



**Mémoire de fin d'études**

---

**Construction d'une table de mortalité  
prospective à partir d'un échantillon réduit**

—  
*Application à la population  
d'un régime de retraite marocain*

---

**Candy N. WENDJA**

Master 2 d'Actuariat  
Promotion 2007

# Remerciements

Je remercie tout particulièrement M. Frédéric PLANCHET, M. Patrice PLOUVIER et M. Olivier de MONPLANET, pour leur dévouement et leur soutien pédagogique tout au long de la rédaction de ce mémoire.

Je remercie également ma famille pour son soutien et ses encouragements permanents.

## Résumé

L'objectif de ce travail est d'étudier la mise en place de tables de mortalité prospectives à partir de faibles effectifs soumis au risque. Les modèles présentés sont les modèles de Lee- Carter et log-Poisson respectivement. La faible taille des effectifs étudiés, particulièrement observée aux âges avancés, implique l'utilisation d'une méthode d'extrapolation des taux de mortalité aux âges élevés.

Les méthodes de Lee-Carter et log-Poisson constituent des modèles bi-dimensionnels prenant en compte à la fois l'année et l'âge pour déterminer les taux de mortalité. Les méthodes proposées sont appliquées à un portefeuille réel. Les tables prospectives ainsi construites permettent de projeter l'évolution des taux dans le futur, par extrapolation de la composante temporelle.

On détermine la meilleure méthode par la proximité des taux lissés par rapport aux taux bruts et surtout par la prudence de ces modèles pour le calcul de rentes viagères.

MOTS-CLEFS : Tables prospectives, extrapolation, lissage, rentes viagères.

## **Abstract**

The aim of this dissertation is to study the construction of prospective mortality tables from a small sample of people. The presented models are the Lee-Carter and log-Poisson methods respectively. The low size of the sample, particularly for elderly people, implies the use of an extrapolation method for the mortality rates.

The Lee-Carter and log-Poisson methods constitute two-dimensional models, taking the year and the age into account to calculate the mortality rates. The methods suggested are applied to a real data set. The prospective tables, built in this way, allow to project the rates evolution in the future, extrapolating the temporal constituent.

The best method of determination is through the nearness of the smoothed rates in comparison with the raw rates and essentially through the caution of these models for the life annuities calculation.

**KEYWORDS:** Prospective tables, extrapolation, adjustment, life annuities.

# Table des matières

<b>Introduction .....</b>	<b>7</b>
<b>Les tables de mortalité prospectives .....</b>	<b>9</b>
1.1 Nécessité des tables prospectives.....	9
1.1.1 Les tables transversales ou tables du moment .....	9
1.1.2 Les tables longitudinales ou tables prospectives .....	10
1.2 La construction de tables prospectives.....	10
1.2.1 Quotients et taux de mortalité .....	10
1.2.1.1 Les quotients de mortalité .....	11
1.2.1.2 Les taux de mortalité.....	11
1.2.2 Répartition des décès dans l'année .....	12
1.2.3 Les modèles d'ajustement .....	12
1.2.3.1 Les modèles à référence interne ou endogène .....	12
1.2.3.2 Les modèles à référence externe ou exogène .....	13
1.2.4 La fermeture des tables.....	13
<b>Présentation, traitement et description des données .....</b>	<b>15</b>
2.1 Présentation des données.....	15
2.2 Traitement des données .....	16
2.3 Description des données .....	16
<b>Les taux bruts de mortalité.....</b>	<b>19</b>
3.1 Calcul des taux bruts .....	19
3.2 Espérance de vie.....	21
<b>Ajustement des données d'expérience .....</b>	<b>23</b>
4.1 Ajustement par le modèle de Lee-Carter et extrapolation des taux de mortalité par la méthode de Coale et Kisker.....	23
4.1.1 Le modèle de Lee Carter.....	23
4.1.2 Estimation des paramètres du modèle .....	24
4.1.3 La méthode d'extrapolation de Coale et Kisker .....	26
4.1.4 Application du modèle de Lee-Carter .....	26
4.1.4.1 Estimation des paramètres .....	26
4.1.4.2 Les taux de mortalité lissés par le modèle de Lee-Carter.....	28
4.1.5 Intervalles de confiance asymptotiques des taux bruts et taux lissés par le modèle de Lee-Carter.....	30
4.1.6 Synthèse et limites de la méthode de Lee-Carter .....	33
4.2 Ajustement par le modèle log-Poisson.....	35
4.2.1 Présentation du modèle .....	35
4.2.2 Résolution du maximum de vraisemblance .....	35
4.2.3 Application du modèle log-Poisson.....	36
4.2.3.1 Estimation des paramètres et comparaison avec les résultats du modèle Lee-Carter .....	36
4.2.3.2 Les taux de mortalité lissés par le modèle log-Poisson .....	38

4.2.4 Application du modèle log-Poisson dans le cas où les paramètres $\alpha$ , $\beta$ et $k$ ont une forme polynomiale .....	40
4.2.4.1 Approximation polynomiale des paramètres .....	40
4.2.4.2 Les taux de mortalité lissés .....	42
4.2.4.3 Intervalle de confiance asymptotique des taux bruts et taux lissés par le modèle log-Poisson .....	45
4.3 Autre méthode d'extrapolation des taux de mortalité aux grands âges : l'extrapolation des polynômes des coefficients $\alpha$ , $\beta$ et $k$ .....	47
4.3.1 Présentation .....	47
4.3.2 Les résultats obtenus .....	47
<b>Détermination du meilleur modèle de lissage.....</b>	<b>50</b>
5.1 Le test du Chi-Deux d'ajustement .....	50
5.1.1 Statistiques du Chi-Deux avec le modèle log-Poisson .....	50
5.1.2 Statistiques du Chi-Deux avec le modèle de Lee-Carter .....	53
5.2 Comparaison des capitaux constitutifs de rente .....	54
<b>Estimation de la mortalité future .....</b>	<b>56</b>
6.1. Théorie .....	56
6.2 Prévision des taux de mortalité futurs par le modèle logPoisson .....	57
6.3. Evolution de l'espérance de vie à 60 ans .....	58
<b>Conclusion.....</b>	<b>60</b>
<b>Annexes.....</b>	<b>62</b>
<b>Bibliographie .....</b>	<b>67</b>
<b>Table des figures.....</b>	<b>69</b>

# Introduction

En France, la liquidation de capitaux sous forme de rentes n'a qu'un faible attrait pour les ménages. Pourtant, les rentes viagères constituent un mode de financement intéressant de la période de retraite, car elles garantissent contre le risque de longévité.

Pour la tarification de ces rentes viagères, des tables de mortalité réglementaires, les tables TV 88-90, établies à partir des données de l'INSEE, sont applicables aux contrats d'assurance-vie souscrits depuis le 1<sup>er</sup> juillet 1993.

La nécessité d'utiliser des tables de mortalité prospectives pour les rentes viagères a été prise en compte par le législateur et, des tables de génération ont été homologuées. Ces tables, obtenues sur la base de la mortalité de la population féminine sur la période 1961-1987, servent depuis le 1<sup>er</sup> juillet 1993 à la tarification et au provisionnement de contrats de rentes viagères (immédiates ou différées).

Cependant, la construction de tables de mortalité d'expérience, relative à la population assurée, s'avère fréquemment indispensable pour tarifer le montant des rentes viagères. Ces tables permettent de cerner tout comportement de la population assurée qui serait significativement différent des tables réglementaires. Le danger issu de l'utilisation de tables officielles est le risque de sous provisionnement des rentes viagères.

Pour l'estimation du coût des retraites ou des rentes viagères dans le futur, deux méthodes principales de construction de tables de mortalité d'expérience sont possibles. Les deux méthodes utilisent les données des années antérieures sur la mortalité du groupe pour estimer la mortalité future.

L'approche usuelle consiste à ajuster les taux bruts par âge à un modèle paramétrique classique de type Makeham. Les tables construites sont unidimensionnelles (une par année), seule la composante des âges fait l'objet du lissage. Dans ce cas, on essaie de mettre en évidence l'évolution de la mortalité sur l'ensemble des années de l'étude puis on détermine les taux de mortalité correspondant aux années futures par prolongement de cette tendance.

Parallèlement, des approches récentes déterminant des tables de mortalité prospectives proposent des modèles bi-dimensionnels. Il s'agit du modèle de Lee Carter, de ses variantes de type log-Poisson. Ces méthodes prennent en compte les deux composantes - l'âge et l'année - pour l'estimation des taux de mortalité sur l'ensemble des années considérées. Les taux de mortalité pour les années futures se déduisent alors directement de la composante paramétrique temporelle du modèle prospectif retenu. Ces méthodes ont été initialement développées sur de très grands échantillons de population, à l'échelle d'un pays.

L'objectif du présent travail est de créer des tables de mortalité prospectives, utilisant les modèles de Lee-Carter et log-poisson, dans le cas de portefeuilles de taille réduite. Le faible effectif de la population soumise au risque implique une forte volatilité des taux de mortalité. On met en évidence les problèmes sur les coefficients des paramétrisations par Lee-Carter et log-Poisson, impliqués par ce faible effectif soumis au risque. De plus, nous proposons un moyen de résoudre cette difficulté.

En outre, une autre conséquence du faible effectif soumis au risque de décès se manifeste aux âges les plus avancés. En effet, les effectifs décroissant avec l'âge, les taux de mortalité aux âges les plus élevés sont très volatiles et aucune analyse ne peut être menée sur les données brutes au-delà de l'âge de 90 ans.

En conséquence, on est amené à trouver des méthodes d'extrapolation des taux de mortalité aux âges les plus avancés. La méthode d'extrapolation de Coale-Kisker est la principale méthode d'extrapolation utilisée dans cette étude.

La détermination du meilleur modèle se basera sur le test statistique du Chi-Deux, la proximité des taux avec ceux issus des données brutes et la prudence du coût du provisionnement des rentes viagères qu'il implique. On comparera également le coût des rentes viagères versées aux retraités, avec les tables brutes et les différentes tables lissées.



# Chapitre 1

## Les tables de mortalité prospectives

### 1.1 Nécessité des tables prospectives

La table de mortalité est un modèle qui permet de rendre compte de la mortalité vécue par une population depuis la naissance jusqu'à l'extinction de l'ensemble des individus. Cette population est supposée fermée à toute migration.

Du point de vue de l'assureur, on peut distinguer deux types de tables : les tables réglementaires, imposées par l'autorité de tutelle et les tables d'expérience, dressées par les compagnies d'assurance sur la base de la sinistralité observée au sein de leur propre portefeuille d'assurance-vie. Elles présentent souvent l'inconvénient que le groupe de référence est restreint, et que, dès lors, les fluctuations observées quant à la mortalité peuvent être importantes.

D'un point de vue technique, on peut distinguer deux types de tables :

#### 1.1.1 Les tables transversales ou tables du moment

L'approche transversale ou statique est sans nul la plus connue. Dans ce cadre, les individus sont observés pendant une période variant de un à cinq ans en général, et fournissent les estimations des quotients de mortalité qui seront utilisés pour générer le tarif d'expérience. Si la période d'observation est l'an 2000, les individus de  $x$  années révolues en 2000 serviront à estimer la probabilité  $q_x$  de décéder à cet âge. La table de mortalité ainsi obtenue mélangera donc des individus nés au cours de différentes années, et ne décrira la mortalité d'aucun d'eux. Baser une tarification sur une table transversale revient donc implicitement

à considérer que la mortalité est figée (c'est à dire que les taux de mortalité n'évoluent pas dans le temps) ce qui est clairement démenti par les observations. Si cette manière de faire est synonyme de sécurité pour les opérations en cas de décès (dans une optique d'allongement de la durée de vie humaine), elle constitue un réel danger pour les opérations en cas de vie, et tout particulièrement pour les rentes viagères.

### **1.1.2 Les tables longitudinales ou tables prospectives**

Contrairement à l'approche transversale, dans l'approche longitudinale ou dynamique, on anticipe l'évolution probable de la mortalité des individus ; l'allongement de la durée de vie est pris en compte. On tente ici de suivre les différentes générations, ce qui oblige bien entendu à prédire la mortalité future des cohortes non éteintes. Pour les produits où le risque viager est fort important, comme les rentes, il s'agit de la seule approche fournissant une tarification réaliste (dans le cas de figure où la mortalité évolue de façon significative).

## **1.2 La construction de tables prospectives**

L'objectif de tables prospectives est de tenir compte des évolutions à venir de la mortalité ; les méthodes usuelles cherchent tout d'abord à ajuster les tendances passées, puis à les extrapoler à l'avenir.

La démarche de construction d'une table prospective comporte systématiquement deux étapes : tout d'abord l'estimation de taux bruts, par âge, ou par âge et génération, puis ensuite l'ajustement de ces taux bruts à un modèle, permettant d'une part de lisser les fluctuations d'échantillonnage et d'autre part de projeter l'évolution des taux dans le futur, par extrapolation.

### **1.2.1 Quotients et taux de mortalité**

La première étape de construction d'une table de mortalité consiste en l'estimation à partir des données disponibles sur la population soumise au risque et du nombre de décès par âge, des quotients de mortalité et/ou des taux de mortalité. Cette population sera supposée homogène, c'est-à-dire présentant des caractéristiques de risques identiques à l'exception de l'âge.

### 1.2.1.1 Les quotients de mortalité

Soit la fonction de survie  $x \rightarrow l_x$  décrivant le nombre moyen de survivants d'une cohorte de  $l_0$  individus tous nés à une même date ; théoriquement, partant d'un nombre  $l_0$  de nouveaux nés,  $l_x$  est le nombre moyen de survivants à l'âge  $x$ , i.e.

$$l_x = l_0 p_x$$

${}_t p_x$  étant la probabilité pour qu'un individu vivant à l'âge  $x$  survive jusqu'à l'âge  $x+t$ .

La différence  $d_x = l_x - l_{x+1}$  représente le nombre moyen de décès observés parmi les  $l_x$  individus âgés de  $x$  années. On a alors :

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

${}_t q_x$  étant la probabilité qu'un individu ayant atteint l'âge  $x$  décède avant l'âge  $x+t$ .

Si on note  ${}_t d_x = l_x - l_{x+t}$ , le nombre d'individus décédant entre  $x$  et  $x+t$ , on a alors :

$${}_t q_x = \frac{{}_t d_x}{{}_t l_x}$$

${}_t q_x$  est appelé quotient de mortalité (car il apparaît comme le rapport entre le nombre de décès sur l'intervalle  $(x, x+t)$  et le nombre d'individus vivants en début de période.

### 1.2.1.2 Les taux de mortalité

Le taux de mortalité sur  $(x, x+t)$  diffère du quotient en ce sens que le nombre de décès est ramené à l'effectif moyen, au lieu de l'effectif initial. Le taux de mortalité sur  $(x, x+t)$ , noté  ${}_t m_x$ , est donné par :

$${}_t m_x = \frac{{}_t d_x}{({}_t l_x + {}_{t+1} l_x) / 2}$$

Les quotients de mortalité sont des probabilités (nombres sans dimension) alors que les taux de décès sont exprimés en inverse de l'unité de temps et comptabilisent des décès par personne sous risque et par unité de temps.

## 1.2.2 Répartition des décès dans l'année

Les données disponibles sont souvent des données regroupées dans lesquelles l'unité de temps est l'année. Il convient alors de se donner une règle de répartition des décès dans l'année ; deux hypothèses sont classiquement proposées :

- la constance des taux instantanés de décès entre 2 âges non entiers,
- la répartition uniforme des décès au cours de l'année ;

Le choix entre ces 2 hypothèses n'est pas neutre sur l'appréciation que l'on aura du niveau de la mortalité. En effet, l'hypothèse de constance des taux instantanés conduit à des durées de vie inférieures : de ce fait, il s'agit d'une hypothèse prudente dans le cas de garanties en cas de décès, moins prudente pour des contrats de rentes. Toutefois, l'écart entre les deux approches est faible. S'agissant ici d'un régime de retraite, et donc de rentes viagères, nous retiendrons dans le cadre de cette étude l'hypothèse de répartition des décès dans l'année.

## 1.2.3 Les modèles d'ajustement

Comme nous l'avons souligné précédemment, après avoir calculé les taux bruts, il convient de les ajuster.

Deux approches utilisant les données des années antérieures sur la mortalité du groupe pour construire un ajustement des taux de décès bruts passés, puis des projections de la mortalité future, sont envisageables :

### 1.2.3.1 Les modèles à référence interne ou endogène

La démarche consiste ici à exploiter l'information contenue dans les taux bruts de mortalité pour ajuster un modèle représentant correctement les données et permettant une projection «réaliste». Les effectifs relativement restreints à disposition peuvent toutefois nuire à l'identification de la dérive de mortalité dans une proportion délicate à quantifier.

Les modèles à référence interne récents reposent sur des modèles bi-dimensionnels, qui « capturent » conjointement les deux composantes - l'âge et l'année - pour l'ajustement des taux de mortalité sur l'ensemble des années considérées. Il s'agit notamment du modèle de Lee-Carter et de ses variantes de type log-Poisson, que nous présenterons par la suite.

Les taux de mortalité pour les années futures se déduisent dans un second temps directement de la composante temporelle du modèle prospectif retenu. Ces méthodes ont été initialement développées sur de très grands échantillons de population, à l'échelle d'un pays.

Nous présenterons les ajustements qui doivent leur être apportés dans le cas de population de taille plus réduite.

### **1.2.3.2 Les modèles à référence externe ou exogène**

Une solution pour pallier aux difficultés associées à des échantillons de taille réduite est de positionner la mortalité du groupe considéré par rapport à une référence externe ; on pourrait retenir par exemple un ajustement logistique, ce modèle ayant été proposé par Daniel Serant. Disposant d'un ensemble de tables de moments historiques et prospectives, l'idée est alors de « positionner » les tables du moment d'expérience par rapport à cet ensemble de tables. L'extrapolation est ensuite immédiate en utilisant les coefficients de la régression sur les tables prospectives exogènes.

Notre attention se portera essentiellement sur la première classe de modèle.

### **1.2.4 La fermeture des tables**

Après avoir ajusté les taux bruts, il convient de « fermer » la table ; en effet, bien souvent, les données font défaut au-delà de 90 ans, elles ne sont pas assez nombreuses ou sont de piètre qualité. Il devient donc nécessaire de développer des techniques, dites de fermeture de tables, qui permettent d'extrapoler la mortalité aux grands âges et ainsi d'obtenir une meilleure estimation de la mortalité aux âges élevés.

Le besoin de fermeture des tables de mortalité soulève le problème de l'existence d'un éventuel âge ultime, ou âge biologique maximal. D'éminents biologistes et médecins croient à une limite supérieure de la durée de vie humaine qui serait inscrite dans nos gènes. Dans son étude, Vaupel met en évidence la décroissance de la mortalité à partir de 80 ans. Les modèles de fermeture des tables sont compatibles avec cette caractéristique.

Certaines méthodes de fermeture de tables nécessitent une contrainte de fermeture comme le modèle de Denuit & Goderniaux dans lequel l'âge limite est 130 ans.

D'autres méthodes, comme celle de Coale & Kisker, que nous retiendrons ici et présenterons plus en détail, ne limitent pas la durée de vie humaine mais pose simplement une contrainte sur les taux de mortalité à un âge donné.

# Chapitre 2

## Présentation, traitement et description des données

### 2.1 Présentation des données

Les données utilisées dans cette étude proviennent d'un régime de retraite marocain.

Un décalage important entre la mortalité constatée et la mortalité escomptée avait été observé lors de surveillances menées antérieurement. La table de mortalité d'expérience prospective établie dans le cadre de cette étude pourrait donc éventuellement permettre un meilleur pilotage du régime.

La base de données initiale est une base historique des pensionnés du régime sur une période d'observation de 15 ans, de 1989 à 2003. Elle se présente sous format texte ; la conversion et le traitement de cette base ont été réalisés au moyen du logiciel ACCESS.

Cette base présente les huit informations suivantes :

- un code d'enregistrement (0 si vivant, et 1 si décédé au 31/12/2003)
- un n° d'identification
- la date de naissance (format JJMMAAAA)
- la date éventuelle de décès (format JJMMAAA)
- la nature de la pension ('D' pour Directe, 'I' pour Indirecte)
- le sexe ('M' ou 'F')
- le trimestre d'effet de la pension
- l'année d'effet de la pension (format AAAA)

Notons que la qualité de cette base (nombre de lignes > 90.000) est assez remarquable, ce que confirme le peu d'anomalies décelées lors du traitement des données.

## 2.2 Traitement des données

La première étape du traitement de données consiste en l'extraction et la conversion de celles-ci, puis de la détection des anomalies. Au moyen de requêtes simples, nous avons vérifié la complétude des différents champs, et détecter deux anomalies :

- des individus indiqués décédés mais ayant une date de décès manquante.  
Il s'est avéré que la plupart de ces personnes étaient décédées avant 1986 et ne rentraient donc pas dans le périmètre de l'étude.
- des individus ayant une date de naissance manquante,
- des individus n'ayant pas d'année d'effet de la pension.

La deuxième étape a consisté en un découpage de ladite base (de 1989 à 2003), effectué en vue de disposer de quinze périodes d'observation et de pouvoir dégager des statistiques annuelles ainsi qu'une tendance éventuelle d'augmentation de l'espérance de vie.

Nous avons effectué ce découpage pour chaque année, en utilisant la règle suivante :

$$\left\{ \begin{array}{l} \text{date de décès} > 31/12/199(X-1) \text{ (ou } 31/12/200(X-1)) \text{ ou 'vide'} \\ \text{et} \\ \text{année d'effet de la pension} < 199(X+1) \text{ ( ou } 200(X-1)) \end{array} \right.$$

Ce qui revient à isoler les effectifs des personnes pensionnées de l'année 199(X) et 200(X).

## 2.3 Description des données

Tout au long cette étude, nous ne distinguerons pas les populations masculine et féminine ; en effet, la proportion de femmes varie de 13,5% à 14,2% sur l'ensemble de la période d'observation. Sachant que l'effectif de l'ensemble de la population est déjà faible, les résultats obtenus en considérant uniquement la population féminine ne seraient en aucun cas significatifs.



De 1989 à 2003 la taille de la population soumise au risque et les décès observés évolue comme le montre le graphique suivant :

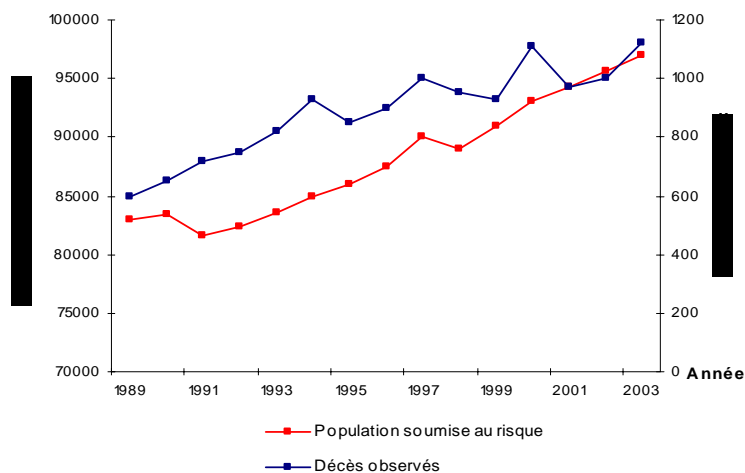


Fig. (1) : Effectifs soumis au risque et décès observés de 1989 à 2003

On peut noter que la taille de la population étudiée est relativement restreinte, puisqu'elle avoisine les 90 000 personnes (avec une variation de plus ou moins 7%) sur chaque année de l'étude. Comme cet effectif est réparti en plus de 100 classes d'âges, l'effectif étudié pour une année et un âge considérés est en moyenne inférieur à 900 personnes. Ce nombre relativement faible sera à l'origine la volatilité des taux bruts de mortalité.

La répartition moyenne par âge des effectifs soumis au risque sur la période d'observation est illustrée par le graphique ci-dessous :

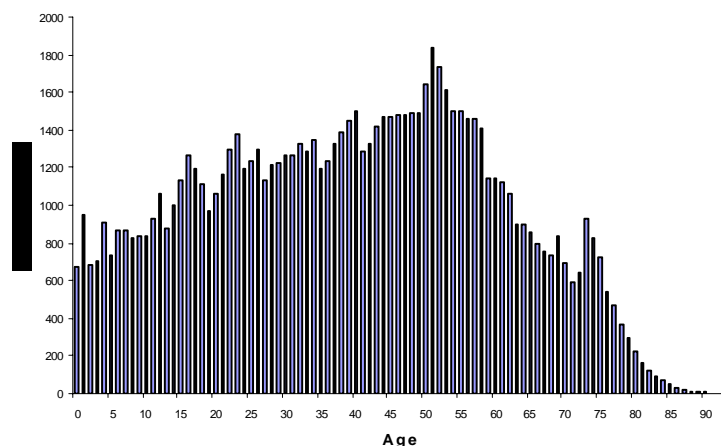


Fig. (2) : Répartition moyenne par âge des effectifs soumis au risque

On constate la faiblesse de l'effectif au-delà de 90 ans. La fermeture de la table par la méthode de Coale et Kisker nous permettra de résoudre le problème suscité par ce nombre trop restreint d'observations aux âges avancés.

# Chapitre 3

## Les taux bruts de mortalité

### 3.1 Calcul des taux bruts

S'agissant ici d'une table de mortalité d'expérience, nous avons directement calculé les taux bruts par âge, pour chacune des années calendaires (de 1989 à 2003). Nous avons calculé ces taux bruts, à l'année  $t$  et à l'âge  $x$  comme suit :

$${}_t m_x = \frac{{}_t d_x}{({}_t l_x + {}_{t+1} l_x) / 2}$$

Il s'agit du taux de mortalité, noté  ${}_t m_x$ , sur  $(x, x+t)$  ;  ${}_t d_x$  étant le nombre d'individus décédant entre  $x$  et  $x+t$  et  ${}_t l_x$  (respectivement  ${}_{t+1} l_x$ ) l'effectif d'âge  $x$  en  $t$  (respectivement en en  $t+1$ ).

Le graphe ci-dessous montre l'évolution des taux de mortalité bruts observés sur la période d'observation, de 1989 à 2003 :

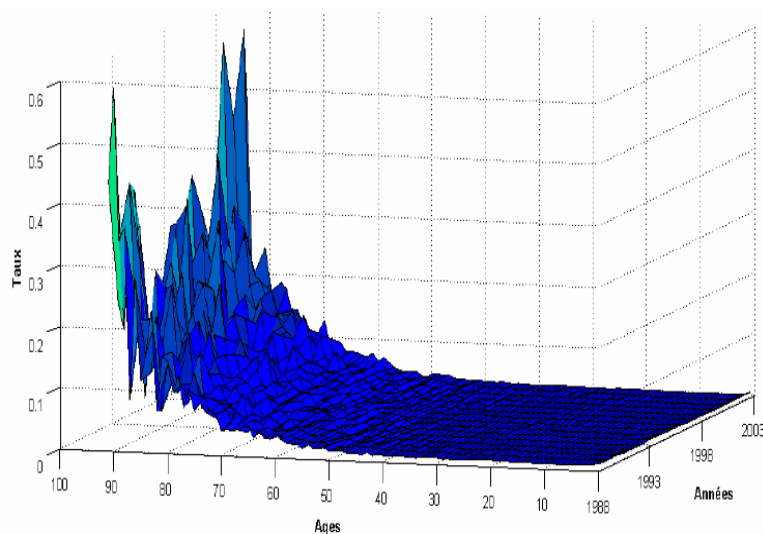


Fig. (3) : Taux bruts de mortalité de 1989 à 2003

On constate que les taux deviennent visiblement très volatiles à partir de l'âge de 60 ans. De plus, à partir de l'âge de 50 ans, on observe aussi de fortes variations des taux de mortalité entre les années. Cela justifie la nécessité de lissage des taux de mortalité.

Le graphique suivant permet de mieux apprécier ces taux de mortalité :

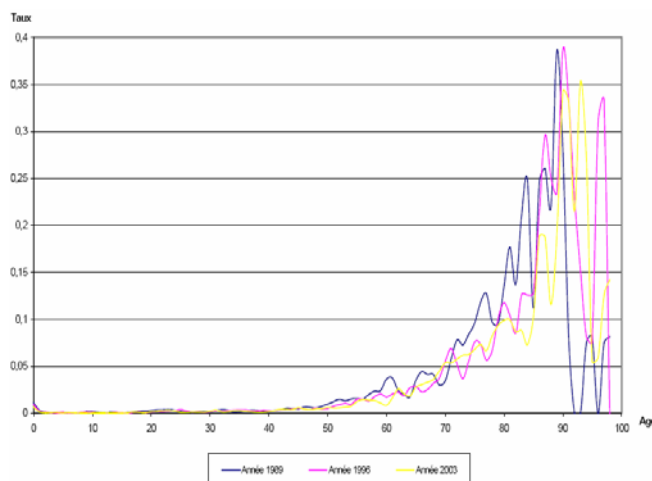


Fig. (4) : Comparaison des taux bruts de mortalité de 1989, 1996 et 2003

Nous avons présenté les taux pour trois années calendaires dans un souci de lisibilité du graphe. De plus, l'évolution observée sur ces trois années, qui correspondent respectivement à la première année d'observation (1989), à l'année médiane (1996) et à la dernière année (2003), est représentative de l'évolution générale sur l'ensemble des quinze années.

On observe que plus la date d'observation est récente, plus les taux de mortalité à âge fixé diminuent entre 50 et 90 ans. Cependant pour les âges inférieurs à 45 ans, on peut observer que les courbes sont pratiquement confondues.

A partir de 90 ans, les données apparaissent difficilement interprétables, on constate en effet des chutes de mortalité soudaines – le taux pouvant même s'annuler (au-delà de 90 ans pour l'année 1989) – ce qui ne correspond sans doute pas à la réalité. Ainsi, cette partie de la table sera lissée grâce à la méthode de fermeture des tables de Coale et Kisker, que nous présenterons dans la suite de cette étude.

## 3.2 Espérance de vie

L'espérance de vie à l'âge  $x$  est la durée moyenne restant à vivre après  $x$ , ou durée de vie moyenne au-delà de  $x$ , pour les survivants à l'âge  $x$ . Une personne en vie à l'âge  $x$  décèdera en moyenne à l'âge  $x + e_x$ .

Par définition, l'espérance de vie résiduelle à l'âge  $x$  est:

$$e_x = \frac{1}{l_x} \int_0^+ l_{x+u} du$$

$l_x$  désignant la fonction de survie à l'âge  $x$ . La version discrète de cette formule est simplement :

$$e_x = \frac{1}{l_x} \sum_{j=1}^w l_{x+j}$$

$w$  étant l'âge limite de la vie biologiquement envisageable.

Les graphiques suivants présentent l'évolution de l'espérance sur la période d'observation :

➤ *Espérance de vie à la naissance :*

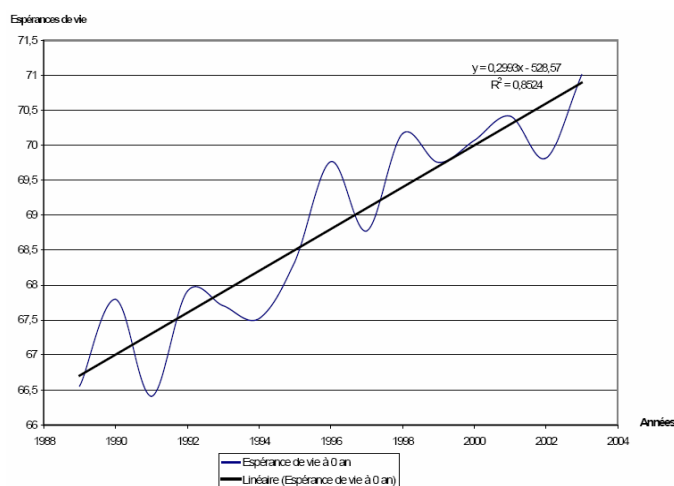


Fig. (5) : Evolution de l'espérance de vie à la naissance de 1989 à 2003

On peut constater une augmentation globale très sensible de l'espérance de vie à la naissance puisqu'en 15 ans (de 1989 à 2003), l'espérance de vie a augmenté de

4,5 ans (hausse de 6,69 %, de 65 ans et demi à 71 ans) soit une progression annuelle moyenne de près de 4 mois.

➤ *Espérance de vie à 60 ans :*

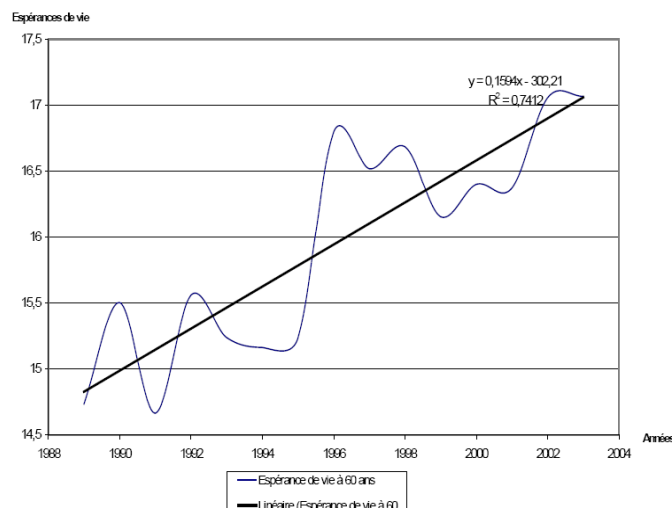


Fig. (6) : Evolution de l'espérance de vie à 60 ans de 1989 à 2003

La tendance générale est à une forte augmentation de l'espérance de vie à 60 ans, même si on constate comme à la naissance et pour certaines années des phénomènes de chute de l'espérance de vie. Globalement, entre 1989 et 2003, l'espérance de vie à 60 ans est passée de 14,73 ans à 17,07 ans, soit une progression de 2 ans et 4 mois (hausse de 15,8 %).

Cette augmentation de l'espérance de vie résiduelle à 60 ans est donc plus marquée que celle à la naissance.

Afin de réduire la variabilité des taux due aux fluctuations d'échantillonnage, il apparaît nécessaire de lisser les taux de mortalité bruts. Rappelons que l'idée d'un ajustement des taux bruts de mortalité est qu'il existe une structure régulière sous-jacente, et que les données brutes fournissent une information partielle sur cette structure. La méthode d'ajustement consiste précisément à dégager cette structure inobservable.

# Chapitre 4

## Ajustement des données d'expérience

### 4.1 Ajustement par le modèle de Lee-Carter et extrapolation des taux de mortalité par la méthode de Coale et Kisker

#### 4.1.1 Le modèle de Lee Carter

Il s'agit d'une méthode d'extrapolation des tendances passées initialement utilisée sur des données américaines, qui est devenue rapidement un standard. La modélisation retenue pour le taux instantané de mortalité est la suivante :

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

où :

- $\mu_{xt}$  représente le taux instantané de mortalité à la date  $t$  pour l'âge  $x$ ,
- $\alpha_x$  est la composante spécifique à l'âge  $x$ , elle s'interprète comme la valeur moyenne des  $\ln(\mu_{xt})$  au cours du temps,
- $k_t$  décrit l'évolution générale de la mortalité,
- $\beta_x$  traduit la sensibilité de la mortalité instantanée à l'âge  $x$  par rapport à

l'évolution générale  $k_t$  au sens où  $\frac{d \ln(\mu_{xt})}{dk_t} = \beta_x$ .

Ainsi, les âges pour lesquels les  $\beta_x$  sont importants sont les plus sensibles à l'évolution générale de la mortalité. Le modèle de Lee-Carter suppose la constance au cours du temps de cette sensibilité.

- $\varepsilon_{xt}$  est une variable aléatoire iid distribuée selon une loi  $N(0, \sigma^2)$ .

L'idée du modèle est donc d'ajuster à la série (doublement indicée par  $x$  et  $t$ ) des logarithmes des taux instantanés de décès une structure paramétrique

(déterministe) à laquelle s'ajoute un phénomène aléatoire ; le critère d'optimisation retenu va consister à maximiser la variance expliquée par le modèle, ce qui revient à minimiser la variance des erreurs.

Afin de rendre le modèle identifiable, il convient d'ajouter des contraintes sur les paramètres ; on retient en général les contraintes suivantes :

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

avec  $t_m$  désignant la première année d'observation des taux de mortalité et  $t_M$  la dernière année de l'étude.

On obtient alors les paramètres par un critère de moindres carrés :

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

Il convient alors de résoudre ce programme d'optimisation, sous les contraintes d'identifiabilité. Le nombre de paramètres à estimer est élevé, il est égal à

$$2 \times (x_M - x_m + 1) + t_M - t_m - 1.$$

#### 4.1.2 Estimation des paramètres du modèle

Les paramètres du modèle à estimer sont les  $\alpha_x$ ,  $\beta_x$  et les  $k_t$ . Leur estimation se fait en 3 étapes :

##### 1<sup>ère</sup> Etape

Après avoir estimé les taux bruts par la formule  $\hat{\mu}_{xt} = \frac{{}_t d_x}{({}_t l_x + {}_{t+1} l_x)/2}$ , on trouve en tenant compte de la contrainte d'identifiabilité sur les  $k_t$  que :

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

En effet, on a :  $\sum_{t=t_m}^{t_M} \ln \hat{\mu}_{xt} = \sum_{t=t_m}^{t_M} \alpha_x + \beta_x \sum_{t=t_m}^{t_M} k_t + \sum_{t=t_m}^{t_M} \varepsilon_{xt}$

Les contraintes  $\sum_{t=t_m}^{t_M} k_t = 0$  et  $\sum_{t=t_m}^{t_M} \varepsilon_t = 0$  impliquent alors l'estimation obtenue pour

le paramètre  $\hat{\alpha}_x$ . En d'autres termes,  $\hat{\alpha}_x$  est la moyenne temporelle, à l'âge  $x$ , des taux instantanés de décès (sur l'échelle logarithmique).



## 2<sup>ème</sup> Etape

Afin d'estimer les paramètres  $k_t$  et  $\beta_x$ , on construit la matrice  $Z = (z_{xt})$  définie par :

$$z_{xt} = \ln \hat{\mu}_{xt} - \hat{\alpha}_x.$$

On décompose ensuite en valeurs singulières la matrice  $Z$ . Si  $u_1$  (respectivement  $v_1$ ) désigne le vecteur propre de  $Z'Z$  (respectivement  $ZZ'$ ) associé à la plus grande valeur propre  $\lambda_1$ , les estimations de  $k_t$  et  $\beta_x$  sont:

$$\hat{\beta} = \frac{v_1}{\sum_j v_{1j}} \text{ et } \hat{k} = \sqrt{\lambda_1} * \sum_j v_{1j} * u_1$$

## 3<sup>ème</sup> Etape

Cette étape est facultative. On peut en effet ajuster les  $\hat{k}_t$  pour que le nombre total de décès enregistrés chaque année corresponde à celui prédit par le modèle, étant donnés les  $\hat{\alpha}_x$  et les  $\hat{\beta}_x$ .

Ainsi, on cherche les  $\hat{k}_t$  vérifiant :

$$\sum_{x=x_m}^{x_M} {}_t d_x = \sum_{x=x_m}^{x_M} {}_t l_x \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{k}_t)$$

Afin de vérifier les contraintes du modèle de Lee-Carter, les estimations retenues pour les paramètres  $\alpha_x$  et  $k_t$  sont :

$$k_t^* = \hat{k}_t - \bar{k} \text{ et } \alpha_x^* = \hat{\alpha}_x + \hat{\beta}_x * \bar{k}$$

où  $\bar{k}$  désigne la moyenne des  $\hat{k}_t$  dans le temps.

Ainsi, les taux de mortalité lissés à partir des données brutes sont directement obtenus par la méthode de lissage de Lee-Carter, mais uniquement pour les âges inférieurs à 80 ans. En effet, pour les âges supérieurs ou égaux à 80 ans, comme l'échantillonnage n'est pas assez conséquent, on ne peut les obtenir directement par le modèle de Lee-Carter.

On a alors recourt à une méthode de fermeture des table, telle que celle de Coale et Kisker : celle-ci construit les taux de mortalité aux grands âges à partir des taux lissés (par Lee-Carter) aux âges de 65 et 80 ans.

### 4.1.3 La méthode d'extrapolation de Coale et Kisker

La méthode consiste à extrapoler les taux de mortalité aux grands âges (jusqu'à  $x=110$  ans par exemple) en se basant sur la formule :

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

$g_x$  désignant le taux moyen de croissance de  $\mu_x$  entre 65 et  $x$  ans. On calcule ainsi ainsi de les coefficients  $g_x$  jusqu'à un certain âge, puis on les extrapole afin de pouvoir recomposer les taux  $\mu_x$ .

Coale et Kisker ont en effet remarqué empiriquement que les courbes des  $g_x$  possèdent en général un pic aux alentours de 80 ans avant de décroître linéairement. Ils ont par conséquent proposé l'équation :

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

On peut alors utiliser la formule suivante pour extrapoler au-delà de 80 ans les taux instantanés de mortalité :

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

On utilise les valeurs de paramètres suivantes :

$$s = -\frac{\ln(\mu_{79} + 31 * g_{80})}{465} \text{ et } g_{80} = \frac{\ln(\frac{\hat{\mu}_{80}}{\hat{\mu}_{65}})}{15}$$

### 4.1.4 Application du modèle de Lee-Carter

#### 4.1.4.1 Estimation des paramètres

Rappelons que la fonction objectif vérifiée par les paramètres du modèle de Lee-Carter est la suivante :

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

Le graphique suivant présente les estimations des  $\alpha_x$  obtenues par le modèle :

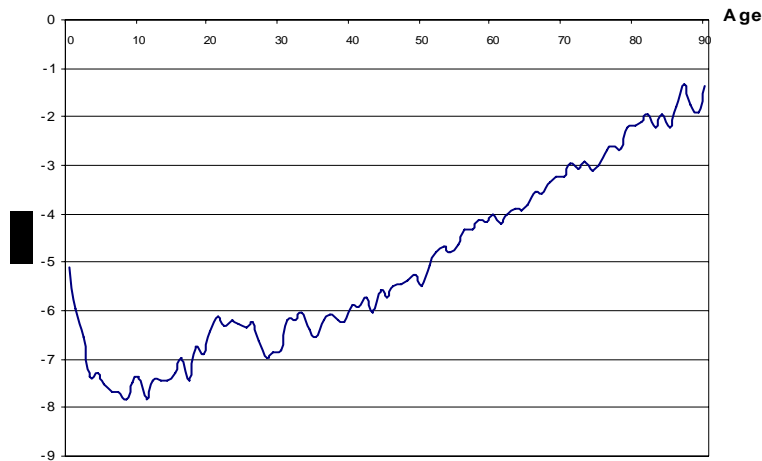


Fig. (7) : Estimation des paramètres  $\alpha_x$  par la modèle de Lee-Carter

Comme nous l'avons dit précédemment,  $\alpha_x$  est la composante spécifique à l'âge  $x$ , elle décrit le comportement moyen des  $\ln(\mu_{xt})$  au cours du temps.

Le graphique suivant illustre l'évolution des paramètres estimés  $\hat{\beta}_x$  en fonction de l'âge :

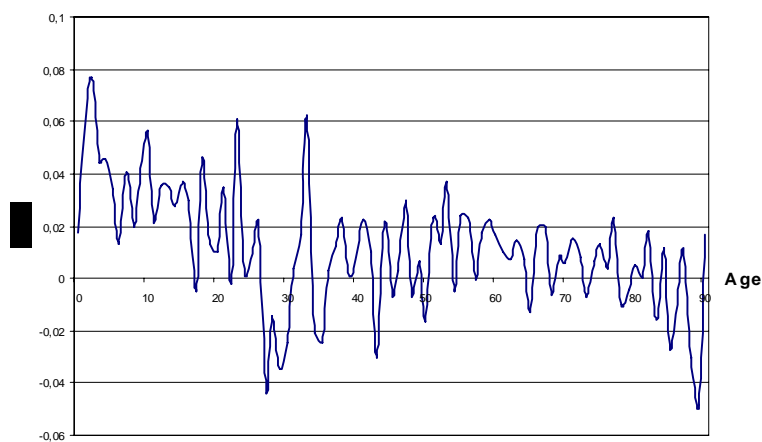


Fig. (8) : Estimation des paramètres  $\beta_x$  par la modèle de Lee-Carter

Rappelons que  $\beta_x$  indique la sensibilité de la mortalité instantanée par rapport à l'évolution générale de la mortalité. Si on se situe à partir de 20 ans, on constate que les âges les plus sensibles à l'évolution temporelle de la mortalité sont ceux entre 20 et 35 ans, et un peu avant 90 ans. On atteint en effet des pics sur ces tranches d'âges. On constate de façon générale la très grande variabilité des coefficients trouvés par la modélisation effectuée à partir des données issues de la population marocaine.

Le graphique suivant illustre l'évolution des paramètres estimés  $\hat{k}_t$  en fonction de l'année :

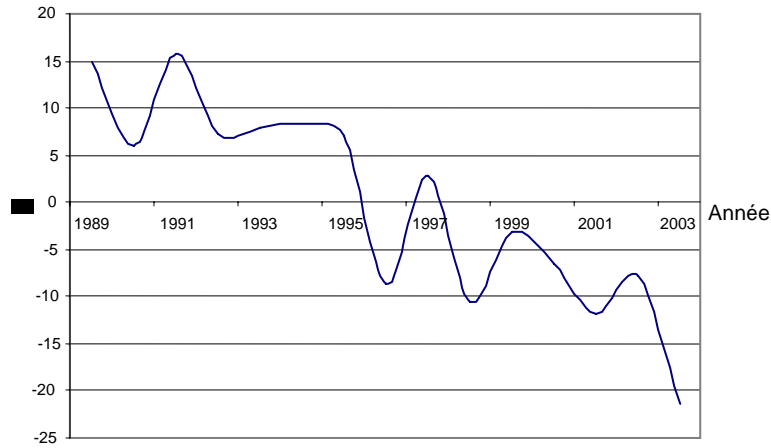


Fig. (9) : Estimation des paramètres  $k_t$  par le modèle de Lee-Carter

Rappelons que  $k_t$  désigne l'évolution générale de la mortalité dans le temps. On constate une tendance linéaire à la décroissance des  $k_t$  entre 1989 et 2003 apparaît. Cette tendance à la décroissance du paramètre  $k_t$ , qui devient négatif au cours de la période, associée à la positivité moyenne du paramètre  $\beta$  implique d'après la formule de Lee-Carter, une diminution des taux instantanés de mortalité. En conséquence, on assiste à un rallongement de la durée de vie sur la période observée.

#### 4.1.4.2 Les taux de mortalité lissés par le modèle de Lee-Carter

Comme le montre le graphique suivant, le modèle de Lee-Carter ne permet de lisser que partiellement les taux de mortalité de la population observée :

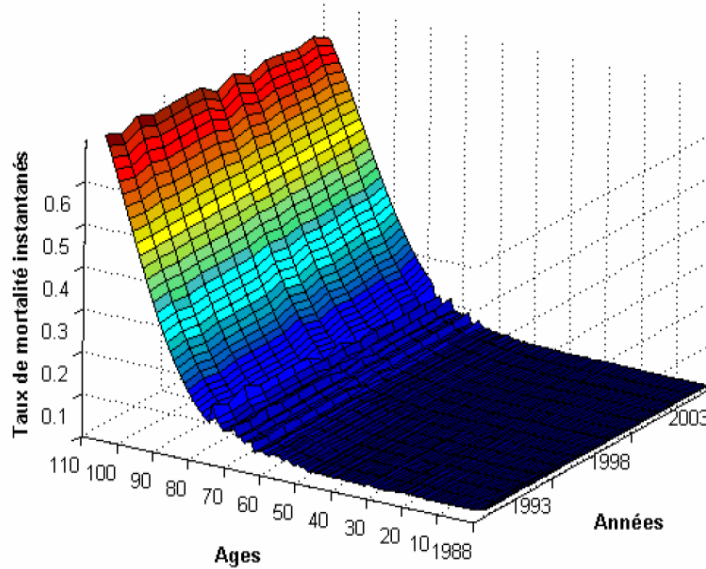


Fig. (10) : Taux de mortalité lissés par le modèle de Lee-Carter

On constate encore une certaine volatilité des taux lissés entre les années et entre les âges. Cette volatilité est particulièrement visible à partir de 65 ans. Ces variations montrent que le lissage des taux de mortalité par le modèle de Lee-Carter n'est pas très efficace. On remarquera tout de même une nette amélioration par rapport aux taux bruts.

Le graphique suivant permet de comparer les taux de 1989, lissés par le modèle aux taux bruts de la même année.

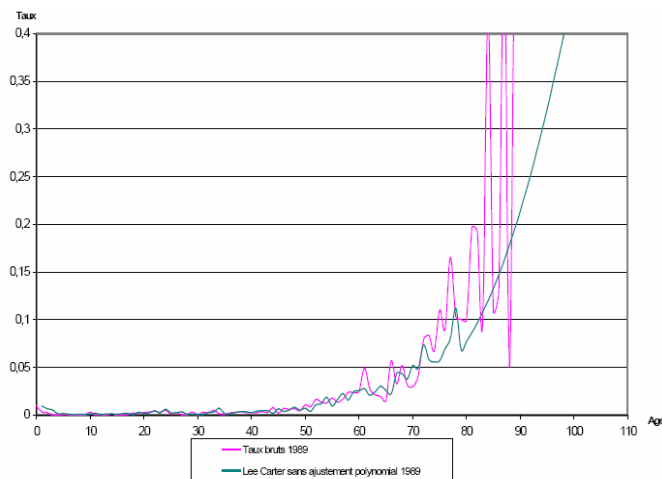


Fig. (11) : Taux de mortalité lissés par le modèle de Lee-Carter pour l'année 1989

Pour cette année 1989, les résultats obtenus par la méthode de Lee-Carter ne semblent pas optimaux. En effet, les taux sont non seulement faiblement lissés mais il apparaît également une forte discontinuité à 80 ans, âge de départ de l'extrapolation par la méthode de Coale-Kisker.

Représentons le même graphique que ci avant pour l'année 2003 :

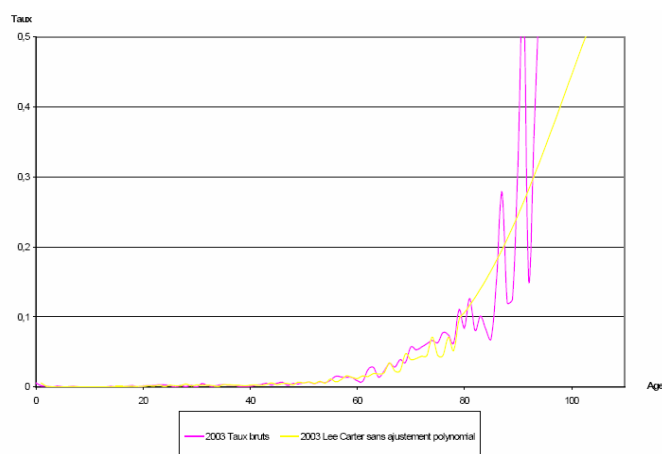


Fig. (12) : Taux de mortalité lissés par le modèle de Lee-Carter pour l'année 2003

Même si le lissage n'est pas parfait, on observe une amélioration de celui-ci sur les données de 2003 par rapport à celles de 1989. On note malgré tout des variations de taux entre 60 et 80 ans.

Les courbes des paramètres  $\alpha$ ,  $\beta$  et  $k$  montrent une forte variabilité de ces coefficients, cette dernière entraînant une instabilité des taux de mortalité obtenus par le modèle.

#### 4.1.5 Intervalles de confiance asymptotiques des taux bruts et taux lissés par le modèle de Lee-Carter

##### ➤ Théorie

Soit  $N_x$ , l'effectif sous risque de notre échantillon à l'âge  $x$  et  $q_x$  le taux de mortalité lissé à l'âge  $x$  (probabilité de décéder entre les âges  $x$  et  $x+1$ ). On a estimé  $q_x$  par  $\hat{q}_x$ . Par application du théorème central limite, il vient l'équation suivante :

$$\sqrt{N_x} \frac{q_x - \hat{q}_x}{\sqrt{\hat{q}_x(1 - \hat{q}_x)}} \xrightarrow{N \rightarrow \infty} N(0,1)$$

Ainsi, l'intervalle de confiance de niveau  $1 - \alpha$  pour  $q_x$  est donné par :

$$\left[ \hat{q}_x - u_{1-\frac{\alpha}{2}} \frac{\sqrt{\hat{q}_x(1-\hat{q}_x)}}{\sqrt{N_x}}, \hat{q}_x + u_{1-\frac{\alpha}{2}} \frac{\sqrt{\hat{q}_x(1-\hat{q}_x)}}{\sqrt{N_x}} \right]$$

Où  $u_{1-\frac{\alpha}{2}}$  désigne le quantile de niveau  $1 - \frac{\alpha}{2}$  de la loi normale centrée réduite.

➤ *Application*

Nous avons choisi un intervalle de confiance de niveau  $1 - \alpha = 95\%$ . On a alors

$$1 - \frac{\alpha}{2} = 97,5\%; \quad u_{0,975} = 1,96.$$

Montrons pour les années 1989, 1996 et 2003 les intervalles de confiance asymptotiques obtenus pour les taux bruts.

- 1989

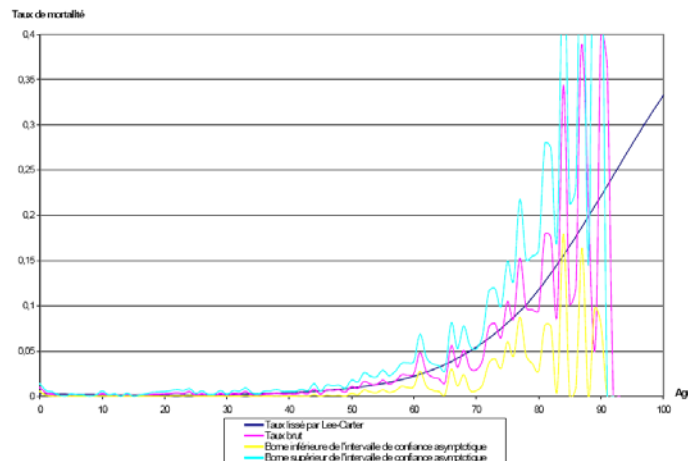


Fig. (13) : Intervalles de confiance asymptotiques pour les taux bruts en 1989 et taux lissés par le modèle de Lee-Carter

On constate que globalement, les taux lissés sont compris dans l'intervalle de confiance de niveau 95% déterminé à partir des taux bruts pour l'année 1989. Il existe seulement quelques exceptions, pour certains âges, tel qu'à 65 ans.

- 1996

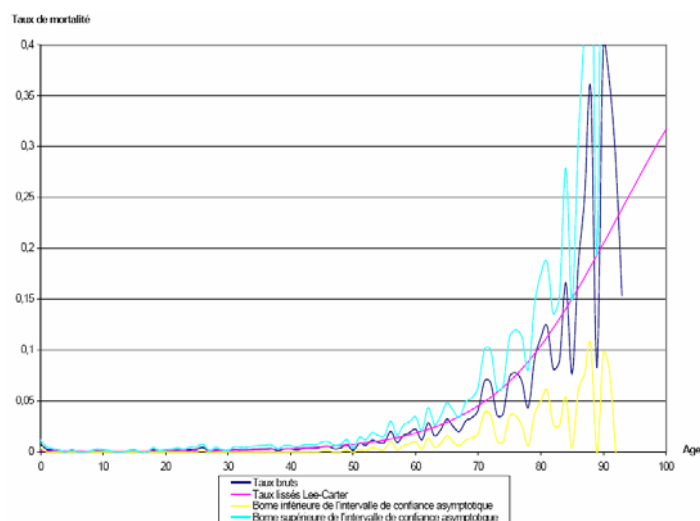


Fig. (14) : Intervalles de confiance asymptotiques pour les taux bruts en 1996 et taux lissés par le modèle de Lee-Carter

On constate à nouveau que les taux estimés appartiennent quasiment pour tous les âges à l'intervalle de confiance de niveau 95% centré sur les taux bruts. On observe après 75 ans, une tendance à la surestimation de la mortalité par les taux lissés (rappelons que les taux bruts sont très variables et s'appuient sur un faible échantillonnage). En effet, pour ces âges, les taux lissés sont fréquemment relativement proches de la borne supérieure.

- 2003

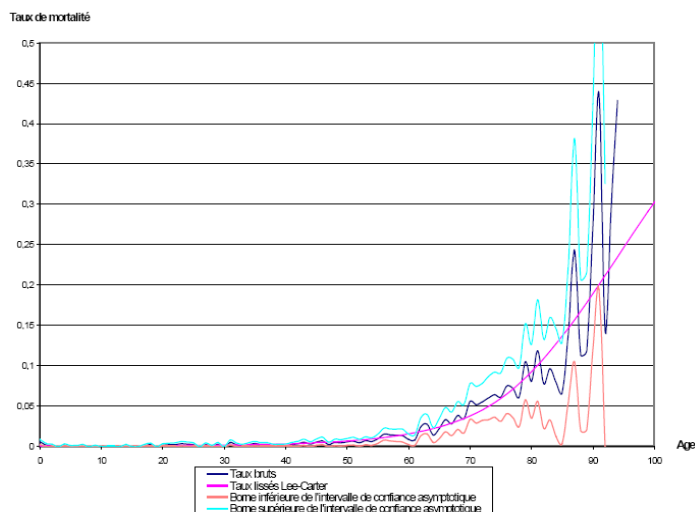


Fig. (15) : Intervalles de confiance asymptotiques pour les taux bruts en 2003 et taux lissés par le modèle de Lee-Carter



Pour l'année 2003, on peut faire la même analyse que pour l'année 1996. Néanmoins, les résultats sont meilleurs et il n'y a pas de surestimation apparente de la mortalité aux âges avancés (supérieurs à 75 ans).

On peut donc conclure que le modèle de Lee-Carter tient compte de l'amélioration de la mortalité visualisée au cours du temps sur la population étudiée. Cependant, le modèle de Lee-Carter, associé à l'extrapolation de Coale et Kisker, présente une légère tendance à la surestimation de la mortalité aux âges avancés pour les dernières années de l'étude.

#### **4.1.6 Synthèse et limites de la méthode de Lee-Carter**

Nous avons mis en évidence certaines limites du modèle de Lee-Carter dans le cas de faibles échantillons. En effet, ces paramètres servant au lissage sont très volatiles. En conséquence, on retrouve cette volatilité dans les taux de mortalité calculés par le modèle ; le lissage ne fonctionne donc que partiellement avec ce modèle.

Gutterman et Vanderhoof [1999] ont émis une critique sur le modèle de Lee-Carter portant sur l'approche purement extrapolative de leur méthode. Seule l'histoire passée peut être introduite dans le modèle. Le modèle ne peut prendre en compte les nouveaux progrès de la médecine, ce qui biaise les projections de mortalité au-delà de la période d'étude.

Le modèle repose sur les moindres carrés ordinaires, ce qui nécessite que les résidus vérifient l'hypothèse d'homoscédasticité. Or, c'est une hypothèse peu réaliste. Le graphique suivant représente les résidus de l'ajustement par Lee-Carter pour les années 1989, 1996 et 2003 :

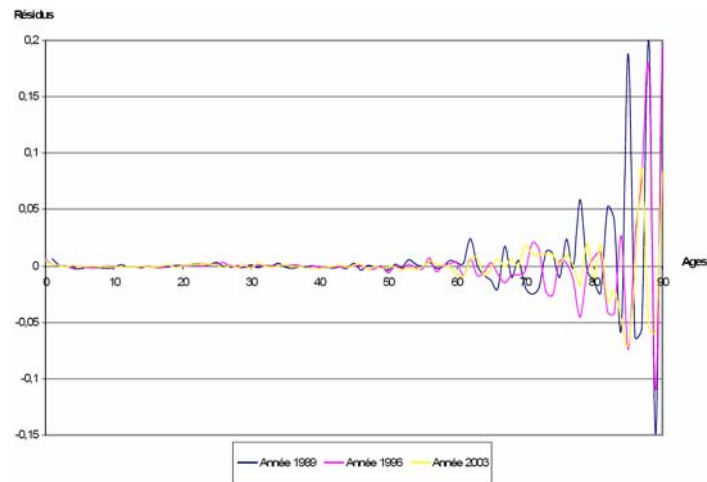


Fig. (16) : Analyse des résidus dans le modèle de Lee-Carter

On constate que la dispersion des résidus n'est pas homogène. Les résidus sont bien hétéroscédastiques.

On peut par ailleurs noter que le critère de sélection des paramètres optimaux dans le modèle de Lee-Carter n'a pas de justification probabiliste, et chercher alors à modifier l'approche pour obtenir un critère de détermination des paramètres de type « maximum de vraisemblance », qui permettra de bénéficier des « bonnes propriétés » de cette classe d'estimateurs (convergence, efficacité asymptotique, normalité asymptotique, etc.).

Tout ceci incite à rechercher un autre modèle palliant ces inconvénients. Le modèle log- Poisson, proposé par BROUHNS et al., est une adaptation du modèle de Lee-Carter précisément destinée à pallier ces deux dernières limites.

## 4.2 Ajustement par le modèle log-Poisson

### 4.2.1 Présentation du modèle

L'idée est de modéliser le nombre de décès à l'âge  $x$  et à l'année  $t$ ,  $D_{xt}$ , par une loi de Poisson, en supposant que  $D_{xt}$  suit une loi de Poisson de paramètre  $L_{xt} \cdot \mu_{xt}$  avec

$$\mu_{xt} = \exp(\alpha_x + \beta_x k_t)$$

L'expression du taux de décès instantané est identique à celle proposée dans le modèle de Lee-Carter, avec la même interprétation des différents paramètres. En particulier, le modèle ne sera identifiable qu'avec des contraintes sur les paramètres, et on peut retenir les mêmes que celles utilisées par Lee et Carter. Afin d'estimer les paramètres du modèle, on utilise la méthode du maximum de vraisemblance. La log-vraisemblance du modèle s'écrit :

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

### 4.2.2 Résolution du maximum de vraisemblance

On cherche à maximiser l'expression ci-dessus. On utilise pour cela un algorithme unidimensionnel de Newton. La méthode est itérative. On dispose de trois paramètres:  $\alpha$ ,  $\beta$  et  $k$ . Pour appliquer l'itération, on les initialise de façon libre, en veillant toutefois à choisir des valeurs  $\beta_x$  non nulles pour éviter les divisions par 0. Comme on souhaite maximiser la log-vraisemblance, il convient d'annuler les dérivées par rapport à chacun des 3 paramètres  $\alpha$ ,  $\beta$  et  $k$ .

Or,

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

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

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

On obtient donc les relations de récurrence suivantes :

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

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

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

Ensuite, il faut ajuster les paramètres ainsi estimés, afin de vérifier les contraintes d'identifiabilité. On pose alors:

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

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

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

### 4.2.3 Application du modèle log-Poisson

#### 4.2.3.1 Estimation des paramètres et comparaison avec les résultats du modèle Lee-Carter

Le graphique suivant présente les estimations des  $\alpha_x$  obtenues par le modèle :

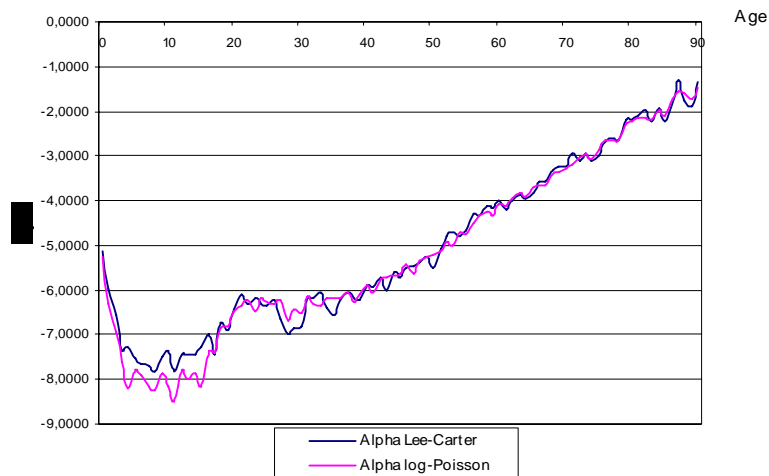


Fig. (17) : Comparaison des paramètres  $\alpha$  obtenus par les modèles de Lee-Carter et log-Poisson

Au-delà de 35 ans, les deux modèles présentent des valeurs de  $\alpha_x$  très voisines. Avant cet âge les différences de valeur sont plus nettes sans être toutefois excessivement élevées. Globalement, on peut cependant observer que les  $\alpha_x$  obtenus par la régression de Poisson sont moins volatiles que ceux obtenus par la méthode de Lee-Carter.

Le graphique suivant illustre l'évolution des paramètres estimés  $\hat{\beta}_x$  en fonction de l'âge :

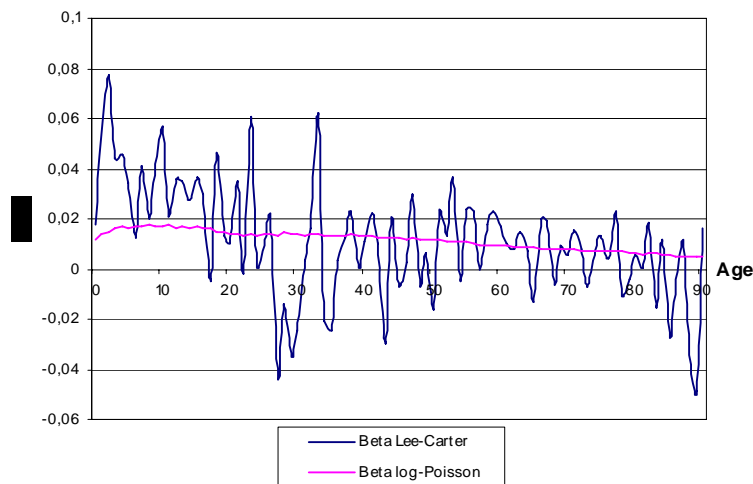


Fig. (18) : Comparaison des paramètres  $\beta_x$  obtenus par les modèles de Lee-Carter et log-Poisson

On n'observe pratiquement pas de points communs entre les deux évolutions. Cependant, ce paramètre indiquant les âges les plus sensibles à l'évolution de la

mortalité, il y a concordance sur le fait que les âges entre 0 et 20 ans sont les plus sensibles à l'évolution de la mortalité sur la période.

Le graphique suivant illustre l'évolution des paramètres estimés  $\hat{k}_t$  en fonction de l'année :

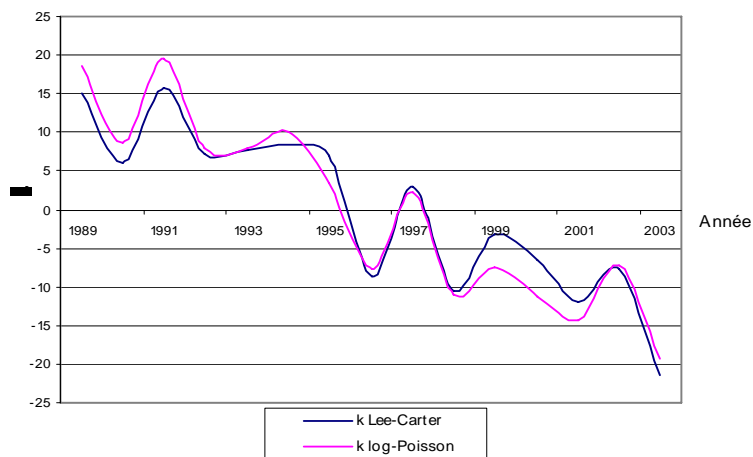


Fig. (19) : Comparaison des paramètres  $k$  obtenus par les modèles de Lee-Carter et log-Poisson

On observe des évolutions analogues pour les deux modèles et des valeurs souvent assez proches pour ces derniers.

#### 4.2.3.2 Les taux de mortalité lissés par le modèle log-Poisson

Comme le montre le graphique suivant, le modèle log-Poisson ne permet de lisser que partiellement les taux de mortalité de la population observée :

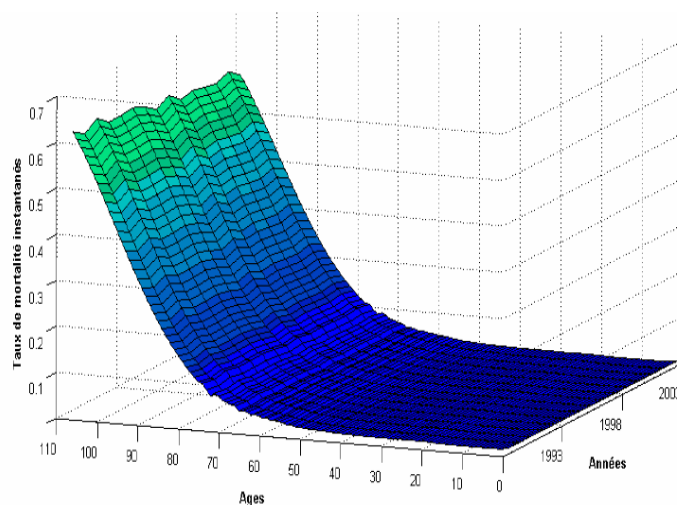


Fig. (20) : Taux de mortalité lissés par le modèle log-Poisson

Avec nos données, le modèle de Poisson présente l'avantage de mieux lisser la mortalité que le modèle de Lee-Carter, et cela pour tous les âges. Néanmoins, les taux restent assez volatiles, notamment entre 70 et 80 ans. Afin de réduire cette variabilité, on effectuera des transformations polynomiales sur les paramètres ; en effet, d'après les représentations graphiques de ces paramètres, on constate que l'on peut approximer les courbes obtenues par des fonctions polynomiales dont l'évolution est proche de celle de ces paramètres. Nous allons donc imposer une forme polynomiale aux paramètres afin de réduire leur nombre de degrés de liberté et améliorer ainsi le lissage des taux de mortalité.

Le graphique suivant permet de comparer les taux de 1990, lissés par les modèles de Lee-Carter à ceux lissés par le modèle log-Poisson :

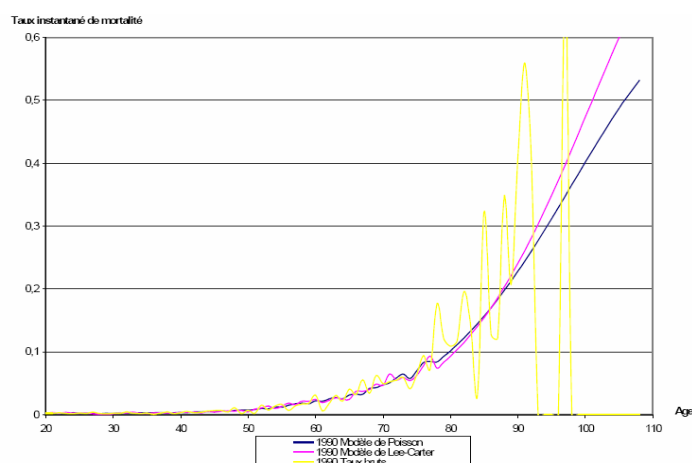


Fig. (21) : Comparaison des taux lissés par les modèles de Lee-Carter et log-Poisson pour l'année 1990

Pour l'année 1990, on observe une différence tendancielle entre les 2 modèles à partir de 78 ans. En fait, ce sont les résultats de l'extrapolation par la méthode de Coale-Kisker qui diffèrent entre les modèles de Poisson et Lee-Carter.

Le graphique suivant permet de comparer les taux de 2000, lissés par les modèles de Lee-Carter à ceux lissés par le modèle log-Poisson :

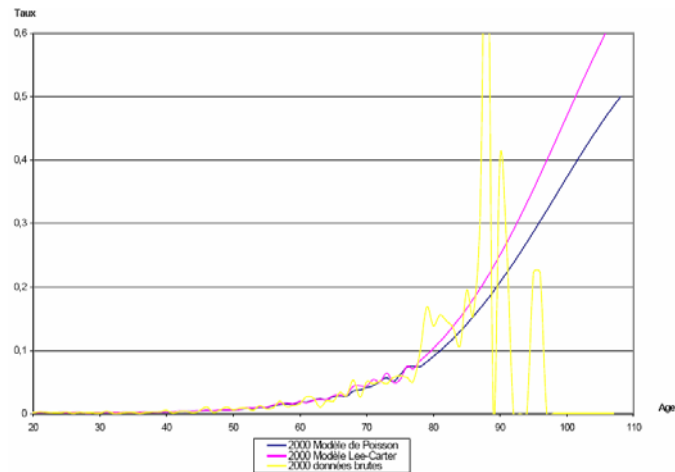


Fig. (22) : Comparaison des taux lissés par les modèles de Lee-Carter et log-Poisson pour l'année 2000

Par rapport à 1990, on constate à nouveau que l'extrapolation à partir des taux de Lee-Carter a tendance à produire des taux de mortalité plus élevés que celle basée sur les taux log-poisson.

Nous retiendrons que les taux obtenus par le modèle de Poisson sans imposer de contraintes aux paramètres  $\alpha$ ,  $\beta$  et  $k$  sont plus réguliers (la courbe est plus lisse) que ceux obtenus dans les mêmes conditions par le modèle de Lee-Carter.

La réduction de la variabilité des paramètres permettra d'améliorer le lissage obtenu par la méthode de Poisson. Cette diminution de la volatilité des paramètres sera obtenue par la réduction du nombre de degrés de liberté. Jusqu'à présent, aucune contrainte n'avait été retenue sur  $\alpha$ ,  $\beta$  et  $k$ , chacun d'eux disposant de 92 degrés de liberté (un pour chaque âge).

#### 4.2.4 Application du modèle log-Poisson dans le cas où les paramètres $\alpha$ , $\beta$ et $k$ ont une forme polynomiale

##### 4.2.4.1 Approximation polynomiale des paramètres

Dans le but d'améliorer le lissage, nous avons réduit le nombre de degrés de liberté des paramètres du modèle.

Nous avons retenu une fonction polynomiale de degré 3 pour approximer le paramètre  $\alpha$ . Le graphique suivant présente les résultats obtenus :



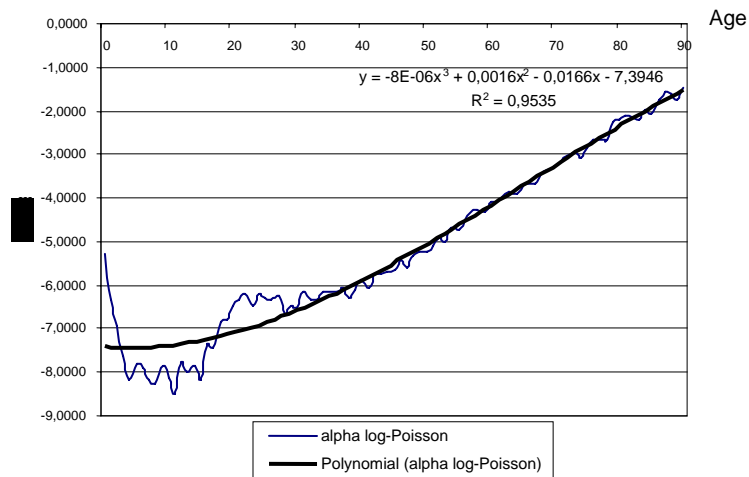


Fig. (23) : Approximation polynomiale des paramètres  $\alpha$  obtenus par le modèle log-Poisson

On a :

$$\alpha_x = -8.10^{-6} x^3 + 0,0016x^2 - 0,0166x - 7,3946$$

Avec un  $R^2=0,9535$ , cette approximation est plutôt satisfaisante.

De même, nous avons retenu une structure polynomiale de degré 3 au paramètre  $\beta$ . Le résultat obtenu est le suivant:

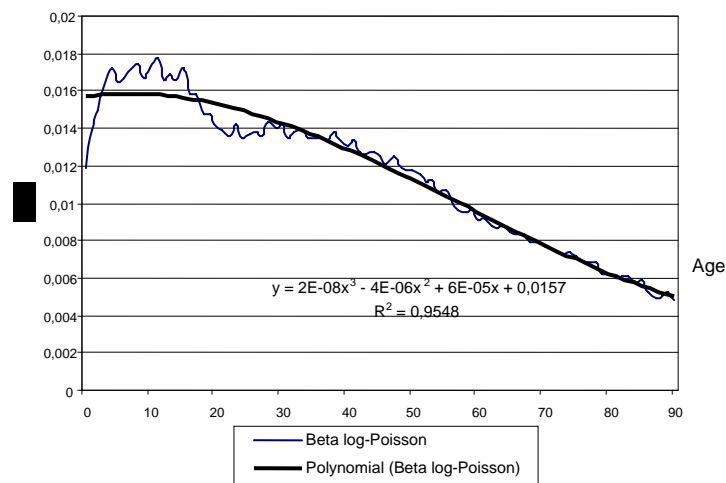


Fig. (24) : Approximation polynomiale des paramètres  $\beta$  obtenus par le modèle log-Poisson

$$\beta_x = 2.10^{-8} x^3 - 4.10^{-6} x^2 + 6.10^{-5} x + 0,0157$$

Avec un  $R^2=0,9548$ , cette approximation est nettement meilleure que celle réalisée précédemment avec la modèle de Lee-Carter.

Pour le paramètre  $k$ , nous avons retenu une régression linéaire simple :

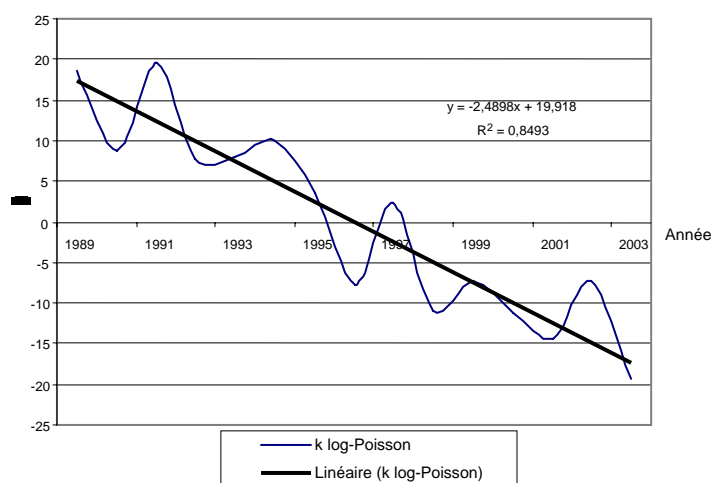


Fig. (25) : Approximation linéaire des paramètres  $k$  obtenus par le modèle log-Poisson

$$k_t = -2,4898t + 19,918 \text{ et } R^2=0,8493.$$

#### 4.2.4.2 Les taux de mortalité lissés

En donnant une structure polynomiale aux différents paramètres, on obtient les résultats suivants :

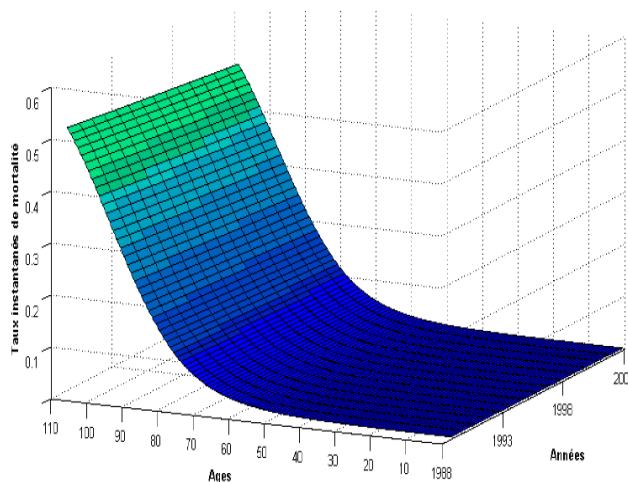


Fig. (26) : Taux de mortalité lissés par le modèle log-Poisson avec approximation polynomiale des paramètres

On constate que les taux obtenus avec le modèle de Poisson avec approximation polynomiale des paramètres, lissent de façon satisfaisante les taux bruts de la population étudiée.

Les graphes suivants nous permettront de comparer en termes de valeur le lissage par la méthode de log-Poisson lorsque les paramètres ont une structure polynomiale et les taux bruts.

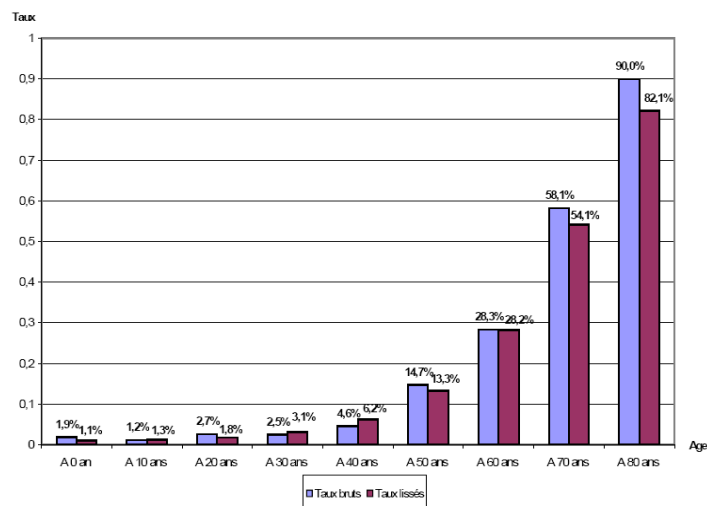


Fig. (27) : Comparaison pour l'année 1989 des risques de décéder dans la décennie issus des données brutes et de l'ajustement de log-Poisson

Pour l'année 1989, le lissage par le modèle log- Poisson a légèrement tendance à sous-estimer la mortalité observée, principalement après 50 ans. On note en effet qu'au-delà de cet âge, la probabilité de décéder au cours de la prochaine décennie est toujours inférieure avec les taux lissés par rapport aux taux bruts.

L'utilisation du modèle de Poisson apparaît par conséquent prudente pour cette année 1989 dans l'optique d'une tarification de rentes viagères.

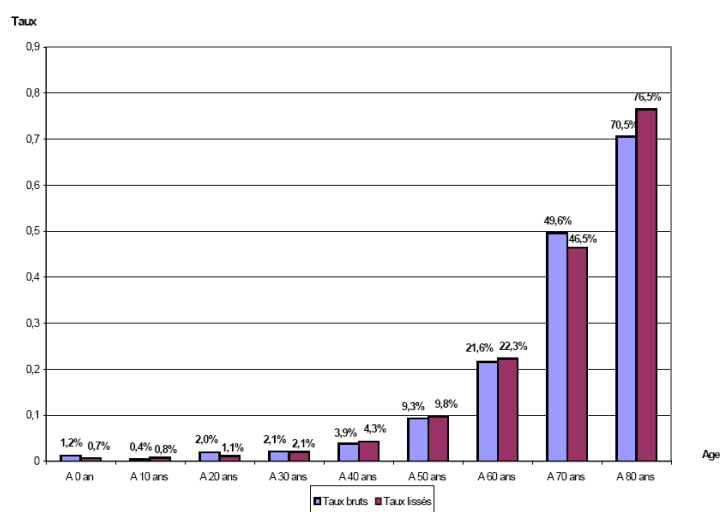


Fig. (28) : Comparaison pour l'année 2003 des risques de décéder dans la décennie issus des données brutes et de l'ajustement de log-Poisson

Les résultats obtenus pour l'année 2003 ne sont pas semblables à ceux obtenus en 1989. Avant l'âge de 60 ans, on ne distingue pas de tendance nette à la sous-estimation ou à la surestimation de la mortalité par le modèle de Poisson.

A l'âge de 60 ans, le risque de décès calculé par le lissage de Poisson est similaire (légèrement supérieur) à celui déterminé à partir des données brutes. A 70 ans, le même risque est inférieur de 3,1 % lorsqu'on effectue le calcul à l'aide des taux lissés (par rapport au risque déterminé avec les taux bruts).

Le principal changement intervient à 80 ans, puisque le risque de décéder dans la décennie calculé avec le lissage de Poisson (extrapolé par Coale-Kisker) est alors supérieur de 6 points au risque déterminé à partir des données brutes.

Cette inversion du résultat observé à 80 ans en 1989 semble montrer un défaut de la méthode d'extrapolation de Coale-Kisker, puisque ces résultats montrent qu'elle ne tient pas suffisamment compte de l'amélioration temporelle de la mortalité.

Ce graphe montre finalement que le modèle de Poisson associé à l'extrapolation de Coale et Kisker n'est pas aussi prudent en 2003 qu'en 1989 pour la détermination du coût des rentes viagères. Néanmoins, il ne peut être considéré comme dangereux, assertion que l'on justifiera dans la section consacrée à la comparaison du coût des rentes viagères.

#### 4.2.4.3 Intervalle de confiance asymptotique des taux bruts et taux lissés par le modèle log-Poisson

Application avec un intervalle de confiance de niveau  $1-\alpha = 95\%$ .

➤ 1989

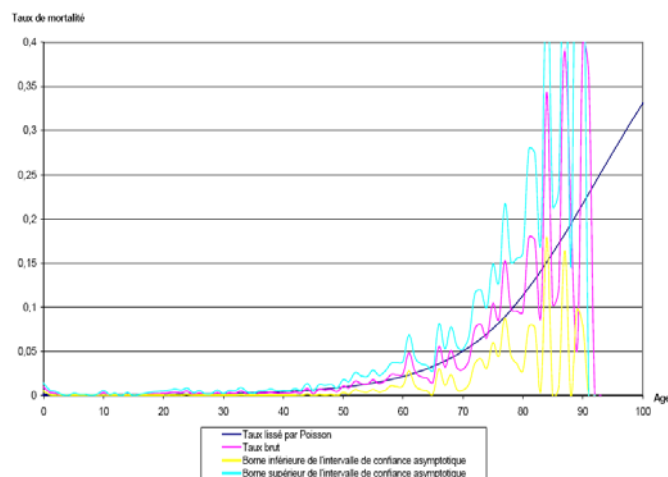


Fig. (29) : Intervalles de confiance asymptotiques pour les taux bruts en 1989 et taux lissés par le modèle log-Poisson

On observe que pour tous les âges étudiés (hormis 65 ans), les taux lissés par le modèle de Poisson sont compris dans l'intervalle de confiance asymptotique de niveau 95% déterminé à partir des taux bruts. Pour l'année 1989, les taux lissés par le modèle de Poisson ne surestiment et ne sous-estiment pas de manière apparente la mortalité de la population étudiée. En conséquence, le lissage est satisfaisant pour cette année.

➤ 1996

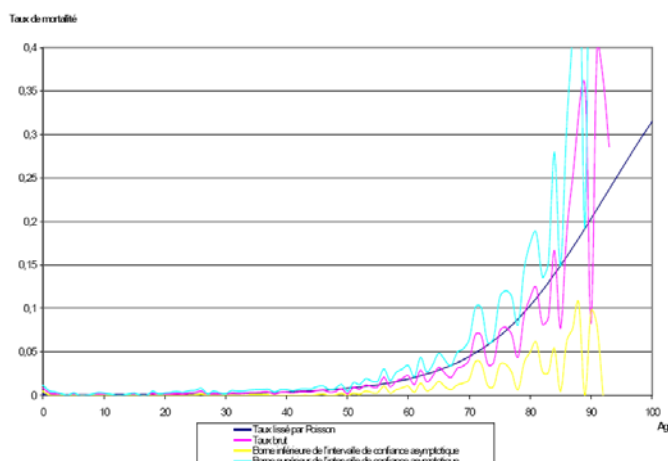


Fig. (30) : Intervalles de confiance asymptotiques pour les taux bruts en 1996 et taux lissés par le modèle log-Poisson

Comme pour l'année 1989, le lissage par le modèle log-Poisson fournit des résultats satisfaisants en ce qui concerne les intervalles de confiance asymptotiques. Il n'y a que pour un âge, 78 ans, que les taux lissés n'appartiennent pas à l'intervalle de confiance de niveau 95% calculé à partir des données brutes. Après 70 ans, on distingue un taux de mortalité issu du lissage proche de la borne supérieure de l'intervalle de confiance asymptotique, ce qui montre une légère surestimation de la mortalité par le lissage.

➤ 2003

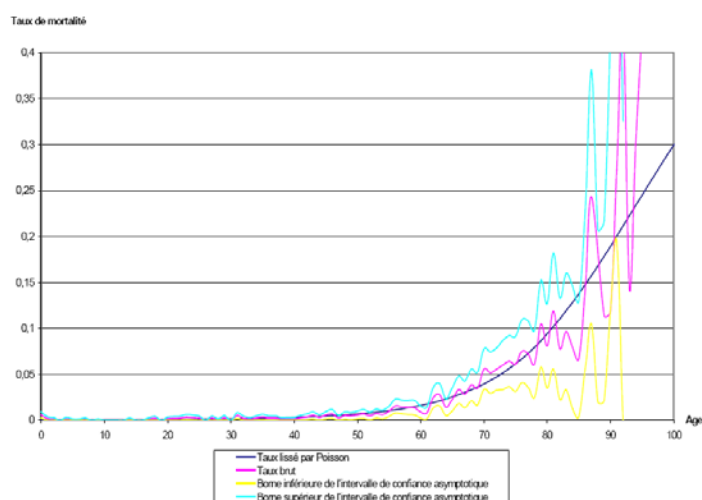


Fig. (31) : Intervalles de confiance asymptotiques pour les taux bruts en 2003 et taux lissés par le modèle log-Poisson

En 2003, les taux lissés appartiennent à nouveau pour tous les âges à l'intervalle de confiance de niveau 95%, calculé à partir des données brutes. Il existe deux exceptions (à 61 ans et 85 ans) où les taux lissés excèdent légèrement la borne supérieure de l'intervalle de confiance.

Ainsi, le modèle de Poisson a effectivement tendance à surestimer légèrement la mortalité aux grands âges sur les données de la mortalité de la population marocaine étudiée. De plus, cette tendance apparaît sur les dernières années de l'étude, ce qui semble indiquer que l'amélioration de la mortalité (au regard des données brutes) n'est pas assez prise en compte par le modèle log-poisson combiné à l'extrapolation Coale et Kisker.

## 4.3 Autre méthode d'extrapolation des taux de mortalité aux grands âges : l'extrapolation des polynômes des coefficients $\alpha$ , $\beta$ et $k$

### 4.3.1 Présentation

Après avoir imposé une structure polynomiale aux paramètres  $\alpha$ ,  $\beta$  (sur la plage d'âges 0-90 ans pour la méthode de Poisson) et  $k$  (sur les générations 1989 à 2003), on conçoit une deuxième méthode d'extrapolation des taux de mortalité aux âges avancés. Celle-ci consiste à extrapoler les polynômes sur les paramètres  $\alpha$  et  $\beta$  pour les âges allant de 91 à 108 ans. A l'aide de ces extrapolations, on pourra déterminer les taux de mortalité pour ces âges avancés, par application du modèle de Poisson.

### 4.3.2 Les résultats obtenus

Le graphique suivant présente les résultats obtenus pour le paramètre  $\alpha$  :

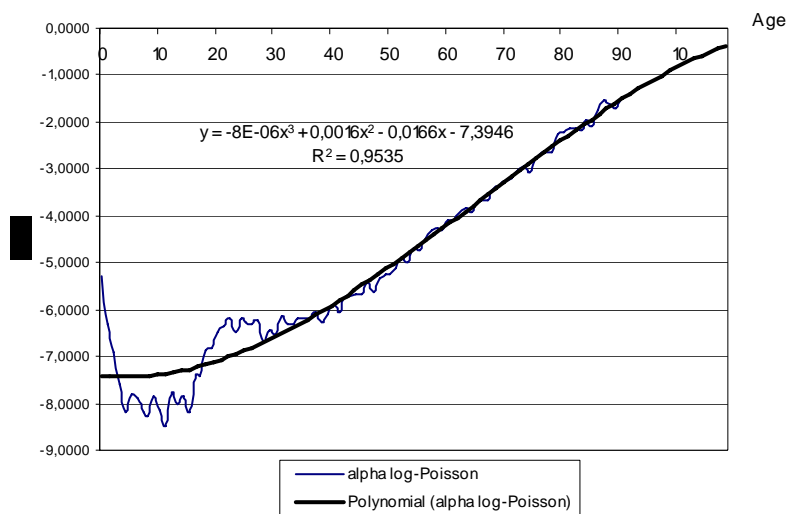


Fig. (32) : Extrapolation aux grands âges du paramètre  $\alpha$  de structure polynomiale

Pour le paramètre  $\beta$ , le graphique suivant présente les résultats obtenus :

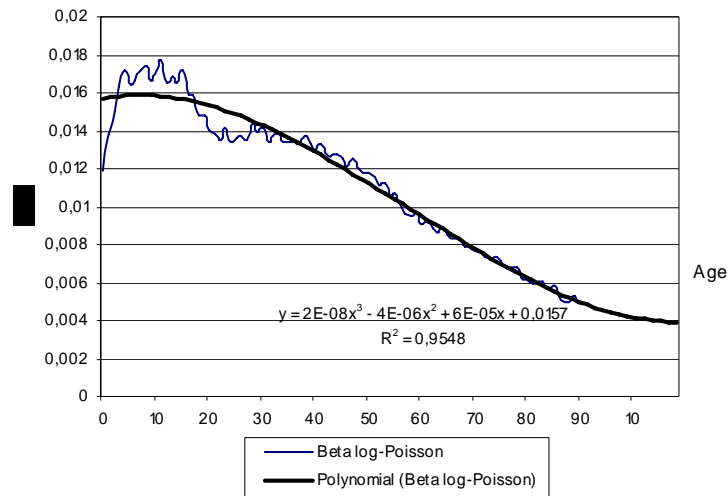


Fig. (33) : Extrapolation aux grands âges du paramètre  $\beta$  de structure polynomiale

Ces extrapolations conduisent aux taux de mortalité représentés ci-après :

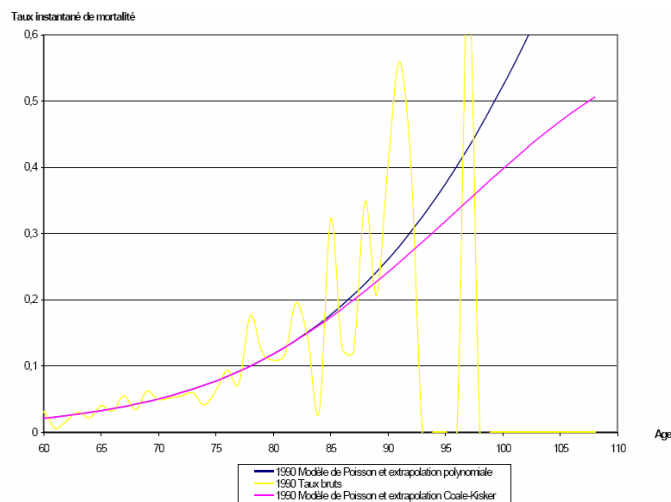


Fig. (34) : Comparaison des taux de mortalité 1990 aux âges avancés obtenus par extrapolation polynomiale avec ceux obtenus par l'extrapolation de Coale-Kisker.

En 1990, le modèle d'extrapolation polynomiale appliqué sur le modèle log-Poisson conduit à des taux de mortalité instantanés plus élevés aux grands âges. L'écart grandit progressivement, et il est très net à partir de 90 ans.



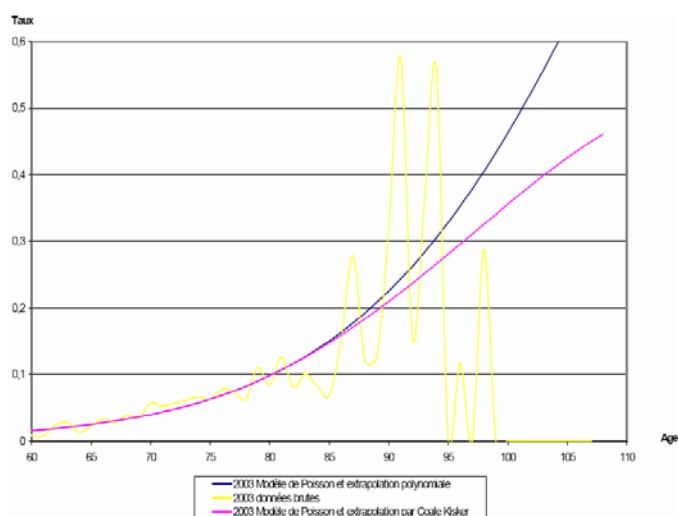


Fig. (35): Comparaison des taux de mortalité 2003 aux âges avancés obtenus par extrapolation polynomiale avec ceux obtenus par l'extrapolation de Coale-Kisker.

Les résultats obtenus en 2003 sont analogues à ceux de 1990 : l'extrapolation polynomiale conduit à des taux instantanés de mortalité plus élevés que la méthode d'extrapolation de Coale-Kisker.

Il n'est pas facile de déterminer la méthode d'extrapolation qui correspond le mieux à la réalité, puisqu'on a eu recours à l'extrapolation en raison du faible échantillonnage (impliquant des taux très variables) aux grands âges. Cependant, par soucis de prudence pour le calcul des rentes viagères, on adoptera la méthode d'extrapolation de Coale-Kisker; celle-ci présentant des taux de mortalité plus faibles aux grands âges (sur nos données).

# Chapitre 5

## Détermination du meilleur modèle de lissage

### 5.1 Le test du Chi-Deux d'ajustement

Il s'agit ici de calculer la déviation entre les décès observés sur les données brutes et les *décès théoriques* (obtenus par les modèles de Lee-Carter ou Poisson) pour chaque génération (de 1989 à 2003) et chaque classe d'âge (de 0 à 89 ans).

Le calcul cette déviation pour un âge  $x$  et une génération  $t$  s'effectue grâce à la formule suivante :

$$\chi_{x,t}^2 = \frac{(D_{x,t} - D_{x,t}^*)^2}{D_{x,t}^*},$$

où  $D_{x,t}$  représente le nombre de décès observés (issus directement des taux bruts) à l'âge  $x$  et pour l'année, et  $D_{x,t}^*$  représente les *décès théoriques* obtenus par application de la méthode de lissage de Lee-Carter ou de Poisson.

La valeur globale de la déviation entre les décès observés et théoriques vaut alors la somme de tous les Chi-2 locaux:

$$\chi_t = \sum_x \chi_{x,t}^2 = \sum_x \frac{(D_{x,t} - D_{x,t}^*)^2}{D_{x,t}^*}$$

#### 5.1.1 Statistiques du Chi-Deux avec le modèle log-Poisson

- *Sans ajustement polynomial des paramètres*

On considère les décès sur la plage d'étude suivante: 0 an - 89 ans. Les statistiques du  $\chi^2$  sont effectuées sur 9 classes comprenant chacune dix années. Puis, on mesure les décès entre 0 et 79 ans, répartis en 8 classes d'âges. On obtient les résultats suivants :

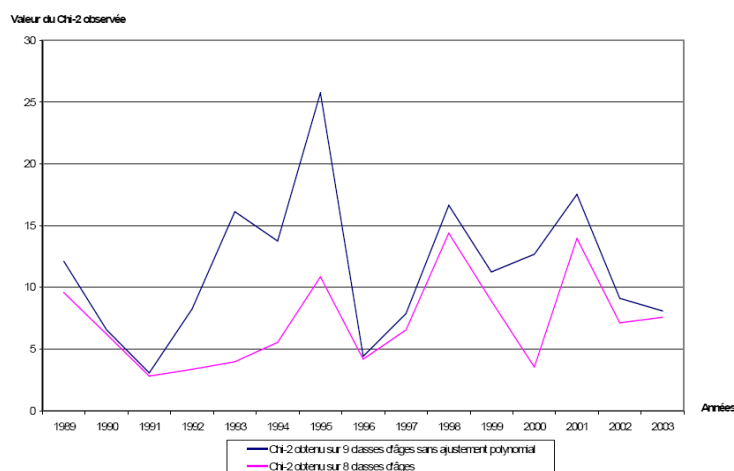


Fig. (36) : Statistiques du Chi-2 réalisées sur les données brutes et l'ajustement par le modèle log-Poisson (sans ajustement polynomial des paramètres).

En enlevant la classe d'âge [80 ans, 89 ans] de notre étude, les valeurs du Khi-2 sont en forte diminution pour les années 1992, 1993, 1994, 1995 et 2000. Cela montre certaines limites de la méthode d'extrapolation par Coale-Kisker (du moins lorsqu' aucun ajustement polynomial des paramètres n'est effectué).

➤ *Avec ajustement polynomial des paramètres :*

On réduit la plage d'études à 30-89 ans, car l'ajustement polynomial des paramètres, notamment pour le paramètre alpha ôte de la précision à l'estimation des taux de mortalité sur les premières années de la vie. En outre, notre modèle de lissage doit surtout être adapté au provisionnement des rentes viagères, donc cet obstacle impliqué par le lissage des paramètres peut être admis si les valeurs du  $\chi^2$  ne sont pas trop importantes sur la tranche d'âge 30-89 ans.

On constitue alors 6 classes sur notre plage d'étude, chaque classe comportant 10 années :



Fig. (37): Statistiques du Chi-2 réalisées sur les données brutes et l'ajustement par la méthode log-Poisson (avec et sans ajustement polynomial des paramètres).

On constate que le fait d'imposer une structure polynomiale aux paramètres de la méthode n'enlève pas de précision à l'estimation sur la plage d'âge 30 ans - 89 ans (les valeurs du Chi-2 sont même parfois inférieures avec l'ajustement polynomial). En effet, lorsqu'on effectue un ajustement par le modèle log-Poisson, les valeurs du  $\chi^2$  observé restent proches, que le lissage soit réalisé avec ou sans ajustement polynomial des paramètres.

En enlevant la classe d'âge [80 ans ; 89 ans], on obtient les résultats suivants :

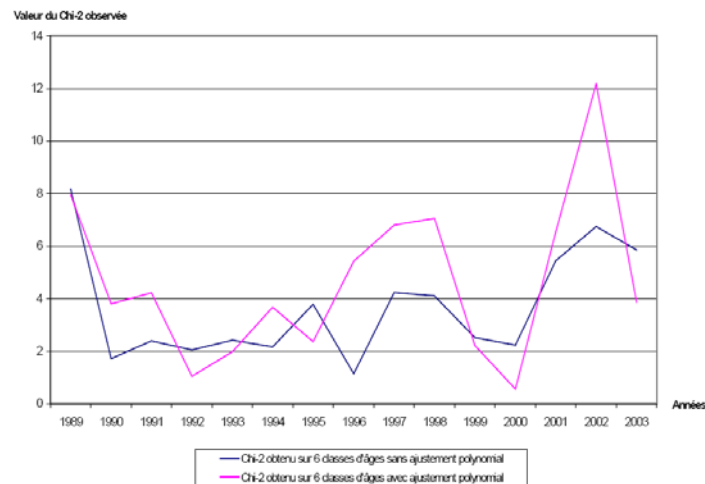


Fig. (38): Statistiques du Chi-2 réalisées sur les données brutes et l'ajustement par la méthode log-Poisson (avec et sans ajustement polynomial des paramètres).

Les classes d'âges considérées pour faire le test du Chi-2 sont les mêmes que pour le graphe précédent, hormis le fait qu'on ne considère plus la classe [80 ans ; 89

ans]. On observe que pour la plupart des années, et souvent de façon très nette, les valeurs du  $\chi^2$  observé sont supérieures lorsque les paramètres sont ajustés polynomialement (courbe rose). Comme on ne retrouve pas cette différence entre les  $\chi^2$  observés lorsque la classe [80 ans, 89 ans] est intégrée à l'étude, on en déduit que la méthode d'extrapolation de Coale-Kisker des taux de mortalité, appliquée à partir des taux lissés par le modèle de Poisson, produit des résultats beaucoup plus proches des taux bruts lorsque les paramètres du modèle ont une forme polynomiale.

### 5.1.2 Statistiques du Chi-Deux avec le modèle de Lee-Carter

Le graphe suivant montre les statistiques du Chi-2 obtenues d'une part avec les décès lissés suivant la méthode de Lee-Carter et les données brutes (en rose), et d'autre part, avec le nombre de décès lissés suivant la méthode log-Poisson, et les données brutes (en bleu), ceci sur la tranche d'âges allant de 20 à 79 ans.

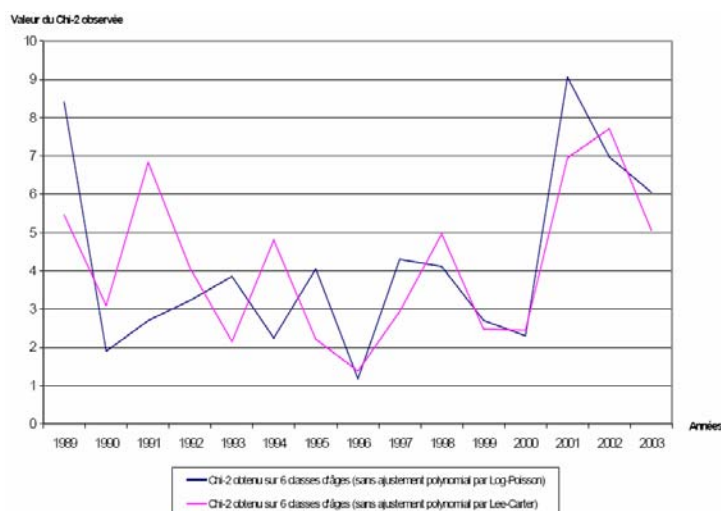


Fig. (39): Statistiques du Chi-2 sur l'ajustement par le modèle de Lee-Carter (sans ajustement polynomial des paramètres) comparées aux statistiques obtenues à partir du modèle log-Poisson

Ce graphe montre que les valeurs de  $\chi^2$  restent faibles (comprises entre 1 et 9) pour toutes les années, avec les deux méthodes. De plus, il apparaît que le lissage (sans ajustement des paramètres) obtenu par le modèle log-poisson est légèrement plus proche des données brutes que celui obtenu par le modèle de Lee-Carter (valeurs du  $\chi^2$  en général légèrement inférieures lorsque les décès lissés sont déterminés par le modèle log-poisson).

## 5.2 Comparaison des capitaux constitutifs de rente

On verse 1€ annuellement à un retraité à partir de l'âge de 60 ans (on considère cet âge comme celui d'entrée à la retraite). Ce versement se poursuit tant que la personne est en vie. L'objectif est de comparer le coût de ce versement avec les deux méthodes de lissage et avec les taux bruts.

Le montant des capitaux à constituer est calculé comme suit :

$$\text{Coût} = \sum_{k=0}^{\infty} v^k {}_k p_x$$

où  ${}_k p_x$  désigne la probabilité pour un individu d'âge  $x$  d'être en vie dans  $k$  années et  $v^k$  le facteur d'actualisation. On suppose que le taux d'actualisation est constant et égal à 2,25%.

Les capitaux constitutifs moyens à constituer, obtenus sur l'ensemble de la période d'observation sont les suivants :

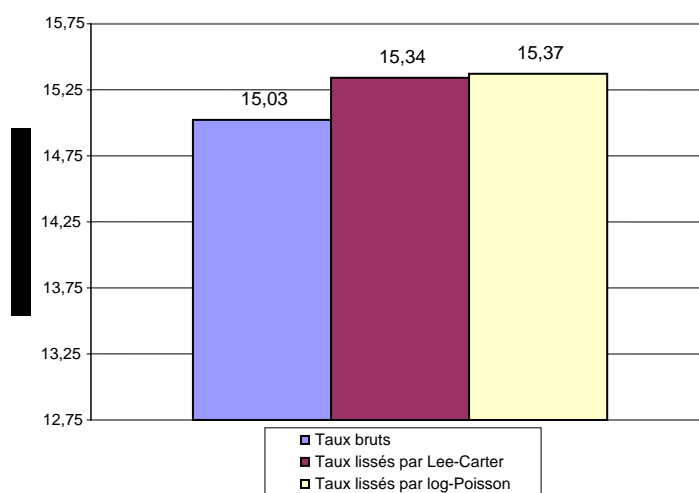


Fig. (40) : Comparaison des capitaux constitutifs obtenus avec les taux bruts et les taux lissés respectivement par les modèles de Lee-Carter et log-Poisson

Ce graphique confirme le fait qu'en moyenne, les modèles de Lee-Carter et log-Poisson conduisent à des résultats très proches. En effet, les deux méthodes prospectives d'ajustement impliquent un provisionnement très proche pour le versement de rentes viagères, ceci pour toutes les années et tous les âges. En fait,

il n'y a jamais plus de 1% d'écart entre le provisionnement issu des tables produites par ces modèles.

Une analyse menée sur les résultats obtenus conduit à constater que l'écart entre les deux provisionnements varie fréquemment entre 3% et 10% pour les âges compris entre 60 ans et 70 ans. Au-delà de cet âge, l'écart entre les deux provisionnements a tendance à devenir plus important (souvent supérieur à 15%) pour les âges compris entre 75 ans et 85 ans.

Néanmoins, il apparaît plus prudent de déterminer le coût des rentes viagères à l'aide du modèle log-Poisson par rapport à celui de Lee-Carter. Nous remarquerons toute fois qu'en moyenne, ces 2 modèles conduisent à provisionnement supérieur à celui des taux bruts.

# Chapitre 6

## Estimation de la mortalité future

### 6.1. Théorie

La projection des valeurs du paramètre  $k_t$  pour les années futures va nous permettre en appliquant la méthode d'ajustement par le modèle log-Poisson d'estimer la mortalité future de notre population. Il se pose naturellement le problème lié au faible nombre d'années d'observations (15). De ce fait, on ne peut effectuer aisément des projections sur un horizon lointain.

Comme nous l'avons précédemment mis en évidence, le paramètre  $k_t$  a une évolution linéaire. Le graphique suivant présente la projection des  $k_t$  à l'horizon 2020 (même si nos 15 années d'observation ne nous autorisent normalement pas à effectuer des projections sur un horizon aussi lointain) :

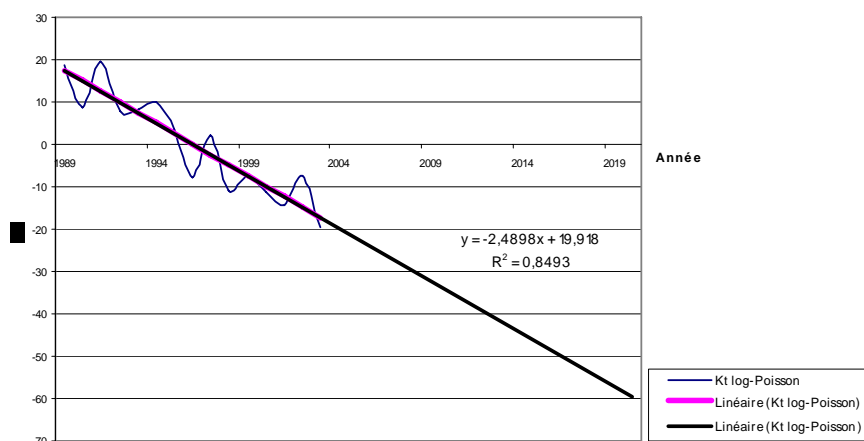


Fig. (41) : Extrapolation de la composante temporelle du modèle log-Poisson.



## 6.2 Prédiction des taux de mortalité futurs par le modèle logPoisson

On utilise simplement les projections de  $k_t$  pour obtenir les tables de mortalité prospectives. Une fois la projection  $k_{2003+t}$ ,  $t=1,2,\dots,17$  obtenue, nous en déduisons :

$$\hat{\mu}_x(2003+t) = e^{\alpha_x + \beta_x * k_{2003+t}}, \text{ pour } t \text{ allant de } 1 \text{ à } 17.$$

Les résultats obtenus sont les suivants :

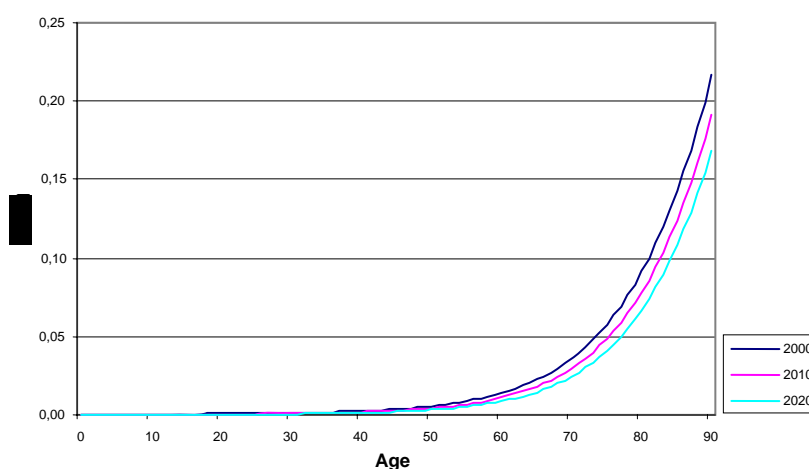


Fig. (42) : Projection des taux de mortalité instantanés futurs par le modèle log-Poisson

La projection des taux de mortalité montre une amélioration de ces taux au cours des trois prochaines décennies, on distingue nettement la diminution des taux de mortalité au-delà de 60 ans.

Afin d'apprécier en valeur cette amélioration de la mortalité, nous avons calculé pour les années 2000 et 2010 le rapport des taux instantanés. Nous avons obtenu les résultats suivants :

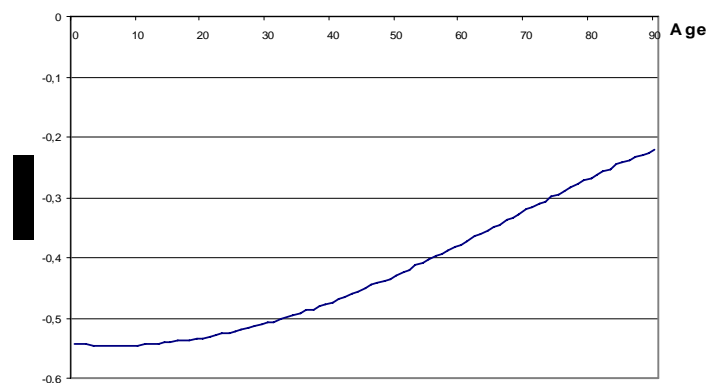


Fig. (43): Rapport entre les taux instantanés 2010 et les taux instantanés 2000, lissés par le modèle log-Poisson

L'évolution des taux de mortalité instantanés entre 2000 et 2010 (l'évolution entre 2010 et 2020 est semblable puisqu'on extrapole linéairement l'évolution du paramètre  $kt$ ) montre une amélioration importante des taux de mortalité aux âges inférieurs à 30 ans. Plus l'âge augmente, moins l'amélioration est forte en pourcentage.

Cependant, même aux âges les plus avancés, une diminution sensible des taux de mortalité entre 2000 et 2010 devrait être observée. Ainsi, au-delà de 80 ans, l'amélioration moyenne des taux de mortalité est proche de 25%.

Une analyse des taux futurs estimés par le modèle de Lee-Carter montre que les évolutions projetées des taux de mortalité instantanés sont différentes selon qu'on emploie ce dernier modèle ou celui de Poisson. On observe qu'entre 20 ans et 55 ans, les taux de mortalité futurs projetés par le modèle de Lee-Carter sont supérieurs à ceux projetés par le modèle log-poisson. Au-delà de 55 ans, et jusqu'à 95 ans, c'est l'inverse qui se produit. On présume ainsi qu'à partir de 2000, le modèle de Lee-Carter est plus prudent pour l'estimation de rentes viagères versées à des retraités de plus de 60 ans.

### 6.3. Evolution de l'espérance de vie à 60 ans

On calcule les espérances de vie résiduelles à 60 ans, cet âge étant choisi dans l'optique du calcul des rentes viagères.

Rappelons que l'espérance de vie résiduelle à 60 ans de notre population atteignait 16,40 ans en 2000, d'après les données brutes (en ne considérant pas la mortalité au-delà de 90 ans, faute de données suffisante. Le lissage par le modèle de Poisson (avec les paramètres sous forme polynomiale et extrapolés aux grands âges) indique une espérance de vie – la même année et au même âge, de 16,70

ans. On constate donc une légère sous-estimation de la mortalité par le lissage de Poisson ; il y a en effet une surestimation de 1,8% de l'espérance de vie par rapport aux données brutes – sans comptabiliser la mortalité au-delà de 90 ans.

Pour les années postérieures à 2003, le graphe suivant montre la progression temporelle, linéaire en raison de l'extrapolation linéaire réalisée sur le coefficient  $k$ , de l'espérance de vie à 60 ans:

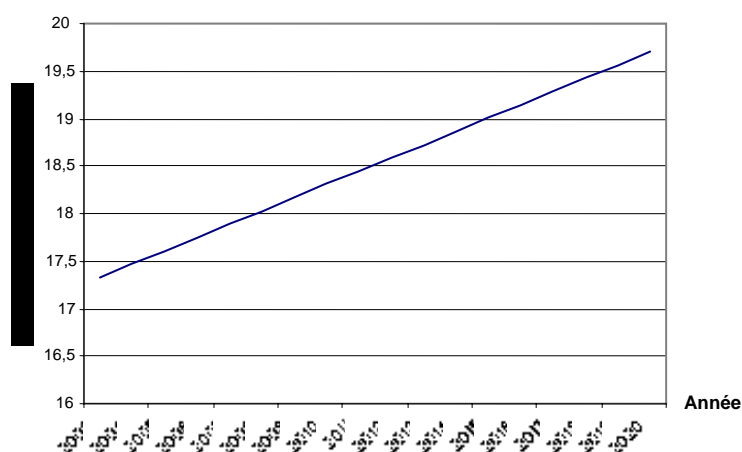


Fig. (44) : Projection de l'espérance de vie future à 60 ans par le modèle Log-Poisson.

D'après ces résultats, l'espérance de vie à 60 ans va augmenter de 2 ans et 5 mois entre 2003 et 2020, soit une augmentation annuelle supérieure à un mois et demi. L'espérance de vie à 60 ans serait voisine de 20 ans en 2020 (en 2003, elle était de 17 ans et 4 mois).

**Remarque :** le modèle de Lee-Carter conduit à une croissance plus forte de l'espérance de vie à 60 ans, par rapport à celle déterminée par le modèle log-Poisson. En effet, avec ce modèle, l'espérance de vie augmente d'un peu plus de deux mois par an et atteindrait 21 ans et demi en 2020.

# Conclusion

A partir de données brutes, la création de tables de mortalité prospectives prend en compte les deux composantes, âge et année, pour l'estimation des taux de mortalité, et permet la projection de la mortalité future de la population. Deux méthodes différentes, Lee-Carter et log-Poisson, ont été appliquées dans ce document. Le modèle de Lee-Carter ne prend pas en compte l'hétéroscédasticité des résidus alors que sa variante log-Poisson permet de s'affranchir de l'hypothèse d'hétéroscédasticité et met de surcroît en œuvre une estimation par maximum de vraisemblance.

La construction de ces tables prospectives présente la particularité dans cette étude de s'effectuer sur des portefeuilles de taille réduite. Cela implique des difficultés non rencontrées sur les précédents travaux, lorsque ces modèles étaient utilisés sur des populations ayant une échelle nationale.

D'une part, la volatilité des taux de mortalité bruts observés, liée aux fluctuations d'échantillonnage, entraîne une instabilité des trois coefficients permettant le calcul des taux de mortalité par les méthodes prospectives. On est ainsi conduit à diminuer le nombre de degrés de liberté de ces paramètres. Le fait d'imposer une structure polynomiale aux paramètres a finalement permis un meilleur lissage de la mortalité observée, en réduisant de manière drastique le nombre de paramètres du modèle.

D'autre part, aux âges les plus élevés, au-delà de 90 ans, on ne dispose plus de données brutes sur une population de taille suffisante. Cette caractéristique explique l'intervention de la méthode d'extrapolation de Coale-Kisker pour obtenir les taux de mortalité aux âges avancés.

Les résultats obtenus par l'utilisation de ces deux modèles montrent qu'en présence de l'ajustement polynomial des paramètres et lorsqu'on applique les méthodes sur les effectifs bruts, les taux de mortalité obtenus sont très proches.

Le modèle log-Poisson, associé à l'extrapolation de Coale-Kisker, apparaît meilleur sur la période 1989-2003 (période d'étude) que le modèle de Lee-Carter. De surcroît, il est plus prudent que celui de Lee-Carter pour l'estimation de rentes viagères sur cette période.

Mais le modèle de Lee-Carter prenant davantage en compte l'amélioration temporelle de la mortalité, il devient progressivement plus prudent pour l'estimation de rentes viagères que la méthode log-Poisson. Ainsi, nous préconisons l'utilisation de Lee-Carter pour une estimation prudente du coût des rentes viagères futures.

Enfin, d'autres techniques de lissage que celles présentées ici auraient pu être appliquées, on pense notamment aux modèles à référence externe, aux ajustements par splines ou encore à la méthode non paramétrique de Whittaker-Henderson.

# Annexes

## Annexe 1 : Coefficients alpha sans ajustement polynomial

Age	Alpha Lee-Carter	Alpha log-Poisson	Age	Alpha Lee-Carter	Alpha log-Poisson
0	-5,1191	-5,2725	46	-5,5101	-5,4087
1	-5,9758	-6,3276	47	-5,4495	-5,6161
2	-6,5649	-6,9240	48	-5,3691	-5,3331
3	-7,3715	-7,6127	49	-5,2702	-5,2549
4	-7,2904	-8,1892	50	-5,5119	-5,2348
5	-7,5140	-7,7879	51	-5,0770	-5,1146
6	-7,6748	-7,8941	52	-4,8040	-4,9122
7	-7,6760	-8,1740	53	-4,6947	-4,9844
8	-7,8337	-8,2564	54	-4,7952	-4,6973
9	-7,4761	-7,8671	55	-4,6510	-4,7314
10	-7,3484	-8,1390	56	-4,3079	-4,4902
11	-7,8196	-8,4879	57	-4,3428	-4,3268
12	-7,4340	-7,7654	58	-4,1396	-4,2659
13	-7,4391	-8,0105	59	-4,1706	-4,3216
14	-7,4240	-7,8518	60	-4,0042	-4,0761
15	-7,2853	-8,1719	61	-4,1957	-4,1019
16	-6,9886	-7,3632	62	-4,0029	-3,9423
17	-7,4564	-7,4172	63	-3,8794	-3,8423
18	-6,7225	-6,8583	64	-3,9347	-3,9226
19	-6,9039	-6,8154	65	-3,8069	-3,7031
20	-6,4353	-6,4738	66	-3,5639	-3,6512
21	-6,1113	-6,3371	67	-3,5663	-3,6682
22	-6,3161	-6,2049	68	-3,3380	-3,4001
23	-6,1991	-6,4848	69	-3,2250	-3,3635
24	-6,2888	-6,1783	70	-3,2474	-3,2627
25	-6,3618	-6,2785	71	-2,9633	-3,1903
26	-6,2148	-6,3209	72	-3,0912	-3,0382
27	-6,6239	-6,2398	73	-2,9245	-2,9340
28	-6,9657	-6,6759	74	-3,0989	-3,0549
29	-6,8400	-6,4546	75	-2,9833	-2,8545
30	-6,8311	-6,5256	76	-2,7009	-2,6682
31	-6,2030	-6,1570	77	-2,6268	-2,6475
32	-6,1776	-6,3233	78	-2,6666	-2,6544
33	-6,0510	-6,3298	79	-2,2029	-2,2710
34	-6,3853	-6,1663	80	-2,2000	-2,2140
35	-6,5511	-6,1723	81	-2,0958	-2,1322
36	-6,2076	-6,1731	82	-1,9627	-2,1391
37	-6,0657	-6,0619	83	-2,2158	-2,1957
38	-6,1975	-6,2763	84	-1,9459	-1,9811
39	-6,2263	-6,0912	85	-2,2334	-2,0871
40	-5,8699	-5,8750	86	-1,8025	-1,7384
41	-5,9235	-6,0510	87	-1,3173	-1,5467
42	-5,7365	-5,7511	88	-1,7660	-1,6072
43	-6,0338	-5,7326	89	-1,9026	-1,7268
44	-5,5840	-5,6935	90	-1,3462	-1,4799
45	-5,7159	-5,6550			

## Annexe 2 : Coefficients alpha avec ajustement polynomial

Age	Alpha log-Poisson	Age	Alpha log-Poisson
0	-7,4160	46	-5,4067
1	-7,4267	47	-5,3233
2	-7,4342	48	-5,2390
3	-7,4387	49	-5,1538
4	-7,4401	50	-5,0677
5	-7,4386	51	-4,9809
6	-7,4342	52	-4,8933
7	-7,4269	53	-4,8050
8	-7,4168	54	-4,7161
9	-7,4039	55	-4,6266
10	-7,3882	56	-4,5365
11	-7,3700	57	-4,4459
12	-7,3490	58	-4,3549
13	-7,3254	59	-4,2634
14	-7,2994	60	-4,1715
15	-7,2708	61	-4,0794
16	-7,2398	62	-3,9870
17	-7,2064	63	-3,8943
18	-7,1707	64	-3,8015
19	-7,1325	65	-3,7085
20	-7,0922	66	-3,6155
21	-7,0497	67	-3,5224
22	-7,0050	68	-3,4294
23	-6,9582	69	-3,3364
24	-6,9093	70	-3,2435
25	-6,8584	71	-3,1507
26	-6,8056	72	-3,0581
27	-6,7508	73	-2,9658
28	-6,6941	74	-2,8739
29	-6,6356	75	-2,7822
30	-6,5753	76	-2,6909
31	-6,5133	77	-2,6001
32	-6,4496	78	-2,5097
33	-6,3842	79	-2,4199
34	-6,3173	80	-2,3307
35	-6,2488	81	-2,2421
36	-6,1788	82	-2,1542
37	-6,1073	83	-2,0669
38	-6,0345	84	-1,9805
39	-5,9602	85	-1,8948
40	-5,8847	86	-1,8101
41	-5,8080	87	-1,7262
42	-5,7300	88	-1,6433
43	-5,6508	89	-1,5614
44	-5,5705	90	-1,4805
45	-5,4891		

### Annexe 3 : Coefficients béta sans ajustement polynomial

Age	Béta Lee-Carter	Béta log-Poisson	Age	Béta Lee-Carter	Béta log-Poisson
0	0,0175	0,0119	46	0,0026	0,0121
1	0,0529	0,0139	47	0,0296	0,0125
2	0,0772	0,0150	48	-0,0066	0,0120
3	0,0441	0,0163	49	0,0067	0,0118
4	0,0456	0,0172	50	-0,0164	0,0118
5	0,0343	0,0165	51	0,0232	0,0116
6	0,0128	0,0167	52	0,0129	0,0111
7	0,0408	0,0172	53	0,0370	0,0112
8	0,0196	0,0174	54	-0,0047	0,0106
9	0,0395	0,0167	55	0,0240	0,0107
10	0,0566	0,0172	56	0,0231	0,0101
11	0,0215	0,0177	57	-0,0004	0,0097
12	0,0357	0,0166	58	0,0175	0,0096
13	0,0351	0,0169	59	0,0228	0,0098
14	0,0276	0,0166	60	0,0169	0,0091
15	0,0369	0,0172	61	0,0100	0,0092
16	0,0297	0,0159	62	0,0077	0,0088
17	-0,0049	0,0159	63	0,0148	0,0086
18	0,0466	0,0148	64	0,0075	0,0088
19	0,0129	0,0148	65	-0,0130	0,0084
20	0,0101	0,0141	66	0,0196	0,0083
21	0,0348	0,0139	67	0,0193	0,0083
22	-0,0022	0,0135	68	-0,0064	0,0079
23	0,0606	0,0142	69	0,0087	0,0079
24	0,0015	0,0134	70	0,0057	0,0077
25	0,0105	0,0137	71	0,0156	0,0076
26	0,0208	0,0138	72	0,0079	0,0074
27	-0,0435	0,0135	73	-0,0074	0,0072
28	-0,0142	0,0144	74	0,0068	0,0074
29	-0,0349	0,0140	75	0,0130	0,0071
30	-0,0244	0,0142	76	0,0039	0,0068
31	0,0037	0,0134	77	0,0233	0,0068
32	0,0165	0,0138	78	-0,0103	0,0068
33	0,0616	0,0139	79	-0,0044	0,0062
34	-0,0208	0,0134	80	0,0054	0,0062
35	-0,0243	0,0134	81	0,0005	0,0060
36	0,0030	0,0134	82	0,0185	0,0061
37	0,0127	0,0133	83	-0,0155	0,0061
38	0,0232	0,0138	84	0,0113	0,0058
39	0,0005	0,0133	85	-0,0273	0,0059
40	0,0107	0,0130	86	-0,0070	0,0053
41	0,0224	0,0133	87	0,0111	0,0049
42	0,0099	0,0127	88	-0,0303	0,0050
43	-0,0300	0,0126	89	-0,0487	0,0053
44	0,0209	0,0127	90	0,0166	0,0048
45	-0,0062	0,0125			



## Annexe 4 : Coefficients béta avec ajustement polynomial

Age	Béta log-Poisson	Age	Béta log-Poisson
0	0,0158	46	0,0119
1	0,0158	47	0,0118
2	0,0158	48	0,0116
3	0,0159	49	0,0114
4	0,0159	50	0,0112
5	0,0159	51	0,0110
6	0,0159	52	0,0109
7	0,0159	53	0,0107
8	0,0159	54	0,0106
9	0,0159	55	0,0104
10	0,0159	56	0,0102
11	0,0158	57	0,0101
12	0,0158	58	0,0099
13	0,0158	59	0,0097
14	0,0156	60	0,0096
15	0,0156	61	0,0093
16	0,0155	62	0,0091
17	0,0155	63	0,0090
18	0,0154	64	0,0088
19	0,0153	65	0,0086
20	0,0153	66	0,0085
21	0,0152	67	0,0083
22	0,0151	68	0,0082
23	0,0150	69	0,0080
24	0,0149	70	0,0078
25	0,0148	71	0,0077
26	0,0147	72	0,0075
27	0,0146	73	0,0074
28	0,0145	74	0,0071
29	0,0144	75	0,0070
30	0,0143	76	0,0068
31	0,0142	77	0,0067
32	0,0140	78	0,0065
33	0,0139	79	0,0064
34	0,0138	80	0,0063
35	0,0137	81	0,0061
36	0,0134	82	0,0060
37	0,0133	83	0,0059
38	0,0131	84	0,0057
39	0,0130	85	0,0056
40	0,0129	86	0,0055
41	0,0127	87	0,0054
42	0,0126	88	0,0053
43	0,0124	89	0,0051
44	0,0123	90	0,0050
45	0,0121		

## Annexe 5 : Coefficients k sans ajustement polynomial

Année	k Lee-Carter	k log-Poisson
1989	14,9710	18,5550
1990	6,0235	8,6416
1991	15,6708	19,5390
1992	7,3075	7,8371
1993	7,8139	8,0568
1994	8,3748	10,0813
1995	6,9580	3,4504
1996	-8,6193	-7,6407
1997	2,9064	2,2809
1998	-10,6002	-11,0461
1999	-3,1330	-7,3718
2000	-6,5783	-11,2988
2001	-11,8924	-14,3982
2002	-7,8104	-7,3367
2003	-21,3921	-19,3495

## Annexe 6 : Coefficients k avec ajustement polynomial

Année	k log-Poisson
1989	17,5174
1990	15,0276
1991	12,5378
1992	10,0481
1993	7,5583
1994	5,0686
1995	2,5788
1996	0,0890
1997	-2,4007
1998	-4,8905
1999	-7,3802
2000	-9,8700
2001	-12,3598
2002	-14,8495
2003	-17,3393

# Bibliographie

BROUHNS N., DELWARDE A., DENUIT M. [2002]. *Elaboration de tables de mortalité prospectives. Consulting report n°02-03*. Institut de Statistique, Université catholique de Louvain, Louvain-la-Neuve, Belgique.

COALE A., KISKER E.E. [1990]. *Modèles d'extrapolation de la mortalité aux grands âges*. Institut des Sciences Actuarielles et Institut de Statistique Université Catholique de Louvain, Louvain-la-Neuve, Belgique.

COALE A., KISKER E.E. [1990]. *Defects in data on old age mortality in the United States: New procedures for calculating approximately accurate mortality schedules and life tables at the highest ages*. Asian and Pacific Population Forum 4,-31.

DENUIT M., BROUHNS N., VERMUNT J. [2002]. *A poisson log-bilinear regression approach to the construction of projected lifetables*. Insurance: Mathematics and Economics, vol.31, 373-393.

GUTTERMAN S., VANDERHOOFT I.T. [1999]. *Forecasting changes in mortality: a search for a law of causes and effects*. North American Actuarial Journal 2, 135-138.

LEE R.D.[2000]. *The Lee-Carter method of forecasting mortality, with various extensions and applications*. North American Actuarial Journal.

LEE R.D., CARTER L. [1992]. *Modelling and forecasting mortality, with various extensions and applications*. Journal of the American Statistical Association 87, 659-671.

MESLE F., VALLIN J. [2002]. *Comment améliorer la précision des tables de mortalité aux grands âges ? Le cas de la France*. INED, Population n°4, page 603.

PETAUTON P. [1996]. *Théorie et pratique de l'assurance vie*. DUNOD.

PLANCHET F. [2005]. *Tables de mortalité, construction de tables de mortalité d'expérience*. Support de cours, modèles de durée, Institut de Sciences Financières et d'Assurances Lyon.

SAPORTA G. [1990]. *Probabilités, analyse des données et statistique*. Paris: Editions Technip.

CORFIAS T. [2003]. *Assurance vie : techniques et produits*. Editions l'Argus de l'Assurance.

LELIEUR V. [2005]. *Utilisation des méthodes de Lee-Carter et Log-Poisson pour l'ajustement de tables de mortalité dans le cas de petits échantillons*. ISFA, Mémoire d'Actuariat.

# Table des figures

Fig. (1) : Effectifs soumis au risque et décès observés de 1989 à 2003.....	17
Fig. (2) : Répartition moyenne par âge des effectifs soumis au risque .....	17
Fig. (3) : Taux bruts de mortalité de 1989 à 2003.....	19
Fig. (4) : Comparaison des taux bruts de mortalité de 1989, 1996 et 2003.....	20
Fig. (5) : Evolution de l'espérance de vie à la naissance de 1989 à 2003 .....	21
Fig. (6) : Evolution de l'espérance de vie à 60 ans de 1989 à 2003 .....	22
Fig. (7) : Estimation des paramètres $\alpha_x$ par la modèle de Lee-Carter .....	27
Fig. (8) : Estimation des paramètres $\beta_x$ par la modèle de Lee-Carter .....	27
Fig. (9) : Estimation des paramètres $k_t$ par le modèle de Lee-Carter .....	28
Fig. (10) : Taux de mortalité lissés par le modèle de Lee-Carter.....	29
Fig. (11) : Taux de mortalité lissés par le modèle de Lee-Carter pour l'année 1989 .....	29
Fig. (12) : Taux de mortalité lissés par le modèle de Lee-Carter pour l'année 2003 .....	30
Fig. (13) : Intervalles de confiance asymptotiques pour les taux bruts en 1989 et taux lissés par le modèle de Lee-Carter .....	31
Fig. (14) : Intervalles de confiance asymptotiques pour les taux bruts en 1996 et taux lissés par le modèle de Lee-Carter .....	32
Fig. (15) : Intervalles de confiance asymptotiques pour les taux bruts en 2003 et taux lissés par le modèle de Lee-Carter .....	32
Fig. (16) : Analyse des résidus dans le modèle de Lee-Carter .....	34
Fig. (17) : Comparaison des paramètres $\alpha$ obtenus par les modèles de Lee-Carter et log-Poisson.....	37
Fig. (18) : Comparaison des paramètres $\beta_x$ obtenus par les modèles de Lee-Carter et log-Poisson.....	37
Fig. (19) : Comparaison des paramètres $k$ obtenus par les modèles de Lee-Carter et log-Poisson.....	38
Fig. (20) : Taux de mortalité lissés par le modèle log-Poisson .....	38
Fig. (21) : Comparaison des taux lissés par les modèles de Lee-Carter et log-Poisson pour l'année 1990.....	39
Fig. (22) : Comparaison des taux lissés par les modèles de Lee-Carter et log-Poisson pour l'année 2000.....	40
Fig. (23) : Approximation polynomiale des paramètres $\alpha$ obtenus par le modèle log-Poisson .....	41
Fig. (24) : Approximation polynomiale des paramètres $\beta$ obtenus par le modèle log-Poisson .....	41
Fig. (25) : Approximation linéaire des paramètres $k$ obtenus par le modèle log-Poisson .	42
Fig. (26) : Taux de mortalité lissés par le modèle log-Poisson avec approximation polynomiale des paramètres .....	42
Fig. (27) : Comparaison pour l'année 1989 des risques de décéder dans la décennie issus des données brutes et de l'ajustement de log-Poisson.....	43
Fig. (28) : Comparaison pour l'année 2003 des risques de décéder dans la décennie issus des données brutes et de l'ajustement de log-Poisson.....	43
Fig. (29) : Intervalles de confiance asymptotiques pour les taux bruts en 1989 et taux lissés par le modèle log-Poisson.....	45

Fig. (30) : Intervalles de confiance asymptotiques pour les taux bruts en 1996 et taux lissés par le modèle log-Poisson.....	46
Fig. (31) : Intervalles de confiance asymptotiques pour les taux bruts en 2003 et taux lissés par le modèle log-Poisson.....	46
Fig. (32) : Extrapolation aux grands âges du paramètre $\alpha$ de structure polynomiale .....	47
Fig. (33) : Extrapolation aux grands âges du paramètre $\beta$ de structure polynomiale .....	48
Fig. (34) : Comparaison des taux de mortalité 1990 aux âges avancés obtenus par extrapolation polynomiale avec ceux obtenus par l'extrapolation de Coale-Kisker. ....	48
Fig. (35): Comparaison des taux de mortalité 2003 aux âges avancés obtenus par extrapolation polynomiale avec ceux obtenus par l'extrapolation de Coale-Kisker. ....	49
Fig. (36) : Statistiques du Chi-2 réalisées sur les données brutes et l'ajustement par le modèle log-Poisson (sans ajustement polynomial des paramètres). ....	51
Fig. (37): Statistiques du Chi-2 réalisées sur les données brutes et l'ajustement par la méthode log-Poisson (avec et sans ajustement polynomial des paramètres).....	52
Fig. (38): Statistiques du Chi-2 réalisées sur les données brutes et l'ajustement par la méthode log-Poisson (avec et sans ajustement polynomial des paramètres).....	52
Fig. (39): Statistiques du Chi-2 sur l'ajustement par le modèle de Lee-Carter (sans ajustement polynomial des paramètres) comparées aux statistiques obtenues à partir du modèle log-Poisson.....	53
Fig. (40) : Comparaison des capitaux constitutifs obtenus avec les taux bruts et les taux lissés respectivement par les modèles de Lee-Carter et log-Poisson .....	54
Fig. (41) : Extrapolation de la composante temporelle du modèle log-Poisson. ....	56
Fig. (42) : Projection des taux de mortalité instantanés futurs par le modèle log-Poisson	57
Fig. (43): Rapport entre les taux instantanés 2010 et les taux instantanés 2000, lissés par le modèle log-Poisson.....	58
Fig. (44) : Projection de l'espérance de vie future à 60 ans par le modèle Log-Poisson. ....	59