HAL
open science

Prise en compte de l'hétérogénéité des individus dans le modèle de Lee-Carter

Cédric Boureau

To cite this version:

Cédric Boureau. Prise en compte de l'hétérogénéité des individus dans le modèle de Lee-Carter. Gestion des risques [q-fin.RM]. 2014. dumas-01073547

HAL Id: dumas-01073547

https://dumas.ccsd.cnrs.fr/dumas-01073547

Submitted on 10 Oct 2014

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

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

Mémoire présenté devant

l'UFR de Mathématique et d'Informatique

pour l'obtention du Diplôme Universitaire d'Actuaire de Strasbourg

et l'admission à l'Institut des Actuaires

le Mercredi 1er Octobre 2014

Par : BOUREAU Cédric

Titre: Prise en compte de l'hétérogénéité des individus dans le modèle de Lee-Carter

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

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

Membres du jury de l'Institut des Actuaires signature

M. Dubois

M. You

Membres du jury de l'UdS :

M. Franchi

M. Eisele

M. Netzer

M. Bérard

Invités :

M. Fitouchi

M. Dubois

M. Modry

Mme Kelle-Vignon

M. You

Mme. Foata

M. Gadenne

M. Hess

Mme Maumy-Bertrand

M. Vigon

Secrétariat : Mme. Richard

Bibliothèque : Mme. Disdier

Entreprise :

Nom : Active Asset Allocation

Signature :

Directeur de mémoire en entreprise :

Nom : GRIGORIU Adina

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

Remerciements

J'adresse mes remerciements aux personnes qui m'ont aidé dans la réalisation de ce mémoire, tout au long de mon stage et de mon parcours universitaire.

En premier lieu, je remercie Mme. Adina GRIGORIU et toute l'équipe d'Active Asset Allocation de m'avoir accueilli au sein de leur entreprise pour mon stage de fin d'études.

Mes pensées vont également à mon tuteur de stage, M. Pierre DEVOLDER, qui a su m'encadrer et trouver les mots justes à mes interrogations.

Mes remerciements s'adressent par ailleurs à l'ensemble de mes collègues de promotion 2014 ainsi qu'aux professeurs de l'Université de Strasbourg, qui nous ont transmis tous leurs savoirs et leurs connaissances, mais qui ont aussi su nous faire découvrir et aimer l'actuariat à travers toutes ses facettes.

Résumé

Depuis de nombreuses années, l'espérance de vie ne cesse d'augmenter. Bien que ce phénomène puisse paraître attrayant pour la population, il engendre des conséquences importantes sur l'économie, et notamment celle des rentes viagères. Le risque de longévité mérite donc une attention et une évaluation toute particulière.

La modélisation et la projection de la mortalité a fait l'objet de nombreux sujets de recherches. De par sa simplicité de mise en place et son explication dans les paramètres, le modèle le plus utilisé est certainement celui de Lee-Carter. Pourtant, ce modèle atteste de plusieurs inconvénients, comme l'hypothèse d'homoscédascité des résidus ou bien les étroits intervalles de confiance des projections. Le principal modèle développé ici est une extension de ce dernier. Il suppose que le nombre de décès est distribué selon une loi Négative Binomiale, qui permet de soulager le modèle de l'inconvénient d'homoscédascité, mais également de prendre en compte l'hétérogénéité qui peut exister chez les individus.
Cette extension permet d'obtenir un meilleur ajustement sur les données utilisées avec des intervalles de confiance plus adéquats et moins prudents lors de la construction de backtests.

Mots-clefs : Risque de longévité, Modélisation de la mortalité, Projection de la mortalité, Modèle de Lee-Carter, Modèle Log-Poisson, Modèle Binomiale Négatif, Bootstrap, Hétérogénéité, Modèles de fragilité, Rentes viagères

Abstract

Looking at the mortality rates in time, one can check that people use to live older. This phenomenon leads to important matters on financial institutions such as pension fund system or life annuities business. This is why, actuaries are getting used to model and forecast mortality in order to assess the longevity risk.
The most famous model is the Lee-Carter one, which has been applied to several studies and american institutions, and most of times, gave consistent and attractive results. However, this model gives rise to several inconsistencies, like an homoskedascticity of errors terms and narrow confidence bound. The model we developed here is an extension of that one, based on the assumption that the number of deaths can be model with a Binomiale Negative distribution. According to that hypothesis, we remove the homoskedasticity assumption and take into account the degree of heterogeneity that arise from the population. The perks of this model lead to wider confidence intervals for the projection of the mortality rates and a better goodness-of-fit.

Key-Words : Longevity risk, Mortality modeling, Mortality forecasting, Lee-Carter model, Log-Poisson model, Binomiale Negative Model, Heterogeneity, Frailty models, Life annuities

Table des matières

Introduction **1**

1 Mathématiques de l'assurance vie **4**
1.1 La durée de vie vue comme une variable aléatoire 4
1.2 La force de mortalité 5
1.3 Les notations actuarielles 6
1.4 La répartition intra-annuelle des décès 8
1.5 Les tables de mortalité 10

2 Présentation des données **12**
2.1 *Human Mortality Database* 12
2.2 Evolution de la mortalité en France 13

3 Etat de l'art sur les modèles de mortalités **20**
3.1 Les modèles paramétriques 20
3.2 Les modèles non-paramétriques 21
3.2.1 Les modèles Age-Période 22
3.2.2 Les modèles Age-Période-Cohorte 24
3.3 Les modèles stochastique en temps continu 25

4 Modéliser la mortalité **26**
4.1 Le modèle de Lee-Carter 27
4.1.1 Estimation des paramètres 28
4.1.2 Ajustement de κ_t 31
4.2 Le modèle Log-Poisson 35
4.2.1 Estimation des paramètres 35
4.3 La dispersion dans les données 39
4.4 Les facteurs d'hétérogénéité 40
4.4.1 Les facteurs d'hétérogénéité observables 40
4.4.2 Les facteurs d'hétérogénéité inobservables 41

4.4.3 Quelles distributions pour la fragilité ? 41
4.5 Le modèle Binomial Négatif spécifique à l'âge 43
4.5.1 Estimations des paramètres 46
4.6 Ajustement aux données 50
4.6.1 Taux d'inertie 50
4.6.2 Variance temporelle 51
4.6.3 AIC et SBC 53
4.6.4 Analyse des résidus 53
4.7 Modélisation aux âges extrêmes 59
4.7.1 La méthode de Coale-Kisker 59
4.7.2 Méthode de Denuit-Goderniaux 61
4.8 Conclusion du chapitre 63

5 Projeter la mortalité 65
5.1 Correction du saut dans la projection 65
5.1.1 Modèle log-Poisson 66
5.1.2 Modèle Négatif Binomial 66
5.2 Modélisation de κ_t 69
5.2.1 Analyse de la série 69
5.2.2 Analyse de la série différenciée 70
5.2.3 Estimation des paramètres du modèle 73
5.2.4 Analyse des résidus 74
5.3 Projection de κ_t 80
5.4 Conclusion du chapitre 83

6 Intervalles de confiance des modèles et backtest 85
6.1 Intervalles de confiance par bootstrap 85
6.2 Backtest 93

Conclusion Générale 96

Table des figures 99

Liste des tableaux 101

Bibliographie 103

Annexes 114

Introduction

Depuis la fin du 18^{e}siécle, l'espérance de vie n'a cessé de progresser de manière spectaculaire et dépasse les limites biologiques annoncées il y a quelques décennies. Plus l'espérance de vie augmente, plus elle entraîne une diminution intensive de la mortalité des populations les plus âgées et donc, la longévité d'un humain, qui est la durée de vie pour laquelle il est "programmé", continue d'augmenter au cours du temps.
Les produits sous forme de rentes viagères (pensions de retraite, bien immobilier en viager, rentes de retraite supplémentaire, dommage corporel en droit commun, jeux de hasard avec une rente "à vie") sont des contrats qui garantissent le versement périodique d'une somme d'argent à une personne, jusqu'à son décès. Le risque de longévité y est alors fortement présent. Malgré leurs attractivité, ces produits sont trop peu utilisés dans la plupart des pays, que ce soit à cause des préférences pour les liquidations en capitaux ou bien par méconnaissance de l'assurance-vie.

Lors du calcul des montants de primes d'assurance, les actuaires utilisent traditionnellement les tables de mortalité transversales, ce qui constitue une abérration lorsque l'on calcule des rentes viagères. Tarifer ou provisionner avec des tables apporte de grands risques pour l'assureur qui n'aura pas chercher à prendre en compte le risque de longévité, étant donné que les primes calculées ne reflètent le risque d'aucun assuré réel. Il est alors préférable d'utiliser des tables de mortalité prospectives, ayant deux entrées : l'âge et l'année calendaire. Elles sont établies grâce à des modèles mathématiques estimant les tendances de la mortalité, et les projetant dans le futur.

Nous venons de voir que l'espérance de vie des individus ne cesse d'augmenter en France, mais quels sont les tendances, dans le temps, mais également selon l'âge, des taux de décès ? Quels sont les modèles mathématiques existants aujourd'hui et permettant de modéliser et projeter la mortalité dans le

temps ? Ces modèles fournissent-ils les mêmes approximations ? Les projections sont-elles comparables ?

L'un des modèles les plus utilisés et connus en terme de mortalité prospective est le modèle de Lee-Carter [85]. Ce modèle stochastique à trois paramètres, dépendant de l'âge ou du temps, a servi avec succès dans de nombreuses études de populations. Depuis son élaboration, de nombreux auteurs y ont eu recours et l'ont amélioré, que ce soit par des rajouts de termes temporels ou de termes de cohorte. L'une des variantes de ce modèle, supposé par BROUHNS et *al.* [72], est la modélisation du nombre de décès par une loi de Poisson. Mais cette distribution s'avère être trop rigide de par sa propriété d'équipersion. Cette sur-dispersion dans les données s'explique par l'hétérogénéité qui réside chez les individus. La mise en place d'un paramètre de sur-dispersion spécifique à l'âge (LI et *al.* [81]) permet de soulager le modèle de cette hypothèse.

En effet, la quasi-totalité des modèles existants suppose uniquement que l'âge, le temps et le sexe influent sur les probabilités de décès. D'autres facteurs, comme l'effet de l'alcool, du tabac ou la pratique de certains sports jouent sur celles-ci. On parle alors de facteurs observables. Ceux-ci sont la plupart du temps pris en compte lors de la souscription de contrats d'assurance-vie, où une visite médicale est également possible. Cependant, certains facteurs ne sont pas observables, les individus ne naissent pas avec les mêmes comportements psychologiques ou les mêmes aptitudes physiologiques. On parle alors d'hétérogénéité de la population.

Tout en gardant la simplicité de compréhension des paramètres du modèle de Lee-Carter, le modèle de LI et *al.* [81] nous a permis sur des données de mortalités françaises, d'obtenir un meilleur ajustement sur les décès que les modèles plus classiques. La projection est également meilleure : les intervalles de confiance sont moins prudents ce qui permet à la quasi-totalité des taux de décès de rester à l'intérieur de l'intervalle.

Ce mémoire s'articule en plusieurs chapitres. Nous reviendrons tout d'abord sur les bases des mathématiques actuarielles liées à l'assurance vie, qui constitue une notion importante, avant de pouvoir entrer dans le vif du sujet. Puis, nous présenterons nos données, qui sont celles de la population en France.

Une analyse de la mortalité française sera ainsi effectuée, de manière à bien saisir les dynamiques qui s'en dégagent. Nous avons pensé intéressant d'effectuer un état de l'art sur les différentes techniques statistiques existantes pour la modélisation et la projection de la mortalité. Nous aborderons ensuite l'aspect de la modélisation de la mortalité, avec les modèles de Lee-Carter, log-Poisson et Négatif Binomiale. Ces derniers étant très sensibles aux âges extrêmes, à cause de la forte variabilité qui en résulte, nous verrons comment modéliser cette classe d'âge, en comparant deux méthodes. Viendra ensuite la question de la projection de la mortalité. Nous terminerons le mémoire sur les techniques de bootstrap, qui permettent d'obtenir des intervalles de confiance sur les modèles et sur les projections, ainsi que des indicateurs intéressants comme les espérances de vie résiduelle ou les prix des rentes.

Chapitre 1

Mathématiques de l'assurance vie

1.1 La durée de vie vue comme une variable aléatoire

En assurance-vie, lorsque qu'un assuré contracte une police, que ce soit sur un produit décès ou vie, l'assureur doit pouvoir évaluer les probabilités de survie ou de décès dans le temps des individus. Dans la suite, nous modélisons la durée de vie restante d'un individu d'âge x par la variable aléatoire T_x avec $x > 0$. On note alors $F_x(t)$ la fonction de répartition de T_x, c'est à dire :

$$F_x(t) = \Pr[T_x \leq t]$$

$F_x(t)$ représente la probabilité que l'individu décède entre l'âge x et $x + t$. L'intêret se porte aussi sur la fonction de survie de T_x :

$$S_x(t) = 1 - F_x(t) = \Pr[T_x > t]$$

$S_x(t)$ représente la probabilité qu'un individu survive encore t années au minimum, donc qu'il survive au moins jusqu'à l'âge $x + t$.

En supposant indépendante chaque probabilité de décèder une année sur l'autre, on obtient :

$$\Pr[T_x \leq t] = \Pr[T_0 \leq x + t | T_0 > x] = \frac{\Pr[x < T_0 \leq x + t]}{\Pr[T_x > x]}$$

d'où :

$$F_x(t) = \frac{F_0(x+t) - F_0(x)}{S_0(x)}$$

et en utilisant le fait que $S_x(t) = 1 - F_x(t)$:

$$S_0(x+t) = S_0(x) \times S_x(t)$$

Ce résultat permet d'interpréter la probabilité de survie de l'âge x à l'âge $x+t$ comme le produit entre, la probabilité de survie de la naissance jusqu'à l'âge x et la probabilité d'avoir survécut jusqu'à x, puis de survivre jusqu'à l'âge $x+t$.

Les fonctions de survie doivent obligatoirement répondre à plusieurs propriétés :

- $S_x(0) = 1$
- $\lim_{x \to +\infty} S_x(t) = 0$
- $S_x(t)$ ne peut pas être croissante en fonction t

1.2 La force de mortalité

La force de mortalité (ou taux de mortalité instantané) est un concept fondamental pour modéliser la durée de vie future. On la note :

$$\begin{aligned} \mu_x &= \lim_{dx \to 0^+} \frac{1}{dx} \Pr[T_0 \leq x + dx | T_0 > x] \\ &= \lim_{dx \to 0^+} \frac{1}{dx} \Pr[T_x \leq dx] \end{aligned}$$

Elle peut être interprétée comme la probabilité qu'un individu d'âge x décède avant d'atteindre l'âge $x + dx$ (avec dx infiniment petit), soit qu'il décède "instantanément".

La force de mortalité peut aussi se déduire à partir de la fonction de survie comme suit :

$$\begin{aligned}
\mu_x &= \lim_{dx \to 0^+} \frac{1}{dx}(1 - S_x(dx)) \\
&= \frac{1}{S_0(x)} \lim_{dx \to 0^+} \frac{S_0(x) - S_0(x + dx)}{dx} \\
&= -\frac{1}{S_0(x)} \frac{d}{dx} S_0(x) \\
&= \frac{f_0(x)}{S_0(x)}
\end{aligned}$$

d'après la relation :

$$f_x(t) = \frac{d}{dt} F_x(t) = -\frac{d}{dt} S_x(t)$$

1.3 Les notations actuarielles

Les précédentes notations sont les notations standard en statistique. Les sciences actuarielles ont développé les leurs, qui sont internationales. Parmi celles-ci, on retrouve :

$${}_tp_x = \Pr[T_x > t] = S_x(t)$$

$${}_tq_x = \Pr[T_x \leq t] = F_x(t)$$

$${}_{u|t}q_x = \Pr[u < T_x \leq x + t] = S_x(u)$$

$$\mu_x(t) = -\frac{1}{S_x(t)} \frac{d}{dx} S_x(t) = -\frac{d}{dt} \ln({}_tp_x)$$

En passant par la dérivée logarithmique, on obtient :

$$ {}_tp_x = \exp\left(\int_0^t -\mu_x(t)dt\right) $$

Dans de nombreuses applications actuarielles, la durée de vie restante d'un individu d'âge x en années entières[1] est nécessaire. On s'intéresse alors à :

$$ K_x = \lfloor T_x \rfloor $$

qui a pour fonction de probabilité :

$$ \Pr[K_x = k] = \Pr[x < T_x \leq x + 1] = {}_tp_x \times q_{x+k} $$

Egalement, on note L_x la population totale vivante, d'âge x. L'exposition au risque, qui représente le nombre total d'années vécues par les individus à l'âge x, se note alors :

$$ E_x = \int_0^1 L_x \times {}_up_x du = \int_0^1 L_{x+u} du $$

Un indicateur essentiel en assurance-vie est l'espérance de vie résiduelle à l'âge x (ou espérance de survie à l'âge x), soit la durée de vie moyenne restante à l'individu d'âge x, qui se calcule comme l'espérance de K_x :

$$ \begin{aligned} e_x &= \mathbb{E}[K_x] \\ &= \sum_{k \geq 0} k \Pr[K_x = k] \\ &= \sum_{k \geq 1} {}_kp_x \end{aligned} $$

[1]Curtate future lifetime en anglais

1.4 La répartition intra-annuelle des décès

Les tables de mortalités sont des données annuelles, il est donc nécessaire de déterminer les probabilités pour les décès entre les âges $x+t$ avec $t<1$. Plusieurs hypothèses existent :

- Hypothèse de linéarité de la mortalité : $\Pr[k < T_x \leq x+t | K_x = k] = u$, d'où, ${}_tq_x = t \times q_x$ pour $0 < t < 1$
- Hypothèse de Balducci : ${}_tp_x = \frac{p_x}{1-{}_tp_{x+t}} = \frac{p_x}{p_x + t \times q_x}$ pour $0 < t < 1$
- Hypothèse de mortalité constante : $\mu_{x+\epsilon_1}(t+\epsilon_2) = \mu_x(t)$ pour $0 \leq \epsilon_1, \epsilon_2 \leq 1$

Dans toute la suite du mémoire, nous considérerons la dernière hypothèse, soit une mortalité constante entre deux dates successives d'un an.

A partir de cette hypothèse retenue, nous pouvons en déduire différentes propriétés (avec $0 \leq \xi \leq 1$) :

$$\begin{aligned}
{}_\xi p_x(t) &= \exp\left(-\int_0^\xi \mu_{x+\tau}(t+\tau)d\tau\right) \\
&= \exp\left(-\int_0^\xi \mu_x(t)d\tau\right) \\
&= \exp\left(-\xi\mu_x(t)\right)
\end{aligned}$$

$$\begin{aligned}
1 - {}_\xi q_x(t) &= [\exp(-\mu_x(t))]^\xi \\
&= (1-q_x(t))^\xi
\end{aligned}$$

$$\begin{aligned}
{}_\xi E_x(t) &= \int_{u=0}^\xi L_{x+u,t+u}du \\
&= L_{x,t}\int_0^\xi {}_up_x(t)du \\
&= L_{x,t}\int_0^\xi [p_x(t)]^u\,du \\
&= \frac{-L_{x,t}\left[1-(1-q_x(t))^\xi\right]}{\ln(1-q_x(t))}
\end{aligned}$$

L'espérance de vie restante d'un individu d'âge x durant l'année t est alors (en notant ω l'âge limite) :

$$\begin{aligned}
e_x(t) &= \mathbb{E}[T_x(t)] \\
&= \int_{\xi \geq 0} {}_{\xi}p_x(t) d\xi \\
&= \sum_{k \geq 0}^{\omega} {}_{k}p_x(t) \int_0^1 {}_{\xi}p_{x+k}(t+k) d\xi \\
&= \sum_{k \geq 0}^{\omega} \exp(-k\mu_x(t)) \left[\frac{1 - \exp(-\mu_{x+k}(t+k))}{\mu_{x+k}(t+k)} \right] \\
&= \frac{1 - \exp(-\mu_x(t))}{\mu_x(t)} + \sum_{k \geq 1}^{\omega} \left[\prod_{j=0}^{k-1} \exp(-\mu_{x+j}(t+j)) \right] \frac{1 - \exp(-\mu_{x+k}(t+k))}{\mu_{x+k}(t+k)}
\end{aligned}$$

Nous allons à présent chercher l'estimateur au maximum de vraisemblance pour la force de mortalité. En considérant chaque individu de $L_{x,t}$ vivant au début de l'année t, on note D_i une variable aléatoire dichotomique de Bernouilli tel que :

$$D_i = \begin{cases} 1, & \text{si l'individu } i \text{ décède à l'âge } x \\ 0, & \text{sinon} \end{cases}$$

En notant u_i la fraction de l'année vécue par l'individu i, on a :

$$\sum_{i=1}^{L_{x,t}} D_i = D_{xt} \quad \text{et} \quad \sum_{i=1}^{L_{x,t}} u_i = E_{xt}$$

La contribution de chaque individu à la vraisemblance est alors :

$$\begin{cases} p_x(t) = \exp(-\mu_x(t)), & \text{s'il survit} \\ {}_{u_i}p_x(t)\mu_{x+u_i}(t+u_i) = \mu_x(t)\exp(-u_i\mu_x(t)), & \text{s'il décède à la date } u_i \text{ au cours de l'année } t \end{cases}$$

En supposant une indépendance des individus, la vraisemblance s'écrit alors :

$$\mathcal{L}(\mu_x(t)) = \prod_{j=1}^{L_{x,t}} (\mu_x(t))^{D_i} \exp(-u_i \mu_x(t)) = (\mu_x(t))^{D_{x,t}} \exp(-E_{x,t}\mu_x(t))$$

L'estimateur au maximum de vraisemblance de la force de mortalité est alors $\hat{\mu}_x(t) = \frac{D_{x,t}}{E_{x,t}}$, ce qui correspond au taux de mortalité brut à l'âge x, l'année t.

1.5 Les tables de mortalité

Une table de mortalité permet de décrire l'évolution de la mortalité au sein d'une population regroupant, en général, initialement 100000 individus à l'âge 0 (naissance) et qui indique le nombre de survivant par âge x. Elles permettent de déterminer les probabilités de survie à chaque âge :

$$_tp_x = \frac{L_{x+t}}{L_x}$$

Une réglementation encadre les tables de mortalité utilisées par les assureurs, que ce soit pour tarifer ou provisionner. Il est courant que des tables de la population générale soit utilisées par défaut, mais un assureur peut tout de même utiliser ses propres tables. L'article A335-1 du code des Assurances évoque le contexte dans lequel l'assureur peut utiliser ses propres tables.

Il existe différents types de tables de mortalité.

Les tables de mortalité instantanées

Une table instantanée (ou du moment) est une table qui donne sur une période donnée la mortalité observée. Elles sont généralement construites sur de courtes périodes, entre 1 et 3 ans, et sont utilisées pour les contrats en capital, le paiement n'intervenant qu'une fois.

En France, la table de mortalité instantanée actuellement en vigueur pour les hommes (resp. femmes) est la TH 00-02 (resp. TF 00-02). Ces tables sont homologuées par l'arrêté du 20 décembre 2005 et modifie l'article A.335-1. Elles ont été établies à l'aide de données de la population générale récoltées par

l'INSEE entre 2000 et 2002.

Les tables de mortalité prospectives

Ce type de table donne la mortalité observée et projetée d'un génération (cohorte ayant la même année de naissance). Elles permettent d'observer les caractéristiques démographiques d'une génération, puisque certains facteurs affectent l'espérance de vie future. Ces tables sont apparues dans les années 1990, lorsque l'on a pris conscience que l'allongement de la vie humaine dépendait de l'année de naissance, et, ont alors un intêret certain pour les actuaires lors de la tarification des rentes viagères.

L'arrêté du 01/08/2006 modifie l'article A.335-1 et homologue les tables TGH 05 et TGF 05 comme tables de mortalités prospectives officielles. Elles sont construites avec des données récoltées par l'INSEE entre 1962 et 2000 et des données de la population ayant touchées des rentes entre 1993 et 2005. La construction de ces tables est expliquée par SERANT [23] ou PLANCHET [52].

Les tables de mortalité d'expérience

Ces tables sont établies par les compagnies d'assurances elles-mêmes, en se basant sur la mortalité observée de leur propre portefeuille. Le principal inconvénient est que le nombre d'assurés est assez restreint ce qui provoque des fluctuations conséquentes sur la mortalité.

Chapitre 2

Présentation des données

2.1 *Human Mortality Database*

La "*Human Mortality Database*" a été lancée en 2002 pour fournir les données de mortalité des différents pays. Elle a été crée par le département de démographie de l'Université de Berkeley aux Etats-Unis et le l'institut Max Planck à Rostock en Allemagne. Cette base recense de nombreuses données comme le nombre de naissances, de décès, la taille de population,... dans 37 pays et vise principalement les chercheurs, scientifiques, étudiants ou journalistes. La plupart des auteurs travaillant sur la mortalité ou la longévité utilisent cette base.

Les données que nous utiliserons sont françaises, tirées de cette base, proviennent de l'INSEE ou de l'INED (Institut National d'Etudes Demographiques). Elles rassemblent toutes les publications officielles effectuées. Sur certaines périodes, les nombres de décès sont approximatifs, c'est le cas notamment des périodes de guerre.

Les dates disponibles s'étendent de 1816 à 2012 et concernent des âges compris entre 0 et 110 ans. Les données sont tronquées, tous les individus âgés de plus de 110 ans sont notés "110^+". Pour certains âges les plus extrêmes, aucune observation n'est recensée sur certaines périodes. Ce problème nous amène à travailler avec les données allant de 0 à 105 ans dans ce chapitre. Plus d'informations sur le recensement de ces données ou les estimations faite sur certaines périodes sont disponible sur le site www.mortality.org.

Dans ce mémoire, nous nous servirons de deux bases de données. La première

présente les expositions au risque, notée E_{xt} en fonction de l'âge x et de l'année t, en contenant l'effectif de la population totale au premier janvier de chaque année. La seconde décrit le nombre total de décès, notée D_{xt}, avec x l'âge de l'individu et t l'année.

Nous allons à présent analyser l'évolution de la mortalité en France, de manière à savoir pourquoi les rentes viagères nécessitent la création de tables prospectives ainsi que pour comprendre le phénomène de longévité.

2.2 Evolution de la mortalité en France

Dans le but de comprendre la nécessité d'une analyse dynamique de la démographie, nous pouvons commencer par visualiser la surface de mortalité. Comme nous l'avons vu dans le chapitre précédent, le taux de mortalité instantané se note $\hat{\mu}_{xt} = \frac{D_{xt}}{E_{xt}}$. Les Figures 2.1 et 2.2 montrent les surfaces de mortalité des individus français entre 1816 et 2012, agés de 0 à 100 ans. On aperçoit plusieurs tendances, notamment pendant les périodes de guerres (1914-1918, 1939-1945) ou bien les épidémies (1830, 1850, 1870). Pour mieux les comprendre, nous allons effectuer différentes coupes sur cette surface.

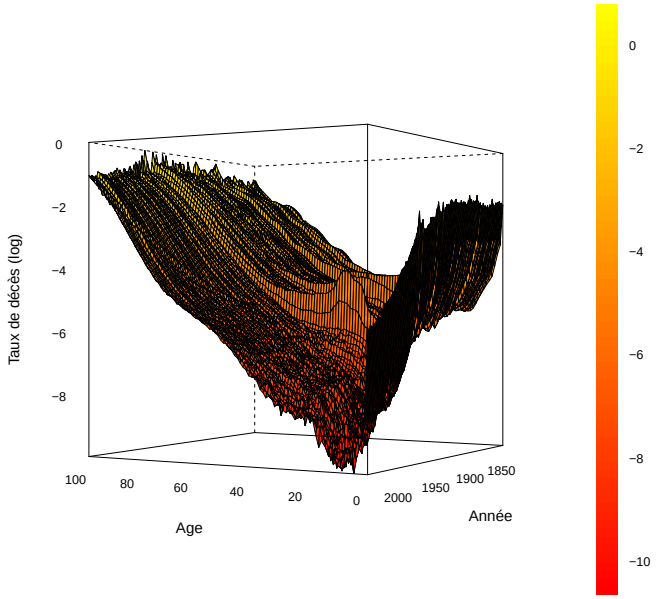

FIGURE 2.1: Surface de mortalité française de 1816 à 2012 pour des femmes de 0 à 100 ans

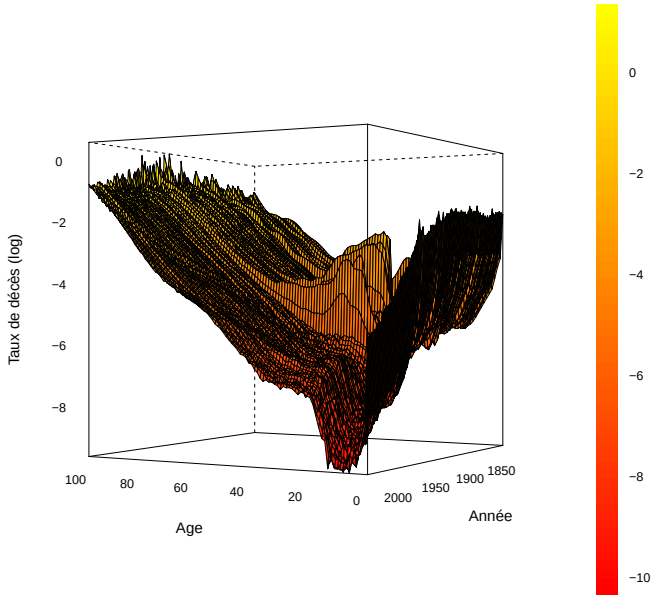

FIGURE 2.2: Surface de mortalité française de 1816 à 2012 pour des hommes de 0 à 100 ans

En fixant une certaine date sur la surface, on étudie la mortalité pour des personnes d'âges différents à un moment donné. Dans le temps, ces taux gardent la même allure comme le montrent les Figures 2.3 et 2.4.
Premièrement, on peut constater une mortalité infantile sur 3-4 premières années de vie, avant de remarquer que ce taux de mortalité décroît. Mais la mortalité subit un bond entre 15 et 20 ans : on parle de "bosse de l'accident"[1], causée par les accidents automobiles des jeunes conducteurs ou les actes inconscients de certains. Cette tendance est plus prononcée chez les hommes que chez les femmes. Ensuite, les taux progressent avec constance selon l'âge. On pourra tout de même noter une assez forte variabilité aux âges supérieurs à 90 ans. Evidemment, les périodes de guerre provoquent des sauts importants dans les taux de décès, notamment pour les hommes.

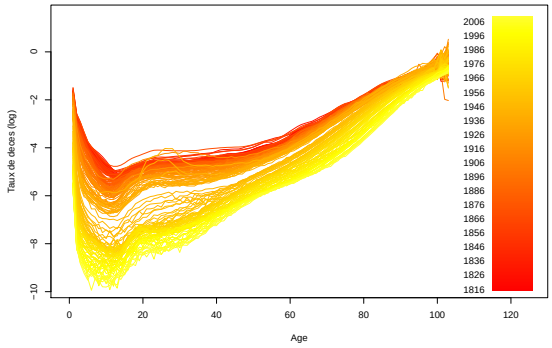

FIGURE 2.3: Coupe longitudinale de la surface de mortalité des femmes

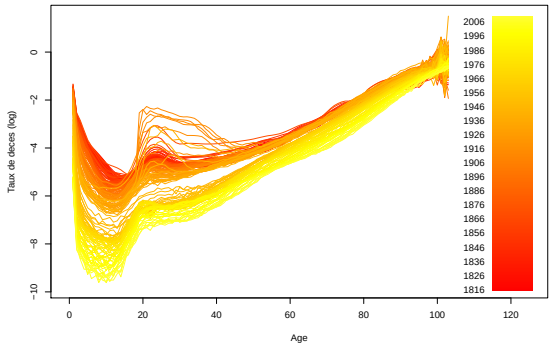

FIGURE 2.4: Coupe longitudinale de la surface de mortalité des hommes

[1]Accident Hump en anglais

En fixant un certain âge sur la surface, on étudie la mortalité pour des personnes d'âge identique sur différentes dates. Entre 1816 et 1900, les taux de décès sont restés stables sur tous les âges, en subissant quelques légères variations. Après cette période, ils n'ont cessé de diminuer.
La chute la plus importante concerne les décès infantiles, qui ont plongé lors des cinquantes dernières années. Auparavant, ils étaient proches, voir supérieurs à ceux de 80 ans. Cette forte baisse implique évidemment la progression de l'espérance de vie à la naissance. Cette diminution a été rapide chez les plus jeunes tandis qu'elle est plus lente chez les personnes agées. On remarquera également de gros sauts sur les périodes de guerres, qui sont moins importants pour les classes d'âges les plus vieilles. (Figures 2.5 et 2.6).

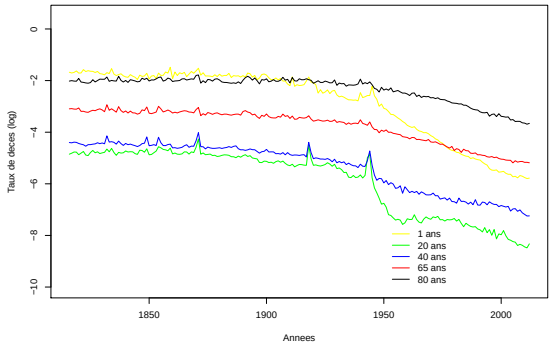

FIGURE 2.5: Taux de décès pour des femmes à des âges différents

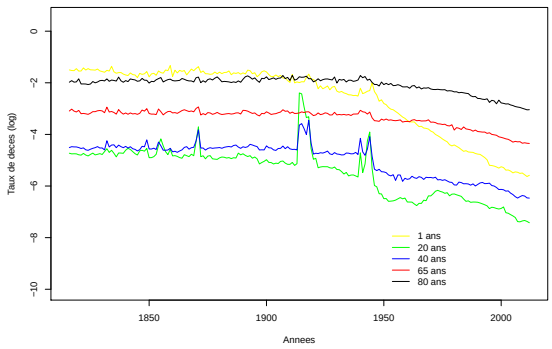

FIGURE 2.6: Taux de décès pour des hommes à des âges différents

A partir de ces surfaces de mortalités, il est également possible d'analyser l'évolution des fonctions de survie dans le temps à la naissance (Figures 2.7 et 2.8). Comme on peut le voir, les probabilités de survie n'ont cessé de s'améliorer au fil des années. Malgré quelques brusques sauts durant les deux guerres mondiales, on note une "rectangulisation" des fonctions de survie, causée principalement par l'amélioration de la survie aux jeunes âges. Cette conséquence implique une augmentation de l'âge médian des décès, avant 1900, 50% des décès avaient lieu avant 60 ans, tandis qu'aujourd'hui, ces 50% ont lieu après 80 ans.

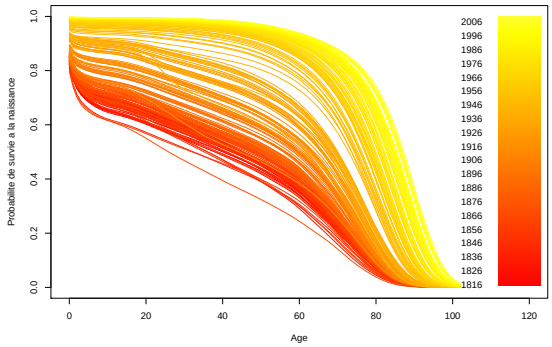

FIGURE 2.7: Probabilités de survie à la naissance des femmes

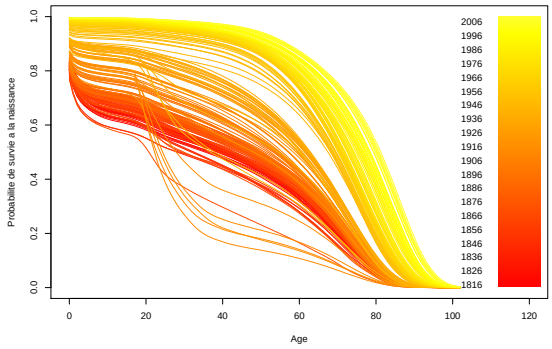

FIGURE 2.8: Probabilités de survie à la naissance des hommes

L'espérance de vie dans le temps à certains âges peut également être analysée (Figures 2.9 et 2.10). En presque 200 ans, l'espérance de vie à la naissance a doublé en France. Elle a suivit une progression rapide depuis 1900, mais a subit quelques sauts aux 19^{e}siècle à cause des différentes épidemies (varioles et choléra) et au siècle dernier à cause des deux guerres.
L'amélioration de l'espérance de vie à 20 ans a été plus plate jusque dans les années 1930 où elle a commencé à augmenter. Un individu de 20 ans a gagné 20 ans d'espérance de vie en 200 ans environ. On peut noter les mêmes sauts que précédemment expliqué.
Ensuite, plus les âges augmentent, moins le gain de l'espérance de vie est grand. En effet, un individu de 40 ans a gagné environ 5 ans de vie, tandis qu'un individu de 65 ans environ 3 ans et un autre de 80 ans, environ 2 ans. L'espérance de vie à âges différents est resté le même jusqu'à la deuxième guerre mondiale, à peu près.

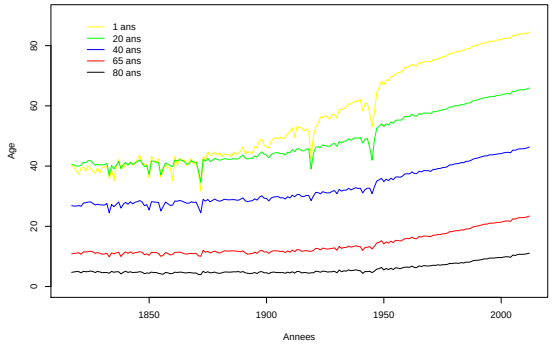

FIGURE 2.9: Evolution de l'espérance de vie des femmes

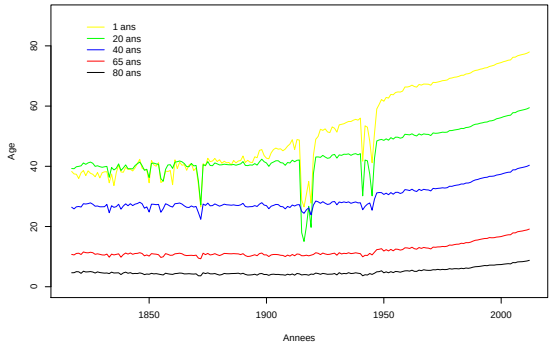

FIGURE 2.10: Evolution de l'espérance de vie des hommes

Cette manière de regarder les surfaces de mortalité nous permet uniquement de nous intéresser aux décès soit à une certaine date, soit à un certain âge. Ici, nous sommes motivés par la longévité des individus. La découpe de la surface de mortalité ne s'effectue plus en fixant un âge ou une date, mais en suivant directement une cohorte. On peut alors comparer les taux de mortalité d'une cohorte née à une certaine date avec les taux de mortalité à cette date (Figure 2.11).

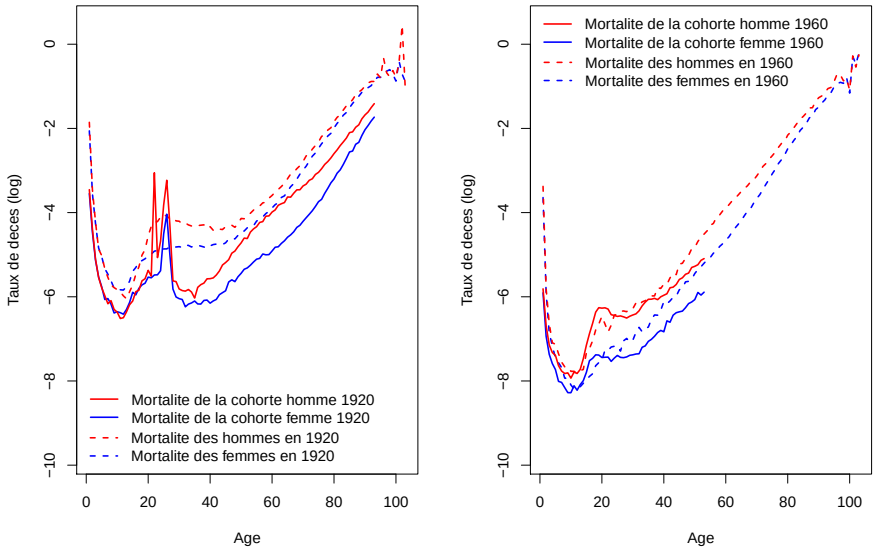

FIGURE 2.11: Comparaison des taux de décès (log) pour différentes cohortes

On constate un certain biais entre ces deux taux, par conséquent, si un assureur tarifie ou provisionne avec le taux de mortalité instantané, cela amènera des incohérences. Il faut donc prédire les taux de mortalité des personnes les plus jeunes pour les années à venir.

Avant de pouvoir prédire la mortalité, il est nécessaire de trouver un modèle adéquat pour celle-ci. La prochaine partie du mémoire recense un état de l'art sur les différents modèles que l'on peut trouver concernant la modélisation et la projection de la mortalité.

Chapitre 3

Etat de l'art sur les modèles de mortalités

Suite à la peste londonienne, on doit à GRAUNT [66] la première table de mortalité. HALLEY [32] fit de même en Pologne en étudiant la ville de Breslau et en classant les individus par classe d'âge. Mais la première table de mortalité utilisée par les actuaires a été établie par MILNE [69] sur la base statistiques de deux paroisses de Carlisle, en Angleterre.
Cependant, ces tables n'étaient basées que sur des observations.

3.1 Les modèles paramétriques

Le premier modèle paramétrique, le plus simpliste par ailleurs, est celui de DE MOIVRE, qui suppose que la durée de vie résiduelle d'un individu d'âge x est uniformément distribué entre 0 et w-x, avec w la durée de vie limite. GOMPERTZ [13] propose plus tard un autre modèle, en supposant que la force de mortalité augmente exponentiellement avec l'âge. Ce dernier est repris par MAKEHAM [119] qui le développe en tenant compte des causes de mortalités indépendantes de l'âge. FORFAR et *al.* [27] proposent une généralisation des modèles de la classe Gompertz-Makeham, mais MELNIKOV et *al.* [123] précisent toutefois que l'utilisation de ces deux lois pour projeter la mortalité n'est pas correcte et risque d'engendrer de sérieux problèmes financiers à un assureur ou un fond de pensions qui s'en serviraient.

PERKS [117], KANNISTO [115] et THATCHER [11] modélisent la force de mortalité par un modèle logistique, qui permet un meilleur ajustement que les

autres modèles, aux âges les plus extrêmes. Cependant, l'estimation des paramètres peut paraître assez compliqué comme l'explique DORAY [1]. BONGAARTS [65] reprend le modèle de THATCHER [11] en rendant les taux de décès non constants dans le temps. Ce dernier est repris par PLANCHET et *al.* [5] pour modéliser la mortalité d'expérience future en l'abscence de données historiques suffisantes à partir d'informations externes.

Ces modèles expliquent la force de mortalité. D'autres sont directement basés sur la probabilité de décès.

C'est le cas de BEARD [106], de BARNETT [16] ou encore d'HELIGMAN et *al.* [71]. Ces derniers suggèrent que la probabilité de décès peut-être représentée par trois composantes différentes avec huit paramètres dépendants du temps : MCNOWN et *al.* [2] se servent alors des processus ARIMA pour projeter ces paramètres.

Les paramètres de ces lois de mortalité sont déterminés selon les données observées, taux de mortalité ou nombre de décès. Or, ces méthodes peuvent être intéressantes dans certains cas, notamment lorsque l'on dispose d'un nombre de données assez restreint, ou bien que l'on regarde la mortalité sur une certaine période. Ici, on s'intéresse au risque de longévité des individus et l'extrapolation de la mortalité par ces modèles n'est pas assez fiable. En revanche, la fiabilité des modèles stochastiques utilisés depuis le début des années 1990 est supérieure.

3.2 Les modèles non-paramétriques

Deux types de modèles non-paramétriques s'opposent, les modèles Age-Période et les modèles Age-Période-Cohorte. Le facteur période explique les effets globaux sévissant sur la population comme l'amélioration des conditions sanitaires ou de la médecine tandis que le facteur cohorte explique les effets historiques, comme les guerres, les pandémies, les habitudes de certaines générations...

3.2.1 Les modèles Age-Période

Le modèle Age-Période le plus connu et utilisé est celui de LEE et CARTER [85]. Ce dernier décompose l'évolution des taux instantanés de mortalité en trois composantes : l'une est spécifique à l'âge tandis que le produit des deux autres décrit l'évolution de la mortalité dans le temps selon l'âge. Ce modèle souffre cependant de plusieurs inconvénients, dont le plus important est de considérer que les résidus du modèle sont homoscédastiques.
Sur les recommandations de BRILLINGER [28], ALHO [74] et WILMOTH [75], BROUHNS et *al.* [73] utilisent le modèle de Lee-Carter pour mettre au point un modèle Log-Poisson, qui permet alors de travailler dans un cadre probabiliste donnant directement une estimation du nombre de décès, pour un âge et une année précise. L'hypothèse d'homoscédasticité des résidus est par ailleurs soulagée.
Cependant, la propriété d'équidispersion de la distribution de Poisson est rarement vérifié. RENSHAW et *al.* ([110], [109] et [111]) préfèrent une distribution de Poisson surdispersée. Pour rajouter une part d'hétérogénéité chez les individus, DELWARDE et *al.* [17] modélisent le nombre de décès par une distribution négative binomiale, ce qui revient à supposer que le paramètre de surdispersion de la distribution de Poisson surdispersée suit une loi Gamma. LI et *al.* [81] font de même en supposant que ce paramètre est spécifique à l'âge.
COSSETTE et *al.* [31] et RENSHAW et *al.* [112] optent eux pour une régression binomiale afin de déterminer les probabilités de décès.

Le modèle de Lee-Carter n'utilise qu'un seul paramètre pour la dimension temporelle, ce qui qualitativement peu s'avérer trop limité et trop rigide. Des termes bilinéaires sont alors rajoutés par RENSHAW et *al.* [109] et BOOTH et *al.* [82]. Les résultats sont meilleurs, mais peut-être pas assez convaincants par rapport à l'ajout de la difficulté supplémentaire, notamment par la modélisation d'une série multivariée lors de la projection ([10]).

Un autre modèle Age-Période, plus récent, est celui de CAIRNS-BLAKE-DOWD (CBD) [78]. Ces derniers utilisent un cas particulier du modèle de PERKS [117], en incluant deux facteurs de période, et via une transformation logit sur les probabilités de décès, lissent l'effet de l'âge. Deux facteurs de périodes permettent la prise en compte de la corrélation des taux de mortalités entre deux âges consécutifs.

En Angleterre, la méthode des P-Splines est très souvent appliquée. DURBAN et *al.* [102] lissent les taux de mortalité via cette technique et utilisent la méthode des B-Splines de EILERS et *al.* [14] pour ajuster les taux de mortalité. Une caractéristique attrayante de cette méthode concerne la projection, qui est une conséquence du processus de lissage.

Plus récemment, des processus AR(1)-ARCH(1) et BEKK-GARCH ont été utilisé respectivement par GIACOMETTI et *al.* [48] et GAO et *al.* [18].

Dans le papier original, le modèle de Lee-Carter implique la projection d'un terme représentant l'évolution temporelle de la mortalité par un processus ARIMA(0,1,0). Cependant, ces types de processus ne permettent pas de prendre en compte les sauts dus à une éventuelle pandémie ou guerre. Dans leur article original, Lee et Carter ont supposé la grippe espagnole de 1918 comme un évènement exceptionnel et ne le prennent pas en compte dans leur modèle. Des processus avec sauts existent et peuvent être utilisés pour la projection de l'indice temporel du modèle de Lee-Carter, on peut citer par exemple les processus de Levy (HAINAULT et *al.* [41]) , le processus Normal Inverse Gaussien (MITCHELL et *al.* [79]), les sauts exponentiels doubles (DENG et *al.* [103]). Bien que les résultats soient convaincants, la calibration des paramètres reste assez compliquée. Le mémoire de KOUADIO et PONCELET [46] recense l'application de différents processus sur des données françaises, avec en plus de ceux cités, le processus de Merton, de Kou et le variance Gamma.

Cette projection de la mortalité via le modèle Lee-Carter se fait en deux étapes : le nombre de décès est d'abord ajusté au modèle puis le terme temporel est ensuite extrapolé via une série temporelle ARIMA. Cette suite d'étapes semble illogique pour CZADO et *al.* [91] qui décide de travailler dans un cadre bayésien, via les méthodes MCMC, en utilisant le modèle log-Poisson. GRECO et *al.* [49] travaillent également avec un modèle bayésien hiérarchique pour comparer la mortalité à travers les différentes régions de l'Italie. LI [67] propose une alternative MCMC au modèle de Lee-Carter lorsque le nombre de données disponibles est restreint.

De manière à prendre en compte une forte hausse (pandémie, guerre) ou chute des taux de décès (découverte médicale sur le cancer, maladies cardio-

vasculaires), certains auteurs préfèrent modéliser non pas le niveau du taux de décès, mais par les changements dans ces taux : le terme du taux de décès de l'année N dépend de celui de l'année N-1., MITCHELL et *al.* [79] utilisent ce type de principe, repris par CHAI et *al.* [121] qui s'essaient à un modèle AR-GARCH, pour au final obtenir une meilleure qualité prédictive que le modèle de Lee-Carter sur un backtest. Cette méthode permet un bien meilleur ajustement aux taux de mortalité passés, mais peut également être appliquée à n'importe quel modèle modélisant la force de mortalité.

3.2.2 Les modèles Age-Période-Cohorte

RENSHAW et *al.* [111] reprennent le modèle de Lee-Carter pour y ajouter l'effet de cohorte à travers de nouveaux paramètres, dont l'évolution est difficilement interprétable. Ce modèle obtient un meilleur ajustement dans les pays où un effet de cohorte est observé dans le passé, mais souffre d'un manque de robustesse. CURRIE [63] se sert de ce même modèle et le simplifie pour supprimer l'instabilité d'ajustement qui pouvait résider dans les nouveaux paramètres et suppose indépendant les facteurs d'âge, de temps et de cohorte. Cette amélioration peut sembler excessive car elle suppose que la baisse de mortalité est la même quelque soit l'âge de la personne.
CAIRNS *al.* [62] rajoute également un terme au modèle CBD pour capturer l'effet de cohorte. Cependant, ce modèle est intéressant aux âges élevés seulement, car lorsqu'il est utilisé pour toutes les classes d'âges, sa qualité d'ajustement est faite.
PLAT [104] développe un nouveau modèle stochastique principalement fondé sur celui de CBD ou de Currie permettant un meilleur ajustement, une corrélation dans la structure du modèle et une application pour toutes les classes d'âges sauf les extrêmes. En développant ce modèle, BORGER *al.* [98] utilise le logit des probabilités de décès et rajoute un terme pour prendre en compte l'effet des âges élevés.

D'un point de vue pratique, les modèles de Age-Période-Cohorte souffrent du fait que les données doivent être étendues sur de longues périodes. Tout comme les modèles Age-Période, ils considèrent également une force de mortalité constante par morceaux, soit entre les âges, et font des prévisions sur l'évolution future grâce aux observations passées. Les modèles stochastiques en temps continu sont alors plus réalistes mais beaucoup plus compliqué.

3.3 Les modèles stochastique en temps continu

Pour être plus réaliste, certains auteurs modélisent directement l'intensité de mortalité en supposant que cette dernière est un processus stochastique adapté à une filtration.

La modélisation des décès peut se faire via un processus de comptage. DAHL [92] utilise un processus de diffusion pour modéliser l'intensité de mortalité. Ce dernier modèle n'est pas très réaliste car tous les changements de l'intensité de mortalité sont similaires et touchent toute la cohorte, alors qu'elle devrait être affectée par plusieurs facteurs de mortalité de manières différentes. Mis à part les processus de comptage ou de diffusion, ceux à structure affine sont également couramment utilisés. DAHL [92] postule que les modèles de taux d'intêret peuvent être appliqués pour modéliser la mortalité et emprunte le modèle de Cox-Ingersoll-Ross. Les modèles de taux forward sont également utilisables.
L'inconvénient de ces modèles, bien que leurs résultats soient très satisfaisants, est qu'ils sont très sophistiqués et qu'ils nécessitent la connaissance d'outils mathématiques puissants. L'auteur peut retrouver une bonne revue et explication de ces modèles chez BEN DBABIS [25] ou PLANCHET [53].

Chapitre 4

Modéliser la mortalité

Dans toute cette partie et les suivantes, nous notons :

- Le vecteur contenant les âges : $x = \{x_1, \ldots, x_m\}$
- Le vecteur contenant les années : $t = \{t_1, \ldots, t_n\}$
- $\boldsymbol{\alpha} = \{\alpha_{x_1}, \ldots, \alpha_{x_m}\}$
- $\boldsymbol{\beta} = \{\beta_{x_1}, \ldots, \beta_{x_m}\}$
- $\boldsymbol{\kappa} = \{\kappa_{t_1}, \ldots, \kappa_{t_n}\}$
- $\boldsymbol{\phi} = \{\phi_{x_1}, \ldots, \phi_{x_m}\}$
- D_{xt} : le nombre de décès d'âge x durant l'année t
- E_{xt} : l'exposition au risque à l'âge x durant l'année t
- $\mu_x(t)$: la force de mortalité à l'âge x durant l'année t

De plus, la période de temps que nous considérons pour travailler s'étale de 1950 à 2012. Nous considérons cette période pour plusieurs raisons :

- La plupart des auteurs prennent ces mêmes périodes (voir à partir de 1970)
- Les taux de mortalité ont beaucoup baissé depuis 1950 et la qualité des données n'est pas complétement sûre avant cette date
- Aujourd'hui, les causes de décès les plus communes sont les maladies cardiaques ou les cancers, ce qui n'était pas le cas auparavant

Dans ce chapitre, nous aborderons trois modèles de mortalité stochastique age-période, le modèle de Lee-Carter, le modèle log-Poisson et le modèle Binomiale négatif. Nous introduirons également la notion d'hétérogénéité dans la population. L'ajustement des modèles aux données est également un point important avant de passer à la projeciton de la mortalité. Puis, nous traiterons le point sensible de la fermeture des tables avec la modélisation des décès aux âges extrêmes.

4.1 Le modèle de Lee-Carter

Le modèle proposé par Lee et Carter [85] est assez simple, il spécifie que la force de mortalité a une forme, tel que :

$$\ln \hat{\mu}_x(t) = \alpha_x \ + \beta_x \kappa_t \ + \epsilon_{xt}$$

avec les contraintes :

$$\sum_{x=x_1}^{x_n} \beta_x = 1 \qquad \sum_{t=t_1}^{t_m} \kappa_t = 0$$

Sans elles, le modèle n'est pas identifiable, il existerai une quantité infinie de valeurs possibles pour β_x et κ_t.

Sur l'échelle logarithmique, le taux de mortalité à l'âge x pour l'année t est donc décomposé en la somme d'une composante spécifique à l'âge x et d'un produit entre un paramètre temporel décrivant l'effet de période et un paramètre spécifique à l'âge pur. L'interprétation des paramètres est simple :

- α_x : correspond à la forme de la force de mortalité au cours du temps (sur l'échelle logarithmique).
- β_x : correspond à l'écart de la force de mortalité par rapport au comportement moyen α_x car :

$$\frac{d}{dt} \ln \mu_x(t) = \beta_x \frac{d}{dt} \kappa_t$$

Les âges x pour lesquels β_x est grand seront très sensible à l'évolution temporelle.

- κ_t : décrit l'évolution de la mortalité au cours du temps.

- ϵ_{xt} : sont les résidus du modèle, supposés homoscédastique, de moyenne 0 et de variance σ_ϵ^2

4.1.1 Estimation des paramètres

Etant donné qu'il n'y a pas de variables observables dans le terme de droite du modèle, une régression linéaire simple est impossible. Les paramètres s'estiment alors via la méthode des moindres carrés ordinaires, en résolvant le problème d'optimisation :

$$(\hat{\alpha}_x, \hat{\beta}_x, \hat{\kappa}_t) = \arg\min_{\alpha_x, \beta_x, \kappa_t} \mathcal{O}_{MC}$$

avec

$$\mathcal{O}_{MC} = \sum_{x=x_1}^{x_m} \sum_{t=t_1}^{t_n} (\ln \hat{\mu}_{xt} - \alpha_x - \beta_x \kappa_t)^2$$

On commence par estimer α_x :

$$\frac{d}{d\alpha_x} \mathcal{O}_{MC} = 0 \iff \sum_{t=t_1}^{t_n} \ln \hat{\mu}_{xt} = (t_n - t_1 + 1)\alpha_x + \beta_x \sum_{t=t_1}^{t_n} \kappa_t$$

La contrainte $\sum_{t=t_1}^{t_n} \kappa_t = 0$ amène à :

$$\hat{\alpha}_x = \frac{1}{t_n - t_1 + 1} \sum_{t=t_1}^{t_n} \ln \hat{\mu}_{xt}$$

L'estimation de β_x et κ_t se fait via le premier terme de la décomposition en valeurs singulières de la matrice $\ln \hat{m}_x(t) - \hat{\alpha}_x$. On note $\mathcal{M}$ la matrice des taux de décès :

$$\mathcal{M} = \begin{pmatrix} m_{x_1}(t_1) & \cdots & m_{x_1}(t_n) \\ \vdots & \ddots & \vdots \\ m_{x_m}(t_1) & \cdots & m_{x_m}(t_n) \end{pmatrix}$$

Le modèle de Lee-Carter est donc ajusté pour être le plus proche possible de $\mathcal{M}$.
On définit également la matrice $\mathcal{Z}$ tel que :

$$\mathcal{Z} = \ln \hat{\mathcal{M}} - \hat{\boldsymbol{\alpha}} = \begin{pmatrix} z_{x_1 t_1} = \ln \hat{m}_{x_1}(t_1) - \hat{\alpha}_{x_1} & \cdots & z_{x_1 t_n} = \ln \hat{m}_{x_1}(t_n) - \hat{\alpha}_{x_1} \\ \vdots & \ddots & \vdots \\ z_{x_m t_1} = \ln \hat{m}_{x_m}(t_1) - \hat{\alpha}_{x_m} & \cdots & z_{x_m t_n} = \ln \hat{m}_{x_m}(t_n) - \hat{\alpha}_{x_m} \end{pmatrix}$$

On cherche alors $\hat{\beta}_x$ et $\hat{\kappa}_t$ tels qu'ils minimisent :

$$\mathcal{O}'_{MC}(\beta, \kappa) = \sum_{x=t_1}^{x_m} \sum_{t=t_1}^{t_n} (z_{xt} - \beta_x \kappa_t)^2$$

La minimisation de cette fonction s'effectue par une décomposition en valeurs singulières de $\mathcal{Z}$.

On définit u_i comme le i-ème vecteur propre de la matrice $\mathcal{Z}^{\mathrm{T}}\mathcal{Z}$ de dimensions $(t_n - t_1 + 1) \times (t_n - t_1 + 1)$ correspondant à la valeur propre λ_i.
De même, on définit v_i comme le i-ème vecteur propre de la matrice $\mathcal{Z}\mathcal{Z}^{\mathrm{T}}$ de dimensions $(x_m - x_1 + 1) \times (x_m - x_1 + 1)$ correspondant à la valeur propre λ_i.

On a alors pour $\lambda_i \neq 0$:

$$v_i = \frac{1}{\sqrt{\lambda_i}} \mathcal{Z} u_i$$

$$u_i = \frac{1}{\sqrt{\lambda_i}} \mathcal{Z}^{\mathrm{T}} v_i$$

Ce qui nous amène à :

$$\mathcal{Z}u_i = v_i\sqrt{\lambda_i}$$

Puis, en multipliant des deux côtés par u_i^{T} et en sommant sur les valeurs propres de $\mathcal{Z}\mathcal{Z}^{\mathrm{T}}$, on obtient :

$$\mathcal{Z}\left(\sum_{i\geq 1} u_i u_i^T\right) = \sum_{i\geq 1}\sqrt{\lambda_i}v_i u_i^T$$

Ce qui mène à $\mathcal{Z} = \sum_{i\geq 1}\sqrt{\mu_i}v_1 u_1^T$, appellé décomposition en valeurs singulières.

Lorsque la première valeur propre μ_1 est bien supérieure aux autres, la meilleure approximation de $\mathcal{Z}$ est alors donnée par :

$$\mathcal{Z} \approx \sqrt{\lambda_1}v_1 u_1^T$$

On peut en déduire que :

$$\hat{\beta}_x = \frac{v_1}{\sum_{j=t_1}^{x_m-x_1+1} v_{1j}} \text{ et } \hat{\kappa}_t = \sqrt{\lambda_1}\left(\sum_{j=t_1}^{x_m-x_1+1} v_{1j}\right)u_1$$

Sous la contrainte que $\sum_{j=t_1}^{x_m-x_1+1} v_{1j} \neq 0$, ce qui implique le respect des contraintes du modèle.

La décomposition en valeurs singulières n'est pas l'unique méthode d'estimation des paramètres :

- La décomposition en valeurs singulières se fait uniquement sur des matrices rectangulaires, la mise en place d'un algorithme de Newton-Raphson est envisageable et permet de soulager cet inconvénient.

- WILMOTH [75] utilise la méthode des moindres carrés pondérés de manière à pouvoir attribuer un poids (nombre de décès observés) aux observations. De cette manière, les couples (x, t) dont le taux bruts instantané est peu varibale aura moins d'importance que les plus volatiles.

- Les paramètres ne sont pas complétement lisses (surtout β_x), ce qui peut causer des problèmes lors de mise en place d'une table de mortalité. Les probabilités de décès doivent être strictement croissantes dans le temps et ne doivent pas subir de quelconque variation soudaine. DELWARDE et *al.* [100] remedie à ce problème via la méthode des moindres carrés pénalisés.

4.1.2 Ajustement de κ_t

Pour reproduire exactement le nombre de décès lors de l'année t, il est nécessaire de réajuster les paramètres. Pour cela, nous re-calibrons $\hat{\kappa}_t$ tel que :

$$\sum_{x=x_1}^{x_m} D_{xt} = \sum_{x=x_1}^{x_m} E_{xt} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t)$$

Le nouvel estimateur $\hat{\hat{\kappa}}_t$ vérifie alors :

$$\sum_{x=x_1}^{x_m} E_{xt} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\hat{\kappa}}_t) - D_{xt} = 0$$

Une méthode de Newton-Raphson peut donc être mise en place. En notant θ le paramètre à estimer, f la fonction à minimiser et f' sa dérivée, k la k-ième itération de l'algorithme, on fixe une valeur initiale pour $\theta^{(0)}$, puis on obtient :

$$\theta^{(k+1)} = \theta^{(k)} - \frac{f(\theta^{(k)})}{f'(\theta^{(k)})}$$

En fixant $\hat{\hat{\kappa}}_t^{(0)} = \hat{\kappa}_t$, l'estimation du nouveau paramètre se fait en quelques itérations :

$$\hat{\hat{\kappa}}_t^{(k+1)} = \hat{\hat{\kappa}}_t^{(k)} - \frac{\sum_{x=x_1}^{x_m} E_{xt} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\hat{\kappa}}_t^{(k)}) - D_{xt}}{\sum_{x=x_1}^{x_m} E_{xt} \hat{\beta}_x \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\hat{\kappa}}_t^{(k)})}$$

Il faut ensuite procéder à une transformation pour vérifier les contraintes du modèle :

$$\hat{\alpha}_x \leftarrow \hat{\alpha}_x + \hat{\beta}_x \hat{\hat{\bar{\kappa}}}_t$$

$$\hat{\hat{\kappa}}_t \leftarrow (\hat{\hat{\kappa}}_t - \hat{\hat{\bar{\kappa}}}_t) \times \sum_{x=x_1}^{x_m} \hat{\beta}_x$$

avec $\hat{\hat{\bar{\kappa}}}_t$ la moyenne des $\hat{\hat{\kappa}}_t$

Ici, les β_x du modèle sont positif quelque soit l'âge. Si ce paramètre vient à changer de signe et passer négatif, cette méthode de réajustement offrirait plusieurs solutions pour κ_t, ce qui n'est pas cohérent. D'autres méthodes existent :

- LEE [114] réajuste κ_t en se servant de l'espérance de vie future des individus à un certain âge, ce qui permet de n'utiliser ni l'exposition au risque, ni le nombre de décès observés. Ici encore, le problème de signe sur β_x persiste.
- BOOTH et *al.* [82] suppose que le nombre de décès observés suit une distribution de Poisson puis, maximime la fonction de vraisemblance par rapport à κ_t (ce qui revient à minimiser la déviance). Le problème de réajustement de κ_t n'existe plus.

La Figure 4.1 représente les paramètres du modèle, qui sont conformes aux observations démographiques faite dans le Chapitre 2. Comme on peut le voir, $\boldsymbol{\alpha}$ traduit la croissance des taux de décès avec l'âge, la forme de $\boldsymbol{\beta}$ montre que l'amélioration relative à la mortalité est la plus importante sur la période 0-20 ans puis aux alentours de 80 ans et, $\boldsymbol{\kappa}$ est décroissant ce qui montre bien que le niveau général de la mortalité chute dans le temps.

La principale limite du modèle de Lee-Carter est l'hypothèse d'homoscédasticité des résidus, nécessaire lors de l'estimations par moindres carrés. Cette hypothèse n'est pas vérifiée en pratique car la variance du logarithme des taux de décès croît avec l'âge, à cause de la baisse de la population vivante aux âges extrêmes. Cette forte variabilité est également présente lors des toutes premières années de vie des individus. La Figure 4.2 montre la variance des

résidus du modèle selon l'âge. De plus, il peut être intéressant de modifier l'approche de ce modèle de manière à obtenir des estimateurs de type "maximum de vraisemblance" pour profiter des propriétés avantageuses de ceux-là, comme la convergence des estimateurs, leurs efficacité et normalité asymptotique...

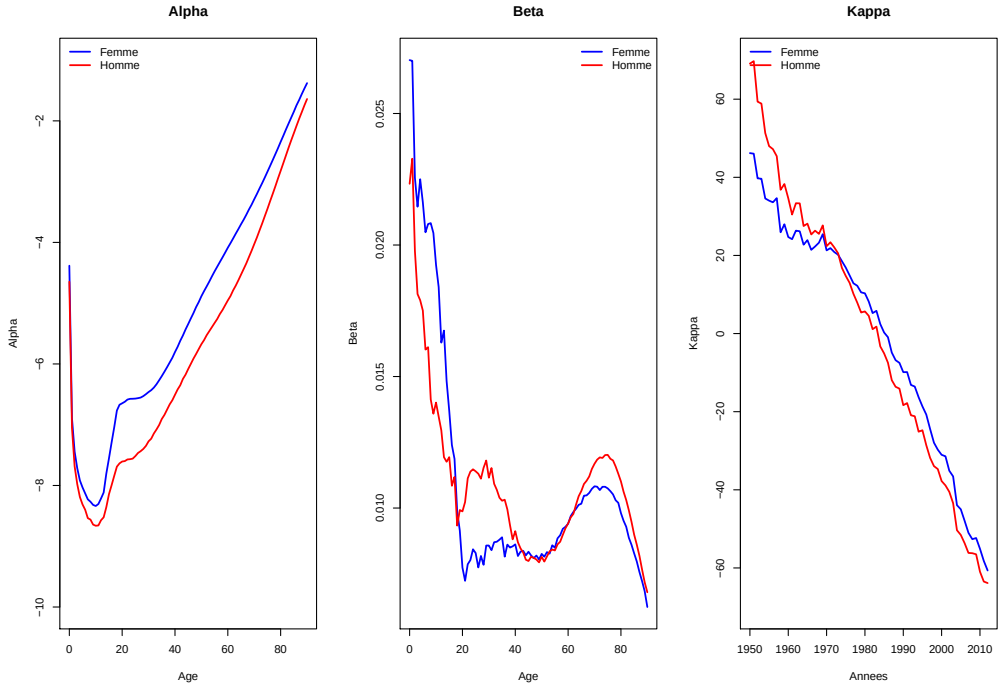

FIGURE 4.1: Paramètres du modèle de Lee-Carter

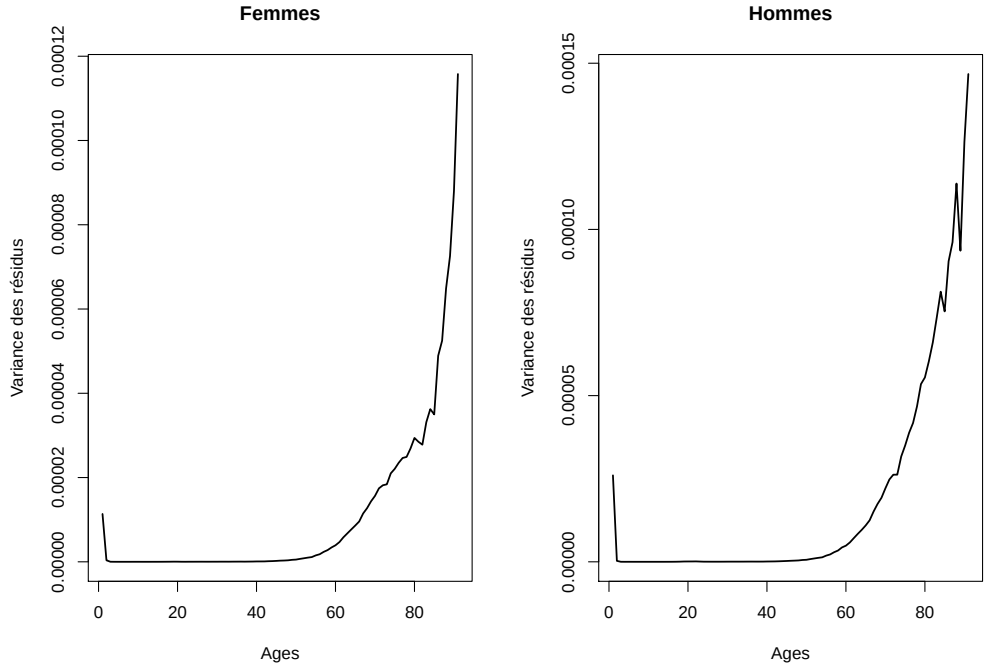

FIGURE 4.2: Variance des résidus du modèle de Lee-Carter selon l'âge

4.2 Le modèle Log-Poisson

BRILLINGER [28] estime qu'il «serait satisfaisant d'approcher la distribution du nombre de morts conditionnellement à la taille de la population, par une loi de Poisson de moyenne proportionnelle à la taille de la population». De manière à améliorer la projection dans le modèle originale de Lee-Carter, WILMOTH [75] et ALHO [74] recommandent également l'utilisation de cette distribution. La méthode mise au point par BROUHNS et *al.* [73] repose sur cette hypothèse :

$$D_{xt} \sim \mathcal{P}(E_{xt}\mu_{xt})$$

avec μ_{xt} = $\exp(\alpha_x + \beta_x\kappa_t)$ la force de mortalité à l'âge x et l'année t. Les paramètres ont la même signification et les mêmes contraintes que le modèle original de Lee-Carter, $\sum_{x=x_1}^{x_m} \beta_x = 1$ et $\sum_{t=t_1}^{t_n} \kappa_t = 0$.

4.2.1 Estimation des paramètres

La fonction de log-vraisemblance du modèle se note :

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}, d_{xt}) &= \ln \prod_{x,t} f(d_{xt}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}) \\
&= \ln \prod_{x,t} \exp(-E_{xt}\mu_{xt}) \times \frac{(E_{xt}\mu_{xt})^{d_{xt}}}{d_{xt}!} \\
&= \sum_{x,t} [-E_{xt}\mu_{xt} + d_{xt} \ln(E_{xt}\mu_{xt}) - \ln d_{xt}!]
\end{aligned}$$

L'estimation des paramètres s'obtiant en résolvant le problème d'optimisation de cette fonction. A première vue, la résolution de ce problème peut sembler trivial. Pourtant, la présence du terme bilinéaire $\beta_x\kappa_t$ rend impossible l'utilisation d'une régression poissonienne classique via un simple algorithme de Newton-Raphson. Il faut alors utiliser l'algorithme de GOODMAN [84], qui est le premier a proposer une méthode itérative pour estimer les modèles log-linéaire contenant des termes bilinéaires.

En notant θ les paramètres à estimer, f la fonction à optimiser, f' et f'' respectivement la dérivée première et la dérivée seconde de f et k la k-ième itération de l'algorithme, le paramètre θ_j est estimé par :

$$\hat{\theta}_j^{(k+1)} = \hat{\theta}_j^{(k)} - \frac{f'(\hat{\theta}_j^{(k)})}{f''(\hat{\theta}_j^{(k)})}$$

Il est évidemment nécessaire de fixer une valeur d'initialisation des paramètres. Ici, nous prenons $\alpha_x^{(0)} = 0$, $\beta_x^{(0)} = 1$ et $\kappa_t^{(0)} = 0$. L'algorithme est le suivant :

Etape 1 : Mise à jour des α_x

$$\begin{aligned}
\hat{\alpha}_x^{(k+1)} &= \hat{\alpha}_x^{(k)} - \frac{\displaystyle\sum_{t=t_1}^{t_n} \left[D_{xt} - E_{xt}\exp(\hat{\alpha}_x^{(k)} + \hat{\beta}_x^{(k)}\hat{\kappa}_t^{(k)})\right]}{\displaystyle -\sum_{t=t_1}^{t_n} \left[E_{xt}\exp(\hat{\alpha}_x^{(k)} + \hat{\beta}_x^{(k)}\hat{\kappa}_t^{(k)})\right]} \\
\hat{\beta}_x^{(k+1)} &= \hat{\beta}_x^{(k)}, \quad \hat{\kappa}_t^{(k+1)} = \hat{\kappa}_t^{(k)}
\end{aligned}$$

Etape 2 : Mise à jour des κ_t

$$\begin{aligned}
\hat{\kappa}_t^{(k+2)} &= \hat{\kappa}_t^{(k+1)} - \frac{\displaystyle\sum_{x=x_1}^{x_m} \hat{\beta}_x^{(k+1)} \left[D_{xt} - E_{xt}\exp(\hat{\alpha}_x^{(k+1)} + \hat{\beta}_x^{(k+1)}\hat{\kappa}_t^{(k+1)})\right]}{\displaystyle -\sum_{x=x_1}^{x_m} (\hat{\beta}_x^{(k+1)})^2 \left[E_{xt}\exp(\hat{\alpha}_x^{(k+1)} + \hat{\beta}_x^{(k+1)}\hat{\kappa}_t^{(k+1)})\right]} \\
\hat{\alpha}_x^{(k+2)} &= \hat{\alpha}_x^{(k+1)}, \quad \hat{\beta}_x^{(k+2)} = \hat{\beta}_x^{(k+1)}
\end{aligned}$$

Etape 3 : Contraintes

$$\begin{aligned}
\hat{\alpha}_x^{(k+2)} &= \hat{\alpha}_x^{(k+2)} + \bar{\kappa}_t^{(k+2)} \times \hat{\beta}_x^{(k+2)} \\
\hat{\kappa}_t^{(k+2)} &= \hat{\kappa}_t^{(k+2)} - \bar{\kappa}_t^{(k+2)}
\end{aligned}$$

Etape 4 : Mise à jour des β_x

$$\hat{\beta}_x^{(k+3)} = \hat{\beta}_x^{(k+2)} - \frac{\sum_{t=t_1}^{t_n} \hat{\kappa}_x^{(k+2)} \left[D_{xt} - E_{xt} \exp(\hat{\alpha}_x^{(k+2)} + \hat{\beta}_x^{(k+2)} \hat{\kappa}_t^{(k+2)}) \right]}{-\sum_{t=t_1}^{t_n} (\hat{\kappa}_x^{(k+2)})^2 \left[E_{xt} \exp(\hat{\alpha}_x^{(k+2)} + \hat{\beta}_x^{(k+2)} \hat{\kappa}_t^{(k+2)}) \right]}$$

$$\hat{\alpha}_x^{(k+3)} = \hat{\alpha}_x^{(k+2)}, \quad \hat{\kappa}_t^{(k+3)} = \hat{\kappa}_t^{(k+2)}$$

Etape 5 : Contraintes

$$\hat{\beta}_x^{(k+3)} = \frac{\hat{\beta}_x^{(k+3)}}{\sum_{x=x_1}^{x_m} \hat{\beta}_x^{(k+3)}}$$

$$\hat{\kappa}_t^{(k+3)} = \hat{\kappa}_t^{(k+3)} - \bar{\kappa}_t^{(k+3)}$$

Malgré leur préconisation de stopper l'algorithme lorsque l'augmentation de la valeur de la fonction de log-vraisemblance est supérieure à 10^{-10} (ce que nous faisons ici), BROUHNS et *al.* [73] l'arrêtent lorsque cette valeur est supérieur à 10^{-6}.

Pour une écriture sous forme matricielle de cette algorithme, le lecteur peut se référer à DELWARDE [10].

Ici, étant donné que l'on modélise directement le nombre de décès, il n'est pas nécessaire de procéder à un réajustement des κ_t comme dans le modèle original de Lee-Carter. La tendance des paramètres obtenus (Figure 4.3) est similaire à ceux du modèle précédent et les mêmes remarques peuvent être faites.

En considérant que le nombre de décès peut être approché par une distribution de Poisson, on considère ici que la variance des décès équivaut à leurs moyenne, d'après la propriété d'équidispersion ($\mathbb{V}ar[D_{xt}] = \mathbb{E}[D_{xt}]$) de la loi de Poisson. Cette propriété est très criticable étant donné qu'elle n'est que très rarement respectée, comme le montre la Figure 4.4. La prise en compte d'un paramètre de dispersion semble nécessaire.

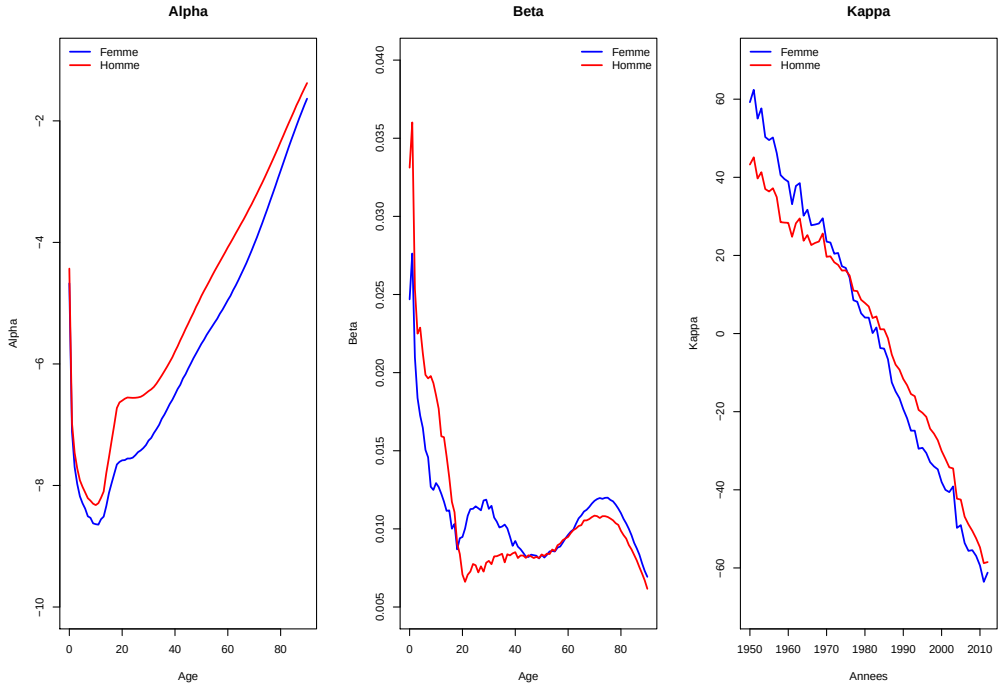

FIGURE 4.3: Paramètres du modèle de log-Poisson

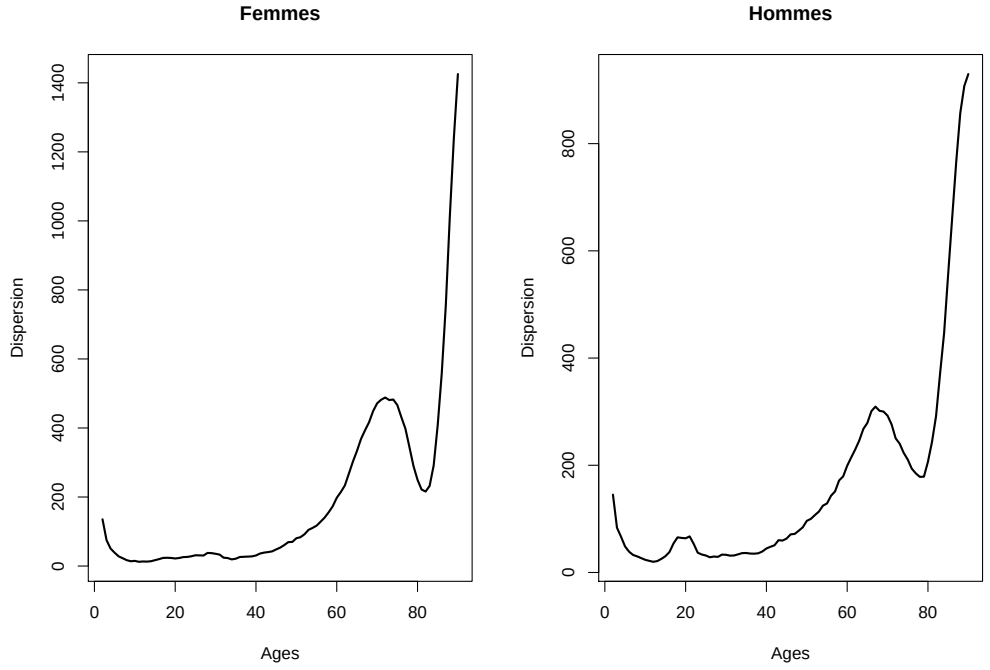

FIGURE 4.4: Dispersion des décès selon l'âge

4.3 La dispersion dans les données

Nous avons vu que le modèle log-Poisson de BROUHNS et *al.* [73] est une alternative cohérente au modèle de Lee-Carter. Mais, en cas de données hétérogènes, la propriété d'équidispersion de la loi de Poisson est trop restrictive. Une manière de soulager le modèle est de considérer une distribution de Poisson sur-dispersée pour le nombre de décès :

$$\mathbb{V}ar[D_{xt}] = \phi\mathbb{E}[D_{xt}]$$

avec ϕ un paramètre qui mesure le degrès de surdispersion. Si celui-ci est égal à 1, on retombe sur une distribution de Poisson standard. Cette méthode a été utilisé par RENSHAW et *al.* ([110], [109] et [111]).

En général, ce paramètre est estimé par :

$$\phi = \frac{\chi^2}{(n-p)}$$

avec $\chi^2 = \sum_{x,t} \frac{(D_{xt} - \hat{D}_{xt})^2}{\hat{D}_{xt}}$, la statistique du χ^2 de Pearson, n le nombre d'observations et p le nombre de paramètres. Ici, nous avons alors $\phi = 6,24$ pour les femmes et $\phi = 11,41$ pour les hommes. Cette technique possède néanmoins un inconvénient majeur, elle n'autorise aucune inférence sur le paramètre de dispersion. Il est difficile de calculer la variance de l'estimateur et donc sa significativité.

Pour être sûr de la significativité de la sur-dispersion, nous pouvons appliquer un test sur nos données. DEAN [21] a proposé différents tests statistiques permettant d'évaluer la surdispersion liée à un modèle de Poisson. Nous testerons l'un d'entre eux :

$$T_1 = \frac{\sum_{x,t} \left[(D_{xt} - \hat{D}_{xt})^2 - D_{xt} \right]}{\sqrt{2 \sum_{x,t} \hat{D}_{xt}^2}} \sim \mathcal{N}(0,1)$$

On teste ici H_0 : "On ne rejette pas le modèle de Poisson" contre H_1 : "On rejette le modèle de Poisson". La valeur de la statistique T_1 est notée dans la Table 4.1. Les p-values obtenues sont inférieures à 10^{-6}, ce qui nous fait rejeter l'hypothèse de Poisson et nous indique la présence de surdispersion dans les données, comme l'on pouvait s'en douter.

Statistique	Femme	Homme
T_1	269,83	541,90

TABLE 4.1: Valeurs de la statististique du test de dispersion

Cette sur-dispersion dans les données de mortalité est causée par la présence d'un degré d'hétérogénéité chez les individus. En effet, dans une population, chaque personne de même âge et de même sexe peut présenter un risque de mortalité différent.

4.4 Les facteurs d'hétérogénéité

L'identification et la prise en compte de ces facteurs sont utiles lors de la projection de la mortalité. Ces facteurs d'hétérogénéité peuvent être observables tandis que d'autres ne le sont pas.

4.4.1 Les facteurs d'hétérogénéité observables

- Les facteurs biologiques ou physiologiques (âge, sexe, ...)
- Les caractéristiques de l'environnement de vie (climat, pollution, conditions sanitaires,...)
- L'emploi exercé (probabilités d'accidents plus ou moins fortes)
- Le mode de vie (alcool, tabagisme, drogue, activité physique, ...)
- Les conditions de santé actuelles (antécédents médicaux familiaux, statut social, ...)

Mis à part le deuxième point, qui touche la plupart de la population d'un pays (c'est notamment pour cela que l'on considère des zones géographiques pour établir des tables de mortalité), les autres concernent les individus eux-même et peuvent être observés : on parle de sélection individuelle d'un risque.

Dans la grande majorité des cas, en assurance sur la vie, l'assureur va appliquer une sélection aux personnes qui sollicitent une assurance pour savoir si le risque de mortalité de l'individu est cohérent avec ses hypothèses.
Cette sélection a pour but de prévenir le risque d'anti-sélection qui correspond à l'écart d'information entre l'assuré et l'assureur, sur le risque que porte l'assuré. La plupart du temps, l'évaluation de ce risque est effectuée via un questionnaire, voir même une évaluation médicale. En cas de mensonges de la part de l'assuré (qui viole ainsi l'article L113-2 du code des Assurances), il tomberait sous le coup des articles L113-8 et L113-9 qui définissent les conséquences de cet acte.

4.4.2 Les facteurs d'hétérogénéité inobservables

Les facteurs inobservables sont souvent comportementaux (attitude de l'individu envers sa propre santé, problèmes médicaux non diagnostiqués, ...).
Ces facteurs peuvent être modélisés dans une approche continue, via une variable aléatoire strictement positive, appelé fragilité[1], ayant pour but d'inclure tous les facteurs inobservables influant sur la mortalité. La fragilité est définit comme une quantité strictement positive, dont le niveau exprime le risque de mortalité inobservable. Un individu ayant une fragilité supérieure à la moyenne décèdera, en moyenne, plus tôt qu'un autre, et inversement, un individu ayant une fragilité inférieure à la moyenne décèdera, en moyenne, plus tard qu'un autre.

4.4.3 Quelles distributions pour la fragilité ?

Pour modéliser cette fragilité, toute distribution de probabilité ayant un support strictement positif peut être utilisée.

La distribution la plus répandue dans les modèles de fragilité est la distribution Gamma ([35], [64] ou [3]).

Soit $\mathcal{Z} \sim \Gamma(\delta, \theta)$, alors la fonction de probabilité de $\mathcal{Z}$ est :

$$f(z) = \frac{\theta^{\delta} z^{\delta-1}}{\Gamma(\delta)} e^{-\delta z}$$

[1]frailty en anglais

Avec $z > 0$, $\delta > 0$ représentant le paramètre d'intensité, $\theta > 0$ représentant le paramètre d'échelle et $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt$. De plus,

$$\mathbb{E}[\mathcal{Z}] = \frac{\delta}{\theta} \quad \text{et} \quad \mathbb{V}ar[\mathcal{Z}] = \frac{\delta}{\theta^2} \quad \text{et} \quad \mathbb{CV}[\mathcal{Z}] = \frac{\sqrt{\mathbb{V}ar[\mathcal{Z}]}}{\mathbb{E}[\mathcal{Z}]} = \frac{1}{\sqrt{\delta}}$$

Le coefficient de variation montre que δ joue un rôle de mesure du niveau d'hétérogeneité dans la population. Plus celui ci est grand, plus le coefficient de variation est faible (proche de zéro), ce qui implique un homogénéité de la population. Réciproquement, plus δ est faible, plus le coefficient de variation est grand (beaucoup de dispersion), ce qui implique une hétérogénéité de la population.

Une autre distribution assez répandue dans les modèles de fragilité est la loi Inverse-Gaussienne ([97] ou [101]).

Soit $\mathcal{S} \sim \mathcal{IG}(\mu, \lambda)$, alors la fonction de probabilité de $\mathcal{S}$ est :

$$f(s) = \left[\frac{\lambda}{2\pi s^3}\right]^{1/2} \times \exp\left(\frac{-\lambda(s-\mu)^2}{2\mu^2 s}\right)$$

Avec $s > 0$, $\mu > 0$ représentant la moyenne et $\lambda > 0$ représentant le paramètre d'échelle. De plus,

$$\mathbb{E}[\mathcal{S}] = \mu \quad \text{et} \quad \mathbb{V}ar[\mathcal{S}] = \frac{\mu^3}{\lambda} \quad \text{et} \quad \mathbb{CV}[\mathcal{S}] = \frac{\sqrt{\mathbb{V}ar[\mathcal{S}]}}{\mathbb{E}[\mathcal{S}]} = \sqrt{\frac{\mu}{\lambda}}$$

Comme on peut le voir, le coefficient de variation ne dépend plus d'un seul paramètre mais de deux. Il apparait donc plus difficile de pouvoir examiner l'hétérogénéité avec cette distribution. Pourtant, cette loi peut sembler plus appropriée. En effet, en l'utilisant, la population devient de plus en plus homogène dans le temps, tandis qu'avec la distribution Gamma, elle reste constante [38].
Ce point est soulevé par BUTT et *al.* [107] et MANTON et *al.* [34] qui estime que celle-ci ne change pas dans le temps, mais diffère avec l'âge, puisque les

individus les plus "faibles" sont censés décéder avant les plus "forts", ce qui a pour conséquence de baisser le niveau d'hétérogénéité .

Dans la suite du mémoire, nous considérerons alors la distribution Gamma comme référence pour modéliser la fragilité pour plusieurs propriétés de cette fonction :

- Sa propriété d'échelle (i.e si $X \sim \Gamma(\delta, \theta)$ alors $aX \sim \Gamma(\delta, \frac{\theta}{a})$)
- Son coefficient de variation ne dépend que d'un paramètre
- Elle a de bonnes propriétés lorsqu'elle est associée à la loi de Poisson
- Elle est flexible étant donné que la forme de sa distribution est très variable selon son paramètre d'intensité.

Les modèles de fragilité les plus communs sont des extensions des modèles de hazard proportionnel, le plus connu étant celui de COX [29], repris par TOMAS et *al.*[54] pour développer un modèle intégrant explicitement la prise en compte de l'hétérogéneité avec données *a priori* issues de groupes hétérogènes. Pour une bonne revue des différents modèles paramétriques multiplicatif existants, mais aussi des différentes possibilités qu'en à la modélisation de la fragilité, le lecteur peut se réferer à WIENKE [6] et OLIVIEIRI [3].

Ici, nous cherchons à inclure ce degrès d'hétérogéneité, présent à travers les populations, dans le modèle de Lee-Carter.

4.5 Le modèle Binomial Négatif spécifique à l'âge

Comme nous l'avons précisé, la propriété d'équidispersion du modèle log-Poisson n'est pas très raisonnable et utiliser une distribution de Poisson surdispersée semble plus cohérent. Il est alors possible de considérer que ce paramètre de dispersion suit lui même une loi de probabilité. On se place alors dans le cas d'un modèle de Poisson Mélange.
En utilisant une loi Gamma pour le paramètre de dispersion dans son mélange, DELWARDE et *al.* [17] modélisent le nombre de décès par une loi Négative Binomiale. Cependant, ce dernier laisse constant le paramètre de dispersion. Comme expliqué plus haut, BUTT et *al.* [107] et MANTON et *al.* [34] postulent que la fragilité des individus augmente avec l'âge. LI et *al.* [81] se

sont basé sur ce modèle en prenant un paramètre de dispersion spécifique à l'âge.

Si on se réfère au diagramme de Lexis, le modèle original de Lee-Carter [85] et le log-Poisson [73] considèrent que les taux de mortalité varient entre les cellules , mais pas à l'intérieur de la cellule elle-même. Comme nous l'avons expliqué, en réalité, des facteurs d'hétérogénéité, observables ou non, divisent chaqu'une de ces cellules en plusieurs cellules.

On note N_x le nombre de groupes compris dans chaque cellule initiale, x étant l'âge des individus. Par conséquent, chaque cellule aura une exposition au risque équivalente à $\frac{E_{xt}}{N_x}$, et le nombre de décès de chaque groupe sera alors noté $D_{xt}(i)$ avec $i = 1, 2, ..., N_x$.

Le nombre de décès total dans une cellule sera alors :

$$D_{xt} = \sum_{i=1}^{N_x} D_{xt}(i)$$

On suppose alors que $D_{xt}(i)$ et $D_{xt}(j)$ pour $i \neq j$ sont indépendants, et que :

$$D_{xt}(i)|z_x(i) \sim \mathcal{P}\left(z_x(i) \times \frac{E_{xt}\mu_{xt}}{N_x}\right)$$
$$z_x(i) \sim \Gamma(\delta_x, \theta_x)$$

avec $\mu_{xt} = \exp(\alpha_x + \beta_x \kappa_t)$.

Ici, $z_x(i)$ représente la fragilité de l'individu. On suppose que $\mathbb{E}[z_x] = 1$. De cette façon, en moyenne, le groupe d'individus i a le même niveau de mortalité que la cellule du diagramme de Lexis à laquelle il appartient. Cela dit, $z_x(i)$ varie d'un groupe à l'autre. Lorsque $z_x(i) > 1$, les individus du groupe i sont plus fragiles en moyenne, et inversement, lorsque $z_x(i) < 1$, les individus du groupe i sont moins fragiles en moyenne. Toute distribution ayant un support positif peut donc être choisi. Ici, nous avons retenu la loi Gamma comme expliqué précédemment mais également grâce à ses bonnes propriétés avec la loi de Poisson.

En ayant supposé que $\mathbb{E}[z_x(i)] = 1$, les paramètres de la loi de $z_x(i)$ sont donc égaux. On a alors $z_x(i) \sim \Gamma(\phi_x^{-1}, \phi_x^{-1})$.

La distribution de D_{xt} est donc :

$$\begin{aligned} D_{xt}(i)|z_x(1), ..., z_x(N_x) = \sum_{i=1}^{N_x} D_{xt}(i)|z_x(i) &\sim \mathcal{P}\left(\sum_{i=1}^{N_x} z_x(i) \times \frac{E_{xt}\mu_{xt}}{N_x}\right) \\ &\sim \mathcal{P}\left(\frac{E_{xt}\mu_{xt}}{N_x} \times \sum_{i=1}^{N_x} z_x(i)\right) \\ &\sim \mathcal{P}\left(E_{xt}\mu_{xt}\bar{z}_x\right) \end{aligned}$$

Avec $\bar{z}_x$ la moyenne des $z_x(i)$:

$$\begin{aligned} \bar{z}_x = \frac{\sum_{i=1}^{N_x} z_x(i)}{N_x} &\sim \Gamma\left(N_x\phi_x^{-1}, N_x\phi_x^{-1}\right) \\ &\sim \Gamma\left(\bar{\phi}_x^{-1}, \bar{\phi}_x^{-1}\right) \end{aligned}$$

Pour conclure, le modèle se caractérise comme :

$$\begin{aligned} D_{xt}|\bar{z}_x &\sim \mathcal{P}(E_{xt}\mu_{xt}\bar{z}_x) \\ \bar{z}_x &\sim \Gamma\left(\bar{\phi}_x^{-1}, \bar{\phi}_x^{-1}\right) \end{aligned}$$

Pour en déduire la distribution de D_{xt}, nous allons passer par les fonctions génératrice des moments, en rapellant tout d'abord que la fonction génératrice des moments d'une variable aléatoire Xest définie par $M_X(h) = \mathbb{E}\left[e^{tX}\right]$, avec $h \in \mathbb{R}$. On a alors :

$$\begin{aligned}
M_{D_{xt}|\bar{z}_x}(h) &= \mathbb{E}\left[e^{h\times D_{xt}|\bar{z}_x}\right] \\
&= \mathbb{E}\left[\mathbb{E}[e^{h\times D_{xt}|\bar{z}_x}|\bar{z}_x]\right] \\
&= \exp\left(\bar{z}_x(e^h-1)\right) \\
&= M_{\bar{z}_x}(e^h-1) \\
&= \left(\frac{\bar{\phi}_x^{-1}}{\bar{\phi}_x^{-1}-(e^h-1)}\right)^{\bar{\phi}_x^{-1}}
\end{aligned}$$

En utilisant le fait que $M_{\bar{z}_x}(h) = \left(\frac{\bar{\phi}_x^{-1}}{\bar{\phi}_x^{-1}-h}\right)^{\bar{\phi}_x^{-1}}$. On reconnaît ici la fonction génératrice d'un loi Négative Binomiale. Alors :

$$\Pr\left[D_{xt}=d_{xt}\right] = \frac{\Gamma(d_{xt}+\bar{\phi}_x^{-1})}{\Gamma(\bar{\phi}_x^{-1})d_{xt}!} \times \left(\frac{1}{1+E_{xt}\mu_{xt}\bar{\phi}_x}\right)^{\bar{\phi}_x^{-1}} \times \left(\frac{E_{xt}\mu_{xt}\bar{\phi}_x}{1+E_{xt}\mu_{xt}\bar{\phi}_x}\right)^{d_{xt}}$$

$$D_{xt} \sim \mathcal{NB}\left(\bar{\phi}_x^{-1}, \frac{1}{1+E_{xt}\mu_{xt}\bar{\phi}_x}\right) \tag{4.1}$$

Et :

$$\begin{aligned}
\mathbb{E}[D_{xt}] &= \mathbb{E}[\mathbb{E}[D_{xt}|\bar{z}_x]] = E_{xt}\mu_{xt} \\
\mathbb{V}ar[D_{xt}] &= \mathbb{E}[\mathbb{V}ar[D_{xt}|\bar{z}_x]] + \mathbb{V}ar[\mathbb{E}[D_{xt}|\bar{z}_x]] = \mathbb{E}[D_{xt}] + \bar{\phi}_x\left[\mathbb{E}(D_{xt})\right]^2
\end{aligned}$$

On s'aperçoit que si $\bar{\phi}_x \to 0$, $\mathbb{E}[D_{xt}] = \mathbb{V}ar[D_{xt}]$, ce qui revient à la propriété d'équidispersion de la distribution de Poisson. La distribution de Poisson est donc un cas limite de loi Négative Binomiale.

Pour toutes démonstrations concernant la loi mélange Poisson-Gamma, le lecteur peut se réferer aux ouvrages classique de théorie du risque ([57] ou [95]).

4.5.1 Estimations des paramètres

La fonction de log-vraisemblance du modèle est alors donné par :

$$
\begin{aligned}
\mathcal{L}(\alpha,\beta,\kappa,\phi_x;d_{xt}) &= \ln\prod_{x,t} f(d_{xt};\alpha,\beta,\kappa,\bar{\phi}) \\
&= \ln\prod_{x,t}\frac{\Gamma(d_{xt}+\bar{\phi}_x^{-1})}{\Gamma(\bar{\phi}_x^{-1})d_{xt}!}\times\left(\frac{1}{1+E_{xt}\mu_{xt}\bar{\phi}_x}\right)^{\bar{\phi}_x^{-1}}\times\left(\frac{E_{xt}\mu_{xt}\bar{\phi}_x}{1+E_{xt}\mu_{xt}\bar{\phi}_x}\right)^{d_{xt}} \\
&= \sum_{x,t}\sum_{i=0}^{d_{xt}-1}\ln(\bar{\phi}_x^{-1}-i)-\ln(d_{xt}!)+d_{xt}\ln(\bar{\phi}_x E_{xt}\mu_{xt}) \\
&- (d_{xt}+\bar{\phi}_x^{-1})\ln(1+\bar{\phi}_x E_{xt}\mu_{xt})
\end{aligned}
$$

En utilisant le fait que $\frac{\Gamma(d_{xt}+\bar{\phi}_x^{-1})}{\Gamma(\bar{\phi}_x^{-1})d_{xt}!}=\prod_{k=1}^{d_{xt}}(d_{xt}+\bar{\phi}_x^{-1}-k)$. Ici encore, pour maximiser la log-vraisemblance du modèle, on utilise un algorithme de GOODMAN [84]. On note k la k-ième itération de l'algorithme et, pour simplifier, nous faisons comme LI et *al.* [81] et utilisons les fonctions suivantes :

$$\hat{D}_{xt}^{(k)} = E_{xt}\exp(\hat{\alpha}_x^{(k)}+\hat{\beta}_x^{(k)}\hat{\kappa}_t^{(k)})$$

$$f_{xt}^{(k)} = \sum_{i=0}^{D_{xt}-1}\left(\frac{i}{1+\hat{\bar{\phi}}_x^{(k)}\times i}-\frac{1}{\hat{\bar{\phi}}_x^{(k)}}\right)$$

$$g_{xt}^{(k)} = \hat{D}_{xt}^{(k)}\left(D_{xt}+\frac{1}{\hat{\bar{\phi}}_x^{(k)}}\right)\left(1+\hat{\bar{\phi}}_x^{(k)}\hat{D}_{xt}^{(k)}\right)^{-1}$$

$$h_{xt}^{(k)} = \sum_{i=0}^{D_{xt}-1}\left(\frac{-i^2}{(1+\hat{\bar{\phi}}_x^{(k)}\times i)^2}-\frac{1}{(\hat{\bar{\phi}}_x^{(k)})^2}\right)$$

$$r_{xt}^{(k)} = \frac{\ln(1+\hat{D}_{xt}^{(k+3)}\hat{\bar{\phi}}_x^{(k+3)})}{\hat{\bar{\phi}}_x^{(k+3)^2}}$$

Il est évidemment nécessaire de fixer une valeur d'initialisation des paramètres. Ici, les paramètres $\alpha_x^{(0)}$, $\beta_x^{(0)}$ et $\kappa_t^{(0)}$ prennent les valeurs finales obtenues du modèle log-Poisson et $\phi_x^{(0)}=10^{-3}$. L'algorithme pour déterminer les

paramètres du modèle est alors :

Etape 1 : Mise à jour des α_x

$$\hat{\alpha}_x^{(k+1)} = \hat{\alpha}_x^{(k)} - \frac{\sum_{t=t_1}^{t_n} \left(D_{xt} - \hat{\bar{\phi}}_x^{(k)} g_{xt}^{(k)} \right)}{\sum_{t=t_1}^{t_n} \left(-\hat{\bar{\phi}}_x^{(k)} g_{xt}^{(k)} \right) / \left(1 + \hat{\bar{\phi}}_x^{(k)} \hat{D}_{xt}^{(k)} \right)}$$

$$\hat{\beta}_x^{(k+1)} = \hat{\beta}_x^{(k)}, \quad \hat{\kappa}_t^{(k+1)} = \hat{\kappa}_t^{(k)}, \quad \hat{\bar{\phi}}_x^{(k+1)} = \hat{\bar{\phi}}_x^{(k)}$$

Etape 2 : Mise à jour des β_x

$$\hat{\beta}_x^{(k+2)} = \hat{\beta}_x^{(k+1)} - \frac{\sum_{t=t_1}^{t_n} \left(D_{xt} \hat{\kappa}_t^{(k+1)} - \hat{\bar{\phi}}_x^{(k+1)} \hat{\kappa}_t^{(k+1)} g_{xt}^{(k+1)} \right)}{\sum_{t=t_1}^{t_n} \left(\hat{\bar{\phi}}_x^{(k+1)} \hat{\kappa}_t^{(k+1)^2} g_{xt}^{(k+1)} \right) / \left(1 + \hat{\bar{\phi}}_x^{(k+1)} \hat{D}_{xt}^{(k+1)} \right)}$$

$$\hat{\alpha}_x^{(k+2)} = \hat{\alpha}_x^{(k+1)}, \quad \hat{\kappa}_t^{(k+2)} = \hat{\kappa}_t^{(k+1)}, \quad \hat{\bar{\phi}}_x^{(k+2)} = \hat{\bar{\phi}}_x^{(k+1)}$$

Etape 3 : Contraintes d'identification

$$\hat{\beta}_x^{(k+2)} = \frac{\hat{\beta}_x^{(k+2)}}{\sum_{x=x_1}^{x_m} \hat{\beta}_x^{(k+2)}}$$

$$\hat{\kappa}_t^{(k+2)} = \hat{\kappa}_t^{(k+2)} \times \sum_{x=x_1}^{x_m} \hat{\beta}_x^{(k+2)}$$

Etape 4 : Mise à jour des κ_t

$$\hat{\kappa}_t^{(k+3)} = \hat{\kappa}_t^{(k+2)} - \frac{\sum_{x=x_1}^{x_m} \left(D_{xt} \hat{\beta}_x^{(k+2)} - \hat{\bar{\phi}}_x^{(k+2)} \hat{\beta}_x^{(k+2)} g_{xt}^{(k+2)} \right)}{\sum_{x=x_1}^{x_m} \left(-\hat{\bar{\phi}}_x^{(k+2)} \hat{\beta}_x^{(k+2)^2} g_{xt}^{(k+2)} \right) / \left(1 + \hat{\bar{\phi}}_x^{(k+2)} \hat{D}_{xt}^{(k+2)} \right)}$$

$$\hat{\alpha}_x^{(k+3)} = \hat{\alpha}_x^{(k+2)}, \quad \hat{\beta}_x^{(k+3)} = \hat{\beta}_x^{(k+2)}, \quad \hat{\bar{\phi}}_x^{(k+3)} = \hat{\bar{\phi}}_x^{(k+2)}$$

Etape 5 : Contraintes d'identification

$$\begin{aligned}\hat{\alpha}_x^{(k+3)} &= \hat{\alpha}_x^{(k+3)} + \bar{\kappa}_t^{(k+3)} \times \hat{\beta}_x^{(k+2)} \\ \hat{\kappa}_t^{(k+3)} &= \hat{\kappa}_t^{(k+3)} - \bar{\kappa}_t^{(k+3)}\end{aligned}$$

Etape 6 : Mise à jour des $\bar{\phi}_x$

$$\hat{\bar{\phi}}_x^{(k+4)} = \hat{\bar{\phi}}_x^{(k+3)} - \frac{\displaystyle\sum_{t=t_1}^{t_n}\left(f_{xt}^{(k+3)} - g_{xt}^{(k+3)} + r_{xt}^{(k+3)} + \frac{D_{xt}}{\hat{\bar{\phi}}_x^{(k+3)}}\right)}{\displaystyle\sum_{t=t_1}^{t_n}\left[h_{xt}^{(k+3)} - \frac{D_{xt} + 2r_{xt}^{(k+3)}\hat{\bar{\phi}}_x^{(k+3)}}{\hat{\bar{\phi}}_x^{(k+3)2}} + \left(\frac{2}{(\hat{\bar{\phi}}_x^{(k+3)})^2} + g_{xt}^{(k+3)}\right)\left(\frac{\hat{D}_{xt}^{(k+3)}}{1+\hat{\bar{\phi}}_x^{(k+3)}\hat{D}_{xt}^{(k+3)}}\right)\right]}$$

$$\hat{\alpha}_x^{(k+4)} = \hat{\alpha}_x^{(k+3)}, \quad \hat{\beta}_x^{(k+4)} = \hat{\beta}_x^{(k+3)}, \quad \hat{\kappa}_t^{(k+4)} = \hat{\kappa}_t^{(k+3)}$$

Cette algorithme, par rapport au précédent nécessite un grand nombre d'itérations pour voir converger le paramètre de dispersion. La procédure est donc très longue. Dans un soucis de temps, nous avons décider d'arrêter l'algorithme lorsque la croissance entre deux log-vraisemblance est inférieur à 10^{-6}.

L'interprétation des paramètres α_x, β_x et κ_t est similaire aux modèles précédents. Comme on peut le voir, le paramètre d'hétérogénéité est très variable entre 0 et 40 ans surtout chez les femmes, où il est plus important. Nous aurions pu nous attendre à l'inverse. Au delà de 40 ans, le paramètre varie très peu et est très proche de zéro.

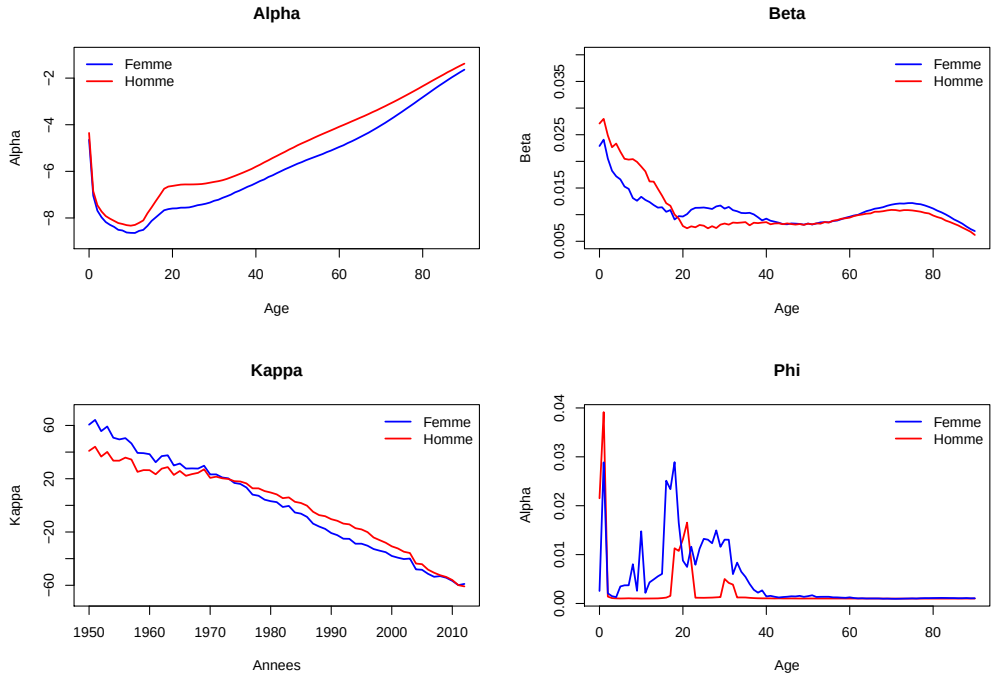

FIGURE 4.5: Paramètres du modèle Binomiale Négatif

4.6 Ajustement aux données

Avant de continuer vers la projection de la mortalité, il est important de savoir si les taux de décès sont correctement modélisés et s'ils sont en adéquation avec les taux bruts ou le nombre de décès, selon le modèle.

Ici le modèle de Lee-Carter n'aura pas les mêmes indicateurs d'ajustement que les autres, ces derniers étant estimés par maximum de vraimsemblance.

4.6.1 Taux d'inertie

Le taux d'inertie ne concerne que le modèle de Lee-Carter. Cette indicateur n'est calculable qu'en passant par la méthode de décomposition en valeurs singulières. Le taux d'inertie se définit par :

$$\tau_1 = \frac{\lambda_1}{\sum_{i \geq 1} \lambda_i}$$

En plus de déterminer la pertinence de l'approximation de la matrice $\mathcal{Z}$, ce

pourcentage permet d'aide à la décision dans le rajout ou non d'un terme bilinéaire temporel. Ici, nous obtenu un taux d'inertie de 43,34% pour les femmes et de 44,65% pour les hommes. Ces résultats ne sont pas très satisfaisants et une meilleure approximation peut être trouvéee en suivant les conseils de BOOTH et al. [82] ou RENSHAW et al. [109].

4.6.2 Variance temporelle

Comme dans tous modèles, on souhaite que les résidus, ici ϵ_{xt}, soit le plus petit possible. La majeure partie de la variance des $\mu_x(t)$ à chaque âge sera expliquée par κ_t. La variance temporelle expliquée par le modèle à un âge x donnée est alors le rapport :

$$\mathcal{V}_{TEMP}(x) = 1 - \frac{\frac{1}{t_n - t_1 + 1}\sum_{t=t_1}^{t_n}(\hat{\zeta}_x(t))^2 - \left[\frac{1}{t_n - t_1 + 1}\sum_{t=t_1}^{t_n}(\hat{\zeta}_x(t))\right]^2}{\frac{1}{t_n - t_1 + 1}\sum_{t=t_1}^{t_n}(\hat{\mu}_x(t))^2 - \left[\frac{1}{t_n - t_1 + 1}\sum_{t=t_1}^{t_n}\hat{\mu}_x(t)\right]^2}$$

avec $\hat{\zeta}_x(t) = \mu_x(t) - \exp(\hat{\alpha}_x + \hat{\beta}_x\hat{\kappa}_t)$ et $\mu_x(t)$ les taux bruts. Cette variance peut se calculer sur les trois modèles (Figures 4.6 et 4.7).
Ici, les parts de variance expliquées par le modèle de Lee-Carter sont très variables et fluctuent beaucoup, surtout pour les femmes. Pour ce sexe, elle ne dépassent pas 90%, mais reste inférieur à 80% sur la plupart des âges avec notamment de fortes parts non expliquées entre 16 et 20 ans puis entre 45 et 65 ans où elle descend en dessous des 50%. Chez les hommes, cette variance fluctue moins mais connaît également un pic vers 20 ans.
Les deux autres modèles, log-Poisson et Négatif Binomiale, sont assez proche, surtout pour les femmes. Chez les hommes, le modèle Négatif Binomiale est le moins performant lors des premières années de vie. Dans les deux modèles, on note d'ailleurs une forte part de variance non-expliquée entre 16 et 23 ans. Nous aurions pu penser qu'étant donné la valeur du paramètre d'hétérogéneité sur ces âges, le modèle Négatif Binomiale s'ajusterai mieux.

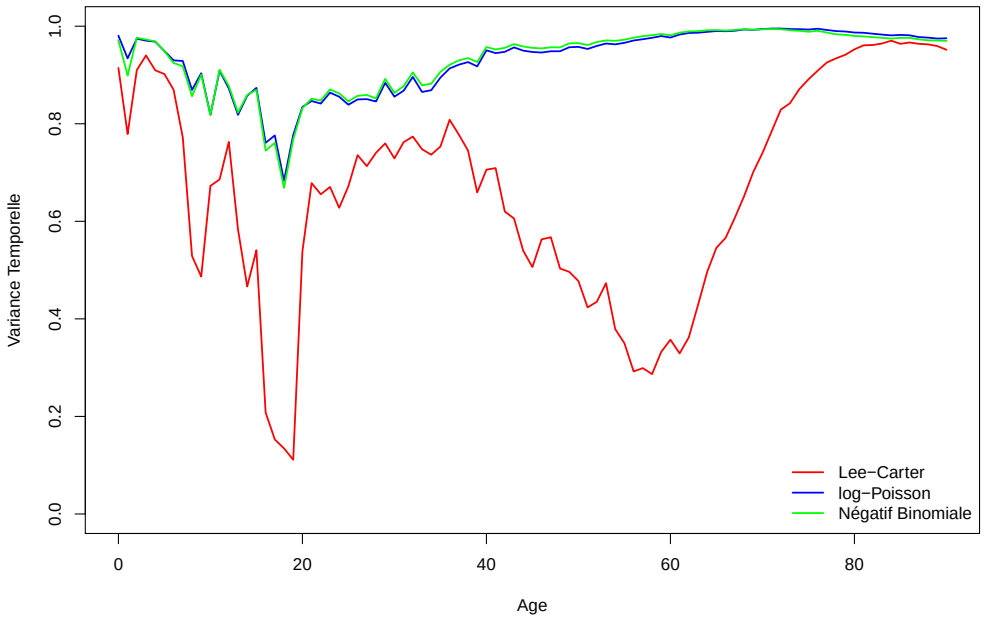

FIGURE 4.6: Variance Temporelle des différents modèles pour les femmes

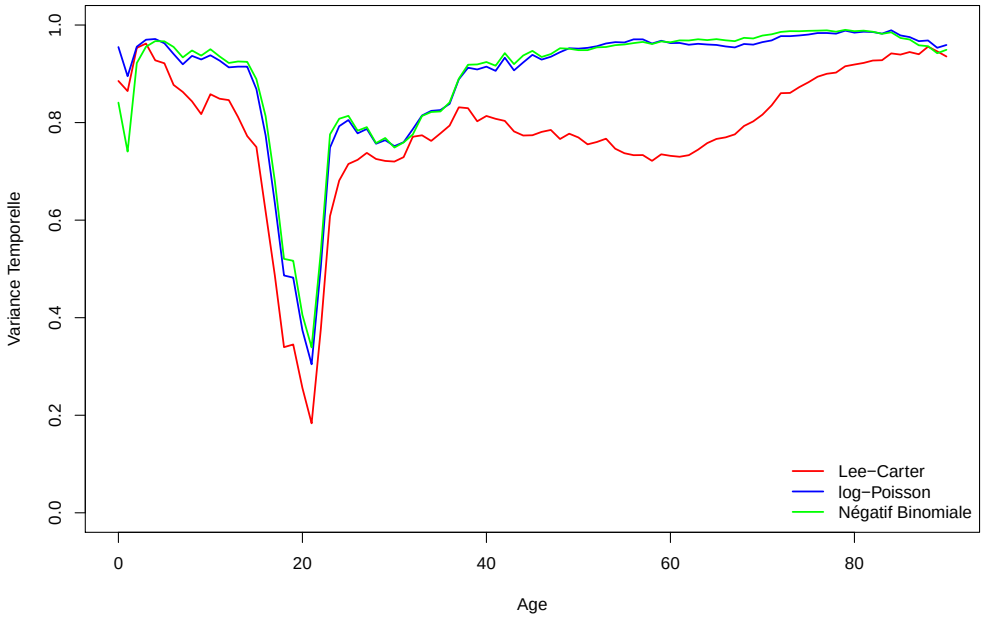

FIGURE 4.7: Variance Temporelle des différents modèles pour les hommes

4.6.3 AIC et SBC

Les modèles log-Poisson et Négatif Binomiale sont estimés par maximum de vraisemblance et certaines statistiques utilisent justement la vraisemblance pour mesurer l'ajustement aux données :

- AIC [58] (*Akaike Information Criterion*) : $k - \mathcal{L}$. Il représente un compromis entre le biais, qui diminue avec le nombre de paramètre libres, et la volonté de décrire les données en utilisant le moins de paramètres.
- SBC (ou BIC) [56] (*Schwarz Bayesian Criterion* ou *Bayesian Information Criterion*) : $-\mathcal{L} + 0,5k\ln(n)$. Il est plus pénalisant que le AIC puisqu'il prend en compte le nombre de variables présent dans le modèle.

Avec $\mathcal{L}$ la log-vraisemblance du modèle, k le nombre de paramètres du modèle et n le nombre d'observations. Ici, il n'est pas nécessaire de calculer l'AIC corrigé, qui s'applique lorsque le nombre de données est 40 fois plus grand que le nombre de paramètres.

Dans le modèle log-Poisson, nous avons $91(\alpha_x)+91(\beta_x)+63(\kappa_t)-2(\text{contraintes}) = 243$ paramètres. Le modèle Négatif Binomiale contient 91 paramètres de plus(ϕ_x), soit 334 paramètres. Dans tous les cas, nous avons $91 \times 63 = 5733$ observations.

Les résultats obtenus se trouvent dans le tableau 4.2.

Modèle	**Femme**			**Homme**		
	$\mathcal{L}$	AIC	SBC	$\mathcal{L}$	AIC	SBC
$\mathcal{P}$	-41576,14	41819,14	42627,6	-56722,744	56965,74	57774,2
$\mathcal{NB}$	-16906,31	17240,31	18351,53	-31513,08	31847,08	32958,3

TABLE 4.2: Indicateurs de la qualité des modèles log-Poisson et Binomial Négatif

4.6.4 Analyse des résidus

Les résidus fournissent également une analyse pertinente. Il ne doivent représenter aucune tendance, aucune structure, autrement cela signifierait que le modèle ignore certains mécanismes lié à la mortalité. On note y la variable modélisée. Plusieurs type de résidus existent :

- Les résidus en ligne, représentés par la simple différence entre les valeurs observées et estimées : $r_{xt}^{ligne} = y_{xt} - \hat{y}_{xt}$

- Les résidus de Pearson : $r_{xt}^{Pearson} = \frac{y_{xt}-\hat{y}_{xt}}{\sqrt{\mathbb{V}\text{ar}(\hat{y}_{xt})}}$
- Les résidus de Déviance : $r_{xt}^{Dviance} = \sqrt{a_i}\text{signe}(y_{xt} - \hat{y}_{xt})$ avec a_i la contribution de l'observation i à la déviance $\mathcal{D}$.

Pour le modèle de Lee-Carter, nous utilisons les résidus en ligne :

$$r_{xt}^{ligne} = \mu_{xt} - \hat{\mu}_{xt}$$

Etant donné que nous avons effectué des régressions de Poisson et Négative Binomiale pour les autres modèles, nous allons regarder les résidus de Pearson, exprimés par :

$$r_{xt}^{Pearson\mathcal{P}} = \frac{D_{xt} - \hat{D}_{xt}}{\sqrt{\hat{D}_{xt}}}$$

$$r_{xt}^{Pearson\mathcal{NB}} = \frac{D_{xt} - \hat{D}_{xt}}{\sqrt{\hat{D}_{xt} + \bar{\phi}_x \hat{D}_{xt}}}$$

Nous aurions également pu analyser les résidus de Déviance, mais ceux-ci donne une tendance assez similaire ceux de Pearson. Nous donnons tout de même les formules de ces valeurs :

$$r_{xt}^{Deviance\mathcal{P}} = \sqrt{2}\sum_{t=t_1}^{t_n}\sum_{x=x_1}^{x_m}\sqrt{D_{xt}\ln\frac{D_{xt}}{\hat{D}_{xt}} - (D_{xt} - \hat{D}_{xt})}$$

$$r_{xt}^{Deviance\mathcal{NB}} = \sqrt{2}\sum_{t=t_1}^{t_n}\sum_{x=x_1}^{x_m}\sqrt{D_{xt}\ln\frac{D_{xt}}{\hat{D}_{xt}} - (D_{xt} + \hat{\bar{\phi}}_x)\ln\left(\frac{D_{xt} + \hat{\bar{\phi}}_x}{\hat{D}_{xt} + \hat{\bar{\phi}}_x}\right)}$$

Les résidus étant trop élevés aux âges infantiles pour le modèle de Lee-Carter (0 à 4 ans) et log-Poisson (0 à 2 ans), nous avons décidé de ne pas les représenter sur les Figures 4.8, 4.9, 4.10 et 4.11 par soucis de représentation.

La première remarque concernant le modèle de Lee-Carter vient du fait qu'il sur-estime les taux de mortalité chez les hommes (Figure 4.9) mais qu'il les sous-estime chez les femmes (Figure 4.8). L'ajustement semble très bien entre 4 et 15 ans pour les deux sexes avant de considérablement se dégrader. La plus grosse tendance concerne la "bosse de l'accident", soit les âges entre 16 et 20 ans. A noter également l'effet de cohorte, qui est visible via des diagonales sur les résidus.

Pour le modèle log-Poisson (Figures 4.10 et 4.11) , on s'aperçoit bien ici que les résidus de Pearson décèlent plusieurs tendances. On peut noter l'effet de cohorte qui est bien présent, avec une sur-estimation du nombre de décès suivit directement d'une sous estimation. Chez les femmes, le modèle varie également beaucoup entre sur et sous-estimation tandis que chez les hommes, mis à part au début des années 1950 pour les âges les plus élevées, il sur-estime le nombre de décès.

Pour le modèle Négatif Binomiale (Figures 4.12 et 4.13), l'effet de cohorte est toujours présent. Une certaine partie des résidus de Pearson chez les hommes (entre 22 et 40 ans aux alentours du milieu des années 1990) est surestimée, comme c'était déjà le cas avec le modèle log-Poisson. On retrouve néanmoins une variance moins importante de ces résidus comparé à ceux du modèle précédent.

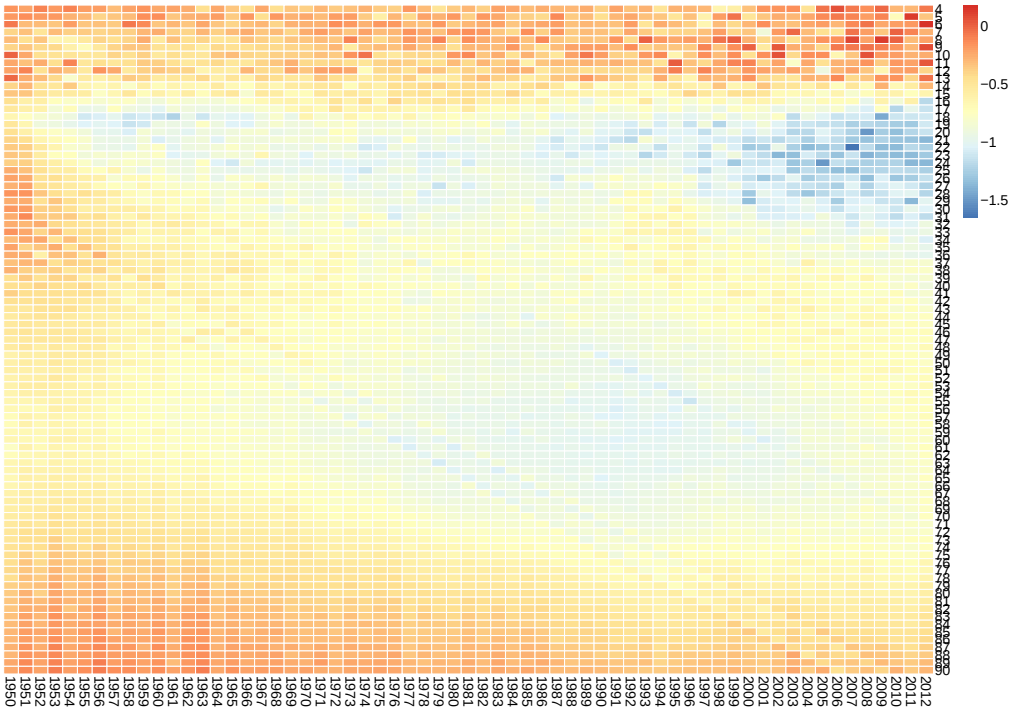

FIGURE 4.8: Résidus du modèle de Lee-Carter pour les femmes

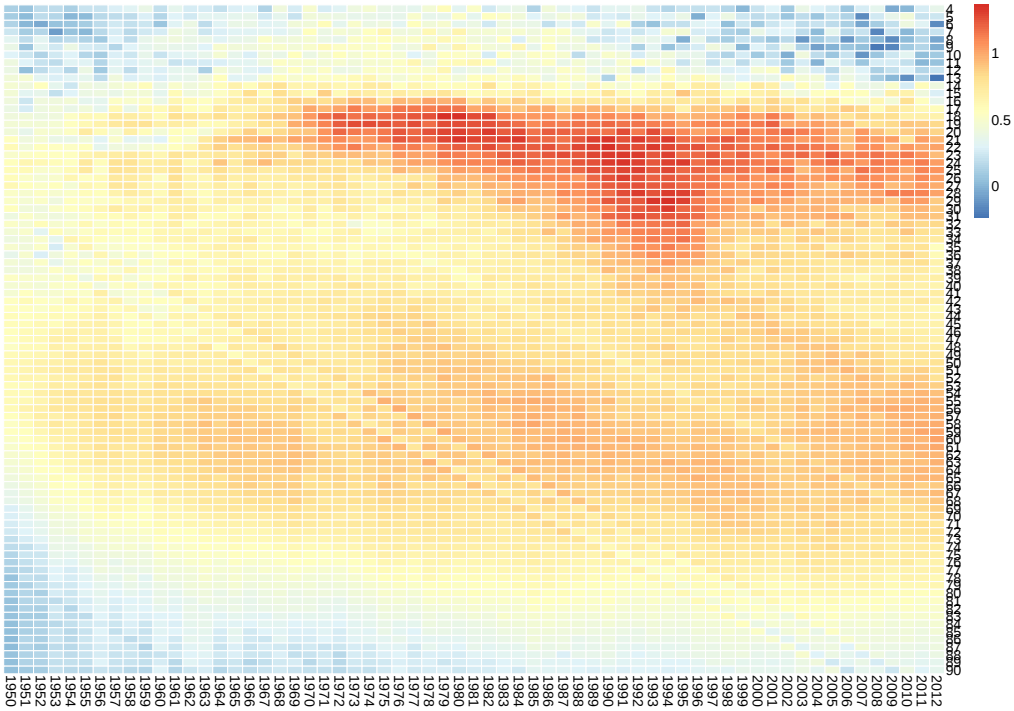

FIGURE 4.9: Résidus du modèle de Lee-Carter pour les hommes

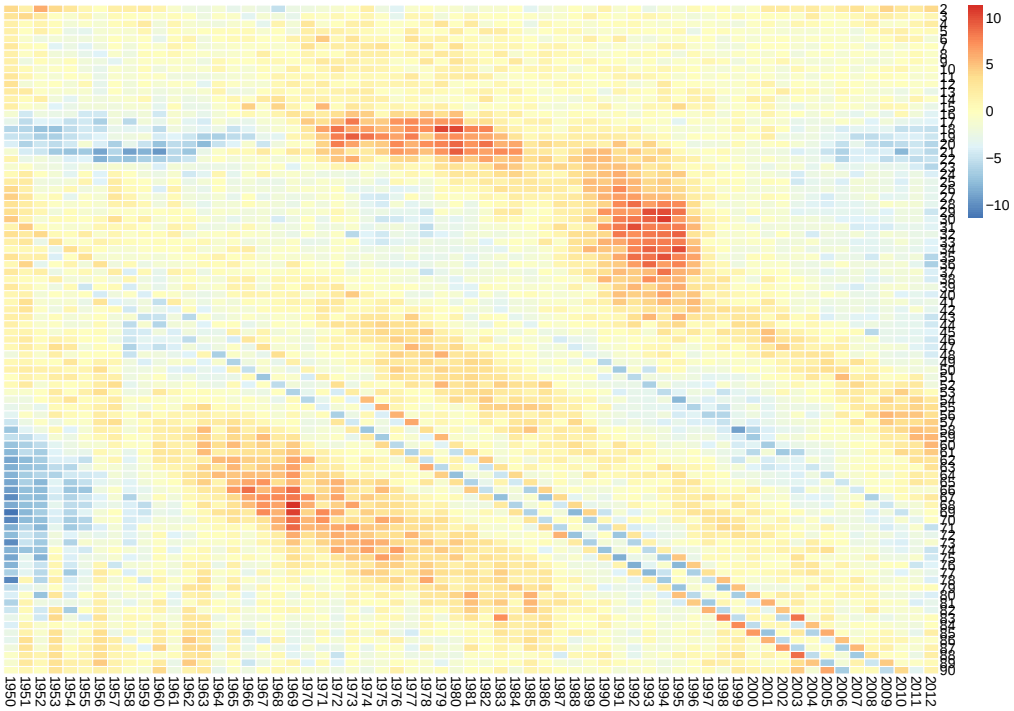

FIGURE 4.10: Résidus de Pearson du modèle log-Poisson des femmes

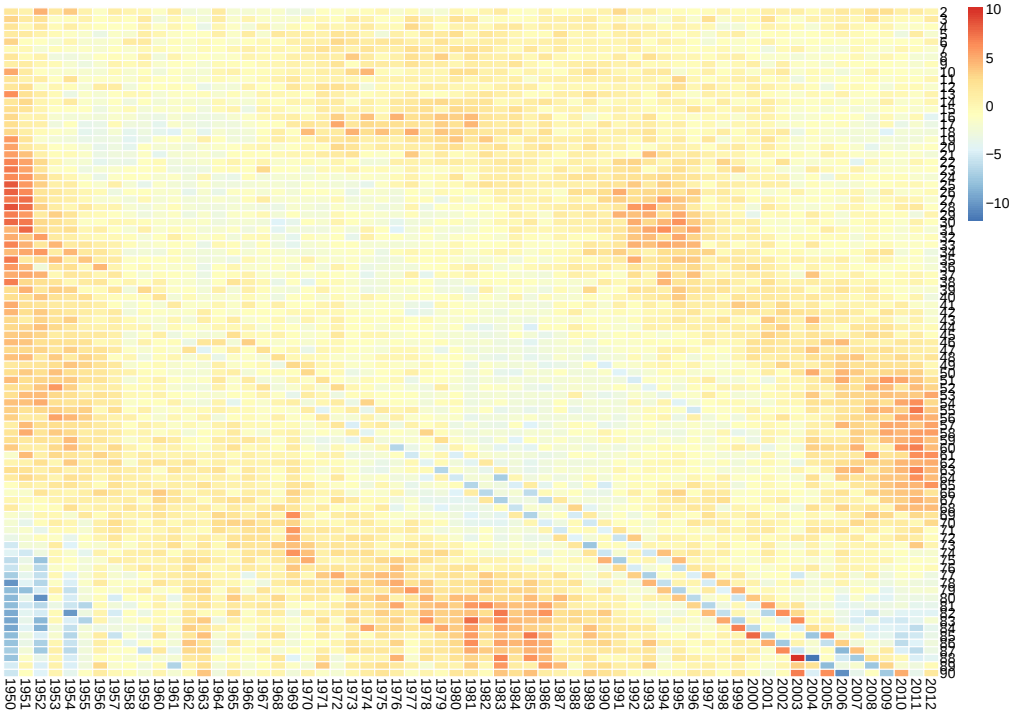

FIGURE 4.11: Résidus de Pearson du modèle log-Poisson des hommes

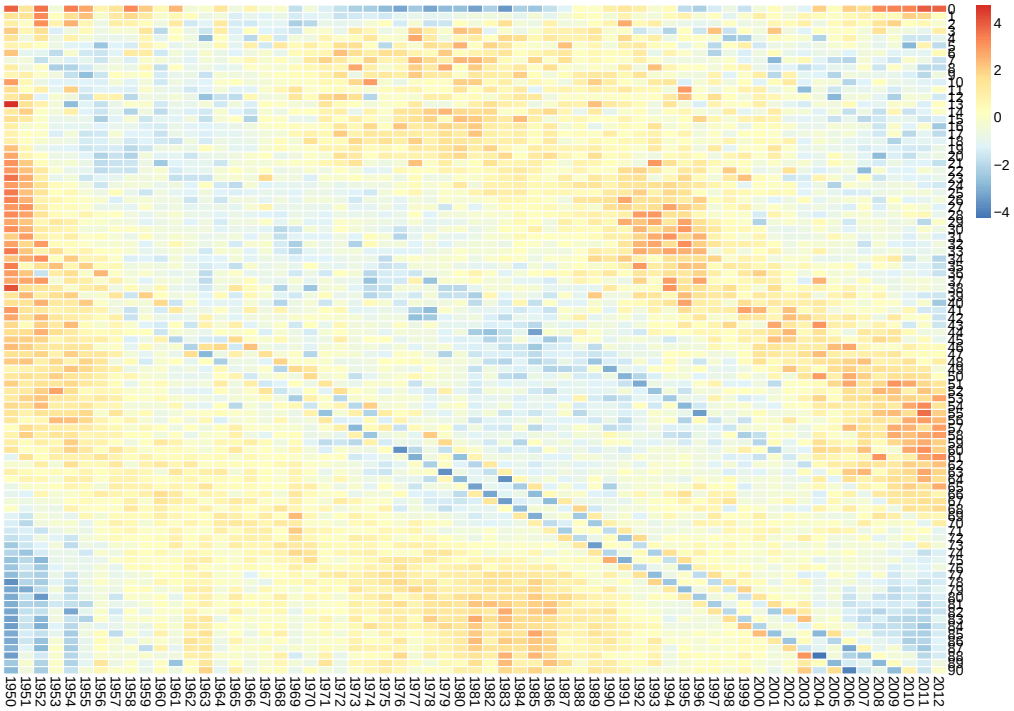

FIGURE 4.12: Résidus de Pearson du modèle Negatif Binomiale des femmes

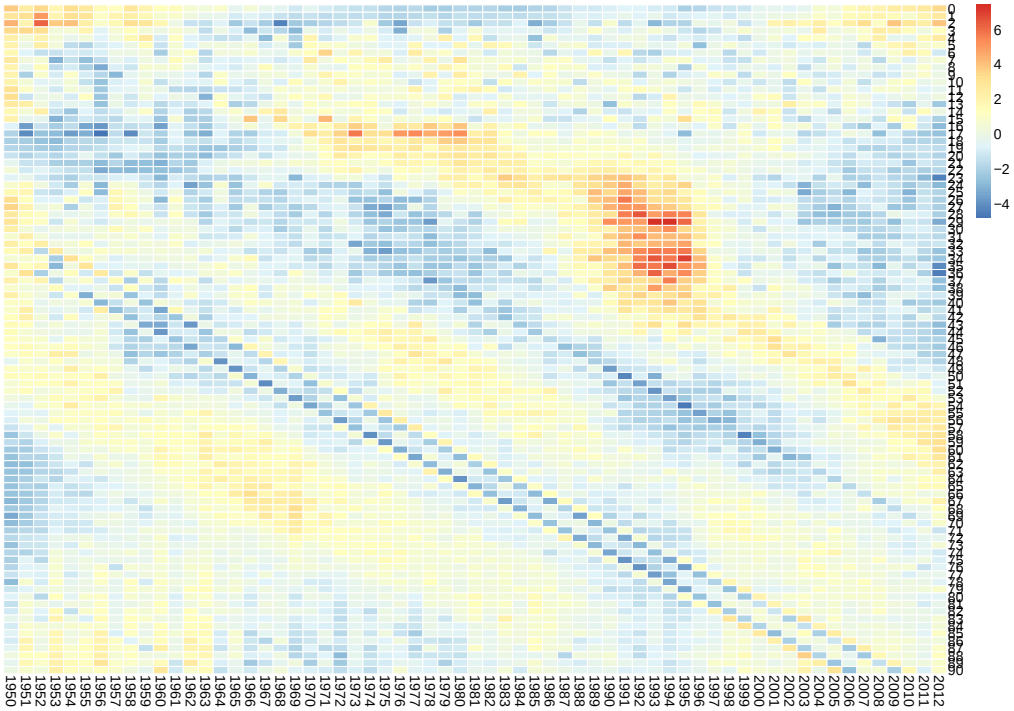

FIGURE 4.13: Résidus de Pearson du modèle Negatif Binomiale des hommes

4.7 Modélisation aux âges extrêmes

Comme nous l'avons évoqué auparavant, les données dont nous disposons sont très variable aux âges extrêmes et risquent d'être de mauvaise qualité. Les modèles que nous avons utilisés sont trop sensibles aux fluctuations des taux de décès résidant aux âges extrèmes. Des extrapolations des paramètres α_x et β_x peuvent être mises au point, par des méthodes de lissage ou de régression. Mais l'allure des paramètres est trop aléatoire et cette méthode ne donne pas des résultats très satisfaisants. Les actuaires et démographes ont établis certaines techniques permettant de meilleurs ajustements ; on parle de méthodes de "fermeture de table de mortalité". Pour une revue des différentes techniques existantes, le lecteur peut se réferer à QUASHIE et *al.* [40].

4.7.1 La méthode de Coale-Kisker

Cette méthode suppose que :

$$\hat{\mu}_x = \hat{\mu}_{65} \exp\left(k_x(x-65)\right)$$

Avec $x \leq 65$ ans et k_x le taux de croissance de la mortalité à l'âge x. En examinant les taux de croissance de la mortalité pour plusieurs pays, les auteurs ont remarqué un pic des k_x aux alentours de 80 ans avant de décroître linéairement. Ils supposent alors :

$$k_x = k_{80} + s \times (x-80), \quad \text{avec } x \leq 65 \text{ ans}$$

Le taux de mortalité pour un âge ultime est ensuite fixé arbitrairement à 110 ans :

$$\hat{\mu}_{110}(t) = \begin{cases} 0.8 & \text{pour les femmes} \\ 1 & \text{pour les hommes} \end{cases}$$

On en déduit que :

$$\begin{aligned}
\hat{\mu}_{110}(t) &= \hat{\mu}_{79}(t) \exp\left(\sum_{x=80}^{110} k_x\right) \\
&= \hat{\mu}_{79}(t) \exp\left(\sum_{x=80}^{110} k_{80} + s \times (x-80)\right) \\
&= \hat{\mu}_{79}(t) \exp\left(31 \times k_{80} + s \times \sum_{x=80}^{110} (x-80)\right) \\
&= \hat{\mu}_{79}(t) \exp\left(31 \times k_{80} + s \times \frac{30 \times 31}{2}\right)
\end{aligned}$$

Alors, $s = -\frac{\ln\left(\frac{\hat{\mu}_{79}(t)}{\hat{\mu}_{110}(t)}\right)+31k_{80}}{465}$. Le calcul des taux de mortalité s'en déduit aisément :

$$\hat{\mu}_x(t) = \hat{\mu}_{79}(t) \exp\left(\sum_{y=80}^{x} k_{80} + s \times (y-80)\right) \qquad \text{avec } x \in \{80, 81, \ldots, 109\}$$

Les Figures 4.14 et 4.15 représentent l'extrapolation via la méthode Coale-Kisker sur plusieurs dates, 1980, 1990, 2000 et 2010. On se rend bien compte à la vue des graphiques, que la méthode semble de moins en moins performante au fil des ans, les taux de décès bruts étants supérieurs aux taux de décès calculés par la méthode. En effet, l'ajustement semble plus performant pour 1980 et 1990 que pour 2000 et 2010. Cette tendance est également plus voyante chez les femmes que chez les hommes. Nous choisissons de ne pas modifier la méthode originale mais plusieurs possibilités existent pour l'améliorer :

- Le pic de croissance des taux de mortalité (k_x) est fixé ici à 80 ans comme dans la méthode originale. Aujourd'hui, celui-ci doit être un petit peu plus élevée.
- Comme le montre QUASHIE et *al.* [40], la contrainte sur l'âge limite joue également un rôle et le choix de $\hat{\mu}_{110}$ peut avoir un impact significatif sur les courbes brutes.

Les mauvaises performances obtenues nous amène à essayer la méthode de DENUIT et GODERNIAUX [42].

4.7.2 Méthode de Denuit-Goderniaux

Cette méthode s'applique directement aux probabilités de décès. Elle consiste à faire une régression sur ces probabilités :

$$\ln \hat{q_x} = a + bx + cx^2 + \epsilon_x$$

Avec x l'age de l'individu et ϵ_x des bruits blancs de variance σ^2. De manière à assurer une concativité des probabilités de décès aux âges extrêmes, on suppose un âge limite x_{max} avec $q_{x_{max}} = 1$ et $q'_{x_{max}} = 0$, ce qui implique :

$$\ln \hat{q}_x = c(x_{max}^2 - 2x \times x_{max} + x^2) + \epsilon_x$$

Ici, nous supposerons comme dans la méthode originale que l'âge limite est de 130 ans, d'où :

$$\ln q_x = c(130^2 - 260x + x^2)$$

Pour déterminer à partir de quel âge nous utiliserons cette modélisation à la place des probabilités de décès brutes, nous cherchons à maximiser le R^2 ajusté sur la période d'âge de 70 à 90 ans. Une attention doit être portée à la discontinuité qui réside lors de "l'âge raccord". Pour lisser cette extrapolation, nous remplaçons la valeur de "l'âge raccord" $\hat{q}_x$ par une moyenne géométrique de 5 ans sur les probabilités de décès.
Ici, étant donné que les modèles de mortalité que nous utilisons modélisent directement les taux de décès, il suffira de repasser de la probabilité de décès à la force de mortalité :

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

Les résultats sont bien meilleurs que ceux du modèle de Coale-Kisker, nous retiendrons alors la méthode de Denuit-Goderniaux. Comme le souligne QUASHIE et *al.* [40], celle-ci est une des meilleures pour mettre en évidence l'exposition aux risques qui réside aux âges extrêmes.

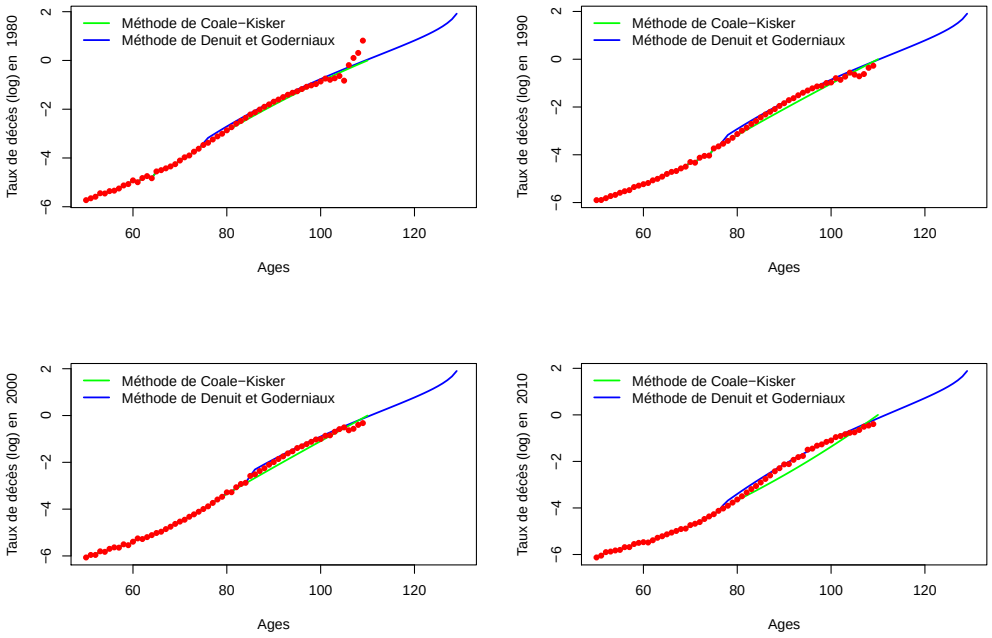

FIGURE 4.14: Quelques taux de décès (log) aux âges extrêmes chez les femmes

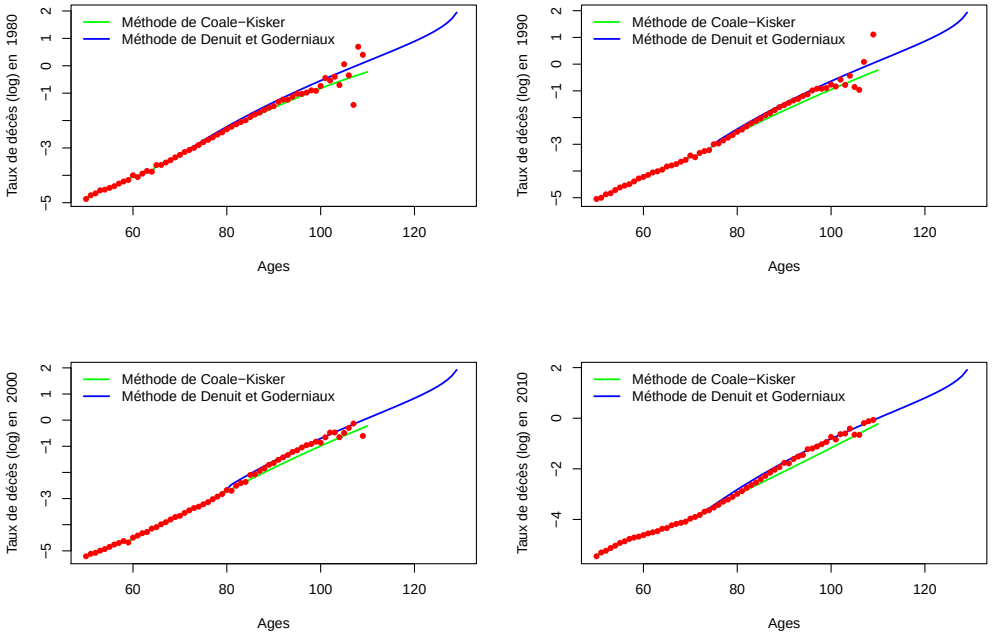

FIGURE 4.15: Quelques taux de décès (log) aux âges extrêmes chez les hommes

4.8 Conclusion du chapitre

Le modèle de Lee-Carter est un modèle simple dont les composantes sont facilement interprétables et estimables. Plusieurs techniques existent pour estimer les paramètres, nous avons utilisé ici la décomposition en valeurs singulières, qui suppose l'utilisation de données dans une matrice rectangulaire. Un réajustement du paramètre temporel est ensuite nécessaire pour modéliser au mieux le nombre de décès intervenant chaque année. Le principal inconvénient de ce modèle est l'hypothèse d'homoscédasticité des résidus des taux de décès. En effet, cette hypothèse est incohérente puisque la variance des taux de décès augmente avec l'âge, étant donné que la population décroît.

Le modèle log-Poisson permet de supposer une hétéroscédascité de ces résidus. Nous avons vu que l'estimation des paramètres se fait via l'algorithme de Goodman, basé sur celui de Newton-Raphson, mais est spécifique aux modèles log bilinéaire. Ce modèle a plusieurs avantages par rapport à celui de Lee-Carter : il permet de déterminer directement le nombre de décès intervenant chaque année pour chaque âge, de soulager le modèle de l'hypothèse d'homoscédascité des résidus, et d'avoir une estimation des paramètres par maximum de vraisemblance, apportant toutes les bonnes propriétés de ces estimateurs. La distribution de Poisson suppose une équidispersion, qui est rarement vérifiée. Ici, la variance des décès est souvent supérieure à son espérance, à cause du degré d'hétérogéneité de la population.

La prise en compte d'un paramètre de dispersion permet de représenter ce degrès d'hétérogénéité. Nous avons supposé ici qu'il était spécifique à chaque âge et distribué selon une loi Gamma. Cette hypothèse nous amène à supposer que le nombre de décès est distribué selon une distribution Négative Binomiale. L'estimation des paramètres du modèle s'effectue encore par un algorithme de Goodman. Ce modèle est certes plus compliqué d'un point de vue estimation des paramètres, mais il permet d'obtenir une meilleure qualité d'ajustement aux données comme nous l'avons vu sur le AIC ou le BIC pour les deux sexes. Le degré d'hétérogéneité est très variable selon l'âge, ce qui remet en cause l'hypothèse de DELWARDE et *al.* [17] de le laisser général.

Néanmoins, ces modèles étant tous dérivés de celui de Lee-Carter, ils ignorent les mécanismes sous-jacent à la mortalité, comme les avancées médicales ou changements environnementaux, ce qui engendre une mauvaise modélisation

lors des guerres (effets de cohorte), ou des évènements "catastrophes" (canicule de 2003 sur les personnes agées), comme le montrent les résidus.

La modélisation aux âges extrêmes est très volatile et les modèles de mortalité classiques reflètent mal cette tranche d'âges. La méthode de DENUIT et GODERNIAUX est la plus adaptée. Elle se base sur une régression polynomiale des probabilités de décès et, par une transformation, permet d'approcher quasi-naturellement les taux de décès bruts.

Chapitre 5

Projeter la mortalité

La projection des modèles se fait par l'extrapolation du paramètre κ_t. En effet, les taux de décès à la date $t+h$ seront calculés via $\mu_{x,t+h} = \exp(\alpha_x + \beta_x \kappa_{t+h})$. Cependant, ces derniers sont projetés à partir du dernier point du modèle, et non de la dernière donnée observée, ce qui crée un saut. Un réajustement des paramètres est alors nécessaire.

Dans cette partie, nous verrons tout d'abord comment régler le problème de saut dans la projection pour les deux modèles sur lesquels nous travaillons, puis nous analyserons la série κ_t pour déterminer un processus cohérent à ceux-là. Comme précisé dans le chapitre 3, plusieurs façons existent pour projeter κ_t (ARIMA(p,d,q), loi de sauts exponentielle double, processus de Lévy, processus NIG,...). Ici, étant donné que nous travaillons avec des données n'ayant pas vraiment de variabilité (nous ne notons pas de sauts dans la série), nous traiterons uniquement avec un processus ARIMA(p,d,q). Nous effectuerons plusieurs étapes pour obtenir un processus cohérent qui répondra aux différentes hypothèses émises.

5.1 Correction du saut dans la projection

Comme expliqué, les taux de mortalité établis par les modèles ne sont pas identiques à la dernière observation. Une condition s'impose alors :

$$\ln \mu_x(t_n) = \ln \hat{\mu}_x(t_n)$$

Si cette condition n'est pas respectée, on notera alors une incohérence dans la projection. Dans l'article original, Lee et Carter [85], l'accepte en expliquant que ce saut n'a pas un impact conséquent lors du calcul des espérances de vie. Cette explication peut sembler cohérente lorsque le modèle s'ajuste très bien aux données, ce qui n'est pas le cas pour tous les âges. BELL [120] et LEE et *al.* [114] propose une méthode pour corriger ce saut, ce qui améliore la projection. Deux contraintes s'imposent :

$$\begin{aligned}\hat{\alpha}_x &= \ln \mu_x(t_n)\\ \kappa_{t_n} &= 0\end{aligned}$$

Elles signifient que le paramètre α_x correspond au dernier taux de mortalité de l'âge x, et le recalibrage de κ_t provoquera alors l'ajustement du modèle sur les derniers taux de mortalité observé.

Certaines étapes des algorithmes précédent doivent alors être modifiées.

5.1.1 Modèle log-Poisson

Dans le modèle log-Poisson, l'étape 1 change et devient :

$$\begin{aligned}\hat{\alpha}_x^{(k+1)} &= \ln \mu_x(t_n)\\ \hat{\beta}_x^{(k+1)} &= \hat{\beta}_x^{(k)}, \quad \hat{\kappa}_t^{(k+1)} = \hat{\kappa}_t^{(k)}\end{aligned}$$

Idem pour les contraintes de l'étape 3 :

$$\begin{aligned}\hat{\alpha}_x^{(k+2)} &= \hat{\alpha}_x^{(k+2)} + \hat{\kappa}_{t_n}^{(k+2)} \times \hat{\beta}_x^{(k+2)}\\ \hat{\kappa}_t^{(k+2)} &= \hat{\kappa}_t^{(k+2)} - \hat{\kappa}_{t_n}^{(k+2)}\end{aligned}$$

Les nouveaux paramètres sont représentés dans la Figure 5.1.

5.1.2 Modèle Négatif Binomial

Dans le modèle Négatif Binomial, l'étape 1 change et devient :

$$\hat{\alpha}_x^{(k+1)} = \ln \mu_x(t_n)$$
$$\hat{\beta}_x^{(k+1)} = \hat{\beta}_x^{(k)}, \quad \hat{\kappa}_t^{(k+1)} = \hat{\kappa}_t^{(k)}, \quad \hat{\bar{\phi}}_x^{(k+1)} = \hat{\bar{\phi}}_x^{(k)}$$

Idem pour les contraintes de l'étape 5 :

$$\hat{\alpha}_x^{(k+2)} = \hat{\alpha}_x^{(k+2)} + \hat{\kappa}_{t_n}^{(k+2)} \times \hat{\beta}_x^{(k+2)}$$
$$\hat{\kappa}_t^{(k+2)} = \hat{\kappa}_t^{(k+2)} - \hat{\kappa}_{t_n}^{(k+2)}$$

Les nouveaux paramètres sont représentés sur les Figures 5.1 et 5.2. On peut constater que le paramètre β_x, quelque soit le modèle et le sexe est très volatil, ce qui est causé par le réajustement des κ_t. L'interprétation effectuée auparavant n'est plus possible. On peut cependant noter que le paramètre $\bar{\phi}_x$ est cette fois plus elevé pour les hommes.

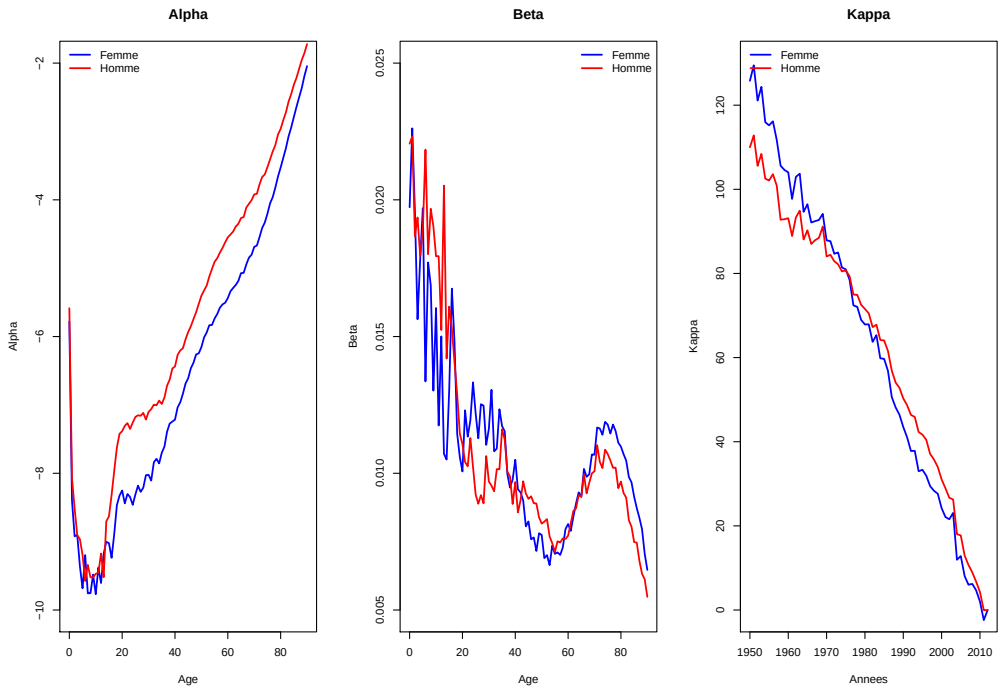

FIGURE 5.1: Paramètres du modèle log-Poisson pour la projection

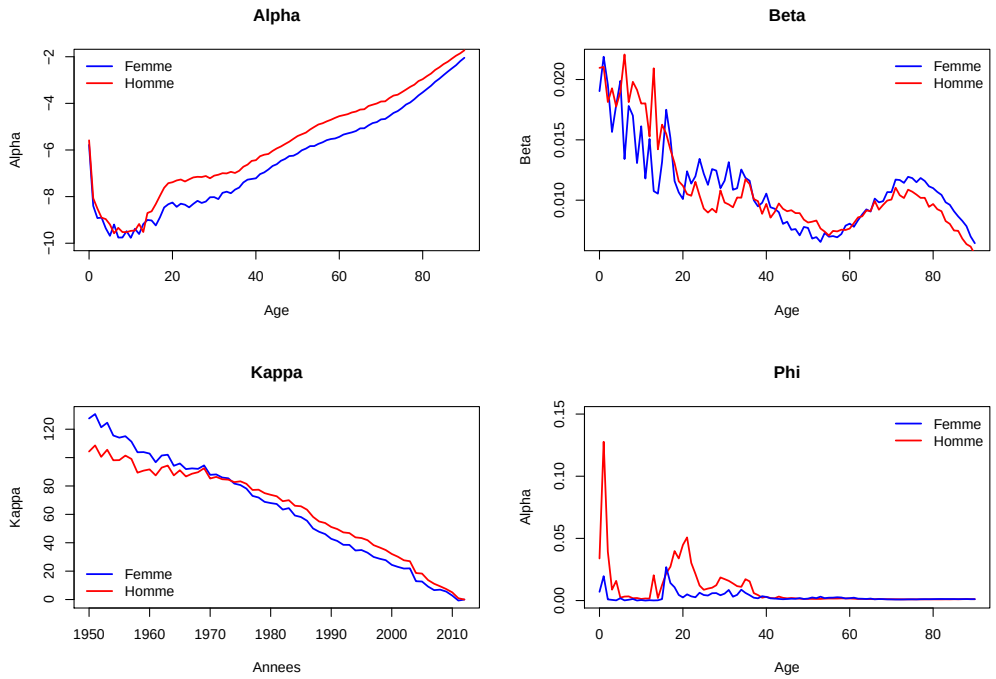

FIGURE 5.2: Paramètres du modèle Négatif Binomiale pour la projection

5.2 Modélisation de κ_t

Le paramètre κ_t se définit comme une série temporelle puisqu'il s'agit d'un jeu de données spécifiques au temps. Box et Jenkins [43] ont proposé une démarche de prévision pour ces séries, basée sur les processus ARIMA(p,d,q). L'analyse doit se faire par étapes :

- Etape 1 : Examination de la série de manière à déceler une éventuelle tendance, saisonnalité ou valeur aberrante voir des changements brutaux
- Etape 2 : Enlever la tendance ou la saisonnalité de manière à rendre les résidus de la série stationnaire
- Etape 3 : Choisir un modèle pour ajuster les données
- Etape 4 : Analyser les résidus du modèle pour vérifier que celui-ci est adéquat
- Etape 5 : Projection de série

Pour plus de précision sur les séries temporelles, le lecteur peut se réferer à BROCKWELL et *al.* [39].

5.2.1 Analyse de la série

Les Figures 5.1 et 5.2 contiennent les paramètres κ_t des deux modèles. On se rend compte que ceux-là décroissent linéairement dans le temps, il y a alors une tendance. Aucun signe de saisonnalité n'est à signaler.

Une série X_t est dite stationnaire si elle répond a deux propriétés statistique :

- $\mathbb{E}[X_t]$ est indépendant de t
- $Cov(X_{t+h}, X_t)$ est indépendant de t pour tout h

La fonction d'autocorrélation et la fonction partielle d'autocorrélation peuvent être analysées pour vérifier la corrélation entre les observations. L'ACF a une décroissance linéaire et toutes les observations jusqu'à l'ordre 15 sont corrélées au seuil de 95%, il est alors certain que la série n'est pas stationnaire. Pour enlever cette stationnarité, nous allons passer par une différenciation de la série κ_t.

5.2.2 Analyse de la série différenciée

Pour estimer cette composante de tendance, plusieurs méthodes existent, comme les moyennes mobile, le lissage exponentiel ou l'ajustement polynomial. Ici, plutôt que de chercher à estimer cette tendance, nous chercherons a l'éliminer en différenciant la série.

On note ∇^d l'opérateur qui définit la différenciation au temps d :

$$\nabla^d \kappa_t = \kappa_t - \kappa_{t-1} - \ldots - \kappa_{t-d}$$

Pour rappel, un processus ARIMA(p,d,q) se note :

$$\nabla^d \kappa_t = \sum_{k=1}^{p} \phi_k \kappa_{t-k} + \sum_{k=1}^{q} \theta_k \epsilon_{t-k} + \epsilon_t$$

Avec $\epsilon_t \sim \mathcal{BB}(0, \sigma^2)$, les résidus du modèle

Ici, en différenciant à l'ordre 1, on obtient les séries des Figures 5.3 et 5.4, qui n'offrent pas de structure particulière, leurs comportements étant eratique.

Pour déterminer les processus adéquats à la modélisation de κ_t différenciée, nous calculons les fonctions d'autocorrélation et les fonctions partielles d'autocorrélation, représentées sur les Figures 5.5 et 5.6. La Table 5.1 contient le numéro des observations sortant de l'intervalle de confiance à 95% de ces deux fonctions.

	Femme		Homme	
Fonction	$\mathcal{P}$	$\mathcal{NB}$	$\mathcal{P}$	$\mathcal{NB}$
ACF	1 et 8	1 et 8	1, 6, 8 et 12	1,6 et 12
PACF	1, 6 et 8	1, 6 et 8	6	5 et 6

TABLE 5.1: Numéros de l'observations sortant de l'IC à 95% de l'ACF et de la PACF

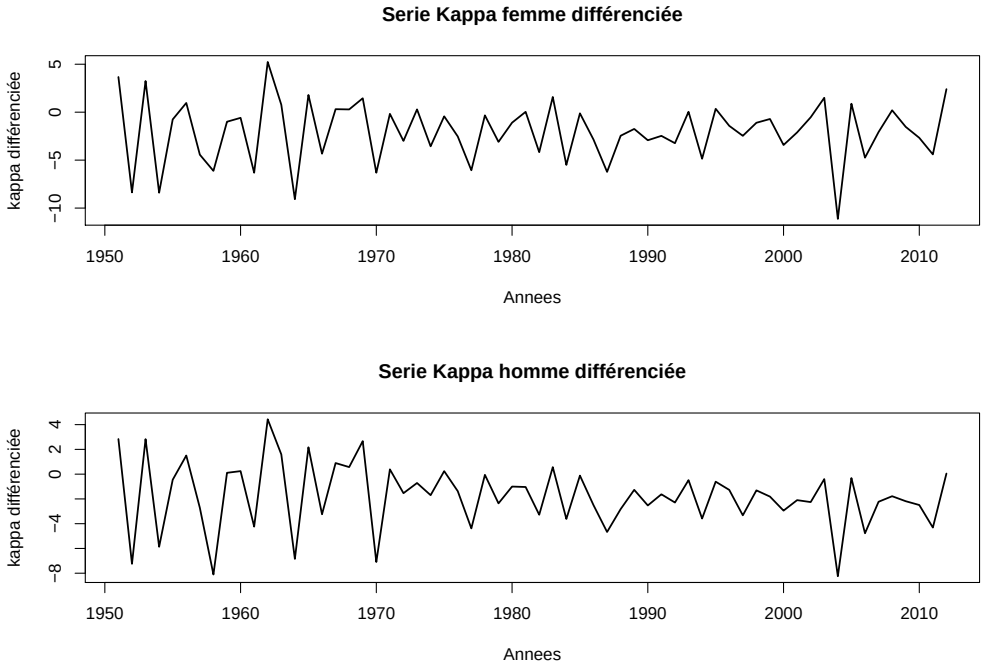

FIGURE 5.3: Série κ_t différenciée à l'ordre 1 du modèle log-Poisson

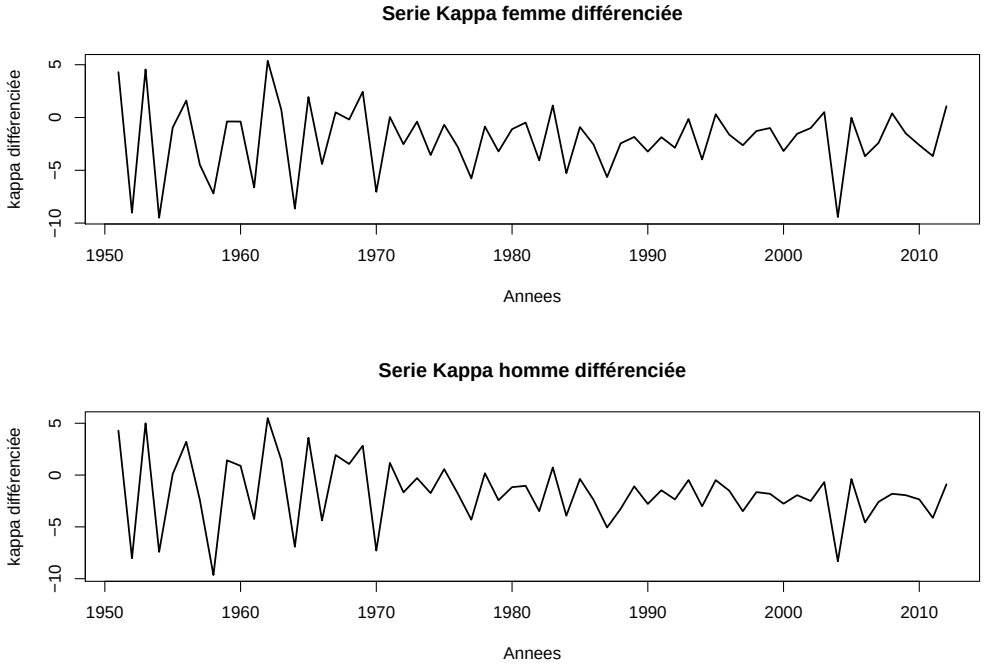

FIGURE 5.4: Série κ_t différenciée à l'ordre 1 du modèle Négatif Binomiale

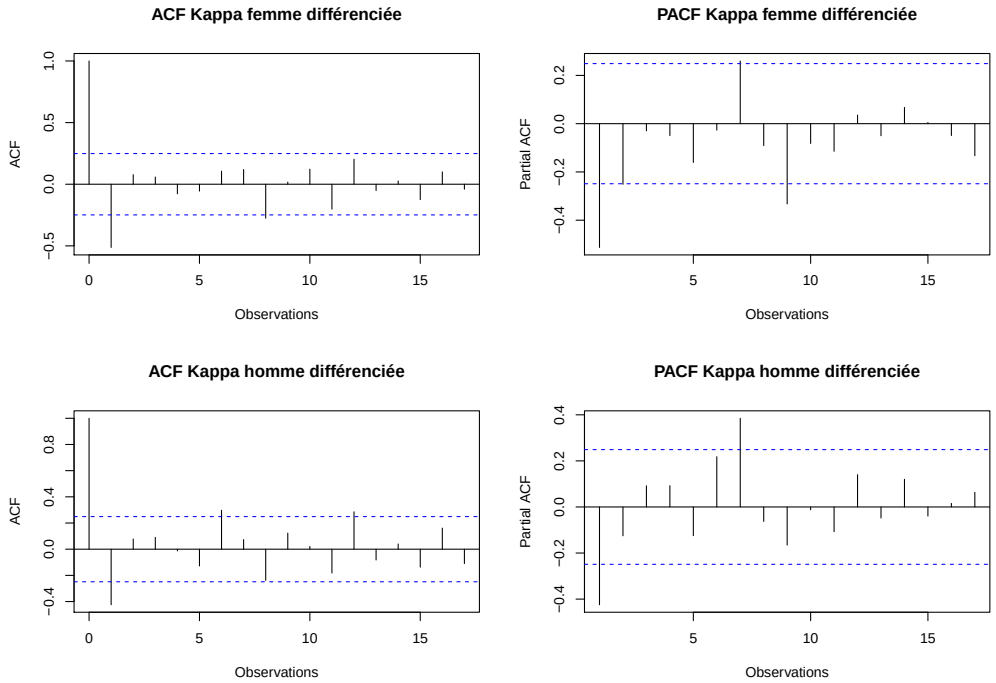

FIGURE 5.5: ACF et PACF de la série κ_t différenciée à l'ordre 1 du modèle log-Poisson

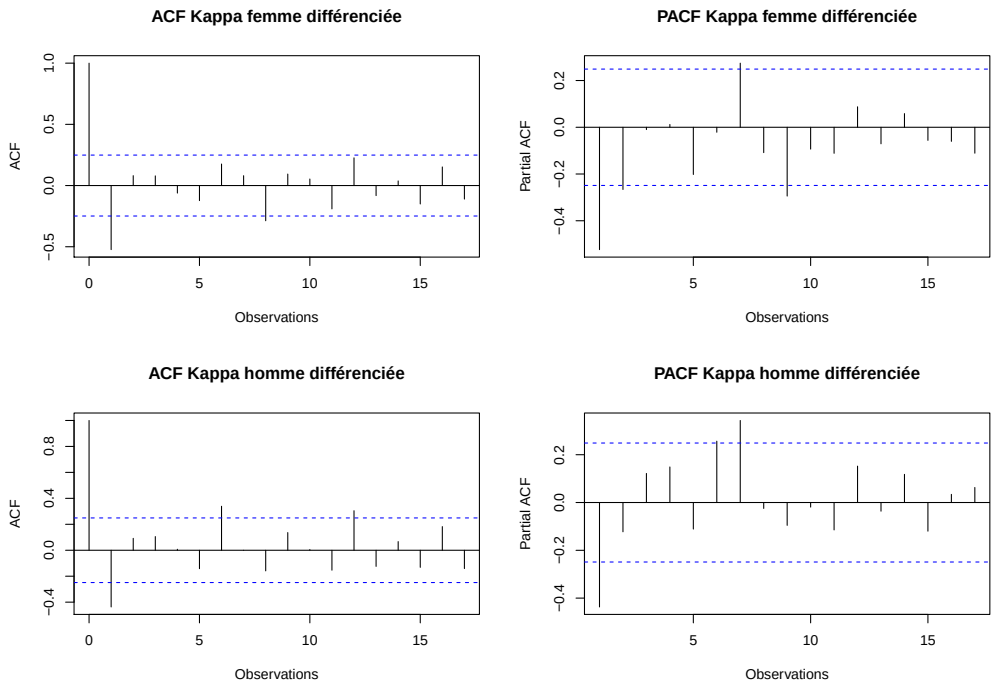

FIGURE 5.6: ACF et PACF de la série κ_t différenciée à l'ordre 1 du modèle Négatif Binomiale

5.2.3 Estimation des paramètres du modèle

En supposant les résidus normalement distribués et indépendants, $(\epsilon_{t_1}, \ldots, \epsilon_{t_n})$ est un vecteur gaussien. $\boldsymbol{\kappa}$ est ici différencié à l'ordre 1, les composantes du vecteur $\boldsymbol{\kappa}$ forment alors un vecteur gaussien. La vraisemblance du modèle à maximiser est donc :

$$\mathcal{L}(\boldsymbol{\kappa}, \phi, \theta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \times \frac{1}{\sqrt{\det\Gamma}} \times \exp\left(-\frac{1}{2\sigma^2}\boldsymbol{\kappa}'\Gamma^{-1}\boldsymbol{\kappa}\right)$$

Avec Γ la matrice des covariances de $\boldsymbol{\kappa}$. Sans rentrer dans les détails de l'optimisation :

$$\hat{\sigma}^2 = \frac{\sum_{t=t_1}^{t_n}(\kappa_t - \hat{\kappa}_t)^2}{t_n - t_1}$$

Et, les estimateurs au maximum de vraisemblance de ϕ et θ sont tels qu'ils minimisent la fonction :

$$l(\phi, \theta) = \ln\left(\frac{\sum_{t=t_1}^{t_n}(\kappa_t - \hat{\kappa}_t)^2}{t_n - t_1}\right) + \frac{\sum_{j=1}^{t_n - t_1}\ln(\epsilon_{j-1})}{t_n - t_1}$$

Les estimateurs au maximum de vraisemblance convergent vers des distributions normales :

$$\begin{pmatrix} \hat{\phi}_1 \\ \vdots \\ \hat{\phi}_p \\ \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_q \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \phi_1 \\ \vdots \\ \phi_p \\ \theta_1 \\ \vdots \\ \theta_q \end{pmatrix}, \frac{\sigma^2}{t_n - t_1}\Gamma_n\right)$$

Cette estimation permet de calculer des intervalles de confiance pour les paramètres. Dans la suite, nous avons utilisé la fonction *arima* du package

forecast de R pour déterminer la valeur des paramètres ainsi que leur erreur associée.
Les tables en annexe donnent les valeurs de l'AIC et du BIC pour les processus testés. Comme d'habitude, nous cherchons a minimiser ces indicateurs.

5.2.4 Analyse des résidus

Pour valider un modèle, les résidus des modèles doivent répondre à plusieurs hypothèses : ils doivent être centrés, normalement distribués et indépendants. Plusieurs tests sont applicables :

- Le T-Test (1) pour tester si les résidus sont centrés, c'est à dire leur moyenne est nulle (hypothèse H_0 : les résidus ne sont pas centrés contre l'hypothèse H_1 : les résidus sont centrés)
- Le test de Box-Pierce (2) pour tester la non corrélation des résidus (hypothèse H_0 : les résidus sont corrélés contre hypothèse H_1 : les résidus ne sont pas corrélés)
- Les tests de normalité pour tester l'allure normale des résidus (hypothèse H_0 : les résidus ne sont pas distribués selon une loi normale contre l'hypothèse H_1 : les résidus sont distribués selon une loi normale)
  - Anderson-Darling (3)
  - Lilliefors (4), qui est une adaptation du test de Kolmogorov-Smirnov
  - Jarques-Bera (5), qui se base sur le kurtosis et le coefficient d'asymétrie
  - Shapiro-Wilk (6)

Les test utilisés proviennent des packages *nortest* et *tseries* de R. Les p-values des tests sont données en annexe.

Le modèle log-Poisson

Pour les femmes, le processus minimisant l'AIC est un ARIMA(1,1,8), celui minimisant le BIC est un ARIMA(1,1,1). Au vu des tests sur les résidus, de sa bonne qualité et de sa simplicité, nous décidons de choisir le processus ARIMA(1,1,1). La Figure 5.7 montre quelques graphiques sur les résidus du modèle :

- Les résidus semblent centrés, leur moyenne est 0,032
- La fonction d'autocorrélation montre que les résidus sont bien non corrélés au seuil de 5%, une seule observation sort de l'intervalle de confiance (la 8me).
- Les résidus semblent distribués normalement, commme le montre l'histogramme, le QQ-Plot et la droite d'Henry.

Le modèle pour les femmes s'écrit alors :

$$\begin{aligned}
\kappa_t - \kappa_{t-1} &= \xi + \phi_1 \kappa_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t \\
\kappa_t - \kappa_{t-1} &= -2,0919 - 0,2462\kappa_{t-1} - 0,4496\epsilon_{t-1} + \epsilon_t
\end{aligned}$$

Avec ξ une constante et $\epsilon_t \sim \mathcal{BB}(0; 6,751)$, indépendant et identiquement distribué.

Pour les hommes, le processus minimisant l'AIC est un ARIMA (6,1,6), celui minimisant le BIC est un ARIMA (6,1,1). Malgré ces critères de sélection non-optimal, nous décidons d'opter pour le processus ARIMA (6,1,8) au vu des résultats des tests sur les résidus. La Figure 5.8 montre quelques graphiques sur les résidus du modèle :

- Les résidus semblent centrés, leur moyenne est -0,05
- La fonction d'autocorrélation montre que les résidus sont bien non corrélés au seuil de 5%.
- Les résidus semblent distribués normalement, commme le montre l'histogramme, le QQ-Plot et la droite d'Henry.

Le modèle pour les hommes s'écrit alors :

$$\begin{aligned}
\kappa_t - \kappa_{t-1} &= \xi + \sum_{k=1}^{6} \phi_k \kappa_{t-k} + \sum_{k=1}^{8} \theta_k \epsilon_{t-k} + \epsilon_t \\
\kappa_t - \kappa_{t-1} &= -1,784 - 0,156\kappa_{t-1} - 0,0003\kappa_{t-2} - 0,086\kappa_{t-3} - 0,0302\kappa_{t-4} \\
&\quad + 0,054\kappa_{t-5} + 0,808\kappa_{t-6} - 0,272\epsilon_{t-1} + 0,31\epsilon_{t-2} + 0,215\epsilon_{t-3} \\
&\quad + 0,09\epsilon_{t-4} - 0,3305\epsilon_{t-5} - 0,4141\epsilon_{t-6} + 0,5876\epsilon_{t-7} - 0,4774\epsilon_{t-8} + \epsilon_t
\end{aligned}$$

Avec ξ une constante et $\epsilon_t \sim \mathcal{BB}(0; 2,468)$, indépendant et identiquement distribué.

Le modèle Binomiale Négatif

Pour les femmes, le processus minimisant l'AIC et le BIC est un ARIMA (1,1,1), mais celui-ci ne répond pas assez bien aux hypothèses sur les résidus pour être adéquat. Nous préférons le processus ARIMA (1,1,8). La Figure 5.9 montre quelques graphiques sur les résidus du modèle :

- Les résidus ne sont pas correctement centrés, leur moyenne est 0,18
- La fonction d'autocorrélation montre que les résidus sont bien non corrélés au seuil de 5%.
- Les résidus ne semblent pas normalement distribués à cause des valeurs extrêmes.

Le modèle pour les femmes s'écrit alors :

$$\kappa_t - \kappa_{t-1} = \xi + \theta_1 \kappa_{t-1} \sum_{k=1}^{8} \phi_k \epsilon_{t-k} + \epsilon_t$$
$$\begin{aligned}\kappa_t - \kappa_{t-1} = & -2,0712 - 0,453\kappa_{t-1} - 0,3161\epsilon_{t-1} - 0,0244\epsilon_{t-2} - 0,0746\epsilon_{t-3} - 0,0776\epsilon_{t-4} \\ & - 0,132\epsilon_{t-5} + 0,1522\epsilon_{t-6} - 0,1056\epsilon_{t-7} - 0,422\epsilon_{t-8} + \epsilon_t\end{aligned}$$

Avec ξ une constante et $\epsilon_t \sim \mathcal{BB}(0; 5)$, indépendant et identiquement distribué.

Pour les hommes, le processus minimisant l'AIC est un ARIMA(5,1,6), celui minimisant le BIC est un ARIMA(5,1,1). Au vu des tests effectués sur les résidus, nous choisissons l'ARIMA(5,1,6). La Figure 5.10 montre quelques graphiques sur les résidus du modèle :

- Les résidus semblent centrés, leur moyenne est -0,07
- La fonction d'autocorrélation montre que les résidus sont bien non corrélés au seuil de 5%.
- Les résidus semblent suivre une distribution normale mis à part quelques valeurs extrêmes.

Le modèle pour les hommes s'écrit alors :

$$\kappa_t - \kappa_{t-1} = \xi + \sum_{k=1}^{5} \phi_k \kappa_{t-k} + \sum_{k=1}^{6} \theta_k \epsilon_{t-k} + \epsilon_t$$

$$\begin{aligned}\kappa_t - \kappa_{t-1} = & -1,6613 + 0,0673\kappa_{t-1} - 0,1368\kappa_{t-2} + 0,1210\kappa_{t-3} - 0,2336\kappa_{t-4} + 0,1635\kappa_{t-5} \\ & - 0,7285\epsilon_{t-1} + 0,5876\epsilon_{t-2} - 0,2306\epsilon_{t-3} + 0,3768\epsilon_{t-4} - 0,4626\epsilon_{t-5} \\ & + 0,875\epsilon_{t-6} + \epsilon_t\end{aligned}$$

Avec ξ une constante et $\epsilon_t \sim \mathcal{BB}(0; 4, 71)$, indépendant et identiquement distribué.

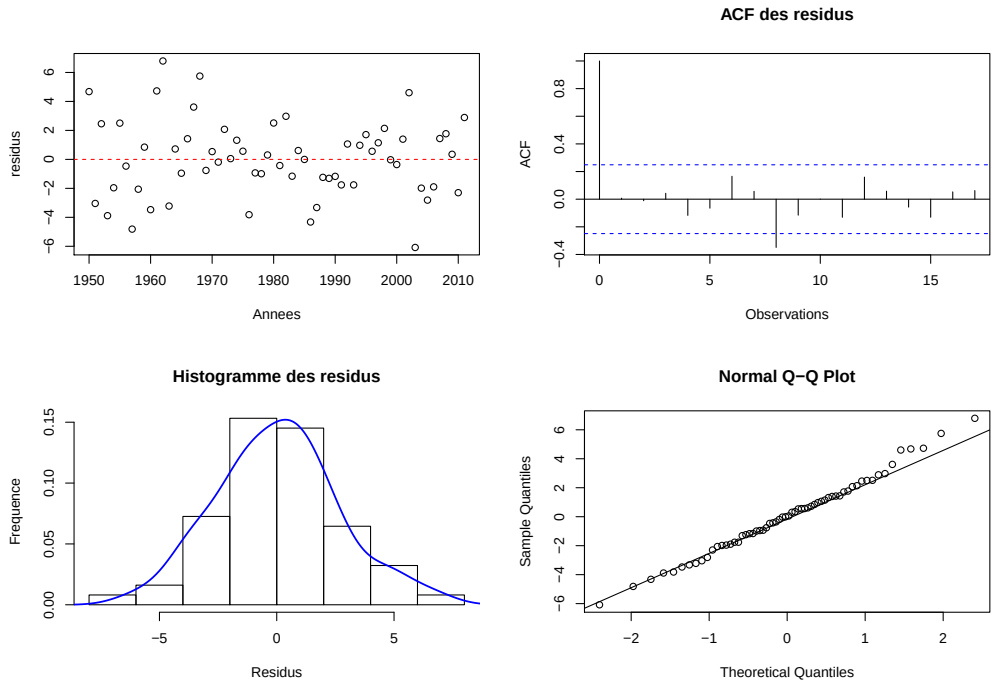

FIGURE 5.7: Modèle log-Poisson femme : Analyse des résidus du ARIMA(1,1,1)

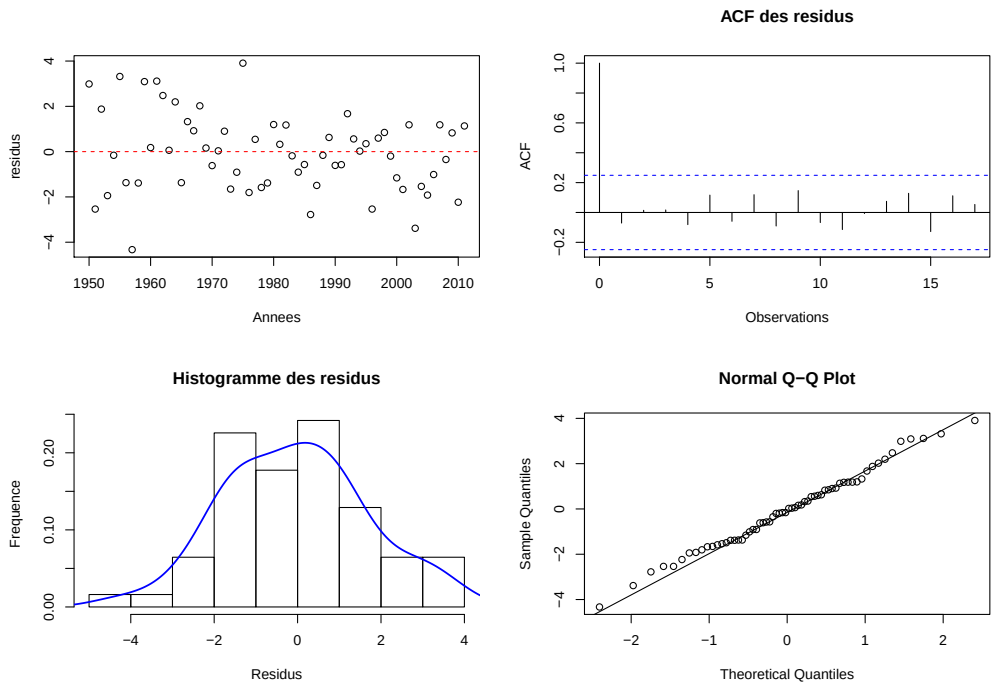

FIGURE 5.8: Modèle log-Poisson homme : Analyse des résidus du ARIMA(6,1,8)

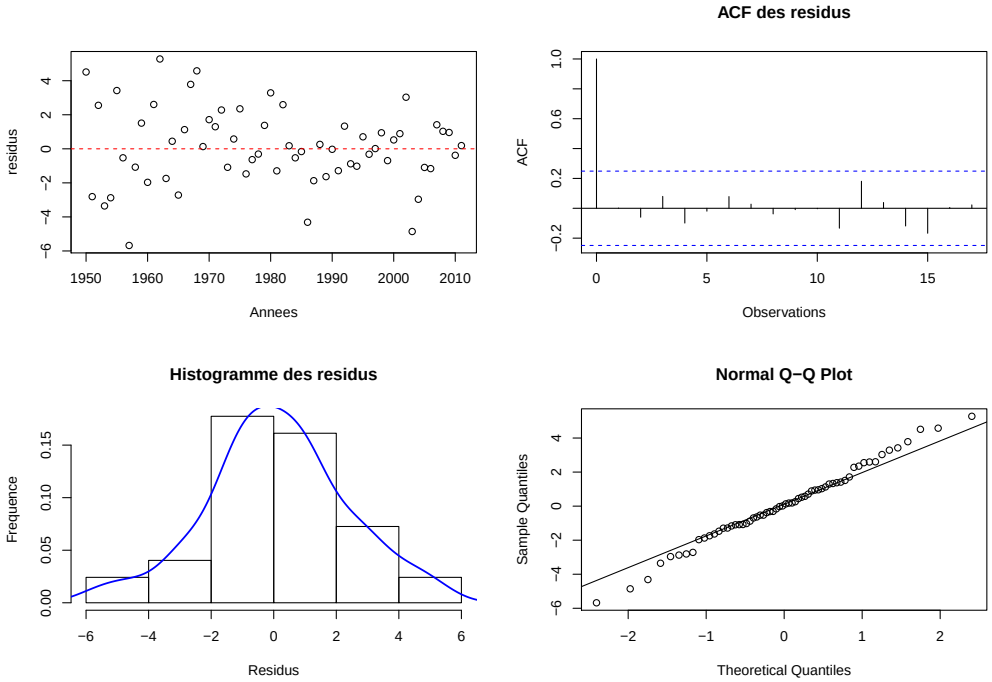

FIGURE 5.9: Modèle Négatif Binomiale femme : Analyse des résidus du ARIMA(1,1,8)

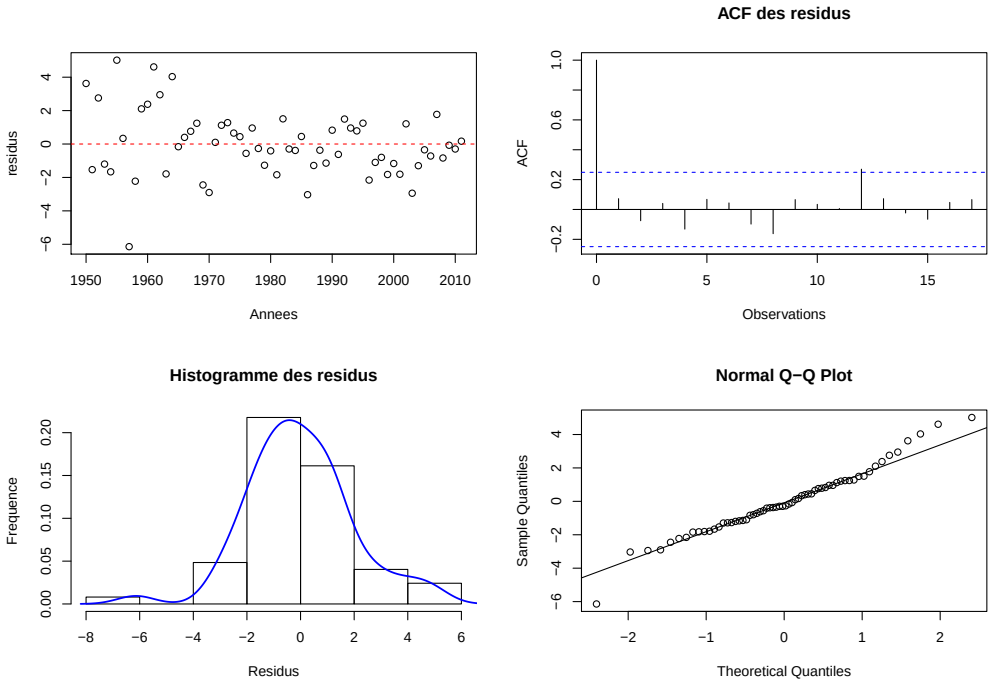

FIGURE 5.10: Modèle Négatif Binomiale homme : Analyse des résidus du ARIMA(5,1,6)

5.3 Projection de κ_t

Nous entrons à présent dans la dernière étape du processus, la projection des κ_t. Dans tous les cas, les processus que nous utilisons sont des ARIMA(p,1,q) dont la prédiction à l'ordre $t+h$ sera :

$$\kappa_{t+h} - \kappa_{t+h-1} = \xi + \sum_{k=1}^{p} \phi_k \kappa_{t+h-k} + \sum_{k=1}^{q} \theta_k \epsilon_{t+h-k} + \epsilon_{t+h}$$

La meilleure prédiction à la date $t+h$ est notée :

$$\begin{aligned}\hat{\kappa}_{t+h} &= \mathbb{E}\left[\kappa_{t+h} | \kappa_1, ..., \kappa_t\right] \\ &= \begin{cases} \hat{\xi} + \hat{\kappa}_{t+h-1} + \sum_{k=1}^{p} \hat{\phi}_k \hat{\kappa}_{t+h-k} & \text{pour } h > q \\ \hat{\xi} + \hat{\kappa}_{t+h-1} + \sum_{k=1}^{p} \hat{\phi}_k \hat{\kappa}_{t+h-k} + \sum_{k=1}^{q} \hat{\theta}_k \epsilon_{t+h-k} & \text{pour } 1 \leq h \leq q \end{cases}\end{aligned}$$

En se servant de l'hypothèse de bruit blanc des résidus, on peut déterminer un intervalle de confiance pour κ_{t+h} au niveau $1-\alpha$:

$$\kappa_{t+h} - \hat{\kappa}_{t+h} \sim \mathcal{N}\left(0, \sigma^2 \sum_{k=0}^{h} \theta_k^2\right)$$

$$\kappa_{t+h} \in \left[\hat{\kappa}_{t+h} - q_{1-\frac{\alpha}{2}} \times \hat{\sigma}^2 \sqrt{\sum_{k=0}^{h} \hat{\theta}_k^2} \quad ; \quad \hat{\kappa}_{t+h} + q_{1-\frac{\alpha}{2}} \times \hat{\sigma}^2 \sqrt{\sum_{k=0}^{h} \hat{\theta}_k^2}\right]$$

Les Figures 5.11 et 5.12 montrent les meilleures prédictions ainsi que leurs intervalles de confiance à 95%, qui sont plus larges pour le modèle Négatif Binomiale. Pour les femmes, la décroissance est quasi identique entre les modèles tandis qu'elle est différente chez les hommes. On notera une très forte décroissance pour le modèle log-Poisson tandis que dans l'autre, l'extrapolation du paramètre est légèrement incurvée.

Bien que cette méthode d'extrapolation soit simple et séduisante, on peut douter de sa performance sur le long terme. La stricte décroissance de κ_t in-

dique que le niveau général de la mortalité ne sera strictement décroissant dans le temps, alors qu'il semble vraisemblable qu'à une certaine date, ce paramètre devrait rester stable.

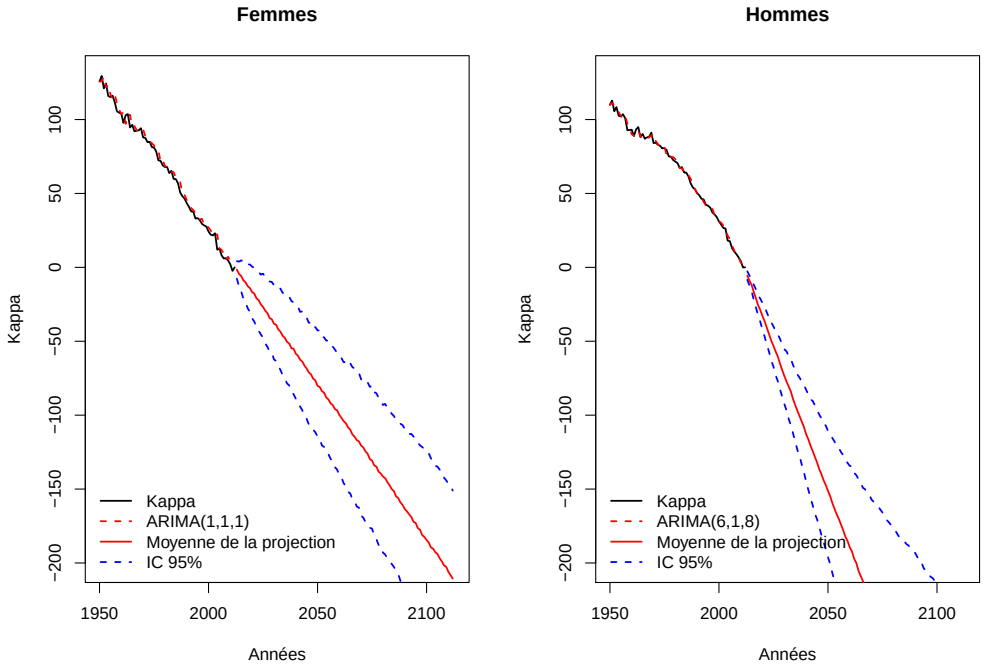

FIGURE 5.11: Projection de κ_t du modèle log-Poisson

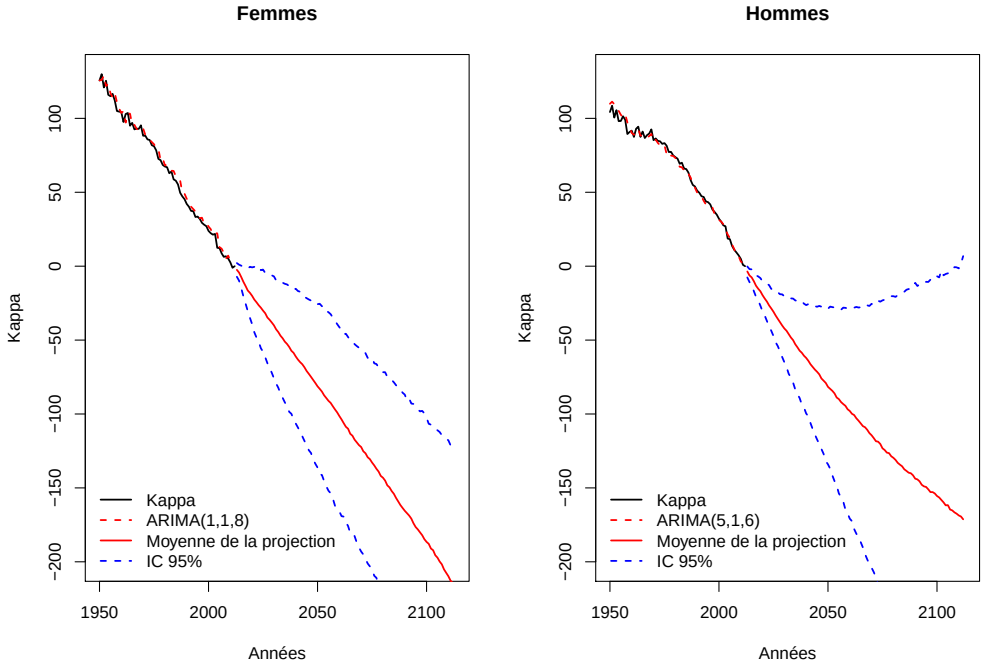

FIGURE 5.12: Projection de κ_t du modèle Négatif Binomiale

5.4 Conclusion du chapitre

Avant de projeter le paramètre, nous avons recalculé les différents estimateurs du modèle, de manière à éviter le "saut" de la première année de projection, qui est préjudiciable dans le calcul d'indicateur comme l'espérance de vie. Ce réajustement s'obtient en posant certaines conditions et modifie complètement les paramètres utilisés précédemment dans la modélisation.

Etant donné que κ_t est un paramètre dépendant du temps, sa modélisation et sa projection peut s'effectuer via une série temporelle. Pour définir cette série, nous avons utilisé la procédure de Box-Jenkins [43]. Il a été nécessaire de procéder à une différenciation de la série κ_t pour la rendre stationnaire, celle-ci ayant une tendance quasi-linéaire. Les paramètres sont estimés à leur maximum de vraisemblance, ce qui nous a permis de calculer des indicateurs d'ajustement du modèle à la série, l'AIC et le BIC. Choisir à ce moment donné le modèle minimisant l'un des deux critères n'est pas optimal puisque les résidus ne répondent pas nécessairement à certaines hypothèses. Qu'importe le sexe, l'article original de Lee-Carter [85], préconise l'utilisation d'un ARIMA(0,1,0), ce qui est repris par de nombreux auteurs. Concernant le modèle log-Poisson, BROUHNS et *al.* [73] préfèrent un ARIMA(0,1,1). Ici, nous avons donc cherché des processus plus complexes, mais plus adéquats.

Les estimateurs des processus choisis sont estimés par maximum de vraisemblance, il est donc possible de déterminer des intervalles de confiance. Dans les projections, nous avons choisi de montrer celui à 95%. Les intervalles obtenus sont plus larges pour le modèle Négative Binomiale. Nous avons également noté que la forme légèrement quadratique de la série des hommes amène à une décroissance importante du paramètre pour le modèle log-Poisson.

Les données sur lesquelles nous travaillons pour la prédiction de la mortalité vont de 1950 à 2012 pour des raisons évoquées en début de chapitre, mais des techniques de sélection de périodes optimales pour calibrer la projection existent :

- BOOTH et *al* [82] utilise la plus longue période sur lesquelles les $\hat{\kappa}_t$ sont linéaire
- DENUIT et *al.* [42] cherche une date de départ sur les dernières années

d'observations, de manière à ce que $\hat{\kappa}_t$ soit représenté par une droite.

- VAN BERKUM et *al.* [37] utilise les changements dans l'indice de mortalité.

Malgré tout, cette approche ne tient pas compte des potentielles améliorations de la longévité apportées par une importante découverte médicale ou de nouveaux traitements. Réciproquement, des détériorations de la longévité humaine à cause d'une épidémie, de nouvelles maladies ou de facteurs inobservables comme la pollution ou les ondes ne sont donc pas pris en compte dans la projection.

Chapitre 6

Intervalles de confiance des modèles et backtest

6.1 Intervalles de confiance par bootstrap

Calculer un indicateur soulève trois sources d'incertitude :

- L'incertitude sur la nature stochastique du modèle, soit ici l'incertitude sur la projection des κ_t
- L'incertitude sur les valeurs des paramètres du modèle
- L'incertitude liée à la sélection du modèle

L'analyse de ces risques est primordial pour l'actuaire car elle permettra, dans notre cas, de tarifer une prime adéquate pour les rentes viagères, tout en considérant une marge de sécurité.
D'après Lee et Carter, la première source d'incertitude citée est la plus importante concernant le risque de longévité, cependant, prendre en compte les deux premiers types de risque s'avère nécessaire.

La complexité des indicateurs des modèles étudiés nous invite à utiliser des simulations. En effet, deux sources différentes d'incertitude sont combinées, l'erreur sur les paramètres des modèles et l'erreur de projection de la série ARIMA(p,d,q). BROUHNS et *al.* [72] ont proposé d'estimer les intervalles de confiance via des simulations de Monte-Carlo, en utilisant le fait que les paramètres du modèle Log-Poisson sont estimés par maximum de vraisemblance et qu'ils sont donc distribués asymptotiquement selon une loi normale multivariée. Cette technique suggère deux choses, que le modèle soit correctement

spécifié et que la taille de l'échantillon soit assez grand pour pouvoir utiliser la loi asymptotique des estimateurs. Mais, RENSHAW et *al.* [113] soulignent que le résultat de cette approche repose beaucoup sur les contraintes du modèle. Etant donné que celles-ci ne sont pas uniques et que ce choix engendre des conséquences sur les résultats des simulations, il est préférable de ne pas avoir recours à cette technique. Nous procéderons alors par Bootstrap.

Plusieurs méthodes de Bootstrap existent, ici nous nous intéressons uniquement au bootstrap paramétrique, qui est introduit par BROUHNS et *al.* [61], et utilise le simple fait que les décès sont distribués selon leur loi supposée (Poisson ou Négative Binomiale). Après avoir ajusté le modèle sur les données et obtenus les paramètres, on répète N fois la procédure suivante :

- **Etape 1** : Simulation de D_{xt} en utilisant le fait que : $D_{xt} \sim \mathcal{P}(E_{xt}\mu_{xt})$ ou $D_{xt} \sim \mathcal{NB}\left(\hat{\hat{\phi}}_x^{-1}, \frac{1}{1+E_{xt}\mu_{xt}\hat{\hat{\phi}}_x}\right)$
- **Etape 2** : Estimation des paramètres du modèle sur les décès simulés
- **Etape 3** : Estimation d'une série chronologique sur les $\hat{\kappa}_t$ et projection
- **Etape 4** : Extrapolation de la surface des taux de décès via la méthode Coale-Kisker
- **Etape 5** : Calcul de l'indicateur souhaité

Deux indicateurs attirent notre attention ici, l'espérance de vie pour un individu de 65 ans en 2012 $(e_{65}(2012))$ et le prix en valeur de marché d'une rente pour un individu de 65 ans en 2012 $(a_{65}(2012))$. Nous choisissons 65 ans à cause de l'âge de la retraite :

$$\begin{aligned} e_{65}(2012) =& \frac{1-\exp(-\mu_{65}(2012))}{\mu_{65}(2012)} \\ &+ \sum_{k=1}^{65}\left\{\prod_{j=0}^{k-1}\exp(-\mu_{65+j}(2012+j))\right\}\frac{1-\exp(-\mu_{65+k}(2012+k))}{\mu_{65+k}(2012+k)} \end{aligned}$$

$$a_{65}(2012) = \sum_{k=1}^{65}\left\{\prod_{j=0}^{k-1}\exp(-\mu_{65+j}(2012+j))\right\} \times \left(\frac{1}{1+r_{2012+k}}\right)^k$$

Avec r_t le taux d'une obligation zéro coupon de maturité t. Ici, étant donné que nous calculons le prix d'une rente d'un individu de 65 ans et que l'âge limite est de 130 ans dans notre modélisation, nous avons besoin d'utiliser les taux de toutes les maturités jusqu'en 2077. La courbe des taux zéro-coupons utilisées est celle des taux swaps fournit par le Comité de Normalisation Obligataire (Figure 6.1). Cette dernière ne proposant pas de taux pour des maturités supérieures à 60 ans, nous fixons un *Ultimate Forward Rate* (i.e $r_{61,\ldots,65}^{31/12/2012} = r_{60}^{31/12/2012}$).

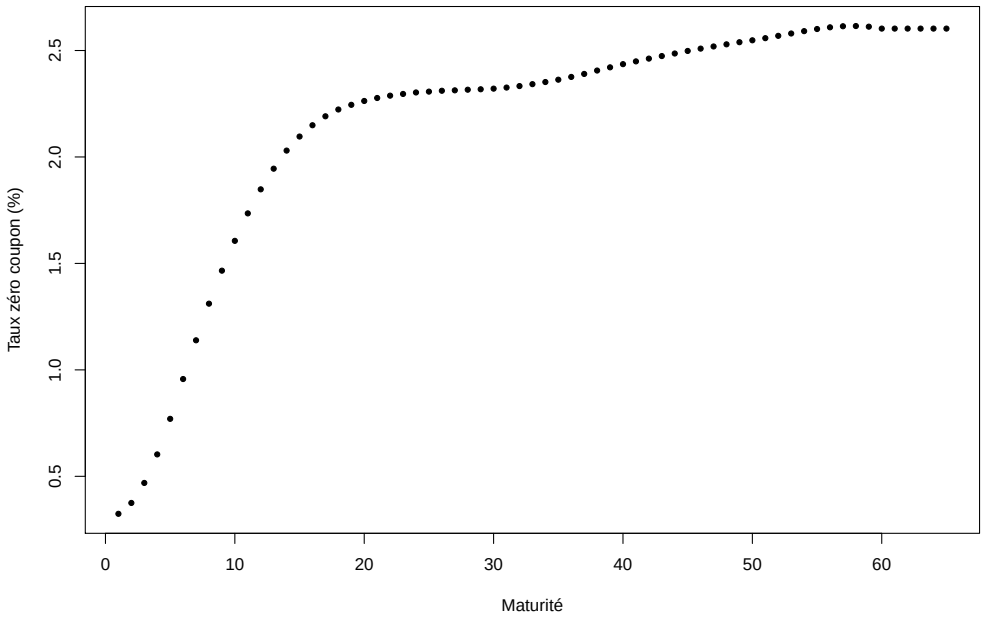

FIGURE 6.1: Taux zéro-coupon au 31/12/2012

Les Figures 6.2 et 6.3 représentent le résultat des 2000 simulations de bootstrap paramétriques sur les paramètres pour du modèle log-Poisson. Chez les femmes, les paramètres α_x et β_x sont plus volatiles que chez les hommes, surtout entre 0 et 40 ans. Les Figures 6.4 et 6.5 sont les simulations pour le modèle Négatif Binomiale. On note la même tendance que pour le précédent modèle sauf que la volatilité est moindre. Ceci s'explique vraisemblablement par la présence du paramètre $\bar{\phi}_x$, qui ce dernier a un intervalle de confiance très étroit. Comme attendu, les intervalles de confiance des paramètres κ_t sont très proches des projections des Figures 5.11 et 5.12.
Les Figures 6.6 et 6.7 représentent les espérances de vie, tandis que les Figures 6.8 et 6.9 représentent le prix d'une rente sous forme d'histogramme. Comme l'on pouvait s'y attendre par rapport à la projection du paramètre κ_t,

les intervalles de confiance des indicateurs du modèle Négatif Binomiale sont plus larges que ceux du modèle log-Poisson. Les quantiles à 0%, 5%, 20%, 50%, 80%, 95% et 100% sont donnés dans des tables en annexes.

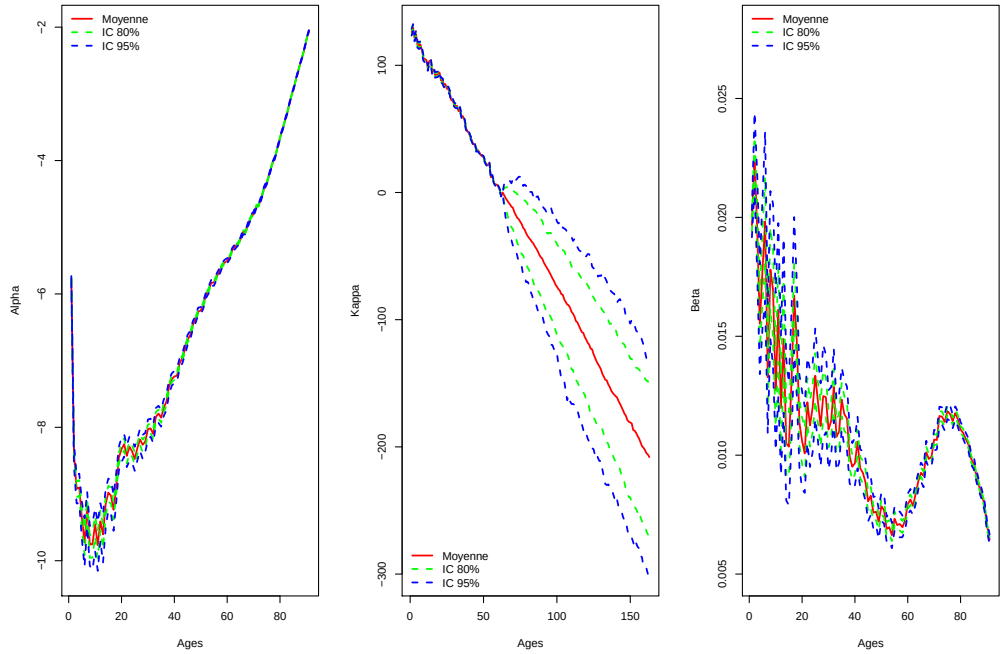

FIGURE 6.2: 2000 simulations de bootstrap paramétriques sur le modèle log-Poisson (femme)

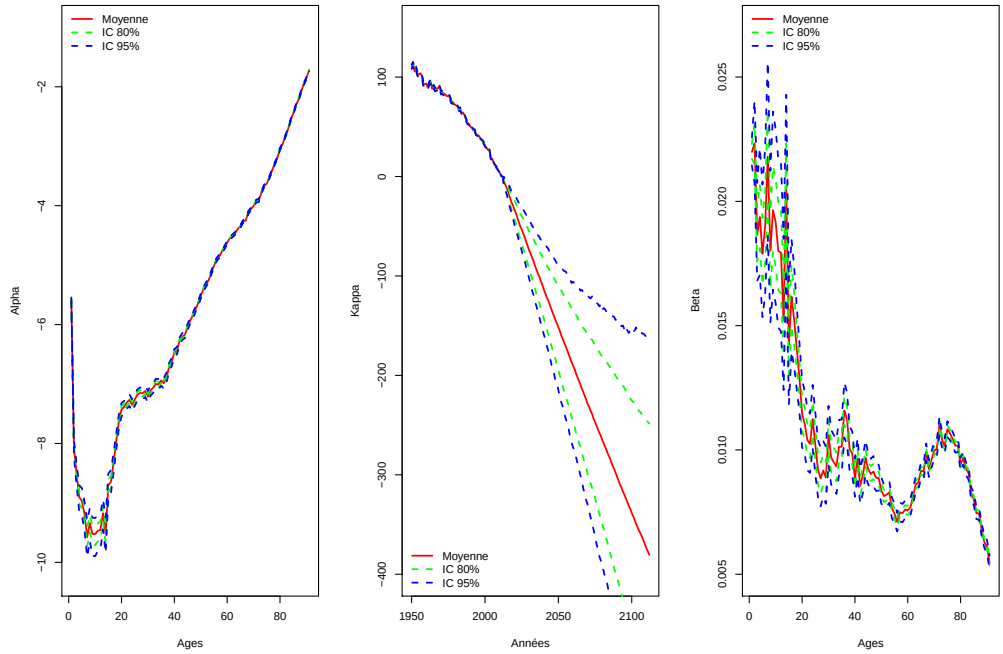

FIGURE 6.3: 2000 simulations de bootstrap paramétriques sur le modèle log-Poisson (homme)

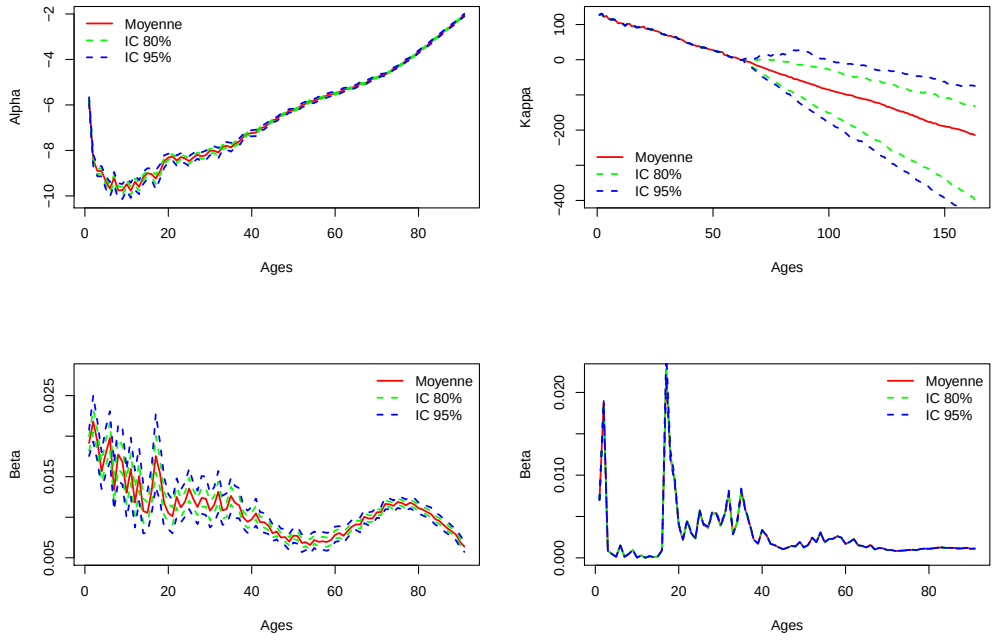

FIGURE 6.4: 2000 simulations de bootstrap paramétriques sur le modèle Négatif Binomiale (femme)

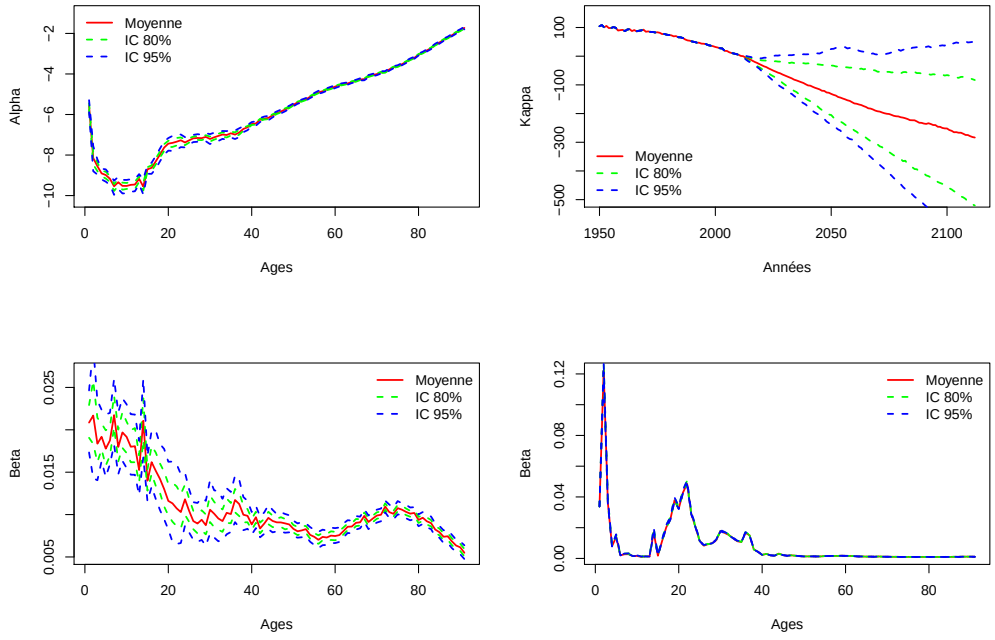

FIGURE 6.5: 2000 simulations de bootstrap paramétriques sur le modèle Négatif Binomiale (homme)

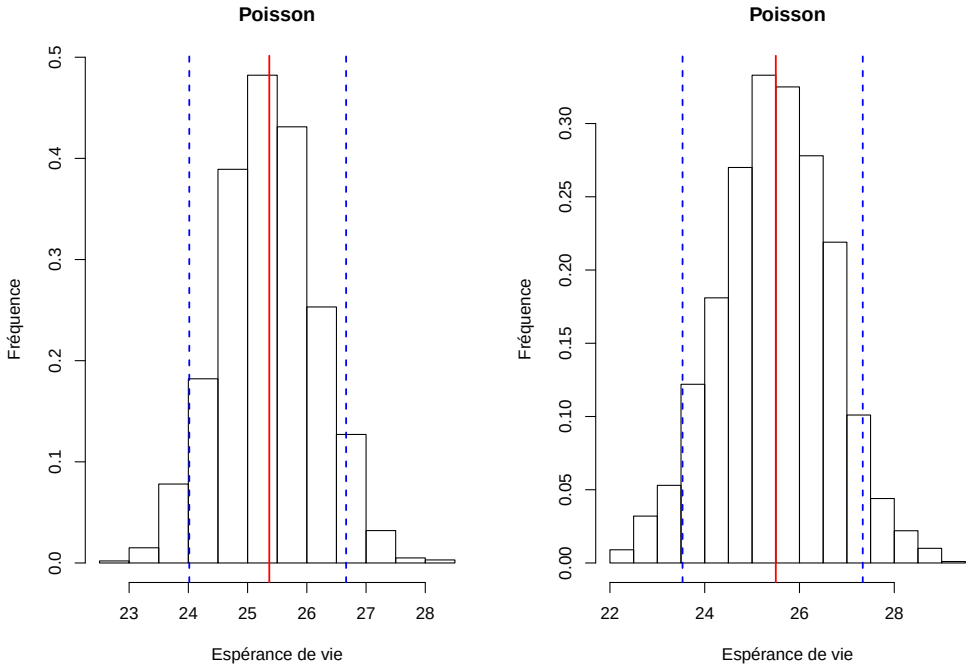

FIGURE 6.6: Histogramme de $e_{65}(2012)$ pour les femmes

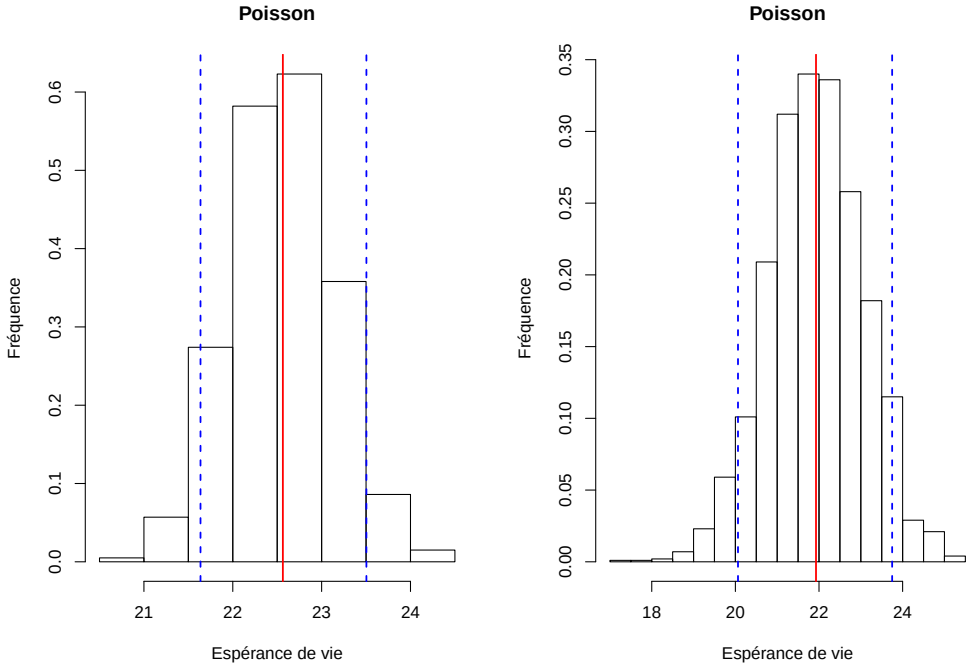

FIGURE 6.7: Histogramme de $e_{65}(2012)$ pour les hommes

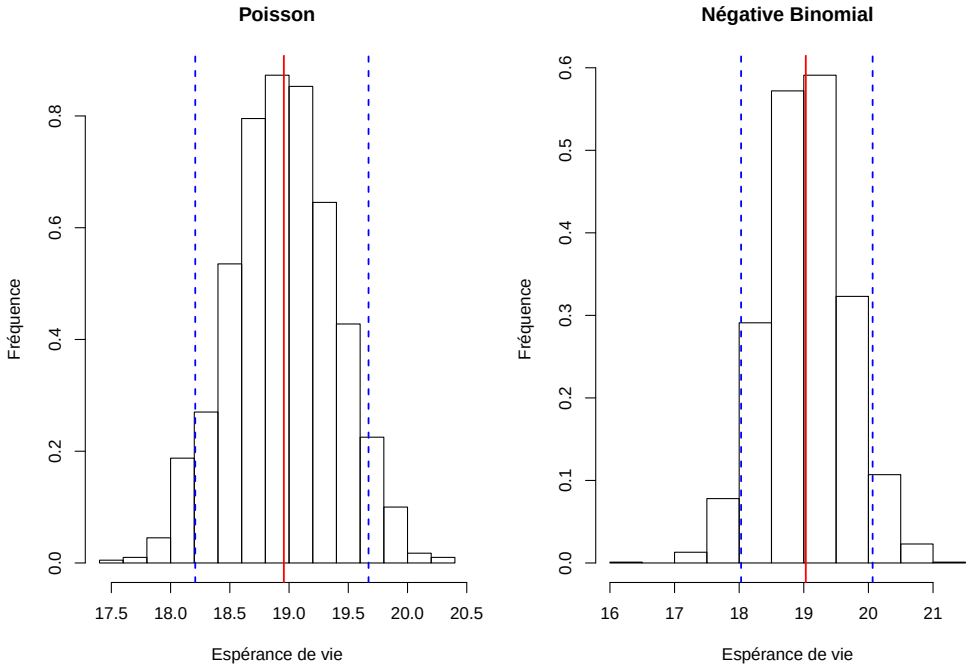

FIGURE 6.8: Histogramme de $a_{65}(2012)$ pour les femmes

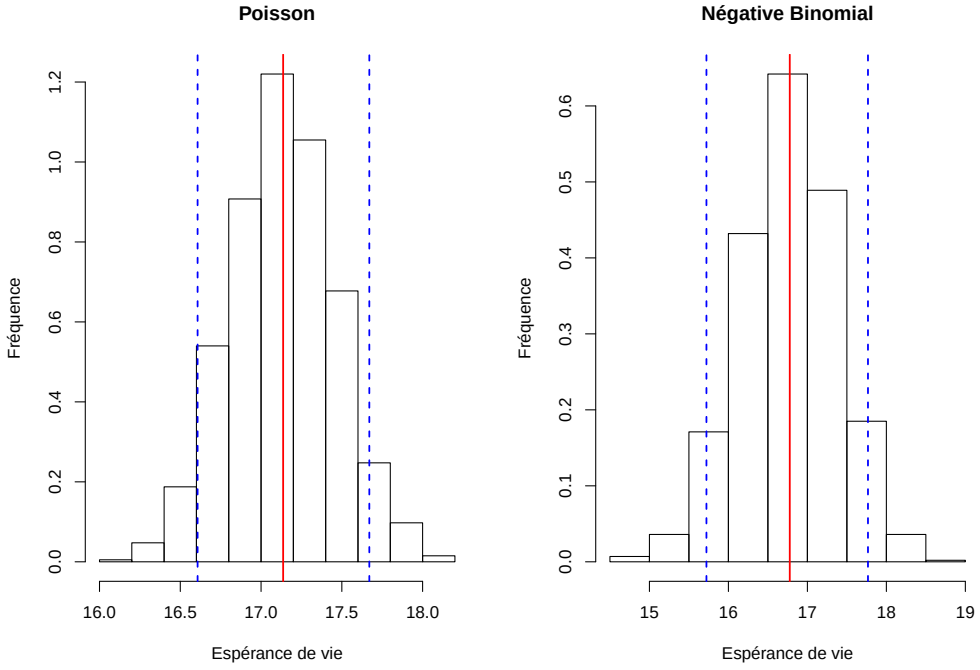

FIGURE 6.9: Histogramme de $a_{65}(2012)$ pour les hommes

Le principal inconvénient du Bootstrap paramétrique est qu'il est applicable uniquement en présence d'une éventuelle distribution. Le modèle original de Lee-Carter n'est donc pas utilisable avec cette méthode.
D'autres type de bootstrap ont été appliqués aux modèles de mortalité :

- BROUHNS et *al.* [61] introduisent le bootstrap non paramétrique, en supposant que les décès peuvent être approchés par une loi multinomiale. Une bonne explication sur la manière de procéder peut être trouvé dans DELWARDE [10].

- KOISSI et *al.* [59] utilisent un bootstrap résiduel, qui consiste a simuler de nouvelles matrices de décès, à l'aide des résidus de Pearson ou de déviance.

- LIU et *al.* [68] postulent que l'hypothèse d'indépendance et d'identique distribution des résidus n'est pas tout à fait correcte à cause de l'existence de certaines corrélations entre l'âge et le temps, ce qui entraîne une sous estimation du risque de mortalité. Ils se servent alors du bloc bootstrap résiduel pour répondre à cet inconvénient en gardant la structure de dépendance sous-jacente des résidus, et en générant des échantillons plus réalistes. La dépendance spatiale des résidus est donc prise en compte, au moins localement. La principale difficulté de ce type de simulation réside dans la détermination de la taille des blocs.

- Plus récemment, pour réduire la variance qui réside dans les techniques par simulations, D'AMATO et *al.* [24] utilisent un bootstrap stratifié, quasi similaire à un block bootstrap.

6.2 Backtest

La meilleure façon de tester la prédiction des modèles est d'établir un backtest, comme le souligne DOWD et *al.* [93]. Ici, nous ferons un backtest de 21 ans, sur la période de 1991 à 2012. Nous regarderons la qualité de prédiction du modèle sur différents taux de décès de manière à pouvoir déterminer jusqu'à quel horizon de risque le modèle semble applicable. La procédure est la suivante :

- Simulations de 1000 bootstrap paramétriques sur une période de 1950 à 1990

- Détermination des surfaces de mortalité à différents quantiles (5%, 50% et 95%)
- Détermination de la surface de mortalité bruts entre 1950 et 2012.
- Analyse des taux de décès

Ici, nous ne réeffectuons pas toute l'analyse concernant κ_t et supposons que ce paramètre peut être modélisé de la même manière que sur la période d'observation de 1950 à 2012.

Les résultats sont présentés sur les Figures 6.10 et 6.11. Nous avons choisi de montrer les taux de décès (log) pour différents âges (15, 40, 60 et 70 ans). Comme on peut le voir, les intervalles de confiance du modèle Négatif Binomiale sont plus larges que ceux du log-Poisson.
Chez les femmes, les modèles choisis ne sont pas très performants lors des premières années pour les âges de 15 et 40 ans mais, les intervalles de confiance des âges plus élevés (60 et 70 ans), sont plus performants étant donné la décroissance quasi-linéaire des taux de décès.
Chez les hommes, on retrouve le même problème pour les premières années de projection, mais ici, pour les âges les plus élevés, la décroissance des taux n'est pas quasi-linéaire, ce qui entraîne la sortie de quelques observations de l'intervalle de confiance à 95% pour le modèle log-Poisson (à 40 et 70 ans entre autre).

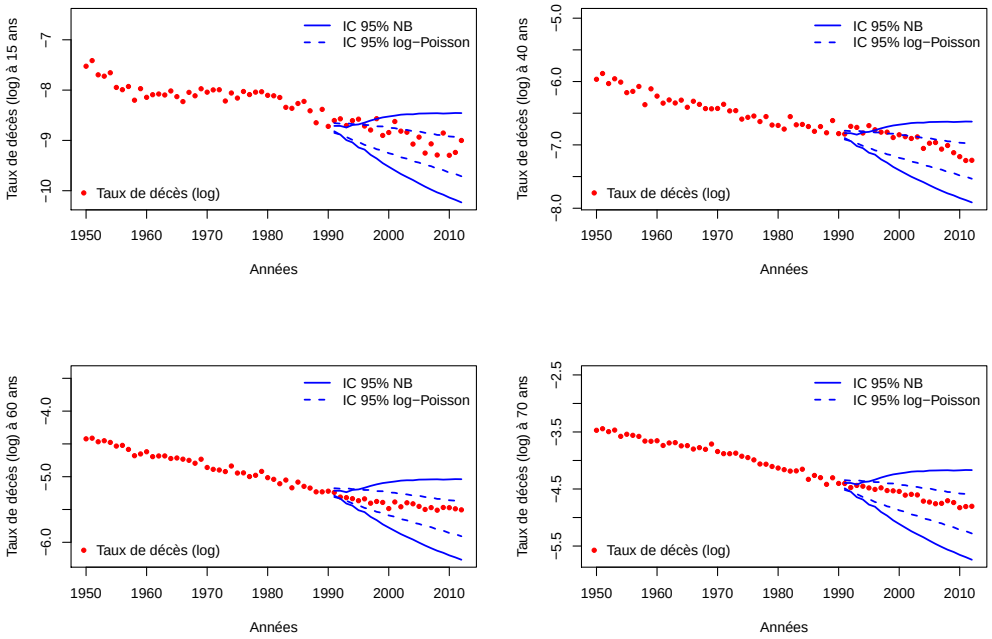

FIGURE 6.10: Quelques taux de décès (log) chez les femmes

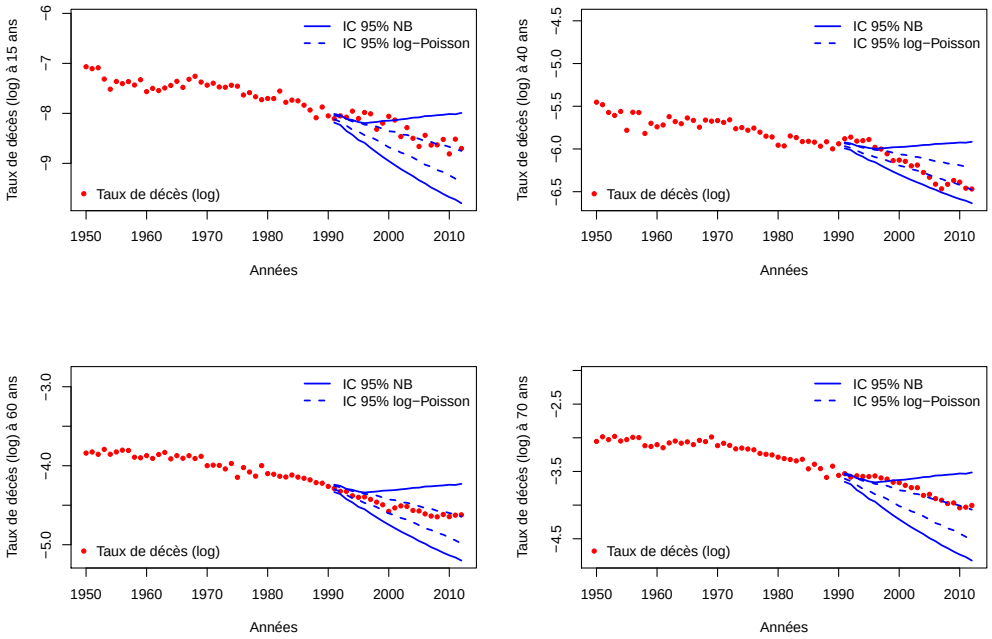

FIGURE 6.11: Quelques taux de décès (log) chez les hommes

Conclusion Générale

La longévité des individus est un risque important pour les assureurs-vie, principalement pour les produits viagers, et sa modélisation est primordiale. Pour tarifer un produit comme les rentes viagères, la création de tables prospectives s'avère nécessaire car l'utilisation de tables instantanées provoquerait une mauvaise tarification et pourrait engendrer un sous-provisionnement.

Depuis le 17^{e}siècle, les démographes et les actuaires s'intéressent alors à modéliser et projeter la mortalité. On recense aujourd'hui un grand nombre de modèles, qu'ils soient paramétriques ou stochastiques. Le modèle de LEE et CARTER [85] est certainement le plus connu puisque qu'il a fait ses preuves dans de nombreuses études de populations [20]. En modélisant la force de mortalité, il s'avère très pratique grâce à son faible nombre de paramètres facilement interprétables et sa simplicité ; il est aujourd'hui utilisé par certaines administrations publiques des Etats-Unis. Néanmoins, ce modèle dispose de plusieurs inconvénients, comme la supposition de l'homoscédasticité des taux de mortalité.
En cherchant directement à modéliser le nombre de décès par une distribution de Poisson pour chaque âge et chaque année, BROUHNS et *al* [73] soulage le modèle de cette hypothèse. Les résultats sont plus performants. Cependant, la loi de Poisson est trop restrictive à cause de la propriété d'équidispersion. En effet, il est rare que celle-ci soit vérifiée. Considérer un facteur de dispersion ayant pour but de représenter l'hétérogénéité dans la population s'avère être bénéfique.
Supposer les individus de même risque peut paraître incohérent étant donné que des facteurs observables ou inobservables, autres que l'âge et le temps, influencent le risque de mortalité. Ce paramètre est très variable selon l'âge et sa prise en compte provoque une amélioration de la qualité d'ajustement selon le AIC et le BIC, pour les deux sexes. On suppose alors que les décès sont distribués selon une distribution négative binomiale.

La variabilité du nombre de décès aux âges extrêmes dans le temps provoque une mauvaise adéquation des modèles aux alentours des 80 ans. Il est alors important de passer par une méthode dites de "fermeture de tables". Nous en avons ici comparé deux et choisi d'utiliser celle de DENUIT et GODERNIAUX [42] qui s'ajuste mieux à nos données.

La projection de la mortalité par le modèle de Lee-Carter peut s'effectuer de plusieurs manières. On retrouve dans la littérature plusieurs méthodes d'extrapolation du paramètre représentant le niveau de la mortalité dans le temps, que ce soit par simple régression linéaire ou bien par des complexes modèles de sauts. La manière la plus répandue est un processus ARIMA(p,d,q). Lee et Carter utilise une marche aléatoire simple, mais ce modèle viole les diverses hypothèses faites concernant les séries temporelles. Nous avons chercher de meilleurs processus répondant correctement à ces dernières.

Le calcul d'indicateur soulève plusieurs sources d'incertitudes. Ici, nous nous sommes intéressés à celles concernant la projection du paramètre dépendant du temps et la valeur des paramètres des modèles. En passant par des simulations de bootstrap, nous avons pu déterminer des intervalles de confiance aussi bien sur les paramètres du modèle que sur des indicateurs comme l'espérance de vie ou le prix d'une rente. De manière à tester la qualité prédictive des deux modèles, nous avons procédé à un backtest. Le fait que les intervalles de confiance soient plus larges pour le modèle Négatif Binomiale, moins d'observations sortent de ce dernier.

Les modèles que nous avons abordés ne sont pas directement applicables à un portefeuille d'assurés. En effet, ceux-ci détiennent trop peu de données mais également des périodes d'observations beaucoup plus courtes. Les assureurs ne peuvent également pas directement prendre les données de la population nationale sur celles de leur portefeuille, étant donné que le risque n'est pas le même. On parle de risque d'antisélection, qui découle d'une asymétrie entre l'assureur et l'assuré, et est très persceptible sur le marché des rentes viagères, car la plupart des contractants espèrent vivre plus longtemps que la moyenne. Pour prendre en compte ce type de risque, il est préférable d'utiliser des modèles relationnels, comme celui de COX ou de BRASS, qui utilise tout de même les taux de mortalité de la population nationale. Sur le modèle Négatif Binomiale, HOK KWAN [80] utilise une approche bayésienne,

qui a l'avantage de pouvoir être appliqué en cas de données limitées ou manquantes.
On note souvent la présence de sur-dispersion dans les portefeuilles d'assurance-vie, pas uniquement à cause de l'hétérogénéité, mais également dû à la présence de contrats "doubles" dans les portefeuilles. En effet, il est courant qu'un contractant détienne plusieurs contrats d'assurance, ce qui a pour but d'augmenter la variance du nombre de décès [17].

Dans ce mémoire, nous n'avons pas cherché à améliorer le modèle de LI et *al.* [81], mais plusieurs pistes pourraient être explorées, comme la prise en compte de paramètres servant à expliquer l'effet de cohorte ou bien rajouter des termes extra-bilinéaires temporels. Nous avons émis l'hypothèse que l'hétérogénéité des individus pouvait être approchée par une loi Gamma. Nous avons introduit la distribution Inverse-Gaussienne, qui peut être une très bonne alternative, car la distribution Poisson Inverse-Gausienne permet de modéliser aussi bien la sur-dispersion que la sous-dispersion dans les données. Cependant, les calculs semblent plus laborieux. Une approche MCMC, qui permettrait d'estimer les paramètres et d'effectuer la projection en même temps, est possible en utilisant l'approche de CZADO et *al.* [91] sur le modèle log-Poisson.

Table des figures

2.1 Surface de mortalité française de 1816 à 2012 pour des femmes de 0 à 100 ans 14
2.2 Surface de mortalité française de 1816 à 2012 pour des hommes de 0 à 100 ans 14
2.3 Coupe longitudinale de la surface de mortalité des femmes . . . 15
2.4 Coupe longitudinale de la surface de mortalité des hommes . . 15
2.5 Taux de décès pour des femmes à des âges différents 16
2.6 Taux de décès pour des hommes à des âges différents 16
2.7 Probabilités de survie à la naissance des femmes 17
2.8 Probabilités de survie à la naissance des hommes 17
2.9 Evolution de l'espérance de vie des femmes 18
2.10 Evolution de l'espérance de vie des hommes 18
2.11 Comparaison des taux de décès (log) pour différentes cohortes . 19

4.1 Paramètres du modèle de Lee-Carter 34
4.2 Variance des résidus du modèle de Lee-Carter selon l'âge 34
4.3 Paramètres du modèle de log-Poisson 38
4.4 Dispersion des décès selon l'âge 38
4.5 Paramètres du modèle Binomiale Négatif 50
4.6 Variance Temporelle des différents modèles pour les femmes . . 52
4.7 Variance Temporelle des différents modèles pour les hommes . 52
4.8 Résidus du modèle de Lee-Carter pour les femmes 56
4.9 Résidus du modèle de Lee-Carter pour les hommes 56
4.10 Résidus de Pearson du modèle log-Poisson des femmes 57
4.11 Résidus de Pearson du modèle log-Poisson des hommes 57
4.12 Résidus de Pearson du modèle Negatif Binomiale des femmes . 58
4.13 Résidus de Pearson du modèle Negatif Binomiale des hommes . 58
4.14 Quelques taux de décès (log) aux âges extrêmes chez les femmes 62
4.15 Quelques taux de décès (log) aux âges extrêmes chez les hommes 62

5.1 Paramètres du modèle log-Poisson pour la projection 68
5.2 Paramètres du modèle Négatif Binomiale pour la projection . . 68
5.3 Série κ_t différenciée à l'ordre 1 du modèle log-Poisson 71
5.4 Série κ_t différenciée à l'ordre 1 du modèle Négatif Binomiale . . 71
5.5 ACF et PACF de la série κ_t différenciée à l'ordre 1 du modèle log-Poisson 72
5.6 ACF et PACF de la série κ_t différenciée à l'ordre 1 du modèle Négatif Binomiale 72
5.7 Modèle log-Poisson femme : Analyse des résidus du ARIMA(1,1,1) 78
5.8 Modèle log-Poisson homme : Analyse des résidus du ARIMA(6,1,8) 78
5.9 Modèle Négatif Binomiale femme : Analyse des résidus du ARIMA(1,1,8) 79
5.10 Modèle Négatif Binomiale homme : Analyse des résidus du ARIMA(5,1,6) 79
5.11 Projection de κ_t du modèle log-Poisson 82
5.12 Projection de κ_t du modèle Négatif Binomiale 82

6.1 Taux zéro-coupon au 31/12/2012 87
6.2 2000 simulations de bootstrap paramétriques sur le modèle log-Poisson (femme) 89
6.3 2000 simulations de bootstrap paramétriques sur le modèle log-Poisson (homme) 89
6.4 2000 simulations de bootstrap paramétriques sur le modèle Négatif Binomiale (femme) 90
6.5 2000 simulations de bootstrap paramétriques sur le modèle Négatif Binomiale (homme) 90
6.6 Histogramme de $e_{65}(2012)$ pour les femmes 91
6.7 Histogramme de $e_{65}(2012)$ pour les hommes 91
6.8 Histogramme de $a_{65}(2012)$ pour les femmes 92
6.9 Histogramme de $a_{65}(2012)$ pour les hommes 92
6.10 Quelques taux de décès (log) chez les femmes 95
6.11 Quelques taux de décès (log) chez les hommes 95
12 ACF et PACF de la série κ_t modèle log-Poisson 115
13 ACF et PACF de la série κ_t modèle Négative Binomiale 115

Liste des tableaux

4.1 Valeurs de la statististique du test de dispersion 40
4.2 Indicateurs de la qualité des modèles log-Poisson et Binomial Négatif 53

5.1 Numéros de l'observations sortant de l'IC à 95% de l'ACF et de la PACF 70

1 Qualité des modèles ARIMA pour le κ_t du modèle log-Poisson (femme) 116
2 Qualité des modèles ARIMA pour le κ_t du modèle log-Poisson (homme) 116
3 Qualité des modèles ARIMA pour le κ_t du modèle Négative Binomiale (femme) 116
4 Qualité des modèles ARIMA pour le κ_t du modèle Négative Binomiale (homme) 116
5 Comparaison des p-values des tests sur les résidus des ARIMA du modèle log-Poisson (femme) 117
6 Comparaison des p-values des tests sur les résidus des ARIMA du modèle log-Poisson (homme) 117
7 Comparaison des p-values des tests sur les résidus des ARIMA du modèle Négative Binomiale (femme) 117
8 Comparaison des p-values des tests sur les résidus des ARIMA du modèle Négatif Binomiale (homme) 118
9 Quantile de $e_{65}(2012)$ pour les femmes 119
10 Quantile de $e_{65}(2012)$ pour les hommes 119
11 Quantile de $a_{65}(2012)$ pour les femmes 119
12 Quantile de $a_{65}(2012)$ pour les hommes 119
13 Paramètres α_x, β_x et ϕ_x des différents modèles (femmes) pour la modĺisation Chapitre 4 121

14 Paramètres α_x, β_x et ϕ_x des différents modèles (hommes)pour la modĺisation Chapitre 4 123
15 Paramètres κ_t des différents modèles (hommes) pour la modĺisation Chapitre 4 125
16 Paramètres α_x, β_x et ϕ_x des différents modèles (femmes) pour la projection Chapitre 4 127
17 Paramètres α_x, β_x et ϕ_x des différents modèles (hommes) pour la projection Chapitre 5 129
18 Paramètres κ_t des différents modèles (hommes) pour la projection du Chapitre 5 131

Bibliographie

[1] DORAY A. Inference for logistic-type models for the force of mortality. *Présentation au "Living to 100" Survival to Advanced Ages international symposium Society of Actuaries Orlando Florida*, 2008.

[2] MCNOWN R ET ROGERS A. Forecasting mortality and parametrized time series approach. *Demography*, 26, 1989.

[3] OLIVIEIRI A. Heterogeneity in survival models applications to pensions and life annuities. *Belgian Actuarial Bulletin*, 6, 2006.

[4] PITTACO G ET DENUIT M ET HABERMAN S ET OLIVIERI A. *Modelling longevity dynamics for pensions and annuity business*. Oxford University Press, 2009.

[5] PLANCHET F ET KAMEGA A. Mortalité prospectives en cas de petits échantillon : modélisation à partir d'informations externes en utilisant l'approche de bongaarts. *Assurances et gestion des risques*, 80, 2011.

[6] WIENKE A. *Frailty models in survival analysis*. Chapman and Hall / CRC Biostatistics Series, 2010.

[7] HABERMAN S ET RENSHAW AE. Measurement of longevity risk using bootstrapping for lee-carter and generalised linear poisson models of mortality. *Computational Applied Probability*, 11, 2009.

[8] HABERMAN S ET RENSHAW AE. Modelling and projecting mortality improvement rates using a cohort perspective. *Insurance : Mathematics and Economics*, 53, 2012.

[9] CAIRNS AJG. A discussion of parameter and model uncertainty in insurance. *Insurance : Mathematics and Economics*, 27, 2000.

[10] DELWARDE Antoine. *Modèle log-bilinéaire pour l'élaboration de tables de mortalité prospectives*. PhD thesis, Université de Louvain, 2005.

[11] THATCHER AR. The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical SociETy*, 162, 1999.

[12] EFRON B. Bootstrap methods : Another look at the jacknife. *Annals of Statistics*, 7, 1979.

[13] GOMPERTZ B. On the nature of the function expressive of the law of human mortality and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, 115, 1825.

[14] EILERS PHC ET MARX BD. Flexible smoothing with b-splines and penalties. *Statistical Science*, 11, 1996.

[15] Harry BENSUSAN. *Risque de de taux et de longévité : Modélisation dynamique et applications aux produits dérivés et à l'assurance vie*. PhD thesis, Ecole Polytechnique, 2010.

[16] CMI Bureau. Considerations affecting the preparation of tandard tables of mortality. *Journal of the Institute of Actuaries*, 101, 1974.

[17] DELWARDE A ET DENUIT M ET PARTRAT C. Negative binomial version of the lee-carter model for mortality forecasting. *Applied Stochastic Models in Business*, 23, 2007.

[18] GAO Q ET HU C. Dynamic mortality factor model with conditional heteroskedascity. *Insurance : Mathematics en Economics*, 40, 2009.

[19] PEDROZA C. A bayesian forecasting model : Predicting us male mortality. *Biostatistics*, 7, 2006.

[20] TULJAPURKAR S ET LI N ET BOE C. A universal pattern of mortality decline in the g7 countries. *Nature*, 2000.

[21] DEAN CB. Testing for overdispersion in poisson and binomial regression models. *Journal of American Statistical Association*, 87, 1992.

[22] HAINAUT D. Multidimensional lee-carter model with switching mortality processes. *Insurance : Mathematics and Economics*, 50, 2011.

[23] SERANT D. Construction de tables prospectives de mortalité. *Document interne FFSA (Confidentiel)*, 2005.

[24] et HABERMAN S et RUSSOLILLO M D'AMATO V. The stratified sampling bootstrap for measuring the uncertainty in mortality forecasts. *Computational Applied Probability*, 14, 2012.

[25] Makram BEN DBABIS. *Mortality table, prospective approach and stochastic mortality*. PhD thesis, Université Paris Dauphine, 2009.

[26] Departement de démographie de l'UniversitÃ© de Berkeley et Max Planck Institute. Human mortality database. http ://www.mortality.org, Mis à jour le 10 janvier 2014.

[27] FORFAR DO ET SMITH DM. The changing shape of english life tables. *Transactions of the Faculty of Actuaries*, 40, 1988.

[28] BRILLINGER DR. The natural variability of vital rates and associated statistics. *Biometrics*, 42, 1986.

[29] COX DR. Regression models and life-tables. *Journal of the Royal Society Series B (Methodological)*, 34, 1972.

[30] BIFFIS E. Affine processes for dynamic mortality and actuarial valuations. *Working paper*, 2004.

[31] COSSETTE H ET DELWARDE A ET DENUIT M ET GUILLOT F ET MARCEAU E. Pension plan valuation and mortality projection : a case study with mortality data. *North American Actuarial Journall*, 11, 2007.

[32] HALLEY E. An estimate of the degrees of the mortality of mankond drawn from curious tables of the births and funerals at the city of breslaw ; with an attempt to ascertain the price of annuities upon lives. *Philosophical Transactions*, 17, 2011.

[33] JEVTIC P ET LUCIANO E ET VIGNA E. Mortality surface by means of continuous time cohort models. *Insurance : Mathematics and Economics*, 53, 2012.

[34] MANTON KG ET STALLARD E. Recent trend in mortality analysis. *Academic Press*, 1986.

[35] VAUPEL JW ET MANTON KG ET STALLARD E. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16, 1979.

[36] MCGILCHRIST CA et AISBETT CW. Regression with frailty in survival analysis. *Biometrics*, 47, 1991.

[37] VAN BERKUM F. et ANTONIO K. et VELLEKOP M. Structural changes in mortality rates with an application to dutch and belgian data. Disponible sur https ://lirias.kuleuven.be, 2013.

[38] HANAGAL DD et DABADE AD. A comparative study of shared frailty models for kidney infection data with generalized exponential baseline distribution. *Journal of Data Science*, 1, 2013.

[39] BROCKWELL PJ et DAVIS RA. *Introduction to time series and forecasting*. Springer, 2002.

[40] QUASHIE A. et DENUIT M. Modèles d'extrapolation de la mortalité aux grand âges. Disponible sur https ://www.uclouvain.be, 2005.

[41] HAINAULT D et DEVOLDER P. Mortality modelling with levy processes. *Insurance : Mathematics and Economics*, 42, 2008.

[42] DENUIT M. et GODERNIAUX A-C. Closing and projecting lifetables using log-linear models. *Bulletin de l'Association Suisse des Actuaires*, 2005.

[43] BOX G et JENKINS G. *Time series analysis : Forecasting and control*. San Francisco : Holden-day, 1970.

[44] COALE A. et KISKER E.E. Defects in data on old age mortality in the united states : New procedures for calculatind approximately accurate schedules and life tables at the highest ages. *Asian and Pacific Poupaltion Forum*, 1, 1990.

[45] NORDMAN DJ et LAHIRI SN et FRIDLEY BL. Optimal block size for variance estimation by a spatial block bootstrap method. *The Indian Journal of Statistics*, 69, 2007.

[46] Eugène KOUADIO et Vincent PONCELET. *Modélisation de la mortalité à l'aide processus de Levy*. PhD thesis, ENSAE, 2011.

[47] DICKEY D.A. et WAYNE A.F. Distribtuion of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 1979.

[48] GIACOMETTI R ET BERTOCCHI M ET RACHEV ST ET FABOZZI F. A comparison of the lee-carter model and ar-arch model for forecasting mortality rates. *Insurance : Mathematics and Economics*, 50, 2012.

[49] GRECO F ET SCALONE F. A space-time extension of the lee-carter model in a hierarchical bayesian. Disponible sur http ://www.unece.org/, 2013.

[50] GUIBERT Q ET PLANCHET F. Quels sont les risques associés à un régime de retraite ? *La lettre de l'Observatoire des Retraites*, 2013.

[51] KAMEGA A ET PLANCHET F. Hétérogénéité : mesure du risque d'estimation dans le cas d'une modélisation intégrant des facteurs observales. *Bulletin FranÃ§ais d'Actuariat*, 21, 2011.

[52] PLANCHET F. Tables de mortalité d'expérience pour des portefeuilles de rentiers. *Note mÉThodologiques de l'Institut des Actuaires*, 2005.

[53] PLANCHET F. Modèle stochastiques de mortalité. *Support de Cours 2012-2013 ISFA sur www.ressources-actuarielles.net*, 2013.

[54] TOMAS J ET PLANCHET F. Prospective mortality tables : taking heterogeneity into account. *Les cahiers de recherches de l'ISFA*, 2013.

[55] PLANCHET Frédéric. *Risque de longévité et détermination du besoin en capital : travaux en cours*. PhD thesis, ISFA, 2008.

[56] SCHWARZ G. Estimating the dimension of a model. *Annals of Statistics*, 6, 1978.

[57] KLUGMAN SA ET PANJER HJ ET WILLMOT GE. *Loss Models : from data to decisions*. New York : Wiley, 2012.

[58] AKAIKE H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 :716–723, 1974.

[59] KOISSI MC ET SHAPIRO AF ET HOGNES. Evaluating and extending the lee-carter model for mortality forecasting : Bootstrap confidence interval. *Insurance : Mathematics and Economics*, 38, 2006.

[60] DICKSON DCM ET HARDY MR ET WATERS HR. *Actuarial mathematics for life contingent risk*. Cambridge University Press, 2009.

[61] BROUHNS N ET DENUIT M ET VAN KEILEGOM I. Boostrapping the poisson log-bilinear model for mortality forecasting. *Scandinavian Actuarial Journal*, 42, 2005.

[62] CAIRNS AJG ET BLAKE D ET DOWD K ET COUGHLAN GD ET EPSTEIN D ET ONG A ET BALEVICH I. A quantitative comparison of stochastic mortality models using data from england and wales and the united states. *North American Actuarial Journal*, 13, 2009.

[63] CURRIE ID. Smoothing and forecasting mortality rates with p-splines. *Présentation à l'Institute of Actuaries*, 2006.

[64] BARBI E ET CASELLI G ET VALLIN J. Hétérogénéité des générations et âge extrêmes de la vie. *Population*, 58, 2003.

[65] BONGAARTS J. Long-range trends in adult mortality : models and projections methods. *Population Council*, 2004.

[66] GRAUNT J. Natural and political observations mentioned in a following index and made upon the bills of mortality. *John Martyn and James Allestry*, 1662.

[67] LI J. An application of MCMC simulation in motality projection for populations with limited data. *Demographic Research*, 30, 2014.

[68] LIU X ET BRAUN J. Investigating mortality uncertainty using the block bootstrap. *Journal of Probability and Statistics*, 2010.

[69] MILNE J. *A treatise on the valuation of annuities and assurances on lives and survivorships ; on the construction of tables of mortality ; and on the probabilities and expectations of life*. Longman, 1815.

[70] CARRIERE JF. Parametric models for life tables. *Transactions of society of actuaries*, 44, 1992.

[71] HELIGMAN L ET POLLARD JH. The age pattern of mortality. *Journal of the Institute of Actuaries*, 107, 1980.

[72] BROUHNS N ET DENUIT M ET VERMUNT JK. Measuring the longevity risk in mortality projections. *Bulletin of the Swiss Association of Actuaries*, 2002.

[73] BROUHNS N ET DENUIT M ET VERMUNT JK. A poisson log-bilinear regression approach to the construction of projected life-tables. *Insurance : Mathematics and Economics*, 31, 2002.

[74] ALHO JM. Discussion of lee (2000). *North American Actuarial Journal*, 4 :91–93, 2000.

[75] WILMOTH JR. Computational methods for fitting and extrapolating the lee-carter model of mortality change. *Technical Report*, 1993.

[76] WILMOTH JR. Demography of longevity : past present and future trends. *Experimental Gerontology*, 35, 2000.

[77] VAUPEL JW. Life expentacy at current rates vs current conditions : a reflexion stimulated by bongaarts and feeney's "how long do we live ?". *Demographic Research*, 7, 2002.

[78] CAIRNS AJG ET BLAKE D ET DOWD K. A two-factor for stochastic mortality with parameter uncertainty : therory and calibration. *The Journal of Risk and Insurance*, 73, 2006.

[79] MITCHELL D ET BROCKETT P ET MENDOZA-ARRIAGA R ET MUTHURAMAN K. Modeling and forecasting mortality rates. *Insurance : Mathematics and Economics*, 52, 2013.

[80] HOK KWAN Kan. *A Bayesian mortalitÃ© forecasting framework for population and portfolio mortality*. PhD thesis, University of Amsterdam, 2012.

[81] LI JSH ET HARDY MR ET TAN KS. Uncertainty in mortality forecasting : An extension to the classical lee-carter approach. *ASTIN Bulletin*, 39, 2009.

[82] BOOTH H ET MAINDONALD J ET SMITH L. Applying lee-carter under conditions of variable mortality decline. *Population Studies*, 3, 2002.

[83] BOOTH H ET TICKLE L. Mortality modelling and forecasting : A review of methods. *Annals of Actuarial Science*, 3, 2008.

[84] GOODMAN LA. Simple models for the analysis of association in cross-classifications having ordered categories. *Biometrika*, 74, 1979.

[85] LEE RD ET CARTER LR. Modeling et forecasting us mortality. *Journal of the American Statistical Association*, 87, 1992.

[86] BEN DBABIS M. Modélisation stochastique de la mortalité d'expérience et effets de sélection adverse. *Bulletin FranÃ$ais d'Actuariat*, 10, 2010.

[87] BROUHNS N ET DENUIT M. Risque de longévité et rentes viagères 1 evolution de la mortalité en belgique de 1880 à nos jours. *Belgian Actuarial Bulletin*, 2, 2002.

[88] BROUHNS N ET DENUIT M. Risque de longévité et rentes viagères 2 tables de mortalité prospectives pour la population belge. *Belgian Actuarial Bulletin*, 2, 2002.

[89] BROUHNS N ET DENUIT M. Risque de longévité et rentes viagères 3 elaboration de tables de mortalité prospectives pour la population assurée belge et évaluation du cout. *Belgian Actuarial Bulletin*, 2, 2002.

[90] CAIRNS AJG ET BLAKE D ET DOWD K ET COUGHLAN GD ET EPSTEIN D ET KHALAF-ALLAH M. Mortality density forecasts : An analysis of six stochastic mortality models. *Insurance : Mathematics and Economics*, 48, 2011.

[91] CZADO C ET DELWARDE A ET DENUIT M. Bayesian poisson log-bilinear mortality projections. *Insurance : Mathematics and Economics*, 36, 2005.

[92] DAHL M. Stochastic mortality in life insurance : market reserves and mortality-linked insurance contracts. *Insurance : Mathematics and Economics*, 35, 2004.

[93] DOWD K ET CAIRNS AJG ET BLAKE D ET COUGHLAN GD ET EPSTEIN D ET KHALAF-ALLAH M. Backtesting stochastic mortality models : An ex-post evaluation of multi-period-ahead density forecasts. *North American Actuarial Journal*, 14, 2010.

[94] DOWD K ET CAIRNS AJG ET BLAKE D ET COUGHLAN GD ET EPSTEIN D ET KHALAF-ALLAH M. Evaluating the goodness of fit of stochastic mortality models. *Insurance : Mathematics and Economics*, 47, 2010.

[95] KAAS R ET GOOVAERTS M ET DHAENE J ET DENUIT M. *Modern actuarial risk theory*. Springer, 2008.

[96] MEYRICKE R ET SHERRIS M. The determinants of mortality heterogeneity and implications for pricing annuities. *Insurance : Mathematics and Economics*, 53, 2013.

[97] SHU S ET SHERRIS M. Heterogeneity of australian population mortality and implications for a viable life annuity market. *Insurance : Mathematics and Economics*, 51, 2012.

[98] BORGER M ET FLEISCHER D ET KUKSIN N. Modeling the mortality trend under modern solvency regimes. *ASTIN Bulletin*, 44, 2014.

[99] COUGHLAN G ET EPSTEIN D ET ONG A ET HEVIA-PORTOCARRERO J ET GINGRICH E ET KHALAF-ALLAH M ET JOSEPH P. Lifemetrics : A toolkit for measuring and managing longevity and mortality risks. *Technical document*.

[100] DELWARDE A ET DENUIT M ET EILERS P. Smoothing the lee-carter and poisson log-bilinear models for mortality forecasting : a penalised log-likelihood approach. *Statistical Modelling*, 7, 2007.

[101] HOUGAAR P. Life table methods for heterogeneous populations distributions describing the heterogeneity. *Biometrika*, 71, 1984.

[102] DURBAN M ET CURRIE ID ET EILERS PHC. Smoothing and forecasting mortality rates. *Statistical Modelling*, 4, 2004.

[103] DENG Y ET BROCKETT P ET MACMINN R. Journal of risk et insurance. *Applied Stochastic Models in Business*, 79, 2012.

[104] PLAT R. Stochastic portfolio specific mortality and the quantification of mortality basis risk. *Insurance : Mathematics and Economics*, 45, 2009.

[105] LEE RD. The lee-carter method of forecasting mortality with various extensions and applications (with discussion). *North American Actuarial Journal*, 4, 2000.

[106] BEARD RE. Some aspects of theories of mortality models. *Biological aspects of demography*, 1971.

[107] BUTT Z ET HABERMAN S. Application of frailty-based mortality models using generalized linear models. *ASTIN Bullerin*, 34, 2004.

[108] RENSHAW AE ET HABERMAN S. Lee-carter mortality forecasting : a parallel generalised linear modelling approach for england and wales mortality. *Applied Statistics*, 52, 2003.

[109] RENSHAW AE ET HABERMAN S. Lee-carter mortality forecasting with age-specific enhancement. *Insurance : Mathematics and Economics*, 33, 2003.

[110] RENSHAW AE ET HABERMAN S. On the forecasting of mortality reduction factors. *Insurance : Mathematics and Economics*, 32, 2003.

[111] RENSHAW AE ET HABERMAN S. A cohort-based extension to the lee-carter model for mortality reduction factors. *Insurance : Mathematics and Economics*, 38, 2006.

[112] RENSHAW AE ET HABERMAN S. On simulation-based approaches to risk measurement in mortality with specific reference to binomial lee-carter modelling. *Présentation au "Living to 100" Survival to Advanced Ages international symposium Society of Actuaries Orlando Florida*, 2008.

[113] RENSHAW AE ET HABERMAN S. On simulation-based approaches to risk measurement in mortality with specific reference to poisson lee-carter modelling. *Insurance : Mathematics and Economics*, 42, 2008.

[114] LEE RD ET MILLER T. Evaluating the performance of lee-carter mortality forecasts. *Demography*, 28, 2001.

[115] KANNISTO V. Presentation at a workshop on old-age mortality. *Université d'Odense, Danemark*, 1992.

[116] PLANCHET F ET LELIEUR V. Utilisation des méthodes de lee-carter et log-poisson pour l'ajustement de tables de mortalité dans le cas de petits échantillons. *Bulletin FranÃ§ais d'Actuariat*, 8, 2007.

[117] PERKS W. On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries*, 63, 1932.

[118] OUBURG Wilbert. *Single and multi-population mortality models for Dutch data*. PhD thesis, University of Amsterdam, 2013.

[119] MAKEHAM WM. On the law of mortality and the construction of annuity tables. *The assurance magazine and Journal of the Institute of Actuaries*, 8, 1860.

[120] BELL WR. Comparing and assessing time series methods for forecasting age-specific fertility and mortality rates. *Journal of Official Statistics*, 13 :279–303, 1997.

[121] CHAI CMH ET SIU TK ET ZHOU X. A double-exponential garch model for stochastic mortality. *European Actuarial Journal*, 3, 2013.

[122] BARRIEU P ET BENSUSAN H ET EL KAROUI N ET HILLAIRET C ET LOISEL S ET RAVANELLI C ET SALHI Y. Understanding modelling and managing longevity risk : key issues and main challenges. *Scandinavian Actuarial Journal*, 2012.

[123] MELNIKOV A ET ROMANIUK Y. Evaluating the performance of gompertz makeham and lee-carter mortality models for risk management with the unit-linked contracts. *Insurance : Mathematics and Economics*, 39, 2006.

[124] LIN DY ET YING Z. Semiparametric analysis of the additive risk model. *Biometrika*, 81, 1994.

Annexe

ACF et PACF de κ_t

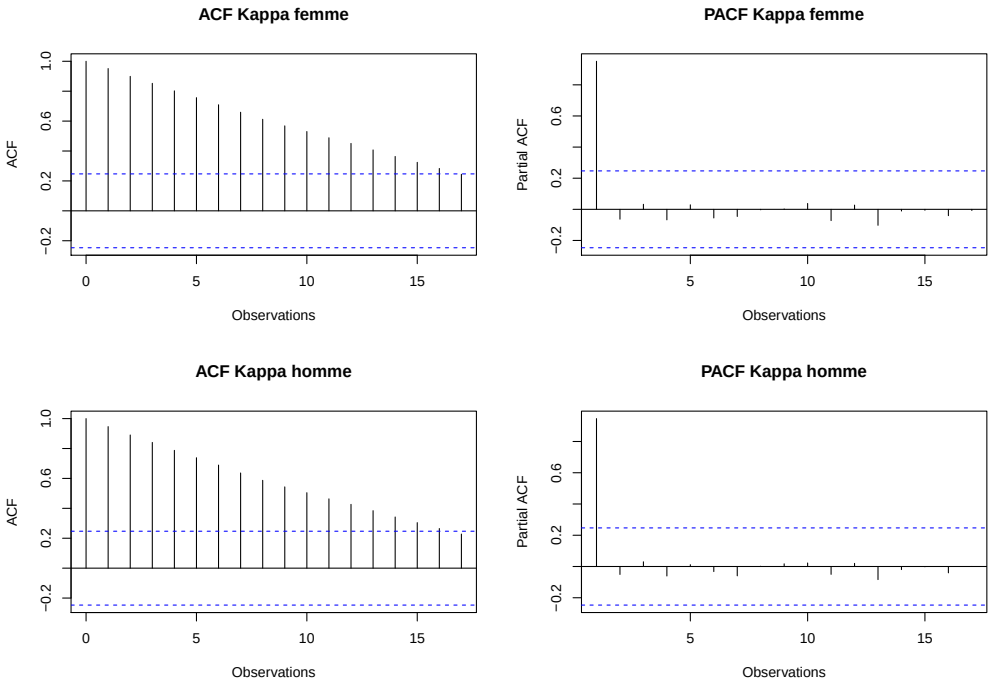

FIGURE 12: ACF et PACF de la série κ_t modèle log-Poisson

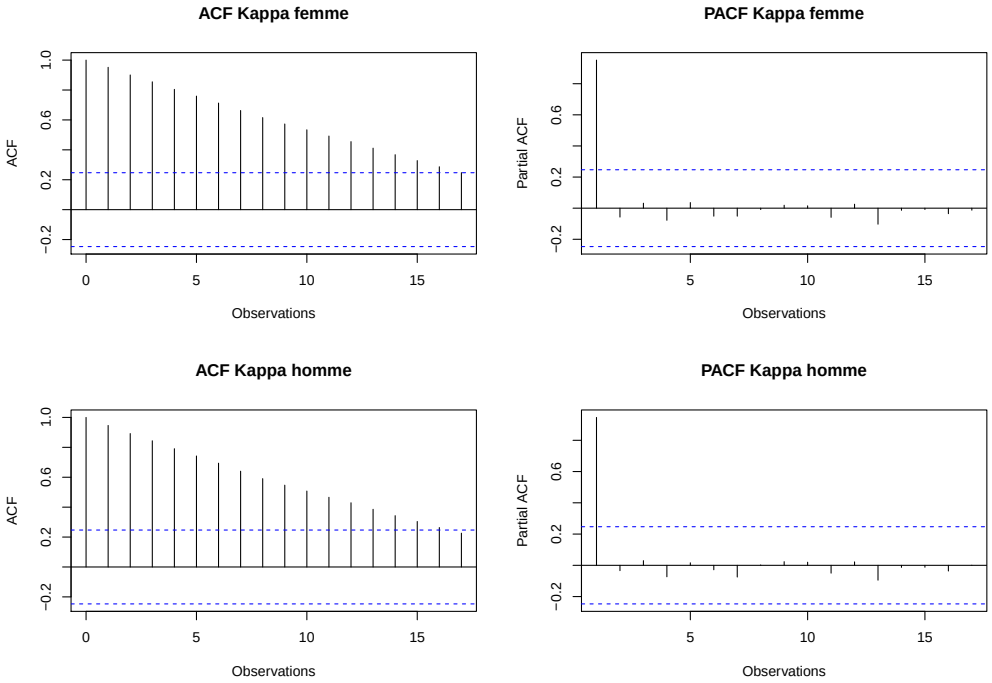

FIGURE 13: ACF et PACF de la série κ_t modèle Négative Binomiale

Qualité des modèles ARIMA(p,d,q)

ARIMA	(1,1,1)	(1,1,8)	(6,1,1)	(6,1,8)	(8,1,1)	(8,1,8)
AIC	302,84	302,75	309,53	304,09	304,75	310,6
BIC	311,35	326,15	328,67	338,12	328,15	348,89

TABLE 1: Qualité des modèles ARIMA pour le κ_t du modèle log-Poisson (femme)

ARIMA	(6,1,1)	(6,1,6)	(6,1,8)	(6,1,12)
AIC	291,06	283,53	286,94	289,34
BIC	310,21	313,31	320,97	331,88

TABLE 2: Qualité des modèles ARIMA pour le κ_t du modèle log-Poisson (homme)

ARIMA	(1,1,1)	(1,1,8)	(6,1,1)	(6,1,8)	(8,1,1)	(8,1,8)
AIC	301,27	302,04	306,56	302,32	300,03	305,92
BIC	309,78	325,44	325,7	336,35	323,43	344,21

TABLE 3: Qualité des modèles ARIMA pour le κ_t du modèle Négative Binomiale (femme)

ARIMA	(5,1,1)	(5,1,6)	(5,1,12)	(6,1,1)	(6,1,6)	(6,1,12)
AIC	313,47	308,67	311,26	307,46	305,7	313,27
BIC	330,49	336,32	351,67	326,6	353,48	355,82

TABLE 4: Qualité des modèles ARIMA pour le κ_t du modèle Négative Binomiale (homme)

Tests sur les résidus des modèles ARIMA(p,d,q)

ARIMA	(1,1,1)	(1,1,8)	(6,1,1)	(6,1,8)	(8,1,1)	(8,1,8)
(1)	0,922	0,725	0,947	0,725	0,950	0,789
(2)	0,953	0,925	0,722	0,968	0,648	0,970
(3)	0,956	0,975	0,820	0,945	0,1496	0,5372
(4)	0,929	0,980	0,6084	0,821	0,272	0,350
(5)	0,840	0,945	0,937	0,831	0,601	0,916
(6)	0,978	0,992	0,917	0,887	0,342	0,680

TABLE 5: Comparaison des p-values des tests sur les résidus des ARIMA du modèle log-Poisson (femme)

ARIMA	(6,1,1)	(6,1,6)	(6,1,8)	(6,1,12)
(1)	0,977	0,870	0,799	0,616
(2)	0,683	0,715	0,577	0,197
(3)	0,117	0,443	0,8741	0,911
(4)	0,113	0,367	0,693	0,917
(5)	0,777	0,973	0,890	0,785
(6)	0,411	0,661	0,93	0,928

TABLE 6: Comparaison des p-values des tests sur les résidus des ARIMA du modèle log-Poisson (homme)

ARIMA	(1,1,1)	(1,1,8)	(6,1,1)	(6,1,8)	(8,1,1)	(8,1,8)
(1)	0,906	0,729	0,937	0,782	0,634	0,737
(2)	0,944	0,989	0,638	0,972	0,940	0,940
(3)	0,712	0,872	0,6737	0,08	0,202	0,199
(4)	0,607	0,866	0,727	0,075	0,3185	0,3531
(5)	0,353	0,615	0,964	0,381	0,417	0,466
(6)	0,734	0,938	0,863	0,1501	0,475	0,271

TABLE 7: Comparaison des p-values des tests sur les résidus des ARIMA du modèle Négative Binomiale (femme)

ARIMA	(5,1,1)	(5,1,6)	(5,1,12)	(6,1,1)	(6,1,6)	(6,1,12)
(1)	0,951	0,941	0,900	0,875	0,840	0,641
(2)	0,857	0,732	0,898	0,741	0,865	0,511
(3)	0,042	0,570	0,157	0,109	0,199	0,079
(4)	0,026	0,851	0,392	0,120	0,216	0,093
(5)	0,576	0,725	0,380	0,656	0,542	0,441
(6)	0,106	0,372	0,308	0,293	0,250	0,190

TABLE 8: Comparaison des p-values des tests sur les résidus des ARIMA du modèle Négatif Binomiale (homme)

Quantiles des indicateurs des 2000 simulations de bootstrap

	Quantile						
Modèle	0%	5%	20%	50%	80%	95%	100%
$\mathcal{P}$	22,8	24,02	24,69	25,37	26,03	26,66	28,15
$\mathcal{NB}$	22,14	23,53	24,5	25,5	26,49	27,34	29,19

TABLE 9: Quantile de $e_{65}(2012)$ pour les femmes

	Quantile						
Modèle	0%	5%	20%	50%	80%	95%	100%
$\mathcal{P}$	20,82	21,64	22,06	22,56	23,06	23,5	24,42
$\mathcal{NB}$	17,49	20,06	21	21,93	22,89	23,75	25,3

TABLE 10: Quantile de $e_{65}(2012)$ pour les hommes

	Quantile						
Modèle	0%	5%	20%	50%	80%	95%	100%
$\mathcal{P}$	17,51	18,21	18,58	18,96	19,32	19,67	20,4
$\mathcal{NB}$	16,47	18,03	18,52	19,03	19,56	20,07	21,23

TABLE 11: Quantile de $a_{65}(2012)$ pour les femmes

	Quantile						
Modèle	0%	5%	20%	50%	80%	95%	100%
$\mathcal{P}$	16,12	16,61	16,86	17,14	17,41	17,67	18,15
$\mathcal{NB}$	14,64	15,72	16,26	16,78	17,27	17,77	18,75

TABLE 12: Quantile de $a_{65}(2012)$ pour les hommes

	Lee-Carter		$\mathcal{P}$		$\mathcal{NB}$		
Age	α_x	β_x	α_x	β_x	α_x	β_x	ϕ_x
0	-4.3832	0.0270	-4.6727	0.0247	-4.6426	0.0229	0.0026
1	-6.8936	0.0270	-7.1095	0.0276	-7.0407	0.0241	0.0289
2	-7.4426	0.0226	-7.6957	0.0210	-7.6865	0.0205	0.0021
3	-7.7171	0.0215	-7.9764	0.0184	-7.9723	0.0182	0.0015
4	-7.9171	0.0225	-8.1780	0.0172	-8.1750	0.0172	0.0013
5	-8.0324	0.0216	-8.2929	0.0165	-8.2920	0.0166	0.0035
6	-8.1307	0.0205	-8.3816	0.0151	-8.3815	0.0153	0.0037
7	-8.2282	0.0208	-8.5069	0.0146	-8.5080	0.0148	0.0037
8	-8.2674	0.0208	-8.5326	0.0127	-8.5354	0.0131	0.0080
9	-8.3196	0.0205	-8.6233	0.0125	-8.6224	0.0126	0.0026
10	-8.3377	0.0192	-8.6373	0.0129	-8.6391	0.0133	0.0148
11	-8.3051	0.0184	-8.6444	0.0127	-8.6426	0.0128	0.0022
12	-8.2140	0.0163	-8.5548	0.0122	-8.5537	0.0124	0.0043
13	-8.1157	0.0167	-8.5123	0.0117	-8.5106	0.0118	0.0049
14	-7.8057	0.0148	-8.3385	0.0112	-8.3393	0.0113	0.0056
15	-7.5548	0.0137	-8.1233	0.0112	-8.1256	0.0114	0.0061
16	-7.2932	0.0124	-7.9633	0.0100	-7.9745	0.0105	0.0251
17	-7.0415	0.0119	-7.8084	0.0103	-7.8208	0.0109	0.0234
18	-6.7669	0.0099	-7.6542	0.0087	-7.6687	0.0091	0.0289
19	-6.6715	0.0091	-7.6105	0.0094	-7.6201	0.0097	0.0165
20	-6.6476	0.0077	-7.5859	0.0095	-7.5920	0.0096	0.0088
21	-6.6226	0.0072	-7.5824	0.0100	-7.5869	0.0101	0.0075
22	-6.5863	0.0079	-7.5561	0.0108	-7.5589	0.0109	0.0116
23	-6.5742	0.0080	-7.5563	0.0113	-7.5574	0.0113	0.0079
24	-6.5729	0.0084	-7.5422	0.0113	-7.5453	0.0113	0.0112
25	-6.5701	0.0083	-7.5018	0.0114	-7.5045	0.0113	0.0132
26	-6.5629	0.0077	-7.4514	0.0113	-7.4534	0.0112	0.0130
27	-6.5533	0.0082	-7.4237	0.0112	-7.4249	0.0111	0.0123
28	-6.5295	0.0078	-7.3849	0.0118	-7.3872	0.0115	0.0149
29	-6.4977	0.0086	-7.3355	0.0119	-7.3372	0.0117	0.0116
30	-6.4621	0.0086	-7.2604	0.0113	-7.2625	0.0111	0.0131
31	-6.4317	0.0084	-7.2176	0.0115	-7.2190	0.0115	0.0130
32	-6.3892	0.0087	-7.1381	0.0107	-7.1371	0.0108	0.0060
33	-6.3343	0.0087	-7.0724	0.0105	-7.0708	0.0106	0.0084
34	-6.2683	0.0088	-7.0019	0.0101	-6.9990	0.0103	0.0065
35	-6.2006	0.0089	-6.9036	0.0101	-6.9015	0.0103	0.0054
36	-6.1280	0.0082	-6.8331	0.0103	-6.8332	0.0103	0.0040
37	-6.0534	0.0086	-6.7489	0.0100	-6.7481	0.0101	0.0028
38	-5.9759	0.0085	-6.6609	0.0095	-6.6611	0.0095	0.0022
39	-5.8984	0.0085	-6.5903	0.0089	-6.5918	0.0089	0.0027
40	-5.8040	0.0086	-6.5036	0.0092	-6.5044	0.0093	0.0014
41	-5.7183	0.0082	-6.4126	0.0089	-6.4150	0.0089	0.0016
42	-5.6179	0.0083	-6.3434	0.0087	-6.3457	0.0087	0.0014
43	-5.5315	0.0084	-6.2402	0.0085	-6.2425	0.0085	0.0012
44	-5.4326	0.0082	-6.1695	0.0082	-6.1720	0.0082	0.0013
45	-5.3437	0.0083	-6.0768	0.0082	-6.0798	0.0082	0.0014

	Lee-Carter		$\mathcal{P}$		$\mathcal{NB}$		
Age	α_x	β_x	α_x	β_x	α_x	β_x	ϕ_x
46	-5.2514	0.0082	-5.9930	0.0084	-5.9954	0.0084	0.0015
47	-5.1626	0.0081	-5.9054	0.0083	-5.9084	0.0083	0.0014
48	-5.0639	0.0082	-5.8278	0.0083	-5.8313	0.0083	0.0016
49	-4.9808	0.0080	-5.7480	0.0081	-5.7508	0.0081	0.0014
50	-4.8850	0.0083	-5.6679	0.0083	-5.6707	0.0083	0.0015
51	-4.8020	0.0081	-5.6012	0.0082	-5.6041	0.0082	0.0017
52	-4.7238	0.0083	-5.5214	0.0083	-5.5239	0.0083	0.0013
53	-4.6429	0.0083	-5.4539	0.0086	-5.4566	0.0085	0.0014
54	-4.5576	0.0086	-5.3876	0.0086	-5.3899	0.0086	0.0014
55	-4.4759	0.0085	-5.3199	0.0086	-5.3225	0.0086	0.0014
56	-4.3986	0.0088	-5.2550	0.0088	-5.2571	0.0088	0.0013
57	-4.3220	0.0090	-5.1734	0.0089	-5.1758	0.0089	0.0012
58	-4.2460	0.0092	-5.1070	0.0091	-5.1095	0.0091	0.0012
59	-4.1674	0.0093	-5.0291	0.0094	-5.0310	0.0094	0.0012
60	-4.0869	0.0094	-4.9496	0.0096	-4.9521	0.0096	0.0013
61	-4.0133	0.0097	-4.8794	0.0098	-4.8815	0.0098	0.0011
62	-3.9360	0.0099	-4.7873	0.0100	-4.7887	0.0100	0.0010
63	-3.8564	0.0100	-4.7083	0.0103	-4.7097	0.0103	0.0011
64	-3.7805	0.0101	-4.6206	0.0107	-4.6216	0.0107	0.0010
65	-3.7012	0.0102	-4.5280	0.0109	-4.5289	0.0109	0.0010
66	-3.6260	0.0105	-4.4360	0.0111	-4.4366	0.0111	0.0010
67	-3.5446	0.0105	-4.3432	0.0112	-4.3439	0.0112	0.0010
68	-3.4580	0.0106	-4.2401	0.0114	-4.2397	0.0114	0.0010
69	-3.3792	0.0107	-4.1363	0.0116	-4.1368	0.0117	0.0010
70	-3.2897	0.0108	-4.0266	0.0118	-4.0272	0.0119	0.0010
71	-3.2039	0.0108	-3.9191	0.0119	-3.9197	0.0120	0.0010
72	-3.1126	0.0107	-3.8022	0.0120	-3.8030	0.0121	0.0010
73	-3.0246	0.0108	-3.6868	0.0119	-3.6880	0.0121	0.0010
74	-2.9322	0.0108	-3.5656	0.0120	-3.5669	0.0122	0.0010
75	-2.8373	0.0107	-3.4466	0.0120	-3.4477	0.0122	0.0010
76	-2.7419	0.0106	-3.3207	0.0119	-3.3221	0.0120	0.0010
77	-2.6452	0.0105	-3.1986	0.0118	-3.2002	0.0120	0.0011
78	-2.5475	0.0103	-3.0742	0.0116	-3.0756	0.0117	0.0011
79	-2.4490	0.0102	-2.9443	0.0113	-2.9458	0.0115	0.0011
80	-2.3451	0.0098	-2.8192	0.0110	-2.8212	0.0111	0.0011
81	-2.2461	0.0095	-2.6932	0.0106	-2.6953	0.0108	0.0011
82	-2.1429	0.0093	-2.5664	0.0103	-2.5685	0.0104	0.0011
83	-2.0440	0.0089	-2.4392	0.0100	-2.4412	0.0100	0.0011
84	-1.9462	0.0086	-2.3171	0.0096	-2.3191	0.0096	0.0011
85	-1.8489	0.0083	-2.1991	0.0091	-2.2013	0.0091	0.0011
86	-1.7445	0.0080	-2.0794	0.0088	-2.0818	0.0087	0.0011
87	-1.6560	0.0076	-1.9620	0.0083	-1.9647	0.0083	0.0011
88	-1.5589	0.0072	-1.8514	0.0078	-1.8549	0.0078	0.0011
89	-1.4672	0.0068	-1.7399	0.0073	-1.7433	0.0073	0.0011
90	-1.3785	0.0062	-1.6349	0.0069	-1.6381	0.0069	0.0011

TABLE 13: Paramètres α_x, β_x et ϕ_x des différents modèles (femmes) pour la modĺisation Chapitre 4

	Lee-Carter		$\mathcal{P}$		$\mathcal{NB}$		
Age	α_x	β_x	α_x	β_x	α_x	β_x	ϕ_x
0	-4.6465	0.0223	-4.4308	0.0331	-4.3521	0.0271	0.0215
1	-7.0534	0.0233	-6.9708	0.0360	-6.8508	0.0280	0.0391
2	-7.6840	0.0198	-7.4596	0.0254	-7.4517	0.0249	0.0014
3	-7.9785	0.0181	-7.7221	0.0225	-7.7228	0.0226	0.0011
4	-8.1903	0.0179	-7.9141	0.0229	-7.9189	0.0233	0.0010
5	-8.3121	0.0175	-8.0227	0.0212	-8.0291	0.0218	0.0010
6	-8.3994	0.0160	-8.1154	0.0199	-8.1231	0.0205	0.0010
7	-8.5377	0.0161	-8.2090	0.0196	-8.2174	0.0203	0.0011
8	-8.5613	0.0141	-8.2498	0.0198	-8.2580	0.0204	0.0010
9	-8.6438	0.0136	-8.2996	0.0193	-8.3073	0.0199	0.0010
10	-8.6627	0.0140	-8.3230	0.0186	-8.3300	0.0190	0.0010
11	-8.6581	0.0135	-8.2911	0.0177	-8.2973	0.0181	0.0010
12	-8.5720	0.0130	-8.2026	0.0159	-8.2076	0.0162	0.0010
13	-8.5237	0.0119	-8.0992	0.0159	-8.1046	0.0162	0.0010
14	-8.3573	0.0118	-7.7946	0.0146	-7.7991	0.0149	0.0010
15	-8.1440	0.0119	-7.5382	0.0133	-7.5436	0.0136	0.0011
16	-7.9959	0.0108	-7.2664	0.0117	-7.2744	0.0121	0.0012
17	-7.8412	0.0112	-7.0041	0.0111	-7.0159	0.0117	0.0016
18	-7.6902	0.0093	-6.7250	0.0091	-6.7463	0.0100	0.0113
19	-7.6361	0.0099	-6.6326	0.0084	-6.6522	0.0093	0.0108
20	-7.6057	0.0099	-6.6038	0.0071	-6.6255	0.0079	0.0131
21	-7.5989	0.0102	-6.5727	0.0066	-6.5965	0.0074	0.0165
22	-7.5717	0.0111	-6.5522	0.0071	-6.5691	0.0078	0.0085
23	-7.5675	0.0114	-6.5555	0.0073	-6.5614	0.0076	0.0012
24	-7.5580	0.0115	-6.5582	0.0078	-6.5624	0.0081	0.0012
25	-7.5163	0.0114	-6.5566	0.0077	-6.5603	0.0079	0.0012
26	-7.4651	0.0113	-6.5512	0.0072	-6.5540	0.0074	0.0012
27	-7.4356	0.0111	-6.5405	0.0076	-6.5435	0.0078	0.0012
28	-7.3991	0.0115	-6.5165	0.0073	-6.5191	0.0075	0.0013
29	-7.3483	0.0118	-6.4824	0.0078	-6.4858	0.0081	0.0013
30	-7.2734	0.0112	-6.4474	0.0080	-6.4518	0.0083	0.0050
31	-7.2302	0.0115	-6.4179	0.0077	-6.4219	0.0081	0.0042
32	-7.1444	0.0109	-6.3782	0.0082	-6.3805	0.0085	0.0039
33	-7.0784	0.0107	-6.3231	0.0083	-6.3256	0.0085	0.0012
34	-7.0050	0.0104	-6.2571	0.0083	-6.2597	0.0085	0.0012
35	-6.9066	0.0103	-6.1897	0.0084	-6.1922	0.0086	0.0012
36	-6.8384	0.0103	-6.1178	0.0079	-6.1207	0.0080	0.0012
37	-6.7518	0.0099	-6.0455	0.0084	-6.0479	0.0085	0.0011
38	-6.6646	0.0094	-5.9685	0.0083	-5.9711	0.0084	0.0011
39	-6.5968	0.0088	-5.8907	0.0084	-5.8936	0.0085	0.0011
40	-6.5075	0.0091	-5.7957	0.0085	-5.7990	0.0086	0.0011
41	-6.4199	0.0087	-5.7097	0.0081	-5.7132	0.0082	0.0010
42	-6.3500	0.0085	-5.6108	0.0083	-5.6139	0.0083	0.0010
43	-6.2467	0.0083	-5.5218	0.0083	-5.5260	0.0084	0.0010
44	-6.1759	0.0080	-5.4256	0.0081	-5.4289	0.0082	0.0010
45	-6.0843	0.0080	-5.3375	0.0083	-5.3403	0.0084	0.0010

	Lee-Carter		$\mathcal{P}$		$\mathcal{NB}$		
Age	α_x	β_x	α_x	β_x	α_x	β_x	ϕ_x
46	-5.9986	0.0082	-5.2445	0.0082	-5.2477	0.0082	0.0010
47	-5.9122	0.0081	-5.1577	0.0081	-5.1598	0.0081	0.0010
48	-5.8356	0.0080	-5.0595	0.0082	-5.0614	0.0082	0.0010
49	-5.7543	0.0079	-4.9764	0.0081	-4.9783	0.0081	0.0010
50	-5.6743	0.0081	-4.8806	0.0084	-4.8824	0.0083	0.0010
51	-5.6078	0.0080	-4.7976	0.0083	-4.7994	0.0082	0.0010
52	-5.5269	0.0082	-4.7194	0.0084	-4.7212	0.0084	0.0010
53	-5.4597	0.0084	-4.6382	0.0084	-4.6403	0.0083	0.0010
54	-5.3927	0.0084	-4.5530	0.0087	-4.5550	0.0086	0.0010
55	-5.3255	0.0084	-4.4715	0.0086	-4.4735	0.0085	0.0010
56	-5.2595	0.0086	-4.3936	0.0089	-4.3960	0.0089	0.0010
57	-5.1783	0.0087	-4.3173	0.0091	-4.3198	0.0090	0.0010
58	-5.1120	0.0090	-4.2399	0.0093	-4.2433	0.0092	0.0010
59	-5.0330	0.0092	-4.1622	0.0094	-4.1653	0.0093	0.0010
60	-4.9547	0.0094	-4.0810	0.0095	-4.0847	0.0094	0.0010
61	-4.8836	0.0096	-4.0073	0.0097	-4.0113	0.0097	0.0010
62	-4.7903	0.0098	-3.9309	0.0099	-3.9344	0.0099	0.0010
63	-4.7114	0.0101	-3.8508	0.0100	-3.8548	0.0100	0.0010
64	-4.6230	0.0104	-3.7743	0.0102	-3.7788	0.0102	0.0010
65	-4.5302	0.0106	-3.6954	0.0102	-3.6997	0.0102	0.0010
66	-4.4378	0.0109	-3.6202	0.0105	-3.6245	0.0105	0.0010
67	-4.3451	0.0110	-3.5383	0.0105	-3.5429	0.0106	0.0010
68	-4.2404	0.0112	-3.4527	0.0106	-3.4568	0.0106	0.0010
69	-4.1376	0.0115	-3.3737	0.0108	-3.3778	0.0108	0.0010
70	-4.0280	0.0117	-3.2849	0.0108	-3.2886	0.0109	0.0010
71	-3.9204	0.0118	-3.1993	0.0108	-3.2028	0.0109	0.0010
72	-3.8038	0.0119	-3.1085	0.0107	-3.1116	0.0107	0.0010
73	-3.6889	0.0119	-3.0208	0.0108	-3.0237	0.0109	0.0010
74	-3.5678	0.0120	-2.9289	0.0108	-2.9313	0.0108	0.0010
75	-3.4487	0.0120	-2.8339	0.0108	-2.8364	0.0108	0.0010
76	-3.3230	0.0119	-2.7388	0.0107	-2.7410	0.0107	0.0010
77	-3.2013	0.0118	-2.6422	0.0105	-2.6443	0.0106	0.0010
78	-3.0768	0.0116	-2.5452	0.0104	-2.5467	0.0103	0.0010
79	-2.9469	0.0113	-2.4472	0.0103	-2.4483	0.0102	0.0010
80	-2.8225	0.0110	-2.3429	0.0099	-2.3444	0.0098	0.0010
81	-2.6967	0.0106	-2.2442	0.0096	-2.2455	0.0095	0.0010
82	-2.5699	0.0103	-2.1413	0.0094	-2.1424	0.0093	0.0010
83	-2.4425	0.0099	-2.0422	0.0090	-2.0434	0.0089	0.0010
84	-2.3205	0.0095	-1.9450	0.0087	-1.9458	0.0086	0.0010
85	-2.2027	0.0090	-1.8472	0.0083	-1.8482	0.0083	0.0010
86	-2.0831	0.0086	-1.7434	0.0080	-1.7440	0.0080	0.0010
87	-1.9660	0.0082	-1.6548	0.0075	-1.6553	0.0076	0.0010
88	-1.8566	0.0077	-1.5579	0.0071	-1.5585	0.0072	0.0010
89	-1.7447	0.0072	-1.4671	0.0067	-1.4671	0.0068	0.0010
90	-1.6396	0.0068	-1.3779	0.0062	-1.3785	0.0062	0.0010

TABLE 14: Paramètres α_x, β_x et ϕ_x des différents modèles (hommes)pour la modĺisation Chapitre 4

	Femmes			Hommes		
	Lee-Carter	$\mathcal{P}$	$\mathcal{NB}$	Lee-Carter	$\mathcal{P}$	$\mathcal{NB}$
Année	κ_t	κ_t	κ_t	κ_t	κ_t	κ_t
1950	46.1919	59.2286	60.5707	69.1044	43.2693	40.8892
1951	46.0424	62.3973	64.1525	69.7319	45.1181	44.0199
1952	39.7422	55.0320	55.7157	59.3641	39.7024	36.6333
1953	39.6201	57.6438	59.2122	58.8629	41.3024	40.1121
1954	34.5786	50.3097	50.7965	51.3301	36.9965	33.5619
1955	34.0290	49.5277	49.5095	47.9884	36.3645	33.5430
1956	33.5943	50.1901	50.4695	47.2092	37.1918	35.8463
1957	34.6620	46.2020	46.4795	45.4046	34.9255	34.4024
1958	25.9303	40.5625	39.3759	36.8042	28.5229	25.1627
1959	27.9736	39.5643	39.1846	38.2812	28.4053	26.4744
1960	24.7144	38.8853	38.4097	34.6466	28.3393	26.4101
1961	24.1589	33.1115	32.4014	30.4721	24.7752	23.3326
1962	26.3449	37.7811	36.9751	33.3638	28.2272	27.5683
1963	26.1847	38.5145	37.6048	33.3232	29.4708	28.7199
1964	22.7482	30.1720	29.9451	27.5281	23.7579	22.8508
1965	23.9153	31.6966	31.4301	28.1317	25.2117	25.8090
1966	21.4324	27.7252	27.6048	25.3696	22.6642	22.1868
1967	22.2891	27.9078	27.7475	26.3401	23.1963	23.4605
1968	23.2540	28.1727	27.6126	25.5160	23.5605	24.4055
1969	25.4355	29.5096	29.7713	27.6880	25.6124	27.0735
1970	21.2999	23.5253	23.2358	22.3714	19.6523	20.6388
1971	21.8742	23.3020	23.2550	23.3725	19.7814	21.5334
1972	20.8715	20.4334	20.9397	22.0597	18.2513	20.3101
1973	20.1803	20.6251	20.2851	20.5488	17.5795	19.8089
1974	18.5338	17.2294	16.8387	16.8090	16.1159	18.1658
1975	16.8859	16.7950	16.0721	14.7466	16.1929	17.9104
1976	14.9005	14.2947	13.4009	12.9874	14.8613	16.4701
1977	12.8697	8.5248	8.0734	10.1703	10.9889	12.7954
1978	12.2366	8.1438	7.2136	7.9124	10.8616	12.7708
1979	10.5457	5.1771	4.2649	5.3934	8.6483	10.7326
1980	10.2976	4.0937	3.1296	5.6710	7.8297	9.6536
1981	8.1856	4.0702	2.4212	4.5299	6.9281	8.2316

	Femmes			Hommes		
	Lee-Carter	$\mathcal{P}$	$\mathcal{NB}$	Lee-Carter	$\mathcal{P}$	$\mathcal{NB}$
Année	κ_t	κ_t	κ_t	κ_t	κ_t	κ_t
1982	5.2420	0.1217	-1.2088	1.1226	3.9748	5.4137
1983	5.8173	1.5494	-0.4104	1.7786	4.3771	5.9812
1984	2.4308	-3.7192	-5.3458	-3.2899	1.0961	2.6792
1985	0.2712	-3.8476	-6.2189	-5.0743	1.1002	1.7660
1986	-0.9324	-6.5877	-8.6496	-7.4701	-1.2076	-0.2409
1987	-4.8592	-12.4980	-13.7189	-11.9263	-5.3627	-4.8050
1988	-6.8166	-14.8239	-15.8894	-13.6250	-7.9572	-7.3242
1989	-7.5051	-16.5438	-17.6610	-14.0932	-9.2327	-8.1388
1990	-9.8846	-19.3547	-20.7707	-18.3380	-11.6294	-10.3656
1991	-9.8573	-21.7035	-22.3620	-17.7948	-13.2334	-11.5813
1992	-13.1559	-24.8611	-25.0284	-20.9190	-15.4702	-13.6818
1993	-13.5696	-24.8597	-25.1348	-21.1707	-16.0396	-14.2679
1994	-16.2983	-29.4868	-28.7894	-25.1199	-19.5513	-17.1453
1995	-18.6515	-29.2426	-28.7834	-24.7537	-20.2862	-18.0339
1996	-20.7426	-30.5759	-30.2883	-28.5883	-21.3208	-20.1002
1997	-24.4376	-32.9186	-32.7820	-31.8241	-24.3575	-24.1109
1998	-27.9395	-34.0045	-34.0287	-33.9341	-25.6117	-26.1115
1999	-29.7153	-34.7581	-35.1437	-34.6868	-27.3152	-28.0595
2000	-31.0431	-37.9955	-37.9769	-37.7559	-30.0407	-30.7473
2001	-31.4142	-40.0064	-39.2619	-38.8710	-32.0902	-32.3292
2002	-35.1096	-40.5874	-40.3158	-40.5090	-34.2431	-34.8080
2003	-36.5700	-39.1409	-39.9224	-43.4382	-34.5216	-35.7774
2004	-43.9304	-49.7466	-48.0488	-50.3446	-42.2690	-43.6868
2005	-44.9887	-49.0409	-48.3605	-51.5494	-42.5036	-44.1598
2006	-47.9458	-53.6108	-51.4230	-53.6654	-46.8678	-48.2698
2007	-51.0002	-55.6260	-53.6236	-56.1580	-48.8441	-50.5916
2008	-52.6339	-55.4450	-53.1093	-56.2236	-50.4865	-52.4526
2009	-52.3420	-56.9168	-54.3407	-56.5058	-52.4218	-53.7648
2010	-55.0933	-59.3656	-56.6671	-60.9556	-54.7168	-56.0384
2011	-58.1970	-63.5550	-59.8684	-63.4798	-58.7812	-59.8960
2012	-60.6507	-61.1971	-58.9687	-63.8691	-58.4918	-60.8353

TABLE 15: Paramètres κ_t des différents modèles (hommes) pour la modĺisation Chapitre 4

	$\mathcal{P}$		$\mathcal{NB}$		
Age	α_x	β_x	α_x	β_x	ϕ_x
0	-5.7789	-5.7789	-5.7789	0.0190	0.0070
1	-8.3971	-8.3971	-8.3971	0.0219	0.0196
2	-8.9215	-8.9215	-8.9215	0.0196	0.0009
3	-8.9014	-8.9014	-8.9014	0.0157	0.0005
4	-9.3573	-9.3573	-9.3573	0.0179	0.0002
5	-9.6822	-9.6822	-9.6822	0.0199	0.0019
6	-9.1966	-9.1966	-9.1966	0.0134	0.0001
7	-9.7556	-9.7556	-9.7556	0.0178	0.0006
8	-9.7539	-9.7539	-9.7539	0.0170	0.0013
9	-9.4825	-9.4825	-9.4825	0.0131	0.0000
10	-9.7685	-9.7685	-9.7685	0.0161	0.0004
11	-9.3794	-9.3794	-9.3794	0.0118	0.0000
12	-9.6052	-9.6052	-9.6052	0.0151	0.0003
13	-9.1776	-9.1776	-9.1776	0.0108	0.0001
14	-9.0010	-9.0010	-9.0010	0.0105	0.0002
15	-9.0235	-9.0235	-9.0235	0.0132	0.0011
16	-9.2367	-9.2367	-9.2367	0.0175	0.0269
17	-8.8880	-8.8880	-8.8880	0.0152	0.0141
18	-8.4659	-8.4659	-8.4659	0.0116	0.0107
19	-8.3269	-8.3269	-8.3269	0.0106	0.0045
20	-8.2523	-8.2523	-8.2523	0.0101	0.0025
21	-8.4398	-8.4398	-8.4398	0.0124	0.0050
22	-8.3036	-8.3036	-8.3036	0.0114	0.0034
23	-8.3478	-8.3478	-8.3478	0.0120	0.0026
24	-8.4616	-8.4616	-8.4616	0.0134	0.0063
25	-8.3111	-8.3111	-8.3111	0.0122	0.0044
26	-8.1807	-8.1807	-8.1807	0.0113	0.0040
27	-8.2718	-8.2718	-8.2718	0.0126	0.0059
28	-8.2113	-8.2113	-8.2113	0.0125	0.0060
29	-8.0276	-8.0276	-8.0276	0.0110	0.0043
30	-8.0226	-8.0226	-8.0226	0.0116	0.0056
31	-8.1060	-8.1060	-8.1060	0.0132	0.0085
32	-7.8388	-7.8388	-7.8388	0.0109	0.0031
33	-7.7882	-7.7882	-7.7882	0.0110	0.0046
34	-7.8556	-7.8556	-7.8556	0.0125	0.0087
35	-7.7005	-7.7005	-7.7005	0.0118	0.0061
36	-7.6104	-7.6104	-7.6104	0.0116	0.0043
37	-7.3921	-7.3921	-7.3921	0.0100	0.0023
38	-7.2728	-7.2728	-7.2728	0.0095	0.0018
39	-7.2432	-7.2432	-7.2432	0.0098	0.0035
40	-7.2152	-7.2152	-7.2152	0.0105	0.0029
41	-7.0359	-7.0359	-7.0359	0.0094	0.0018
42	-6.9608	-6.9608	-6.9608	0.0093	0.0016
43	-6.8363	-6.8363	-6.8363	0.0090	0.0013
44	-6.6865	-6.6865	-6.6865	0.0080	0.0011
45	-6.6108	-6.6108	-6.6108	0.0082	0.0013

	$\mathcal{P}$		$\mathcal{NB}$		
Age	α_x	β_x	α_x	β_x	$\bar{\phi}_x$
46	-6.4640	-6.4640	-6.4640	0.0075	0.0015
47	-6.3836	-6.3836	-6.3836	0.0076	0.0014
48	-6.2620	-6.2620	-6.2620	0.0071	0.0020
49	-6.2455	-6.2455	-6.2455	0.0078	0.0013
50	-6.1536	-6.1536	-6.1536	0.0077	0.0016
51	-6.0145	-6.0145	-6.0145	0.0068	0.0025
52	-5.9400	-5.9400	-5.9400	0.0069	0.0019
53	-5.8335	-5.8335	-5.8335	0.0065	0.0031
54	-5.8308	-5.8308	-5.8308	0.0073	0.0019
55	-5.7367	-5.7367	-5.7367	0.0070	0.0023
56	-5.6701	-5.6701	-5.6701	0.0070	0.0023
57	-5.5780	-5.5780	-5.5780	0.0069	0.0026
58	-5.5296	-5.5296	-5.5296	0.0072	0.0025
59	-5.5054	-5.5054	-5.5054	0.0079	0.0017
60	-5.4371	-5.4371	-5.4371	0.0081	0.0019
61	-5.3384	-5.3384	-5.3384	0.0078	0.0023
62	-5.2899	-5.2899	-5.2899	0.0083	0.0015
63	-5.2451	-5.2451	-5.2451	0.0088	0.0015
64	-5.1853	-5.1853	-5.1853	0.0092	0.0013
65	-5.0726	-5.0726	-5.0726	0.0091	0.0016
66	-5.0681	-5.0681	-5.0681	0.0101	0.0010
67	-4.9456	-4.9456	-4.9456	0.0098	0.0012
68	-4.8461	-4.8461	-4.8461	0.0099	0.0012
69	-4.8030	-4.8030	-4.8030	0.0107	0.0009
70	-4.6880	-4.6880	-4.6880	0.0107	0.0009
71	-4.6688	-4.6688	-4.6688	0.0117	0.0008
72	-4.5468	-4.5468	-4.5468	0.0117	0.0009
73	-4.4119	-4.4119	-4.4119	0.0114	0.0009
74	-4.3321	-4.3321	-4.3321	0.0119	0.0010
75	-4.2034	-4.2034	-4.2034	0.0118	0.0010
76	-4.0516	-4.0516	-4.0516	0.0115	0.0009
77	-3.9613	-3.9613	-3.9613	0.0118	0.0011
78	-3.8201	-3.8201	-3.8201	0.0116	0.0011
79	-3.6585	-3.6585	-3.6585	0.0111	0.0011
80	-3.5272	-3.5272	-3.5272	0.0110	0.0012
81	-3.3851	-3.3851	-3.3851	0.0107	0.0012
82	-3.2439	-3.2439	-3.2439	0.0104	0.0013
83	-3.0723	-3.0723	-3.0723	0.0098	0.0012
84	-2.9412	-2.9412	-2.9412	0.0096	0.0012
85	-2.7883	-2.7883	-2.7883	0.0091	0.0012
86	-2.6400	-2.6400	-2.6400	0.0087	0.0011
87	-2.4999	-2.4999	-2.4999	0.0083	0.0012
88	-2.3635	-2.3635	-2.3635	0.0078	0.0012
89	-2.1893	-2.1893	-2.1893	0.0070	0.0011
90	-2.0444	-2.0444	-2.0444	0.0064	0.0011

TABLE 16: Paramètres α_x, β_x et ϕ_x des différents modèles (femmes) pour la projection Chapitre 4

	$\mathcal{P}$		$\mathcal{NB}$		
Age	α_x	β_x	α_x	β_x	ϕ_x
0	-5.5852	0.0221	-5.5852	0.0210	0.0339
1	-8.0712	0.0223	-8.0712	0.0210	0.1277
2	-8.5127	0.0187	-8.5127	0.0181	0.0393
3	-8.9045	0.0194	-8.9045	0.0193	0.0088
4	-8.9643	0.0179	-8.9643	0.0178	0.0158
5	-9.1941	0.0189	-9.1941	0.0190	0.0024
6	-9.5757	0.0218	-9.5757	0.0221	0.0032
7	-9.3400	0.0180	-9.3400	0.0181	0.0033
8	-9.5210	0.0197	-9.5210	0.0198	0.0019
9	-9.5243	0.0190	-9.5243	0.0191	0.0020
10	-9.4699	0.0179	-9.4699	0.0180	0.0013
11	-9.4555	0.0179	-9.4555	0.0180	0.0017
12	-9.1731	0.0152	-9.1731	0.0153	0.0015
13	-9.5194	0.0205	-9.5194	0.0209	0.0203
14	-8.7021	0.0142	-8.7021	0.0142	0.0021
15	-8.6298	0.0161	-8.6298	0.0163	0.0125
16	-8.3185	0.0153	-8.3185	0.0155	0.0219
17	-7.9643	0.0141	-7.9643	0.0142	0.0272
18	-7.6221	0.0129	-7.6221	0.0131	0.0398
19	-7.4229	0.0115	-7.4229	0.0116	0.0338
20	-7.3835	0.0111	-7.3835	0.0112	0.0447
21	-7.3065	0.0104	-7.3065	0.0105	0.0508
22	-7.2663	0.0103	-7.2663	0.0104	0.0303
23	-7.3514	0.0113	-7.3514	0.0115	0.0215
24	-7.2593	0.0102	-7.2593	0.0104	0.0121
25	-7.1764	0.0092	-7.1764	0.0093	0.0086
26	-7.1509	0.0089	-7.1509	0.0090	0.0096
27	-7.1594	0.0092	-7.1594	0.0093	0.0103
28	-7.1165	0.0089	-7.1165	0.0090	0.0123
29	-7.2145	0.0106	-7.2145	0.0108	0.0186
30	-7.1010	0.0097	-7.1010	0.0098	0.0173
31	-7.0611	0.0095	-7.0611	0.0096	0.0159
32	-6.9964	0.0093	-6.9964	0.0094	0.0138
33	-7.0051	0.0102	-7.0051	0.0102	0.0116
34	-6.9368	0.0101	-6.9368	0.0102	0.0110
35	-6.9858	0.0116	-6.9858	0.0117	0.0172
36	-6.8922	0.0112	-6.8922	0.0113	0.0154
37	-6.7181	0.0101	-6.7181	0.0101	0.0060
38	-6.6287	0.0099	-6.6287	0.0099	0.0044
39	-6.4683	0.0089	-6.4683	0.0089	0.0023
40	-6.4372	0.0097	-6.4372	0.0097	0.0030
41	-6.2678	0.0086	-6.2678	0.0086	0.0021
42	-6.2026	0.0090	-6.2026	0.0090	0.0018
43	-6.1704	0.0097	-6.1704	0.0097	0.0032
44	-6.0403	0.0093	-6.0403	0.0093	0.0022
45	-5.9336	0.0091	-5.9336	0.0091	0.0018

	$\mathcal{P}$		$\mathcal{NB}$		
Age	α_x	β_x	α_x	β_x	$\bar{\phi}_x$
46	-5.8497	0.0092	-5.8497	0.0092	0.0020
47	-5.7443	0.0089	-5.7443	0.0089	0.0018
48	-5.6441	0.0089	-5.6441	0.0089	0.0015
49	-5.5209	0.0084	-5.5209	0.0084	0.0013
50	-5.4027	0.0082	-5.4027	0.0081	0.0013
51	-5.3269	0.0082	-5.3269	0.0082	0.0014
52	-5.2539	0.0083	-5.2539	0.0083	0.0013
53	-5.1213	0.0077	-5.1213	0.0077	0.0013
54	-5.0113	0.0074	-5.0113	0.0074	0.0015
55	-4.9057	0.0071	-4.9057	0.0071	0.0016
56	-4.8531	0.0075	-4.8531	0.0075	0.0016
57	-4.7717	0.0075	-4.7717	0.0074	0.0016
58	-4.7028	0.0076	-4.7028	0.0076	0.0018
59	-4.6216	0.0076	-4.6216	0.0075	0.0017
60	-4.5486	0.0077	-4.5486	0.0076	0.0017
61	-4.5071	0.0082	-4.5071	0.0081	0.0015
62	-4.4645	0.0086	-4.4645	0.0086	0.0013
63	-4.3904	0.0087	-4.3904	0.0087	0.0013
64	-4.3501	0.0092	-4.3501	0.0091	0.0012
65	-4.2666	0.0092	-4.2666	0.0091	0.0012
66	-4.2517	0.0100	-4.2517	0.0099	0.0011
67	-4.1138	0.0093	-4.1138	0.0092	0.0013
68	-4.0576	0.0096	-4.0576	0.0096	0.0011
69	-4.0047	0.0100	-4.0047	0.0100	0.0011
70	-3.9207	0.0101	-3.9207	0.0101	0.0010
71	-3.9126	0.0110	-3.9126	0.0110	0.0011
72	-3.7742	0.0104	-3.7742	0.0104	0.0010
73	-3.6664	0.0102	-3.6664	0.0102	0.0010
74	-3.6285	0.0109	-3.6285	0.0109	0.0010
75	-3.5213	0.0107	-3.5213	0.0107	0.0010
76	-3.4086	0.0105	-3.4086	0.0105	0.0010
77	-3.2923	0.0102	-3.2923	0.0102	0.0010
78	-3.1980	0.0102	-3.1980	0.0102	0.0010
79	-3.0427	0.0095	-3.0427	0.0094	0.0010
80	-2.9625	0.0097	-2.9625	0.0097	0.0010
81	-2.8346	0.0093	-2.8346	0.0093	0.0010
82	-2.7211	0.0091	-2.7211	0.0091	0.0010
83	-2.5644	0.0083	-2.5644	0.0083	0.0010
84	-2.4523	0.0080	-2.4523	0.0080	0.0010
85	-2.3165	0.0075	-2.3165	0.0075	0.0011
86	-2.2149	0.0075	-2.2149	0.0075	0.0010
87	-2.0813	0.0068	-2.0813	0.0068	0.0011
88	-1.9544	0.0063	-1.9544	0.0064	0.0012
89	-1.8529	0.0061	-1.8529	0.0062	0.0012
90	-1.7225	0.0055	-1.7225	0.0055	0.0011

TABLE 17: Paramètres α_x, β_x et ϕ_x des différents modèles (hommes) pour la projection Chapitre 5

	Femmes		Hommes	
	$\mathcal{P}$	$\mathcal{NB}$	$\mathcal{P}$	$\mathcal{NB}$
Année	κ_t	κ_t	κ_t	κ_t
1950	125.8001	127.5710	109.9878	104.2861
1951	129.4679	130.6729	112.8237	108.5726
1952	121.1053	121.3675	105.5828	100.5448
1953	124.3551	124.5678	108.4098	105.5392
1954	115.9586	115.5731	102.5454	98.1250
1955	115.2037	114.0544	102.0946	98.2180
1956	116.1665	115.0846	103.5986	101.4330
1957	111.7333	111.2470	100.8804	99.0341
1958	105.6127	103.7498	92.7773	89.3861
1959	104.6178	103.8860	92.8866	90.8096
1960	104.0274	102.8342	93.1321	91.6985
1961	97.7136	96.7367	88.8947	87.4565
1962	102.9489	101.4615	93.3254	92.9519
1963	103.7213	102.0453	94.9209	94.4013
1964	94.6484	94.2345	88.0777	87.4781
1965	96.4432	95.8831	90.2467	91.0788
1966	92.1154	91.8873	87.0070	86.7002
1967	92.4276	92.3979	87.9033	88.6365
1968	92.7147	92.0707	88.4742	89.7078
1969	94.1638	94.4974	91.1413	92.5375
1970	87.8529	87.8541	84.0478	85.2472
1971	87.6791	88.1500	84.4360	86.4222
1972	84.6845	85.9585	82.8989	84.7568
1973	84.9752	85.4092	82.1849	84.4639
1974	81.4158	81.6824	80.4874	82.7239
1975	80.9834	80.6504	80.7299	83.2922
1976	78.4539	78.1298	79.3557	81.5157
1977	72.3971	73.0056	74.9757	77.2087
1978	72.0651	71.8729	74.9206	77.3855
1979	68.9700	68.8162	72.5621	74.9552
1980	67.8753	67.9689	71.5607	73.7890
1981	67.9124	67.3006	70.5237	72.7467

	Femmes		Hommes	
	$\mathcal{P}$	$\mathcal{NB}$	$\mathcal{P}$	$\mathcal{NB}$
Année	κ_t	κ_t	κ_t	κ_t
1982	63.7429	63.3651	67.2463	69.2598
1983	65.3290	64.3940	67.8135	69.9940
1984	59.8323	59.1914	64.1960	66.0722
1985	59.7125	58.1263	64.0805	65.6992
1986	56.8575	55.5243	61.5776	63.3122
1987	50.6318	50.0727	56.9080	58.2512
1988	48.1756	47.7384	54.0860	54.9940
1989	46.4285	46.1165	52.8197	53.9048
1990	43.5000	42.7736	50.2929	51.1185
1991	41.0305	41.2746	48.6665	49.6510
1992	37.7968	38.5453	46.3749	47.3018
1993	37.8392	38.4875	45.8914	46.8238
1994	32.9847	34.5458	42.3089	43.8095
1995	33.3388	34.8762	41.7007	43.3204
1996	31.9151	33.0010	40.4298	41.8140
1997	29.4501	30.0061	37.1044	38.3244
1998	28.3448	28.7321	35.7951	36.6770
1999	27.6352	27.7251	33.9828	34.8693
2000	24.2135	24.4260	31.0374	32.1054
2001	22.1046	23.0127	28.9394	30.1679
2002	21.5736	21.8285	26.6834	27.6643
2003	23.0735	21.9590	26.2837	26.9868
2004	11.9468	12.9058	18.0421	18.6627
2005	12.8274	12.6879	17.7265	18.2806
2006	8.0850	9.1249	12.9510	13.7052
2007	6.0124	6.6177	10.7220	11.1188
2008	6.2080	6.9535	8.9402	9.3120
2009	4.6776	5.6000	6.7553	7.3703
2010	2.0019	2.7790	4.2632	5.0213
2011	-2.4021	-0.7252	-0.0491	0.8941
2012	0.0000	0.0000	0.0000	0.0000

TABLE 18: Paramètres κ_t des différents modèles (hommes) pour la projection du Chapitre 5