



**Mémoire présenté  
devant l'Institut de Science Financière et d'Assurances  
pour l'obtention  
du diplôme d'Actuaire de l'Université de Lyon**

le \_\_\_\_\_

Par : Marc Juillard

Titre : Identification et quantification des risques techniques pour un régime de rentes

Confidentialité : \_\_\_\_\_

*Membre du jury I.A.*

*Entreprise : Lynxial*

*Membres du jury I.S.F.A.*

*Directeur de mémoire :*

M. Frédéric Planchet

*Invité :*

*Secrétariat :*

Mme GARCIA Marie-José

Mme BARTHELEMY Diane

Mme BRUNET Marie-Christine

Mme MOUCHON Marie-Claude

*Bibliothèque :*

Mme SONNIER Michèle

## Résumé

Pour être en accord avec les nouvelles normes comptables et les futures réglementations de solvabilité, il devient de plus en plus important d'identifier toutes les sources d'aléa dans les contrats d'assurance. En ce qui concerne les régimes de rentes, le risque systématique de mortalité mis en évidence depuis quelques années nécessite des investigations.

Le modèle choisi est celui de Lee-Carter (1992). Ce modèle, régulièrement utilisé pour la création de tables prospectives, permet de séparer la composante « temporelle » et la composante « âge » des taux de mortalité. Ainsi, après calage de la composante « âge » sur les données du passé, on modélise de manière aléatoire la composante temporelle. Cette dernière permettra de prendre en compte le caractère aléatoire de l'évolution de la mortalité dans le futur.

La modélisation de la dérive temporelle est effectuée alternativement sous deux hypothèses :

- Les oscillations aléatoires de la mortalité au cours du temps, sont indépendantes entre elles et peuvent donc être modélisées à l'aide d'une régression linéaire ;
- Les oscillations aléatoires de la mortalité au cours du temps, sont liées entre elles par une relation temporelle et doivent donc être modélisées par une série temporelle.

L'analyse de la distribution de l'engagement dans les deux modèles tend à montrer que le risque auquel est soumis un régime de rentes est plutôt un risque de modèle, par nature peu "capturable" par la seule analyse des évolutions passées.

Mots clés : Tables prospectives, extrapolation, ajustement, rentes, mortalité stochastique.

## Abstract

In order to respect the new accounting standards and the futur regulations of solvency, it is getting more and more important to identify all the sources of unknown quantities in insurance policies.

Concerning annuity plans, the systematic risk of mortality underlined for few years requires investigations.

We chose to use the model of Lee-Carter (1992). This model, used for creating stochastic prospective tables, make it possible to separate the “temporal” component and the « age » component of mortality rates. Once the « age » component adjusted on data of the past, we model in a random way the « temporal » component. The temporal component permit to take into account the random nature of the evolution of the mortality in the future.

The modelling of the temporal drift is made with two hypothesis:

Random oscillations of the mortality during time are independent and could be modelled by a linear regression.

Random oscillations of the mortality during time are linked with a temporal relation and have to be modelled by a temporal series.

The analysis of the commitment distribution in the two models tends to show that the risk which is submit an annuity plan is rather a risk of model, by nature difficult to catch by the analysis of the past evolutions.

KEYWORDS: Prospective tables, extrapolation, adjustment, life annuities, stochastic mortality.

## Sommaire

<b>Introduction</b>	<b>6</b>
---------------------	----------

<b>Présentation du modèle</b>	<b>11</b>
-------------------------------	-----------

<b>1. Problématique</b>	<b>11</b>
1.1 Représentation de la durée de vie d'un individu	11
1.2 Notations	14
1.3 Hypothèse de répartition des décès	15
1.4 Caractéristiques des données	17
1.5 Engagement	18
<b>2. Les différents types de Modèles</b>	<b>19</b>
2.1 Modèles d'inspiration financière	19
2.1.1 Modèle stochastique affine	19
2.1.2 Cadres des modèles	21
2.1.3 Le processus d'Orstein-Uhlenbeck	22
2.1.4 Le processus de Feller	23
2.2 Modèle de type financier	24
2.3 Modèles d'inspiration démographique	26
2.3.1 Le modèle de Lee Carter	26
2.3.1.1 Présentation du modèle	26
2.3.1.2 Estimation des paramètres	27
2.3.1.3 Extrapolation de la composante temporelle	30
2.3.2 Le modèle de Poisson	30
<b>3. Méthode de simulation</b>	<b>32</b>
3.1 Algorithme principal	32
3.2 Modélisation de la sortie des individus	34
3.3 Générateur aléatoire	35
3.4.1 Générateur de lois uniformes	35
3.4.2 Générateur de lois normales	36

<b>Construction du modèle</b>	<b>38</b>
-------------------------------	-----------

<b>I Etude du risque systématique à l'aide d'une régression linéaire</b>	<b>43</b>
--	-----------

<b>1. Cas d'une mortalité déterministe</b>	<b>43</b>
1.1 Ajustement du modèle	43
1.2 Applications numériques	45
<b>2. Cas d'une mortalité stochastique</b>	<b>47</b>
2.1 Modèle de mortalité	47
2.2 Analyse de la répartition du risque pour le régime de rentes	49
2.3 Applications numériques	50
2.3.1 Analyse de l'engagement	50
2.3.2 Importance de la correction de biais	54
2.3.3 Analyse de la part de variance expliquée	55
<b>3. Mesure du risque d'erreur de dérive</b>	<b>60</b>
3.1 Le modèle de mortalité	61
3.2 Applications numériques	62
3.2.1 Application à un régime de rentiers	66
3.2.2 Application du modèle à la TPG 1993	71

<b>II</b>	<b><i>Etude du Risque systématique à l'aide de séries temporelles</i></b>	<b>73</b>
<b>1.</b>	<b><i>Ajustement du modèle.</i></b>	<b>73</b>
1.1	Définition de la non stationnarité	73
1.1.1	Non stationnarité déterministe	74
1.1.2	Non stationnarité stochastique.	76
1.1.3	Etude de la non stationnarité.	77
1.1.4	Estimation des paramètres du modèle.	78
<b>2.</b>	<b><i>Applications numériques</i></b>	<b>85</b>
<b>III</b>	<b><i>Impact de l'ancienneté maximale de l'historique sur l'engagement</i></b>	<b>91</b>
	<b><i>Conclusion</i></b>	<b>97</b>
	<b><i>Bibliographie</i></b>	<b>99</b>
	<b><i>Annexe</i></b>	<b>101</b>

## Introduction

Le contrat de rente viagère suppose le paiement d'une rente au bénéficiaire jusqu'au moment de son décès. L'augmentation du nombre de retraités s'accompagnant d'une diminution du nombre d'actifs, les régimes de retraite deviennent de moins en moins avantageux. Le principe de mutualisation rendant la retraite que recevra le salarié lors de son départ à la retraite dépendante du rapport actifs sur retraités, il ne peut estimer de manière précise le montant qu'il recevra. Ainsi, investir dans un contrat de rente viagère pour s'assurer un revenu minimum est une pratique de plus en plus courante.

D'une manière générale, un régime de rente est soumise à trois risques. En premier lieu, le taux d'intérêt technique qui permet d'actualiser les flux financiers générés par le contrat. Ensuite une table de mortalité décrivant la loi de survenance des décès, laquelle permet d'évaluer le coût moyen des contrats souscrits par la compagnie. Enfin, les taux de chargement dont le but est de couvrir les divers frais créés par les contrats.

Le principe de mutualisation nous incite à penser, que le seul risque portant sur la tarification, et sur le provisionnement des contrats de rente, est le risque financier.

Cependant depuis 1946 l'espérance en vie est en constante augmentation. On constate ainsi en France depuis une trentaine d'années une augmentation de l'espérance de vie d'environ un trimestre par an et le seuil des 80 ans a été franchi en 2006.

Depuis une quinzaine d'années, les tables de mortalité utilisées pour le provisionnement et la tarification des rentes viagères s'attachent à prendre en compte la dérive tendancielle de la mortalité, qui se traduit par une croissance régulière de l'espérance de vie.

La nécessité d'utiliser des tables de mortalité prospectives intégrant ce phénomène de dérive a été prise en compte par le législateur et des tables « de génération » ont été homologuées il y a une quinzaine d'années. 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).

De nouvelles tables générationnelles ont été créées en juillet 2006. Séparant la mortalité masculine (TGH 05) de la mortalité féminine (TGF 05), elles ont été établies sur la base de la population française couvrant la période 1993 – 2005 du point de vue des données et intègre la structure de la mortalité couvrant la période 1962 -2000.

Cette actualisation a mis en lumière l'insuffisance des provisions mathématiques établies sur la base de la TPG 93. Or en tant que table générationnelle, la TPG 93 présente une vision des taux futurs, laquelle a été ajusté en tenant compte de l'évolution de la mortalité.

Cette insuffisance des provisions établies vient donc de cette évolution estimée à partir de la vision de la mortalité que les actuaires avaient en 1993, et qui, vue en 2006 n'est plus exacte.

Estimer correctement la dérive de la mortalité semble donc être un problème.

Cependant, à la vue de l'évolution de l'espérance de vie à 60 ans, il semblerait que cette dernière se fasse de manière linéaire et très peu volatile à partir de 1985. Il serait donc facile d'appréhender l'évolution future de la mortalité, et les modèles de construction de tables prospectives pourraient alors tenir compte facilement de la dérive de la mortalité.



Fig. 1 : *Evolution de l'espérance de vie à 60 ans en France*

Mais des conclusions différentes ont été obtenues sur les taux instantanés de mortalité.

Des études récentes (CURRIE et al. [2004]) font apparaître que l'évolution du taux instantané de mortalité présente, aux différents âges, des variations erratiques autour de la tendance qui se dégage, variations non expliquées par les fluctuations d'échantillonnage. CAIRNS et al. [2004] présentent une analyse détaillée de ce phénomène. Sur les données INED utilisées dans le présent travail (données reprises de MESLE et VALLIN [2002]) on constate le même phénomène, illustré par la figure 2.

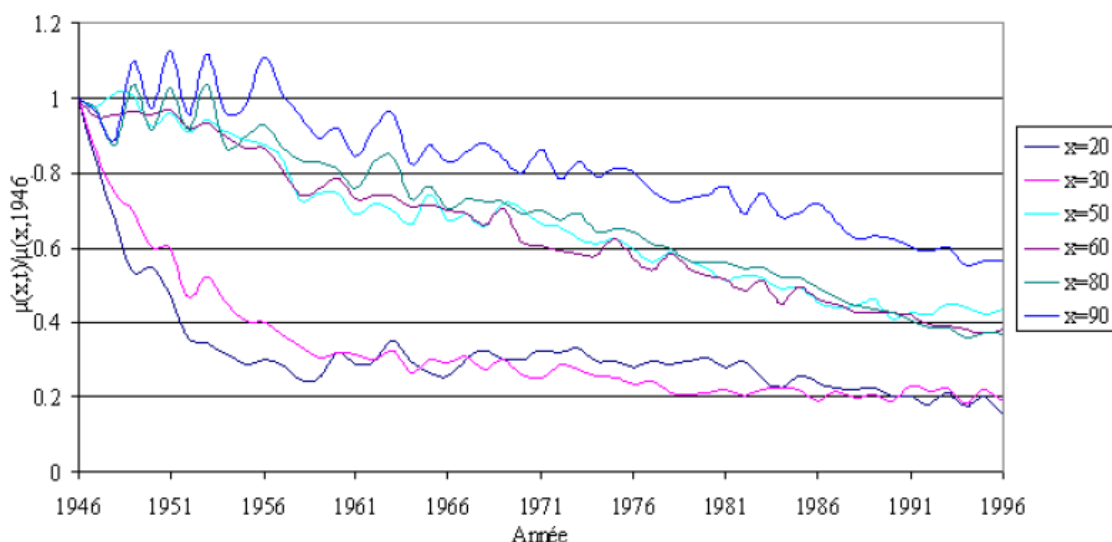


Fig. 2 : *Fluctuations des taux de décès autour de la tendance de long terme*

De même, les nouvelles tables générationnelles (TGF 05 et TGH 05) mettent à mal l'idée selon laquelle prédire l'évolution de la mortalité est un mécanisme aisé.

Ainsi, en France, la récente actualisation des tables utilisées par les assureurs pour le provisionnement des rentes viagères illustre les difficultés d'une telle anticipation et les enjeux financiers associés. Par rapport aux tables TPG<sup>1</sup> 1993 en vigueur jusqu'au 31/12/2006, les nouvelles tables TGH 05 et TGF 05 qui entrent en vigueur le 01/01/2007 conduisent à des majorations des provisions parfois supérieures à 20 %, comme l'illustre le tableau suivant :

Age	Génération	TPG 1993 Femmes	TPG 1993 Hommes	Femmes / TPG	Hommes / TPG
50	1955	26,81647	28,40552	5,9 %	26,75507
55	1950	24,26368	25,95575	7,0 %	24,07474
60	1945	21,50832	23,30185	8,3 %	21,25828
65	1940	18,53412	20,39677	10,0 %	18,22126
70	1935	15,39467	17,28922	12,3 %	15,08772
75	1930	12,25679	14,08680	14,9 %	12,05698
80	1925	9,35194	10,96271	17,2 %	9,12890
85	1920	6,88306	8,15548	18,5 %	6,64827
90	1915	4,93310	5,89309	19,5 %	4,73880
95	1910	3,46780	4,29408	23,8 %	3,40109

Tab:1 Comparaison des coefficients de provisionnement TPG 1993 et TGH/TGF 05

<sup>1</sup> Tables obtenues sur la base de la mortalité de la population féminine sur la période 1961-1987, utilisées depuis le 1<sup>er</sup> juillet 1993.



Les variations de la dérive de la mortalité affectant de manière systématique l'ensemble des individus, elles ne se mutualisent pas. Elles font donc courir un risque potentiellement important à un régime de rentes, dont l'équilibre technique est construit sur la mutualisation du risque de survie de ses membres.

Il est à noter que les sociétés d'assurance sont soumises à de nouvelles contraintes réglementaires.

- L'entrée en vigueur des nouvelles normes comptables IFRS oblige les entreprises à provisionner leurs régimes supplémentaires d'entreprises.
- L'entrée en vigueur du projet Solvabilité II oblige les assureurs à modifier leur technique d'estimation des provisions techniques. Les sociétés d'assurance doivent prendre en compte l'ensemble des risques portant sur leur portefeuille de contrats.

Les sociétés vont donc être obligées de quantifier l'impact de la dérive stochastique de la mortalité sur leurs provisions mathématiques afin d'être en accord avec les nouvelles normes prudentielles. L'utilisation de tables de mortalité déterministes ne permettant pas d'appréhender le phénomène de dérive aléatoire de la mortalité, il sera alors nécessaire de construire des tables de mortalité stochastiques.

On est donc conduit à rechercher une modélisation capable de rendre compte de ces fluctuations autour de la valeur tendancielle et d'en tirer des conséquences sur le niveau des provisions mathématiques que doit constituer le régime pour assurer son équilibre technique. De manière plus précise, il importe de quantifier la part, dans le risque total auquel est exposé le régime (selon une mesure à définir), de la composante non mutualisable.

Les modèles stochastiques de mortalité fournissent un outil bien adapté à cette analyse. Ils proposent en effet de considérer que le taux de mortalité futur  $\mu(x,t)$  est lui-même aléatoire, et que donc  $\mu(x,t)$  est un processus stochastique (comme fonction de  $t$  à  $x$  fixé). Le taux de mortalité effectivement observé pour un âge et une année donnés est alors la réalisation d'une variable aléatoire : on peut noter l'analogie avec les méthodes de lissage bayésiennes (sur ces méthodes de lissage, on pourra consulter TAYLOR [1992]).

Les modèles récents de construction de tables prospectives, comme le modèle de Lee-Carter (voir notamment LEE et CARTER [1992], LEE [2000], SITHOLE et al. [2000]) ou les modèles poissoniens (*cf.* BROUHNS et al. [2002] et PLANCHET et THEROND [2006] pour une présentation et une discussion de ces modèles), sont également des cas particuliers de modèles

stochastiques, bien qu'ils soient à l'origine élaborés pour construire des extrapolations (temporelles) de la surface  $\mu(x, t)$  déterministe. Après un ajustement des taux passés, les taux de mortalité pour les années futures se déduisent directement de l'extrapolation de la composante temporelle (paramétrique ou non paramétrique) du modèle prospectif retenu (on peut évidemment critiquer cette approche purement extrapolative ; on pourra par exemple consulter GUTTERMAN et VANDERHOOF [1999] sur ces questions).

On retient dans cette étude le modèle de Lee-Carter, qui permet aisément de produire des surfaces de mortalité stochastiques et qui est, par ailleurs, devenu un standard pour la construction de tables prospectives. La variante log-Poisson (*cf.* BROUHNS et al. [2002]) conduit à des résultats très proches.

Après avoir construit un jeu de tables prospectives sur des données nationales à l'aide de ce modèle, nous l'utilisons pour calibrer la composante aléatoire pour le risque systématique et appliquons le modèle ainsi obtenu pour déterminer le niveau de l'engagement d'un régime de rentes.

Les données utilisées pour les applications numériques sont reprises de PLANCHET et THEROND [2004].

# Présentation du modèle

## 1. PROBLEMATIQUE

### 1.1 REPRESENTATION DE LA DUREE DE VIE D'UN INDIVIDU

La mortalité d'un individu dépend de trois paramètres ; son âge, l'instant d'observation et la génération à laquelle l'individu appartient. Si ces trois paramètres peuvent se résumer à l'aide de deux d'entre eux, chacun possède son importance dans la détermination du niveau de mortalité :

- L'âge est évidemment le paramètre le plus important.
- La date permet d'appréhender les courtes, mais brutales, modifications de l'indice de mortalité. Ainsi durant l'année 2003, la canicule a augmenté les probabilités de décès aux âges élevés.
- La génération permet d'appréhender les progrès de la médecine mais également les événements qui ont durement touché certaines générations. Ainsi, les probabilités de décès d'une génération affectée par une épidémie dans le passé seront moindres que celles des générations non touchées. En effet, l'épidémie aura opéré une sélection importante, ne laissant en vie que les individus les plus forts.

La mortalité dépend donc de trois coordonnées. Le diagramme de Lexis, du nom du statisticien démographe allemand Wilhelm LEXIS (1837-1914), permet d'appréhender ce phénomène.

Il s'agit d'un graphique qui permet de caractériser une cohorte à partir de 3 variables démographiques sur un plan à deux dimensions. Ces trois dimensions sont l'âge, le moment de naissance et une date quelconque, par exemple la date de décès ou celle du mariage.

Dans sa version la plus habituelle, l'axe des abscisses supporte le temps (calendrier) et celui des ordonnées, l'âge. Ces deux axes dessinent un diagramme cartésien des plus classiques. Toutefois, sur ce diagramme, une troisième coordonnée s'utilise également, à savoir le moment de naissance.

Si les coordonnées en temps (abscisse) et en âge (ordonnée) se traduisent de manière tout à fait habituelle par des perpendiculaires à leurs axes respectifs, la coordonnée du moment de naissance va engendrer des diagonales, chacune étant une ligne de vie qui montre l'évolution de l'âge en fonction du temps pour une date de naissance donnée.

Pour un emploi aisé, le diagramme de Lexis est complété par des réseaux de droites : les horizontales qui délimitent les âges révolus ; les verticales qui délimitent les années et les obliques qui délimitent les générations. Cette façon de le présenter peut se dénommer "diagramme muet" dans le sens où il est prêt à recevoir les données à localiser.

La vie d'un individu  $y$  est représentée par une ligne parallèle à la première bissectrice et s'arrêtant au jour du décès, appelé "point mortuaire".

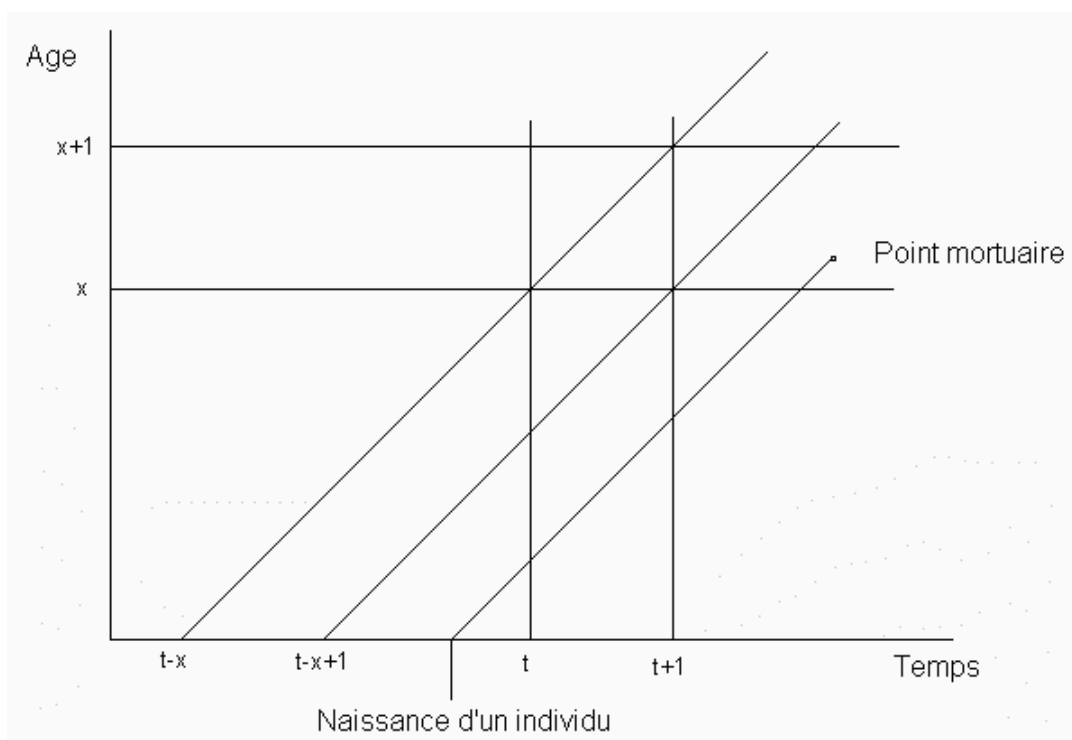


Fig. 3 : *Diagramme de Lexis*

Pour modéliser la mortalité, les assureurs ont recours aux tables de mortalité réglementaires. Il en existe deux types, les tables du moment et les tables générationnelles. Pour bien visualiser la différence existant entre ces deux tables, il suffit d'observer le diagramme de Lexis :

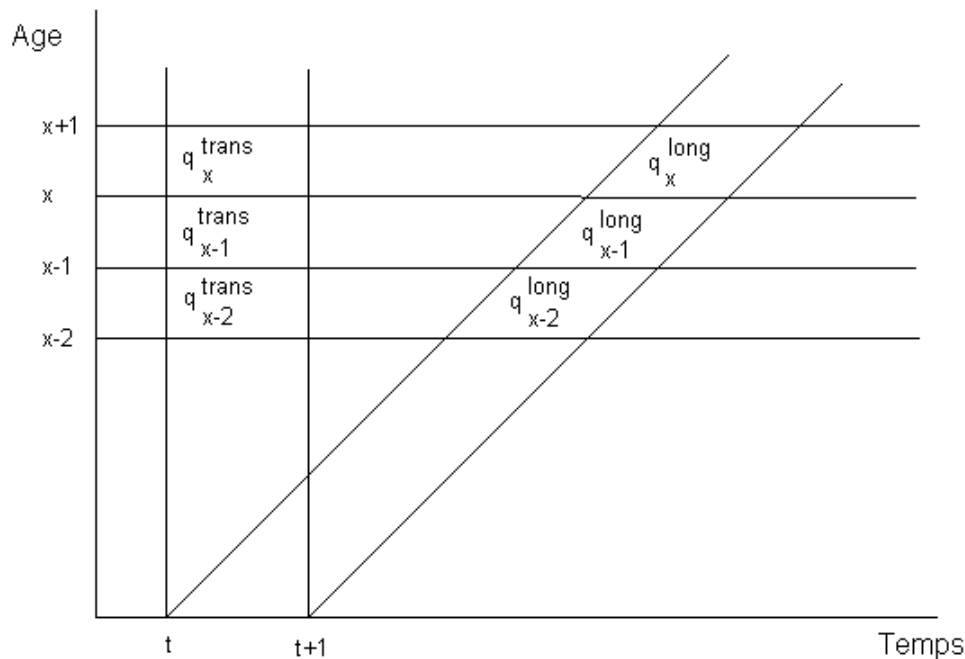


Fig. 4 : Table longitudinale et transversale

Les tables de mortalité classiques, ou tables du moment, sont des tables transversales, c'est à dire qu'elles ne représentent pas la mortalité d'une génération mais sont construites avec les observations d'une année fixée. Ces tables ne décrivent la mortalité d'aucun assuré du portefeuille dans la mesure où chaque quotient de mortalité a été obtenu avec des individus de générations différentes. Les tables prospectives, ou générationnelles, sont créées pour résoudre ce problème.

Les tables prospectives, seront nécessairement plus prudentes que les tables du moment. Ce sont les seules tables qui peuvent prendre en compte le phénomène de dérive de la mortalité, mais seulement d'une manière déterministe.

La construction de tables de mortalité prenant en compte la dérive de la mortalité ne peut donc se faire qu'à partir de tables générationnelles.

Après avoir fixé le type de table que nous allons utiliser, il convient de retenir un modèle afin d'analyser puis de modéliser l'impact de la mortalité stochastique sur le portefeuille de rentiers que nous étudions.

Dans un premier temps, nous allons d'écrire quelques modèles d'inspiration financière puis, nous analyserons le modèle d'inspiration démographique que nous avons retenu : le modèle de Lee-Carter.

## 1.2 NOTATIONS

Nous allons décrire dans cette partie l'ensemble des notations qui seront utilisées par la suite.

Considérons tout d'abord la variable aléatoire  $T$  représentant la durée de vie d'un individu. Il est alors naturel de s'intéresser à la variable aléatoire  $T_x$  qui est identique à  $T$  conditionnellement au fait que l'individu est vivant à l'âge  $x$ .

On a alors :  $T_x =_d [T-x|T>x]$ .

Grâce à cette définition, il nous est alors possible d'introduire la probabilité de survie pendant  $t$  années pour un individu d'âge  $x$  :

$${}_t p_x = P(T_x > t) = P(T > x+t | T > x), \quad (1)$$

On obtient alors la probabilité de décès dans les  $t$  prochaines années pour un individu d'âge  $x$  en posant simplement :

$${}_t q_x = 1 - {}_t p_x = P(T_x \leq t) = P(T \leq t+x | T > x), \quad (2)$$

Dans la suite, nous noterons  $q_x = {}_1 q_x$  et  $p_x = {}_1 p_x$  respectivement la probabilité de mourir dans l'année et la probabilité de survivre durant l'année.

En introduisant la fonction de survie  $S$  de la variable  $T$  définie par  $S(x) = P(T > x)$ , on obtient la relation :

$${}_t p_x = \frac{S(x+t)}{S(x)}, \quad (3)$$

D'une manière générale, les modèles font appel à la fonction de hasard  $\mu$ . Dans le cadre de la mortalité elle est appelée taux de décès instantané et est définie par :

$$\mu_{x+t} = \lim_{h \rightarrow 0} \frac{P(t < T_x \leq t+h | T_x > t)}{h} = \lim_{h \rightarrow 0} \frac{P(t < T_x \leq t+h)}{h P(T_x > t)}, \quad (4)$$

D'où l'on tire que :

$${}_t p_x = \exp\left(-\int_0^t \mu_{x+s} ds\right), \quad (5)$$

### 1.3 HYPOTHESE DE REPARTITION DES DECES

Les tables de mortalité indiquent les taux de mortalité à des âges entiers. Or, en pratique l'âge des rentiers au moment du calcul de la provision mathématique n'est pas entier. Il peut donc s'avérer nécessaire de calculer les taux de décès à n'importe quel âge. Il est pour cela indispensable de formuler une hypothèse qui permette de passer d'une expression discrète de la loi de mortalité à une expression continue. Il faut donc choisir de quelle manière se répartissent les décès entre deux dates anniversaires. Trois hypothèses sont classiquement utilisées :

- l'hypothèse de constance des taux instantanés de mortalité entre 2 âges entiers ;
- l'hypothèse d'uniforme répartition des décès entre deux âges, également appelée linéarisation de la fonction de survie (hypothèse de DUD) ;
- l'hypothèse de Balducci postulant pour tout âge entier  $x$  :  ${}_{1-t}q_{x+t} = (1-t)q_x$  (avec  $t \in [0;1[$ ).

On peut tout d'abord noter que l'hypothèse de Balducci peut être écartée dans la mesure où elle conduit à des taux instantanés de mortalité décroissants entre 2 âges entiers. En effet, on a

$${}_tq_x = 1 - {}_t p_x = 1 - \frac{p_x}{1-t p_{x+t}} = 1 - \frac{p_x}{1-t q_{x+t}} = \frac{t \cdot q_x}{p_x + t \cdot q_x}, \quad (6)$$

Donc  $\mu_{x+t} = -\frac{\delta}{\delta \cdot t} {}_t p_x = \frac{q_x}{p_x + t q_x}$ , ce qui va à l'encontre de ce qu'on peut espérer d'un taux instantané de mortalité, au moins aux âges élevés.

Examinons à présent les deux choix restants. En notant  $T_x^c$  et  $T_x^l$  les durées de vie résiduelles respectivement du modèle de constance des taux et du modèle de répartition linéaire. Notons  $t=k+r$  avec  $k=[t]$  et  $0 \leq r < 1$ , on a alors avec les notations habituelles :

$$S_x^l(t) = P(T_x^l > k+r) = {}_k p_x (1-r \cdot q_{x+k}), \quad (7)$$

Et

$$S_x^c(t) = P(T_x^c > k+r) = {}_k p_x \cdot p_{x+k}^r, \quad (8)$$

Ceci implique que :

$$S_x^l(t) \geq S_x^c(t) \Leftrightarrow 1 - r(1 - p_{x+k}) \geq p_{x+k}^r, \quad (9)$$

Or ce dernier point est clair puisque si  $r \in [0;1]$  on a  $a(1+x)^r \leq 1+rx$ . Il vient alors  $S_x^I(t) \geq S_x^C(t)$ . L'hypothèse de constance des taux conduit donc à des durées de vie inférieures : il s'agit alors d'une hypothèse prudente pour des garanties en cas de décès, mais moins prudente pour des garanties de type rentes viagères.

Dans la suite de l'étude, on considèrera l'hypothèse de constance des taux instantanés de mortalité entre deux âges entiers, de sorte que l'on a  ${}_t q_x = 1 - (1 - q_x)^t$ . On considère donc constant le taux instantané de mortalité dans chaque carré du diagramme de Lexis, comme représenté sur la figure 5.

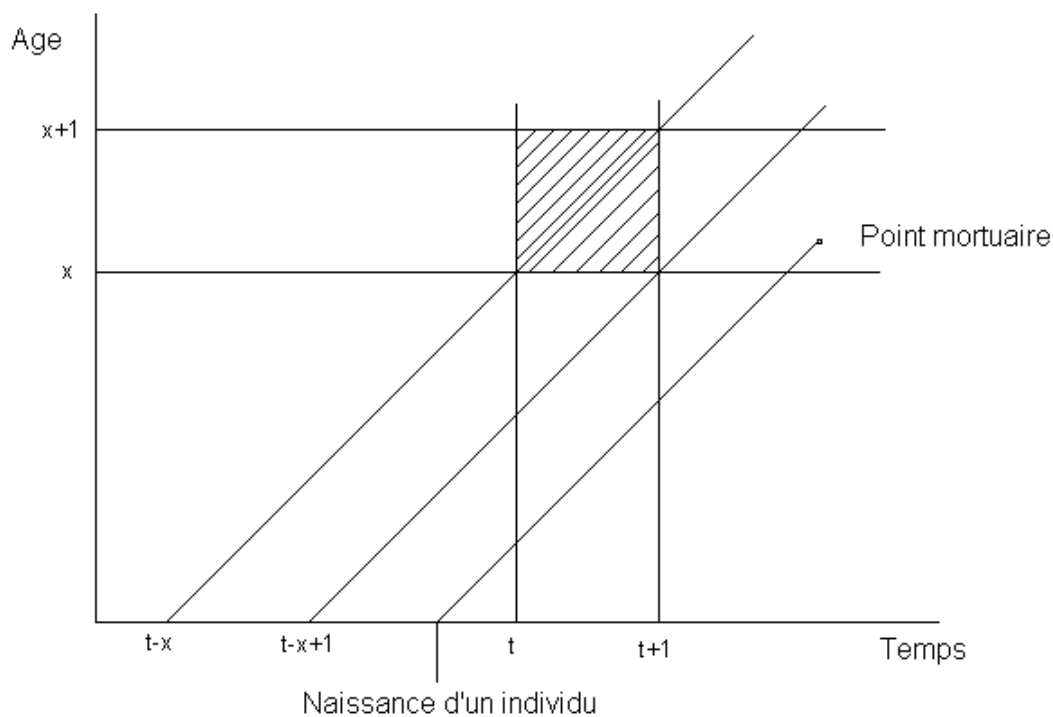


Fig. 5 : Constance des taux instantanés entre deux âges

Les données disponibles font souvent état des quotients de mortalité bruts, c'est à dire des probabilités de décès dans l'année, on rappelle que :

$$q_x = \frac{l_x - l_{x+1}}{l_x}, \quad (10)$$

Avec  $l_x$  l'effectif sous risque d'âge  $x$ . On a donc :

$$\hat{\mu}_x(t) = \ln(1 - \hat{q}_x(t)), \quad (11)$$



## 1.4 CARACTERISTIQUES DES DONNEES

Dans la suite, nous utiliserons pour les applications numériques un portefeuille constitué de 374 rentiers de sexe féminin, âgés en moyenne de 63,8 ans au 31/12/2005. La rente annuelle moyenne s'élève à 5,5 k€.

La figure 6 présente les flux de prestations espérées en fonction du temps obtenus à partir de la table de mortalité TV 2000.

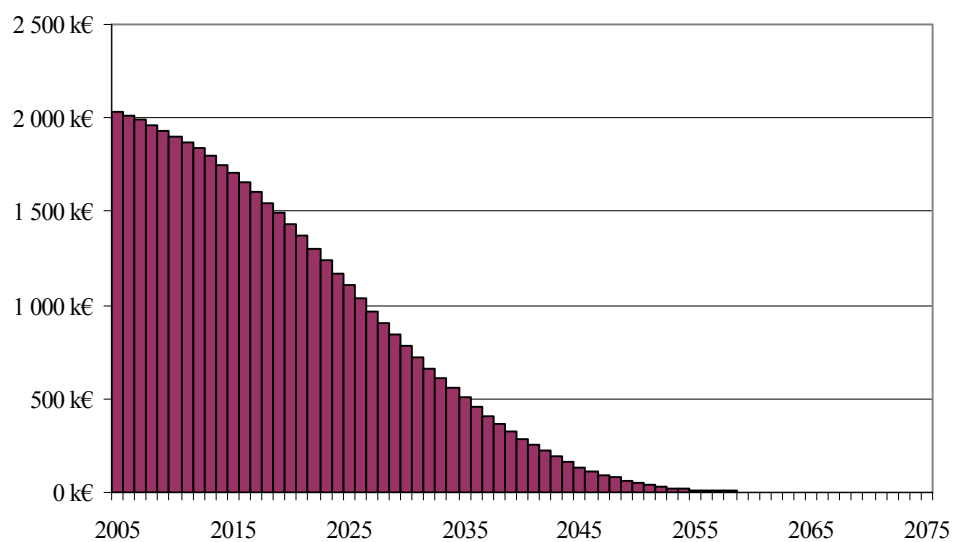


Fig. 6 : *Flux de prestations espérées*

Avec un taux d'escompte des provisions de 2,5 %, la provision mathématique initiale calculée avec la table prospective construite dans la deuxième partie, s'élève à 37,9 M€. On peut noter que dans PLANCHET et THEROND [2004], les calculs menés avec la table du moment féminine TV2000 conduisaient à une provision de 32,8 M€. Cela met incidemment en évidence l'importance d'une approche prospective de la mortalité pour déterminer le niveau de l'engagement sans le sous-estimer exagérément.

Nous noterons dans la suite de l'article :

- $L_0$  le montant des provisions mathématiques à la date initiale,
- $\tilde{F}_t$  le flux de prestation (aléatoire) à payer à la date  $t$ ,
- $(\Phi_t)$  la filtration décrivant l'information disponible à la date  $t$ ,
- $i$  le taux (discret) d'escompte des provisions mathématiques,
- $\mathbf{J}$  l'ensemble des individus,
- $x(j)$  l'âge en 0 de l'individu  $j$  et  $r_j$  le montant de sa rente annuelle.

## 1.5 ENGAGEMENT

Le portefeuille est exclusivement constitué de rentes en cours de service supposées non réversibles. En 0, l'assureur estime la suite de flux probables de sinistres que nous noterons  $(F_t)_{t \geq 1}$  :

$$F_t = \mathbf{E}[\tilde{F}_t \mid \Phi_0], \quad (12)$$

avec :

$$\tilde{F}_t = \sum_{j \in \mathcal{J}} r_j * \mathbf{1}_{]t; \infty[}(T_{x(j)}), \quad (13)$$

où  $T_{x(j)}$  désigne la date de décès (aléatoire) de la tête d'âge  $x(j)$ . Avec les notations classiques de l'assurance vie (cf. PETAUTON [1996]) :

$$F_t = \sum_{j \in \mathcal{J}} r_j * \frac{l_{x(j)+t}}{l_{x(j)}}. \quad (14)$$

A partir de cette estimation, l'assureur détermine la provision mathématique  $L_0$  :

$$L_0 = \sum_{t=1}^{\infty} F_t (1+i)^{-t}. \quad (15)$$

Dans la suite, nous analyserons le montant aléatoire de l'engagement, soit la variable aléatoire :

$$\Lambda = \sum_{t=1}^{\infty} \tilde{F}_t (1+i)^{-t} = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \sum_{j \in \mathcal{J}} r_j * \mathbf{1}_{]t; \infty[}(T_{x(j)}) \quad (16)$$

Cette variable aléatoire est telle que  $E(\Lambda) = L_0$ . Pour une analyse détaillée de la loi de  $\Lambda$  lorsque la mortalité est connue (avec éventuellement des taux d'intérêt stochastiques), on se reportera à MAGNIN et PLANCHET [2000].

Les propriétés de  $\Lambda$  seront examinées dans un premier temps avec l'hypothèse classique que la mortalité future est connue (déterministe), via l'utilisation d'une table prospective, puis, dans un second temps, dans un contexte de mortalité stochastique. Dans ce second cas, on aura donc à considérer des surfaces de mortalité stochastiques  $\Pi$  et, une surface étant donnée, la loi conditionnelle de l'engagement  $\Lambda \mid \Pi$ .

## 2. LES DIFFERENTS TYPES DE MODELES

### 2.1 MODELES D'INSPIRATION FINANCIERE

#### 2.1.1 Modèle stochastique affine

Ce type de modélisation stochastique repose sur le principe suivant : si le taux instantané de mortalité est une fonction affine des paramètres, alors les probabilités de décès sont connues de manière fermée. Son principal intérêt est de proposer une structure simple permettant de nombreuses analyses. Elle donne une interprétation des paramètres et est cohérente avec les modèles d'évaluation des dérivés de mortalité. Par ailleurs, ce modèle considère l'intensité de mortalité pour tous les âges simultanément, contrairement aux modèles du type DAHL développés dans DAHL [2004]. Il permet également de prendre en compte les risques de longévité, d'échantillonnage et le risque systématique.

Enfin, en supposant l'indépendance du risque de mortalité et du risque financier, ce modèle peut permettre d'évaluer certains contrats d'assurance en prenant en compte la mortalité stochastique et en la combinant avec un modèle de type « *term structure model* ».

Si l'on note  $G_t$  la filtration contenant l'ensemble des informations de l'intervalle  $[0, t]$ , alors l'instant de décès  $\tau$  est un temps d'arrêt pour cette filtration. On note alors  $G_\tau$  la filtration engendrée par ce temps d'arrêt. En notant  $H_t$  la filtration contenant toutes les informations indépendamment du fait que l'individu soit en vie ou non, on peut écrire:

$$G_t = G_\tau \vee H_t. \quad (17)$$

Par ailleurs, en notant  $M_t$  la filtration des informations concernant la mortalité et  $F_t$  celle concernant les informations du marché financier, on obtient :

$$H_t = M_t \vee F_t. \quad (18)$$

La forme générale du taux instantané de mortalité est alors :

$$\mu_x(t) = g_0(x) + \sum_{i=1}^M Y_i(t)g_i(x). \quad (19)$$

où les  $g_i : x \rightarrow \mathfrak{R}^+$  et les  $Y_i(t)$  mesurent l'aléa sur le taux de mortalité. Le facteur dynamique de dimension M est donné par :

$$dY(t) = A(\theta - Y(t))dt + \Sigma\sqrt{V_t} dW_t^P. \quad (20)$$

où  $W_t^P$  est un mouvement brownien standard de dimension M sous la mesure historique, A et  $\Sigma$  sont des matrices carrées de taille M,  $\theta$  un vecteur de taille M et  $V_t$  une matrice diagonale telle que :

$$\forall i = 1, \dots, M; (V_t)_i = \alpha_i + \beta_i' Y(t). \quad (21)$$

où les  $\alpha_i$  et les  $\beta_i$  sont des vecteurs de taille M.

Avec les notations précédentes et en notant  $g(x) = (g_1(x), \dots, g_M(x))'$  on dispose alors des probabilités de survie suivantes :

$${}_{T-t}P_{x+t} = E^P [1_{[\tau_{\infty} > T]} | H_t \cup [\tau_x > t]]. \quad (22)$$

$${}_{T-t}P_{x+t} = E^P \left[ \exp \left( - \int_t^T \mu_{x+s}(s) ds \right) | M_t \right]. \quad (23)$$

En introduisant  $\gamma(x, t, T)$  et  $\delta(x, t, T)$  solutions du système suivant :

$$\frac{d\delta(x, t, T)}{dt} = -g(x+t) + A'\delta(x, t, T) + \frac{1}{2} \sum_{i=1}^M (\Sigma'\delta(x, t, T))_i^2 \beta_i. \quad (24)$$

$$\frac{d\gamma(x, t, T)}{dt} = -g_0(x+t) + \theta'A'\delta(x, t, T) + \frac{1}{2} \sum_{i=1}^M (\Sigma'\delta(x, t, T))_i^2 \alpha_i. \quad (25)$$

on obtient l'écriture simplifiée :

$${}_{T-t}P_{x+t} = \exp(\gamma(x, t, T) - \delta(x, t, T)Y_t). \quad (26)$$

Avec  $\gamma(x, T, T) = \delta(x, T, T) = 0$

### 2.1.2 Cadres des modèles

Ces modèles adoptent la simplification de modéliser le processus d'intensité  $\mu_x(t)$  pour un  $x$  fixé, c'est-à-dire comme seule fonction de  $t$ . On suppose alors que ce processus, étant non observable, s'écrit :  $\mu_x(t) = \Phi(X_t)$  où le processus  $X$  est solution de l'Equation Différentielle Stochastique suivante :

$$dX_t = \lambda(X_t)dt + \sigma(X_t)dB_t. \quad (27)$$

où  $B$  est un mouvement brownien standard. On suppose par ailleurs que la dépendance des coefficients  $\mu$  et  $\sigma$  en fonction de  $x$  est affine.

Les modèles de mortalité stochastique utilisent de manière intensive les processus à structure affine, qui conduisent à des formules fermées dans un grand nombre de cas. Lorsque l'on considère les modèles à structure affine à un facteur, on peut montrer que deux situations sont possibles : le processus d'Ornstein-Uhlenbeck associé au modèle de Vasicek dans la littérature portant sur les taux d'intérêt, et le processus de Feller associé quant à lui au modèle CIR (Du nom des auteurs COX, INGERSOLL et ROSS). Par ailleurs, ces deux processus apparaissent de manière naturelle dans des modèles physiques simples.

Afin de simplifier les écritures, on considère que l'on fixe un âge  $x$  et on cherche donc à modéliser le processus d'intensité à cet âge, vu comme une seule fonction de  $t$ . Cette approche est bien entendu assez restrictive puisqu'elle n'intègre pas explicitement la prise en charge de la surface de mortalité  $\mu_x(t)$ .

### 2.1.3 Le processus d'Orstein-Uhlenbeck

Le processus d'Ornstein-Uhlenbeck ne peut être a priori utilisé en l'état pour modéliser l'intensité du décès, puisqu'il autorise des valeurs négatives. Toutefois, la probabilité d'observer des valeurs négatives peut être rendue faible, et l'importance de ce processus dans la littérature nous a amené à le présenter ici.

D'un point de vue heuristique, le mouvement brownien permet de modéliser le mouvement d'une particule soumise à l'agitation thermique. Toutefois, en faisant l'hypothèse que la position de la particule est un processus de Markov à accroissements indépendants, on néglige certains paramètres. En effet possédant une masse, la particule possède une inertie. De ce fait sa position à l'instant  $t + h$  ne dépend pas uniquement de sa position à l'instant  $t$ , mais également de sa vitesse à cet instant.

Le processus d'Ornstein-Uhlenbeck permet de rendre compte de ce phénomène. Plus précisément, si la masse de la particule est  $m$  et que sa vitesse est  $x_t$ , on peut écrire que la variation de la quantité de mouvement de la particule entre  $t$  et  $t + dt$  est de la forme :

$$mdx_t = -rx_t dt + dM_t. \quad (28)$$

où  $r$  représente un coefficient de viscosité. Des considérations physiques relatives au terme  $dM_t$  représentant la part de la variation conséquence des chocs moléculaires conduisent à proposer que  $dM_t = \sigma dB_t$ , avec  $B$  un mouvement brownien. On obtient ainsi l'équation de Langevin :

$$mdx_t = -rx_t dt + \sigma dB_t. \quad (29)$$

Le processus solution de cette équation s'appelle le processus d'Ornstein-Uhlenbeck. Usuellement on utilise la présentation à partir de l'équation suivante, qui servira de référence par la suite :

$$dx_t = k(\theta - x_t)dt + \sigma dB_t. \quad (30)$$

Il résulte immédiatement de la définition que ce processus est gaussien, comme intégrale par rapport au mouvement brownien d'une fonction déterministe. En effet, on vérifie directement que la solution de l'équation précédente est donnée par :

$$x_t = x_0 e^{-kt} + \theta \left(1 - e^{-kt}\right) + \sigma \int_0^t e^{-k(t-s)} dB_s. \quad (31)$$

On en déduit en particulier que la variable  $Z(t, T) = \int_t^T x(s) ds$  est gaussienne. On a alors :

$$E \left[ \exp \left( - \int_s^t x(u) du \right) \middle| G_s \right] = \exp \left( - m(s, t) + \frac{1}{2} v(s, t) \right). \quad (32)$$

où  $m(s, t)$  et  $v(s, t)$  sont respectivement l'espérance et la variance conditionnelle de  $Z(s, t)$ . Leur calcul est le suivant :

$$m(s, t) = \theta(t - s) + (x_0 - \theta) \frac{e^{-ks}}{k!} \left(1 - e^{-k(t-s)}\right). \quad (33)$$

$$v(s, t) = \frac{\sigma^2}{2k} \left(1 - e^{-k(t-s)}\right)^2 + \frac{\sigma^2}{k} \left( (t - s) - \frac{1 - e^{-k(t-s)}}{k} \right). \quad (34)$$

#### 2.1.4 Le processus de Feller

Le processus d'Ornstein-Uhlenbeck présente l'inconvénient, comme on l'a vu précédemment, de prendre des valeurs négatives ou nulles avec une probabilité strictement positive. En modifiant légèrement l'équation différentielle qui le définit, on introduit ainsi un nouveau processus dont les trajectoires sont presque sûrement positives. C'est le processus de Feller.

Le processus de Feller est défini par l'équation différentielle stochastique suivante :

$$dx_t = k(\theta - x_t)dt + \sigma \sqrt{x(t)} dB_t. \quad (35)$$

Cette équation admet une solution unique pour  $k\theta > 0$ . La solution n'admet pas de représentation explicite, comme dans le cas du processus d'Ornstein-Uhlenbeck. Si la condition  $2k\theta > \sigma^2$  est de plus satisfaite, alors presque sûrement  $x(t) > 0$  pour tout  $t > 0$ .

Comme dans le cas du processus d'Ornstein-Uhlenbeck, par le calcul de l'espérance et de la variance on peut obtenir la probabilité de survie.

## 2.2 MODELE DE TYPE FINANCIER

Dans le cadre stochastique, on dispose de la fonction de survie du type :

$$S(u, y) = \exp\left(-\int_0^u \mu(t, y+t)dt\right). \quad (36)$$

On obtient alors aisément la probabilité qu'un individu d'âge  $x$  à l'origine et vivant en  $t$  soit encore vivant en  $T$  :

$$p(t, T, x) = E\left[\frac{S(T, x)}{S(t, x)} \mid M_t\right]. \quad (37)$$

où  $(M_t)$  représente la filtration associée à la structure  $\mu_x(t)$ . On définit alors une obligation  $B$  qui paie  $S(t,x)$  en  $T$ , on obtient donc le prix de l'obligation :

$$B(t, T, x) = E[S(T, x) \mid M_t]. \quad (38)$$

qui est un processus solution de l'équation différentielle stochastique suivante :

$$dB(t, T, x) = B(t, T, x)V'(t, T, x)d\tilde{W}(t). \quad (39)$$

où  $V(t,T,x)$  représente la volatilité de l'obligation et  $\tilde{W}(t)$  un processus brownien standard.



On peut alors définir un taux de mortalité instantané "forward" par :

$$\bar{\mu}(t, T, x + T) = -\frac{\mathcal{G}}{\mathcal{G}T} \log(B(t, T, x)). \quad (40)$$

et donc le taux de mortalité instantané "spot" par :

$$\mu(t, x + t) = \bar{\mu}(t, t, x + t). \quad (41)$$

En utilisant à présent les définitions ci-dessus, on peut appliquer ces modèles au taux de mortalité stochastique :

- Modèle zéro-coupon,
- Modèle de taux stochastique,
- Modèle "Survivor Credit Offer Rate",
- Les dérivés de mortalité stochastique.

Tous ces modèles sont davantage détaillés dans Cairns et al. [2004].

## 2.3 MODELES D'INSPIRATION DEMOGRAPHIQUE

### 2.3.1 Le modèle de Lee Carter

#### 2.3.1.1 Présentation du modèle

La méthode développée par Lee et Carter [1992] est une méthode d'extrapolation des tendances passées. La modélisation retenue pour le taux instantané de mortalité est :

$$\ln \mu_x(t) = \alpha_x + \beta_x k_t + \varepsilon_x(t). \quad (42)$$

où les variables aléatoires  $\varepsilon_x(t)$  sont indépendantes et identiquement distribuées suivant une loi  $N(0, \sigma^2)$ . Elle exprime le logarithme du taux instantané de mortalité comme dépendant de quatre paramètres. Les deux premiers ( $\alpha_x$  et  $\beta_x$ ) représentent l'impact de l'âge sur le taux instantané de mortalité. Le troisième ( $k_t$ ) représente l'impact de l'année d'observation sur le taux instantané. Enfin le dernier paramètre ( $\varepsilon_x(t)$ ) représente l'erreur de modèle qui peut être faite durant l'estimation des paramètres.

Il s'agit donc d'ajuster à la série des logarithmes des taux instantanés une structure paramétrique combinée à un phénomène aléatoire.

Afin de rendre les paramètres identifiables, il est nécessaire de fixer des contraintes. De plus cela permet une description plus fine des paramètres. Nous appliquerons ici les contraintes habituelles.

$$\sum_{x=x_m}^{x_M} \beta_x = 1. \quad (43)$$

$$\sum_{t=t_m}^{t_M} k_t = 0. \quad (44)$$

Ainsi, en forçant la somme des  $k_t$  à être nulle, on force  $\alpha_x$  à représenter la valeur moyenne du logarithme du taux de mortalité sur l'intervalle d'étude.

En forçant la somme des  $\beta_x$  à être égale à 1, on traduit le caractère croissant ou décroissant de la série  $k_t$ . En effet considérons deux années  $t$  et  $t'$ . Supposons qu'en moyenne la mortalité durant l'année  $t$  ait été plus faible que durant l'année  $t'$ .

Les taux de mortalité durant cette année sont en moyenne plus faibles et donc la somme  $\sum_{i=x_m}^{x_M} \ln(\mu_{x_i})$  qui est égale à  $(k_t - k_{t'}) \sum_{i=x_m}^{x_M} \beta_x + \sum_{i=x_m}^{x_M} \ln(\mu_{x_i'})$  est inférieure à  $\sum_{i=x_m}^{x_M} \ln(\mu_{x_i})$ . Ainsi,  $k_t$  est inférieur à  $k_{t'}$ . Une décroissance de la série  $(k_t)$  représente donc une diminution globale (autrement dit à presque tous les âges) de la mortalité.

Comme  $\frac{\delta \ln(\mu_x(t))}{\delta x} = \beta_x \frac{dk_t}{dt}$ ,  $\beta_x$  représente la sensibilité du taux instantané de mortalité à l'âge  $x$  à la décroissance globale de la mortalité. Si  $\beta_x$  est négatif, alors  $\mu_x$  réagira d'une manière opposée à l'évolution générale de la mortalité. Plus la valeur absolue du coefficient  $\beta_x$  sera élevée, plus  $\mu_x$  prendra en compte les modifications de la mortalité.

### 2.3.1.2 Estimation des paramètres

Le critère d'optimisation de la modélisation est la minimisation de la variance des erreurs. Par le critère des moindres carrés, on obtient les paramètres :

$$\left( \hat{\alpha}_x, \hat{\beta}_x, \hat{k}_t \right) = \arg \min_{\alpha_x, \beta_x, k_t} \left[ \sum_{x,t} \left( \ln \mu_x^*(t) - \alpha_x + \beta_x k_t \right)^2 \right]. \quad (45)$$

Estimons tout d'abord  $\alpha_x$ .

Par application de la méthode des moindres carrés,  $\alpha_x$  est solution de l'équation suivante :

$$\frac{\delta}{\delta \alpha_x} \sum_{y,t} \left( \ln(\mu_y^*(t)) - \alpha_y + \beta_y k_y \right)^2 = 0. \quad (46)$$

Ce qui peut se réécrire :  $2(t_M - t_m + 1) \hat{\alpha}_x - 2 \sum_{t=t_m}^{t_M} \left( \ln(\mu_x^*(t)) + \beta_x k_t \right) = 0$

De plus, en tenant compte de la contrainte 44, on obtient :

$$\hat{\alpha}_x = \frac{1}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \ln(\mu_x^*(t)). \quad (47)$$

On considère alors la matrice  $Z = (z_{xt})$  des taux centrés par rapport à la dimension temporelle, c'est à dire telle que :

$$z_{xt} = \ln(\mu_x^*(t)) - \hat{\alpha}_x . \quad (48)$$

On recherche alors une approximation de  $Z$  en un produit de deux vecteurs de la forme  $Z \approx \hat{\beta} \hat{k}'$ , de sorte que la décomposition soit optimale au sens du critère des moindres carrés, c'est à dire qui explique le mieux possible la variance totale. Pour résoudre ce problème, il faut passer par la décomposition en valeurs propres de la matrice  $Z$ , sous la forme

$$Z = \sum_{i \geq 1} \sqrt{\lambda_i} v_i u_i' . \quad (49)$$

où  $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$  sont les valeurs propres de  $Z'Z$ ,  $u_i$  le vecteur propre normé de  $Z'Z$  associé à la valeur propre  $\lambda_i$  et  $v_i$  le vecteur propre normé associé à la même valeur propre mais pour la matrice  $ZZ'$ . En effet, comme  $Z'Z u_i = \lambda_i u_i$  on a alors  $ZZ'(z u_i) = \lambda_i Z u_i$  ; donc les 2 matrices ont les mêmes valeurs propres avec les mêmes ordres de multiplicité. En posant  $v_i = \frac{1}{\sqrt{\lambda_i}} Z u_i$ , on remarque que  $Z u_i u_i' = \sqrt{\lambda_i} v_i u_i'$ , donc en sommant et en tenant compte de l'orthogonalité des vecteurs propres on retrouve bien la formule  $Z$ . On pose alors naturellement l'approximation  $Z \approx \sqrt{\lambda_1} v_1 u_1'$  dont la qualité est mesurée par  $\frac{\lambda_1}{\sum \lambda_i}$ . On obtient alors les estimateurs suivant pour  $\beta$  et pour  $k$  :

$$\hat{\beta} = \frac{1}{\sum v_{1j}} v_1$$

$$\hat{k} = \sqrt{\lambda_1} \sum v_{1j} u_{1j}$$

En vue d'extrapoler les taux de mortalité pour les  $t \geq t_M$ , on est donc amené à analyser la série des  $\left( \hat{k}_t \right)$  afin de lui ajuster un modèle de série temporelle. Avant toute chose, il paraît raisonnable d'ajuster les paramètres du modèle afin que le nombre de décès prévu soit égal au nombre de décès observé chaque année. La population sous risque étant exprimée par  $L_{xt} = \frac{D_{xt}}{\mu_x^*(t)}$  avec  $D_{xt} = D_{xt}(t-x) + D_{xt}(t-x-1)$ , l'ajustement s'exprime alors par :

$$\sum_{x=x_m}^{x_M} D_{xt} = \sum_{x=x_m}^{x_M} L_{xt} \exp\left(\hat{\alpha}_x + \hat{\beta}_x \hat{k}_t\right). \quad (50)$$

de variable  $\hat{k}_t$ . On introduit alors la fonction  $F(x) = \sum_{x=x_m}^{x_M} L_{xt} \exp\left(\hat{\alpha}_x + \hat{\beta}_x \hat{k}_t\right) - \sum_{x=x_m}^{x_M} D_{xt}$ , ainsi la contrainte  $F(k)=0$ . Les caractéristiques de la fonction  $F$  nous assurent l'unicité de la racine lorsqu'elle existe. L'obtention de la racine peut se faire par un algorithme de type Newton-Raphson. Pour cela, il suffit de poser :

$$k_{i+1} = k_i - \frac{F(k_i)}{F'(k_i)}. \quad (51)$$

et avec le critère initial  $k_0 = \hat{k}_t$  et le critère d'arrêt  $\left| \frac{k_{i+1} - k_i}{k_i} \right| \leq \varepsilon$  où  $\varepsilon$  est petit.

Une fois la série des  $\left(\hat{k}_t\right)$  ajustée, elle doit encore être corrigée afin de remplir la condition d'identifiabilité 44, ce qui nous conduit à poser :

$$k_t^* = \hat{k}_t - \frac{1}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \hat{k}_t. \quad (52)$$

Il faut à présent corriger les  $\left(\hat{\alpha}_x\right)$  pour assurer que l'égalité entre le nombre de décès prédits et le nombre de décès observés reste valable, ce qui nous conduit à poser :

$$\alpha_x^* = \hat{\alpha}_x + \frac{\hat{\beta}_x}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \hat{k}_t. \quad (53)$$

### 2.3.1.3 Extrapolation de la composante temporelle

La modélisation de la série des  $(k_t)$  se fait généralement par un modèle ARIMA en suivant la technique de Box et Jenkins, mais toute série temporelle peut convenir. Il faut prendre en compte l'adéquation aux données dont on dispose, et tenir compte également de la population que l'on veut modéliser. On peut, par exemple, utiliser un modèle très simple comme une régression linéaire des coefficients, de sorte que :

$$k_t^* = a + bt + \varepsilon_t. \quad (54)$$

où  $(\varepsilon_t)$  est un bruit blanc gaussien.

### 2.3.2 Le modèle de Poisson

Le modèle de Lee-Carter repose sur l'hypothèse forte mais peu réaliste de l'homoscédasticité des taux de mortalité. En effet, les valeurs estimées des taux bruts sont très instables pour les grands âges compte tenu des faibles effectifs, d'où une variance importante pour ces âges. Vient alors l'idée intuitive de modéliser le nombre de décès à l'âge  $x$  l'année  $t$  par une loi de Poisson, en supposant que  $D_{xt}$  suit une loi de Poisson de paramètre  $L_{xt}\mu_x(t)$  avec  $\ln \mu_x(t) = \alpha_x + \beta_x k_t + \varepsilon_x(t)$  comme pour le modèle de Lee-Carter. Cette approche est plus largement détaillée dans Brouhns et Denuit [2002]. Le lecteur intéressé pourra également se référer à Haderer [2003] pour une application. Les paramètres ont alors les mêmes significations et subissent les mêmes contraintes. On dispose alors de la log-vraisemblance du modèle :

$$\ln L(\alpha, \beta, k) = \sum_{x,t} [D_{xt}(\alpha_x + \beta_x k_t) - L_{xt} \exp(\alpha_x + \beta_x k_t)] + \text{constante} \quad (55)$$

Les équations de vraisemblance peuvent être résolues grâce à l'expression simple de la log-vraisemblance. On peut par exemple utiliser le procédé itératif proposé par Goodman [1979] et l'algorithme de Newton-Raphson dans le but d'annuler le vecteur des scores. Cette méthode conduit aux relations de récurrences suivantes :

$$\hat{\alpha}_x^{i+1} = \hat{\alpha}_x^i - \frac{\sum_t \left( D_{xt} - L_{xt} \exp\left(\hat{\alpha}_x^i + \hat{\beta}_x^i \hat{k}_t^i\right) \right)}{-\sum_t \left( L_{xt} \exp\left(\hat{\alpha}_x^i + \hat{\beta}_x^i \hat{k}_t^i\right) \right)}. \quad (56)$$

$$\hat{k}_t^{i+1} = \hat{k}_t^i - \frac{\sum_t \left( D_{xt} - L_{xt} \exp\left(\hat{\alpha}_x^i + \hat{\beta}_x^i \hat{k}_t^i\right) \right) \hat{\beta}_x^i}{-\sum_t \left( L_{xt} \exp\left(\hat{\alpha}_x^i + \hat{\beta}_x^i \hat{k}_t^i\right) \right) \left(\hat{\beta}_x^i\right)^2}. \quad (57)$$

$$\hat{\beta}_x^{i+1} = \hat{\beta}_x^i - \frac{\sum_t \left( D_{xt} - L_{xt} \exp\left(\hat{\alpha}_x^i + \hat{\beta}_x^i \hat{k}_t^i\right) \right) \hat{k}_t^i}{-\sum_t \left( L_{xt} \exp\left(\hat{\alpha}_x^i + \hat{\beta}_x^i \hat{k}_t^i\right) \right) \left(\hat{k}_t^i\right)^2}. \quad (58)$$

Les valeurs initiales sont a priori libres. Il suffit uniquement de spécifier les valeurs  $\hat{\beta}_x^0 \neq 0$ . Enfin, pour soustraire les contraintes d'identifiabilité, il convient d'ajuster les paramètres estimés précédemment en posant :

$$k_t^* = \left( \begin{pmatrix} \hat{k} \end{pmatrix} \right) \frac{1}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \left( \begin{pmatrix} \hat{k} \end{pmatrix} \right) \sum_x \hat{\beta}_x. \quad (59)$$

$$\beta_x^* = \frac{\hat{\beta}_x}{\sum_x \hat{\beta}_x}. \quad (60)$$

$$\alpha_t^* = \left( \begin{pmatrix} \hat{\alpha}_x \end{pmatrix} \right) \frac{\hat{\beta}_x}{t_M - t_m + 1} \sum_{t=t_m}^{t_M} \left( \begin{pmatrix} \hat{k} \end{pmatrix} \right). \quad (61)$$

Les résultats obtenus sont assez proches des résultats du modèle de Lee-Carter. En ce qui concerne l'extrapolation de la composante temporelle, elle est réalisée de la même manière que dans le modèle de Lee-Carter.

### 3. METHODE DE SIMULATION

L'ensemble de l'étude de la mortalité stochastique présentée ici est empirique. Elle repose donc en grande partie sur la technique de Monte-Carlo, qui repose elle-même sur la Loi des Grands Nombres.

Son principe est simple : si l'on possède un certain nombre de réalisations indépendantes d'une variable aléatoire, alors l'estimateur empirique de sa moyenne convergera presque sûrement vers l'espérance de la variable aléatoire source. Cependant, ce résultat n'est qu'asymptotique. Pour une plus grande précision, il faut alors appliquer le théorème de La Limite Centrale, qui nous permet d'obtenir un intervalle de confiance dont les bornes dépendent de la taille de l'échantillon.

Afin d'obtenir des estimations relativement précises, il est nécessaire de simuler un grand nombre de fois le montant de l'engagement. Les résidus du modèle de Lee-Carter étant gaussiens, il est nécessaire de déterminer l'algorithme permettant d'obtenir le plus rapidement possible une suite de gaussiennes centrées réduites indépendantes.

Nous aborderons dans cette partie le principe de la simulation et les aspects qui entrent en jeu, notamment la modélisation de la sortie des individus. Nous verrons ensuite le choix du générateur pseudo aléatoire rendu nécessaire pour les simulations, tout en gardant à l'esprit l'aspect temps de calcul afin de conserver une dimension applicable à cette étude.

#### 3.1 ALGORITHME PRINCIPAL

L'algorithme permettant d'obtenir la distribution de l'engagement moyen, ainsi que la part de variance expliquée par la mortalité stochastique, est basé sur un scénario central qui permet de simuler une surface de mortalité.

Ainsi après avoir calibré les paramètres  $(\alpha_x), (\beta_x), (k_t)$  du modèle de Lee-Carter, et ajusté un modèle permettant de simuler de manière stochastique la dérive de la mortalité dans le futur, il ne reste plus qu'à simuler des variables aléatoires gaussiennes indépendantes et identiquement distribuées afin de simuler une première surface de mortalité. La surface étant obtenue, on simule N scénarios de survie des rentiers, et ce jusqu'à extinction de la cohorte. Pour chaque scénario on calcule la valeur de l'engagement, puis on stocke l'ensemble de ces valeurs.

Le scénario central qui vient d'être décrit est réitéré M fois. Les valeurs de N et de M sont calibrées de manière empirique, de façon à obtenir, grâce à la technique de Monte Carlo, des résultats robustes.



Une fois les simulations terminées, nous sommes en possession de N.M données. Les traitements restant à effectuer peuvent être résumés dans le tableau suivant.

Simulation de la surface de mortalité	Simulation du montant de l'engagement	Espérance de l'engagement	Variance de l'engagement
$\Sigma_1$	$\lambda_1$ ... $\lambda_N$	$E[\Lambda]$	$V[\Lambda]$
...	...	...	...
$\Sigma_M$	$\lambda_1$ ... $\lambda_N$	$E[\Lambda]$	$V[\Lambda]$
Résultat		Variance stochastique : $V[E[\Lambda \Sigma]]$	Variance mutualisable : $E[V[\Lambda \Sigma]]$
		Variance de l'engagement : $V[E[\Lambda \Sigma]] + E[V[\Lambda \Sigma]]$	

Tab:2 Calcul de la volatilité de l'engagement

Où  $\lambda_i$  représente une réalisation de la variable aléatoire  $\Lambda$ . Cette dernière modélise l'engagement du portefeuille.

$\Sigma_i$  représente une réalisation de la variable aléatoire  $\Sigma$ . Cette dernière modélise les simulations de surface de mortalité stochastique.

### 3.2 MODELISATION DE LA SORTIE DES INDIVIDUS

La technique la plus simple afin de simuler la durée de vie d'un individu est la technique par récurrence.

Après avoir obtenu les probabilités annuelles de décès d'un individu d'âge  $x$ , on tire une première variable aléatoire  $U$  suivant une distribution uniforme sur  $[0,1]$ . Si  $U$  est inférieur à la probabilité de décéder entre  $x$  et  $x+1$ , alors l'individu vit au moins une année. Il a maintenant l'âge  $x+1$ .

On tire une nouvelle variable aléatoire  $U$  et on la compare avec la probabilité de décéder entre  $x+1$  et  $x+2$ . On itère ce procédé tant que la variable uniforme tirée est inférieure à la probabilité de décès à laquelle on la compare.

Supposons que cet événement arrive lors de la  $i^{\text{ème}}$  itération. L'individu vivra alors  $i-1$  année.

Cet algorithme est très simple à mettre en œuvre mais n'est pas optimal pour le portefeuille que nous étudions. En effet, l'âge moyen étant de 63,8 ans, on peut s'attendre en moyenne à observer 16 itérations par individu. Soit en moyenne 5984 simulations de variables aléatoires uniformes pour chaque calcul d'engagement.

Pour optimiser le calcul de la durée de survie des individus, nous allons avoir recours à la technique dite d'inversion de la fonction de répartition.

On calcule dans un premier temps la distribution de l'engagement de la durée de vie de l'individu. Notons  $F_x$  la fonction de répartition de la durée de survie d'un individu d'âge  $x$ . Afin de simuler la durée de survie d'un individu d'âge  $x$ , on tire une variable aléatoire uniforme  $U$ . L'individu vivra alors  $t$  année avec  $F_x(t) \leq U < F_x(t+1)$ . Ainsi en moyenne, le nombre de variables aléatoires uniformes à simuler se verra alors divisé par 16.

Si l'on cherche à optimiser d'avantage l'algorithme de sortie des individus, une technique de dichotomie peut être appliquée afin de déterminer le paramètre  $t$ .

### 3.3 GENERATEUR ALEATOIRE

Les résultats de cette étude étant basés sur la simulation et l'utilisation de la méthode de Monte-Carlo, la méthode utilisée pour simuler des nombres aléatoires revêt une importance particulière. Cette étude requiert la génération de lois uniformes sur  $[0,1]$  et celle de lois normales centrées réduites. On pourra notamment se reporter à Planchet et al. [2005].

#### 3.4.1 Générateur de lois uniformes

Les générateurs de nombres aléatoires peuvent être classés en deux catégories :

- Les générateurs pseudo-aléatoires
- Les générateurs quasi-aléatoires.

Les générateurs pseudo-aléatoires sont les plus classiquement connus. Ils sont dits pseudo-aléatoires dans le sens où l'on peut prédire la suite des nombres qui sera simulée, sous certaines conditions. Ils possèdent de mauvaises propriétés d'équi-répartition sur l'intervalle  $[0,1]$ . A l'inverse, les générateurs quasi-aléatoires s'appuient sur la construction de suites à discrédance faible. Ils possèdent de ce fait de bonnes propriétés d'équi-répartition sur l'intervalle unité. Ce sont donc ces générateurs qui seront préférés lors des simulations de Monte-Carlo.

Mais si les générateurs quasi-aléatoires sont performants, ils s'appuient souvent sur des algorithmes de programmation complexes, pénalisant ainsi les temps de calculs. Afin de limiter les temps de calculs, il est donc important de choisir le générateur quasi-aléatoire le plus rapide.

C'est la méthode de translation irrationnelle du Tore qui a été choisie ici. La suite implémentée est la suivante :

$$x_n = \left( \text{frac} \left( n \sqrt{P_1} \right), \dots, \text{frac} \left( n \sqrt{P_s} \right) \right)$$

où  $p_i$  est le  $i^{\text{ème}}$  nombre premier, et où  $\text{frac}(y)$  représente la partie fractionnaire de  $y$ , c'est-à-dire  $y$  moins sa partie entière. Le principal inconvénient de cette méthode réside dans la forte dépendance existant entre les nombres simulés. Ceci peut engendrer des biais dans notre étude puisque les individus sont supposés indépendants.

Pour solutionner ce problème, nous allons recourir à l'algorithme du Tore mélangé. Cet algorithme consiste à simuler une suite de nombres à l'aide de la translation irrationnelle du Tore puis de mélanger l'ordre de cette suite à l'aide d'une permutation notée  $\varphi$ .

Ainsi, si l'on note  $u_n$  la  $n^{\text{ième}}$  réalisation de l'algorithme du Tore à l'aide d'un nombre premier  $p$ , l'algorithme du Tore mélangé remplacera  $u_n$  par  $u_{\varphi(n)}$  où  $\varphi$  est définie par :

$$\varphi : [0; T] \rightarrow [0; T]$$

$$n \rightarrow \left[ \alpha T \tilde{u} + 1 \right]$$

où :

- $[x]$  représente la partie entière de  $x$ ,
- $\alpha$  représente un nombre quelconque permettant de contrôler la probabilité de tirer deux fois le même nombre aléatoire,
- $T$  représente le nombre de variables aléatoires que l'on souhaite générer
- $\tilde{u}$  représente un nombre aléatoire.

Dans notre étude nous prendrons  $\alpha T = 10^8$ .

Le problème de la suite des nombres premiers est ainsi également résolu. Cependant, un temps de calcul supplémentaire dû à la génération d'un nombre aléatoire par une méthode implémentée a été introduit.

### 3.4.2 Générateur de lois normales

La fonction de répartition de la loi normale n'étant pas facilement inversible, il est nécessaire de recourir à des algorithmes spécifiques. Par exemple la méthode de Box et Muller qui retourne le couple de variables aléatoires  $(x_1, x_2)$ , avec :

$$x_1 = \sqrt{-2 \ln(u)} \cos(2\pi v)$$

$$x_2 = \sqrt{-2 \ln(u)} \sin(2\pi v)$$

où  $(x_1, x_2)$  est un couple de gaussiennes centrées réduites indépendantes, et  $u$  et  $v$  deux variables aléatoires indépendantes uniformément réparties sur  $[0,1]$ . Cet algorithme, très simple à mettre en œuvre, nécessite de posséder un générateur de nombres aléatoires de bonne qualité. Générer deux nombres aléatoires indépendants à chaque passage dans l'algorithme peut s'avérer coûteux en terme de temps. Il est donc préférable d'utiliser un algorithme ne nécessitant qu'un seul nombre aléatoire. L'algorithme de Moro répond à ce critère.

Son algorithme est le suivant :

1. Soit  $U$  un nombre généré suivant une loi uniforme sur  $[0,1]$ . On pose «  $x$  » la réalisation d'une gaussienne centrée réduite telle que  $y = \Phi(x)$ , où  $\Phi$  est la fonction de répartition d'une gaussienne centrée réduite.

2. On pose  $Z=U-0.5$ .

– Si  $|Z| \leq 0.42$ , on approche  $x$  par :

$$x = Z \frac{\sum_{i=0}^3 a_i Z^{2i}}{\sum_{j=0}^4 b_j Z^{2j}}$$

– Si  $|Z| > 0.42$ , on approche  $x$  par :

$$x = \text{sg}(Z) \left( \sum_{i=0}^8 c_i T_i(t) \right) - \text{sg}(Z) \frac{c_0}{2}$$

$$\text{où } \text{sg}(Z) \text{ est le signe de } Z \text{ et } t = k_1 \left\{ 2 \ln \left( - \ln \left( \frac{1}{2} - |Z| \right) \right) - k_2 \right\}$$

La fonction  $f(t)$ , quant à elle, peut être approchée par l'algorithme suivant :

1. Posons  $d_{10} = d_9 = 0$

2. Soient  $d_i$  les réels déterminés par la relation récursive :

$$d_i = 2td_{i+1} - d_{i+2} + c_i \text{ pour tout } i=8,7,\dots,1$$

3. On obtient alors :  $f(t) = td_1 - d_2 + \frac{c_0}{2}$

Les valeurs des constantes  $(a_i, b_j, k_t)$  sont données en annexe.

## Construction du modèle

Dans la littérature, les modélisations stochastiques des phénomènes de mortalité sont nombreuses. On ne considère pas ici les modèles d'inspiration financière nés de la problématique de tarification de dérivés de mortalité. Le lecteur intéressé par ces approches pourra consulter DAHL [2004] et SCHRAGER [2006]. On notera simplement que l'ajustement de ces modèles à des données de mortalité n'est pas immédiat et que, pour la situation qui nous préoccupe ici, la non prise en compte des dépendances entre les âges est problématique (ces modèles proposent une dynamique d'évolution en fonction du temps pour chaque âge sans prendre en compte les corrélations existant entre les âges proches).

Plusieurs modèles classiques sont de fait des modèles stochastiques ; en premier lieu, les lissages bayésiens, et en particulier le modèle de Kimeldorf-Jones (KIMELDORF et JONES [1967]) entrent dans cette catégorie. Toutefois, ils sont élaborés essentiellement dans la perspective du lissage d'une table du moment, sans intégrer naturellement la dimension prospective essentielle dans l'analyse de l'engagement d'un régime de rentes.

On retient dans cette étude le modèle de Lee-Carter, qui permet aisément de produire des surfaces de mortalité stochastiques et qui est, par ailleurs, devenu un standard pour la construction de tables prospectives. La variante log-Poisson (*cf.* BROUHNS et al. [2002]) conduit à des résultats très proches.

Après avoir construit un jeu de tables prospectives sur des données nationales à l'aide de ce modèle, nous l'utilisons pour calibrer la composante aléatoire pour le risque systématique et appliquons le modèle ainsi obtenu pour déterminer le niveau de l'engagement d'un régime de rentes.

Les données utilisées pour les applications numériques sont reprises de PLANCHET et THEROND [2004].

L'ajustement du modèle de Lee-Carter se fait en deux étapes.

Tout d'abord il faut ajuster les suites  $(\alpha_x)$  et  $(\beta_x)$ . Ces deux suites sont déterministes et seront utilisées dans toutes les créations de tables du moment, quelle que soit l'année de la table.

Ensuite, il faut ajuster la série  $(k_t)$  représentant l'évolution de la mortalité dans le passé. Cette série ne sera pas utilisée telle quelle lors des créations de tables du moment dans le futur. C'est la tendance de cette série qui sera utilisée afin de projeter l'évolution future de la mortalité.

## Ajustement des paramètres $(\alpha_x), (\beta_x), (k_t)$

Après application de la technique décrite dans la partie I.3, nous obtenons les données suivantes pour la série  $(k_t)$  :

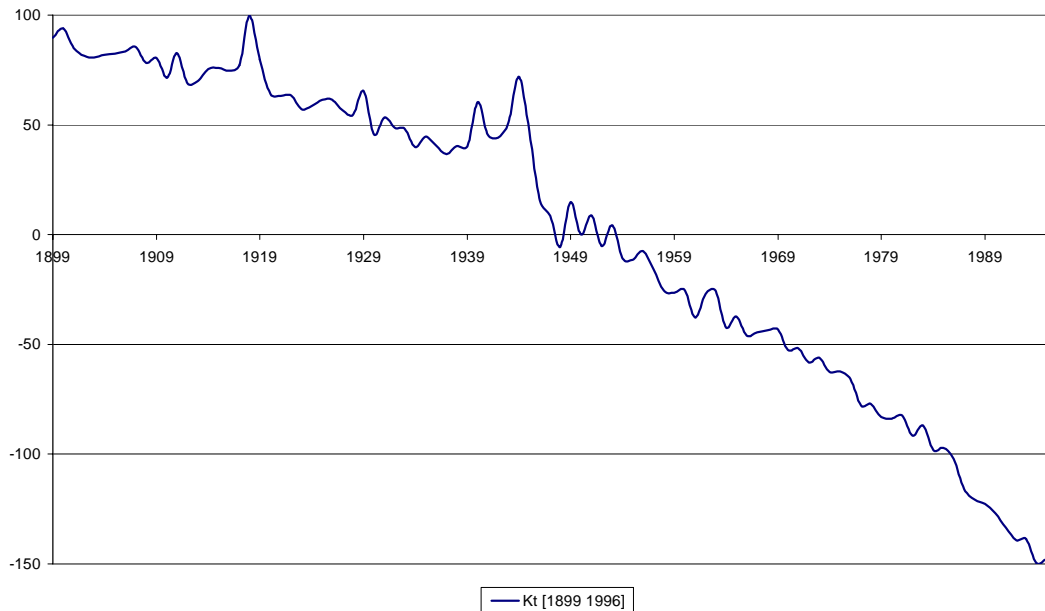


Fig. 7 : Evolution de la série  $k(t)$  entre 1899 et 1996

On remarque que le graphique représentant l'évolution de la dérive temporelle au cours du 20<sup>ème</sup> siècle présente deux pics. Il s'agit des deux Guerres Mondiales. Or la période couvrant les deux guerres ne représente pas l'évolution normale de la mortalité de la population française. Il ne faut donc pas prendre en compte ces deux périodes, au risque d'augmenter artificiellement la volatilité de la dérive de la mortalité et de sous-estimer la décroissance de la mortalité.

Les calculs sont relancés sur la période 1950-1996. Nous obtenons les données suivantes pour les séries  $(\alpha_x), (\beta_x), (k_t)$  :

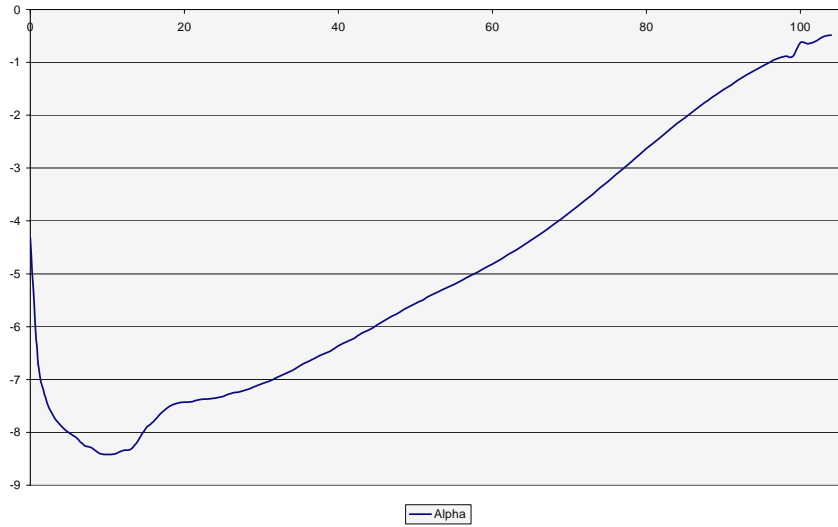


Fig. 8 : Alpha

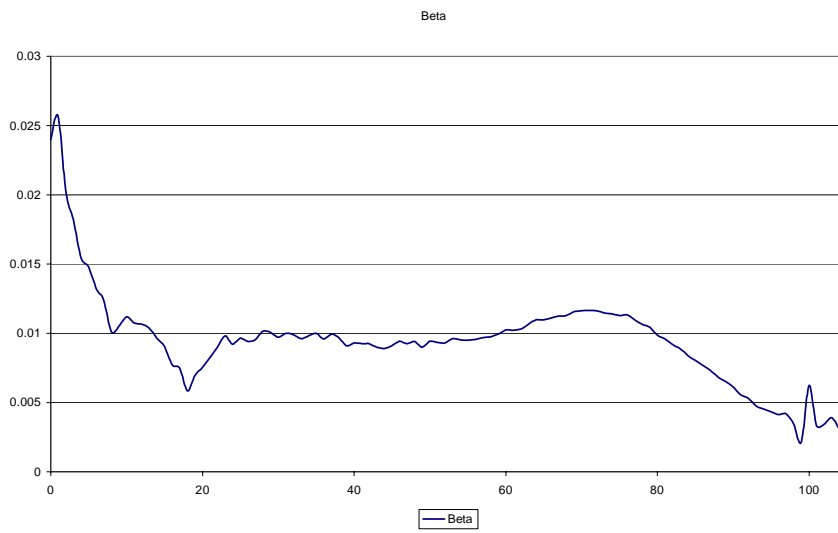


Fig. 9 : Bêta

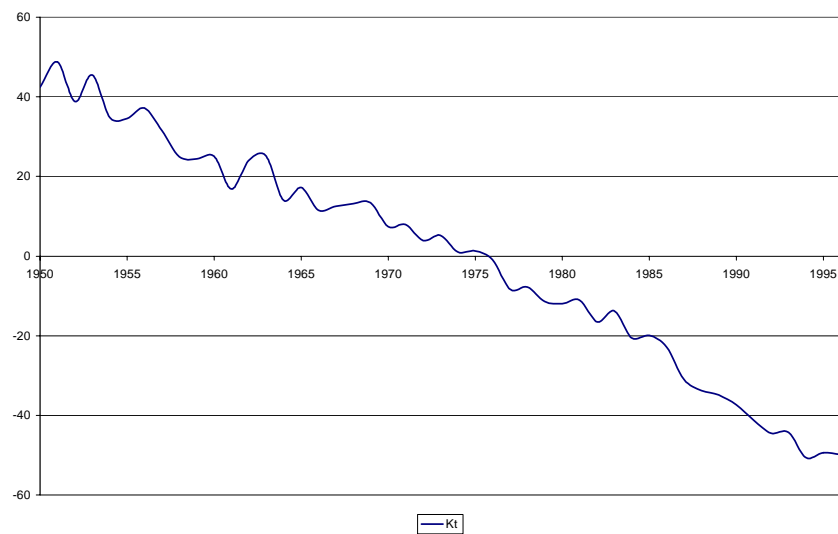


Fig. 10 : Evolution de la série k(t) entre 1950 et 1996



## Analyse de la projection de la dérive temporelle ( $k_t$ )

Une fois la surface de mortalité ajustée sur les données passées, il reste à modéliser la série ( $k_t$ ) pour extrapoler les taux futurs.

Deux scénarios sont envisageables :

- ✓ On peut ajuster sur la série ( $k_t$ ) une courbe de tendance de type linéaire.
- ✓ On ne peut pas ajuster sur la série ( $k_t$ ) une courbe de tendance de type linéaire.

Si l'on se trouve dans la deuxième hypothèse, il faut alors utiliser un modèle temporel de type ARIMA, mais toute autre modélisation de série temporelle peut être utilisée.

Si l'on trouve dans la première hypothèse il faut envisager deux types de modélisation :

- ✓ Une régression linéaire simple
- ✓ Une analyse temporelle.

Cependant le choix entre une modélisation linéaire et une modélisation temporelle n'est pas un choix sans conséquence. En effet ajuster un processus à l'aide d'une régression linéaire revient à supposer que les sauts aléatoires n'ont de conséquences qu'au moment où ils apparaissent. Du point de vue de la mortalité stochastique, cela revient à supposer que la dérive de la mortalité est un phénomène linéaire soumis à de légères perturbations annuelles, lesquelles n'ont aucune conséquence à long terme. A l'inverse modéliser la dérive de la mortalité à l'aide d'une série temporelle revient à supposer que son évolution est perturbée par des sauts annuels impactant sa trajectoire d'évolution future.

Si des tests permettent de trancher entre ces deux types de modèles, ils ne permettent pas de décider si un modèle transversale, type « risque de dérive » ne pourrait constituer un modèle acceptable. De plus, lorsque la volatilité ajustée des sauts est relativement faible, une modélisation linéaire peut constituer un processus tout à fait acceptable.

Ainsi dans un souci de créer le modèle le plus facilement applicable, nous analyserons les trois modèles dont nous venons de parler. Nous nous appuierons sur les résultats obtenus ainsi que sur les tests statistiques afin de choisir quel modèle est préférable pour analyser la mortalité stochastique. En effet si modèle temporel et modèle linéaire donnent des résultats sensiblement identiques, l'emploi d'une régression linéaire pourrait être préféré par souci de simplicité et de rapidité.

Dans un premier temps, nous allons donc étudier si la série  $(k_t)$  que nous avons ajustée présente une tendance linéaire :

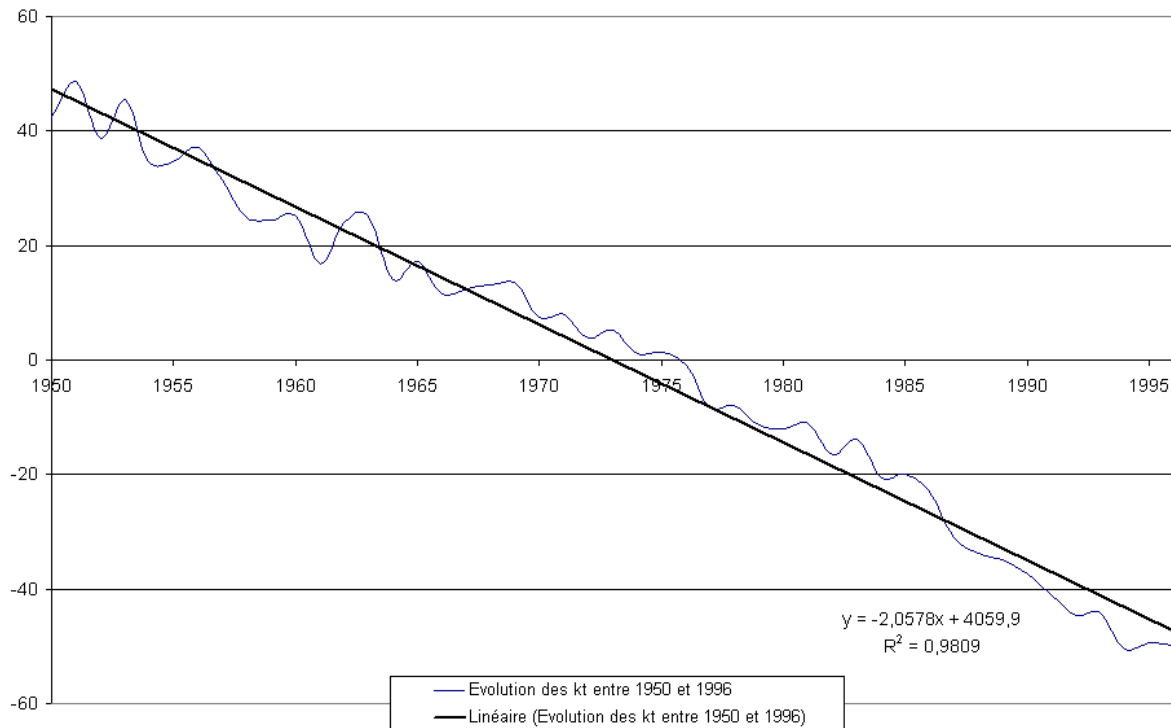


Fig. 11 : Evolution du paramètre  $k(t)$  avec le temps

Compte tenu de l'allure obtenue sur les données disponibles, une régression linéaire semble être un bon ajustement. En effet, la décroissance de la série  $(k_t)$  semble linéaire, et les oscillations autour de sa tendance ressemblent à des bruits blancs.

Comme nous ne pouvons pas rejeter le modèle linéaire, nous allons devoir tester le modèle linéaire et le modèle temporel.

# I Etude du risque systématique à l'aide d'une régression linéaire

## 1. CAS D'UNE MORTALITE DETERMINISTE

### 1.1 AJUSTEMENT DU MODELE

On choisit donc de modéliser la dérive de la mortalité à l'aide du modèle linéaire suivant :

$$k_t^* = at + b + \gamma_t, \quad (62)$$

avec  $(\gamma_t)$  un bruit blanc gaussien de variance  $\sigma_\gamma$ . On obtient ainsi des estimateurs  $\hat{a}$  et  $\hat{b}$  qui permettent de construire des surfaces projetées en utilisant simplement  $\hat{k}_t = \hat{a}t + \hat{b}$ . C'est ce modèle qui sera utilisé par la suite.

Il est précisé que l'objectif du présent travail étant de quantifier, dans le risque total auquel est exposé un régime rentier, la part issue de l'incertitude sur la mortalité future, nous privilégions le choix d'un modèle stochastique simple et opérationnel. Les limites du modèle de Lee-Carter, si elles peuvent en effet avoir un impact sur le niveau absolu de la provision obtenue, doivent avoir un impact moindre sur la répartition entre la part mutualisable et la part non mutualisable du risque de mortalité.

On notera également que, compte tenu de la répartition actuelle par âge du groupe étudié, la méthode de fermeture de la table n'est pas déterminante dans la formation du résultat. Plus généralement, on peut considérer que l'importance du choix d'une méthode de fermeture peut être relativisée dans le cas d'un régime de rentes. On pourra consulter sur ce point PLANCHET et THEROND [2006]. La question de la fermeture des tables est également discutée dans DENUIT et QUASHIE [2005] avec un point de vue sensiblement différent.

Les estimateurs  $\hat{a}$  et  $\hat{b}$  sont obtenus à partir de la procédure « Reg » sous SAS :

Résultats estimés des paramètres					
Variable	DF	Résultat estimé des paramètres	Erreur std	Valeur du test t	Pr >  t
Intercept	1	4059.94439	84.49378	48.05	<.0001
AN	1	-2.05775	0.04282	-48.05	<.0001

Tab:3 Estimation des paramètres  $\hat{a}$  et  $\hat{b}$

Les tests de Student de significativité des paramètres  $\hat{a}$  et  $\hat{b}$  permettent d'accepter l'hypothèse de non nullité des coefficients. De plus, le  $R^2$  de la régression égal à 0.9809 confirme la qualité de la régression effectuée.

Root MSE	3.98240	R-Square	0.9809
Dependent Mean	1.51179E-16	Adj R-Sq	0.9805
Coeff Var	2.634225E18		

Tab:4 Analyse de la régression

Il ne reste plus qu'à étudier le caractère gaussien des résidus. Ceci est effectué grâce au test de Shapiro-Wilk. Ce test a été préféré à celui de Kolmogorov-Smirnov à cause du faible nombre de données (50). Les résultats du test sont les suivants :

Tests de normalité				
Test	Statistique		p Value	
Shapiro-Wilk	W	0.951951	Pr < W	0.0517

Tab:5 Analyse de la normalité des résidus

Ainsi, au seuil 5 %, on accepte l'hypothèse de normalité des résidus de la régression linéaire.

La volatilité de ces résidus est estimée grâce à la formule :

$$\hat{\sigma}_\gamma = \frac{1}{n - p} \sum_{i=1}^{50} (\varepsilon_i - \bar{\varepsilon})^2 \quad (63)$$

où p représente le nombre de paramètres estimés dans la régression linéaire.

Ainsi, en ce qui concerne le volet prédictif du modèle, nous obtenons sur nos données

$$\hat{b} = 4059,94439, \hat{a} = -2,05775 \text{ et } \hat{\sigma}_\gamma = 3.98227882$$

L'ajustement sur les données historiques conduit à la surface de mortalité suivante :

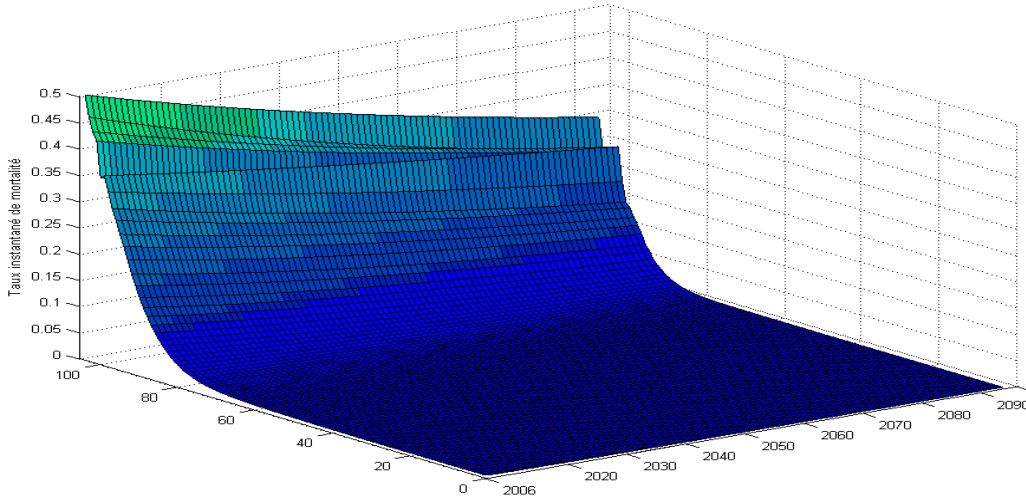


Fig. 12 : Surface de mortalité ajustée

## 1.2 APPLICATIONS NUMERIQUES

Au-delà des moments d'ordre un et deux de la distribution de  $\Lambda$ , qui peuvent être obtenus explicitement (cf. Magnin et Planchet [2000]), on s'intéresse ici à la distribution de l'engagement du régime. La méthode retenue consiste à simuler les durées de survie des rentiers,  $T_{x(j)}$ ,  $j \in J$  puis à calculer :

$$\Lambda = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \sum_{j \in J} r_j * \mathbf{1}_{]t; \infty[}(T_{x(j)}). \quad (64)$$

sur la base des réalisations des  $T_{x(j)}$ ,  $j \in J$  obtenues. On obtient ainsi des réalisations  $\lambda_1, \dots, \lambda_n$  de  $\Lambda$  et on détermine la fonction de répartition empirique de l'engagement, représentée ci-dessous (dans le cas de 20 000 simulations) :

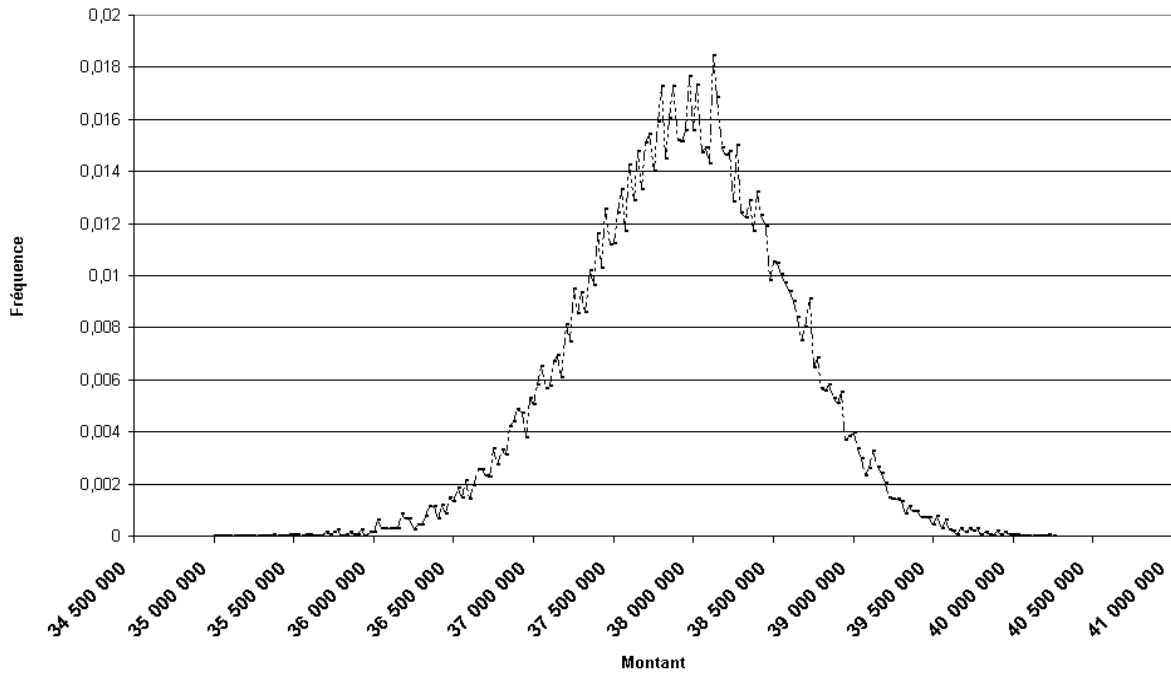


Fig. 13 : *Distribution empirique de l'engagement*

La provision  $L_0$  est approchée par  $\bar{\lambda} = \frac{1}{N} \sum_{n=1}^N \lambda_n$ .

La variance de l'engagement est estimée par  $\frac{1}{N-1} \sum_{n=1}^N (\lambda_n - L_0)^2$  ;

On s'intéresse plus particulièrement au coefficient de variation empirique :

$$cv = \frac{\sqrt{\frac{1}{N-1} \sum_{n=1}^N (\lambda_n - L_0)^2}}{\frac{1}{N} \sum_{n=1}^N \lambda_n}, \quad (65)$$

qui fournit un indicateur de la dispersion de l'engagement et dans une certaine mesure de sa « dangerosité ». On obtient, pour ces différents indicateurs, les valeurs suivantes :

	<b>Déterministe</b>
<b>Espérance</b>	37 973 994
<b>Ecart-type</b>	631 525
<b>Borne inférieure de l'intervalle de confiance</b>	36 700 000
<b>Borne supérieure de l'intervalle de confiance</b>	39 175 000
<b>Coefficient de variation</b>	1,66 %

On note, malgré la taille modeste de la population, que l'engagement est évalué avec une précision relative de plus ou moins 6,1 % (rapport de la demi-largeur de l'intervalle de confiance à l'espérance). Le risque démographique est, de ce point de vue, assez bien maîtrisé.

## 2. CAS D'UNE MORTALITE STOCHASTIQUE

On intègre ici au modèle une composante de risque systématique au travers d'une incertitude sur la surface de mortalité future, la surface construite à la section précédente définissant alors la tendance de référence autour de laquelle la mortalité observée sera supposée fluctuer.

### 2.1 MODELE DE MORTALITE

On utilise l'équation de régression qui a permis d'obtenir la tendance de mortalité future :

$$k_t^* = at + b + \gamma_t, \quad (66)$$

et l'on obtient des réalisations de la mortalité future en effectuant des tirages dans la loi du résidu  $(\gamma_t)$ ,  $N(0, \sigma_\gamma^2)$ . La variable  $k_t^*$  est donc gaussienne telle que  $E(k_t^*) = k_t$ . On obtient ainsi des réalisations des taux instantanés de sortie *via* :

$$\mu_{xt}^* = \exp(\alpha_x + \beta_x k_t^*). \quad (67)$$

Puisque  $k_t^*$  est gaussienne,  $\mu_{xt}^*$  est log-normale et son espérance se calcule selon

$$E(\mu_{xt}^*) = E \exp(\alpha_x + \beta_x k_t^*) = \exp\left(\alpha_x + \beta_x k_t + \frac{\beta_x^2 \sigma_\gamma^2}{2}\right) \text{ et donc :}$$

$$E\left(\mu_{xt}^*\right) = \mu_{xt} \exp\left(\frac{\beta_x^2 \sigma_\gamma^2}{2}\right) > \mu_{xt}. \quad (68)$$

Le modèle stochastique a donc tendance à surestimer les taux de sortie. Mais, comme l'illustre la figure 14, en pratique ce biais est faible et peu pénalisant pour le modèle, car peu significatif aux âges élevés. On représente sur la figure 14  $\frac{E\left(\mu_{xt}^*\right)}{\mu_{xt}}$  :

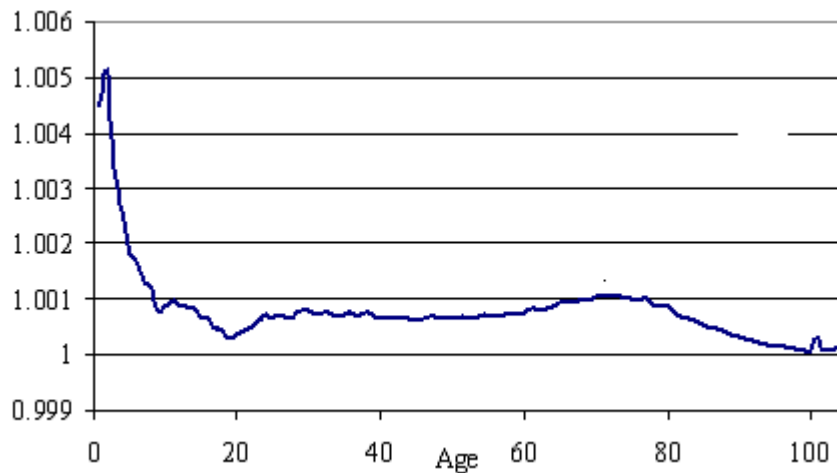


Fig. 14 : *Analyse du biais relatif en fonction de l'âge*

Incidentement, on ne peut montrer simplement si le biais qui en résulte sur les taux discrets  $q_x$  est positif ou négatif. En effet, l'inégalité de Jensen ( $f(EX) \leq Ef(X)$ ) appliquée à  $f(x) = -(1 - e^{-x})$  conduit à  $E(q_x^*) \leq 1 - \exp(-E\mu_x^*)$ , ce qui ne permet pas de conclure simplement sur le sens du biais. Cela indique toutefois que la relation entre  $q_x$  et  $\mu_x$  n'amplifie pas le biais, et donc que le biais sur le taux discret doit être également faible.

On observera que la volatilité  $\sigma_\gamma$  est un paramètre permettant de contrôler simplement le degré d'incertitude attaché à la mortalité future. Dans le cas particulier  $\sigma_\gamma = 0$ , on retrouve le modèle déterministe analysé à la section précédente.

Le biais généré sur les taux de sortie par le mode de simulation des surfaces de mortalité, s'il s'avère peu pénalisant au voisinage de la volatilité  $\sigma_\gamma$  estimée sur les données disponibles, pourrait être potentiellement gênant dans des situations de plus forte volatilité. Ce point sera



illustré par une application numérique. Aussi, nous utiliserons dans la suite une version corrigée du biais du modèle proposé définie par :

$$\mu_{xt}^* = \exp \left( \alpha_x - \frac{\beta_x^2 \sigma_\gamma^2}{2} + \beta_x k_t^* \right). \quad (69)$$

Cette version du modèle satisfait par construction  $E(\mu_{xt}^*) = \mu_{xt}$ . Elle apparaît donc plus cohérente avec l'objectif recherché de « perturber » la surface de mortalité, mais sous l'hypothèse que celle-ci définit correctement la tendance future des taux instantanés de décès.

## 2.2 ANALYSE DE LA REPARTITION DU RISQUE POUR LE REGIME DE RENTES

On retient comme mesure du risque la variance de la somme des flux futurs actualisés  $\Lambda$ . En conditionnant par la surface de mortalité  $\Pi$  et en utilisant l'équation de décomposition de la variance on obtient :

$$V[\Lambda] = E[V(\Lambda|\Pi)] + V[E(\Lambda|\Pi)]. \quad (70)$$

Le second terme du membre de droite de l'expression ci-dessus représente le risque systématique associé au régime de rentes ; le premier terme le risque technique, *i.e.* le risque de mortalité mutualisable. En pratique on retiendra comme indicateur la part de variance expliquée par la composante de risque systématique, soit :

$$\omega(\sigma_\gamma) = \frac{V[E(\Lambda|\Pi)]}{V[\Lambda]}. \quad (71)$$

Lorsque la taille du groupe tend vers l'infini,  $\omega(\sigma_\gamma)$  converge vers 1. En d'autres termes, pour un groupe parfaitement mutualisé, toute la variance est expliquée par la composante systématique. Le calcul direct de  $\omega(\sigma_\gamma)$  n'est pas aisé, aussi on utilise ici une approche par simulation. On sélectionne dans un premier temps, au moyen d'un tirage dans une loi  $N(0, \sigma_\gamma^2)$  une surface de mortalité (en pratique les réalisations de la gaussienne sont obtenues par inversion de la fonction de répartition en utilisant l'approximation de Moro présentée dans PLANCHET et al. [2005]) puis, conditionnellement à cette hypothèse de mortalité, on simule la survie des rentiers (« le passif »).

De manière plus formelle, si  $\lambda_{n,m}$  est la réalisation de  $\Lambda$  résultant de la  $n$ -ième trajectoire de la mortalité et de la  $m$ -ième trajectoire du passif, on note :

$$\bar{\lambda}_n = \frac{1}{M} \sum_{m=1}^M \lambda_{n,m} \quad \text{et} \quad \bar{\bar{\lambda}} = \frac{1}{N} \sum_{n=1}^N \bar{\lambda}_n = \frac{1}{NM} \sum_{n=1}^N \sum_{m=1}^M \lambda_{n,m}. \quad (72)$$

Les quantités ci-après sont des estimateurs sans biais et convergents respectivement de  $E[V(\Lambda|M)]$  et de  $V[E(\Lambda|M)]$  :

$$\hat{E}[V(\Lambda|M)] = \frac{1}{N} \sum_{n=1}^N \frac{1}{M-1} \sum_{m=1}^M (\lambda_{n,m} - \bar{\lambda}_n)^2, \quad (73)$$

$$\hat{V}[E(\Lambda|M)] = \frac{1}{N-1} \sum_{n=1}^N (\bar{\lambda}_n - \bar{\bar{\lambda}})^2. \quad (74)$$

Le calibrage du nombre de simulations est effectué de manière empirique en arrêtant l'algorithme lorsque les résultats se stabilisent (écart entre deux résultats successifs  $< 10^{-3}$ ).

## 2.3 APPLICATIONS NUMERIQUES

Les applications numériques proposées sont à deux niveaux :

- dans un premier temps, on analyse l'impact de la prise en compte de la mortalité stochastique sur la loi de l'engagement du régime considéré ;
- dans un second temps, on détermine l'évolution en fonction de la volatilité de la surface de mortalité (et donc de l'incertitude attachée à la prévision de cette dernière) de la part de variance expliquée par cette source de risque dans la variance totale.

### 2.3.1 Analyse de l'engagement

On détermine tout d'abord pour la valeur estimée de la volatilité de la surface de mortalité ( $\sigma_\gamma=3.98$ ) la distribution empirique de l'engagement, représentée ici avec la distribution de référence obtenue précédemment dans le cas déterministe (avec 20 000 tirages) :

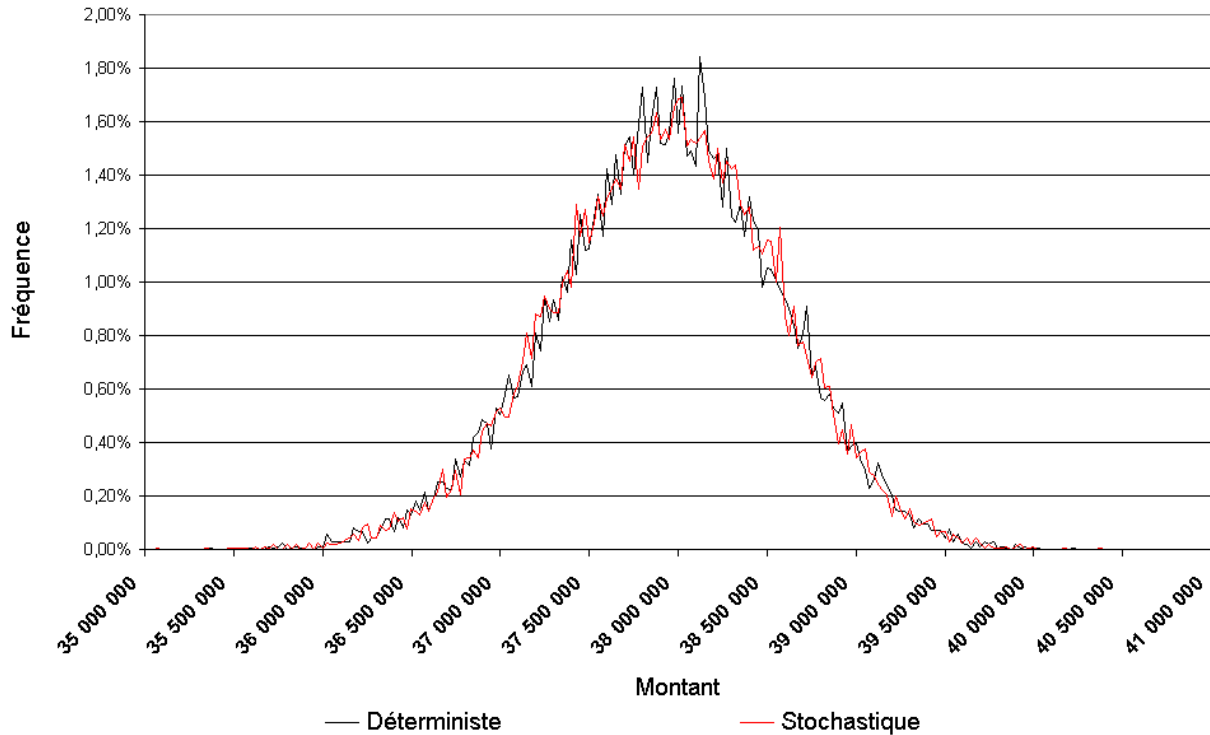


Fig. 15 : *Distribution empirique de l'engagement : approche déterministe VS approche stochastique*

On peut noter que le coefficient de variation de la distribution de l'engagement est estimé à 1,66 % dans les deux situations (mortalité déterministe et mortalité stochastique), alors que l'on aurait pu s'attendre à ce que la prise en compte de ce facteur de risque supplémentaire augmente sensiblement la dispersion, et donc la « dangerosité » de l'engagement. On constate ici une quasi-stabilité. Les résultats détaillés sont repris ci-après :

	<b>Déterministe</b>	<b>Stochastique</b>
Espérance	37 973 994	37 939 862
Ecart-type	631 525	628 266
Borne inférieure de l'intervalle de confiance	36 700 000	36 650 000
Borne supérieure de l'intervalle de confiance	39 175 000	39 100 000
Coefficient de variation	1.66 %	1.66 %

On note en particulier que l'espérance de l'engagement est identique dans le cas déterministe et dans le cas stochastique, ce qui illustre le faible biais induit par le mode de sélection des tables aléatoires au travers de  $(k_t)$ . Dans le cas présent, les données de référence utilisées pour calibrer la mortalité de référence fluctuent faiblement, ce qui conduit à un risque associé marginal au regard du risque d'échantillonnage. Ceci est d'autant plus marqué que la population concernée est d'effectif relativement faible.

Dans une approche « valeur à risque » (*VaR*), on trouve que le quantile à 75 % de la distribution de l'engagement est de 38.3M€ dans le cas stochastique, ce qui est quasiment identique à la valeur obtenue dans le cas déterministe.

Ces résultats sont, bien entendu, la conséquence directe de la très faible volatilité estimée sur la série ( $k_t$ ). Dans le cas où cette volatilité augmente, la distribution de l'engagement intégrant le risque systématique diverge rapidement de la situation de référence constituée par le modèle déterministe. A titre d'illustration, on présente ci-après les résultats avec une volatilité multipliée par 10 (le nombre de tirages reste fixé à 20 000) :

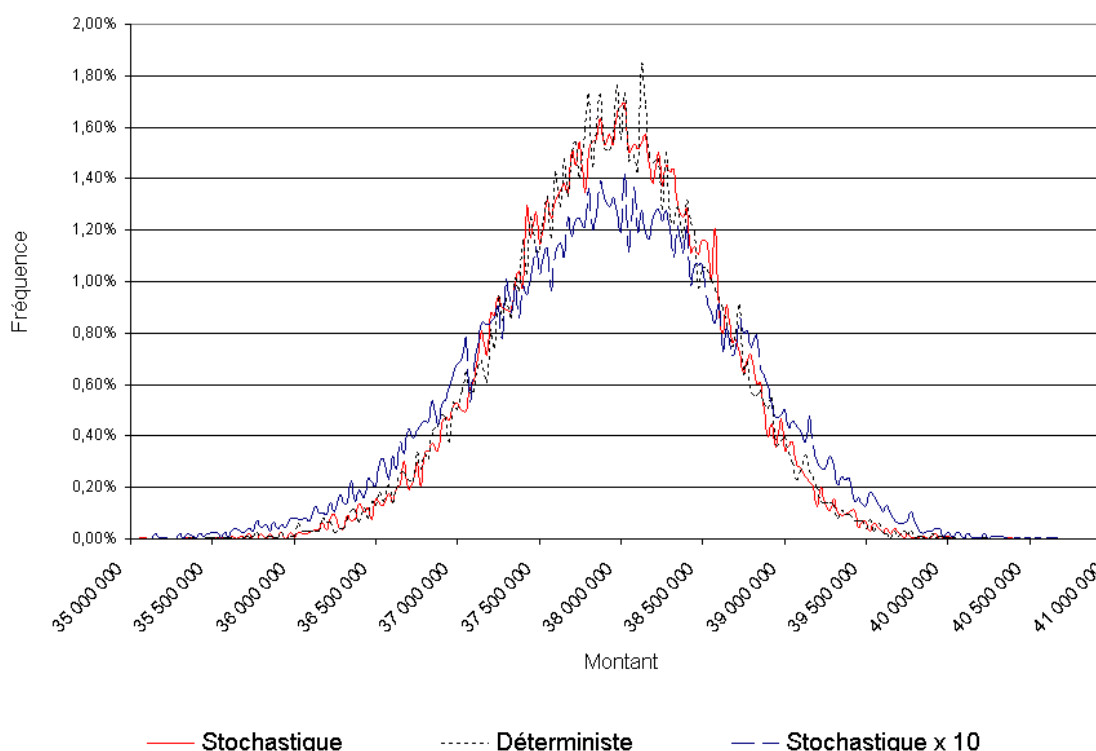


Fig. 16 : *Distribution empirique de l'engagement : approche déterministe VS approche stochastique*

On peut remarquer que le coefficient de variation de la distribution de l'engagement dans le cas de la prise en compte d'un caractère stochastique de la mortalité à forte volatilité est estimé à 2,06 %, contre 1,66 % dans la situation de référence. Le modèle, comme on pouvait s'y attendre, est très sensible à la volatilité de processus de génération des surfaces de mortalité stochastiques. Les résultats détaillés sont repris ci-après :

	<b>Déterministe</b>	<b>Stochastique</b>	<b>Stochastique x10</b>
Espérance	37 973 994	37 939 862	37 925 061
Ecart-type	631 525	628 266	778 276
Borne inférieure de l'intervalle de confiance	36 700 000	36 650 000	35 600 000
Borne supérieure de l'intervalle de confiance	39 175 000	39 100 000	39 375 000
Coefficient de variation	1.66 %	1.66 %	2.06 %

Bien entendu, toutes choses égales par ailleurs, l'impact sur la loi de l'engagement de la prise en compte de la composante non mutualisable du risque est d'autant plus important que les fluctuations d'échantillonnage sont faibles, et donc que le groupe est de taille importante. Afin d'illustrer ce point, les calculs de la distribution de l'engagement ont été menés sur un groupe fictif composé de 100 répliques du groupe de base, ce qui constitue un groupe de 37 400 rentiers. On obtient le graphique suivant :

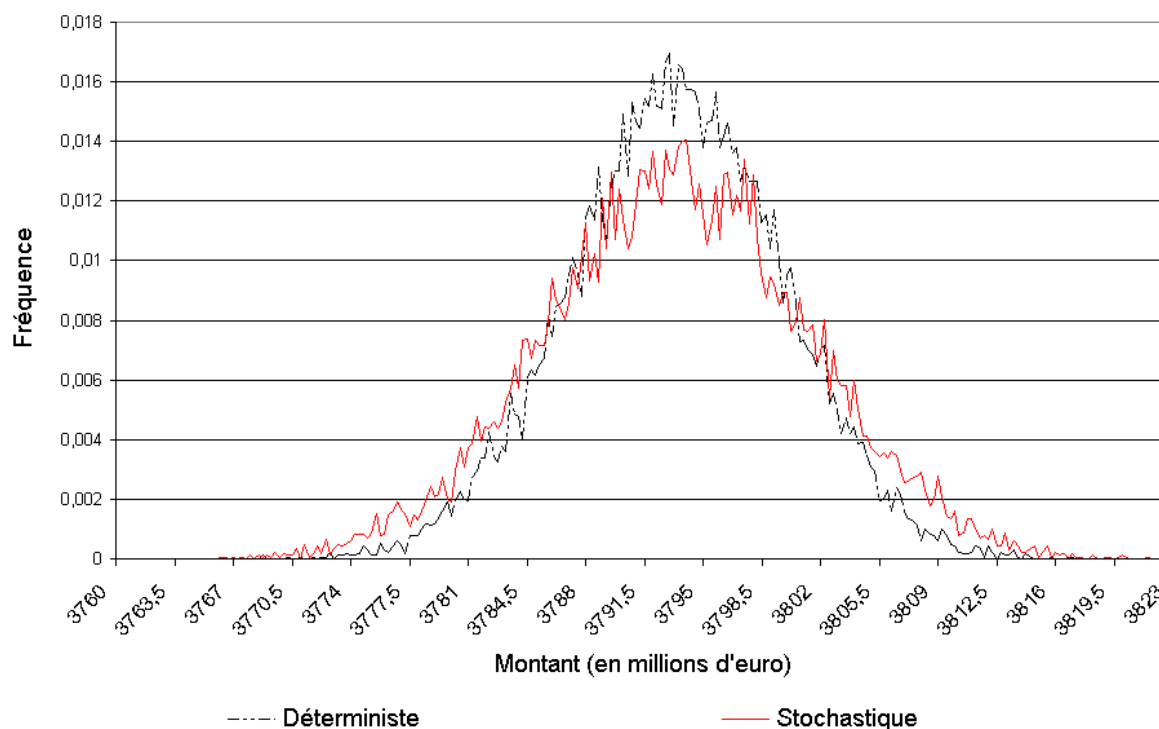


Fig. 17 : Distribution empirique de l'engagement : taille du groupe x100

On constate sur le graphique l'aplatissement de la distribution de l'engagement, illustrant une situation *a priori* plus risquée pour le régime lorsque la part non mutualisable du risque est prise en compte. Les résultats synthétiques sont détaillés ci-dessous :

	<b>Déterministe</b>	<b>Stochastique</b>
Espérance	3 793 577 165	3 793 566 113
Ecart-type	6 324 333	7 771 749
Quantile à 5 %	3 780 750 000	3 777 750 000
Quantile à 95 %	3 803 750 000	3 806 250 000
Coefficient de variation	0.167 %	0.205 %

On constate que si l'espérance de l'engagement est simplement multipliée par 100 par rapport à la situation de référence, on voit apparaître une différence sensible entre l'approche déterministe classique et l'approche stochastique : l'écart-type augmente en effet de 23 % lorsque l'on prend en compte le risque non mutualisable.

On peut toutefois noter que le risque démographique, même en présence du facteur de risque systématique, reste bien maîtrisé, l'intervalle de confiance à 95 % pour le niveau de l'engagement étant (en millions d'euros) [3778;3807], ce qui donne une précision dans la mesure de l'engagement supérieure à 99 %.

Dans une approche « valeur à risque » (*VaR*), on trouve que le quantile à 75 % de la distribution de l'engagement est de 3 798 M€ dans le cas stochastique, contre 3 797 M€ dans le cas déterministe. Là encore, on note un très faible impact absolu sur le régime.

### 2.3.2 Importance de la correction de biais

Dès lors que la volatilité du processus de mortalité sous-jacent devient significative, la correction de biais s'avère indispensable. Le graphique suivant, qui représente la distribution empirique avec une volatilité égale à 10 fois, puis 20 fois la volatilité de référence du modèle sans correction du biais, illustre ce point :

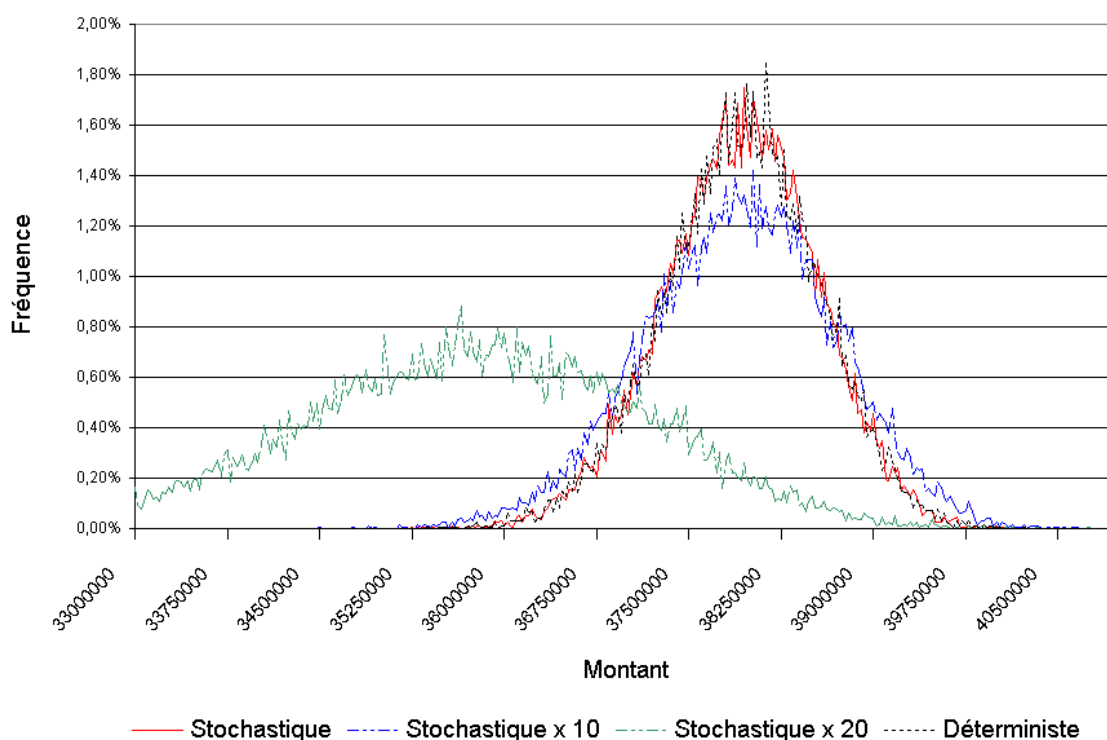


Fig. 18 : Distribution empirique de l'engagement : approche déterministe VS approche stochastique

On note que la distribution intégrant le risque systématique se trouve décalée vers la gauche du fait du biais sur les taux de sortie et qu'elle conduit ainsi à une vision minorée de l'engagement.

	<b>Stochastique</b>	<b>Stochastique x 10</b>	<b>Stochastique x 20</b>
Espérance	37 939 862	37 925 061	35 746 549
Ecart-type	628 266	778 276	1 403 005
Borne inférieure de l'intervalle de confiance	36 650 000	35 600 000	31 625 000
Borne supérieure de l'intervalle de confiance	39 100 000	39 400 000	38 325 000
Coefficient de variation	1.66 %	2.05 %	3.92 %

### 2.3.3 Analyse de la part de variance expliquée

Comme les résultats précédents l'ont montré, le comportement de l'engagement est fonction de la volatilité du processus stochastique de mortalité sous-jacent, qu'il importe donc de quantifier. Les estimateurs présentés *supra* nous permettent d'obtenir le graphique d'évolution de la part du risque systématique dans le risque global  $\omega(\sigma_\gamma)$  en fonction de la volatilité  $\sigma_\gamma$  du générateur de tables prospectives.

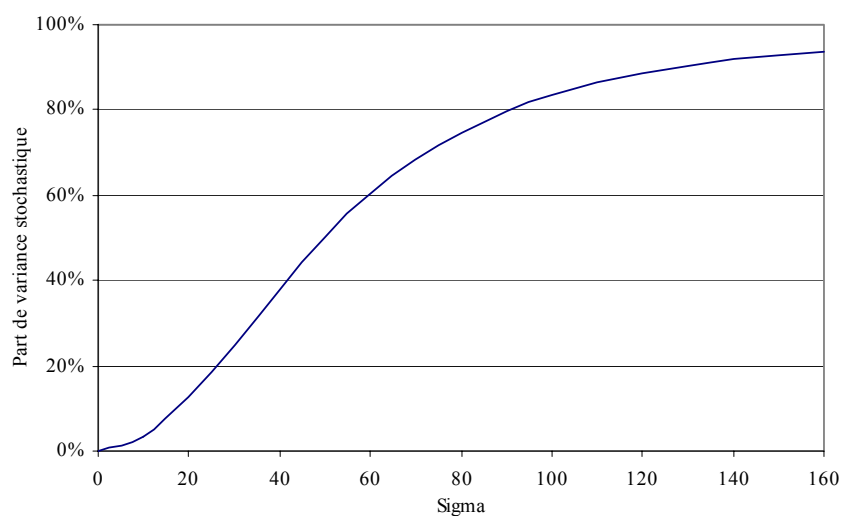


Fig. 19 : Evolution de la part du risque systématique dans le risque global en fonction de  $\sigma_\gamma$

On note que la croissance de la courbe est lente au début et ne s'accélère que pour les valeurs de la volatilité importantes.

La mortalité stochastique étant un risque systématique, son impact sera d'autant plus grand que la taille du portefeuille est grande.

A  $\sigma_\gamma$  fixé, la part de variance expliquée par la composante stochastique de la mortalité augmente avec la taille du portefeuille ; on obtient ainsi (la taille du portefeuille est exprimée en nombre de fois la taille du portefeuille de référence) :

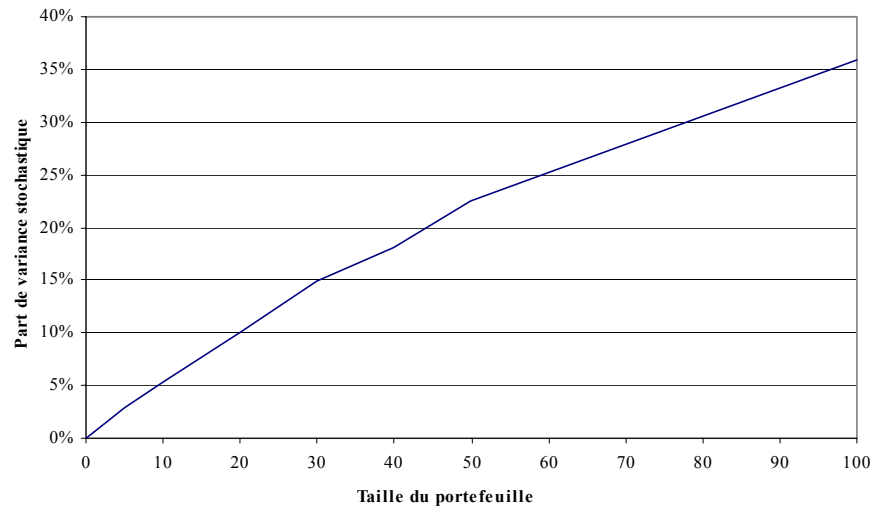


Fig. 20 : Evolution de la part du risque systématique dans le risque global en fonction de la taille du portefeuille

Le graphique précédent indique que pour un portefeuille de plus de 30 000 rentiers, l'impact de la mortalité stochastique représente plus de 30 % de la volatilité du portefeuille.

L'impact de la taille du portefeuille sur l'indicateur  $\omega(\sigma_\gamma)$  est un élément essentiel.

Cependant, il peut s'avérer très long à obtenir par simulation. Nous allons donc rechercher une formule analytique permettant d'obtenir directement le graphique 20.

Partons de l'égalité :

$$V[\Lambda] = E[VAR[\Lambda|\Sigma]] + VAR[E[\Lambda|\Sigma]]$$

avec

$$\Lambda = \sum_{t=1}^{\infty} \tilde{F}_t (1+i)^{-t} = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \sum_{j \in J} r_j * \mathbf{1}_{]t; \infty[}(T_{x(j)})$$

Lorsque l'on multiplie la taille du portefeuille par n,  $\Lambda$  devient  $n\Lambda$ . Comme les individus sont considérés comme indépendants on a :

- $VAR[n\Lambda|\Sigma] = nVAR[\Lambda|\Sigma]$
- $E[n\Lambda|\Sigma] = nE[\Lambda|\Sigma]$



Ainsi en notant  $f$  la fonction donnant l'évolution de  $\omega(\sigma_\gamma)$  en fonction de la taille du portefeuille, on a :

$$f(n) = \frac{n^2 \text{VAR}[E[\Lambda|\Sigma]]}{n^2 \text{VAR}[E[\Lambda|\Sigma]] + nE[\text{VAR}[\Lambda|\Sigma]]} = \frac{n \text{VAR}[E[\Lambda|\Sigma]]}{n \text{VAR}[E[\Lambda|\Sigma]] + E[\text{VAR}[\Lambda|\Sigma]]}, \quad (75)$$

Le graphique 20 représentant l'évolution de la part du risque systématique dans la variance du portefeuille peut s'obtenir par application de la fonction  $f$ .

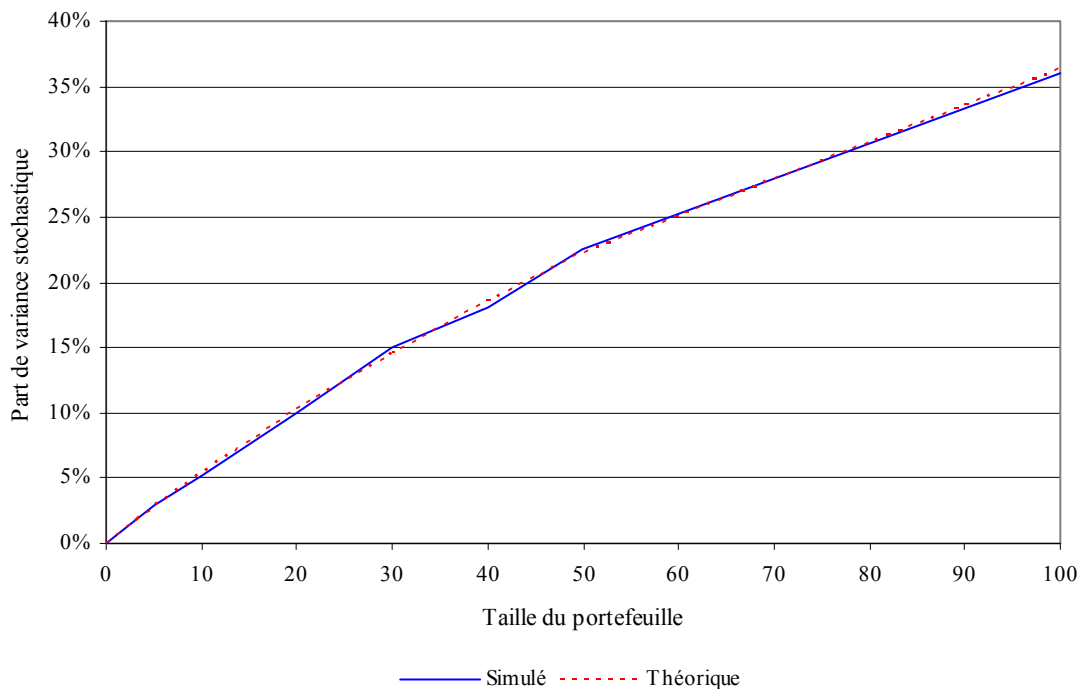


Fig. 21 : Evolution théorique de la part de variance stochastique

Ainsi, à partir du calcul de  $\omega(\sigma_\gamma)$  pour le portefeuille de taille un, on peut calculer l'évolution de la part de variance stochastique en fonction de la taille du portefeuille. La variance du portefeuille sera expliquée à plus de 90 % par la mortalité stochastique pour un portefeuille de plus de 510 000 rentiers.

Cependant, le coefficient  $\omega(\sigma_\gamma)$  est un indicateur fiable uniquement sur les portefeuilles de petite taille. En effet, plus le portefeuille est important plus l'impact de la volatilité sur l'intervalle de confiance à 95 % devient faible.

Le graphique suivant indique l'impact de la volatilité globale du portefeuille sur l'intervalle de confiance à 95 %. Passé 2300 rentiers, la largeur de l'intervalle de confiance à 95 % est inférieure à 2.81 %.

Ainsi, pour les portefeuilles de rentiers de faible taille, la non prise en compte de la mortalité stochastique n'a qu'un faible impact sur l'intervalle de confiance à 95 %.

Lorsque la taille du portefeuille augmente, la mortalité stochastique a un impact certain sur l'intervalle de confiance.

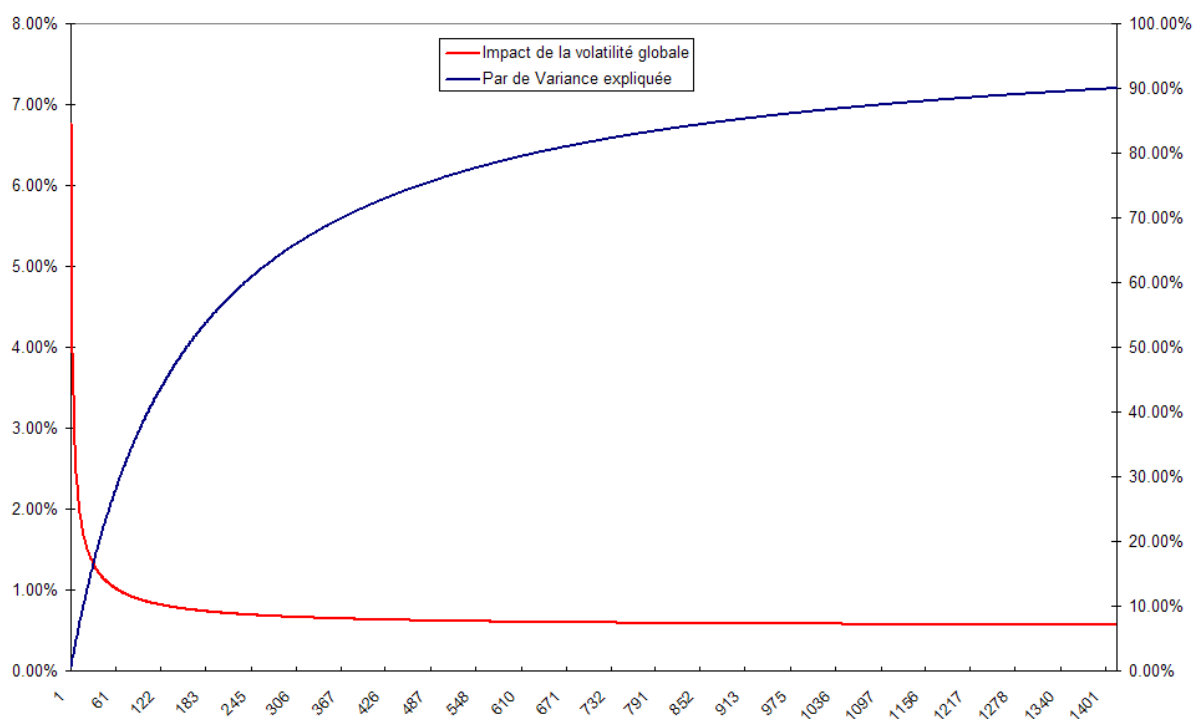


Fig. 22 : *Impact de la volatilité sur l'intervalle de confiance en fonction de la taille du portefeuille*

Ainsi, modéliser la mortalité stochastique à l'aide d'une simple régression linéaire conduit à conclure que la mortalité stochastique n'a pas d'impact sur le niveau de provisionnement. La dérive de la mortalité serait donc totalement linéaire, ce qui s'oppose aux constatations que l'on peut faire si l'on compare l'évolution du montant de provision lors du passage de la TPG 93 à la TGF 05.

On peut cependant remarquer que l'impact de la mortalité stochastique n'est pas identique à tout âge. Le graphique suivant nous montre qu'un portefeuille de taille identique au précédent mais constitué de rentiers de 60 ans ou plus serait plus sensible à la mortalité stochastique.

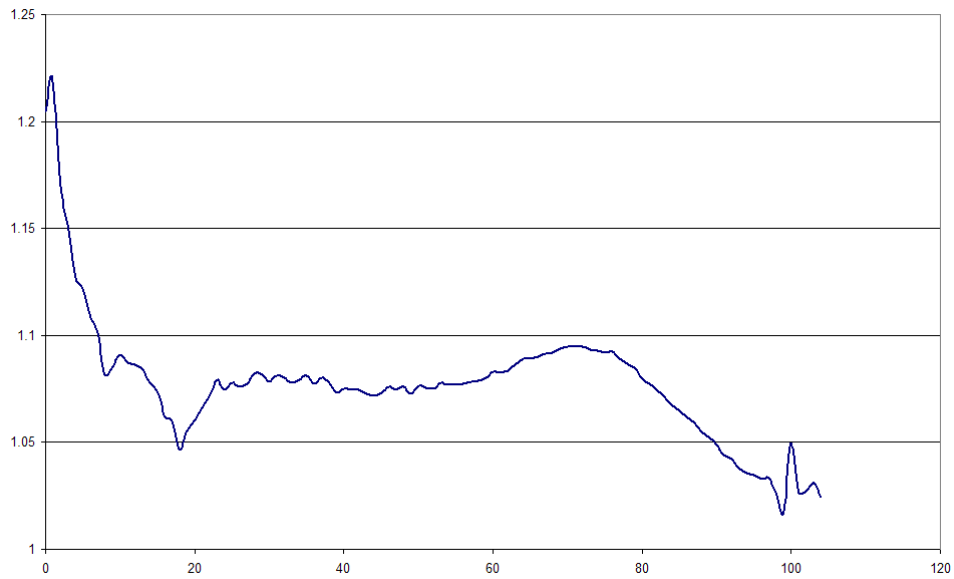


Fig. 23 : *Impact de la mortalité stochastique en fonction de l'âge*

Quoi qu'il en soit, les résultats précédents nous incitent à considérer que si la dérive de la mortalité présente une tendance linéaire, on ne peut expliquer la volatilité en ne considérant que de simples oscillations autour de sa tendance.

Cela nous conduit à envisager l'hypothèse d'une erreur de modèle au niveau de la régression linéaire.

### 3. MESURE DU RISQUE D'ERREUR DE DERIVE

Les analyses prospectives de mortalité conduisent à anticiper les évolutions futures des taux de décès aux différents âges. Dans les modèles maintenant classiques de construction de tables prospectives, comme le modèle de Lee-Carter (voir notamment LEE et CARTER [1992], LEE [2000], SITHOLE et al. [2000]) ou les modèles poissonniens (*cf.* BROUHNS et al. [2002] et PLANCHET et THEROND [2006] pour une présentation et une discussion de ces modèles), la dérive de mortalité future est anticipée à partir des observations passées.

Même en admettant qu'il est légitime de prolonger dans les années à venir les tendances observées par le passé (on pourra se reporter à CAREY et TULAPURKAR [2003] pour des analyses intégrant des considérations biologiques et environnementales, ainsi que GUTTERMAN et VANDERHOOF [1999] pour une discussion sur ce point), plusieurs sources d'incertitude viennent perturber la détermination de la tendance future : choix de la période d'observation, fluctuations stochastiques des taux de mortalité, événements exceptionnels, etc.

Cette incertitude fait peser sur les assureurs de rentes viagères et les régimes de retraite un risque systématique (non mutualisable) dont l'impact financier peut être très important.

Le régime considéré ici est en fait soumis à deux risques distincts. Le premier, dont l'étude a été effectuée dans la partie précédente, est la conséquence des fluctuations aléatoires des taux de mortalité futurs autour de la tendance définie par la table prospective. Le second est, quant à lui, associé à l'incertitude sur cette tendance.

Il est délicat d'évaluer *a priori* l'erreur de modèle ; une première approche consiste par exemple à tester la robustesse du modèle proposé en effectuant les estimations sur différentes plages temporelles. Ces estimations conduisent à des estimations de la mortalité future qui peuvent être sensiblement divergentes.

L'imprécision attachée aux coefficients de l'interpolation eux-mêmes est plus aisée à quantifier. En effet, dans le modèle :

$$k_t^* = at + b + \gamma_t, \quad (76)$$

les coefficients  $a$  et  $b$  sont estimés par la méthode des moindres carrés ordinaires, qui conduit à ce que le couple  $(\hat{a}, \hat{b})$  soit un vecteur gaussien. On peut donc simuler différentes droites d'extrapolation et intégrer cette source de variabilité dans le modèle.

### 3.1 LE MODELE DE MORTALITE

Afin de simplifier l'écriture des formules à venir, on pose  $\tau = t - t_m + 1$  et  $T = t_M - t_m + 1$ , ce qui conduit aux expressions :

$$\hat{a} = \frac{\frac{1}{T} \sum \tau k_\tau - \frac{T+1}{2} \bar{k}}{T^2 - 1} \text{ et } \hat{b} = \bar{k} - \frac{T+1}{2} \hat{a}, \quad (77)$$

avec  $\bar{k} = \frac{1}{T} \sum k_\tau = \frac{1}{T} \sum k_t$ . De plus, le vecteur  $(\hat{a}, \hat{b})$  est distribué selon une loi normale d'espérance  $(a, b)$  et de variance :

$$\Sigma = \frac{12\sigma_\gamma^2}{T(T^2 - 1)} \begin{bmatrix} 1 & -\frac{T+1}{2} \\ -\frac{T+1}{2} & \frac{(T+1)(2T+1)}{6} \end{bmatrix} \quad (78)$$

On peut donc construire des réalisations de la mortalité future en effectuant des tirages dans la loi du vecteur  $(\hat{a}, \hat{b})$ . La variable  $k_t^*$  ainsi obtenue est telle que  $E(k_t^*) = k_t$ . On obtient alors des réalisations des taux instantanés de sortie *via* :

$$\mu_{xt}^* = \mathbf{exp}(\alpha_x + \beta_x k_t^*). \quad (79)$$

Comme  $k_t^* = \hat{a}t + \hat{b}$  est une variable gaussienne d'espérance  $k_t = at + b$  et de variance :

$$\sigma_\tau^2 = \frac{12\sigma_\gamma^2}{T(T^2 - 1)} \left( \tau^2 - \tau(T+1) + \frac{(T+1)(2T+1)}{2} \right), \quad (80)$$

on note que :

$$E(\mu_{xt}^*) = E(\mathbf{exp}(\alpha_x + \beta_x k_t^*)) = \mathbf{exp} \left( \alpha_x + \beta_x k_t + \frac{\beta_x^2 \sigma_t^2}{2} \right), \quad (81)$$

et donc :

$$E\left(\mu_{xt}^*\right) = \mu_{xt} \exp\left(\frac{\beta_x^2 \sigma_t^2}{2}\right) > \mu_{xt}. \quad (82)$$

Le modèle stochastique a donc tendance à surestimer les taux de sortie par rapport à la surface de référence fournie par le modèle de Lee-Carter. Compte tenu de notre objectif de « perturber » la surface de mortalité, mais sous l'hypothèse que celle-ci définit correctement la tendance future espérée des taux instantanés de décès, cette propriété du modèle est pénalisante et il convient d'adapter l'approche proposée.

Une version corrigée du biais du modèle définie par :

$$\mu_{xt}^* = \exp\left(\alpha_x - \frac{\beta_x^2 \sigma_t^2}{2} + \beta_x k_t^*\right) \quad (83)$$

Cette version du modèle satisfait par construction  $E\left(\mu_{xt}^*\right) = \mu_{xt}$ . Elle apparaît donc cohérente avec l'objectif recherché. Toutefois, on peut lui reprocher de déformer de manière arbitraire la distribution des taux stochastiques. En effet, si le modèle  $k_t^* = \hat{a}t + \hat{b}$  est pertinent, alors les taux de sortie  $\mu_{xt}^*$  réellement observés seront bien issus du modèle  $\mu_{xt}^* = \exp\left(\alpha_x + \beta_x k_t^*\right)$  et non de la version corrigée du biais.

On utilise ici une approche différente et *a priori* plus naturelle consistant à utiliser comme surface de référence déterministe la surface moyenne du modèle stochastique,

soit  $E\left(\mu_{xt}^*\right) = \exp\left(\alpha_x + \beta_x k_t + \frac{\beta_x^2 \sigma_t^2}{2}\right)$ . En effet, si le mécanisme aléatoire dont sont issus les

taux de décès est bien associé à  $k_t^* = \hat{a}t + \hat{b}$ , alors la mortalité de référence déterministe est bien définie par la surface moyenne ci-dessus, et non plus par la surface de Lee-Carter.

### 3.2 APPLICATIONS NUMERIQUES

On présente ici les résultats obtenus tout d'abord sur la famille de tables prospectives proposée puis, dans un second temps, les conséquences en terme de valorisation de l'engagement du régime de rentes.

Avant toute chose, il est nécessaire de comparer la surface de mortalité moyenne qui va être utilisée avec celle que nous utilisons jusqu'à présent. En effet il faut s'assurer que nous n'allons pas (dans le but de mieux appréhender la volatilité de l'engagement) biaiser fortement l'engagement moyen de notre portefeuille.

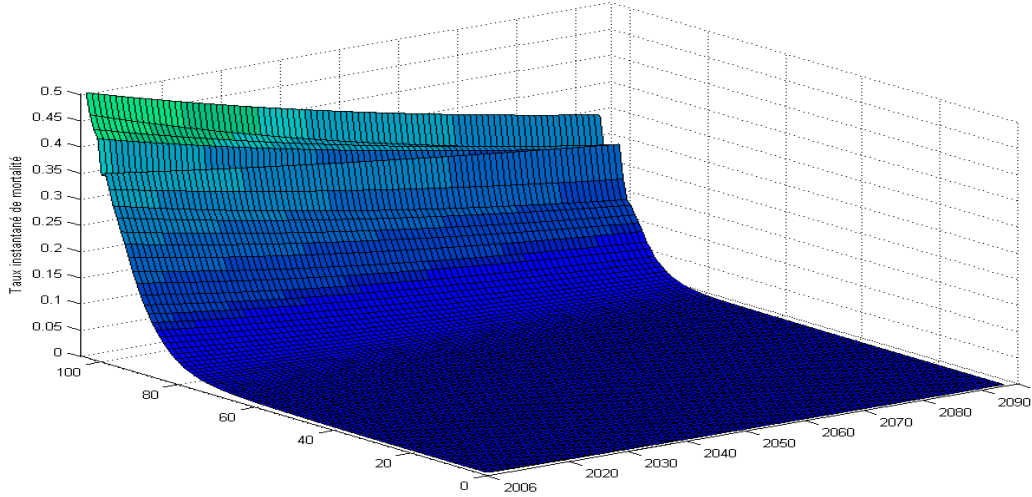


Fig. 24 : *Surface de mortalité ajustée par Lee-Carter*

La surface de référence ajustée de notre modèle stochastique est présentée sur le graphe ci-dessous :

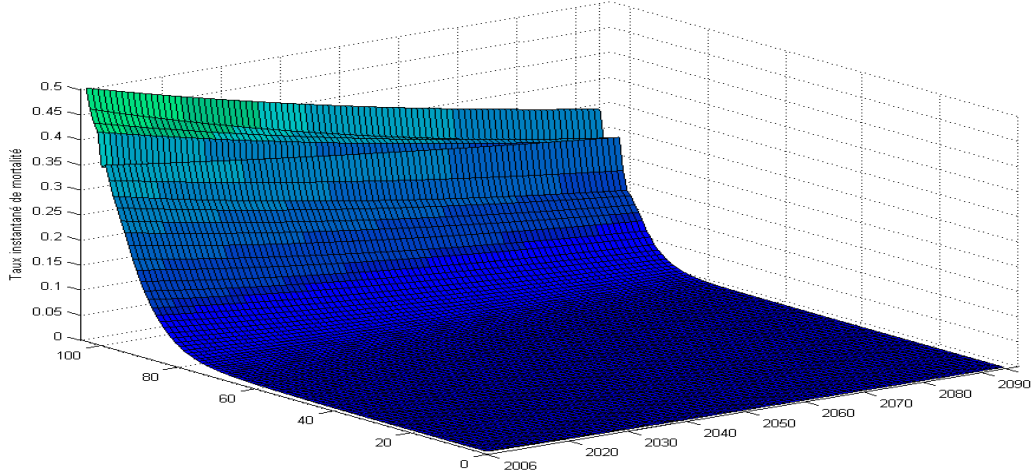


Fig. 25 : *Surface de mortalité de référence du modèle stochastique*

On remarque que la différence entre les deux graphes est très faible. En effet, dû à la faible valeur des coefficients  $\beta(x)$ ,  $\exp\left(\frac{\beta_x^2 \sigma_t^2}{2}\right)$  est proche de un. Ainsi, la modification de la surface de référence n'entraînera pas une différence notable au niveau de l'engagement déterministe.

En ce qui concerne le volet prédictif du modèle, nous obtenons sur nos données

$$\hat{b} = 49.38604, \hat{a} = -2.05775 \text{ et } \hat{\sigma}_\gamma = 3.98227882$$

$$\hat{\sigma}_b = 1.18058, \hat{\sigma}_a = 0.04282 \text{ et } \hat{\sigma}_t^2 = 1.39367413 + 0.00183378t^2 - 0.08802152t$$

Une centaine de tirages de trajectoires de  $k_t$  permet d'obtenir le graphe suivant :

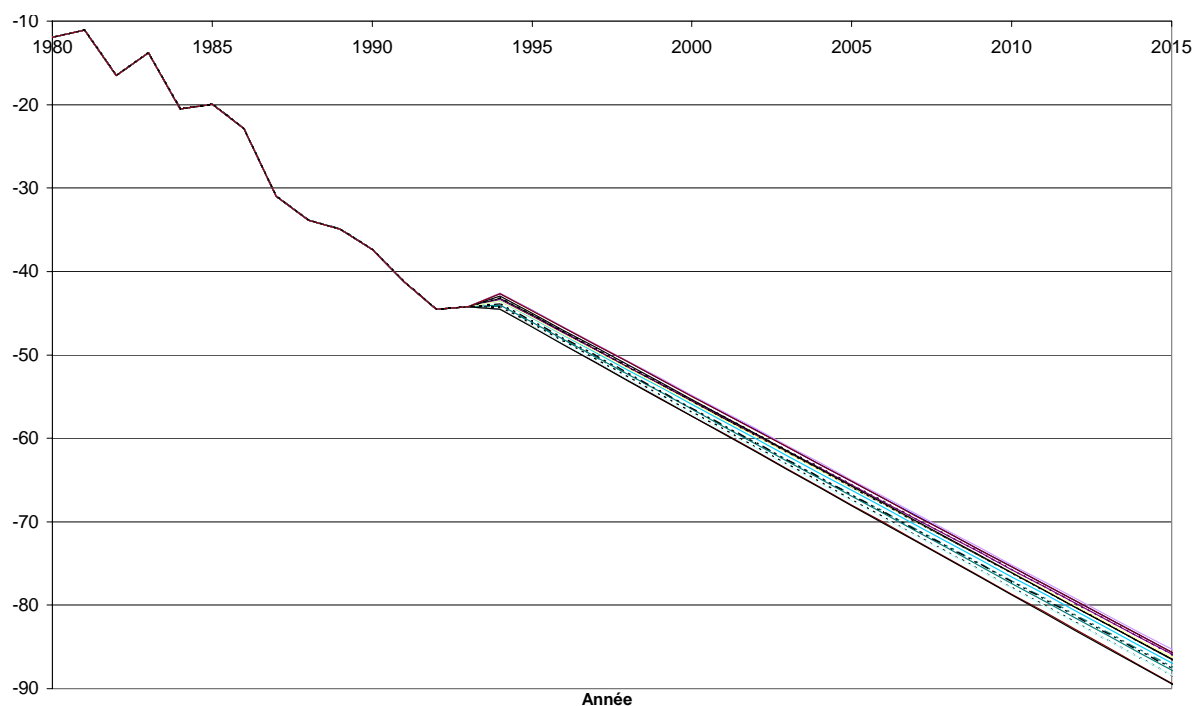


Fig. 26 : Simulation de trajectoires de la tendance

On constate sur le graphique que les différentes trajectoires simulées de  $k_t$  prennent quasiment la même valeur en 2006. Ceci montre que le modèle tire sa volatilité de la dérive de la mortalité et non d'un saut brutal de mortalité entre 2005 et 2006. Ainsi, la volatilité de l'engagement stochastique que nous allons observer, ne sera pas due à de brutales translations



de la mortalité (le modèle mis en place ne peut modéliser ce type de phénomène). Le graphique précédent permet donc de valider le modèle de « risque de dérive de la tendance ».

Si l'on cherche à comparer la volatilité de la dérive temporelle modélisée dans cette partie avec celle du modèle utilisé dans la partie précédente, on obtient le graphique suivant :

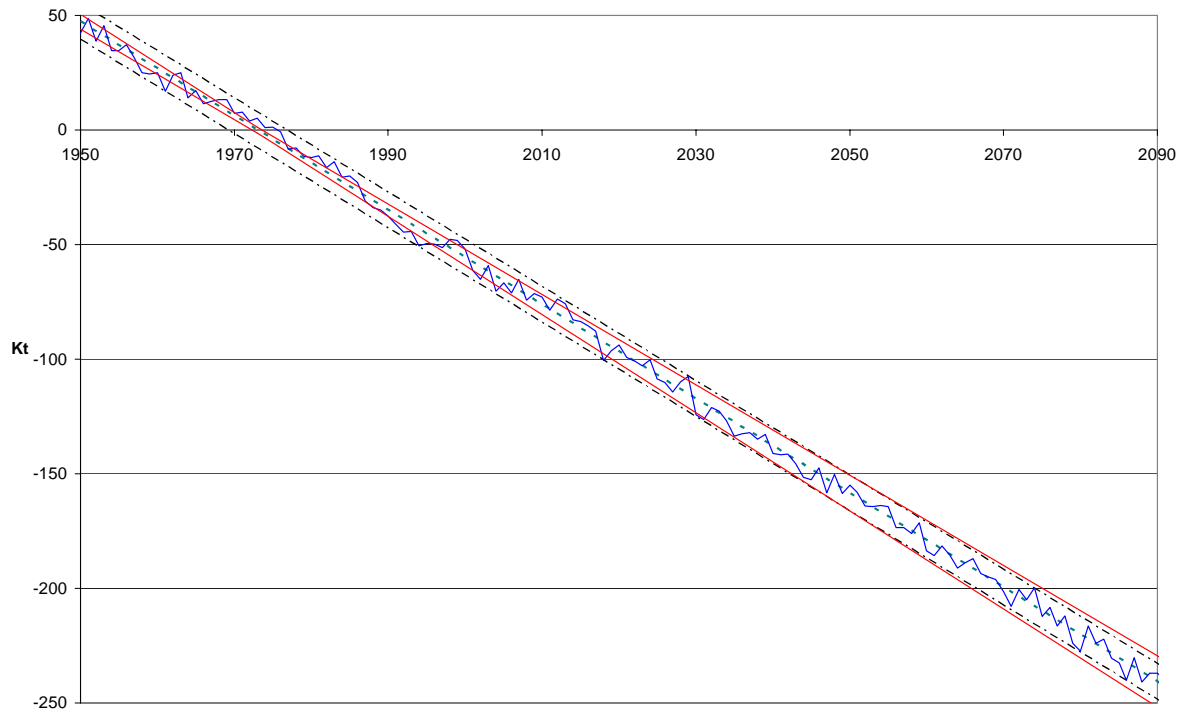


Fig. 27 : « couloir de variation » de la tendance

Alors que dans la partie précédente, le modèle utilisé ne prenait en compte que l'oscillation de la dérive temporelle ( $k_t$ ) autour de sa moyenne, le modèle actuel simule l'erreur d'ajustement qui a pu être faite sur la dérive temporelle passée.

Si dans le modèle linéaire précédent, les différentes trajectoires de  $k_t$  étaient situées entre deux droites parallèles délimitant le couloir de variation de la tendance à 95 %, on constate sur le graphique ci-dessus que les trajectoires de la dérive temporelle sont situées entre deux droites s'éloignant l'une de l'autre avec le temps. Les variations de la tendance temporelle et celles du taux instantané de mortalité étant comparables, le graphique ci-dessus traduit l'hypothèse du modèle : la mortalité stochastique tire sa volatilité non pas des fluctuations annuelles qui, comme nous l'avons démontré précédemment restent faibles, pour le moins anecdotiques, mais plutôt de la dérive aléatoire de la mortalité qui n'est vraiment observable que dans le futur.

### 3.2.1 Application à un régime de rentiers

La distribution empirique de l'engagement, représentée ici avec la distribution de référence dans le cas déterministe (avec 20 000 tirages) est présentée ci-dessous :

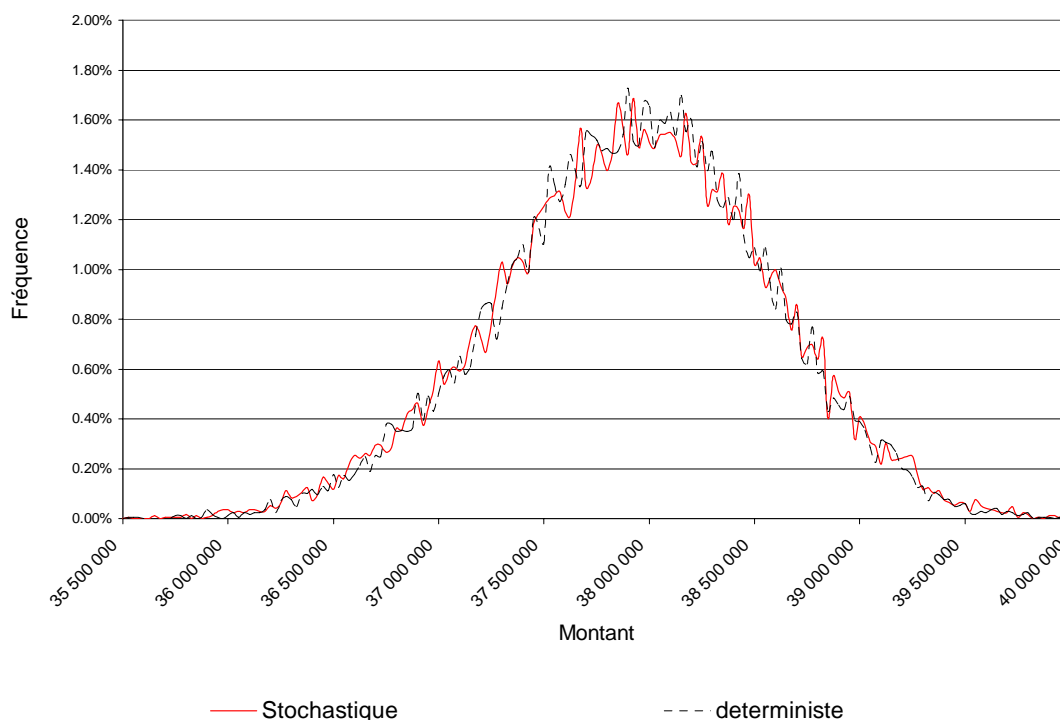


Fig. 28 : Distribution empirique de l'engagement

Les résultats détaillés sont repris ci-après :

	Déterministe surface Lee Carter	Déterministe	Stochastique
Espérance	37 937 707	37 937 707	37 937 720
Ecart-type	626 918	626 918	645 601
Borne inférieure de l'intervalle de confiance	36 625 000	36 625 000	36 600 000
Borne supérieure de l'intervalle de confiance	39 075 000	39 075 000	39 150 000
Coefficient de variation	1.65 %	1.65 %	1.70 %

On remarque tout d'abord que la modification de la surface de référence que nous utilisons jusqu'alors n'a pas d'impact. Ceci est dû au fait que pour tout  $x$  et pour tout  $t$ ,  $\exp\left(\frac{\beta_x^2 \sigma_t^2}{2}\right)$  est proche de un.

Sur un portefeuille de petite taille, l'impact de la mortalité stochastique sur l'engagement ne semble pas très important. En effet, le coefficient de variation de l'engagement stochastique n'est que 2.9 % plus élevé que celui de l'engagement déterministe.

Cependant, comme nous l'avons montré dans le chapitre précédent, la taille du portefeuille est un paramètre important à prendre en compte. En effet, si le risque systématique ne dépend pas de la taille du portefeuille, il n'en va pas de même pour le risque mutualisable.

En supposant tous les individus du portefeuille indépendants et les montants de rente unitaires, la volatilité de l'engagement stochastique de l'engagement non actualisé est donnée par la formule suivante :

$$E[VAR[\Lambda|\Sigma]] + VAR[E[\Lambda|\Sigma]] \quad (84)$$

Or si l'on pose  $P_{x,t}(\Sigma) = E[\mathbf{1}_{]0;+\infty[} T(xt) | \Sigma]$  et  $\Lambda(i) = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} \sum_{j \in J} r_j \mathbf{1}_{]t;+\infty[}(T_{x(j)})$ , cette formule peut se réécrire :

$$E[VAR[\Lambda(i) | \Sigma] + VAR\left[\sum_{i \in J} \left(\sum_{k=0}^{\infty} P_{x(i),j}(\Sigma)\right)\right]] \quad (85)$$

De plus on a :

$$P_{x,t}(\Sigma) = E\left[\frac{P_{x,t}(\Sigma)}{\exp(\beta_x \sigma_t^2)}\right]^{\varepsilon_{x,t}} \quad (86)$$

avec  $\varepsilon_{x,t}$  variable aléatoire suivant une loi lognormale  $LN(0, \beta_x \sigma_t^2)$ , représentant la dérive de la probabilité de survie annuelle.

Comme les graphiques 24 et 25 l'ont montré,  $\exp\left(\frac{\beta_x \sigma_t^2}{2}\right)$  est proche de un, donc la formule

86 peut se réécrire :

$$P_{x,t}(\Sigma) = E \left[ P_{x,t}(\Sigma) \right]^{\varepsilon_{x,t}} \quad (87)$$

Or à 95 %,  $\varepsilon_{x,t}$  est située dans l'intervalle de confiance suivant :

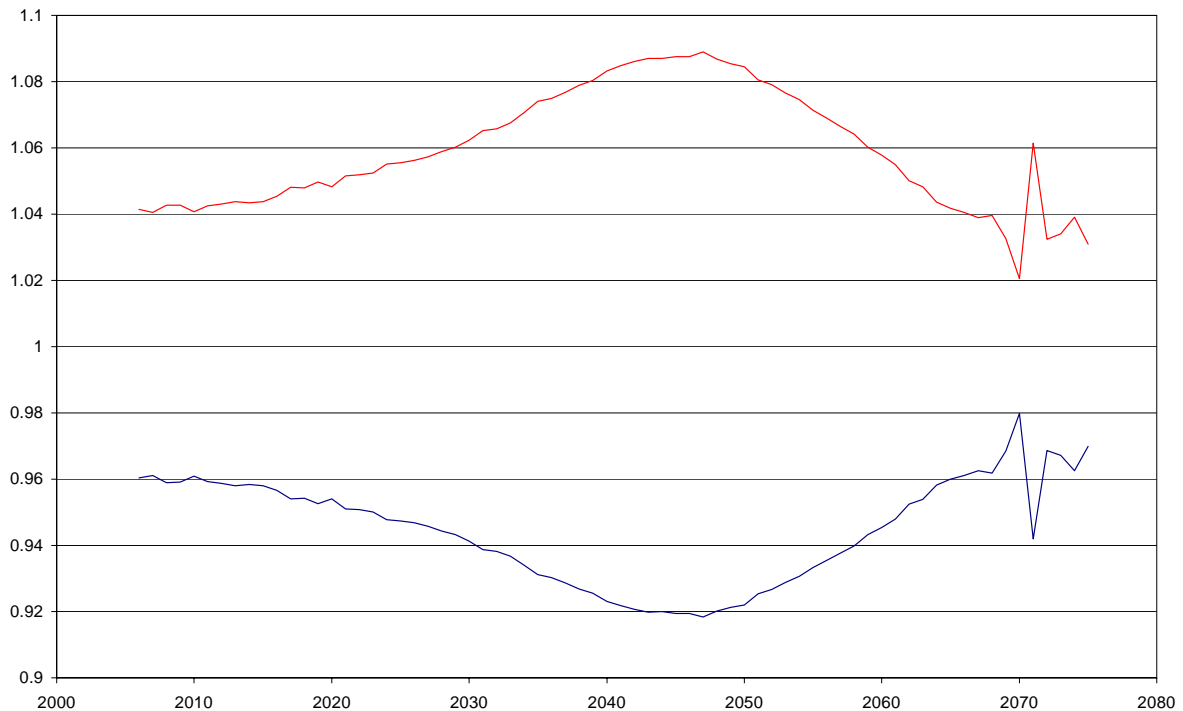


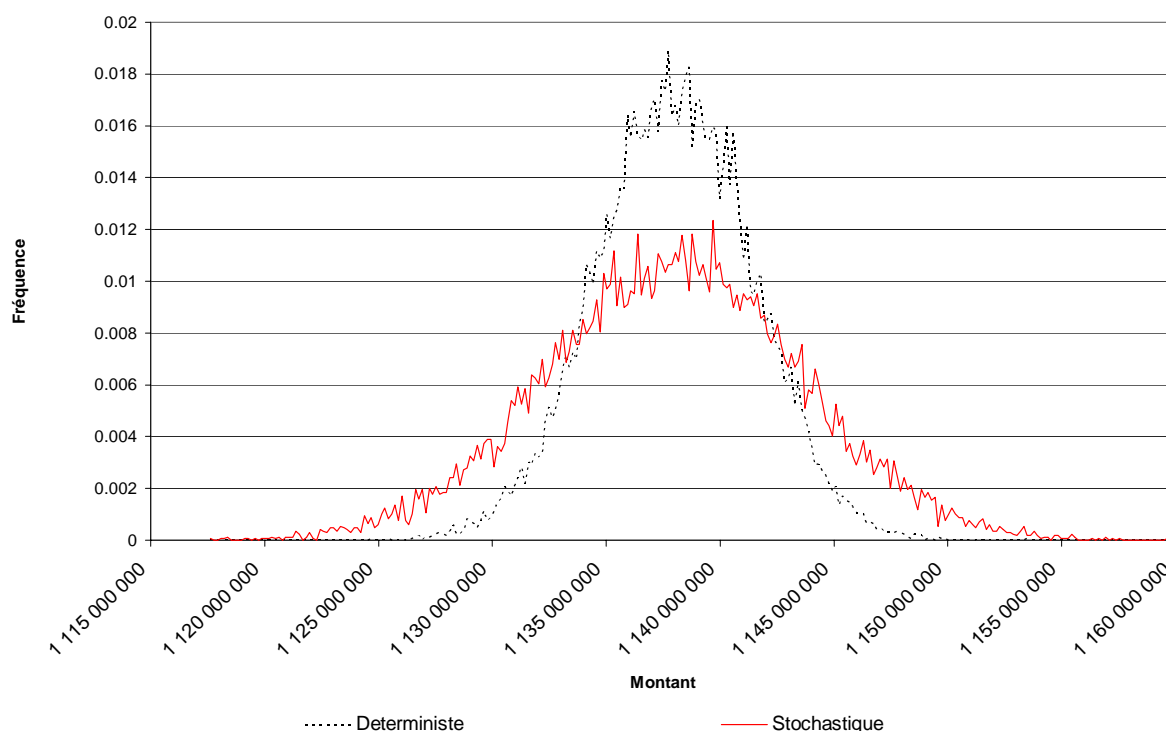
Fig. 29 : Intervalle de confiance de la variation de la dérive de la probabilité de survie annuelle

Les variations des deux courbes sont dues aux variations des coefficients  $\beta(x)$ .

Le graphique précédent montre que l'amplitude de la suite  $\left( P_{x,t}(\Sigma) \right)_{x,t}$  est très faible.

Ainsi, tant que le portefeuille n'est pas correctement mutualisé, l'impact du risque systématique est trop faible comparé à celui du risque mutualisable. La volatilité de l'engagement stochastique vaut alors  $E[VAR[\Lambda(i) | \Sigma]]$ , ce qui est égal à la volatilité de l'engagement déterministe. Ainsi, la faible taille du portefeuille ne permet pas de mesurer l'impact de la dérive de la mortalité sur la distribution de l'engagement.

Si l'on multiplie le portefeuille par 30, on obtient les résultats suivants :



Les résultats détaillés sont repris ci-après :

	<b>Déterministe</b>	<b>Stochastique</b>
Espérance	1 138 076 960	1 138 008 113
Ecart-type	3 410 560	3 410 560
Borne inférieure de l'intervalle de confiance	1 131 130 658	1 126 780 658
Borne supérieure de l'intervalle de confiance	1 144 480 658	1 148 830 658
Coefficient de variation	0.30 %	0.49 %

La prise en compte de la mortalité stochastique augmente le coefficient de variation de 63 %.

Dans une approche « valeur à risque » (*VaR*), on trouve que le quantile à 75 % de la distribution de l'engagement est de 1 142 M€ dans le cas stochastique, ce qui est supérieur à la valeur obtenue dans le cas déterministe, soit 1 140 M€. En d'autres termes, la prise en compte du risque de dérive conduit ici (en suivant une approche *VaR* pour le calcul de la provision) à augmenter le montant provisionné de 0,17 %.

La prise en compte du risque de dérive a également pour conséquence de doubler l'imprécision dans l'évaluation de l'engagement<sup>2</sup> qui, au niveau de confiance de 95 %, passe de 0,7 % environ à près de 1,3 %

<sup>2</sup> L'imprécision est mesurée par la demi-longueur relative de l'intervalle de confiance à 95 %.

Le graphique du « couloir de variation » de la tendance ayant mis en évidence le faible impact de la dérive de la mortalité dans le futur proche, le coefficient d'actualisation va jouer un rôle important dans la volatilité de la distribution de l'engagement.

En effet, l'engagement stochastique tire sa volatilité des rentes qui vont être servies pendant longtemps. Or plus la rente à une durée de vie importante, plus le coefficient d'actualisation qui s'y applique est important.

En fixant le taux d'actualisation à 0, on obtient les résultats suivants :

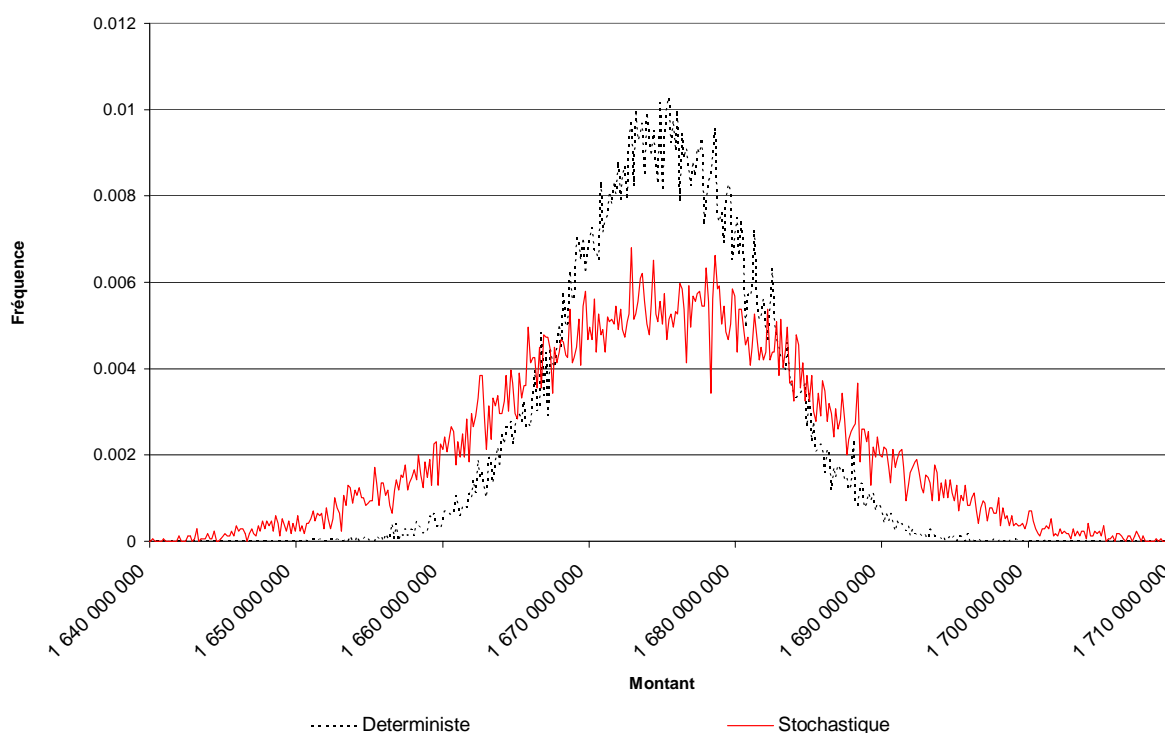


Fig. 30 : Distribution empirique de l'engagement (taux technique nul)

Les résultats détaillés sont repris ci-après :

	<b>Déterministe</b>	<b>Stochastique</b>
Espérance	1 675 256 646	1 675 129 393
Ecart-type	6 389 153	10 893 496
Borne inférieure de l'intervalle de confiance	1 662 554 600	1 653 554 600
Borne supérieure de l'intervalle de confiance	1 687 754 600	1 696 454 600
Coefficient de variation	0.38 %	0.65 %

En prenant un taux technique nul, on augmente le coefficient de variation de l'engagement déterministe de 27 % et le coefficient de variation de l'engagement stochastique de 33 %.

Ainsi, la prise en compte de la mortalité stochastique introduit une plus grande sensibilité de l'engagement au taux d'actualisation.

### 3.2.2 Application du modèle à la TPG 1993

Afin de vérifier la robustesse de l'approche proposée, nous avons appliqué ce modèle aux tables de référence en France depuis 1993, les TPG 1993.

Intuitivement le montant moyen de l'engagement stochastique devrait être égal au montant de la provision obtenue grâce à la TPG 93.

	<b>Déterministe</b>	<b>Stochastique</b>
Espérance	36 412 868	36 412 886
Ecart-type	640 549	640 563
VaR 75 %	36 825 000	36 825 000
Coefficient de variation	1,76 %	1,76 %

On remarque que l'engagement déterministe et l'engagement stochastique sont identiques. Ceci s'explique par la très faible volatilité de la série  $(k_t)$  ajustée sur la TPG 93. Ce phénomène est dû à la nature de la TPG 93. En effet, en tant que table générationnelle, la TPG 93 prend déjà en compte l'évolution de la mortalité dans le futur. Ainsi, le modèle de Lee-Carter permet de faire évoluer la TPG 93 dans le temps, mais ne permet pas d'appréhender une dérive de l'évolution de la mortalité dans le futur.

Les mêmes remarques peuvent être effectuées si l'on applique le modèle aux nouvelles tables générationnelles féminines, les TGF 05.

	<b>Déterministe</b>	<b>Stochastique</b>
Espérance	40 486 954	40 486 944
Ecart-type	633 355	633 309
VaR 75 %	40 900 000	40 900 000
Coefficient de variation	1,56 %	1,56 %

Ainsi, on ne peut pas prendre en compte la mortalité stochastique à l'aide de tables générationnelles. Les tables générationnelles étant construites à l'aide d'une hypothèse sur l'évolution de la mortalité dans le futur, le modèle de Lee-Carter appliqué à une table générationnelle reprendra cette hypothèse d'évolution sans prendre en compte la possibilité d'une dérive de cette évolution.

Ceci se vérifie sur les résultats que nous obtenons. Projeter la TPG 93 en 2006 à l'aide du modèle de Lee-Carter ou utiliser directement la TPG 93 en 2006 conduit (à 2 % près) au même montant de provision.

Le montant de la provision obtenue à l'aide de la surface de mortalité obtenue par le modèle de Lee-Carter appliqué sur la TPG 93 est 10,3 % plus faible que le montant obtenu si l'on applique le modèle sur la TGF 05. En utilisant la surface stochastique obtenue à l'aide de l'algorithme de Lee-Carter appliqué sur les tables INED, l'augmentation n'est plus que de 6,2 %.

Les mêmes conclusions peuvent être obtenues si l'on détermine la provision à l'aide du quantile à 75 %.

Ainsi, nous venons de démontrer que la mortalité française évolue en moyenne de façon linéaire. Sa volatilité demeure difficilement appréhendable, n'étant pas due à des sauts annuels indépendants les uns des autres. Il est nécessaire d'étudier la modélisation de la dérive de la mortalité à l'aide d'une série temporelle, et tout porte à croire que c'est le modèle temporel qui sera privilégié.



## II Etude du Risque systématique à l'aide de séries temporelles

Comme nous l'avons montré dans la partie précédente, les résultats obtenus en modélisant la dérive de la mortalité à l'aide d'une régression linéaire ne sont pas totalement concluants. La volatilité de l'engagement stochastique comparée à celle de l'engagement déterministe reste somme toute faible. L'évolution de l'intervalle de confiance à 95 % de l'engagement stochastique nous incite à penser qu'une mortalité déterministe suffit à provisionner de manière prudente les engagements de rente.

Or le passage de la table générationnelle TPG 93 à la table générationnelle TGF 05 s'accompagne d'une augmentation de la provision à constituer de 10,3 %. Ceci s'oppose à l'hypothèse selon laquelle le risque de mortalité est avant tout un risque mutualisable. En effet, si une telle conclusion était exacte, il serait alors aisé de prédire d'une manière satisfaisante l'évolution de la mortalité dans le futur. L'augmentation de 10,3 % n'aurait donc pas lieu d'être.

Nous allons donc chercher un modèle permettant d'appréhender le risque systématique de mortalité. Modéliser la dérive de la mortalité à l'aide d'une régression linéaire revient à supposer que les fluctuations annuelles de la mortalité sont indépendantes. Cette hypothèse peut être levée par application d'un modèle temporel.

### 1. Ajustement du modèle.

#### 1.1 DEFINITION DE LA NON STATIONNARITE

La série  $(k_t)_{t \geq 1}$  représentant la dérive de la mortalité au cours de la deuxième moitié du 20<sup>ième</sup> siècle est non stationnaire.

En effet, une série  $(y_t)_{t \geq 1}$  sera dite faiblement stationnaire si et seulement si elle vérifie les trois conditions suivantes :

1.  $E(y_t) = \text{constante} \quad \forall t$
  2.  $V(y_t) = \text{constante} \quad \forall t$
  3.  $\text{cov}(y_t, y_{t+h}) = \text{cov}(y_t, y_{t'}) \quad \forall t$
- (88)

Or l'espérance de  $(y_t)_{t \geq 1}$  n'étant pas constante, la série  $(k_t)_{t \geq 1}$  n'est pas stationnaire.

Cependant, la non stationnarité de la série  $(k_t)_{t \geq 1}$  peut être soit déterministe, soit stochastique.

### 1.1.1 Non stationnarité déterministe

On dit que le processus  $(y_t)_{t \geq 1}$  est caractérisé par une non stationnarité déterministe, ou encore que le processus  $(y_t)_{t \geq 1}$  est TS (Trend stationary) s'il peut s'écrire :

$$y_t = f(t) + Z_t \quad (89)$$

où  $f(t)$  est une fonction déterministe et  $Z_t$  est un processus stationnaire. Ajuster une série temporelle à l'aide d'une régression linéaire revient à supposer que la série est caractérisée par une non stationnarité déterministe.

Un processus est caractérisé par une non stationnarité déterministe si  $Z_t = y_t - f(t)$  est un processus stationnaire. Il est important de vérifier cette condition si l'on s'intéresse à la volatilité du processus  $(y_t)_{t \geq 1}$ . En modélisant  $(y_t)_{t \geq 1}$  à l'aide d'un processus TS, on présuppose qu'un choc imprévu  $\varepsilon_t$  n'impactera pas la tendance moyenne de la série  $(y_t)_{t \geq 1}$ . Au niveau de l'analyse de la dérive de la mortalité, cela revient à supposer que les sauts de mortalité annuels sont des phénomènes anecdotiques du point de vue de l'évolution de la mortalité dans le temps. Cette hypothèse peut sembler forte, surtout dans un processus modélisant l'évolution « normale » de la mortalité et non un processus modélisant des sauts de mortalité dus à des épidémies.

Ainsi, afin de valider le modèle linéaire que nous avons mis en place dans la partie précédente, nous allons analyser la stationnarité des résidus obtenus lors de la régression linéaire. En effet, le  $R^2$  obtenu lors de la régression linéaire à beau être proche de un, il indique uniquement que la régression ajustée modélise correctement la dérive moyenne de la mortalité.

A la vue du tableau 6 renseignant l'autocorrélogramme des résidus, il est évident qu'il ne s'agit pas d'une série stationnaire.

Ainsi, si une régression linéaire constitue un modèle adéquat pour modéliser la dérive de la mortalité, il ne permettra pas de prendre en compte correctement la volatilité de la mortalité. Il est donc nécessaire de se tourner vers un processus étudiant d'une manière différente la non stationnarité de la série  $(k_t)_{t \geq 1}$ .

Autocorrelations				
Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1	Std Error
0	15.184660	1.00000	*****	0
1	5.651543	0.37219	.*****	0.145865
2	7.268756	0.47869	.*****	0.164837
3	6.011378	0.39588	.*****	0.192151
4	3.825251	0.25192	.**** .	0.208785
5	3.511029	0.23122	.**** .	0.215155
6	3.086381	0.20326	.**** .	0.220379
7	2.209920	0.14554	.*** .	0.224332
8	-0.839650	-0.05530	. * .	0.226332
9	-0.475838	-0.03134	. * .	0.226619
10	-1.150457	-0.07576	. ** .	0.226711
11	-2.060009	-0.13566	. *** .	0.227249
12	-0.810400	-0.05337	. * .	0.228966
13	-2.867795	-0.18886	. **** .	0.229230
14	-2.499755	-0.16462	. *** .	0.232518
15	-3.628310	-0.23895	. **** .	0.234984
16	-2.630616	-0.17324	. *** .	0.240098
17	-4.114645	-0.27097	. **** .	0.242743
18	-3.106698	-0.20459	. **** .	0.249096
19	-5.366379	-0.35341	. ***** .	0.252646
20	-3.514800	-0.23147	. **** .	0.262954
21	-4.489523	-0.29566	. ***** .	0.267254
22	-3.635913	-0.23945	. **** .	0.274125
23	-3.112007	-0.20494	. **** .	0.278540
24	-3.509506	-0.23112	. **** .	0.281730
25	-3.077669	-0.20268	. **** .	0.285736
26	-2.570414	-0.16928	. *** .	0.288778
27	-0.746454	-0.04916	. * .	0.290882

Tab:6 Autocorrélogramme des résidus

### 1.1.2 Non stationnarité stochastique.

On dit que le processus  $(y_t)_{t \geq 1}$  est caractérisé par une non stationnarité stochastique, ou encore que le processus  $(y_t)_{t \geq 1}$  est DS (Difference stationary) si le processus différencié d fois est stationnaire. En notant  $L$  le processus d'intégration d'une série temporelle ( $L(y_t) = y_{t-1}$ ), on peut écrire :

$$(1 - L)^d y_t = Z_t \quad (90)$$

où  $Z_t$  est un processus stationnaire. On dira également que le processus  $(y_t)_{t \geq 1}$  est  $I(d)$  où intégré.

Il existe trois types de processus intégrés :

1. La marche aléatoire, caractérisée par l'équation :  $y_t = y_{t-1} + \varepsilon_t$  (où  $\varepsilon_t$  est un bruit blanc).
2. La marche aléatoire avec dérive, caractérisée par l'équation :  $y_t = y_{t-1} + c + \varepsilon_t$  (où  $\varepsilon_t$  est un bruit blanc et  $c$  une constante).
3. La marche aléatoire avec trend, caractérisée par l'équation :  $y_t = y_{t-1} + c + bt + \varepsilon_t$  (où  $\varepsilon_t$  est un bruit blanc,  $c$  et  $b$  des constantes).

Modéliser un processus non stationnaire à l'aide d'un des trois modèles que nous venons de définir revient à supposer que les sauts aléatoires impacteront de manière permanente la dérive de la série. Autrement dit un choc temporaire à une date donnée aura un effet permanent sur le niveau du processus puisque ce dernier ne rejoindra jamais sa valeur initiale à la suite de ce choc.

### 1.1.3 Etude de la non stationnarité.

Afin de pouvoir modéliser la série  $(k_t)_{t \geq 1}$  il est tout d'abord nécessaire d'étudier la nature de sa non stationnarité. Pour cela nous avons recours au test de Dickey-Fuller augmenté. Il permet de tester les 3 modèles suivants :

$$\begin{aligned}
 [1] \quad k_t &= \rho k_{t-1} + a_t \\
 [2] \quad k_t &= \rho k_{t-1} + a_t + c \\
 [3] \quad k_t &= \rho k_{t-1} + a_t + c + bt
 \end{aligned}
 \tag{91}$$

Dans chacun des modèles, on teste en premier lieu l'hypothèse  $H_0 : \rho = 1$  contre l'hypothèse  $H_1 : \rho \neq 1$ . Si l'hypothèse  $H_0$  est acceptée (elle le sera alors dans les trois modèles) on pratique des tests dits d'hypothèses jointes. Ils diffèrent alors selon le modèle étudié. En effet, dans le modèle [3] on teste l'hypothèse  $H_0 : b = 0$  contre l'hypothèse  $H_1 : b \neq 0$ . Dans le modèle [2] on teste l'hypothèse  $H_0 : c = 0$  contre l'hypothèse  $H_1 : c \neq 0$ .

En pratique, on teste donc l'hypothèse jointe sur le modèle [3].

Si on refuse l'hypothèse nulle, alors la série étudiée est de type [3], à savoir que sa différence première suit une tendance linéaire.

Si on rejette l'hypothèse nulle alors on teste l'hypothèse jointe sur le modèle [2], à savoir que sa différence première est en moyenne constante. Si on refuse l'hypothèse nulle, alors la série étudiée est de type [2]. Sinon elle est obligatoirement de type [1], à savoir une marche aléatoire.

L'application de ce processus de test à la série  $(k_t)_{t \geq 1}$  donne les résultats suivants :

Augmented Dickey-Fuller Unit Root Tests							
Type	Retards	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	1	-0.2413	0.6238	-0.26	0.5860		
Single Mean	1	0.0462	0.9561	0.08	0.9608	16.68	0.0010
Trend	1	-10.7824	0.3364	-2.16	0.5010	2.39	0.7081

Tab:7 Test de Dickey-Fuller sur la série kt

La p-valeur associée au test de racine unitaire (0.3364) sur le modèle [3] (*trend*) nous indique la présence d'une racine unitaire au niveau de la série  $(k_t)_{t \geq 1}$ . Cependant le test associé à l'hypothèse jointe présente également une p-valeur très élevée (0.7081). Ainsi, on ne peut accepter le modèle [3] comme ajustement de la dérive temporelle.

On étudie alors le test associé au modèle [2] (*single mean*). La p-valeur associée à ce test (0.9561) ne permet pas de refuser l'hypothèse de racine unitaire. On étudie alors le test associé permettant de mettre en lumière la présence d'une dérive déterministe. La p-valeur de 0.001 permet, quand à elle, de refuser l'hypothèse de nullité de la tendance déterministe.

Ainsi la dérive temporelle est un processus intégré d'ordre un, c'est-à-dire non stationnaire de type DS avec une dérive, c'est à dire avec un taux de croissance stationnaire.

#### 1.1.4 Estimation des paramètres du modèle.

La série  $(k_t)_{t \geq 1}$  suit donc le processus suivant :

$$k_t = k_{t-1} + c + a_t \quad (92)$$

où  $(a_t)_{t \geq 1}$  est un processus stationnaire de moyenne nulle.

Afin de vérifier la stationnarité du processus  $(a_t)_{t \geq 1}$  il suffit de différencier la série  $(k_t)_{t \geq 1}$  et d'étudier la présence ou non d'une racine unitaire au sein de la série différenciée (le test précédent prouve que le processus  $(k_t)_{t \geq 1}$  est intégré, mais il ne nous donne aucune indication sur son ordre d'intégration).

Nous obtenons alors les résultats suivants :

Augmented Dickey-Fuller Unit Root Tests							
Type	Retards	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	1	-15.0402	0.0046	-2.70	0.0080		
Single Mean	1	-91.1050	0.0003	-6.40	0.0001	20.47	0.0010
Trend	1	-96.5428	<.0001	-6.43	<.0001	20.74	0.0010

Tab:8 Test de Dickey-Fuller sur la série kt intégrée

L'étude des résultats de ce test est immédiate puisque le test de racine unitaire possède une p-valeur inférieure à 1 % dans les trois modèles. La série des  $(k_t)_{t \geq 1}$  différenciée est donc stationnaire et de moyenne nulle (on force sa moyenne à être nulle par la présence de la constante c).

L'étude de la constante c se fait à l'aide de l'estimateur empirique de la moyenne de la série  $(k_t)_{t \geq 1}$  différenciée. On trouve comme résultat :  $c = -2.006$ . On pourra noter que cette valeur est proche du coefficient de la dérive estimée dans l'étude linéaire de la série  $(k_t)_{t \geq 1}$ .

Il reste maintenant à ajuster la série  $(a_t)_{t \geq 1}$ . Pour ce faire nous appliquons la procédure ARIMA de SAS qui reprend l'algorithme de Durbin-Watson. Les autocorrélogrammes obtenus sont les suivants :

Autocorrelations				
Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1	Std Error
0	18.818064	1.00000	*****	0
1	-10.583433	-.56241	*****  .	0.147442
2	1.923941	0.10224	.  ** .	0.188392
3	2.121572	0.11274	.  ** .	0.189594
4	-2.964250	-.15752	. ***  .	0.191046
5	0.382074	0.02030	.   .	0.193849
6	1.066264	0.05666	.  * .	0.193895

Tab:9 Autocorrélation

Inverse Autocorrelations		
Lag	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
1	0.49819	.  *****
2	0.06842	.  * .
3	-0.03321	. *  .
4	0.15430	.  *** .
5	0.09300	.  ** .
6	-0.03532	. *  .

Tab:10 Autocorrélation inverse

Partial Autocorrelations		
Lag	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
1	-0.56241	*****  .
2	-0.31310	*****  .
3	0.01971	.   .
4	-0.05410	. *  .
5	-0.14285	. ***  .
6	-0.05841	. *  .

Tab:11 Autocorrélation partielle



Afin de choisir quels types de modèles permettront d'ajuster au mieux la série  $(a_t)_{t \geq 1}$ , nous appliquons les méthodes SCAN, ESACF et BIC. Nous obtenons les résultats suivants :

ARMA(p+d,q) Tests de sélection d'ordre provisoire					
SCAN			ESACF		
p+d	q	BIC	p+d	q	BIC
0	1	2.173614	2	0	1.987392
2	0	1.987392	0	1	2.173614
			1	1	2.166753
			3	0	1.873967
			5	1	1.937318

Tab:12 Résultat des tests de sélection de modèle

Les modèles autorégressifs d'ordre 2 et moyenne mobile d'ordre un sont détectés par les tests SCAN et ESACF. Le test BIC favorise lui un ARMA (5,1).

On teste alors ces trois modèles :

Pour le modèle autorégressif d'ordre 2, on obtient les résultats suivants :

Maximum Likelihood Estimation					
Paramètre	Estimation	Erreur standard	Valeur du test t	Pr. Approx. >  t	Retard
AR1,1	-0.78247	0.14394	-5.44	<.0001	1
AR1,2	-0.33379	0.14293	-2.34	0.0195	2

Variance Estimate	11.52503
Std Error Estimate	3.394853
AIC	245.6036
SBC	249.2609
Number of Residuals	46

Tab:13 Estimation des paramètres du modèle AR (2)

Pour le modèle moyenne mobile d'ordre un, on obtient les résultats suivants :

Maximum Likelihood Estimation					
Paramètre	Estimation	Erreur standard	Valeur du test t	Pr. Approx. >  t	Retard
MA1,1	0.65172	0.11513	5.66	<.0001	1

Variance Estimate	11.99618
Std Error Estimate	3.46355
AIC	246.3753
SBC	248.2039
Number of Residuals	46

Tab:14 Estimation des paramètres du modèle MA (1)

Pour le modèle ARMA (5,1), on obtient les résultats suivants :

Maximum Likelihood Estimation					
Paramètre	Estimation	Erreur standard	Valeur du test t	Pr. Approx. >  t	Retard
MA1,1	0.11463	1.25329	0.09	0.9271	1
AR1,1	-0.67637	1.24459	-0.54	0.5868	1
AR1,2	-0.26977	0.98905	-0.27	0.7850	2
AR1,3	-0.05129	0.48235	-0.11	0.9153	3
AR1,4	-0.13149	0.20318	-0.65	0.5175	4
AR1,5	-0.12157	0.17066	-0.71	0.4763	5

Variance Estimate	12.44519
Std Error Estimate	3.527774
AIC	252.8296
SBC	263.8015
Number of Residuals	46

Tab:15 Estimation des paramètres du modèle ARMA (5,1)

Comme on pouvait s'y attendre, le modèle ARMA (5,1) n'est pas un modèle réaliste. Reste à choisir entre le modèle autorégressif d'ordre 2 et la moyenne mobile d'ordre un. Pour ce faire nous allons utiliser le critère AIC. Le modèle autorégressif d'ordre 2 minimisant ce critère, c'est ce modèle qui sera conservé.

Ainsi, l'étude que nous venons de réaliser nous permet de simuler la série  $(k_t)_{t \geq 1}$  à l'aide du processus suivant :

$$k_t = c + (1 + a_1)k_{t-1} + (a_2 - a_1)k_{t-2} - a_3k_{t-3} + \varepsilon_t \quad (93)$$

avec :

- ✓  $c = -2.006422194$
- ✓  $a_1 = -0.78247$
- ✓  $a_2 = -0.33379$
- ✓  $\varepsilon_t = 3.394853$

Une centaine de simulations de la série temporelle  $(k_t)_{t \geq 1}$  permet d'obtenir le graphique suivant :

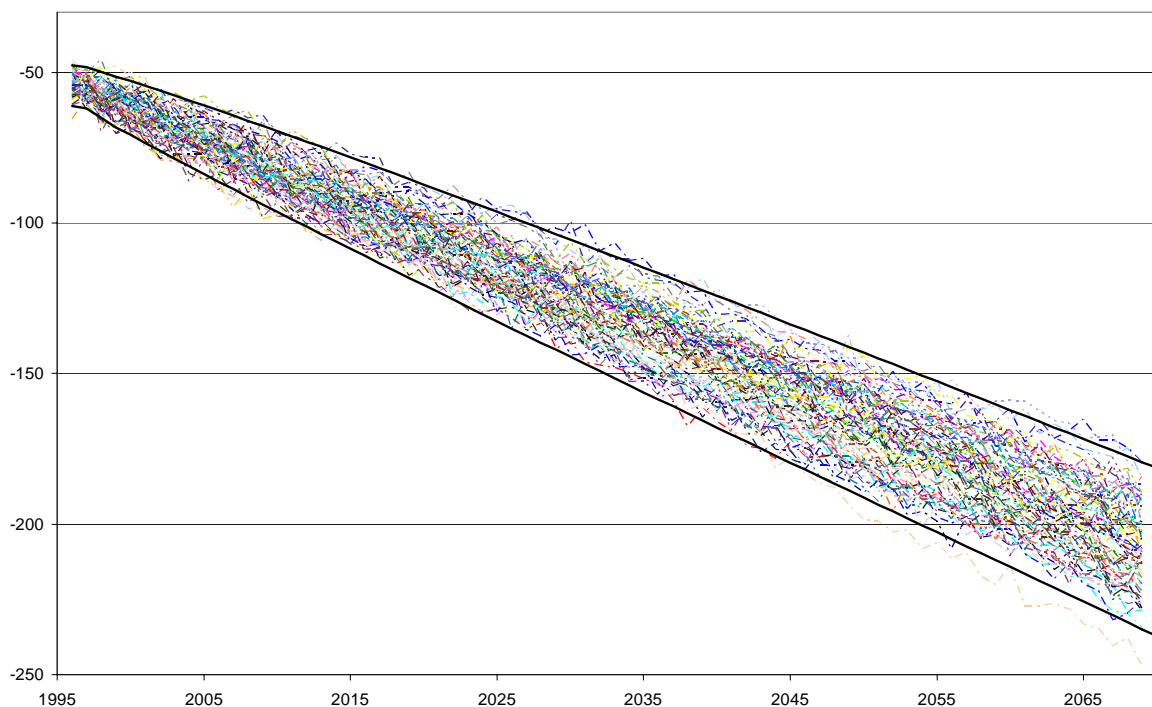


Fig. 31 : Intervalle de confiance à 95 % de la dérive temporelle

Si l'on compare ce graphique à celui de la figure 27 « couloir de variation de la tendance du modèle linéaire », on s'aperçoit de la grande volatilité de la série  $(k_t)_{t \geq 1}$  ainsi modélisée.

Il est alors intéressant de modéliser le volume à 95 % de la surface de mortalité stochastique :

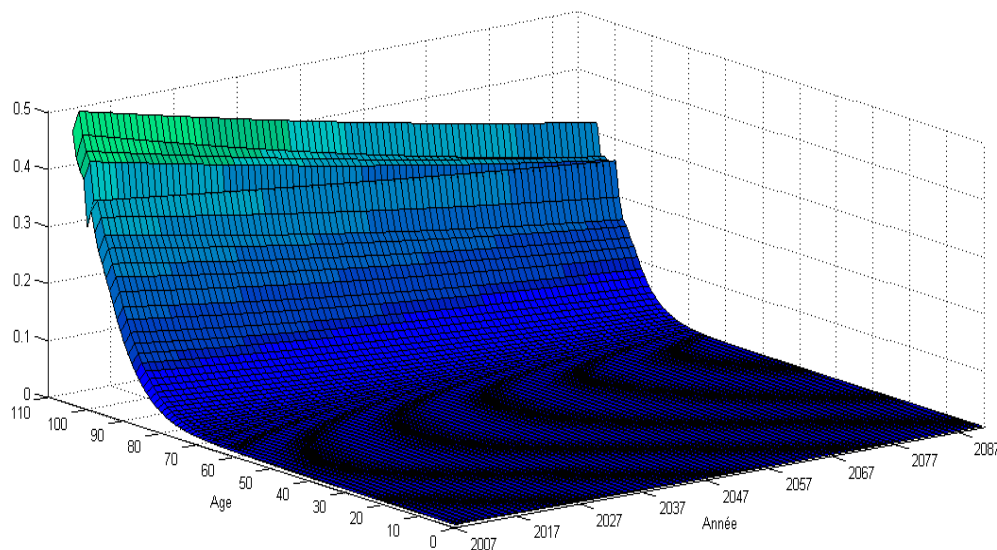


Fig. 32 : Volume à 95 % de la surface de mortalité

On constate que le volume possède une épaisseur non négligeable. Il n'en est pas de même lorsque l'on modélise la dérive de la mortalité à l'aide du modèle « risque de dérive ». Cette dernière n'est pas présentée dans ce modèle étant donné que le volume n'est pas différentiable de la surface moyenne.

Nous allons maintenant analyser la distribution de l'engagement que l'on obtient avec ce nouveau modèle.

## 2. Applications numériques

Une fois le modèle ajusté, on simule la distribution de l'engagement.

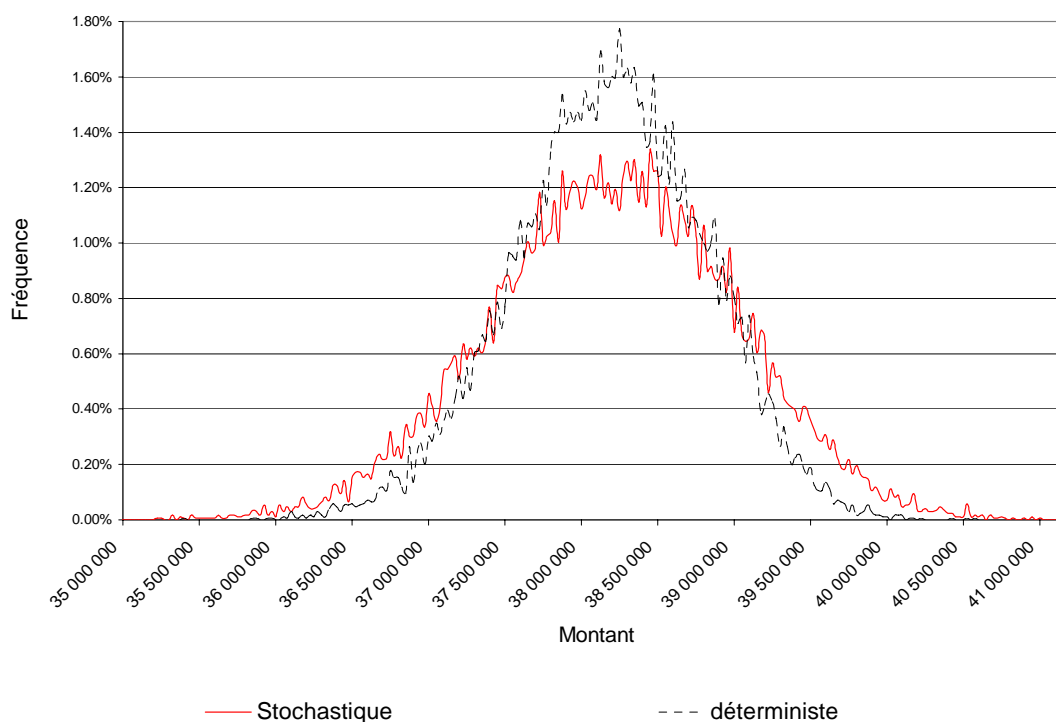


Fig. 33 : *Distribution empirique de l'engagement*

Les résultats détaillés sont repris ci-après :

	<b>Déterministe</b>	<b>Stochastique</b>
Espérance	38 198 606	38 218 125
Ecart-type	626 397	803 479
Borne inférieure de l'intervalle de confiance	36 900 000	36 600 000
Borne supérieure de l'intervalle de confiance	39 350 000	39 700 000
Coefficient de variation	1.64 %	2.10 %

On remarque tout d'abord que l'engagement déterministe obtenu avec ce modèle est identique à celui obtenu avec le modèle précédent. Ceci nous montre la qualité du modèle que nous venons d'ajuster.

L'unique différence se situe au niveau du coefficient de variation de l'engagement stochastique. En modélisant la dérive temporelle à l'aide d'une série temporelle, nous obtenons un coefficient de variation 24 % plus élevé.

Ainsi, le modèle que nous venons d'ajuster simule correctement la volatilité de la dérive stochastique de la mortalité tout en conservant les qualités du modèle linéaire ajusté précédemment (modélisation précise de la dérive moyenne).

Au niveau des résultats, sur un portefeuille de petite taille le coefficient de variation de l'engagement stochastique est 28 % plus élevé que celui de l'engagement déterministe.

Le quantile à 75 % vaut 38.8 Millions euros en tenant compte de la mortalité stochastique lorsqu'il vaut 38.6 Millions euros si l'on travaille avec une mortalité déterministe. Le quantile à 75 % augmente donc de 0.5 %.

En augmentant la taille du portefeuille par 30, nous obtenons pour la distribution de l'engagement :

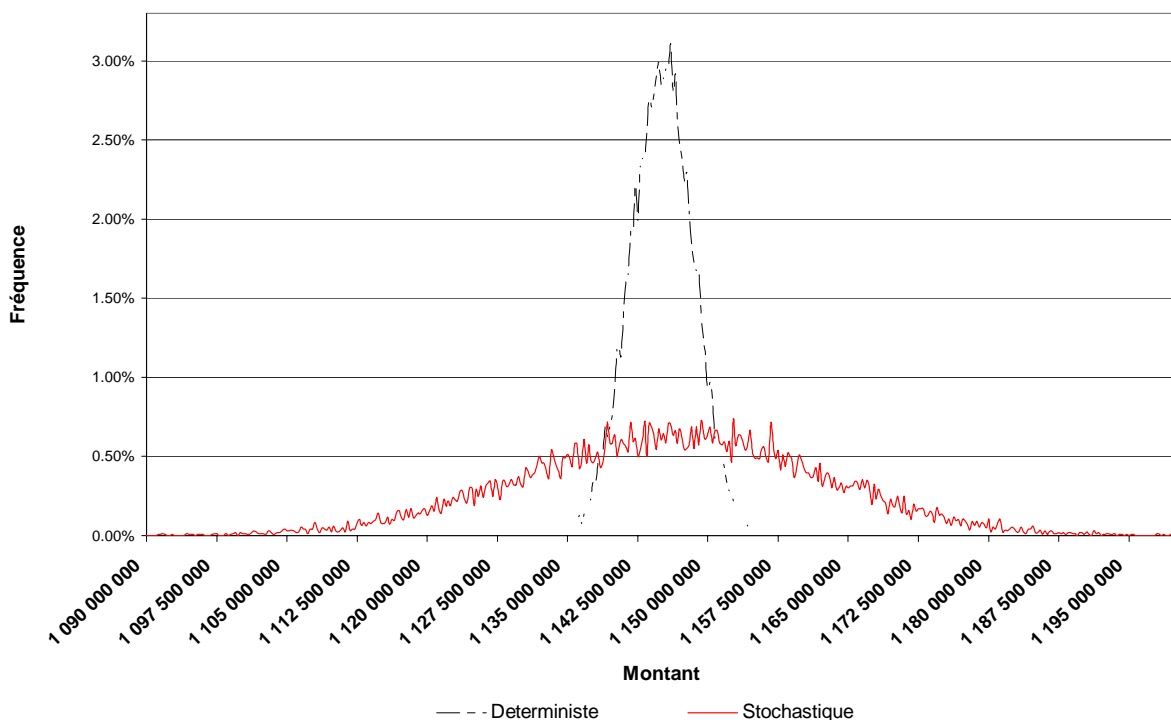


Fig. 34 : Distribution empirique de l'engagement portefeuille taille 30

	<b>Déterministe</b>	<b>Stochastique</b>
Espérance	1 145 232 071	1 146 211 145
Ecart-type	3 374 739	15 793 259
Borne inférieure de l'intervalle de confiance	1 135 750 000	1 101 000 000
Borne supérieure de l'intervalle de confiance	1 151 750 000	1 176 250 000
Coefficient de variation	0.29 %	1.38 %

On constate que sur un portefeuille de grande taille (11220 rentiers), le coefficient de variation de l'engagement stochastique est presque 400 % plus élevé que celui de l'engagement déterministe.

La largeur de l'intervalle de confiance à 95 % passe de 16 M euros dans le cas déterministe à 75 M euros dans le cas stochastique, soit une augmentation de plus de 400 % de la taille de l'intervalle de confiance à 95 %. Ainsi, considérer que la mortalité n'est pas un phénomène déterministe mais stochastique conduit à augmenter l'intervalle de confiance à 95 % de presque 6 %.

Le quantile à 75 % vaut 1 157 M euros en tenant compte de la mortalité stochastique lorsqu'il vaut 1 147 M euros si l'on travaille avec une mortalité déterministe. Le quantile à 75 % augmente lui de 1 %.

Ainsi, en modélisant la dérive temporelle à l'aide d'une série temporelle, la volatilité de l'engagement devient plus importante. Il devient alors important de prendre en compte le risque systématique si l'on ne veut pas sous estimer l'engagement. Comme nous l'avons fait lors de l'étude de la dérive temporelle à l'aide d'un modèle linéaire, il est important d'analyser l'impact d'une augmentation de la taille du portefeuille sur l'intervalle de confiance à 95 %.

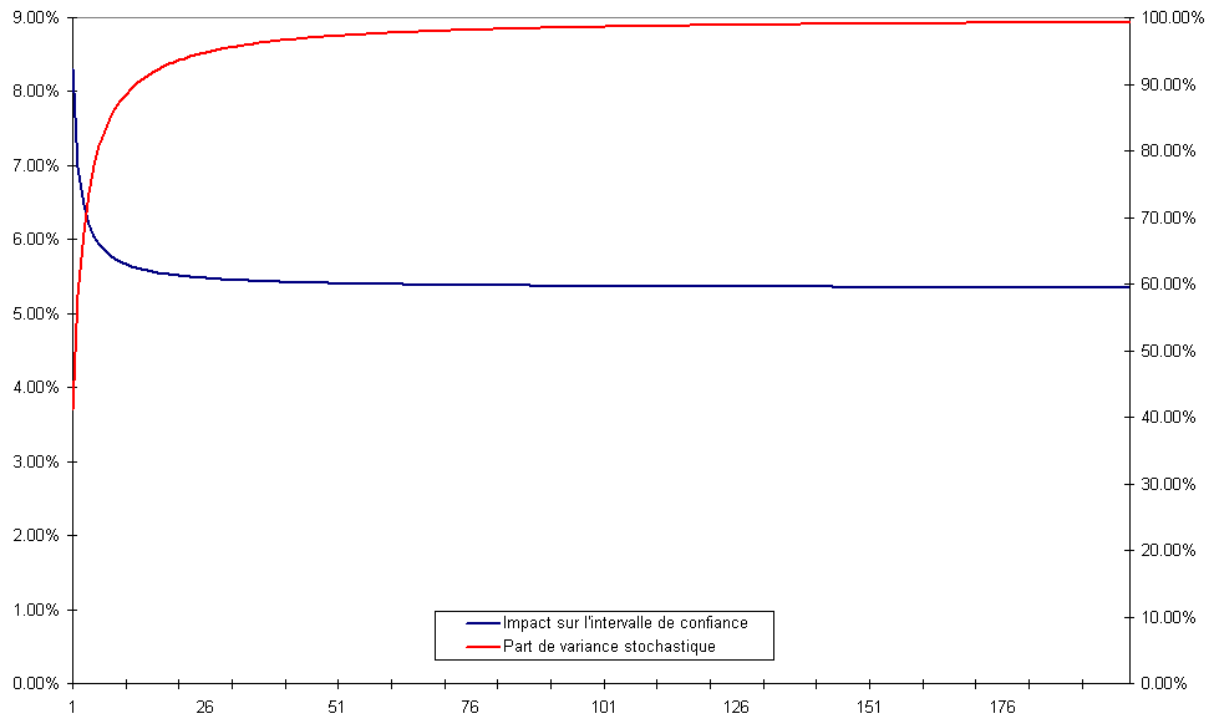


Fig. 35 : Impact de la volatilité sur l'intervalle de confiance en fonction de la taille du portefeuille

Le modèle que nous avons mis en place nous montre donc que la mortalité stochastique présente un impact important sur l'intervalle de confiance à 95 %. Ainsi pour des portefeuilles de taille faible, plus de 50 % de la volatilité sera due au risque systématique. Donc sur un portefeuille de 3000 rentiers la volatilité de l'engagement sera à 90 % expliquée par la mortalité stochastique, sachant que son impact sur l'intervalle de confiance à 95 % sera au minimum de 5,5 %. Il est donc important de tenir compte de ce paramètre.

On peut donc conclure que la dérive de la mortalité possède des oscillations annuelles qui sont dépendantes, elle ne possède pas une tendance linéaire avec des oscillations indépendantes, ni une tendance linéaire présentant un risque de modèle. Tout saut annuel de mortalité aura donc un impact sur les taux de mortalité futurs.

Quoiqu'il en soit, ce n'est pas la volatilité de l'engagement qui nous importe avant tout, mais plutôt la provision que nous allons passer grâce au modèle que nous avons construit. Il est donc important d'analyser les différentes probabilités de ruine qui sont attachées à chaque modèle.



## Conclusion

Nous venons donc de démontrer que la dérive de la mortalité présente une tendance linéaire dont la volatilité ne peut être expliquée qu'en ayant recours à un modèle temporel.

Cependant, les trois modèles que nous venons d'étudier sont tous applicables à une analyse de la mortalité stochastique. Le choix du modèle dépend de la population étudiée. Si sur la population française nous concluons que la mortalité présente une dérive linéaire associée à des sauts de mortalité corrélés, une toute autre conclusion pourrait être obtenue si l'on étudiait la population d'un autre pays.

Si les valeurs moyennes des taux de mortalité futurs estimées par ces trois modèles sont identiques, l'estimation de l'erreur commise sur les vrais taux de mortalité futurs est différente.

Le modèle linéaire présuppose que l'incertitude sur le taux de mortalité de l'année 2007 est identique à l'incertitude portant sur le taux de mortalité en 2090. Cette hypothèse relativement forte sera préférée lors d'études portant sur la mortalité en présence d'épidémie

Le modèle avec « risque de dérive » présuppose qu'il existe une droite permettant de prédire d'une manière exacte la mortalité future. L'incertitude repose alors sur l'estimation de cette droite. Etant donné que l'impact d'une modification de la pente sera d'autant plus grand que l'on s'éloignera dans le futur, l'incertitude portant sur les taux de mortalité du futur lointain sera plus grande que celle portant sur les taux de mortalité du futur proche.

Le modèle ajusté à l'aide d'une série temporelle présuppose que les taux futurs dépendent des taux passés et présents. Ainsi l'incertitude portant sur les taux futurs intégrera l'incertitude portant sur les taux présents.

Les différents modèles que nous venons de présenter peuvent être résumés grâce au tableau synoptique suivant :

	Régression linéaire	Risque de dérive	Série temporelle
Hypothèse :	On suppose que la dérive de la mortalité possède une tendance linéaire associée à des sauts annuels indépendants et sans impact sur les taux futurs.	On suppose que la dérive de la mortalité possède une tendance linéaire dont le caractère stochastique est associé à un risque de modèle et non à des sauts annuels	On suppose que la dérive de la mortalité possède une tendance linéaire associée à des sauts annuels indépendants ayant un impact sur les taux futurs.
Impact de la mortalité passée sur la mortalité future :	Les sauts de mortalité sont indépendants les uns des autres. La mortalité passée n'impacte pas sur la mortalité future : $k_t = E[k_{t-1}] + a + \varepsilon_t$	La régression linéaire ayant montré que les sauts annuels de mortalité n'ont pas d'impact, ils ne sont pas pris en compte : $k_t = E[k_{t-1}] + a$	La mortalité future dépend des valeurs de la mortalité passée. Les sauts passés impactent sur la valeur de la mortalité future: $k_t = \mu + \sum_{i=1}^n a_i k_{t-i} + \varepsilon_t$
Estimation des paramètres :	L'intercepte et la pente sont obtenus grâce à la <b>Proc Reg</b> de SAS	L'intercepte et la pente sont obtenus grâce à la <b>Proc Reg</b> de SAS La matrice de variance covariance est obtenue par application de la formule 78	Après avoir stationnarisé la série par différenciation, on ajuste les coefficients par la <b>Proc ARIMA</b> sous SAS appliquant la méthode de Box-Jenkins.
Impact de la volatilité de la mortalité stochastique ajustée par le modèle :	Sur un portefeuille de faible taille la volatilité due à la mortalité stochastique est faible en comparaison de celle de la mortalité mutualisable.  Sur un portefeuille de grande taille, le coefficient de variation de l'engagement est faible.	Sur un portefeuille de faible taille la volatilité due à la mortalité stochastique est faible en comparaison de celle de la mortalité mutualisable.  Sur un portefeuille de grande taille, le coefficient de variation de l'engagement est faible.	Sur un portefeuille de faible taille la volatilité due à la mortalité stochastique est aussi grande que celle de la mortalité mutualisable.  Sur un portefeuille de grande taille, l'impact de la mortalité stochastique est non négligeable :  Coefficient de variation supérieur à 1 %. Largeur de l'intervalle de confiance à 95 % supérieur à 5 % de l'engagement
Conclusion	La dérive de la mortalité possède une tendance linéaire associée à des sauts annuels ayant un impact sur les taux futurs.		

Tab:16 Tableau synoptique des modèles stochastiques d'ajustement de la mortalité

### **III Impact de l'ancienneté maximale de l'historique sur l'engagement**

Nous venons de détecter quel modèle nous devons appliquer. Il s'agit maintenant de calibrer ce modèle en analysant l'engagement moyen.

Jusqu'à présent l'engagement moyen nous a servi à contrôler la cohérence des différents modèles que nous avons construits. La TGF 05 étant une table prospective qui ne date que de quelques mois, elle nous indique l'engagement moyen que doit retourner notre modèle lorsque nous l'appliquons sur notre portefeuille de rentiers. Cet engagement doit être de 40.4 Millions d'euros, soit 2.5 Millions de plus que l'engagement que nous avons pour l'instant.

Pour solutionner ce problème, nous allons étudier l'historique sur lequel nous calibrons le modèle. Jusque là, nous utilisons l'ensemble des tables du moments couvrant la période [1950 – 1996]. Nous avons décidé au début de ne pas descendre en dessous de l'année 1950, et donc de ne pas prendre en compte les tables du moment relatives aux deux guerres, afin de ne pas biaiser par le bas l'engagement de notre modèle. L'historique a été pris délibérément grand afin de capter au mieux la volatilité de la mortalité. Cependant, le modèle de Lee-Carter ne distingue pas les données issues des tables de mortalité anciennes des données issues des tables de mortalité récentes. Le modèle ainsi calibré sera plus proche des taux de mortalité des années [1950 – 1970] que celui des années [1970 – 1996].

Il est donc nécessaire de diminuer l'historique permettant de calibrer le modèle, tout en conservant un historique suffisamment grand pour appréhender correctement la volatilité de la mortalité stochastique, ainsi que pour calibrer de manière suffisamment stable les paramètres du modèle.

De plus, en diminuant l'intervalle d'étude, est on sûr que l'analyse de la dérive temporelle que nous avons effectuée préalablement reste valable ?

Pour savoir si la dérive de la mortalité possède une structure stable durant le 20<sup>ème</sup> siècle nous allons déterminer l'évolution de la dérive temporelle dans le passé sur les périodes [1960-1996], [1970-1996] et [1980-1996].

Nous obtenons le graphique suivant :

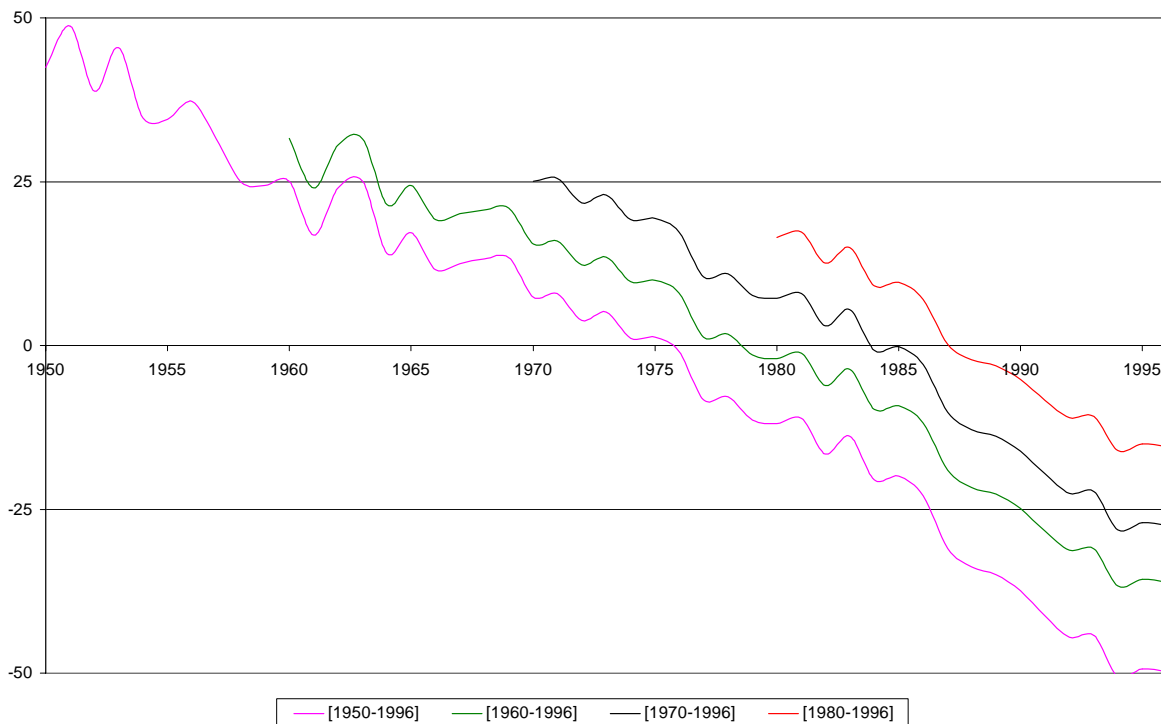


Fig. 36 : Impact de la diminution de l'intervalle d'étude sur la dérive de mortalité

Ainsi, diminuer l'intervalle d'étude revient à translater la série  $(k_t)_{t \geq 1}$ . La valeur de la translation, qu'elle soit positive ou négative importe peu, étant combinée avec les suites  $(\alpha_x)$  et  $(\beta_x)$ , dont les valeurs varient à chaque simulation. Le modèle que nous avons construit sera donc toujours le meilleur. Dans le cas où la dérive de la mortalité présente une tendance linéaire, cette conclusion sera toujours valable.

L'allure de l'évolution de la dérive temporelle nous incite également à penser que la volatilité de la mortalité ne diminuera pas lorsque la taille de l'historique diminuera.

La taille de l'historique sera donc calibrée en tenant compte de l'engagement moyen du modèle. La table générationnelle TGF 05 jouera donc un rôle essentiel dans le calibrage du modèle.

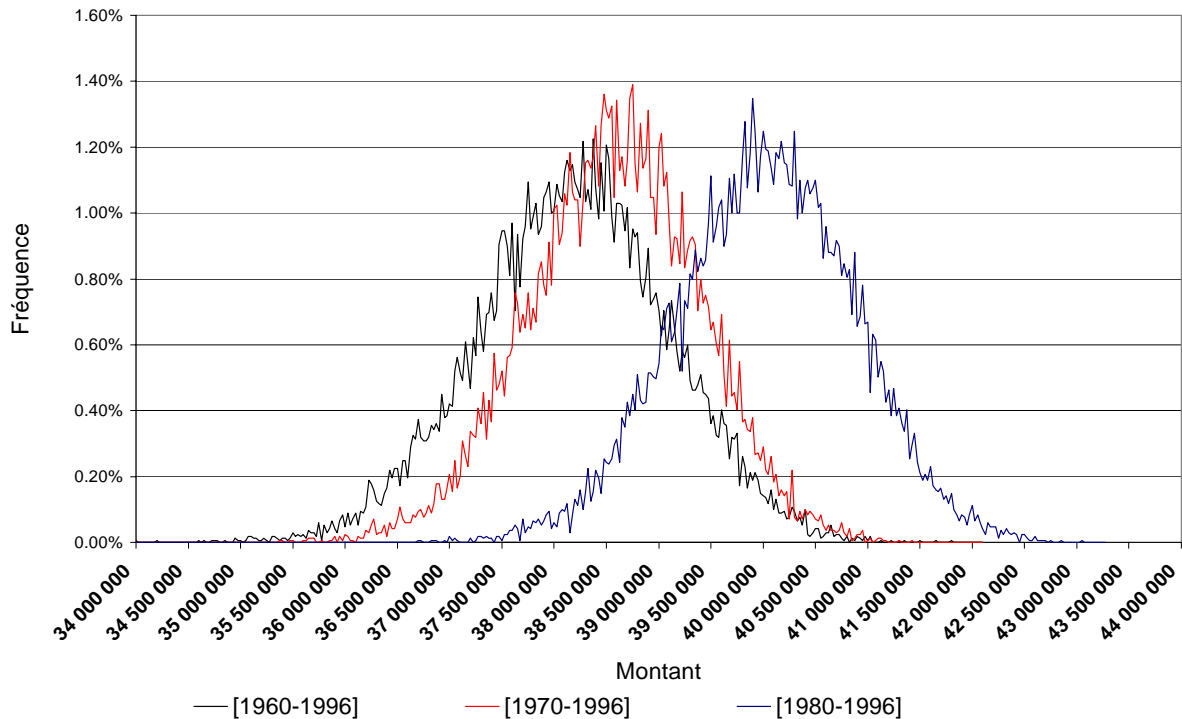


Fig. 37 : Evolution de l'engagement moyen lors de la diminution de l'intervalle d'étude

A la vue du graphique précédent, deux remarques peuvent être faites. Tout d'abord, comme nous le présentions, la volatilité de l'engagement ne diminue pas lorsque l'intervalle d'étude se réduit. Ensuite, diminuer l'intervalle d'étude a bien solutionné le problème de l'engagement moyen.

En effet, si nous ajustons notre modèle grâce aux tables du moment de la période [1980-1996] nous obtenons un engagement moyen de 40.02 Millions d'euros soit 460 000 euros de moins que le montant obtenu en utilisant la TGF 05. Il est à noter que si l'on réduit l'intervalle d'étude à la période [1987-1996] nous ne possédons que 10 points pour ajuster la série temporelle (ce qui n'est pas suffisant pour être précis) et nous obtenons un engagement moyen de 39.9 Millions d'euros.

Nous venons donc de mettre en lumière le principal problème lié à l'étude de la mortalité stochastique. Son caractère aléatoire étant difficilement prévisible (notamment au niveau des variations brutales de la dérive moyenne), il est très difficile de déterminer le nombre de tables du moment qu'il faut considérer afin d'effectuer une étude correcte.

En ajustant le modèle sur l'intervalle [1990-2006] nous obtiendrons alors un modèle applicable, présentant un engagement moyen en accord avec celui que nous pouvons obtenir avec la table générationnelle TGF 05 et possédant une volatilité relativement élevée. De plus, lorsque nous travaillons sur un portefeuille de quelques dizaines de rentiers, ce n'est pas tant la volatilité de l'engagement stochastique qui nous intéresse, mais le couple :

(Volatilité, part de variance stochastique)

la part de variance stochastique nous renseignant sur l'évolution du caractère risqué du portefeuille lorsque sa taille augmentera.

En effet, pour analyser l'impact de l'augmentation de la taille du portefeuille, il est inutile de faire de nouvelles simulations, et ce quelque soit la méthode utilisée pour modéliser la dérive de la mortalité.

Partons de la volatilité de l'engagement pour un portefeuille de taille n:

$$V[n\Lambda] = nE[\text{VAR}[\Lambda|\Sigma]] + n^2 \text{VAR}[E[\Lambda|\Sigma]] \quad (94)$$

Comme la part de variance stochastique vaut :

$$P_{sto} = \frac{\text{VAR}[E[\Lambda|\Sigma]]}{\text{VAR}[E[\Lambda|\Sigma]] + E[\text{VAR}[\Lambda|\Sigma]]} \quad (95)$$

L'évolution de la volatilité en fonction de la taille du portefeuille peut donc se réécrire :

$$V[n\Lambda] = V[\Lambda][n(1 - P_{sto}) + n^2 P_{sto}] \quad (96)$$

Ainsi, plus la part de variance stochastique sur un portefeuille de taille un sera proche de un, plus l'engagement sera volatile lorsque le portefeuille grossira, et donc plus on considérera que le portefeuille est risqué.

Le coefficient de variation sera également d'autant plus important que le coefficient de variation et la part de variance stochastique sur un portefeuille de taille un seront grands.

En effet, par définition du coefficient de variation, on peut écrire :

$$Cv_n = \frac{\sqrt{V[n\Lambda]}}{nE[\Lambda]} = \frac{\sqrt{V[\Lambda][n(1 - Psto) + n^2 Psto]}}{nE[\Lambda]} = \frac{Cv}{\sqrt{n}} \sqrt{(1 - Psto) + n Psto} \quad (97)$$

Ainsi :

$$\lim_{n \rightarrow \infty} Cv_n = Psto Cv \quad (98)$$

On s'aperçoit que le choix du modèle de dérive doit se faire en tenant compte de deux paramètres : la volatilité globale de l'engagement qui est un indicateur de premier ordre du caractère risqué de l'engagement, et la part de variance stochastique qui est un indicateur de deuxième ordre puisqu'il nous indique de quelle manière va évoluer la volatilité de l'engagement.

Les résultats obtenus sont renseignés dans le tableau suivant

	<b>Déterministe</b>	<b>Stochastique</b>
Espérance	40 006 749	40 027 175
Ecart-type	648543	851 234
Borne inférieure de l'intervalle de confiance	38 675 000	39 512 177
Borne supérieure de l'intervalle de confiance	41 225 000	41 625 000
Coefficient de variation	1.62 %	2.13 %

De plus le modèle ainsi calibré possède un coefficient de variation de 2.13 % et une part de variance expliquée sur un portefeuille de taille un de 42 %. L'impact de la mortalité stochastique sera donc relativement important sur l'intervalle de confiance à 95 % lorsque la taille du portefeuille augmentera. Les résultats obtenus dans la partie précédente seront donc toujours exacts.

Ce dernier résultat sera particulièrement important sous solvabilité 2. En effet, sous solvabilité I, la volatilité de l'engagement importe peu lors du calcul de la marge de solvabilité réglementaire. La volatilité de l'engagement intervenant uniquement au niveau de la probabilité que le montant « provision plus marge de solvabilité » soit insuffisant.

Le projet solvabilité 2 prend en compte la volatilité de l'engagement lors de l'établissement de la provision et du niveau des fonds propres nécessaires. En effet, l'une des mesures envisagée consiste à fixer le montant de la provision au quantile à 75 % de l'engagement.

Ainsi, plus l'engagement sera volatile, plus la provision à passer sera grande. L'impact de la mortalité stochastique sur la provision sera alors d'autant plus significatif que le portefeuille sera grand. En effet, en présence d'une mortalité déterministe, donc mutualisable, lorsque le portefeuille, et donc l'engagement moyen, augmente de manière linéaire, la volatilité augmente en racine carré. De cette façon, sur des portefeuilles de grande taille, l'impact de la volatilité sur l'engagement devient très faible. L'engagement tend vers une masse de Dirac et le quantile à 75 % est presque confondu avec l'engagement moyen.

En présence d'une mortalité stochastique ceci n'est plus exact. En effet, la volatilité stochastique évolue de manière linéaire lorsque le portefeuille augmente. Ainsi l'engagement du portefeuille aura tendance à se resserrer (à cause de la diminution de la volatilité due à la part mutualisable), mais elle tendra vers un engagement dont le coefficient de variation vaudra :

$$\lim_{n \rightarrow \infty} Cv_n = PstoCv \quad (99)$$

Dans le cadre de notre étude, nous seront donc confrontés à un engagement dont le coefficient de variation vaudra : 0.82 %. Ce coefficient de variation peut paraître faible, mais lorsqu'il est associé à un engagement de plusieurs centaines de millions d'euros, il n'en est rien.

Le quantile à 75 % suivant une évolution comparable à celle du coefficient de variation, on peut donc conclure que la mortalité stochastique aura un rôle lors de l'application du projet solvabilité 2.



## Conclusion

Lors de l'étude que nous venons de réaliser, nous avons construit un modèle permettant de modéliser la mortalité stochastique. Nous avons également montré combien il est difficile d'appréhender et de modéliser ce phénomène. En effet, un modèle permettant de simuler de manière efficace la mortalité stochastique doit vérifier deux conditions :

- L'engagement moyen déterminé doit être suffisamment prudent.
- La volatilité de la surface de mortalité simulée doit être suffisamment importante.

Le modèle que nous avons construit permet de respecter ces deux conditions uniquement si l'historique utilisé pour calibrer le modèle est adéquat. Il ne doit pas être trop grand, sinon il risquerait de biaiser par le bas l'engagement que nous simulerions. Mais il ne doit pas être trop petit sous peine d'obtenir un modèle dont les paramètres ne seraient pas fiables.

Ainsi, en restreignant l'intervalle d'étude à la période [1987 – 1996] nous obtenons un engagement moyen inférieur à 40 Millions d'euros. Ceci s'explique par la présence d'une légère diminution de la dérive de la mortalité en 1994, 1995, et 1996. L'intervalle d'étude ne présentant que 10 points, toute valeur aberrante possédera un poids important dans l'estimation du modèle.

La véritable difficulté ne réside pas dans la mise en place d'un modèle opérationnel, mais plutôt dans la détermination de l'intervalle sur lequel nous devons calibrer le modèle. Aux vues de l'étude que nous avons réalisée, nous nous apercevons que sans la TGF 05 il est presque impossible de déterminer correctement l'intervalle d'étude.

De plus, un processus d'analyse de la mortalité stochastique uniquement basé sur des données issues des tables du moment passées, ajuste la dérive (évolution moyenne et volatilité) en fonction de l'évolution dite « normale » de la mortalité, et en fonction de l'impact de certaines variables exogènes qui a été constaté par le passé. Cependant, nous nous contentons d'analyser les conséquences visibles de ces variables exogènes mais nous ne pouvons prédire de quelle manière elles impacteront réellement la mortalité future, et à quelles dates.

Une manière aisée de prouver ce phénomène consiste à ajuster ce modèle à l'aide des tables du moment couvrant la période [1960-1976]. Nous avons alors un modèle tenant compte de la mortalité telle qu'elle pouvait être vue à la fin des années 70. En appliquant ce modèle en 2006 sur le portefeuille de rentier que nous étudions, nous trouvons un engagement moyen de plus de 50 millions. Ainsi la diminution de la décroissance de la dérive de la mortalité (qui pourrait s'expliquer par un ralentissement des progrès de la médecine) au cours des années 80-90 n'a pu être appréhendée à l'avance.

Ainsi, lorsque nous étudions de quelle façon elle a évolué au cours du 20<sup>ème</sup> siècle, nous prenons en compte une décroissance de la mortalité due à l'évolution de la médecine, des conditions de vie. Puis nous projetons des évolutions stochastiques de la mortalité en considérant que l'évolution passée de ces paramètres sera identique dans le futur.

Ainsi, si nous désirions construire un modèle permettant de prédire de manière tout à fait exacte l'évolution de la mortalité dans le futur, une technique consisterait à calibrer le modèle sur un intervalle d'étude suffisamment large pour appréhender la volatilité de la dérive de la mortalité, puis de calibrer la dérive moyenne à l'aide de coefficients exogènes relatifs à l'évolution de la durée de vie humaine. On peut alors penser à utiliser des données médicales nous permettant d'estimer l'évolution de l'espérance de vie à 60 ans, puis de forcer le modèle à vérifier cette évolution.

Ainsi, une analogie entre la mortalité stochastique et la finance peut être faite : l'étude des évolutions passées de la mortalité ne permet pas de prédire à long terme son évolution de manière précise. Il est donc nécessaire de mettre à jour le modèle régulièrement, et ce à l'aide de nouvelles tables du moment ou de nouvelles tables générationnelles. Le modèle mis en place permettra d'obtenir un « best estimate » du portefeuille.

## Bibliographie

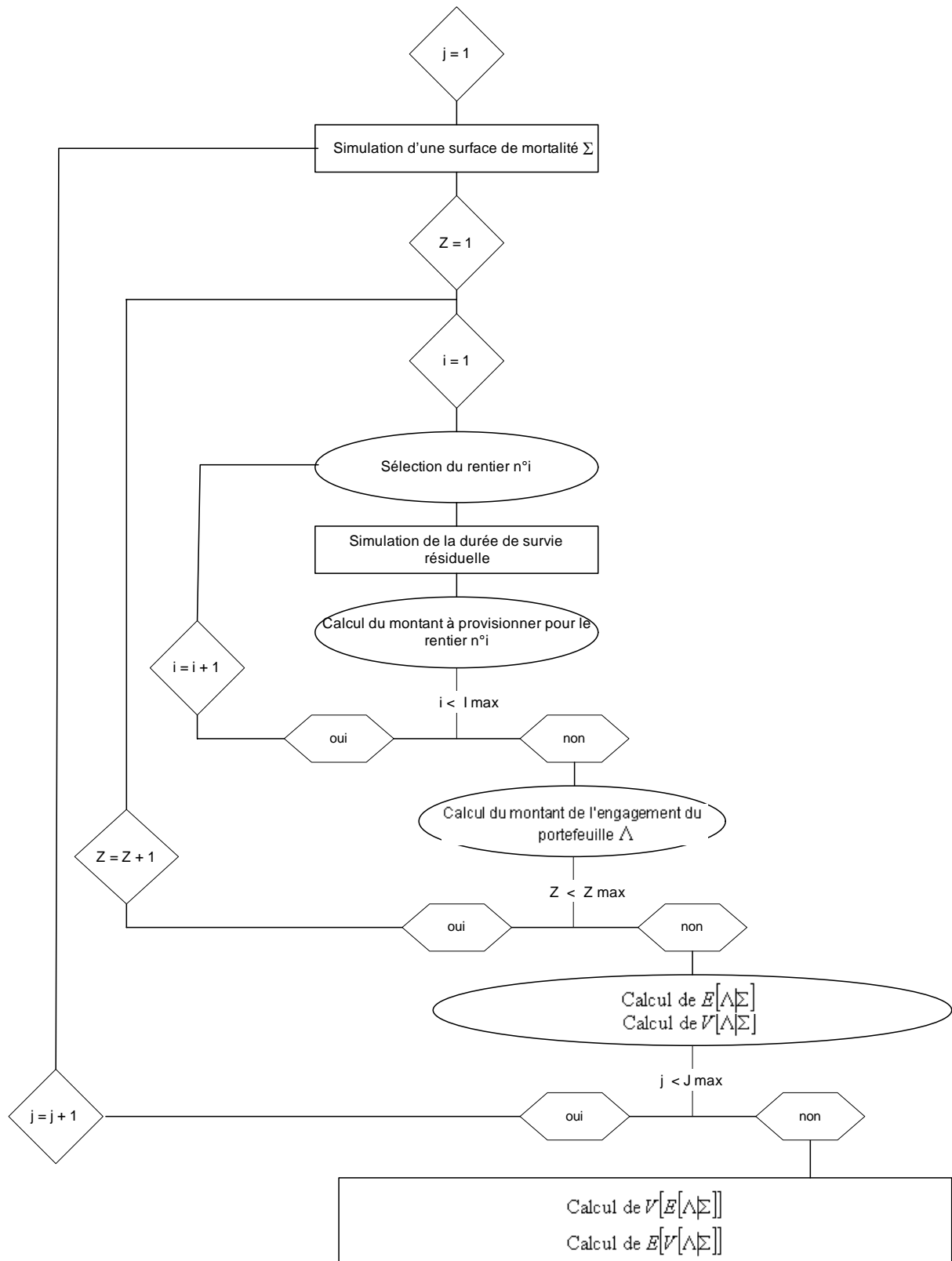
- BROUHNS N., DENUIT M., VERMUNT J.K. [2002] « A POISSON LOG-BILINEAR REGRESSION APPROACH TO THE CONSTRUCTION OF PROJECTED LIFETABLES », INSURANCE: MATHEMATICS AND ECONOMICS, VOL. 31, 373-393.
- CAREY J.R., TULAPURKAR S. ED. [2003] LIFE SPAN, EVOLUTIONNARY, ECOLOGICAL AND DEMOGRAPHIC PERSPECTIVES, POPULATION AND DEVELOPMENT REVUE (SUP. TO VOL. 29) , POPULATION COUNCIL.
- CEIOPS [2006], « QUANTITATIVE IMPACT STUDY 2 – TECHNICAL SPECIFICATION », CEIOPS ([HTTP://WWW.CEIOPS.COM](http://www.ceiops.com))
- DENUIT M [2006], CONSTRUCTION DE TABLES DE MORTALITÉ PÉRIODIQUES ET PROSPECTIVES : ECONOMICA
- GUTTERMAN S., VANDERHOOF I.T. [1999] « FORECASTING CHANGES IN MORTALITY: A SEARCH FOR A LAW OF CAUSES AND EFFECTS », NORTH AMERICAN ACTUARIAL JOURNAL, VOL. 2, 135-138.
- HAAS S. [2006] MÉTHODOLOGIE D'ÉVALUATION ÉCONOMIQUE DES TRAITÉS PROPORTIONNELS EN RÉASSURANCE VIE - APPLICATION AU SWAP DE MORTALITÉ, MÉMOIRE D'ACTUAIRAIT, ISFA.
- LEE R.D., CARTER L. [1992] « MODELLING AND FORECASTING THE TIME SERIES OF US MORTALITY », JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION, VOL. 87, 659–671.
- LEE R.D. [2000] « THE LEE–CARTER METHOD OF FORECASTING MORTALITY, WITH VARIOUS EXTENSIONS AND APPLICATIONS », NORTH AMERICAN ACTUARIAL JOURNAL, VOL. 4, 80–93.
- MESLE F., VALLIN J. [2002] « COMMENT AMÉLIORER LA PRÉCISION DES TABLES DE MORTALITÉ AUX GRANDS ÂGES ? LE CAS DE LA FRANCE », POPULATION N°4, INED, 603.
- PLANCHET F., JUILLARD M., FAUCILLON L. [2006], « QUANTIFICATION DU RISQUE SYSTEMATIQUE DE MORTALITE POUR UN REGIME DE RENTES EN COURS DE SERVICE », INSURANCE AND RISK MANAGEMENT, .
- PLANCHET F., JUILLARD M. « MESURE DE L'INCERTITUDE TENDANCIELLE SUR LA MORTALITE», INSURANCE AND RISK MANAGEMENT, SOUMIS.
- PLANCHET F., LELIEUR V. [2006] « CONSTRUCTION DE TABLES DE MORTALITÉ PROSPECTIVES : LE CAS DES PETITES POPULATIONS », À PARAÎTRE DANS LE BULLETIN FRANÇAIS D'ACTUARIAT, VOL. 6, N°13.

PLANCHET F., THÉRON P.E. [2006] MODÈLES DE DURÉE – APPLICATIONS ACTUARIELLES, PARIS : ECONOMICA.

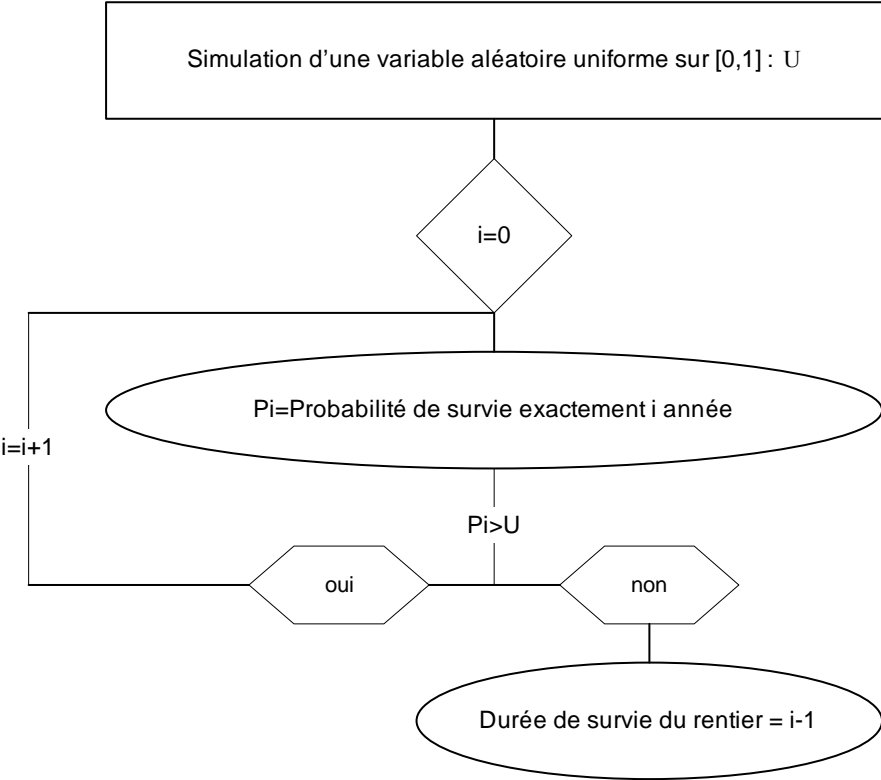
SITHOLE T., HABERMAN S., VERRALL R.J. [2000] « AN INVESTIGATION INTO PARAMETRIC MODELS FOR MORTALITY PROJECTIONS, WITH APPLICATIONS TO IMMEDIATE ANNUITANTS AND LIFE OFFICE PENSIONERS », INSURANCE: MATHEMATICS AND ECONOMICS, VOL. 27, 285–312.

# Annexe

## Algorithme Principal



# Durée de survie résiduelle



# Simulation variable gaussienne à l'aide de la méthode de l'inverse de Moro

