Heterogeneity: measure integrating risk estimation in the case of a modeling of the observable factors

Application to the measurement of mortality in sub-Saharan Africa

AFIR 2011, Madrid

Frédéric PLANCHET *Partner Actuary*
fplanchet@winter-associes.fr

Aymric KAMEGA *Actuary*
akamega@winter-associes.fr

Introduction and Background

As part of the calculation of insurance liabilities, best estimate, it should take into account the heterogeneity of the portfolio.

Indeed, with a comprehensive approach too, you risk not having a detailed assessment on the significant sub-populations in the portfolio and see the table become inappropriate when the portfolio becomes distorted with time.

Whatever the sophistication of the model, taking into account the heterogeneity leads to segment the population into subpopulations experience.

Conversely, a too fine segmentation indicates a risk estimation and model risk (not mutualizable and potentially dangerous to the insurer).

A reflection then is essential for choose the optimal level of segmentation. Quantification of risk estimates associated with the table construction experience may be a useful criterion for this purpose.

CONTENTS

Heterogeneity: measure integrating risk estimation in the case of a modeling of the observable factors

1. Impact of a reduction in sample in the presence of heterogeneity

Presentation of the insured population

Data are available of population UEMOA (West African francophone, WAEMU). These data include data from three countries, presented in the table below (these are data on 30-55 years).

Man (insured population)	Population at risk	Average age	Average death rate	Average death rate (lower limit at 95%)	Average death rate (upper limit at 95%)
CI	549 656	43,9 years old	0,40%	0,38%	0,41%
ML	12 114	42,5 years old	0,22%	0,14%	0,31%
TG	133 779	43,2 years old	0,42%	0,39%	0,46%
UEMOA (CI-ML-TG)	695 549	43,8 years old	0,40%	0,38%	0,41%

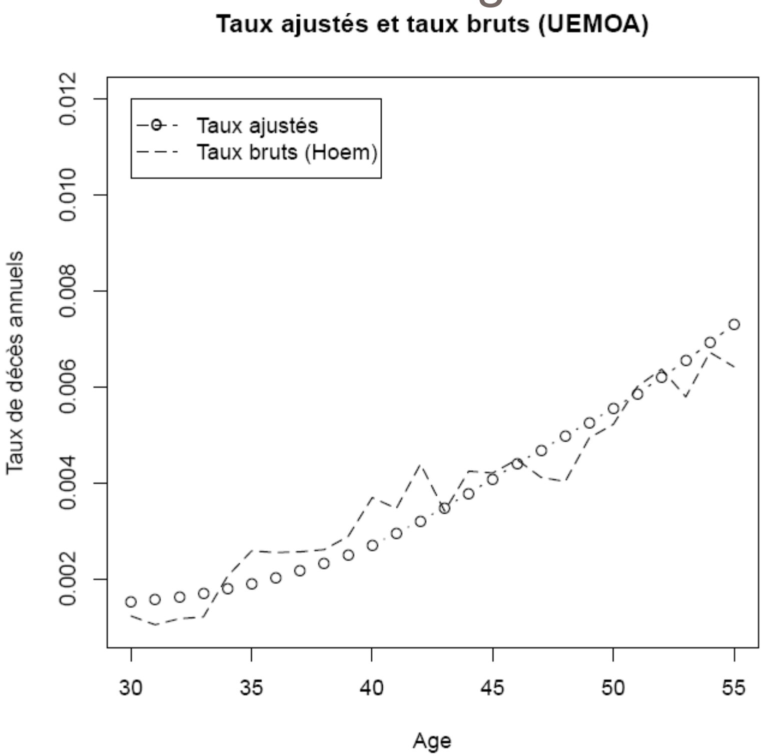
The objective here is then to present the evolution of systematic risk estimation when one moves from the population level to that of UEMOA (WAEMU) countries.

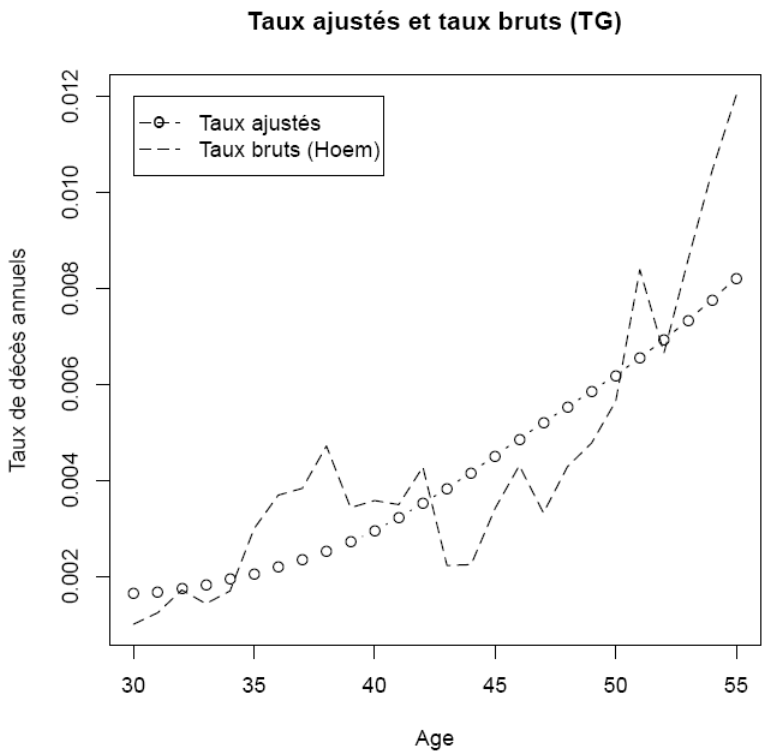
The countries chosen for this illustration is Togo (TG), intermediate country between Côte d'Ivoire (CI) and Mali (ML) in terms of risk exposure.

1. Impact of a reduction in sample in the presence of heterogeneity

Presentation of crude mortality rate and ajusted mortality rate for UEMOA and Togo

It appears on the graphic below that sampling fluctuations are consistently greater with the data of Togo.





1. Impact of a reduction in sample in the presence of heterogeneity

Measure of risk estimation on ajusted mortality rates

We consider a model of Brass, that is to say an external reference model as $y_x = a \times z_x + b + \varepsilon_x$, with $y_x = \ln\left(\hat{q}_x / \left(1 - \hat{q}_x\right)\right)$ and $z_x = \ln\left(q_x^{ref} / \left(1 - q_x^{ref}\right)\right)$.

We consider crude death rates estimed by Hoem, either: $Q_x \sim N\left(\hat{q}_x ; \sqrt{\frac{\hat{q}_x\left(1-\hat{q}_x\right)}{R_x}}\right)$.

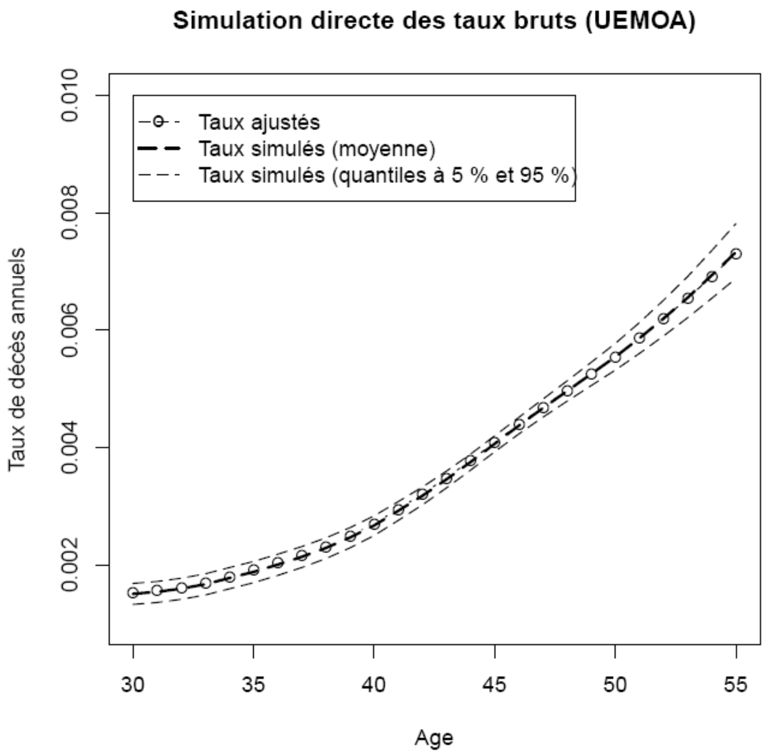
It then generates *k* simulations of crude death rates according to this law, and for each simulation *k*, determining an estimate of the parameters $\theta^k = \left(a^k, b^k\right)$.

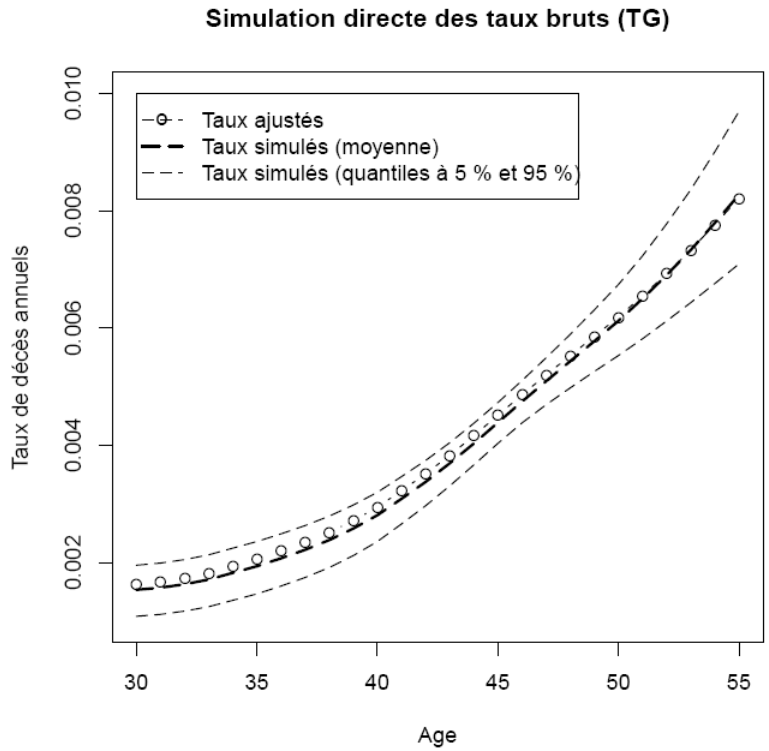
We deduce then *k* achievement rates simulated: $q_x\left(\hat{\theta}^k\right) = \frac{\exp\left(\hat{a}^k z_x + \hat{b}^k\right)}{1 + \exp\left(\hat{a}^k z_x + \hat{b}^k\right)}$.

1. Impact of a reduction in sample in the presence of heterogeneity

Measure of risk estimation on ajusted mortality rates

The graphs below show ajusted death rates and simulated death rates with data of UEMOA (WAEMU) and Togo. It appears that the dispersion of death rates from simulated data Togo is consistently more important.





CONTENTS

Heterogeneity: measure integrating risk estimation in the case of a modeling of the observable factors

2. Reduction of risk estimation in the presence of heterogeneity

Choice of model of heterogeneity

Three possible approaches are quoted here to model the heterogeneity:

- the first approach consists in modeling the behavior of each subpopulation in an independent way ;
- the second approach consists in turning to models integrating the observable factors of heterogeneity starting from explanatory variables ;
- the third approach consists in turning to models integrating the unobservable factors of heterogeneity.

2. Reduction of risk estimation in the presence of heterogeneity

Choice of model of heterogeneity (2nd approach)

In statistics, when a phenomenon can be explained by several explanatory variables, one can turn to purely parametric regressions such as linear regressions. The advantage is that, in this case, one can easily find consistent estimators. The disadvantage of these models is that they are based on many assumptions on the behavior of the phenomenon observed and thus present a significant risk of not being faithful to the experience.

An alternative then consists in turning to nonparametric regressions, which are based on a limited number of assumptions for the behavior of the observed phenomenon and, thus, are less constraining. However, these models present a well-known disadvantage under the term of “curse of dimensionality” (by mathematician Richard Bellman), relating to the problem caused by an exponential increase in volume associated with adding extra dimensions to a (mathematical) space. According to this curse, the nonparametric estimators of a function of regression behave badly when the number of variables is important.

2. Reduction of risk estimation in the presence of heterogeneity

Choix du modèle d'hétérogénéité (2ème démarche)

It is thus necessary to reduce the dimension of the models. The method which is considered here is the method adopted by Lopez [2007], it is the single index model :

$$m(z) = E(Y \mid Z = z) = f\left(\delta^T z\right).$$

In practice, it appears that the multiplicative Cox [1972] model and the additive Lin and Ying [1994] model are typical cases of SIM, in which the assumptions do not relate on conditional expectancy but they do to the conditional instantaneous hazard rate :

- Cox : $\lambda(t \mid Z = z) = \lambda_0(t) e^{\delta^T z}$

- Lin et Ying : $\lambda(t \mid Z = z) = \lambda_0(t) + \gamma^T z$

2. Reduction of risk estimation in the presence of heterogeneity

Cox model : adjustment

We initially determine the adjusted mortality rates for the subpopulation of the Côte d'Ivoire (basic subpopulation in our applications) starting from the Brass model with external reference. These ajusted mortality rates are noted : $q_{x,CI}\left(\hat{\theta}\right)$, with $\hat{\theta}=\left(\hat{a},\hat{b}\right)$.

To fit the Cox model, we can look to the Breslow approximation whose log-likelihood is:

$$L(\delta)=\sum_{i=1}^{D}\delta^{T}s_{(i)}-\sum_{i=1}^{D}d_{i}\times\ln\left(\sum_{j\in R_i}\exp\left\{\delta^{T}z_{j(i)}\right\}\right).$$

Next, we deduce the adjusted mortality rates of Mali and Togo, starting from the parameters of the Cox model by the following relations (by retaining the hypothesis that the rates of instantaneous hazard are constant between two entire ages):

$$q_{x,ML}\left(\hat{\theta};\hat{\delta}_{ML}\right)=1-\left(1-q_{x,CI}\left(\hat{\theta}\right)\right)^{\exp\left(\hat{\delta}_{ML}\right)} \qquad q_{x,TG}\left(\hat{\theta};\hat{\delta}_{TG}\right)=1-\left(1-q_{x,CI}\left(\hat{\theta}\right)\right)^{\exp\left(\hat{\delta}_{TG}\right)}$$

2. Reduction of risk estimation in the presence of heterogeneity

Lin and Ying model: adjustment

At first we determine the adjusted mortality rates for the sub-Ivorian population, as before.

Lin and Ying [1994] and Klein and Moeschberger [2005] show that the estimate of the coefficients of the model is:

$$\hat{\gamma} = \mathrm{A}^{-1}\mathrm{B}$$

where : $\mathrm{A} = \sum_{i=1}^{D} \sum_{j \in R_i} \left(z_{j(i)} - \bar{z}_{j(i)} \right)^T \left(z_{j(i)} - \bar{z}_{j(i)} \right)$, $\mathrm{B} = \sum_{i=1}^{D} \sum_{h} d_{i,h} \left(z_{(i),h} - \bar{z}_{(i)} \right)$ and $\bar{z}_{(i)} = \frac{1}{R_i} \sum_{j \in R_i} z_{j(i)}$

We deduce the adjusted mortality rates of Mali and Togo, starting from the parameters of the Lin and Ying model, by the following relationship (under the assumption that the instantaneous rates of hazard are constant between two entire ages):

$$q_{x,ML}\left(\hat{\theta}; \hat{\gamma}_{ML}\right) = 1 - \left(1 - q_{x,CI}\left(\hat{\theta}\right)\right) \mathbf{exp}\left(-\hat{\gamma}_{ML}\right) \qquad q_{x,TG}\left(\hat{\theta}; \hat{\gamma}_{TG}\right) = 1 - \left(1 - q_{x,CI}\left(\hat{\theta}\right)\right) \mathbf{exp}\left(-\hat{\gamma}_{TG}\right)$$

2. Reduction of risk estimation in the presence of heterogeneity

Comparison of adjustments et *backtesting* on heterogeneity

Global model UEMOA - H (Brass global)			
Country	Observed deaths	Predicted deaths	Différence
Côte d'Ivoire	2 188	2 203	0,7%
Mali	27	44	63,8%
Togo	565	511	-9,6%

Integrating heterogeneity models (without and with obs. fact.)			
Country (model)	Observed deaths	Predicted deaths	Différence
Côte d'Ivoire (Brass)	2 188	2 144	-2,0%
Mali (Brass(*))	27	29	8,4%
Mali (Cox)		26	-3,2%
Mali (Lin et Ying)		26	-4,9%
Togo (Brass)	565	565	-0,1%
Togo (Cox)		548	-3,0%
Togo (Lin et Ying)		550	-2,7%

(*) agreement with the treatment of crude death rate equal to zero.

The first sub-table presents the differences when the adjusted mortality rates are given in total without taking into account heterogeneity between subpopulations. The second sub-table presents the differences when the adjusted mortality rates are given from the models integrating heterogeneity, either from the independent models for each subpopulation (as with the Brass model), or from the models integrating heterogeneity from observable factors.

2. Reduction of risk estimation in the presence of heterogeneity

Cox model: simulation

For Côte d'Ivoire (CI), for each simulation : $q_{x,CI}\left(\hat{\theta}^{k}\right)=\dfrac{\mathbf{exp}\left(\hat{a}^{k} z_{x}+\hat{b}^{k}\right)}{1+\mathbf{exp}\left(\hat{a}^{k} z_{x}+\hat{b}^{k}\right)}$

In the case of Mali (ML) and Togo (TG), adjusted mortality rates were obtained from the Cox model. For each *k* simulation we estimated the parameters with the Breslow method, in which case the log-likelihood reads:

$$L\left(\delta^{k}\right)=\sum_{i=1}^{D}\left(\delta^{k}\right)^{T} s_{(i)}^{k}-\sum_{i=1}^{D}\hat{d}_{i}^{k}\times\mathbf{ln}\left(\sum_{j\in R_i}\mathbf{exp}\left\{\left(\delta^{k}\right)^{T} z_{j(i)}\right\}\right),$$

where $\hat{d}_{i}^{k}=\sum_{h}\hat{d}_{i,h}^{k}$ and $\delta^{k}=\left(\delta_{ML}^{k};\delta_{TG}^{k}\right)$.

We deduce for each simulation *k* adjusted mortality rates of Mali and Togo by:

$$q_{x,ML}\left(\hat{\theta}^{k};\hat{\delta}_{ML}^{k}\right)=1-\left(1-q_{x,CI}^{k}\left(\hat{\theta}^{k}\right)\right)^{\mathbf{exp}\left(\hat{\delta}_{ML}^{k}\right)} \qquad q_{x,TG}\left(\hat{\theta}^{k};\hat{\delta}_{TG}^{k}\right)=1-\left(1-q_{x,CI}^{k}\left(\hat{\theta}^{k}\right)\right)^{\mathbf{exp}\left(\hat{\delta}_{TG}^{k}\right)}$$

2. Reduction of risk estimation in the presence of heterogeneity

Lin and Ying model : simulation

The first steps in the simulation of the mortality rates for the Lin and Ying model are identical to those are carried out for simulations of the Cox model. We have therefore simulated the number of deaths by country $\hat{d}^{k}_{x,h}$ and simulated mortality rates for Côte d'Ivoire $q_{x,CI}\left(\hat{\theta}^{k}\right)$.

To détermine the mortality rates of Mali and Togo for each simulation *k*, we are based on the Lin and Ying model and we estimate $\gamma^{k}=\left(\gamma^{k}_{ML};\gamma^{k}_{TG}\right)$

For this purpose, we use the relation : $\hat{\gamma}=\mathrm{A}^{-1}\mathrm{B}^{k}$, where $\mathrm{B}^{k}=\sum_{i=1}^{D}\sum_{h}\hat{d}^{k}_{i,h}\left(z_{(i),h}-\bar{z}(i)\right)$.

We deduce for each simulation *k* adjusted mortality rates of Mali and Togo by:

$$q_{x,ML}\left(\hat{\theta}^{k};\hat{\gamma}^{k}_{ML}\right)=1-\left(1-q^{k}_{x,CI}\left(\hat{\theta}^{k}\right)\right)\mathbf{exp}\left(-\hat{\gamma}^{k}_{ML}\right) \qquad q_{x,TG}\left(\hat{\theta}^{k};\hat{\gamma}^{k}_{TG}\right)=1-\left(1-q^{k}_{x,CI}\left(\hat{\theta}^{k}\right)\right)\mathbf{exp}\left(-\hat{\gamma}^{k}_{TG}\right)$$

2. Reduction of risk estimation in the presence of heterogeneity

Evolution of risk estimation by model (liabilities)

Statistic	Brass		Cox		Lin et Ying	
	Liab. and adjusted rate	Liab. and simulated rate	Liab. and adjusted rate	Liab. and simulated rate	Liab. and adjusted rate	Liab. and simulated rate
Mean	$4,18.10^{-2}$	$4,03.10^{-2}$	$4,06.10^{-2}$	$4,01.10^{-2}$	$4,22.10^{-2}$	$4,17.10^{-2}$
Quantile at 0,5 %	NA	$3,18.10^{-2}$	NA	$3,52.10^{-2}$	NA	$3,54.10^{-2}$
Quantile at 5 %	NA	$3,52.10^{-2}$	NA	$3,70.10^{-2}$	NA	$3,74.10^{-2}$
Quantile at 95 %	NA	$4,50.10^{-2}$	NA	$4,34.10^{-2}$	NA	$4,61.10^{-2}$
Quantile a 99,5 %	NA	$4,74.10^{-2}$	NA	$4,50.10^{-2}$	NA	$4,83.10^{-2}$

It arises that with the data of Togo, accounting for risk estimation decreases the calculated liability by 3.6 % when the Brass model is retained, where as the impact is weaker with the Cox model and Lin and Ying model (it falls to 1.2 %).

In our example, it appears however that the weight of the risk estimation in the evaluation of a liability is comparable to the weight of the risk of model.

Conclusion and outlook

In the presence of heterogeneity, we are left with small samples and special attention must be given to measuring the uncertainty of the estimators.

A general framework to make this measurement is presented here from a resampling method of crude mortality rates. This measurement tool shows that modeling the heterogeneous models from independent sub-populations is not appropriate.

It should then retain models directly integrating heterogeneity of observable factors.

Theses models has several advantages, among which we emphasize the capacity to model the mortality rates when the data are noticeably limited.

The choice of theses models presents however some disadvantages, like the general potential impact in terms of model risk.

References

Cao H. [2005] *A comparison Between the Additive and Multiplicative Risk Models*, Mémoire d'actuariat, Université Laval.

Cox D. R. [1972] « Regression Models and Life-Tables », *Journal of the Royal Society. Series B (Methodological), Vol. 34, No. 2*.

Droesbeke J.-J., Fichet B., Tassi P. [1989] (éditeurs) « Analyse statistique des durées de vie », *Economica*.

Hill C., Com-Nougué C., Kramar A., Moreau T., O'Quigley J., Senoussi R., Chastang C. [1990] « Analyse statistique des données de survie », *Inserm, Médecine – Sciences, Flammarion*.

Kamega A., Planchet F. [2010], « Mesure du risque d'estimation associé à une table d'expérience », *Cahiers de recherche de l'ISFA, WP2136.*

Kamega A., Planchet F. [2011a], « Analyse et comparaison des populations générale et assurée en Afrique subsaharienne francophone pour anticiper la mortalité future », *Cahiers de recherche de l'ISFA, WP 2138.*

Kamega A., Planchet F. [2011b], « Hétérogénéité : mesure du risque d'estimation dans le cas d'une modélisation intégrant des facteurs observables », *Bulletin Français d'Actuariat, Vol. 11, No 21.*

Klein J. P., Moeschberger M. L. [2005] « Survival Analysis – Techniques for Censored and Truncated Data », *Springer, 2nd edition*.

Kone M.G.K. [2007] « État des lieux de l'assurance vie dans les pays de la CIMA », *CIMA*.

Lin D. Y., Ying Z. [1994] « Semiparametric analysis of the additive risk model », *Biometrika, n. 81*.

Lopez O. [2007] *Réduction de dimension en présence de données censurées*, Thèse de doctorat, Université de Rennes 1.

Martinussen T., Scheike T. H. [2006] « Dynamic Regression Models for Survival Data », *Springer*.

Planchet F. [2006], « Tables de mortalité d'expérience pour les portefeuilles de rentiers – (Tables TGH 05 et TGF 05) – Notice de présentation », *Institut des actuaires.*

Planchet F., Kamega A., Ziguélé M. [2010], « Confection des tables de mortalité réglementaires pour la zone CIMA », *CIMA/WINTER & Associés (confidentiel).*

Planchet F., Leroy G. [2009] « Quel niveau de segmentation pertinent ? », *La Tribune de l'Assurance, n. 142*.

Planchet F., Thérond P. [2006] « Modèles de Durée - Applications actuarielles », *Economica*.

Vaupel J. W., Manton K. G. Stallard E. [1979] « The impact of heterogeneity in individual frailty on the dynamics of mortality », *Demography, Vol. 16, No. 3*.

Contacts

Frédéric Planchet

fplanchet@ressources-actuarielles.net

Aymric Kamega

akamega@ressources-actuarielles.net

Institut de Science Financière et d'Assurance (ISFA)

50 avenue T. Garnier
69007 LYON (FRANCE)

http://www.ressources-actuarielles.net/
http://blog.ressources-actuarielles.net/
http://afrique.ressources-actuarielles.net/