



Mémoire présenté devant le jury de l'EURIA en vue de l'obtention du
Diplôme d'Actuaire EURIA
et de l'admission à l'Institut des Actuaires

le 19 Septembre 2018

Par : Coralie CAPPE

Titre : Établissement d'une table de mortalité spécifique pour un régime de retraite existant
et évaluation des impacts financiers

Confidentialité : Non

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

**Membres présents du jury de l'Institut
des Actuaires :**

Aliou SOW

Signature :

Marine HABART

Signature :

Stéphane JASSON

Signature :

Entreprise :

SPAC Actuaires

Signature :

Directeur de mémoire en entreprise :

Olivier SEGUIN

Signature :

Membre présent du jury de l'EURIA : **Invité :**

Daniel BOIVIN

Signature :

Signature :

**Autorisation de publication et de mise en ligne sur un site de diffusion
de documents actuariels**

(après expiration de l'éventuel délai de confidentialité)

Signature du responsable entreprise :

Signature du candidat :

Secrétariat :

Bibliothèque :

Résumé

Mots clés : régime de retraite, ouvrant droit, ayant droit, mortalité, table de mortalité, ajustement table de référence, taux de mortalité, taux d'abattement, lissage, décalage d'âges, ELT, espérance de vie, flux de pension, droits, impacts financiers

L'augmentation régulière de l'espérance de vie de la population, accompagnée du fait que les *baby-boomers* liquident progressivement leur retraite, constitue un risque de longévité de plus en plus important et menaçant pour les régimes des retraites sur le plan financier. Ces régimes favorisent l'utilisation de tables de mortalité propres à leur portefeuille d'assurés plutôt que les tables réglementaires afin d'éviter d'éventuelles pertes financières. En effet, une simple sur- ou sous-estimation de la mortalité des individus peut vite avoir de forts impacts sur les engagements ou les plans de financement.

Ce mémoire propose alors d'établir une loi de mortalité à partir des observations de la population des affiliés du régime de retraite étudié. Cette loi est établie à l'aide d'un ajustement effectué sur la table de référence INSEE 2013-2070 à partir des observations des années de 2013 à 2017. Elle est ensuite comparée à la loi utilisée jusqu'à présent par le régime et qui relève de la mortalité retournée par la table de référence INSEE 2007-2060 ajustée de coefficients d'abattement retenus lors d'une étude établie en 2012.

L'étude se décompose en plusieurs parties avec notamment une première étape consacrée à l'analyse de statistiques descriptives et démographiques sur la population qui compose le régime en question. Des sous-populations se distinguent et l'étude est appliquée à chacune d'entre elles. Ces analyses révèlent également que les individus sont parfois peu nombreux sur certaines tranches d'âges, ce qui a une incidence sur les taux bruts de mortalité qui reflètent alors quelques fluctuations nécessitant un lissage des taux observés. Différentes méthodes sont testées de sorte à retenir celle qui tient compte au mieux de la fidélité et de la régularité à nos données.

L'étape suivante de ce travail est la mise en application de deux méthodes principales de construction de table de mortalité par ajustement d'une table de référence : d'un côté une méthode par décalage d'âges, et d'un autre côté une méthode par abattement de la table de référence par tranches d'âges. L'ensemble des calculs est effectué à partir des taux de mortalité moyens calculés sur la période 2013-2017. Ces méthodes sont ensuite comparées à des méthodes de positionnement proposées par un *package* propre à la construction de tables prospectives disponible sous *RStudio*. De nombreux tests et calculs d'erreurs d'ajustement sont effectués dans l'optique de rechercher le meilleur modèle adapté à nos données brutes. Ce modèle est ensuite comparé avec celui actuellement utilisé par le régime de sorte à ne retenir que celui qui concorde le mieux avec la population de nos affiliés, et donc celui qui est plus apte à évaluer le risque viager auquel le régime est soumis.

Enfin, un changement d'utilisation de table de mortalité peut avoir des répercussions financières, c'est pourquoi après avoir choisi la méthode d'ajustement la plus pertinente, projeté les taux ajustés et fermé la table établie, ce mémoire se termine par une évaluation des impacts financiers engendrés par l'utilisation de la nouvelle loi de mortalité sur le calcul des montants de pensions.

Abstract

Keys words : *pension plan, right owner, beneficiary, mortality, mortality table, reference table's adjustment, mortality rate, abatement rate, smooth, age difference, ELT, life expectancy, pension flows, rights, financial impact*

In France, there are two main trends that are financially threatening pension plans. Firstly, the steady increase in life expectancy, and secondly, baby-boomers are now drawing their pension. Both trends constitute an increasingly important longevity risk for pension plans. These pension plans thus promote mortality tables following their own populations' mortality, in order to avoid potential financial losses. These losses can become important if there is a significant difference between the pension liability and its calculations, due to an over/under-estimation of mortality.

This memoir conducts a study that establishes a law of mortality, based on the pension plan's population. This law is built through an adjustment from 2013-2070 INSEE prospective mortality tables, themselves being based on observations from 2013 to 2017. Then, this table is compared to the current law used by the pension plan. This law follows 2007-2060 INSEE prospective mortality tables, adjusted with abatement rates adopted in a previous study.

The study is structured into several parts with a first section analysing descriptive and demographic statistics on the pension scheme's population. Some sub-populations distinguish themselves and thus each one of them is studied separately. This analysis also highlights the fact that certain age sections only count a small number of people, consequently impacting the crude mortality rates. It reflects some fluctuations that need a rate smoothing. Different smoothing methods are therefore tested in order to determine the one that best respects accuracy and regularity to the data.

The next section provides two methods for constructing mortality tables based on the adjustment of a reference table. On the one hand, a method shifting all ages from the reference table, and on the other hand, a method enforcing abatements on the reference table, and this by age sections. Basically, average rates on years from 2013 to 2017 are used for all calculations. These applications are then compared to other methods offered in a computer package for RStudio, dedicated to constructing experience life tables. Many tests and error of adjustment calculations are carried out in order to find which model is the most suitable to the crude data. This model is finally compared with the current one - used by the pension plan - to obtain the best evaluation of the life risks of the studied population.

Eventually, a change of use for mortality tables can have financial impacts. Therefore, after choosing the most relevant adjustment method, projecting the adjusted rates, and closing the built table, this memoir ends with a valuation of financial impacts generated by using the new law of mortality on pension amounts.

Synthèse

Les différences de sexe, de catégorie socio-professionnelle, ou encore de situation familiale sont à l'origine d'écart non négligeables entre les espérances de vie. Ces disparités sont accompagnées d'une augmentation régulière de l'espérance de vie conjuguée à des liquidations de plus en plus importantes ces dernières années (*baby-boomers*). Cela entraîne un impact sur la durée de vie passée en retraite, et a des répercussions sur le montant total des charges de pensions à verser par les régimes.

Ainsi, pour mesurer ce risque de longévité et éviter d'éventuelles pertes financières, les régimes de retraite favorisent l'utilisation de tables de mortalité propres à leur portefeuille d'assurés plutôt que des tables réglementaires. En effet, l'utilisation d'une table mal adaptée à la population engendre une sur- ou sous-estimation de la mortalité des individus qui peut rapidement avoir de fortes conséquences au niveau des engagements.

Objectif : Ce mémoire a pour but principal d'établir une table de mortalité propre à la population d'un régime de retraite spécial à partir de l'ajustement de la table de référence INSEE prospective 2013-2070. Cette méthode est adaptée lorsque l'on souhaite établir une table de mortalité générationnelle alors que les effectifs disponibles pour l'observation ne sont pas assez fournies. Les enjeux de l'utilisation de cette nouvelle table sont ensuite chiffrés.

Démarche mise en place : L'ensemble de l'étude s'est effectué au travers d'une démarche que nous pouvons représenter de la manière suivante :



FIGURE 1 – Démarche mise en application

Lors de l'analyse des statistiques descriptives, quatre sous-populations se distinguent selon le sexe et le type de prestation des individus : les femmes ouvrants droit, les hommes ouvrants droit, les femmes ayants droit, ainsi que les hommes ayants droit. Par conséquent, les étapes qui ont suivi ont été effectuées pour chacune d'entre elles.

Étude de la population

La période d'observation de la population du régime considéré s'étend sur les cinq dernières années, c'est-à-dire sur les années de 2013 à 2017. Pour une étude de construction de table de mortalité, l'analyse de statistiques descriptives et démographiques conduisent au calcul des taux de mortalité bruts moyens sur la période d'observation.

Statistiques descriptives :

Les variables relatives au sexe, à l'âge (date de naissance/génération), au type de prestation (ouvrant droit, noté "OD", ou ayant droit, noté "AD"), au caractère décédé ou non, à la date d'entrée au régime ou encore à la date d'entrée en retraite d'un individu, sont des données essentielles à une étude démographique pour un régime de retraite. Ce sont alors sur ces variables que nos statistiques reposent.

L'effectif sous risque observé, ainsi que le nombre de décès observés constituent les éléments principaux d'une construction de table de mortalité. Le schéma suivant présente les résultats qui concernent la population étudiée, exprimés en proportion de population d'effectif sous risque (en bleu) et de personnes décédées (en rouge) sur l'ensemble de la période d'observation.

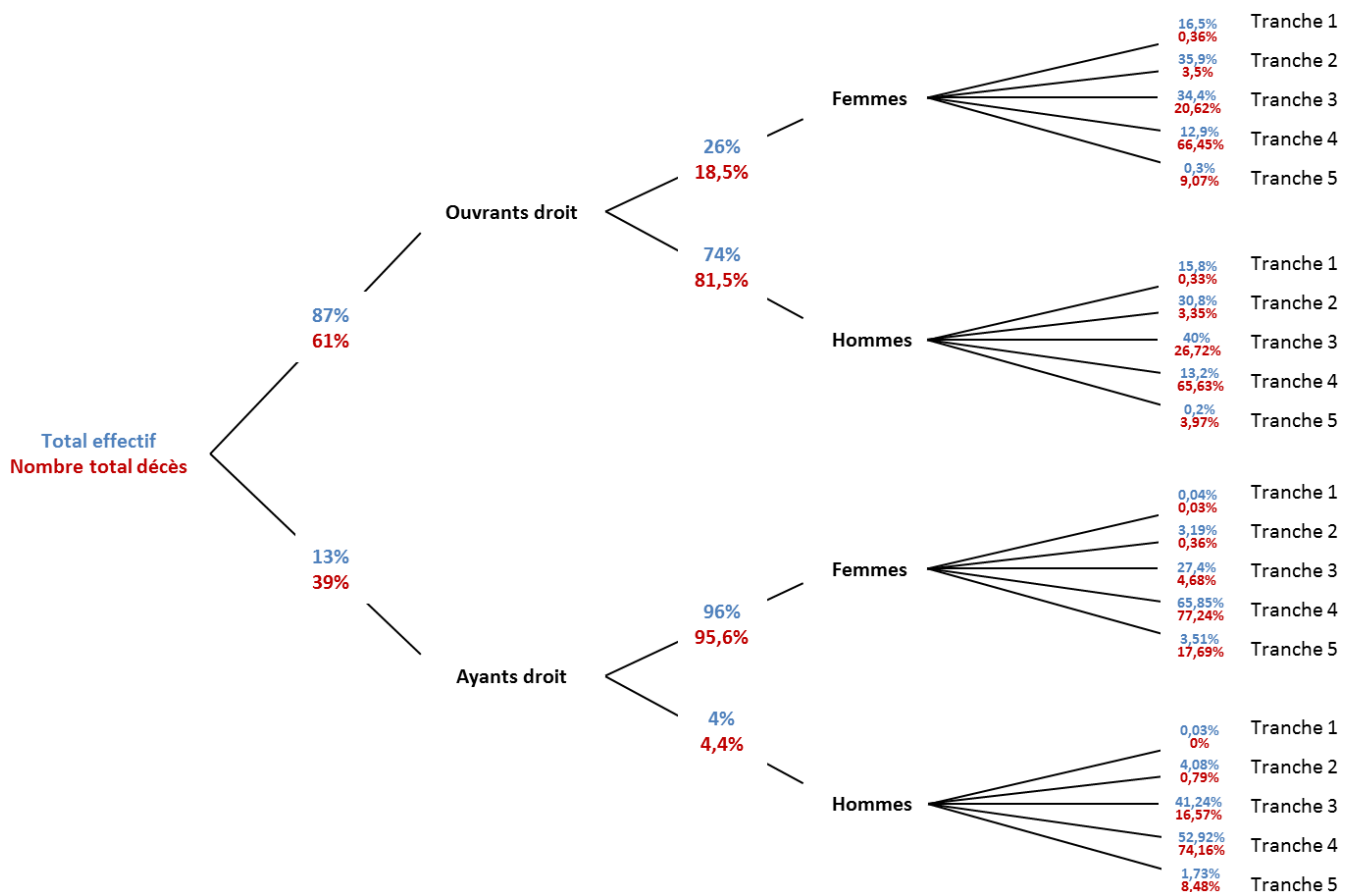


FIGURE 2 – Statistiques de populations

Les individus sont répartis selon leur type de prestation, leur sexe et leur âge par tranches d'âges (Tranche 1 : âges jusqu'à 35 ans, Tranche 2 : de 36 à 55 ans, Tranche 3 : de 56 à 75 ans, Tranche 4 : de 76 à 95 ans, Tranche 5 : au-delà de 95 ans).

Calcul des taux bruts de mortalité :

Les effectifs sous risque et nombres de décès à chaque âge permettent de calculer les taux bruts de mortalité moyens pour chaque âge sur la période d'observation. Nous rappelons que les quotients de

mortalité calculés sur la période de référence pour chaque cohorte correspond au rapport suivant :

$$\text{Pour une cohorte de génération } n - x - 1 \text{ d'âge } x \text{ l'année } n : \quad q_x = \frac{\text{Décès enregistrés l'année } n}{\text{Effectif exposé au risque l'année } n}$$

Ces taux constituent la base de construction de la table de mortalité.

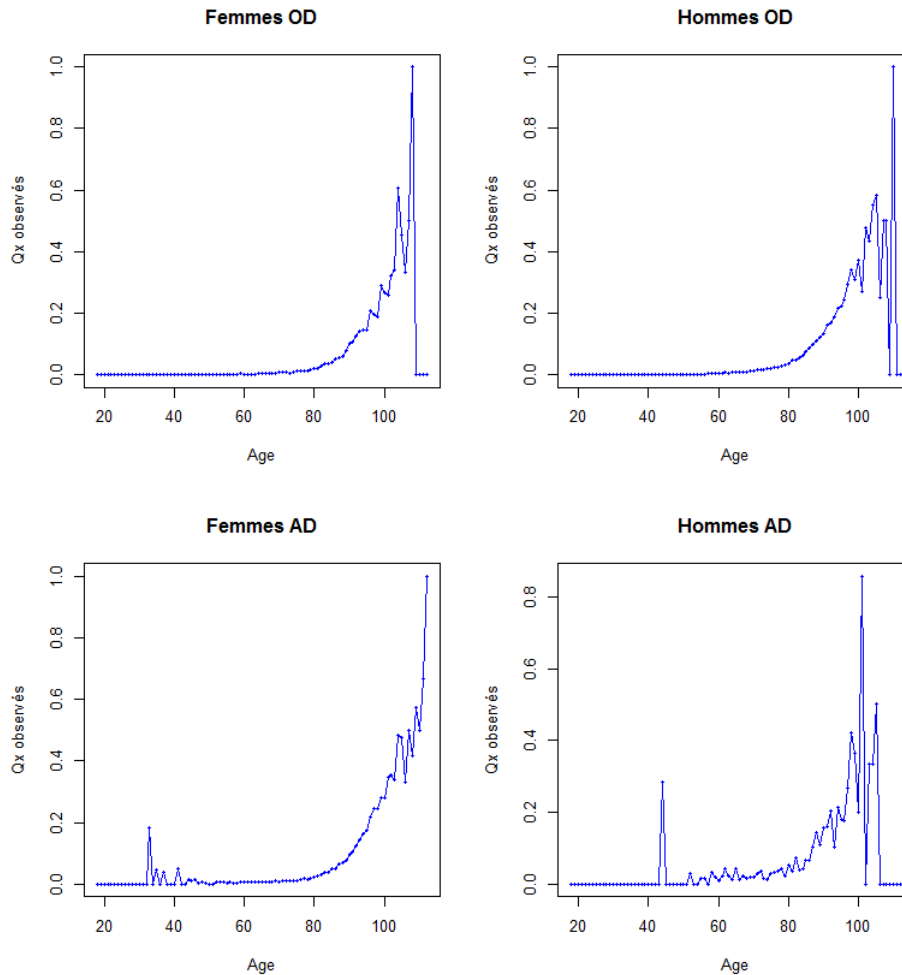


FIGURE 3 – Taux de mortalité observés pour chaque population

La croissance exponentielle naturelle de la mortalité avec l'âge est constatée pour nos quatre populations. Cependant, les statistiques et calculs réalisés révèlent que les observations relatives aux individus de plus de 95 ans sont très peu significatives du fait de leur représentation extrêmement faible au sein du régime. Par ailleurs, certains taux aberrants apparaissent aux bas âges chez les deux populations des ayants droit. Par conséquent, ces taux devront être lissés, et les âges compris au-delà de 95 ans ne seront pas considérés dans la suite de l'étude pour éviter les erreurs de modélisation.

Construction de la table de mortalité

Lissage des taux bruts :

Plusieurs méthodes de lissage sont effectuées de sorte à pallier les aberrations observées sur les taux bruts qui pourraient nuire à un ajustement précis et réel de la table.

Plusieurs méthodes de lissage ont été mises en place : la méthode d'ajustement par les moindres carrés avec maximisation du R^2 , la méthode de Makeham, la méthode de Loess, ainsi que la méthode de Whittaker-Henderson. Elles sont ensuite comparées entre elles graphiquement, puis par le calcul de trois

valeurs d'intérêt : le R^2 (à maximiser), le χ^2 (à minimiser), et la $MAPE$ (à minimiser). Ces trois valeurs s'expriment de la manière suivante :

$$R^2 = 1 - \frac{\sum_x (\hat{q}_x - q_x)^2}{\sum_x \left(\hat{q}_x - \frac{\sum_x \hat{q}_x}{n} \right)^2} \quad \chi^2 = \sum_x E_x \frac{(\hat{q}_x - q_x)^2}{q_x} \quad MAPE = \frac{\sum_x |(\hat{q}_x - q_x) / \hat{q}_x|}{n} \times 100$$

avec :

- \hat{q}_x : les taux de mortalité observés à chaque âge x
- q_x : les taux de mortalité lissés à chaque âge x
- n : le nombre de taux à lisser
- E_x : l'exposition à l'âge x

La méthode de lissage finalement retenue est celle de Whittaker-Henderson, qui repose sur un principe de combinaison d'un critère de fidélité avec un critère de régularité, l'objectif étant de rechercher les valeurs qui minimisent une combinaison linéaire de ces deux critères. Le poids attribué à chacun de ces deux critères est déterminé par un paramètre représentant l'importance souhaitée être donnée à la régularité.

Ajustement de la table de référence :

- **Méthode par décalage d'âges** : il s'agit d'une méthode qui consiste à appliquer un vieillissement ou un rajeunissement sur la table de référence pour l'adapter à la mortalité de la population d'étude. Le décalage est recherché optimal, et pour cela, des décalages sont appliqués sur différentes tranches d'âges.

Il est alors recherché $i \in \mathbb{Z}$ tel que la quantité suivante soit minimisée pour chaque âge x :

$$(Taux_{lissé_x} - Taux_{ref_{x+i}})^2$$

Dans cette étude, les décalages sont des rajeunissements qui vont de 1 à 9 ans pour les femmes OD, de 1 à 18 ans pour les hommes OD, et de 0 à 3 ans pour les AD.

- **Méthode par application de coefficients d'abattement** : il s'agit d'une méthode qui consiste à appliquer un coefficient d'abattement sur les taux de mortalité de la table de référence pour l'adapter à la mortalité de la population d'étude. Le coefficient est recherché optimal, et pour cela, des décalages sont appliqués sur différentes tranches d'âges.

Il est alors recherché $i \in [0; 1]$ tel que la quantité suivante soit minimisée pour chaque âge x :

$$(Taux_{lissé_x} - Taux_{ref_x} \times (1 - i))^2$$

Dans cette étude, les coefficients d'abattement vont de 15% à 60% pour les femmes OD, de 12% à 64% pour les hommes OD, et de 0% à 19% pour les AD.

Pour ces deux méthodes, rassembler les individus par tranches d'âges permet de limiter l'importance donnée aux âges pour lesquels le nombre de personnes est trop faible pour obtenir un bon ajustement.

- **Méthodes de positionnement** : il s'agit de méthodes proposées par un *package* informatique appelé *ELT* et élaboré par Frédéric PLANCHET et Julien TOMAS. Trois méthodes sont étudiées ici.

- Une approche à un paramètre avec le SMR (*Standardized Mortality Ratio*) :

$$\tilde{q}_{x,t} = SMR \times q_{x,t}^{ref} \quad \text{avec} \quad SMR = \frac{\sum_{(x^*, t^*)} d_{x^*, t^*}}{\sum_{(x^*, t^*)} E_{x^*, t^*} \times q_{x^*, t^*}^{ref}}$$

- Une approche à deux paramètres avec un modèle relationnel :

$$\forall x \in x^*, t \in t^* : \tilde{q}_{x,t} = \frac{\exp(\hat{\alpha} + \hat{\beta} \times \lg(q_{x,t}^{ref}))}{1 + \exp(\hat{\alpha} + \hat{\beta} \times \lg(q_{x,t}^{ref}))} \quad \text{avec} \quad \lg(\tilde{q}_{x^*, t^*}) = \alpha + \beta \lg(q_{x^*, t^*}^{ref}) + \varepsilon_{x^*, t^*}$$

— Une approche par modèle linéaire généralisé de Poisson incluant des interactions avec l'âge et l'année calendaire :

$$\tilde{q}_{x,t} = \exp\left(\hat{\beta}_0 + \hat{\beta}_1 \log q_{x,t}^{ref} + \hat{\beta}_2 x + \hat{\beta}_3 t + \hat{\beta}_4 x t\right)$$

avec $D_{x^*,t^*} \sim \mathcal{P}(E_{x^*,t^*} \mu_{x^*,t^*})$ et $\mu_{x^*,t^*} = \beta_0 + \beta_1 \log q_{x^*,t^*}^{ref} + \beta_2 x^* + \beta_3 t^* + \beta_4 x^* t^*$

L'objectif de ces trois méthodes est d'estimer les paramètres des modèles sur la période d'observation pour ensuite pouvoir estimer les taux de mortalité ajustés sur cette même période.

Choix de la méthode d'ajustement :

Les moyens de comparaison de toutes ces méthodes sont les mêmes valeurs d'intérêt que calculées à l'étape précédente (R^2 , χ^2 , et $MAPE$).

La méthode de construction de table retenue est la méthode d'ajustement par abattement de la table de référence. Les résultats graphiques sont présentés ci-après :

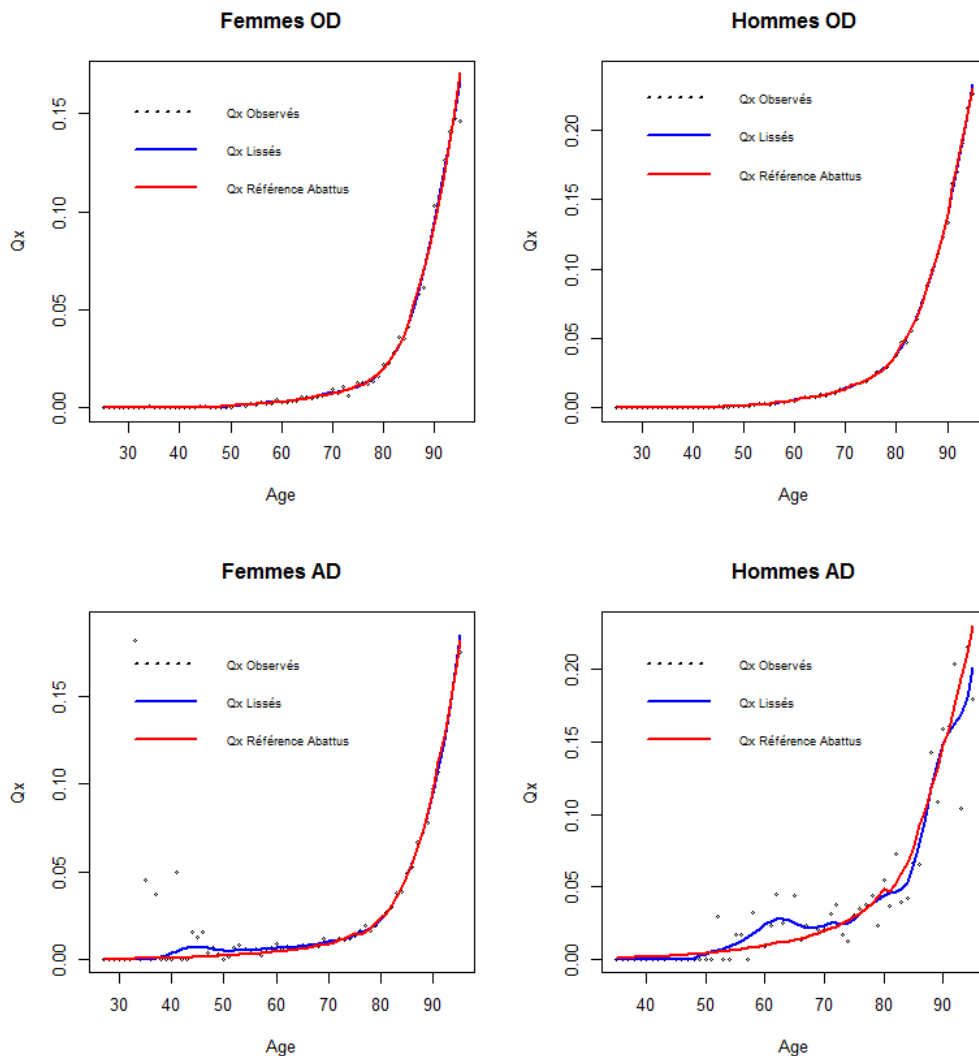


FIGURE 4 – Taux de mortalité observés, lissés, et abattus pour chaque population

Pour les OD, les taux bruts, les taux lissés et les taux abattus sont alignés. En ce qui concerne les AD, de légers écarts sont constatés. Ces écarts sont dus à la faible présence d'individus sur certaines tranches d'âges qui peuvent venir perturber la régularité de la mortalité de la population. L'ajustement réalisé est donc globalement considéré comme un bon ajustement de la mortalité de ces quatre populations.

La comparaison par les valeurs d'intérêt de cette nouvelle loi avec celle actuellement utilisée par le régime prouve que la table créée est bien plus adaptée à la mortalité des populations du régime.

Fermeture de la table :

La méthode utilisée pour la fermeture de la table est la méthode de Denuit et Goderniaux qui consiste en l'ajustement d'un modèle log-quadratique en imposant deux contraintes aux grands âges.

La première contrainte stipule une probabilité de décès de 100% à 120 ans ainsi que la présence d'une tangente horizontale à cet âge. Par ailleurs, la seconde contrainte traduit une concavité aux âges élevés qui empêche une décroissance des taux de mortalité à ces âges.

Évaluation des impacts financiers

Le deuxième objectif de cette étude est d'évaluer les impacts financiers causés par le changement d'utilisation de table de mortalité. Cette évaluation est réalisée sur le montant des flux de pensions probables en droits acquis projetés du régime en groupe fermé.

Deux scénarios sont étudiés ici :

- Un premier scénario où la loi de mortalité utilisée pour les calculs d'engagements est identique pour les ouvriers droit et les ayants droit (loi des ouvriers droit également appliquée aux ayants droit).
- Un second scénario où la loi de mortalité utilisée pour les calculs d'engagements est différente pour les ouvriers droit et les ayants droit.

Ces deux cas sont présentés de manière à laisser le choix aux instances dirigeantes du régime, de celui qu'elles préféreront conserver.

Trois populations sont étudiées : les actifs seuls, les inactifs seuls, ainsi que les actifs et les inactifs rassemblés. Ici, les résultats sont présentés pour la population globale du régime (c'est-à-dire actifs et inactifs rassemblés).

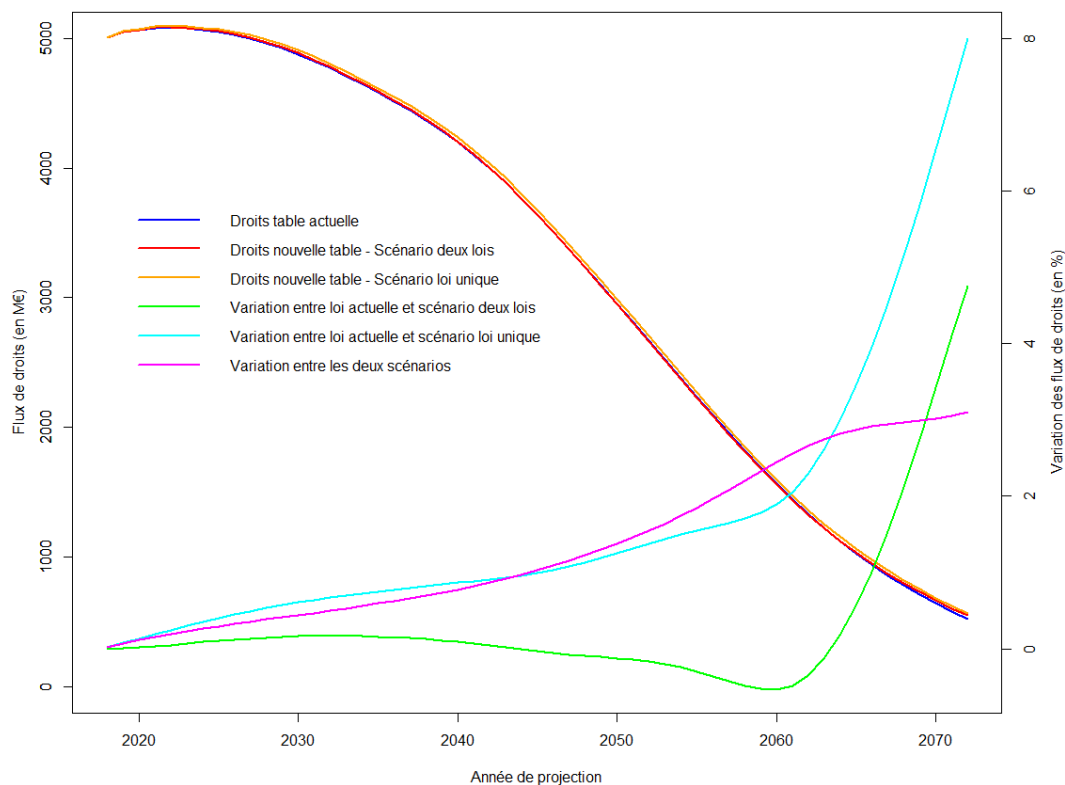


FIGURE 5 – Flux de droits projetés et pourcentages de variation selon les tables considérées

Les comparaisons des projections de flux montrent que les montants sont très proches, qu'ils soient calculés avec la table actuelle ou avec la nouvelle loi, et ce, quel que soit le scénario considéré. De légers écarts à la hausse sont constatés à partir de 40 années de projection (lorsque la masse des pensions servies a fortement diminuée).

Les résultats montrent également que les flux de droits projetés calculés avec la loi unique sont légèrement supérieurs à ceux calculés avec deux lois différentes. C'est pourquoi ce scénario est généralement considéré comme plus prudent que le scénario utilisant deux lois de mortalité différentes selon la population considérée.

Conclusion

Cette étude conclut sur le fait que la nouvelle table construite à partir d'un ajustement par coefficients d'abattement de la table de référence INSEE prospective 2013-2070 est beaucoup plus adaptée à la mortalité de la population du régime étudiée que celle utilisée jusqu'à présent, mais que ce changement d'utilisation de table a peu d'impact sur les montants des droits.

La mise à jour régulière des tables de mortalité permet donc de limiter l'ampleur de l'impact financier causé par le changement d'utilisation de table. Toutefois, la taille de la population de ce régime est suffisante pour une étude de la mortalité, mais pas assez pour créer une table de mortalité générationnelle qui projette les espérances de vie sur le long terme.

Synthesis

Differences of gender, socio-professional category, or family status can create deviations in people's life expectancies. Furthermore, there is a steady increase in average life expectancy, and baby-boomers are now drawing their pension. That impacts the lifetime spent retiring, and in consequence the global amount due by pension plans.

Thus, in order to reduce the increasingly longevity risk, and to avoid potential technical losses, pension plans promote mortality tables following their own populations' mortality rather than reference tables. Indeed, using a maladjusted table can lead to a significant difference between the actual pension liabilities and their valuations, due to an over/under-estimation of mortality.

Goal : The aim of this memoir is the construction of a mortality table adapted to a pension plan's own population, thanks to an adjustment of the 2013-2070 INSEE prospective reference mortality table. Another goal of this study is to evaluate the financial impacts caused by a change of mortality tables used for calculation.

Approach used : The entire study has been carried out on an approach that can be represented by this following schematic :

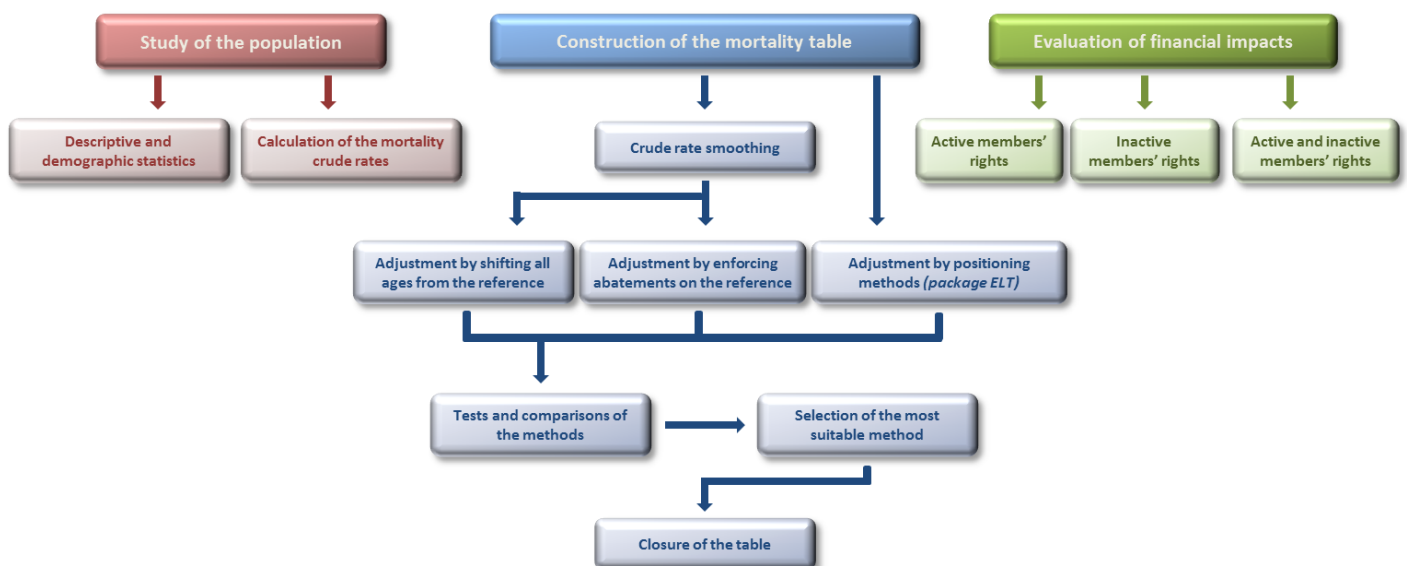


FIGURE 6 – Approach used

Throughout the descriptive statistics, four sub-populations distinguished themselves and thus each one of them has been studied separately : female right owners, male right owners, female beneficiaries, and male beneficiaries.

Study of the population

The pension scheme's population is studied on observations from the past five-year period, so from 2013 to 2017. In order to construct a mortality table, analysing descriptive and demographic statistics leads to the average crude mortality rates' calculation on the observation period.

Descriptive and demographic statistics :

Variables concerning sex, age (birthdate or generation), benefit type (right owner or beneficiary), deceased person or not, the date of entrance in the pension scheme or in retirement of people, are essential data for a pension scheme's demographic study. The following statistics are thus based on these variables.

The under risk staff and number of deaths constitute the main elements for a life table's construction. The following schematic presents results that concern the studied population. They are expressed as percentages of under risk staff (in blue), and deceased people (in red) on the whole period.

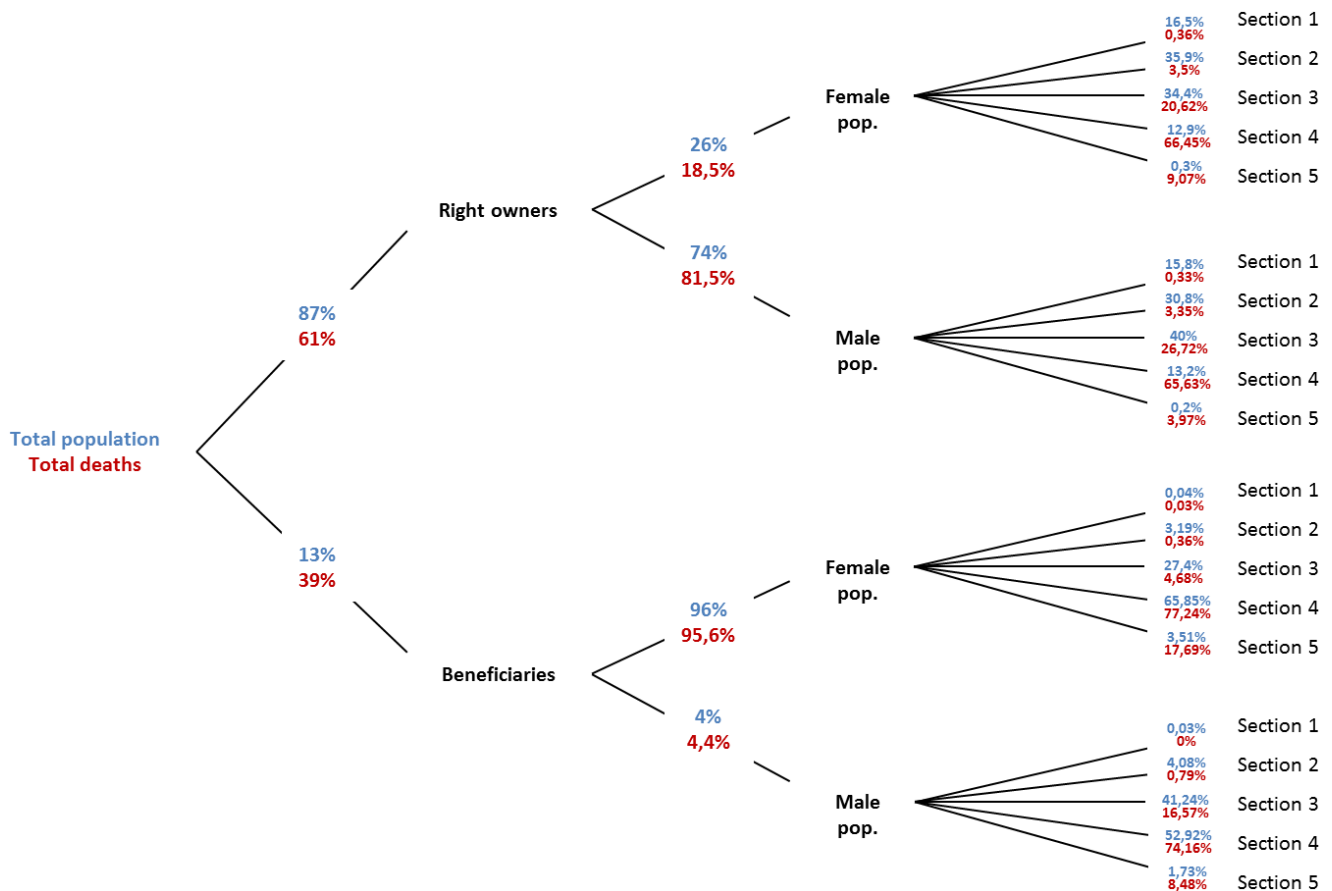


FIGURE 7 – Statistics of the populations

People are divided into several groups, depending on benefit type, sex, and age into age sections (Section 1 : ages under 35 and 35, Section 2 : ages from 36 to 55, Section 3 : ages from 56 to 75, Section 4 : ages from 76 to 95, Section 5 : ages over 95).

Calculation of the crude mortality rates :

The under risk staff and number of deaths for each age allow to calculate the average crude mortality rates for each age during the observation period. During the reference period, mortality rates calculated for each generation correspond to the following relation :

$$\text{For a generation } n - x - 1 \text{ aged } x \text{ during year } n : \quad q_x = \frac{\text{Number of registered deaths during year } n}{\text{Under risk staff during year } n}$$

The mortality table's construction is based on these rates.

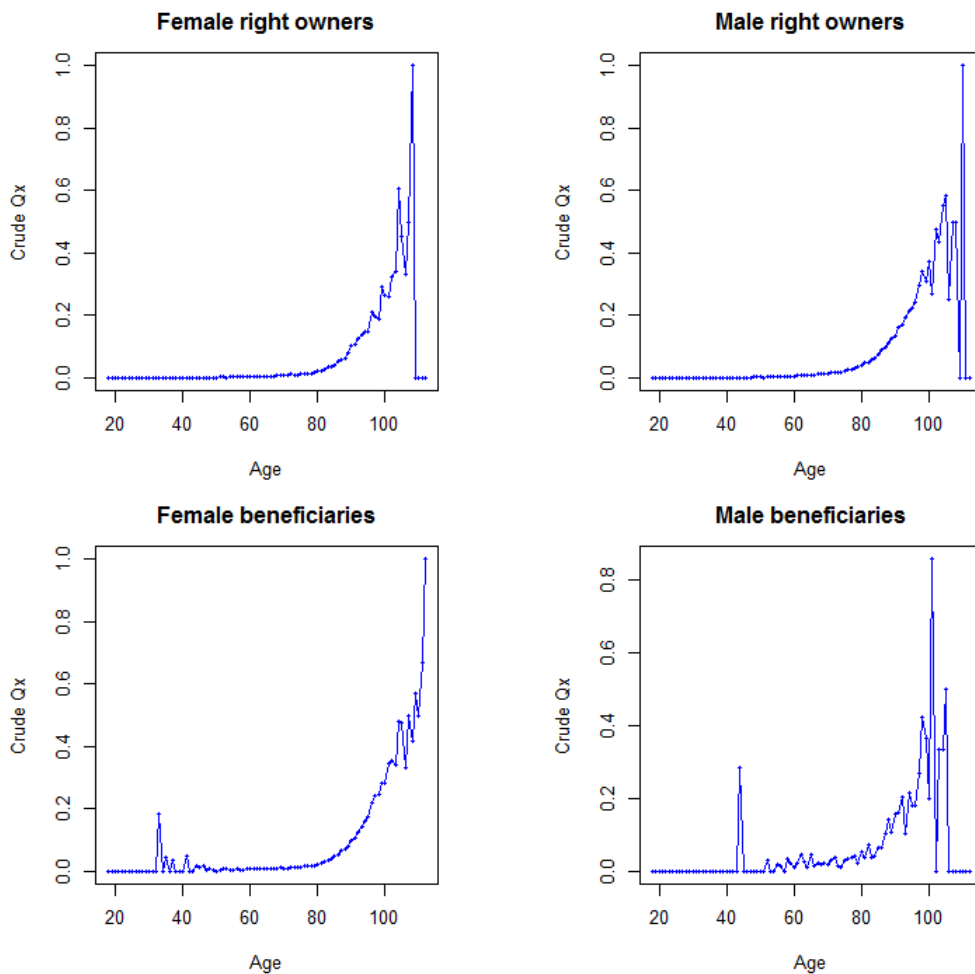


FIGURE 8 – Crude mortality rates of each population

The natural exponential increasing mortality with age is noticed for each one of the four sub-populations. However, statistics and calculations reflect that people over 95 are not significant because of an under-representation within the global population. Furthermore, some outrageous rates for young ages can be noticed amongst the beneficiaries'. Therefore, a crude rate smoothing is required, and ages under 95 are removed from the study in order to avoid modelling errors.

Construction of the mortality table

Crude rate smoothing :

In order to avoid outrageous rates that could impact the model, several smoothing methods are carried out.

Several smoothing methods have been studied : the least square approach with a maximisation of the coefficient of determination, the Makeham model, the Loess model, and the Whittaker-Henderson method. They are then graphically compared to each other, and with statistical measurements : the R^2 (to maximize), the χ^2 (to minimize), and the Mean Absolute Percentage Error (*MAPE*, to minimize too). These quantities are expressed like following :

$$R^2 = 1 - \frac{\sum_x (\hat{q}_x - q_x)^2}{\sum_x \left(\hat{q}_x - \frac{\sum_x \hat{q}_x}{n} \right)^2} \quad \chi^2 = \sum_x E_x \frac{(\hat{q}_x - q_x)^2}{q_x} \quad MAPE = \frac{\sum_x |(\hat{q}_x - q_x) / \hat{q}_x|}{n} \times 100$$

with :

- \hat{q}_x : the crude mortality rates for each age x
- q_x : the smoothed mortality rates for each age x
- n : the number of rates to smooth
- E_x : the exposure at x

The smoothing method selected is Whittaker-Henderson's, based on a combination between accuracy and regularity. The aim is to find values minimizing a linear combination of both. A parameter representing the importance given to regularity determines the weight assigned to each criterion.

Reference table's adjustment :

- **Adjustment by shifting ages** : the method consists in enforcing a rejuvenation or an ageing population on the reference table to adapt it to the studied population's mortality. The shift has to be the best, so shifting are applied on different age sections.

Thus, the researched shift is $i \in \mathbb{Z}$ such as the following quantity is minimized for each age x :

$$(Rate_{smooth_x} - Rate_{ref_{x+i}})^2$$

In this study, the shifts are rejuvenations from 1 to 9 years for female right owners, from 1 to 18 years for male right owners, and from 0 to 3 years for beneficiaries.

- **Adjustment by enforcing abatements** : the method consists in enforcing a coefficient of abatement on the reference table to adapt it to the studied population's mortality. The abatement has to be the best, so coefficients are applied on different age sections.

Thus, the researched coefficient is $i \in \mathbb{Z}$ such as the following quantity is minimized for each age x :

$$(Taux_{lissé_x} - Taux_{ref_x} \times (1 - i))^2$$

In this study, coefficients range from 15% to 60% for female right owners, from 12% to 64% for male right owners, and from 0% to 19% for beneficiaries.

Gathering people into age sections allows to limit the importance given to under-represented ages.

- **Positioning methods** : these methods are offered in a computer package created by Frédéric PLANCHET and Julien TOMAS, *ELT*. Three methods are carried out here.

- A one-parameter approach with SMR (*Standardized Mortality Ratio*) :

$$\tilde{q}_{x,t} = SMR \times q_{x,t}^{ref} \quad \text{with} \quad SMR = \frac{\sum_{(x^*, t^*)} d_{x^*, t^*}}{\sum_{(x^*, t^*)} E_{x^*, t^*} \times q_{x^*, t^*}^{ref}}$$

- A two-parameters approach with a relational model :

$$\forall x \in x^*, t \in t^* : \tilde{q}_{x,t} = \frac{\exp(\hat{\alpha} + \hat{\beta} \times \lg(q_{x,t}^{ref}))}{1 + \exp(\hat{\alpha} + \hat{\beta} \times \lg(q_{x,t}^{ref}))} \quad \text{with} \quad \lg(\tilde{q}_{x^*, t^*}) = \alpha + \beta \lg(q_{x^*, t^*}^{ref}) + \varepsilon_{x^*, t^*}$$

- A Poisson generalized linear model including interactions with ages and calendar years :

$$\tilde{q}_{x,t} = \exp(\hat{\beta}_0 + \hat{\beta}_1 \log q_{x,t}^{ref} + \hat{\beta}_2 x + \hat{\beta}_3 t + \hat{\beta}_4 x t)$$

$$\text{with} \quad D_{x^*, t^*} \sim \mathcal{P}(E_{x^*, t^*} \mu_{x^*, t^*}) \quad \text{and} \quad \mu_{x^*, t^*} = \beta_0 + \beta_1 \log q_{x^*, t^*}^{ref} + \beta_2 x^* + \beta_3 t^* + \beta_4 x^* t^*$$

The main objective of these three methods is estimating the models' parameters during the observation period. The adjusted mortality can be estimated during the same period then.

Selection of the most suitable method :

These methods' comparisons are carried out throughout the same statistical measurements that previously (R^2 , χ^2 , and $MAPE$).

The selected method of construction for the mortality table is adjustment by enforcing abatement on the reference table. Graphic results are following :

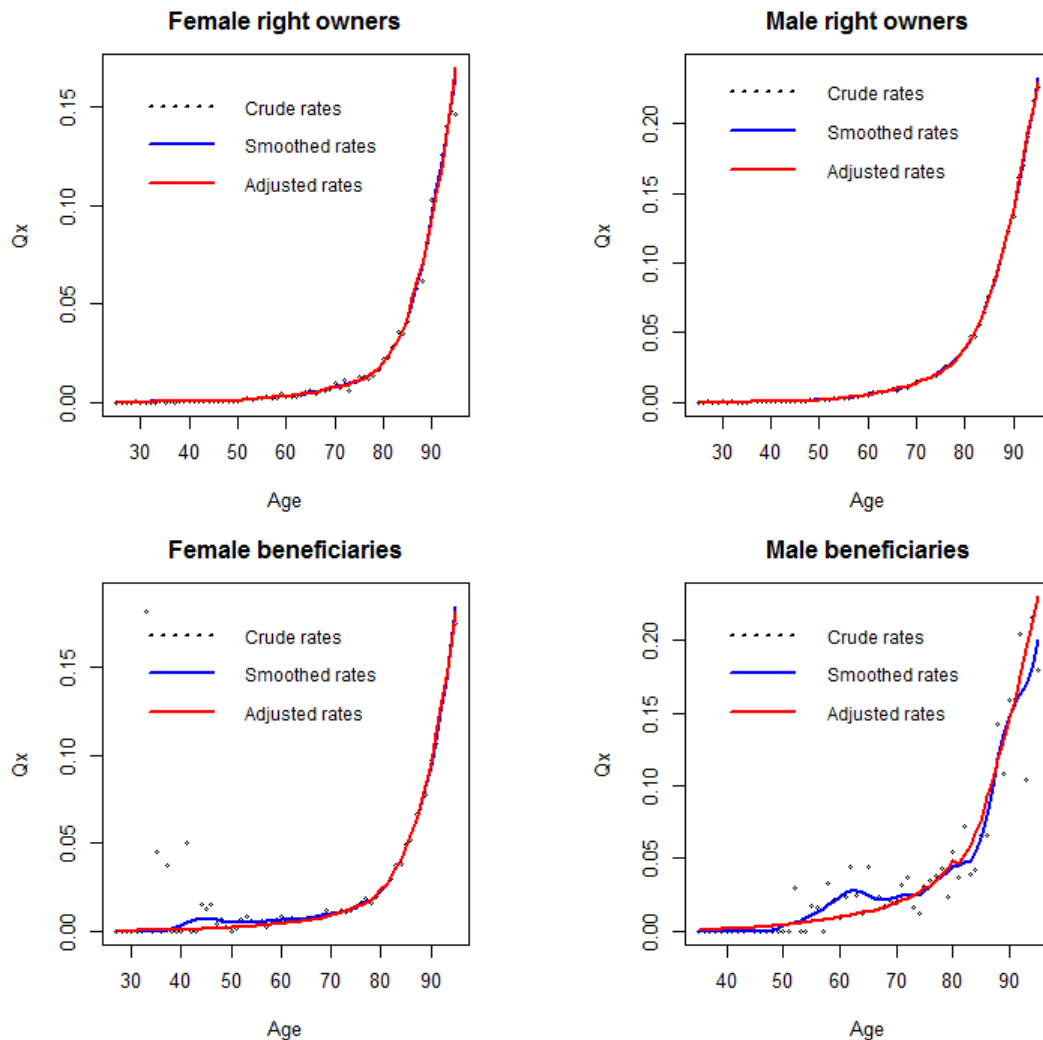


FIGURE 9 – Crude, smoothed, and adjusted mortality rates for each population

Crude rates, smoothed rates, and adjusted rates are aligned for right owners. For beneficiaries, small differences are noticed. These differences are due to an under-representation into some age sections, that can impact the regularity of mortality. The adjustment is therefore considered as quite well.

Statistic measurements between the new law and the actual law used by the pension plan reflect that the new law is more suitable to the pension scheme's population than the actual one.

Closure of the table :

The method used to close the table is the Denuit and Goderniaux's, that consists of the adjustment of a log-quadratic model imposing two conditions.

The first one specifies a probability of dying equal to 100% for people aged 120, and an horizontal tangent at 120. Furthermore, the second one reflects a concavity avoiding a decreasing mortality in old ages.

Evaluation of financial impacts

The second aim of this study is evaluating the financial impacts created by a change of use for mortality tables. This evaluation is carried out throughout projected pension amounts.

Two scenarios are distinguished here :

- A first one that uses the same mortality law for right owners and beneficiaries in pension amounts' calculations (the law used for both is the right owners').
- A second one that uses two different mortality laws for right owners and beneficiaries in pension amounts' calculations.

Both are presented to allow the pension scheme's managers to select the one they prefer maintaining.

Three populations are studied : only active members, only inactive members, and active and inactive members both. The results are presented for the last one.

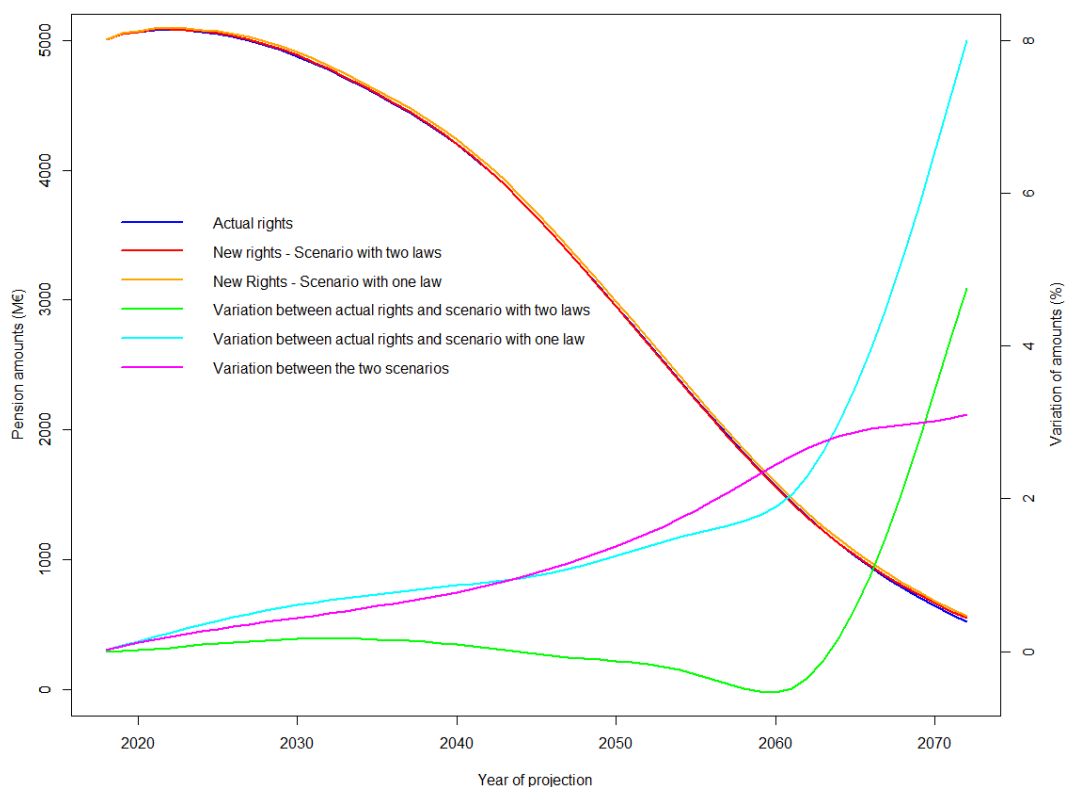


FIGURE 10 – Projected pension amounts and percentages of variation between laws

Whatever the law and scenario considered, the projected cash flows are very alike. Small positive deviations are noticed after forty years of projection (when pensions have decreased substantially).

Results also reflect that the liabilities calculated with the single law are slightly higher than those calculated with two different laws. Thus, the first scenario is considered as more prudent than the second one.

Conclusion

Eventually, this memoir concludes that the new table constructed with an adjustment of the abatement rates on the 2013-2017 INSEE prospective reference mortality table is more suitable to the mortality of the pension plan's population, than the one used until today. Nevertheless, this change of use doesn't have a significant impact on pension obligation.

The update of tables allows to limit the financial impacts due to a change of mortality tables used for calculation . However, this population's size is large enough to study its mortality, but not to create a generational mortality table that can project life expectancies on long-term periods.

Remerciements

Avant toute chose, je tiens à adresser mes plus sincères remerciements à tous les membres de l'équipe du cabinet SPAC Actuaire, qui m'ont permis de réaliser mon stage dans les meilleures conditions.

En premier lieu, je souhaite remercier Olivier SEGUIN - tuteur entreprise, Actuaire et Associé - pour avoir eu la patience et pris le temps de m'accompagner tout au long de ce stage.

Je remercie Brigitte ECARY - Directrice Générale et Actuaire - ainsi que Louis LENGLIN - Actuaire et Associé - pour m'avoir fait confiance en me permettant d'effectuer un stage au sein du cabinet SPAC Actuaire.

Je tiens, par ailleurs, à exprimer ma plus grande reconnaissance à Swan BROUTARD - tutrice universitaire et Actuaire chez BNP Paribas Cardif - et Alexis MERX - tuteur universitaire et Actuaire Sénior Consultant chez Fixage - pour le suivi de mon mémoire et les précieux conseils qu'ils m'ont donnés; ainsi qu'à tout le corps professoral de l'EURIA pour la qualité de l'enseignement et la formation qu'il m'ont permis de suivre au cours de ces trois dernières années.

Enfin, un grand merci à tous mes proches qui m'ont soutenue tout au long de la réalisation de mon mémoire et sans qui cette épreuve aurait été plus difficile.

Merci à toutes ces personnes pour m'avoir apporté de l'aide et de nombreux conseils à chaque fois que j'en avais besoin, toujours dans la bonne humeur même dans les moments les plus délicats, pour leur disponibilité et les moments qui ont été passés ensemble au cours de ces derniers mois. Sans elles, je n'aurais pas pris autant de plaisir à participer à cette expérience.

Table des matières

1	Introduction	20
2	La mortalité à l'usage dans les régimes de retraite	22
2.1	Étude de la population	22
2.1.1	Évolution de la mortalité en France	22
2.1.1.1	Évolution de l'espérance de vie à la naissance	22
2.1.1.2	Évolution du nombre de décès	23
2.1.2	Rappels sur les tables de mortalité	27
2.1.2.1	Quelques notions importantes	27
2.1.2.2	Histoire et tables de mortalité usuelles	29
2.2	Utilisation des tables de mortalité dans les régimes de retraite	32
2.2.1	Une nécessité contextuelle	32
2.2.1.1	Des réformes et un risque de longévité de plus en plus important	32
2.2.1.2	Intervention des tables de mortalité	37
2.2.2	Construction de tables de mortalité propres aux régimes de retraite	37
3	Les données utilisées	39
3.1	Présentation des données	39
3.2	Quelques statistiques descriptives et démographiques	41
4	Création d'une table de mortalité propre à notre régime de retraite	48
4.1	Création d'une table de mortalité par des méthodes non-paramétriques	48
4.1.1	Les différentes méthodes de lissage utilisées	48
4.1.1.1	Lissage par la méthode des moindres carrés et maximisation du R^2	48
4.1.1.2	Lissage par la méthode de Makeham	54
4.1.1.3	Lissage par la méthode de Loess	58
4.1.1.4	Lissage par la méthode de Whittaker-Henderson	62
4.1.1.5	Comparaison des résultats obtenus	66
4.1.1.6	Validation de la méthode de lissage retenue	74
4.1.2	Étude par décalage d'âges	77
4.1.3	Étude par abattement de la table de référence	79
4.2	Création d'une table de mortalité à l'aide du <i>package ELT</i>	81
4.2.1	Les différentes méthodes proposées : principe et application	82
4.2.1.1	Méthode 1 : Approche à un paramètre avec le SMR	82
4.2.1.2	Méthode 2 : Approche à deux paramètres avec un modèle relationnel	83
4.2.1.3	Méthode 3 : Approche par modèle linéaire généralisé de Poisson incluant des interactions avec l'âge et l'année calendaire	85
4.2.2	Comparaison des méthodes proposées	85
4.2.2.1	Methodologie des tests appliqués	86
4.2.2.2	Comparaison des résultats obtenus	90
4.2.2.3	Choix de la méthode retenue pour la construction de notre table de mortalité	93
4.3	Fermeture de la table	97

5	Évaluation des impacts financiers	99
6	Conclusion	102
A	Le diagramme de Lexis	112
B	Tables de mortalité TGF-05 et TGH-05	115
C	Compléments sur les réformes des retraites	116
C.1	La réforme de 2003	116
C.2	La réforme de 2010	117
C.3	La réforme de 2014	117
D	Parts en pourcentage d'individus au sein des populations	118
D.1	Population globale	118
D.2	Ouvrants droit par tranche d'âges	118
D.3	Ayants droit par tranche d'âges	118
E	Comparaison des différentes méthodes de lissage	119
E.1	Population des Hommes OD	119
E.2	Population des Femmes AD	120
E.3	Population des Hommes AD	120
E.4	Valeurs du χ^2 sur la tranche 60-89 ans	121
E.5	Valeurs de la <i>MAPE</i> sur la tranche 60-89 ans	121
F	Méthodologie des tests appliqués sur les résidus	122
F.1	Test de Shapiro	122
F.2	Test de Box-Pierce	122
F.3	Test de White	123
G	Diagrammes <i>Quantile-Quantile</i> des résidus entre les taux bruts et les taux lissés	124
G.1	Hommes OD	124
G.2	Femmes AD	125
G.3	Hommes AD	125
H	Résultats complémentaires obtenus à partir des méthodes du <i>package ELT</i>	126
H.1	Méthode 1 appliquée à la population des ouvrants droit : complément sur les SMR	126
H.2	Méthodes de positionnement appliquées sur la population des ayants droit	126
H.2.1	Méthode 1 : Approche à un paramètre avec le SMR	126
H.2.2	Méthode 2 : Approche à deux paramètres avec un modèle relationnel	127
H.2.3	Méthode 3 : Approche par un modèle linéaire généralisé de Poisson incluant des interactions avec l'âge et l'année calendaire	127
I	Méthodes de validation d'ajustement	128
I.1	χ^2	128
I.2	R^2	128
I.3	<i>MAPE</i>	128
J	Résultats des tests du <i>package ELT</i> et comparaisons sur la population des ayants droit	129
K	Comparaison de la nouvelle loi avec les observations annuelles	132
L	Espérance de vie	137
L.1	Comparaison des espérances de vie par sexe	137
L.2	Comparaison des espérances de vie par qualité	138
M	Tables de mortalités ajustées	139

Chapitre 1

Introduction

On constate de nombreuses disparités en ce qui concerne l'espérance de vie chez les individus. En effet des écarts non négligeables peuvent être observés entre les espérances de vie au même âge de deux personnes de sexes opposés, de catégories socio-professionnelles différentes, ou encore en fonction de leur situation de famille. Ces disparités sont accompagnées d'une évolution régulière de l'espérance de vie qui entraîne un impact sur la durée de vie passée en retraite et a des répercussions sur le montant des charges de pensions à verser.

L'usage de tables de mortalité est indispensable en retraite dans le cadre de la projection de charges et de ressources financières. L'utilisation d'une table mal adaptée à la population peut entraîner de gros écarts au niveau du calcul des moyens de financement. Les régimes cherchent alors à évaluer leurs produits et charges de façon la plus précise possible en tenant compte de chacun des risques auxquels leurs populations d'affiliés et eux-mêmes sont soumis. Pour cela, ils essaient d'utiliser des tables qui reflètent au plus près la mortalité des populations qui les composent. Ces tables serviront ensuite à projeter les populations dans le futur afin de connaître, ou du moins d'estimer leur durée de vie restante.

L'usage de nouvelles tables différentes de celles fournies par l'Institut National de la Statistique et des Études Économiques (INSEE) aura des conséquences financières pour les régimes. L'enjeu de ce mémoire est de construire une table de mortalité d'expérience qui suit la population des affiliés d'un régime de retraite particulier, puis d'en évaluer les impacts financiers suite à son utilisation.

Pour atteindre ces objectifs, ce mémoire se décomposera en plusieurs sections regroupées au sein de quatre grandes parties.

Une première partie, de nature assez descriptive, sera consacrée à la mise en place du contexte et des notions importantes utiles à la bonne compréhension et au suivi du travail effectué tout au long de cette étude. Nous réaliserons une analyse de l'évolution de l'espérance de vie au cours des derniers siècles en France avant de poursuivre sur la description des différentes tables de mortalité existantes, et enfin terminer sur la mise en relation de ces deux grands thèmes que sont la retraite et la mortalité : la façon dont les tables de mortalité sont utilisées par les régimes de retraite.

Nous proposons de consacrer la seconde partie aux données utilisées. Elles seront présentées dans le détail mais dans la limite de ce qui nous est utile pour notre étude. Des statistiques descriptives et démographiques seront également effectuées, ce qui permettra une première analyse de la population qui compose le régime en question.

La construction d'une table de mortalité propre à la population du régime étudié fera l'objet de la troisième partie. Cette table sera établie au travers de l'exploitation de diverses méthodes mais résultera néanmoins de l'ajustement d'une table de mortalité de référence. La période d'observation se résumera aux années 2013 à 2017, une plage de cinq années qui semble suffisante pour pallier les éventuels évènements

nements accidentels survenus exceptionnellement au cours d'une certaine année. Nous constaterons également que l'utilisation des méthodes évoquées ci-dessus, consistant en l'abattement ou au décalage d'âges des données d'une table de référence, requerront un lissage préalable des observations de sorte à limiter l'importance accordée à d'éventuelles aberrations dues à une taille de population trop réduite. Plusieurs techniques de lissage seront alors mises en oeuvre afin de retenir la plus pertinente vis à vis des données dont nous disposons. Une fois les deux méthodes d'ajustement appliquées à notre population, nous les comparerons à des méthodes statistiques proposées par un *package* spécifique à la construction de tables de mortalité prospectives disponible sous le logiciel *RStudio*. La clôture de ce chapitre se fera au travers d'une analyse qui permettra de valider la cohérence de la nouvelle table ajustée avec les statistiques démographiques observées.

Enfin, ce mémoire aura pour dernier chapitre l'évaluation des impacts financiers engendrés par l'utilisation de la nouvelle table de mortalité. Cette évaluation se fera au travers de la comparaison des flux de pensions probables en droits acquis calculés à l'aide de chacune des deux tables et pour différentes populations au sein du régime (actifs, inactifs, actifs et inactifs réunis).

Chapitre 2

La mortalité à l'usage dans les régimes de retraite

Étudier la mortalité est très important en retraite, notamment dans le cadre de calculs de rente viagères où la valeur de la rente est estimée à partir des tables de mortalité établies par l'INSEE ou par des compagnies d'assurance. Une étude régulière de cette mortalité est nécessaire du fait de l'évolution de l'espérance de vie (surtout après 60 ans) qui aura un impact sur le financement des régimes de retraite.

2.1 Étude de la population

2.1.1 Évolution de la mortalité en France

La démographie de la population française s'est montrée très évolutive au cours des derniers siècles. Nous proposons de présenter dans cette section les principales évolutions observées en termes d'espérance de vie, de décès et de vieillissement de la population en France au cours du temps.

2.1.1.1 Évolution de l'espérance de vie à la naissance

L'espérance de vie d'une population correspond à sa durée de vie moyenne. Il s'agit du meilleur indicateur permettant de suivre l'évolution de la mortalité d'une population. Cette espérance de vie n'a jamais cessé d'augmenter au cours des derniers siècles, notamment du fait des progrès scientifiques et médicaux.

Le graphique 2.1, communiqué par l'INED, présente l'évolution de l'espérance de vie à la naissance des hommes et des femmes entre 1740 et 2010 :

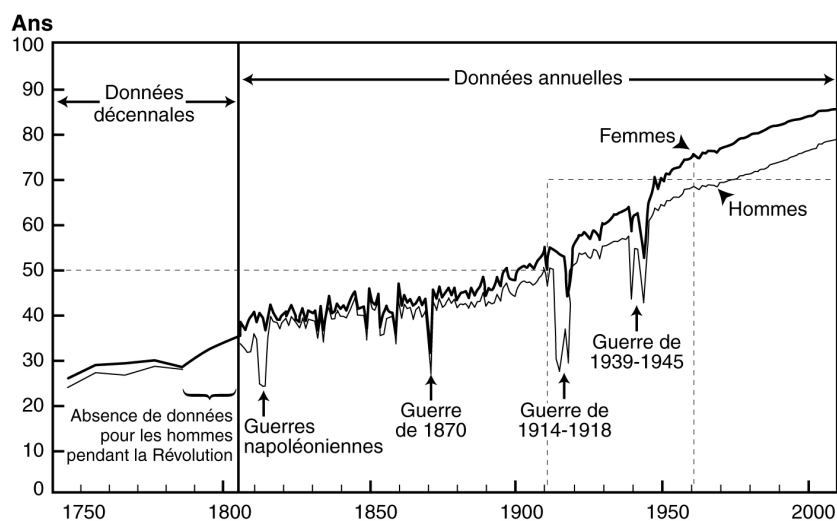


FIGURE 2.1 – Espérance de vie à la naissance en France entre 1740 et 2010

La mortalité était plus que conséquente au cours du XVIIIe siècle et l'espérance de vie était alors évaluée autour des 27 ans, mais au début du siècle suivant, la population française atteint une espérance de vie de 37 ans, ceci notamment grâce à la découverte du tout premier vaccin ¹.

Il est ensuite observé une progression faible mais constante tout au long du XIXe siècle, l'espérance de vie atteint alors un âge de 45 ans en 1900. Cette espérance de vie augmente beaucoup plus rapidement au cours du siècle suivant et atteint alors un âge de 79,4 ans pour les hommes et 85,4 ans pour les femmes en 2016.

Notons par ailleurs que des "pics descendants" sont présents au niveau des périodes de guerre, mais retenons que l'espérance de vie a plus que doublé en deux siècles et demi. En deux siècles, l'espérance de vie a augmenté de 35 ans pour les hommes et de 38 ans pour les femmes. À la sortie de la Seconde Guerre Mondiale, l'espérance de vie était de 60 ans pour les hommes et de 65 ans pour les femmes.

La progression de l'espérance de vie s'explique par les progrès scientifiques qui permettent de plus en plus d'élaborer de nouveaux traitements contre les maladies (découvertes de nouveaux vaccins, création de nouveaux médicaments et antibiotiques...), mais aussi par une baisse de la mortalité infantile lors des accouchements, ce qui permet d'atteindre une espérance de vie de 82,7 ans pour les femmes et 75,2 ans pour les hommes (et donc en moyenne 79 ans sur toute la population) au début du XXIe siècle.

Par ailleurs, remarquons que l'espérance de vie des femmes est supérieure à celle des hommes. Ces différences s'expliquent par des phénomènes biologiques, mais aussi par des différences de modes de vie. En effet, les hommes sont plus sujets au tabac, à l'alcool, à la violence, ou encore aux risques professionnels... L'homme est alors plus exposé au risque de mortalité que la femme.

Toutefois, ces écarts se resserrent peu à peu du fait du rapprochement du mode de vie des femmes à celui des hommes (par exemple du fait de l'augmentation du tabagisme chez les femmes).

Les écarts observés sont de l'ordre de +5,3 ans pour les femmes en 1945, de +8,2 ans en 1990, et de +7,5 ans en 2000. Les études démographiques prévoient la continuité de ce rapprochement dans le temps.

2.1.1.2 Évolution du nombre de décès

Tout comme l'espérance de vie, la mortalité et plus particulièrement le nombre de décès en France par an a connu certaines évolutions que nous proposons de détailler dans cette sous-section.

— Mortalité infantile :

Comme énoncé précédemment, la mortalité infantile a connu d'importantes évolutions. En effet, la mortalité infantile était beaucoup plus élevée autrefois qu'aujourd'hui, et c'est ce que nous constatons sur le graphique 2.2 (également communiqué par l'INED), qui présente l'évolution de la mortalité infantile en France entre 1740 et 2010.

Au cours du XVIIIe siècle, il est constaté que près d'un enfant sur trois décède avant d'avoir atteint son premier anniversaire. Cette mortalité était très souvent la conséquence de maladies infectieuses.

Durant le siècle suivant, le développement des traitements contre les maladies infectieuses, la découverte des nouveaux vaccins, ainsi que l'amélioration des conditions d'accouchement et de vie, permettent de réduire cette mortalité à un enfant sur six en 1850.

Cependant, courant la seconde moitié du XIXe siècle, la mortalité infantile remonte du fait de l'industrialisation et de l'amas des populations dans les villes qui développent de nombreuses épidémies telles que le choléra. En 1900, 15% des enfants meurent avant l'âge d'un an.

Lors de la deuxième décennie du XXe siècle, la mortalité infantile connaît deux pics dus à la canicule de 1911 et à l'épidémie de grippe espagnole de 1918.

1. 1796 : découverte du tout premier vaccin. Il s'agit du vaccin contre la variole, découvert par le scientifique et médecin anglais Edward JENNER (1749-1823).

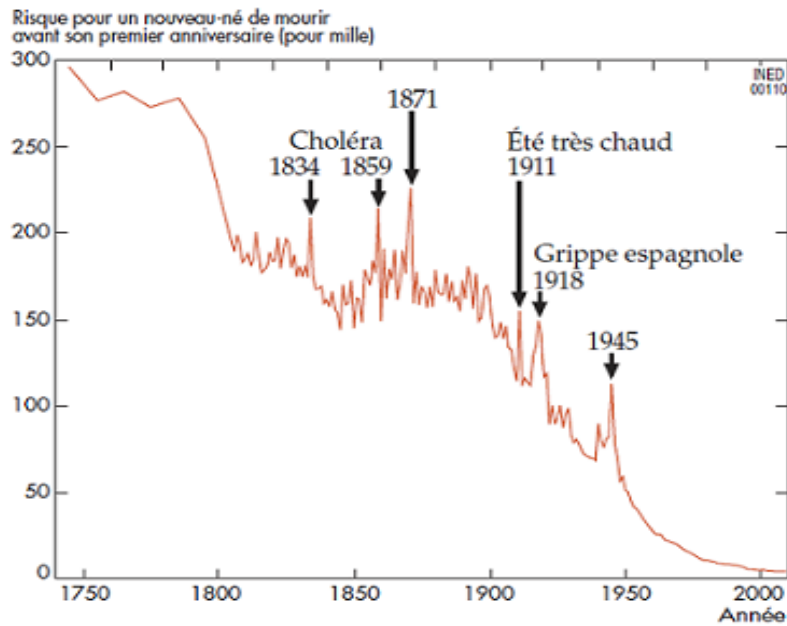


FIGURE 2.2 – Évolution de la mortalité infantile en France entre 1740 et 2010

En 1950, plus que 5% des enfants décèdent avant d’avoir atteint leur premier anniversaire, avec toutefois un nouveau pic en 1945 des suites d’une désorganisation des circuits de distribution de lait qui a suivi la Libération. Enfin en 2015, le taux de mortalité chez les enfants de moins d’un an est évalué à 3,5‰.

Le développement des traitements contre les maladies infectieuses a grandement permis de reculer le niveau de mortalité chez les enfants. De nos jours, cette mortalité se montre très faible, il s’agit d’événements rares et souvent de nature accidentelle.

— Nombre de décès sur la population totale :

L’évolution du nombre de décès par an de la population globale en France depuis 1920 est représentée sur le graphique 2.3 :

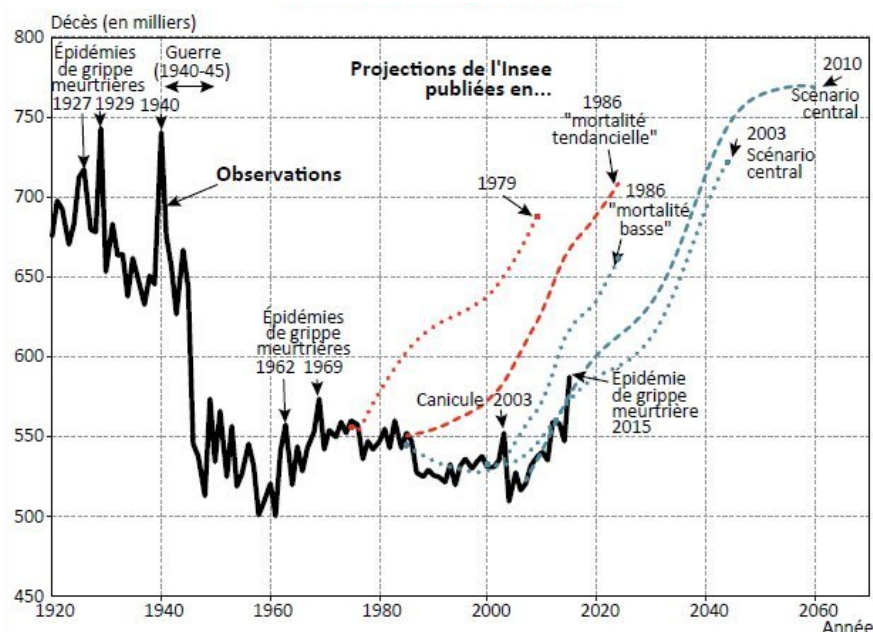


FIGURE 2.3 – Évolution annuelle des décès en France depuis 1920²

2.1. ÉTUDE DE LA POPULATION

De même que pour la mortalité infantile, les progrès scientifiques et médicaux ont permis à la mortalité globale de diminuer de manière conséquente jusqu'au début du XXe siècle. Toutefois, les épidémies de grippe et la Seconde Guerre Mondiale intervenue dans la première moitié de ce même siècle ont fait que le nombre de décès resta toujours conséquent avec près de 700 000 décès en moyenne par an sur la période 1920-1945.

Pendant la période d'après guerre, le nombre de décès a considérablement diminué, et nous constatons que celui-ci a peu évolué depuis cette période. En effet, le nombre de décès par an est plutôt stable depuis ces 70 dernières années avec environ 550 000 décès par an, sauf exceptions de quelques pics lors d'événements environnementaux tels que la canicule de 2003 ou encore lors des épidémies de grippe de 1962, 1969 et 2015.

Cette stabilité du nombre de décès par an au cours des dernières décennies peut paraître surprenante à première vue du fait que l'espérance de vie augmente, et donc que nous devrions observer une baisse de la mortalité. Ce phénomène s'explique par un accroissement constant de la population d'année en année. En effet, analysons le graphique 2.4 qui représente l'évolution de la population et du nombre de décès observés en France sur la période 1946-2000 :

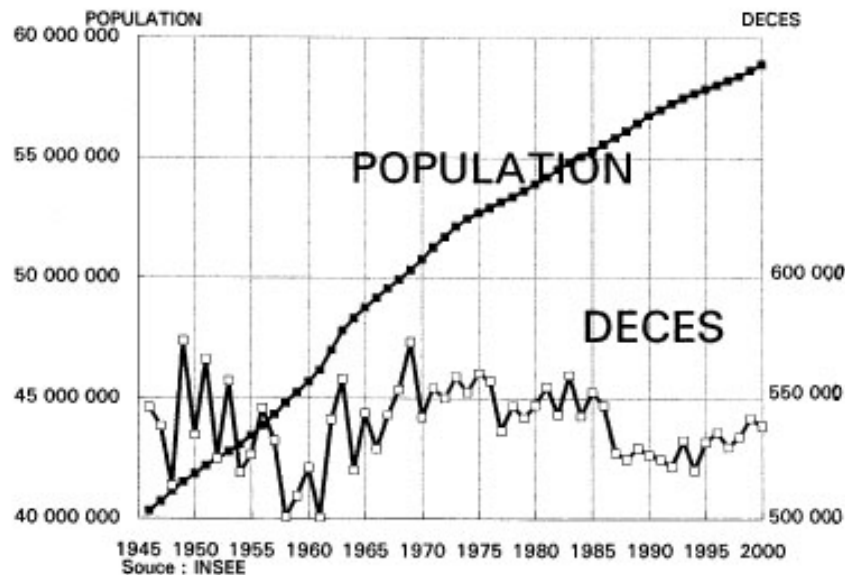


FIGURE 2.4 – Évolution de la population et du nombre de décès en France entre 1946 et 2000

Nous constatons effectivement que la population ne cesse d'augmenter mais que le nombre de décès reste plutôt stable. Toutefois, en termes de rapport entre le nombre de décès observés et l'effectif de la population, il convient de remarquer une baisse d'au moins 50% de la mortalité entre les années 1945 et 2000. Autrement dit, si la population était restée constante sur cette période, le nombre de décès observés aurait diminué d'au moins 50% entre les années 1945 et 2000. C'est pourquoi nous pouvons parler de baisse de la mortalité mais pas de baisse du nombre de décès.

Par ailleurs, des études démographiques prévoient une augmentation du nombre de décès dans les années à venir, notamment du fait que les *baby-boomers* arrivent à l'âge d'entrée dans la catégorie des "personnes âgées" ainsi que du fait que l'espérance de vie augmente régulièrement. Cela entraîne alors une augmentation du nombre de personnes qui sont soumises au risque de décès avec une probabilité plus importante.

Ce risque est qualifié de "vieillesse de la population". Nous pouvons remarquer cela sur les pyramides des âges proposées ci-après³.

2. Source : INSEE

3. Source : Pison, 2014

2.1. ÉTUDE DE LA POPULATION

À chaque numéro correspond un événement qui a eu une influence sur la démographie française :

- 1 : Déficit de naissances dû à la Guerre de 1870
- 2 : Été 1911 : Canicule qui a entraîné une forte mortalité chez les nouveaux nés
- 3 : Pertes militaires de la Première Guerre Mondiale
- 4 : Déficit de naissances dû à la Première Guerre Mondiale (il s'agit de classes creuses)
- 5 : Passage des classes creuses à l'âge de la fécondité
- 6 : Déficit de naissances dû à la Seconde Guerre Mondiale
- 7 : Commencement du *Baby-Boom*
- 8 : Fin du *Baby-Boom*

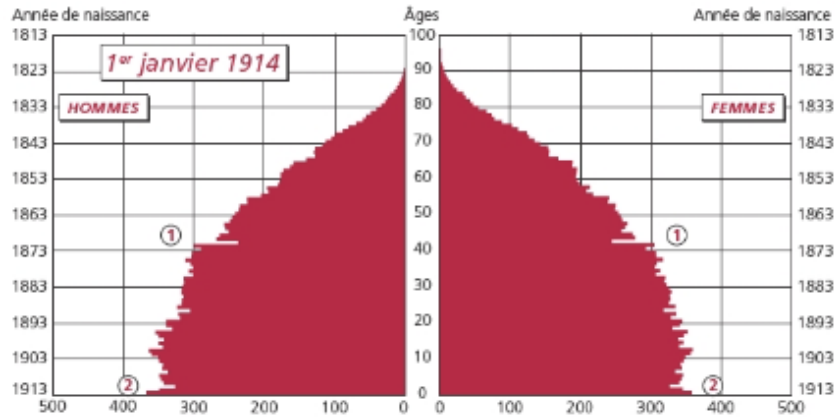


FIGURE 2.5 – Pyramide des âges en 1914

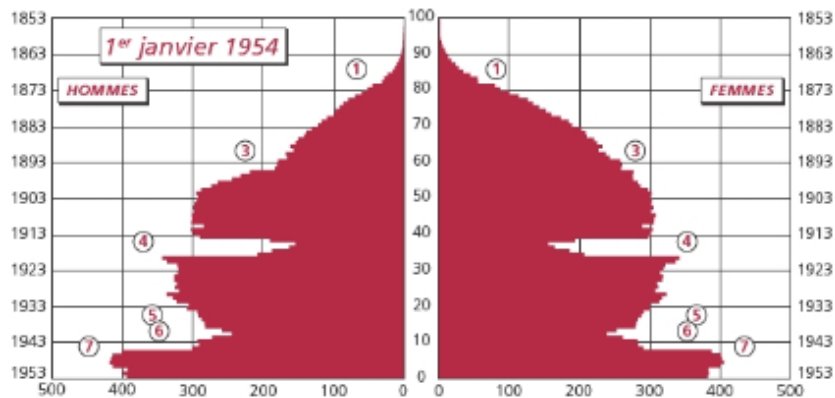


FIGURE 2.6 – Pyramide des âges en 1954

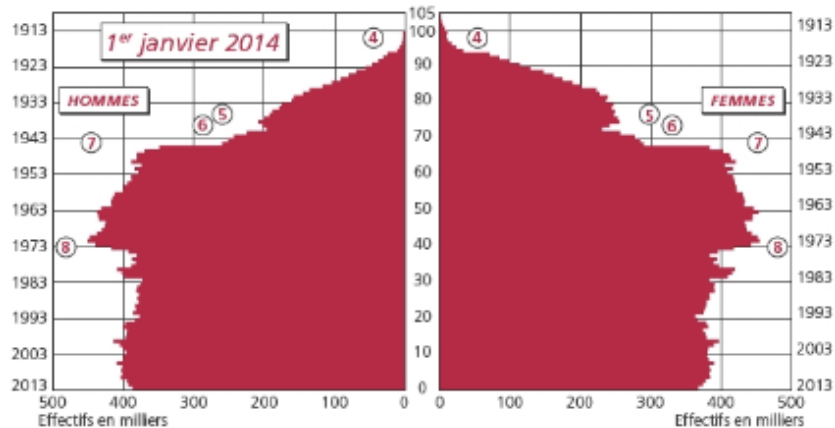


FIGURE 2.7 – Pyramide des âges en 2014

Nous constatons alors bien que les effectifs des classes de personnes âgées ont beaucoup augmenté en un siècle, et même en un demi-siècle.

En effet, de manière générale, les pyramides se sont élargies au fil du siècle du fait de l'augmentation de l'espérance de vie. Cet élargissement est notamment visible sur les tranches d'âges élevées. Par exemple, en 2014, le nombre d'hommes âgés de 80 ans dépasse 150 000 individus et celui des femmes atteint presque 250 000 personnes; tandis que nous avons environ 50 000 hommes âgés de 80 ans en 1954 et 100 000 femmes, contre moins de 50 000 personnes dans les deux populations en 1914.

Il est également possible de remarquer qu'en surcroît de l'augmentation de l'espérance de vie, la classe des *baby-boomers*⁴ se situe dans la tranche d'âges 50-70 ans en 2014. Cette classe sera alors amenée à "vieillir" et à commencer à atteindre des âges élevés dans les prochaines années et par conséquent à être exposée plus grandement au risque de décès.

Ainsi, il convient de bien distinguer l'évolution de l'espérance de vie de l'évolution du nombre de décès, car une augmentation de l'espérance de vie ne sera pas reflétée par une baisse du nombre de décès. Toutefois, les effectifs des populations seront amenés à croître avec l'espérance de vie, ce qui entraînera qu'un nombre de décès constant se traduira par une baisse du rapport décès sur population.

2.1.2 Rappels sur les tables de mortalité

2.1.2.1 Quelques notions importantes

En démographie et plus particulièrement dans les études de la mortalité, certaines notions sont incontournables.

Avant de poursuivre sur la présentation des différentes tables de mortalité, nous proposons dans cette courte section, de définir brièvement ces notions qui pourront être utiles tout au long de notre étude.

Tout d'abord, définissons T comme la **durée de vie** d'un individu de la population de référence, et T_x comme la **durée de vie restante** d'un individu sachant qu'il a atteint l'âge x (nous pouvons nous restreindre à $x \in \mathbb{N}$). Nous avons de ce fait :

$$\mathcal{L}(T_x) = \mathcal{L}(T - x \mid T > x)$$

Ensuite, nous définissons la **probabilité de survie** ${}_t p_x$ comme la probabilité qu'un individu d'âge x soit encore en vie à l'âge $x + t$:

$${}_t p_x = \mathbb{P}(T_x > t) = \mathbb{P}(T > x + t \mid T > x)$$

Par opposition, la **probabilité de décès** ${}_t q_x$ correspond à la probabilité qu'un individu d'âge x n'atteigne pas l'âge $x + t$:

$${}_t q_x = 1 - {}_t p_x = \mathbb{P}(T \leq x + t \mid T > x)$$

La **fonction de survie** $x \mapsto l_x$ correspond au nombre moyen de survivants à l'âge x d'une cohorte de l_0 individus tous nés à une même date.

$l_x - l_{x+t}$ donne alors parmi les l_x individus âgés de x années, le nombre de décès survenus entre les âges x et $x + t$. Il est important de noter que l_x correspond à l'**effectif initial** de la population d'âge x présent sur la période d'observation. Par conséquent, le terme l_{x+t} s'identifie au nombre d'individus initialement âgés de x années et toujours vivants à la fin de la période t , soit au nombre d'individus âgés de $x + t$ années au début d'une nouvelle période d'observation.

4. Les générations appartenant à la classe des *baby-boomers* sont les générations comprises entre 1946 et 1964.

Donc si nous définissons ${}_t d_x$ comme étant le **nombre de décès survenus entre x et $x+t$** , nous avons :

$${}_t d_x = l_x - l_{x+t}$$

et nous pouvons alors donner des valeurs aux probabilités de décès :

$${}_t q_x = \frac{{}_t d_x}{l_x} = \frac{l_x - l_{x+t}}{l_x}$$

En démographie, ce nombre est appelé **quotient de mortalité**.

Le nombre ${}_t d_x$ peut également être calculé à l'aide d'une indicatrice δ_i associée à la mortalité de chaque individu i :

$$\delta_i = \begin{cases} 1 & \text{si l'individu } i \text{ est décédé} \\ 0 & \text{sinon} \end{cases}$$

Le nombre de décès survenus entre les âges x et $x+t$ s'exprime alors :

$${}_t d_x = \sum_{i=0}^{l_{x+t}} \delta_i$$

Nous définissons par ailleurs le **taux de mortalité** sur l'intervalle $[x, x+t]$:

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

avec :

$${}_t L_x = \int_0^t l_{x+s} ds$$

Il convient de préciser que dans le cas du taux de mortalité, le rapport se fait à l'effectif moyen sur l'intervalle de temps, tandis que pour le quotient de mortalité, le rapport se fait à l'effectif initial sur la période.

Le critère ${}_t L_x$ mesure le **nombre d'années vécues par tous les individus** de la population entre x et $x+t$, et se calcule en général par une somme. Cette somme est alors appelée **exposition au risque à l'âge x sur la période $(x, x+t)$** , se note ${}_t E_x$.

Soit, en considérant un individu i , le temps τ_i durant lequel il est observé sur la période. Ce temps s'agit de l'**exposition au risque** de cet individu.

L'exposition au risque sur l'ensemble de la période se calcule alors :

$${}_t E_x = \sum_{i=0}^{l_{x+t}} \tau_i$$

La plupart de ces notions seront utilisées au cours de ce mémoire, dans la partie permettant de modéliser la mortalité de la population étudiée.

2.1.2.2 Histoire et tables de mortalité usuelles

Les tables de mortalité sont apparues au cours du XVII^e siècle. Ces tables reposaient alors sur l'analyse du nombre de décès hebdomadaires observés, et indiquaient à chaque âge un nombre de décès moyen, un nombre de survivants moyen, ainsi que des probabilités de décès et de survie. De nos jours, elles sont établies à partir de l'observation d'une population de taille assez importante et ce, sur plusieurs exercices de sorte à lisser les éventuels phénomènes conjoncturels tels que les épidémies par exemple.

Ces tables présentent des indicateurs qui chiffrent la mortalité des populations et permettent la tarification de contrats d'assurance (assurance vie, assurance décès ou encore rentes viagères...) en offrant des prévisions sur l'âge de décès des individus. Ces tables sont révisées de manière périodique, de façon à tenir compte de l'évolution de l'espérance de vie de la population au fil des années.

Nous distinguons deux catégories de tables de mortalité : d'un côté les **tables réglementaires**, et d'un autre côté les **tables d'expérience**.

— Les tables réglementaires :

Il s'agit de tables élaborées le plus souvent par l'INSEE ou par des organismes privés à partir des statistiques de décès observées pour pouvoir disposer de tables de références qualifiées de "prudentes", souvent utilisées dans le cas de tarification de contrats de rentes viagères.

Parmi les plus connues, nous retrouvons par exemple les tables TH-00-02 (ou TF-00-02) ou encore TGH-05 (ou TGF-05), qui sont des tables assurantielles et dont nous détaillerons le contenu un peu plus loin dans ce rapport.

— Les tables d'expérience :

Il s'agit de tables de mortalité calculées à partir d'observations constatées sur une population donnée. Souvent, ce sont des professionnels (principalement dans le secteur de l'assurance) qui construisent ce genre de tables, à partir des données historiques dont ils disposent sur leur propre population.

Toutefois, malgré le fait que ces tables ne sont utilisables que sur les populations à partir desquelles elles sont conçues, elles doivent être certifiées par des actuaires agréés.

Parmi les tables présentées ci-dessus, certaines correspondent à des **tables du moment** (ou **tables instantanées**) et d'autres à des **tables générationnelles**.

— Les tables du moment :

Il s'agit de tables établies à partir de la mortalité observée sur une population entière d'individus sur une période de courte durée (de l'ordre de quelques années).

Ces tables sont généralement composées de trois colonnes :

- La première colonne correspond à la variable contenant les différents âges x .
- La deuxième colonne correspond au nombre de survivants à chaque âge x . Cette variable est généralement notée l_x .
- La troisième colonne correspond aux probabilités de décéder à chaque âge x , notées q_x .

Notons que le nombre moyen de décès à l'âge x s'obtient via la relation suivante :

$$d_x = l_x \times q_x$$

avec :

$$q_x = \frac{l_x - l_{x+1}}{l_x}$$

2.1. ÉTUDE DE LA POPULATION

Parmi les tables du moment réglementaires les plus utilisées, nous retrouvons les tables TH-00-02 et TF-00-02 énoncées précédemment, qui présentent la mortalité chez les hommes et les femmes entre les années 2000 et 2002. Elles sont utilisées en assurance vie dans le cas d'un versement de capital décès par exemple. La tranche d'âges considérée dans ces tables est limitée à la tranche 0 - 112 ans.

TF 00-02		
Âge	l_x	q_x
0	100 000	0.3840%
1	99 616	0.0331%
2	99 583	0.0211%
3	99 583	0.0171%
4	99 545	0.0141%
...
40	98 242	0.1140%
41	98 130	0.1253%
42	98 007	0.1377%
43	97 872	0.1512%
44	97 724	0.1647%
...
108	44	54.5455%
109	20	55%
110	9	55.5556%
111	4	75%
112	1	100%

TH 00-02		
Âge	l_x	q_x
0	100 000	0.4890%
1	99 511	0.0382%
2	99 473	0.0271%
3	99 446	0.0221%
4	99 424	0.0181%
...
40	96 369	0.2366%
41	96 141	0.2642%
42	95 887	0.2931%
43	95 606	0.3253%
44	95 295	0.3599%
...
108	4	50%
109	2	50%
110	1	100%
111	0	100%
112	0	100%

TABLE 2.1 – Extrait de la table du moment TF 00-02 TABLE 2.2 – Extrait de la table du moment TH 00-02

— Les tables générationnelles :

Il s'agit de tables établies de la même manière que les tables du moment, c'est-à-dire à partir de la mortalité d'une population observée, néanmoins, la période d'observation est beaucoup plus longue car dans ces tables, la mortalité dépend de l'année de naissance et du sexe des individus constituant la population.

Initialement, les tables générationnelles étaient construites après chaque extinction de génération dans la population observée, mais les études qui nous intéressent surtout de nos jours, sont de savoir comment va évoluer la mortalité dans le futur. C'est pourquoi sous diverses hypothèses d'évolution de la mortalité dans le futur ainsi qu'à partir d'autres tables de mortalité, il est possible d'utiliser des **tables dites prospectives** construites à partir d'exercices de projections de populations.

Ces tables ont une allure un peu différente de celle des tables du moment. En effet, les tables générationnelles se lisent en diagonale de sorte à avoir la mortalité d'une population pour une année donnée :

- La première colonne correspond aux différentes générations t .
- La première ligne correspond à la variable contenant les différents âges x (de projection).
- Le tableau central peut contenir différents éléments : soit le nombre de survivants de chaque génération projeté à chaque âge, soit le nombre de décès de chaque génération projeté à chaque âge, soit les probabilités projetées de décéder entre deux âges pour chaque génération.

Dans le cas de tables générationnelles "classiques", les âges se trouvent en colonne et les générations en ligne. Cela ne change pas le principe d'utilisation de la table, il s'agit seulement d'une différence de mode de lecture par rapport à une table générationnelle prospective.

2.1. ÉTUDE DE LA POPULATION

Parmi les tables générationnelles réglementaires les plus utilisées, nous retrouvons les tables TGH05 et TGF05 énoncées précédemment, qui présentent les nombres de survivants âge par âge respectivement chez les hommes et les femmes pour chaque génération comprise entre 1900 et 2005.⁵

D'autres tables prospectives existent, comme la TPG-93, ou encore les tables INSEE 2007-2060 et 2013-2070. La première énoncée est une table prospective de rente viagère (ou table prospective par génération) construite à partir de l'étude de la mortalité des générations de 1887 à 1993, et les tables prospectives INSEE tiennent compte respectivement des générations de 1888 à 2006 et 1893 à 2012.

TP INSEE 2013 - 2070 Femmes										
Âge Gén.	0	1	2	...	59	60	...	118	119	120
1893	-	-	-	...	-	-	...	-	100 000	58 596
1894	-	-	-	...	-	-	...	100 000	58 596	35 475
1895	-	-	-	...	-	-	...	58 596	35 475	20 549
...
1952	-	-	-	...	-	100 000	...	1.365	0.740	0.402
1953	-	-	-	...	100 000	99 566	...	1.455	0.789	0.428
1954	-	-	-	...	99 592	99 177	...	1.544	0.837	0.454
...
2010	-	-	100 000	...	98 495	98 391	...	3.718	2.017	1.094
2011	-	100 000	99 979	...	98 512	98 409	...	3.719	2.017	1.094
2012	100 000	99 948	99 925	...	98 498	98 394	...	3.719	2.017	1.094

TABLE 2.3 – Extrait de la table INSEE prospective 2013 - 2070 des Femmes

TP INSEE 2013 - 2070 Hommes										
Âge Gén.	0	1	2	...	59	60	...	118	119	120
1893	-	-	-	...	-	-	...	-	100 000	63 928
1894	-	-	-	...	-	-	...	100 000	63 928	40 660
1895	-	-	-	...	-	-	...	63 928	40 660	26 268
...
1952	-	-	-	...	-	100 000	...	0.359	0.193	0.103
1953	-	-	-	...	100 000	98 993	...	0.385	0.207	0.111
1954	-	-	-	...	99 088	98 157	...	0.417	0.224	0.120
...
2010	-	-	100 000	...	97 304	97 135	...	1.282	0.688	0.369
2011	-	100 000	99 978	...	97 371	97 202	...	1.283	0.688	0.369
2012	100 000	99 935	99 911	...	97 383	97 215	...	1.283	0.688	0.369

TABLE 2.4 – Extrait de la table INSEE prospective 2013 - 2070 des Hommes

Par convention, une table de mortalité s'initialise toujours avec une population de 100 000 individus, l'âge 0 correspond à la naissance d'un individu, et la table se termine à l'âge pour lequel nous observons pour la dernière fois des individus (cet âge est souvent compris entre 110 et 130 ans pour les tables réglementaires).

Il est à noter que l'ensemble de ces tables sont révisées régulièrement afin de tenir compte de l'évolution de l'espérance de vie au cours du temps.

5. Un extrait de chacune de ces tables pour chaque sexe est disponible en Annexe B

2.2 Utilisation des tables de mortalité dans les régimes de retraite

2.2.1 Une nécessité contextuelle

2.2.1.1 Des réformes et un risque de longévité de plus en plus important

Au cours de ces dernières années, les diverses réformes des retraites ont mené à décupler les exercices de projection à long terme de la situation des régimes de retraite en matière démographique et financière. Afin d'évaluer l'efficacité des mesures adoptées, cet exercice de prévision sera encouragé à se consolider dans le futur par un suivi régulier des réformes engagées.

Regardons de plus près les différents impacts causés par la mise en place de réformes. Nous nous concentrons sur les réformes de 2003, 2010 et 2014⁶ qui avaient pour objet :

- Réforme de 2003 : instauration d'une règle de partage de gain d'espérance de vie à 60 ans. Deux tiers de ces gains devaient se refléter sur la durée d'activité, le dernier tiers sur la durée de retraite, ceci afin de "stabiliser à horizon 2020 le rapport entre le temps de travail et le temps de retraite afin d'assurer la pérennité des régimes par répartition et l'équité entre générations"⁷.
- Réforme de 2010 : recul des bornes d'âge de liquidation.
- Réforme de 2014 : augmentation de la durée de la période de cotisation nécessaire pour un départ à taux plein.

Ces trois réformes ont eu un impact fort sur les projections d'âge de départ et de durée passée à la retraite pour les générations notamment comprises entre 1943 et 1990. En effet, sans réforme, trois quarts des gains d'espérance de vie auraient dû se traduire par du temps supplémentaire passé à la retraite. Avec les réformes, l'augmentation de la durée de retraite correspond à environ un tiers des gains d'espérance de vie, ce qui confirme ce qui avait été annoncé par la réforme de 2003. Cette hausse de durée passée à la retraite s'explique par la compensation de l'allongement de l'espérance de vie avec le décalage de l'âge de départ à la retraite. L'allongement seul de l'espérance de vie aurait eu pour conséquence une durée de retraite augmentée d'un peu plus de la moitié des gains d'espérance de vie.

Dans les statistiques qui vont suivre, nous appelons "âge de départ" l'âge atteint pour un départ à la retraite à taux plein. Par ailleurs, nous considérons quatre scénarios que nous comparerons :

- Un scénario "Contrefactuel" : scénario de référence (sans réforme) avec comme paramètres une durée requise pour le taux plein de 37.5 années, un âge minimum d'ouverture des droits à 60 ans et un âge maximum d'annulation de décote à 65 ans.
- Un scénario "2003 + Part EV" : scénario tenant compte de la législation mise en place par la réforme de 2003 avec toutefois des projections qui vont au-delà de 2020 pour pouvoir étudier les populations issues des générations les plus jeunes.
- Un scénario "2014 + Part EV" : scénario tenant compte du partage entre le gain d'espérance de vie et la durée de retraite de la réforme de 2003, mais également des modifications mises en place par la réforme de 2010.
- Un scénario "Leg. 2014" : autre scénario de référence qui s'appuie sur la législation actuelle (en date de l'année 2014, année dont les statistiques seront issues), avec l'application de la réforme de 2014.

Les projections d'espérance de vie à 60 ans⁸ et d'âge de départ à la retraite permettent le calcul de la durée moyenne passée en retraite par génération :

$$\text{Durée de retraite} = \text{Espérance de vie à 60 ans}_n - (\text{Âge de départ à la retraite} - 60)$$

6. Des compléments sur chacune de ces trois réformes sont explicités en Annexe C

7. Exposé des motivations de la loi n°2003-775 du 21 août 2003 portant réforme des retraites.

8. Projections selon la table INSEE prospective 2007 - 2060

avec n la génération considérée.

Il convient de noter que ce calcul suppose une négligence de la mortalité entre 60 ans et l'âge de départ à la retraite.

La durée de retraite reflète donc à la fois l'augmentation de l'espérance de vie à 60 ans et le décalage de l'âge de départ à la retraite. Cela signifie que lorsque l'espérance de vie à 60 ans augmente moins que l'âge de départ, la durée passée à la retraite diminue, et inversement.

Les âges de liquidation par génération sont représentés pour chacun des quatre scénarios sur les graphiques suivants⁹ :

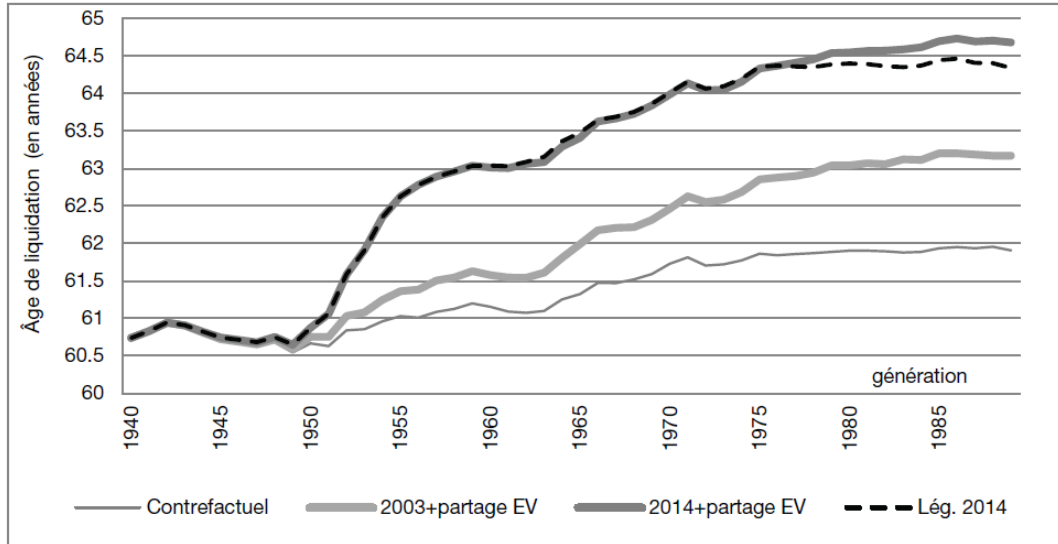


FIGURE 2.8 – Âges moyens de départ à la retraite - Ensemble de la population

Nous constatons que les générations à partir de 1950, et plus particulièrement les générations comprises entre 1950 et 1955, sont assez fortement impactées par la réforme de 2010, et ce quelque soit le sexe des populations.

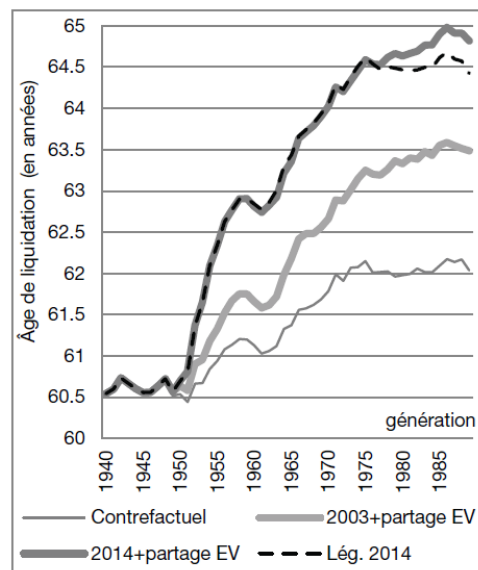
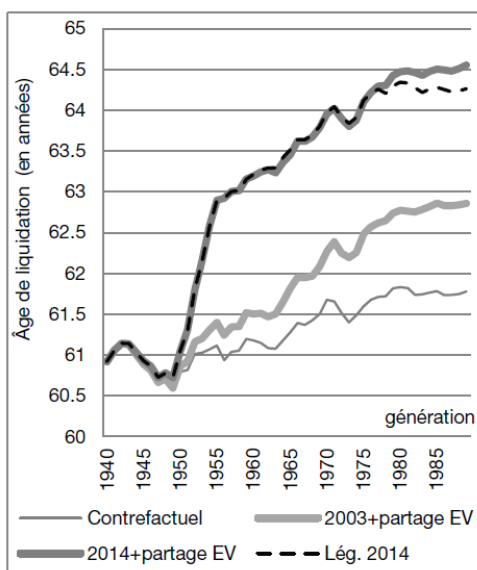


FIGURE 2.9 – Âges moyens de départ à la retraite - Femmes
 FIGURE 2.10 – Âges moyens de départ à la retraite - Hommes

9. L'ensemble des graphiques de cette rubrique sont issus du modèle de microsimulation *Destinie 2* de l'INSEE

2.2. UTILISATION DES TABLES DE MORTALITÉ DANS LES RÉGIMES DE RETRAITE

En effet, prenons par exemple la génération 1965 chez les femmes, l'âge de liquidation moyen est de 61.8 années pour le scénario "2003 + Part EV" contre 63.6 années pour le scénario "2014 + Part EV", une différence de +1.8 année imposée par la réforme de 2010. Il en est de même pour les hommes avec une augmentation de 1.2 année (+1.5 au global sur l'ensemble de la population).

Regardons maintenant les évolutions de durée de retraite par génération :

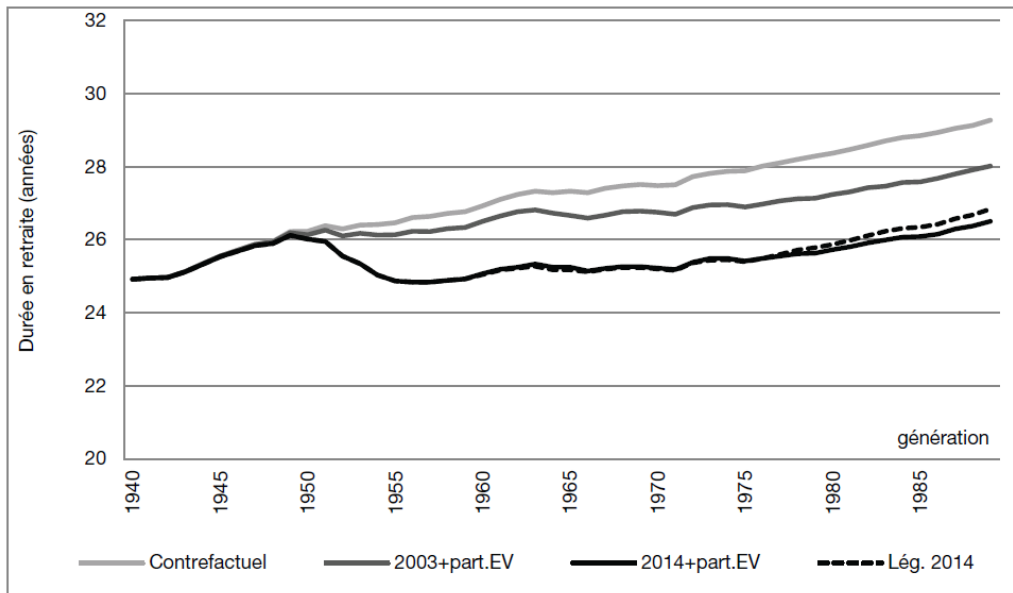


FIGURE 2.11 – Durée moyenne passée en retraite par génération - Ensemble de la population

Nous constatons une augmentation progressive de la durée passée en retraite pour le scénario "2003 + Part EV", en effet, une différence de +2.9 années est observée sur l'ensemble de la population entre les générations 1943 et 1949.

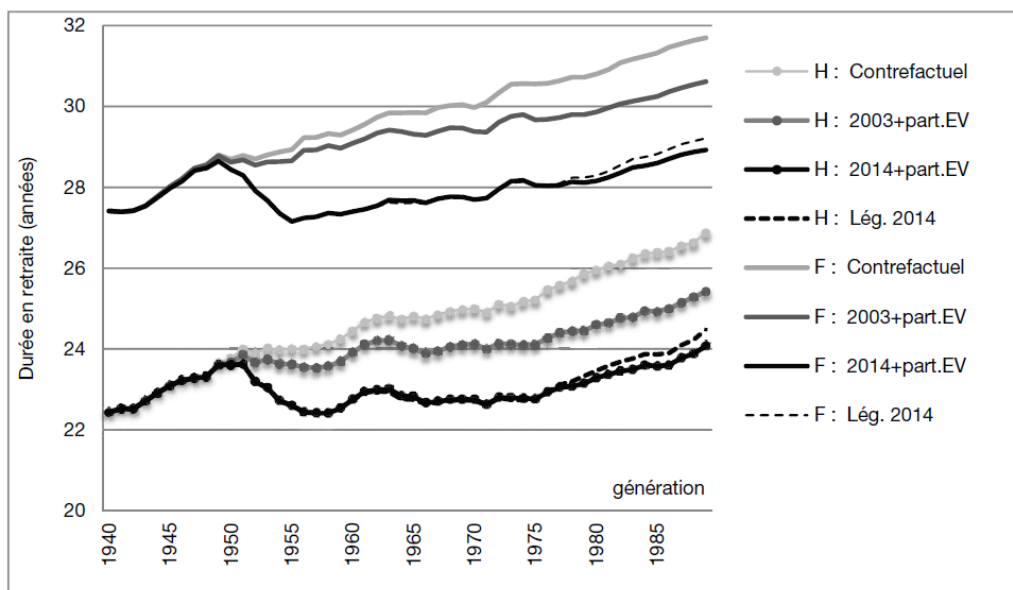


FIGURE 2.12 – Durée moyenne passée en retraite par génération - Par sexe

Cette même évolution est observée pour chacun des sexes : une différence de +3.1 années est constatée chez les femmes contre une différence de +2.7 années chez les hommes entre les générations 1943 et 1989.

2.2. UTILISATION DES TABLES DE MORTALITÉ DANS LES RÉGIMES DE RETRAITE

Nous constatons par ailleurs, que la réforme de 2010 défavorise, en termes de durée de retraite, les générations fortement impactées par la hausse des âges de départ à la retraite par rapport aux générations qui les précèdent. De manière générale, cette réforme limite beaucoup l'allongement de la durée de retraite, seulement une différence de +1.4 année est constatée entre les générations 1943 et 1989.

Comparons l'évolution de la durée de retraite par rapport à l'évolution de l'espérance de vie à l'aide du tableau suivant¹⁰ :

Scénario	Population	Variation de la durée de retraite (a)	Variation de l'espérance de vie (b)	Rapport (a)/(b)
Contrefactuel	Tous	+4.2 ans	+5.2 ans	81%
	Femmes	+4.1 ans	+4.8 ans	86%
	Hommes	+4.1 ans	+5.5 ans	75%
2003 + Part EV	Tous	+2.9 ans	+5.2 ans	56%
	Femmes	+3.1 ans	+4.8 ans	64%
	Hommes	+2.7 ans	+5.5 ans	49%
2014 + Part EV	Tous	+1.4 ans	+5.2 ans	27%
	Femmes	+1.4 ans	+4.8 ans	29%
	Hommes	+1.4 ans	+5.5 ans	25%
Lég. 2014	Tous	+1.7 ans	+5.2 ans	33%
	Femmes	+1.7 ans	+4.8 ans	35%
	Hommes	+1.8 ans	+5.5 ans	32%

TABLE 2.5 – Évolution de l'espérance de vie à 60 ans et de la durée de retraite entre les générations 1943 et 1989

Il est possible de constater qu'après la réforme de 2003 (scénario "2003 + Part EV"), les gains en durée de retraite dépassent le tiers des gains d'espérance de vie (environ 56% sur l'ensemble de la population); que suite à la réforme de 2010 (scénario "2014 + Part EV"), les gains en durée de retraite deviennent inférieurs au tiers des gains d'espérance de vie (environ 27% sur l'ensemble de la population); mais que la législation de 2014 (scénario "Lég. 2014") permet de rétablir l'équilibre entre les gains en durée de retraite et le tiers des gains en espérance de vie. Par ailleurs, nous pouvons remarquer que les femmes bénéficient d'un partage plus avantageux quel que soit le scénario considéré.

Toutefois, il est à noter que ces résultats ne sont valables qu'en comparaison des deux générations 1943 et 1989, et que les résultats seraient tout à fait différents si nous comparions d'autres générations, et notamment des générations dont le départ à la retraite s'est effectué après l'année 2003.

Regardons alors une représentation de l'évolution relative de la durée de retraite et de l'espérance de vie (cf. FIGURE 2.13 et FIGURE 2.14). L'espérance de vie sur ces graphique est à 20 ans, mais approximée avec celle à 60 ans à laquelle sont ajoutées 40 années (ceci suppose une certaine négligence de la mortalité entre 20 ans et 60 ans); ce choix est dû au fait que 20 ans est un âge moyen d'entrée sur le marché du travail.

Nous constatons que la part de l'espérance de vie passée à la retraite est d'un niveau initial de 38% au global pour chacun des quatre scénarios (40% pour les femmes contre un peu moins de 36% chez les hommes), ce qui est légèrement supérieur à un tiers. La réforme de 2003 entraîne un partage plus favorable aux générations les plus récentes, en effet, la part d'espérance de vie passée à la retraite pour le scénario "2003 + Part EV" augmente progressivement au cours du temps, se rapprochant de 40% pour la population totale. En ce qui concerne le scénario "2014 + Part EV", nous observons une baisse du rapport pour les générations les plus touchées par la hausse des âges de liquidation due à la réforme de 2010, pour atteindre finalement une stabilisation autour de 37% à partir de la génération 1955. Ces évolutions sont observables pour l'ensemble de la population, que ce soit pour les femmes ou pour les hommes.

10. Données issues du modèle de microsimulation *Destinie 2* de l'INSEE

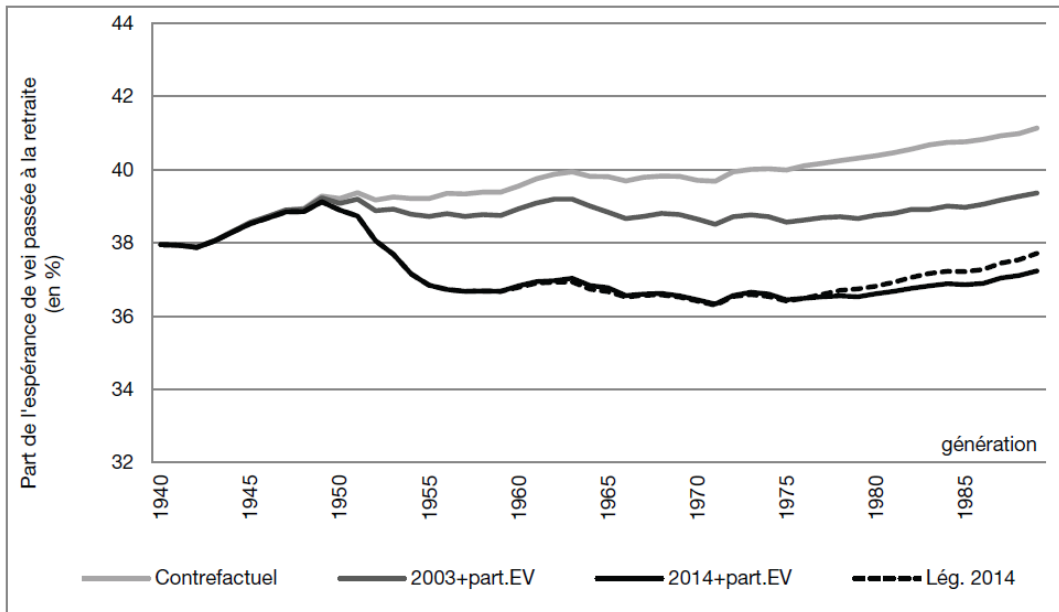


FIGURE 2.13 – Rapport entre la durée de retraite et l'espérance de vie par génération

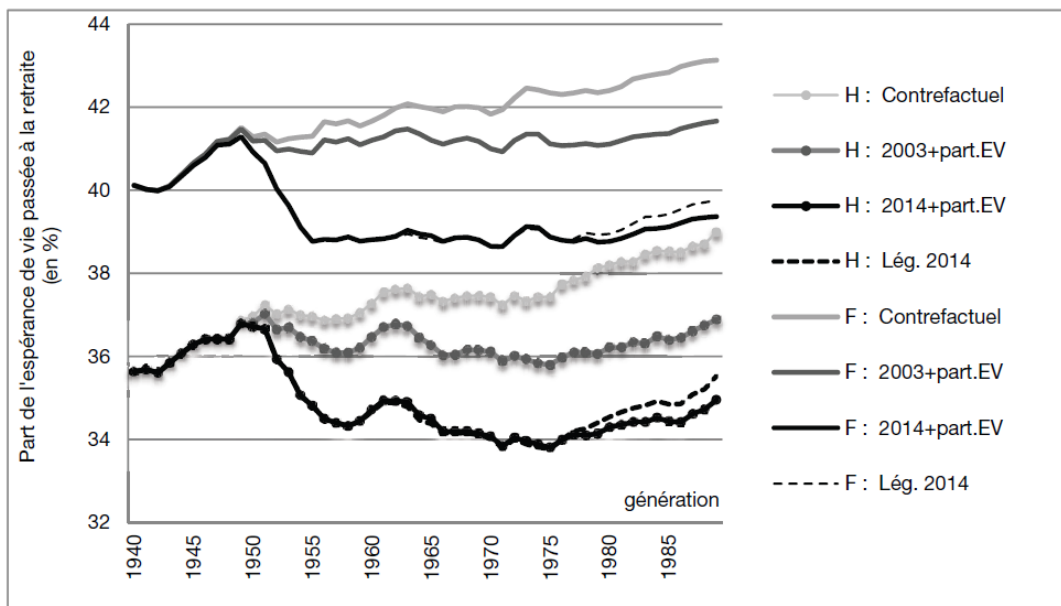


FIGURE 2.14 – Rapport entre la durée de retraite et l'espérance de vie par génération et par sexe

Les générations futures auraient été grandement favorisées au vu du scénario "Contrefactuel", mais la mise en place des réformes a apporté des modifications à cette situation. Suite à ces réformes, les générations les plus favorisées sont celles des années 1940 en termes de durée de retraite par rapport à l'espérance de vie, ce qui vient à menacer l'équilibre financier des systèmes de retraite car ces générations sont celles du *baby-boom* et dont la liquidation a déjà bien avancé.

La longévité s'approprie alors une place particulière au sein des différents facteurs qui sont à l'origine d'un accroissement des besoins de financement pour les régimes de retraite. En effet, l'allongement progressif et régulier de l'espérance de vie contribue à l'augmentation du nombre de personnes retraitées, à l'allongement du temps passé en retraite, et par conséquent à l'augmentation du montant des charges de pensions. Ainsi, des coefficients dits "de longévité" destinés à adapter les montants de pensions liqui-

dées ou à augmenter la durée de la période de cotisation, sont institués par les réformes des régimes de retraite par répartition afin d'intégrer les gains d'espérance de vie dans les calculs.

Par ailleurs, les régimes d'épargne-retraite, les régimes individuels, ainsi que les régimes collectifs, doivent également faire face à ce risque de longévité. Le calcul de provisionnement des engagements, de même que la tarification des annuités et rentes à verser, doivent dorénavant tenir compte des gains d'espérance de vie futurs afin d'éviter d'occasionner des pertes pour ces régimes qui pourraient avoir pour conséquence de les amener à reconsidérer leurs promesses de rentes. Ainsi, une recherche d'actifs financiers en mesure de proposer une protection contre ce risque de longévité est en cours par les politiques de placement des provisions.

2.2.1.2 Intervention des tables de mortalité

Suite aux différentes réformes parues au cours de cette dernière décennie, le pilotage des régimes publics de retraite par répartition nécessite de plus en plus de techniques de projections élaborées et approfondies qui permettraient de déterminer avec précision la participation des différents éléments responsables du besoin de financement des régimes.

Dans ce cadre, afin d'être en mesure d'établir une évaluation précise de la situation financière à long terme du régime, le choix d'utilisation de tables de mortalité adaptées pour la projection des populations des régimes devient une étape incontournable.

D'ordinaire, les tables de mortalité de la population d'un régime de retraite peuvent être comparées à deux types de tables :

- **Les tables assurantielles générationnelles ou instantanées** : généralement utilisées par les entreprises d'assurance pour la tarification et le provisionnement des contrats de rentes viagères immédiates ou différées, tels que les régimes supplémentaires facultatifs de retraite en capitalisation.
- **Les tables de mortalité prospectives ou passées** (par génération et par année), construites et rendues publiques par l'INSEE. Ces tables sont établies pour chaque âge atteint dans l'année et permettent de reconstituer la mortalité des générations. Elles sont principalement utilisées pour les régimes de retraite par répartition.

Dans le cas où les régimes ont été conçus sur une base professionnelle dans les systèmes de retraite, il est récurrent de constater que la mortalité des affiliés à un régime de retraite peut s'écarter de celle représentée par les tables de mortalité précédemment évoquées, et ce de façon significative.

Les catégories socio-professionnelles présentes au sein de la population considérée peuvent être des facteurs illustratifs des écarts traduits par les tables. D'autre part, la taille de la population influe sur la régularité des tables de ces dernières. Les écarts constatés peuvent se traduire aussi bien au niveau de l'espérance de vie qu'au niveau de son évolution dans le temps.

Par conséquent, afin d'améliorer la qualité des projections de sorte à minimiser le risque de sous-évaluation des charges futures, il est fortement préconisé d'utiliser des tables représentatives de la population des affiliés du régime en question.

2.2.2 Construction de tables de mortalité propres aux régimes de retraite

Pour la construction de tables de mortalité propres aux régimes de retraite, plusieurs options s'offrent à nous.

En effet, premièrement, si nous disposons d'un grand nombre de données avec un grand nombre d'observations, les statistiques relatives à la population du régime en question (notamment le nombre d'entrées, le nombre de sorties, et le nombre de décès) enregistrées chaque année permettent de calculer des quotients de mortalité qui conduisent à l'élaboration de tables de mortalité d'expérience utilisables lors des projections d'exercices.

Ces tables requièrent toutefois d'être actualisées très régulièrement et nécessitent l'assurance d'une possession d'un grand nombre d'informations concernant les évolutions de population.

Une autre méthode consiste alors à comparer les quotients de mortalité bruts de la population du régime calculés sur une période de référence avec les quotients de mortalité bruts des tables de mortalité de référence, le but étant de rechercher un décalage d'années ou d'âges, ou encore un abattement à appliquer sur la table de référence qui permettrait de réduire les écarts au maximum.

Nous rappelons que les quotients de mortalité calculés sur la période de référence pour chaque cohorte correspondent au rapport suivant :

Pour une cohorte de génération $n - x - 1$ d'âge x l'année n :

$$q_x = \frac{\text{Décès enregistrés l'année } n}{\text{Effectif exposé au risque l'année } n}$$

Pour projeter les populations, les tables utilisées seront alors des tables de référence auxquelles nous aurons appliqué les décalages ou abattements observés.

L'objectif de ce mémoire est donc dans un premier temps de déterminer quels sont les décalages ou abattements qui ajustent le mieux la mortalité des tables de référence avec celle observée dans le régime.

Afin de construire la meilleure table adaptée à notre régime de retraite, plusieurs méthodes seront exploitées au cours de cette étude : une recherche de décalage d'âges sera établie en premier lieu, puis une recherche de taux d'abattement à appliquer par tranche d'âges sur la table de référence en utilisant des méthodes non paramétriques de lissage des taux bruts de mortalité, et enfin nous appliquerons les méthodes d'ajustement de tables de mortalité contenues dans le *package* **ELT** disponible sous *RStudio*, que nous détaillerons un peu plus loin dans ce rapport.

Notre période d'observation se déroulant sur les années 2013 à 2017, l'ensemble des ajustements exploités seront réalisés à partir de la table **INSEE prospective 2013-2070**.

Chapitre 3

Les données utilisées

Actuellement, le régime utilise pour ses calculs, une loi de mortalité établie en 2012 à partir d'une étude de la population constituant le régime entre les années 2007 et 2011, ceci en appliquant des coefficients d'abattement sur la table de référence INSEE 2007-2060¹

Rappelons que l'objectif de notre étude est de construire une nouvelle loi à partir des données de la table INSEE prospective 2013-2070 ainsi que des données propres à la population du régime sur les dernières années (2013 à 2017), afin de les comparer avec la loi actuelle de sorte à choisir celle qui semble la mieux adaptée au risque viager de notre population².

Nous proposons alors, de sorte à introduire ce travail, de présenter les différentes données qui nous sont mises à disposition pour la construction de notre table de mortalité. Ensuite, nous étudierons quelques statistiques descriptives et démographiques réalisées sur nos affiliés qui nous permettront de faciliter l'analyse des résultats que nous obtiendrons au cours de notre étude.

Avant toute chose, il semble important d'introduire deux notions qui seront fondamentales tout au long de notre travail : les notions d'ouvrant droit et d'ayant droit.

Un **ouvrant droit** (noté OD par la suite) correspond à un assuré de droit direct. Son affiliation à un régime lui permet de percevoir des droits au moment de son départ en retraite, mais permet également d'ouvrir des droits à ses bénéficiaires (conjoint, enfants,...) en cas de décès. Ces bénéficiaires sont alors qualifiés d'**ayants droit** (AD par la suite).

3.1 Présentation des données

Pour la réalisation de notre travail, nous disposons de différents types de données : les données de référence issues des tables INSEE, les données propres à la population des affiliés actifs du régime, ainsi que les données propres à la population des affiliés inactifs du régime.

Comme données de référence, nous disposons des tables INSEE prospectives 2013-2070 pour chaque sexe, qui correspondent aux tables de référence sur lesquelles notre travail va s'appuyer et que nous allons ajuster. Nous disposons également des tables INSEE prospectives 2007-2060 pour chaque sexe, ainsi que de ces mêmes tables ajustées par des coefficients d'abattement déterminés à partir de l'étude de la population réalisée en 2012. Il s'agit de la table actuellement utilisée par le régime.

Concernant les données propres à la population des inactifs du régime, nous disposons de cinq fichiers de données (un pour chaque année de la période d'observation) dans lesquels nous retrouvons

1. La méthode d'ajustement de la table de référence par abattement de celle-ci est décrite au cours du chapitre suivant (SECTION 4.1.3).

2. Notre population totale sera divisée en plusieurs "sous-populations" pour lesquelles une loi de mortalité différente sera établie pour chacune.

pour chaque année :

- Une base de données contenant l'ensemble des affiliés inactifs du régime enregistrés en date du 30 septembre de l'année considérée.
- Une base de données contenant les sorties des inactifs au cours de l'année considérée.

Notons que l'entrée en inactivité d'un ouvrant droit correspond à son départ à la retraite, et que sa sortie de la population des inactifs ne peut être due que pour cause de décès.

En ce qui concerne les ayants droit, ils sont nécessairement considérés comme inactifs. L'entrée de l'un d'entre eux au régime se fait au moment du décès de l'affilié qui lui a ouvert des droits, et la sortie de l'ayant droit est causée par le décès de celui-ci pour un conjoint ou l'atteinte de la limite d'âge pour les enfants.

D'autre part, en ce qui concerne les données propres à la population des actifs du régime, nous disposons également de cinq fichiers de données (un pour chaque année de la période d'observation) dans lesquels nous retrouvons pour chaque année :

- Une base de données contenant l'ensemble des affiliés actifs du régime enregistrés en date du 30 septembre de l'année considérée.
- Une base de données contenant les nouveaux actifs entrés dans le régime au cours de l'année considérée.
- Une base de données contenant les sorties des actifs au cours de l'année considérée.

Notons que l'entrée au régime d'un individu se fait au moment de son embauche, et que sa sortie en tant qu'actif est due soit à son départ en retraite soit à son décès (il y a très peu de sorties pour une autre cause telle que le changement d'emploi).

Si un individu est compris dans la base des sorties d'actifs d'une année mais qu'il n'est pas compris dans la base des inactifs enregistrés la même année, alors cet individu est considéré comme décédé.

Pour terminer, nous avons à notre disposition une dernière base de données contenant les dates de décès des personnes affiliées sorties du régime. Ce fichier nous sera utile pour vérifier les dates de sortie des individus décédés.

Toutes ces bases de données contiennent un certain nombre de variables dont seulement quelques unes d'entre elles nous seront utiles pour notre étude. L'objectif est de travailler sur des bases de données plus "légères" qui ne contiennent que les variables utiles. Nous retenons parmi elles :

- **Matricule** : identifiant d'un individu au sein du régime, il permet de croiser les données entre elles.
- **Sexe** : femme ou homme, la population sera divisée par sexe lors de la création des tables de mortalité.
- **Date de naissance** : permet de déterminer la génération d'un individu et l'âge qu'il a au moment de son entrée et de sa sortie du régime.
- **Type de prestation** : ouvrant droit, veuf/veuve ou ex-époux/ex-épouse non remarié/remariée (ayants droit), cette variable nous permettra de séparer les ouvrants droit des ayants droit lors de la construction des tables de mortalité.
- **Date d'embauche** : lorsqu'elle est renseignée, elle nous donne la date d'entrée au régime de l'affilié considéré.
- **Date de cessation** : il s'agit de la date d'entrée d'un actif en inactivité, à noter que la cessation de toute activité professionnelle est une condition requise pour percevoir sa retraite à un régime de base.

Par des travaux réalisés antérieurement, il a été constaté que la mortalité des ayants droit différait de la mortalité des ouvrants droit, c'est pourquoi nous faisons le choix d'étudier les deux populations séparément³. La population des ouvrants droit que nous considérerons et que nous noterons "OD" dans

3. Le contenu des données n'est pas assez conséquent pour envisager un critère de séparation des individus autre que le

3.2. QUELQUES STATISTIQUES DESCRIPTIVES ET DÉMOGRAPHIQUES

toute la suite de notre étude regroupera l'ensemble des inactifs ouvrants droit et des actifs, et la population des ayants droit se composera des inactifs veufs et des inactifs ex-époux. Par la suite, nous noterons cette population "AD".

Pour la population des OD, nous étudions les actifs et les inactifs séparément puis rassemblés, ceci en veillant à séparer les ouvrants droit des ayants droit chez les inactifs et en vérifiant l'absence de doublons qui pourraient apparaître au moment de la "réunion" des deux populations. Le travail effectué sur chacune de ces populations est ensuite reproduit sur la population des ayants droit.

À partir de toutes ces données, nous avons la possibilité de relever le nombre d'entrées, le nombre de sorties, le nombre de décès et l'effectif sous risque par génération (donc par âge) et par sexe pour chacune des populations et pour chaque année de la période d'observation (qui pour rappel, s'étend entre 2013 et 2017). Ces statistiques sont détaillées dans la section suivante et permettent de calculer les taux de décès moyens sur la période d'observation à partir desquels nous pourrions ajuster nos tables de mortalité.

3.2 Quelques statistiques descriptives et démographiques

Nous consacrons cette section à quelques statistiques descriptives et démographiques de chacune de nos quatre populations étudiées (Femmes OD, Hommes OD, Femmes AD, Hommes AD)⁴.

Dans un premier temps, regardons la répartition de nos individus au sein de la population totale en terme de qualité d'OD ou d'AD :

Répartition des individus selon leur qualité

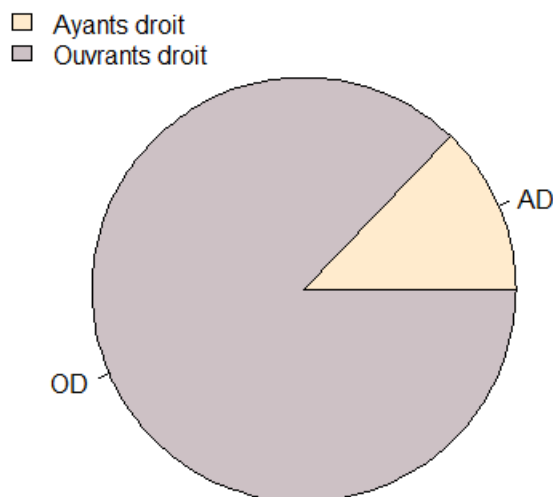


FIGURE 3.1 – Répartition de la population selon le type de prestation

Sans grande surprise, les ouvrants droit sont beaucoup plus présents que les ayants droit au sein de notre population totale (87% d'OD contre seulement 13% d'AD). Ceci s'explique pas le fait qu'un AD n'entre dans le régime qu'au moment du décès de son OD.

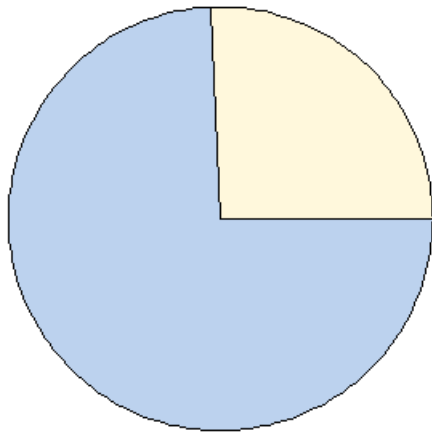
Regardons maintenant la répartition femmes/hommes au sein de chacune de nos deux populations distinguées ci dessus (AD et OD) sur les deux diagrammes ci-après.

sexe ou le type de prestation des affiliés.

4. Les parts en pourcentage de population selon différents critères sont disponibles en ANNEXE D

Répartition des OD selon leur sexe

■ Femmes
■ Hommes



Répartition des AD selon leur sexe

■ Femmes
■ Hommes

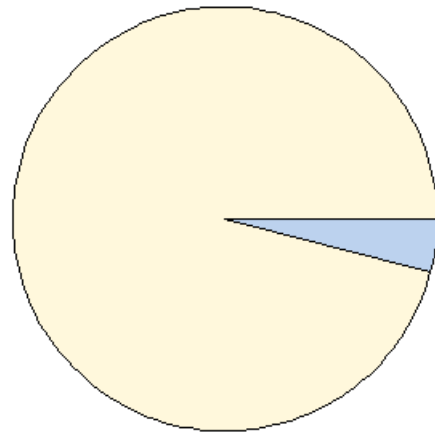


FIGURE 3.2 – Répartition des populations des OD et des AD selon leur sexe

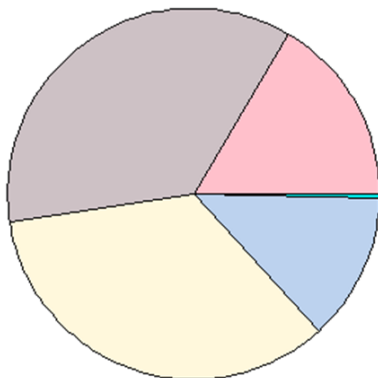
Nous constatons que les hommes représentent environ trois quarts des OD (contre un quart pour les femmes). Les ayants droits étant principalement composés de veufs ou de veuves, il semble donc normal de retrouver une grande majorité de femmes parmi les AD (96% de femmes contre seulement 4% d'hommes).

Analysons la répartition des individus des quatre populations distinguées (Femmes OD, Hommes OD, Femmes AD, Hommes AD) selon leurs âges. Il s'agit d'une répartition moyenne par tranches d'âges sur l'ensemble de la période d'observation.

Nous considérons les tranches d'âges suivantes : les âges inférieurs ou égaux à 35 ans (Tranche 1), les âges allant de 36 à 55 ans inclus (Tranche 2), les âges allant de 56 à 75 ans inclus (Tranche 3), les âges allant de 76 à 95 ans inclus (Tranche 4), et enfin les âges au-delà de 95 ans (Tranche 5).

Répartition des femmes OD par tranches d'âges

■ Tranche1
■ Tranche2
■ Tranche3
■ Tranche4
■ Tranche5



Répartition des hommes OD par tranches d'âges

■ Tranche1
■ Tranche2
■ Tranche3
■ Tranche4
■ Tranche5

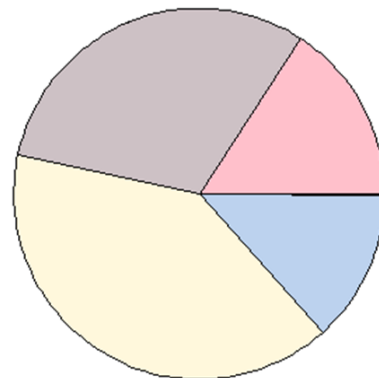


FIGURE 3.3 – Répartition de la population des ouvrants droit selon leur âge

La répartition est similaire chez les femmes et chez les hommes OD. Les tranches 2 et 3 sont les plus importantes, tandis que la Tranche 5 n'est quasiment pas représentée. Ceci est peu étonnant du fait que les espérances de vie des femmes et des hommes sont inférieures à 95 ans de manière générale, les individus âgés de plus de 95 ans sont conséquemment très peu nombreux par rapport au reste de la population.

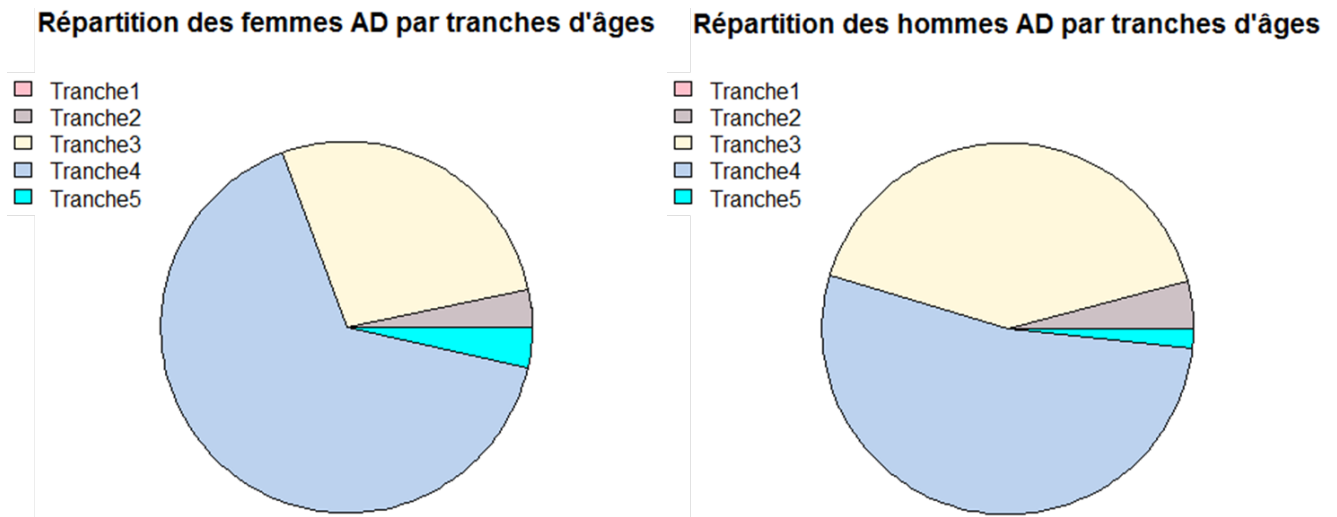


FIGURE 3.4 – Répartition de la population des ayants droit selon leur âge

La répartition est similaire chez les femmes et chez les hommes AD : la Tranche 4 (76-95 ans) est de loin la plus représentée, suivie de la Tranche 3 (56-75 ans). Les proportions des trois autres tranches sont très faibles, voire presque nulle pour la Tranche 1. L'entrée des ayants droit au régime ayant lieu au moment du décès du conjoint pour un veuf ou de l'ex-conjoint pour un ex-époux, accompagnée du fait que les décès (sauf accidentels) et départs en retraite surviennent généralement après 55 ans, il y a par conséquent peu d'ayants droit dans les tranches d'âges faibles. La tranche 5 est peu représentée pour les mêmes raisons que chez les OD.

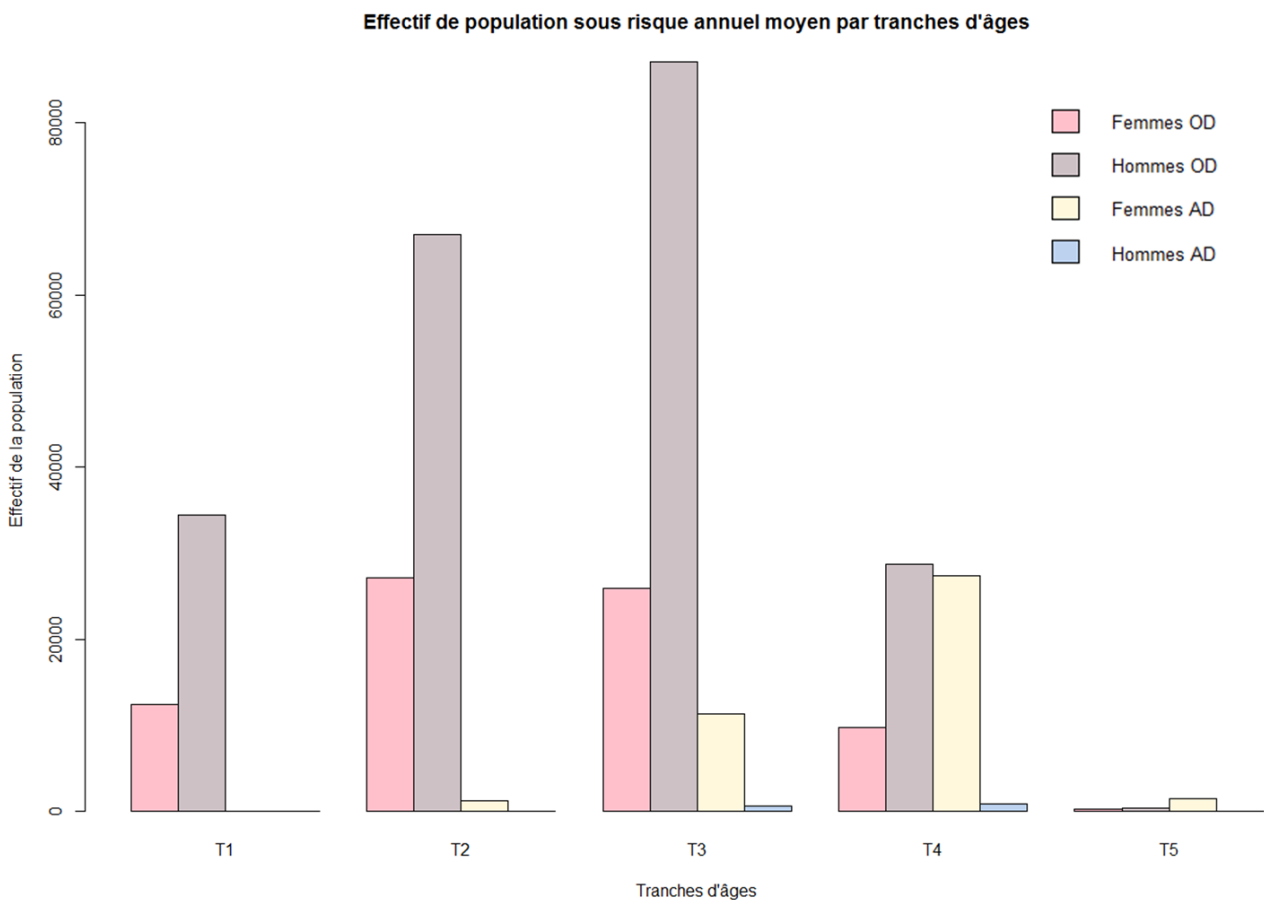


FIGURE 3.5 – Effectifs annuels moyens de population par tranche d'âges sur la période d'observation

3.2. QUELQUES STATISTIQUES DESCRIPTIVES ET DÉMOGRAPHIQUES

Sur la FIGURE 3.5 située en bas de page précédente et représentant les effectifs annuels moyens de population sur la période d'observation par tranche d'âges, nous retrouvons les ouvrants droit majoritaires sur les quatre premières tranches d'âges. Nous constatons par ailleurs, que les femmes AD sont très présentes sur la Tranche 4 et majoritaires sur la dernière tranche. Ceci peut s'expliquer du fait que les femmes AD entrent au régime au moment du décès d'hommes OD (de manière générale), qui sont majoritaires sur les autres tranches d'âges. Par ailleurs, l'espérance de vie des femmes étant plus longue que celle des hommes, il apparaît peu surprenant de retrouver les femmes majoritaires sur des âges très élevés.

La figure qui suit représente le nombre de décès observé par année d'observation selon chacune des quatre populations :

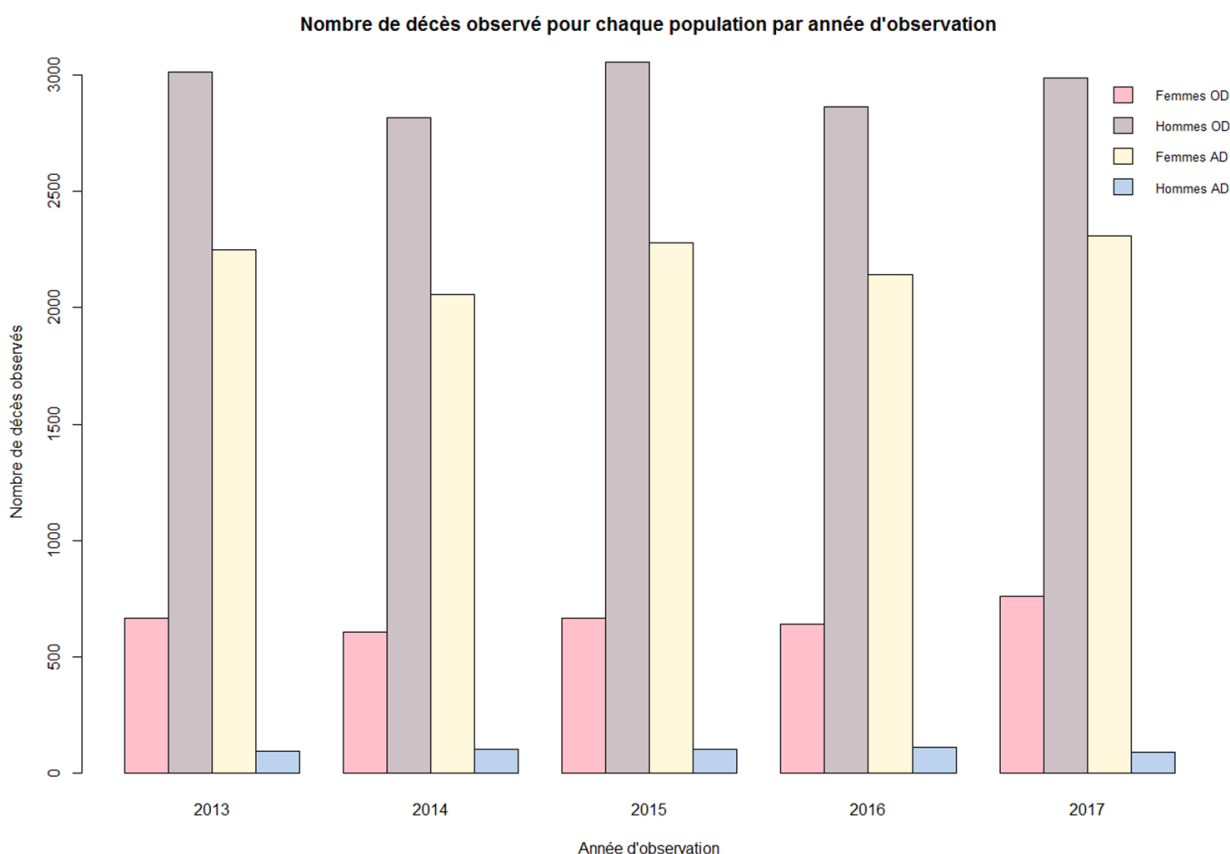


FIGURE 3.6 – Décès observés par année de la période d'observation selon la qualité et le sexe des individus

Nous pouvons constater une majorité de décès chez les hommes OD, quelle que soit l'année de la période d'observation considérée. L'analyse précédente justifie ce constat. En effet, la population des hommes OD étant très nombreuse par rapport aux trois autres, il apparaît normal que nous retrouvions cette même population en nombre de décès le plus important.

Nous remarquons d'autre part, que les femmes AD se situent juste derrière les hommes OD en nombre de décès par an, ce qui peut sembler troublant étant donné la faible densité de cette population. Ce constat peut toutefois se justifier du fait de la "vieillesse" de la population : les femmes AD étant principalement âgées de plus de 55 ans, elles sont plus sujettes au risque de décès que les femmes OD. Le nombre d'individus hommes AD est trop faible pour que nous puissions établir les mêmes remarques sur cette population.

La FIGURE 3.7 disposée juste ci-après représente le nombre de décès total observé par tranche d'âges sur la période d'observation selon la population des individus.

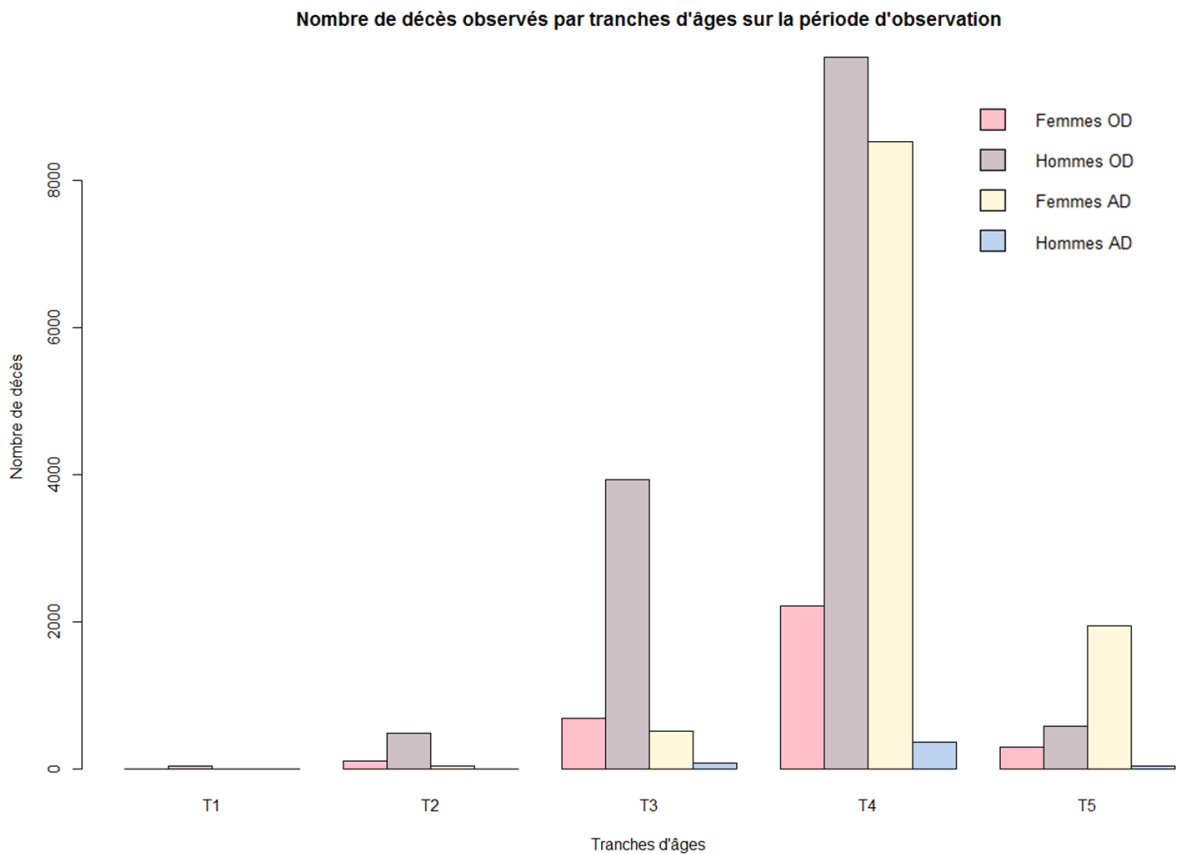


FIGURE 3.7 – Décès observés par tranche d'âges selon la qualité et le sexe des individus sur la période d'observation

Pour les quatre populations, la quatrième tranche d'âges est celle qui présente le plus de décès enregistrés sur la période d'observation. Chez les AD, il s'agit de la tranche la plus nombreuse en terme d'effectif, c'est pourquoi nous pouvons établir ce constat. Toutefois, un facteur pertinent pour chacune des populations est à souligner pour expliquer la forte mortalité sur la Tranche 4 : cette plage d'âges contient les individus âgés de plus de 75 ans, ce qui fait que la population est sujette à un risque de décès beaucoup plus important que les tranches d'âges plus jeunes. Les personnes âgées de plus de 95 ans (Tranche 5) sont soumis à un risque de décès encore plus important, mais trop peu d'individus composent cette tranche pour observer un nombre de décès plus important.

Par ailleurs, il est possible de constater un nombre de décès particulièrement élevé chez les hommes OD sur la Tranche 3. Ceci est une conséquence de la présence majoritaire d'hommes OD sur cette plage d'âges.

L'ensemble de ces statistiques ont montré que les observations relatives aux individus âgés de plus de 95 ans sont peu significatives du fait de leur représentation extrêmement faible au sein du régime. Par conséquent, nous envisageons de ne considérer que les âges inférieurs à 95 ans dans la construction de notre table de mortalité afin de représenter au mieux la réalité.

Les figures qui vont suivre représentent les taux de décès moyens observés sur la période d'observation. Il s'agit des taux à partir desquels notre étude va se construire.

Les FIGURES 3.8 et 3.9 sont propres à la population des ouvrants droit, tandis que les FIGURES 3.10 et 3.11 sont propres à la population des ayants droit.

Sur les deux premières figures, nous pouvons distinguer une croissance exponentielle jusqu'à l'âge de 95 ans, ce qui ne semble pas présenter d'aberration car le risque décès et donc la mortalité, croît exponentiellement avec l'âge.

3.2. QUELQUES STATISTIQUES DESCRIPTIVES ET DÉMOGRAPHIQUES

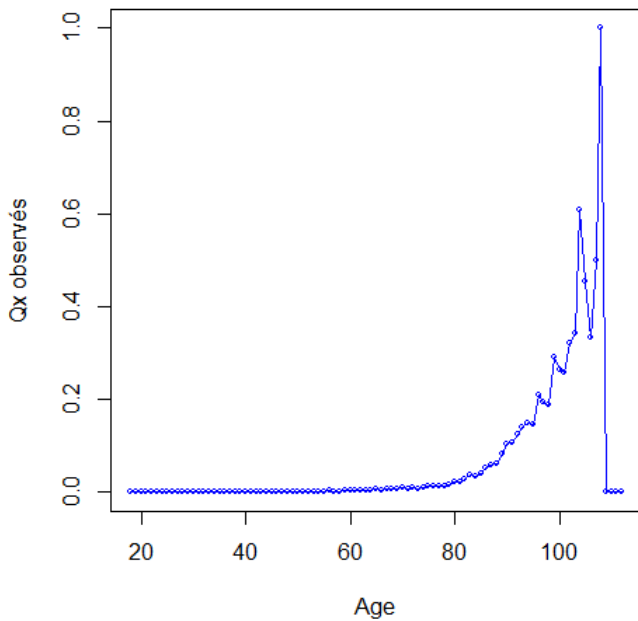


FIGURE 3.8 – Qx Observés - Femmes OD

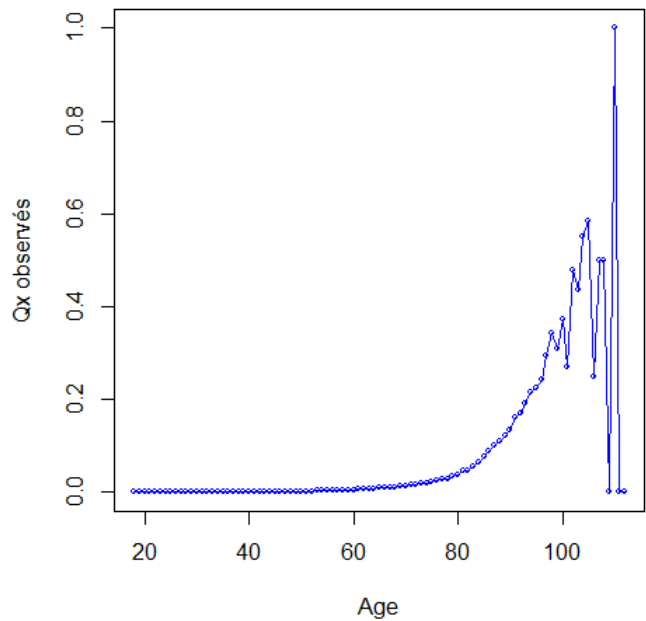


FIGURE 3.9 – Qx Observés - Hommes OD

Au-delà de 95 ans, de fortes fluctuations de mortalité sont observables à cause de la population très peu représentée à ces âges. Cette tranche à partir de 95 ans semble donc très peu significative et pourrait causer des écarts d'ajustement importants au moment de notre construction de table.

Chez les AD, de la même manière que chez les OD, une croissance exponentielle se distingue sur les âges inférieurs à 95 ans. De plus nombreuses fluctuations peuvent être remarquées pour la population des hommes AD du fait qu'ils sont moins représentés. Les mêmes constats que pour les OD sont effectués sur les âges élevés (au-delà de 95 ans) chez les AD.

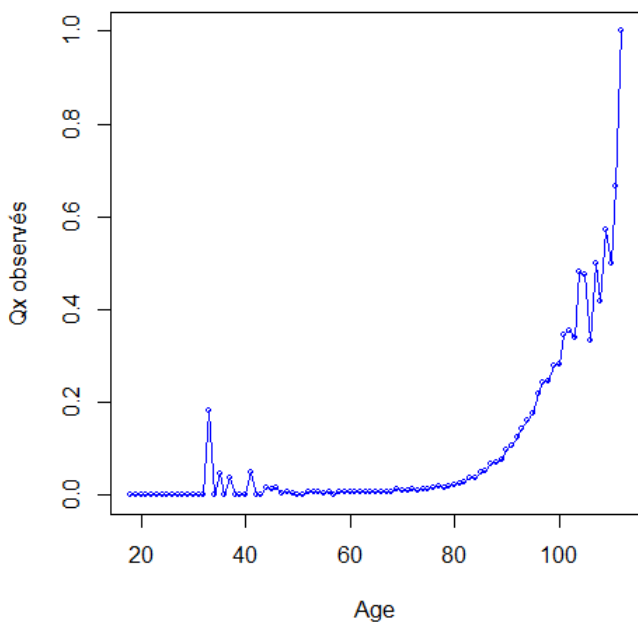


FIGURE 3.10 – Qx Observés - Femmes AD

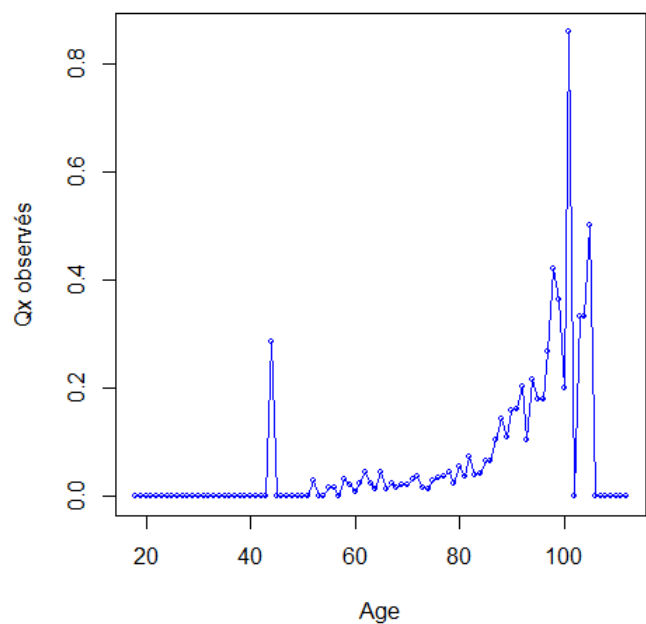


FIGURE 3.11 – Qx Observés - Hommes AD

Aux âges inférieurs à 50 ans, des pics de mortalité se distinguent, que ce soit chez les femmes ou chez les hommes. Il s'agit d'une mortalité accidentelle survenue aux bas âges où la population est faible. Cette

3.2. QUELQUES STATISTIQUES DESCRIPTIVES ET DÉMOGRAPHIQUES

mortalité est considérée comme aberrante dans notre étude, ces points seront donc supprimés dans la suite pour éviter d'éventuelles aberrations dans les résultats lors de la construction des tables.

Enfin, la FIGURE 3.12 permet de comparer les mortalité de nos quatre populations entre elles :

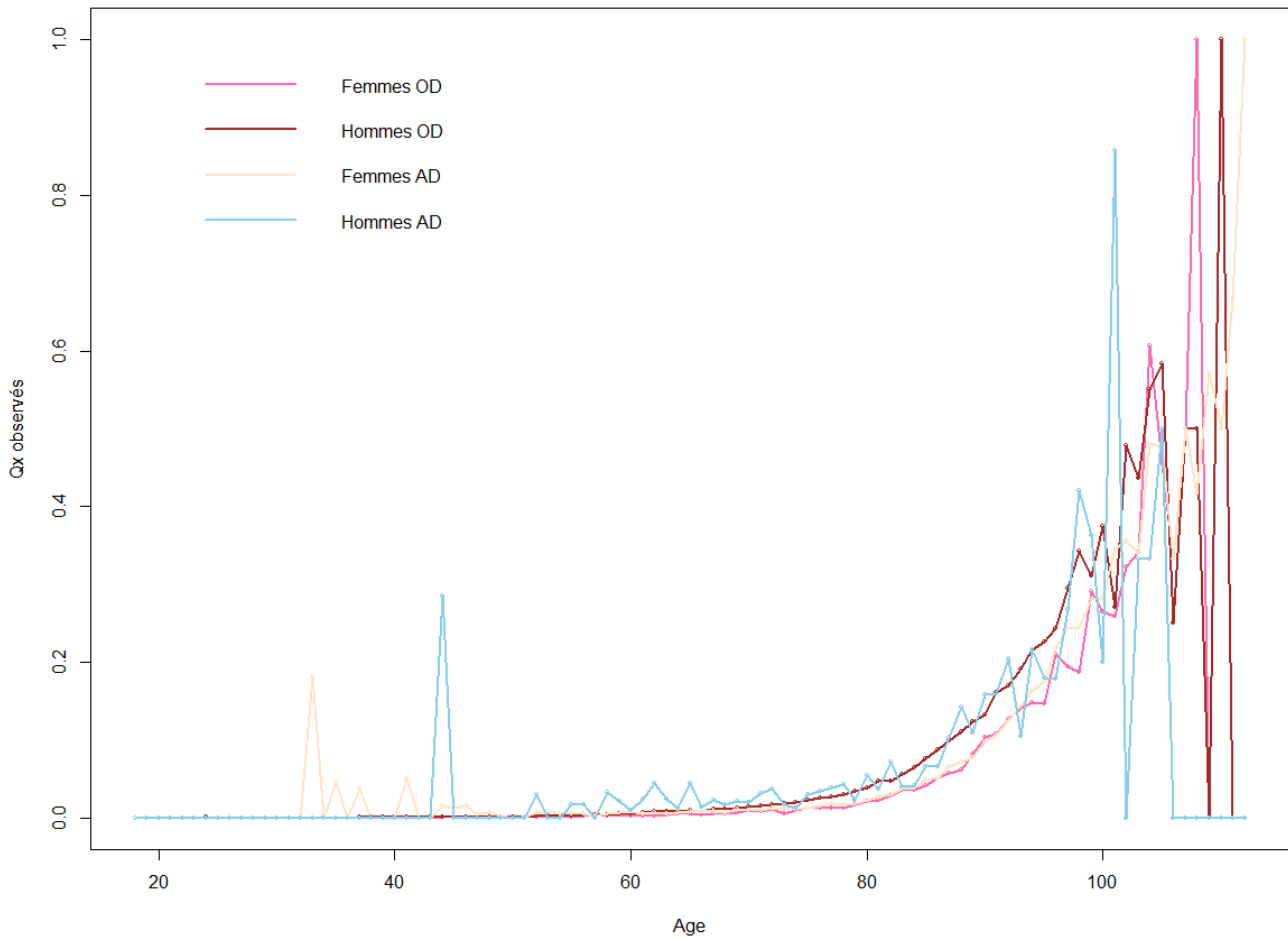


FIGURE 3.12 – Qx moyens observés par population sur la période d'observation

Ce graphique révèle une mortalité très basse aux âges inférieurs à 75 ans. Les pics de mortalité des AD aux bas âges semblent d'autant plus se révéler aberrants quand nous les comparons aux autres populations.

En ce qui concerne les âges compris entre 75 et 95 ans, nous distinguons une surmortalité des populations masculines par rapport aux populations féminines, ceci du fait que les femmes ont une espérance de vie plus longue que les hommes.

Enfin, de nouveau, pour les âges compris au-delà de 95 ans, il est difficile de distinguer une tendance de la mortalité, les fluctuations sont trop nombreuses et trop importantes. Ceci confirme alors notre choix de ne considérer que les âges inférieurs à 95 ans pour notre étude.

Chapitre 4

Création d'une table de mortalité propre à notre régime de retraite

4.1 Création d'une table de mortalité par des méthodes non-paramétriques

4.1.1 Les différentes méthodes de lissage utilisées

Afin de pallier les éventuelles aberrations qui pourraient être dues au fait que peu d'individus de notre population appartiennent aux classes d'âges extrêmes, ou encore à la survenance de décès accidentels aux jeunes âges qui pourraient nuire à un ajustement précis et réel de notre table, nous avons fait le choix d'appliquer plusieurs méthodes de lissage sur les données observées, de manière à sélectionner la meilleure méthode et donc les données lissées les plus proches de la réalité.

Compte tenu de la répartition de nos individus au sein de nos populations et les âges compris au-delà de 95 ans étant très peu significatifs, nous établissons nos lissages et notre table de mortalité à partir des données des plages 25-95 ans pour les OD, 27-95 ans pour les femmes AD et 35-95 ans pour les hommes AD.

Les différentes méthodes et résultats obtenus seront présentés tout au long de cette section, il s'agit de l'application d'une méthode d'ajustement par les moindres carrés avec maximisation du R^2 , de la méthode de Makeham, de la méthode de Loess, puis de la méthode de Whittaker-Henderson.

4.1.1.1 Lissage par la méthode des moindres carrés et maximisation du R^2

La première méthode de lissage consiste en une méthode assez peu technique qui est la recherche du coefficient de détermination maximal sur *Excel*, ayant recours à l'application de la méthode des moindres carrés pour la détermination des ajustements.

- Description de la méthode :

Lorsque nous disposons d'un nuage de points que nous souhaitons ajuster par une courbe de fonction mathématique, la méthode utilisée pour rechercher les coefficients de la fonction la plus pertinente au vu des données est la méthode des moindres carrés (méthode développée par Gauss¹ en 1809) qui consiste à minimiser la somme des carrés des distances entre les points de nos observations y_i et les points retournés par la fonction d'ajustement $f(x_i)$.

Le caractère à minimiser est alors :

$$\Delta = \sum_{i=1}^n (y_i - f(x_i))^2$$

1. Carl Friedrich GAUSS (1777-1855) : mathématicien, astronome et physicien allemand

avec :

- n : le nombre d'observations à ajuster
- x_i : les abscisses des points d'observation
- y_i : les ordonnées des points d'observation
- f : la fonction d'ajustement appliquée en chaque point x_i .

Cela revient alors à minimiser la somme des carrés des résidus ε_i :

$$\Delta = \sum_{i=1}^n \varepsilon_i^2 \quad \text{avec } \varepsilon_i = y_i - f(x_i)$$

Dans notre étude, les régressions effectuées sont soit polynomiales soit exponentielles. Nous faisons donc le choix de présenter uniquement ces deux types d'ajustement ici.

Regardons dans un premier temps le cas d'une régression polynomiale de degré p . Dans un tel cas, l'ajustement se fait par une fonction de la forme suivante :

$$f(x) = \alpha_0 + \alpha_1 x + \dots + \alpha_p x^p$$

avec p le degré du polynôme considéré, et $(\alpha_i)_{1 \leq i \leq p}$ les coefficients à rechercher.

Il faut alors chercher à minimiser :

$$\Delta = \sum_{i=1}^n (y_i - f(x_i))^2 = \sum_{i=1}^n \left[y_i - \sum_{j=0}^p \alpha_j x_i^j \right]^2$$

Nous supposons que les $p+1$ dérivées partielles de Δ au point $(\alpha_0, \dots, \alpha_p)$ s'annulent en ce point.

Nous avons alors $\forall 1 \leq j \leq p$:

$$\begin{aligned} \frac{\partial \Delta}{\partial \alpha_j} &= \sum_{i=1}^n \frac{\partial}{\partial \alpha_j} \left[\left(y_i - \sum_{j=0}^p \alpha_j x_i^j \right)^2 \right] \\ &= \sum_{i=1}^n 2 \left(y_i - \sum_{j=0}^p \alpha_j x_i^j \right) \frac{\partial}{\partial \alpha_j} \left[\left(y_i - \sum_{j=0}^p \alpha_j x_i^j \right) \right] \\ &= 2 \sum_{i=1}^n \left[(y_i - f(x_i)) \times (-x_i^j) \right] = 0 \end{aligned}$$

Ce qui fait apparaître un système de $p+1$ équations de la forme :

$$\forall 1 \leq j \leq p, \quad \sum_{i=1}^n x_i^j f(x_i) = \sum_{i=1}^n y_i x_i^j$$

En simplifiant :

$$S_j = \sum_{i=1}^n x_i^j \quad \text{et} \quad W_j = \sum_{i=1}^n y_i x_i^j$$

nous obtenons un système d'équation de la forme :

$$\begin{cases} n\alpha_0 + \alpha_1 S_1 + \dots + \alpha_p S_p = W_0, & j = 0 \\ \alpha_0 S_1 + \alpha_1 S_2 + \dots + \alpha_p S_{p+1} = W_1, & j = 1 \\ \dots \\ \alpha_0 S_p + \alpha_1 S_{p+1} + \dots + \alpha_p S_{2p} = W_p, & j = p \end{cases}$$

4.1. CRÉATION D'UNE TABLE DE MORTALITÉ PAR DES MÉTHODES NON-PARAMÉTRIQUES

La résolution de ce système (que nous constatons symétrique) conduit alors à la détermination de chacun des coefficients $(\alpha_0, \dots, \alpha_p)$ de l'ajustement polynomial.

Maintenant, regardons le cas d'une régression exponentielle. Dans un tel cas, l'ajustement se fait par une fonction de la forme suivante :

$$f(x) = \lambda e^{\alpha x}$$

avec λ et α les paramètres à déterminer.

L'ajustement exponentiel relève de l'ajustement linéaire des points ayant pour coordonnées $(x_i, \ln y_i)$, c'est-à-dire que comme la méthode des moindres carrés construit une droite de type $y = \alpha x + \beta$, nous aurons une équation de la forme :

$$\ln y = \alpha x + \beta \quad \text{donc : } y = \lambda e^{\alpha x} \quad \text{avec : } \lambda = e^\beta$$

Il est possible que plusieurs fonctions soient adaptées à l'ajustement d'un nuage de points, le meilleur ajustement est alors déterminé par le coefficient de détermination noté R^2 . Ce coefficient est associé à chacune des fonctions d'ajustement, il est compris entre 0 et 1 et estime la part de variation de nos observations expliquée par le modèle d'ajustement.

Sa valeur est déterminée par :

$$R^2 = 1 - \frac{\sum_x (\hat{q}_x - q_x)^2}{\sum_x \left(\hat{q}_x - \frac{\sum_x \hat{q}_x}{n} \right)^2}$$

avec :

- \hat{q}_x : les taux de mortalité observés à chaque âge x
- q_x : les taux de mortalité lissés à chaque âge x
- n : le nombre de taux à lisser

Notons que plus ce coefficient se rapproche de 1, plus le modèle est en adéquation avec nos observations, a contrario, plus il se rapproche de 0, moins le modèle est bon.

- Application à nos observations :

Les observations dont nous disposons ont une répartition très irrégulière, c'est pourquoi nous appliquons cette méthode sur différentes tranches d'âges de taille réduite pour ensuite extrapoler les résultats aux âges extrêmes.

- Cas des actifs et des inactifs ouvrants droit :

Pour les femmes de la catégorie OD, l'équation considérée sur la plage d'âges 25-95 ans est la relation exponentielle suivante :

$$y = 3 \times 10^{-6} \times e^{0.1123x}$$

avec x correspondant à l'âge.

Le coefficient de détermination obtenu avec cette équation de lissage pour la population des femmes OD a une valeur de 0.975.

Sur le graphe suivant, les points **bleus** correspondent aux taux bruts, et la courbe **rouge** correspond au lissage des taux par la méthode des moindres carrés dans le cas des femmes OD.

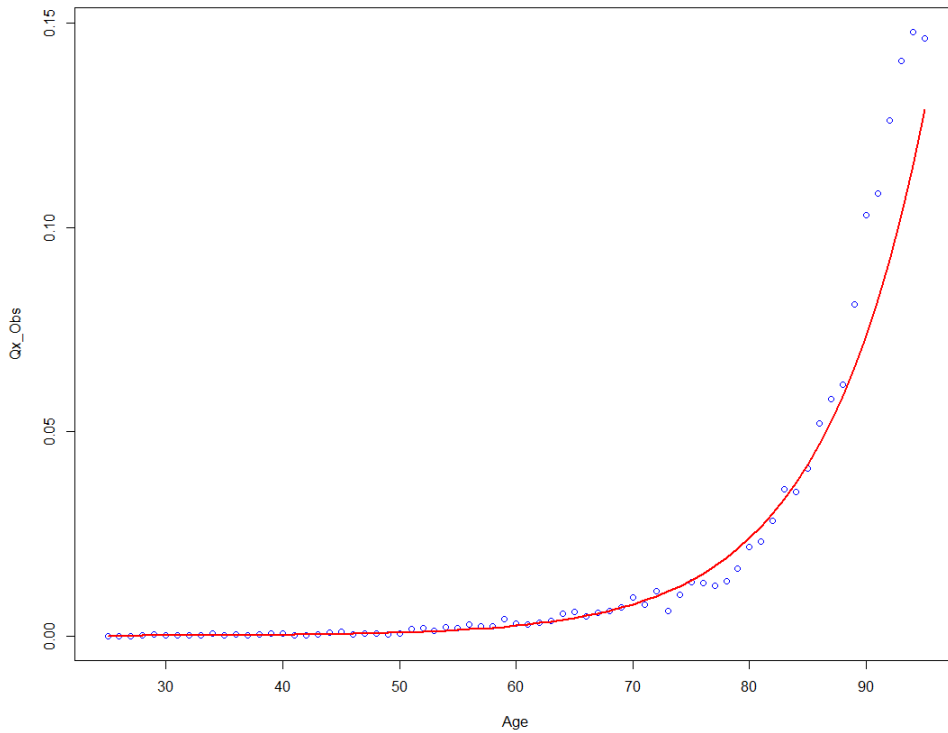


FIGURE 4.1 – Lissage des taux bruts par la méthode des moindres carrés - Femmes OD

Nous constatons que le lissage semble plutôt cohérent aux observations, d'où un coefficient de détermination proche de 1. Une légère sur-estimation de la mortalité entre les âges 70 et 85 ans et une sous-estimation de la mortalité à partir de 85 ans sont toutefois observables.

Regardons maintenant le cas des hommes OD. Nous considérons cette fois-ci, deux tranches d'âges : la première comprend les âges entre 25 et 59 ans (Tranche 1), la deuxième ceux entre 60 et 95 ans (Tranche 2).

Pour la tranche entre 25 et 59 ans, l'équation qui maximise le coefficient de détermination est la relation exponentielle suivante :

$$y = 0.0000144 \times e^{0.0956423x}$$

et pour la tranche comprise entre 60 et 95 ans, l'équation considérée est la relation exponentielle suivante :

$$y = 0.00000746 \times e^{0.10838103x}$$

avec x correspondant à l'âge.

Le tableau suivant présente les valeurs des coefficients de détermination R^2 obtenus dans le cas des hommes OD pour chacune des deux plages d'âges précédentes :

Tranche d'âges	25 : 59	60 : 95
R^2	0.902	0.994

TABLE 4.1 – Valeurs des coefficients de détermination pour la population des hommes OD

Sur les graphes suivants, les points bleus correspondent aux taux bruts, et la courbe rouge correspond au lissage des taux par la méthode des moindres carrés dans le cas des hommes OD.

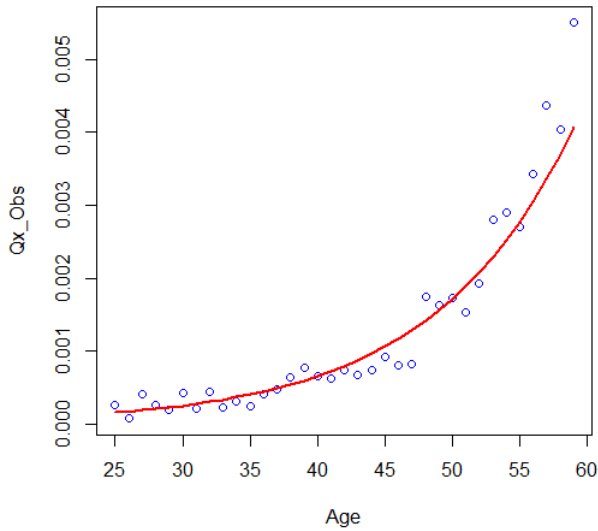


FIGURE 4.2 – Lissage des taux bruts par la méthode des moindres carrés - Hommes OD Tranche 1

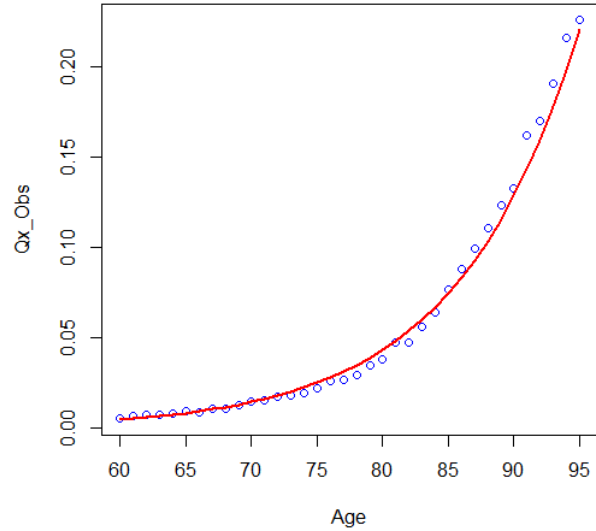


FIGURE 4.3 – Lissage des taux bruts par la méthode des moindres carrés - Hommes OD Tranche 2

Le lissage par cette méthode semble mieux s'adapter à la population des hommes, le coefficient de détermination pour les âges au-delà de 60 ans est très bon et graphiquement, la courbe de lissage se rapproche mieux des taux bruts.

Dans la suite de notre étude, nous aurons besoin de récupérer les taux de mortalité lissés sur l'ensemble de la population. C'est pourquoi il semble important d'étudier la jonction de ces deux courbes pour la population des hommes OD, de sorte à éviter d'avoir un saut important dans notre lissage au niveau du point de raccord.

Dans le cas présent, le travail s'effectue au niveau des taux à 59 ans et à 60 ans. Nous retiendrons comme taux lissé à 59 ans le taux $q_{x_{59}}$ suivant :

$$q_{x_{59}} = \frac{q_{x_{59}}^1 - q_{x_{59}}^2}{2}$$

avec :

- $q_{x_{59}}^1$: le taux de mortalité lissé à 59 ans obtenu avec l'équation de lissage de la Tranche 1.
- $q_{x_{59}}^2$: le taux de mortalité lissé à 59 ans obtenu avec l'équation de lissage de la Tranche 2.

D'autre part, nous retiendrons comme taux lissé à 60 ans le taux $q_{x_{60}}$ suivant :

$$q_{x_{60}} = \frac{q_{x_{60}}^1 - q_{x_{60}}^2}{2}$$

avec :

- $q_{x_{60}}^1$: le taux de mortalité lissé à 60 ans obtenu avec l'équation de lissage de la Tranche 1.
- $q_{x_{60}}^2$: le taux de mortalité lissé à 60 ans obtenu avec l'équation de lissage de la Tranche 2.

Ces taux "moyens" seront ceux utilisés dans les calculs par la suite.

— Cas des ayants droit :

Pour les femmes de la catégorie AD, l'équation qui maximise le coefficient de détermination sur la plage 27-95 ans est la relation exponentielle suivante :

$$y = 0.0000137 \times e^{0.0970532x}$$

avec x correspondant à l'âge.

Le coefficient de détermination obtenu avec cette équation de lissage pour la population des femmes AD a une valeur de 0.943.

Sur le graphe suivant, les points bleus correspondent aux taux bruts, et la courbe rouge correspond au lissage des taux par la méthode des moindres carrés dans le cas des femmes AD.

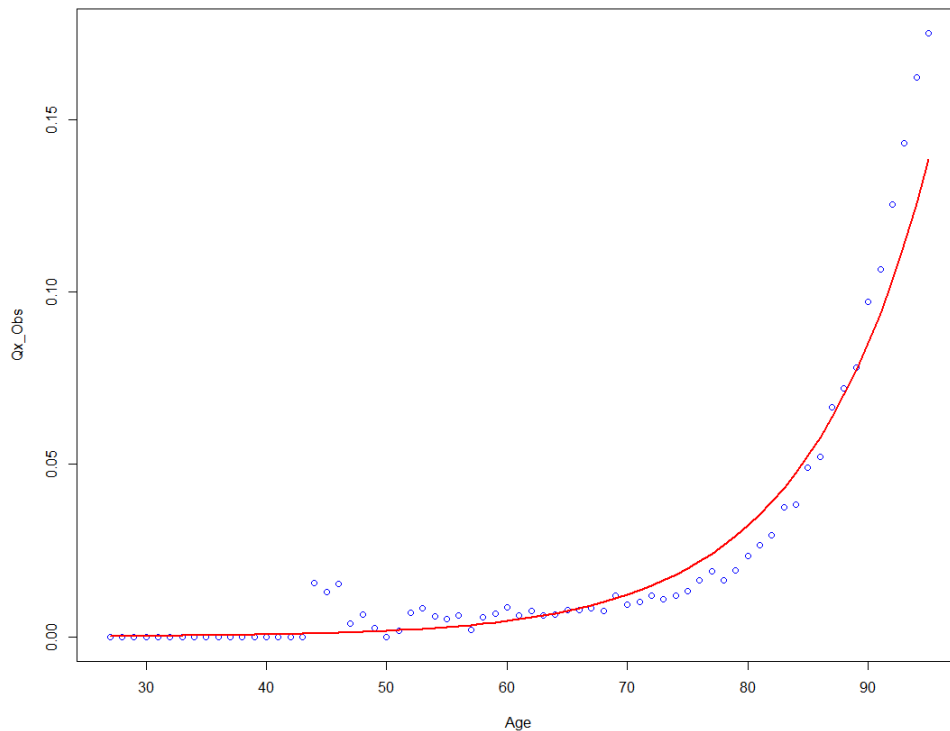


FIGURE 4.4 – Lissage des taux bruts par la méthode des moindres carrés - Femmes AD

Ce lissage semble moyennement s'adapter aux taux bruts. En effet, il sous-estime la mortalité des individus d'âges compris entre 45 et 60 ans, ainsi qu'au-delà de 90 ans. Une légère sur-estimation est également observable entre 70 et 85 ans. Le coefficient de détermination reste néanmoins largement acceptable.

Regardons maintenant le cas des hommes AD. L'équation qui maximise le coefficient de détermination sur la plage 35-95 ans est la relation exponentielle suivante :

$$y = 0.00011298 \times e^{0.07734021x}$$

avec x correspondant à l'âge.

Le coefficient de détermination obtenu avec cette équation de lissage pour la population des hommes AD a une valeur de 0.804.

Sur le graphe suivant, les points bleus correspondent aux taux bruts, et la courbe rouge correspond au lissage des taux par la méthode des moindres carrés dans le cas des hommes AD.

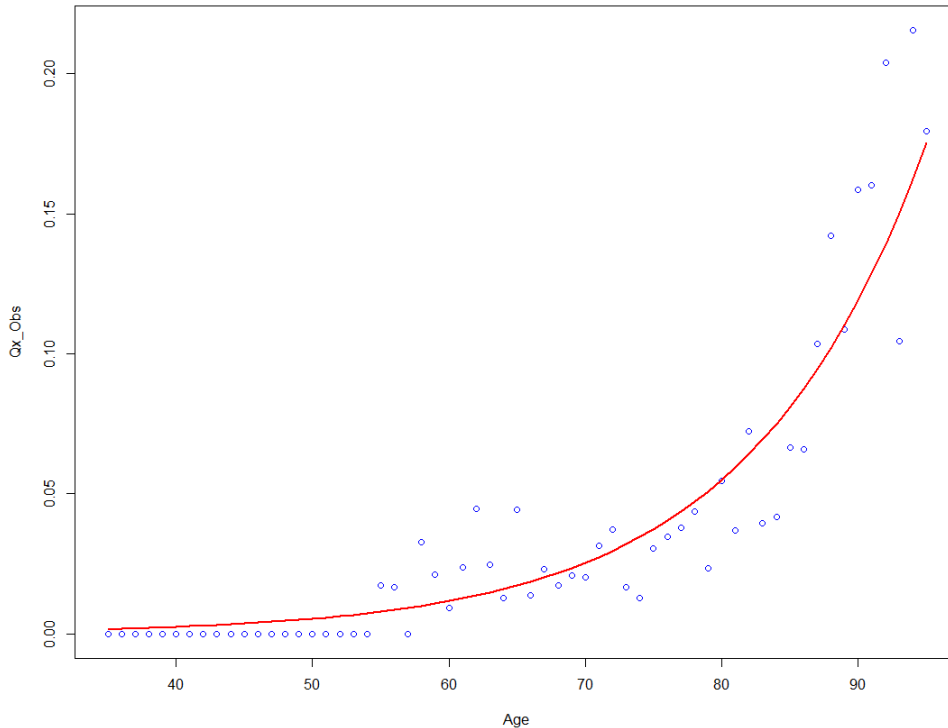


FIGURE 4.5 – Lissage des taux bruts par la méthode des moindres carrés - Hommes AD

Dans ce cas aussi, la méthode de lissage par les moindres carrés semble convenablement moyenniser nos taux bruts. Le coefficient de détermination est toutefois moins élevé que les précédents du fait de disparités plus nombreuses au niveau de la répartition des taux bruts de mortalité.

4.1.1.2 Lissage par la méthode de Makeham

La seconde méthode de lissage appliquée est la méthode de Makeham² (1860), il s'agit d'un lissage paramétrique.

- *Description de la méthode :*

Le modèle de Gompertz-Makeham est le modèle de référence pour la construction des tables de mortalité. De par sa simplicité, il s'agit du modèle le plus souvent utilisé.

Makeham met en évidence le fait que sur de nombreuses populations, les taux de mortalité instantanés observés croissent de manière quasi-exponentielle avec l'âge (modèle de Gompertz³ (1825)), mais tient également compte d'un critère qui refléterait le caractère accidentel à l'origine de la majorité des cas de mortalité aux jeunes âges.

La formule de Makeham s'exprime alors de la façon suivante :

$$\mu_x = \alpha + \beta\gamma^x$$

avec :

- μ_x : le taux instantané de mortalité à l'âge x
- α : le taux de décès accidentel (indépendant de l'âge) ($\alpha > 0$)
- $\beta\gamma^x$: le taux de décès modélisant le vieillissement exponentiel de la population (avec ($\beta > 0$ et $\gamma > 1$))

2. William MAKEHAM (1826-1891) : actuaire et mathématicien britannique

3. Benjamin GOMPertz (1779-1865) : actuaire et mathématicien britannique

4.1. CRÉATION D'UNE TABLE DE MORTALITÉ PAR DES MÉTHODES NON-PARAMÉTRIQUES

En pratique, le modèle que nous souhaitons ajuster est celui correspondant aux taux de mortalité bruts q_x calculés.

Pour cela, nous utilisons une "relation de passage" entre les taux de mortalité bruts et les taux de mortalité instantanés qui se caractérise de la sorte :

$$q_x = 1 - \exp \left[- \int_x^{x+1} \mu_y dy \right]$$

Nous avons alors :

$$p_x = \exp \left[- \int_x^{x+1} \mu_y dy \right] = \exp \left[- \int_x^{x+1} (\alpha + \beta \gamma^y) dy \right] = \exp(-\alpha) \exp \left[- \frac{\beta}{\ln \gamma} \gamma^x (\gamma - 1) \right]$$

Posons :

$$s = \exp(-\alpha) \quad \text{et} \quad g = \exp \left(- \frac{\beta}{\ln \gamma} \right)$$

Nous obtenons une fonction d'ajustement des taux de décès qui se traduit par l'expression suivante :

$$q_x = 1 - p_x = 1 - s g^{\gamma^x (\gamma - 1)}$$

Afin de déterminer des valeurs initiales convenables, nous utilisons la relation suivante en faisant au préalable l'approximation $\ln(1 - q_x) \approx -q_x$:

$$\ln(q_{x+1} - q_x) = x \ln \gamma + \ln \left(\frac{\beta}{\ln \gamma} (\gamma - 1)^2 \right) = ax + b$$

Les points $(x, y = \ln(q_{x+1} - q_x))$ sont donc approximativement alignés sur une droite de pente $\ln \gamma$.

Toutefois, il est à noter que ce modèle n'est pas adapté aux âges élevés, il convient mieux à des modélisations sur des intervalles d'âges raisonnables.

- Application à nos observations :

Les paramètres du modèle de Makeham sont estimés via l'ajustement des taux bruts par la fonction *fitGM()* disponible sous le logiciel *RStudio*.

— Cas des actifs et des inactifs ouvrants droit :

Le tableau suivant regroupe les valeurs des paramètres du modèle obtenues lors de l'ajustement des taux bruts pour chacune des populations des OD⁴ :

	s	g	γ
Femmes	1.0050	0.9993	1.1075
Hommes	1.0145	0.9809	1.0708

TABLE 4.2 – Paramètres du lissage par la méthode de Makeham pour la population des OD

Sur les graphes suivants, les points **bleus** correspondent aux taux bruts, et la courbe **rouge** correspond au lissage des taux par la méthode de Makeham dans le cas des OD.

4. Les valeurs des paramètres sont arrondies à 4 décimales.

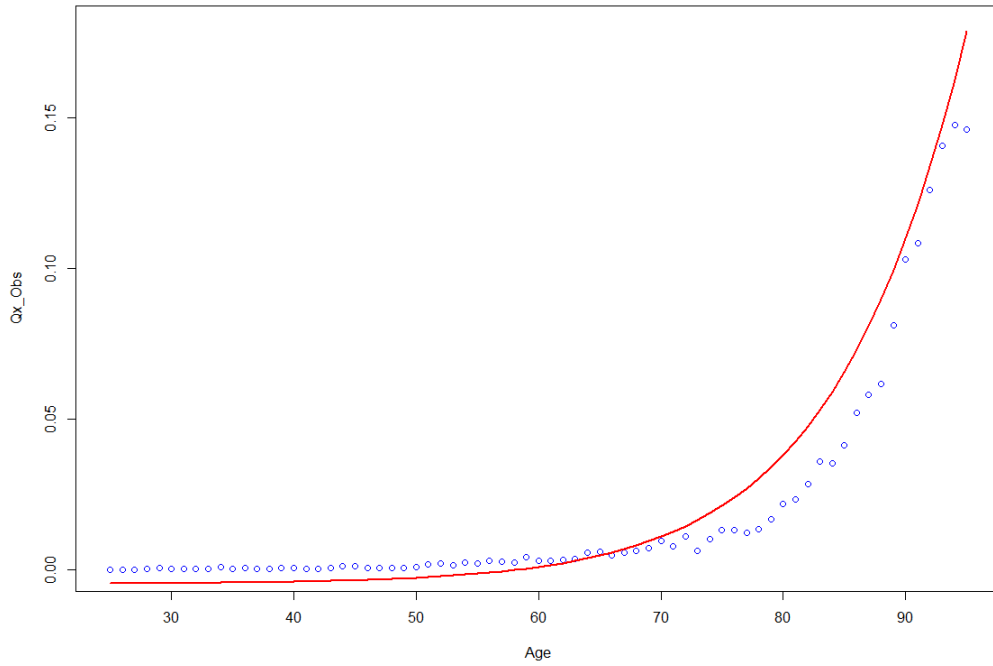


FIGURE 4.6 – Lissage des taux bruts par la méthode de Makeham - Femmes OD

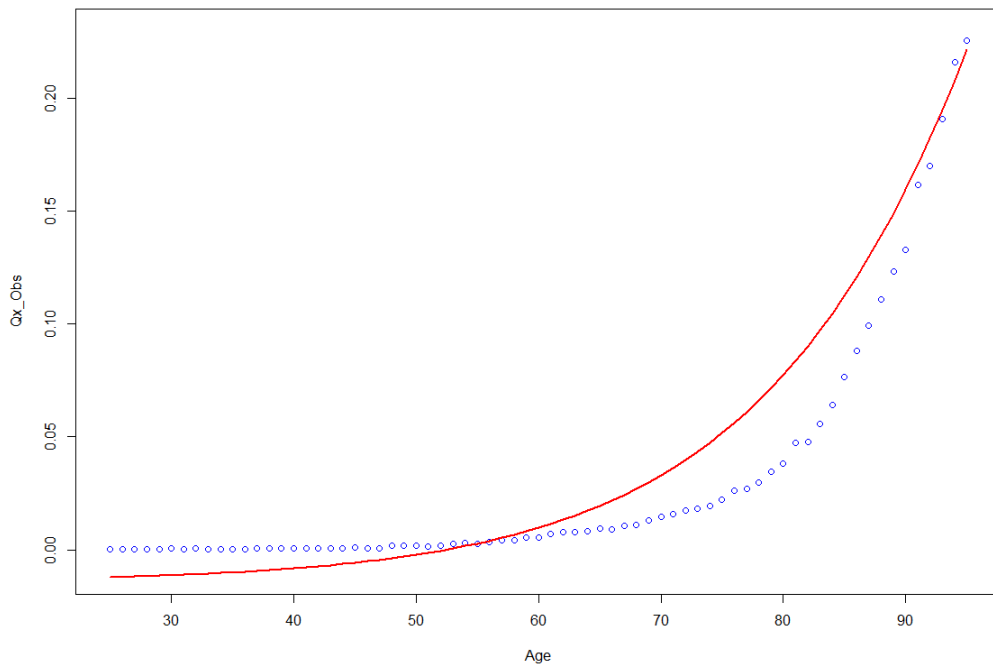


FIGURE 4.7 – Lissage des taux bruts par la méthode de Makeham - Hommes OD

Nous constatons, quelque soit le sexe de la population, que le lissage par la méthode de Makeham est peu adapté à nos populations d'ouvrants droit. En effet, il sous-estime les taux bruts de mortalité jusqu'à 65 ans chez les femmes et 55 ans chez les hommes, et les sur-estime au-delà de ces âges. Ceci est toutefois peu étonnant car ce modèle a souvent tendance à sur-estimer la mortalité aux âges élevés, comme énoncé précédemment, il est peu adapté aux âges élevés.

— Cas des ayants droit :

Le tableau suivant regroupe les valeurs des paramètres du modèle obtenues lors de l'ajustement des taux bruts pour chacune des populations des AD :

	s	g	γ
Femmes	0.9966	0.9983	1.1080
Hommes	0.9940	0.9956	1.0975

TABLE 4.3 – Paramètres du lissage par la méthode de Makeham pour la population des AD

Sur les graphes suivants, les points **bleus** correspondent aux taux bruts, et la courbe **rouge** correspond au lissage des taux par la méthode de Makeham dans le cas des AD.

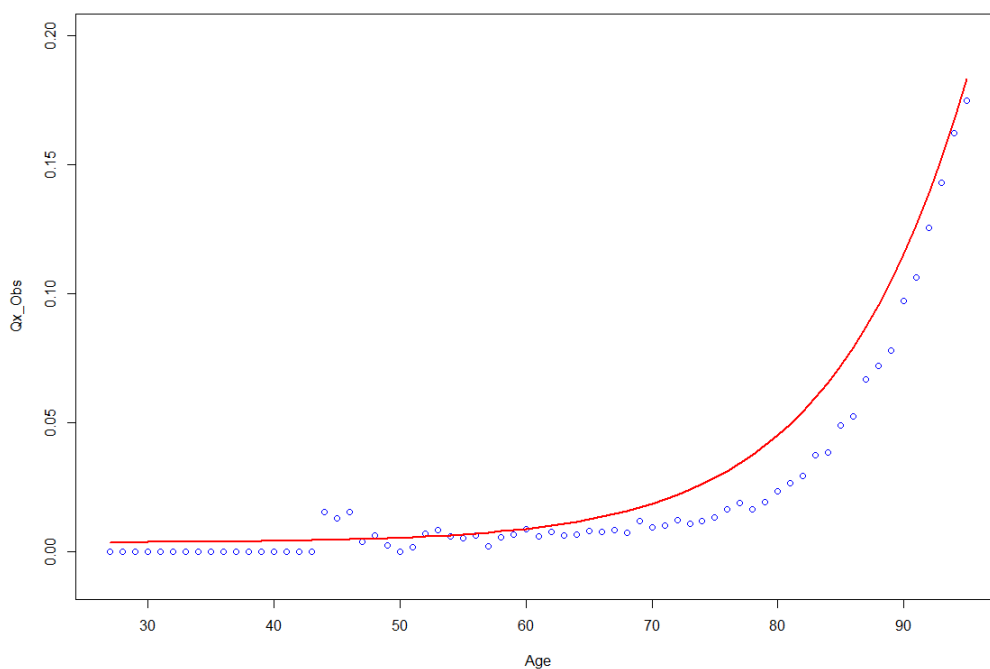


FIGURE 4.8 – Lissage des taux bruts par la méthode de Makeham - Femmes AD

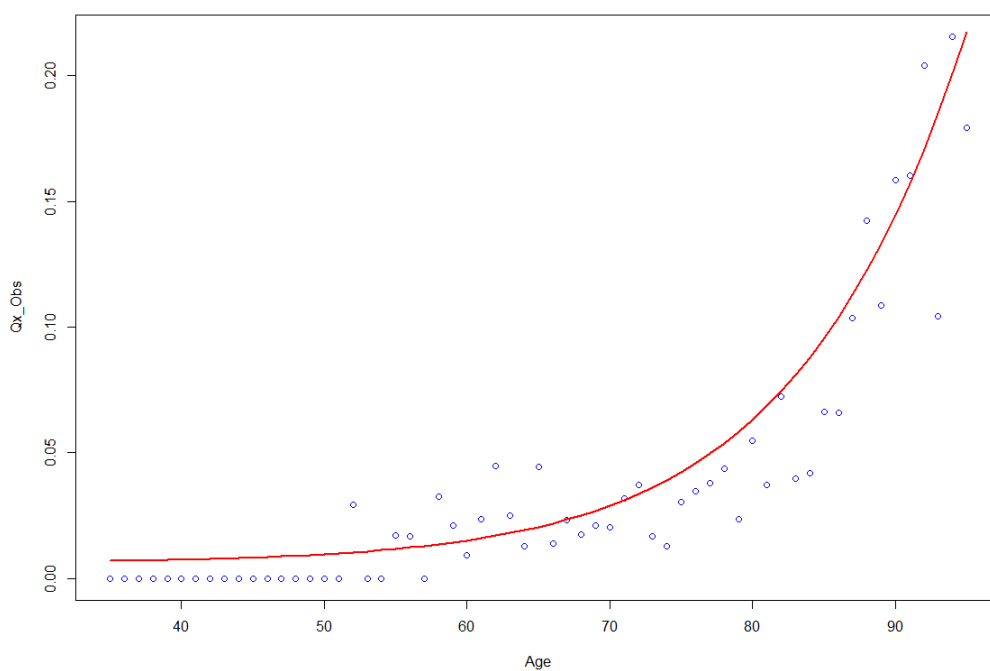


FIGURE 4.9 – Lissage des taux bruts par la méthode de Makeham - Hommes AD

Là encore, la mortalité de nos populations est sur-estimée par le lissage réalisé aux âges élevés (à partir de 55 ans pour les femmes et à partir de 70 ans pour les hommes), mais également aux âges inférieurs à 45 ans pour les femmes et à 55 ans pour les hommes.

Une sur-estimation de la mortalité n'est pas acceptable en régime de retraite, c'est pourquoi nous pouvons rejeter cette méthode dès à présent.

4.1.1.3 Lissage par la méthode de Loess

La troisième méthode de lissage appliquée est la méthode de LOESS (*LOcally weighted Scatterplot Smoother*) ou régression locale, il s'agit d'une méthode de régression non-paramétrique qui repose sur le principe de la méthode des k plus proches voisins en combinant plusieurs modèles de régression multiple.

- Description de la méthode :

La régression locale est équivalente à une régression polynomiale avec pondération locale. Deux notions sont alors importantes dans le principe de cette méthode de lissage : le voisinage et la pondération. En effet, pour chaque point d'abscisse x_i du jeu d'observations, un polynôme de degré 1 ou 2 est ajusté aux points du voisinage de x_i afin de déterminer l'ordonnée du point de lissage correspondant. D'autre part, les points les plus proches des points x_i auront plus de poids lors de l'ajustement.

En ce qui concerne le voisinage, sa taille est déterminée par un paramètre noté α :

- Si $\alpha < 1$, seule une proportion α des points du jeu d'observations sont inclus dans le voisinage.
- Si $\alpha \geq 1$, l'ensemble des points est utilisé, α n'a alors pas d'importance au niveau du voisinage mais en aura au niveau de la pondération.

Par ailleurs, en ce qui concerne la pondération, le poids w d'un point x_j situé au voisinage d'un point x_i aura d'autant plus d'importance que sa distance d_{ij} au point x_i est petite.

Les poids sont déterminés par la formule suivante :

$$w = \left(1 - \left(\frac{d_{ij}}{d_{max}} \right)^3 \right)^3$$

- Si $\alpha < 1$, d_{max} correspond à la distance maximale qui sépare deux points du voisinage considéré.
- Si $\alpha \geq 1$, d_{max} correspond au produit de α avec la distance maximale observée entre deux points du jeu d'observations.

À partir du voisinage et des poids, les données peuvent être ajustées par un polynôme de faible degré (1 ou 2), la valeur prédite en x_i par la méthode de LOESS correspond à la valeur prédite par l'ajustement local du polynôme considéré.

L'avantage de cette méthode est qu'elle n'impose pas de fonction globale pour l'ajustement du modèle à la totalité des observations à lisser. Par ailleurs elle a l'intérêt d'étendre la méthode des moindres carrés tout en étant assez simple d'implémentation. Cependant, ce lissage est assez sensible aux valeurs aberrantes et nécessite une taille d'échantillon assez conséquente pour ajuster un modèle de bonne qualité.

- Application à nos observations :

La fonction `loess()` disponible sous *RStudio* construit des courbes lissées et ajustées à un nuage de points (d'observations) en appliquant la méthode de Loess.

Dans notre application, la valeur de α est de 0.75 par défaut, cette valeur semble adaptée à notre modèle car elle permet de rester proche des données observées tout en limitant l'importance donnée aux valeurs qui pourraient être aberrantes.

— Cas des actifs et des inactifs ouvrants droit :

Les points bleus correspondent aux taux bruts, et la courbe rouge correspond au lissage des taux par la méthode de Loess dans le cas des OD, à gauche pour ce qui concerne les femmes, à droite pour ce qui concerne les hommes.

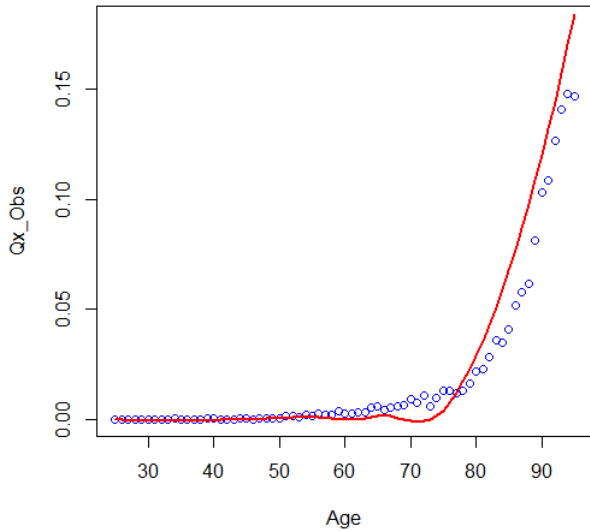


FIGURE 4.10 – Lissage des taux bruts par la méthode de Loess - Femmes OD

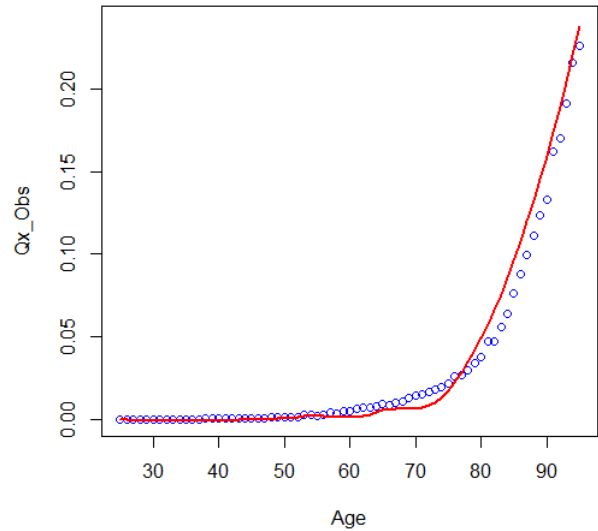


FIGURE 4.11 – Lissage des taux bruts par la méthode de Loess - Hommes OD

Sur les deux graphiques précédents, nous remarquons que le lissage effectué est assez cohérent avec les observations réelles pour les âges jusqu'à 65 ans, que le lissage sur-estime la mortalité de notre population au-delà de 75 ans, mais qu'il la sous-estime entre 70 et 75 ans, quelque soit le sexe de la population.

Il semble alors intéressant de découper notre population en deux tranches d'âges et d'effectuer un lissage par la méthode de Loess propre à chacune de ces deux tranches qui sont la tranche inférieure à 65 ans (Tranche 1), la tranche supérieure à 65 ans (Tranche 2).

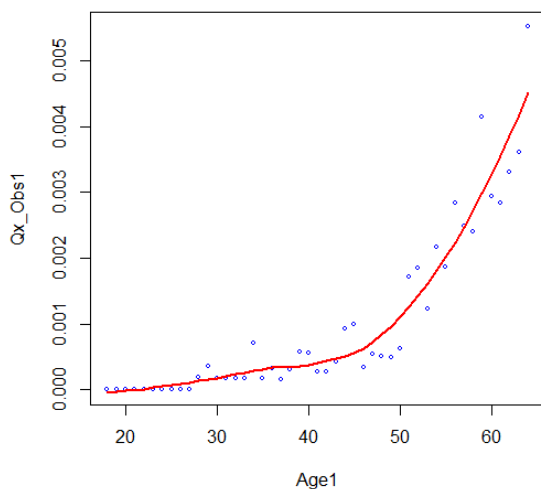


FIGURE 4.12 – Lissage des taux bruts par la méthode de Loess - Femmes OD Tranche 1

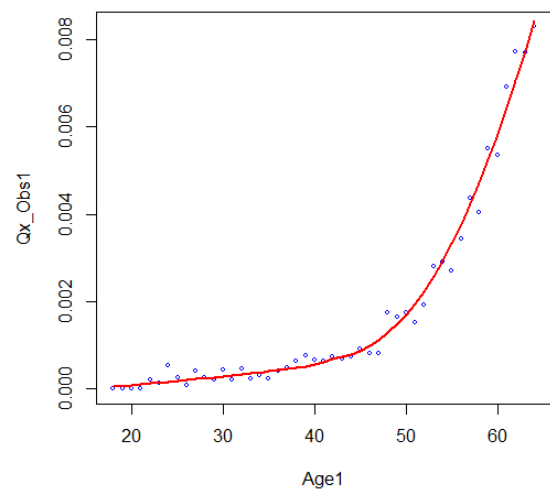


FIGURE 4.13 – Lissage des taux bruts par la méthode de Loess - Hommes OD Tranche 1

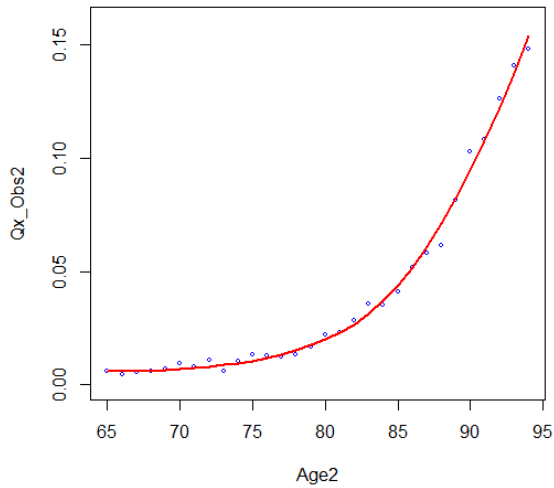


FIGURE 4.14 – Lissage des taux bruts par la méthode de Loess - Femmes OD Tranche 2

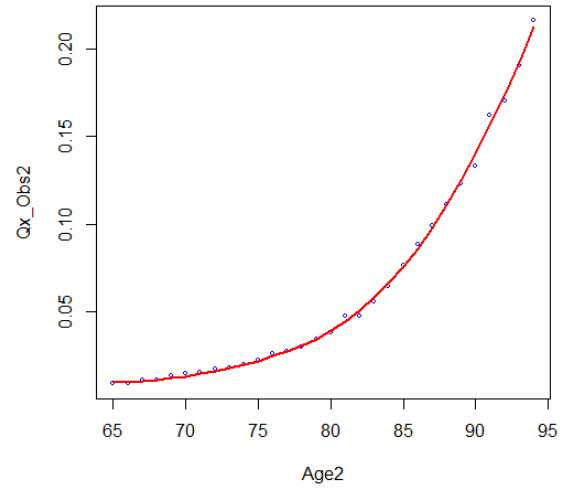


FIGURE 4.15 – Lissage des taux bruts par la méthode de Loess - Hommes OD Tranche 2

Nous constatons que pour les deux tranches, le lissage est assez fidèle aux observations. Nous ne remarquons pas de grands écarts d'ajustement comme dans les cas précédents.

Tout comme dans le cas des hommes OD lors de l'ajustement par la méthode des moindres carrés, nous devons étudier la jonction entre les courbes de lissage pour nos deux populations. Cette fois-ci, le taux de raccord sera le taux à 65 ans $q_{x_{65}}$ retenu comme suit quel que soit le sexe :

$$q_{x_{65}} = \frac{q_{x_{65}}^1 - q_{x_{65}}^2}{2}$$

avec :

- $q_{x_{65}}^1$: le taux de mortalité lissé à 65 ans obtenu avec le lissage de Loess sur la Tranche 1.
- $q_{x_{65}}^2$: le taux de mortalité lissé à 65 ans obtenu avec le lissage de Loess sur la Tranche 2.

— Cas des ayants droit :

Les points bleus correspondent aux taux bruts, et la courbe rouge correspond au lissage des taux par la méthode de Loess dans le cas des AD, à gauche pour ce qui concerne les femmes, à droite pour ce qui concerne les hommes.

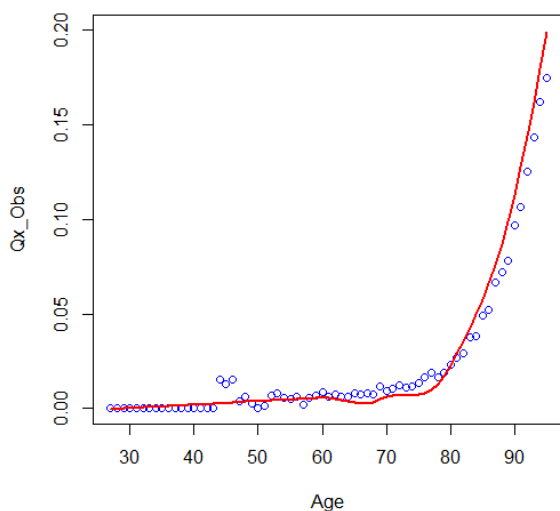


FIGURE 4.16 – Lissage des taux bruts par la méthode de Loess - Femmes AD

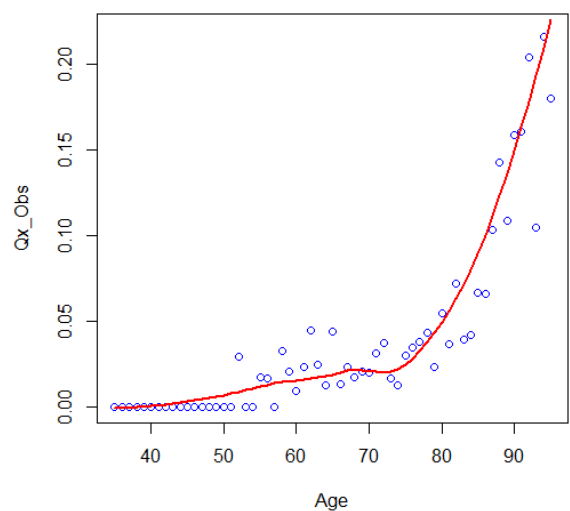


FIGURE 4.17 – Lissage des taux bruts par la méthode de Loess - Hommes AD

4.1. CRÉATION D'UNE TABLE DE MORTALITÉ PAR DES MÉTHODES NON-PARAMÉTRIQUES

Cette fois-ci, le lissage semble plutôt bien correspondre aux données d'observation avec toutefois une légère sous-estimation entre les âges 60 et 80 ans chez les femmes.

Découpons notre population des AD de la même manière que les OD et effectuons un lissage par la méthode de Loess sur chacune des deux tranches d'âges, ceci afin d'obtenir un lissage d'une meilleure précision.

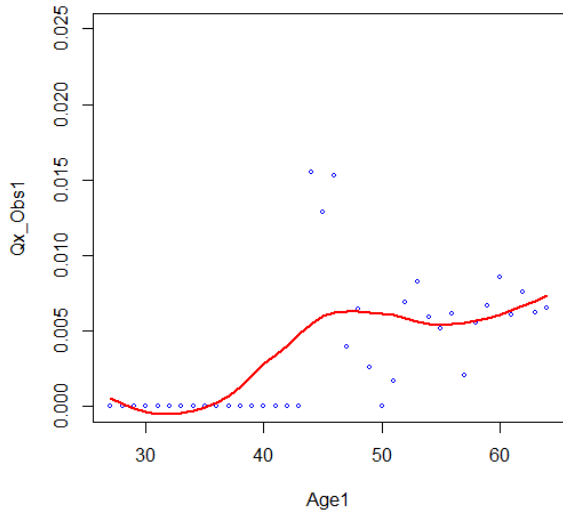


FIGURE 4.18 – Lissage des taux bruts par la méthode de Loess - Femmes AD Tranche 1

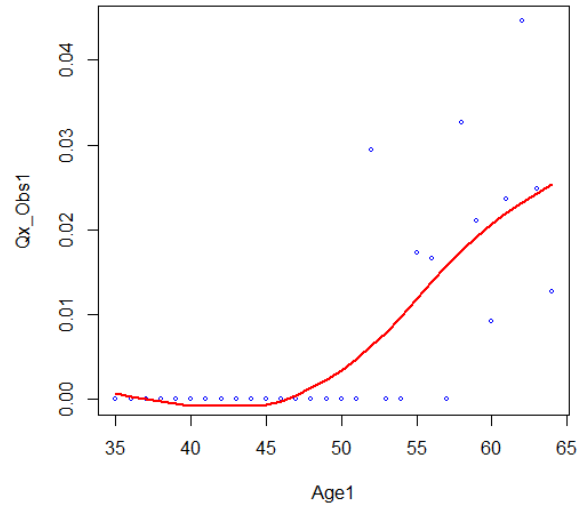


FIGURE 4.19 – Lissage des taux bruts par la méthode de Loess - Hommes AD Tranche 1

Nous constatons que pour la première tranche d'âges (FIGURES 4.18 et 4.19), le lissage n'est pas adapté. En effet, la mortalité estimée est négative entre 30 et 35 ans chez les femmes, et entre 35 et 40 chez les hommes, ce qui ne peut correspondre à la réalité. Pour les âges compris au-delà de cette plage, le lissage semble convenablement moyenner les taux bruts de mortalité. Toutefois, des écarts assez importants entre la courbe de lissage et les taux bruts peuvent être constatés du fait de la grande disparité des taux bruts à ces âges.

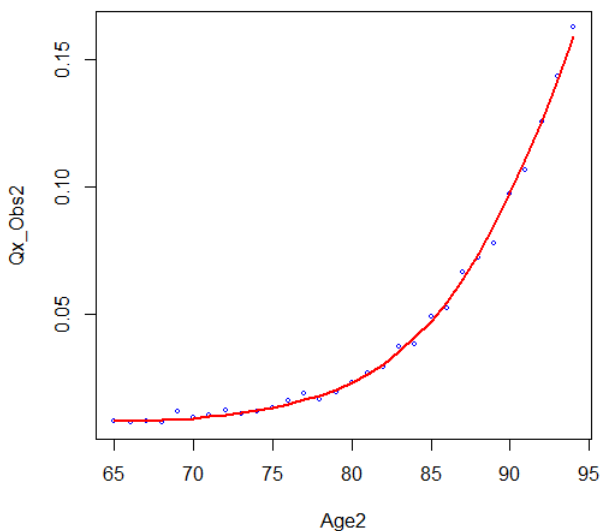


FIGURE 4.20 – Lissage des taux bruts par la méthode de Loess - Femmes AD Tranche 2

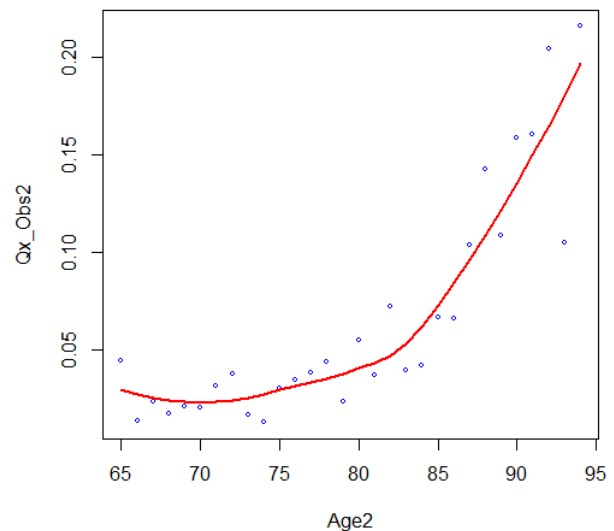


FIGURE 4.21 – Lissage des taux bruts par la méthode de Loess - Hommes AD Tranche 2

En ce qui concerne la deuxième tranche d'âges, le lissage semble plutôt bien adapté aux observations des femmes. En revanche, il est à rejeter pour les hommes du fait d'une décroissance des taux lissés entre 65 et 70 ans.

4.1.1.4 Lissage par la méthode de Whittaker-Henderson

La dernière méthode de lissage appliquée est la méthode de Whittaker-Henderson^{5 6} (1923-1924), il s'agit d'une méthode de lissage non paramétrique.

• Description de la méthode :

La méthode de Whittaker-Henderson repose sur un principe de combinaison d'un critère de fidélité avec un critère de régularité, l'objectif étant de rechercher les valeurs qui minimisent une combinaison linéaire de ces deux critères.

Les taux de mortalité que nous cherchons à lisser ne dépendent que d'un seul paramètre (l'âge), donc le lissage effectué ici sera de dimension 1.

Le **critère de fidélité** est défini comme suit :

$$F = \sum_{i=1}^n w_i (\hat{q}_i - q_i)^2$$

avec :

- n : le nombre de taux de mortalité à lisser
- w_i : les poids attribués à chaque taux de mortalité
- \hat{q}_i : les taux de mortalité bruts observés à lisser
- q_i : les taux de mortalité lissés

Le poids des taux peut être défini de plusieurs manières. Pour notre étude, afin de limiter le poids donné aux points aberrants, nous choisissons d'effectuer une pondération par l'effectif à chaque âge x :

$$w_x = \frac{\text{Effectif}_x}{\text{Effectif moyen tous âges}}$$

avec :

$$\text{Effectif moyen tous âges} = \frac{\sum_{i=\text{Age}_{\min}}^{\text{Age}_{\max}} \text{Effectif}_i}{\text{Age}_{\max} - \text{Age}_{\min} + 1}$$

Le **critère de régularité** se définit comme suit :

$$S = \sum_{i=1}^{n-z} (\Delta^z q_i)^2$$

avec :

- n : le nombre de taux de mortalité à lisser
- z : un paramètre du modèle
- q_i : les taux de mortalité lissés
- Δ^z : désigne une différence avant d'ordre z

Notons qu'une différence avant d'ordre 1, pour toute fonction f et pour tout x entier, équivaut à :

$$\Delta f(x) = f(x+1) - f(x)$$

Donc une différence avant d'ordre z , pour toute fonction f et pour tout x entier, équivaut à :

$$\Delta^z f(x) = \sum_{j=0}^z \binom{z}{j} (-1)^{z-j} f(x+j)$$

5. Edmund Taylor WHITTAKER (1873-1956) : astronome, mathématicien et historien des sciences britannique

6. Robert HENDERSON (1871-1942) : actuair et mathématicien américano-canadien

Ce qu'il faut chercher à minimiser est alors une combinaison linéaire de la fidélité et de la régularité. Le poids attribué entre ces deux critères est déterminé par un paramètre h représentant l'importance que nous souhaitons donner à la régularité (plus h sera grand, plus la courbe sera lisse, inversement, plus h sera petit, plus la courbe se rapprochera des données brutes).

La relation à minimiser est alors de la forme :

$$M = F + h \times S$$

Ce qui nous amène à résoudre les n équations suivantes :

$$\forall 1 \leq i \leq n, \quad \frac{\partial M}{\partial q_i} = 0$$

Ce système d'équations peut être résolu à l'aide de formes matricielles. En effet, les critères de fidélité et de régularité peuvent s'écrire matriciellement :

$$F = (q - \hat{q})' w (q - \hat{q})$$

$$S = (\Delta^z q)' (\Delta^z q)$$

avec les notations $q = (q_i)_{1 \leq i \leq n}$, $\hat{q} = (\hat{q}_i)_{1 \leq i \leq n}$, $\Delta^z q = (\Delta^z q_i)_{1 \leq i \leq n}$ et $w = \text{diag}(w_i)_{1 \leq i \leq n}$

Par ailleurs, considérons une matrice K_z de taille $(n - z, n)$ qui a pour termes les coefficients binomiaux d'ordre z et dont les signes sont alternés en commençant par un signe positif si z est pair. La matrice K_z est alors telle que $\Delta^z q = K_z q$.

La forme de M peut alors s'écrire :

$$M = (q - \hat{q})' w (q - \hat{q}) + h q' K_z' K_z q$$

D'où par développement :

$$M = q' w q - 2q' w \hat{q} + \hat{q}' w \hat{q} + h q' K_z' K_z q$$

Puis par dérivation :

$$\frac{\partial M}{\partial q} = 2wq - 2w\hat{q} + 2hK_z' K_z q$$

Donc en résolvant $\frac{\partial M}{\partial q} = 0$, nous obtenons comme expression pour les taux de mortalité ajustés :

$$q^* = (w + hK_z' K_z)^{-1} w \hat{q}$$

Cette méthode est avantageuse dans la mesure où elle permet une certaine liberté de choix d'ajustement selon que nous souhaitons un ajustement plus proche des données observées ou plus lisse qui viendrait pallier les aberrations.

- Application à nos observations :

Dans notre étude, du fait de l'irrégularité des données, nous favorisons le choix d'accorder plus d'importance à la régularité qu'à la fidélité. Les résultats présentés ci-dessous sont obtenus avec $h = 20$.

— Cas des actifs et des inactifs ouvrants droit :

Sur les graphes suivants, les points bleus correspondent aux taux bruts, et la courbe rouge correspond au lissage des taux par la méthode de Whittaker-Henderson dans le cas des OD.

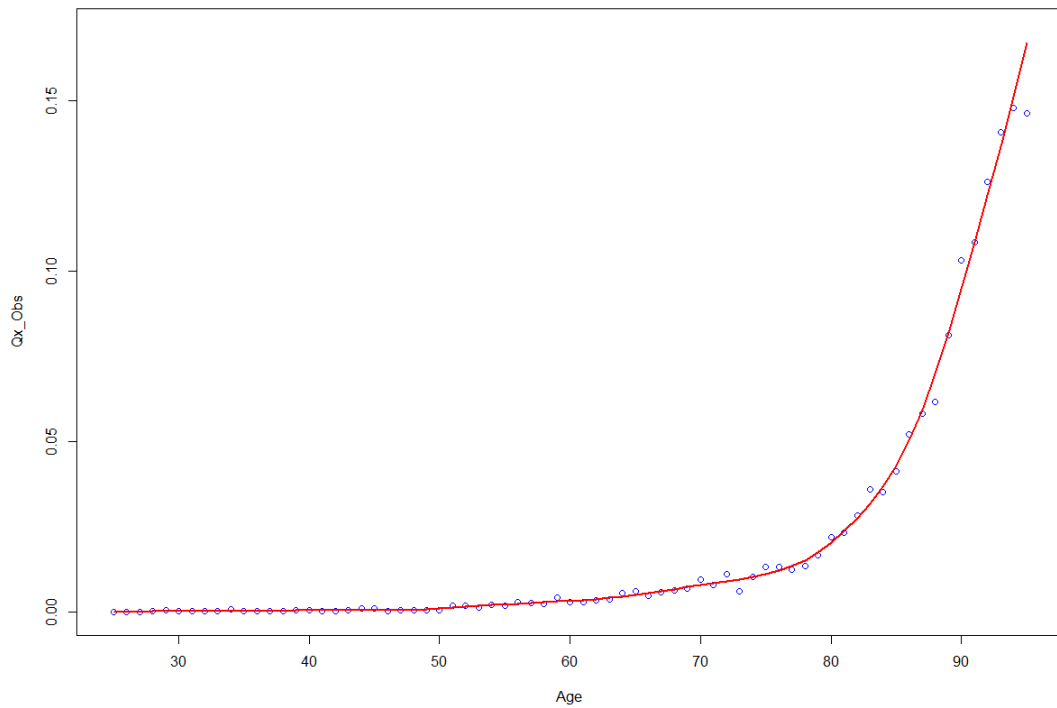


FIGURE 4.22 – Lissage des taux bruts par la méthode de Whittaker-Henderson - Femmes OD

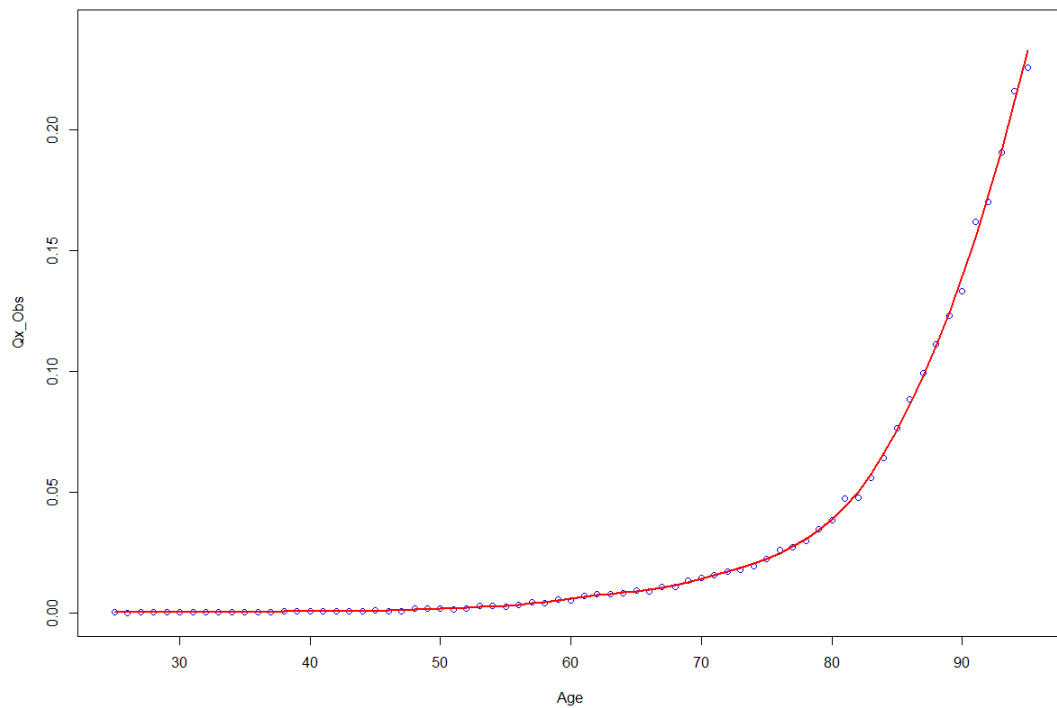


FIGURE 4.23 – Lissage des taux bruts par la méthode de Whittaker-Henderson - Hommes OD

Que ce soit la population des femmes ou la population des hommes, le lissage par la méthode de Whittaker-Henderson s'accorde très bien à nos taux bruts de mortalité. Les écarts observés entre ces derniers et la courbe de lissage sont très faibles.

— Cas des ayants droit :

Sur les graphes suivants, les points bleus correspondent aux taux bruts, et la courbe rouge correspond au lissage des taux par la méthode de Whittaker-Henderson dans le cas des AD.

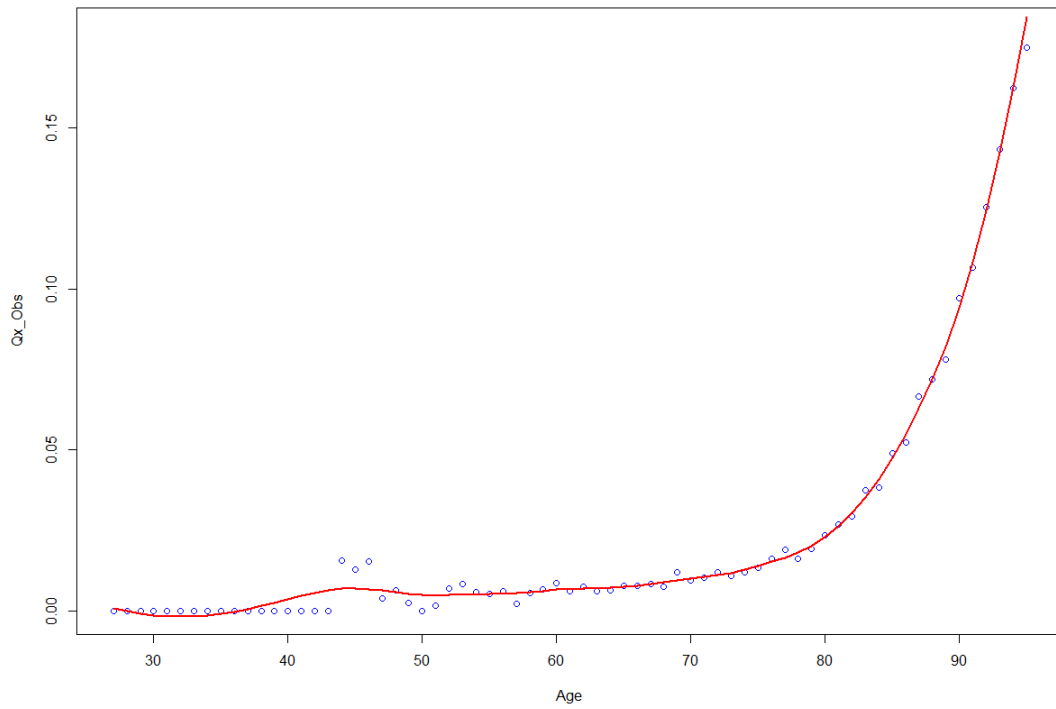


FIGURE 4.24 – Lissage des taux bruts par la méthode de Whittaker-Henderson - Femmes AD

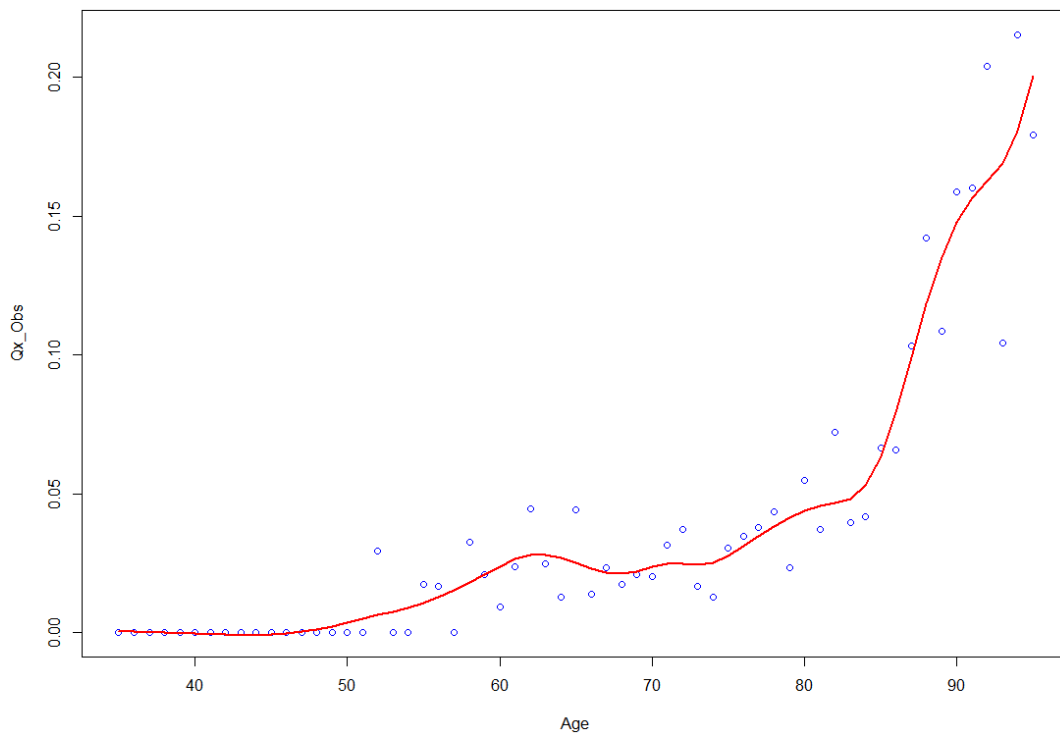


FIGURE 4.25 – Lissage des taux bruts par la méthode de Whittaker-Henderson - Hommes AD

De même que dans le cas des ouvrants droit, dans le cas des ayants droit, le lissage par la méthode de Whittaker-Henderson semble plutôt bien adaptée aux observations et particulièrement chez les femmes.

4.1.1.5 Comparaison des résultats obtenus

Une fois toutes ces méthodes de lissage mises en application et étudiées individuellement, il convient de les comparer entre elles afin de n'en garder qu'une seule qui nous fournira les taux de mortalité sur lesquels notre table sera ajustée. Pour sélectionner la meilleure d'entre elles, nous réalisons quatre comparaisons : dans un premier temps une comparaison graphique qui nous donnera une première idée des résultats attendus, puis nous calculerons le R^2 , le χ^2 et la *Mean Absolute Percentage Error (MAPE)* pour chacun des lissages par rapport aux taux bruts.

Commençons par évaluer la qualité des lissages graphiquement. Pour chacune des populations, nous réalisons les comparaisons sur l'ensemble de la population, puis sur des morceaux de population par tranches d'âges (25-49 ans (Tranche 1), 50-69 ans (Tranche 2), 70-95 ans (Tranche 3), et 60-89 (Tranche 4) qui est la tranche contenant le plus d'individus et donc la tranche d'âges la plus pertinente en terme de choix de lissage)⁷.

— Cas des actifs et des inactifs ouvrants droit :

Regardons le cas des femmes OD pour chacune des plages d'âges énoncées ci-dessus.

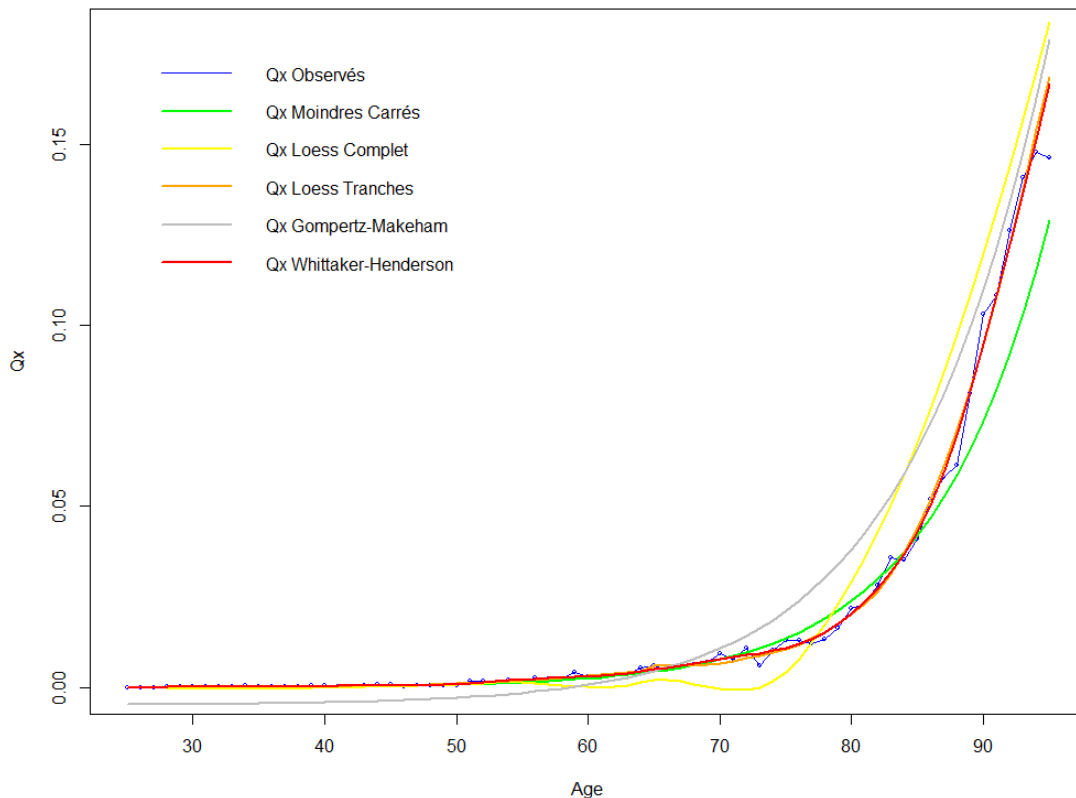


FIGURE 4.26 – Comparaison des lissages réalisés - Femmes OD Complet

Sur l'ensemble de la population des femmes OD, les lissages qui semblent le mieux suivre les taux bruts sont les lissages par la méthode de Loess par tranches ou la méthode de Whittaker-Henderson. Ces observations se distinguent à partir de 60 ans, avant cet âge, seul le lissage par la méthode de Makeham est rejetable graphiquement. c'est pourquoi il semble intéressant de se tourner vers des tranches d'observations plus restreintes.

7. Nous présentons dans cette section l'ensemble des graphiques pour la catégories des femmes OD, pour ce qui concerne les autres catégories, seules certaines tranches qui semblent pertinentes seront présentées dans cette section, les autres graphes se trouveront en ANNEXES E.1, E.2 et E.3.

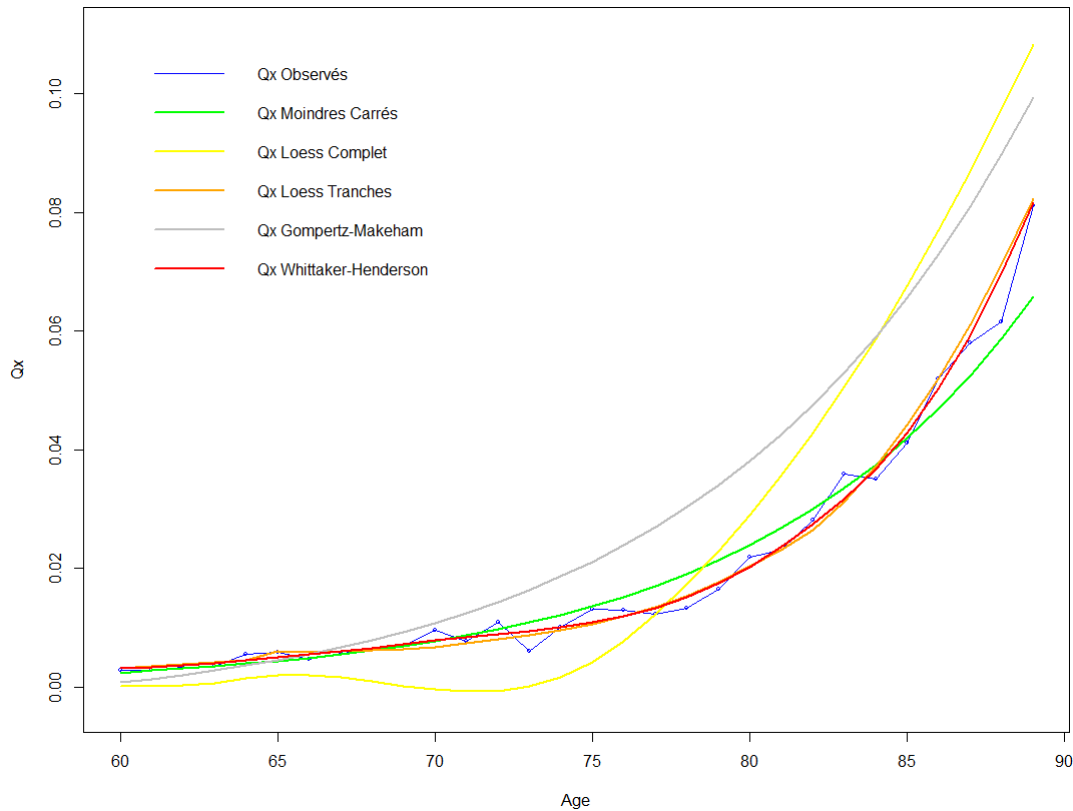


FIGURE 4.27 – Comparaison des lissages réalisés - Femmes OD Tranche 4

La FIGURE 4.27 contenant les observations de la tranche 60 - 89 ans, plage d'âges où la population est la plus importante, nous permet d'écarter les autres méthodes de lissage. Cependant, les deux courbes retenues (Loess par tranches et Whittaker-Henderson) sont trop rapprochées sur cette plage pour déterminer laquelle des deux est la plus pertinente.

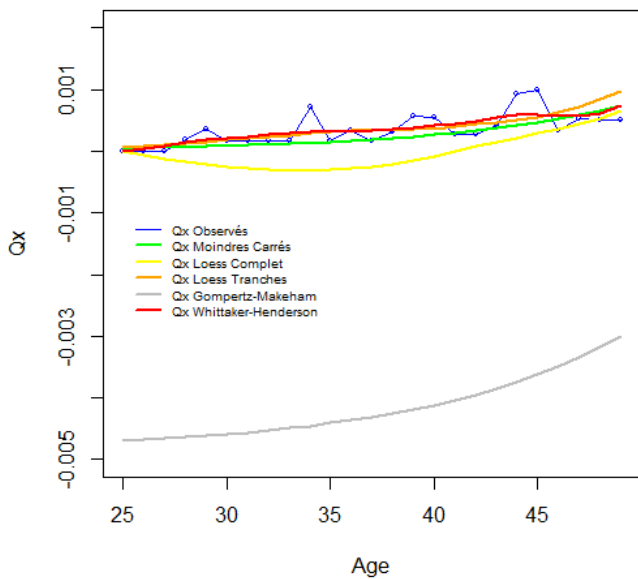


FIGURE 4.28 – Comparaison des lissages réalisés - Femmes OD Tranche 1

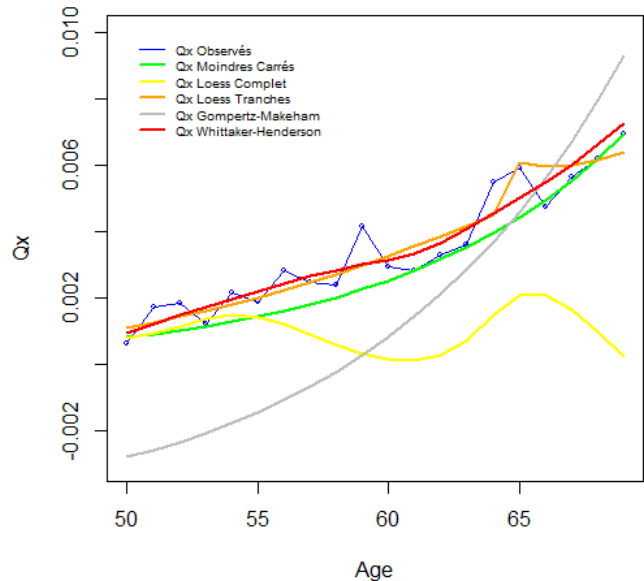


FIGURE 4.29 – Comparaison des lissages réalisés - Femmes OD Tranche 2

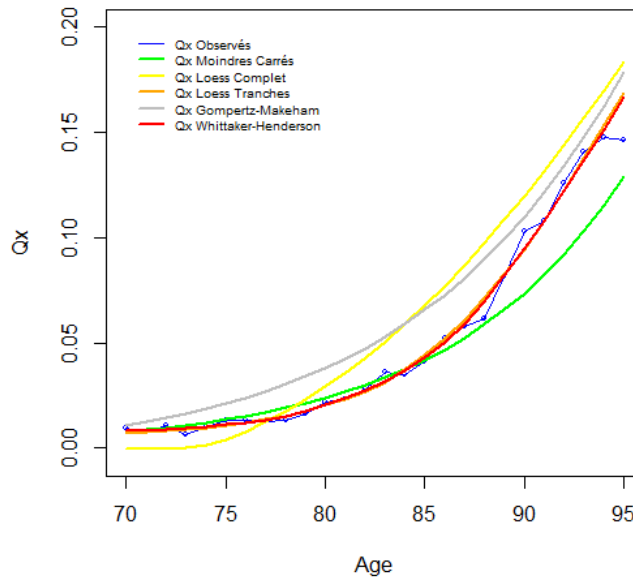


FIGURE 4.30 – Comparaison des lissages réalisés - Femmes OD Tranche 3

En découpant la population en tranches d'âges plus raisonnables, nous constatons que les lissages de Loess par tranches et de Whittaker-Henderson sont plus ou moins superposés sauf pour ce qui est des âges compris entre 65 et 70 ans où les taux obtenus avec la méthode de Whittaker-Henderson ont une courbe plus lissée que celle des taux obtenus par la méthode de Loess par tranches. C'est pourquoi graphiquement nous pouvons faire l'hypothèse que le lissage de Whittaker-Henderson est le mieux adapté à cette population.

Regardons maintenant le cas des hommes OD :

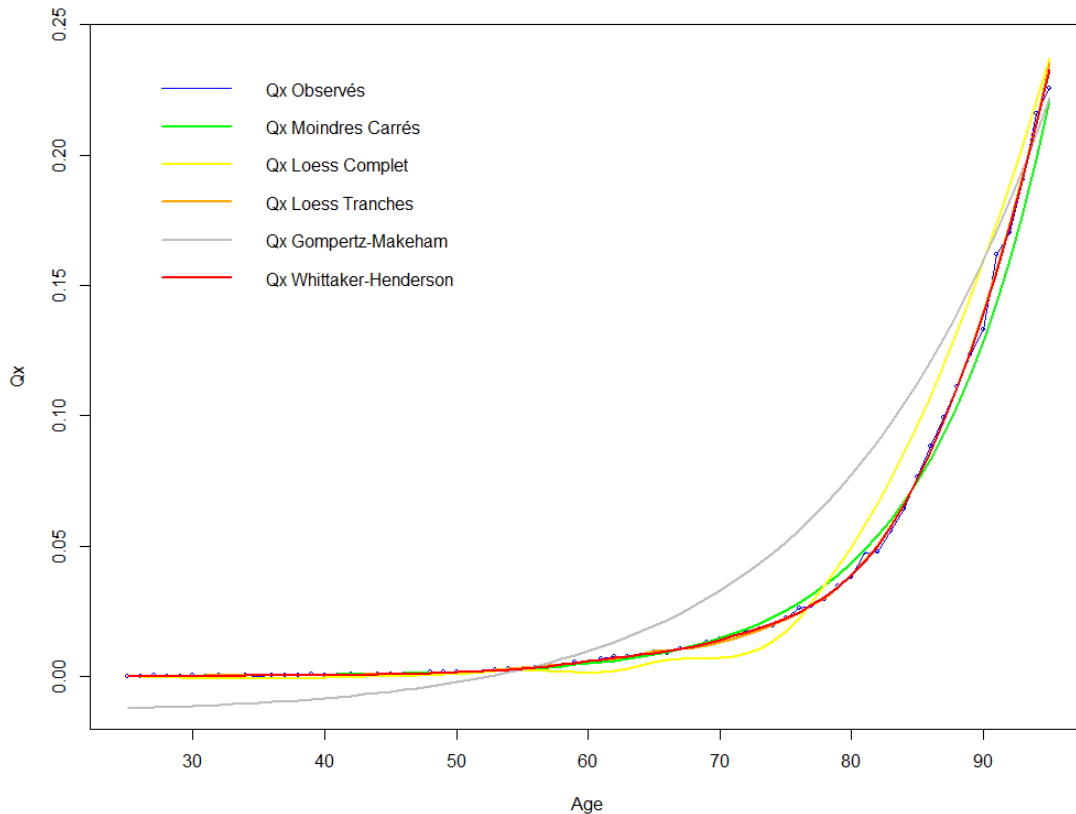


FIGURE 4.31 – Comparaison des lissages réalisés - Hommes OD Complet

4.1. CRÉATION D'UNE TABLE DE MORTALITÉ PAR DES MÉTHODES NON-PARAMÉTRIQUES

Nous constatons que les lissages par les méthodes des moindres carrés, Loess par tranches et Whittaker-Henderson semblent être les plus proches des taux d'observation. Toutefois il est difficile de juger laquelle de ces trois méthodes est la plus pertinente au vu de la FIGURE 4.31.

Regardons alors de plus près les plages d'observations restreintes :

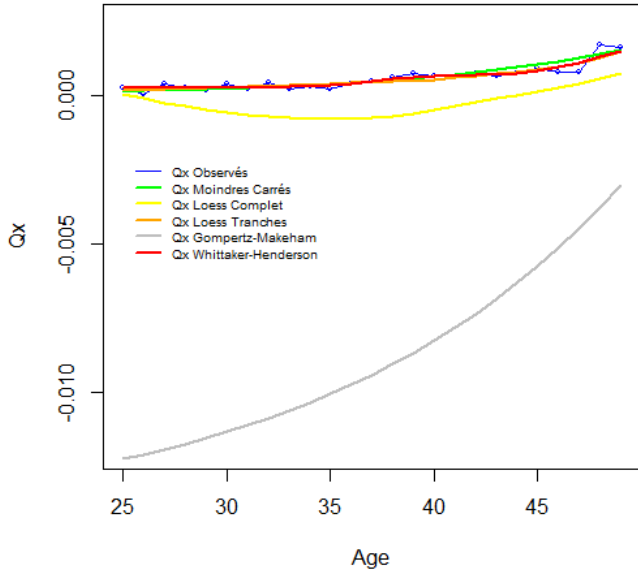


FIGURE 4.32 – Comparaison des lissages réalisés - Hommes OD Tranche 1

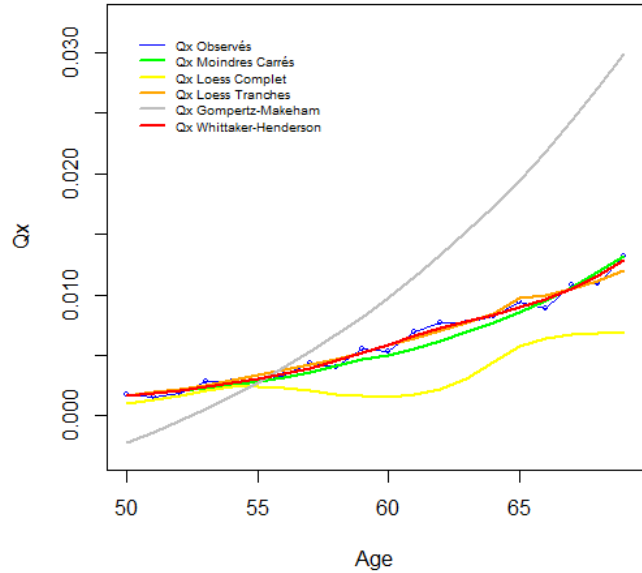


FIGURE 4.33 – Comparaison des lissages réalisés - Hommes OD Tranche 2

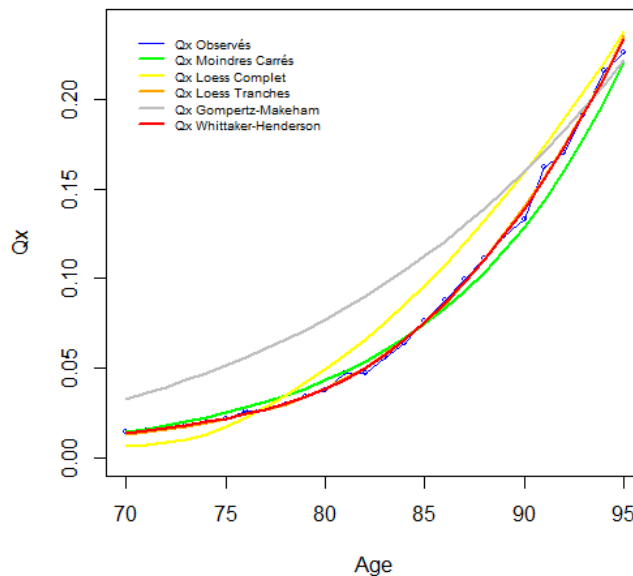


FIGURE 4.34 – Comparaison des lissages réalisés - Hommes OD Tranche 3

Les figures ci-dessus permettent d'écarter l'hypothèse que le lissage par les moindres carrés est le plus proche des taux bruts du fait de petits écarts observables à plusieurs endroits entre la courbe de lissage et les taux bruts, mais ne nous permet pas de distinguer le lissage de Loess par tranches du lissage de Whittaker-Henderson. Toutefois, sur la FIGURE 4.33, le lissage de Whittaker-Henderson semble plus lisse que celui de Loess par tranches, ce qui semble plus cohérent avec la réalité. C'est pourquoi comme dans le cas des femmes OD, le lissage qui semble le mieux adapté à la population des hommes OD est le lissage par la méthode de Whittaker-Henderson.

— Cas des ayants droit :

Maintenant étudions le cas des femmes AD. Cette fois-ci, la tranche 1 commence à 27 ans au lieu de 25 dans le cas des OD car le nombre d'individus d'âges précédents est nul.

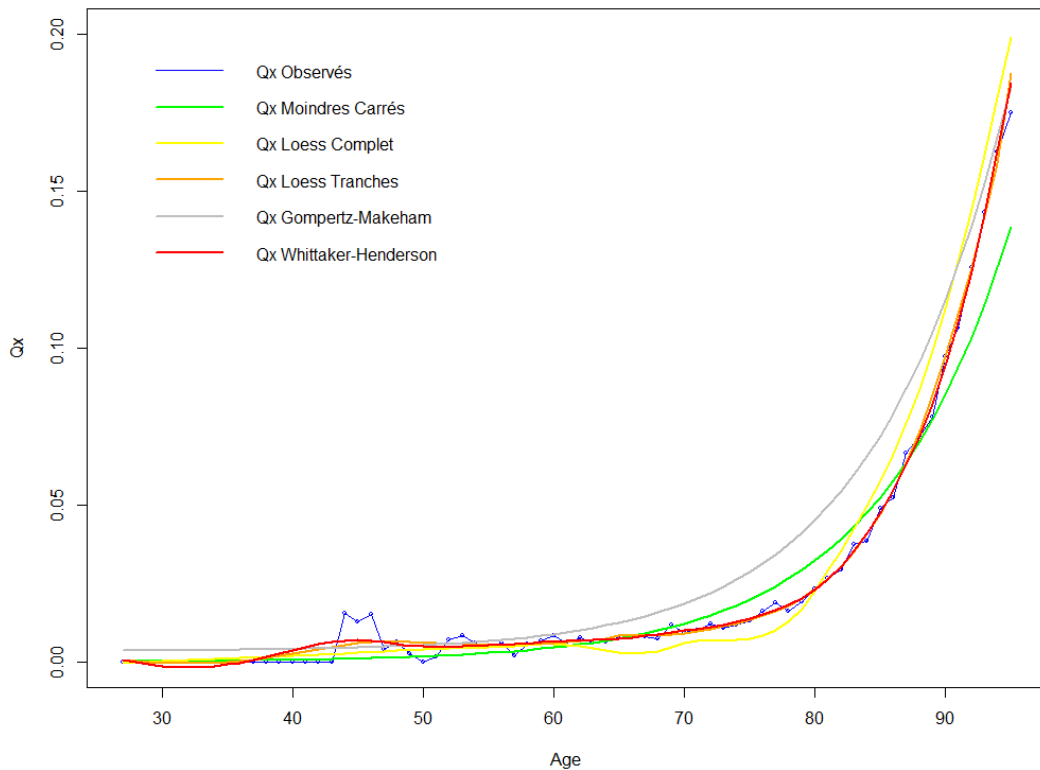


FIGURE 4.35 – Comparaison des lissages réalisés - Femmes AD Complet

Sur la FIGURE 4.35, les méthodes de Loess par tranches et de Whittaker-Henderson semblent à nouveau être les plus pertinentes aux taux bruts, mais sont trop proches l'une de l'autre pour en distinguer une plus que l'autre. Il est alors nécessaire de se tourner vers des plages d'observations plus restreintes afin de ne retenir qu'une seule des différentes méthodes proposées.

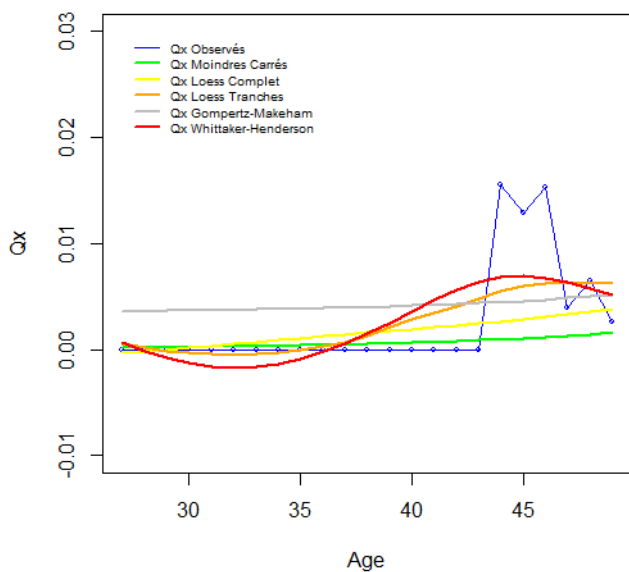


FIGURE 4.36 – Comparaison des lissages réalisés - Femmes AD Tranche 1

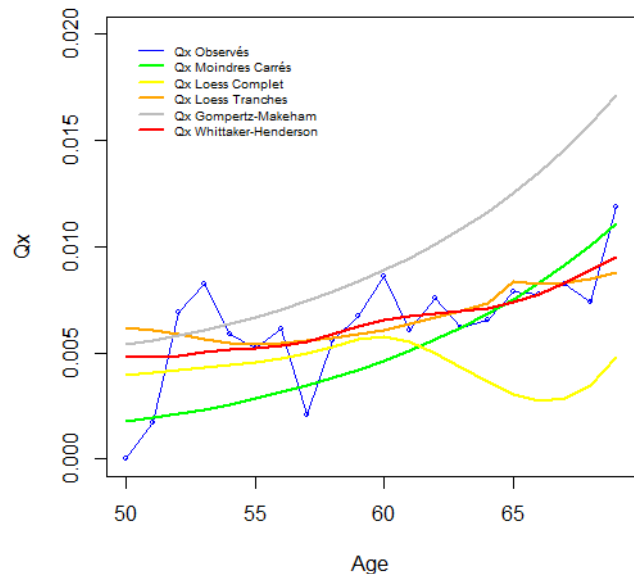


FIGURE 4.37 – Comparaison des lissages réalisés - Femmes AD Tranche 2

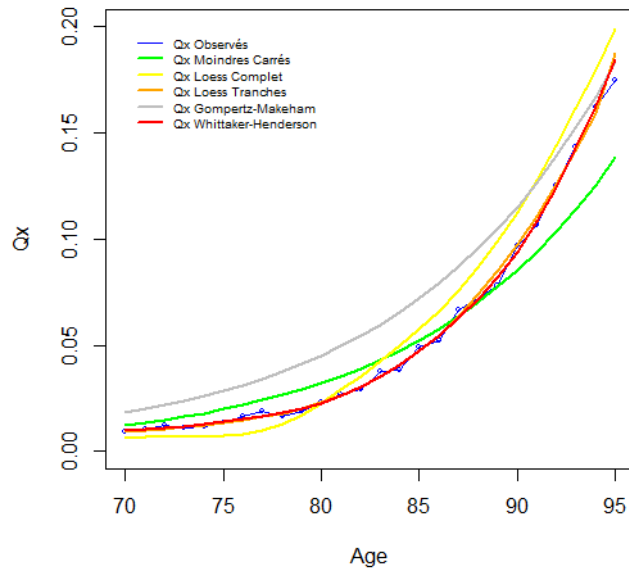


FIGURE 4.38 – Comparaison des lissages réalisés - Femmes AD Tranche 3

Les trois figures précédentes ne nous permettent pas d'évaluer une méthode meilleure qu'une autre parmi les deux sélectionnées, même si la courbe de Loess par tranches semble moins lisse que celle de Whittaker-Henderson, cette dernière semble moins bien s'accorder aux taux bruts aux bas âges. Dans ce cas-ci, les autres méthodes de comparaison explicitées un peu plus loin dans ce rapport nous éclaireront mieux sur la méthode à retenir.

Maintenant, tournons-nous vers les hommes AD. Notons que la tranche 1 commence à 35 ans au lieu de 25 dans le cas des OD car le nombre d'individus d'âges inférieurs est nul. Cette dernière n'est pas représentée ici car les taux bruts sont tous nuls sur cette plage d'âges, et donc il est plus pertinent de regarder les deux autres.

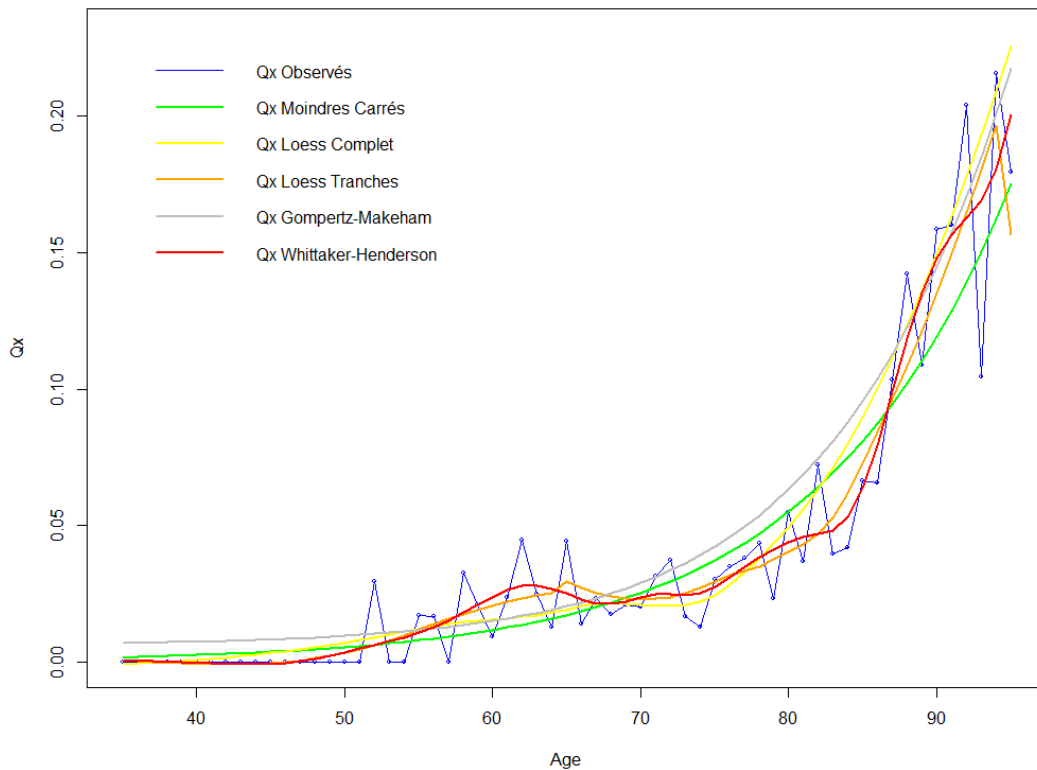


FIGURE 4.39 – Comparaison des lissages réalisés - Hommes AD Complet

En étudiant la population entière des hommes AD, il est à nouveau difficile de distinguer une méthode de lissage en particulier, les méthodes des moindres carrés, de Loess complet et de Gompertz-Makeham semblent un peu moins bien respecter la fidélité aux observations à côté des méthodes de Loess par tranches et Whittaker-Henderson.

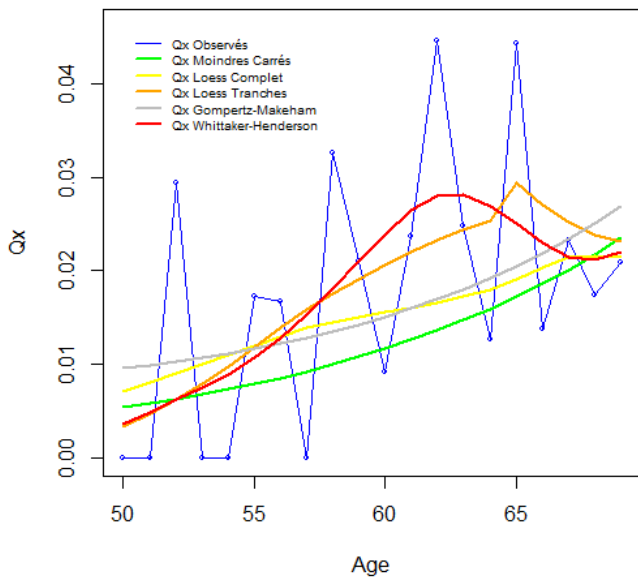


FIGURE 4.40 – Comparaison des lissages réalisés - Hommes AD Tranche 2

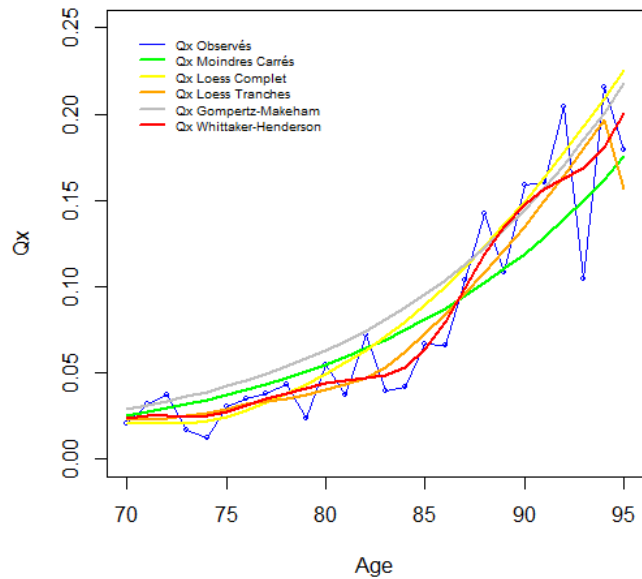


FIGURE 4.41 – Comparaison des lissages réalisés - Hommes AD Tranche 3

La Tranche 3 (FIGURE 4.41) nous permet de confirmer l'écartement des méthodes des moindres carrés, de Loess complet et de Gompertz-Makeham car ces lissages font l'objet d'une sur-estimation des taux de mortalité sur cette plage d'âges. Par ailleurs, d'après les résultats de la Tranche 2 (FIGURE 4.40), le lissage de Whittaker-Henderson semble mieux adapté aux taux bruts que le lissage de Loess par tranches car ce dernier présente de brusques variations qui apparaissent peu conformes à la réalité.

Les comparaisons graphiques réalisées précédemment nous permettent de constater que quelle que soit la population considérée (Femmes, Hommes, OD, AD), les méthodes de lissages qui semblent les mieux adaptées sont les méthodes de Loess par tranches et de Whittaker-Henderson. Cependant, il n'est pas toujours évident de déterminer laquelle des deux est la meilleure pour notre étude, c'est pourquoi nous réalisons maintenant des comparaisons statistiques permettant d'évaluer la qualité de chacun des lissages réalisés. Les comparaisons statistiques sont effectuées sur l'ensemble de chacune des quatre populations, puis sur la tranche 60-89 ans qui correspond à la plage d'âges pour laquelle la population est la plus nombreuse dans chacune des catégories et donc où le lissage doit être le mieux adapté.

La première grandeur calculée est le coefficient de détermination de chacun des lissages à nos taux bruts. La formule utilisée pour le calcul de ce critère est la même que celle utilisée lors du lissage par la méthode des moindres carrés en SECTION 4.1.1.1, l'objectif étant de rechercher une valeur du R^2 la plus proche de 1.

Les résultats obtenus pour chacune des populations sur l'ensemble des individus pour chacun des lissages sont présentés dans le TABLEAU 4.4 présent à la page suivante.

4.1. CRÉATION D'UNE TABLE DE MORTALITÉ PAR DES MÉTHODES NON-PARAMÉTRIQUES

	Moindres carrés	Makeham	Loess	Loess Tranches	Whittaker-Henderson
Femmes OD	0.939	0.918	0.903	0.992	0.993
Hommes OD	0.993	0.877	0.973	0.9988	0.9991
Femmes AD	0.850	0.905	0.914	0.939	0.935
Hommes AD	0.844	0.856	0.859	0.860	0.855

TABLE 4.4 – Coefficients de détermination R^2 obtenus pour chaque lissage et chaque population

Selon ces résultats, les coefficients de détermination sont plutôt bons (au-delà de 0.9) à l'exception des hommes AD (catégorie pour laquelle le nombre d'individus est très faible), mais nous relevons que le lissage par la méthode de Whittaker semble mieux adapté à l'ensemble de la population des OD tandis que le lissage par la méthode de Loess par tranches paraît plus correspondre à l'ensemble de la population des AD.

Regardons maintenant les résultats obtenus sur la tranche 60-89 ans :

	Moindres carrés	Makeham	Loess	Loess Tranches	Whittaker-Henderson
Femmes OD	0.962	0.522	0.495	0.985	0.989
Hommes OD	0.988	0.321	0.874	0.9985	0.9989
Femmes AD	0.932	0.342	0.862	0.990	0.994
Hommes AD	0.690	0.541	0.690	0.806	0.839

TABLE 4.5 – Coefficients de détermination R^2 obtenus pour chaque lissage et chaque population pour la tranche 60-89 ans

Dans ce cas, il est possible de constater que les méthodes de Loess par tranches et de Whittaker-Henderson sont clairement plus performantes vis à vis de nos taux bruts, avec un léger avantage à la méthode de Whittaker-Henderson pour l'ensemble des quatre populations.

La deuxième grandeur que nous calculons est le χ^2 pour chacune de nos méthodes de lissage. Il s'agit d'un critère permettant d'évaluer la qualité de l'ajustement du modèle effectué. Sa valeur est donnée par la relation suivante :

$$\chi^2 = \sum_x E_x \frac{(\hat{q}_x - q_x)^2}{q_x}$$

avec :

- E_x : l'exposition à l'âge x
- \hat{q}_x : le taux de mortalité brut observé à l'âge x
- q_x : le taux de mortalité lissé à l'âge x

La valeur du χ^2 est recherchée de sorte à se rapprocher le plus possible de 0.

Les valeurs du χ^2 obtenues pour chacun des lissages dans le cas de chacune des populations étudiées sont rassemblées dans le tableau suivant :

	Moindres carrés	Makeham	Loess	Loess Tranches	Whittaker-Henderson
Femmes OD	251.03	-824.44	2626.33	84.69	63.62
Hommes OD	145.17	575.93	2678.30	50.38	34.93
Femmes AD	396.46	1189.38	643.99	41.49	30.38
Hommes AD	62.74	66.91	54.37	41.65	35.63

TABLE 4.6 – Valeurs du χ^2 obtenues pour chaque lissage et chaque population

4.1. CRÉATION D'UNE TABLE DE MORTALITÉ PAR DES MÉTHODES NON-PARAMÉTRIQUES

Les valeurs du χ^2 obtenues sur chacune des populations révèlent que les taux lissés de Whittaker-Henderson sont les plus proches des taux bruts, bien que les résultats du lissage de Loess par tranches soient assez proches de ceux de Whittaker-Henderson. Ces déductions sont également observées sur la tranche 60-89 dont le tableau récapitulatif se trouve en ANNEXE E.4.

La dernière grandeur que nous calculons est la *MAPE*. Il s'agit d'un pourcentage d'erreur qui mesure l'exactitude d'un ajustement par rapport aux observations. Sa valeur est donnée par la relation suivante :

$$MAPE = \frac{\sum_x |(\hat{q}_x - q_x) / \hat{q}_x|}{n} \times 100$$

avec :

- \hat{q}_x : les taux de mortalité bruts observés
- q_x : les taux de mortalité lissés
- n : le nombre d'observations

La valeur de la *MAPE* est recherchée de sorte à être la plus faible possible.

Les valeurs de la *MAPE* obtenues pour chacune des méthodes de lissage dans le cas de chaque population étudiée sont regroupées dans le tableau suivant :

	Moindres carrés	Makeham	Loess	Loess Tranches	Whittaker-Henderson
Femmes OD	45.99	529.16	79.83	23.06	21.64
Hommes OD	16.63	1070.10	90.48	13.89	13.01
Femmes AD	59.83	47.97	32.99	15.77	13.87
Hommes AD	25.39	30.96	21.97	24.34	21.75

TABLE 4.7 – Valeurs de la *MAPE* obtenues pour chaque lissage et chaque population

Cette fois encore, les méthodes de Loess par tranches et de Whittaker-Henderson retournent des résultats bien meilleurs que les autres, avec un léger avantage pour Whittaker-Henderson. Ces déductions sont également observées sur la tranche 60-89 dont le tableau récapitulatif se trouve en ANNEXE E.5.

D'après tous ces constats, nous en déduisons que le lissage à adopter pour les taux bruts de chacune de nos populations est la **méthode de Whittaker-Henderson**. La mortalité ne pouvant pas être négative, les taux lissés retenus inférieurs à 0 seront considérés comme nuls dans la suite de l'étude.

4.1.1.6 Validation de la méthode de lissage retenue

Au cours de la section précédente, nous avons déterminé quelle méthode de lissage était plus adaptée au lissage de nos taux bruts. Nous souhaitons maintenant valider la méthode de lissage retenue. Pour cela, nous proposons de réaliser une analyse des résidus.

En effet, pour valider un modèle, les résidus calculés doivent répondre à certaines conditions :

- Les résidus doivent être normalement distribués.
- Il ne doit pas y avoir d'auto-corrélation des résidus.
- L'homoscédasticité des résidus doit être vérifiée.

Dans la perspective de vérifier chacune de ces conditions, nous appliquons trois tests : le test de Shapiro⁸, le test de Box-Pierce ainsi que le test de White.

Chacun de ces tests a un objectif bien précis : le test de Shapiro permet de tester l'hypothèse de normalité des résidus, le test de Box-Pierce permet de tester l'hypothèse d'indépendance de ces derniers, et enfin le test de White permet de tester l'homoscédasticité des résidus. Nous les implémentons sous

8. Samuel Sanford SHAPIRO (1930-) : statisticien et ingénieur américain

4.1. CRÉATION D'UNE TABLE DE MORTALITÉ PAR DES MÉTHODES NON-PARAMÉTRIQUES

le logiciel *RStudio* à l'aide des fonctions *shapiro.test()*, *Box.test()* et *whites.test()*. Nous retrouvons les aspects théoriques de ces trois tests en ANNEXE F.

Commençons par rappeler comment sont déterminés les résidus. Il s'agit des écarts constatés entre les valeurs observées et les valeurs prédites. Dans notre cas, la relation utilisée pour les déterminer est alors :

$$\text{Résidu}_x = \hat{q}_x - q_x$$

avec :

- Résidu_x : le résidu à l'âge x
- \hat{q}_x : le taux de mortalité brut observé à l'âge x
- q_x : le taux de mortalité lissé à l'âge x

Nous relevons pour chacune des populations les valeurs des *p-values* retournées par les tests. Nous considérons que le test est accepté à partir du moment où la *p-value* est supérieure à 5%.

La TABLE 4.8 rassemble les valeurs des *p-values* retournées par les tests appliqués à l'ensemble des données des plages d'étude, ainsi que la statistique du test de Shapiro *W* pour chacune de nos populations :

	Femmes OD	Hommes OD	Femmes AD	Hommes AD
<i>p-values</i> Shapiro	2.427e-13	1.109e-10	0.0027	4.834e-05
Statistique <i>W</i>	0.555	0.7035	0.9407	0.8894
<i>p-values</i> Box-Pierce	0.4019	0.0002	0.115	2.671e-06
<i>p-values</i> White	0.0000	0.0000	0.0000	0.0002

TABLE 4.8 – Résultats retournés par les tests pour chaque population

En appliquant les tests à l'ensemble de la population de chaque catégorie, nous constatons que les hypothèses ne sont acceptées que pour le test de Box-Pierce pour les femmes. Ceci s'explique par le fait que les valeurs extrêmes empêchent les résidus d'être normaux et homoscedastiques. En effet, pour ce qui concerne la normalité des résidus, c'est ce que nous constatons sur les diagrammes *Quantile-Quantile* (ou *QQ-Plot*) des résidus de chaque population (les cas non présentés ici sont en ANNEXE G).

Ce diagramme *Quantile-Quantile* permet de confirmer que les valeurs extrêmes sont à l'origine de la non normalité des résidus sur la population des femmes OD.

QQ Plot des résidus - Femmes OD 25:95 ans

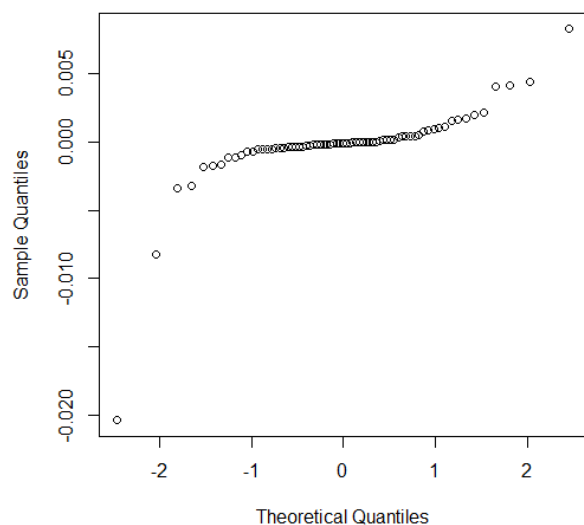


FIGURE 4.42 – Diagramme *Quantile-Quantile* de la population des femmes OD

4.1. CRÉATION D'UNE TABLE DE MORTALITÉ PAR DES MÉTHODES NON-PARAMÉTRIQUES

Par ailleurs, nous constatons que la statistique du test de Shapiro est inférieure à 95% pour l'ensemble des populations, ce qui confirme la non normalité des résidus. Celle de la population des femmes AD est celle qui est la plus grande, les résidus de cette population sont donc ceux dont la distribution se rapproche le plus d'une loi normale.

Si nous sélectionnons la plage d'âges sur laquelle la population est la plus abondante (c'est-à-dire 60-84 ans pour les populations des femmes OD, des femmes AD et des hommes AD, et 60 :87 ans pour la population des hommes OD), nous constatons graphiquement que les résidus approchent une loi normale :

Ce diagramme *Quantile-Quantile* montre une certaine normalité des résidus pour la population des femmes OD sur la tranche d'âges 60-84 ans.

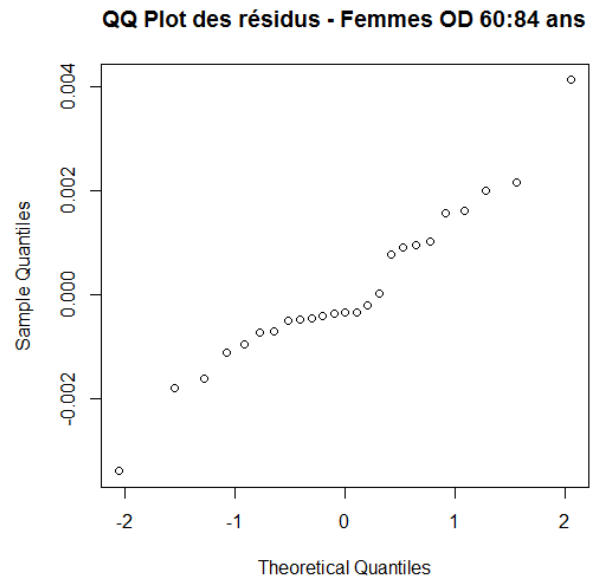


FIGURE 4.43 – Diagramme *Quantile-Quantile* de la population des femmes OD pour la tranche 60-84 ans

En appliquant de nouveau les tests sur les plages d'âges sélectionnées, nous obtenons les résultats suivants :

	Femmes OD	Hommes OD	Femmes AD	Hommes AD
<i>p-values</i> Shapiro	0.2846	0.2006	0.3057	0.4802
Statistique <i>W</i>	0.9525	0.95022	0.9539	0.96313
<i>p-values</i> Box-Pierce	0.3375	0.8259	0.0540	0.0611
<i>p-values</i> White	0.4258	0.1117	0.1428	0.1661

TABLE 4.9 – Résultats retournés par les tests pour chaque population sur les plages d'âges sélectionnées

Les *p-values* sont peu élevées dans l'ensemble mais sont toutes acceptées car supérieures à 5%. Les tests sont donc validés sur les plages d'âges pour lesquelles la population est la plus abondante. Par ailleurs, la statistique de Shapiro a une valeur supérieure à 95% pour l'ensemble des populations, la normalité des résidus est donc acceptée dans chacun des cas.

D'après l'ensemble des résultats retournés par les tests, nous pouvons valider la méthode de Whittaker-Henderson comme lissage des taux bruts de mortalité des populations du régime considéré.

4.1.2 Étude par décalage d'âges

La première méthode d'ajustement que nous appliquons est une méthode par décalage d'âges sur la table de référence. Il s'agit d'appliquer un vieillissement ou un rajeunissement sur la table de référence pour l'adapter à la mortalité de notre population d'étude. Le décalage recherché doit être optimal, et pour cela, nous appliquons des décalages sur différentes tranches d'âges, afin de ne pas généraliser le vieillissement ou le rajeunissement de la table à l'ensemble de la population mais uniquement sur les âges concernés. Par ailleurs, rassembler certains âges pour construire des tranches permet de limiter l'importance donnée aux âges pour lesquels le nombre d'individus est trop faible pour obtenir un bon ajustement.

L'objectif de cet ajustement est donc pour chaque âge, de minimiser le carré de la différence entre le taux brut lissé et le taux de référence "décalé" pour un âge donné. En d'autres termes, nous recherchons $i \in \mathbb{Z}$ tel que la quantité suivante soit minimisée pour chaque âge x :

$$(Taux_{lissé_x} - Taux_{ref_{x+i}})^2$$

Pour une tranche d'âges T , nous cherchons à minimiser la moyenne de ces quantités à chaque âge x qui compose T .

Une fois cette méthode d'ajustement appliquée à nos taux de mortalité, nous obtenons les décalages rassemblés dans les tableaux présentés ci-dessous pour chacune de nos quatre populations. Les graphiques représentent les taux bruts lissés et les taux de référence avec application du décalage optimal sur les différentes tranches d'âges.

Tranche d'âges	Décalage optimal
25 :30 ans	-7
31 :35 ans	-3
36 :40 ans	-5
41 :45 ans	-6
46 :50 ans	-9
51 :55 ans	-5
56 :65 ans	-4
66 :70 ans	-2
71 :75 ans	-3
76 :80 ans	-1
81 :95 ans	-2

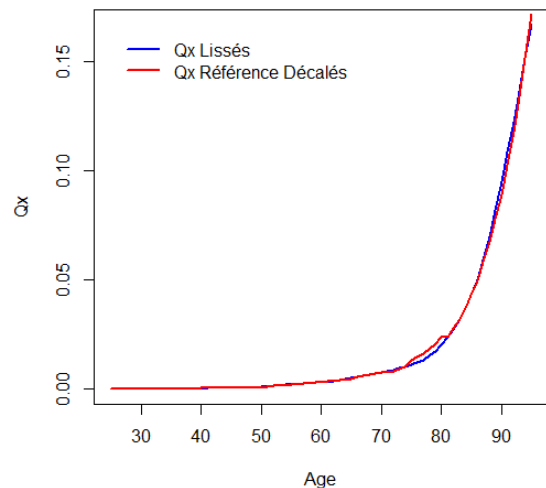


FIGURE 4.44 – Décalages optimaux - Femmes OD

Tranche d'âges	Décalage optimal
25 :30 ans	-10
31 :35 ans	-16
36 :40 ans	-18
41 :45 ans	-16
46 :50 ans	-10
51 :55 ans	-8
56 :60 ans	-7
61 :70 ans	-6
71 :75 ans	-3
76 :85 ans	-2
86 :95 ans	-1

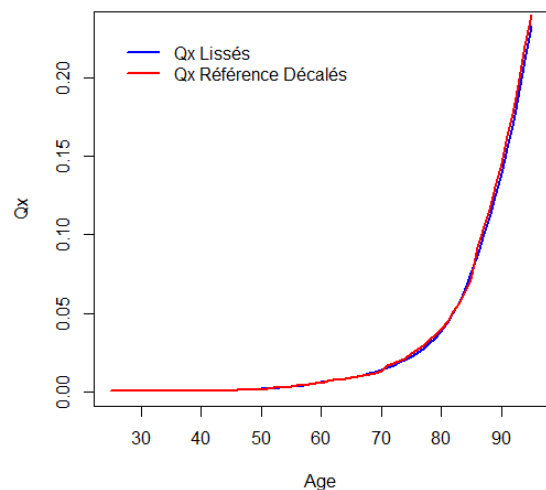


FIGURE 4.45 – Décalages optimaux - Hommes OD

Tranche d'âges	Décalage optimal
27 :74 ans	0
75 :85 ans	-1
86 :90 ans	-2
91 :95 ans	-1

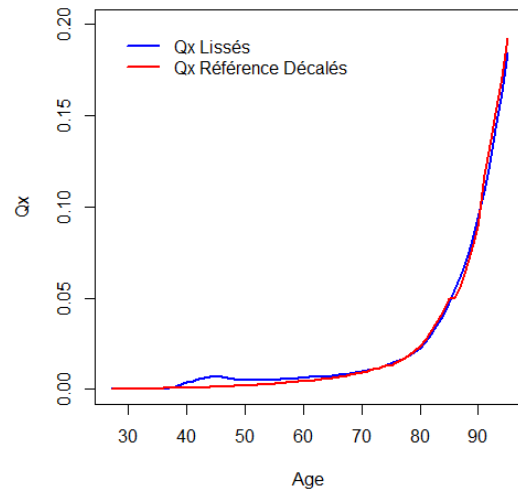


FIGURE 4.46 – Décalages optimaux - Femmes AD

Tranche d'âges	Décalage optimal
35 :75 ans	0
76 :80 ans	-1
81 :85 ans	-3
86 :90 ans	-1
91 :95 ans	-2

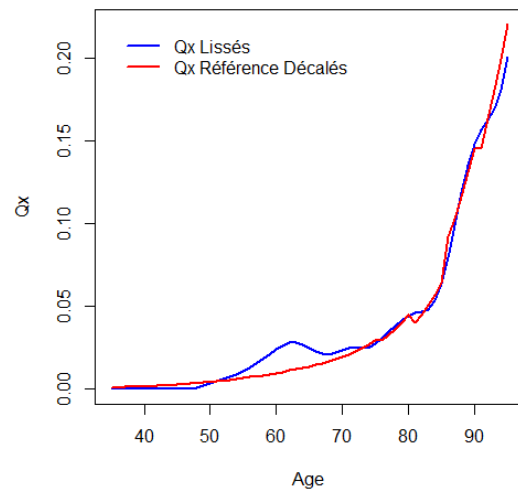


FIGURE 4.47 – Décalages optimaux - Hommes AD

Un décalage négatif signifie que nous "rajeunissons" la table de référence, c'est-à-dire que le nouveau taux de mortalité d'un âge x décalé de $-n$ ans correspondra au taux de mortalité de l'âge $x - n$ de la table de référence initiale.

Nous remarquons que les décalages optimaux sont beaucoup moins élevés dans les cas des populations des ayants droit que pour celles des ouvrants droit. La mortalité de nos AD est alors plus proche de la mortalité de référence que la mortalité de nos OD. Cependant, les taux de mortalité de référence décalés de façon optimale se rapprochent beaucoup mieux des taux lissés dans le cas des OD que dans le cas des AD. En effet, pour les populations des ayants droit, de légers écarts sont observables à presque tous les âges entre les taux lissés et les taux de référence décalés, notamment chez les hommes. Une autre méthode d'ajustement pourrait peut-être mieux s'adapter à ces taux.

Dans le tableau suivant, nous regroupons les valeurs du χ^2 , de la $MAPE$ et du R^2 calculées pour chacune de nos quatre populations (valeurs calculées entre les taux lissés et les taux de référence décalés) :

	χ^2	$MAPE$	R^2
Femmes OD	21.6498	57.4089	0.9980
Hommes OD	43.4073	10.1535	0.9975
Femmes AD	352.3914	41.5027	0.9293
Hommes AD	49.1934	33.3443	0.9097

TABLE 4.10 – Valeurs d'intérêts obtenues sur les taux de référence décalés

Les constatations effectuées un peu plus haut semblent se confirmer : mise à part la MAPE chez les femmes OD, les valeurs d'intérêts retournées pour les populations des ouvrants droit sont globalement meilleures que celles retournées pour les ayants droit; de plus, la population des hommes AD a le R^2 le plus faible, il s'agit donc bien de la population dont les taux de référence décalés se rapprochent le moins des taux lissés.

4.1.3 Étude par abattement de la table de référence

La deuxième méthode d'ajustement que nous appliquons est une méthode par abattement de la table de référence. Il s'agit d'appliquer un coefficient d'abattement sur la table de référence pour l'adapter à la mortalité de notre population d'étude. Le coefficient recherché doit être optimal, et pour cela, nous appliquons des abattements sur différentes tranches d'âges de la même manière que lors de l'application de la première méthode.

L'objectif de cet ajustement est donc pour chaque âge, de minimiser le carré de la différence entre le taux brut lissé et le taux de référence "abattu" pour un âge donné. En d'autres termes, nous recherchons $i \in [0; 1]$ tel que la quantité suivante soit minimisée pour chaque âge x :

$$(Taux_{lissé_x} - Taux_{ref_x} \times (1 - i))^2$$

Pour une tranche d'âges T , nous cherchons à minimiser la moyenne de ces quantités à chaque âge x qui compose T .

Une fois cette méthode d'ajustement appliquée à nos taux de mortalité, nous obtenons les coefficients d'abattement rassemblés dans les tableaux présentés ci-dessous pour chacune de nos quatre populations. Les graphiques représentent les taux bruts lissés et les taux de référence avec application de l'abattement optimal sur les différentes tranches d'âges.

Tranche d'âges	Abattement optimal
25 :30 ans	53%
31 :35 ans	23%
36 :40 ans	43%
41 :45 ans	50%
46 :50 ans	60%
51 :55 ans	34%
56 :60 ans	27%
61 :65 ans	23%
66 :70 ans	15%
71 :75 ans	21%
76 :80 ans	27%
81 :85 ans	26%
86 :95 ans	21%

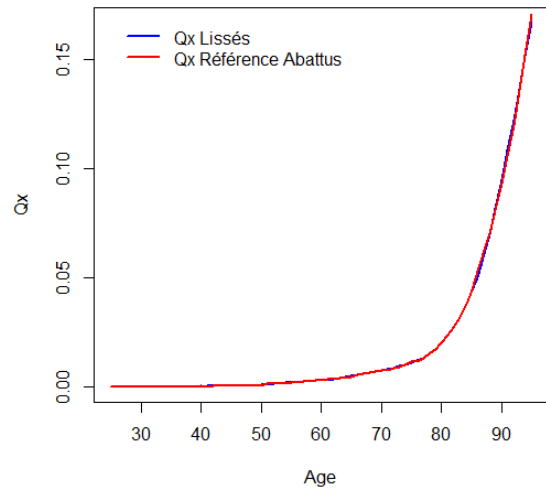


FIGURE 4.48 – Abattements optimaux - Femmes OD

Les résultats constatés sur les FIGURES 4.1.3 et 4.1.3 sont à peu près équivalents à ceux de l'ajustement par la première méthode, les taux de référence abattus semblent plutôt bien s'accorder aux taux lissés dans le cas des ouvrants droits.

Cette fois encore, les abattements à appliquer sont plus élevés chez les OD que chez les AD, ce qui est cohérent avec l'analyse par décalage d'âges : la mortalité de la population des AD est plus proche de la mortalité de référence que celle des OD.

De même, les taux de référence abattus pour les populations des AD semblent moins proches des taux lissés que les taux de référence abattus pour les populations des OD, et particulièrement dans le cas des hommes.

Tranche d'âges	Abattement optimal
25 :30 ans	60%
31 :35 ans	64%
36 :40 ans	56%
41 :45 ans	64%
46 :50 ans	60%
51 :55 ans	55%
56 :60 ans	45%
61 :65 ans	36%
66 :70 ans	30%
71 :75 ans	22%
76 :80 ans	21%
81 :85 ans	18%
86 :90 ans	14%
91 :95 ans	12%

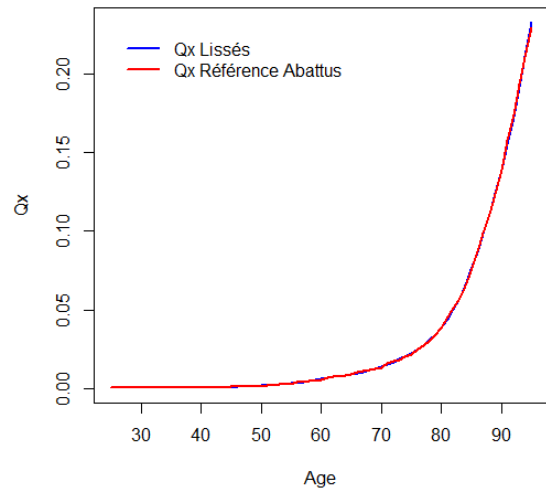


FIGURE 4.49 – Abattements optimaux - Hommes OD

Tranche d'âges	Abattement optimal
27 :70 ans	0%
71 :75 ans	1%
76 :80 ans	14%
81 :85 ans	18%
86 :95 ans	19%

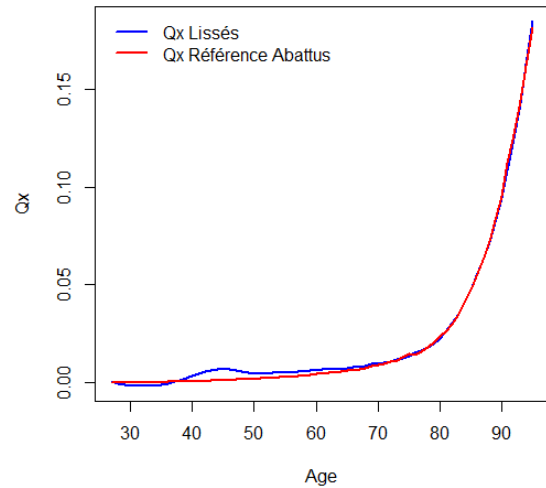


FIGURE 4.50 – Abattements optimaux - Femmes AD

Tranche d'âges	Abattement optimal
35 :75 ans	0%
76 :80 ans	4%
81 :85 ans	18%
86 :90 ans	10%
91 :95 ans	12%

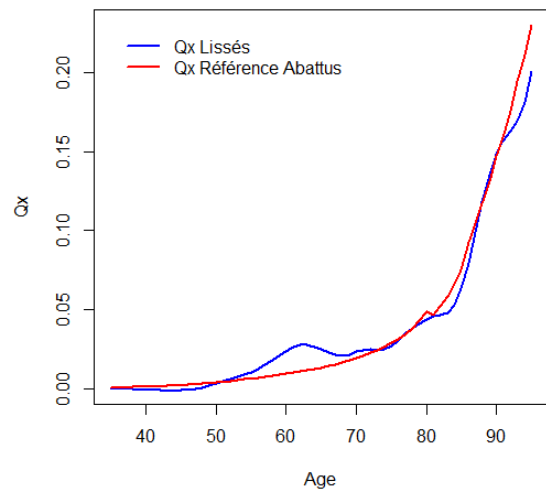


FIGURE 4.51 – Abattements optimaux - Hommes AD

4.2. CRÉATION D'UNE TABLE DE MORTALITÉ À L'AIDE DU PACKAGE ELT

Dans le tableau suivant, nous regroupons les valeurs du χ^2 , de la *MAPE* et du R^2 calculées pour chacune de nos quatre populations (valeurs calculées entre les taux lissés et les taux de référence abattus) :

	χ^2	<i>MAPE</i>	R^2
Femmes OD	5.1430	34.5246	0.9995
Hommes OD	9.0588	3.1117	0.9998
Femmes AD	321.1149	40.4689	0.9341
Hommes AD	51.7299	34.4265	0.8960

TABLE 4.11 – Valeurs d'intérêts obtenues sur les taux de référence abattus

Les constatations effectuées un peu plus haut semblent se confirmer : les valeurs d'intérêts retournées pour les populations des ouvrants droit sont globalement meilleures que celles retournées pour les ayants droit; de plus, la population des hommes AD a le R^2 le plus faible, il s'agit donc bien de la population dont les taux de référence abattus se rapprochent le moins des taux lissés.

Si maintenant nous comparons les valeurs d'intérêt obtenues pour chacune des deux méthodes, en rappelant que le χ^2 et la *MAPE* doivent être recherchés de sorte à être les plus faibles possible et le R^2 de sorte à être le plus proche de 1 possible, nous constatons pour les populations des femmes OD, des hommes OD, et des femmes AD, des résultats qui retournent unanimement la méthode par abattement de la table de référence comme meilleur ajustement. En revanche, pour ce qui est de la population des hommes AD, selon les trois valeurs d'intérêt, c'est la méthode par décalage d'âges sur la table de référence qui semble la mieux adaptée.

Cependant, les différences étant très peu élevées et par souci de conformité, nous choisissons de retenir la méthode par abattement de la table de référence comme ajustement, quelle que soit la population considérée.

4.2 Création d'une table de mortalité à l'aide du *package ELT*

Le *package ELT* disponible sous le logiciel *RStudio* permet la construction de tables de mortalité prospectives par des méthodes de positionnement par rapport à une table de référence. Cette section se voit présenter les différentes méthodes que nous appliquerons ensuite à nos données. Il est à noter que le *package* prévoit quatre méthodes différentes mais nous ne nous concentrerons que sur trois d'entre elles dans cette étude, que nous comparerons ensuite à l'aide de tests statistiques.

L'ensemble des aspects théoriques présentés ci-dessous est issu des notes de travail II1291-12 v1.4 (méthodologie des tests statistiques), ainsi que II1291-12 v1.3 et II1291-14 v1.7 (méthodologie et application des méthodes de positionnement) de l'Institut des Actuariers rédigée par Frédéric PLANCHET et Julien TOMAS, et la table de référence utilisée est toujours la **Table INSEE prospective 2013-2070**.

L'utilisation de ce *package* requiert un format de données spécifique, qui doit alors correspondre au modèle suivant :

Id	Gender	DateOfBirth	DateIn	DateOut	Status
100001	Male	24/05/1945	03/10/1968	31/12/2017	other
100002	Female	22/01/1939	22/01/1957	29/03/2013	deceased
100003	Male	28/09/1951	15/07/1974	01/07/2013	deceased
100004	Male	22/01/1946	08/11/1965	31/12/2017	other
100005	Female	16/07/1953	01/02/1977	31/12/2017	other

TABLE 4.12 – Format des données

avec :

- **Id** : un identifiant quelconque
- **Gender** : le sexe de l'individu (*Male* pour les hommes et *Female* pour les femmes)
- **DateOfBirth** : la date de naissance de l'individu
- **DateIn** : la date d'entrée au régime
- **DateOut** : la date de sortie du régime
- **Status** : le statut de l'individu (*deceased* si l'individu est décédé, *other* sinon)

Du fait du contenu des données dont nous disposons initialement, nous considérons comme date d'entrée au régime la date d'embauche de l'individu. Par ailleurs, la seule façon de "sortir" du régime est de décéder, c'est pourquoi lorsque le statut de l'individu est *deceased*, la date de sortie correspond à la date de décès de l'individu; dans le cas où nous ne disposons pas de la date de décès de l'individu mais que nous savons qu'il est décédé, nous considérons la date de sortie en milieu d'année (c'est-à-dire au 1^{er} juillet de l'année de sortie du régime). D'autre part, lorsque le statut de l'individu est *other*, nous considérons comme date de sortie le dernier jour de la période d'observation, c'est-à-dire le 31 décembre 2017.

4.2.1 Les différentes méthodes proposées : principe et application

Dans cette section, nous reprenons les notations introduites en partie 2.2.3.1 en les adaptant à notre étude.

Notre période d'observation s'étend sur les années 2013 à 2017, par conséquent, nous définissons t une variable correspondant aux différentes années d'observation.

Nous avons alors :

- $l_{x,t}$: le nombre d'individus d'âge x vivants au début de l'année t
- $d_{x,t}$: le nombre d'individus d'âge x décédés au cours de l'année t
- $E_{x,t}$: l'exposition au risque à l'âge x au cours de l'année t

Nous avons les relations adaptées suivantes :

$$d_{x,t} = \sum_{i=1}^{l_{x,t}} \delta_i \quad \text{et} \quad E_{x,t} = \sum_{i=1}^{l_{x,t}} \tau_i$$

L'estimateur de Hoem permet de calculer les taux bruts de mortalité à partir de ces quantités :

$$\hat{q}_{x,t} = \frac{d_{x,t}}{E_{x,t}}$$

où $\hat{q}_{x,t}$ correspond à la probabilité qu'un individu d'âge x décède au cours de l'année t .

Par ailleurs, introduisons les notations suivantes qui seront utiles dans les interprétations qui vont suivre :

- x^* : la plage d'âge commune avec la table de mortalité de référence
- t^* : la période d'observation commune avec la table de mortalité de référence
- ε_{x^*,t^*} : un terme d'erreur
- $q_{x,t}^{ref}$: la probabilité de décéder à l'âge x au cours de l'année t de la table de référence

4.2.1.1 Méthode 1 : Approche à un paramètre avec le SMR

La première méthode de positionnement est une approche à un paramètre avec le *Standardized Mortality Ratio* (SMR).

- *Description de la méthode :*

Il s'agit d'une démarche assez simple consistant en l'application d'un coefficient de réduction ou de majoration aux probabilités conditionnelles de décès de la table de référence, ce de manière multiplicative. Ce coefficient correspond au SMR, qui se calcule de la sorte :

$$SMR = \frac{\sum_{(x^*, t^*)} d_{x^*, t^*}}{\sum_{(x^*, t^*)} E_{x^*, t^*} \times q_{x^*, t^*}^{ref}}$$

Les probabilités de décès ajustées de notre population s'obtiennent de la manière suivante :

$$\tilde{q}_{x, t} = SMR \times q_{x, t}^{ref}$$

La mortalité de notre population est alors ajustée uniquement avec le paramètre SMR, qui traduit la déviation constatée entre les décès observés au sein de notre population et ceux estimés par la table de référence.

- *Application à nos observations :*

La plage d'âges x^* est un choix très important car la table de mortalité qui sera construite par positionnement n'aura de validité que sur la plage d'âges que nous aurons choisie.

C'est pourquoi nous appliquons le calcul du SMR à différentes plages d'âges afin de sélectionner celle qui semble la plus pertinente pour la suite :

Plage d'âges	60 : 100 ans	65 : 100 ans	65 : 95 ans	65 : 90 ans
Femmes	0.80615	0.80657	0.80811	0.80738
Hommes	0.81836	0.83458	0.83141	0.82355

TABLE 4.13 – SMR obtenus pour différentes plages d'âges

Plus la valeur du SMR se rapproche de 1, plus la mortalité de la population étudiée se rapproche de celle prédite par la table de référence.

Nous pouvons constater que du fait que les valeurs du SMR sont inférieures à 1, la population de notre régime a une mortalité plus faible que la population représentée par la table de référence.

Du fait d'une exposition plus large et d'un SMR proche de 1, nous retenons la plage d'âges 65-100 ans.

4.2.1.2 Méthode 2 : Approche à deux paramètres avec un modèle relationnel

La deuxième méthode de positionnement proposée dans cette étude est une approche à deux paramètres avec un modèle relationnel inspiré du modèle de Brass (1971), dont l'ajustement se fait par l'intermédiaire d'une régression logistique.

- *Description de la méthode :*

Dans ce positionnement, les taux de mortalité ajustés sont reliés aux taux de mortalité de référence via la fonction *logit* plus couramment notée *lg* :

$$lg(\tilde{q}_{x^*, t^*}) = \alpha + \beta lg(q_{x^*, t^*}^{ref}) + \varepsilon_{x^*, t^*}$$

Les *logits* des probabilités conditionnelles de décès sont donc ajustés à partir des *logits* des taux de mortalité de référence.

Il est par ailleurs possible de constater que le paramètre α a un impact sur tous les âges tandis que le paramètre β modifie cet impact en fonction de l'âge.

L'ajustement est alors réalisé de sorte à minimiser la distance pondérée par l'exposition qui se trouve entre les probabilités de décès observées \hat{q}_{x^*,t^*} et les probabilités de décès ajustées \tilde{q}_{x^*,t^*} , soit de sorte à minimiser l'expression suivante :

$$\sum_{(x^*,t^*)} |E_{x^*,t^*} \times (\hat{q}_{x^*,t^*} - \tilde{q}_{x^*,t^*})|$$

C'est ainsi que nous obtenons des estimations $\hat{\alpha}$ et $\hat{\beta}$ des paramètres α et β , et les probabilités de décès ajustées de la population observée s'obtiennent via la relation suivante :

$$\forall x \in x^*, t \in t^* : \tilde{q}_{x,t} = \frac{\exp(\hat{\alpha} + \hat{\beta} \times \lg(q_{x,t}^{ref}))}{1 + \exp(\hat{\alpha} + \hat{\beta} \times \lg(q_{x,t}^{ref}))}$$

Notons que ce modèle doit être utilisé avec précaution dans le cas de l'étude du risque de décès car il a tendance à sur- ou sous-estimer les probabilités de décès.

En effet, rappelons que la fonction *logit* est définie comme suit :

$$\forall x \in]0; 1[, \lg(x) = \log\left(\frac{x}{1-x}\right)$$

La fonction *logit* a des valeurs dans l'intervalle $] -\infty ; +\infty[$.

Du fait que nous ayons :

$$\frac{d}{dx} \lg(x) = \frac{1}{x(1-x)} > 0 \quad \forall x \in]0; 1[$$

la fonction *logit* est croissante sur $]0; 1[$.

Par ailleurs, sur l'intervalle $]0; \frac{1}{2}[$, nous avons :

$$\frac{d^2}{dx^2} \lg(x) = \frac{2x-1}{x^2(1-x)^2} < 0$$

La fonction *logit* est concave sur cet intervalle. Cette concavité nous permet d'appliquer l'égalité de Jensen^{9 10} à notre fonction. Les taux de mortalité estimés $\hat{q}_{x,t}$ supposés sans biais vérifient alors :

$$\lg(\hat{q}_{x,t}) \geq \mathbb{E}[\lg(\hat{q}_{x,t})]$$

Ainsi, sur l'intervalle $]0; \frac{1}{2}[$, les *logits* obtenus sous-estiment les *logits* réels.

À l'inverse, sur l'intervalle $]\frac{1}{2}; 1[$, la fonction *logit* est convexe, et donc l'inégalité de Jensen nous permet d'affirmer que les *logits* obtenus sur-estiment les *logits* réels sur cet intervalle.

Toutefois, ces affirmations sont valables dans le cas de l'étude du risque de décès, mais en ce qui concerne notre étude, c'est-à-dire l'étude du risque de survie, ce modèle est source d'une marge de sécurité.

- Application à nos observations :

Nous considérons la plage d'âges 65-100 ans comme déterminée à la méthode précédente.

Les estimations des paramètres α et β pour la plage d'âges considérée sont rassemblées dans le tableau suivant :

9. Johan JENSEN (1859-1925) : mathématicien et ingénieur danois

10. Inégalité de Jensen : pour toute fonction convexe $f : f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)]$, et à l'inverse, pour toute fonction concave $g : g(\mathbb{E}[X]) \geq \mathbb{E}[g(X)]$

	$\hat{\alpha}$	$\hat{\beta}$
Femmes	-0.2626894	0.9870621
Hommes	0.08044813	1.09399952

TABLE 4.14 – Estimations des paramètres α et β pour chacune des populations des OD

4.2.1.3 Méthode 3 : Approche par modèle linéaire généralisé de Poisson incluant des interactions avec l'âge et l'année calendaire

La troisième méthode de positionnement appliquée est un modèle linéaire généralisé de Poisson¹¹ dans lequel nous incluons des interactions avec l'âge et l'année calendaire.

- Description de la méthode :

Notons D_{x^*,t^*} le nombre de décès observés au sein de la population de l'organisme à l'âge x^* et l'année calendaire t^* . Nous supposons alors :

$$D_{x^*,t^*} \sim \mathcal{P}(E_{x^*,t^*} \mu_{x^*,t^*})$$

avec :

$$\mu_{x^*,t^*} = \beta_0 + \beta_1 \log q_{x^*,t^*}^{ref} + \beta_2 x^* + \beta_3 t^* + \beta_4 x^* t^*$$

Les probabilités de décès de la population de notre organisme ajustées $\tilde{q}_{x,t}$ sont obtenues pour tout x dans x^* et pour tout t dans t^* de la sorte :

$$\tilde{q}_{x,t} = \exp(\hat{\beta}_0 + \hat{\beta}_1 \log q_{x,t}^{ref} + \hat{\beta}_2 x + \hat{\beta}_3 t + \hat{\beta}_4 x t)$$

- Application à nos observations :

Cette fois encore, nous considérons la plage d'âges déterminée lors de l'exécution de la première méthode de positionnement, c'est-à-dire la plage 65-100 ans.

La période d'observation reste 2013-2017, la période commune avec la référence est alors inférieure à 10 ans, c'est pourquoi les interactions avec l'année calendaire ne sont pas prises en compte. Par conséquent, seuls trois paramètres sont estimés : $\hat{\beta}_0$, $\hat{\beta}_1$ et $\hat{\beta}_2$. Les résultats sont regroupés dans le tableau suivant :

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
Femmes	2.33898	1.19220	-0.02363
Hommes	-0.457567	1.041609	0.004714

TABLE 4.15 – Estimations des paramètres β_0 , β_1 et β_2 pour chacune des populations des OD

4.2.2 Comparaison des méthodes proposées

Dans un premier temps, comparons les trois méthodes appliquées juste ci-dessus. Nous comparons ensuite les résultats du modèle retenu avec ceux obtenus par abattement de la table de référence.

11. Siméon Denis POISSON (1781-1840) : mathématicien, géomètre et physicien français

4.2.2.1 Méthodologie des tests appliqués

Afin de comparer les trois méthodes proposées ci-dessus, le *package ELT* propose l'application de différents tests statistiques, tels que des tests de proximité ou des tests de régularité d'ajustement. Cette section propose la présentation de ces différents tests d'un point de vue théorique, nous permettant de comprendre et d'interpréter au mieux les résultats qui seront obtenus et présentés dans la section suivante.

Commençons par regarder les **critères de proximité** :

- Test du rapport de vraisemblance :

Le premier test appliqué est un test du rapport de vraisemblance. Dans ce cas, nous testons les hypothèses suivantes :

- H_0 : l'ajustement effectué correspond à la loi de mortalité sous-jacente ;
- H_1 : l'ajustement effectué ne correspond pas à la loi de mortalité sous-jacente.

En faisant l'hypothèse que les décès suivent une loi de Poisson $D_{x,t} \sim \mathcal{P}(E_{x,t} q_{x,t})$, la statistique du test du rapport de vraisemblance se traduit comme suit :

$$\xi^{LR} = \sum_{(x,t)} \left(D_{x,t} \ln \left(\frac{D_{x,t}}{E_{x,t} \tilde{q}_{x,t}} \right) - (D_{x,t} - E_{x,t} \tilde{q}_{x,t}) \right)$$

avec $\tilde{q}_{x,t}$ les taux de mortalité ajustés.

Si H_0 est acceptée, la statistique du test est telle que :

$$\xi^{LR} \sim \chi^2(n)$$

avec n correspondant au nombre d'observations.

H_0 sera alors rejetée dans le cas où :

$$\xi^{LR} > \chi^2_{(1-\alpha)}(n)$$

avec $\chi^2_{(1-\alpha)}(n)$ correspondant au quantile $(1 - \alpha)$ de la distribution du χ^2 à n degrés de liberté.

La p -valeur¹² (ou *p-value* en anglais), de formule ci-dessous, est recherchée de sorte à être la plus proche possible de 1 :

$$p\text{-valeur} = \mathbb{P}[\chi^2_{(1-\alpha)}(n) > \xi^{LR}] = 1 - F_{\chi^2(n)}(\xi^{LR})$$

avec $F_{\chi^2(n)}(\xi^{LR})$ la fonction de répartition de la loi $\chi^2(n)$ au point ξ^{LR} .

- Test du SMR :

Le deuxième test est basé sur le rapport entre le nombre de décès observés et le nombre de décès obtenu par ajustement. Il s'agit du test du SMR, en gardant l'hypothèse que les décès suivent une loi de Poisson $D_{x,t} \sim \mathcal{P}(E_{x,t} q_{x,t})$:

$$SMR = \frac{\sum_{(x,t)} D_{x,t}}{\sum_{(x,t)} E_{x,t} \tilde{q}_{x,t}}$$

avec $\tilde{q}_{x,t}$ les taux de mortalité ajustés.

Ce qui signifie une sous-estimation des taux de décès ajustés dans le cas où le SMR est supérieur à 1, et une sur-estimation dans le cas où le SMR est inférieur à 1.

12. La p -valeur correspond à la plus petite valeur de α pour laquelle l'hypothèse H_0 est rejetée

Les hypothèses que nous testons sont donc :

- H_0 : le SMR n'est pas significativement différent de 1 ;
- H_1 : le SMR est significativement différent de 1.

La statistique du test du SMR calculée est :

$$\left\{ \begin{array}{l} \xi^{SMR} = 3 \times D^{\frac{1}{2}} \left(1 - (9D)^{-1} - \left(\frac{D}{E} \right)^{\frac{1}{3}} \right), \text{ si } SMR > 1 \\ \xi^{SMR} = 3 \times D^{*\frac{1}{2}} \left(\left(\frac{D^*}{E} \right)^{\frac{1}{3}} + (9D^*)^{-1} - 1 \right), \text{ si } SMR < 1 \end{array} \right.$$

avec :

- $D = \sum_{(x,t)} D_{x,t}$
- $D^* = \sum_{(x,t)} D_{x,t} + 1$
- $E = \sum_{(x,t)} -E_{x,t} \tilde{q}_{x,t}$

Si H_0 est acceptée, la statistique du test est telle que :

$$\xi^{SMR} \sim \mathcal{N}(0, 1)$$

H_0 sera alors rejetée dans le cas où :

$$\xi^{SMR} > \mathcal{N}_{1-\alpha}(0, 1)$$

avec $\mathcal{N}_{1-\alpha}(0, 1)$ correspondant au quantile $(1 - \alpha)$ de la distribution normale centrée réduite.

La p -valeur (ou p -value en anglais), de formule ci-dessous, est recherchée de sorte à être la plus proche possible de 1 :

$$p\text{-valeur} = \mathbb{P}[\mathcal{N}_{1-\alpha}(0, 1) > \xi^{SMR}] = 1 - F_{\mathcal{N}(0,1)}(\xi^{SMR})$$

avec $F_{\mathcal{N}(0,1)}(\xi^{SMR})$ la fonction de répartition de la loi normale centrée réduite au point ξ^{SMR} .

- Test de Wilcoxon :

Le test de Wilcoxon est le dernier test de proximité appliqué à nos données. Il utilise la direction et la grandeur des différences observées entre les paires composées par chaque taux de décès observé avec le taux de décès ajusté qui lui correspond.

Les hypothèses que nous testons sont donc :

- H_0 : la médiane de la différence entre chacune des paires est nulle ;
- H_1 : la médiane de la différence entre chacune des paires n'est pas nulle.

Après calcul des différences entre chaque taux de décès observé et son taux de décès ajusté, ces différences sont ordonnées en valeurs absolues par ordre croissant, ceci en faisant abstraction des différences nulles. À chaque différence classée (c'est à dire non nulle), est attribué son rang dans le classement effectué.

Notons w_+ la somme des rangs des différences strictement positives, w_- la somme des rangs des différences strictement négatives, et w la plus grande de ces deux sommes : $w = \max(w_+; w_-)$.

La statistique du test de Wilcoxon est alors :

$$\xi^{WIL} = \frac{(w - \frac{1}{2} - n(n+1)) / 4}{\sqrt{n(n+1)(2n+1)/24}}$$

Si H_0 est acceptée, la statistique du test est telle que :

$$\xi^{WIL} \sim \mathcal{N}(0, 1)$$

H_0 sera alors rejetée dans le cas où :

$$|\xi^{WIL}| > \mathcal{N}_{1-\frac{\alpha}{2}}(0, 1)$$

avec $\mathcal{N}_{1-\frac{\alpha}{2}}(0, 1)$ correspondant au quantile $(1 - \frac{\alpha}{2})$ de la distribution normale centrée réduite.

La p -valeur (ou p -value en anglais), de formule ci-dessous, est recherchée de sorte à être la plus proche possible de 1 :

$$p\text{-valeur} = \mathbb{P} \left[\mathcal{N}_{1-\frac{\alpha}{2}}(0, 1) > |\xi^{WIL}| \right] = 2 \times \left(1 - F_{\mathcal{N}(0,1)}(|\xi^{WIL}|) \right)$$

avec $F_{\mathcal{N}(0,1)}(|\xi^{WIL}|)$ la fonction de répartition de la loi normale centrée réduite au point $|\xi^{WIL}|$.

- Valeurs d'intérêt :

Une fois les trois tests précédents mis à exécution, il sera possible d'analyser les valeurs d'intérêt suivantes : le χ^2 , le R^2 , la *MAPE* (*Mean Absolute Percentage Error* ou Erreur Absolue Moyenne en Pourcentage) et la déviance.

Le χ^2 , le R^2 et la *MAPE* ont déjà été définis et présentés dans la Section 4.1.1.5, un rappel sur les formules utilisées avec les notation de la section en cours sont disponibles en Annexe I.

En ce qui concerne la déviance, elle mesure la qualité de l'ajustement. Toujours en supposant que les décès suivent une loi de Poisson $D_{x,t} \sim \mathcal{P}(E_{x,t} q_{x,t})$, nous avons la déviance qui s'exprime de la façon suivante :

$$\begin{cases} Dev_{x,t} = 2 \times \left(D_{x,t} \ln \left(\frac{\hat{q}_{x,t}}{\tilde{q}_{x,t}} \right) - (D_{x,t} - E_{x,t} \tilde{q}_{x,t}) \right), & \text{si } D_{x,t} > 0 \\ Dev_{x,t} = 2 \times E_{x,t} \tilde{q}_{x,t}, & \text{si } D_{x,t} = 0 \end{cases}$$

avec $Dev_{x,t}$ la déviance observée à l'âge x l'année t , et $\tilde{q}_{x,t}$ les taux de mortalité ajustés.

La déviance totale Dev_{tot} s'obtient par :

$$Dev_{tot} = \sum_{(x,t)} Dev_{x,t}$$

La valeur de Dev_{tot} est recherchée de sorte à être la plus faible possible.

Maintenant regardons les **critères de régularité d'ajustement** :

- Test des *runs* :

Le premier test de régularité appliqué est un test des *runs*. Ce test est non-paramétrique et permet de déterminer si les éléments d'une séquence sont mutuellement indépendants ou non. Un *run* est défini comme étant le segment d'une séquence, construit avec des éléments avoisinants qui sont égaux.

Considérons par exemple la séquence suivante :

$$\{++++-----+ +----++++\}$$

Cette séquence est composée de 16 éléments et de 5 *runs*, dont 3 *runs* composés de "+" et 2 *runs* composés de "-".

Les hypothèses que nous testons sont alors :

- H_0 : En considérant une séquence de n éléments, le nombre de *runs* de cette séquence est une variable aléatoire dont la distribution conditionnelle connaissant les nombres n_+ et n_- de signes respectivement positifs et négatifs, avec $n = n_+ + n_-$, est approximativement normale avec comme paramètres :

$$\mu = \frac{2n_+n_-}{n_+ + n_-} + 1 \quad \text{et} \quad \sigma^2 = \frac{2n_+n_-(2n_+n_- - (n_+ + n_-))}{(n_+ + n_-)^2 (n_+ + n_- - 1)} \quad ;$$

- H_1 : Le nombre de *runs* de la séquence n'est pas une variable aléatoire de distribution conditionnelle normale.

La statistique du test des *runs* se traduit comme suit :

$$\xi^{RUN} = \frac{NbRuns - \mu}{\sigma}$$

avec :

- $NbRuns$: le nombre de *runs* contenus par la séquence
- μ et σ : les paramètres de la distribution conditionnelle normale

Si H_0 est acceptée, la statistique du test est telle que :

$$\xi^{RUN} \sim \mathcal{N}(0, 1)$$

H_0 sera alors rejetée dans le cas où :

$$|\xi^{RUN}| > \mathcal{N}_{1-\frac{\alpha}{2}}(0, 1)$$

avec $\mathcal{N}_{1-\frac{\alpha}{2}}(0, 1)$ correspondant au quantile $(1 - \frac{\alpha}{2})$ de la distribution normale centrée réduite.

La *p*-valeur (ou *p-value* en anglais), de formule ci-dessous, est recherchée de sorte à être la plus proche possible de 1 :

$$p\text{-valeur} = \mathbb{P} \left[\mathcal{N}_{1-\frac{\alpha}{2}}(0, 1) > |\xi^{RUN}| \right] = 2 \times \left(1 - F_{\mathcal{N}(0,1)}(|\xi^{RUN}|) \right)$$

avec $F_{\mathcal{N}(0,1)}(|\xi^{RUN}|)$ la fonction de répartition de la loi normale centrée réduite au point $|\xi^{RUN}|$.

- Test des signes :

Le deuxième test de régularité appliqué ici est un test des signes. Ce test est très semblable au test de Wilcoxon vu auparavant, il réalise une analyse de la fréquence des changements de signes des différences observées entre les paires composées par chaque taux de décès observé avec le taux de décès qui lui correspond. Ici, la grandeur de ces différences n'est pas prise en compte.

Les hypothèses que nous testons sont alors :

- H_0 : la médiane entre les signes positifs et négatifs de la différence entre chacune des paires est nulle;
- H_1 : la médiane entre les signes positifs et négatifs de la différence entre chacune des paires n'est pas nulle.

Comme pour le test des *runs*, nous considérons n_+ et n_- comme étant respectivement le nombre de signes positifs et le nombre de signes négatifs, ainsi que $n = n_+ + n_-$

La statistique du test des signes se traduit comme suit :

$$\xi^{SIG} = \frac{|n_+ + n_-| - 1}{\sqrt{n}}$$

4.2. CRÉATION D'UNE TABLE DE MORTALITÉ À L'AIDE DU PACKAGE ELT

Si H_0 est acceptée, la statistique du test est telle que :

$$\xi^{SIG} \sim \mathcal{N}(0, 1)$$

H_0 sera alors rejetée dans le cas où :

$$|\xi^{SIG}| > \mathcal{N}_{1-\frac{\alpha}{2}}(0, 1)$$

avec $\mathcal{N}_{1-\frac{\alpha}{2}}(0, 1)$ correspondant au quantile $(1 - \frac{\alpha}{2})$ de la distribution normale centrée réduite.

La p -valeur (ou p -value en anglais), de formule ci-dessous, est recherchée de sorte à être la plus proche possible de 1 :

$$p\text{-valeur} = \mathbb{P} \left[\mathcal{N}_{1-\frac{\alpha}{2}}(0, 1) > |\xi^{SIG}| \right] = 2 \times \left(1 - F_{\mathcal{N}(0,1)}(|\xi^{SIG}|) \right)$$

avec $F_{\mathcal{N}(0,1)}(|\xi^{SIG}|)$ la fonction de répartition de la loi normale centrée réduite au point $|\xi^{SIG}|$.

4.2.2.2 Comparaison des résultats obtenus

Nous présentons ici l'ensemble des résultats obtenus pour chacun des tests présentés ci-dessus, appliqués aux trois méthodes de positionnement mises en application sur la population des OD (les résultats de la population des AD sont présentés en ANNEXE J).

- Test du rapport de vraisemblance :

TEST DE VRAISEMBLANCE	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
χ^2	218.26	199.2	217.3	439.18	52.49	373.59
Seuil	212.3	212.3	212.3	212.3	212.3	212.3
p-value	0.0272	0.1555	0.0301	0	1	0
Hypothèse retenue	H_1	H_0	H_1	H_1	H_0	H_1

TABLE 4.16 – Résultats obtenus pour le test du rapport de vraisemblance par sexe pour chaque méthode

Pour chacune des deux populations, seule l'utilisation de la deuxième méthode est acceptée. Pour ce qui concerne les hommes, la p -value est excellente, l'ajustement par la méthode 2 semble alors parfaitement s'ajuster à notre population masculine. En revanche, en ce qui concerne la population des femmes, la méthode 2 est acceptée mais la p -value reste très faible. La grande dispersion des données pour cette population est un facteur pouvant expliquer cette faible adéquation de l'ajustement à nos données.

- Test du SMR :

TEST DU SMR	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
SMR	1	0.9928	1	1	0.9723	1
χ^2	0.0064	0.3651	0.0128	0.0031	3.0512	0.0062
Seuil	1.6449	1.6449	1.6449	1.6449	1.6449	1.6449
p-value	0.4975	0.3575	0.4949	0.4988	0.0011	0.4975
Hypothèse retenue	H_0	H_0	H_0	H_0	H_1	H_0

TABLE 4.17 – Résultats obtenus pour le test du SMR par sexe pour chaque méthode

4.2. CRÉATION D'UNE TABLE DE MORTALITÉ À L'AIDE DU PACKAGE ELT

Pour ce test-ci, l'ensemble des *p-values* est acceptable, à l'exception de celle de la méthode 2 pour les hommes, ce qui semble contredire les résultats du test précédent. Par ailleurs, le SMR traduisant le rapport entre le nombre de décès observés et le nombre de décès ajustés, il découle des résultats de ce test que pour les tables construites par les méthodes 1 et 3, le nombre de décès ajusté représente assez bien la réalité.

- Test de Wilcoxon :

TEST DE WILCOXON	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
<i>w</i>	8224	8429	8264	9272	9068	8873
χ^2	0.1121	0.405	0.1693	1.6092	1.3178	1.0392
Seuil	1.96	1.96	1.96	1.96	1.96	1.96
<i>p-value</i>	0.9107	0.6855	0.8656	0.1076	0.1876	0.2987
Hypothèse retenue	H_0	H_0	H_0	H_0	H_0	H_0

TABLE 4.18 – Résultats obtenus pour le test de Wilcoxon par sexe pour chaque méthode

Nous constatons que les *p-values* sont bien meilleures pour la population des femmes que pour celle des hommes, quelle que soit la méthode utilisée. La méthode 1 pour les femmes et la méthode 3 pour les hommes semblent les mieux adaptées à ces populations selon ce test.

- Valeurs d'intérêt :

VALEURS D'INTÉRÊT	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
χ^2	231.58	230.75	231.69	472.66	409.88	402.55
R^2	0.8578	0.8539	0.8598	0.8474	0.8596	0.8657
MAPE	30	30.38	29.86	19.53	18.11	17.62
Déviante	436.53	398.4	434.59	878.36	104.98	747.19

TABLE 4.19 – Valeurs d'intérêt obtenues par sexe pour chaque méthode

Rappelons que le χ^2 , *MAPE* et la déviante doivent être minimisés tandis que le R^2 doit être maximisé et au plus proche de 1.

Nous constatons d'après ces résultats, que les valeurs des R^2 sont à peu près égales pour chacune des méthodes de positionnement et pour chacune des populations. Même s'il s'avère que la troisième méthode donne des résultats légèrement meilleurs que les deux précédentes, la différence n'est pas assez importante pour valider définitivement cette méthode. Cependant, le χ^2 et la *MAPE* donnent également l'avantage à la troisième méthode dans le cas des hommes. Pour ce qui concerne les femmes, la *MAPE* donne aussi un avantage à la méthode 3, mais le χ^2 favorise légèrement la deuxième. En revanche, le critère de la déviante montre de grands écarts entre les méthodes 1 et 3 avec la méthode 2, en donnant l'avantage à cette dernière. Ces résultats semblent cohérents avec ceux retournés par le test du rapport de vraisemblance qui utilise la déviante dans son application et qui préconise de même l'utilisation de la méthode 2.

À première vue, les test précédents mesurant le critère de proximité, ne permettent pas de distinguer une méthode à retenir parmi les trois proposées. Regardons alors ce qu'il en est des tests de régularité d'ajustement.

4.2. CRÉATION D'UNE TABLE DE MORTALITÉ À L'AIDE DU PACKAGE ELT

- Test des runs :

TEST DES RUNS	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
Nombre de runs	83	87	83	58	70	64
Signes (-)	86	92	88	90	94	83
Signes (+)	94	88	92	90	86	97
$ \chi^2 $	1.1717	0.5916	1.1899	4.9331	3.1189	3.979
Seuil	1.96	1.96	1.96	1.96	1.96	1.96
p-value	0.2413	0.5541	0.2341	0	0.0018	0.0001
Hypothèse retenue	H_0	H_0	H_0	H_1	H_1	H_1

TABLE 4.20 – Résultats obtenus pour le test des runs par sexe pour chaque méthode

L'ensemble des *p-values* est acceptable pour les femmes, bien que celles des méthodes 1 et 3 restent assez peu élevées; tandis que pour les hommes, l'ensemble des *p-values* est rejetable, la plus élevée étant de 0.18% pour la deuxième méthode de positionnement.

- Test des signes :

TEST DES SIGNES	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
Signes (+)	94	88	92	90	86	97
Signes (-)	86	92	88	90	94	83
χ^2	0.5217	0.2236	0.2236	-0.0745	0.5217	0.969
Seuil	1.96	1.96	1.96	1.96	1.96	1.96
p-value	0.6018	0.8231	0.8231	0.9506	0.6018	0.3326
Hypothèse retenue	H_0	H_0	H_0	H_0	H_0	H_0

TABLE 4.21 – Résultats obtenus pour le test des signes par sexe pour chaque méthode

Pour ce test-ci, les *p-values* sont toutes très bonnes et donc acceptables. Pour les femmes, il est cependant difficile de distinguer une méthode parmi les trois car leurs *p-values* sont très rapprochées; en ce qui concerne les hommes, la première méthode semble la plus adaptée selon ce test.

Une fois encore, il est difficile de savoir laquelle de ces trois méthodes de positionnement est la mieux adaptée à notre jeu de données. Afin d'éclaircir ce point, nous décidons de regrouper les résultats en mettant en place un système de poids selon l'ordre d'arrivée de la méthode en terme de qualité retournée par les tests (les méthodes sont ordonnées en fonction de leurs résultats par test, un poids de 3 est donné à la meilleure méthode, un poids de 2 à celle qui arrive en deuxième position, et enfin un poids de 1 à la moins bonne des méthodes), la moyenne de ces poids (recherchée la plus grande possible) permettra de distinguer quelle méthode apparaît être mieux adaptée à nos populations.

Nous constatons, de par cette pondération (TABLEAU 4.22), que l'utilisation de la méthode 3 semble plus favorable que les deux précédentes à la population des hommes. En revanche, pour les femmes, il est toujours difficile de différencier la qualité des méthodes 2 et 3. Néanmoins, si nous regardons les résultats retournés par les tests de proximité et valeurs d'intérêt, qui caractérisent la concordance des taux ajustés aux taux observés, la troisième méthode semble prendre un léger avantage sur la deuxième. Ainsi, en accordant plus d'importance à la fidélité aux données qu'à la régularité des taux observés, la méthode 3 semble celle à retenir pour cette étude.

4.2. CRÉATION D'UNE TABLE DE MORTALITÉ À L'AIDE DU PACKAGE ELT

SYNTHÈSE DES TESTS	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
Vraisemblance	1	3	2	1	3	2
SMR	3	1	2	3	1	2
Wilcoxon	3	1	2	1	2	3
χ^2	2	3	1	1	2	3
R^2	2	1	3	1	2	3
MAPE	2	1	3	1	2	3
Déviante	1	3	2	1	3	2
Runs	2	3	1	1	3	2
Signes	1	2	2	3	2	1
MOYENNE	1.89	2	2	1.44	2.22	2.33

TABLE 4.22 – Poids attribués par les tests à chaque méthode de positionnement selon leur qualité d'ajustement

Toutefois, cette méthode ne présentant pas de distinction majeure par rapport aux deux autres, nous comparons dans la section qui suit, les valeurs d'intérêt obtenues pour chacune des méthodes avec celles calculées sur les taux ajustés par abattement de la table de référence obtenus en SECTION 4.1.3 avant de choisir définitivement quelle semble être la méthode la mieux adaptée à notre jeu de données.

4.2.2.3 Choix de la méthode retenue pour la construction de notre table de mortalité

Comme énoncé juste auparavant, comparons les méthodes par positionnement et la méthode par abattement de la table de référence à l'aide des différentes valeurs d'intérêt calculées sur la plage d'âges 65-100 ans pour chacune de nos populations d'OD, et sur la plage 60-90 ans pour nos populations d'AD.

— Population des femmes OD :

	Abattement	Méthode 1	Méthode 2	Méthode 3
χ^2	2.1486	231.58	230.75	231.69
R^2	0.9992	0.8578	0.8539	0.8598
MAPE	1.7795	30	30.38	29.86

TABLE 4.23 – Comparaison des valeurs d'intérêt - Femmes OD

— Population des hommes OD :

	Abattement	Méthode 1	Méthode 2	Méthode 3
χ^2	4.9856	472.66	409.88	402.55
R^2	0.9963	0.8474	0.8596	0.8657
MAPE	1.9475	19.53	18.11	17.62

TABLE 4.24 – Comparaison des valeurs d'intérêt - Hommes OD

— Population des femmes AD :

	Abattement	Méthode 1	Méthode 2	Méthode 3
χ^2	12.101	238.28	240.52	215.82
R^2	0.9983	0.9458	0.9453	0.9458
MAPE	8.0224	26.58	26.34	27.93

TABLE 4.25 – Comparaison des valeurs d'intérêt - Femmes AD

4.2. CRÉATION D'UNE TABLE DE MORTALITÉ À L'AIDE DU PACKAGE ELT

— Population des hommes AD :

	Abattement	Méthode 1	Méthode 2	Méthode 3
χ^2	22.1186	233.5	268.61	173.99
R^2	0.9376	0.5538	0.5526	0.5679
MAPE	19.8502	54.5	53.21	45.73

TABLE 4.26 – Comparaison des valeurs d'intérêt - Hommes AD

Pour l'ensemble de nos populations, nous remarquons que la méthode d'ajustement par abattement de la table de référence donne de bien meilleurs résultats que les méthodes de positionnement proposées par le *package ELT*, quelle que soit la valeur d'intérêt considérée. Nous retenons alors comme méthode pour la construction de notre table de mortalité pour chacune de nos populations étudiées, la méthode d'ajustement par abattement de la table de référence.

Afin de valider la proximité de notre table à la réalité de la mortalité de notre population, nous comparons les données démographiques observées chaque année de la période d'observation avec celles attendues par la nouvelle loi de mortalité. Nous présentons ici les résultats obtenus sur l'ensemble de la période d'observation ainsi que pour l'année 2015. Les résultats des années 2013, 2014, 2016 et 2017 sont présentés en ANNEXE K.

Pour chacune de nos populations, nous présentons les résultats démographiques obtenus sur l'ensemble de la période d'observation (2013-2017) dans les tableaux suivants :

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	377 896	54.54	82.65	3 272	3 341	+2%
Hommes	1 088 863	55.58	79.79	15 391	14 729	-4%
Total	1 466 759	55.31	80.30	18 663	18 070	-3%

TABLE 4.27 – Résultats démographiques de la population des ouvriers droit sur la période 2013-2017

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	208 066	79.70	89.30	9 984	11 028	+10%
Hommes	8 633	76.24	84.67	484	507	+5%
Total	216 699	79.56	89.09	10 468	11 535	+10%

TABLE 4.28 – Résultats démographiques de la population des ayants droit sur la période 2013-2017

Sur l'ensemble de la période d'observation, nous constatons des écarts très faibles entre les valeurs observées et les valeurs attendues par la nouvelle table pour les OD. Des écarts légèrement plus importants sont constatés chez les AD, ceci vient du fait que la population des AD est peu nombreuse et qu'il est difficile d'ajuster une table de mortalité à une population de petite taille.

En ce qui concerne l'année d'observation 2015, nous obtenons les résultats démographiques suivants pour chacune de nos populations :

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	75 916	53.94	82.65	653	667	+2%
Hommes	218 131	55.09	79.84	3 069	3 053	-1%
Total	294 047	54.79	80.33	3 722	3 720	0%

TABLE 4.29 – Résultats démographiques de la population des ayants droit sur l'année 2015

4.2. CRÉATION D'UNE TABLE DE MORTALITÉ À L'AIDE DU PACKAGE ELT

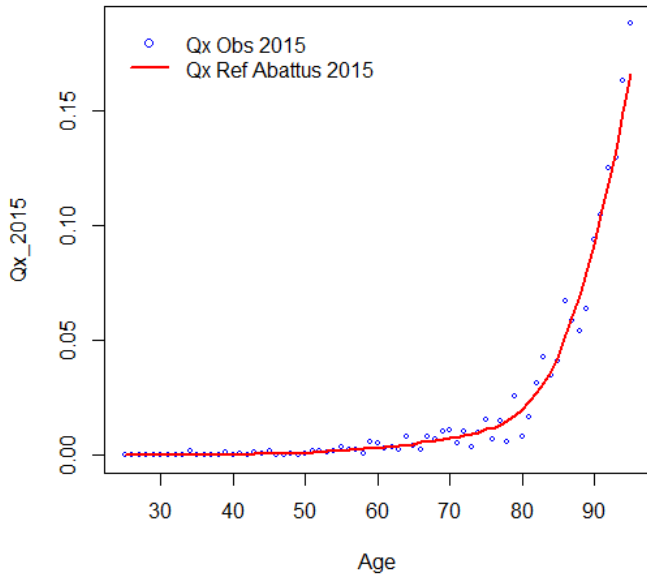


FIGURE 4.52 – Femmes OD - 2015

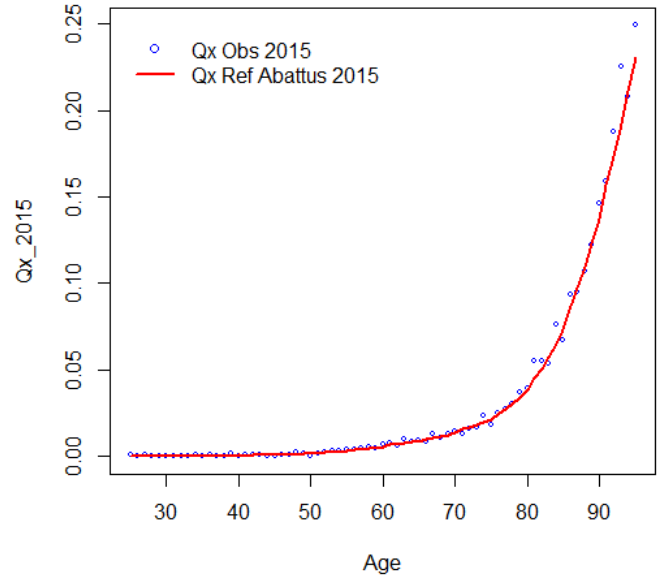


FIGURE 4.53 – Hommes OD - 2015

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	41 739	79.74	89.31	2 008	2 276	+13%
Hommes	1 733	76.25	86.16	98	105	+7%
Total	43 472	79.60	89.16	2 106	2 381	+13%

TABLE 4.30 – Résultats démographiques de la population des ayants droit sur l'année 2015

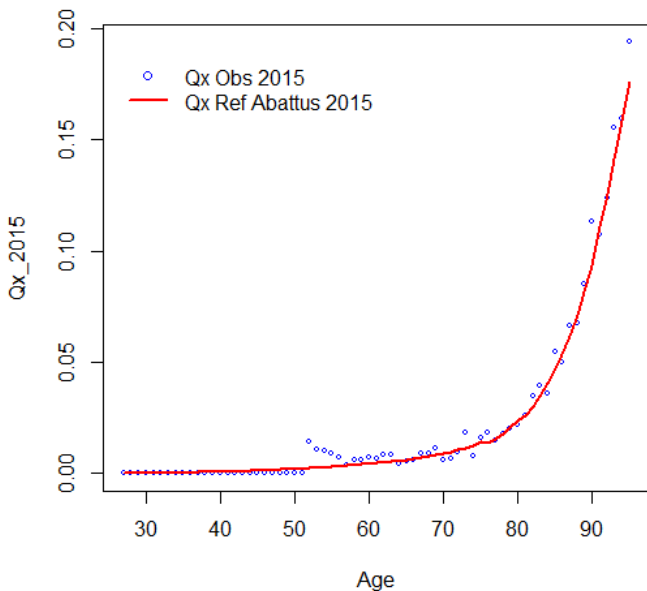


FIGURE 4.54 – Femmes AD - 2015

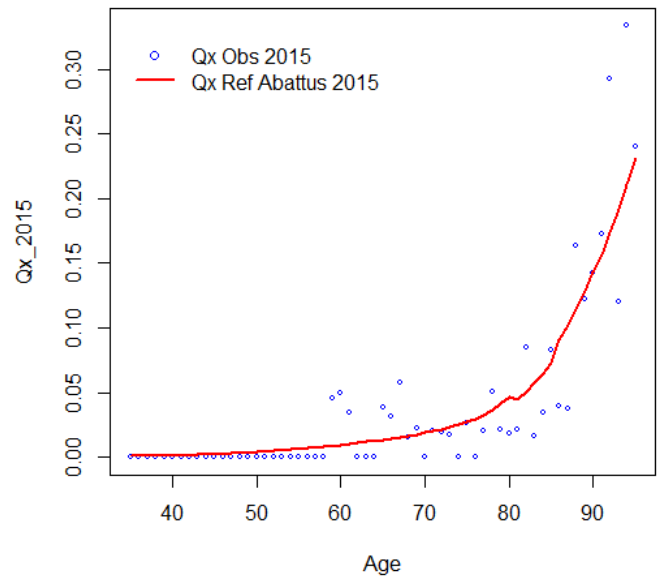


FIGURE 4.55 – Hommes AD - 2015

Un écart moyen de 0% est constaté pour les OD, la table est donc parfaitement adaptée à cette population sur l'année 2015. Du côté des AD, l'écart moyen constaté est de 13%, un écart un peu plus élevé pour les mêmes raisons que celles énoncées lors du cas précédent (taille de population faible). Les tables semblent donc plutôt bien adaptées à la mortalité de nos populations.

4.2. CRÉATION D'UNE TABLE DE MORTALITÉ À L'AIDE DU PACKAGE ELT

De plus, nous remarquons graphiquement que les taux des nouvelles tables suivent relativement bien les taux bruts et ce, particulièrement chez les OD ce qui confirme bien les écarts constatés.

Par ailleurs, nous remarquons que l'âge moyen au décès est plus élevé chez les femmes que chez les hommes, quelle que soit la population considérée. Ceci reflète la supériorité de l'espérance de vie chez les femmes, et nous pouvons le constater sur les deux graphiques suivants :

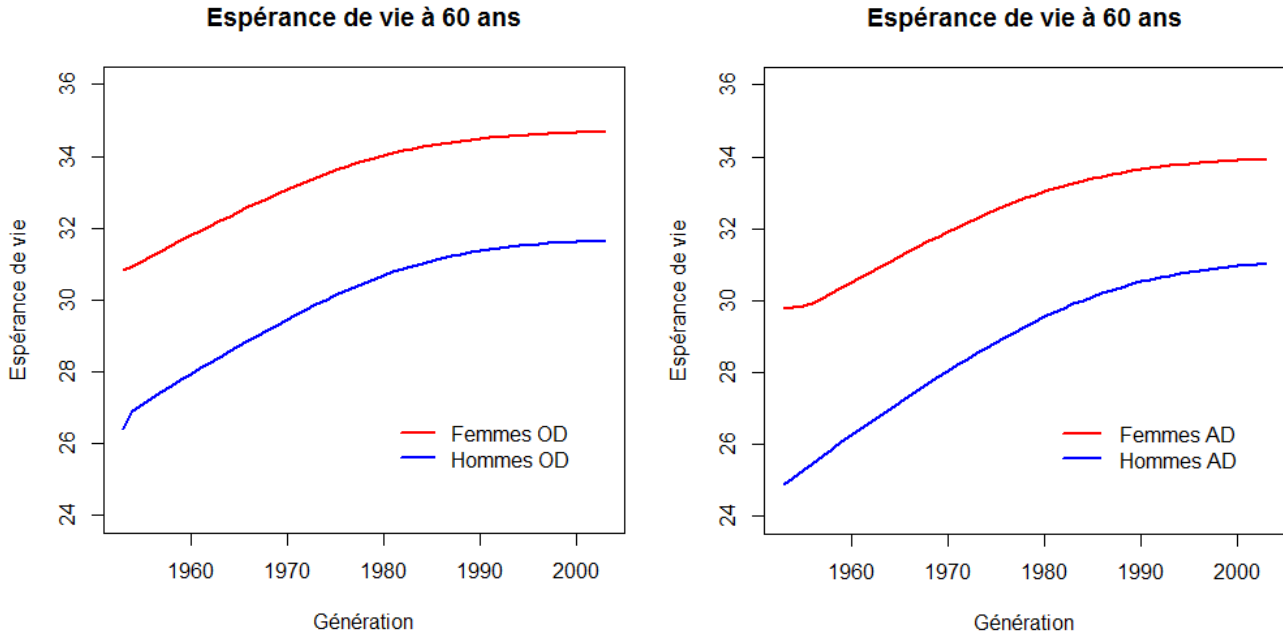


FIGURE 4.56 – Comparaison des espérances de vie à 60 ans des populations en fonction de leur sexe

Enfin, la comparaison des espérances de vie à 60 ans, 70 ans et 80 ans entre les populations des OD et des AD montrent qu'en moyenne, les affiliés qui possèdent la qualité d'ouvrant droit vivent plus longtemps que celles considérées comme ayant droit. En effet, c'est ce que nous constatons sur les graphiques suivants¹³ :

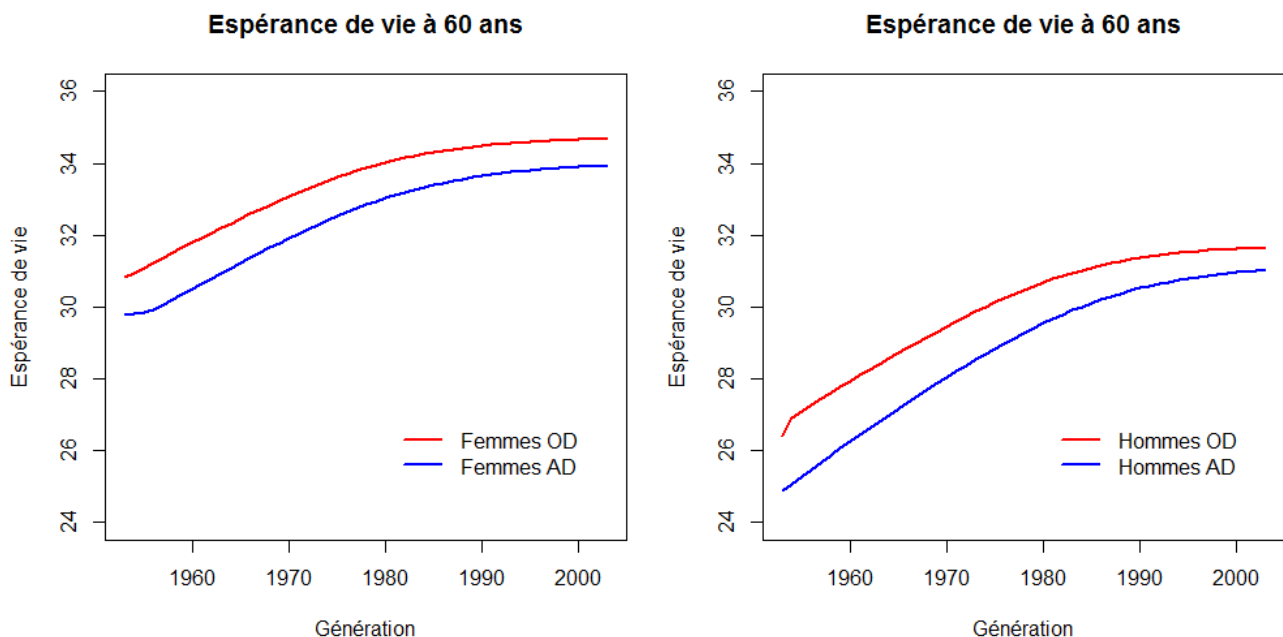


FIGURE 4.57 – Comparaison des espérances de vie à 60 ans des populations en fonction de leur qualité

13. Les graphiques des espérances de vie à 70 et 80 ans sont présentés en ANNEXE L.1

4.3. FERMETURE DE LA TABLE

Maintenant que nous avons construit notre table de mortalité, il convient de la comparer avec celle utilisée actuellement par le régime de retraite afin d'évaluer si elle est plus adaptée ou non à notre population. Pour cela, nous appliquons à nouveau les trois valeurs d'intérêt utilisées jusqu'à présent, sur les taux obtenus avec la table actuelle, de sorte à les comparer avec celles calculées précédemment dans le cadre de la nouvelle table.

	Table actuelle	Nouvelle table
χ^2	40.060	5.143
R^2	0.9961	0.9995
MAPE	55.939	34.525

TABLE 4.31 – Évaluation des tables - Femmes OD

	Table actuelle	Nouvelle table
χ^2	308.610	9.059
R^2	0.9903	0.9998
MAPE	22.066	3.112

TABLE 4.32 – Évaluation des tables - Hommes OD

	Table actuelle	Nouvelle table
χ^2	605.297	321.115
R^2	0.930	0.929
MAPE	47.222	40.469

TABLE 4.33 – Évaluation des tables - Femmes AD

	Table actuelle	Nouvelle table
χ^2	122.150	51.730
R^2	0.856	0.896
MAPE	45.014	34.426

TABLE 4.34 – Évaluation des tables - Hommes AD

L'ensemble des valeurs d'intérêt conduit à la conclusion que la nouvelle table ajustée est la plus adaptée à chacune de nos populations du régime. C'est donc cette dernière que nous retenons.

4.3 Fermeture de la table

La méthode que nous utilisons ici dans le cadre de la fermeture de la table de mortalité est la méthode de Denuit et Goderniaux (2005). Cette méthode impose une contrainte sur la fermeture des tables de mortalité aux grands âges.

Elle consiste en l'ajustement du modèle log-quadratique suivant par la méthode des moindres carrés :

$$\ln(q_x) = a + bx + cx^2 + \varepsilon_x$$

avec ε_x les résidus de loi normale centrée.

Les deux contraintes suivantes sont imposées :

$$\begin{cases} q_{120} = 1 \\ q'_{120} = 0 \end{cases}$$

avec q'_x la dérivée de q_x à l'âge x .

La première contrainte stipule une probabilité de décès de 100% à 120 ans ainsi que la présence d'une tangente horizontale à cet âge.

Par ailleurs, la seconde contrainte traduit une concavité aux âges élevés qui empêche les taux de mortalité de décroître aux grands âges.

Ces deux contraintes nous conduisent alors aux deux équations suivantes :

$$\begin{cases} a + 120b + 120^2c = 0 \\ b + 2c = 0 \end{cases}$$

soit :

$$\ln(q_x) = c(120^2 - 240x + x^2) + \varepsilon_x$$

4.3. FERMETURE DE LA TABLE

Il convient maintenant de déterminer l'âge de raccord à partir duquel les taux bruts sont remplacés par les taux retournés par le modèle précédent. La façon "classique" de déterminer ce dernier est la maximisation du coefficient de détermination R^2 en effectuant le test à plusieurs âges.

Ici, les projections s'effectuant pour plusieurs années, il semble difficile de considérer des âges de raccord différents pour chacune d'entre elles au risque d'obtenir des tables non lisses. C'est pourquoi nous déterminons un âge de raccord commun à chacune des années, à partir duquel les taux de décès ajustés par abattement de la table de référence seront remplacés par les taux issus de la méthode de fermeture des tables. L'âge en question est alors 95 ans.

Les tables finales ajustées par abattement de la table de référence obtenues pour chacune de nos quatre populations sont représentées sur les graphiques suivants (sur la plage d'âges 60-120 ans¹⁴) :

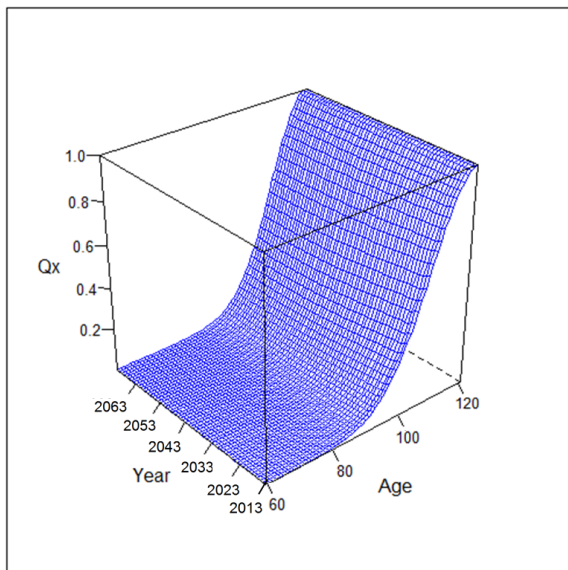


FIGURE 4.58 – Table de mortalité ajustée des femmes OD

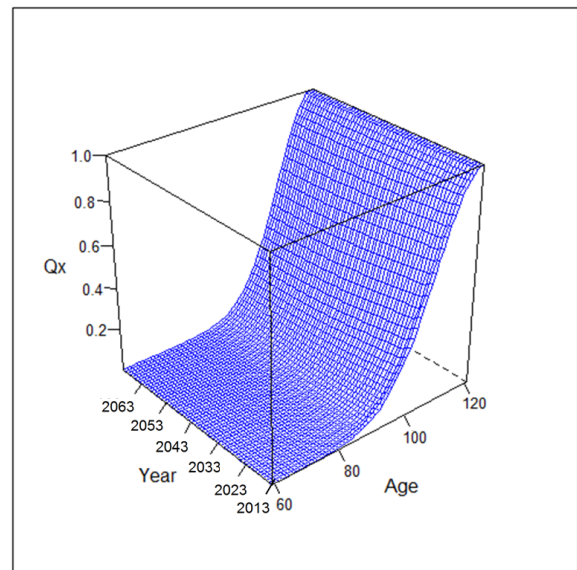


FIGURE 4.59 – Table de mortalité ajustée des hommes OD

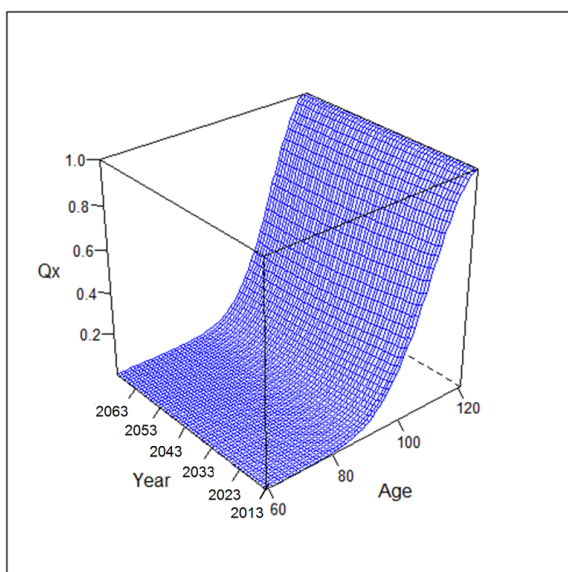


FIGURE 4.60 – Table de mortalité ajustée des femmes AD

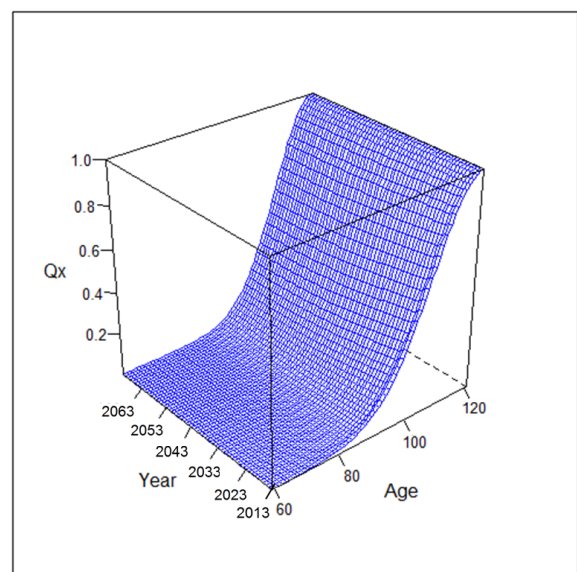


FIGURE 4.61 – Table de mortalité ajustée des hommes AD

14. Les tables complètes sont disponibles en ANNEXE M.

Chapitre 5

Évaluation des impacts financiers

Il existe différentes manières d'évaluer les impacts financiers faisant suite à un changement d'utilisation de table de mortalité dans les calculs établis par un régime de retraite. Nous présentons ici les résultats qui concernent les flux de pensions probables en droits acquis projetés du régime de retraite en groupe fermé.

L'objectif de ce mémoire étant principalement fondé sur la construction de la table de mortalité et non sur l'élaboration d'un outil de calcul de pensions de retraite, nous ne présenterons pas le détail des calculs effectués ici, mais nous nous contenterons de présenter les résultats obtenus afin de les comparer en fonction des tables de mortalité utilisées.

Deux cas sont étudiés ici :

- Un premier cas où la loi de mortalité utilisée pour les calculs d'engagements est identique pour les ouvriers et les ayants droit (loi des ouvriers également appliquée aux ayants droit).
- Un second cas où la loi de mortalité utilisée pour les calculs d'engagements est différente pour les ouvriers et les ayants droit.

Ces deux scénarios seront présentés aux instances dirigeantes du régime qui choisiront si elles préfèrent conserver par mesure de prudence, une seule et même table pour les OD et les AD.

Les résultats sont présentés pour trois types de population qui sont la population des actifs seuls, la population des inactifs seuls, et la population des actifs et inactifs rassemblés.

Dans un premier temps, comparons les flux de pension pour la population des actifs seuls :

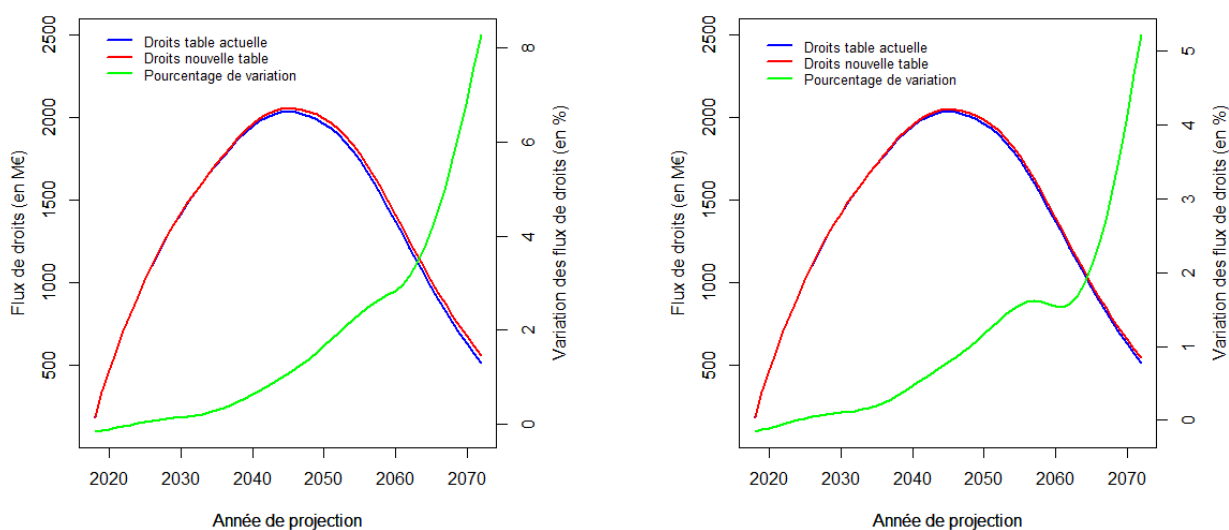


FIGURE 5.1 – Comparaison des flux de droits pour les actifs seuls avec une table unique (à gauche) et deux tables différentes (à droite)

Nous constatons une légère différence de projection de flux de pensions probables en droits acquis selon l'utilisation de la table pour la population des actifs sur le long terme. En effet entre 2% et 8% de différence peuvent être remarqués sur les projections de flux pour les années au-delà de 2050 dans le cas du scénario avec loi unique, et entre 1% et 5% pour le scénario avec deux lois différentes. L'utilisation de la nouvelle loi de mortalité sur la population des actifs seuls a pour conséquence une légère augmentation des flux de pensions estimés que le régime aura à verser dans les années à venir.

Ensuite, comparons les flux de pension pour la population des inactifs seuls :

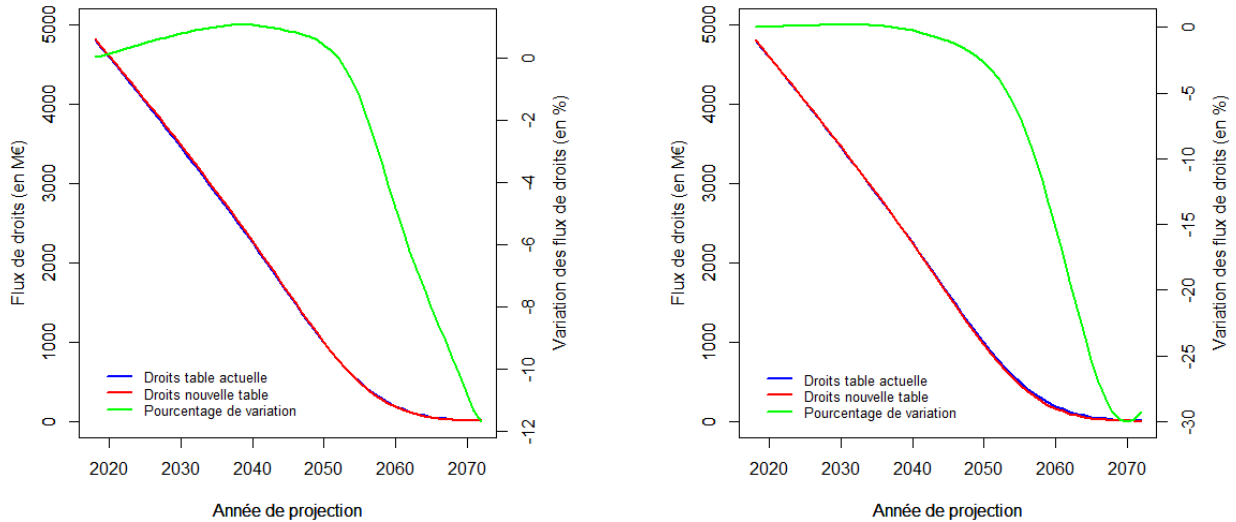


FIGURE 5.2 – Comparaison des flux de droits pour les inactifs seuls avec une table unique (à gauche) et deux tables différentes (à droite)

Cette fois encore, une différence de projection de flux de pensions probables en droits acquis est constatée selon l'utilisation de la table pour la population des inactifs (sur le long terme). Les projections sont quasiment équivalentes sur les trente premières années de projections mais des écarts pouvant aller jusqu'à -12% (scénario loi unique) et -30% (scénario deux lois) sont observables en queue de projection. Dans ce cas-ci, l'utilisation de la nouvelle loi de mortalité sur la population des inactifs seuls a pour conséquence une baisse des flux de pensions estimés que le régime aura à verser dans les années à venir.

Enfin, comparons les flux de pension pour la population des actifs et inactifs rassemblés (c'est-à-dire l'ensemble de la population).

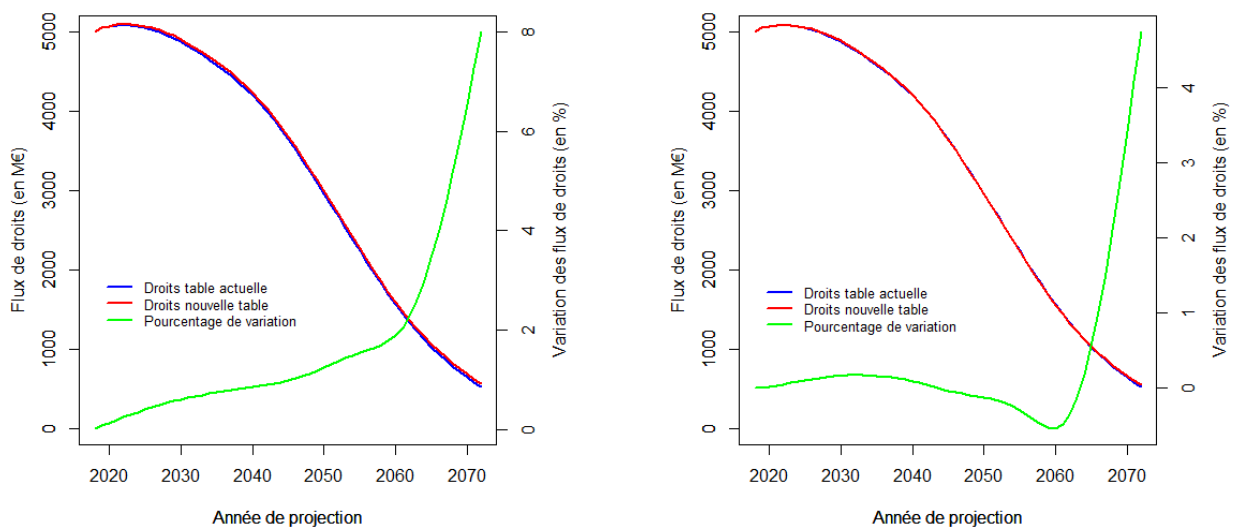


FIGURE 5.3 – Comparaison des flux de droits pour l'ensemble des actifs et des inactifs avec une table unique (à gauche) et deux tables différentes (à droite)

En ce qui concerne la population totale des affiliés du régime, les projections de flux de pension sont très proches quelle que soit la table utilisée. De légers écarts sont constatés à partir de 40 années de projections (lorsque la masse des pensions servies a fortement diminuée), mais ces variations ne dépassent pas 8% et 5% respectivement pour le scénario loi unique et le scénario deux lois. Une légère augmentation des flux de pensions estimés à verser par le régime est alors conséquence de l'utilisation de la nouvelle loi de mortalité établie.

Nous venons d'analyser les comparaisons de flux de pension calculés avec les différentes lois de mortalité entre les différentes populations qui composent notre régime de retraite. Maintenant, analysons les différences entre les deux scénarios proposés un peu plus haut : le cas où la loi de mortalité utilisée est la même pour les deux populations (ouvrants droit et ayants droit), et le cas où la loi utilisée est différente pour chaque population.

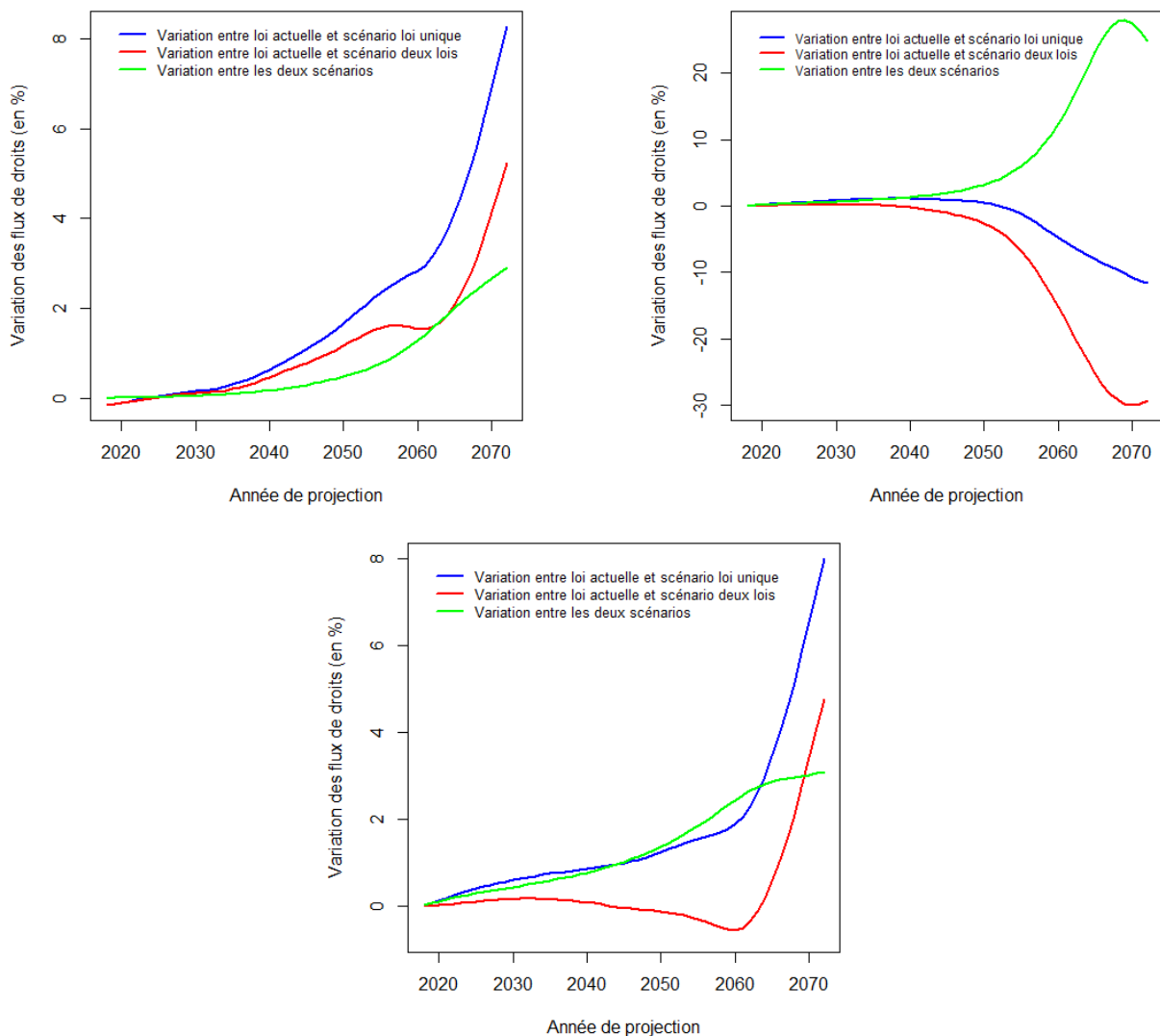


FIGURE 5.4 – Comparaison des flux de droits selon le scénario pour les actifs seuls (à gauche), les inactifs seuls (à droite), et l'ensemble de la population (en bas)

Sur les trois graphiques de la figure précédente, les courbes **bleue**, **rouge**, et **verte**, correspondent respectivement aux pourcentages de variation entre les projections des flux de droits obtenus avec la table actuelle et la nouvelle table pour le scénario loi unique, entre la table actuelle et la nouvelle table pour le scénario deux lois, et entre les deux scénarios proposés. Il est alors possible de constater que dans chacun des cas, les flux de droits projetés calculés avec la loi unique sont légèrement supérieurs à ceux calculés avec deux lois différentes. C'est pourquoi ce scénario est généralement considéré comme plus prudent que le scénario utilisant deux lois de mortalité différentes selon la population considérée.

Chapitre 6

Conclusion

En retraite, éviter les erreurs de financement constitue un objectif primordial pour assurer la pérennité d'un régime, et les tables de mortalité deviennent l'outil principal pour y parvenir. Les tables utilisées doivent donc régulièrement être mises à jour ou ajustées du fait de l'évolution progressive de l'espérance de vie en France.

L'objectif principal de ce mémoire était d'établir une loi de mortalité par ajustement d'une table de référence qui fournit une représentation de la mortalité de la population d'un régime de retraite spécifique et ce, en appliquant différentes méthodes d'ajustement. Pour une analyse plus juste et plus réelle, la population du régime a été segmentée en quatre "sous-populations" selon le sexe et le type de prestation des individus. Chacune de ces populations a alors fait l'objet d'une construction de table de mortalité spécifique.

L'analyse des statistiques descriptives a révélé des disparités dans la répartition des individus au sein de chaque population. L'application d'un lissage sur les taux bruts de mortalité s'est alors montré nécessaire du fait de certaines irrégularités. Un lissage bien adapté permet à la fois d'éviter d'accorder trop d'importance à des données présentant un caractère peu significatif tout en tenant rigueur de la fidélité aux données brutes. La méthode de Whittaker-Henderson tient compte de ces deux paramètres en nous accordant la liberté de choisir l'importance attribuée à chacun des deux critères de fidélité et de régularité, en prenant également en considération les poids des données brutes. Par la suite, des tests statistiques ont été mis en application afin de valider ce lissage effectué. Les données lissées retenues ont alors pu constituer l'élément d'appui principal de notre ajustement de table. Toutefois, bien que le lissage ait été validé et que les résultats se soient montrés satisfaisants, nous avons retenu comme paramètres de notre lissage des paramètres par défaut. Ainsi, la variation de ces paramètres aurait peut-être conduit à une modification des tendances et donc à une loi finale encore plus adaptée que celle obtenue.

L'application de décalages d'âges ou de coefficients d'abattement sur les taux de mortalité renseignés par la table de référence sont deux méthodes qui ont permis l'obtention d'ajustements proches de la réalité. La procédure s'appliquant tranche d'âges par tranche d'âges, la fidélité aux données brutes est respectée à tout âge. D'autre part, les méthodes de positionnement proposées par le *package* informatique *ELT* dédié à la construction de tables de mortalité prospectives (et plus particulièrement la troisième méthode consistant à approcher le modèle par un modèle linéaire généralisé de Poisson) peuvent être acceptables. Cependant, ces applications ont montré des erreurs d'ajustement avec les données lissées plus importantes que les deux méthodes précédentes. Ceci peut s'expliquer notamment par le fait que l'ajustement est réalisé directement à partir des données brutes et ce, sur l'ensemble des données. La méthode par application de coefficients d'abattement sur la table de référence est alors celle qui se distingue le mieux car s'avère plus précise que celle par décalages d'âges.

Par ailleurs, la comparaison des lois actuelles et nouvelles a révélé que la nouvelle loi est plus fidèle aux observations que la loi actuelle. De plus, elle conserve la supériorité de l'espérance de vie des femmes sur celle des hommes, ainsi que celle des ouvriers droit sur celle des ayants droit. L'objectif est donc atteint, la nouvelle loi que nous venons d'établir est plus adaptée au risque viager que les tables utilisées par le régime jusqu'à présent.

L'étape finale de cette étude consistait à évaluer les impacts financiers dus à l'utilisation de la nouvelle table. Il a été conclu que l'application de la nouvelle loi de mortalité n'a pas d'impact significatif sur la projection des droits totaux (la nouvelle table améliore l'espérance de vie aux grands âges, mais la réduit aux âges moins élevés par rapport à la table précédente). La mise à jour régulière des tables de mortalité permet donc de limiter l'ampleur de l'impact financier causé par le changement d'utilisation de table. D'autre part, la très légère augmentation des montants de pension qui a pu être observée était valable pour les projections de plus de 30 années. Cela permet de conclure sur le fait que la taille de notre population est suffisante pour une étude de la mortalité, mais pas assez pour créer une table de mortalité générationnelle qui projette les espérances de vie sur le long terme. Néanmoins, toujours dans l'optique d'évaluer les impacts financiers qui pourraient avoir lieu à la suite d'un changement d'utilisation de table de mortalité, il pourrait être intéressant de réitérer l'étude sur une segmentation différente de la population. Par exemple, les espérances de vie pouvant se montrer très diverses selon la profession des individus, une segmentation par catégories socio-professionnelles pourrait donc faire l'objet d'une étude approfondie.

Table des figures

1	Démarche mise en application	3
2	Statistiques de populations	4
3	Taux de mortalité observés pour chaque population	5
4	Taux de mortalité observés, lissés, et abattus pour chaque population	7
5	Flux de droits projetés et pourcentages de variation selon les tables considérées	8
6	Approach used	10
7	Statistics of the populations	11
8	Crude mortality rates of each population	12
9	Crude, smoothed, and adjusted mortality rates for each population	14
10	Projected pension amounts and percentages of variation between laws	15
2.1	Espérance de vie à la naissance en France entre 1740 et 2010	22
2.2	Évolution de la mortalité infantile en France entre 1740 et 2010	24
2.3	Évolution annuelle des décès en France depuis 1920	24
2.4	Évolution de la population et du nombre de décès en France entre 1946 et 2000	25
2.5	Pyramide des âges en 1914	26
2.6	Pyramide des âges en 1954	26
2.7	Pyramide des âges en 2014	26
2.8	Âges moyens de départ à la retraite - Ensemble de la population	33
2.9	Âges moyens de départ à la retraite - Femmes	33
2.10	Âges moyens de départ à la retraite - Hommes	33
2.11	Durée moyenne passée en retraite par génération - Ensemble de la population	34
2.12	Durée moyenne passée en retraite par génération - Par sexe	34
2.13	Rapport entre la durée de retraite et l'espérance de vie par génération	36
2.14	Rapport entre la durée de retraite et l'espérance de vie par génération et par sexe	36
3.1	Répartition de la population selon le type de prestation	41
3.2	Répartition des populations des OD et des AD selon leur sexe	42
3.3	Répartition de la population des ouvriers droit selon leur âge	42
3.4	Répartition de la population des ayants droit selon leur âge	43
3.5	Effectifs annuels moyens de population par tranche d'âges sur la période d'observation	43
3.6	Décès observés par année de la période d'observation selon la qualité et le sexe des individus	44
3.7	Décès observés par tranche d'âges selon la qualité et le sexe des individus sur la période d'observation	45
3.8	Qx Observés - Femmes OD	46
3.9	Qx Observés - Hommes OD	46
3.10	Qx Observés - Femmes AD	46
3.11	Qx Observés - Hommes AD	46
3.12	Qx moyens observés par population sur la période d'observation	47
4.1	Lissage des taux bruts par la méthode des moindres carrés - Femmes OD	51
4.2	Lissage des taux bruts par la méthode des moindres carrés - Hommes OD Tranche 1	52
4.3	Lissage des taux bruts par la méthode des moindres carrés - Hommes OD Tranche 2	52
4.4	Lissage des taux bruts par la méthode des moindres carrés - Femmes AD	53

4.5	Lissage des taux bruts par la méthode des moindres carrés - Hommes AD	54
4.6	Lissage des taux bruts par la méthode de Makeham - Femmes OD	56
4.7	Lissage des taux bruts par la méthode de Makeham - Hommes OD	56
4.8	Lissage des taux bruts par la méthode de Makeham - Femmes AD	57
4.9	Lissage des taux bruts par la méthode de Makeham - Hommes AD	57
4.10	Lissage des taux bruts par la méthode de Loess - Femmes OD	59
4.11	Lissage des taux bruts par la méthode de Loess - Hommes OD	59
4.12	Lissage des taux bruts par la méthode de Loess - Femmes OD Tranche 1	59
4.13	Lissage des taux bruts par la méthode de Loess - Hommes OD Tranche 1	59
4.14	Lissage des taux bruts par la méthode de Loess - Femmes OD Tranche 2	60
4.15	Lissage des taux bruts par la méthode de Loess - Hommes OD Tranche 2	60
4.16	Lissage des taux bruts par la méthode de Loess - Femmes AD	60
4.17	Lissage des taux bruts par la méthode de Loess - Hommes AD	60
4.18	Lissage des taux bruts par la méthode de Loess - Femmes AD Tranche 1	61
4.19	Lissage des taux bruts par la méthode de Loess - Hommes AD Tranche 1	61
4.20	Lissage des taux bruts par la méthode de Loess - Femmes AD Tranche 2	61
4.21	Lissage des taux bruts par la méthode de Loess - Hommes AD Tranche 2	61
4.22	Lissage des taux bruts par la méthode de Whittaker-Henderson - Femmes OD	64
4.23	Lissage des taux bruts par la méthode de Whittaker-Henderson - Hommes OD	64
4.24	Lissage des taux bruts par la méthode de Whittaker-Henderson - Femmes AD	65
4.25	Lissage des taux bruts par la méthode de Whittaker-Henderson - Hommes AD	65
4.26	Comparaison des lissages réalisés - Femmes OD Complet	66
4.27	Comparaison des lissages réalisés - Femmes OD Tranche 4	67
4.28	Comparaison des lissages réalisés - Femmes OD Tranche 1	67
4.29	Comparaison des lissages réalisés - Femmes OD Tranche 2	67
4.30	Comparaison des lissages réalisés - Femmes OD Tranche 3	68
4.31	Comparaison des lissages réalisés - Hommes OD Complet	68
4.32	Comparaison des lissages réalisés - Hommes OD Tranche 1	69
4.33	Comparaison des lissages réalisés - Hommes OD Tranche 2	69
4.34	Comparaison des lissages réalisés - Hommes OD Tranche 3	69
4.35	Comparaison des lissages réalisés - Femmes AD Complet	70
4.36	Comparaison des lissages réalisés - Femmes AD Tranche 1	70
4.37	Comparaison des lissages réalisés - Femmes AD Tranche 2	70
4.38	Comparaison des lissages réalisés - Femmes AD Tranche 3	71
4.39	Comparaison des lissages réalisés - Hommes AD Complet	71
4.40	Comparaison des lissages réalisés - Hommes AD Tranche 2	72
4.41	Comparaison des lissages réalisés - Hommes AD Tranche 3	72
4.42	Diagramme <i>Quantile-Quantile</i> de la population des femmes OD	75
4.43	Diagramme <i>Quantile-Quantile</i> de la population des femmes OD pour la tranche 60-84 ans	76
4.44	Décalages optimaux - Femmes OD	77
4.45	Décalages optimaux - Hommes OD	77
4.46	Décalages optimaux - Femmes AD	78
4.47	Décalages optimaux - Hommes AD	78
4.48	Abattements optimaux - Femmes OD	79
4.49	Abattements optimaux - Hommes OD	80
4.50	Abattements optimaux - Femmes AD	80
4.51	Abattements optimaux - Hommes AD	80
4.52	Femmes OD - 2015	95
4.53	Hommes OD - 2015	95
4.54	Femmes AD - 2015	95
4.55	Hommes AD - 2015	95
4.56	Comparaison des espérances de vie à 60 ans des populations en fonction de leur sexe	96
4.57	Comparaison des espérances de vie à 60 ans des populations en fonction de leur qualité	96

TABLE DES FIGURES

4.58	Table de mortalité ajustée des femmes OD	98
4.59	Table de mortalité ajustée des hommes OD	98
4.60	Table de mortalité ajustée des femmes AD	98
4.61	Table de mortalité ajustée des hommes AD	98
5.1	Comparaison des flux de droits pour les actifs seuls avec une table unique (à gauche) et deux tables différentes (à droite)	99
5.2	Comparaison des flux de droits pour les inactifs seuls avec une table unique (à gauche) et deux tables différentes (à droite)	100
5.3	Comparaison des flux de droits pour l'ensemble des actifs et des inactifs avec une table unique (à gauche) et deux tables différentes (à droite)	100
5.4	Comparaison des flux de droits selon le scénario pour les actifs seuls (à gauche), les inactifs seuls (à droite), et l'ensemble de la population (en bas)	101
A.1	Représentation simpliste d'un diagramme de Lexis	112
A.2	Représentation simpliste d'âges révolus sur un diagramme de Lexis	113
A.3	Représentation simpliste d'âges atteints sur un diagramme de Lexis	113
A.4	Représentation simpliste du nombre de décès par âge atteint sur un diagramme de Lexis	114
E.1	Comparaison des lissages réalisés - Hommes OD Tranche 4	119
E.2	Comparaison des lissages réalisés - Femmes AD Tranche 4	120
E.3	Comparaison des lissages réalisés - Hommes AD Tranche 4	120
G.1	Diagramme <i>Quantile-Quantile</i> de la population des hommes OD	124
G.2	Diagramme <i>Quantile-Quantile</i> de la population des hommes OD pour la tranche 60-87 ans	124
G.3	Diagramme <i>Quantile-Quantile</i> de la population des femmes AD	125
G.4	Diagramme <i>Quantile-Quantile</i> de la population des femmes AD pour la tranche 60-84 ans	125
G.5	Diagramme <i>Quantile-Quantile</i> de la population des hommes AD	125
G.6	Diagramme <i>Quantile-Quantile</i> de la population des hommes AD pour la tranche 60-84 ans	125
K.1	Femmes OD - 2013	132
K.2	Hommes OD - 2013	132
K.3	Femmes AD - 2013	133
K.4	Hommes AD - 2013	133
K.5	Femmes OD - 2014	133
K.6	Hommes OD - 2014	133
K.7	Femmes AD - 2014	134
K.8	Hommes AD - 2014	134
K.9	Femmes OD - 2016	135
K.10	Hommes OD - 2016	135
K.11	Femmes AD - 2016	135
K.12	Hommes AD - 2016	135
K.13	Femmes OD - 2017	136
K.14	Hommes OD - 2017	136
K.15	Femmes AD - 2017	136
K.16	Hommes AD - 2017	136
L.1	Comparaison des espérances de vie à 70 ans des populations en fonction de leur sexe	137
L.2	Comparaison des espérances de vie à 70 ans des populations en fonction de leur sexe	137
L.3	Comparaison des espérances de vie à 80 ans des populations en fonction de leur qualité	138
L.4	Comparaison des espérances de vie à 80 ans des populations en fonction de leur qualité	138
M.1	Table de mortalité ajustée des femmes OD	139
M.2	Table de mortalité ajustée des hommes OD	139

TABLE DES FIGURES

M.3 Table de mortalité ajustée des femmes AD	139
M.4 Table de mortalité ajustée des hommes AD	139

Liste des tableaux

2.1	Extrait de la table du moment TF 00-02	30
2.2	Extrait de la table du moment TH 00-02	30
2.3	Extrait de la table INSEE prospective 2013 - 2070 des Femmes	31
2.4	Extrait de la table INSEE prospective 2013 - 2070 des Hommes	31
2.5	Évolution de l'espérance de vie à 60 ans et de la durée de retraite entre les générations 1943 et 1989	35
4.1	Valeurs des coefficients de détermination pour la population des hommes OD	51
4.2	Paramètres du lissage par la méthode de Makeham pour la population des OD	55
4.3	Paramètres du lissage par la méthode de Makeham pour la population des AD	57
4.4	Coefficients de détermination R^2 obtenus pour chaque lissage et chaque population	73
4.5	Coefficients de détermination R^2 obtenus pour chaque lissage et chaque population pour la tranche 60-89 ans	73
4.6	Valeurs du χ^2 obtenues pour chaque lissage et chaque population	73
4.7	Valeurs de la <i>MAPE</i> obtenues pour chaque lissage et chaque population	74
4.8	Résultats retournés par les tests pour chaque population	75
4.9	Résultats retournés par les tests pour chaque population sur les plages d'âges sélectionnées	76
4.10	Valeurs d'intérêts obtenues sur les taux de référence décalés	78
4.11	Valeurs d'intérêts obtenues sur les taux de référence abattus	81
4.12	Format des données	81
4.13	SMR obtenus pour différentes plages d'âges	83
4.14	Estimations des paramètres α et β pour chacune des populations des OD	85
4.15	Estimations des paramètres β_0 , β_1 et β_2 pour chacune des populations des OD	85
4.16	Résultats obtenus pour le test du rapport de vraisemblance par sexe pour chaque méthode	90
4.17	Résultats obtenus pour le test du SMR par sexe pour chaque méthode	90
4.18	Résultats obtenus pour le test de Wilcoxon par sexe pour chaque méthode	91
4.19	Valeurs d'intérêt obtenues par sexe pour chaque méthode	91
4.20	Résultats obtenus pour le test des <i>runs</i> par sexe pour chaque méthode	92
4.21	Résultats obtenus pour le test des signes par sexe pour chaque méthode	92
4.22	Poids attribués par les tests à chaque méthode de positionnement selon leur qualité d'ajustement	93
4.23	Comparaison des valeurs d'intérêt - Femmes OD	93
4.24	Comparaison des valeurs d'intérêt - Hommes OD	93
4.25	Comparaison des valeurs d'intérêt - Femmes AD	93
4.26	Comparaison des valeurs d'intérêt - Hommes AD	94
4.27	Résultats démographiques de la population des ouvrants droit sur la période 2013-2017	94
4.28	Résultats démographiques de la population des ayants droit sur la période 2013-2017	94
4.29	Résultats démographiques de la population des ouvrants droit sur l'année 2015	94
4.30	Résultats démographiques de la population des ayants droit sur l'année 2015	95
4.31	Évaluation des tables - Femmes OD	97
4.32	Évaluation des tables - Hommes OD	97
4.33	Évaluation des tables - Femmes AD	97

4.34	Évaluation des tables - Hommes AD	97
B.1	Extrait de la table TGF-05	115
B.2	Extrait de la table TGH-05	115
D.1	Parts d'individus au sein de chaque population	118
D.2	Parts de chaque population au sein de la population totale des affiliés du régime	118
D.3	Parts d'OD par tranche d'âges selon le sexe des individus	118
D.4	Parts d'AD par tranche d'âges selon le sexe des individus	118
E.1	Valeurs du χ^2 obtenues pour chaque lissage et chaque population pour la tranche 60-89 ans	121
E.2	MAPE obtenues pour chaque lissage et chaque population pour la tranche 60-89 ans	121
H.1	Valeurs du SMR pour différentes plages d'âges sur la population des OD	126
H.2	Valeurs du SMR pour différentes plages d'âges sur la population des AD	126
H.3	Valeurs du SMR pour différentes plages d'âges sur la population des AD	126
H.4	Estimations des paramètres α et β pour chacune des populations des AD	127
H.5	Estimations des paramètres β_0 , β_1 et β_2 pour chacune des populations des AD	127
J.1	Résultats obtenus pour le test du rapport de vraisemblance par sexe pour chaque méthode - AD	129
J.2	Résultats obtenus pour le test du SMR par sexe pour chaque méthode - AD	129
J.3	Résultats obtenus pour le test de Wilcoxon par sexe pour chaque méthode - AD	130
J.4	Valeurs d'intérêt obtenues par sexe pour chaque méthode - AD	130
J.5	Résultats obtenus pour le test des <i>runs</i> par sexe pour chaque méthode - AD	130
J.6	Résultats obtenus pour le test des signes par sexe pour chaque méthode - AD	131
J.7	Poids attribués par les tests à chaque méthode de positionnement selon leur qualité d'ajus- tement - AD	131
K.1	Résultats démographiques de la population des ouvriers droit sur l'année 2013	132
K.2	Résultats démographiques de la population des ayants droit sur l'année 2013	132
K.3	Résultats démographiques de la population des ouvriers droit sur l'année 2014	133
K.4	Résultats démographiques de la population des ayants droit sur l'année 2014	134
K.5	Résultats démographiques de la population des ouvriers droit sur l'année 2016	134
K.6	Résultats démographiques de la population des ayants droit sur l'année 2016	134
K.7	Résultats démographiques de la population des ouvriers droit sur l'année 2017	135
K.8	Résultats démographiques de la population des ayants droit sur l'année 2017	136

Liste des abréviations, des sigles et des symboles

MAPE Mean Absolute Percentage Error

AD Ayants Droit

AGIRC Association Générale des Institutions de Retraite des Cadres

ARRCO Association pour le Régime de Retraite Complémentaire

CAVIMAC Caisse d'Assurance Vieillesse, Invalidité, Maladie et Accidents des Cultes

CNAV Caisse Nationale d'Assurance Vieillesse

CNAVPL Caisse Nationale d'Assurance Vieillesse des Professions Libérales

CNBF Caisse Nationale des barreaux Français

CNIEG Caisse Nationale des Industries Électriques et Gazières

CNRACL Caisse Nationale de Retraites des Agents des Collectivités Locales

COR Conseil d'Orientation des Retraites

CRPCEN Caisse de Retraite et de Prévoyance des Clercs et Employés de Notaires

CRPCF Caisse de Retraite du Personnel de la Comédie Française

CRPN Caisse de Retraite du Personnel Naviguant

CRPRATP Caisse de Retraite du Personnel du Régime Autonome des Transports Parisiens

CRPSNCF Caisse de Retraite du Personnel de la Société Nationale des Chemins de Fer

CSR Comité de Surveillance des Retraites

ENIM Établissement National des Invalides de la Marine

FRR Fonds de Réserve pour les Retraites

FSPOEIE Fonds Spécial des Pensions des Ouvriers des Établissements Industriels de l'État

FSV Fonds de Solidarité Vieillesse

INED Institut National d'Études Démographiques

INSEE Institut National de la Statistique et des Études Économiques

IRCANTEC Institution de Retraite Complémentaire des Agents non Titulaires de l'État et des Collectivités publiques

IRCEC Institution de Retraite Complémentaire de l'Enseignement et de la Création

LOESS *LOcally weighted Scatterplot Smoother*

MCO Moindres Carrés Ordinaires

MSA Mutuelle Sociale Agricole

LISTE DES TABLEAUX

OD	Ouvrants Droit
PASS	Plafond Annuel de la Sécurité Sociale
PERCO	Plan d'Épargne pour la Retraite Collectif
PERE	Plan d'Épargne Retraite Entreprise
PERP	Plan d'Épargne Retraite Populaire
PSS	Plafond de la Sécurité Sociale
RAFP	Retraite Additionnelle de la Fonction Publique
RGSS	Régime Général de la Sécurité Sociale
RSI	Régime Social des Indépendants
SAM	Salaire Annuel Moyen
SMIC	Salaire Minimum Interprofessionnel de Croissance
SMR	<i>Standardized Mortality Ratio</i>

Annexe A

Le diagramme de Lexis

Dans l'étude de la mortalité, il est important de noter l'intervention du temps sous trois formes :

- l'âge des individus
- l'instant des observations
- la génération dont l'individu fait partie

Afin de représenter les évolutions de chaque individu ou génération aux âges successifs, nous utilisons généralement un diagramme de Lexis.

Trois notions importantes sont mises en avant par le diagramme de Lexis.

Premièrement, il permet le suivi de l'effectif des différentes populations dans le temps, ce qui est également généralement étudié en retraite.

Le diagramme de Lexis est tracé en choisissant les temps d'observation en abscisses et les âges en ordonnées :

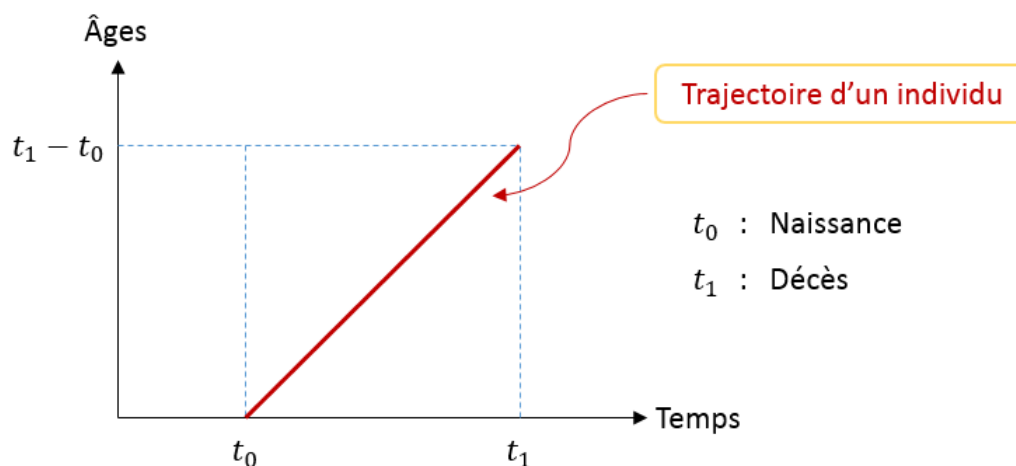


FIGURE A.1 – Représentation simpliste d'un diagramme de Lexis

L'effectif de la population correspond à environ une centaine de cohortes pour lesquelles sera estimé le risque de décès à chaque âge et pour chaque année.

Définissons alors ce que nous caractérisons de cohorte : une cohorte correspond à un groupe d'individus ayant vécu une même expérience telle que l'exposition à un facteur de risque par exemple, et dont le suivi est effectué depuis la date de cette expérience.

Deuxièmement, les enregistrements des populations se font généralement à trois âges. Définissons alors chacun de ces derniers :

- **Âge exact de décès** : différence entre l'instant de décès et l'instant de naissance.
- **Âge révolu** : âge exact à l'instant d'observation.
- **Âge atteint** : différence entre le millésime de décès et le millésime de naissance.

L'âge révolu sur un diagramme de Lexis est représenté comme suit :

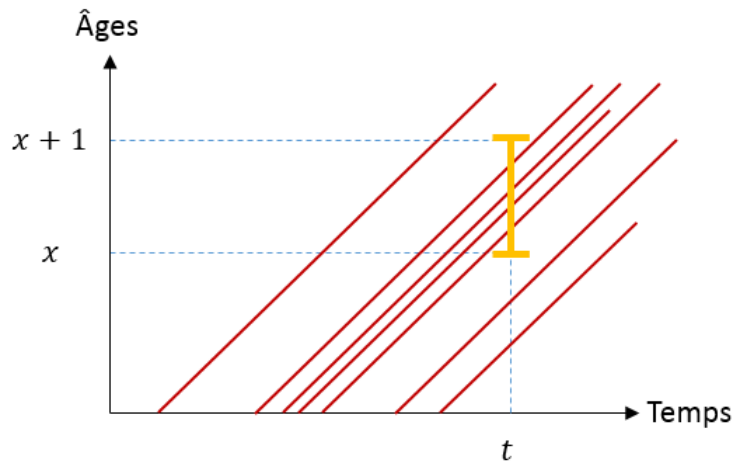


FIGURE A.2 – Représentation simpliste d'âges révolus sur un diagramme de Lexis

À un instant t , le nombre de lignes ayant traversé le segment vertical de couleur orange donne le nombre d'individus ayant x années à l'instant t .

L'âge atteint sur un diagramme de Lexis est représenté comme suit :

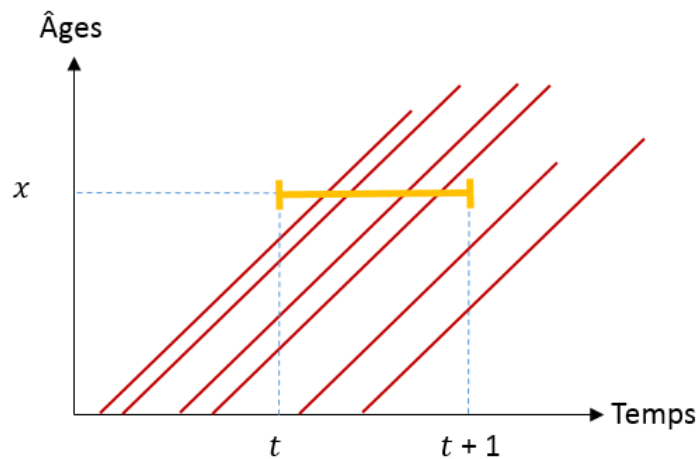


FIGURE A.3 – Représentation simpliste d'âges atteints sur un diagramme de Lexis

Le nombre de lignes traversant le segment horizontal de couleur orange donne le nombre d'individus ayant l'âge x entre t et $t+1$.

Il est également possible de représenter le nombre de décès par âge atteint :

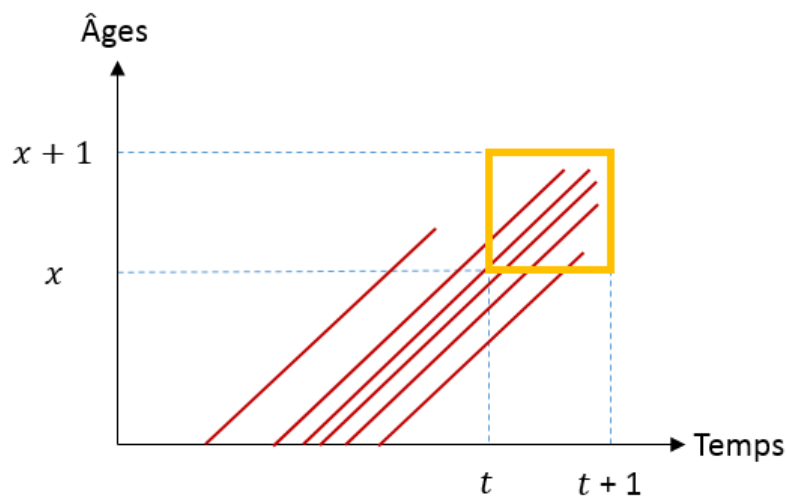


FIGURE A.4 – Représentation simpliste du nombre de décès par âge atteint sur un diagramme de Lexis

Le nombre de lignes s'arrêtant dans le carré de couleur orange correspond au nombre de décès d'individus d'âge x entre t et $t + 1$.

Généralement, dans les régimes de retraite, les effectifs comptabilisés sont des effectifs d'individus d'âges révolus. Dans cette étude, par manque de certaines informations sur les individus constituant la population étudiée, les âges considérés seront des âges atteints.

Enfin dernièrement, la période d'enregistrement correspond toujours à l'année civile et les effectifs sont la plupart du temps mesurés au 31 décembre.

Annexe B

Tables de mortalité TGF-05 et TGH-05

Les tables présentées ci-dessous correspondent à des extraits des tables générationnelles réglementaires TGF-05 et TGH-05 proposées par le Code des Assurances. La tranche d'âges considérée dans ces tables est limitée à la tranche 0 - 120 ans¹.

		TGF-05							
Âge	Gén.	1900	1901	...	1950	1951	...	2004	2005
	0		-	-	...	-	-	...	100 000
1		-	-	...	-	-	...	99 744	99 745
2		-	-	...	-	-	...	99 662	99 663
...	
46		-	-	...	100 000	99 903	...	98 509	98 513
47		-	-	...	99 894	99 797	...	98 435	98 441
48		-	-	...	99 777	99 681	...	98 357	98 363
...	
118		-	-	...	22	24	...	1 104	1 154
119		-	-	...	11	12	...	746	782
120		-	-	...	5	6	...	491	516

TABLE B.1 – Extrait de la table TGF-05

		TGH-05							
Âge	Gén.	1900	1901	...	1950	1951	...	2004	2005
	0		-	-	...	-	-	...	100 000
1		-	-	...	-	-	...	99 676	99 683
2		-	-	...	-	-	...	99 574	99 583
...	
46		-	-	...	100 000	99 868	...	98 745	98 773
47		-	-	...	99 851	99 723	...	98 702	98 732
48		-	-	...	99 687	99 563	...	98 665	98 685
...	
118		-	-	...	2	3	...	408	432
119		-	-	...	1	1	...	266	283
120		-	-	...	-	1	...	169	181

TABLE B.2 – Extrait de la table TGH-05

1. Les valeurs comprises dans les tables ci-dessus sont arrondies à l'individu près

Annexe C

Compléments sur les réformes des retraites

Chacune des réformes présentées ci-dessous ont été établies dans le but d'assurer la pérennité du système de retraite par répartition.

C.1 La réforme de 2003

La réforme 2003 fait suite au premier rapport établi par le COR (qui rappelons-nous, a vu le jour pour la première fois en 2000), qui avait préconisé une réforme pour pallier une situation financière désastreuse annoncée avec les départs à la retraite des *papy-boomers*.

L'augmentation des ressources se ferait à la fois au niveau des taux de cotisation et de la durée d'activité, tandis que la baisse des pensions se ferait au travers une diminution des montants de pension versés et du nombre de pensionnés.

La "Loi Fillon sur les retraites" est signée le 21 août 2003 dans le but de rééquilibrer le système et de grandement atténuer le problème de financement des retraites, en prévoyant et en anticipant les éventuels déficits auxquels les systèmes de retraite devront faire face, ceci sans pour autant impacter les contraintes d'équité qui sont imposées par le principe du système par répartition.

Cette loi est reliée aux gains d'espérance de vie et prévoit différentes mesures pour la pérennité des régimes de retraite :

- Allongement de la durée de cotisation nécessaire pour liquider une pension à taux plein à 60 ans, qui évolue sur les gains de l'espérance de vie dans l'objectif d'une durée de retraite $\frac{1}{3}$ / $\frac{2}{3}$.
- Un système de départs à la retraite anticipés est instauré dans le cas de carrières longues de sorte à ne pas défavoriser les périodes ayant commencé de travailler avant 17 ans (possibilité de partir avant 60 ans si 42 années de cotisation minimum).
- Revalorisation du minimum contributif : un assuré ayant cotisé au SMIC tout au long de sa carrière ne peut percevoir une pension de retraite inférieure à 85% du SMIC.
- Nouveau système de surcote : permet à un assuré qui atteint l'âge de la retraite et la durée légale de cotisation mais qui continue de travailler au-delà, d'obtenir une majoration sur le montant de sa pension.
- Mise en place de nouveaux produits d'épargne retraite individuelle : le Plan d'Épargne Retraite Populaire (PERP) ; mais également collective : le Plan d'Épargne pour la Retraite Collectif (PERCO) et le Plan d'Épargne Retraite Entreprise (PERE), afin d'aider les assurés à compléter leur pension de retraite par les revenus de leur épargne.
- Possibilité de rachat sur des trimestres d'études ou sur des trimestres dont les cotisations étaient insuffisantes.
- Mise en place d'un régime complémentaire par points obligatoire par capitalisation pour les fonctionnaires : la Retraite Additionnelle de la Fonction Publique (RAFP).
- Naissance de l'organisme *GIP Info-retraite* pour faciliter les démarches permettant d'informer tous les actifs sur leurs droits à la retraite.

Toutes ces mesures ont été adoptées dans le but de maintenir l'équilibre des régimes de retraite par répartition en assurant un certain équilibre entre temps de travail et temps de retraite.

C.2 La réforme de 2010

La réforme 2010 fait suite à la crise économique de 2008 qui a eu de grosses répercussions sur le financement des régimes de retraite. En effet elle a entraîné l'augmentation rapide du nombre de chômeurs et par conséquent une baisse importante des cotisations, d'autant plus que les générations du *baby-boom* commencent à liquider leur retraite à cette période. La réforme est alors promulguée le 10 novembre 2010 pour une entrée en vigueur le 1^{er} juillet 2011 avec pour objet l'augmentation progressive de la durée passée en activité afin d'assurer le financement du système sans détériorer le niveau de vie des individus (en augmentant le montant des cotisations ou en diminuant le montant des pensions).

Différentes mesures ont alors été adoptées telles que :

- Le relèvement progressif de l'âge de départ à la retraite sur six années : passage d'ouverture des droits de 60 ans à 62 ans, soit un relèvement de 4 mois par an à partir de 2011, et passage de 65 à 67 ans pour un départ à taux plein à compter de 2016.
- Conservation des dispositions dédiées aux carrières longues mais avec toutefois augmentation de la période de cotisation de deux ans : passage de 41 à 43 années de cotisation pour les personnes ayant commencé à travailler avant 17 ans.
- Prise en compte de la pénibilité : le droit de départ à la retraite est maintenu à 60 ans au taux plein pour les personnes assurées dont l'état de santé a été dégradé des suites des conditions de leur travail.
- Amélioration du calcul des retraites des femmes : les mères de plus de 55 ans avec trois enfants peuvent bénéficier du taux plein à 65 ans si leur carrière a été interrompue dans le cadre de l'éducation de leurs enfants, et resserrage des écarts entre les pensions des hommes et les pensions des femmes.
- Rapprochement des régimes publics et des régimes privés de retraite : les fonctionnaires et les régimes spéciaux sont également concernés par le relèvement de 2 ans de l'âge de départ à la retraite.

C.3 La réforme de 2014

La réforme 2014 est signée le 20 janvier 2014 pour une entrée en vigueur le 1^{er} janvier 2015. Elle a mis en place ou modifié certaines mesures qui sont les suivantes :

- L'augmentation des cotisations de 0.3 points entre 2014 et 2017 sur les cotisations des actifs et des entreprises.
- La majoration de 10% accordée aux retraités ayant élevé trois enfants minimum est dès lors soumise à l'impôt sur le revenu.
- Allongement des durées de cotisation : la durée de cotisation requise pour bénéficier du taux plein augmentera d'un trimestre tous les trois ans à partir de 2020 dans l'objectif d'atteindre 43 années de cotisations en 2035 pour les générations à compter de 1973. Suppression de la règle d'automatisabilité $\frac{2}{3} / \frac{1}{3}$.
- Possibilité de bénéficier d'une retraite anticipée à partir de 60 ans : si poursuite de l'activité à temps partiel après cet âge-là, une fraction de la pension peut être perçue, tout en continuant de payer des cotisations afin d'augmenter le montant de la pension qui sera perçue au moment du départ à la retraite définitif.
- Le Comité de Surveillance des Retraites (CSR) est créé afin de faire face aux aléas financiers subis par les régimes de retraite.

Annexe D

Parts en pourcentage d'individus au sein des populations

D.1 Population globale

	Ouvrants droit	Ayants droit
Population globale	87%	13%
Femmes	26%	96%
Hommes	74%	4%

TABLE D.1 – Parts d'individus au sein de chaque population

	Ouvrants droit	Ayants droit
Femmes	22.62%	12.48%
Hommes	64.38%	0.52%

TABLE D.2 – Parts de chaque population au sein de la population totale des affiliés du régime

D.2 Ouvrants droit par tranche d'âges

	Tranche 1	Tranche 2	Tranche 3	Tranche 4	Tranche 5
Femmes	16.5%	35.9%	34.4%	12.9%	0.3%
Hommes	15.8%	30.8%	40%	13.2%	0.2%

TABLE D.3 – Parts d'OD par tranche d'âges selon le sexe des individus

D.3 Ayants droit par tranche d'âges

	Tranche 1	Tranche 2	Tranche 3	Tranche 4	Tranche 5
Femmes	0.04%	3.19%	27.4%	65.85%	3.51%
Hommes	0.03%	4.08%	41.24%	52.92%	1.73%

TABLE D.4 – Parts d'AD par tranche d'âges selon le sexe des individus

Annexe E

Comparaison des différentes méthodes de lissage

Nous présentons dans cette section les graphes de comparaison des méthodes de lissages appliquées à nos taux bruts pour chacune des populations considérées sur les plages d'âges qui ne sont pas les plus pertinentes à la détermination de la méthode de lissage la plus adaptée à notre modèle, mais qui peuvent néanmoins confirmer les hypothèses émises suite à l'analyse des graphiques précédents. Nous présentons par ailleurs les valeurs des comparaisons statistiques pour la tranche 60-89 ans qui confirment les résultats obtenus sur la population globale.

E.1 Population des Hommes OD

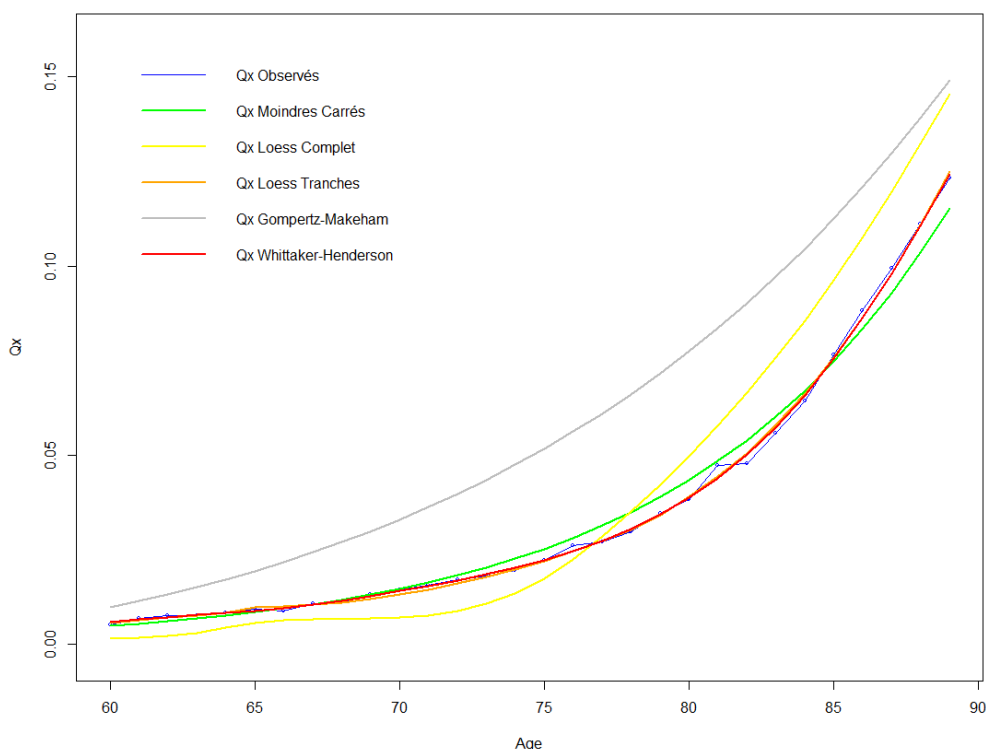


FIGURE E.1 – Comparaison des lissages réalisés - Hommes OD Tranche 4

Le graphique ci-dessus permet de distinguer les lissages de Loess par tranches et de Whittaker-Henderson des autres lissages mais ne permet pas de déterminer laquelle de ces deux méthodes est la meilleure.

E.2 Population des Femmes AD

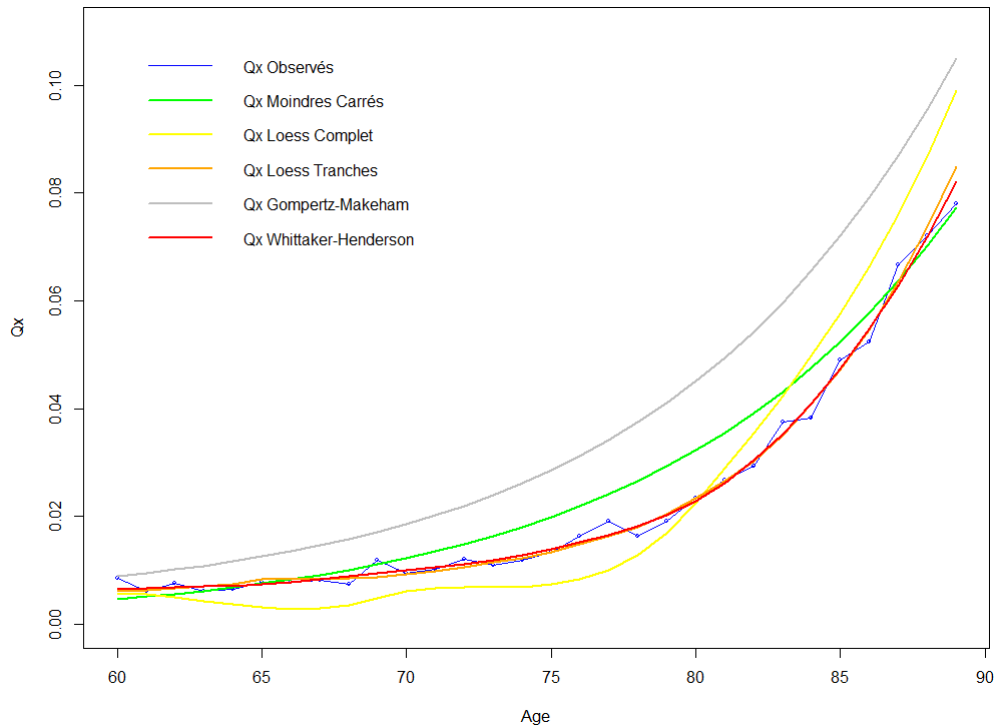


FIGURE E.2 – Comparaison des lissages réalisés - Femmes AD Tranche 4

Les mêmes constats que pour les hommes OD sont établis.

E.3 Population des Hommes AD

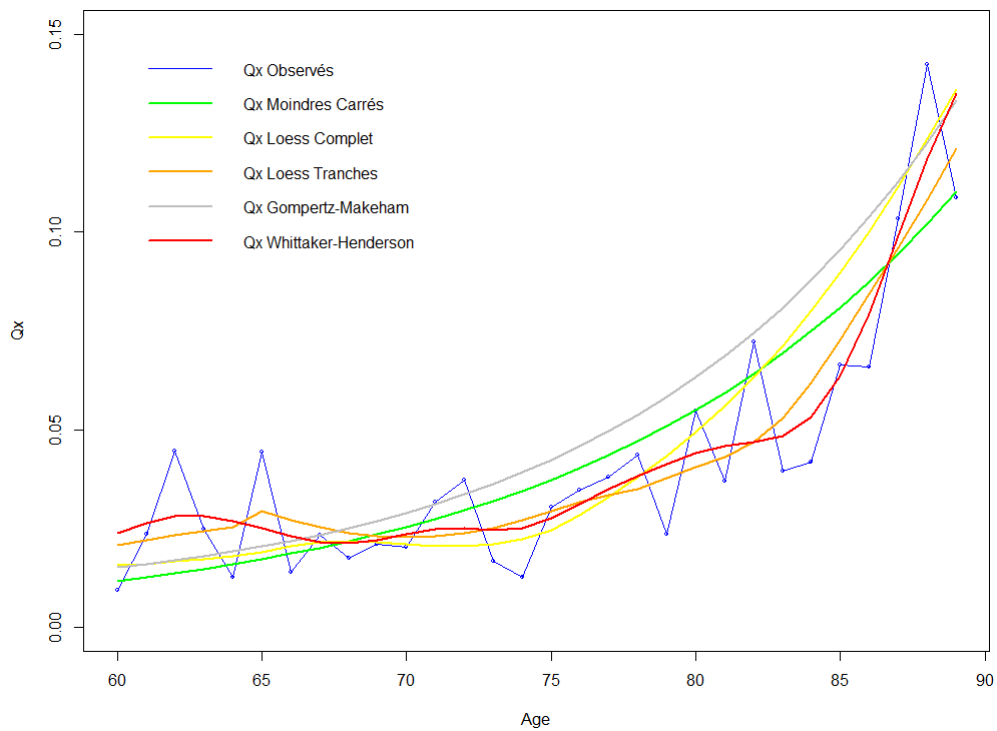


FIGURE E.3 – Comparaison des lissages réalisés - Hommes AD Tranche 4

E.4. VALEURS DU χ^2 SUR LA TRANCHE 60-89 ANS

Sur la FIGURE E.3, une sur-estimation des taux de mortalité par les méthodes des moindres carrés, de Gompertz-Makeham et de Loess complet est de nouveau constatable. Les deux autres méthodes paraissent plus ou moins équivalentes.

E.4 Valeurs du χ^2 sur la tranche 60-89 ans

	Moindres carrés	Makeham	Loess	Loess Tranches	Whittaker-Henderson
Femmes OD	71.81	626.60	2238.84	44.11	30.68
Hommes OD	94.80	4216.39	2355.60	21.48	13.00
Femmes AD	182.79	1112.53	549.12	24.80	17.84
Hommes AD	49.48	56.06	46.20	28.56	23.18

TABLE E.1 – Valeurs du χ^2 obtenues pour chaque lissage et chaque population pour la tranche 60-89 ans

E.5 Valeurs de la *MAPE* sur la tranche 60-89 ans

	Moindres carrés	Makeham	Loess	Loess Tranches	Whittaker-Henderson
Femmes OD	14.87	57.81	66.33	12.07	10.57
Hommes OD	8.91	97.27	35.98	4.09	3.19
Femmes AD	24.60	75.01	38.34	8.36	7.49
Hommes AD	41.55	51.40	36.40	36.23	33.68

TABLE E.2 – *MAPE* obtenues pour chaque lissage et chaque population pour la tranche 60-89 ans

Annexe F

Méthodologie des tests appliqués sur les résidus

F.1 Test de Shapiro

Le premier test appliqué à nos résidus est un test de Shapiro permettant d'évaluer s'ils suivent une loi normale.

Les hypothèses testées sont alors :

- H_0 : les résidus sont normalement distribués;
- H_1 : les résidus ne sont pas normalement distribués.

La statistique du test de Shapiro se traduit comme suit :

$$W = \frac{\left(\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} a_i (x_{(n-i+1)} - x_{(i)}) \right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

avec :

- $x_{(i)}$: la série des résidus triée
- n : le nombre de résidus dans la série
- $\lfloor \frac{n}{2} \rfloor$: la partie entière de $\frac{n}{2}$
- a_i : constantes fournies par des tables spécifiques et qui sont générées à partir de la moyenne et de la matrice des variances co-variances des quantiles d'un échantillon de taille n qui suit une loi normale.

Il est possible d'interpréter la statistique W comme le coefficient de détermination entre la série des quantiles empiriques obtenue à partir des observations et la série des quantiles issue de la loi normale.

Par conséquent, plus la valeur de W sera élevée, plus la distribution des résidus sera proche de celle d'une loi normale.

F.2 Test de Box-Pierce

Le deuxième test appliqué à nos résidus est un test de Box-Pierce consistant à regarder si l'auto-corrélation des résidus n'est pas d'ordre r supérieure à 1.

Nous testons les hypothèses suivantes :

- H_0 : il n'y a pas d'auto-corrélation des résidus d'ordre 1 à r ;
- H_1 : il y a auto-corrélation des résidus d'ordre compris entre 1 et r .

En posant $(\varepsilon_x)_{1 \leq x \leq n}$ la série des résidus, nous construisons le modèle suivant $\forall 1 \leq x \leq n$:

$$\varepsilon_x = \rho_1 \varepsilon_{x-1} + \rho_2 \varepsilon_{x-2} + \dots + \rho_r \varepsilon_{x-r} + \mu_x$$

avec :

- $(\rho_x)_{1 \leq i \leq r}$: les coefficients d'auto-corrélation d'ordre 1 à r
- μ : un bruit blanc

H_0 sera donc acceptée dans le cas où l'ensemble des ρ_i est nul, par conséquent, H_0 sera rejetée dans le cas où au moins un des ρ_i est non nul.

La statistique du test de Box-Pierce est la suivante :

$$Q = n \times \sum_{i=1}^r \hat{\rho}_i^2$$

avec $\hat{\rho}_i$ des estimations des ρ_i par la méthode des Moindres Carrés Ordinaires (MCO) et n le nombre de résidus.

Si H_0 est acceptée, alors la statistique du test est telle que Q suit asymptotiquement une loi du χ_r^2 avec r le nombre de degrés de libertés.

E.3 Test de White

Le dernier test appliqué à nos résidus est un test de White permettant d'évaluer l'homoscédasticité de ces derniers.

Ce test tient compte de l'ensemble des variables explicatives d'un modèle de même que du produit et du carré de ces variables explicatives entre elles.

En posant $(\varepsilon_x)_{1 \leq x \leq n}$ la série des résidus et Y une variable explicative du modèle, la régression suivante est réalisée $\forall 1 \leq x \leq n$:

$$\varepsilon_x^2 = \alpha_0 + \alpha_1 Y_x + \alpha_2 Y_x^2$$

Les hypothèses testées s'écrivent :

- H_0 : il y a homoscédasticité des résidus, c'est-à-dire que $Var(\varepsilon_x)$ est constante et égale à α_0 et les paramètres α_1 et α_2 sont nuls;
- H_1 : il y a hétéroscédasticité des résidus, c'est-à-dire que $Var(\varepsilon_x)$ n'est pas constante et est égale à $\alpha_0 + \alpha_1 Y_x + \alpha_2 Y_x^2$.

La statistique du test de White est traduite de la sorte :

$$S = n \times R^2$$

avec :

- n : le nombre d'observations
- R^2 : le coefficient de détermination de la régression

Si H_0 est acceptée, alors la statistique du test est telle que S suit asymptotiquement une loi du χ_2^2 (le degré de liberté est égal à 2 du fait des deux contraintes suivantes : $\alpha_1 = 0$ et $\alpha_2 = 0$).

Annexe G

Diagrammes *Quantile-Quantile* des résidus entre les taux bruts et les taux lissés

Cette annexe a pour but de montrer les différences de normalité des résidus obtenus entre les taux bruts et les taux lissés retenus en fonction de la sélection de plage d'âge pour chacune des populations (population totale ou plage d'âges contenant le plus d'individus).

G.1 Hommes OD

QQ Plot des résidus - Hommes OD 25:95 ans

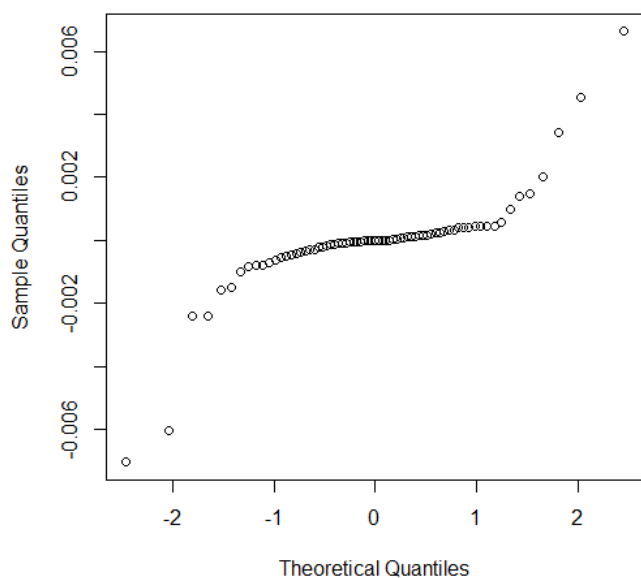


FIGURE G.1 – Diagramme *Quantile-Quantile* de la population des hommes OD

QQ Plot des résidus - Hommes OD 60:87 ans

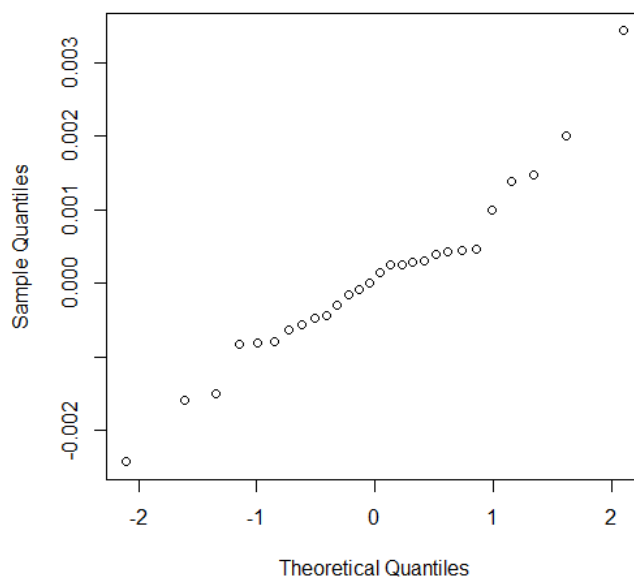


FIGURE G.2 – Diagramme *Quantile-Quantile* de la population des hommes OD pour la tranche 60-87 ans

G.2 Femmes AD

QQ Plot des résidus - Femmes AD 27:95 ans

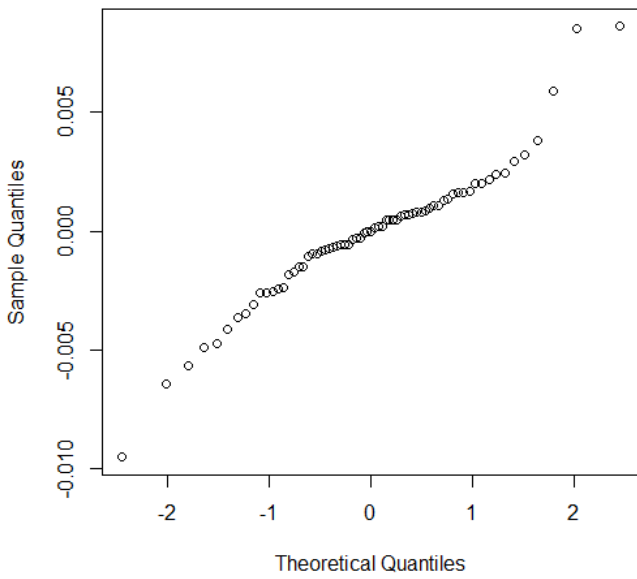


FIGURE G.3 – Diagramme *Quantile-Quantile* de la population des femmes AD

QQ Plot des résidus - Femmes AD 60:84 ans

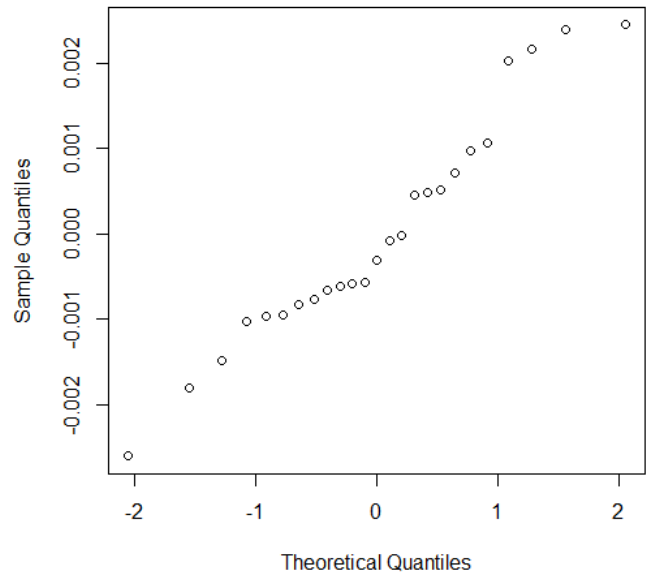


FIGURE G.4 – Diagramme *Quantile-Quantile* de la population des femmes AD pour la tranche 60-84 ans

G.3 Hommes AD

QQ Plot des résidus - Hommes AD 35:95 ans

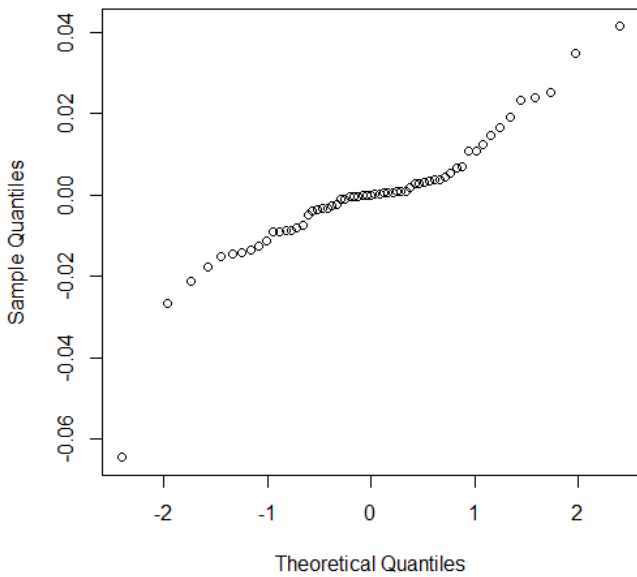


FIGURE G.5 – Diagramme *Quantile-Quantile* de la population des hommes AD

QQ Plot des résidus - Hommes AD 60:84 ans

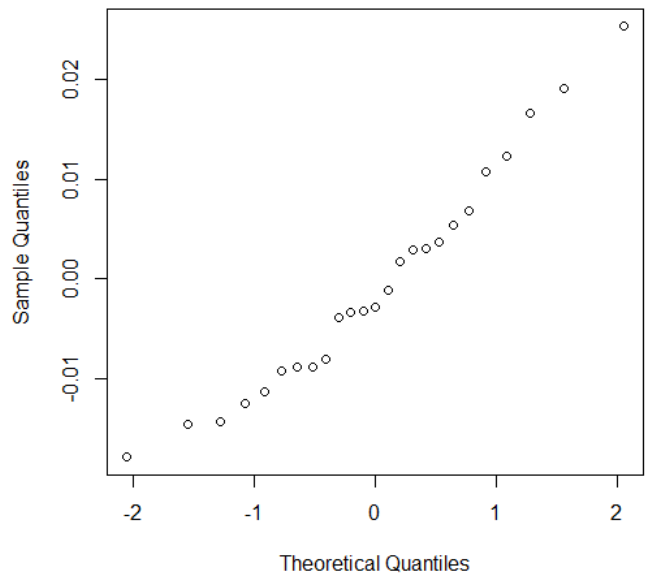


FIGURE G.6 – Diagramme *Quantile-Quantile* de la population des hommes AD pour la tranche 60-84 ans

Annexe H

Résultats complémentaires obtenus à partir des méthodes du *package ELT*

H.1 Méthode 1 appliquée à la population des ouvriers droit : complément sur les SMR

Plage d'âges	60 : 90 ans	60 : 95 ans	70 : 95 ans	70 : 100 ans
Femmes	0.80675	0.80758	0.80658	0.80503
Hommes	0.80569	0.81508	0.84938	0.85242

TABLE H.1 – Valeurs du SMR pour différentes plages d'âges sur la population des OD

H.2 Méthodes de positionnement appliquées sur la population des ayants droit

H.2.1 Méthode 1 : Approche à un paramètre avec le SMR

Nous appliquons le calcul du SMR à différentes plages d'âges afin de sélectionner celle qui semble la plus pertinente à notre population pour la suite de notre étude :

Plage d'âges	60 : 100 ans	65 : 100 ans	65 : 95 ans	65 : 90 ans
Femmes	0.84054	0.83898	0.82978	0.82006
Hommes	0.80951	0.78594	0.77542	0.79253

TABLE H.2 – Valeurs du SMR pour différentes plages d'âges sur la population des AD

Plage d'âges	60 : 90 ans	60 : 95 ans	70 : 95 ans	70 : 100 ans
Femmes	0.82291	0.83159	0.82765	0.83718
Hommes	0.82497	0.80049	0.76428	0.77583

TABLE H.3 – Valeurs du SMR pour différentes plages d'âges sur la population des AD

Nous pouvons constater, tout comme pour la population des OD, que du fait que les valeurs du SMR sont inférieures à 1, la population de notre régime a une mortalité plus faible que la population représentée par la table de référence.

La plage d'âges que nous retenons pour notre population d'ayants droit est la tranche 60-90.

H.2.2 Méthode 2 : Approche à deux paramètres avec un modèle relationnel

Nous considérons la plage d'âges 60-90 ans comme sélectionnée lors de l'application de la méthode précédente.

Les estimations des paramètres α et β pour la plage d'âges considérée sont rassemblées dans le tableau suivant :

	$\hat{\alpha}$	$\hat{\beta}$
Femmes	-0.2946325	0.9782125
Hommes	-0.2275678	1.0264050

TABLE H.4 – Estimations des paramètres α et β pour chacune des populations des AD

H.2.3 Méthode 3 : Approche par un modèle linéaire généralisé de Poisson incluant des interactions avec l'âge et l'année calendaire

Cette fois encore, nous considérons la plage d'âges déterminée lors de l'exécution de la première méthode de positionnement, c'est-à-dire la plage 60-90 ans.

De la même façon que pour la population des OD, la période d'observation reste 2013-2017, la période commune avec la référence est alors inférieure à 10 ans, c'est pourquoi les interactions avec l'année calendaire ne sont pas prises en compte. Par conséquent, seuls trois paramètres sont estimés : β_0 , β_1 et β_2 . Les résultats sont regroupés dans le tableau suivant :

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
Femmes	4.64186	1.31769	-0.04635
Hommes	18.9574	2.5880	-0.1811

TABLE H.5 – Estimations des paramètres β_0 , β_1 et β_2 pour chacune des populations des AD

Annexe I

Méthodes de validation d'ajustement

I.1 χ^2

Rappelons que le χ^2 est un indicateur permettant d'évaluer la qualité de l'ajustement du modèle effectué. Sa valeur est ici déterminée par :

$$\chi^2 = \sum_{(x,t)} \frac{(d_{x,t} - E_{x,t} \tilde{q}_{x,t})^2}{E_{x,t} \tilde{q}_{x,t} (1 - \tilde{q}_{x,t})}$$

avec $\tilde{q}_{x,t}$ les taux de mortalité ajustés.

La valeur du χ^2 est recherchée de sorte à être la plus faible possible.

I.2 R^2

Le R^2 , qui rappelle nous correspond au coefficient de détermination et donne une évaluation de l'adéquation du modèle d'ajustement aux données observées, est calculé par la formule suivante :

$$R^2 = 1 - \frac{\sum_{(x,t)} (\hat{q}_{x,t} - \tilde{q}_{x,t})^2}{\sum_{(x,t)} \left(\hat{q}_{x,t} - \frac{\sum_{(x,t)} \hat{q}_{x,t}}{n} \right)^2}$$

avec n correspondant au nombre d'observations et $\tilde{q}_{x,t}$ les taux de mortalité ajustés.

La valeur du R^2 est recherchée de sorte à être la plus proche possible de 1.

I.3 *MAPE*

La *MAPE* est un pourcentage d'erreur qui pour rappel, mesure l'exactitude de l'ajustement par rapport aux observations :

$$MAPE = \frac{\sum_{(x,t)} |(\hat{q}_{x,t} - \tilde{q}_{x,t}) / \hat{q}_{x,t}|}{\sum_{(x,t)} d_{x,t}} \times 100$$

avec $\tilde{q}_{x,t}$ les taux de mortalité ajustés.

La valeur de la *MAPE* est recherchée de sorte à être la plus faible possible.

Annexe J

Résultats des tests du *package ELT* et comparaisons sur la population des ayants droit

Nous présentons ici l'ensemble des résultats obtenus pour chacun des tests présentés, appliqués aux trois méthodes de positionnement mises en application sur la population des AD.

- Test du rapport de vraisemblance :

TEST DE VRAISEMBLANCE	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
χ^2	230.98	363.17	221.18	236.26	264.29	224.72
Seuil	185.05	185.05	185.05	185.05	185.05	185.05
<i>p-value</i>	0.0001	0	0.0004	0	0	0.0002
Hypothèse retenue	H_1	H_1	H_1	H_1	H_1	H_1

TABLE J.1 – Résultats obtenus pour le test du rapport de vraisemblance par sexe pour chaque méthode - AD

Pour chacune des deux populations, aucune des trois méthodes n'est acceptée. Les *p-values* sont très faibles voire nulles, la grande dispersion des données pour cette population est un facteur pouvant expliquer cette faible adéquation de l'ajustement à nos données.

- Test du SMR :

TEST DU SMR	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
SMR	1	1.0237	1	1	1.0867	1
χ^2	0.0044	1.7613	0.0088	0.0181	1.4958	0.0361
Seuil	1.6449	1.6449	1.6449	1.6449	1.6449	1.6449
<i>p-value</i>	0.4982	0.0391	0.4965	0.4928	0.0674	0.4856
Hypothèse retenue	H_0	H_1	H_0	H_0	H_0	H_0

TABLE J.2 – Résultats obtenus pour le test du SMR par sexe pour chaque méthode - AD

Pour ce test-ci, la plupart des *p-values* est acceptable à l'exception de celle de la méthode 2 pour les femmes, celle retournée par cette même méthode dans le cas des hommes reste néanmoins très faible. Les autres, aux alentours de 50%, reflètent une représentation de la réalité de qualité assez moyenne par les taux ajustés.

- Test de Wilcoxon :

TEST DE WILCOXON	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
w	6656	7030	6113	6092	6857	6520
χ^2	1.0906	1.7588	0.1206	0.0831	1.4497	0.8477
Seuil	1.96	1.96	1.96	1.96	1.96	1.96
p-value	0.2754	0.0786	0.904	0.9338	0.1471	0.3966
Hypothèse retenue	H_0	H_0	H_0	H_0	H_0	H_0

TABLE J.3 – Résultats obtenus pour le test de Wilcoxon par sexe pour chaque méthode - AD

Nous constatons que l'ensemble des *p-values* est acceptable, et que deux d'entre elles se distinguent : celles de la troisième méthode pour les femmes et de la première méthode pour les hommes. ces deux méthodes semblent très bien s'adapter à nos populations respectives selon ce test.

- Valeurs d'intérêt :

VALEURS D'INTÉRÊT	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
χ^2	238.28	240.52	215.82	233.5	268.61	173.99
R^2	0.9458	0.9453	0.9458	0.5538	0.5526	0.5679
MAPE	26.58	26.34	27.93	54.5	53.21	45.73
Déviante	461.96	726.33	442.36	472.51	528.58	449.44

TABLE J.4 – Valeurs d'intérêt obtenues par sexe pour chaque méthode - AD

Rappelons que le χ^2 , la *MAPE* et la déviante doivent être minimisés tandis que le R^2 doit être maximisé et au plus proche de 1. Nous constatons d'après ces résultats, que les valeurs des R^2 sont à peu près égales les trois méthodes de positionnement pour chacune des populations. De plus, ces valeurs reflètent la répartition hommes-femmes au sein de la population, en effet les R^2 sont élevés pour la population des femmes, dont l'abondance est plus importante que chez les hommes sur la plage d'âges sélectionnée, pour des R^2 plus faibles chez les hommes. Les résultats des autres valeurs d'intérêt semblent se tenir entre les méthodes même si nous pourrions laisser un léger avantage à la méthode 3.

- Test des runs :

TEST DES RUNS	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
Nombre de runs	76	77	81	73	67	69
Signes (-)	69	70	77	85	74	84
Signes (+)	86	85	78	70	81	71
$ \chi^2 $	0.2558	0.126	0.4035	0.7768	1.8317	1.4535
Seuil	1.96	1.96	1.96	1.96	1.96	1.96
p-value	0.7981	0.8998	0.6866	0.4373	0.067	0.1461
Hypothèse retenue	H_0	H_0	H_0	H_0	H_0	H_0

TABLE J.5 – Résultats obtenus pour le test des runs par sexe pour chaque méthode - AD

L'ensemble des *p-values* est acceptable pour les deux populations, bien que celle de la méthode 2 reste très faible pour les hommes alors qu'il s'agit de la meilleure méthode retournée pour la population des femmes.

- Test des signes :

TEST DES SIGNES	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
Signes (+)	86	85	78	70	81	71
Signes (-)	69	70	77	85	74	84
χ^2	1.2852	1.1245	0	1.1245	0.4819	0.9639
Seuil	1.96	1.96	1.96	1.96	1.96	1.96
<i>p-value</i>	0.1987	0.2608	1	0.2608	0.6299	0.3351
Hypothèse retenue	H_0	H_0	H_0	H_0	H_0	H_0

TABLE J.6 – Résultats obtenus pour le test des signes par sexe pour chaque méthode - AD

Pour ce test-ci, les *p-values* sont toutes acceptables et très bonnes dans le cas de la méthode 3 pour les femmes (selon ce test, la méthode 3 est parfaitement adapté comme ajustement de la population des femmes AD de notre jeu de données), et pour la méthode 2 chez les hommes.

Nous rencontrons à la vue de ces résultats, une certaine difficulté à distinguer laquelle de ces trois méthodes de positionnement est la mieux adaptée à notre jeu de données. Afin d'éclaircir ce point, nous décidons de regrouper les résultats en mettant en place un système de poids selon l'ordre d'arrivée de la méthode en terme de qualité retournée par les tests (comme dans le cas des OD, les méthodes sont ordonnées en fonction de leurs résultats par test, un poids de 3 est donnée à la meilleure méthode, un poids de 2 à celle qui arrive en deuxième position, et enfin un poids de 1 à la moins bonne des méthodes), la moyenne de ces poids (recherchée la plus grande possible) permettra de distinguer quelle méthode apparaît être mieux adaptée a nos populations.

SYNTHÈSE DES TESTS	Femmes			Hommes		
	Méthode 1	Méthode 2	Méthode 3	Méthode 1	Méthode 2	Méthode 3
Vraisemblance	2	1	3	2	1	3
SMR	3	1	2	3	1	2
Wilcoxon	2	1	3	3	1	2
χ^2	2	1	3	2	1	3
R^2	2	1	3	2	1	3
MAPE	2	3	1	1	2	3
Déviance	2	1	3	2	1	3
Runs	2	3	1	3	1	2
Signes	1	2	3	1	3	2
MOYENNE	2	1.56	2.44	2.11	1.33	2.56

TABLE J.7 – Poids attribués par les tests à chaque méthode de positionnement selon leur qualité d'ajustement - AD

Nous constatons, de par cette pondération, que l'utilisation de la méthode 3 semble plus favorable que les deux précédentes à chacune de nos deux populations d'ayants droit.

Les valeurs d'intérêts calculées pour cette méthode seront comparées à celles calculées sur la même plage d'âges sur les taux ajustés retournés par la méthode d'ajustement par abattement de la table de référence.

Annexe K

Comparaison de la nouvelle loi avec les observations annuelles

En ce qui concerne l'année d'observation 2013, nous obtenons les résultats démographiques suivants pour chacune de nos populations :

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	72 741	54.33	82.19	612	665	+9%
Hommes	213 568	55.52	79.59	2 993	3 010	+1%
Total	286 309	55.21	80.03	3 605	3 675	+2%

TABLE K.1 – Résultats démographiques de la population des ouvriers droit sur l'année 2013

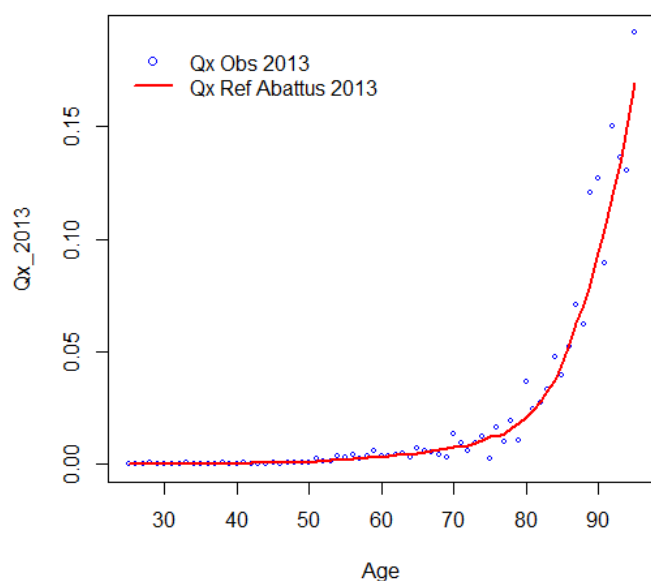


FIGURE K.1 – Femmes OD - 2013

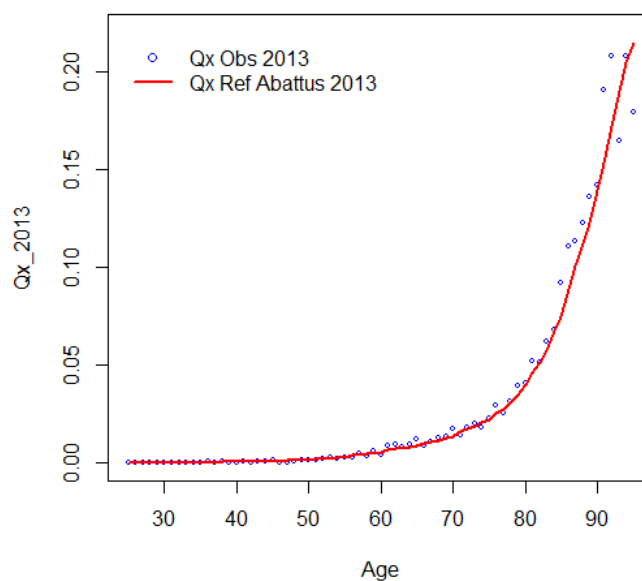


FIGURE K.2 – Hommes OD - 2013

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	41 765	79.29	89.03	1 933	2 248	+16%
Hommes	1 575	76.07	85.88	88	96	+9%
Total	43 340	79.17	88.89	2 021	2 344	+16%

TABLE K.2 – Résultats démographiques de la population des ayants droit sur l'année 2013

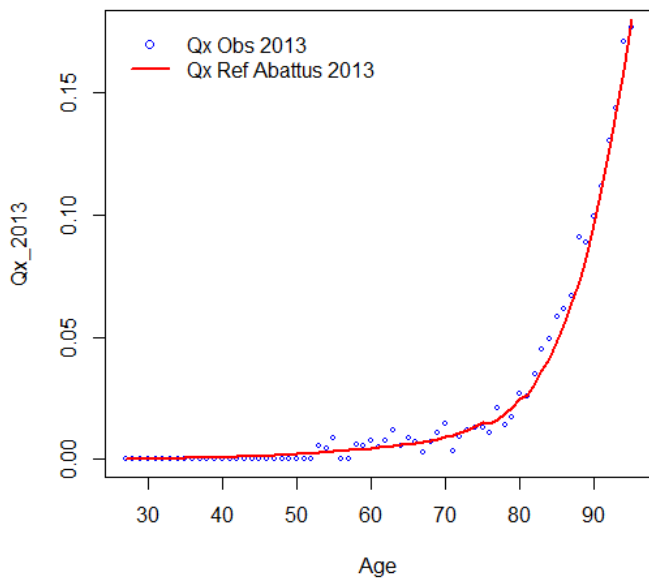


FIGURE K.3 – Femmes AD - 2013

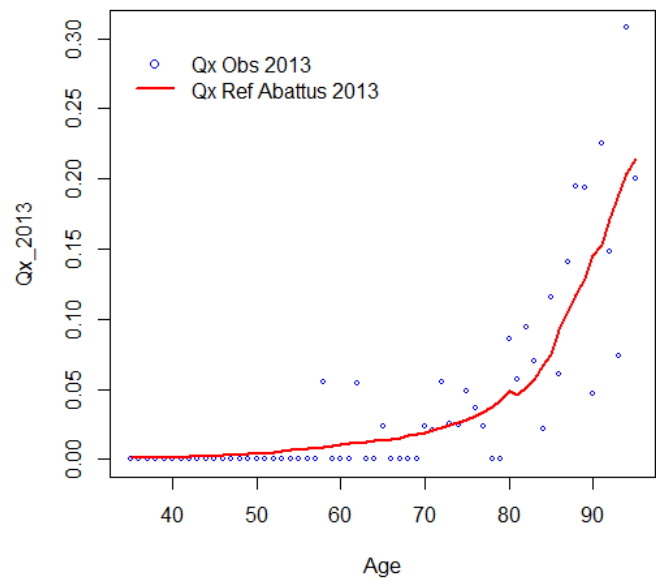


FIGURE K.4 – Hommes AD - 2013

En ce qui concerne l'année d'observation 2014, nous obtenons les résultats démographiques suivants pour chacune de nos populations :

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	74 401	54.38	81.81	630	608	-4%
Hommes	216 331	55.53	79.34	3 032	2 816	-7%
Total	290 732	55.24	79.76	3 663	3 424	-7%

TABLE K.3 – Résultats démographiques de la population des ouvrants droit sur l'année 2014

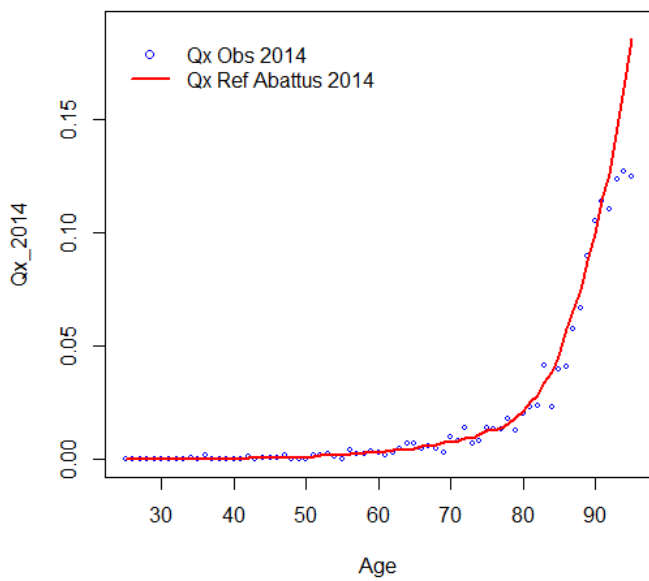


FIGURE K.5 – Femmes OD - 2014

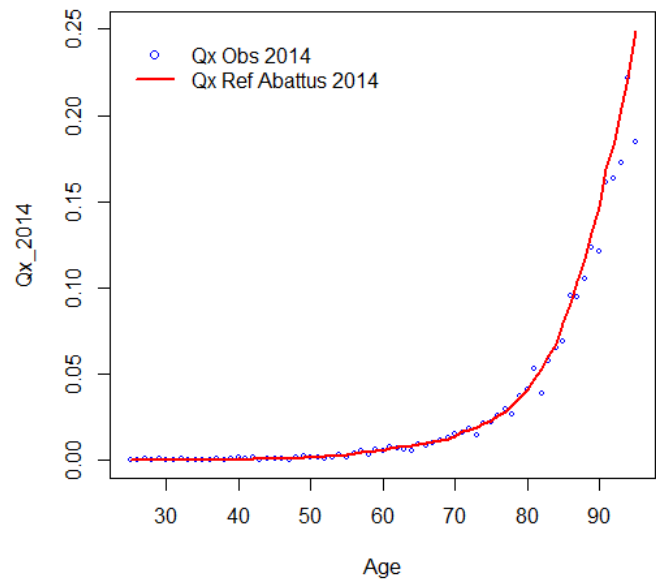


FIGURE K.6 – Hommes OD - 2014

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	41 649	79.47	88.47	1 952	2 057	+5%
Hommes	1 672	75.88	81.41	92	102	+11%
Total	43 321	79.33	88.15	2 044	2 159	+6%

TABLE K.4 – Résultats démographiques de la population des ayants droit sur l'année 2014

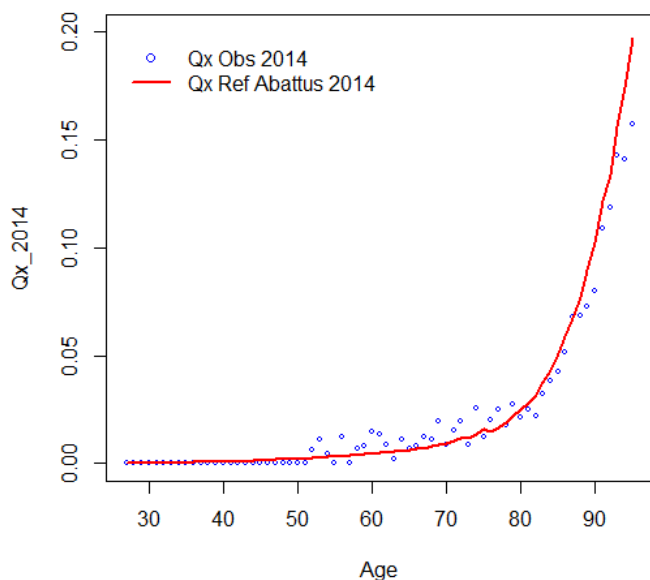


FIGURE K.7 – Femmes AD - 2014

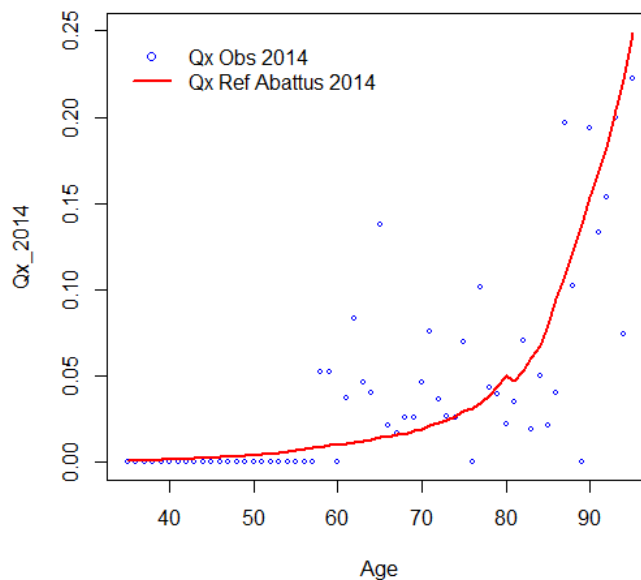


FIGURE K.8 – Hommes AD - 2014

En ce qui concerne l'année d'observation 2016, nous obtenons les résultats démographiques suivants pour chacune de nos populations :

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	76 741	54.84	83.56	677	640	-6%
Hommes	219 551	55.76	79.89	3 122	2 864	-8%
Total	296 292	55.52	80.54	3 799	3 504	-8%

TABLE K.5 – Résultats démographiques de la population des ouvrants droit sur l'année 2016

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	41 501	79.90	89.85	2 028	2 140	+6%
Hommes	1 799	76.31	84.24	101	112	+11%
Total	43 300	79.75	89.59	2 128	2 252	+6%

TABLE K.6 – Résultats démographiques de la population des ayants droit sur l'année 2016

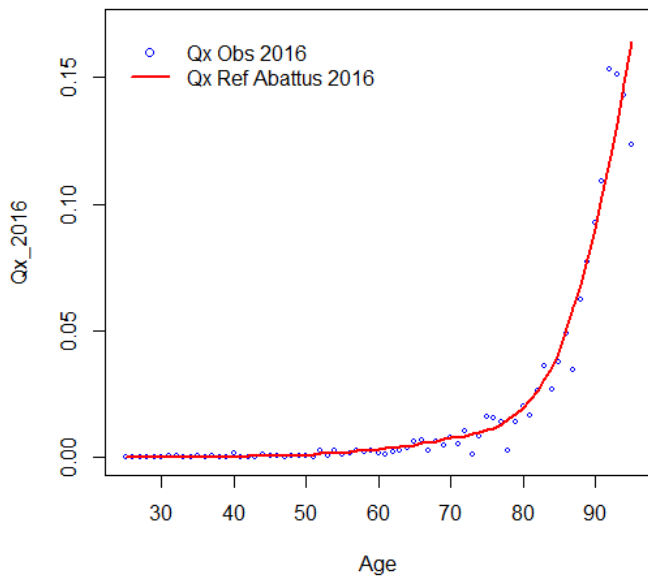


FIGURE K.9 – Femmes OD - 2016

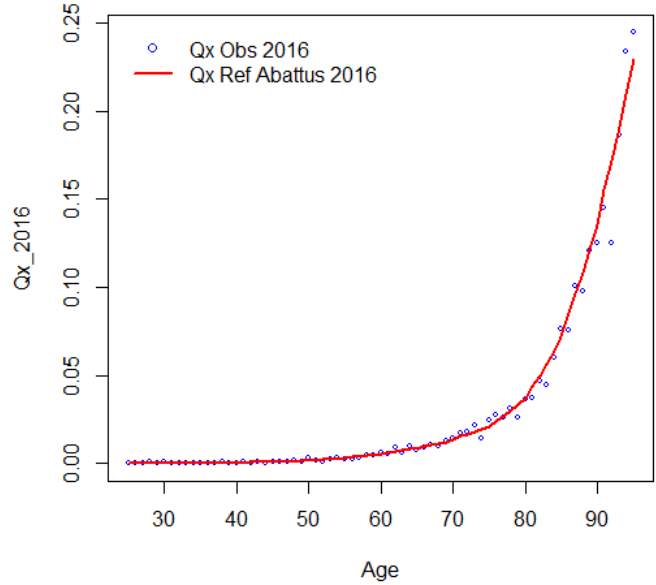


FIGURE K.10 – Hommes OD - 2016

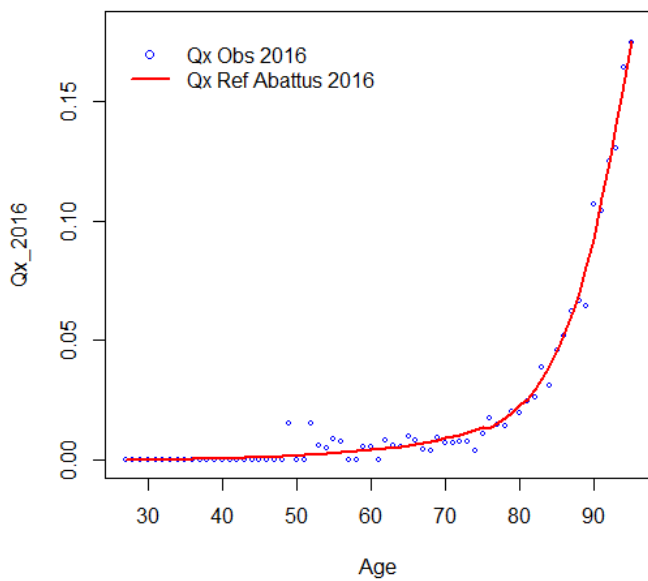


FIGURE K.11 – Femmes AD - 2016

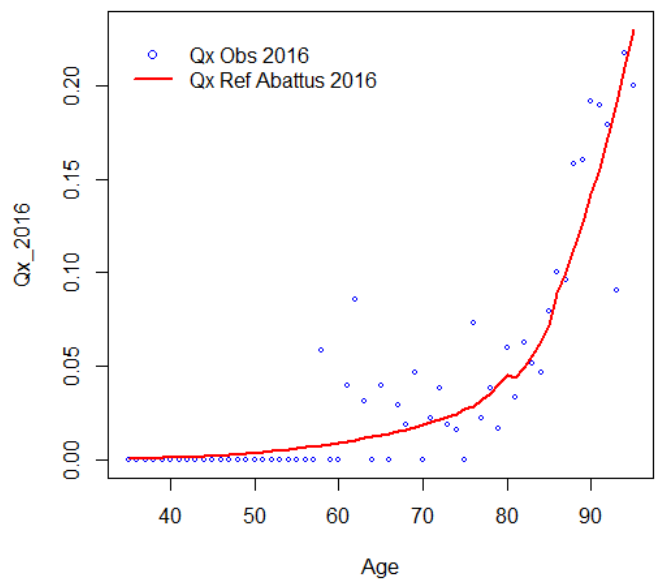


FIGURE K.12 – Hommes AD - 2016

En ce qui concerne l'année d'observation 2017, nous obtenons les résultats démographiques suivants pour chacune de nos populations :

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	78 097	55.19	83.03	700	761	+9%
Hommes	221 282	55.98	80.32	3 175	2 986	-6%
Total	299 379	55.77	80.81	3 874	3 747	-3%

TABLE K.7 – Résultats démographiques de la population des ouvriers droit sur l'année 2017

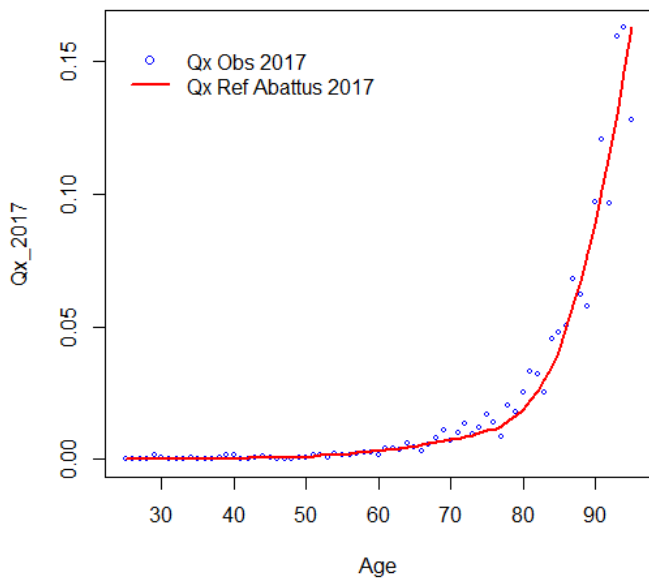


FIGURE K.13 – Femmes OD - 2017

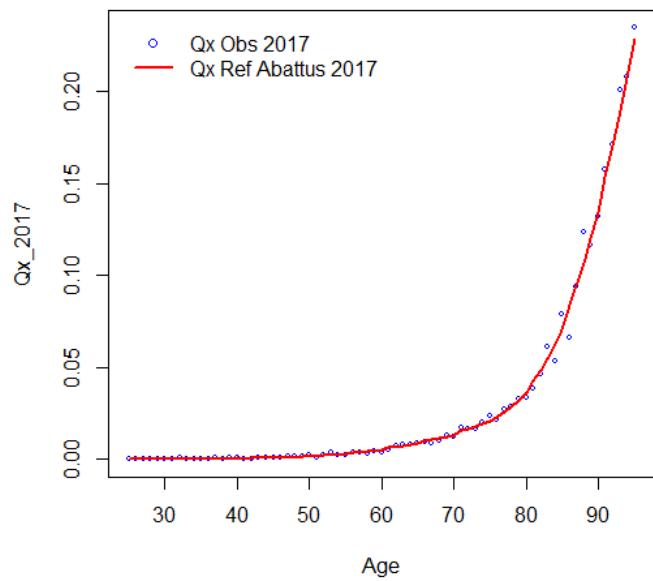


FIGURE K.14 – Hommes OD - 2017

	Population sous risque	Âge moyen population	Âge moyen au décès	Nombre décès attendus	Nombre décès observés	Écart constaté
Femmes	41 412	80.10	89.85	2 063	2 307	+12%
Hommes	1 854	76.61	85.65	105	92	-13%
Total	43 266	79.95	89.65	2 169	2 399	+11%

TABLE K.8 – Résultats démographiques de la population des ayants droit sur l'année 2017

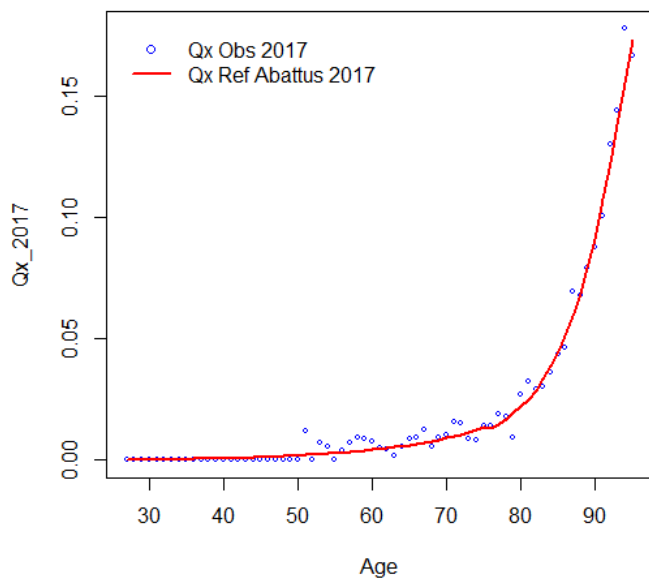


FIGURE K.15 – Femmes AD - 2017

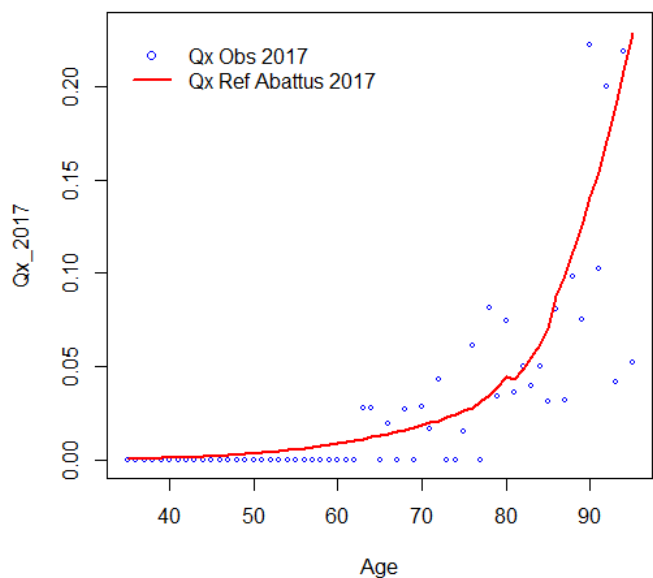


FIGURE K.16 – Hommes AD - 2017

Globalement, les écarts constatés entre ce qui est observé et ce qui est attendu par les nouvelles tables sont acceptables (écarts inférieurs à 10%). ceci peut également être constaté graphiquement. En effet, les taux des nouvelles tables se rapprochent plutôt bien des taux bruts et ce, particulièrement chez les OD.

Annexe L

Espérance de vie

L.1 Comparaison des espérances de vie par sexe

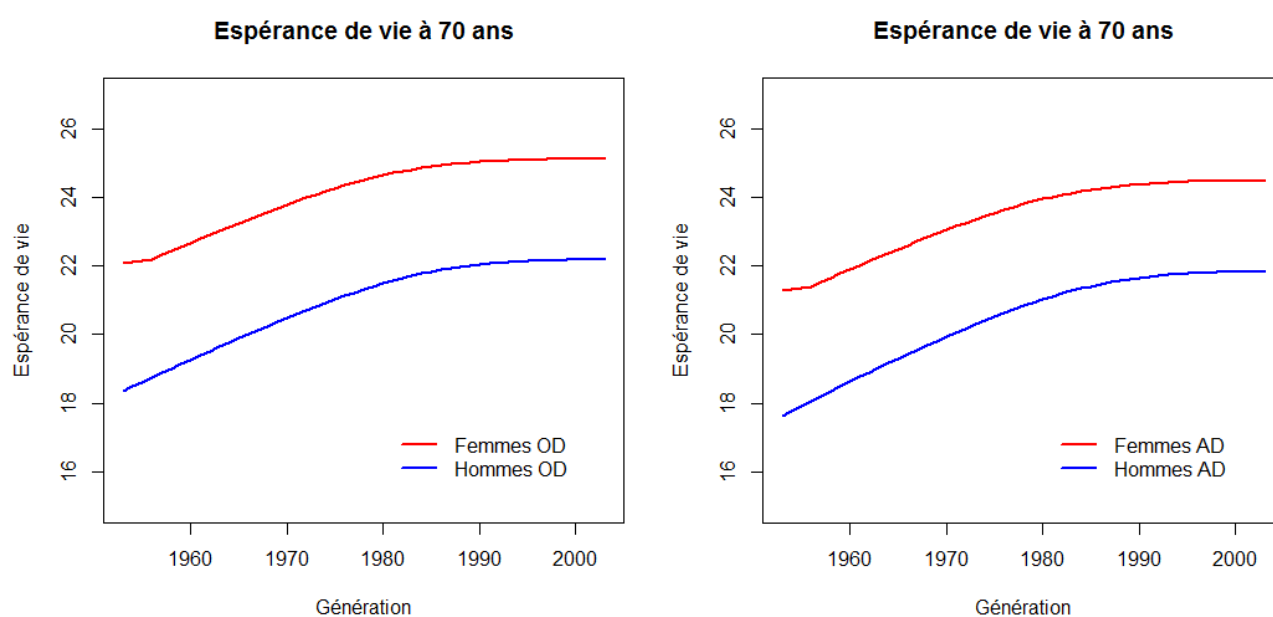


FIGURE L.1 – Comparaison des espérances de vie à 70 ans des populations en fonction de leur sexe

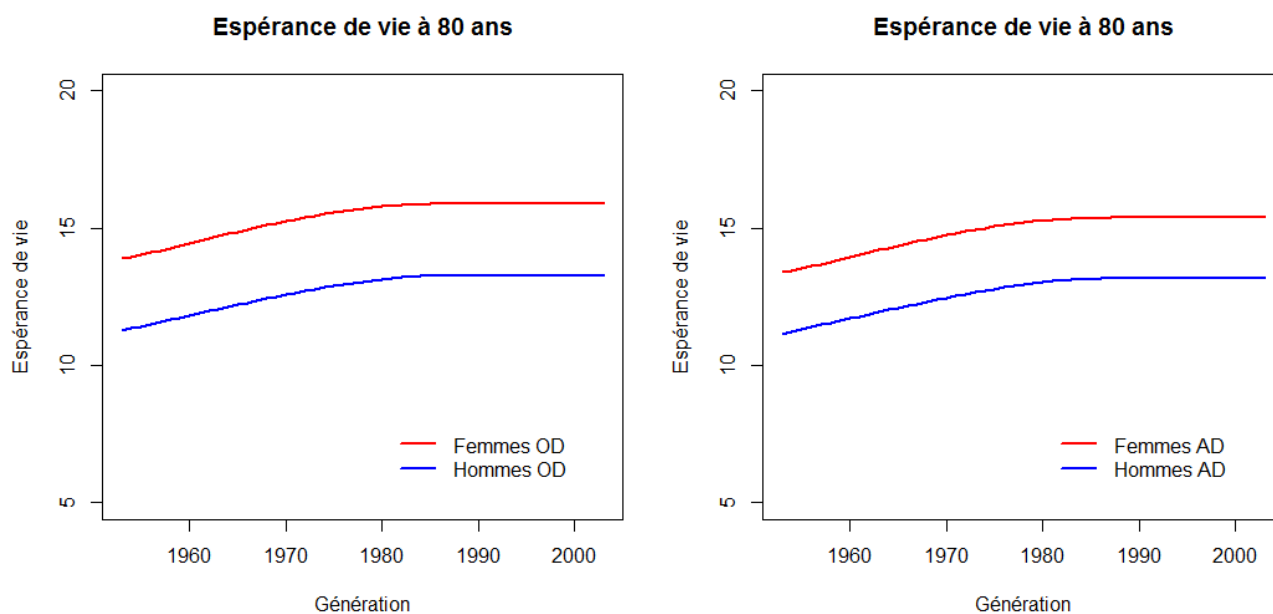


FIGURE L.2 – Comparaison des espérances de vie à 70 ans des populations en fonction de leur sexe

L.2 Comparaison des espérances de vie par qualité

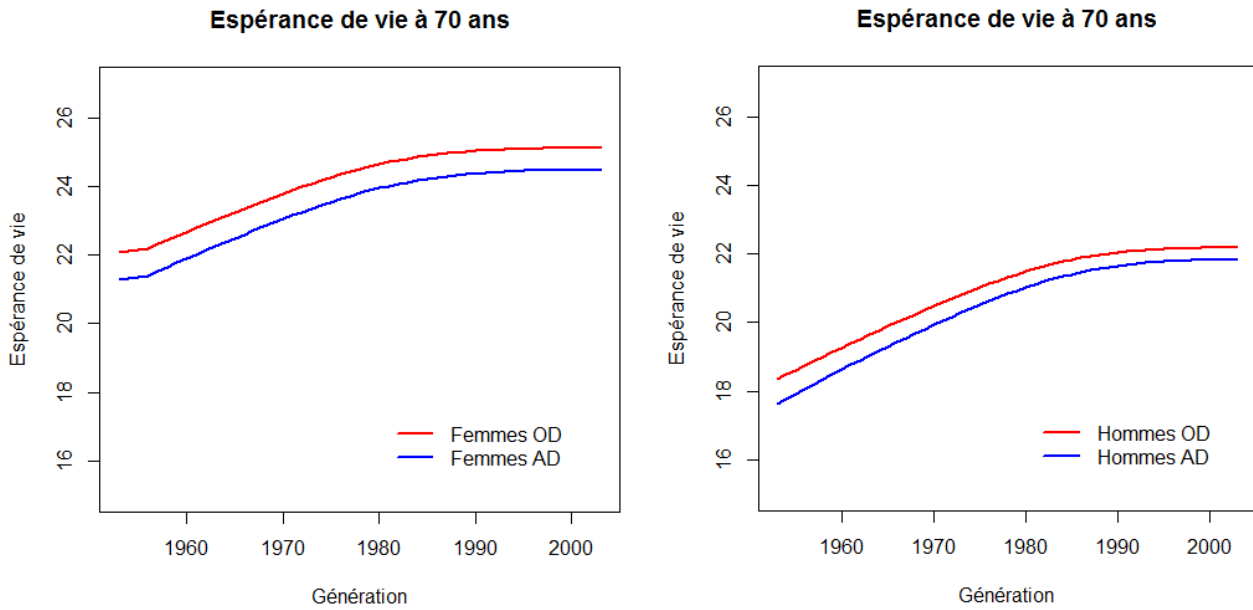


FIGURE L.3 – Comparaison des espérances de vie à 80 ans des populations en fonction de leur qualité

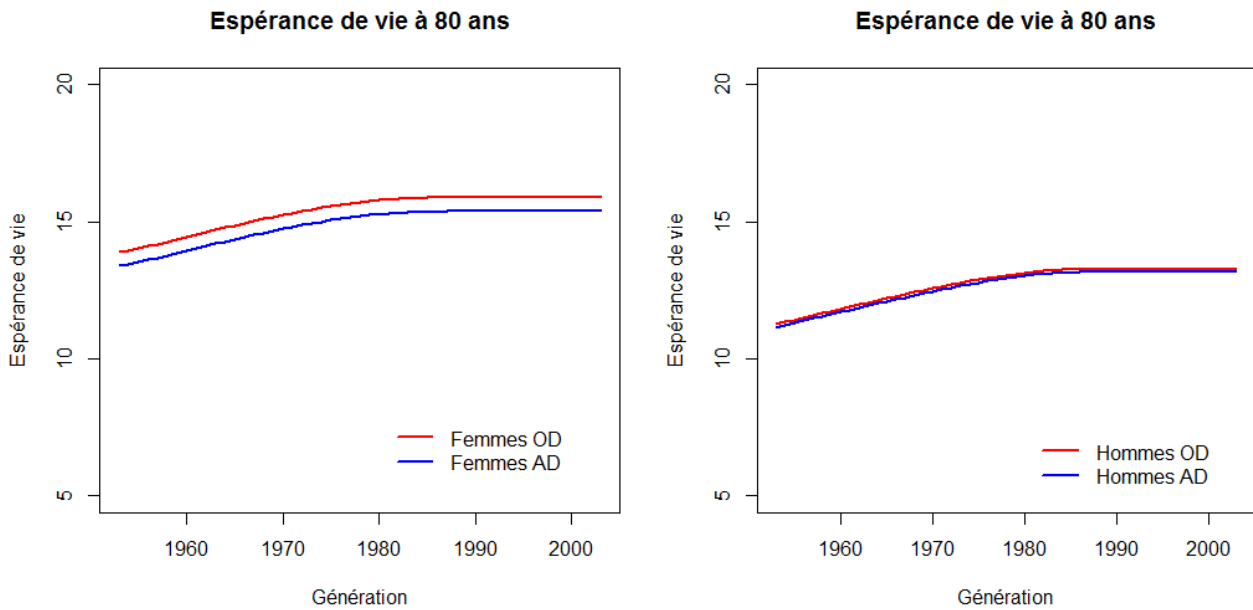


FIGURE L.4 – Comparaison des espérances de vie à 80 ans des populations en fonction de leur qualité

Annexe M

Tables de mortalités ajustées

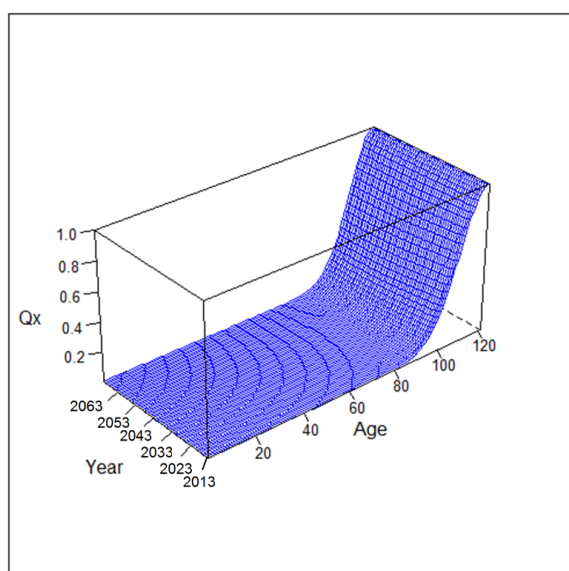


FIGURE M.1 – Table de mortalité ajustée des femmes OD

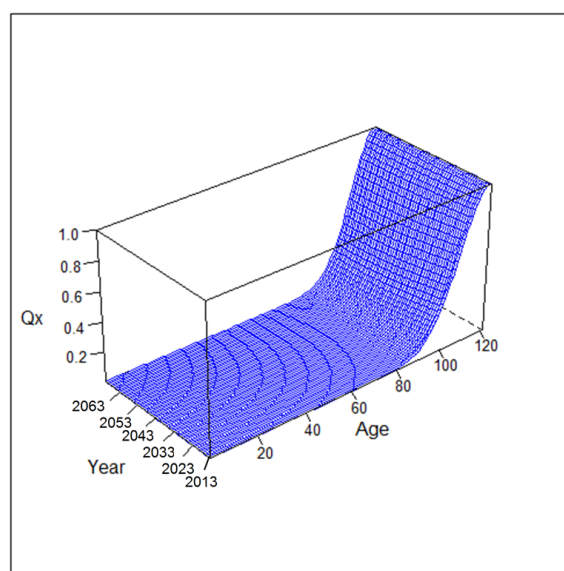


FIGURE M.2 – Table de mortalité ajustée des hommes OD

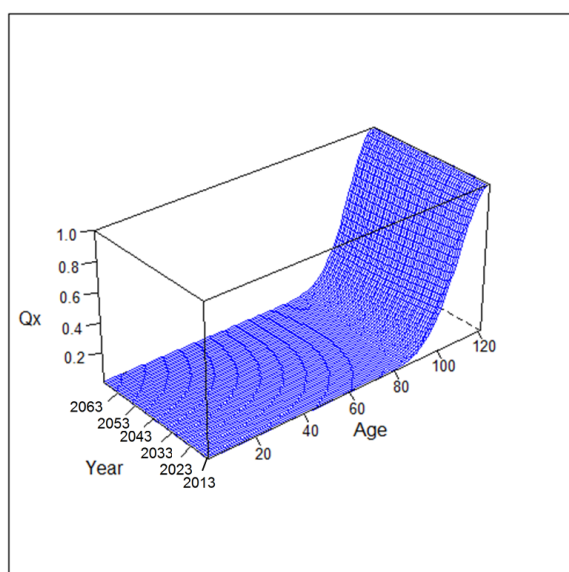


FIGURE M.3 – Table de mortalité ajustée des femmes AD

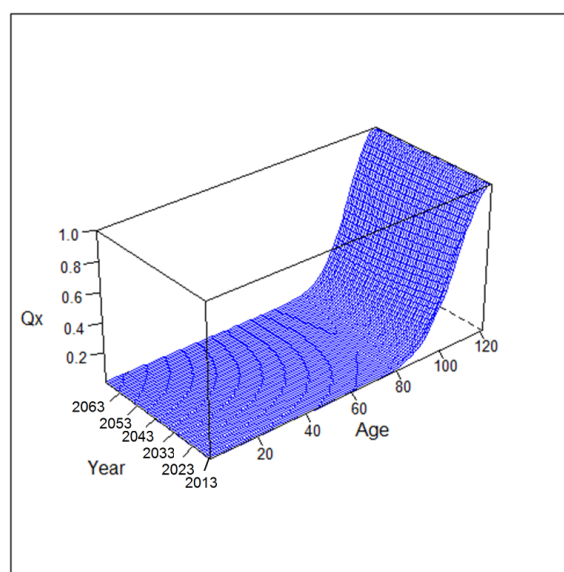


FIGURE M.4 – Table de mortalité ajustée des hommes AD

Bibliographie

- [1] AUBERT P., RABATÉ S. Durée passée en carrière et durée de vie en retraite : quel partage des gains d'espérance de vie? *Économie et statistiques*, (n°474), 2014.
- [2] BÉHAR J. L'ajustement des tables de mortalité des régimes de retraite. *Questions retraite*, (n°67 - Avril 2005), 2005.
- [3] ECARY B. Panorama de la retraite en france. *Support de cours*, 2017-2018. ISFA Lyon.
- [4] HAMISULTANE H. Tests d'autocorrélation. *Archives Ouvertes*, 2002.
- [5] INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES. La mortalité infantile en france. *Tout savoir sur la population*, 2017.
- [6] INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES. L'espérance de vie en france. *Tout savoir sur la population*, 2017.
- [7] LE BRETON M. Pourquoi le nombre de décès en france va augmenter (et pourquoi ce n'est pas arrivé avant). *C'est la vie*, 2016.
- [8] LE FOL S. Réforme des retraites : les principales mesures. *Le Point*, 2015.
- [9] LEGROS F. Les principales mesures de la réforme fillon de 2003. *La retraite en clair*, 2010.
- [10] MEHL S. Régression par la méthode des moindres carrés. *ChronoMath*, 2018.
- [11] MOMAFI P. Mathématiques financières et actuarielles –partie ii : Introduction aux risques d'assurance. *Support de cours*, 2016-2017. EURIA Brest.
- [12] MONTESINOS P. Réflexion sur l'équité intergénérationnelle. *Mémoire présenté devant l'Institut des Actuaire Français*, 2017.
- [13] NIZARD A. La mortalité en france au cours des cinquante dernières années. *Études sur la mort*, (n°121), 2002.
- [14] PLANCHET F. Modèles de durée : Méthodes de lissage et d'ajustement. *Support de cours*, 2017-2018. ISFA Lyon.
- [15] PLANCHET F., TOMAS J. Construction d'une table de mortalité par positionnement : Mode d'emploi. *Note de travail de l'Institut des Actuaire III291-15 v1.3*, 2014.
- [16] PLANCHET F., TOMAS J. Critères de validation : Aspects méthodologiques. *Note de travail de l'Institut des Actuaire III291-15 v1.4*, 2014.
- [17] PLANCHET F., TOMAS J. Méthodes de positionnement : Aspects méthodologiques. *Note de travail de l'Institut des Actuaire III291-15 v1.7*, 2014.
- [18] POIRAUD L. Construction et validation d'une table de mortalité d'expérience par des méthodes non paramétriques. *Mémoire présenté devant l'Institut des Actuaire Français*, 2013.
- [19] RAKOTOMALALA R. Test de normalité. *Support de cours*, 2015. Université Lyon 2 Internet.
- [20] RÉGIME GÉNÉRAL DE LA SÉCURITÉ SOCIALE. Les grandes mesures de la réforme des retraites de 2010 (entrée en vigueur le 1^{er} juillet 2011). *Portail du service public de la Sécurité Sociale*, 2012.
- [21] SECRÉTARIAT GÉNÉRAL DU COR. Espérance de vie et durée de retraite. *Séance plénière du COR*, 2014.
- [22] SILUÉ Z. Elaboration d'une table de mortalité spécifique au régime d'assurance vieillesse de la cnps. *Mémoire Online*, 2007.

BIBLIOGRAPHIE

- [23] VAUDOR L. Régression loess. *R-atique*, 2015.
- [24] VERMET F. Mathématiques des populations. *Support de cours*, 2016-2017. EURIA Brest.
- [25] WANDJI A. Construction d'une table de mortalité prospective sur un portefeuille de taille modeste et historique limité. *Mémoire présenté devant l'Institut des Actuaire Français*, 2016.
- [26] ÉCONOMÉTRIE. Tests d'autocorrélation. *Support de cours*, 2015.