat{\lambda} = \frac{D_x}{\sum_{i=1..N_x} (1 - D_i)(t_i - a_i)}$; $\hat{q}_x = 1 - e^{-\hat{\lambda}}$	*supérieure à la solution exacte*
équation de cohérence (2)	$D_x + \sum_{i=1..N_x} (1 - D_i)(1 - e^{-\lambda(b_i - t_i)}) = \sum_{i=1..N_x} (1 - e^{-\lambda(b_i - a_i)})$	*résolution numérique*
solution approchée	$\hat{\lambda} = \frac{D_x}{\sum_{i=1..N_x} (1 - D_i)(t_i - a_i) + D_i(b_i - a_i)}$; $\hat{q}_x = 1 - e^{-\hat{\lambda}}$	*supérieure à la solution exacte*
équation de cohérence (3)	$D_x + \sum_{i=1..N_x} (1 - D_i)(1 - e^{-\lambda(1 - t_i)}) = \sum_{i=1..N_x} (1 - e^{-\lambda(1 - a_i)})$	*résolution numérique*
solution approchée	$\hat{\lambda} = \frac{D_x}{\sum_{i=1..N_x} (1 - D_i)(t_i - a_i) + D_i(1 - a_i)}$; $\hat{q}_x = 1 - e^{-\hat{\lambda}}$	*supérieure à la solution exacte*
équation MV avec âges au décès	$\frac{D_x}{\lambda} = \sum_{i=1..N_x} (t_i - a_i)$	***solution explicite*** $\hat{\lambda} = \frac{D_x}{\sum_{i=1...N_x} (t_i - a_i)}$
solution approchée	$\hat{\lambda} = \frac{D_x}{\sum_{i=1..N_x} t_i - a_i}$; $\hat{q}_x = 1 - e^{-\hat{\lambda}}$	***pas d'erreur***

III.4 Hypothèse de Balducci :

Définition	$_{1-t}q_{x+t} = (1-t)q_x$	#
Densité de loi de S sur [x,x+1]	$\frac{S(x+1)q_x}{(p_x + tq_x)^2}$	#
Taux de hasard	$\frac{q_x}{p_x + tq_x}$ *en x+t*	*L'individu rajeunit sur [x,x+1] !!!*
Loi conditionnelle de S-x sachant $S \in [x, x+1]$	$S_C(t) = \frac{(1-t)p_x}{p_x + tq_x}$	*dépend encore de* q_x
$_{b-a}q_{x+a}$	$\frac{(b-a)q_x}{p_x + bq_x}$	#
approximation de $_{b-a}q_{x+a}$	$(b-a)q_x$	*commune aux trois modèles*
équation MV de type Bernoulli	$\frac{D_x}{q} = \sum_{i=1..N_x}\left[(1-D_i)\frac{t_i - a_i}{(p + a_i q)(p + t_i q)} - D_i\frac{1-b_i}{p+b_i q}\right]$	*résolution numérique*
solution approchée	$\hat{q}_x = \frac{D_x}{\sum_{i=1...N_x}(1-D_i)(t_i - a_i) - \sum_{i=1...N_x} D_i(1-b_i)}$	*supérieure à la solution exacte*
équation de cohérence (2)	$\frac{D_x}{q} = \sum_{i=1..N_x}\left[\frac{b_i - a_i}{p_x + b_i q_x} - (1-D_i)\frac{b_i - t_i}{p_x + b_i q_x}\right]$	*résolution numérique*
solution approchée	$\hat{q}_x = \frac{D_x}{\sum_{i=1..N_x}(b_i - a_i) - \sum_{i=1..N_x}(1-D_i)(b_i - t_i)}$	*supérieure à la solution exacte*
équation de cohérence (3)	$\frac{D_x}{q} = \sum_{i=1..N_x}\left[1 - a_i - (1-D_i)(1-t_i)\right]$	***solution explicite***
solution approchée	$\hat{q}_x = \frac{D_x}{\sum_{i=1..N_x}(1-D_i)(t_i - a_i) + \sum_{i=1..N_x} D_i(1-a_i)}$	***pas d'erreur***
équation MV avec âges au décès	$\frac{D_x}{q} = \sum_{i=1..N_x}\left[\frac{t_i - a_i}{(p + a_i q)(p + t_i q)} - D_i\frac{1-t_i}{p+t_i q}\right]$	*résolution numérique*
solution approchée	$\hat{q}_x = \frac{D_x}{\sum_{i=1..N_x}(t_i - a_i) - \sum_{i=1..N_x} D_i(1-t_i)}$	*supérieure à la solution exacte*

III.5 Bilan

1° - Pour les trois modèles présentés, qui sont les modèles les plus fréquemment utilisés, il y a uniquement deux scénarios où l'estimateur prend une forme explicite :

a- Modèle taux de hasard constant, estimateur maximum de vraisemblance avec prise en compte des âges au décès dont nous rappelons l'expression :

$$\hat{q}_x = 1 - exp(-\frac{D_x}{\sum_{i=1..N_x} t_i - a_i})$$

Ce couple modèle-méthode est donc particulièrement intéressant. Il convient de noter néanmoins que cet estimateur est biaisé (asymtotiquement sans biais) et que le modèle de taux de hasard constant exprime que l'individu a une potentialité de décéder constante sur l'intervalle [*x,x*+1] ce qui, pour des âges élevés, peut être une hypothèse contestée.

Cet estimateur est également obtenu par un modèle dit Poissonien qui suppose que nombre total de décès, dont on suppose qu'il suit une loi de Poisson de paramètre $\lambda_x . \sum_i (t_i - a_i)$ dont la justification est empruntée à la théorie des processus de Poisson et au fait que $\sum_i (t_i - a_i)$ représente un exposition « moyenne » au risque de décès.

b- Modèle de Balducci, estimateur cohérent deuxième version.

$$\hat{q}_x = \frac{D_x}{\sum_{i=1..N_x} (1 - D_i)(t_i - a_i) + \sum_{i=1..N_x} D_i (1 - a_i)}$$

Pour ce couple modèle-méthode il reste intéressant d'avoir une expression explicite. Cependant l'hypothèse de Balducci a pour conséquence que sur la plage [x,x+1] l'individu a une potentialité de décéder qui diminue !!

2°- Les solutions approchées sont toutes égales à des termes de second ordre près (par rapport à q_x). Elles ne diffèrent que dans l'expression des dénominateurs. Ces dénominateurs s'interprètent en termes d'exposition au risque et diffèrent essentiellement sur la prise en compte, dans cette exposition, des individus décédés. Ces constatations se comprennent aussi en remarquant que les équivalents des taux ${}_{b-a}q_{x+a}$ sont tous égaux à *(b-a)*q_x.

Ces solutions approchées sur évaluent toujours la solution associée exacte. Il semble donc important que le certificateur puisse avoir connaissance des solutions exactes afin d'évaluer la pénalisation éventuellement due à l'utilisation de la solution approchée au moins pour le modèle « répartition uniforme des décès ». Rappelons que nous préconisons aussi la comparaison de ces estimateurs à l'estimateur Kaplan Meier.

3°- **Pour le calcul des intervalle de confiance,** qui comme nous l'avons dit est moins exigeant en précision, l'estimateur « traditionnel » peut, de notre point de vue, suffire :

$$\hat{q}_x = \frac{D_x}{\sum_{i=1..N_x} (t_i - a_i)} = \frac{\sum_{i=1..N_x} D_i}{\sum_{i=1..N_x} (t_i - a_i)}$$

Cet estimateur repose sur l'approximation commune: ${}_{b-a}q_{x+a} = (b-a)q_x$.
Son dénominateur est en fait équivalent à tous les dénominateurs obtenus précédemment et a l'immense avantage de ne pas être aléatoire.

Il est, à l'approximation près non biaisé, de variance : $\frac{q_x \sum (t_i - a_i) \times (1 - [t_i - a_i] q_x)}{\left[\sum (t_i - a_i)\right]^2}$ elle-même estimée en substituant dans l'expression ci-dessus q_x par son estimation.
Les intervalle de confiance se déduisent à partir du comportement gaussien supposé pour $\hat{q}_x$.

4°- Des simulations faites, notamment pour des âges élevés, on peut dégager une sensibilité relativement faible au modèle de répartition choisi et à la méthode utilisée, mais des différences notables entre les solutions « exactes » et approchées.

Chapitre 3 : Lissage des taux annuels de mortalité

I. Introduction

Les estimations âge par âge des taux annuels de décès forment une courbe de mortalité qui se révèle, en général, assez irrégulière. Ces aspérités sont dues aux fluctuations d'échantillonnage et ne sont pas représentatives de la réalité : nous savons en effet que les taux de décès évoluent graduellement avec l'âge[4]. Pour améliorer les estimations brutes et se rapprocher encore plus des véritables taux de mortalité, il est possible d'utiliser les connaissances que nous avons a priori sur la forme des courbes de mortalité. De nombreuses méthodes permettent de répondre à cet objectif. Les estimations corrigées qu'elles produisent progressent régulièrement avec l'âge et, pour cette raison, sont appelées **estimations « lissées » des taux de décès**.

Nous présentons quatre catégories de méthodes permettant d'obtenir des estimations lissées des taux de décès et, ainsi, de construire une table de mortalité :

- les modèles paramétriques,
- les lissages paramétriques,
- les lissages non-paramétriques,
- les modèles relationnels.

Avec un modèle paramétrique, une hypothèse est posée a priori sur la forme de la courbe de mortalité. Pour cette raison, la fonction mathématique qui exprime le taux de mortalité en fonction de l'âge doit être une fonction dont la capacité à retracer la courbe de mortalité a déjà été éprouvée sur d'autres populations. Elle doit permettre de capturer des caractéristiques fondamentales et persistantes des courbes de mortalité, ce qui conduit à privilégier les fonctions contenant peu de paramètres. Cette particularité conduit à un certain manque de souplesse dans la fidélité aux données, en contrepartie elle permet théoriquement d'étendre l'estimation des taux de mortalité à des âges où il n'y a pas encore d'observations. La partie II de ce chapitre présente plusieurs modèles paramétriques de référence.

Les méthodes de lissage (paramétriques ou non-paramétriques) permettent un ajustement assez fidèle aux données d'expérience. Contrairement aux modèles paramétriques, elles ne reposent pas sur l'hypothèse que la courbe de mortalité a une forme connue a priori et, à ce titre, ne sont pas prévues pour obtenir une estimation des taux de mortalité en dehors de la plage de lissage. On peut même noter qu'une extrapolation est par nature impossible pour les méthodes de lissages non-paramétriques étant donné que les taux de mortalité ne sont pas représentés à l'aide d'une fonction mathématique. Les parties III et IV de ce chapitre sont consacrées successivement aux lissages paramétriques puis non-paramétriques.

Les modèles relationnels partent du même principe que la modélisation paramétrique, à la seule différence que le taux de mortalité est désormais exprimé en fonction non plus de l'âge, mais du taux de mortalité donné par une autre table. Ainsi, une table de mortalité connue est prise comme référence et il est supposé que l'on peut, à l'aide d'une fonction comprenant un petit nombre de paramètres, transformer cette table de mortalité de référence pour obtenir celle de la population étudiée. Ces modèles sont abordés dans la partie V.

Les modélisations paramétriques et les méthodes relationnelles permettent d'estimer les taux de mortalité même en dehors des plages d'âge d'expérience, ce qui est une propriété intéressante. Il faut toutefois rappeler que ces approches font prendre un risque de modèle. Pour limiter ce risque, il est recommandé de toujours commencer par une analyse graphique des taux bruts de mortalité avant de sélectionner un modèle paramétrique ou relationnel. Quand il y a trop de fluctuations dans les estimations brutes, il est intéressant d'effectuer un

[4] à quelques exceptions près, notamment la bosse accidents aux alentours de 18-25 ans

lissage non-paramétrique préalablement à l'analyse graphique. A l'inverse, les méthodes de lissage paramétrique ou non-paramétrique permettent une plus grande fidélité dans l'ajustement aux taux bruts mais ne sont pas conçues pour être utilisées en dehors de la plage d'estimation. Soulignons également l'intérêt des méthodes de lissages paramétriques ou non paramétriques pour le lissage en deux dimensions.

Le choix de la méthode sera guidé par les besoins d'application de la table (au-delà, ou pas, des plages d'âge observées, une ou deux dimensions) et par la quantité d'observations disponibles : les méthodes de lissage présentent une souplesse intéressante quand le portefeuille d'expérience est volumineux, les modèles paramétriques ou relationnels sont particulièrement utiles pour les portefeuilles plus réduits. Dans tous les cas, il est intéressant de comparer les résultats donnés par plusieurs méthodes avant de choisir la plus satisfaisante. Signalons enfin que ces méthodes peuvent aussi être combinées entre elles (par exemple lissage non-paramétrique suivi d'une modélisation paramétrique).

Notations :

n_x : nombre de personnes d'âge x exposées au risque

q_x : taux annuel de décès

$\hat{Q}_x$: estimateur brut (ou initial) du taux annuel de décès, variable aléatoire de réalisation $\hat{q}_x$

G_x : estimateur lissé (ou revu) du taux annuel de décès, variable aléatoire de réalisation g_x

U_x : erreur de l'estimation brute, $U_x = \hat{Q}_x - q_x$, variable aléatoire de réalisation $u_x = \hat{q}_x - q_x$.

U'_x : erreur de l'estimation lissée, $U'_x = G_x - q_x$ variable aléatoire de réalisation $u'_x = g_x - q_x$

Remarque préliminaire :

Un estimateur peut être vu comme étant la somme de la vraie valeur et d'un terme résiduel (positif ou négatif) que l'on appelle "erreur d'estimation". Soit : $\hat{Q}_x = q_x + U_x$.

Dans la plupart des cas, sauf dans la théorie Bayesienne, q_x, la vraie valeur sous jacente, n'est pas une variable aléatoire. Par contre, les termes d'erreur en sont et peuvent être définis par : $U_x = \hat{Q}_x - q_x$.

D. London (1985) suggère alors le point du vue suivant sur le lissage :

Soit G un opérateur ou une procédure de lissage que l'on va appliquer aux valeurs initiales $\hat{q}_x$. L' idée est alors que G s'applique à q_x et à u_x séparément, en s'additionnant. Ainsi :

$$G(\hat{q}_x) = G(q_x + u_x) = G(q_x) + G(u_x)$$

Pour que cette dernière propriété soit vérifiée, il faut que G soit linéaire. Si on ajoute à cela le fait de vouloir que G, appliquée à la vraie valeur, soit invariante ($G(q_x) = q_x$), on obtient :

$$G(\hat{q}_x) = q_x + G(u_x) = q_x + u'_x = g_x .$$

Si de plus, l'opérateur G est tel $G(u_x) = u'_x < u_x$ (ce que l'on souhaite), alors g_x est un meilleur estimateur de q_x que $\hat{q}_x$.

II. La modélisation paramétrique

La modélisation paramétrique repose sur l'hypothèse que la courbe de mortalité peut être représentée par une fonction mathématique de quelques paramètres. Le choix du modèle est donc déterminant. Démographes et actuaires ont étudié de nombreux modèles potentiels et identifié ceux qui arrivent le mieux à retracer les caractéristiques fondamentales et permanentes des courbes de mortalité.

Un avantage des modèles paramétriques est qu'ils permettent, de par leur construction, d'étendre l'estimation des taux de mortalité aux âges situés en dehors de la plage d'observation (à condition toutefois que la fonction a été correctement choisie, c'est-à-dire si des études ont montré qu'elle était adaptée à la plage d'âge cible). En contrepartie les modèles paramétriques ne permettent pas toujours un ajustement très fidèle aux données brutes, ils s'avèrent donc souvent plus pratiques pour des portefeuilles de taille réduite ou en complément d'autres méthodes pour extrapoler la mortalité à des âges non observés.

Une attention particulière doit être accordée au nombre de paramètres contenus dans le modèle. En effet, si l'augmentation du nombre de paramètres permet un meilleur ajustement aux taux bruts, elle se fait au détriment de la robustesse du modèle, c'est-à-dire de sa capacité à refléter des caractéristiques générales des courbes de mortalité. Un modèle qui n'est pas robuste donne de bons estimateurs s'il est adapté aux données, en revanche, s'il n'est pas approprié il peut donner de très mauvais estimateurs. Les modèles qui comportent beaucoup de paramètres doivent donc être maniés avec prudence.

Nous présentons dans ce chapitre quelques modèles paramétriques, en commençant par la très ancienne et classique loi de Gompertz. Viennent ensuite les formules de Makeham, de Weibull, de Heligman-Pollard et, finalement la fonction logistique et la formule de Kannisto.

Ces modèles paramétriques ont été largement validés pour des données de population générale. Pour une utilisation sur des données d'assurance, il est important de s'assurer au préalable que la population étudiée est relativement homogène. Citons, comme exemple de source d'hétérogénéité, la sélection effectuée à la souscription du contrat : la sélection peut influencer le risque de décès différemment selon l'âge de l'assuré. Ainsi la courbe de mortalité sera déformée de façon différente selon les âges. Plus généralement, les modèles paramétriques décrits ici s'appliquent à des tables unidimensionnelles. Nous ne présentons pas de modèle pour les tables de mortalité sélectionnée-finale car il n'existe pas de consensus général sur le sujet. Le lecteur intéressé pourra par exemple consulter les articles de Tenenbein et Vanderhoof (1980)**.**

Enfin, il ne faut pas négliger le risque qu'un modèle ne soit pas adapté aux données. Pour éviter cet écueil, il est nécessaire de procéder à toutes les vérifications usuelles, à commencer par une analyse graphique des estimations brutes des taux de décès en fonction de l'âge pour déterminer si la fonction choisie pour la modélisation paramétrique semble acceptable. Il est également nécessaire de procéder aux vérifications usuelles de la qualité d'une régression.

Remarque : Modèles ad hoc

Le constructeur de la table de mortalité peut choisir de ne pas se limiter aux modèles paramétriques connus et classiquement utilisés pour représenter la courbe de mortalité. Il peut construire un modèle paramétrique ad hoc, par exemple à l'aide d'une méthode de sélection des variables explicatives. Il est alors légitime de se demander si ce type d'ajustement peut permettre d'étendre l'estimation à des âges situés en dehors de la plage d'observation. Il est dans ce cas conseillé de s'assurer de la proximité des résultats obtenus à l'aide de ce modèle ad hoc avec ceux donnés par des modèles classiques tels que présentés ci-après.

II.1 Loi de Gompertz (2 paramètres)

Sur de nombreuses populations, il a été observé que le taux instantané de mortalité augmente d'une manière quasi-exponentielle avec l'âge. Gompertz (1825) a proposé un modèle paramétrique simple qui traduit cette tendance :

$$\mu_x = BC^x, \qquad \text{avec } B > 0, C > 1$$

- *B* varie en fonction du niveau de mortalité,
- *C* mesure l'augmentation du risque de décès avec l'âge.

Cette fonction peut permettre de modéliser la courbe de mortalité au-delà de 30 ans environ. Il faut cependant savoir qu'elle tend à sous-estimer la mortalité avant 40 ans et à la surestimer au-delà de 80 ans.

Remarque :

Comme $p_x = e^{-\int_x^{x+1}\mu_s ds}$ et $\int_x^{x+1} BC^s ds = \frac{B}{\ln C} C^x (C-1)$, cela revient à : $p_x = e^{-\left(\frac{B}{\ln C} C^x (C-1)\right)}$

En notant $b = -\frac{B}{\ln C}(C-1)$, on obtient alors : $p_x = e^{bC^x} = 1 - q_x$

II.2 Loi de Makeham (3 paramètres)

Pour améliorer l'évaluation de la mortalité des jeunes adultes (avant 40 ans environ), Makeham (1960) a enrichi la formule de Gompertz d'un paramètre :

$$\mu_x = A + BC^x, \qquad \text{avec } A > 0, B > 0, C > 1$$

On considère usuellement que le paramètre A rend compte de la mortalité environnementale (parfois appelée mortalité extérieure à l'individu), indépendante de l'âge. Cette interprétation peut poser question car il arrive d'obtenir des valeurs négatives pour *A*.

La formule de Makeham ne résout pas le problème de la surestimation du risque aux âges supérieurs à 80 ans déjà rencontré avec la formule de Gompertz. L'opportunité de l'utilisation de ce modèle sera donc liée à l'âge limite d'assurance du produit étudié.

Remarque :

On trouve, de façon similaire à la loi de Gompertz, que :

$$p_x = e^{-\left(A + \frac{B}{\ln C} C^x (C-1)\right)}$$

Notons $s = e^{-A} < 1$ et $g = e^{-\frac{B}{\ln C}} < 1$, on obtient alors : $p_x = s g^{C^x (C-1)} = 1 - q_x$

II.3 La loi de Weibull (2 paramètres)

Weibull (1951) a proposé un modèle pour décrire les défaillances techniques d'un système. Des analogies pouvant être faites entre les défaillances du corps humain et celle d'un système, ce modèle a été transposé à l'étude de la mortalité. On pose alors :

$\mu_x = ax^b$, où $a > 0, b > 1$

ce qui équivaut à $\ln \mu_x = \ln(a) + b \ln x$ (la relation entre $\ln \mu_x$ et $\ln x$ est linéaire).

II.4 La formule de Heligman-Pollard (8 paramètres)

Cette loi a été introduite par L. Heligman et J.H. Pollard en 1980. Elle présente l'avantage de comporter un terme spécifique à la modélisation de la mortalité infantile.

La formule est : $\frac{q_x}{p_x} = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + GH^x$

où : $A^{(x+B)^C}$ modélise la mortalité infantile,

$De^{-E(\ln x - \ln F)^2}$ modélise la mortalité dite accidentelle ou environnementale,

GH^X modélise la mortalité aux âges adultes, hors mortalité environnementale, c'est-à-dire la mortalité due au vieillissement.

On peut montrer que, dans ce modèle, la force de mortalité tend asymptotiquement vers une droite, alors qu'elle a une forme exponentielle dans le modèle de Gompertz et en puissance de l'âge dans le modèle de Weibull. Ainsi, aux âges les plus élevés, le modèle de Gompertz donnera usuellement des estimations supérieures à celles du modèle de Weibull, qui seront elles-mêmes plus élevées que celles du modèle de Heligman-Pollard.

Cette formule, qui contient déjà 8 paramètres, a été raffinée par certains auteurs qui ont rajouté des termes pour obtenir un meilleur ajustement à leurs données. Ces extensions diminuent la robustesse de la formule : les différents paramètres deviennent difficilement interprétables et le modèle ne donne plus une description générale et stable des courbes de mortalité.

II.5 Le modèle logistique et l'approximation de Kannisto (de 2 à 4 paramètres)

Les quatre premiers modèles présentés précédemment (Gompertz, Makeham, Weilbull et Heligman-Pollard) impliquent que la probabilité de décès tende asymptotiquement vers 1 quand l'âge augmente. Une autre possibilité est que la probabilité de décès augmente avec l'âge mais tende vers une limite inférieure à 1 : c'est l'hypothèse qui est contenue dans le modèle logistique (le modèle de Kannisto est une simplification du modèle logistique). D'après ce modèle, il n'existe pas de limite maximale à la durée de vie humaine étant donné que, à aucun âge, la probabilité de survivre jusqu'à l'âge suivant ne devient négligeable. Le modèle logistique présente donc une approche relativement différente des précédents quant à la mortalité aux âges les plus élevés. Ainsi une divergence entre les modèles est toujours observée aux âges très élevés (même avec un ajustement très similaire sur la plage 80-100 ans, on observe une divergence au-delà de 100 ans, voir Thatcher et ali. (1998)).

Le modèle logistique repose sur l'hypothèse que le taux de hasard est une fonction logistique de l'âge. Le modèle le plus général comporte 4 paramètres :

$$\mu_x = a + \frac{b\exp(cx)}{1 + d\exp(cx)}$$

Il inclut le modèle de Makeham, que l'on retrouve quand $d = 0$.

Le modèle logistique a été initialement utilisé par Perks (1932). Depuis, plusieurs justifications théoriques ont été élaborées[5], ce qui lui donne un intérêt particulier même si ces théories ont chacune leurs limites et sont forcément incomplètes :

- La première théorie est celle du « fixed frailty model » - Beard (1971), Vaupel et al. (1979). Elle montre que le modèle logistique découle d'une modélisation simple de la mortalité dans une population hétérogène : la mortalité de chaque individu est

[5] Source : Thatcher et al. (1998)

supposée suivre une loi de Makeham, avec un paramètre *B* qui varie d'une personne à l'autre et qui est distribué selon une loi gamma à la naissance. Sous ces hypothèses il est démontré que le taux de hasard moyen des survivants à l'âge x a une forme logistique (modèle Gamma-Makeham).

- La deuxième explication est celle du processus stochastique de Le Bras (1976). Le Bras part d'une cohorte homogène à la naissance du point de vue de son état de santé. Les individus évoluent ensuite d'un état de santé vers un autre, ce qui crée une hétérogénéité. En utilisant des hypothèses assez simples, Le Bras trouve une formule pour le taux de hasard des survivants à l'âge *x*. Cette formule a été réétudiée par Yashin et al (1994), qui ont montré qu'elle était très proche de celle du modèle Gamma-Makeham (alors que les hypothèses initiales sont très différentes).
- Le théories biologiques du vieillissement constituent le dernier groupe de justifications du modèle. La littérature sur le sujet est très abondante. Le lecteur intéressé peut se référer à Stearns (1992).

Remarque : Ces théories permettent de facto d'expliquer pourquoi les modèles de Makeham et de Gompertz fonctionnent si bien sur certaines tranches d'âge.

Il a été montré que les paramètres *b* et *d* sont à peu près égaux, ce qui permet d'utiliser une forme simplifiée du modèle : $\mu_x = a + \frac{b\exp(cx)}{1+b\exp(cx)}$.

Aux âges les plus élevés (plus de 80 ans), plusieurs auteurs (Kannisto (1992), Himes et ali. (1994)) ont noté que le modèle peut encore être simplifié, en supprimant la constante. Cette version épurée est usuellement appelée le modèle de Kannisto : $\mu_x = \frac{b\exp(cx)}{1+b\exp(cx)}$.

Une étude conduite par Thatcher, Kannisto et Vaupel (1998) conduite sur la mortalité aux âges de 80 ans et plus a montré que ces deux modèles sont ceux qui donnent les meilleurs ajustements parmi un ensemble d'autres modèles paramétriques (dont les modèles de Weibull et de Heligman-Pollard). Ils peuvent donc être d'une utilisation pratique sur les études portant sur des âges élevés ou pour prolonger les courbes de mortalité.

III. Lissages paramétriques

Le lissage paramétrique consiste à trouver une courbe paramétrique qui représente bien l'évolution des taux de décès en fonction de l'âge. Contrairement à la modélisation paramétrique, il ne s'agit pas de postuler a priori une forme bien précise pour la courbe de mortalité mais de déterminer, parmi une famille de fonctions mathématiques, celle qui s'adapte le mieux aux taux bruts. Les fonctions utilisées doivent être suffisamment flexibles pour permettre d'obtenir une courbe de taux lissés qui soit proche de celle des taux bruts.

Nous présentons ici deux familles de fonctions qui peuvent être utilisées pour effectuer un lissage paramétrique. Nous commençons par les splines, qui peuvent aussi servir pour les lissages de données à deux dimensions. Nous nous intéressons ensuite aux lois de la famille Gompertz-Makeham. Nous avons classé ces lois dans la famille des lissages paramétriques car elles n'imposent pas une forme absolue à la courbe de mortalité. Il faut toutefois noter qu'elles se situent à la frontière entre le lissage et la modélisation paramétriques puisqu'elles dérivent de la fameuse courbe de Makeham, qui, elle, relève de la modélisation paramétrique.

Pour effectuer un lissage paramétrique, on a recours aux techniques de régression adaptées aux modèles linéaires généralisés quand la variable expliquée est une fonction linéaire des variables explicatives et aux techniques de régression non linéaire quand cette fonction n'est pas linéaire. Dans ce dernier cas, il est fait appel à des techniques itératives qui nécessitent de choisir des valeurs initiales pour les paramètres de la fonction. Le choix de ces valeurs initiales est un élément crucial car elles doivent être suffisamment proches de leurs vraies valeurs pour que la procédure de régression converge.

III.1 Méthode des splines

III.1.a Présentation générale

Le terme spline tire son origine d'une technique utilisée autrefois pour construire les coques des navires. Le spline était une longue latte de bois fixée en plusieurs endroits. Cette technique permettait d'obtenir, entre les points d'attache, la forme la plus lisse possible. Les mathématiciens ont étudié cette forme à partir de 1946 et en ont dérivé la fonction spline.

Définition

Une fonction spline de degré d (d>0) à l nœuds $k_1,\ldots,k_l$ est une fonction de classe C^{d-1} dont les restrictions aux intervalles $(k_i,k_{i+1}),\ i=0,\ldots,l$ sont des polynômes de degré d. Par convention, on pose $k_0=-\infty$ et $k_{l+1}=+\infty$.

Pour mémoire, une fonction de classe C^{d-1} est une fonction continue, différentiable $d-1$ fois et dont les dérivées d'ordre 1 à d-1 sont continues.

Lissage par un spline

L'ajustement d'une fonction spline aux estimations brutes des taux de décès se fait par la méthode des moindres carrés pondérés (minimisation de la somme pondérée des écarts quadratiques entre les estimations brutes et lissées : $SC=\sum_{x=x\inf}^{x\sup} w_x.(g_x-\hat{q}_x)^2$).

Pour les poids, on peut par exemple utiliser les effectifs sous risque ou l'inverse des variances des estimateurs bruts.

Il n'existe pas de procédure pré-établie pour déterminer le nombre de noeuds optimal et leur valeur. On peut cependant souligner que :

- Plus le nombre de noeuds est élevé, plus les valeurs lissées seront proches des estimations brutes.
- Dans l'objectif de minimiser *SC*, le déplacement d'un noeud est parfois plus intéressant que l'ajout d'un noeud supplémentaire.

Pour déterminer les valeurs des noeuds, plusieurs approches sont envisageables. Une première consiste à les fixer préalablement à la minimisation et à comparer les résultats obtenus avec plusieurs choix différents. Une analyse graphique de la courbe de mortalité peut aider à déterminer les valeurs et le nombre de noeuds. Une approche alternative serait d'inclure les valeurs des noeuds dans les paramètres de la minimisation.

III.1.b Exemple : spline cubique à deux noeuds

Pour lisser la courbe de mortalité par un spline cubique à deux noeuds sur l'intervalle d'âge $[x\inf, x\sup]$, on détermine les valeurs lissées des $\hat{q}_x$ par :

$$\begin{aligned} g_x &= P_1(x) && \text{si } x\inf \le x \le k_1 \\ &= P_2(x) && \text{si } k_1 \le x \le k_2 \\ &= P_3(x) && \text{si } k_2 \le x \le x\sup \end{aligned}$$

où k_1, k_2 sont les 2 nœuds et $P_i(x), i = 1, 2, 3$ des polynômes de degré 3.

Le spline est une fonction de classe C^2 par conséquent, le système suivant doit être vérifié :

$$\begin{aligned} P_1(k_1) &= P_2(k_1) & P_2(k_2) &= P_3(k_2) \\ P_1'(k_1) &= P_2'(k_1) & P_2'(k_2) &= P_3'(k_2) \\ P_1''(k_1) &= P_2''(k_1) & P_2''(k_2) &= P_3''(k_2) \end{aligned}$$

ce qui est revient à prendre :

$$\begin{aligned} P_1(x) &= c_0 + c_1 x + c_2 x^2 + c_3 x^3 \\ P_2(x) &= P_1(x) + c_4 (x - k_1)^3 \\ P_3(x) &= P_2(x) + c_5 (x - k_2)^3 \end{aligned}$$

III.2 Lois de la famille Gompertz-Makeham

III.2.a Présentation générale

On appelle loi de Gompertz-Makeham de paramètres *r* et *s*, $GM(r,s)$, une fonction de la forme : $P_1 + \exp(P_2)$ où P_1 et P_2 sont des polynômes de degrés *r*-1 et *s*-1 (en d'autres termes, r et s représentent les nombres de termes des polynômes P_1 et P_2).

Les lois de Gompertz et de Makeham sont des cas particuliers de cette famille de fonctions.

Un lissage de la courbe de mortalité par une loi de la famille Gompertz-Makeham de paramètres *r* et *s* se fait posant $\mu_x = P_1 + \exp(P_2)$ puis en déterminant les coefficients des polynômes P_1 et P_2 par la méthode des moindres carrés pondérés.

Usuellement, des lissages sont effectués pour plusieurs couples de paramètres (r,s). Une analyse comparative des courbes lissées permet ensuite de déterminer celle qui donne les meilleurs résultats.

III.2.b Exemples

$GM(0,2)$: $\mu_x = e^{b_0 + b_1 x}$ (formule de Gompertz)

$GM(2,2)$: $\mu_x = a_0 + a_1 x + e^{b_0 + b_1 x}$ (dite "2e loi de Makeham")

$GM(1,3)$: $\mu_x = a_0 + e^{b_0 + b_1 x + b_2 (2x^2 - 1)}$

Dans la formule $GM(2,2)$, aux âges élevés, la partie "Gompertz" (exponentielle du polynôme) domine largement les deux premiers termes. Il en va de même avec la formule $GM(1,3)$, où la première constante représente une très faible fraction du taux de hasard aux âges élevés.

IV. Lissages non-paramétriques

Dans ce chapitre, nous présentons deux méthodes de lissage non-paramétriques, celle des moyennes mobiles pondérées et celle de Whittaker-Henderson. La première fait partie des plus anciennes mais elle n'est plus tellement utilisée aujourd'hui en raison de ses faiblesses aux bornes de l'intervalle de lissage. La méthode de Whittaker-Henderson, au contraire, connaît beaucoup de succès, notamment en Amérique du Nord. Nous décrivons la méthode d'origine puis quelques variantes ou extensions intéressantes, avec, en particulier, un paragraphe sur son application au lissage de données à deux dimensions (tables de mortalité sélectionnée-finale).

IV.1 Méthode des moyennes mobiles pondérées (MMP)

IV.1.a Formule R_0-minimum

La méthode des moyennes mobiles centrées symétriques pondérées est l'une des premières méthodes de lissage à avoir été développée ; citons, en particulier, les travaux de E.L. DeForest dans les années 1870. La méthode MMP est assez peu utilisée aujourd'hui car des méthodes plus efficaces lui ont succédé.

La valeur lissée du taux de mortalité est obtenue en prenant la moyenne mobile pondérée de 2k+1 estimations initiales consécutives (taux de décès estimés bruts), indicées de $x-i$ à $x+i$. La formule générale pour un lissage de degré 2k+1 est :

$$g_x = \sum_{i=-k}^{k} c_i \hat{q}_{x+i}$$

Quelles valeurs utiliser pour les c_i ?

Classiquement, les MMP sont choisies symétriques : $c_i = c_{-i}$ pour $i = 1,...,k$.

Pour calculer les coefficients c_i, il est possible de suivre le raisonnement suggéré dans la remarque préliminaire de ce chapitre. On cherche alors à obtenir les coefficients qui reproduisent les q_x et minimisent l'erreur résiduelle de g_x, $u'_x = \sum_{i=-k}^{k} c_i u_{x+i}$:

- En faisant l'hypothèse que q_x a une forme cubique sur la plage d'application de la formule (forme cubique par morceaux), la tranformation par MMP reproduit bien les q_x, c'est-à-dire que $q_x = \sum_{i=-k}^{k} c_i q_{x+i}$, si les deux contraintes suivantes sont respectées : $c_0 + 2\sum_{1}^{k} c_i = 1$ et $\sum_{1}^{k} i^2 c_i = 0$.
- Toujours sous l'hypothèse d'une forme cubique par morceaux, l'espérance de l'erreur résiduelle est nulle (en effet : $E(G_x) = \sum c_i E(\hat{Q}_x) = \sum c_i q_x = q_x$). La minimisation de l'erreur résiduelle revient donc à faire une minimisation de la variance.
- Sous les hypothèses supplémentaires que les erreurs d'estimations (U_x) aux différents âges sont non corrélées et ont la même variance, on peut montrer que la minimisation de la variance revient à minimiser $R_0^2 = \sum_{i=-k}^{k} c_i^2$ (voir, par exemple, D. London, 1985)

Cette minimisation sous contrainte, dite « **formule R_0-Minimum** », donne les valeurs suivantes : $c_i = \frac{3\left(3k^2 + 3k - 1\right) - 15i^2}{(2k+1)(2k-1)(2k+3)}$.

Influence du paramètre k

Le choix de *k* détermine la taille de la plage de lissage et permet d'arbitrer entre lissage et proximité aux données brutes : une augmentation de *k* permettra d'accroître le lissage aux dépens de la proximité entre les taux lissés et les taux bruts.

IV.1.b Limites de la méthodes et extensions

Problème aux extrémités de l'intervalle d'âge

Une limite de la méthode des moyennes mobiles pondérées tient à l'impossibilité de calculer les valeurs lissées aux extrémités de la plage d'âge étudiée. La formule est, en effet, centrée autour de l'âge x. Si $x\inf$ et $x\sup$ sont respectivement le plus petit et le plus grand âges, alors la plus petite et la plus grande valeurs lissées g_x obtenues par une MMP de degré k seront $g_{x\inf+k}$ et $g_{x\sup-k}$ respectivement (et non $g_{x\inf}$ et $g_{x\sup}$). Des méthodes d'extension du lissage jusqu'aux bornes ont été proposées (Greville, 1981).

Remise en cause de l'hypothèse d'égalité des variances des erreurs d'estimation

Une autre limite de la méthode tient aux deux hypothèses utilisées pour déterminer les coefficients :

- la non corrélation des erreurs d'estimation (U_x) aux différents âges
- l'égalité des variances des erreurs d'estimation (U_x) à tous les âges.

La validité de ces hypothèses ne va pas de soi. Le deuxième hypothèse est particulièrement difficile à accepter dans de nombreux cas pratiques. Supposons, en effet, que $\hat{Q}_x$ suive une loi binomiale, alors $Var(U_x) = Var\left(\hat{Q}_x\right) = \frac{q_x(1-q_x)}{n_x}$. L'exposition n_x peut varier considérablement d'un âge à l'autre sur la plage de lissage. Plutôt que $Var(U_x) = Var(U_{x+r}) = \sigma^2$ (0<*r*<*k*), il serait plus correct de supposer que $Var(U_x).n_x = Var(U_{x+r}).n_{x+r} = \sigma^2$. Pour plus de détails, le lecteur intéressé peut se référer à l'article de R.L. London (1981) qui explore cette modification de la formule.

Variante : formule R_z-Minimum

R_0 représente le ratio des variances des estimateurs lissés et bruts :

$$R_0^2 = \frac{Var(G_x)}{Var\left(\hat{Q}_x\right)} \text{ car } Var\left(\hat{Q}_x\right) = Var(U_x) = \sigma ?.$$

Cette observation est à l'origine d'une variante méthodologique qui consiste à minimiser $R_z^2 = \frac{Var\left(\Delta^z G_x\right)}{Var\left(\Delta^z \hat{Q}_x\right)}$ (formule « R_z-Minimum »). Le lecteur intéressé trouvera plus d'informations en annexe ainsi que dans l'ouvrage de D. London (1985).

Soulignons que le raisonnement utilisé pour justifier la formule R_0-Minimum ne tient plus : la formule R_z-Minimum (avec $z > 0$) ne peut s'expliquer que par un simple objectif de lissage de la courbe de mortalité.

Comparaison de la formule standard et de ses variantes

D'après R.L. London (1981) les meilleurs résultats de lissage par les MMP sont obtenus par la formule R_0-Minimum avec modification de l'hypothèse sur l'égalité des variances des erreurs, puis par la formule R_0-Minimum standard. Les résultats donnés par des formules « R_z-Minimum » sont moins satisfaisants.

IV.2 Méthode de Whittaker-Henderson

IV.2.a Méthode d'origine

Cette méthode de lissage doit son nom à E. T. Whittaker (1923), qui l'a introduite, et à R. Henderson (1924), qui a montré comment passer de la théorie à la pratique (bien que les méthodes actuellement utilisées soient très éloignées de ses propositions).

La méthode de Whittaker-Henderson consiste à rechercher le meilleur compromis entre l'adéquation aux données brutes et la régularité de la courbe de mortalité. Les taux de mortalité lissés sont obtenus en minimisant la mesure $M = F + hS$, où :

- $F = \sum_{x=x\,\inf}^{x\,\sup} w_x \left(g_x - \hat{q}_x\right)^2$: F est la somme pondérée des carrés des écarts entre les valeurs lissées et les valeurs brutes, elle mesure la **fidélité** des taux de mortalité lissés aux taux bruts . Plus les taux lissés se rapprochent des taux bruts, plus la valeur de F diminue.
- $S = \sum_{x=x\,\inf}^{x\,\sup - z} \left(\Delta^z g_x\right)^2$: *S* est la somme des carrés des différences d'ordre *z* des taux lissés, elle permet d'évaluer la **régularité** de la courbe lissée. Plus l'aspect de la courbe est régulier, plus la valeur de *S* diminue.

  z est un entier positif. En pratique, les valeurs les plus utilisées sont *z* = 2, 3 ou 4. Pour mémoire, rappelons que : $\Delta^z g_x = \Delta^{z-1} g_{x+1} - \Delta^{z-1} g_x$ pour $z > 0$; $\Delta^0 g_x = g_x$
- *h* est un réel positif qui permet de contrôler l'influence que l'on souhaite donner à chacun des deux critères précédents. Plus h est grand, plus la minimisation porte sur le terme S et impose à la courbe une allure régulière. Plus h est petit, plus la minimisation accorde de l'importance au terme F et impose à la courbe lissée de se rapprocher des données brutes. Soulignons que si $h = 0$, aucun lissage n'est effectué.
- w_x sont les poids donnés à chaque âge ($w_x \geq 0$).

On appelle méthode de type A celle où l'on attribue le même poids à tous les âges ($\forall x$, $w_x = 1$) et méthode de type B celle où l'on donne des poids différenciés en fonction de l'âge.

Une hypothèse implicite de la méthode de Whittaker-Henderson est que la forme de la courbe de mortalité est proche de celle d'un polynôme de degré *z*-1 au plus : en effet, sur une courbe polynômiale d'ordre z-1, les différences d'ordre z sont nulles. Les observations de Miller (1946) ont conduit à donner une nouvelle interprétation au paramètre *h* :

- plus *h* est élevé, plus les valeurs lissées approchent celles données par un polynôme en *x* de degré *z*-1 ajusté aux données brutes par la méthode des moindres carrés pondérés,
- plus *h* est petit, plus la courbe lissée ressemble aux données brutes.

Choix des pondérations

Plusieurs formules peuvent être choisies pour les poids, w_x. Il est fréquent d'utiliser une pondération par l'effectif à l'âge *x* rapporté à l'effectif moyen pour tous les âges :

$$w_x = \frac{n_x}{\overline{n}}, \text{ où } \overline{n} = \frac{\sum_{x=x\inf}^{x\sup} n_x}{x\sup - x\inf + 1}.$$

Ce choix permet de limiter le poids donné aux points aberrants (voir glossaire). Il permet également d'avoir deux propriétés intéressantes : on obtient le même nombre total de décès et le même âge moyen au décès avec les taux bruts et les taux lissés.

De façon générale, **il est important de regarder si la pondération introduit un changement d'ordre de grandeur de la valeur de F par rapport à celle de S car cela aura une influence sur le choix du paramètre *h***.

Application pratique

La minimisation de *M* fait appel à l'algèbre linéaire. Le lecteur intéressé trouvera en annexe des indications sur le passage à une expression matricielle du problème et sur la forme de la solution analytique. L'obtention de la solution numérique nécessite l'inversion d'une matrice qui est presque singulière et doit donc être conduite avec soin.

IV.2.b Autres calculs de la distance : modification de Schuette, norme de Chebyshev

Un manque de robustesse a été reproché à la formule de Whittaker-Henderson. En effet, si la distribution du risque à l'âge *x* génère des points aberrants plus fréquemment que la loi normale, alors les méthodes utilisant les moindres carrés ne donnent pas forcément les meilleurs résultats. Elles sont trop sensibles aux points aberrants. En 1978, Donald R. Schuette propose une alternative à la formule de Whittaker-Henderson : la distance utilisée n'est plus la distance quadratique mais la valeur absolue. Ce changement conduit à minimiser la mesure suivante :

$$M = \sum_{x\inf}^{x\sup} w_x \left| g_x - \hat{q}_x \right| + h \sum_{x\inf}^{x\sup - z} \left| \Delta^z g_x \right|$$

La formule de D.R. Schuette est moins sensible aux points aberrants que celle de Whittaker-Henderson ; elle fournit une estimation plus robuste. Par contre elle présente une particularité peu élégante : elle produit des valeurs lissées qui évoluent par palier avec *h*.

L'article de D.R. Schuette a ouvert la voie à l'utilisation d'autres distances dans la formule de Whittaker-Henderson. Citons notamment F.Y. Chan, L.K. Chan et M.H. Yu (1984), qui étudient le passage aux normes ℓ_p et à la norme de Chebyshev, ℓ_∞ (définitions données dans le glossaire).

Le choix de la norme est un arbitrage entre avantages et inconvénients de chacune :

- Pour diminuer l'influence des points aberrants, il est préférable d'utiliser, dans le critère de fidélité, une norme ℓ_p où *p* est petit. Le choix de D.R. Schuette est alors le plus adapté.
- Au contraire, dans le critère de régularité, il vaut mieux choisir un *p* grand afin d'avoir une courbe qui varie le plus uniformément possible. La norme de Chebyshev ($p = \infty$) est la plus adaptée de ce point de vue : *S* est alors égal à la plus grande valeur absolue des différences d'ordre *z* des valeurs lissées.

IV.2.c Modification du critère de régularité : autres fonctions de référence

On appelle fonction de référence une fonction qui donne une valeur nulle pour la mesure de régularité S. Le critère de régularité impose aux valeurs lissées de se rapprocher des valeurs données par une fonction de cette forme.

Nous avons vu que, dans la formule originale de Whittaker-Henderson, la fonction de référence est un polynôme de degré *z*-1 au plus, c'est-à-dire de la forme : $a_0 + a_1 x + a_2 x^2 + ... + a_{z-1} x^{z-1}$.

Walter B. Lowrie (1982) modifie ce critère et impose aux valeurs lissées d'être proches d'une fonction de référence égale à la somme d'une exponentielle et d'un polynôme de degré *z*-2 au plus. La fonction de référence prend la forme :

$$a_0 + a_1 x + a_2 x^2 + ... + a_{z-2} x^{z-2} + (1+r)^x \text{, où } r > -1 \text{ (} r \text{ est un réel) ;}$$

et le critère de régularité devient : $S = \sum_{x=x\inf}^{x\sup - z} \left(\Delta^z g_x - r\Delta^{z-1} g_x\right)^2$.

Pour le choix de *r*, il a été suggéré de procéder par itération et de retenir la valeur qui minimise la quantité *S*. On peut remarquer que $1 + r = \frac{\Delta^{z-1} g_{x+1}}{\Delta^{z-1} g_x}$, ce qui peut donner l'idée de prendre comme une valeur de départ pour l'itération la moyenne des $\frac{\Delta^{z-1} g_{x+1}}{\Delta^{z-1} g_x}$.

IV.2.d Introduction d'un critère de proximité à une table de référence

Walter B. Lowrie (1982) a enrichi la formule de Whittaker-Henderson d'un deuxième critère de fidélité, qui mesure la proximité à une table de référence. Ce critère permet de prendre en compte une connaissance a priori sur la forme de la courbe de mortalité et de contraindre les taux de mortalité lissés à ne pas trop s'en éloigner.

En pratique, cela conduit à ajouter un troisième terme dans la mesure de Whittaker-Henderson. La mesure à minimiser est désormais :

$$M = (1-l)F + lF' + hS \text{, où :}$$

- F, S et h sont définis comme précédemment ;
- $F' = \sum_{x=x\inf}^{x\sup} w_x^s (g_x - s_x)^2$: F' mesure l'écart entre les taux lissés et les taux donnés par la table de référence ;
- Les s_x sont les taux de mortalité par âge donnés par la table de référence ;
- Les w_x^s sont les poids attribués aux taux de mortalité de référence (w_x^s>0) ;
- Le paramètre *l* ($0 \le l \le 1$) fixe l'importance relative donnée au critère de proximité aux taux de référence par rapport au critère de proximité aux taux bruts.

W. B. Lowrie fait remarquer que la formule peut être utilisée également pour l'interpolation de données. Il faut pour cela donner des poids positifs aux points pivots (points où les taux bruts ont été estimés) et des poids nuls aux points pour lesquels la méthode doit fournir des valeurs. Une condition suffisante pour que la minimisation génère un résultat est que le nombre de points pivots soit supérieur ou égal à *z.*

Cette modification fait perdre une propriété intéressante de la méthode de Whittaker-Henderson, pour laquelle $\sum_{x=x\inf}^{x\sup} w_x g_x = \sum_{x=x\inf}^{x\sup} w_x \hat{q}_x$ et $\sum_{x=x\inf}^{x\sup} x.w_x.g_x = \sum_{x=x\inf}^{x\sup} x.w_x.\hat{q}_x$. Si les poids choisis sont les effectifs (ou les effectifs par âge rapportés à l'effectif total), cette

propriété permet de préserver, lors du lissage, le nombre de décès total et l'âge moyen au décès.
Pour avoir un lissage qui conserve le nombre de décès total, il est possible de prendre une transformation linéaire des taux de mortalité de référence ($s'_x = as_x + b$).

IV.2.e Extension à deux dimensions et ajout de contraintes

Des développements de la méthode de Whittaker-Henderson ont permis de l'appliquer au lissage de taux de mortalité dépendant de deux variables (classiquement l'âge et la durée écoulée). Il s'agit :

- du passage à deux dimensions,
- et de la prise en compte de contraintes linéaires, qui permettent notamment d'imposer une augmentation des probabilités de décès avec l'âge et la durée écoulée.

Nous considérons ici le lissage de taux de mortalité dépendant de l'âge et de la durée écoulée. Usuellement les estimateurs bruts sont notés $\hat{q}_{\lfloor x^{sous} \rfloor_d}$ et les valeurs lissées $g_{\lfloor x^{sous} \rfloor_d}$ où x^{sous} est l'âge en début d'assurance et *d* la durée écoulée. Pour simplifier les notations, nous choisissons, dans cette partie, de noter $\hat{q}_{x,d} = \hat{q}_{\lfloor x^{sous} \rfloor_d}$ et $g_{x,d} = g_{\lfloor x^{sous} \rfloor_d}$ avec $x = x^{sous} + d$.

Le lissage s'effectue par une minimisation sous contrainte.

- La mesure à minimiser conserve une forme similaire au cas à une dimension :
  - $M = F + S(h)$ où *h* est un vecteur de paramètres
  - ou $M = (1-l)F + lF' + S(h)$ si un critère de proximité à une table de référence est introduit
- Les contraintes linéaires sont exprimées sous la forme : $Eg^{vect} \leq b$ où $g^{vect} = (g_{x\inf,d\inf}, g_{x\inf,d\inf+1}, \ldots, g_{x\inf,d\sup}, g_{x\inf+1,d\inf}, g_{x\inf+1,d\inf+1}, \ldots, g_{x\sup,d\sup})^T$, *E* est une matrice de dimensions (*e*,(*xsup-xinf*+1).(*dsup-dinf*+1)) contenant les coefficients des *e* contraintes et *b* est un vecteur de dimension *e* contenant les bornes des contraintes.

Le passage à deux dimensions donne :

$$F = \sum_{x=x\inf}^{x\sup} \sum_{d=d\inf}^{d\sup} w_{x,d} \left(g_{x,d} - \hat{q}_{x,d}\right)^2 , \; F' = \sum_{x=x\inf}^{x\sup} \sum_{d=d\inf}^{d\sup} w^s_{x,d} \left(g_{x,d} - s_{x,d}\right)^2 .$$

Pour *S*, plusieurs formules sont possibles selon la fonction de référence retenue pour le critère de régularité.

Nous détaillons la valeur de *S* pour la fonction de référence suivante :

$$f(x,d) = P_1(x) + P_2(d) + a_1(1+r_1)^x + a_2(1+r_2)^d ,$$

où $P_1(x)$ est un polynôme de degré z_1,

et $P_2(d)$ est un polynôme de degré z_2.

W.B. Lowrie considère que cette fonction constitue un « excellent critère de régularité pour un problème en deux dimensions », mais il est, bien évidemment, possible de se limiter à une forme plus simple (par exemple du type $f(x,d) = P_1(x) + P_2(d)$).

On a alors :

$$S(h) = h_1 \sum_{x=x\inf}^{x\sup - z_1} \sum_{d=d\inf}^{d\sup} \left(\Delta_1^{z_1} g_{x,d} - r_1 . \Delta_1^{z_1 - 1} g_{x,d}\right)^2 + h_2 \sum_{x=x\inf}^{x\sup} \sum_{d=d\inf}^{d\sup - z_2} \left(\Delta_2^{z_2} g_{x,d} - r_2 . \Delta_2^{z_2 - 1} g_{x,d}\right)^2$$

avec les notations suivantes :

- $h = (h_1, h_2)$

- $\underset{1}{\Delta}^{n} g_{x,d} = \underset{1}{\Delta}^{n-1} g_{x+1,d} - \underset{1}{\Delta}^{n-1} g_{x,d}$, avec n>0 et $\underset{1}{\Delta}^{0} g_{x,d} = g_{x,d}$
- $\underset{2}{\Delta}^{n} g_{x,d} = \underset{2}{\Delta}^{n-1} g_{x,d+1} - \underset{2}{\Delta}^{n-1} g_{x,d}$, avec n>0 et $\underset{2}{\Delta}^{0} g_{x,d} = g_{x,d}$

Elias S. W. Shiu (discussion de F.E. Knorr, 1984) a fait remarqué que ce critère peut être incomplet dans la mesure où il n'introduit pas de termes croisés ($x.d$ ou $(1+r_1)^x.(1+r_2)^d$). Il est possible de l'enrichir, en ajoutant des différences « mixtes » ($\underset{1,2}{\Delta}^{z_1,z_2} g_{x,d} = \underset{1}{\Delta}^{z_1}\left(\underset{2}{\Delta}^{z_2} g_{x,d}\right)$), le lecteur intéressé pourra, par exemple, se référer à l'article de W.B. Lowrie (1993).

La minimisation de *M* fait appel à l'algèbre linéaire ; plus de détails sont données en annexe.

V. Les modèles relationnels

Les modèles relationnels sont des modèles dans lesquels le taux de mortalité n'est pas fonction uniquement de l'âge, mais du taux de mortalité donné par une table de référence et/ou de l'âge.

Les modèles relationnels ont été initialement développés par des démographes (citons les modèles de Cox, de Brass et de Hannerz). Ces modèles partent de l'hypothèse qu'il existe un lien mathématique simple entre la mortalité de la population étudiée et celle d'une population de référence. Leur application à des données d'assurance s'avère fort intéressante, dans un format un peu modifié par rapport aux modèles proposés par les démographes.

Michel Denuit et ali. (ouvrage à paraître[6]) ont proposé une méthodologie consistant à sélectionner les variables explicatives les plus pertinentes par comparaison de plusieurs lissages non-paramétriques. Pour plus de régularité et/ou pour étendre les estimations à d'autres âges, il reste ensuite à déterminer un modèle paramétrique adapté qui permettra de formaliser le lien entre la mortalité d'expérience et celle donnée par la table de référence.

Pour obtenir des résultats satisfaisants, il est important de choisir une population de référence dont les caractéristiques sont proches de celle retenue pour la table de mortalité d'expérience (par exemple, il vaut mieux prendre une table masculine si l'on étudie des assurés hommes).

Ces méthodes sont particulièrement intéressantes dans le cas où le volume de données n'est pas très important.

[6] Pour cette raison, nous ne consacrons pas plus que ces quelques lignes aux modèles relationnels.

Chapitre 4 : Validation de la table construite

I. Introduction

Une table de mortalité d'expérience doit être construite à partir de toutes les informations disponibles, ce qui inclut les connaissances que le constructeur a, a priori, sur la forme de la courbe de mortalité et celles issues de l'observation de la mortalité d'expérience sur le portefeuille étudié. L'objectif de l'étape de validation est de contrôler la cohérence entre la table de mortalité obtenue et ces informations initiales :

- Les connaissances a priori sur la mortalité ont-elles bien été prises en considération ?
- Les estimations sont-elles en accord avec la mortalité observée dans le passé sur le portefeuille ?

Si plusieurs tables ont été construites selon des méthodes différentes, les tests présentés dans cette partie peuvent aussi être utilisés pour choisir celle qui présente les meilleures propriétés.

Nous proposons trois types de contrôles. La deuxième partie de ce chapitre est consacrée aux contrôles de la cohérence des taux. La troisième présente des méthodes permettant d'évaluer la régularité de la courbe de mortalité et la proximité entre les taux estimés et la mortalité observée sur le portefeuille.

Rappel des notations utilisées :

Les notations sont les mêmes que celles utilisées dans les chapitres précédents.

n_x : nombre de personnes d'âge x exposées au risque

q_x : taux annuel de décès

$\hat{Q}_x$: estimateur brut du taux annuel de décès, variable aléatoire de réalisation $\hat{q}_x$

G_x : estimateur lissé du taux annuel de décès, variable aléatoire de réalisation g_x

Pour les tables à deux dimensions :

Si *x* est l'âge à la souscription du contrat et *t* la durée (nombre entier) écoulée depuis la souscription (en faisant commencer les durées écoulées à 0), on note :

$q_{[x]+t}$ le taux annuel de décès,

$\hat{q}_{[x]+t}$ son estimation brute,

$g_{[x]+t}$ son estimation lissée.

II. Vérification de la cohérence des taux

Lors de la validation d'une table de mortalité, il faut s'assurer que les taux obtenus sont cohérents entre eux. Il s'agit de contrôler que l'on observe bien une croissance des taux avec l'âge et la durée écoulée. Dans le cas où des tables de mortalité distinctes ont été construites pour différents groupes d'assurés, il est essentiel de vérifier que les résultats obtenus présentent des écarts en accord avec notre connaissance a priori du risque.

II.1 Table à une dimension

A l'exception de la mortalité aux premiers âges de la vie et de la bosse accident aux alentours de 16-25 ans, les taux de mortalité devraient être croissants avec l'âge, ce qui revient à vérifier $g_x \leq g_{x+1}$ pour tout âge *x*.

Il peut également être intéressant de vérifier si les différences premières des taux de décès augmentent avec l'âge, ce qui est généralement le cas. Les exceptions les plus classiques sont :

- les premiers âges de la vie,
- la plage d'âge correspondant à la bosse accident,
- les âges très élevés.

Cela revient à contrôler si l'inégalité $g_{x+2} - g_{x+1} \geq g_{x+1} - g_x$ est vérifiée.

II.2 Table de mortalité sélectionnée-finale

Dans une table de mortalité sélectionnée-finale, les taux de mortalité, qui dépendent de l'âge de l'assuré et de la durée écoulée, doivent respecter une certaine logique dans leur évolution selon chacun de ces deux axes et en diagonale.

Avec les mêmes exceptions que pour une table à une dimension, il faut que :

- Pour un âge de souscription donné, le taux de décès augmente avec la durée écoulée du contrat du contrat :

$$\forall x, \quad g_{[x]} \leq g_{[x]+1} \leq g_{[x]+2} \leq \ldots$$

- Pour une durée écoulée donnée, le taux de décès augmente avec l'âge de l'assuré :

$$\forall x, \forall t \quad g_{[x]+t} \leq g_{[x+1]+t} \leq g_{[x+2]+t} \leq \ldots$$

- Pour un âge atteint donné, le taux de décès augmente avec la durée écoulée du contrat :

$$\forall x \quad g_{[x]} \leq g_{[x-1]+1} \leq g_{[x-2]+2} \leq \ldots$$

Remarque :
On peut noter que certains de ces contrôles sont redondants. Par exemple, la troisième condition précédente donne $g_{[x]} \leq g_{[x-1]+1}$ et la deuxième $g_{[x-1]+1} \leq g_{[(x-1)+1]+1} = g_{[x]+1}$. Si ces deux conditions sont vérifiées, alors la première condition l'est aussi de façon automatique ($g_{[x]} \leq g_{[x]+1}$).

Comme pour les tables à une dimension, il est intéressant de vérifier la croissance des différences premières. Il a été remarqué (American Academy of Actuaries' CSO Task Force, 2001) que l'effet, sur le taux de décès, de l'augmentation de l'âge est en général supérieur à celui de l'augmentation de la durée écoulée, on s'attend alors à observer :

$$g_{[x]+t+2} - g_{[x]+t+1} \geq g_{[x]+t+1} - g_{[x]+t}$$

II.3 Tables par segment de population

Si la population a été segmentée et des tables construites pour chacun des segments, il faut vérifier que les tables de mortalité obtenues respectent bien nos connaissances a priori sur la mortalité. A défaut, il est important de rechercher des explications et de comprendre pourquoi la population étudiée ne présente pas une mortalité aux propriétés classiques.

Il serait difficile de lister tous les cas de segmentation possible. Citons en deux assez courants : les tables hommes / femmes et les tables fumeurs / non-fumeurs.

Le taux de mortalité des fumeurs excède largement celui des non-fumeurs : il est important de vérifier que ce critère est bien respecté à tous les âges sur les tables construites.

A tout âge, dans une population homogène, les femmes ont un risque de décès inférieur à celui des hommes. Cette propriété devrait être retrouvée dans les données d'assurance si les groupes constitués par les hommes et les femmes ont des distributions comparables vis-à-vis des autres critères influant sur le risque de mortalité.

III. Critères de régularité et de fidélité

La dernière étape du travail de validation concerne le contrôle de la régularité de la courbe de mortalité et la vérification de sa fidélité aux données observées. Nous l'avons vu tout au long des chapitres précédents, la construction d'une table de mortalité est nécessairement un compromis entre ces deux critères. Pour cette raison, et sauf erreur de construction manifeste, les résultats de ces tests ne permettront pas de rejeter la table construite, ils donneront, en revanche, des éléments d'appréciation de sa qualité.

L'importance à accorder à chacun de ces critères dépendra notamment du volume de données utilisées pour la construction de la table et de leurs caractéristiques (homogénéité, écarts par rapport aux caractéristiques attendues du portefeuille pour lequel la table certifiée sera utilisée).

III.1 Evaluation de la régularité de la courbe de mortalité

Il est généralement admis que les vrais taux de mortalité ont une croissance avec l'âge assez régulière (à l'exception de quelques tranches d'âge bien spécifiques). Il est souhaitable que les estimations aient aussi cette propriété, c'est pourquoi il est utile de vérifier la régularité des taux de mortalité obtenus.

La régularité, telle qu'elle a été abordée dans la méthode de Whittaker-Henderson, peut également constituer un critère d'évaluation.

Pour mémoire, on rappelle qu'il y a régularité des taux lissés si :

$$\sum_{x=x\inf}^{x\sup-z} \left(\Delta^z g_x\right)^2 \to 0 \quad , \quad \text{soit} \quad \text{pour } z=1 : \sum_{x=x\inf}^{x\sup-1} \left(g_{x+1}-g_x\right)^2 \to 0$$

$$\text{pour } z=2 : \sum_{x=x\inf}^{x\sup-2} \left(g_{x+2}-2g_{x+1}+g_x\right)^2 \to 0$$

Il faut noter que ce critère ne prend pas en compte la possibilité que la courbe de mortalité ait une forme exponentielle, du type $f(x)+(1+r)^x$. Il faudrait pour cela un critère adapté, similaire à la modification proposée W. B. Lowrie pour la formule de Whittaker-Henderson, à savoir : $\sum_{x=x\inf}^{x\sup-z} \left(\Delta^z g_x - r\Delta^{z-1} g_x\right)^2 \to 0$. Ceci pose le problème du choix de *r* (voir la partie consacrée au lissage par la méthode Whittaker-Henderson).

III.2 Vérification de la fidélité à la mortalité observée

Nous ne citons ici que les critères qui nous semblent les plus pertinents. L'actuaire certificateur peut cependant demander que d'autres tests soient effectués s'il considère qu'ils présentent un intérêt pour compléter la validation de la table (citons, par exemple, le test du signe).
Nous nous limitons à des critères généraux et ne rappelons pas ici les vérifications usuelles à effectuer pour s'assurer de la qualité des ajustements paramétriques.

III.2.a Distance en valeur absolue entre les taux lissés et les taux bruts

Lorsqu'il y a fidélité des taux bruts aux taux lissés, on a : $\sum_{x\,\inf}^{x\,\sup} |\hat{q}_x - g_x| \to 0$.

Plus la somme des distances entre taux bruts et taux lissés se rapproche de 0, plus la fidélité est grande.

III.2.b Ratio du nombre de décès observés sur ceux attendus

Une vérification simple et assez pragmatique de la fidélité de la table construite à la mortalité observée consiste à appliquer cette table de mortalité au portefeuille étudié : ceci permet de calculer le nombre de décès prédits par la table sur la période d'étude. Ce nombre de décès attendus, noté A, est ensuite comparé au nombre de décès effectivement observés, noté O. Un ratio O/A proche de 1 montre une bonne capacité de la table à prédire la mortalité d'expérience. Nous conseillons d'effectuer systématiquement ce type de contrôle lors d'une certification.

Il peut aussi être intéressant de calculer le même ratio en pondérant les sinistres attendus et observés par le montant assuré. Ce ratio O/A en montants montre la capacité de la table à refléter les flux de sinistres payés, il faut toutefois garder en mémoire que ce ratio est souvent sensiblement plus variable que le ratio O/A en nombres. Sa variabilité est d'autant plus grande que les capitaux assurés sont hétérogènes et que le portefeuille est de taille réduite.

L'analyse du ratio O/A peut être faite d'abord au niveau global du portefeuille, notamment sans distinction par âge ni par sexe. Elle donne une idée de la qualité globale de la table construite. Le ratio doit normalement être très proche de 1.

Puis, il est intéressant d'analyser ce ratio, de façon plus fine, sur différents sous-groupes : par tranche d'âge, par sexe, par tranche de durée écoulée d'assurance et selon tous les autres facteurs explicatifs disponibles. La proximité à 1 des ratios obtenus sur ces sous-groupes dépendra fortement de la taille du portefeuille étudié. L'intérêt d'une telle analyse est de permettre d'identifier les sous-groupes sur lesquels la table tend à surestimer (O/A<1) ou à sous-estimer le risque (O/A>1).

Définitions

Durée écoulée :
La « durée écoulée » est une abréviation pour désigner la durée écoulée depuis la souscription du contrat. Elle est usuellement mesurée en nombres entiers : en général, il s'agit de la durée réelle arrondie à l'entier supérieur le plus proche, ou, plus rarement, à l'entier inférieur le plus proche. Ainsi, la durée écoulée est égale à 1 entre la souscription et le premier anniversaire de la souscription, à 2 entre le premier et le deuxième anniversaires de la souscription et ainsi de suite. On peut noter que la durée écoulée n'est autre que l'âge du contrat d'assurance arrondi à l'entier immédiatement supérieur.

Factorisation de Cholesky :
Une matrice symétrique A est dite définie positive si, pour tout vecteur x, le produit $x^T Ax$ est positif. Pour une telle matrice, on montre qu'il est possible de trouver une matrice triangulaire inférieure L telle que : $A = LL^T$
Cette décomposition permet notamment de calculer la matrice inverse de A.

Normes ℓ_p et ℓ_∞ :

Rappelons que, pour le vecteur $X = (x_1, \ ..., \ x_n)^T$:

- la norme ℓ_p est définie par $\|X\|_p = \left(|x_1|^p + ... + |x_n|^p\right)^{1/p}$
- la norme ℓ_∞, ou norme de Chebyshev, est définie par : $\|X\|_\infty = \max_{1 \le i \le n} |x_i|$;
  elle vérifie : $\|X\|_\infty = \lim_{p \to \infty} \|X\|_p$.

Méthodes robustes - Points aberrants :
L'ajustement par la méthode des moindres carrés pour obtenir le maximum de vraisemblance suppose une distribution normale des erreurs. Sous cette hypothèse, certaines réalisations statistiques ont une très faible probabilité d'être observées, elles sont appelées points aberrants. Même dans les meilleures conditions d'observation, il arrive d'avoir des points aberrants : leur présence modifie beaucoup le résultat de l'ajustement par les moindres carrés.
L'hypothèse de distribution normale des erreurs est difficile à vérifier, il s'agit souvent d'une approximation de la réalité. Or l'approximation de la loi Binomiale ou de la loi de Poisson par une loi normale n'est pas forcément aussi satisfaisante sur les queues de distribution que sur le coeur de la distribution. Ces lois prédisent plus points aberrants qu'une loi normale : dans un ajustement par les moindres carrés, ces points aberrants peuvent peser beaucoup trop sur le résultat. Il faut donc éviter les points aberrants ou avoir recours à des procédures d'estimation qui sont y peu sensibles ; ces dernières sont appelées méthodes robustes.

Table de mortalité sélectionnée-finale
La sélection médicale effectuée à la souscription de certains contrats d'assurance influe sur le niveau du risque de décès. Une table de mortalité sélectionnée-finale est une table à deux dimensions (l'âge de l'assuré et la durée écoulée du contrat) qui reflète l'influence de la sélection médicale. Tant que cette influence est présente, on observe des niveaux de « **mortalité sélectionnée** ». La « **mortalité finale** » représente le niveau de mortalité atteint quand l'effet de la sélection médicale a disparu.

Usuellement, on note $q_{[x]+t}$ le taux annuel de décès à l'âge *x+t* quand la souscription a été effectuée à l'âge *x* : la durée écoulée du contrat est alors *t+1* ou *t* (selon la convention choisie d'arrondir la durée écoulée à l'entier supérieur ou inférieur).

Notons d^f la durée écoulée à partir de laquelle le niveau de mortalité finale est atteint. Au-delà de cette durée écoulée la mortalité évolue seulement en fonction de l'âge, avec les notations utilisées, cela équivaut à : $\forall t \geq d^f, \quad q_{[x]+t} = q_{[x+t-d^f]+d^f}$