

Mémoire présenté le : 17 mai 2017

**pour l'obtention du Diplôme Universitaire d'actuariat de l'ISFA
et l'admission à l'Institut des Actuaire**

Par : Laureline HERBAUT

Titre : Étude de la mortalité de personnes victimes d'accidents
médicaux

Confidentialité : ☐ NON ☒ OUI (Durée : ☐ 1 an ☒ 2 ans)

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

Membres présents du jury de l'IA

M. Olivier CAYOT

M. Frédéric PLANCHET

Membres présents du jury de l'ISFA

M. Didier RULLIERE

M. Christian ROBERT

Signature

Entreprise

Nom : Sham

Signature :

Directeur de mémoire en entreprise

Nom : Sylvain CARACO

Signature :

Invité

Nom :

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 :

Mme Christine DRIGUZZI

Bibliothèque :

Mme Patricia BARTOLO

INSTITUT DE SCIENCE FINANCIÈRE ET
D'ASSURANCES

MÉMOIRE D'ACTUARIAT

**Étude de la mortalité des personnes
victimes d'accidents médicaux**

Auteur :
Laureline HERBAUT

Tuteurs :
Sylvain CARACO
Alexis BIENVENÜE

*Mémoire présenté dans le cadre de l'obtention du statut d'actuaire
proposé par*

Institut des Actuaire

2016 – 2017

Résumé

Dans ce mémoire, nous nous intéressons au provisionnement des rentes d'invalidité de personnes victimes d'accidents médicaux et présentes dans le portefeuille de Sham, société d'assurance mutuelle spécialisée dans l'assurance et le management des risques des acteurs de la santé, du social et du médico-social.

En effet, la question se pose de savoir si la TD 88-90, table de mortalité imposée par le Code des Assurances pour provisionner le risque de survie en invalidité et homologuée en 1993, est toujours prudente à ce jour.

Du fait du faible nombre de rentiers dans le portefeuille, des tables de mortalité sont construites à partir d'une base de données comportant la totalité des victimes d'accidents médicaux, et qui est ensuite réduite en fonction de certains critères afin de se rapprocher au mieux de la mortalité des rentiers.

Nous verrons que plus nous nous rapprochons de la mortalité des rentiers, plus la table réglementaire TD 88-90 semble adaptée aux calculs de provisions mathématiques, mais le nombre de données utilisées devient également plus faible, ce qui augmente l'incertitude autour de la table construite.

Il serait intéressant de poursuivre cette étude dans les prochaines années, pour avoir davantage de données exploitables, afin de dégager une tendance sur plusieurs années.

Mots-clés : Mortalité, Accidents médicaux, Kaplan-Meier, Modèle de Brass, Régression exponentielle

Abstract

In this thesis, the issue of the reserving of invalidity annuities for victims of medical accidents is tackled. These subjects are present in Sham's portfolio, a mutual insurance company specialised in the insurance and management of risks for health and social care providers.

The question arises indeed to know whether the TD 88-90 is still prudent today. This mortality table is imposed by the French Insurance Code for the reserving of the survival risk in invalidity. It was however certified in 1993.

Because of the small number of annuitants in the portfolio, mortality tables are constructed from a database of all the medical accident victims. This database is then reduced according to certain criteria, in order to get closer to the mortality of the annuitants.

The more the mortality of the annuitants is approached, the more adapted to the mathematical reserves calculations the TD 88-90 seems to be. However, the number of data used gets also smaller, which increases the uncertainty around the constructed table.

It would be interesting to pursue this study in the following years, in order to have more working data, so as to reveal a trend over several years.

Keywords : Mortality, Medical accidents, Kaplan-Meier, Brass' relational model, Exponential regression

Table des matières

Résumé	iii
Abstract	v
Introduction	1
1 Sham	3
1.1 Présentation	3
1.2 Historique	3
1.3 La Responsabilité Civile Médicale	4
1.4 Les modes de gestion des sinistres en RC	4
2 Analyse de la base de données	7
2.1 Objectifs	7
2.2 Présentation de la base de données	8
2.3 Présentation des clients	9
2.4 Étude descriptive des victimes	9
2.4.1 Population totale	9
2.4.2 Population en rente	11
2.4.3 Conclusion	13
3 Construction de la table de mortalité	15
3.1 Notations actuarielles	15
3.2 Taux bruts de Hoem	16
3.2.1 Définition	16
3.2.2 Application aux données	16
3.3 Critère de Cochran	18
3.4 Évolution de la mortalité sur la période d'étude	20
3.4.1 Tests de comparaison des courbes de survie	21
3.4.2 Application aux données	23
3.5 Traitement des âges intermédiaires	26
3.5.1 Estimateur de Kaplan-Meier de la fonction de survie	26
3.5.2 Ajustement de Makeham	28
3.5.3 Lissage de Whittaker-Henderson	31
3.5.4 Validation des ajustements et choix de la méthode	33
3.6 Traitement des âges jeunes (avant 49 ans)	37
3.6.1 Abattement par le ratio SMR	37
3.6.2 Modèle de Brass	41
3.6.3 Validation des ajustements	45
3.7 Traitement des âges élevés	46
3.7.1 Fermeture de Coale & Kisker	46
3.7.2 Fermeture par régression exponentielle	48
3.7.3 Validation des fermetures et choix de la méthode	50

3.8 Conclusion	51
4 Approximation de la mortalité des rentiers	55
4.1 Mortalité par coûts	55
4.1.1 Dossiers avec un coût total supérieur à 1 000 €	56
4.1.2 Dossiers avec un coût total supérieur à 10 000 €	58
4.1.3 Dossiers avec un coût total supérieur à 100 000 €	60
4.2 Mortalité par IPP	63
4.2.1 Taux d'IPP supérieur à 8	63
4.2.2 Taux d'IPP supérieur à 20	65
4.3 Mortalité par spécialité mise en cause	67
4.4 Mortalité des rentiers	67
4.5 Sensibilité de la mortalité aux hypothèses et aux modèles	69
4.5.1 Sensibilité aux hypothèses	69
4.5.2 Sensibilité aux modèles	71
4.6 Provisions mathématiques pour les rentes	73
Conclusion	75
Bibliographie	77
A Décès et exposition pour les principales spécialités mises en cause	79
A.1 Accidents en neurochirurgie	79
A.2 Accidents en obstétrique	80
A.3 Accidents en orthopédie	80
A.4 Accidents en chirurgie viscérale	81
A.5 Accidents aux urgences	81
A.6 Accidents en pédiatrie et néonatalogie	81

Table des figures

1.1 Répartition des réclamations reçues en 2014	4
2.1 Répartition des sinistres par spécialités mises en cause	10
2.2 Répartition des sinistres par causes médicales	11
2.3 Répartition des sinistres par spécialités mises en cause	12
2.4 Répartition des sinistres par causes médicales	13
3.1 Données d'exposition par âge	17
3.2 Taux bruts de Hoem par sexe	18
3.3 Décès lissés par âge	19
3.4 Taux bruts de Hoem par sexe	20
3.5 Courbes de hasard pour deux sous-populations indépendantes de la population féminine	24
3.6 Courbes de hasard pour deux sous-populations indépendantes de la population masculine	25
3.7 Estimateur de Kaplan-Meier des taux bruts de décès par sexe	27
3.8 Approximation par sexe des points $\ln(q_{x+1} - q_x)$ par une droite	29
3.9 Ajustement de Makeham par sexe	30
3.10 Lissage de Whittaker-Henderson par sexe	32
3.11 Comparaison des décès observés et des décès théoriques	33
3.12 Comparaison des taux lissés par Whittaker-Henderson avec les taux de Hoem	36
3.13 Taux abattus par le ratio SMR pour la population féminine	38
3.14 Taux abattus par le ratio SMR et lissés pour la population féminine	39
3.15 Taux abattus par le ratio SMR pour la population masculine	40
3.16 Taux abattus par le ratio SMR et lissés pour la population masculine	40
3.17 Taux ajustés par le modèle de Brass pour la population féminine	43
3.18 Taux ajustés par le modèle de Brass et lissés pour la population féminine	43
3.19 Taux ajustés par le modèle de Brass pour la population masculine	44
3.20 Taux ajustés par le modèle de Brass et lissés pour la population masculine	45
3.21 Comparaison des taux ajustés pour les âges inférieurs à 48 ans	45
3.22 Tables obtenues par sexe pour les âges inférieurs à 89-87 ans	46
3.23 Évolution par sexe des taux de croissance de mortalité	47
3.24 Tables extrapolées par sexe par la méthode de Coale & Kisker	49
3.25 Table féminine extrapolée par régression exponentielle	49
3.26 Table masculine extrapolée par régression exponentielle	50
3.27 Comparaison des tables construites par sexe	51
3.28 Table mixte construite pour la population totale	53
4.1 Étude de la mortalité par coûts - Exposition et nombre de décès	56
4.2 Dossiers ayant un coût supérieur à 1 000 €- Exposition et nombre de décès	57

4.3	Dossiers ayant un coût supérieur à 1 000 €- Comparaison des méthodes	58
4.4	Dossiers ayant un coût supérieur à 10 000 €- Exposition et nombre de décès	59
4.5	Dossiers ayant un coût supérieur à 10 000 €- Comparaison des méthodes	60
4.6	Dossiers ayant un coût supérieur à 10 000 €- Comparaison avec les tables réglementaires	60
4.7	Dossiers ayant un coût supérieur à 100 000 €- Exposition et nombre de décès	61
4.8	Dossiers ayant un coût supérieur à 100 000 €- Taux de Hoem	61
4.9	Dossiers ayant un coût supérieur à 1 000 €- Comparaison avec les tables réglementaires	62
4.10	Taux d'IPP supérieur à 8 - Exposition et nombre de décès	63
4.11	Taux d'IPP supérieur à 8 - Comparaison des modèles	64
4.12	Taux d'IPP supérieur à 20 - Exposition et nombre de décès	65
4.13	Taux d'IPP supérieur à 20 - Taux de Hoem	66
4.14	Taux d'IPP supérieur à 20 - Comparaison avec les tables réglementaires	66
4.15	Rentiers - Exposition et nombre de décès	68
4.16	Rentiers - Comparaison avec les tables réglementaires	69
4.17	Sensibilité des ajustements aux hypothèses	71
4.18	Sensibilité des ajustements aux modèles	72
4.19	Répartition des provisions mathématiques	74
A.1	Accidents en neurochirurgie - Exposition et nombre de décès	79
A.2	Accidents en obstétrique - Exposition et nombre de décès	80
A.3	Accidents en orthopédie - Exposition et nombre de décès	80
A.4	Accidents en chirurgie viscérale - Exposition et nombre de décès . . .	81
A.5	Accidents aux urgences - Exposition et nombre de décès	82
A.6	Accidents en pédiatrie et néonatalogie - Exposition et nombre de décès	82

Liste des tableaux

2.1	Répartition des différentes variables décrivant les victimes	10
2.2	Répartition des différentes variables décrivant les victimes	12
2.3	Comparaison des âges des décès parmi la population totale et les rentiers	13
3.1	Ratios SMR pour la population féminine	38
3.2	Ratios SMR pour la population masculine	39
3.3	Valeurs des paramètres du modèle de Brass pour la population féminine	42
3.4	Valeurs des paramètres du modèle de Brass pour la population masculine	44
3.5	Étapes de la construction de la table de mortalité mixte	52
4.1	Étude de la mortalité par coûts - Proportion conservée pour la population, les décès et les rentiers	56
4.2	Étapes de la construction de la table de mortalité pour les dossiers ayant un coût supérieur à 1 000€	58
4.3	Étapes de la construction de la table de mortalité pour les dossiers ayant un coût supérieur à 10 000€	59
4.4	Taux d'IPP supérieur à 8 - Proportion conservée pour la population, les décès et les rentiers	63
4.5	Étapes de la construction de la table de mortalité pour les personnes ayant un taux d'IPP supérieur à 8	64
4.6	Taux d'IPP supérieur à 20 - Proportion conservée pour la population, les décès et les rentiers	65
4.7	Étude de la mortalité par spécialité mise en cause - Proportion conservée pour la population, les décès et les rentiers	67
4.8	Ratios SMR de la population des rentiers selon différentes tables de référence	69
4.9	Ratios SMR des sous-populations étudiées en prenant pour référence la population totale mixte	70
4.10	Paramètres du modèle de Brass des sous-populations étudiées en prenant pour référence la population totale mixte	70
4.11	Provisions mathématiques au 31/12/2015 des rentes selon les tables de mortalité construites (en base 100)	73

Introduction

Le risque le plus important pour Sham, société d'assurance pour les établissements hospitaliers, est la Responsabilité Civile médicale. Cette garantie permet de protéger les professionnels de santé et d'indemniser les éventuelles victimes d'accidents médicaux au sein des structures assurées.

Les accidents les plus graves peuvent causer l'invalidité de la victime, qui devient alors bénéficiaire d'une rente d'invalidité. Le Code des Assurances impose aux assureurs de provisionner le risque de survie en invalidité avec la table de mortalité réglementaire TD 88-90. Cependant, cette table a été homologuée en 1993, et il devient normal de se demander si cette table est toujours représentative, ou du moins suffisamment prudente, du risque de survie des rentiers en invalidité.

L'objectif de cette étude est donc d'estimer la mortalité réelle du portefeuille des rentiers de Sham et de la positionner par rapport à la table TD 88-90, et par rapport aux tables plus récentes : les tables générationnelles TGF05 et TGH05 utilisées pour la tarification et le provisionnement des rentes viagères.

Du fait du très faible nombre de rentiers dans le portefeuille de Sham (moins de 500), nous allons dans un premier temps estimer la mortalité de l'ensemble des victimes d'accidents médicaux, puis tenter de restreindre l'étude à une population plus petite en nombre, mais plus représentative de la mortalité des rentiers.

Après une brève présentation de Sham et de la garantie en Responsabilité Civile médicale, nous allons réaliser une analyse de la base de données afin de mieux appréhender la population totale et les caractéristiques des rentiers.

Nous construisons ensuite la table de mortalité par sexe, puis mixte, sur l'ensemble de la population disponible, et nous concluons quant à sa position par rapport aux tables TD 88-90 et TGF05/TGH05.

Par la suite, nous utiliserons la partie d'analyse de la base de données afin de restreindre l'étude aux victimes proches de la population des rentiers. Les tables de mortalité construites seront alors comparées à la table mixte construite plus tôt, ainsi qu'aux tables réglementaires.

Nous concluons finalement sur la mortalité des rentiers de Sham et la prudence de la TD 88-90.

Chapitre 1

Sham

1.1 Présentation

Fondée en 1927 par des directeurs d'hôpitaux, la Société Hospitalière d'Assurances Mutuelles (Sham) est une société d'assurance mutuelle spécialisée dans l'assurance et le management des risques des acteurs de la santé, du social et du médico-social. Elle est aujourd'hui leader de la Responsabilité Civile médicale en France et un expert de l'environnement de la santé, du social et du médico-social.

C'est une société de personnes à statut mutualiste et qui vise la satisfaction de ses sociétaires.

À la fin de l'année 2015, Sham disposait d'un chiffre d'affaires de 306.3 M€ et d'un portefeuille clients de près de 10 000 sociétaires. Elle emploie aujourd'hui environ 350 collaborateurs.

1.2 Historique

Forte de son expérience en matière d'évaluation, de prévention, d'assurance et de gestion des risques, Sham a progressivement étendu son domaine d'activité de l'assurance du personnel hospitalier à celle des organismes médicaux, établissements d'hébergement, cliniques privées, structures du soin et du maintien à domicile. Elle a également élargi son offre aux professionnels de santé libéraux, aux établissements privés et aux établissements sociaux et médico-sociaux.

En 2013, Sham fait l'acquisition du courtier Sofaxis et forment ensemble le Groupe Sham. Créé 28 ans plus tôt, Sofaxis est l'un des principaux courtiers français en assurance de personnes auprès des collectivités territoriales et des établissements hospitaliers.

Ils ont pour ambition de se positionner comme un acteur multi spécialiste de référence en matière d'assurance, de gestion des risques et de protection sociale, au service du monde territorial et de l'ensemble des acteurs de la filière de soins à la personne.

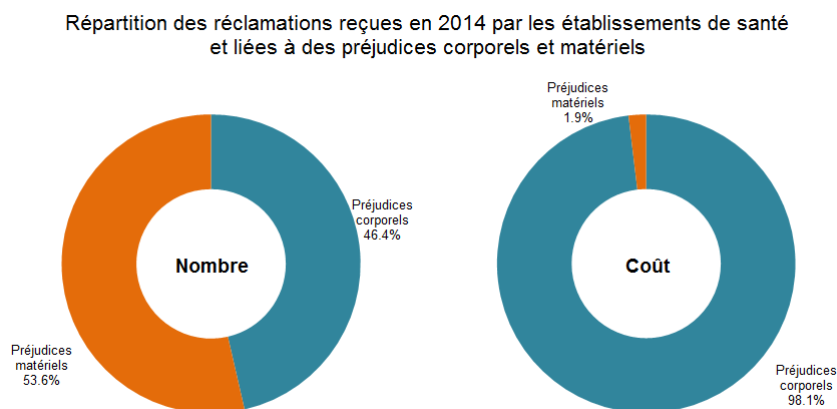
Sham a également pour objectif la diversification géographique de ses activités d'assurance responsabilité civile médicale. Après avoir entamé sa diversification géographique en 2014 avec l'ouverture d'un bureau de distribution et de gestion en Espagne, à Madrid, la société a ouvert fin 2015 un bureau en Italie, à Rome, et fait l'acquisition d'un acteur majeur de la gestion de sinistres en Italie, basé à Turin.

1.3 La Responsabilité Civile Médicale

Les contrats proposés par Sham en Responsabilité Civile peuvent être des contrats de Responsabilité Civile Médicale, Professionnelle ou des Dirigeants. Ces contrats permettent de se prémunir contre les conséquences pécuniaires des risques que le sociétaire peut encourir à l'égard des tiers à l'occasion de l'exercice de son activité.

La Responsabilité Civile médicale est la principale activité de Sham, et représentait 68.5% de son chiffre d'affaires en 2015.

Au sein des différentes garanties proposées, on peut distinguer notamment les garanties concernant des accidents corporels ou matériels. Bien que les sinistres matériels soient les plus nombreux, ce sont bien les accidents corporels qui coûtent le plus à Sham (1.1).



GRAPHIQUE 1.1 – Répartition des réclamations reçues en 2014

1.4 Les modes de gestion des sinistres en RC

Lorsqu'un patient fait une réclamation suite à une prise en charge médicale, le client se tourne alors vers Sham, qui étudie si la RC du client est engagée.

À l'ouverture du dossier, les gestionnaires étudient les garanties (type de contrat RC, classification en fonction de la spécialité mise en cause lors du sinistre et de la cause médicale) et évaluent le dossier. Par la suite, ils pilotent la stratégie de défense et l'analyse du dossier en collaboration avec les différents interlocuteurs du dossier, et revoient régulièrement l'évaluation du dossier.

Pour savoir si la RC du client est engagée, l'article L1142-1 du Code de la Santé Publique stipule que : « hors les cas où leur responsabilité est encourue en raison d'un défaut de produit de santé de santé, les professionnels de santé (...) ne sont responsables des conséquences dommageables d'acte de prévention, de diagnostic ou de soins qu'en cas de faute ». Il faut donc qu'il y ait une faute, un préjudice et un lien de causalité direct et certain entre les deux.

Un dossier fait donc l'objet d'une analyse médico-légale et d'une analyse juridique plus ou moins poussées selon les dossiers. Celles-ci aident à revoir l'évaluation du dossier et rentrent en compte dans le provisionnement.

Du point de vue technique, les sinistres étaient auparavant gérés sous l'AS400. Suite à un changement de logiciel de gestion entre 2006 et 2008, les dossiers clos depuis un certain temps n'ont pas été migrés. Il manque donc probablement de nombreux dossiers sur l'historique ancien dans le nouveau logiciel, et l'étude de la mortalité ne devra pas s'étendre sur une période antérieure aux années 2000 pour éviter ce manque de données.

Chapitre 2

Analyse de la base de données

2.1 Objectifs

L'article A331-10 du Code des Assurances stipule que « le calcul des provisions techniques de prestations d'incapacité de travail et d'invalidité est effectué à partir des éléments suivants :

1° Une loi de survie en invalidité définie par la table TD 88-90 homologuée par l'arrêté du 27 avril 1993 réactualisant les tables de mortalité (...) » [8]

Sham utilise donc la table TD 88-90 pour le calcul de provisions pour les rentiers victimes d'accidents médicaux présents en portefeuille.

La TD 88-90 est une table du moment construite par l'Insee à partir d'observations réalisées entre 1988 et 1990 sur les hommes. Elle décrit la mortalité d'une génération fictive à une période précise dans le temps, ici 1988-1990.

Les tables recommandées pour la tarification et le provisionnement des contrats de rentes viagères sont depuis le 1er janvier 2007 les tables TGH05 et TGF05. Il s'agit de tables générationnelles, créées soit sur la population masculine (Table Générationnelle pour les Hommes), soit sur la population féminine (Table Générationnelle pour les Femmes), et à l'aide de méthodes prospectives. Elles décrivent les niveaux de mortalité pour chaque génération, et sont donc plus proches de la réalité. Suite à une polémique sur le principe d'égalité hommes/femmes, le calcul des primes et des prestations pour les contrats de rente souscrits après le 20 décembre 2012 doit se faire à partir de la TGF05 uniquement, et plus la TGH05. Cependant, il est toujours possible d'utiliser l'une ou l'autre des deux tables dans le calcul de provisions.

Du fait de l'ancienneté de la table TD 88-90, l'ACPR demande de vérifier si cette table décrit correctement la mortalité historique des rentiers dans le portefeuille de Sham, et si les provisions calculées avec cette table sont suffisamment prudentes. Dans le cas contraire, les nouvelles provisions devront être calculées à partir des tables plus récentes TGH05 et TGF05.

L'objectif de ce travail est donc d'étudier la mortalité des personnes victimes d'accidents en RC médicale et qui perçoivent aujourd'hui des rentes. Du fait du très faible nombre de rentes en cours (environ 400) et du nombre total de rentes présentes dans les bases de données, l'étude est élargie dans un premier temps à l'ensemble des victimes d'accidents médicaux.

Il faut donc pour cela étudier toutes les données disponibles pour cette garantie. Seuls les sinistres corporels seront pris en compte, puisque c'est sur ce type de sinistres que porte le risque. Ils constituent un tiers de l'ensemble des sinistres déclarés en RC sur la totalité du portefeuille, les autres étant matériels.

Une fois la mortalité de cette population établie, il sera possible de réduire la population étudiée aux personnes ayant des dossiers significatifs (personnes victimes d'accidents à la naissance, ou dossiers supérieurs à un certain montant par exemple) afin de supprimer les risques mineurs qui réduisent la mortalité de la population globale.

Il sera ensuite possible de fournir l'écart entre les tables retenues dans le modèle de calcul de provisions et la mortalité historique observée.

2.2 Présentation de la base de données

La base de données construite contient dans un premier temps l'ensemble des dossiers de réclamations en RC médicale présents dans les bases de sinistres de Sham, avec notamment les différents coûts des sinistres. Les montants de règlements, provisions et coûts totaux (qui sont la somme des règlements et des provisions pour chaque dossier) sont gardés en vision au 31/12/2015.

Pour chaque sinistre, les informations de la victime ont ensuite été rajoutées, puis la base est retraitée pour avoir une vision par victime et non par sinistre. Pour chaque victime, on dispose donc des informations suivantes :

- Un identifiant unique
- Date de naissance
- Date de décès (si renseignée)
- Date d'entrée en observation
- Date de sortie d'observation
- Le montant des règlements / provisions / coûts totaux, correspondant à la somme des coûts de chaque sinistre déclaré pour cette victime
- Taux d'IPP (si renseigné)

Cette base a été retraitée des anomalies suivantes :

- Date de naissance non renseignée
- Naissance au XVII^e siècle ou après 2016
- Décès après la date du jour
- Décès avant naissance
- Dates de naissance ou de décès différentes pour une même personne (s'il existe un écart significatif de plus d'un mois entre les dates)
- Décès avant la date des faits

Par ailleurs, pour l'étude de la mortalité, il faut que les dossiers soient ouverts avant la date de décès : cela supprime plus de 7 000 dossiers sur une base d'environ 77 000 dossiers.

La clôture puis la réouverture d'un dossier a également été prise en compte en créant dans ce cas deux observations différentes. La date d'entrée en observation considérée dans l'étude de la mortalité est en effet la date d'ouverture du contrat.

2.3 Présentation des clients

La base contient des dossiers ouverts entre 1955 et 2015, toutefois les premiers décès ont lieu à partir de 1996, et ces décès deviennent conséquents à partir de 2004. Cela vient du fait du changement de logiciel évoqué précédemment, ayant laissé les dossiers clos dans l'AS400 lors de la migration. On décide donc de réduire l'étude aux personnes présentes entre le 01/01/2004 et le 21/12/2015.

Sur cette période, le nombre de clients dont les patients ont déclaré des accidents en RC médicale varie autour de 550 par an, avec un nombre plus important entre 2010 et 2012 (autour de 700). Cela représente un total de 1 795 établissements distincts, toutes années confondues, la majorité étant des établissements publics hospitaliers.

Ces déclarations ont principalement lieu chez des établissements de médecine et chirurgie-obstétrique (MCO). Ce secteur correspond à des séjours hospitaliers de courte durée, et prend donc en charge les hospitalisations, les interventions chirurgicales, ainsi que les grossesses et accouchements.

La base contient un nombre de données très faible pour une étude de la mortalité. Cependant, il faut savoir que Sham assure un établissement hospitalier sur deux en France en RC médicale, ainsi que 2 naissances sur 3. Cela représente donc une part importante du risque de la France en RC médicale. Pour obtenir plus de données sur ce risque, il faudrait par exemple obtenir des données au niveau européen.

2.4 Étude descriptive des victimes

2.4.1 Population totale

L'étude concerne 52 273 victimes présentes entre 2004 et 2015, correspondant à 57 599 observations dans la base de données. Plus de 3 200 dossiers ont effectivement été clos puis rouverts au moins une fois. Certaines personnes ont par ailleurs été victimes de plusieurs sinistres : 1 916 individus, dont plus de 80% mettant en cause des pathologies différentes à chaque sinistre.

Parmi ces victimes, on dénombre 757 décès. Ceux-ci ont lieu entre 0 et 102 ans, mais il y a certains âges manquants, notamment avant 30 ans ou après 95 ans, où il peut arriver qu'il n'y ait aucun décès pour quelques âges.

Enfin, le Sex Ratio (correspondant à la proportion d'hommes dans la population) est de 47.1%. Il augmente jusqu'à 55.7% parmi la population décédée.

Pour l'ensemble des victimes présentes sur la période d'étude, les sinistres ont coûté plus de 2 milliards d'euros à Sham, en regardant l'évolution des coûts jusqu'au 31/12/2015, dont 65% sont des provisions à cette date. Le coût moyen total par victime s'élève à 38 961€, mais le coût médian total n'est que de 2 336€. En effet, seulement un nombre réduit de dossiers, 7% des dossiers totaux, représente à la fois 82% des coûts totaux, 97% des rentiers et 26% du nombre de décès totaux.

Cela montre bien que seule une faible proportion de la population est représentative de la mortalité, et c'est dans cette proportion que sont a priori présents les dossiers en rente.

En regardant les répartitions de certaines variables décrivant les victimes (Table 2.1), on remarque que les âges sont légèrement concentrés autour des 50 ans.

Pour seulement la moitié des personnes, le taux d'IPP est renseigné. Cette variable est donc pour l'instant laissée de côté dans l'étude de la mortalité, mais elle est toujours intéressante pour connaître le portefeuille. Lorsqu'ils sont renseignés, les taux d'IPP sont alors en revanche plutôt faibles.

	Min	1 ^{er} quartile	Médiane	Moyenne	3 ^{ème} quartile	Max
Âge d'entrée	0	35	50	48	63	108
Âge de sortie	0	38	53	51	66	116
IPP renseignés	0	5	10	16	17	100

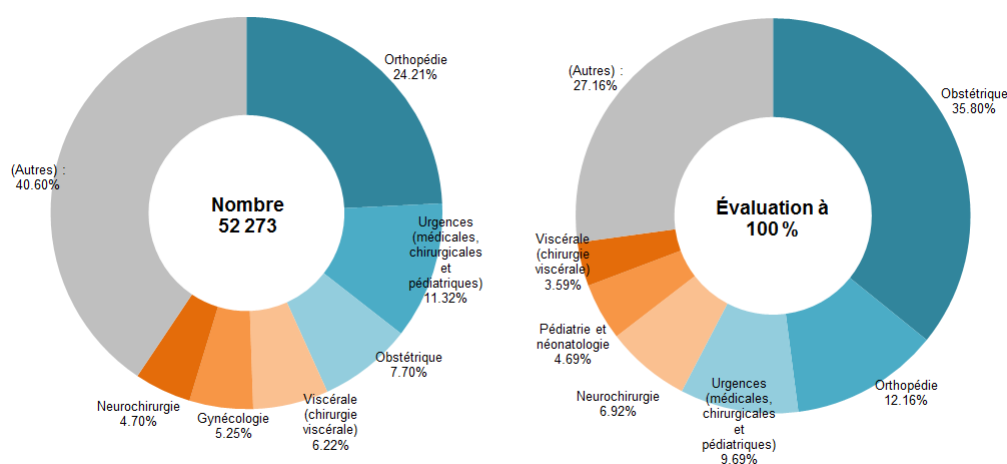
TABLE 2.1 – Répartition des différentes variables décrivant les victimes

On s'intéresse maintenant à la répartition des sinistres par spécialités mises en cause et par causes médicales. Celle-ci est présentée ci-dessous dans deux séries de graphiques, qui ne reprennent que les catégories les plus importantes afin de ne pas les surcharger.

Plus de 65% des sinistres (Graphique 2.1) mettent en cause comme spécialité les services de chirurgie (comprenant notamment l'obstétrique, l'orthopédie, la neurochirurgie et la gynécologie), et 13% mettent en cause les services d'urgences et de réanimation.

Bien que près d'un quart des sinistres ont pour spécialité mise en cause l'orthopédie, c'est l'obstétrique qui représente plus d'un tiers des coûts.

En effet, l'obstétrique est une spécialité qui prend en charge les grossesses et les accouchements. Les sinistres qui surviennent dans cette spécialité peuvent donc être très importants en termes de gravité.

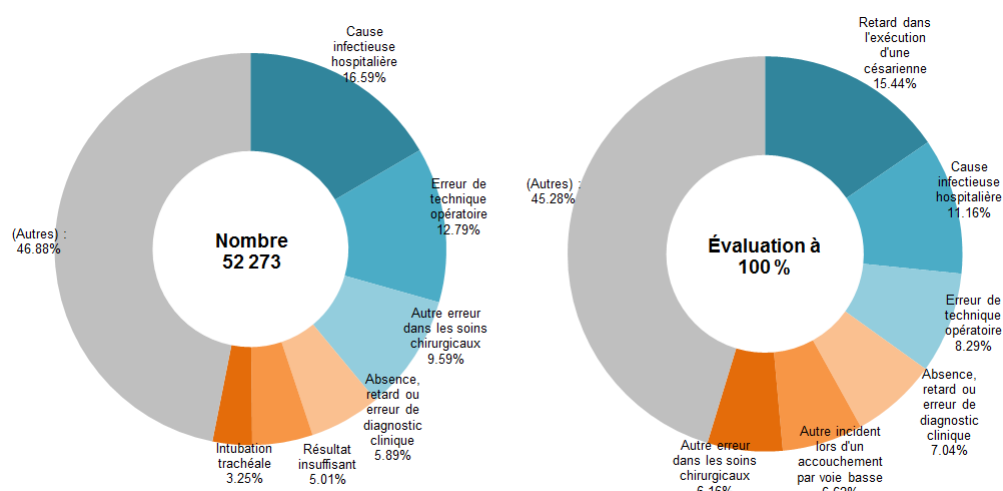


GRAPHIQUE 2.1 – Répartition des sinistres par spécialités mises en cause

Près de la moitié des causes médicales entrant en jeu dans un sinistre (Graphique 2.2) sont des erreurs dans les soins. Viennent ensuite les infections (18%) et les erreurs de diagnostic (11%).

En revanche, en termes de montants, les erreurs dans les soins ne représentent plus qu'un tiers des coûts globaux, autant que la procréation (qui comprend les retards dans l'exécution d'une césarienne et les autres incidents lors d'un accouchement par voie basse, mais également d'autres causes comme l'absence de diagnostic d'une malformation, etc.)

Bien que la procréation représente moins de 5% des causes médicales en nombre de dossiers, elle représente près d'un tiers des coûts pour Sham, tout comme les causes liées à des erreurs dans les soins.



GRAPHIQUE 2.2 – Répartition des sinistres par causes médicales

2.4.2 Population en rente

L'objectif de ce travail est d'étudier la mortalité des rentiers assurés par Sham. Il est donc nécessaire de positionner les rentiers par rapport à la population totale, afin de savoir quels critères utiliser plus tard pour réduire la population étudiée et mieux cerner la mortalité des rentiers.

Il y a dans la base de données seulement 474 dossiers évalués en rente, avec un Sex Ratio de 51.5%. Ce rapport est plus élevé que celui de la population totale. Il augmente jusqu'à 63.1% en ce qui concerne les décès. Près de deux tiers des décès parmi les rentiers sont donc des hommes, et il y eu 38 décès parmi les rentiers à ce jour.

Les coûts totaux pour Sham des dossiers évalués en rente représentent la moitié (49%) des coûts pour l'ensemble de la RC médicale, et sont constitués à 79% de provisions. Par ailleurs, 59% des provisions totales pour la RC médicale correspondent à des provisions pour les rentiers.

Enfin, les coûts des rentiers présentent une variance beaucoup plus faible, et le coût moyen total par victime est d'environ 2 millions d'euros pour un coût médian d'environ 1,1 million d'euros.

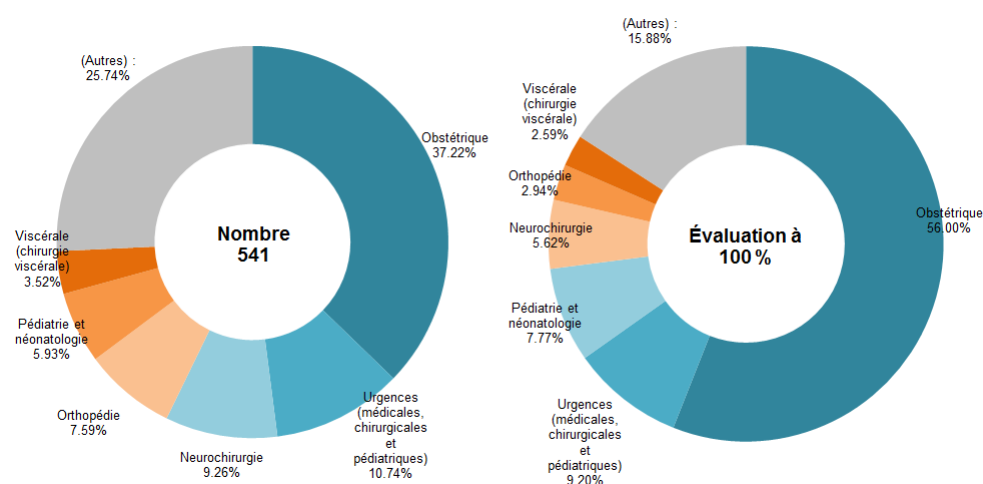
Les dossiers évalués en rente représentent donc une part importante des provisions et des coûts pour Sham.

Concernant la répartition des âges des rentiers (Table 2.2), la population des rentiers est globalement très jeune, et plus jeune que la population totale (majoritairement moins de 50 ans). Par ailleurs, ils ont des taux d'IPP assez élevé (plus de 60% pour la plupart, lorsqu'il est renseigné).

	Min	1 ^{er} quartile	Médiane	Moyenne	3 ^{ème} quartile	Max
Âge d'entrée	0	6	20	27	46	83
Âge de sortie	2	14	29	33	51	92
IPP renseignés	8	60	80	71	85	100

TABLE 2.2 – Répartition des différentes variables décrivant les victimes

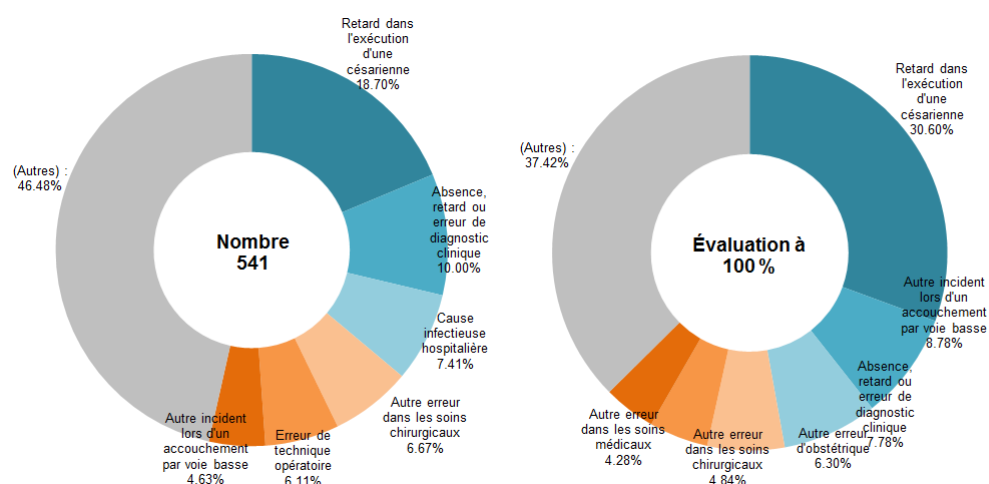
Dans la répartition par spécialité mise en cause lors du sinistre (Graphique 2.3), on observe parmi les rentiers un bond de l'obstétrique, à la fois en nombre de dossiers et en coût pour Sham. Cette spécialité représente en effet plus d'un tiers des dossiers évalués en rente, et plus de la moitié des coûts pour Sham. Les services de chirurgie représentent environ 70% des spécialités mises en cause lors des sinistres déclarés en RC médicale, à la fois en répartition en nombre et en coût. C'est donc une spécialité importante dans les dossiers des rentiers.



GRAPHIQUE 2.3 – Répartition des sinistres par spécialités mises en cause

Concernant les causes médicales des sinistres (Graphique 2.4), la procréation prend également une part très importante. Elle représente désormais un tiers du nombre de dossiers, alors qu'elle ne représente que 5% des causes de sinistres pour la population totale.

En termes de coût pour Sham, près de la moitié des causes médicales appartiennent à la catégorie procréation, alors que celle-ci représente moins d'un tiers des causes médicales pour la population totale.



GRAPHIQUE 2.4 – Répartition des sinistres par causes médicales

2.4.3 Conclusion

La population des rentiers de Sham est une population assez jeune. Les accidents médicaux qu'ils ont subis ont en effet régulièrement lieu dans des services tels que l'obstétrique, la gynécologie, la pédiatrie ou la néonatalogie. Cependant, les accidents résultant en des évaluations en rente sont peu nombreux, et la population globale ne présente pas la même distribution des âges que les rentiers.

En revanche, la population des rentiers présente des décès répartis à tous les âges, alors que les décès parmi la population globale ont généralement lieu après 50 ans (Table 2.3).

	Min	1 ^{er} quartile	Médiane	Moyenne	3 ^{ème} quartile	Max
Âge de décès (population totale)	0	53	65	62	77	102
Âge de décès (rentiers)	8	26	55	50	65	92

TABLE 2.3 – Comparaison des âges des décès parmi la population totale et les rentiers

Cela complique l'étude de la mortalité des rentiers : ceux-ci n'ont pas du tout la même répartition d'exposition et de décès (Annexe A) que la population totale disponible. Cependant, l'étude que nous allons pouvoir réaliser ici ne peut a priori que sous-estimer la mortalité des rentiers, puisque la population totale présente 1.45% de décès contre 8.02% parmi les rentiers. Cela représente un comportement prudent dans l'étude du risque de survie mais entraînera une forte sur-évaluation des provisions. Par ailleurs, il n'est pas possible d'estimer directement la mortalité des rentiers, et le seul moyen que nous avons à notre disposition est de s'en approcher autant que possible par l'étude d'une population plus nombreuse.

Chapitre 3

Construction de la table de mortalité

Nous allons dans un premier temps étudier la mortalité de la population totale, car le nombre de données pour réaliser cette étude est déjà faible. Cela permettra par la suite d'éventuellement positionner une part restreinte de la population par rapport à cette table afin de mieux approximer la mortalité des rentiers.

L'étude présentée ici s'inspire des travaux de Planchet et al. [6], Reynaud [7] et Zhang [9].

3.1 Notations actuarielles

Afin d'étudier la mortalité des victimes d'accidents médicaux, il est avant tout nécessaire d'introduire certaines notations.

Si T est une variable aléatoire à valeurs dans $[0, +\infty[$ décrivant la durée de vie d'un individu, on note alors dans la suite $S(t) = \mathbb{P}(T > t)$ la fonction de survie. Elle décrit la probabilité qu'un individu survive jusqu'à l'âge t . C'est donc une fonction décroissante telle que $S(0) = 1$ si $P(T = 0) = 0$ et telle que $\lim_{t \rightarrow +\infty} S(t) = 0$.

On définit la fonction de survie conditionnelle $S_x(t)$ comme étant la probabilité pour l'individu de survivre encore t années sachant qu'il est en vie à l'âge x : $S_x(t) = \mathbb{P}(T > t + x | T > x)$.

Comme la densité de la variable est définie par $f(t) = -\frac{dS(t)}{dt}$, on peut définir la fonction de hasard par $h(t) = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)} = -\frac{d}{dt} \ln S(t)$. Elle représente le risque instantané, c'est-à-dire la probabilité de mourir dans un petit intervalle de temps dt après t conditionnellement à la survie jusqu'en t .

On peut donc en déduire une nouvelle expression de la fonction de survie : $S(t) = \exp\left(-\int_0^t h(u)du\right)$.

Enfin, le taux de décès q_x est la probabilité pour un individu de décéder à l'âge x :

$$q_x = P(T < x + 1 | T > x) = 1 - \frac{S(x+1)}{S(x)} = 1 - S_x(1)$$

C'est donc la quantité que nous allons essayer d'estimer pour notre population.

Cependant, notre base de données contient certaines données incomplètes. Par exemple, certaines personnes sortent d'observation car leur dossier est clos. On ne

sait alors pas ce qui leur arrive par la suite, à savoir si elles meurent et à quel âge. On sait juste que leur durée de vie est supérieure à leur âge au moment de la sortie. On dit qu'elles sont censurées à droite.

Par ailleurs, les personnes qui sont décédées avant d'être entrées en observation ne sont pas prises en compte dans l'étude. On dit alors qu'il y a troncature à gauche.

Il faudra donc intégrer ces informations dans l'étude.

3.2 Taux bruts de Hoem

3.2.1 Définition

Un estimateur sans biais, convergent et asymptotiquement normal des taux bruts de mortalité est l'estimateur de Hoem. À un âge x donné, notons D_x le nombre de décès et E_x l'exposition totale, c'est-à-dire la somme des durées de présence (exprimées en fraction d'année) de chaque individu à l'âge x . L'estimateur de Hoem est alors défini par : $\hat{q}_x = \frac{D_x}{E_x}$.

L'intervalle de confiance à 95% de cet estimateur est le suivant :

$$IC_{5\%} = \left(\hat{q}_x \pm 1.96 \sqrt{\frac{\hat{q}_x(1-\hat{q}_x)}{E_x}} \right)$$

3.2.2 Application aux données

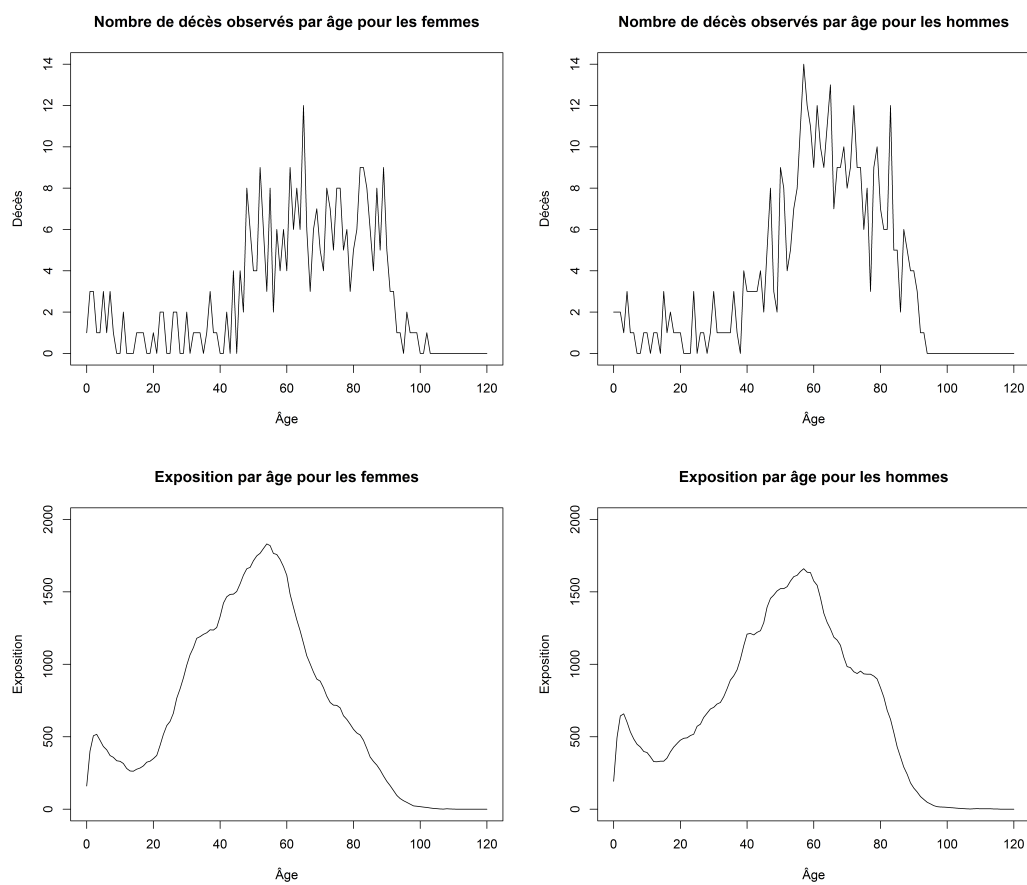
Pour l'étude de la mortalité, il est conseillé de distinguer les populations des hommes et des femmes car elles présentent souvent une mortalité différente (les femmes ont généralement une espérance de vie plus élevée). Nous allons donc étudier les données d'exposition pour chaque sexe.

L'exposition au risque de décès à l'âge x est calculée pour chaque individu comme étant la durée de présence entre deux âges $[x, x + 1[$, exprimée en fraction d'année. Cette durée de présence est calculée en prenant en compte l'entrée en observation et les censures. Un individu décédé est compté comme ayant une présence de 1 à l'âge de son décès.

On remarque (Graphique 3.1) qu'il y a le plus de données pour les âges compris entre 45 et 90 ans environ pour les décès, et entre 28 et 80 ans pour l'exposition. À cause de la faible taille de la base, il y a très peu de données pour les âges inférieurs, et il arrive régulièrement qu'il n'y ait aucun décès entre 10 et 30 ans. Cela va poser problème lors de la construction des taux de mortalité, qu'il faudra calculer différemment à ces âges-là.

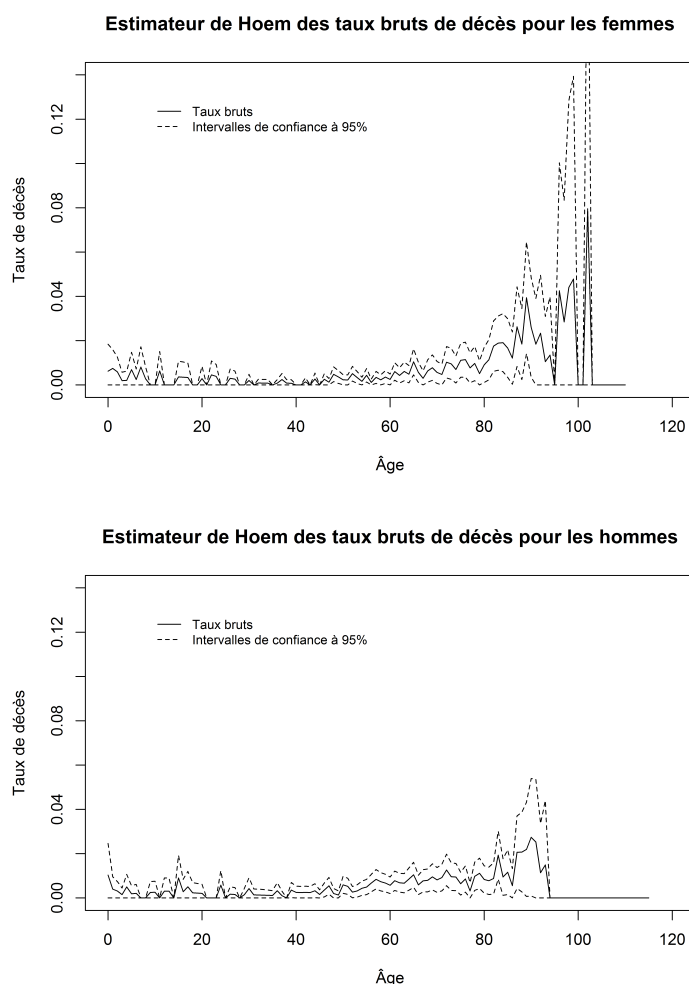
On remarque également que pour les âges compris entre 45 et 90 ans, la population féminine présente moins de décès que la population masculine. Par ailleurs, l'exposition de la population féminine est plus importante que celle de la population masculine jusqu'à environ 70 ans, puis c'est le contraire pour les âges supérieurs et jusqu'à 90 ans environ.

Cela nous permet d'avoir une première idée de la comparaison des mortalités des deux populations.



GRAPHIQUE 3.1 – Données d'exposition par âge

À partir des données de décès et d'exposition, nous pouvons désormais calculer les taux bruts de décès par la méthode de Hoem. Les intervalles de confiance sont calculés de façon à être toujours positifs, en prenant la valeur maximale entre la borne inférieure de l'intervalle et 0.



GRAPHIQUE 3.2 – Taux bruts de Hoem par sexe

On remarque que les taux de décès présentent une forte volatilité au-delà de 85 ans, et que les intervalles de confiance prennent souvent la valeur 0 en-dessous de 40 ans. La population masculine ne présentant aucun décès au-delà de 93 ans, les taux bruts de Hoem ne peuvent pas être estimés après cet âge. Pour la population féminine, l'estimation s'arrête quant à elle à 102 ans.

Il est donc nécessaire d'appliquer un critère afin de vérifier la pertinence du calcul des taux bruts pour les différents âges.

3.3 Critère de Cochran

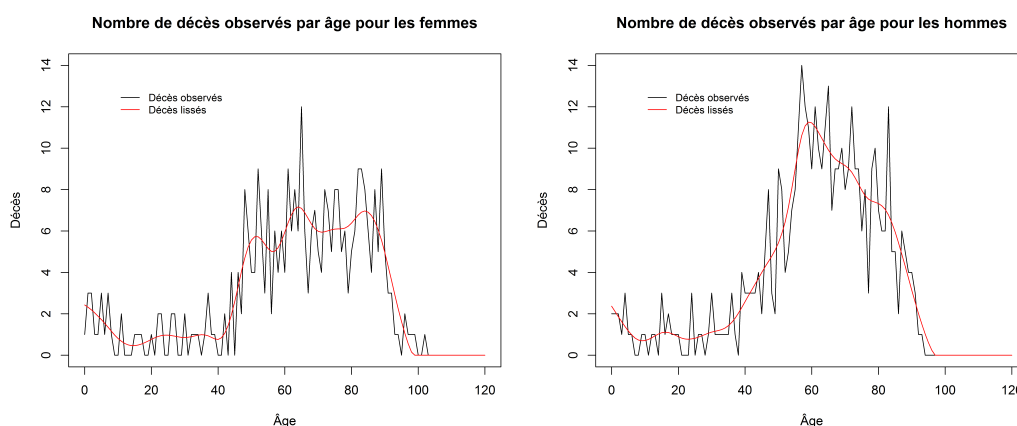
À cause du nombre faible de données, notamment avant 45 ans et aux âges élevés, nous devons vérifier que les données sont suffisantes pour construire une

table de mortalité. Pour cela, nous allons appliquer le critère de Cochran à nos données. Ce critère oblige à avoir au moins 5 décès par âge, et au moins 5 personnes non décédées par âge :

$$E_x \hat{q}_x \geq 5 \text{ et } E_x (1 - \hat{q}_x) \geq 5$$

Pour notre base d'études, les données sont clairement insuffisantes avant 45 ans, et au-delà de 90 ans, et cela est validé par le critère de Cochran.

Cependant, les résultats pour les deux populations ne donnent jamais de plage d'étude suffisamment grande. Cela est dû au fait que, les données étant très faibles, il arrive à tout âge que le nombre de décès observé soit inférieur à 5. Pour y remédier, nous pouvons lisser les nombres de décès grâce à la méthode de Whittaker-Henderson (voir 3.5.3) et appliquer le critère de Cochran ensuite. Nous obtenons le lissage suivant.



GRAPHIQUE 3.3 – Décès lissés par âge

Grâce à ce lissage, le nombre de décès observé ne descend pas en dessous de 5 sur une plage plus étendue, entre environ 50 et 90 ans. Il sera donc possible d'étudier la mortalité sur cette plage.

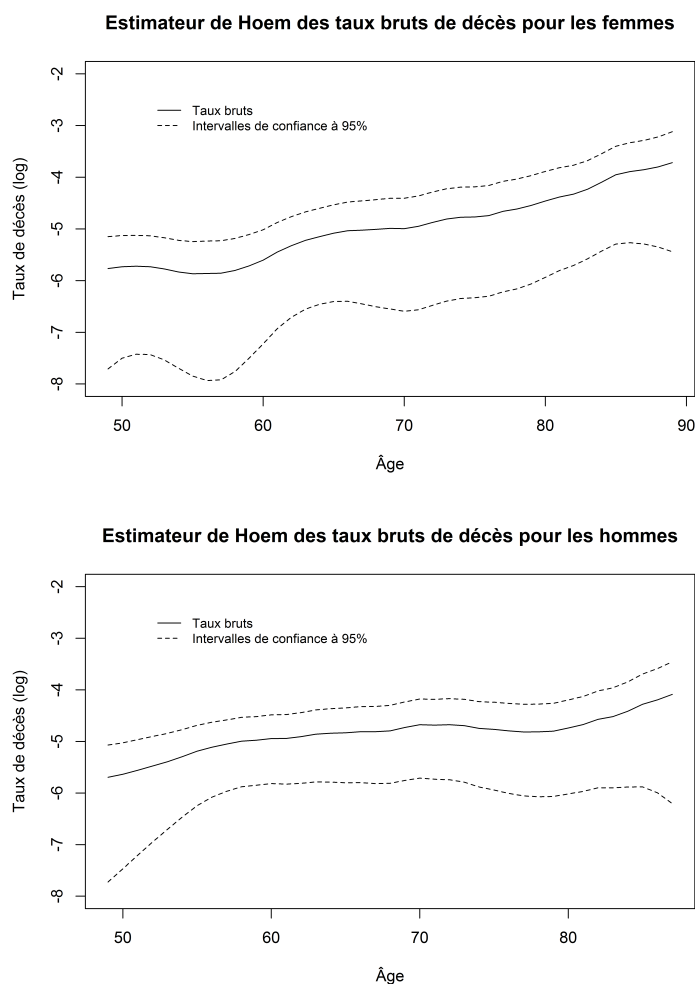
En effet, la construction de taux de mortalité se fait en deux temps. Des taux de décès bruts sont d'abord calculés à partir des observations sur la population. Vient ensuite une étape d'ajustement ou de lissage afin d'effacer les variations dues à l'échantillonnage.

Le fait de lisser le nombre de décès en amont, lors de l'application du critère de Cochran et du calcul des taux de Hoem, nous permet d'obtenir des taux bruts de Hoem plus lisses qui seront comparables non plus aux taux de décès bruts, mais aux taux de décès après lissage. Cela nous permet également de ne pas trop réduire la plage d'étude de la mortalité.

Pour la population féminine, seules les valeurs pour les âges compris entre 49 et 89 ans vérifient le critère de Cochran, après calcul des taux de Hoem à partir des décès lissés. Concernant la population masculine, la tranche d'âge est un peu plus courte et comprise entre 49 et 87 ans.

Nous étudierons donc la mortalité de chaque population pour les tranches d'âge déterminées, et les autres âges feront l'effet d'un traitement à part.

Les taux de Hoem finaux sont présentés dans les graphiques ci-dessous pour chaque population et selon la plage d'étude déterminée.



GRAPHIQUE 3.4 – Taux bruts de Hoem par sexe

3.4 Évolution de la mortalité sur la période d'étude

Comme notre étude s'étend sur une plage d'observation élevée (12 ans entre 2004 et 2015), il est important de vérifier qu'il n'y a pas d'évolution sensible de la mortalité entre le début et la fin de la période d'observation. Pour cela, nous pouvons appliquer des tests permettant de comparer la mortalité de deux populations, en comparant la population dont les premiers dossiers ont été ouverts avant le milieu de la période d'observation, c'est-à-dire avant le 01/01/2010, avec la population dont les premiers dossiers ont été ouverts après cette date.

Cette sous-partie s'inspire du livre *Survival analysis : techniques for censored and truncated data* [4].

3.4.1 Tests de comparaison des courbes de survie

Tests de rang

Pour analyser l'évolution de la mortalité sur la période d'étude, nous allons utiliser des tests à deux échantillons, qui ont pour hypothèse nulle que les fonctions de survies des deux échantillons sont égales.

Ces tests sont adaptés à la présence de censures et troncatures dans les données, et les deux populations étudiées doivent être indépendantes.

Pour que cette hypothèse soit vérifiée, nous construisons les deux bases de données à comparer de la manière suivante : on sépare les personnes dont la date d'ouverture du premier dossier est avant 2010, et celles dont les premiers dossiers ont été ouverts après 2010.

Notons A et B les deux échantillons créés pour la réalisation du test. Soient $t_1 < t_2 < \dots < t_D$ les dates de décès distinctes dans l'échantillon total.

À chaque instant t_i , et pour un échantillon $j = A$ ou B , on note :

- d_{ij} le nombre de décès dans l'échantillon j
- r_{ij} l'effectif sous risque dans l'échantillon j
- $d_i = d_{iA} + d_{iB}$ le nombre de décès dans l'échantillon total
- $r_i = r_{iA} + r_{iB}$ l'effectif sous risque dans l'échantillon total.

Le nombre de décès estimés en t_i dans la sous-population j est donc $D_{ij}^{\text{théo}} = d_i \frac{r_{ij}}{r_i}$.

Le nombre de décès estimés dans la sous-population j est alors $D_j^{\text{théo}} = \sum_i D_{ij}^{\text{théo}}$ et le nombre de décès observés dans cette sous-population est $D_j^{\text{obs}} = \sum_i d_{ij}$.

Le test de H_0 est fondé sur la comparaison pondérée des taux de hasard estimés pour une des populations. On peut ramener ce test à la comparaison pondérée du nombre de décès observé et du nombre de décès estimé à chaque date dans un des échantillons.

Les statistiques de chaque échantillon j sont donc :

$$Z_j = \sum_{i=1}^D w(t_i) \left(d_{ij} - D_{ij}^{\text{théo}} \right)$$

où $w(t_i)$ est un poids commun à chaque échantillon.

Comme nous sommes dans un cas avec deux échantillons, alors les nombres de décès observés et estimés dans chaque échantillon vérifient $d_{iA} + d_{iB} = D_{iA}^{\text{théo}} + D_{iB}^{\text{théo}}$.

Nous avons donc $Z_A = -Z_B$, et il n'est donc utile d'utiliser qu'une seule statistique, par exemple celle de la population A .

La variance de la statistique Z_A est donnée par :

$$\sigma_A^2 = \sum_i w(t_i)^2 d_i \frac{r_i - d_i}{r_i - 1} \frac{r_{iA} r_{iB}}{r_i^2}$$

La statistique du test peut alors s'écrire :

$$Z = \frac{\sum_{i=1}^D w(t_i)(d_{iA} - D_{iA}^{\text{théo}})}{\sqrt{\sum_i w(t_i)^2 d_i \frac{r_i - d_i}{r_i - 1} \frac{r_{iA} r_{iB}}{r_i^2}}}$$

et elle suit une loi normale centrée réduite pour des échantillons suffisamment larges sous l'hypothèse H_0 .

Selon les poids choisis, on construit différents tests :

- Test du Log Rank : $w(t_i) = 1$
- Test de Gehan (ou de Wilcoxon) : $w(t_i) = r_i$
- Test de Tarone-Ware : $w(t_i) = \sqrt{r_i}$
- Test de Peto-Peto : $w(t_i) = \hat{S}(t_i)$
- Test de Peto-Peto modifié : $w(t_i) = \hat{S}(t_i) \frac{r_i}{r_i + 1}$
- Test de Fleming-Harrington : $w(t_i) = \hat{S}(t_{i-1})^p (1 - \hat{S}(t_{i-1}))^q$

Chaque test insiste sur différents aspects de la courbe de survie. Par exemple, le test de Gehan donne plus de poids aux décès lorsque la population soumise au risque est importante, tandis que le test de Peto-Peto donne plus de poids aux décès précoces. D'après Peto et Peto (1972), ce test est plus robuste lorsque la censure entre les deux échantillons suit des schémas différents.

Cependant, ces tests atteignent leur limite lorsque les courbes de hasard des deux échantillons se croisent. En effet, dans ce cas, les différences qui existent avant le croisement des deux courbes se retrouvent annulées par les différences plus tardives, et les tests perdent de leur pouvoir.

Tests de type Renyi

Les tests de type Renyi sont utilisés pour les problèmes de comparaison dans le cas où les deux courbes des taux de hasard se croisent à une date de l'étude.

Lors de la construction des tests, une statistique de l'échantillon est calculée pour toutes les dates antérieures aux dates de décès $t_i, i = 1, \dots, D$. Il existe alors une date pour laquelle la valeur absolue de ces statistiques aura atteint une valeur maximale. C'est à partir de cette statistique que la statistique du test de type Renyi est construite.

En reprenant les notations de la partie précédente pour deux échantillons indépendants A et B , nous calculons pour chaque date $t_i, i = 1, \dots, D$, la statistique suivante :

$$Z_{t_i} = \sum_{t_k \leq t_i} w(t_k) (d_{kA} - D_{kA}^{\text{théo}})$$

La variance de la statistique $Z(t_D)$ est donnée par :

$$\sigma_{t_D}^2 = \sum_{t_k \leq t_D} w(t_k)^2 d_k \frac{r_k - d_k}{r_k - 1} \frac{r_{kA} r_{kB}}{r_k^2}$$

La statistique du test est alors donnée par :

$$Q = \frac{\sup\{|Z(t)|, t \leq t_D\}}{\sigma(t_D)}$$

Sous l'hypothèse H_0 , la distribution de Q peut-être approximée par la distribution de $\sup\{|B(x)|, 0 \leq x \leq 1\}$ où B est un mouvement Brownien standard. La p-value du test peut donc être calculée par :

$$\mathbb{P}(\sup\{|B(x)|, 0 \leq x \leq 1\} > y) = 1 - \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \exp\left(\frac{-\pi^2(2k+1)^2}{8y^2}\right)$$

Pour ces tests, il est possible d'utiliser les poids présentés précédemment.

3.4.2 Application aux données

Choix du seuil de signification

Avant de réaliser un test d'hypothèse, il faut choisir un seuil de signification pour le test. Ce seuil de signification est utilisé pour déterminer si les résultats du test sont statistiquement significatifs, ainsi que la probabilité d'erreur inhérente au test.

Dans notre étude, nous allons utiliser un seuil de signification de 0.01. L'hypothèse H_0 du test n'est pas rejetée si la p-value est supérieure à cette valeur.

Choix des poids utilisés

Nous n'allons étudier des tests de rang qu'avec deux poids différents :

- Le test du Log Rank : il s'agit en effet du test le plus souvent utilisé.
- Le test de Gehan : nous souhaitons réduire l'incertitude due au faible nombre de données. Ce test est donc adapté puisqu'il donne plus de poids aux âges où la population à risque est plus importante.

Population féminine

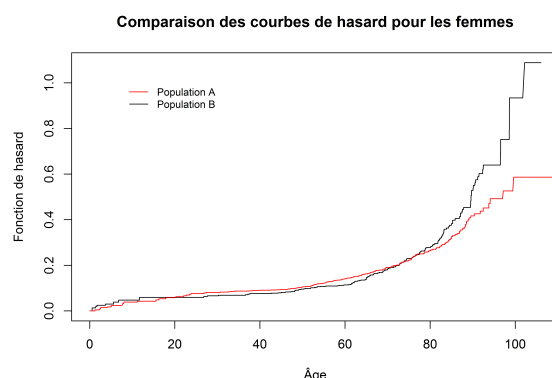
Tests de rangs Lorsque l'on applique le test du Log Rank à la population féminine séparée en deux en fonction de la première date d'ouverture du premier dossier, nous obtenons une statistique de 2.84, avec une p-value de 0.0919. Cette p-value est supérieure à 0.01, donc on ne rejette pas l'hypothèse H_0 . On peut donc considérer que les deux populations ont la même mortalité, et donc que la mortalité de la population féminine totale est stable dans le temps.

De la même manière, lorsque l'on applique le test de Gehan nous obtenons une statistique de 0.0884, avec une p-value de 0.7662. Cela confirme le résultat précédent, mais de façon beaucoup plus prononcée, puisque la p-value est beaucoup plus importante tout en restant négligeable devant 0.01.

En effet, regardons les courbes de hasards (Graphique 3.5) utilisées pour le test.

La population A est la population entrée en observation pour la première fois avant 2010.

On remarque que les courbes sont sensiblement proches jusqu'à un certain âge (elles commencent à s'éloigner à partir de l'âge de 80 ans environ), et notamment aux âges où l'exposition est la plus importante (entre 30 et 70 ans pour les femmes).



GRAPHIQUE 3.5 – Courbes de hasard pour deux sous-populations indépendantes de la population féminine

On remarque également que les courbes de hasard se croisent à deux reprises. Bien que les courbes ne s'éloignent pas fortement, nous pouvons appliquer le test de type Renyi pour vérifier le résultat obtenu.

Par ailleurs, il pourra arriver dans la suite de l'étude que les courbes soient plus distinctes ou que les résultats des tests de rang ne soient pas satisfaisants. Dans ces cas également, nous appliquerons le test de type Renyi.

Tests de type Renyi Le test du Log Rank appliqué à la population féminine confirme les résultats précédents puisque la p-value obtenue est de 0.1838. De la même manière, la p-value obtenue en appliquant le test de Gehan est de 0.3857.

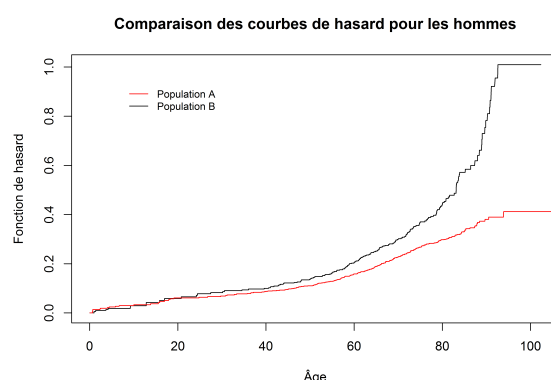
On remarque toutefois que les valeurs des p-value sont plus ou moins importantes que celles des tests de rang.

La mortalité de la population féminine est donc stable au cours du temps sur la période d'étude.

Population masculine

Tests de rangs De la même manière, appliquons le test du Log Rank à la population masculine séparée en deux sous-populations en fonction de la première date d'ouverture du premier dossier. Nous obtenons cette fois-ci une statistique de 32.68, avec une p-value de 1.089×10^{-8} . Celle-ci étant inférieure à 0.01, on rejette l'hypothèse H_0 et on ne peut pas supposer que la mortalité de la population masculine est stable dans le temps. Cependant, il est possible que ce test ne fonctionne pas car il n'est pas fiable lorsque les fonctions de hasard se croisent, ce qui est le cas dans nos données (Graphique 3.6), de manière cependant moins marquée que pour la population féminine.

De la même manière, lorsque l'on applique le test de Gehan nous obtenons une statistique de 19.10, avec une p-value de 1.2405×10^{-5} . Cela confirme le résultat précédent, mais de façon beaucoup moins prononcée, puisque la p-value est beaucoup plus importante.



GRAPHIQUE 3.6 – Courbes de hasard pour deux sous-populations indépendantes de la population masculine

On remarque que les courbes de hasard s'éloignent beaucoup plus rapidement que précédemment pour la population féminine (à partir de l'âge de 25 ans environ).

Essayons de voir si les tests de type Renyi améliorent les résultats.

Tests de type Renyi Le test du Log Rank et le test de Gehan appliqués à la population masculine permettent d'obtenir des p-value deux fois plus importantes, puisqu'elles sont de 1.892×10^{-8} et de 2.468×10^{-5} respectivement.

Cela ne permet toujours pas de supposer que la mortalité de la population masculine est stable dans le temps.

Cependant, lorsque l'on observe les courbes de hasard, nous remarquons que celles-ci deviennent significativement différentes à partir de 70 ans environ.

En effet, les p-value des tests deviennent supérieures au seuil de signification lorsque l'on ne considère que la population âgée entre 0 et 71 ans. Pour le test du Log Rank, la p-value est de 0.01386, tandis que pour le test de Gehan elle est de 0.01003.

Nous ne rejetons donc pas l'hypothèse H_0 pour cette tranche d'âge.

Pour les âges supérieurs à 71 ans, supposer que la mortalité reste stable sur la période d'étude introduit un comportement prudent dans l'étude. En effet, la mortalité semble augmenter avec le temps, et ne pas prendre en compte cette augmentation conduit à sous-estimer la mortalité pour les personnes entrées dans le portefeuille plus récemment.

Ce raisonnement sera conservé pour la suite de l'étude, à savoir que pour les prochaines populations étudiées, si la mortalité augmente avec le temps pour les âges élevés, nous supposons tout de même de manière prudente que la mortalité reste stable pour ces âges.

Nous pouvons donc réaliser notre étude de la mortalité sur la période d'observation 2004-2015 pour les deux populations féminine et masculine.

3.5 Traitement des âges intermédiaires

Pour étudier la mortalité sur la plage d'âge comprise entre 49 et 89 ans pour les femmes, et entre 49 et 87 ans pour les hommes, nous allons d'abord estimer des taux bruts de décès à partir de l'estimateur de Kaplan-Meier de la fonction de survie, puis comparer différentes méthodes de lissage et d'ajustement des taux bruts afin de réduire les irrégularités.

3.5.1 Estimateur de Kaplan-Meier de la fonction de survie

L'estimateur de Kaplan-Meier [3] est un estimateur non paramétrique de la fonction de survie. Il est par ailleurs adapté aux bases comportant des censures à droite.

Pour le calculer, il suffit de remarquer qu'il existe une relation de récurrence entre les fonctions de survie à deux dates $t > s$. En effet,

$$S(t) = \mathbb{P}(T > t) = \mathbb{P}(T > t | T > s) \mathbb{P}(T > s) = \mathbb{P}(T > t | T > s) S(s)$$

En notant T_i les dates de décès ou de censure ordonnées, on obtient :

$$S(t) = \prod_{T_i < t} p_i$$

où $p_i = P(T > T_i | T > T_{i-1})$ est la probabilité pour un individu de rester en vie entre T_{i-1} et T_i sachant qu'il était vivant en T_{i-1} .

On peut alors estimer p_i par $1 - \frac{d_i}{n_i}$, d'où l'estimateur de Kaplan-Meier de la fonction de survie :

$$\hat{S}(t) = \prod_{T_i \leq t} \left(1 - \frac{d_i}{n_i}\right)$$

où : d_i est le nombre de décès à la date i n_i est le nombre de survivants à la date i , prenant en compte les personnes en vie, les entrées, les sorties et les décès

On en déduit facilement les taux de mortalité à l'âge x :

$$\hat{q}_x = 1 - \mathbb{P}(T < x + 1 | T > x) = 1 - \prod_{x \leq T_i < x+1} \left(1 - \frac{d_i}{n_i}\right)$$

L'estimateur de la variance de $\hat{S}(t)$ est l'estimateur de Greenwood :

$$Var(\hat{S}(t)) = \hat{S}(t)^2 \sum_{T_i < t} \frac{d_i}{n_i(n_i - d_i)}$$

Cet estimateur se fonde sur l'hypothèse que les décès suivent une loi binomiale : $d_i \text{ Bin}(n_i, p_i)$, d'où :

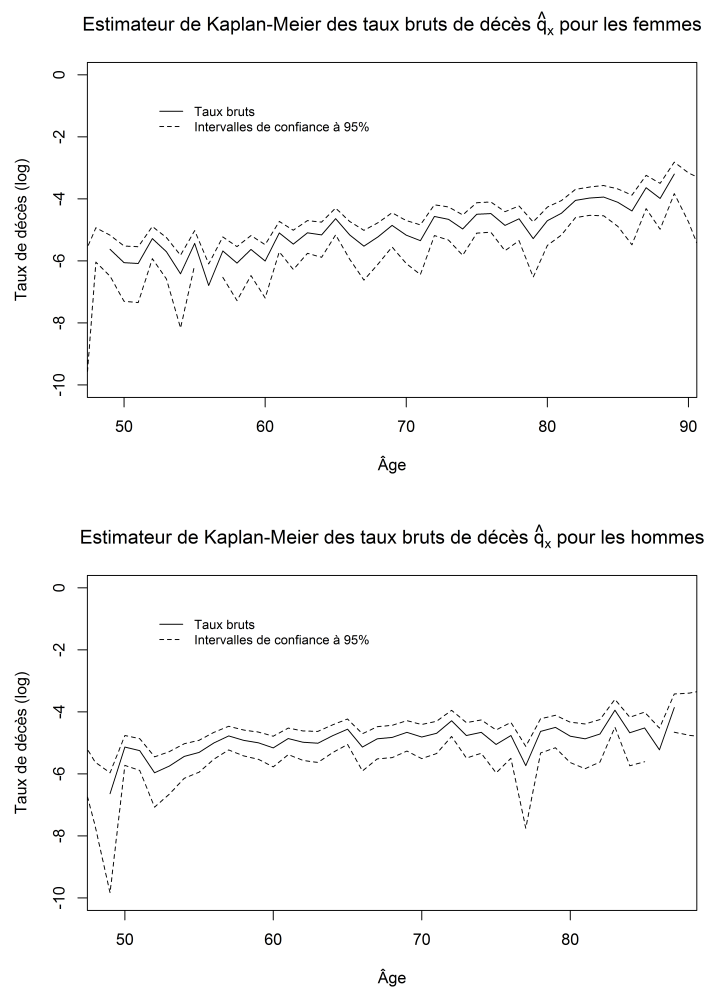
$$Var(\hat{q}_x) = \frac{\hat{q}_x(1 - \hat{q}_x)}{n_i}$$

On peut alors en déduire les intervalles de confiance des taux de mortalité de Kaplan-Meier :

$$IC_{5\%} = \left(\hat{q}_x \pm 1.96 \sqrt{\frac{\hat{q}_x(1-\hat{q}_x)}{E_x}} \right)$$

Pour le calcul sur notre base de données, les âges ont été calculés de manière exacte. Il a été pris comme hypothèse que deux personnes ne peuvent pas décéder exactement au même âge T_i , donc que $d_i = 1$ dans le calcul de $\hat{q}_x$. Dans le calcul de l'estimateur de la mortalité, les âges T_i sont en effet calculés de manière exacte, au jour près.

On obtient donc les taux bruts de décès suivants, après passage au logarithme pour plus de lisibilité :



GRAPHIQUE 3.7 – Estimateur de Kaplan-Meier des taux bruts de décès par sexe

On remarque que les taux de décès dans la population féminine sont plus variables que ceux de la population masculine. Nous parlons là à la fois des variations dues à l'échantillonnage, mais également du fait que l'augmentation des taux de décès avec l'âge est plus forte au sein de la population féminine. Les variations dues à l'échantillonnage seront lissées dans la prochaine partie.

3.5.2 Ajustement de Makeham

La méthode de Makeham est un ajustement paramétrique, qui part de l'hypothèse que le taux instantané de mortalité s'exprime à partir d'un certain âge de la façon suivante :

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

avec $\alpha > 0, \beta > 0, c > 1$.

Le coefficient α est un terme de mortalité accidentelle, qui est indépendant de l'âge, tandis que le deuxième terme croît exponentiellement avec l'âge et traduit le vieillissement.

On peut déduire de cette formule une expression de la probabilité de rester en vie à l'âge x :

$$\begin{aligned} p_x &= \exp\left(-\int_x^{x+1} \mu_y dy\right) \\ &= \exp\left(-\int_x^{x+1} (\alpha + \beta c^y) dy\right) \\ &= \exp(-\alpha) \exp\left(-\frac{\beta}{\ln c} c^x (c - 1)\right) \\ &= s g^{c^x (c-1)} \end{aligned}$$

avec $s = \exp(-\alpha)$ et $g = \exp(-\frac{\beta}{\ln(c)})$ pour simplifier l'écriture.

L'expression des taux de décès s'en déduit alors facilement.

Par ailleurs, si les taux instantanés de mortalité suivent une loi de Makeham, alors les termes $\ln(|q_{x+1} - q_x|)$ sont linéaires avec l'âge et alignés sur une droite de pente $\ln(c)$.

En effet, on peut faire l'approximation suivante :

$$\ln(1 - q_x) \simeq -q_x$$

et donc

$$q_x \simeq \ln(s) + \ln(g) c^x (c - 1)$$

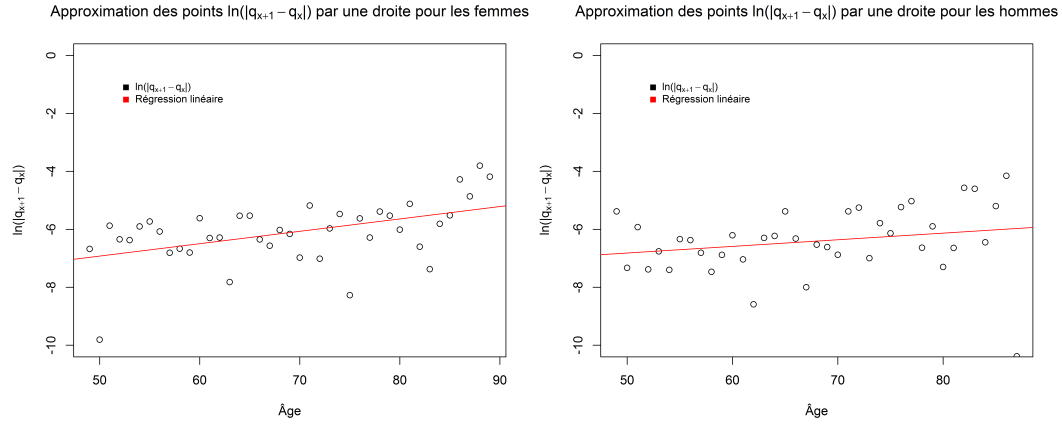
d'où :

$$\ln(|q_{x+1} - q_x|) = x \ln(c) + \ln(\ln(g)(c - 1)^2)$$

Il s'agit bien d'une équation linéaire.

Vérifions sur nos données que les points $\ln(|q_{x+1} - q_x|)$ peuvent être approximés par une droite.

Cela semble être le cas. Une première estimation de cette droite est calculée à l'aide d'une régression linéaire. On obtient alors :



GRAPHIQUE 3.8 – Approximation par sexe des points $\ln(|q_{x+1} - q_x|)$ par une droite

$$y = -9.06 + 0.0427x \text{ pour les femmes}$$

$$y = -7.96 + 0.0229x \text{ pour les hommes}$$

Nous pouvons désormais effectuer un ajustement de Makeham sur nos données. Pour cela, on utilise la méthode des moindres carrés pondérés.

$$\min_{(s,g,c)} \sum_x \frac{E_x}{(q_x(1-q_x))} (q_x^{bruts} - (1 - sg^{c^x(c-1)}))^2$$

sous les contraintes :

$$\begin{cases} 0 < s < 1 \\ 0 < g < 1 \\ c > 1 \end{cases}$$

Pour que l'algorithme converge vers les bonnes valeurs, nous pouvons prendre comme valeurs initiales pour g et c les valeurs issues de la régression linéaire précédente. En notant $\ln(|q_{x+1} - q_x|) = m + px$, on déduit en effet des équations précédentes :

$$\begin{cases} c_0 = e^p \\ g_0 = \exp\left(-\left|\frac{e^m}{(c-1)^2}\right|\right) \end{cases}$$

En appliquant ces équations à nos données, on obtient pour les femmes :

$$\begin{cases} c_0 = 1.0436 \\ g_0 = 0.9406 \end{cases}$$

Et les valeurs pour les hommes sont :

$$\begin{cases} c_0 = 1.0232 \\ g_0 = 0.5219 \end{cases}$$

Pour le dernier paramètre s , il n'est pas possible d'obtenir une valeur initiale dont l'expression est indépendante de x :

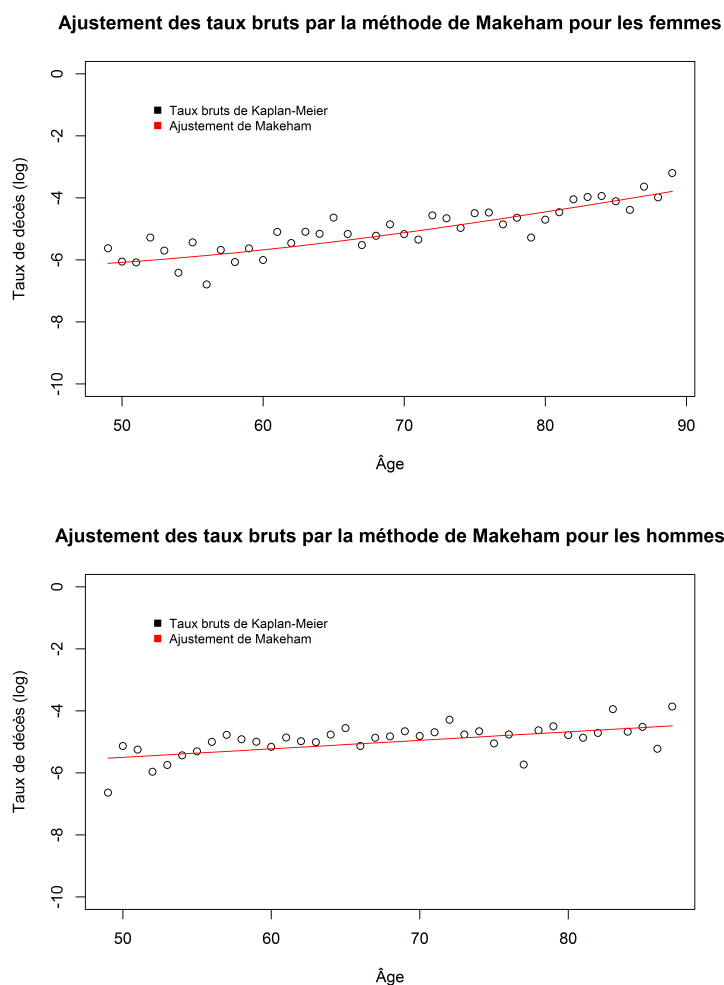
$$s_0 = \exp(\ln(1 - q_x^{\text{lissés}} - c^x(c-1)\ln(g)))$$

Nous allons donc arbitrairement prendre pour valeur initiale de s la valeur $s_0 = 0.9999$, qui est une valeur proche des résultats finaux. Il a été vérifié que les résultats restaient identiques en prenant comme valeur initiale une valeur éloignée, par exemple 0.0001.

Avec ces valeurs initiales, l'algorithme converge vers les points optimaux suivants :

$$\begin{cases} (s, g, c) = (0.998, 0.999, 1.110) \text{ pour les femmes} \\ (s, g, c) = (0.999, 0.942, 1.025) \text{ pour les hommes} \end{cases}$$

On obtient l'ajustement présenté dans le graphique 3.9.



GRAPHIQUE 3.9 – Ajustement de Makeham par sexe

On remarque que la courbe lissée correspond bien aux données, mais semble légèrement surestimer la mortalité autour de 50 ans pour les hommes, et la sous-estimer autour entre 60 et 70 ans pour les deux populations.

Cependant, l'ajustement nous paraît satisfaisant pour le moment.

3.5.3 Lissage de Whittaker-Henderson

Nous allons également appliquer la méthode de lissage non-paramétrique de Whittaker-Henderson afin de comparer deux méthodes et choisir la plus appropriée.

Le principe de cette méthode est de combiner un critère de fidélité et un critère de régularité, et de rechercher les valeurs ajustées qui minimisent la somme des deux critères. Ainsi, en dimension un, on se fixe des poids (w_i) et on pose pour le critère de fidélité :

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

et pour critère de régularité :

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

Ici, z est un paramètre du modèle et p est le nombre d'observations des $\hat{q}_i$

Le critère à minimiser est une combinaison linéaire de la fidélité et de la régularité, le poids de chacun des deux termes étant contrôlé par un second paramètre h :

$$M = F + hS$$

Posons $q = (q_i)_{1 \leq i \leq p}$, $\hat{q} = (\hat{q}_i)_{1 \leq i \leq p}$, $w = \text{diag}(w_i)$ et $\Delta^z q = (\Delta^z q_i)_{(1 \leq i \leq p-z)}$. Comme :

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

Alors, en introduisant la matrice K_z de taille $(p-z, p)$, dont les termes sont les coefficients binomiaux d'ordre z dont le signe alterne et commence positivement pour z pair, le critère à minimiser peut s'écrire sous la forme :

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

La résolution de $\frac{\partial M}{\partial q} = 0$ conduit alors à l'expression des taux ajustés :

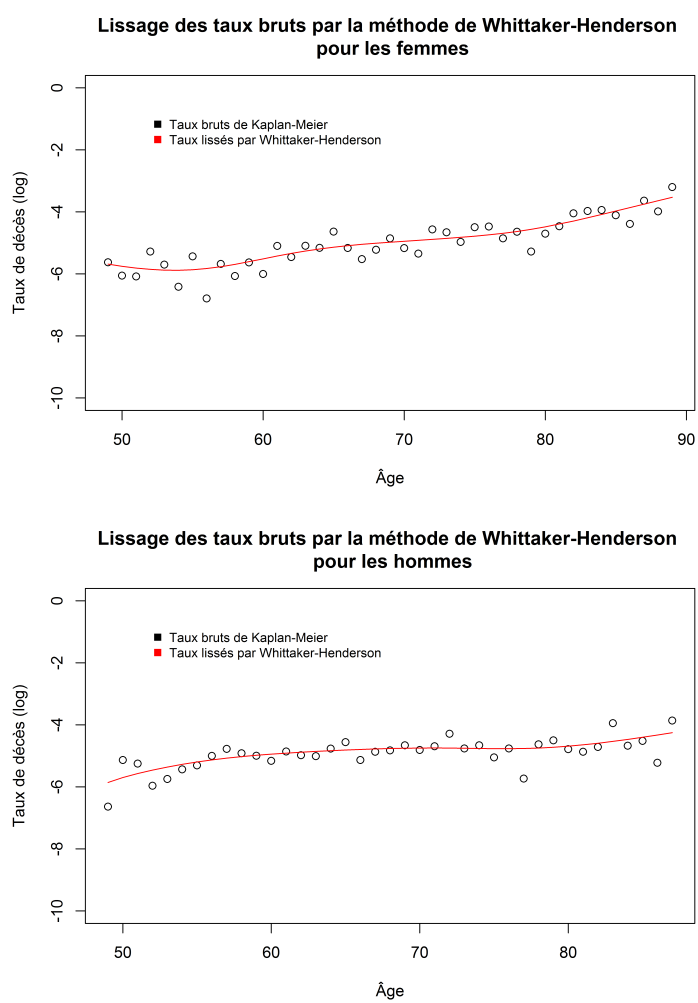
$$q^* = (w + h K_z' K_z q)^{-1} w \hat{q}$$

Pour notre application, les poids choisis correspondent à l'exposition à chaque âge.

Le paramètre de lissage h choisi vaut 1000000 et le paramètre z vaut 3.

Avec les paramètres choisis, le lissage par la méthode de Whittaker-Henderson semble mieux approximer les taux bruts que l'ajustement de Makeham (Graphique 3.10).

Cependant, pour valider cette hypothèse, nous allons devoir nous appuyer sur certains critères et tests.



GRAPHIQUE 3.10 – Lissage de Whittaker-Henderson par sexe

3.5.4 Validation des ajustements et choix de la méthode

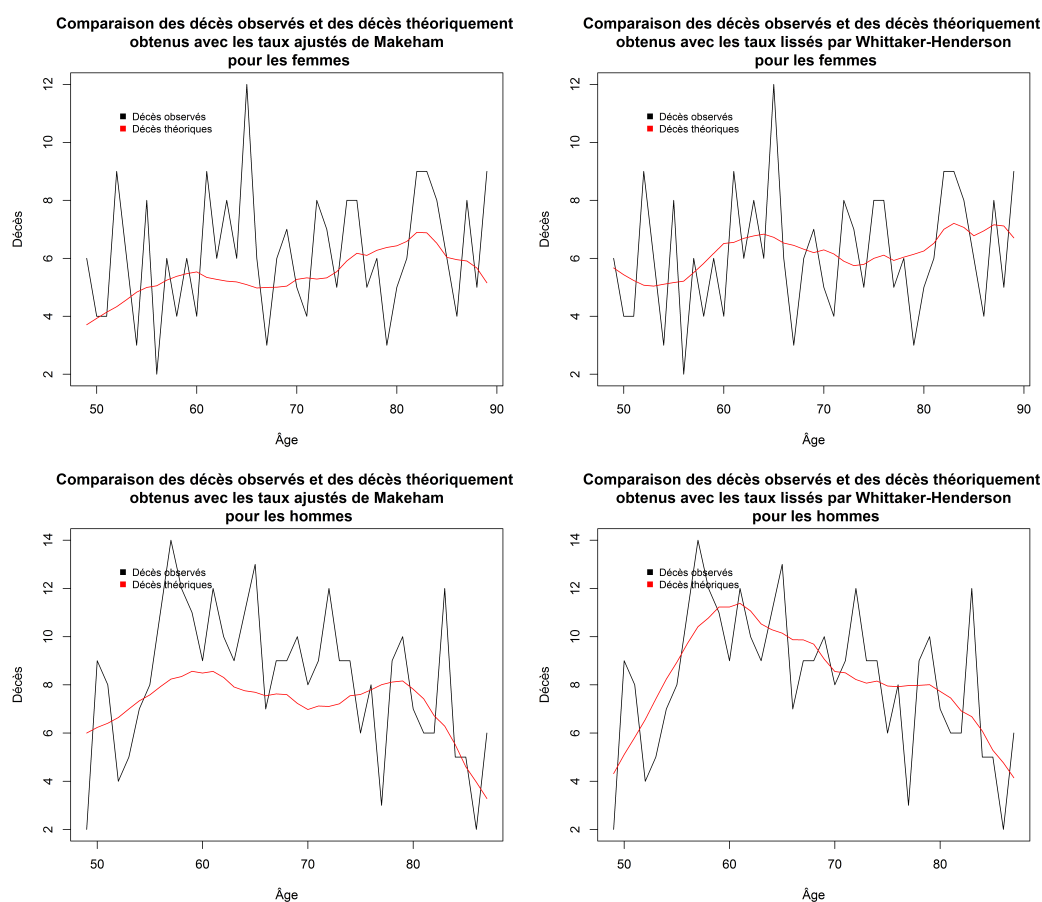
Comparaison des décès observés et des décès théoriques

Afin d'avoir une première idée de la qualité de l'ajustement, on peut comparer les décès observés et les décès théoriques qui seraient obtenus avec les tables lissées.

Les décès théoriques sont calculés en faisant le produit des taux de décès lissés et de l'exposition pour chaque âge :

$$D_x^{\text{théo}} = q_x^{\text{lissés}} E_x$$

En appliquant cela à nos deux ajustements possibles, nous obtenons les graphiques de comparaison suivants :



GRAPHIQUE 3.11 – Comparaison des décès observés et des décès théoriques

Comme remarqué précédemment, les taux ajustés par Makeham estiment globalement mal les décès. En revanche, le lissage de Whittaker-Henderson semble plutôt bien prédire nos données.

Critère de fidélité aux taux bruts

Ce critère consiste à comparer les taux bruts et les taux ajustés. Afin de mesurer la qualité de l'ajustement, on regarde la somme des carrés des différences entre les taux :

$$\sum_x (q_x^{\text{ajustés}} - q_x^{\text{bruts}})^2$$

Plus ce critère est proche de 0, meilleur est l'ajustement.

En faisant les calculs pour la population féminine avec les taux ajustés par Makeham, on obtient comme valeur de ce critère 6.1310^{-4} . En revanche pour les données lissées par la méthode de Whittaker-Henderson, ce critère vaut 4.3010^{-4} .

Cela confirme donc que le lissage de Whittaker-Henderson se rapproche mieux des données brutes que l'ajustement de Makeham.

De la même manière, le critère calculé au sein de la population masculine vaut 3.8410^{-4} pour l'ajustement de Makeham, et 3.0010^{-4} pour le lissage de Whittaker-Henderson. C'est donc cette dernière méthode qui est de nouveau appropriée.

Critère de régularité des taux lissés

Il est également possible de mesurer la régularité des taux lissés. Cela consiste à vérifier que la différence entre les taux consécutifs n'est pas trop importante, c'est-à-dire qu'il faut minimiser le critère suivant :

$$\sum_x (q_x - q_{x+1})^2$$

Cette fois-ci, c'est le lissage de Whittaker-Henderson qui donne une valeur plus élevée du critère, valant respectivement 4.3410^{-5} contre 1.8010^{-5} pour l'ajustement de Makeham dans le cadre de la population féminine, et 6.1010^{-6} contre 1.5410^{-6} pour l'ajustement de Makeham dans le cadre de la population masculine.

En effet, on voit bien sur les graphiques précédents que les taux ajustés par Makeham augmentent de façon régulière, tandis que les taux lissés par Whittaker-Henderson présentent plusieurs points d'inflexion qui traduisent une certaine irrégularité.

Test du Khi-deux

Les différents critères que nous venons de présenter ne permettent pas de choisir le meilleur ajustement, mais simplement de se donner une idée de la qualité de l'ajustement. Pour l'instant, c'est le lissage de Whittaker-Henderson qui semble le plus approprié, mais nous allons vérifier cela avec un test statistique.

Le test du Khi-deux consiste à tester l'hypothèse nulle selon laquelle les taux ajustés ne sont pas significativement différents des taux bruts estimés. Il s'agit ici de comparer le nombre de décès espérés aux décès réels, grâce à la statistique suivante :

$$\chi^2 = \sum_{x_{min}}^{x_{max}} \left(1 - \frac{D_x^{\text{obs}}}{D_x^{\text{théo}}} \right)$$

Cette statistique suit une loi du Khi-deux à $x_{max} - x_{min} - n$ degrés de liberté, où n est le nombre de paramètres de la méthode de lissage utilisée. Ainsi, pour l'ajustement de Makeham, $n = 3$, tandis que pour le lissage de Whittaker-Henderson, $n = 0$.

L'ajustement est alors accepté si la statistique est inférieure au quantile à 95% de la loi du Khi-deux correspondante.

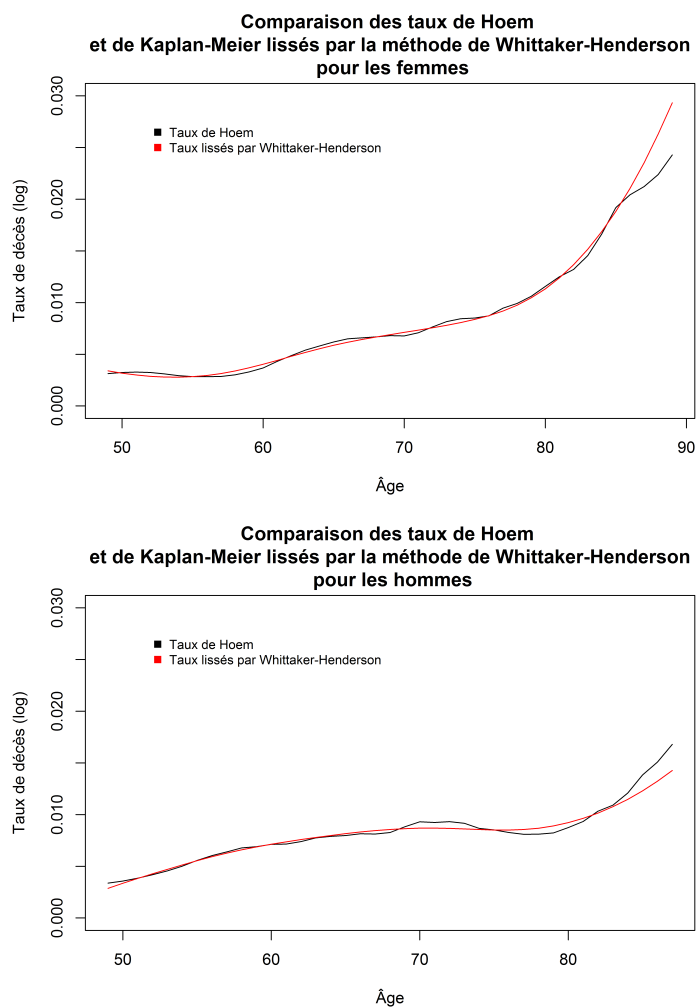
Concernant l'ajustement de Makeham, la statistique du test pour la population féminine vaut 37.77, et le quantile d'ordre 95% d'une loi du χ^2 à 37 degrés de libertés vaut 52.19. La statistique étant inférieure à la valeur du quantile, nous pouvons accepter cet ajustement. Pour la population masculine, l'ajustement est également validé, puisque la statistique du test vaut alors 41.33, et le quantile d'ordre 95% d'une loi du χ^2 à 35 degrés de libertés vaut 49.80.

Pour le lissage de Whittaker-Henderson, la statistique du test pour la population féminine vaut 24.92 tandis que le quantile à 95% d'une loi du χ^2 à 40 degrés de libertés est de 55.76. La statistique étant de nouveau inférieure à la valeur du quantile, nous pouvons également accepter ce lissage. Enfin, concernant la population masculine, le lissage est validé car la statistique du test est de 26.18 et est donc inférieure au quantile d'ordre 95% d'une loi du χ^2 à 38 degrés de libertés, qui vaut 53.38.

Nous allons donc conserver les taux bruts lissés par Whittaker-Henderson pour la suite de l'étude.

Comparaison des taux lissés par Whittaker-Henderson avec les taux de Hoem

Il ne nous reste plus qu'à vérifier que les taux lissés ne sont pas éloignés des taux de Hoem calculés au début de cette étude. Cela est bien le cas pour les deux populations féminine et masculine (Graphique 3.12).



GRAPHIQUE 3.12 – Comparaison des taux lissés par Whittaker-Henderson avec les taux de Hoem

3.6 Traitement des âges jeunes (avant 49 ans)

Pour les âges compris entre 0 et 49 ans, le nombre de données était insuffisant pour calculer les taux bruts et auraient résulté en une forte variance des résultats.

Nous allons donc les traiter différemment en réalisant des ajustements par rapport à une table de référence.

Pour la population féminine, c'est la TGF05 qui sera la table de référence pour les ajustements, tandis que la TGH05 sera prise en compte pour la population masculine. Ces tables seront prises en vision par année courante et non par génération, pour correspondre à nos données. Celles-ci sont obtenues par une simple transformation linéaire avec l'âge x : génération = année courante - x .

À l'issue de cette transformation, nous considérons les tables de référence pour les années courantes en commun avec notre période d'étude, à savoir la période 2004-2015. Parmi cette période, seules les années courantes 2004 et 2005 présentent des taux de décès pour chaque âge avant 49 ans, tandis qu'il n'y a pas d'informations pour les jeunes âges les années suivantes. Au cours de ces deux années courantes, la mortalité des tables de référence n'a pas subi de fortes modifications, et la table pour l'année courante 2005 peut être considérée comme présentant une mortalité plus proche de celle des années suivantes dans la période d'étude que la table pour l'année courante 2004. Nous allons donc utiliser les tables TGF05 et TGH05 pour l'année courante 2005 pour notre base de données, dont les individus sont observés entre 2004 et 2015.

Par ailleurs, les tables réglementaires traditionnelles présentent un forte baisse de la mortalité après la naissance et jusqu'au début de l'adolescence, puis une surmortalité pour les jeunes adultes autour de 20 ans.

Ce comportement ne semble pas être présent dans notre base de données. Au contraire, les 0-18 ans semblent présenter une mortalité identique voire plus élevée que les plus de 20 ans.

Pour prendre cela en compte, nous allons ajuster séparément les tables pour les 0-17 ans, puis pour les 18-48 ans. Cela permettra par ailleurs d'évaluer correctement les provisions pour les rentes temporaires jusqu'à 18 ans.

3.6.1 Abattement par le ratio SMR

La méthode d'abattement par le ratio SMR (Standardized Mortality Ratio) est un modèle à un paramètre, qui suppose que les taux de décès de la population observée sont proportionnels à ceux de la population de référence.

Les taux de décès des âges jeunes vont ainsi être remplacés par les taux des tables réglementaires TGH/F05, auxquelles aura été appliqué un abattement avec le ratio SMR.

Ce ratio est calculé en faisant le rapport entre le nombre total de décès observés dans la population et le nombre total de décès théoriques qui seraient attendus si la population observée avait une mortalité identique à la table de référence.

Ainsi, les taux de mortalité de la population observée et de la table de référence sont quasiment identiques lorsque le ratio SMR est proche de 1.

On a donc :

$$SMR = \frac{D^{\text{obs}}}{D^{\text{théo}}}$$

$$\text{avec : } \begin{cases} D^{\text{obs}} = \sum_x D_x^{\text{obs}} \\ D^{\text{théo}} = \sum_x D_x^{\text{théo}} = \sum_x q_x^{\text{réf}} E_x \end{cases}$$

Population féminine

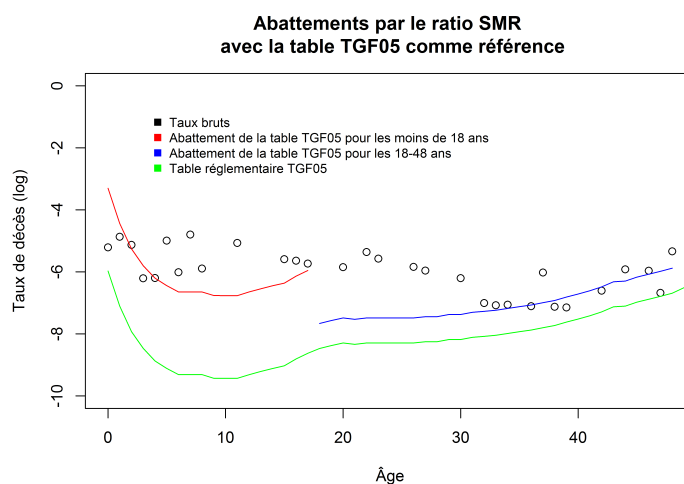
En calculant le ratio sur les données de la population féminine et la TGF05, nous obtenons les ratios SMR suivants :

	0-17 ans	18-48 ans
SMR	14.404	2.241

TABLE 3.1 – Ratios SMR pour la population féminine

Les taux de mortalité de notre population sont donc très éloignés de ceux de la table TGF05 pour les moins de 18 ans, et également un peu éloignés pour les plus de 18 ans. De manière générale, la table de référence TGF05 sous-estime la mortalité de la population féminine, puisque le ratio SMR est supérieur à 1 pour les deux populations.

On obtient alors les taux présentés dans le graphique suivant, que l'on compare avec les taux bruts et avec la table de référence.

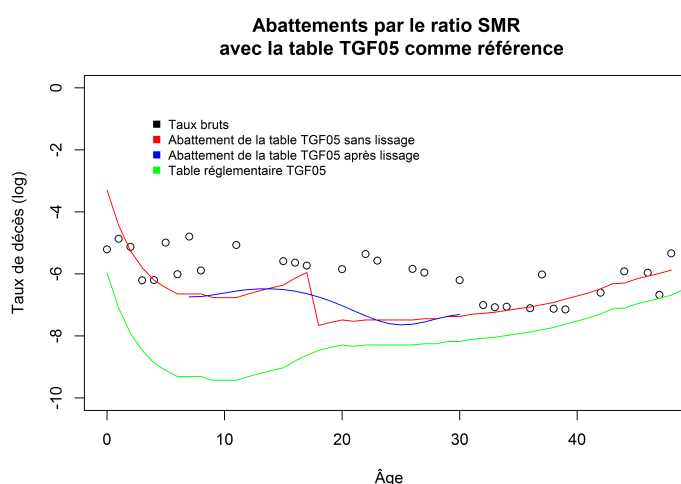


GRAPHIQUE 3.13 – Taux abattus par le ratio SMR pour la population féminine

On observe un saut important des taux de décès entre 17 et 18 ans. Nous allons donc lisser les taux autour de ces âges. Nous obtenons alors la table suivante :

Les taux de décès estiment correctement la mortalité de la population aux âges extrêmes, et la sous-estiment de 10 à 30 ans du fait du lissage.

Ce comportement, bien que non idéal, n'est pas inquiétant puisqu'il introduit une certaine prudence dans l'estimation de la mortalité des rentiers. Cette sous-estimation pourra être éventuellement réduite par l'utilisation d'autres méthodes



GRAPHIQUE 3.14 – Taux abattus par le ratio SMR et lissés pour la population féminine

d'ajustement à une table de référence que celle du ratio SMR, comme le modèle de Brass présenté par la suite.

Population masculine

De la même manière, les calculs sur les données de la population masculine et la TGH05 renvoient les ratios SMR suivants :

	0-17 ans	18-48 ans
SMR	9.217	3.284

TABLE 3.2 – Ratios SMR pour la population masculine

Les taux de mortalité de la population masculine sont également très éloignés de ceux de la table TGH05 pour les moins de 18 ans, mais sont plus rapprochés entre eux, puisque les ratios SMR sont plus proches. Encore une fois, la table TGH05 sous-estime la mortalité des moins de 49 ans.

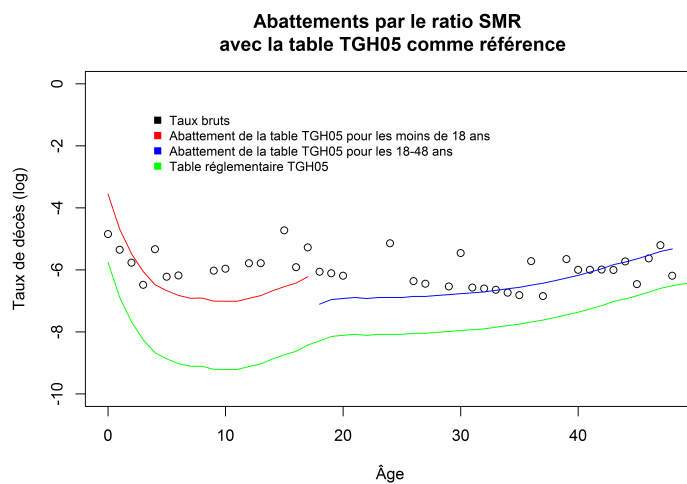
On obtient alors les taux présentés dans le graphique suivant, que l'on compare avec les taux bruts et avec la table de référence.

On observe là encore un saut important des taux de décès entre 17 et 18 ans. La table après lissage est la suivante :

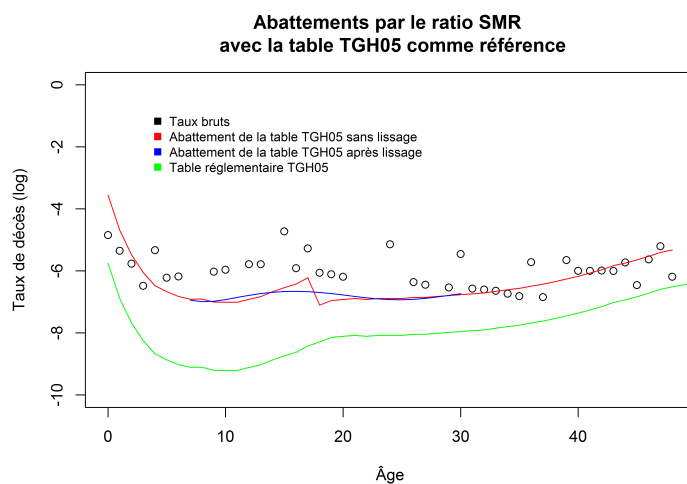
Les taux de décès sous-estiment globalement la mortalité de la population, sauf éventuellement pour les moins de 5 ans et les plus de 35 ans.

De la même manière que pour la population féminine, cette table introduit donc un comportement prudent pour les moins de 30 ans du fait de la sous-estimation des taux bruts de décès.

Nous allons dès lors essayer d'utiliser une autre méthode d'ajustement, le modèle de Brass, dans l'espoir de mieux approximer la mortalité de la population observée.



GRAPHIQUE 3.15 – Taux abattus par le ratio SMR pour la population masculine



GRAPHIQUE 3.16 – Taux abattus par le ratio SMR et lissés pour la population masculine

3.6.2 Modèle de Brass

L'inconvénient d'un abattement par le ratio SMR est qu'il s'agit d'une méthode à un seul paramètre et s'appliquant directement aux taux de décès. Il suppose que la mortalité de la population étudiée est proportionnelle à la mortalité de référence, et pourrait alors permettre d'obtenir des taux supérieurs à 1.

En pratique, notre population observée ne semble pas présenter le même comportement que la population de référence en terme de mortalité, même après avoir séparé la population des moins de 49 ans en deux populations constituées des moins de 18 ans et des plus de 18 ans.

En effet, cette première population des moins de 18 ans semble présenter une mortalité plus élevée que la population plus âgée, et on n'observe pas le comportement de la table de référence correspondant à un pic de mortalité à 20 après une forte baisse des taux de décès lors de l'enfance.

Nous cherchons donc à appliquer un modèle qui modifie à la fois le niveau de mortalité global, mais qui permet également de diminuer la proportion de survivants aux âges jeunes.

En 1971, Brass introduit un modèle relationnel à deux paramètres après avoir découvert qu'il est possible d'obtenir une représentation linéaire entre les logits des probabilités de décès ${}_xq_0$ de la population étudiée et ceux de la population de référence, où les ${}_xq_0$ définissent les probabilités de décéder x années après la naissance. Par définition, nous avons donc ${}_xq_0 = 1 - \mathbb{P}(T > x | T > 0) = 1 - \mathbb{P}(T > x) = 1 - S(x)$ car $\mathbb{P}(T > 0) = 1$.

La fonction logit est une fonction de $[0, 1]$ dans $\mathbb{R}$, définie de la manière suivante :

$$\text{logit}(x) = \frac{1}{2} \ln\left(\frac{x}{1-x}\right)$$

Le modèle de Brass, défini par la relation suivante :

$$\text{logit}({}_xq_0) = \alpha + \beta \text{logit}({}_xq_0^{\text{ref}})$$

permet donc par inversion des logits d'obtenir des probabilités comprises entre 0 et 1.

En pratique, c'est la relation suivante qui sera utilisée, par souci de simplification et en prenant en compte la relation entre les taux de décès et les probabilités de survie :

$$\ln\left(\frac{1-S(x)}{S(x)}\right) = \alpha + \beta \ln\left(\frac{1-S^{\text{ref}}(x)}{S^{\text{ref}}(x)}\right)$$

Elle est bien évidemment identique à la relation précédente, seules les valeurs estimées des coefficients seront modifiées.

Notons :

- $y_x = \ln\left(\frac{1-S(x)}{S(x)}\right)$ pour la population observée
- $z_x = \ln\left(\frac{1-S^{\text{ref}}(x)}{S^{\text{ref}}(x)}\right)$ pour la population de référence

Ces coefficients seront estimés par la méthode des moindres carrés pondérés, en minimisant le critère suivant :

$$\min_{\alpha, \beta} \sum_{x=x_{\min}}^{x_{\max}} w_x [y_x - (\alpha + \beta z_x)]^2$$

où w_x représente le poids associé à l'âge x .

La minimisation du critère permet alors d'obtenir les valeurs estimées suivantes des paramètres :

$$\begin{cases} \hat{\alpha} = \bar{y}_w - \beta \bar{z}_w \\ \hat{\beta} = \frac{\sum_{x=x_{min}}^{x_{max}} w_x (z_x - \bar{z}_w)(y_x - \bar{y}_w)}{\sum_{x=x_{min}}^{x_{max}} w_x (z_x - \bar{z}_w)^2} \end{cases}$$

Le paramètre α est dans ce contexte un indicateur du niveau de mortalité qui affecte tous les âges de manière égale, tandis que β influence la mortalité selon les âges. Selon que α est négatif ou positif, le niveau de mortalité est respectivement inférieur ou supérieur au niveau de mortalité de la table de référence. En revanche, le positionnement de β par rapport à 1 modifie la balance de la mortalité aux jeunes âges et aux âges élevés par rapport à la table de référence : plus la valeur de β est grande, plus la proportion de survivants est grande parmi les âges jeunes et donc plus la mortalité est faible à ces âges et importante parmi les âges élevés, toujours relativement à la population de référence. La constatation inverse est faite si β est inférieur à 1.

Dans le cas de nos deux populations, nous avons déjà observé une surmortalité globale de la population, et plus particulièrement une mortalité plus élevée pour les âges jeunes. Dans le cadre du modèle de Brass, il faut cependant comparer la mortalité de la population observée à celle de la population de référence.

Nous nous attendons donc à obtenir une valeur de β supérieure à 1 lors de l'ajustement pour les moins de 18 ans. Ceux-ci présentent effectivement une mortalité fortement éloignée de la table de référence, et cette différence augmente avec l'âge. Au contraire, la différence de mortalité pour les 18-48 ans s'amenuise avec l'âge, donc nous nous attendons à avoir un paramètre β inférieur à 1 pour cette population.

Population féminine

Nous appliquons dans un premier temps le modèle de Brass à la population féminine avec comme table de référence la TGF05.

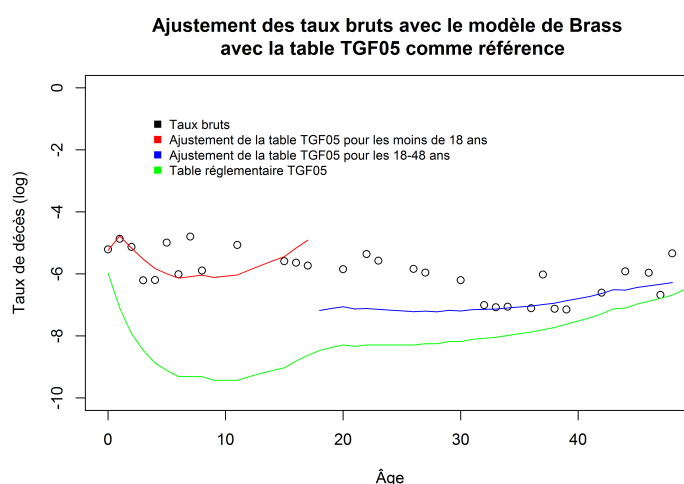
Les valeurs estimées des paramètres sont les suivantes :

	α	β
0-17 ans	15.22	3.43
18-48 ans	-1.06	0.31

TABLE 3.3 – Valeurs des paramètres du modèle de Brass pour la population féminine

Les valeurs estimées des paramètres correspondent bien à ce que nous avons déduit du comportement de la population féminine observée. Nous avons effectivement un paramètre β estimé à une valeur supérieure à 1 pour les 0-17 ans, et inférieure à 1 pour les 18-48 ans.

Après inversion de la fonction logit et calcul des taux de décès, nous obtenons les taux ajustés présentés dans le graphique 3.17.

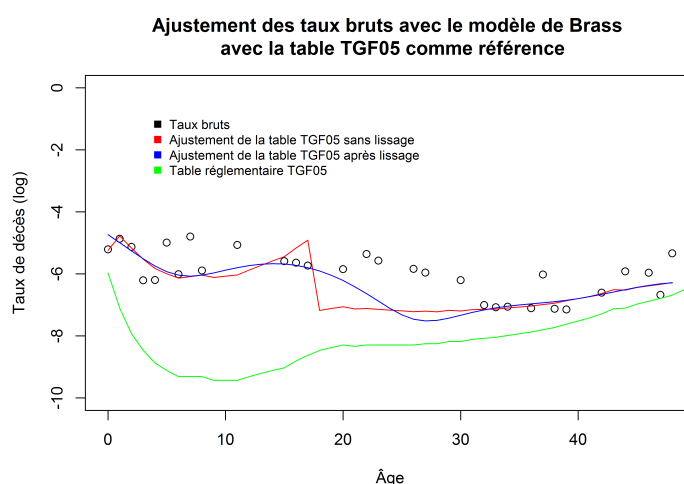


GRAPHIQUE 3.17 – Taux ajustés par le modèle de Brass pour la population féminine

Contrairement à la méthode d'abattement par le ratio SMR, ce modèle estime correctement la mortalité pour les 0-17 ans. La différence pour les 18-45 est moins importante, mais on voit bien également que la « pente » des taux de décès ajustés sur cette tranche d'âge est plus faible qu'avec la méthode du ratio SMR, et que les taux ajustés sont plus proches des taux bruts.

Cependant, il est encore une fois nécessaire de lisser ces taux autour de 18 ans, ainsi qu'à la naissance pour éviter les décrochages à ces âges.

Les taux lissés obtenus sont présentés dans le graphique 3.18.



GRAPHIQUE 3.18 – Taux ajustés par le modèle de Brass et lissés pour la population féminine

Ceux-ci semblent de nouveau sous-estimer la mortalité de la population, mais sur une tranche d'âge plus réduite, à savoir entre 20 et 30 ans. En dehors de cette tranche d'âge, les taux bruts sont correctement approximés.

Population masculine

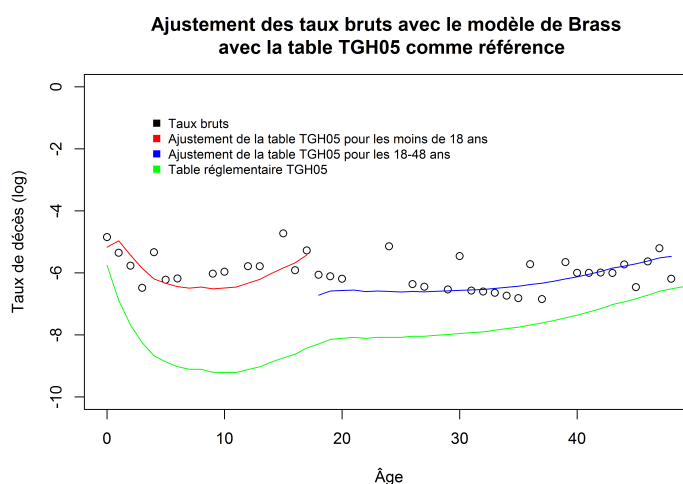
Le modèle de Brass appliqué à la population masculine nous renvoie les valeurs estimées suivantes des paramètres :

	α	β
0-17 ans	11.34	2.87
18-48 ans	0.15	0.60

TABLE 3.4 – Valeurs des paramètres du modèle de Brass pour la population masculine

Ces estimations sont du même ordre que celles pour la population féminine, avec notamment un paramètre β estimé à une valeur supérieure à 1 pour les 0-17 ans, et inférieure à 1 pour les 18-48 ans.

Les taux ajustés obtenus après inversion de la fonction logit sont présentés dans le graphique 3.19.

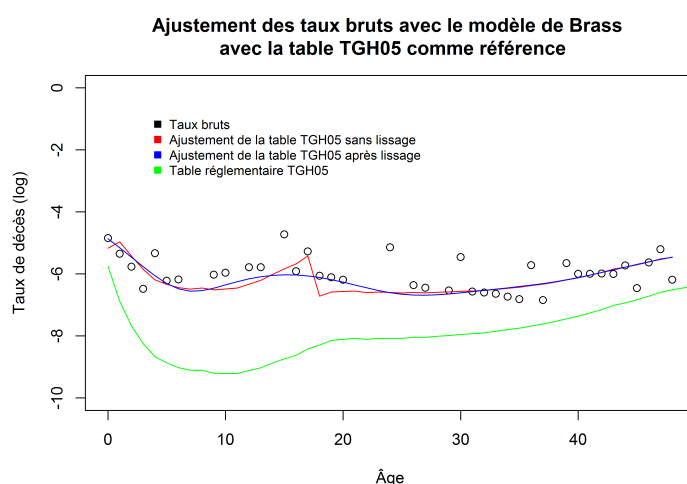


GRAPHIQUE 3.19 – Taux ajustés par le modèle de Brass pour la population masculine

Pour la population masculine, les taux ajustés sont très proches des taux bruts, pour les deux tranches d'âge.

Après lissage des taux, on obtient la table 3.20.

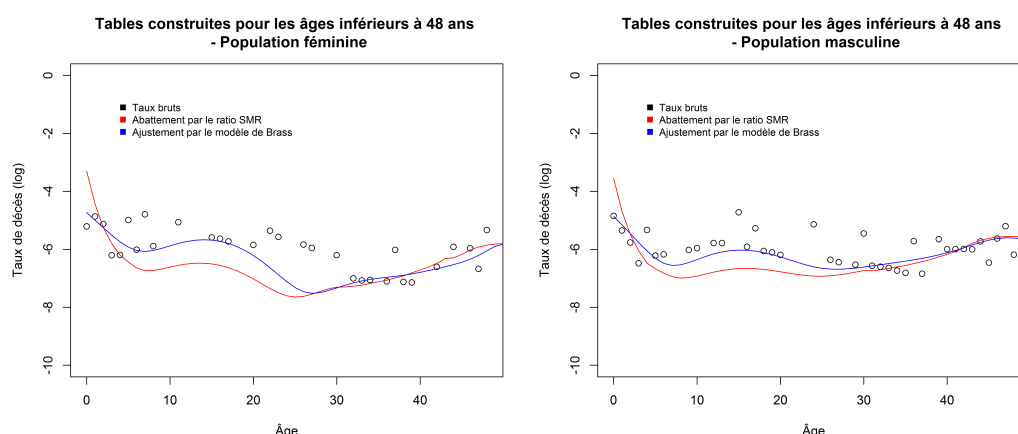
Les taux lissés ne sous-estiment pas les taux autour de 18, contrairement à la méthode du ratio SMR. Cette méthode semble donc plus appropriée pour l'estimation des taux avant 49 ans pour la population masculine.



GRAPHIQUE 3.20 – Taux ajustés par le modèle de Brass et lissés pour la population masculine

3.6.3 Validation des ajustements

Le graphique 3.21 compare les taux ajustés par les deux méthodes du ratio SMR et de Brass, pour chaque population étudiée.

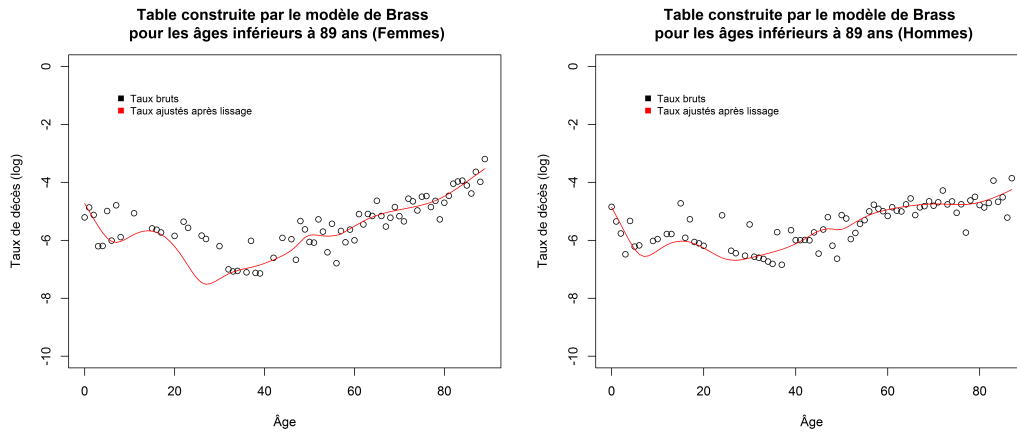


GRAPHIQUE 3.21 – Comparaison des taux ajustés pour les âges inférieurs à 48 ans

Pour les deux populations, c'est le modèle de Brass qui réalise le meilleur ajustement. On observe une surestimation de la mortalité à la naissance par la méthode du ratio SMR, qui disparaît avec le modèle de Brass. Par ailleurs, les taux de décès sont beaucoup moins sous-estimés entre 5 et 25 ans par le modèle de Brass que par la méthode du ratio SMR.

Nous choisissons donc le modèle de Brass pour l'ajustement de la mortalité des populations féminines et masculines pour les moins de 48 ans.

Les tables obtenues après lissage autour de l'âge de raccord sont présentées dans le graphique 3.22.



GRAPHIQUE 3.22 – Tables obtenues par sexe pour les âges inférieurs à 89-87 ans

Il ne reste désormais plus qu'à clore les deux tables, en extrapolant les taux pour les âges élevés.

3.7 Traitement des âges élevés

À cause du manque de données aux âges élevés, il est nécessaire d'extrapoler les taux de décès. En effet, la base ne comprend que très peu de décès après 90 ans, et il n'y a plus aucune donnée au-delà de 105 ans.

Nous allons donc comparer quelques méthodes de fermeture et choisir la plus appropriée.

3.7.1 Fermeture de Coale & Kisker

Coale et Kisker [1] [2] ont remarqué que les courbes des taux de croissance de la mortalité (notés g_x à l'âge x) présentent un pic à 80 ans, avant de décroître linéairement. Ils ont donc posé l'équation suivante :

$$g_x = g_{80} + s(x - 80) \text{ pour } x > 80$$

Les auteurs supposent ensuite que les taux instantanés de mortalité prennent pour valeur 1 à l'âge de fermeture de 110 ans pour les hommes, et valent 0.8 pour les femmes au même âge.

$$\mu_{110}^{\text{hommes}} = 1 \text{ et } \mu_{110}^{\text{femmes}} = 0.8$$

On peut alors déduire la valeur de s de l'équation suivante :

$$\mu_{110} = \mu_{79} \exp \left(\sum_{x=80}^{110} [g_{80} + s(x - 80)] \right)$$

$$\text{soit } s = - \frac{\ln \left(\frac{\mu_{79}}{\mu_{110}} \right) + 31g_{80}}{465}$$

Si les données ne sont pas disponibles ou peu fiables pour déterminer g_{80} , les auteurs proposent également de partir du postulat que $\mu_x = \mu_{65} \exp(g_x(x - 65))$ pour les âges supérieurs à 65 ans. Cela permet de calculer la valeur de g_{80} :

$$g_{80} = \frac{1}{15} \ln \left(\frac{\mu_{80}}{\mu_{65}} \right)$$

On peut ensuite calculer les taux instantanés de mortalité pour les âges compris entre 80 et 110 ans :

$$\mu_x = \mu_{79} \exp \left(\sum_{y=80}^x [g_{80} + s(y - 80)] \right)$$

Pour adapter le modèle à nos données et à nos choix, nous prendrons un âge de fermeture de 120 ans, avec l'hypothèse que $\mu_{120} = 1$.

La relation qui lie les taux instantanés de mortalité aux taux de décès est la suivante :

$$q_x = 1 - \exp(-\mu_x)$$

Avant d'appliquer le modèle d'extrapolation de Coale & Kisker à nos données, nous devons vérifier que les hypothèses du modèle sont bien vérifiées.

Tout d'abord, nous devons vérifier que les taux de croissance de la mortalité g_x présentent effectivement un pic autour de 80 ans.

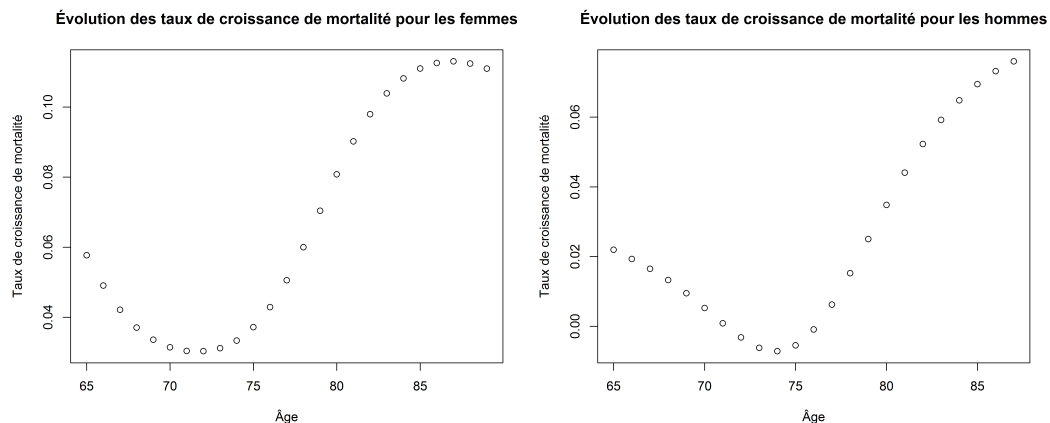
Pour cela, nous commençons par calculer les taux instantanés de mortalité grâce à la relation

$$\mu_x = -\ln(1 - q_x)$$

On peut ensuite calculer les taux de croissance de la mortalité, qui se déduisent de la manière suivante des μ_x :

$$g_x = \ln \left(\frac{\mu_x}{\mu_{x-1}} \right)$$

On obtient alors l'évolution suivante (Graphique 3.23) de ces taux pour les deux populations.



GRAPHIQUE 3.23 – Évolution par sexe des taux de croissance de mortalité

Sur nos données, le pic semble plutôt survenir à 87 ans pour la population féminine. Concernant la population masculine, nous ne pouvons pas être certains de

l'évolution des taux de croissance au-delà de 87 ans, mais nous allons également supposer que les taux atteignent leur valeur maximale à l'âge de 87 ans. Cela permettra de simplifier les calculs par la suite, et n'induit aucune différence notable dans les résultats.

On prend donc plutôt comme hypothèse que $g_x = g_{87} + s(x - 87)$ pour $x > 87$.
On obtient alors : $s = -\frac{\ln\left(\frac{\mu_{86}}{\mu_{120}}\right) + 34g_{87}}{561}$.

Ensuite, comme nos données étaient suffisantes pour étudier la mortalité jusqu'à 87 ans, nous pouvons utiliser la valeur de g_{87} calculée directement, sans ajouter d'hypothèses. On utilise donc la valeur $g_{87} = 0.113$ pour les femmes, et $g_{87} = 0.076$ pour les hommes. On en déduit alors :

$$\begin{cases} s = 1.3910^{-5} & \text{pour les femmes} \\ s = 3.0910^{-3} & \text{pour les hommes} \end{cases}$$

On peut alors en déduire les valeurs des taux instantanés de mortalité pour les âges supérieurs à 87 ans :

$$\mu_x = \mu_{86} \exp\left(\sum_{y=87}^x [g_{87} + s(y - 87)]\right)$$

On obtient alors les tables de mortalité finale présentées dans le graphique 3.24.

Les tables extrapolées par la méthode de Coale & Kisker semblent bien estimer la mortalité au-delà de 90 ans, et sont plutôt proches des taux de décès observés avant 102 ans.

Cependant, cette méthode ne permet pas d'obtenir un taux de mortalité de 1 à l'âge de fermeture de la table.

3.7.2 Fermeture par régression exponentielle

La méthode d'extrapolation des taux de décès par régression exponentielle consiste simplement à supposer que les $\hat{q}_x$ vérifient la relation $\hat{q}_x = ae^{bx}$ après un certain âge de raccord x_0 .

Pour ajuster les paramètres a et b , on suppose que :

$$\begin{cases} \hat{q}_{x_0} = q_{x_0} \\ \hat{q}_{120} = 1 \end{cases}$$

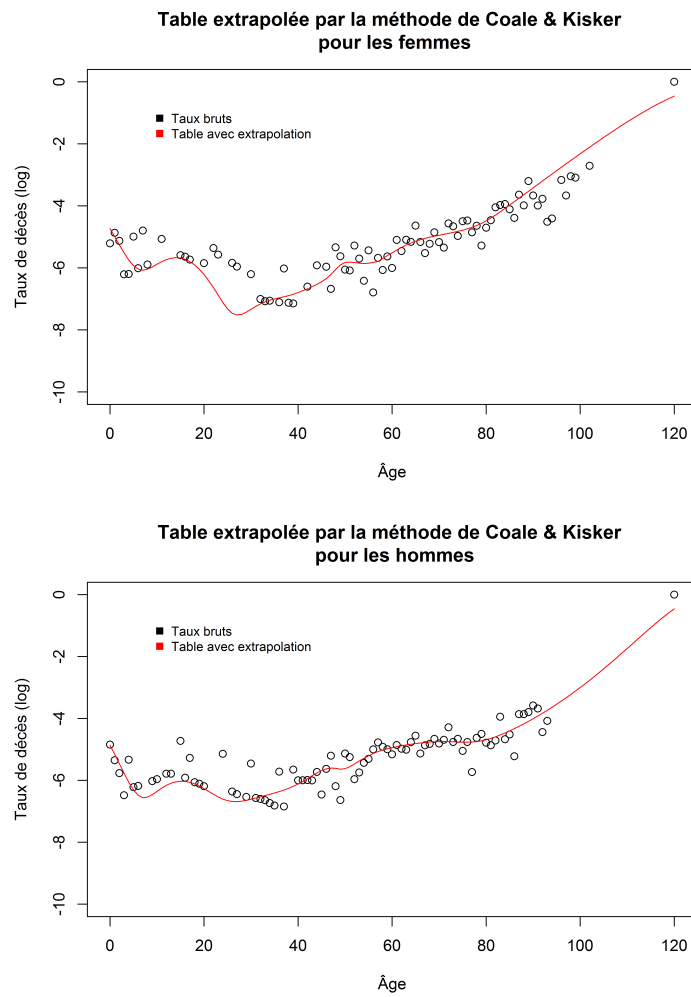
Population féminine

Pour nos données, nous supposons de nouveau que l'extrapolation se fait à partir de 90 ans.

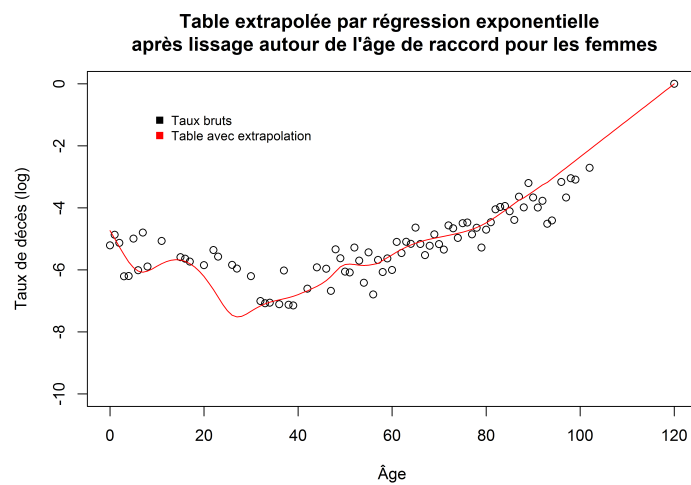
On obtient alors

$$\begin{cases} a = 7.40 \times 10^{-7} \\ b = 0.118 \end{cases}$$

Les taux présentent là aussi un léger saut entre 89 et 90 ans, qui est facilement lissé. Les taux finaux sont présentés dans le graphique 3.25.



GRAPHIQUE 3.24 – Tables extrapolées par sexe par la méthode de Coale & Kisker



GRAPHIQUE 3.25 – Table féminine extrapolée par régression exponentielle

Avec cette méthode, les $\ln(\hat{q}_x)$ sont linéaires après 91 ans. De la même manière que la méthode de Coale & Kisker, la table extrapolée semble bien estimer la mortalité observée au-delà de 91 ans, mais présente en plus l'avantage d'avoir un taux de décès qui vaut 1 à l'âge de fermeture.

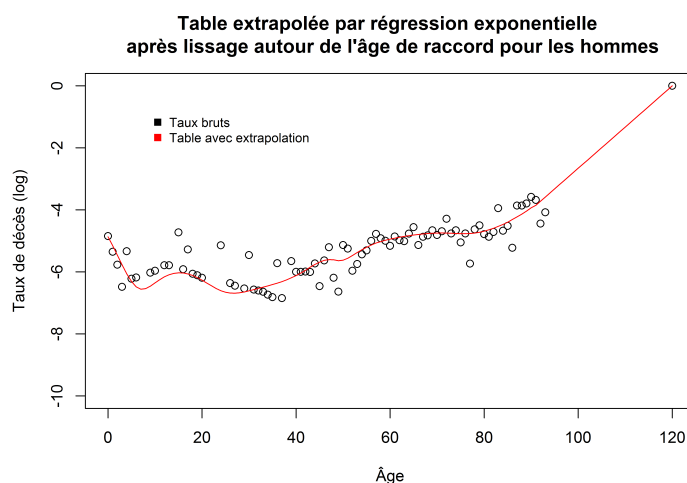
Population masculine

Concernant la population masculine, nous supposons de nouveau que l'extrapolation se fait à partir de 88 ans.

On obtient alors

$$\begin{cases} a = 1.20 \times 10^{-7} \\ b = 0.133 \end{cases}$$

Les taux obtenus après lissage autour de l'âge de raccord sont présentés dans le graphique 3.26.



GRAPHIQUE 3.26 – Table masculine extrapolée par régression exponentielle

3.7.3 Validation des fermetures et choix de la méthode

Les méthodes de Coale & Kisker et de régression exponentielle semblent toutes les deux adaptées à nos données car proches de la mortalité observée entre 90 et 102 ans environ. C'est cependant cette dernière que nous allons conserver car elle vérifie par construction $\hat{q}_{120} = 1$.

D'autres méthodes d'extrapolation ont été testées, notamment celle de Denuit & Goderniaux [2], cependant les résultats obtenus n'étaient pas satisfaisants pour nos données, et les méthodes n'ont pas été conservées. En effet, du fait de la contrainte de concavité et de tangente horizontale à l'âge de fermeture, cette méthode surestimait rapidement les taux bruts au-delà de 90 ans.

3.8 Conclusion

Les tables finales obtenues ont été créées différemment selon les âges étudiés. Pour les âges jeunes, moins de 49 ans, la table est issue d'un ajustement par le modèle de Brass des tables de référence TGF05 et TFH05, selon la population étudiée. Pour les âges compris entre 49 ans et 89 ans pour les femmes, et 49 et 87 ans pour les hommes, les taux de Kaplan-Meier ont été lissés par la méthode de Whittaker-Henderson. Enfin, les taux pour les âges élevés sont extrapolés par une régression exponentielle entre 90 et 120 ans pour la population féminine, et entre 88 et 120 pour la population masculine.

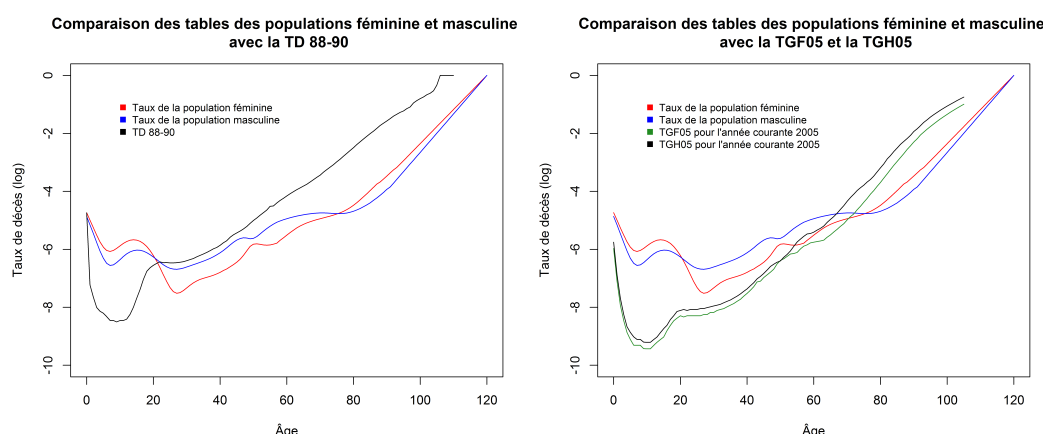
En comparant les tables finales de la population féminine et de la population masculine entre elles (Graphique 3.27), on remarque qu'elles se croisent plusieurs fois.

Entre 0 et 18 ans environ, la population féminine présente une surmortalité par rapport à la population masculine. Cela correspond à ce que nous avons remarqué sur les données brutes de décès et d'exposition.

En effet, l'exposition pour les femmes était plus faible en dessous de 18 ans, pour un nombre équivalent de décès.

La tendance s'inverse ensuite entre 18 et 75 ans environ, où les hommes ont une mortalité supérieure, puis les courbes des taux de mortalité se croisent de nouveau au-delà. Nous avons effectivement remarqué qu'après ces âges, l'exposition des femmes était de nouveau plus faible, pour un nombre de décès équivalent.

Le comportement de ces courbes n'est pas habituel, puisque généralement les tables de la population féminine présentent une sous-mortalité par rapport à celles des hommes à tous les âges, car elles ont une espérance de vie plus élevée. Cependant, en prenant en compte les intervalles de confiance, nous ne pouvons pas être certains que ces tables se croisent aux points indiqués, car l'écart à chaque âge entre les tables est inférieur à l'incertitude des méthodes, pour les moins de 20 ans.



GRAPHIQUE 3.27 – Comparaison des tables construites par sexe

Toujours est-il que la table réglementaire TD 88-90 surestime la mortalité des deux populations à partir de 18 ans environ, tandis que les tables TGF05 et TGH05 ne la surestime qu'à partir de 70 ans environ. Cependant, il faut garder en tête que les deux populations présentes ici ne sont pas les rentiers de Sham, mais l'ensemble des personnes victimes d'accidents en RC médicale. *A priori*, les rentiers présentent

une mortalité supérieure à la population totale. Donc nous ne pouvons pas conclure sur le caractère approprié de la TD 88-90 pour le calcul des provisions pour les rentes dans le portefeuille de Sham, cependant nous pouvons supposer que la TGF05 et la TGH05 seront *a priori* prudentes.

Maintenant que nous avons réussi à estimer la mortalité des personnes victimes d'accidents médicaux, nous devons réussir à nous rapprocher de la mortalité réelle des rentiers. En effet, beaucoup d'accidents médicaux sont « mineurs » comparés à la gravité des accidents subis par les rentiers, et ne devraient pas être pris en compte. Cela conduit à une sous-estimation importante de la mortalité réelle des rentiers et ne nous permet pas de conclure directement quant à la prudence de la TD 88-90 dans le calcul des provisions. Nous avons en effet obtenus des tables féminine et masculine qui présentent une sous-mortalité par rapport à cette table de référence. Cela semble être moins le cas pour les tables TGF05 et TGH05 présentées par année courante, dont on pourrait conclure directement qu'elles sont prudentes au moins jusqu'à 70 ans.

Comme la suite de l'étude sera réalisée sur une population plus restreinte, nous n'aurons plus la possibilité de réaliser une étude par sexe. Pour permettre une meilleure comparaison des résultats futurs avec l'étude que nous venons de réaliser, nous avons donc également construit une table de mortalité pour la population totale, sans distinction par sexe. Les méthodes utilisées lors de la construction sont identiques à celles retenues pour les populations féminine et masculine, à savoir :

- Pour les tranches d'âge 0-17 ans et 18-45 ans : modèle de Brass
- Pour la tranche d'âge 46-90 ans : taux bruts de Kaplan-Meier lissés par la méthode de Whittaker-Henderson
- Pour la tranche d'âge 91-120 ans : extrapolation par régression exponentielle.

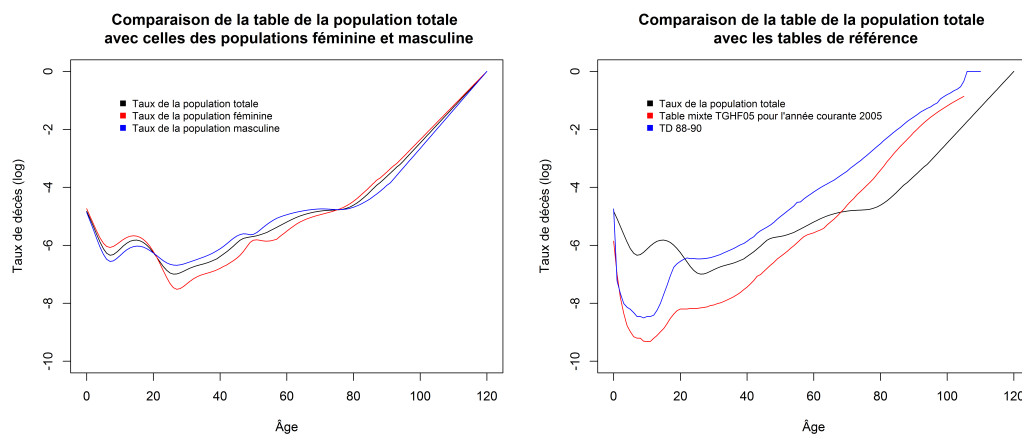
Les résultats obtenus avec ces différentes méthodes sont résumés dans la tableau ci-dessous.

Étapes	Résultats pour l'ensemble de la base de données
Critère de Cochran	Vérifié entre 46 et 90 ans Pas de lissage des décès
Whittaker-Henderson	Paramètres : $z = 3$ et $h = 10^7$ Critère de fidélité aux taux bruts : 2.65×10^{-4} Critère de régularité des taux lissés : 3.93×10^{-5} Statistique du Khi-deux : 26.43 (Quantile à 44 ddl : 60.48)
Modèle de Brass	Paramètres pour les 0-17 ans : $a = 13.26$ et $b = 3.15$ Paramètres pour les 18 - 45 ans : $a = -0.38$ et $b = 0.47$
Régression exponentielle	Paramètres : $a = 3.68 \times 10^{-7}$ et $b = 0.123$

TABLE 3.5 – Étapes de la construction de la table de mortalité mixte

La table obtenue est comparée avec la TD 88-90 et la moyenne simple des tables TGH05 et TGF05, notée table mixte TGHF05. Cela est rendu possible grâce au Sex Ratio de 47.1% au sein de la population. Autrement il aurait été possible de pondérer la moyenne par le sexe ratio.

Cette table est présentée dans le graphique 3.28.



GRAPHIQUE 3.28 – Table mixte construite pour la population totale

Cette table se situe entre les tables féminine et masculine, et possède donc sensiblement les mêmes comportements vis-à-vis de la TD 88-90 et de la TGHF05. La TD 88-90 sous-estime ainsi la mortalité de la population totale à partir de 20 ans environ, tandis que la TGHF05 est prudente jusqu'à 67 ans environ.

Nous pouvons désormais réduire la population étudiée selon certains critères, et étudier la mortalité de la population ainsi obtenue par comparaison à celle de la population totale ci-dessus.

Chapitre 4

Approximation de la mortalité des rentiers

Dans la partie précédente, nous avons établi une table de mortalité par âge et par sexe pour l'ensemble de la population des personnes victimes d'accidents médicaux présentes entre 2004 et 2015, avec et sans distinction par sexe.

Dans le cadre de cette nouvelle étude, nous allons essayer de réduire la population étudiée afin de mieux approcher la mortalité des rentiers. Du fait du faible nombre de données, nous n'allons plus réaliser une étude par sexe

Afin de réduire la population étudiée, nous avons à notre disposition des informations sur les coûts des dossiers, les taux d'IPP des victimes, et les spécialités mises en cause lors des accidents.

Dans cette partie, nous allons donc essayer de réduire la population selon certains critères fondés sur ces variables. Des tables de mortalité seront alors créées, soit en appliquant la méthodologie de la partie précédente si les données sont suffisantes, soit en effectuant un ajustement de la table de mortalité créée pour la population totale à l'aide du modèle de Brass appliqué à tous les âges, lorsque le nombre de données est trop faible selon le critère de Cochran pour une plage d'âge suffisamment importante.

4.1 Mortalité par coûts

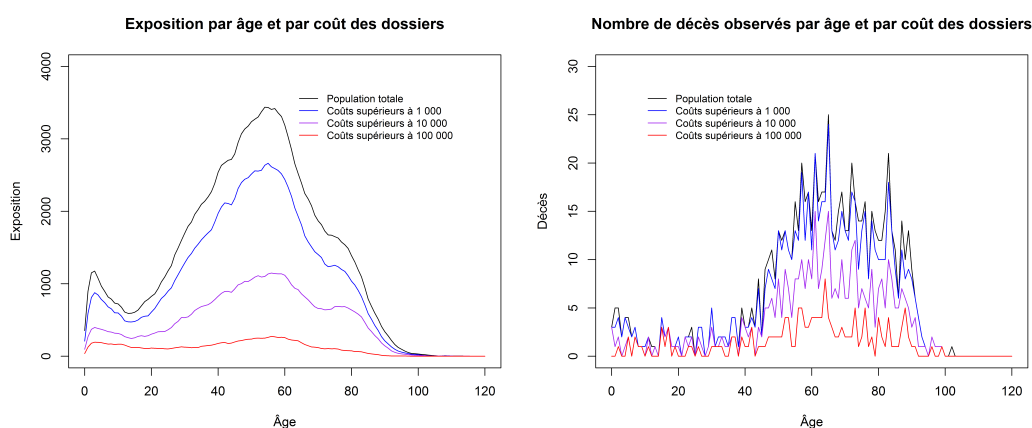
Afin de réduire le nombre de dossiers étudiés, une première idée est d'utiliser comme critère les coûts totaux des dossiers.

En effet, nous avons vu que le coût médian des dossiers de la population totale est de 2 336€, tandis que le coût moyen est de 38 961€. En revanche, le coût médian pour la population en rente s'élève à environ 1.1 million d'euros. Ce montant élevé est dû aux importantes provisions qui sont passées

Il est donc possible de réduire la population étudiée en considérant les dossiers ayant des coûts totaux supérieurs à un seuil de plus en plus élevé.

Au vu des distributions précédentes pour les coûts totaux des dossiers, nous allons appliquer trois critères :

- Dossiers ayant un coût total supérieur à 1 000 €
- Dossiers ayant un coût total supérieur à 10 000 €
- Dossiers ayant un coût total supérieur à 100 000 €



GRAPHIQUE 4.1 – Étude de la mortalité par coûts - Exposition et nombre de décès

D'après le graphique 4.1, l'exposition diminue avec les coûts plus rapidement que le nombre de décès. En effet, les décès ont en moyenne un coût trois fois plus élevé que les autres dossiers.

La portion de la population totale conservée pour chaque seuil est présentée dans le tableau 4.1.

	Proportion en nombre de la population totale conservée	Proportion en nombre des décès totaux conservés	Proportion en nombre des rentiers conservés	Proportion en coût des rentiers conservés
Coûts supérieurs à 1 000€	69,30%	89.70%	100%	100%
Coûts supérieurs à 10 000€	27.20%	53.90%	99.40%	100.00%
Coûts supérieurs à 100 000€	5.30%	21.30%	97.90%	99.96%

TABLE 4.1 – Étude de la mortalité par coûts - Proportion conservée pour la population, les décès et les rentiers

On remarque ici que plus le seuil est élevé, plus la mortalité est importante, mais aussi plus l'estimation de la mortalité des rentiers est juste.

Nous allons donc étudier la mortalité pour chacune des populations construites à partir de ces seuils.

4.1.1 Dossiers avec un coût total supérieur à 1 000 €

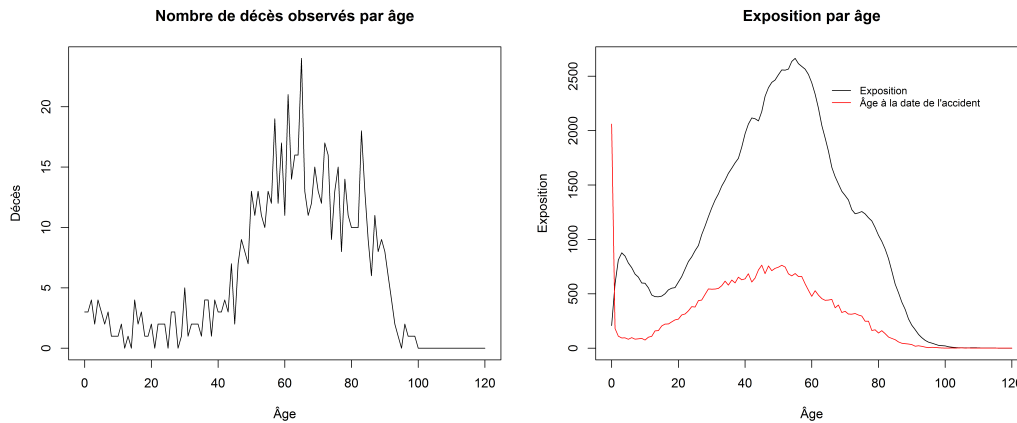
La répartition des décès et de l'exposition pour les dossiers ayant des coûts supérieurs à 1 000€ est présentée dans le graphique 4.2.

Globalement, la répartition par âge est semblable à celle de la population totale. Mais du fait de la proportion des décès légèrement supérieure, nous nous attendons à obtenir des taux de décès légèrement supérieurs.

Cela est confirmé par le ratio SMR, calculé par rapport à la table de mortalité de la population totale, et qui vaut 1.124. Ce ratio a été calculé par rapport à la mortalité de la population totale et non par rapport à celle de la TGHF05 car il y a de fortes différences entre les deux, et la table que nous avons construite est

davantage susceptible d'être proche de la mortalité de nos populations réduites, et plus particulièrement de la mortalité des rentiers.

Par ailleurs, on remarque que cette population contient des personnes ayant été victimes d'accidents médicaux avant 1 an ou entre 15 et 80 ans environ.



GRAPHIQUE 4.2 – Dossiers ayant un coût supérieur à 1 000 €- Exposition et nombre de décès

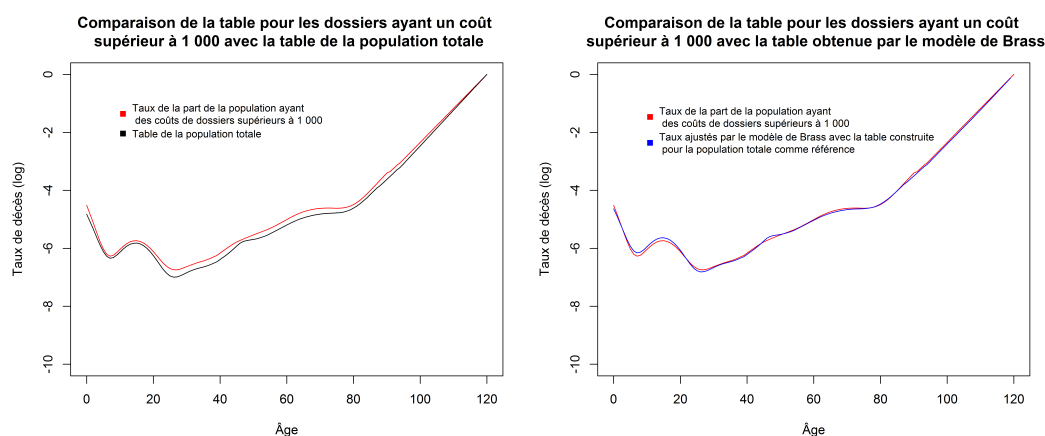
Les résultats obtenus lors de la construction de la table sont résumés dans le tableau 4.2.

Étapes	Résultats pour l'ensemble de la base de données
Critère de Cochran	Vérifié entre 46 et 90 ans Pas de lissage des décès
Whittaker-Henderson	Paramètres : $z = 3$ et $h = 10^7$ Critère de fidélité aux taux bruts : 3.14×10^{-4} Critère de régularité des taux lissés : 6.07×10^{-5} Statistique du Khi-deux : 26.84 (Quantile à 44 ddl : 60.48)
Modèle de Brass	Paramètres pour les 0-17 ans : $a = 11.77$ et $b = 2.83$ Paramètres pour les 18 - 45 ans : $a = -0.13$ et $b = 0.48$
Régression exponentielle	Paramètres : $a = 7.70 \times 10^{-7}$ et $b = 0.48$

TABLE 4.2 – Étapes de la construction de la table de mortalité pour les dossiers ayant un coût supérieur à 1 000€

La table de mortalité obtenue est présentée dans le graphique 4.3. On remarque que la mortalité de la population réduite est effectivement légèrement plus élevée que celle de la population, notamment pour les âges compris entre 18 et 80 ans.

La table construite est également comparée à la table que l'on aurait obtenue si nous avions choisi d'appliquer le modèle de Brass pour tous les âges en prenant comme table de référence la table construite pour la population totale. Les deux tables semblent être quasiment identiques. Le modèle de Brass surestime très légèrement la mortalité entre 5 et 20 ans.



GRAPHIQUE 4.3 – Dossiers ayant un coût supérieur à 1 000 €- Comparaison des méthodes

Cependant, la courbe n'est pas suffisamment éloignée de la nouvelle table de référence (i.e. la table créée pour la population totale), et la TD 88-90 n'est donc pas suffisamment prudente avec ce critère.

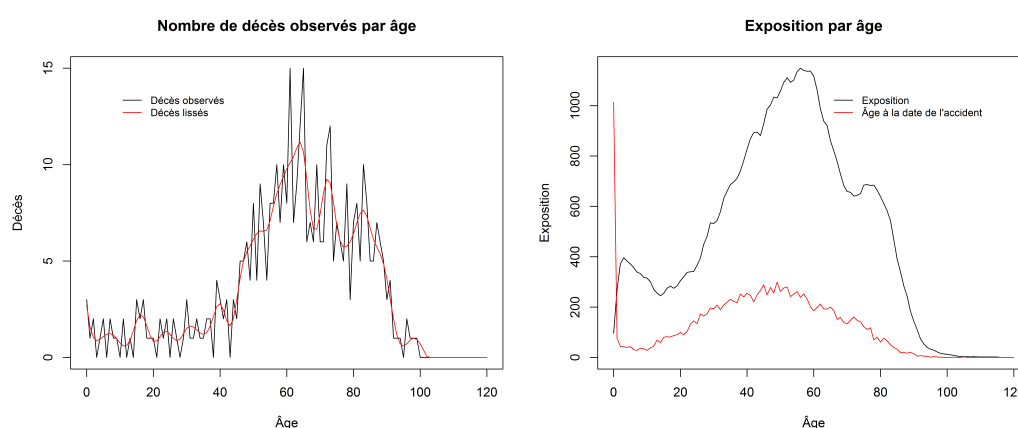
4.1.2 Dossiers avec un coût total supérieur à 10 000 €

Afin d'obtenir une table de mortalité plus éloignée de la table de référence, nous augmentons le seuil pour les coûts des dossiers étudiés à 10 000€. Le ratio SMR calculé par rapport à cette table est de 1.317 : il est supérieur à celui calculé pour le seuil de 1 000€. Nous nous attendons donc à avoir une mortalité encore plus

éloignée de la table de référence, *a priori* sur la même tranche d'âge que pour le seuil de 1 000€, à savoir entre 18 et 80 ans.

La répartition des décès, de l'exposition et du nombre d'accidents en fonction de l'âge est présentée dans le graphique 4.4. Les décès deviennent plus variables, et sont donc lissés afin de respecter le critère de Cochran sur une plage d'âge suffisamment importante.

Encore une fois, un nombre important d'accidents ont lieu à la naissance, puis le reste des accidents est réparti entre 20 et 80 ans.



GRAPHIQUE 4.4 – Dossiers ayant un coût supérieur à 10 000 €- Exposition et nombre de décès

Les résultats obtenus lors de la construction de la table sont résumés dans le tableau 4.3.

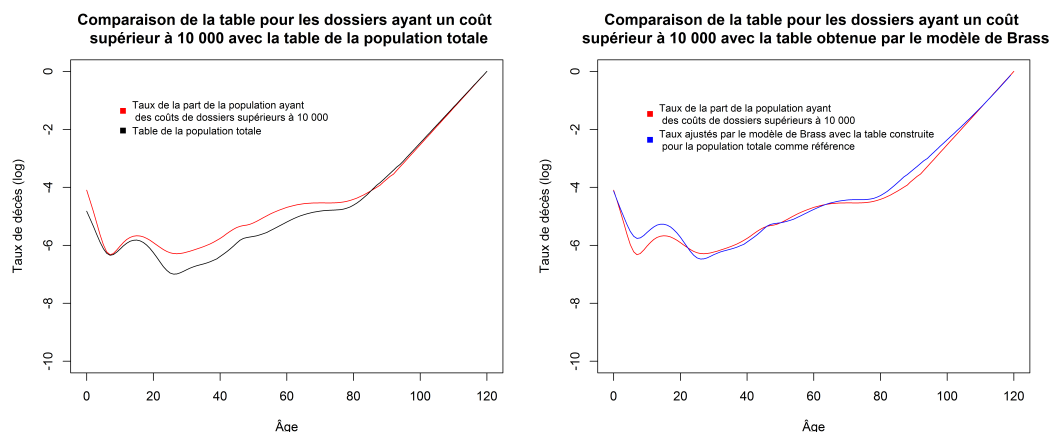
Étapes	Résultats pour l'ensemble de la base de données
Critère de Cochran	Vérifié entre 48 et 88 ans Un lissage des décès a été réalisé
Whittaker-Henderson	Paramètres : $z = 3$ et $h = 10^7$ Critère de fidélité aux taux bruts : 4.90×10^{-4} Critère de régularité des taux lissés : 1.06×10^{-5} Statistique du Khi-deux : 31.43 (Quantile à 40 ddl : 55.76)
Modèle de Brass	Paramètres pour les 0-17 ans : $a = 8.91$ et $b = 2.23$ Paramètres pour les 18 - 45 ans : $a = 0.37$ et $b = 0.50$
Régression exponentielle	Paramètres : $a = 2.55 \times 10^{-7}$ et $b = 0.127$

TABLE 4.3 – Étapes de la construction de la table de mortalité pour les dossiers ayant un coût supérieur à 10 000€

La table obtenue est présentée dans le graphique 4.5. L'écart avec la mortalité de la population totale est effectivement plus important entre 18 et 80 ans, mais également à la naissance.

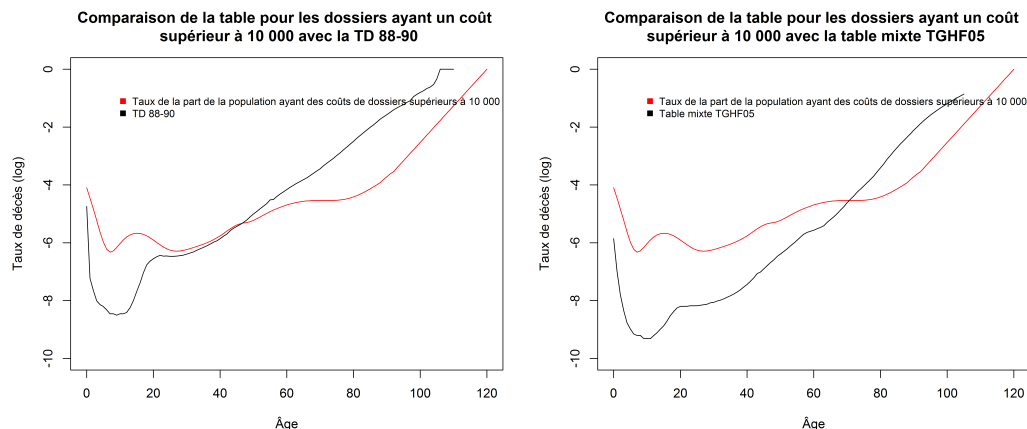
La table construite est également comparée à la table que l'on aurait obtenue si nous avions choisi d'appliquer le modèle de Brass pour tous les âges en prenant comme table de référence la table construite pour la population totale. Les deux

tables semblent être quasiment identiques entre 30 et 75 ans. Au-delà de cet intervalle, le modèle de Brass surestime globalement la mortalité, sauf avant 5 ans où elle semble bien estimée.



GRAPHIQUE 4.5 – Dossiers ayant un coût supérieur à 10 000 €- Comparaison des méthodes

D'après le graphique 4.6 de comparaison avec les tables réglementaires, la TD 88-90 est cette fois-ci prudente jusqu'à 48 ans environ, mais c'est *a priori* toujours la TGHF05 qui est la table la plus prudente pour le calcul de provisions.

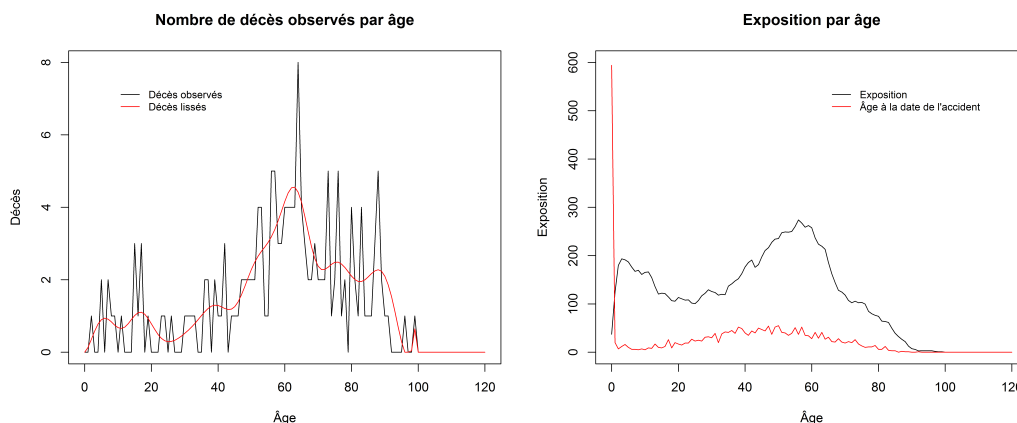


GRAPHIQUE 4.6 – Dossiers ayant un coût supérieur à 10 000 €- Comparaison avec les tables réglementaires

4.1.3 Dossiers avec un coût total supérieur à 100 000 €

Pour les dossiers présentant des coûts supérieurs à 100 000€, la population est fortement réduite. Le nombre de personnes présentes (2 787) et le nombre de décès (161) sont dans ce cas beaucoup trop faibles pour construire une table de mortalité, même en lissant les décès (Graphique 4.7). Nous allons donc ajuster les taux bruts de Hoem à l'aide du modèle de Brass en prenant comme table de référence la table créée pour la population totale.

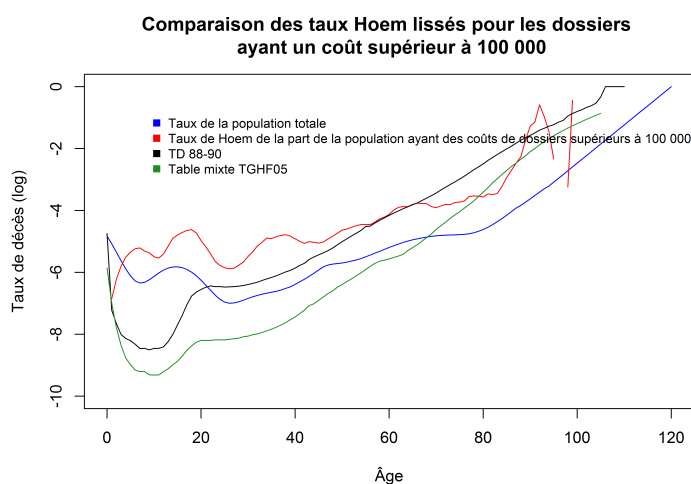
Le ratio SMR pour cette population, calculé par rapport à la table de la population totale, est de 2.904. Pour ce seuil de 100 000€, la mortalité est donc fortement supérieure à celle de la population de référence.



GRAPHIQUE 4.7 – Dossiers ayant un coût supérieur à 100 000 €- Exposition et nombre de décès

Encore une fois, un nombre très important des accidents (environ 21%) a eu lieu à la naissance.

Les taux bruts de Hoem après lissage des décès sont présentés dans le graphique 4.8, cependant les intervalles de confiance ne sont pas significatifs et valent régulièrement 0 pour la borne inférieure. Ces taux nous donc permettent juste de supposer que la mortalité de cette population est effectivement très supérieure à celles de toutes les populations précédentes, et que cela nous permettrait peut-être de considérer que la TD 88-90 est prudente. Nous allons donc appliquer le modèle de Brass pour confirmer cette intuition.

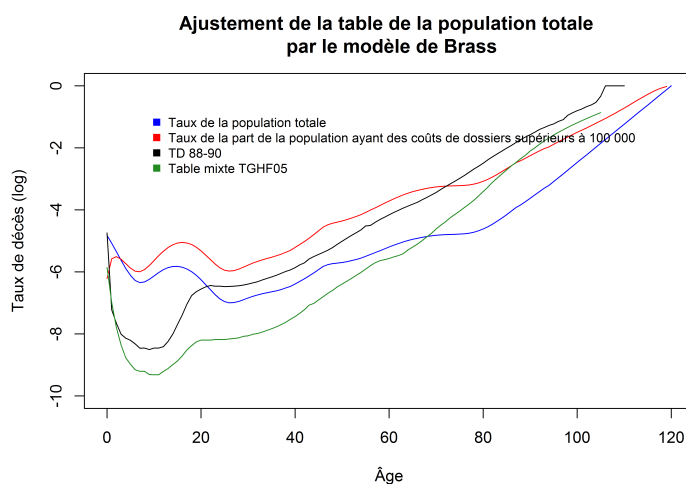


GRAPHIQUE 4.8 – Dossiers ayant un coût supérieur à 100 000 €- Taux de Hoem

En réalisant un ajustement de la table de mortalité de la population totale à l'aide des données d'exposition et de décès pour la population dont les coûts totaux

des dossiers sont supérieurs à 100 000€, on obtient les résultats présentés dans le graphique 4.9.

Comme observé dans les deux cas précédents, on peut supposer que la mortalité est surestimée au-delà d'un certain âge et pour les âges jeunes supérieurs à 5 ans. La sous-mortalité par rapport à la table construite pour la population totale que l'on observe pour les moins de 5 ans est donc *a priori* une estimation correcte de la mortalité de cette population.



GRAPHIQUE 4.9 – Dossiers ayant un coût supérieur à 1 000 €- Comparaison avec les tables réglementaires

Avec le modèle de Brass, on trouve que la TD 88-90 est prudente jusqu'à 75 ans environ. On remarque cependant que la mortalité avant l'âge de 2 ans environ est plus faible de celle de la TD 88-90. Cela est dû au fait que dans nos données brutes, il n'y a eu qu'un seul décès parmi la tranche d'âge 0-4 ans. Les valeurs estimées des paramètres du modèle de Brass sont les suivantes :

$$\begin{cases} a = 2.653 \\ b = 1.840 \end{cases}$$

Une valeur du paramètre b supérieure à 1 implique une mortalité plus élevée aux grands âges (relativement à la population de référence), et plus faible aux âges jeunes. Le paramètre a est quant à lui un indicateur du niveau de mortalité qui affecte tous les âges de manière égale.

Pour cette population, nous avons donc une mortalité supérieure à celle de la population totale, et plus élevée aux grands âges relativement à cette population de référence.

Cependant, au vu de la faible contribution des âges élevés, il est légitime de se demander si les provisions ne sont dès lors pas suffisamment prudentes avec cette table.

4.2 Mortalité par IPP

Un autre critère possible pour réduire la population étudiée est le taux d'Incapacité Permanente Partielle (IPP). En effet, seulement la moitié de la population a un taux d'IPP renseigné, et on peut supposer que cela correspond à un accident grave. Il est effectivement possible que certaines victimes d'accidents médicaux se retrouvent en incapacité ou en invalidité suite à l'accident. Ce critère de sélection permettra peut-être d'obtenir une mortalité plus élevée.

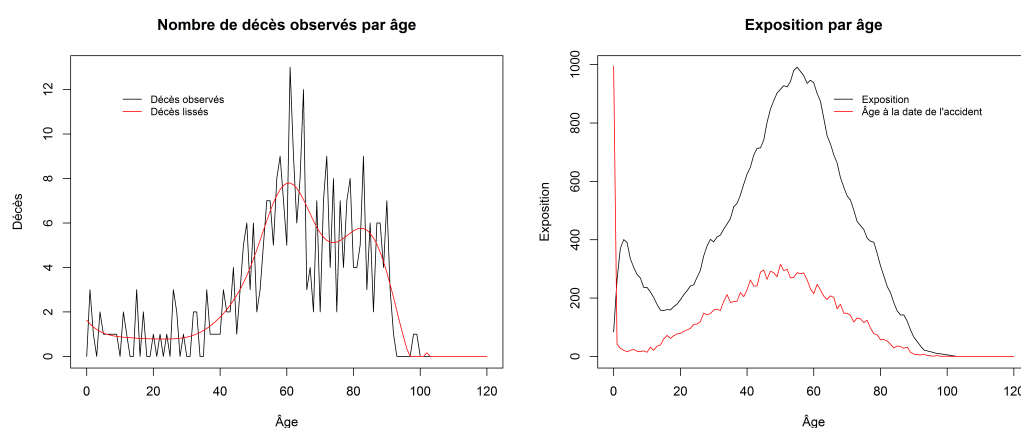
4.2.1 Taux d'IPP supérieur à 8

Lorsqu'il est renseigné, le taux d'IPP parmi les rentiers est au moins égal à 8. Nous choisissons donc comme premier critère de conserver les personnes ayant un taux d'IPP renseigné supérieur à 8. Par ailleurs, beaucoup de personnes ont un taux d'IPP faible et proche de 0, et ce critère permet de les retirer.

La proportion conservée, présentée dans le tableau 4.4, est très faible. En effet, seulement la moitié de la population totale a un taux d'IPP renseigné, et parmi ceux-là seulement la moitié à un taux d'IPP supérieur à 8. Cependant, la proportion des décès conservés est élevée, plus de 40%, et cela permet d'affirmer que la mortalité de la population étudiée ici est supérieure à celle de la population totale. Le ratio SMR calculé par rapport à la population totale vaut en effet 1.516.

Proportion en nombre de la population totale conservée	Proportion en nombre des décès totaux conservés	Proportion en nombre des rentiers conservés	Proportion en coût des rentiers conservés
26.2%	41.9%	65.6%	66.3%

TABLE 4.4 – Taux d'IPP supérieur à 8 - Proportion conservée pour la population, les décès et les rentiers



GRAPHIQUE 4.10 – Taux d'IPP supérieur à 8 - Exposition et nombre de décès

Le graphique 4.10 représente la répartition des décès, de l'exposition et du nombre d'accidents par âge. Les répartitions suivent de nouveau le même schéma que lors de l'étude de la population totale. Les décès sont plus importants entre 45

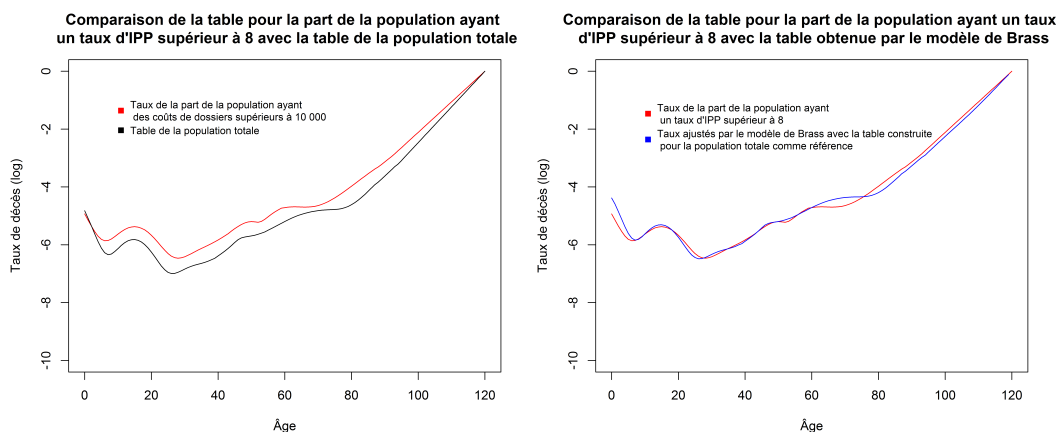
et 90 ans, et l'exposition est forte avant 10 ans et entre 25 et 90 ans. Encore une fois, de nombreux accidents (7.3%) ont eu lieu à la naissance au sein de cette population.

Les résultats obtenus lors de la création de la table sont résumés dans le tableau 4.5.

Étapes	Résultats pour l'ensemble de la base de données
Critère de Cochran	Vérifié entre 52 et 87 ans Un lissage des décès a été réalisé
Whittaker-Henderson	Paramètres : $z = 3$ et $h = 10^7$ Critère de fidélité aux taux bruts : 1.32×10^{-3} Critère de régularité des taux lissés : 4.72×10^{-5} Statistique du Khi-deux : 31.49 (Quantile à 35 ddl : 49.70)
Modèle de Brass	Paramètres pour les 0-17 ans : $a = 20.23$ et $b = 4.46$ Paramètres pour les 18 - 45 ans : $a = 0.31$ et $b = 0.52$
Régression exponentielle	Paramètres : $a = 3.12 \times 10^{-6}$ et $b = 0.106$

TABLE 4.5 – Étapes de la construction de la table de mortalité pour les personnes ayant un taux d'IPP supérieur à 8

On obtient alors la table de mortalité 4.11. Les taux de décès sont cette fois-ci supérieurs à ceux de la population totale pour tous les âges. La table construite est également comparée à la table que l'on aurait obtenue si nous avions choisi d'appliquer le modèle de Brass pour tous les âges en prenant comme table de référence la table construite pour la population totale. Les deux tables semblent être quasiment identiques entre 7 et 60 ans. Pour les moins de 7 ans, le modèle de Brass surestime globalement la mortalité, et on observe une certaine variation au-delà de 60 ans.



GRAPHIQUE 4.11 – Taux d'IPP supérieur à 8 - Comparaison des modèles

Cette table présente à peu près le même comportement entre 20 et 70 ans que la table construite à partir des victimes ayant un coût total de dossier supérieur à 10 000€. Elle présente toutefois une mortalité supérieure à cette table après 70 ans et entre 5 et 20 ans. Avant l'âge de 5 ans, la mortalité rejoint celle de la population totale. De la même manière que la table construite pour les dossiers ayant un coût supérieur à 10 000€, la TD 88-90 est prudente jusqu'à 50 ans environ avec cette nouvelle table.

4.2.2 Taux d'IPP supérieur à 20

La majorité des rentiers ont un taux d'IPP très élevé lorsqu'il est renseigné. Nous allons donc réduire encore plus la population en n'étudiant que les personnes qui ont un taux d'IPP renseigné supérieur à 20.

On ne conserve alors qu'une faible part de la population et des décès totaux. En revanche, près des deux tiers des rentiers sont conservés (Table 4.6).

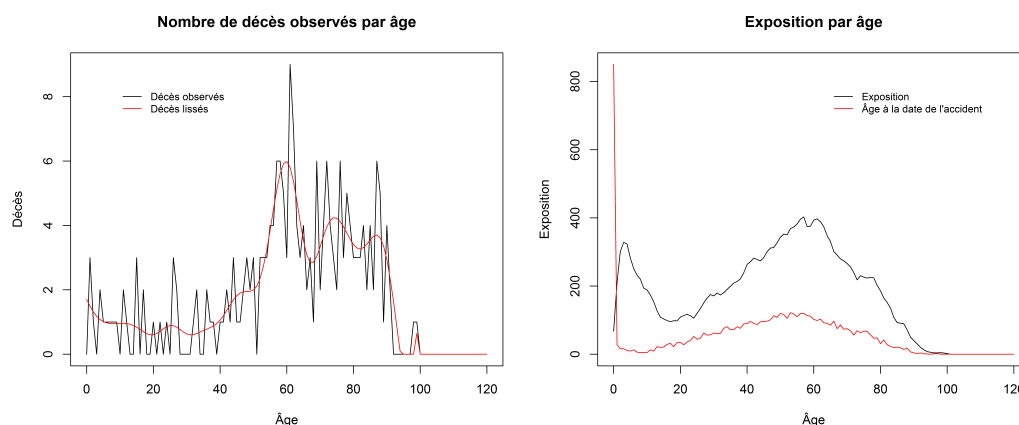
Encore une fois, la proportion de décès est supérieure à celle de la population totale, et le ratio SMR calculé par rapport à la table de la population totale est de 2.070.

Proportion en nombre de la population totale conservée	Proportion en nombre des décès totaux conservés	Proportion en nombre des rentiers conservés	Proportion en coût des rentiers conservés
11.9%	27.8%	63.1%	65.7%

TABLE 4.6 – Taux d'IPP supérieur à 20 - Proportion conservée pour la population, les décès et les rentiers

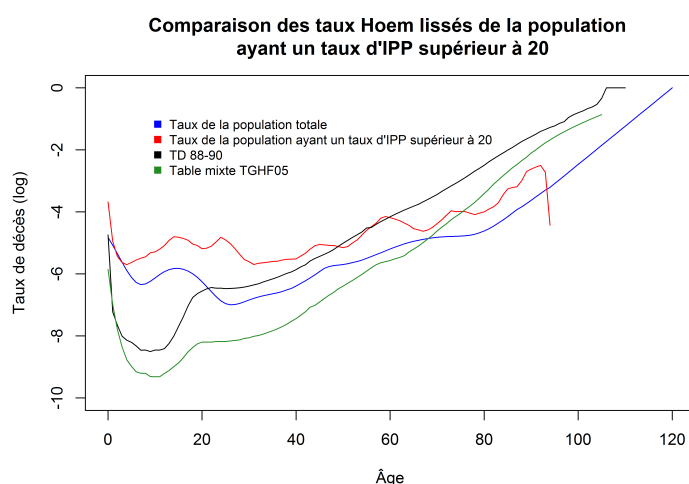
Une faible part de la population est conservée avec ce critère. Il n'est donc pas possible de construire une table de mortalité à partir de ces données. Cependant, il est possible de réaliser un ajustement à partir de la table de la population totale.

Les répartitions présentées dans le graphique 4.12 confirment que le nombre de données est trop faible : le critère de Cochran n'est quasiment jamais respecté, et l'exposition est peu élevée. Le nombre d'accidents est de nouveau très important à la naissance, puisqu'il correspond à plus de 13.6% des accidents survenus au sein de cette population.



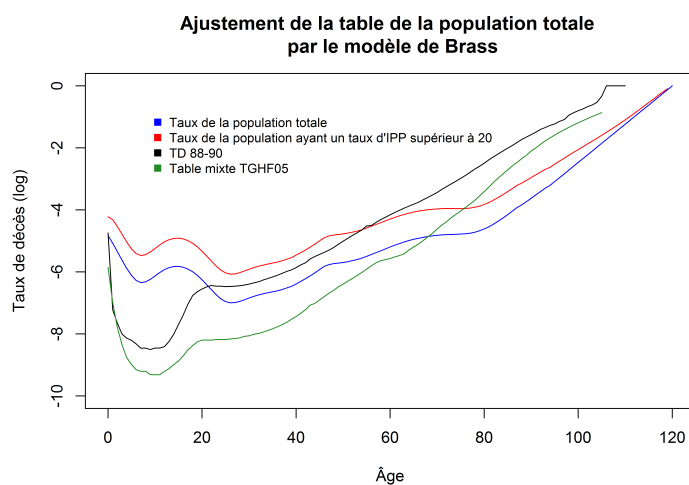
GRAPHIQUE 4.12 – Taux d'IPP supérieur à 20 - Exposition et nombre de décès

Les taux bruts de Hoem après lissage des décès sont présentés dans le graphique 4.13 et donnent une idée de l'évolution de la mortalité avec l'âge. Cependant, les intervalles de confiance ne sont pas significatifs et ne permettent pas de tirer une conclusion.



GRAPHIQUE 4.13 – Taux d'IPP supérieur à 20 - Taux de Hoem

L'application du modèle de Brass par rapport à la population totale indique que la table TD 88-90 est prudente jusqu'à 52 ans environ (Graphique 4.14). Cette fois-ci, la table obtenue ne croise pas la TD 88-90 pour les moins de 2 ans, mais reste bien au-dessus de la table de la population totale.



GRAPHIQUE 4.14 – Taux d'IPP supérieur à 20 - Comparaison avec les tables réglementaires

Les paramètres du modèle de Brass ont été estimés aux valeurs suivantes :

$$\begin{cases} a = 1.237 \\ b = 1.130 \end{cases}$$

Globalement, la mortalité de cette table est inférieure à celle de la table construite pour les dossiers ayant des coûts supérieurs à 100 000€. Le ratio SMR et les valeurs des paramètres du modèle de Brass sont beaucoup plus faibles que pour la table créée avec le seuil de 100 000€.

Cependant, elle a l'avantage d'avoir été créée avec deux fois plus de données, même si le nombre de décès est équivalent.

Néanmoins, la proportion des rentiers conservée est de presque 98% avec la part de la population ayant des coûts de dossiers supérieurs à 100 000€, tandis qu'elle est à peine de 63% pour la table créée pour les personnes ayant un taux d'IPP supérieur à 20. On peut donc supposer que la première table est une meilleure approximation de la mortalité des rentiers, bien que moins fiable du fait du faible nombre de données utilisées.

4.3 Mortalité par spécialité mise en cause

Considérons maintenant la population formée des personnes ayant été victimes d'accidents médicaux dans l'ensemble des 6 spécialités mises en cause les plus importantes en termes de nombre et de coûts (Annexe A), à savoir :

- Neurochirurgie
- Obstétrique
- Orthopédie
- Chirurgie viscérale
- Urgences
- Pédiatrie et néonatalogie

La proportion de la population totale conservée est la suivante :

Proportion en nombre de la population totale conservée	Proportion en nombre des décès totaux conservés	Proportion en nombre des rentiers conservés	Proportion en coût des rentiers conservés
56.7%	47.3%	75.8%	85.0%

TABLE 4.7 – Étude de la mortalité par spécialité mise en cause - Proportion conservée pour la population, les décès et les rentiers

On remarque que cette population conserve environ la moitié de la population et des décès. Il n'est pas nécessaire d'étudier la mortalité de cette base de données, car elle ne sera pas très différente de celle de la mortalité totale. Elle aura même pour certains âges une mortalité inférieure à celle de la population totale. Cela est confirmé par le ratio SMR calculé par rapport à la table de la population totale, et qui vaut 0.920. Ce ratio est inférieur à 1, la mortalité de cette population est donc globalement plus faible que celle de la population totale.

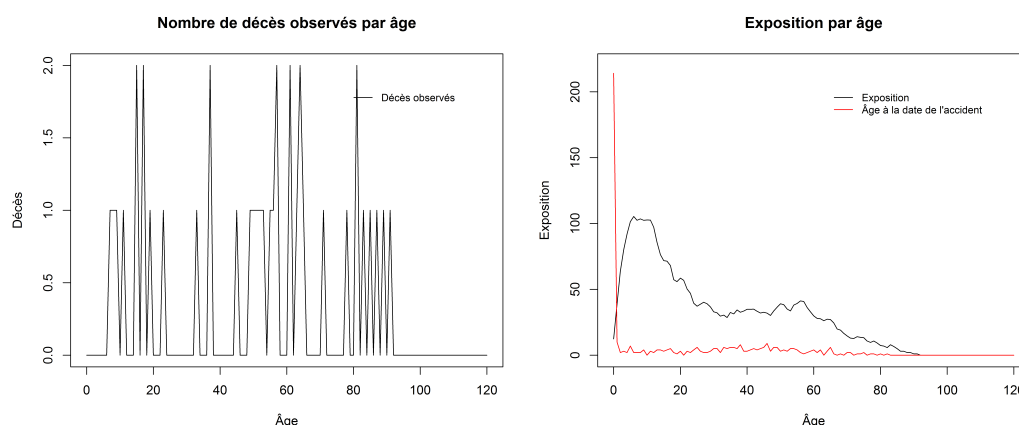
En effet, cette population conserve moins de la moitié des décès. Cela signifie donc que plus de la moitié des décès a lieu lorsque d'autres spécialités sont mises en cause, comme par exemple la cardio-chirurgie. Cette population ne conserve par ailleurs qu'à peine 60% du nombre de décès parmi les rentiers. Il n'est donc pas utile d'étudier la mortalité de cette population.

4.4 Mortalité des rentiers

Il n'est pas possible d'estimer directement la mortalité des rentiers. En effet, il y a seulement 474 dossiers évalués en rente dans la base de données, parmi lesquels ont eu lieu 38 décès. Nous allons cependant utiliser les informations sur les rentiers

pour valider le choix de la population ayant la mortalité la plus proche de celle des rentiers.

La répartition des décès et de l'exposition parmi les rentiers entre 2004 et 2015 est présentée dans le graphique 4.15.



GRAPHIQUE 4.15 – Rentiers - Exposition et nombre de décès

On remarque que les rentiers sont davantage présents avant 18 ans, du fait des graves accidents ayant eu lieu en obstétrique. Plus de 45% des accidents ont effectivement eu lieu à la naissance pour les rentiers. Au-delà, ceux-ci sont en nombre plus faible mais le type de spécialité mise en cause est plus hétérogène. C'est en effet à l'âge adulte que les consultations et donc les accidents aux urgences, en chirurgie, etc. sont plus importants.

En revanche, le nombre de décès reste constant et faible selon les âges. On peut donc en déduire que les taux bruts de décès sont a priori plus faibles avant 18 ans qu'après, et qu'ils augmentent avec l'âge.

Il n'est pas possible d'établir une table de mortalité à partir d'un nombre aussi faible de données.

Nous pouvons cependant nous donner une idée du comportement des rentiers en effectuant un ajustement d'une table existante, dans notre cas la table que nous avons construite pour la population totale, à l'aide des données de décès et d'exposition.

Les valeurs des ratios SMR calculés en prenant comme mortalités de référence les différentes tables construites jusqu'alors sont présentées dans le tableau 4.8. C'est la table construite avec les dossiers ayant un coût supérieur à 100 000€ qui présente la mortalité la plus proche de celle des rentiers, suivie de la table construite avec les victimes ayant un taux d'IPP supérieur à 20.

Les résultats obtenus en réalisant un ajustement par le modèle de Brass sont présentés dans le graphique 4.16. L'ajustement est réalisé par rapport à la table créée pour la population totale, car elle est a priori plus fiable que les tables construites avec une population plus réduite. Cela permet de plus de comparer les résultats avec ceux obtenus précédemment.

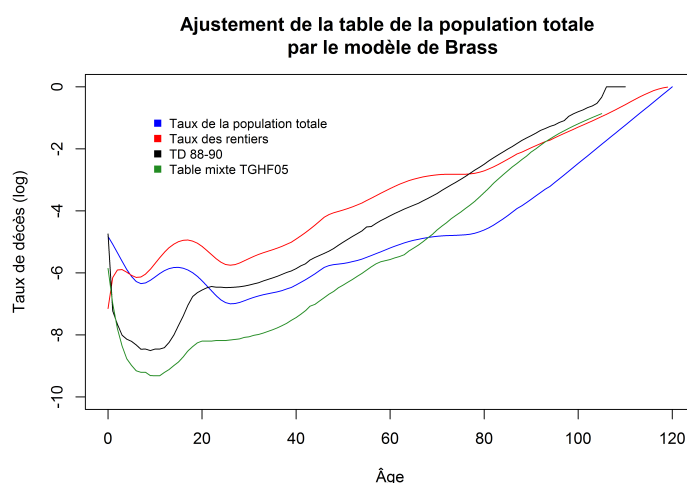
Les valeurs estimées des paramètres du modèle de Brass calculé par rapport à la population totale sont les suivantes :

Mortalité de référence	Ratio SMR
Population totale	3.438
Dossiers ayant un coût supérieur à 1 000€	2.937
Dossiers ayant un coût supérieur à 10 000€	2.404
Dossiers ayant un coût supérieur à 100 000€	1.101
Victimes ayant un taux d'IPP supérieur à 8	2.261
Victimes ayant un taux d'IPP supérieur à 20	1.435

TABLE 4.8 – Ratios SMR de la population des rentiers selon différentes tables de référence

$$\begin{cases} a = 3.799 \\ b = 2.272 \end{cases}$$

Ces valeurs sont très élevées, et les plus proches que nous ayons obtenues sont effectivement celles calculées pour la part de la population ayant des coûts de dossiers supérieurs à 100 000€.



GRAPHIQUE 4.16 – Rentiers - Comparaison avec les tables réglementaires

Avec l'ajustement par le modèle de Brass, la TD 88-90 est prudente jusqu'à 77 ans, tandis que la TGHF05 reste prudente pour tous les âges.

Cependant, les provisions sont tout de même probablement prudentes lorsqu'elles sont calculées avec la TD 88-90. En effet, la contribution des âges élevés est faible au calcul des provisions totales, surtout lorsque la plupart des accidents ont lieu à un âge jeune, comme cela est notre cas ici.

4.5 Sensibilité de la mortalité aux hypothèses et aux modèles

4.5.1 Sensibilité aux hypothèses

Nous avons étudié dans les parties précédentes la sensibilité de la mortalité aux différentes hypothèses réalisées sur la population, c'est-à-dire que nous avons étudié la mortalité pour différentes sous-populations de la base de données, dans

l'objectif d'essayer d'approcher la mortalité des rentiers ayant été victimes d'accidents médicaux.

Pour résumer le niveau de mortalité total de chaque sous-population, les ratios SMR, calculés pour chacune de ces sous-populations en prenant pour table de référence la table créée à partir de l'ensemble de la base de données sans distinction par sexe, sont présentés dans le tableau 4.9.

Sous-population étudiée	Ratio SMR
Population totale	1
Dossiers ayant un coût supérieur à 1 000€	1.124
Dossiers ayant un coût supérieur à 10 000€	1.317
Dossiers ayant un coût supérieur à 100 000€	2.904
Victimes ayant un taux d'IPP supérieur à 8	1.516
Victimes ayant un taux d'IPP supérieur à 20	2.070
Rentiers	3.438

TABLE 4.9 – Ratios SMR des sous-populations étudiées en prenant pour référence la population totale mixte

On remarque que plus les hypothèses sur la population sont restrictives, plus la mortalité est élevée et supérieure à la mortalité de la population totale.

Ce sont les rentiers qui ont le niveau de mortalité le plus élevé.

De la même manière, les paramètres du modèle de Brass appliqués à ces sous-populations en prenant comme mortalité de référence celle de la population totale sont résumés dans le tableau 4.10. Ces paramètres permettent une interprétation plus fine de la mortalité car ils sont au nombre de deux, contrairement à la méthode du ratio SMR qui est un modèle à un seul paramètre. Les paramètres du modèle de Brass permettent d'interpréter séparément le niveau global de mortalité de la sous-population étudiée, ainsi que la mortalité selon les âges. Pour rappel, le modèle de Brass établit une relation linéaire entre les logits des probabilités de décéder dans les x années suivant la naissance, sous la forme : $y = \alpha + \beta x$.

Sous-population étudiée	Paramètre α	Paramètre β
Population totale	0	1
Dossiers ayant un coût supérieur à 1 000€	0.198	1.002
Dossiers ayant un coût supérieur à 10 000€	0.514	0.961
Dossiers ayant un coût supérieur à 100 000€	2.653	1.840
Victimes ayant un taux d'IPP supérieur à 8	0.613	1.035
Victimes ayant un taux d'IPP supérieur à 20	1.237	1.126
Rentiers	3.799	2.272

TABLE 4.10 – Paramètres du modèle de Brass des sous-populations étudiées en prenant pour référence la population totale mixte

Les paramètres indiquent globalement les mêmes comportements par rapport à la table de référence créée à partir de la base de données complète.

Le paramètre α est toujours positif, ce qui signifie que pour toutes les sous-populations, le niveau de mortalité est supérieur au niveau de mortalité de la table

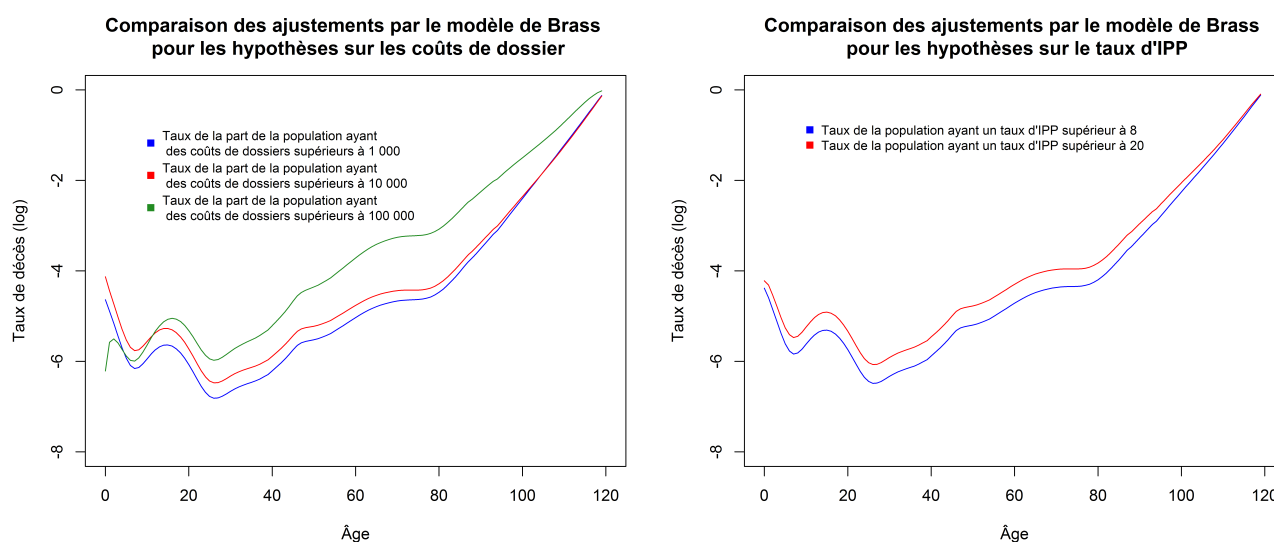
de référence. Par ailleurs, le niveau de mortalité augmente avec le niveau de restriction des hypothèses sur la population.

Le paramètre β est quant à lui presque toujours supérieur à 1. Pour les hypothèses les moins restrictives, sa valeur est assez proche de 1, ce qui signifie que la proportion des survivants aux âges jeunes ou aux âges élevés est globalement identique à celle de la table de référence. En revanche, pour les rentiers et pour la table des dossiers ayant un coût supérieur 10 000€, le paramètre β est nettement supérieur à 1, donc la proportion de survivants est plus grande parmi les âges jeunes que parmi les âges élevés par rapport à la répartition au sein de la table de référence.

Ces études nous permettent d'établir quelles sous-populations ont la mortalité la plus proche des rentiers, avec toutefois une incertitude de plus en plus importante du fait du faible nombre de données.

Ce sont a priori les tables créées pour les dossiers ayant un coût supérieur à 1 000€ et à 10 000€ qui ont une mortalité la plus éloignée de celle des rentiers, après bien sûr la table créée pour la population totale. Viennent ensuite les tables créées pour les victimes ayant un taux d'IPP supérieur à 8, puis supérieur à 20. C'est enfin la table créée pour les dossiers ayant un coût supérieur à 100 000€ qui est la plus proche de la mortalité des rentiers, mais c'est également celle qui a été créée avec le moins de données.

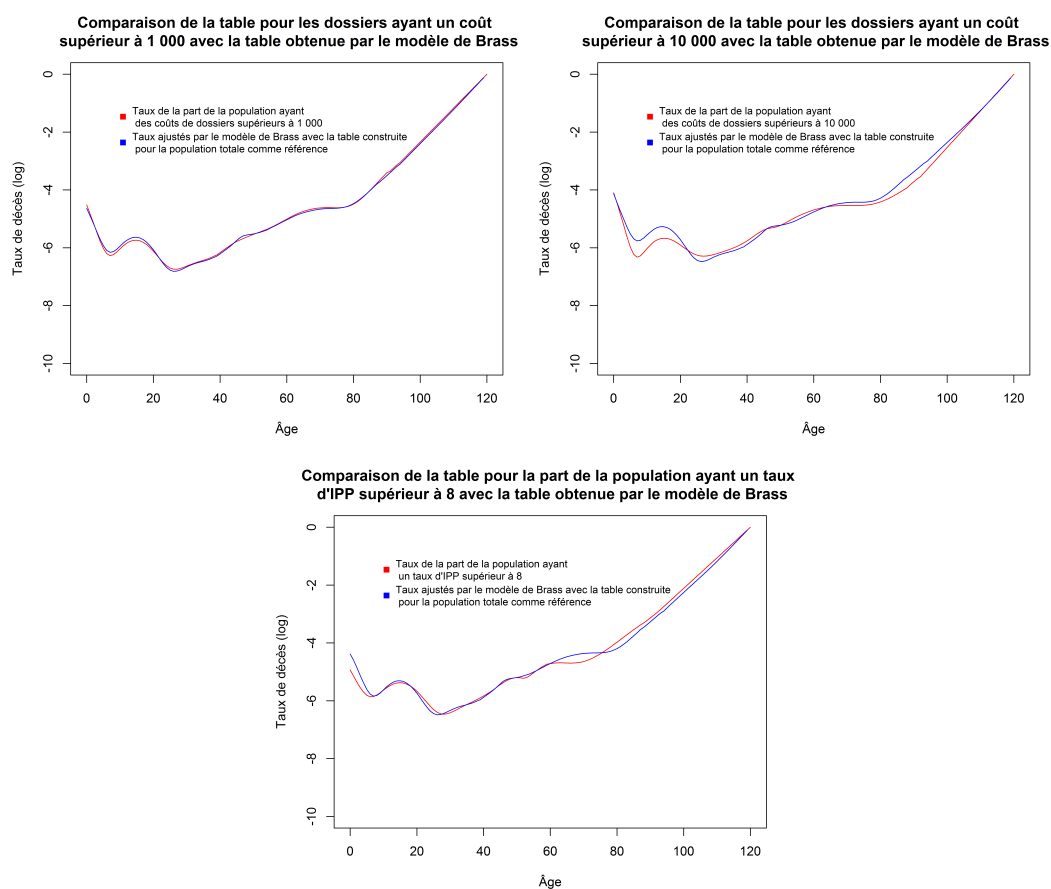
Les différentes tables créées par le modèle de Brass sont comparées dans le graphique 4.17, en fonction du type d'hypothèse (sur les coûts des dossiers ou sur les taux d'IPP).



GRAPHIQUE 4.17 – Sensibilité des ajustements aux hypothèses

4.5.2 Sensibilité aux modèles

Lors de l'étude, nous avons également comparé les résultats obtenus par le modèle de Brass avec les tables construites lorsque cela était possible. Cela nous permet effectivement d'avoir une idée de l'erreur faite lorsque le nombre de données est trop faible pour utiliser un autre modèle que l'ajustement à une référence par le modèle de Brass.



GRAPHIQUE 4.18 – Sensibilité des ajustements aux modèles

Globalement (Graphique 4.18), les deux tables construites diffèrent pour les moins de 25 ans, et autour de 80 ans, plages d'âges pour lesquelles le modèle de Brass surestime légèrement la mortalité. Autrement, les taux de décès sont sensiblement identiques pour les deux modèles de construction utilisés.

4.6 Provisions mathématiques pour les rentes

Les provisions mathématiques des rentes en invalidité sont calculées avec la TD 88-90, conformément à l'article A331-10 du Code des Assurances.

L'objectif de cette partie est de calculer les provisions mathématiques en s'appuyant sur les tables construites, et d'estimer les écarts entre les différentes provisions.

Nous nous appuyerons sur les provisions calculées au 31/12/2015 avec la TD 88-90, qui seront exprimées en base 100. Ce sont les valeurs actuellement utilisées par Sham, et qui serviront de base pour le calcul des provisions avec les autres tables de mortalité. Les autres valeurs de provisions seront exprimées relativement à cette valeur de référence (Table 4.11).

Table utilisée	Provisions mathématiques
TD 88-90	100
TGHF05	142
Population totale	137
Dossiers ayant un coût supérieur à 1 000€	132
Dossiers ayant un coût supérieur à 10 000€	130
Dossiers ayant un coût supérieur à 100 000€	94
Victimes ayant un taux d'IPP supérieur à 8	124
Victimes ayant un taux d'IPP supérieur à 20	111
Rentiers	82

TABLE 4.11 – Provisions mathématiques au 31/12/2015 des rentes selon les tables de mortalité construites (en base 100)

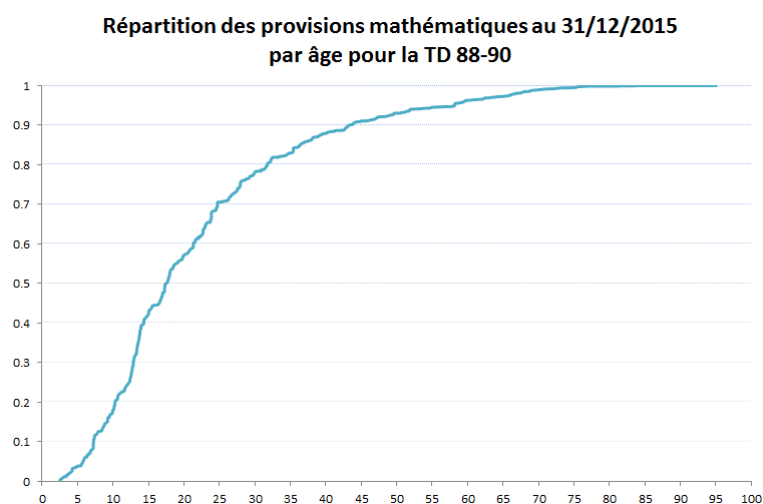
On remarque dans un premier temps que si les provisions mathématiques des rentes devaient être calculées avec la TGHF05, leur montant augmenterait de 42% par rapport à leur niveau actuel.

Les provisions calculées avec toutes les autres tables sont également supérieures à la référence, sauf celles calculées avec les tables des rentiers et des dossiers ayant un coût supérieur à 100 000€, qui sont inférieures respectivement de 18% et 6% par rapport aux provisions actuelles.

Il est cependant difficile de conclure quant à la prudence de la TD 88-90, car ces deux dernières tables ont été construites avec un nombre de données très insuffisant pour nous assurer de leur significativité.

En revanche, les provisions calculées avec la TGHF05 sont toujours supérieures à celles calculées avec n'importe laquelle des tables, et nous permettent a priori d'assurer une forte prudence vis-à-vis de la mortalité des rentiers.

La répartition par âge des provisions mathématiques calculées par la TD 88-90 est présentée dans le graphique 4.19.



GRAPHIQUE 4.19 – Répartition des provisions mathématiques

Ce graphique permet de constater que 50% des provisions sont constituées pour les moins de 17 ans, et que 75% concernent les moins de 28 ans. 90% des provisions sont constituées pour les moins de 43 ans.

Conclusion

L'étude de la base de personnes victimes d'accidents médicaux de Sham nous a permis de situer les rentiers parmi cette population, et de déduire certaines caractéristiques permettant de mieux approximer leur mortalité.

La population des rentiers est effectivement une population jeune, mais ayant des coûts de dossier et des taux d'IPP renseignés supérieurs à ceux du reste de la population totale.

Une première étude de mortalité sur la population globale nous a permis de montrer que les hommes et les femmes ont une mortalité semblable. Comme plus d'un tiers des sinistres ont lieu en obstétrique, il était intéressant de savoir si les femmes présentes dans le portefeuille révèlent de ce fait une surmortalité entre 20 et 35 ans. Ce n'est pas le cas car les accidents médicaux ayant lieu dans cette spécialité médicale concernent principalement les nouveaux nés.

Nous avons également pu choisir durant cette étude la méthode de construction de table de mortalité la plus adaptée à nos données. Si le critère de Cochran est vérifié sur une plage d'âge suffisamment importante, la table est construite de la manière suivante :

- Ajustement par le modèle de Brass pour les âges jeunes.
- Lissage des taux bruts de Kaplan-Meier par la méthode de Whittaker-Henderson sur la plage d'âge où le critère de Cochran est vérifié.
- Fermeture par régression exponentielle aux âges élevés.

Dans le cas contraire, la table mixte construite pour la population totale sert de référence pour un ajustement par le modèle de Brass entre 0 et 120 ans.

À partir de là, il est désormais possible de réduire la population étudiée.

En utilisant notamment des critères concernant les coûts des dossiers et les taux d'IPP des victimes lorsque ceux-ci étaient renseignés, nous avons réussi à nous approcher de mieux en mieux de la mortalité de rentiers en augmentant la valeur de ces critères. Ce faisant, nous réduisons cependant la taille de la population étudiée et augmentons l'incertitude autour de ces calculs. Les résultats doivent donc être considérés avec d'autant plus de précautions.

Les tables construites les plus proches de la mortalité des rentiers indiquent que la TD 88-90 est très prudente jusqu'à 18 ans puis encore assez prudente jusqu'à 70 ans environ. Au global, les provisions calculées au 31/12/2015 avec cette table réglementaire sont insuffisantes pour la population complète de victimes d'accidents médicaux, mais deviennent cependant progressivement suffisantes et prudentes lorsque l'on se rapproche de la population de rentiers étudiée.

Il serait intéressant de poursuivre cette étude dans les prochaines années, pour avoir davantage de données exploitables, afin de dégager une tendance sur plusieurs années.

Bibliographie

- [1] A. COALE et E.E. KISKER. « Defects in data on old-age mortality in the United States : New procedures for calculating schedules and life tables at the higher ages ». In : *Asian and Pacific Population Forum* (1990), 4 :1–31.
- [2] M. DENUIT et A. QUASHIE. *Modèles d'extrapolation de la mortalité aux grands âges*. Louvain-la-Neuve, Belgium, Institut des Sciences Actuarielles et Institut de Statistique, Université Catholique de Louvain, 2005.
- [3] E.L. KAPLAN et P. MEIER. « Nonparametric Estimation from Incomplete Observations ». In : *Journal of the American Statistical Association* 53.282 (1958), p. 457–481.
- [4] J.P. KLEIN et M.L. MOESCHBERGER. *Survival analysis : techniques for censored and truncated data*. Springer Science & Business Media, 2005.
- [5] P.N. MARI BHAT. *A Generalized Relational Model for Demographic Analysis*. Delhi, India, Institute of Economic Growth, 2004.
- [6] F. PLANCHET et P. THÉRON. *Modèles de durée – Applications actuarielles*. Paris, Economica, 2006.
- [7] T. REYNAUD. « Analyse du risque de mortalité par la construction de tables d'expérience sur des populations spécifiques en réponse aux exigences de Solvabilité II ». Paris, Institut de Statistique de l'Université Pierre et Marie Curie (Mémoire présenté devant l'Institut des Actuaire), 2011.
- [8] P. THÉRON. *Sépia « Certification des tables d'expérience »*. Le Point de vue du certificateur, Paris, 2010.
- [9] X. ZHANG. « Construction des tables de mortalité d'expérience en cas de décès ». Lyon, Institut de Science Financière et d'Assurances (Mémoire présenté devant l'Institut des Actuaire), 2011.

Annexe A

Décès et exposition pour les principales spécialités mises en cause

Nous allons regarder la contribution au nombre de décès et à l'exposition de chacune des 6 spécialités mises en cause les plus importantes en terme de nombre et de coût total, à savoir :

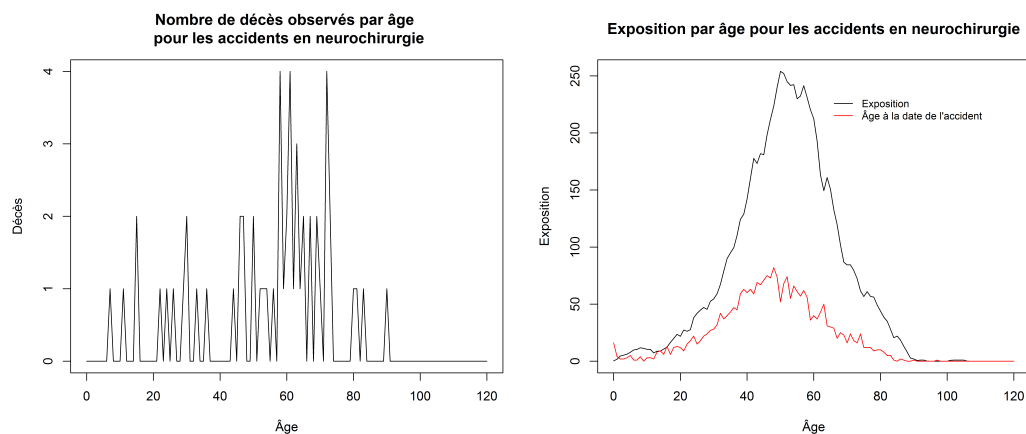
- Neurochirurgie
- Obstétrique
- Orthopédie
- Chirurgie viscérale
- Urgences
- Pédiatrie et néonatalogie

A.1 Accidents en neurochirurgie

La neurochirurgie est une spécialisation de la chirurgie qui se consacre au diagnostic et aux opérations des maladies du système et des centres nerveux (par exemple une tumeur cérébrale).

On remarque (Graphique A.1) que les accidents en neurochirurgie ont majoritairement lieu entre 30 et 70 ans, et les décès sont les plus importants entre 59 et 74 ans.

Cette population conserve 8.9% des rentiers.



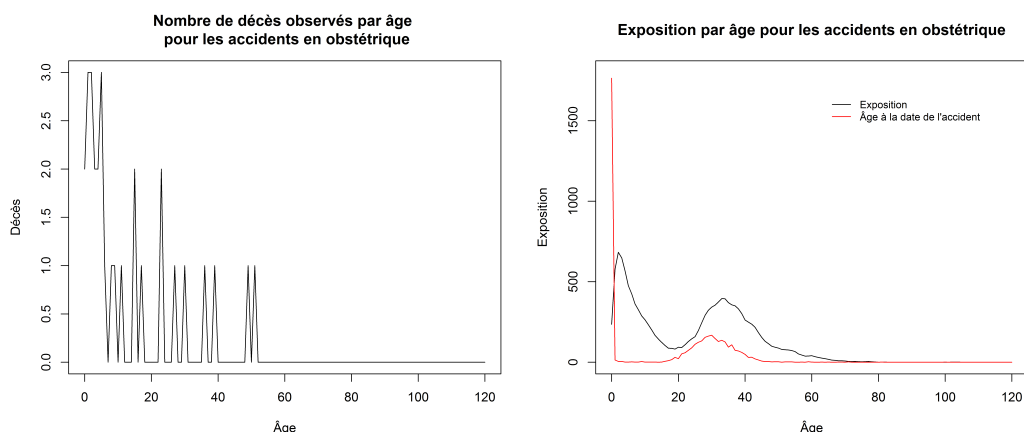
GRAPHIQUE A.1 – Accidents en neurochirurgie - Exposition et nombre de décès

A.2 Accidents en obstétrique

L'obstétrique est la branche de la médecine qui prend en charge la grossesse, l'accouchement et les suites de couches.

Les accidents en obstétrique (Graphique A.2) ont principalement lieu à la naissance, et entre 18 et 45 ans. Cela correspond effectivement aux nouveau-nés et aux femmes enceintes ou qui ont accouché. Le nombre de décès est plus important avant 5 ans, puis est stable jusqu'à 50 ans.

Cette population conserve 38.2% des rentiers.



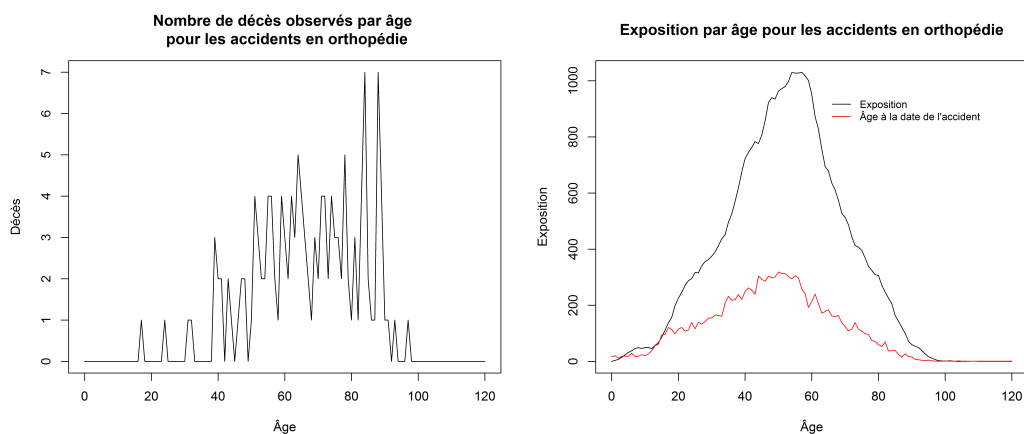
GRAPHIQUE A.2 – Accidents en obstétrique - Exposition et nombre de décès

A.3 Accidents en orthopédie

L'orthopédie est la spécialité chirurgicale qui traite les problèmes osseux, ceux des articulations, des tendons, des ligaments, et des muscles.

Le nombre d'accidents augmente fortement à l'adolescence (Graphique A.3), puis redevient faible après 80 ans. En revanche, les décès sont plus importants entre 50 et 90 ans.

Parmi la population des personnes ayant eu un ou plusieurs accidents médicaux en orthopédie conserve 8.1% des rentiers.



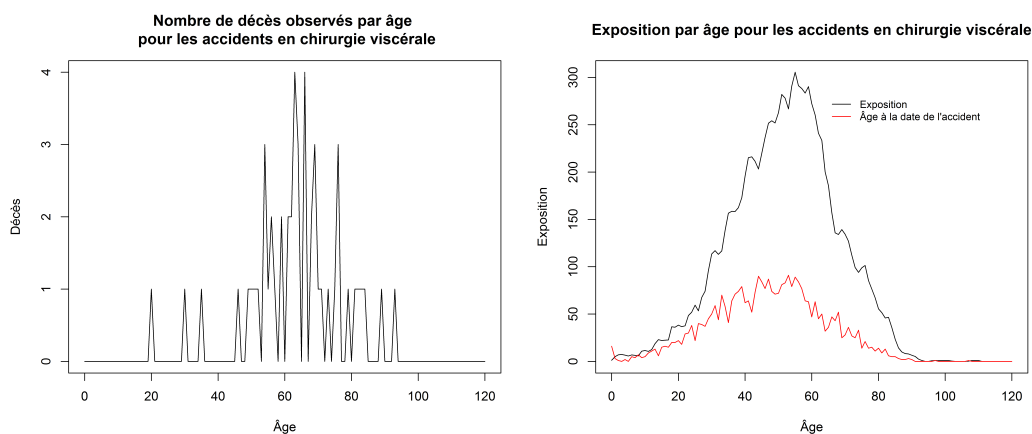
GRAPHIQUE A.3 – Accidents en orthopédie - Exposition et nombre de décès

A.4 Accidents en chirurgie viscérale

La chirurgie viscérale est la spécialité chirurgicale destinée à traiter les pathologies qui touchent la zone abdominale, notamment l'appareil digestif, la paroi abdominale, et les organes intra-abdominaux non-digestifs. Les affections les plus courantes sont l'appendicite, l'occlusion intestinale, les traumatismes abdominaux, les hémorroïdes et les hernies.

Les accidents en chirurgie viscérale ont principalement lieu entre 20 et 75 ans (Graphique A.4). Les décès sont quant à eux plus importants entre 55 et 80 ans.

Cette population conserve 3.4% des rentiers



GRAPHIQUE A.4 – Accidents en chirurgie viscérale - Exposition et nombre de décès

A.5 Accidents aux urgences

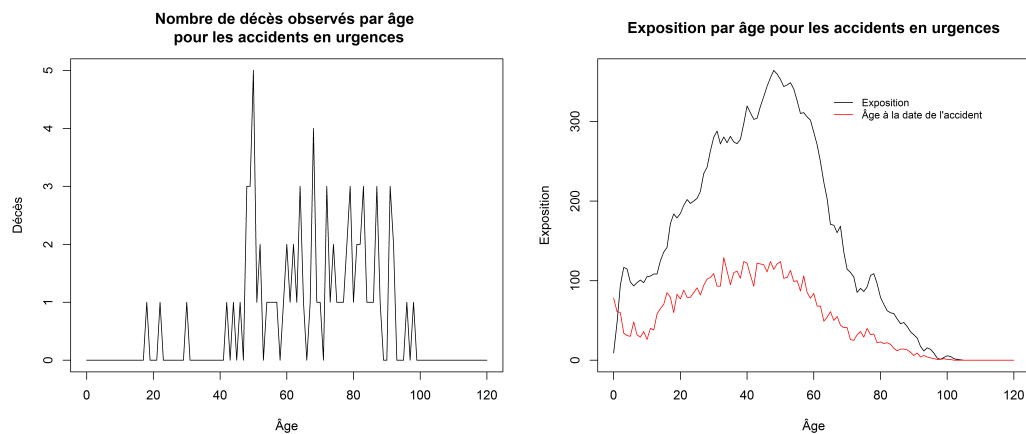
La médecine d'urgence est une spécialité médicale et chirurgicale qui traite les maladies ou les accidents qui occasionnent ou risquent d'entraîner rapidement une altération d'une fonction essentielle ou la mort.

Les accidents aux urgences ont lieu à tous les âges, et principalement avant 80 ans (Graphique A.5). Cependant, les décès ont généralement lieu entre 50 et 90 ans, et cette population conserve 11.7% des rentiers.

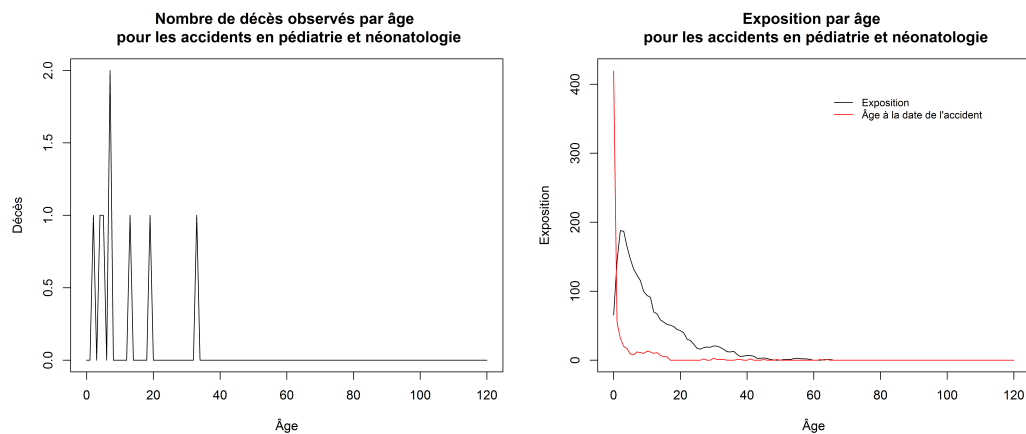
A.6 Accidents en pédiatrie et néonatalogie

La pédiatrie est une branche de la médecine qui a pour objet l'étude, le diagnostic, le traitement, la prévention des maladies infantiles et la protection de l'enfance. La néonatalogie est une branche de la pédiatrie qui se consacre aux soins et au traitement des nouveau-nés, définis par un âge inférieur à 28 jours.

Les accidents en pédiatrie et néonatalogie ont lieu à la naissance, principalement. Cependant, les décès sont très faibles, et ont généralement lieu avant 20 ans (Graphique A.6). Cette population conserve 6.4% des rentiers.



GRAPHIQUE A.5 – Accidents aux urgences - Exposition et nombre de décès



GRAPHIQUE A.6 – Accidents en pédiatrie et néonatalogie - Exposition et nombre de décès