

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t, \infty[}(T_x)$$

ressources-actuarielles.net



MODÈLES DE DURÉE

Support de cours 2023-2024

Tables de mortalité

Frédéric PLANCHET

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t, \infty[}(T_x)$$

SOMMAIRE

1. Introduction	3
1.1. Les tables réglementaires.....	3
1.2. Les tables d'expérience	4
2. L'analyse de la mortalité	5
2.1. Notations	5
2.2. Le diagramme de Lexis	7
2.3. Mortalité longitudinale et mortalité transversale.....	10
2.4. Taux de décès par génération et par année	11
2.5. Répartition des décès dans l'année	12
2.6. Les indicateurs synthétiques du niveau de la mortalité.....	13
2.6.1. Espérance de vie résiduelle	13
2.6.2. Entropie	14
3. Quelques indicateurs.....	15
3.1. Données générales.....	16
3.2. Impact du tabagisme	16
3.2.1. Étude de tables canadiennes.....	16
3.2.2. Étude de Ministère de la Santé.....	18
4. La construction de tables de mortalité d'expérience.....	19
4.1. Cadre général	19
4.2. Tables prospectives.....	22
4.3. Le modèle de Lee-Carter.....	23
4.3.1. Estimation des paramètres.....	24
4.3.2. Extrapolation de la composante temporelle.....	27
4.3.3. Fermeture de la table.....	27
4.4. Le modèle log-Poisson	28
4.4.1. Estimation des paramètres.....	31
4.4.2. Obtention d'intervalles de confiance.....	32
4.5. Les modèles log-linéaires.....	33
4.6. Le modèle logistique décalé	35
4.7. Utilisation des séries chronologiques	37
4.8. Les modèles à référence externe.....	38
4.8.1. Régression logistique.....	38
4.8.2. Positionnement par rapport à une référence externe	39
5. Les critères de validation du modèle	39
5.1. La fidélité aux données	40
5.2. La comparaison des valeurs modélisées et des observations.....	41
5.3. La stabilité des estimations	41
5.4. La capacité prospective	42
6. Le décès comme premier instant d'atteinte d'un seuil par un processus.....	42
6.1. Présentation du modèle	42
6.2. Estimation des paramètres.....	43
7. Références.....	43
8. Annexe : comparaison des modèles de Lee-Carter et de Poisson	45

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[} (T_x)$$

1. Introduction

La construction d'une table de mortalité dans le cadre paramétrique standard a déjà été évoquée précédemment dans ce cours¹ ; l'objectif du présent support est de détailler les outils d'analyse des tables de mortalité, d'une part, et d'aborder la question de l'évolution de la mortalité au cours du temps et des modèles propres à en rendre compte.

Les tables de mortalité utilisées par les assureurs pour leurs tarifs et leurs provisions sont encadrées par la réglementation. En pratique, des tables de la population générale sont utilisables par défaut, et la réglementation prévoit les conditions dans lesquelles l'organisme peut utiliser ses propres tables. Ce contexte est défini par les articles A132-18 du Code des Assurances (ex A335-1) Les différents types de tables de mortalité

Du point de vue de l'assureur, on peut distinguer les tables réglementaires, qui jouent un rôle particulier dans la détermination du tarif et des provisions, et les tables d'expérience ; d'un point de vue technique, on distingue les tables transversales, ou « tables du moment » et les tables prospectives, intégrant l'aspect dynamique de la mortalité.

1.1. Les tables réglementaires

Les tables réglementaires comportent deux volets :

- Les tables TH et TF 00-02 pour les assurances en cas de décès ;
- Les tables ci-dessus utilisées avec des décalages d'âges pour les assurances en cas de vie (à l'exclusion des rentes).

Homologuées par l'arrêté du 20 décembre 2005, les tables TH et TF 00-02 ont été établies à partir des données de l'INSEE issues d'observations réalisées entre 2000 et 2002 et sont applicables aux contrats d'assurance vie souscrits depuis le 1^{er} juillet 1993. La table TF décrit la mortalité féminine. La table TH est construite à partir de la population masculine.

De plus, la nécessité d'utiliser des tables de mortalité prospectives pour les rentes viagères a été prise en compte par le législateur et des tables par génération (TGH et TGF 05) ont été homologuées par un arrêté du 01/08/2006. Celles-ci ont été obtenues² sur base de la mortalité de la population des bénéficiaires de contrats de rentes observée sur la période 1993-2005 et de données sur la population générale (INSEE) de 1962 à 2000. Ces tables servent depuis le 1^{er} janvier 2007 à la tarification et au provisionnement des contrats de rentes viagères immédiates ou différées. Elles imposent un tarif minimal.

On peut compléter ce panorama en citant les tables « IA 2013 », construites par l'Institut des Actuaire en 2014 pour fournir aux entités qui le souhaitent des références et outils de calcul pour ajuster des tables de mortalité prospectives sans biais sur leur portefeuille³.

¹ Voir le support « [Statistique des modèles paramétriques](#) ».

² Des informations plus complètes sur ces tables se trouvent [ici](#).

³ Une présentation de ces tables et des outils et méthodologies associés est proposée dans [cette page](#).

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t;\infty[}(T_x)$$

1.2. Les tables d'expérience

Dans le cadre du suivi technique de ses produits et au regard de l'article A. 132-18 du Code des assurances⁴, un assureur peut souhaiter utiliser des tables de mortalité d'expérience en lieu et place des tables officiellement en vigueur pour justifier du niveau de la prime pure dans les contrats qu'il couvre et du niveau de provisionnement des engagements dans les comptes sociaux⁵. Il apparaît en effet opportun, dans ce cadre, de cerner au mieux tout « comportement » de la population assurée qui serait significativement différent des tables réglementaires.

La procédure d'agrément des actuaires indépendants habilités à certifier et à suivre les tables de mortalité (et les lois de maintien en incapacité de travail et en invalidité) est définie par l'Institut des Actuaires, après avis de la Commission de Contrôle des assurances et de la Commission de Contrôle des mutuelles et des institutions de prévoyance :

- dans le cadre des arrêtés du 19 mars 1993 (entreprises d'assurances), du 13 octobre 1993 (mutuelles), du 21 décembre 1993 (institutions de prévoyance) concernant les lois de mortalité,
- dans le cadre de l'arrêté du 28 mars 1996 (entreprises d'assurances, mutuelles et institutions de prévoyance), concernant les lois de maintien en incapacité de travail et en invalidité.

Cette procédure comprend la mise en place d'une Commission d'Agrément indépendante et souveraine dans ses missions d'habilitation des Actuaires à certifier et à suivre les tables de mortalité et les lois de maintien en incapacité de travail et en invalidité. Elle a été approuvée par les membres de la Commission d'Agrément le 3 décembre 2002. Elle a été ratifiée par le Conseil d'administration de l'Institut des Actuaires le 11 décembre 2002 et transmise aux autorités de tutelle le 18 décembre 2002.

En pratique la mise en place, et l'autorisation d'utilisation, d'une table d'expérience comporte 3 étapes :

- La construction de la table ;
- La certification initiale ;
- Le suivi annuel destiné à assurer la pérennité du droit d'utilisation de la table.

Le rapport final de certification doit s'assurer que la table permet la « constitution de provisions suffisantes et prudentes ». Ce document doit en particulier :

- « valider les données utilisées et leurs sources, qu'elles soient internes ou externes à l'entreprise,

⁴ L'article équivalent pour l'arrêt de travail est l'article 143-12 du règlement ANC 2015-11 du 26/11/2015

⁵ Les comptes prudentiels relèvent d'une autre logique, avec le calcul « sans biais » des engagements. Dans ce cadre, l'utilisation de lois de mortalité d'expérience n'est pas soumise à certification.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[}(T_x)$$

- vérifier les hypothèses de travail et les modalités utilisées pour construire les tables de mortalité ou les lois de maintien en incapacité de travail ou en invalidité
- s'assurer que les principes de prudence communément admis ont été respectés, eu égard aux risques induits (en particulier stabilité des tables ou des lois de maintien),
- définir précisément les conditions d'application et de validité des éléments certifiés, les statistiques ou tableaux de bord à préparer périodiquement par l'entreprise pour permettre le suivi des résultats d'expérience. »

Le suivi doit être annuel. En l'absence de suivi, la validité des tables (et des lois de maintien) cesse deux ans après leur certification. La validité des tables de mortalité est limitée à cinq ans (celle des lois de maintien en incapacité et en invalidité à quatre ans).

Le point important que l'on peut retenir est que la certification ne concerne pas une table dans l'absolu, mais une table utilisée pour un contrat ou un groupe de contrats particuliers, au regard notamment du risque induit par le contrat considéré.

2. L'analyse de la mortalité

On s'intéresse à la variable aléatoire T représentant la durée de vie d'un individu ; on suppose les individus de la population dans un premier temps identiques, de sorte qu'on pourra disposer d'échantillons issus de la loi de T .

2.1. Notations

Il est commode de considérer les variables T_x représentant la durée de vie résiduelle d'un individu conditionnellement au fait qu'il soit vivant à l'âge x , i.e. $T_x =_d [T - x | T > x]$. On peut alors définir la probabilité de survie entre x et $x+t$ par :

$${}_t p_x = P(T_x > t) = P(T > x+t | T > x),$$

et le quotient de mortalité entre x et $x+t$:

$${}_t q_x = 1 - {}_t p_x = P(T_x \leq t) = P(T \leq x+t | T > x).$$

Lorsque $t=1$ il est omis dans les notations, et on écrit plus simplement $q_x = {}_1 q_x$ et $p_x = {}_1 p_x$.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[}(T_x)$$

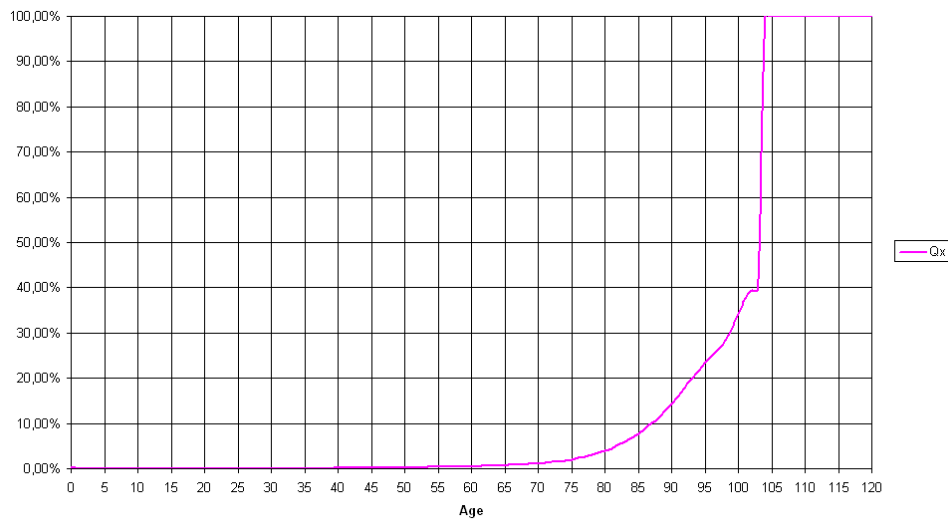


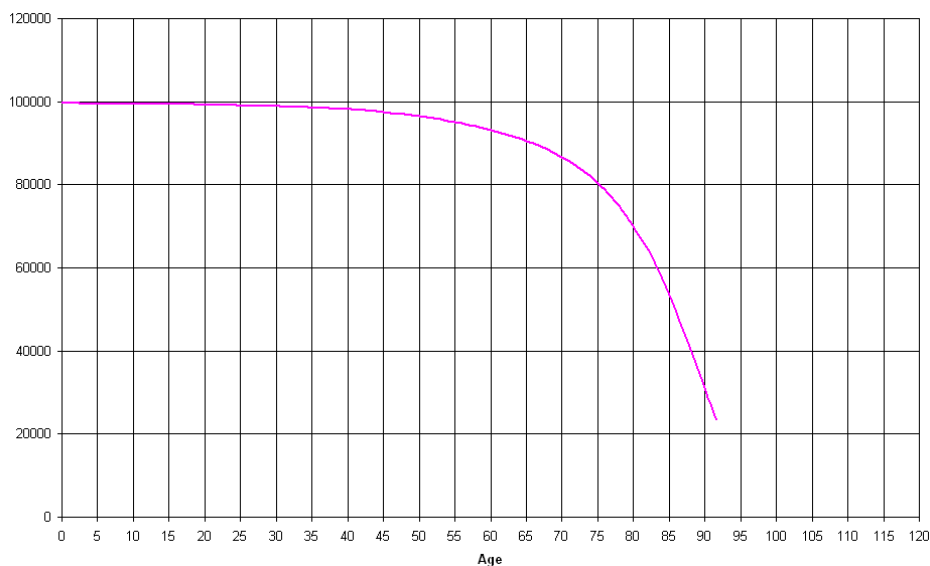
Fig. 1 : Taux instantané de mortalité en fonction de l'âge

Ces quotients s'expriment simplement à l'aide de la fonction de survie de T :

$${}_t p_x = \frac{S(x+t)}{S(x)}.$$

Il est usuel de noter⁶ $l_x = S(x)$; le nombre de décès entre x et $x+t$ est noté ${}_t d_x = l_x - l_{x+t}$; dans le cadre de l'analyse statistique de la mortalité d'une cohorte on mesure le temps vécu par les individus de la cohorte entre x et $x+t$, défini par :

$${}_t L_x = \int_0^t l_{x+u} du.$$



⁶ À une constante multiplicative de normalisation près.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[}(T_x)$$

ressources-actuarielles.net

Fig. 2 : Représentation de la fonction de survie

À partir de cet indicateur on peut définir la durée de vie résiduelle, qui est un indicateur caractéristique de la table de mortalité :

$$E_x = \int_0^{+\infty} l_{x+u} du = \sum_{i=x}^{\infty} L_x.$$

Le quotient de mortalité ${}_t q_x$ est calculé en rapportant un nombre de décès sur la période à l'effectif en début de période ; on calcule également le taux de mortalité, obtenu en rapportant le nombre de décès à l'effectif moyen sur la période, soit :

$${}_t m_x = \frac{{}_t d_x}{{}_t L_x}.$$

Les quotients de mortalité sont des probabilités (nombres sans dimension) alors que les taux de décès sont exprimés en inverse de l'unité de temps et décomptent des décès par personne sous risque et par unité de temps. Cette différence conduit aux relations suivantes avec la fonction de hasard, appelée dans ce contexte « taux instantané de mortalité » :

$$\mu_{x+t} = \lim_{h \rightarrow 0} h^{-1} P(t < T_x \leq t+h | T_x > t) = \frac{1}{{}_t p_x} \frac{\partial} {\partial t} {}_t q_x,$$

car $h^{-1} P(t < T_x \leq t+h | T_x > t) = \frac{{}_{t+h} q_x - {}_t q_x}{{}_t p_x}$; donc lorsque h est petit, ${}_h q_x \approx h \mu_x$ et

${}_h p_x \approx 1 - h \mu_x$. Le lien entre le taux instantané de mortalité et le taux de mortalité est direct : $\mu_x = \lim_{h \rightarrow 0} {}_h m_x$, ce qui justifie *ex-post* la terminologie.

La relation entre fonction de survie conditionnelle et fonction de hasard s'écrit avec les notations utilisées ici :

$${}_t p_x = \exp\left(-\int_0^t \mu_{x+s} ds\right).$$

2.2. Le diagramme de Lexis

Lors des études de mortalité, il est rare que l'on dispose d'une information exacte sur les âges au décès et les dates de décès ; ces données sont le plus souvent disponibles sous forme arrondie, en âge entier et année entière. Afin de déterminer correctement les taux bruts de mortalité dans ce contexte, on utilise un formalisme particulier, le diagramme de Lexis⁷.

L'analyse de la mortalité d'un groupe donné fait intervenir trois mesures de temps : l'âge des individus, leur génération (date de naissance) et la date d'observation ; bien entendu ces 3 informations sont liées et la connaissance de 2 d'entre elles détermine la troisième.

⁷ Du nom du statisticien et démographe allemand Wilhelm LEXIS (1837-1914).

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t, \infty[}(T_x)$$

Chacune de ces dimensions a toutefois son importance dans la détermination du niveau de la mortalité :

- L'âge : cette variable influence évidemment le risque de décès ;
- La date d'observation : le risque de décès peut varier en fonction de circonstances comme une épidémie, un événement exceptionnel (la canicule de l'été 2003 par exemple), etc.
- La génération : des phénomènes tels que l'amélioration des conditions sanitaires, les progrès de la médecine conduisent à modifier le risque de mortalité à un âge donné au cours du temps ; de plus, on peut imaginer que le passé d'une génération donnée puisse modifier le niveau de sa mortalité future : typiquement, une épidémie intervenant à une date t et touchant les gens d'âge x à cette date peut contribuer à diminuer les taux de décès aux âges supérieurs à x pour cette génération, en entraînant la mort prématurée des individus les moins résistants.

Il est alors commode de représenter la vie d'un individu dans un système d'axes rectangulaire appelé « diagramme de Lexis », de la manière suivante :

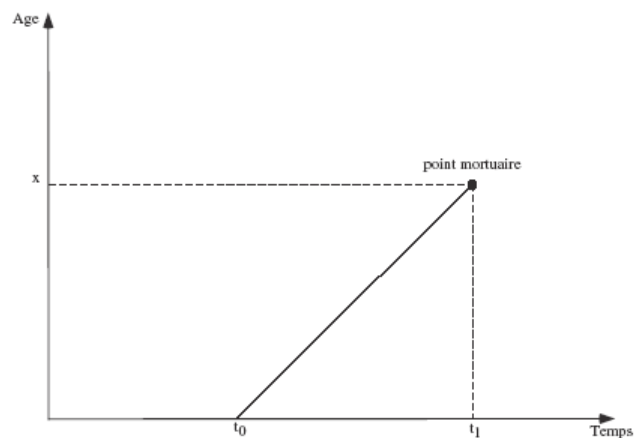


Fig. 3 : Diagramme de Lexis

La vie d'un individu est donc représentée par une ligne parallèle à la première bissectrice, qui coupe l'axe des abscisses l'année de la naissance et s'arrête au « point mortuaire » au jour du décès. En traçant une bande horizontale entre x et $x + 1$ on isole les individus décédés à l'âge x , et en traçant une bande verticale entre g et $g + 1$, on isole les décès des individus de la génération g . Dans ce formalisme, x et g sont entiers, et x mesure l'âge en années révolues.

Les points mortuaires qui se situent dans le carré ci-dessous sont associés aux décès à l'âge x au cours de l'année t :

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[} (T_x)$$

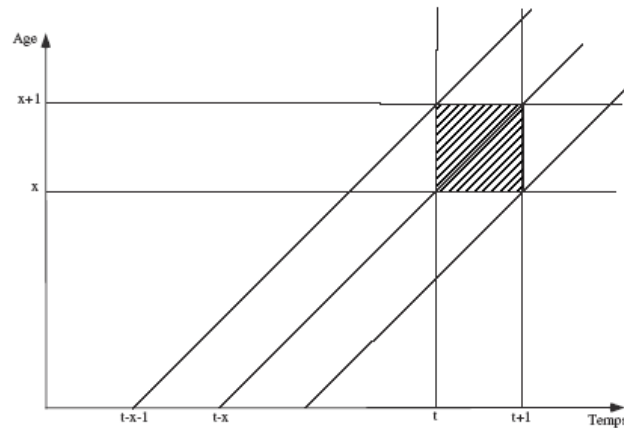


Fig. 4 : Identification des décès à l'âge x l'année t

Les individus concernés appartiennent aux générations $t-x$ et $t-x-1$. De même on peut décomposer le nombre de décès à l'âge x parmi la génération g :

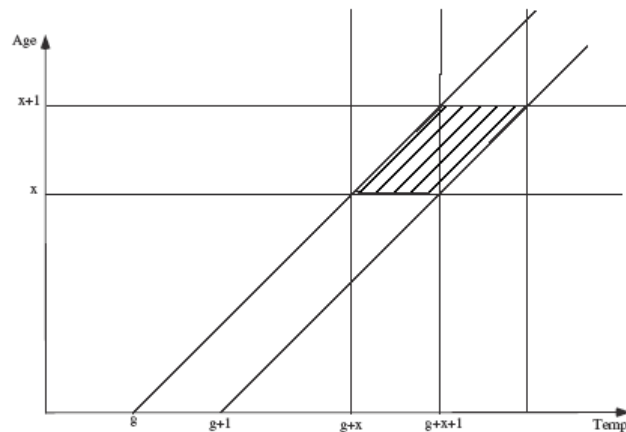
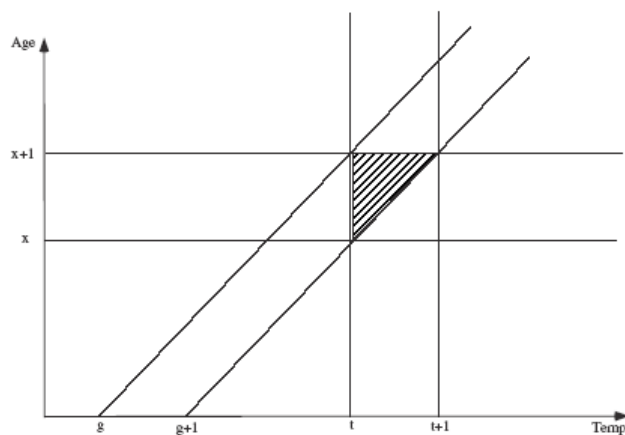


Fig. 5 : Identification des décès à l'âge x dans la génération g

Ces décès se sont produits au cours des années $g+x$ et $g+x+1$. On obtient également le nombre de décès à l'âge x parmi la génération g au cours de l'année t :



$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t, \infty[}(T_x)$$

Fig. 6 : *Identification des décès à l'âge x l'année t dans la génération g*

Enfin, on peut représenter de la même manière le nombre de décès au cours de l'année t parmi les individus de la génération g :

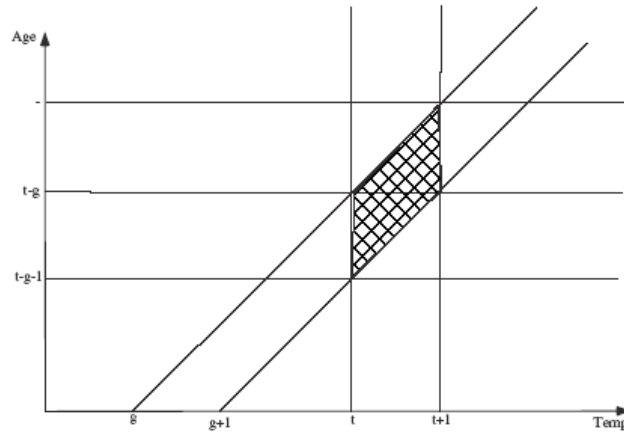


Fig. 7 : *Identification des décès dans la génération g l'année t*

2.3. Mortalité longitudinale et mortalité transversale

La mesure « naturelle » de la mortalité consiste à comptabiliser les décès survenus au cours d'une période donnée (une année par exemple), puis à calculer les taux de décès par âge en rapportant ce nombre de décès à l'effectif sous risque. Cela revient à considérer une bande verticale du diagramme de Lexis.

On voit que si la mortalité évolue au fil du temps, cette approche biaise la mesure de la mortalité, plus précisément, dans une période de baisse tendancielle de la mortalité, elle conduit à sous-estimer les durées de vie (ou à surestimer les taux de décès). En effet, dans cette approche on considère des individus de générations différentes pour calculer les taux de décès, la table obtenue ne représente donc la mortalité d'aucune génération réelle.

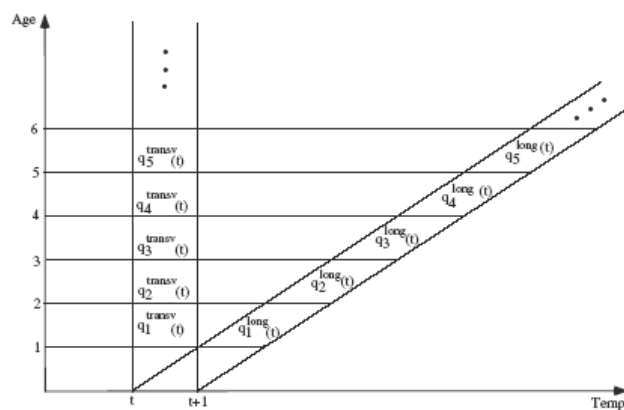


Fig. 8 : *Mortalité longitudinale et mortalité transversale*

La mortalité réelle d'une génération s'obtient en considérant les taux le long d'une bande comme ci-dessus.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t; \infty[}(T_x)$$

On aura besoin par la suite de calculer le quotient de mortalité à l'âge x pour l'année t ; comme on l'a vu en 2.2, ce quotient fait intervenir deux générations, $t-x$ et $t-x-1$. On détermine donc les « quotients partiels de mortalité » suivants, en notant $D_{xt}(g)$ le nombre de décès à l'âge x pour la génération g intervenus l'année t :

$$\hat{q}_{xt}^1 = \frac{D_{xt}(t-x)}{l_{x,t+1} + D_{xt}(t-x)}.$$

($l_{x,t+1}$ désigne le nombre de personnes d'âge x au 01/01/t+1); ce quotient approche donc la probabilité pour les individus de la génération $t-x$ de décéder à l'âge x l'année t . On estime de même la probabilité pour les individus de la génération $t-x-1$ de décéder à l'âge x l'année t :

$$\hat{q}_{xt}^2 = \frac{D_{xt}(t-x-1)}{l_{xt}}.$$

Le quotient cherché résulte alors de l'agrégation de ces 2 quotients partiels: pour survivre entre son $x^{\text{ème}}$ et son $(x+1)^{\text{ème}}$ anniversaire, il faut survivre de son $x^{\text{ème}}$ anniversaire à la fin de l'année civile, puis de la fin de l'année civile à son $(x+1)^{\text{ème}}$ anniversaire, soit:

$$1 - \hat{q}_{xt} = (1 - \hat{q}_{xt}^1)(1 - \hat{q}_{xt}^2).$$

Lorsque l'on veut déterminer le taux de mortalité à l'âge x pour l'année t , on calcule classiquement, avec des notations évidentes:

$$\hat{m}_x = \frac{D_{xt}}{(l_{x,t} + l_{x,t+1})/2}.$$

2.4. Taux de décès par génération et par année

On note x l'âge entier considéré, t l'année courante et $g=t-x$ la génération correspondante. On a avec des notations évidentes pour l'exposition au risque sur $[x, x+1[\times [t, t+1[$:

$$E(x, t) = E(x, t, g) + E(x, t, g-1).$$

$E(x, t, g)$ correspond au temps passé dans $[x, x+1[\times [t, t+1[$ par les individus nés en g après leur anniversaire de l'année t et $E(x, t, g-1)$ correspond au temps passé dans $[x, x+1[\times [t, t+1[$ par les individus nés en $g-1$ avant leur anniversaire de l'année t .

De même, le nombre de décès s'écrit:

$$D(x, t) = D(x, t, g) + D(x, t, g-1).$$

De manière symétrique, les individus de la génération g peuvent décéder à l'âge x l'année t ou l'année $t+1$ (leur anniversaire $x+1$ est dans l'année $t+1$):

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t, \infty[}(T_x)$$

$$E(x, g) = E(x, t, g) + E(x, t+1, g) \text{ et } D(x, g) = D(x, t, g) + D(x, t+1, g).$$

Les estimateurs de Hoem des probabilités conditionnelles de décès $\hat{q} = \frac{D}{E}$ en vision annuelle ou générationnelle ne sont donc pas liés par une relation simple et en particulier

$$\hat{q}_{an}(x, t) \neq \hat{q}_{gen}(x, t-x).$$

Dans l'hypothèse où chaque génération dispose de sa propre table de mortalité, il est donc préférable de calculer $\hat{q}_{gen}(x, g) = \frac{D(x, g)}{E(x, g)}$ puis de travailler ensuite pour l'extrapolation avec $\hat{q}(x, t) = \hat{q}_{gen}(x, t-x)$ plutôt que de calculer directement $\hat{q}_{an}(x, t)$.

2.5. Répartition des décès dans l'année

Les données disponibles sont souvent des données regroupées dans lesquelles l'unité de temps est l'année. Il convient alors de se donner une règle de répartition des décès dans l'année. Ce point a été abordé précédemment ; trois hypothèses sont classiquement proposées :

- la constance des taux instantanés de décès entre 2 âges non entiers (hypothèse exponentielle) : ${}_t q_x = 1 - (1 - q_x)^t$;
- la répartition linéaire des décès au cours de l'année : ${}_t q_x = 1 - t \times q_x$;
- l'hypothèse de Balducci, qui postule que ${}_t q_x = \frac{t \times q_x}{1 - (1-t)q_x}$.

L'hypothèse de Balducci peut être écartée d'emblée car elle conduit à des taux instantanés de mortalité décroissants entre 2 âges entiers ; en effet, on trouve dans ce modèle que :

$$\mu_{x+t} = -\frac{\partial}{\partial t} \ln({}_t p_x) = \frac{q_x}{p_x + tq_x},$$

ce qui résulte de ${}_t p_x = 1 - {}_t q_x = 1 - \frac{tq_x}{p_x + tq_x} = \frac{p_x}{p_x + tq_x}$. Le choix entre les 2 hypothèses

restantes n'est pas neutre sur l'appréciation que l'on aura du niveau de la mortalité. En effet, si T_x^c et T_x^l sont les durées de vie résiduelles respectivement dans le modèle de constance des taux instantanés et dans le modèle de répartition linéaire des décès, on a, avec des notations évidentes : $S_x^l(t) \geq S_x^c(t)$, ce qui implique en particulier que $e_x^l(t) \geq e_x^c(t)$; l'hypothèse de constance des taux instantanés conduit donc à des durées de vie inférieures : de ce fait, il s'agit d'une hypothèse prudente dans le cas de garanties en cas de décès, moins prudente pour des contrats de rentes. Toutefois, l'écart entre les 2 approches est faible.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t, \infty[}(T_x)$$

ressources-actuarielles.net

Modèles de durée

Pour prouver l'inégalité $S_x^l(t) \geq S_x^c(t)$, on fixe $t = k + r$, avec $k = [t]$ et $0 \leq r < 1$ et on note que :

$$S_x^l(t) = P(T_x^l > k + r) = {}_k p_x (1 - r q_{x+k})$$

et

$$S_x^c(t) = P(T_x^c > k + r) = {}_k p_x p_{x+k}^r.$$

L'inégalité à démontrer est donc équivalente à $1 - r(1 - p_{x+k}) \geq p_{x+k}^r$, et cette dernière inégalité est la conséquence directe⁸ de $(1+x)^r \leq 1+rx$ pour tout $0 \leq r < 1$. Dans les modèles présentés ci-après, l'hypothèse de constance du taux instantané de mortalité entre 2 âges entiers est effectuée, de sorte que l'on a ${}_t q_x = 1 - (1 - q_x)^t$.

2.6. Les indicateurs synthétiques du niveau de la mortalité

Les caractéristiques d'une table de mortalité sont usuellement résumées au travers d'un certain nombre d'indicateurs : l'espérance de vie et l'entropie sont deux indicateurs importants, présentés ci-après.

2.6.1. Espérance de vie résiduelle

L'espérance de vie résiduelle est par définition $e_x = E(T_x) = \frac{E_x}{l_x}$; on a donc $e_x = \frac{1}{l_x} \int_0^{+\infty} l_{x+u} du$;

on déduit en particulier de cette expression que :

$$\frac{de_x}{dx} = -1 + \mu_x e_x.$$

En effet $\frac{d}{dx} e_x = \frac{-l_x^2 - \left(\frac{d}{dx} l_x\right) \int_x^{+\infty} l_u du}{l_x^2}$, et comme $\mu_x = -\frac{d}{dx} \ln l_x$, on a bien l'égalité ci-

dessus. La version discrète de cette formule est simplement $e_x = \frac{1}{L_x} \sum_{h>0} L_{x+h}$.

Cette expression signifie que lorsque le taux de mortalité est petit, l'espérance de vie résiduelle diminue d'environ un an chaque année ; en revanche, lorsque le taux de mortalité est grand, on peut avoir une espérance de vie résiduelle qui augmente.

D'un point de vue pratique, cela signifie que le graphe des e_x est à peu près aligné sur une droite de pente -1 jusque vers 75 ans, pour s'incurver ensuite, comme on le constate sur le graphique ci-dessous :

⁸ On a même l'inégalité stricte si $r > 0$.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t; \infty} [(T_x)]$$

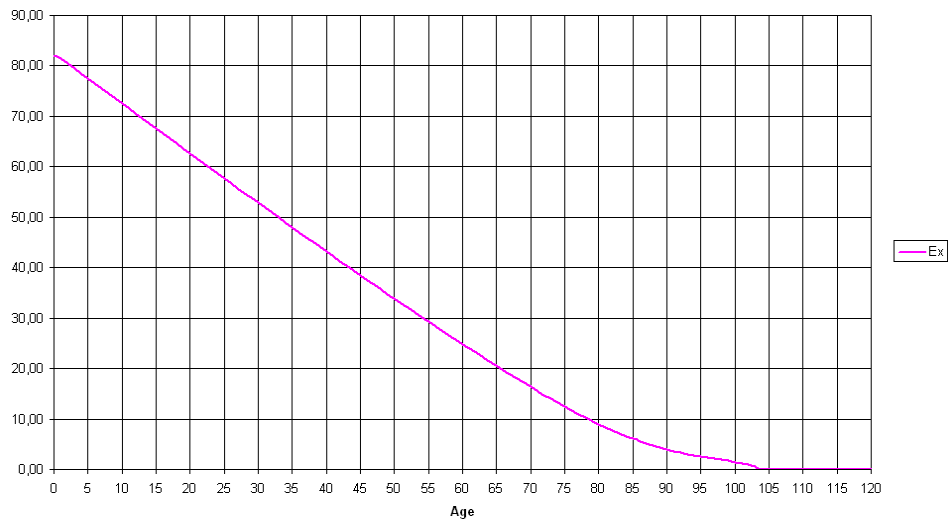


Fig. 9 : Espérance de vie résiduelle en fonction de l'âge

À partir de 75 ans, un ajustement polynomial d'ordre 2 fonctionne en général correctement (ce qui fournit une paramétrisation simple d'une table de mortalité du moment). On peut noter que l'espérance de vie résiduelle peut s'interpréter comme le prix d'une rente viagère continue actualisée à taux 0.

2.6.2. Entropie

La baisse des taux de mortalités aux âges jeunes, sans pour autant que l'âge ultime de vie semble évoluer sensiblement, a pour conséquence un phénomène d'« orthogonalisation » des tables de mortalité, de plus en plus de personnes décédant à un âge élevé⁹ :

⁹ Ce phénomène s'accompagne d'une baisse de la variance de la durée de vie au cours du temps.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[}(T_x)$$

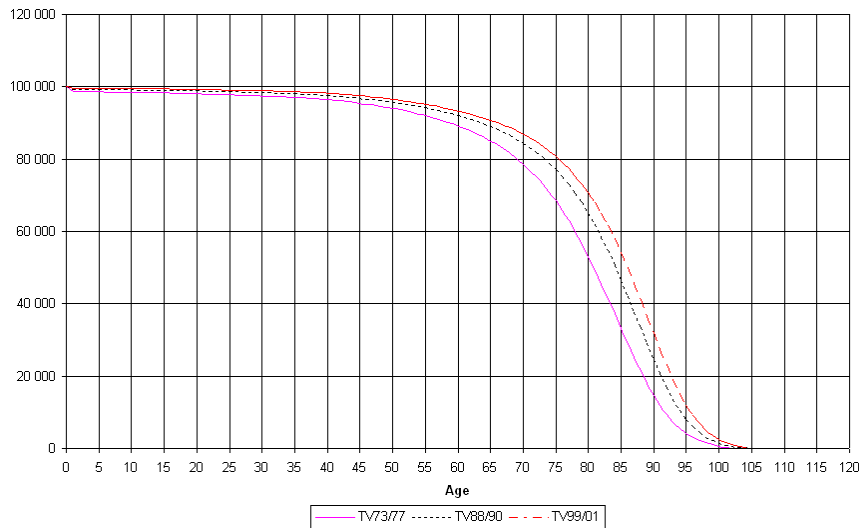


Fig. 10 : Illustration du phénomène d'orthogonalisation des tables de mortalité

L'entropie se propose de mesurer ce phénomène ; on la définit par :

$$H = - \frac{\int_0^{+\infty} l_x \ln(l_x) dx}{\int_0^{+\infty} l_x dx}.$$

Comme on a $\mu_x = -\frac{d}{dx} \ln(l_x)$, on peut réécrire cette quantité sous la forme :

$$H = \frac{\int_0^{+\infty} l_x \mu_x e^{-\int_0^x \mu_x dx} dx}{l_0 e_0}.$$

La version discrète de cette formule est $H = -\frac{\sum_{h>0} L_{x+h} \ln(L_{x+h})}{\sum_{h>0} L_{x+h}}$. L'entropie rapporte donc

le nombre moyen d'« années perdues » du fait du décès au nombre d'années possibles « en stock » à la date 0.

On peut remarquer que $H = 0$ si et seulement si tous les décès se produisent au même âge et que $H = 1$ correspond à la situation extrême opposée dans laquelle le taux instantané de mortalité est constant : cette grandeur mesure est donc bien adaptée à la mesure du phénomène d'orthogonalisation. L'entropie est passée d'environ 50 % à la fin du 19^{ième} siècle à 15 % aujourd'hui.

3. Quelques indicateurs

L'objectif de cette section est de fournir quelques ordres de grandeur utiles sur le niveau de la mortalité. On illustre également la manière de quantifier l'impact sur la mortalité de

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t;\infty} [(T_x)]$$

caractéristiques particulières de la population, en prenant l'exemple du critère fumeur / non-fumeur.

3.1. Données générales¹⁰

Des espérances de vie à la naissance et à 60 ans, ainsi que le taux de décès à cet âge, sont indiqués dans le tableau ci-dessous :

	Femmes			Hommes		
	TV73/77	TV88/90	TV99/01	TD73/77	TD88/90	TD99/01
Naissance	76,5	80,2	82,2	68,6	72,0	74,7
60 ans	20,9	23,5	25	16,1	18,3	19,9
q60	0,77%	0,57%	0,48%	1,90%	1,57%	1,18%

Ce tableau fait clairement apparaître des disparités entre les hommes et les femmes :

	Femmes / hommes		
	TV73/77	TV88/90	TV99/01
Naissance	112 %	111 %	110 %
60 ans	130 %	128 %	126 %
q60	41 %	37 %	41 %

On lit également la baisse tendancielle de la mortalité :

	Femmes		Hommes	
	TV88/90 / TV73/77	TV99/01 / TV88/90	TD88/90 / TD73/77	TD99/01 / TD88/90
Naissance	105 %	102 %	105 %	104 %
60 ans	112 %	106 %	114 %	109 %
q60	74 %	85 %	82 %	75 %

L'écart de mortalité entre les hommes et les femmes se traduit par le fait que, dans les pays développés¹¹, on a 70 hommes pour 100 femmes au sein des plus de 60 ans et 44 hommes pour 100 femmes au sein des plus de 80 ans.

3.2. Impact du tabagisme

L'impact du tabagisme sur la mortalité est illustré sur la base de :

- L'étude de tables homme fumeur / homme non-fumeur canadiennes.
- Des études épidémiologiques menées par le ministère de la Santé en France.

3.2.1. Étude de tables canadiennes

Le caractère non-fumeur des assurés a un impact favorable sur leur mortalité. Au Canada, des études ont conduit à l'établissement de tables de mortalité pour les

¹⁰ Des éléments sur la mortalité en France sont fournis dans [ce post de blog](#).

¹¹ D'après une étude du US bureau of the census de 1998.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t;\infty[}(T_x)$$

hommes âgés de plus de 30 ans différenciées en fonction de cette caractéristique. Leur étude nous permet de quantifier l'impact du tabagisme sur la mortalité.

Le graphique suivant reprend ainsi l'évolution du taux de mortalité annuel en fonction de l'âge selon que l'homme fume ou ne fume pas.

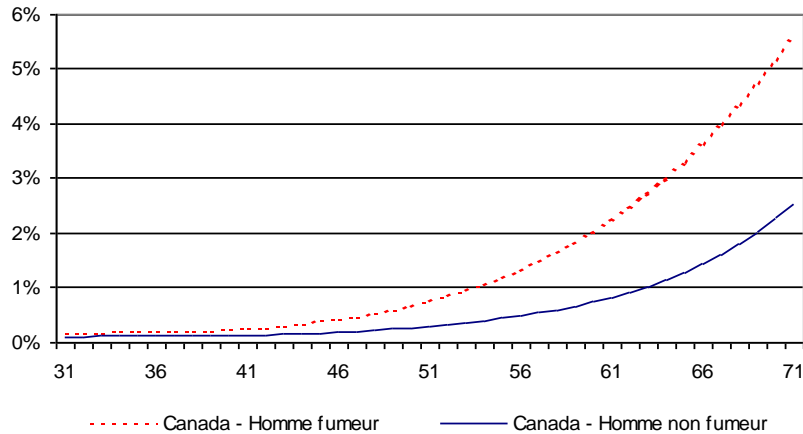


Fig. 11 : Comparaison des taux de décès fumeur et non-fumeur

Comme on pouvait s'y attendre, les taux de mortalité des fumeurs sont systématiquement supérieurs à ceux des non-fumeurs. L'abattement de la mortalité des non-fumeurs par rapport à celle des fumeurs connaît un maximum à 59 ans (63,35 %).

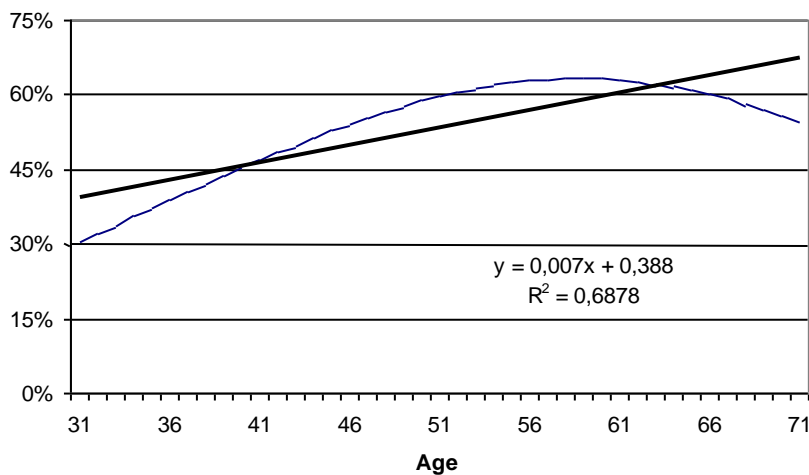


Fig. 12 : Mortalité des fumeurs par rapport à la population des fumeurs (hommes)

L'abattement moyen entre 31 ans et 71 ans ressort à 53,4 %. Cette analyse permet de mesurer l'impact du tabagisme sur la mortalité. Par rapport à l'ensemble de la population, l'abattement des non-fumeurs ressort en moyenne à 20 % entre 31 et 70 ans.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t; \infty[}(T_x)$$

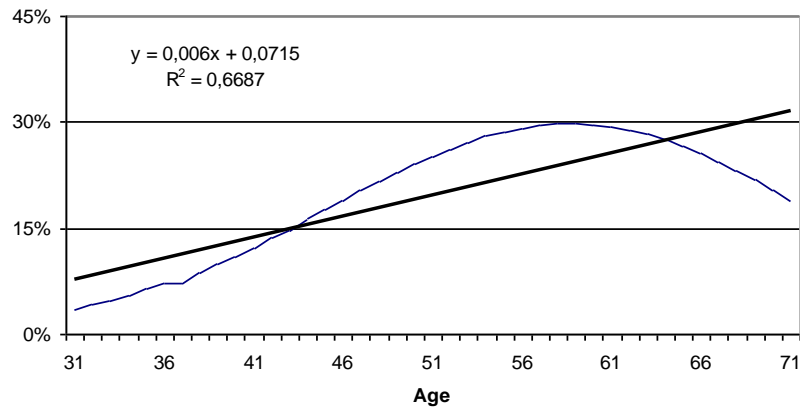


Fig. 13 : *Mortalité des non-fumeurs par rapport à la population générale (hommes)*

Un maximum est atteint à 58 ans avec un taux d'abattement de près de 30 %.

3.2.2. Étude de ministère de la Santé

L'étude « Tabagisme et mortalité : aspects épidémiologiques » fournit des indicateurs intéressants permettant de quantifier la sous-mortalité des non-fumeurs. Ainsi les éléments clés peuvent être résumés comme suit :

- Entre 39 et 65 ans, 1 décès sur 3 chez les hommes est attribuable au tabac et 1 décès sur 16 chez les femmes.
- Entre 35 et 49 ans, 40 % des hommes et 29 % des femmes sont des fumeurs réguliers. Ces pourcentages diminuent respectivement à 28 % et 14 % entre 50 et 64 ans.

En faisant l'hypothèse que ces proportions sont homogènes sur les plages d'âges indiquées, il est possible d'estimer la sous mortalité des non-fumeurs par rapport aux fumeurs. Notons :

- q_{tabac} le taux de sur mortalité lié au tabagisme,
- q le taux de mortalité hors tabagisme,
- π_F la proportion de fumeurs,
- π_{NF} la proportion de non-fumeurs,
- δ la proportion de décès dus au tabagisme.

$$\frac{\pi_F \times q_{\text{tabac}}}{\pi_F \times (q + q_{\text{tabac}}) + \pi_{NF} \times q} = \delta,$$

donc :

$$\frac{q_{\text{tabac}}}{q} = \frac{\delta}{\pi_F \times (1 - \delta)}$$

Le taux de sous mortalité des non-fumeurs par rapport au fumeur s'écrit donc :

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[}(T_x)$$

$$1 - \frac{q}{q + q_{\text{tabac}}} = 1 - \frac{\pi_F \times (1 - \delta)}{\delta + \pi_F \times (1 - \delta)}$$

Les taux calculés à partir des éléments de l'étude du ministère de la Santé sont résumés dans le tableau suivant :

	Hommes	Femmes
35-49 ans	55,56 %	20,41 %
50-64 ans	64,10 %	32,26 %

Concernant les hommes, les taux de sous mortalité des non-fumeurs par rapport aux fumeurs (56 % et 64 %) sont comparables à ceux issus des tables canadiennes (53 %). Ces mêmes taux sont inférieurs pour les femmes ; toutefois le tabagisme féminin est plus récent et son impact moins bien cerné que celui des hommes.

En supposant que les proportions de fumeurs citées plus haut sont homogènes sur toutes les tranches d'âges, les taux de sous-mortalité des non-fumeurs par rapport à la population dans son ensemble sont donnés par :

	Hommes	Femmes
35-49 ans	33,33 %	6,25 %
50-64 ans	31,71 %	6,25 %

Les taux masculins sont légèrement supérieurs à ce qui est observé avec les tables canadiennes. Les taux féminins sont nettement inférieurs aux taux masculins.

4. La construction de tables de mortalité d'expérience

4.1. Cadre général

La démarche de construction d'une table comporte globalement deux étapes :

- tout d'abord l'estimation de taux bruts, par âge, ou par âge et génération dans le cas de tables prospectives,
- puis ensuite l'ajustement de ces taux bruts à un modèle paramétrique ou leur régularisation dans le cadre semi ou non paramétrique.

En pratique on peut distinguer deux situations : tout d'abord, la situation « de référence » dans laquelle on dispose de données en quantité suffisante pour construire une table fiable. Mais dans certains cas il se peut que les données disponibles ne soient pas suffisantes pour déterminer de manière suffisamment précise la structure de la table, et on pourra alors chercher à positionner simplement la mortalité du groupe étudié par rapport à une mortalité de référence, qui fournira la structure générale.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t;\infty[}(T_x)$$

La démarche standard de construction d'une table de mortalité du moment dans un cadre paramétrique a été décrite précédemment¹², elle n'est donc pas reprise ici. On retiendra simplement qu'elle s'appuie sur le choix d'une forme paramétrique pour la fonction de hasard, avec comme modèle de référence le modèle de Makeham, l'estimation des paramètres s'effectuant par la méthode du maximum de vraisemblance.

Dans certaines situations particulières, on pourra toutefois se tourner vers d'autres modèles, tels que les régressions de type Poisson ; l'exemple type d'application de tels modèles est l'analyse de la mortalité d'un événement rare, comme les conséquences de l'exposition à l'amiante. La justification de l'intérêt de la loi de Poisson pour des événements « rares » provient de l'observation que la distribution binomiale $B\left(n, \frac{\lambda}{n}\right)$ converge en loi vers $P(\lambda)$ lorsque $n \rightarrow \infty$, ce qui se vérifie facilement avec les transformées de Laplace.

Comme le nombre de décès est très faible en regard des effectifs sous risque, on peut utiliser une loi de Poisson comme modèle pour le nombre de décès par âge et par période.

À titre d'exemple, la table suivante donne le nombre de décès par mésothéliome¹³ constaté par classe d'âge, pendant cinq périodes, ainsi que la population à risque pendant cette période :

		1970-74		1975-79		1980-84		1985-89		1990-95
25-29	0	10041742	1	10978690	1	10602254	1	10680272	1	10791607
30-34	2	7720583	1	10038396	0	11005461	1	10651073	1	10837520
35-39	2	8074903	3	7589268	5	9904593	5	10922900	6	10657919
40-44	5	8510762	7	7879250	7	7457766	10	9761988	11	10853140
45-49	9	8211522	9	8220829	14	7662805	14	7265550	17	9464014
50-54	10	7173352	18	7821153	22	7866442	26	7354438	24	7022582
55-59	16	4824443	20	6743790	32	7372021	41	7446988	41	7071006
60-64	28	6069611	26	4404567	38	6213936	58	6813378	68	6988969
65-69	33	5371770	42	5298248	41	3889820	63	5575185	84	6148376
70-74	34	4157113	49	4371284	56	4387290	56	3277849	72	4829840
75-79	24	2432745	37	3018047	53	3254297	73	3391145	64	2511709
80-84	10	1229739	22	1467570	35	1878692	54	2112437	63	2362417
85-89	7	527277	11	560756	16	691452	23	927740	31	1123450

Si on veut expliquer les décès en fonction de l'âge et de la période, on peut choisir deux séries de paramètres, (a_i) et (c_j) décrivant chacun l'effet d'une tranche d'âge donnée et d'une cohorte donnée. Pour satisfaire les contraintes de positivité (les décès sont un nombre positif), on peut proposer un modèle multiplicatif $a_i c_j$. On choisira par exemple de modéliser le nombre de décès espéré $\lambda_{ij} = E(d_{ij})$ avec un modèle de la forme

¹² Voir [ce support](#).

¹³ Cancer de la plèvre conséquence de l'exposition à l'amiante.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[}(T_x)$$

$\ln\left(\frac{\lambda_{ij}}{N_{ij}}\right) = a_i + c_j$ ou, de manière équivalente $\lambda_{ij} = N_{ij} \exp(a_i + c_j)$. Dans une cellule, on a finalement une vraisemblance liée à la loi de Poisson : $(N_{ij} \exp(a_i + c_j))^{d_{ij}} \exp(-N_{ij} \exp(a_i + c_j)) / d_{ij}!$, et la vraisemblance globale s'obtient en multipliant les vraisemblances de chaque cellule.

L'utilisation d'une référence externe consiste à rechercher un « positionnement » de la table d'expérience par rapport à une table de référence donnée ; la table de référence peut être par exemple une table INSEE.

De nombreux modèles sont possibles, mais l'approche la plus courante consiste à appliquer un taux d'abattement (ou de majoration) aux taux de la table de référence, ce qui consiste à rechercher un coefficient α tel que $q_x^{ex} = \alpha \times q_x^{ref}$. En se souvenant que le quotient de mortalité est la version discrète du taux de hasard μ_x (avec la relation $\mu_x = -\ln(1 - q_x)$ si on fait l'hypothèse de constance de la fonction de hasard entre deux âges entiers), on remarque que ce modèle est donc un modèle à hasard proportionnel dans lequel on suppose connue la fonction de hasard de base. Plus précisément, si on suppose que $\mu_x^{ex} = \alpha \times \mu_x^{ref}$ d'une part et que $\mu_x = -\ln(1 - q_x)$ d'autre part, on obtient la relation suivante entre les quotients de mortalité :

$$1 - q_x^{ex} = (1 - q_x^{ref})^\alpha,$$

relation qui au premier ordre lorsque les taux sont petits est équivalente à $q_x^{ex} \approx \alpha \times q_x^{ref}$. On a vu que dans ce contexte un estimateur de type « moindres carrés ordinaires » pouvait être proposé pour $\theta = -\ln(\alpha)$. On obtient ainsi l'estimateur :

$$\hat{\alpha} = \exp\left(\gamma - \frac{1}{n} \sum_{i=1}^n \ln H_{ref}(x_i)\right) = e^\gamma \left[\prod_{i=1}^n H_{ref}(x_i) \right]^{1/n}$$

avec γ la constante d'Euler¹⁴ et $H_{ref}(x) = -\ln S_{ref}(x) = -\ln L_x^{ref}$ la fonction de hasard cumulée.

On peut également considérer comme critère de choix du paramètre α l'écart entre le nombre de décès observés et le nombre de décès théorique associé à la table abattue. En notant L_x^{obs} l'effectif sous risque à l'âge x dans la population considérée, le nombre de décès prédit par la table abattue à l'âge x est $\alpha \times q_x^{ref} \times L_x^{obs}$. Si on contraint le nombre total de décès prédits à évaluer le nombre observé, on obtient l'estimation suivante de α :

¹⁴ Dont la valeur est approximativement 0,577215665.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[}(T_x)$$

$$\hat{\alpha} = \frac{\sum_x D_x^{obs}}{\sum_x q_x^{ref} \times L_x^{obs}}$$

Une approche alternative consiste à raisonner âge par âge et à considérer une statistique de type Khi-2 définie par :

$$\chi(\alpha) = \sum_{i=1}^n L_x^{obs} \frac{(q_x^{obs} - \alpha \times q_x^{ref})^2}{\alpha \times q_x^{ref}}$$

et à chercher la valeur de α qui rend minimale cette distance.

4.2. Tables prospectives

L'objectif de tables prospectives est de tenir compte des évolutions à venir de la mortalité ; les méthodes usuelles cherchent tout d'abord à ajuster les tendances passées, puis à les extrapoler à l'avenir. L'approche prospective consistant à intégrer dans l'avenir l'effet de progrès médicaux futurs n'est pas examinée ici.

Les modèles utilisés se proposent d'ajuster les taux bruts calculés par des méthodes telles que celle présentée en 2.3 ci-dessus à un modèle paramétrique, permettant d'une part de lisser les fluctuations d'échantillonnage et d'autre part de projeter l'évolution des taux dans le futur, par extrapolation.

On dispose de taux de taux bruts indicés par l'âge x et l'année calendaire t , qui ont typiquement l'allure suivante :

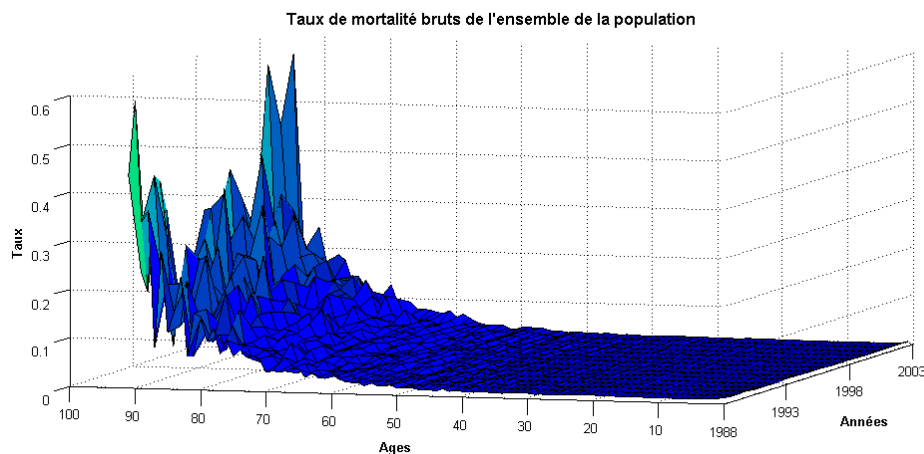


Fig. 14 : *Taux de décès bruts par année*

Le passage des quotients de mortalité bruts au taux instantané de mortalité, qui est la variable modélisée dans certaines approches, s'effectue *via* une hypothèse sur la répartition des décès dans l'année (voir 2.5 ci-dessus) ; dans le cas où l'on fait l'hypothèse de constance du taux instantané dans chaque carré du diagramme de Lexis, on obtient l'estimateur suivant :

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t;\infty} [(T_x)]$$

$$\mu_{xt}^* = -\ln(1 - \hat{q}_{xt}).$$

Une attention particulière doit être apportée au calcul de l'estimateur \hat{q}_{xt} pour éviter le mélange de générations différentes et les biais qui en résultent (cf. GUIBERT et PLANCHET [2017])

4.3. Le modèle de Lee-Carter

Il s'agit d'une méthode d'extrapolation des tendances passées initialement utilisée sur des données américaines, qui est devenue rapidement un standard (voir l'article original LEE et CARTER [1992]). La modélisation retenue pour le taux instantané de mortalité est la suivante :

$$\ln \mu_{xt} = \alpha_x + \beta_x k_t,$$

ou, de manière équivalente

$$\ln \mu_{xt}^* = \alpha_x + \beta_x k_t + \varepsilon_{xt}$$

L'idée du modèle est donc d'ajuster à la série (doublement indicée par x et t) des logarithmes des taux instantanés de décès une structure paramétrique (déterministe) à laquelle s'ajoute un phénomène aléatoire ; le critère d'optimisation retenu va consister à maximiser la variance expliquée par le modèle, ce qui revient à minimiser la variance des erreurs.

Le paramètre α_x s'interprète comme la valeur moyenne des $\ln(\mu_{xt})$ au cours du temps.

On vérifie que $\frac{d \ln(\mu_{xt})}{dt} = \beta_x \frac{dk_t}{dt}$ et on en déduit que le coefficient β_x traduit la sensibilité de la mortalité instantanée à l'âge x par rapport à l'évolution générale k_t , au sens où $\frac{d \ln(\mu_{xt})}{dk_t} = \beta_x$. En particulier, le modèle de Lee-Carter suppose la constance au cours du temps de cette sensibilité. Cette contrainte du modèle peut apparaître relativement forte :

- Pour tout âge x les quotients des variations relatives des taux de mortalité à des dates différentes ne dépendent pas de l'âge x . Si la variation relative du taux de mortalité à 50 ans en 2000 était 80 % de ce qu'elle était en 1990 ce coefficient de 80 % est retenu pour tous les âges ;
- Pour une même date t les quotients des variations relatives des taux de mortalité à des âges différents ne dépendent pas de la date t . Si en 2000 la variation relative du taux de mortalité à 20 ans est 50 % de la variation relative du taux à 50 ans ce coefficient de 50 % s'appliquera à toute date future ou passée.

Enfin, on peut remarquer que la forme du modèle implique l'homoscédasticité des taux de mortalité, ce qui est manifestement faux en pratique. Cet inconvénient sera examiné plus en détails en 4.4 ci-dessous.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t, \infty[}(T_x)$$

Afin de rendre le modèle identifiable, il convient d'ajouter des contraintes sur les paramètres ; en effet, pour toute constante c non nulle le modèle est invariant par les transformations suivantes :

$$\begin{aligned} (\alpha_x, \beta_x, k_t) &\rightarrow \left(\alpha_x, \frac{\beta_x}{c}, c \times k_t \right) \\ (\alpha_x, \beta_x, k_t) &\rightarrow (\alpha_x - c\beta_x, \beta_x, k_t + c) \end{aligned}$$

Il convient donc d'imposer deux contraintes sur les paramètres. On retient en général les contraintes suivantes :

$$\sum_{x=x_m}^{x_M} \beta_x = 1 \text{ et } \sum_{t=t_m}^{t_M} k_t = 0.$$

On obtient alors les paramètres par un critère de moindres carrés (non linéaire) :

$$(\hat{\alpha}_x, \hat{\beta}_x, \hat{k}_t) = \mathbf{arg\,min} \sum_{x,t} (\ln \mu_{xt}^* - \alpha_x - \beta_x k_t)^2$$

Il convient donc de résoudre ce programme d'optimisation, sous les contraintes d'identifiabilité. Le nombre de paramètres à estimer est élevé, il est égal à $2 \times (x_M - x_m + 1) + t_M - t_m - 1$.

4.3.1. Estimation des paramètres

Vis à vis de (α_x) , comme :

$$\frac{\partial}{\partial \alpha_x} \sum_{y,t} (\ln \mu_{yt}^* - \alpha_y - \beta_y k_t)^2 = 2(t_M - t_m + 1)\alpha_x - 2 \sum_{t=t_m}^{t_M} (\ln \mu_{xt}^* - \beta_x k_t),$$

on trouve en tenant compte de la contrainte $\sum_{t=t_m}^{t_M} k_t = 0$ que :

$$\hat{\alpha}_x = \frac{1}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \ln \mu_{xt}^*.$$

En d'autres termes, $\hat{\alpha}_x$ est la moyenne temporelle, à l'âge x , des taux instantanés de décès (sur l'échelle logarithmique). On considère alors la matrice $Z = (z_{xt})$ des taux centrés par rapport à la dimension temporelle :

$$z_{xt} = \ln \mu_{xt}^* - \hat{\alpha}_x.$$

Z est une matrice de dimension $(x_M - x_m + 1, t_M - t_m + 1)$. La forme du modèle revient à chercher une approximation de Z en produit de 2 vecteurs de la forme $Z \approx \hat{\beta} \hat{k}'$, de sorte que la décomposition soit optimale au sens du critère des moindres carrés, i.e. explique la plus grande part possible de la variance totale. La résolution de ce problème passe par la décomposition en valeurs propres de la matrice Z , que l'on met sous la forme :

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[}(T_x)$$

$$Z = \sum_{i \geq 1} \sqrt{\lambda_i} v_i u_i'$$

avec $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ les valeurs propres de $Z'Z$, u_i le vecteur propre normé¹⁵ de $Z'Z$ associé à λ_i , et v_i le vecteur propre associé à la même valeur propre pour ZZ' . Ceci est justifié par le fait que, comme $Z'Zu_i = \lambda_i u_i$, on a $ZZ'(Zu_i) = \lambda_i Zu_i$, et donc les deux matrices transposées ont les mêmes valeurs propres avec des ordres de multiplicité identiques. De plus, si u_i est un vecteur propre de $Z'Z$ alors Zu_i est un vecteur propre de ZZ' associé à la même valeur propre. Si on pose $v_i = \frac{1}{\sqrt{\lambda_i}} Zu_i$, on voit que $Zu_i u_i' = \sqrt{\lambda_i} v_i u_i'$, ce qui en sommant et en tenant compte de l'ortho normalité des vecteurs propres, conduit à la décomposition de Z .

On est ainsi conduit à proposer comme approximation $Z \approx \sqrt{\lambda_1} v_1 u_1'$, avec comme mesure de la qualité de cette approximation la part d'inertie expliquée, $\frac{\lambda_1}{\sum \lambda_i}$. On obtient finalement les estimateurs de β et k suivants :

$$\hat{\beta} = \frac{1}{\sum v_{1j}} v_1 \text{ et } \hat{k} = \sqrt{\lambda_1} \sum v_{1j} u_{1j}$$

L'objectif est d'utiliser les résultats de cet ajustement pour extrapoler les taux de mortalité pour $t > t_M$; l'idée est d'analyser la série des (\hat{k}_t) , qui capture l'information sur l'évolution temporelle de ces taux pour lui ajuster un modèle de type ARIMA.

À ce stade on dispose d'une première estimation des paramètres du modèle ; toutefois, si on se trouve dans une situation dans laquelle l'effectif soumis au risque est très important, on peut se dire que les fluctuations d'échantillonnage sur le nombre total de décès par âge doivent être très faibles. Il apparaît alors souhaitable d'ajuster auparavant les paramètres du modèle pour que le nombre de décès prévus par le modèle chaque année soit égal au nombre de décès observés. Comme l'exposition au risque est $L_{xt} = \frac{D_{xt}}{\mu_{xt}^*}$

avec $D_{xt} = D_{xt}(t-x) + D_{xt}(t-x-1)$, cette contrainte s'exprime par :

$$\sum_{x=x_m}^{x_M} D_{xt} = \sum_{x=x_m}^{x_M} L_{xt} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{k}_t),$$

la variable étant \hat{k}_t . On introduit la fonction $F(k) = \sum_{x=x_m}^{x_M} L_{xt} \exp(\hat{\alpha}_x + \hat{\beta}_x k) - \sum_{x=x_m}^{x_M} D_{xt}$, de sorte que la contrainte ci-dessus s'exprime par $F(k) = 0$. La forme de la fonction F assure

¹⁵ $u_i' u_i = 1$.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[}(T_x)$$

l'unicité de la racine si elle existe. La recherche de la racine peut se faire par un algorithme de type Newton-Raphson, en posant :

$$k_{i+1} = k_i - \frac{F(k_i)}{F'(k_i)}$$

avec la valeur initiale $k_0 = \hat{k}_t$ et le critère d'arrêt $\frac{k_{i+1} - k_i}{k_i} \leq \varepsilon$ en prenant par exemple

$\varepsilon = 10^{-7}$. On obtient ainsi un nouvel estimateur \hat{k}_t . Mais la série (\hat{k}_t) ainsi ajustée doit

être corrigée pour respecter la contrainte d'identifiabilité $\sum_{t=t_m}^{t_M} k_t = 0$, ce qui conduit à poser :

$$k_t^* = \hat{k}_t - \frac{1}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \hat{k}_t.$$

Il reste alors à corriger les $(\hat{\alpha}_x)$ pour que l'égalité entre décès prédits par le modèle et décès observés reste valide, ce qui conduit à :

$$\alpha_x^* = \hat{\alpha}_x + \frac{\hat{\beta}_x}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \hat{k}_t.$$

L'allure « typique » des paramètres obtenus est la suivante, tout d'abord pour les paramètres fonction de l'âge :

Paramètres α_x

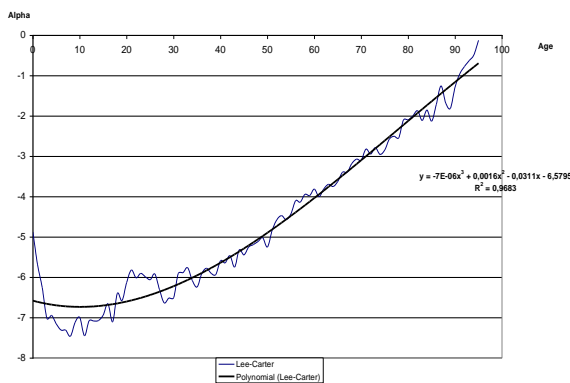


Fig. 15 : Coefficients alpha

Paramètres β_x

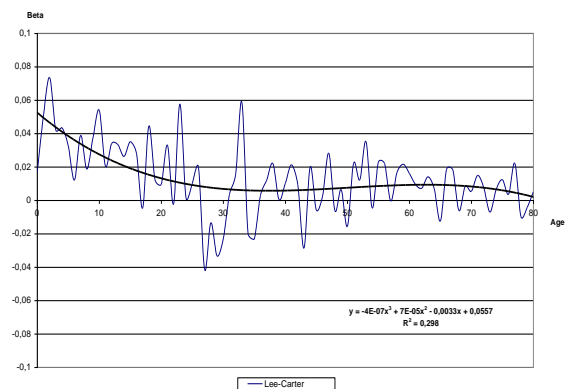


Fig. 16 : Coefficients beta

puis la composante temporelle :

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t; \infty} [(T_x)]$$

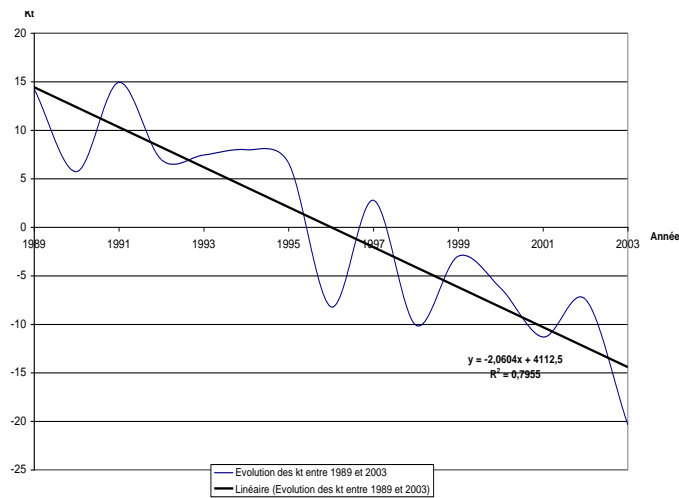


Fig. 17 : *Évolution de la tendance temporelle*

4.3.2. Extrapolation de la composante temporelle

Il reste alors à modéliser la série (k_t^*) pour extrapoler les taux futurs ; pour cela, on utilise en général un modèle ARIMA¹⁶, mais toute autre modélisation de série temporelle peut être utilisée. Toutefois, compte tenu de l'allure du graphe ci-dessus, la modélisation la plus simple que l'on puisse imaginer est par exemple une régression linéaire de ces coefficients :

$$k_t^* = at + b + \varepsilon_t$$

avec (ε_t) un bruit blanc gaussien.

4.3.3. Fermeture de la table

L'estimation des paramètres du modèle de Lee-Carter nécessite que l'on dispose d'une matrice rectangulaire complète de taux de décès (μ_{xt}^*) ; en pratique les valeurs brutes estimées présentent une grande instabilité aux âges élevés, du fait du faible effectif disponible. Au surplus, il peut arriver que les données ne soient plus disponibles au-delà d'un âge limite. Plusieurs méthodes existent pour compléter la table avant d'effectuer l'ajustement, ou *ex-post* (on parle de « fermeture de la table de mortalité »). On pourra notamment consulter sur la sujet DENUIT et QUASHIE [2005].

À titre d'illustration, on présente ici la méthode de Coale et Kisker (COALE et KISKER [1990]) ; la méthode consiste à extrapoler les taux de mortalité aux grands âges (jusqu'à $x=110$ ans par exemple) en se basant sur la formule¹⁷ :

$$\hat{\mu}_x = \hat{\mu}_{65} \times e^{g_x(x-65)}$$

¹⁶ En suivant la démarche de Box et Jenkins.

¹⁷ On omet ici l'indice t pour alléger les notations.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t;\infty}[(T_x)]$$

ressources-actuarielles.net

g_x désignant le taux moyen de croissance de μ_x entre 65 et x ans. On calcule ainsi les coefficients g_x jusqu'à un certain âge, puis on les extrapole afin de pouvoir recomposer les taux μ_x . Coale et Kisker ont en effet remarqué empiriquement que les courbes des g_x possèdent en général un pic aux alentours de 80 ans avant de décroître linéairement. Ils ont par conséquent proposé l'équation :

$$g_x = g_{80} + s \times (x - 80), \quad x \geq 80.$$

Finalement, on peut utiliser la formule suivante pour extrapoler au-delà de 80 ans les taux instantanés de mortalité :

$$\hat{\mu}_x = \hat{\mu}_{x-1} \times e^{g_{80} + s(x-80)}, \quad x \geq 80.$$

On utilise les valeurs de paramètres suivantes :

$$s = -\frac{\ln(\hat{\mu}_{79} + 31 \times g_{80})}{465} \quad \text{et} \quad g_{80} = \frac{\ln\left(\frac{\hat{\mu}_{80}}{\hat{\mu}_{65}}\right)}{15}.$$

Ainsi, les taux de mortalité lissés à partir des données brutes sont directement obtenus par la méthode de lissage de Lee-Carter pour les âges inférieurs à 80 ans. Pour les âges supérieurs ou égaux à 80 ans, si l'échantillonnage n'est pas assez conséquent, on recourt à la méthode de Coale et Kisker : celle-ci construit les taux de mortalité aux grands âges à partir des taux lissés (par Lee-Carter) aux âges de 65 et 80 ans.

La question de la fermeture de la table est importante dans le cas de la construction d'une table pour des provisionnements de rentes viagères. On pourra toutefois noter que cette importance doit être relativisée si les rentiers d'âge très élevé sont en proportion modeste dans le portefeuille.

En effet, considérons l'exemple simple dans lequel on évalue un capital constitutif d'une rente viagère sur une tête avec la table TF00-02 ; on compare le calcul réalisé avec la table complète et celui réalisé avec la même table fermée de manière prudente en figeant le taux de décès à 95 ans. Ainsi, si pour évaluer le capital constitutif d'une rente viagère sur une tête à 75 ans au taux de 2,5 % (et avec la TH00-02), on considère que le taux de décès est stable à partir de 95 ans (et que les survivants sortent brutalement à 120 ans), on ne majore la provision que de 0,7 % (et environ 2,5 % à 85 ans).

L'écart entre deux méthodes de fermeture en termes de provisionnement n'est véritablement significatif qu'à des âges très élevés (voir par exemple DELWARDE et DENUIT [2006]).

4.4. Le modèle log-Poisson

Le modèle de Lee-Carter repose sur l'hypothèse d'homoscédasticité des taux de mortalité, ce qui constitue une hypothèse forte et peu réaliste : en effet, la variance des taux de décès croît aux âges élevés, du fait notamment de la baisse des effectifs de survivants. On peut illustrer ce fait de deux manières ; tout d'abord, on considère la population française au 01/01/2005, que l'on suppose mourir selon la table TV1999/2001.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[}(T_x)$$

La variance des taux de décès bruts que l'on observerait est approchée par $\frac{q_x(1-q_x)}{L_x}$, et on constate l'évolution suivante :

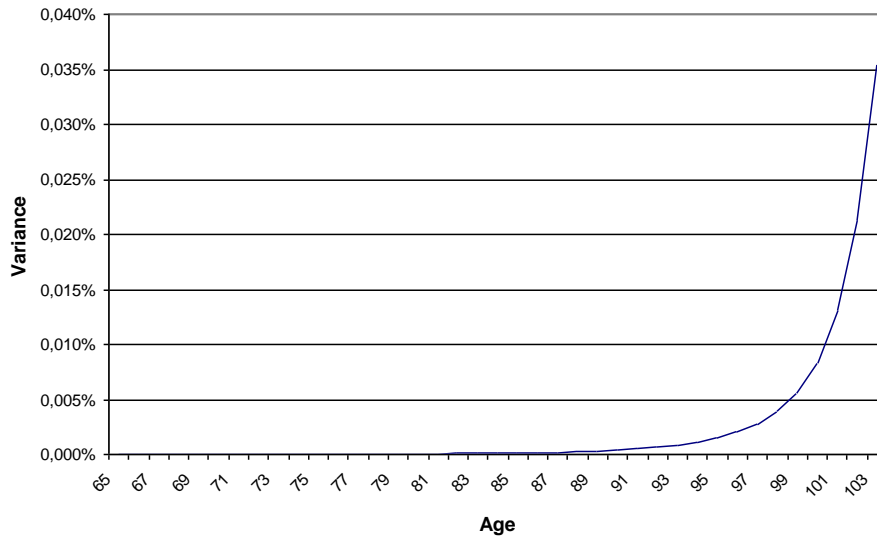


Fig. 18 : *Variance du taux de décès avec l'âge*

On note une très forte augmentation après 85 ans. De manière plus directe, lorsque l'on effectue un ajustement par la méthode de Lee-Carter, on peut analyser la variance des résidus, et confronter les observations à l'hypothèse d'hétéroscédasticité. On obtient des graphiques à l'allure suivante¹⁸ :

¹⁸ Voir LELIEUR [2005].

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t, \infty[}(T_x)$$

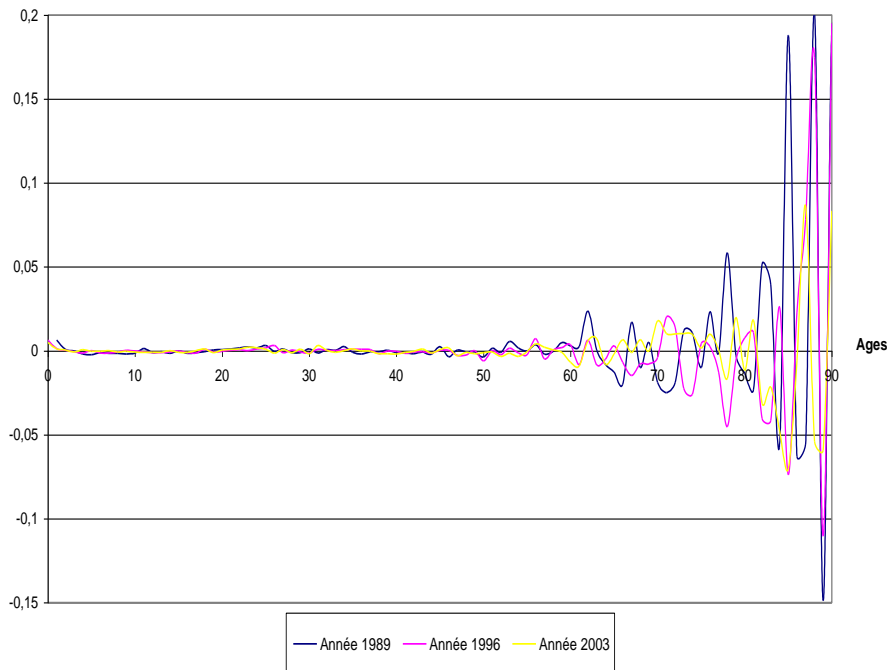


Fig. 19 : Analyse des résidus dans le modèle de Lee-Carter

De plus, le critère retenu dans la méthode de Lee-Carter pour estimer les paramètres n'est pas de type « maximum de vraisemblance ».

Au surplus, il peut sembler naturel de modéliser directement le nombre de décès plutôt que le taux instantané de mortalité. Si D_{xt} désigne le nombre de décès à l'âge x l'année t , et L_{xt} l'exposition au risque, D_{xt} est alors une variable aléatoire dont on va modéliser l'espérance en posant :

$$E(D_{xt}) = L_{xt} \times \mu_{xt}$$

Le modèle log-Poisson, proposé par BROUHNS et al. [2002], est une adaptation du modèle de Lee-Carter qui intègre ces différents éléments. On notera que l'égalité ci-dessus est la conséquence directe de l'hypothèse de constance de μ_{xt} sur chaque carré du diagramme de Lexis. En effet, comme l'exposition au risque est égale à :

$$L_{xt} = \int_0^1 S(x+u, t+u) du$$

et que $S(x+u, t+u) = S(x, t) \exp\left(-\int_0^u \mu(x+v, t+v) dv\right)$, la constance de μ_{xt} conduit à :

$$S(x+u, t+u) = S(x, t) \exp(-u \times \mu(x, t))$$

puis $L_{xt} = S(x, t) \int_0^1 \exp(-u \times \mu(x, t)) du$ et donc :

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t, \infty[}(T_x)$$

$$L_{xt} = \frac{S(x,t)}{\mu(x,t)} (1 - \exp(-\mu(x,t))) = \frac{S(x,t)}{\mu(x,t)} q(x,t),$$

ce qui établit le résultat.

L'idée est de modéliser le nombre de décès à l'âge x l'année t D_{xt} par une loi de Poisson de paramètre $L_{xt}\mu_{xt}$ avec $\mu_{xt} = \exp(\alpha_x + \beta_x k_t)$. L'expression du taux de décès instantané est identique à celle proposée dans le modèle de Lee-Carter, avec la même interprétation des différents paramètres. En particulier, le modèle ne sera identifiable qu'avec des contraintes sur les paramètres, et on peut retenir les mêmes que celles utilisées par Lee et Carter. Enfin, on peut noter que passer du modèle de Lee-Carter à ce modèle poissonien revient à passer d'un modèle linéaire à un modèle linéaire généralisé avec le logarithme comme fonction de lien¹⁹.

4.4.1. Estimation des paramètres

Comme on a $P(D_{xt} = d) = \frac{(L_{xt}\mu_{xt})^d}{d!} \exp(-L_{xt}\mu_{xt})$ avec $\mu_{xt} = \exp(\alpha_x + \beta_x k_t)$, la log-vraisemblance²⁰ du modèle s'écrit (à une constante additive près) :

$$\ln L(\alpha, \beta, k) = \sum_{x,t} \{D_{xt}(\alpha_x + \beta_x k_t) - L_{xt} \exp(\alpha_x + \beta_x k_t)\}$$

On dispose donc d'une expression simple de la log-vraisemblance ; les équations de vraisemblance n'ont pas de solution analytique du fait de la présence du terme non linéaire $\beta_x k_t$ et doivent être résolues numériquement ; on peut par exemple utiliser un algorithme de Newton-Raphson et utiliser le schéma proposé en 4.3.1 ci-dessus avec pour fonction objectif F à annuler le vecteur des scores $\left(\frac{\partial L}{\partial \alpha}, \frac{\partial L}{\partial \beta}, \frac{\partial L}{\partial k} \right)$; cela conduit ici aux relations de récurrence suivantes :

$$\hat{\alpha}_x^{i+1} = \hat{\alpha}_x^i - \frac{\sum_t (D_{xt} - L_{xt} \exp(\hat{\alpha}_x^i + \hat{\beta}_x^i k_t^i))}{-\sum_t (L_{xt} \exp(\hat{\alpha}_x^i + \hat{\beta}_x^i k_t^i))}$$

$$\hat{k}_t^{i+1} = \hat{k}_t^i - \frac{\sum_x (D_{xt} - L_{xt} \exp(\hat{\alpha}_x^{i+1} + \hat{\beta}_x^i k_t^i)) \hat{\beta}_x^i}{-\sum_x (L_{xt} \exp(\hat{\alpha}_x^{i+1} + \hat{\beta}_x^i k_t^i)) (\hat{\beta}_x^i)^2}$$

¹⁹ On pourra se reporter à RENSHAW [1991].

²⁰ Il ne s'agit d'une vraisemblance que si on utilise les effectifs sous risque réels ; si on normalise les effectifs en partant d'un effectif initial de L_0 , on obtient une pseudo-vraisemblance.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[} (T_x)$$

$$\hat{\beta}_x^{i+1} = \hat{\beta}_x^i - \frac{\sum_t (D_{xt} - L_{xt} \exp(\hat{\alpha}_x^{i+1} + \hat{\beta}_x^i \hat{k}_t^{i+1})) \hat{k}_t^i}{-\sum_t (L_{xt} \exp(\hat{\alpha}_x^{i+1} + \hat{\beta}_x^i \hat{k}_t^{i+1})) (\hat{k}_t^{i+1})^2}$$

Les valeurs initiales sont libres, on choisira simplement des valeurs $\hat{\beta}_x^0 \neq 0$ pour éviter des divisions par 0. Pour que les contraintes d'identifiabilité soient vérifiées, il convient ensuite d'ajuster les paramètres ainsi estimés, en posant :

$$k_t^* = \left(\hat{k}_t - \frac{1}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \hat{k}_t \right) \sum_x \hat{\beta}_x$$

$$\beta_x^* = \frac{\hat{\beta}_x}{\sum_x \hat{\beta}_x}$$

$$\alpha_x^* = \hat{\alpha}_x + \frac{\hat{\beta}_x}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \hat{k}_t$$

Les valeurs estimées des paramètres sont assez proches de celles obtenues par le modèle de Lee-Carter, comme on peut le constater sur les graphiques repris en annexe²¹. L'extrapolation de la composante temporelle est ensuite conduite de la même manière que dans le modèle de Lee-Carter.

4.4.2. Obtention d'intervalles de confiance²²

En pratique les tables ainsi construites vont en général servir à calculer des espérances de vie résiduelle, pour obtenir des durées de vie de rentiers (actuels et futurs); plus précisément elles pourront être utilisées pour calculer des capitaux constitutifs de rentes viagères, de la forme :

$$a_{xt} = \sum_{i \geq 0} v^{i+1} \prod_{j=0}^i \exp(-\mu_{x+j,t+j})$$

avec $v = \frac{1}{1+r}$ le facteur d'actualisation. Au-delà de l'estimation ponctuelle de a_{xt} qui découle de la modélisation des μ_{xt} , on souhaite mesurer la précision associée, et donc obtenir des intervalles de confiance. Deux sources d'aléa se combinent ici, d'une part les fluctuations d'échantillonnage du modèle de régression poissonien, et d'autre part l'incertitude liée à la prédiction des k_t pour $t > t_M$.

²¹ Ces graphiques sont repris de BROUHNS et al. [2002].

²² Seul le principe de la méthode est décrit ici ; pour l'approche détaillée on pourra se reporter à HADERER [2003].

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[} (T_x)$$

L'estimation des paramètres du modèle (pour $t \leq t_M$) par la méthode du maximum de vraisemblance permet de conclure que le vecteur $(\hat{\alpha}_x, \hat{\beta}_x, \hat{k}_t)$ est asymptotiquement distribué selon une loi normale. On peut alors construire alors des intervalles de confiance pour des fonctionnelles telles que a_{xt} par la méthode de simulation suivante :

- on génère une réalisation $(\hat{\alpha}_x, \hat{\beta}_x, \hat{k}_t)$ à partir de la loi normale ;
- à partir de la réalisation ci-dessus, on estime les paramètres de projection du modèle ARIMA associé aux k_t ;
- on simule une trajectoire de k_t pour $t > t_M$
- à partir des éléments ainsi calculés, on détermine une réalisation de la variable d'intérêt (par exemple a_{xt})

En renouvelant l'opération on obtient une distribution empirique de la variable d'intérêt, puis, en particulier, un intervalle de confiance.

Lorsque la taille de l'échantillon est très importante on peut considérer que les fluctuations d'échantillonnage deviennent négligeables, et supprimer la première étape de l'algorithme.

Ce type d'application sera développé dans le cadre des modèles stochastiques de mortalité²³.

4.5. Les modèles log-linéaires

Dans le choix d'un modèle susceptible de structurer un jeu de données historiques, la « flexibilité » du modèle et par la même sa fidélité aux données est directement liée aux nombres de paramètres introduits. Le choix d'un modèle très flexible se fait le plus souvent au détriment des qualités prédictives de celui-ci (un modèle totalement non paramétrique n'autorise aucune prédiction).

Les modèles de Lee-Carter ou Log-Poisson peuvent de ce fait paraître très paramétrés. Au surplus, dans le contexte de données de portefeuilles, dont le volume est sensiblement inférieur à ce que l'on peut obtenir comme taille de population sous risque à l'échelle d'un pays, le nombre élevé de paramètres du modèle peut conduire à des irrégularités consécutives de fluctuations d'échantillonnage. Ce phénomène est mis en évidence dans LELIEUR [2005].

Dans ce contexte il peut être utile de se tourner vers des modèles alternatifs moins paramétrés mettant en jeu des expressions analytiques portant sur les âges ou sur les années (ou les deux).

Par ailleurs les influences de l'âge x et de l'année t sur les taux de mortalité $q_x(t)$ sont exprimées *via* l'introduction du *logit* :

²³ Voir [ce support](#).

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t, \infty[}(T_x)$$

$$\mathbf{lg}(q_{xt}) = \mathbf{lg}_x(t) = \ln(q_{xt} / (1 - q_{xt})).$$

Le logit pour des taux de mortalités faibles est peu différent de la variable $\ln(\mu_{xt})$ du modèle de Lee-Carter mais il peut être sensiblement différent pour des âges élevés. Il présente l'avantage de varier dans $] -\infty, +\infty[$, ce qui simplifie la mise en œuvre de modèle de régression. La forme typique d'un logit est la suivante (obtenue avec la TV 1999/2001) :

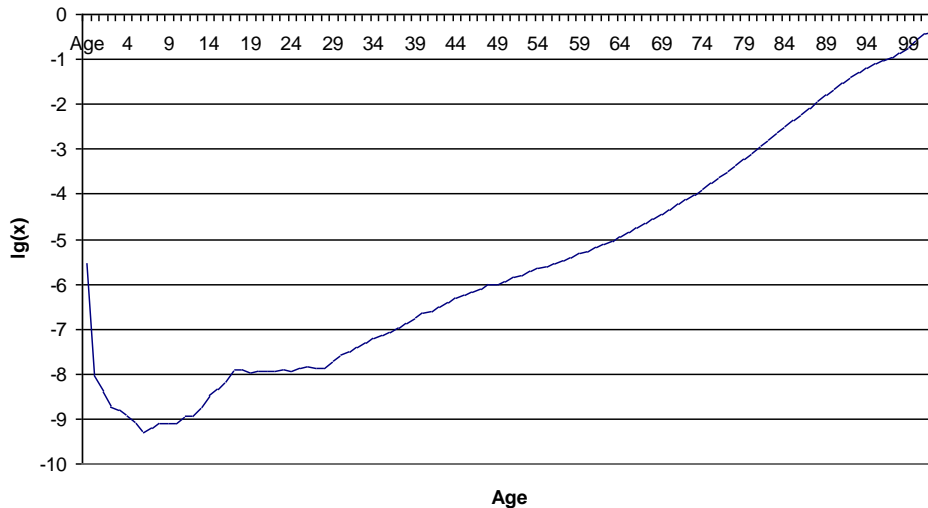


Fig. 20 : Logit du taux de décès en fonction de l'âge (table TV 99-01)

On est ainsi conduit à introduire les modèles log-linéaires. Le modèle de base de cette famille impose une tendance linéaire en fonction du temps :

$$\mathbf{lg}_x(t) = \alpha_x + \beta_x \times t + \varepsilon_{xt}.$$

On suppose les résidus iid (et donc homoscédastiques), ce qui permet d'utiliser les résultats standards du modèle linéaire ordinaire (avec l'année calendaire t comme variable explicative, à x fixé). Cette paramétrisation est proche de celle du modèle de Lee-Carter dans lequel on aurait supposé $k_t = t$ et remplacé $\ln(\mu_{xt})$ par $\mathbf{lg}_x(t)$. Ce modèle est en particulier utilisé pour la construction des tables TPG 1993. On dispose d'expressions explicites pour les paramètres. En effet, en se souvenant que dans le modèle linéaire $y_i = ax_i + b + \varepsilon_i$ on a :

$$\hat{a} = \frac{\frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \times \bar{y}}{\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2} = \frac{\mathbf{cov}(x, y)}{\mathbf{var}(x)} \text{ et } \hat{b} = \bar{y} - \hat{a}\bar{x},$$

avec $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ on obtient facilement les expressions des coefficients α_x et β_x .

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t;\infty[}(T_x)$$

On constate empiriquement une très forte corrélation entre les séries (α_x) et (β_x) , ce qui conduit à proposer une variante du modèle dans laquelle ces deux coefficients sont liés par une relation affine ; cela conduit au modèle suivant :

$$\mathbf{lg}_x(t) = a\beta_x + b + \beta_x \times t + \varepsilon_{xt} = (a+t)\beta_x + b + \varepsilon_{xt}$$

Le nombre de paramètres à estimer diminue sensiblement pour s'établir à $2 + x_M - x_m + 1$ au lieu de $2 \times (x_M - x_m + 1)$ dans le modèle précédent et de $2 \times (x_M - x_m + 1) + t_M - t_m - 1$ dans le modèle de Lee-Carter. Cependant le problème de moindres carrés devient non linéaire, ce qui complique un peu l'estimation des paramètres²⁴. En pratique on doit avoir recours à des méthodes numériques alors que dans la première version du modèle on dispose d'une expression explicite directe des paramètres.

La dérive linéaire peut apparaître irréaliste sur le long terme, et on constate par exemple sur des données américaines un ralentissement de la tendance. On peut alors chercher des modélisations permettant d'introduire au niveau des prévisions de très long terme des informations exogènes traduisant un ralentissement prévisible de la dérive. Ceci peut être réalisé avec les modèles suivants :

$$\mathbf{lg}_x(t) = \alpha_x + \beta_x \times t + \gamma_x \times t^\delta + \varepsilon_{xt}.$$

Dans ces modèles les estimations font également apparaître une très forte corrélation entre les estimations des paramètres (α_x) , (β_x) et (γ_x) , ce qui incite à proposer deux nouveaux modèles en posant $\alpha_x = a\gamma_x + b$ et $\beta_x = c\gamma_x + d$ et conduit à la spécification :

$$\mathbf{lg}_x(t) = b + d t + \gamma_x (t^\delta + c t + a) + \varepsilon_{xt}.$$

La résolution numérique du critère de moindres carrés associé n'appelle pas de commentaire particulier.

4.6. Le modèle logistique décalé

On considère ici le modèle proposé par BONGAART [2004] et défini par :

$$\mu_{xt} = \frac{\alpha(t) \exp(\beta x)}{1 + \alpha(t) \exp(\beta x)} + \gamma(t).$$

Comme l'inverse de la fonction logistique $\mathbf{lg}(x) = \ln\left(\frac{x}{1-x}\right)$ est $y \rightarrow \frac{e^y}{1+e^y}$, on en déduit en écrivant :

$$\mu_{xt} - \gamma(t) = \frac{\exp(\beta x + \delta(t))}{1 + \exp(\beta x + \delta(t))}$$

²⁴ Qui doit être effectuée globalement et non plus âge par âge.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t; \infty[}(T_x)$$

ressources-actuarielles.net

Modèles de durée

avec $\delta(t) = \ln(\alpha(t))$ que ce modèle peut également s'écrire :

$$\lg(\mu_{xt} - \gamma(t)) = \beta x + \ln(\alpha(t)).$$

Ce modèle est en fait une généralisation du modèle de Makeham (MAKEHAM [1860])

$\mu_x = \alpha \exp(\beta x) + \gamma$ proposée par THATCHER [1999] en posant $\mu_x = \frac{\alpha \exp(\beta x)}{1 + \alpha \exp(\beta x)} + \gamma$ que

l'on adapte au cas de taux de décès non constants au cours du temps. Cet ajustement du modèle de Makeham est motivé originellement par la volonté de corriger la surestimation des taux de décès conditionnels aux âges élevés observée en pratique.

Le fait β soit indépendant du temps est la conséquence du fait qu'on constate empiriquement que ce paramètre dépend peu du temps.

L'estimation des paramètres peut être effectuée par une méthode de moindres carrés

non linéaires en minimisant $\varepsilon = \sum_{x,t} n_{xt} \frac{(\hat{q}_{xt} - q_{xt})^2}{q_{xt}}$ avec n_{xt} l'exposition au risque pour l'âge

et l'année considérés²⁵. Une fois le modèle ajusté sur les valeurs passées, l'extrapolation de la mortalité future se ramène à une extrapolation, *via* des techniques de séries temporelles, des coefficients $\alpha(t)$ et $\gamma(t)$. Cette paramétrisation présente l'intérêt d'être moins contrainte dans la dimension temporelle que Lee-Carter ou log-Poisson, l'extrapolation reposant sur 2 paramètres et non un seul.

Le calcul de q_x en fonction de $\mu_x = \frac{\alpha \exp(\beta x)}{1 + \alpha \exp(\beta x)} + \gamma$ est effectué *via*

$$q_x = 1 - \exp\left(-\int_x^{x+1} \mu(u) du\right) = 1 - \exp\left(-\int_x^{x+1} \left(\frac{\alpha \exp(\beta u)}{1 + \alpha \exp(\beta u)} + \gamma\right) du\right).$$

En posant $v_{\alpha, \beta}(u) = 1 + \alpha \exp(\beta u)$ on remarque que $\frac{\alpha \exp(\beta u)}{1 + \alpha \exp(\beta u)} du = \frac{1}{\beta} \frac{dv}{v}$, ce qui conduit après quelques manipulations à :

$$q_x = 1 - e^{-\gamma} \left[\frac{v_{\alpha, \beta}(x)}{v_{\alpha, \beta}(x+1)} \right]^{\frac{1}{\beta}}.$$

En d'autres termes, la fonction de survie de ce modèle est $S(x) = e^{-\gamma x} \left(\frac{v_{\alpha, \beta}(x)}{v_{\alpha, \beta}(0)} \right)^{-\frac{1}{\beta}}$. Le

terme de « modèle décalé » est motivé par l'observation suivante : si on ne considère que la composante du taux de mortalité associée au vieillissement,

²⁵ En pratique ce critère est proche d'un maximum de vraisemblance discrétisé (voir le support sur les modèles paramétrique).

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t, \infty} [(T_x)]$$

$$\mu_{xt}^s = \frac{\alpha(t) \exp(\beta x)}{1 + \alpha(t) \exp(\beta x)}$$

alors pour une année t_0 fixée on peut écrire pour $t \geq t_0$:

$$\mu_{xt}^s = \frac{\alpha(t_0) \exp(\beta(x - \Lambda(t)))}{1 + \alpha(t_0) \exp(\beta(x - \Lambda(t)))}$$

avec $\Lambda(t) = -\frac{1}{\beta} \ln \left(\frac{\alpha(t)}{\alpha(t_0)} \right)$.

4.7. Utilisation des séries chronologiques

Le modèle de Lee-Carter, après avoir ajusté sur les données historiques les paramètres α , β et k , propose de considérer la suite des k_t comme une série chronologique pour obtenir les valeurs prospectives des taux. On est ainsi conduit à poser :

$$k_t^* = at + b + \varepsilon_t$$

Cette approche peut être transposée dans le cadre des modèles logistiques, dans le but de réduire le nombre de paramètres. On cherche alors à paramétrer la fonction $x \rightarrow \mathbf{lg}_x(t)$ pour prendre en compte l'influence de l'année t de manière non paramétrique, puis à modéliser dans un second temps les séries chronologiques introduites. On considère ainsi un modèle de la forme :

$$\mathbf{lg}_x(t) = f(x, \theta_t) + \varepsilon_{xt}$$

où la fonction $f(x, \theta_t)$ est choisie, pour des arguments de simplicité de mise en œuvre, linéaire par rapport au paramètre (vectoriel) θ_t . Dans une deuxième phase la série (θ_t) est modélisée.

La forme retenue pour f est celle d'une spline cubique avec des nœuds aux âges $(x_i, i = 1, \dots, p)$. La forme de la fonction f avec p nœuds est alors la suivante :

$$f(x, a, b, c, d, e_1, \dots, e_p) = a_t + b_t x + c_t x^2 + d_t x^3 + \sum_{i=1}^p e_{i,t} \left[(x - x_i)^+ \right]^3$$

En pratique, une version simplifiée de ce modèle dans laquelle seul le paramètre a_t dépend du temps fournit des résultats fiables. En observant que la modélisation de a_t au travers d'une régression linéaire analogue à celle menée pour k_t , on peut construire une version entièrement paramétrique du modèle en proposant :

$$f(x, a, b, c, d, e_1, \dots, e_p) = a + \lambda t + bx + cx^2 + dx^3 + \sum_{i=1}^p e_i \left[(x - x_i)^+ \right]^3$$

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t; \infty} [(T_x)]$$

4.8. Les modèles à référence externe

Si on ne dispose pas de données suffisantes pour structurer correctement la table complète, on peut imaginer d'utiliser la structure d'une table de référence existante et de simplement positionner la mortalité du groupe considéré par rapport à cette référence.

Deux approches sont envisageables pour atteindre cet objectif, elles sont présentées succinctement ci-après.

4.8.1. Régression logistique

Lorsque l'on souhaite positionner une table par rapport à une autre, il peut apparaître naturel d'effectuer la régression des logits des taux bruts sur les logits de la table de référence²⁶, ce qui conduit au modèle suivant, proposé initialement dans BRASS [1971] :

$$\ln(q_{xt} / (1 - q_{xt})) = a \times \ln(q_{xt}^{ref} / (1 - q_{xt}^{ref})) + b,$$

avec $q_{xt} = q_{xt}(a, b)$ le taux de mortalité d'expérience, ou encore :

$$\mathbf{lg}_x(t) = a \times \mathbf{lg}_x^{ref}(t) + b.$$

L'estimation des paramètres a et b est effectuée en minimisant une fonction de perte comme :

$$\varphi(a, b) = \sum_{x,t} E_{x,t} \times (q_{xt}(a, b) - \hat{q}_{x,t})^2,$$

qui est la généralisation naturelle du critère de maximum de vraisemblance dans un modèle discret, ou d'autres critères, comme :

$$\varphi(a, b) = \sum_x \left| \sum_t E_{x,t} \times q_{xt}(a, b) - D_{x,t} \right| + \sum_t \left| \sum_x E_{x,t} \times q_{xt}(a, b) - D_{x,t} \right|.$$

Elle permet, au surplus, une extrapolation aisée des logits des taux d'expérience dans les plages d'âge pour lesquelles les données d'expérience seraient insuffisantes.

On peut adapter le critère d'optimisation utilisé pour tenir compte du contexte d'utilisation des tables en retenant plutôt :

$$(\hat{a}, \hat{b}) = \mathbf{arg\,min} \left[e_{60}^{lissé}(a, b) - e_{60}^{nonlissé} \right],$$

sous la contrainte suivante :

$$e_{60}^{lissé}(a, b) - e_{60}^{nonlissé} > 0.$$

²⁶ Voir notamment l'utilisation de ce modèle dans <http://www.ressources-actuarielles.net/gtmortalite>, qui fournit aussi des tables de référence usuelles : TGH/F 05, INSEE 2060 et IA 2013.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{]t, \infty[}(T_x)$$

ressources-actuarielles.net

Modèles de durée

où $e_{60}^{\text{lissé}}(a, b)$ désigne l'espérance de vie résiduelle à 60 ans, fonction des paramètres a et b , calculée à partir de la régression sur les logits et $e_{60}^{\text{nonlissé}}$ désignant l'espérance de vie résiduelle à 60 ans calculée à partir des données brutes. On perd alors le caractère explicite de l'expression des paramètres. Le détail de l'approche est présenté dans LELIEUR [2005].

On peut également retenir comme variante $\mathbf{lg}_x(t) = a_x \mathbf{lg}_x^{\text{ref}}(t) + b_x$ avec des coefficients dépendant de l'âge (ou de l'année). C'est un modèle de ce type qui a été utilisé pour construire les tables TGH et TGF 05 (cf. PLANCHET [2006]).

Enfin, on peut observer que lorsque $\ln(q_{xt}^{\text{ref}} / (1 - q_{xt}^{\text{ref}})) = \alpha_x + \beta_x \times k_t$, c'est à dire si la structure de mortalité sous-jacente est décrite par un modèle de type Lee-Carter, alors :

$$\ln(q_{xt} / (1 - q_{xt})) = a \times (\alpha_x + \beta_x \times k_t) + b = (a \times \alpha_x + b) + a \times \beta_x \times k_t$$

et donc le modèle ajusté est également de type Lee-Carter avec la même tendance temporelle. Seul le coefficient de sensibilité β_x est transformé en $a \times \beta_x$. On effectue donc par ce biais un positionnement en niveau de la mortalité d'expérience, la tendance de la référence étant réutilisée directement.

4.8.2. Positionnement par rapport à une référence externe

On peut également rechercher, dans un ensemble de tables prospectives exogènes disponibles la période des tables de référence $[t, t + h]$ la plus « proche » de la période $[t^{\text{ex}}, t^{\text{ex}} + h]$ issue des données d'expériences. Cela conduit à utiliser comme tables d'expérience les tables exogènes décalées.

La notion de « la plus proche » suppose l'utilisation d'une distance entre deux tables. Différentes approches sont possibles à ce niveau : Khi-2 sur les $q_x(t)$, distance déduite des espérances résiduelles ou de leurs intégrales (qui représente à une unité monétaire près l'engagement d'un portefeuille de rentes où tous les âges sont équi-représentés et à taux technique nul) l'avantage de cet indicateur est le « gommage » des fluctuations.

Ces modèles ne seront pas développés ici.

5. Les critères de validation du modèle

Les critères de validation de modèle fournissent des aides à la décision dans le cadre de la sélection du modèle le plus pertinent. La pertinence est ici appréciée en regard du contexte d'utilisation des tables proposées : souvent l'évaluation des engagements au titre de rentes viagères pour des tables prospectives.

Cela conduit notamment à porter une attention particulière à la représentation des espérances de vie résiduelles.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t; \infty[}(T_x)$$

ressources-actuarielles.net

5.1. La fidélité aux données

La première des exigences que doit satisfaire un modèle est d'être fidèle aux données qui ont servi à le calibrer. Cette fidélité peut être examinée *a priori* de deux manières :

- Au travers des taux de mortalité $q_x(t)$, $(x, t) \in [x_0, x_1] \times [a_0, a_1]$;
- Au travers de l'espérance de survie résiduelle dans la plage $[x_0, x_1]$, définie par :

$$e_t(x, x_1) = E[\mathbf{min}(X - x, x_1 - x) | X > x]$$

Le second critère est motivé par le fait que l'utilisation des tables prospectives est principalement orientée vers les calculs des engagements des rentes viagères. Les espérances de survie résiduelles représentent les engagements associés au calcul des rentes avec un taux d'actualisation nul. L'audit des espérances conditionnelles est donc incontournable.

Les modèles les moins paramétrés sont en principe et en général les plus fidèles. Néanmoins cette logique statistique n'est pas toujours respectée le calibrage se faisant sur $\mathbf{lg}_x(t)$ ou sur $\mathbf{ln}(\mu_{xt})$ et non sur les éléments retenus pour apprécier la fidélité du modèle (taux de mortalité $q_x(t)$ et espérance résiduelle $e_t(x, x_1)$). On peut toutefois, et avant la mise en œuvre proprement dite, faire les remarques suivantes :

- Le modèle Lee-Carter peut conduire à sous-évaluer notablement les taux de mortalité des âges élevés (à partir de 85-90 ans). En effet, l'algorithme de référence construit sur une approche maximum de vraisemblance favorise les premiers âges (les plus « jeunes) et par ailleurs la relation $\mu_{xt} = -\mathbf{ln}(1 - q_{xt})$ repose que l'hypothèse de constance du taux instantané de décès entre deux âges entiers, hypothèse discutable aux âges élevés.
- On peut s'attendre à ce que les modèles les moins paramétrés épousent mieux les « irrégularités résiduelles » des données brutes, ce qui constitue un handicap d'autant plus important que le volume de données est restreint.
- Au niveau des espérances résiduelles les irrégularités des tables brutes et des tables ajustées sont classiquement écrasées et ne ressortent que les dérives éventuelles et systématiques des modèles sur les q_x .

On peut noter à ce stade que pour éviter d'avoir à utiliser la « formule de passage » $\mu_{xt} = -\mathbf{ln}(1 - q_{xt})$ dans le modèle de Lee-Carter, il est possible de modéliser directement $\mathbf{lg}_x(t)$ plutôt que $\mathbf{ln}(\mu_{xt})$ en écrivant :

$$\mathbf{ln}\left(\frac{q_{xt}}{1 - q_{xt}}\right) = \alpha_x + \beta_x k_t + \varepsilon_{xt}.$$

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t; \infty} [(T_x)]$$

5.2. La comparaison des valeurs modélisées et des observations

La comparaison entre le nombre de sorties observé et le nombre de sorties estimé par le modèle sur un intervalle $[x, x+1[$ ou un pavé $[x, x+1[\times [t, t+1[$ est un outil central de la validation d'une loi d'expérience et de l'appréciation de son niveau de prudence.

Le décompte des sorties observé ne pose pas de difficulté et implique simplement de définir une convention d'arrondi. La partie entière de l'instant de sortie est le plus souvent utilisée.

Les sorties modélisées sont le plus souvent approchées par l'expression $E(D) \approx q(x) \times E(x)$, avec $E(x) = \sum_{i=1}^n [t_i \wedge (x+1) - e_i \vee x]^+$.

5.3. La stabilité des estimations

Le choix de la plage d'âges et de la plage d'années à partir desquelles on doit générer les prévisions est important dans la mesure où les estimations des paramètres dépendent sensiblement de ce choix. En effet il est possible que ces différences, si elles existent, engendrent des prévisions différentes.

En ce qui concerne les estimations des âges le choix de la plage d'âges ne doit pas avoir d'incidence notable sur les estimations. Ainsi si l'on retient par exemple la plage [50 ans-70 ans] et la plage [60 ans-80 ans] les âges communs (de 60 à 70 ans) doivent avoir des estimations voisines. Cela n'est pas le cas dans les modèles de Lee-Carter et log-Poisson :

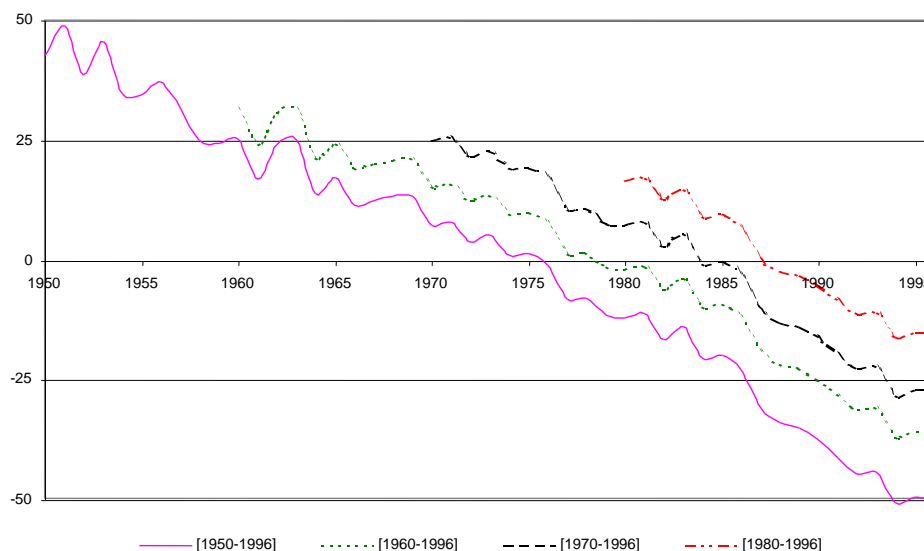


Fig. 21 : *Tendances estimées dans Lee-Carter en fonction de la période d'observation*

Pour les modèles où les estimations se font âge par âge la nature même des modèles assure l'égalité des estimations quand on fait varier la plage utilisée. Pour les autres modèles qui tiennent compte conjointement de tous les âges constituant une plage on peut craindre que les estimations des paramètres dépendent (plus ou moins fortement) de la plage d'étude choisie. Toutefois, on peut espérer que cette instabilité ne perturbe

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t;\infty} [(T_x)]$$

ressources-actuarielles.net

pas les estimations des logits, taux de décès et espérances résiduelles compte tenu des qualités de fidélité des modèles.

Pour le modèle de Lee-Carter (le moins paramétré en âges) la contrainte d'identification $\sum \beta_x = 1$ a pour conséquence mécanique des différences entre les estimations limitées cependant à des translations.

5.4. La capacité prospective

On peut noter que, d'une manière générale, la capacité d'un modèle à une utilisation prospective est d'autant plus importante que le modèle est fortement paramétrique. Cette remarque conduit à privilégier les approches paramétriques.

6. Le décès comme premier instant d'atteinte d'un seuil par un processus

Une approche alternative est parfois utilisée pour modéliser la survie ; cette approche est utilisée en général pour des populations non humaines (insectes notamment). Elle consiste à modéliser un « processus vital » par un processus de diffusion, et à considérer que le décès survient lorsque le niveau du processus vital diminue trop et franchit un certain seuil ; à un changement d'échelle près, on peut toujours supposer que ce seuil est zéro²⁷.

6.1. Présentation du modèle²⁸

Le modèle le plus simple que l'on puisse imaginer est alors :

$$dS_i(t) = -\mu dt + \sigma dW_i(t)$$

où l'indice i se rapporte à la tête considérée, et les W_i sont des mouvements browniens indépendants. Étant donné le niveau de la « viabilité » initiale, $S_0 = x > 0$, la probabilité de mourir entre t et $t + dt$ est égale à la probabilité que le brownien avec dérive ci-dessus atteigne l'origine pour la première fois à l'instant t , soit :

$$p_x(t) = \frac{x}{\sqrt{2\pi\sigma^2 t^3}} \exp\left(-\frac{(x-\mu t)^2}{2\sigma^2 t}\right)$$

Pour une distribution initiale des viabilités de densité ω_0 , on obtient que le nombre de décès à la date t est $D(t) = \int_0^{+\infty} p_x(t) \omega_0(x) dx$, ce qui conduit à l'expression suivante du taux de décès instantané :

²⁷ On peut imaginer la modélisation duale d'un processus de « morbidité » qui, lorsqu'il dépasse un certain seuil, déclenche le décès.

²⁸ Voir FRASER et WEITZ [2003] pour le détail de l'approche et une application numérique.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t;\infty}[(T_x)]$$

ressources-actuarielles.net

$$\mu(t) = \frac{D(t)}{N_0 - \int_0^t D(s) ds}$$

On peut montrer que dans le cas d'une population initialement homogène ($x = 1$ pour tous les individus de la population), alors le taux de mortalité instantané admet une expression analytique.

6.2. Estimation des paramètres

L'estimation par la méthode du maximum de vraisemblance ne pose pas de difficulté particulière et conduit ici à :

$$\hat{\mu} = \left(\frac{1}{N} \sum_{i=1}^N T_i \right)^{-1} = \frac{1}{\bar{T}}$$

$$\hat{\sigma} = \left(\frac{1}{N} \sum_{i=1}^N \left(\frac{1}{T_i} - \frac{1}{\bar{T}} \right) \right)^{1/2}$$

Ce type d'approche permet d'exploiter les nombreux résultats existants sur les temps d'atteinte d'un seuil par un mouvement brownien avec dérive.

7. Références

- BRASS W. [1971] « *On the Scale of Mortality* », Biological aspects of demography, W. Brass (ed.), Londres, Taylor and Francis, 69-110.
- BONGAARTS J. [2004] « *Long-Range Trends in Adult Mortality : Models and Projection Methods* », Population Council, n°194.
- BROUHNS N., DENUIT M. [2002], « *Risque de longévité et rentes viagères : II. Tables de mortalité prospectives pour la population belge* », BAB.
- BROUHNS N., DENUIT M. [2002], « *Risque de longévité et rentes viagères : I. Evolution de la mortalité en Belgique de 1880 à nos jours* », BAB.
- BROUHNS N., DENUIT M., VERMUNT J.K. [2002], [A Poisson log-bilinear regression approach to the construction of projected lifetables](#), Insurance, Mathematic and Economics, vol. 31.
- COALE A., KISKER E.E. [1990]. « *Defects in data on old age mortality in the united states : New procedures for calculating approximatively accurate mortality schedules and life tables at the highest ages* ». Asian and Pacific Population Forum 4,1-31.
- CZADO C., DELWARDE A., DENUIT M. [2004] « [Bayesian Poisson Log-bilinear Mortality Projections](#) », Working Paper.
- DEBON A., MARTINEZ-RUIZ F., MONTES F. [2004], « *Dynamic Life Tables: A Geostatistical Approach* », IME Congress

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t;\infty}[(T_x)]$$

ressources-actuarielles.net

Modèles de durée

- DENUIT M., QUASHIE A. [2005], « *Modèles d'extrapolation de la mortalité aux grands âges* ». Institut des Sciences Actuarielles et Institut de Statistique Université Catholique de Louvain, Louvain-la-Neuve, Belgique.
- DENUIT M., MAGIS C., WALHIN F. [2004] « *La mortalité, un phénomène en pleine mutation: quelle solution pour le marché des rentes ?* », Working Paper.
- FRASER H.B., WEITZ J. S. [2003] « *Explaining Mortality Rate Plateaus* », Proc. Nat. Acad. Sci.
- GUIBERT Q., PLANCHET F. [2017] « [Utilisation des estimateurs de Kaplan-Meier par génération et de Hoem pour la construction de tables de mortalité prospectives](#) », Document de travail, ISFA.
- HADERER M. [2003] « *Réassurance du risque de longévité* », ISFA, Mémoire d'actuariat.
- LELIEUR V. [2005] « *Construction de tables de mortalité prospectives: le cas des petites populations* », ISFA, Mémoire d'actuariat.
- PLANCHET F. [2006] « *Tables de mortalité d'expérience pour des portefeuilles de rentes – note de présentation* », Publication de l'Institut des Actuaire.
- RENSHAW A.E. [1991] « *Actuarial graduation practice and generalized linear and non-linear models* », Journal of the Institute of Actuaries, vol. 118, pp. 295–312.
- SITHOLE T., HABERMAN S., VERRALL R.J. [2000] « *An investigation into parametric models for mortality projections, with applications to immediate annuitants and life office pensioners* », Insurance: Mathematics and Economics, vol. 27, pp. 285–312.
- THATCHER A.R. [1999] « [The Long-term Pattern of Adult Mortality and the Highest Attained Age](#) », Journal of the Royal Statistical Society 162 Part 1: 5-43.
- THATCHER A.R., KANNISTO V., VAUPEL J.W. [1998]. *The Force of Mortality at Ages 80 to 120*. Odense: Odense University Press.

$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \mathbf{1}_{t;\infty} (T_x)$$

8. Annexe : comparaison des modèles de Lee-Carter et de Poisson

