

QUELQUES RELATIONS ENTRE LES PRINCIPAUX MOMENTS REPRÉSENTATIFS DU RISQUE EN ASSURANCE AUTOMOBILE, SUR LA BASE D'UN ÉCHANTILLON SEGMENTÉ

Charles HELBRONNER
Institut des Actuaire¹

I. INTRODUCTION

La connaissance du risque relatif à toute branche de l'assurance IARD nécessite d'évaluer, par garantie, la moyenne et l'écart type de la variable aléatoire que constitue la charge annuelle des sinistres. Ces éléments découlent généralement de paramètres fondamentaux tels qu'en particulier les deux premiers moments associés à ses différentes composantes. L'appréhension, sinon la mesure exacte de telles grandeurs donne alors accès à un outil utile d'aide à la décision aux entreprises d'assurance, notamment en matière de tarification.

Cet article concerne, plus particulièrement, le cas de l'assurance automobile. Les composantes de la charge annuelle des sinistres sont, dans ce domaine :

- le nombre annuel de sinistres par police,
- le coût par sinistre.

Les valeurs moyennes empiriques de ces variables aléatoires sont des données facilement accessibles, souvent par grande catégorie de conducteurs, ne serait-ce que par simple consultation de documents statistiques généralement édités par la profession. En revanche, les écarts types ne connaissent guère un degré de diffusion aussi large. Leur évaluation nécessite ainsi la mise en évidence et l'étude des distributions statistiques correspondantes. La possibilité d'estimer, du moins approximativement dans un premier temps, ces grandeurs de façon rapide et simple, à partir des valeurs moyennes, constitue donc un avantage certain.

Dans cet objectif, à travers une étude globale du risque automobile réalisée par nos soins, nous avons tenté de dégager des résultats qui tendent effectivement à mettre en évidence, pour chacune des principales garanties de l'assurance automobile, des relations invariantes entre les deux premiers moments (espérance mathématique et variance), aussi bien pour le nombre annuel de sinistres par police que pour le coût d'un sinistre. L'approche a été menée par garantie, à partir d'un échantillon de polices assurant des véhicules de première catégorie, segmenté en classes plus homogènes que la population globale des assurés.

Les résultats de ce travail expérimental nous donnent en effet, selon le mot de Condorcet, un *motif de croire* à l'existence de telles relations. Afin de présenter ces éléments dans un ordre logique et progressif, nous avons adopté le plan suivant :

¹ Note de l'auteur : Cet article a été réalisé à partir de mon mémoire d'actuariat, intitulé *Etude du risque et évaluation des besoins en fonds propres en assurance automobile, sur un échantillon segmenté*. Je tiens particulièrement à remercier M. Jean-Luc Besson, de la FFSA, qui a été mon directeur de mémoire, pour ses conseils fort utiles. Je remercie également les rapporteurs qui, selon l'usage, ont pris connaissance du présent article.

Le chapitre qui suit (**chapitre II**) présente les hypothèses retenues, les données utilisées ainsi que la méthodologie conçue pour modéliser les lois de probabilité correspondant au nombre annuel de sinistres par police, comme au coût des sinistres pour chaque garantie.

Le **chapitre III** met en évidence une relation entre la moyenne et la variance associés au nombre annuel de sinistre par police. Nous verrons notamment à cette occasion que cette relation s'avère très simple et d'une nature semblable quelle que soit la garantie.

Le **chapitre IV** tente de mettre en évidence des relations entre la moyenne et l'écart type associés aux coûts par sinistre. Il est alors possible de dégager des relations relativement simples, mais dont la nature peut différer d'une garantie à l'autre.

Le **chapitre V** opère une synthèse des conclusions des deux chapitres précédents. Il s'agit en effet de pouvoir évaluer l'écart type de la charge annuelle des sinistres par police en fonction des résultats ainsi obtenus. Nous tentons alors de montrer dans quelles conditions cela peut être réalisé, puis de modéliser, à partir de l'acquis des relations mises en évidence au cours des précédents chapitres, l'évaluation de cet écart type (noté σ). Nous abordons également quelques propriétés de la fonction qui permet de relier σ , et en particulier $\frac{\sigma}{P}$, (P étant la prime pure) aux deux variables de base que sont la fréquence moyenne f et le coût moyen C .

Enfin, en guise de conclusion le **chapitre VI**, tout en rappelant l'intérêt opérationnel d'une telle démarche, suggère quelques pistes pour la pérenniser.

II. HYPOTHESES RETENUES – DONNEES UTILISEES – AJUSTEMENTS REALISES

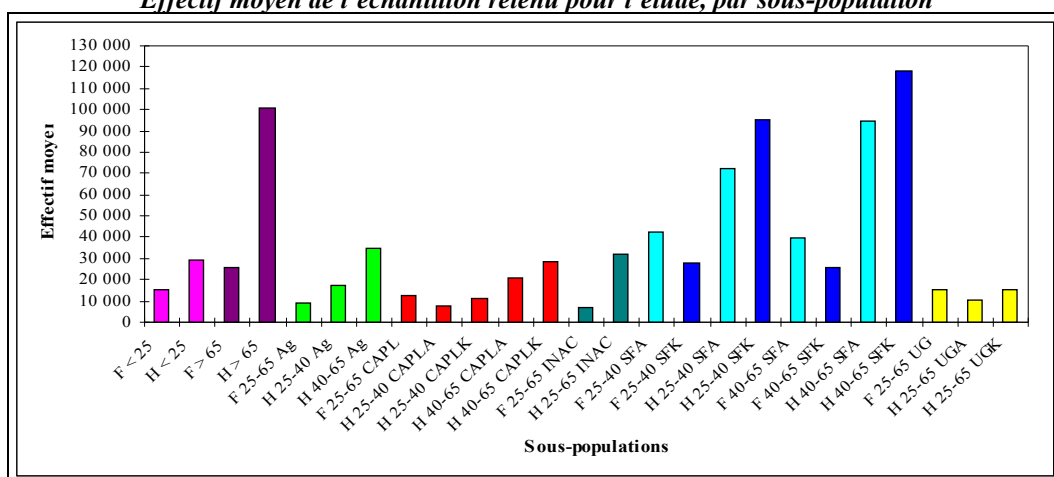
II.1. Hypothèses retenues et données utilisées

L'étude a été menée sur la base d'un échantillon représentant environ 3,6% du parc des véhicules de première catégorie, recueilli auprès de sociétés avec intermédiaires. Il nous a paru nécessaire, pour affiner notre approche et pour disposer d'un nombre suffisant d'observations, de mener nos travaux sur la base de sous-populations plus homogènes que la population globale de l'échantillon. Dans ce but, ce dernier a été segmenté selon quatre critères socio-démographiques :

- l'âge, en distinguant quatre tranches d'âge : moins de 25 ans, 25 à 40 ans, 40 à 65 ans, plus de 65 ans,
- le sexe,
- la catégorie socio-professionnelle, en distinguant les agriculteurs (Ag), les commerçants, artisans et professions libérales (CAPL), les inactifs (INAC), les salariés et fonctionnaires (SF) et les usages généraux (UG),
- la puissance tarifaire du véhicule, en distinguant deux classes : puissance faible (désignée ici par A) et puissance élevée (désignée ici par K).

La segmentation aboutit, après certains regroupements de catégories d'effectif trop faible et peu significatives si on les considère individuellement, à 25 sous-populations, indiquées dans le graphique ci-dessous.

Graphique n° 1
Effectif moyen de l'échantillon retenu pour l'étude, par sous-population



L'étude est réalisée de façon individualisée pour chacune des principales garanties de l'assurance automobile :

- la responsabilité civile, en distinguant les sinistres matériels d'une part et les sinistres corporels d'autre part,
- la garantie dommages accidents,
- la garantie vol,
- la garantie bris de glaces,

sur une période couvrant les exercices de 1992 à 1997.

Pour chacune des sous-populations, chaque garantie et chaque exercice, ont été identifiées au niveau de l'échantillon les distributions statistiques empiriques en nombre de sinistres, et en coût des sinistres par tranche de coût. Les coûts ont été évalués en francs constants de 1997. Les distributions, aussi bien en nombre de sinistres que par tranche de coût, ont été ajustées à des lois de probabilité usuelles, testées selon le critère du χ^2 .

Ainsi, ont été testées :

- pour modéliser le nombre annuel de sinistres par police des lois discrètes : loi binomiale, loi de Poisson, loi binomiale négative,
- pour modéliser les coûts par sinistre des lois continues : loi gamma, loi log-normale et loi de Weibull.

Dans chaque cas, la loi et les paramètres associés donnant le meilleur ajustement (valeur la plus faible du χ^2) a été retenue.

En ce qui concerne le cas particulier des sinistres corporels en responsabilité civile, les ajustements ont également été réalisés sur la base d'un écrêtement des coûts de sinistres au-delà de

500 KF, seuil qui nous a semblé le plus pertinent, compte tenu des distributions observées. Il s'avère en fait que les résultats ne sont pas fondamentalement différents de ce que l'on obtient en exploitant les données non écrêtées.

Les sous-populations qui ont été constituées sont de taille variable ; certaines d'entre elles présentent un effectif beaucoup moins fourni que d'autres. Pour ces dernières, les propriétés découlant de la loi des grands nombres pourraient éventuellement en être affectées. Il apparaît toutefois que la taille de ces sous-populations leur conserve encore un niveau de fiabilité satisfaisant, pour ce qui concerne les relations entre les principaux moments. En effet, les résultats qui seraient obtenus en ôtant ces sous-populations du dispositif expérimental ne varient pas de façon significative en ordre de grandeur (au niveau des valeurs de $\frac{V_f}{f}$ et de $\frac{\sum C}{C}$) par rapport à ceux qui découlent de la prise en compte de l'ensemble des sous-populations. Les données liées aux sous-populations dont l'effectif est le moins élevé ne semblent pas déséquilibrer les résultats expérimentaux ; elles tendent plutôt, en s'ajoutant au nombre d'observations, à conforter ces résultats.

II.2. Résultat des ajustements

Les ajustements réalisés font notamment apparaître que le type de loi correspondant au meilleur ajustement est, pour chaque garantie, invariant dans l'espace (entre les différentes sous-populations) et dans le temps – à l'exception toutefois du cas particulier de la garantie bris de glaces, pour le nombre annuel de sinistres.

Les résultats de ces ajustements sont présentés dans le tableau ci-dessous.

Tableau n° 1
Récapitulation des lois de probabilité retenues pour chaque garantie,
pour modéliser le nombre annuel de sinistres par police et le coût des sinistres

Garantie	Nombre annuel de sinistres	Coût d'un sinistre
Responsabilité civile	Sinistres matériels	Binomiale négative
	Sinistres corporels non écrêtés	Binomiale négative
	Sinistres corporels écrêtés après 500 KF	Binomiale négative
Dommages accidents	Binomiale négative	Gamma
Vol	Binomiale négative	Log-normale
Bris de glaces	Binomiale négative -> Binomiale	Gamma

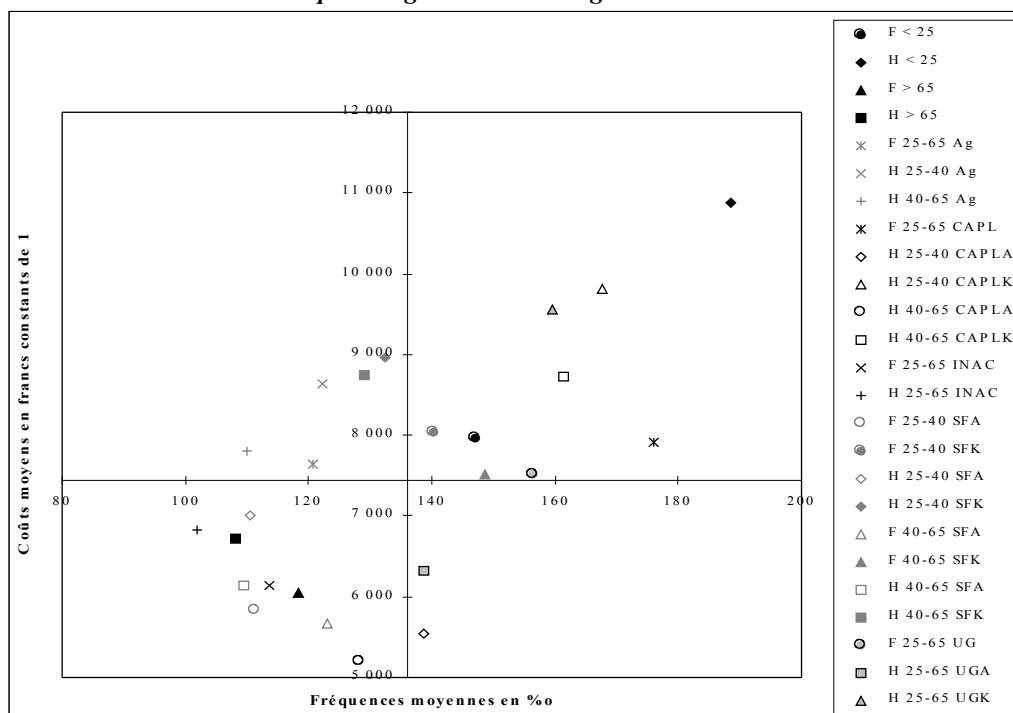
La suite de cet article porte sur les moments d'ordre 1 et 2 correspondant aux distributions ajustées par rapport aux lois de probabilité indiquées ci-dessus.

Nous pouvons noter en particulier qu'en ce qui concerne le nombre annuel de sinistres, le meilleur ajustement que nous avons pu réaliser, à partir de lois discrètes usuelles, correspond généralement à des lois binomiales négatives - plutôt qu'à une loi de Poisson, même si parfois (cas

de la responsabilité civile des sinistres corporels), les lois binomiales négatives obtenues sont relativement proches de lois de Poisson.

Il faut également remarquer que l'on observe, pour chaque garantie, une importante dispersion des fréquences et des coûts moyens entre les sous-populations, ainsi qu'une tendance générale à la baisse de ces valeurs entre 1992 et 1997. Le graphique suivant montre *pour l'exemple de la garantie dommages-accidents*, qui illustre une situation commune à l'ensemble des garanties faisant l'objet de notre étude, la diversité des situations qui peuvent se présenter en fonction des sous-populations, au niveau de la fréquence et du coût moyens des sinistres

Graphique n° 2
Croisement des fréquences moyennes et des coûts moyens des sinistres par sous-population, pour la garantie Dommages-accidents



Malgré la disparité observée, nous verrons dans les chapitres qui suivent que les deux premiers moments des lois de probabilité retenues peuvent en première approximation être reliés, pour une garantie donnée, par des **relations invariantes**, indépendantes de la sous-population et de l'exercice considérés. En effet, V_f étant la variance associée au nombre annuel de sinistres par police et ΣC l'écart type du coût des sinistres, il est possible d'identifier un regroupement des coefficients $\frac{V_f}{f}$ et $\frac{\Sigma C}{C}$, qui se prêtent particulièrement bien à cette étude, autour de droites ou de courbes, à partir de nuages de points qui permettent d'établir une formule de régression - linéaire ou non - relativement pertinente.

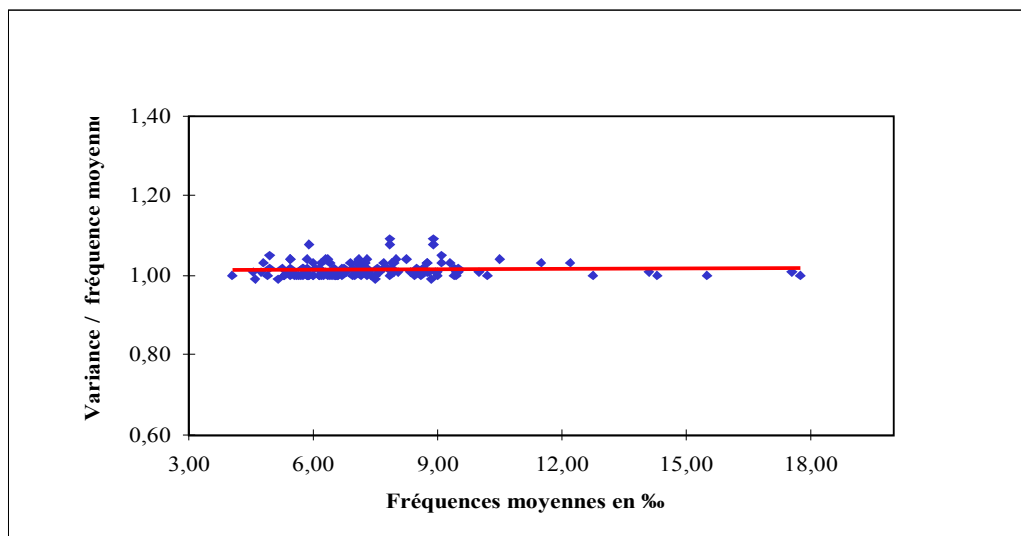
III. IDENTIFICATION DES RELATIONS ENTRE MOYENNE ET VARIANCE POUR LE NOMBRE ANNUEL DE SINISTRES PAR POLICE

A partir des éléments précédents, nous avons tenté d'estimer, à partir d'une approche empirique, s'il était possible de dégager des relations entre variance et espérance mathématique pour ce qui concerne le nombre annuel de sinistres par police.

Cette recherche nous a conduit à constater que, pour chaque garantie, existe une relation invariante entre variance et moyenne. En effet, comme l'illustre l'exemple du graphique ci-après (il s'agit du cas de la garantie responsabilité civile des sinistres corporels, mais la situation est analogue pour les autres garanties), les points d'abscisse f (fréquence moyenne) et d'ordonnée $\frac{V_f}{f}$ (variance associée / fréquence moyenne) se regroupent, pour chaque garantie, autour d'une droite horizontale.

Dans l'hypothèse où les lois de distribution du nombre de sinistres annuel par police seraient ajustables à des lois de Poisson, la propriété $\frac{V_f}{f} = 1$ (après approximation) en découlerait immédiatement, par définition, pour toutes les sous-populations. Dans notre cas de figure, la propriété que $\frac{V_f}{f}$ soit une constante indépendamment de l'exercice et de la sous-population reste vérifiée, malgré le caractère binomial négatif des lois de probabilité correspondant aux ajustements.

Graphique n° 3
Garantie : responsabilité civile des sinistres corporels
Représentation du ratio $\frac{V_f}{f}$ en fonction de f



Ainsi, pour chaque garantie, après ajustement par rapport à des droites de régression, se dessine approximativement une relation entre V_f et f , qui s'avère être de type :

$$V_f = v_i f$$

i étant un indice représentatif de la garantie considérée.

N.B. : Dans la suite de cet article, les indices 0,1,2,3 et 4 correspondent respectivement à la responsabilité civile des sinistres matériels, à la responsabilité civile des sinistres corporels, à la garantie dommages-accidents, à la garantie vol et à la garantie bris de glaces. L'indice 1e sera par ailleurs utilisé à partir du chapitre IV pour désigner les paramètres correspondant aux sinistres corporels de responsabilité civile, après écrêtement des coûts au-delà du seuil de 500 MF.

Le tableau suivant indique l'estimation des coefficients de proportionnalité v_i que l'étude a permis de réaliser pour chaque garantie, ainsi que la demi-amplitude des intervalles de confiance associés, à un niveau de risque de 5% .

Tableau n° 2

Valeurs estimées de $\frac{V_f}{f}$

et demi-amplitude ΔIC des intervalles de confiance au risque 5% associés par garantie

Garantie		$\frac{V_f}{f}$	Demi-amplitude ΔIC	$\frac{\Delta IC}{v_i}$
Responsabilité civile	Sinistres matériels	1,054	0,396 10 ⁻²	0,38%
	Sinistres corporels	1,016	0,308 10 ⁻²	0,30%
Dommages accidents		1,050	0,466 10 ⁻²	0,44%
Vol		1,057	0,473 10 ⁻²	0,45%
Bris de glaces		1,007	0,384 10 ⁻²	0,38%

D'après l'amplitude des intervalles de confiance, la qualité de l'ajustement s'avère ici tout à fait satisfaisante.

En conclusion, si l'on utilise les résultats ainsi obtenus par régression, la connaissance des fréquences moyennes *observées* pour chaque garantie et chaque sous-population suffit, pour peu évidemment que l'effectif des sous-populations étudiées soit assez élevé, à estimer la fréquence moyenne par la méthode des moments, et la variance associée à l'aide des relations mises en évidence précédemment.

IV. IDENTIFICATION DES RELATIONS ENTRE MOMENTS D'ORDRE 1 ET 2 POUR LES COÛTS

De même que pour le nombre annuel de sinistres, nous avons également tenté d'identifier à partir des résultats obtenus s'il était possible de mettre en évidence des relations entre la moyenne et l'écart type pour ce qui concerne les coûts des sinistres.

Nous observons dans ce cas que, pour chaque garantie, existe – approximativement - une relation invariante entre l'écart type et la moyenne. En effet, les points d'abscisse C (coût moyen) et d'ordonnée $\frac{\Sigma C}{C}$ (écart type associé / coût moyen) se regroupent, pour chaque garantie, autour d'une droite ou d'une courbe croissante concave, selon les garanties.

En ce sens, $\frac{\Sigma C}{C}$ peut être approché par ajustement, par une relation de type :

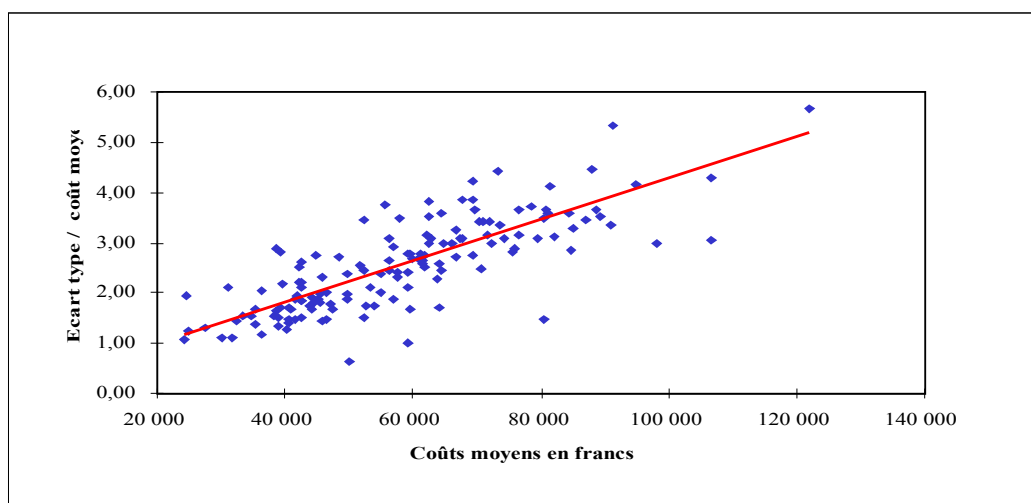
$$\frac{\Sigma C}{C} = a_i C + b_i$$

ou bien

$$\frac{\Sigma C}{C} = a_j \ln(C) + b_j$$

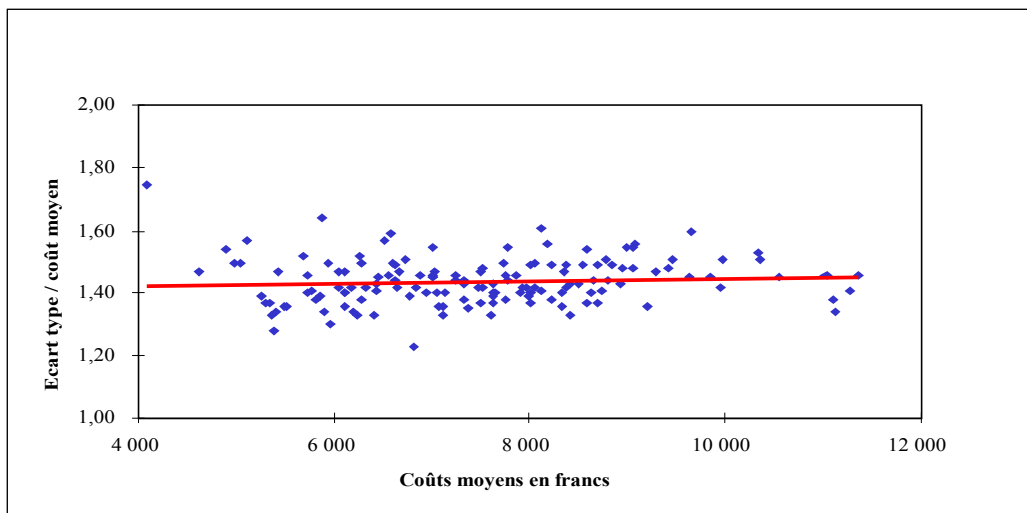
selon la garantie considérée, comme le montrent les graphiques qui suivent. Ces graphiques illustrent successivement l'exemple des garanties responsabilité civile des sinistres corporels, dommages-accidents et vol.

Graphique n° 4
Garantie : responsabilité civile des sinistres corporels
Représentation du ratio $\frac{\Sigma C}{C}$ en fonction de C



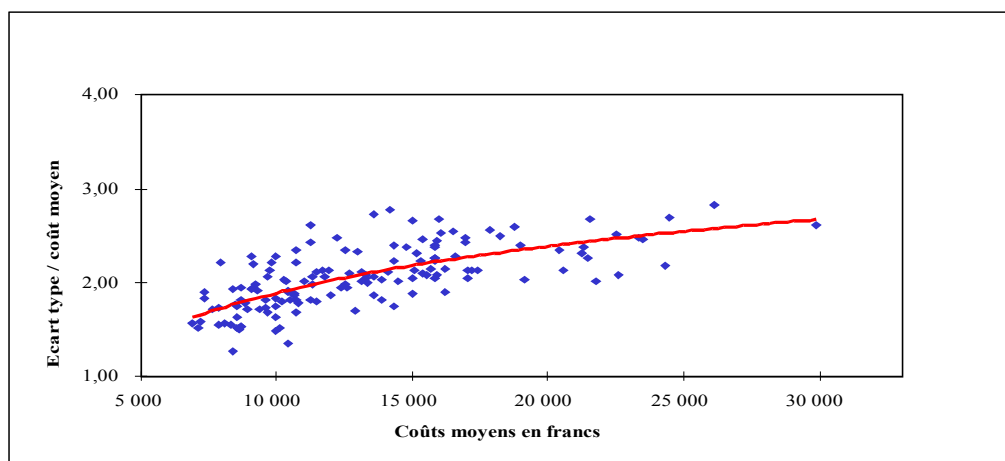
Au niveau de la garantie responsabilité civile des sinistres corporels, le nuage de points se rassemble autour d'une droite de pente positive, dans le cas des sinistres non écrêtés comme dans le cas des sinistres écrêtés au-delà du seuil de 500 KF. Une configuration de type semblable apparaît pour la responsabilité civile des sinistres matériels.

Graphique n° 5
Garantie : dommages accidents
Représentation du ratio $\frac{\Sigma C}{C}$ en fonction de C



Dans le cas de la garantie dommages accidents, le nuage de points se regroupe autour d'une droite quasiment horizontale (la pente de la droite obtenue par régression linéaire est pratiquement négligeable). Il s'agit en fait d'un cas particulier extrême de la situation que l'on peut observer en responsabilité civile (relation affine entre $\frac{\Sigma C}{C}$ et C).

Graphique n° 6
Garantie : vol
Représentation du ratio $\frac{\Sigma C}{C}$ en fonction de C



Pour la garantie vol, le nuage de points se regroupe autour d'une courbe logarithmique. C'est une situation de même nature qui se produit pour la garantie bris de glaces.

Il faut remarquer, par ailleurs, que, compte tenu des effets de dispersion observés sur notre échantillon, une relation affine pourrait également être admissible pour modéliser pour ces deux garanties le lien entre $\frac{\Sigma C}{C}$ et C . La relation logarithmique s'ajuste cependant mieux, tout en s'exprimant de manière simple, dans le cadre de notre échantillon, aux observations relevées, surtout pour la garantie Vol.

Le tableau qui suit présente les relations qui peuvent être obtenues par ajustement pour chaque garantie, avec les valeurs correspondantes des coefficients, a_i , b_i , a_j et b_j des droites ou des courbes de régression.

Tableau n° 3
Expressions estimées de $\frac{\Sigma C}{C}$ en fonction de C (exprimé en F) par garantie

Garantie		$\frac{\Sigma C}{C}$	a_i, a_j	b_i, b_j
Responsabilité civile	Sinistres matériels	$\Sigma c/C = a_0 C + b_0$	$6,28 \cdot 10^{-5}$	0,7242
	Sinistres corporels non écrêtés	$\Sigma c/C = a_1 C + b_1$	$4,13 \cdot 10^{-5}$	0,1517
	Sinistres corporels écrêtés après 500 KF	$\Sigma c/C = a_1 e C + b_1 e$	$4,71 \cdot 10^{-5}$	-0,1560
Dommages accidents		$\Sigma c/C = a_2 C + b_2$	$0,36 \cdot 10^{-5}$	1,4126
Vol		$\Sigma c/C = a_3 \ln(C) + b_3$	0,7152	-4,6916
Bris de glaces		$\Sigma c/C = a_4 \ln(C) + b_4$	0,2637	-1,3554

Pour la garantie dommages accidents, le coefficient a_2 de la droite de régression est particulièrement faible. Le test de Student, appliqué par rapport à la moyenne des observations de $\frac{\Sigma C}{C}$ montre que l'on peut adopter une valeur de a_2 nulle et une valeur de b_2 égale à la moyenne observée (soit 1,439), au risque de 5%. La demi-amplitude de l'intervalle de confiance correspondant à l'estimation de la constante b_2 est alors de 0,84% de la moyenne observée de $\frac{\Sigma C}{C}$, valeur retenue pour b_2 .

De façon générale, la qualité de l'ajustement est nettement moins bonne pour les coûts que pour le nombre annuel de sinistres, comme le montre le tableau ci-dessous (n désignant, pour chaque garantie, le coefficient de corrélation relatif à la courbe de régression). Cependant, le niveau d'approximation retenu reste relativement acceptable.

Tableau n° 4

**Amplitude des intervalles de confiance au risque de 5% correspondant aux valeurs retenues
pour l'estimation de $\frac{\Sigma C}{C}$ en fonction de C**

Garantie		$\frac{\Sigma C}{C}$ moy	Demi-amplitude ΔIC	$\Delta IC / \frac{\Sigma C}{C}$ moy
Responsabilité civile	Sinistres matériels	1,135	$0,972 \cdot 10^{-2}$	0,86%
	Sinistres corporels non écrêtés	2,567	$9,590 \cdot 10^{-2}$	3,74%
	Sinistres corporels écrêtés après 500 KF	2,536	$9,033 \cdot 10^{-2}$	3,56%
Dommages accidents		1,439	$1,216 \cdot 10^{-2}$	0,84%
Vol		2,060	$3,841 \cdot 10^{-2}$	1,86%
Bris de glaces		0,602	$0,464 \cdot 10^{-2}$	0,77%

Précisons, d'une façon générale, que dans la mesure où nous ne travaillons ici, dans le cadre de l'échantillon, que sur les sinistres survenus -et non sur l'ensemble des polices de l'échantillon-, le nombre d'observations disponibles pour la modélisation des distributions statistiques des coûts est nettement inférieur au nombre des polices, ce qui est de nature à détériorer la précision des ajustements par rapport aux lois de probabilité retenues, au regard de celle que l'on peut obtenir pour le nombre annuel de sinistres par police, où l'on peut effectivement travailler sur l'ensemble des assurés. La dispersion des points autour des courbes d'ajustement constatée en ce qui concerne les coûts de sinistres peut ainsi s'expliquer.

Néanmoins, les éléments qui se dégagent des résultats recueillis suggèrent tout de même, pour chaque garantie, une tendance. Il en ressort en tout état de cause un certain regroupement des résultats, obtenus sur la base de classes de populations plus homogènes que la population globale, autour de droites ou de courbes très classiques. Cela incite, sinon à établir trop hâtivement des propriétés absolues et immuables, à approfondir du moins la recherche sur la base d'échantillons plus étendus.

Enfin, il est à remarquer, comme le montre le tableau suivant, que, par garantie, la nature des relations identifiées entre l'écart type et la moyenne des coûts par sinistre apparaît comme a priori indépendante de la loi de probabilité ayant servi à l'ajustement des distributions statistiques.

Tableau n° 5
Rapprochement par garantie entre la loi de probabilité retenue pour modéliser le coût
d'un sinistre et la formule permettant d'estimer $\frac{\Sigma C}{C}$ en fonction de C

Garantie		Loi du coût d'un sinistre	$\frac{\Sigma C}{C}$
Responsabilité civile	Sinistres matériels	Gamma	$\Sigma C/C = a_0 C + b_0$
	Sinistres corporels non écrêtés	Log-normale	$\Sigma C/C = a_1 C + b_1$
	Sinistres corporels écrêtés après 500 KF	Log-normale	$\Sigma C/C = a_1 e C + b_1 e$
Dommages accidents		Gamma	$\Sigma C/C = a_2 C + b_2$
Vol		Log-normale	$\Sigma C/C = a_3 \ln(C) + b_3$
Bris de glaces		Gamma	$\Sigma C/C = a_4 \ln(C) + b_4$

En conclusion de ce chapitre, de la même façon que pour le nombre annuel de sinistres par police, la connaissance des coûts moyens *observés* pour chaque garantie et chaque sous-population peut suffire en toute première approximation pour estimer le coût moyen par la méthode des moments, et l'écart type associé à l'aide des relations identifiées précédemment.

V. EVALUATION DE L'ECART TYPE DE LA CHARGE ANNUELLE DES SINISTRES PAR POLICE A PARTIR DE LA FREQUENCE ET DU COUT MOYENS

L'utilisation des relations mises en évidence aux chapitres précédents permettent une estimation – du moins théorique – de l'écart type σ de la charge annuelle des sinistres par police, en fonction de f et de C .

En effet, σ est donné, comme cela est démontré dans l'ouvrage de M. Petauton «Théorie de l'assurance dommage» par la relation :

$$\sigma^2 = \Sigma_j^2 C^2 + f \Sigma_C^2.$$

Les relations permettant d'estimer Σ_f en fonction de f et Σ_C en fonction de C permettent donc d'estimer σ , et en particulier $\frac{\sigma}{P}$, en fonction des deux seules variables f et C . En transposant ces relations à la formule ci-dessus, nous obtenons alors les expressions suivantes, les coefficients a_i , b_i , a_j et b_j dépendant, comme nous l'avons vu, de chaque garantie :

- Pour les garanties de responsabilité civile et de dommages accidents :

$$\frac{\sigma}{P} = \frac{\sqrt{v_i + (a_i C + b_i)^2}}{\sqrt{f}}$$

(avec $a_2 = 0$ pour la garantie dommages accidents).

- Pour les garanties vol et bris de glaces :

$$\frac{\sigma}{P} = \frac{\sqrt{v_j + (a_j \ln(C) + b_j)^2}}{\sqrt{f}}$$

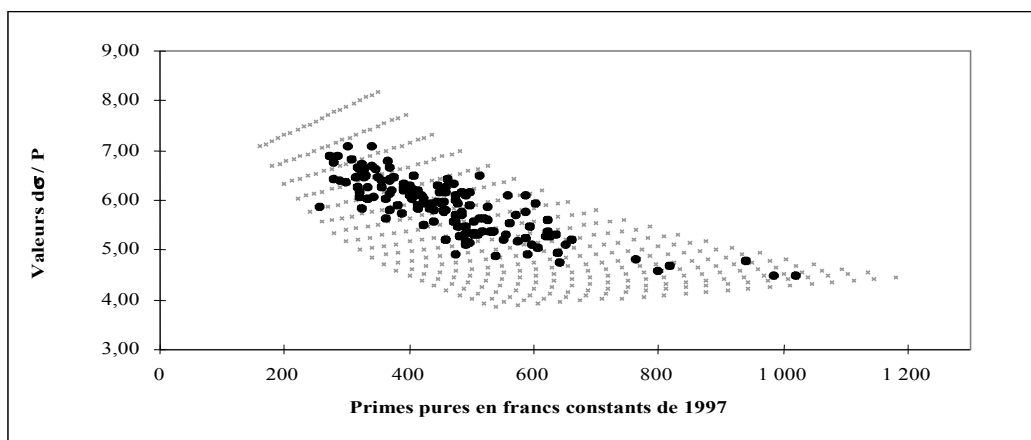
$\frac{\sigma}{P}$ s'écrit alors, dans tous les cas, sous la forme $\frac{\sigma}{P} = \frac{\varphi_C(C)}{\varphi_f(f)}$, φ_f et φ_C étant des fonctions croissantes.

A ce stade, on peut être tenté de dégager une relation directe entre les deux premiers moments de la charge annuelle des sinistres par police, qui se traduirait par une relation de type $\sigma = \varphi(P)$, P étant la prime pure ($P = fC$). Une telle entreprise n'est pas du tout aisée, dans le cadre de notre modélisation. En effet, les valeurs de $\frac{\sigma}{P}$ obtenues empiriquement au niveau des sous-populations de l'échantillon correspondent généralement à des nuages de points par rapport à P , plutôt qu'à des courbes franchement affirmées.

Les graphiques ci-dessous présentent de façon superposée, à titre d'exemple, pour la responsabilité civile des sinistres matériels, la responsabilité civile des sinistres corporels et le vol :

- les nuages de points, d'abscisse P et d'ordonnée $\frac{\sigma}{P}$, correspondant au résultat théorique des relations entre les deux premiers moments, identifiées précédemment, dans la limite des spectres des primes pures observées par garantie au niveau de notre échantillon (points de teinte claire sur le graphique) ; à travers ces nuages de points peuvent se distinguer des courbes iso-fréquence (qui apparaissent comme un faisceau de courbes croissantes) et des courbes iso-coût (qui apparaissent comme un faisceau de courbes décroissantes convexes).
- les mêmes nuages de points obtenus de façon empirique, avec les résultats dégagés à partir de notre échantillon (points de teinte foncée sur le graphique).

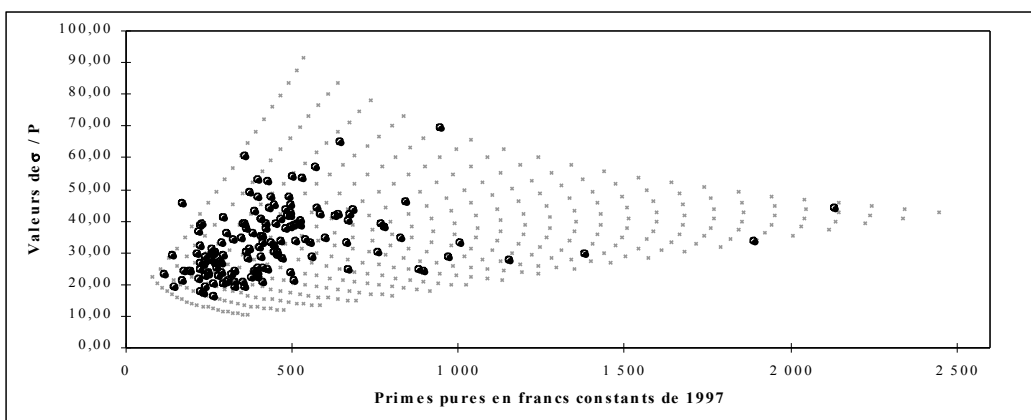
Graphique n° 7
Garantie : Responsabilité civile des sinistres matériels
Superposition des nuages de points théoriques et empiriques
représentant les valeurs de $\frac{\sigma}{P}$ en fonction de P



Les courbes iso-fréquence du graphique des points théoriques sont des courbes croissantes en fonction de la prime pure (et donc ici du coût moyen), d'allure proche de celle de droites affines.

Les courbes iso-coût du graphique des points théoriques sont des courbes décroissantes et convexes.

Graphique n° 8
Garantie : Responsabilité civile des sinistres corporels
Superposition des nuages de points théoriques et empiriques
représentant les valeurs de $\frac{\sigma}{P}$ en fonction de P

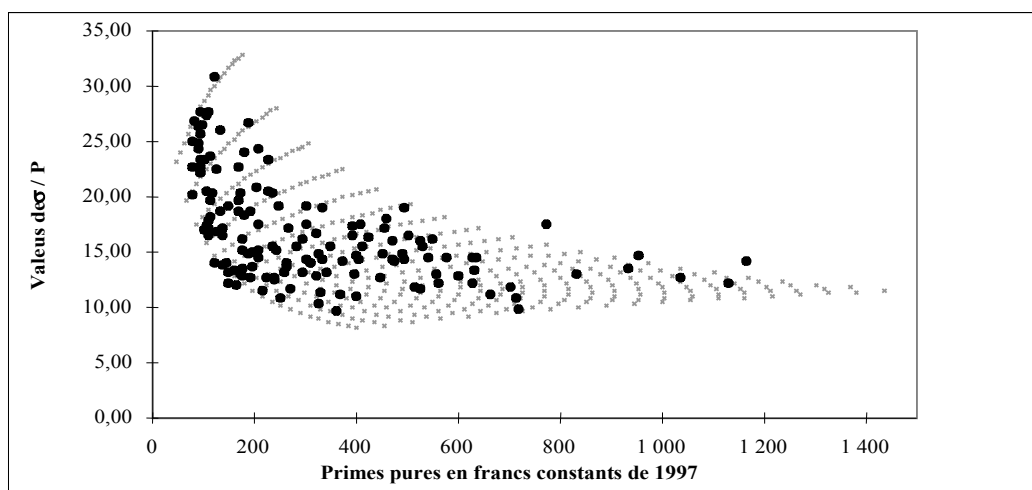


Ici, comme pour la responsabilité civile des sinistres matériels, les courbes iso-fréquence du graphique des points théoriques sont des courbes croissantes en fonction de la prime pure (et donc ici du coût moyen), d'allure proche de celle de droites affines. La pente de ces courbes, en apparence beaucoup plus affirmée sur le graphique que pour la responsabilité civile des sinistres matériels, tient (en partie en raison de l'échelle du graphique) à ce que le spectre des primes pures et surtout des valeurs de $\frac{\sigma}{P}$ est nettement plus étendu pour la responsabilité civile des sinistres corporels.

Les courbes iso-coût du graphique des points théoriques sont également des courbes décroissantes et convexes.

En ce qui concerne la garantie dommages accidents, le nuage de points théorique correspondant présente une configuration analogue à celle de la responsabilité civile des sinistres matériels, les courbes iso-fréquence étant alors des droites horizontales.

Graphique n° 9
Garantie : vol
Superposition des nuages de points théoriques et empiriques
représentant les valeurs de $\frac{\sigma}{P}$ en fonction de P



Les courbes iso-fréquence du graphique des points théoriques sont des courbes croissantes et concaves en fonction de la prime pure (et donc ici du coût moyen), en raison de l'expression de $\frac{\sigma}{P}$ en fonction de $\ln(C)$.

Les courbes iso-coût du graphique des points théoriques sont, ici encore, des courbes décroissantes et convexes.

Le nuage de points théorique correspondant à la garantie bris de glaces présente une configuration analogue à celui de la garantie vol.

Il apparaît ainsi que :

- les nuages de points empiriques s'insèrent dans le domaine délimité par les nuages de points théoriques, ce qui montre que, dans le cas de notre échantillon, l'expérience et la théorie ne se contredisent pas,
- les nuages de points, théoriques comme empiriques, montrent qu'il est possible d'obtenir plusieurs valeurs de σ avec le même niveau P de prime pure, ce qui infirme l'hypothèse d'une relation directe entre P et σ .

En fait, une relation directe entre σ et P supposerait, au niveau des valeurs obtenues par ajustement des distributions statistiques, une forte corrélation entre f et C , ce qui n'apparaît guère au vu des résultats de notre étude.

Dans ces conditions, $\frac{\sigma}{P}$ apparaît bien comme une fonction de deux variables, à savoir f et C , qu'il est plus opportune de représenter par une surface.

Les graphiques qui suivent reproduisent à titre d'exemple de telles surfaces, pour la responsabilité civile des sinistres matériels, la responsabilité civile des sinistres corporels et le vol, qui illustrent les propriétés des fonctions $\frac{\sigma}{P}$ des deux variables f et C . Nous allons évoquer ci-après ces propriétés.

Nous remarquons que, quelle que soit la garantie, $\frac{\sigma}{P}$ est une fonction décroissante de f , en $\frac{1}{\sqrt{f}}$, et croissante de C , selon des modalités différentes en fonction des garanties (à l'exception de la garantie dommages accidents, où $\frac{\sigma}{P}$ apparaît comme indépendant de C). Dans le cadre de chaque garantie, les valeurs de $\frac{\sigma}{P}$ sont donc d'autant plus faibles que l'on se trouve dans une configuration de fréquences élevées et de coûts moyens faibles, et donc en revanche d'autant plus élevées que les fréquences sont basses et les coûts moyens élevés.

De plus, l'élasticité E_f de $\frac{\sigma}{P}$ par rapport à f est toujours égale à $-\frac{1}{2}$, ce qui découle directement de la relation de dépendance, dans tous les cas, de $\frac{\sigma}{P}$, par rapport à $\frac{1}{\sqrt{f}}$.

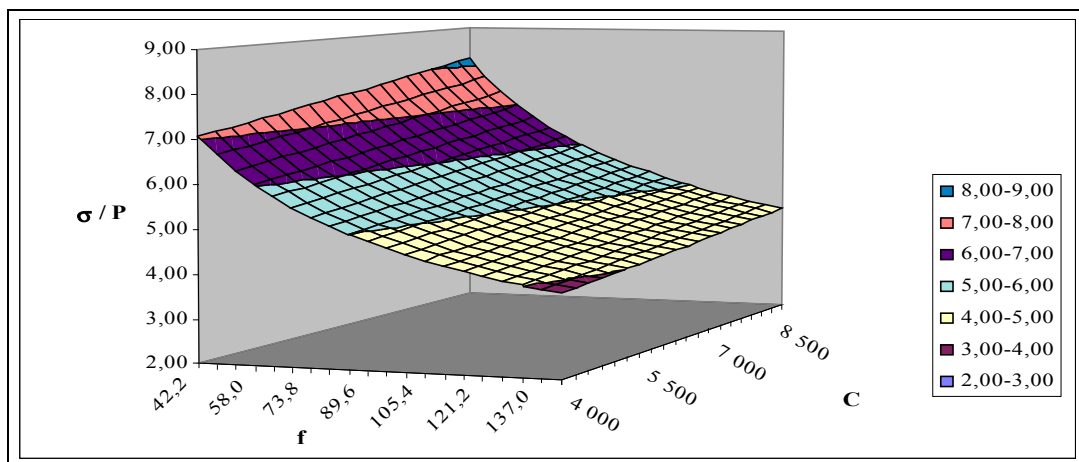
$$E_f = \frac{f}{\sigma/P} \frac{\partial(\sigma/P)}{\partial f} = -\frac{1}{2}.$$

Enfin, nous pouvons noter que la décroissance de $\frac{\sigma}{P}$ en fonction de f est, dans tous les cas, d'autant plus accentuée que C est élevé, comme le montre de façon immédiate l'expression de $\frac{\sigma}{P}$ mise en évidence précédemment.

Graphique n° 10

Garantie : Responsabilité civile des sinistres matériels

Représentation estimée de $\frac{\sigma}{P}$ en fonction de f et de C



Pour la responsabilité civile des sinistres matériels, les valeurs théoriques de $\frac{\sigma}{P}$ se situent entre 3,8 et 8,2. Nous nous trouvons dans un cas où les fréquences sont relativement élevées et les coûts relativement faibles (par rapport à d'autres garanties).

Les dérivées partielles de $\frac{\sigma}{P}$ par rapport aux deux variables f et C s'expriment comme suit, avec $i = 0$, selon la codification que nous avons adoptée pour les garanties (cf. tableaux n° 2 et 3) et $v_0 = 1,054$, $a_0 = 6,28 \cdot 10^{-5}$, $b_0 = 0,7272$.

$$\frac{\partial(\sigma/P)}{\partial f} = - \frac{\sqrt{v_i + (a_i C + b_i)^2}}{2(\sqrt{f})^3}$$

$$\frac{\partial(\sigma/P)}{\partial C} = \frac{a_i(a_i C + b_i)}{\sqrt{f(v_i + (a_i C + b_i)^2)}}.$$

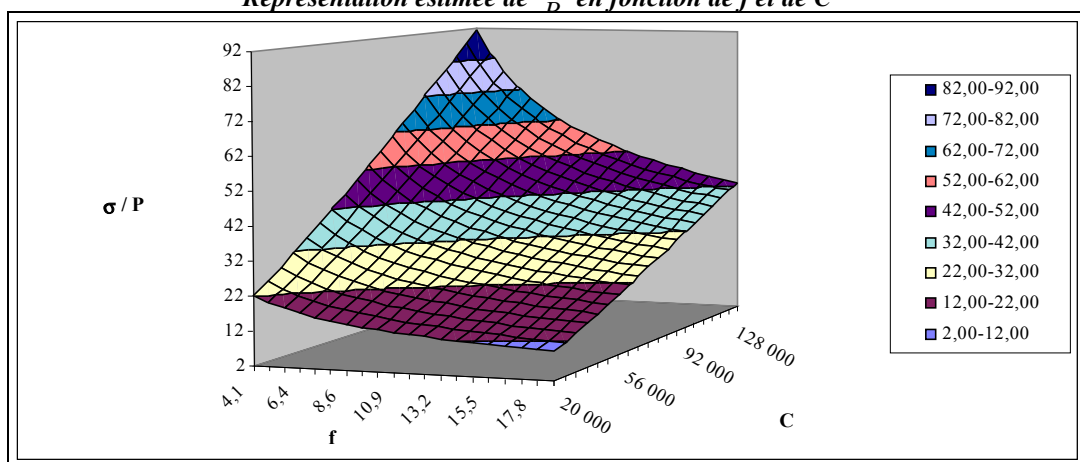
La croissance de $\frac{\sigma}{P}$ en fonction de C s'atténue à mesure que f augmente. L'expression de $\frac{\partial(\sigma/P)}{\partial C}$ est une fonction croissante de C , qui prend la valeur $\frac{a_i b_i}{\sqrt{f(v_i + b_i^2)}}$ pour $C = 0$ et tend vers $\frac{1}{\sqrt{f}}$ quand C tend vers l'infini.

L'élasticité Ec de $\frac{\sigma}{P}$ par rapport à C s'écrit comme suit.

$$EC = \frac{C}{\sigma/P} \frac{\partial(\sigma/P)}{\partial C} = \frac{a_i C (a_i C + b_i)}{v_i + (a_i C + b_i)^2}.$$

Cette expression, indépendante de f , croît avec C . L'expression de EC est égale à 0 pour $C = 0$ et tend vers 1 quand C tend vers l'infini. Dans le cas de la responsabilité civile des sinistres matériels, compte tenu du spectre des coûts moyens observés, EC croît de 0,14 à 0,27.

Graphique n° 11
Garantie : Responsabilité civile des sinistres corporels
Représentation estimée de $\frac{\sigma}{P}$ en fonction de f et de C



Dans le cas de la responsabilité civile des sinistres corporels, les valeurs théoriques de $\frac{\sigma}{P}$ se situent entre 10 et 92. Les fréquences sont basses ; les coûts moyens sont à la fois relativement élevés et répartis sur un large spectre.

Les dérivées partielles de $\frac{\sigma}{P}$ par rapport à f et à C s'écrivent de la même façon que dans le cas de la responsabilité civile des sinistres matériels, avec ici $i = 1$, $v_1 = 1,016$, $a_1 = 4,13 \cdot 10^{-5}$, $b_1 = 0,1517$ (cf. tableaux n° 2 et 3).

L'évolution de $\frac{\partial(\sigma/P)}{\partial C}$ en fonction de f et de C est analogue à celle que l'on constate pour la responsabilité civile des sinistres matériels. Elle paraît nettement plus spectaculaire sur le graphique ci-dessus, du fait de l'étendue du spectre des coûts moyens, et de la faiblesse des fréquences, alors que le coefficient a_1 est pourtant moins élevé que dans le cas de la responsabilité civile des sinistres matériels.

La variation de l'élasticité EC est analogue à celle que l'on observe pour la responsabilité civile des sinistres matériels, aux coefficients v_i , a_i et b_i près. Compte tenu de la valeur de ces

coefficients et du spectre des coûts moyens observés, elle croît ici de 0,48 à 0,93 à mesure que C augmente.

Dans le cas de la garantie dommages accidents, les valeurs théoriques de $\frac{\sigma}{P}$ se situent entre 3,6 et 6. Les fréquences sont presque deux fois plus élevées qu'en responsabilité civile des sinistres matériels, alors que les coûts, bien que plus élevés, ne suivent pas cette proportion. Cela explique, en partie, des valeurs inférieures à celles que l'on obtient en responsabilité civile des sinistres matériels.

L'approximation qui consiste à retenir un coefficient a_2 égal à 0 permet d'écrire en fait, avec $i = 2$ selon la codification que nous avons adoptée pour les garanties, $v_2 = 1,050$, $b_2 = 1,439$:

$$\frac{\partial(\sigma/P)}{\partial f} = -\frac{\sqrt{v_i + b_i^2}}{2(\sqrt{f})^3}$$

qui est donc indépendant de C , la décroissance du ratio σ / P en fonction de f (qui correspond à la valeur absolue de la dérivée partielle exprimée ci-dessus) s'atténuant à mesure que f augmente.

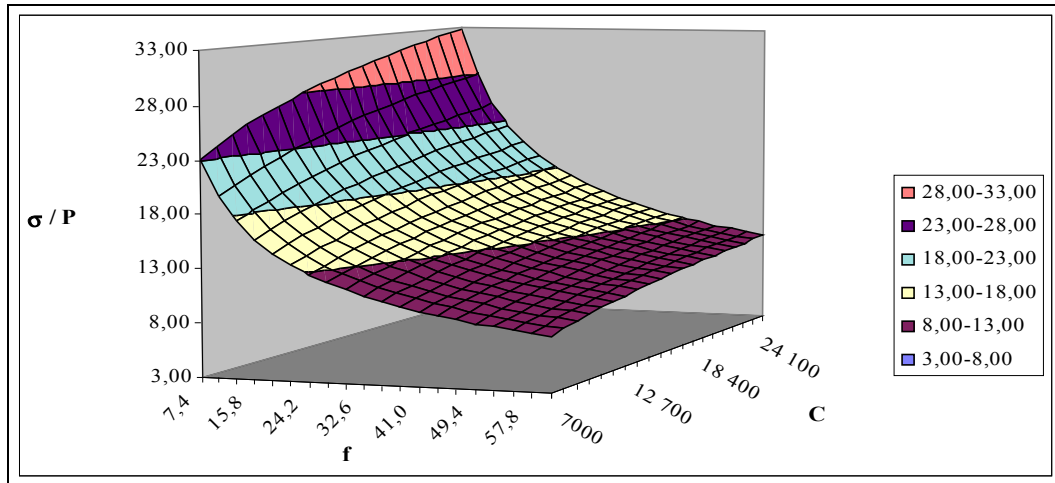
Par ailleurs, $\frac{\partial(\sigma/P)}{\partial C}$ est dans ce cas égal à 0.

Pour la garantie dommages accidents, l'élasticité E_C est nulle, dans la mesure où l'on considère que ΣC est indépendant de C .

Graphique n° 12

Garantie : Vol

Représentation estimée de $\frac{\sigma}{P}$ en fonction de f et de C



Pour la garantie vol, les valeurs de $\frac{\sigma}{P}$ se situent entre 8 et 33. Nous nous trouvons ici dans un contexte de fréquences plutôt basses et de coûts relativement élevés, avec un spectre étendu, de façon toutefois plus atténuée que dans le cas de la responsabilité civile des sinistres corporels. Les dérivées partielles de $\frac{\sigma}{P}$ s'écrivent comme suit, avec $j = 3$ selon la codification que nous avons adoptée pour les garanties, $v_3 = 1,057$, $a_3 = 0,7152$, $b_3 = -4,6916$:

$$\frac{\partial(\sigma/P)}{\partial f} = -\frac{\sqrt{v_j + (a_j \ln(C) + b_j)^2}}{2(\sqrt{f})^3}$$

$$\frac{\partial(\sigma/P)}{\partial C} = \frac{a_j(a_j \ln(C) + b_j)}{C\sqrt{f(v_j + (a_j \ln(C) + b_j)^2)}}$$

Dans le cas de la garantie vol, $\frac{\partial(\sigma/P)}{\partial C}$ décroît en fonction de f . On montre également que, en tout état de cause, pour des valeurs de C supérieures à 2 000 F, ce qui inclut largement le spectre des coûts moyens observés, $\frac{\partial(\sigma/P)}{\partial C}$ décroît également en fonction de C . Ce dernier point explique notamment la concavité des courbes iso-fréquence.

$$EC = \frac{C}{(\sigma/P)} \frac{\partial(\sigma/P)}{\partial C} = \frac{a_j(a_j \ln(C) + b_j)}{v_j + (a_j \ln(C) + b_j)^2}.$$

Ici, l'expression de l'élasticité de $\frac{\sigma}{P}$ par rapport à C est, compte tenu notamment de la négativité de b_3 , une fonction décroissante de C , jusqu'à une valeur de C égale à $C_{3-} = e\left(\frac{-\sqrt{v_3} - b_3}{a_3}\right)$, croissante jusqu'à une valeur de C égale à $C_{3+} = e\left(\frac{\sqrt{v_3} - b_3}{a_3}\right)$, et décroissante ensuite, en tendant vers 0 lorsque C tend vers l'infini.

Dans le cas de la garantie vol, nous avons $C_{3-} = 172$ F et $C_{3+} = 2\,973$ F. Les coûts moyens observés étant ici tous supérieurs à C_{3+} , l'élasticité EC apparaît dans ce contexte comme une fonction décroissante de C . Compte tenu du spectre des coûts moyens, elle décroît ainsi de 0,32 à 0,23.

La forme du graphique correspondant à la garantie bris de glaces est, en gros, analogue à celle du graphique de la garantie vol. Pour la garantie bris de glaces, les valeurs de $\frac{\sigma}{P}$ se situent

entre 2,5 et 6,2. Les fréquences sont d'un ordre de grandeur comparable à celles qui correspondent à la responsabilité civile des sinistres matériels, alors que les coûts moyens sont très faibles par rapport aux autres garanties.

Les commentaires concernant la variation des dérivées partielles de $\frac{\sigma}{P}$ pour la garantie bris de glaces sont analogues à ceux qui se rapportent à la garantie vol, aux coefficients v_j, a_j et b_j près. En particulier, on montre également que, en tout état de cause, pour des valeurs de C supérieures à 500 F, ce qui inclut largement le spectre des coûts moyens observés, $\frac{\partial(\sigma/P)}{\partial C}$ décroît également en fonction de C .

En ce qui concerne l'élasticité de $\frac{\sigma}{P}$ par rapport à C , compte tenu également de la négativité du coefficient b_4 , la fonction Ec présente des variations analogues à celle qui se rapporte à la garantie vol. De la même façon peuvent être identifiées des valeurs de C , C_{4-} et C_{4+} correspondant au changement de sens de la variation. Ici, $C_{4-} = 4$ F et $C_{4+} = 7\ 672$ F. Le spectre des coûts moyens observés au niveau de la garantie bris de glaces (1 200 F à 2 500 F) se situe dans la partie croissante de la fonction, qui augmente ainsi de 0,107 à 0,124 en fonction de C .

De façon générale, nous voyons que les effets comparés des variations de f et de C diffèrent selon les garanties. Ainsi, par exemple, en responsabilité civile des sinistres matériels, le spectre des fréquences est, en valeur relative, plus étendu que celui des coûts ; l'effet «fréquence» est alors prédominant. En revanche, en responsabilité civile des sinistres corporels, le spectre des coûts moyens est plus étendu que celui des fréquences ; c'est alors l'effet «coût» qui est prédominant, d'autant plus que les fréquences moyennes sont dans ce cas peu élevées – la valeur des intervalles Δf et ΔC (cf. tableau n° 6) étant, respectivement, beaucoup plus faible et beaucoup plus élevée que pour la responsabilité civile des sinistres matériels. Les élasticités respectives de $\frac{\sigma}{P}$ par rapport à f et à C jouent également un rôle sur les effets «fréquence» et «coût». En effet, toujours dans le cas de la responsabilité civile des sinistres corporels, le niveau élevé par rapport aux autres garanties ainsi que la croissance par rapport à C de la fonction Ec accentue encore l'augmentation de $\frac{\sigma}{P}$ par rapport à C , et ce à mesure que les coûts deviennent élevés.

Au demeurant, l'allure des surfaces et des nuages de points relatifs à chaque garantie dépend en particulier du rapport entre les étendues des spectres des fréquences et des coût moyens, ainsi que du niveau des valeurs prises par ces grandeurs.

Le tableau ci-dessous indique l'étendue des spectres des fréquences et des coûts moyens correspondant aux résultats obtenus pour les différentes sous-populations pour chaque garantie.

Tableau n° 6
Etendue des spectres observés par garantie entre les différentes sous-populations,
pour les fréquences et les coûts moyens

Garantie		f_{\min}	f_{\max}	Δf	$\frac{f_{\max}}{f_{\min}}$	C_{\min}	C_{\max}	ΔC	$\frac{C_{\max}}{C_{\min}}$
Responsabilité civile	Sinistres matériels	45,3	127,9	82,6	2,82	4 600	8 800	4 200	1,91
	Sinistres corporels non écrêtés	4,0	17,8	13,8	4,45	24 300	121 900	97 600	5,02
	Sinistres corporels écrêtés > 500 KF	4,0	17,8	13,8	4,45	26 100	101 400	75 300	3 89
Dommages accidents		90,0	242,0	152,0	2,69	4 000	11 400	7 400	2,85
Vol		7,0	57,4	50,4	8,20	6 800	29 900	23 100	4,40
Bris de glaces		40,0	193,9	153,9	4,85	1 200	2 500	1 300	2,08

VI. CONCLUSION

Comme nous l'avons vu au cours de cet article, l'application de relations entre les deux premiers moments associés au nombre annuel de sinistres par police et au coût par sinistre permet d'évaluer, en première approximation, les écarts types associés, ainsi que l'écart type de la charge annuelle des sinistres par police, pour une garantie donnée et une sous-population suffisamment homogène, à partir de la fréquence et du coût moyen observés. Ces relations peuvent donc présenter un intérêt opérationnel significatif. Nous tenons cependant à remarquer qu'en tout état de cause, l'ambition de cet article s'en tient à des notions d'ordre de grandeur estimatifs sans aller jusqu'à établir des règles de calcul à appliquer à la lettre, dans une logique de tarification par exemple.

L'étude réalisée dégage ce type de résultats expérimentaux sur une période de six ans, et à partir d'un échantillon de la population globale des assurés. Il serait, au demeurant, intéressant de confirmer de tels éléments par des études réalisées sur d'autres échantillons. De plus malgré la stabilité, sur une période couvrant six exercices, de nos résultats qui permet de suggérer des relations invariantes, la question de leur pérennité dans le temps reste posée : les relations mises en évidence et les coefficients correspondants sont-ils stables par rapport à une évolution sur le long terme de la sinistralité ?

Il semble donc utile, pour s'en assurer ou, dans le cas contraire, pour voir si d'autres relations apparaissent, de réactualiser ce type d'étude régulièrement, dans les conditions les plus simples possibles. Pour ce faire, il est nécessaire de pouvoir lever quelques difficultés telles que, en particulier, l'extraction et le traitement de données fiables et homogènes, à partir d'un échantillon suffisamment représentatif de la population globale des assurés.

Dans cette perspective, une solution pourrait consister à ce que la profession se dote d'outils permettant d'appréhender de façon éclairée - sinon de résoudre exactement - le problème de l'évaluation des composantes du risque (notamment les moments d'ordre 2 des variables aléatoires en jeu) d'un portefeuille pour une garantie déterminée, à partir de données facilement accessibles telles que les fréquences et les coûts moyens par garantie et par segment de clientèle, dans le cas de l'assurance automobile.

REFERENCES BIBLIOGRAPHIQUES

Pierre PETAUTON, 2000, *Théorie de l'assurance dommages*, Dunod, Paris.