



Mémoire présenté
devant l'Institut de Science Financière et d'Assurances
pour l'obtention du diplôme d'Actuaire de l'Université de Lyon

le 18 Septembre 2013

Par : TRIEU Thi Diep

Titre: Une analyse exploratoire et une tarification spécifique aux seniors sur le
portefeuille de Swiss Life dans le cadre du programme de recherche
« Conséquence du vieillissement de la population sur la sinistralité des
assurances »

Confidentialité : ☒ ☐ NON ☐ OUI (Durée : ☐ 1 an ☐ 2 ans)

Membres du jury de l'Institut des Actuaire

M. Olivier CAYOT

M. Gérald MONCHANIN

Entreprise :

Swiss Life et ESSEC Business School

Membres du jury I.S.F.A.

M. Alexis BIENVENÜE

M. Areski COUSIN

Mme Diana DOROBANTU

Mme Anne EYRAUD-LOISEL

M. Nicolas LEBOISNE

M. Stéphane LOISEL

Mlle Esterina MASIELLO

Mme Véronique MAUME-DESCHAMPS

M. Frédéric PLANCHET

Mme Béatrice REY-FOURNIER

M. Pierre RIBEREAU

M. Christian-Yann ROBERT

M. Didier RULLIERE

M. Pierre THEROND

Directeur de mémoire en entreprise :
Mme. Marie KRATZ

Invité :

Autorisation de mise en ligne sur
un site de diffusion de documents
actuariels (après expiration de
l'éventuel délai de confidentialité)

Signature du responsable entreprise

Signature du candidat

Secrétariat

Mme Marie-Claude MOUCHON

Bibliothèque :

Mme Patricia BARTOLO

Résumé

L'allongement de l'espérance de vie et l'arrivée de la génération du baby-boom à l'âge de la retraite provoquent une augmentation de l'effectif des populations âgées en France. Ce mémoire a pour objet l'étude des risques liés au vieillissement de la population et l'étude de ses conséquences sur la sinistralité des assurances. Nous allons nous intéresser à l'évolution de la population française avant d'entrevoir les différentes étapes à suivre pour l'analyse et la modélisation du risque liés au vieillissement dans l'assurance automobile. Les premières notions abordées ici sont la comparaison des résultats sur la population française versus la population Swiss Life et l'analyse descriptive des données de Swiss Life. Nous détaillerons les principes théoriques et les résultats sur des données de Swiss Life de différents types d'analyses factorielles telles que l'analyse en composantes principales (ACP) et l'analyse des correspondances multiples (ACM). Nous nous attacherons ensuite plus particulièrement à quelques études préliminaires telles que la fréquence, le coût moyen, la prime pure et le ratio S/C (sinistres/primes) sur les croisements de variables, notamment l'âge du premier conducteur (les seniors) et la région commerciale. Ensuite, nous avons utilisé une méthode de tarification de Swiss Life (basée sur S/C) qui nous permet de introduire des nouveaux coefficients tarifaires pour les personnes âgées chez Swiss Life. Nous avons également appliqué une méthode classique de tarification (basée sur le modèle linéaire généralisé) pour déterminer ces coefficients. Enfin, nous comparerons l'évolution des primes de ces deux différentes méthodes avant de choisir des nouveaux tarifs pour une segmentation plus fine par l'âge du conducteur.

Abstract

The extension of life expectancy and the arrival of the baby boomers in retirement cause an increase in the number of elderly people in France. This thesis aims to study the risks associated with an aging population and its impact on insurance claims. We will look at the evolution of the French population before getting a glimpse on different steps in the analysis and modeling of risk associated with aging in automobile insurance. First, we present the comparison results on the French population and Swiss Life population and some descriptive analysis about the Swiss Life's database. We also present the details of theoretical results on data provided by Swiss Life for different types of factorial analysis such as principal component analysis (ACP), multiple correspondence analysis (ACM). We then continue with some preliminary analyses such as the frequency, the average cost, the pure premium and the ratio S / C (claims / premiums) on the combination of variables, including age of first conductor (seniors) and commercial regions. Then, we used a method of pricing from Swiss Life (based on S / C) that allow us to introduce new rate coefficients for the elderly. We also applied a standard method of pricing (based on generalized linear model) to determine these coefficients. Finally, we compare the evolution of the premiums of these two different methods before selecting the new rates for finer segmentation by age of drivers.

Remerciements

Le présent rapport a été réalisé au sein de l'équipe de l'ESSEC Business School en collaboration avec Swiss Life. Que cette équipe au grand complet, trouve ici l'expression de ma reconnaissance pour la bonne ambiance ainsi que le professionnalisme dans laquelle s'est déroulé ce stage et pour la grande solidarité entre ses membres. Je tiens à remercier plus particulièrement la Professeure Marie KRATZ, responsable de notre projet, qui a su par sa patience et sa disponibilité, me guider tout au long de ce projet, et m'a apporté son aide précieuse dans l'élaboration de l'étude et la rédaction du présent mémoire. Par ailleurs, je souhaite aussi remercier Messieurs Olivier Stul à l'ESSEC et Meitner CADENA (stagiaire à l'ISFA) pour leurs conseils avisés : j'ai apprécié leur écoute et les aides apportées quant à l'élaboration du projet.

J'aimerais également remercier Messieurs Pierre François (Directeur Général de SwissLife), Alexandre Vercher (Directeur Technique & Produits IARD), Bruno Laporte (Direction Technique et Produits Dommages), Madame Valérie Louin (Direction Technique et Produits Dommages), Philippe Drapier (Directeur Technique, SwissLife Prévoyance et Santé), Madame Fatou Fall, Jean-Emmanuel Platel (Responsable Département Santé), l'ensemble des collaborateurs de Jean-Emmanuel Platel, Didier Liebart (Direction des Systèmes d'Information), Guy Saison (Direction des Systèmes d'Information), pour l'assistance et la disponibilité dont ils ont fait preuve.

Enfin, je tiens à remercier l'équipe des professeurs de l'Institut de Science Financière et d'Assurances et en particulier mon tuteur pédagogique Monsieur Didier RULLIERE.

Table des matières

Résumé	1
Abstract	2
Remerciements	3
Introduction	7
I Contribution des conducteurs âgés au bilan de l'insécurité routière en France	9
1 Le vieillissement de la population	9
1.1 Facteurs démographiques et besoins de déplacements	10
1.2 Vieillesse et handicap	11
1.3 Conclusion	11
2 Caractéristiques des déplacements et style de vie de différents groupes de personnes âgées	12
2.1 Accès à l'automobile et permis de conduire	12
2.2 Comportement du conducteur et cessation de la conduite	14
3 La sécurité des usagers de la route âgés	15
3.1 Les problèmes de sécurité des conducteurs âgés	15
3.2 Facteurs de risque	16
3.2.1 Les problèmes de conduite des conducteurs âgés	16
3.2.2 Etat de santé du conducteur/Pathologie	17
3.2.3 Evénements de vie "Stressants"	19
4 Prévention routière	19
4.1 Législation française	19
4.2 Législation dans les autres pays	20
4.3 Prévention de l'accident	21
II Présentation de l'étude et analyse des résultats sur l'ensemble du portefeuille de Swiss Life	24
5 Objectifs de l'étude et description des données	24
5.1 Problématique et enjeu de l'étude	24
5.2 Description des données	25

6	Etude préliminaire	26
6.1	L'évolution du nombre de contrats chez Swiss Life	26
6.2	Etude de la fréquence et du coût moyen des sinistres et analyse du ratio S / C	28
6.3	Analyse en Composantes Principales (ACP)	32
6.3.1	Méthodologie	32
6.3.2	Application aux données Swiss Life en 2011	37
6.4	Analyse des correspondances multiples (ACM)	42
6.4.1	Introduction	42
6.4.2	Méthodologie	42
6.4.3	Application aux données Swiss Life en 2011	46
III	La segmentation et la tarification	95
7	Etude préliminaire sur les variables croisées	98
7.1	Répartition du portefeuille, fréquence et coût moyen des sinistres, par âge et par région commerciale	104
7.2	La prime pure	128
7.3	L'analyse du ratio S/C pour chaque type de garantie	139
8	La tarification sur une segmentation en classes homogènes de l'âge chez les seniors et du zonage selon la méthode basée sur S/C	152
8.1	La méthode actuelle et sa limite	152
8.2	Grille tarifaire pour les garanties RC - Tous Accidents, Vol - Incen- die, Bris de Glace	162
8.3	Analyse des résultats	165
9	Présentation du modèle de tarification classique : modèle linéaire généralisé	172
9.1	Rappel théorique : présentation des modèles linéaires généralisés . .	172
9.2	Définition	177
9.3	Choix d'une fonction lien	180
9.4	Choix d'une loi de probabilité	191
9.5	Méthodologie	194
9.5.1	Choix d'un modèle	197
9.5.2	Estimer les coefficients de la régression pour la garantie RC- Tous Accidents	262
9.5.3	Validation du modèle	271
9.6	Comparaison des primes entre les deux méthodes de tarification (S/C et MLG)	291

Conclusion	310
Bibliography	314
Annexe 1 : Méthode des multiplicateurs de Lagrange	317
Annexe 2 : Détails ACP	324
Annexe 3 : Détails ACM de 2008 à 2010	331
Annexe 4 : Les tables de données de Swiss Life	335

Introduction

Swiss Life est un intervenant majeur sur les marchés de l'assurance vie et de la santé-prévoyance en France. A côté de ces secteurs, Swiss Life offre également à ses clients des produits d'assurance dommages. On trouve donc chez Swiss Life des clients à la fois présents dans le portefeuille automobile et dans le portefeuille santé. Suite à l'augmentation de la population âgée du fait de l'allongement de l'espérance de vie et l'arrivée de la génération du baby-boom à l'âge de la retraite, Swiss Life en collaboration avec l'ESSEC Business School (ESSEC) réalise un projet de recherche sur le thème : "Conséquences du vieillissement de la population sur la sinistralité des assurances et impacts sur la prévention automobile". L'objectif de ce projet de recherche est l'analyse des risques liés au vieillissement de la population et l'étude de ses conséquences sur la sinistralité des assurances, en contribuant en particulier à renforcer la prévention automobile. Les seniors sont des usagers de la route avec des comportements de mobilité bien spécifiques et cela a des conséquences sur les accidents. Swiss Life voudrait réaliser ce projet sur ses propres données pour savoir si les insuffisances et déficiences liées au vieillissement peuvent contribuer à accroître les accidents de la route. Les insuffisances et déficiences sont notamment la fatigue, les troubles du sommeil, la baisse d'acuité visuelle et auditive, la prise de médicaments,... Notre travail est de chercher à savoir ce qui a le plus d'incident sur les accidents parmi ces facteurs, et plus généralement à déterminer quels sont les facteurs de risque potentiels pour ce secteur, en tenant compte de l'évolution de la population. Pour parvenir à cette fin, la première étape consiste en l'analyse sur la population française et sur les clients du portefeuille automobile de Swiss Life. Ce mémoire présente les analyses suivantes :

- Dans la première partie, nous nous concentrons sur l'évolution de la population française ; ainsi nous mettons en évidence les spécificités des conducteurs âgés en termes de mobilité, d'accidents et de risque routier, à travers une approche quantitative. Pour cela il est fait état de différents travaux déjà menés sur cette question, en réactualisant certains résultats à partir de données récentes. De plus, on présente la législation française actuelle et quelques exemples de législations dans les autres pays, liées à la prévention automobile.

- Dans la deuxième partie, une motivation est d'étudier si les assurés chez Swiss Life ont une même tendance à évoluer que la population française. Cela permettrait d'utiliser certains résultats obtenus sur la population française ou de comparer les résultats sur la population française versus la population Swiss Life. Nous nous consacrerons tout d'abord à analyser l'évolution actuelle de la fréquence et du coût moyen des sinistres ainsi que du ratio S / C (charge de sinistres / primes). Nous considérons dans cette étude tous les contrats présents dans le portefeuille

automobile de Swiss Life en 2011. Nous comparons aussi ces facteurs lors des années précédentes pour avoir une analyse plus pertinente.

Ensuite, après quelques rappels théoriques sur certaines méthodes d'analyse statistique exploratoire, nous réalisons une analyse en composantes principales (ACP) sur les données de Swiss Life. Nous nous concentrons sur les clients présents à la fois dans le portefeuille automobile et dans le portefeuille santé pour l'analyse statistique exploratoire. Cette analyse nous permet de retenir les facteurs quantitatifs les plus importants. Ensuite, pour les facteurs qualitatifs nous effectuons une analyse des correspondances multiples (ACM). Les données en 2011 ne nous permettent pas de préciser quelques résultats obtenus sur la population française, nous réalisons donc l'ACM sur les données des années précédentes. Ces analyses conduisent à des résultats assez importants. Enfin, nous faisons les calculs du pourcentage du nombre de sinistres corporels et non-corporels, du pourcentage de montant de sinistres corporels et non-corporels (par rapport au nombre total et montant total, respectivement, de sinistres). Ces chiffres nous permettent de déduire comme principaux résultats : la proportion de sinistres corporels parmi les sinistres annuels est très faible, alors que leurs coûts sont très élevés. Sans que l'on puisse établir de lien entre la fréquence des sinistres corporels et l'âge, il est clair que le coût moyen des sinistres corporels augmente, lui, avec le vieillissement, et donc le coût total. Si les résultats de la population âgée de Swiss Life sont similaires à ceux sur la population française (moins d'accidents que chez les jeunes mais plus graves), ils sont par contre différents en terme de sinistres corporels (plus fréquents que chez les jeunes pour Swiss Life, contrairement à l'étude sur la population française). Nous réalisons ces analyses pendant 4 années successives, de 2008 à 2011. Les résultats restent presque identiques d'une année à l'autre.

- Dans la troisième partie, nous effectuons quelques études préliminaires sur les croisements de variables, notamment l'âge du premier conducteur et la région commerciale. Ensuite, nous avons utilisé une méthode de tarification de Swiss Life (basée sur S/C) qui nous permet de introduire des nouveaux coefficients tarifaires pour les personnes âgées chez Swiss Life. Nous avons également appliqué une méthode classique de tarification (basée sur le modèle linéaire généralisé) pour déterminer ces coefficients. Enfin, nous comparerons l'évolution des primes de ces deux différentes méthodes avant de choisir des nouveaux tarifs pour une segmentation plus fine par l'âge du conducteur.

Notons qu'un autre stagiaire, Meitner Cadena, a aussi participé à ce programme de recherche ; il a travaillé sur le secteur santé des conducteurs âgés et longévité. Nous avons mené notre étude en parallèle mais avons pu bien échanger sur nos sujets respectifs.

Première partie

Contribution des conducteurs âgés au bilan de l'insécurité routière en France

Nous explorons dans ce chapitre l'impact de l'âge des conducteurs de voiture dans la survenue d'un accident corporel et de ses conséquences. L'Observatoire National Interministériel de Sécurité Routière (ONISR) donne la définition suivante de l'accident corporel : "l'accident provoque au moins une victime, survient sur une voie ouverte à la circulation publique, et il implique au moins un véhicule. Les impliqués d'un accident corporel peuvent donc être indemnes, blessés ou tués".

Les premières questions que l'on se pose, dès lors qu'on souhaite établir un bilan des accidents selon l'âge des conducteurs, sont les suivantes : les conducteurs âgés ont-ils plus d'accidents que les conducteurs plus jeunes ? Leurs accidents sont-ils plus graves ? Les blessures des conducteurs âgés présentent-elles des spécificités dans leur gravité, leur localisation ou encore leurs conséquences à long terme ? Existe-t-il des particularités dans les accidents des conducteurs âgés qui suggèreraient un problème spécifiquement lié à l'âge ? Enfin, les conducteurs âgés sont-ils plus dangereux pour les autres usagers que les conducteurs plus jeunes ? Questions de plus en plus orientées vers une approche "santé" qui s'intéresse à l'individu et au fait qu'il existe une grande hétérogénéité au sein d'une même classe d'âge de conducteurs âgés.

Dans cette première partie, nous nous intéressons à la population française. Cela permet d'avoir une vision générale des évolutions de la population dans les 30-50 prochaines années.

1 Le vieillissement de la population

L'allongement de l'espérance de vie et l'arrivée de la génération du baby-boom à l'âge de la retraite provoquent une augmentation de l'effectif des populations âgées en France. Nous donnons ci-dessous (Figure 1) le pourcentage actuel et prévu de la population âgée de 65 ans et plus en France sur la période 2000-2050.

On voit sur cette figure 1 qu'il y a une tendance croissante et presque linéaire du nombre de personnes âgées en France. Cette tendance pourrait-elle entraîner une présence plus forte des personnes âgées sur la route ?

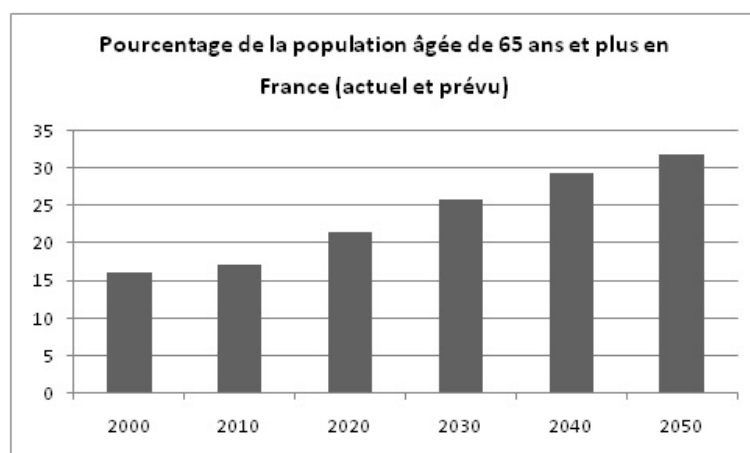


FIGURE 1 – Le pourcentage de la population âgée de 65 ans et plus en France (Source : OCDE, 2001)

D'après l'OCDE, on s'attend d'ici 2030 à un doublement de la population des 80 ans et plus et à son triplement d'ici 2050 dans la plupart des pays membres de l'OCDE.

1.1 Facteurs démographiques et besoins de déplacements

Les évolutions des modes de vie telles les améliorations des infrastructures et des moyens techniques, l'augmentation du recours à la voiture, la disparition de certains services de proximité (gare, postes, structures de soins...) notamment dans les campagnes, sont des facteurs déterminants des besoins de transport. Les statistiques montrent, d'année en année, que les distances moyennes parcourues vont augmenter. La proportion de kilomètres parcourus par les 65 ans et plus passe de 8,7 % en 1990 à 13,5 % en 2000 et à 14,9% en 2010 (ONISR). Est-ce une tendance qui va réellement se poursuivre ? Ou le rapport à la voiture va-t-il changer ? Notons que de nombreuses villes en France adoptent une nouvelle politique de transport, avec la construction de transports tels les tramways reliant villes et périphéries. De plus, les commerces de proximité réapparaissent tels Monop', Carrefour city, ... Il serait intéressant de comparer le rapport à la voiture (ou au transport public) de la génération 30-50 ans à celui touchant les seniors lorsqu'ils avaient cet âge là. Cela permettrait de voir s'il est raisonnable de s'attendre à une même tendance.

Le changement de la structure familiale affecte également les déplacements des personnes âgées. Actuellement, les personnes âgées ont en général deux enfants ou plus qui vivent à proximité. Il est fréquent que l'un d'entre eux soit disposé

à s'occuper d'un parent âgé. Ceci va changer, il ne sera pas rare que les futures seniors soient seuls car célibataires ou avec au plus un enfant. Autrement dit, les baby-boomers auront une probabilité plus faible d'avoir un membre de leur proche famille qui s'occupera d'eux dans leur grand âge.

1.2 Vieillesse et handicap

Depuis longtemps l'association entre vieillissement et l'apparition du handicap est reconnue. Parmi les facteurs de risque de handicap associés au vieillissement ayant un impact sur la performance de conduite et/ou risque d'accident, nous pouvons citer les pathologies suivantes : l'état de la vue (cataractes, glaucomes, acuité visuelle, etc.), la perte de mémoire, la diminution des capacités et des processus cognitifs, la démence (notamment la maladie d'Alzheimer), les maladies mentales, les troubles musculo-squelettiques graves (notamment l'arthrite, le mal de dos), etc.

Une proportion non négligeable des personnes âgées est atteinte de handicap et maladies. Selon l'OCDE, en 2008 on estime que environ 12% des adultes souffre un handicap en ce moment et que parmi ces personnes, les deux tiers sont des personnes âgées de 65 ans et plus. En revanche, la santé des personnes âgées devrait être meilleure dans la future par rapport à celle à notre époque. Les seniors pourraient donc retarder leur départ à la retraite et poursuivre une activité salariée, devenir plus actifs pendant leur temps de loisirs et être davantage capables de conduire sans problème.

Toutes ces études démographiques nous incitent à prendre comme âge de référence pour la population âgée, non pas un âge fixé, mais l'âge de retraite en vue d'une étude (modélisation) dynamique, l'âge de la retraite étant amené à augmenter progressivement ces prochaines années.

1.3 Conclusion

D'ici 2030, lorsque les derniers baby-boomers seront parvenus à l'âge de la retraite, la proportion des plus de 65 ans aura presque doublée. Le taux de croissance de la proportion des personnes âgées de 80 ans ou plus sera davantage marqué.

En raison de la diminution de la taille des familles et de la mondialisation du travail, les personnes âgées auront moins sinon pas de membres de la famille proche disponibles pour s'occuper d'elles. Ceci devrait constituer une pression sur la collectivité pour fournir des services et des équipements en remplacement.

Les futures cohortes de personnes âgées seront vraisemblablement en meilleure santé et plus actives que celles d’aujourd’hui. Toutefois, en raison de l’augmentation du nombre de personnes âgées de 80 ans ou plus, plus de personnes ayant dépassé les 80 ans pourraient souffrir de handicaps fonctionnels. Cette évolution a des implications directes sur la mobilité et la sécurité.

2 Caractéristiques des déplacements et style de vie de différents groupes de personnes âgées

Les transports collectifs, la marche à pied et la bicyclette sont des modes de déplacement courant dans presque tous les pays européens. Selon une enquête nationale sur les transports, réalisée par l’Institut National de Recherche sur les Transports et les Sécurité (INRETS), en France, les comportements de déplacement varient fortement selon l’âge et l’activité des individus. L’analyse des budgets-temps dans une requête réalisée par l’INRETS, montre que les hommes actifs passent, par jour, plus d’une heure (70 min) en déplacements tous modes confondus, alors que les retraités consacrent un peu plus d’une demi-heure (36 min) à leurs déplacements quotidiens. Chez les hommes de 65 à 74 ans, sept trajets sur dix se font en voiture en tant que conducteur ou passager et chez les femmes du même âge, la voiture est utilisée dans un peu moins de la moitié des déplacements. Au-delà de 75 ans, cette part des trajets en voiture tombe à 36 % chez les hommes et 29 % chez les femmes, les déplacements à pied étant alors plus importants.

2.1 Accès à l’automobile et permis de conduire

L’accès à l’automobile et la possession du permis de conduire constituent deux déterminants majeurs du choix de mode de déplacement. Si actuellement, la plupart des personnes d’âge moyen ont le permis de conduire, il n’en pas de même pour les personnes âgées aujourd’hui. Actuellement une proportion substantielle de personnes âgées (les femmes en particulier) n’a pas ou n’a jamais eu de permis de conduire. La Figure 2 donne une comparaison des taux actuels de possession du permis de conduire selon les tranches d’âge, ce taux est calculé par la division de nombre des titulaires du permis et la sous population selon les tranches d’âge.

Selon la Figure 2, le taux de possession du permis de conduire est maximal chez les personnes d’âge de 25 à 34 ans. Il y a une tendance décroissante du pourcentage de possession du permis de conduire à partir de l’âge 25 ans. Le taux chez les personnes âgées est relativement assez modéré (78,3%) par rapport à celui des personnes d’âge moyen. Avoir un permis de conduire n’est pas nécessairement sy-

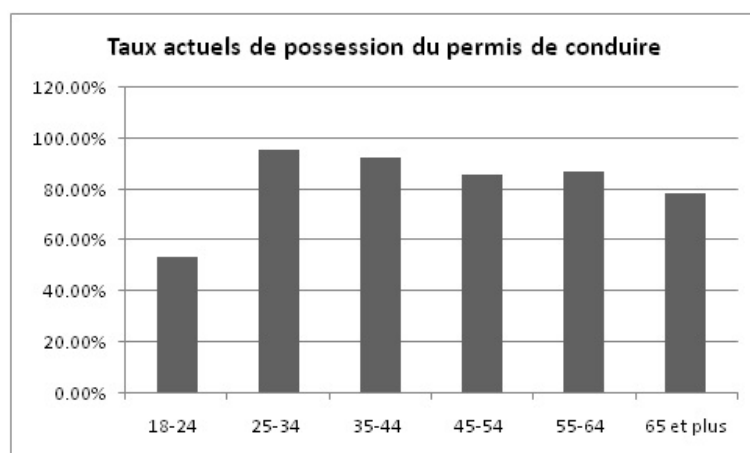


FIGURE 2 – Taux actuels de possession du permis de conduire
(Source : enquête SOFRES 2008)

nonyme d'avoir accès à une voiture. Donnons ci-dessous (Figure 3) une répartition des titulaires du permis de conduire par âge selon la fréquence de conduite.

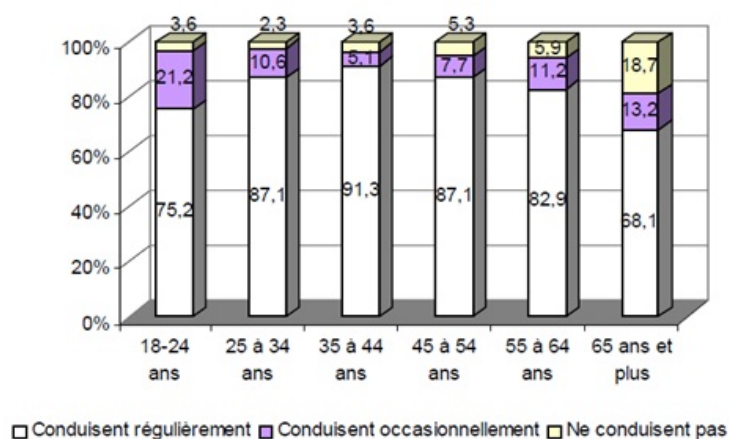


FIGURE 3 – La répartition des titulaires du permis de conduire par âge
(Source : enquête SOFRES 2008)

On remarque que la part des personnes âgées n'ayant pas accès à une voiture est largement supérieure aux autres personnes. C'est-à-dire, une proportion très grande de cette population a cessé de conduire. L'écart entre les conducteurs et les non-conducteurs est particulièrement marqué pour les femmes. Actuellement, les femmes conduisent moins que les hommes.

En confrontant les niveaux actuels de possession du permis et les projections démographiques présentées dans la partie précédente, il est possible d'estimer le nombre de personnes âgées qui auront leur permis en 2030. Les projections pour lesquelles on a obtenu des statistiques sur les permis, sont présentées dans la Table 1 suivant :

Pourcentage de conducteurs avec permis âgés de 65+ en 2008	Pourcentage de conducteurs avec permis âgés de 65+ en 2030	Pourcentage d'augmentation des conducteurs avec permis ages de 65+ entre 2008 et 2030
19,8%	25,8%	30,3 %

TABLE 1 – Projection du taux de possession du permis de conduire pour les personnes âgées en 2030. (Source : enquête SOFRES et OCDE)

2.2 Comportement du conducteur et cessation de la conduite

Les conducteurs âgés modifient leurs habitudes de conduite selon leurs capacités. Ils évitent de conduire dans les conditions difficiles par exemple conduisant moins la nuit, aux heures de pointe ou de fort trafic, ou encore dans des zones de circulation trop complexes ; ils roulent moins vite que les autres. Cela peut signifier un comportement plus prudent chez les conducteurs âgés. Donnons la Figure 4 l'évolution du nombre annuel de kilomètres parcourus depuis 1999.

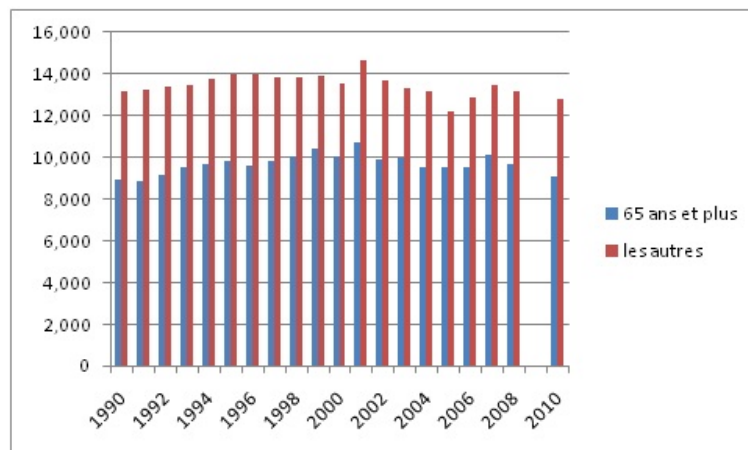


FIGURE 4 – Le nombre annuel de kilomètres parcourus (Source : ONISR)

Les personnes âgées conduisent moins que l'ensemble des conducteurs bien que depuis quelques années, le kilométrage effectué par cette population tend à s'accroître. Le nombre annuel de kilomètres parcourus par les personnes âgées de 65 ans et plus qui évoluait régulièrement aux environs de 10000 kilomètres par an entre 1999 et 2003 est stabilisé depuis autour de 9500 kilomètres. D'une année à l'autre, la différence de parcours entre les personnes âgées et l'ensemble de la population reste globalement stable, de l'ordre de 1,3 à 1,4 fois moins.

3 La sécurité des usagers de la route âgés

La sécurité est peut-être la question qui a le plus l'influence sur la formulation des pratiques et des politiques qui touchent la mobilité et le transport des personnes âgées. On peut affirmer que le problème de sécurité le plus important pour les usagers de la route âgée réside dans leur plus grande fragilité, ce qui les rend plus susceptibles d'être gravement blessés. Bien que les personnes âgées aient moins d'accidents que les autres tranches d'âge, elles sont sur-représentées parmi les tués et les blessés graves.

3.1 Les problèmes de sécurité des conducteurs âgés

La gravité des accidents de la route augmente avec l'âge des usagers. Les résultats concernant les personnes âgées recensées par le Registre dans le document de travail réalisé par l'ONISR sont les suivants (Table 2) sur la période 2006-2008 :

	Tués	Blessés avec séquelles lourdes
25-64 ans	11,3 %	11,3 %
65 ans et plus	49,3 %	28,1 %

TABLE 2 – La gravité des accidents de la route sur la période 2006-2008 (étude ONISR)

Sur ce tableau, nous pouvons remarquer qu'il y a 4 fois plus de tués chez les seniors que chez les jeunes, et qu'il y a 2,5 fois plus de blessés avec séquelles lourdes chez les seniors que chez les jeunes.

Présentons des résultats plus détaillés :

	Décédés	Survivants avec séquelles lourdes	Survivants avec séquelles mineures ou modérées	Victimes sans séquelles	Total toutes gravités
	Effectif et %	Effectif et %	Effectif et %	Effectif et %	Effectif et %
65 à 69 ans	9 2,6 %	6 1,8 %	122 35,9 %	203 59,7 %	340 100%
70 à 74 ans	12 3,7 %	5 1,5 %	121 37,0 %	189 57,8 %	327 100%
75 à 79 ans	8 3,0 %	6 2,2 %	82 30,4 %	174 64,4 %	270 100%
80 à 84 ans	10 5,2 %	4 2,1 %	54 28,3 %	123 64,4 %	191 100%
85 ans et plus	5 4,1 %	3 2,5 %	38 31,1 %	76 62,3 %	122 100%
Total	44 3,5 %	24 1,9 %	417 33,4 %	765 61,2 %	1250 100 %

TABLE 3 – La gravité des accidents de la route chez les personnes âgées (2006-2008)

Selon la Table 3, le nombre de décédés et de survivants avec séquelles lourdes est le plus élevé chez les personnes à 80 ans et plus. Ces données montrent que les personnes de 80 ans et plus sont davantage exposées que les autres.

3.2 Facteurs de risque

3.2.1 Les problèmes de conduite des conducteurs âgés

Dans la vie quotidienne, les conducteurs âgés rencontrent un certain nombre de difficultés, notamment les trois principaux types de problèmes de conduite suivants :

- **Les erreurs** (échec d'une action prévue à atteindre son objectif initial) :
Elles recouvrent des erreurs dans l'observation et dans le jugement, elles ont des conséquences potentiellement dangereuses. On ne note pas de différences de sexe. On liste quelques actions qui présentent les erreurs suivantes : sous-estimer la vitesse d'un véhicule en face en dépassant ; freiner trop fortement sur une route glissante, ou braquer dans le mauvais sens lors d'un dérapage ; attendre en file pour tourner à gauche sur voie prioritaire, trop préoccupé par le trafic prioritaire, percuter le véhicule de devant ; échec à vérifier le rétro intérieur lors d'un changement de file ; en tournant dans une rue, échec à percevoir un piéton traversant ; échec à percevoir la signalisation de priorité, et en conséquence frôler l'accident avec le véhicule prioritaire ; en tournant à gauche, manquer de renverser un cycliste ; essayer de doubler quelqu'un et ne pas voir qu'il avait mis son clignotant pour tourner à droite.

Lorsqu'une erreur est commise, les conducteurs âgés ont sans doute plus de mal à la rattraper et à empêcher l'évolution vers l'accident ; l'une des raisons à ce phénomène est qu'ils sont moins alertes que plus jeunes et qu'ils ont passé le permis il y a très longtemps, à une époque où l'environnement routier était très différent, moins congestionné et qu'ils n'ont pas eu de « rafraîchissement » depuis.

- **Les oublis** (faiblesses de l'attention) :

Ils sont plus souvent rapportés par les femmes et les personnes âgées. On liste quelques actions qui présentent les oublis suivants : mal lire la signalisation et prendre une mauvaise route à un carrefour ; se mettre dans une mauvaise voie en approchant d'une intersection ; démarrer au feu avec une vitesse trop élevée.

- **Les infractions** (conduites risquées volontaires, délibérées) :

On donne exemple des infractions suivantes : vitesse ; infraction-alcool.

D'une manière générale, les infractions sont liés plus souvent aux accidents chez les jeunes que chez les conducteurs âgés, du fait par exemple que l'alcool soit moins présent dans les accidents des conducteurs âgés. En plus, les conducteurs âgés roulent moins vite que les autres.

3.2.2 Etat de santé du conducteur/Pathologie

La prévalence de nombreuses pathologies augmente avec l'âge. La question de l'état de santé des conducteurs risque de se poser de plus en plus avec le vieillissement de la population.

Les conducteurs âgés sont souvent des conducteurs expérimentés, mais c'est dans cette classe d'âge que les pathologies médicales vont le plus majorer le risque d'accident. Plusieurs études suggèrent que les personnes âgées sont exposées à un risque accru d'accidents de la route. Cependant, l'hétérogénéité de la catégorie "personnes âgées" empêche toute généralisation, et il faut préciser quelles sont les composantes du vieillissement qui peuvent interférer avec l'activité de conduite. En passant en revue la littérature existante sur le sujet, nous présentons ci-dessous les principales affections pour lesquelles l'association avec les accidents de la route a été étudiée.

- **La maladie d'Alzheimer** : Les personnes présentant une démence de type Alzheimer à un stade débutant, avaient des difficultés à engager et désengager leur attention, par rapport à des personnes non démentes. Ils pourraient expliquer le sur risque d'accident observé chez les personnes qui présentent une démence de type Alzheimer [4] [23] [15]. Un constat important par rapport à la maladie d'Alzheimer est son lien avec d'autres maladies chroniques,

en effet l'hypertension artérielle et l'athérosclérose semblent être associées à un risque accru de maladie d'Alzheimer et le traitement de l'hypertension artérielle pourrait entraîner une réduction importante de l'incidence de la maladie. Le diabète majorerait également le risque de survenue de maladie d'Alzheimer [26].

- **La flexibilité, la force musculaire, la dextérité manuelle** : ces fonctions déclinent avec l'âge et ces changements peuvent influencer la capacité d'entrer et de sortir d'une voiture, conduire le véhicule, et peut influer sur les blessures et la récupération [7].
- **Cardiologie** : plusieurs études retrouvent une augmentation du risque d'accident chez les personnes atteintes de pathologies cardio-vasculaires, et plus précisément de pathologies coronaires ou de troubles du rythme [23]. En revanche, d'autres ne retrouvent pas de risque supérieur à celui de la population générale [20].
- **Œil et vision** : les principales pathologies responsables de malvoyance chez l'adulte, en Europe, sont la dégénérescence maculaire liée à l'âge, la cataracte sénile, le glaucome et les rétinopathies. La cataracte et le glaucome ont été retrouvés comme facteurs associés aux accidents de la circulation, de même qu'une diminution de l'acuité visuelle ou une altération du champ visuel [23] [4].
- **Audition** : l'audition est également sollicitée lors de la conduite. Il existe pourtant peu d'études dans ce domaine. Une étude réalisée en Australie, où l'on conduit à gauche, retrouve un risque d'accident associé à une perte auditive de l'oreille droite uniquement, et suggère le rôle de l'audition dans la détection d'événements survenant du côté droit de la route [23] [4].
- **L'épilepsie** et la conduite ont fait l'objet de nombreuses études et publications. Deux causes principales peuvent être responsables d'accident chez un épileptique : les pertes de connaissance liées aux crises et les troubles psychologiques ou troubles du comportement inter-critique. Certaines études retrouvent des taux d'accidents 1,3 à 2 fois plus élevés que dans une population témoin [21], alors que d'autres ne montrent pas de risque supérieur à celui de la population générale [22].
- **La fatigue et des troubles du sommeil** : la somnolence au volant est un facteur fréquent de survenue d'accidents et la première cause d'accidents

mortels sur autoroute en France notamment chez les seniors [2] [23].

- **La pathologie d’arthrite** : une étude parmi des conducteurs de plus de 65 ans retrouve un risque d’accident près de deux fois plus élevé chez les femmes atteintes d’arthrite [23].
- **Le diabète** est une pathologie fréquente dans la population. Il y a quelques études retrouvant une augmentation du risque d’accidents chez les diabétiques [23].

3.2.3 Événements de vie "Stressants"

Certains événements, tels que des changements d’ordre professionnels, financiers, relationnels ou familiaux, peuvent être considérés comme "stressants", dans la mesure où ils entraînent des modifications importantes du mode de vie, et nécessitent un temps d’adaptation. Quelques études ont montré que ces événements pouvaient être associés à la survenue de pathologies organiques [23]. D’autres suggèrent que ces événements sont également susceptibles d’induire une prédisposition temporaire aux accidents de la circulation [23]. En 1981, une étude a rapporté un nombre total d’événements de vie plus élevé dans les 12 mois précédant un accident pour des conducteurs déclarés en faute, par rapport aux conducteurs accidentés non en faute [23]. Auparavant, une étude américaine avait montré que le taux d’accident était multiplié par deux au cours des 6 mois précédant un divorce, ainsi qu’au cours des 6 mois suivants [23]. Il s’agissait cependant d’une étude très descriptive, sans tests statistiques des hypothèses avancées.

4 Prévention routière

4.1 Législation française

Depuis les débuts de l’automobile, la bonne santé a toujours été considérée comme une condition de la capacité à conduire, voire, bonne santé et capacité à conduire ont pu être considérées comme identiques. Ainsi, dans de nombreux pays, une visite médicale est requise pour la passation du permis. De même, le dépistage médical est apparu « naturellement » comme la mesure de sécurité routière pertinente pour gérer le problème des conducteurs âgés.

En France, les titulaires du permis de conduire pour les véhicules autres que automobiles (catégories C, D et E et conducteurs professionnels) sont soumis à un examen médical périodique auprès des commissions départementales de la sécurité

routière. Cet examen médical est pratiqué avec la délivrance initiale du permis, puis tous les 5 ans pour les conducteurs de moins de 60 ans, tous les 2 ans pour les conducteurs dont l'âge est compris entre 60 et 76 ans, tous les ans pour les conducteurs ayant dépassé l'âge de 76 ans. Pour les usagers d'automobiles, seules quelques situations peuvent donner lieu à une visite médicale : la déclaration spontanée d'une affection dans le questionnaire médical lors de l'inscription au permis de conduire, la décision de l'inspecteur lors de l'examen, et, en aval, une suspension de permis de plus d'un mois ou une annulation de permis. L'arrêté ministériel du 7 mai 1997 fixe une liste de pathologies reconnues comme des conditions incompatibles avec l'obtention ou le renouvellement du permis de conduire. Cette liste comprend 6 grandes catégories, classées par spécialités : "cardiologie", "oeil et vision", "oto-rhino-laryngologie et pneumologie", "neurologie et psychiatrie", "appareil locomoteur" et "divers". L'incompatibilité (temporaire ou permanente) avec la conduite dépend à la fois de la catégorie du permis de conduire et de la gravité de l'affection considérée. Pour de nombreuses affections, l'avis d'un spécialiste est recommandé. Les médecins manquent cependant de critères objectifs pour évaluer l'aptitude médicale à la conduite. Cette liste a été établie essentiellement sur des considérations cliniques (risques de malaises, de pertes de connaissance, limitations des fonctions motrices ou sensorielles) et l'excès de risque provoqué par telle ou telle pathologie n'a pas été établi scientifiquement.

Dans divers pays européens, le renouvellement de la validité du permis après 65 ou 70 ans est soumis à une visite médicale obligatoire. On va présenter dans la section suivante les systèmes de quelques pays européens pour comparaison avec la France. Si ces systèmes sont prometteurs, cela inciterait à les voir proposés en France ; ils pourraient à défaut servir de base à de nouveaux modèles.

4.2 Législation dans les autres pays

Nous donnons ci-dessous les systèmes de législation de quelques pays.

▷ L'exemple finlandais

En Finlande, un dépistage médical très strict des conducteurs âgés a été introduit aux débuts des années 70 ; à ce jour, aucune analyse coûts-bénéfices de cette mesure n'a encore été effectuée.

Selon la loi finlandaise, le droit de conduire une automobile est valide jusqu'à 70 ans ; toutefois, à partir de 45 ans, ce droit devient conditionnel, et le conducteur doit passer un examen médical et visuel tous les 5 ans. A l'âge de 70 ans, le permis n'est plus valide et les personnes qui désirent continuer à conduire doivent passer une visite médicale et solliciter un nouveau permis. De plus, lorsqu'ils le sollicitent, ils doivent présenter un document signé par « deux personnes de foi » stipulant

qu'ils ont bien conservé leurs savoir-faire en continuant de conduire régulièrement. Au delà de 70 ans, le permis n'est accordé que pour une période de 5 ans, voire pour une période plus courte selon l'avis du médecin. Ces périodes de validité diminuent nettement après 80 ans. Une fois que le conducteur âgé a satisfait à ces obligations, il doit encore aller au commissariat pour déposer ou reprendre son permis, puis payer les frais et taxes afférents.

On l'aura compris , le système finlandais est un modèle contraignant.

▷ **L'exemple suédois**

La Suède a un système très libéral et permissif vis à vis des conducteurs âgés. Il y a une récente comparaison entre la Finlande et la Suède qui a montré que les taux d'accidents de voiture des personnes âgées par tête d'habitant étaient semblables dans les deux pays, mais les usagers vulnérables (piétons, cyclistes, cyclomotoristes) âgés finlandais présentaient un taux d'accident deux fois supérieur à celui de leurs homologues suédois.

L'efficacité du dépistage a donc été mis en doute.

▷ **L'exemple australien**

A la pointe des innovations politiques, l'Australie est également un pays vers lequel se tournent les regards des experts en sécurité routière. Aucun effet positif du dépistage n'a pu être observé dans ce pays : ainsi, l'état de Victoria, qui est le seul à ne pas avoir de dépistage médical des conducteurs âgés , présente les meilleurs résultats de sécurité routière pour le groupe concerné (Hull, 1991).

En conclusion, le dépistage médical traditionnel semble de plus en plus considéré comme étant « archaïque » ; toutefois, nous ne disposons pas encore d'une alternative pour un dépistage qui serait plus efficace.

Le diagnostic des conducteurs âgés semble donc être un domaine de décisions sociétales où le besoin urgent de solutions pratiques devance bien trop tôt les preuves scientifiques sur lesquelles on pourrait fonder une décision.

4.3 Prévention de l'accident

Le manque d'efficacité du dépistage conduit les chercheurs à proposer les quelques méthodes suivantes pour la prévention routière.

1) *L'enquête de routine par le médecin de famille :*

Même lorsqu'une visite médicale n'est pas obligatoire, un conducteur âgé consulte en fait régulièrement son médecin ; de telles consultations devraient en fait fournir au généraliste l'occasion de tester la situation de leur patient quant à la conduite automobile. Donc, de manière routinière, le médecin de famille (ou d'autres per-

sonnels de santé en relation avec la personne âgée) devrait systématiquement s'enquérir si cette dernière conduit encore.

Si un déficit est détecté, le généraliste devrait déterminer quelle fonction cognitive est affectée (savoir-faire perceptifs, attention ou jugement) ; il devrait aussi envisager s'il existe un effet cumulatif des diverses altérations¹.

Par ailleurs, les effets secondaires des traitements médicamenteux doivent être sérieusement considérés à cet âge de la vie.

2) L'utilisation d'informations sur le fonctionnement du patient dans les activités de la vie quotidienne

On prend en compte les aspects de la vie quotidienne suivants : les difficultés de l'alimentation, de l'habillement ; les témoignages de la famille sur le comportement de conduite automobile (les expériences de passagers du sujet, les perturbations repérées, les tendances à la prise de risque, les infractions et les accidents).

Les performances du patient lors des activités simples de la vie quotidienne pourraient en fait constituer un meilleur indicateur de la conduite automobile que les résultats des tests cognitifs.

Toutefois, la tâche de conduite automobile est une tâche complexe qui requiert parfois intégration et adaptation ; en miroir dans la vie quotidienne du patient, l'on peut évaluer les performances lors des activités complexes de la dite vie quotidienne (les courses, les déplacements, la gestion de son budget, etc.) ; ces dernières sont des tâches plus complexes, reflétant une organisation hiérarchique des comportements, des habitudes, des planifications.

Les performances lors des activités complexes de la vie quotidienne pourraient donc être un meilleur indicateur de la performance de conduite que les performances lors des activités simples de la vie quotidienne ou les performances cognitives ; en effet, les accidents « typiques » des conducteurs déments semblent moins souvent être dûs à des déficits dans les manœuvres basiques du véhicule, et plus souvent être dus à des déficits de savoir-faire plus élaborés, ou du moins à la coordination entre savoir-faire « inférieurs » et « supérieurs ».

Pour synthétiser, on peut faire l'hypothèse d'une corrélation entre la compétence de conduite et l'indépendance lors des activités complexes de la vie quotidienne.

3) L'épreuve pratique de conduite

L'épreuve pratique du permis de conduire n'est sans doute pas assez « pointue » pour révéler des déficits, car la « tâche standard » de conduite est très automatisée ; pour être adaptée aux conducteurs déments, cette épreuve devrait évaluer

1. en effet, un seul déficit en soi n'est pas nécessairement alarmant, mais une combinaison peut l'être beaucoup plus ...

plus spécifiquement des fonctions comme l'attention, la capacité à suivre une séquence de directions, ou la capacité à gérer un niveau minimal de stress. Dans cette perspective, les simulateurs pourraient se révéler utiles. Finalement, nous ne savons pas très bien encore si l'on doit tester un niveau minimum de compétences ou les limites de la performance du sujet.

Deuxième partie

Présentation de l'étude et analyse des résultats sur l'ensemble du portefeuille de Swiss Life

5 Objectifs de l'étude et description des données

5.1 Problématique et enjeu de l'étude

Le marché de l'assurance automobile en France est un marché quasiment saturé à cause de la faible croissance du parc automobile. Dans ce contexte, les assureurs qui désirent tous accroître leur part de marché se mènent une guerre commerciale très intense. Pour pouvoir rester compétitif, le tarif doit nécessairement être basé sur une segmentation² performante. Une segmentation fine permet d'avoir une meilleure connaissance du risque afin de se prémunir au maximum contre le risque d'anti-sélection³. Il est donc nécessaire de segmenter au mieux les assurés en définissant des critères de risque qui permettent de scinder la population initiale en populations de risques homogènes. L'étude des personnes âgées participe à la conception de cette segmentation.

La prime d'assurance d'un contrat automobile dépend d'un certain nombre de caractéristiques : caractéristiques du véhicule (puissance, vitesse, marque...), caractéristiques du conducteur (âge, sexe, CSP...), caractéristiques du contrat (garanties souscrites,...). En fonction des montants des sinistres et de leurs fréquences observés sur le portefeuille, il est possible de déterminer l'influence de chacune des modalités prises par ces caractéristiques au travers d'un poids affecté à la prime. Ainsi, les modalités prises par un contrat donné permettent de calculer la prime grâce aux poids calculés précédemment.

On va s'intéresser plus particulièrement aux deux caractéristiques de l'âge et de la région commerciale. En effet, en fonction de son âge et de son lieu habitation, l'assuré peut présenter un profil plus ou moins risqué (intuitivement lié à son état de santé, son comportement de conduite,...). A partir de ce constat, l'enjeu est de

2. Une segmentation est le découpage en sous-ensembles homogènes, significatifs et accessibles

3. L'anti-sélection est une expression désigne les dysfonctionnements des marchés d'assurance qui résultent de l'information cachée dont les assurés peuvent disposer sur leurs propres risques et qui n'est pas accessible aux assureurs.

mieux étudier la sinistralité des conducteurs âgés pour chaque région commerciale, fonction du profil de l'assuré.

5.2 Description des données

- Notre population de référence est en priorité la population des retraités chez Swiss Life puisque notre objectif est d'étudier les personnes âgées. Mais dans un premier temps, nous observons l'ensemble du portefeuille d'assurés.
- Nous considérons dans un premier temps tous les contrats présents dans le portefeuille automobile de Swiss Life sur les années 2008, 2009, 2010 et 2011. Cette période permet à la fois d'avoir un nombre important de données pour pouvoir tarifier (le portefeuille automobile Swiss Life comporte sur 4 ans l'équivalent de plus de 200 000 d'assurés) et d'avoir un recul suffisant afin de les analyser. Cette période permet aussi d'avoir une comparaison approfondie des évolutions de la population française.
- Le périmètre correspond à Swiss Life, qui est divisé en 7 régions commerciales :
 1. Ile de France (Paris, Seine et Marne, Yvelines, Essonne, Haut de Seine, Seine Saint Denis, Val de Marne, Val d'Oise).
 2. Région Nord (Calvados, Eure, Manche, Nord, Oise, Orne, Pas de Calais, Seine Maritime, Somme).
 3. Région Ouest (Cher, Côtes d'Armor, Eure et Loir, Finistère, Ile et Vilaine, Indre, Indre et Loire, Loir et Cher, Loire Atlantique, Loiret, Maine et Loire, Mayenne, Morbihan, Sarthe, Deux Sèvres, Vendée, Vienne).
 4. Région Sud-ouest (Ariège, Aude, Aveyron, Charente, Charente Maritime, Corrèze, Creuse, Dordogne, Haute Garonne, Gers, Gironde, Landes, Lot, Lot et Garonne, Pyrénées Atlantiques, Hautes Pyrénées, Pyrénées Orientales, Tarn, Tarn et Garonne, Haute Vienne).
 5. Région Sud (Bouches du Rhône, Corse du Sud, Haute Corse, Gard, Hérault, Lozère, Var, Vaucluse).
 6. Région Est (Ardennes, Aube, Côte d'Or, Doubs, Jura, Marne, Haute Marne, Meurthe et Moselle, Meuse, Moselle, Nièvre, Bas Rhin, Haut Rhin, Haute Saône, Saône et Loire, Vosges, Yonne, Territoire de Belford).
 7. Région Rhône-Alpes (Cantal, Drôme, Isère, Loire, Haute Loire, Puy de Dôme, Rhône, Savoie, Haute Savoie).

- Nous nous concentrons également sur les clients présents à la fois dans le portefeuille automobile et dans le portefeuille santé de Swiss Life pendant la période 2009, 2010, 2011 (il y a environ 30000 de ces clients, soit 10% des assurés automobiles chez Swiss Life). Cette représentativité de l'échantillon peut entraîner des biais au sein études. (En ce moment, les données "santé" ne sont pas disponibles et nous nous concentrons donc juste sur les données automobiles).
- Nous ne considérons que les voitures légères et ne nous intéressons pas aux camping-cars et véhicules de collection pour lesquels une tarification particulière a été créée.

6 Etude préliminaire

6.1 L'évolution du nombre de contrats chez Swiss Life

Donnons ci-dessous (Figure 5) l'évolution du nombre de contrats chez les personnes âgées selon chaque tranche d'âge pendant 3 années successives 2009, 2010, 2011.

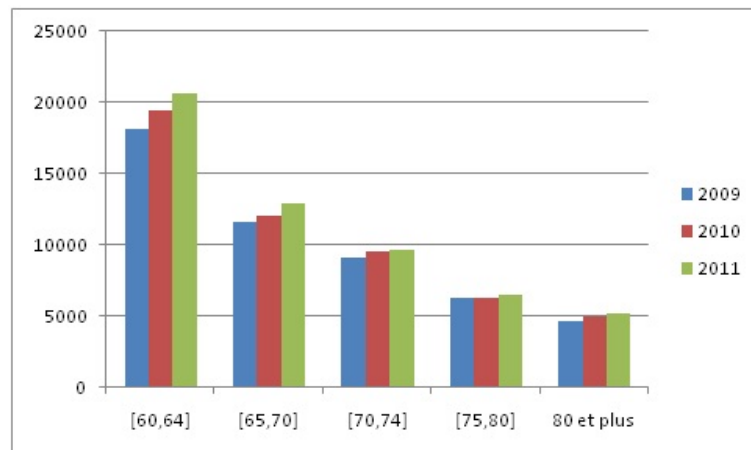


FIGURE 5 – L'évolution du nombre de contrats chez les seniors

On voit que le nombre de contrats de la "population âgée" évolue ces 3 dernières années. Il y a une forte augmentation dans la tranche d'âge de 60 à 64 ans (6,7%) et une augmentation moins marquée pour les autres tranches d'âge. Pour les personnes âgées de plus de 80 ans, l'accroissement est de 5,5% en moyenne chaque année (Table 4).

	[60,64]	[65,70]	[70,74]	[75,80]	80 ans et plus
Le pourcentage d'accroissement moyen par an	6,7 %	5,1 %	3,2 %	1,7 %	5,5 %

TABLE 4 – Le pourcentage d'accroissement moyen par an du nombre de contrats chez les personnes âgées

Une question se pose : accroissement vient il des nouveaux contrats chaque année, ou vient du changement de tranche d'âge du conducteur âgé ?

La table 5 présente la différence entre le nombre de nouveaux contrats et celui de contrats résiliés, de 2009 à 2010 et de 2010 à 2011, selon les tranches d'âge :

	[60,64]	[65,70]	[70,74]	[75,80]	80 ans et plus
2009-2010	-227	-128	-168	-528	-781
2010-2011	-6	-294	-110	-488	-800

TABLE 5 – La différence entre le nombre de nouveaux contrats et celui de contrats résiliés

Nous voyons que le nombre de contrats résiliés chaque année est plus élevé que celui de nouveaux contrats. C'est-à-dire il y a beaucoup de personnes âgées qui ont cessé de conduire et c'est assez évident que le nombre des personnes âgées qui souscrivent un nouveau contrat est faible. La tendance d'accroissement provient donc essentiellement du changement des tranches d'âge du conducteur âgé, comme on pouvait s'en douter du fait du vieillissement de la population.

6.2 Etude de la fréquence et du coût moyen des sinistres et analyse du ratio S / C

Nous avons calculé la fréquence moyenne et le coût moyen des sinistres sur 3 années 2009, 2010, 2011 et le ratio S/C selon les formules suivantes :

$$\begin{aligned}\text{Fréquence} &= \frac{\text{nombre de sinistres}}{\text{nombre d'années police}} \times 100 \\ \text{Coût moyen} &= \frac{\text{charge des sinistres}}{\text{nombre de sinistres}} \\ \text{S/C} &= \frac{\text{charge des sinistres}}{\text{primes}}\end{aligned}$$

Le ratio S/C permet de mettre en évidence d'éventuelles sur-tarifications ou sous-tarifications selon les tranches d'âge.

La Table 6 présente les résultats de ces 3 années que l'on considère :

Age du conducteur	Coût moyen sinistres (euro)	S/C	Fréquence
19 ans et moins	3 259	151,6%	31,46
20 ans à 24 ans	2 701	114,1%	26,32
25 ans à 29 ans	1 404	59,6%	22,86
30 ans à 34 ans	1 415	64,4%	22,04
35 ans à 39 ans	1 444	66,6%	20,43
40 ans à 44 ans	1 604	75,8%	19,79
45 ans à 49 ans	1 406	68,5%	19,72
50 ans à 54 ans	2 000	96,4%	18,92
55 ans à 59 ans	1 506	67,8%	17,12
60 ans à 64 ans	1 269	56,8%	16,29
65 ans à 70 ans	1 358	60,5%	15,63
70 ans à 74 ans	1 566	71,5%	15,37
75 ans à 80 ans	3 030	126,0%	13,49
80 ans et plus	3 258	116,7%	12,05

TABLE 6 – Le coût moyen et la fréquence des sinistres et le ratio S/C en 2011

Pour avoir une comparaison plus concrète entre les tranches d'âge, on donne en Figure 6 le graphique du ratio S/C.

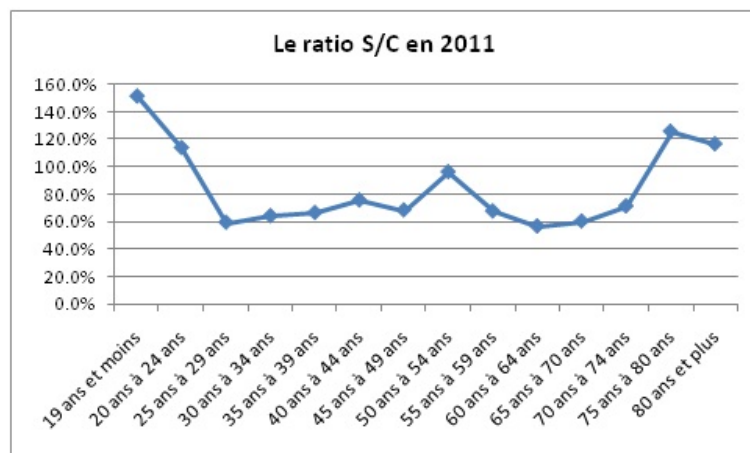


FIGURE 6 – Le ratio S/C en 2011

Selon la Table 6, nous y observons un coût moyen de sinistres largement supérieur dans les tranches d'âge [75,80], 80 ans et plus. Nous pouvons expliquer cette tendance par le fait que les sinistres corporels sont souvent plus graves et donc plus coûteux que dans les autres tranches d'âge. On voit aussi que les coûts

moyens des sinistres chez les personnes de tranches d'âge [19 ans et moins], [20,24], [50,54] sont importants par rapport à ceux des personnes d'âge moyen. Pour les 2 tranches d'âge [19 ans et moins] et [20,24], les coûts moyens sont les plus importants, cela se comprend du fait du manque d'expérience de conduite des jeunes. Les accidents sont plus fréquents. En revanche, pour les personnes d'âge de 50 à 54 ans, le coût des sinistres est étonnamment important en 2011. Par la suite, on donne un graphique des coûts des sinistres pendant 3 années afin de comparer les résultats (Figure 7) :

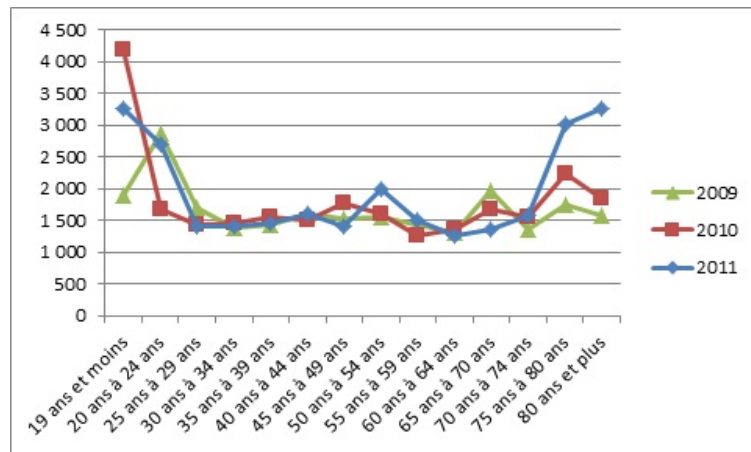


FIGURE 7 – Les coûts moyens des sinistres pendant 3 années 2009, 2010, 2011

A travers l'analyse de Figure 7, on voit que c'est uniquement en 2011 que le coût moyen des sinistres des personnes d'âge de 50 à 54 ans présente un pic, sans doute dû à un ou plusieurs sinistres exceptionnels.

En étudiant la Table 6 et la Figure 8, nous remarquons que les fréquences des accidents chez les personnes âgées sont moins élevées que chez les jeunes, alors que c'est le rapport inverse pour les coûts moyens sinistres. Cela se comprend du fait d'une part de la modification des habitudes de conduite des personnes âgées selon leurs capacités (ils conduisent moins dans les conditions difficiles : moins la nuit, aux heures de pointe,... ; ils routent moins vite,...). D'une autre part, au niveau de la sécurité, les conducteurs âgés sont les plus fragiles, ce qui les rend plus susceptibles d'être gravement blessés. Les résultats sur les données de Swiss Life donc une fois encore confirment que les accidents et agressions corporelles chez les personnes âgées sont moins fréquents que chez les jeunes, mais plus graves.

Les ratios S/C pour les tranches d'âge de 60 à 74 ans sont inférieurs à 100%. Ceci signifie que les primes perçues par Swiss Life au titre de ces tranches d'âge

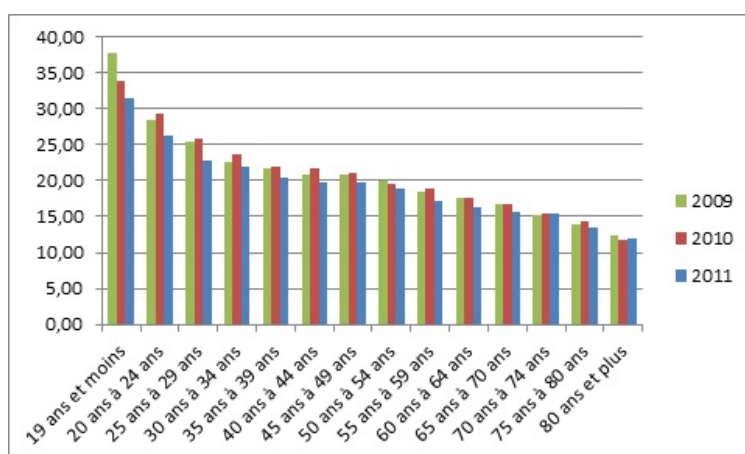


FIGURE 8 – La fréquence des sinistres pendant 3 années 2009, 2010, 2011

permettent de régler l'ensemble des sinistres. La partie des primes non utilisée pour régler les sinistres sert ensuite à payer les frais généraux, les coûts fixes et à rémunérer le réseau (le tout représentant entre 25% et 30% des primes) ainsi que l'actionnaire. En revanche, les ratios S/C pour les deux tranches d'âge de 75 ans à plus de 80 ans atteignent plus de 100%. Ceci signifie que les primes perçues par Swiss Life au titre de ces 2 tranches d'âge ne suffisent pas à couvrir l'ensemble des sinistres de cette sous-population.

L'analyse de Table 6 permet de dire que les deux tranches d'âge [75,80] et 80 ans et plus sont largement déficitaires pour Swiss Life. En revanche, Swiss Life est bénéficiaire sur les autres tranches d'âge (à partir de 60 ans), et notamment sur la tranche 60 à 64 ans, 65 à 70 ans et 70 à 74 ans du conducteur.

Nous donnons ci-dessous une évolution du ratio S/C selon les tranches d'âge pendant les années 2009, 2010, 2011 (Figure 9) :

Selon la Figure 9, nous pouvons y remarquer que les ratios S/C évoluent chaque année. Les ratios S/C de la tranche d'âge [75,80], 80 ans et plus sont augmentés fortement pendant 3 années. Les autres tranches d'âge évoluent non-régulièrement. Il est difficile de prévoir leur évolution pour les années suivantes. On remarque aussi que les ratios S/C de la tranche d'âge [75,80] sont presque plus élevés (pas en 2009) pendant ces 3 années considérées. C'est-à-dire, les gens appartenant à cette tranche d'âge coûtent plus que les autres.

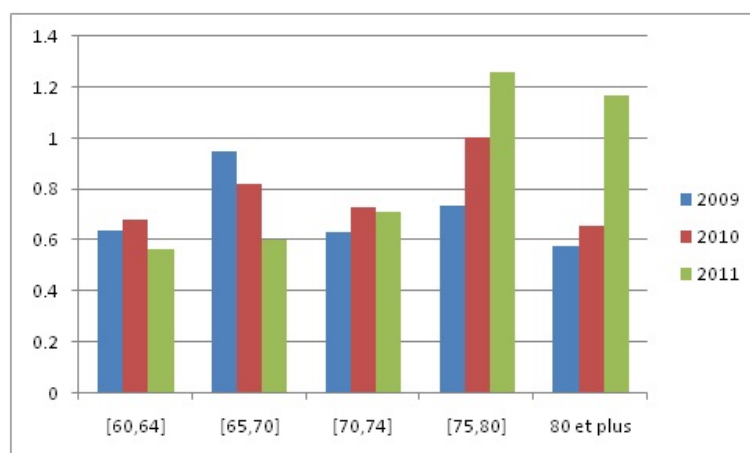


FIGURE 9 – L'évolution du ratio S/C selon les tranches d'âge

6.3 Analyse en Composantes Principales (ACP)

L'analyse en Composantes principales (ACP) fait partie du groupe des méthodes descriptives multidimensionnelles appelée méthodes factorielles. Ces méthodes, sont apparues au début des années 30, ont été surtout développées en France dans les années 60, en particulier par Jean-Paul Benzécri qui a introduit une approche géométrique et exploité les représentations graphiques. Dans la mesure où ce sont des méthodes descriptives, elles ne s'appuient pas sur un modèle probabiliste, mais elles dépendent d'un modèle géométrique. L'ACP propose, à partir d'un tableau rectangulaire de données comportant les valeurs de p variables quantitatives pour n unités (appelées aussi individus), des représentations géométriques de ces unités et de ces variables. Ces données peuvent être issues d'une procédure d'échantillonnage ou bien de l'observation d'une population toute entière. Les représentations des unités permettent de voir s'il existe une structure, non connue a priori, sur cet ensemble d'unités. De façon analogue, les représentations des variables permettent d'étudier les structures de liaisons linéaires sur l'ensemble des variables considérées. Ainsi, on cherchera si l'on peut distinguer des groupes dans l'ensemble des unités en regardant quelles sont les unités qui se ressemblent, celles qui se distinguent des autres, etc. Pour les variables, on cherchera quelles sont celles qui sont très corrélées entre elles, celles qui, au contraire ne sont pas corrélées aux autres, etc. Rappelons brièvement la méthodologie de l'ACP.

6.3.1 Méthodologie

▷ Tableau de données

Les données sont les mesures effectuées sur n unités $\{u_1, u_2, \dots, u_i, \dots, u_n\}$. Les p variables quantitatives qui représentent ces mesures sont $\{v_1, v_2, \dots, v_j, \dots, v_p\}$.

Le tableau des données brutes à partir duquel on va faire l'analyse est noté X et a la forme suivante :

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2j} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{i1} & x_{i2} & \dots & x_{ij} & \dots & x_{ip} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nj} & \dots & x_{np} \end{bmatrix} \quad (1)$$

On peut représenter chaque unité par le vecteur de ses mesures sur les p variables :

$${}^tU_i = [x_{i1} \ x_{i2} \dots x_{ij} \dots x_{ip}] \quad (2)$$

U_i est un vecteur de R^p .

De façon analogue, on peut représenter chaque variable par un vecteur de R^n dont les composantes sont les valeurs de la variable pour les n unités :

$${}^tV_j = [x_{1j} \ x_{2j} \dots x_{ij} \dots x_{nj}] \quad (3)$$

Pour avoir une image de l'ensemble des unités, on se place dans un espace affine en choisissant comme origine un vecteur particulier de R^p , par exemple le vecteur dont toutes les coordonnées sont nulles. Alors, chaque unité sera représentée par un point dans cet espace. L'ensemble des points qui représentent les unités est appelé traditionnellement "nuage des individus".

En faisant de même dans R^n , chaque variable pourra être représentée par un point de l'espace affine correspondant. L'ensemble des points qui représentent les variables est appelé "nuage des variables".

On constate que ces espaces étant de dimension supérieure en général à 2 et même 3, on ne peut visualiser ces représentations.

L'idée générale des méthodes factorielles est de trouver un système d'axes et de plans tels que les projections de ces nuages de points sur ces axes et ces plans permettent de reconstituer les positions des points les uns par rapport aux autres, c'est-à-dire avoir des images les moins déformées possible.

▷ Choix de l'origine

$\forall i = 1, \dots, n$, on note p_i le poids de l'unité u_i tel que $\sum_{i=1}^n p_i = 1$. Par définition le centre de gravité G est défini comme le point tel que :

$$\sum_{i=1}^n p_i \overrightarrow{Gu_i} = \overrightarrow{0}$$

Pour l'ACP on choisit de donner le même poids $\frac{1}{n}$ à tous les individus.

Le centre de gravité G du nuage des individus est alors le point dont les coordonnées sont les valeurs moyennes des variables :

$$G = \begin{pmatrix} \frac{1}{n} \sum_{i=1}^n x_{i1} \\ \dots \\ \frac{1}{n} \sum_{i=1}^n x_{ij} \\ \dots \\ \frac{1}{n} \sum_{i=1}^n x_{ip} \end{pmatrix} = \begin{pmatrix} x_{\bullet 1} \\ \dots \\ x_{\bullet j} \\ \dots \\ x_{\bullet p} \end{pmatrix} \quad (4)$$

Prendre G comme origine, conformément à la figure suivante, revient alors à travailler sur le tableau des données centrées :

$$X_c = \begin{bmatrix} x_{11} - x_{\bullet 1} & \dots & x_{1j} - x_{\bullet j} & \dots & x_{1p} - x_{\bullet p} \\ \dots & \dots & \dots & \dots & \dots \\ x_{i1} - x_{\bullet 1} & \dots & x_{ij} - x_{\bullet j} & \dots & x_{ip} - x_{\bullet p} \\ \dots & \dots & \dots & \dots & \dots \\ x_{n1} - x_{\bullet 1} & \dots & x_{nj} - x_{\bullet j} & \dots & x_{np} - x_{\bullet p} \end{bmatrix} \quad (5)$$

et le vecteur des coordonnées centrées de l'unité u_i est :

$$U_{ci} = [x_{i1} - x_{\bullet 1} \dots x_{ij} - x_{\bullet j} \dots x_{ip} - x_{\bullet p}] \quad (6)$$

celui des coordonnées centrées de la variable v_j est :

$$V_{cj} = [x_{1j} - x_{\bullet j} \dots x_{ij} - x_{\bullet j} \dots x_{nj} - x_{\bullet j}] \quad (7)$$

▷ Inertie totale du nuage des individus

On note I_G le moment d'inertie du nuage des individus par rapport au centre de gravité

$$I_G = \frac{1}{n} \sum_{i=1}^n d^2(G, u_i) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^p (x_{ij} - x_{\bullet j})^2 = \frac{1}{n} \sum_{i=1}^n {}^t U_{ci} U_{ci}.$$

Ce moment d'inertie totale est intéressant car c'est une mesure de la dispersion du nuage des individus par rapport à son centre de gravité. Si ce moment d'inertie est grand, cela signifie que le nuage est très dispersé, tandis que s'il est petit, alors le nuage est très concentré sur son centre de gravité.

L'inertie du nuage des individus par rapport à un axe passant par G est égale, par définition, à :

$$I_{\Delta} = \frac{1}{n} \sum_{i=1}^n d^2(h_{\Delta i}, u_i)$$

ou $h_{\Delta i}$ est la projection orthogonale de u_i sur l'axe Δ et d est la distance euclidienne classique.

Soit V un sous-espace de R^n . Cette inertie est, par définition, égale à :

$$I_V = \frac{1}{n} \sum_{i=1}^n d^2(h_{V i}, u_i)$$

ou $h_{V i}$ est la projection orthogonale de u_i sur le sous-espace V.

Soit V^* le complément orthogonal de V, on a

$$d^2(h_{V i}, u_i) + d^2(h_{V^* i}, u_i) = d^2(G, u_i)$$

donc

$$I_G = I_V + I_{V^*}$$

Si $\mathbb{R}^n = \Delta_1 \oplus \dots \oplus \Delta_p$ on a :

$$I_G = I_{\Delta_1} + I_{\Delta_2} + \dots + I_{\Delta_p}.$$

▷ Recherche de l'axe Δ_1 passant par G d'inertie minimal

On cherche un axe Δ_1 passant par G d'inertie I_{Δ_1} minimum car c'est l'axe le plus proche de l'ensemble des points du nuage des individus, et donc, si l'on doit projeter ce nuage sur cet axe, c'est lui qui donnera l'image la moins déformée du nuage. Si on utilise la relation entre les inerties donnée au paragraphe précédent, rechercher Δ_1 tel que I_{Δ_1} est minimum, est équivalent à chercher Δ_1 tel que $I_{\Delta_1^*}$ est maximum.

On définit l'axe Δ_1 par son vecteur directeur unitaire $\overrightarrow{Ga_1}$.

Il faut donc trouver $\overrightarrow{Ga_1}$ tel que I_{Δ_1} est maximum sous la contrainte que $\|\overrightarrow{Ga_1}\| = 1$

$$d^2(G, h_{\Delta_1^* i}) = {}^t a_1 U_{ci} {}^t U_{ci} a_1$$

en utilisant la symétrie du produit scalaire. On en déduit :

$$I_{\Delta_1^*} = \frac{1}{n} \sum_{i=1}^n {}^t a_1 U_{ci} {}^t U_{ci} a_1 = {}^t a_1 \left[\frac{1}{n} \sum_{i=1}^n U_{ci} {}^t U_{ci} \right] a_1.$$

Entre crochets on reconnaît la matrice de covariance empirique Σ des p variables

$$I_{\Delta_1^*} = {}^t a_1 \Sigma a_1$$

et

$$\|\overrightarrow{Ga_1}\| = {}^t a_1 a_1.$$

Le problème à résoudre est trouver a_1 tel que ${}^t a_1 \Sigma a_1$ soit maximum avec la contrainte ${}^t a_1 a_1 = 1$ où Σ est la matrice de covariance.

Il s'agit du problème de la recherche d'un optimum d'une fonction de plusieurs variables liées par une contrainte (les inconnues sont les composantes de a_1). La méthode des multiplicateurs de Lagrange peut alors être utilisée (cf. annexe page 55).

Dans le cas de la recherche de a_1 , il faut calculer les dérivées partielles de :

$$g(a_1) = g(a_{11}, a_{12}, \dots, a_{1p}) = {}^t a_1 \Sigma a_1 - \lambda_1 ({}^t a_1 a_1 - 1).$$

En utilisant la dérivée matricielle, on obtient :

$$\frac{\partial g(a_1)}{\partial a_1} = 2 \Sigma a_1 - 2\lambda_1 a_1 = 0.$$

Le système à résoudre est :

$$\begin{cases} \Sigma a_1 - \lambda_1 a_1 = 0 \\ {}^t a_1 a_1 - 1 = 0 \end{cases} \quad (8)$$

De l'équation matricielle de ce système on déduit que a_1 est vecteur propre de la matrice Σ associé à la valeur propre λ_1 . En multipliant à gauche par ${}^t a_1$ les deux membres de l'équation 8 on obtient :

$${}^t a_1 \Sigma a_1 = \lambda_1.$$

On reconnaît que le premier membre de l'équation précédente est égal à l'inertie $I_{\Delta_1^*}$ qui doit être maximum. Cela signifie que la valeur propre λ_1 est la plus grande valeur propre de la matrice de covariance Σ et que cette valeur propre est égale à l'inertie portée par l'axe Δ_1 .

L'axe Δ_1 pour lequel le nuage des individus a l'inertie minimal a comme vecteur directeur unitaire le premier vecteur propre associé à la plus grande valeur propre de la matrice de covariance Σ .

On présente ici une étude de l'ACP sur les données brutes en 2011 chez Swiss Life. Les intitulés complets des variables sont :

Police : le numéro de police

Age_cond : l'âge du conducteur

Durée_contrat : la durée du contrat

An_cond : le nombre d'années de conduite du conducteur à partir de la date d'obtention du permis de conduire

An_route : l'âge du véhicule à partir de la date de première mise en circulation
Sex : le sexe du conducteur. On remarque que le sexe est une variable binaire, on la considère donc comme une variable quantitative.

Nbre_sin : le nombre de sinistre

Cout_sinistre : le montant de sinistres

Coefbm : le coefficient bonus-malus du composant contrat auto⁴

Frmalus : le franchise malus du composant⁵

Ctnet : le coefficient tarifaire

6.3.2 Application aux données Swiss Life en 2011

Les listings et les graphiques présentés de la page 56 à la page 64 sont les résultats d'une ACP normée sur données brutes chez Swiss Life en 2011. Le fait de normer l'ACP permet d'éliminer les effets de masse qui induisaient la prédominance des variables sur les autres.

On note sur la sortie page 56 que SAS utilise la matrice des corrélations entre les variables. On lit bien, page 56, que la somme des valeurs propres vaut 12 (colonne Fréquence cumulé du Tableau des valeurs propres), c'est-à-dire le nombre de variables explicatives. Ici les trois premiers axes ne portent que 46,65% de l'inertie totale et il faut aller jusqu'au septième axe pour obtenir 80.13% (colonne Pctage cumulé du Tableau des valeurs propres).

Nous donnons ci-dessous la Table 7 de qualité de la présentation et contributions des variables.

Le Table 7 détaille d'abord la contribution (CTR) (en proportion de variance) des descripteurs à l'établissement des axes. Les descripteurs : le coefficient bonus-malus, le franchise malus ont les contributions de variance les plus fortes.

Il donne ensuite la proportion de variance de chaque variable qui est expliquée par l'axe, noté COS2, et les QLT correspondant aux sommes cumulées des COS2.

4. Le coefficient bonus-malus applicable chez Swiss Life est celui utilisé à la précédente échéance réduit de 5 % arrêté à la deuxième décimale et arrondi par défaut après chaque période annuelle d'assurance sans sinistre. Le coefficient d'origine est de 1,00 et le coefficient bonus-malus ne peut être inférieur à 0,50. Aucune majoration n'est appliquée pour le premier sinistre survenu après une première période d'au moins 3 ans au cours de laquelle le coefficient bonus-malus a été égal à 0,50. Un sinistre survenu au cours de la période annuelle d'assurance majore le coefficient de 25%, un second sinistre majore le coefficient obtenu de 25% ; et il en est de même pour chaque sinistre supplémentaire. Le coefficient bonus-malus ne peut être supérieur à 3,50.

5. Le franchise malus du composant est un indice pour savoir pendant combien d'années le sinistre n'est pas survenu. Cet indice prend 3 valeurs : 1,2,3. S'il n'y a pas de sinistre pendant une période plus 3 ans, cet indice est égal à 3.

Name	CTR1	CTR2	CTR3	CTR4	CO21	CO22	CO23	CO24	QLT12	QLT13	QLT14
Sex	0.36	0.64	9.11	0.07	0.91	1.17	11.09	0.07	2.08	13.18	13.25
age_cond	9.94	39.07	0.41	0.18	25.27	71.82	0.50	0.21	97.09	97.59	97.80
an_cond	9.69	39.40	0.50	0.25	24.63	72.42	0.61	0.28	97.05	97.67	97.94
an_route	0.03	0.17	35.77	1.72	0.06	0.31	43.58	1.95	0.38	43.95	45.90
Coefbm	23.57	6.53	0.06	0.15	59.91	12.00	0.08	0.16	71.91	71.98	72.15
cout_sinistre	0.02	0.00	5.85	7.83	0.04	0.00	7.13	8.86	0.05	7.18	16.04
Ctnet	1.75	1.45	17.76	22.21	4.46	2.66	21.64	25.12	7.12	28.76	53.88
durée_contrat	4.69	0.19	5.06	20.87	11.91	0.35	6.17	23.61	12.27	18.43	42.04
Frmaalus	23.67	5.96	0.01	0.09	60.16	10.96	0.01	0.10	71.12	71.13	71.24
nbre_sin	0.44	0.07	25.30	14.21	1.11	0.12	30.82	16.08	1.23	32.05	48.13

TABLE 7 – La qualité de la présentation et contributions des variables

Donc, l'axe 1 est expliqué par le coefficient bonus-malus, et le franchise malus du composant. L'axe 2 est bien expliqué par les deux variables l'âge du conducteur et le nombre d'années de conduite du conducteur. Les descripteurs l'âge du conducteur, le nombre d'années de conduite du conducteur, le coefficient bonus-malus, et le franchise malus du composant sont pratiquement expliqué dans le plan des deux premiers axes ($QLT12 > 70\%$). Les variables le sexe, l'âge du véhicule, le coût de sinistre, le nombre de sinistres sont très mal représenté dans le plan des deux premiers axes ($QLT12 < 5\%$). Il faut étudier les autres axes pour avoir une analyse plus concrète.

Le graphique du cercle des corrélations est présenté ci-dessous (Figure 16).

Le cercle des corrélations des descripteurs avec les 2 premiers axes principaux s'inscrit dans ce rectangle. Plus un descripteur est proche d'un axe, plus il y contribue. Plus il est proche de la circonférence, plus il est expliqué par l'axe. Les 5 descripteurs : l'âge du conducteur, le nombre d'années de conduite du conducteur, le coefficient bonus-malus, et le franchise malus du composant sont donc bien expliqués par les axes.

Sur le graphique du cercle de corrélation, on peut aussi interpréter les positions des anciennes variables les unes par rapport aux autres en termes de corrélations. Deux points très proches du cercle des corrélations, donc bien représentées dans le plan, seront très corrélés positivement entre elles. Si elles sont proches du cercle, mais dans des positions symétriques par rapport à l'origine, elles seront très corrélées négativement.

Deux variables proches du cercle des corrélations et dont les vecteurs qui les

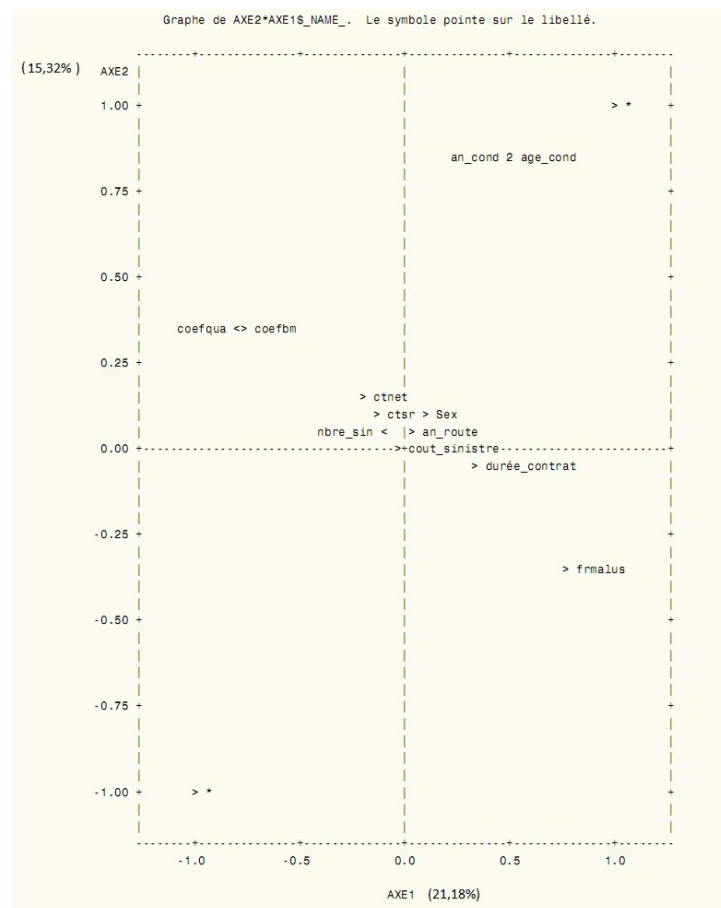


FIGURE 10 – Le résultat de l'ACP pour les axes 1 et 2

joignent à l'origine forment un angle droit, ne seront pas corrélées entre elles.

Il faut, pour interpréter correctement ces graphiques de cercle de corrélations, se souvenir qu'un coefficient de corrélation est une mesure de liaison linéaire entre deux variables, et qu'il existe d'autre type de dépendance, il peut arriver que deux variables très fortement liées aient un coefficient de corrélation nul ou très faible, si leur liaison n'est pas linéaire.

Sur le graphique du cercle de corrélations (Figure 16) des 2 premiers axes principaux, on remarque qu'il y a une forte corrélation positive entre les 2 variables l'âge du conducteur et le nombre d'années de conduite du conducteur alors qu'une corrélation négative entre le coefficient bonus-malus avec le franchise malus du composant.

Des cercles sont produits pour les principales combinaisons des 3 premiers axes.

Le cercle de corrélations des descripteurs avec le premier axe principal et le troisième est donné dans la Figure 11.

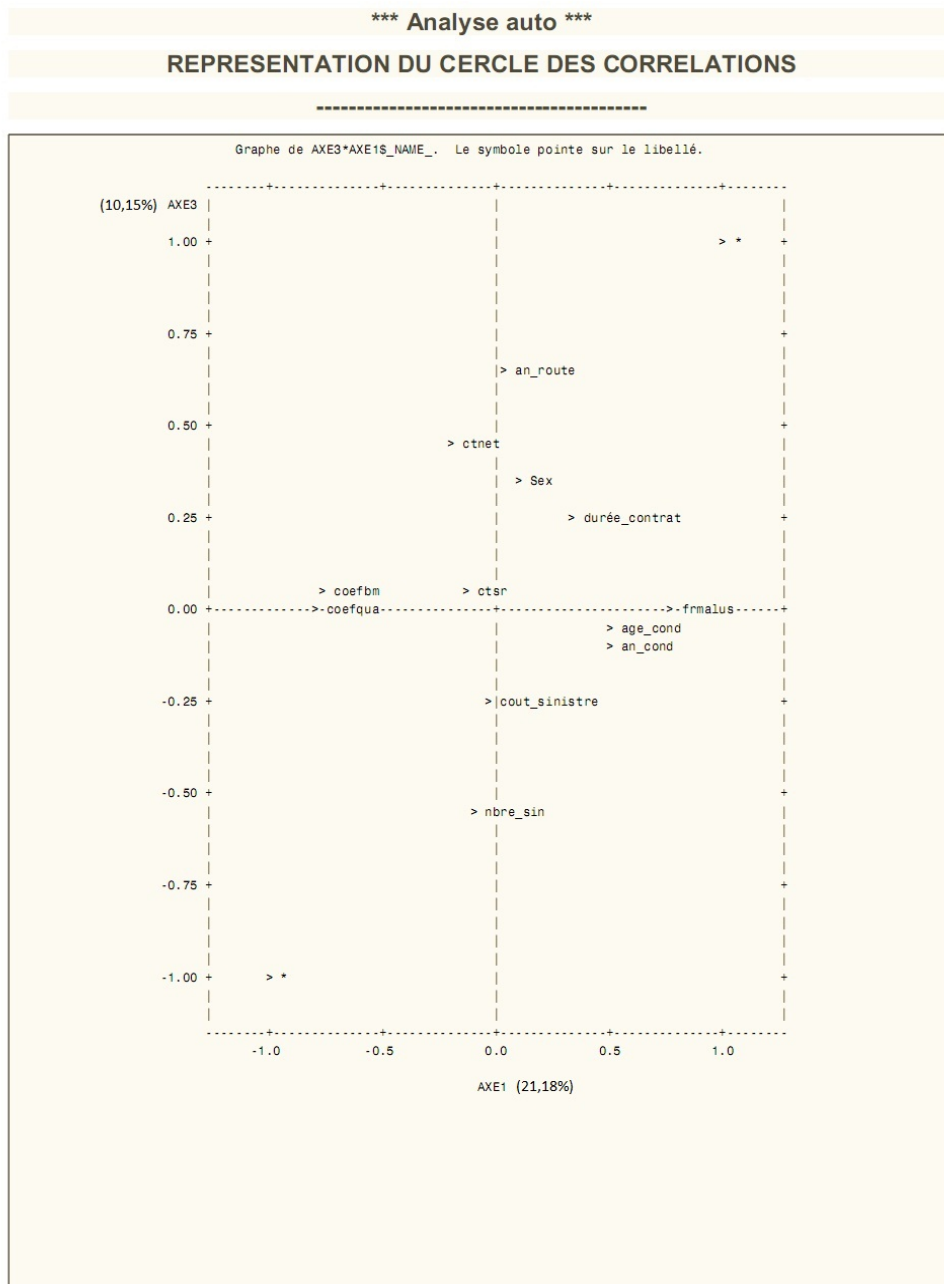


FIGURE 11 – Le résultat de l'ACP pour les axes 1 et 3

Sur le Figure 11, on remarque que les descripteurs le nombre de sinistres, le coût de sinistre, l'âge du véhicule, le coefficient tarifaire, le sexe, la durée du contrat

sont assez bien expliqués par le troisième axe principal. On voit aussi une forte corrélation négative entre 2 groupes de variables : l'âge du véhicule, la sexe, la durée du contrat, le coefficient tarifaire avec le nombre de sinistres, le coût de sinistre.

En résumé, après l'ACP fait sur les données de Swiss Life, on note que :

- Le nombre ou la fréquence de sinistres et le coût de sinistres ne semblent pas liés linéairement à l'âge du conducteur sur les données chez Swiss Life contrairement à l'étude sur la population française.
- Les facteurs les plus importants apparaissant dans l'ACP, et donc à conserver, sont les suivants : l'âge du conducteur, le nombre d'années de conduite du conducteur, le coefficient bonus-malus, la sexe, le nombre de sinistres, le coût de sinistre, le coefficient tarifaire, l'âge du véhicule, la durée du contrat.

6.4 Analyse des correspondances multiples (ACM)

6.4.1 Introduction

Tout d'abord, il faut noter que la technique de l'ACP présentée dans la section précédente ne s'applique que à des variables quantitatives. Nous allons donc utiliser des techniques spécifiques pour variables qualitatives. Une de ces techniques se nomme l'Analyse des Correspondances Multiples (ACM). Le but de l'ACM est le même que celui de l'ACP. Il s'agit de représenter graphiquement les individus par des points dans un sous-espace de manière à ce que le nuage de points ressemble le plus au nuage de points de l'espace original. On verra que l'on peut procéder exactement comme pour l'ACP, à ceci près qu'on ne mesure pas les distances entre les individus avec la métrique canonique (euclidienne), mais à l'aide de la «distance du Khi-deux».

6.4.2 Méthodologie

▷ Table des données

L'ACM permet l'étude de tableaux décrivant une population de I individus et J variables qualitatives. Une variable qualitative (ou nominale) peut être décrite par une application de l'ensemble des I individus dans un ensemble fini non structuré, par exemple non ordonné. Ces variables qualitatives peuvent être codées par un codage condensé qui attribue une valeur à chaque modalité. Les données peuvent donc être représentées sous la forme d'une matrice $X = (x_{ij})_{I \times J}$ où I représente le nombre d'individus, J présente le nombre de variables et x_{ij} est le codage condensé de l'individu i pour la variable j .

Les x_{ij} représentant une codification, en prendre la moyenne n'a aucun sens. Ces données ne peuvent donc pas être traitées par l'ACP précédemment étudiées.

Ce tableau présente donc des spécificités dont l'analyse factorielle doit tenir compte par une méthode spécifique.

▷ La transformation des données

Diagram illustrating the structure of a 2D array (matrix) with dimensions I (rows) and K (columns). The array is partitioned into four blocks:

- Top-left block: 0100000 (row i , column 1)
- Top-right block: 0000100 (row i , column K)
- Bottom-left block: x_{ik} (row i , column k)
- Bottom-right block: x_{ik} (row i , column k)

The array is also partitioned into four blocks along the diagonal:

- Top-left block: I_1 (row 1, column 1)
- Top-right block: I_k (row k , column k)
- Bottom-left block: I_k (row k , column k)
- Bottom-right block: I_K (row K , column K)

FIGURE 12 – La représentation des données transformées

$$x_{ik} = \begin{cases} 1 & \text{si l'individu } i \text{ possède la modalité } k \\ 0 & \text{sinon} \end{cases} \quad (9)$$

où x_{ik} est à l'intersection de la ligne i avec la colonne k .

Ce tableau porte le nom de disjonctif complet, car l'ensemble des valeurs x_{ik} d'un même individu pour les modalités d'une même variable comporte la valeur 1 une fois (complet) et une fois seulement (disjonctif). Chaque modalité k est reliée à une variable j . Nous avons ainsi trois familles d'éléments : les individus, les variables et les modalités.

Notons K_j le nombre de modalités de la variable j . Ainsi $K = \sum_{j=1}^p K_j$ est le nombre de modalités, toutes variables confondues. Nous avons donc les égalités

suivantes :

$$\begin{aligned} \sum_{k=1}^{K_j} x_{ik} = 1 \quad \forall(i, j); & \quad \sum_{k=1}^K x_{ik} = J \quad \forall i; \\ \sum_{i=1}^I x_{ik} = I_k \quad \forall k; & \quad \sum_{k=1}^{K_j} I_k = I \quad \forall j. \end{aligned}$$

Les propriétés intéressantes de l'ACM sont essentiellement dues aux propriétés des tableaux disjonctifs complets. Notons surtout que c'est un tableau binaire dont les lignes ont la somme de leur coefficients constantes égale à J et dont les colonnes sont regroupées par paquet correspondant à une variable avec pour somme par ligne égale à 1.

▷ **L'analyse factorielle des correspondances du tableau disjonctif complet**

Nous allons considérer le tableau disjonctif complet en profils-lignes et en profils-colonnes. Pour ce faire, nous modifions ce tableau afin de considérer les fréquences (Figure 13).

	1		k		K	marge
1	$\frac{x_{ik}}{IJ}$					$\frac{1}{I}$
⋮						
⋮						
i						
⋮						
I						1
marge	$\frac{I_k}{IJ}$					

FIGURE 13 – Le tableau conjoint complet

Il faut noter que la marge totale de ce tableau est égale à IJ. La fréquence f_{ik} de chaque élément binaire x_{ik} est données par :

$$f_{ik} = \frac{x_{ik}}{IJ}.$$

De plus les marges sont données par :

$$f_{i\bullet} = \sum_{k \in K} \frac{x_{ik}}{IJ} = \frac{1}{I}$$

$$f_{\bullet k} = \sum_{i \in I} \frac{x_{ik}}{IJ} = \frac{1}{J}$$

Une fois les fréquences calculées, il faut considérer le tableau en profils-lignes et profils-colonnes comme le tableau suivant (Figure 14).

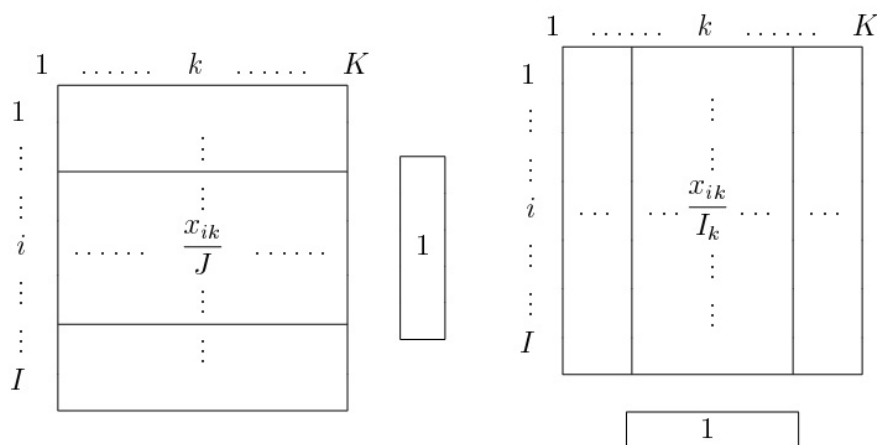


FIGURE 14 – Le tableau conjoint complet modifié

Ainsi le tableau est de nouveau modifié de façon à ce que pour les profils-lignes la marge des lignes soit 1 et pour les profils-colonnes la marge des colonnes soit 1. Ainsi chaque case est composée respectivement de $\frac{x_{ik}}{J}$ et $\frac{x_{ik}}{I_k}$.

▷ L'analyse des nuages

Chaque individu du nuage des individus, noté par N_I , est représenté par les modalités qu'il possède. La marge étant constante, la transformation en profils-lignes ne modifie en rien les données. Ainsi le nuage N_I appartient à un hypercube noté H_I d'arrête $\frac{1}{J}$, puisque le profil d'une ligne est soit 0 soit $\frac{1}{J}$. Un individu i est un point de R^K qui a pour coordonnée sur l'axe k la valeur $\frac{x_{ik}}{J}$ avec un poids identique pour chaque individu (car la marge est constante) de $\frac{1}{J}$. Le barycentre, ou centre de gravité, G_I du nuage N_I a pour coordonnée $\frac{I_k}{IJ}$ sur l'axe k . La ressemblance entre deux individus est définie par les modalités de chacun des individus. Si les deux individus présentent globalement les mêmes modalités, alors ils se ressemblent. La

distance Khi-deux qui caractérise cette ressemblance entre deux individus i et l est définie par :

$$d^2(i, l) = \sum_{k=1}^K \frac{IJ}{I_k} \left(\frac{x_{ik}}{J} - \frac{x_{lk}}{J} \right)^2 = \frac{1}{J} \sum_{k=1}^K \frac{I}{I_k} (x_{ik} - x_{lk})^2 \quad (10)$$

Cette expression est remarquable car

$$(x_{ik} - x_{lk})^2 = \begin{cases} 1 & \text{si un seul individu possède la modalité } k \\ 0 & \text{sinon} \end{cases}$$

Cette distance croît logiquement avec le nombre de modalités qui diffèrent pour les individus i et l , ce qui est recherché. Le poids de la modalité k dans la distance est l'inverse de sa fréquence $\frac{I}{I_k}$. Ainsi si un individu possède une modalité rare, il sera éloigné de tous les autres individus et du centre de gravité.

La ressemblance entre deux modalités k et h est donnée par la distance :

$$d^2(k, h) = \sum_{i \in I} I \left(\frac{x_{ik}}{I_k} - \frac{x_{ih}}{I_h} \right)^2 \quad (11)$$

En notant que $(x_{ik})^2 = x_{ik}$ qui ne prennent que les valeurs 1 ou 0, cette distance peut s'écrire :

$$d^2(k, h) = \frac{I}{I_k I_h} (I_k + I_h - 2 \sum_{i \in I} x_{ik} x_{ih}) \quad (12)$$

ce qui est le nombre d'individus possédant une et une seule des deux modalités h ou k multiplié par $\frac{I}{I_k I_h}$. Cette distance croît donc avec le nombre d'individus possédant une et une seule des deux modalités k et h et décroît avec l'effectif de chacune de ces modalités. Ainsi, par construction, deux modalités d'une même variable sont éloignées l'une de l'autre (puisqu'elles ne peuvent pas être possédées par le même individu). Deux modalités possédées par exactement les mêmes individus sont confondues, tandis que les modalités rares sont éloignées de toutes les autres et du centre de gravité G_K .

6.4.3 Application aux données Swiss Life en 2011

Le but de cette partie est de mieux comprendre l'influence de certains caractères des assurés sur le type de sinistre, notamment le sinistre corporel et non corporel.

▷ Présentation des données étudiées

La structure des données est présentée dans la Table 8.

Variables	N°	Modalités	Abréviation
Classe d'âge	1	19 ans et moins	< 19
	2	20 ans à 24 ans	[20,24]
	3	25 ans à 29 ans	[25,29]
	4	30 ans à 34 ans	[30,34]
	5	35 ans à 39 ans	[35,39]
	6	40 ans à 44 ans	[40,44]
	7	45 ans à 49 ans	[45,49]
	8	50 ans à 54 ans	[50,54]
	9	55 ans à 59 ans	[55,59]
	10	60 ans à 64 ans	[60,64]
	11	65 ans à 70 ans	[65,70]
	12	70 ans à 74 ans	[70,74]
	13	75 ans à 80 ans	[75,80]
	14	80 ans et plus	> 80
Coefficient bonus malus	15	50%	bm50
	16	51% à 59%	bm51-59
	17	60% à 69%	bm60-69
	18	70% à 79%	bm70-79
	19	80% à 89%	bm80-89
	20	90% à 99%	bm90-99
	21	100%	bm100
	22	plus de 100%	bm>100
Sexe du conducteur	23	Homme	H
	24	Femme	F
Situation professionnelle	25	Exploitant Agricole	Agri
	26	Artisan	Art
	27	Commerçant	Commerc
	28	Chef d'Entreprise	Chef
	29	Profession libérale	Pro.lib
	30	Cadre, Profession intellectuelle, scientifique	Cadre
	31	Technicien	Technic
	32	Employé	Employé
	33	Ouvrier	Ouvrier
	34	Retraité	Retraité
	35	Autre sans profession	Autre
	36	Non réponse	Cnon_reponse
Etat-civil	37	Exploitant Agricole	Marié_contrat
	38	Marié avec séparation de biens	Marié_sép_biens
	39	Célibataire	Célibataire
	40	Divorcé	Divorcé
	41	Concubin	Concubin
	42	Marié	Marié
	43	Pacsé	Pacsé
	44	Séparé	Séparé
	45	Union Libre	
	46	Veuf	Veuf
Type de sinistre	47	Corporel	OUI
	48	Non-Corporel	NON

TABLE 8 – La structure des données

On sait qu'il vaut mieux utiliser des variables comportant à peu près le même nombre de modalités. Dans ces données, le nombre de modalités de chaque variable est presque identique (plus ou moins 10) sauf les 2 variables : la sexe et le type de sinistre.

▷ Résultats relatifs aux modalités

On donne ci-dessous le tableau de qualité de la présentation et contributions des modalités (Table 9)

Name	Qualité	Inertie Dim.1	Inertie Dim.2	Cosinus2 Dim.1	Cosinus2 Dim.2
< 19	0,0018	0,0007	0,0000	0,0018	0,0000
> 80	0,2486	0,0282	0,1018	0,0682	0,1803
[20,24]	0,2425	0,0722	0,0399	0,1727	0,0698
[25,29]	0,3727	0,1319	0,0328	0,3153	0,0574
[30,34]	0,1830	0,0693	0,0044	0,1749	0,0081
[35,39]	0,0677	0,0246	0,0010	0,0656	0,0020
[40,44]	0,0463	0,0036	0,0179	0,0101	0,0363
[45,49]	0,0641	0,0011	0,0296	0,0030	0,0611
[50,54]	0,0748	0,0059	0,0293	0,0162	0,0586
[55,59]	0,0621	0,0134	0,0132	0,0360	0,0261
[60,64]	0,0950	0,0248	0,0138	0,0675	0,0275
[65,70]	0,0732	0,0276	0,0000	0,0732	0,0000
[70,74]	0,2014	0,0400	0,0551	0,1002	0,1012
[75,80]	0,2954	0,0348	0,1190	0,0843	0,2111
F	0,0371	0,0037	0,0085	0,0137	0,0234
M	0,0371	0,0022	0,0052	0,0137	0,0234
NON	0,0261	0,0003	0,0001	0,0215	0,0046
OUI	0,0261	0,0085	0,0025	0,0215	0,0046
Agri	0,0055	0,0013	0,0012	0,0033	0,0022
Art	0,0092	0,0033	0,0005	0,0083	0,0010
Autre	0,0056	0,0022	0,0000	0,0055	0,0001
Cadre	0,0014	0,0004	0,0003	0,0010	0,0005
Chef	0,0036	0,0014	0,0000	0,0036	0,0000
Cnon_rep	0,0777	0,0020	0,0001	0,0741	0,0036
Commerc	0,0245	0,0097	0,0003	0,0240	0,0005
Employé	0,0323	0,0044	0,0112	0,0113	0,0210
Ouvrier	0,0079	0,0001	0,0042	0,0003	0,0076
Pro.lib	0,0143	0,0029	0,0039	0,0072	0,0071
Retraité	0,4170	0,0398	0,1890	0,0931	0,3238
Technic	0,0073	0,0001	0,0039	0,0002	0,0071
Concubin	0,1741	0,0614	0,0059	0,1626	0,0115
Célibata	0,6602	0,2219	0,0614	0,5490	0,1112
Divorcé	0,0631	0,0090	0,0204	0,0238	0,0394
Libre	0,0124	0,0049	0,0001	0,0122	0,0002
Marié	0,4821	0,0877	0,0167	0,4230	0,0591
Marié_co	0,0004	0,0001	0,0001	0,0003	0,0001
Pacsé	0,0182	0,0073	0,0000	0,0182	0,0000
Snon_rep	0,0139	0,0000	0,0076	0,0000	0,0139
Séparé	0,0177	0,0002	0,0093	0,0006	0,0171
Veuf	0,4377	0,0469	0,1898	0,1103	0,3274

TABLE 9 – La qualité de la présentation et contributions des modalités

Comme dans la partie de l'ACP, ce tableau détaille d'abord la contribution (colonne "Qualité") (en proportion de variance) des modalités dans le premier

plan factoriel ; il donne ensuite l'inertie relative (contribution en proportion de variance) par rapport à chaque axe (colonne "Inertie") ; enfin il donne la proportion de variance de chaque modalité qui est expliquée par l'axe, noté "Cosinus2".

Les modalités qui ont les meilleures qualités de représentation selon chacun des deux axes factoriels (colonnes Cosinus2) ou dans le premier plan factoriel (colonne "Qualité") vont être retenues, On voit le tableau précédent et on a décidé à retenir les modalités jusqu'à un seuil 0,25.

Donc, les modalités qui ont les meilleurs qualités de représentation sont : les assurés de 25 à 29 ans, les retraités, les célibataires, les mariés et les veuves.

Enfin, l'inertie relative par rapport à chaque axe permettra de retenir les modalités qui ont le plus fortement contribué à la formation de cet axe. On pourra retenir les modalités dont l'inertie relative par rapport à un axe dépasse 0,04.

En conclusion, l'axe 1 nous permet de retenir les modalités suivantes : les assurés de 20 ans à 34 ans, les assurés de 70 à 75 ans, les concubins, les célibataires, les mariés et les veuves alors que l'axe 2 nous permet de retenir les modalités suivantes : les assurés plus de 70 ans, les retraités, les célibataires et les veuves.

▷ **Résultats graphiques**

La Figure 15 représente les modalités sur le premier plan factoriel.

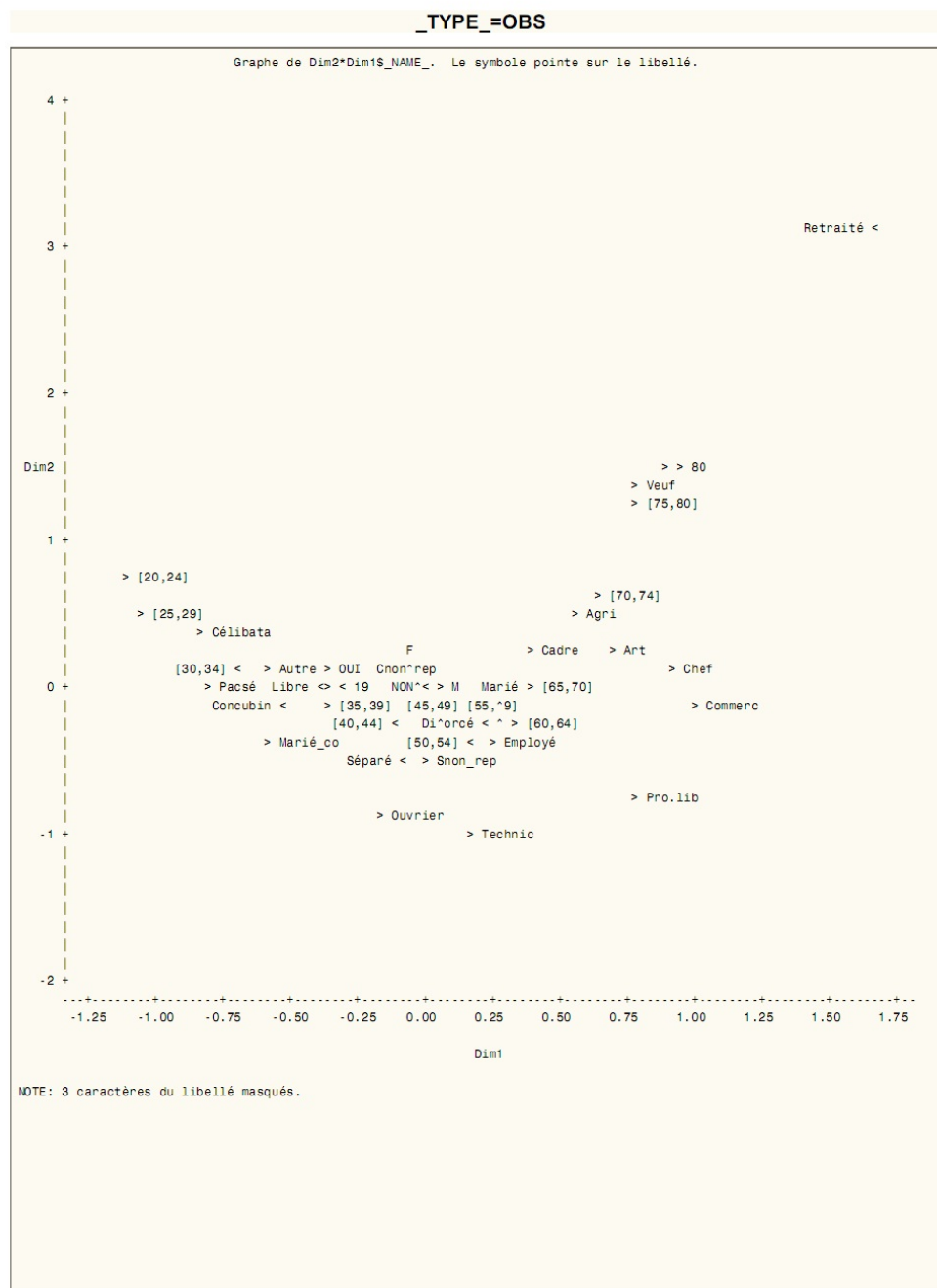


FIGURE 15 – Le résultat de l'ACM

On doit rappeler le but de l'ACM, à savoir qu'on cherche d'une part à interpréter les oppositions entre modalités, et d'autre part à interpréter les proximités entre modalités.

Ainsi, dans nos données, on note de fortes oppositions entre d'un côté : "plus 75 ans", "veuf" et de l'autre côté : "marié avec contrat de mariage", "de 35 à 44 ans". Cette opposition résume assez bien le premier axe factoriel.

Le deuxième axe marque une opposition entre "profession libérale", "technicien", "ouvrier", "de 50 à 54 ans" d'une part et d'autre part "de 20 à 29 ans", "célibataire", "autre type de situation professionnelle".

De plus l'interprétation des proximités entre les modalités devra tenir compte de la remarque suivante :

- Si deux modalités d'une même variable sont proches, cela signifie que les individus qui possèdent l'une des modalités et ceux qui possèdent l'autre sont globalement similaires du point de vue des autres variables ;
- Si deux modalités de deux variables différentes sont proches, cela peut signifier que ce sont globalement les mêmes individus qui possèdent l'une et l'autre.

Donc, les modalités "profession libérale", "technicien", "ouvrier" de la variable " situation professionnelle" sont très proches l'une de l'autre. Il n'y a pas de différence du point de vue des autres variables : c'est la même pour les modalités "célibataire", "concubin", et "pacsé" de la variable "état civil".

Ensuite, nous étudions la relation entre les modalités de l'âge du conducteur et celles de coefficient bonus-malus. Le résultat est présenté en Figure 16.

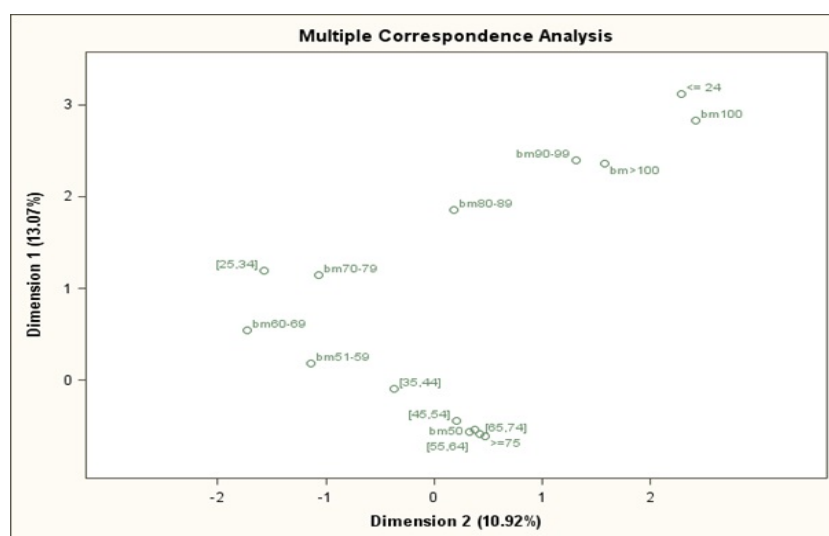


FIGURE 16 – Le résultat de l'ACM pour les tranches d'âge et coefficient bonus-malus

On voit que les gens à partir de 45 ans ont un coefficient bonus-malus de 50% ; les gens de 25 à 44 ans ont des coefficients bonus-malus plus élevés (de 51% à 79%) et les gens de moins de 25 ans ont les coefficients bonus-malus les plus élevés (plus de 80%). Cela se comprend du fait du manque d'expérience de conduite des jeunes. Ainsi l'âge du conducteur apparaît négativement corrélé au coefficient bonus-malus.

A présent, afin d'avoir une bonne vision de la relation entre les modalités de l'âge du conducteur et le type de sinistre, réalisons une analyse entre ces 2 variables présentées dans la Figure 17.

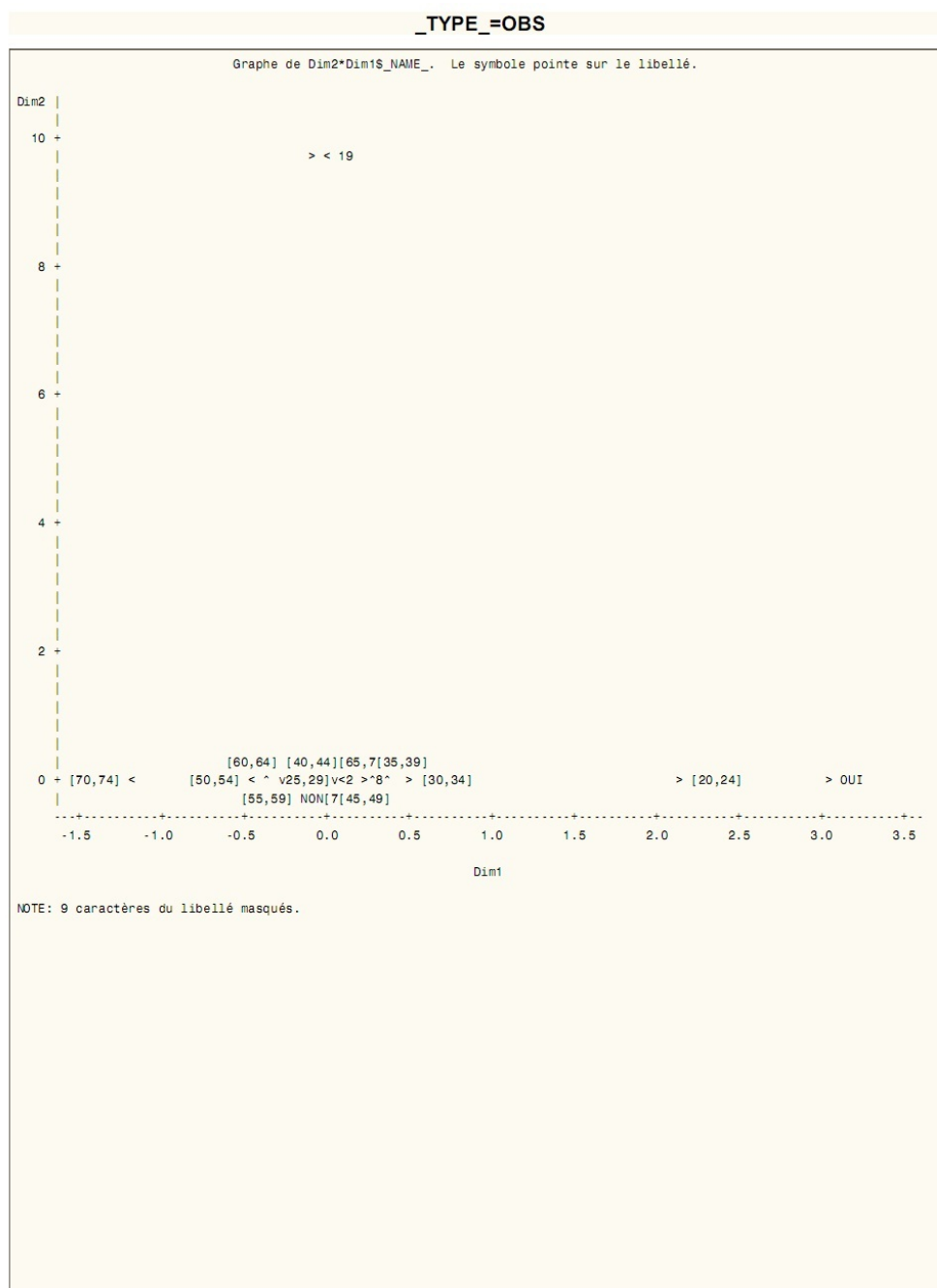


FIGURE 17 – Le résultat de l’ACM pour les tranches d’âge et le type de sinistres

On voit que les sinistres corporels (modalité OUI) touchent plus fréquemment la tranches d'âge 20-24 ans que les autres tranches d'âge. Ceci, une fois encore, explique pourquoi le coût de sinistres de cette tranche d'âge est souvent plus élevé par rapport aux autres tranches d'âge (on a fait cette analyse lors de l'étude de la fréquence, et du coût des sinistres et l'analyse du ratio S/C). Sur les données 2011, on voit que les sinistres corporels touchent moins fréquemment les personnes de 65 à 70 ans. Pour avoir une analyse plus pertinente, on a aussi effectué l'ACM sur les données des années précédentes 2008, 2009, 2010, dont les graphiques seront donnés en annexe. Quelques résultats intéressants de dégagent de ces analyses :

- Pendant les 3 années de 2008 à 2010, les sinistres corporels chez les personnes de 20 à 29 ans sont toujours plus fréquents que chez les autres personnes. On voit bien aussi que les conducteurs d'âge moyen ont des sinistres corporels plus fréquemment que les conducteurs âgés. Dans la classe des personnes au delà de 65 ans, on remarque que les conducteurs de 65 à 70 ans ont plus de chance avoir un sinistre corporel que les personnes âgées au delà de 70 ans.

- Pour les tranches d'âge 35-39 ans, 50-59 ans les conducteurs n'ont pas eu de sinistre corporel pendant les années de 2008 à 2011.

On doit noter que cette analyse répond à la question pour quelles tranches d'âge les sinistres corporels sont-ils les plus fréquents ? Par la suite, nous nous intéressons à la gravité des sinistres corporels. C'est un facteur qui nous permettra d'affiner la suite de notre étude.

Nous donnons ci-dessous la Table 10 du nombre et de la gravité de sinistres corporels pendant les 4 années successives de 2008 à 2011.

		Nb sinistres	Cout sinistres	Cout moyen	Ecart type	Cout min	Cout max
2008	Corporel (%)	147 3,30 %	1 581 583,29 28,64 %	10 759,07	8 387,24	-643,00	80 000
	Non_corporel (%)	4 310 96,70 %	3 941 107,10 71,36 %	914,41	1 281,88	-2 243,89	22 028
2009	Corporel (%)	181 3,26 %	4 920 772,79 49,56 %	27 186,59	155 072,02	0,00	2 000 000
	Non_corporel (%)	5 372 96,74 %	5 008 201,88 50,44 %	932,29	1 409,37	-1 236,95	50 000
2010	Corporel (%)	205 3,14 %	2 396 320,85 28,68 %	11 689,37	15 760,49	-7 112,99	180 000
	Non_corporel (%)	6 330 96,86 %	5 958 175,80 71,32 %	941,26	1 486,75	-4 415,44	29 120
2011	Corporel (%)	209 2,99 %	3 180 702,03 33,05 %	15 218,67	43 145,18	-117,56	600 000
	Non_corporel (%)	6 788 97,01 %	6 442 422,92 66,95 %	949,09	1 585,09	-2 003,28	68 940

TABLE 10 – Nombre et gravité de sinistres corporels

Sur la Table 10, on voit que le nombre de sinistres corporels parmi l'ensemble de sinistres est très faible, environ de 3% alors que la proportion du coût de sinistres corporels est très importante, environ de 30% et allant même jusqu'à 50 %. En 2009, le coût de sinistres corporels a augmenté jusqu'à 49,56% dû à un sinistre corporel d'un montant 2 000 000 €. On remarque aussi que le coût moyen de sinistres corporels est plus élevé que celui des sinistres non-corporels, environ 10 fois plus élevé. En 2011, il y a eu un sinistre corporel coûtant 600 000 €, et plus remarquable, ce sinistre a touché une personne âgée. Les données de Swiss Life montrent que les sinistres corporels sont souvent très coûteux chez les personnes âgées. On voit bien qu'il y a une tendance décroissante légère du nombre de sinistres corporels, alors que les montants de sinistres corporels augmentent chaque année. Cela est peut-être dû soit à l'accroissement du coût de la vie, soit au fait que les sinistres corporels sont de plus en plus graves. Pour mieux voir la variation du montant de sinistres selon chaque type de sinistres, nous utilisons 2 "boxplot" donnés en Figure 18. Pour le premier "boxplot", on garde toutes les données et pour le deuxième, on enlève les valeurs extrêmes pour voir mieux la médiane, le premier quartile et le troisième quartile.

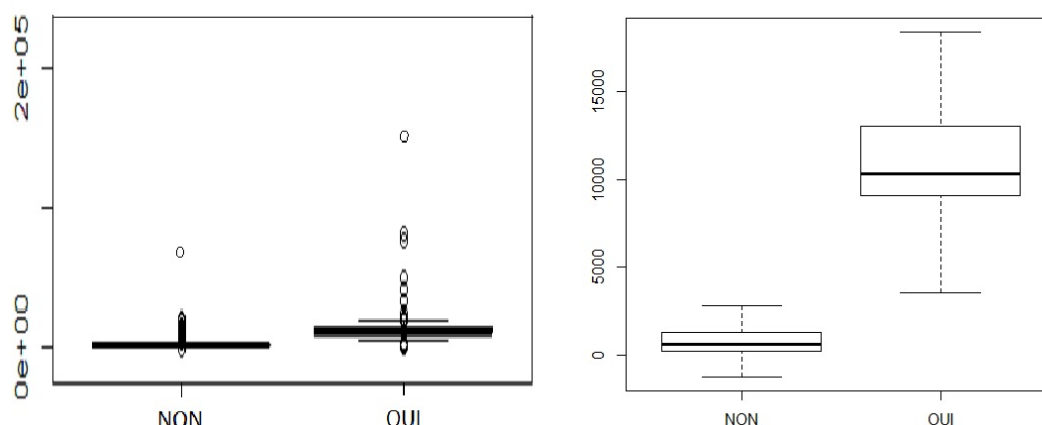


FIGURE 18 – Le Boxplot des montants de sinistres corporels et non-corporels

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
NON	-2003	231	586	949	1 267	68 940
OUI	-118	9 064	10 300	15 218	13 036	600 000

TABLE 11 – Le résumé des données de sinistres corporels et non-corporels

Ce boxplot et la Table 11 indiquent que :

- 50% des données se trouvent entre les limites inférieure et supérieure du "boxplot", à savoir, entre les premier et troisième quartiles situés à 9064 et 13036 pour les sinistres corporels, respectivement ; à 231 et 949 respectivement pour les sinistres non-corporels.
- 50% des données se trouvent en dessous de la médiane, les montants de sinistres sont inférieurs à 10300 et 586 pour les sinistres corporels et non-corporels, respectivement. De même, 50% des données se trouvent au-dessus de la ligne médiane dont les montants de sinistres sont supérieures à 10300 et 586 pour les sinistres corporels et non-corporels, respectivement.
- 25% des données se trouvent entre -118 et 9064 pour les sinistres corporels et entre -2003 et 231 pour les sinistres non-corporels. De même, 25% des données se trouvent au-dessus de 13036 et 1267 pour les sinistres corporels et non-corporels, respectivement.

Le boxplot montre une forte variation du montant de sinistres corporels par rapport au montant de sinistres non-corporels. De plus, il y a beaucoup de valeurs "outliers" parmi les montants de sinistres corporels. Ces valeurs affectent beaucoup le résultat de l'étude. On étudiera ces valeurs de façon plus approfondie dans la suite.

On présente ensuite la Table 12 de la fréquence et de la gravité des sinistres corporels chez les personnes de plus de 65 ans pendant les 4 années de 2008 à 2011.

		nb de sinistre	% nb de sinistre	% montant de sinistre	Cout moyen	Ecart type
2008	Corporel	20	3,13%	32,52%	13 935,95	9 358,69
	Non_corporel	619	96,87 %	67,48 %	934,24	1 061,74
2009	Corporel	23	2,56 %	29,60 %	15 050,99	11 156,55
	Non_corporel	875	97,44 %	70,40 %	941,11	1 044,86
2010	Corporel	42	3,66 %	35,69 %	13 976,61	16 227,35
	Non_corporel	1 104	96,34 %	64,31 %	957,93	1 309,80
2011	Corporel	37	2,66 %	49,65 %	33 012,09	97 881,80
	Non_corporel	1 352	97,34 %	50,35 %	916,03	1 222,76

TABLE 12 – Fréquence et montant de sinistres corporels chez les seniors

En considérant la population de personnes âgées de plus de 65 ans, on remarque que la proportion de sinistres corporels sur l'ensemble des sinistres est

presque identique à celle évaluée sur la population entière (tous âges confondus), alors que le coût moyen de sinistres corporels chez les personnes âgées est plus élevé que pour les autres âges sauf en 2009 à cause d'un sinistre corporel de 2 000 000 € pour un conducteur de 22 ans, mais ce constat reste le même dès que l'on enlève cette valeur extrême.

Donnons ensuite le "boxplot" qui décrit la variation du montant de sinistres corporels (Figure 19).

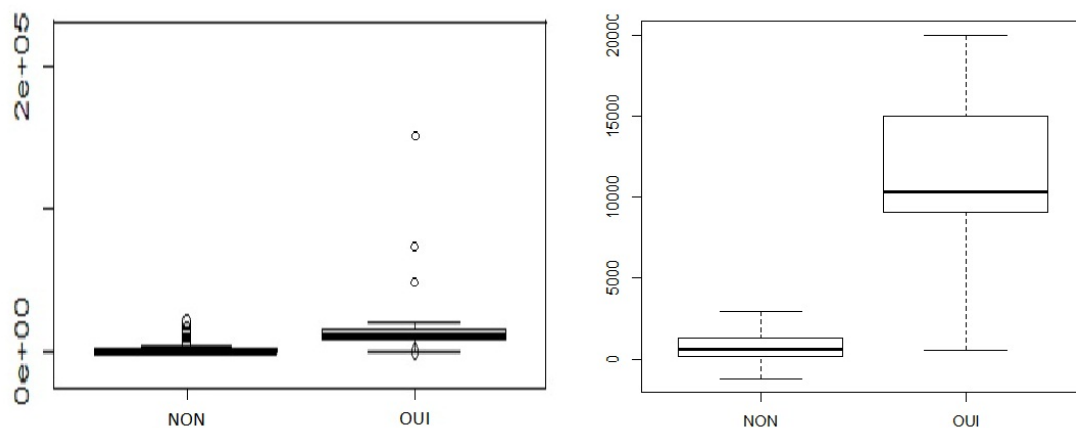


FIGURE 19 – Le boxplot des montants de sinistres corporels et non-corporels chez les seniors

	Min	1st Quantile	Median	Mean	3rd Quantile	Max
NON	-1 189	209	594	916	1 296	19 999
OUI	43	9 064	10 300	33 012	15 000	600 000

TABLE 13 – Le résumé des données de sinistres corporels et non-corporels chez les seniors

Le "boxplot" et la Table 13 indiquent que :

- 50% des données se trouvent entre les limites inférieure et supérieure du "boxplot", à savoir, entre les premier et troisième quartiles situés à 9064 et 15000 pour les sinistres corporels, respectivement ; à 209 et 1296 respectivement pour les sinistres non-corporels.

- 50% des données se trouvent en dessous de la médiane, les montants de sinistres sont inférieurs à 10300 et 594 pour les sinistres corporels et non-corporels, respectivement. De même, 50% des données se trouvent au-dessus de la ligne médiane dont les montants de sinistres sont supérieures à 10300 et 594 pour les sinistres corporels et non-corporels, respectivement.
- 25% des données se trouvent entre 43 et 9064 pour les sinistres corporels et entre -1189 et 209 pour les sinistres non-corporels. De même, 25% des données se trouvent au-dessus de 15000 et 1296 pour les sinistres corporels et non-corporels, respectivement.

On voit que pour les personnes âgées, la variation du montant de sinistres corporels est assez forte en 2011. C'est dû aux valeurs dites "outliers". Il y a aussi beaucoup de valeurs "outliers" pour les sinistres non-corporels. De plus le montant de sinistres corporels est plus élevé que le montant de sinistres non-corporels.

Ensuite, nous étudions la fréquence de sinistre corporel et non corporel parmi l'ensemble de sinistres pour les conducteurs seniors (plus de 65 ans) et les conducteurs non seniors (moins de 65 ans). Nous assumons également que le sinistre corporel suit la distribution Binomial $B(n,p)$ où n étant le nombre de sinistres observé et p étant la fréquence de sinistre corporel. Nous donnons ci-dessous la Table 14 de la fréquence et de la distribution de sinistre corporel et non corporel pour les deux types de références de la population :

		Nombre de sinistre		Fréquence de sinistre		Distribution	
		Seniors	Non seniors	Seniors	Non seniors	Seniors	Non seniors
2008	Corporel	20	127	3,13%	3,33%	B(639 ; 3,13%)	B(3 818 ; 3,33%)
	Non _corporel	619	3 691	96,87%	96,67%	B(639 ; 96,87%)	B(3 818 ; 96,67%)
2009	Corporel	23	158	2,56 %	3,39%	B(898 ; 2,56%)	B(4 655 ; 3,39%)
	Non _corporel	875	4 497	97,44%	96,61%	B(898 ; 97,44%)	B(4 655 ; 96,61%)
2010	Corporel	42	163	3,66%	3,02%	B(1 146 ; 3,66%)	B(5 389 ; 3,02%)
	Non _corporel	1 104	5 226	96,34%	96,98%	B(1 146 ; 3,66%)	B(5 389 ; 3,02%)
2011	Corporel	37	172	2,66%	3,07%	B(1 389 ; 2,66%)	B(5 608 ; 3,07%)
	Non _corporel	1 352	5 436	97,34%	96,93%	B(1 389 ; 97,34%)	B(5 608 ; 96,93%)

TABLE 14 – Fréquence et distribution de sinistre corporel et non-corporel chez les seniors et non seniors

Nous pouvons calculer les intervalles de confiance pour les fréquences de deux

types de sinistre selon la formule suivante :

$$\hat{p} \in \left[p - z_{1-\frac{1}{2}\alpha} \sqrt{\frac{1}{n}p(1-p)}, p + z_{1-\frac{1}{2}\alpha} \sqrt{\frac{1}{n}p(1-p)} \right]$$

où $z_{1-\frac{1}{2}\alpha}$ est le quantile d'ordre α de la loi Normale.

L'intervalle de confiance à 95% de la fréquence de sinistre corporel et non corporel est donnée dans la Table 15 suivante :

		Intervalle de confiance	
		Seniors	Non seniors
2008	Corporel	[1,78% ; 4,48%]	[2,76% ; 3,90%]
	Non_corporel	[95,52% ; 98,22%]	[96,10% ; 97,24%]
2009	Corporel	[1,53% ; 3,59%]	[2,87% ; 3,91%]
	Non_corporel	[96,41% ; 98,47%]	[96,09% ; 97,13%]
2010	Corporel	[2,57% ; 4,75%]	[2,57% ; 3,48%]
	Non_corporel	[95,25% ; 97,43%]	[96,52% ; 97,43%]
2011	Corporel	[1,81% ; 3,51%]	[2,62% ; 3,52%]
	Non_corporel	[96,49% ; 98,19%]	[96,48% ; 97,38%]

TABLE 15 – Intervalle de confiance de la fréquence de sinistre corporel et non-corporel chez les seniors et non seniors

Nous voyons que l'intervalle de confiance de la fréquence de sinistre corporel chez les seniors est souvent plus large que celle chez les non seniors pendant les 4 années de 2008 à 2010.

En conclusion, les sinistres corporels sont très importants parmi l'ensemble de sinistres. De plus, chaque année, il y a des sinistres corporels très coûteux, ils perturbent notre résultat technique. Pour avoir une étude plus pertinente, il mérite d'enlever ces sinistres extrêmes dans la partie suivante.

Troisième partie

La segmentation et la tarification

Pour pouvoir déterminer les coefficients tarifaires, il semble nécessaire d'analyser et de mesurer la performance et le comportement de l'assuré en fonction de certaines caractéristiques notamment l'âge du 1^{er} conducteur et la région qu'il habite (appelée en assurance région commerciale). Il faut noter que dans cette partie notre objectif est toujours l'étude de la population des seniors.

7 Etude préliminaire sur les variables croisées

Nous avons étudié la fréquence et le coût moyen des sinistres chez les personnes âgées, maintenant nous nous intéressons à savoir si ces facteurs varient en fonction de la région commerciale au sein d'une même tranche d'âge. Dans cette section, nous considérons l'ensemble des garanties sans la garantie "bris de glace". En effet, la proportion du nombre de sinistres pour la garantie bris de glace est en moyenne de 40 % de l'ensemble du nombre de sinistres chaque année, alors que le montant de sinistres est seulement environ de 9 % du coût total de sinistres chaque année. Ce résultat montre que les sinistres de type bris de glace sont très fréquents, alors que le coût total de ces sinistres est faible. Nous décidons donc de ne pas tenir compte de ce type de sinistre bris de glace pour avoir une analyse plus pertinente.

7.1 Répartition du portefeuille, fréquence et coût moyen des sinistres, par âge et par région commerciale

La fréquence et le coût moyen des sinistres sont calculés selon la méthode présentée dans la deuxième partie, la répartition du portefeuille ou le poids de chaque tranche d'âge est calculé dans chaque région commerciale.

Notre étude se concentre sur les retraités chez Swiss Life, répartis en 4 tranches d'âge suivantes : [65,69], [70,74], [75,79], [80 ans et plus]. La Figure 20 présente des histogrammes décrivant, pour les variables de l'âge et de la région commerciale, la répartition du portefeuille entre les différents niveaux de la variable et, pour chacun de ces niveaux, la fréquence et le coût moyen (en %) de sinistres.

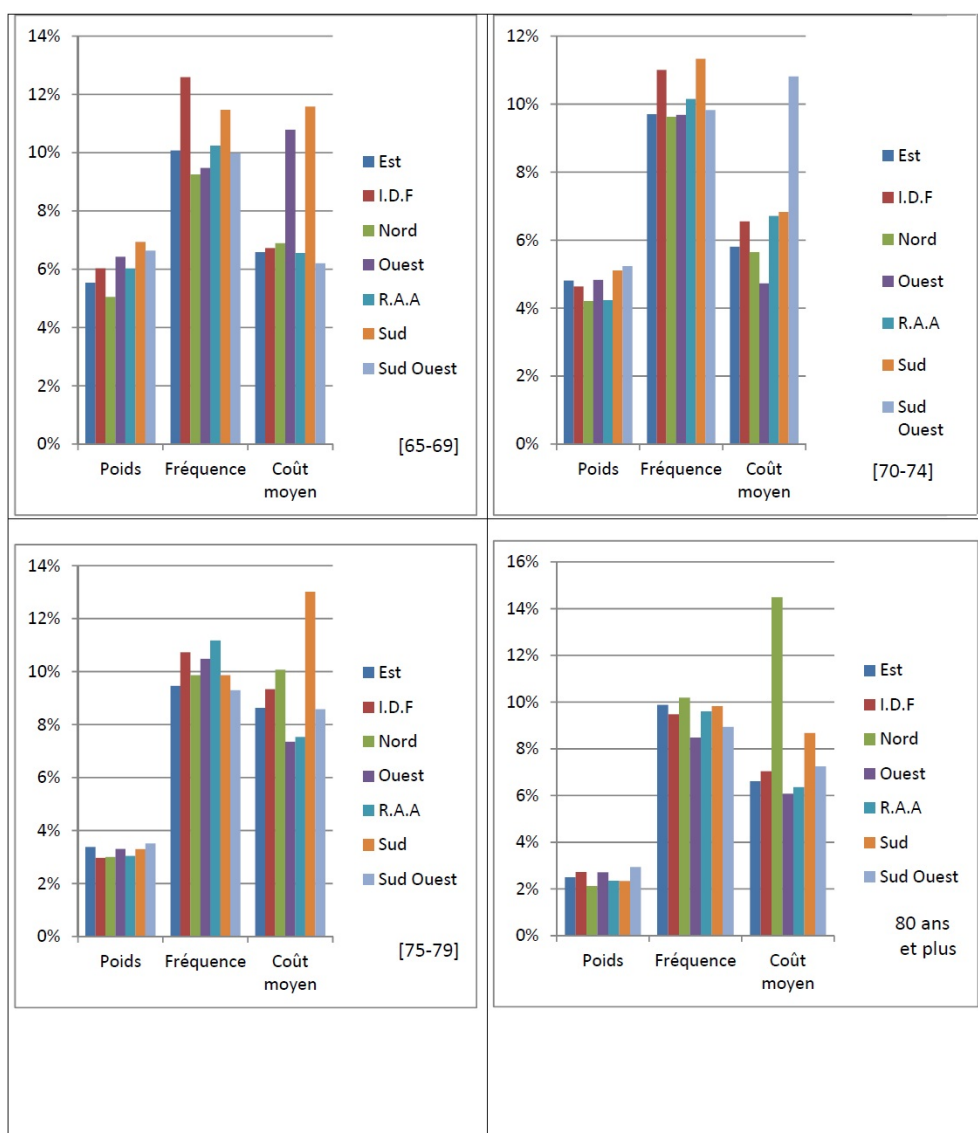


FIGURE 20 – Histogrammes des variables poids, fréquence, coût moyen selon les tranches d'âge et les régions commerciales

Nous remarquons qu'il y a des hétérogénéités entre les régions commerciales à l'intérieur d'une même tranche d'âge. Commentons brièvement les histogrammes obtenus selon les tranches d'âge :

▷ **Tranche [65,69]** : les trois régions Ile de France, Sud et Ouest sont plus risquées. La fréquence des sinistres dans la région Ile de France est la plus élevée, alors que le coût moyen des sinistres dans cette région n'est pas trop élevé. Cela peut

se comprendre du fait de la densité importante du trafic, mais les conséquences des sinistres devraient être faibles du fait de la vitesse réduite. Au contraire, dans l'Ouest, la fréquence des sinistres sont moins élevés qu'en Ile de France, du fait de la plus faible densité du trafic, alors que le coût moyen des sinistres est le plus grave, du fait de la vitesse plus élevée des véhicules. Dans la région de Sud, la fréquence et le coût moyen sont très élevés, c'est-à-dire, les personnes âgées dans cette région sont très risquées par rapport aux autres. Pour les autres régions, la fréquence et le coût moyen des sinistres sont presque identiques, il n'y a pas de grosses différences. Nous voyons aussi que c'est dans le Sud que la densité de population de 65 à 69 ans est la plus forte.

▷ **Tranche [70,74]** : les trois régions Ile de France, Sud et Sud-Ouest sont plus risquées. Dans la région Ile de France, la fréquence des sinistres est très élevés, alors que le coût moyen est modéré. Au contraire, dans le Sud, les sinistres sont les plus fréquents et le coût moyen est assez grave. La région la plus évoluée est le Sud-Ouest. La densité des seniors de 70 à 74 ans est la plus forte dans cette région, les sinistres ne sont pas très fréquents alors que le coût des sinistres est le plus élevé. C'est-à-dire, les sinistres des personnes âgées dans cette région sont très coûteux.

▷ **Tranche [75,79]** : dans la région Rhône Alpes, la fréquence des sinistres est la plus élevée, alors que le coût moyen des sinistres n'est pas élevé. Cela signifie que les sinistres ne coûtent pas très chers. Au contraire, dans le Sud, les sinistres ne sont pas très fréquents, alors que le coût moyen des sinistres est le plus élevé. C'est-à-dire les sinistres des seniors dans cette région sont très importants. Dans la région du Nord, les sinistres ne sont pas très fréquents pour la tranche [75,79], alors que le coût moyen des sinistres est assez important. La fréquence des sinistres dans la région Ile de France est toujours élevée, alors que le coût moyen n'est pas élevé. La densité de population de 75 à 79 ans est la plus importante dans la région Sud-Ouest.

▷ **Tranche [80 ans et plus]** : il y a deux types de comportement l'un regroupant les 2 régions Nord et Sud et l'autre les 5 autres régions. Le premier groupe apparaît plus risqué au sens de la fréquence et du coût moyen des sinistres. Pour le deuxième groupe, les comportements sont presque identiques. La densité des gens de plus de 80 ans est la plus forte dans la région Sud-Ouest.

Nous donnons ci-dessous (Table 16) les effectifs de la population d'une tranche par une région commerciale :

	Est	I.D.F	Nord	Ouest	R.A.A	Sud	Sud Ouest
[65,69]	5 239	4 187	9 909	3 715	5 850	10 847	7 906
[70,74]	4 555	3 218	8 249	2 789	4 107	7 983	6 238
[75,79]	3 195	2 060	5 889	1 908	2 952	5 159	4 186
[80 ans et plus]	2 369	1 888	4 180	1 569	2 280	3 665	3 503

TABLE 16 – Les effectifs de la population d’une tranche d’âge par une région commerciale

7.2 La prime pure

Dans la section précédente, nous avons analysé séparément la fréquence et le coût moyen des sinistres. Dans cette section, nous allons calculer la prime pure à partir de la fréquence et le coût moyen observés des sinistres sur l’ensemble des données de Swiss Life.

Le principe général de calcul de la prime pure correspond à l’espérance mathématique du coût annuel des sinistres déclarés à l’assureur. Cette prime pure se décompose toutefois en deux termes : la fréquence et le coût moyen.

En effet, soit X_i la variable aléatoire représentant le montant du $i^{\text{ème}}$ sinistre, N la variable aléatoire représentant le nombre de sinistres. La charge total de sinistres est donc :

$$S_N = \sum_{i=1}^N X_i$$

On suppose de plus que :

- les X_i sont indépendantes et de même loi,
- $(X_i)_{i \geq 1}$ est indépendante de N (indépendance coûts-fréquence).

Du fait de l’indépendance, on peut calculer facilement par conditionnement en N la prime pure selon la formule suivante :

$$E[S_N] = E\left[\sum_{i=1}^N X_i\right] = E[N] \times E[X_i]$$

La Figure 21 montre la prime pure pour chaque niveau de la variable :

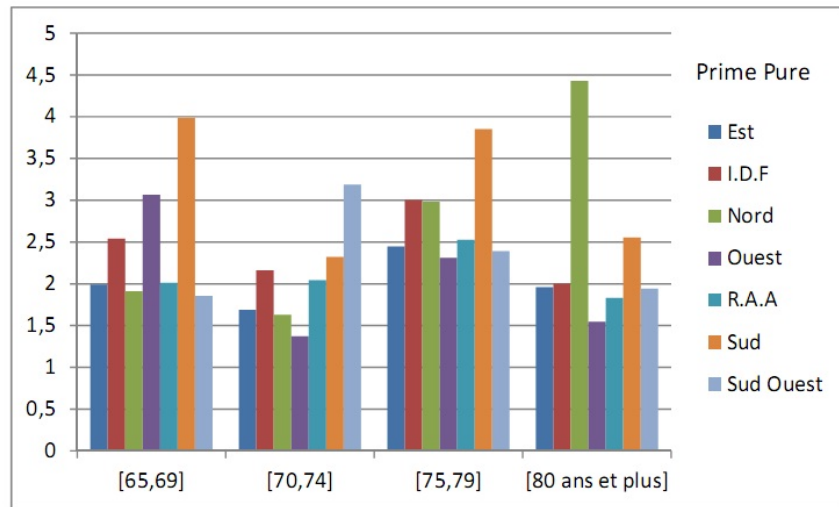


FIGURE 21 – La prime pure

Nous voyons que la prime pure observée sur l'ensemble des données de 2008 à 2011 varie selon l'âge du conducteur et de la région qu'il habite. Les seniors dans le Sud sont "plus risqués" (au sens où la prime pure est la plus élevée) que dans les autres régions. Les résultats sont presque identiques à ceux de la partie de l'analyse de la fréquence et du coût moyen de sinistres. Nous avons aussi effectué une ACM pour mieux connaître la relation entre l'âge du conducteur, sa région et la fréquence des sinistres (Figure 22).

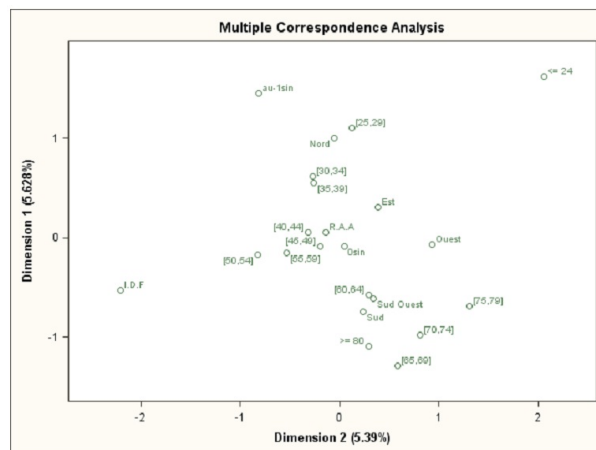


FIGURE 22 – Le résultat de l'ACM

Nous constatons que les personnes âgées ont une tendance à habiter dans le

Sud ou le Sud-Ouest. Ils ont moins de sinistres que les plus jeunes habitant dans le Nord.

7.3 L'analyse du ratio S/C pour chaque type de garantie

Dans la section précédente, en considérant le total des garanties, nous avons enlevé la garantie de type bris de glace. Dans cette section nous étudions le ratio S/C pour chaque type de garantie, nous gardons donc la garantie de bris de glace.

Notre étude a porté sur les garanties suivantes :

- Responsabilité civile (RC) et tous accidents : cette garantie comporte la garantie de responsabilité civile et la garantie de dommage. En fait, il n'y a pas de distinction des montants de sinistres de type RC et Tous-Accidents dans les données de Swiss Life. La garantie de responsabilité civile est une obligation légale. Elle couvre les dommages matériels ou corporels qu'un assuré est susceptible de causer à autrui. Cette garantie peut être divisée en responsabilité civile matérielle et responsabilité civile corporelle. La garantie de dommage couvre les dégâts que le véhicule de l'assuré pourrait subir.
- Bris de glace : cette garantie entre en jeu si des éléments vitrés du véhicule sont endommagés.
- Incendie-vol : cette garantie s'applique comme son nom l'indique en cas d'incendie ou de vol du véhicule.

Le ratio S/C est important dans notre étude, car il permet de définir les garanties rentables pour Swiss Life. Ce même ratio en base 100 permet aussi de mettre en évidence d'éventuelles sur-tarifications ou sous-tarifications selon les régions commerciales et l'âge du 1^{er} conducteur, qui pourraient entraîner des phénomènes d'antisélection. Une sous-tarification dans les régions commerciales et dans les tranches d'âge les plus risquées pourrait engendrer une augmentation du nombre de "mauvais" risques. De la même manière, une sur-tarification dans des régions commerciales et dans des tranches d'âge peu risquées entraînerait une diminution du nombre de "bons" risques. Pour éviter des phénomènes d'antisélection, nous avons besoin d'introduire des coefficients tarifaires qui dépendent des tranches d'âge et des régions commerciales dans les formules tarifaires de Swiss Life.

Calculons le ratio S/P des assurés de plus de 65 ans selon 4 tranches d'âge : [65,69], [70,74], [75,79], [80 ans et plus] dans la Table 17.

Age du 1 ^{er} conducteur	Zone Commerciale	RC-Tous Accidents	Vol-Incendie	Bris de Glace
[65,70]	Est	96,86	30,42	92,16
[65,70]	I.D.F	77,97	58,52	99,81
[65,70]	Nord	84,51	42,13	104,84
[65,70]	Ouest	162,26	36,95	89,04
[65,70]	R.A.A	83,96	32,93	77,47
[65,70]	Sud	155,57	37,97	67,31
[65,70]	Sud Ouest	85,34	46,01	76,98
Total		107,47	41,50	85,53
[70,74]	Est	82,53	33,30	73,15
[70,74]	I.D.F	76,08	24,75	79,61
[70,74]	Nord	73,89	27,60	99,70
[70,74]	Ouest	76,28	15,41	68,41
[70,74]	R.A.A	85,22	49,64	73,97
[70,74]	Sud	98,78	30,07	64,33
[70,74]	Sud Ouest	156,97	38,82	68,46
Total		95,04	31,81	76,89
[75,80]	Est	97,93	42,17	58,86
[75,80]	I.D.F	108,72	10,61	76,36
[75,80]	Nord	142,42	23,20	78,01
[75,80]	Ouest	121,11	27,83	54,26
[75,80]	R.A.A	112,41	10,38	57,76
[75,80]	Sud	170,22	23,76	56,83
[75,80]	Sud Ouest	122,86	47,09	61,51
Total		131,40	25,46	64,39
> 80 ans	Est	101,64	0,11	42,93
> 80 ans	I.D.F	61,86	19,40	44,13
> 80 ans	Nord	201,89	30,56	54,44
> 80 ans	Ouest	80,07	20,90	36,66
> 80 ans	R.A.A	79,92	25,79	49,87
> 80 ans	Sud	105,20	26,59	40,61
> 80 ans	Sud Ouest	97,33	29,21	47,21
Total		113,65	23,55	46,31

TABLE 17 – Le ratio S/P des assurés pour chaque type de garantie

L'analyse de ces résultats donnés dans la Table 17 permet de dire que la garantie RC-Tous Accidents est largement déficitaire pour Swiss Life pour presque toutes les tranches d'âge âgées et pour toutes les régions commerciales. En revanche, Swiss Life est bénéficiaire sur toutes les autres garanties, et notamment sur la garantie incendie-vol et sur la garantie bris de glace.

8 La tarification sur une segmentation en classes homogènes de l'âge chez les seniors et du zonage selon la méthode basée sur S/C

8.1 La méthode actuelle et sa limite

Avant cette étude, les coefficients tarifaires dépendaient entre autres, de la région commerciale, mais pas de la classe d'âge (si non de façon indirecte par le système de bonus, malus, les jeunes conducteurs étant plus touchés par ce système). Il s'agit donc d'introduire dans cette étude la classification par âge afin d'étudier la classe spécifique des seniors. Nous utilisons la méthode basée sur le calcul du ratio S/C (sinistres/coûts) comme suivant :

Soit $S_{i,j}$, $C_{i,j}$ représentant respectivement le montant total de sinistres et la prime totale de la tranche d'âge i dans la région commerciale j . Le ratio S/C est donc pour la tranche d'âge i et la région commerciale j égale à $\frac{S_{i,j}}{C_{i,j}}$. Nous avons donc calculer le charge sinistre, la prime et le ratio S/C de chaque région commerciale comme suivants :

$$S_j = \sum_i S_{i,j}$$

$$C_j = \sum_i C_{i,j}$$

$$\frac{S}{C} = \frac{S_j}{C_j}$$

Notre but est d'introduire des nouveaux coefficients tarifaires (de l'âge et de la région commerciale) tel que après d'appliquer ces coefficients dans chaque région commerciale le ratio S/C pour chaque tranche d'âge est toujours égale au ratio S/C de cette région, autrement dit si on note $K_{i,j}$ étant le coefficient tarifaire pour la tranche d'âge i et la région commerciale j :

$$K_{i,j} = \frac{S_j}{C_j} \times \left(\frac{S_{i,j}}{C_{i,j}}\right)^{-1} = \frac{C_{i,j} S_j}{S_{i,j} C_j}$$

Cependant cette méthode possède des limites qui viennent de la nature des données :

- Nous avons considéré des données d’une période de 2008 à 2011, les coûts de sinistres sont les valeurs estimées parce que certains sinistres sont non clôturés.
- La méthode basée sur S/C introduit de nouveaux tarifs pour des clients de Swiss Life. Notons que dans l’assurance des primes sont traditionnellement déterminées selon la théorie de l’espérance mathématique du montant des sinistres. La méthode actuelle de Swiss Life peut donc entraîner des biais dans l’étude.

8.2 Grille tarifaire pour les garanties RC - Tous Accidents, Vol - Incendie, Bris de Glace

Les résultats dans la section précédente nous conduisent à considérer les deux tranches d’âge 65 à 74 ans et plus de 75 ans. Nous introduisons ci-dessous (Table 18) les coefficients tarifaires calculés selon le modèle de tarification de Swiss Life (basé sur le ratio S/C).

Age du 1 ^{er} conducteur	Région Commerciale	RC-Tous Accidents	Vol-Incendie	Bris de Glace
[65,74]	Est	0,92	0,63	0,98
[65,74]	I.D.F	0,92	0,86	1,07
[65,74]	Nord	0,78	0,68	1,11
[65,74]	Ouest	1,25	0,68	1,00
[65,74]	R.A.A	0,80	0,79	1,08
[65,74]	Sud	1,13	0,59	1,10
[65,74]	Sud Ouest	1,07	0,68	1,05
≥ 75 ans	Est	1,02	0,49	0,61
≥ 75 ans	I.D.F	1,01	0,29	0,71
≥ 75 ans	Nord	1,63	0,50	0,74
≥ 75 ans	Ouest	1,01	0,59	0,58
≥ 75 ans	R.A.A	0,93	0,34	0,78
≥ 75 ans	Sud	1,21	0,43	0,83
≥ 75 ans	Sud Ouest	1,02	0,62	0,79

TABLE 18 – Les coefficients tarifaires chez les seniors, par classe d’âge et région commerciale, pour chaque type de garanties

8.3 Analyse des résultats

Afin d’interpréter les résultats obtenus, nous complétons l’étude en considérant les taux des évolutions des primes avant et après application des coefficients tarifaires. Le premier tableau (Table 19) représente les évolutions des primes en fonction de l’âge, de la région commerciale et du type de garantie. Le second (Table 20) représente les évolutions des ratios S/C selon les régions, avant et après avoir pris en compte les tranches d’âge.

Age du 1 ^{er} conducteur	Région Commerciale	RC-Tous Accidents	Vol-Incendie	Bris de Glace
[65,74]	Est	-8%	-37%	-2%
[65,74]	I.D.F	-8%	-14%	+7%
[65,74]	Nord	-22%	-32%	+11%
[65,74]	Ouest	+25%	-32%	-
[65,74]	R.A.A	-20%	-21%	+8%
[65,74]	Sud	+13%	-41%	+10%
[65,74]	Sud Ouest	+7%	-32%	+5%
≥ 75 ans	Est	+2%	-51%	-39%
≥ 75 ans	I.D.F	+1%	-71%	-29%
≥ 75 ans	Nord	+63%	-50%	-26%
≥ 75 ans	Ouest	+1%	-41%	-42%
≥ 75 ans	R.A.A	-7%	-66%	-22%
≥ 75 ans	Sud	+21%	-57%	-17%
≥ 75 ans	Sud Ouest	+2%	-38%	-21%

TABLE 19 – Les évolution des primes pour chaque type de garanties

Nous pouvons remarquer que :

- **Pour la garantie RC-Tous Accidents** : les personnes de 65 à 74 ans habitant dans les régions commerciales Est, I.D.F, Nord, R.A.A et les personnes de plus de 75 ans habitant dans le R.A.A apparaissent comme étant sur-tarifés (ils nécessitent une réduction de tarif). Nous notons également que comme signalé précédemment, les seniors de 65 à 74 ans dans le Sud, l'Ouest et Sud-Ouest sont réellement sous-tarifés (ils nécessitent une augmentation de tarif). Nous remarquons également que les plus de 75 ans sont presque sous-tarifés pour toutes les régions commerciales (sauf dans le R.A.A), nous avons donc vraiment besoin d'une augmentation de prime.
- **Pour la garantie Vol-Incendie** : il apparait qu'il y ait une sur-tarification pour toutes les personnes de plus de 65 ans, quelle que soit leur région commerciale. Aussi nécessiteraient-ils une réduction de tarif pour cette garantie.
- **Pour la garantie Bris de Glace** : les seniors de 65 à 74 ans apparaissent comme étant sous-tarifés dans presque toutes les régions commerciales (sauf dans l'Est), alors que les seniors de plus de 75 ans apparaissent comme étant sur tarifés dans toutes les régions commerciales. Nous avons donc besoin d'un ajustement raisonnable des coefficients tarifaires.

Age du 1 ^{er} conducteur	Région Commerciale	RC-Tous Accidents	Vol-Incendie	Bris de Glace
[65,74]	Est	-8%	-19%	-2%
[65,74]	I.D.F	-7%	-8%	+6%
[65,74]	Nord	-23%	-17%	+10%
[65,74]	Ouest	+25%	-13%	-
[65,74]	R.A.A	-21%	-11%	+6%
[65,74]	Sud	+15%	-24%	+6%
[65,74]	Sud Ouest	+7%	-20%	+3%
≥ 75 ans	Est	+2%	-26%	-33%
≥ 75 ans	I.D.F	+1%	-37%	-25%
≥ 75 ans	Nord	+65%	-27%	-24%
≥ 75 ans	Ouest	+1%	-17%	-34%
≥ 75 ans	R.A.A	-7%	-33%	-16%
≥ 75 ans	Sud	+25%	-33%	-10%
≥ 75 ans	Sud Ouest	+2%	-24%	-15%

TABLE 20 – Les évolution des ratios S/P pour chaque type de garanties

Le tableau de l'évolution des ratios S/P (Table 20) illustre une fois encore la sur-tarification et la sous-tarification pour les segmentations tarifaires de l'âge du premier conducteur et de la région commerciale pour quelques types de garanties.

9 Présentation du modèle de tarification classique : modèle linéaire généralisé

9.1 Rappel théorique : présentation des modèles linéaires généralisés

Lorsque nous cherchons à expliquer linéairement une variable Y par un ensemble de variables $X_1, \dots, X_p$, le modèle linéaire classique possède l'avantage d'être facilement mis en œuvre. Son principal inconvénient est le suivant : nous supposons que la variable à expliquer suit une distribution gaussienne. Le modèle linéaire généralisé est une extension du modèle linéaire classique qui permet de modéliser des phénomènes dont la distribution n'est pas gaussienne et d'étendre le modèle gaussien à une famille de lois particulière, appelée famille exponentielle. Ainsi, nous nous tournons vers des modèles tenant mieux compte de la réalité de l'assurance que ne le fait le modèle linéaire.

9.2 Définition

Le modèle linéaire généralisé (MLG) consiste à exprimer, par un prédicteur linéaire, l'influence des variables $X_1, X_2, \dots, X_p$ sur une fonction (appelée fonction lien) de l'espérance de Y . Aussi, Y relève du modèle linéaire généralisé si la loi de Y sachant $\{X_1 = x_1, \dots, X_p = x_p\}$ est telle que :

- Il existe une fonction g monotone de $\mathbb{R}$ dans $\mathbb{R}$.
- Il existe des coefficients $(\beta_1, \dots, \beta_p)$ tels que :

$$g(E(Y)) = \beta_0 + \beta_1 \times x_1 + \dots + \beta_p \times x_p$$

- La loi de probabilité de Y doit appartenir à la famille exponentielle de la forme

$$f(y; \theta; \varphi) = \exp\left\{\frac{\theta y - b(\theta)}{a(\varphi)} + c(y, \varphi)\right\}$$

où :

a est une fonction non nulle définie sur $\mathbb{R}$

b est une fonction définie sur $\mathbb{R}$, deux fois dérivable

c est une fonction définie sur $\mathbb{R}^2$

θ paramètre canonique

φ paramètre de dispersion

Conséquences :

- L'espérance de Y s'écrit $\mu = E(Y) = b'(\theta)$
- La variance de Y s'écrit $Var(Y) = b''(\theta) * a(\varphi)$

On pose $b''(\theta) = V(\mu)$ appelée fonction variance, pour mettre en valeur le fait que la variance de Y dépend de la moyenne μ .

9.3 Choix d'une fonction lien

Dans le cadre du modèle linéaire généralisé, on introduit la notion de fonction de lien canonique. Si pour tout i de 1 à n , il existe une fonction b définie sur $\mathbb{R}$, telle que $b_i(x) = b(x)$, avec b dérivable, de dérivée b' inversible, alors il existe une fonction g telle que :

$$\theta_i = (b')^{-1}(E(Y_i)) = g(E(Y_i))$$

g est la fonction inverse de b' , elle est appelée fonction de lien canonique.

L'intérêt d'une telle fonction est de rendre le vecteur $Y = (Y_1, Y_2, \dots, Y_n)$ exhaustif pour le prédicteur linéaire $(g(E(Y_1)), \dots, g(E(Y_n))) = a_0 + a_1 X_1 + \dots + a_n X_n$.

Bien que la fonction de lien canonique fournisse des propriétés statistiques intéressantes, il est tout à fait possible de choisir d'autres fonctions liens qui répondent à d'autres critères, par exemple concernant la structure du modèle.

Le choix s'effectue entre un modèle additif, pour lequel le risque est modélisé par une somme d'effets explicatifs, et un modèle multiplicatif, pour lequel le risque est modélisé par un produit d'effets multiplicatifs.

Nous désirons obtenir un tarif multiplicatif, pour deux raisons principales :

- des raisons pratiques : cette structure est traditionnellement utilisée en Assurance Automobile.
- des raisons de positivité : les modèles additifs ont le défaut de pouvoir estimer des fréquences négatives.

Dans le cas du modèle multiplicatif, si λ_k est le risque à modéliser pour la classe tarifaire k et λ_{ij} le coefficient estimé associé à la modalité j de la variable i :

$$\lambda_k = \prod_{i=1}^{n_v} \prod_{j=1}^{n_i} \lambda_{ij}^{c_{ij}^k}$$

avec :

- n_v le nombre de variables
- n_i le nombre de modalités de la variable i

- c_{ij}^k l'indicatrice qui vaut 1 lorsque la classe tarifaire k présente la modalité j de la variable i .

Notons qu'on modélise $g(E(Y))$ de manière nécessairement additive.

Choisissons la fonction lien "log". Les effets additifs pour $\log(E(Y))$ se transforment alors, par inversion de la fonction log (= fonction exponentielle), en effets multiplicatifs pour $E(Y)$. La fonction "log" permet de répondre aux besoins d'obtenir au final un modèle multiplicatif.

9.4 Choix d'une loi de probabilité

L'application du modèle linéaire généralisé nécessite le choix d'une distribution de la variable à expliquer. Comme nous l'avons vu, la distribution de la variable que l'on modélise doit appartenir à la famille exponentielle. Seules certaines lois appartiennent à cette famille, notamment la loi normale, la loi gamma, la loi inverse gaussienne, la loi binomiale négative et la loi poisson.

La Table 21 résume les principales caractéristiques de chacune de ces lois :

Loi $E(Y) = \mu$	Densité	Paramètre θ	Dispersion φ	Fonction variance V	Lien canonique
Normal(μ, σ^2)	$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{1}{2} \frac{(y-\mu)^2}{\sigma^2}$	μ	σ^2	1	Identité
Inverse Gaussienne(μ, σ^2)	$f(y) = \frac{1}{\sigma\sqrt{2\pi y^3}} \exp -\frac{1}{2} \frac{(y-\mu)^2}{\mu\sigma^2 y}$	$\frac{-1}{2\mu}$	σ^2	μ^3	Inverse carré
Gamma(α, β) $\mu = \frac{\alpha}{\beta}$	$f(y) = \frac{1}{\Gamma(\alpha)} \beta^{(\alpha)} y^{(\alpha-1)} \exp -\beta y$	$\frac{-1}{mu}$	$\frac{1}{\alpha}$	μ^2	Inverse
Binomiale Négative(n, p) $\mu = \frac{1-p}{p}$	$P(Y = y) = \binom{ny}{ny+n-1} p^n (1-p)^{ny}$	$\log(\frac{\mu}{\mu+1})$	1	$\mu(1 + \mu)$	
Poisson(m)	$P(Y = y) = \frac{\mu^y \exp -\mu}{y!}$	$\log(m)$	1	μ	log

TABLE 21 – Lois de la famille exponentielle classiquement utilisée en modélisation linéaire généralisée

9.5 Méthodologie

Le processus de modélisation se décompose en 3 phases successives :

- ▷ La première étape consiste à faire le choix d'un modèle, en définissant :
 - la variable à expliquer Y

- la loi de probabilité de Y
- la fonction lien g
- les variables explicatives $X_1, \dots, X_p$

▷ Une fois la forme du modèle définie, nous devons estimer les coefficients de la régression (grâce à une procédure GENMOD sous SAS).

▷ La dernière étape consiste à s'assurer de la validité du modèle, ce qui dans le cas d'un modèle linéaire généralisé, revient à :

- s'assurer de la cohérence des coefficients calculés,
- examiner les écarts entre les fréquences modélisées et les fréquences observées, dans chaque classe tarifaire.

9.5.1 Choix d'un modèle

▷ Choix des variables explicatives

Avant de choisir les variables explicatives, nous effectuons une analyse sur l'ensemble du portefeuille dommage chez Swiss Life. Parmi toutes les variables à notre disposition (environ 150), une vingtaine, utiles pour l'étude, a été sélectionnée selon les critères de Swiss life.

police : numéro de police.

version : numéro de la version d'une police.

dnais1_ : date de naissance du 1^{er} conducteur.

dperm1_ : date de permis du 1^{er} conducteur.

datemec : date de 1^{er} mise en circulation du véhicule à assurer.

marquemp : marque du véhicule.

condex : conduite exclusive ou non (conduite du véhicule par l'assuré ou son conjoint).

usage : usage du véhicule.

cpgarage : code postal de garage habituel du véhicule.

zone : zone de tarification.

groupe : groupe du véhicule (déterminé à partir de la puissance du véhicule).

classe : classe du véhicule (déterminée à partir de la valeur du véhicule).

coefbm : coefficient de bonus/malus ou de réduction/majoration (de 0,50 à 3,50).

frmalus : franchise malus - nombre d'années au bonus 50 (maximum 3).

gtaant : code de la compagnie d'assurance précédente.

prime : montant de la cotisation hors tax (HT) technique d'une garantie.

fract : type de fractionnement : M, T, S ou A.

prelev : prélèvement automatique ou non de la prime.

Afin de déterminer les coefficients tarifaires par application du modèle linéaire généralisé, nous avons dû dans un premier temps mettre en classes plusieurs variables présentant initialement de nombreuses modalités.

Date de naissance

Pour le premier conducteur, la date de naissance a été mise en classe de la façon suivante (Table 22) :

Modalité	Nom modalité	Modalité	Nom modalité	Modalité	Nom modalité
modalité 1	≤ 19	modalité 2	[20,24]	modalité 3	[25,29]
modalité 4	[30,34]	modalité 5	[35,39]	modalité 6	[40,44]
modalité 7	[45,49]	modalité 8	[50,54]	modalité 9	[55,59]
modalité 10	[60,64]	modalité 11	[65,69]	modalité 12	[70,74]
modalité 13	[75,79]	modalité 14	≤ 80		

TABLE 22 – Les modalités de l'âge du 1^{er} conducteur

Age du véhicule

La variable "datemec" a été regroupée en 6 classes. Un découpage par tranches de deux ans a été choisi car le contrôle technique est obligatoire tous les deux ans (après quatre ans d'ancienneté du véhicule) (Table 23).

Modalité	Nom modalité	Age du véhicule
modalité 1	mec 10 ou 11	Véhicule de moins de 2 ans exclu
modalité 2	mec 08 ou 09	Véhicule de 2 à 4 ans exclu
modalité 3	mec 06 ou 07	Véhicule de 4 à 6 ans exclu
modalité 4	mec 04 ou 05	Véhicule de 6 à 8 ans exclu
modalité 5	mec 02 ou 03	Véhicule de 8 à 10 ans exclu
modalité 6	mec avant 01	Véhicule de plus de 10 ans

TABLE 23 – Les modalités de l'âge du véhicule

Marque du véhicule

Les marques ont été regroupées en 15 modalités selon la marque ou le pays (Table 24).

Modalité	Nom modalité	Marque du véhicule
modalité 1	Citroën	Citroën
modalité 2	Renault	Renault
modalité 3	Peugeot	Peugeot, Talbot, Simca, Matra
modalité 4	Volkswagen	Volkswagen, Seat, Skoda
modalité 5	All Eco	Opel, Ford
modalité 6	Allemagne	Audi, BMW, Mercedes, Porsche, Saab, Volvo
modalité 7	Rover	Rover, BLMC
modalité 8	Italie	Alfa Romeo, Lancia, Fiat, Autobianchi
modalité 9	Américain	Buick, Chevrolet, Chrysler, Pontiac
modalité 10	Asie	Daewoo, Datsun, Mitsubishi, Nissan, Suzuki, Toyota, Honda, Mazda, Hyundai,...
modalité 11	Anglais	Daimler, Rayton, Jaguar, Lotus
modalité 12	Est	Lada, Polski, Tata, Zaz, ...
modalité 13	4x4	Jeep, Land Rover, Willys, Santana, Isuzu, Cherokee, ...
modalité 14	Utilitaire	Avion, Bedford, Iveco, GME, Betaillière, Avia, Hymer, ...
modalité 15	Voiturette	Axiam, Microcar, AutoStar, Erad, Teilhol, Gateau ...

TABLE 24 – Les modalités du marque du véhicule

Usage du véhicule

L'usage est un critère demandé au souscripteur pour le calcul de sa prime automobile. Il dépend de la profession et des déplacements effectués.

Les différents usages sont les suivants :

U01 : tournées (tournées ou visites régulières de clientèle, chantiers, agences, ...)

U03 : affaires (sans tournées).

Déplacements privés et professionnels (sans tournées) :

P43 : agriculteurs et professions annexes.

P44 : exploitations familiales agricole.

P30 : artisans.

M30 : artisans en société.

P41 : commerçants.

M41 : commerçants en société.

P42 : ecclésiastiques.

P40 : officiers ministériels.

P50 : professions libérales sédentaires.

Déplacements privés et trajet :

P10 : salariés.

P45 : salariés agricoles.

P46 : étudiants.

P20 : fonctionnaires.

Déplacements privés :

P60 : retraités.

U04 : sans profession.

Pour diminuer le nombre de modalités de cette variable, des usages ont été regroupés par domaine d'activités :

M30 - P30 : artisans.

M41 - P41 : commerçants.

P43 - P44 - P45 : domaine agricole.

P10 - U04 - P46 - P60 - P50 : salariés et sans profession.

U01 - U03 : commerciaux.

P40 - P42 - P20 : publique.

Région commerciale

La variable "cpgarage" a été regroupée en 7 classes suivant la ville d'habitation. Les différents zonages sont les suivants (Table 25) :

Modalité	Nom modalité	Modalité	Nom modalité
modalité 1	Est	modalité 2	I.D.F
modalité 3	Nord	modalité 4	Ouest
modalité 5	R.A.A	modalité 6	Sud
modalité 7	Sud Ouest		

TABLE 25 – Les modalités de région commerciale

Afin de diminuer le nombres de variables, nous avons créé de nouvelles variables croisant des variables initiales.

Prélèvement et fractionnement du paiement

Quatre types de fractionnement sont proposés aux assurés : mensuel, trimestriel, semestriel ou annuel. Le fractionnement mensuel implique forcément le prélèvement automatique, pour les autres cas, l'assuré choisit ou non le prélèvement.

Une nouvelle variable "paiem" a été créé en croisant les variables "fract" et "prelev". Les modalités de "paiem" sont construites de la manière suivante (Table 26) :

Modalité	Fract	Prelev
modalité 1	M	O
modalité 2	T, S ou A	O
modalité 3	T ou S	N
modalité 4	A	N

TABLE 26 – Les modalités du prélèvement automatique ou non de la prime

Coefficient réduction-majoration et franchise malus

Tout assuré sans antécédent d'assurance démarre avec un coefficient réduction-majoration à 1. Après une année sans sinistre responsable, celui-ci est réduit de 5 % ; dans le cas contraire, il est augmenté de 25 %. Il vaut 0,50 au minimum 3,50 au maximum.

Ce coefficient (multiplié par 100) a été mis en classe de la façon suivante :

Principe de la franchise malus : aucune majoration n'est appliquée pour le premier sinistre survenu après une première période d'au moins trois ans au cours de laquelle le coefficient réduction-majoration a été égal à 0,50.

La variable "frmalus" nous renseigne sur le nombre d'années qu'un assuré est à un coefficient réduction-majoration de 0,50 (nombre variant de 0 à 3). La modalité "50" créée précédemment a donc été croisée avec la variable "frmalus", on obtient donc quatre nouvelles modalités :

coef50 - fr 3	coef50 - fr 2
coef50 - fr 1	coef50 - fr 0.

Tris à plat

La tarification a été faite en tenant compte du nombre d'années police. En effet, l'année police représente la durée pendant laquelle court la police d'assurance. Dès qu'un mouvement de gestion a lieu sur un contrat, une nouvelle image apparaît. A chaque image est associée une proportion d'année police et à chaque année police est associée une prime. Au cours d'une année, suite à la résiliation d'un contrat ou à la souscription d'une affaire nouvelle, la proportion d'année police alors indiquée sera inférieure à 1.

Vous trouverez ci-dessous la Table 27 mettant en évidence le pourcentage d'années police dans chaque tranche d'âge pour l'ensemble du portefeuille.

Tranche d'âge	Poids de tranche d'âge	Pourcentage	Histogramme
<= 24	34 606	4,4%	***
[25,29]	43 809	5,6%	***
[30,34]	57 608	7,4%	****
[35,39]	79 890	10,3%	*****
[40,44]	93 218	12,0%	*****
[45,49]	94 082	12,1%	*****
[50,54]	86 708	11,1%	*****
[55,59]	82 852	10,7%	*****
[60,64]	75 174	9,7%	****
[65,69]	47 730	6,1%	***
[70,74]	37 223	4,8%	***
[75,79]	25 382	3,3%	**
<= 80	19 494	2,5%	**

TABLE 27 – Date de naissance du 1^{er} conducteur

Vous trouverez ci-dessous la Table 28 mettant en évidence le pourcentage d'années police dans chaque région commerciale pour l'ensemble du portefeuille.

Région commerciale	Poids de région	Pourcentage	Histogramme
Est	94 681	12,0%	*****
I.D.F	69 446	8,8%	****
Nord	196 102	24,8%	*****
Ouest	57 764	7,3%	****
R.A.A	97 125	12,3%	*****
Sud	156 383	19,8%	*****
Sud Ouest	119 153	15,1%	*****

TABLE 28 – Région commerciale

Vous trouverez ci-dessous la Table 29 mettant en évidence le pourcentage d'années police dans chaque tranche d'âge du véhicule pour l'ensemble du portefeuille.

Age du véhicule	Poids de l'âge du véhicule	Pourcentage	Histogramme
mec<2	84 732	10,7%	*****
mec[2,4[	81 391	10,3%	*****
mec[4,6[	84 852	10,7%	*****
mec[6,8[	84 391	10,7%	*****
mec[8,10[	90 183	11,4%	*****
mec ≥ 10	365 870	46,2%	*****

TABLE 29 – Age du véhicule

58% du portefeuille sont des voitures de plus de 8 ans.

Vous trouverez ci-dessous la Table 30 mettant en évidence le pourcentage d'années police dans chaque mode de paiement pour l'ensemble du portefeuille.

Mode de paiement	Poids du mode de paiement	Pourcentage	Histogramme
non_prelev_A	217 697	27,5%	*****
non_prelev_TS	76 471	9,7%	*****
mensuel	467 161	59,0%	*****
prelev_TSA	30 065	3,8%	***

TABLE 30 – Mode de paiement

63% des paiements se font par prélèvement automatique.

Vous trouverez ci-dessous la Table 31 mettant en évidence le pourcentage d'années police dans chaque marque du véhicule pour l'ensemble du portefeuille.

Marque du véhicule	Poids de marque	Pourcentage	Histogramme
4x4	6 957	0,9%	*
All Eco	74 089	9,5%	*****
Allemagne	69 355	8,9%	*****
Americain	37 99	0,5%	*
Anglais	3 984	0,5%	*
Asie	58 724	7,5%	****
Citroen	103 478	13,3%	*****
Pays de l'est	881	0,1%	*
Italie	33 919	4,3%	**
Peugeot	146 423	18,8%	*****
Regnault	211 744	27,1%	*****
Rover	3 745	0,5%	*
Utilitaire	3 685	0,5%	*
Voiturette	487	0,1%	*
Volkswagen	58 969	7,6%	****

TABLE 31 – Marque du véhicule

Les marques françaises (Citroën, Peugeot, Renault) représentent environ 59% du portefeuille.

Vous trouverez ci-dessous la Table 32 mettant en évidence le pourcentage d'années police ayant au moins un sinistre ou non pour l'ensemble du portefeuille.

Sinistre ou non	Poids du sinistre ou non	Pourcentage	Histogramme
0sin	717 602	90,7%	*****
au-1sin	73 816	9,3%	*****

TABLE 32 – Sinistre ou non

Nous avons décidé de faire l'analyse des correspondances multiples sur le total des garanties, pour cela nous avons tout d'abord sélectionné les variables sur

lesquelles nous effectuerons l'ACM. En effet, nous étions au départ en possession d'une vingtaine de variables. Si nous avions intégré l'ensemble de ces variables dans notre ACM, l'interprétation risquait d'être difficile. Nous avons décidé d'effectuer l'ACM sur un nombre assez restreint de variables qui semblent, à la vue des sélections faites précédemment, expliquer fortement la sinistralité. Nous avons finalement sélectionné 5 variables : marque du véhicule (15 modalités), âge du véhicule (1^{er} mise en circulation - 6 modalités), âge du 1^{er} conducteur (12 modalités), région commerciale (7 modalités), mode de paiement (4 modalités).

En effectuant l'ACM nous avons besoin de remarquer les points principaux :

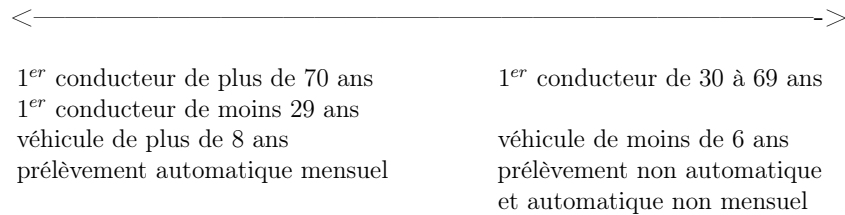
- La contribution d'une modalité à l'inertie est inversement proportionnelle à son effectif. Il est donc préférable d'éviter de travailler avec des modalités d'effectif trop petit, ces modalités risquant de perturber les résultats de part leur trop grande influence. On essaiera donc, dans la mesure du possible, de regrouper les modalités de faible effectif avec d'autres modalités.
- Le fait de travailler avec un nombre réduit de modalités permet de ne pas surcharger les représentations graphiques ce qui nuirait à leur lisibilité.

Les modalités supprimées, pour la construction des axes (faible effectif) sont les modalités de variable marque du véhicule : 4 × 4, voiturette, américain, anglais, utilitaire, rover, est.

Notre ACM a permis de réaliser la Figure 23 suivante :

Le premier facteur oppose les 1^{er} conducteurs de 35 à 49 ans avec ceux de plus de 70 ans et les prélèvements mensuels avec le non-prélèvement et les prélèvements trimestriels, semestriels et annuels.

L'axe 2 sépare les individus suivant les variables âge du 1^{er} conducteur, âge du véhicule, marque du véhicule et région commerciale. Ces résultats sont reportés sur le graphique suivant :



L'ACM induirait de regrouper la variable âge du 1^{er} conducteur de la façon suivante : ≤ 24 ans, [25,34], [35,44], [45,54], [55,64], [65,74] et 75 ans et plus.

▷ Choix de variable à expliquer, loi de probabilité et fonction lien

Suite à la sélection des variables, puis à l'étude de ces mêmes variables d'autres part, nous avons cherché à découper notre population en classes de risque homogènes à partir de la modélisation de la fréquence. Notre analyse relative aux fréquences et coûts moyens dans la partie précédente nous permet de déterminer quelle sera la base de notre tarification : nous ratifierons en fonction de la fréquence, ce qui nous permettra ensuite de calculer une prime pure (= fréquence $\times$ coût moyen). Nous avons donc créé des cases tarifaires à partir d'une fréquence estimée et non à partir d'une fréquence observée. Notons que le coût moyen des sinistres est une notion plus biaisée que la fréquence, et ceci pour 2 raisons :

- Le coût moyen est un indicateur plus fluctuant que la fréquence d'une année sur l'autre, surtout s'il y a un sinistre corporel d'un montant très élevé qui pourrait entraîner l'apparition d'un « point aberrant ».
- Le coût moyen est un indicateur biaisé en raison de la convention qui lie les assureurs entre eux. En effet, lors d'un sinistre matériel responsable, quelque soit le montant réel du sinistre, un forfait sera payé. Le coût moyen sera donc toujours le même.

Nous allons utiliser le modèle linéaire généralisé (loi poisson, fonction lien log) sur les 4 variables les plus explicatives (âge du véhicule, âge du 1^{er} conducteur, région commerciale, mode de paiement) par rapport à la fréquence que nous observons. Il est donc possible de déterminer des coefficients pour chacune de ces

variables à l'aide de la structure multiplicative suivante :

$$\log(E(freq)) = \text{constante} + \text{coef}(\text{\AA ge du v\AA hicule}) + \text{coef}(\text{\AA ge du 1}^{er}\text{ conducteur}) + \text{coef}(\text{r\AA gion commerciale}) + \text{coef}(\text{mode de paiement})$$

Nous obtenons donc :

$$E(freq) = \exp(\text{constante} + \sum \text{coef}).$$

9.5.2 Estimer les coefficients de la r gression pour la garantie RC-Tous Accidents

Gr ce   une proc dure GENMOD sous SAS, nous avons les coefficients pour chacune des variables suivantes (Figure 24 et Figure 25) :

Age du conducteur	Param�tre estim�	Intervalle de confiance � 95%		Region commerciale	Param�tre estim�	Intervalle de confiance � 95%	
<= 24	1,719442	1,661798	1,779264	Est	0,964833	0,939037	0,991338
>=75	1,065666	1,026752	1,106055	I.D.F	1,169411	1,136894	1,202858
[25,34]	1,095598	1,064282	1,127722	Nord	1,003105	0,980591	1,026033
[35,44]	1,02511	0,998301	1,052533	Ouest	0,972486	0,942141	1,003908
[45,54]	1,05001	1,022755	1,077992	R.A.A	1,06375	1,036241	1,091988
[55,64]	0,977556	0,952181	1,003506	Sud	1,1126	1,086977	1,138828
[65,74]	1	1	1	Sud Ouest	1	1	1

FIGURE 24 – Les coefficients estim s de l' ge du conducteur et de la r gion commerciale par le MLG

Age du v�hicule	Param�tre estim�	Intervalle de confiance � 95%		Mode de paiement	Param�tre estim�	Intervalle de confiance � 95%	
Mec < 2	1,659639	1,617206	1,703015	Mensuel	1,179157	1,135417	1,224583
Mec >=10	0,645132	0,630211	0,660472	non_prelev_A	0,956571	0,919891	0,994714
Mec [2,4[	1,495858	1,456738	1,535875	non_prelev_TS	1,141565	1,093955	1,191246
mec[4,6[	1,312456	1,277621	1,34824	prelev_TSA	1	1	1
mec[6,8[	1,172924	1,141108	1,205748				
mec[8,10[	1	1	1				

FIGURE 25 – Les coefficients estim s de l' ge du v hicule et de mode de paiement par le MLG

A partir de ces coefficients, nous pouvons calculer les coefficients pour une segmentation croisée des personnes de plus de 65 ans avec leur région commerciale. Nous avons les résultats suivants (Figure 26) :

Age du conducteur	Region commerciale	Paramètre estimé	Intervalle de confiance à 95%	
[65,74]	Est	0,964833	0,939037	0,991338
	I.D.F	1,169411	1,136894	1,202858
	Nord	1,003105	0,980591	1,026033
	Ouest	0,972486	0,942141	1,003908
	R.A.A	1,06375	1,036241	1,091988
	Sud	1,1126	1,086977	1,138828
	Sud Ouest	1	1	1

Age du conducteur	Region commerciale	Paramètre estimé	Intervalle de confiance à 95%	
>= [75]	Est	1,02819	0,964158	1,096474
	I.D.F	1,246201	1,167308	1,330427
	Nord	1,068975	1,006823	1,134849
	Ouest	1,036345	0,967345	1,110377
	R.A.A	1,133602	1,063962	1,207799
	Sud	1,18566	1,116055	1,259607
	Sud Ouest	1,065666	1,026752	1,106055

FIGURE 26 – Les coefficients estimés d’une segmentation croisée des seniors et de région commerciale par le MLG

D’une manière générale, les coefficients évoluent d’une manière assez cohérente et conforme à ce qu’on attendait à priori.

Les variables "Age du conducteur" et "Région commerciale" sont plutôt bien modélisées. La modélisation sanctionne assez durement sur les fréquences les assurés de plus de 65 ans résidant dans l’Ile-De-France et dans le Sud. Ceci est cohérent car nous avons déjà observé que dans ces deux régions commerciales, la fréquence des sinistres est la plus élevée.

S’agissant de l’âge du véhicule, on observe que le coefficient est plus élevé pour les véhicules plus neufs que pour les plus anciens. Concernant le mode de paiement, le coefficient est plus élevé pour le prélèvement automatique mensuel que pour le prélèvement non automatique.

Enfin, nous pouvons calculer la fréquence et la prime pure estimées suivant (Figure 27) en utilisant le MLG :

Age du conducteur	Region commerciale	Fréquence estimée	Intervalle de confiance à 95%		Prime pure	Intervalle de confiance à 95%	
[65,74]	Est	0,09	0,08	0,09	155	144	168
	I.D.F	0,10	0,10	0,11	197	183	213
	Nord	0,09	0,08	0,10	163	152	175
	Ouest	0,09	0,08	0,09	212	195	229
	R.A.A	0,09	0,09	0,10	180	168	194
	Sud	0,10	0,09	0,11	272	253	292
	Sud Ouest	0,09	0,08	0,09	218	208	229

Age du conducteur	Region commerciale	Fréquence estimée	Intervalle de confiance à 95%		Prime pure	Intervalle de confiance à 95%	
>=75	Est	0,09	0,08	0,10	189	168	211
	I,D,F	0,11	0,10	0,12	284	254	319
	Nord	0,09	0,08	0,11	342	307	381
	Ouest	0,09	0,08	0,10	188	167	212
	R,A,A	0,10	0,09	0,11	213	191	239
	Sud	0,10	0,09	0,12	346	311	386
	Sud Ouest	0,09	0,09	0,10	222	204	242

FIGURE 27 – La fréquence et la prime pure estimées par le MLG

9.5.3 Validation du modèle

▷ **Principe** La validité du modèle se mesure en premier lieu en vérifiant les hypothèses concernant les résidus du modèle, en particulier les résidus de Pearson standardisés observés définis par :

$$r^P_i = \frac{y_i - \hat{\mu}_i}{\sqrt{a(\varphi)V(\mu)}}$$

Ces résidus ont une espérance proche de 0 et une variance approximativement constante.

La validité de ces hypothèses s'effectue en traçant le nuage des résidus, en plaçant en ordonnées les résidus en fonction du numéro d'observation, ce qui permet d'identifier les observations conduisant à de grands résidus (et donc contribuant à l'inadéquation du modèle) . L'hypothèse "espérance nulle" s'apprécie par la répartition des résidus autour de l'axe des abscisses. L'hypothèse "variance constante" s'apprécie par la forme du nuage des résidus : si la variance est constante, le nuage ne doit pas présenter de forme particulière, de type « entonnoir » par exemple.

D'autre part, on propose souvent, pour mesurer la qualité du modèle, de com-

parer dans chaque case tarifaire les fréquences estimées aux fréquences observées. Si les estimations $\hat{\mu}_i$ sont proches des observations y_i , on conclut alors que le modèle est acceptable.

La précision de l'estimation se mesure également par les résidus de Pearson. On estime asymptotiquement que les r^P_i ont des distributions normales $N(0,1)$. Ainsi un résidu de Pearson supérieur à 1,96 en valeur absolue indique une observation critique.

Cette deuxième approche est parfaitement applicable lorsque la modélisation est effectuée sur la base d'un faible nombre de variables explicatives. En effet, si à partir d'un nombre de paramètres restreint, on arrive pour chacune des classes tarifaires à bien approcher les valeurs observées, on peut raisonnablement conclure que le modèle est valide, puisqu'il réalise alors le compromis faible nombre de paramètres - bonne qualité d'ajustement qui définit un bon modèle.

▷ Résultats

Nous avons pu calculer les résidus de Pearson, ce qui nous donne le graphique suivant (Figure 28) :

N	Moyenne	Ecart-type	Err. type	Minimum	Maximum
1176	0.0208	1.1369	0.0332	-4.1039	5.7544

Moyenne	Moyenne de l'IC à 95%	Ecart-type	Ecart-type de l'IC à 95%	
0.0208	-0.0442	0.0859	1.1369	
			1.0928	1.1848

DDL	Valeur du test t	Pr > t
1175	0.63	0.5295

FIGURE 28 – Résidus de Pearson de la fréquence

L'étude du nuage ne fait pas apparaître de forme particulière, du type "entonnoir", et on peut considérer que l'hypothèse "variance constante" est vérifiée.

Nous avons effectué le t-test (le test de l'hypothèse de "espérance nulle"), nous obtenons les résultats suivants (Figure 29 et Figure 30) :

La moyenne empirique des résidus est égale à 0,0208 et la p-value de t-test, égale à 0,5295, est supérieure à 0,05. Nous acceptons donc l'hypothèse de "espérance nulle". De plus, le test nous montre également que la distribution des résidus de Pearson est normale $N(0,1)$. Le graphique Q-Q des résidus de Pearson confirme une fois encore la normalité des résidus de Pearson. Le modèle est donc bien validé.

N	Moyenne	Ecart-type	Err. type	Minimum	Maximum
1176	0.0208	1.1369	0.0332	-4.1039	5.7544

Moyenne	Moyenne de l'IC à 95%	Ecart-type	Ecart-type de l'IC à 95%
0.0208	-0.0442	0.0859	1.1369
	1.0928	1.1848	

DDL	Valeur du test t	Pr > t
1175	0.63	0.5295

FIGURE 29 – T-test des résidus de Pearson

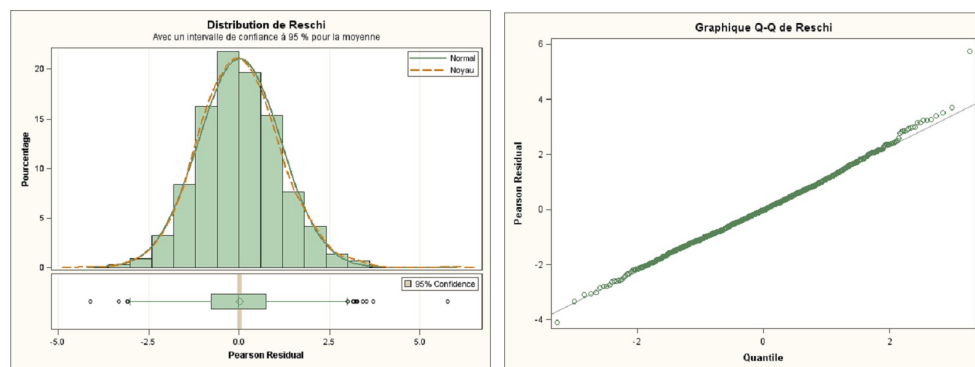


FIGURE 30 – La distribution et le graphique Q-Q des résidus de Pearson

9.6 Comparaison des primes entre les deux méthodes de tarification (S/C et MLG)

En prenant le ratio de la prime pure théorique par le MLG et la prime pure initiale, on pourrait déterminer des nouveaux coefficients tarifaires pour la segmentation croisée de l'âge et de région commerciale sur chaque type de garanties par le MLG suivant (Table 33, Table 34 et Table 35) :

Age du conducteur	Région commerciale	Prime pure initiale (données complètes)	Prime pure théorique (MLG)	Coefficient tarifaire
[65,74]	Est	172	155	0,91
	I.D.F	211	197	0,93
	Nord	165	163	0,99
	Ouest	225	212	0,94
	R.A.A	188	180	0,96
	Sud	300	272	0,91
>= 75	Sud Ouest	228	218	0,96
	Est	193	189	0,98
	I.D.F	244	284	1,17
	Nord	350	342	0,98
	Ouest	190	188	0,99
	R.A.A	216	213	0,99
	Sud	311	346	1,11
	Sud Ouest	205	222	1,08

TABLE 33 – Les coefficients tarifaires pour la garantie RC-Tous Accidents

Age du conducteur	Région commerciale	Prime pure initiale (données complètes)	Prime pure théorique (MLG)	Coefficient tarifaire
[65,74]	Est	9	7	0,77
	I.D.F	29	19	0,63
	Nord	12	10	0,79
	Ouest	8	5	0,66
	R.A.A	16	14	0,88
	Sud	17	14	0,78
>= 75	Sud Ouest	15	10	0,70
	Est	7	6	0,89
	I.D.F	10	6	0,58
	Nord	8	6	0,77
	Ouest	7	4	0,60
	R.A.A	7	5	0,81
	Sud	12	11	0,93
	Sud Ouest	12	11	0,89

TABLE 34 – Les coefficients tarifaires pour la garantie Vol-Incendie

Age du conducteur	Région commerciale	Prime pure initiale (données complètes)	Prime pure théorique (MLG)	Coefficient tarifaire
[65,74]	Est	23	23	1,01
	I.D.F	24	24	1,01
	Nord	26	26	1,00
	Ouest	21	21	1,00
	R.A.A	21	21	0,99
	Sud	17	16	0,94
	Sud Ouest	21	21	0,97
≥ 75	Est	12	14	1,23
	I.D.F	13	14	1,10
	Nord	14	14	1,04
	Ouest	10	13	1,28
	R.A.A	11	11	0,94
	Sud	10	9	0,97
	Sud Ouest	12	11	0,92

TABLE 35 – Les coefficients tarifaires pour la garantie Bris de Glace

Nous pouvons remarquer que :

- **Pour la garantie RC-Tous Accidents** : les personnes de 65 à 74 ans sont toutes sur-tarifées pour toutes les régions commerciales, est donc envisageable une réduction de prime. Nous remarquons également que les gens de plus de 75 ans habitant dans l'Est, le Nord, l'Ouest et le Rhône-Alpes apparaissent comme étant sur tarifés (ils nécessitent une réduction de tarif). Nous notons également que comme signalé précédemment, les seniors de 75 ans et plus dans le Sud, le Sud-Ouest et l'Ile-De-France sont réellement sous tarifés (ils nécessitent une augmentation de tarif).
- **Pour la garantie Vol-Incendie** : il apparait qu'il y ait une sur-tarification pour toutes les personnes de plus de 65 ans, quelle que soit leur région commerciale. Aussi nécessiteraient-ils une réduction de tarif pour cette garantie.
- **Pour la garantie Bris de Glace** : les seniors de 65 à 74 ans apparaissent comme étant bien tarifés dans presque toutes les régions commerciales (sauf dans le Sud et Sud-Ouest, il y a une légèrement sur-tarification), alors que les seniors de plus de 75 ans apparaissent comme étant sous tarifés dans presque toutes les régions commerciales (sauf dans le Rhône-Alpes, le Sud et le Sud-Ouest, il y a une sur-tarification). Nous avons donc besoin d'un ajustement raisonnable des coefficients tarifaires.

Enfin, nous avons comparé des primes déterminées par le MLG et par la méthode basée sur S/C avec la prime initiale. Nous avons aussi travaillé sur des

données écrêtées dans cette partie. En effet, en RC-Tous Accidents, la charge des sinistres peut être très élevée ; c'est pourquoi, afin de ne pas tenir compte du caractère exceptionnel de cette charge, il peut être décidé de l'écrêter : on définit pour cela un seuil au-dessus duquel la charge est mutualisée. En d'autres termes au-dessus de ce seuil, la charge restante est répartie de manière proportionnelle sur les autres contrats.

Nous remarquons que si nous utilisons les données complètes et les données écrêtées, les coefficients tarifaires sont différents selon la méthode basée sur S/C, alors qu'ils sont identiques selon le MLG. Plus précisément, rappelons que le coefficient tarifaire basé sur le MLG est calculé selon la formule suivante :

$$K = \frac{f_1 \times C_1}{f_0 \times C_0}$$

où f_0, C_0 sont, respectivement, la fréquence observée et le coût moyen initial, f_1, C_1 sont, respectivement, la fréquence estimée et le coût moyen selon le MLG.

On utilise le MLG pour estimer uniquement la fréquence des sinistres et non pas le coût moyen qui reste donc identique avant et après application de méthode. Les coefficients tarifaires déterminés par le MLG ne dépendent pas du fait d'avoir utilisé des données complètes ou écrêtées. Les Table 36, Table 37, Table 38 présentent des comparaisons des primes obtenues par les 2 méthodes (S/C et MLG) pour chaque type de garantie.

Age du conducteur	Région commerciale	Coefficient tarifaire (données complètes)	Evolution des primes	Coefficient tarifaire (données écrêtées 50K)	Evolution des primes	Coefficient tarifaire (MLG)	Evolution des primes
[65,74]	Est	0,92	-8 %	1,01	1 %	0,91	-9 %
	I.D.F	0,92	-8 %	0,99	-1 %	0,93	-7 %
	Nord	0,78	-22 %	0,95	-5 %	0,99	-1 %
	Ouest	1,25	25 %	1,01	1 %	0,94	-6 %
	R.A.A	0,80	-20 %	0,90	-10 %	0,96	-4 %
	Sud	1,13	13 %	1,05	5 %	0,91	-9 %
	Sud Ouest	1,07	7 %	1,05	5 %	0,96	-4 %
>= 75	Est	1,02	2 %	1,12	12 %	0,98	-2 %
	I.D.F	1,01	1 %	0,92	-8 %	1,17	17 %
	Nord	1,63	63 %	1,15	15 %	0,98	-2 %
	Ouest	1,01	1 %	1,20	20 %	0,99	-1 %
	R.A.A	0,93	-7 %	1,01	1 %	0,99	-1 %
	Sud	1,21	21 %	1,08	8 %	1,11	11 %
	Sud Ouest	1,02	2 %	1,06	6 %	1,08	8 %

TABLE 36 – Comparaisons des primes obtenues par les 2 méthodes (S/C et MLG) pour la RC-Tous Accidents

Age du conducteur	Région commerciale	Coefficient tarifaire (données complètes)	Evolution des primes	Coefficient tarifaire (données écrêtées 50K)	Evolution des primes	Coefficient tarifaire (MLG)	Evolution des primes
[65,74]	Est	0,63	-37 %	0,63	-37 %	0,77	-23 %
	I.D.F	0,86	-14 %	0,86	-14 %	0,63	-37 %
	Nord	0,68	-32 %	0,69	-31 %	0,79	-21 %
	Ouest	0,68	-32 %	0,68	-32 %	0,66	-34 %
	R.A.A	0,79	-21 %	0,78	-22 %	0,88	-12 %
	Sud	0,59	-41 %	0,59	-41 %	0,78	-22 %
	Sud Ouest	0,68	-32 %	0,68	-32 %	0,70	-30 %
>= 75	Est	0,49	-51 %	0,49	-51 %	0,89	-11 %
	I.D.F	0,29	-71 %	0,29	-71 %	0,58	-42 %
	Nord	0,50	-50 %	0,51	-49 %	0,77	-23 %
	Ouest	0,59	-41 %	0,59	-41 %	0,60	-40 %
	R.A.A	0,34	-66 %	0,34	-66 %	0,81	-19 %
	Sud	0,43	-57 %	0,43	-57 %	0,93	-7 %
	Sud Ouest	0,62	-38 %	0,61	-39 %	0,89	-11 %

TABLE 37 – Comparaisons des primes obtenues par les 2 méthodes (S/C et MLG) pour la Vol-Incendie

Age du conducteur	Région commerciale	Coefficient tarifaire (données complètes)	Evolution des primes	Coefficient tarifaire (données écrêtées 50K)	Evolution des primes	Coefficient tarifaire (MLG)	Evolution des primes
[65,74]	Est	0,98	-2 %	0,99	-1 %	1,01	1 %
	I.D.F	1,07	7 %	1,06	6 %	1,01	1 %
	Nord	1,11	11 %	1,11	11 %	1,00	0 %
	Ouest	1,00	0 %	1,00	0 %	1,00	0 %
	R.A.A	1,08	8 %	1,09	9 %	0,99	-1 %
	Sud	1,10	10 %	1,10	10 %	0,94	-6 %
	Sud Ouest	1,05	5 %	1,04	4 %	1,00	0 %
>= 75	Est	0,61	-39 %	0,62	-38 %	1,23	23 %
	I.D.F	0,71	-29 %	0,71	-29 %	1,10	10 %
	Nord	0,74	-26 %	0,74	-26 %	1,04	4 %
	Ouest	0,58	-42 %	0,58	-42 %	1,28	28 %
	R.A.A	0,78	-22 %	0,78	-22 %	0,94	-6 %
	Sud	0,83	-17 %	0,83	-17 %	0,97	-3 %
	Sud Ouest	0,79	-21 %	0,79	-21 %	0,92	-8 %

TABLE 38 – Comparaisons des primes obtenues par les 2 méthodes (S/C et MLG) pour la Bris de Glace

Nous pouvons remarquer que :

- **Pour la garantie RC-Tous Accidents** : selon la méthode basée sur S/C, on observe généralement une évolution moins élevée avec les données écrêtées que avec les données complètes. Selon le MLG, les primes ne dépendent pas de type de données et elles évoluent plus linéairement. Pour les seniors de 65 à 74 ans, dans l’Ile-De-France, le Nord, le Rhône-Alpes, la tendance

de l'évolution des primes est identique (on peut envisager de diminuer des primes) pour l'une au l'autre des méthodes. Pour les seniors de plus de 75 ans, les primes peuvent être augmentées dans le Sud et le Sud-Ouest, quelles que soient les méthodes utilisées.

- **Pour la garantie Vol-Incendie** : des primes déterminées par la méthode basée sur S/C sont presque identiques pour des données écrêtées et des données complètes. C'est assez évident parce que le montant des sinistres de type vol et incendie n'est pas très grave. La tendance de l'évolution des primes est toute identique (il faut diminuer des primes) pour l'une au l'autre de méthode, alors que l'évolution des primes selon le MLG est moins forte que selon la méthode basée sur S/C.
- **Pour la garantie Bris de Glace** : des primes déterminées par la méthode basée sur S/C sont presque identiques pour des données écrêtées et des données complètes. C'est assez évident parce que le montant des sinistres de type bris de glace n'est pas très grave. L'évolution des primes déterminées par le MLG est moins forte que par la méthode basée sur S/C.

Si Swiss Life appliquait des nouvelles grilles tarifaires pour les gens de plus de 65 ans, voici les nouvelles primes que Swiss Life recevrait (Table 39) :

	Nouvelles primes totales	Évolution des primes
Méthode basée sur S/C (données complètes)	260 579 070	-0,25%
Méthode basée sur S/C (données écrêtées)	260 126 570	-0,43%
Le MLG	259 679 190	-0,60 %

TABLE 39 – Les nouvelles primes totales

Avec l'introduction d'une classification par l'âge, le résultat montre que des personnes âgées sont actuellement sur-tarifées quelles que soient les méthodes utilisées. Une stratégie de Swiss Life, pour être plus compétitif, pourrait être de réduire les primes des seniors.

Conclusion

Notre étude axée sur le cas des seniors (de plus de 65 ans) se compose deux parties principales : la première consiste en une analyse exploratoire sur des données de Swiss Life, et la seconde porte sur la tarification lorsqu'on introduit une nouvelle classification par l'âge.

Nous présentons une synthèse de ce que nous avons obtenu :

- L'analyse exploratoire permet de dégager les principaux résultats suivants :
 1. Le nombre de sinistres et le montant de sinistres ne semblent pas liés linéairement à l'âge du conducteur alors qu'ils sont liés linéairement au coefficient tarifaire. De plus, l'âge du conducteur est lié linéairement au coefficient de bonus-malus.
 2. On ne peut déduire que les sinistres corporels sont fortement liés aux personnes âgées sur les données de Swiss Life. La proportion de sinistres corporels correspond seulement à 3% de l'ensemble des sinistres. Les résultats de l'ACM en sont donc affectés.
 3. La fréquence de sinistres corporels est faible dans toutes les tranches d'âge, alors que la gravité de sinistres corporels est très importante. Ces résultats sont affectés fortement par les valeurs extrêmes. Les montants de sinistres corporels sont très dispersés, il y a quelques sinistres qui coûtent très chers et qui dominent les autres.
- Comparaison avec les études menées sur la population française :
 1. Le nombre de conducteurs âgés présents dans le portefeuille automobile a également tendance à augmenter les années précédentes ; c'est un résultat que l'on retrouve sur l'étude faite sur la population française.
 2. Les conducteurs âgés chez Swiss Life ont moins d'accidents que les conducteurs plus jeunes et leurs accidents sont plus graves. Ces résultats sont similaires à ceux obtenus sur la population française.
 3. On ne peut conclure que les sinistres corporels pour les personnes âgées soient plus fréquents que pour les jeunes conducteurs, contrairement à l'étude menée sur la population française.
- Les méthodes de tarification (S/C et MLG) mènent aux résultats suivants :
 1. La segmentation selon des profils de risques (l'âge et la région commerciale) est pertinente au niveau de la tarification. Les données de Swiss

Life présentent des hétérogénéités selon l'âge du conducteur et sa région commerciale.

2. La méthode basée sur S/C et le MLG montrent qu'il y a des surtarifications et sous-tarifications par rapport aux calculs initiaux des primes de Swiss Life pour quelques tranches d'âge et quelques régions commerciales. Pour être plus compétitif dans le marché d'assurance automobile, Swiss Life pourrait envisager de réduire les primes selon des profils de risque (l'âge et la région commerciale) des clients.
3. En utilisant le MLG, nous avons modélisé des primes plus pertinentes et plus compétitives pour les seniors.

Références

- [1] Hélène Fontaine : *Âge des conducteurs de voiture et accidents de la route. Quel risque pour les seniors*, Recherche - Transports - Sécurité, pp. 107 - 120, 2003.
- [2] Sandrine Gruau, Annick Pottier, Damien Davenne, and Pierre Denise : *facteurs d'accidents de la route par somnolence chez les conducteurs âgés. Prévention par l'activité physique*, Recherche - Transports - Sécurité (rts), pp. 134 - 144, 2003.
- [3] Catherine Gabaude : *Exploration des capacités visuelles et attentionnelles des conducteurs âgés. Intérêts et techniques*, Recherche - Transports - Sécurité (rts), pp. 165 - 176, 203.
- [4] Sylviane Lafont : *Sécurité Routière et Conducteurs Âgés : Estimation des enjeux, et contributions des facteurs cognitifs et de la démence*, PhD Thèse, 2003.
- [5] Sandy Leproust : *Inaptitude médicale à la conduite automobile : évaluation de l'opportunité d'un programme de dépistage*, PhD Thèse, 2007.
- [6] Jan Roche : *Conduite automobile et maladie d'Alzheimer*, Psychologie & NeuroPsychiatrie du vieillissement, p. Vol 3 (163 168), 2005.
- [7] Ragnhild Davidse : *Assisting the older driver : Intersection design and in-car devices to improve the safety of the older driver*, PhD Thèse, 2006.
- [8] Sara Carmel : *Driving in Old Age : Risks, Behaviors and Interventions*, PhD thesis,.
- [9] James Peter Thompson : *The Self-Regulation of the Driving Behaviour of Older Drivers : A Longitudinal Assessment*, Bachelor thesis, October 2007.
- [10] Anabela Simoes : *The cognitive training needs of older drivers*, Recherche - Transports - Sécurité (rts), p. 145 155, Vol 79 (2003).
- [11] Richard Lefrancois and Monia D'amours : *Exposure and risk factors among elderly drivers - A case-control study*, Accident Analysis and Prevention - Accid Anal Prev, pp. 267-275, Vol 29 no 3 (1997).
- [12] Silvia Ravera : *Psychotropic medications and traffic safety - Contributions to risk assessment and risk communication*, PhD thesis, 2003.
- [13] Pierre Van Elslande : *Les erreurs des conducteurs âgés*, Recherche - Transports - Sécurité (rts), pp. 190-202, Vol 81 (2003).
- [14] INRETS, *Seniors : concilier mobilité et sécurité*, Rapport 2003.
- [15] OCDE, *Vieillissement et transports : concilier mobilité et sécurité*, Rapport 2001.

- [16] Observatoire National Interministériel de Sécurité Routière, *La sécurité des conducteurs de VL âgés*, Rapport 2001.
- [17] VicRoads Patricia Williams : *Survey of older licence holders : Are they still driving ? What is the association between mileage driven and crash risk ?*, in Proceedings of the Australasian Road Safety Research, Policing and Education Conference , 2005.
- [18] Status Report : *Older drivers up close : they aren't dangerous, except maybe to themselves*, Vol. 36, No. 8, September 8, 2001.
- [19] McGwin G, Sims RV, Pulley L, Roseman JM : *Relations among chronic medical conditions, medications, and automobile crashes in the elderly : a population-based case-control study*, Rapport, 2000.
- [20] Gislason T, Tomasson K, Reynisdottir H, Bjornsson JK, Kristbjarnarson H : *Medical risk factors amongst drivers in single-car accidents*, Rapport, 2007.
- [21] Hansotia P, Broste SK : *The effect of epilepsy or diabetes mellitus on the risk of automobile accidents*, Rapport, 1991.
- [22] Simonson E, Baker C, Burns N, Keiper C, Schmitt OH, Stackhouse S : *Cardiovascular stress (electrocardiographic changes) produced by driving an automobile*, Rapport, 1968.
- [23] La Fondation MAIF : *Etude épidémiologique des accidents de la circulation dans une cohorte de 20 000 employés d'une entreprise publique française*, Rapport, 22 janvier 2002.
- [24] Direction de la recherche, des études, de l'évaluation et des statistiques DREES, *Document de travail Handicaps-Incapacités-Dépendance*, 2000.
- [25] Observatoire National Interministériel de la Sécurité Routière, *La sécurité routière en France Bilan de l'année 2010*, 2010.
- [26] Jennifer Oxley, Judith Charlton, and Brian Fildes : *The effect of cognitive impairment on older pedestrian behaviour and crash risk*, 2005.
- [27] Observatoire National Interministériel de la Sécurité Routière, *Document de travail - Accidents corporels de la circulation routiere annee 2011*,.
- [28] Liisa Hakamies-Blomqvist : *Ageing Europe : The challenges and opportunities for transport safety*, Janvier 2003.
- [29] Joëlle Gaymu : *Aspects démographiques du vieillissement*, Février 2006.
- [30] Joël Ankri : *Épidémiologie des démences et de la maladie d'Alzheimer*, Février 2006.
- [31] CNRS, *Recherche sur les relations pouvant exister entre la qualité du vieillissement et la fréquence des accidents de la route*, Avril 2003.

- [32] Serge Lesourd : *La conduite automobile après 60 ans : représentations et vécu subjectif des personnes concernées*, Avril 2003.
- [33] Iain S. Begg, Jean-François Yale, Robyn L. Houlden, Richard C. Rowe, and James McSherry : *Canadian Diabetes Association's Clinical Practice Guidelines for Diabetes and Private and Commercial Driving*, 2003.
- [34] Judith Charlton et al : *Influence of chronic illness on crash involvement of motor vehicle drivers*, Avril 2004.
- [35] Association Prévention Routière, FFSA : *Seniors, mobilité et risque routier*, Mars 2009.
- [36] Michael Rytz : *Les seniors et la sécurité routière – de l'analyse à la prévention*, 2006.
- [37] Gaultier & Associés : *Les personnes âgées et le risque routier*, Novembre 2008.
- [38] Association Prévention Routière : *Rapport d'activité 2011, 2012*.
- [39] ONISR : *Victimes dans accidents corporels de la circulation routière*, 1985 à 2010.
- [40] S. Bonin-Guillaume : *Elderly drivers : Assessing performance or predicting driving safety ?*, 2010.
- [41] Joanne M. Wood, Kaarin J. Anstey, Graham K. Kerr, Philippe F. Lacherez, and Stephen Lord : *A Multidomain Approach for Predicting Older Driver Safety Under In-Traffic Road Conditions*, 2008.
- [42] Huang Helai, Chin Hoong Chor, and Md. Mazharul Haque : *Severity of driver injury and vehicle damage in traffic crashes at intersections : A Bayesian hierarchical analysis*, 2008.
- [43] Nicol Korner-Bitensky, Johanne Desrosiers, Malcolm Man-Son-Hing, Shawn Marshall Ailene Kua : *Older driver retraining : A systematic review of evidence of effectiveness*, 2007.
- [44] J. Langford and S. Koppel : *Epidemiology of older driver crashes – Identifying older driver risk factors and exposure patterns*, 2006.
- [45] Sylvie Bonin-Guillaume, Aymé Mohr, Arlette Parola, Raphaël Grandjean, Yves-Michel Frances Jean-Pascal Assailly : *Les conducteurs âgés en bonne santé font plus d'erreurs et d'oublis que d'infractions*, 2006.
- [46] Jean Pascal : *ASSAILLY, Le suivi des conducteurs agés*, 2004.

Annexe 1 : Méthode des multiplicateurs de Lagrange

Pour chercher les optimums d'une fonction

$$f(t_1, t_2, \dots, t_p)$$

de p variables liées par une relation

$$l(t_1, t_2, \dots, t_p) = cste$$

on calcule les dérivées partielles de la fonction

$$g(t_1, t_2, \dots, t_p) = f(t_1, t_2, \dots, t_p) - \lambda(l(t_1, t_2, \dots, t_p) - cste)$$

par rapport à chacune des variables. En annulant ces p dérivées partielles et en ajoutant la contrainte, on obtient un système de $p + 1$ équations à $p + 1$ inconnues. Les $p + 1$ inconnues sont les valeurs des variables t_i ($i = 1, \dots, n$) et de λ appelé le “multiplicateur de Lagrange”. L'existence de solutions à ce système est une condition nécessaire mais pas suffisante à l'existence d'un optimum pour la fonction f .

On peut généraliser ce problème au cas où les p variables sont soumises à c contraintes. On construit la fonction $g(t_1, t_2, \dots, t_p)$ en rajoutant une combinaison linéaire des c contraintes, dont les coefficients $\lambda_1, \lambda_2, \dots, \lambda_c$ sont les multiplicateurs de Lagrange. On doit alors résoudre un système de $p + c$ équations à $p + c$ inconnues.

Annexe 2 : Détails ACP

*** Analyse auto ***

REPRESENTATION DES POINTS DANS LE NOUVEL ESPACE

Procédure PRINCOMP

Observations	100507
Variables	12

Simple Statistics						
	age_cond	durée_contrat	an_cond	Sex	an_route	nbre_sin
Mean	49.46044835	7.423575731	28.00980980	0.6438855005	10.09144517	0.1923149631
STD	65.23455718	6.928229905	66.56465607	0.4788520065	7.05707495	0.4625955339

Simple Statistics						
cout_sinistre	coefbm	frmalus	ctnet	ctsr	coefqua	
313.677365	55.96272896	1.794113843	8871.529565	10027.35269	59.12351379	
7388.738811	12.48589851	1.406575627	2015.894139	238.57267	13.59641580	

Correlation Matrix							
		age_cond	durée_contrat	an_cond	Sex	an_route	nbre_sin
age_cond		1.0000	0.0812	0.9628	0.0619	0.0236	-.0152
durée_contrat		0.0812	1.0000	0.0703	0.0560	0.0156	-.0088
an_cond		0.9628	0.0703	1.0000	0.0724	0.0168	-.0127
Sex		0.0619	0.0560	0.0724	1.0000	0.0834	-.0168
an_route		0.0236	0.0156	0.0168	0.0834	1.0000	-.1285
nbre_sin		-.0152	-.0088	-.0127	-.0168	-.1285	1.0000
cout_sinistre		-.0027	-.0028	-.0022	-.0094	-.0064	0.0929
coefbm	coefficient bonus malus du composant contrat auto	-.0994	-.1744	-.0950	-.0198	0.0221	0.0538
frmalus	franchise malus du composant	0.1083	0.1944	0.1034	0.0336	0.0158	-.0681
ctnet	coefficient tarifaire	-.0063	0.0750	-.0146	0.0051	0.0925	-.0193
ctsr	Coefficient surveillance des risques	-.0008	0.0117	-.0001	-.0031	-.0332	0.0478
coefqua	coefficient qualite du contrat	-.1236	-.2272	-.1145	-.0197	-.0143	0.0520

*** Analyse auto ***

REPRESENTATION DES POINTS DANS LE NOUVEL ESPACE

Procédure PRINCOMP

Correlation Matrix					
cout_sinistre	coefbm	frmalus	ctnet	ctsr	coefqua
-0.0027	-0.0994	0.1083	-0.0063	-0.0008	-0.1236
-0.0028	-0.1744	0.1944	0.0750	0.0117	-0.2272
-0.0022	-0.0950	0.1034	-0.0146	-0.0001	-0.1145
-0.0094	-0.0198	0.0336	0.0051	-0.0031	-0.0197
-0.0064	0.0221	0.0158	0.0925	-0.0332	-0.0143
0.0929	0.0538	-0.0681	-0.0193	0.0478	0.0520
1.0000	0.0079	-0.0049	0.0013	0.0047	0.0085
0.0079	1.0000	-0.6091	0.1208	0.1014	0.6471
-0.0049	-0.6091	1.0000	-0.1400	-0.0910	-0.6242
0.0013	0.1208	-0.1400	1.0000	0.1179	0.1594
0.0047	0.1014	-0.0910	0.1179	1.0000	0.0496
0.0085	0.6471	-0.6242	0.1594	0.0496	1.0000

Eigenvalues of the Correlation Matrix				
	Valeur propre	Différence	Proportion	Cumulé
1	2.54172662	0.70356724	0.2118	0.2118
2	1.83815938	0.61996664	0.1532	0.3650
3	1.21819274	0.08684441	0.1015	0.4665
4	1.13134833	0.11088332	0.0943	0.5608
5	1.02046501	0.05528528	0.0850	0.6458
6	0.96517973	0.06465486	0.0804	0.7263
7	0.90052487	0.06931229	0.0750	0.8013
8	0.83121257	0.05164470	0.0693	0.8706
9	0.77956787	0.38561276	0.0650	0.9355
10	0.39395511	0.05129686	0.0328	0.9684
11	0.34265825	0.30564871	0.0286	0.9969
12	0.03700954		0.0031	1.0000

*** Analyse auto ***

REPRESENTATION DES POINTS DANS LE NOUVEL ESPACE

Procédure PRINCOMP

Eigenvectors								
		axe1	axe2	axe3	axe4	axe5	axe6	axe7
age_cond		0.315312	0.625086	-.064174	-.042954	-.032539	-.048412	-.006964
durée_contrat		0.216478	-.043940	0.225016	0.456808	-.022417	0.080997	-.557238
an_cond		0.311317	0.627672	-.070985	-.049593	-.026958	-.035859	0.000570
Sex		0.059794	0.079932	0.301756	0.025551	0.502740	0.745347	0.122436
an_route		0.015834	0.041256	0.598094	-.131274	0.319723	-.261785	0.175605
nbre_sin		-.066060	0.025536	-.502978	0.377023	0.233560	0.230539	-.200898
cout_sinistre		-.012994	0.005077	-.241965	0.279851	0.682290	-.487820	0.184428
coefbm	coefficient bonus malus du composant contrat auto	-.485489	0.255474	0.025354	-.038155	0.027984	0.035642	-.069166
frmalus	franchise malus du composant	0.486516	-.244175	0.009841	0.030365	-.001967	-.030673	0.088396
ctnet	coefficient tarifaire	-.132441	0.120353	0.421472	0.471236	-.112988	-.236975	-.233994
ctsr	Coefficient surveillance des risques	-.087973	0.086866	0.038413	0.564348	-.328650	0.125664	0.699668
coefqua	coefficient qualite du contrat	-.500760	0.240161	0.007070	-.075783	0.033264	0.021449	-.111127

*** Analyse auto ***

REPRESENTATION DES POINTS DANS LE NOUVEL ESPACE

Procédure PRINCOMP

Eigenvectors				
axe8	axe9	axe10	axe11	axe12
0.008287	-.001408	-.001468	0.002316	0.707357
-.258271	0.550198	-.009754	0.069008	-.006613
0.004093	-.005516	0.000703	0.012425	-.706680
-.176657	-.208203	-.013889	-.014913	0.008658
0.548930	0.341835	-.030569	0.048733	-.004680
0.671884	0.007647	0.016055	0.006837	0.000842
-.354015	-.007492	-.007164	-.002274	0.000287
-.068275	0.212187	0.518029	-.609963	-.002025
0.065666	-.118173	0.810474	0.136958	0.000553
0.091883	-.656234	0.018255	-.071676	-.006295
-.022844	0.214827	0.006611	0.068525	0.001063
-.077370	0.033959	0.269939	0.769275	0.007950

*** Analyse auto ***

TABLEAU DES VALEURS PROPRES :

Procédure FREQ

Name of Variable or Component				
NAME_	Fréquence	Pourcentage	Fréquence cumulée	Pctage. cumulé
AXE1	2.541727	21.18	2.541727	21.18
AXE2	1.838159	15.32	4.379886	36.50
AXE3	1.218193	10.15	5.598079	46.65
AXE4	1.131348	9.43	6.729427	56.08
AXE5	1.020465	8.50	7.749892	64.58
AXE6	0.96518	8.04	8.715072	72.63
AXE7	0.900525	7.50	9.615597	80.13
AXE8	0.831213	6.93	10.44681	87.06
AXE9	0.779568	6.50	11.22638	93.55
AXE10	0.393955	3.28	11.62033	96.84
AXE11	0.342658	2.86	11.96299	99.69
AXE12	0.03701	0.31	12	100.00

*** Analyse auto ***

TABLEAU DES VECTEURS PROPRES :

Obs.	NAME_	AXE1	AXE2	AXE3	AXE4
1	age_cond	0.31531	0.62509	-0.06417	-0.04295
2	durée_contrat	0.21648	-0.04394	0.22502	0.45681
3	an_cond	0.31132	0.62767	-0.07098	-0.04959
4	Sex	0.05979	0.07993	0.30176	0.02555
5	an_route	0.01583	0.04126	0.59809	-0.13127
6	nbre_sin	-0.06606	0.02554	-0.50298	0.37702
7	cout_sinistre	-0.01299	0.00508	-0.24197	0.27985
8	coefbm	-0.48549	0.25547	0.02535	-0.03815
9	frmalus	0.48652	-0.24417	0.00984	0.03036
10	ctnet	-0.13244	0.12035	0.42147	0.47124
11	ctsr	-0.08797	0.08687	0.03841	0.56435
12	coefqua	-0.50076	0.24016	0.00707	-0.07578

*** Analyse auto ***

QUALITE DE LA REPRESENTATION ET CONTRIBUTIONS DES VARIABLES

Obs.	_NAME_	CTR1	CTR2	CTR3	CTR4	CO21	CO22	CO23	CO24	QLT12	QLT13
1	Sex	0.36	0.64	9.11	0.07	0.91	1.17	11.09	0.07	2.08	13.18
2	age_cond	9.94	39.07	0.41	0.18	25.27	71.82	0.50	0.21	97.09	97.59
3	an_cond	9.69	39.40	0.50	0.25	24.63	72.42	0.61	0.28	97.05	97.67
4	an_route	0.03	0.17	35.77	1.72	0.06	0.31	43.58	1.95	0.38	43.95
5	coefbm	23.57	6.53	0.06	0.15	59.91	12.00	0.08	0.16	71.91	71.98
6	coefqua	25.08	5.77	0.00	0.57	63.74	10.60	0.01	0.65	74.34	74.34
7	cout_sinistre	0.02	0.00	5.85	7.83	0.04	0.00	7.13	8.86	0.05	7.18
8	ctnet	1.75	1.45	17.76	22.21	4.46	2.66	21.64	25.12	7.12	28.76
9	ctsr	0.77	0.75	0.15	31.85	1.97	1.39	0.18	36.03	3.35	3.53
10	durée_contrat	4.69	0.19	5.06	20.87	11.91	0.35	6.17	23.61	12.27	18.43
11	frmalus	23.67	5.96	0.01	0.09	60.16	10.96	0.01	0.10	71.12	71.13
12	nbre_sin	0.44	0.07	25.30	14.21	1.11	0.12	30.82	16.08	1.23	32.05
		100.01	100.00	99.98	100.00	254.17	183.80	121.82	113.12	437.99	559.79

QLT14
13.25
97.80
97.94
45.90
72.15
74.99
16.04
53.88
39.57
42.04
71.24
48.13
672.93

*** Analyse auto ***

COORDONNEES DES VARIABLES

Obs.	_NAME_	AXE1	AXE2	AXE3	AXE4
1	Sex	0.09533	0.10837	0.33305	0.02718
2	age_cond	0.50269	0.84748	-0.07083	-0.04569
3	an_cond	0.49633	0.85099	-0.07835	-0.05275
4	an_route	0.02524	0.05594	0.66013	-0.13963
5	coefbm	-0.77401	0.34637	0.02798	-0.04058
6	coefqua	-0.79835	0.32561	0.00780	-0.08061
7	cout_sinistre	-0.02072	0.00688	-0.26706	0.29766
8	ctnet	-0.21115	0.16317	0.46519	0.50123
9	ctsr	-0.14025	0.11777	0.04240	0.60027
10	durée_contrat	0.34513	-0.05957	0.24835	0.48588
11	frmalus	0.77564	-0.33105	0.01086	0.03230
12	nbre_sin	-0.10532	0.03462	-0.55515	0.40102
13	*	1.00000	1.00000	1.00000	1.00000
14	*	-1.00000	-1.00000	-1.00000	-1.00000

*** Analyse auto ***

CONTRIBUTIONS (> 5%) ET QUALITE DE LA REPRESENTATION DES POINTS

Obs.	police	CTR1	CTR2	CTR3	CTR4	CO21	CO22	CO23	CO24	QLT12	QLT13	QLT14
1	012457366	1.07	5.94	0.09	0.05	.	.	.	.	.	.	.
2	012457366	1.07	5.94	0.09	0.05	.	.	.	.	.	.	.
3	012457366	1.07	5.94	0.09	0.05	.	.	.	.	.	.	.
4	012457366	1.07	5.94	0.09	0.05	.	.	.	.	.	.	.
5	012457366	1.07	5.94	0.09	0.05	.	.	.	.	.	.	.
6	012457366	1.07	5.94	0.09	0.05	.	.	.	.	.	.	.
7	012457366	1.07	5.94	0.09	0.05	.	.	.	.	.	.	.
8	012457366	1.07	5.94	0.09	0.05	.	.	.	.	.	.	.
9	012457366	1.07	5.94	0.09	0.05	.	.	.	.	.	.	.
		9.63	53.46	0.81	0.45	0	0	0	0	0	0	0

Annexe 3 : Détails ACM de 2008 à 2010

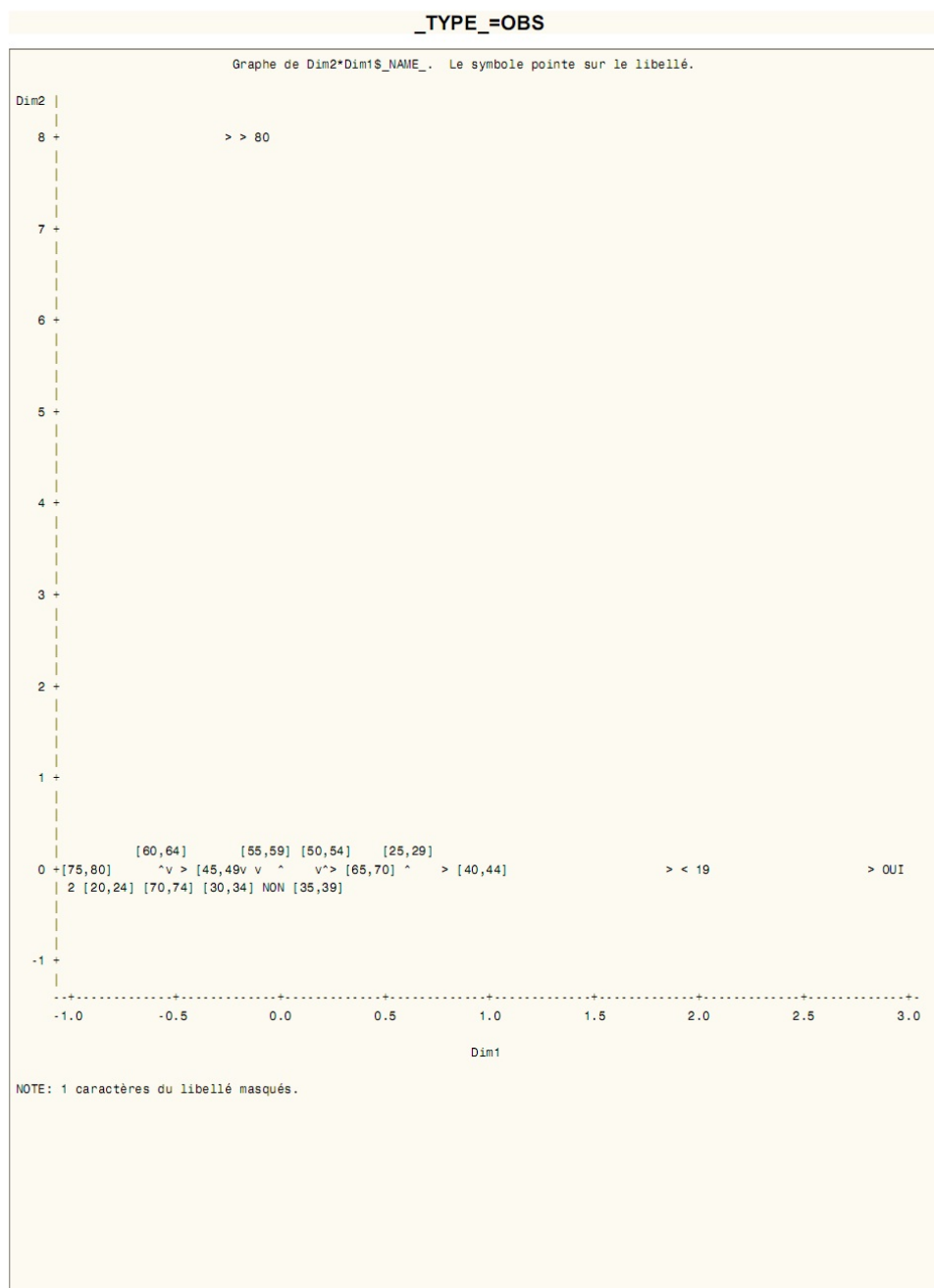


FIGURE 31 – Le résultat de l'ACM en 2008

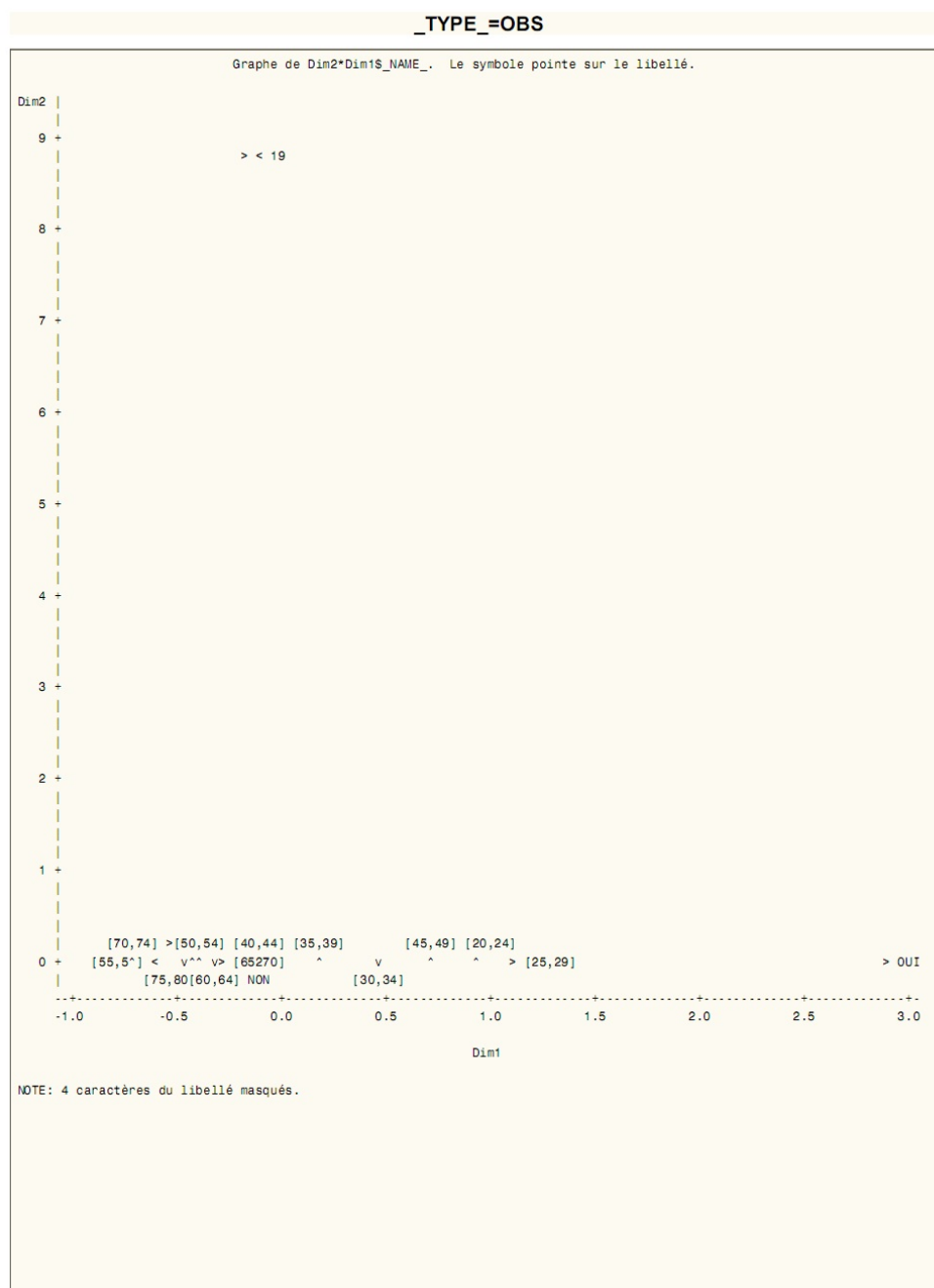


FIGURE 32 – Le résultat de l'ACM en 2009

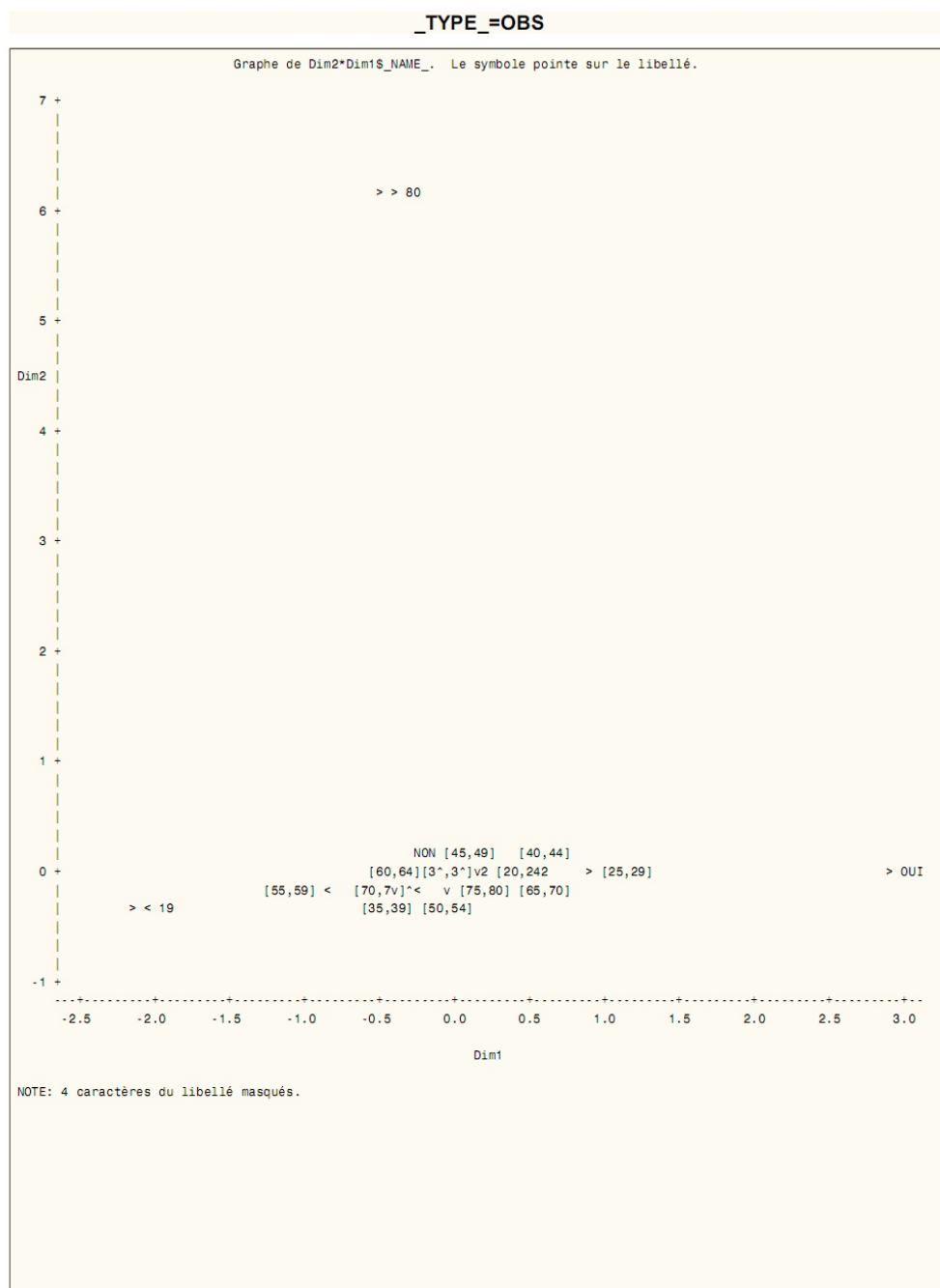


FIGURE 33 – Le résultat de l'ACM en 2010

Annexe 4 : Les tables de données de Swiss Life

Age	Zone	RC-TsAcc	Vol-Inc	BG
[20,24]	Est	3721	2753	2741
[20,24]	I.D.F	1678	1229	1226
[20,24]	Nord	8735	6659	6636
[20,24]	Ouest	2229	1741	1735
[20,24]	R.A.A	3548	2524	2518
[20,24]	Sud	5189	3046	3035
[20,24]	Sud Ouest	3819	2597	2586
[25,29]	Est	5537	4410	4392
[25,29]	I.D.F	3453	2738	2734
[25,29]	Nord	12460	10375	10351
[25,29]	Ouest	3190	2648	2643
[25,29]	R.A.A	5372	4164	4150
[25,29]	Sud	7518	5113	5097
[25,29]	Sud Ouest	6279	4778	4751
[30,34]	Est	7008	5751	5742
[30,34]	I.D.F	4836	3838	3835
[30,34]	Nord	15655	13354	13339
[30,34]	Ouest	4290	3699	3693
[30,34]	R.A.A	7029	5680	5666
[30,34]	Sud	10349	7245	7231
[30,34]	Sud Ouest	8415	6513	6483
[35,39]	Est	9604	8032	8028
[35,39]	I.D.F	6721	5485	5489
[35,39]	Nord	21287	18541	18542
[35,39]	Ouest	5836	5026	5024
[35,39]	R.A.A	10081	8413	8393
[35,39]	Sud	14637	10733	10723
[35,39]	Sud Ouest	11631	9416	9392
[40,44]	Est	11324	9594	9609
[40,44]	I.D.F	8210	6731	6725
[40,44]	Nord	23018	19995	19999
[40,44]	Ouest	6771	5902	5899
[40,44]	R.A.A	11730	9855	9852
[40,44]	Sud	18667	13738	13736
[40,44]	Sud Ouest	13411	10930	10946
[45,49]	Est	11331	9616	9616
[45,49]	I.D.F	8275	6855	6848
[45,49]	Nord	23049	19999	20022
[45,49]	Ouest	6885	5953	5957
[45,49]	R.A.A	11470	9670	9658
[45,49]	Sud	19063	14245	14232
[45,49]	Sud Ouest	13904	11258	11248
[50,54]	Est	9974	8525	8526
[50,54]	I.D.F	8009	6596	6596
[50,54]	Nord	21458	19012	19020
[50,54]	Ouest	5932	5205	5208
[50,54]	R.A.A	10539	8804	8795
[50,54]	Sud	17656	13183	13179
[50,54]	Sud Ouest	13055	10856	10854
[55,59]	Est	10205	8844	8862
[55,59]	I.D.F	7677	6346	6342
[55,59]	Nord	20516	18313	18344
[55,59]	Ouest	5983	5223	5239
[55,59]	R.A.A	9611	8135	8140
[55,59]	Sud	15941	11951	11969
[55,59]	Sud Ouest	12875	10817	10821
[60,64]	Est	8810	7749	7757
[60,64]	I.D.F	6466	5494	5506
[60,64]	Nord	17448	15693	15725
[60,64]	Ouest	5627	5026	5047
[60,64]	R.A.A	9207	7825	7840
[60,64]	Sud	15643	12084	12092
[60,64]	Sud Ouest	11901	10094	10117
[65,70]	Est	5239	4638	4661
[65,70]	I.D.F	4183	3482	3492
[65,70]	Nord	9907	9046	9093
[65,70]	Ouest	3715	3302	3319
[65,70]	R.A.A	5850	5082	5085
[65,70]	Sud	10847	8603	8615
[65,70]	Sud Ouest	7906	6789	6830
[70,74]	Est	4555	4047	4078
[70,74]	I.D.F	3214	2734	2745
[70,74]	Nord	8249	7681	7713
[70,74]	Ouest	2787	2537	2559
[70,74]	R.A.A	4107	3602	3609
[70,74]	Sud	7981	6402	6416
[70,74]	Sud Ouest	6237	5380	5399
[75,80]	Est	3195	2860	2874
[75,80]	I.D.F	2053	1812	1824
[75,80]	Nord	5888	5487	5514
[75,80]	Ouest	1908	1723	1736
[75,80]	R.A.A	2950	2634	2646
[75,80]	Sud	5159	4153	4161
[75,80]	Sud Ouest	4185	3636	3648
> 80 ans	Est	2359	2153	2160
> 80 ans	I.D.F	1872	1686	1676
> 80 ans	Nord	4163	3896	3918
> 80 ans	Ouest	1567	1402	1421
> 80 ans	R.A.A	2274	1972	1975
> 80 ans	Sud	3653	3009	3031
> 80 ans	Sud Ouest	3499	3032	3063
<= 19 ans	Est	731	463	463
<= 19 ans	I.D.F	215	132	136
<= 19 ans	Nord	1508	980	975
<= 19 ans	Ouest	487	347	346
<= 19 ans	R.A.A	723	427	426
<= 19 ans	Sud	1106	518	517
<= 19 ans	Sud Ouest	914	556	554

FIGURE 34 – Le nombre des années risques pendant les 4 années 2008 à 2011

Age	Zone	RC-TsAcc	Vol-Inc	BG
[20,24]	Est	1480425	208129	228947
[20,24]	I.D.F	767231	141382	83221
[20,24]	Nord	3542042	575840	492914
[20,24]	Ouest	789399	121741	130915
[20,24]	R.A.A	1457923	225694	192174
[20,24]	Sud	2223371	329551	225165
[20,24]	Sud Ouest	1498861	209019	208009
[25,29]	Est	1842464	272108	309262
[25,29]	I.D.F	1395779	280660	158087
[25,29]	Nord	4221495	743592	672496
[25,29]	Ouest	950350	155582	174843
[25,29]	R.A.A	1860414	307680	277991
[25,29]	Sud	2691938	460072	320991
[25,29]	Sud Ouest	2000751	306187	321785
[30,34]	Est	1982270	311972	349980
[30,34]	I.D.F	1791201	379032	211015
[30,34]	Nord	4589717	830690	755857
[30,34]	Ouest	1112230	186947	212821
[30,34]	R.A.A	2173561	376359	330564
[30,34]	Sud	3228880	588740	409898
[30,34]	Sud Ouest	2321158	373901	393209
[35,39]	Est	2365387	390162	430981
[35,39]	I.D.F	2325271	507994	277177
[35,39]	Nord	5581025	1039092	941576
[35,39]	Ouest	1343806	232672	259022
[35,39]	R.A.A	2746895	510609	436371
[35,39]	Sud	4212615	793103	541327
[35,39]	Sud Ouest	2836633	473949	499902
[40,44]	Est	2633560	437047	481504
[40,44]	I.D.F	2686398	591048	320271
[40,44]	Nord	5600931	1050563	942974
[40,44]	Ouest	1435100	252647	279130
[40,44]	R.A.A	3091975	586780	490677
[40,44]	Sud	5032375	953406	637131
[40,44]	Sud Ouest	3097657	531509	540815
[45,49]	Est	2481962	418567	460539
[45,49]	I.D.F	2648603	592269	313829
[45,49]	Nord	5311104	986647	883659
[45,49]	Ouest	1360482	237698	259222
[45,49]	R.A.A	2926090	545529	452208
[45,49]	Sud	5009557	971982	636092
[45,49]	Sud Ouest	3033280	516463	524502
[50,54]	Est	2118379	347086	370204
[50,54]	I.D.F	2474751	559998	282604
[50,54]	Nord	4863748	903586	785165
[50,54]	Ouest	1188362	202492	214262
[50,54]	R.A.A	2633015	482515	396838
[50,54]	Sud	4490424	852551	556493
[50,54]	Sud Ouest	2784813	478095	467459
[55,59]	Est	2110664	338058	351548
[55,59]	I.D.F	2353730	523326	258871
[55,59]	Nord	4661255	849904	707729
[55,59]	Ouest	1179060	194191	198576
[55,59]	R.A.A	2323591	422639	333973
[55,59]	Sud	3890979	746103	465346
[55,59]	Sud Ouest	2716417	457468	436011
[60,64]	Est	1757054	260692	268293
[60,64]	I.D.F	1920949	410187	193351
[60,64]	Nord	3819759	627457	507410
[60,64]	Ouest	1034171	159155	163280
[60,64]	R.A.A	2101140	351737	269790
[60,64]	Sud	3698415	674425	394955
[60,64]	Sud Ouest	2400928	368208	356220
[65,70]	Est	1006912	139950	136305
[65,70]	I.D.F	1168806	241899	99047
[65,70]	Nord	2054388	321381	243345
[65,70]	Ouest	673470	98569	92496
[65,70]	R.A.A	1310281	211795	150873
[65,70]	Sud	2518474	448429	236889
[65,70]	Sud Ouest	1577413	238466	211151
[70,74]	Est	854356	113099	103511
[70,74]	I.D.F	853209	166993	64039
[70,74]	Nord	1706039	248210	176615
[70,74]	Ouest	486043	65901	58873
[70,74]	R.A.A	900696	137677	87209
[70,74]	Sud	1747510	306514	146583
[70,74]	Sud Ouest	1200725	173446	144772
[75,80]	Est	625084	85074	66380
[75,80]	I.D.F	555081	113117	39051
[75,80]	Nord	1196443	171148	112888
[75,80]	Ouest	351591	50635	38942
[75,80]	R.A.A	651255	103202	57781
[75,80]	Sud	1102768	187515	82819
[75,80]	Sud Ouest	760339	111047	80710
> 80 ans	Est	454061	60852	44472
> 80 ans	I.D.F	570771	122796	35782
> 80 ans	Nord	896705	125719	77429
> 80 ans	Ouest	291285	40637	27877
> 80 ans	R.A.A	494309	77369	37910
> 80 ans	Sud	824576	148889	57017
> 80 ans	Sud Ouest	655355	95180	63001
<= 19 ans	Est	307217	39767	40227
<= 19 ans	I.D.F	109439	16673	10511
<= 19 ans	Nord	647959	96480	78106
<= 19 ans	Ouest	186108	29056	28546
<= 19 ans	R.A.A	314883	43246	35086
<= 19 ans	Sud	523549	65447	42057
<= 19 ans	Sud Ouest	398569	51058	47277

FIGURE 35 – Les primes émises pendant les 4 années 2008 à 2011

Age	Zone	RC-TsAcc	Vol-Inc	BG
[20,24]	Est	1259963	122424	156247
[20,24]	I.D.F	859121	94797	60011
[20,24]	Nord	5207114	370859	378492
[20,24]	Ouest	658158	59199	81582
[20,24]	R.A.A	2001408	130785	104473
[20,24]	Sud	4241482	229086	100269
[20,24]	Sud Ouest	3622729	174045	113293
[25,29]	Est	1224770	177254	245689
[25,29]	I.D.F	1113475	280378	131001
[25,29]	Nord	3755943	467758	594324
[25,29]	Ouest	835322	81793	123046
[25,29]	R.A.A	2746976	212197	173338
[25,29]	Sud	2429868	471173	177109
[25,29]	Sud Ouest	1987061	277883	205671
[30,34]	Est	2666227	143253	283798
[30,34]	I.D.F	1220259	198656	181448
[30,34]	Nord	3426348	500925	732030
[30,34]	Ouest	863155	95048	175045
[30,34]	R.A.A	2196395	226714	235344
[30,34]	Sud	3118110	460976	237747
[30,34]	Sud Ouest	1818984	254560	300122
[35,39]	Est	2022174	220225	401499
[35,39]	I.D.F	1547313	402632	235382
[35,39]	Nord	5084346	578718	941006
[35,39]	Ouest	2205522	82120	216619
[35,39]	R.A.A	2967339	333092	310536
[35,39]	Sud	3704201	510222	336612
[35,39]	Sud Ouest	2623576	309172	321855
[40,44]	Est	3013926	164566	442011
[40,44]	I.D.F	2333164	344467	280735
[40,44]	Nord	5531370	771646	961907
[40,44]	Ouest	1394629	122096	233697
[40,44]	R.A.A	2852366	216057	376359
[40,44]	Sud	5979803	703855	382622
[40,44]	Sud Ouest	3086765	340986	398305
[45,49]	Est	2354169	330843	399993
[45,49]	I.D.F	2298913	401454	272340
[45,49]	Nord	6984375	335675	886508
[45,49]	Ouest	1416157	96709	237666
[45,49]	R.A.A	2660203	339186	337682
[45,49]	Sud	5321619	561451	382793
[45,49]	Sud Ouest	3336028	337855	389888
[50,54]	Est	2377970	173805	323091

[55,59]	Est	1829072	155042	326419
[55,59]	I.D.F	1880688	237407	224716
[55,59]	Nord	3905123	377908	626544
[55,59]	Ouest	852851	43188	167884
[55,59]	R.A.A	2238904	183984	231541
[55,59]	Sud	4420306	299054	285600
[55,59]	Sud Ouest	2520540	254985	338439
[60,64]	Est	1421852	100429	238704
[60,64]	I.D.F	1432502	128002	189325
[60,64]	Nord	2688029	313484	457857
[60,64]	Ouest	735258	46932	147451
[60,64]	R.A.A	1741203	148903	176716
[60,64]	Sud	4293912	332404	264689
[60,64]	Sud Ouest	2092834	227945	251492
[65,70]	Est	975305	42572	125625
[65,70]	I.D.F	911332	141559	98857
[65,70]	Nord	1736214	135408	255124
[65,70]	Ouest	1092772	36421	82356
[65,70]	R.A.A	1100089	69740	116881
[65,70]	Sud	3918072	170276	159458
[65,70]	Sud Ouest	1346115	109714	162537
[70,74]	Est	705084	37659	75720
[70,74]	I.D.F	649154	41337	50979
[70,74]	Nord	1260545	68517	176090
[70,74]	Ouest	370743	10155	40276
[70,74]	R.A.A	767568	68339	64511
[70,74]	Sud	1726230	92155	94296
[70,74]	Sud Ouest	1884720	67332	99108
[75,80]	Est	612121	35874	39070
[75,80]	I.D.F	603485	12003	29819
[75,80]	Nord	1703991	39703	88059
[75,80]	Ouest	425802	14092	21130
[75,80]	R.A.A	732084	10712	33373
[75,80]	Sud	1877155	44559	47064
[75,80]	Sud Ouest	934140	52290	49648
> 80 ans	Est	461503	64	19093
> 80 ans	I.D.F	353098	23826	15792
> 80 ans	Nord	1810315	38417	42153
> 80 ans	Ouest	233233	8495	10219
> 80 ans	R.A.A	395063	19950	18905
> 80 ans	Sud	867487	39597	23157
> 80 ans	Sud Ouest	637869	27799	29742
<= 19 ans	Est	716750	10749	26988

FIGURE 36 – Les charges sinistres complètes pendant les 4 années 2008 à 2011

Age	Zone	RC-TsAcc	Vol-Inc	BG
[20,24]	Est	1470472	123582	156247
[20,24]	I.D.F	745903	95694	60011
[20,24]	Nord	3630811	374367	378492
[20,24]	Ouest	768121	59759	81582
[20,24]	R.A.A	1688935	132023	104473
[20,24]	Sud	2574683	231253	100269
[20,24]	Sud Ouest	1502199	175691	113293
[25,29]	Est	1429400	178931	245689
[25,29]	I.D.F	1299510	283030	131001
[25,29]	Nord	3799247	469549	594324
[25,29]	Ouest	936398	82567	123046
[25,29]	R.A.A	1951471	214204	173338
[25,29]	Sud	2735458	475630	177109
[25,29]	Sud Ouest	2086865	280512	205671
[30,34]	Est	1656916	144608	283798
[30,34]	I.D.F	1424135	200535	181448
[30,34]	Nord	3963797	505663	732030
[30,34]	Ouest	993091	95948	175045
[30,34]	R.A.A	2334936	228858	235344
[30,34]	Sud	3389878	465337	237747
[30,34]	Sud Ouest	2067652	256967	300122
[35,39]	Est	2357859	222309	401499
[35,39]	I.D.F	1805832	406441	235382
[35,39]	Nord	5625573	560931	941006
[35,39]	Ouest	1261088	82897	216619
[35,39]	R.A.A	3184199	336243	310536
[35,39]	Sud	4294463	511883	336612
[35,39]	Sud Ouest	3034818	312096	321855
[40,44]	Est	2834175	166122	442011
[40,44]	I.D.F	2587898	347726	280735
[40,44]	Nord	5799453	777134	961907
[40,44]	Ouest	1407517	123251	233697
[40,44]	R.A.A	3326833	218100	376359
[40,44]	Sud	6273915	687473	382622
[40,44]	Sud Ouest	3431644	344212	398305
[45,49]	Est	2684797	333972	399993
[45,49]	I.D.F	2486439	334049	272340
[45,49]	Nord	5428963	338850	886508
[45,49]	Ouest	1516996	97624	237666
[45,49]	R.A.A	3104659	342395	337682
[45,49]	Sud	6052373	566762	382793
[45,49]	Sud Ouest	3550674	341051	389888
[50,54]	Est	2396596	175449	323091
[50,54]	I.D.F	2613588	185889	251817
[50,54]	Nord	4670150	457824	615345
[50,54]	Ouest	1252353	134170	176329
[50,54]	R.A.A	3238355	248447	277537
[50,54]	Sud	5431380	514207	343062
[50,54]	Sud Ouest	3233197	331421	339867
[55,59]	Est	2057694	156509	326419
[55,59]	I.D.F	2111299	237774	224716
[55,59]	Nord	4414643	336421	626544
[55,59]	Ouest	995343	43596	167884
[55,59]	R.A.A	2552189	185725	231541
[55,59]	Sud	4482371	301883	285600
[55,59]	Sud Ouest	2712435	257397	338439
[60,64]	Est	1659410	101379	238704
[60,64]	I.D.F	1671838	116500	189325
[60,64]	Nord	3137134	316449	457857
[60,64]	Ouest	858102	47376	147451
[60,64]	R.A.A	2020597	150311	176716
[60,64]	Sud	4428778	335548	264689
[60,64]	Sud Ouest	2230876	230101	251492
[65,70]	Est	1072123	42975	125625
[65,70]	I.D.F	1063594	142898	98857
[65,70]	Nord	2010803	136689	255124
[65,70]	Ouest	738022	36766	82356
[65,70]	R.A.A	1283887	70400	116881
[65,70]	Sud	3198340	171887	159458
[65,70]	Sud Ouest	1568847	110752	162537
[70,74]	Est	822887	38015	75720
[70,74]	I.D.F	757612	41728	50979
[70,74]	Nord	1471152	69165	176090
[70,74]	Ouest	432686	10251	40276
[70,74]	R.A.A	895811	68985	64511
[70,74]	Sud	2014642	93026	94296
[70,74]	Sud Ouest	1527843	67969	99108
[75,80]	Est	679380	36214	39070
[75,80]	I.D.F	529252	12117	29819
[75,80]	Nord	1336088	40078	88059
[75,80]	Ouest	496944	14225	21130
[75,80]	R.A.A	815768	10814	33373
[75,80]	Sud	1432459	44980	47064
[75,80]	Sud Ouest	858051	52785	49648
> 80 ans	Est	538609	65	19093
> 80 ans	I.D.F	412093	24052	15792
> 80 ans	Nord	1004053	38780	42153
> 80 ans	Ouest	272201	8575	10219
> 80 ans	R.A.A	461069	20139	18905
> 80 ans	Sud	1011209	39971	23157
> 80 ans	Sud Ouest	744441	28062	29742
<= 19 ans	Est	463037	10850	26988
<= 19 ans	I.D.F	152700	6290	5916
<= 19 ans	Nord	970537	67613	45163
<= 19 ans	Ouest	418730	17229	15913
<= 19 ans	R.A.A	648523	10291	17766
<= 19 ans	Sud	847387	24047	27949
<= 19 ans	Sud Ouest	514397	18874	21502

FIGURE 37 – Les charges sinistres écrêtées pendant les 4 années 2008 à 2011