TD du 28/11/2023

Ajout de chocs à un modèle de mortalité prospective

1. Introduction

La construction de projections d'espérances de vie fait l'objet de nombreux travaux depuis l'article fondateur de Lee et Carter (Lee et Carter [1992]).

Dans la perspective d'extrapoler dans le futur les tendances observées dans le passé, la majorité des approches proposées se basent sur une « surface de mortalité », mesurant les forces de mortalité par âge et année du moment, qu'il s'agit donc d'extrapoler dans la dimension temporelle.

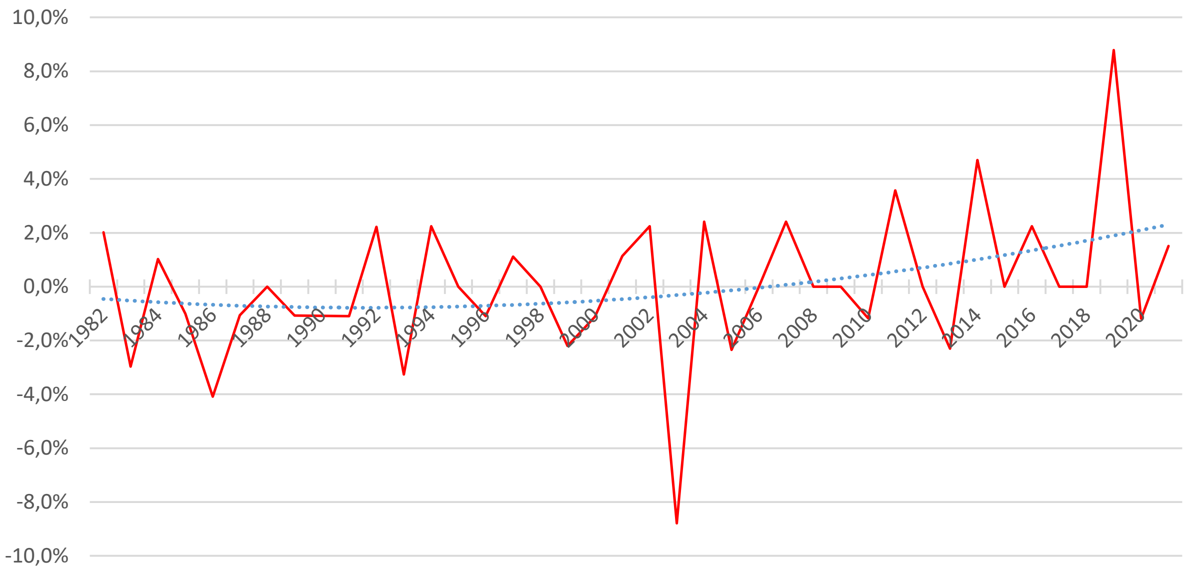
Les modèles inspirés par Lee et Carter commencent par réduire la dimension en réalisant un ACP puis en extrapolant une ou deux séries temporelles associées à la projection sur les axes principaux.

Bongaarts (Bongaarts [2004]) a proposé une démarche différente basée sur des ajustements paramétriques par année du moment et de l'extrapolation des coefficients estimés chaque année.

Dans Bongaarts [2004], l'auteur utilise toutefois une représentation paramétrique assez frustre (modèle de Thatcher) qui ne permet pas d'englober tous les âges. De plus, il limite son extrapolation à 2 paramètres sur 3 en les traitant de manière indépendante, ce qui est une approximation discutable.

Les modèles de ce type projettent des séries $t \to \mu(x,t)$ régulières. Pourtant, lorsque l'on examine les variations annuelles du taux de mortalité en France par exemple, on observe une volatilité assez forte :

Fig. 1 : **Variation annuelle du taux de mortalité[1]**



Les modèles classiques ci-dessus ne peuvent pas rendre compte de ces variations de court terme. Des propositions d'approches ont été formulées par exemple dans GUETTE [2010] ou CURRIE et al. [2003], mais avec un objectif un peu différent, ces travaux se proposant de modéliser des catastrophes comme des guerres ou des épidémies. Plus récemment, une approche utilisant des modèles à changement de régime a été proposée dans ROBBEN et ANTONIO [2023].

Notre propos ici n'est pas de modéliser des catastrophes, mais d'intégrer la volatilité ci-dessus dans la modélisation pour fournir une évaluation plus précise des espérances de vie résiduelles lorsque l'on dispose d'une estimation sans biais des taux de mortalité.

On propose donc ici une approche spécifique avec pour objectif de rendre compte de cette volatilité de court terme et de son impact dans l'anticipation des espérances de vie résiduelles prospectives.

2. Introduction de chocs sur un modèle de mortalité prospective

On s'inspire ici des modèles de fragilité proposés par VAUPEL et al. [1979], en affectant une fonction de hasard de base régulière d'un choc ne dépendant que du temps, avec un modèle à hasard proportionnel.

On décrit ci-après la spécification proposée, puis une méthode d'estimation des paramètres dans le cadre d'un maximum de vraisemblance conditionnel.

[1] Source : https://actudactuaires.typepad.com/laboratoire/2021/01/taux-de-mortalit%C3%A9-en-france-depuis-1982.html

a. Spécification

On considère la spécification suivante de la fonction de hasard de l'année du moment *t* :

$$\mu(x,t) = Z_t \times \mu_0(x,t)$$

Significat° cf en dessous

avec la forme semi-paramétrique de la fonction de hasard de base $\ln \mu_0(x,t) = \alpha_x + \beta_x k_t$. On fait l'hypothèse que les chocs sont en moyenne centrés, soit $E(Z_t) = 1$ et on impose les conditions d'identifiabilité classiques pour la fonction de hasard de base (cf. BROUHNS et al. [2002]) :

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

Il s'agit d'estimer les paramètres a et la matrice (α, β, k), puis d'extrapoler ensuite la série temporelle $t \to k(t)$.

b. Détermination de la log-vraisemblance

Pour l'estimation par maximum de vraisemblance, on peut montrer que tout se passe comme si le nombre de décès observés suivait une loi de Poisson,

$$D_{x,t} \sim P\left(E_{x,t} \times \mu(x,t)\right),$$

ce qui conduit à l'expression suivante de la vraisemblance conditionnelle pour une observation, en notant $\lambda = E_{x,t} \times \mu_0(x,t)$:

$$P(D = d | Z) = e^{-\lambda Z} \frac{\lambda^d}{d!} Z^d .$$

La vraisemblance s'en déduit aisément :

$$P(D = d) = E_Z\left[P(D = d | Z)\right] = \int e^{-\lambda z} z^d \frac{\lambda^d}{d!} dF_Z(z).$$

On fait alors le choix d'une loi Gamma de paramètres a et b pour la distribution de Z, soit $f_Z(z) = z^{a-1} \frac{b^a e^{-b \times z}}{\Gamma(a)}$, ce qui conduit à $P(D = d) = \frac{\lambda^d}{d!} \frac{b^a}{\Gamma(a)} \int_0^{+\infty} e^{-(\lambda+b)z} z^{d+a-1} dz$. À l'aide du changement de variable $u = (\lambda + b) z$, on obtient l'expression suivante de la vraisemblance pour une observation

$$P(D=d) = \frac{\lambda^d}{d!} \frac{b^a}{\Gamma(a)} \frac{1}{(\lambda+b)^{d+a}} \int_0^{+\infty} e^{-u} u^{d+a-1} dz = \frac{\lambda^d}{d!} \frac{b^a}{\Gamma(a)} \frac{1}{(\lambda+b)^{d+a}} \Gamma(d+a),$$

ce qui donne en passant à la log-vraisemblance

$$\ln P(D=d) = f(a,b) + d \ln(\lambda) - (d+a) \ln(\lambda+b),$$

avec $f(a,b) = \ln\left(b^a \frac{\Gamma(d+a)}{\Gamma(d+1)\Gamma(a)} \right)$.

En fonction des paramètres (α,β,k) et conditionnellement à (a,b), la log-vraisemblance pour une observations est de la forme $l(\alpha,\beta,k) = \ln P(D=d)$ avec $\lambda = E_{x,t} \times \mu_0(x,t) = E_{x,t} \times e^{\alpha_x + \beta_x k_t}$, ce dont on déduit que les dérivées partielles par rapport aux différents paramètres sont égales à :

$$\frac{\partial}{\partial p} \ln P(D=d) = d \frac{1}{\lambda} \frac{\partial \lambda}{\partial p} - (d+a) \frac{1}{\lambda+b} \frac{\partial \lambda}{\partial p} = \left(\frac{d}{\lambda} - \frac{d+a}{\lambda+b} \right) \frac{\partial \lambda}{\partial p},$$

avec p l'un des paramètres (α,β,k). On a par ailleurs $\frac{\partial \lambda}{\partial \alpha} = \lambda$, $\frac{\partial \lambda}{\partial \beta} = k\lambda$ et $\frac{\partial \lambda}{\partial k} = \beta\lambda$. La log-vraisemblance a donc, conditionnellement à (a,b) l'allure suivante (à une constante additive près) :

$$\ln L = \sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \ln P(D_{x,t} = d_{x,t} | a,b) = \sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \left[d_{x,t} \ln(\lambda_{x,t}) - (d_{x,t} + a) \ln(\lambda_{x,t} + b) \right].$$

Il s'agit donc de maximiser la fonction ci-dessus en (α,β,k) sous la contrainte que $\sum_{x=x_m}^{x_M} \beta_x = 1$ et $\sum_{t=t_m}^{t_M} k_t = 0$. ← cf en bas pourquoi ?

c. Estimation des paramètres

L'estimation des paramètres peut être réalisée en deux temps : dans une première étape, on estime le paramètres de la fragilité, puis, dans une seconde étape, on maximise en (α,β,k) la log-vraisemblance ci-dessus.

La condition $E(Z_t) = 1$ implique $a = b$. On a par ailleurs $V(Z_t) = \frac{a}{b^2} = \frac{1}{a}$, donc le paramètre de contrôle de la perturbation Z_t est l'inverse de la variance $a = \sigma_Z^{-2}$. L'estimation directe de ce paramètre peut être effectuée de la manière suivante en observant que les intensités annuelles moyennes de sortie sont de la forme

$$\overline{\mu}(t) = Z_t \times \overline{\mu}_0(t) \text{ avec } \overline{\mu}_0(t) = \frac{\sum_{x=x_m}^{x_M} E_{x,t}\mu_0(x,t)}{\sum_{x=x_m}^{x_M} E_{x,t}}$$

d'où l'on tire $E\left(\overline{\mu}(t)\right) = \overline{\mu}_0(t)$, $V\left(\overline{\mu}(t)\right) = V(Z_t)\overline{\mu}_0^2(t)$ puis $V(Z_t) = \frac{V\left(\overline{\mu}(t)\right)}{E\left(\overline{\mu}(t)\right)^2}$. Il est alors direct de proposer comme estimateur

$$\hat{\sigma}_Z^2 = \frac{\frac{1}{t_M - t_m + 1}\sum_{t=t_m}^{t_M}\left(\hat{\mu}(t) - \frac{1}{t_M - t_m + 1}\sum_{t=t_m}^{t_M}\hat{\mu}(t)\right)^2}{\left(\frac{1}{t_M - t_m + 1}\sum_{t=t_m}^{t_M}\hat{\mu}(t)\right)^2}.$$

avec comme estimateur $\hat{\mu}(t) = \frac{\sum_{x=x_m}^{x_M} E_{x,t}\hat{\mu}(x,t)}{\sum_{x=x_m}^{x_M} E_{x,t}}$ et $\hat{\mu}(x,t) = \frac{D_{x,t}}{E_{x,t}}$ l'estimateur de Hoem de la fonction de hasard.

À l'issue de cette première étape, il reste à estimer (α, β, k), qui est solution des conditions du premier ordre :

$$\frac{\partial}{\partial \alpha_x}\ln L = \sum_{t=t_m}^{t_M}\left(\frac{d}{\lambda_{x,t}} - \frac{d+a}{\lambda_{x,t}+b}\right)\lambda_{x,t} = 0$$

$$\frac{\partial}{\partial \beta_x}\ln L = \sum_{t=t_m}^{t_M}\left(\frac{d}{\lambda_{x,t}} - \frac{d+a}{\lambda_{x,t}+b}\right)k_t\lambda_{x,t} = 0$$

$$\frac{\partial}{\partial k_t}\ln L = \sum_{x=x_m}^{x_M}\left(\frac{d}{\lambda_{x,t}} - \frac{d+a}{\lambda_{x,t}+b}\right)\beta_x\lambda_{x,t} = 0$$

Ce système est non linéaire

3. Calcul des espérances de vie résiduelles prospectives

Dans le modèle proposé, le calcul de l'espérance de vie prospective

$$e(x,t) = \sum_{i \geq 0} \prod_{j=0}^{i} E\left[\mathbf{exp}\left(-\mu_{x+j,t+j}\right)\right]$$

se met sous la forme

$$e(x,t) = \sum_{i \geq 0} \prod_{j=0}^{i} \left(\frac{b}{b+\mu_0(x+j,t+j)}\right)^a$$

car la Transformée de Laplace d'une loi Gamma s'écrit $E\left(e^{-xZ_t}\right) = \left(\frac{b}{b+x}\right)^a$, et comme $E\left(e^{-\mu(x,t)}\right) = E\left(e^{-Z_t \times \mu_0(x,t)}\right)$, on en déduit que :

$$E\left(e^{-\mu(x,t)}\right) = E\left(e^{-Z_t \times \mu_0(x,t)}\right) = \left(\frac{b}{b+\mu_0(x,t)}\right)^a .$$

On peut remarquer que, lorsque $b = a \to +\infty$, $E\left(e^{-\mu(x,t)}\right) \to \mu_0(x,t)$ et on retrouve alors la formule classique $e(x,t) = \sum_{i \geq 0} \prod_{j=0}^{i} \mathbf{exp}\left(-\mu_0(x+j,t+j)\right)$.

4. Références

AGALVA E., BLANPAIN N. [2021] Projections de population 2021-2070, Insee Résultats.

BONGAARTS J. [2004] "Long-range trends in adults mortality : Models and projection methods", Population Council .

BROUHNS N., DENUIT M., VERMUNT J.K. [2002] A Poisson log-bilinear regression approach to the construction of projected lifetables, *Insurance, Mathematic and Economics*, vol. 31.

CURRIE I.; DURBAN M. ; EILERS P. [2003] Using P-splines to extrapolate two-dimensional Poisson data, P roceedings of 18th International Workshop on Statistical Modelling, Leuven, Belgium.

DEBONNEUIL E. [2015] Modèle paramétrique de mortalité en fonction de l'âge, pour des applications à des portefeuilles de retraite, Mémoire d'actuaire, ISFA.

GUILBAUD C. [2018] Nouveaux modèles d'analyse et de projection de la mortalité, application a la population française, Mémoire d'actuaire, Dauphine.

GUETTE V. |2010] La prise en compte des catastrophes dans la modélisation de la mortalité, Mémoire d'actuaire, ISFA.

LEE R. D., CARTER L. [1992] "Modeling and forecasting us mortality", *Journal of the American statistical association* 87(419), 659{671.

PLANCHET F., THÉROND P.E. [2011] *Modélisation statistique des phénomènes de durée – applications actuarielles*, Paris : Economica.

ROBBEN J., ANTONIO K. [2023] Catastrophe risk in a stochastic multi-population mortality model, Document de travail, arXiv:2306.15271.

THATCHER A.R. [1999] The Long-term Pattern of Adult Mortality and the Highest Attained Age., *Journal of the Royal Statistical Society*, 162, Part 1: 5-43.

VAUPEL J. W., MANTON K., STALLARD E. [1979] « The impact of heterogeneity in individual frailty on the dynamics of mortality », *Demography*, 16, p. 439-454.

Modèles de durée

Ajout de chocs à un modèle de mortalité prospective

+++ Pour partiel

2-a - Spécification

Fonction de hasard : $\mu(x,t) = Z_t \, \mu_0(x,t)$

avec forme semi-paramétrique de la fonct° de hasard de base :

$$\ln(\mu_0(x,t)) = \alpha_x + \beta_x k_t$$

α_x : niveau moyen du log des taux instantanés à chaque âge. (ils sont généralem^t ↗ sauf aux âges jeunes).

k_t : Ils décrivent l'évolut°, en log, de la mortalité au cours du temps. Les coeff k_t sont globalem^t ↘ et suivent une tendance approximativem^t linéaire. Cette évolut° du coeff traduit l'accroissem^t régulier de l'$\mathbb{E}$ de vie avec le temps (amélioràt° de l'hygiène, du syst de santé...)

β_x : sensibilité par âge à cette tendance d'évolut° (coeff le + souvent positifs)

* On fait l'hypothèse $\mathbb{E}[Z_t] = 1$ (i.e les chocs sont en moyenne centrés).
* Conditions d'identifiabilité classiques pour la fonct° de hasard de base :

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

car la modification $(\alpha_x, \beta_x, k_t) \to (\alpha_x + c\beta_x, \frac{\beta_x}{d}, d(k_t - c))$ n'affecte pas la valeur de $\ln(\mu_0(x,t))$

2-b - Déterminat° de la log-vraisemblance

On considère que le nombre de décès observés suit une loi de Poisson :

$$D_{x,t} \sim \mathcal{P}(E_{x,t}\, \mu(x,t))$$

facteur de l'exposit° au risque de la population étudiée

=> En notant $\lambda = E_{x,t}\, \mu_0(x,t)$ on obtient l'expression suivante de la vraissemblance conditionnelle pour une observation : $\to D_{x,t} \sim \mathcal{P}(\lambda Z_t)$

$$\mathbb{P}(D = d \mid Z) = e^{-\lambda Z} \frac{\lambda^d}{d!} Z^d$$

$$\Rightarrow \mathbb{P}(D = d) = \mathbb{E}_Z[\mathbb{P}(D = d \mid Z)] = \int e^{-\lambda z} z^d \frac{\lambda^d}{d!} \, dF_Z(z)$$

On suppose que $Z \sim \Gamma(a,b) \Rightarrow f_Z(z) = z^{a-1} \frac{b^a e^{-bz}}{\Gamma(a)}$

$$\Rightarrow \mathbb{P}(D = d) = \frac{\lambda^d}{d!} \frac{b^a}{\Gamma(a)} \int_0^\infty e^{-\lambda z} e^{-bz} z^{a-1} z^d \, dz$$

$$= \frac{\lambda^d}{d!} \frac{b^a}{\Gamma(a)} \int_0^\infty e^{-(\lambda+b)z} z^{a+d-1} \, dz$$

$$= \frac{\lambda^d}{d!} \frac{b^a}{\Gamma(a)} \left(\frac{1}{\lambda+b}\right)^{a+d} \int_0^\infty e^{-u} u^{a+d-1} \, du$$

$$= \frac{\lambda^d}{d!} \left(\frac{1}{\lambda+b}\right)^{a+d} \frac{b^a}{\Gamma(a)} \Gamma(a+d) = b^a \frac{\Gamma(a+d)}{\Gamma(d+1)\Gamma(a)} \frac{\lambda^d}{(\lambda+b)^{a+d}}$$

Changem^t de var.
$u = (\lambda+b)z$
$\to du = (\lambda+b)dz$

$$\Rightarrow \ln(\mathbb{P}(D = d)) = f(a,b) + d\ln(\lambda) - (d+a)\ln(\lambda+b)$$

avec $f(a,b) = \ln\left(b^a \frac{\Gamma(d+a)}{\Gamma(d+1)\Gamma(a)}\right)$

On rappelle que $\lambda = E_{x,t}\, \mu_0(x,t)$ et $\ln(\mu_0(x,t)) = \alpha_x + \beta_x k_t$

=> En fonct° de (α, β, k) et conditionnellement à (a,b), la log-vraisemblance pour une observat° est de la forme :

$$\ell(\alpha, \beta, k) = \ln(\mathbb{P}(D = d)) \quad \text{avec } \lambda = E_{x,t}\, \mu_0(x,t) = E_{x,t}\, e^{\alpha_x + \beta_x k_t}$$

$$\Rightarrow p\in\{\alpha,\beta,k\},\ \frac{\partial}{\partial p}\ln(P(D=d)) = d\,\frac{1}{\lambda}\frac{\partial\lambda}{\partial p} - (d+a)\,\frac{1}{\lambda+b}\frac{\partial\lambda}{\partial p}$$

$$= \left[\frac{d}{\lambda} - \frac{d+a}{\lambda+b}\right]\frac{\partial\lambda}{\partial p} \quad *$$

$$\text{avec } \frac{\partial\lambda}{\partial p} = \begin{cases}\lambda & \text{si } p=\alpha\\ k\lambda & \text{si } p=\beta\\ \beta\lambda & \text{si } p=k\end{cases}$$

Donc, conditionnellement à (a,b), la log-vraisemblance est de la forme (à une cte additive près):

$$\ln(\mathcal{L}) = \sum_{x=x_m}^{x_M}\sum_{t=t_m}^{t_M}\ln(P(D=d|a,b)) = \sum_{x=x_m}^{x_M}\sum_{t=t_m}^{t_M}\left[d_{x,t}\ln(\lambda_{x,t}) - (d_{x,t}+a)\ln(\lambda_{x,t}+b)\right]$$

→ Il s'agit donc de maximiser $\ln(\mathcal{L})$ en (α,β,k) sous la contrainte que $\sum_{x=x_m}^{x_M}\beta_x = 1$ et $\sum_{t=t_m}^{t_M}k_t = 0$.

2-c- Estimation des paramètres

L'estimation peut être réalisée en 2 temps:

* Estimat° du paramètre de fragilité (ici Z)
* On maximise en (α,β,k) la log-vraisemblance.

→ La condition $\mathbb{E}[Z_t]=1 \Rightarrow a=b$

car $Z_t \sim \Gamma(a,b) \Rightarrow$ si $\mathbb{E}[Z_t]=1 \Rightarrow 1=\frac{a}{b}$

$\Rightarrow \mathbb{V}[Z_t] = \frac{a}{b^2} = \frac{1}{a}$

$\boxed{\text{Donc le paramètre de contrôle de la perturbat° } Z \text{ est l'inverse de la variance } a=\sigma_Z^{-2}}$

On peut directement estimer ce paramètre en observant que les intensités annuelles moyennes de sortie sont de la forme

$$\bar\mu(t) = Z_t\,\bar\mu_0(t) \quad \text{avec } \bar\mu_0(t) = \frac{\sum_{x=x_m}^{x_M}E_{x,t}\,\mu_0(x,t)}{\sum_{x=x_m}^{x_M}E_{x,t}}$$

d'où on tire $\mathbb{E}[\bar\mu(t)] = \bar\mu_0(t)$

$$\mathbb{V}[\bar\mu(t)] = \mathbb{V}[Z_t]\,\bar\mu_0^2(t)$$

$$\rightarrow \mathbb{V}[Z_t] = \frac{\mathbb{V}[\bar\mu(t)]}{\mathbb{E}[\bar\mu(t)]^2}$$

$$\Rightarrow \hat\sigma_Z^2 = \frac{\frac{1}{t_M-t_m+1}\sum_{t=t_m}^{t_M}\left(\hat\mu(t) - \frac{1}{t_M-t_m+1}\sum_{t=t_m}^{t_M}\hat\mu(t)\right)^2}{\left(\frac{1}{t_M-t_m+1}\sum_{t=t_m}^{t_M}\hat\mu(t)\right)^2}$$

avec $\hat\mu(t) = \frac{\sum_{x=x_m}^{x_M}E_{x,t}\,\hat\mu(x,t)}{\sum_{x=x_m}^{x_M}E_{x,t}}$ et $\hat\mu(x,t) = \frac{D_{x,t}}{E_{x,t}}$ l'estimateur de Hoem de la fonct° de hasard.

Il reste à estimer (α,β,k) solution des CPO:

Par * il vient:

$$\frac{\partial}{\partial\alpha_x}\ln(\mathcal{L}) = \sum_{t=t_m}^{t_M}\left[\frac{d}{\lambda_{x,t}} - \frac{d+a}{\lambda_{x,t}+b}\right]\lambda_{x,t} = 0$$

$$\frac{\partial}{\partial\beta_x}\ln(\mathcal{L}) = \sum_{t=t_m}^{t_M}\left[\frac{d}{\lambda_{x,t}} - \frac{d+a}{\lambda_{x,t}+b}\right]k_t\,\lambda_{x,t} = 0$$

$$\frac{\partial}{\partial k_t}\ln(\mathcal{L}) = \sum_{t=t_m}^{t_M}\left[\frac{d}{\lambda_{x,t}} - \frac{d+a}{\lambda_{x,t}+b}\right]\beta_x\,\lambda_{x,t} = 0$$

Ce système est non-linéaire.

3- Calcul des espérances de vie résiduelles prospectives.

Espérance de vie prospective : $e(x,t) = \sum_{i \geq 0} \prod_{j=0}^{i} \mathbb{E}[e^{-\mu_{x+j,t+j}}]$ ← *L'espérance de vie résiduelle prospective à un âge x et au temps t, notée e(x,t), est calculée en sommant sur tous les âges futurs i les produits des probabilités de survie depuis l'âge x+i.*

Dans notre cas $e(x,t)$ est de la forme :

$$e(x,t) = \sum_{i \geq 0} \prod_{j=0}^{i} \left[\frac{b}{b+\mu_0(x+j,t+j)}\right]^a$$

Car la transformée de Laplace d'une $\Gamma(a,b)$ s'écrit $\mathbb{E}[e^{-xZ_t}] = \left(\frac{b}{b+x}\right)^a$

Rappel : $\mathcal{L}(f)(p) = \int_0^{\infty} e^{-pt} f(t)\,dt = \mathbb{E}[e^{-X}]$

Et comme $\mathbb{E}[e^{-\mu(x,t)}] = \mathbb{E}[e^{-Z_t \mu_0(x,t)}]$

$$\Rightarrow \mathbb{E}[e^{-\mu(x,t)}] = \mathbb{E}[e^{-Z_t \mu_0(x,t)}] = \left(\frac{b}{b+\mu_0(x,t)}\right)^a = e^{-a \ln\left(1+\frac{\mu_0(x,t)}{b}\right)}$$

$\Rightarrow$ lorsque $b = a \to \infty$, $\mathbb{E}[e^{-\mu(x,t)}] \to e^{-\mu_0(x,t)}$

et on retrouve $e(x,t) = \sum_{i \geq 0} \prod_{j=0}^{i} e^{-\mu_0(x+j,t+j)}$

Newton Raphson

$$\frac{\partial}{\partial \theta} \ln(\mathcal{L}_\theta) = 0$$

Rappel : $f(x) = 0$

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} \qquad \text{car } f(x_{i+1}) = f(x_i) + (x_{i+1} - x_i) f'(x_i) + \dots$$

$$\theta_{i+1} = \theta_i - \left[\frac{\partial^2 \ln(\theta)}{\partial \theta_i \partial \theta_j}\right]^{-1} \frac{\partial \ln(\theta_i)}{\partial \theta}$$

On a vu que $p \in \{\alpha, \beta, k\}$, $\frac{\partial}{\partial p} \ln(\mathbb{P}(D=d)) = d \frac{1}{\lambda} \frac{\partial \lambda}{\partial p} - (d+a) \frac{1}{\lambda+b} \frac{\partial \lambda}{\partial p}$

$$= \left[\frac{d}{\lambda} - \frac{d+a}{\lambda+b}\right] \frac{\partial \lambda}{\partial p}$$

$$\Rightarrow p \in \{\alpha, \beta, k\}, \frac{\partial^2}{\partial p^2} \ln(\mathbb{P}(D=d)) = \left[-\frac{d}{\lambda^2} \times \frac{\partial \lambda}{\partial p} + \frac{d+a}{(\lambda+b)^2} \times \frac{\partial \lambda}{\partial p}\right] \frac{\partial \lambda}{\partial p} + \left[\frac{d}{\lambda} - \frac{d+a}{\lambda+b}\right] \frac{\partial^2 \lambda}{\partial p^2}$$

De plus $\frac{\partial \lambda}{\partial p} = \begin{cases} \lambda & \text{si } p = \alpha \\ k\lambda & \text{si } p = \beta \\ \beta\lambda & \text{si } p = k \end{cases} \Rightarrow \frac{\partial^2 \lambda}{\partial p^2} = \begin{cases} \lambda & \text{si } p = \alpha \\ k^2\lambda & \text{si } p = \beta \\ \beta^2\lambda & \text{si } p = k \end{cases}$

$$\frac{\partial^2}{\partial \alpha^2} \ln(\mathbb{P}(D=d)) = \left[-\frac{d}{\lambda^2} \times \lambda + \frac{d+a}{(\lambda+b)^2} \times \lambda\right] \times \lambda + \left[\frac{d}{\lambda} - \frac{d+a}{\lambda+b}\right] \times \lambda$$
$$= -\frac{(d+a)b\lambda}{(\lambda+b)^2} = -\frac{(d+a)b}{(\lambda+b)^2} \lambda$$

$$\frac{\partial^2}{\partial \beta^2} \ln(\mathbb{P}(D=d)) = \left[-\frac{d}{\lambda^2} \times k\lambda + \frac{d+a}{(\lambda+b)^2} \times k\lambda\right] k\lambda + \left[\frac{d}{\lambda} - \frac{d+a}{\lambda+b}\right] \times k^2\lambda$$
$$= -\frac{(d+a)bk^2\lambda}{(\lambda+b)^2} = -\frac{(d+a)b}{(\lambda+b)^2} k^2\lambda$$

$$\frac{\partial^2}{\partial k^2} \ln(\mathbb{P}(D=d)) = \left[-\frac{d}{\lambda^2} \times \beta\lambda + \frac{d+a}{(\lambda+b)^2} \times \beta\lambda\right] \beta\lambda + \left[\frac{d}{\lambda} - \frac{d+a}{\lambda+b}\right] \times \beta^2\lambda$$
$$= -\frac{(d+a)b\beta^2\lambda}{(\lambda+b)^2} = -\frac{(d+a)b}{(\lambda+b)^2} \beta^2\lambda$$

$$\Rightarrow \hat{\alpha}_x^{i+1} = \hat{\alpha}_x^i - \left[\frac{\partial^2}{\partial \alpha^2} \ln(\mathcal{L})\right]^{-1} \times \frac{\partial}{\partial \alpha} \ln(\mathcal{L})$$
$$= \hat{\alpha}_x^i - \left[-\sum_t \frac{(d_{x,t}+a)b}{(\lambda_{x,t}+b)^2} \lambda_{x,t}\right]^{-1} \times \sum_t \left[\frac{d_{x,t}}{\lambda_{x,t}} - \frac{d_{x,t}+a}{\lambda_{x,t}+b}\right] \lambda_{x,t}$$

$$\Rightarrow \hat{\beta}_x^{i+1} = \hat{\beta}_x^i - \left[\frac{\partial^2}{\partial \beta^2} \ln(\mathcal{L})\right]^{-1} \times \frac{\partial}{\partial \beta} \ln(\mathcal{L})$$

$$= \hat{\beta}_x^i - \left[- \sum_t \frac{(d_{x,t}+a)b}{(\lambda_{x,t}+b)^2} k_t^2 \lambda_{x,t}\right]^{-1} \times \sum_t \left[\frac{d_{x,t}}{\lambda_{x,t}} - \frac{(d_{x,t}+a)}{\lambda_{x,t}+b}\right] k_t \lambda_{x,t}$$

$$\Rightarrow \hat{k}_t^{i+1} = \hat{k}_t^i - \left[\frac{\partial^2}{\partial k^2} \ln(\mathcal{L})\right]^{-1} \times \frac{\partial}{\partial k} \ln(\mathcal{L})$$

$$= \hat{k}_t^i - \left[- \sum_x \frac{(d_{x,t}+a)b}{(\lambda_{x,t}+b)^2} \beta_x^2 \lambda_{x,t}\right]^{-1} \times \sum_x \left[\frac{d_{x,t}}{\lambda_{x,t}} - \frac{(d_{x,t}+a)}{\lambda_{x,t}+b}\right] \beta_x \lambda_{x,t}$$

$$\Longrightarrow \quad \hat{\alpha}_x^{i+1} = \hat{\alpha}_x^i -$$

$$\ln(\mathcal{L}) = \sum_{x=x_m}^{x_M} \sum_{r=r_m}^{r_M} \ln(P(D = d \mid a,b)) = \sum_{x=x_m}^{x_M} \sum_{r=r_m}^{r_M} \left[d_{x,r} \ln(\lambda_{x,r}) - (d_{x,r}+a)\ln(\lambda_{x,r}+b) \right]$$

relations de récurrence suivantes :

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

$$\hat{k}_t^{i+1} = \hat{k}_t^i - \frac{\sum_x \left(D_{xt} - L_{xt}\,\mathbf{exp}\left(\hat{\alpha}_x^{i+1} + \hat{\beta}_x^i \hat{k}_t^i\right)\right)\hat{\beta}_x^i}{-\sum_x \left(L_{xt}\,\mathbf{exp}\left(\hat{\alpha}_x^{i+1} + \hat{\beta}_x^i \hat{k}_t^i\right)\right)\left(\hat{\beta}_x^i\right)^2}$$

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

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

troduit la fonction $F(k) = \sum_{x=x_m}^{x_M} L_{xt}\,\mathbf{exp}(\hat{\alpha}_x + \hat{\beta}_x k) - \sum_{x=x_m}^{x_M} D_{xt}$,

$$F(\alpha) = \sum_r \left(D_{xr} - L_{xr} e^{\alpha + \hat{\beta}_x \hat{k}_r}\right)$$

$$\rightarrow F'(\alpha) = \sum -L_{xr} e^{\alpha + \hat{\beta}\hat{k}}$$

$$\rightarrow \hat{\alpha}_x^{i+1} = \hat{\alpha}_x^i - \frac{\sum D_{xr} - L_{xr} e^{\alpha + \hat{\beta}\hat{k}}}{-\sum L_{xr} e^{\alpha + \beta k}}$$

$$F(k) = \sum_r \left(D_{xr} - L_{xr} e^{\hat{\alpha} + \hat{\beta} k}\right)$$

$$\rightarrow F'(k) = -\sum$$

$$\frac{\partial}{\partial \alpha_x} \ln(\mathcal{L}) = \sum_{r=r_m}^{r_M} \left[\frac{d}{\lambda_{x,r}} - \frac{d+a}{\lambda_{x,r}+b} \right] \lambda_{x,r} = 0$$

$$\frac{\partial}{\partial \beta_x} \ln(\mathcal{L}) = \sum_{r=r_m}^{r_M} \left[\frac{d}{\lambda_{xr}} - \frac{d+a}{\lambda_{x,r}+b} \right] k_r \lambda_{x,r} = 0$$

$$\frac{\partial}{\partial k_r} \ln(\mathcal{L}) = \sum_{r=r_m}^{r_M} \left[\frac{d}{\lambda_{xr}} - \frac{d+a}{\lambda_{xr}+b} \right] \beta_x \lambda_{x,r} = 0$$

$$d_{x,r} \ln\left(e^{\alpha+\beta k}\right) - (d_{x,r}+a)\ln\left(e^{\alpha+\beta k}+b\right)$$

$$d_{x,r}\,(\alpha + \beta k)$$

$$\rightarrow \hat{\alpha}_x^{i+1} = \hat{\alpha}_x^i - \frac{\frac{\partial \ln(\mathcal{L})}{\partial \alpha}}{\frac{\partial^2 \ln(\mathcal{L})}{\partial \alpha^2}}$$

$$= \hat{\alpha}_x^i - \frac{\sum_r \left[\frac{d}{\lambda_{xr}} - \frac{d+a}{\lambda_{xr}+b} \right] \lambda_{xr}}{\sum \left[\frac{2d}{\lambda_{xr}} - \frac{(d+a)b}{(\lambda_{xr}+b)^2} \right] \lambda_{xr}}$$

$$\lambda_{xr} = E_{xr}\, e^{\alpha+\beta k} \quad \Longrightarrow \frac{\partial^2}{\partial \alpha^2} \ln(\mathcal{L}) = \sum \left[\left(\frac{d}{E_{xr} e^{\alpha+\beta k}} - \frac{-(d+a) E_{x,r} e^{\alpha+\beta k}}{(E_{xr} e^{\alpha+\beta k}+b)^2} \right] e^{\alpha+\beta k} + \left(\frac{d}{\lambda_{xr}} - \frac{d+a}{\lambda_{x,r}+b} \right) \lambda_{x,r}$$

$$\left(\frac{d}{E_{xr} e^{\alpha+\beta k}} - \frac{d+a}{E_{xr} e^{\alpha+\beta k}+b} \right) e^{\alpha+\beta k}$$

$$\sum \left[\left(\frac{d}{\lambda_{xr}} + \frac{(d+a)\lambda_{x,r}}{(\lambda_{x,r}+b)^2} \right) \lambda_{xr} + \left(\frac{d}{\lambda_{xr}} - \frac{d+a}{\lambda_{x,r}+b} \right) \lambda_{x,r} \right]$$

$$\frac{2d}{\lambda_{xr}} \lambda_{xr} + \frac{(d+a)\lambda_{x,r} - (d+a)(\lambda_{x,r}+b)}{(\lambda_{xr}+b)^2} \lambda_{xr}$$

$$\frac{2d}{\lambda_{xr}} \lambda_{xr} - \frac{(d+a)b}{(\lambda_{xr}+b)^2} \lambda_{xr}$$

$$\Longrightarrow \frac{\partial^2}{\partial \alpha^2} \ln\mathcal{L} = \sum_r \left[\frac{2d}{\lambda_{xr}} - \frac{(d+a)b}{(\lambda_{xr}+b)^2} \right] \lambda_{xr}$$

$$\frac{\partial \ln(\mathcal{L})}{\partial \beta} = \sum_r \left[\frac{d}{\lambda_{xr}} - \frac{(d+a)}{\lambda_{xr}+b} \right] k_r \lambda_{xr}$$

$$\rightarrow \frac{\partial^2 \ln(\mathcal{L})}{\partial \beta^2} = \sum_r \left[\frac{d}{k_r \lambda_{xr}} + \frac{(d+a) k_r \lambda_{x,r}}{(\lambda_{xr}+b)^2} \right] k_r \lambda_{x,r} + \left[\frac{d}{\lambda_{xr}} - \frac{(d+a)}{\lambda_{xr}+b} \right] k_r^2 \lambda_{x,r}$$

$$= \sum_r \left[\frac{d(1+k_r^2)}{\lambda_{x,r}} - \frac{(d+a) b k_r^2}{(\lambda_{xr}+b)^2} \right] \lambda_{xr}$$

$$\rightarrow \hat{\beta}_x^{i+1} = \hat{\beta}_x^i - \frac{\sum_r \left[\frac{d}{\lambda_{xr}} - \frac{(d+a)}{\lambda_{xr}+b} \right] k_r \lambda_{xr}}{\sum_r \left[\frac{d}{\lambda_{xr}}(1+k_r^2) - \frac{(d+a) b k_r^2}{(\lambda_{xr}+b)^2} \right] \lambda_{xr}}$$

$$\lambda = e^{\alpha+\beta k}$$

$$\frac{\partial \ln(\mathcal{L})}{\partial k} = \sum_x \left[\frac{d}{\lambda_{xr}} - \frac{(d+a)}{\lambda_{xr}+b} \right] \beta_x \lambda_{xr}$$

$$\rightarrow \frac{\partial^2 \ln(\mathcal{L})}{\partial k^2} = \sum_x \Big[$$

On va essayer d'estimer α, β et k

→ On a vu que la log-vraisemblance du modèle est de la forme (à une cte add. près) :

$$\ln(\mathcal{L}) = \sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \ln(P(D=d \mid a,b)) = \sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \left[d_{x,t} \ln(\lambda_{x,t}) - (d_{x,t}+a) \ln(\lambda_{x,t}+b) \right]$$

de plus on a $\lambda_{x,t} = E_{x,t}\, \mu_0(x,t) = E_{x,t}\, e^{\alpha_x + \beta_x k_t}$

Donc, conditionnellement à (a,b), la log-vraisemblance est de la forme (à une cte additive près) :

$$\ln(\mathcal{L}) = \sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \ln(P(D=d \mid a,b)) = \sum_{x=x_m}^{x_M} \sum_{t=t_m}^{t_M} \left[d_{x,t} \ln(\lambda_{x,t}) - (d_{x,t}+a) \ln(\lambda_{x,t}+b) \right]$$

→ Il s'agit donc de maximiser $\ln(\mathcal{L})$ en (α, β, k) sous la contrainte que $\sum_{x=x_m}^{x_M} \beta_x = 1$ et $\sum_{t=t_m}^{t_M} k_t = 0$.