Etude de différentes méthodes d’ajustement de tables de mortalité : application aux données d’une compagnie d’assurance

Robert Langmeier

Travail de mémoire réalisé sous la direction
du professeur François Dufresne

Ecole des HEC
Université de Lausanne

Octobre 2000

TABLE DES MATIERES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES GRAPHIQUES

LISTE DES ANNEXES

REMERCIEMENTS

Je tiens à remercier le Professeur François DUFRESNE pour ses avis éclairés ainsi que Corinne ANTOINE et Nathalie PALEY pour leurs conseils toujours judicieux permettant d'améliorer la compréhension du texte.

1 Introduction

Nous allons commencer, après un bref survol sur les tables de mortalité, par rappeler les hypothèses classiques pour le traitement des âges non entiers.

Les résultats donnés par la méthode des moments seront également répertoriés avant d'approfondir l'estimation par la méthode du maximum de vraisemblance.

Différentes méthodes d'ajustement seront ensuite évoquées, celles-ci permettent d'obtenir une révision des estimés qui devrait aboutir à une courbe, par exemple de mortalité, lissée.

Nous allons également nous intéresser aux mesures permettant de juger de la qualité du lissage produit.

Le cadre théorique posé, nous allons passer à une application avec des données réelles. Ces données ont pu être obtenues à partir du système d'informations des Rentes Genevoises.

Les assurés des Rentes Genevoises proviennent, d'une part, des contrats individuels et, d'autre part, des contrats collectifs.

Les assurés individuels contractent des polices immédiates ou différées de 3$^{\text{ème}}$ pilier libre ou lié ou alors des polices de libre passage et sont soumis au risque d'anti-sélection.

Les assurés collectifs proviennent pour leur part de la reprise de caisses de pensions et ils reçoivent des rentes de vieillesse, de conjoint survivant, d'invalidité, d'enfant de retraités, d'enfant d'invalides et enfin des rentes d'orphelins.

Dans notre étude, nous nous limiterons à l'analyse des tables de mortalité de second ordre dans le cadre de l'assurance individuelle.

2 Remarques préliminaires

2.1 Terminologie

Le but de ce paragraphe est de donner un ensemble de vocabulaire et de définitions. Il est principalement destiné aux lecteurs n'ayant pas encore eu l'occasion de se familiariser avec les concepts liés aux tables de mortalité.

Nous pouvons classer les modèles de survie utilisés en deux catégories:

- Les modèles tabulaires dont les probabilités sont énumérées par âge: table de mortalité, table d'invalidité, …
- Les modèles paramétriques dont les probabilités s'obtiennent en utilisant une formule mathématique: Gompertz, Makeham, …

Les études de mortalité peuvent être de deux types:

- Dans les études *transversales* ("*cross sectional*") sont observés tous les décès survenus durant une période de temps limitée, quel que soit l'âge des personnes qui composent l'effectif. Les modèles tabulaires ou paramétriques établis sur cette base sont dénommés *modèle à données incomplètes*.
- Dans les études *longitudinales* est observée une cohorte d'individus depuis leur naissance jusqu'à leur décès. Les modèles tabulaires ou paramétriques établis sur cette base sont dénommés *modèle à données complètes*. Ces modèles sont surtout utilisés dans les études cliniques.

Les tables de mortalité peuvent être classées en plusieurs catégories:

- Les tables du moment qui décrivent la mortalité telle qu'elle est observée sur une période de temps donnée.
- Les tables de génération décrivent la mortalité d'une cohorte d'individus depuis leur naissance jusqu'à leur décès.
- Les tables de sélection ("*select life tables*"), dans lesquelles pendant une durée appelée période de sélection, le risque de mortalité, par exemple, est aggravé pour finalement rejoindre la table ultime ("*ultimate life table*").

Les tables de premier ordre sont utilisées pour la tarification des primes et sont, en général, établies grâce à la mise en commun des données, des assureurs privés par exemple, afin que la loi des grands nombres puisse s'appliquer. De plus, des marges de sécurité sont prises en compte pour respecter le principe de prudence. A contrario, une table de second ordre représente la mortalité telle qu'elle est observée pour un effectif particulier.

Les études de mortalités à données complètes sont très souvent utilisées dans les études cliniques. Dans ces études, l'observateur pourra examiner une unité jusqu'à ce qu'elle tombe en panne, car la durée est relativement courte. Il est également à noter qu'aucune sortie volontaire n'est possible et que l'on attendra jusqu'à ce qu'il y ait une panne de l'unité. Nous imaginons aisément que ces conditions ne peuvent pas être réunies pour les études de mortalité dans le contexte des caisses de pensions ou des assurances.

2.2 Méthodologie

Pour construire une table de mortalité, qu'elle soit de premier ou de second ordre, nous devons en premier lieu obtenir les données suivantes de l'effectif sous revue; les entrées (nouveaux assurés), les décès, les sorties (départs volontaires) et les survivants.

Si nous modélisons l'invalidité, nous aurons également besoin de données plus précises sur les sorties de l'effectif pour cause d'invalidité ainsi que les entrées dues à la réactivité. Dans notre étude, nous ne pourrons pas étudier cet aspect, car les Rentes Genevoises ne peuvent pas, de par leur statut, reprendre les cotisants d'une caisse de pensions.

Selon la taille de l'effectif, nous aurons besoin de mener la récolte des données sur une ou plusieurs années, une récolte sur dix années étant un maximum, car au-delà l'accroissement de la longévité faussera les résultats de manière sensible. En général, la récolte se fait pour 3, 5 ou 10 ans en fonction de l'effectif sous revue. Pour notre cas, nous travaillerons avec une récolte sur 10 années.

Les données brutes recueillies nous permettent alors de calculer des probabilités observées qui seront très irrégulières. Nous devons donc lisser les probabilités obtenues au moyen d'une technique de "*graduation*" dont le résultat sera une révision des estimés. La "*graduation*" est un compromis entre la régularité ("*smoothing*") et la fidélité ("*fit*") par rapport aux observations.

Nous procéderons ensuite à un diagnostic du lissage de la graduation produite afin de valider la qualité de celle-ci.

2.3 Description du problème

En principe, l'estimation de la probabilité du décès d'un individu entre l'âge x et $x + 1$ à partir des données statistiques devrait être simple. Pour obtenir ce taux, il suffit en effet de diviser le nombre de décès observés pour un âge x donné par la population (du même âge) exposée au risque décès.

Toutefois, en pratique, la réalité est toujours complexe à appréhender. Le diagramme de Lexis permet de représenter les problèmes auxquels nous pouvons être confrontés pour l'estimation des probabilités de décès, plus précisément de l'exposition au risque. Dans ce diagramme, on représente chaque individu sous observation par une diagonale et les axes représentent l'évolution du temps et de l'âge.

Dans la figure 1 ci-dessous, nous avons explicité plusieurs cas, non exhaustifs, qui peuvent se présenter:

1. L'individu est observé complètement entre les âges x et $x + 1$.
2. L'individu est observé alors qu'il est déjà plus âgé que x, il décède ou il quitte l'effectif avant d'avoir atteint l'âge $x + 1$.
3. L'individu entre dans l'effectif pendant la période d'observation alors que son âge est déjà compris entre x et $x + 1$.
4. On se retrouve dans le même cas de figure que 3 et l'individu décède ou quitte l'effectif avant d'avoir atteint l'âge $x + 1$.

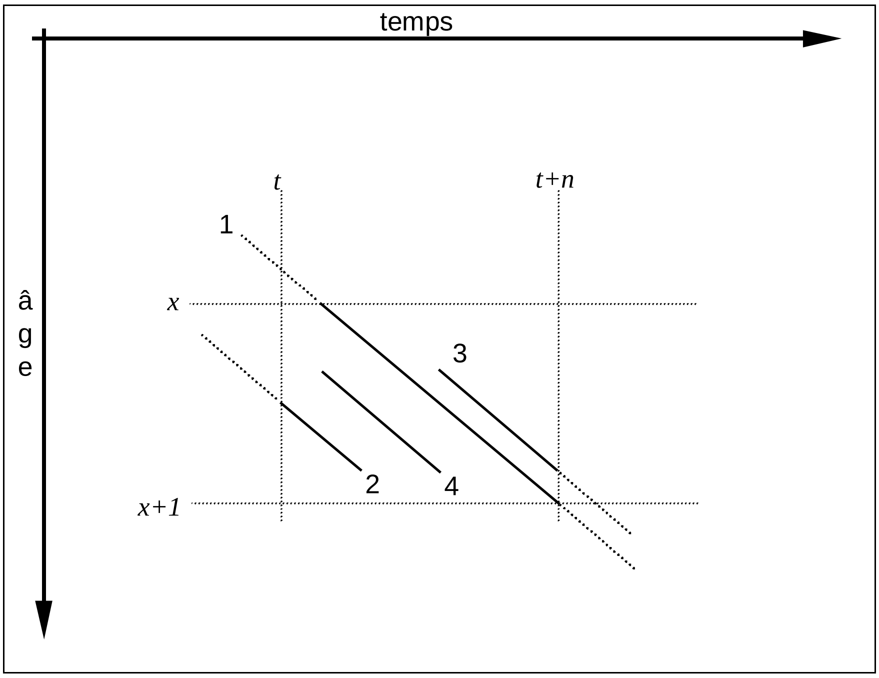


Figure 1 : diagramme de Lexis

Dans chacun des cas de figure ci-dessus, il y a une exposition plus ou moins longue au risque, on parle alors de contribution à l'exposition. Dans le point 3.1 ci-dessous, nous allons examiner comment prendre en compte la contribution à l'exposition.

3 Modèles tabulaires

3.1 Etude à données incomplètes

La notation (x, x + 1] signifie qu'un décès à x + 1 appartient à cet intervalle alors qu'un décès à x appartient à l'intervalle (x - 1,x]. L'indice i implique que nous considérons la i^{ème} personne.

Pour mener une étude de mortalité nous avons besoin des données suivantes:

1. Dates de début et de fin de la période d'observation.

Et pour chaque individu:

2. Date de naissance.
3. Date d'entrée dans l'effectif.
4. Date de sortie ou date de décès.

A partir des données des points 1 à 3 ci-dessus, nous sommes à même de calculer l'âge à l'entrée dans l'étude et l'âge prévu à la terminaison de l'étude. Nous obtenons ainsi pour chaque individu une paire ordonnée (y_i, z_i).

y_i	Age à l'entrée dans l'étude.
z_i	Age prévu à la terminaison de l'étude.

Nous allons également définir l'âge à l'entrée et l'âge prévu à la terminaison de l'intervalle d'observation (x,x + 1]. Pour chaque individu, nous obtenons ainsi une paire ordonnée ($x+r_i$, $x+s_i$) ou plus simplement (r_i, s_i).

$x+r_i$	Age d'entrée de la i^{ème} personne dans l'intervalle d'estimation où $0 \le r_i < 1$.
$x+s_i$	Age de sortie prévu de la i^{ème} personne dans l'intervalle d'estimation où $0 < s_i \le 1$.

La contribution d'un individu dans l'intervalle d'observation (x,x + 1] est déterminée en outre par les règles suivantes :

$$\left.\begin{aligned} z_i &\le x \\ y_i &\ge x+1 \end{aligned}\right\} : \text{ pas de contribution à } (x,x+1].$$

$$y_i \le x \text{ et } z_i \ge x+1 : \text{ contribution pleine à } (x,x+1].$$

$$\left.\begin{array}{c} x < y_i < x+1 \\ \text{et / ou} \\ x < z_i < x+1 \end{array}\right\} : \text{ contribution partielle à } (x,x+1].$$

Si la contribution à l'intervalle $(x,x + 1]$ est nulle; r_i et s_i ne sont pas définis. Si ces derniers sont définis, nous avons également que $y_i \leq x + r_i < x + s_i \leq z_i$.

Les moyens de traitements informatiques actuels permettent de calculer les paires ordonnées (r_i, s_i) individuellement de manière précise; il n’est dès lors pas nécessaire de grouper les données.

Pour l’estimation de q_x, nous serons en présence des trois cas spéciaux définis ci-dessous. Nous évoquerons brièvement ces cas dans l’estimation par la méthode des moments.

Cas spécial	
A	$r_i = 0$, $s_i = 1$ toutes les personnes entrent à x et terminent à $x + 1$.
B	$0 \leq r_i < 1$, $s_i = 1$ tous les individus terminent à $x + 1$, mais au moins un des $r_i > 0$ (Hypothèse de Balducci).
C	$r_i = 0$, $0 < s_i \leq 1$ toutes les personnes entrent à x, mais au moins un des $s_i < 1$ (Hypothèse de distribution uniforme des décès ou DUD).

En plus des âges définis ci-dessus, nous définissons encore les âges suivants:

e_i Age lors du décès, avec $u_i = e_i - x \quad , \quad 0 < u_i < 1$.

f_i Age de la sortie volontaire, avec $t_i = f_i - x \quad , \quad 0 < t_i < 1$.

alors nous pouvons illustrer les cas de figures auxquels nous serons confrontés dans l'intervalle d'estimation $(x,x + 1]$:

1. Pas de décès ni de sortie volontaire durant l’intervalle d'observation.

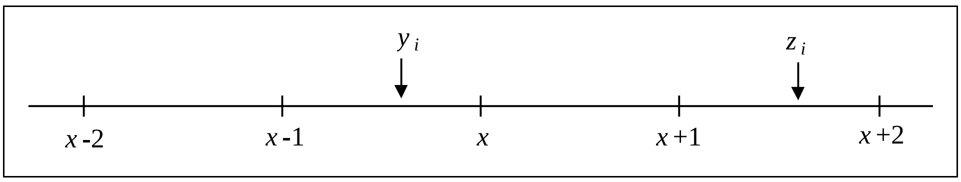


Figure 2 : pas de décès ni de sortie volontaire durant l’intervalle d'observation

DC : Temps du décès
SV : Temps de la sortie volontaire

Intervalle	(r_i,s_i)	DC	SV	Exposition prévue	exacte	actuarielle
$(x-1,x]$	$(y_i,1)$			s_i - r_i	s_i - r_i	s_i - r_i
$(x,x+1]$	$(0,1)$			s_i - r_i	s_i - r_i	s_i - r_i
$(x+1,x+2]$	$(0,z_i)$			s_i - r_i	s_i - r_i	s_i - r_i

Tableau 1 : pas de décès ni de sortie volontaire durant l'intervalle d'observation

2. Décès durant l'intervalle d'observation

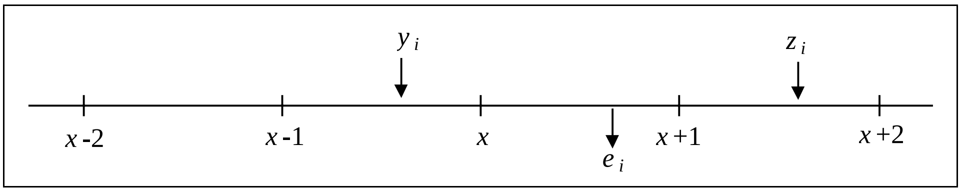


Figure 3 : décès durant l'intervalle d'observation

Intervalle	(r_i,s_i)	DC	SV	Exposition prévue	exacte	actuarielle
$(x-1,x]$	$(y_i,1)$			s_i - r_i	s_i - r_i	s_i - r_i
$(x,x+1]$	$(0,1)$	e_i		s_i - r_i	u_i - r_i	1 - r_i
$(x+1,x+2]$	-			0	0	0

Tableau 2 : décès durant l'intervalle d'observation

Nous pouvons voir que l'exposition actuarielle introduit un biais s'il est prévu que l'individu sorte de l'observation entre x et $x+1$.

3. Sortie volontaire durant l'intervalle d'observation

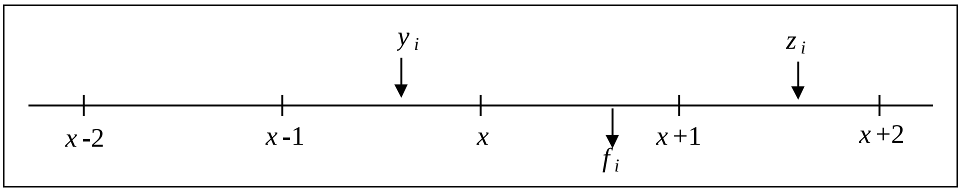


Figure 4 : sortie volontaire durant l'intervalle d'observation

Intervalle	(r_i,s_i)	DC	SV	Exposition prévue	exacte	actuarielle
$(x-1,x]$	$(y_i,1)$			s_i - r_i	s_i - r_i	$s_i - r_i$
$(x,x+1]$	$(0,1)$		f_i	t_i - r_i	t_i - r_i	$t_i - r_i$
$(x+1,x+2]$	-			0	0	0

Tableau 3 : sortie volontaire durant l'intervalle d'observation

L'estimation par la méthode des moments utilise l'exposition prévue, tandis que l'estimation par la méthode du maximum de vraisemblance utilise l'exposition exacte.

3.2 Modèles de survie

Définissons les variables aléatoires suivantes:

X Age au décès d'un nouveau-né
$T(x) = X - x$ Temps de vie futur d'une personne d'âge x

Soit la fonction de répartition

$$F(x) = \Pr[X \le x] = 1 - S(x) \quad , \quad x > 0 \tag{1}$$

et la fonction de survie correspondante

$$S(x) = 1 - F(x) = \Pr[X > x] \tag{2}$$

Probabilité conditionnelle qu'un nouveau-né décède entre x et $x + n$, sachant qu'il a survécu jusqu'à l'âge x

$$\Pr[x < X \le x + n \mid X > x] = \frac{F(x+n) - F(x)}{1 - F(x)} = \frac{S(x) - S(x+n)}{S(x)} \tag{3}$$

Probabilité de décès entre x et $x + t$ pour une personne d'âge x

$$_t q_x = \Pr[T(x) \le t] = 1 - \frac{S(x+t)}{S(x)} \quad , \quad t \ge 0 \tag{4}$$

Probabilité de survie jusqu'à l'âge $x + t$ pour une personne d'âge x

$$_t p_x = 1 - {}_t q_x = \Pr[T(x) > t] = \frac{S(x+t)}{S(x)} \quad , \quad t \ge 0 \tag{5}$$

Taux instantané de décès

$$\Pr[x < X < x + dx \mid X > x] = \frac{F(x+dx) - F(x)}{1 - F(x)} \cong \frac{f(x) \cdot dx}{1 - F(x)} \tag{6}$$

La densité de probabilité conditionnelle est définie par

$$\mu_x = \frac{f(x)}{1 - F(x)} = \frac{\frac{d}{dx}F(x)}{S(x)} = \frac{-\frac{d}{dx}S(x)}{S(x)} = -\frac{d}{dx}\ln S(x) \tag{7}$$

3.3 Hypothèses pour les âges non entiers

3.3.1 Formes pour l_{x+t}

Nous allons récapituler les résultats des trois hypothèses suivantes:

1. Forme linéaire, hypothèse de distribution uniforme des décès (DUD),
2. Forme exponentielle, hypothèse de force constante de mortalité,
3. Forme hyperbolique, aussi appelée hypothèse de Balducci.

La dérivation de toutes les autres fonctions requiert que la valeur de l_{x+t} soit connue pour tous les t , $0 \le t \le 1$.

3.3.2 Résumé des résultats

Fonction	Linéaire (DUD)	Exponentielle (Force constante)	Hyperbolique (Balducci)	
l_{x+t}	$l_x - t \cdot d_x$	$l_x \cdot (p_x)^t = (l_{x+1})^t \cdot (l_x)^{1-t}$	$\left[(1-t)\cdot\frac{1}{l_x} + t \cdot \frac{1}{l_{x+1}}\right]^{-1}$	(8,9,10)
${}_t p_x$	$1 - t \cdot q_x$	$(p_x)^t = e^{-\mu \cdot t}$	$\frac{px}{1-(1-t)\cdot q_x}$	(11,12,13)
${}_t q_x$	$t \cdot q_x$	$1-(1-q_x)^t$	$\frac{t \cdot q_x}{1-(1-t)\cdot q_x}$	(14,15,16)
${}_{1-t} p_{x+t}$	$\frac{p_x}{1-t\cdot q_x}$	$(p_x)^{1-t} = e^{-\mu\cdot(1-t)}$	$1-(1-t)\cdot q_x$	(17,18,19)
${}_{1-t} q_{x+t}$	$\frac{(1-t)\cdot q_x}{1-t\cdot q_x}$	$1-(1-q_x)^{1-t}$	$(1-t)\cdot q_x$	(20,21,22)
μ_{x+t}	$\frac{q_x}{1-t\cdot q_x}$	$\mu = -\ln p_x$	$\frac{q_x}{1-(1-t)\cdot q_x}$	(23,24,25)
${}_t p_x \cdot \mu_{x+t}$	q_x	$\mu \cdot e^{-\mu \cdot t}$	$\frac{q_x \cdot (1-q_x)}{[1-(1-t)\cdot q_x]^2}$	(26,27,28)
L_x	$l_x - \frac{1}{2}\cdot d_x = l_{x+1} + \frac{1}{2}\cdot d_x$	$\frac{d_x}{\mu}$	$-l_{x+1}\cdot\frac{\ln p_x}{q_x}$	(29,30,31)
m_x	$\frac{q_x}{1-\frac{1}{2}\cdot q_x}$	μ	$\frac{(q_x)^2}{-p_x \cdot \ln p_x}$	(32,33,34)

Tableau 4 : résultats pour différentes formes de l_{x+t}

3.4 Estimation par la méthode des moments

Le principe de base qui sous-tend dans cette méthode est le principe statistique qui veut que le nombre attendu de décès entre $(x,x+1]$ pour un échantillonnage donné est égal au nombre de décès observés.

d_x Nombre de décès entre $(x,x+1]$.

D_x Variable aléatoire des décès entre $(x,x+1]$.

n_x Nombre de personnes contribuant à $(x,x+1]$, on parle également de l’exposition.

La résolution de l'équation des moments

$$E[D_x]=\sum_{i=1}^{n_x} {}_{s_i-r_i}q_{x+r_i} \ , \tag{35}$$

nous permet de trouver la forme générale de l'estimateur

$$\hat{q}_x=\frac{d_x}{\sum_{i=1}^{n_x}(s_i-r_i)} \ . \tag{36}$$

ainsi que les résultats suivants

Cas spécial	A	B	C
Estimateur	$\hat{q}_x=\frac{d_x}{n_x}$	$\hat{q}_x=\frac{d_x}{\sum_{i=1}^{n_x}(1-r_i)}$	$\hat{q}_x=\frac{d_x}{\sum_{i=1}^{n_x}s_i}$

(37,38,39)

Tableau 5 : estimateurs pour les trois cas spéciaux par la méthode des moments

3.5 Estimation par la méthode de l’estimateur du maximum de vraisemblance

La méthode de l'estimateur du maximum de vraisemblance (EMV) est une alternative à la méthode des moments. Elle possède de meilleures propriétés statistiques que cette dernière, en particulier dans le cas asymptotique.

Lorsque l'âge précis du décès est connu, nous pouvons utiliser cette information dans la procédure d'estimation. Nous pouvons donc distinguer deux subdivisions de l'EMV: la situation dans laquelle l'âge précis du décès est connu est désignée "situation à données complètes"; lorsque seul le nombre de décès est connu dans l'intervalle $(x,x+1]$, on parle alors de "situation à données partielles". Notons que la méthode des moments est par nature basée sur une "situation à données partielles".

Pour le point de départ, il est nécessaire de fixer la méthode à utiliser, elle sera déterminée par les données disponibles. Cette information sera ensuite utile pour déterminer la fonction de vraisemblance. Nous serons ensuite amenés à poser des hypothèses simplificatrices, comme par exemple la distribution des décès, de telle manière qu'il soit possible de résoudre l'équation de vraisemblance afin d'en extraire l'estimateur de mortalité $\hat{q}_x$.

Le concept de l'âge de sortie prévu de l'intervalle ne jouera pas un grand rôle dans cette méthode.

3.5.1 Données partielles, cas spécial A

Pour la suite de notre développement, nous allons omettre l'indice *x* pour les quantités *n*, *d* et *q* afin d'en faciliter l'écriture.

Sur *n* vies, exactement *d* décèdent dans l'intervalle $(x,x+1]$ et $n-d$ survivent à l'âge $x+1$. Nous reconnaissons ici une distribution binomiale, la fonction de vraisemblance est simplement la probabilité binomiale de l'échantillon, soit

$$L = \frac{n!}{d!(n-d)!}(q)^d(1-q)^{n-d} \quad . \tag{40}$$

Une des propriétés de base de l'EMV est que les constantes multiplicatives peuvent être ignorées, les estimés q_x résultants seront identiques. Donc nous pouvons écrire

$$L \propto (q)^d(1-q)^{n-d} \quad , \tag{41}$$

où $\propto$ signifie "est proportionnel à". Comme nous l'avons écrit précédemment, la solution ne dépend pas des constantes multiplicatives, nous pouvons donc écrire plus simplement

$$L = (q)^d(1-q)^{n-d} \quad . \tag{42}$$

Nous devons maintenant trouver la valeur de *q*, appelée $\hat{q}$, qui maximise l'équation (44). Formellement, si $\hat{q}$ existe tel que $L(\hat{q}) \geq L(q)$, $\forall q$, alors $\hat{q}$ est l'estimateur du maximum de vraisemblance (EMV) de *q*.

Pour trouver $\hat{q}$, nous devons résoudre l'équation $\frac{dL}{dq} = 0$.

Comme nous avons à faire à un produit, il serait nécessaire d'appliquer la règle de dérivation d'un produit, ce qui dans le cas de la formule (42) ne pose pas de problème particulier.

Toutefois, les fonctions de vraisemblance peuvent devenir très complexes. Pour faciliter la résolution, nous posons $l = \ln L$ avant la dérivation, ce qui nous permet de trouver la fonction de log-vraisemblance, ce qui dans notre cas nous donne

$$l = d \cdot \ln q + (n-d) \cdot \ln(1-q), \tag{43}$$

puis nous posons $\frac{dl}{dq} = 0$, d'où

$$\frac{dl}{dq} = \frac{d}{q} - \frac{n-d}{1-q} = 0, \tag{44}$$

qui est l'équation de vraisemblance qui nous permet finalement de trouver

$$\hat{q} = \frac{d}{n}, \tag{45}$$

qui est le même estimateur que celui trouvé par la méthode des moments.

3.5.2 Données complètes

Si nous disposons de l'âge exact x_i du décès pour chaque individu dans l'intervalle d'observation $(x,x + 1]$, alors nous prenons en compte chaque décès individuellement et nous prenons le produit de chaque contribution aux décès dans la fonction de vraisemblance.

La vraisemblance pour le i$^{\text{ème}}$ décès est donnée par la densité de probabilité de décéder à l'âge x_i sachant que l'individu était en vie à l'âge x. Soit la contribution à L du i$^{\text{ème}}$ décès

$$L_i = f(x_i \mid X > x) = \frac{f(x_i)}{S(x)} = \frac{S(x_i) \cdot \mu_{x_i}}{S(x)}. \tag{46}$$

Soit $s_i = x_i - x$, le moment du i$^{\text{ème}}$ décès dans l'intervalle $(x,x + 1]$ où $0 < s_i \leq 1$, alors nous avons

$$L_i = \frac{S(x+s_i) \cdot \mu_{x+s_i}}{S(x)} = {}_{s_i}p_x \cdot \mu_{x+s_i}. \tag{47}$$

La contribution à L de la combinaison de tous les décès est

$$\prod_{i=1}^{d_x} {}_{s_i}p_x \cdot \mu_{x+s_i}. \tag{48}$$

Nous devons bien entendu encore prendre en compte les survivants à la contribution de L. Comme il reste $n_x - d_x$ survivants, la contribution s'écrit comme

$$\left(p_x\right)^{n_x-d_x} = \left(1-q_x\right)^{n_x-d_x} . \tag{49}$$

La vraisemblance totale est alors le produit de ces contributions, soit

$$L = \left(1-q_x\right)^{n_x-d_x} \cdot \prod_{i=1}^{d_x} {}_{s_i}p_x \cdot \mu_{x+s_i} \tag{50}$$

pour la situation à données complètes, cas spécial A.

Pour résoudre l'équation (50) en $\hat{q}$, il est nécessaire de faire des hypothèses qui nous permettront d'exprimer ${}_{s_i}p_x \cdot \mu_{x+s_i}$ en termes de q_x. Nous allons considérer deux hypothèses de ce type.

Sous l'hypothèse de distribution uniforme des décès, nous déduisons de (26) que

$${}_{s_i}p_x \cdot \mu_{x+s_i} = q_x \text{ , } 0 < s_i \leq 1 .$$

La contribution à L va être q_x pour les d_x décès, soit $\left(q_x\right)^{d_x}$, d'où (50) devient

$$L = \left(1-q_x\right)^{n_x-d_x} \cdot \left(q_x\right)^{d_x} , \tag{51}$$

qui est équivalent à (42). Ce qui nous permet de déduire que l'estimateur est le même que celui trouvé en (45). La situation à données complètes, cas spécial A, sous l'hypothèse DUD nous donne le même résultat que la situation à données partielles, cas spécial A, qui lui ne requiert pas d'hypothèse pour son évaluation.

Sous l'hypothèse exponentielle, nous nous souviendrons que μ_{x+s} est une constante, plus précisément $\mu = -\ln p_x$, et de (12), ${}_{s_i}p_x = \left(p_x\right)^{s_i} = e^{-\mu \cdot s_i}$, d'où (50) devient

$$\begin{aligned} L &= \left(p_x\right)^{n_x-d_x} \cdot \prod_{i=1}^{d_x} \left(p_x\right)^{s_i} \cdot \mu \\ &= \mu^{d_x} \cdot \exp\left[-\mu\left(n_x-d_x\right) - \mu \cdot \sum_{i=1}^{d_x} s_i\right] . \end{aligned} \tag{52}$$

Nous en tirons la log-vraisemblance

$$l = d_x \cdot \ln \mu - \mu\left[\left(n_x-d_x\right) + \sum_{i=1}^{d_x} s_i\right], \tag{53}$$

d'où

$$\frac{dl}{d\mu} = \frac{d_x}{\mu} - \left[(n_x - d_x) + \sum_{i=1}^{d_x} s_i \right] = 0\,, \tag{54}$$

qui peut être aisément résolu pour

$$\hat{\mu} = \frac{d_x}{(n_x - d_x) + \sum_{i=1}^{d_x} s_i}\,. \tag{55}$$

Le dénominateur de (55) est l'exposition exacte, ce qui implique que nous estimons m_x qui est identique à μ sous l'hypothèse exponentielle. Donc nous obtenons que

$$\hat{q}_x = 1 - e^{-\hat{\mu}} \tag{56}$$

est l'estimateur du maximum de vraisemblance de q_x.

3.5.3 Données complètes, forme générale

Nous allons définir δ_i comme étant une variable indicatrice où

$$\delta_i = \begin{cases} 1 & \text{si la i}^{\text{ème}} \text{ personne décède entre } (x, x+1]\,, \\ 0 & \text{sinon}\,. \end{cases} \tag{57}$$

Nous allons poser $x + t_i$, l'âge auquel la i^ème personne cesse d'être observée, soit qu'elle est survivante, soit qu'elle décède ou alors par suite de sortie volontaire ou de fin d'observation.

Nous avons alors

- $t_i = 1$ et $\delta_i = 0$, la personne i est survivante,
- $t_i < 1$ et $\delta_i = 0$, la personne i cesse d'être observée,
- $t_i \leq 1$ et $\delta_i = 1$, la personne i est décédée

A l'âge $x + r_i$, il est prévu que la personne i cesse d'être observée à un âge quelconque pas plus grand que $x + 1$, mais la possibilité d'un décès prématuré implique que l'âge auquel l'observation cessera est aléatoire. Donc t_i est la réalisation de la variable aléatoire T_i. En d'autres termes, r_i, le début de l'observation, n'est pas une variable aléatoire.

Nous pouvons voir que si $r_i = 0$ pour tous les i, nous avons soit le cas spécial A ou le cas spécial C. De manière similaire, nous pouvons voir que si $t_i = 1$ pour tous les i pour lesquels $\delta_i = 0$ (c.-à-d. pour tous les i qui ne décèdent pas), nous avons soit le cas spécial A ou le cas spécial B. Nous avons résumé ces cas de figure dans le tableau ci-dessous.

	$t_i = 1$ pour tous les i pour lesquels $\delta_i = 0$	$t_i < 1$ pour tous les i pour lesquels $\delta_i = 0$
$r_i = 0$ pour tous les i	Cas spécial A	Cas spécial C
$r_i > 0$ pour certains i	Cas spécial B	Cas général

Tableau 6 : résumé des cas spéciaux issus de la forme générale

La forme générale de la contribution à L de la i$^{\text{ème}}$ personne est

$$L_i = {}_{t_i - r_i} p_{x+r_i} \cdot \left(\mu_{x+t_i}\right)^{\delta_i} \ . \tag{58}$$

Si $\delta_i = 0$, la vraisemblance se résume à la probabilité de survie entre $x + r_i$, et $x + t_i$, et si $\delta_i = 1$, c'est la densité de probabilité de décéder à $x + t_i$ sachant que cette personne est en vie à $x + r_i$,. La vraisemblance totale s'écrit comme

$$L = \prod_{i=1}^{n_x} {}_{t_i - r_i} p_{x+r_i} \cdot \left(\mu_{x+t_i}\right)^{\delta_i} \ . \tag{59}$$

Pour évaluer (59), nous devons faire des hypothèses sur la distribution des décès. Nous allons à nouveau considérer deux cas.

3.5.4 Données complètes, cas général, distribution exponentielle

Sous l'hypothèse exponentielle (force de mortalité constante), nous trouvons que

$$L = (\mu)^{d_x} \cdot \prod_{i=1}^{n_x} e^{-(t_i - r_i)\cdot\mu} \ , \tag{60}$$

d'où nous obtenons la log-vraisemblance

$$l = \ln L = d_x \ln \mu - \mu \cdot \sum_{i=1}^{n_x} (t_i - r_i) \, , \tag{61}$$

et il est aisé de voir que $\frac{dl}{dq} = 0$ produit

$$\hat{\mu} = \frac{d_x}{\sum_{i=1}^{n_x} (t_i - r_i)} \, . \tag{62}$$

L'estimateur (62) est de la même forme que l'estimateur (55), plus précisément le rapport entre d_x avec l'exposition exacte de l'échantillon qui est le taux central de décès de l'échantillon.

Nous avons que (62) est l'EMV du cas général à données complètes, les cas spéciaux A, B et C sont compris dans ce résultat.

3.5.5 Données complètes, cas général, distribution uniforme

Pour évaluer (59) sous l'hypothèse de distribution uniforme des décès, notons en premier lieu que

$$ {}_{t_i - r_i} p_{x+r_i} = 1 - {}_{t_i - r_i} q_{x+r_i} = 1 - \frac{(t_i - r_i) \cdot q_x}{1 - r_i \cdot q_x} = \frac{1 - t_i \cdot q_x}{1 - r_i \cdot q_x}, \tag{63} $$

et

$$ \mu_{x+t_i} = \frac{q_x}{1 - t_i \cdot q_x}. \tag{64} $$

Lorsque (63) et (64) sont substitués dans (59), nous trouvons que (59) contient q_x pour chaque décès, $(1 - t_i \cdot q_x)$ pour chaque survivant de l'intervalle ($t_i = 1$) et pour chaque cessation d'observation ($t_i < 1$), et $(1 - r_i \cdot q_x)^{-1}$ pour tout le monde. Donc (59) devient

$$ L = (q_x)^{d_x} \cdot \prod_{i=1}^{n_x} (1 - r_i \cdot q_x)^{-1} \cdot \prod_{S\&E} (1 - t_i \cdot q_x). \tag{65} $$

Alors $\frac{dl}{dq} = \frac{d}{dq} \ln L = 0$ produit

$$ \frac{d_x}{q_x} + \sum_{i=1}^{n_x} \frac{r_i}{1 - r_i \cdot q_x} - \sum_{S\&E} \frac{t_i}{1 - t_i \cdot q_x} = 0. \tag{66} $$

En général, l'équation (66) doit être résolue en $\hat{q}_x$ par itérations. Notons que si r_i est différent pour chacune des n_x personnes, et que si t_i est différent pour tous les n_x - d_x personnes qui ne décèdent pas, alors il va y avoir $1 + n_x + (n_x - d_x) = 2 \cdot n_x - d_x + 1$ dénominateurs distincts dans (66) qui vont produire une équation en q_x de degré $2 \cdot n_x - d_x$. Nous pouvons néanmoins nous attendre à ce qu'il y ait beaucoup de $r_i = 0$ et beaucoup de $t_i = 1$, ce qui fera que le degré du polynôme ne sera pas si élevé que cela.

Comme l'équation (66) est un polynôme en q_x nous pouvons voir qu'une solution commode existe si le polynôme est de degré un ou deux.

Si $r_i = 0$ pour toutes les n_x personnes dans l'échantillon, et si $t_i = 1$ pour toutes les n_x - d_x personnes qui ne décèdent pas, alors nous avons à faire au cas spécial A et (66) devient (44), qui a deux termes et qui se résoud linéairement pour obtenir (45).

Si $t_i = 1$ pour toutes les n_x - d_x personnes qui ne décèdent pas, $r_i = 0$ pour b_x des n_x personnes dans l'échantillon, $r_i > 0$ pour le reste des $k_x = n_x - b_x$ personnes, alors nous avons le cas spécial B. Si $r_i = r$, une constante pour les k_x personnes nous avons le cas spécial B groupé. Il s'ensuit que (66) devient

$$\frac{d_x}{q_x} + k_x \cdot \frac{r}{1 - r \cdot q_x} - (n_x - d_x) \cdot \frac{1}{1 - q_x} = 0 , \tag{67}$$

de laquelle nous pouvons obtenir l'équation quadratique

$$r \cdot (n_x - k_x) \cdot (q_x)^2 - (n_x - r \cdot k_x + r \cdot d_x) \cdot q_x + d = 0 , \tag{68}$$

ce qui nous amène au résultat

$$\hat{q}_x = \frac{b - \sqrt{b^2 - 4 \cdot r \cdot (n_x - k_x) \cdot d_x}}{2 \cdot r \cdot (n_x - k_x)} , \tag{69}$$

où $b = n_x - r \cdot k_x + r \cdot d_x$. Le signe négatif est utilisé, car l'utilisation du signe positif entraînerait $\hat{q}_x > 1$.

De manière similaire, pour le cas spécial C groupé avec âge de fin $x + s$ pour toutes les personnes dont il est prévu qu'elles cessent d'être observées, l'équation (66) est également quadratique. Dans ce cas, $r_i = 0$ pour les n_x personnes de l'échantillon, $t_i = s$ pour les e_x personnes cessant d'être observées, $t_i = 1$ pour les n_x - e_x - d_x survivants de l'intervalle. Alors (66) devient

$$\frac{d_x}{q_x} - e_x \cdot \frac{s}{1 - s \cdot q_x} - (n_x - e_x - d_x) \cdot \frac{1}{1 - q_x} = 0 , \tag{70}$$

qui nous amène au résultat

$$\hat{q}_x = \frac{b - \sqrt{b^2 - 4 \cdot s \cdot n_x \cdot d_x}}{2 \cdot s \cdot n_x} , \tag{71}$$

où $b = n_x - (1 - s) \cdot e_x - s \cdot d_x$.

Pour compléter notre analyse de l'équation (66), notons que dans le cas général, si k_x personnes entrent à l'âge commun $x + r$ et e_x personnes cessent d'être observées à un âge commun $x + s$, la présence d'un groupe de $n_x - k_x$ à $r_i = 0$ et d'un groupe de $n_x - e_x - d_x$ à

$t_i = 1$ implique qu'il y aura quatre différents dénominateurs dans l'équation (66), cela produira une équation cubique en q_x.

3.5.6 Données partielles, cas spécial C

Nous supposons que toutes les personnes entrent à l'âge exact x dans l'intervalle $(x,x + 1]$ ($r_i = 0$) et que c_x de celles-ci ont une cessation d'observation prévue commune ou moyenne à $x + s$, $0 < s < 1$. Le solde des $n_x – c_x$ reste dans l'échantillon à $x + 1$.

Notons que c_x et $n_x – c_x$ forment des échantillonnages binomiaux séparés. Un membre de c_x peut décéder avant l'âge $x + s$, ou peut survivre jusqu'à cet âge et sortir de l'observation à ce moment.

Si nous posons $d_x^{'}$ le nombre de décès observés parmi l'échantillon de c_x et $d_x^{''}$ le nombre de décès observés dans l'échantillon $n_x – c_x$, alors le nombre total de décès observé est $d_x = d_x^{'} + d_x^{''}$. Notons également que $e_x = c_x - d_x^{'}$ est le nombre observé des individus qui sortent de l'observation à $x + s$.

La contribution à la vraisemblance de l'échantillon c_x est

$$L_{c_x} = \left({}_s q_x\right)^{d_x^{'}} \cdot \left(1 - {}_s q_x\right)^{c_x - d_x^{'}} , \tag{72}$$

et la contribution de l'échantillon $n_x – c_x$ est

$$L_{n_x - c_x} = \left(q_x\right)^{d_x^{''}} \cdot \left(1 - q_x\right)^{n_x - c_x - d_x^{''}} . \tag{73}$$

Comme les deux groupes sont des échantillons binomiaux indépendants, la vraisemblance totale est le produit de (72) et (73), donc nous avons

$$L = \left({}_s q_x\right)^{d_x^{'}} \cdot \left(1 - {}_s q_x\right)^{c_x - d_x^{'}} \cdot \left(q_x\right)^{d_x^{''}} \cdot \left(1 - q_x\right)^{n_x - c_x - d_x^{''}} . \tag{74}$$

Si cette vraisemblance est évaluée sous l'hypothèse de distribution uniforme des décès, nous avons ${}_s q_x = s \cdot q_x$, et nous trouvons

$$L = (s)^{d_x^{'}} \cdot \left(q_x\right)^{d_x^{'}} \cdot \left(1 - s \cdot q_x\right)^{c_x - d_x^{'}} \cdot \left(q_x\right)^{d_x^{''}} \cdot \left(1 - q_x\right)^{n_x - c_x - d_x^{''}} . \tag{75}$$

Si nous omettons la constante multiplicative $s^{d_x^{'}}$ et nous combinons les deux termes $q_x^{d_x^{'}}$ et $q_x^{d_x^{''}}$, nous avons

$$L = q_x^{d_x} \cdot \left(1 - s \cdot q_x\right)^{c_x - d_x^{'}} \cdot \left(1 - q_x\right)^{n_x - c_x - d_x^{''}} , \tag{76}$$

et de là nous pouvons tirer

$$\frac{d}{dq}\ln L = \frac{d_x}{q_x} - \left(c_x - d_x^{'}\right)\frac{s}{1 - s \cdot q_x} - \left(n_x - c_x - d_x^{''}\right)\frac{1}{1 - q_x} = 0\,. \tag{77}$$

Mais $c_x - d_x^{'} = e_x$, et $n_x - c_x - d_x^{''} = n_x - c_x - \left(d_x - d_x^{'}\right) = n_x - d_x - e_x$, d'où nous pouvons voir que (77) est identique à (70), donc il va en résulter l'estimateur (71). Nous avons trouvé que l'EMV du cas spécial C à données groupées sous l'hypothèse de distribution uniforme des décès est identique, que les données complètes ou partielles soient disponibles.

Si la vraisemblance (74) est évaluée sous l'hypothèse de distribution exponentielle des décès, avec ${}_s p_x = (p_x)^s$, nous trouvons

$$L = \left(1 - p_x^s\right)^{d_x^{'}} \cdot \left(p_x^s\right)^{c_x - d_x^{'}} \cdot (q_x) d_x^{''} \cdot (1 - q_x)^{n_x - c_x - d_x^{''}}\,. \tag{78}$$

En utilisant $e_x = c_x - d_x^{'}$, nous pouvons voir que le deuxième facteur de (78) peut être écrit comme $(p_x)^{s \cdot e_x}$, et ensuite combiné avec le dernier facteur pour résulter en

$$L = p^{s \cdot e_x + n_x - c_x - d_x^{''}} \cdot \left(1 - p_x\right)^{d_x^{''}} \cdot \left(1 - p_x^s\right)^{d_x^{'}}\,, \tag{79}$$

d'où nous obtenons

$$\frac{d}{dp}\ln L = \frac{s \cdot e_x + n_x - c_x - d_x^{''}}{p_x} - \frac{d_x^{''}}{1 - p_x} - \frac{d_x^{'} \cdot s \cdot p_x^{s-1}}{1 - p_x^s} = 0\,. \tag{80}$$

L'équation de vraisemblance obtenue en (80) doit, en général, être résolue au moyen de méthodes numériques pour trouver l'EMV de $\hat{p}_x$ (et ensuite obtenir $\hat{q}_x = 1 - \hat{p}_x$). Un cas spécial existe pour $s = 1/2$; l'équation (80) est alors quadratique en $x = p^{1/2}$ et peut être résolue en $\hat{p}_x$.

4 Méthodes d'ajustement

4.1 Rappels

4.1.1 Différences avant

Les différences avant d'ordre 1 à *z* sont définies, pour un intervalle unitaire, comme suit:

$$\Delta f(x) = f(x+1) - f(x) \tag{81}$$

$$\begin{aligned}\Delta^2 f(x) &= \Delta(\Delta f(x)) = \Delta(f(x+1) - f(x)) \\ &= f(x+2) - 2 \cdot f(x+1) + f(x)\end{aligned} \tag{82}$$

$$\Delta^z f(x) = \Delta(\Delta^{z-1} f(x)) = \Delta(\Delta(\Delta^{z-2} f(x))) = \ldots \tag{83}$$

4.1.2 Différences arrière

Les différences arrière d'ordre 1 à *z* sont définies, pour un intervalle unitaire, comme suit:

$$\nabla f(x) = f(x) - f(x-1) \tag{84}$$

$$\begin{aligned}\nabla^2 f(x) &= \nabla(\nabla f(x)) = \nabla(f(x) - f(x-1)) \\ &= f(x) - 2 \cdot f(x-1) + f(x-2)\end{aligned} \tag{85}$$

$$\nabla^z f(x) = \nabla(\nabla^{z-1} f(x)) = \nabla(\nabla(\nabla^{z-2} f(x))) = \ldots \tag{86}$$

4.1.3 Différences centrales

Les différences centrales d'ordre 1 à *z* sont définies, pour un intervalle unitaire, comme suit:

$$\delta f(x) = f\left(x + \tfrac{1}{2}\right) - f\left(x - \tfrac{1}{2}\right) \tag{87}$$

$$\begin{aligned}\delta^2 f(x) &= \delta(\delta f(x)) = \delta\left(f\left(x + \tfrac{1}{2}\right) - f\left(x - \tfrac{1}{2}\right)\right) \\ &= f(x+1) - 2 \cdot f(x) + f(x-1)\end{aligned} \tag{88}$$

$$\delta^z f(x) = \delta(\delta^{z-1} f(x)) = \delta(\delta(\delta^{z-2} f(x))) = \ldots \tag{89}$$

4.1.4 Symboles utilisés

u_x Observations pour l'âge x.
w_x Poids associés aux observations.
v_x Valeurs ajustées.

4.2 Whittaker-Henderson

4.2.1 Critère de fidélité "Fit"

$$F = \sum_{x=1}^{n} w_x (v_x - u_x)^2 \tag{90}$$

4.2.2 Critère de régularité "Smoothness"

$$S = \sum_{x=1}^{n-z} (\Delta^z v_x)^2 \tag{91}$$

où z fixe le degré du polynôme utilisé pour le critère de régularité, z est généralement compris entre 2 et 4.

Nous prenons une combinaison linéaire de la fidélité et de la régularité en mettant plus ou moins l'accent sur la régularité au moyen du paramètre h.

$$M = F + h \cdot S = \sum_{x=1}^{n} w_x (v_x - u_x)^2 + h \cdot \sum_{x=1}^{n-z} (\Delta^z v_x)^2 \tag{92}$$

Les valeurs ajustées v_x pour x = 1, 2, . . . , n, seront celles qui minimisent la mesure composite M qui est fonction des n valeurs inconnues de v_x.

Pour trouver les v_x, il nous faut résoudre les n équations provenant des dérivées partielles de M par rapport à chacun des v_x, tel que le résultat est nul.

Nous pouvons donc écrire

$$\frac{\partial M}{\partial v_x} = 0 \quad , \quad x = 0,1,2,\ldots,n. \tag{93}$$

Pour trouver la solution, nous allons utiliser la notation matricielle. Nous posons les vecteurs et matrices suivants:

$$U = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \quad , \quad V = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \quad , \quad W = \begin{bmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{bmatrix}.$$

Nous pouvons réécrire *F* en notation matricielle

$$F = (V - U)^T \cdot W \cdot (V - U) \tag{94}$$

Nous devons également obtenir le vecteur suivant, qui représente les différences avant d'ordre *z* de *V*, nous avons donc un vecteur colonne de *n*-*z* lignes

$$\Delta^z V = \begin{bmatrix} \Delta^z v_1 \\ \vdots \\ \Delta^z v_{n-z} \end{bmatrix} \tag{95}$$

ce qui nous permet d'écrire *S* en notation matricielle

$$S = (\Delta^z V)^T \cdot (\Delta^z V) \tag{96}$$

Pour trouver $\Delta^z V$, réécrivons ce vecteur en utilisant une matrice spéciale K_z qui contient les coefficients binomiaux d'ordre *z*. Le signe des coefficients alternera et commencera positivement pour *z* pair. La dimension de cette matrice est de $(n-z)\times z$.

$$\Delta^z V = K_z \cdot V \tag{97}$$

La mesure *M* devient

$$\begin{aligned} M &= F + h \cdot S \\ &= (V - U)^T \cdot W \cdot (V - U) + h \cdot (\Delta^z V)^T \cdot (\Delta^z V) \\ &= (V - U)^T \cdot W \cdot (V - U) + h \cdot V^T \cdot K_z^T \cdot K_z \cdot V \end{aligned} \tag{98}$$

En développant, nous trouvons

$$\begin{aligned} M &= V^T \cdot W \cdot V - V^T \cdot W \cdot U - U^T \cdot W \cdot V + U^T \cdot W \cdot U \\ &\quad + h \cdot V^T \cdot K_z^T \cdot K_z \cdot V \end{aligned} \tag{99}$$

Ce qui donne après regroupement

$$M = V^T \cdot W \cdot V - 2 \cdot V^T \cdot W \cdot U + U^T \cdot W \cdot U + h \cdot V^T \cdot K_z^T \cdot K_z \cdot V \tag{100}$$

Dérivons *M* vectoriellement

$$\frac{\partial M}{\partial V} = 2 \cdot W \cdot V - 2 \cdot W \cdot U + 2 \cdot h \cdot K_z^T \cdot K_z \cdot V = 0 \qquad (101)$$

Après résolution, nous trouvons

$$W \cdot V + h \cdot K_z^T \cdot K_z \cdot V = W \cdot U \qquad (102)$$

En posant $C = W + h \cdot K_z^T \cdot K_z$, nous pouvons écrire

$$C \cdot V = W \cdot U \qquad (103)$$

Si *C* est régulière, nous pouvons obtenir la solution suivante

$$V = C^{-1} \cdot W \cdot U \qquad (104)$$

4.3 Interpolation à jonctions lisses

Toutes les formules d'interpolations peuvent être écrites sous la forme générale d'Everett, soit

$$v_{x+s} = F(s) \cdot u_{x+1} + F(1-s) \cdot u_x \quad , \quad 0 \le s \le 1 \qquad (105)$$

où

$$F(s) = A(s) + B(s) \cdot \delta^2 + C(s) \cdot \delta^4 \qquad (106)$$

- $A(s)$, $B(s)$, $C(s)$, $\ldots$ sont des polynômes en *s*.
- $F(s)$ n'est pas une fonction, mais un opérateur sur u_x dont les fonctions en *s* sont des coefficients sur les différences centrales.
- δ^{2m} sont les différences centrales d'ordre 2*m*.
- La formule est symétrique.

L'interpolation linéaire est un cas particulier de cette formule générale avec $F(s) = s$, donc $A(s) = s$, $B(s) = C(s) = \ldots = 0$.

4.3.1 Les propriétés des jonctions lisses

1. Continuité des valeurs ajustées au point *x*

$$v_{x-1+s}\Big|_{s=1} = v_{x+s}\Big|_{s=0} \tag{107}$$

Cette condition n'est remplie pour toute suite $\{u_x\}$ que si

$$F(0)=0 \quad \Rightarrow \quad A(0)=B(0)=C(0)=\ldots=0 \tag{108}$$

2. Continuité de la dérivée première au point *x*

$$v'_{x-1+s}\Big|_{s=1} = v'_{x+s}\Big|_{s=0} \tag{109}$$

La condition n'est remplie pour toute suite $\{u_x\}$ que si

$$2\cdot F'(1) = F'(0)\cdot\left(2+\delta^2\right) \tag{110}$$

Si les conditions 1 et 2 sont remplies, la formule est dite "tangentielle".

3. Continuité de la dérivée seconde au point *x*

$$v''_{x-1+s}\Big|_{s=1} = v''_{x+s}\Big|_{s=0} \tag{111}$$

La condition n'est remplie pour toute suite $\{u_x\}$ que si

$$F''(0)=0 \quad \Rightarrow \quad A''(0)=B''(0)=C''(0)=\ldots=0 \tag{112}$$

Si les conditions 1, 2 et 3 sont remplies, la formule est dite "osculatoire".

4.3.2 Fonction interpolante ou lissante

Comme (108) doit être satisfait pour qu'il y ait continuité des valeurs ajustées, nous avons que

$$v_{x+s}\Big|_{s=0} = A(1)u_x + B(1)\delta^2 u_x + \cdots. \tag{113}$$

Nous pouvons voir que la fonction est interpolante ($v_x = u_x$ pour x entier) si et seulement si

$$F(1) = 1 \quad \Rightarrow \quad A(1) = 1 \quad , \quad = B(1) = C(1) = \ldots = 0 \,, \tag{114}$$

La fonction sera lissante lorsque $v_x \neq u_x$. Nous voyons clairement que les valeurs de $B(1)$, $C(1)$, … vont jouer un rôle déterminant dans l'écart de v_x par rapport à u_x.

4.3.3 Degré de fidélité

Une caractéristique importante d'une formule d'interpolation est le degré du polynôme qui la compose. Si le degré du polynôme à interpoler est plus petit ou égal à z et que la formule d'interpolation est exacte pour le degré z, alors le résultat sera exact.

Cette caractéristique n'a rien à voir avec les différences d'ordres supérieures apparaissant dans $F(s)$, mais dépend uniquement du choix des fonctions $A(s)$, $B(s)$, etc.

Degré ***d***	**Conditions cumulatives**
0	$A(s) + A(1-s) = 1$
1	$A(s) = s$
2	$B(s) + B(1-s) = \frac{1}{2} s \cdot (s-1)$
3	$B(s) = \frac{1}{6} s \cdot (s^2 - 1)$
4	$C(s) + C(1-s) = \frac{1}{24} s \cdot (s^2 - 1) \cdot (s-2)$
$\vdots$	$\vdots$

Tableau 7 : Conditions de reproduction d'un polynôme de degré ≤ *d*

4.3.4 Famille de formules à quatre points

Nous connaissons $u_{x-1}, u_x, u_{x+1}, u_{x+2}$ et nous cherchons à interpoler v_{x+s} , $0 \le s \le 1$. Nous devons alors déterminer $F(s) = A(s) + B(s) \cdot \delta^2$.

La reproduction des fonctions linéaires implique que $A(s) = s$.
La continuité des valeurs ajustées implique que l'on pose $B(0) = 0$.
La reproduction des valeurs observées, $B(1) = L$, dépend d'un paramètre de contrôle.
La continuité de la dérivée première nous permet de trouver $B'(0) = 0$ et $B'(1) = \frac{1}{2}$.

En résolvant $B(s)$ pour un degré de fidélité égal à 3, nous trouvons

$$B(s) = \left(3 \cdot L - \frac{1}{2}\right) \cdot s^2 + \left(\frac{1}{2} - 2 \cdot L\right) \cdot s^3 \ . \qquad (115)$$

De cette formule, trois cas peuvent être tirés, ceux-ci sont récapitulés dans le tableau ci-dessous.

Cas	*L*	
Karup-King	0	$B(s) = \frac{1}{2} \cdot s^2 \cdot (s-1)$
seul membre de la famille qui est quadratique	$\frac{1}{4}$	$B(s) = \frac{1}{4} \cdot s^2$
seul membre de la famile qui est osculatoire	$\frac{1}{6}$	$B(s) = \frac{1}{6} \cdot s^3$

Tableau 8 : Membres spéciaux de la famille à quatre points

Nous utiliserons la formule de Karup-King dans l'application que nous aborderons dans l'application pratique.

4.4 Splines avec points de jonctions

Il n'est pas toujours possible, si la quantité de données est grande, d'obtenir une "graduation" satisfaisante avec une seule fonction. Dans ces cas, il est possible de lisser la fonction en la découpant en sous-intervalles en prenant des précautions particulières aux points de jonctions.

Une des caractéristiques fondamentales de la "graduation" par "splines" est que les fonctions utilisées sur les sous-intervalles peuvent être beaucoup plus simples qu'une fonction qui devrait couvrir l'intervalle complet.

Nous allons considérer uniquement les fonctions de polynômes du troisième degré et ajuster le "spline" cubique aux estimés initiaux u_x par les moindres carrés.

Observations préliminaires sur cette méthode de "graduation":

1. La séquence de révision des estimés v_x est donnée par une fonction en x qui est appelée "spline".
2. Le "spline" est composé par deux ou plusieurs courbes cubiques, aussi dénommés arcs, qui se rejoignent de manière continue et régulière (même pente et courbure).
3. Il n'est pas nécessaire que $v_x = u_x$ pour une valeur particulière de x.
4. Les valeurs de u_x n'ont pas besoin d'être connues pour toutes les valeurs de x pour lesquelles v_x est désiré. Si c'est le cas, le "spline" est interpolant.
5. Les paramètres du "spline" vont être déterminés par une régression par les moindres carrés sur les estimés initiaux u_x.

4.4.1 Spline cubique à deux arcs

On choisit un nœud k afin de partager l'intervalle $[a,b]$.

$$v_x = \begin{cases} p_0(x) & a \le x \le k \\ p_1(x) & k \le x \le b \end{cases}, \tag{116}$$

où $p_i(x)$ est un polynôme du 3$^{\text{ème}}$ degré.

Nous cherchons à minimiser

$$S = \sum_{x=a}^{b} w_x \cdot (u_x - v_x)^2 = \sum_{x=a}^{h} w_x \cdot (u_x - p_0(x))^2 + \sum_{x=h+1}^{b} w_x \cdot (u_x - p_1(x))^2 \tag{117}$$

où h est la plus grande valeur de $x \le k$ pour laquelle il existe une valeur de u_x.

Pour obtenir une jonction lisse entre $p_0(x)$ et $p_1(x)$, nous devons remplir les conditions suivantes et comme $p_0(x)$ et $p_1(x)$ sont cubiques, elles sont différentiables deux fois.

$$\begin{aligned} p_0(k) &= p_1(k), \\ p_0'(k) &= p_1'(k), \\ p_0''(k) &= p_1''(k). \end{aligned} \tag{118}$$

Posons

$$p_0(x) = c_1 + c_2 \cdot x + c_3 \cdot x^2 + c_4 \cdot x^3 \tag{119}$$

et

$$\begin{aligned} p_1(x) &= p_0(x) + c_5 \cdot (x-k)^3 \\ &= c_1 + c_2 \cdot x + c_3 \cdot x^2 + c_4 \cdot x^3 + c_5 \cdot (x-k)^3 . \end{aligned} \tag{120}$$

En substituant (119) et (120) dans (118), nous trouvons

$$\begin{aligned} S = &\sum_{x=a}^{h} w_x \cdot \left(u_x - c_1 - c_2 \cdot x - c_3 \cdot x^2 - c_4 \cdot x^3\right)^2 \\ &+ \sum_{x=h+1}^{b} w_x \cdot \left(u_x - c_1 - c_2 \cdot x - c_3 \cdot x^2 - c_4 \cdot x^3 - c_5 \cdot (x-k)^3\right)^2 \end{aligned} \tag{121}$$

Il suffit alors de prendre les dérivées partielles de cette équation par rapport à c_1, c_2, c_3, c_4, égaliser à zéro et résoudre les équations normales résultantes.

4.4.2 Spline cubique, cas général

Nous disposons de m observations de u_x aux points $x = a, \ . \ . \ . \ , b$. Si nous divisons l'intervalle $[a,b]$ en $(n + 1)$ par n nœuds aux points $x = k_1, k_2, \ . \ . \ . \ , k_n$, alors

$$v_x = \begin{cases} p_0(x) & a \le x \le k_1 \\ p_i(x) & k_i \le x \le k_{i+1} \\ p_n(x) & k_n \le x \le b \end{cases} . \tag{122}$$

Les conditions imposées sur les $p_i(x)$ sont

$$\left.\begin{aligned} p_{i-1}(k) &= p_i(k) \\ p'_{i-1}(k) &= p'_i(k) \\ p''_{i-1}(k) &= p''_i(k) \end{aligned}\right\} \quad , \quad i = 1,2,...,n \quad . \tag{123}$$

Nous pouvons écrire les équations des $p_i(x)$ comme suit

$$p_i(x) = p_0(x) + c_5(x-k_1)^3 + \cdots + c_{i+4}(x-k_i)^3 \quad , \quad i = 1,2,...,n \quad . \tag{124}$$

En substituant les équations des $p_i(x)$, nous obtenons la somme des erreurs au carré pondérées

$$\begin{aligned} S = & \sum_{x=a}^{h_1} w_x \cdot \left(u_x - c_1 - c_2 \cdot x - c_3 \cdot x^2 - c_4 \cdot x^3\right)^2 \\ & + \sum_{x=h+1}^{h_2} w_x \cdot \left(u_x - c_1 - c_2 \cdot x - c_3 \cdot x^2 - c_4 \cdot x^3 - c_5 \cdot (x-k_1)^3\right)^2 + \cdots \\ & + \sum_{x=h+1}^{b} w_x \cdot \left(u_x - c_1 - \cdots - c_5 \cdot (x-k_1)^3 - \cdots - c_{n+4} \cdot (x-k_n)^3\right)^2 . \end{aligned} \tag{125}$$

Les valeurs de $h_1, h_2, \ldots$ sont définies de la même manière qu'au point 4.4.1.

Le résultat de la minimisation matricielle par rapport à C est

$$X^T \cdot W \cdot X \cdot C = X^T \cdot W \cdot U \ , \tag{126}$$

où

$$C = \underset{(n+4)\cdot 1}{\begin{bmatrix} c_1 \\ c_2 \\ \\ c_{n+4} \end{bmatrix}}, \quad U = \underset{m\cdot 1}{\begin{bmatrix} u_a \\ u_{a+1} \\ \vdots \\ u_b \end{bmatrix}},$$

et

$$W = \begin{bmatrix} w_a & 0 & \cdots & 0 \\ 0 & w_{a+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_b \end{bmatrix}.$$

Nous donnons un exemple de X pour le cas à deux nœuds ($m \cdot 6$)

$$X = \underset{m\cdot(n+4)}{\begin{bmatrix} 1 & a & a^2 & & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & h_1 & h_1^2 & h_1^3 & 0 & 0 \\ 1 & h_1+1 & (h_1+1)^2 & (h_1+1)^3 & (h_1+1-k_1)^3 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & h_2 & h_2^2 & h_2^3 & (h_2-k_1)^3 & 0 \\ 1 & h_2+1 & (h_2+1)^2 & (h_2+1)^3 & (h_2+1-k_1)^3 & (h_2+1-k_2)^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & b & b^2 & b^3 & (b-k_1)^3 & (b-k_2)^3 \end{bmatrix}}$$

Nous pouvons également écrire le résultat sous une forme plus explicite

$$C = \left(X^T \cdot W \cdot X\right)^{-1} \cdot X^T \cdot W \cdot U \,. \tag{127}$$

Les valeurs des coefficients nous permettent alors de calculer v_x en n'importe quel point.

5 Considérations statistiques

5.1 Introduction

Les tests statistiques utilisés pour comparer les observations avec une table de mortalité de référence peuvent aussi être utilisés pour examiner la fidélité de la graduation avec toutefois quelques difficultés lorsqu'ils sont utilisés à cette fin. Les tests statistiques doivent être vus comme une aide à l'évaluation de la graduation et ne pas être interprétés de manière trop rigide.

Dans le choix d'une méthode de graduation, des considérations spéciales peuvent s'appliquer, comme l'application du principe de prudence sur toute la table ou sur un intervalle particulier. Il est, par exemple, important que la mortalité d'une assurance décès ne soit pas sous-estimée. A contrario, on ne devra pas surestimer la mortalité d'une assurance de rente.

5.2 Test du khi-carré

Dans le domaine des tests statistiques, le test du khi-carré est un des plus utilisé pour vérifier l'hypothèse nulle que les décès dans une expérience de mortalité proviennent d'une population selon un taux de mortalité défini. Toutefois il existe certaines limitations sérieuses lorsqu'il est appliqué aux données de mortalités. Un certain nombre de tests moins sophistiqués ont été conçus pour palier ces défauts et sont en général effectués avant le calcul du khi-carré, qui peut être omis si les calculs précédents révèlent une fidélité inadéquate ou une extrêmement bonne fidélité.

Pour apprécier l'utilité des autres tests, il est nécessaire de comprendre le test du khi-carré et ses limitations. Sous l'hypothèse nulle que l'expérience de mortalité est issue d'une population avec des taux de mortalité connus $\{q_x\}$, le nombre de décès θ_x à l'âge x est distribué selon une loi binomiale de paramètres E_x et q_x. Si le nombre de décès prévu $E_x \cdot q_x$ est suffisamment grand (plus grand que 5), la distribution de θ_x est approximativement normale avec une moyenne de $E_x \cdot q_x$, une variance de $E_x \cdot p_x \cdot q_x$ et

$$Y = \sum_x \frac{(\theta_x - E_x \cdot q_x)^2}{E_x \cdot p_x \cdot q_x} \tag{128}$$

a une distribution khi-carré avec n degrés de liberté, où n est le nombre d'âges ou de groupes d'âges différents utilisés dans le calcul. L'hypothèse nulle que la mortalité correspond à la table de référence est usuellement rejetée si la valeur du khi-carré tombe dans la région des 5 pour-cent supérieurs de la distribution.

5.3 Limitations du test khi-carré

Même lorsque le test khi-carré indique qu'il y a une adhérence raisonnable à la table de référence, un examen plus précis des écarts individuels révèle quelquefois de sérieuses discordances entre les observations et la table de référence. Le test statistique (128) comprend la somme des carrés des écarts-types pour des âges individuels ou des groupes d'âges. Il n'est donc pas surprenant qu'il échoue parfois de détecter

1. l'existence d'un certain nombre de grands écarts contrebalancés par un très grand nombre de petits écarts;
2. un grand écart cumulé sur une partie ou sur la totalité de l'intervalle;
3. un excès d'écarts positifs (ou négatifs) sur une partie ou sur la totalité de l'intervalle;
4. un grand excès d'écarts du même signe.

Ces divergences ne sont pas nécessairement mutuellement exclusives.

5.4 Quelques tests complémentaires

Malgré que le test du khi-carré fourni une aide précieuse pour la comparaison entre les observations et la table de référence, d'autres tests sont nécessaires, et certains d'entre eux peuvent être plus importants que le test du khi-carré. Nous allons examiner six tests complémentaires qui sont couramment appliqués. Ceux-ci permettent d'examiner respectivement

1. écart-types standardisés individuels;
2. écart-types absolus;
3. cumul des écart-types;
4. signes des écart-types;
5. groupement des signes;
6. changements des signes.

Les tests 1, 2 et 4 permettent d'examiner la "normalité" de l'écart-type, tandis que les tests 3, 5 et 6 vérifient que la distribution des écart-types est aléatoire par rapport à l'âge.

5.4.1 Ecart-type standardisé individuel

Sous l'hypothèse nulle que la table de référence représente la mortalité réelle, le nombre observé de décès à l'âge *x*, θ_x est une variable aléatoire binomiale de paramètres E_x et q_x, et sa distribution peut être approximée par une variable aléatoire normale de moyenne $E_x \cdot q_x$ et de variance de $E_x \cdot p_x \cdot q_x$. Les décès aux âges successifs sont indépendants, et il en découle que l'écart-type

$$\frac{(\theta_x - E_x \cdot q_x)}{\sqrt{E_x \cdot p_x \cdot q_x}} \tag{129}$$

devrait ressembler à des observations indépendantes selon une loi normale standard.

Il convient donc de vérifier que pas plus d'environ 5 pour-cent de l'écart-type normé excède 2 en valeur absolue. Selon la distribution de la loi normale standard, les pourcentages des écarts-types attendus dans les intervalles sont les suivants

(-∞ , -3)	(-3 , -2)	(-2 , -1)	(-1 , 0)	(0 , 1)	(1 , 2)	(2 , 3)	(3 , ∞)
0%	2%	14%	34%	34%	14%	2%	0%

Les nombres des écarts-types tombant dans ces intervalles devraient être comparés aux nombres attendus calculés selon ces bases.

5.4.2 Ecart-type absolu

La probabilité qu'une loi normale standard soit comprise entre –2/3 et 2/3 est pratiquement égale à 0.5. Il en découle qu'approximativement la moitié de la valeur absolue des écarts standardisés

$$\left|\frac{(\theta_x - E_x \cdot q_x)}{\sqrt{E_x \cdot p_x \cdot q_x}}\right| \tag{130}$$

devrait être inférieure à 2/3 et qu'une autre moitié devrait être supérieure.

Dans le cas de n âges ou groupes d'âges, le nombre N d'écarts-types absolus excédent ⅔ doit être une variable aléatoire binomiale de paramètre n et ½. Nous ne serons généralement pas concernés si la distribution des écarts-types est concentrée de manière plus proche de zéro que ce que la loi normale ne le prédit. Nous n'aurons seulement des doutes de la pertinence de la table de référence si N tombe dans la région des 5 pour-cent supérieure de la distribution binomiale (n, ½). Pour n plus grand que 20, l'approximation normale peut être utilisée, dans ce cas la pertinence de la table de référence sera mise en doute si

$$T = \frac{(2 \cdot N - n)}{\sqrt{n}} \tag{131}$$

tombe dans la région des 5 pour-cent supérieurs de la distribution normale standard (plus grand qu'environ 1.65).

5.4.3 Cumul des écart-types

Les décès aux différents âges sont indépendants, et sous l'hypothèse nulle que la mortalité suit la table de référence, θ_x est approximativement une variable aléatoire normale de moyenne $E_x \cdot q_x$ et de variance de $E_x \cdot p_x \cdot q_x$. Il en découle que l'écart accumulé entre les âges x_1 et x_2

$$\sum_{x=x1}^{x2} \left(\theta_x - E_x \cdot q_x\right) \tag{132}$$

est une variable aléatoire normale de moyenne nulle et de variance

$$\sum_{x=x1}^{x2} E_x \cdot p_x \cdot q_x \tag{133}$$

L'utilisation de la table de référence peut être vue comme suspecte si la valeur absolue du cumul de l'écart-type sur la totalité de la table est plus grand qu'environ le double de la racine carrée de la variance donnée par (133) ou que l'écart-type accumulé sur des parties de la table est trop grand en valeur absolue.

Il n'est pas correct d'appliquer ce test sur un intervalle déterminé par les données, l'intervalle doit être choisi sans référence aux données. Les tests effectués sur différents intervalles ne seront indépendants que si les intervalles ne se recouvrent pas, sinon ils seront positivement corrélés et la proportion des tests produisant un résultat significatif pour l'hypothèse nulle ne sera pas forcément au niveau de 5 pour-cent.

5.4.4 Test du signe

Sous l'hypothèse nulle que la mortalité est conforme à la table de référence, les écart-types des observations de décès par rapport aux prévisions sont des variables aléatoires normales indépendantes. Le signe d'un écart-type individuel a autant de chance d'être positif ou négatif.

Nous serons concernés si l'examen de l'expérience révèle un nombre anormalement élevé d'écarts positifs ou négatifs. Dans le cas de n groupes d'âges, nous allons commencer à émettre des doutes sur l'hypothèse nulle si le nombre d'écarts positifs N tombe dans la région supérieure ou inférieure des 2 ½ pour-cent de la distribution binomiale (n, ½).

Pour n plus grand que 20, l'approximation normale (131) peut être utilisée; l'hypothèse nulle est rejetée si T tombe dans la région supérieure ou inférieure des 2 ½ pour-cent de la distribution normale standard (plus grand qu'approximativement 2 en valeur absolue).

5.4.5 Groupement des signes (Test de Stevens)

Sous l'hypothèse nulle que la mortalité est conforme à la table de référence, les signes des écarts individuels sont indépendants et ont autant de chance d'être positifs ou négatifs. Le nombre d'écarts positifs et négatifs peut être raisonnable selon le test du signe du point 5.4.4, mais nous aurons toujours des doutes sur l'utilisation de la table de référence si le nombre de groupes positifs (et en conséquence le nombre de groupes négatifs) est faible, signifiant que la taille des suites de nombres positifs et négatifs est grande.

Soit

n_+	nombre d'écarts positifs
n_-	nombre d'écarts négatifs
g	nombre de groupes d'écarts positifs

et

$$m = n_+ \cdot \frac{n_- + 1}{n_+ + n_-}, \tag{134}$$

$$v \approx \frac{(n_+ \cdot n_-)^2}{(n_+ + n_-)^3}. \tag{135}$$

Le test de Stevens peut être réalisé approximativement en comparant

$$G = \frac{g - m}{\sqrt{v}} \tag{136}$$

avec la région inférieure des 5 pour-cent de la loi normale standard (plus petit qu'approximativement –1.65).

5.4.6 Test binomial des changements de signes

Sous l'hypothèse nulle que la mortalité est conforme à la table de référence, le signe des écarts successifs a autant de chances d'être de même signe que du signe opposé. Le nombre de changement de signes dans une suite de n écarts est donc binomial de paramètres $n - 1$ et ½.

Si nous examinons une expérience qui révèle trop peu de changement de signes, nous rejetterons la table de référence si le nombre de changement de signes N tombe dans la région des 5 pour-cent inférieurs de la distribution binomiale ($n - 1$, ½).

Ce test permet d'examiner à la fois le nombre de signes positifs et négatifs ainsi que leur groupement, dans un sens, il regroupe le test du signe et le test de Stevens.

Lorsque $n-1$ est plus grand que 20, l'approximation normale de la distribution binomiale peut être utilisée; nous rejetterons la table de référence si

$$S = \frac{2 \cdot N - n + 1}{\sqrt{n-1}} \tag{137}$$

tombe dans la région des 5 pour-cent inférieurs de la distribution normale standard.

Test sur des sections de la table

Les tests que nous avons décrits sont souvent appliqués sur des sections de table de mortalité, tout comme sur la totalité de la table.

6 Application aux données d’une compagnie d’assurance

6.1 Structure des données du système d’information nécessaire

Le système d'informations des Rentes Genevoises contient, entre autre, les entités "Client" et "Police". La notion de client recouvre plusieurs rôles possibles, à savoir:

- Preneur d'assurance
- Assuré(s) (première ou deuxième tête selon le type de police)
- Bénéficiaire(s) en cas de vie ou en cas de décès
- Payeur de prime

Dans notre étude sur la mortalité, notre intérêt se portera uniquement sur les clients ayant un rôle d'assuré.

Nous allons décrire ci-dessous les données qui sont extraites de la base de données de production ainsi que les traitements que nous avons dû mettre en œuvre pour extraire et "nettoyer" les données.

6.1.1 Entité "Client"

Les données qui nous sont utiles sont extraites ou calculées à partir des informations disponibles. Pour l'entité "Client", les données suivantes sont nécessaires ou utiles:

Zone	Source
No assuré	Base de données
Sexe	Base de données
Nom	Base de données
Prénom	Base de données
Date de naissance	Base de données
Date de décès	Base de données
Date d'entrée	Calculée
Date de sortie	Calculée

Tableau 9 : extrait de l'entité "Client" du système d'informations

Les noms et prénoms serviront dans la procédure de détection des doublons.

La date d'entrée correspond à la date d'effet de la plus ancienne des polices de l'assuré, de même la date de sortie correspond à la date d'extinction de la plus récente des polices de l'assuré.

Un assuré principal d'une police sur une tête peut également être le second assuré sur une police sur deux têtes.

6.1.2 Entité "Police"

Pour l'entité "Police", les données suivantes sont nécessaires:

Zone	Source
No police	Base de données
Indicateur individuel/collectif	Base de données
Date d'effet	Base de données
Date d'échéance	Base de données
Date d'extinction	Base de données
No assuré 1	Base de données
No assuré 2	Base de données

Tableau 10 : extrait de l'entité "Police" du système d'informations

L'indicateur "individuel/collectif" permet de différencier:

- les polices conclues par des assurés individuels avec les risques d'anti-sélection que cela présuppose,
- les polices conclues par des contrats avec des caisses de pensions pour la reprise du service des rentes.

La date d'extinction correspond soit à la date de décès de l'assuré, soit à la date de la sortie volontaire.

6.2 Extraction et traitement des données

6.2.1 Qualité des données

Une condition très importante pour le calcul des mortalités brutes concerne la qualité des données collectées.

Les assurés qui sont créés plusieurs fois dans la base de données, alors même que le modèle relationnel adopté dans la conception de la base de données a été prévu pour éviter ce genre de problème, seront appelés "*doublons*". Les assurés ayant les mêmes noms et prénoms seront appelés "*homonymes*".

Une différence dans le calcul de l'exposition avec un "*doublon*" est déjà gênante en soi, mais c'est extrêmement ennuyeux lorsqu'il s'agit d'assurés décédés, car il y a un risque non négligeable que l'on continue à verser la rente si le décès n'est pas saisi sur tous les assurés.

Une série de requêtes a été écrite pour faciliter le travail de détection des doublons. Ces requêtes permettent de détecter visuellement les problèmes suivants:

- Même nom, même prénom et même date de naissance. Ces assurés peuvent être des homonymes, mais la majorité des cas recensés sont des saisies à double,
- Nom pratiquement identique, une lettre diffère,
- Nom de la femme mariée sans et avec le nom de jeune fille,
- Nom de la femme suivi du nom de jeune fille avec et sans tiret.

La vérification approfondie doit se faire sur le système AS/400 et nécessite, selon les cas, une vérification dans les dossiers physiques.

Toutes ces vérifications permettent d'alimenter deux tables, l'une permettant d'identifier les doublons et l'autre les homonymes. Lorsque nous refaisons une extraction, une procédure de mise à jour élimine automatiquement les doublons et évite de présenter à nouveau les homonymes dans la liste des doublons potentiels. Ces mises à jour sont faites localement dans notre base de données et ne perturbent pas la base réelle.

Nous avons passé un temps relativement important sur l'aspect qualitatif et nous pensons que le système d'informations doit être amélioré à ce niveau afin d'obtenir des données correctes, de minimiser les erreurs de saisie et de limiter les risques financiers liés aux problèmes évoqués précédemment.

6.2.2 Autres problèmes rencontrés avec les données

Nous avons trouvé des polices sur deux têtes pour lesquelles il n'existe plus qu'un seul assuré. Ces polices proviennent du premier système informatique, lequel était moins pointu au niveau de la cohérence des données.

Nous avons également détecté des assurés dont la date de sortie est inférieure à la date d'entrée et également des assurés dont la date de décès est inférieure à la date d'entrée.

6.2.3 Extraction des données

L'extraction des données est effectuée depuis le logiciel Access au moyen de requêtes SQL dans des tables liées par ODBC sur le système AS/400 de production.

Les données extraites sont d'abord copiées avec un premier traitement dans des tables locales. Ce premier traitement consiste à transformer les dates enregistrées en format numérique inversé en format "Date/Heure" dans Access. Lors de l'exécution de ces requêtes, toutes les polices et tous les assurés rattachés à ces polices sont pris en compte. Cette phase est appelée "Extraction AS/400". Elle requiert la connexion au système central pour être effectuée; toutes les autres phases peuvent être effectuées sur un poste autonome.

6.2.4 Traitement des données

L'extraction des seuls assurés individuels ou collectifs sera traitée dans la phase suivante, cette phase est appelée "Préparation de l'extraction". Lors de cette phase, de nombreux traitements de "nettoyage" des données sont effectués, plus précisément:

- Suppression des polices sur deux têtes dont seul l'assuré numéro un est défini.
- Elimination des doublons déjà inventoriés.
- Suppression des polices dont la date de sortie est inférieure à la date d'entrée.
- Suppression des polices dont la date de décès est inférieure à la date d'entrée.
- Suppression des assurés qui ne sont plus rattachés à au moins une police.
- Calcul des dates d'entrées.
- Calcul des dates de sorties.

Après cette phase, il faut exécuter les requêtes permettant de vérifier qu'il n'y ait plus de "*doublons*" dans la base de données. Cette problématique est évoquée au point 6.2.1 ci-dessus. Dans le cas où d'autres "*doublons*" ou "*homonymes*" seraient repérés et répertoriés dans les tables adéquates, nous devrions à nouveau exécuter la phase de "Préparation de l'extraction".

Le modèle relationnel des tables utilisées lors de cette phase figure ci-dessous, par contre les tables "ASclient" et "ASpolice" qui ont exactement la même structure que "CRclient" et "CRpolice" n'y figurent pas.

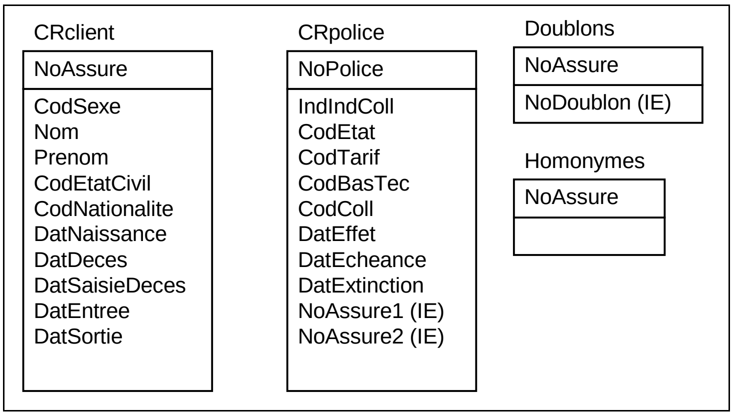


Figure 5 : modèle relationnel – tables de préparation

La troisième phase appelée "Nouvelle extraction" effectue un archivage des données recueillies lors de la préparation. Ainsi nous pouvons faire autant d'estimations de mortalité que nous le souhaitons sur différentes périodes d'observations.

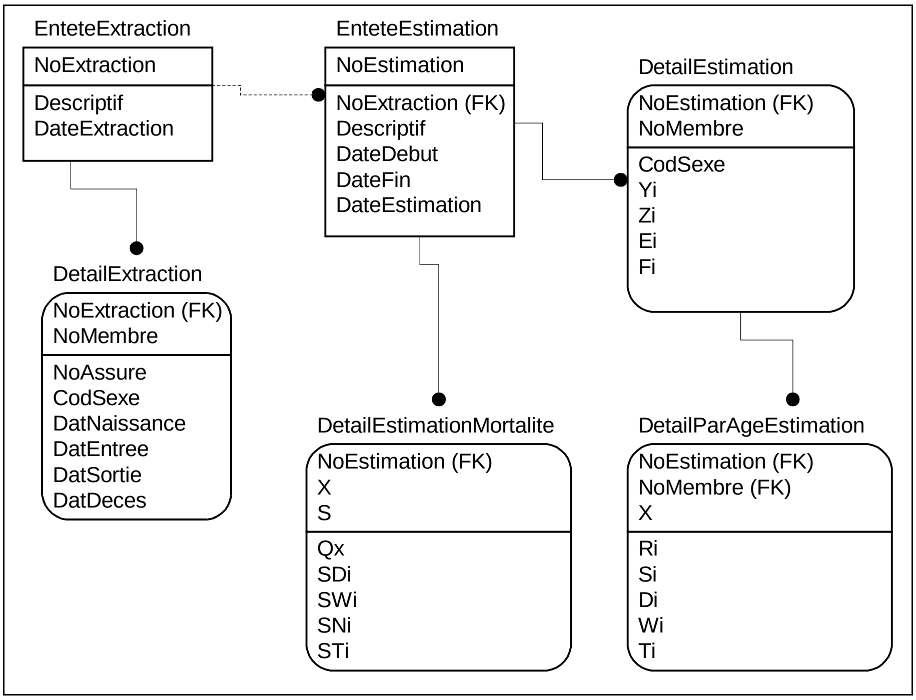


Figure 6 : modèle relationnel – tables d'extractions et d'estimations

La quatrième phase dénommée "Nouvelle observation" génère, à partir des dates de début et de fin d'observation, une estimation de la mortalité par sexe et par âge. Cette phase peut être répétée plusieurs fois avec la même extraction, de façon à pouvoir analyser l'évolution de la mortalité sur une période donnée.

6.3 Probabilités brutes

Les probabilités brutes obtenues pour la période d'observation allant du 01/01/1990 au 31/12/1999 figurent aux annexes A et B.

Nous constatons qu'il y a passablement de lacunes (probabilités nulles) dans les valeurs observées, aucun décès n'ayant pu être observé pour ces âges, car l'effectif sous observation est faible.

Dans le cas des hommes pour des âges compris entre 51 et 97 ans, nous avons trois lacunes à 62, 68 et 73 ans alors que pour les femmes entre 55 et 102 ans, nous avons quatre années consécutives de lacunes entre 64 et 67 ans ainsi qu'à 56, 60 et 72 ans.

Notons également que nous n'observons aucun décès en dessous de 30 ans pour les hommes et 25 ans pour les femmes. D'autre part, nous constatons beaucoup de lacunes entre 25 et 55 ans pour les femmes, ce qui rend le lissage très difficile dans cette tranche.

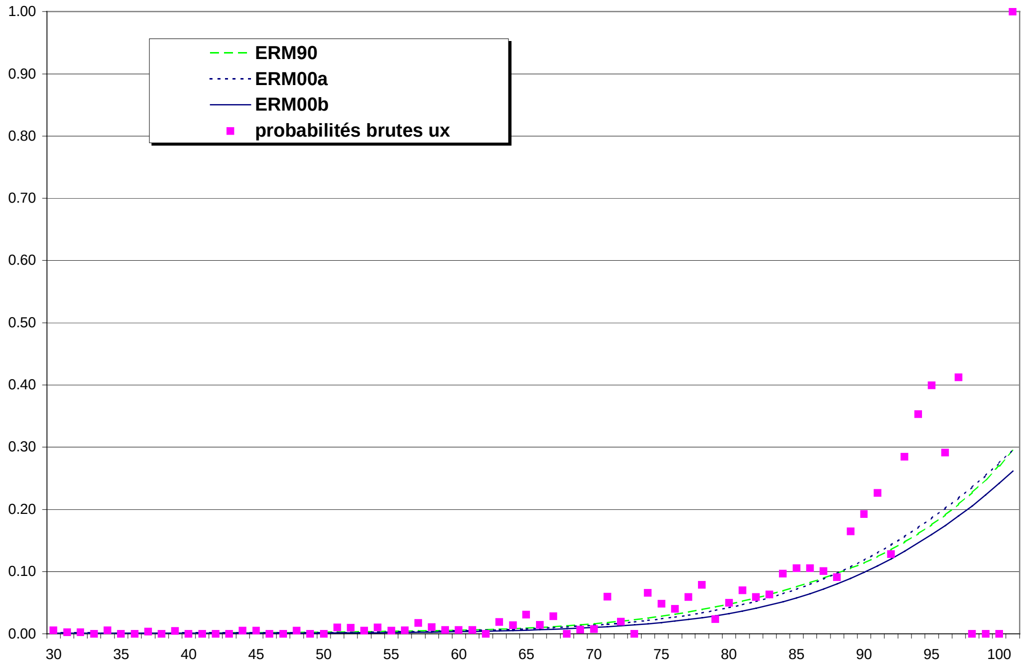
6.4 Comparaisons de la mortalité obtenue avec les tables de premier ordre

6.4.1 ERM/F 1990

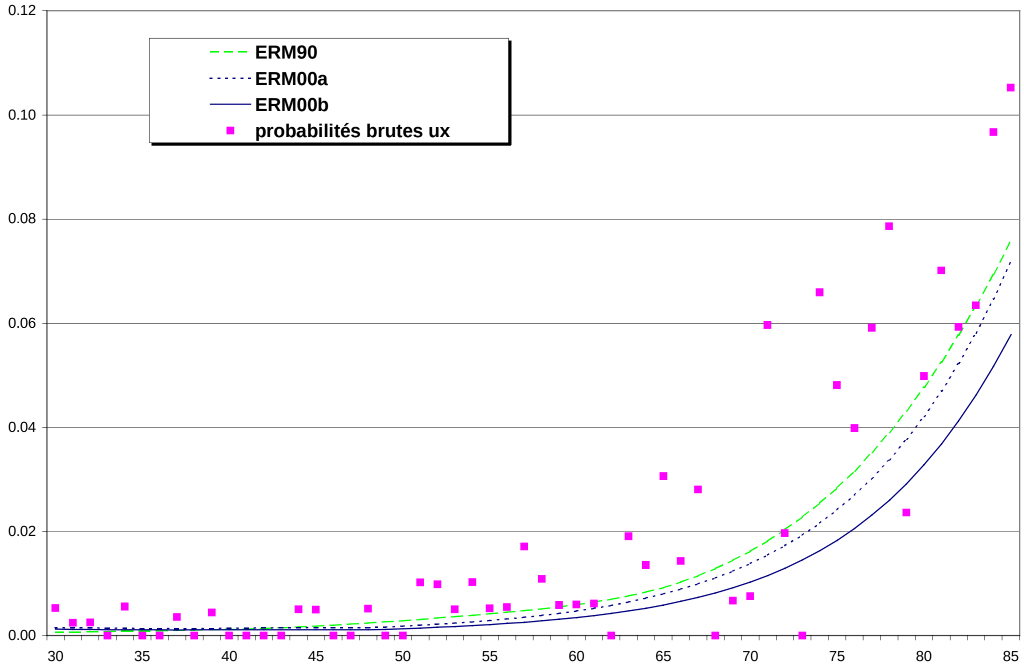
Les probabilités brutes obtenues pour les hommes sont pratiquement toutes au-dessus de la table ou se confondent avec celle-ci. Cette surmortalité se manifeste plus fortement pour les âges élevés. La mortalité effective présente donc encore de la marge par rapport à la mortalité attendue.

Les symboles suivants seront utilisés dans les graphiques ci-dessous:

- ERM90 ou ERF90: Probabilités de décès selon la table ERM/F 1990.
- ERM00a ou ERF00a: Probabilités de décès selon la table de génération ERM/F 2000 valable en l'an 2000.
- ERM00b ou ERF00b: Probabilités de décès selon la table de génération ERM/F 2000 valable en l'an 2010.

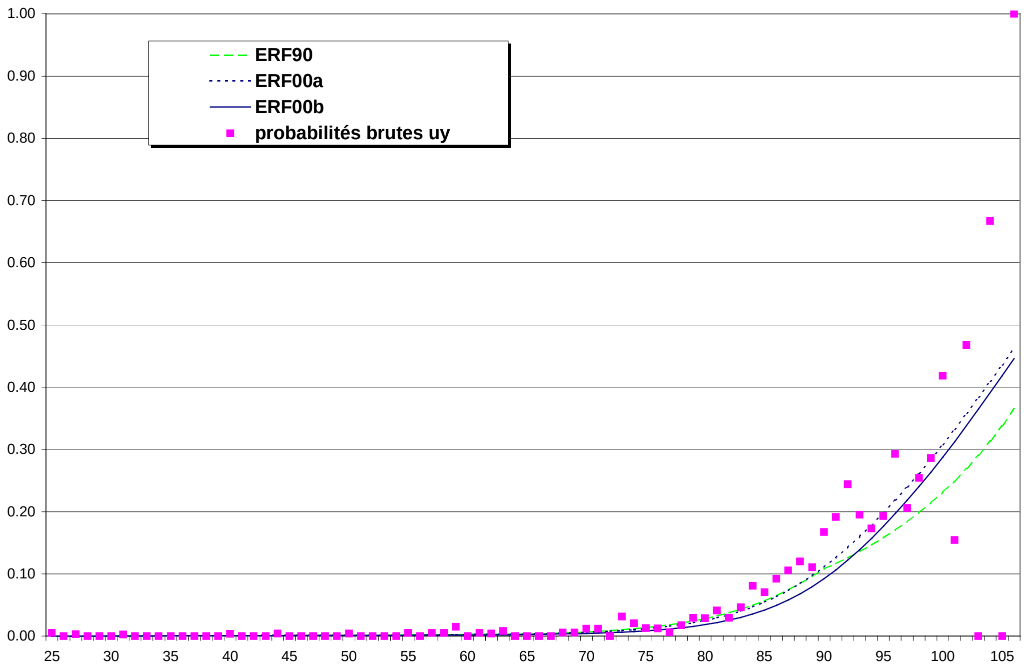


Graphique 1 : mortalité observée des hommes entre 30 et 100 ans par rapport à ERM90 et ERM2000

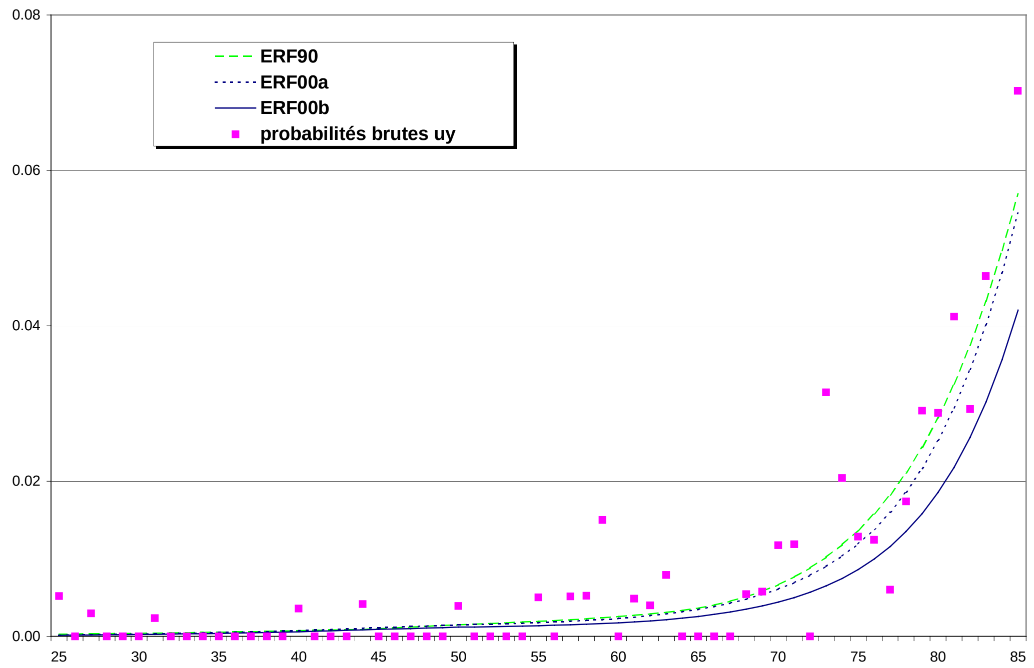


Graphique 2 : mortalité observée des hommes entre 30 et 85 ans par rapport à ERM90 et ERM2000

Pour les femmes, nous constatons que la mortalité observée est plus proche de la table de référence, la surmortalité croissant de manière significative à partir de 100 ans.



Graphique 3 : mortalité observée des femmes entre 25 et 105 ans par rapport à ERF90 et ERF2000



Graphique 4 : mortalité observée des femmes entre 25 et 85 ans par rapport à ERF90 et ERF2000

6.4.2 ERM/F 2000

La table ERM/F 2000, contrairement aux tables ERM/F 1980 et ERM/F 1990 a été construite en prenant en compte l'évolution de la longévité future. Il s'agit en effet d'une table de générations. La détermination de la probabilité de mortalité d'un individu dépend, conformément à la formule ci-dessous, de l'âge *x* atteint l'année *t* où $t_0 = 1993$.

$$q_{x,t} = q_{x,t_0} \cdot \begin{cases} \exp\left(-\lambda_x \cdot (t - t_0)\right), & \text{si } t \geq t_0, \\ 1, & \text{sinon} \end{cases} \tag{138}$$

Pour nos comparaisons, nous utiliserons les valeurs des tables $q_{x,2000}$ et $q_{x,2010}$ obtenues à partir de la formule ci-dessus.

Dans les graphiques 1 et 2 ci-dessus, nous constatons que les probabilités de décès sont plus faibles que la table ERM90 jusqu'à 88 ans pour la table 2000, au-delà ces probabilités sont légèrement plus fortes jusqu'à 101 ans. Pour la table 2010, ce phénomène n'apparaît pas. Nous pouvons constater qu'il y a encore une marge entre la mortalité observée et ces tables.

Les graphiques 3 et 4 ci-dessus mettent également en évidence le même phénomène que nous avons constaté pour les hommes, à savoir une mortalité plus faible que la table ERF90 jusqu'à un âge de 88 ans, puis une mortalité plus forte jusqu'à 113 ans pour la table 2000. La table 2010 nous permet également d'observer une surmortalité entre 93 et 113 ans. Nous pouvons constater que la marge par rapport aux observations entre 90 et

100 ans a diminué et est inexistante pour la table 2000 puisque nous nous situons en dessous.

6.5 Commentaires sur les résultats du lissage

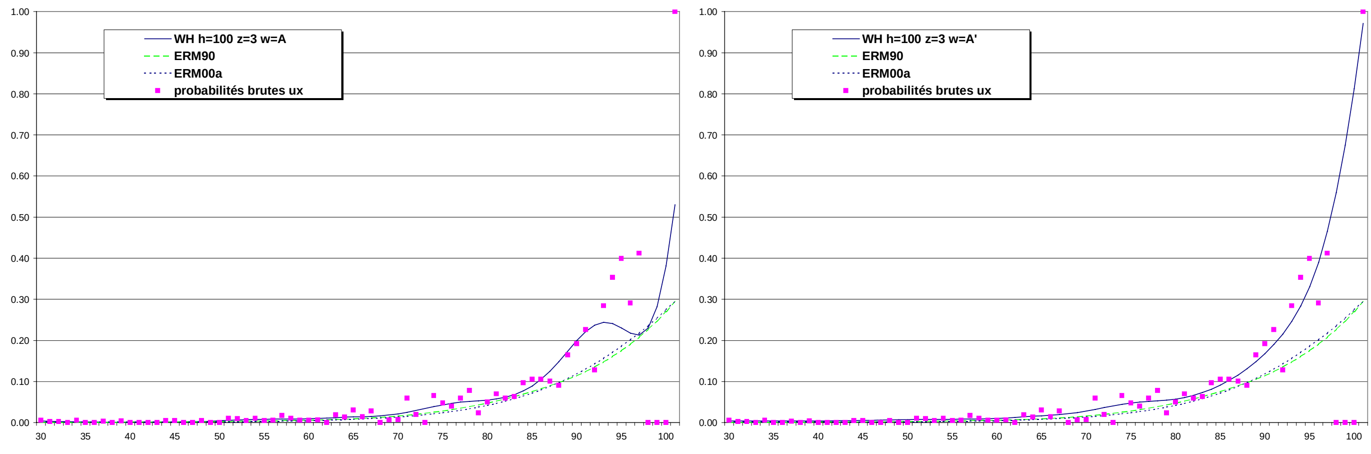
6.5.1 Whittaker-Henderson

Les types des formules utilisées sont les suivants :

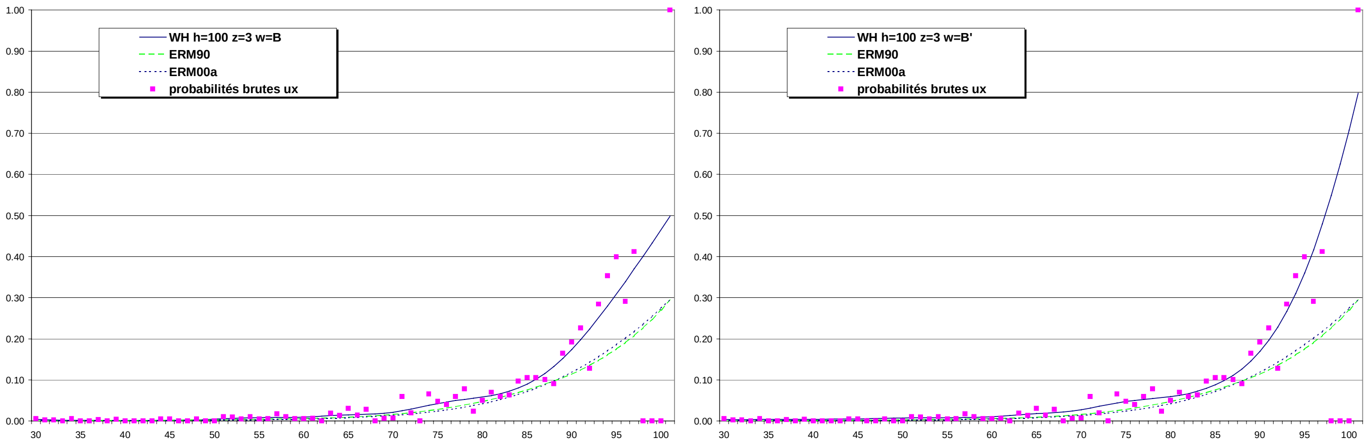
A	Tous les $w_i = 1$.
A’	Tous les $w_i = 1$ sauf si $q_i = 0$ alors $w_i = 0$.
B	Tous les $w_i = N \cdot \sum_{i=1}^{N} \frac{n_i}{\overline{n}}$, de cette façon nous obtenons $\sum_{i=1}^{N} w_i = N$ comme dans **A**.
B’	Tous les $w_i = N \cdot \sum_{i=1}^{N} \frac{n_i}{\overline{n}}$ sauf si $q_i = 0$ alors $w_i = 0$ et $\sum w_i$ est identique à **A'**.

Nous allons voir les incidences des lacunes évoquées précédemment au point 6.3 sur la révision des estimés par la méthode de Whittaker-Henderson.

Pour commencer, nous allons observer l'effet du type de formule utilisée pour la révision des estimés.



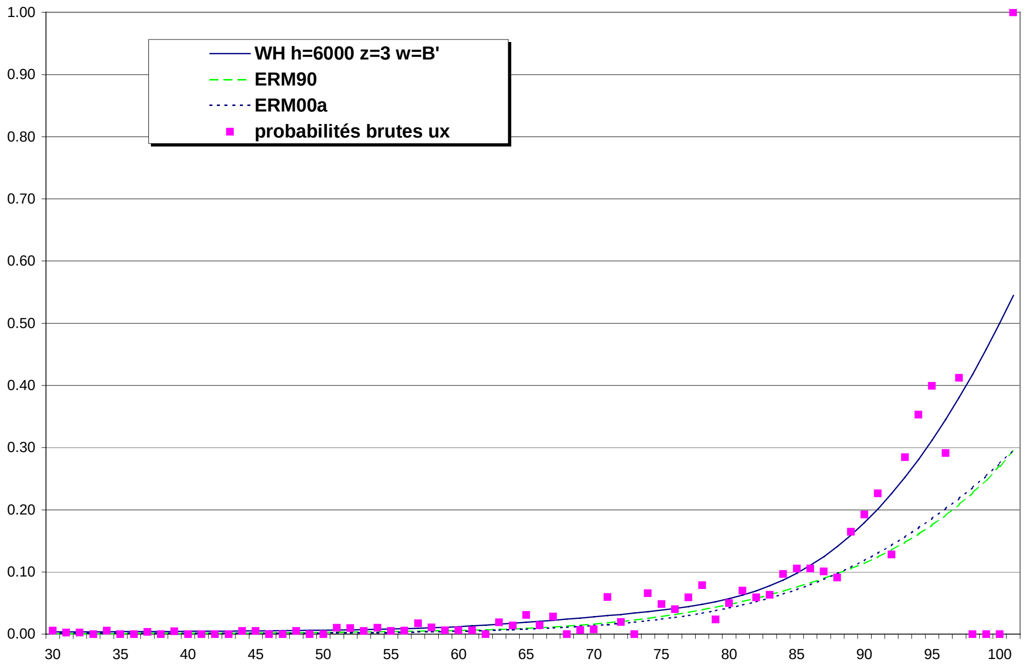
Graphique 5 : W-H, effet du type de formule avec h=100 et z=3, types A et A'



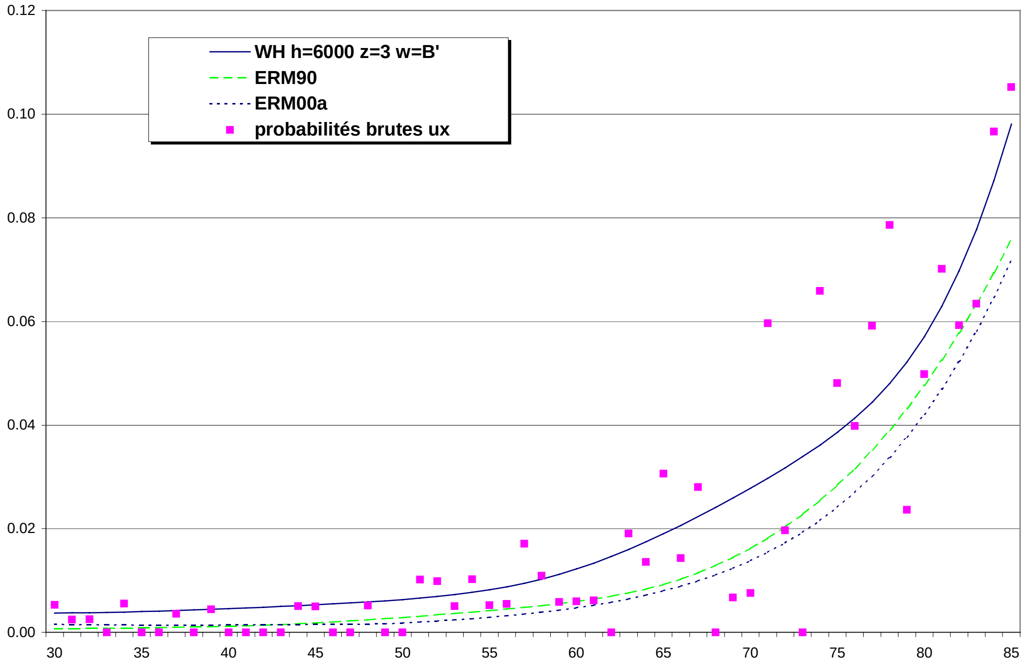
Graphique 6 : W-H, effet du type de formule avec h=100 et z=3, types B et B'

Nous constatons visuellement que les lacunes de probabilités perturbent le lissage de manière assez importante pour les cas A et B, elles "tirent" les courbes vers le bas.

Nous avons effectué le lissage pour plusieurs valeurs de h comprises entre 100 et 100'000. Nous avons retenu une valeur de h=6'000 qui donne un bon compromis pour le lissage aussi bien pour les hommes que les femmes.

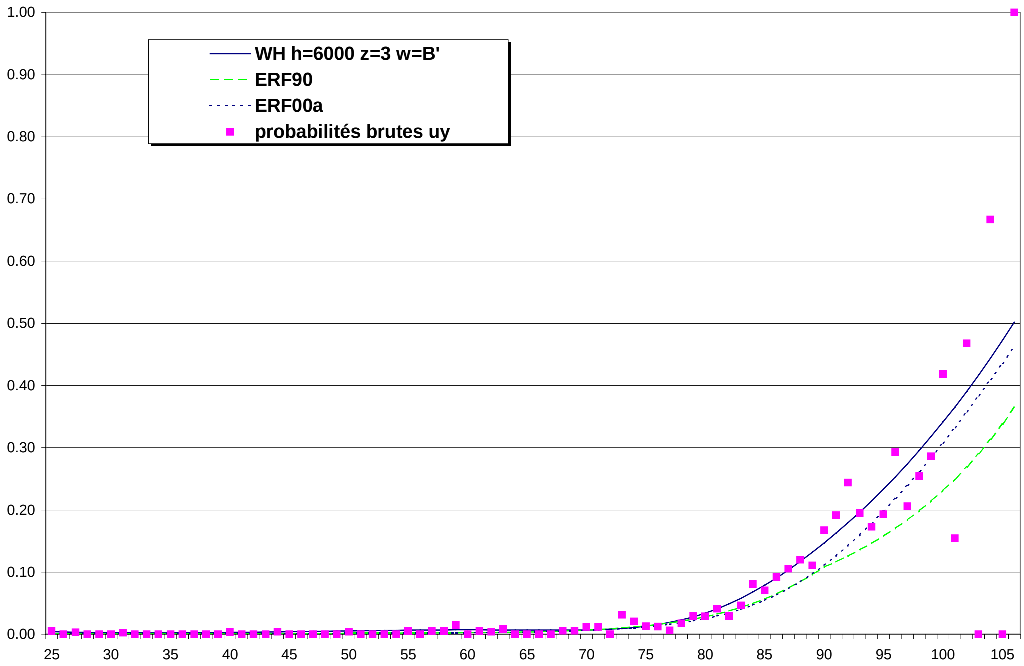


Graphique 7 : W-H, hommes entre 30 et 101 ans, h=6'000 type B'

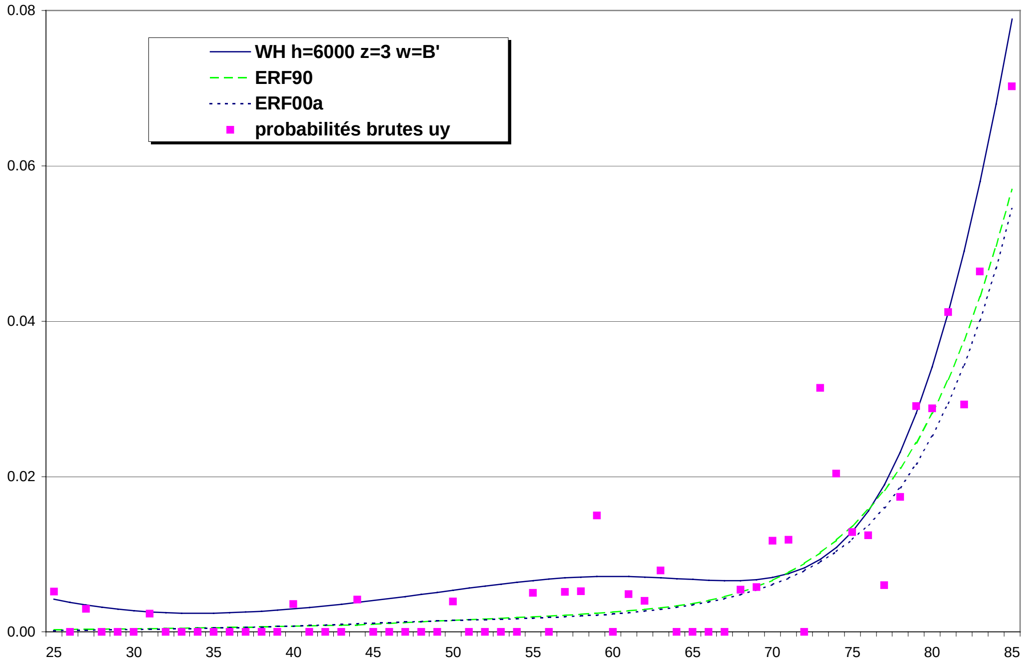


Graphique 8 : W-H, hommes entre 30 et 85 ans, h=6'000 type B'

Nous pouvons constater dans les graphiques ci-dessous que les nombreuses lacunes de probabilités des femmes donnent une courbe dont le lissage est assez ondulé. Le h relativement élevé revient à privilégier fortement la régularité par rapport à la fidélité.



Graphique 9 : W-H, femmes entre 30 et 106 ans, h=6'000 type B'



Graphique 10 : W-H, femmes entre 30 et 85 ans, h=6'000 type B'

Les valeurs détaillées de la révision des estimés par la méthode de Whittaker-Henderson se trouvent à l'annexe E.

Smoothness					Smoothness				
h	**A**	**A'**	**B**	**B'**	**f**	**A**	**A'**	**B**	**B'**
100	0.000812	0.000149	0.000147	0.000107	**100**	0.000347	0.000260	0.000091	0.000178
500	0.000174	0.000113	0.000121	0.000095	**500**	0.000123	0.000190	0.000068	0.000147
1'000	0.000133	0.000106	0.000116	0.000094	**1'000**	0.000090	0.000164	0.000059	0.000136
3'000	0.000115	0.000090	0.000112	0.000087	**3'000**	0.000060	0.000143	0.000045	0.000121
6'000	0.000107	0.000080	0.000108	0.000084	**6'000**	0.000053	0.000131	0.000039	0.000117
10'000	0.000099	0.000090	0.000105	0.000086	**10'000**	0.000051	0.000112	0.000038	0.000114
50'000	0.000060	0.000137	0.000076	0.000082	**50'000**	0.000025	0.000094	0.000023	0.000046
100'000	0.000038	0.000022	0.000054	0.000070	**100'000**	0.000010	0.000044	0.000006	0.000015

Tableau 11 : évolution du critère de régularité en fonction de h et du type

Fit					Fit				
h	**A**	**A'**	**B**	**B'**	**f**	**A**	**A'**	**B**	**B'**
100	0.610044	0.041180	0.015404	0.008775	**100**	0.644702	0.069841	0.014755	0.012214
500	0.725219	0.048442	0.015833	0.009263	**500**	0.684472	0.081768	0.015439	0.014123
1'000	0.752770	0.052499	0.015992	0.009581	**1'000**	0.700045	0.092636	0.015716	0.014826
3'000	0.777309	0.061173	0.016294	0.010433	**3'000**	0.722901	0.117833	0.016125	0.015615
6'000	0.785023	0.069202	0.016556	0.011225	**6'000**	0.735119	0.137074	0.016402	0.015960
10'000	0.788460	0.077078	0.016793	0.011910	**10'000**	0.742576	0.151805	0.016620	0.016171
50'000	0.793283	0.116597	0.017888	0.014424	**50'000**	0.757272	0.196043	0.017346	0.016814
100'000	0.794303	0.141223	0.018636	0.015641	**100'000**	0.760520	0.213998	0.017814	0.017328

Tableau 12 : évolution du critère de fidélité en fonction de h et du type

Différence v_x-u_x				
h	**A**	**A'**	**B**	**B'**
100	0.000000	2.183574	0.694141	1.973472
500	0.000000	2.211212	0.727819	1.675796
1'000	0.000000	2.211188	0.695932	1.508992
3'000	0.000000	2.194939	0.584653	1.201828
6'000	0.000000	2.168703	0.486960	0.994895
10'000	0.000000	2.136350	0.406653	0.843715
50'000	0.000000	1.984191	0.122370	0.408550
100'000	0.000000	1.931747	-0.013361	0.240897

Différence v_y-u_y				
f	**A**	**A'**	**B**	**B'**
100	0.000000	1.478821	0.066486	0.946178
500	0.000000	1.491502	-0.044752	0.610637
1'000	0.000000	1.485925	-0.048200	0.521325
3'000	0.000000	1.465335	0.005692	0.467896
6'000	0.000000	1.453942	0.059164	0.472396
10'000	0.000000	1.446807	0.096231	0.481366
50'000	0.000000	1.415539	0.118308	0.455719
100'000	0.000000	1.425695	0.057056	0.388143

Tableau 13 : évolution de la somme des différences entre les estimés et les observations en fonction de h et du type

6.5.2 Interpolations avec jonctions lisses

6.5.2.1 Formule de Karup-king

Nous avons regroupé les données des décès et de l'exposition par groupes quinquennaux en sommant cinq valeurs centrées sur l'âge.

$$E_x^{(5)} = \sum_{i=x-2}^{x+2} E_i \quad , \quad d_x^{(5)} = \sum_{i=x-2}^{x+2} d_i \tag{139}$$

De ces données groupées, nous obtenons les points-pivots en utilisant la formule de King qui est correcte jusqu'à la cinquième différence.

$$\begin{aligned} u_x^P &= 0.2 \cdot w_x - 0.008 \cdot \left(w_{x-5} - 2 \cdot w_x + w_{x+5}\right) \\ &\quad + 0.000896 \cdot \left(w_{x-10} - 4 \cdot w_{x-5} + 6 \cdot w_x - 4 \cdot w_{x+5} + w_{x+10}\right) \end{aligned} \tag{140}$$

En pratique, nous n'utiliserons que les 2 premiers termes.

$$u_x^P = 0.2 \cdot w_x - 0.008 \cdot \left(w_{x-5} - 2 \cdot w_x + w_{x+5}\right) \tag{141}$$

Un traitement particulier est appliqué aux extrémités.

Si w_a est le premier groupe quinquennal, on obtient

$$u_a^P = 0.2 \cdot w_a - 0.008 \cdot \left(w_a - 2 \cdot w_{a+5} + w_{a+10}\right) \tag{142}$$

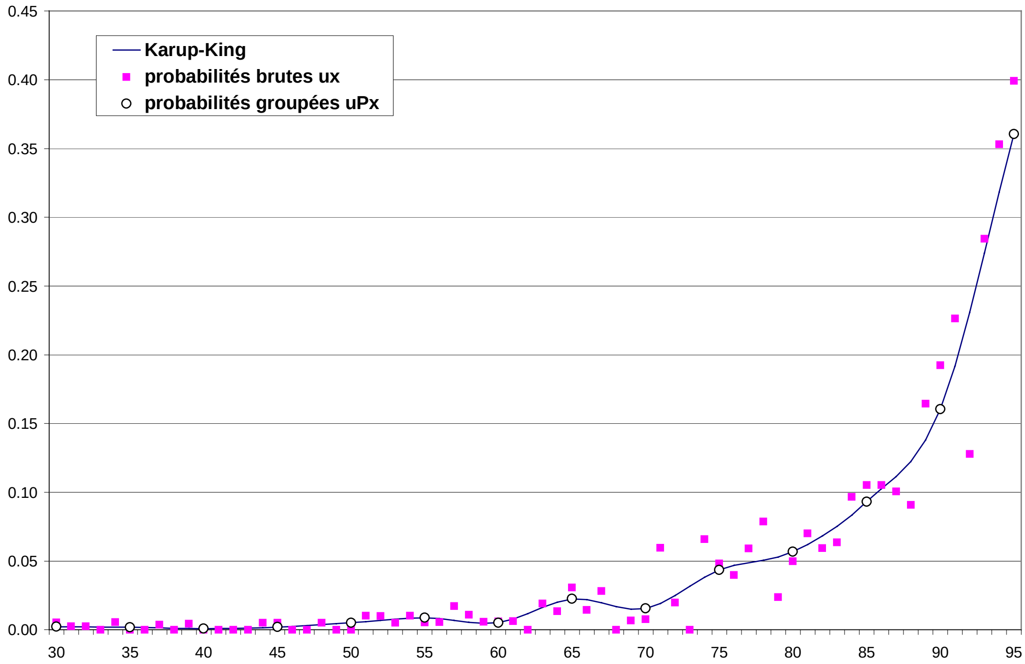
Si w_b est le dernier groupe quinquennal, on a que

$$u_b^P = 0.2 \cdot w_b - 0.008 \cdot \left(w_b - 2 \cdot w_{b-5} + w_{b-10}\right) \tag{143}$$

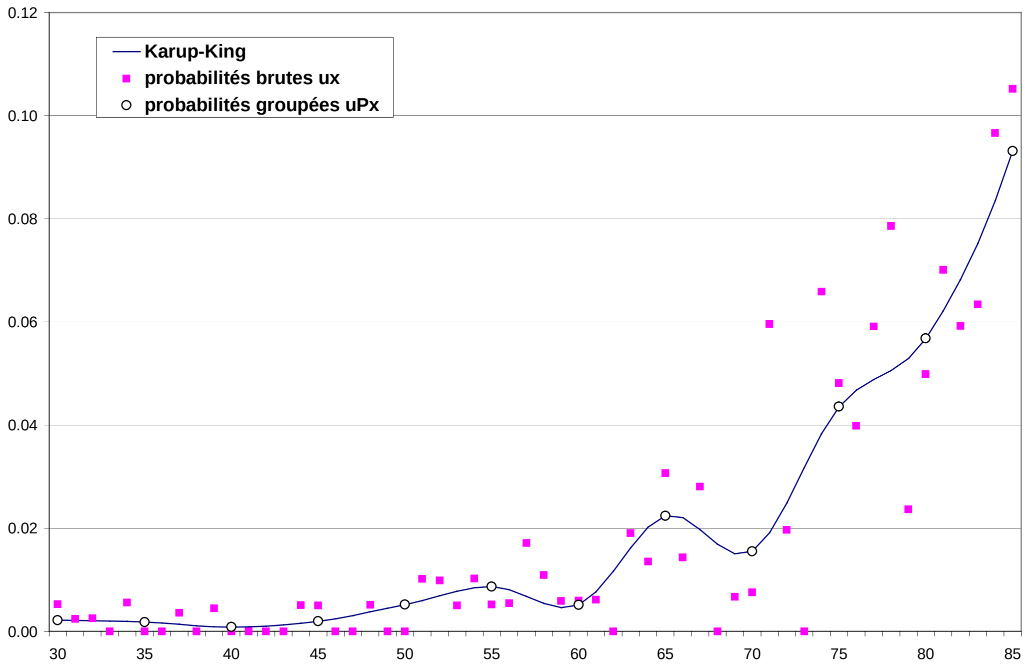
Les probabilités brutes qui figurent aux annexes A et B ont été regroupées en utilisant la formule de King et figurent à l'annexe C.

Nous pouvons constater que l'application de cette formule nous donne un résultat négatif pour le nombre de décès pour les femmes appartenant au groupe quinquennal x=35. Une probabilité négative n'étant pas possible, nous avons posé $u_{35}^P = 0$. Cette valeur est également justifiée par le fait que nous n'observons aucun décès pour les femmes entre 32 et 39 ans.

Nous avons un effet analogue en fin de table avec des expositions négatives, nous avons donc délibérément limité nos calculs d'interpolations jusqu'aux avant-derniers groupes quinquennaux.

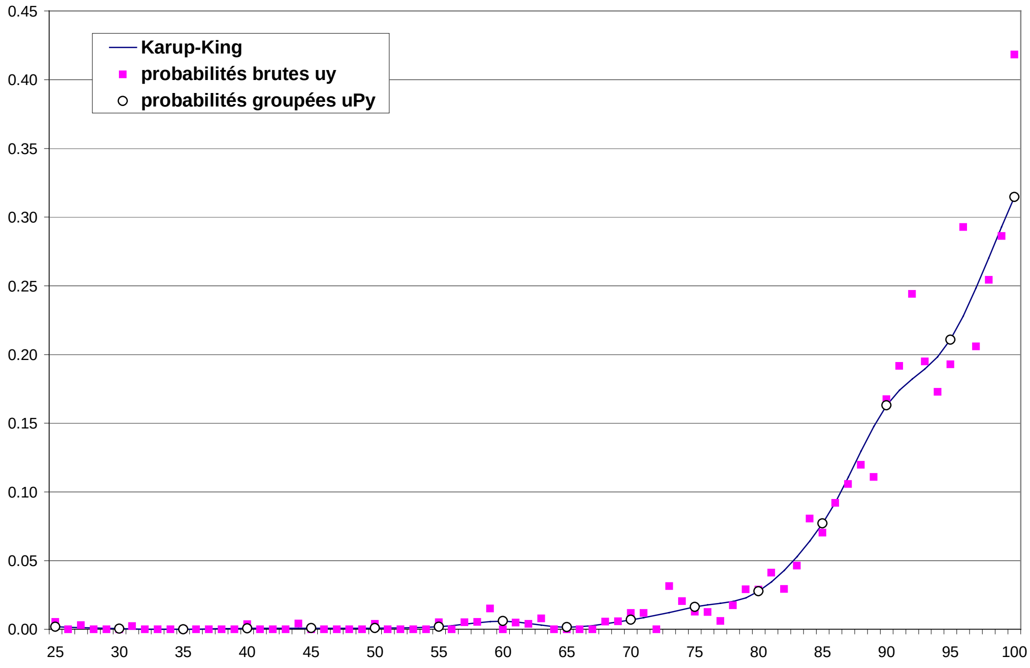


Graphique 11 : Karup-King, groupes quinquennaux, hommes entre 30 et 95 ans

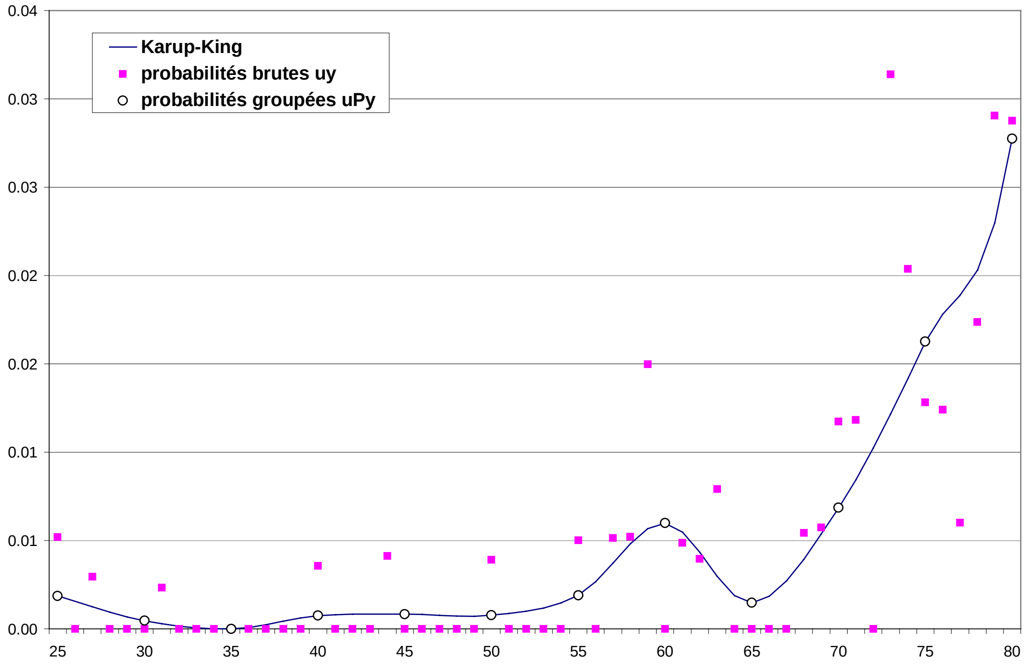


Graphique 12 : Karup-King, groupes quinquennaux, hommes entre 30 et 85 ans

Le lissage produit à partir de 75 ans est d'une régularité suffisante, par contre en dessous de cet âge, nous observons des variations que le regroupement quinquennal n'a pas réussi à éliminer. Nous constatons également que les groupes quinquennaux sont "tirés" vers le bas lorsqu'il existe d'observations.



Graphique 13 : Karup-King, groupes quinquennaux, femmes entre 25 et 100 ans

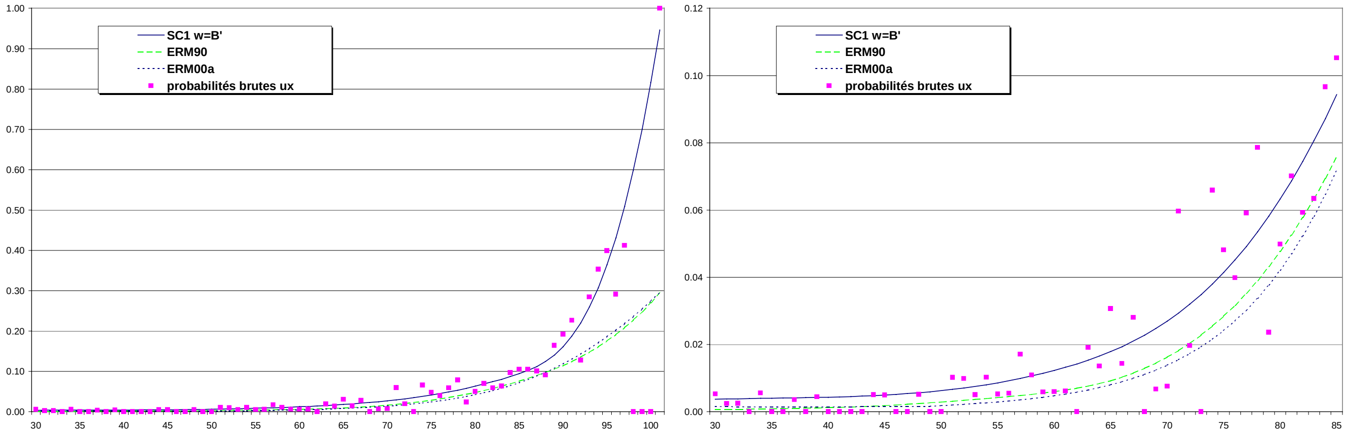


Graphique 14 : Karup-King, groupes quinquennaux, femmes entre 25 et 80 ans

Vu le grand nombre de lacunes jusqu'à 55 ans pour les femmes, les groupes quinquennaux sont très fortement "tirés" vers le bas et nous observons une très grande variabilité jusqu'à 70 ans.

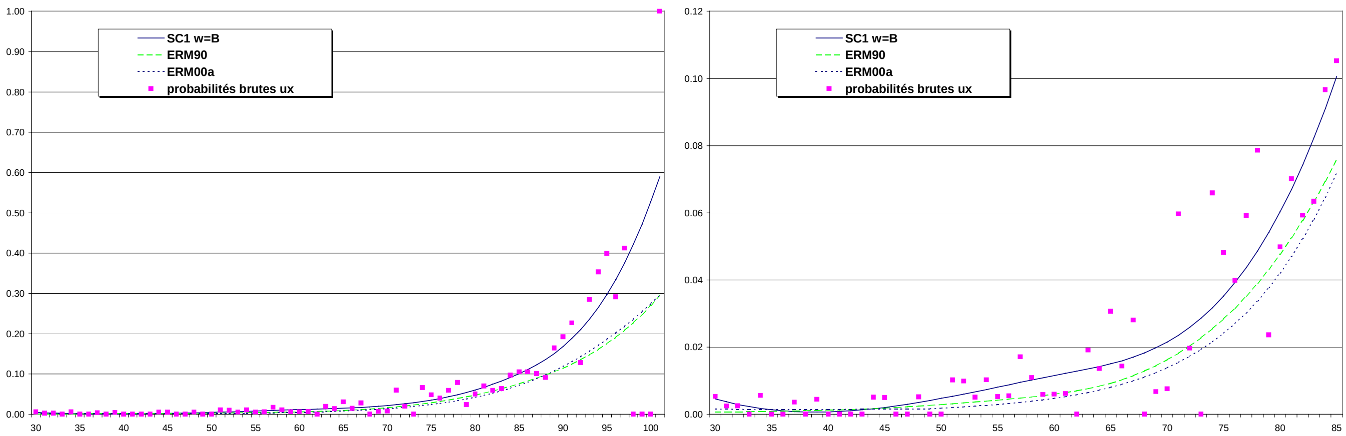
6.5.3 Splines cubiques lissantes

Pour les hommes, nous avons utilisé des nœuds à 60.5 et 84.5 ans. Dans l'exemple ci-dessous, nous avons utilisé des poids nuls lorsqu'aucune probabilité de décès n'a pu être observée.



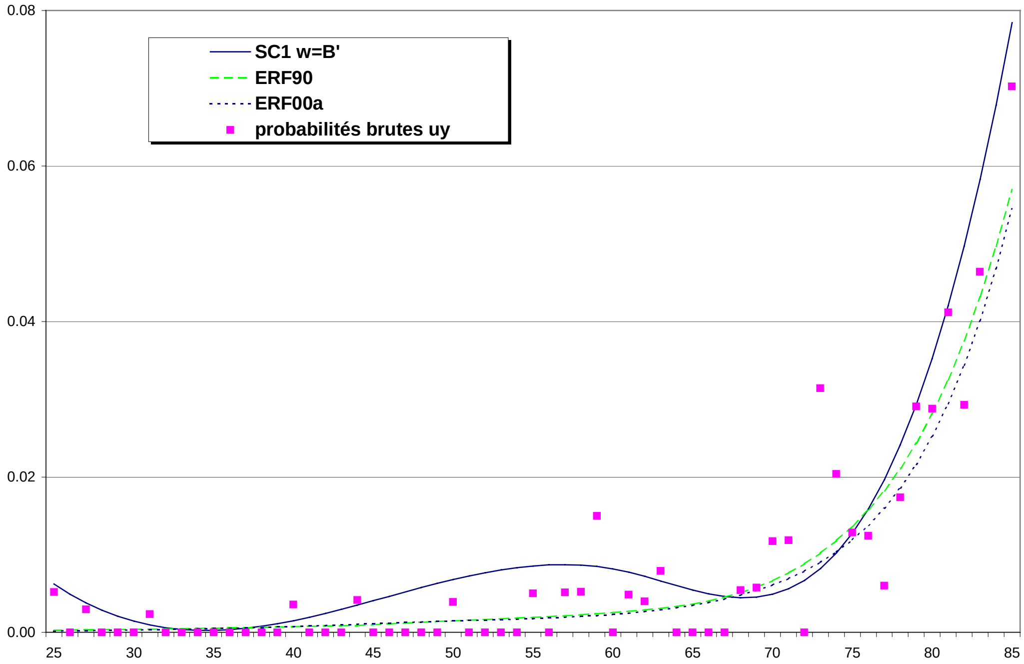
Graphique 15 : Splines MCP, nœuds à 60.5 et 84.5, hommes

Dans le graphique ci-dessous, nous pouvons constater l'effet des lacunes d'observations sur le lissage produit lorsque nous n'affectons pas de poids nuls à celles-ci. Le lissage produit dans le graphique 15 nous semble plus adéquat.



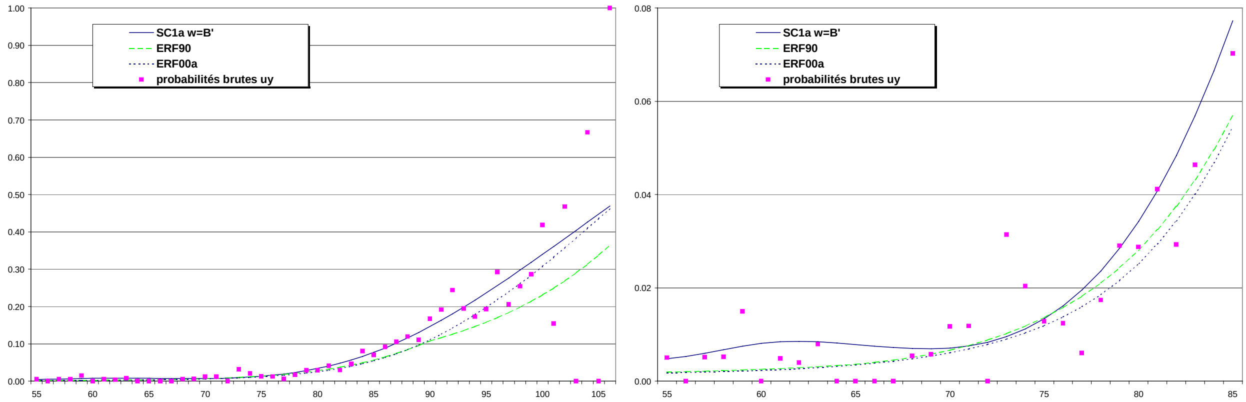
Graphique 16 : Splines MCP, nœuds à 60.5 et 84.5, hommes de 30 à 101 ans

En utilisant les mêmes nœuds pour les femmes, les très nombreuses lacunes dans la tranche d'âge de 25 à 55 ans nous donnent la courbe qui est représentée dans le graphique ci-dessous. Nous pouvons immédiatement constater que l'intervalle de 25 à 55 ans va nous poser des problèmes.



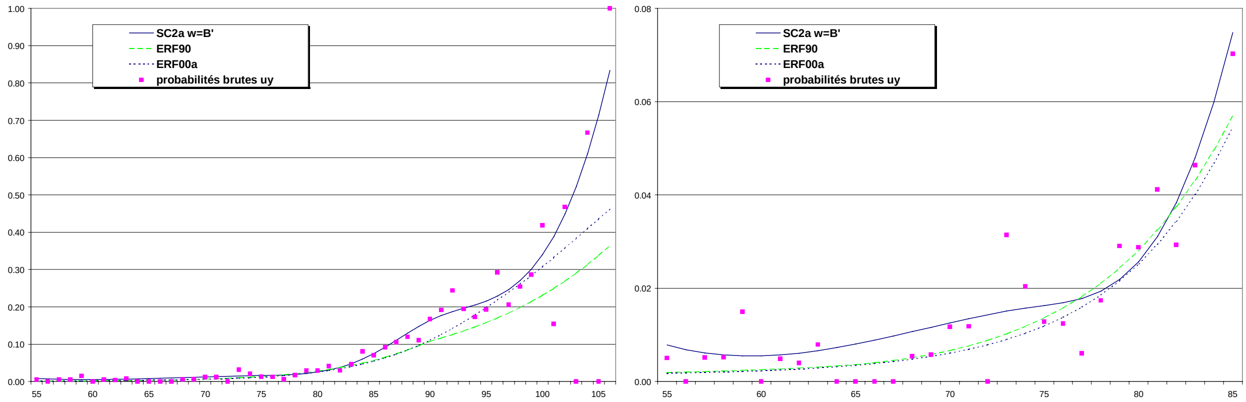
Graphique 17 : Splines MCP, nœuds à 60.5 et 84.5, femmes de 25 à 85 ans

Vu les très nombreuses lacunes, nous n'avons utilisé pour la révision des estimés que les observations de 55 à 106 ans pour cette méthode. Nous voyons l'effet de cette restriction de l'intervalle des valeurs observées dans la représentation graphique ci-dessous.



Graphique 18 : Splines MCP, nœuds à 60.5 et 84.5, femmes

Le lissage produit pour des âges élevés ne nous satifaisant pas entièrement, nous avons utilisé les trois nœuds 74.5, 84.5 et 90.5. Nous constatons que les âges élevés sont plus représentatifs par rapport aux observations, mais la bosse à 60 ans est devenue un creux.



Graphique 19 : Splines MCP avec poids nuls, nœuds à 74.5, 84.5 et 90.5, femmes

Nous avons fait des essais avec différents jeux de nœuds qui donnaient parfois des résultats très surprenants.

7 Conclusion

Nous avons pu constater que l'effectif relativement faible des assurés conduit rapidement à des variations de taux de mortalité importantes et se traduit ensuite par des difficultés pour obtenir un lissage régulier.

Cette faiblesse de l'effectif donne également lieu à de nombreuses lacunes d'observations de décès, un effectif plus important comblerait d'une part celles-ci et permettrait d'obtenir une meilleure régularité.

Nous avons pu constater que le choix de la méthode de "graduation" ainsi que des paramètres influent sur la régularité et la qualité du lissage produit. Il est nécessaire d'apprécier le lissage produit et changer les paramètres ou la méthode jusqu'à obtenir la courbe qui nous paraît la plus adéquate. On ne peut donc pas appliquer une méthode telle quelle, cette étape se rapproche plus d'un art car le jugement humain est très important.

Nous avons mis en œuvre un programme permettant d'effectuer des extractions de données dans le système d'informations des Rentes Genevoises, afin de pouvoir déterminer périodiquement les taux de mortalité bruts, cette phase étant moins sensible au jugement humain.

Nous n'avons pas abordé une méthode de "graduation" qui utilise une table de référence et qui peut se révéler intéressante pour de faibles effectifs. Cette méthode permet d'obtenir une courbe ayant la même allure que la table de référence, mais les taux de mortalités dépendent des observations. Il est quand même nécessaire que la table de référence sélectionnée épouse l'idée que l'on se fait a priori sur l'allure de la courbe.

8 Références

- Benjamin, B. et Pollard, J. H. (1993) *The analysis of mortality and other actuarial statistics*
- Gerber, H. U. (1995) Life insurance mathematics, Springer p. 109-118
- London, D. (1985) *Graduation: the revision of estimates*, Actex publications
- London, D. (1997) *Survival models and their estimation*, Actex publications

9 Annexes

A Probabilités brutes, hommes, observations du 01/01/90 au 31/12/99

Age	u_x	# décès	# sorties	# indiv-an	Exposition
0				8	4.46
1				15	8.28
2				13	8.83
3				15	7.76
4				15	8.83
5				12	8.21
6				14	7.61
7				11	5.97
8				16	10.69
9				19	11.24
10				15	8.20
11				12	9.64
12				12	7.46
13				10	6.59
14				10	5.78
15				10	7.43
16				10	7.64
17			1	9	7.27
18				4	3.33
19				4	2.63
20				11	4.21
21				29	14.41
22			3	65	43.64
23			2	98	75.05
24			4	135	108.11
25			10	194	145.77
26			13	263	206.02
27			25	332	258.88
28			30	389	295.83
29			61	460	350.10
30	0.005 293	2	56	508	377.85
31	0.002 426	1	56	528	412.21
32	0.002 511	1	65	519	398.21
33			54	513	396.44
34	0.005 545	2	43	477	360.70
35			52	435	326.06
36			49	394	304.29
37	0.003 580	1	30	362	279.31
38			28	328	261.56
39	0.004 434	1	24	306	225.52
40			29	274	216.22
41			16	255	194.74
42			16	248	188.56
43			18	253	195.83
44	0.005 079	1	10	258	196.90
45	0.004 985	1	19	253	200.62
46			14	260	196.87
47			16	256	191.46
48	0.005 146	1	15	241	194.31
49			11	238	186.57
50			19	235	184.31
51	0.010 178	2	12	246	196.49
52	0.009 846	2	13	261	203.12
53	0.005 023	1	9	253	199.07
54	0.010 245	2	10	246	195.23

Age	u_x	# décès	# sorties	# indiv-an	Exposition
55	0.005 208	1	7	237	192.01
56	0.005 451	1	13	230	183.46
57	0.017 117	3	6	225	175.26
58	0.010 907	2	12	222	183.37
59	0.005 869	1	23	217	170.38
60	0.005 961	1	14	214	167.76
61	0.006 156	1	4	197	162.43
62			3	196	157.27
63	0.019 070	3	9	183	157.31
64	0.013 556	2	7	192	147.54
65	0.030 642	7	10	285	228.44
66	0.014 321	3	1	251	209.49
67	0.028 037	5		205	178.34
68				184	160.84
69	0.006 704	1		178	149.16
70	0.007 578	1	1	163	131.96
71	0.059 621	7	1	134	117.41
72	0.019 674	2		115	101.66
73				116	103.01
74	0.065 910	7		125	106.20
75	0.048 110	5		122	103.93
76	0.039 843	4		115	100.39
77	0.059 142	6		114	101.45
78	0.078 596	8		115	101.79
79	0.023 620	2		105	84.68
80	0.049 835	4		93	80.27
81	0.070 122	6		100	85.57
82	0.059 260	5		95	84.37
83	0.063 423	5		90	78.84
84	0.096 659	7		80	72.42
85	0.105 230	7		77	66.52
86	0.105 170	7		75	66.56
87	0.100 675	6		67	59.60
88	0.090 891	5		66	55.01
89	0.164 340	9		58	54.76
90	0.192 274	9		51	46.81
91	0.226 440	9		43	39.75
92	0.127 835	4		32	31.29
93	0.284 437	7		28	24.61
94	0.352 941	6		17	17.00
95	0.399 237	4		11	10.02
96	0.291 301	1		5	3.43
97	0.412 376	1		3	2.42
98				1	1.00
99				1	1.00
100				1	1.00
101	1.000 000	1		1	1.00

B Probabilités brutes, femmes, observations du 01/01/90 au 31/12/99

Age	u_y	# décès	# sorties	# indiv-an	Exposition
0				9	4.62
1				12	6.16
2				14	9.25
3				14	11.33
4			1	18	11.08
5				17	13.32
6				18	11.03
7				14	8.41
8				13	8.04
9				9	7.24
10				15	6.92
11				12	8.10
12				8	6.34
13				10	6.61
14				9	6.21
15				6	4.95
16				4	2.62
17				3	2.39
18				5	2.87
19				11	5.75
20			1	17	10.93
21			2	38	21.11
22			4	72	49.82
23			13	120	80.35
24			6	170	130.70
25	0.005 190	1	15	261	192.68
26			15	345	274.40
27	0.002 955	1	37	438	338.46
28			35	500	386.73
29			46	560	423.91
30			66	580	436.10
31	0.002 325	1	44	552	430.18
32			53	556	437.76
33			42	532	422.59
34			51	500	384.86
35			47	472	355.35
36			33	420	333.01
37			33	412	316.03
38			40	393	296.65
39			24	371	282.83
40	0.003 563	1	26	366	280.66
41			19	348	276.62
42			35	338	261.81
43			17	318	242.55
44	0.004 129	1	19	326	242.20
45			23	330	249.13
46			24	309	241.66
47			17	292	233.36
48			12	283	228.99
49			15	314	245.96
50	0.003 902	1	13	329	256.30
51			17	316	248.94
52			7	293	240.85
53			15	289	214.35
54			6	264	201.22

Age	u_y	# décès	# sorties	# indiv-an	Exposition
55	0.005 013	1	15	260	199.46
56			15	257	188.93
57	0.005 141	1	7	254	194.50
58	0.005 204	1	8	243	192.15
59	0.014 971	3	5	246	200.39
60			5	249	206.34
61	0.004 868	1	4	257	205.41
62	0.003 971	1	15	317	251.80
63	0.007 915	2	3	305	252.69
64			1	286	239.59
65			1	273	222.63
66			1	264	213.72
67				237	196.71
68	0.005 424	1	1	216	184.35
69	0.005 747	1	1	209	174.00
70	0.011 733	2		201	170.45
71	0.011 827	2		187	169.10
72				193	166.54
73	0.031 388	5		186	159.30
74	0.020 375	3		170	147.24
75	0.012 822	2		181	155.98
76	0.012 410	2		195	161.15
77	0.006 001	1		197	166.63
78	0.017 368	3		200	172.73
79	0.029 050	5		199	172.12
80	0.028 761	5		207	173.85
81	0.041 143	8		215	194.44
82	0.029 285	6		237	204.89
83	0.046 376	10		230	215.63
84	0.080 553	17		230	211.04
85	0.070 250	14		218	199.29
86	0.091 874	17		210	185.04
87	0.105 612	20		221	189.37
88	0.119 565	23		208	192.36
89	0.110 717	19		192	171.61
90	0.167 299	28		188	167.36
91	0.191 567	28		160	146.16
92	0.243 979	30		136	122.96
93	0.194 901	20		110	102.62
94	0.172 653	15		94	86.88
95	0.192 839	15		83	77.79
96	0.292 691	19		69	64.91
97	0.205 860	9		47	43.72
98	0.254 336	9		39	35.39
99	0.286 210	7		27	24.46
100	0.418 235	7		18	16.74
101	0.154 465	1		7	6.47
102	0.467 649	2		5	4.28
103				4	2.95
104	0.666 667	2		3	3.00
105				1	1.00
106	1.000 000	1		1	1.00

C Probabilités groupées quinquennales

Age	$E^{(5)}_x$	$D^{(5)}_x$	E^P_x	D^P_x	u^P_x
25					
30	1834.21	4	370.14	0.81	0.002 183
35	1666.81	3	336.66	0.61	0.001 806
40	1086.60	1	213.52	0.18	0.000 824
45	981.69	2	195.63	0.38	0.001 963
50	964.81	5	192.99	1.00	0.005 182
55	945.02	8	189.68	1.65	0.008 688
60	841.22	5	166.77	0.86	0.005 133
65	921.12	20	186.94	4.19	0.022 424
70	661.03	11	131.29	2.04	0.015 538
75	514.99	22	102.46	4.46	0.043 570
80	436.67	25	87.45	4.97	0.056 811
85	343.93	32	68.97	6.42	0.093 135
90	227.62	36	45.95	7.37	0.160 332
95	57.49	19	10.56	3.81	0.360 462
100	4.00	1	négatif	0.21	négatif
105					

Age	$E^{(5)}_y$	$D^{(5)}_y$	E^P_y	D^P_y	u^P_y
25	1016.59	2	214.52	0.40	0.001 865
30	2114.69	1	434.15	0.20	0.000 461
35	1811.84		363.25	négatif	négatif
40	1398.57	1	277.93	0.21	0.000 748
45	1208.90	1	240.17	0.20	0.000 833
50	1221.05	1	246.09	0.19	0.000 780
55	998.47	2	197.45	0.38	0.001 904
60	1056.08	6	211.12	1.26	0.005 987
65	1125.33	2	227.71	0.34	0.001 476
70	864.45	6	171.40	1.18	0.006 861
75	790.30	13	156.45	2.54	0.016 261
80	918.02	27	183.97	5.10	0.027 744
85	1000.37	78	202.33	15.61	0.077 141
90	800.46	128	161.89	26.40	0.163 074
95	375.91	78	74.10	15.62	0.210 756
100	87.33	26	15.79	4.97	0.314 577
105	7.95	3	négatif	0.37	négatif

D Interpolations à jonctions lisses, formule de Karup-King

Age	u^P_x	v_x
25		
26		
27		
28		
29		
30	0.002 183	0.002 183
31		0.002 117
32		0.002 061
33		0.002 000
34		0.001 920
35	0.001 806	0.001 806
36		0.001 614
37		0.001 355
38		0.001 093
39		0.000 895
40	0.000 824	0.000 824
41		0.000 883
42		0.001 027
43		0.001 256
44		0.001 568
45	0.001 963	0.001 963
46		0.002 469
47		0.003 087
48		0.003 774
49		0.004 486
50	0.005 182	0.005 182
51		0.005 978
52		0.006 903
53		0.007 780
54		0.008 435
55	0.008 688	0.008 688
56		0.008 096
57		0.006 774
58		0.005 393
59		0.004 623
60	0.005 133	0.005 133
61		0.007 644
62		0.011 709

Age	u^P_x	v_x
63		0.000 000
64		0.000 000
65	0.000 000	0.000 000
66		0.000 000
67		0.000 000
68		0.002 000
69		0.001 920
70	0.002 183	0.002 183
71		0.002 117
72		0.002 061
73		0.001 093
74		0.000 895
75	0.001 806	0.001 806
76		0.001 614
77		0.001 355
78		0.001 256
79		0.001 568
80	0.000 824	0.000 824
81		0.000 883
82		0.001 027
83		0.003 774
84		0.004 486
85	0.001 963	0.001 963
86		0.002 469
87		0.003 087
88		0.007 780
89		0.008 435
90	0.005 182	0.005 182
91		0.005 978
92		0.006 903
93		0.005 393
94		0.004 623
95	0.008 688	0.008 688

Age	u^P_y	v_y
25	0.001 865	0.001 865
26		0.001 569
27		0.001 258
28		0.000 954
29		0.000 681
30	0.000 461	0.000 461
31		0.000 289
32		0.000 150
33		0.000 052
34		0.000 000
35	0.000 000	0.000 000
36		0.000 083
37		0.000 244
38		0.000 439
39		0.000 622
40	0.000 748	0.000 748
41		0.000 810
42		0.000 837
43		0.000 841
44		0.000 835
45	0.000 833	0.000 833
46		0.000 812
47		0.000 765
48		0.000 723
49		0.000 718
50	0.000 780	0.000 780
51		0.000 882
52		0.001 003
53		0.001 185
54		0.001 471
55	0.001 904	0.001 904
56		0.002 669
57		0.003 737
58		0.004 831
59		0.005 673
60	0.005 987	0.005 987
61		0.005 476
62		0.004 326

Age	u^P_y	v_y
63		0.000 954
64		0.000 681
65	0.001 865	0.001 865
66		0.001 569
67		0.001 258
68		0.000 052
69		0.000 000
70	0.000 461	0.000 461
71		0.000 289
72		0.000 150
73		0.000 439
74		0.000 622
75	0.000 000	0.000 000
76		0.000 083
77		0.000 244
78		0.000 841
79		0.000 835
80	0.000 748	0.000 748
81		0.000 810
82		0.000 837
83		0.000 723
84		0.000 718
85	0.000 833	0.000 833
86		0.000 812
87		0.000 765
88		0.001 185
89		0.001 471
90	0.000 780	0.000 780
91		0.000 882
92		0.001 003
93		0.004 831
94		0.005 673
95	0.001 904	0.001 904
96		0.002 669
97		0.003 737
98		0.002 980
99		0.001 882
100	0.005 987	0.005 987

E Révision des estimés Whittaker-Henderson, h=6'000, z=3, type B'

Age	u_x	v_x	u_y	v_y
25			0.005 190	0.004 210
26				0.003 811
27			0.002 955	0.003 463
28				0.003 166
29				0.002 920
30	0.005 293	0.003 738		0.002 722
31	0.002 426	0.003 761	0.002 325	0.002 573
32	0.002 511	0.003 801		0.002 470
33		0.003 856		0.002 410
34	0.005 545	0.003 926		0.002 391
35		0.004 007		0.002 411
36		0.004 100		0.002 465
37	0.003 580	0.004 203		0.002 551
38		0.004 315		0.002 666
39	0.004 434	0.004 436		0.002 806
40		0.004 565	0.003 563	0.002 968
41		0.004 701		0.003 151
42		0.004 844		0.003 350
43		0.004 994		0.003 566
44	0.005 079	0.005 151	0.004 129	0.003 795
45	0.004 985	0.005 316		0.004 037
46		0.005 491		0.004 289
47		0.005 677		0.004 550
48	0.005 146	0.005 878		0.004 819
49		0.006 098		0.005 092
50		0.006 341	0.003 902	0.005 367
51	0.010 178	0.006 615		0.005 641
52	0.009 846	0.006 927		0.005 908
53	0.005 023	0.007 289		0.006 163
54	0.010 245	0.007 713		0.006 401
55	0.005 208	0.008 211	0.005 013	0.006 616
56	0.005 451	0.008 798		0.006 801
57	0.017 117	0.009 484	0.005 141	0.006 951
58	0.010 907	0.010 279	0.005 204	0.007 061
59	0.005 869	0.011 190	0.014 971	0.007 127
60	0.005 961	0.012 223		0.007 147
61	0.006 156	0.013 377	0.004 868	0.007 124
62		0.014 648	0.003 971	0.007 063
63	0.019 070	0.016 024	0.007 915	0.006 972
64	0.013 556	0.017 495		0.006 861
65	0.030 642	0.019 049		0.006 747

Age	u_x	v_x	u_y	v_y
66	0.014 321	0.020 674		0.006 649
67	0.028 037	0.022 364		0.006 591
68		0.024 115	0.005 424	0.006 605
69	0.006 704	0.025 926	0.005 747	0.006 729
70	0.007 578	0.027 801	0.011 733	0.007 007
71	0.059 621	0.029 746	0.011 827	0.007 493
72	0.019 674	0.031 770		0.008 251
73		0.033 897	0.031 388	0.009 353
74	0.065 910	0.036 164	0.020 375	0.010 881
75	0.048 110	0.038 623	0.012 822	0.012 926
76	0.039 843	0.041 349	0.012 410	0.015 583
77	0.059 142	0.044 435	0.006 001	0.018 947
78	0.078 596	0.047 994	0.017 368	0.023 105
79	0.023 620	0.052 155	0.029 050	0.028 135
80	0.049 835	0.057 061	0.028 761	0.034 099
81	0.070 122	0.062 864	0.041 143	0.041 043
82	0.059 260	0.069 716	0.029 285	0.048 995
83	0.063 423	0.077 772	0.046 376	0.057 966
84	0.096 659	0.087 182	0.080 553	0.067 944
85	0.105 230	0.098 087	0.070 250	0.078 903
86	0.105 170	0.110 622	0.091 874	0.090 804
87	0.100 675	0.124 910	0.105 612	0.103 603
88	0.090 891	0.141 060	0.119 565	0.117 249
89	0.164 340	0.159 163	0.110 717	0.131 696
90	0.192 274	0.179 295	0.167 299	0.146 902
91	0.226 440	0.201 515	0.191 567	0.162 835
92	0.127 835	0.225 870	0.243 979	0.179 478
93	0.284 437	0.252 394	0.194 901	0.196 837
94	0.352 941	0.281 112	0.172 653	0.214 939
95	0.399 237	0.312 040	0.192 839	0.233 821
96	0.291 301	0.345 188	0.292 691	0.253 523
97	0.412 376	0.380 566	0.205 860	0.274 086
98		0.418 178	0.254 336	0.295 546
99		0.458 029	0.286 210	0.317 937
100		0.500 118	0.418 235	0.341 280
101	1.000 000	0.544 448	0.154 465	0.365 593
102			0.467 649	0.390 889
103				0.417 174
104			0.666 667	0.444 453
105				0.472 727
106			1.000 000	0.501 997

F Conversion du format de la date

```
' Conversion d'une date enregistrée dans l'AS/400 en format date Access
'
Function DateFromASv(vDate) As Variant
Dim dDate As Long, iJour As Integer, iMois As Integer, iAn As Integer

  If IsNull(vDate) Then
    DateFromASv = Null
    Exit Function
  End If
  dDate = CLng(vDate)
  If dDate = 0 Then
    DateFromASv = Null
    Exit Function
  End If
  iAn = vDate \ 10000
  iMois = (vDate \ 100) Mod 100
  iJour = vDate Mod 100
  DateFromASv = DateSerial(iAn, iMois, iJour)
End Function
```

```
' Même chose que DateFromASv mais la valeur Null est remplacée par 0
'
Function DateFromAS(vDate) As Date
Dim vRes

  vRes = DateFromASv(vDate)
  If IsNull(vRes) Then
    DateFromAS = 0
  Else
    DateFromAS = vRes
  End If
End Function
```

```
' Conversion en date numérique 0 de la valeur Null
'
Function Zdate(vDate) As Date
  If IsNull(vDate) Then
    Zdate = 0
  Else
    Zdate = vDate
  End If
End Function
```

```
' Conversion en Null d'une date numérique 0
'
Function Ndate(vDate) As Variant
  If IsNull(vDate) Then
    Ndate = Null
  ElseIf vDate = 0 Then
    Ndate = Null
  Else
    Ndate = vDate
  End If
End Function
```

G Fonctions de calculs sur les âges

```
' Calcul de l'âge exact entre 2 dates
'
Function Age(dDeb As Date, _
            dFin As Date) _
            As Double
Dim d1 As Double, d2 As Double
a
  d1 = Year(dDeb) + (dDeb - DateSerial(Year(dDeb), 1, 1)) / _
    (DateSerial(Year(dDeb) + 1, 1, 1) - DateSerial(Year(dDeb), 1, 1))
  d2 = Year(dFin) + (dFin - DateSerial(Year(dFin), 1, 1)) / _
    (DateSerial(Year(dFin) + 1, 1, 1) - DateSerial(Year(dFin), 1, 1))
  Age = d2 - d1
End Function
```

```
' Retourne l'âge à l'entrée de l'étude
' ou Null si en dehors de la période d'observation
'
Function Yi(dDebObs As Date, _
            dFinObs As Date, _
            dNaiss As Date, _
            dEntree As Variant, _
            dSortie As Variant, _
            dDeces As Variant) _
            As Variant

  If Not IsNull(dSortie) And dSortie < dDebObs Or _
     Not IsNull(dDeces) And dDeces < dDebObs Or _
     Not IsNull(dEntree) And dEntree > dFinObs Then
    Yi = Null
    Exit Function
  End If
  If IsNull(dEntree) Or _
     dEntree <= dDebObs Then
    Yi = Age(dNaiss, dDebObs)
  Else
    Yi = Age(dNaiss, (dEntree))
  End If
End Function
```

```
' Retourne l'âge prévu à la terminaison de l'étude
' ou Null si en dehors de la période d'observation
'
Function Zi(dDebObs As Date, _
            dFinObs As Date, _
            dNaiss As Date, _
            dEntree As Variant, _
            dSortie As Variant, _
            dDeces As Variant) _
            As Variant

  If Not IsNull(dSortie) And dSortie < dDebObs Or _
     Not IsNull(dDeces) And dDeces < dDebObs Or _
     Not IsNull(dEntree) And dEntree > dFinObs Then
    Zi = Null
    Exit Function
  End If
  Zi = Age(dNaiss, dFinObs)
End Function
```

```
' Retourne l'âge du décès s'il a lieu dans la période
' d'observation
'
Function Ei(dDebObs As Date, _
            dFinObs As Date, _
            dNaiss As Date, _
            dEntree As Variant, _
            dSortie As Variant, _
            dDeces As Variant) _
            As Variant

  If Not IsNull(dDeces) And dDeces >= dDebObs And dDeces <= dFinObs Then
    Ei = Age(dNaiss, (dDeces))
  Else
    Ei = Null
  End If
End Function
```

```
' Retourne l'âge de sortie volontaire si elle a lieu dans la période
' d'observation
'
Function Fi(dDebObs As Date, _
            dFinObs As Date, _
            dNaiss As Date, _
            dEntree As Variant, _
            dSortie As Variant, _
            dDeces As Variant) _
            As Variant

  If Not IsNull(dSortie) And dSortie >= dDebObs And dSortie <= dFinObs Then
    Fi = Age(dNaiss, (dSortie))
  Else
    Fi = Null
  End If
End Function
```

H Fonctions diverses

```
Function Maxi(m1, m2)
  If m1 > m2 Then Maxi = m1 Else Maxi = m2
End Function
```

```
Function Mini(m1, m2)
  If m1 < m2 Then Mini = m1 Else Mini = m2
End Function
```

I Fonctions de calculs sur l'intervalle d'estimation

```
' Age d'entrée prévu dans l'intervalle d'estimation
'
'
Function Ri(iXi As Integer, _
            vYi As Variant, _
            vZi As Variant, _
            vEi As Variant, _
            vFi As Variant) _
            As Variant

  If iXi > vZi Or _
     Not IsNull(vEi) And iXi > vEi Or _
     Not IsNull(vFi) And iXi > vFi Then
    Ri = Null
```

```
  Else
    Ri = Maxi(0, vYi - iXi)
  End If
End Function
```

```
' Age de terminaison prévu dans l'intervalle d'estimation
'
'
Function Si(iXi As Integer, _
            vYi As Variant, _
            vZi As Variant, _
            vEi As Variant, _
            vFi As Variant) _
            As Variant

  If iXi > vZi Or _
     Not IsNull(vEi) And iXi > vEi Or _
     Not IsNull(vFi) And iXi > vFi Then
    Si = Null
  Else
    Si = Mini(1, vZi - iXi)
  End If
End Function
```

```
' Nombre de décès dans l'intervalle d'estimation
'
'
Function Di(iXi As Integer, _
            vYi As Variant, _
            vZi As Variant, _
            vEi As Variant, _
            vFi As Variant) _
            As Variant

  If IsNull(vEi) Then
    Di = 0
  Else
    If vEi >= vYi And vEi <= vZi And vEi >= iXi And vEi < iXi + 1 Then
      Di = 1
    Else
      Di = 0
    End If
  End If
End Function
```

```
' Nombre de sortie volontaire dans l'intervalle d'estimation
'
'
Function Wi(iXi As Integer, _
            vYi As Variant, _
            vZi As Variant, _
            vEi As Variant, _
            vFi As Variant) _
            As Variant

  If IsNull(vFi) Then
    Wi = 0
  Else
    If vFi >= vYi And vFi <= vZi And vFi >= iXi And vFi < iXi + 1 Then
      Wi = 1
    Else
      Wi = 0
    End If
  End If
End Function
```

```
' Exposition dans l'inervalle d'estimation
'
```

```
'
Function Ti(iXi As Integer, _
            vYi As Variant, _
            vZi As Variant, _
            vEi As Variant, _
            vFi As Variant, _
            vRi As Variant, _
            vSi As Variant) _
            As Variant

  If IsNull(vRi) Or IsNull(vSi) Then
    Ti = Null
  Else
    If Not IsNull(vFi) And vFi >= iXi And vFi < iXi + 1 Then
      Ti = vFi - (iXi + vRi)
    Else
      Ti = vSi - vRi
    End If
  End If
End Function
```

J Manipulations des tables

```
' Effacement d'une table
'
Sub DropTable(sTable As String)
Dim sSql As String

  sSql = "DROP TABLE " & sTable
  On Error Resume Next
  DoCmd.RunSQL sSql, False
  On Error GoTo 0
End Sub
```

```
' Copie d'une table (sans index)
'
Sub CopyTable(sSrcTable As String, sDstTable As String)
Dim sSql As String

  sSql = "SELECT * INTO " & sDstTable & " FROM " & sSrcTable & ";"
  DoCmd.RunSQL sSql, False
End Sub
```

```
' Création d'un index
'
Sub CreateIndex(sTable As String, sIndex As String, sFields As String, fPrimary As Boolean)
Dim sSql As String

  If fPrimary Then
    sSql = "CREATE UNIQUE INDEX " & sIndex & " ON " & sTable & "(" & sFields & ")  WITH
PRIMARY;"
  Else
    sSql = "CREATE INDEX " & sIndex & " ON " & sTable & "(" & sFields & ")  WITH IGNORE
NULL;"
  End If
  DoCmd.RunSQL sSql, False
End Sub
```

K Extraction depuis l'AS/400 et pré-traitement

```
#Const AS400_VERIFICATION = 1
' 0: n'inclus pas les zones de contrôle
' 1: inclus les zones de contrôle
#Const SUB_SELECT_CLIENT = 0
' 0: depuis la base locale
' 1: depuis l'AS/400

' Cette procédure s'occupe de l'extraction des données depuis l'AS/400
' et effectue une copie avec pré-traitement dans la base Access
'
Sub ExtraitAS400()
Dim sSql As String, sSql1 As String, sSql2 As String
Dim fPolice As Boolean, fClient As Boolean

' Test de la disponibilité du système AS/400
'--------------------------------------------

  DropTable "AStabent"
  sSql = "SELECT TE.* INTO AStabent "
  sSql = sSql & "FROM RGINDEFIC_TABENTP AS TE"
  On Error GoTo ErreurExtraction
  DoCmd.RunSQL sSql, False
  On Error GoTo 0
  DropTable "AStabent"

' Polices
'---------

' Traitement des polices

  DropTable "ASpolice"
  fPolice = True

  sSql = "CREATE TABLE ASpolice ("
  sSql = sSql & "NoPolice TEXT(12) "
  sSql = sSql & "CONSTRAINT PrimaryKey PRIMARY KEY, "
  sSql = sSql & "IndIndColl TEXT(1), "
  sSql = sSql & "CodEtat TEXT(2), "
  sSql = sSql & "CodTarif TEXT(10), "
  sSql = sSql & "CodBasTec TEXT(10), "
  sSql = sSql & "CodColl TEXT(10), "
  sSql = sSql & "DatEffet DATE, "
  sSql = sSql & "DatEcheance DATE, "
  sSql = sSql & "DatExtinction DATE, "
  sSql = sSql & "NoAssure1 TEXT(10), "
  sSql = sSql & "NoAssure2 TEXT(10) "
  sSql = sSql & ")"
  DoCmd.RunSQL sSql, False

  CreateIndex "ASpolice", "NoAssure1", "NoAssure1", False
  CreateIndex "ASpolice", "NoAssure2", "NoAssure2", False

' Extraction des données polices
'--------------------------------

  sSql = "INSERT INTO ASpolice ("
  sSql = sSql & "NoPolice, "
  sSql = sSql & "IndIndColl, "
  sSql = sSql & "CodEtat, "
  sSql = sSql & "CodTarif, "
  sSql = sSql & "CodBasTec, "
  sSql = sSql & "CodColl, "
  sSql = sSql & "DatEffet, "
  sSql = sSql & "DatEcheance, "
  sSql = sSql & "DatExtinction, "
```

```
    sSql = sSql & "NoAssure1, "
    sSql = sSql & "NoAssure2 "
    sSql = sSql & ") "
    sSql = sSql & "SELECT PO.POCPOL AS NoPolice, "
    sSql = sSql & "PO.POIIC AS IndIndColl, "
    sSql = sSql & "PO.POCETP AS CodEtat, "
    sSql = sSql & "PO.POCTAR AS CodTarif, "
    sSql = sSql & "PO.POCBTE AS CodBasTec, "
    sSql = sSql & "PO.POCCMT AS CodColl, "
    sSql = sSql & "DateFromASv([PO].[PODEFF]) AS DatEffet, "
    sSql = sSql & "DateFromASv([PO].[PODECH]) AS DatEcheance, "
    sSql = sSql & "DateFromASv([PO].[PODEXT]) AS DatExtinction, "
    sSql = sSql & "PO.POCCA1 AS NoAssure1, "
    sSql = sSql & "PO.POCCA2 AS NoAssure2  "
    sSql = sSql & "FROM RGINDEFIC_POLICEP AS PO "
    sSql = sSql & "WHERE PO.POCETP<>'99';"
    DoCmd.RunSQL sSql, False

  ' Clients
  '---------

  ' Traitement des clients

    DropTable "ASclient"
    fClient = True

    sSql = "CREATE TABLE ASclient ("
    sSql = sSql & "NoAssure TEXT(10) "
    sSql = sSql & "CONSTRAINT PrimaryKey PRIMARY KEY, "
    sSql = sSql & "CodSexe TEXT(1), "
  #If AS400_VERIFICATION = 1 Then
    sSql = sSql & "Nom TEXT(22), "
    sSql = sSql & "Prenom TEXT(18), "
    sSql = sSql & "CodEtatCivil TEXT(1), "
    sSql = sSql & "CodNationalite TEXT(4), "
  #End If
    sSql = sSql & "DatNaissance DATE, "
    sSql = sSql & "DatDeces DATE, "
    sSql = sSql & "DatSaisieDeces DATE, "
    sSql = sSql & "DatEntree DATE, "
    sSql = sSql & "DatSortie DATE "
    sSql = sSql & ")"
    DoCmd.RunSQL sSql, False

  ' Extraction des données clients
  '---------------------------------

  #If SUB_SELECT_CLIENT = 0 Then
    sSql1 = "SELECT PO.NoAssure1 FROM ASpolice AS PO"
    sSql2 = "SELECT PO.NoAssure2 FROM ASpolice AS PO"
  #ElseIf SUB_SELECT_CLIENT = 1 Then
    sSql1 = "SELECT PO.POCCA1 FROM RGINDEFIC_POLICEP AS PO "
    sSql1 = sSql1 & "WHERE PO.POCETP<>'99'"
    sSql2 = "SELECT PO.POCCA2 FROM RGINDEFIC_POLICEP AS PO "
    sSql2 = sSql2 & "WHERE PO.POCETP<>'99'"
  #End If

    sSql = "INSERT INTO ASclient ("
    sSql = sSql & "NoAssure, "
    sSql = sSql & "CodSexe, "
  #If AS400_VERIFICATION = 1 Then
    sSql = sSql & "Nom, "
    sSql = sSql & "Prenom, "
    sSql = sSql & "CodEtatCivil, "
    sSql = sSql & "CodNationalite, "
  #End If
    sSql = sSql & "DatNaissance, "
    sSql = sSql & "DatDeces, "
    sSql = sSql & "DatSaisieDeces, "
    sSql = sSql & "DatEntree, "
```

```
  sSql = sSql & "DatSortie"
  sSql = sSql & ") "
  sSql = sSql & "SELECT CL.CLCCLI AS NoAssure, "
  sSql = sSql & "CL.CLCSEX AS CodSexe, "
#If AS400_VERIFICATION = 1 Then
  sSql = sSql & "CL.CLENOM AS Nom, "
  sSql = sSql & "CL.CLEPRN AS Prenom, "
  sSql = sSql & "CL.CLCCIV AS CodEtatCivil, "
  sSql = sSql & "CL.CLCNAT AS CodNationalite, "
#End If
  sSql = sSql & "DateFromASv([CL].[CLDNAI]) AS DatNaissance, "
  sSql = sSql & "DateFromASv([CL].[CLDDEC]) AS DatDeces, "
  sSql = sSql & "DateFromASv([CL].[CLDEDC]) AS DatSaisieDeces, "
  sSql = sSql & "DateFromASv(0) AS DatEntree, "
  sSql = sSql & "DateFromASv(0) AS DatSortie "
  sSql = sSql & "FROM RGINDEFIC_CLIENTP AS CL "
  sSql = sSql & "WHERE CL.CLCCLI In (" & sSql1 & ") Or CL.CLCCLI In (" & sSql2 & ");"
  DoCmd.RunSQL sSql, False

  DropTable "ASpolice_copie"
  DropTable "ASclient_copie"

  MsgBox "Extraction des données de l'AS/400 terminée", vbOKOnly, "Extraction"
  Exit Sub

ErreurExtraction:
  MsgBox "Extraction des données de l'AS/400 abandonnée suite à un problème", vbOKOnly +
vbCritical, "Extraction"
End Sub
```


```
Private Sub cmdOK_Click()
  DoCmd.Hourglass True
  ExtraitAS400
  DoCmd.Hourglass False
  DoCmd.Close acForm, Me.Name
End Sub
```

L Préparation d'une extraction

```
' Prépare une extraction "nettoyée" des incohérences
' sIndColl = I : Polices individuelles
'          = C : Polices collectives
'          = * : Toutes les polices
'
Sub PrepareExtraction(sIndColl As String)
Dim sSql As String, sSql1 As String, sSql2 As String

' Extraction des polices
'------------------------

  DropTable "CRpolice"

  sSql = "SELECT PO.* INTO CRpolice "
  sSql = sSql & "FROM ASpolice AS PO "
  If sIndColl <> "*" Then
```

```
    sSql = sSql & "WHERE PO.IndIndColl = '" & sIndColl & "';"
  End If
  DoCmd.RunSQL sSql, False

  CreateIndex "CRpolice", "PrimaryKey", "NoPolice", True
  CreateIndex "CRpolice", "NoAssure1", "NoAssure1", False
  CreateIndex "CRpolice", "NoAssure2", "NoAssure2", False

' Extraction des clients
'-----------------------

  DropTable "CRclient"

  sSql1 = "SELECT PO.NoAssure1 FROM CRpolice AS PO"
  sSql2 = "SELECT PO.NoAssure2 FROM CRpolice AS PO"
  sSql = "SELECT CL.* INTO CRclient "
  sSql = sSql & "FROM ASclient AS CL "
  sSql = sSql & "WHERE CL.NoAssure In (" & sSql1 & ") Or CL.NoAssure In (" & sSql2 & ");"
  DoCmd.RunSQL sSql, False

  CreateIndex "CRclient", "PrimaryKey", "NoAssure", True

  TraitePolices2T
  TraiteDoublons
  TraiteAssuresOrphelin

  RazDateEntreeSortie
  CalculDateEntree sIndColl
  CalculDateSortie sIndColl

  TraiteSortieLessThanEntree
  TraiteDecesLessThanEntree
  TraiteAssuresOrphelin

  MsgBox "Préparation des données terminée", vbOKOnly, "Préparation"
End Sub
```

```
' Elimination des doublons
'
Sub TraiteDoublons()
Dim sSql As String

' DoCmd.OpenQuery "RQ 0501 UPD NoAssure1 (Elimine doublon)"
  sSql = "UPDATE CRpolice AS PO INNER JOIN Doublons AS DB "
  sSql = sSql & "ON PO.NoAssure1 = DB.NoDoublon SET PO.NoAssure2 = [DB].[NoAssure];"
  DoCmd.RunSQL sSql, False

' DoCmd.OpenQuery "RQ 0501 UPD NoAssure2 (Elimine doublon)"
  sSql = "UPDATE CRpolice AS PO INNER JOIN Doublons AS DB "
  sSql = sSql & "ON PO.NoAssure2 = DB.NoDoublon SET PO.NoAssure2 = [DB].[NoAssure];"
  DoCmd.RunSQL sSql, False

' DoCmd.OpenQuery "RQ 0503 DEL NoAssure Doublons"
  sSql = "DELETE * "
  sSql = sSql & "FROM CRclient AS CL "
  sSql = sSql & "WHERE CL.NoAssure In (SELECT NoDoublon FROM Doublons);"
  DoCmd.RunSQL sSql, False
End Sub
```

```
' Mise à Null des dates d'entrées et de sortie
'
Sub RazDateEntreeSortie()
Dim sSql As String

  sSql = "UPDATE CRclient AS CL SET CL.DatEntree = Null, CL.DatSortie = Null;"
  DoCmd.RunSQL sSql, False
End Sub
```

```
' Détermination de la date d'entrée
'
Sub CalculDateEntree(sIndColl As String)
Dim sSql As String

' DoCmd.OpenQuery "RQ 0202 CRT CRdmin"
  sSql = "SELECT CL.NoAssure, Min(Zdate(PO.DatEffet)) AS DatEntree INTO CRdmin "
  sSql = sSql & "FROM CRclient AS CL INNER JOIN CRpolice AS PO ON CL.NoAssure =
PO.NoAssure1 "
  If sIndColl <> "*" Then
    sSql = sSql & "WHERE PO.IndIndColl = '" & sIndColl & "' "
  End If
  sSql = sSql & "GROUP BY CL.NoAssure "
  sSql = sSql & "ORDER BY CL.NoAssure;"
  DoCmd.RunSQL sSql, False

' DoCmd.OpenQuery "RQ 0203 ADD CRdmin"
  sSql = "INSERT INTO CRdmin ( NoAssure, DatEntree ) "
  sSql = sSql & "SELECT CL.NoAssure, Min(Zdate(PO.DatEffet)) AS DatEntree "
  sSql = sSql & "FROM CRclient AS CL INNER JOIN CRpolice AS PO ON CL.NoAssure =
PO.NoAssure2 "
  If sIndColl <> "*" Then
    sSql = sSql & "WHERE PO.IndIndColl = '" & sIndColl & "' "
  End If
  sSql = sSql & "GROUP BY CL.NoAssure "
  sSql = sSql & "ORDER BY CL.NoAssure;"
  DoCmd.RunSQL sSql, False

' DoCmd.OpenQuery "RQ 0204 CRT CRdmingrp"
  sSql = "SELECT DM.NoAssure, Min(Zdate(DM.DatEntree)) AS DatEntree INTO CRdmingrp "
  sSql = sSql & "FROM CRdmin AS DM "
  sSql = sSql & "GROUP BY DM.NoAssure;"
  DoCmd.RunSQL sSql, False

' DoCmd.OpenQuery "RQ 0205 UPD DatEntree"
  sSql = "UPDATE CRclient AS CL INNER JOIN CRdmingrp AS DM "
  sSql = sSql & "ON CL.NoAssure = DM.NoAssure SET CL.DatEntree = Ndate(DM.DatEntree) "
  sSql = sSql & "WHERE CL.NoAssure=[DM].[NoAssure];"
  DoCmd.RunSQL sSql, False

  DropTable "CRdmin"
  DropTable "CRdmingrp"
End Sub
```

```
' Détermination de la date de sortie
'
Sub CalculDateSortie(sIndColl As String)
Dim sSql As String

' DoCmd.OpenQuery "RQ 0207 CRT CRdmax"
  sSql = "SELECT CL.NoAssure, Max(Zdate(PO.DatExtinction)) AS DatSortieMax, "
  sSql = sSql & "Min(Zdate(PO.DatExtinction)) AS DatSortieMin INTO CRdmax "
  sSql = sSql & "FROM CRclient AS CL INNER JOIN CRpolice AS PO ON CL.NoAssure =
PO.NoAssure1 "
  If sIndColl <> "*" Then
    sSql = sSql & "WHERE PO.IndIndColl = '" & sIndColl & "' "
  End If
  sSql = sSql & "GROUP BY CL.NoAssure "
  sSql = sSql & "ORDER BY CL.NoAssure;"
  DoCmd.RunSQL sSql, False

' DoCmd.OpenQuery "RQ 0208 ADD CRdmax"
  sSql = "INSERT INTO CRdmax ( NoAssure, DatSortieMax, DatSortieMin ) "
  sSql = sSql & "SELECT CL.NoAssure, Max(Zdate(PO.DatExtinction)) AS DatSortieMax, "
  sSql = sSql & "Min(Zdate(PO.DatExtinction)) AS DatSortieMin "
  sSql = sSql & "FROM CRclient AS CL INNER JOIN CRpolice AS PO ON CL.NoAssure =
PO.NoAssure2 "
  If sIndColl <> "*" Then
    sSql = sSql & "WHERE PO.IndIndColl = '" & sIndColl & "' "
  End If
```

```
  sSql = sSql & "GROUP BY CL.NoAssure "
  sSql = sSql & "ORDER BY CL.NoAssure;"
  DoCmd.RunSQL sSql, False

' DoCmd.OpenQuery "RQ 0209 CRT CRdmaxgrp"
  sSql = "SELECT DM.NoAssure, Max(Zdate(DM.DatSortieMax)) AS DatSortieMax, "
  sSql = sSql & "Min(Zdate(DM.DatSortieMin)) AS DatSortieMin INTO CRdmaxgrp "
  sSql = sSql & "FROM CRdmax AS DM "
  sSql = sSql & "GROUP BY DM.NoAssure;"
  DoCmd.RunSQL sSql, False

' DoCmd.OpenQuery "RQ 0210 UPD DatSortie"
  sSql = "UPDATE CRclient AS CL INNER JOIN CRdmaxgrp AS DM "
  sSql = sSql & "ON CL.NoAssure = DM.NoAssure SET CL.DatSortie = Ndate(DM.DatSortieMax) "
  sSql = sSql & "WHERE (((CL.NoAssure)=[DM].[NoAssure])) AND DM.DatSortieMin<>0;"
  DoCmd.RunSQL sSql, False

' Si l'assuré est décédé, on n'utilise pas la date de sortie
  sSql = "UPDATE CRclient AS CL SET CL.DatSortie = Null "
  sSql = sSql & "WHERE CL.DatSortie Is Not Null AND CL.DatDeces Is Not Null;"
  DoCmd.RunSQL sSql, False

  DropTable "CRdmax"
  DropTable "CRdmaxgrp"
End Sub
```

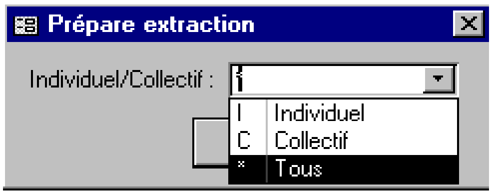
```
' Supprime les polices 2 têtes qui n'ont que l'assuré 1 de défini
' Ces polices ont étés éteintes avant 1989, la saisie n'avait pas
' été faite correctement
'
Sub TraitePolices2T()
Dim sSql As String

' DoCmd.OpenQuery "RQ 0213 DEL Polices I2LS 1 assuré"
  sSql = "DELETE * "
  sSql = sSql & "FROM CRpolice AS PO "
  sSql = sSql & "WHERE PO.NoAssure2='          ' AND PO.CodTarif Like 'I2*' OR "
  sSql = sSql & "PO.NoAssure2='          ' AND PO.CodTarif Like 'D2*';"
  DoCmd.RunSQL sSql, False
End Sub
```

```
' Supprime les assurés qui sont orphélins
'
Sub TraiteAssuresOrphelin()
Dim sSql As String, sSql1 As String, sSql2 As String

' Recherche les assurés qui ont un lien avec au moins une police
' DoCmd.OpenQuery "RQ 0214 CRT CRclient_util"
  sSql1 = "SELECT PO.NoAssure1 FROM CRpolice AS PO"
  sSql2 = "SELECT PO.NoAssure2 FROM CRpolice AS PO"
  sSql = "SELECT CL.NoAssure INTO CRclient_util "
  sSql = sSql & "FROM CRclient AS CL "
  sSql = sSql & "WHERE CL.NoAssure In (" & sSql1 & ") Or CL.NoAssure In (" & sSql2 & ");"
  DoCmd.RunSQL sSql, False

  CreateIndex "CRclient_util", "PrimaryKey", "NoAssure", True

' Supprime les assurés qui n'ont de lien avec aucune police
  sSql = "DELETE CL.NoAssure "
  sSql = sSql & "FROM CRclient AS CL "
  sSql = sSql & "WHERE CL.NoAssure Not In (SELECT CL1.NoAssure FROM CRclient_util AS CL1);"
  DoCmd.RunSQL sSql, False

  DropTable "CRclient_util"
End Sub
```

```
' Supprime les polices des assurés dont la date de sortie
' est inférieure à la date d'entrée
```

```
'
Sub TraiteSortieLessThanEntree()
Dim sSql As String

  sSql = "DELETE PO.* "
  sSql = sSql & "FROM (CRpolice AS PO INNER JOIN CRclient AS CL ON PO.NoAssure1 =
CL.NoAssure) "
  sSql = sSql & "LEFT JOIN CRclient AS CL1 ON PO.NoAssure2 = CL1.NoAssure "
  sSql = sSql & "WHERE CL.DatSortie<CL.DatEntree OR CL1.DatSortie<CL1.DatEntree;"
  DoCmd.RunSQL sSql, False
End Sub
```

```
' Supprime les polices des assurés dont la date de décès
' est inférieure à la date d'entrée
'
Sub TraiteDecesLessThanEntree()
Dim sSql As String

  sSql = "DELETE PO.* "
  sSql = sSql & "FROM (CRpolice AS PO INNER JOIN CRclient AS CL ON PO.NoAssure1 =
CL.NoAssure) "
  sSql = sSql & "LEFT JOIN CRclient AS CL1 ON PO.NoAssure2 = CL1.NoAssure "
  sSql = sSql & "WHERE CL.DatDeces<CL.DatEntree OR CL1.DatDeces<CL1.DatEntree;"
  DoCmd.RunSQL sSql, False
End Sub
```




```
Private Sub cmdOK_Click()
  DoCmd.Hourglass True
  PrepareExtraction txtIndColl
  DoCmd.Hourglass False
  DoCmd.Close acForm, Me.Name
End Sub
```

M Archive une extraction

```
' Enregistre l'extraction effectuée afin de pouvoir
' générer autant d'estimation que l'on veut sur la
' même base
'
Sub NouvelleExtraction(sDesc As String)
Dim dbsCur As Database, rstEnteteExtraction As Recordset
Dim lNoExtraction As Long, sSql As String

  Set dbsCur = CurrentDb
  Set rstEnteteExtraction = dbsCur.OpenRecordset("EnteteExtraction", dbOpenDynaset)

' Création de l'entête de l'extraction

  With rstEnteteExtraction
    .AddNew
    ![Descriptif] = sDesc
    ![DateExtraction] = Now
    lNoExtraction = ![NoExtraction]
    .Update
  End With

  sSql = "INSERT INTO DetailExtraction "
```

```
  sSql = sSql & "( NoExtraction, NoAssure, CodSexe, "
  sSql = sSql & "DatNaissance, DatEntree, DatSortie, DatDeces ) "
  sSql = sSql & "SELECT " & CStr(lNoExtraction) & " AS NoExtraction, "
  sSql = sSql & "CL.NoAssure, CL.CodSexe, "
  sSql = sSql & "CL.DatNaissance, CL.DatEntree, "
  sSql = sSql & "CL.DatSortie, CL.DatDeces "
  sSql = sSql & "FROM CRclient AS CL;"
  DoCmd.RunSQL sSql, False

  rstEnteteExtraction.Close
  dbsCur.Close

  MsgBox "Extraction enregistrée sous le numéro " & lNoExtraction, vbOKOnly
End Sub
```


```
Private Sub cmdOK_Click()
  If Not IsNull(txtDesc) Then
    DoCmd.Hourglass True
    NouvelleExtraction txtDesc
    DoCmd.Hourglass False
    DoCmd.Close acForm, Me.Name
  Else
    MsgBox "Tous les champs doivent être remplis", vbOKOnly + vbExclamation
  End If
End Sub
```

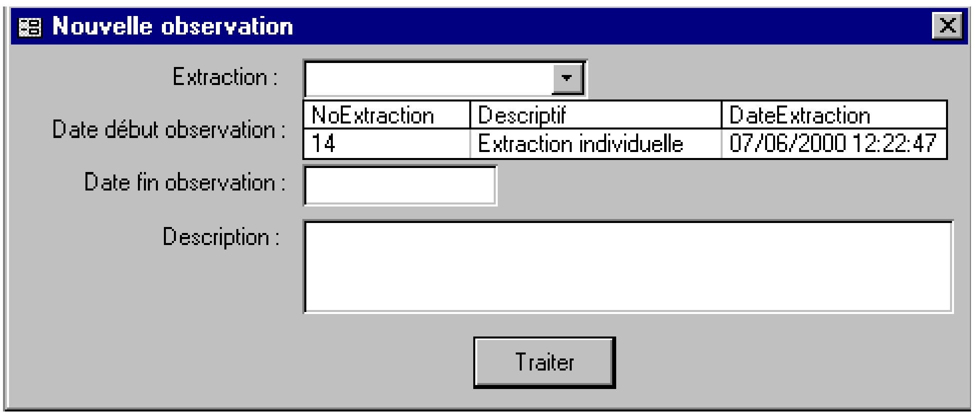
N Génère une estimation

```
' Génère une estimation à partir des données de l'extraction
' pour une période d'observation
'
Sub TraiteObservation(lNoExtraction As Long, sDesc As String, dDeb As Date, dFin As Date)
Dim dbsCur As Database
Dim rstEnteteExtraction As Recordset, rstDetailExtraction As Recordset
Dim rstEnteteEstimation As Recordset, rstDetailEstimation As Recordset
Dim rstDetailParAgeEstimation As Recordset
Dim lNoEstimation As Long
Dim vYi As Variant, vZi As Variant, vEi As Variant, vFi As Variant
Dim vRi As Variant, vSi As Variant, vDi As Variant, vWi As Variant
Dim vTi As Variant
Dim iXi As Integer, iYi As Integer, iZi As Integer, iEi As Integer, iFi As Integer
Dim sSql As String, sSearch As String

  Set dbsCur = CurrentDb
  Set rstEnteteExtraction = dbsCur.OpenRecordset("EnteteExtraction", dbOpenDynaset)
  Set rstDetailExtraction = dbsCur.OpenRecordset("DetailExtraction ", dbOpenDynaset)
  Set rstEnteteEstimation = dbsCur.OpenRecordset("EnteteEstimation", dbOpenDynaset)
  Set rstDetailEstimation = dbsCur.OpenRecordset("DetailEstimation", dbOpenDynaset)
  Set rstDetailParAgeEstimation = dbsCur.OpenRecordset("DetailParAgeEstimation",
dbOpenDynaset)

' Vérifie l'existence de l'extraction

  With rstEnteteExtraction
    sSearch = "NoExtraction = " & CStr(lNoExtraction)
```

```
      .FindFirst sSearch
      If .NoMatch Then
        MsgBox "Extraction non trouvée dans la base de données", vbCritical + vbOKOnly
        Exit Sub
      End If
    End With

  ' Création de l'entête de l'estimation

    With rstEnteteEstimation
      .AddNew
      ![NoExtraction] = lNoExtraction
      ![Descriptif] = sDesc
      ![DateDebut] = dDeb
      ![DateFin] = dFin
      lNoEstimation = ![NoEstimation]
      .Update
    End With

  ' Création du détail de l'estimation avec filtrage des enregistrements

    sSql = "SELECT * FROM DetailExtraction "
    sSql = sSql & "WHERE NoExtraction = " & CStr(lNoExtraction) & " AND "
    sSql = sSql & "DatEntree <= " & CStr(CDbl(dFin)) & " AND "
    sSql = sSql & "(DatDeces >= " & CStr(CDbl(dDeb)) & " OR DatDeces IS NULL) AND "
    sSql = sSql & "(DatSortie >= " & CStr(CDbl(dDeb)) & " OR DatSortie IS NULL)"
    Set rstDetailExtraction = dbsCur.OpenRecordset(sSql, dbOpenDynaset)

    With rstDetailExtraction
      While Not rstDetailExtraction.EOF
        vYi = Yi(dDeb, dFin, ![DatNaissance], ![DatEntree], ![DatSortie], ![DatDeces])
        vZi = Zi(dDeb, dFin, ![DatNaissance], ![DatEntree], ![DatSortie], ![DatDeces])
        vEi = Ei(dDeb, dFin, ![DatNaissance], ![DatEntree], ![DatSortie], ![DatDeces])
        vFi = Fi(dDeb, dFin, ![DatNaissance], ![DatEntree], ![DatSortie], ![DatDeces])

  ' Conserve le détail des calculs des âges

        With rstDetailEstimation
          .AddNew
          ![NoEstimation] = lNoEstimation
          ![NoMembre] = rstDetailExtraction![NoMembre]
          ![CodSexe] = rstDetailExtraction![CodSexe]
          ![Yi] = vYi
          ![Zi] = vZi
          ![Ei] = vEi
          ![Fi] = vFi
          .Update
        End With

        iYi = Fix(vYi)       ' Age prévu à l'entrée dans l'étude
        iZi = Fix(vZi) + 1   ' Age prévu à la terminaison

  ' Calcul de l'exposition par tranche d'âge

        With rstDetailParAgeEstimation
          For iXi = iYi To iZi - 1
            vRi = Ri(iXi, vYi, vZi, vEi, vFi)
            vSi = Si(iXi, vYi, vZi, vEi, vFi)
            vDi = Di(iXi, vYi, vZi, vEi, vFi)
            vWi = Wi(iXi, vYi, vZi, vEi, vFi)
            vTi = Ti(iXi, vYi, vZi, vEi, vFi, vRi, vSi)
            If IsNull(vRi) And IsNull(vSi) Then Exit For
            .AddNew
            ![NoEstimation] = lNoEstimation
            ![NoMembre] = rstDetailExtraction![NoMembre]
            ![X] = iXi
            ![Ri] = vRi
            ![Si] = vSi
            ![Di] = vDi
            ![Wi] = vWi
```

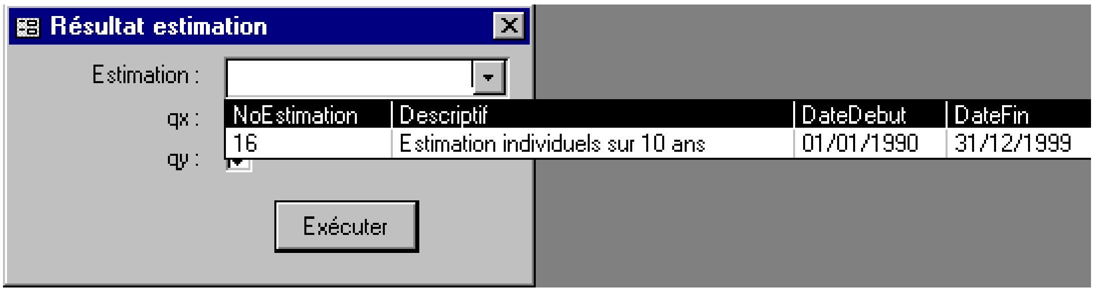
```
            ![Ti] = vTi
            .Update
          Next
        End With

        .MoveNext
      Wend

    End With

    rstDetailParAgeEstimation.Close
    rstDetailEstimation.Close
    rstEnteteEstimation.Close
    rstDetailExtraction.Close
    rstEnteteExtraction.Close
    dbsCur.Close

    GenerateQx lNoEstimation

    MsgBox "Estimation enregistrée sous le numéro " & CStr(lNoEstimation), vbOKOnly
  End Sub
```

```
' Calcul des probabilités de décès brutes
'
Sub GenerateQx(lNoEstimation As Long)
Dim sSql As String

  sSql = "DELETE DEM.NoEstimation "
  sSql = sSql & "FROM DetailEstimationMortalite AS DEM "
  sSql = sSql & "WHERE DEM.NoEstimation=" & CStr(lNoEstimation)
  DoCmd.RunSQL sSql, False

  sSql = "INSERT INTO DetailEstimationMortalite ( SNi, X, S, SDi, SWi, STi, NoEstimation )
"
  sSql = sSql & "SELECT Count(DPAE.NoEstimation) AS SNi, DPAE.X, DE.CodSexe AS S, "
  sSql = sSql & "Sum(DPAE.Di) AS SDi, Sum(DPAE.Wi) AS SWi, "
  sSql = sSql & "Sum(DPAE.Ti) AS STi, DPAE.NoEstimation "
  sSql = sSql & "FROM DetailEstimation AS DE INNER JOIN DetailParAgeEstimation AS DPAE ON
(DE.NoMembre = DPAE.NoMembre) AND "
  sSql = sSql & "(DE.NoEstimation = DPAE.NoEstimation) "
  sSql = sSql & "GROUP BY DPAE.X, DE.CodSexe, DPAE.NoEstimation "
  sSql = sSql & "HAVING DPAE.NoEstimation=" & CStr(lNoEstimation)
  DoCmd.RunSQL sSql, False

  sSql = "UPDATE DetailEstimationMortalite AS DEM "
  sSql = sSql & "SET Qx = SDi / STi "
  sSql = sSql & "WHERE DEM.NoEstimation=" & CStr(lNoEstimation)
  DoCmd.RunSQL sSql, False

End Sub
```



```
Private Sub cmdOK_Click()
  If Not IsNull(lstExtraction) And _
     Not IsNull(txtDesc) And _
     Not IsNull(txtDatDebut) And _
     Not IsNull(txtDatFin) Then
    DoCmd.Hourglass True
    TraiteObservation lstExtraction, txtDesc, txtDatDebut, txtDatFin
    DoCmd.Hourglass False
    DoCmd.Close acForm, Me.Name
  Else
    MsgBox "Tous les champs doivent être remplis", vbOKOnly + vbExclamation
  End If
End Sub
```

O Affichage du résultat de l'estimation




```
Private Sub cmdOK_Click()
  If Not IsNull(lstEstimation) And _
     Not IsNull(cbxQx) And _
     Not IsNull(cbxQy) Then
    If cbxQx.Value Then DoCmd.OpenQuery "RQ 1001 Résultat Qx"
    If cbxQy.Value Then DoCmd.OpenQuery "RQ 1001 Résultat Qy"
    DoCmd.Close acForm, Me.Name
  Else
    MsgBox "Tous les champs doivent être remplis", vbOKOnly + vbExclamation
  End If
End Sub
```