



**HAL**  
open science

## Stratégie optimale de réduction de l'intervalle de confiance pour l'estimateur de la prime ajustée.

### Application en assurance automobile

Kmar Fersi, Kamel Boukhetala, Samir Ben Ammou

► **To cite this version:**

Kmar Fersi, Kamel Boukhetala, Samir Ben Ammou. Stratégie optimale de réduction de l'intervalle de confiance pour l'estimateur de la prime ajustée. Application en assurance automobile. 2011. hal-00625684v2

**HAL Id: hal-00625684**

**<https://hal.science/hal-00625684v2>**

Submitted on 26 Feb 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Stratégie optimale de réduction de l'intervalle de confiance pour l'estimateur de la prime ajustée. Application en assurance automobile

Kmar Fersi\*, Kamel Boukhetala\*\*  
Samir Ben Ammou\*\*\*

\*Institut Supérieur de Gestion Sousse, & Computational Mathematics Laboratory,  
Route de Kairouan, 5019 Monastir, Tunisia  
fersi.kmar@gmail.com

\*\* Faculté de Mathématiques Bp. 32, El-Alia,  
Bab-Ezzouar, USTHB Alger, Algeria

\*\*\* Computational Mathematics Laboratory, Route de Kairouan, 5019 Monastir, Tunisia

**Abstract.** The estimator of the adjusted premium developed by Necir and Boukhetala (2004) is considered. To minimize the variance of this estimator we proposed an original strategy by formulating a non-linear optimization problem with stochastic constraints. For finding global optimal solutions we used an hybrid genetic algorithm. An illustration of this strategy on real data is developed. This paper provides insurers with a key pricing tool in non-life insurance.

## 1 Introduction

Les risques extrêmes sont, aujourd'hui, caractérisés par le fait qu'ils sont abstraits et complexes (Picard (2002) et Chemarin, Henry & Michel-Kerjan (2005)). Partrat et Besson (2005) ont montré que les événements naturels ou les risques industriels n'ont pas le monopole des sinistres extrêmes. Ceux-ci sont aussi présents, relativement, dans les branches RC automobile qui sont touchées par de très graves sinistres corporels. Donc les assureurs se retrouvent devant une situation où coexiste une sous population d'individus à haut risque et une large majorité à bas risque. Dans ce cas, on propose une différenciation tarifaire efficace et adéquate qui permet d'attirer à la fois hauts risques et bas risques. L'objectif de tarification équité peut être atteint, tout en préservant l'efficacité des mécanismes de l'estimation de la prime des risques extrêmes.

Dans ce travail, on s'intéresse à la recherche d'une stratégie d'optimalité visant la réduction de l'intervalle de confiance de cet estimateur. Cette optimalité est obtenue par l'utilisation d'algorithme approprié à la résolution du problème par la minimisation d'un critère lié à un ensemble de contraintes qui caractérisent les risques extrêmes.

Cette approche permet de contribuer à la réalisation des objectifs stratégiques d'optimalité et d'adaptation de la tarification "a posteriori" aux exigences de l'assureur et l'assuré. La démarche de calcul du critère d'optimalité contribue aussi à la mise en place d'un reporting

efficace et également au respect de la conformité des lois et réglementations en vigueur. Cet article se présente de la manière suivante: dans une première section, on traite l'utilité de la différenciation tarifaire en matière d'assurabilité des risques extrêmes. En deuxième section, on calcule l'estimateur de la prime ajustée proposée par Necir & Boukhetala (2004). En troisième section, on définit le problème général de la réduction de l'étendu de l'intervalle de confiance de l'estimateur de la prime ajustée sous forme d'un programme d'optimisation non linéaire avec contraintes stochastiques. Le caractère aléatoire complexe de ce problème nous emmène à adapter, dans une quatrième section, un algorithme hybride et efficace pour la recherche de solutions optimales, statistiquement exploitables. Enfin, on propose l'ébauche d'un modèle de calcul d'une prime ajustée raisonnable et nécessaire à la couverture des risques extrêmes en assurance non vie (assurance RC-Automobile). Cette démarche de résolution est ensuite analysée en introduisant les coûts des sinistres extrêmes obtenus par l'application d'une approche appropriée issue de la théorie des valeurs extrêmes. Enfin, on propose une stratégie de calcul d'une prime estimée par intervalle de confiance assez réduit.

## **2 Sinistralités extrêmes et tarification a posteriori**

La présence des sinistres extrêmes perturbe l'homogénéité du portefeuille et leurs détection est donc essentielle pour éviter les erreurs de tarification et d'interprétation qui peuvent avoir pour conséquences des modifications tarifaires utiles ou à l'inverse inutiles. L'intuition en est que la tarification a posteriori qui lie la prime payée à la sinistralité passée de l'assuré a une double justification. D'une part, elle tient compte de la différenciation tarifaire entre les sinistres extrêmes et les sinistres standards. Et d'autre part, elle présente un mécanisme d'incitation à la prudence.

### **2.1 Risques extrêmes et iniquité de la tarification**

L'hétérogénéité au sein d'un portefeuille d'assurance pose un grand nombre de problèmes, en particulier de tarification. Lorsque la même prime est appliquée à l'ensemble du portefeuille, les mauvais risques s'assureront, mais les bons pourraient être découragés par le tarif excessif; iniquité de la tarification.

La raison en est que les individus à bas risque ne peuvent pas faire leurs choix d'assurance correspondants à la nature du risque, dans la mesure où les contrats qui leur conviendraient seraient également achetés par les individus à haut risque et se trouveraient ainsi financièrement déséquilibrés.

Pour résoudre ce problème, les assureurs seraient conduits à identifier une stratégie de tarification permettant de concilier les principes d'efficacité et d'équité.

*sinistres graves en assurance non vie:*

Le classement des risques selon leurs fréquence et leur gravité trace des frontières entre deux catégories de sinistres;

- Les bas risques très fréquents et de petites ampleurs
- Les hauts risques rares et de coûts très élevés

En effet, les outils statistiques traditionnels qui étudient un phénomène à travers son comportement normal, caractérisent mal ces extrêmes. Elles fournissent uniquement des indications sur la distribution générale d'un phénomène. McNeil & al. (1997) et Rootzen & Tajvidi (1997) montrent que l'application de la théorie des valeurs extrêmes donne un cadre plus systématique, rigoureux et cohérent que les méthodes classiques.

En effet, cette méthode permet de déduire la distribution des extrêmes à partir de la distribution totale, sans traiter séparément les extrêmes des risques récurrents. Ces événements appartiennent à ce que l'on appelle "la queue de distribution", ce qui veut dire que l'on dispose de très peu d'observations pour déterminer analytiquement les distributions de probabilité. Donc il est important de bien modéliser les queues de distribution afin d'éviter une tarification inéquité.

Généralement, l'activité d'assurance est basée sur la loi des grands nombres qui nous assure que la moyenne empirique du coût de sinistres indépendants et identiquement distribués tend vers le coût moyen pour un nombre de polices importants. La loi de grands nombres peut ne pas fonctionner correctement en cas des risques hétérogènes, et donc non identiquement distribués. L'assureur doit, dès lors, réaliser des compensations dans le temps, dans l'espace, et limiter ses souscriptions de manière à conserver un système de tarification homogène.

En pratique, il n'est pas aisé d'obtenir ces compensations, ni de limiter sa souscription. Donc la différenciation tarifaire peut offrir une solution attrayante à ces problèmes.

#### *Différenciation tarifaire et incitation à la prévention:*

En assurance, notamment en non vie, le système actuel est basé sur le principe de solidarité où chaque assuré paie le même taux de prime et les assureurs indemnisent le montant des dommages à leurs assurés (François & Constant (2009)). Ce système permet donc une couverture complète des risques assurés. Dans ce cas, l'intérêt immédiat de l'assuré est de ne faire aucun effort et l'assureur est donc contraint d'établir une prime actuarielle correspondante au risque élevé afin d'éviter des situations de ruine (Kunreuther & al. (1995)).

La différenciation tarifaire a posteriori n'est rien d'autre que la recherche d'une tarification plus juste traduisant la réalité des risques. Elle s'interprète comme une distorsion de la structure de la prime des risques par rapport à la structure des coûts. En première lieu, des tarifs différenciés peuvent être appliqués car les risques représentent des coûts différents. En second lieu, des tarifs plus au moins élevés peuvent traduire des différences de pouvoir de marché que les assureurs sont susceptibles d'exercer à l'égard de différentes catégories de clientèles.

Cependant, le critère uniforme de la prime d'assurance crée une distorsion qui lui induit une moindre incitation à la prévention et génère un surcoût pour la société. En effet, l'utilisation d'une différenciation tarifaire peut inciter les assurés à adopter un comportement prudent et donc diminuer les charges de l'assureur qui peut ainsi pratiquer des primes plus faibles (voir par exemple Wills & Sherris (2010)). D'où, la pratique de la prime des sinistres graves permet d'opérer une différenciation des produits par le biais d'un jeu de niveaux de couverture différents. Autrement dit, elle suggère de discriminer entre les assurés au moyen d'une tarification non linéaire reliant la prime et la quantité d'assurance achetée. Ce qui nous permet de constater que là crée une bonne incitation transmise aux assurés par l'intermédiaire des as-

Prime des risques extrêmes

sureurs.

## 2.2 stratégie de tarification pour les risques majeurs

L'avenir des assurances dépend-il des risques extrêmes? A l'évidence oui. L'augmentation du coût cumulé des grands sinistres est plus importante depuis 20 ans que les quarante années précédentes. C'est donc l'intérêt des assureurs et des réassureurs de recourir à des modèles scientifiques pour élaborer un système de tarification adéquat et efficace. S'inspirant du principe de Wang (1996) et Lin (2005), Necir & Boukhetala (2004) ont introduit une nouvelle approche du calcul de prime des risques extrêmes. Cette approche est mise en œuvre sous forme d'un logiciel permettant un calcul automatique de l'estimateur de la prime ajustée par intervalle de confiance.

*Exploration des sinistres extrêmes:*

L'utilisation de la théorie des extrêmes permet de répartir la population des risques en deux catégories: les hauts risques et les bas risques. Les marchés étant concurrentiels, les primes d'assurance demandées pour couvrir le risque en question devront logiquement différer entre les deux catégories: les hauts risques payent une prime supérieure aux bas risques pour un même niveau de couverture.

Dans la tarification à posteriori, l'idée est de détecter les sinistres extrêmes qui perturbent l'homogénéité d'un portefeuille d'assurance. Cependant, plusieurs auteurs ont montré que la méthode de dépassement du seuil (Peak Over Threshold (POT)) issue de la théorie des extrêmes fournit les outils adéquats pour détecter et traiter les sinistres graves. En effet, la mise en œuvre et l'utilisation de cette approche pose le problème de détermination du seuil. Théoriquement, il existe plusieurs techniques de détermination du seuil au-delà duquel les sinistres sont qualifiés extrêmes. En pratique, la méthode graphique a illustré des possibilités d'exploration tout à fait adéquate.

Pour améliorer cette politique de tarification, il est nécessaire de bien caractériser le comportement asymptotique des sinistres au-delà d'un seuil fixe.

Soit  $X_1, X_2, \dots, X_n$  une suite de variable aléatoires de risque, indépendantes et identiquement distribuées avec une fonction de répartition commune  $F(x) = P(X \leq x)$ ,  $x \in R$ . Pickands (1975) a introduit la loi de Pareto généralisée (GPD) pour analyser le comportement asymptotique des excès qui sont définis par  $y_i = x_i - u/X > u$ .

Le théorème de Pickands affirme que la fonction de distribution GPD est:

$$G_{\xi, \beta}(y) = \begin{cases} 1 - (1 + \frac{\xi y}{\beta})^{-\frac{1}{\xi}} & \text{si } \xi \neq 0 \\ 1 - \exp(-\frac{y}{\beta}) & \text{si } \xi = 0 \end{cases}$$

avec  $y \in [0, x_F - u]$  si  $\xi \geq 0$ ,  $y \in [0, \text{Min}(-\beta/\xi, x_F - u)]$ , si  $\xi < 0$ , et si  $\xi \rightarrow 0$ ,  $G_{\xi, \beta}(y)$  est approximée par la distribution exponentielle avec  $\beta$  correspond à la moyenne.

Afin d'estimer les paramètres  $\beta$  et  $\xi$  de la distribution  $G_{\xi, \beta}(y)$ , on détermine la fonction vraisemblance comme:

$$l(\beta, \xi) = -n \log \beta - (1 - 1/\xi) \sum_{i=1}^k \log(1 - \xi y_i / \xi),$$

Les estimateurs du maximum de vraisemblance de  $\beta$  et  $\xi$  sont déterminés par les solutions du système d'équations suivant:

$$\frac{1}{n} \sum_{i=1}^n \log\left(1 + \xi \frac{x_i}{\beta}\right) = \xi \quad (1)$$

$$\frac{1}{n} \sum_{i=1}^n \frac{x_i / \beta}{1 + \xi x_i / \beta} = \frac{1}{1 + \xi} \quad (2)$$

Pour plus de détails voir Hosking & Wallis (1978), Embrechets, Klüpperberg & Mikosh (1997). Ces équations sont également utilisées comme contraintes dans la stratégie alternative de minimisation de l'étendu de l'intervalle de confiance de l'estimateur de la prime ajustée développée selon le principe de Wang. Dans le paragraphe suivant, on présente en détail une formulation mathématique de ce problème d'optimisation en se focalisant sur la pertinence de l'algorithme génétique hybride pour la recherche simultanée des solutions optimales des paramètres  $\xi$ ,  $\beta$  et le coefficient de distorsion de Wang  $p$ . Ce qui donne à l'utilisateur l'espoir d'améliorer la performance des techniques de détermination de ces paramètres selon la nature du risque observé et constitue également une démarche rationnelle pour le calcul des primes de risque extrême.

*Principe de Wang de la prime ajustée:*

Pour une tarification raisonnable des risques extrêmes, Wang (1996) propose un nouveau procédé de calcul de primes en appliquant une transformation à la fonction de survie. Il obtient alors la prime ajustée au risque, suivante:

$$\pi(X) = \int_0^{\infty} (1 - F(x))^{1/p} dx.$$

Le paramètre  $p \geq 1$  est appelé "coefficient de distorsion" ou "coefficient d'aversion au risque". Donc la prime ajustée pour un seuil élevé  $u > 0$  est donnée par:

$$\pi_u(X) = \int_u^{\infty} (1 - F(x))^{1/p} dx \quad (3)$$

$$= \int_u^{\infty} (\bar{F}(x))^{1/p} dx \quad (4)$$

Il faut noter que ce principe est celui de la prime pure appliqué à la distribution d'un risque rendu "plus dangereux". En effet, l'estimation de ces mesures, à l'aide d'une sinistralité historique adaptée, mettra en jeu des approches paramétriques puis semi-paramétriques.

*Estimation paramétrique de la prime ajustée:*

## Prime des risques extrêmes

Soit  $X_1, \dots, X_n$  une suite de risques indépendants et identiquement distribués, de distribution commune  $F$ , telle que

$$\bar{F}(x) = cx^{-\frac{1}{\xi}}(1 + x^{-\beta}L(x)) \quad \text{quand } x \rightarrow \infty \quad (5)$$

pour  $\xi \in [0, 1]$ ,  $\beta > 0$ , une constante  $c$  et  $L(\cdot)$  fonction à variation lente, on peut écrire la prime ajustée qu'est définie dans l'équation (3), comme:

$$\pi_u(X) = \int_u^\infty (1 - F(y+u))^{1/p} dy,$$

on peut montrer que

$$\bar{F}(y+u) = \bar{F}(u)\bar{F}_u(y),$$

et on déduit

$$\pi_u(X) = \int_u^\infty (\bar{F}(u)\bar{F}_u(y))^{1/p} dy.$$

Donc, l'estimateur de la prime ajustée s'écrit:

$$\hat{\pi}_u(X) = \int_u^\infty (\hat{F}(u)\hat{F}_u(y))^{1/p} dy.$$

Un estimateur naturel de  $\bar{F}(u)$  est donné en fonction de la distribution empirique par:

$$\hat{\lambda} = \hat{F}(u) = \bar{F}_n(u) = \frac{1}{n} \sum_{i=1}^n 1_{X_i > u} = \frac{N_u}{n},$$

où  $N_u$  est le nombre des excès au-delà du seuil  $u$ . D'autre part, le théorème de Pickands (1975) montre que la queue des excès est approximée par la distribution de Pareto généralisée

$$\hat{F}_u(y) \approx \bar{G}_{\hat{\xi}, \hat{\beta}}(y).$$

Alors

$$\hat{\pi}_u(X) = \hat{\lambda}^{\frac{1}{p}} \int_0^\infty (\bar{G}_{\hat{\xi}, \hat{\beta}}(y))^{1/p} dy.$$

Ce qui revient à écrire,

$$\hat{\pi}_u(X) = \hat{\lambda}^{\frac{1}{p}} \int_0^\infty [(1 + \xi \frac{y}{\hat{\beta}})^{-\frac{1}{\xi}}]^{1/p} dy$$

On peut écrire cet estimateur sous forme d'une espérance

$$\hat{\pi}_u(X) = \hat{\lambda}^{1/p} \int_0^\infty \frac{1}{\hat{\beta}} \left(1 + \hat{\xi} \frac{y}{\hat{\beta}}\right)^{\frac{-1}{\hat{\xi}} - 1} \hat{\beta} \left(1 + \hat{\xi} \frac{y}{\hat{\beta}}\right)^{\frac{-1}{\hat{\xi}} (\frac{1}{p} - 1) + 1} dy$$

D'où

$$\hat{\pi}_u(X) = \hat{\beta} \hat{\lambda}^{1/p} E\left(\left(1 + \hat{\xi} \frac{Y}{\hat{\beta}}\right)^{-r}\right),$$

où  $Y \sim GPD(\hat{\xi}, \hat{\beta})$  et  $r = \frac{1}{\hat{\xi}}(\frac{1}{p} - 1) - 1$ .

Sous la condition  $r > -\frac{1}{\hat{\xi}}$ , on a:

$$E\left(\left(1 + \hat{\xi} \frac{Y}{\hat{\beta}}\right)^{-r}\right) = \frac{1}{1 + \hat{\xi} r}$$

Pour  $\hat{\xi} \leq \frac{1}{p} \leq 1$ , l'estimateur de la prime ajustée peut être déterminé par

$$\hat{\pi}_u(X) = \hat{\lambda}^{\frac{1}{p}} \frac{\hat{\beta}}{\frac{1}{p} - \hat{\xi}}. \quad (6)$$

Le théorème suivant décrit la normalité asymptotique de  $\hat{\pi}_u(X)$

**Théorème (Necir & Boukhetala (2004))**

Supposons que (4) est vraie pour  $\xi \in [1/2, 1]$  et  $\delta > 0$ , si  $x \rightarrow x^{-\delta} L(x)$  est une fonction décroissante. Soit  $u = u_n$  tel que  $u \rightarrow \infty$  et  $\sqrt{nu^{-\delta}} L(u) \rightarrow 0$  quand  $n \rightarrow \infty$ . Alors

$$\frac{\sqrt{n}(\hat{\pi}_u(X) - \pi_u(X))}{\sqrt{\lambda^{2/p-1}(1-\lambda)}} \rightarrow \mathcal{N}(0, \sigma_\pi^2), \text{ quand } n \rightarrow \infty.$$

où

$$\sigma_\pi^2 = \frac{(1+\xi)\beta^2}{(\frac{1}{p}-\xi)^2} \left( \frac{(1/p-1)^2}{1+\xi} + \frac{(1+\xi)}{(\frac{1}{p}-\xi)^2} - \frac{2}{(\frac{1}{p}-\xi)} + 2 \right) \quad (7)$$

Après avoir détaillé la construction de notre estimateur, on présente maintenant l'algorithme de calcul de cet estimateur:

**Algorithme de calcul direct de la prime ajustée**

1. Choisir le nombre des excès  $N$  d'un échantillon  $(x_1, \dots, x_n)$ .
2. Estimer les paramètres  $(\xi, \beta)$  de la loi GPD à partir de l'échantillon des excès.
3. Calculer l'estimateur de la prime ajustée en fonction de  $p$  en utilisant la formule suivante:

$$\hat{\pi}_u(X) = \hat{\lambda}^{\frac{1}{p}} \frac{\hat{\beta}}{\frac{1}{p} - \hat{\xi}}$$

La Théorie des Valeurs Extrêmes permet d'effectuer des interpolations dans la queue des distributions expérimentales, mais les extrapolations apparaissent discutables. Aussi doit-on se



méfier des estimations effectuées  $(\hat{\xi}, \hat{\beta})$ , en cas d'insuffisance d'information, notamment quand elles concernent des problématiques relatives à la sécurité des personnes et des biens, même si des intervalles de confiance peuvent être calculés par inversion de la matrice de Fisher.

L'ajustement par la méthode du maximum de vraisemblance de la loi généralisée de Pareto (GPD) peut s'effectuer de manière très efficace et précise au moyen d'un outil d'optimisation globale. Donc on propose une stratégie d'optimalité visant la recherche des valeurs optimales des paramètres qui minimise la variance  $\sigma_\pi^2$  (équation 6) sous contraintes dérivées de la fonction de vraisemblance (équations 1 et 2). Cette stratégie, qui permet d'estimer conjointement les valeurs optimales de ces paramètres, fait l'objet de la démarche présentée dans le paragraphe suivant. Elle permet également de déterminer des intervalles de confiance plus réduits pour l'estimateur de la prime ajustée et fournit à l'utilisateur un outil d'aide à la tarification en assurance.

### 3 Approche alternative de calcul de la prime ajustée

L'objectif de tarification équité peut être atteint tout en préservant l'efficacité des mécanismes de l'estimation de la prime des risques extrêmes. Donc on propose une stratégie de réduction de l'intervalle de confiance de cet estimateur. Cette stratégie vise à minimiser la variance de l'étendu  $\sigma_\pi^2$  sous contrainte d'égalités et d'inégalités. Cette démarche donne un outil souple et interactif pour la conception d'un système d'aide à la tarification pour des situations de risque exceptionnel.

#### 3.1 Formulation mathématique

Notre objectif est de minimiser la variance de l'étendu de l'intervalle de confiance de l'estimateur de la prime ajustée (équation 6) qui se présente sous forme d'une fonction non linéaire, en tenant compte de certaines contraintes de type égalités et inégalités. Les fonctions des contraintes d'égalités sont propres aux paramètres de la loi des excédants des sinistres. Ces fonctions se sont présentées par le système d'équation (équation 1 et 2), qui est déterminé à l'aide de la méthode du maximum de vraisemblance, alors que les contraintes d'inégalités se sont présentées par les hypothèses du théorème de calcul de la variance  $\sigma_\pi^2$ .

Cependant, on propose la formule globale de ce problème d'optimisation comme suit:

$$(Q) = \begin{cases} \min_{\xi, \beta, p} \sigma^2(\xi, \beta, p) = \frac{(1+\xi)\beta^2}{(1/p-\xi)^2} \left[ \frac{(1/p-1)^2}{1+\xi} + \frac{1+\xi}{(1/p-\xi)^2} - \frac{2}{1/p-\xi} + 2 \right] \\ \frac{1}{M} \sum_{i=1}^M \log(1 + \xi \frac{y_i}{\beta}) = \xi \\ \frac{1}{M} \sum_{i=1}^M \log \frac{y_i/\beta}{1+y_i/\beta} = \frac{1}{1+\xi} \\ \frac{1}{2} \leq \xi < \frac{1}{p}, \quad 2 > p \geq 1, \quad \beta > 0 \end{cases} \quad (8)$$

avec  $y_1, \dots, y_M$ : sont les observations réelles  $x_i$  qui dépassent le seuil  $u$ .

La méthode des extrema liés peut être utilisée pour trouver un minimum local.

### 3.2 Recherche d'une solution globale

Notre problème d'optimisation (Q) peut être résolu de manière globale, en supposant que les trois paramètres  $(\xi, \beta, p)$  sont inconnus simultanément. Dans ce cas, c'est un problème d'optimisation ordinaire où on peut utiliser un algorithme génétique adapté pour la recherche d'une solution optimale globale  $(\xi^*, \beta^*, p^*)$ . On peut également utiliser d'autres algorithmes comme par exemple le recuit simulé, la recherche Tabou, etc...

En pratique, surtout pour les événements extrêmes, on dispose peu d'information pour la détermination du paramètre d'aversion au risque. Cette incertitude est liée à l'ambiguïté et la complexité qui caractérisent les risques extrêmes. En effet, on peut supposer, conditionnellement aux paramètres  $(\xi, \beta)$  de la loi des valeurs extrêmes (Pareto généralisée), que  $p$  est un paramètre aléatoire de loi de probabilité connue et de support  $[1, 2[$ .

Supposant:

$$p = \psi_\alpha(a) = 1 + a^\alpha$$

$\alpha > 0$  et  $a$  variable aléatoire uniformément distribuée sur l'intervalle  $[0, 1[$ .

Dans ce cas, on utilise une approche par simulation qui intègre des données réelles de coûts de sinistres de valeurs extrêmes en assurance automobile. Ceci permet à l'utilisateur d'avoir une batterie de scénarii de situations complexes, où manque de l'information pour la détermination exacte du paramètre  $p$ .

Le problème (Q) devient un problème d'optimisation avec contraintes stochastiques:

$$(Q') = \begin{cases} \min_{\xi, \beta, p} \sigma^2(\xi, \beta, p) = \frac{(1+\xi)\beta^2}{(1/p-\xi)^2} \left[ \frac{(1/p-1)^2}{1+\xi} + \frac{1+\xi}{(1/p-\xi)^2} - \frac{2}{1/p-\xi} + 2 \right] \\ \frac{1}{M} \sum_{i=1}^M \log(1 + \xi \frac{y_i}{\beta}) = \xi \\ \frac{1}{M} \sum_{i=1}^M \log \frac{y_i/\beta}{1+y_i/\beta} = \frac{1}{1+\xi} \\ p = \psi_\alpha(a) \\ g_1 = \frac{1}{2} - \xi \leq 0 \\ g_2 = \xi - \frac{1}{p} \leq 0 \\ g_3 = 1 - p \leq 0 \\ g_4 = p - 2 < 0 \\ g_5 = \beta > 0 \end{cases} \quad (9)$$

Bien que les problèmes d'optimisations combinatoires soient souvent faciles à définir, ils sont généralement difficiles à résoudre. En effet, ce problème appartient à la classe des problèmes NP-difficile et ne peut pas être résolu par les méthodes classiques (Gard (1979)).

## 4 Approche de résolution

Etant donnée l'importance du problème de minimisation de la variance  $\sigma_\pi^2$  avec contraintes d'égalités et d'inégalités, on propose une nouvelle méthode d'optimisation reposant sur un algorithme génétique. Ces dernières sont fondées principalement sur diverses heuristiques, souvent spécifiques à un type de problème.

## 4.1 Approche combinatoire d'optimisation

Les méta-heuristiques sont représentées essentiellement par les méthodes de voisinage et les algorithmes évolutifs. Ils constituent une classe de méthodes approchées adaptables à un très grand nombre de problèmes combinatoires et de problèmes d'affectation sous contraintes. Ces méta-heuristiques sont généralement rapides mais fournissent le plus souvent des solutions moins satisfaisantes. En appliquant certains algorithmes évolutifs de types génétiques, la majorité des problèmes d'optimisation combinatoire engendrent des éléments de la population relativement de mauvaise qualité, ce qui peut conduire à une convergence prématurée vers un optimum local. Afin de remédier à ce problème, certains de ces éléments doivent être convenablement mutés, par comparaison à leurs parents. Kirkpatrick & al. (1983) ont proposé d'associer la règle d'acceptation de Métropolis de l'algorithme Recuit Simulé (RS) (Youssef et Sadiq & Adiche (2001)), à celui de l'AG. Ceci donne une nouvelle variante de l'AG appelée Algorithmes Génétique recuit Simulé, AGS, où la règle de Métropolis opère en parallèle, pour le choix des individus candidats à une opération de mutation. Ce qui réduit considérablement le risque de convergence prématurée vers un optimum local. Un tel algorithme a une grande popularité en optimisation, dans le cas où la forme de la fonction objectif ne se prête pas bien aux méthodes classiques connues, telles que les méthodes: Quasi-Newton, extrema liés et autres.

### *Principe de l'AGS*

La particularité de l'algorithme AGS est l'exploration de l'espace de solutions réalisables, par la conception d'opérateurs de reproduction, croisement et mutation. Ces opérateurs agissent, à chaque génération, sur la population formée de solutions. Ce qui apporte des modifications qualitatives pouvant aller jusqu'à la satisfaction d'un critère de convergence prédéfini (convergence vers une meilleure solution). Afin d'introduire l'algorithme AGS, on présente certains concepts de base qui sont nécessaires à la compréhension de ses mécanismes de fonctionnement.

### *Amélioration de l'opérateur mutation*

Comme on l'a indiqué dans le paragraphe précédent, cette opération peut être couplée à la règle d'acceptation de Métropolis de l'algorithme RS. Ce qui permet d'éviter une convergence prématurée vers une solution locale. Cette règle permet d'aider à choisir d'une façon plus cohérente les individus pour mutation, selon leur degré de contribution à l'ajout de l'information permettant l'amélioration de la qualité de la population courante. En effet, dans un voisinage de deux chromosomes parents sélectionnés, deux autres chromosomes seront produits en utilisant les opérateurs de croisement et de mutation. Un enfant et un parent sont résultats de quatre chromosomes arbitraires.

On peut observer les deux cas suivants:

La valeur d'aptitude de l'enfant est meilleure que celle du parent, alors l'enfant est choisi et utilisé pour générer un nouveau chromosome pour la génération suivante.

La valeur d'aptitude de l'enfant est inférieure à celle du parent, l'enfant peut être choisi, si pour un nombre aléatoire  $r \in [0, 1]$  est inférieur à  $P_T = \exp\left(\frac{(F_{Parent}^* - F_{Enfant}^*)}{T}\right)$ , où  $F^*$  est la fonction critère et  $T$  est le paramètre "température courante" utilisé par l'AGS. Cette règle est connue sous le nom de "Règle de Métropolis".

Cette procédure est également appliquée aux autres paires d'enfants et de parents.

D'une génération à une autre, le paramètre température évolue suivant la formule  $T_k = \rho T_{k-1}$ ,

$k = 1, \dots, NG$ , où  $NG$  est le nombre de générations prédéterminé et  $\rho$  est le taux de refroidissement relatif à l'algorithme AGS.

La température initiale peut être choisie par la formule  $T_0 = \frac{1}{L_d} \sum_{i=1}^{L_d} \frac{(F_i^* - F_{i-1}^*)}{\log(P_0)}$ , où  $L_d$  (rep.  $L_m$ ) est le nombre de descentes (rep. montées) relatif au mouvement au voisinage d'une solution courante.  $P_0 \approx 1$ ,  $P_0 = 0.999$  (correspond à  $T_0$ ).

**Critère d'arrêt:** L'algorithme de recherche globale s'arrête lorsque le nombre de générations prédéfini  $NG$ , est entièrement épuisé, ou bien la température finale  $T_{NG} < 0.001$ .

#### **Algorithme d'optimisation AGS**

Donc, on est en mesure de décrire l'algorithme AGS suivant:

*Initialiser*  $P(0)$  de taille  $M$ , *initialiser*  $T_0, NG = 0$ ,

**Tant Que** critère d'arrêt non satisfait *faire*

*Evaluer l'aptitude de la population*  $P(NG)$

**Pour**  $j = 1$  à  $NG$  **Faire**

*Selectionner aléatoirement deux chromosomes*  $P^*(NG)$  à partir de  $P(NG)$

*Appliquer l'opérateur de croisement sur*  $P^*(NG)$  pour produire  $P^{**}(NG)$

*Executer l'opérateur de mutation sur*  $P^{**}(NG)$  pour produire  $P^{***}(NG)$

*Evaluer l'aptitude des chromosomes*  $P^{***}(NG)$

*Excuter la règle de Metropolis pour*  $P^{***}(NG)$  et  $P^*(NG)$  pour générer deux chromosomes de la génération suivante  $P(NG+1)$ .

**Fin Pour**

$NG = NG + 1$

$T_{NG} = \rho T_{NG-1}$

**Fin Tant Que.**

#### **Algorithme alternatif de calcul de la prime ajustée**

1. Choisir le nombre des excès  $M$  dépassants le seuil  $u$  de l'échantillon  $(x_1, \dots, x_n)$ .
2. Déterminer les paramètres d'optimalités  $(\xi^*, \beta^*, p^*)$  par l'AGS.
3. Calculer l'estimateur de la prime ajustée en utilisant la formule suivante:

$$\hat{\pi}_u(X) = \hat{\lambda}^{\frac{1}{p^*}} \frac{\beta^*}{\frac{1}{p^*} - \xi^*} \quad \text{avec} \quad \hat{\lambda} = \frac{M_u}{n}$$

où

$M_u$  est le nombre des excées au-delà du seuil  $u$

$n$  est la taille de l'échantillon de la population assurée

Une implimentation informatique a été effectuée à l'aide des logiciels adaptés.

## **5 Application numérique en assurance automobile**

*Description de l'échantillon de données*

## Prime des risques extrêmes

La base de données concerne un échantillon aléatoire de 50000 observations pour des véhicules 4 roues de tourisme durant l'année 2004, issu d'un portefeuille d'assurance française. La variable étudiée correspond aux coûts des sinistres.

Rappelons que le montant d'un sinistre inclut l'indemnisation directe de la victime, les frais de gestion internes à la société ainsi que les frais externes (expertise, frais judiciaires) afférents à ce sinistre. Il n'inclut pas les frais d'acquisition de contrats.

En assurance automobile, ce montant est inconnu. Il est donc la réalisation d'une variable aléatoire réelle positive ou nulle.

On remarque aussi que le montant du sinistre est différent du coût à la charge de l'assureur puisque l'assureur déduit le montant de la franchise fixé dans le contrat. Cette franchise est presque nulle pour un contrat tous risques. Dans ce travail on utilise le montant d'évaluation cumulé qui reflète le mieux la charge de sinistres extrêmes détectée par la théorie des extrêmes.

### *Exploration et ajustement des sinistres extrêmes*

L'idée naturelle est de partitionner le portefeuille afin de constituer des sous portefeuilles sur lesquels les risques peuvent être considérés comme indépendants et de même loi. Il est donc nécessaire de détecter les sinistres extrêmes afin d'analyser leurs coûts. En général, on utilise une méthode graphique basée sur l'espérance conditionnelle des excès (fonction moyenne des excès):

$$E[X - u | X > u] = \frac{\sigma + \xi u}{1 - \xi} \quad (10)$$

L'estimation graphique du seuil  $u$  consiste à repérer les plages de linéarité qui donne une stabilité à la fonction moyenne des excès de l'échantillon étudié (figure 1). Cette méthode détecte toujours moins de valeurs extrêmes que des autres

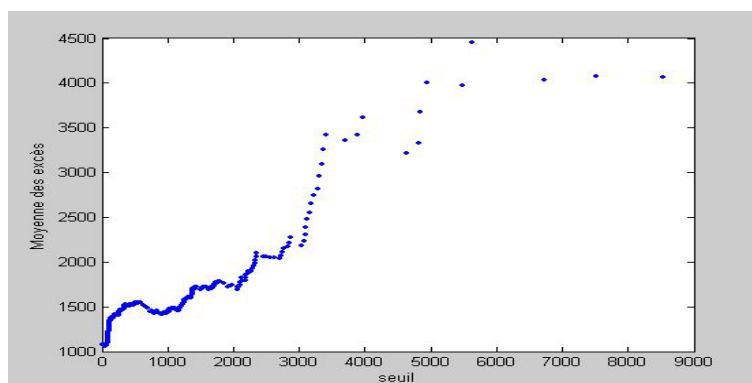


FIG. 1 – La fonction moyenne des excès de l'échantillon total des assurés

méthodes classiques. Plus le seuil est grand, plus on s'intéresse aux événements extrêmes mais comme la prime ajustée sera estimée avec peu de données, sa variance sera grande. Cependant la détermination du seuil par la méthode POT permet de calculer une prévision du coût d'un sinistre extrême pour une probabilité d'occurrence de 99,9% d'être une valeur extrême et avec un niveau de confiance de 95%.

La figure 2 présente une estimation ponctuelle du quantile extrême ainsi que la

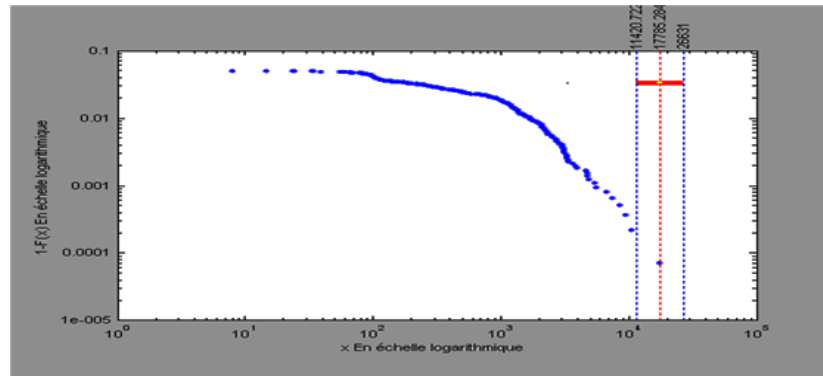


FIG. 2 – La prévision d'un coût extrême

limite inférieure et supérieure de la perte maximale pour la compagnie d'assurance en cas de survenance d'un sinistre grave avec une probabilité de 0,1%.

#### Recherche des solutions d'optimalités $(\xi^*, \beta^*, p^*)$ et interprétations

La recherche d'optimums globaux pour le problème général stochastique (Q') sous contraintes, en utilisant l'algorithme AGS, donne des résultats raisonnables de la variance  $\sigma_{\pi}^*$ . Mais, les valeurs optimales obtenues sont relativement élevées, comme l'indique la figure 3, pour une seule exécution. Pour 100 réalisations effectuées par l'algorithme AGS, on engendre un échantillon de solutions optimales  $\{\xi_i^*, \beta_i^*, p_i^*\}_{i=1..100}$  et de moindres variances  $\sigma_{\pi_i}^*$ ,  $i = 1..100$ , donné par le tableau 1. Ceci correspond à une centaine de situations réelles identiques, conditionnellement à un choix aléatoire du paramètre  $p$ . Un tel choix doit tenir compte de l'asymétrie de l'information assuré-assureur. Il peut être également lié à un environnement économique incertain ne permettant pas de déterminer avec exactitude les valeurs de ce paramètre. C'est ce qui a motivé notre choix de supposer que  $p$  est aléatoire, prenant ses valeurs possibles dans un ensemble vérifiant les hypothèses du théorème de Nicer & Boukhtela (2004).

Dans certains cas rares, la meilleure configuration trouvée peut ne pas vérifier les contraintes et engendrer des résultats négatifs (tableau 1), du fait que le problème d'optimisation stochastique (Q') est traité en mode  $P$ -presque sûr ( $P$ : indique

Prime des risques extrêmes

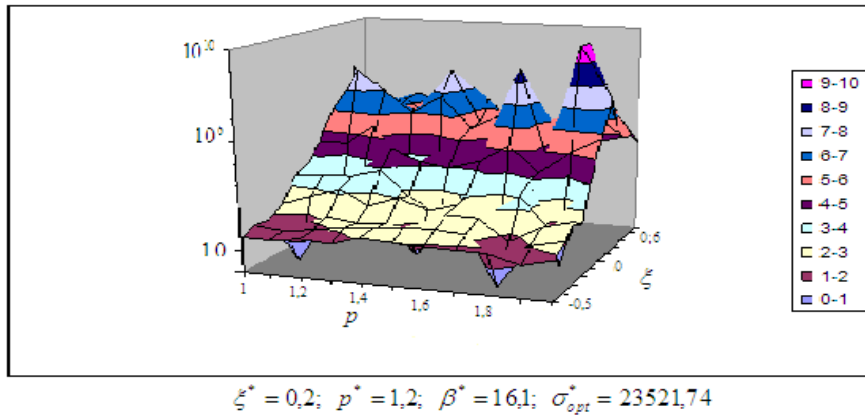


FIG. 3 – Courbe de la variance  $\sigma_{\pi_i}^*$  conditionnellement à un comportement aléatoire du paramètre ;  $p = \psi(u)$ ,  $u \sim \mathcal{L}(U^{[0,1]})$ ;  $U^{[0,1]}$ : est la loi uniforme sur  $[0, 1]$

N° réalisation	$\xi^*$	$\beta^*$	$p^*$	$\sigma_{\pi_i}^*$
1	0,89	2,49	1,3	18077,56
2	0,55	1,19	1,14	1129073,25
3	0,85	2,74	1,3	34163,63
4	0,51	2,69	1,01	-409226,48
5	0,73	1,93	1,74	1107202,66
6	0,68	2,31	1,2	55253,22
7	0,59	1,21	1,96	3985069,54
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
100	0,72	2,19	1,27	363670,29

TAB. 1 – Résultats des variances optimales de 100 réalisations obtenus par l'AGS

Probabilité 1). Pour remédier à ce problème, on utilise des techniques de filtrage ou de sélection de résultats obtenus par l'AGS, donnant que des valeurs positives de la variance. On remarque pendant l'exécution de l'AGS que très peu de valeurs négatives sont obtenues, ce qui correspond à un ensemble d'éventualités provenant de l'environnement hasardeux, de pourcentage très faible. La performance de cette procédure de filtrage (nettoyage des résultats) dépend largement du schéma d'évolution de la variance  $\sigma_{\pi_i}^*$  en fonction de  $p$ , voir la figure 4:

Cette variabilité élevée de la variance produit une certaine incertitude dans le

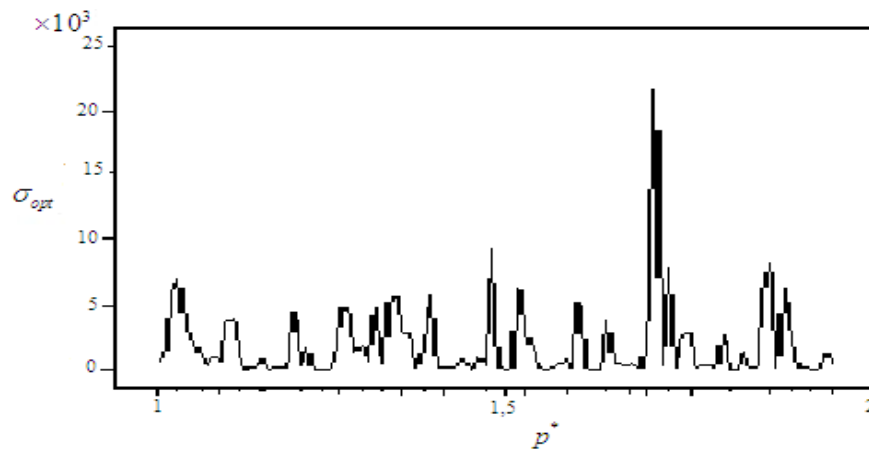


FIG. 4 – Evolution des optima de la variance en fonction des optima du paramètre aléatoire  $p$

calcul de la prime et pourrait créer une certaine méfiance chez l'utilisateur (assureur/réassureur). Donc ce dernier doit être plus prudent par le biais d'une telle stratégie, requise par l'usage du régulateur aléatoire, et pouvant conduire à un calcul de prime de risque répondant à des normes économiquement souhaitables en matière de tarification des risques extrêmes. De fait que la prise en compte comme contrainte, dans le programme (Q'), de la charge de sinistres extrêmes risque d'augmenter la variance  $\sigma_{\pi}^*$ , on a décidé de multiplier ces valeurs par un coefficient de pondération afin de condenser les échelles et de pouvoir normaliser la modélisation de ces événements (exceptionnels). En pratique, ceci peut être vu comme conséquence de la mise en place d'une stratégie de prévention adéquate; conduisant à une sinistralité relativement réduite, où on remplace  $y_i$  dans (Q') par  $y'_i = C_i y_i$ , les  $C_i$  sont estimés en présence de l'information statistique ou supposés aléatoires en absence de cette information. Ce dernier cas est supposé pour notre problème, où les  $C_i$  sont de loi uniforme sur l'intervalle  $[0,1]$ ; c'est le cas de situations se comportant dans un environnement économique complexe et en absence de systèmes d'information statistique approprié. Comme résultats de cette approche obtenus par 100 réalisations du problème (Q') par l'AGS, on présente les distributions en "box plot" des variables aléatoires  $\xi_i^*$ ,  $\beta_i^*$ ,  $p_i^*$  et  $\sigma_{\pi_i}^*$ , ainsi que les résultats de l'application du test d'équation Kolmogorov-Smirnov pour cha-



Prime des risques extrêmes

cune de ces variables.

Représentation en Box-plot de la distribution des paramètres:  $\xi_i^*$ ,  $\beta_i^*$ ,  $p_i^*$  et  $\sigma_{\pi_i}^*$

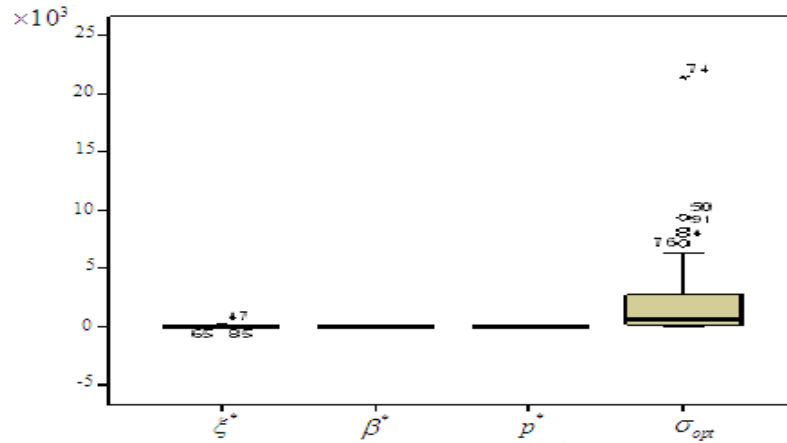


FIG. 5 – Box-plot de la distribution des paramètres

		$\xi^*$	$\beta^*$	$p^*$	Variance
N		100	100	100	100
	Mean	,3648	1,8995	1,5379	1890,94122
Normal Parameters(a,b)	Std. Deviation	,30946	,60891	,28529	2970,400882
Most Extreme differences	Absolute	,088	,108	,093	,262
	Positive	,088	,108	,064	,217
	Negative	-,082	-,085	-,093	-,262
Kolmogorov-Smirnov Z		,880	1,082	,928	2,622
Asymp. Sig. (2-tailed)		,420	,192	,356	,000

a: Test distribution is Normal.

b: Calculated from data.

TAB. 2 – Résultats des tests statistiques

On peut tracer l'histogramme (figure 6) des résultats de la variance  $\sigma_{\pi_i}^*$  obtenus par l'AGS et la densité de probabilité pour mieux visualiser la loi de probabilité associée. La variabilité intrinsèque des résultats de la variance  $\sigma_{\pi}^*$  n'est pas significativement de la loi normale, comme en témoigne le qq-plot. Cela qualifie l'incertitude sur le paramètre  $p$  qui pourra donner une représentation plus adaptée à la réalité, en présence des événements "rares" (tableau 2). Quand on trouve une déformation anormale de la courbe, c'est aussi une information importante

FIG. 6 – Répartition de la variance  $\sigma_{\pi_i}^*$ 

qui nous dirige à déterminer une stratégie implicite de prévention. En se basant sur l'analyse empirique, on prouve que la répartition statistique de la variance suit la loi de Pareto. Intuitivement cela se justifie par la figure 7: Ces graphiques

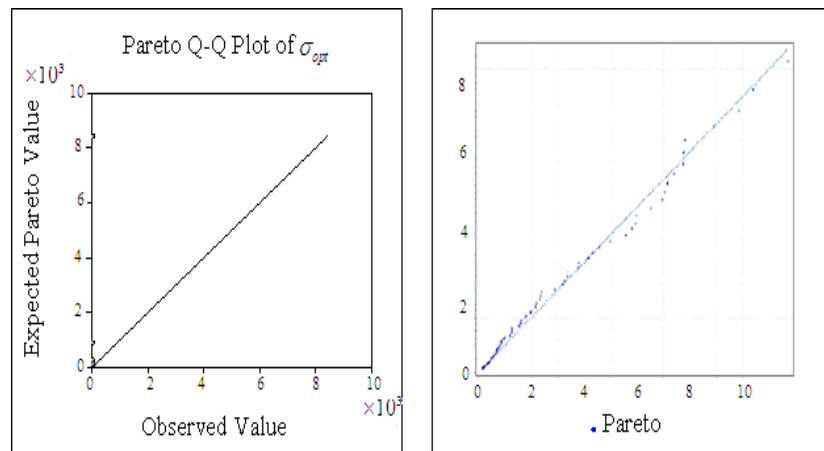


FIG. 7 – Pareto Q-Qplot

montrent que l'ajustement est licite en ce qui concerne le test graphique. Le diagramme de Pareto (figure 8) est utilisé pour résumer et afficher graphiquement l'importance relative des différences entre groupes de données. Ce diagramme est construit en segmentant les données en groupes. Il permet d'identifier les catégories de risques qui nécessitent la mise en place de stratégies de préventions et modes de calculs de la prime ajustés distincts. Cette loi permet aussi d'analyser l'évolution des degrés de l'incertitude des assureurs en fonction des valeurs de la variance. En effet, la représentation graphique montre qu'au-dessus d'une certaine valeur  $e_0$ , la répartition cumulative des résultats de la variance obtenus par l'AGS, une fois tracée sur un graphique Log-Log, forme une ligne presque droite. Cela conforte l'idée qu'il y a plus de risques avec une faible variance que de risques avec une variance élevée. Ce qui corrobore aussi la remarque de Gustav (1925); la répartition de Pareto se retrouve curieusement sur les grands risques, où 20% des sinistres pouvaient représenter 80% de la charge totale. Dans ce cas, les conclusions qu'on a obtenu de l'analyse statistique sont claires et montrent qu'il y a une situation favorable si les valeurs de la variance sont faibles et une situation défavorable dans le cas contraire. Ces résultats (variances élevées) peuvent ne pas intéresser l'utilisateur, en matière de calcul de prime et suscitent l'utilisation d'une autre stratégie de réduction de celle-ci, plus adéquate. Étant donnée l'importance de ce

Prime des risques extrêmes

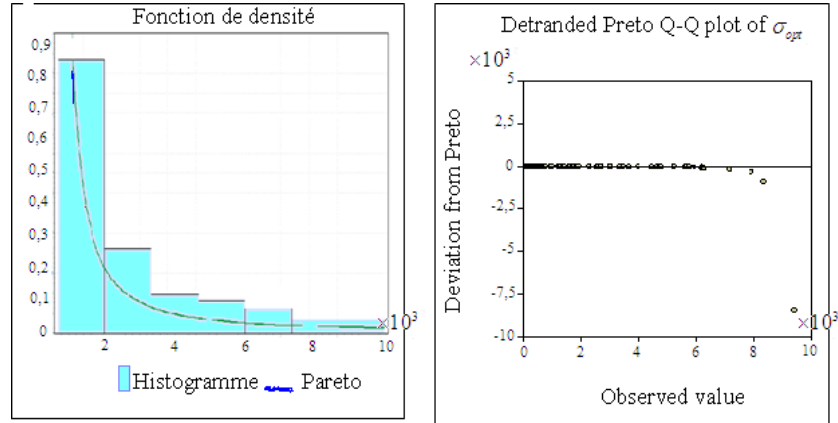


FIG. 8 – Fonction de densité de la loi de Paréto

problème, on propose une stratégie avantageuse de réduction de la variance afin de minimiser le souci de l'utilisateur. Cependant, la réduction de la variance peut être déterminée en introduisant une variance ciblée (qui correspond à la valeur retenue à partir de laquelle la courbe de la distribution forme presque une droite), en cherchant à minimiser l'écart entre la vraie variance et une variance attractive déterminée selon le principe de Paréto, dans notre cas, et par lecture direct sur le graphique de Pareto, on a retenu la valeur  $e_0 = 7900$ . Puis, on conditionne la valeur  $p_i^*$  par les écarts réduits qui ne dépassant pas un certain seuil fixé  $e_0$  déterminé selon des considérations statistiques et économiques. Cette stratégie est présentée au cours du paragraphe suivant.

*Stratégie de choix optimal du paramètre  $p_i$  pour une variance réduite  $\sigma_{\pi_i}^*$ :*

Afin d'obtenir des optima globaux correspondants à une variance  $\sigma_{\pi_i}^*$ , en présence d'une valeur ciblée désirable, on propose de résoudre un problème d'optimisation stochastique, constituant une autre variante du problème (Q'), qui est donné par:

$$\min_{\xi, \beta, p} \tilde{\sigma}_{\pi_i}^*(\xi, \beta, p) = |\sigma_{\pi_i}^*(\xi, \beta, p) - V_{attractive}|$$

où  $V_{attractive}$  est une valeur de la variance désirable. Les contraintes restent inchangées.

Cette formule commode permet de déterminer la variation minimale entre la vraie variance et la variance ciblée, quand on trouve des résultats de la variance  $\sigma_{\pi_i}^*$  très élevés. Si, par exemple, on suppose que la variance ciblée est égale à  $e_0$ , une réalisation probable de la solution optimale est donnée par: ( $p^* = 1.06$ ,  $\xi^* = -0.54$ ,

et  $\beta^* = 0,18$ ) et  $\tilde{\sigma}_{\pi_i}^*$  est de l'ordre de 7890 qui est relativement réduite par comparaison à la valeur ciblée de l'ordre de 7900, comme le montre la figure 9. Ceci signifie que le problème stochastique de minimisation de l'écart absolu rend

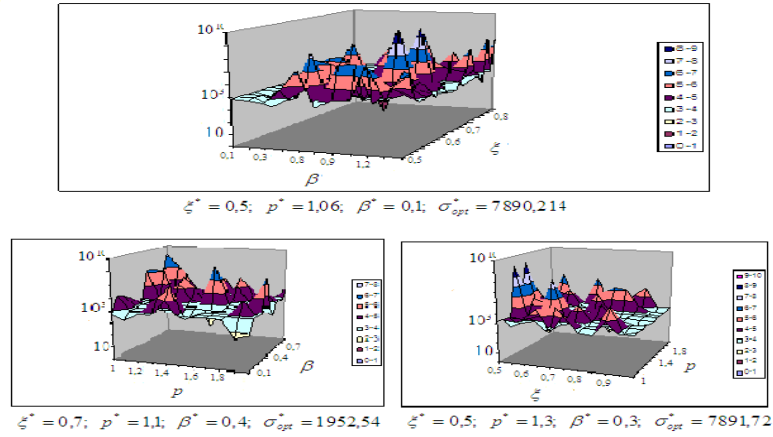


FIG. 9 – Courbes de configurations optimales des écarts stochastiques  $\tilde{\sigma}_{\pi_i}^*(\xi^*, \beta^*, p^*)$  pour la variance ciblée à l'ordre de  $e_0$

en même temps toutes les variances minimums presque stables et limitées par le seuil  $e_0$ , compte tenu des contraintes imposées.

Bien que les formules soient partiellement exactes, il est clair que  $\tilde{\sigma}_{\pi_i}^*$  est suff-

$N^\circ$ réalisation	$\xi^*$	$\beta^*$	$p^*$	$\tilde{\sigma}_{\pi_i}^*$	$\log(\log(\tilde{\sigma}_{\pi_i}^*))$
1	0,6	0,41	1,9	7570,55	0,58
2	0,58	0,43	1,98	7829,58	0,58
3	0,95	0,33	1,49	7846,94	0,57
4	0,64	0,1	1,64	7875,34	0,59
5	0,96	0,39	1,97	7812,07	0,58
6	0,54	0,35	1,98	57639,29	0,49
7	0,67	0,76	1,27	6431,72	0,53
8	0,8	0,85	1,34	7375,97	0,57
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
100	0,74	0,33	1,99	7738,82	0,58

TAB. 3 – Résultats des écarts absolus optimaux de 100 réalisations obtenus par l'AGS

isamment petit (tableau 3), c'est la valeur de la variance ciblée qui est désirable et l'on voit qu'on peut trouver des valeurs d'écarts qui tendent à faire diminuer et

Prime des risques extrêmes

presque stabiliser le niveau de l'incertitude de l'utilisateur. Il y a d'ailleurs un optimum que notre problème général ne permet pas de calculer, et il est intéressant de noter ici l'incidence de l'introduction de la variance attractive sur la stratégie du calcul de la prime ajustée pour les risques exceptionnels.

Il est utile d'étudier en détail les résultats empiriques obtenus (figure 10) pour une

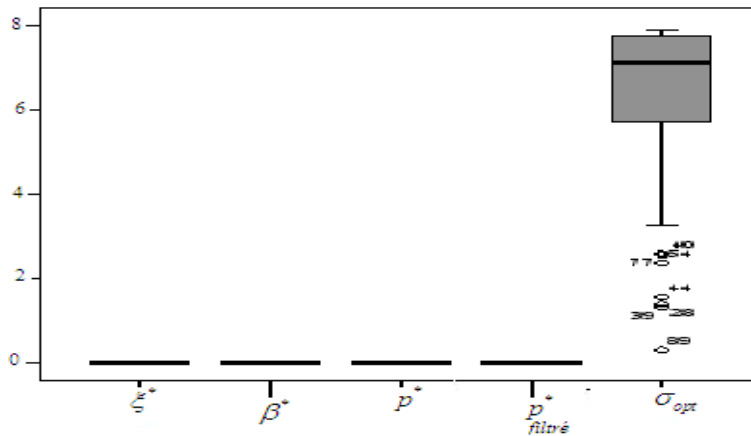


FIG. 10 – Box-plot de la distribution de l'écart absolu

stratégie du calcul des variances minimales suffisantes. Donc, graphiquement, on peut mentionner que les paramètres qui nous permettent d'arriver à des résultats optimaux, sont distribués sur une loi normale. Cela peut amener à effectuer un test de Kolmogorov-Smirnov pour cet ajustement (tableau 4).

Il est crucial de pouvoir montrer une bonne significativité (presque 42%, 20%,

		$\xi^*$	$\beta^*$	$p^*$	$\tilde{\sigma}_{\pi_i}^*$
N		100	100	100	100
Normal Parameters(a,b)	Mean	,7217	,4930	1,5341	6446,8684
	Std. Deviation	,12235	,22857	,30810	1826,10275
Most Extreme differences	Absolute	,087	,079	,091	,214
	Positive	,087	,067	,091	,214
	Negative	-,064	-,079	-,078	-,201
Kolmogorov-Smirnov Z		,868	,790	,862	2,026
Asymp. Sig. (2-tailed)		,438	,560	,447	,001

a: Test distribution is Normal.

b: Calculated from data.

TAB. 4 – Résultats des tests statistiques

35% respectivement pour les paramètres  $\xi_i^*$ ,  $\beta_i^*$ ,  $p_i^*$ ). Cela signifie que notre

stratégie est performante du fait que les solutions optimales, d'un problème stochastique combinatoire de complexité exponentielle, sont approximativement distribuées sur une loi Gaussienne.

*Représentation graphique de l'écart absolu:*

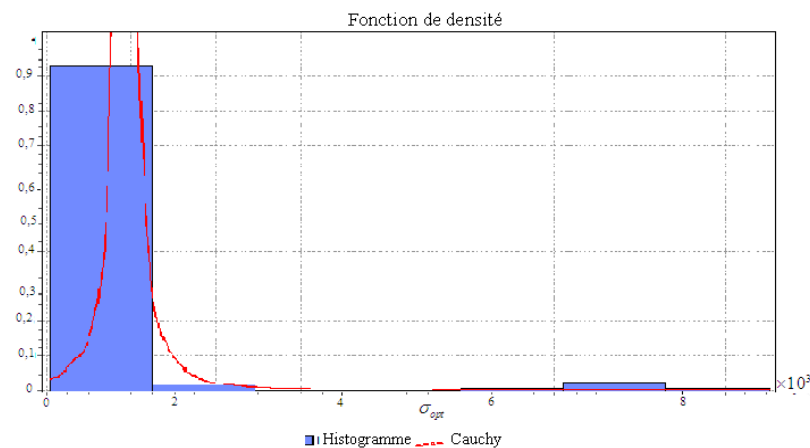


FIG. 11 – *Fonction de distribution de la loi de Cauchy*

La figure 11 montre un ajustement de la variabilité des résultats de l'écart absolu, obtenus par l'AGS, par une loi de Cauchy généralisée. Cet ajustement est très intéressant car il peut donner une bonne représentation de données réelles. De plus, le test de KHI-DEUX a justifié un bon ajustement par la loi de Cauchy, presque de l'ordre 24% de niveau de significativité.

D'une manière générale, on constate que les résultats de l'écart sont limités par la valeur de la variance ciblée. On peut donc noter que les variations des résultats de l'écart entre la vraie valeur et la valeur visée par l'assureur sont au plus inférieures que la valeur ciblée. En pratique, ces résultats soulignent la robustesse de la stratégie de réduction de la variance de l'étendu  $\sigma_{\pi}^*$ . La simplicité et la rapidité de mise en œuvre ainsi que la plus grande maîtrise de la méthodologie confèrent à la procédure proposée ici un avantage indiscutable pour le réglage d'une fonction d'écart de type Cauchy.

Cette loi a une grande importance dans le système de prévention en présence des événements extrêmes, puisque les distributions stables peuvent être caractérisées à partir du théorème central limité. Une généralisation par Gnedenko & Kolmogorov (1954) indique que si la condition de variance finie n'est pas respectée, la seule loi limite possible de la somme de  $n$  variables aléatoires, est une loi stable.

Concrètement, cela signifie que le vrai problème économique ici est celui du

## Prime des risques extrêmes

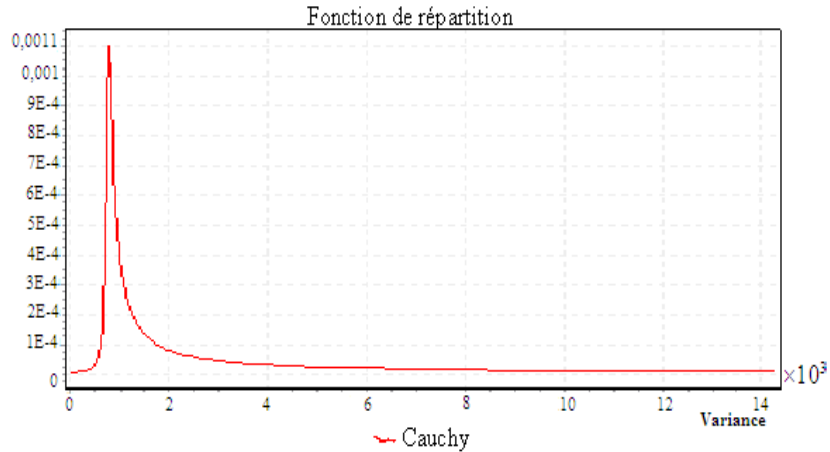


FIG. 12 – *Fonction de répartition de l'écart absolu*

niveau de protection. Ce qui est important à souligner, en revanche, c'est une relation beaucoup plus fondamentale entre cet écart et le niveau de la protection. On voit ainsi que le risque, ayant le plus grand écart en valeur absolue, peut être qualifié d'exceptionnel. La valeur de l'écart absolu est fortement croissante en fonction du degré de protection (figure 12). Tout ceci rend délicate la mise en place d'un système de prévention, dont le volet "équité" ne doit pas être négligé. Cependant, la réussite de cette démarche permet de qualifier le risque et de conduire à une meilleure adaptation aux exigences des utilisateurs. Autrement dit, le calcul de la variance de l'étendu des intervalles de confiance de l'estimateur de la prime ajustée incite les assureurs à être plus protégés contre les risques exceptionnels par le biais d'un régulateur stochastique (le paramètre de poids). Cela permet donc à l'assureur de discriminer ces produits par le biais d'un jeu de couvertures différentes, ce qui incite aussi les assurés à être plus prudents.

### *Comportement stochastique pour $p_{cond}$*

Le fait que les conditions de minimisation de la variance soient remplies, la performance de notre stratégie peut être vérifiée à partir de la réduction de l'écart absolu de la vraie variance et celle ciblée et il peut être utile éventuellement d'effectuer une transformation de manière à rendre l'écart presque stable et limité par la valeur de la variance ciblée. Cependant, la résolution de ce problème nous amène en quelque sorte à régulariser le comportement du paramètre du poids conditionné par une variabilité inférieure à la variabilité relative à la variance ciblée de l'étendu des intervalles de confiance, en considérant  $p_i$  au lieu de  $p'_i$ , supposé toujours fonction stochastique définie par la relation:

$$p_i = \psi(u_i) = p'_i u \quad \text{avec} \quad p'_i \geq 1$$

L'optimum aléatoire  $p_i^*$  conditionné par une variabilité faible du critère d'écart présente une variabilité normale justifiée par l'analyse statistique d'un échantillon d'optima globaux obtenus par l'AGS. La distribution bootstrap (c'est-à-dire la distribution des moyennes empiriques obtenues à partir des 10000 échantillons bootstrap) est représentée par la figure 13:

Le caractère tout à fait normal de la distribution du paramètre optimal  $p_i^*$ ;  $i =$

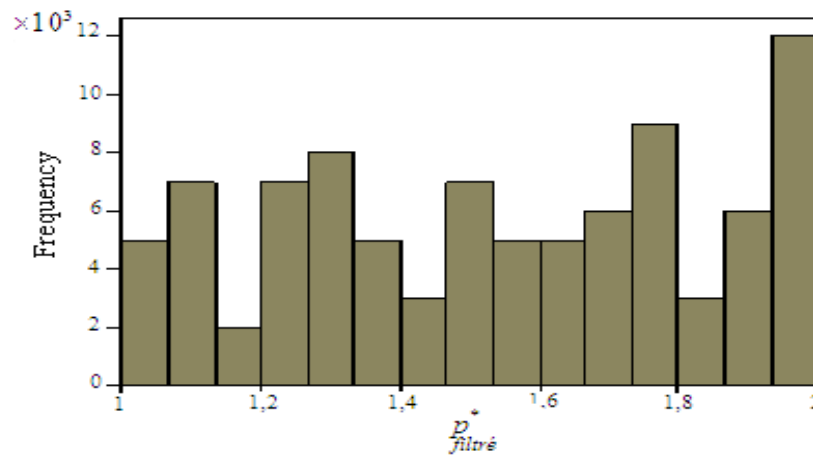


FIG. 13 – Histogramme du paramètre stochastique conditionnel  $p_{cond}$  pour 10000 échantillons bootstrap

1..100, conditionnelle aux écarts optimaux obtenus par l'AGS, nous conduit cependant à minimiser le doute sur les résultats de la variance de l'étendu des intervalles de confiance. Il est donc facile de proposer le test de Kolmogorov-Smirnov pour justifier l'ajustement de la variable aléatoire de ce paramètre  $(p_{cond})_i$ ;  $i = 1..100$ .

Les résultats du tableau 5 montrent une approximation asymptotique significative. Il est donc très intéressant d'obtenir des valeurs de  $p_i^*$  qui se répartissent selon une courbe de Gauss. On remarque en effet que la variance est presque faible donc il s'agit d'une précision appréciable et l'intervalle de confiance bootstrap est de faible amplitude.

#### Prime d'optimalité des risques extrêmes

Il semble que la difficulté d'interprétation du caractère exceptionnel d'un risque a également un impact sur le comportement des organisations. Par exemple, les assureurs sont le plus souvent allergiques aux situations d'ambiguïté (univers incertain), dans lesquelles le montant des pertes éventuelles est difficilement calculable. Ils ont alors tendance à pratiquer des tarifs largement supérieurs à ceux



Prime des risques extrêmes

		$p^*$	
N		90	
	Mean	1,5341	
Normal Parameters(a,b)	Std. Deviation	,30810	
Most Extreme differences	Absolute	,091	
	Positive	,091	
	Negative	-,078	
Kolmogorov-Smirnov Z		,862	
Asymp. Sig. (2-tailed)		,447	
Monte Carlo Sig. (2-tailed)	Sig.	,424(c)	
	99% Confidence Interval	Lower Bound	,411
		Upper Bound	,437

a: Test distribution is Normal.  
b: Calculated from data.  
c: Based on 10000 sampled tables with starting seed 2000000.

TAB. 5 – Résultats des tests statistiques

relatifs à des risques clairement spécifiés. Par ailleurs, le système de préventions, considéré dans notre démarche, prises pour obtenir une réduction de l'ampleur des sinistres, quelle que soit leur origine, nous permet de réaliser un tableau de bord pour le calcul d'une prime ajustée à variance réduite, conditionnellement à un paramètre aléatoire  $p$ . Les lignes du tableau 6 comprennent des réalisations des vecteurs de solution optimale de la fonction écart  $\{\xi_i^*, \beta_i^*, p_i^*\}, \bar{\sigma}_{\pi_i}^*; i = 1...53$  et la prime ajustée optimale  $\hat{\pi}_u$  qui est calculée à l'aide de la formule suivante:

$$\hat{\pi}_u(X) = \hat{\lambda}^{\frac{1}{p^*}} \frac{\beta^*}{\frac{1}{p^*} - \xi^*} \quad \text{avec} \quad \hat{\lambda} = \frac{M_u}{n}$$

où

$M_u$  est le nombre des excées au-delà du seuil  $u$

$n$  est la taille de l'échantillon de la population assurée

Les résultats du tableau 6 montrent que plus le point d'attachement se situe à un niveau élevé de sinistralité, plus la prime de risque exigée par l'assureur ou le réassureur augmente. D'où, pour pouvoir faire face à des sinistralités importantes, l'assureur peut proposer des contrats à prime variable. Tant que les utilisateurs se sont contentés de modéliser les risques, ils ont été brillants. C'est d'ailleurs ce que disait Emile Borel lorsqu'il écrivait, au sujet des compagnies d'assurances : "nous partons ainsi d'une base pratique assez solide pour que nous ayons dans la théorie la confiance qui peut être nécessaire pour ne pas avoir à tenir compte du scepticisme qui peut toujours être opposé à toute tentative d'explication rationnelle".

N° réalisation	$\xi^*$	$\beta^*$	$p^*$	$\tilde{\sigma}_{\pi_i}^*$	Prime pure
1	0,53	0,1	1,05	7768,56	1,84
2	0,85	0,1	1,03	2496,72	2,01
3	0,56	0,2	1,06	7691,03	4,49
4	0,56	0,37	1,27	7725,01	52,07
5	0,62	0,27	1,28	7853,98	54,99
6	0,54	0,36	1,32	5435,71	66,37
7	0,67	1,41	1,32	2673,73	738,32
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
53	0,57	1,17	1,7	6978,06	15661,65

TAB. 6 – Résultats de la prime optimale correspondante à des écarts absolus conditionnels obtenus par l'AGS

On présente ici le cas d'un échantillon de primes ajustées optimales obtenues par l'AGS (figure 14).

Bien que le système de prévention réduise significativement la variabilité du ré-

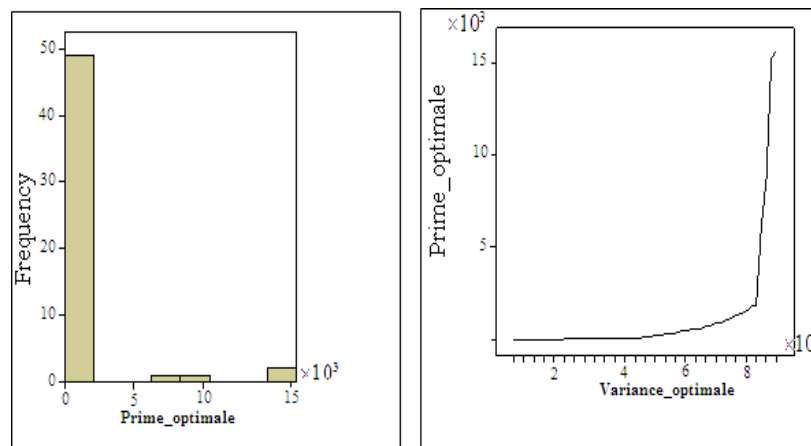


FIG. 14 – (a) Prime ajustée d'optimalité distribuée, (b) Réalisation de la Prime ajustée d'optimalité en fonction de l'écart conditionné par le paramètre  $p$

sultats, l'histogramme de la prime présente trois classes distinctes  $cl_i; i = 1, \dots, 3$ . Parallèlement, la figure 14 montre que l'évolution de la prime variable (par écart variable) est limitée par une droite. Dans ce cas, l'assureur doit veiller à ne pas céder trop de prime pour le niveau de risque à assurer.

Mais il s'agit là d'un autre problème, qui est l'adéquation entre la couverture et le

## Prime des risques extrêmes

risque: une sous-estimation dans les tarifs peut entraîner de graves difficultés pour l'assureur. A l'opposé, une surévaluation des tarifs est généralement à l'origine de l'incertitude, et le manque d'information peut nuire au rendement financier. L'assureur doit donc veiller à avoir des tarifs adéquats et une gestion adaptée de ses provisions techniques.

Cependant, on a constitué une prime corrigée d'équilibre appliquée à la catégorie des conducteurs extrêmement sinistrés. Après la mise en place du système de prévention peut être calculée selon la formule suivante:

$$\pi_{corrg} = \sum_{i=1}^3 f_{cl_i} (\hat{\pi}_{max}^* - \hat{\pi}_{min}^*)_{cl_i}$$

Avec

$f_{cl_i}$  est fréquence de la classe  $cl_i$ ,

L'intérêt de cette formule est de diminuer le montant des primes d'assurance, ce qui permet à l'assureur de proposer des contrats plus adaptés à un moindre coût. Pour cela, la question de l'adéquation à la couverture du risque extrême est abordée de manière très synthétique. A noter qu'il est important de raisonner en termes de couverture moyenne. En effet, dans la mesure où l'utilisateur n'est pas satisfait économiquement par cette prime, il peut générer éventuellement des primes simulées, à l'aide de sa distribution obtenue.

### **Conclusion**

L'analyse des risques extrêmes en assurance a permis de mettre en évidence l'importance des différences de modélisations statistiques entre les risques "standards" et ceux "dangereux".

En effet, on constate que la stratégie de pondération des sinistres extrêmes présente des phases de corrections afin d'obtenir une normalisation de ces événements, tout en conservant plus efficacement la présence de l'incertitude par le biais d'un régulateur aléatoire. Par ailleurs, cette stratégie tient compte davantage du système de prévention de la prime des risques extrêmes qui est basée sur la minimisation de la variance de l'étendu des intervalles de confiance de l'estimateur de la prime ajustée, par la nouvelle approche d'algorithme hybride. Ainsi, l'utilisation de l'AGS pour la résolution de ce problème de minimisation de la variance sous contraintes liées aux caractéristiques des événements extrêmes, présente plusieurs avantages: la fiabilité, la précision, le temps, la robustesse et surtout la garantie de convergence en un point stationnaire (à un point global). Cette méthode a révélé une certaine performance pour fournir des solutions approchées de bonne qualité pour des problèmes NP-difficiles.

L'application du système de prévention des risques extrêmes, au travers de la minimisation de la variance de l'étendu par l'AGS, offre une efficacité pour le calcul d'un tarif raisonnable et acceptable aux assurés et assureurs. Si on peut généraliser l'efficacité de cette approche pour un nombre de sinistres dangereux tendant vers

l'infini, il existe en revanche, dans notre application, très peu d'observations permettant de comprendre la raison de cette efficacité. Cette question constitue sans doute un défi important qui dessine de nouvelles perspectives de recherches, notamment en matière de gestion dynamique de la prime de couverture des risques extrêmes en assurance non vie.

## References

- Blondeau J. & Partrat C. (2003). *La Réassurance : approche technique*. Economica, Paris, France
- Camp C., Pezeshk S. & Cao G. (1998). Optimized design of two-dimensional structure using genetic algorithm, *Journal of Structural Engineering*, Vol. 124, p.551-559
- Centeno, M.L.& Guerra, M. (2010). The optimal reinsurance strategy: the individual claim case. *Insurance: Mathematics and Economics*, Vol. 46, p.450-460
- Charpentier A. (2009). *Les modèles en réassurance*. *Risque*, Vol. 80, p.36-41
- Chemarin S. (2005). Couverture des risques catastrophiques potentiels: Vers une théorie économique de l'assurabilité en incertitude. *Assurances et gestion des risques*, Vol. 73
- Chemarin S., Henry C. & E. Michel-Kerjan E. (2005). *Incertitude, Précaution et Assurabilité*. Commissariat Général au Plan.
- Chen S.Y. & Rajan S.D. (2000). A robust genetic algorithm for structural optimization. *Structural Engineering & Mechanics Journal*, Vol. 10, p.313-336
- Chiappori P.A. (1996). *Risque et assurance*. Dominos, Flammarion, Paris
- Cummins J.D. (1992). *Financial Pricing Of Property And Liability Insurance*. Dans Dionne G.(ed.) *Contributions to Insurance Economics*, Boston, Kluwer Academic Press, p141-169
- David C.M. Dickson (2005). *Insurance Risk and Ruin* . Cambridge University Press (CUP). ISBN0-521-846404
- De Haan L. & Ferreira A. (2006). *Extreme Value Theory. An Introduction*. Springer, New York.
- Deelstra G. & Plantin G. (2006). *Théorie du Risque et de la Réassurance*. Economica, Paris
- Denuit M. & Charpentier A. (2005). *Mathématiques de l'assurance non-vie : Tarification et provisionnement*. Economica, Paris
- Embrechts P., Claudia K. & Thomas M. (1997). *Modelling Extremal Events For Insurance and Finance*. Springer. 1st ed, New York.
- Fersi K., Benlagha N. & Ben Ammou S. (submitted (2011)). The insurability of risks: A quantitative approach applied to motor insurance. *Insurance: Mathematics and Economics*

- François C. & Constant E. (2009). Les grands principes de l'assurance. Argus de l'assurance, France
- Gnedenko B.V. & Kolmogorov A.N. (1954). Limit Distributions for Sums of Independent Random Variables. Addison-Wesley, MA, Cambridge
- Goldberg D. (1989). Genetic Algorithm In Search, Optimization And Machine Learning. Addison-Wesley
- Goovaerts M., De Vylder F. & Haezendonck J. (1984). Insurance Premiums. North-Holland, VII + 398
- Gollier, C., Hilton D.J. and Raufaste E. (2003). Daniel Kahneman et l'analyse de la décision face au risque, Revue d'Economie Politique, N°113. Vol 3, p. 295-307
- Griselda D. & Guillaume P. (2006). Théorie du risque et réassurance. Economica, Paris
- Hageman J.A., Wehrens R., Van Sprang H.A. & Buydens L.M.C. (2003). Hybrid genetic algorithm tabu search approach for optimising multilayer optical coatings. Analytica Chimica Acta, Vol. 490, p.211-222
- Jean M. R. (1995). Algorithmes génétiques & réseaux de neurones. Hermes, Paris
- Krishnamoorthy C.S., Venkatesh P.P. & Sudarshan R. (2002). Object-oriented framework for genetic algorithms with application to space truss optimization. Journal of Computing in Civil Engineering, Vol. 16, p.66-75
- McNeil A.J. (1997). Estimating the tails of loss severity distributions using extreme value theory. ASTIN Bulletin n°27, p. 117-137
- Necir A. & Boukhetala K. (2004). Estimating the risk-adjusted premium for the largest claims reinsurance covers. COMPSTAT'2004. Proceeding in Computational Statistics. ISBN 3-7908-1554-3, Physica-Verlag, Heidelberg, New York. Springer
- Necir A., Boukhetala K. & Cherfi M. (2005). Estimating the Risk-Adjusted Premium of Loss Using the POT Model. Prépublication, Faculté de Mathématiques d'Alger, Algérie.
- Ohlsson E. & Johansson B. (2010). Non-Life Insurance Pricing with Generalized Linear Models. EAA series
- Picard P. (2002). L'assureur face à la théorie du risque. Risque, Vol. 49, p.61-63
- Partrat C. & Besson J.L. (2005). Assurance non-vie, modélisation, simulation. Economica. Paris
- Petauton P. (2004). Théorie et Pratique de l'Assurance Vie. Dunod
- Rajaev S. & Krishnamoorthy C.S. (1992). Discrete optimization of structures using genetic algorithm. Journal of Structural Engineering, Vol. 118, p.1233-1250
- Rootzen H. & Tajvidi N. (1997). Extreme value statistics and wind storm losses: a case study. Scandinavian Actuarial Journal, Vol. 1, p.70-94. This paper has also appeared in the book "Extremes and Integrated Risk Management" published by Risk Books, Risk Waters Group, (2000), p.215-234

- Tassi P. (2004). Méthodes Statistiques. ESA, Economica, Paris
- Togan V. & Daloglu A. (2006). Optimization of 3D Trusses with Adaptive Approach in Genetic Algorithms. Engineering Structures, Vol. 28, p.1019-1027
- Tosseti A. & al. (2002). Assurance : Comptabilité, Réglementation, Actuariat. Economica, Paris
- Wang S. (1996). Premium calculation by transforming the layer premium density. ASTIN Bulletin, Vol. 26, p.71-92
- Wang S. (2000). A Class of Distorsion Operators for Pricing Financial and Insurance Risks. Journal of Risk and Insurance, Vol. 67, p.15-36
- Wang S. (2002) Pricing of Catastrophe Bonds. In Alternative Risk Strategies, Lane M. (Ed), Risk Books (2003)
- Wang S. (2004). Cat Bond Pricing Using Probability Transforms. p.19-29, Insurance and the State of the Art in Cat Bond Pricing, Geneva Papers 278 (2004), Switzerland
- Wills S.& Sherris M. (2010). Securitization, structuring and pricing of longevity risk. Insurance: Mathematics and Economics, Vol. 46, p.173-185
- Yu X. & Gen M. (2010). Introduction to Evolutionary Algorithms. Springer, Berlin
- Young V. (1999) Discussion of Christofides's Conjecture regarding Wang's Premium Principle. ASTIN Bulletin, Vol. 29, p.191-195
- Youssef H., Sadiq M.S. & Adiche H. (2001). Evolutionary algorithms, simulated annealing and tabu search: a comparative study, Engng. Appl. Artificial Intelligence, Vol. 14, p.167-181
- Zajdenweber D. (2000). Economie des extrêmes, Flammarion, Paris.

Prime des risques extrêmes

#### Biblioweb

Site du Committee of Sponsoring Organizations of the Treadway Commission.

COSO, (2009), *www.coso.org*

Site d'EMB France - Solvabilité II, (2009-2010), *www.emb.com*

Site des dossiers techniques d'information Optimind "Solvabilité II et les modèles internes", *www.optimind.fr*

## Résumé

L'estimateur de la prime ajustée développée par Necir & Boukhetala (2004) est considéré. Afin de minimiser la variance de cet estimateur on a proposé une stratégie originale en formulant un problème d'optimisation non linéaire sous contraintes stochastiques. Pour faire face à la complexité de ce problème on a utilisé un algorithme génétique hybride visant la recherche des solutions globales. Une illustration de cette stratégie sur des données française est développée. Les résultats obtenus fournissent à l'assureur un outil d'aide à la tarification en assurance non vie.