Olympio, Anani Ayodéle

Thesis

Contributions au provisionnement en assurance de personnes et à la gestion des risques

Contributions au provisionnement en assurance de personnes et à la gestion des risques

Anani Ayodele Olympio

N°d'ordre NNT : 2019LYSE1148

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale ED 486
SCIENCES ÉCONOMIQUES ET DE GESTION

Spécialité de doctorat : Science de Gestion
Discipline : Actuariat

Soutenue publiquement le 25 septembre 2019, par :
Anani Ayodélé OLYMPIO

Contributions au provisionnement en assurance de personnes et à la gestion des risques

---

Devant le jury composé de :

Nom, prénom grade/qualité établissement/entreprise	
Frédéric Planchet, Professeur à l'Université Claude Bernard Lyon 1	Président
Michel Bera, Professeur au Conservatoire National des Arts et Métiers	Rapporteur
Mercè Claramunt, Professeur à University of Barcelona	Rapporteure
Frédéric Planchet, Professeur à l'Université Claude Bernard Lyon 1	Examinateur
Annamaria Olivieri, Professeur à l'Università di Parma	Examinatrice
Stéphane Loisel, Professeur à l'Université Claude Bernard Lyon 1	Directeur de thèse

N°d'ordre NNT : 2019LYSE1148

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale ED 486
SCIENCES ÉCONOMIQUES ET DE GESTION

Spécialité de doctorat : Science de Gestion
Discipline : Actuariat

Soutenue publiquement le 25 septembre 2019, par :
Anani Ayodélé OLYMPIO

Contributions to insurance reserving and Risk Management

Devant le jury composé de :

Nom, prénom grade/qualité établissement/entreprise	
Frédéric Planchet, Professeur à l'Université Claude Bernard Lyon 1	Président
Michel Bera, Professeur au Conservatoire National des Arts et Métiers	Rapporteur
Mercè Claramunt, Professeur à University of Barcelona	Rapporteure
Frédéric Planchet, Professeur à l'Université Claude Bernard Lyon 1	Examinateur
Annamaria Olivieri, Professeur à l'Università di Parma	Examinatrice
Stéphane Loisel, Professeur à l'Université Claude Bernard Lyon 1	Directeur de thèse

REMERCIEMENTS

Je formule mes sincères remerciements à mon directeur de thèse Stéphane Loisel, sans qui ce projet n'aurait pas été possible. D'abord, il m'a progressivement guidé avec pédagogie dans le monde de la recherche universitaire. Ensuite, il m'a permis d'apprendre de son expérience de chercheur, de son expertise scientifique et technique. Enfin, il m'a transmis sa passion pour la recherche théorique et opérationnelle, ainsi que le souci du travail méticuleux et rigoureux. Je souhaite lui exprimer toute ma gratitude, pour son soutien et sa disponibilité lors de mes périodes de doutes et de questionnements. Ses précieux conseils, sa patience et sa bonne humeur m'ont aidé à franchir bien des difficultés durant mon parcours de doctorant. J'ai beaucoup appris à ses côtés au cours des nombreux projets scientifiques, en lien ou non avec cette thèse, auxquels il m'a associé. J'espère poursuivre mes travaux de recherche en cours et futurs avec la même passion qu'il a su me transmettre. J'associe à mes remerciements sa magnifique famille, son épouse Anne Eyraud-Loisel et leurs deux merveilleux enfants Maud et Arthur.

Je tiens à exprimer ma reconnaissance à Mercè Claramunt et Michel Bera qui ont accepté de rapporter sur mes travaux. Un grand merci à eux pour le temps consacré à lire ce manuscrit et leurs remarques pertinentes qui m'ont permis d'améliorer la qualité de l'ouvrage. Je suis également très honoré que Annamaria Olivieri et Frédéric Planchet aient accepté d'être membre du jury de soutenance de ma thèse.

Je souhaite également remercier particulièrement la direction générale de CNP Assurances. Tout d'abord, je remercie Antoine Lissowski, Directeur général de CNP Assurances, pour l'intérêt qu'il a porté à mes thématiques de recherche dès la genèse de ce projet. Je remercie également Thomas Béhar, Directeur Financier Groupe de CNP Assurances, merci de m'avoir encouragé à entreprendre cette thèse et de m'avoir suivi sans relâche comme tuteur en entreprise pendant toutes ces années de recherche.

Je souhaite remercier tout particulièrement David Ingram. Son rôle fut très important au début de ma thèse. Je le remercie pour le temps qu'il m'a consacré à New York afin de m'expliquer ces différents travaux sur l'attitude face au risque. Cette rencontre m'a aidé à la compréhension de la méthodologie et des données américaines utilisées pour ses études. Merci d'avoir fortement contribué à mon initiation à la théorie socioculturelle face au risque.

J'ai également eu l'opportunité de travailler avec certains co-auteurs, sur des sujets de recherche dont certains résultats sont présentés dans ce manuscrit. Je veux remercier : Yacine Gaïd (pour sa contribution au chapitre 1), Jérémie Zozime (pour sa contribution au chapitre 2), Camille Gutknecht (pour sa contribution au chapitre 3) et enfin Denis Clot (pour sa contribution au chapitre 4).

Mes remerciements s'adressent à Gilles Kué Gaba, que j'ai entraîné avec moi dans le monde de la recherche malgré nos emplois du temps professionnels déjà bien chargés. C'était une bonne idée, car toutes les heures passées à travailler la matière statistique et à confronter nos points de vue et expériences de praticiens m'ont permis d'avancer sur des bases solides. Merci à Vohirana Rananaiv, à Sophie Michon et à Florence Picard pour vos encouragements. Une pensée toute particulière à Nicolas Leboisne pour ses encouragements, Claude Lefèvre pour la richesse de ses conseils, à Christian Robert pour ses retours d'expérience en tant que chercheur et à Olivier Lopez pour la clarté de ses explications sur la construction de l'algorithme Tree base censored.

Je tiens également à remercier l'ensemble de l'équipe de recherche et des membres du laboratoire SAF de l'ISFA pour m'avoir toujours très bien accueilli et de faire du laboratoire un lieu où il est plaisant de travailler. Toutefois au laboratoire SAF, sachez que les places sont chères. J'y ai malgré tout trouvé une place provisoire lors de mes nombreux passages à l'ISFA. Un grand merci donc à mes bienveillants

hôtes, à Yahia Salhi et à Xavier Milhaud. J'ai apprécié leur amitié et générosité, toujours prêts à apporter leur aide et soutien lors de mes séjours dans leur bureau. Je remercie vivement Pierre Thérond pour nos échanges sur ses sujets de prédilection, les normes règlementaires européennes. Ses réflexions sur les biais liés à l'interprétation et aux traductions des textes règlementaires (par exemple de l'anglais vers le français) ont été de précieux éclairages. Je ne manquerai pas d'adresser un petit clin d'œil aux doctorants et anciens doctorants. Merci à Julien Védani pour ses conseils de vieux routards et à Claire Mouminoux pour ses encouragements. Merci à Pierre Montesinos et à Sarah Bensalem pour toutes les fois où ils ont pris le temps d'imprimer mes supports de présentation. Vous êtes nombreux et je n'oublierai jamais vos sourires et vos gestes d'amitié. A tous, ne perdez rien de la bonne ambiance qui règne à l'ISFA.

Un grand merci à tous mes collègues chez CNP Assurances avec qui j'ai eu l'occasion d'échanger sur mes projets de recherche. Je tiens à remercier plus particulièrement Josselin Kalifa et son équipe, mais aussi Jean-Baptiste Nessi, Nicolas Legrand, Magali Noé, Sophie Wittmer et Stéphanie Cariou-Hellec. Un grand merci également à Cédric Laurans, à Cédric Atchama, à Babacar Sow et à Baptiste Dieltiens. Merci à mes anciennes équipes de R&D Data'Lab, Réassurance, Méthodes et Normes actuarielles et Outils actuariels.

Je remercie très chaleureusement ma famille, mes parents Nitou et Angèle Olympio qui m'ont donné la vie. Vous m'avez appris la valeur du travail et de la persévérance, de l'effort et de la patience… J'ose croire que ce travail fera aussi votre fierté. A mes trois frères et leur famille, Orphéo, Josélito et Roméo Olympio, quel que soit la distance qui nous sépare, vous n'êtes jamais loin de mon cœur.

Mes sincères remerciements à mes parents de cœur de Marseille, Nelly et Richard Olympio. Ils m'ont adopté depuis mes premières années dans la cité Phocéenne. D'une très grande hospitalité, ils n'ont cessé d'être une source d'encouragement et d'inspiration pour moi. Merci également à mes quatre cousins : Yohan, Neil (un autre actuaire de la famille), Loïs et Edwin. Toujours à Marseille, j'ai eu la joie de rencontrer Fernando Mensah et sa famille. Je tiens à vous remercier pour votre amitié et vos encouragements. Un grand merci également à Jean-Marc Mensah.

Un merci particulier à mes beaux-parents Daniel et Viviane Landes, pour leur soutien et leur grande disponibilité. J'ai également une pensée affectueuse pour Jonathan Landes, Séphora et Sébastien Brunel, Ruth et Christophe Papacalodouca.

Je n'oublie pas mes plus proches amis Cyril et Glenda Petit, Franklin et Céline Nwachukwu, Maryse Petershmit, Aurélie Boucaca, Gueneth Dumond, Thierry et Sarah Jeanmaire, Guillaume et Aurélie Levy, Christian et Isabelle Kabala, Sérilo et Elisabeth Looky, Servais et Sophie César... Un très grand merci à Juliette Mathieu pour sa relecture et ses remarques sur ce manuscrit. A tous nos jeunes gens et jeunes filles du groupe de jeunesse S-PASS-J que mon épouse et moi avons le privilège de servir, à vous tous, amis et frères de la famille EBE…

Mes derniers mots de remerciements vont à mon épouse Débora, fidèle compagne et mère dévouée. Je sais que sans son soutien et ses sacrifices je n'aurai pas pu mener ce travail à sa fin. Le fruit de ce travail est aussi le sien. Un grand merci à nos trois grands Jean-Samuel, Timothée et Elnathane… Je suis fier de vous et de votre courage. Merci d'avoir accepté que votre papa soit souvent loin de la maison pour ses travaux de recherche. J'espère que vous êtes, vous aussi, fiers de lui et que cela vous servira de point d'encrage pour atteindre vos propres objectifs afin de réaliser vos rêves. Je vous lance un seul défi maintenant, faites mieux que votre papa.

A mon épouse Débora et à nos trois fils Jean-Samuel, Timothée et Elnathane…

« Tout est possible à celui qui croit. » Marc 9 : 23

RESUME

Dans le secteur de l'assurance, les dernières évolutions de la règlementation et des normes comptables vont dans le sens de la standardisation de la gestion des risques au sein des organismes. Dans ce contexte, l'objectif principal de ma thèse est de proposer différentes méthodologies d'évaluation et d'analyse des risques dans ce secteur. La première partie de ce manuscrit traite de la problématique de provisionnement individuel en non-vie. Je propose des adaptions d'algorithmes d'apprentissage automatique ensemblistes et de certaines métriques de performance pour l'estimation des durées des sinistres ainsi que des charges sinistres ultimes en présence de données censurées à droite. L'application de ces méthodes à des données réelles de contrats de prêts ou de contrats de prévoyance collective conduit à des estimations plus performantes et plus robustes des paramètres considérés. La deuxième partie présente une approche d'estimation de choc à un an sur des paramètres spécifiques à l'entité (*Undertaking Specific Parameters*) du module « *santé assimilable à la vie* » du pilier 1 de la formule standard de la norme Solvabilité II. L'utilisation de la crédibilité américaine (ou crédibilité à variation limitée) permet la prise en compte partielle des contraintes de disponibilité des données d'expérience (volumétrie et profondeur d'historique) lors du calibrage des chocs. A titre d'illustration, j'ai appliqué cette approche aux risques d'incidence et de maintien (ou de rétablissement) des garanties d'incapacité et d'invalidité en arrêt de travail d'un portefeuille de contrats de prêts. Les résultats obtenus montrent des baisses significatives des besoins de capitaux de solvabilité requis (SCR) du risque de souscription par rapport à la formule standard. La troisième partie est une étude descriptive des calculs de la formule standard pour l'évaluation du besoin de fonds propres économiques du risque de dépendance. Elle permet de mettre en évidence les insuffisances de la norme et de proposer des pistes d'améliorations en vue d'une meilleure prise en compte des spécificités de ce risque. Enfin, dans la dernière partie du manuscrit, je propose une étude comparative des préférences d'attitudes face au risque dans le secteur financier, notamment la banque et l'assurance. Il s'agit d'une analyse empirique menée dans trois zones géographiques (Amérique, Europe et Afrique) afin de mesurer les liens et les différences entre les profils d'attitude face au risque et certaines variables sociodémographiques.

Mots clés : *gestion des risques, apprentissage automatique, réserves, prévoyance, crédibilité américaine, paramètre spécifique à l'entité, fonds propres économique, solvabilité II, dépendance, attitude face au risque, théorie socio-culturelle du risque.*

ABSTRACT

In the insurance sector, the latest regulatory developments and accounting standards are in line with the standardization of risk management within organizations. In this context, the main objective of my thesis is to propose different methodologies for risk evaluation and analysis in this sector. The first part of this manuscript deals with the problem of individual non-life reserving. I proposed adaptations of machine learning algorithms and some performance metrics for the estimation of the durations of the claims as well as the ultimate claims in the presence of right censored data. We applied these methods to different contracts: property insurance, consumer loans insurance and group protection. The parameters estimations are better and robustest. The second part presents a one-year shock estimation approach on entity-specific parameters (Undertaking Specific Parameters) of the life-sustaining health module of Pillar 1 of the Solvency II standard formula. The use of American credibility (or limited variation credibility) allows partial consideration of the availability constraints of data (volume and historical depth of data) when calibrating shocks. By way of illustration, I applied this approach to incidence and recovery (or non-recovery) of incapacity and disability risks. The results obtained show significant decreases in solvency capital requirements (SCR) of underwriting risk need compared to the standard formula calculation. The third part is a descriptive study of the calculations of the standard formula for economic solvency capital need of long term care risk. The main purpose is to highlight the inadequacies of the standard formula and to suggest ways of improving them in order to better take into account the specificities of this risk. Finally, in the last part of the manuscript, I proposed a comparative study of risk attitude preferences in the financial sector, including banking and insurance. This is an empirical analysis conducted in three geographical areas (America, Europe and Africa) to measure the links and differences between risk attitude profiles and sociodemographic variables.

Keywords: *risk management, machine learning, reserving, non-life insurance, US credibility, entity-specific parameter, economic capital, Solvency II, Long Term Care, risk attitude, socio-cultural theory of risk.*

Contenu

Introduction générale

INTRODUCTION GENERALE

CONTRIBUTIONS AU PROVISIONNEMENT EN ASSURANCE DE PERSONNES ET A LA GESTION DES RISQUES

La gestion des risques en assurance, et notamment en assurance de personnes, s'est standardisée avec l'avènement de la directive européenne Solvabilité II. Le contexte la rend d'autant plus indispensable, avec : la persistance des taux bas dans la zone euro, une croissance faible depuis plusieurs années, une transition sociodémographique (caractérisée notamment par un taux de fécondité très bas et un vieillissement de la population), des réactions de plus en plus conservatrices des pouvoirs publics (conduisant à une inflation règlementaire), des évolutions de l'environnement (avec des conséquences sur le climat, la pollution et les problèmes de biodiversité…), l'émergence des risques de cyber attaque et la crainte des dérives potentielles liées à l'usage de l'intelligence artificielle…

Pour être en mesure d'honorer leurs engagements, les assureurs constituent des provisions mathématiques afin de faire face à l'inversion du cycle de production spécifique à ce métier. Les assureurs mettent également en place un processus de gestion des différents risques auxquels ils sont confrontés. L'objectif principal de mes travaux de recherche présentés dans cette thèse est de proposer différentes approches et méthodes d'estimation de certains risques en matière de gestion des risques dans le secteur financier, notamment en assurance de personnes. Ainsi, ce manuscrit est composé d'une introduction générale, suivie de quatre chapitres et d'une conclusion générale. Les trois premiers chapitres abordent la gestion des risques dans le cadre du provisionnement et des calculs des capitaux requis sous la norme Solvabilité II. Le dernier chapitre est consacré à l'exploration de méthodes empiriques de mesure des préférences d'attitude face au risque suivant la théorie socioculturelle du risque.

Plus particulièrement, je propose l'un des chapitres de cette thèse une nouvelle méthode de provisionnement. Il s'agit d'adaptations algorithmiques et de métriques de performance applicables aux méthodes d'apprentissage automatique ensemblistes permettant de prédire des paramètres utilisés pour le provisionnement individuel des sinistres en assurance non vie. Nous appliquons nos approches à deux portefeuilles (de prêts et de prévoyance collective) assez différents par leur volumétrie, par les risques sous-jacents aux garanties assurées, par les variables explicatives et la période d'historique disponibles. Les paramètres de provisionnement prédits sont les durées de maintien et des charges sinistres ultimes individuelles. Les indicateurs de validation sont cohérents, plus précis et plus stables que ceux fournis par l'algorithme des arbres de classification et de régression modifié proposé par Lopez et al. (2016) [1].

Dans le cadre de Solvabilité II, je développe une méthode pour les USP (entre formule standard et modèle interne). Cette contribution fournit un cadre opérationnel de calibration d'USP du module « *Health SLT* » pour les risques d'incidence et de maintien (ou le rétablissement) des garanties d'incapacité et d'invalidité en assurance prévoyance sur des données réelles. Ce travail apporte une modélisation alternative aux chocs forfaitaires fixés de manière arbitraire par la formule standard.

Je m'intéresse à l'amélioration de la modélisation de la dépendance par rapport au modèle sous-jacent de Solvabilité II. Je présente les difficultés liées à l'application de la formule standard au calcul des fonds propres économique pour ce risque. Je soulève plus spécifiquement les problématiques de la prise en compte de la faculté de révision des primes, de la revalorisation des contrats et de l'indexation des primes, dans les calculs de la meilleure estimation des provisions.

Enfin, j'étudie les attitudes face au risque dans le secteur de l'assurance et leurs différences dans plusieurs zones géographiques.

Dans cette introduction générale, organisée autour de dix sections, je présente les concepts et outils clés utilisés dans cette thèse et repositionne mes travaux dans leur contexte. J'expose dans la première section quelques éléments de contexte de la gestion et de l'analyse des risques. Dans la seconde section, je rappelle les grands principes de la norme Solvabilité II, ainsi que le cadre d'utilisation des paramètres spécifiques à l'entité (*Undertaking Specific Parameters*, USP). La troisième section aborde la thématique de l'application des modèles à multi-états en assurance de personnes. La section quatre est consacrée aux rappels des risques et méthodes de provisionnement en assurance non-vie. La section cinq expose les techniques d'apprentissage automatique en présence de données censurées. La section six présente les composantes entrant dans le calcul de la meilleure estimation des provisions pour risque. La section sept décrit le calibrage des paramètres spécifiques pour la garantie arrêt de travail. La section huit présente la modélisation du risque de dépendance sous Solvabilité II. La section neuf est consacrée à décrire l'approche socioculturelle des attitudes face au risque, puis la section dix est une synthèse des différentes contributions et des principaux résultats.

1. Gestion et analyse des risques

L'étude théorique de la gestion des risques a débuté après la Deuxième Guerre mondiale. Selon plusieurs sources (voir Crockford, 1982 [2] ; Harrington et Niehaus, 2003 [3] ; Williams et Heins, 1995 [4]), la gestion des risques moderne remonte à la période 1955-1964. Le contenu des deux premiers livres académiques publiés par Mehr et Hedges (1963) [5] et Williams et Heins (1964) [6] portait sur la gestion des risques dits purs, ce qui excluait les risques financiers des entreprises. Parallèlement, les ingénieurs ont développé des modèles de gestion des risques technologiques. Le risque opérationnel couvre en partie les pertes technologiques et il est maintenant géré par toutes les entreprises, y compris par les institutions financières. Les ingénieurs ont également commencé à s'intéresser aux risques politiques des projets.

La gestion des risques a pendant longtemps été associée à l'utilisation de l'assurance de marché pour protéger les individus et les entreprises contre différentes pertes associées à des accidents (voir Harrington et Niehaus, 2003 [7]). En 1982, Crockford [2] écrivait: "*Operational convenience continues to dictate that pure and speculative risks should be handled by different functions within a company, even though theory may argue for them being managed as one. For practical purposes, therefore, the emphasis of risk management continues to be on pure risks*".

Des formes de gestion des risques purs, alternatives à l'assurance de marché, ont pris forme durant les années 1950 lorsque différentes protections d'assurance sont devenues très coûteuses et incomplètes. En effet, plusieurs risques d'entreprise n'étaient pas assurables ou coûtaient très chers à assurer. Avant les années 1970, les produits dérivés n'étaient pas utilisés pour couvrir des produits financiers. Ils étaient limités aux produits agricoles. L'utilisation des produits dérivés comme instruments de gestion de différents risques assurables et non assurables a débuté durant les années 1970 et s'est développée très rapidement durant les années 1980. C'est aussi durant les années 1980 que les entreprises ont commencé à considérer la gestion financière ou portefeuille des risques. La gestion financière des risques est devenue complémentaire à la gestion des risques purs pour beaucoup d'entreprises. Les institutions financières, dont les banques et les compagnies d'assurances, ont intensifié leurs activités de gestion des risques de marché et de crédit durant les années 1980. Les activités de gestion du risque opérationnel et du risque de liquidité sont apparues durant les années 1990 (voir Dionne, 2013 [8])

La réglementation internationale des risques a aussi débuté durant les années 1990 et les entreprises financières ont développé des modèles de gestion des risques internes et des formules de calcul du capital pour se protéger contre les risques non anticipés et pour réduire le capital réglementaire. C'est également durant ces années que la gouvernance de la gestion des risques est devenue essentielle,

que la gestion des risques intégrée a été introduite et que les premiers postes de gestionnaire des risques ont été créés.

Suite à différents scandales et faillites associés à une mauvaise gestion des risques, la réglementation Sarbanes-Oxley a été instaurée aux États-Unis, en 2002, afin d'introduire des règles de gouvernance des entreprises. Des bourses, dont le NYSE en 2002 ont aussi ajouté des règles de gouvernance sur la gestion des risques pour les entreprises inscrites à ces bourses (voir Blanchard et Dionne, 2004 [9]). Mais toutes ces règlementations, règles et méthodes de gestion des risques n'ont pas été suffisantes pour empêcher la crise financière de 2007. Ce ne sont pas nécessairement les réglementations des risques et les règles de gouvernance qui ont fait défaut mais leur application ou leur respect. Il est bien connu que les réglementations et les règles sont généralement appelées à être contournées par divers intervenants dans différents marchés. Mais il semble que ces contournements soient devenus des comportements standards durant les années qui ont précédé la crise financière sans que les autorités réglementaires ne les aient anticipés, vus ou réprimandés.

Les récentes réformes règlementaires prudentielles du secteur financier encouragent à la prise en compte holistique des facteurs de risque dans les calculs de fonds propres des entreprises. Le dispositif règlementaire Solvabilité II, adoptée en 2009 par le Conseil de l'Union Européenne et le Parlement Européen (Directive 2009/138/CE du Parlement européen et du Conseil du 25 novembre 2009), est une réforme réglementaire applicable aux organismes d'assurance et de réassurance au niveau européen. Elaborée pour améliorer l'évaluation et le contrôle des risques, elle modifie en profondeur le régime prudentiel de ces organismes. Dans la lignée de Bâle II, son objectif est de mieux adapter les fonds propres exigés des compagnies d'assurance et de réassurance aux risques que celles-ci encourent dans leur activité. Cette directive est entrée officiellement en vigueur au 1er janvier 2016[1]. Elle incite assureurs et réassureurs à une meilleure gestion et analyse des risques au sein de leur organisation.

La gestion des risques, ou management du risque (*risk management*), est la discipline qui s'attache à identifier, évaluer et prioriser les risques relatifs aux activités d'une organisation, quelle que soit la nature ou l'origine de ces risques, pour les traiter méthodiquement de manière coordonnée et économique, afin de réduire et de contrôler la probabilité des événements redoutés, pour réduire l'impact éventuel de ces événements.

Selon le référentiel ISO Guide 73 – Vocabulaire du management du risque[2] qui a été revu lors du développement de la norme ISO 31000:2009 – Management du risque — Principes et lignes directrices[3], le risque est nouvellement défini comme « l'effet de l'incertitude sur les objectifs » et s'ajoute en note que « Un risque est souvent caractérisé en référence à des événements et des conséquences potentiels ou à une combinaison des deux. ».

La gestion des risques a pour but de créer un cadre de référence aux entreprises afin d'affronter efficacement le risque et l'incertitude. Les risques sont présents dans presque toutes les activités économiques et financières des entreprises. Le processus d'indentification, d'évaluation et de gestion des risques fait partie du développement stratégique de l'entreprise et doit être conçu et planifié au plus haut niveau, soit au conseil d'administration. Une approche intégrée de la gestion des risques doit évaluer, contrôler et faire le suivi de tous les risques auxquels l'entreprise est exposée.

En général, un risque pur est une combinaison de la probabilité ou fréquence d'un événement et de sa conséquence qui peut être positive ou négative. Il peut se mesurer par la déviation (ou la volatilité) par rapport à l'espérance mathématique ou aux résultats anticipés. L'incertitude est moins précise car,

[1] EUR-Lex, Directive 2009/138/CE du Parlement européen et du Conseil du 25 novembre 2009 sur l'accès aux activités de l'assurance et de la réassurance et leur exercice (solvabilité II) (Texte présentant de l'intérêt pour l'EEE)

[2] Référence officielle ISO Guide 73:2009 - Management du risque — Vocabulaire

[3] Référence officielle ISO 31000:2009 – Management du risque — Principes et lignes directrices

souvent, la probabilité d'un événement incertain n'est pas connue, de même que sa conséquence. Dans ce cas, on parlera plus d'activités de précaution plutôt que d'activités de prévention pour se protéger de l'incertitude. Finalement, il y a les risques spéculatifs, qui consistent à entreprendre des activités opportunistes par rapport aux risques futurs.

L'analyse des risques est le processus d'identification et d'analyse de problèmes potentiels pouvant avoir un impact négatif sur des initiatives commerciales ou des projets critiques afin d'aider les organisations à éviter ou à atténuer ces risques. Elle comprend la prise en compte de la probabilité d'événements indésirables causés par des processus naturels, ou d'événements provoqués par des activités humaines malveillantes ou fortuites. Une partie importante de l'analyse des risques consiste à identifier le potentiel de préjudice de ces événements, ainsi que la probabilité qu'ils se produisent.

Enfin, en 2003, la loi de sécurité financière (LSF) contraint les entreprises à communiquer sur leurs procédures de contrôle interne (et leurs procédures de management des risques). Enfin, le 8 décembre 2008, la transposition de la 8ème directive européenne sur la gouvernance d'entreprise fixe les modalités d'adoption des normes internationales d'audit et amène les entreprises à s'interroger sur l'efficacité et le suivi de leurs systèmes de contrôle interne et, de gestion des risques.

Des normes et standards internationaux sont venus s'ajouter à ces textes. On peut notamment citer COSO (*Committee of Sponsoring Organizations of the Treadway Commission*), référentiel de contrôle interne qui aborde, à partir de 2002 et de COSO 2, les règles de la gestion des risques d'entreprise (*Entreprise Risk management*, ou ERM). Et, plus récemment, la norme ISO 31000 : 2009 consacrée au management du risque. L'ERM est infiniment plus ambitieux que le traditionnel *risk management*. Il ne vise plus seulement à identifier et prévenir (ou couvrir) les risques opérationnels, mais à cartographier l'ensemble des risques de toutes natures (y compris exogènes à l'entreprise), à les hiérarchiser et à identifier les risques pouvant avoir un impact sur la stratégie même de l'entreprise.

2. Les grands principes de la norme Solvabilité II : paramètres spécifiques à l'entité

2.1. Principes de la Réforme Solvabilité II

La directive européenne Solvabilité II est entrée en vigueur au 1er janvier 2016. Il s'agit d'une règlementation prudentielle fondée sur trois piliers, comme les accords de Bâle pour le secteur bancaire.

Le premier pilier (pilier I) de Solvabilité II, prévoit de mesurer l'impact sur les capitaux propres de la société un ensemble d'évènements défavorables sur l'horizon d'un an à une probabilité correspondant au quantile 99,5%. De manière synthétique, ceci se traduit d'abord par des calculs des meilleures estimations des provisions (provisions *best estimate*), puis l'utilisation de la formule standard[4] afin de mesurer le capital réglementaire de solvabilité ajusté aux risques encourus. Ce montant de capital devrait permettre à la compagnie de faire face à une ruine économique à horizon 1 an et avec un temps de retour de 200 ans. Une approche alternative consiste à mettre en place un modèle interne partiel ou total, mieux adapté au profil de risque de la compagnie, permettant de déroger à la formule standard pour tout ou partie de ses modules de risques et matrices de corrélations. Enfin, pour certains risques classés dans la catégorie non-vie, l'Autorité Européenne des Assurances et des Pensions Professionnelles (*European Insurance and Occupational Pensions Authority*, ou EIOPA) propose une méthodologie de recalibrage des paramètres spécifiques à l'entité (*Undertaking Specific Parameters* ou *USP*) pour évaluer le capital économique des risques de suscription du sous module des primes et de réserves.

[4] La formule standard repose sur l'application aux différents facteurs de risque de chocs forfaitaires fournis par la norme et l'agrégation de besoins en capital élémentaires à partir de matrices de corrélation.

Le deuxième pilier (pilier II) a pour objectif de fixer des normes qualitatives de suivi des risques en interne aux sociétés et comment l'autorité de contrôle doit exercer ses pouvoirs de surveillance dans ce contexte. La procédure de surveillance de la gestion des fonds propres est destinée à s'assurer que la compagnie est bien gérée et est en mesure de calculer et maîtriser ses risques d'une part, et à s'assurer qu'elle est bien capitalisée d'autre part. L'identification des sociétés les plus risquées est un objectif et les autorités de contrôle ont en leur pouvoir la possibilité de réclamer à ces sociétés de détenir un capital plus élevé (*capital add-on*) que le montant suggéré par le calcul du SCR et/ou de réduire leur exposition aux risques.

Enfin, Le troisième pilier a pour objectif de définir l'ensemble des informations détaillées auxquelles le public aura accès, d'une part, et auxquelles les autorités de contrôle pourront avoir accès pour exercer leur pouvoir de surveillance, d'autre part.

Les trois premiers chapitres du manuscrit s'inscrivent principalement dans la démarche quantitative du pilier I et pilier II. Nous proposons différentes approches et méthodes de modélisation pour l'évaluation de la meilleure estimation des provisions, le calibrage de chocs spécifiques à l'entité et les calculs du capital économique. Pour illustrer nos propos, nous appliquons ces approches à trois familles de contrats en assurance de personnes (contrats d'assurance de prêts, contrats de prévoyance collective et contrats de dépendance) d'une compagnie d'assurance vie. Le dernier chapitre de la thèse se rapproche plutôt de la philosophie plus qualitative de suivi des risques en interne aux sociétés du pilier II de la norme Solvabilité II. Nous apportons des éclairages en matière de gestion des risques d'entreprise, en considérant le risque inhérent aux attitudes face au risque des individus au sein des organisations, notamment les entreprises du secteur financier.

2.2. Règles d'utilisation des paramètres spécifiques à l'entité

Pour le calcul de leur exigence de capital réglementaire, les organismes ou groupes d'assurance peuvent dans certains cas utiliser des paramètres qui leurs sont propres en remplacement des paramètres de la formule standard. Ces paramètres sont appelés couramment USP (*Undertaking specific parameters*), ou bien GSP (*Group specific parameters*) quand ils sont appliqués au niveau d'un groupe pour le calcul de l'exigence de capital sur base combinée ou consolidée. Les paramètres pouvant être personnalisés sont ainsi définis par la règlementation. Ils ne portent que sur le calcul de certains risques de souscription, à la maille du segment comprenant pour chaque ligne d'activité visée les affaires directes et acceptées proportionnellement. L'utilisation par un organisme ou groupe d'assurance de paramètres spécifiques pour le calcul de l'exigence en capital est soumise à autorisation préalable du régulateur national, l'Autorité de Contrôle Prudentiel et de Résolution (ACPR) en France.

La règlementation fixe les règles d'utilisation de paramètres spécifiques. Elles sont définies pour les organismes et pour les groupes, respectivement au V de l'article R. 352-5 et au neuvième alinéa du II de l'article R. 356-19 du code des assurances, applicables aux organismes et groupes relevant des trois codes, qui transposent l'article 104 (7) de la Directive 2009/138/CE de Solvabilité II. Ces dispositions sont complétées par les articles 218 à 220 et l'annexe XVII, ainsi que l'article 338 pour les groupes, du règlement délégué de niveau 2 (UE) n°2015/35. Enfin, le contenu du dossier de candidature et les différentes étapes de la procédure d'approbation sont précisés par le règlement d'exécution (UE) 2015/498 établissant des normes techniques d'exécution ou « ITS » en ce qui concerne la procédure d'approbation par les Autorités de contrôle de l'utilisation de paramètres propres à l'entreprise. Une fois le dossier déposé, l'ACPR se prononcera dans un délai maximal de 6 mois sur l'autorisation de l'utilisation totale ou partielle de paramètres spécifiques, à compter de la date de réception du dossier complet.

L'utilisation de paramètres spécifiques pour le calcul de la solvabilité au niveau individuel par un organisme et sur base consolidée (ou combinée) par un groupe sont deux mesures distinctes et indépendantes sur le plan administratif. Un groupe calculant son exigence de capital réglementaire sur la base des comptes consolidés ne peut tenir compte d'aucun paramètre spécifique approuvé au niveau d'un organisme pris individuellement. Si le groupe souhaite lui aussi utiliser des paramètres spécifiques, il est tenu d'en solliciter préalablement l'autorisation à l'ACPR, en tenant compte de son profil de risque pour l'ensemble de son périmètre, selon la même procédure que pour un organisme pris individuellement. De même, l'autorisation accordée à un groupe d'utiliser des paramètres spécifiques pour le calcul de sa solvabilité ajustée ne vaut pas autorisation pour les organismes membres du groupe pris individuellement.

3. Application des modèles à multi-états en assurance de personnes

Dans de nombreux domaines, décrire l'évolution des phénomènes dans le temps est d'un intérêt capital, en particulier pour aborder les problématiques de la prédiction et de la recherche de facteurs causaux. Les observations correspondent souvent à des mesures d'une même caractéristique faites à plusieurs instants. Ces données, appelées mesures répétées, se distinguent de celles présentes dans les modélisations statistiques traditionnelles. Les modèles multi-états constituent une alternative intéressante pour modéliser des données de type mesures répétées. Par exemples, ces modèles permettent de comprendre l'évolution des patients atteints de maladies chroniques comme le VIH ou virus de l'immunodéficience humaine, les cancers, l'asthme… (voir Philippe Saint Pierre, 2005 [10]). Ils permettent également d'analyser certains risques en assurance de personne (les arrêts de travail, la dépendance, etc…).

3.1. Les modèles à multi-états

Depuis une trentaine d'années, les modèles multi-états ne cessent de connaître un intérêt croissant. Ces modèles utilisent la notion d'« état » et de processus pour décrire un phénomène. La notion de processus est utilisée pour représenter les différents états successivement occupés à chaque temps d'observation. En épidémiologie, ils permettent par exemple, de représenter l'évolution d'un patient à travers les différents stades d'une maladie. Après définition des différents stades, les modèles multi-états permettent d'étudier de nombreuses dynamiques complexes. L'étude de ces modèles consiste à analyser les forces de passage (intensités de transition) entre les différents états.

Un nombre important de publications statistiques concerne les modèles multi-états. Cependant, l'application de ces modèles dépasse rarement le cadre des revues spécialisées. Cette situation s'explique en partie, par l'absence de logiciels adaptés et la méconnaissance des méthodes statistiques. La popularité et la richesse des modèles de survie, en particulier du modèle de Cox (1972) [17], dessert l'utilisation de ces modèles dans le domaine appliqué. Il est pourtant des situations où l'étude d'un délai d'apparition d'un événement ne peut apporter qu'une réponse partielle au problème posé.

Dans les modèles multi-états les plus simples, l'information sur l'état présent renseigne sur les états précédents : par exemple, les modèles progressifs (voir Hougaard, 1999 [11]), les modèles à risques compétitifs (voir Huber-Carol et Pons, 2004 [12] ; Andersen et al., 1993 [13]), ou encore les modèles de survie qui représentent le cas le plus simple avec uniquement deux états : « vivant » et « décès » (voir Therneau et Grambsch, 2000 [14]). Cependant, dès que le modèle comprend des états réversibles (c'est-à-dire que certains événements sont récurrents), il devient nécessaire de faire des hypothèses sur l'histoire de l'individu. Les modèles de type Markovien sont très utiles car ils supposent que l'infor-

mation sur les états précédents est résumée par l'état présent. Le terme de modèle multi-états regroupant de nombreuses problématiques biostatistiques, le nombre de publications sur le sujet est très important. On pourra se référer, par exemple, aux travaux de Andersen et Keiding (2002) [15], Hougaard (1999) [11], Andersen et al. (1993) [13] et Commenges (1999) [16] qui font le point sur l'état de l'art dans ce domaine.

Dans ces modèles de type Markovien, les intensités de transition entre les états peuvent dépendre de différentes échelles de temps, en particulier :

- la durée du suivi (temps depuis l'inclusion dans l'étude),
- le temps depuis la dernière transition (durée dans l'état présent).

Il existe plusieurs possibilités pour définir les intensités de transition $\alpha(t, d)$, où t représente la durée du suivi et d la durée passée dans l'état. Lorsque $\alpha(t, d) = \alpha$, le modèle est dit homogène par rapport au temps t. Lorsque $\alpha(t, d) = \alpha(t)$ le modèle est dit non-homogène. Dans le cas où les intensités de transition dépendent de la durée du suivi, $\alpha(t, d) = \alpha(d)$, le modèle est semi-Markovien homogène par rapport au temps t. Enfin, lorsque $\alpha(t, d)$ dépend des deux échelles de temps, le modèle est semi-Markovien non-homogène.

Dans certaines applications, la durée du suivi n'est pas toujours l'échelle de temps la mieux adaptée. En effet, le temps calendaire et l'âge peuvent également être considérés comme échelle de temps principale. Par exemple, le temps calendaire peut être adapté quand on considère le risque de contracter une maladie qui a une incidence variant beaucoup, comme l'infection par le VIH dans les années 80. Le choix entre les échelles de temps dépend de ce qui est le plus important dans une application donnée.

Plusieurs modèles statistiques sont possibles, on distingue les approches paramétriques, non-paramétrique et semi-paramétrique.

L'**approche paramétrique** stipule que les intensités de transition appartiennent à une classe particulière de fonctions, qui dépendent d'un nombre fini de paramètres. L'avantage de cette approche est la facilitation attendue de la phase d'estimation des paramètres. L'inconvénient est l'inadéquation pouvant exister entre le modèle retenu et le phénomène étudié.

L'**approche non-paramétrique** ne nécessite aucune hypothèse sur la forme des intensités de transition et c'est là son principal avantage. L'inconvénient d'une telle approche est la nécessité de disposer d'un nombre important d'observations. En effet, le problème de l'estimation d'un paramètre fonctionnel est délicat puisqu'il appartient à un espace de dimension infinie.

L'**approche semi-paramétrique** est une sorte de compromis entre les deux approches précédentes. Les intensités de transition appartiennent à une classe de fonctions en partie dépendant de paramètres et en partie s'écrivant sous forme de fonctions non-paramétriques. Cette approche est très répandue en analyse de survie au travers du modèle de régression de Cox (voir Therneau et al., 2000 [14]).

Le modèle peut également faire intervenir un effet aléatoire qui agit de manière multiplicative sur les intensités de transition. Dans les études de survie, ces modèles permettent de tenir compte de la dépendance entre les temps d'événement sont appelés modèle de fragilité ou *frailty model*, (voir Therneau et al. (2000) [14] et Hougaard (1995) [18]). Plus généralement, ces modèles permettent de prendre en compte des variables omises dans la modélisation, par exemple, les variables non observées, celles dont les effets sont déjà bien connus ou celles dont il n'est pas certain qu'elles influencent les intensités (voir Huber-Carol et Vonta, 2004 [19] ; Hougaard, 2000 [20] ; Nielsen et al., 1992 [21] et Andersen et al., 1993 [13]). Ces modèles sont particulièrement intéressants quand on peut distinguer des groupes d'individus. Par exemple de manière théorique, dans une population d'individus en dépendance, on peut différencier des groupes par cause d'entrée en dépendance ou type de pathologie. Cette distinction peut permettre d'avoir un effet aléatoire spécifique à chaque groupe.

Une particularité des données de cohorte réside dans le fait qu'elles ne sont que partiellement observées à cause des différents phénomènes de censure et troncature (à droite, à gauche, par intervalles). Par exemple, le mécanisme de censure à droite est toujours présent car on n'observe pas un phénomène sur un temps infini. Le mécanisme de censure par intervalles intervient quand les temps exacts de transition ne sont pas connus, on sait uniquement que les transitions se sont produites pendant un intervalle de temps (voir Commenges, 2002 [22]). Les méthodes d'estimation varient en fonction du type de données incomplètes.

Les modèles statistiques faisant intervenir des données censurées considèrent le plus souvent que le processus de censure est indépendant du processus d'événement. Dans les études de survie par exemple, cela suppose que le fait qu'un patient soit censuré n'apporte aucune information sur la survenue de l'événement. Dans le cas du VIH par exemple, les patients qui arrêtent le suivi sont souvent ceux dont l'état est le plus grave et dont le moral est affecté. Il est alors nécessaire de proposer des alternatives afin de tenir compte de l'information comprise dans la censure (voir Robins et al., 2000 [23], Minini et al., 2004 [24] et Little, 1995 [25]).

3.2. Utilisation en assurance de personnes

En assurance de personne, le calcul d'une provision *best estimate* et le suivi des risques des risques d'arrêt de travail ou de dépendance (notée garantie multi-état dans la suite), nécessitent de s'intéresser aux différents états parcourus par l'assuré pendant la période de couverture.

Quatre états peuvent être identifiés : 1) actif (vs valide), 2) incapacité de travail (vs dépendance partielle), 3) invalidité de travail (vs dépendance totale) et 4) décès.

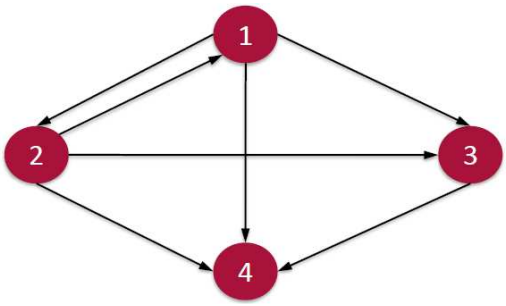


Figure 1 : les liens entre les différents états

Chaque changement d'états donne lieu à des flux financiers différents soit pour l'assureur, soit pour l'assuré. Les approches classiques évaluent le coût d'une garantie multi-état en considérant : a) des lois d'incidence pour décrire la survenance de l'incapacité (vs dépendance partielle), de l'invalidité (vs dépendance totale) et du décès, b) des lois de maintien en incapacité (vs dépendance partielle) et en invalidité (vs dépendance totale). Cependant, pour construire ces lois, les estimations sont réalisées de manière marginale en ne considérant que deux états à la fois (ex : Kaplan-Meier) ce qui biaise l'estimation réalisée. La construction de lois d'expérience *best estimate* se devrait de tenir compte de l'interaction entre les différents états, sous peine de générer un biais dans l'estimation des engagements. L'utilisation de modèles multi-états permettent de tenir compte de ces interactions et d'avoir une représentation détaillée de chaque loi de transition.

Les modèles multi-états constituent une classe particulière de modèle de durée qui s'utilise de manière naturelle dans les situations où l'assuré est susceptible d'évoluer entre plusieurs états. Il permet de déterminer les lois de transition d'un état vers un autre. Toutefois, on distingue plusieurs types de

modèles multi-états : le modèle **markovien homogène** (les intensités de transition sont constantes), le modèle **markovien inhomogène** (les intensités de transition dépendent du temps), le modèle **semi markovien homogène** (les intensités de transition dépendent de la durée passée dans l'état) et le modèle **semi markovien inhomogène** (les intensités de transition dépendent de la durée passée dans l'état et du temps).

Dans le cadre de l'arrêt de travail (vs dépendance), les modèles markoviens non-homogènes ou semi markovien non-homogènes constituent les classes de modèle les plus souvent retenues.

4. Rappels des risques et méthodes de provisionnement en assurance non-vie

En assurance non-vie, on peut distinguer les risques courts des risques longs. Les risques courts se définissent au regard de durations faibles (généralement inférieur à 3 ans). Ils se caractérisent principalement par un risque de fréquence. Les risques communément considérés comme courts sont : les capitaux en cas de décès (en prévoyance), les indemnisations des frais de soins de santé, les paiements des mensualités de prêts suite à une perte d'emploi, ainsi que les compléments de salaire sur l'incapacité (ce dernier étant considéré sous la norme Solvabilité II comme de la non-vie). Pour ces risques, les méthodes classiques utilisées pour le provisionnement sont basées sur l'utilisation des triangles de liquidation. Parmi ces méthodes de provisionnement de niveau agrégé (*macro-level reserving*), nous pouvons citer : la méthode des cadences, la méthode de *Chain Ladder*, la méthode dite liquidative qui s'appuie sur les bonis/malis. Ils existent d'autres méthodes de la même famille : l'approche par coût moyen et nombre de sinistres, la méthode de *Bootstrap* (voir Planchet, 2011 [26]), le modèle de Mack (voir Mack, 1993 [27] ; Mack, 1999 [28], Mack, 2000 [29] ; Mack, 2008 [30]) et la méthode Bornhuetter-Ferguson (1972) [31], etc…

La méthode de constitution des triangles est basée sur les triangles de liquidation constitués à partir des données de gestion ou comptable. L'estimation du nombre de sinistres ultimes est réalisée à partir d'un triangle de « nombre » fois un « coût moyen ». Dans la pratique, la mise en œuvre de cette méthode est rapide, la compréhension et l'interprétation des résultats sont faciles. Elle n'exige pas de disposer de données tête par tête. Les triangles sont aisément utilisables pour les calculs des provisions des comptes sociaux tout en permettant l'appréciation pertinente du niveau de la meilleure estimation des engagements. Pour les évaluations prospectives (*Market Consistent Embedded Value*, Solvabilité II, *International Financial Reporting Standards* ou IFRS) sur les risques courts, il est possible d'utiliser dans les modèles de projection un écoulement des IBNR (*Incurred But Not Reported*) en fonction des triangles de liquidations. Les flux projetés comparés aux flux réalisés restent très proches.

En contrepartie, elles sont sensibles à l'homogénéité des triangles de liquidation et à la stabilité des flux dans le temps. Le problème de données hétérogènes est une difficulté puisque les caractéristiques individuelles ne sont pas prises en compte. De plus, les données utilisées sont trop agrégées pour construire des indicateurs de suivi de sinistralité pertinents à des mailles plus fines. Enfin, ces méthodes ne sont pas adaptées aux risques longs.

Les risques longs se définissent au regard de durations plus importantes et se caractérisent principalement par un aléa sur la durée de paiement des prestations. Les risques communément considérés comme longs sont : les capitaux en cas de décès (en emprunteur), l'arrêt de travail emprunteur, l'invalidité prévoyance collective et les frais d'obsèques vie entière. Pour ces risques, les méthodes classiques utilisées pour le provisionnement s'appuient sur les lois biométriques. Ils existent des lois réglementaires nationales dans le code des assurances. La règlementation permet également aux assureurs l'établissement des lois propres à son portefeuille sur la base de leur expérience (lois d'expérience) et de les faire certifier pour une utilisation dans les comptes sociaux (par exemple la norme comptable française). Les lois réglementaires et d'expérience certifiées incluent une part de prudence qui génère généralement des bonis.

Pour les évaluations prospectives, des lois d'expérience dites de meilleure estimation (*best estimate*) sont généralement utilisées, ou par défaut, des lois règlementaires éventuellement ajustées à la sinistralité du portefeuille. D'une part, l'approche par les lois facilite la compréhension des évolutions des métriques provenant des risques dits techniques. Ces lois permettent d'enrichir les indicateurs avec des informations de duration. La disponibilité des données tête par tête apporte une richesse sur les indicateurs. L'utilisation des données tête par tête permet également de prendre en compte partiellement certaines caractéristiques individuelles (âge, ancienneté, sexe). D'un point de vue pratique, l'élaboration de ces lois est assez bien maîtrisée par les actuaires. D'autre part, l'obligation de disposer des données tête par tête et de bonne qualité est une contrainte incontournable pour la précision des calculs. Lors de la conception des tables d'expérience, les praticiens sont confrontés aux problèmes d'hétérogénéité dans les données. Cette analyse nécessite la prise en compte de toutes les caractéristiques des individus. Enfin, l'existence d'un niveau de prudence dans les lois réglementaires est une difficulté pour leur utilisation dans le cadre des analyses prospectives nécessitant les meilleures estimations des lois biométriques.

Finalement, afin d'améliorer la vision des risques, plusieurs approches ont été développée dans la littérature comme méthodes alternatives (voir paragraphe précédent). Ces approches utilisent pour la plupart des techniques d'apprentissage automatique. D'une part, celles qui suivent la logique *top-down*, par exemple la méthode *Chain-Ladder* non-paramétrique utilisant les réseaux neuronaux proposée par Wüthrich (2018) [32]. D'autre part, celles qui suivent une logique *bottom-up*, par exemple les méthodes dites *micro-level reserving*. Enfin, notons qu'en présence de données généralement censurées (à droite), ces méthodes *micro-level reserving* nécessitent des adaptations algorithmiques et des métriques pour fournir des prédictions performantes des risques (voir Lopez et al., 2016 [1]).

5. Techniques d'apprentissage automatique en présence de données censurées

Des revues systématiques récentes ont décrit plus de 100 modèles de prédiction de risque ont été produits entre 1999 et 2009, y compris les scores de risque bien connus comme le *Framingham risk score*, le *Reynolds risk score* et les récentes équations des cohortes regroupées de l'*American Heart Association* / *American College of cardiology*. La plupart des modèles de prédiction du risque ont été estimés à l'aide de données provenant de cohortes épidémiologiques homogènes et soigneusement sélectionnées. Ces modèles fonctionnent souvent mal lorsqu'ils sont appliqués à des populations diverses et actuelles. Comme exemples nous pouvons citer les modèles de prédiction du risque de maladie cardiovasculaire et les conséquences associées comme les crises cardiaques, les accidents vasculaires cérébraux.

Aux Etats-Unis, la disponibilité croissante de données électroniques sur la santé (EHD) et d'autres sources de données biomédicales majeures représente une opportunité clé pour améliorer les modèles de prévision des risques. EHD est un grand système de santé américain du Midwest. Il est composé de bases de données contenant des dossiers médicaux électroniques (*Electronic Medical Records*, EMRs), des données sur les sinistres d'assurance (*Insurance Claims Data*) et des données sur la mortalité obtenue à partir des registres de décès gouvernementaux (*governmental vital records*). Les bases de données EHD sont de plus en plus disponibles dans les systèmes de soins de santé de grande taille. Elles incluent généralement des données sur des centaines de milliers à des millions de patients avec des informations sur des millions de procédures, de diagnostics et de mesures en laboratoire. L'échelle et la complexité de l'EHD constituent une excellente occasion pour le développement des modèles de prédiction de risque plus précis à l'aide de techniques modernes d'apprentissage automatique.

Les données du système EHD comme d'autres sources de données massives sur les soins de santé ne sont pas collectées pour répondre à une question de recherche spécifique. Dans de nombreux ensembles de données dérivés du système EHD, la durée de collecte des informations sur un individu

particulier varie fortement d'un sujet à l'autre. Par conséquent, au cours d'une période donnée, une grande partie des individus observés ne disposent pas de données de suivi suffisantes pour déterminer s'ils ont ou non subi les conséquences des soins sur leur santé ou l'événement indésirable d'intérêt étudié. Dans le langage de l'analyse statistique de survie, la date de survenance de l'événement adverse (notée date de survenance) est appelé censure à droite si le suivi du sujet se termine avant qu'il ne soit touché par cet événement.

Malheureusement, la plupart des algorithmes d'apprentissage automatique supervisés et des méthodes de classification supposent généralement que le statut de l'événement est connu pour tous les individus, alors que le statut de l'événement est indéterminé pour les sujets dont la date de réalisation de l'événement est censurée. Ces derniers ne sont pas suivis sur toute la période de temps au-delà de laquelle l'on souhaite faire des prédictions.

Pour gérer les statuts d'événement inconnus en raison de la censure à droite, les travaux antérieurs ont proposé d'utiliser des étapes de prétraitement pour compléter ou exclure les données censurées, ou encore adapter des outils spécifiques d'apprentissage automatique aux données censurées. Par exemple, les approches courantes pour traiter les données censurées sont principalement de : (1) rejeter ces observations, (2) de les traiter comme des non-événements, (3) de séparer les observations en deux : une où l'événement se produit et une où l'événement ne se produit pas. Dans ce dernier cas, un poids est attribué à chacune de ces observations en fonction de la probabilité marginale de réalisation possible de l'événement entre la date de censure et la période où leurs statuts seront évalués.

Ces approches simples introduisent des biais dans l'estimation du risque (estimation des probabilités de classe) car, rejeter les observations avec un statut d'événement inconnu ou les traiter comme des non-événements à sous-estimer le risque. Même la troisième approche, bien que plus sophistiquée, produit des estimations de risque mal calibrées, car ces pondérations atténuent la relation entre les co-variables (caractéristiques) et les variables à expliquer (les résultats).

Ainsi, la plupart des approches d'apprentissage automatique qui prennent en compte le cas des données censurées de manière plus précises sont relativement récentes. La majorité de ces techniques sont des cas particuliers d'outils spécifiques d'apprentissage automatique. Par exemple, plusieurs auteurs, dont Hothorn et al. (2004) [33], Ishwaran et al. (2008) [34] et Ibrahim et al. (2008) [35] décrivent des versions d'arbres de classification et des forêts aléatoires (*Random Forest*) pour estimer la probabilité de survie. Lucas et al. (2004) [36] et Bandyopadhyay et al. (2015) [37] présentent une application des réseaux Bayésiens aux données censurées à droite. Peu d'auteurs ont envisagé appliquer les réseaux neuronaux aux données de survie, mais supposent généralement que les censures sont peu nombreuses. En outre, plusieurs ont envisagé adapter les machines à vecteurs de support (*support vector machines*, SVM) aux données censurées en modifiant la fonction de perte (*Loss function*) pour tenir compte des censurés. Ces approches sont toutes basées sur la modification de techniques spécifiques d'apprentissage automatique pour gérer la censure, ce qui limite la possibilité de généraliser l'approche utilisée pour traiter les résultats censurés à droite.

Finalement, Vock et al. (2016) [38] proposent une approche générale pour prendre en compte les données censurées à droite en utilisant la probabilité inverse de la pondération de la censure (*Inverse Probability of Censoring Weighting*, IPCW). Ils illustrent comment l'IPCW peut facilement être incorporé dans un certain nombre d'algorithmes d'apprentissage automatique calibrés sur de grandes données sur les soins de santé, y compris les réseaux bayésiens, les k-voisins les plus proches, les arbres de décision et les modèles additifs généralisés. Ils montrent ensuite que notre approche conduit à des prévisions mieux calibrées que les approches ad hoc appliquées à la prédiction du risque à cinq ans de survenance d'un accident cardiovasculaire, en utilisant l'EHD.

En réalité, ils existent certains travaux antérieurs utilisant la probabilité inverse de la pondération de la censure (IPCW) dans les méthodes d'apprentissage automatique. Par exemple, Bandyopadhyay et al. (2015) [37] ont discuté de la manière d'utiliser l'IPCW en particulier dans le contexte de l'estimation des réseaux bayésiens avec des données censurées à droite. Cependant, il est très probable qu'il s'agisse du premier article à proposer l'IPCW comme technique à usage général pouvant être utilisée

par de nombreuses méthodes d'apprentissage automatique. Ainsi, l'IPCW permet de bien prendre en compte des données censurées. Cette technique peut facilement être intégrée dans de nombreuses techniques d'apprentissage automatique existantes pour l'estimation des probabilités de classe. Elle ouvre la voie à la création de nouveaux outils d'apprentissage automatique pour les données censurées.

Lopez et al. (2016) [1] proposent une procédure d'arbre de régression pour estimer la distribution conditionnelle d'une variable qui n'est pas directement observée en raison de la censure en assurance, incluant l'analyse des durées de survie et le provisionnement des sinistres. Ils adaptent la méthodologie de CART (*Classification And Regression Trees*) de Breiman et al. (1984) [39] dans un contexte de données d'analyse de survie. La procédure fournie des résultats cohérents, et pour la sélection d'un sous arbre optimal utilise une stratégie d'élagage intégrant un schéma de pondération basé sur l'estimateur de Kaplan Meier (1958) [40] de la fonction de survie des données censurées. Ces résultats théoriques s'appuient sur une étude de simulation et deux applications aux données d'assurance respectivement la prévoyance et la responsabilité civile. Dans notre approche, nous proposons de corriger l'une des limites évoquées par les auteurs sur l'utilisation des arbres de régression, c'est-à-dire son instabilité.

C'est donc ces derniers travaux que nous retiendrons comme cadre de référence. Nous adaptons deux algorithmes d'apprentissage automatique qui corrigent l'instabilité de CART (à savoir, *Random Forest Censored* et *Gradient Tree Censored Boosting*) appliquées aux données censurées en intégrant la technique IPCW.

5.1. Méthodologie de calcul de l'IPCW

Nous présentons quelques notations et terminologies statistiques, ainsi que les éléments de calcul de l'IPCW dans le contexte d'observations contenant des données de censurées en assurance non-vie. Pour des justifications plus formelles de l'IPCW, consulter Bang et al. (2000) [41] et Tsitis et al. (2006) [42]. Une généralisation de cette technique aux algorithmes d'apprentissage automatique pour la prédiction de la survenance de conséquences des maladies cardio-vasculaires sur une durée donnée avec applications aux données provenant de l'EHD a été réalisée par Vock et al. (2016) [38]. L'article de Lopez et al. (2016) [1] présente le cadre théorique et l'adaptation de l'algorithme CART aux données censurées en assurance.

5.2. Notations et terminologies

Dans notre article, nous nous intéressons à la durée de maintien en cas d'arrêt de travail ou de chômage d'un individu sinistré couvert par un contrat de prêts (proposant des garanties d'arrêt de travail et de perte d'emploi) et de prévoyance collective (couvrant les arrêts de travail pour cause de maladie).

Nous nous intéressons plus généralement au vecteur aléatoire (M, T, X), où $M \in R^p$, $T \in R^+$ est la variable de durée, et $X \in \mathcal{X} \subset R^p$ le vecteur des co-variables aléatoires explicatives de T et/ou M. La présence de censures dans les données ne permet pas l'observation directe de (M, T), alors que X est toujours observable. Notons également $C \in R^+$ la variable de censure. Dans un souci de simplicité et sans perdre de généralité, supposons que les variables T et C sont des variables continues, et que les composantes de M sont toutes strictement positive. Les variables observables en dehors de (M, T) sont : $Y = min(T, C)$, $\delta = 1_{T \leq C}$ et $N = \delta M$.

Enfin, soit i $(1 \leq i \leq n)$ un individu de la base, il est représenté par $(N_i, Y_i, \delta_i, X_i)$. La base est donc une réplication de variables correspondant à des individus indépendants et identiquement distribués notée $(N_i, Y_i, \delta_i, X_i)_{1 \leq i \leq n}$. La variable M_i correspond à des quantités observables si et seulement si la durée

totale du sinistre de l'individu i est observée. Il s'agit des sinistres clôturés et M la charge ultime correspondant au montant total de prestations versées par l'assureur à l'individu sinistré. Dans certains cas spécifique $M = T$, voir *the censored regression framework* de Stute (1993) [43].

Dans ces conditions, le but de l'apprentissage automatique est de déterminer les impacts de X et si possible de T sur M. Cela revient à estimer une fonction π_0 telle que : $\pi_0 = \underset{\pi \in \Psi}{\text{Argmin}}\ E\big[\varphi\big(M, \pi(T, X)\big)\big]$, où Ψ est un sous ensemble d'un espace fonctionnel approprié et φ une fonction de perte. Lopez et al. (2016) [1] propose différents types de modèles de régression correspondant à différents choix de fonctions de perte dans un sous-ensemble fonctionnel.

5.3. Estimation des poids IPCW

Nous utilisons l'approche IPCW sur les données censurées et récemment utilisée pour des besoins d'apprentissage automatique. Dans la technique IPCW, seul les individus dont la durée totale des sinistres sont connues contribuent directement à l'analyse et aux estimations des probabilités de classes, mais ils sont pondérés pour représenter de manière la plus précise possible les individus censurés. L'avantage de la technique IPCW à tenir compte des données censurées est qu'elle permet de disposer d'une approche généralisable à n'importe qu'elle méthode d'apprentissage automatique pour la prédiction de risque (voir Vock et al. (2016) [38]). L'intérêt de cette démarche est de permettre de sélectionner le modèle le plus performant.

La démarche générale de la méthode IPCW se décompose en trois principales étapes : l'estimation du poids de Kaplan-Meier, la définition du schéma de pondération (probabilité inverse de pondération de la censure) et enfin application à l'algorithme d'apprentissage automatique.

Etape 1 : estimation du poids de Kaplan-Meier

En utilisant les données d'apprentissage, estimer la fonction $G(t) = 1 - P(C > t)$, la probabilité que la date de censure soit supérieur à t, en utilisant l'estimateur de Kaplan-Meier de la fonction de survie avec des données censurées. Nous retiendrons la formulation proposée par Lopez et al. (2016) [1]. En supposant que C est indépendant de (M, T) et $P(T \leq C|M, T, X) = P(T \leq C|T)$, on observe que pour toute fonction $\varphi \in L^1$,

$$E\left[\frac{\delta\varphi(N,Y,X)}{1-G(Y-)}\right] = E[\varphi(M, T, X)].$$

La fonction G qui est généralement inconnue peut cependant être estimée de manière cohérente par l'estimateur de Kaplan-Meier (1958) [40], c'est-à-dire :

$$\hat{G}(t) = 1 - \prod_{Y_i \leq t}\left(1 - \frac{\delta_i}{\sum_{j=1}^{n} 1_{Y_j \geq Y_i}}\right),$$

puisque T et C sont indépendantes, et $P(T = C) = 0$ pour les variables aléatoires continues.

Par conséquent, un estimateur naturel de $F(t, m, x) = P(T \leq t, M \leq m, X \leq x)$ est :

$$\hat{F}(t, m, x) = \frac{1}{n}\sum_{i=1}^{n} \frac{\delta_i 1_{Y_j \leq t, N_i \leq m, X_i \leq x}}{1-\hat{G}(Y_i-)},$$

Stute (1993) [43] démontrent la cohérence de l'estimateur de Kaplan-Meier. Ainsi, sous les conditions appropriées, un estimateur cohérent de $E[\varphi(T, M, X)]$ est :

$$\hat{E}[\varphi(T, M, X)] = \frac{1}{n}\sum_{i=1}^{n} \frac{\delta_i \varphi(Y_i, N_i, X_i)}{1-\hat{G}(Y_i-)}.$$

Etape 2 : définition du schéma de pondération (probabilité inverse de pondération de la censure)

Pour chaque individu i de la base d'apprentissage, on définit sa probabilité inverse de pondération de la censure ω_i tel que :

$$\omega_i = \begin{cases} \frac{1}{\hat{G}(\min(Y_i,t))} & si\ min(T_i,t) < C_i \\ 0 & sinon \end{cases}.$$

Les individus dont le statut d'événement est inconnu à t (c'est-à-dire, sont censurés avant t et ont donc $C_i \leq min(T_i, t)$) ont un poids $\omega_i = 0$, et sont donc exclus de l'analyse, et les individus restants se voient attribuer leurs poids ω_i.

Etape 3 : application à l'algorithme d'apprentissage automatique

Cette étape consiste à appliquer une méthode d'apprentissage automatique existant à la version pondérée des données d'apprentissage où chaque individu i de la base d'apprentissage est pondéré par son poids ω_i. Il faut noter que la manière spécifique dont les poids de l'IPC sont incorporés variera selon les techniques d'apprentissage automatique utilisées. Les mises en œuvre standard de certaines techniques d'apprentissage permettent la spécification directe des poids observés, auquel cas il faut peu de travail supplémentaire pour prédire les risques.

Plus généralement

La plupart des techniques d'apprentissage automatique impliquent une estimation (généralement à l'aide d'estimateurs du maximum de vraisemblance) et l'évaluation de la qualité de l'ajustement / de la pureté (*fit/purity*) du modèle. Vock et al. (2016) [38] montrent que l'incorporation des poids de l'IPC lors de l'estimation et de l'évaluation se font de manière assez simple. Une justification complète de l'utilisation de l'IPCW peut être trouvée dans Tsiatis.

5.4. Paysage simplifié des modèles prédictifs d'apprentissage automatique

Nous pouvons classer ces modèles en trois grandes familles : les modèles linéaires, les arbres de décisions et les autres. Les modèles de chaque famille peuvent être classés par ordre de complexité. Ainsi, la famille des modèles linéaires est composée des : modèles linéaires (*Linear Models*), modèles linéaires généralisés (*Generalized Linear Models*), modèles linéaires régularisés type régressions *ridge* et *elasticnet* (*Regularized Linear Models*). La famille des modèles d'arbres de décision est composée des : modèles d'arbres de classification et de régression (*Classification And Regression Trees*), modèles de forêt aléatoire (*Random Forest*), modèles *Gradient Boosted machines*. Enfin dans une troisième famille les modèles comme : *Nearest Neighbor*, *Naïve Bayes*, *Neural Networks* et *Support Vector machines*.

5.5. Technique d'adaptation des arbres de décision en présence de données censurées

L'apprentissage par arbre de décision est une méthode classique en apprentissage automatique. Son but est de créer un modèle qui prédit la valeur d'une variable-cible depuis la valeur de plusieurs variables d'entrée. Une des variables d'entrée est sélectionnée à chaque nœud intérieur (ou interne, nœud qui n'est pas terminal) de l'arbre selon une méthode qui dépend de l'algorithme. Chaque arête vers un nœud-fils correspond à un ensemble de valeurs d'une variable d'entrée, de manière à ce que l'ensemble des arêtes vers les nœuds-fils couvrent toutes les valeurs possibles de la variable d'entrée. Chaque feuille (ou nœud terminal de l'arbre) représente soit une valeur de la variable-cible, soit une

distribution de probabilité des diverses valeurs possibles de la variable-cible. La combinaison des valeurs des variables d'entrée est représentée par le chemin de la racine jusqu'à la feuille. L'arbre est en général construit en séparant l'ensemble des données en sous-ensembles en fonction de la valeur d'une caractéristique d'entrée. Ce processus est répété sur chaque sous-ensemble obtenu de manière récursive, il s'agit donc d'un partitionnement récursif. Cette famille d'algorithme est largement appliquée notamment en bio statistique ou en médecine.

De nombreuses techniques ont été proposées pour développer des arbres de décision, différant pour la plupart des critères utilisés pour décider comment ou s'il faut diviser un nœud et empêcher le sur-apprentissage. Dans le cas des arbres de classification, le critère d'évaluation des partitions caractérise l'homogénéité (ou le gain en homogénéité) des sous-ensembles obtenus par division de l'ensemble. Ces métriques sont appliquées à chaque sous-ensemble candidat et les résultats sont combinés (par exemple, moyennés) pour produire une mesure de la qualité de la séparation.

La technique très populaire, CART par exemple, utilise l'indice de diversité de Gini. Cet indice mesure avec quelle fréquence un élément aléatoire de l'ensemble serait mal classé si son étiquette était choisie aléatoirement selon la distribution des étiquettes dans le sous-ensemble. Cette métrique permet donc de déterminer à chaque nœud s'il convient de le subdiviser en nœuds-fils. La variation (diminution) de l'indice de diversité de Gini pour une subdivision possible est donné par :

$$\Delta I_G = \hat{\pi}(s)[1-\hat{\pi}(s)] - \frac{1}{N_s}\left\{N_{s_l}\hat{\pi}(s_l)[1-\hat{\pi}(s_l)] + N_{s_r}\hat{\pi}(s_r)[1-\hat{\pi}(s_r)]\right\}$$

où $\hat{\pi}(s)$, $\hat{\pi}(s_l)$ et $\hat{\pi}(s_r)$ sont respectivement la proportion de l'échantillon (d'apprentissage) affectés au nœud s, au nœud-fils à gauche s_l et au nœud-fils droit s_r pour un critère donné de subdivision. N_{s_l} et N_{s_r} sont le nombre d'individus dans chaque nœud-fils (gauche et droit) , tel que : $N_s = N_{s_l} + N_{s_r}$.

Notons E l'indicateur de survenance de l'événement que l'on cherche à prédire (tel que : $E = 1$ si l'événement s'est produit et $E = 0$ sinon).

Dans le cas simple non pondéré, la proportion de l'échantillon $\hat{\pi}(s)$ au nœud s est calculée comme une moyenne des indicateurs de survenance de l'événement des individus affectés à ce nœud. Pour l'étape de test pour un individu représenté par la réalisation des co-variables (*features*) notées x, appartenant au nœud terminal s_T, nous pouvons estimer son risque noté $\pi(x)$, comme la proportion des individus i appartenant à ce nœud terminal lors de l'étape d'apprentissage et pour lesquels l'événement est survenu (c'est-à-dire $E_i = 1$).

Il est simple d'adapter les arbres de décision pour incorporer les poids IPCW. Dans ce cas, il faut affecter aux individus i lors de la phase d'apprentissage les poids ω_i comme décrits ci-dessus. Ces poids sont utilisés pour la pondération des classes dans l'algorithme d'arbre de décision.

Avec les poids IPCW, nous calculons une diminution pondérée de l'indice de diversité de Gini comme suit :

$$\Delta I_G^{\omega} = \hat{\pi}^{\omega}(s)[1-\hat{\pi}^{\omega}(s)] - \frac{1}{N_s^{\omega}}\left\{N_{s_l}^{\omega}\hat{\pi}^{\omega}(s_l)[1-\hat{\pi}^{\omega}(s_l)] + N_{s_r}^{\omega}\hat{\pi}^{\omega}(s_r)[1-\hat{\pi}^{\omega}(s_r)]\right\}$$

avec $N_s^{\omega} = \sum_{i\in s}\omega_i$, $N_{s_l}^{\omega} = \sum_{i\in s_l}\omega_i$, $N_{s_r}^{\omega} = \sum_{i\in s_r}\omega_i$,

et $\hat{\pi}^{\omega}(s) = \frac{\sum_{i\in s}\omega_i E_i}{N_s^{\omega}}$, $\hat{\pi}^{\omega}(s_l) = \frac{\sum_{i\in s_l}\omega_i E_i}{N_{s_l}^{\omega}}$, $\hat{\pi}^{\omega}(s_r) = \frac{\sum_{i\in s_r}\omega_i E_i}{N_{s_r}^{\omega}}$.

Une fois la structure de l'arbre de décision est déterminée, la prédiction lors de la phase de test avec les réalisations x des *features* représentant les individus appartenant au nœud terminal s_T , du risque $\hat{\pi}(x)$, est estimée en utilisant la moyenne pondérée :

$$\hat{\pi}^{\omega}(x) = \frac{\sum_{i\in s_T}\omega_i E_i}{N_{s_T}^{\omega}},$$

où ω_i correspond au poids IPC donné précédemment, et $N_{s_T}^{\omega} = \sum_{i=1}^{n}\omega_i\, I_{\{S_i = s_T\}}$.

Dans le cas des arbres de régression, le même schéma de séparation peut être appliqué, mais au lieu de minimiser le taux d'erreur de classification, on cherche à maximiser la variance interclasse (avoir des sous-ensembles dont les valeurs de la variable-cible soient les plus dispersées possibles). En général, le critère utilise le test du chi carré. Pour l'adaptation de l'algorithme au cas pondéré, une approche similaire à celle présentée ci-dessus pourrait être appliquée pour calculer la version pondérée de la métrique adéquate.

L'une des principales limites des arbres de classification est liée à leur flexibilité, les rendant instables et souvent confrontés à un problème de sur-apprentissage sur les données d'entraînement. Plusieurs procédures d'élagage permettant de contourner ce problème de sur-apprentissage sont proposées avec la plupart impliquant un paramètre de réglage qui limite la complexité de l'arbre.

L'une des stratégies consiste à définir une limite inférieure m pour le nombre d'individus affectés à un nœud terminal. Dans notre notation ci-dessus, le nœud S ne serait pas subdivisé (suivant la règle de subdivision utilisée) à moins que : $min(N_{s_l}, N_{s_r}) \geq m$. Cette stratégie est facilement généralisable au cas des données censurées en retenant comme contrainte : $min(N_{s_l}^{\omega}, N_{s_r}^{\omega}) \geq m$. Cependant nous notons que $N_{s_l} \approx N_{s_l}^{\omega}$ et $N_{s_r} \approx N_{s_r}^{\omega}$ si la valeur de attendue de $\omega_i = 1$, donc, en pratique, fixer une limite inférieure pour $min(N_{s_l}, N_{s_r})$ est généralement suffisante.

Une autre approche consiste à accepter la subdivision d'un nœud dès lors que la diminution de l'indice de diversité de Gini dépasse un certain seuil fixé θ, c'est-à-dire $\Delta I_G(s) \geq \theta$. Substituer $\Delta I_G(s) \geq \theta$ par $\Delta I_G^{\omega}(s) \geq \theta$ permet d'utiliser la même règle dans le paramétrage des données censurées. Finalement, les valeurs optimales des paramètres de réglage peuvent être choisies par des techniques de validation croisée. La validation croisée est un moyen de prédire l'efficacité d'un modèle sur un ensemble de validation hypothétique lorsqu'un ensemble de validation indépendant et explicite n'est pas disponible. Il existe au moins trois variantes : « *tests et validation* » ou « *hold out method* », « *k-fold cross-validation* » et « *leave-one-out cross-validation* ».

Dans ce qui suit, nous reprenons les notations précédentes pour la description des algorithmes mise en œuvre.

5.6. Modèle d'apprentissage « *Tree base Censored* »

Lopez et al. (2016) [1] propose une formalisation de l'adaptation de CART aux données censurées en assurances que nous ne rappelons pas dans cet article. Ils y proposent également une stratégie d'élagage des arbres de régressions intégrant la pénalisation par le poids de Kaplan-Meier. D'un point de vue mise en œuvre pratique, les principales étapes de calculs sont les suivantes : le calcul des poids de Kaplan-Meier, l'obtention de l'arbre maximal et la stratégie d'élagage permettant d'adapter l'algorithme de CART au cas de données censurées.

5.7. Modèle d'apprentissage « *Random Forest Censored* »

Les forêts d'arbres décisionnels ou forêts aléatoires (*Random Forest Classifier*) ont été formellement proposées en 2001 par Leo Breiman et Adèle Cutler (voir Breiman, 2001 [44]). Elles font partie des techniques d'apprentissage automatique. Cet algorithme combine les concepts de sous-espaces aléatoires et de *bagging*. L'algorithme des forêts d'arbres décisionnels effectue un apprentissage sur de multiples arbres de décision entraînés sur des sous-ensembles de données légèrement différents.

La base du calcul repose sur l'apprentissage par arbre de décision. La proposition de Breiman vise à corriger plusieurs inconvénients connus de la méthode initiale, comme la sensibilité des arbres uniques à l'ordre des prédicteurs, en calculant un ensemble de B arbres partiellement indépendants.

Une présentation rapide de la procédure peut s'exprimer comme suit :

1) Créer B nouveaux ensembles d'apprentissage par un double processus d'échantillonnage :
   - sur les observations, en utilisant un tirage avec remise d'un nombre N d'observations identique à celui des données d'origine (technique d'inférence statistique datant de la fin des années 1970 et connue sous le nom de *bootstrap*),
   - et sur les p prédicteurs, en n'en retenant qu'un échantillon de cardinal $q < \sqrt{p}$ (la limite n'est qu'indicative).

2) Sur chaque échantillon, on entraîne un arbre de décision selon une des techniques connues, en limitant sa croissance par validation croisée.

3) On stocke les B prédictions de la variable d'intérêt pour chaque observation d'origine.

4) La prédiction de la forêt aléatoire est alors un simple vote majoritaire (*Ensemble learning*).

Nous commencerons d'abord par définir les algorithmes type *bagging* puis expliquerons en quoi consiste le *Random Forest*. Le *Bagging* pour *bootstrap* et *aggregating* est une méthode pouvant s'appliquer à toute méthode de modélisation (régression, CART, Neural Network, etc...) pour réduire l'erreur de prédiction. La méthode *bagging* est très utile dans le cas de modèle instable comme CART car elle permet de réduire la variance.

Soient $Y \in R$ la variable à prédire (ou réponse) et X_k $(k = 1,2, \dots, p)$ les co-variables (ou variables explicatives) et supposons qu'on a construit un premier modèle $\hat{f}$ sur l'échantillon d'apprentissage. Le *bagging* consiste à choisir un ensemble de *bootstrap* de l'échantillon d'apprentissage puis de créer M modèles pour chaque échantillon *bootstrapé*. Par exemple, si nous considérons le cas où la variable réponse est quantitative, si on note $\hat{f}_m^{(b)}$ le modèle calibré sur les données de l'échantillon d'apprentissage *boostrapé* m $(m = 1, \dots, M)$, alors l'estimateur par la méthode *bagging* dans le cas d'une régression est tout simplement la moyenne des prédictions de chaque modèle :

$$\hat{f}_{bag}(x) = \frac{1}{m} \sum_{m=1}^{M} \hat{f}_m^{(b)}(x)$$

Dans le cas de l'algorithme CART, Breiman propose une amélioration du *bagging* en ajoutant une dimension aléatoire. L'objectif étant de rendre les arbres plus indépendants en sélectionnant les variables explicatives de manière aléatoire.

L'adaptation proposée utilise la même pondération que dans *Tree-Base Censored* (voir Lopez et al., 2016 [1]). Les poids de Kaplan-Meier sont utilisés lors de la phase d'apprentissage de l'algorithme *Random Forest* lors de la création des arbres optimaux élagués.

Dans ces conditions, nous proposons dans ce qui suit une procédure pour l'implémentation de l'algorithme du *Random Forest* pour les données contenant des observations censurées.

5.8. Adaptation de l'algorithme de *Random Forest* aux données censurées

Données **:** *M le nombre d'arbres optimaux élagués prédéfini, on dispose d'une base composée de n individus représentés par $z = \{(N_i, Y_i, \delta_i, X_i)_{i=1,\dots,n}\}$*

Résultats : *Prédiction du risque*

Etape 0 *: Calculer l'estimateur $\hat{G}$ suivant la formule analytique proposée pour les n individus et calculer les poids de Kaplan-Meir*

Etape 1 *: itération de la procédure Tree-base censored*

Pour $m = 1$ à M faire

- un tirage aléatoire dans z d'un échantillon boostrap (avec remise) noté $z_m^{(b)}$

- estimer l'arbre de régression optimal élagué $\hat{f}_{z_m^{(b)}}$ avec l'algorithme Tree-base censored

Etape 2 *: agrégation des modèles*

Calculer $\hat{f}_{bag}(x) = \frac{1}{m}\sum_{m=1}^{M} \hat{f}_{z_m^{(b)}}(x)$

Estimation du risque : *utiliser $\hat{f}_{bag}$ comme l'estimateur final de π_0*

5.9. Modèle d'apprentissage *Gradient Tree Censored Boosting*

Après le *Bagging* et l'algorithme *Random Forest* présenté ci-dessus, nous proposons une adaptation d'un algorithme de boosting avec le cas particulier de l'algorithme *Gradient Tree Boosting Machine* (GBM). Il s'agit là encore d'une méthode d'agrégation de modèles dont le principe de fonctionnement ressemble à celle du *Bagging*. Plutôt que d'utiliser un seul modèle, cet algorithme utilise plusieurs modèles qu'il agrège ensuite pour obtenir un seul résultat.

Dans la construction des modèles, le *Boosting* travaille de manière séquentielle. Le principe de base est de construire une séquence de modèles de sorte que chaque étape, chaque modèle ajouté à la combinaison, apparaisse comme un pas vers une meilleure solution. Il commence par construire un premier modèle qu'il va évaluer. A partir de cette mesure, chaque individu va être pondéré en fonction de la performance de la prédiction.

L'objectif est de donner un poids plus important aux individus pour lesquels la valeur a été mal prédite pour la construction du modèle suivant. Le fait de corriger les poids au fur et à mesure permet de mieux prédire les valeurs difficiles. Pour le cas particulier de GBM, le calcul du poids des individus lors de la construction de chaque nouveau modèle se fait avec le gradient de la fonction de perte. La principale innovation est que ce pas est franchi dans la direction du gradient de la fonction perte, afin d'améliorer les propriétés de convergence. Il s'agit d'une démarche similaire à la descente de gradient pour les réseaux de neurones proposée par Friedman (2002) [45].

GBM utilise généralement des arbres CART et on peut personnaliser l'algorithme en utilisant différents paramètres, différentes fonctions. Enfin, limiter la taille des arbres ou construire les modèles sur des échantillons de la population (on parle de *stochastic gradient boosting*) sont des méthodes utilisées pour éviter le sur-apprentissage.

Le GBM est donc une famille une famille d'algorithmes basés sur une fonction perte supposée convexe et différentiable notée $l^{(P)}$. Le modèle pas à pas est une descente de gradient de la forme suivante :

$$\hat{f}_m = \hat{f}_{m-1} - \gamma_m \sum_{i=1}^{n} \nabla f_{m-1} l^{(p)}\big(y_i, f_{m-1}(x_i)\big).$$

Le problème revient à rechercher un meilleur pas de descente γ tel que :

$$\min_{\gamma} \sum_{i=1}^{n} l^{(p)}\left(y_i, f_{m-1}(x_i) - \gamma \frac{\partial l^{(p)}\big(y_i, f_{m-1}(x_i)\big)}{\partial f_{m-1}(x_i)}\right)$$

Dans ces conditions, nous proposons une procédure pour l'implémentation de l'algorithme du *Gradient Boosting* pour les données contenant des observations censurées.

5.10. Adaptation de l'algorithme de *Gradient Tree Boosting* aux censures

Données **:** *M le nombre d'arbres optimaux élagués prédéfini, on dispose d'une base composée de n individus représentés par : $z = \{(N_i, Y_i, \delta_i, X_i)_{i=1,\dots,n}\}$*

Résultats : *Prédiction du risque*

Etape 0 *: Calculer l''estimateur $\hat{G}$ suivant la formule analytique proposée pour les n individus et calculer les poids de Kaplan-Meir*

Etape 1 *(Initialisation) : $\hat{f}_0 = \underset{\gamma}{argmin} \sum_{i=1}^{n} l^{(p)}(y_i, \gamma)$*

Début de la boucle : Pour $m = 1$ à M faire

- Calculer $r_i^m = -\left[\frac{\partial l^{(p)}(y_i, f(x_i))}{\partial f(x_i)}\right]_{f=f_{m-1}}$ pour $i = 1, \dots, n$

- Ajuster l'arbre de régression optimal élagué δ_m au couple $(x_i, r_i^m)_{i=1,\dots,n}$ avec l'algorithme Treebase censored

- Calculer γ_m en résolvant : $\min_{\gamma} \sum_{i=1}^{n} l^{(p)}\big(y_i, f_{m-1}(x_i) - \gamma \delta_m(x_i)\big)$

- Mise à jour : $\hat{f}_m(x) = \hat{f}_{m-1}(x) - \gamma_m \delta_m(x)$

Fin de la boucle

Estimation du risque : *utiliser $\hat{f}_M$ comme l'estimateur final de π_0*

5.11. Evaluation des modèles de prédiction de risque sur des données censurées

a) *Métriques de sélection : MSE et MSEW*

Afin de choisir le meilleur modèle parmi tous les modèles calibrés à la phase d'apprentissage et de validation, nous utiliserons deux métriques de comparaison des modèles. Il s'agit de la version pondérée de l'erreur quadratique moyenne (*Weighted Mean Square Error*, MSEW) par les poids de Kaplan-Meier.

La formule non pondérée MSE est souvent utilisées dans le cas de la régression lorsque les observations sont complètes. La prise en compte d'observations censurées dans les données conduit à une adaptation de la MSE par l'introduction du poids de Kaplan-Meier. Formellement, ces métriques sont définies par

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - f(X_i)\right)^2,$$

$$MSEW = \frac{1}{n}\sum_{i=1}^{n}\omega_i \times \left(Y_i - f(X_i)\right)^2,$$

avec i les individus (($i = 1, \dots, n$), X_i le vecteur des co-variables, Y_i la variable à prédire (*target*) et $f(.)$ la fonction d'estimation du modèle, ω_i le poids de Kaplan-Meier de l'individu.

b) *Amélioration du MSEW par ajustement à l'unité de la valeur cible*

Pour se ramener à l'unité de Y , on peut prendre la racine de la MSEW. On obtient ainsi la RMSEW, (*Weighted Root Mean Squared Error*). La formule de la RMSEW est

$$RMSEW = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\omega_i \times \left(Y_i - f(X_i)\right)^2}.$$

Mais la RMSEW ne se comporte pas très bien quand les étiquettes peuvent prendre des valeurs qui s'étalent sur plusieurs ordres de grandeur. Pour prendre cela en compte, nous proposons d'appliquer la fonction logarithme aux valeurs prédites et aux vraies valeurs avant de calculer la RMSEW.

La formule de la *RMSLEW* (*Weighted Root Mean Squared Log Error*) est donnée par l'égalité

$$RMSLEW = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\omega_i \times (\log(Y_i + 1) - \log(f(X_i) + 1))^2}.$$

c) *Coefficient de détermination pondérée*

Même si les valeurs à prédire ont toutes le même ordre de grandeur, la RMSEW peut être difficile à interpréter. Nous proposons pour ce faire de compléter notre analyse par une amélioration de l'erreur carrée relative ou RSE (*Relative Squared Error*), notée l'erreur carrée relative pondérée, ou RSEW (*Weighted Relative Squared Error*) plus facilement interprétable.

Cet indicateur est le résultat de la normalisation de la somme pondérée des carrés des résidus non pas par le nombre de points n dans le jeu de données, mais par une mesure de ce qu'il serait raisonnable de faire comme erreur : la somme pondérée des distances entre chacune des valeurs à prédire et leur moyenne. La formule de la RSEW est donnée par

$$RSEW = \frac{\sum_{i=1}^{n}\omega_i \times \left(Y_i - f(X_i)\right)^2}{\sum_{i=1}^{n}\omega_i \times (Y_i - \bar{Y})^2}, \text{ avec } \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i.$$

Le coefficient de détermination pondéré correspond au complémentaire du RSEW dans ce cas est $R_W^2 = 1 - RSEW$.

Ce coefficient de détermination nous indique donc à quel point les valeurs prédites sont corrélées aux vraies valeurs.

d) Métrique de validation

Après avoir sélectionné les modèles candidats pour la prédiction suivant la métrique MSEW, nous validons notre choix avec l'indice de concordance (concordance index).

En effet, l'aire sous la courbe ROC (AUC) est un indicateur synthétique largement utilisée pour mesurer la performance des modèles prédictifs. Lorsque la variable à expliquer (résultat) est complètement observée sur tous les individus, l'AUC est équivalent à l'indice de concordance (C-index), correspondant à la probabilité de classer correctement les résultats pour une paire d'individus choisis au hasard et dont les valeurs des prédites sont différentes.

Comme décrit dans Harrell, le C-index peut être adapté pour les cas de données censurées en considérant la concordance des résultats de survie versus la probabilité de survie prédite entre des paires d'individus dont les résultats de survie peuvent être ordonnés. Dans notre cas, cela revient à dire parmi les paires d'individus non censurés car tous les deux rétablis, ou l'un est rétabli et l'autre est censuré après le rétablissement du premier avant la fin d'observation.

La formule de l'indice C-index adapté au cas de données censurées est donnée par l'égalité suivante

$$C_{cens}(t) = \frac{\sum_{i \neq j} \delta_i I_{\{V_i < V_j\}} I_{\{\hat{\pi}(X_i) < \hat{\pi}(X_j)\}}}{\sum_{i \neq j} \delta_i I_{\{V_i < V_j\}}},$$

où $I_{\{.\}}$ est une fonction indicatrice.

6. Evaluation de composantes entrant dans le calcul de la meilleure estimation des provisions pour sinistre en assurance non-vie

6.1. Les composantes de la provision technique sous solvabilité II

Au niveau du Bilan prudentiel défini par Solvabilité II, l'approche économique a été retenue pour l'évaluation des passifs, visant à assurer une communication financière homogène et transparente au niveau européen. Les passifs sont valorisés au montant pour lequel ils pourraient être transférés ou réglés dans le cadre d'une transaction conclue dans des conditions d'assurance et de réassurance normales, entre des parties informées et consentantes. La valeur des provisions techniques devrait être égale à la somme de la meilleure estimation (la provision dite *Best Estimate*) et d'une Marge de risque.

L'article 77 de la Directive 2009/138/CE définit le cadre général de calcul des provisions techniques à inscrire au Bilan prudentiel.

L'alinéa 4 précise que le *Best Estimate* et la Marge de risque doivent être évalués séparément : « *Les entreprises d'assurance et de réassurance procèdent à une évaluation séparée de la meilleure estimation et de la Marge de risque*. ». Il existe cependant des exceptions (engagements réplicables au moyen d'instruments financiers), mais a priori elles ne concernent pas les entreprises d'assurance Non-Vie.

L'alinéa 2 est particulièrement important, il définit le cadre général de calcul des *Best Estimates* : « *La meilleure estimation correspond à la moyenne pondérée par leur probabilité des flux de trésorerie futurs, compte tenu de la valeur temporelle de l'argent (valeur actuelle attendue des flux de trésorerie*

futurs), estimée sur la base de la courbe des taux sans risque pertinente. Le calcul de la meilleure estimation est fondé sur des informations actualisées et crédibles ainsi que sur des hypothèses réalistes. Ce calcul fait appel à des méthodes actuarielles et statistiques adéquates, applicables et pertinentes.

La projection en matière de flux de trésorerie utilisée dans le calcul de la meilleure estimation tient compte de toutes les entrées et sorties de trésorerie nécessaires pour faire face aux engagements d'assurance et de réassurance pendant toute la durée de ceux-ci.

La meilleure estimation est calculée brute, sans déduction des créances découlant des contrats de réassurance et des véhicules de titrisation. Ces montants sont calculés séparément, conformément à l'article 81. ».

Ainsi, pour valoriser les *Best Estimates*, l'entreprise d'assurance ou de réassurance doit identifier tous les encaissements et décaissements relatifs à ses engagements. Les flux correspondants doivent être probabilisés de manière à calculer une espérance mathématique (ce qui exclut d'intégrer toute marge de prudence) et doivent ensuite être actualisés sur la base de la courbe de taux sans risque (ce qui nécessite d'identifier les dates de tombée de flux) puis sommés pour obtenir le montant de *Best Estimate*.

Les *Best Estimates* sont évalués bruts de réassurance et inscrits au passif du Bilan. Des provisions sont constituées en représentation à l'actif pour prendre en compte les cessions aux réassureurs et aux véhicules de titrisation. Ces dernières provisions sont ajustées « *afin de tenir compte des pertes probables pour défaut de la contrepartie* ».

L'alinéa 3 définit le cadre général de calcul de la Marge de risque : « *La Marge de risque est calculée de manière à garantir une valeur des provisions techniques équivalente au montant que les entreprises d'assurance et de réassurance demanderaient pour reprendre et honorer les engagements d'assurance et de réassurance.* ». En effet, une société en situation de run-off (c'est-à-dire qui ne souscrit plus de nouveaux contrats) doit conserver un capital minimum pour être certaine de couvrir ses engagements (avec une probabilité de 99,5% selon les normes Solvabilité II), car ceux-ci sont aléatoires.

Le coût de ce capital est représenté par la Marge de risque. Le calcul de la Marge de risque est encadré par des règles strictes, il ne s'agit donc pas d'une marge de prudence. Les provisions techniques du Bilan prudentiel correspondent à la somme du *Best Estimate* et de la Marge de risque.

Pour pouvoir valoriser des provisions techniques, il est nécessaire de déterminer avec précision les engagements de l'entreprise envers les assurés et tous autres tiers, et réciproquement de déterminer les engagements des assurés et tiers envers l'entreprise d'assurance ou de réassurance. La notion de frontière des contrats permet d'identifier les engagements à considérer pour constituer le Bilan et ceux à ne pas inclure. Il existe de nombreux paragraphes dans les spécifications techniques pour définir cette notion. La frontière des contrats constitue une différence importante entre la norme comptable française actuelle et la norme Solvabilité II. La norme Solvabilité II requiert de prendre en compte plus de contrats (puisque les engagements doivent être comptabilisés dès lors que l'entreprise n'a plus la possibilité d'agir unilatéralement sur les termes du contrat).

En normes actuelles, les provisions de sinistres sont destinées à couvrir l'ensemble des règlements de sinistres y compris les frais de gestion restant à honorer pour l'ensemble des sinistres déjà survenus, qu'ils soient connus ou non de l'assureur. Il s'agit de la somme des provisions dossier/dossier et des provisions pour tardifs, calculées par les actuaires. L'article 36 du Règlement Délégué dans son point 3 précise que le périmètre de calcul du *Best Estimate* des provisions pour sinistres est strictement identique, il s'agit bien de provisionner l'ensemble des sinistres survenus, connus ou non de l'assureur.

6.2. Estimation des paramètres de provisionnement

Pour les assureurs non-vie, il existe deux principales méthodes issues des statistiques classiques permettant de calculer les provisions pour sinistres. La première méthode couramment utilisée est celle du *Chain Ladder*. Elle s'appuie sur des données agrégées et a pour avantage d'être facile à comprendre, aisée à implémenter et intègre implicitement les sinistres tardifs (*Incurred But Not Reported* ou IBNR). De par sa simplicité, ce modèle présente des inconvénients importants. D'abord, elle admet des hypothèses très fortes sur la stabilité des sinistres dans le temps et l'indépendance des survenances entre les années. Pour la plupart des portefeuilles d'assurance, ces hypothèses ne sont pas vérifiées, ce qui pose des problèmes et nécessites de grandes précautions lors de la mise en œuvre de cette méthode. Ensuite, elle n'utilise pas les données individuelle tête par tête disponible, ce qui ne permet pas de s'assurer que les agrégats forment bien des groupes de risques homogènes. La deuxième méthode est l'approche *Micro-Level Reserving* (comme le *Genaral Linear Model* ou GLM) et l'estimateur de Kaplan-Meier. Cette méthode est également facile à expliquer, avec en plus l'avantage d'utiliser des données individuelles. Cependant, elle présente certaines limites liées à son implémentation, à sa validation et à l'estimation des paramètres. En pratique ce modèle paramétrique est plus difficile à implémenter, sa performance est jugée sur sa qualité asymptotique plutôt que sur son adéquation aux données réelles. Notons également qu'ils existent des risques d'erreur d'estimation des paramètres surtout dans le cas de plusieurs co-variables retenues comme explicatives. Enfin, ces deux approches ne tiennent pas en compte spécifiquement des données censurées. Finalement, à l'heure des données massives, ces modèles n'utilisent pas toute la richesse des données individuelles disponible chez certains assureurs (données identitaires, transactionnelles, relationnelles, démographiques, contractuelles et sur l'état des assurés).

7. Calibrage des paramètres spécifiques pour la garantie arrêt de travail

En assurance de personne, les risques d'incidence et de maintien (ou de rétablissement) associés à la garantie arrêt de travail sont classées dans le module Santé assimilable à la vie (*Health SLT*) de la formule standard. Or ce module ne bénéficie pas d'USP contrairement au module Santé non assimilable à la vie (*Health NSLT*) pour lequel les USP non-vie sont applicables.

De notre analyse des garanties rencontrées et des dispositions réglementaires, les contrats de garanties arrêt de travail peuvent être appréhendés dans le module Santé SLT au sein des risques de longévité et d'incapacité/invalidité. Pour le risque de « longévité », le niveau de choc de la Formule Standard est de -20% des taux de mortalité des invalides en capturant les effets défavorables sur les engagements d'invalidité en cours et d'invalidité en attente. Pour le risque d'« incapacité/invalidité » prise en compte conjoint des effets du choc de +35% la première année (+25 % à partir de la 2ème année) des taux d'incidence en incapacité et en invalidité en capturant les effets défavorables des engagements d'invalidité en attente et du choc de -20% des taux de sortie des incapables (des taux de maintien s'ils sont inférieurs à 50% ou des taux de rétablissement sinon), en capturant les effets défavorables sur les engagements d'incapacité et d'invalidité en attente.

Plusieurs acteurs du marché français, notamment les détenteurs de portefeuille de prêts immobiliers proposant des garanties d'arrêt de travail, ont souvent remontés à la place l'inadéquation des paramètres de chocs fixes proposés par la Formule Standard de la norme Solvabilité II à leurs profils de risque. Ces niveaux de chocs fixes semblent être très pénalisant et coûteux en terme de besoin de fonds propres. A ce propos, l'ACPR (Autorité de Contrôle Prudentiel et de Résolution) a émis une note[5]

[5] ACPR (2011), Solvabilité 2 : Principaux enseignements de la cinquième étude quantitative d'impact (QIS5)

relative aux enseignements du QIS5 où elle a fait part de ses constats quant au traitement du risque d'arrêt de travail parmi les différents participants.

De plus, l'application de chocs uniques exclus de facto la prise en compte des problèmes d'hétérogénéité des risques regroupés dans le module « *Health SLT* ». En effet, le cadre de la Formule Standard ne fait pas de différence a piori entre les contrats proposant par exemple la garantie d'arrêt de travail. Ainsi, même si le risque est le même, par exemple la cohabitation de contrats individuels vs collectifs d'assurance de prêts, et/ou de prévoyance santé proposant tous des garanties d'arrêt de travail, ces contrats sont structurellement très différents par rapport aux caractéristiques suivants : la sélection adverse en assurance (choix entre contrat individuel ou contrat groupe), la sélection médicale (dont les effets impactent la qualité et la gravité des risques), la duration de leur passif, le type et la spécificité des population assurées.

8. Modélisation du risque de dépendance sous Solvabilité II

8.1. La dépendance, un risque biométrique pas comme les autres

L'assurance du risque dépendance constitue une branche particulière en termes de risque en assurance. La frontière entre dépendance et problèmes de santé est poreuse, dans la mesure où ces limitations d'autonomie résultent souvent de problèmes de santé actuels ou passés. Tout comme l'assurance santé, les couvertures dépendances privées en Europe viennent compléter les prestations ou les garanties offertes par le régime public.

En France, les garanties dépendance commercialisées par les assureurs ont leur propre système d'évaluation du risque dépendance. Les montants des prestations et leur versement sont indépendants des prestations versées par la Sécurité Sociale. Les garanties proposées par les différentes couvertures qui existent sur le marché peuvent être de nature différente. Certains produits d'assurance de prévoyance collective entreprise prévoient des garanties annuelles pour couvrir le risque de dépendance dans leur contrat. Lors du passage à la retraite le contrat peut se poursuivre ou non. La majorité des contrats individuels ou collectifs à adhésion facultative proposent des garanties viagères. Au bout d'une période définie au contrat, l'arrêt du paiement des cotisations n'entraine plus la résiliation du contrat mais sa mise en réduction. Les contrats proposent une couverture financière dont le montant est défini et ne prévoient pas l'indemnisation de frais de soin dont le montant pourrait évoluer dans le temps.

La tarification et le provisionnement des contrats dépendent naturellement de la nature de la garantie (viagère ou annuelle) et du type de données disponibles à la tarification. Si les garanties annuelles peuvent être tarifées et provisionnées à partir du suivi du ratio des sinistres survenus ramenés aux primes émises, la tarification des garanties viagères va nécessiter un nombre important d'hypothèses. Lorsque les données disponibles le permettent, les assureurs se fondent sur l'observation des sinistres passés pour tarifier leurs garanties décès ou incapacité. Dans le cas de la dépendance, le caractère très long terme des garanties proposées rend la collecte de ces données longues et entraine la nécessité de prendre en compte d'autres facteurs. En effet, le risque dépendance est lié à de nombreux paramètres sociétaux comme l'évolution de la cellule familiale, du style de vie et des liens intergénérationnels, l'évolution des pathologies, des modes de vie, de l'accès au soin. Les hypothèses posées lors de l'établissement du tarif font intervenir la mortalité de la population assurée, la probabilité de devenir dépendant selon différent niveau de dépendance appelée « état de dépendance », la durée de maintien en dépendance par état, les probabilités de transition entre états. La complexité et le nombre de ces hypothèses font que les assureurs ne peuvent se reposer sur la validité de leur tarif.

La dimension temps prend un rôle tout à fait fondamental dans ce type de garantie. Des progrès médicaux dans le traitement de cette maladie pourraient donc avoir pour effet un changement rapide des

probabilités de dépendance ou au contraire certains styles de vie pourraient entrainer une augmentation de la survenance de certaines pathologies et auraient un impact plusieurs années après. Pour toutes ces raisons, un pilotage régulier et un ajustement des garanties et ou des prestations des contrats dépendance sont indispensables tout au long de la durée des garanties. Le suivi et le pilotage des garanties est indispensable dans un monde en perpétuelle évolution. La nature long terme des garanties offertes par un contrat dépendance et le délai important qui en résulte entre l'encaissement de la première prime et le paiement des sinistres entraine naturellement un coût potentiel élevé du simple fait que le cadre de projection prévu dans le cas de Solvabilité II prévoit le maintien de la situation défavorable jusqu'à extinction des polices. Les garanties long terme comme peuvent l'être dans certains cas les garanties dépendance voient donc leurs chroniques de résultats durablement impactée par les dérives de sinistralité.

En outre, le caractère long terme entraine une dépendance plus grande de la rentabilité à l'environnement de taux. L'équilibre financier devient un enjeu lorsque plusieurs années, voire dizaine d'année s'écoulent entre l'encaissement des primes et le paiement des prestations. Dans la pratique, chez la plupart des acteurs, une fois le risque identifié l'assureur dispose d'un ensemble de mesures possibles. Premièrement, certains contrats collectifs peuvent même être arrêtés sur décision de l'assureur. Deuxièmement, si les dispositions contractuelles le permettent, la possibilité de revoir les primes des assurés cotisants à la hausse pour compenser l'évolution du risque. Troisièmement, la valeur de la rente versée pour les contrats réduits ou en cas d'arrêt de versement des primes pour les contrats en cours, selon les dispositions prévues au contrat, peut être revue à la baisse. Quatrièmement, la tarification des nouveaux contrats peut inclure une marge de prudence permettant d'assurer en vertu du principe de mutualisation un retour à l'équilibre technique. Le cadre prévu par la norme Solvabilité II ne permet pas de tenir compte de ce dernier mécanisme puisque les affaires nouvelles ne font pas parties du cadre de projection. Les autres mesures, peuvent être modélisées et prises en compte dans le calcul des flux projetés s'il existe une stratégie décrite et validée par le Conseil d'Administration pour faire face à ce type de situation. Si cette stratégie est mise en œuvre dans le calcul de la solvabilité de la compagnie, un rapport régulier de son impact sur les résultats et de son déclenchement dans les projections doit être fait.

Récemment, plusieurs auteurs ont réalisé une publication reprenant certaines thématiques traitant du risque de la perte d'autonomie et de dépendance. Ils proposent des bases techniques pour servir de référence en la matière à partir d'une définition médicale de la dépendance et des bases nationale d'hospitalisation (PMSI) 2008-2013. Castaneda et al. (2018) [46] ont réalisé un panorama de l'assurance dépendance en France. Schwarzinger (2018) [47] présente les données source et les retraitements pour l'étude du risque perte d'autonomie. Guibert et al. (2018) [48], [49], [50] présentent différentes études réalisées sur la population en France Métropolitaine de mesure de l'espérance de vie sans dépendance totale, de mesure du risque de perte d'autonomie totale et de mesure de l'espérance de vie en dépendance totale.

8.2. Modélisation du risque de dépendance sous la norme Solvabilité II

La dépendance étant un risque nouveau, croissant avec l'âge et mal connu des assureurs, ces derniers ont généralement recours à la réassurance proportionnelle pour plusieurs raisons. D'abord, ce risque est mal appréhendé du fait du manque de données sur le marché (risque relativement nouveau) et de l'absence de tables réglementaires. Ensuite, l'engagement de l'assureur peut être très élevé du fait de la durée de l'engagement (long terme) et entraine un capital de solvabilité requis important. Enfin, les effets sur la diminution du capital de risque sont relativement aisés à prendre en compte. En termes de réduction du capital de solvabilité induit par la réassurance, il n'existe pas de plafond sous Solvabilité II. Ainsi, la réassurance peut être pleinement prise en compte pour réduire le risque. Toutefois, les portées des clauses contractuelles ne doivent pas être omises. En effet, l'assureur doit accorder de

l'importance aux clauses de résiliations, aux clauses de participation aux bénéfices ainsi qu'aux conditions de transfert des actifs.

Dans le troisième chapitre du manuscrit, nous avons principalement détaillé la méthodologie de modélisation actuarielle et d'évaluation du besoin en capital relatif au risque de souscription du risque dépendance. Nous évoquons également sur les difficultés liées à l'application de la formule standard au calcul des fonds propres économique pour la garantie dépendance. Nous soulevons plus spécifiquement les problématiques de la prise en compte de la faculté de révision des primes, de la revalorisation des contrats et de l'indexation des primes, dans les calculs de la meilleure estimation des provisions. Même si ce chapitre ne traite pas en profondeur des SCR marché (*SCR Market*) et SCR opérationnel, il conviendrait de compléter la vision risque de souscription, par une évaluation des risques financiers et opérationnels. Ainsi, l'assureur disposera d'une évaluation holistique du capital requis pour faire face aux risques sous-jacents d'une couverture dépendance. Comme nous l'avons mentionné dans ce chapitre, les règles de calcul du capital requis sous Solvabilité II, norme prudentielle en vigueur depuis le 1er janvier 2016, n'ont pas été prévues spécifiquement pour le risque dépendance. Elles semblent donc mal adaptées à ce risque, ce qui pose un certains nombres de problèmes de cohérence et de mesure du niveau réel du besoin de capital pour ce risque. Néanmoins, les autorités de contrôles offrent aux organismes assureurs la possibilité de déroger à ces règles en ayant recours à un modèle interne partiel. Il faut noter qu'en pratique, à ce jour la modélisation interne partielle n'est pas envisageable pour la plupart des acteurs, car ils sont confrontés au manque de statistiques et de données d'expérience indispensables à cet exercice. En conséquence, la majorité des organismes français ne sont pas satisfaits du traitement du risque dépendance dans Solvabilité II mais se retrouvent contraints à appliquer la formule standard qui génère un besoin en capital très élevé. Etant donné la duration longue des engagements et le niveau important du SCR souscription associé, il est indispensable pour un assureur de détenir un modèle de projection prenant en compte des actions de management. Ces choix permettent de prendre en compte les leviers de pilotage à déclencher afin de calculer le besoin en capital adéquat. Ainsi, l'intégration de management actions est un moyen pour l'organisme assureur de répondre aux exigences prudentielles en assurant la pérennité d'un régime comportant des engagements viagers. Enfin, l'Autorité Européenne des Assurances et des Pensions Professionnelles a publié le 28 février 2018 de nouvelles recommandations pour amender les actes délégués servant au calcul de la formule standard du SCR. Malheureusement, ces recommandations n'apportent pas de modifications quant au traitement de la dépendance viagère. Toutefois, notons qu'une nouvelle révision de formule standard est prévue pour 2020.

9. Approche socioculturelle des attitudes face au risque

9.1. Cadre théorique

L'une des plus anciennes théories sur l'attitude des agents économiques face au risque est celle qui a fondé le modèle d'Espérance d'utilité (*Expected Utility Theory*) de von Neumann et Morgenstern (1944) [51]. Le grand défaut de ce modèle né de la théorie de l'Utilité espérée est son hypothèse fondée sur la rationalité des individus. Cette hypothèse a été contrecarrée par la théorie des perspectives (*Prospect Theory*) de Kahneman et al. (1979) [52]. Ils exposent et démontrent l'irrationalité des individus face aux choix risqués par l'intégration de la dimension psychologique dans leur analyse. Ces auteurs proposent une alternative basée sur la démarche expérimentale. Cette contribution a permis de mieux comprendre des biais cognitifs qui peuvent influencer l'attitude des agents face au risque. Malgré son éclairage, il n'y a pas de consensus général sur le niveau d'influence des profils individuels sur les décisions effectivement prises par les agents en situation risquée. En effet, lors de prise de décisions en environnement risqué, le profil d'attitude face au risque du décideur n'est pas le seul paramètre

déterminant. Ils existent différentes situations où les choix finaux dépendent de la prise en considération d'indicateurs externes (quantitatifs et qualitatifs), des interactions entre les parties prenantes, ainsi que de l'environnement économique. Finalement c'est le mélange de facteurs endogènes et exogènes qui joue lors des prises de décisions face au risque, lorsque les agents ont la responsabilité de se positionner face à la concurrence, aux changements des cycles économiques et aux incertitudes liées aux contextes des affaires.

D'autres disciplines se sont aussi intéressées à ce phénomène. La théorie socio-culturelle développée par l'anthropologie sociale établie un lien entre la perception du risque et les groupes socio-culturel. Inglehart et al. (2005) [53], (1997) [54], (1998) [55] ont mis en place une enquête mondiale construite sur la base des travaux menés précédemment par Kerkhofs et de Moor en Europe (*European Values Survey*, 1981), nommée la *World Values Survey* et englobant un plus d'une centaine de pays. A partir d'un modèle d'analyse des changements dans le temps entre les multiples vagues de l'enquête, Inglehart et al. ont identifié deux principaux facteurs bipolaires comme base structurelle de la culture. Le premier facteur oppose la survie aux valeurs d'auto-expression et le second facteur oppose l'autorité traditionnelle au séculaire/rationnelle. Ces travaux ont servi de base au développement de la théorie culturelle du risque.

Douglas et al. (1983) [56], co-auteurs des travaux sur le risque et la culture (*Risk and Culture*), ont proposé une catégorisation des groupes socio-culturels face au risque. Ils fournissent des explications sur les archétypes d'attitudes face aux risques et la position socio-culturelle des groupes humains. Ils considèrent que les positions culturelles peuvent se révéler plus influentes que le sexe ou l'âge et agir comme des facteurs de confusion. Ils distinguent ainsi dans toutes les sociétés humaines la présence de quatre groupes sociologiques correspondant à quatre visions du monde et des risques (les individualistes, les égalitaristes, les hiérarchiques et fatalistes).

A partir de ces quatre groupes socio-culturels, ils ont créé une matrice à deux dimensions. La première dimension de cette matrice, correspond à la structure d'organisation du groupe ou la « Grille » (ou en anglais « *Grid* »). Elle mesure le niveau de maturité de la structure organisationnelle du groupe. Dans les groupes sociologiques Individualiste et Egalitaire, l'organisation est très peu structurée et les individus très autonomes. A l'inverse dans les groupes sociologiques Fataliste et Hiérarchique, la structure impose une forme d'organisation. Les individus sont par conséquents plus dépendants du groupe et moins autonomes. La deuxième dimension, correspond au niveau d'incorporation de l'individu et de soutien social à l'individu ou le « Groupe » (« *Group* »). Un faible niveau d'incorporation et de soutien social à l'individu renvoie aux groupes sociologiques Individualiste et Fataliste. A contrario, dans les groupes sociologiques Egalitaire et Hiérarchique, le niveau d'incorporation et de soutien social est très élevé. Ainsi, ces deux dimensions influencent dans une certaine mesure les attitudes face au risque des individus et constituent des facteurs explicatifs acceptés par cette théorie culturelle du risque.

Dake (1991) [57], psychologue américain, a tenté de valider quantitativement cette théorie culturaliste face au risque. Sa thèse de doctorat intègre culture, psychologie et risque. Les observations de Dake s'appuient sur une méthodologie en trois étapes : la détermination des inquiétudes sociétales, celle des biais culturels et l'interprétation des résultats (voir Wildavsky and Dake, 1990 [58]).

9.2. Modèle paramétrique de détection des profils d'attitudes face au risque

C'est dans le cadre théorique des valeurs culturelles et des comportements des groupes sociologiques face au risque que Ingram et al. (2010) [59], [60], [61] proposent un modèle paramétrique liant les groupes socio-culturels à quatre catégories d'attitude face au risque. Dans la lignée de Douglas et al. (1983) [56], ils associent le groupe des « individualistes » aux « Maximisateurs », les « Egalitaires » aux « Conservateurs », les « Fatalistes » aux « Pragmatiques » et enfin les « Hiérarchiques » aux « Managers ».

Ils définissent les Maximisateurs (« *Maximizers* ») comme ceux qui pensent que l'augmentation des profits est plus importante que l'examen des risques. Les Managers (« *Managers* ») préfèrent s'assurer de l'équilibre minutieux entre les risques et les profits. Les Conservateurs (« *Conservators* ») sont ceux pour qui l'accroissement des gains n'est pas plus important que l'évitement des pertes. Enfin les Pragmatiques (« *Pragmatists* ») pensent que l'avenir est très peu prévisible et préfèrent éviter les engagements fermes afin de garder autonomie et flexibilité.

D'autre part, Ingram et al. (2009) [62], (2011) [63], (2012) [64] proposent un modèle paramétrique de détection des profils individuels face au risque, modèle calibré de manière empirique à partir d'avis d'experts et de données issu de sondages menés auprès de professionnels du secteur financier aux Etats-Unis. En plus des cas des individus classés par cette approche à l'une des quatre catégories d'attitude face au risque, ces auteurs découvrent également que le modèle proposé affecte d'autres individus à des profils mixtes. Ces profils mixtes sont des combinaisons entre les profils supposés disjoints des quatre catégories d'attitude face au risque. Ils expliquent la présence des profils mixtes dans la population par la théorie de l'adaptabilité ou des rationalités plurielles ou «*plurial rationalities theory*» (voir Ingram et al., (2011) [65]). Cette théorie admet qu'un individu peut préférer plusieurs catégories d'attitude face au risque en cohérence avec sa capacité à s'adapter à des situations et à des environnements différents.

10. Contributions et principaux résultats

Le chapitre 1 est un article coécrit avec Yassine Gaïd et Stéphane Loisel et soumis en 2019. Il traite de la prédiction des paramètres de provisionnement individuel avec des méthodes d'apprentissage ensembliste en assurance non vie. L'objectif principal de ce travail est de proposer des adaptations algorithmiques et de métriques de performance applicables aux méthodes d'apprentissage automatique ensemblistes permettant de prédire des paramètres utilisés pour le provisionnement individuel des sinistres en assurance non vie. En présence de données de durée de survie et d'observations censurées à droite, les améliorations proposées introduisent un système de pondération fonction des poids de Kaplan-Meier comme hyper paramètres des algorithmes. Ces modifications ont conduit à adapter conjointement les métriques de mesure de la performance des modèles. Pour des raisons évidentes de recherche d'algorithmes permettant d'obtenir des prédictions stables, nous utilisons les techniques de *bagging* (*Bootstrap aggregating*) et de *boosting*. Le calibrage des modèles à partir de ces algorithmes modifiés est réalisé sur des données individuelles d'effectifs suffisants issues des bases de gestion d'un assureur de personnes. Nous appliquons notre méthodologie à deux portefeuilles (de prêts et de prévoyance collective) assez différents par leur volumétrie, par les risques sous-jacents aux garanties assurées, par les variables explicatives et la période d'historique disponibles. La base des contrats de prêts dispose de 879 553 sinistrés sur une période d'observation comprise entre 2004 et 2016. Celle des contrats de prévoyance collective contient 347 904 sinistrés sur une période d'observation allant de 2014 à 2017. Les paramètres de provisionnement prédits sont les durées de maintien et des charges sinistres ultimes individuelles. Les indicateurs de validation sont cohérents, plus précis et plus stables que ceux fournis par l'algorithme des arbres de classification et de régression modifié proposé par Lopez et al. (2015). Ces résultats renforcent l'intérêt général de notre nouvelle méthode.

Le chapitre 2 correspond à l'article coécrit avec Stéphane Loisel et Jérémie Zozime et soumis en 2019. Il propose une approche quantitative d'estimation de chocs des risques d'incidence et de maintien en assurance non-vie sous Solvabilité II. Nous y présentons dans le cadre de la norme Solvabilité II, une méthodologie permettant de calibrer les paramètres spécifiques à l'entité (*Undertaking Specific Parameters*, USP). Dans cette étude, nous traitons le cas des risques d'incidence et de maintien de la garantie arrêt de travail d'un portefeuille de contrats de prêts (composé de prêts immobiliers et à la consommation). Nous illustrons notre approche à partir de données réelles provenant des bases de

gestion d'un assureur de personne. Le volume de données du portefeuille considéré permet de disposer de plusieurs générations de contrats de taille (un total de plus de 12 millions de polices souscrites) et de profondeur d'historique variable (avec des générations de contrats entre 2 années et dépassant 20 années d'ancienneté). La principale contribution de ce chapitre est de fournir un cadre opérationnel de calibration d'USP du module « *Health SLT »* pour les risques d'incidence et de maintien (ou le rétablissement) des garanties d'incapacité et d'invalidité en assurance prévoyance. Ce travail apporte une modélisation alternative aux chocs forfaitaires fixés de manière arbitraire par la formule standard de la norme Solvabilité II.

Pour estimer les niveaux des chocs à un an, nous retiendrons une approche intégrant l'erreur d'estimation et l'erreur *process variance*. L'erreur d'estimation correspond à la variance du risque. L'erreur *process variance* est estimée suivant deux types de modèles. Un modèle basé sur les séries temporelles (modèles Autorégressifs et Moyenne-mobile d'ordres (p, q) ou ARMA(p,q)) et un modèle linéaire généralisé (GLM) de type Log-Normal. Le choix des séries chronologiques est fondé sur leurs applications aux nombreux domaines (la prévision d'indices économiques en économie, l'évolution des cours de la bourse en finance, l'analyse de l'évolution d'une population en démographie, l'analyse de données sismiques en géophysique…). Ces modèles s'adaptent en général bien à la problématique posée. Il conviendrait, suivant le niveau de prudence souhaitée dans les calculs, d'estimer l'erreur d'estimation comme étant la variance de l'estimateur de la loi moyenne retenue. Cette précaution sur l'évaluation de l'erreur d'estimation est particulièrement nécessaire dans le cas où l'approche GLM est retenue.

L'une des principales difficultés d'estimation des paramètres des modèles retenus est la disponibilité de données de taille et d'historique suffisants. Ces données doivent également répondre aux critères de qualité (exhaustivité, exactitude, pertinence) préconisés par la norme Solvabilité II. Du point de vue théorique, pour pallier à l'insuffisance de profondeur d'historique des données, la théorie de crédibilité de Bülhmann (1967) [66], (1969) [67] ou théorie de la crédibilité fondée sur la plus grande exactitude (TCGE) est préférable, car elle est complète. La réalité des informations disponibles, la gestion de leur intégrité dans le temps, les limites et l'insuffisance des données d'expérience de la plupart des compagnies ne permettent pas sa mise en œuvre opérationnelle. Nous proposons donc une variante normalisée de la théorie de la crédibilité à variation limitée (TCVL) ou crédibilité américaine. Il s'agit d'une approche préconisée par l'Institut des Actuaires Canadiens dans sa note éducative (2002) [68], (2014) [69] pour résoudre des problématiques similaires sur les données afin d'établir des hypothèses d'évaluation de mortalité d'expérience lorsque les résultats sont crédibles, mais qu'il n'est pas possible de créer une table pour la compagnie. L'introduction de la crédibilité TCVL dans la modélisation, permet d'estimer des niveaux de chocs par ancienneté de la police dans le portefeuille.

L'application de ces modèles à des données réelles conduit à des réductions de SCR par rapport à la Formule Standard de l'ordre de -60% (pour des contrats disposant un historique de données de plus de 10 années) et de l'ordre de -30% (pour des contrats disposant d'un historique de données de moins de 5 années). L'approche proposée pourrait s'adapter plus généralement aux risques de prévoyance du module « *Health SLT* », ainsi qu'au calibrage de chocs propres à l'entité pour les calculs de type ORSA (*Own Risk and Solvency Assessment*), fournir des paramètres de sensibilité pour les business plans ou la tarification. L'introduction des facteurs de crédibilité est une approche prudente car permettant de tenir compte au moins partiellement de l'expérience de la compagnie. Finalement, les limites et points d'attention que nous relevons concernent la problématique de l'application des coefficients de crédibilité calibrés sur lois centrales aux paramètres de chocs. De plus, l'approche série temporelle nécessite de disposer d'historique conséquent pour son utilisation. Elle peut donc ne pas être adaptée dans la plupart des cas, ce qui conduirait à l'utilisation de l'approche Log Normale qui fournit des paramètres de chocs un peu plus sévère.

Le troisième chapitre correspond à un chapitre de livre (sur la dépendance) publié en juin 2019. Cet article coécrit avec Camille Gutknecht (2019) traite des calculs des fonds propres économique des pro-

duits de dépendance suivant la formule standard de la norme Solvabilité II (« *Long-Term Care : Construction of an economic balance sheet and solvency capital requirement calculation in solvency II* »). Nous avons principalement détaillé la méthodologie de modélisation actuarielle et d'évaluation du besoin en capital relatif au risque de souscription du risque dépendance. Nous évoquons également sur les difficultés liées à l'application de la formule standard au calcul des fonds propres économique pour la garantie dépendance. Nous soulevons plus spécifiquement les problématiques de la prise en compte de la faculté de révision des primes, de la revalorisation des contrats et de l'indexation des primes, dans les calculs de la meilleure estimation des provisions. Même si ce chapitre ne traite pas en profondeur des SCR marché (*SCR Market*) et SCR opérationnel, il conviendrait de compléter la vision risque de souscription, par une évaluation des risques financiers et opérationnels.

Ainsi, l'assureur disposera d'une évaluation holistique du capital requis pour faire face aux risques sous-jacents d'une couverture dépendance. Comme nous l'avons mentionné dans ce chapitre, les règles de calcul du capital requis sous Solvabilité II, norme prudentielle en vigueur depuis le 1er janvier 2016, n'ont pas été prévues spécifiquement pour le risque dépendance. Elles semblent donc mal adaptées à ce risque, ce qui pose un certains nombres de problèmes de cohérence et de mesure du niveau réel du besoin de capital pour ce risque. Néanmoins, les autorités de contrôles offrent aux organismes assureurs la possibilité de déroger à ces règles en ayant recours à un modèle interne partiel. Il faut noter qu'en pratique, à ce jour la modélisation interne partielle n'est pas envisageable pour la plupart des acteurs, car ils sont confrontés au manque de statistiques et de données d'expérience indispensables à cet exercice. En conséquence, la majorité des organismes français ne sont pas satisfaits du traitement du risque dépendance dans Solvabilité II mais se retrouvent contraints à appliquer la formule standard qui génère un besoin en capital très élevé.

Etant donné la duration longue des engagements et le niveau important du SCR souscription associé, il est indispensable pour un assureur de détenir un modèle de projection prenant en compte des actions de management. Ces choix permettent de prendre en compte les leviers de pilotage à déclencher afin de calculer le besoin en capital adéquat. Ainsi, l'intégration de management actions est un moyen pour l'organisme assureur de répondre aux exigences prudentielles en assurant la pérennité d'un régime comportant des engagements viagers. Enfin, l'Autorité Européenne des Assurances et des Pensions Professionnelles a publié le 28 février 2018 de nouvelles recommandations pour amender les actes délégués servant au calcul de la formule standard du SCR. Malheureusement, ces recommandations n'apportent pas de modifications quant au traitement de la dépendance viagère. Toutefois, notons qu'une nouvelle révision de formule standard est prévue pour 2020.

Enfin, le quatrième chapitre correspond à l'article coécrit avec Denis Clot, David Ingram et Stéphane Loisel (2019) est une étude empirique des attitudes face au risque dans le secteur de la banque et assurance. Cet article est en cours de soumission. Le modèle paramétrique proposé par Ingram et al. (2014) sur l'expérience d'un panel issu de la zone Amérique permet de détecter les profils d'attitude face au risque. Partant de cette approche empirique et basée sur des jugements d'experts, les objectifs de nos travaux sont doubles. D'abord, généraliser ce modèle afin de l'appliquer à d'autres zones géographiques. Ensuite, tester statistiquement sur notre panel d'étude l'existence de liens d'association éventuels entre les profils et certaines variables sociodémographiques (secteur d'activité, position occupée dans l'entreprise, le genre, l'ancienneté professionnelle). D'un point de vue méthodologique, nous avons utilisé le même questionnaire conçus par l'équipe Ingram et al. (2014) pour réaliser les enquêtes en Europe et en Afrique auprès de professionnels des secteurs financiers (notamment en Banque et en Assurances) entre 2013 et 2015. L'application du modèle généralisé permet de constater qu'il existe une opposition de préférence d'attitude face au risque entre la zone Europe et la zone Amérique. La zone Europe est liée positivement mais avec une faible intensité avec les catégories de profils « *Managers* » et « *Pragmatists* » contrairement à la zone Amérique. La zone Amérique n'a aucun lien d'association avec les catégories de profils mixtes « *Conservators/Managers* » et « *Conservators/Pragmatists* », mais reste lier positivement avec les autres catégories de profils d'attitude face au

risque contrairement à la zone Europe. Le panel de la zone Afrique composé de salarié d'une banque nationale a un lien d'association positif mais faible avec les catégories « *Conservators/Managers* » et « *Managers/Pragmatists* ». Nous notons que les liens d'association entre les positions occupées dans l'entreprise (respectivement le secteur d'activité) et les catégories de profils d'attitude face au risque sont positives et faibles, mais a contrario il n'existe aucun lien avec les variables sociodémographiques (genre et ancienneté professionnelle) pour les zones Europe et Afrique. Enfin, ces résultats restent partiellement cohérents avec les conclusions de la théorie socioculturelle face au risque.

Nous notons quelques limites dans l'interprétation des résultats de cette étude. La taille des échantillons demeure modeste et nécessite la collecte de nouvelles données. Toutes les informations ne sont pas disponibles (données manquantes sur des variables explicatives), ce qui limite les analyses et l'approfondissement des liens entre attitude et facteurs explicatifs. Enfin, une telle étude nécessite plusieurs vagues de sondages pour valider les principaux résultats dans le temps, notamment l'hypothèse de l'adaptabilité plurielle. Ainsi, pour s'assurer de la validité de ces résultats, à d'autres populations, nous préconisons la réalisation d'une étape de validation préalable sur leurs données par l'application de la méthodologie proposée.

Ce travail pourrait servir d'étape structurante de catégorisation des individus lors des études d'analyse des comportements face au risque, avant la mise en œuvre d'un programme d'expérimentation plus coûteuse. Les résultats obtenus pourraient également aider lors des discussions au sein des Comités de Risque et des Conseils des organisations du secteur financier afin de rendre plus agile leur programme *Entreprise Risk Management* (ERM). Cette agilité est nécessaire pour permettre aux décideurs d'aligner leurs stratégies de prise et de gestion des risques face aux différents cycles économiques.

References

[1] Lopez, O. Milhaud, X. Therond, P., (2016). “Tree-based censored regression with applications in insurance”, Electronic Journal of Statistics, (Oct. 2016) Volume 10 issue 2, pp.2685-2716.

[2] Crockford, G.N. (1982). “The Bibliography and History of Risk Management: Some Preliminary Observations”, The Geneva Papers on Risk and Insurance, 7, 169-179.

[3] Harrington, S., Niehaus, G.R. (2003).” Risk Management and Insurance”, Irwin/McGraw-Hill, USA.

[4] Williams, A., Heins M. H. (1995). “Risk Management and Insurance”, McGraw-Hill, New York.

[5] Mehr, R. I., Hedges, B. A. (1963). “Risk Management in the Business Enterprise”, Irwin, Homewood, Illinois.

[6] Williams, A., Heins M. H. (1964). “Risk Management and Insurance”, McGraw Hill, New York.

[7] Harrington, S., Niehaus, G.R. (2003). “Risk Management and Insurance”, Irwin/McGraw-Hill, USA.

[8] Dionne, G. (2013). “Gestion des risques: histoire, définition et critique”, Cahier de Recherche 13-01

[9] Blanchard, D., Dionne, G. (2004). “The Case for Independent Risk Management Committees”, Risk 17, 5, S19-S21

[10] Philippe Saint Pierre (2005). “Modèles multi-états de type Markovien et application à l’asthme”. Sciences du Vivant [q-bio]. Université Montpellier I, 2005. Français. <tel-00010146>

[11] Hougaard, P. (1999). “Multi-State Models: A Review”. Lifetime Data Analysis, 5, 239-264.

[12] Huber-Carol C. et Pons O. (Mai 2004). “Independent competing risks versus a general semi-Markov model: application to heart transplant data”. Prépublication. URL http://www.math-info.univ-paris5.fr/map5/publis/PUBLIS04/huber-2004-13.pdf.

[13] Andersen P. K., Borgan Ø., Gill R. D. et Keiding N. (1993). “Statistical models based on counting processes”. Springer-Verlag.

[14] Therneau T. M. et Grambsch P. M. (2000). “Modeling Survival Data: Extending the Cox Model”. Springer – Statistics for Biology and Health.

[15] Andersen P. K. et Keiding N. (2002). “Multi-state models for event history analysis”. Statistical Methods in Medical Research, vol. 11, n◦2. pages 91–115.

[16] Commenges D. (1999). “Multi-State Models in Epidemiology”. Lifetime Data Analysis, vol. 5. pages 315–327.

[17] Cox D. R. (1972). “Regression models and life tables (with discussion)”. J Royal Statistical Soc B, vol. 34. pages 187–220.

[18] Hougaard P. (1995). “Frailty models for survival data”. Lifetime Data Analysis, vol. 1. pages 255–273.

[19] Huber-Carol C. et Vonta I. (2004). “Fraitly models for arbitrarily censored and truncated data”. Lifetime Data Analysis, vol. 10. pages 369–388.

[20] Hougaard P. (2000). “Analysis of multivariate survival data”. Springer – Statistics for biology and health.

[21] Nielsen G. G., Gill R. D., Andersen P. K. et Sørensen T. I. A. (1992). “A counting process approach to maximum likelihood estimation in frailty models”. Scandinavian Journal of Statistics, vol. 8. pages 25–43.

[22] Commenges D. (2002). “Inference for multi-state models from interval-censored data”. Statistical Methods in Medical Research, vol. 11, n◦2. pages 167–182.

[23] Robins J. M. et Finkelstein D. M. (2000). “Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests”. Biometrics, vol. 56, n◦3. pages 779–788.

[24] Minini P. et Chavance M. (2004). “Sensitive analysis of longitudinal normal data with drop-outs”. Statistics in Medicine, vol. 23. pages 1039–1054.

[25] Little R. J. A. (1995). “Modelling the drop-out mechanism in repeated-measures studies”. Journal of the American Statistical Association, vol. 90. pages 1112–1121.

[26] Planchet, F. (2011). “Bootstrap et méthodes de triangles”. Actudactuaires.typepad.com/laboratoire/2011/07/bootstrap-et-méthodes-de-triangles.html.

[27] Mack, T. (1993). “Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates”. Astin Bulletin, Vol 23, no. 2, 213-225.

[28] Mack, T. (1999). “The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor”. Astin Bulletin, Vol 29, no. 2, 361-366.

[29] Mack, T. (2000). “Credible Claims Reserves: the Benktander Method”. Astin Bulletin, Vol 30, n°2.

[30] Mack, T. (2008). “The Prediction Error of Bornhuetter-Ferguson”. Casualty Actuarial Society E-Forum, Fall 2008, 222-240.

[31] Bornhuetter, R.L., Ferguson, R.E. (1972). “The actuary and IBNR”. Proc Casualty Actuarial Society, Vol. LIX, 181-195.

[32] Wüthrich (2018), Neural Networks Applied to Chain-Ladder Reserving, European Actuarial Journal, December 2018, Volume 8, Issue 2, pp 407–436

[33] Hothorn, T., Lausen, B., Benner, A., Radespiel-Tröger, M., (2004). “Bagging survival trees”, Stat. Med. 23 (1) 77–91.

[34] Ishwaran, H., Kogalur, U.B., Blackstone, E.H., Lauer, M.S., (2008). “Random survival forests”, Ann. Appl. Stat. 841–860.

[35] Ibrahim, N.A., Abdul Kudus, A., Daud, I., Abu Bakar, M.R., (2008). “Decision tree for competing risks survival probability in breast cancer study”, Int. J. Biol. Med. Sci. 3 (1), 25–29.

[36] Lucas, P.J.F., van der Gaag, L.C., Abu-Hanna, A., (2004). “Bayesian networks in biomedicine and health-care, Artif. Intell”. Med. 30 (3), 201–214.

[37] Bandyopadhyay, S., Wolfson, J., Vock, D.M., Vazquez-Benitez, G., Adomavicius, G., Elidrisi, M., Johnson, P.E., O’Connor, P.J., (2015). “Data mining for censored time-to-event data: a Bayesian network model for predicting cardiovascular risk from electronic health record data”, Data Min. Knowl. Disc. 29 (4) 1033–1069.

[38] Vock, D. M., Wolfson, J., Bandyopadhyay, S., Adomavicius, G., Johnson, P. E., Vazquez-Benitez, G., & O’Connor, P. J. (2016).” Adapting machine learning techniques to censored time-to-event health record data : A general-purpose approach using inverse probability of censoring weighting”. Journal of biomedical informatics, 61, 119-131.

[39] Breiman, L., Friedman, J., Olshen, R., Stone, C., (1984). “Classification and regression trees”. Wadsworth Books, 358.

[40] Kaplan E.L., Meier P., (1958), « Nonparametric Estimation from Incomplete Observations », Journal of the American Statistical Association, vol. 53, n°282, pp. 457 481.

[41] Bang, H., Tsiatis, A.A., (2000). “Estimating medical costs with censored data”, Biometrika 87 (2), 329 -343.

[42] Tsiatis, A.A., (2006). “Semiparametric Theory and Missing Data”, Springer, New York.

[43] Stute, W. (1993). “Consistent estimation under random censorship when covariables are present”. Journal of Multivariate Analysis, 45(1), 89-103.

[44] Breiman, L., (2001). “Random Forests”. Machine Learning 45 (1): 5-32. doi :10.1023/A: 1010933404324

[45] Friedman, J-H., (2002). “Stochastic gradient boosting”, Computational Statisrics and Data Analysis 38.

[46] Castaneda, F., Lusson, F. (2018). “Un panorama de l’assurance dépendance en France”. Bulletin Français d’Actuariat, Vol 18 N°35, janvier-décembre 2018.

[47] Schwarzinger, M., (2018). “Etude QalyDays: données source et retraitements pour l’étude du risque perte d’autonomie”. Bulletin Français d’Actuariat, Vol 18 N°35, janvier-décembre 2018.

[48] Guibert, Q., Planchet, F., Schwarzinger, M. (2018). “Mesure de l’espérance de vie sans dépendance totale en France Métropolitaine”. Bulletin Français d’Actuariat, Vol 18 N°35, janvier-décembre 2018.

[49] Guibert, Q., Planchet, F., Schwarzinger, M. (2018). “Mesure du risque de perte d’autonomie totale en France Métropolitaine”. Bulletin Français d’Actuariat, Vol 18 N°35, janvier-décembre 2018.

[50] Guibert, Q., Planchet, F., Schwarzinger, M. (2018). “Mesure de l’espérance de vie en dépendance totale en France Métropolitaine”. Bulletin Français d’Actuariat, Vol 18 N°35, janvier-décembre 2018.

[51] Neumann, John von; Morgenstern, Oskar (1953) [1944]. “Theory of Games and Economic Behavior” (Third ed.). Princeton, NJ: Princeton University Press.

[52] Kahneman D., Tversky A., (1979). “Prospect theory: an analysis of decision under risk”, Econometrica, 47, 263-91.

[53] Inglehart, R., Welzel, C. (2005). “Modernization, Cultural Change and Democracy”. Cambridge, Cambridge University Press, 2005 (ISBN 978-0-521-84695-0)

[54] Inglehart, R., (1997). “Modernization and postmodernization: Cultural”, economic and political change in 43 societies. Princeton, NJ: Princeton University Press.

[55] Inglehart, R., Basañez, M., & Moreno, A., (1998). “Human Values and Beliefs: A Cross-Cultural”. Sourcebook, Ann Arbor.

[56] Douglas, M., Wildavsky, A., (1983). “Risk and culture, an essay on the selection of technical and environmental dangers”, Berkeley, University of California Press.

[57] Dake, K., (1991), “Orienting Dispositions in the Perception of Risk: An Analysis of contemporary Worldviews and Cultural Biases”.

[58] Wildavsky, A., Dake, K. (1990). “Theories of Risk Perception: Who Fears What and Why?”. The MIT Press, Vol. 119, No. 4, Risk (Fall, 1990), pp. 41-60.

[59] Ingram D., Thompson E., (2010). “A Universe in Four Parts”. Wilmott Magazine, March 2010.

[60] Ingram D., Thompson E., Tayler, (2010). “Eyes Wide Open: Towards Rational Adaptability”. Wilmott Magazine, July 2010.

[61] Ingram D., Underwood A., (2010). “Full Spectrum of Risk Attitude”. Actuary Magazine, August 2010.

[62] Ingram D., (2009). “Four Seasons of Risk Management”. Actuary Magazine, December 2009.

[63] Ingram D., Thompson E., (2011) « Changing Seasons of Risk Attitudes ». Actuary Magazine, April 2011

[64] Ingram D., Thompson E., (2012) « What’s Your Risk Attitude? ». Harvard Business Review Blog, June 2012

[65] Ingram D., Underwood A., (2011) « The Human Dynamics of the Insurance Cycle and Implications for Insurers: An Introduction to the Theory of Plural Rationalities ». ERM Symposium/SOA Monograph, January 2010

[66] Bühlmann, H., (1967). «Experience rating and credibility», ASTIN Bulletin, vol. 4, p. 199–207

[67] Bühlmann, H., (1969). «Experience rating and credibility», ASTIN Bulletin, vol. 5, p. 157–165

[68] Canadian Institute of Actuaries, (2002). “Commission des rapports financiers des compagnies d’assurances vie”, Note éducative, mortalité prévue: polices canadiennes d’assurances vie individuelle avec tarification complète.

Chapitre 1 :
Prédiction des paramètres de provisionnement individuel avec des méthodes d'apprentissage ensemblistes en assurance non vie

CHAPITRE 1 : PREDICTION DES PARAMETRES DE PROVISIONNEMENT INDIVIDUEL AVEC DES METHODES D'APPRENTISSAGE ENSEMBLISTES EN ASSURANCE NON VIE

1. Introduction

En assurance non-vie, ils existent deux principales méthodes de provisionnement. L'approche classique dite *Macro-Level Reserving* et l'approche plus fine dite *Micro-Level Reserving*.

Les méthodes *Macro-Level Reserving* sont des techniques basées sur des triangles de liquidation construits à partir de données agrégées pour la projection des règlements futurs. Parmi ces méthodes agrégées permettant d'estimer de la charge sinistre totale on peut citer les méthodes de *Chain-Ladder*, du *Loss-Ratio*, de Bornhuetter-Ferguson, etc… La plus populaire d'entre elles a été formalisée dans les années 30 et connue sous le nom de la méthode de *Chain Ladder* (voir Astesan, 1938 [1]). L'utilisation de données agrégées a pour avantage de fournir des résultats faciles à comprendre et à expliquer. L'implémentation opérationnelle des calculs est aisée. Ils permettent d'estimer la charge totale de tous les sinistres survenus à une date comptable dites provisions IBNR (*Incurred But Not Reported*, IBNR). Ces provisions sont composées d'une provision complémentaire destinée à couvrir l'insuffisance des provisions dossier par dossier des sinistres ouverts, les IBNeR (*Incurred But Not Enough Reported*) et d'une provision complémentaire pour couvrir les dossiers survenus mais dont l'assureur n'a pas encore connaissance, les IBNyR (*Incurred But Not Yet Reported*). De par sa simplicité, ces méthodes présentent certaines limites importantes. D'une part, elles sont soumises aux fortes hypothèses sous-jacentes, sur la stabilité des sinistres dans le temps et l'indépendance des survenances entre les années. En réalité, ces hypothèses sont très rarement vérifiables, ce qui nécessite de grandes précautions d'usage. D'autre part, elles sont confrontées aux problèmes d'hétérogénéité auxquels sont soumises les méthodes agrégées, puisqu'elles n'exploitent pas la richesse des données individuelles. Enfin, ces méthodes agrégées ne sont pas adaptées à la prise en compte d'observations censurées propres à l'analyse des données de survie.

Plus récemment, il y a eu un intérêt croissant pour les approches ascendantes (*bottom-up*) et plusieurs auteurs comme Arjas (1989) [2], Jewell (1989) [3], Norberg (1993) [4], Hesselager (1994) [5], Antonio-Plat (2014) [6], Badescu et al. (2016) [7, 8], Baudry-Robert (2017) [9], Harej et al. (2017) [10], Jessen et al. (2011) [11], Lopez (2018) [12], Pigeon et al. (2013) [13], Verrall-Wüthrich (2016) [14] et Zarkadoulas (2017) [15] ont proposés des méthodes dites *Micro-Level Reserving*. Ces travaux visent à modéliser les sinistres au niveau micro-level en utilisant des modèles paramétriques et non-paramétriques. Ainsi, les provisions pour sinistres du portefeuille sont obtenues par agrégation des provisions individuelles par sinistre. Parmi ces méthodes, les modèles paramétriques sont les plus faciles à expliquer. Ils ont en plus l'avantage d'utiliser les données individuelles, lorsqu'elles sont disponibles. Cependant, elle présente d'autres limites liées à leurs implémentations, à leurs validations et à l'estimation de leurs paramètres. En pratique les modèles paramétriques sont plus difficiles à implémenter que les méthodes agrégées présentées ci-dessus. De plus, leurs performances sont jugées sur leur qualité asymptotique plutôt que sur leur adéquation aux données réelles. Enfin, ils existent des risques liés aux erreurs d'estimation des paramètres surtout lorsque plusieurs covariables sont retenues comme variables explicatives du risque. Ces modèles, paramétrique et non-paramétriques, peuvent incorporer l'hétérogénéité, les changements structurels, etc. (voir Friedland, 2010 [16]).

Wüthrich (2018) [17] propose une approche descendante (*top down*) consistant à différencier les types sinistres dans le cadre du modèle *Chain Ladder* de Mack (1993) [18]. Il explore le cadre théorique des réseaux de neurones pour la modélisation des différences structurelles, en remplaçant l'hypothèse de régression simplifiée du modèle de Mack par un modèle de régression (non linéaire) de réseaux de

neurones pour expliquer les différences individuelles entre les sinistres. L'objectif principal de ce travail est d'étendre l'hypothèse de régression simplifiée admise dans le modèle de Mack, en permettant l'inclusion d'informations sur les sinistres individuels. Ces informations conduiront à un affinement des calculs des provisions pour sinistres et d'aider à comprendre les différences structurelles entre différents types de sinistres.

L'autre difficulté relevée précédemment est la présence de données censurées à droite dans les observations. Cette problématique n'est pas spécifique à l'assurance, mais concerne plus généralement l'analyse des données de survie. Une littérature abondante en épidémiologie, en biostatistique et en démographie s'intéresse à l'inférence de modèles de survie. Lorsque des données internes sont disponibles avec des informations fiables, les actuaires estiment la loi de survie en appliquant des méthodes tirées de l'analyse de survie unidimensionnelle. Les estimateurs de Hoem (1971) [19] ou de Kaplan et Meier (1958) [20] sont alors classiquement utilisés dans la pratique pour l'évaluation des durées des sinistres. Toutefois, il faut noter qu'ils existent plusieurs approchent plus ou moins efficaces de prise en compte et de gestion des données censurées dans les modèles d'apprentissage automatique.

Dans ce contexte, des revues systématiques récentes décrivant plus de 100 modèles de prédiction de risque ont été produits entre 1999 et 2009 [21,22], y compris les scores de risque comme le *Framingham risk score* [23], le *Reynolds risk score* [24,25] et les récentes équations des cohortes regroupées de l'*American Heart Association / American College of cardiology* [26]. La plupart des modèles de prédiction du risque ont été estimés à l'aide de données provenant de cohortes épidémiologiques homogènes (cohortes soigneusement sélectionnées). Ces modèles fonctionnent souvent mal lorsqu'ils sont appliqués à des populations variées et actuelles [27]. Comme exemples nous pouvons citer les modèles de prédiction du risque de maladie cardiovasculaire et les conséquences associées comme les crises cardiaques, les accidents vasculaires cérébraux. Nous pouvons classer de manière simplifiée les modèles d'apprentissage souvent utilisés en trois grandes familles. La famille des modèles linéaires est composée des modèles linéaires (*Linear Models*), des modèles linéaires généralisés (*Generalized Linear Models*) et des modèles linéaires régularisés (type régressions *ridge* et *elasticnet*, *Regularized Linear Models*). La famille des modèles d'arbres de décision est composée des modèles d'arbres de classification et de régression (*Classification And Regression Trees*), des modèles de forêt aléatoire (*Random Forest*) et des modèles *Gradient Boosted Machines*. Enfin nous pouvons classer dans une troisième famille les modèles comme le *Nearest Neighbor*, le *Naïve Bayes*, le *Neural Networks* et le *Support Vector Machines*.

Dans le langage de l'analyse statistique de survie, la date de survenance de l'événement adverse (notée date de survenance) est appelé censure à droite si le suivi du sujet se termine avant qu'il ne soit touché par cet événement [28]. Malheureusement, la plupart des algorithmes d'apprentissage automatique supervisés et des méthodes de classification supposent généralement que le statut de l'événement est connu pour tous les individus, alors que le statut de l'événement est indéterminé pour les sujets dont la date de réalisation de l'événement est censurée. Ces derniers ne sont pas suivis sur toute la période de temps au-delà de laquelle l'on souhaite faire des prédictions. Pour gérer les statuts d'événements inconnus en raison de la censure à droite, les travaux antérieurs ont proposé d'utiliser des étapes de prétraitement pour compléter ou exclure les données censurées, ou encore adapter des outils spécifiques d'apprentissage automatique aux données censurées. Les approches courantes utilisées pour traiter les données censurées sont principalement : soit de rejeter ces observations [29,30], soit de les traiter comme des non-événements [31,32], ou encore de séparer l'échantillon d'observations en deux (une où l'événement se produit et une où l'événement ne se produit pas). Dans ce dernier cas, un poids est attribué à chacune de ces observations en fonction de la probabilité marginale de réalisation possible de l'événement entre la date de censure et la période où leurs statuts seront évalués [33]. Ces approches simples introduisent des biais dans l'estimation du risque (estimation des probabilités de classe). D'une part, rejeter les observations avec un statut d'événement inconnu ou les traiter comme des non-événements à sous-estimer le risque [31,32]. D'autre part, la troisième approche, bien que plus sophistiquée, conduit à une mauvaise estimation du risque, car ces pondérations atténuent la relation entre les covariables et les variables à expliquer.

Ainsi, la plupart des approches d'apprentissage automatique qui prennent en compte le cas des données censurées de manière plus précises sont relativement récentes. La majorité de ces techniques sont des cas particuliers d'outils spécifiques d'apprentissage automatique. Par exemple, plusieurs auteurs, dont Hothorn et al. (2004) [34], Ishwaran et al. (2008) [35] et Ibrahim et al. (2008) [36] décrivent des versions d'arbres de classification et des forêts aléatoires (*Random Forest*) pour estimer la probabilité de survie. Lucas et al. (2004) [37] et Bandyopadhyay et al. (2015) [38] présentent une application des réseaux Bayésiens aux données censurées à droite. Peu d'auteurs ont envisagé appliquer les réseaux neuronaux aux données de survie, mais supposent généralement que les censures sont peu nombreuses [39,40]. En outre, plusieurs ont envisagé adapter les algorithmes SVM (*support vector machines*) aux données censurées en modifiant la fonction de perte (*Loss function*) pour tenir compte des censurés [41,42]. Malgré l'intérêt de ces différentes approches, elles sont trop spécifiques pour permettre une large généralisation aux techniques d'apprentissage lorsque l'on est en présence des données censurées à droite.

Finalement, Vock et al. (2016) [43] proposent une approche plus générale pour prendre en compte les données censurées à droite en utilisant la méthode IPCW (*Inverse probability of Censoring Weighting*). En réalité, il convient de noter que l'usage de l'IPCW dans les méthodes d'apprentissage automatique n'est pas récent. Par exemple, de précédents travaux de Bandyopadhyay et al. (2015) [38] ont illustré la manière d'utiliser l'IPCW, notamment dans le cas d'estimation des réseaux bayésiens avec des données censurées à droite. Cependant, il est très probable que l'article de Vock et al., soit parmi les premiers à proposer l'utilisation de l'IPCW comme technique de prise en compte des censures à droite généralisable à de nombreuses méthodes d'apprentissage automatique. Ils illustrent comment l'IPCW peut facilement être incorporé dans un certain nombre d'algorithmes d'apprentissage automatique calibrés sur de grandes données sur les soins de santé, y compris les réseaux bayésiens, les k-voisins les plus proches, les arbres de décision et les modèles additifs généralisés. Ils démontrent que cette approche conduit à des prévisions mieux calibrées que les autres approches ad hoc. Ils donnent une illustration en appliquant leur méthode à la prédiction du risque de survenance d'un accident cardiovasculaire dans les cinq prochaines années, à partir des données du système de santé américain du Midwest EHD (*Electronic Healthcare Documentation system*). L'EHD est composé de bases de données contenant des dossiers médicaux électroniques (*Electronic Medical Records*, EMRs), des données sur les sinistres d'assurance (*Insurance Claims Data*) et des données sur la mortalité obtenue à partir des registres de décès gouvernementaux (G*overnmental Vital Records*). Les bases de données EHD sont de plus en plus disponibles dans les systèmes de soins de santé de grande taille. Elles incluent généralement des données sur des centaines de milliers à des millions de patients avec des informations sur des millions de procédures, de diagnostics et de mesures en laboratoire [44,45].

La technique de l'IPCW a été également utilisée par Lopez et al. (2016) [46] pour améliorer l'algorithme d'arbre de classification et de régression (*Classification And Regression Trees*, CART) de Breiman et al. (1984) [47] afin d'estimer la distribution conditionnelle de variables censurées à droite pour le provisionnement des sinistres en assurance non-vie. Ils proposent une stratégie d'élagage des arbres intégrant un schéma de pondération par des poids de Kaplan-Meier (1958) [20] comme technique de sélection d'un sous-arbre optimal. A titre d'illustration, ils testent leur algorithme modifié (*Tree-Base Censored*) sur des données issues de simulations de sinistres de contrats garantissant contre le risque de prévoyance et de responsabilité civile. Le grand avantage de l'algorithme proposé (*Tree-Base Censored*) est de produire des résultats interprétables et sa principale limite c'est l'instabilité des prédictions spécifique à la méthode de CART.

L'objectif principal de cet article est double. D'abord, nous apportons une première réponse en adaptant des méthodes ensemblistes d'apprentissage automatique avec la technique de l'IPCW pour le provisionnement individuel des sinistres en assurance non-vie. Ce type de méthode permet d'isoler le calcul des IBNeR de celui des IBNyR en prenant en compte les caractéristiques de chaque sinistre, ce qui n'est pas le cas des méthodes agrégées. Nous utilisons ces techniques d'apprentissage automatique pour estimer les IBNeR. Les techniques de *bagging* et de *boosting* permettent d'obtenir des résultats plus stables et de valider la robustesse des prédictions dans le temps. Les adaptations réalisées

sur les algorithmes intègrent certaines modifications proposées par Lopez et al. (2016) sur CART pour la prédiction des durées de sinistre. D'autres modifications ont été nécessaires pour permettre à ces algorithmes d'apprentissage ensembliste de prédire à la fois les durées et les charges sinistres ultimes. Enfin, nous avons dû adapter également les métriques de performance afin de gérer les données censurées. Ainsi, notre contribution améliore les prédictions des durées de maintien des sinistres et les charges sinistres ultimes individuelles. Elle apporte de surcroît une réponse à la principale limite de l'approche *Tree-base censored* proposée Lopez et al. Ensuite, à titre d'illustration, nous testons notre approche sur des données individuelles réelles de taille suffisante, issues des bases de gestion d'un assureur de personne. Nous appliquons nos approches à deux portefeuilles assez différents par leur taille d'effectifs, les risques sous-jacents aux garanties assurées, les variables explicatives et la période d'historiques disponibles. Le portefeuille des contrats de prêts est composé de 879 553 sinistrés sur la période comprise entre 2004 et 2016. Celui des contrats de prévoyance collective contient 347 904 sinistrés sur la période 2014 et 2017. Nous utilisons notre approche pour prédire la durée de maintien et la charge sinistre ultime des risques couverts, notamment l'arrêt de travail et la perte d'emploi (ou chômage). Dans la suite nous ne ferons pas de distinction entre charge sinistre ultime la charge sinistre.

La suite de l'article est organisée de la manière suivante. La section 2 est consacrée au cadre méthodologique et à la modélisation. Nous présentons la méthode IPCW (*Inverse Probability of Censoring Weighting*) ainsi que les procédures d'adaptation des techniques d'apprentissage automatique fondées sur les méthodes ensemblistes et permettant de gérer les données censurées pour la prédiction des paramètres d'intérêt. Dans la section 3 nous fournissons quelques applications sur des données issues des bases de contrats d'assurance de prêts et de prévoyance collective. Nous discutons dans la section 4, des résultats obtenus et des limites de l'approche proposée. Enfin, nous terminons cet article par une conclusion à la section 5.

2. Cadre méthodologique et modélisation

2.1. Notations et terminologies

Considérons un vecteur aléatoire (M, T, X), où $M \in R^p$ est la variable d'intérêt, $T \in R^+$ est la variable de durée, et $X \in \mathcal{X} \subset R^p$ le vecteur des covariables aléatoires explicatives de T et/ou M. La présence de censures dans les données ne permet pas l'observation directe de (M, T), alors que X est toujours observable. Notons également $C \in R^+$ la variable de censure. Dans un souci de simplicité et sans perdre de généralité, supposons que les variables T et C sont des variables continues, et que les composantes de M sont toutes strictement positive. Les variables observables en dehors de (M, T) sont : $Y = min(T, C)$, $\delta = 1_{T \leq C}$ et $N = \delta M$.

Enfin, soit i $(1 \leq i \leq n)$ un individu de la base, il est représenté par $(N_i, Y_i, \delta_i, X_i)$. Cette base est une réplication de variables correspondant à des individus indépendants et identiquement distribués notée $(N_i, Y_i, \delta_i, X_i)_{1 \leq i \leq n}$. Lorsque la durée totale du sinistre d'un individu i est observée, la variable M_i est une quantité observable et M correspond à la charge sinistre totale des prestations versées au titre des sinistres clôturés. Dans certains cas spécifiques comme dans le cadre de régression censurée $M = T$, voir Stute (1993) [48].

Dans ces conditions, le but de l'apprentissage automatique est de déterminer les impacts de X et si possible de T sur M. Cela revient à estimer une fonction π_0 telle que

$$\pi_0 = \underset{\pi \in \Psi}{\text{Argmin}}\ E\big[\varphi\big(M, \pi(T, X)\big)\big],$$

où Ψ est un sous-ensemble d'un espace fonctionnel approprié et φ une fonction de perte. Lopez et al. (2016) [46] proposent des types de modèles de régression correspondant à différents choix de fonctions de perte suivant le sous-ensemble fonctionnel considéré.

2.2. Le sur-apprentissage et le sous-apprentissage

Les méthodes d'apprentissage supervisé permettent de construire lors de la phase d'entraînement de modèles à partir d'un échantillon d'apprentissage, en vue de prédire une variable réponse. Pour cela il faut disposer d'observations sur la réalisation de cette variable réponse pour chaque individu de l'échantillon d'apprentissage. La prédiction du modèle est comparée à la réalisation de la variable réponse des individus de l'échantillon d'apprentissage n'est pas suffisant pour évaluer la qualité du modèle. En effet, l'objectif est de créer un modèle qui est généralisable et applicable à des individus qui ne sont pas présents dans la base d'apprentissage. Ainsi, pour mesurer la qualité d'un modèle, nous devons évaluer l'erreur de prédiction sur des données qui n'ont pas servi à la création de ce modèle. En pratique, l'estimation de cette erreur se fait de façon empirique en utilisant un échantillon de test (ou base test).

Le but d'un algorithme d'apprentissage automatique est de capter la tendance d'un échantillon et les relations entre la variable réponse et les covariables. Le sur-apprentissage (*overfitting*) désigne le fait que le modèle prédictif produit par l'algorithme d'apprentissage automatique s'adapte trop bien aux données de l'échantillon d'apprentissage. Par conséquent, le modèle prédictif capturera tous les détails qui caractérisent les données d'apprentissage. Dans ce sens, il capturera toutes les fluctuations et variations aléatoires des données d'apprentissage. En d'autres termes, le modèle prédictif capturera non seulement les corrélations généralisables mais aussi les bruits propres aux données d'apprentissage. On dit que la fonction prédictive se généralise mal. De la même manière, lorsque l'on n'arrive pas à capter les relations entre la variable réponse et les covariables, le modèle ne sera pas très performant. Cela sous-entend que le modèle prédictif généré lors de la phase d'apprentissage, s'adapte mal aux données d'apprentissage. Dans ce cas on parle alors de sous-apprentissage (*underfitting*).

Le sur-apprentissage et le sous-apprentissage sont les causes principales des mauvaises performances des modèles prédictifs générés par les algorithmes d'apprentissage automatique. Le but, dans la mise en place de n'importe quel algorithme, est de trouver le meilleur équilibre entre sous-apprentissage et sur-apprentissage. En d'autres termes, il ne souffre ni d'un grand *biais* ni d'une grande *variance*.

2.3. Optimisation du critère biais-variance

Dans le cas de la régression, l'outil le plus utilisé pour tester un modèle est l'erreur quadratique, aussi appelé MSE (*Mean Squared Error*). Dans la pratique, l'estimation de l'espérance de l'erreur quadratique d'un modèle est réalisée de façon empirique sur un échantillon de validation V de taille *n*. La MSE est souvent utilisées dans le cas de la régression lorsque les observations sont complètes afin de sélectionner le meilleur modèle parmi tous les modèles calibrés à la phase d'apprentissage. Formellement, on a $\hat{E}(\hat{f}, V) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{f}(x_i)\right)^2$, où y_i est la réalisation et $\hat{f}(x_i)$ la prédiction du modèle. Supposons que la fonction à modéliser est $y = f(x) + \varepsilon$, où ε est un bruit d'espérance nulle et de variance σ. Nous créons un estimateur $\hat{f}(.)$ avec l'échantillon d'apprentissage de telle sorte que pour chaque point de l'échantillon de validation y_i, son estimation soit $\hat{f}(x_i)$. L'erreur quadratique peut donc s'écrire

$$MSE = E\left[\left(y - \hat{f}(x)\right)^2\right].$$

L'erreur quadratique peut donc se décomposer comme suit

$$MSE = Biais^2 + Variance + Erreur\ irréductible.$$

La variance correspond à la sensibilité du modèle due aux fluctuations de l'échantillon d'apprentissage. Ainsi, plus un modèle possède une variance importante, plus il changera si on perturbe les données. Les modèles qui ont une grande variance sont des modèles trop complexes car ils modélisent du bruit. Ils créent des relations qui ne sont pas applicables sur un autre échantillon. On est alors dans le cas du

sur-apprentissage. Le biais correspond à l'erreur moyenne de $\hat{f}(x)$. Les modèles qui ont un biais important sont des modèles qui sont trop simples. Ils ne captent pas correctement les relations entre la variable de sortie et les covariables. On est alors dans le cas du sous-apprentissage.

Finalement, un estimateur de la métrique MSE s'écrit

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{f}(X_i)\right)^2,$$

avec i représentant les individus (($i = 1, \dots, n$), X_i le vecteur des co-variables, Y_i la variable à prédire (*target*) et $\hat{f}(.)$ la fonction d'estimation du modèle.

2.4. Découpage des bases de données

Le découpage des bases de données est une étape importante de l'apprentissage automatique. Il est d'usage de décomposer la base en trois parties ou échantillons. L'échantillon d'apprentissage permet de construire l'estimateur. L'échantillon de validation sert à calibrer et à tester la performance du modèle pendant la phase d'apprentissage. Il sert à gérer les problèmes de sur-apprentissage et de sous-apprentissage. L'échantillon de test est utilisé pour mesurer le score final des modèles et permet de mesurer leur qualité prédictive.

Le principe de découpage des bases c'est de disposer d'un maximum de données pour l'échantillon d'apprentissage tout en en gardant suffisamment pour les bases de validation et de test. Pour disposer de proportions suffisantes par échantillon et limiter la perte d'informations pour l'apprentissage, plusieurs techniques sont mise en œuvre dans la pratique. Pour créer un modèle, certaines méthodes de ré-échantillonnage proposent des techniques permettant d'utiliser l'intégralité des données de l'échantillon dit de construction (composé des échantillons d'apprentissage et de validation).

Primo, la technique du découpage apprentissage-validation est la méthode la plus simple qui consiste à découper l'échantillon de construction en deux. Cette méthode est également la méthode la plus rapide en termes de temps de calcul mais la moins précise pour estimer le taux d'erreur d'un modèle.

Secundo, la technique de la validation croisée (ou *k-Fold Cross Validation*) est de découper la base de données en k échantillons de même taille. Un modèle est construit en utilisant tous les échantillons sauf le dernier qui sert d'échantillon de validation. La procédure est répétée en utilisant l'avant-dernier échantillon comme échantillon de validation et le reste comme échantillon d'apprentissage, etc... Les modèles ainsi construits disposent chacun d'un score de performance. La moyenne des scores permet d'obtenir un score de performance moyen. La validation croisée *leave-one-out* correspond au cas particulier où k est égal à la taille de l'échantillon de construction. C'est la technique qui permet d'avoir une mesure de la performance de notre modèle avec le biais le plus faible mais en échange d'un temps d'exécution très long. Cette technique est intéressante lorsque la base de données est petite.

Enfin tertio, la technique du *bootstrap* est une technique de ré-échantillonnage utilisée dans les algorithmes de forêts aléatoires. Soit un échantillon de construction de taille n. Créer un échantillon *bootstrap* à partir de cet échantillon revient à tirer n individus avec remise. Dans chaque échantillon *boostrap* créé, certains individus ont été tirés plusieurs fois alors que d'autres n'ont pas été sélectionnés. Cette procédure est répétée un grand nombre de fois. Les individus non tirés dans chaque échantillon *bootstrap* servent alors de base de validation. Cet ensemble est appelé l'échantillon *Out-Of-Bag* (OOB) qui signifie "en dehors du *bootstrap*". La probabilité pour qu'un individu ne soit pas tiré dans un échantillon *bootstrap* est proche de e^{-1} soit 0,368. Ainsi, en moyenne, sur chaque échantillon *bootstrap* 63,2% des individus sont utilisés pour créer le modèle d'apprentissage alors que 36,8% des individus sont utilisés pour mesurer la performance du modèle. L'échantillon OOB sert souvent d'échantillon de validation pour calibrer les paramètres du modèle.

2.5. Optimisation des hyper paramètres ou tuning du modèle

La plupart des algorithmes d'apprentissage automatique nécessitent d'être calibrés, c'est-à-dire que certains paramètres doivent être fixés. Nous avons vu que l'échantillon de validation est utilisé pour mesurer le score du modèle pour chaque paramètre possible. La procédure pour optimiser les hyper-paramètres d'un modèle se décline en générale en cinq principales étapes. La première étape consiste à définir la liste des valeurs possibles des paramètres du modèle. Lors de la seconde étape l'on procède au ré-échantillonnage de la base de données, puis à la construction du modèle et enfin à l'application sur la ou les bases de validation. La troisième étape sert à l'agrégation des résultats pour déterminer le score estimé du modèle en fonction de la valeur des paramètres. Ensuite la quatrième étape sert au choix des paramètres optimaux du modèle. Enfin la cinquième et dernière étape sert à la construction du modèle final sur la totalité de la base de construction. Finalement, l'échantillon de test ne sera utilisé qu'une fois que l'algorithme sera finalisé pour mesurer le score de performance final du modèle.

2.6. Les algorithmes ensembliste d'apprentissage automatique

2.6.1. Les catégories de l'apprentissage automatique

On distingue trois catégories d'apprentissage automatique. D'abord, les *méthodes d'apprentissage supervisé* regroupant l'ensemble des algorithmes qui utilisent une base de données d'apprentissage afin de réaliser des prédictions. Cette base d'apprentissage inclut des variables explicatives (ou covariables) et la variable à expliquer (ou variable réponse). A partir de cette base, l'algorithme supervisé a pour objectif de construire un modèle qui permet la prédiction des variables réponses d'un nouvel échantillon. On distingue les méthodes de classification (où la variable réponse est catégorielle) et des méthodes de régression (où la variable réponse est continue).

En classification, le but est de prédire la classe d'une variable. L'objectif est ainsi de lier les covariables avec une fonction discrète. Ensuite, les *méthodes d'apprentissage non supervisé* pour lesquelles l'algorithme ne dispose pas de variable réponse. Ce dernier doit trouver la structure présente dans le jeu de données. Le modèle regroupe alors les individus en sous-groupes homogènes. Le *clustering* fait par exemple partie des algorithmes d'apprentissage non supervisés. S'ajoute aussi les *méthodes d'apprentissage par renforcement* pour lesquelles la machine apprend en fonction des conséquences de ses actions en interagissant avec l'environnement. On appelle aussi ce type d'apprentissage l'apprentissage « *trial* and *error* ».

Nous utilisons dans cet article les méthodes ensemblistes d'apprentissage supervisé des arbres de décision, notamment l'algorithme des arbres de classification et de régression (*Classification And Regression Tree,* CART), l'algorithme de la forêt aléatoire (*Random Forest,* RF) et l'algorithme du (*Gradient Tree Boosting Machine,* GB).

2.6.2. L'algorithme CART (Arbres de classification et de régression)

Les arbres de décision sont des techniques d'apprentissage automatique qui servent à construire des modèles prédictifs à partir d'un échantillon de données. Ces techniques consistent à partitionner récursivement l'échantillon de données initial et de construire un modèle prédictif sur chaque partition. Ils existent deux catégories d'arbres de décision, les arbres de classification et les arbres de régression. Les arbres de classification ont pour objectif de prédire une variable qualitative alors que les arbres de régression permettent de prédire une variable quantitative.

Le but d'un algorithme d'arbre de décision c'est de créer un modèle qui prédit une variable réponse (ou variable cible) avec des variables explicatives (ou variables d'entrée) issues de l'échantillon de construction (c'est-à-dire les échantillons d'apprentissage et de validation). Pour expliquer la variable réponse, les modèles par arbre de décision partitionnent récursivement l'échantillon d'apprentissage grâce aux variables explicatives. Les arbres de décision permettent de segmenter une population de manière efficace mais ils peuvent aussi servir à prédire la classe d'une nouvelle. En effet, pour prédire la variable réponse d'un nouvel individu, il suffit de connaître la valeur des réalisations des variables aléatoires explicatives et d'interroger l'arbre.

Il est d'usage d'utiliser un vocabulaire approprié pour illustrer le principe de la répartition récursive. On appelle *racine* l'ensemble de la population. C'est l'échantillon initial qui n'a pas encore été segmenté. Les *branches* sont les règles qui permettent de segmenter la population en deux. Les *nœuds* sont les sous-échantillons qui sont créés dans l'arbre. Les *feuilles* sont les sous-populations homogènes créées qui donnent un estimateur de ce que l'on veut estimer.

L'algorithme CART est un arbre de décision binaire. À chaque fois que l'on partitionne l'échantillon, on le sépare en deux sous-échantillons grâce à une condition sur les variables explicatives. L'algorithme CART s'effectue en trois temps, la création de l'arbre maximal, l'élagage de l'arbre maximal et le choix de l'arbre optimal parmi les arbres élagués. Cette famille d'algorithme est largement appliquée notamment en bio statistique ou en médecine. De nombreuses techniques ont été proposées pour développer des arbres de décision, différant pour la plupart des critères utilisés pour empêcher le surapprentissage (voir [36], [52] et [53]).

Spécificités des arbres de classification

Dans le cas où l'arbre de décision est un arbre de classification, la variable à expliquer est une variable qualitative. Autrement dit, on cherche à regrouper dans la même classe les individus ayant les mêmes caractéristiques (variables explicatives) et ayant des valeurs identiques pour la variable à expliquer.

Le but de l'arbre est de partitionner la population initiale en sous-populations. L'idée est de couper la racine en deux feuilles plus homogènes selon une règle. Le critère d'évaluation des partitions caractérise l'homogénéité (ou le gain en homogénéité) des sous-ensembles obtenus par division de l'ensemble. Ces métriques sont appliquées à chaque sous-ensemble candidat et les résultats sont combinés (par exemple, moyennés) pour produire une mesure de la qualité de la séparation. L'algorithme CART utilise l'indice de diversité de Gini. Cet indice mesure avec quelle fréquence un élément aléatoire de l'ensemble serait mal classé si son étiquette était choisie aléatoirement selon la distribution des étiquettes dans le sous-ensemble. Cette métrique permet donc de déterminer à chaque nœud s'il convient de le subdiviser en nœuds-fils. La variation (diminution) de l'indice de diversité de Gini pour une subdivision possible est donné par

$$\Delta I_G = \hat{\pi}(s)[1-\hat{\pi}(s)] - \frac{1}{N_s}\left\{N_{s_l}\hat{\pi}(s_l)[1-\hat{\pi}(s_l)] + N_{s_r}\hat{\pi}(s_r)[1-\hat{\pi}(s_r)]\right\},$$

où $\hat{\pi}(s)$, $\hat{\pi}(s_l)$ et $\hat{\pi}(s_r)$ sont respectivement la proportion de l'échantillon (d'apprentissage) affectés au nœud s, au nœud-fils à gauche s_l et au nœud-fils droit s_r pour un critère donné de subdivision. Le nombre d'individus N_{s_l} et N_{s_r} dans chaque nœud-fils (gauche et droit) est tel que

$$N_s = N_{s_l} + N_{s_r}.$$

Le processus s'arrête lorsque l'on ne peut plus segmenter les nœuds obtenus. Ceci se produit quand il n'y a plus qu'une seule observation dans la population ou bien si les observations de la population ne peuvent plus être différenciées avec les covariables du modèle. L'algorithme de la création de l'arbre maximal peut être subdivisé en trois étapes. La première étape correspond à l'initialisation, on se place au niveau de la racine de l'arbre. La deuxième étape consiste à minimiser la somme pondérée de l'impureté des deux nœuds fils créés en testant chaque covariable et chaque seuil. La troisième étape sert de test de fin. Si on ne peut plus segmenter, on s'arrête. Sinon, on recommence l'étape 2 sur chacun des deux nœuds fils.

Spécificités des arbres de régression

Dans le cas où l'arbre de décision est un arbre de régression, la variable à expliquer est une variable continue. On cherche ainsi à connaître la valeur d'une quantité d'intérêt. Le même schéma de séparation peut être appliqué, mais au lieu de minimiser le taux d'erreur de classification, on cherche à maximiser la variance interclasse. Il s'agit d'avoir des sous-ensembles dont les valeurs de la variable à expliquer soient les plus dispersées possibles. En général, le critère utilise le test du chi carré. Pour l'adaptation de l'algorithme au cas pondéré, une approche similaire à celle présentée ci-dessus pourrait être appliquée pour calculer la version pondérée de la métrique adéquate.

De la même manière que pour les arbres de classification, pour séparer l'ensemble en deux sous-ensembles plus homogènes, on teste chaque variable explicative et chaque seuil puis on retient le couple qui minimise la variance des deux nouveaux nœuds. On recommence ce processus sur chaque nouveau nœud avec le nouvel ensemble associé. L'algorithme de la création de l'arbre maximal peut également être subdivisé en trois étapes. La première étape correspond à l'initialisation, on se place au niveau de la racine de l'arbre (population initiale). La seconde étape consiste à minimiser la somme pondérée de la variance des deux nœuds fils en testant chaque covariable et chaque seuil. Enfin, la troisième étape sert de test de fin. Si on ne peut plus segmenter, on s'arrête. Sinon, on recommence l'étape 2 sur chaque nœud fils.

2.6.3. L'algorithme de la forêt aléatoire (Random Forest)

Les forêts d'arbres décisionnels ou forêts aléatoires (*Random Forest Classifier*) font parties des techniques d'apprentissage automatique (voir [54]). Elles ont été formellement proposées en 2001 par Leo Breiman et al. [55,56]. La proposition de Breiman vise à corriger plusieurs inconvénients connus de la méthode initiale d'arbres décisionnels, comme la sensibilité des arbres uniques à l'ordre des covariables explicatives (ou prédicteurs). Cet algorithme effectue un apprentissage sur de multiples arbres de décision entraînés sur des sous-ensembles de données (échantillons d'apprentissage) légèrement différents, en calculant un ensemble de M arbres partiellement indépendants. De ce fait, il combine les concepts de sous-espaces aléatoires et la technique du *bagging*.

Technique du *bagging* (contraction de *Bootstrap Aggregating*)

Le *bagging* est une technique utilisée pour réduire l'erreur de prédiction. En général, le *bagging* a pour but de réduire la variance de l'estimateur, en d'autres termes de corriger l'instabilité des arbres de décision. Cette instabilité se traduit par le fait que de petites modifications dans l'échantillon d'apprentissage entraînent des arbres très différents. Pour ce faire, le principe du *bootstrap* (technique d'inférence statistique [36]) permet de créer un ensemble de M échantillons d'apprentissage par tirage aléatoires avec remise dans l'échantillon d'apprentissage initial. L'algorithme est entraîné pour calibrer B modèles partiellement indépendants. Lorsque les données sont quantitatives, cas d'un arbre de régression, les estimateurs ainsi obtenus sont moyennés. Lorsqu'il s'agit de données qualitatives, cas d'un arbre de classification, on utilise un vote à la majorité. C'est la combinaison de ces multiples estimateurs indépendants qui permet de réduire la variance. Cette technique est utilisée pour améliorer les algorithmes des arbres de décision, considérés comme des « classifieurs faibles », c'est-à-dire à peine plus efficaces qu'une classification aléatoire, ou d'autres méthodes de modélisation (régression, *Neural Network*, etc…).

De manière formelle, soient $Y \in R$ la variable à prédire (ou variable réponse) et X_k $(k = 1,2, \dots, p)$ les covariables (ou variables explicatives). Supposons qu'on a construit un premier modèle $\hat{f}$ sur l'échantillon d'apprentissage. Le *bagging* consiste à choisir un ensemble de *bootstrap* de l'échantillon d'apprentissage puis de créer M modèles pour chaque échantillon *bootstrapé*.

Par exemple, si nous considérons le cas où la variable réponse est quantitative, si on note $\hat{f}_m^{(b)}$ le modèle calibré sur les données de l'échantillon d'apprentissage *boostrapé* m ($m = 1, \dots, M$), alors l'estimateur par la méthode *bagging* dans le cas d'une régression est tout simplement la moyenne des prédictions de chaque modèle

$$\hat{f}_{bag}(x) = \frac{1}{m}\sum_{m=1}^{M} \hat{f}_m^{(b)}(x).$$

Réduction Biais Variance

D'après Rouvière (2018) [59], de ce qui précède, nous obtenons le résultat suivant

$$Var\left[\hat{f}_{bag}(x)\right] = \rho(x)\sigma^2(x) + \frac{1-\rho(x)}{M}\sigma^2(x),$$

avec $\sigma^2(x) = Var\left[\hat{f}_m^{(b)}\right]$ la variance de chaque estimateur et $\rho(x) = cor\left[\hat{f}_r^{(b)}, \hat{f}_s^{(b)}\right]$ le coefficient de corrélation entre deux estimateurs différents.

Ainsi, plus M est grand et plus la variance est faible, car

$$\lim_{M\to\infty} Var\left[\hat{f}_{bag}(x)\right] = \rho(x)\sigma^2(x).$$

En ce qui concerne le bais, Rouvière (2018) [59] montre que si nous supposons les estimateurs identiquement distribués alors le biais de l'estimateur *baggé* est le même que celui des estimateurs à agréger. La technique du *bagging* est ainsi adaptée dans le cas où les estimateurs sont très sensibles aux fluctuations de la base d'apprentissage. Ceci est le cas pour les arbres CART et plus particulièrement quand ils ne sont pas élagués.

Compromis Biais-Variance dans le choix du paramètre *mtry*

Le paramètre *mtry* correspond au nombre de variables testées à chaque division. Intuitivement, plus *mtry* est petit, plus la création des arbres est aléatoire. La corrélation entre les arbres ainsi que la variance du modèle sont donc plus faible. En revanche, plus *mtry* sera petit, plus le biais de chaque arbre sera fort. Un *mtry* petit donnera une variance petite (i.e. la corrélation entre les arbres sera plus faible) mais un biais fort. A l'inverse, un *mtry* grand minimisera le bais mais aura une variance élevée. Le choix de ce paramètre est très important car il permet de trouver le meilleur compromis entre le biais et la variance.

Dans la pratique plusieurs approches euristiques sont utilisées. Une première façon est de faire tourner *Random Forest* une dizaine de fois avec des valeurs de *mtry* différentes et choisir celle pour laquelle le *mtry* est minimal et se stabilise. Une autre approche est de tester avec un *mtry* par défaut (entre 2 et 5) et avec la moitié de cette valeur et le double.

L'erreur *Out Of Bag*

Il s'agit d'une procédure permettant de fournir un estimateur des erreurs. De tels estimateurs sont souvent construits à l'aide de méthode apprentissage-validation ou validation croisée. L'avantage de la procédure *Out Of Bag* (OOB) est qu'elle ne nécessite pas de découper l'échantillon de construction en échantillon d'apprentissage et de validation. Elle utilise le fait que les arbres sont construits sur des estimateurs *baggés* et que, par conséquent, ils n'utilisent pas toutes les observations de l'échantillon d'apprentissage. Nous la détaillons dans le cadre des forêts aléatoires mais elle se généralise à l'ensemble des méthodes *bagging*.

Nous avons vu que sur chaque échantillon *boostrap*, en moyenne, 36.8% des individus sont absents et peuvent servir d'échantillon de validation. Pour évaluer le score de performance d'une forêt nous pouvons nous servir de l'échantillon OOB (sans découpage en échantillon d'apprentissage et de validation).

Soit $D_n = \{(X_1, Y_1), \dots, (X_n, Y_n)\}$ l'échantillon de construction de taille n. Soit un individu (X_i, Y_i) de l'échantillon, on désigne par I_B l'ensemble de taille B des arbres de la forêt qui ne contiennent pas cette observation dans leur échantillon *bootstrap*. La prévision de Y_i de l'individu i par la forêt aléatoire calibrée sur cet échantillon OOB est alors égale à

$$\hat{Y}_i = \frac{1}{B}\sum_{m \in I_B} \hat{f}_m^{(b)}(X_i).$$

Si on est dans un contexte de classification, la prévision s'obtient en faisant voter ces arbres à la majorité. L'erreur Out Of Bag est alors définie par $\frac{1}{n}\sum_{i=1}^{n}\left(\hat{Y}_i - Y_i\right)^2$ en régression et $\frac{1}{n}\sum_{i=1}^{n} 1_{\hat{Y}_i \neq Y_i}$ en classification.

L'erreur *Out Of Bag* d'une forêt aléatoire permet donc de donner une approximation de la véritable erreur de la forêt. Ainsi, l'erreur OOB est très utile pour comparer différents modèles entre eux. Par conséquent, le choix des paramètres de la forêt peut se faire grâce à cette estimation de l'erreur.

Importance des variables

Les forêts aléatoires peuvent être utilisées dans le but de mesurer l'importance des variables explicatives. Cette mesure d'importance peut servir à améliorer un modèle car certaines variables explicatives non informatives peuvent dégrader le score de ce modèle. De plus, elle peut aussi permettre d'obtenir un modèle performant avec peu de variables explicatives.

Soit Y la variable réponse et un vecteur aléatoire $X = (X_1, \dots, X_n)$ qui correspond aux covariables. Soit $D_n = \{(X_1, Y_1), \dots, (X_n, Y_n)\}$ l'échantillon d'apprentissage que nous utilisons pour construire l'estimateur $\hat{f}(x) = E[Y|X = x]$ (voir Gregorutti et al. [60]).

Pour estimer l'erreur quadratique du modèle, nous utilisons un échantillon de validation V de taille m_V. Une estimation de l'erreur est égale à

$$\hat{E}\left[\hat{f}, V\right] = \frac{1}{m_V}\sum_{i=1}^{m_V}\left(y_i - \hat{f}(x_i)\right)^2,$$

où y_i est la réalisation et $\hat{f}(x_i)$ la prédiction du modèle.

Avec la mesure d'importance de Breiman (2001) [55], une variable explicative est considérée comme importante si l'erreur de prédiction du modèle augmente lorsque nous perturbons aléatoirement lien entre cette variable et la variable réponse. Pour rompre le lien, Breiman (2001) [55] propose de permuter la variable explicative en question.

Considérons une forêt aléatoire composée de n_{arbres} estimateurs $\hat{f}_1, \dots, \hat{f}_{n_{arbres}}$. Soit V un échantillon de validation, soit $V_j^{(P)}$ l'échantillon après permutation des réalisations de la variable X_j. La mesure d'importance de la variable X_j sur l'échantillon V par permutation est égale à

$$\hat{I}(X_j) = \frac{1}{n_{arbre}}\sum_{t=1}^{n_{arbre}}\left(\hat{E}\left[\hat{f}_t, V_j^{(P)}\right] - \hat{E}\left[\hat{f}_{n_{arbre}}, V\right]\right).$$

Cet indicateur est très intéressant pour comparer l'importance de chaque variable dans le modèle et peut donc permettre d'en faire un classement.

Sensibilité du classement des variables d'importance

Par rapport aux paramètres de la forêt aléatoire, Genuer et al. (2012) [61] montrent que leur choix a un impact sur la mesure d'importance des variables explicatives. D'une part ces auteurs remarquent dans le cadre de leur étude, que l'ampleur des écarts entre les différents indices d'importance augmente lorsque le paramètre *mtry* augmente. En revanche, le choix de *mtry* ne semble pas changer le classement donné par la procédure. D'autre part, ils remarquent que fixer un nombre d'arbres élevé entraîne une meilleure stabilité de la mesure d'importance des covariables.

Par rapport à la corrélation entre les variables explicatives, Gregorutti et al. (2014) [60] montrent que si les variables explicatives sont fortement corrélées cela peut perturber le classement donné par la mesure d'importance par permutation. Ils proposent d'utiliser l'algorithme *Recursive Feature Elimination*. Cet algorithme consiste à construire une forêt aléatoire (étape 1), puis de mesurer l'importance de chaque covariable par permutation (étape 2), ensuite d'enlever du modèle la variable qui obtient le score le moins bon (étape 3) et enfin de répéter les étapes 1 à 3 jusqu'à l'obtention du classement corrigé des effets de la corrélation (étape 4).

Procédure de l'algorithme de la forêt aléatoire

Dans la pratique, la procédure peut se décomposer en quatre principales étapes [57].

La première étape consiste à créer M échantillons d'apprentissage par un double processus d'échantillonnage. Ces échantillons d'apprentissage sont construit en réalisant un tirage avec remise d'un nombre *N* d'observations identiques à celui des données d'origine, puis en retenant q prédicteurs sur les *p* variables explicatives tel que $q < \sqrt{p}$ (la limite n'est qu'indicative).

La seconde étape sert à entraîner sur chaque échantillon un arbre de décision selon une des techniques connues. Pour limiter la croissance des arbres de décisions, on utilise une technique comme celle de la validation croisée.

La troisième étape permet pour chaque observation d'origine de stocker les M prédictions de la variable à expliquer. Enfin la quatrième et dernière étape consiste à réaliser la prédiction de la forêt aléatoire par un simple vote majoritaire (*Ensemble learning*).

2.6.4. L'algorithme du Gradient tree boosting machine

Dans la communauté du learning, le *boosting* (Freund et Schapire (1996) [62]) se révèle être l'une des idées les plus puissantes des 15 dernières et continue de faire l'objet d'une littérature abondante. Le *boosting* est une technique ensembliste qui consiste à agréger des classifieurs (ou modèles) élaborés séquentiellement sur un échantillon d'apprentissage dont les poids des individus sont corrigés au fur et à mesure. Le principe général du *boosting* consiste à construire une famille d'estimateurs qui sont ensuite agrégés par une moyenne pondérée des estimations (en régression) ou un vote à la majorité (en discrimination). Les estimateurs sont construits de manière récursive : chaque estimateur est une version adaptative du précédent en donnant plus de poids aux observations mal ajustées ou mal prédites. L'estimateur construit à l'étape k concentrera donc ses efforts sur les observations mal ajustées par l'estimateur à l'étape $k-1$.

Le terme *boosting* s'applique à des méthodes générales capables de produire des décisions très précises à partir de règles peu précises (qui font légèrement mieux que le hasard). L'algorithme *adaboost* développé par Freund et Schapire (1997) [63] est l'algorithme le plus populaire appartenant à la famille des algorithmes *boosting*.

Soit par $f(x)$ une règle de classification « faible » (*weaklearner*), c'est-à-dire une règle dont le taux d'erreur est légèrement meilleur que celui d'une règle purement aléatoire (penser pile ou face). Soit M le nombre d'itérations. Pour $m = 1, \dots M$, notons e_m le taux d'erreur, α_m la qualité de l'ajustement et $f_m(x)$ l'estimateur issu de l'ajustement.

L'idée consiste à appliquer cette règle plusieurs fois en affectant « judicieusement » un poids différent aux observations à chaque itération. Les poids de chaque observation i ($i = 1, \dots, n$) sont initialisés à $w_i = \frac{1}{n}$ pour l'estimation du premier modèle.

Le taux d'erreur est calculé avec la formule

$$e_m = \frac{\sum_{i=1}^{n} w_i 1_{y_i - fm(x_i)}}{\sum_{i=1}^{n} w_i},$$

ainsi que la qualité de l'ajustement par la formule

$$\alpha_m = log\left[\frac{1-e_m}{e_m}\right].$$

Les poids sont ensuite mis à jour pour chaque itération.

L'importance w_i d'une observation est inchangée si l'observation est bien classée, dans le cas inverse elle croit avec la qualité d'ajustement du modèle mesurée par α_m .

L'agrégation finale est une combinaison des règles f_1,…, f_M pondérée par les qualités d'ajustement de chaque modèle. Se pose bien évidemment la question du choix de la règle faible. L'algorithme suppose que cette règle puisse s'appliquer à un échantillon pondéré. Ce n'est bien évidemment pas le cas de toutes les règles de classification. De nombreux logiciels utilisent par exemple des arbres de classification comme règle faible (voir Rouvière, 2018 [59]).

Un problème classique de l'apprentissage consiste à choisir une règle f à l'intérieur d'une famille donnée G. Ce problème est généralement abordé en minimisant l'espérance d'une fonction de perte

$$l : R \times R \to R \text{ tel que } f^*(x) = \underset{f \in G}{\operatorname{argmin}} E\left[l\left(Y, f(X)\right)\right].$$

La fonction de perte naturelle pour la classification est

$$l\left(Y, f(X)\right) = 1_{f(x) \neq y}.$$

La loi de (X, Y) étant inconnue, on minimise une version empirique de $E\left[l\left(Y, f(X)\right)\right]$. Ce qui revient à

$$f^*(x) = \underset{f \in G}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} l\left(Y, f(X)\right) = \underset{f \in G}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} 1_{f(X_i) \neq Y_i}.$$

Pour de nombreuses familles de règles, ce problème de minimisation est généralement difficile à résoudre numériquement. Une solution consiste à « *convexifier* » la fonction de perte de manière à utiliser des algorithmes de minimisation plus efficaces.

L'algorithme GBM (*Gradient Tree Boosting Machine*) est un cas particulier d'agrégation des classifieurs utilisant plusieurs modèles qu'il agrège ensuite pour obtenir le résultat final (voir Friedman [62], Chen, Y. [63] et Chen, T. [64]). Pour l'algorithme GBM, le calcul du poids des individus lors de la construction de chaque nouveau modèle se fait avec le gradient de la fonction de perte, afin d'améliorer les propriétés de convergence. La descente du gradient est une technique itérative qui permet d'approcher la solution d'un problème d'optimisation.

En apprentissage supervisé, la construction du modèle revient souvent à déterminer les paramètres du modèle qui permettent d'optimiser (maximum ou minimum) une fonction « objectif ». Il s'agit d'une démarche similaire à la descente de gradient pour les réseaux de neurones (voir Friedman, 1996 [62]).

Le GBM est donc une famille d'algorithmes basés sur une fonction de perte supposée convexe et différentiable notée $l^{(P)}$. Le modèle pas à pas est une descente de gradient de la forme

$$\hat{f}_m = \hat{f}_{m-1} - \gamma_m \sum_{i=1}^{n} \nabla f_{m-1} l^{(p)}\left(y_i, f_{m-1}(x_i)\right)$$

Le problème revient à rechercher un meilleur pas de descente γ tel que

$$\min_{\gamma} \sum_{i=1}^{n} l^{(p)} \left(y_i, f_{m-1}(x_i) - \gamma \frac{\partial l^{(p)}(y_i, f_{m-1}(x_i))}{\partial f_{m-1}(x_i)} \right).$$

Notons enfin que l'algorithme GBM utilise généralement des arbres CART. Il peut également être personnalisé en utilisant différents paramètres, différentes fonctions. Pour éviter le sur-apprentissage, il est d'usage d'utiliser des méthodes permettant de limiter la taille des arbres ou de construire les modèles sur des échantillons de la population (on parle de *stochastic gradient boosting*).

2.7. Modélisation des censures par la méthode IPCW

2.7.1. Problématique pour le provisionnement non-vie

L'utilisation d'échantillon de construction composé que de sinistres clôturés pour construire les modèles prédictifs conduit à un biais. En effet, cette contrainte conduit à une surreprésentation des sinistres courts qui sont en moyenne des sinistres à faible coût. Dans cette configuration, les algorithmes fournissent des modèles prédictifs performants sur une population qui ne représente pas la population réelle. En réalité, elle est composée d'un grand nombre de sinistres courts clôturés et de sinistres longs souvent censurés lors de l'estimation. Par conséquent, la prédiction des sinistres par un modèle biaisé aura tendance à être inférieure à la réalité, donc conduisant potentiellement à un sous-provisionnement.

La technique IPCW (*Inverse probability of Censoring Weighting*) est utilisée pour calibrer des modèles d'apprentissage automatique à partir de données contenant des observations censurées. L'estimation des probabilités de classes est réalisée uniquement avec les données individuelles pour lesquelles les durées totales de sinistre sont connues. Pour intégrer la présence des données censurées, les observations complètes sont pondérées par les poids IPCW. Ainsi, l'échantillon de construction fourni une représentation la plus précise possible de l'ensemble des observations. En pratique, en présence d'observations individuelles contenant des données censurées, l'on estime le poids de Kaplan-Meier. Les poids IPCW sont ensuite calculés en fonction des poids de Kaplan-Meier et utilisés comme des paramètres des algorithmes d'apprentissage automatique. Ainsi, le principal avantage de cette technique est de disposer d'une approche généralisable à n'importe qu'elle méthode d'apprentissage automatique pour la prédiction de risque en présence de données censurées à droite. Pour obtenir des justifications plus formelles de l'IPCW, le lecteur pourra consulter Bang et al. (2000) [49], Tsitis et al. (2006) [50] et Vock et al (2016) [43].

2.7.2. Estimation du poids de Kaplan-Meier et calcul des poids IPCW

A partir des données d'apprentissage, la fonction $G(t) = 1 - P(C > t)$ correspondant à la probabilité que la date de censure soit supérieure à t . Cette fonction est estimée avec l'estimateur de Kaplan-Meier de la fonction de survie en présence de données censurées [20].

En supposant que C est indépendant de (M, T) et

$$P(T \leq C | M, T, X) = P(T \leq C | T),$$

alors pour toute fonction $\varphi \in L^1$, on obtient l'égalité

$$E\left[\frac{\delta \varphi(N,Y,X)}{1-G(Y-)}\right] = E[\varphi(M,T,X)].$$

La fonction G qui est généralement inconnue peut cependant être estimée de manière cohérente par l'estimateur de Kaplan-Meier, c'est-à-dire

$$\hat{G}(t) = 1 - \prod_{Y_i \leq t} \left(1 - \frac{\delta_i}{\sum_{j-1}^{n} 1_{Y_j \geq Y_i}}\right),$$

puisque T et C sont indépendantes, et $P(T = C) = 0$ pour les variables aléatoires continues (voir dans Stute et al. (1993) [51] pour les détails sur la cohérence de l'estimateur de Kaplan-Meier).

Par conséquent, un estimateur naturel de $F(t, m, x) = P(T \leq t, M \leq m, X \leq x)$ est

$$\hat{F}(t, m, x) = \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i 1_{Y_j \leq t, N_i \leq m, X_i \leq x}}{1 - \hat{G}(Y_i -)}.$$

D'où, sous les conditions appropriées, un estimateur cohérent de $E[\varphi(T, M, X)]$ est

$$\hat{E}[\varphi(T, M, X)] = \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i \varphi(Y_i, N_i, X_i)}{1 - \hat{G}(Y_i -)}.$$

Finalement, le calcul des poids IPCW correspond à l'étape de définition du schéma de pondération à appliquer aux données. Pour chaque individu i de la base d'apprentissage, on définit sa IPCW par ω_i tel que

$$\omega_i = \begin{cases} \frac{1}{\hat{G}(\min(Y_i, t))} & si\ min(T_i, t) < C_i \\ 0 & sinon \end{cases}.$$

Les individus dont le statut d'événement est inconnu à t (c'est-à-dire, sont censurés avant t et ont donc $C_i \leq min(T_i, t)$) ont un poids $\omega_i = 0$, et sont donc exclus de l'analyse, et les individus restants se voient attribuer leurs poids ω_i.

2.8. Amélioration des techniques d'apprentissage automatique en présence de données censurées

2.8.1. Technique de modification des algorithmes

Cette étape consiste à appliquer une méthode d'apprentissage automatique existant à la version pondérée des données d'apprentissage où chaque individu i de la base d'apprentissage est pondéré par son poids ω_i. L'intégration des poids IPCW dans les algorithmes varie suivant les techniques d'apprentissage automatique utilisées. La mise en œuvre standard de certaines techniques d'apprentissage permettent la spécification directe des poids observés comme paramètres du modèle, auquel cas il faut peu de travail supplémentaire pour prédire les risques. L'incorporation des poids de l'IPCW se fait de manière assez simple lors de la phase d'estimation des hyper-paramètres du modèle (par exemple à l'aide d'estimateurs du maximum de vraisemblance) et de la phase d'évaluation de la qualité de l'ajustement/pureté du modèle (*fit and purity*). Le lecteur pourra trouver des justifications plus complètes de l'utilisation de l'IPCW dans Tsiatis et al. (2006) [50] et Vock et al. (2016) [43].

En pratique, lors de la phase d'apprentissage du modèle, une étape de prétraitement est réalisée. Elle consiste à pondérer par un poids IPCW nul tous les individus ayant un statut d'événement inconnu et par un poids IPCW non nul tous ceux ayant un statut d'événement connu. Cette précaution permet de tenir compte des individus qui auraient eu la même durée de sinistre et pour lesquels les observations sont censurées dans les bases. Par exemple pour l'estimation des durées de sinistre, des poids plus élevés sont attribués aux individus avec des durées de maintien plus importants pour tenir compte du fait qu'ils sont plus susceptibles d'être censurés avant leur rétablissement. Les données ainsi pondérées peuvent ensuite être exploitées par toute technique d'apprentissage automatique pouvant incorporer ces paramètres (des poids IPCW appliqués aux observations).

2.8.2. Modification de l'algorithme CART

Une version d'adaptation de l'algorithme CART aux données censurées (*Tree base censored*) en assurances non-vie est proposée par Lopez et al. (2015) [16]. Ils introduisent une stratégie d'élagage des arbres de régressions intégrant la pénalisation par le poids de Kaplan-Meier permettant de gérer les données censurées et d'obtention de l'arbre maximal.

Notons E l'indicateur de survenance de l'événement que l'on cherche à prédire (tel que : $E = 1$ si l'événement s'est produit et $E = 0$ sinon). Dans le cas simple non pondéré, la proportion de l'échantillon $\hat{\pi}(s)$ au nœud s est calculée comme une moyenne des indicateurs de survenance de l'événement des individus affectés à ce nœud.

A l'étape de test pour un individu représenté par la réalisation des covariables (*features*) notées x, appartenant au nœud terminal s_T, nous pouvons estimer son risque noté $\pi(x)$, comme la proportion des individus i appartenant à ce nœud terminal lors de l'étape d'apprentissage et pour lesquels l'événement est survenu (c'est-à-dire $E_i = 1$). Il est simple d'adapter les arbres de décision pour incorporer les poids IPCW. Dans ce cas, il faut affecter aux individus i lors de la phase d'apprentissage les poids ω_i comme décrits ci-dessus. Ces poids sont utilisés pour la pondération des classes dans l'algorithme d'arbre de décision.

Avec les poids IPCW, nous calculons une diminution pondérée de l'indice de diversité de Gini suivant la formule

$$\Delta I_G^{\omega} = \hat{\pi}^{\omega}(s)[1 - \hat{\pi}^{\omega}(s)] - \frac{1}{N_s^{\omega}}\left\{N_{s_l}^{\omega}\hat{\pi}^{\omega}(s_l)[1 - \hat{\pi}^{\omega}(s_l)] + N_{s_r}^{\omega}\hat{\pi}^{\omega}(s_r)[1 - \hat{\pi}^{\omega}(s_r)]\right\}$$

avec $N_s^{\omega} = \sum_{i \in s} \omega_i$, $N_{s_l}^{\omega} = \sum_{i \in s_l} \omega_i$, $N_{s_r}^{\omega} = \sum_{i \in s_r} \omega_i$, $\hat{\pi}^{\omega}(s) = \frac{\sum_{i \in s} \omega_i E_i}{N_s^{\omega}}$, $\hat{\pi}^{\omega}(s_l) = \frac{\sum_{i \in s_l} \omega_i E_i}{N_{s_l}^{\omega}}$, $\hat{\pi}^{\omega}(s_r) = \frac{\sum_{i \in s_r} \omega_i E_i}{N_{s_r}^{\omega}}$.

Une fois la structure de l'arbre de décision déterminée, la prédiction lors de la phase de test avec les réalisations x des *features* représentant les individus appartenant au nœud terminal s_T , du risque $\hat{\pi}(x)$, est estimée en utilisant la moyenne pondérée

$$\hat{\pi}^{\omega}(x) = \frac{\sum_{i \in s_T} \omega_i E_i}{N_{s_T}^{\omega}},$$

où ω_i correspond au poids IPCW donné précédemment, et $N_{s_T}^{\omega} = \sum_{i=1}^{n} \omega_i \, I_{\{S_i = s_T\}}$.

Notons (Y, δ) les données d'entrée de calcul pour le calcul des poids de Kaplan-Meier comme résultats. On suppose qu'on a une variable T_i de durée dont l'observable est $Y_i = min(T_i, C_i)$ et $\delta = I_{\{T_i \leq C_i\}}$. La formule pour le calcul des poids de Kaplan-Meier ω_i est donnée par

$$\omega_{(i),n} = \frac{\delta_i}{n} \prod_{j=1}^{i-1} \left(\frac{n-j}{n-j-1}\right)^{\delta_i}$$

où $\omega_{(i),n}$ est le poids de Kaplan-Meier associé à la (i)-ème valeur $Y_{(i)}$ classé par ordre croissant.

Notons $(\mathrm{M}, \mathrm{T}, \mathrm{X})$ les données d'entrée d'une base composée de n individus. L'objectif de cet algorithme est de fournir un arbre de régression maximal élagué sur des données contenant des observations censurées et des prédictions des paramètres d'intérêt (par exemple la durée de maintien et la charge sinistre en assurance non-vie).

La construction de l'arbre maximal

Elle consiste d'abord par le calcul de l'estimateur $\widehat{\mathrm{G}}$ suivant la formule analytique proposée pour les n individus. Ensuite une étape d'initialisation de l'algorithme est réalisée en considérant l'arbre avec une feuille unique composée de n_{nc} s individus non censurés ($n_{nc} \leq n$).

Enfin des étapes d'itérations de split ou subdivision des feuilles en nœud. Il s'agit de considérer que l'arbre obtenu à l'étape $s-1$ avec L_{s-1} feuilles où les individus de chaque feuille l correspondant à la classe $T_l^{(s-1)}$ sont distincts de ceux des autres feuilles. Les observations censurées (au nombre de n_c^l) de mêmes caractéristiques, c'est-à-dire $\hat{X} \in T_l^{(s-1)}$ sont assignés à cette feuille.

Pour chaque feuille l, avec $1 \leq l \leq L_{s-1}$ deux cas sont possibles. Si $n_c^l = 1$ ou si toutes les observations ont les même valeurs que $\hat{X}$, alors on ne procède pas à la subdivision (split) de la feuille (cas 1). Si non la feuille devient un nœud du prochain arbre issu de cette étape (cas 2).

Pour le cas 2, le split est réalisé en déterminant les valeurs j_0 et $x_l^{(j_0)}$ qui minimisent $L_l\left(j, x_l^{(j)}\right)$, puis de définir les deux nouveaux sous-ensembles disjoints

$$T_l^{(s-1)} \cap \left\{\hat{X}^{(j_0)} \leq x_l^{(j_0)}\right\} \text{ et } T_l^{(s-1)} \cap \left\{\hat{X}^{(j_0)} > x_l^{(j_0)}\right\}.$$

Le nombre de feuilles devient alors L_s et on passe à l'étape $s+1$ jusqu'à ce que la condition suivante $L_s = L_{s+1}$ soit remplie.

La procédure s'arrête lorsque toutes les feuilles ne peuvent plus être subdivisées.

L'élagage de l'arbre maximal

Cette phase consiste à sélectionner un sous arbre $\hat{S}(\alpha)$ possédant $\hat{K}_\alpha$ feuilles, parmi l'ensemble $\aleph$ des sous arbres de l'arbre maximal (possédant $K_n \leq n$ feuilles) déterminé lors de la phase de construction de l'arbre maximal, tel que

$$\hat{S}(\alpha) = \underset{S \epsilon \aleph}{argmin} \left\{ \int \varphi\left(m, \hat{\pi}^S(t,x)\right) d\hat{F}(m,t,x) + \frac{\alpha K(S)}{n} \right\}$$

avec $K(S)$ le nombre total des feuilles du sous arbre S, α le facteur de pondération du terme de pénalisation $\frac{K(S)}{n}$.

Pour ce faire, l'algorithme est d'abord initialisé en mettant le facteur $\alpha = 0$, puis suivi d'une étape d'incrémentation consistant à augmenter progressivement la valeur de α telle que $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{K_n}$ jusqu'à ce que $\widehat{K}_{\alpha_{j+1}} = \widehat{K}_{\alpha_j}$. Le choix de la valeur optimale α_{j_0} correspond à la valeur minimisant pour un échantillon de taille , noté $(N_i, Y_i, \delta_i, X_i)_{n+1 \leq i \leq n+\tau}$, la formule $\vartheta\left(\alpha_j\right)$ suivante

$$\vartheta\left(\alpha_j\right) = \sum_{i=n+1}^{n+\tau} \frac{\delta_i \varphi\left(N_i, \hat{\pi}^{K\left(\alpha_j\right)}(X_i, Y_i)\right)}{1-\hat{G}(Y_i-)}.$$

L'estimateur de π_0 est donné par la formule suivante

$$\pi^S(t,x) = \sum_{l=1}^{K(S)} \rho_l R_l(t,x).$$

Pour l une feuille du sous arbre optimal est associé un sous-ensemble d'individus T_l (distinct des autres sous-ensemble d'individus et tel que la réunion de tous forme l'ensemble T de tous les n individus) et au critère $R_l(\tilde{x}) = I_{\{\tilde{x} \epsilon T_l\}}$ permettant de déterminer si un individu est affecté ou non à cette feuille.

Le coefficient $\rho_l = \underset{\pi \epsilon X}{argmin}\, E\left[\varphi(M,\pi) \middle| \tilde{X} \in T_l\right]$.

L'estimation du risque est donnée par $\hat{\pi}^{S(\alpha)}$ comme l'estimateur final de π_0.

Limite de l'algorithme *Tree-base censored*

L'une des principales limites de cet algorithme découle de la grande flexibilité de CART qui le rend instable et souvent confronté à un problème de sur-apprentissage sur les données d'entraînement.

Plusieurs procédures d'élagage permettant de contourner ce problème de sur-apprentissage sont proposées avec la plupart impliquant un paramètre de réglage qui limite la complexité de l'arbre [52, 53].

L'une des stratégies consiste à définir une limite inférieure m pour le nombre d'individus affectés à un nœud terminal. Dans notre notation ci-dessus, le nœud S ne serait pas subdivisé (suivant la règle de subdivision utilisée) à moins que $min(N_{s_l}, N_{s_r}) \geq m$.

Cette stratégie est facilement généralisable au cas des données censurées en retenant comme contrainte $min(N_{s_l}^{\omega}, N_{s_r}^{\omega}) \geq m$. Cependant nous notons que $N_{s_l} \approx N_{s_l}^{\omega}$ et $N_{s_r} \approx N_{s_r}^{\omega}$ si la valeur attendue de $\omega_i = 1$. En pratique, fixer une limite inférieure pour $min(N_{s_l}, N_{s_r})$ est généralement suffisante.

Une autre approche consiste à accepter la subdivision d'un nœud dès lors que la diminution de l'indice de diversité de Gini dépasse un certain seuil fixé θ, c'est-à-dire $\Delta I_G(s) \geq \theta$. Substituer $\Delta I_G(s) \geq \theta$ par $\Delta I_G^{\omega}(s) \geq \theta$ permet d'utiliser la même règle dans le paramétrage des données censurées.

Finalement, les valeurs optimales des paramètres de réglage peuvent être choisies par des techniques de validation croisée.

2.8.3. Modification de l'algorithme de forêt aléatoire

Nous proposons une modification de l'algorithme de forêt aléatoire pour l'adapter aux données censurées (*Random Forest Censored*).

Notons M le nombre d'arbres optimaux élagués prédéfini, on dispose d'une base composée de n individus représentés par $z = \{(N_i, Y_i, \delta_i, X_i)_{i=1,\dots,n}\}$. L'objectif de cet algorithme est de fournir des prédictions des paramètres d'intérêt (par exemple la durée de maintien et la charge sinistre en assurance non-vie).

La première étape consiste à initialiser la procédure. Elle permet de calculer de l'estimateur $\widehat{\mathrm{G}}$ suivant la formule analytique proposée pour les n individus et les poids de Kaplan-Meir.

La seconde étape est une itération de l'algorithme *Tree base censred* dans les conditions du *boostrap*. Pour $m = 1$ à M , on fait un tirage aléatoire dans la base z d'un échantillon *boostrap* (avec remise) noté $z_m^{(b)}$ et on estime l'arbre de régression optimal élagué $\hat{f}_{z_m^{(b)}}$ suivant les critères de l'algorithme *Tree base censred* afin d'obtenir M modèles.

La dernière étape consiste à l'agrégation des M modèles, l'estimateur par la méthode *bagging* est tout simplement la moyenne des prédictions de chaque modèle, calculer avec la formule suivante

$$\hat{f}_{bag}(x) = \frac{1}{m}\sum_{m=1}^{M} \hat{f}_{z_m^{(b)}}(x).$$

Ainsi, un estimateur du paramètre d'intérêt ou du risque π_0 serait $\hat{f}_{bag}$.

2.8.4. Modification du Gradient boosting censuré (Gradient tree boosting censored)

La démarche proposée est similaire à la précédente. Nous proposons une modification de l'algorithme du *Gradient Boosting* pour l'adapter aux données censurées (*Gradient tree boosting censored*).

Notons M le nombre d'arbres optimaux élagués prédéfini, on dispose d'une base composée de n individus représentés par $z = \{(N_i, Y_i, \delta_i, X_i)_{i=1,\dots,n}\}$. L'objectif de cet algorithme est de fournir des prédictions des paramètres d'intérêt (par exemple la durée de maintien et la charge sinistre en assurance non-vie).

La première étape consiste à initialiser la procédure. D'une part, elle permet de calculer de l'estimateur $\hat{G}$ suivant la formule analytique proposée pour les n individus et les poids de Kaplan-Meir. D'autre part, de la fonction de perte

$$\hat{f}_0 = \underset{\gamma}{argmin} \sum_{i=1}^{n} l^{(p)}(y_i, \gamma).$$

La seconde étape est récursive. Pour $m = 1$ à M , on calcule

$$r_i^m = -\left[\frac{\partial l^{(p)}(y_i, f(x_i))}{\partial f(x_i)}\right]_{f=f_{m-1}},$$

pour $i = 1, \dots, n$.

Puis on ajuste l'arbre de régression optimal élagué δ_m au couple $(x_i, r_i^m)_{i=1,\dots,n}$ avec l'algorithme *Tree-base censored*.

Ensuite on calcul γ_m en résolvant

$$\min_{\gamma} \sum_{i=1}^{n} l^{(p)}\left(y_i, f_{m-1}(x_i) - \gamma \delta_m(x_i)\right).$$

Enfin on met à jour $\hat{f}_m(x) = \hat{f}_{m-1}(x) - \gamma_m \delta_m(x)$.

Ainsi, un estimateur du paramètre d'intérêt ou du risque π_0 serait la valeur finale $\hat{f}_M$.

2.9. Adaptation des métriques d'évaluation en présence de données censurées

2.9.1. La métrique MSEW

La métrique MSE est souvent utilisées dans le cas de la régression lorsque les observations sont complètes. La prise en compte d'observations censurées dans les données conduit à une adaptation de la MSE par l'introduction du poids de Kaplan-Meier. Il s'agit de la version pondérée de l'erreur quadratique moyenne MSEW (*Weighted Mean Square Error*) par les poids de Kaplan-Meier en présence de données censurées.

Formellement, on écrit

$$MSEW = \frac{1}{n}\sum_{i=1}^{n} \omega_i \times \left(Y_i - f(X_i)\right)^2,$$

avec i les individus ($(i = 1, \dots, n)$, X_i le vecteur des covariables, Y_i la variable à prédire (*target*) et $f(.)$ la fonction d'estimation du modèle, ω_i le poids de Kaplan-Meier de l'individu.

2.9.2. Amélioration du MSEW par ajustement à l'unité de la valeur cible

Pour se ramener à l'unité, on peut prendre la racine de la MWSE et l'on obtient la RMSEW (*Weighted Root Mean Squared Error*),

$$RMSEW = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \omega_i \times \left(Y_i - f(X_i)\right)^2}.$$

La RMSEW ne se comporte pas très bien quand les étiquettes peuvent prendre des valeurs qui s'étalent sur plusieurs ordres de grandeur. Pour prendre cela en compte, nous proposons une transformation logarithmique des valeurs prédites et des vraies valeurs avant de calculer la RMSEW. Nous obtenons ainsi la *RMSLEW* (*Weighted Root Mean Squared Log Error*)

$$RMSLEW = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \omega_i \times (\log(Y_i + 1) - \log(f(X_i) + 1))^2}.$$

2.9.3. Coefficient de détermination pondérée

Même si les valeurs à prédire ont toutes le même ordre de grandeur, la RMSEW peut être difficile à interpréter. Nous proposons pour ce faire de compléter notre analyse par une amélioration de l'erreur carrée relative ou RSE (*Relative Squared Error*), notée l'erreur carrée relative pondérée, ou RSEW (*Weighted Relative Squared Error*) plus facilement interprétable.

Cet indicateur est le résultat de la normalisation de la somme pondérée des carrés des résidus non pas par le nombre de points n dans le jeu de données, mais par une mesure de ce qu'il serait raisonnable de faire comme erreur. Il s'agit de la somme pondérée des distances entre chacune des valeurs à prédire et leur moyenne.

$$RSEW = \frac{\sum_{i=1}^{n} \omega_i \times \left(Y_i - f(X_i)\right)^2}{\sum_{i=1}^{n} \omega_i \times (Y_i - \bar{Y})^2},$$

avec $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

Le coefficient de détermination pondéré, encore plus facile à interprèter, correspond au complémentaire du RSEW est

$$R_W^2 = 1 - RSEW.$$

Ce coefficient de détermination nous indique donc à quel point les valeurs prédites sont corrélées aux vraies valeurs.

2.9.4. Taux de vrais positifs, taux de faux positifs et l'indice de concordance (C-index)

En apprentissage automatique, pour mesurer la performance des modèles, notamment pour les méthodes de classification, il est d'usage d'utiliser des indicateurs comme la courbe ROC (*Receiver Operating Characteristic*) et l'AUC (*Area Under Curve*).

Une courbe ROC est un graphique représentant les performances d'un modèle de classification pour tous les seuils de classification. Cette courbe trace le taux de vrais positifs en fonction du taux de faux positifs. Soient VP le nombre de vrais positifs, FN le nombre de faux négatifs, alors le taux de vrais positifs TVP est défini comme suit

$$TVP = \frac{VP}{VP+FN}.$$

De même, soient FP le nombre de faux positifs, VN le nombre de vrais négatifs, alors le taux de faux positifs TFP est défini comme suit

$$TFP = \frac{FP}{FP+VN}.$$

Une courbe ROC trace les valeurs TVP et TFP pour différents seuils de classification. Diminuer la valeur du seuil de classification permet de classer plus d'éléments comme positifs, ce qui augmente le nombre de faux positifs et de vrais positifs.

Pour calculer les points d'une courbe ROC, nous pourrions par exemple effectuer plusieurs évaluations d'un modèle de régression logistique en variant les seuils de classification, mais ce serait inefficace. Nous pouvons en revanche calculer efficacement l'aire sous cette courbe, ou AUC, grâce à un algorithme de tri. AUC signifie « aire sous la courbe ROC ». Cette valeur mesure l'intégralité de l'aire à deux dimensions située sous l'ensemble de la courbe ROC (par calculs d'intégrales) de (0,0) à (1,1). L'AUC fournit une mesure agrégée des performances pour tous les seuils de classification possibles.

On peut interpréter l'AUC comme une mesure de la probabilité pour que le modèle classe un exemple positif aléatoire au-dessus d'un exemple négatif aléatoire. Les valeurs d'AUC sont comprises dans une

plage de 0 à 1. Un modèle dont 100 % des prédictions sont erronées à un AUC de 0,0. Si toutes ses prédictions sont correctes, son AUC est de 1,0.

L'AUC présente les avantages suivants :

- ✓ L'AUC est invariante d'échelle. Elle mesure la qualité du classement des prédictions, plutôt que leurs valeurs absolues.
- ✓ L'AUC est indépendante des seuils de classification. Elle mesure la qualité des précisions du modèle quel que soit le seuil de classification sélectionné.

Toutefois, ces deux avantages comportent des limites qui peuvent réduire la pertinence de l'AUC dans certains cas d'utilisation.

- ✓ L'invariance d'échelle n'est pas toujours souhaitable. Par exemple, nous avons parfois besoin d'obtenir des probabilités précisément calibrées, ce que l'AUC ne permet pas de déterminer.
- ✓ L'indépendance vis-à-vis des seuils de classification n'est pas toujours souhaitable. Lorsque des disparités importantes de coût existent entre les faux négatifs et les faux positifs, il peut être essentiel de minimiser l'un des types d'erreur de classification. Par exemple, dans un contexte de détection de spam il sera probablement préférable de minimiser en priorité les faux positifs (même si cela entraîne une augmentation significative des faux négatifs). L'AUC n'est pas un critère à retenir pour ce type d'optimisation.

L'indice de concordance (C-index) est un indicateur qui permet de valider et de confirmer le choix du meilleur modèle sur la base des métriques de performance présentées dans les paragraphes précédents. Lorsque la variable à expliquer est complètement observée sur tous les individus, l'indice de concordance (C-index) est l'équivalent de la métrique AUC. Il correspond à la probabilité de classer correctement les résultats pour une paire d'individus choisis au hasard et dont les valeurs prédites sont différentes.

Le C-index peut être adapté pour les cas de données censurées en considérant la concordance des résultats de survie versus la probabilité de survie prédite entre des paires d'individus dont les résultats de survie peuvent être ordonnés.

Dans notre cas, cela revient à dire parmi les paires d'individus non censurés car tous les deux rétablis (observations complètes), ou l'un est rétabli et l'autre est censuré après le rétablissement du premier avant la fin d'observation.

La formule de l'indice C-index adapté au cas de données censurées est la suivante

$$C_{cens}(t) = \frac{\sum_{i \neq j} \delta_i I_{\{V_i < V_j\}} I_{\{\hat{\pi}(X_i) < \hat{\pi}(X_j)\}}}{\sum_{i \neq j} \delta_i I_{\{V_i < V_j\}}},$$

où $I_{\{.\}}$ est une fonction indicatrice.

3. Données et hyper paramètres

Les données proviennent des bases de gestion d'un assureur de personnes pour les contrats d'assurance de prêts et de prévoyance collective. Nous disposons d'informations individuelles quantitatives et qualitatives sur les assurés et les sinistrés. Pour constituer un historique d'observation suffisante, nous avons regroupé les données de plusieurs dates d'arrêté (de 2012 à 2016). Il s'agit des bases utilisées lors des travaux de calculs des provisions lors des travaux d'inventaires.

Enfin, nous considérons les sinistres non clôturés (sinistres ouverts) aux différentes dates d'arrêtés comme des censures à droite.

3.1. Description des données des contrats de prêts

3.1.1. Données disponibles

La base d'étude contient des données de sinistres (ouverts ou clôturés) de plusieurs réseaux de distributions. Elle regroupe les dates d'arrêtés comprises entre 2012 et 2016, et contenant des sinistres observés depuis 2004. La taille totale des effectifs de la population à fin 2016 est de **879 553 d'individus**. Il s'agit d'un portefeuille regroupant des prêts immobiliers, des prêts à la consommation et des prêts professionnels. Les garanties considérées couvrent le remboursement de tout ou d'une partie des échéances des prêts en cas de survenance d'un arrêt de travail (code sinistre 20) et de perte d'emploi (code sinistre 30).

3.1.2. Risques et garanties

Nous analysons dans cette étude trois différentes natures de prêts. Les **Prêts Immobiliers** concernent les crédits accordés dans le but de financier une opération immobilière. Il peut s'agir d'une acquisition d'un bien immobilier, de construction ou de travaux sur une propriété. Les durées de remboursements peuvent s'étaler jusqu'à 50 ans et sa durée moyenne est de 13 ans. Les **Prêts Professionnels** qui sont des prêts réservés aux professionnels pour répondre à des besoins comme le financement d'équipements, la construction ou l'acquisition de locaux. Dans certains cas des besoins de trésorerie. Ces prêts s'adressent généralement aux professions libérales, artisans, commerçants, très petites entreprises etc... Ils sont généralement moins longs que les prêts immobiliers et ont une durée moyenne de 8 ans. Enfin, les **Prêts à la consommation** accordés à des particuliers dans le but de financer des achats de biens et de service (équipement de la maison, achat de voiture ou autre biens...). Ils ont une durée de remboursement relativement courte et des montants empruntés plus faible que les prêts immobiliers. Ils concernent également les prêts pour travaux inférieurs à 140 000 euros. Ces prêts ont une durée moyenne bien plus courte que les prêts immobiliers et sont en moyenne de 5 ans.

3.1.3. Statistiques descriptives

Les différentes variables d'intérêt extraites du système de gestion sont décrites dans le tableau ci-dessous. La population globale est d'âge moyen à la souscription de 37,64 ans et de 43,53 ans à la survenance du sinistre. Elle est composée de 53,42% d'hommes et de 43,62% de femme, contre 0,01% de personnes morales et 2,92% de données manquantes. La variable « Code type échéance » permet de différencier les prêts à taux fixes des prêts à taux variables. Nous comptons 85,79% de prêts à taux fixe, 8,99% de prêts à taux variable et 2,16% de prêts relais contre 3,07% de données manquantes pour cette variable. La variable « Nature de prêt » représente les trois catégories de prêts (immobilier, professionnel et consommation. La base contient 88% de prêts immobilier, presque 10% de prêts professionnels et seulement 2% de prêts à la consommation, contre 0,01% de données manquantes.

Enfin, il faut noter que pour chacune des trois variables « Montant moyen échéance assurance mensuelle », « Montant sinistre » et « Montant des prestations », leur troisième quartile ($75^{ème}$ percentile) est inférieur à leur valeur moyenne. Ce qui indique que 75 % des données sont inférieures à la valeur moyenne. Ce qui est expliqué par le poids du maximum dans le calcul de la moyenne.

Nom des variables	Valeurs maquantes (%)	Minimum	1er quartile	Médiane	Moyenne	3ème quartile	Maximum
Code Clôture (0 : ouvert / 1 : clôturé)	0%						
Code Catégorie socio professionnelle (1 à 8)	3%						
Code Département	0%						
Code Produit	0%						
Code Risque	0%						
Code sexe	2,92%						
Code type échéance	3,07%						
Montant capiltal initial	2,92%	-	12 911	33 000	51 097	71 926	201 012 690
Montant Capiltal Restant Dû	0%	-	-	1 053	20 749	24 201	65 185 763
Montant moyen échéance assurance mensuelle	**17,03%**	-	68	230	**863**	**454**	404 000 000
Montant sinistre	0,00%	-	310	1 251	**4 096**	**3 949**	495 853
Nature de prêt	0,01%						
Taux d'emprunt	2,93%	1,90%	3,60%	4,40%	4,50%	4,90%	10,50%
Date de chargement de la base	0%						
Date d'arrêté	0%						
Date de fin de garantie	0%						
Date de survenance	0%						
Date de clôture sinistre	0%						
Date de fin d'indemnisation	0%						
Date de dernier paiement	0%						
Date de souscription	0%						
Date de début d'indemnisation	0%						
Date de naissance	0%						
Date de fin de prêt	0%						
Age souscription	0%	18	30	37	37,64	45	66
Age survenance	0%	18	37	44	43,53	51	68
Durée du sinistre	0%	0,00	3,43	7,54	13,12	15,45	155,61
Délais avant survenance du sinistre	0%	0,00	2,55	4,99	5,89	8,42	35,33
Montant des prestations	0%	-	144,70	882,00	**3 129,60**	**3 083,70**	495 852,80

Tableau 3.1.3.1 – Statistiques descriptives des variables d'intérêt

Les individus sont classés dans huit différentes catégories de CSP : Agriculteurs exploitant (code n° 1), Artisans, commerçants et chefs d'entreprise (code n° 2), Cadre et professions intellectuelles supérieures (code n° 3), Professions intermédiaires (code n° 4), Employés (code n° 5), Ouvriers (code n° 6), Personnels de service (code n° 7) et Autres catégories (code n° 8).

Finalement, nous remarquons que pour la grande majorité des variables nous disposons de données complètes. Le point d'attention concerne spécifiquement la variable « Montant moyen échéance assurance mensuelle », correspondant à la moyenne des échéances mensuelles payées par l'assuré au titre de sa garantie d'assurance, pour laquelle 17% des données sont manquantes.

L'analyse par nature de prêt fournie des informations plus détaillées sur les variables d'intérêt. La répartition suivant le critère du genre est très différente, 71% d'hommes pour les prêts professionnels et 59% pour les prêts à la consommation, contre 53% pour les prêts immobiliers. Cette répartition plus équilibrée du genre pour les prêts immobiliers est expliquée par la présence de prêts conjoints de personnes vivant en couple sur ce segment.

Les âges moyens à la souscription suivant le critère du genre sont proches pour les prêts à la consommation (proche de 43 ans) et les prêts professionnels (autour de 41,5 ans). Nous notons une différence de 1,88 années de l'âge moyen à la souscription des prêts immobiliers (37,93 ans pour les hommes contre 36,05 ans pour les femmes). Les âges moyens de survenance du sinistre suivant le même critère sont proches pour les prêts à la consommation (proche de 45,5 ans) et les prêts professionnels (autour de 46 ans), contre une différence de 2,65 ans pour les prêts immobiliers (44,32 ans pour les hommes contre 41,67 ans pour les femmes). La situation est un peu différente quand nous regardons la valeur médiane de ces variables. Les âges médians sont identiques pour les prêts à la consommation (44 ans), sont différents pour les prêts immobiliers de 2 ans (en faveur des femmes) et pour les prêts professionnels de 1 an (en faveur des hommes).

			Age à la souscription (années)					Age de survenance du sinistre (années)				
Nature de prêt	Genre	Répartition	Minimum	Médiane	Moyenne	Maximum	Ecart-type	Minimum	Médiane	Moyenne	Maximum	Ecart-type
Consommation	Homme	59%	18	44	42,73	64	9,62	18	47	45,29	64	9,63
	Femme	41%	18	44	43,13	64	9,02	19	47	45,85	64	8,98
Immobilier	Homme	53%	18	37	37,93	63	8,45	18	45	44,32	66	8,62
	Femme	47%	18	35	36,05	66	9	19	42	41,67	68	9,43
Professionnel	Homme	71%	18	42	41,44	63	8,77	20	47	46,05	64	8,47
	Femme	29%	18	43	41,89	62	8,29	21	48	46,46	64	8,14

Tableau 3.1.3.2 – Ages atteints

Les délais d'attente avant survenance des sinistres sont les plus faibles pour les prêts à la consommation, car de durée de garantie en général plus faible que les autres natures de prêts. Ils sont plus élevés pour les prêts immobiliers de durée de garantie en générale plus longue. Les délais d'attente des prêts professionnels se trouvent entre les deux autres délais. Cette variable semble donc dépendre de la durée des garanties. La valeur médiane suit globalement la même logique.

		Délais d'attente avant survenance sinsitre (années)				
Nature de prêt	Genre	Minimum	Médiane	Moyenne	Maximum	Ecart-type
Consommation	Homme	0	2,08	2,55	19,35	2,14
	Femme	0	2,21	2,72	17,79	2,27
Immobilier	Homme	0	5,58	6,39	35,33	4,37
	Femme	0	4,73	5,63	27,43	4,02
Professionnel	Homme	0	3,97	4,6	30,08	3,29
	Femme	0	4,04	4,58	28,91	3,15

Tableau 3.1.3.3 – Délais avant survenance du sinistre

L'analyse des indicateurs de sinistralité comme me montant des sinistres réglés et la durée des sinistres (en mois), les sinistres clôturés (observations complètes) et les sinistres ouverts (observations incomplètes) nous donnent des informations complémentaires sur les risques étudiés.

Pour le montant des sinistres réglés la médiane est autour de 1 150,00 euros pour les différentes natures de prêts. Les valeurs moyennes suivant ce critère sont proches de 4 083,00 euros pour les prêts immobiliers et professionnels contre une valeur autour de 2 500,00 euros pour les prêts à la consommation (dont les montants empruntés sont plus faibles).

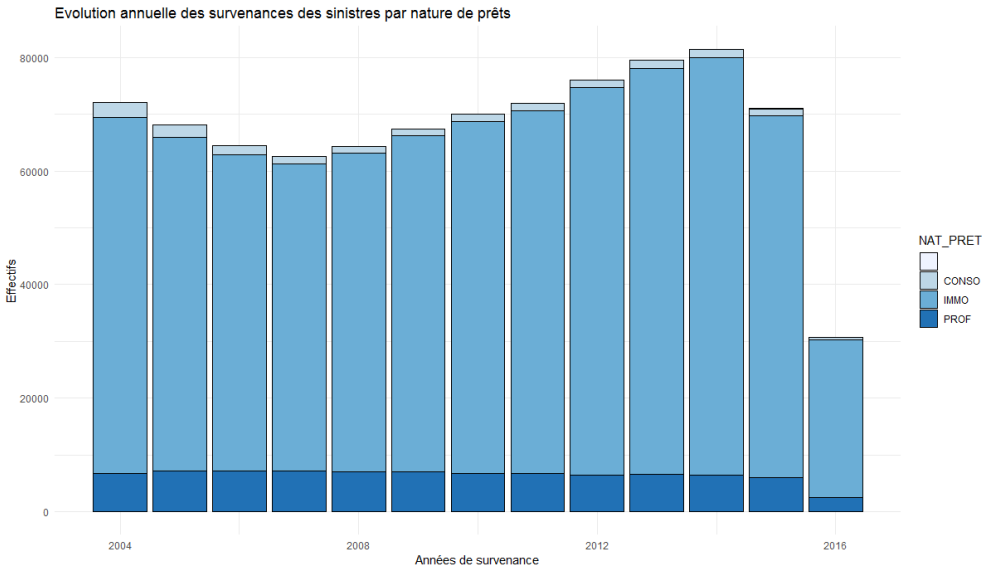
Pour les prêts immobiliers, la différence des valeurs moyennes des durées des sinistres par genre reste faible (13,06 mois pour les hommes et 13,64 ans pour les femmes). Ce constat est identique pour les valeurs médianes pour cette nature de prêt (valeur autour de 7,5 mois). La durée moyenne des sinistres des hommes pour les prêts professionnels est la plus faible (9,4 mois), elle est inférieure à toutes les autres durées moyennes des sinistres par genre et nature de prêt. Ce constat est également conservé pour la valeur médiane.

		Montant sinistre (millier €)					Durée sinistre (mois)				
Nature de prêt	Genre	Minimum	Médiane	Moyenne	Maximum	Ecart-type	Minimum	Médiane	Moyenne	Maximum	Ecart-type
Consommation	Homme	-	1	3	46	4	0	7,55	11,79	111,88	12,59
	Femme	-	1	3	59	4	0	8,94	13,33	112,97	13,96
Immobilier	Homme	-	1	4	423	9	0	7,33	13,06	155,61	17,1
	Femme	-	1	4	324	9	0	7,81	13,64	155,58	17,36
Professionnel	Homme	-	1	4	496	11	0	5,42	9,4	134,42	11,74
	Femme	-	1	4	423	12	0	7	11,25	140,29	13,05

Tableau 3.1.3.4 – Indicateurs de prestations : montant et durée

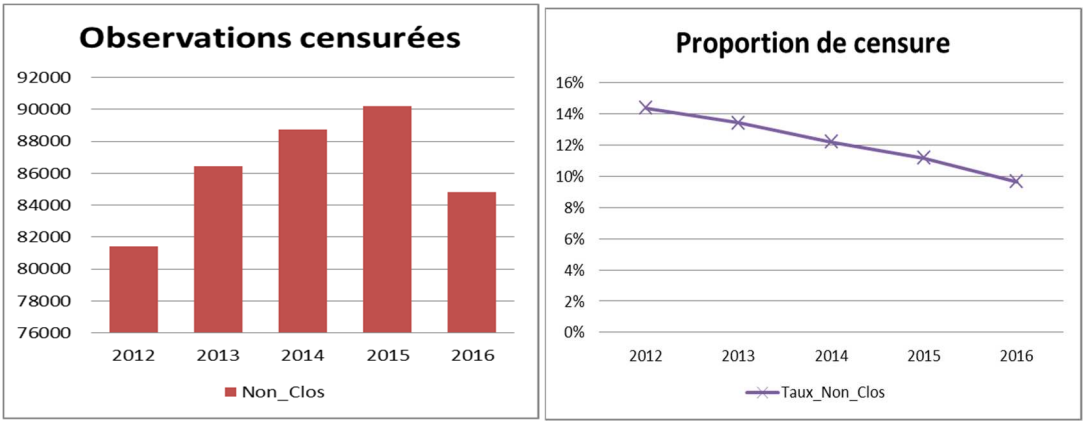
L'analyse du nombre des sinistres ouverts par année de survenance montre des résultats cohérents avec la proportion de chaque nature de prêts dans la base. Les prêts immobiliers représentent la ma-

jorité des contrats. Après une baisse du nombre des sinistres ouverts entre 2004 et 2007, nous observons une évolution contraire depuis 2008, avec une tendance à la hausse atteignant en 2014 un pic de plus de 80 000 sinistres ouverts (dont 73 458 pour les prêts immobiliers). Cette évolution va dans le même sens que celle des expositions (nombre total des assurés dans le portefeuille). La baisse observée en 2015, de l'ordre de 13% par rapport à 2014, est expliquée par les impacts positifs des nouvelles conditions de prise en charge des sinistres des prêts immobiliers mises en place entre 2013 et 2014 pour corriger la tendance haussière de la sinistralité observée les années précédentes. Le niveau des sinistres ouverts en 2016 est quant à lui expliqué par l'effet conjoint des retards de déclarations et de traitements en gestion des sinistres survenus en 2016.



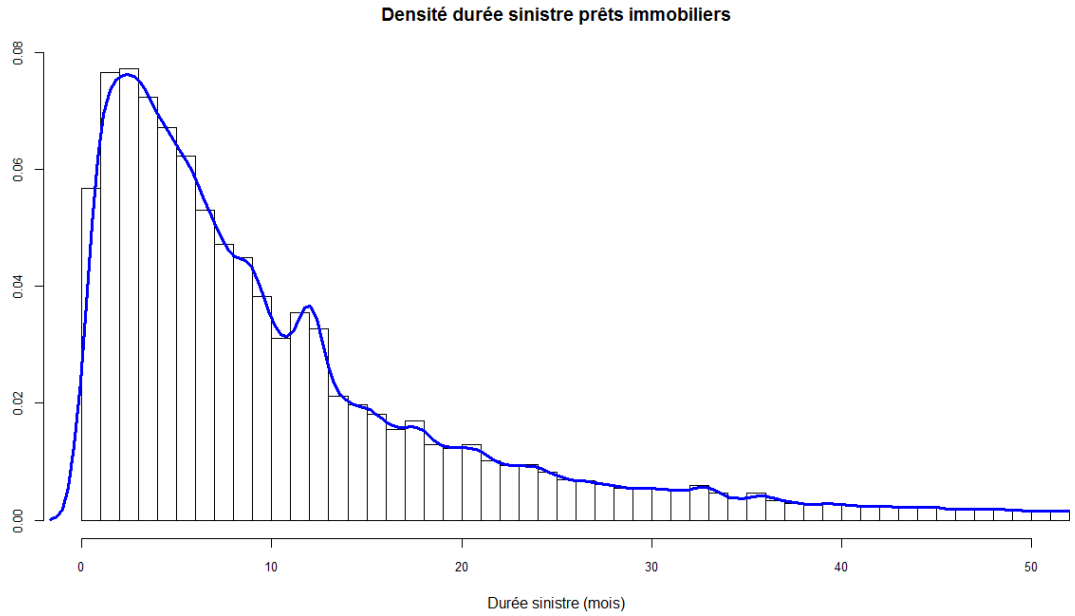
Graphique 3.1.3.1 – Nombre de sinistres ouverts par année de survenance

Les cas de censure à droite correspondent aux sinistres en cours à chaque date d'arrêté, une fois les périodes de franchise dépassées. La tendance depuis 2012 des observations censurées par années d'arrêté est haussière sauf en 2016 où l'on constate une baisse en nombre de l'ordre de 6000 par rapport à 2015. L'augmentation annuelle par rapport aux effectifs totaux des sinistres varie de 14% (2012) à 11% (2015). En 2016, le taux de censure représente un peu moins de 10% de l'effectif total des sinistres.



Graphique 3.1.3.2 – Sinistres ouverts

Dans ce qui suit, afin de disposer d'éléments complémentaires d'analyses sur les variables d'intérêt (la durée des sinistres et la charge sinistre), nous analysons la répartition de ces paramètres dans la base arrêtée à fin 2016 correspondant au stock des sinistres entre 2004 et 2016. Nous concentrons notre analyse sur les indicateurs de sinistralité des prêts immobiliers.



Graphique 3.1.3.3 – Répartition des durées de sinistres prêts immobiliers

Globalement, sans faire de distinction entre les sinistres clos et les censures (sinistres encore ouverts à la date d'arrêté 2016) les effectifs diminuent en fonction de la durée de maintien. Nous notons toutefois des pics de sortie de l'état à 3 mois, 11 mois et 12 mois.

	Durée sinistre					
	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Sinistre clos	0	3,26	7,10	11,83	14,10	151,16
Sinistre ouvert	0,03	8,42	17,77	28,29	38,03	155,61

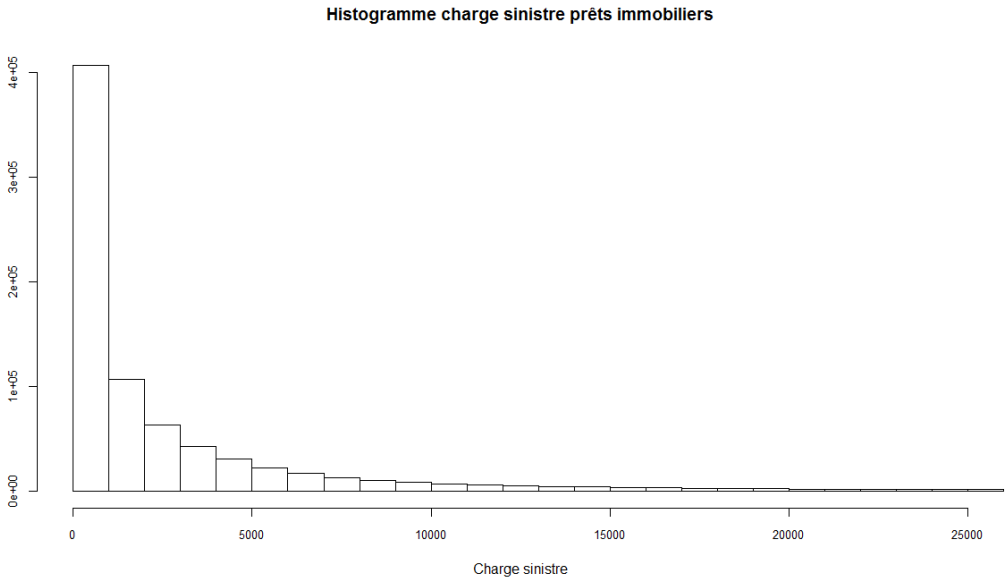
Tableau 3.1.3.5 – Indicateurs de dispersion des durées de sinistre prêts immobiliers

La durée moyenne des sinistres clôturés est de 11,83 mois (presque qu'une année en moyenne), le premier quartile est de moins de 4 mois (0,27 ans) et un maximum à 151,16 mois (presque 12,6 ans). Les sinistres censurés ont des durées de maintien plus élevées que les celles des sinistres clôturés, ce quel que soit l'indicateur de dispersion, avec une valeur moyenne 2,4 fois plus grande.

		Durée sinistre					
Risques	Etat	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Arrêt de travail	Sinistre clos	0,00	3,13	6,76	11,86	14,03	151,16
	Sinistre ouvert	0,03	8,58	18,52	28,95	39,44	155,61
Perte d'emploi	Sinistre clos	0,00	6,79	11,93	11,37	14,63	82,63
	Sinistre ouvert	0,74	5,94	9,65	11,55	13,97	120,19

Tableau 3.1.3.6 – Indicateurs de dispersion par risque des durées de sinistre prêts immobiliers

L'analyse par risque (garantie d'arrêt de travail et garantie perte d'emploi ou chômage), donne des résultats qui s'expliquent par les conditions de prise en charge des sinistres dans les deux cas. Les valeurs maximales sont très différentes (12,6 ans pour l'arrêt de travail contre 6,9 ans pour la perte d'emploi).



Graphique 3.1.3.4 – Répartition de la charge sinistre des prêts immobiliers

La charge sinistre moyenne des sinistres clos est de 3 473,00 euros contre 10 124 euros pour les sinistres censurés.

	Charge sinistre					
	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Sinistre clos	0,00	275,80	1149,80	3472,90	3565,50	423462,30
Sinistre ouvert	0,00	926,10	3646,60	10124,20	11548,80	310355,80

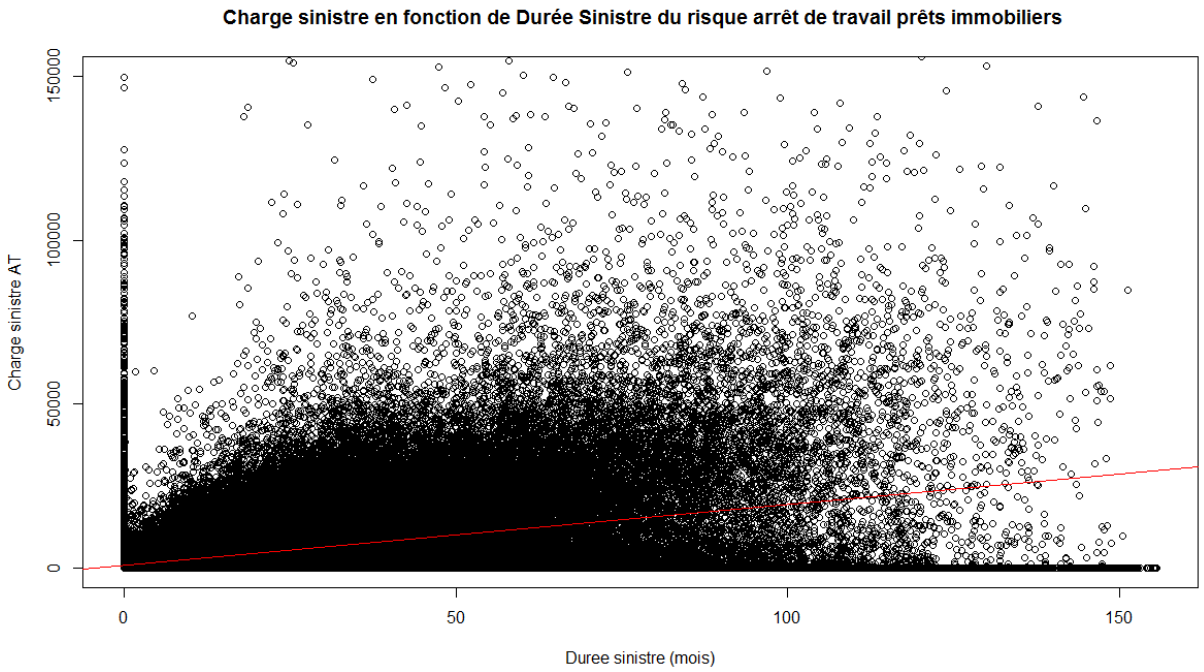
Tableau 3.1.3.7 – Indicateurs de dispersion de la charge sinistre des prêts immobiliers

L'analyse par risque (garantie d'arrêt de travail et garantie perte d'emploi ou chômage), donne des résultats cohérents avec le constat fait pour la durée des sinistres.

		Charge sinistre					
Risques	Etat	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Arrêt de travail	Sinistre clos	0,00	260,20	1102,10	3540,10	3551,50	423462,30
	Sinistre ouvert	0,00	979,20	3890,90	10446,80	12094,90	310355,80
Perte d'emploi	Sinistre clos	0,00	590,10	1737,90	2626,70	3662,30	95679,40
	Sinistre ouvert	0,00	382,20	1199,00	1960,60	2713,30	21992,90

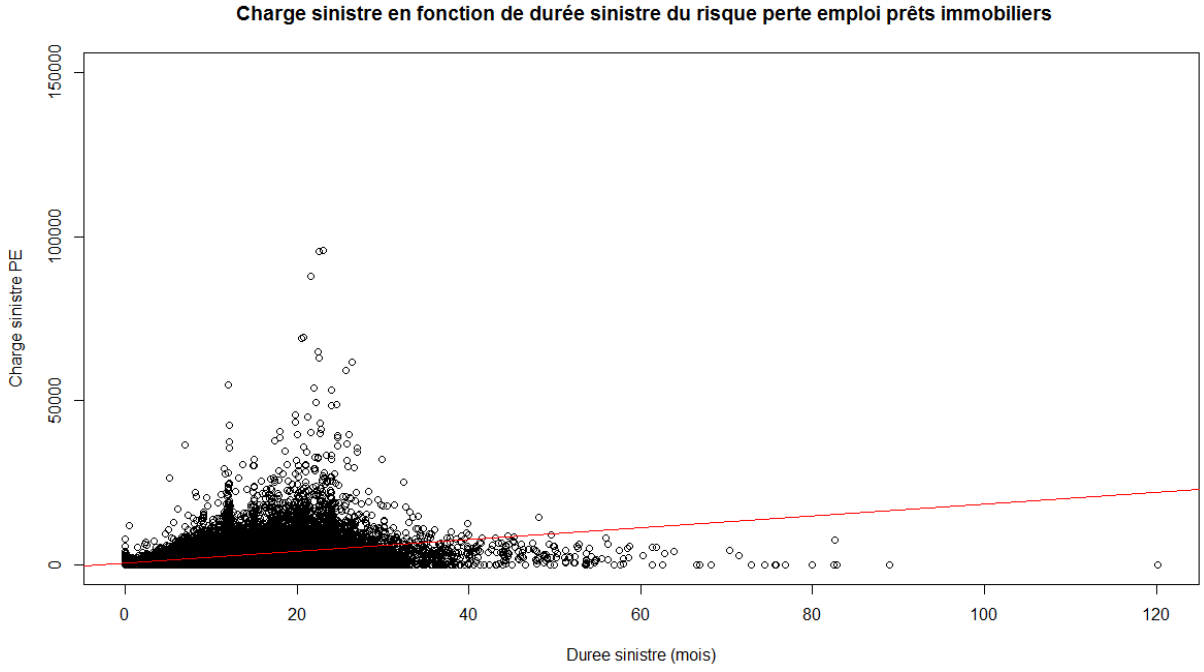
Tableau 3.1.3.8 – Indicateurs de dispersion par risque de la charge sinistre des prêts immobiliers

Enfin, l'analyse croisée entre durée sinistre et charge sinistre du risque arrêt de travail ne permet pas de dégager une structure particulière de la distribution. Nous observons toutefois une concentration de la charge sinistre en dessous du seuil de 50 000,00 euros et une durée de sinistre de 100 mois.



Graphique 3.1.3.5 – Répartition de la charge sinistre en fonction de la durée des sinistres du risque arrêt de travail

La structure est plus concentrée pour le risque perte d'emploi. Nous observons toutefois une concentration de la charge sinistre en dessous du seuil de 15 000,00 euros et une durée de sinistre de 40 mois.



Graphique 3.1.3.6 – Répartition de la charge sinistre en fonction de la durée des sinistres du risque perte d'emploi

Pour terminer cette partie, nous proposons de vérifier statistiquement l'existence ou non de lien entre les variables « durée du sinistre » et « charge sinistre » par un test statistique. Nous avons observé que

les distributions de ces deux ne suivent pas une loi normale. Dans ce cas nous optons pour le teste de Spearman (Rho de Spearman).

Test Rho de Spearman : existe-il un lien entre durée du sinistre et charge sinistre?			
Risques	**Arrêt de travail**	**Perte d'emploi**	**Global**
S	4.0589e+16	1.3921e+13	4.9485e+16
Rho	0.3550038	0.4791637	0.3673198
Conclusion	NON	NON	NON

Tableau 3.1.3.9 – résultats du test statistique de Rho de Spearman

Dans les trois cas étudiés, avec un seuil de confiance fixé à 95%, nous pouvons rejeter l'hypothèse nulle de l'existence d'un lien entre les variables « durée du sinistre » et « charge sinistre » pour le portefeuille composé de contrats de prêts étudié.

3.2. Description des données des contrats de prévoyance collective

3.2.1. Données disponibles

Le portefeuille de contrats de prévoyance collective est assez différent du portefeuille de contrats de prêts par la taille des effectifs sinistrés, les risques sous-jacents aux garanties assurées, les variables explicatives et la période d'historique disponibles. Pour la prévoyance collective, les risques étudiés sont l'arrêt de travail (code risque 27) et l'accident de travail (code risque 32). La population assurée est composée d'agents hospitaliers (code H) et d'agents des collectivités locales territoriales (code P).

La base brute regroupe les dates d'arrêtés comprises entre 2014 et 2017. Elle comprend les flux de prestations par année où chaque ligne correspond à une prestation reçue par l'assuré sinistré durant son sinistre. Ainsi, le nombre de ligne total associé à chaque sinistre correspond au nombre total de prestations perçues à la date de fin d'observation (date d'arrêté de la base). Pour réaliser l'étude, un retraitement par regroupement des lignes par sinistre a été réalisé, permettant d'obtenir pour chaque sinistre une ligne unique de prestations. Nous posons comme critères de regroupement à chaque date d'arrêté deux conditions. La première est telle que la durée du sinistre soit égale à la somme des durées des différentes périodes de prestations. La seconde stipule que le montant total du sinistre est la somme de toutes les prestations payées au cours de ces périodes d'indemnisation. Ces règles permettent de gérer les rechutes. Après retraitement, la taille totale des effectifs de la population en 2017 est de **347 904 individus**.

3.2.2. Risques et garanties

La fonction publique se décompose en trois statuts législatifs dont la fonction publique d'Etat, la fonction publique territoriale et la fonction publique hospitalière. Chaque statut a ses particularités mais ils tendent à s'aligner. Le contrat d'assurance de prévoyance collective ne doit son existence qu'aux obligations statutaires et réglementaires dues par les collectivités locales (territoriales et hospitalières) à leurs agents.

Les différents risques qui concernent les agents hospitaliers et territoriaux sont : la maladie ordinaire, la longue maladie, la maladie longue durée et l'accident de travail. Le positionnement en disponibilité prévoit entre autres, la disponibilité d'office pour maladie. Les différents risques couverts et cités ci-dessus ont chacun des durées et montants de prestations différents.

La **garantie maladie ordinaire** (code MAL) couvre les maladies dûment constatées de nature à mettre l'agent dans l'impossibilité d'exercer ses fonctions, de durée maximale de 12 mois consécutifs. Les prestations sont des indemnités journalières égales à 100% du traitement pendant 3 mois, puis 50% du traitement pendant 9 mois (ce taux est porté à 66.66% si l'agent à au moins 3 enfants à charge). La fin de droit intervient lors de la reprise d'activité à temps pleins, lors de la mise en disponibilité d'office en cas d'inaptitude non définitive ou lors de la mise à la retraite d'office si l'agent rempli les conditions d'âge et de durée de cotisations.

La **garantie longue maladie** (code LM) couvre les maladies rendant nécessaires un traitement et des soins prolongés et présentant un caractère invalidant et de gravité confirmée. La durée maximale est de 3 ans consécutifs. Les prestations sont des indemnités journalières égales à 100% du traitement pendant 1 an, puis 50% du traitement pendant 2 ans (ce taux est porté à 66.66% si l'agent à au moins 3 enfants à charge). La fin de droit intervient lors de la reprise d'activité à temps pleins, de la reprise d'activité à mi-temps thérapeutique, de la mise à la retraite d'office, de la mise en disponibilité d'office en cas d'inaptitude non définitive à toute reprise d'activité.

Le passage à la **couverture maladie longue durée** (LD), s'il doit avoir lieu, a lieu 1 an après une période de longue maladie à plein traitement. La maladie longue Durée est un congé accordé uniquement en cas d'affections bien définies. La durée maximale est de 5 ans ou 8 ans si la maladie a été contractée dans l'exercice des fonctions. Les prestations sont des indemnités journalières égales à 100% du traitement pendant 3 ans, puis 50% du traitement pendant 2 ans (ce taux est porté à 66.66% si l'agent à au moins 3 enfants à charge). La fin de droit fait suite à une reprise d'activité à temps pleins, à une reprise d'activité à mi-temps thérapeutique, à une mise en disponibilité d'office en cas d'inaptitude non définitive à toute reprise d'activité.

La **Disponibilité d'Office pour maladie** (code DO) est assimilée au risque type invalidité car ils font suite, au sens classique du terme (notion de passage), à un risque incapacité. Elle couvre la situation d'un agent inapte à reprendre ses fonctions à l'expiration d'un congé MAL, LM ou LD et dont l'inaptitude n'est pas jugée définitive. L'agent alors « sorti » du positionnement statutaire activité, ne relève plus du régime spécial des agents des collectivités locales. La DO n'est donc sollicitée qu'en dernier recours quand l'agent n'a plus droit à aucun congé maladie. Cet état est donc unique dans la carrière de l'agent. Durée maximale : Le statut distingue la durée du positionnement de la durée d'indemnisation. La durée du positionnement est comptabilisée à l'expiration des droits à congés pour maladies, elle est limitée à 3 ans (par durée d'un an renouvelable), éventuellement 4 ans, si l'agent est jugé apte à reprendre son activité avant la fin de la 4éme année. La durée d'indemnisation (de 3 ans maximum) est fixée à compter de la date de l'interruption initiale du travail. Cette distinction peut donc conduire un agent à ne plus ou ne pas percevoir d'indemnisation pendant sa DO. L'agent peut alors avoir intérêt à être admis à la retraite d'office si les conditions sont réunies ou, en attendant la fin du positionnement DO et sa mise à la retraite pour invalidité, à être reconnu en état d'invalidité temporaire et ainsi bénéficier de l'allocation d'invalidité temporaire. Prestations : indemnité journalière égale à 50% du traitement (supplément familial versé intégralement et indemnité de résidence au prorata du salaire. Le taux est porté à 66,66% si l'agent à au moins 3 enfants à charge). Fin de droits : réintégré, mis à la retraite, radié des cadres.

Nous pouvons également citer les deux autres risques d'invalidité ne faisant pas partie du périmètre de notre étude : le **Mi-Temps Thérapeutique** (code MTT), l'**Allocation d'invalidité temporaire** (INV) et le **Capital décès** (DC).

Enfin, la **garantie l'accident du travail** couvre les accidents survenus par le fait où à l'occasion du travail de toute personne salarié ou travaillant, à quelque titre ou en quel que lieu que ce soit, pour un ou plusieurs employeurs ou chefs d'entreprise. Il s'agit alors d'accident de service, d'accident de trajet ou de maladie professionnelle contractés dans l'exercice des fonctions. La durée maximale s'étend jusqu'à ce que l'agent puisse reprendre ses fonctions ou jusqu'à sa consolidation. Les prestations sont des indemnités journalières égales à 100% du traitement pendant 3 ans et intégralité des frais médicaux et pharmaceutiques.

Type d'arrêt	Nature de l'arrêt	Indemnités journalières STATUT : **obligation de la collectivité**	Indemnités journalières RELAIS AU STATUT : **organisme assureur**
MALADIE ORDINAIRE (Article 57-2°, Loi n° 84-53, 25/01/1984, Fonction publique territoriale)	Maladie sans gravité particulière rendant impossible l'exercice des fonctions	*Du 1er au 90ème jour* : 100 % du salaire *Du 91ème au 360ème jour* : 50 % du salaire	*Du 1er au 90ème jour* : Néant *Du 91ème au 360ème jour* : $(x-50)$ % du salaire
LONGUE MALADIE (Article 57-3°, Loi n° 84-53, 26/01/1984, Fonction publique territoriale)	Maladie (cf. annexe A.1) de gravité confirmée, non imputable au service, à caractère invalidant nécessitant un traitement et des soins prolongés	*Du 1er au 360ème jour* : 100 % du salaire *Du 361ème au 1 080ème jour* : 50 % du salaire	*Du 1er au 360ème jour* : Néant *Du 361ème au 1 080ème jour* : $(x-50)$ % du salaire
MALADIE DE LONGUE DURÉE (Article 57-4°, Loi n° 84-53, 26/01/1984, Fonction publique territoriale)	Maladie (cf. annexe A.2) de gravité confirmée à caractère invalidant nécessitant un traitement et des soins prolongés	*Du 1er au 1 080ème jour* : 100 % du salaire *Du 1 081ème au 1 800ème jour* : 50 % du salaire	*Du 1er au 1 080ème jour* : Néant *Du 1 081ème au 1 800ème jour* : $(x-50)$ % du salaire
DISPONIBILITÉ D'OFFICE (Article 62°, Loi n° 86-53, 09/01/1986, Fonction publique territoriale)	Inaptitude temporaire à la reprise des fonctions lorsque tous les droits statutaires de l'agent aux congés pour la MO, la LM et la LD sont épuisés	Néant (sauf dans le cas exceptionnel du versement d'une indemnité de coordination fixée à 50 % du salaire)	*Du 1er au 360ème jour* : renouvelable 3 fois au plus à x % du salaire ($(x-50)$ % si indemnité de coordination)

Tableau 3.2.2 – Synthèse des principaux arrêts de travail étudiés

3.2.3. Statistiques descriptives

Les variables explicatives extraites du système de gestion sont : le numéro de contrat, la date de début de prestations, la date de fin de prestations, le numéro du sinistre, la collectivité, la date de survenance du sinistre, la cause du sinistre (accident, maladie, etc…), le montant réglé, la date de souscription, la date de naissance, le risque (maladie ordinaire, longue maladie, maladie longue durée, accident de travail). Les différentes variables d'intérêt retenues pour la prédiction des risques et utilisées dans les modèles d'apprentissage automatique sont décrites dans le tableau ci-dessous.

Nom des variables	Description de la variable
Code Clôture (0 : ouvert / 1 : clôturé)	Indicateur de sinsitre clos à date de fin d'observation (0 : sinistre ouvert / 1 : sinistre clôturé)
Code HP (H ou P)	H : personnel hospitalier et P : agent collectivité locale territoriale
Code sexe	1 : Homme (53,42%) / 2 : Femme (43,62%)
Num Collectivité	Numéro de la collectivité locale hospitalière ou territoriale
Code Activité NAF	Code nomenclature d'activité française de la collectivité
Nom Risque	Accident de Travail (AT) / Arrêt de Travail (DO : Disponibilité d'Office pour Maladie, LD : Longue Durée, LM : Longue Maladie, MAL : Maladie Ordinaire)
Code Risque	27 : Accident de Travail / 32 : Arrêt de Travail
Durée du sinistre	Durée de règlement des sinistres
Prestations Payées	Montant total des sinistres réglés à la date du clôture
Age survenance	Age atteint par l'assuré sinistré à la survenance du sinistre (par différence de millésime)
Durée Flux	Durée totale des règlements des sinistres estimée à partir des différentes périodes de prestations
Montant des prestations	Charge sinistre estimée à partir de la variable "Prestations Payées" et de la date de clôture du sinistre

Tableau 3.2.3.1 (a) –Variables d'intérêt

Après retraitement, la base ne contient aucune donnée manquante. La population globale est d'âge moyen de survenance du sinistre de 45,67 ans. L'âge maximum de couverture est de 76 ans et le minimum à 16 ans (l'âge légal de travail dans les collectivités locales). Le portefeuille est composé de 36% d'hommes et de 64% de femme. Il faut noter que pour chacune des variables « Durée du sinistre », « Durée Flux », « Prestations Payées » et « Montant des prestations », leur troisième quartile (75ème percentile) est inférieur à leur valeur moyenne.

Nom des variables	Minimum	1er quartile	Médiane	Moyenne	3ème quartile	Maximum
Code Clôture (0 : ouvert / 1 : clôturé)						
Code HP (H ou P)						
Code sexe						
Num Collectivité						
Code Activité NAF						
Nom Risque						
Code Risque						
Durée du sinistre	-	0,33	0,83	**3,32**	2,60	220,50
Prestations Payées	-	206,20	745,20	**3 544,20**	2 635,30	309 294,30
Age survenance	16	39	47	45,67	53	76
Durée Flux	-	0,33	0,83	**3,37**	2,67	220,50
Montant des prestations	-	206,80	751,50	**3 605,90**	2 698,60	321 004,70

Tableau 3.2.3.1 (b) – Statistiques descriptives des variables d'intérêt

Dans la date d'arrêté fin 2017, la base est composée de 92% des sinistres sont des agents des collectivités territoriales contre 8% de personnel hospitalier. Nous notons que 70% des sinistres sont couverts par la garantie arrêt de travail contre 30% couverts au titre de la garantie accident de travail.

Collectivités	Effectif	Proportion
Hospitalière (H)	29 283	8%
Territoriale (P)	318 621	92%
Total	347 904	100%

Garanties	Effectif	Proportion
Accident de Travail (27)	104 674	30%
Arrêt de Travail (32)	243 230	70%
Total	347 904	100%

Tableau 3.2.3.2 (a) – Répartition des sinistres par collectivité et garanties

Le risque maladie ordinaire (MAL) représente 61,32% des effectifs totaux. Le risque LD/LM/MAL correspondant à 2,47% des effectifs totaux sont des individus qui sont en maladie longue durée (LD) après avoir changé d'état de santé en passant par la maladie ordinaire (MAL) puis par la longue maladie (LM).

Risques	Effectif	Proportion
AT	104 674	30,09%
DO	1 709	0,49%
LD	5 827	1,67%
LM	13 760	3,96%
MAL	213 350	61,32%
LD/LM/MAL	8 584	2,47%
Total	347 904	100%

Tableau 3.2.3.2 (b) – Répartition des sinistres par risque

Les âges moyens de survenance des sinistres par genre varient entre 44 ans et 51 ans suivant le type de risque. Parmi les individus sinistrés, 42,96% sont des femmes couvertes par la garantie arrêt de travail pour maladie ordinaire (contre 18,36% d'hommes) et d'âge moyen de survenance de 45,25 ans (contre 46,20 ans pour les hommes parmi lesquels se trouve l'âge maximal de survenance de 76 ans). Les sinistres au titre de la garantie accident de travail sont représentés par 15,71% de femme (contre 14,37% d'hommes) d'âge moyen de survenance de 45,84 ans (contre 44,10 ans pour les hommes).

			Age de survenance du sinistre (années)				
Garantie / Risque	**Genre**	**Répartition**	**Minimum**	**Médiane**	**Moyenne**	**Maximum**	**Ecart-type**
Accident de travail	Homme	14,37%	16	45	44,1	67	9,83
	Femme	15,71%	17	47	45,84	70	9,45
Arrêt de travail	Homme	0,18%	25	51	49,57	63	7,73
Disponibilité d'office (DO)	Femme	0,31%	22	50	48,78	62	7,95
Arrêt de travail	Homme	0,52%	21	52	50,19	65	7,56
Longue Durée (LD)	Femme	1,15%	24	50	49,1	64	7,83
Arrêt de travail	Homme	1,44%	22	52	50,56	65	7,56
Longue Maladie (LM)	Femme	2,51%	21	50	48,99	66	7,92
Arrêt de travail	Homme	18,36%	17	48	46,2	76	9,66
Maladie Ordinaire (MAL)	Femme	42,96%	17	46	45,25	73	9,95
Arrêt de travail	Homme	0,70%	22	50	48,17	64	7,71
(LD/LM/MAL)	Femme	1,77%	21	48	47,18	64	7,54

Tableau 3.2.3.3 – Ages à la survenance par type de risque

Les durées maximales des indemnisations diffèrent selon le type d'arrêt. Les risques MAL et DO sont des risques à court déroulement i.e. de durée inférieure à 1 an, et que les risques LM et LD sont plutôt des risques à long déroulement i.e. de durée supérieure à 1 an (voir Tableau 3.2.2 – Synthèse des principaux arrêts de travail étudiés).

En moyenne, les sinistres d'arrêt de travail du risque LD sont ceux qui coûtent les plus chers. La durée moyenne des sinistres est de 32,37 mois pour les femmes (31,63 mois pour les hommes) pour un montant sinistre moyen 41 150 € pour les femmes (43 850€ pour les hommes). Le sinistre le plus court et le moins coûteux en moyenne est le sinistre survenu au titre du risque MAL. Pour ce risque, la durée moyenne des sinistres est de 1,73 mois pour les femmes (1,77 mois pour les hommes) pour un montant sinistre moyen de 1350 € pour les femmes (1580 € pour les hommes).

Les valeurs de la durée moyenne (19,50 mois pour les hommes) et du montant sinistre moyen (20 080€ pour les hommes) du risque LM restent cohérentes par rapport aux autres indicateurs. Elles sont comprises entre celles des risques MAL et LD. En effet, comme vu précédemment la durée maximale couverte du risque LM est de 3 ans consécutifs, alors que celle du risque MAL est de 12 mois consécutifs et celle du risque LD est de 5 ans (ou 8 ans si la maladie a été contractée dans l'exercice des fonctions).

Les valeurs de la durée moyenne (10,36 mois pour les hommes) et du montant sinistre moyen (7 140€ pour les hommes) du risque DO restent également cohérentes par rapports à celles des autres risques. En effet, elle couvre la situation d'agent inapte à reprendre ses fonctions à l'expiration d'un congé MAL, LM ou LD et dont l'inaptitude n'est pas jugée définitive. Elle n'est donc sollicitée qu'en dernier recours quand l'agent n'a plus droit à aucun congé maladie.

La durée moyenne (7,21 mois pour les hommes) et le montant sinistre moyen (5 490 € pour les hommes) du risque LD/LM/MAL sont inférieurs à 1 an. Ce statut correspond aux sinistres requalifiés en arrêt de travail maladie longue durée (LD) après avoir changé d'état de santé en passant par la maladie ordinaire (MAL) puis par la longue maladie (LM). L'indicateur de ce risque ne tient pas compte des délais déjà passés dans les états précédents.

		Durée sinistre (mois)					Montant sinistre (millier €)				
Garantie / Risque	**Genre**	**Minimum**	**Médiane**	**Moyenne**	**Maximum**	**Ecart-type**	**Minimum**	**Médiane**	**Moyenne**	**Maximum**	**Ecart-type**
Accident de travail	Homme	-	0,53	2,15	162,37	5,17	-	0,90	3,41	219,91	8,28
	Femme	-	0,63	2,91	220,5	6,33	-	0,96	4,11	321,00	9,48
Arrêt de travail	Homme	-	6,87	10,36	55,8	9,29	-	4,87	7,14	47,80	6,80
Disponibilité d'office (DO)	Femme	-	8,8	11,38	81,93	9,79	-	5,49	7,34	46,96	6,47
Arrêt de travail	Homme	-	27,63	31,63	147,87	19,82	-	39,91	43,85	238,58	29,01
Longue Durée (LD)	Femme	-	28,23	32,37	165,4	20,2	-	37,93	41,15	213,14	26,96
Arrêt de travail	Homme	-	16,47	19,5	119,83	13,16	-	18,81	20,08	114,30	14,86
Longue Maladie (LM)	Femme	-	15,9	19,01	145,73	12,87	-	16,88	17,56	116,42	13,36
Arrêt de travail	Homme	-	0,73	1,77	42,13	2,9	-	0,56	1,58	96,87	2,54
Maladie Ordinaire (MAL)	Femme	-	0,77	1,73	41,67	2,76	-	0,47	1,35	39,10	2,19
Arrêt de travail	Homme	-	5,93	7,21	48,03	5,09	-	4,79	5,49	29,48	3,73
(LD/LM/MAL)	Femme	-	5,97	7,56	41,77	4,94	-	4,63	5,32	29,16	3,57

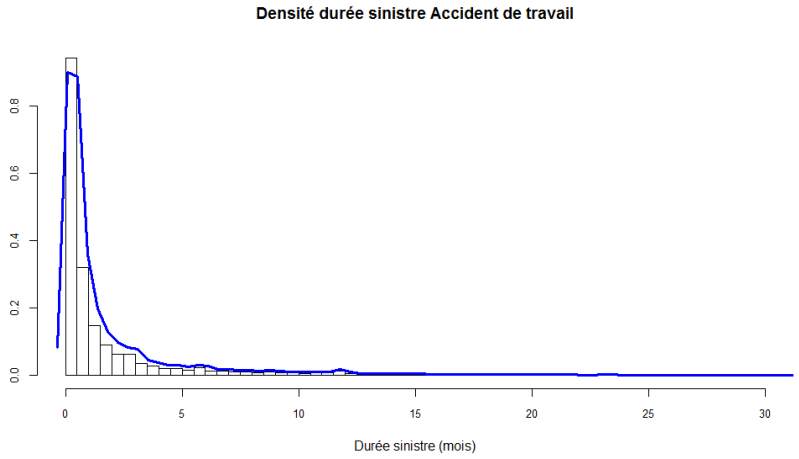
Tableau 3.2.3.4 – Indicateurs de prestations : durée et montant sinistre

Notons que pour la plupart des risques, les valeurs maximales de l'indicateur durée sinistre dépassent la durée totale d'indemnisation prévue dans les contrats. Ceci s'explique par la méthode retenue pour regrouper sur une même ligne, tous les sinistres de même type pour un même assuré. Il s'agit donc d'une vision maximaliste et conservatrice de la durée totale de sinistre, allant dans le sens de la prudence. Les effectifs concernés sont négligeables.

Nous analysons dans ce qui suit les distributions des indicateurs de sinistralité (durée et montant des sinistres) par garanties (accident de travail et arrêt de travail).

- **Garantie accident de travail**

Pour la garantie accident de travail, la distribution de la durée des sinistres (clos et censurés) est globalement décroissante. Ceci s'explique par la présence d'un nombre important de sinistres à faible durée de maintien pour cette garantie. Nous notons toutefois des pics de sortie de l'état (à 1 mois, 11 mois et 12 mois).



Graphique 3.2.3.1 – Densité des durées de sinistre accident de travail

L'analyse des indicateurs de dispersion permet de noter que la durée de sinistres clôturés moyenne est de 2,29 mois avec un maximum à 220,50 mois (presque 18,38 ans). Les sinistres censurés ont une durée moyenne plus élevées à 6,27 mois avec une valeur moyenne maximale plus faible à 162,37 mois (13,53 ans).

Durée sinistre (mois)								
Garantie	Etat	Effectif	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Accident de travail	Sinistre clos	97 959	0,00	0,23	0,53	2,29	1,80	220,50
	Sinistre ouvert	6 715	0,00	0,47	2,27	6,27	7,90	162,37

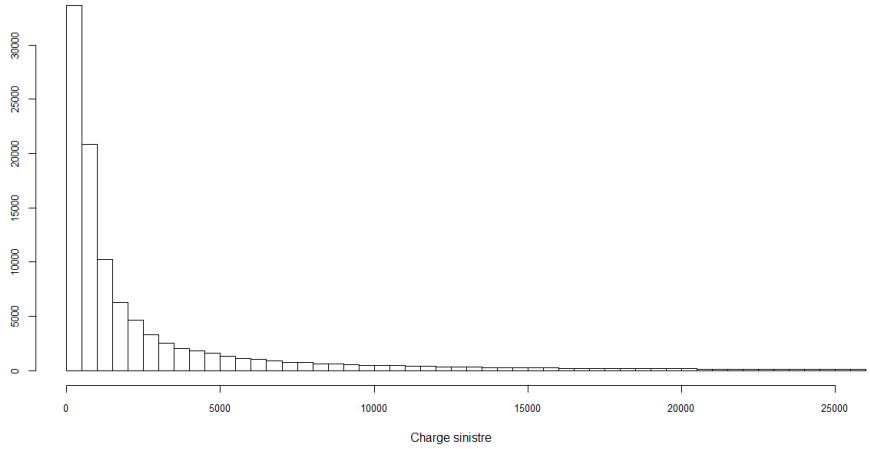
Tableau 3.2.3.5 – Indicateurs de dispersion des durées de sinistre accident de travail

L'analyse des indicateurs de dispersion de la charge sinistre permet de faire des commentaires principalement sur la valeur maximale des sinistres censurés supérieure à celle des sinistres clos, alors que c'est le cas inverse pour la durée de sinistre. Les autres indicateurs vont dans le même sens que celui de la durée de sinistres.

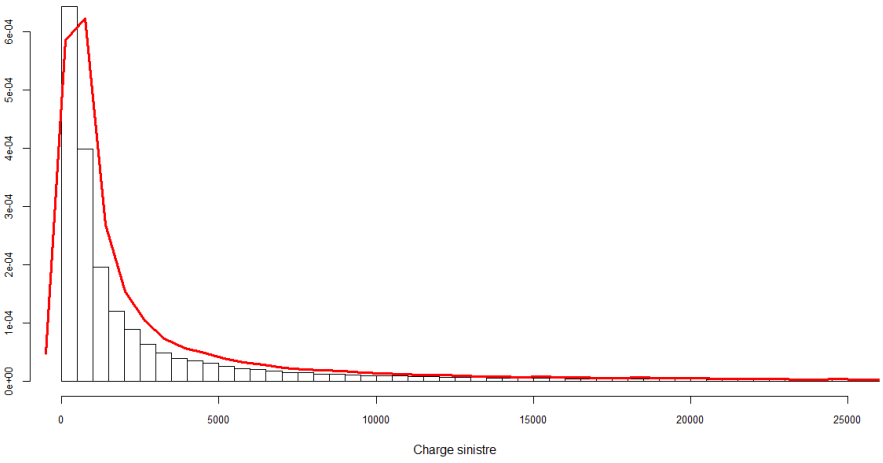
Charge sinistre								
Garantie	Etat	Effectif	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Accident de travail	Sinistre clos	97 959	-	372,30	879,80	3 338,70	2 632,90	219 912,10
	Sinistre ouvert	6 715	-	801,30	3 309,40	10 102,30	11 568,40	321 004,70

Tableau 3.2.3.6 – Indicateurs de dispersion de la charge sinistre accident de travail

L'effectif des sinistres clos est 15 fois supérieur à celui des sinistres censurés. La valeur médiane est à 879,80 euros pour les sinistres clos (3 309,40 euros pour les sinistres censurés). La valeur moyenne est à 3 338,70 euros pour les sinistres clos (11 568,40 euros pour les sinistres censurés). Nous notons également que la distribution de la charge sinistre est également décroissante.

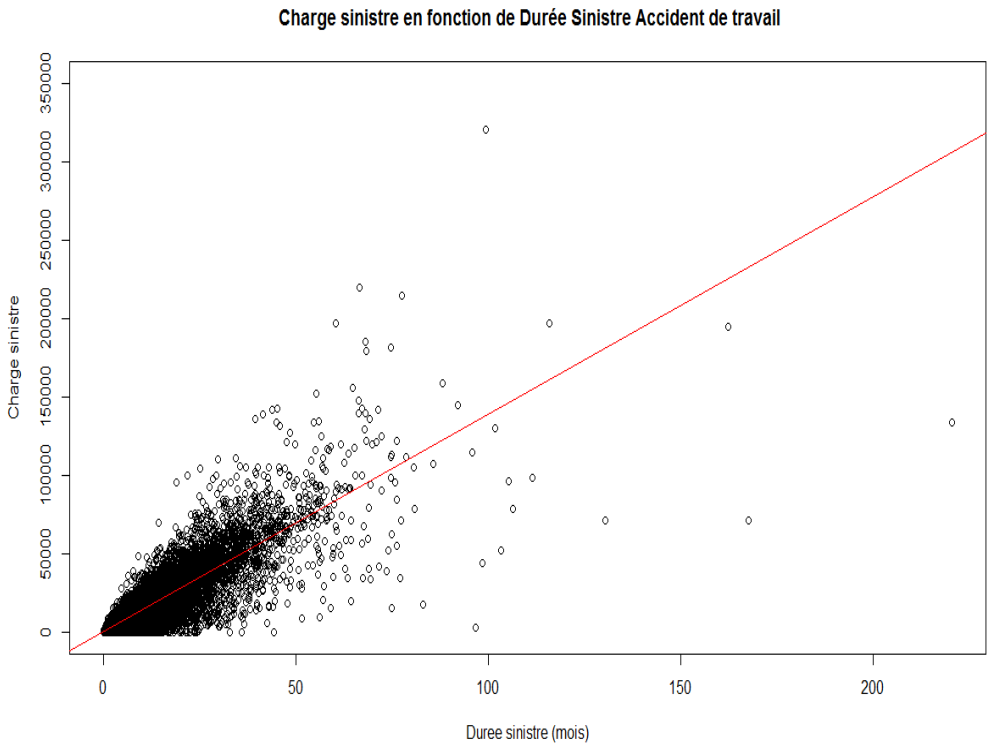


Graphique 3.2.3.2 (a) – Répartition de la charge sinistre accident de travail



Graphique 3.2.3.2 (b) – Densité de la charge sinistre accident de travail

Enfin, l'analyse croisée entre durée sinistre et charge sinistre de la garantie accident de travail permet d'observer une répartition de la distribution presque homogène autour de la diagonale. Nous observons une concentration de la charge sinistre en dessous du seuil de 75 000 euros de charge sinistre et de 50 mois de durée de sinistre.



Graphique 3.2.3.3 – Répartition de la charge sinistre en fonction de la durée de sinistre accident de travail

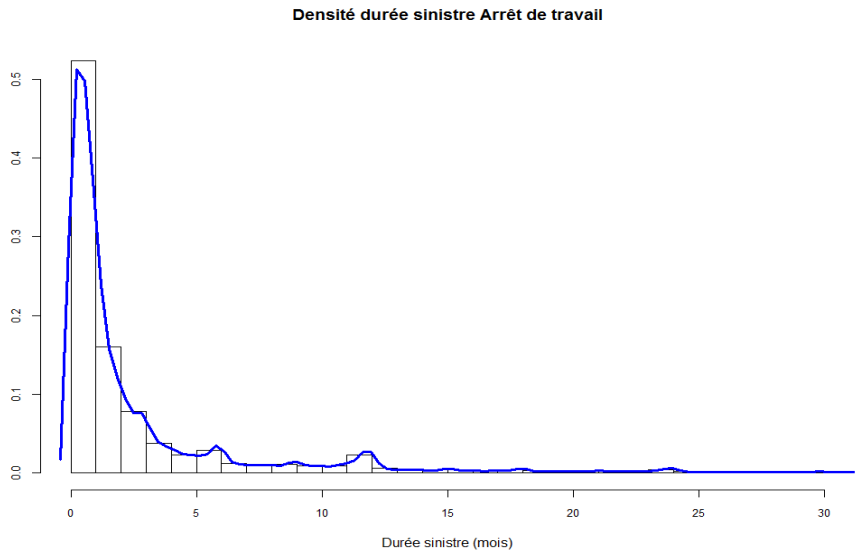
Le test statistique de Spearman conduit, avec un seuil de confiance fixé à 95%, à accepter l'hypothèse alternative d'existence d'un lien entre les variables « durée du sinistre » et « charge sinistre » (Rho est égal à 92,25% donc non nul) pour la garantie accident de travail.

Test Rho de Spearman : existe-il un lien entre durée du sinistre et charge sinistre?	
Risques	Accident de travail
S	1.4805e+13
Rho	0.9225231
Conclusion	OUI

Tableau 3.2.3.7 – résultats du test statistique de Rho de Spearman

- **Garantie arrêt de travail**

Pour la garantie arrêt de travail, la distribution de la durée des sinistres (clos et censurés) est également globalement décroissante. Ceci s'explique par la présence d'un nombre important de sinistres à faible durée de maintien pour cette garantie. Nous notons également plusieurs pics de sorties (à 1 mois, 6 mois et 12 mois).



Graphique 3.2.3.4 – Densité des durées de sinistre arrêt de travail

L'analyse des indicateurs de dispersion permet de noter que les durées de sinistres moyennes de la garantie arrêt de travail sont supérieures à celles de la garantie accident de travail. La durée de sinistres clôturés moyenne est de 3,34 mois (contre 2,29 mois pour l'accident de travail). La valeur maximale de la durée de sinistres est de 160,53 mois (presque 13,37 ans contre 18,37 ans pour l'accident de travail). Les sinistres censurés ont une durée moyenne plus élevées à 8,15 mois avec une valeur moyenne maximale plus faible à 165,40 mois (13,78 ans).

Durée sinistre (mois)								
Garantie	Etat	Effectif	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Arrêt de travail	Sinistre clos	223 882	0,00	0,37	0,90	3,34	2,63	160,53
	Sinistre ouvert	19 348	0,00	0,57	2,43	8,15	10,13	165,40

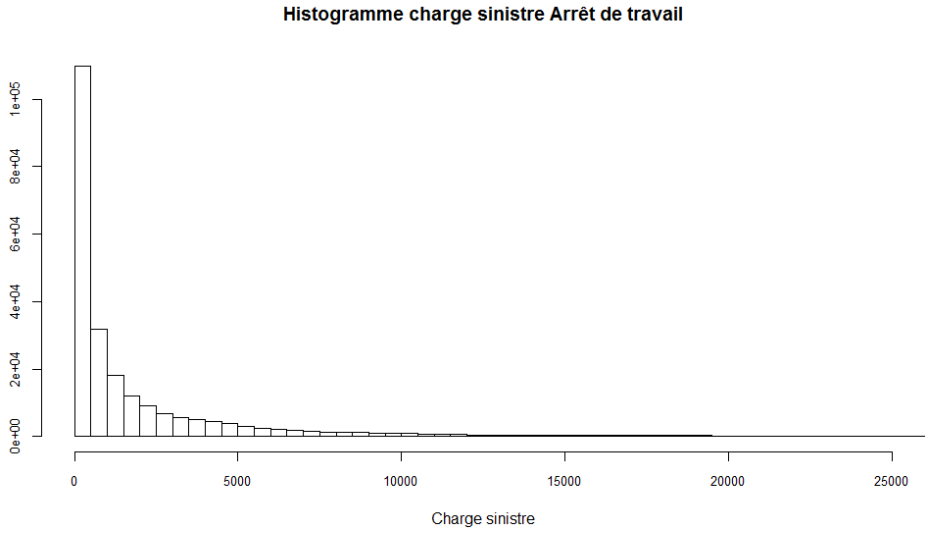
Tableau 3.2.3.8 – Indicateurs de dispersion des durées de sinistre d'arrêt de travail

L'analyse des indicateurs de dispersion de la charge sinistre montre des indicateurs allant dans le même sens que celui de la durée de sinistres.

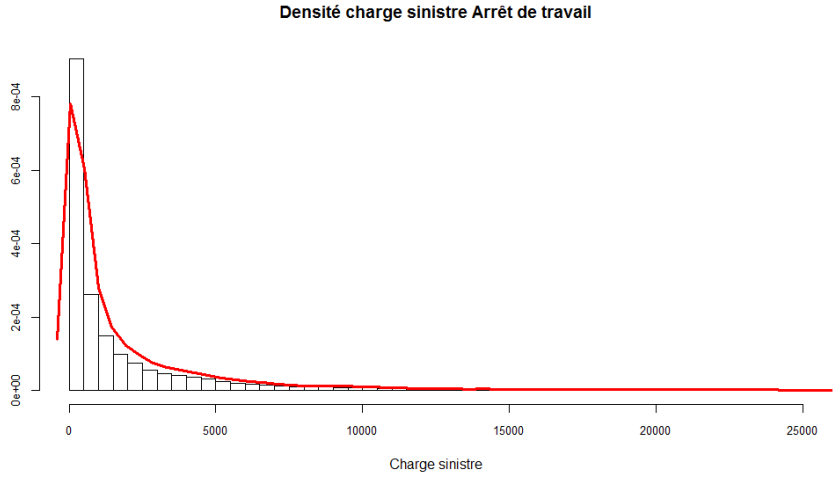
Charge sinistre								
Garantie	Etat	Effectif	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Arrêt de travail	Sinistre clos	223 882	-	133,80	599,50	2 963,50	2 278,20	238 584,60
	Sinistre ouvert	19 348	-	407,40	2 529,40	10 138,00	10 025,60	184 426,50

Tableau 3.2.3.9 – Indicateurs de dispersion de la charge sinistre d'arrêt de travail

L'effectif des sinistres clos est 11,57 fois supérieur à celui des sinistres censurés. La valeur médiane est à 599,50 euros pour les sinistres clos (2 529,4 euros pour les sinistres censurés). Ces valeurs médianes sont plus faibles que celles de la garantie accident de travail. C'est aussi le cas de la valeur moyenne des sinistres clos. Elle est de 2 963,5 euros (contre 3 338,70 euros pour la garantie accident de travail). Par contre, pour les sinistres censurés, les moyennes sont assez proches avec une valeur de 10 138,00 euros (contre 10 102,3 euros pour la garantie accident de travail). Nous notons toujours que la distribution de la charge sinistre est également décroissante.



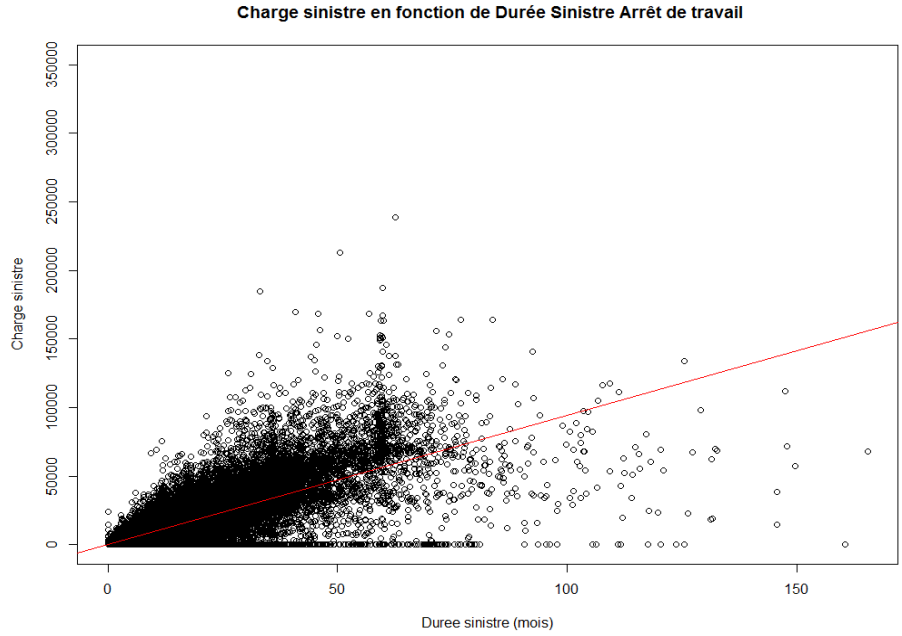
Graphique 3.2.3.5 (a) – Répartition de la charge sinistre arrêt de travail



Graphique 3.2.3.5 (b) – Densité de la charge sinistre arrêt de travail

Enfin, l'analyse croisée entre durée sinistre et charge sinistre de la garantie arrêt de travail est plus dispersée que celle de la garantie accident de travail. La répartition autour de la diagonale est moins nette.

Cette dispersion s'explique par la présence des différents risques couverts par la garantie d'arrêt de travail (MAL, DO, LD, LM). Nous observons toutefois une concentration de la charge sinistre en dessous du seuil de 75 000,00 euros de charge sinistre et 75 mois de durée de sinistre.

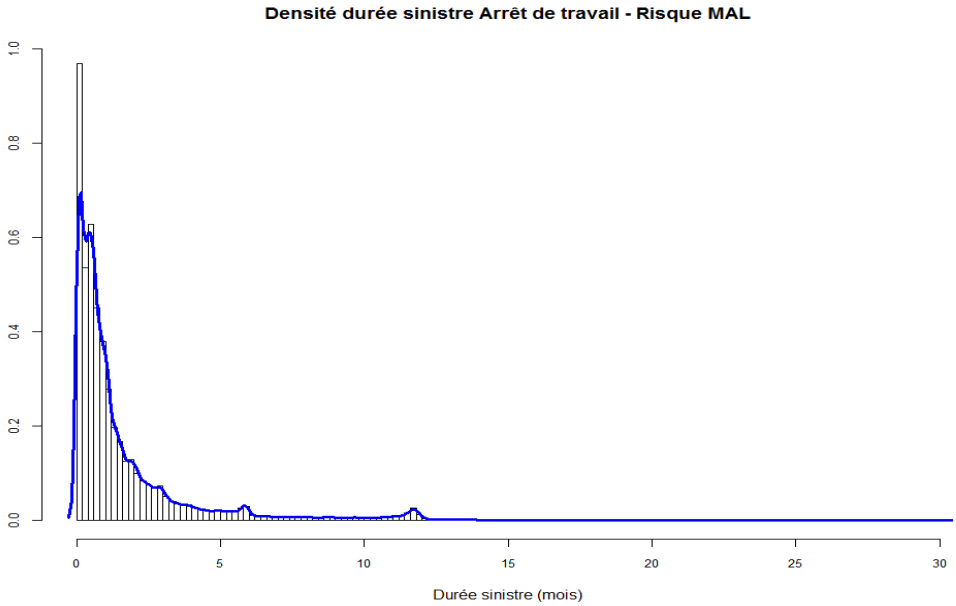


Graphique 3.2.3.6– Répartition de la charge sinistre en fonction de la durée de sinistre arrêt de travail

Partant du constat précédent sur l'éventuelle hétérogénéité des risques assurés pour la garantie arrêt de travail, une analyse par risque semble être nécessaire pour mieux nous éclairer. Nous distinguons donc les sinistrés au titre de la maladie ordinaire (MAL) des autres (DO, LD, LM). Ce regroupement est motivé par une distinction des sinistres de durée courte des sinistres de durée plus longue.

- ✓ **Analyse du risque maladie ordinaire (MAL) de la garantie arrêt de travail**

La distribution de la durée des sinistres (clos et censurés) est toujours globalement décroissante avec deux pics de sorties à moins d'un mois, puis à 6 mois et 12 mois. Les deux pics à moins d'un mois révèlent la présence de sinistre de très courte durée de quelques jours.



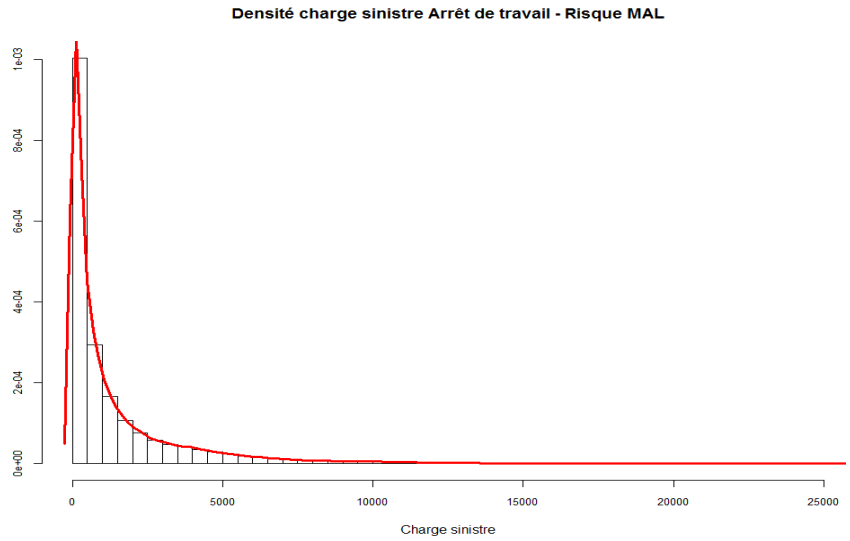
Graphique 3.2.3.6 – Densité des durées de sinistres d'arrêt de travail du risque Maladie Ordinaire

L'analyse des indicateurs de dispersion permet de noter que les durées de sinistres moyennes du risque maladie ordinaire (MAL) sont plus faibles que celles de la garantie arrêt de travail tous les risques confondus (par exemple pour les sinistres clos, la durée de sinistre moyenne est de 1,71 mois pour le risque MAL contre 3,34 mois pour l'ensemble des risques de la garantie arrêt de travail). La valeur maximale est de 42,13 pour les sinistres clos (contre 160,53 mois pour l'ensemble des risques de la garantie arrêt de travail).

Durée sinistre (mois) - Risque MAL								
Garantie	Etat	Effectif	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Accident de travail	Sinistre clos	199 979	0,00	0,33	0,77	1,71	1,77	42,13
	Sinistre ouvert	13 371	0,00	0,37	1,00	2,19	2,87	41,30

Tableau 3.2.3.10 – Indicateurs de dispersion des durées de sinistre arrêt de travail du risque Maladie Ordinaire

L'effectif des sinistres clos est 15 fois supérieur à celui des sinistres censurés (contre 11,57 fois pour l'ensemble des risques de la garantie arrêt de travail). Nous notons toujours que la distribution de la charge sinistre conserve l'allure décroissante.



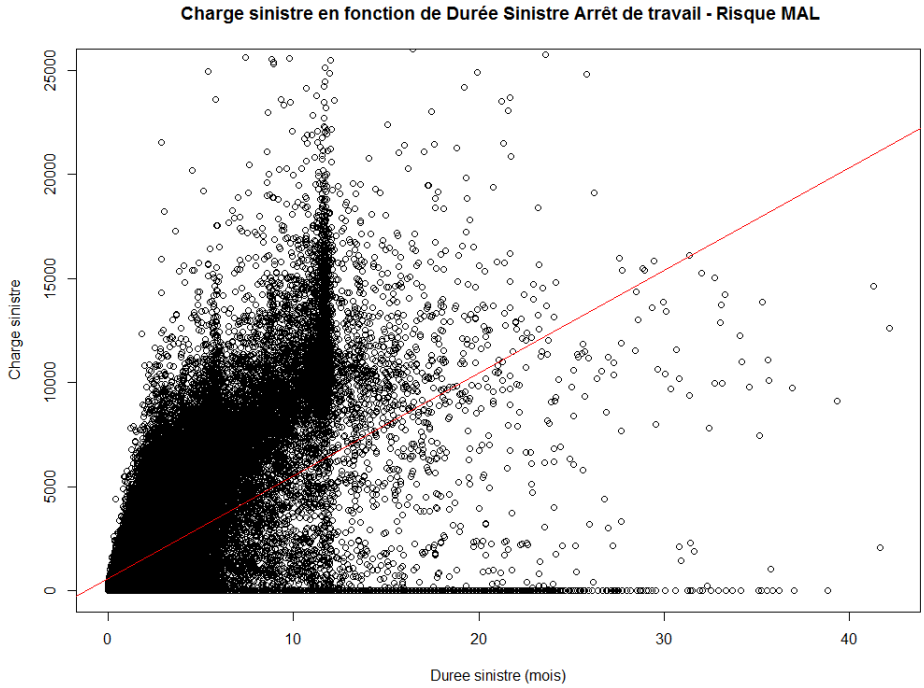
Graphique 3.2.3.7 – Densité de la charge sinistre arrêt de travail du risque Maladie Ordinaire

De manière analogue, les indicateurs de dispersion de la charge sinistre donnent des valeurs plus faibles (par exemple pour les sinistres clos, la charge sinistre moyenne est de 1 360,10 euros pour le risque MAL contre 2 963,50 euros pour l'ensemble des risques). La valeur maximale est de 96 873,00 euros pour les sinistres clos (contre 238 584,60 euros pour l'ensemble des risques de la garantie arrêt de travail).

Charge sinistre - Risque MAL								
Garantie	Etat	Effectif	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Accident de travail	Sinistre clos	199 979	-	115,40	475,40	1 360,10	1 565,60	96 873,00
	Sinistre ouvert	13 371	-	209,90	911,10	2 261,40	3 351,50	39 095,00

Tableau 3.2.3.11 – Indicateurs de dispersion de la charge sinistre arrêt de travail du risque Maladie Ordinaire

Enfin, l'analyse croisée entre durée sinistre et charge sinistre du risque maladie ordinaire permet d'observer une forte concentration en dessous du seuil de 15 000 euros de charge sinistre (contre 75 000 euros pour l'ensemble) et 12 mois de durée de sinistre (contre 75 mois pour l'ensemble des risques de la garantie arrêt de travail).



Graphique 3.2.3.8 – Répartition de la charge sinistre en fonction de la durée de sinistre arrêt de travail du risque Maladie Ordinaire

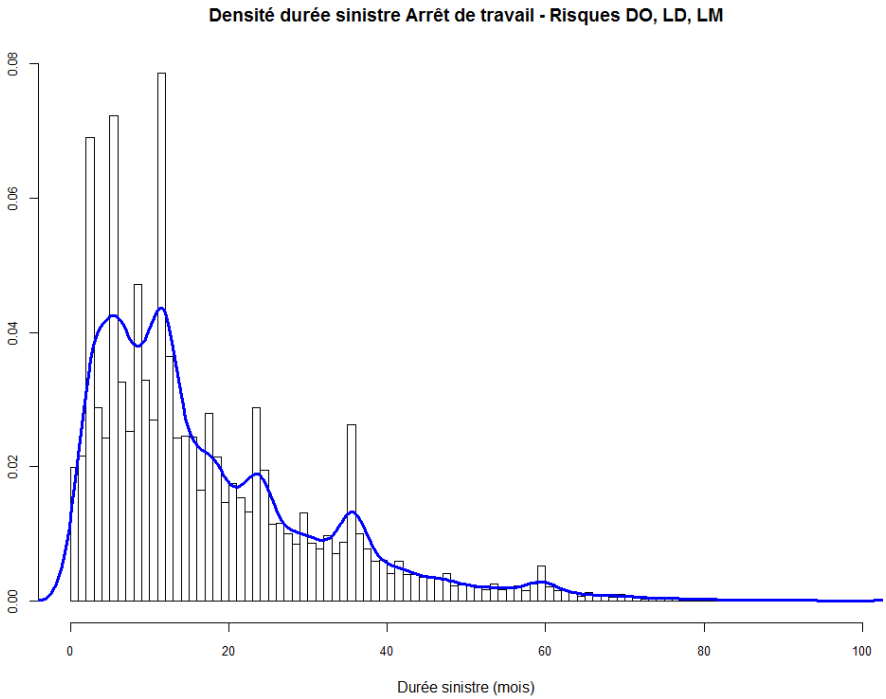
Le test statistique de Spearman conduit, avec un seuil de confiance fixé à 95%, à accepter l'hypothèse alternative d'existence d'un lien entre les variables « durée du sinistre » et « charge sinistre » (Rho est égal à 69,65% donc non nul) pour le risque maladie ordinaire.

Test Rho de Spearman : existe-il un lien entre durée du sinistre et charge sinistre?	
Risques	Risque Maladie Ordinaire
S	4.9119e+14
Rho	0.6965246
Conclusion	OUI

Tableau 3.2.3.12 – résultats du test statistique de Rho de Spearman

✓ **Analyse des autres risques de la garantie arrêt de travail autres risques (DO, LM et LD)**

La distribution de la durée des sinistres (clos et censurés) est plus erratique avec plusieurs pics de sorties et une durée beaucoup plus longue comme attendu.



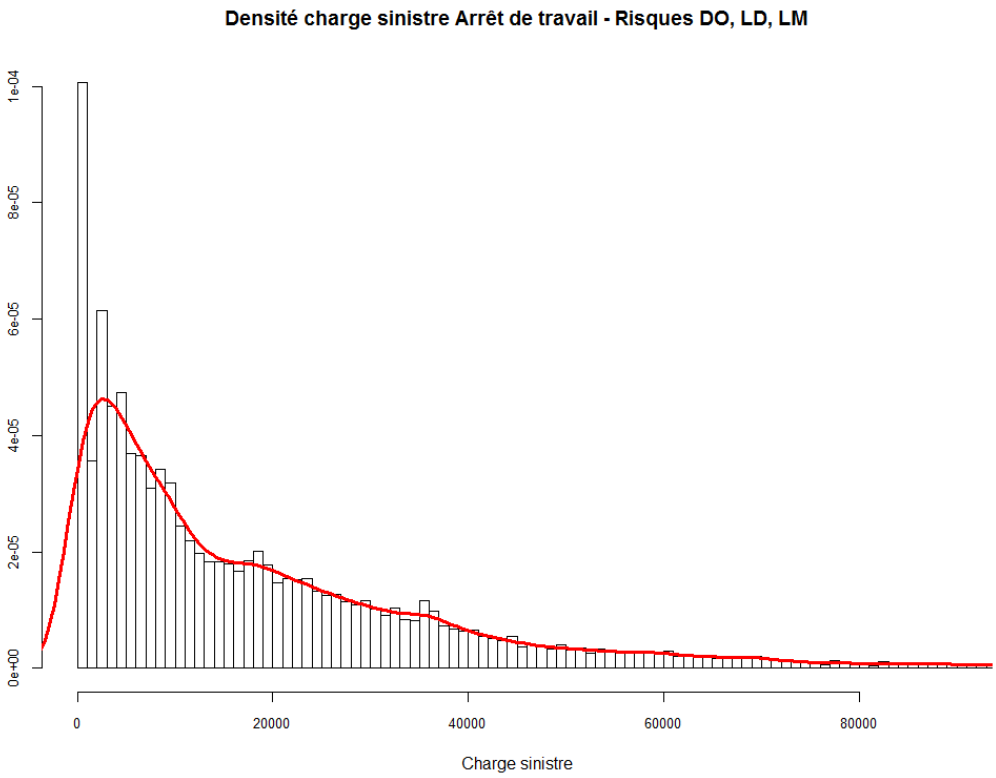
Graphique 3.2.3.9 – Densité des durées de sinistres d'arrêt de travail autres risques (DO, LM, LD)

L'analyse des indicateurs de dispersion permet de noter que les durées de sinistres moyennes des autres risques (DO, LD et LM) sont plus élevées que celles de la garantie arrêt de travail tous les risques confondus (par exemple pour les sinistres clos, la durée de sinistre moyenne est de 23,67 mois contre 1,71 mois pour le risque MAL et 3,34 mois pour l'ensemble des risques de la garantie arrêt de travail). La valeur maximale correspond aux 160,53 mois pour l'ensemble des risques de la garantie arrêt de travail pour les sinistres clos (contre 42,13 mois pour le risque MAL).

Durée sinistre (mois) - Risques DO, LM, LD								
Garantie	Etat	Effectif	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Arrêt de travail	Sinistre clos	23 903	-	6,00	11,97	16,97	23,67	160,53
	Sinistre ouvert	5 977	0,00	9,13	17,37	21,48	30,00	165,40

Tableau 3.2.3.13 – Indicateurs de dispersion des durées de sinistre arrêt de travail autres risques (DO, LM, LD)

L'effectif des sinistres clos est 4 fois supérieur à celui des sinistres censurés (contre 15 fois pour le risque MAL et 11,57 fois pour l'ensemble des risques de la garantie arrêt de travail). Nous notons toujours que la distribution de la charge sinistre conserve l'allure globale décroissante, avec la présence de plusieurs pics en adéquation avec les observations faites précédemment sur les durées de sinistre.



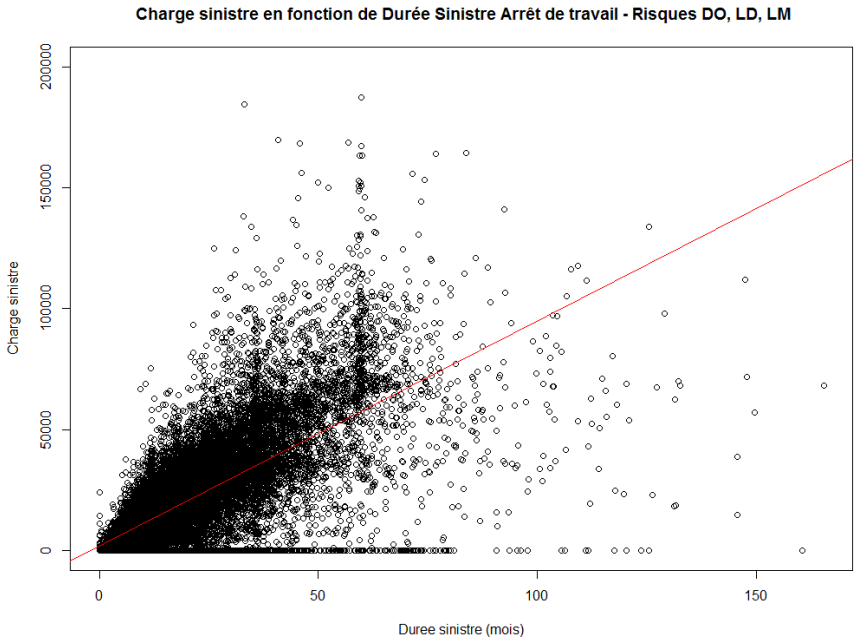
Graphique 3.2.3.10 – Densité de la charge sinistre arrêt de travail autres risques (DO, LM, LD)

De manière analogue, les indicateurs de dispersion de la charge sinistre donnent des valeurs plus importantes (par exemple pour les sinistres clos, la charge sinistre moyenne est de 16 378,00 euros (contre 1 360,10 euros pour le risque MAL et 2 963,50 euros pour l'ensemble des risques). La valeur maximale correspond aux 238 584,60 euros pour l'ensemble des risques de la garantie arrêt de travail pour les sinistres clos (contre 96 873,00 euros pour le risque MAL).

Charge sinistre - Risques DO, LM, LD								
Garantie	Etat	Effectif	Minimum	1er Quartile	Médiane	Moyenne	3ème Quartile	Maximum
Arrêt de travail	Sinistre clos	23 903	-	3 455,00	9 653,00	16 378,00	23 012,00	238 585,00
	Sinistre ouvert	5 977	-	9 429,00	23 114,00	27 729,00	37 972,00	184 427,00

Tableau 3.2.3.14 – Indicateurs de dispersion de la charge sinistre arrêt de travail autres risques (DO, LM, LD)

Enfin, l'analyse croisée entre durée sinistre et charge sinistre des autres risques (DO, LM et LD) permet d'observer une forte concentration conforme à ce que nous avons remarqué pour l'ensemble des risques de la garantie arrêt de travail. En dessous du seuil 75 000 euros de la charge sinistre (contre de 15 000 euros pour le risque maladie ordinaire) et 75 mois de durée de sinistre (contre 12 mois pour le risque maladie ordinaire).



Graphique 3.2.3.11 – Répartition de la charge sinistre en fonction de la durée de sinistre arrêt de travail autres risques (DO, LM, LD)

Le test statistique de Spearman conduit, avec un seuil de confiance fixé à 95%, à accepter l'hypothèse alternative d'existence d'un lien entre les variables « durée du sinistre » et « charge sinistre » (Rho est égal à 67,95% donc non nul) pour le risque maladie ordinaire.

Test Rho de Spearman : existe-il un lien entre durée du sinistre et charge sinistre?	
Risques	Autres risques (DO, LD, LM)
S	1.4261e+12
Rho	0.6795801
Conclusion	OUI

Tableau 3.2.3.15 – résultats du test statistique de Rho de Spearman

En conclusion, la distinction entre le risque maladie ordinaire (MAL) et les autres risques (DO, LM et LD) de la garantie arrêt de travail semble pertinente.

3.3. Découpage des bases et ajustement des hyper paramètres

Un choix judicieux de protocole est nécessaire pour éviter le piège lié au surapprentissage ou sous-apprentissage des modèles lors de la phase d'ajustement aux données. Nous sommes dans le cas où la volumétrie est suffisamment importante pour réaliser les différentes phases de l'apprentissage automatique (apprentissage, validation et test). L'erreur de prédiction entre la phase de validation et la phase de test doit être globalement de même grandeur pour s'assurer de la stabilité des modèles. Finalement, la performance des modèles lors des validations et des tests sont mesurées par rapport aux différentes métriques présentées ci-dessus.

3.3.1. Découpage de la base de données

De manière pratique, nous avons procédé en deux étapes. La première étape est une phase d'apprentissage et de validation afin de choisir les modèles les plus performants au sens des métriques retenues. Le choix du modèle de prédiction finale dépendra des métriques retenues. La seconde étape est une phase de test pour mesurer la qualité des prédictions. Les réglages des paramètres des modèles prédictifs lors de la phase d'apprentissage et validation se font par itération sur les différentes dates d'arrêtés allant de 2012 à 2015. Nous choisissons à chaque étape, 70% des individus de la base comme échantillon d'apprentissage et 30% comme échantillon de validation. Une fois le meilleur modèle déterminé, l'étape de test consiste à l'utiliser pour prédire la durée de maintien et la charge sinistre (sinistres clos et des sinistres non-clos) de la date d'arrêté 2016.

3.3.2. Calibrage du tree base censored

Nous construisons les arbres de régression avec le package ***rpart*** dans R. Cet algorithme utilise l'indice de Gini comme critère de mesure de la pureté des feuilles. A chaque nœud le *split* est construit de manière à maximiser le gain d'information apporté par une des variables explicatives d'importance et de permettre la création de feuilles aussi homogènes que possible. Enfin, l'algorithme CART utilise une procédure de « *surrogate splits* », ou variables-substituts, pour s'affranchir du problème des valeurs manquantes. Cette procédure permet notamment de classer de nouveaux individus inconnus, même lorsqu'ils possèdent des valeurs manquantes pour les variables apparaissant sur l'arbre de décision.

Nous fixons comme paramètres de réglage du modèle : le nombre minimal d'individus à chaque nœud terminal (via le paramètre « *minsplit* » de *rpart*, fixé à 30 individus au minimum), la valeur de complexité de l'arbre pour l'élagage de l'arbre correspondant à la valeur minimale des erreurs de prédiction commises en validation croisée en fonction de la taille de l'arbre (via « *cptable »* de *rpart* , fournissant l'erreur de prédiction « *X-val Relative Error* » en fonction de la taille de l'arbre « *size of tree »*), les données d'apprentissage et de validation (fixés à 70% vs 30% des bases aux années d'arrêté entre 2012 et 2015) et les poids IPWC de Kaplan-Meier pour tenir compte des données censurées (calculés précédemment, via le paramètre «*weights*» de *rpart*).

Ce modèle servira de référence de comparaison avec les modèles provenant des adaptations de *Random forest* et *Gradient tree Boosting*.

3.3.3. Calibrage du random forest censored

Nous construisons un ensemble fini d'arbres de régression avec le package ***randomForest*** dans R. Cet algorithme permet d'optimiser les paramètres du modèle précédent « *modèle tree base censored* » pendant la phase de validation, afin de minimiser les erreurs de prédiction.

Le *tunage* de Random Forest dépend fortement des données. Après plusieurs itérations sous contrainte de temps de traitement et d'évitement du surapprentissage, nous fixons comme paramètres de réglage du modèle : le nombre d'arbre de régressions construit par l'algorithme (via le paramètre « *ntree* », fixé à 500 après une étape de test de plusieurs valeurs entre 100 et 2000), le nombre de variables explicatives testées à chaque divisions (via le paramètre « *mtry* », fixé à la valeur entière 6 correspondant à la partie entière du tiers du nombre de variables explicatives utilisées), la taille minimale des nœuds terminaux (via le paramètre « *nodesize* », fixé à 5 après une étape de test de plusieurs valeurs entre 1 et 30), le nombre maximal de nœuds terminaux (via le paramètre « *maxnodes* », fixé à 25 après une étape de test de plusieurs variables entre 5 et 50), les données d'apprentissage et de validation (70% vs 30% des bases aux années d'arrêté entre 2012 et 2015) et les poids IPWC de Kaplan-

Meier pour tenir compte des données censurées (via le paramètre «*weights*», calculés précédemment).

3.3.4. Calibrage du gradient tree boosting censored

Nous utilisons le package ***gmb*** dans R (*Generalized Boosted Regression Models*). Cet algorithme est une alternative très efficace au *Random Forest* précédent.

La question du paramétrage est particulièrement délicate dans le cadre du *gradient boosting*. En effet, ils sont nombreux, et leur influence est considérable. Malheureusement, même si l'on a à peu près une idée des pistes à explorer pour améliorer la qualité des modèles, l'identification avec exactitude les paramètres à manipuler et fixer les bonnes valeurs est souvent très empirique, surtout lorsqu'ils interagissent entre eux. Ici, plus que pour d'autres méthodes de machine learning, la stratégie essai-erreur prend beaucoup d'importance.

Le tunage de cet algorithme est toujours dépendant des données disponibles. Nous spécifions quand même la distribution gaussienne (« *gaussien* »), qui correspond à la fonction de coût (« déviance multinomiale ») pour que la procédure interprète correctement notre variable cible continue. Après plusieurs itérations sous contrainte de temps de traitement et d'évitement du surapprentissage, nous fixons comme paramètres de réglage du modèle : le nombre d'arbre de régressions construit par l'algorithme (via le paramètre « *n.trees* », fixé à 5000 après une étape de test de plusieurs valeurs entre 1000 et 10000), le nombre minimal d'individus à chaque nœud terminal (via le paramètre « *minobsinnode* », fixé à 30 individus au minimum), la vitesse de convergence gère le surapprentissage et elle est contrôlée par les corrections apportées par chaque itération si on dépasse le seuil d'un taux d'erreur donné (via le paramètre « *shrinkage* », fixé à 1% après une étape de test de plusieurs valeurs entre 0,01% et 1%), la valeur de la profondeur maximale de chaque arbre de régression (via le paramètre « *interaction.depth* », fixée à 3 après une étape de test entre 2 et 5), le nombre de cross-validation à réaliser pour l'optimisation des paramètres (via le paramètre « *cv.folds* », fixé à 3 après une étape de test entre 3 et 5), les données d'apprentissage et de validation (70% vs 30% des bases aux années d'arrêté entre 2012 et 2015) et les poids IPWC de Kaplan-Meier pour tenir compte des données censurées (via le paramètre «*weights*», calculés précédemment).

4. Applications et analyses

Nous illustrons l'application des méthodes d'apprentissage automatique adaptés aux données censurées pour la prédiction du risque au problème de la prédiction de durée de maintien et de charge de prestations de contrats d'assurance de prêts et de garanties au titre de la prévoyance collective. Les données utilisées proviennent de l'entrepôt de données d'un assureur leader du marché français en assurance de personnes. Ce système de gestion contient des données individuelles tête par tête sur la gestion des sinistres, y compris les données identitaires et sociodémographiques, les données transactionnelles et les données sur la gestion des garanties.

4.1. Résultats de l'étude des contrats d'assurance de prêts

L'étude est réalisée sur le portefeuille de prêt. Nous présentons dans cette section les différents éléments d'appréciation de la qualité des prédictions des risques suivant les trois algorithmes : les graphes de comparaison entre les valeurs prédites contre les valeurs réelles, les comparaisons des

durées moyennes de maintien prédites vs durées moyennes de maintien dite « réelle » et les métriques MSE et MSEW. Pour la prédiction des charges sinistre des prestations, nous regarderons la MSEW.

4.1.1. Prédiction de la durée de sinistre des garanties d'arrêt de travail et de perte d'emploi

Quelque soit la méthode utilisée, l'analyse des variables d'importance permet de connaître les variables qui jouent un rôle prépondérant lors de la classification ou la régression. Il aparaît que le montant du sinistre soit la première variable la plus explicative des risques d'arrêt de travail et de perte d'emploi pour les contrats de prêts. Les résultats fournis par les différents modèles d'apprentissage sont assez cohérents et classent parmi les trois variables les plus explicatives : le montant sinistre, le montant des échéances mensuelles, la durée du sinistre. D'autres variables comme le capital intitial, le code clôture (pour le Gradient Boosting et le Random forest modifiés), ainsi que le capital restant dû (pour le CART modifié) ressortent comme ayant du signal mais classées à des positions différentes suivant le modèle.

	Variables d'importance		
Position	CART modifié	Gradient Boosting modifié	Random forest modifié
1ère	Montant sinistre	Montant sinistre	Montant sinistre
2ème	Montant échéances mensuelles	Durée sinistre	Durée sinistre
3ème	Durée sinistre	Montant échéances mensuelles	Montant échéances mensuelles
4ème	Capital initial	Capital initial	Code Clôture (ouvert ou clos)
5ème	Capital restant dû	Code Clôture (ouvert ou clos)	Capital initial

Tableau 4.1.1.1 – Variables d'importance suivant les trois algorithmes

Pour la prédiction des durées de sinistre, nous avons retenu les métriques MSE et MSEW pour tester la performance des modèles. Les différents algorithmes ont été testés sur les différentes dates d'arrêté (de 2012 à 2015). Les meilleurs modèles, quel que soit l'algorithme utilisé, correspondent aux modèles calibrés avec la base d'arrêté 2015, car ils possèdent les plus faibles valeurs des métriques de performance MSE et MSEW. Ces résultats sont cohérents puisque la base d'arrêté 2015 comprend l'historique des sinistres ouverts et clos entre 2004 et 2015. Les analyses précédentes nous ont permis également de nous assurer de la complétude de cette base.

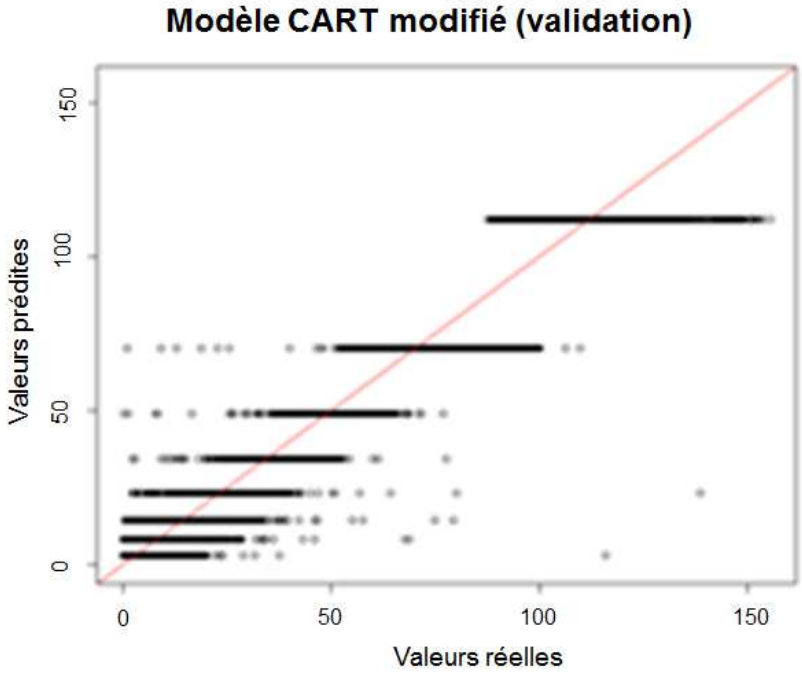
	MSE								
Année d'arrêté	MSE_CART_all	MSE_GB_all	MSE_RF_all	MSE_CART_clos	MSE_GB_clos	MSE_RF_clos	MSE_CART_non_clos	MSE_GB_non_clos	MSE_RF_non_clos
2012	62,8	34,5	33,5	14,5	1,7	0,5	362,7	238,1	217,1
2013	42,7	21,2	20,3	11,7	1,2	0,5	252,1	155,7	140,2
2014	28,7	10,5	10,5	8,6	0,1	0,3	175,4	86,8	78,6
2015	16,4	3,7	4,0	7,0	0,0	0,5	93,0	33,3	30,3

Tableau 4.1.1.2 (a) – Métriques de performance MSE

	MSWE								
Année d'arrêté	MWSE_CART_all	MWSE_GB_all	MWSE_RF_all	MWSE_CT_clos	MWSE_GB_clos	MWSE_RF_clos	MWSE_CART_non_clos	MWSE_GB_non_clos	MWSE_RF_non_clos
2012	70,3	43,9	35,2	11,2	1,0	0,3	59,2	42,8	34,9
2013	46,0	26,4	20,6	8,6	1,2	0,3	37,4	25,1	20,3
2014	29,1	12,2	10,2	6,5	0,1	0,2	22,6	12,1	10,0
2015	15,7	3,9	3,7	5,6	0,0	0,4	10,1	3,9	3,3

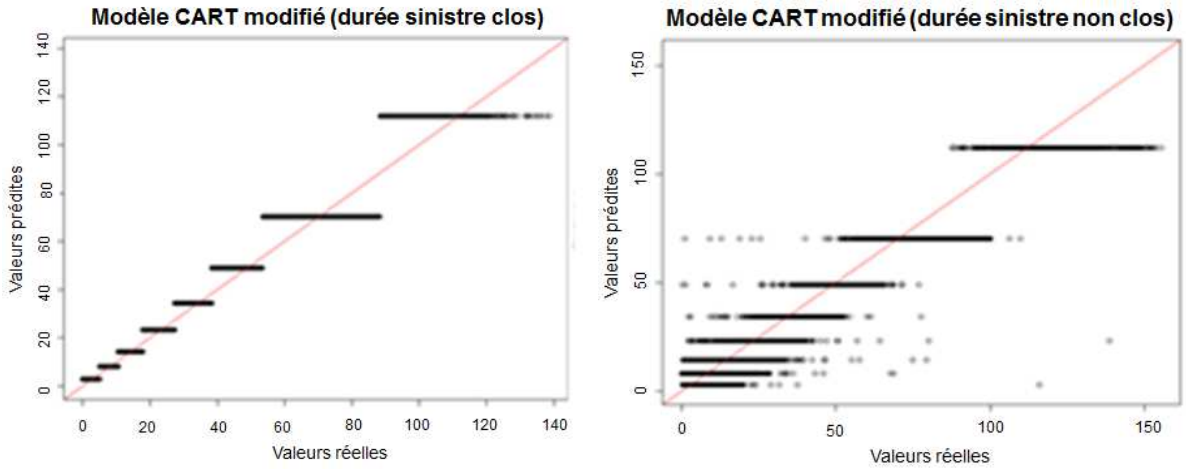
Tableau 4.1.1.2 (b) – Métriques de performance MSWE

Les résultats des durées de sinistre prédites par le modèle calibré avec l'algorithme du CART modifié (*Tree-base censored regression*) ne sont pas très satisfaisants pour des calculs de provisions.



Graphique 4.1.1.1 (a) – Prédiction des durées du sinistre avec le modèle CART modifié (*Tree-base censored regression*)

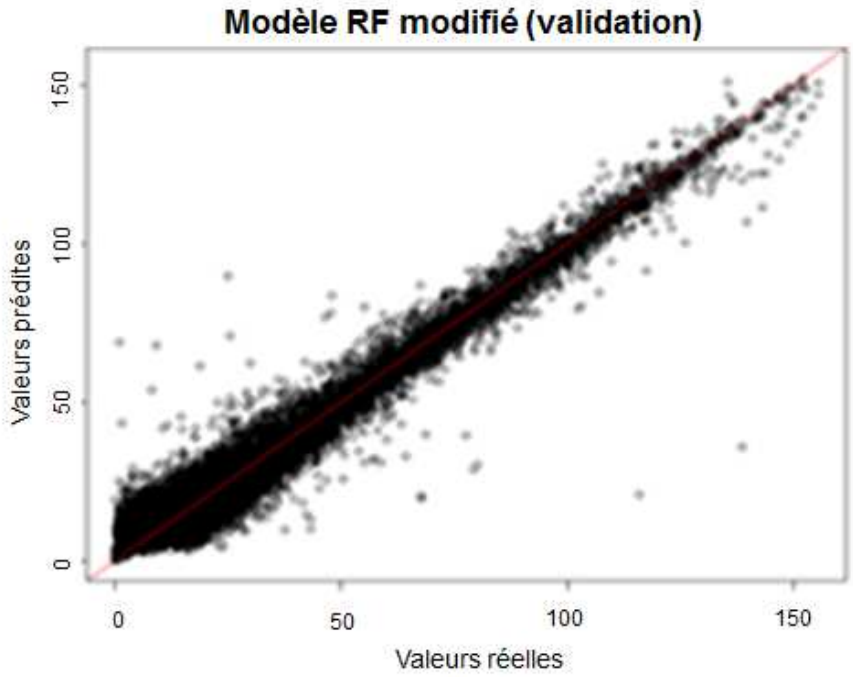
Ce modèle fournit des résultats sous forme de paliers avec des dispersions importantes par rapport à la diagonale. Ce phénomène correspond à une approximation peu précise du risque. Le modèle n'arrive donc pas à bien prédire les valeurs atypiques et à tendance à sous-estimer le risque (valeurs en dessous de la diagonale).



Graphique 4.1.1.1 (b) – Prédiction des durées des sinistres clos et censurés avec le modèle CART modifié

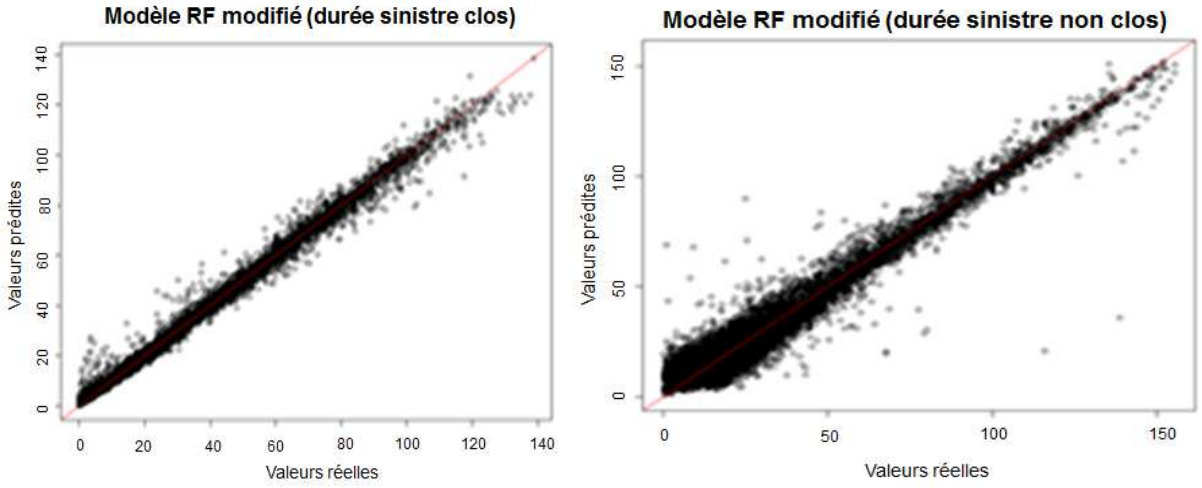
La prédiction des sinistres censurés nécessaires au provisionnement donne des résultats trop dispersés et dans le sens d'une sous-estimation du risque.

Le second modèle proposé est celui obtenu avec l'algorithme du *Random forest* modifié (*Random Forest Censored*). Les résultats de prédiction sont meilleurs par rapport au CART modifié.



Graphique 4.1.1.2 (a) – Prédiction des durées du sinistre avec le modèle *Random forest* modifié (*Random Forest Censored*)

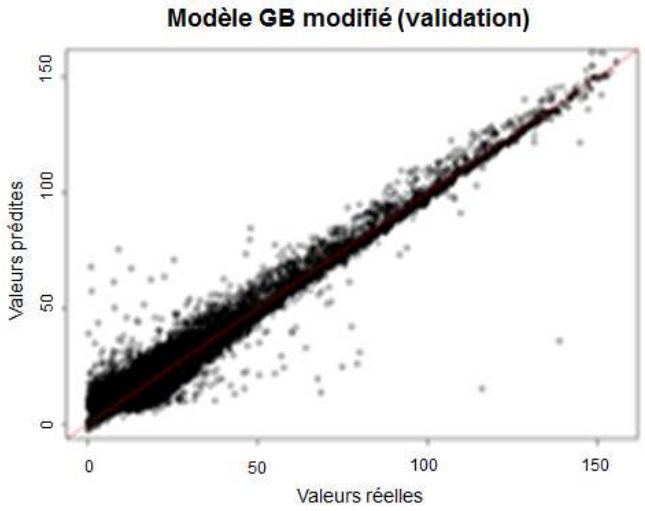
Le modèle calibré fournit des prédictions plus concentrées autour de la diagonale. La majorité des dispersions surtout pour les durées de sinistre élevées. Ce phénomène correspond à une approximation peu précise du risque sur les valeurs atypiques et à tendance à sous-estimer le risque (valeurs en dessous de la diagonale).



Graphique 4.1.1.2 (b) – Prédiction des durées des sinistres clos et censurés avec le modèle *Random forest* modifié

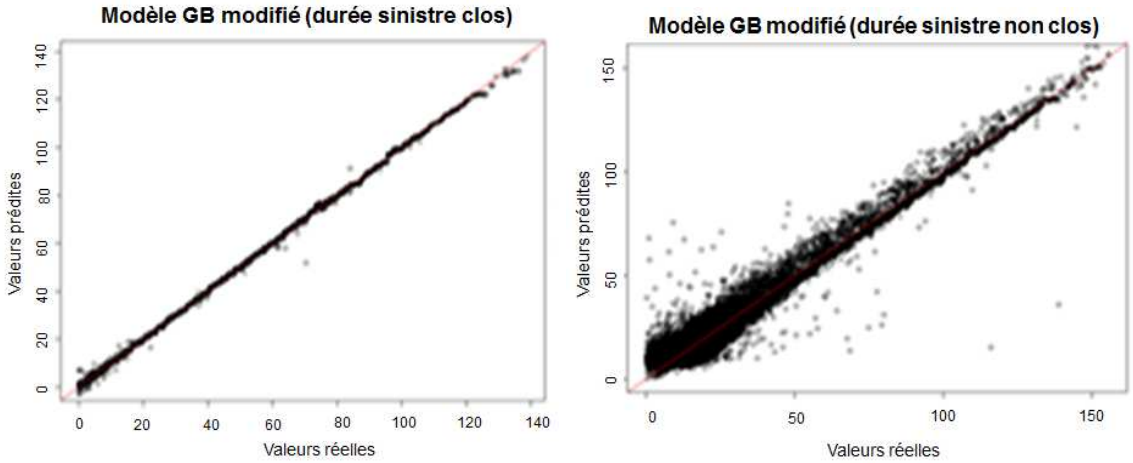
Nous constatons une bonne adéquation entre les durées de sinistre prédites et réelles. Pour les prédictions des sinistres censurés, la quantité de cas mal estimés par le modèle tendant à une sous-estimation du risque semble être plus élevée que les cas de sur-estimation.

Le troisième modèle utilisé est celui calibré avec l'algorithme du *Gradient Boosting* modifié (*Gradient Tree Censored Boosting)*). L'algorithme du *Gradient Tree censored Boosting* donne de meilleurs résultats que les deux autres algorithmes, car la distribution est plus homogène autour de la diagonale.



Graphique 4.1.1.3 (a) – Prédiction des durées du sinistre avec le modèle *Gradient Boosting* modifié (*Gradient Tree Censored Boosting*)

En effet, par rapport au modèle du *Random forest* modifié, les prédictions du risque vont plutôt dans le sens de la prudence, car globalement légèrement au-dessus de la diagonale.

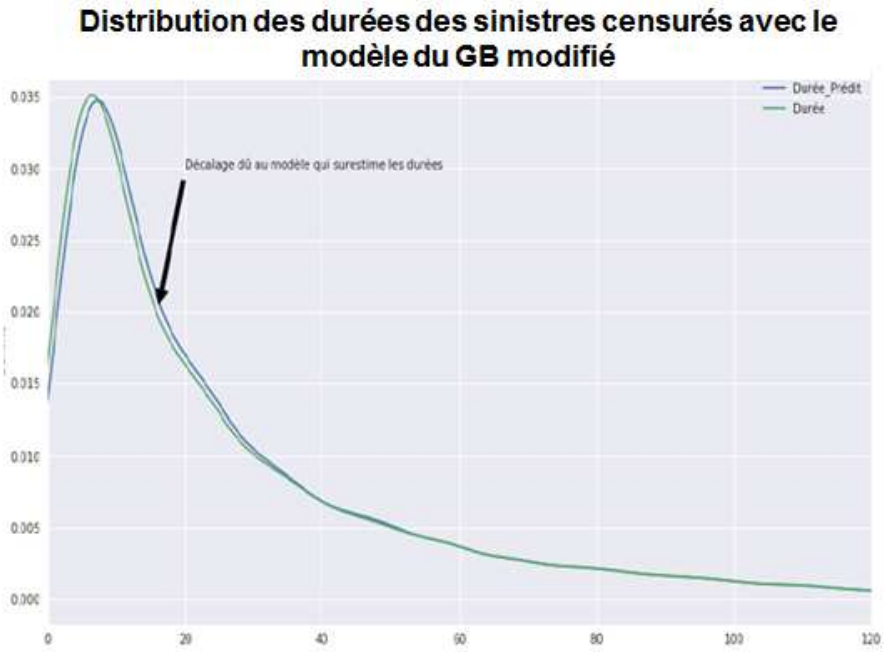


Graphique 4.1.1.2 (b) – Prédiction des durées des sinistres clos et censurés avec le modèle *Gradient Boosting* modifié

La comparaison entre les durées de sinistre clos prédites et réelles montre un alignement presque parfait sur la diagonale. Ce qui signifie la très bonne qualité des prédictions de ce modèle pour les sinistres clos. Concernant la durée prédite pour les sinistres censurés (sinistres non clos), nous remarquons également une concentration autour de la diagonale. Il y a donc une très bonne qualité prédictive du modèle des durées de sinistre par individu. Nous remarquons néanmoins quelques erreurs de classifications.

Il s'agit d'un petit nombre restreint de points qui se détachent de cette diagonale et dont les valeurs réelles sont largement supérieures aux valeurs prédites. Toutefois, globalement, ce modèle donne des prédictions allant dans le sens de la prudence.

Enfin, la comparaison de la distribution des densités des durées prédites et réelles des sinistres censurés (sinistres ouverts), montre un léger décalage entre les deux courbes dans le sens de la prudence. Le modèle issu de l'apprentissage avec l'algorithme *Gradient Tree censored Boosting* aura tendance à surestimer légèrement les durées des sinistres non clos, ce qui conduira donc à des niveaux de provisions prudentes.

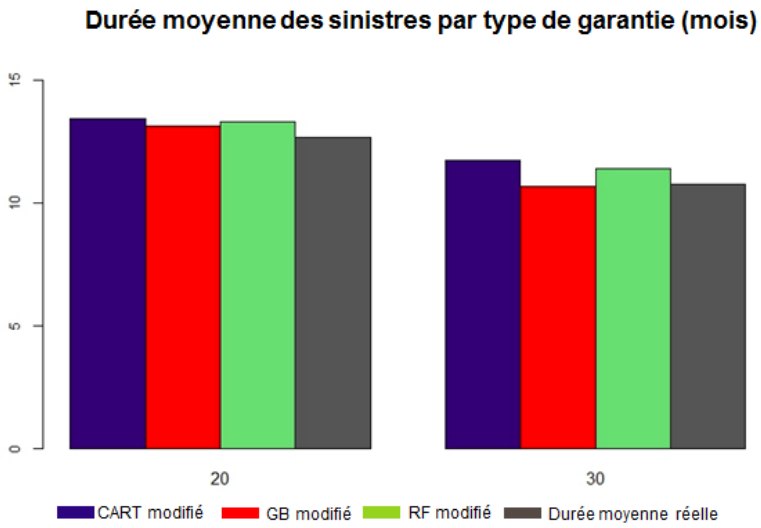


Graphique 4.1.1.3 – Comparaison de la densité des durées réelles des sinistres non clos avec la densité des durées prédites (*Gradient Boosting Censored*)

Afin de valider les résultats précédents, nous allons regarder les prédictions réalisées par les trois modèles de manière plus fine suivant axes d'analyses métier. Nous regarderons plus précisément la qualité des prédictions par type de garantie (20 : Arrêt de travail et 30 : chômage), par réseau de distribution (ou partenaire), par tranche d'âge, et par code clôture (statut à 0 pour les sinistres censurés et à 1 pour les sinistres clos).

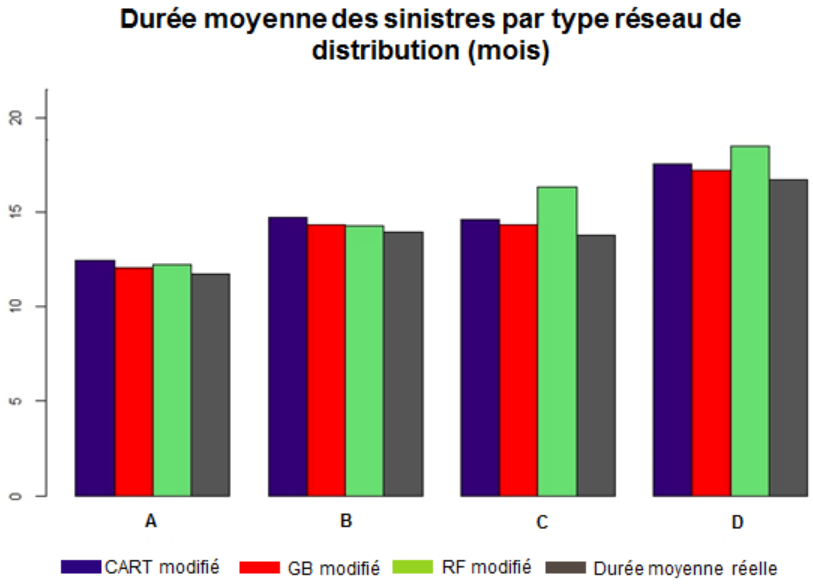
La base de validation retenue dans ce qui suit est la base d'arrêté 2016. Cette base est intéressante car elle n'a pas servi pour le calibrage des modèles lors de l'apprentissage et de la validation. Nous retenons comme base de référence les durées réelles (sinistres clos et sinistres censurés). Nous souhaitons comparer nos modèles à celui calibré avec la méthode de Lopez et al., c'est-à-dire le CART modifié.

La durée moyenne de sinistre par type de garantie prédite par le modèle GB modifié donne la valeur la plus proche de la valeur réelle de ce paramètre.



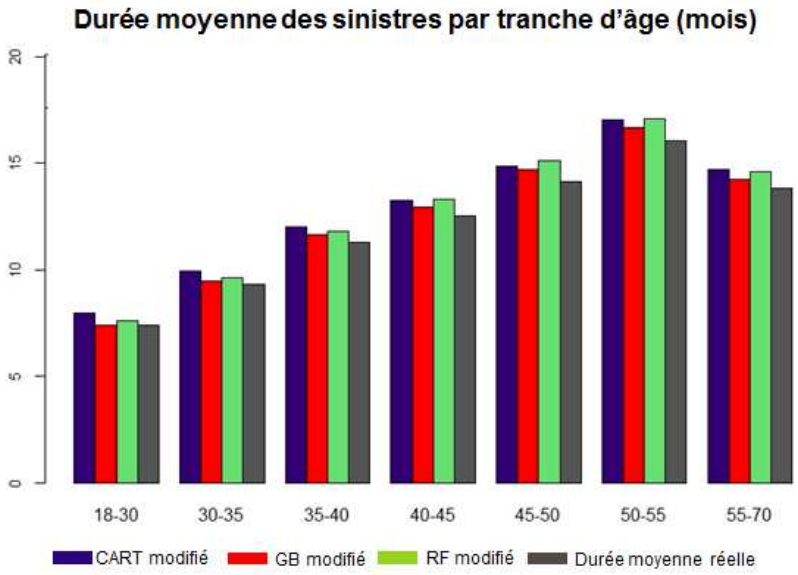
Graphique 4.1.1.4 – Comparaison des durées moyennes de sinistre par type de garantie

Les durées moyennes de sinistre prédites pour les réseaux de distribution représentant les principaux partenaires avec le modèle GB modifié donne encore les valeurs les plus proches des valeurs réelles de ce paramètre. Une différence presque constante entre les prédictions du modèle CART modifié et GB modifié est maintenue quelque soit le partenaire. Par contre si l'écart est quasi nul entre les prédictions de RF modifié et GB modifié pour les partenaires A et B, il se creuse pour les partenaires C et B.



Graphique 4.1.1.5 – Comparaison des durées moyennes de sinistre par réseau de distribution

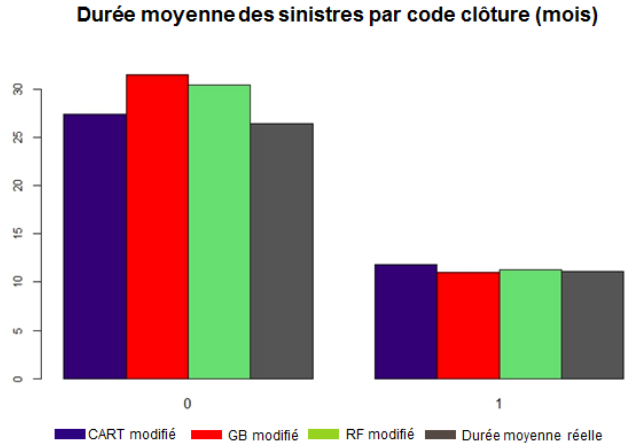
Les durées moyennes de sinistre prédites par tranche d'âge avec le modèle GB modifié donne encore les valeurs les plus proches des valeurs réelles.



Graphique 4.1.1.6 – Comparaison des durées moyennes de sinistre par tranche d'âge

La prédiction des durées moyennes de sinistre par code clôture (sinistre clos et sinistre censuré) est un cas intéressant, car c'est un indicateur de suivi de risque KRI (Key Risk Indicator) pertinent pour le métier. En effet, dans notre cas, nous observons que la durée moyenne des sinistres censurées est sous-estimée par rapport aux prédiction du modèle GB modifié. Ce constat est cohérent avec le fait que la majorité des sinistres censurés sont des sinistres récents. Ainsi, l'estimation de la durée moyenne des sinistres censurés nécessite une modélisation particulière de type modèles de durée.

Le fait que la prédiction de la durée moyenne des sinistres censurés par le modèle CART modifié soit proche du calcul direct, indique et confirme que ce modèle à tendance à sous-estimer le risque. Notons également que ce résultat nous paraît raisonnable, car nous avons signalé dans la partie précédente consacrée aux statistiques descriptives que l'effectif des sinistres ouverts en 2016 semblait être trop faible par rapport aux années précédentes, ce qui correspondrait aux effets croisés des retards de traitements de gestion et de déclarations des sinistres.



Graphique 4.1.1.7 – Comparaison des durées moyennes de sinistre par code clôture

En conclusion, nous observons que globalement l'algorithme *Gradient tree censored Boosting* fournit de meilleures prévisions que l'algorithme *Tree-Base Censored* et l'agorithme *Random Forest censored*.

4.1.2. *Prédiction de la charge sinistre des garanties d'arrêt de travail et de perte d'emploi*

Comme vu dans la partie théorique, le choix des métriques de performance doit être adapté aux données et à la problématique à résoudre. Il s'agit notamment de considérer l'échelle de grandeur et le type de variable (qualitative ou quantitative) que l'on cherche à prédire. La variable charge sinistre est une variable quantitative réelle pouvant prendre de très grandes valeurs. En effet, il s'agit de remboursement de tout ou partie de montants empruntés dans le cadre notamment de prêts immobiliers, professionnels ou à la consommation. Dans ce cas, nous utiliserons les métriques suivantes : RMSLEW, RSEW, R2W et le C-Index. Le meilleur modèle correspond à celui qui donne les valeurs les plus faibles pour le RMSLEW et le RSEW d'une part, et les valeurs les plus élevées pour le R2W et le C-index d'autre part.

	RMWSLE								
Année d'arrêté	RMWSLE_CART_all	RMWSLE_GB_all	RMWSLE_RF_all	RMWSLE_CART_clos	RMWSLE_GB_clos	RMWSLE_RF_clos	RMWSLE_CART_non_clos	RMWSLE_GB_non_clos	RMWSLE_RF_non_clos
2012	0,24%	0,10%	0,14%	0,24%	0,09%	0,14%	0,25%	0,13%	0,14%
2013	0,23%	0,10%	0,14%	0,23%	0,09%	0,13%	0,24%	0,14%	0,13%
2014	0,18%	0,09%	0,13%	0,18%	0,08%	0,13%	0,18%	0,13%	0,13%
2015	0,16%	0,08%	0,13%	0,16%	0,08%	0,12%	0,20%	0,12%	0,12%

Tableau 4.1.2.1 (a) – Métriques de performance RMSLEW (ou RMWSLE)

	RWSE								
Année d'arrêté	RWSE_CART_all	RWSE_GB_all	RWSE_RF_all	RWSE_CART_clos	RWSE_GB_clos	RWSE_RF_clos	RWSE_CART_non_clos	RWSE_GB_non_clos	RWSE_RF_non_clos
2012	20,14%	3,87%	8,23%	12,85%	1,63%	4,87%	32,03%	7,39%	13,65%
2013	19,31%	3,90%	6,28%	12,02%	1,33%	3,22%	30,74%	7,67%	10,91%
2014	19,65%	3,70%	6,15%	13,02%	1,26%	3,30%	31,10%	7,55%	10,84%
2015	15,93%	2,66%	5,97%	9,86%	0,83%	3,48%	28,34%	6,09%	10,97%

Tableau 4.1.2.1 (b) – Métriques de performance RSEW (ou RWSE)

	R2W								
Année d'arrêté	R2W_CART_all	R2W_GB_all	R2W_RF_all	R2W_CART_clos	R2W_GB_clos	R2W_RF_clos	R2W_CART_non_clos	R2W_GB_non_clos	R2W_RF_non_clos
2012	79,86%	96,13%	91,77%	87,15%	98,37%	95,13%	67,97%	92,61%	86,35%
2013	80,69%	96,10%	93,72%	87,98%	98,67%	96,78%	69,26%	92,33%	89,09%
2014	80,35%	96,30%	93,85%	86,98%	98,74%	96,70%	68,90%	92,45%	89,16%
2015	84,07%	97,34%	94,03%	90,14%	99,17%	96,52%	71,66%	93,91%	89,03%

Tableau 4.1.2.1 (c) – Métriques de performance R2W

Globalement, les modèles entraînés sur la base d'arrêté 2015 sont les plus performants. En effet, ils possèdent les RMSLEW et RSEW les plus faibles et les R2W les plus élevés à une ou deux exceptions près dans chaque cas. Il paraît donc légitime de retenir ces modèles. Enfin, parmi les trois modèles candidats, c'est le modèle calibré avec l'algorithme *Gradient Tree Censored Boosting* qui est le plus performant.

Base	C_Cens_CART	C_Cens_GB	C_Cens_RF
Validation	0,4759	0,887	0,9304
Test	0,485	0,8938	0,9357

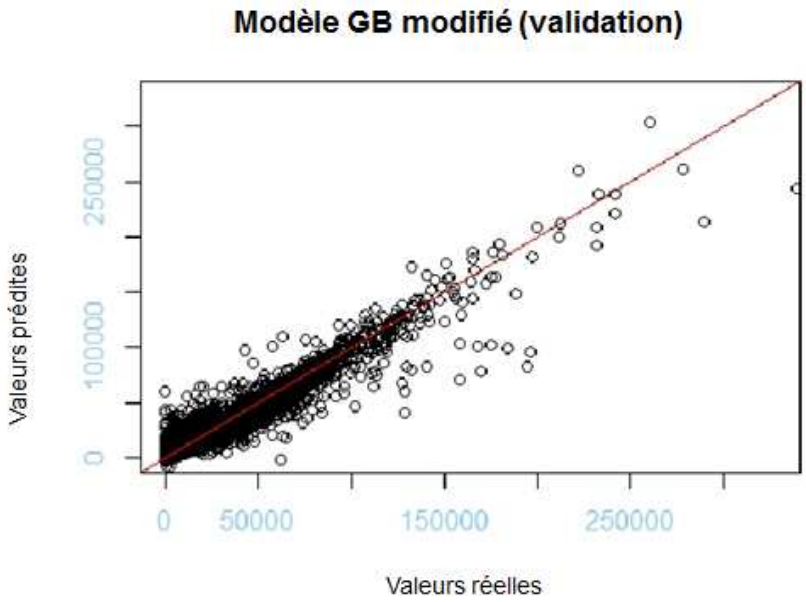
Tableau 4.1.2.1 (d) – Indice de concordance C-Index

L'indice de concordance le plus élevé est celui calculé avec le modèle calibré avec l'algorithme du *Random forest* modifié. Notons toutefois que la valeur du C-Index du modèle calibré avec l'algorithme

Gradient Boosting modifié est très proche. Nous retiendrons comme modèle ce dernier dans la suite de notre analyse.

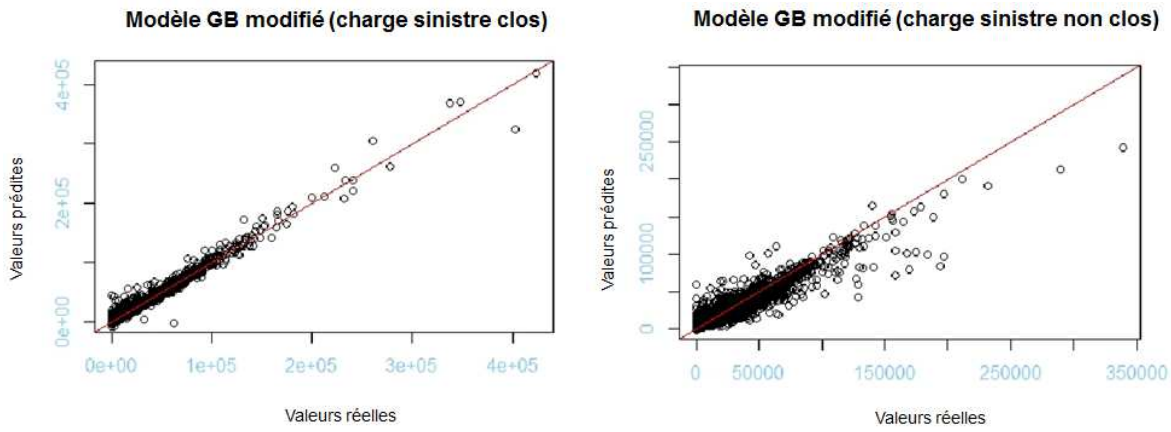
Les cinq variables d'importance utilisées sont : la durée sinistre, le montant sinistre, le code garantie (arrêt de travail ou perte d'emploi), le montant moyen des échéances mensuelles et le poids de Kaplan Meier.

L'analyse des graphiques de comparaison entre les valeurs prédites par le modèle calibré avec l'algorithme du *Gradient Tree censored Boosting* et les valeurs réelles permet d'apprécier la qualité des prédictions suivant la forme de la distribution autour de la diagonale.



Graphique 4.1.2.1 (a) – Prédiction de la charge sinistre avec le modèle *Gradient Boosting* modifié (*Gradient Tree Censored Boosting*)

En effet, par rapport au modèle du *Random forest* modifié, les prédictions du risque vont plutôt dans le sens de la prudence, car globalement légèrement au-dessus de la diagonale.



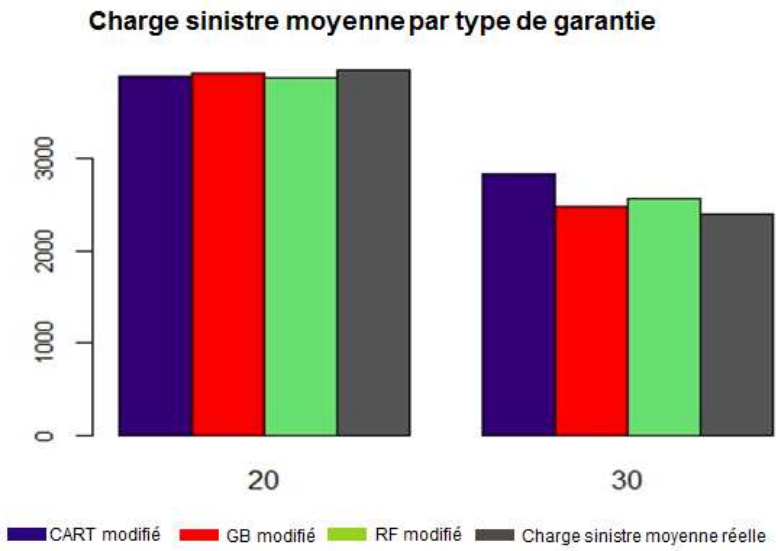
Graphique 4.1.2.1 (b) – Prédiction de la charge sinistre clos et censurés avec le modèle *Gradient Boosting* modifié

La comparaison entre les charges des sinistres clos prédites et réelles montre un alignement presque parfait sur la diagonale. Ce qui signifie la très bonne qualité des prédictions de ce modèle pour les sinistres clos. Concernant la charge sinistre prédite pour les sinistres censurés (sinistres non clos), nous remarquons également une concentration autour de la diagonale. Nous remarquons néanmoins quelques erreurs de classifications. Il s'agit d'un petit nombre restreint de points qui se détachent de cette diagonale et dont les valeurs réelles sont largement supérieures aux valeurs prédites.

Afin de valider les résultats précédents et d'apprécier la prudence des prédictions du modèle, nous allons regarder les prédictions réalisées par les trois modèles de manière plus fine suivant axes d'analyses métier. Nous regarderons plus précisément la qualité des prédictions par type de garantie (20 : Arrêt de travail et 30 : chômage), par réseau de distribution (ou partenaire), par tranche d'âge, et par code clôture (statut à 0 pour les sinistres censurés et à 1 pour les sinistres clos).

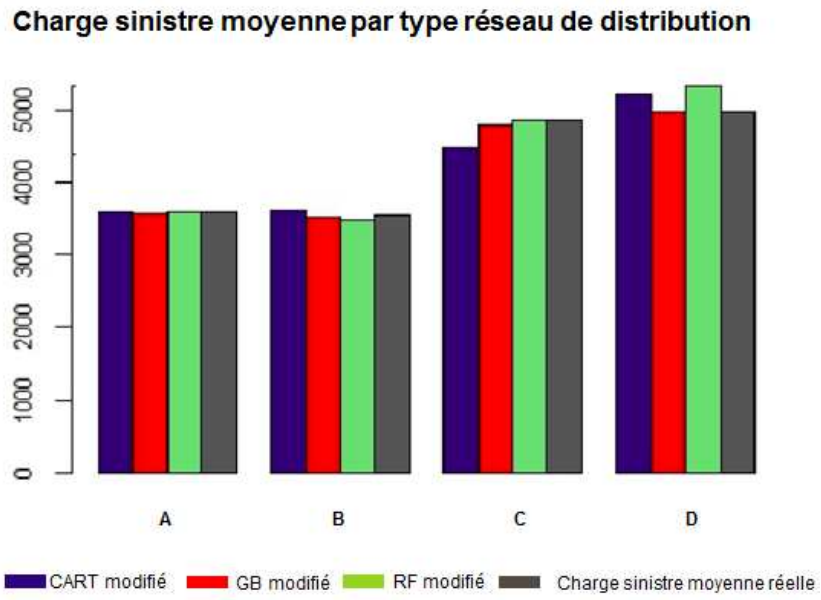
La base de validation retenue dans ce qui suit est la base d'arrêté 2016. Cette base est intéressante car elle n'a pas servi pour le calibrage des modèles lors de l'apprentissage et de la validation. Nous retenons comme base de référence les durées réelles (sinistres clos et sinistres censurés). Nous souhaitons comparer les résultats du modèle GB modifié à ceux des modèles calibrés avec les algorithmes de CART modifié et du RF modifié.

La charge sinistre moyenne par type de garantie prédite par le modèle GB modifié donne la valeur la plus proche de la valeur réelle de ce paramètre.



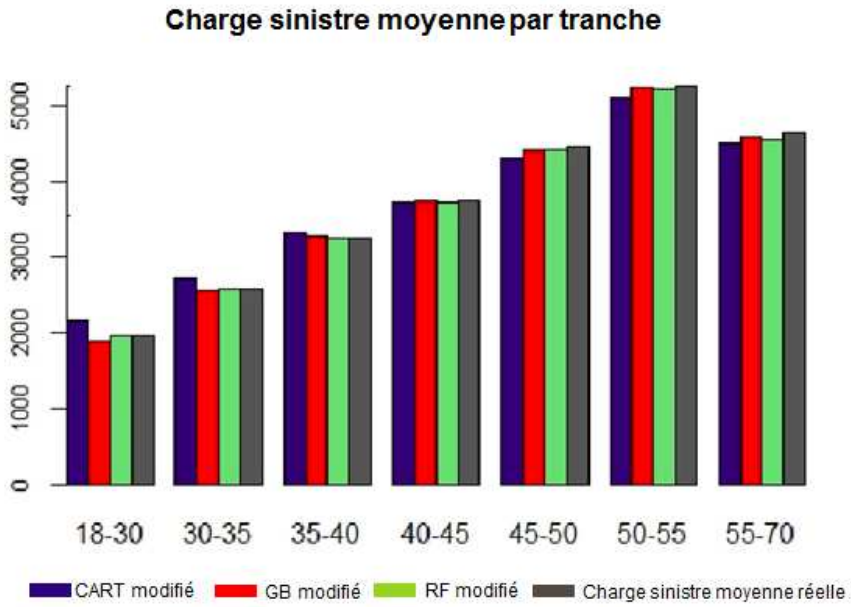
Graphique 4.1.2.2 – Comparaison des charges sinistres moyennes par type de garantie

Les charges sinistres moyennes prédites par réseau de distribution représentant les principaux partenaires avec le modèle GB modifié donne encore les valeurs les plus proches des valeurs réelles de ce paramètre.



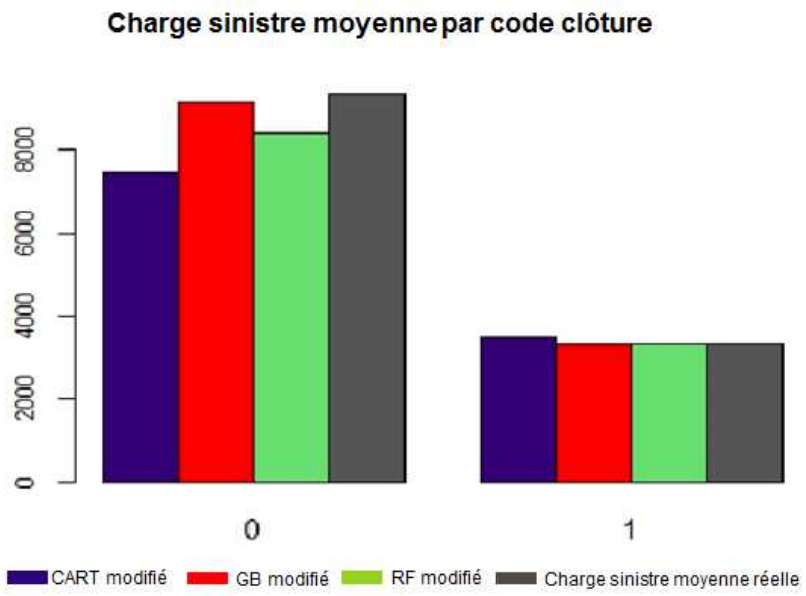
Graphique 4.1.2.3 – Comparaison des charges sinistres moyennes par réseau de distribution

Les charges sinistres moyennes prédites par tranche d'âge avec le modèle GB modifié donne encore les valeurs les plus proches des valeurs réelles.



Graphique 4.1.2.4 – Comparaison des charges sinistres moyennes par tranche d'âge

Enfin, la prédiction des charges sinistres moyennes des sinistres par code clôture (sinistre clos et sinistre censuré) sont meilleures avec le modèle calibré avec le GB modifié.



Graphique 4.1.2.5 – Comparaison des durées moyennes de sinistre par code clôture

En conclusion, nous observons que globalement l'algorithme *Gradient tree censored Boosting* fournit de meilleures prévisions que l'algorithme *Tree-Base Censored* et l'agorithme *Random Forest censored*.

4.2. Résultats de l'étude des contrats de prévoyance collective

L'objectif de cette partie est de valider la méthodologie proposée sur les données d'un portefeuille couvrant le risque arrêt de travail comme les contrats de prêts étudié dans le paragraphe précédent. Sans perdre de généralité, nous limiterons notre analyse à la prédiction de la charge sinistre du risque arrêt de travail (Accident de travail et Indemnités journalières) du portefeuille de prévoyance collective présenté précédemment.

Comme vu précédemment, nous utiliserons également les métriques de performance RMSLEW, RSEW, R2W et le C-Index.

	RMWSLE								
Année d'arrêté	RMWSLE_CART_all	RMWSLE_GB_all	RMWSLE_RF_all	RMWSLE_CART_clos	RMWSLE_GB_clos	RMWSLE_RF_clos	RMWSLE_CART_non_clos	RMWSLE_GB_non_clos	RMWSLE_RF_non_clos
2015	0,48%	0,25%	0,33%	0,47%	0,22%	0,31%	0,52%	0,39%	0,31%
2016	0,41%	0,23%	0,29%	0,40%	0,21%	0,27%	0,47%	0,38%	0,27%

Tableau 4.2.1.1 (a) – Métriques de performance RMSLEW (ou RMWSLE)

	RWSE								
Année d'arrêté	RWSE_CART_all	RWSE_GB_all	RWSE_RF_all	RWSE_CART_clos	RWSE_GB_clos	RWSE_RF_clos	RWSE_CART_non_clos	RWSE_GB_non_clos	RWSE_RF_non_clos
2015	20,58%	15,30%	12,12%	20,22%	16,08%	9,00%	24,91%	18,16%	15,66%
2016	17,55%	9,03%	9,34%	13,78%	3,04%	2,72%	23,67%	14,26%	14,97%

Tableau 4. 2.1.1 (b) – Métriques de performance RSEW (ou RWSE)

	R2W								
Année d'arrêté	R2W_CART_all	R2W_GB_all	R2W_RF_all	R2W_CART_clos	R2W_GB_clos	R2W_RF_clos	R2W_CART_non_clos	R2W_GB_non_clos	R2W_RF_non_clos
2015	79,42%	84,70%	87,88%	79,78%	83,92%	91,00%	75,09%	81,84%	84,34%
2016	82,45%	90,97%	90,66%	86,22%	96,96%	97,28%	76,33%	85,74%	85,03%

Tableau 4. 2.1.1 (c) – Métriques de performance R2W

Globalement, les modèles entraînés sur la base d'arrêté 2016 sont les plus performants. En effet, ils possèdent les RMSLEW et RSEW les plus faibles et les R2W les plus élevés. Il paraît donc légitime de retenir ces modèles. Enfin, parmi les trois modèles candidats, à quelques exceptions près, c'est le modèle calibré avec l'algorithme *Gradient Tree Censored Boosting* qui est le plus performant.

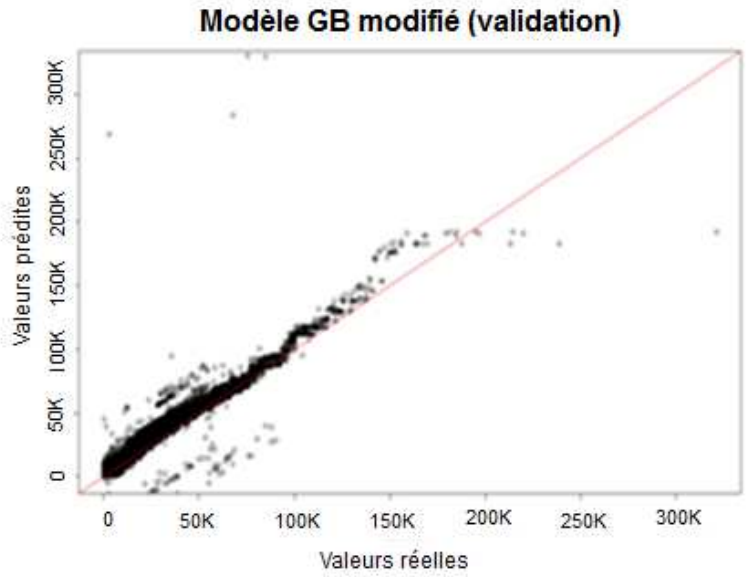
Base	C_Cens_CART	C_Cens_GB	C_Cens_RF
Validation	0,4759	0,8870	0,9304
Test	0,4850	0,8938	0,9357

Tableau 4. 2.1.1 (d) – Indice de concordance C-Index

L'indice de concordance le plus élevé est celui calculé avec le modèle calibré avec l'algorithme du *Random forest* modifié. Notons toutefois que la valeur du C-Index du modèle calibré avec l'algorithme *Gradient Boosting* modifié est très proche. Nous retiendrons comme modèle ce dernier dans la suite de notre analyse.

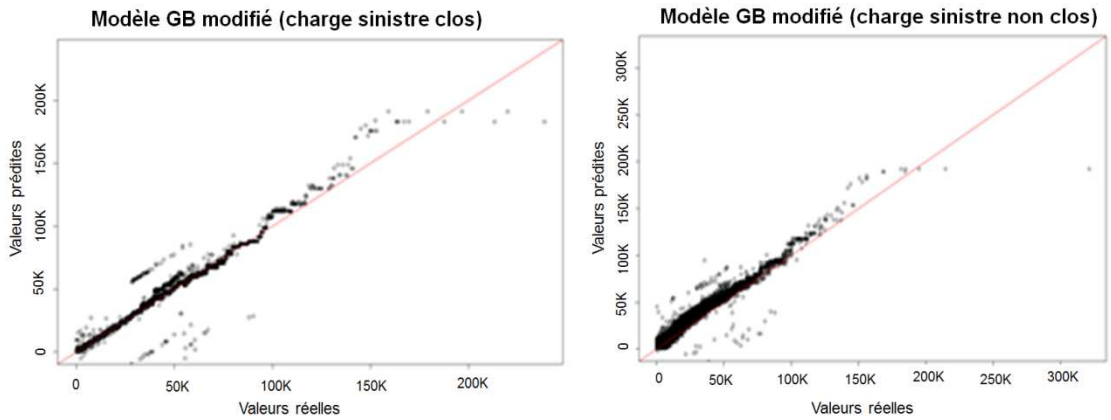
Les principales variables d'importance utilisées par les algorithmes par ordre d'importance sont : le montant des prestations payées, le poids de Kaplan-Meier, l'âge de survenance du sinistre, la catégorie de risque (Maladie Ordinaire, Longue Maladie, Maladie Longue Durée), le code clôture (sinistre clos ou censuré), la durée du sinistre.

L'analyse des graphiques de comparaison entre les valeurs prédites par le modèle calibré avec l'algorithme du *Gradient Tree censored Boosting* et les valeurs réelles permet d'apprécier la qualité des prédictions suivant la forme de la distribution autour de la diagonale.



Graphique 4.1.2.1 (a) – Prédiction de la charge sinistre avec le modèle *Gradient Boosting* modifié (*Gradient Tree Censored Boosting*)

Les prédictions du risque vont plutôt dans le sens de la prudence, car globalement légèrement au-dessus de la diagonale. Nous notons également la présence d'un certain nombre de sinistres sous-estimés. La comparaison des prédictions par code clôture (sinistre clos ou censuré), permet d'identifier ces sinistres afin de les analyser plus finement.



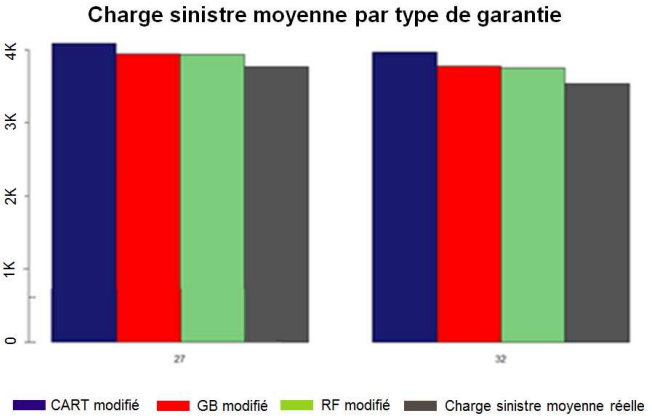
Graphique 4.1.2.1 (b) – Prédiction de la charge sinistre clos et censurés avec le modèle *Gradient Boosting* modifié

La comparaison entre les charges des sinistres clos prédites et réelles montre un bon alignement avec la diagonale. Ce qui signifie la très bonne qualité des prédictions de ce modèle pour les sinistres clos. Concernant la charge sinistre prédite pour les sinistres censurés (sinistres non clos), nous remarquons également une concentration autour de la diagonale. Ces deux comparaisons confirment la présence de quelques erreurs de classifications d'une part et d'autre de la diagonale.

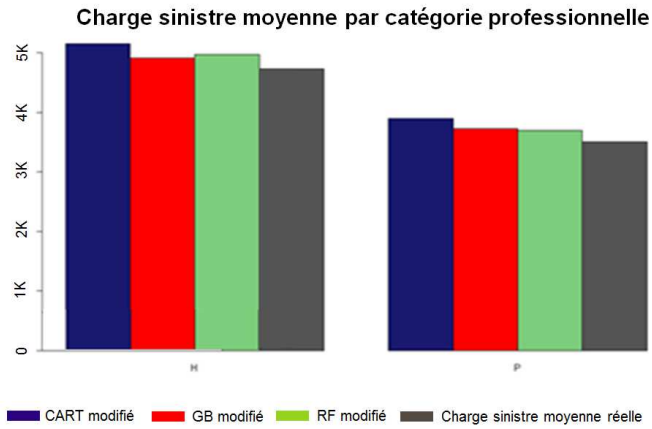
Afin de valider les résultats précédents et d'apprécier la prudence des prédictions du modèle, nous allons regarder les prédictions réalisées par les trois modèles de manière plus fine suivant axes d'analyses métier. Nous regarderons plus précisément la qualité des prédictions par type de garantie suite à l'arrêt de travail (27 pour Accident de travail et 32 pour Indemnités journalières), par catégorie professionnelle (H pour personnel hospitalier et P pour agents de la fonction publique), par tranche d'âge, et par code clôture (statut à 0 pour les sinistres censurés et à 1 pour les sinistres clos).

La base de test retenue dans ce qui suit est la base d'arrêté 2016 n'ayant pas servi pour le calibrage des modèles lors de l'apprentissage et de la validation. Nous comparons les charges sinistres moyennes prédites aux réelles. Nous présentons les résultats des trois modèles proposés (CART modifié, GB modifié et du RF modifié).

Quelque soit la segmentation considérée, nous remarquons que les charges sinistres moyennes prédites par les modèles GB modifié et RF modifié sont assez proches et meilleures aux valeurs prédites par le modèle CART modifié.

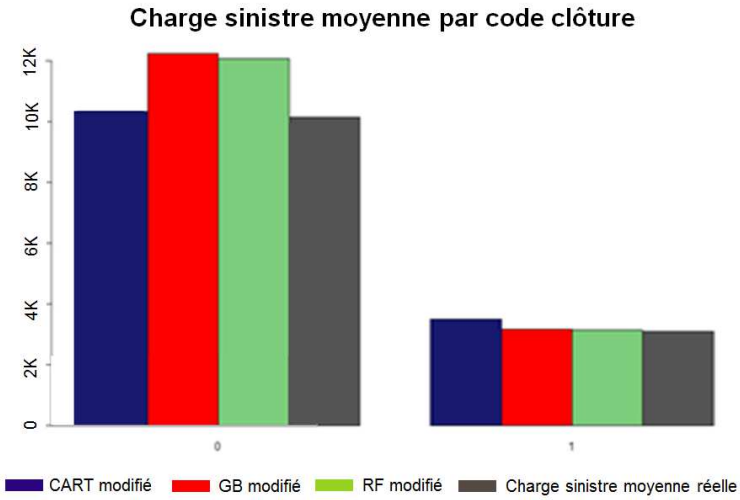


Graphique 4.1.2.2 – Comparaison des charges sinistres moyennes par type de garantie (27 : Accident de travail et 32 : Indemnités journalières)



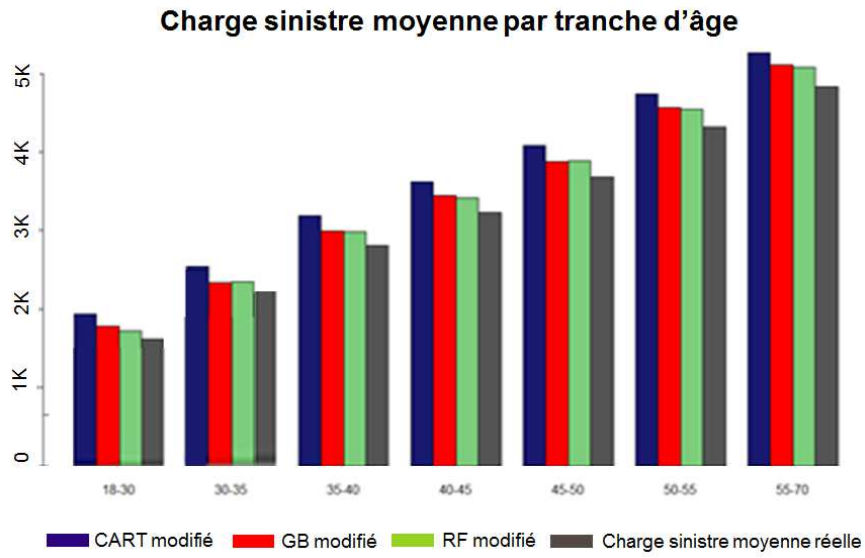
Graphique 4.1.2.3 – Comparaison des charges sinistres moyennes par catégorie professionnelle (H : personnel hospitalier et P : agents de la fonction publique)

Les valeurs prédites par les modèles GB modifié et RF modifié surestiment légèrement les valeurs réelles. Ce phénomène s'explique par le fait que les valeurs réelles moyennes sont calculées sans distinction des sinistres clos et censurés. L'analyse par code clôture (sinistre clos ou censuré) devrait nous renseigner sur l'anticipation des prédictions des modèles GB modifié et RF modifié, dans les calculs des valeurs prédites des sinistres censurés.



Graphique 4.1.2.5 – Comparaison des durées moyennes de sinistre par code clôture

Nous constatons que les prédictions des charges sinistres moyennes pour les sinistres clos sont quasi parfaites. La distinction entre les sinistres clos et les sinistres censurés permet de mettre en évidence l'anticipation des charges sinistres moyennes des sinistres censurés, correspondant à l'écart entre la valeur réelle et la valeur prédite.



Graphique 4.1.2.4 – Comparaison des charges sinistres moyennes par tranche d'âge

Les remarques précédentes sont toujours confirmées sous l'angle de la segmentation par tranche d'âge. En conclusion, puisque nous obtenons quasiment les mêmes résultats, retenir le modèle GB modifié est conseillé, car ce modèle d'un point de vue opérationnel est plus rapide que le RF modifié.

5. Discussion

En présence de données censurées notamment sur des risques de prévoyance en assurances, les méthodes statistiques classiques ne permettent pas une prise en compte adéquate de ces informations, ce qui est source de biais dans les estimations des risques.

Méthodes	Avantages	Inconvénients
Chain Ladder (méthode agrégée)	Facile à implémenter Facile à comprendre Intègre implicitement les IBNR (*Incurred But Not Reported*)	Hypothèses « trop » forte : -> Suppose la stabilité des provisions -> Suppose l'indépendance des survenances entre année N'exploite pas les informations tête par tête
Micro-Level Reserving (GLM)	Exploite les informations tête par tête Facile à expliquer	Performance du modèle jugée sur sa qualité asymptotique plutôt que sur son adéquation aux données « réelles » Modèle paramétrique difficile à implémenter (en pratique) Risques d'erreurs d'estimation des paramètre des modèles (avec plusieurs paramètres)

Graphique 5.1 – Avantages et inconvénients des méthodes Chain Ladder et GLM

Ces méthodes sont souvent utilisées par les praticiens lors du provisionnement des contrats de prévoyance non-vie. Les biais d'estimation des risques conduisent donc à des erreurs d'estimation des provisions.

Il est donc utile de réfléchir sur des approches alternatives permettant de mieux prédire les risques pour les calculs de provisions. Dans ce cadre, l'analyse des critères d'améliorations des prédictions des

risques et de provisionnement permet de déterminer cinq principales caractéristiques « souhaitables » pour l'élaboration d'une méthode alternative. Ces caractéristiques minimums sont complémentaires :

1) Il faut privilégier l'utilisation des données individuelles (informations détaillées sur les sinistres).
2) Il ne faut pas imposer des hypothèses a priori sur la stabilité des provisions et l'indépendance des sinistres.
3) Il est souhaitable d'utiliser un modèle non-paramétrique pour éviter les erreurs d'estimation des paramètres.
4) Le critère de mesure de la performance du modèle doit être fondé sur l'adéquation aux données réelles uniquement.
5) Le modèle doit être capable de gérer les censures (données censurées).

A partir de ces constats, le choix des algorithmes d'apprentissage automatique semble répondre aux différentes caractéristiques « souhaitables ». Il faut noter que l'utilisation de ces algorithmes présente quelques inconvénients. D'une part, le calibrage des modèles par ces algorithmes nécessitent une volumétrie importante de données pour la répartition en base d'apprentissage, de validation et de test. D'autre part, les résultats de ces modèles sont difficilement interprétables, puisque ces algorithmes sont construits dans un objectif de performance et non d'interprétation.

Modèles	Avantages	Inconvénients
CART Random Forest Xg Boost	Pas d'hypothèses La qualité du modèle est jugé sur l'adéquation avec l'échantillon de validation Donne de meilleures prédictions que les méthodes classiques (Chain Ladder, GLM…)	Il faut une volumétrie suffisante de données (taille de l'échantillon d'apprentissage) Modèles sont difficilement interprétables

Graphique 5.2 – Avantages et inconvénients des méthodes d'apprentissage automatique

Notons également que la mise en œuvre des solutions possibles nécessite de procéder à des adaptations de ces algorithmes d'apprentissage automatiques reconnus pour leur performance. En effet, ces algorithmes sont capables de mieux exploiter la richesse des données tête par tête sans a priori et sans faire d'hypothèses (trop forte). Ils savent réaliser les segmentations nécessaires pour la prise en compte de l'hétérogénéité des risques des populations considérées. Par contre, certaines adaptations algorithmiques et ajustements des hyper-paramètres sont nécessaires pour gérer les individus censurées (sinistres non clôturés) contenues dans les données, puisque nous sommes en présence de données de survie.

Modèles	Avantages	Inconvénients
CART modifiée (*Tree Censored Regression*)	Prise en compte de l'hétérogénéité des sinistres pour une meilleure segmentation Utilisation des données tête par tête Prise en compte l'information à la maille contrat (individus) Possibilité de pondérer les observations (prise en compte les sinistres censurés)	Instable à cause de l'algorithme CART (volatilité des prédictions en fonction de l'échantillon d'apprentissage pour une même population)
Random Forest Censored Gradient Tree Censored Boosting	Modèles plus stable que l'algorithme CART Meilleure prédiction par ces algorithmes Prise en compte de plus d'informations Utilisation de presque toutes les variables (features)	Interprètabilité

Graphique 5.3 – Avantages et inconvénients des méthodes d'apprentissage automatique modifiées

Pour illustrer nos propos, après avoir améliorer les algorithmes suivant l'approche proposée, nous les avons testés sur des données réelles de deux portefeuilles d'assurance non-vie (assurance de prêt et prévoyance collective) pour estimer la durée de maintien des sinistres et la charge sinistre des risques d'arrêt de travail et de perte d'emploi. Nous avons obtenu des résultats cohérents et des prédictions de risque plus performant avec les algorithmes RF modifié et GB modifié (respectivement *Random Forest Censored* et *Gradient Tree Censored Boosting*) que l'algorithme du CART modifié de Lopez et al. (*Tree Censored Regression*). Par notre approche, nous corrigeons également le biais d'instabilité spécifique à la méthode du CART modifié.

Le modèle calibré sur les données de prévoyance collective peut être amélioré afin de réduire la taille des sinistres mal prédits en augmentant la taille de l'échantillon. Une solution serait d'enrichir les données de l'assureur est l'utilisation des données publiques (Données INSEE, Base DAMIR…).

6. Conclusion

En assurance non-vie, pour la tarification des affaires ou le provisionnement des sinistres, il est d'usage d'avoir recourt à des modèles statistiques pour prédire les probabilités d'incidence et de maintien sur une période à partir de données de durée. Ces données sont souvent soumises au phénomène de censure à droite, caractérisant les individus partiellement observés durant la période d'estimation des probabilités. Ainsi, en matière de provisionnement non-vie, différentes méthodes classiques telles que le *Chain Ladder*, le Micro *Level Reserving* (GLM) et l'estimateur de Kaplan-Meier sont couramment utilisées par les praticiens. Malgré leur facilité de mise en œuvre, elles présentent certaines limites et ne permettent pas d'utiliser toute la richesse des données individuelles disponibles pouvant porter du signal donc contribuer à expliquer le risque. Ces données peuvent être des données identitaires, des informations transactionnelles et relationnelles, des données démographiques, des variables contractuelles et sur l'état des assurés.

Récemment, le modèle d'arbres de régression censurées (*Tree base censored*) une adaptation du modèle CART a été proposée comme une alternative gérant à la fois les censures à droite et utilisant différentes variables explicatives disponibles afin d'extraire les informations pertinentes pour la prédiction. La principale limite de cette solution est son instabilité héritée du modèle CART. De plus les tests ont été réalisés sur des données simulées.

La valeur ajoutée principale de notre étude est double. D'une part, nous proposons de corriger l'instabilité de CART par l'utilisation de deux algorithmes d'apprentissage automatique (*Random Forest* et *Gradient Boosting*) fournissant des prédictions plus robustes et plus stables. Ces algorithmes ont été adaptés pour prendre en compte les données censurées à droite par l'introduction comme paramètre des poids de Kaplan-Meier (*Inverse Probability of Censoring Weighting*). Nous avons également adapté les métriques de performance pour tenir compte de la présence des censures à droite dans les données et de l'échelle de grandeur des variables à expliquer. D'autre part, l'illustration par des tests sur deux jeux de données réelles, montre que notre approche semble robuste et améliore les résultats. Nous appliquons ces algorithmes aux données de portefeuilles d'assurance de prêts et de prévoyance collective d'une compagnie d'assurance de personnes pour estimer les durées de maintien et les charges sinistres. Les différentes méthodes proposées fournissent des résultats plus performants qu'à ceux des modèles classiques. Elles apportent des améliorations significatives à la méthode de Kaplan-Meier, et permettent de faire du provisionnement tête par tête, en exploitant au mieux toutes les informations disponibles. Nous avons également constaté que les prédictions des risques en prévoyance collective pouvaient être améliorées en complétant les bases de l'assureur. Une manière d'enrichir les données de l'assureur est l'utilisation des données publiques (Données INSEE, Base DAMIR…). Les adaptations proposées pourraient s'adapter facilement à la prédiction des risques de maintien en dépendance et de prévoyance individuelle.

Enfin, notons que ces méthodes de Machine Learning ont également certaines limites. Elles ont besoin de quantité importante de données pour le calibrage des modèles. Elles produisent des résultats certes performants mais difficilement interprétables.

Références

[1] Astesan, E. Les réserves techniques des sociétés d'assurances contre les accidents automobiles ; Librairie générale de droit et de jurisprudence, 1938.

[2] Arjas, E. (1989). The claims reserving problem in non-life insurance: some structural ideas. ASTIN Bulletin 19/2, 139-152.

[3] Jewell, W.S. (1989). Jewell, W.S. Predicting IBNYR events and delays I. Continuous time. ASTIN Bulletin 1989, 19(1), 25-55.

[4] Norberg, R. (1993). Prediction of outstanding liabilities in non-life insurance. ASTIN Bulletin 1993, 23(1), 95-115.

[5] Hesselager, O. A Markov model for loss reserving. ASTIN Bulletin 1994, 24(2), 183-193.

[6] Antonio, K., Plat, R. (2014). Micro-level stochastic loss reserving for general insurance. Scandinavian Actuarial Journal 2014/7, 649-669.

[7] Badescu, A.L., Lin, X.S., Tang, D. (2016). A marked Cox model for the number of IBNR claims: theory. Insurance: Mathematics & Economics 69, 29-37.

[8] Badescu, A.L., Lin, X.S., Tang, D. (2016). A marked Cox model for the number of IBNR claims: estimation and application. Version March 14, 2016. SSRN Manuscript 2747223.

[9] Baudry, M., Robert, C.Y. (2017). Non parametric individual claim reserving in insurance. Preprint.

[10] Harej, B., Gachter, R., Jamal, S. (2017). Individual claim development with machine learning. ASTIN Report.

[11] Jessen, A.H., Mikosch, T., Samorodnitsky, G. (2011). Prediction of outstanding payments in a Poisson cluster model. Scandinavian Actuarial Journal 2011/3, 214-237.

[12] Lopez, O. (2018). A censored copula model for micro-level claim reserving. HAL Id: hal-01706935.

[13] Pigeon, M., Antonio, K., Denuit, M. (2013). Individual loss reserving with the multivariate skew normal framework. ASTIN Bulletin 43/3, 399-428.

[14] Verrall, R.J., Wüthrich, M.V. (2016). Understanding reporting delay in general insurance. Risks 4/3, 25.

[15] Zarkadoulas, A. (2017). Neural network algorithms for the development of individual losses. MSc thesis, University of Lausanne.

[16] Friedland, J. (2010). Estimating Unpaid Claims Using Basic Techniques; Estimating Unpaid Claims Using Basic Techniques.

[17] Wüthrich (2018), Neural Networks Applied to Chain-Ladder Reserving, European Actuarial Journal, December 2018, Volume 8, Issue 2, pp 407–436.

[18] Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. ASTIN Bulletin 23/2, 213-225.

[19] Hoem J.M. (1971), « Point Estimation of Forces of Transition in Demographic Models », Journal of the Royal Statistical Society. Series B (Methodological), vol. 33, n°2, pp. 275−289.

[20] Kaplan, E.L. et Meier P. (1958), « Nonparametric Estimation from Incomplete Observations », Journal of the American Statistical Association, vol. 53, n°282, pp. 457−481.

[21] Cooney, M.T., Dudina, A.L., Graham, I.M., (2009), Value and limitations of existing scores for the assessment of cardiovascular risk: a review for clinicians, J. Am. Coll. Cardiol. 54 (14) 1209–1227.

[22] Matheny, M., McPheeters, M.L., Glasser, A., Mercaldo, N., Weaver, R.B., Jerome, R.N., Walden, R., McKoy, J.N., Pritchett, J., Tsai, C., (2011), Systematic review of cardiovascular disease risk assessment tools, Tech. Rep., Agency for Healthcare Research and Quality (US).

[23] D'Agostino, R.B, Vasan, R.S., Pencina, M.J., Wolf, P.A., Cobain, M., Massaro, J.M., Kannel, W. B., (2008), General cardiovascular risk profile for use in primary care: the Framingham heart study, Circulation 118 (4) E86.

[24] Ridker, P.M., Buring, N., Rifai, J.E., Cook, N.R., (2007), Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds risk score, JAMA: J. Am. Med. Assoc. 297 (6) 611–619.

[25] Ridker, P.M., Paynter, N.P., Rifai, N., Gaziano, J.M., Cook, N.R., (2008), C-Reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds risk score for men, Circulation 118 (18) S1145.

[26] Goff, D.C., Lloyd-Jones, D.M., Bennett, S., Coady, G., D'Agostino, R.B., Gibbons, R., Greenland, P., Lackland, D.T., Levy, D., O'Donnell, C.J., Robinson, J.G., Schwartz, J.S., Shero, S. T., Smith, S.C., Sorlie, P., Stone, N.J., Wilson, P.W., (2014), ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association task force on practice guidelines, J. Am. Coll. Cardiol. 63 (25) 2935–2959.

[27] Collins, G.S., Altman, D.G., (2009), An independent external validation and evaluation of QRISK cardiovascular risk prediction: a prospective open cohort study, Br. Med. J. 339 b2584.

[28] Kalbfleisch, J.D., Prentice, R.L., (2002), The Statistical Analysis of Failure Time Data, Wiley, Hoboken, NJ.

[29] Sierra, B., Larranaga, P., (1998), Predicting survival in malignant skill melanoma using Bayesian networks automatically induced by genetic algorithms: an empirical comparison between different approaches, Artif. Intell. Med. 14 (1-2) 215 - 230.

[30] Blanco, R., Inza, I., Merino, M., Quiroga, J., Larrañaga, P., (2005), Feature selection in Bayesian classifiers for the prognosis of survival of cirrhotic patients treated with TIPS, J. Biomed. Inform. 38 (5) 376 -388.

[31] Kattan, M.W., Hess, K.R., Beck, J.R., (1998), Experiments to determine whether recursive partitioning (CART) or an artificial neural network overcomes theoretical limitations of Cox proportional hazards regression, Comput. Biomed. Res. 31 (5) 363–373.

[32] Štajduhar, I., Dalbelo-Bašic´, B., Bogunovic´, N., (2009), Impact of censoring on learning Bayesian networks in survival modelling, Artif. Intell. Med. 47 (3) 199 - 217.

[33] Štajduhar, I., Dalbelo-Bašic´, B., (2010), Learning Bayesian networks from survival data using weighting censored instances, J. Biomed. Inform. 43 (4) 613–622.

[34] Hothorn, T., Lausen, B., Benner, A., Radespiel-Tröger, M., (2004), Bagging survival trees, Stat. Med. 23 (1) 77–91.

[35] Ishwaran, H., Kogalur, U.B., Blackstone, E.H., Lauer, M.S., (2008), Random survival forests, Ann. Appl. Stat. 841–860.

[36] Ibrahim, N.A., Abdul Kudus, A., Daud, I., Abu Bakar, M.R., (2008), Decision tree for competing risks survival probability in breast cancer study, Int. J. Biol. Med. Sci. 3 (1) 25–29.

[37] Lucas, P.J.F., van der Gaag, L.C., Abu-Hanna, A., (2004), Bayesian networks in biomedicine and health-care, Artif. Intell. Med. 30 (3) 201–214.

[38] Bandyopadhyay, S., Wolfson, J., Vock, D.M., Vazquez-Benitez, G., Adomavicius, G., Elidrisi, M., Johnson, P.E., O'Connor, P.J., (2015), Data mining for censored time-to-event data: a Bayesian network model for predicting cardiovascular risk from electronic health record data, Data Min. Knowl. Disc. 29 (4) 1033–1069.

[39] Biganzoli, E., Boracchi, P., Mariani, L., Marubini, E., (1998), Feed forward neural networks for the analysis of censored survival data: a partial logistic regression approach, Stat. Med. 17 (10) 1169–1186.

[40] Ripley, B.D., Ripley, R.M., (2001), Neural networks as statistical methods in survival analysis, Clin. Appl. Artif. Neural Networks 237–255.

[41] Shivaswamy, P.K., Chu, W., Jansche, M., (2007), A support vector approach to censored targets, in: Seventh IEEE International Conference on Data Mining. ICDM 2007, IEEE, 2007, pp. 655–660.

[42] Khan, F.M., Zubek, V.B., (2008), Support vector regression for censored data (SVRc): a novel tool for survival analysis, in: Eighth IEEE International Conference on Data Mining (ICDM 2008), IEEE, pp. 863–868.

[43] Vock, D. M., Wolfson, J., Bandyopadhyay, S., Adomavicius, G., Johnson, P. E., Vazquez-Benitez, G., & O'Connor, P. J. (2016), Adapting machine learning techniques to censored time-to-event health record data: A general-purpose approach using inverse probability of censoring weighting. Journal of biomedical informatics, 61, 119-131.

[44] Wu, J., Roy, J., Stewart, W.F., (2010), Prediction modeling using EHR data: challenges, strategies, and a comparison of machine learning approaches, Med. Care 48 (6) S106–S113.

[45] Lin, Y.-K., Chen, H., Brown, R.A., Li, S.-H., Yang, H.-J., (2014), Predictive analytics for chronic care: a time-to-event modeling framework using electronic health records, Published by the IEEE Computer Society.

[46] Lopez, O., Milhaud, X., Therond, P.E., (2016), Tree-based censored regression with applications in insurance, Electronic Journal of Statistics, Volume 10 issue 2, pp.2685-2716.

[47] Breiman, L., Friedman, J., Olshen, R., Stone, C., (1984) : Classification and regression trees. Wadsworth Books, 358.

[48] Stute, W. (1999). Nonlinear censored regression. Statistica Sinica, 1089-1102.

[49] Bang, H., Tsiatis, A.A., (2000), Estimating medical costs with censored data, Biometrika 87 (2) 329 -343.

[50] Tsiatis, A.A., (2006), Semiparametric Theory and Missing Data, Springer, New York.

[51] Stute, W. (1993). Consistent estimation under random censorship when covariables are present. Journal of Multivariate Analysis, 45(1), 89-103.

[52] Hothorn, T., Hornik, K., Zeileis, A., (2006), Unbiased Recursive Partitioning: A Conditional Inference Framework. Journal of Computational and Graphical Statistics, vol. 15, no 3, p. 651–674 (DOI 10.1198/106186006X133933, JSTOR 27594202)

[53] Strobl, C., Malley, J., Tutz, G., (2009), An Introduction to Recursive Partitioning: Rationale, Application and Characteristics of Classification and Regression Trees, Bagging and Random Forests. Psychological Methods, vol. 14, n°4, p. 323–348 (DOI 10.1037/a0016973)

[54] Nisbet, R., Elder, J., Miner, G, (2009), Handbook for Statistical Analysis and Data Mining, Academic Press, Page 247 Edition.

[55] Breiman, L., (2001), Random Forests. Machine Learning 45 (1): 5-32. Doi: 10.1023/A: 1010933404324

[56] Breiman, L., (1996). Bagging predictors. Machine learning, 24 (2), 123-140.

[57] Lemberger, P., Batty, M, Morel, M., Raffaëlli, J-L., (2015), Big Data et Machine Learning, Dunod, pp 130-131.

[58] Davison, A. C., Hinkley, D. V., (1997), Bootstrap Methods and Their Application, Cambridge University Press. ISBN 0-521-57471-4.

[59] Rouviere, L., (2018), Introduction aux méthodes d'agrégation : boosting, bagging et forêts aléatoires. Illustrations avec R.

[60] Gregorutti, B., Michel, B., Saint-Pierre, O., (2014), Correlation et importance des variables dans les forêts aléatoires.

[61] Genuer, R., Poggi, J-M., Tuleau-Malot, C., (2012), Variable selection using Random Forests.

[62] Freund Y., Schapire R. (1996), Experiments with a new boosting algorithm. In Proceedings of the Thirteenth International Conference on Machine Learning.

[63] Friedman, J-H., (2002), Stochastic gradient boosting, Computational Statisrics and Data Analysis 38.

[64] Chen, Y., Jia, Z., Mercola, D., & Xie, X. (2013), A Gradient Boosting Algorithm for Survival Analysis via Direct Optimization of Concordance Index. Computational and Mathematical Methods in Medicine, 2013, 873595. http://doi.org/10.1155/2013/873595

[65] Chen, T., Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System, Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, p. 785–794.

Annexes

A. Etude des données du portefeuille des contrats de prêts : variables d'intérêt

Nom des variables	Description de la variable
Code Clôture (0 : ouvert / 1 : clôturé)	Indicateur de sinsitre clos à date de fin d'observation (0 : sinistre ouvert / 1 : sinistre clôturé)
Code Catégorie socio professionnelle (1 à 8)	CSP de l'assuré sinistré
Code Département	Département d'habitation de l'assuré sinistré
Code Produit	Réseau de distribution (Partenaire)
Code Risque	Arrêt de travail (20) / Perte d'emploi (30)
Code sexe	1 : Homme (53,42%) / 2 : Femme (43,62%) / 3 : Personne morale (0,04%)
Code type échéance	C : Prêts à taux fixe (85,79%) / R : Prêts relais (2,16%) / V : Prêts à taux variable (8,99%)
Montant capiltal initial	Montant de la somme empruntée
Montant Capiltal Restant Dû	Montant du CRD à la date de fin d'observation (31/12/2016)
Montant moyen échéance assurance mensuelle	Moyenne des échéances de prêt payées par l'assuré au titre de sa garantie d'assurance
Montant sinistre	Montant cumulé des paiements de sinistres depuis l'ouverture du sinistre
Nature de prêt	Prêt immobilier (88,35%) / Prêt professionnel (9,51%) / Prêt consommation (2,13%)
Taux d'emprunt	Taux d'emprunt réel du prêt
Date de chargement de la base	Date de chargement de la base dans le datawarehouse de l'entreprise
Date d'arrêté	Date d'arrêté de la base permettant de distinguer les différentes dates des clôture
Date de fin de garantie	Conditions maximales de remboursement des sinistres pour la garantie (durée et âge maximal)
Date de survenance	Date de survenance du sinistre
Date de clôture sinistre	Date de clôture du sinistre par l'assureur
Date de fin d'indemnisation	Date de fin d'indemnisation par l'assureur au titre du sinistre
Date de dernier paiement	Date de dernier règlement effectué par l'assureur au titre du sinistre
Date de souscription	Date de souscription du contrat d'assurance
Date de début d'indemnisation	Date de début d'indemnisation du sinistre
Date de naissance	Date de naissance de l'assuré sinistré
Date de fin de prêt	Date de fin du prêt et d'arrêt de toutes les garanties souscrites
Age souscription	Age atteint par l'assuré sinistré à la souscription du contrat (par différence de millésime)
Age survenance	Age atteint par l'assuré sinistré à la survenance du sinistre (par différence de millésime)
Durée du sinistre	Durée de maintien du sinistre calculée à partir des variables "Date_Survenance" et "Date_Clo_Sin"
Délais avant survenance du sinistre	Délais passée en assurance depuis la souscription avant survenance
Montant des prestations	Charge sinistre estimée à partir de la variable "MNT_SINISTRE" et "Date_Clo_Sin"

Tableau 3.1.3.1 (a) – Variables d'intérêt

Nom des variables	Valeurs maquantes (%)	Minimum	1er quartile	Médiane	Moyenne	3ème quartile	Maximum
Code Clôture (0 : ouvert / 1 : clôturé)	0%						
Code Catégorie socio professionnelle (1 à 8)	3%						
Code Département	0%						
Code Produit	0%						
Code Risque	0%						
Code sexe	2,92%						
Code type échéance	3,07%						
Montant capiltal initial	2,92%	-	12 911	33 000	51 097	71 926	201 012 690
Montant Capiltal Restant Dû	0%	-	-	1 053	20 749	24 201	65 185 763
Montant moyen échéance assurance mensuelle	**17,03%**	-	68	230	**863**	**454**	404 000 000
Montant sinistre	0,00%	-	310	1 251	**4 096**	**3 949**	495 853
Nature de prêt	0,01%						
Taux d'emprunt	2,93%	1,90%	3,60%	4,40%	4,50%	4,90%	10,50%
Date de chargement de la base	0%						
Date d'arrêté	0%						
Date de fin de garantie	0%						
Date de survenance	0%						
Date de clôture sinistre	0%						
Date de fin d'indemnisation	0%						
Date de dernier paiement	0%						
Date de souscription	0%						
Date de début d'indemnisation	0%						
Date de naissance	0%						
Date de fin de prêt	0%						
Age souscription	0%	18	30	37	37,64	45	66
Age survenance	0%	18	37	44	43,53	51	68
Durée du sinistre	0%	0,00	3,43	7,54	13,12	15,45	155,61
Délais avant survenance du sinistre	0%	0,00	2,55	4,99	5,89	8,42	35,33
Montant des prestations	0%	-	144,70	882,00	**3 129,60**	**3 083,70**	495 852,80

Tableau 3.1.3.1 (b) – Statistiques descriptives des variables d'intérêt

B. Algorithmes d'apprentissage automatique en présence de données censurées

 # Algorithme 1 : Algorithme pour le calcul des poids de Kaplan-Meier

Données *:* (Y, δ)

Résultats *: Poids de Kaplan-Meier*

On suppose qu'on a une variable T_i *de durée dont l'observable est :*

- $Y_i = min(T_i, C_i)$
- $\delta = I_{\{T_i \leq C_i\}}$

Pour le calcul des poids de Kaplan-Meier ω_i *:*

$$\omega_{(i),n} = \frac{\delta_i}{n} \prod_{j=1}^{i-1} \left(\frac{n-j}{n-j-1} \right)^{\delta_i}$$

où $\omega_{(i),n}$ *est le poids de Kaplan-Meier associé à la* (i)*-ème valeur* $Y_{(i)}$ *classé par ordre croissant.*

 # Algorithme 2 : Algorithme CART sur données censurées

Données *:* (M, T, X)*, on dispose d'une base composée de* n *individus*

Résultats *: Arbre de régression élagué sur données contenant des observations censurées et prédiction du risque*

Phase 1 : Construction de l'arbre maximal

Etape 0 *: Calculer l''estimateur* $\hat{G}$ *suivant la formule analytique proposée pour les* n *individus*

Etape 1 *: Initialisation*

Considérer l'arbre avec une feuille unique composée den_{nc}*s individus non censurés (*$n_{nc} \leq n$ *)*

Etape s *: split ou subdivision d'un nœud*

Considérer l'arbre obtenu à l'étape $s-1$ *avec* L_{s-1} *feuilles où les individus de chaque feuille* l *correspondant à la classe* $T_l^{(s-1)}$ *sont distincts de ceux des autres feuilles. Les observations censurées (au nombre de* n_c^l *) de mêmes caractéristiques, c'est-à-dire* $\hat{X} \in T_l^{(s-1)}$ *sont assignés à cette feuille.*

Pour chaque feuille l, avec $1 \leq l \leq L_{s-1}$ *:*

Cas 1 *: si* $n_c^l = 1$ *ou si toutes les observations ont les même valeurs que* $\hat{X}$*, alors on ne procède pas à la subdivision (split) de la feuille.*

Cas 2 *: si non la feuille devient un nœud du prochain arbre issu de cette étape : déterminer les valeurs* j_0 *et* $x_l^{(j_0)}$ *qui minimisent* $L_l\left(j, x_l^{(j)}\right)$*, puis définir les deux nouveaux sous-ensembles disjoints :*

$T_l^{(s-1)} \cap \left\{\hat{X}^{(j_0)} \leq x_l^{(j_0)}\right\}$ *et* $T_l^{(s-1)} \cap \left\{\hat{X}^{(j_0)} > x_l^{(j_0)}\right\}$

Le nombre de feuilles devient alors L_s *et on passe à l'étape* $s+1$ *jusqu'à ce que la condition* $L_s = L_{s+1}$ *soit remplie. La procédure s'arrête lorsque toutes les feuilles ne peuvent plus être subdivisées (cas 2)*

Phase 2 : élagage de l'arbre maximal

Sélectionner un sous arbre $\hat{S}(\alpha)$ possédant $\widehat{K}_\alpha$ feuilles, parmi l'ensemble $\aleph$ des sous arbres de l'arbre-maximal (possédant $K_n \leq n$ feuilles) déterminé lors de la phase 1, tel que :

$\hat{S}(\alpha) = \underset{S\epsilon\aleph}{argmin}\left\{\int \varphi\big(m, \hat{\pi}^S(t,x)\big) d\hat{F}(m,t,x) + \frac{\alpha K(S)}{n}\right\}$, *avec : $K(S)$ le nombre total des feuilles du sous arbre S, α facteur de pondération du terme de pénalisation $\frac{K(S)}{n}$.*

***Initialisation** : $\alpha = 0$*

***Incrémentation** : augmenter progressivement la valeur de α telle que $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{K_n}$ jusqu'à ce que $\widehat{K}_{\alpha_{j+1}} = \widehat{K}_{\alpha_j}$*

***Choix de α_{j_0}** : la valeur optimale correspond à la valeur minimisant la formule $\vartheta(\alpha_j)$ pour un échantillon de taille , noté $(N_i, Y_i, \delta_i, X_i)_{n+1\leq i\leq n+\tau}$, suivante :*

$$\vartheta(\alpha_j) = \sum_{i=n+1}^{n+\tau} \frac{\delta_i \varphi\left(N_i, \hat{\pi}^{K(\alpha_j)}(X_i, Y_i)\right)}{1 - \hat{G}(Y_i -)}$$

Calcul de l'estimateur de π_0 :

$$\pi^S(t,x) = \sum_{l=1}^{K(S)} \rho_l R_l(t,x)$$

Pour l une feuille du sous arbre optimal est associé un sous ensemble d'individus T_l (distinct des autres sous ensemble d'individus et tel que la réunion de tous forme l'ensemble T de tous les n individus) et au critère $R_l(\tilde{x}) = I_{\{\tilde{x}\,\epsilon T_l\}}$permettant de déterminer si un individu est affecté ou non à cette feuille. Le coefficient $\rho_l = \underset{\pi\epsilon X}{argmin}\, E\big[\varphi(M,\pi)\big|\tilde{X} \in T_l\big]$.

***Estimation du risque** : utiliser $\hat{\pi}^{S(\alpha)}$ comme l'estimateur final de π_0*

 # Algorithme 3 : Adaptation de l'algorithme de *Random Forest* aux données censurées

Données **:** *M le nombre d'arbres optimaux élagués prédéfini, on dispose d'une base composée de n individus représentés par $z = \{(N_i, Y_i, \delta_i, X_i)_{i=1,\ldots,n}\}$*

Résultats : *Prédiction du risque*

***Etape 0** : Calculer l''estimateur $\hat{G}$ suivant la formule analytique proposée pour les n individus et calculer les poids de Kaplan-Meier*

***Etape 1** : itération de la procédure Tree-base censored*

Pour $m = 1$ à M faire

- *un tirage aléatoire dans z d'un échantillon boostrap (avec remise) noté $z_m^{(b)}$*
- *estimer l'arbre de régression optimal élagué $\hat{f}_{z_m^{(b)}}$ avec l'algorithme Tree-base censored*

***Etape 2** : agrégation des modèles*

Calculer $\hat{f}_{bag}(x) = \frac{1}{m}\sum_{m=1}^{M} \hat{f}_{z_m^{(b)}}(x)$, puis utiliser $\hat{f}_{bag}$ comme l'estimateur final de π_0

 # Algorithme 4 : Adaptation de l'algorithme de *Gradient Tree Boosting* aux données censurées

Données **:** *M le nombre d'arbres optimaux élagués prédéfini, on dispose d'une base composée de n individus représentés par $z = \{(N_i, Y_i, \delta_i, X_i)_{i=1,\dots,n}\}$*

Résultats : *Prédiction du risque*

Etape 0 *: Calculer l''estimateur $\hat{G}$ suivant la formule analytique proposée pour les n individus et calculer les poids de Kaplan-Meier*

Etape 1 *: Initialisation : $\hat{f}_0 = \underset{\gamma}{argmin} \sum_{i=1}^{n} l^{(p)}(y_i, \gamma)$*

Début de la boucle :

Pour $m = 1$ à M faire

- Calculer $r_i^m = -\left[\frac{\partial l^{(p)}(y_i, f(x_i))}{\partial f(x_i)}\right]_{f=f_{m-1}}$ pour $i = 1, \dots, n$

- Ajuster l'arbre de régression optimal élagué δ_m au couple $(x_i, r_i^m)_{i=1,\dots,n}$ avec l'algorithme Tree-base censored

- Calculer γ_m en résolvant :

$$\min_{\gamma} \sum_{i=1}^{n} l^{(p)}\left(y_i, f_{m-1}(x_i) - \gamma \delta_m(x_i)\right)$$

- Mise à jour :

$$\hat{f}_m(x) = \hat{f}_{m-1}(x) - \gamma_m \delta_m(x)$$

Fin de la boucle

Estimation du risque : *utiliser $\hat{f}_M$ comme l'estimateur final de π_0*

Chapitre 2 :
Approche quantitative d'estimation de chocs des risques d'incidence et de maintien en assurance non vie sous la norme Solvabilité II

CHAPITRE 2 : APPROCHE QUANTITATIVE D'ESTIMATION DE CHOCS DES RISQUES D'INCIDENCE ET DE MAINTIEN EN ASSURANCE NON VIE SOUS LA NORME SOLVABILITE II

1. Introduction

Depuis le 1er janvier 2016, les organismes d'assurances et de réassurance de la zone union européenne appliquent la formule standard de la réglementation Solvabilité II pour l'évaluation de leurs besoins de fonds propres économiques, tel qu'il est promulgué dans les textes de lois européens et les mesures d'implémentation.

Le pilier I du nouveau dispositif réglementaire Solvabilité II, impose à ces organismes l'utilisation de la formule standard. Cette dernière repose sur l'agrégation de besoins en capitaux élémentaires à partir de matrices de corrélation. Ce montant de capitaux permet à la compagnie de faire face à une ruine économique à horizon 1 an et avec un temps de retour de 200 ans. Ces compagnies peuvent également développer un modèle interne, mieux adapté à leurs profils de risque propres, afin de mesurer le capital réglementaire de solvabilité ajusté aux risques encourus. Une approche alternative consiste à mettre en place un modèle interne partiel permettant de déroger à la formule standard pour certains de ses modules de risques et matrices de corrélations.

Dans ce même contexte s'agissant du calcul du besoin de fonds propres lié au risque de réserve et de prime du module de souscription non-vie et sous certaines conditions, *l'European Insurance and Occupational Pensions Authority* (EIOPA) laisse la possibilité aux compagnies d'utiliser leurs propres données et expériences pour le calibrage des paramètres spécifiques à leurs entités. Il s'agit des *Undertaking Specific Parameters* également connu sous le nom d'USP non-vie (*Undertaking Specific Parameters*, cf. 7. de l'article 104 de la directive Solvabilité II et la section SCR.10 des spécifications techniques du QIS5).

Les méthodes de calcul de ces paramètres spécifiques doivent s'adapter aux nouvelles problématiques de Solvabilité II, comme l'estimation des réserves à un horizon d'un an plutôt qu'à l'ultime. Ainsi, pour les compagnies disposant de données de qualité, de volumétrie et d'historique suffisantes, les textes prévoient, par exemple pour les risques non-vie, l'intégration de l'expérience du portefeuille dans les paramètres de calcul du SCR (*Solvency Capital Requirement*). L'utilisation de ces paramètres spécifiques est autorisée pour :

- Les risques de tarification (ou *premium risk*) et de provisionnement (ou *reserve risk*) des modules d'assurance non-vie et de santé non assimilée à la vie (*Health NSLT*).
- Le risque de révision des modules assurance vie et de santé assimilée à la vie.
- Le facteur d'ajustement dans le cadre de la réassurance non proportionnelle.

L'autorisation de l'utilisation totale ou partielle de ces paramètres spécifiques par le superviseur local est conditionnée au dépôt d'un dossier. Parmi les éléments du dossier de candidature, le superviseur attend une note de justifiant, segment par segment, que le choix de l'organisme ou du groupe d'utiliser ou non des USP ou GSP (*Group Specifics Parameters*) et les méthodes associées sont pertinents eu égard au profil de risque de l'organisme ou du groupe, y compris pour les segments non concernés par la candidature. Cette demande peut se comprendre comme une précaution supplémentaire souhaitée par le superviseur pour limiter une démarche d'optimisation des fonds propres sur certains périmètres ou l'ensemble du business de la part des acteurs.

Ces contraintes soulèvent les épineuses questions de la formalisation, de l'argumentation et de la justification des choix méthodologiques retenus, ainsi que du domaine d'application de l'approche proposée par les experts de l'entreprise. Il s'agit donc d'une démarche nécessitant rigueur et pédagogie attendue par le Conseil d'Administration ou l'Autorité de Contrôle Prudentiel.

Parmi les nombreux travaux académiques, mémoires et publications réalisés, plusieurs contributions apportent des éclairages sur les démarches d'implémentation de modèles internes et des USP non-vie proposés par l'EIOPA. Perrin (2012) a mené des études pour tester les différentes méthodes de calcul des USP proposées par l'EIOPA sur des portefeuilles d'assurance non-vie des branches *Marine Aviation and Transport* et *Other motor* et de la branche *Third-party liability*. Ces analyses l'on conduit à proposer une alternative aux méthodes de calcul du risque de réserve. Cerchiara et al. (2014) font une comparaison entre les paramètres USP non-vie proposées par l'EIOPA et les paramètres issus d'un modèle interne partiel pour le risque de prime.

Planchet et al. (2010) proposent un cadre général pour construire un modèle interne ou un modèle interne partiel dans un contexte d'assurance de personnes. Cette contribution consiste à montrer que dans ce contexte, du fait du caractère gaussien des engagements conditionnellement aux facteurs de risque systématique, il est possible d'obtenir des expressions des valeurs de référence (*best estimate*, marge pour risque et SCR) dans le cadre d'une approche mêlant calculs analytiques et simulation. Cambon (2011) propose une démarche de construction d'un modèle interne partiel pour le risque de souscription non-vie. Il applique son approche aux calculs des besoins de capitaux requis d'une société spécialisée dans les branches longues comme l'assurance construction afin tenir compte des spécificités propres à ces risques. En considérant un horizon d'un an, la variabilité entre le montant cumulé ultime calculé d'une année à l'autre est déterminée par une approche stochastique basée sur la méthode du *Booststrap*. Finalement, la différence entre la charge ultime vue à l'exercice N+1 et celle vue à l'exercice N permet de déterminer le besoin en capital relatif au risque de prime et de réserve. La répétition du processus un grand nombre de fois permet d'obtenir une distribution représentative des risques de prime et de réserve. Le besoin en capital relatif au risque de souscription non vie sera obtenu en prenant le quantile à 99.5% de la distribution de la différence entre la charge ultime vue à fin N+1 et vue à fin N ainsi obtenue.

Aguir (2012) étudie dans le cadre de ses travaux de mémoire d'actuariat l'impact de l'utilisation d'une approche « modèle interne partiel » pour la mesure du capital réglementaire sur le périmètre de l'assurance des emprunteurs et sur le risque de souscription. Elle retient l'approche « formule standard » pour l'ensemble des risques sauf pour le risque de perte d'emploi. En effet, le risque perte d'emploi ne bénéficiant pas d'une place bien définie avec cette approche. Ainsi, elle propose une méthode alternative à la formule standard pour quantifier ce risque à partir de la calibration du choc perte d'emploi. Cette calibration se fait à partir de la calibration d'un scénario central et d'un scénario adverse pour le calcul de la *Value-At-Risk* à 99.5%. Elle fait le constat que le risque associé aux garanties de perte d'emploi a une forte composante économique comme la crise de 2008 l'a démontrée. Ce risque est très sensible aux différentes crises qui peuvent survenir et affecter l'emploi. C'est pourquoi, un scénario adverse cohérent avec une *Value-At-Risk* à 99.5% à horizon 1 an pour le calcul du SCR est d'abord recherché. Puis, les effets de ce scénario sont traduits sur le portefeuille étudié en termes de taux d'entrée et de loi de maintien utilisés dans l'approche modèle interne partiel. Plusieurs modélisations du chômage comme le modèle d'Alghrim, les séries temporelles via les modèles ARIMA, ainsi que la théorie des valeurs extrêmes ont été étudiées. Le processus finalement retenu pour modéliser le nombre de chômeurs est un processus ARIMA. D'après cette étude, les processus ARIMA sont les plus adaptés pour ce risque et ce notamment du fait de leur propriété de stationnarité. Ainsi, grâce à leur robustesse statistique, ces modèles ARIMA permettent de modéliser le chômage en utilisant des lois à queue lourde pour modéliser les résidus. Les données utilisées dans cette étude sont des données de l'INSEE, car c'est le seul organisme à disposer d'un échantillon significatif d'années d'occurrences. L'historique utilisé (entre 1975 et 2010) permet de capter les valeurs extrêmes.

L'objectif principal de notre contribution est de mener une réflexion méthodologique afin de proposer une démarche de calibrage des paramètres spécifiques (chocs USP) du risque de souscription sous-jacent à la garantie arrêt de travail d'un contrat de prêts (immobilier et consommation) d'un assureur vie disposant d'un portefeuille de plus de 12 millions d'assurés, composé de génération de souscription de différente volumétrie de données et de profondeur d'historique. Nous proposons un cadre opérationnel d'évaluation des chocs USP des risques d'incidence et du maintien (ou de rétablissement) en incapacité et en invalidité du risque arrêt de travail suivant les hypothèses de solvabilité requises par la norme Solvabilité II.

Les risques d'incidence et de maintien (ou de rétablissement) associés à cette garantie sont classés dans le module Santé assimilable à la vie (*Health SLT*) de la formule standard. Or ce module ne bénéficie pas de méthodologie de recalibrage des paramètres spécifiques à l'entité contrairement au module Santé non assimilable à la vie (*Health NSLT*) pour lequel les USP non-vie sont applicables. De plus, il apparaît que le module *Health SLT* recouvre une hétérogénéité importante de situations (contrat collectif/contrat individuel, avec sélection à l'entrée/sans sélection, emprunteur/prévoyance, duration très variable des engagements). Ainsi, un paramétrage de chocs forfaitaires comme proposé par la formule standard paraît inadapté pour capter la réalité économique des risques sous-jacents. A l'instar des USP non-vie, notre approche permet de déterminer des paramètres de chocs spécifiques à l'entité associés aux risques considérés pour un portefeuille de prévoyance classé dans le module *Health SLT*, tenant compte de l'expérience et des données disponibles propres à la compagnie et candidats pouvant remplacer les chocs proposés par la formule standard.

La suite de l'article est organisée de la manière suivante. En section 2, nous rappelons les références règlementaires ainsi que le contenu du dossier de candidature attendu par l'Autorité de Contrôle Prudentiel. La section 3 est consacrée à la description des données du portefeuille et du risque étudié. La section 4, présente le cadre méthodologique. La section 5 fournit les applications des modèles proposés, de la démarche de calibrage jusqu'aux résultats obtenus. Dans la section 6, nous présentons des éléments de discussion et de comparaison. Nous finirons par une conclusion à la section 7.

2. Références règlementaires et dossier de candidature

Présentation de la mesure

Pour le calcul de leur exigence de capital réglementaire, les organismes ou groupes d'assurances peuvent dans certains cas utiliser des paramètres qui leur sont propres en remplacement des paramètres de la formule standard. Ces paramètres sont appelés couramment USP *(Undertaking specific parameters* en anglais), ou bien GSP (*Group specific parameters*) quand ils sont appliqués au niveau d'un groupe pour le calcul de l'exigence de capital sur base combinée ou consolidée.

Les paramètres pouvant ainsi être personnalisés sont définis par la règlementation ; ils ne portent que sur le calcul de certains risques de souscription, à la maille du segment comprenant pour chaque ligne d'activité visée les affaires directes et acceptées proportionnellement.

L'utilisation par un organisme ou groupe d'assurance de paramètres spécifiques pour le calcul de l'exigence en capital est soumise à autorisation préalable de l'ACPR.

Références réglementaires

L'utilisation de paramètres spécifiques est définie, pour les organismes et pour les groupes, respectivement au V de l'article R. 352-5 et au neuvième alinéa du II de l'article R. 356-19 du code des assurances, applicables aux organismes et groupes relevant des trois codes, qui transposent l'article 104

(7) de la Directive 2009/138/CE, dite « Solvabilité II ». Ces dispositions sont complétées par les articles 218 à 220 et l'annexe XVII, ainsi que l'article 338 pour les groupes, du règlement délégué (UE) n°2015/35 dit « niveau 2 ».

Enfin, le contenu du dossier de candidature et les différentes étapes de la procédure d'approbation sont précisés par le règlement d'exécution (UE) 2015/498 établissant des normes techniques d'exécution ou « ITS » en ce qui concerne la procédure d'approbation par les Autorités de contrôle de l'utilisation de paramètres propres à l'entreprise.

Contenu du dossier de candidature

Le contenu minimum du dossier de candidature est défini à l'article 1 du règlement d'exécution mentionné ci-avant.

De manière synthétique, le dossier devra contenir les éléments suivants :

- Une lettre formelle de candidature précisant la date d'utilisation souhaitée, la liste exhaustive des paramètres que l'organisme souhaite modifier dans la formule standard avec leur nouvelle valeur à la date d'utilisation, et la méthode utilisée lorsqu'il s'agit d'un écart-type de risque de provisions.
- Une note justifiant, segment par segment, que le choix de l'organisme ou du groupe d'utiliser ou non des USP ou GSP et les méthodes associées est pertinent eu égard au profil de risque de l'organisme ou du groupe, y compris pour les segments non concernés par la candidature.
- Les fichiers de calcul permettant de reproduire le calcul contenant, paramètre par paramètre, les données utilisées, les démarches de calculs effectuées conformément aux méthodes réglementaires, et la valeur des nouveaux paramètres obtenus après calcul.
- La documentation démontrant la satisfaction des exigences de qualité des données stipulées aux articles 19 et 219 du règlement délégué ainsi qu'à son annexe XVII. En particulier, le dossier de candidature devra comporter :
    - Les éléments du point e) de l'article 219 du règlement délégué intégrant un répertoire des données listées à l'annexe XVII.
    - Une description de la gouvernance autour de la qualité des données (comitologie éventuelle, dispositif de contrôle interne, processus de validation, etc.).
    - Les plans d'amélioration de la qualité des données dont l'évaluation n'a pas atteint les critères requis.

Une fois le dossier déposé, l'ACPR se prononcera sur l'autorisation de l'utilisation totale ou partielle de paramètres spécifiques dans un délai maximal de 6 mois à compter de la date de réception du dossier complet.

Conséquences pour le groupe

L'utilisation de paramètres spécifiques pour le calcul de la solvabilité au niveau individuel par un organisme et sur base consolidée ou combinée par un groupe, sont deux mesures distinctes et indépendantes sur le plan administratif.

Un groupe calculant son exigence de capital réglementaire sur la base des comptes consolidés ou combinés ne peut tenir compte d'aucun paramètre spécifique approuvé au niveau d'un organisme pris individuellement. Si le groupe souhaite lui aussi utiliser des paramètres spécifiques, il est tenu d'en solliciter préalablement l'autorisation à l'ACPR, en tenant compte de son profil de risque pour l'ensemble de son périmètre, selon la même procédure que pour un organisme pris individuellement. De

même, l'autorisation accordée à un groupe d'utiliser des paramètres spécifiques pour le calcul de sa solvabilité ajustée ne vaut pas autorisation pour les organismes membres du groupe pris individuellement.

3. Description des données

Nos travaux ont porté sur un périmètre constitué de portefeuilles de contrats de couverture de prêt d'une compagnie d'assurances distribués par plusieurs partenaires. Ces contrats proposent plusieurs garanties dont une garantie couvrant les assurés contre le risque d'arrêt de travail, que ce soit en cas d'incapacité ou d'invalidité.

Notre étude porte sur un portefeuille d'assurances des emprunteurs comprenant essentiellement des contrats de prêts immobiliers et représentatif du marché des prêts immobiliers français. Les données proviennent d'un infocentre de données qualifiées et certifiées issues des systèmes de gestion opérant. Nous disposons de plus de 12 millions de polices, l'âge moyen de souscription est de 40 ans, pour un capital assuré moyen de 48 000,00 € et un montant de sinistre moyen de 12 000,00 €. Les garanties principales sont l'arrêt de travail et le décès PTIA (Perte Totale et Irréversible d'Autonomie).

A partir de cet infocentre nous disposons :

- De données détaillées constituées de bases tête par tête des assurés (de profondeur d'historique variable suivant les partenaires, certains disposant d'une profondeur de moins de 5 ans).
- De bases de données agrégées par tranche d'âges des assurés (en général d'historique plus profond).
- De données détaillées tête par tête des sinistrés avec une profondeur d'historique conséquente (le contrat le plus ancien dispose d'un historique de plus de 20 ans).
- D'une volumétrie importante de données sur les assurés et les sinistrés.

Bases de données disponibles	Taux d'incidence	Choc d'incidence USP	Taux de rétablissement	Choc de rétablissement USP
Bases des assurés « tête par tête » *2 à 5 années d'historique*	✓			
Bases des assurés « agrégées » *16 années d'historique*		✓		
Bases des sinistrés « tête par tête » *23 années d'historique*	✓	✓	✓	✓

Tableau 3 : Structure des bases de données et historiques retenus pour l'étude

Les données contenues dans les bases d'étude respectent les principaux critères de qualité et ont été validées par un processus de qualification des données.

La principale limite de bases de données disponibles concerne l'identification de chaque assuré d'une base à l'autre. En effet, les bases des assurés sont transmises par les partenaires bancaires à la compagnie et les bases des sinistrés sont gérées au sein de la compagnie par le service de gestion intégrée. De ce fait, la clé d'identification des assurés est différente d'une base à l'autre.

Comme nous l'avons souligné ci-dessus, le peu de profondeur de l'historique est une deuxième limite pour le calibrage des chocs associés au risque d'incidence. En effet, les modèles proposés en vue de

l'évaluation des erreurs d'estimation et de *process* variance (voir paragraphe III), nécessitent de disposer d'un historique de données suffisamment profond pour bien capter les variations et les évolutions du risque dans le temps. Pour pallier à cette limite, une approche par crédibilité a été mise en œuvre.

De plus, pour pallier l'insuffisance de profondeur d'historique, les alternatives proposées sont :

- Soit conserver les paramètres de la formule standard.
- Soit regrouper ce portefeuille avec un autre de même structure plus conséquent, dès lors que nous savons prouver que nous sommes en présence de risques homogènes.

Enfin, notons que le risque de maintien en arrêt de travail n'a pas posé de difficulté particulière aussi bien pour estimer la loi de maintien et de rétablissement que pour calibrer les chocs associés. Les modélisations s'effectuent à partir de la seule base des sinistrés, qui de plus est détaillée et d'historique très profond.

4. Cadre méthodologique et modèles

4.1. Définitions, notation et hypothèses

Le risque d'incidence correspond à l'incertitude sur la probabilité d'entrer en état d'incapacité (ou d'invalidité) et le risque de maintien (ou de rétablissement) est l'incertitude sur la durée de l'incapacité (ou d'invalidité).

Nous proposons une démarche de calcul des niveaux de chocs pour chacun de ces risques en intégrant à la fois l'*erreur d'estimation* et la *process variance*, considérées comme principales sources de volatilité des risques considérés.

L'erreur d'estimation, correspondant à l'écart entre la loi estimée et la loi intrinsèque, et qui traduit notamment les fluctuations d'échantillonnage. La *process variance* ou *erreur de process* correspond à la fluctuation temporelle des lois de sinistralité liée à des conditions endogènes ou exogènes. Nous pouvons noter comme sources d'évolutions du risque :

- des conditions internes de souscription et de sélection médicale.
- des conditions d'indemnisation des sinistres,
- de l'état de santé générale,
- des règles externes et légales dans la mesure où la définition des prestations et de l'intervention de l'assureur est alignée sur la reconnaissance de l'état d'incapacité de l'assuré par un organisme externe comme la Sécurité Sociale en France,
- de la structure de la population assurée (reprise de portefeuille, effets économiques, changement de cible de distribution, campagne de commercialisation spécifique…),
- …

Nous admettons dans notre approche, que ni la loi d'incidence, ni la loi de maintien ne distinguent l'état d'incapacité de celui d'invalidité. Sans perdre de généralité, cette hypothèse nous permet de ne pas tenir compte de manière spécifique d'une loi de passage de l'état d'incapacité à l'état d'invalidité. Notre mémoire de recherche ne traitera donc pas le calibrage des chocs associés à ces lois de passage. Nous considérons donc les trois différentes causes de sortie, supposées indépendantes, suivantes : le rétablissement, le décès et la fin contractuelle de la garantie (ou de la couverture).

Dans la suite du document, l'appellation « segment » fait référence à la maille de calibrage du choc. Dans le cas général, il s'agit de la ligne d'affaires ou LoB (*Line Of Business*). Cependant il apparaît que la LoB « *Health SLT* » recouvre une variété importante de garanties et de conditions d'assurance. Ces

différences génèrent des risques sous-jacents très hétérogènes. Il pourra en conséquence être envisagé de considérer des sous-segments de cette LoB afin de modéliser des niveaux de chocs spécifiques. Cela suppose que le profil des risques ainsi regroupés soit jugé homogène au sein de chaque sous-segment.

D'autre part, dans la mesure où l'on peut démontrer par des études statistiques et sur la base de jugements d'experts que l'évolution intrinsèque du risque entre deux sous-segments est similaire, l'erreur de *process* calibrée sur l'un des sous-segments peut être utilisée pour déduire celle de l'autre par l'intermédiaire d'une fonction de lien appropriée. En pratique ce cas de figure se rencontre lorsqu'un sous-segment ne dispose pas d'un historique profond ou d'une volumétrie suffisante de données. Cette précaution est nécessaire pour préserver une certaine robustesse dans l'estimation des paramètres spécifiques.

Nous ne traitons pas non plus des effets de diversification éventuels entre les sous-segments. Ainsi, une approche totalement comonotone est retenue pour l'agrégation des niveaux de volatilité calibrés sur les différents sous-segments de la LoB.

Pour un LoB noté *s* (ou un sous-segment), nous analyserons les évolutions sur l'historique de la période d'observation retenue :

- des taux d'incidence (en incapacité ou invalidité) noté $\mathrm{i}(\mathrm{x},\mathrm{s},\mathrm{t})$ en fonction de l'âge x (ou tranche d'âges) atteint et de la période d'estimation t,
- des taux de maintien (respectivement des taux de rétablissement) noté $\mathrm{m}(\mathrm{x},\mathrm{a},\mathrm{s},\mathrm{t})$ (respectivement $\mathrm{r}(\mathrm{x},\mathrm{a},\mathrm{s},\mathrm{t})$) en fonction de l'âge x (ou tranche d'âges) à l'entrée dans l'état, de l'ancienneté atteinte a et de la période t .

D'un point de vue pratique, il conviendrait de vérifier avec les tests statistiques appropriés que les erreurs d'estimation peuvent être considérées comme décorrélées entre elles. Nous supposerons également que les erreurs d'estimation sont décorélées de toutes les autres sources d'erreurs, et en particulier de l'erreur de *process*.

Afin de proposer une estimation acceptable, nous aurons recours à deux types d'approches pour le calcul de ces paramètres spécifiques :

- une modélisation basée sur les séries chronologiques de type ARMA(p,q),
- et un modèle linéaire généralisé (ou GLM) de type un modèle Log-normal.

En plus de l'approche GLM, le choix des séries chronologiques est fondésur leurs applications aux nombreux domaines comme la prévision d'indices économiques en économie, l'évolution des cours de la bourse en finance, l'analyse de l'évolution d'une population en démographie, l'analyse de données sismiques en géophysique. Ces modèles s'adaptent en général bien à la problématique posée. Il conviendrait, suivant le niveau de prudence souhaitée dans les calculs, d'estimer l'erreur d'estimation comme étant la variance de l'estimateur de la loi moyenne retenue. Cette précaution sur l'évaluation de l'erreur d'estimation est particulièrement nécessaire dans le cas où l'approche GLM est retenue.

Le choix de la série temporelle à retenir résultera en toute rigueur des résultats des tests menés sur les données du portefeuille. Parmi les processus ARMA (*AutoRegressive Moving-Average*) testés sur les données du portefeuille utilisé, il apparaît que le processus AR(1) est celui qui revêt un caractère prudent. En effet, parmi les processus ARMA(p,q) qui minimisent le critère d'information Akaïke (AIC), l'AR(1) est celui qui proposait les niveaux de chocs les plus conservateurs. Des tests d'adéquation du modèle fondés sur le caractère « bruit blanc » des résidus, ont par ailleurs permis de conforter ce choix. Ainsi, dans la suite de notre exposé, nous retiendrons le processus AR(1) comme modèle de référence.

L'une des principales difficultés de calibrage des paramètres des modèles retenus est la disponibilité de données de taille et d'historique suffisants. Ces données doivent également répondre aux critères de qualité (exhaustivité, exactitude, pertinence) préconisés par la norme Solvabilité II. Pour pallier cette problématique de profondeur d'historique, nous proposons une démarche basée sur la théorie

de crédibilité. Nous retiendrons un cadre similaire et ajusté au profil des risques étudiés de la théorie de la crédibilité à variation limité (TCVL) ou crédibilité américaine proposée dans la note éducative de l'Institut Canadien des Actuaires de juillet 2012[6]. Nous développons dans le paragraphe suivant cette méthodologie.

Ainsi, de façon générale que ce soit pour le risque d'incidence ou de maintien, la démarche proposée s'appuie sur l'estimation des écarts-types associés aux deux sources d'erreurs (estimation et *process*). Elle repose sur la modélisation respective des taux d'évolution des taux d'incidence (de maintien ou de rétablissement) pour l'erreur *process* et la détermination de la variance de l'estimateur retenu pour le calcul des taux d'incidence (de maintien ou de rétablissement) pour l'erreur d'estimation. Les modèles présentés permettent de projeter l'évolution temporelle des taux d'incidence (de maintien ou de rétablissement) à horizon d'un an (t+1), puis d'en déduire l'espérance, la variance et le coefficient de variation associé à cette distribution. Le choc marginal correspondant à l'erreur de *process* est approché par le produit du coefficient de variation estimé à t+1 et du taux d'incidence (de maintien ou de rétablissement) calculé à t.

Le choc marginal correspondant à l'erreur d'estimation est défini comme la variance de l'estimateur des taux d'incidence (de maintien ou de rétablissement) estimé à date t. Sous les hypothèses que les erreurs d'estimation peuvent être supposées décorrélées entre elles et de toute autre source d'aléa, le niveau de choc global associé au risque d'incidence (de maintien ou de rétablissement) est obtenu par agrégation des chocs de l'erreur de *process* et de l'erreur d'estimation. Enfin, de façon similaire à la formule standard, l'approche proposée est basée sur le principe d'application d'un choc conjoint aux risques d'incidence et de maintien (ou rétablissement).

Le choc global obtenu pour chaque risque après intégration des facteurs de crédibilité sera comparé aux chocs de la formule standard dont nous rappelons les paramètres dans le tableau 1 ci-dessous.

RISQUE	FORMULE STANDARD	PROPOSITION USP
Incidence en incapacité ou invalidité	Année de projection t=1 : 35% Année de projection t>1 : 25%	$\Delta_{(Inc,s,USP)}$
Maintien en incapacité ou invalidité	Si taux de maintien < 50% : augmentation de 20% des taux de maintien Si taux de maintien > 50% : diminution de 20% des taux de rétablissement	$\Delta_{(recov,s,USP)}$

Tableau 1 : Récapitulatif des chocs Formule Standard

4.2. Modèles de calibration des niveaux de chocs spécifiques

Modélisation des chocs des lois d'incidence

Disposant de données répondant au critère de profondeur d'historique mentionné ci-dessus, pour chaque date $t \epsilon [t_{min}; t_{max}]$ nous pouvons calibrer sur les périodes $[t - T; t]$ (où $T \geq 1$ est l'amplitude des fenêtres d'estimation) des lois d'incidence d'expérience. Nous obtenons ainsi pour tout x donné, la série des taux d'incidence d'expérience $(i(x, s, t))_t$.

[6] CANADIAN INSTITUTE OF ACTUARIES (2002), Commission des rapports financiers des compagnies d'assurances vie, Note éducative, mortalité prévue : polices canadiennes d'assurances vie individuelle avec tarification complète

Le risque étant à la hausse des taux d'incidence, nous présentons dans ce qui suit les méthodologies permettant de disposer des chocs d'incidence au quantile à 99.5% en intégrant les deux types d'erreur décrites précédemment.

Modélisation de l'erreur d'estimation

Suivant la structure et la précision des données disponibles, plusieurs estimateurs sont disponibles (Kaplan-Meier, actuariel, binomial, des moments de Hoem, quotient partiel) chacun avec ses limites, voir Planchet et al. (2011). Nous retenons comme estimateur des taux d'incidence l'estimateur actuariel introduit par Bömer (1912). Cet estimateur permet d'utiliser des données agrégées et tient compte des censures et des troncatures. Il est plus simple à mettre en œuvre que l'estimateur de Kaplan-Meier. De plus, les censures et les troncatures sont supposées indépendantes et uniformément réparties dans l'intervalle.

On fait l'hypothèse que le nombre d'occurrences d'entrée en incapacité/invalidité à l'âge x suit une loi binomiale dont les paramètres sont le nombre d'individus d'âge x valides en début de période et le taux d'incidence correspondant. Enfin, si cet estimateur est adapté aux échantillons de grande taille, l'utilisation de données agrégées conduit à s'interroger sur son adéquation aux données individuelles et sur sa précision.

a. Estimation des taux d'incidence par l'estimateur actuariel

Dans ce paragraphe, nous rappelons la formule de l'estimateur actuariel utilisée pour calculer le taux d'incidence dit central $i(x,s,t)$. Le taux d'incidence sur le segment s, à l'âge ou tranche d'âge x et à la date t (année ou tranche d'années) est donné par

$$i(x,s,t) = \frac{Inc(x,s,t)}{Er(x,s,t)},$$

où $Inc(x,s,t)$ correspond au nombre d'entrées en incapacité/invalidité observé pour le segment s des individus d'âge x durant l'année t, $Er(x,s,t)$ est l'exposition au risque pour le segment s des individus d'âge x durant à la date t.

b. Estimation de l'erreur d'estimation

L'erreur d'estimation est estimée pour le segment s en déterminant la variance de l'estimateur utilisé pour calibrer le taux d'incidence central $i(x,s,t)$. Nous notons $\sigma_{estim,s}$ l'écart-type associé à cette erreur.

Dans le cas où les taux d'incidence sont calibrés en utilisant l'estimateur actuariel, l'erreur d'estimation correspond à la variance de l'estimateur actuariel.

Sous l'hypothèse que le nombre d'occurrences d'entrée en incapacité/invalidité suit une loi binomiale de la forme $\sum_{t=t_{max}-T+1}^{t_{max}} Inc(x,s,t) \sim Bin\left(Er(x,s,t_{max}), i(x,s,t_{max})\right)$ à la date t_{max} , nous obtenons :

$$\sigma^2_{estim,s} = \frac{i(x,s,t_{max})\times[1-i(x,s,t_{max})]}{Er(x,s,t_{max})}.$$

Modélisation de l'erreur de process

Nous recherchons à capter l'erreur de *process*. Pour cela, il conviendrait dans certains cas de réaliser des regroupements suivant une maille d'agrégation s' contenant suffisamment d'individus afin de minimiser l'erreur d'estimation dans le calcul des taux d'incidence. Les calculs pourraient par exemple être réalisés à une maille tout sexe confondu, regroupant des tranches d'âges, des portefeuilles du même segment, dès lors que le sous-segment constitué est composé de risques homogènes. Ainsi, disposant des taux d'incidence $\mathrm{i}(\mathrm{x},\mathrm{s},\mathrm{t})$ calculés précédemment, on en déduira par exemple les calculs des taux d'incidence agrégés à la maille s' noté $\mathrm{i}_{(\mathrm{s}',\mathrm{t})}$ en faisant une moyenne pondérée par les expositions moyennes (expositions moyennes calculées sur tout l'historique de calibrage).

a. Approche par modélisation AR(1)

Cette méthodologie repose sur le suivi de la variation des lois d'incidence d'expérience afin d'étudier l'évolution du risque dans le temps. Considérons la série chronologique des taux d'évolution des taux d'incidence $\left(\mathrm{W}_{(\mathrm{s}',\mathrm{t})}\right)_{\mathrm{t}}$ pour un sous-segment s' donné du segment s, où $\mathrm{W}_{(\mathrm{s}',\mathrm{t})}$ est défini de la manière suivante :

$$\mathrm{W}_{(\mathrm{s}',\mathrm{t})} = \frac{\mathrm{i}_{(\mathrm{s}',\mathrm{t}+1)} - \mathrm{i}_{(\mathrm{s}',\mathrm{t})}}{\mathrm{i}_{(\mathrm{s}',\mathrm{t})}}.$$

Cette variable est modélisée par un processus autorégressif d'ordre 1, dont la formule mathématique est

$$W_{(s',t)} = \alpha_{(s')} + \beta_{(s')} \times W_{(s',t-1)} + \sigma_{(s')} \times \varepsilon_{(s',t)},$$

où $(\varepsilon_{(s',t)})_{s',t}$ est un ensemble de variables aléatoires gaussiennes centrées réduites indépendantes. Les coefficients $\alpha_{(s')}$ et $\beta_{(s')}$ sont estimés à l'aide de l'estimateur des moindres carrés ordinaires (ou pondérés par les expositions).

Le facteur $\sigma_{(s')}$ correspond à l'écart-type empirique des résidus. Ces résidus notés $r_{(s',t)}$ sont obtenus par différence entre les espérances conditionnelles théoriques des variables $W_{(s',t)}$ et leurs réalisations :

$$r_{(s',t)} = W_{(s',t)} - E\left\{W_{(s',t)} \middle| W_{(s',t-1)}\right\},$$

soit

$$r_{(s',t)} = W_{(s',t)} - \left(\alpha_{(s')} + \beta_{(s')} \times W_{(s',t-1)}\right).$$

Considérant cette modélisation, nous obtenons la relation suivante

$$i_{(s',t_{\max}+1)} = i_{(s',t_{\max})} \times (1 + W_{(s',t_{\max})}),$$

soit

$$i_{(s',t_{\max}+1)} = i_{(s',t_{\max})} \times \left(1 + \alpha_{(s')} + \beta_{(s')} \times W_{(s',t_{\max}-1)} + \sigma_{(s')} \times \varepsilon_{(s',t_{\max})}\right)$$

où $\varepsilon_{(\mathrm{s}',\mathrm{t}_{\max})} \sim \mathrm{N}(0,1)$ suit une loi normale centrée réduite.

On en déduit

$$\mathrm{var}\left\{i_{(s',t_{\max}+1)}\right\} = i_{(s',t_{\max})}{}^{2} \times \sigma_{(s')}{}^{2}$$

et

$$E\left\{i_{(s',t_{\max}+1)}\right\} = i_{(s',t_{\max})} \times (1 + \alpha_{(s')} + \beta_{(s')} \times W_{(s',t_{\max}-1)})$$

où var{X} la variance de la variable aléatoire X et E{X} l'espérance de la variable aléatoireX.
Le coefficient de variation associé à l'erreur de process $CV_{(s')}$ se définit de la façon suivante

$$CV_{(s')} = \frac{\sqrt{\mathrm{var}\left\{i_{(s',t_{\max}+1)}\right\}}}{E\left\{i_{(s',t_{\max}+1)}\right\}},$$

d'où

$$CV_{(s')} = \frac{i_{(s',t_{\max})} \times \sigma_{(s')}}{i_{(s',t_{\max})} \times (1 + \alpha_{(s')} + \beta_{(s')} \times W_{(s',t_{\max}-1)})}.$$

A partir de ce coefficient de variation, nous pouvons estimer sur le segment s et à la date $\mathrm{t_{max}}$, l'écart-type associé à l'erreur de *process* sur les taux d'incidence en état d'incapacité/invalidité $\mathrm{i(x, s, t_{max})}$ de la façon suivante:

$$\sigma_{(process,s)} = CV_{(s')} \times i(x, s, t_{max}) .$$

b. Approche par modélisation log-normale

La seconde approche proposée consiste à mesurer l'erreur *process* du risque d'incidence par une modélisation log-normale de la série des taux d'incidence. Dans ce cas, pour un segment s' donné du segment s, nous supposons que le taux d'incidence suit la dynamique

$$\ln\left(\frac{i_{(s',t)}}{i_{(s',t-1)}}\right) = \alpha_{(s')} + \sigma_{(s')} \times \varepsilon_{(s',t)}$$

où $\varepsilon_{(s',t)}$ est distribué selon une loi normale centrée réduite (notée $\mathrm{N(0,1)}$) et le coefficient $\alpha_{(s')}$est estimé à l'aide de l'estimateur des moindres carrés ordinaires (ou pondérés par les expositions).

Le facteur $\sigma_{(s')}$ correspond à l'écart-type empirique des résidus. Ces résidus $r_{(s',t)}$ sont ensuite calculés par différence entre les espérances théoriques (conditionnelles) de

$$\ln\left(\frac{i_{(s',t)}}{i_{(s',t-1)}}\right)$$

et leurs réalisations

$$r_{(s',t)} = \ln\left(\frac{i_{(s',t)}}{i_{(s',t-1)}}\right) - E\left[\ln\left(\frac{i_{(s',t)}}{i_{(s',t-1)}}\right) \middle| \ln\left(\frac{i_{(s',t-1)}}{i_{(s',t-2)}}\right)\right],$$

soit

$$r_{(s',t)} = \ln\left(\frac{i_{(s',t)}}{i_{(s',t-1)}}\right) - \alpha_{(s')}$$

.

Considérant cette modélisation, nous obtenons la relation suivante

$$i_{(s',t_{max}+1)} = i_{(s',t_{max})} \times \exp(\alpha_{(s')} + \sigma_{(s')} \times \varepsilon_{(s',t_{max})})$$

,

où $\varepsilon_{(s',t_{max})} \sim N(0,1)$ et exp( ) : représente la fonction exponentielle.

On en déduit

$$\mathrm{var}\left\{i_{(s',t_{max}+1)}\right\} = i_{(s',t_{max})}^{2} \times \left[\exp(\sigma_{(s')}^{2}) - 1\right] \times \exp(2 \times \alpha_{(s')} + \sigma_{(s')}^{2})$$

,

et

$$E\left\{i_{(s',t_{max}+1)}\right\} = i_{(s',t_{max})} \times \exp(\alpha_{(s')} + \frac{\sigma_{(s')}^{2}}{2})$$

.

Le coefficient de variation associé à l'erreur de *process* $CV_{(s')}$ se définit de la façon suivante

$$CV_{(s')} = \frac{\sqrt{\mathrm{var}\left\{i_{(s',t_{max}+1)}\right\}}}{E\left\{i_{(s',t_{max}+1)}\right\}}$$

d'où

$$CV_{(s')} = \frac{i_{(s',t_{max})} \times \sqrt{\left[\exp(\sigma_{(s')}^{2}) - 1\right] \times \exp(2 \times \alpha_{(s')} + \sigma_{(s')}^{2})}}{i_{(s',t_{max})} \times \exp(\alpha_{(s')} + \frac{\sigma_{(s')}^{2}}{2})} = \sqrt{\exp(\sigma_{(s')}^{2}) - 1}$$

.

A partir de ce coefficient de variation, nous pouvons estimer sur le segment s et à la date t_{max}, l'écart-type associé à l'erreur de *process* sur les taux d'incidence en état d'incapacité/invalidité $i(x, s, t_{max})$ de la façon suivante : $\sigma_{(process,s)} = CV_{(s')} \times i(x, s, t_{max})$.

Formule du niveau de choc d'incidence

Sous hypothèse de non corrélation des erreurs d'estimation et de process, le niveau de choc global proposé est défini par la formule suivante :

$$\Delta_{(inc,s,USP)} = \frac{\sqrt{\sigma_{(process,s)}^{2} + \sigma_{(estim,s)}^{2}} \times VaR_{99,5\%}(\varepsilon)}{i_{(s,t_{max})}}$$

où $i_{(s,t_{max})}$ représente le taux d'incidence en incapacité/invalidité estimé sur l'historique de calibrage le plus récent se terminant à t_{max} pour le segment s et $VaR_{99,5\%}(\varepsilon)$ représente le quantile à 99,5% d'une loi normale centrée réduite ($\varepsilon \sim N(0,1)$).

Modélisation des chocs des lois de maintien ou de rétablissement

Nous retiendrons comme estimateur des taux de maintien ou de rétablissement, l'estimateur de Kaplan Meier. Il s'agit d'un estimateur non-paramétrique permettant d'approcher la forme empirique prise par le risque de sortie de l'état sans adopter une quelconque spécification de loi. Cet estimateur nécessite de connaître exactement toutes les dates d'entrée et de sortie du portefeuille, y compris les dates d'entrée en incapacité/invalidité. Ceci peut constituer un inconvénient dans le cas où le fichier de données serait mal renseigné.

Enfin, la prise en compte de l'effet des caractéristiques individuelles de la population étudiée par cette méthode nécessite la décomposition préalable en sous-populations suffisamment homogène par rapport à ces caractéristiques. Les problématiques de mise en œuvre opérationnelle de ce modèle sont présentées dans Planchet et al. (2011).

Dans ce qui suit, nous décrivons uniquement la méthodologie permettant de disposer des chocs de rétablissement de l'état d'incapacité/invalidité en intégrant les deux niveaux d'erreur requis. Le risque étant à la baisse des taux de rétablissement, nous présentons dans ce qui suit les méthodologies permettant de disposer des chocs de rétablissement au quantile à 0,5%.

De manière triviale, la même démarche s'applique au calibrage des chocs de maintien. Le risque dans ce cas étant à la hausse des taux de maintien, les chocs de maintien correspondent au niveau de quantile à 99,5%.

Estimation des taux de rétablissement

La variable de durée de la fonction de survie de Kaplan-Meier pour estimer les taux de rétablissement conditionnel est l'ancienneté.

La formule mathématique de cet estimateur sur un segment s, pour une ancienneté a en état d'incapacité/invalidité et à la date $t_{\max}$ est

$$R_{(a,s,t_{\max})} = 1 - \frac{S_{(a,s,t_{\max})}}{S_{(a-1,s,t_{\max})}},$$

où $S_{(a-1,s,t_{\max})}$ et $S_{(a,s,t_{\max})}$ représentent les taux de maintien en incapacité/invalidité estimés par la méthode de Kaplan Meier de la façon suivante

$$S_{(a,s,t_{\max})} = \prod_{\substack{i \\ \gamma_i \leq a}} \left(1 - \frac{d_{i,x}}{n_{i,x}}\right)$$

où $n_{i,x}$ représente le nombre d'assurés en incapacité/invalidité ayant passé plus de γ_i jours en incapacité/invalidité pour le segment s, $d_{i-1,x}$ représente le nombre de rétablissements survenus entre γ_{i-1} et γ_i jours d'ancienneté en incapacité/invalidité pour le segment s, $n_{i,x}$ représente l'exposition au risque intégrant les observations censurées et tronquées.

Estimation de l'erreur d'estimation

Dans le cas de l'utilisation de Kaplan-Meier pour estimer les taux de rétablissement $R_{(a,s,t_{\max})}$, l'erreur d'estimation associée peut-être approchée par l'estimateur de la variance de Greenwood. La formule mathématique de cet estimateur est

$$\sigma_{(estim,t_{\max})}{}^2 = (1-R_{(a,s,t_{\max})})^2 \times \sum_{\substack{i \\ a<\gamma_i \le a+1}} \frac{d_{i,x,t_{\max}}}{n_{i,x,t_{\max}} \times \left(n_{i,x,t_{\max}} - d_{i,x,t_{\max}}\right)}.$$

Modélisation de l'erreur de process

La méthodologie repose sur le suivi temporel de la variation des taux de rétablissement afin d'étudier l'évolution de ce risque. En effet, sur un historique de données suffisamment profond, il est possible de construire des lois d'expérience pour chaque année t (calibrées sur les périodes $[t-T;t]$ où $T \geq 1$ est l'amplitude des fenêtres d'estimation). Nous obtenons ainsi la série des taux de rétablissement de l'état d'incapacité/invalidité $\left(R_{(s,t)}\right)_t$.

Nous proposons de considérer la série chronologique des taux d'évolution des taux de rétablissement d'expérience $\left(Z_{(s',t)}\right)_t$ pour un segment s' donné, où $Z_{(s',t)}$ est défini comme une variation relative

$$Z_{(s',t)} = \frac{R_{(s',t+1)} - R_{(s',t)}}{R_{(s',t)}}.$$

Nous modélisons cette variable par un processus autorégressif d'ordre 1, dont la formule mathématique est

$$Z_{(s',t)} = \mu_{(s')} + \alpha_{(s')} \times Z_{(s',t-1)} + \sigma_{(s')} \times \varepsilon_{(s',t)},$$

où $\varepsilon_{(s',t)} \sim N(0,1)$, les coefficients $\mu_{(s')}$ et $\alpha_{(s')}$ sont estimés à l'aide de l'estimateur des moindres carrés ordinaires (ou pondérés par les expositions).

Les résidus $r_{(s',t)}$ sont obtenus par différence entre les espérances conditionnelles théoriques des variables $Z_{(s',t)}$ et leurs réalisations

$$r_{(s',t)} = Z_{(s',t)} - E\left\{Z_{(s',t)} \middle| Z_{(s',t-1)}\right\},$$

soit

$$r_{(s',t)} = Z_{(s',t)} - \left(\mu_{(s')} + \alpha_{(s')} \times Z_{(s',t-1)}\right),$$

où $\sigma_{(s')}$ correspond à l'écart-type empirique des résidus.

Considérant cette modélisation, on obtient la relation

$$R_{(s',t_{\max}+1)} = R_{(s',t_{\max})} \times \left(1 + Z_{(s',t_{\max})}\right),$$

soit

$$R_{(s',t_{\max}+1)} = R_{(s',t_{\max})} \times \left(1 + \mu_{(s')} + \alpha_{(s')} \times Z_{(s',t_{\max}-1)} + \sigma_{(s')} \times \varepsilon_{(s',t_{\max})}\right),$$

où $\varepsilon_{(s',t_{\max})} \sim N(0,1)$ suit une loi normale centrée réduite.

On en déduit

$$\mathrm{var}\left\{R_{(s',t_{\max}+1)}\right\} = R_{(s',t_{\max})}{}^2 \times \sigma_{(s')}{}^2$$

et

$$E\left\{R_{(s',t_{\max}+1)}\right\} = R_{(s',t_{\max})} \times (1 + \mu_{(s')} + \alpha_{(s')} \times Z_{(s',t_{\max}-1)}).$$

Le coefficient de variation associé à l'erreur de process $CV_{(s')}$ se définit comme la racine carrée d'une variance sur l'espérance mathématique, de formule

$$CV_{(s')} = \frac{\sqrt{\mathrm{var}\left\{R_{(s',t_{\max}+1)}\right\}}}{E\left\{R_{(s',t_{\max}+1)}\right\}}$$

d'où

$$CV_{(s')} = \frac{R_{(s',t_{\max})} \times \sigma_{(s')}}{R_{(s',t_{\max})} \times (1 + \mu_{(s')} + \alpha_{(s')} \times Z_{(s',t_{\max}-1)})}$$

L'écart-type associé à l'erreur de process sur le taux de rétablissement de l'état d'incapacité et invalidité

$R_{(s,t_{\max})}$ correspond à $\sigma_{(process,s)} = CV_{(s')} \times R_{(s,t_{\max})}$.

Formule du niveau de choc de rétablissement

Sous hypothèse de non-corrélation des erreurs d'estimation et de process, le niveau de choc global proposé est défini par la formule suivante :

$$\Delta_{(recov,s,USP)} = \frac{\sqrt{\sigma_{(process,s)}{}^2 + \sigma_{(estim,s)}{}^2} \times VaR_{0,5\%}(\varepsilon)}{R_{(s,t_{\max})}}$$

où $R_{(s,t_{max})}$ représente le taux de rétablissement de l'état d'incapacité/invalidité estimé sur l'historique de calibrage le plus récent se terminant à $t_{\max}$ pour le segment s et $VaR_{0,5\%}(\varepsilon)$ représente le quantile à 0,5% d'une loi normale centrée réduite ($\varepsilon \sim N(0,1)$).

Prise en compte des limites des données par la crédibilité

Nous avons signalé que la pertinence des paramètres (propres à l'entité étudiée) estimés par les différentes méthodes proposées dans les paragraphes précédents, dépend d'une part de la qualité des données, d'autre part de la profondeur d'historique disponible. Ainsi, nous proposons d'adopter une « approche crédibilité » pour tenir compte de l'expérience propre à l'entité dans la même logique que les USP non-vie actuels.

Du point de vue théorique, la théorie de crédibilité de Bülhmann ou théorie de la crédibilité fondée sur la plus grande exactitude (TCGE) est préférable, car elle est complète. La réalité des informations disponibles, la gestion de leur intégrité dans le temps, les limites et l'insuffisance des données d'expérience de la plupart des compagnies ne permettent pas sa mise en œuvre opérationnelle. Nous proposons donc une variante normalisée de la théorie de la crédibilité à variation limitée (TCVL) ou crédibilité américaine. Il s'agit d'une approche préconisée par l'Institut des Actuaires Canadiens dans sa note éducative (2002) pour résoudre des problématiques similaires sur les données afin d'établir des hypothèses d'évaluation de mortalité d'expérience lorsque les résultats sont crédibles, mais qu'il n'est pas possible de créer une table pour la compagnie.

Conditions sur les données d'expérience

Le terme « crédible » s'entend de données fiables au plan statistique. Une population d'assurés sera considérée « homogène » si sa structure ou sa composition est uniforme pour l'ensemble des individus face au risque étudié.

Lorsque les données d'expérience vérifient une taille suffisante validée d'après les critères de Cochran et d'exposition minimale pour assurer le calibrage robustes des lois centrales ; lorsque la profondeur d'historique est également jugée suffisante au sens de la stabilité des niveaux de chocs globaux (c'est-à-dire les niveaux de chocs ne sont pas modifiés en augmentant la profondeur d'historique d'une nouvelle année d'historique), alors les données sont jugées crédibles à 100 %, la compagnie peut estimer des paramètres spécifiques en se fondant exclusivement sur ses propres données.

Si le critère d'exposition minimale n'est pas vérifié, nous préconisons de retenir les paramètres de la formule standard, car les lois centrales estimées à partir de ces données seront trop volatile. Par contre, si le critère de profondeur d'historique seul n'est pas vérifié, nous parlerons de crédibilité partielle. En cas de crédibilité partielle, le choc global est la somme du choc estimé à partir des données d'expérience de la compagnie pondéré par le facteur de crédibilité et du paramètre de référence proposé par la formule standard (pour chacun des risques considérés) pondéré par un moins le facteur de crédibilité.

Critères d'une bonne méthode de crédibilité

Dans la pratique, les critères associés à une bonne méthode de crédibilité sont à la fois qualitatifs et quantitatifs. Ainsi une bonne méthode de crédibilité doit être d'application facile, utilisée tous les renseignements pertinents disponibles, l'application des paramètres issus de l'utilisation des facteurs de crédibilité obtenus doivent permettre de valider le niveau de sinistralité du portefeuille suivant différentes mailles de segmentation pertinente.

Il existe deux types principaux de théorie de crédibilité : la théorie de la crédibilité à variation limitée (TCVL) et la théorie de la crédibilité fondée sur la plus grande exactitude (TCGE). La méthode normalisée, qui représente un type de théorie de la crédibilité à variation limitée est celle qui répond opérationnellement le mieux aux critères d'une bonne méthode de crédibilité.

La théorie de la crédibilité fondée sur la plus grande exactitude (TCGE) ou « crédibilité européenne » repose sur les travaux de Bühlmann. La TCGE est mieux fondée au plan théorique que la TCVL et elle fait en sorte que les résultats sont « équilibrés », ce qui permet d'éviter de les normaliser. La théorie de la crédibilité fondée sur la plus grande exactitude permet d'établir une estimation des sources de variation à l'intérieur et entre les sous-catégories (de la population d'assurés du portefeuille étudié). Au plan théorique, la TCGE est complète et satisfait aux critères d'une bonne méthode de crédibilité mais elle comporte une lacune en ce qu'elle exige des renseignements supplémentaires au sujet des résultats de l'industrie (au-delà des données habituellement recueillies et diffusées). Abstraction faite de ces difficultés d'ordre pratique, la TCGE serait probablement la méthode de crédibilité privilégiée à utiliser pour établir l'hypothèse d'évaluation des paramètres des risques étudiés. Il existe plusieurs versions de la TCGE. Le modèle plus simple de Bühlmann et le modèle un peu plus complexe de Bühlmann-Straub.

Méthode normalisée de la théorie de crédibilité à variation limitée

La théorie de la crédibilité à variation limitée ou TCVL propose un critère de pleine crédibilité fondé sur la taille du portefeuille. On entend par pleine crédibilité, le cas où les données du portefeuille sont suffisantes pour le calibrage des paramètres d'expérience sans tenir compte des données de l'ensemble de l'industrie. En outre, cette théorie prévoit une méthodologie spéciale pour établir la crédibilité partielle, à savoir que les résultats du portefeuille et ceux de l'industrie sont pondérés.

A l'instar des USP non-vie et par application de la TCVL, une profondeur d'historique des données d'expérience correspondant au cas de pleine crédibilité sera déterminée pour chaque risque (incidence ou maintien vs rétablissement). Dans ce cas les paramètres issus du calibrage spécifique des chocs sont retenus. A contrario, sous contrainte de la profondeur minimale d'historique de données disponibles et lorsque la profondeur d'historique de pleine crédibilité n'est pas atteinte, il s'agira du cas de la crédibilité partielle et le niveau de choc global final sera une somme du paramètre spécifique (coefficient de crédibilité appliqué au choc déterminé à partir des données d'expérience) et du paramètre de la formule standard de Solvabilité II pondéré par un moins ce coefficient de crédibilité.

Nous retiendrons comme facteurs de crédibilité à appliquer aux chocs, les mêmes valeurs déterminées à partir des lois centrales pour chaque risque. Ce choix implique toute fois la remarque suivante : est-il raisonnable d'utiliser comme facteurs de crédibilité estimés sur lois centrales aux lois choquées, donc sur les chocs ? Nous acceptons cette hypothèse sans toutefois la démontrer dans ce mémoire. Sur cette base, l'hypothèse prévue pour le montant global des sinistres d'une société à l'égard d'une année t peut être exprimée de la manière suivante :

$$X_t = C\bar{X} + (1 - C)\mu,$$

où X_t désigne le montant global prévu des sinistres de l'année t, C désigne le facteur de crédibilité ou le facteur de pondération attribué à un échantillon de données, n est le nombre d'années d'expérience et correspond au nombre d'années d'historique de données disponibles utilisées pour l'estimation du facteur de crédibilité, $\{X_1, X_2, \dots, X_n\}$ les montants des sinistres survenus sur les n années d'observation, $\bar{X}$ représente la moyenne de ces montants des sinistres survenus sur les n années d'observation et μ représente la moyenne de la distribution sous-jacente correspondant au montant prévu de sinistre d'après les données de l'industrie pour le même portefeuille.

En pratique, nous appliquerons le facteur de crédibilité au nombre de sinistre ou au ratio nombre de sinistre réel à prévu, plutôt qu'aux montants prévus des sinistres. De plus, il convient de faire quelques remarques complémentaires. En effet, l'industrie ne fournit pas toujours des données pour la construction de lois de place pouvant servir de référence pour le calcul de la moyenne μ. C'est le cas par exemple pour le risque d'incidence en arrêt de travail de notre étude. Se pose la question du calibrage de ce paramètre utile à nos calculs. Si l'entreprise dispose d'un portefeuille (ou par regroupement de portefeuilles de risques homogènes) de taille et de profondeur d'historique de données suffisantes,

elle peut utiliser ces informations pour réaliser ces calculs. Dans ce cas le facteur de crédibilité de 100% sur ce périmètre. Dans le cas contraire, l'entreprise pourra s'appuyer sur des jugements d'experts objectivement justifiés. Enfin, le fait que cette formule de crédibilité moyenne pondérée comporte un attrait intuitif, la TCVL ne prévoit pas un modèle théorique sous-jacent de distribution des valeurs $(X_i)_i$ qui soit conforme à la formule.

Ainsi, en vertu de la TCVL, on obtient X_t en sélectionnant un paramètre d'écart r ($r > 0$) et un niveau de probabilité p ($0 < p < 1$), de sorte que l'écart entre X_t et sa moyenne μ est faible. Dans le cas d'un modèle de Poisson de paramètre γ, ce critère peut s'exprimer de la manière suivante : $Pr\{|X_t - \mu| \leq -r\gamma\} \geq p$, où r représente la marge d'erreur et p l'intervalle de confiance. En d'autres termes, X_t représente une bonne estimation de la sinistralité future prévue s'il y a une probabilité élevée que l'écart entre X_t et sa moyenne μ soit faible relativement à μ.

Dans bien des cas, il est raisonnable d'établir la distribution de $\bar{X}$ (moyenne de ces montants des sinistres survenus sur les n années d'observation) comme étant une distribution normale. Dans ces cas, le nombre de sinistres correspondant aux différentes valeurs des paramètres p et r peut être obtenu à partir de tables normales types. Le tableau suivant énonce le nombre de sinistres requis pour obtenir une pleine crédibilité selon diverses valeurs de p et r.

Table de la loi normale standard

Nombre de sinistres requis pour une crédibilité totale

Paramètre de probabilité (p)	Paramètre d'écart (r)				
	5%	4%	3%	2%	1%
90%	1 082	1 691	**3 007**	6 765	27 060
95%	1 537	2 401	4 268	9 604	38 416
99%	2 654	4 147	**7 373**	16 589	66 538
99,9%	4 331	6 767	12 030	27 068	108 274

Tableau 2 : Table normale standard – Paramètres d'écart et de probabilité

Lorsque l'on détermine l'hypothèse d'évaluation de la sinistralité centrale prévue, on pourrait vouloir utiliser un seuil élevé pour établir une pleine crédibilité, notamment p = 90 % et r = 3 % pour lequel la valeur provenant de cette distribution est 3 007 sinistres. Dans le cas du choix d'un seuil pour déterminer les facteurs de crédibilité à appliquer à des chocs, nous suggérons de faire le choix d'un seuil plus élevé notamment p = 99 % et r = 3 % pour lequel la valeur est 7 373 sinistres.

Bien que la distribution théorique de l'incapacité ou invalidité d'une garantie d'arrêt de travail soit binomiale, lorsque les probabilités de survenance d'un événement (incidence, représenté par la variable aléatoire X_t dans les formules ci-dessus) sont faibles, la distribution de Poisson offre une approximation raisonnable d'une distribution binomiale. Dans le modèle de Poisson, la seule variable aléatoire est le nombre de sinistres, qui est réputée correspondre au modèle de Poisson. Les variations quant à la taille des sinistres ne sont pas prises en compte. S'il y a dispersion importante du montant net des sinistres pour chaque police du portefeuille étudié, le recours au modèle de Poisson simple pourrait ne pas convenir. Le modèle de Poisson composé intègre l'effet de la variante de taille des sinistres, et il se traduirait normalement par un relèvement du seuil des sinistres requis pour atteindre le même niveau de crédibilité.

Ainsi, faisant l'hypothèse que l'ensemble des sinistres d'un portefeuille de polices peut être décrit comme un modèle de Poisson composé reflétant le nombre et le montant des sinistres, lorsqu'on y

ajoute la variabilité de la taille des sinistres, le seuil de pleine crédibilité est majoré par rapport au modèle de Poisson. Supposons que N est une variable aléatoire représentant le nombre de sinistres d'un portefeuille et qu'il suit une distribution de Poisson dont la moyenne et la variance sont γ. Le nombre de sinistres observés est m. Pour k = 1, 2, 3, … m, supposons que Y_k est la variable aléatoire représentant le montant du kième sinistre. Supposons que les variables Y_k sont indépendantes et ont une distribution dont la moyenne est μ_Y et la variance est σ_Y^2. Supposons également que le nombre de sinistres N est indépendant du montant des sinistres Y_k. Le montant total de sinistres $X = Y_1 + Y_2 + \cdots + Y_N$ a une distribution de Poisson combinée.

En résumé, la valeur X_i a une distribution de Poisson composée avec paramètre de Poisson γ et une distribution du montant des sinistres ayant une moyenne μ_Y et une variance σ_Y^2. À l'aide du modèle de Poisson composé assorti des valeurs paramétriques r = 3 % et p = 99 %, on peut constater que le nombre de sinistres requis pour une pleine crédibilité est obtenu par la formule suivante :

$$Seuil = 7373 \times \frac{\left(\sum_{i=1}^{m} i_{\left(s_i',t\right)} V_i^2\right)\left(\sum_{i=1}^{m} i_{\left(s_i',t\right)}\right)}{\left(\sum_{i=1}^{m} i_{\left(s_i',t\right)} V_i\right)^2},$$

où $i = 1,2,\ldots,m$ le nombre d'individus sinistrés observés, s_i' est la maille de segmentation retenue pour le calibrage de la loi centrale, t correspond à l'année de calcul, $\delta_{(s_i',t)}$ est la loi de sinistralité centrale, V_i correspond au montant de la provision du sinistre ou au capital restant dû de l'individu i.

Pour établir la valeur du facteur de crédibilité c à l'égard du modèle de Poisson composé, il convient de calculer la moyenne μ_Y et l'écart-type σ_Y de la distribution du montant des sinistres. Ces valeurs peuvent être calculées à partir du risque total ou faire l'objet d'une estimation à l'aide des sinistres réels. Pour que la crédibilité soit totale, le nombre de sinistre figurant dans les résultats du portefeuille doit dépasser le nombre $Seuil$.

Si le critère de pleine crédibilité n'est pas respecté, on peut appliquer une crédibilité partielle. Ainsi, pour les valeurs p = 99 % et r = 3 %, l'application de la règle de la « racine carrée » se traduit par un facteur de crédibilité c calculé avec la formule suivante :

$$C = Min\left\{\sqrt{\frac{X}{Seuil}}, 1\right\},$$

où $Seuil$ représente le critère de pleine crédibilité et X le nombre de sinistres observés dans le portefeuille. On peut remarquer que si l'on ne tient compte que du nombre de sinistre la valeur du $Seuil$ est 7373.

L'exemple de modèle de Poisson composé peut être élargi pour tenir compte des données portant sur plus d'une période ou année, où : N_j est une variable aléatoire représentant le nombre de sinistres au cours de la période j (j = 1,2,3,…,n), $Y_{j,k}$ est la variable aléatoire représentant le montant du k-ième sinistre au cours de la période j, X_j est la variable aléatoire qui représente le montant global des sinistres pour la société au cours de la période j (avec : $X_j = Y_{j,1} + Y_{j,2} + \cdots + Y_{j,m}$). Cependant, le nombre d'années serait limité pour assurer l'homogénéité du portefeuille au fil des années d'observation disponibles, voir Klugman, Panjer et Willmot (1998).

Si la compagnie désire tenir compte des résultats des sous-catégories (comme le sexe, la catégorie socio-professionnelle, la durée d'indemnisation, les franchises…), mais que les résultats de ces sous-catégories ne sont pas crédibles à 100%, elle pourrait décider d'utiliser un facteur de crédibilité global ou le facteur de crédibilité moins élevé qui convient aux résultats de la sous-catégorie.

On peut regrouper des distributions disparates à l'intérieur des données globales à certaines conditions. Supposons que le portefeuille comporte six sous-catégories différentes : hommes et femmes ainsi répartis : avec examen médical, sans examen médical et avec examen paramédical. Pour chaque

sous-catégorie, la distribution du nombre de sinistres se conforme à une distribution de Poisson assortie d'un paramètre différent (à cette fin, supposons que la variable aléatoire à l'étude est le ratio sinistre réel à prévu calculé pour chaque catégorie). Si les proportions relatives des sous-catégories sont stables au fil du temps, on peut utiliser le facteur de crédibilité fondé sur la distribution globale de ces distributions hétérogènes de Poisson (c'est-à-dire le facteur de crédibilité global de la société) pour chacune des sous-catégories. L'exigence selon laquelle la composition du portefeuille doit être stable au fil des années en ce qui touche les principales sous-catégories peut limiter l'applicabilité de ce résultat. La composition du portefeuille peut être considérée comme stable au fil du temps si les proportions des sous-catégories pertinentes sont constantes (tant pour la période à l'étude que pour la période de projection future). Si les proportions relatives des sous-catégories ne sont pas stables au fil du temps, d'où l'invalidité des hypothèses, il pourrait convenir de tenir compte de la crédibilité des sous-catégories dans le cadre de l'établissement des hypothèses de sinistralité prévue. Pour déterminer si les conditions sont valables, l'on doit faire preuve de jugement.

La méthode normalisée est l'approche que nous avons privilégiée dans le cadre de notre étude. Cette méthode utilise la crédibilité et les ratios sinistre réel à prévu (notés R/P) des sous-catégories. Cependant, ces ratios R/P sont rajustés pour reproduire le niveau des sinistres prévus découlant du ratio total R/P de la compagnie ajusté pour tenir compte de la crédibilité de l'ensemble des portefeuilles étudiés de la société.

La méthode normalisée offre les avantages suivants :

- La somme des sinistres prévus pour les sous-catégories correspond au nombre total de sinistres prévus, d'après le ratio R/P, calculé à l'échelle de la société (c'est-à-dire que le nombre de sous-catégories sélectionnées n'a pas d'effet sur le résultat global).
- Tous les renseignements sont utilisés : les ratios R/P de l'ensemble de la société et ceux des sous-catégories, de même que les facteurs de crédibilité.
- Les résultats sont raisonnables dans des cas extrêmes ou limitatifs.
- Les ratios R/P des sous-catégories demeurent à l'intérieur de la fourchette initiale.
- Les effets interactifs entre les sous-catégories peuvent être saisis ; et il s'agit d'une méthode facile à mettre en pratique.

Bien que cette méthode ne comporte pas un fondement solide au plan théorique, elle est pratique et satisfait aux critères associés à une bonne méthode de crédibilité.

5. Applications et résultats empiriques

5.1. Lois centrales d'incidence et de maintien d'expérience

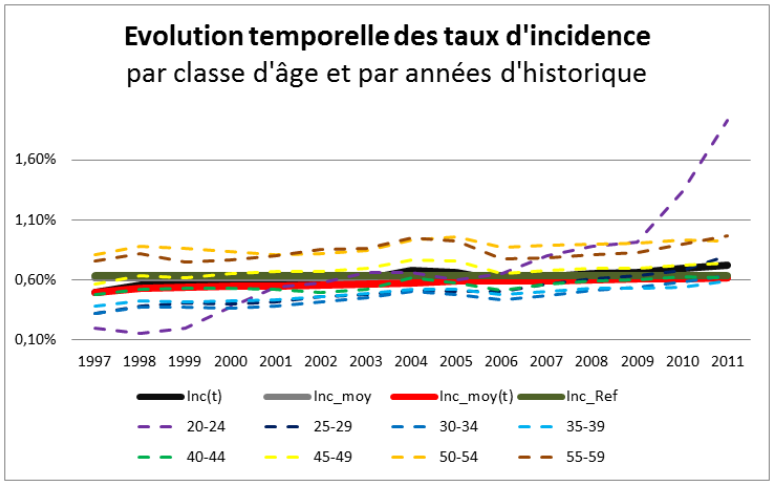
Pour estimer les lois d'incidence, nous avons besoin à la fois des bases des assurés et des bases des sinistrés. La meilleure approximation des lois est réalisée avec les bases détaillées (tête par tête). Même si l'historique est parfois peu profond pour certains partenaires (de l'ordre de 3 – 5 ans), il demeure suffisant pour la plupart des portefeuilles.

Dans cette étude, afin de disposer de la volumétrie la plus importante possible et fournir un jeu de données représentatif du marché français, nous avons fait le choix de regrouper les données des différents partenaires. Nous avons fait attention à sélectionner les produits de mêmes caractéristiques techniques ou très proches. Enfin pour calculer les taux d'incidence en incapacité/invalidité, nous utilisons l'estimateur actuariel car rappelons que les bases des assurés et des sinistrés ne communiquent pas entre elles, en l'absence de clé d'identification commune pour un même assuré.

Disposant de données de qualité, avec une taille et un historique suffisant, nous pouvons calculer les paramètres suivants :

- Les taux d'incidence centraux par tranche d'âge de 5 ans (entre 20 ans et 59 ans) calculés sur un historique de 2 ans entre 1997 et 2011 (notés 20-24, 25-29, …, 55-59), pour apprécier l'évolution temporelle de des lois d'incidence centrale.
- Les taux d'incidence centraux tout âge confondus année par année (Inc(t)).
- Les taux d'incidence moyens centraux tout âge confondus calculés année par année comme moyenne pondérée par les expositions des taux précédents (Inc_moy(t)).
- Les taux d'incidence moyens globaux tout âge et toute année d'historique (Inc_moy).
- Les taux d'incidence centraux de référence 2012 par tranche d'âge calculés sur un historique de 5 ans entre 2007 et 2011 (Inc_Ref).

Sans surprise, en comparant les taux d'incidence centraux par tranche d'âge nous observons qu'ils augmentent globalement avec l'âge. Nous observons en particulier que la tranche 20-24 ans est la plus instable dans le temps ; ceci est dû à la faible exposition de cette tranche. Il faut noter également une hausse de la sinistralité dans le temps. Les tranches 25-29 ans à 44-49 ans sont quant à elles moins volatiles même si elles connaissent aussi une hausse globale de la sinistralité dans le temps, tout en gardant des niveaux crissant en fonction de l'âge. Nous observons des taux d'incidence globalement assez regroupés par rapport aux autres tranches, avec des comportements assez similaires. Enfin, les tranches 50-54 ans et 55-59 ans représentent les taux d'incidence les plus élevés et le plus stables dans le temps.



Graphique 1 : Evolution des taux d'incidence dans le temps

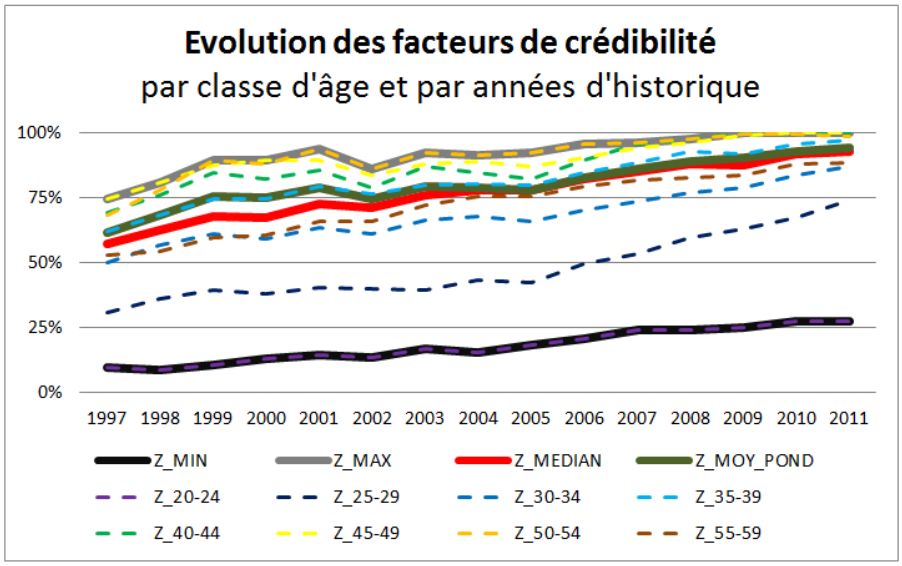
L'objectif de l'étude est de calibrer des facteurs de crédibilité tout âge confondu et à différentes années. Sauf le cas particulier de la tranche 20-24 ans, les différentes tranches évoluent dans le temps de manière similaire, tout en conservant pour la plupart leur niveau d'incidence correspondant à l'effet âge. En effet, il est naturel que le taux d'incidence augmente avec l'âge car le vieillissement a un impact sur l'état de santé et est un facteur de sévérisation du risque arrêt de travail.

Ainsi, nous faisons le choix de faire une analyse tout âge confondu. Dans ces conditions, nous observons une légère hausse dans le temps des taux d'incidence, ce qui se traduit par une augmentation dans le temps des taux Inc(t) et Inc_moy(t). Nous observons également que les taux Inc_moy et Inc_Ref. Ces différents taux sont compris entre 0,50% et 0,7%.

5.2. Estimation des facteurs de crédibilité

En appliquant cette démarche aux données disponible sur la LoB étudiée (disposant d'une profondeur d'historique supérieure à 15 ans), pour une marge d'erreur acceptable $r = 3\%$ et un intervalle de confiance de niveau $p = 99\%$, le nombre de sinistre minimum requis pour une pleine crédibilité est de 7373 (valeur issue de la table de la loi normale standard).

En utilisant les résultats de la TCVL normalisée sous l'hypothèse d'une loi de Poisson, pour un seuil de pleine crédibilité à 7373, nous pouvons obtenir un premier jeu de facteur de crédibilité en fonction des tranches d'âges et de la profondeur d'historique.

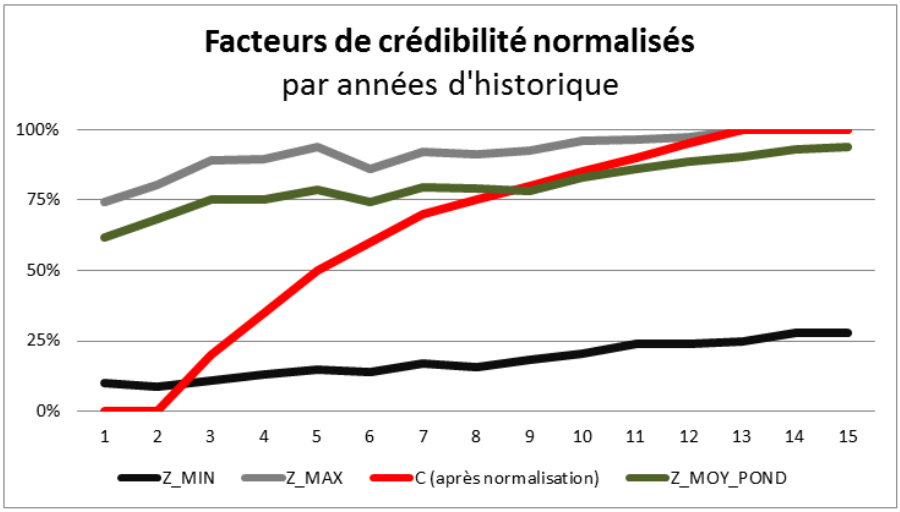


Graphique 2 : Evolution des facteurs de crédibilité en fonction des tranches d'âges et de l'historique

En première remarque nous observons que les coefficients de crédibilités augmentent globalement à la fois avec les tranches d'âges et l'historique. Ce qui traduit à la fois l'effet âge et l'effet volume des données disponibles. L'effet volume est de deux ordres, d'une part pour les contrats de prêts considérés, les expositions sont plus faibles aux âges jeunes, d'autre part la croissance normale du portefeuille joue également sur l'augmentation des expositions avec le temps.

Nous proposons de retenir avant normalisation et ajustements, comme facteurs de crédibilité les coefficients moyens pondérés par les expositions notés Z_Moy_POND. A partir de ces facteurs, nous procédons à une normalisation par réajustement des coefficients pour reproduire le niveau global des sinistres du portefeuille étudié.

Même si dans notre exemple nous obtenons des coefficients de crédibilité avant normalisation significatifs (62% la première année et 68% la seconde année) à cause de la taille de ce portefeuille, nous proposons une limite minimale de 3 ans d'historique en-deçà de laquelle les coefficients de crédibilité sont nuls. En effet, il est souvent difficile de calibrer des lois d'expérience robustes sur un historique de moins de trois ans.



Graphique 3 : Facteurs de crédibilité en fonction de l'historique

Les coefficients de crédibilité estimés par année d'ancienneté dans le portefeuille de la loi d'incidence sont les suivants :

Facteurs de crédibilité par années d'historique disponibles de la loi d'incidence											
Année	3	4	5	6	7	8	9	10	11	12	13 et +
C	20%	35%	50%	60%	70%	75%	80%	85%	90%	95%	100%

Tableau 4 : Facteurs de crédibilité de la loi d'incidence

où « Année » représente le nombre d'années d'historique disponible et « C » le facteur de crédibilité.

De manière analogue, nous obtenons les coefficients de crédibilité de la loi de maintien (vs rétablissement).

Facteurs de crédibilité par années d'historique disponibles de la loi de maintien vs rétablissement											
années	5	6	7	8	9	10	11	12	13	14	15 et +
C	30%	40%	50%	60%	70%	75%	80%	85%	90%	95%	100%

Tableau 5 : Facteurs de crédibilité de la loi de maintien vs rétablissement

Dans le cas de la loi d'incidence, le nombre d'historique minimal est de 3 ans alors que pour la loi de maintien (vs rétablissement) il est de 5. Dans ces conditions, il conviendrait d'appliquer pour les contrats d'anciennetés inférieures les chocs de la formule standard. Enfin, comme nous l'avons proposé dans la partie théorique, nous retiendrons ces coefficients calibrés sur distribution des lois centrales sous-jacentes comme facteurs de crédibilité des paramètres de chocs. Notons qu'il s'agit d'une première approximation de ces paramètres. En effet, sous cette hypothèse, nous acceptons sans le prouver dans cette étude, l'invariance des coefficients de crédibilité par les transformations conduisant à passer des lois centrales aux lois choquées.

Estimation des chocs d'incidence

Disposant des coefficients de crédibilité et des paramètres de chocs calibrés avec les méthodes proposées dans la partie théorique, nous pouvons estimer les chocs USP suivants différentes années d'historique.

Application du modèle AR(1) sur un portefeuille d'historique inférieur ou égal à 16 années (avec C=100% à 16 ans d'ancienneté) :

	CHOCS USP INCIDENCE – AVEC MODELE AR(1)		
	Sans facteur de crédibilité	Avec facteur de crédibilité (1ère année)	Avec facteur de crédibilité (année > 1)
20	68%	68%	68%
38	12%	12%	12%
59	18%	18%	18%

Tableau 6 (a) : Chocs USP avec le modèle AR(1)

Application du modèle Log-normal sur un portefeuille d'historique inférieur ou égal à 16 années (avec C=100% à 16 ans d'ancienneté) :

	CHOCS USP INCIDENCE – AVEC MODELE LOG NORMAL		
	Sans facteur de crédibilité	Avec facteur de crédibilité (1ère année)	Avec facteur de crédibilité (année > 1)
20	111%	111%	111%
38	17%	17%	17%
59	22%	22%	22%

Tableau 6 (b) : Chocs USP avec le modèle Log Normal

Quelle que soit la méthode retenue, nous observons une grande différence entre les chocs issus du calibrage propre à l'entité et les chocs proposés par la Formule Standard. Lorsque la profondeur de l'historique est importante et que le facteur de crédibilité est de 100%, les chocs issus du calibrage spécifique à l'entité restent toujours plus intéressants par rapport aux chocs de la Formule Standard (les niveaux de chocs de la loi d'incidence sont de 35% la 1ère année et de 25% au-delà).

En effet, contrairement aux chocs de la Formule Standard qui sont constants quel que soit l'âge, les chocs issus du calibrage propre à l'entité dépendent de l'âge avec de fortes variations entre les niveaux de chocs appliqués aux différents âges. Même si pour les âges jeunes nous observons des niveaux de choc spécifique très élevés (âges où les effectifs sont faibles), ils restent plus faibles que les chocs de la Formule Standard aux âges où les expositions sont les plus importantes sur le portefeuille d'étude.

Nous observons également que les niveaux de chocs sont différents suivant la méthode utilisée. Ils sont en effet plus conservateurs avec la méthode Log Normale qu'avec l'approche AR(1) dans notre exemple.

Application du modèle AR(1) sur un portefeuille d'historique inférieur ou égal à 5 années (avec C=50% à 5 ans d'ancienneté) :

	CHOCS USP INCIDENCE – AVEC MODELE AR(1)		
	Sans facteur de crédibilité	Avec facteur de crédibilité (1ère année)	Avec facteur de crédibilité (année > 1)
20	49%	42%	37%
38	9%	22%	17%
59	7%	21%	16%

Tableau 7 : Chocs USP avec le modèle AR(1)

Avec la méthode AR(1), les niveaux de chocs observés pour les âges plus jeunes (à 20 ans) sont les plus élevées et sont supérieurs aux chocs de la Formule Standard. L'application du coefficient de crédibilité dans ce cas permet de corriger l'effet d'une sur-estimation des niveaux des chocs liés à la taille des effectifs. Le niveau de choc final crédible passe de 49% à 42% la 1ère année et à 37% au-delà. Cependant les effectifs auxquels sont appliqués ces chocs sont relativement faibles ce qui induira un faible impact sur le calcul de SCR. Les paramètres spécifiques dans ce cas restent toujours plus intéressants.

Estimation des chocs de maintien et de rétablissement

Nous présentons dans cette partie les résultats des paramètres de chocs d'expérience sur les lois de maintien et de rétablissement pour différentes années d'historique.

Exemple des niveaux des chocs USP de rétablissement à la 3ème année d'ancienneté en incapacité invalidité. Application du modèle AR(1) sur un portefeuille d'historique inférieur ou égal à 23 années (avec C=100% à 23 ans d'ancienneté) :

	CHOCS DE RETABLISSEMENT – 3ème ANNEE D'ANCIENNETE		
	Taux de maintien	Sans facteur de crédibilité	Avec facteur de crédibilité
20-30	< 50%	+ 15 % sur taux de maintien	+ 15 % sur taux de maintien
45-50	> 50%	- 10 % sur taux de rétablissement	- 10 % sur taux de rétablissement
55-65	> 50%	- 13 % sur taux de rétablissement	- 13 % sur taux de rétablissement

Tableau 8 : Chocs USP avec le modèle AR(1)

La première grande différence avec les chocs proposés par la Formule Standard est que l'on observe de fortes variations entre les niveaux de chocs appliqués aux différentes anciennetés et classes d'âge. Les niveaux de chocs USP restent encore très compétitifs dans ce cas par rapport aux chocs de la Formule Standard (+20% sur taux de maintien si taux de maintien inférieur à 50% vs -20% sur taux de rétablissement sinon).

Toutefois, les niveaux de chocs observés pour les anciennetés les plus élevées sont supérieurs aux chocs de la Formule Standard. Cependant les taux auxquels ils sont appliqués sont relativement faibles ce qui induira un faible impact lors du calcul de SCR.

Exemple des niveaux des chocs USP de rétablissement à la 3ème année d'ancienneté en incapacité invalidité. Application du modèle AR(1) sur un portefeuille d'historique inférieur ou égal à 7 années (avec C=50% à 7 ans d'ancienneté) :

	CHOCS DE RETABLISSEMENT – 3ème ANNEE D'ANCIENNETE		
	Taux de maintien	Sans facteur de crédibilité	Avec facteur de crédibilité
20-30	< 50%	+ 14 % sur taux de maintien	+ 17 % sur taux de maintien
45-50	> 50%	- 9 % sur taux de rétablissement	- 15 % sur taux de rétablissement
55-65	> 50%	- 12 % sur taux de rétablissement	- 16 % sur taux de rétablissement

Tableau 9 : Chocs USP avec le modèle AR(1)

On observe qu'avant application des facteurs de crédibilité, les chocs USP sur historique de 7 ans sont plus faibles que ceux de l'historique de 23 ans (voir Tableau 8). Ceci découle du fait que l'écart-type associé à l'erreur de *process* est plus faible dans ce cas que lorsque l'on considère un historique plus long (tableau 8). En effet, le faible nombre de données conduit à estimer un modèle AR(1) qui « fit » mieux les données observées. L'écart-type associé au modèle AR(1) qui découle de cette estimation est donc d'autant plus faible. Dans notre exemple, l'introduction du facteur de crédibilité permet de compenser cet effet indu.

Au global, les niveaux de chocs obtenus avec l'historique le plus court sont plus sévères que ceux obtenus avec l'historique long mais reste néanmoins une alternative à la Formule Standard avantageuse dans l'illustration proposée.

Impacts économiques et capital de solvabilité requis

Les impacts économiques du calibrage des USP sur les lois d'incidence et de maintien (vs rétablissement) dépendent de la profondeur d'années d'historique. En effet, nous constatons des impacts à la baisse sur le SCR de souscription. Cette baisse passe de -60% avec les paramètres calibrés sur un historique de 16 ans à -30% pour un historique de 5 ans.

	SCR impact – long history of data			
	BE	SCR - USP	SCR - FS	Evolution
Total	1000	261	856	- 60 %

Tableau 10 : Impacts sur SCR de souscription sur un historique long (16 ans)

	SCR impact – long history of data			
	BE	SCR - USP	SCR - FS	Evolution
Total	1000	500	856	- 30 %

Tableau 11 : Impacts sur SCR de souscription sur un historique court (5 ans)

Nous observons très naturellement que l'impact en termes d'économie de SCR est très important et dépend de la profondeur d'historique disponible.

6. Discussions

Certains experts pensent à juste titre que les USP constituent principalement une opportunité pour les assureurs de niche, qui pourront ainsi profiter de l'effet de levier induit par leur forte expertise de leur secteur d'activité tout en contournant la complexité relative à la mise en place d'un modèle interne partiel.

La méthode proposée est applicable à la fois au risque d'incidence et au risque de maintien (ou de rétablissement) et n'est pas remise en cause par la spécificité de chacun de ces risques. De plus, si nous considérons le risque de maintien, même si les lois de maintien en incapacité et en invalidité présentent des caractéristiques communes, il n'en demeure pas moins qu'elles ont chacune leurs particularités. Outre les différences de durée et de fréquence, les statuts d'invalides et d'incapables sont clairement définis par la loi, avec des degrés différents et une nomenclature bien précise.

Dans notre illustration, la loi d'incidence dispose d'un historique minimal de 3 ans alors que pour la loi de maintien (vs rétablissement) il est de 5. Dans ces conditions, il conviendrait d'appliquer pour les contrats d'anciennetés inférieures les chocs de la formule standard par défaut.

La principale limite à l'approche proposée demeure celle liée aux données disponibles. En effet, la qualité des résultats issus des modèles dépend de la qualité des données, du volume et de la profondeur d'historique. S'agissant d'une estimation de chocs correspondant à des déviations extrêmes des risques, il est évident qu'un historique peut profond ne permet pas de disposer d'estimation robuste des paramètres de volatilité sous-jacents au modèle. L'introduction de la crédibilité (TVCL), certes permet d'apporter des ajustements cohérents avec les paramètres de la formule standard, mais ne règle pas tout, puisqu'elle fait une autre hypothèse forte sur l'adéquation des paramètres de chocs de la formule standard Solvabilité II au risque étudié.

En réalité, le calibrage des modèles théoriques n'est pas toujours aisé à partir des données disponibles pour de multiples raisons. Le superviseur est conscient des problématiques que posent les données et demande un certain nombre de justificatifs sur la qualité des données et sur les choix des périmètres retenus pour le calibrage de paramètres spécifiques dans le dossier de candidature. Ces justifications sont nécessaires pour inciter les praticiens à l'usage de modèles parcimonieux, tout en veillant à fournir aux instances dirigeantes et aux fonctions clés de l'entreprise les éléments de compréhension et d'explication des résultats issus des modèles.

7. Conclusion

L'approche proposée a servi à l'estimation des chocs des risques d'incidence et de maintien (ou de rétablissement) de la garantie arrêt de travail d'un contrat d'assurance de prêts (ou emprunteurs) d'une compagnie d'assurance vie. Ce portefeuille dispose de plus de 12 millions de têtes assurées et de générations de souscription dépassant 20 années d'ancienneté. L'une des principales difficultés de calibrage des paramètres des modèles retenus est la disponibilité de données de taille et de profondeur d'historique suffisantes. L'introduction d'une variante normalisée de la théorie de la crédibilité à variation limitée ou crédibilité américaine (TCVL) dans la modélisation permet d'estimer des niveaux de chocs spécifiques par ancienneté des polices du portefeuille. Les niveaux de chocs obtenus conduisent à des besoins de capitaux requis (SCR) en moyenne de l'ordre de 50% à 60% plus faibles que les ceux estimés avec les paramètres forfaitaires de la formule standard. La mise en place d'une approche USP pourrait donc être source d'économie de besoin de capital requis dans certains cas.

Le cadre méthodologique proposé dans cet article pourrait servir de base de discussion et d'implémentation chez les compagnies disposant de données suffisantes et de qualité pour estimer les paramètres spécifiques des risques d'incidence et de maintien en assurance de personne. Ce cadre est également adaptable aux travaux de gestion des risques notamment lors des évaluations ORSA de la

norme Solvabilité II. Ainsi, l'usage d'une modélisation très précise des risques permet de détecter de manière anticipée les dérives éventuelles.

L'introduction des facteurs de crédibilité est une approche prudente car permettant de tenir compte au moins partiellement de l'expérience de la compagnie, mais il reste à approfondir la problématique du choix des coefficients de crédibilité, à savoir, l'hypothèse de l'application des coefficients de crédibilité calibrés sur lois centrales à des paramètres de chocs. De plus, l'approche série temporelle nécessite de disposer d'historique conséquent pour son utilisation. Elle peut donc ne pas être adaptée dans la plupart des cas, ce qui conduirait à l'utilisation de l'approche modèle linéaire généralisé de type Log Normale qui fournit des paramètres un peu plus sévères.

Références

1. ACP (2012), La revue de l'Autorité de contrôle prudentiel, octobre-novembre 2012
2. AGUIR N. (2012), Impact de la modélisation multi-états sur le SCR en prévoyance, Mémoire de fin d'étude présenté devant l'Institut de Science Financière et d'Assurances pour l'obtention du diplôme d'Actuaire de l'Université de Lyon le 26 Septembre 2012
3. BAILEY, A. L. (1945), «A generalized theory of credibility», Proceedings of the Casualty Actuarial Society, vol. 32, p. 13–20.
4. BAILEY, A. L. (1950), «Credibility procedures, Laplace's generalization of Bayes' rule and the combination of collateral knowledge with observed data», Proceedings of the Casualty Actuarial Society, vol. 37, p.7–23.
5. BOURDONNAIS R., M. TERRAZA (2010), Analyse des séries temporelles, 3ème édition Dunod
6. BOX G.E.P. et G.M. JENKINS, Time Series Analysis: forecasting and control, Ed. Holden-Day
7. BROCKWELL P.J., R.A. DAVIS (1987), Time Series: Theory and Method, Ed. Springer-Verlag
8. BROCKWELL P. J., R. A. DAVIS (2003), Introduction to Time Series and Forecasting, 2nd Ed. Springer
9. BÜHLMANN, H. (1967), «Experience rating and credibility», ASTIN Bulletin, vol. 4, p. 199–207
10. BÜHLMANN, H. (1969), «Experience rating and credibility», ASTIN Bulletin,vol. 5, p. 157–165
11. CAMBON A. (2011), Elaboration d'un modèle interne partiel concernant le risque de souscription non-vie pour tenir compte des spécificités d'une société spécialisée dans les branches longues, Mémoire présenté devant l'Institut de Statistique de l'Université Pierre et Marie Curie pour l'obtention du diplôme de statisticien mention actuariat
12. CANADIAN INSTITUTE OF ACTUARIES (2002), Commission des rapports financiers des compagnies d'assurances vie, Note éducative, mortalité prévue : polices canadiennes d'assurances vie individuelle avec tarification complète
13. CANADIAN INSTITUTE OF ACTUARIES (2014), Commission des rapports financiers des régimes de retraite, Note éducative révisée : sélection des hypothèses de mortalité aux fins des évaluations actuarielles des régimes de retraite
14. CERCHIARA R. R., MAGATTI V., Undertaking Specific Parameters or a Partial Internal Model under Solvency II?, 30th International Congress of Actuaries, 30 March to 4 April 2014, Washington, D.C., USA
15. De VYLDER, F. (1981), «Practical credibility theory with emphasis on parameter estimation», ASTIN Bulletin, vol. 12, p. 115–131
16. DOITTEAU M., La réglementation et la modélisation stochastique de l'incapacité, Mémoire de fin d'étude présenté devant l'Institut de Science Financière et d'Assurances pour l'obtention du diplôme d'Actuaire de l'Université de Lyon le 4 janvier 2011
17. EIOPA (2011), Draft proposal for Implementing Technical Standard on Undertaking Specific Parameters: Methods, EIOPA-FinReq-11/023
18. EIOPA (2013), Technical Specification on the Long Term Guarantee Assessment (Part 1)
19. GOEL, P. K. (1982), «On implications of credible means being exact bayesian», Scandinavian Actuarial Journal, p.41–46.
20. GOOVAERTS, M. J. et W. J. HOOGSTAD (1987), Credibility Theory, no 4, Surveys of actuarial studies, Nationale-Nederlanden N.V., Netherlands
21. GOULET V. (2010), ACT 2008 Mathématiques actuarielles IARD II (Théorie de la crédibilité), école d'actuariat université Laval
22. GUIBERT Q., M. JUILLARD, O. N. TEUGUIA, F. PLANCHET (2014), Solvabilité Prospective en Assurance Méthodes quantitatives pour l'ORSA, Ed. Economica
23. HACHEMEISTER, C. A. (1975), «Credibility for regression models with application to trend», dans Credibility, theory and applications, Proceedings of the Berkeley actuarial research conference on credibility, Academic Press, New York, p. 129–163

24. JEWELL, W. S. (1974), «Credible means are exact bayesian for exponential families», ASTIN Bulletin, vol. 8, p.77–90
25. MACK T. (1993), Distribution-free calculation of the standard error of chain ladder reserve estimates, Astin Bulletin 23 (2), 213-225
26. MERZ M., M. WÜTHRICH (2008), Modelling the claims development result for solvency purposes, CAS E-Forum, Fall 2008, 542-568
27. NORBERG, R. (1979), «The credibility approach to ratemaking», Scandinavian Actuarial Journal, vol. 1979, p. 181–221
28. OLYMPIO A., RANAIVOZANANY V., WITTMER S. (2012), Gestion des risques d'entreprise : qualité des données, levier de pilotage stratégique, article congrès ERM AFIR Lyon juin 2012
29. OSTROW Jr. C.W. (1982), Time Series Analysis: Regression Techniques. Serie: Quantitative Applications in the Social Sciences, Ed. Sage Publications
30. PERRIN M., Calibration des Undertaking Specific Parameters et leurs impacts sur les fonds propres, Mémoire présenté devant le jury de l'EURIA en vue de l'obtention du diplôme d'Actuaire EURIA et de l'admission à l'Institut des Actuaires le 28 septembre 2012
31. PLANCHET F., P. THEROND (2011), Modèles de durée, Applications actuarielles, Ed. Economica
32. PLANCHET F., G. LEROY et M. JUILLARD (2012), Modèle standard, quelle utilisation des paramètres spécifiques à l'entité, La Tribune de l'assurance, avril 2012, n°168
33. PLANCHET F., Segmentation tarifaire et suivi du risque en incapacité, Journées d'étude IA / SACEI du 17 septembre 2009
34. PLANCHET F., P. THEROND, KAMEGA A. (2009), Scénarios économiques en assurance modélisation et simulation, Ed. Economica
35. PLANCHET F., Q. GUIBERT, M. JUILLARD (2010), Un cadre de référence pour un modèle interne partiel en assurance de personnes : application à un contrat de rentes viagères, Bulletin Français d'Actuariat, Institut des Actuaires, 2010, 10 (20), pp. 5-34. <hal-00530864>
36. SANDERS D.H., A.F. MURPH, R.J. ENG (1984), Les statistiques : une approche nouvelle. Ed. MacGraw Hill
37. WHITNEY, A. W. (1918), «The theory of experience rating», Proceedings of the Casualty Actuarial Society, vol. 4, p. 275–293.

Site internet :

https://acpr.banque-france.fr/autoriser/procedures-secteur-assurance/solvabilite/parametres-propres

Annexes

Annexe 1 : Exemple du calibrage du niveau de choc de maintien (rétablissement) sur un portefeuille disposant d'un historique entre 1973 et 2010 et choix du modèle (série chronologique)

La base de sinistrés des contrats du partenaire sélectionner comme test dispose d'un historique de profondeur [1973,2010] et de volumétrie suffisante de données.

Pour chaque sinistre on dispose des éléments suivants :

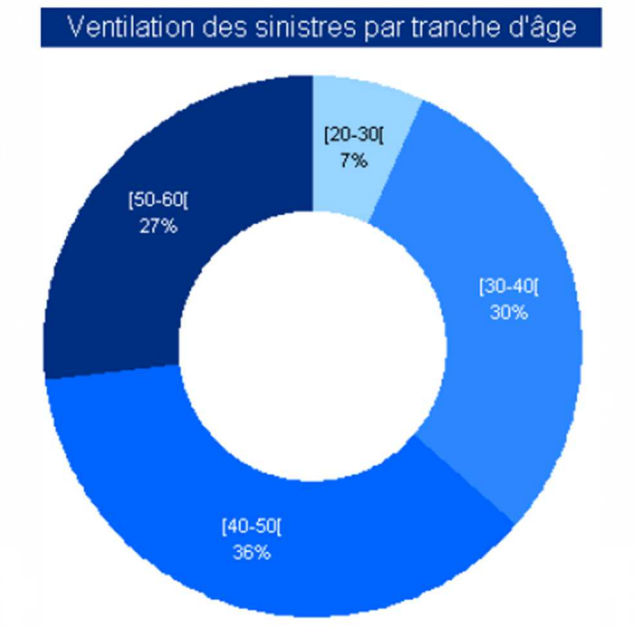
- Âge de l'assuré au moment de l'entrée en arrêt de travail,
- Durée de l'arrêt de travail (pouvant excéder 36 mois),
- Nature de l'information : censurée ou pas,
- Année de survenance du sinistre.

L'objectif est de choisir le modèle parcimonieux permettant d'estimation d'un choc de maintien « *Entity Specific* » *Solvency II* au niveau 0,5%, calculé directement sur les taux de rétablissement annuels.

1) Quelques statistiques sur la base de données étudiée :

Années de survenance observées : de 1973 à 2010

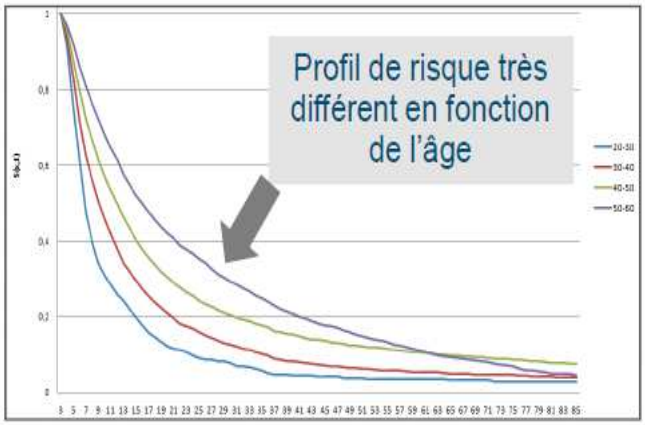
- Pour garantir une bonne robustesse d'estimation, nous avons utilisé des données correspondant à la plage [1983,2007].
- Âge des sinistrés compris entre 20 et 60 ans.
- Âge moyen de 43 ans.
- Nombre de sinistres total : 177 178 lignes sans doublons.
- Nombre de sinistres sur la plage [1983,2007] : 166 647 lignes



2) La construction des tables de maintien

Notons que les lois de maintien varient significativement selon les âges.
Ainsi :

- Nous avons considéré quatre catégories d'âges pour regrouper les sinistrés : [20,30[, [30,40[, [40,50[et [50,60[
- Le graphique ci-contre présente les tables de maintien estimées sur la plage [2003,2007] pour chaque tranche d'âges



Deux niveaux d'erreurs sont en général captés pour le calibrage de chocs sur les risques techniques :

- L'erreur d'estimation : écart entre l'estimateur et la probabilité de survie
- La *process variance* : évolution aléatoire de la probabilité de survie en fonction des années

En pratique pour chaque année, il faut estimer les taux de maintien et les variances associées.

Limite :
Seule l'erreur d'estimation peut être calculée en pratique. La prise en compte de la *process variance* nécessite de disposer d'un historique de données relativement profond (> 20 à 30 ans).

3) Evolution temporelle des lois de maintien :

- Choix de la fenêtre glissante

Zoom sur la taille de la fenêtre glissante : exemple d'un sinistre survenu le 01/06/1996 vu le 01/01/2005

- Estimation des lois de maintien avec l'estimateur de Kaplan-Meier

Construction via Kaplan-Meier de tables de maintien d'expérience par survenance sur [1987,2007] en considérant des fenêtres glissantes de 5 années de survenance

Soient a l'année d'observation, x l'âge ou la classe d'âges, t correspond à l'ancienneté dans l'état en mois. L'estimation du taux de maintien va dépendre de différentes hypothèses de calcul : la classe d'âge considérée, la taille de la fenêtre d'observation.

Pour chaque agrégat de survenances [a-4, a], on calcul l'estimateur de Kaplan-Meier :

$$\hat{S}(x,t,a) = \prod_{i=1}^{t}\left(1 - \frac{d_i^a}{n_{i-1}^a - d_{i-1}^a - c_{i-1}^a}\right)$$

avec n_{i-1}^a : nombre de présents (sur les données [a-4,a]) dans l'état d'incapacité au cours de la ième période, d_i^a : le nombre de sorties (sur les données [a-4,a]) de l'état d'incapacité au cours de la ième période, c_{i-1}^a : le nombre de présents (sur les données [a-4,a]) dans l'état d'incapacité ayant une ancienneté $i-1$ en l'état (correction de l'estimateur pour prendre en compte les censures c).

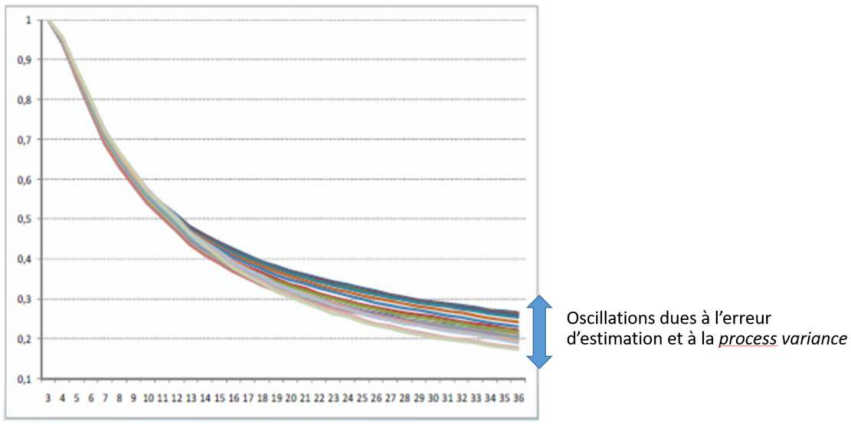
Cet estimateur généralise l'estimateur empirique de S en présence de censure. Un individu est censuré si son rétablissement n'est pas observé.

Propriétés de cet estimateur :
- il est convergent et asymptotiquement gaussien,
- sa variance est minimale et est donnée par l'estimateur de Greenwood :

$$\widehat{Var}\left[\hat{S}(x,t)\right] = \hat{S}(x,t)^2 \times \sum_{i=1}^{t_{max}} \frac{d_i}{n_i \times (n_i - d_i)}.$$

Les principales causes de la censure sont :
- le rachat du contrat par l'assuré,
- la sortie automatique de l'arrêt de travail à l'âge de 60 ans,
- la sortie non observée dans la fenêtre d'années d'observation (illustration ci-dessous).

Ci-dessous les lois de maintien (construites sur 5 années glissantes) des survenances [1987,2007] sur la tranche d'âges [40,50[:



4) Estimation du niveau de choc directement sur taux de rétablissement

- Le taux de sortie annuels de l'état d'incapacité par anciennetés et tranches d'âges

A partir des lois de maintien (construites sur 5 années glissantes) relatives aux survenances [1987,2007], on déduction des taux de rétablissement annuels conditionnels (sortie de l'état d'incapacité) par anciennetés et tranches d'âges :

$$R(x,t,a) = 1 - \frac{S(x,t+12,a)}{S(x,t,a)}$$

avec $R(x,t,a)$: probabilité (vue en « a ») de sortie annuelle de l'état d'incapacité pour un assuré en arrêt de travail depuis t mois, $\frac{S(x,t+12,a)}{S(x,t,a)}$: probabilité de maintien (vue en « a ») estimées sur les données [a-4,a].

- Le taux d'évolution, noté Z(x,t,a), d'une année à l'autre des R(x,t,a) :

$$Z(x,t,a) = \frac{R(x,t,a+1) - R(x,t,a)}{R(x,t,a)}$$

$$Z(x,t,a) = \frac{\frac{S(x,t+12,a)}{S(x,t,a)} - \frac{S(x,t+12,a+1)}{S(x,t,a+1)}}{1 - \frac{S(x,t+12,a)}{S(x,t,a)}}$$

- Estimation du niveau des chocs :

Pour tenir compte de l'autocorrélation de la série (Z(x,t,a)), a=1987,…,2006, nous faisons recours à un processus autorégressif :

$$Z(x,t,a+1) = \mu(x,t) + \alpha(x,t) \times Z(x,t,a) + \sigma(x,t) \times \varepsilon_a^{x,t}$$

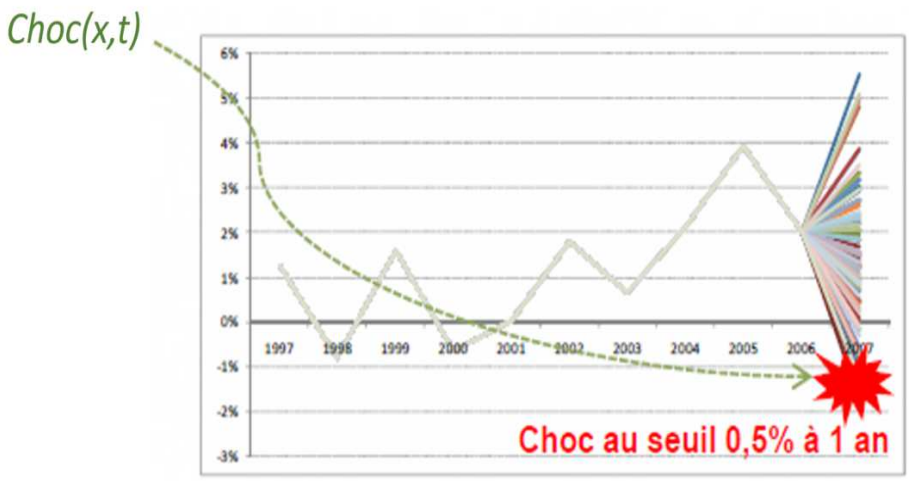
avec $\mu(x,t) + \alpha(x,t) \times Z(x,t,a)$: correspond au *drift*, $\sigma(x,t)$: la volatilité, $\varepsilon_a^{x,t}$:suit une loi normale centrée réduite notée *N(0,1)*.

Les paramètres du modèle sont estimés par MCO (Moindre Carré Ordinaire).

Nous obtenons le choc à 1 an au seuil 0,5% en utilisant le quantile à 0,5% d'une loi *N(0,1)* :

$$Choc(x,t) = \alpha(x,t) \times Z(x,t,A) + \sigma(x,t) \times q_{0,5\%}(\varepsilon)$$

avec A la dernière survenance considérée, i.e. A=2006.



- Durée moyenne de l'arrêt de travail et niveau de choc :

En notant Tx la durée de l'arrêt de travail pour la tranche d'âges « x », l'espérance de maintien en incapacité correspond à E[min(Tx,36)] (car la durée maximale couverte en arrêt de travail pour l'état d'incapacité est de 3 ans, donc 36 mois)

5) Sensibilité des résultats au type de modèle de séries temporelles

L'objectif de cette étude de sensibilité est de mesurer de la robustesse du choc estimé par modélisation AR(1) de la série Z en recourant à des séries temporelles d'ordres supérieurs ARMA(p,q).

Suivant les critères Akaïke (AIC), il ressort de l'étude que le processus ARMA(2,1) est celui qui s'ajuste le mieux aux séries étudiées. La structure d'une modélisation ARMA(2,1) appliquée à Z :

$$Z(x,t,a) = m(x,t) + \alpha_1(x,t) \times \left(Z(x,t,a-1) - m(x,t)\right) + \alpha_2(x,t) \times \left(Z(x,t,a-2) - m(x,t)\right)$$
$$+ \beta_1(x,t)\varepsilon_{a-1}^{x,t} + \varepsilon_a^{x,t}$$
$$Avec \ \ \varepsilon_a^{x,t} \approx N\left(0, \sigma(x,t)^2\right)$$

Finalement, nous retiendrons le processus AR(1) est celui qui revêt un caractère prudent.

Chapitre 3:
Long-Term Care: Construction of an Economic Balance Sheet and Solvency Capital Requirement Calculation in Solvency II

CHAPITER 3: LONG-TERM CARE: CONSTRUCTION OF AN ECONOMIC BALANCE SHEET AND SOLVENCY CAPITAL REQUIREMENT IN SOLVENCY II

1. Introduction

The Solvency II Directive, adopted in 2009 by the Council of the European Union and the European Parliament, officially became effective on January 1, 2016. Developed to improve the assessment and control of risks, it modifies in-depth the prudential guidelines applicable to insurance and reinsurance companies present in the European Union.

Long-Term Care insurance is a particular branch in terms of insurance risk. The borderline between dependence and health problems is porous, since limitations of autonomy often result from current or past health conditions. Like health insurance, private Long-Term Care insurance in Europe supplements the benefits or guarantees offered by public plans. In France, "Long-Term Care" guarantees marketed by insurers have their own risk assessment system for dependence. Benefit amounts and their payment are independent of benefits paid by social insurance benefits. Benefits offered by the various plans that exist on the market may be of a different type. Some insurance products of group pension plans provide yearly Long-Term Care coverage in their contract. During the transition to retirement the contract can continue optionally. The majority of individual or group contracts with optional participation offer lifetime benefits. At the end of a period defined in the contract, premium payment lapse no longer causes termination of the contract, but its reduction. The contracts offer cash benefits, the amount of which is defined and does not provide for the reimbursement of the cost of care, an amount which may evolve over time.

Pricing and reserving of the contracts depend naturally on the type of coverage (lifetime or yearly) and the type of data available when premium rates are calculated. If yearly coverage can be priced and accrued from tracking of the ratio of incurred claims to premium issued, the rate of lifetime benefits will require a significant number of assumptions. Where available data permit, insurers rely on past claim experience to rate their death or disability benefits.

In the case of Long-Term Care, the very long-term nature of the proposed guarantees makes the collection of these data spread over time and causes the need to take into account other factors. Indeed, the Long-Term Care risk is linked to many social parameters:

- The evolution of family unit, lifestyle and intergenerational bounds.
- The evolution of medical science, daily living, access to care.

Assumptions made in the establishment of premium rates involve the mortality of the insured population, "active state", the likelihood of becoming dependent depending on different levels of dependence called "dependent states", continuance in the dependent states, and transition probabilities between states. The complexity and number of these assumptions mean that insurers cannot rely on the validity of their premium rates. The time dimension takes a very fundamental role in this type of coverage.

Monitoring and control of the coverage is indispensable in a world in perpetual evolution. So, if the life span is lengthened in France for men and women since 2006, the latest studies of DRESS[7] indicate that the healthy life expectancy (without disability) has stagnated for 10 years (2006–2016).

[7] Study 1046-The French live longer, but their healthy life expectancy remains stagnant (Jan. 2018, French). Solidarity and health report N ° 22- elderly people in institutions (2011, French). For DREES, see notes 11 in 'The Long Term Care Risk' chapter.

The most common diagnostics encountered in retirement homes are neuropsychiatric disorders according to the DREES report 22,1 in 2007, 36% of EHPAD[8] residents suffered from Alzheimer's.

Medical advances in the treatment of this condition could therefore have the effect of a rapid change in the likelihood of dependence or on the contrary certain lifestyles could result in an increase in the occurrence of certain ailments and would have an impact many years later. For all these reasons, regular monitoring and adjustments of the guarantees and/or benefits of Long-Term Care contracts are indispensable throughout the duration of the coverage.

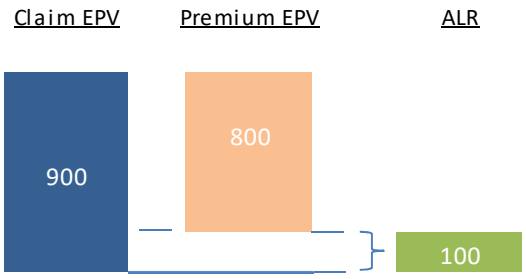
The calculation of the capital requirement as envisaged in Solvency II provides for the measurement of the impact on the sufficiency of a company's capital from a set of adverse events over a one-year horizon to a probability corresponding to the 99.5% quantile.

The long-term nature of the coverage offered by a Long-Term Care contract and the consequent significant delay between the collection of the first premium and the payment of claims naturally leads to a high potential cost of the mere fact that the projection framework foreseen in the case of Solvency II provides for the maintenance of the unfavorable situation until the policies are matured. Long-term guarantees in the case of Long-Term Care insurance thus see their unfolding experience impacted for a long time by adverse actual to expected claims.

In addition, the long-term character leads to a greater dependence of profitability on the interest rate environment. The financial soundness becomes an issue when several years or even a few decades elapse between collection of premiums and benefit payment.

For example, we can highlight the intrinsic leverage effect to lifetime products. Suppose the portfolio's Active Life Reserve (ALR) is €100m. It decomposes into:

- a premium Expected Present Value (EPV) of €800M
- and claim EPV of €900M.



Assuming that the Claim EPV can be written as

$$Claim\ EPV \sim \alpha \cdot Benefit^{\ Annual} \cdot duration^{dependent\ state}$$

If the average duration of the benefit period is 4 years, then a 20% decrease in the mortality of the dependents, all things being equal otherwise, would increase the duration by 0.8 years.

This would represent an estimated cost of around €180M [(.8 /4) x 900M], or 1.6 times the central ALR [180 / (900-800)], using a zero reserve interest rate.
As we have just seen, in the case of Long Term Care, an event occurring today can have very important effects much later. Thus, a 15% longevity increase per year that would be maintained over time has a major impact on the soundness of a Long Term Care insurance product.

[8] Établissement d'hébergement pour personnes âgées dépendantes, or Establishments to accommodate dependent elderly. Since 2002, a national law defines and regulates these types of nursing homes

However, insurers monitor their risk. Thus, once the risk has been identified the insurer has a set of possible measures, for example:

- Some group contracts may be terminated by the insurer.
- If contractual dispositions permit, revising the premium of the insured upward to compensate for the deterioration of the risk.
- The value of the benefits paid for reduced contracts or in the event of a premium lapse for current contracts, in accordance with the contractual dispositions, may be revised downward.
- Pricing of new contracts may include a prudential margin to ensure a return to the underwriting soundness under the principle of mutualization.

The framework provided by the Solvency II standard does not allow the latter mechanism to be taken into account since new cases are not part of the projection framework. Other measures can be modeled and taken into account when calculating projected cash flows if there is a strategy described and validated by the board of directors to account for this type of situation. If this strategy is implemented in calculating the solvency of the company, a regular report of its impact on the results and its occurrence in the projections must be performed (in accordance with article 236 of the Commission Delegated Regulation (EU) 2015/35).

Long Term Care is a newly recognized risk, increasing with age and poorly known to insurers, they generally reinsure using coinsurance for several reasons:

- Risk is not well understood due to lack of market data (relatively new risk) and lack of regulatory tables.
- The insurer's commitment can be very high due to the length of the coverage (long-term) and requires a significant amount of solvency capital.
- The impact of coinsurance on risk capital reduction are relatively easy to take into account

In terms of reinsurance-induced reduction of solvency capital, there is no ceiling under Solvency II. So, reinsurance can be fully taken into account to reduce the risk.

However, contractual clauses must not be omitted. Indeed, the insurer must take account to cessation clauses, profit-sharing clauses, and asset transfer provisions in the best estimate of reserves valuation.

2. Main principles for the economic balance sheet calculations

2.1. Summary of Solvency II provisions

Solvency II is based on three pillars:

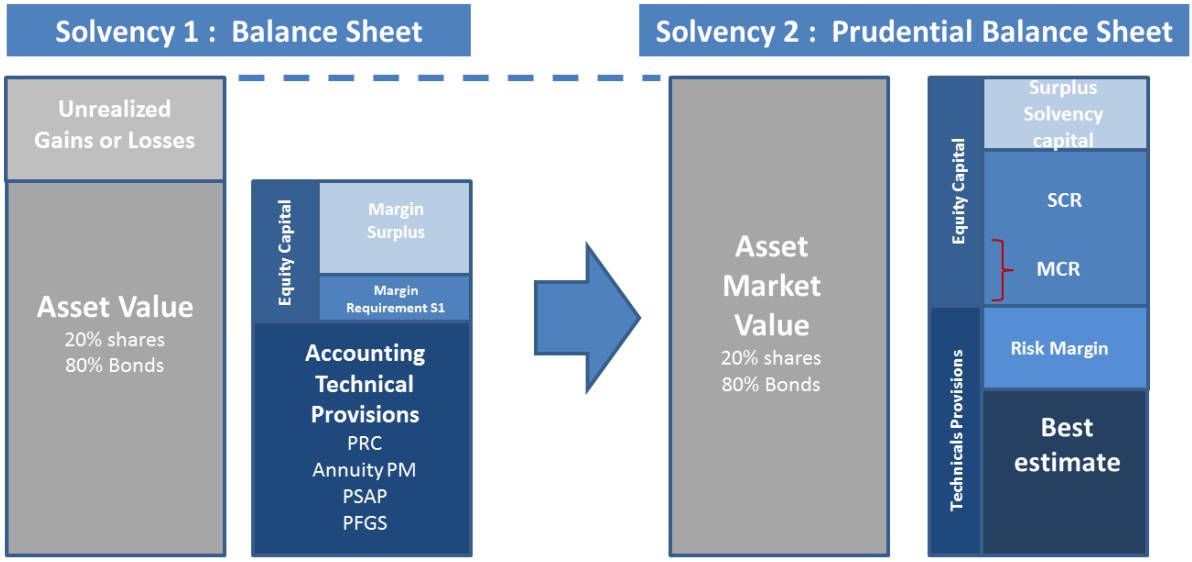
- **Pillar 1** contains quantitative financial requirements, for both balance sheet and solvency.
- **Pillar 2** deals with more qualitative aspects, such as governance and risk management in the broad sense, but also the role of the regulator and the regulatory process.
- **Pillar 3** describes the regulatory reports and information to be published.

The calculation of capital requirements, reserves in the sense of Solvency II and the contribution of Long Term Care insurance contracts to eligible equity for the purposes of capital requirement will be detailed in the following section.

2.2. Elements of an economic balance sheet standard

The economic balance sheet as defined in the Solvency II Directive is based on the concept of the best estimate of the value of balance sheet entries. The best estimate is obtained when the **different entries in the balance sheet are valued at their market value**. When such a market value is not directly available or observable in a reasonably liquid market, the valuation must be based on a model that uses the information available in the markets in an optimal and coherent manner, or on an approach based on the valuation of the income generated. The market value is thus considered as the most relevant indicator of a realistic economic value at any time.

Tab. 1. Statutory Balance sheet and Prudential Balance sheet



The above diagram shows the main differences between the Solvency 1 and Solvency II balance sheets for a typical Long Term Care insurer.

In French statutory accounting, assets are recorded at book value, i.e. on a value based on the depreciated acquisition price. Inversely, under Solvency II, assets are estimated on the basis of their market value at the valuation date.

Under Solvency 1, reserves consisting mainly of Active Life Reserves (ALR) and Claim Reserves (CR) shall be determined from:

- Characteristics of current policyholders and their contracts.
- Generally prudent experience tables due to the lack of regulatory tables and the limited data available.
- A conservative non-life valuation interest rate.

The accounting provisions correspond to the value of the insurer's guaranties toward the insured. Under Solvency II, the reserves shall consist of:

- The Best Estimate (BE) which represents the economic value of future flows relating to contracts in stock as of the valuation date.
- The Risk Margin (RM) which creates a supplement added to BE.

Reserves incorporate all the liabilities of the insurer engendered by the contracts: these include, in particular, beyond policyholder guarantees, liabilities to third parties (commission paid to the sales forces, various costs paid to Third Party Administrators, wages and taxes arising from the contracts...).

The calculation of the capital eligible be applied to the capital requirement under Solvency II derives from the economic balance sheet. In fact, the total market value of assets decreased by reserves, potential deferred tax liabilities and the value of any other low-balance liabilities makes it possible to obtain the elements allowed to be included in the capital before application of any limits arising from the classification rules related to Solvency Capital Requirement (SCR) or Minimum Capital Requirement (MCR).

The balance sheet based on market values is a "snapshot", a representation at a precise time of the financial situation of the organization.

This is subject to a range of risks, such as the underwriting risk (or insurance technical risk) or the market risk (or investment). To ensure that the organization will be able to fulfill its commitments with policyholders, the Solvency II framework sets out to **identify these risks and measure their effects on the balance sheet** and on the eligible capital.

Under Solvency II, the effects of significant quantifiable risks are determined and aggregated within SCR to estimate the cost of the maximum loss arriving once every 200 years. MCR corresponds to the amount of Solvency II capital under which the organization is not authorized to go below in order to continue its activities (withdrawal of approval by the [regulator] [French insurance regulatory authority, *Autorité de Contrôle Prudentiel et de Résolution*, ACPR).

The Solvency II framework offers the possibility of estimating these capital requirements either by the so-called standard formula or by setting up an Internal Model (IM) by the insurer.

In the following, we will assume that the standard formula option has been retained. The use of an internal model could indeed be interesting on the perimeter of Long Term Care risks. However, this approach requires a large amount of data and a history sufficient enough to be able to observe the experience of insureds at extreme old age. As these conditions are difficult to meet for this type of risk, the point is not dealt with in this chapter.

The standard formula provided for in Solvency II requires calculating the impact on the company's capital of a set of risks corresponding to the financial, technical and operational risks to which an insurance undertaking is subject. Aggregation of estimated cost of these risks using diversification factors allows the estimation the capital cost of the 99.5% quantile for the SCR calculation.

3. Description of Solvency II balance sheet basics

3.1. Calculation main elements

Assets

For the balance sheet assets, or at least some of them, market prices directly observable are available: this is the case of sovereign bonds or quoted stocks. For other investments, valuation models may be necessary. For real estate, the value derived from recent assessments can be used. Reinsurer reimbursements are shown as asset and reinsurance is not deducted from liability reserves, which improves the transparency of the process and of the *Reporting*.

Reserves

Under Solvency II, the concept of coherence with the market extends to reserves in principle, but also for a question of coherence in the construction of the balance sheet. However, there is no liquid market that would provide directly observable market prices for insurance portfolio transfers. Under Solvency II, the value of reserves is conceived as the sum of two elements

$$S2\ Reserves = Best\ Estimate + Risk\ Margin$$

The Best Estimate (of the current portfolio of liabilities) is mathematically the expected value of future payments flow less future income flow. These are estimated using realistic probabilistic assumptions about the risk factors that may affect these future flows and updating them with the risk-free rate curve relevant at the valuation date. More simply, the Best Estimate represents the expectation of the cost to the organization to pay its contractual commitments as expected on the expected due date, while taking into account premiums yet to be received. In calculating the best estimate, the underlying assumptions used correspond to the best estimate of future flows and should not contain additional risk margins. Another difference with the statutory reserves in French regulation is that the discount rate applies to all reserves, using the market rate curve without regard to the particular risk. For Non-Life organizations, a separate Best Estimate must be calculated for Active Life Reserve and Claim Reserve.

The contract limit defined in the standard formula (Commission Delegated Regulation (EU) 2015/35 article 18) defines the projection horizon of premiums for inforce contracts (boundary). In the event that the insurer has the unilateral right to refuse the premium, premiums subsequent to the date on which that right may be exercised are not taken into account in the projection. This situation can occur in the case of yearly cancellable group insurance contracts. In this case the insurer is freed from its commitments relating to contracts which are not in claim status.

For lifetime benefit contracts, the insurer's ability to revise the premium rate can lead to limiting the projection of premiums. However, the relevant acts provide that the projection of premiums stops when[9]:

"(c) the future date where the insurance or reinsurance undertaking has a unilateral right to amend the premiums or the benefits payable under the contract in such a way that the premiums fully reflect the risks.

Point (c) shall be deemed to apply where an insurance or reinsurance undertaking has a unilateral right to amend at a future date the premiums or benefits of a portfolio of insurance or reinsurance obligations in such a way that the premiums of the portfolio fully reflect the risks covered by the portfolio.

However, in the case of life insurance obligations where an individual risk assessment of the obligations relating to the insured person of the contract is carried out at the inception of the contract and that assessment cannot be repeated before amending the premiums or benefits, insurance and reinsurance undertakings shall assess at the level of the contract whether the premiums fully reflect the risk for the purposes of point (c)."

Depending on the contract, internal limits to the company in terms of premium revision policy or the inability to know precisely the fair cost of individual risks could lead to the view that premiums must be projected on a lifetime basis.

This approach, depending on the contracts, may be prudent and probably leads to a price closer to that which could constitute a cash value of the contract (or exit value). In fact, a premium lapse may result, depending on the contracts and their duration, in releasing the insurer from its contract or in a

[9] Official Journal of the European Union, Regulation 10/10/2014, Article 18 3 (c)

reduction in benefits. These situations may create margins for the insurer without connection to the likely future experience of the contracts.

Therefore, the limit of Long Term Care contracts depends on the nature of the contract and the interpretation of the regulation. This is a major issue that has a significant impact on the value of reserves and the risk profile within the meaning of Solvency II.

A risk margin is added to the Best Estimate to obtain a value of the reserves consistent with the market. It is calculated using the "Cost of Capital" method (CoC). This method is based on the idea that the organization (or those who finance it) hopes to be paid for the risks taken. The method can be summarized as follows: Suppose the organization wants to transfer or acquire an existing portfolio, the assignee will not accept the transfer if the value of the assets transferred with the portfolio is equal to the Best Estimate. Actual outgoing future payments can significantly defer, and even exceed, future flows associated with the Best Estimate, and the assignee will not accept the risk without an expected positive yield. The assignee will therefore request an additional amount (in addition to the Best Estimate), the risk margin.

The amount of the risk margin may be determined as follows: The organization must not only hold reserves; it must also have a risk capital to absorb unforeseen losses and comply with the regulations. The organization (or those who finance it) wants to be paid for the capital asset. The risk margin is simply equal to the product of the capital required and the cost of capital. The required capital used in calculating the risk margin contains only the SCR inherent in the current insurance portfolio; it also includes the operational risk and the default risk for the transferred reserves. The important thing is that the "avoidable" market risk, because it depends on the investment policy of the insurer and not on its insurance business, not to be included. The amount of the risk margin therefore depends on the characteristics of the insurance portfolio. The risks that can be covered in the markets are not taken into account in the risk margin.

The cost of capital is set at 6% (Commission Delegated Regulation (EU) 2015/35, Section 3, Article 39). The Best Estimate will suffice, on average, to exactly ensure the run-off of the portfolio. The risk margin will then be released gradually and will be sufficient to ensure a 6% return (beyond the risk-free rate) for those who fund the activity. Only the excess yield beyond the risk-free rate is included in the risk margin. The risk capital made available to the organization will be able to be invested in products which earn the riskless interest rate at the time.

Risk segmentation

For reserve calculation, insurance and reinsurance liabilities must be segmented at a minimum for each activity line. This segmentation applies to both the Best Estimate and the Risk Margin. Finer segmentation in homogeneous risk groups can be used if it improves the accuracy of the reserve and capital cost valuation. Guarantees are allocated to the segment of activity that best represents the nature of the underlying risk.

It should therefore be noted that this segmentation may differ from the branches of activity as defined to reach agreements or other accounting elements. For example, guarantees that are technically treated as life contracts (reserving made with assumptions of underlying distributions as opposed to a claim run-off triangle approach or LR ratio) should be considered Life coverage, even if they arise from non-life contracts. Conversely, certain life insurance contracts could give rise to non-life liabilities if the risk selection rules are similar to those of non-life risks. In particular, annuities paid under non-life contracts are considered life insurance.

Depending on the technical nature of the risks, health (in the meaning of Solvency II i.e. all coverage of lifetime care, except for death) is split into two types of coverage: similar to life health (so-called Similar Life Techniques, or SLT, Health), based on techniques similar to that used in life insurance; and

not similar to life health (Non- Similar Life Techniques, or NSLT, Health) in accordance with article 55 of the Commission Delegated Regulation (EU) 2015/35.

Discounting

As part of the Best estimate evaluation, cash flow discounting is carried out at risk-free interest rate: adjusted swap rate curve adjusts to the credit risk (provided by EIOPA), with the possibility of:

- Extrapolate the rate curve to evaluate long term coverage.
- Take into account a Volatility Adjustment to the risk-free rate curve or a Matching Premium on this rate curve (not possible to combine the two). These adjustments help to confront the volatility of the equity induced by the liquidity of assets backing the insurance coverage[10].

Contract maturity

As part of the determination of the incoming and outgoing flows constituting the BE, the strong hypothesis adopted by the Directive is to value the contracts to maturity while assuming the continuation of the insurer's activities in determining projection assumptions. This economic vision is a very different vision of the insurer's future liabilities, depending on its ability to cancel its contracts.

Capital Requirements: SCR, MCR

Solvency II provides for two different capital requirements, the SCR (Solvency Capital Requirement) and the MCR (Minimum Capital Requirement). Although it is not possible to compare them exactly with solvency 1 requirements, the role of SCR can be roughly compared to the solvency margin requirement and MCR to the Guarantee Fund (1/3 of the margin requirement).

As in Solvency 1 guideline, MCR will also be subject to an absolute floor (Absolute Minimum Capital Requirement, or AMCR).

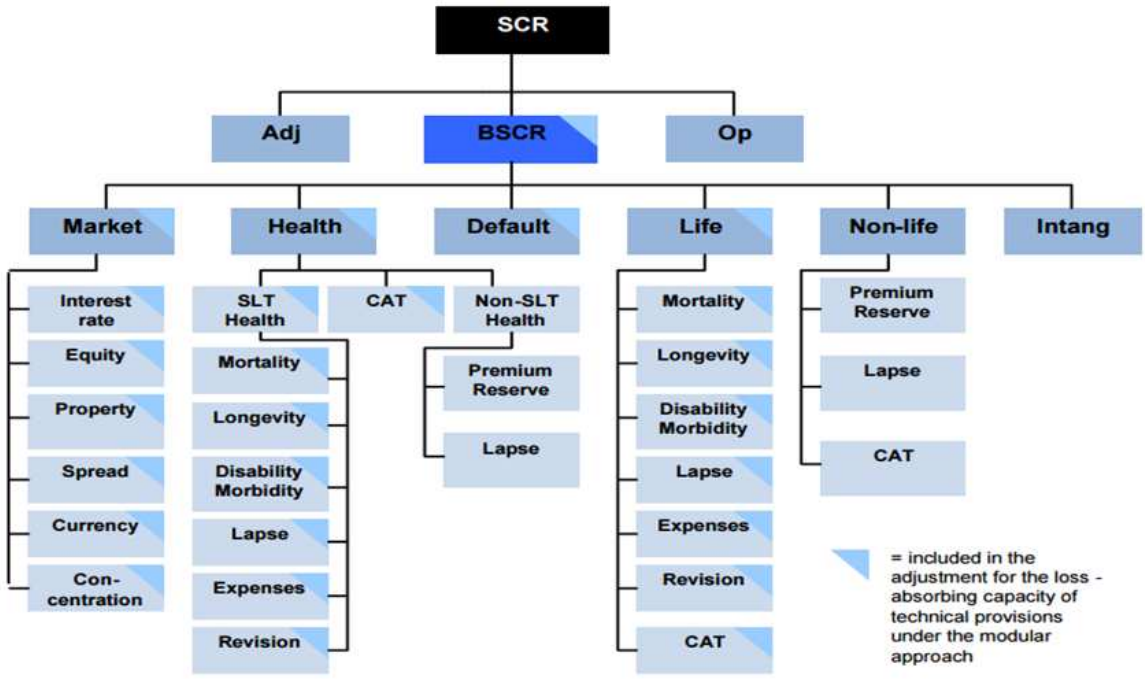
- **SCR**: The SCR is the regulatory capital to be held by an insurance organization to limit the insolvency of the underlying risk to a maximum of one in 200 over a one-year horizon. Insurers have the possibility to calculate the SCR either by using the standard formula (the use of a methodology and calibration parameters are described in the technical specifications) or by using an internal model.
- **MCR**: This is the minimum level of capital that the organization must hold permanently, under penalty of immediate action by the regulatory authority which may lead to withdrawal of approval.

[10] Taking into account an adjustment to the volatility of the risk-free rate curve impacts the Best Estimate valuation, the impact is neutral on the SCR. For the calculation of the risk margin, the rate curve is taken without adjusting for volatility.

3.2. Block Structure for calculating economic capital in the standard formula

The SCR of an insurance organization corresponds to the aggregation of all the following standard risks:

Tab. 2. Analysis Grid of Solvency II risks



The SCR calculation in the standard formula is expressed as follows $SCR = BSCR + Adj + Op$.

The basic SCR (BSCR) corresponds to the gross SCR of the absorption capacity of losses by profit sharing and deferred taxes. It is obtained by aggregating the SCR associated with the 6 standard modules (Market, Health, Counterparty Default, Life, Non-Life, Intangible)

$$BSCR = \sqrt{\sum_{i,j} Corr[i,j].SCR_i.SCR_j} + SCR_{intangible}.$$

This BSCR takes into account the correlation coefficients between risks (i.e. between SCR), defined by the correlation matrix Corr (annex IV, point 1) of Directive 2009/138/EC):

Tab. 3. Correlation Structure

SCR GLOBAL	Market	Default	Life	Health	Non-Life
Market	100%				
Default	25%	100%			
Life	25%	25%	100%		
Health	25%	25%	25%	100%	
Non-Life	25%	50%	0%	0%	100%

4. Application to dependence risk

4.1. Classification of Long Term Care insurance in Solvency II standard

Long Term Care is a perfect illustration of the difference in classification between the French standards and the standard Solvency II.

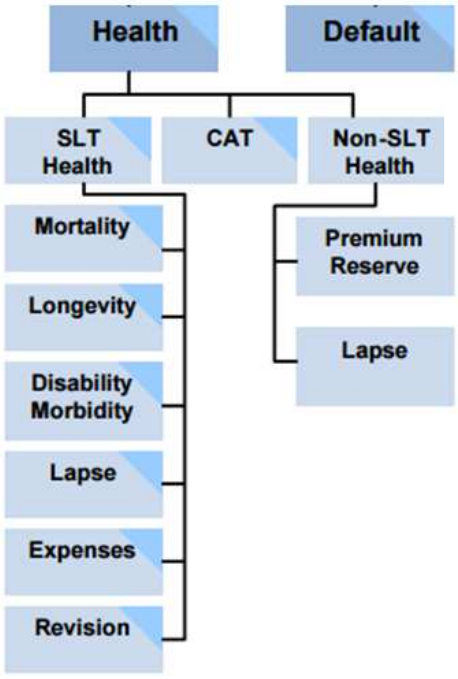
In French standard, Long Term Care can be considered as a non-life risk because the covered hazard is not life nor a life risk.

In Solvency II, Long Term Care is classified as health because it is a contingency risk different from death. Within this module, Long Term Care can be classified into two sub-modules depending on the nature of the coverage and the technical basis:

- Sub-module « Health Similar To Life (SLT)» In the event that the Long Term Care coverage is permanent, this requires modeling close to techniques similar to life insurance.
- Sub-module « Health Non-Similar To Life (non-SLT) » In the event that the Long Term Care coverage is yearly renewable.

Hence, the risks of the corresponding standard formula are as follows:

Tab. 4. Risk Identification



4.2. Characteristics of a Long Term Care insurance and assumptions used

In the case of Long Term Care insurance under the "Non-SLT Health" line, the SCR calculation would be based on a fixed cash benefit and depends on Best Estimate (BE) premium and reserve. These contracts fall under the income protection category. The calculation of the SCR in this case results from the simple application of coefficients provided for in the standard formula, the case considered in the following is that of a Life-like product.

For a given Long Term Care insurance product, the calculation of each SCR in the "Health SLT" module requires calculating a "shocked" BE, i.e. a BE with adverse projection assumptions. This methodology requires modeling the Long Term Care risk: we will therefore define a Long Term Care product with its modeling framework.

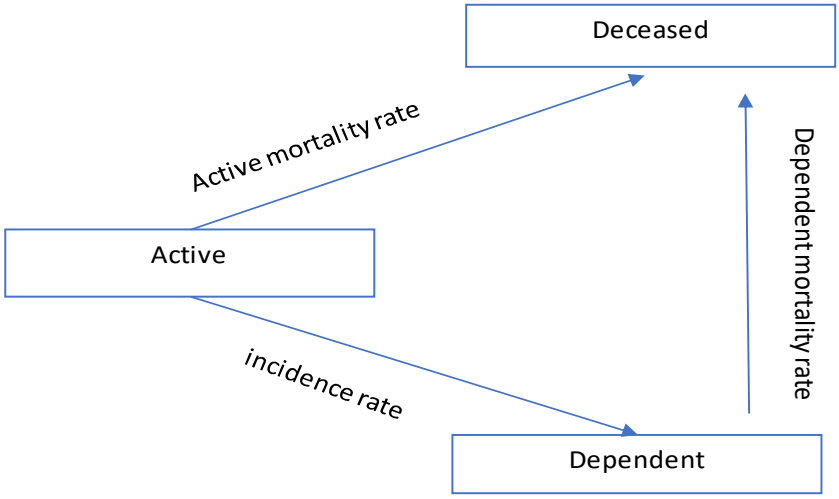
The choice is for a lifetime product which, therefore, does not result in a Health Non-SLT SCR. The modelled product provides for the payment of a lump sum and a monthly annuity in the case of Partial Dependence as recognized by a medical consultant, according to the Iso-Resource Grid. The recognition of a state of Total Dependence[5] would double the amount of the annuity provided in the case of Partial Dependence. The premium is monthly and depends on the issue age.

In the modeled product: each year, an active policyholder can remain active, become dependent, or die. A dependent is supposed to remain dependent or dies.

In this scheme, entry into a state of Partial or Total Dependence is considered definitive without possible recovery; the condition of an individual can only deteriorate. Indeed, the reversibility of the insured's dependent status is a real modeling challenge, linked to the lack of data. It differs from the temporary inability to work and permanent disability with an improvement of the insured's health. The simplifying assumption in this product may be questioned in accordance with the definition given to the Partial Dependence state.

The diagram below illustrates the transitions between the different states of the product:

Figure. 1 **Description of the states**



Finally, we note that the SCR calculations presented below do not take into account the specific treatments related to reinsurance. We do not model treaties in this chapter, and the interested reader is referred to the reinsurance chapter.

4.3. Actuarial modeling of a Long Term Care insurance contract

In the context of Solvency II, it is necessary to have a deterministic model of cash flow projections allowing to dynamically evaluate the solvency of the contract, incorporating the different assumptions. These different streams make it possible to project the balance sheet over several years, to estimate the *Best Estimate* reserves and calculate the economic capital of the contract.

Simplified projection model

Generally speaking, for a given insured with age X at the valuation date, the cash flow projection is described by a system describing successively:

- The number of active premium-paying actives $N_{x,t}^{a} = N_{x,t-1}^{a} \times (1 - q_{x+t-1}^{a}) \times (1 - i_{x+t-1})$.
- Premium income $P_{x,t} = \pi_{x,t} \times N_{x,t}^{a}$.
- Incidence $C_{x,t} = i_{x+t-1} \times N_{x,t-1}^{a}$.
- Claims $N_{x,t}^{d} = \sum_{k=0}^{t-1}\left[C_{x,k} \times \prod_{s=0}^{t-k-1} 1 - q_{x+k,s}^{d}\right]$.
- Finally, benefits (annuity payments can vary depending on a parameter or can remain constant) $B_{x,t} = r_{\theta} \times N_{x,t}^{d}$.

The preceding equations allow to calculate immediately the expected cash flows by the following formula $\rho(t) = \sum_{x}\left(P_{x,t} - B_{x,t}\right)$.

Thus, on any date T, a projection of the Best Estimate Liabilities (BEL) is given by the following formula

$$E[BEL(t)] = \sum_{s \geq t} \frac{\delta(s)}{\delta(t)} \times \rho(s),$$

with $\delta(t) = (1 - r_t)^t$ the discount factor (deterministic).

As $BEL(t) = E_t\left(\sum_{s \geq t} \frac{\delta(s)}{\delta(t)} \times \tilde{\rho}(s)\right) = \sum_{s \geq t} \frac{\delta(s)}{\delta(t)} \times E_t\left(\tilde{\rho}(s)\right)$, we reach the equality

$$E[BEL(t)] = E\left[E_t\left(\sum_{s \geq t} \frac{\delta(s)}{\delta(t)} \times \tilde{\rho}(s)\right)\right] = \sum_{s \geq t} \frac{\delta(s)}{\delta(t)} \times E\left(\tilde{\rho}(s)\right) = \sum_{s \geq t} \frac{\delta(s)}{\delta(t)} \times \rho(s).$$

In the context of Solvency II, to dispose of these cash flows allows to calculate the margin requirement on each date, applying to the calculation assumptions the shocks at the different dates to deduct, by measuring the Best Estimate variation, the required capital as part of the standard formula. It should be noted that when the Best Estimate is negative, which corresponds to anticipation of future gains, the insurer cannot use them as shock reduction. It will be appropriate to add the different loads, commissions and expenses to calculate the BEL value.

The annual approach presented here for simplification purpose should be taken into account in the drafting of distributions if it is retained. Indeed, given the nature of the risks, the order in the application of the probabilities has a significant impact on reserve values.

Finally, the need for unit capital in the event of a shock having no impact on the value of the assets is therefore of the form $SCR = [BE_{choc} - BE_{central}]^{+}$.

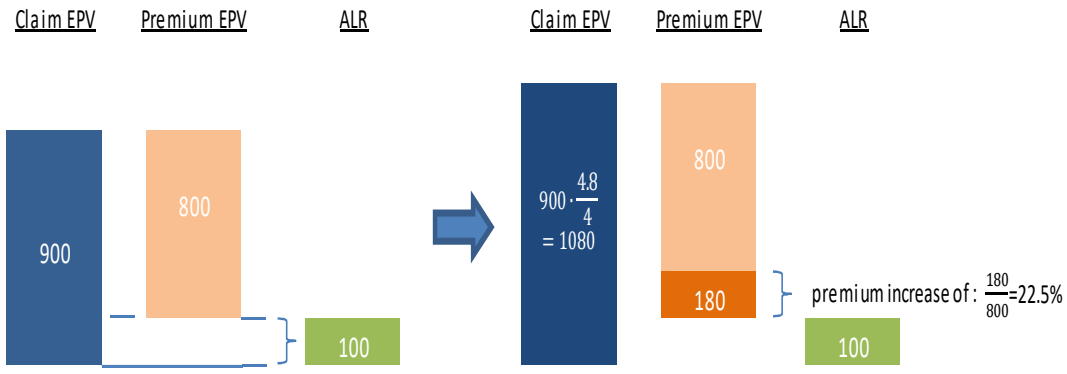
The calculation of a Long Term Care insurance product SCR requires three calculations for underwriting, counterparty, and market risks. The counterparty risk arises when liabilities are transferred to a reinsurer or from certain types of assets held. Market risk would require modeling the company's general assets and, at the start, the cash flows from other marketed products. In the absence of these elements, the market risk was calculated for a single product, more for illustration than to estimate minimum capital. Future premiums are taken into account, which may eventually be revised as contractually planned. If the absence of a projection of premiums can be entertained for an existing portfolio, and in view of a transfer, in a pricing approach and to measure the required capital, the shareholder will wish to know its future remuneration from the perspective of retention to maturity.

Underwriting gain is not modeled in the basic scenario (insureds behave as expected in the premium rate calculation, except in the context of shocks), or administrative gain (premium loads are equal to expenses, except in the expense risk shock, and in the sensitivity scenarios.)

In the Solvency II standard formula, the prescribed shocks considered are as follows:

- Longevity risk (20% reduction in active and dependent mortality)
- Morbidity risk (35% incidence increase in first year and 25% for subsequent years). The 20% shock on either the recovery or the persistency rates should be taken into account only if, in the valuation of the best estimate, another option than death can lead to a change of state in case if dependency. This would be the case in multi states model for example.
- Expense Risk (10% expense increase).

This simplistic model would lead, in some stressed scenarios, to collect premiums that generate underwriting losses without ever revising premium rates, which is clearly unrealistic. As a result, it is important to propose improvements to the model to make it more realistic by taking into account certain risk reduction factors.
To illustrate the impact, we can look at the very simple example shown in the introduction section 1: The cost of a 20% decrease in the mortality of the dependent people has a cost of €180M. Assuming that the cost is transferred to the premiums, they should increase by 22,5%. Practically, the increase can be done on several years.



The main sources of improvement of the proposed model are:

- Taking into account revalorization and financial assets.
- Premium rate revision.
- And taking reinsurance into account.

Revalorization and financial assets

French Long Term Care Insurance contracts generally provide for the creation of a “revalorization fund", fed by possible gains (underwriting/financial), in order to revalorize the guarantees over time. It may be envisaged that the fund thus constituted should be resumed following a deterioration of underwriting results (linked to an “external shock”). The recovery of the established financial reserve has the effect of mitigating the impact shocks on the Best Estimate, and therefore reduces the required SCR.

The modeling of this measure for the SCR calculation implies projection, operational and management assumptions, which must be justified. It also passes through the explanation of the results on the projection horizon.

- **Projection assumptions:**

Increasing the revalorization fund by projecting underwriting gains is hardly justifiable. On the other hand, it is possible to add to this fund through the projection of financial gains and in this case be able to reach over time higher financial return rates than the reserve interest rate.

- **Operational assumptions:**

This assumes that the insurer has the means to detect an Actual to Expected (A/E) deterioration of underwriting results. This assumption is reasonable once the insurer implements a recurring system of risk monitoring better/worse than expected analysis to detect deteriorations in the risk.

- **Management assumptions:**

The assumption of revalorization funds recapture is justifiable if the insurer is contractually allowed to do so.

Financial gains will be projected in the framework of this model. Formally, this amounts to set up the following formula:

$$FR_t = \beta_t - IR_t \times (P_t + B_t) \times \left(ror_{fin} - i\_val\right)$$

$$FDR_t = \sum_{s=0}^{t} FR_t \times \left(1 + ror_{fin}\right)^{t-s}$$

with FR_t : financial result, β_t : financial asset income, IR_t : valuation interest income, P_t : Premium, B_t : Benefit, ror_{fin} : financial returns, i_val : valuation interest rate, FDR_t : revalorization fund.

The introduction of financial assets into the model logically decreases the SCR. Moreover, the establishment of a revalorization fund (increased by financial gains) used to mitigate underwriting losses in stressed scenarios, allows a steeper SCR decrease.

Modeling future decisions: premium rate revisions

Premium rate revision is a possible action in case of reserve deficiency and is generally foreseen in the general provisions of [French] Long Term Care insurance contracts. In the application of the standard formula, the projections made involve collecting premiums that generate underwriting losses without ever revising premium rates again, which is obviously unrealistic. According to the procedures for setting up premium rate revisions, all or part of the additional margin linked to the stressed scenarios is transferred to policyholders, thereby reducing the required SCR.

Modeling this provision for the SCR calculation implies operational assumptions which must be justified.

- **Detection of risk deterioration assumption:**

It is justified if a risk monitoring process and analysis of better/worse than expected analysis is set up to detect worsening risks.

- **Claim follow-up assumption:**

This consists in setting up ways to measure the magnitude of shocks on biometric assumptions (incidence, continuance and severity shocks) and management costs (claim expense shock). It is justified if the insurer defines a method of constructing/adjusting biometric assumptions in order to account for the changes in risks.

- **Premium rate revision assumption:**

This is justified if the insurer is contractually able to carry out premium rate revisions, taking into account any ceilings.

- **Implementation assumption:**

This deals with the time to identify and set up a risk monitoring and adjustment process. In addition to the strategies for identifying and measuring shocks, it is reasonable to make an assumption about the time (months/years) required for its implementation.

At least four categories of insured persons are to be taken into account:

- Premium paying actives.
- Future insured.
- Paid Up (limited premium) and Reduced Paid Up (lapsed premium) actives.
- Claimants.

The transfer of risk through premium rate revisions only occurs through the first two categories, whereas the shocks apply to all entire categories. The restoration of underwriting balance is limited to the additional premium that can reasonably be transferred to premium paying participants (current or future) and is sensitive to the maturity of the portfolio and assumptions about new issues. If these conditions are fulfilled, taking into account future decisions in the projection model may reduce the magnitude of continuance and incidence SCRs.
For this it is necessary to model:

- The time at which the insurer will be able to measure the risk and implement a premium rate revision.
- The decision on the magnitude and duration of premium rate revisions.
- The impact over time on the outcome of the implementation of the strategy.

The model will need to be able to handle two claim assumptions. One, called "experience" that allows to model the cash flow of pricing benefits (those that make up the Best Estimate); one, so-called "reserve", which allows the calculation of statutory reserves and corresponds to a distribution known to the insurer. The time required for the insurer to be able to measure the deterioration of premium rates can thus be estimated in the projection model by tracking the numbers of deaths or the incidence count.
The time limit can thus be calculated from the following elements:

- The period from which the difference between the expected numbers (calculated with the so-called valuation distribution) and the actual numbers ("Best estimate" distribution) exceeds an action threshold in relation to thresholds defined in the current management action policy.
- The time required for the insurer to analyze the data.
- The time limit to effectively implement the premium rate revision, the latter includes policyholder communication.

To determine the duration and value of the premium rate revision, an indicator that can be retained in the model is a deficient statutory reserve. The latter corresponds to the difference between the

statutory reserve calculated under the valuation distribution (or before shock in the case of the SCR calculation) and that calculated with the experience distribution (or shocked distribution in the case of an SCR calculation). It will then be necessary to determine the duration and the annual value of successive revisions. These elements are based on the maximum threshold of revisions practiced by the company or contractually practical. To estimate the amount of the premium rate revision to be used, the rate of increase to be applied to the probable present value insured in the ALR calculation allows to measure the revision to be applied to the net premium to restore balance at the statutory reserve level. The implementation of this increase will also have an impact on the future administrative margins. It will be appropriate to determine whether the insurer's strategy in the case of a claim deterioration is aimed to recover its margins or to cancel losses.

Once these elements are determined, the model must implement the revision by modifying the premium flows after the start of management action and recalculating the ALR taking into account these changes in order to establish the successive results and updating that result. The difficulty in implementing such a strategy often does not lie in modeling.

The question is to clarify precisely the strategy that would be applied in a deterioration situation and to ensure that the company governance provides:

- The effective monitoring of risk indicators, that these indicators be published, analyzed and communicated to management to make an informed decision.
- A clear policy of premium rate review practices incorporating both marketing and actuarial considerations.
- The possibility for management to make decisions about premium rate revisions over several years.

These conditions fulfilled, the premium rate revision policy implemented in the model will constitute a "management action" which must meet criteria laid down in the delegated activities (in accordance with the article 23 of the Commission Delegated Regulation (EU) 2015/35: management validation, proof of "use test" in the event of a real situation and consistency with internal policies, regular communication of the use of the strategy and its effect on the SCR).

Reinsurance

The Best Estimate must be calculated gross of reinsurance, without deducting liabilities arising from reinsurance contracts and securitization tools and without taking into account the amounts covered by reinsurance. In return, a reinsurance asset equal to the difference between the Best Estimate gross and net of reinsurance is recognized. This asset generates a capital charge for the counterparty risk. Moreover, in the case of the standard formula, the risk margin must be calculated net of reinsurance, which is consistent with the article 101-5 of the Directive stipulating that the SCR must be calculated net of reinsurance.

The reinsurance asset must be calculated using the same principles as the Best Estimate calculation by incorporating the probability of the reinsurer default. This consideration results in an adjustment of the probable present value of future cash flows which is calculated on the basis of the probability of default of the counterparties and the resulting average amount of losses.

When the counterparty benefits from high ratings, the matching default adjustment should be fairly low compared to reinsurance recovery. In this case, the adjustment for the counterparty default can be evaluated in a simplified manner by the following formula

$$Adj_{CD} = -\max\left((1-RR)\times BE_{Rec}\times Dur\times\frac{PD}{1-PD};0\right)$$

with *RR* the recovery rate, *PD* the one-year probability of default, D_{ur} is the duration and BE_{Rec} the best estimate of recoverable amounts.

- **Parameter Determination**

Probability of Default (PD) and Recovery Rate (RR) are set by the regulator:

Tab. 5. Probability of Default by credit quality

Credit quality step	0	1	2	3	4	5	6
Probability of default P_iPD_i (%)	0.002	0.01	0.05	0.24	1.20	4.175	4.175

The recovery rate represents the portion of the debt that will be recovered in the event of default of the reinsurer and is set at 50% for reinsurance (SCR.6.30). All risks assumed by the reinsurer from the insurer (collateral, etc.) result in deduction of the exposure.

- **The risk of failure of reinsurance counterparties**

The counterparty risk is the risk of losses resulting from unforeseen failure or downgrade of the credit rating of counterparties or debtors in risk reduction contracts, such as reinsurance treaties. It must cover risk reduction contracts, such as:

- Reinsurance treaties;
- Derivative and securitization products;
- Claims from intermediaries;
- Any other credit exposure not covered in the spread risk sub-module.

In the case where the Loss-Given-Default (LGD) for counterparty *i* (non-random) is denoted as γ_i, the expectation and variance of the overall loss are calculated with

$$M = \sum_{i\in I} p_i y_i \qquad V = \sum_{i,j\in I} \omega_{ij} y_i y_j \,,$$

with the following notations

$$\omega_{ij} = \frac{\theta\left(1-b\left(\theta,p_i\right)\right)\left(1-b\left(\theta,p_j\right)\right)}{\theta+\frac{1}{b\left(\theta,p_i\right)}+\frac{1}{b\left(\theta,p_j\right)}} - \left(p_i-b\left(\theta,p_i\right)\right)\left(p_j-b\left(\theta,p_j\right)\right) \qquad b\left(\theta,p\right) = \frac{p}{1+\theta\left(1-p\right)}$$

By assuming that the distribution of global losses follows a Lognormal distribution Parameters and is defined by

$$\sigma = \sqrt{\ln\left(1+\frac{V}{M^2}\right)} \qquad \mu = \ln\left(M\right) - \frac{\sigma^2}{2}$$

We derive the target capital as follows

$$SCR = \mathbf{min}\left(\sum_{i=1}^{N} Y_i, q \times \sqrt{V}\right)$$

with q = 3.

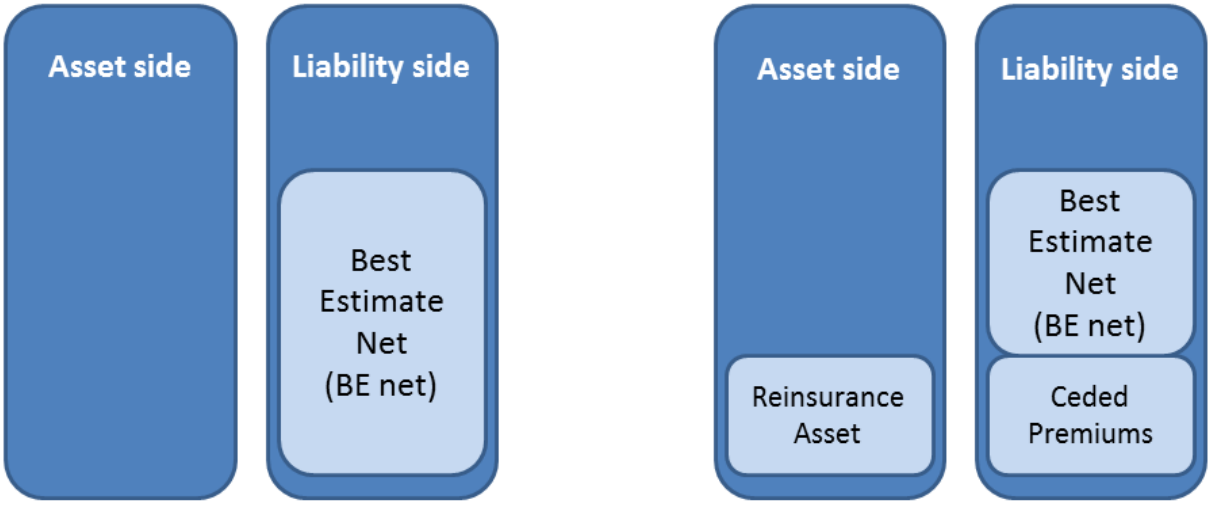
In the case of a single counterparty, the calculation is simplified to

$$SCR = \mathbf{min}\left(1, q \times \sqrt{p \times (1-p)}\right) \times Y$$

For an AAA counterparty it is therefore found that the SCR amounts to 2.9% of the present value of the ceded cash flows (compared to the 58% measured on average for the underwriting risk).

It is possible to synthesize the logic behind incorporating the reinsurance counterparty risk as follows: without reinsurance, we have the best estimate on liabilities and after implementation of a coinsurance treaty, there are the receivables from the reinsurer as an asset and the payables to the reinsurer as a liability.

Tab. 6. Balance sheets with and without reinsurance



- **SCR determination implications**

With respect to the calculations to be carried out to determine the SCR level:

- In the absence of reinsurance, the underwriting SCR is calculated by measuring the variation of the best estimate as a function of applied shocks;
- In the presence of reinsurance, the underwriting SCR is calculated by measuring the change in the value of the best estimate net of the reinsurance asset (i.e. one explicitly manipulates here a variation of Net Asset Value, or NAV) and this calculation is completed by the determination of a counterparty SCR.

SCR calculations for each sub-module

The calculation of the capital required for the components of *Health SLT* underwriting risk for Long Term Care is not based on the nature of the risk but on the application of shocks to the underlying

risks: the required capital corresponds to the impact on the level of the eligible surplus (*BOF, Basic Own Funds*) in a shocked scenario. The impact on BOF for the underwriting risk is measured as the impact on the *Best Estimate* reserves.

$$\Delta BOF = \max\left(0; BOF^{central} - BOF^{shock}\right) = \max(0; Reserves_{BE}^{shock} - Reserves_{BE}^{Central})$$

The calibration of the shocks to be applied reflects the SCR definition, namely $VaR_{99,5\%}$ of the BOFs on a one year horizon. With the many shocks to be applied in the standard formula and their interactions on the active and dependent funds, a detailed and a prospective model is required. It must be able to calculate a robust SCR for each sub-module, a real challenge for the insurer. Moreover, special care must be taken on the assumptions for premium and benefit revalorization, if any, which could lead to particularly high capital requirements.

The standard formula alone does not take into account all Long Term Care underlying risks.
We detail below the shocks according to the standard formula for a cash benefit "Long Term Care" product.

Mortality Shock - EIOPA [2015] 01/2015 – Article 152:

"The capital requirement for health mortality risk shall be equal to the loss in basic own funds of insurance and reinsurance undertakings that would result from an instantaneous permanent increase of 15 % in the mortality rates used for the calculation of technical provisions."

Given the nature of the Long Term Care contract, the mortality shock has 2 opposite effects:

- Reduction of the insurer's liabilities, which would result in a BE decrease, in particular due to the increase in the claimant mortality.
- Reduction in premiums received by the insurer, with the decrease in the active population, thus increasing the BE amount.

The SCR will therefore be non-zero, if the second effect is the most important.

Two product characteristics will be decisive for the weight of this second effect:

- The first element is the duration of the product. The more recent the product, the greater the amount of expected premium.
- The second element is the reserve margin of the product. The more important this margin is, the more premature the lapse, the more significant the loss of profits.

Health Longevity Shock - EIOPA [2015] 01/2015 – Article 153:

"The capital requirement for health longevity risk shall be equal to the loss in basic own funds of insurance and reinsurance undertakings that would result from an instantaneous permanent decrease of 20 % in the mortality rates used for the calculation of technical provisions."

In order to determine the amount of capital required to cope with a decrease in mortality rates, the impact applied is similar to the longevity shock of the module "life" or a permanent 20% decrease of mortality rates for all ages.
In the case of Long Term Care, this decrease affects both the actives and the dependents.
As with the mortality shock, this shock has 2 opposite effects for the BE calculation: the increase of the liabilities of the insurer which increases the BE and the increase of expected premium, which decreases the BE.

Note:
The longevity shock has a very strong weight with an increase of BE which is related to:

- Premium paying active population living longer.
- More active lives reach ages where the dependence incidence is the highest.
- More benefits following the heavier shift from Partial Dependents to Total Dependents[11].

Health disability-morbidity shock - EIOPA [2015] 01/2015 – Article 154:

"The capital requirement for health disability-morbidity risk shall be equal to the sum of the following: (a) the capital requirement for medical expense disability-morbidity risk; (b) the capital requirement for income protection disability-morbidity risk."

Long Term Care insurance products with expense reimbursement benefits would be subject to requirement (a). Long Term Care insurance products with cash benefits are income protection insurance and are subject to requirement (b). This chapter concentrates on cash benefits.

EIOPA [2015] 01/2015 – Article 156:
"The capital requirement for income protection disability-morbidity risk shall be equal to the loss in basic own funds of insurance and reinsurance undertakings that would result from the following combination of instantaneous permanent changes:

(a) an increase of 35 % in the disability and morbidity rates which are used in the calculation of technical provisions to reflect the disability and morbidity in the following 12 months;

(b) an increase of 25 % in the disability and morbidity rates which are used in the calculation of technical provisions to reflect the disability and morbidity in the years after the following 12 months;

(c) where the disability and morbidity recovery rates used in the calculation of technical provisions are lower than 50 %, a decrease of 20 % in those rates;

(d) where the disability and morbidity persistency rates used in the calculation of technical provisions are equal or lower than 50 %, an increase of 20 % in those rates."

The morbidity/disability shock consists of 2 combined sub-shocks: an increase in incidence, which means a deterioration in health status and a decrease in the recovery rate. This recovery rate in the case of dependence can be interpreted as the exit from a dependence state to a dependence-free state or to a lower dependence level. But a dependent state termination due to mortality must not be subject to the application of this shock since it is already shocked in the submodule mortality and longevity. If other cases where changes between states are considered, the transition to an active (healthy) state is often not supported by models. Indeed, a return to the active state after a stay in a dependence state, given the levels of dependence covered by the coverage is rarely accounted for, if at all.

For the case of change between states of dependence, the application of the recovery shock differs significantly according to the approach taken in terms of modeling. Thus, in the case of contracts offering different guarantees depending on the level of dependence, if the model provides for passage distributions between the different states, a shock will have to be applied to these transition probabilities. If the model does not foresee a change of state, the shock cannot not be applied.

[11] Total and Partial Dependence. Total Dependence: 3 ADL out of 5 or Level 1 or 2 for IRG. Partial Dependence: 2 ADL out of 5 or Level 3 or 4 for IRG. IRG (Iso Resource Grid) see chapter 3, section 3.1

Note:
This shock increases the liabilities of the insurer and decreases the insured's.

Lapse shock - EIOPA [2015] 01/2015 – Article 159:

"The capital requirement for SLT health lapse risk referred to in Article 151(1)(f) shall be equal to the largest of the following capital requirements:

- Capital requirement for the risk of a permanent increase in SLT Health lapse rates.
- Capital requirement for the risk of a permanent decrease in SLT Health lapse rates.
- Capital requirement for SLT health mass lapse risk."

The lapse shock is used to assess the amount of capital necessary to cover the risk of unfavorable changes in insured's behavior in terms of voluntary termination.

The SCR amount of SCR retained for this shock is the maximum of the 3 SCR calculated under the following scenarios:

- The permanent increase scenario is applied by increasing lapse rates by 50%.
- The permanent decrease scenario is applied by lowering lapse rates by 50%.
- The massive termination scenario corresponds to a total lapse rate of 40% which replaces the total first year lapse rate in the first year of the projection.

Remarks:
From the modeling point of view, the construction of a biometric distribution on the insureds of the reduced portfolio can be a real challenge given the lack of data and the sometimes unpredictable behavior of the insured. On the other hand, taking into account future claim reductions requires additional projections to take into account a smaller portfolio.

This shock leads to a change in the liabilities of the insurer and those of the insured. In the case of lifetime coverage, the voluntary exit of the insured can result in a reduction in benefits or even an end of coverage. It should be noted that the level of underwriting margin included in the gross premium will have a strong influence on the costliest scenario. In particular, the massive lapse scenario deprives a large portion of future premiums and therefore leads to a very high SCR for products with high underwriting margin.

Warning: the choice of the scenario chosen for the SCR calculation does not depend solely on the results of the Long Term Care product. The scenario chosen is the one with the highest SCR for all contracts and risks covered by the insurance company. This means that the lapse SCR of this product will depend, among other things, on the profile of the insurer's portfolio.

Expense shock - EIOPA [2015] 01/2015 – Article 157:

"The capital requirement for health expense risk shall be equal to the loss in basic own funds of insurance and reinsurance undertakings that would result from the following combination of instantaneous permanent changes:

- (a) An increase of 10 % in the amount of expenses taken into account in the calculation of technical provisions.
- (b) An increase by 1 percentage point to the expense inflation rate (expressed as a percentage) used for the calculation of technical provisions."

The expense shock is intended to estimate the impact on the insurer's liabilities of an inadequate valuation of the expenses associated with the insurance contract. The shock calibration is similar to the expense shock of the "Life" module.

Note:
The methodology for calculating this shock mechanically increases the insurer's liabilities, so the SCR is strictly positive.

Revision Shock - EIOPA [2015] 01/2015 – Article 158:

"The capital requirement for health revision risk shall be equal to the loss in basic own funds of insurance and reinsurance undertakings that would result from an instantaneous permanent increase of 4% in the amount of annuity benefits, only on annuity insurance and reinsurance obligations where the benefits payable under the underlying insurance policies could increase as a result of changes in inflation, the legal environment or the state of health of the person insured."

The revision shock aims to assess the impact on the insurer's commitments of an unanticipated change in the amount of benefit payments due to an evolution of:

- Inflation;
- Regulation: regulatory changes are only taken into account when the regulation is already passed on the valuation date;
- The deterioration of the claimants' health status.

In the case of a pure cash benefit Long Term Care product, annuity payments cannot vary. This shock could therefore be considered as not applicable. However, benefits that would increase with worsening in the dependent status could be part of the application of this shock.

Catastrophe (CAT) Shock - EIOPA [2015] 01/2015 – Article 160:

"The capital requirement for the health catastrophe risk sub-module shall be equal to the following

$$SCR_{HealthCAT} = \sqrt{SCR_{ma}^2 + SCR_{ac}^2 + SCR_p^2},$$

- SCR_{ma} denotes the capital requirement of the mass accident risk sub-module;
- SCR_{ac} denotes the capital requirement of the accident concentration risk sub-module;
- SCR_p denotes the capital requirement of the pandemic risk sub-module."

For a risk classified in the "SLT Health" module, the dependence is subject to CAT shock.

The case of dependence is not explicitly dealt with in the texts. Thus, the transition from active state to dependent state, following the occurrence of one of the three scenarios constituting the SCR CAT, could be considered not to be part of the assumptions contained in the Solvency II regulation. Hence, the SCR CAT would be zero.

Another approach would be to consider the transition to dependence as a form of disability that could correspond to a state reached due to an accident. In this case the dependence risk should be included in the concentration and mass accident SCR calculations.

4.4. Global SCR calculation

From the SCR calculation of each of the sub-modules described in the previous section, we will determine the steps to calculate the SCR.

The following 2 points are essential to fully understand the aggregation mechanism under the Solvency II standard:

- The final SCR that is sought to be determined is the SCR of all the risks born by the insurance company which has a Long Term Care insurance portfolio. Therefore, at each stage of aggregation in the "Octopus" (See Table 2), one must sum the SCR calculated for the Long Term Care portfolio with the SCR of the insurer's other risks.
- At each stage of aggregation in the octopus, i.e. at each stage where one wishes to calculate an SCR using lower level SCR in the octopus, a correlation matrix must be used. This octopus should not be seen as a simple sum of successive calculations.

Step 1:

For each of the sub-modules which are the subject of an SCR calculation (mortality, longevity, incidence, lapse, expenses), the SCR obtained on the Long Term Care portfolio with the SCR of the other risks insured by the company are calculated. The SCR of each of the 5 SCR Health-SLT sub-modules is obtained.

Step 2:

By adding the SCR of the Revision sub-module, and applying the following formula and matrix, we obtain the SLT-Health SCR

$$Health\ SLT\ SCR = \sqrt{\sum_{i,j} Corr[i,j].SCR_i.SCR_j}$$

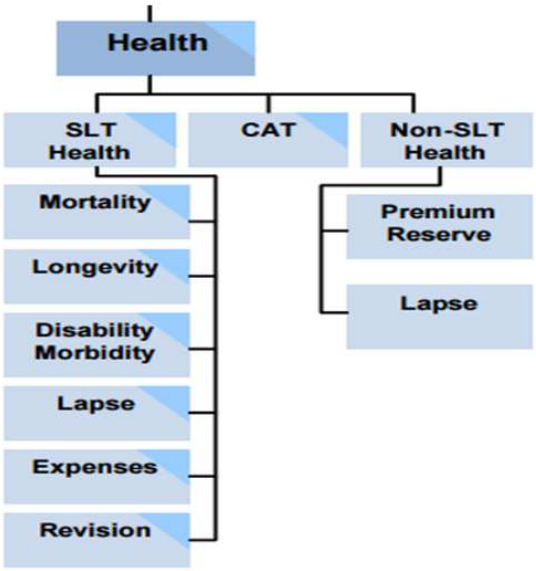
Tab. 7. Aggregation matrix

SCR HEALTH SLT	Mortality	Longevity	Disability/ Morbidity	Lapse	Expenses	Revision
Mortality	100%					
Longevity	-25%	100%				
Disability/Morbidity	25%	0%	100%			
Lapse	0%	25%	0%	100%		
Expenses	25%	25%	50%	50%	100%	
Revision	0%	25%	0%	0%	50%	100%

Step 3

We still go up in the octopus chart and we determine the Health SCR:

Tab. 8. Structure of the "health" SCR



To do this, we add the CAT SCR and the Non-SLT Health SCR, calculated elsewhere on the company's portfolio, using the formula and the following matrix

$$Health\ SCR = \sqrt{\sum_{i,j} Corr[i,j].SCR_i.SCR_j}$$

Tab. 9. Sub-module Aggregation

SCR HEALTH	SLT	Non-SLT	CAT
SLT	100%		
Non-SLT	50%	100%	
CAT	25%	25%	100%

Step 4:

The formula and correlation matrix of part I.B. 4 is applied by adding all SCRs except the Health SCR that we determined in the previous step. This gives the Basic Solvency Capital Requirement (BSCR).

Note:

Before applying the formula and the matrix, we must add the impact of our portfolio in the Market risk module.

The “Market” module "reflects the risk associated with the level or volatility of the market value of the financial instruments impacting the value of the assets and liabilities of the company. It adequately reflects any structural mismatch between assets and liabilities, particularly in terms of their duration. In fact, with an average dependent age of 83[12] years on average (France), the insurer may be required to compensate the insured many years after the asset duration. However, it is sometimes difficult for a Long Term Care insurer to achieve its financial and liability cash flows over a long period which can cause a high liquidity risk. Even if sovereign bonds are of long maturity, they rarely exceed 20 years.

[12] OCIRP (2017) Baromètre OCIRP Autonomie 2017. OCIRP: Organisme Commun Institutions Rente Prévoyance.

Hence, we have 2 contributions to the calculation of this SCR market module:

- Assets and liabilities of the portfolio.
- Impact of a change in discount rates on the BE.

The description of the method of calculating this Market SCR is quite long and is only partially part of this work because the bulk of this module depends on the financial securities in relation to the Long Term Care portfolio. However, the nature of these securities depends essentially on the investment strategy and not on the Long Term Care portfolio.

Step 5:

In this last step, the following formula is applied

$$SCR = BSCR + Adj + Op.$$

Operational (Op):

The OP SCR is a calculation based on the company's premium volumes and statutory reserves. Premium income and reserves of a Long Term Care insurance portfolio are thus added to those of the company's other portfolios to calculate the operational SCR.

Adjustment (Adj):

The ADJ component corresponds to the loss mitigation effect.

Insurers have the opportunity to reduce the basic SCR for a loss-absorption capacity, linked to a profit-sharing mechanism and deferred taxes.

Indeed, if the risks underlying the SCR calculation turn out as expected, variations may still result due to:

- The fiscal profile of the organization and therefore on balance sheet deferred taxes.
- The value of future optional benefits.

The deferred tax represents future tax gain or charge on the result reached by the variation brought to the balance sheet of the company. In the case of a gain, we add the amount of deferred tax to the asset, then it is referred to as Deferred Tax Asset (DTA). If it does not, it is assigned to liabilities and is called Deferred Tax Liability (DTL).

These variations are likely to absorb some of the capital losses.

5. Discussions on the difficulties associated with Solvency II implementation

The technical Solvency II specifications have not been specifically defined for the Long Term Care risk. As a result, insurers must in some cases make their own interpretations to measure the risk associated with this type of portfolio.

In order to better measure the Long Term Care risk, the Solvency II Directive authorizes insurers to use a Partial Internal Model (PIM) which allows them to opt for the standard formula by derogating from it for certain risk modules. However, lack of statistics and history on the Long Term Care risk makes calibrating such a model fastidious.

5.1. Premium revision

The interpretation of premium revisions under Solvency II is subject to question. In the event that future premiums are reviewable "without restrictions" and that the criterium of the section 3 of article 18 of the Commission Delegated Regulation (EU) 2015/35 are fulfilled, they should not be taken into account in the projections beyond one year for the calculations of the standard formula. However, such a revision allows the correction of worse than expected experience, which seems necessary for the long term risks.

If one considers that the ability to revise premiums leads to the retention of a contract maturity of one year, then the contracts become reduced paid up. This situation is often favorable to the insurer, as the premiums can no longer amortize the SCR. On the other hand (see contract maturity), it should be noted that the ability to revise premiums is not the only criterion to be retained in the case of a Long Term Care insurance contract.

5.2. Indexing or Revalorization

Under Solvency II, a revalorization of premiums or benefits, if it exists, is taken into account in the Best Estimate. Often conceived as a profit sharing mechanism, it could require stochastic modeling, which can become costly in calculation time and complexity in the interpretation of the results obtained.

However, contract revalorizations should be dissociated from indexing. In fact, the latter increases premium and benefits in the same proportions and therefore can lead to worsening actual to expected results. Benefit revalorization, on the other hand, can generate a cost for the insurer especially if they are guaranteed in the contract.

6. Conclusion

In this chapter, we have mainly detailed the methodology for assessing the capital requirement for the risk of underwriting the Long Term Care risk. Even if this chapter does not deal in depth with Market SCR and operational SCR, one should supplement the issue risk consideration, by an assessment of financial and operational risks. Thus, the insurer will have a holistic assessment of the required capital to deal with the underlying risks of the Long Term Care coverage.

As mentioned in this chapter, the rules for calculating capital required under Solvency II, the prudential standard in force since January 1, 2016, have not been specifically designed for the Long Term Care risk. They seem to be poorly adapted to this risk, which poses a number of problems of coherence and measurement of the actual level of capital needed for this risk. Nevertheless, regulators offer the insurers the opportunity to deviate from these rules by using a Partial Internal Model. As mentioned above, designing PIM presents its own challenges. As a result, the majority of insurers subject to this standard are not satisfied with the treatment of the Long Term Care risk under Solvency II but are forced to apply the standard formula that generates a very high capital requirement.

Given the long duration of liabilities and the significant level of the associated SCR, it is indispensable for an insurer to hold a projection model taking into account management actions. These choices allow to take into account monitoring tools to be used in order to calculate the appropriate capital requirement. Thus, the integration of management actions is a means for the insurer to meet the prudential requirements by ensuring the continuity of a product with lifetime liabilities.

Finally, on February 28, 2018, the European Insurance and Occupational Pensions Authority (EIOPA) published new recommendations to amend the rules used to calculate the standard SCR formula. Unfortunately, these recommendations do not make any changes to the treatment of lifetime Long Term Care insurance benefits. However, it should be noted that a new standard formula revision is planned for 2020.

Références

CEIOPS (2009), Consultation Paper no.51-Draft Level 2 Advice on SCR Standard Formula-Counterparty Default Risk

COURBAGE C., ROUDAUT N. (2008), Empirical Evidence on Long-term Care Insurance Purchase in France, *The Geneva Papers on Risk and Insurance - Issues and Practice,* October 2008, Volume 33, Issue 4, pp 645–658.

CROIX J.C., PLANCHET F., THÉROND P.E. (2015), Mortality: a statistical approach to detect model misspecification, Bulletin Français d'Actuariat, vol. 15, n°29.

Dupourqué, E., Planchet, F., Sator, N., (2019), Actuarial Aspects of Long Term Care, Springer Actuarial, Springer.

DREES (2018), Étude 1046 - Les Français vivent plus longtemps, mais leur espérance de vie en bonne santé reste stable.

EIOPA (2015), Règlement délégué n°2015-35 complétant la directive 2009/138/CE du Parlement européen et du Conseil sur l'accès aux activités de l'assurance et de la réassurance et leur exercice (solvabilité II).

GUIBERT Q., JUILLARD M., PLANCHET F. (2010), Un cadre de référence pour un modèle interne partiel en assurance de personnes, Bulletin Français d'Actuariat, vol.10, n°20.

LUSSON F. (2013), L'équilibre actuariel de long terme en assurance dépendance en France, Revue d'Analyse Financière, n°47.

OCIRP (2017) Baromètre OCIRP Autonomie 2017

PLANCHET F., TOMAS J. (2014a), Uncertainty on Survival Probabilities and Solvency Capital Requirement: Application to LTC Insurance, Scandinavian Actuarial Journal, doi: 10.1080/03461238.2014.925496.

PLANCHET F., TOMAS J. (2013), Multidimensional smoothing by adaptive local kernel-weighted log-likelihood with application to long-term care insurance, Insurance: Mathematics and Economics, Vol. 52, pp. 573–589. http://dx.doi.org/10.1016/j.insmatheco.2013.03.009

SATOR N., SOTHER G. (2013), Approche Solvabilité II et ERM du risque Dépendance, Proceedings of the AFIR Colloquium

Chapitre 4 :
Etude empirique des attitudes face au risque dans le secteur de la banque et assurance

CHAPITRE 4 : ETUDE EMPIRIQUE DES ATTITUDES FACE AU RISQUE DANS LE SECTEUR DE LA BANQUE ET DE L'ASSURANCE

1. Introduction et état de l'art

Les tentatives de prise en compte des comportements humains face au risque dans les processus de décision ne sont pas récentes. La découverte par Fermat et Pascal de la probabilité et de l'espérance mathématique a permis une rationalisation de certains choix, avec la limite de ne pas rendre compte de certains comportements. C'est à Bernoulli (1738) [1] que revient le mérite d'avoir découvert la notion de l'utilité espérée à partir du paradoxe de Saint-Petersbourg. Le paradoxe de Saint-Pétersbourg se résume à la question suivante : pourquoi, alors que mathématiquement l'espérance de gain est infinie à un jeu, les joueurs refusent-ils de jouer tout leur argent ? Il s'agit donc non d'un problème purement mathématique mais d'un paradoxe du comportement des êtres humains face aux événements d'une variable aléatoire dont la valeur est probablement petite, mais dont l'espérance est infinie. Dans cette situation, la théorie des probabilités dicte une décision qu'aucun acteur raisonnable ne prendrait. La théorie de l'utilité espérée a été développée par John von Neumann et Oskar Morgenstern (Von Neumann et al., 1944 [2]) dans leur ouvrage sur la théorie des jeux. La règle de décision développée par ces auteurs, connue sous le nom « d'utilité espérée », repose sur quatre hypothèses qui sont appelées axiomes (axiome de pré-ordre total, axiome de monotonie, axiome de continuité et axiome d'indépendance). Ils appliquent cette théorie pour prédire le comportement des joueurs dans les jeux non coopératifs. Cette théorie a eu de nombreuses applications, principalement dans les assurances ou la gestion de portefeuille.

L'utilité espérée a été élargie dès 1954 par Savage [3] aux situations incertaines. Son axiomatisation comporte six axiomes lorsque l'ensemble d'états du monde est fini. Il introduit un axiome « clé », le principe de la chose sûre. Bien que séduisant par sa simplicité et sa relative souplesse d'utilisation comparée à d'autres approches, le modèle de l'utilité espérée dans l'incertain a fait l'objet de plusieurs critiques expérimentales. La principale critique est liée au principe de la chose sûre, qui neutralise l'impact de l'ambiguïté sur les préférences, comme le suggère l'exemple du paradoxe d'Ellsberg (1961) [4]. Le paradoxe d'Ellsberg montre que la majorité des agents sont averses à l'incertitude sur les probabilités. D'autres modèles, appelées « *Non-Expected Utility* » modèles, ou modèles non-additifs, ont donc été axiomatisés dans l'incertain pour résoudre ce type de paradoxe.

En 1979, Daniel Kahneman et Amos Tversky (Kahneman et al., 1979 [5]) publient une théorie alternative à la théorie de l'espérance d'utilité. Ces auteurs posent les fondements d'une analyse des décisions en incertitude comme une alternative de la théorie économique standard, dominée jusque-là par le modèle de l'utilité espérée. Cette approche n'épargne aucun des éléments qui concernent le concept du risque, depuis les composantes de sa perception jusqu'à son évaluation en termes de probabilités. À l'inverse de la théorie de l'espérance d'utilité qui a des fondements axiomatiques, la théorie des perspectives cherche à prendre en compte les principaux faits stylisés observés au cours des expériences de laboratoires : 1) les individus ont un point de référence c'est-à-dire qu'ils sont averses au risque lorsqu'ils sont au-dessus de leur point de référence et recherchent le risque lorsqu'ils sont en dessous de leur point de référence, 2) les individus ont une aversion aux pertes c'est-à-dire la perte d'utilité liée à la perte de quelque chose est plus grande que l'utilité liée à son gain, enfin 3) les individus déforment les probabilités c'est-à-dire qu'ils surestiment l'importance des événements rares et sous-estiment l'importance des événements presque certains. Précurseurs de l'économie expérimentale, s'inscrivent dans un courant de psychologues expérimentalistes. Les rangs de ce courant n'ont cessé de grossir depuis plus de vingt-cinq ans, et parmi lesquels on compte Slovic, Fischoff, et Lichtenstein (1982) [6], pour ne citer que les plus connus. Forts d'une base solide de données expérimentales, Kahneman et ses collègues mettent en question la portée positive de la théorie économique classique

des choix en avenir incertain, dont les résultats se trouvent contredits de manière récurrente par les comportements observés. Cette distance par rapport à la théorie économique orthodoxe, confortait la position prise par plusieurs économistes depuis les célèbres expériences de Maurice Allais (1952, 1953) [6], dont Allais lui-même. Ces derniers en effet, avaient déjà exprimé de sérieux doutes sur la validité empirique des axiomes de la théorie de l'utilité espérée et amendé, en conséquence, le corpus de la théorie dominante (Machina, 1982 [7] ; Fishburn, 1988 [8]).

Malgré les avancées apportées par l'introduction de la psychologie expérimentale, l'observation des comportements humains en situation face à l'incertitude pose un certain nombre de problèmes. En effet, mener et réussir une expérimentation nécessite une organisation coûteuse et une méthodologie très contraignante. L'exploitation des résultats d'une expérimentation est également soumise à certaines limites. Baptiste Campion et al. (2013) [9] soulignent que pour assurer à la fois la validité interne et externe des résultats produits lors d'une expérimentation, certains compromis et adaptations sont nécessaires. Lemaine et al. (1969) [10] exposent les contraintes contradictoires qui tiraillent le chercheur entre le contrôle de toutes les variables de l'expérience et celle de s'assurer que cette expérience correspond à une situation réelle ou généralisable. Matalon (1995, 1969) [11] précise que le travail de sélection des aspects les plus pertinents de la situation que l'on veut étudier et du fait du haut niveau de contrôle, conduirait à une situation miniaturisée et donc artificielle par nature. Ce qui semble amener les participants aux expériences à poser des actions et des raisonnements qu'ils n'auraient peut-être jamais effectués en situation naturelle. Ainsi, sur le plan de la validité externe des résultats, se pose la question de leur transposabilité, c'est-à-dire la validité en-dehors de la situation de l'expérience. A ce paradoxe, s'ajoutent d'autres complexités comme le souci de désirabilité sociale relevé par Delory (2003) [12] et l'effet « Hawthorne » qui révèle que se savoir l'objet d'une attention spécifique de la part de scientifiques peut affecter le comportement des participants, voir Adair (1984) [13].

Les chercheurs se sont également longtemps intéressés aux attitudes dans l'espoir de mieux expliquer le comportement. En psychologie, Bloch et al. (1997, p.119) [14] définissent l'attitude comme un « état de préparation dans lequel se trouve un individu qui va recevoir un stimulus ou donner une réponse et qui oriente de façon momentanée ou durable certaines réponses motrices ou perceptives, certaines activités intellectuelles ». Selon ces auteurs, l'attitude désigne « la disposition interne durable qui sous-tend les réponses favorables ou défavorables de l'individu à un objet ou à une classe d'objet du monde social ». Pour Stoetzel [15], « L'attitude consiste en une position (plus ou moins cristallisée) d'un agent (individuel ou collectif) envers un objet (personne, groupe, situation, valeur). Elle s'exprime plus ou moins ouvertement à travers divers symptômes ou indicateurs (parole, ton, geste, acte, choix ou en leur absence). Elle exerce une fonction à la fois cognitive, énergétique et régulatrice sous les conduites qu'elle sous-tend. » (Stoetzel cité par Maisonneuve, 1982, p.111). Il précise que les attitudes sont acquises et non pas innées, elles sont plus ou moins susceptibles de changements sous l'effet d'influences extérieures. Les attitudes sont définies selon Eagly et Chaiken (Eagly et al., 1993 [16]) comme des tendances à évaluer une entité avec un certain degré de faveur ou de défaveur, habituellement exprimées dans des réponses cognitives, affectives et comportementales. Une tendance relativement stable à répondre à quelqu'un ou à quelque chose de manière qui reflète une évaluation (positive ou négative) de cette personne ou chose. Les réponses comportementales font partie des manières par lesquelles l'individu peut exprimer son évaluation. En psychologie sociale, le terme d'attitude est employé pour désigner « un état mental prédisposant à agir d'une certaine manière, lorsque la situation implique la présence réelle ou symbolique de l'objet d'attitude » (Thomas, Alaphilippe, 1993, p.5 [17]). L'attitude désigne aussi « une prédisposition à agir dans un certain sens » (Mathieu & Thomas, 1995, p.393). Allport (1935) [18], cité par Vallerand (1994, p.331) proposa la définition suivante : « Une attitude représente un état mental et neuropsychologique de préparation à répondre, organisé à la suite de l'expérience et qui exerce une influence directrice et dynamique sur la réponse de l'individu à tous les objets et à toutes les situations qui s'y rapportent. ». En dehors du rôle de l'attitude sur la prédisposition à agir, certains auteurs définissent l'attitude comme la manière dont une personne se situe par rapport à un objet exerçant une influence sur le comportement. Au vu de ces différentes définitions, nous comprenons la complexité du concept d'attitude. Moscovici (1960), cité par Ebalé Moneze

(2001, p.4) [19], propose une définition qui tient compte de deux grandes sphères de réflexion sur les attitudes. Il définit l'attitude comme « un schéma dynamique de l'activité psychique, schéma cohérent et sélectif, relativement autonome, résultant de l'interprétation et de la transformation des mobiles sociaux et de l'expérience de l'individu. ». Ce schéma psychologique appelé attitude est fortement corrélé à une opinion ou à un jugement de valeur. L'attitude est donc une variable intermédiaire entre la situation et la réponse à cette situation. Elle permet d'expliquer que, parmi les comportements possibles d'un sujet soumis à un stimulus, celui-ci adopte tel comportement et non tel autre. Une attitude est donc une prédisposition à réagir d'une façon systématiquement favorable ou défavorable face à certains aspects du monde qui nous entoure. L'attitude, est considérée comme porteuse de sens sinon d'intention. Ce que nous pouvons retenir de ces multiples définitions est que l'attitude permet d'avoir une vue d'ensemble sur les comportements des individus face à des objets précis, des situations précises, des personnes précises et des phénomènes précis. L'attitude a un rôle de variable intermédiaire entre l'environnement et les réponses d'une personne. Nous ne pouvons pas l'observer directement, nous l'inférons des réponses de la personne lorsqu'elle est confrontée physiquement ou symboliquement à l'objet.

L'attitude est le plus souvent abordée dans le cadre d'un modèle tripartite composé d'une dimension cognitive, d'une dimension affective et d'une dimension comportementale distinctes (voir Breckler, 1984 ; Eagly et al., 2007 ; Olson et al., 2008 ; Zanna et al., 1988). La dimension cognitive de l'attitude prend appui sur les croyances, pensées, attributs associés à l'objet d'attitude. La dimension affective prend appui sur les sentiments ou émotions associés à l'objet d'attitude (voir Bornstein, 1989 ; Zajonc, 2001). Les attitudes peuvent avoir aussi une origine comportementale qui prend appui sur les comportements passés envers l'objet. Nous nous situons souvent à l'égard d'un objet, surtout lorsque nous ne sommes pas sûrs de nos sentiments (voir Nisbett et al., 1977), en fonction de ce que nous faisons ou avons fait dans le passé (voir Albarracín et al., 2000 ; Bem, 1972 ; Dolinski, 2000 ; Hofmann et al., 2010). Bien que distinctes, les dimensions cognitives et affectives varient dans le même sens. Posséder des croyances positives envers un objet est associé à des réponses affectives positives sur ce même objet, posséder des croyances négatives est, à l'inverse, associé à des réponses négatives affectives.

Selon Allport (1935), la relation attitude-comportement est partie intégrante de la définition de l'attitude, notamment dans la composante se rapportant à la volonté et à l'effort. De plus, le postulat de la consistance attitude-comportement est sous-jacent aux théories du changement d'attitude. L'intérêt considérable accordé aux attitudes provient en grande partie de ce qu'elles devraient normalement permettre de prédire les comportements des individus. Mais il ne faut pas oublier que les attitudes ne sont pas les seuls déterminants des comportements. C'est ainsi que plusieurs études, notamment celle de LaPiere (1934, pp.230-237) [20], ont démontré que la relation attitude-comportement est souvent moins forte que ce que l'on avait cru autrefois. Les études menées par Deutscher (1966) et Wicker (1969) [21] concluent qu'en moyenne, l'attitude n'expliquerait qu'environ 10% de la variable comportementale. A la même période, Mischel (1968) [22] rassembla aussi les recherches concernant la valeur du trait de personnalité comme facteur prédictif du comportement pour conclure que la corrélation moyenne était approximativement de 0,30 entre le trait et la conduite. Pourtant, sur le plan empirique, on retrouve des indices fiables de la validité prédictive de l'attitude au regard du comportement (voir Rajecki, 1990 [23]). Mais, cette causalité linéaire entre attitude et comportement n'est pas vérifiée dans des situations où le comportement est entièrement influencé par des variables intermédiaires entre attitude et comportement (voir Bickman, 1972 [24]).

Ainsi, plusieurs facteurs influencent le comportement humain à différents niveaux et de manière complexe. Le premier niveau correspond aux facteurs contextuels qui sont : les facteurs individuels (personnalité, ambiance, émotionnel, intellectuel, valeurs, stéréotypes, attitudes générales, expérience), les facteurs sociaux (éducation, âge, genre, revenus, religion, race, appartenance ethnique, culture) et les facteurs informatifs (connaissance, média, intervention). Ce premier niveau influence le second niveau composé de trois niveaux : les croyances comportementales qui induisent l'attitude envers le comportement, les croyances normatives qui induisent la norme subjective, et enfin les croyances par rapport au contrôle qui induisent la perception du contrôle. Le troisième niveau est la formation de

l'intention qui découle des influences des facteurs induits. Enfin, sous le contrôle du comportement actuel découlera le comportement.

Les travaux sur les variétés humaines et des groupes sociologiques ont donné lieu à la théorie culturelle face au risque. Elle décrit l'attitude favorable ou défavorable face au risque à partir des modalités anthropologiques et socio-culturelles. Les trois principaux modèles culturels face au risque souvent cité dans la littérature sont ceux de Hofstede (1980, 2001) [25] [26], Inglehart (1977, 1997) [27] [28] et Schwartz (1994, 2004) [29] [30]. Ces modèles tentent d'expliquer que les différences des valeurs culturelles entre zones géographiques influencent les attitudes face au risque et donc les comportements. Hsu, Woodwide et Marshall (2013) [31] ont établi une comparaison approximative entre ces modèles.

Comparaison des modèles de valeurs culturelles sur le comportement face au risque

Dimensions \ Auteurs	HOFSTEDE	SCHWARTZ	INGLEHART
Rapport à l'autorité	Distance au pouvoir	Egalitarisme vs hiérarchie	Séculier-rationnelle vs traditionnelle
Rapport à soi et au groupe	Individualisme vs collectivisme	Autonomie vs enchâssement	Expression de soi vs survie
Lien avec l'environnement social et naturel	Féminité vs masculinité	Harmonie vs maîtrise	
Rapport à l'incertitude	Evitement de l'incertitude		

Tableau 1.1 : comparaison des modèles de valeurs culturelles sur le comportement face au risque

Plusieurs études transculturelles ont exploré sur la base du paradigme psychométrique le lien entre les valeurs culturelles et l'attitude face au risque (voir Fischhoff et al., 1978 [32] ; Slovic et al., 1980 [33] ; Slovic, 1987 [34] ; Slovic, 2000 [35]). Le paradigme psychométrique repose sur une série de travaux remarquables sur la perception des risques effectués en psychologie cognitive et sociale à partir de la fin des années 70. Les plus connus sont sans aucun doute ceux réalisés autour de Paul Slovic à l'université d'Oregon. Le principal apport de cette approche est d'avoir pu montrer que les individus apprécient moins le risque sur des critères quantitatifs que sur des critères qualitatifs. Historiquement, les travaux psychométriques s'inscrivent dans la continuité des études expérimentales de la théorie axiomatique.

Toutefois, même si elle présente une certaine proximité méthodologique avec l'approche axiomatique, la problématique que commence à traiter Slovic dans un article publié en 1978 est très différente de celle d'auteurs comme Tversky et Kahneman. En effet, un important changement théorique s'est opéré : il se traduit notamment par l'abandon des expériences en laboratoire sur la prise de décisions – le plus souvent monétaires – en situation d'incertitude au profit de l'étude empirique de l'attitude des populations face aux risques sanitaires ou écologiques. Le paradigme psychométrique est une classification des différents risques potentiels en fonction de leurs scores perçus sur 19 caractéristiques. L'analyse factorielle réalisée sur ces caractéristiques, a révélé deux dimensions principales : le caractère effrayant des risques et les connaissances acquises sur le risque. Une troisième dimension avec un impact plus faible, est liée au nombre de personnes perçues comme étant exposés au risque. Ces facteurs reflètent le jugement et l'utilisation d'heuristiques par des sujets profanes lors de l'évaluation biaisée de probabilités (voir Kahneman et Tversky, 1979).

La principale thèse développée par Douglas et Wildavsky est qu'il existe pour chaque type d'organisation sociale un « portefeuille de risques » spécifique correspondant aux valeurs dominantes de la communauté. Pour Mary Douglas, les nombreux paradoxes et « points aveugles », mis en évidence dans le cadre des recherches en économie ou en psychologie sur les comportements et les attitudes face au risque, résultent pour l'essentiel de la posture épistémologique que constitue, selon elle, l'adhésion à l'individualisme méthodologique. Mary Douglas apparaît particulièrement critique vis-à-vis des protocoles de recherche sur la perception du risque mis en œuvre par la plupart de ses collègues psychologues (Mary Douglas [36]). Pour Mary Douglas, les études psychologiques ne mettent pas suffisamment l'accent sur l'intersubjectivité, la concertation ou les influences sociales sur les décisions individuelles. Elle exprime toutefois une certaine sympathie pour les travaux d'Herbert Simon sur la rationalité limitée, dans la mesure où l'auteur traite non seulement des facteurs cognitifs individuels, mais

aussi de l'influence du contexte environnemental sur les agents économiques. Pour Mary Douglas, il convient toutefois de dépasser le modèle de Simon en supposant que l'environnement social dans lequel les individus évoluent traite une partie de l'information à leur place. Ainsi, le fait que certains risques sont considérés comme plus ou moins importants selon les cultures peut être envisagé sous l'angle du traitement d'informations sélectionnées par l'environnement social ou organisationnel.

L'autre idée développée par Douglas et Wildavsky c'est que les individus faisant partie d'un système définissent leurs risques, réagissent violemment à certains, en ignorant d'autres, d'une manière compatible avec le maintien de ce système. Au début des années 80, Mary Douglas et Aaron Wildavsky se sont attachés à développer un cadre théorique permettant d'analyser les préférences communautaires et les attitudes vis à vis du risque à partir d'une série d'hypothèses sur les rapports existant entre la forme d'une société et ses valeurs culturelles. Comme le souligne Mary Douglas, chaque type de communauté est un monde de pensée, qui s'exprime dans son propre style de pensée, pénètre la pensée de ses membres, définit leur expérience et met en place les repères de leur conscience morale. Selon les promoteurs de l'approche socioculturelle, les organisations sont dotées de systèmes de valeurs et de croyances qui déterminent la manière dont leurs membres perçoivent l'univers qui les entourent, sélectionnent et interprètent les informations qu'ils reçoivent, mais aussi réagissent aux risques qui les menacent. La thèse développée par ces auteurs est qu'il existe pour chaque type d'organisation sociale un portefeuille de risques spécifique, conforme aux valeurs spécifiques du groupe.

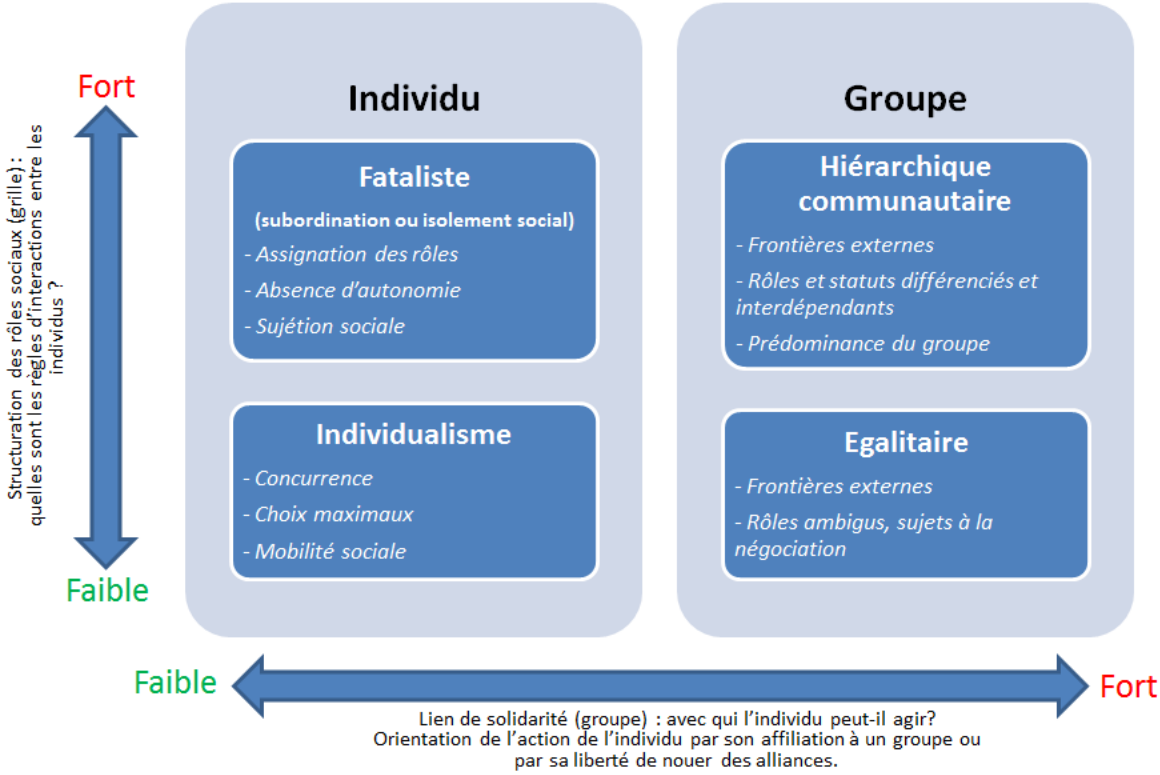
Ainsi, pour Douglas et Wildavsky, en choisissant un mode de vie, nous choisissons également de courir certains risques. Chaque forme de vie sociale a son propre portefeuille de risques. Partager les mêmes valeurs, c'est aussi partager les mêmes craintes, et inversement les mêmes certitudes. Ce cadre théorique permettrait d'expliquer la diversité des comportements face au risque en fonction de la structure institutionnelle dont les individus dépendent. Douglas et al. (1983) [37] considèrent que les positions culturelles peuvent se révéler plus influentes que le sexe ou l'âge et agir comme des facteurs de confusion.

Dans *Risk and Culture*, Douglas et Wildavsky introduisent également la thèse selon laquelle la variabilité culturelle tend à être structurée sur la base de deux variables d'inspiration durkheimienne que sont 1) le degré d'ouverture / fermeture du groupe par rapport au reste de la société (*group*) et 2) le degré d'autonomie dont disposent les membres à l'intérieur du groupe (*grid*). Le croisement de ces deux variables permettrait de dégager quatre principaux modèles organisationnels, qui peuvent être identifiés dans toutes les sociétés humaines (la communauté centrale, les communautés dissidentes, les individualistes et les exclus) et qui correspondent à quatre cultures idéal-typiques. Ils distinguent dans toutes les sociétés humaines la présence de quatre visions du monde face aux risques : les *individualistes*, les *égalitaristes*, les *hiérarchiques* et *fatalistes*. Mary Douglas propose un repère à deux dimensions constituant les facteurs d'influence des attitudes face au risque des groupes sociologiques (voir graphique 1.1 ci-dessous).

Dans les groupes sociologiques *Individualiste* et *Fataliste*, le niveau d'incorporation des individus et de soutien social est faible. Dans les groupes sociologiques *Egalitaire* et *Hiérarchique*, le niveau d'incorporation et de soutien social est très élevé. Dans les groupes sociologiques *Individualiste* et *Egalitaire*, l'organisation est très peu structurée et les individus très autonomes. Enfin, dans les groupes sociologiques *Fataliste* et *Hiérarchique*, la structure impose une forme plus structurée d'organisation et les individus sont par conséquents plus dépendants du groupe et moins autonomes.

D'autres auteurs ne partagent pas la théorie de la dimension culturelle de l'attitude face au risque n'est pas partagé. Rohrmann (2000) [38] explique qu'une variation transnationale considérable dans la perception du risque qui existe. Boholm (1998) [39] dans une étude très approfondie et détaillée, fait des mises en garde méthodologiques sur les comparaisons de la perception du risque entre pays et conteste l'interprétation de ces différences comme reflétant les influences culturelles. Il relève également que de telles différences peuvent être liées à des spécificités contextuelles, plutôt que culturelles, et donc déterminées par des facteurs externes. En effet, la perception du risque est sous l'influence du biais de disponibilité ou en fonction de la couverture médiatique. Il est possible que

certains risques semblent plus disponibles dans un pays en fonction de sa taille ou de la densité de la population. La fréquence de survenance d'un risque peut être corrélée à la superficie ou la densité de la population du pays.



Graphique 1.1 : Typologie grid-group des institutions sociales adapté de Douglas (1983)

A partir de ces différents constats, la question est de savoir comment mesurer l'attitude face au risque ? Notons que l'attitude est une construction hypothétique, c'est-à-dire une entité dont on déduit l'existence mais qui n'est pas directement observable. En effet, l'attitude résulte d'influences de différents facteurs (cognitif, affectif et comportemental) sur l'objet de l'attitude. Pour la mesurer, ils existent des mesures explicites et des mesures implicites.

Les techniques de mesure de l'attitude explicite les plus courantes sont les réponses autorapportées (Thurstone, 1928 ; Thurstone & Chave, 1929 ; Likert, 1932). Par exemple, l'utilisation d'une échelle de Likert implique que le répondant exprime son degré d'accord ou de désaccord sur un objet d'attitude. L'échelle, composée d'items, contient en général cinq, voire sept, niveaux permettant de nuancer le degré d'accord (par exemple : de 1 « pas du tout d'accord » à 5 « tout à fait d'accord »). L'attitude correspond à la moyenne des réponses du répondant à l'ensemble des items de l'échelle. L'échelle de différenciation sémantique (Osgood, Suci, & Tannenbaum, 1957) permet d'évaluer un objet d'attitude à partir d'items en 7 points concernant une série d'adjectifs bipolaires (e.g., bon-mauvais, favorable-défavorable, agréable-désagréable). La somme des moyennes des items permet de recueillir l'attitude explicite du répondant. Ces techniques de mesure sont particulièrement sensibles au contexte. Des variations mineures dans la formulation, l'ordre, le format ou encore dans le nombre de points de l'échelle peuvent affecter les réponses (Krosnick & Fabrigar, sous presse ; Schwarz, 2007b, 2008 ; Tourangeau, Rips, & Rasinski, 2000). Quelles que soient les échelles de mesure autorapportées, les répondants peuvent déformer leurs réponses par souci de désirabilité sociale, c'est-à-dire de se présenter sous un jour favorable à ses interlocuteurs (Eagly & Chaiken, 2005). Les répondants sont aussi limités dans ce qu'ils peuvent dire ou penser d'eux-mêmes (difficultés d'accès à nos déterminants internes :

Joule, 1989 ; Nisbett & Wilson, 1977 ; Wilson, 2002). Afin d'atténuer, voire de supprimer la désirabilité sociale, certains chercheurs ont eu recours à un faux détecteur de mensonges (bogus pipeline, Jones & Sigall, 1971). On fait croire aux répondants qu'un appareillage apparemment sophistiqué, mais factice, permet de savoir ce qu'ils pensent vraiment. Vingt ans de recherches ont montré que le faux détecteur de mensonges permet d'obtenir des réponses plus sincères (Tourangeau, Smith, & Rasinski, 1997 ; Roese & Jamieson, 1993).

D'autres mesures indirectes de l'attitude ont été développées, au nombre desquelles le test d'associations implicites (Implicit Association Test ou IAT, Greenwald, McGhee & Schwartz, 1998). Il s'agit de l'outil qui a suscité le plus d'intérêt et de travaux. L'IAT est une mesure implicite de l'attitude (cf. Blaison, Chassard, Kop, & Gana, 2006). Une attitude implicite est une attitude dont l'individu n'est pas conscient et qu'il ne contrôle donc pas (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009 ; Fazio & Olson, 2003 ; Gawronski & Bodenhausen, 2010 ; Petty, Fazio, & Briñol, 2009). Elle mesure l'association automatique existant entre un objet d'attitude et son évaluation dans la mémoire. La mesure de l'attitude implicite permet de réduire et même d'éliminer le biais de désirabilité (Ferguson, Hassin, & Bargh, 2008). L'IAT a pour objectif de comparer l'écart des temps de réaction entre des associations congruentes, d'une part, et des associations non congruentes, d'autre part. Greenwald et al. (1998) ont montré que l'association compatible débouche sur des temps de réponse plus petits que l'association incompatible. L'écart, en temps, entre ces deux combinaisons rend compte d'une association implicite ou de la force associative entre les concepts-cibles et les pôles évaluatifs (e.g., Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005 ; Greenwald, Nosek, & Banaji, 2003). Plusieurs instruments implicites ont été spécialement créés afin de compléter et renforcer la mesure IAT (e.g., Extrinsic Affective Simon Task ou EAST, De Houwer, 2003 ; Go/No go Association Task ou GNAT, Nosek & Banaji, 2001).

Les mesures physiologiques peuvent également être prises en compte dans l'évaluation des attitudes explicites ou implicites. Par exemple, de faibles activations ou des contractions minimes des muscles faciaux imperceptibles à l'œil nu peuvent refléter la valence (négative ou positive) et l'intensité de l'attitude. La mesure de l'activité cérébrale est, elle aussi, utilisée (Cunningham, Packer, Kesek, & van Bavel, 2009 ; Ito & Cacioppo, 2007).

Ingram et al. (2010) [40] ont mené des études auprès d'un panel de professionnels du secteur d'assurance en Amérique (*United State Of America*). Ils proposent une correspondance entre les archétypes des groupes socio-culturels et les catégories d'attitude face au risque. Ils associent le groupe des « *individualistes* » à la catégorie des « *Maximisateurs* », les « *Egalitaires* » à la catégorie des « *Conservateurs* », les « *Fatalistes* » à la catégorie des « *Pragmatiques* » et enfin les « *Hiérarchiques* » à la catégorie des « *Manager*s ». Ces auteurs utilisent des notions de gains et de profits pour décrire chaque catégorie d'attitude face au risque (D. Ingram et al., 2010 [41] [43] ; D. Ingram et al., 2009 [42] ; D. Ingram et al., 2011 [44] [45] ; D. Ingram et al., 2012 [46]). Les *Maximisateurs* (« *Maximizers* ») sont ceux qui pensent que l'augmentation des profits est plus importante que l'examen des risques. Les *Managers* (« *Managers* ») préfèrent s'assurer de l'équilibre minutieux entre les risques et les profits. Les *Conservateurs* (« *Conservators* ») pensent que l'accroissement des gains n'est pas plus important que l'évitement des pertes. Enfin les *Pragmatiques* (« *Pragmatists* ») croient que l'avenir est très peu prévisible et préfèrent éviter les engagements fermes afin de garder leur autonomie et flexibilité. Ils admettent que chaque catégorie est une dimension des attitudes face au risque. Elles sont différentes et contradictoires, chacune étant adaptée à l'un des quatre états du cycle économique. L'attitude « *Maximizers* » est la plus adaptée à la période de « *Boom économique* ». En effet, en situation de *Boom économique »* les décisions risquées payent généreusement et ne pas acheter des couvertures est plus rentable que l'achat de couverture contre le risque. L'attitude « *Conservators* » est adapté à la période de *Crise économique*, car de nombreuses prises de risque sont dangereuses et conduisent à des pertes, la gestion des risques se concentre sur la survie et l'évitement des pertes. L'attitude « *Pragmatists* » est adapté à la période d'*Incertitude économique*, car l'avenir devient soudainement imprévisible et toutes les actions semblent imparfaites. L'attitude « *Managers* » est adapté à la période de « *Stabilité économique* », car les prévisions de tendance sont fiables et l'achat de couverture contre le risque est efficace.

Ces auteurs proposent un modèle paramétrique d'affectation des individus suivant leur catégorie d'attitude face au risque. Il a été calibré à partir des données d'un sondage réalisé sur un panel de professionnels du secteur d'assurance exerçant dans la zone Amérique. La liste de 40 questions posées aux répondants est une adaptation du modèle de la « *survey of risk attitude* » de Karl Dake (1990) [47] (voir Wildavsky, A. et Dake, K., 1990 [48]).

Le choix du secteur financier, notamment les activités d'assurance et de banque, a été motivé principalement par la volonté de comparaison avec les résultats l'équipe David Ingram. En effet, ces auteurs ont réalisé leurs travaux sur un échantillon de population dans la zone US et travaillant en assurance. Nous avons introduit le secteur bancaire pour élargir le périmètre et enrichir notre analyse. Notons enfin que les professionnels (des secteurs d'activité association, courtiers, finances, régulateurs, réassurances, universités...) qui ont accepté de participer à l'étude exercent leurs activités professionnelles en liens avec l'assurance ou la banque.

La question de recherche posée par cette approche paramétrique sont nombreuses. Est-il possible sous certaines conditions de la généraliser à d'autres pays et entreprises du secteur financier ? Le modèle paramétrique proposé par Ingram et al. (2014) sur l'expérience d'un panel issu de la zone Amérique permet de détection les profils d'attitude face au risque. Partant de cette approche empirique et basée sur des jugements d'experts, les objectifs de nos travaux sont doubles. D'abord, généraliser ce modèle afin de l'appliquer à d'autres zones géographiques. Ensuite, tester statistiquement sur notre panel d'étude l'existence de liens d'association éventuels entre les profils et certaines variables sociodémographiques (secteur d'activité, position occupée dans l'entreprise, le genre, l'ancienneté professionnelle). D'un point de vue méthodologique, nous avons utilisé le même questionnaire conçus par l'équipe Ingram et al. (2009) pour réaliser les enquêtes en Europe et en Afrique auprès de professionnels des secteurs financiers (notamment en Banque et en Assurances) entre 2013 et 2015.

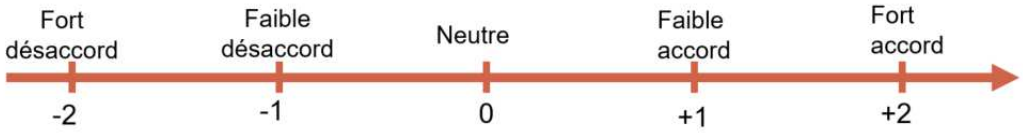
Cet article est organisé en six parties. Dans la Section 2, nous décrivons les données et quelques statistiques descriptives du panel. Dans la Section 3, nous présentons le cadre méthodologique de l'étude. Nous rappelons l'approche paramétrique et les aménagements nécessaires pour une généralisation. Les applications et les analyses des résultats empiriques sont résumées en Section 4. Les discussions sur la robustesse des résultats sont abordées en Section 5. Enfin, nous concluons en Section 6.

2. Description des données

Nous avons complété le panel de la zone Amérique construit par l'équipe David Ingram avec les zones Afrique et Europe. Les sondages ont été réalisés entre 2013 et 2015 en Europe et en Afrique via le site internet « *SurveyMonkey®* ». Ce site propose des outils gratuits pour mener de telles enquêtes auprès d'un public international. Enfin, nous avons organisé des entretiens avec certains participants du panel pour la phase de validation des réponses.

Au total un jeu de quarante questions a été proposé aux participants. Pour respecter les contraintes des échelles de mesure autorapportées, cinq réponses sont possibles pour permettre de mesurer l'adhésion plus ou moins forte ou l'opposition plus ou moins fort des répondants à la question posée.

L'intensité de la préférence (l'adhésion ou l'opposition) se traduit par : « fortement d'accord » pour une note de +2, « d'accord » pour une note de +1, « neutre » pour une note de 0, « en désaccord » pour une note de -1 et « fortement en désaccord » pour une note de -2.

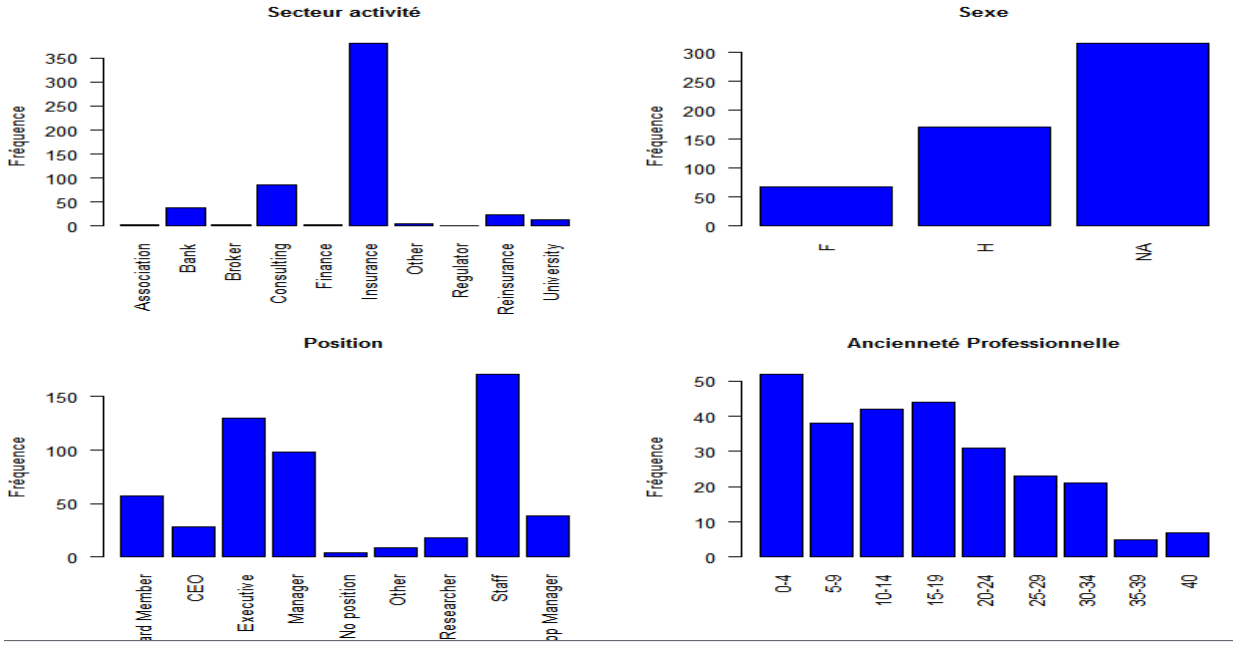


Graphique 2.1 : Intensité d'accord ou désaccord

A partir des notes obtenues aux différentes questions, nous calculons le score de chacune des quatre catégories d'attitude face au risque. Enfin, différentes précautions d'usages propres aux traitements des données de sondage ont été mises en œuvre pour limiter les biais bien connus pouvant provenir d'une telle démarche (voir Llosa Sylvie, 1997 [49] ; Patrice Tremblay et Benjamin Beauregard, 2006 [50]).

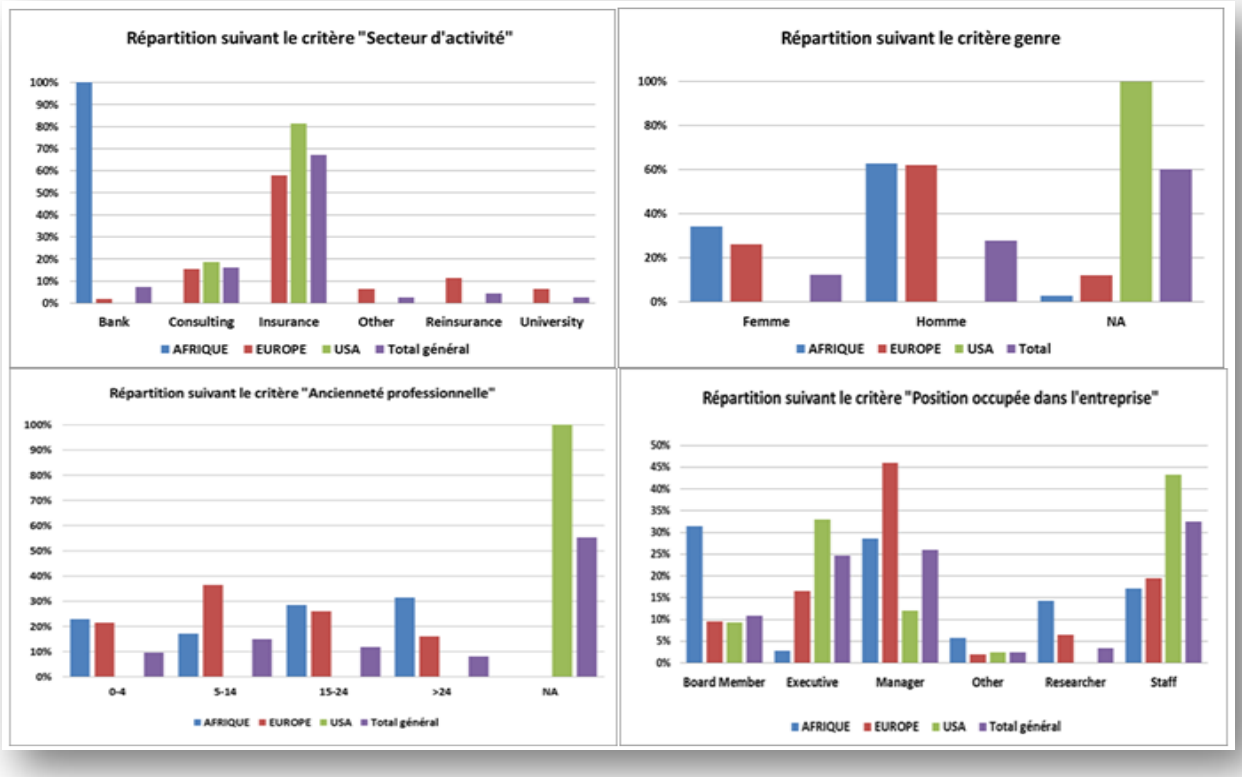
Nous disposons d'un échantillon de 554 répondants, répartis dans trois zones géographiques (Afrique 6%, Europe 40% et Amérique 54%). Les autres variables sociodémographiques collectées sont : Secteur d'activité, Genre, Ancienneté professionnelle et Position occupée dans l'entreprise.

Un travail de contrôle et de complétude de certaines données manquantes a été réalisé, ce qui permet de disposer d'une base comportant des informations de qualité. A noter que les répondants de la zone Amérique n'ont fourni d'information sur leur ancienneté professionnelle et leur sexe.



Graphique 2.2 : Statistiques descriptives du panel

Le secteur assurance est le plus fortement représenté dans l'échantillon. Hors les répondants de la zone Amérique, la proportion des répondants hommes est le triple de celle des femmes. L'échantillon est composé majoritairement d'individus (nombre d'individus supérieur à 50) occupant les postes de membre du « *Staff* », de « *Managers* », des « *Executives* » et des « *Board Members* ». Enfin, même si toutes les anciennetés professionnelles sont représentées dans l'échantillon, il faut noter que les anciennetés professionnelles supérieures à 25 ans sont sous représentées (effectifs inférieur à 30).



Graphique 2.3 : Statistiques descriptives par zone géographique

3. Cadre méthodologique et modèles

3.1. Méthode de calcul des scores des catégories de profil d'attitude face au risque

Cette approche a été initialement proposée par l'équipe de Ingram et al. (2010). Les calculs des scores par catégorie d'attitude face au risque sont réalisés à partir d'une somme algébrique des notes données par les répondants aux différentes questions (10 notes par catégorie).

Un traitement particulier est apporté au score de la catégorie « *Pragmatist* ». Il découle à la fois de l'analyse empirique de la distribution du score des « *Pragmatists* » et de l'hypothèse que les quatre catégories correspondent à des attitudes spécifiques et contradictoires. Ainsi, comme un pur « *Pragmatist* » devrait n'appartient à aucune des trois premières catégories, les auteurs déduisent que le score de la catégorie « *Pragmatist* » est une différence entre un score de référence ($Score_{Ref}$) estimé sur l'ensemble de la population étudiée (ici il s'agissait du panel américain) et les trois scores obtenus sur les autres catégories.

Sur la base d'observation empirique, ils retiennent comme score de référence la moyenne des scores maximum des purs « *Pragmatist* ». Ce score de référence (proposé dans l'étude de D. Ingram et al.) a pour valeur **14** pour le panel américain.

Formelle, pour un individu i quelconque, nous pouvons écrire les relations entre ces notes et ces scores de la manière suivante:

$$Score_{MAX_i} = \sum_{j=1}^{10} Note_{ij}^{Max}, \quad (1)$$

$$Score_{MGR_i} = \sum_{j=1}^{10} Note_{ij}^{MGR}, \quad (2)$$

$$Score_{CONS_i} = \sum_{j=1}^{10} Note_{ij}^{CONS}, \quad (3)$$

$$Score_{PRAG_i} = Score_{Ref} - Max\{Score_{MAX_i}; 0\} - Max\{Score_{MGR_i}; 0\} - Max\{Score_{CONS_i}; 0\}, \quad (4)$$

où $Score_{MAX_i}$ est le score obtenu par l'individu i dans la catégorie d'attitude face au risque « *Maximizers* » et $Note_{ij}^{Max}$ correspond à la note obtenu par l'individu i à la question j dans la catégorie d'attitude face au risque « Maximizers » (notations similaires aux catégories « *Managers* », « *Conservators* » et « *Pragmatists* »). Les valeurs des scores sont comprises entre -20 (cas où la note attribuée à chaque question de la catégorie est de -4) et +20 (cas où chacune des 5 questions a obtenu +4 comme note).

3.2. Règles d'affectation des individus aux profils d'attitude face au risque

Un individu dont le score est compris entre les valeurs -20 et -5 est considéré avoir un « *avis défavorable* » pour cette catégorie d'attitude face au risque. Si le score est compris entre -4 et 4 alors il est considéré avoir un « *avis neutre* ». Enfin, si le score est compris entre 5 et 20, alors il est considéré comme ayant un « *avis favorable* ». Les auteurs admettent que les individus appartiennent à la catégorie où ils donnent un « *avis favorable* ». Pour traiter le cas des individus ayant un « *avis favorable* » dans plusieurs catégories, ils posent un certain nombre de contraintes permettant de concilier les principes d'opposition entre les différentes catégories d'attitude face au risque et de rationalités plurielles. Pour ce faire, ils admettent sous certaines conditions très restrictives, d'écarts entre les scores, qu'un individu puisse avoir un profil mixte (« *blended profil* »).

Ainsi, les trois critères proposés par cette approche afin d'associer un individu à une catégorie d'attitude face au risque sont les suivants :

- **Critère 1** : Les individus considérés comme appartenant à la catégorie « *Pragmatist* » sont ceux qui ont obtenu un « avis favorable » dans cette catégorie (uniquement), ou ayant obtenu des « avis favorables » pour trois catégories ou encre ayant obtenu des avis défavorables pour trois catégories.
- **Critères 2** : si l'avis est favorable pour une seule catégorie, alors l'individu appartient à cette catégorie (unique).
- **Critère 3** : si l'avis est favorable pour deux catégories, tel que l'écart absolu entre les scores de ces deux catégories est inférieur ou égal à 2, l'individu est considéré comme de profil mixte (« *blended profil* »).

Ces critères ainsi que les règles d'affectation d'un individu à une catégorie ou un profil mixte sont également déterminants pour la robustesse des résultats. Des choix trop larges ou trop restrictifs sont de nature à modifier les résultats finaux. Nous avons encadré ces choix lors de la phase d'échanges post-sondage et de validation des résultats auprès de quelques répondants.

3.3. Contrôle et validation du paramètre de calcul du score

L'une des principales critiques que nous avons formulées au modèle paramétrique proposé est la dépendance du score de référence utilisé pour le calcul du score des « *Pragmatists* » aux données. En effet, le fait de déduire ce paramètre de l'échantillon peut être contraint par la taille de l'échantillon d'une part, et soulever le problème de généralisation de cette valeur empirique sur d'autres panel d'autre part. Il n'est pas certain, sans vérification, que la moyenne des scores maximum des purs « *Pragmatists* » américains soit la même que celle des autres zones géographiques considérées (Europe et Afrique). Pour contrôler la stabilité de ce paramètre, nous l'avons ré-estimé de trois façons complémentaires : sans distinction de la zone géographique d'appartenance, par zone géographique et enfin par une approche « *boostrap* » réalisée sur le panel. Nous avons noté que cette étude de sensibilité impact très peu les résultats finaux, car nous trouvons des valeurs autour de 14 (voir annexe 3). Ce qui indique que le choix de ce paramètre sur la base d'un jugement d'expert reste pertinent pour cette étude.

3.4. Test statistique d'association

En statistique, pour avoir le droit d'utiliser les tests paramétriques, il faut vérifier certaines conditions. Quand ces conditions ne sont pas remplies, les tests de remplacement utilisés sont appelés "tests non-paramétriques". Le tableau ci-dessous résume le type de variable, le test paramétrique à utiliser, les conditions d'applicabilité et le test non-paramétrique (de remplacement).

Tests statistiques

Type de variables	Tests paramétriques	Conditions d'application	Test non paramétrique de remplacement
Qualitative et Qualitative	Khi2	Les valeurs de toutes les cases du tableau des effectifs attendus (deuxième tableau) doivent être supérieures ou égales à 5.	Test exact de Fisher
Qualitative et Numérique	T de Student	1) Les écart-types sont égaux(*) 2) Pour chaque groupe, la variable numérique suit une loi normale OU les effectifs sont supérieurs égal à 15	Test des rangs de Wilcoxon
Qualitative et Numérique	F de Fisher (ANOVA)	1) Les écart-types sont égaux(*) 2) Pour chaque groupe, la variable numérique suit une loi normale OU les effectifs sont supérieurs égal à 15	Test de Kruskal-Wallis
Numérique et Numérique	R de Pearson (coefficient de corrélation)	Au moins une des deux variables suit une loi normale	R des rangs de Spearman

Tableau 3.4.1 (a) : Tests statistiques et conditions d'application

(*): les écarts types sont égaux si le plus grand divisé par le plus petit donne un résultat inférieur ou égal à 2

Le test exact de Fisher

Publié en 1922 par Ronald Aylmer Fisher, le test des probabilités exactes de Fisher est une approche non-paramétrique permettant de tester si deux variables qualitatives (nominales ou ordinales) distinctes (X^1, X^2) à deux modalités sont indépendantes. C'est en 1951 que le test exact de Fisher a été étendu aux tableaux $L \times C$, $\forall L, C \geq 2$ par G. H. Freeman et J. H. Halton devenant le test de Fisher-Freeman-Halton (nom est très peu utilisé). Le test exact de Fisher est une alternative au test du Chi2

lorsque la condition sur les effectifs théoriques n'est pas respectée ou bien lorsque le tableau des effectifs croisés est de taille 2×2 du fait des degrés de liberté (ddl) qui valent alors 1.

En général ce test est évité sur des tableaux d'effectifs croisés trop grand du fait du trop grand nombre de calculs que cela implique. Néanmoins certains logiciels ont implémenté des algorithmes permettant de l'exécuter quel que soit la dimension du tableau croisé associé ou quel que soit la taille de l'échantillon.

L'hypothèse préliminaire c'est la présence de variables qualitatives non appariées. L'idée du test exact de Fisher est de comparer la distribution observée avec l'ensemble des combinatoires possibles et issues (au sens statistique) d'une distribution aléatoire. Le test exact de Fisher est, à la base, développé pour les tableaux de taille 2×2 mais nous verrons ensuite sa version étendue.

Nous pouvons alors présenter les effectifs croisés entre (X^1, X^2) :

	Variable	**X1**	
Variables	**Modalités**	**M11**	**M12**
X2	**M21**	a	b
	M22	c	d

Tableau 3.4.1 (b): Tableau 2x2 du test exact de Fisher

Nous sommes donc dans le cadre d'une distribution hyper-géométrique et la formule générale pour la statistique du test exact de Fisher est donnée par,

$$p = \frac{(a+b)!\,(c+d)!\,(a+c)!\,(b+d)!}{a!\,b!\,c!\,d!\,n!}$$

Le calcul de la p-valeur est assez laborieux à obtenir. En effet, il faut comparer la statistique de test associée aux données observées à celles obtenues sur chacune des K configurations conservant les mêmes effectifs marginaux (a + b), (c + d), (a + c), (b + d) que ceux observés et de produit d'effectifs croisés (nommé également « éloignement »), on obtient la p-valeur,

$$p + \sum_{k=1}^{K} p_k$$

où p_k désigne la valeur de la statistique de test calculée sur la configuration numéro k. La p-valeur est comparée directement au seuil de confiance fixé α. L'hypothèse H_0 est: « Il y a indépendance entre les deux variables ».

Le test de Fisher-Freeman-Halton

Freeman et Halton ont proposé une généralisation du test exact de Fisher. En partant du tableau croisé obtenus sur (X^1, X^2) a, respectivement, p, q modalités:

$X_1 \backslash X_2$	x_1	$\cdots$	x_j	$\cdots$	x_q	Totaux
x_1	$n_{1,1}$	$\cdots$	$n_{1,j}$	$\cdots$	$n_{1,q}$	$n_{1,\bullet}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_i	$n_{i,1}$	$\cdots$	$n_{i,j}$	$\cdots$	$n_{i,q}$	$n_{i,\bullet}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_p	$n_{p,1}$	$\cdots$	$n_{p,j}$	$\cdots$	$n_{p,q}$	$n_{p,\bullet}$
Totaux	$n_{\bullet,1}$	$\cdots$	$n_{\bullet,j}$	$\cdots$	$n_{\bullet,q}$	$n_{\bullet,\bullet}$

Tableau 3.4.1 (c) : Tableau pxq de la généralisation du test exact de Fisher

La statistique de test vaut alors,

$$p = \frac{n_{1,.}! \times \ldots \times n_{p,.}! \times n_{.,1}! \times \ldots \times n_{.,q}!}{n_{1,1}! \times \ldots \times n_{p,q}! \times n_{.,.}!}$$

La mesure de l'éloignement pour le calcul de la p-valeur revient alors à calculer le déterminant associé au tableau croisé.

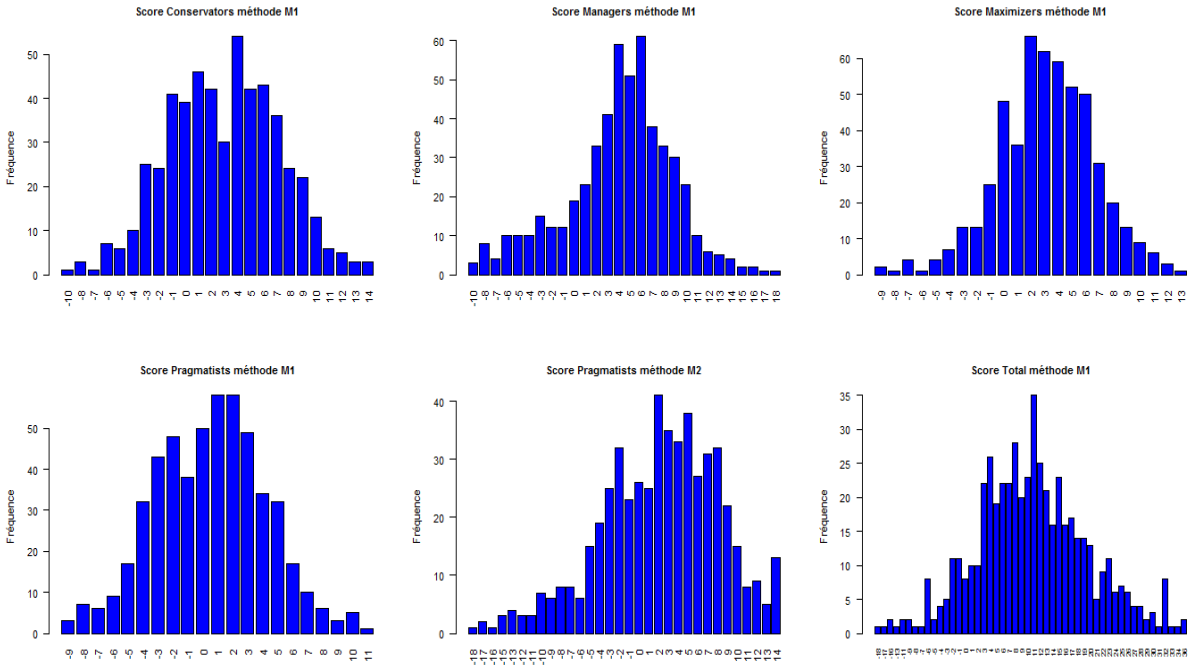
Le test de Fisher-Freeman-Halton est reconnu comme conservateur du fait de sa méthode de calcul faisant intervenir un plus grand nombre de configurations que dans le cas du test exact de Fisher.

4. Applications

4.1. Distribution des scores obtenus suivant l'approche paramétrique

Analyse des distributions des scores de l'approche paramétrique

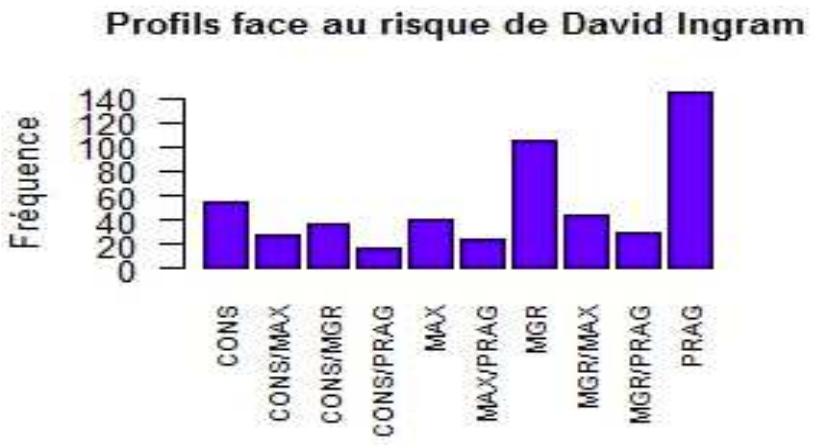
Les distributions des scores ne présentent pas de structures particulières, et très différentes les unes des autres.



Graphique 4.4.1 : Distribution des scores des catégories d'attitude face au risque suivant les méthodes M1 et M2

La composante paramétrique du score de la catégorie « *Pragmatists* » conduit à des résultats très différents en termes de distribution. En effet, la distribution des calculés avec la méthode M2 est plus dispersée surtout sur les valeurs négatives (minimum à -20 contre -10 pour les autres préférences culturelles). La proportion des répondants ayant obtenus des scores supérieurs à 5 est également plus élevée, mais ne dépassant par le seuil de 15. Avec un calcul de score réalisé par la méthode M1, on observe par contre un niveau de scores globalement plus écrasé, avec le score minimum de -10 (vs -20 pour M2) et le score maximum au-delà de 15 (vs 15 pour M2). Finalement, la méthode M2 semble

détecter plus d'individus « d'avis favorable » à la catégorie « *Pragmatists* », avec une répartition de scores plus homogène et caractérisé par un score supérieur ou égal à 5.



Graphique 4.4.2 : Répartition des profils suivant le modèle paramétrique

L'utilisation du modèle paramétrique de Ingram et al., permet de distinguer différents profils individuels appartenant aux quatre catégories d'attitude face au risque, ainsi que des profils mixtes, soit un total de 10 profils différents. Le profil des *Pragmatists* est majoritaire et représente 28% des effectifs, suivi en deuxième position du profil des *Managers* avec 20% des effectifs. Nous trouvons également le profil *Conservators* représentant 10% des effectifs. La faible proportion du profil *Maximizers* (8% des effectifs).

Analyse des répartitions des profils d'attitude face au risque par zone géographique

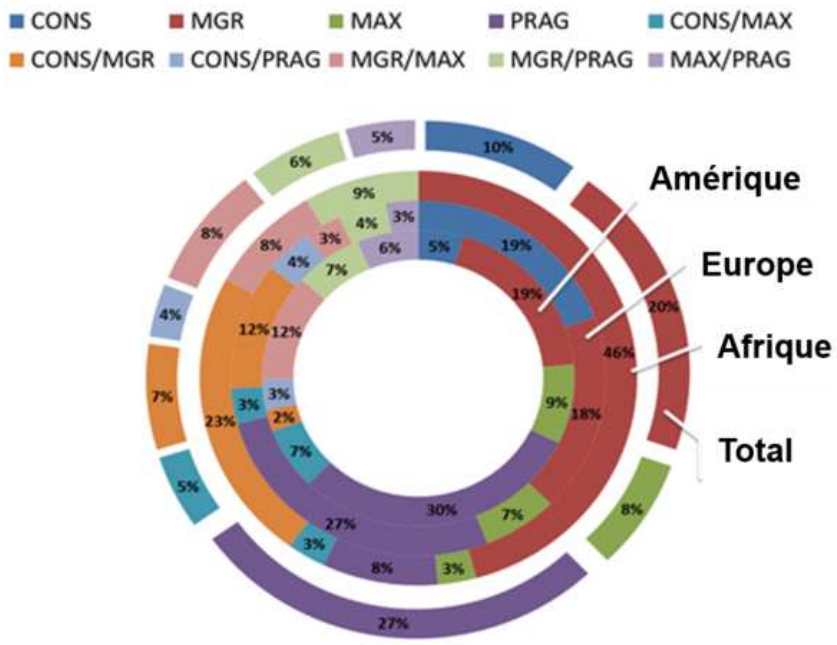
La taille des populations européennes et américaines du panel est identique, nous pouvons donc faire des comparaisons de manière raisonnable. D'une part, nous constatons la proportion prépondérante de la catégorie « *Pragmatists* » dans ces deux populations, en cohérence avec le taux de représentativité de cette catégorie dans le panel (27% des effectifs totaux). Chez les répondants de la zone Europe et Amérique c'est la catégorie « *Pragmatists* » (« *PRAG* ») qui est majoritaire avec 30% des effectifs du panel américain contre 27% du panel européen.

D'autre part, nous observons également de grandes différences dans la répartition des autres catégories d'attitudes face au risque et profils mixtes chez ces deux populations. La deuxième place est occupée chez les américains par la catégorie « *Managers* » (« *MGR* ») représentant 19% des effectifs, alors que chez les européens c'est la catégorie « *Conservators* » (« *CONS* ») représentant également 19% des effectifs. La troisième place est occupée par le profil mixte « *MGR/MAX* » chez les américains et représentant 12% des effectifs, contre la catégorie « *Managers* » chez les européens représentant 18% des effectifs. La quatrième place est occupée par le profil mixte « *CONS/MGR* » chez les européens et représentant 12% des effectifs, contre la catégorie « *Maximizers* » (« *MAX* ») chez les américains représentant 9% des effectifs.

Enfin, la catégorie « *Conservators* » est très faiblement représentée chez les américains (correspond à 5% des effectifs) en 8ème place (après respectivement les profils mixtes « CONS/MAX », « *MGR/PRAG* », « *MAX/PRAG* »), alors que c'est la catégorie « *Maximizers* » qui l'est chez les européens (correspond à 7% des effectifs) en 5ème place.

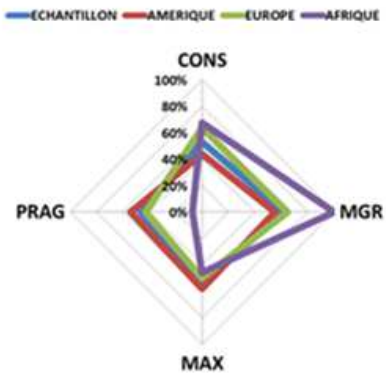
Les résultats de la zone africaine sont regardés de manière spécifique pour deux principales raisons. La première raison est statistique car la taille de ses effectifs de la zone africaine est beaucoup plus faible (6% de l'échantillon contre 54% pour la zone Amérique et 40% pour la zone Europe). La deuxième raison, c'est que les individus interrogés sont tous des salariés d'une banque nationale du même pays. Ainsi, nous observons que la catégorie « *Managers* » représente presque 50% des effectifs des répondants de la zone Afrique (46% des effectifs africains contre presque 19% des effectifs américains et européens).

La deuxième place est occupée par le profil mixte « *CONS/MGR* » représentant 23% des effectifs des répondants de cette zone. Les places suivantes sont respectivement occupées les profils mixtes « *MGR/PRAG* » (9% des effectifs en 3ème place), « *MGR/MAX* » (8% des effectifs en 4ème place) et « *PRAG* » (8% des effectifs en 5ème place).



Graphique 4.4.3 : Répartition des différents profils face au risque par zone géographique

Notons également que les différentes catégories et les différends profils mixtes sont présents dans les différentes zones géographiques, mais en proportions très différentes.

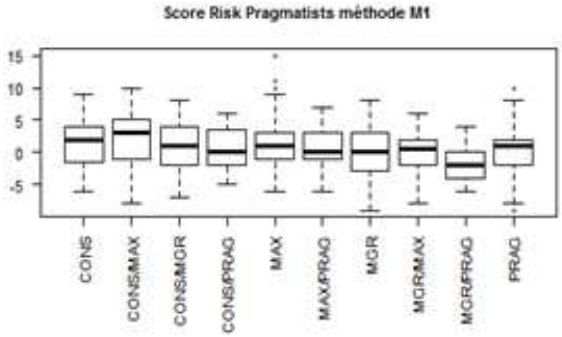


Graphique 4.4.4 : Scores moyens par zone géographique

Ces différents résultats semblent indiquer un faisceau de faibles signaux conduisant à penser qu'il pourrait exister certaines spécifiques géographiques.

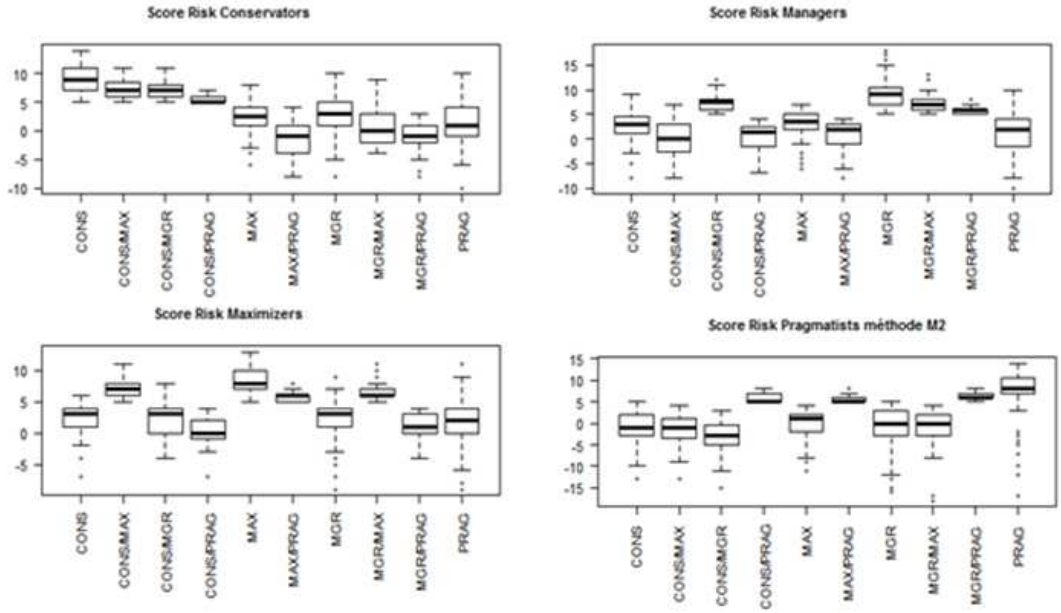
Contrôle de la cohérence du scores « Pragmatist » suivant les méthodes M1 et M2

Avec la méthode M1, le score moyen des individus de la catégorie « *Pragmatists* » n'est pas le plus élevé par rapport à la plupart des autres catégories, il est presque nul avec une dispersion vers des valeurs négatives. Nous ne validons donc pas l'utilisation de cette méthode de calcul direct correspondant à la somme algébrique des notes obtenues aux 10 questions de cette catégorie.



Graphique 4.4.5 : Boîte à moustache (moyenne et dispersion) des scores obtenus par les différents profils face au risque dans la catégorie « *Pragmatists* » avec la méthode M1

L'analyse comparative de la distribution des scores obtenus donne des résultats cohérents par catégorie d'attitude face au risque avec la méthode M2. Par exemple, on observe que les scores moyens des individus appartenant à la catégorie « *Conservator* » ou de profil mixte (*Conservator/Maximizer, Conservator/Manager* et *Conservator/Pragmatist*) sont plus élevés suivant l'axe « *Score risk Conservator* » que ceux des autres catégories.



Graphique 4.4.6 : Boîte à moustache (moyenne et dispersion) des scores obtenus par les différents profils face au risque dans les 4 catégories avec la méthode M2

Nous observons plusieurs individus classés dans différentes catégories d'attitudes face au risque et pour lesquels les scores sont atypiques par rapport à leur catégorie. D'une part, certains scores atypiques sont cohérents avec les différents axes d'analyse. Nous comptons 6 individus (1 « *PRAG* », 2 « *MGR/PRAG* », 1 « *MGR* » et 2 « *MAX* ») qui sont défavorables à la catégorie « *Conservators* ». Nous dénombrons 8 individus (2 « *CONS* », 4 « *MAX* », 1 « *MAX/PRAG* » et 1 « *PRAG* ») qui sont défavorables à la catégorie « *Managers* ». Il y a également 9 individus (2 « *CONS* », 1 « *CONS/PRAG* », 4 « *MGR* » et 2 « *PRAG* ») qui sont défavorables à la catégorie « *Maximizers* ».

Enfin, nous relevons le plus grand nombre d'individus de scores atypiques par rapport à la catégorie « *Pragmatists* ». Nous comptons 10 individus (1 « *CONS* », 2 « *CONS/MAX* », 1 « *CONS/MGR* », 1 « *MAX* », 3 « *MGR* » et 2 « *MGR/MAX* ») qui sont défavorables à cette catégorie, ce qui paraît cohérent. Par contre, nous comptons également 8 individus classés dans la catégorie « *Pragmatists* » avec des scores négatifs. Ceci s'explique essentiellement par l'impact des règles retenues pour affecter les individus dans cette catégorie. En effet, le critère 1 considère que comme « *Pragmatists* » les individus qui ont obtenu un « avis favorable » dans cette catégorie (uniquement), ou ayant obtenu des « avis favorables » pour trois catégories ou encore ayant obtenu des avis défavorables pour trois catégories. De ce fait, l'affectation d'un individu dans la catégorie « *Pragmatists* » ne tient pas compte uniquement de l'avis favorable donné par ce dernier sur cette catégorie, mais également de ses avis sur les autres catégories, d'où cette dispersion des résultats des scores de la catégorie « *Pragmatists* » et cette apparente contradiction qui de fait n'en est pas une.

L'approche paramétrique proposée par David Ingram et al., conduit donc à des résultats satisfaisants par rapport à la répartition des individus dans les différents groupes (y compris les individus affectés à des profils mixtes) et apporte un redressement cohérent du niveau moyen des scores de la catégorie « *Pragmatists* » avec la méthode d'affectation notée M2.

4.2. Analyses des effets des variables sociodémographiques sur les préférences d'attitude face au risque

La méthode paramétrique nous permet de déterminer pour chaque individu sa préférence d'attitude face au risque. Disposant également d'informations sociodémographiques sur les individus observés, nous avons testé dans un premier temps par des tests statistiques adéquats les liens potentiels pouvant exister entre ces différentes variables et les profils d'attitude face au risque. Nous réalisons ensuite différentes analyses croisées entre les variables d'intérêts. Rappelons que les variables ancienneté professionnelle, genre, sont indisponibles pour les individus de la zone américaine. Ce qui limitera notre analyse sur ces variables.

Nous sommes en présence de variables qualitatives et disposons de tableaux de contingences de grandes dimensions. Dans certains cas, les effectifs de certaines modalités sont inférieurs à 5, ce qui nous conduit à choisir le test exact de Fisher pour mesurer les liens potentiels pouvant exister entre les variables testées et leur force. Nous posons comme hypothèse nulle H_0 : « la préférence d'attitude face au risque est indépendante la variable sociodémographique considérée », au seuil de significativité de 5%. Nous utilisons la procédure de test « fisher.test » du logiciel R. Pour interpréter le V de Cramer selon l'échelle de Cohen (voir Gravetter et al, 2013 ; Cohen, 1988), le degré de liberté est défini comme la plus petite valeur entre la dimension des lignes moins 1 (i.e R-1) et la dimension des colonnes moins 1 (i.e C-1).

Analyse macro des associations entre le profil d'attitude face au risque et les variables sociodémographiques

Pour ce premier cas nous croisons les variables profil d'attitude face au risque et zone géographique et nous disposons d'un tableau de contingence de dimension de 30 (10 x 3). La PValue du test est 0,050% très inférieure au seuil de 5%, donc nous rejetons l'hypothèse H_0. Le degré de liberté à retenir pour interpréter le V de Cramer est de valeur 2 (i.e min{R-1, C-1} = C-1= 3-1). Ainsi, d'après l'échelle de Cohen (1988) [54] comme la valeur du V de Cramer est de 40,78% est supérieure à la limite de 35%, nous pouvons donc conclure qu'il existe un lien fort entre ces deux variables.

Pour ce premier cas nous croisons les variables profil d'attitude face au risque et le secteur d'activité et nous disposons d'un tableau de contingence de dimension de 60 (10 x 6). La PValue du test est 0,050% très inférieure à 5%, donc nous rejetons l'hypothèse H_0. Le degré de liberté à retenir est de valeur 5 (i.e min{R-1, C-1} = C-1= 6-1). Ainsi, d'après l'échelle de Cohen (1988) comme la valeur du V de Cramer est de 21,67% est proche mais inférieure à la limite de 22%, nous pouvons donc conclure qu'il existe un lien modéré entre ces deux variables.

Pour ce premier cas nous croisons les variables profil d'attitude face au risque et position dans l'entreprise (correspondant au niveau hiérarchique) et nous disposons d'un tableau de contingence de dimension de 60 (10 x 6). La PValue du test est 0,20% inférieure à 5%, donc nous rejetons l'hypothèse H_0. Le degré de liberté à retenir est de 5 (i.e min{R-1, C-1} = C-1= 6-1). Ainsi, d'après l'échelle de Cohen (1988) comme la valeur du V de Cramer est 17,66%, donc supérieure à la limite du 13% et inférieure à celle du 22%, nous pouvons donc conclure qu'il existe un lien modéré entre ces deux variables.

Test exact de Fisher : Liens de dépendance entre la variable profil d'attitude face au risque et les variables sociodémographiques

Variables d'intérêt	Dimension tableau de contingence	Degré de liberté (V Cramer)	PValue	Hypothèse H_0	V de Cramer	Lien de dépendance	Intensité de la force de l'association	Commentaires (échelle de Cramer)
Zone géographique	10 x 3	2	0,05%	Rejetée	40,78%	Oui	Fort	V Cramer > 35%
Secteur d'activité	10 x 6	5	0,05%	Rejetée	21,67%	Oui	Modéré	13% > V Cramer > 22%
Position hiérarchique	10 x 6	5	0,20%	Rejetée	17,66%	Oui	Modéré	13% > V Cramer > 22%

Tableau 4.2.1 : Liens d'association entre les variables profil d'attitude face au risque et sociodémographiques

Tests alternatifs de comparaisons des zones géographiques

Nous souhaitons valider les résultats précédents spécifiquement sur les zones Europe et Amériques (populations de taille significative). Nous réalisons le test global du Chi-deux d'homogénéité, afin de vérifier si les répartitions de différents effectifs suivant les profils d'attitude face au risque sont équivalentes ou non pour ces deux populations. Nous posons comme hypothèse nulle H0 : « il n'y a pas de différence significative de profils d'attitude face au risque entre ces deux populations » pour un seuil de significativité fixé à 5%.

Les résultats du test donnent une p-Value est inférieure à 5% (X-*squared* = 62.74, df = 9, *p-value* = 3.963×10^{-10}). Nous rejetons donc l'hypothèse nulle à un seuil d'acceptation de 5%. Il y a donc une différence significative de profils d'attitude face au risque entre ces deux zones géographiques.

Analyse de l'intensité de l'association entre zone géographique et les profils d'attitude face au risque

Il existe un lien de dépendance fort entre les préférences d'attitude face au risque et la zone géographique. Nous testons dans ce paragraphe les liens positifs ou négatifs ainsi que l'intensité de la force de l'association entre les différentes catégories de profils d'attitude face au risque et chacune des zones géographiques.

- **Cas de la zone géographique Europe :**

Nous constatons qu'ils existent des liens d'association positifs mais faibles entre le panel de la zone européenne et les attitudes face au risque des catégories de profils « *Managers* » et « *Pragmatists* ». Cette population n'a aucun lien d'association avec les catégories de profils mixtes « *Conservators/Managers* » et « *Conservators/Pragmatists* ». Enfin, elle est liée négativement, c'est-à-dire en opposition, avec les autres catégories de profils d'attitude face au risque.

Test exact de Fisher : Nature de l'association entre les facteurs de la variable profil d'attitude face au risque et *zone Europe*

Catégorie de profil d'attitude face au risque	Dimension tableau de contingence	PValue	Significativité du test	Hypothèse H_0	V de Cramer	Sens de l'association	Intensité de la force de l'association	Commentaires (échelle de Cramer)
Conservateur	2 x 2	2,20%	Oui (PValue < 5%)	Rejetée	-0,1	Négative	Faible	0,3 > |V Cramer| > 0,1
Maximiseur	2 x 2	$2{,}35\times10^{-5}$	Oui (PValue << 5%)	Rejetée	-0,18	Négative	Faible	0,3 > |V Cramer| > 0,1
Manager	2 x 2	$6{,}47\times10^{-6}$	Oui (PValue << 5%)	Rejetée	0,2	Positive	Faible	0,3 > |V Cramer| > 0,1
Pragmatique	2 x 2	$1{,}83\times10^{-8}$	Oui (PValue << 5%)	Rejetée	0,25	Positive	Faible	0,3 > |V Cramer| > 0,1
Conservateur & Maximiseur	2 x 2	$4{,}01\times10^{-5}$	Oui (PValue << 5%)	Rejetée	-0,16	Négative	Faible	0,3 > |V Cramer| > 0,1
Conservateur & Manager	2 x 2	34,24%	Non (PValue >> 5%)	Acceptée				
Conservateur & Pragmatique	2 x 2	30,28%	Non (PValue >> 5%)	Acceptée				
Maximiseur & Pragmatique	2 x 2	2,69%	Oui (PValue << 5%)	Rejetée	-0,1	Négative	Faible	0,3 > |V Cramer| > 0,1
Manager & Maximiseur	2 x 2	$7{,}02\times10^{-9}$	Oui (PValue << 5%)	Rejetée	-0,22	Négative	Faible	0,3 > |V Cramer| > 0,1
Manager & Pragmatique	2 x 2	$7{,}10\times10^{-5}$	Oui (PValue << 5%)	Rejetée	-0,16	Négative	Faible	0,3 > |V Cramer| > 0,1

Tableau 4.2.2 : Liens d'association entre les facteurs de la variable profil d'attitude face au risque et la zone Europe

- **Cas de la zone géographique Amérique :**

Nous constatons qu'ils existent une opposition de préférence d'attitude face au risque entre la zone Europe et la zone Amérique. Cette opposition est traduite par le fait que les mêmes facteurs, les catégories de profils « *Managers* » et « *Pragmatists* », ayant des liens d'association positifs pour la zone Europe sont en opposition pour la zone Amérique comme vu précédemment.

Test exact de Fisher : Nature de l'association entre les facteurs de la variable profil d'attitude face au risque et *zone USA*

Catégorie de profil d'attitude face au risque	Dimension tableau de contingence	PValue	Significativité du test	Hypothèse H_0	V de Cramer	Sens de l'association	Intensité de la force de l'association	Commentaires (échelle de Cramer)
Conservateur	2 x 2	0,58%	Oui (PValue < 5%)	Rejetée	0,12	Positive	Faible	0,3 > |V Cramer| > 0,1
Maximiseur	2 x 2	$1,82x10^{-6}$	Oui (PValue << 5%)	Rejetée	0,2	Positive	Faible	0,3 > |V Cramer| > 0,1
Manager	2 x 2	$4,58x10^{-8}$	Oui (PValue << 5%)	Rejetée	**-0,24**	Négative	Faible	0,3 > |V Cramer| > 0,1
Pragmatique	2 x 2	0,04%	Oui (PValue << 5%)	Rejetée	**-0,16**	Négative	Faible	0,3 > |V Cramer| > 0,1
Conservateur & Maximiseur	2 x 2	$6,59x10^{-5}$	Oui (PValue << 5%)	Rejetée	0,16	Positive	Faible	0,3 > |V Cramer| > 0,1
Conservateur & Manager	2 x 2	5,70%	Non (PValue > 5%)	Acceptée				
Conservateur & Pragmatique	2 x 2	13,21%	Non (PValue >> 5%)	Acceptée				
Maximiseur & Pragmatique	2 x 2	0,98%	Oui (PValue << 5%)	Rejetée	0,11	Positive	Faible	0,3 > |V Cramer| > 0,1
Manager & Maximiseur	2 x 2	$6,17x10^{-7}$	Oui (PValue << 5%)	Rejetée	0,21	Positive	Faible	0,3 > |V Cramer| > 0,1
Manager & Pragmatique	2 x 2	3,65%	Oui (PValue < 5%)	Rejetée	0,1	Positive	Faible	0,3 > |V Cramer| > 0,1

Tableau 4.2.3 : Liens d'association entre les facteurs de la variable profil d'attitude face au risque et la zone USA

Notons également que dans le cas cette zone géographique, les liens d'association positifs comme négatifs sont de faibles intensité. Comme pour la zone Europe, cette population n'a aucun lien d'association avec les catégories de profils mixtes « *Conservators/Managers* » et « *Conservators/Pragmatists* », tout en étant liée positivement avec les autres catégories de profils d'attitude face au risque. Ce qui confirme l'opposition de préférence face au risque de ces deux populations.

- **Cas de la zone géographique Afrique :**

Le panel de la zone Afrique est atypique par rapport aux autres zones géographiques comme nous l'avons vu, puisqu'il s'agit de salariés d'une banque nationale. Cette spécificité se traduit dans les résultats du test exact de Fisher. Pour cette population le lien d'association est positive et reste toujours faible avec la catégorie « *Conservators/Managers* » en premier, puis la catégorie « *Managers/Pragmatists* ». Nous constatons également que ce panel a un lien d'association négatif avec l'attitude face au risque de la catégorie de profil « Pragmatique ». Par contre aucun lien d'association avec les autres catégories.

Test exact de Fisher : Nature de l'association entre les facteurs de la variable profil d'attitude face au risque et *zone Afrique*

Catégorie de profil d'attitude face au risque	Dimension tableau de contingence	PValue	Significativité du test	Hypothèse H_0	V de Cramer	Sens de l'association	Intensité de la force de l'association	Commentaires (échelle de Cramer)
Conservateur	2 x 2	38,60%	Non (PValue >> 5%)	Acceptée				
Maximiseur	2 x 2	34,50%	Non (PValue >> 5%)	Acceptée				
Manager	2 x 2	5,68%	Non (PValue >> 5%)	Acceptée				
Pragmatique	2 x 2	$3,17x10^{-5}$	Oui (PValue << 5%)	Rejetée	-0,17	Négative	Faible	0,3 > |V Cramer| > 0,1
Conservateur & Maximiseur	2 x 2	100,00%	Non (PValue >> 5%)	Acceptée				
Conservateur & Manager	2 x 2	$7,64x10^{-6}$	Oui (PValue << 5%)	Rejetée	**0,28**	Positive	Faible	0,3 > |V Cramer| > 0,1
Conservateur & Pragmatique	2 x 2	100,00%	Non (PValue >> 5%)	Acceptée				
Maximiseur & Pragmatique	2 x 2	100,00%	Non (PValue >> 5%)	Acceptée				
Manager & Maximiseur	2 x 2	73,24%	Non (PValue >> 5%)	Acceptée				
Manager & Pragmatique	2 x 2	3,71%	Oui (PValue << 5%)	Rejetée	**0,11**	Positive	Faible	0,3 > |V Cramer| > 0,1

Tableau 4.2.4 : Liens d'association entre les facteurs de la variable profil d'attitude face au risque et la zone Afrique

Analyse de l'intensité de l'association entre les profils d'attitude face au risque et les secteurs d'activité

Dans la majorité des cas, il n'y a pas de lien entre le secteur d'activité et les catégories de profils d'attitude face au risque, sauf quelques exceptions relevées ci-dessous. Pour les cas d'existence de lien d'association (positif ou négatif), la dépendance est modérée et l'intensité des liens sont faibles entre les préférences d'attitude face au risque et le secteur d'activité.

Test exact de Fisher : Nature de l'association entre les facteurs de la variable profil d'attitude face au risque et *le secteur d'activité*

Catégorie de profil d'attitude face au risque	Secteur d'activité	PValue	Significativité du test	Hypothèse H_0	V de Cramer	Sens de l'association	Intensité de la force de l'association	Commentaires (échelle de Cramer)
Conservateur	Banque	39%	Non (PValue >> 5%)	Acceptée				
	Conseil	**1,13%**	**Oui (PValue < 5%)**	**Rejetée**	**0,12**	**Positive**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Assurance	35%	Non (PValue >> 5%)	Acceptée				
	Université	100%	Non (PValue >> 5%)	Acceptée				
	Réassurance	100%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Maximiseur	Banque	35%	Non (PValue >> 5%)	Acceptée				
	Conseil	67%	Non (PValue >> 5%)	Acceptée				
	Assurance	32%	Non (PValue >> 5%)	Acceptée				
	Université	61%	Non (PValue >> 5%)	Acceptée				
	Réassurance	100%	Non (PValue >> 5%)	Acceptée				
	Autres	61%	Non (PValue >> 5%)	Acceptée				
Manager	Banque	15%	Non (PValue >> 5%)	Acceptée				
	Conseil	**1%**	**Oui (PValue < 5%)**	**Rejetée**	**-0,11**	**Négative**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Assurance	100%	Non (PValue >> 5%)	Acceptée				
	Université	7%	Non (PValue > 5%)	Acceptée				
	Réassurance	65%	Non (PValue >> 5%)	Acceptée				
	Autres	23%	Non (PValue >> 5%)	Acceptée				
Pragmatique	**Banque**	**0,5%**	**Oui (PValue << 5%)**	**Rejetée**	**-0,12**	**Négative**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Conseil	46%	Non (PValue >> 5%)	Acceptée				
	Assurance	70%	Non (PValue >> 5%)	Acceptée				
	Université	56%	Non (PValue >> 5%)	Acceptée				
	Réassurance	12%	Non (PValue >> 5%)	Acceptée				
	Autres	24%	Non (PValue >> 5%)	Acceptée				
Conservateur & Maximiseur	Banque	100%	Non (PValue >> 5%)	Acceptée				
	Conseil	**$1{,}32\times10^{-5}$**	**Oui (PValue < 5%)**	**Rejetée**	**0,23**	**Positive**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Assurance	**0,74%**	**Oui (PValue << 5%)**	**Rejetée**	**-0,13**	**Négative**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Université	100%	Non (PValue >> 5%)	Acceptée				
	Réassurance	100%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Conservateur & Manager	**Banque**	**$1{,}84\times10^{-5}$**	**Oui (PValue < 5%)**	**Rejetée**	**0,26**	**Positive**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Conseil	75%	Non (PValue >> 5%)	Acceptée				
	Assurance	8%	Non (PValue >> 5%)	Acceptée				
	Université	100%	Non (PValue >> 5%)	Acceptée				
	Réassurance	100%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Conservateur & Pragmatique	Banque	100%	Non (PValue >> 5%)	Acceptée				
	Conseil	**1,45%**	**Oui (PValue < 5%)**	**Rejetée**	**0,14**	**Positive**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Assurance	11%	Non (PValue >> 5%)	Acceptée				
	Université	100%	Non (PValue >> 5%)	Acceptée				
	Réassurance	100%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Maximiseur & Pragmatique	Banque	100%	Non (PValue >> 5%)	Acceptée				
	Conseil	37%	Non (PValue >> 5%)	Acceptée				
	Assurance	6%	Non (PValue > 5%)	Acceptée				
	Université	100%	Non (PValue >> 5%)	Acceptée				
	Réassurance	100%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Manager & Maximiseur	Banque	75%	Non (PValue >> 5%)	Acceptée				
	Conseil	**1,92%**	**Oui (PValue < 5%)**	**Rejetée**	**-0,1**	**Négative**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Assurance	**0,19%**	**Oui (PValue << 5%)**	**Rejetée**	**0,13**	**Positive**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Université	61%	Non (PValue >> 5%)	Acceptée				
	Réassurance	40%	Non (PValue >> 5%)	Acceptée				
	Autres	61%	Non (PValue >> 5%)	Acceptée				
Manager & Pragmatique	Banque	5,30%	Non (Pvalue > 5%)	Acceptée				
	Conseil	22%	Non (PValue >> 5%)	Acceptée				
	Assurance	63%	Non (PValue >> 5%)	Acceptée				
	Université	100%	Non (PValue >> 5%)	Acceptée				
	Réassurance	100%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				

Tableau 4.2.5 : Liens d'association entre les facteurs de la variable profil d'attitude face au risque et le secteur d'activité

Nous constatons qu'ils existent des liens d'association positifs mais faibles entre le secteur d'activité « *Conseil* » et les attitudes face au risque des catégories de profils « *Conservators* », « *Conservators/Maximizers* », et « *Conservators/Pragmatists* ». Ce secteur est lié négativement, c'est-à-dire en opposition, avec les catégories de profils d'attitude face au risque « *Managers* » et « *Managers/Maximizers* ».

Ils existent un lien d'association positif mais faible entre le secteur d'activité « *Assurance* » et l'attitude face au risque de la catégorie de profil « *Managers/Maximizers* ». Ce secteur est lié négativement, c'est-à-dire en opposition, avec les catégories de profil d'attitude face au risque « *Conservators/Maximizers* ».

Enfin, ils existent un lien d'association positif mais faible entre le secteur d'activité « *Banque* » et l'attitude face au risque de la catégorie de profil « *Conservators/Managers* ». Ce secteur est lié négativement, c'est-à-dire en opposition, avec les catégories de profil d'attitude face au risque « *Pragmatists* ».

Analyse de l'intensité de l'association entre les profils d'attitude face au risque et la position dans l'entreprise

Dans la majorité des cas, il n'y a pas de lien entre les positions occupées dans l'entreprise et les catégories de profils d'attitude face au risque, sauf quelques exceptions relevées ci-dessous. Pour les cas d'existence de lien d'association (positif ou négatif), la dépendance est toujours modérée et l'intensité des liens est faible entre les préférences d'attitude face au risque et la position occupée dans l'entreprise.

La position membre du conseil (« *Board members* ») a un lien faible et positif avec la catégorie de profil d'attitude face au risque « *Conservators/Managers* ».

Nous constatons qu'il existe un lien d'association positif mais faible entre la position professionnelle « *Managers* » et l'attitude face au risque de la catégorie « *Managers* ». Cette position est liée négativement, c'est-à-dire en opposition, avec les catégories de profils d'attitude face au risque « *Maximizers* » et « *Conservators/Maximizers* ».

La position « *Chercheur* » a un lien faible mais positif avec la catégorie de profil d'attitude face au risque « *Managers* ».

La position « *Executive* » est liée négativement, c'est-à-dire en opposition, avec la catégorie de profil d'attitude face au risque « *Managers* ».

Enfin, la position d'employé (« Staff ») a un lien faible et positif avec la catégorie de profil d'attitude face au risque « *Conservators/Maximizers* ».

Test exact de Fisher : Nature de l'association entre les facteurs de la variable profil d'attitude face au risque et *la position dans l'entreprise*

Catégorie de profil d'attitude face au risque	Position occupée dans l'entreprise	PValue	Significativité du test	Hypothèse H_0	V de Cramer	Sens de l'association	Intensité de la force de l'association	Commentaires (échelle de Cramer)
Conservateur	Membre du Board	72%	Non (PValue >> 5%)	Acceptée				
	Exécutive	19%	Non (PValue >> 5%)	Acceptée				
	Manager	12%	Non (PValue >> 5%)	Acceptée				
	Employé (Staff)	15%	Non (PValue >> 5%)	Acceptée				
	Chercheur (Université)	100%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Maximiseur	Membre du Board	80%	Non (PValue >> 5%)	Acceptée				
	Exécutive	6,34%	Non (PValue >> 5%)	Acceptée				
	Manager	**2,82%**	**Oui (PValue < 5%)**	**Rejetée**	**-0,1**	**Négative**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Employé (Staff)	7,36%	Non (PValue >> 5%)	Acceptée				
	Chercheur (Université)	100%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Manager	Membre du Board	65%	Non (PValue >> 5%)	Acceptée				
	Exécutive	**0,59%**	**Oui (PValue < 5%)**	**Rejetée**	**-0,12**	**Négative**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Manager	**0,68%**	**Oui (PValue < 5%)**	**Rejetée**	**0,12**	**Négative**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Employé (Staff)	23%	Non (PValue > 5%)	Acceptée				
	Chercheur (Université)	**3,35%**	**Oui (PValue < 5%)**	**Rejetée**	**0,12**	**Négative**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Pragmatique	Membre du Board	66%	Non (PValue >> 5%)	Acceptée				
	Exécutive	34%	Non (PValue >> 5%)	Acceptée				
	Manager	54%	Non (PValue >> 5%)	Acceptée				
	Employé (Staff)	33%	Non (PValue >> 5%)	Acceptée				
	Chercheur (Université)	100%	Non (PValue >> 5%)	Acceptée				
	Autres	78%	Non (PValue >> 5%)	Acceptée				
Conservateur & Maximiseur	Membre du Board	15%	Non (PValue >> 5%)	Acceptée				
	Exécutive	12%	Non (PValue >> 5%)	Acceptée				
	Manager	**2,12%**	**Oui (PValue << 5%)**	**Rejetée**	**-0,1**	**Négative**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Employé (Staff)	**$4{,}24\times10^{-7}$**	**Oui (PValue << 5%)**	**Rejetée**	**0,23**	**Négative**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Chercheur (Université)	100%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Conservateur & Manager	**Membre du Board**	**1,10%**	**Oui (PValue < 5%)**	**Rejetée**	**0,13**	**Positive**	**Faible**	**0,3 > |V Cramer| > 0,1**
	Exécutive	43%	Non (PValue >> 5%)	Acceptée				
	Manager	43%	Non (PValue >> 5%)	Acceptée				
	Employé (Staff)	63%	Non (PValue >> 5%)	Acceptée				
	Chercheur (Université)	100%	Non (PValue >> 5%)	Acceptée				
	Autres	8%	Non (PValue >> 5%)	Acceptée				
Conservateur & Pragmatique	Membre du Board	100%	Non (PValue >> 5%)	Acceptée				
	Exécutive	100%	Non (PValue >> 5%)	Acceptée				
	Manager	58%	Non (PValue >> 5%)	Acceptée				
	Employé (Staff)	10%	Non (PValue >> 5%)	Acceptée				
	Chercheur (Université)	100%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Maximiseur & Pragmatique	Membre du Board	60%	Non (PValue >> 5%)	Acceptée				
	Exécutive	42%	Non (PValue >> 5%)	Acceptée				
	Manager	68%	Non (PValue > 5%)	Acceptée				
	Employé (Staff)	72%	Non (PValue >> 5%)	Acceptée				
	Chercheur (Université)	100%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Manager & Maximiseur	Membre du Board	79%	Non (PValue >> 5%)	Acceptée				
	Exécutive	8%	Non (PValue >> 5%)	Acceptée				
	Manager	70%	Non (PValue >> 5%)	Acceptée				
	Employé (Staff)	28%	Non (PValue >> 5%)	Acceptée				
	Chercheur (Université)	63%	Non (PValue >> 5%)	Acceptée				
	Autres	100%	Non (PValue >> 5%)	Acceptée				
Manager & Pragmatique	Membre du Board	47%	Non (Pvalue > 5%)	Acceptée				
	Exécutive	79%	Non (PValue >> 5%)	Acceptée				
	Manager	80%	Non (PValue >> 5%)	Acceptée				
	Employé (Staff)	63%	Non (PValue >> 5%)	Acceptée				
	Chercheur (Université)	51%	Non (PValue >> 5%)	Acceptée				
	Autres	40%	Non (PValue >> 5%)	Acceptée				

Tableau 4.2.6 : Liens d'association entre les facteurs de la variable profil d'attitude face au risque et la position dans l'entreprise

Test statistique d'association entre le profil d'attitude face au risque et les variables complémentaires (genre et ancienneté professionnelle)

Ces variables sont disponibles pour les zones Europe et Afrique. Nous limiterons nos analyses à ses deux zones. Il existe un lien de dépendance fort entre les préférences d'attitude face au risque et la zone géographique. Les tests exacts de Fisher réalisé pour chacune de ces zones géographiques, entre

les variables du genre (respectivement de l'ancienneté professionnelle) et les profils d'attitude face au risque donnent des PValue supérieures à 5%. Nous concluons qu'il n'existe pas pour ces deux populations de lien d'association entre ces variables sociodémographiques et la préférence d'attitude face au risque.

Test exact de Fisher					
Variables d'intérêt		**Dimension tableau de contingence**	**Degré de liberté (V Cramer)**	**Pvalue (seuil 5%)**	**Hypothèse H_0**
Genre		10 x 2	1	76,96%	Acceptée
	EUROPE			9,05%	Acceptée
	AFRIQUE			45,38%	Acceptée
Ancienneté Professionnelle		10 x 4	3	73,11%	Acceptée
	EUROPE			72,56%	Acceptée
	AFRIQUE			93,85%	Acceptée

Tableau 4.2.7 : Liens d'association entre les facteurs de la variable profil d'attitude face au risque et la zone Afrique

5. Discussions

Le modèle paramétrique de détection des profils d'attitude face au risque proposé par Ingram et al. (2014) et testé sur un panel issu de la zone Amérique. Partant de cette approche empirique et basée sur des jugements d'experts, les objectifs de nos travaux étaient doubles. D'abord, nous avons voulu généraliser ce modèle afin de l'appliquer à d'autres zones géographiques. Ensuite, nous avons souhaité tester statistiquement sur notre panel d'étude l'existence ou non de liens d'association entre les profils d'attitude face au risque et certaines variables sociodémographiques complémentaires (secteur d'activité, position occupée dans l'entreprise, le genre, l'ancienneté professionnelle).

D'un point de vue méthodologique, nous avons utilisé le même questionnaire conçus par l'équipe Ingram et al. (2014) pour réaliser des sondages en Europe et en Afrique auprès de professionnels des secteurs financiers (notamment en Banque et en Assurances) entre 2013 et 2015 via le site internet «*SurveyMonkey®*». Cette collecte a été suivie de quelques ateliers et entretiens d'échanges auprès d'un sous-échantillon des individus interrogés. Cette étape nous a permis de faire les validations complémentaires des réponses. La généralisation de l'approche paramétrique a nécessité de s'intéresser aux paramètres du modèle.

Le premier paramètre est celui utilisé dans le calcul du score de la catégorie « *Pragmatists* ». Ce paramètre (score de référence des purs « *Pragmatists* ») a été fixé par les auteurs à 14 sur les données issues de la population de la zone Amérique et sur la base d'un jugement d'expert. Pour vérifier l'adéquation de ce paramètre aux zones géographiques hors Amérique, nous l'avons ré-estimé de trois façons complémentaires, à savoir : sans distinction de la zone géographique d'appartenance, par zone géographique et enfin par une approche *boostrap* réalisée sur le panel. Nous avons noté que cette étude de sensibilité impact très peu les résultats finaux, car nous retrouvons des valeurs autour de 14. Ce qui indique que le choix de ce paramètre sur la base d'un jugement d'expert reste pertinent pour cette étude.

Le second paramètre est le seuil minimal retenu sur les scores des différentes catégories d'attitude face au risque comme critère d'affectation d'un individu à une catégorie d'attitude face au risque donnée. Ce seuil fixé à 5 a également été maintenu suite aux tests de sensibilité réalisés sur ce paramètre. Ce paramètre est également déterminant pour la détection des profils mixtes, car pour associer deux

catégories d'attitude face au risque, leurs scores respectifs doivent obligatoirement dépasser ce seuil minimal.

L'obtention d'un modèle généralisé sur les données issues des zones Europe et Afrique a permis de réaliser la classification des individus de ces zones, puis de réaliser les études statistiques permettant d'atteindre le second objectif de cette étude.

Nous constatons qu'ils existent des liens d'association positifs mais faibles entre le panel de la zone européenne et les attitudes face au risque des catégories de profils « *Managers* » et « *Pragmatists* ». Cette population n'a aucun lien d'association avec les catégories de profils mixtes « *Conservators/Managers* » et « *Conservators/Pragmatists* ». Enfin, elle est liée négativement, c'est-à-dire en opposition, avec les autres catégories de profils d'attitude face au risque. Ceci conduit à une première remarque par rapport aux associations des groupes socio-culturels face au risque (Inglehart et Welzel, Mary Douglas). Dans ce cadre, les européens devraient être classés dans la catégorie des « *Conservators* », profil correspondant au groupe socio-culturel « Egalitariste et Autonomie ». Cette conclusion d'ordre général, ne permet donc pas de mettre en évidence la présence dans la zone Europe des groupes de personnes ayant des attitudes face au risque différentes à ce qui semble être la normalité suivant cette théorie. Ainsi, nous constatons qu'il existe parmi les professionnels du secteur financier de la zone Europe, notamment dans les métiers de l'assurance et de la banque de la zone Europe, des individus opposés à la préférence au profil « Conservateur » d'attitude face au risque.

Nous constatons qu'ils existent une opposition de préférence d'attitude face au risque entre la zone Europe et la zone Amérique. Cette opposition est traduite par le fait que les mêmes facteurs, les catégories de profils « *Managers* » et « *Pragmatists* », ayant des liens d'association positifs pour la zone Europe sont en opposition pour la zone Amérique comme vu précédemment. Notons également que dans le cas cette zone géographique, les liens d'association positifs comme négatifs sont de faibles intensités. Comme pour la zone Europe, cette population n'a aucun lien d'association avec les catégories de profils mixtes « *Conservators/Managers* » et « *Conservators/Pragmatists* », tout en étant liée positivement avec les autres catégories de profils d'attitude face au risque. Ce qui confirme l'opposition de préférence face au risque de ces deux populations. Ceci conduit à une deuxième remarque par rapport aux associations des groupes socio-culturels face au risque (Inglehart et Welzel, Mary Douglas). Dans ce cadre, le panel de la zone Amérique étudié devrait être majoritairement classé dans la catégorie des « *Maximizers* » correspondant au groupe socio-culturel « Autonomie et individualiste ». Cette conclusion d'ordre général, ce confirme sur notre panel. En effet, les catégories « *Maximizers* » et « *Maximizers/Managers* » possèdent les coefficients de Cramer positifs les plus élevés (respectivement 0.21 et 0.20).

Le panel de la zone Afrique est atypique par rapport aux autres zones géographiques comme nous l'avons vu, puisqu'il s'agit de salariés d'une banque nationale. Cette spécificité se traduit dans les résultats du test exact de Fisher. Pour cette population le lien d'association est positive et reste toujours faible avec la catégorie « *Conservators/Managers* » en premier, puis la catégorie « *Managers/Pragmatists* ». Ce résultat combinant le profil « Managers » avec d'autres profils, est cohérent avec l'association des populations africaines à la catégorie « *Managers* » correspondant au groupe socio-culturel « Hiérarchique et Enchâssement » (Inglehart et Welzel, Mary Douglas). Nous constatons également que ce panel a un lien d'association négatif avec l'attitude face au risque de la catégorie de profil « Pragmatique ». Par contre aucun lien d'association avec les autres catégories.

Dans la majorité des cas, il n'y a pas de lien entre les positions occupées dans l'entreprise (respectivement le secteur d'activité) et les catégories de profils d'attitude face au risque, sauf quelques exceptions. Pour les cas d'existence de lien d'association (positif ou négatif), la dépendance est toujours modérée et l'intensité des liens sont faibles entre les préférences d'attitude face au risque et ces variables.

Enfin, les tests exacts de Fisher réalisés entre les variables sociodémographiques (genre et ancienneté professionnelle) montrent qu'ils n'existent pas pour les panels issus des zones Europe et Afrique de lien d'association entre ces variables sociodémographiques et la préférence d'attitude face au risque.

Notons que l'approche paramétrique a pour grand avantage sa simplicité de mise en œuvre et une lecture très agrégée des préférences de catégories d'attitude face au risque. La première limite de cette approche porte sur la détermination des paramètres du modèle, notamment l'estimation du score de référence nécessaire au calcul du score de la catégorie « *Pragmatists* ». Le choix de ce paramètre nécessite d'être justifié lorsque la population change, nous proposons l'utilisation d'une approche par *bootstrap* en estimant la moyenne des scores maximum obtenus sur l'ensemble des différentes catégories (hors la catégorie « *Pragmatists* ») comme présenté dans cette étude. De plus, la méthode de calcul du score de la catégorie « *Pragmatists* » conduit à de la perte d'informations, puisque les notes données par les répondants aux questions de cette catégorie ne sont pas utilisées.

Les critères et des règles d'affectation d'un individu à une catégorie ou un profil mixte sont également déterminants pour la robustesse des résultats. Des choix trop larges ou trop restrictifs sont de nature à modifier les résultats finaux. Il est donc nécessaire d'encadrer ces choix lors de la phase d'échanges post-sondage et de validation des résultats auprès des répondants. L'approche paramétrique utilisée se base sur un mélange de résultats empirique et de jugement d'expert. Ceci conduit à l'obtention d'un profilage des individus avec des règles de décision fixées à priori.

Notre deuxième remarque porte sur les résultats de la zone Europe de profil « *Managers* » et « *Pragmatists* » et n'ayant aucun lien d'association avec les catégories de profils mixtes « *Conservators/Managers* » et « *Conservators/Pragmatists* ». Cette divergence de la classification « Egalitariste et Autonomie » correspondant plutôt à la catégorie de profil « *Conservators* » (Inglehart et Welzel, Mary Douglas) nous conduit toutefois à nous interroger sur l'impact de la traduction du questionnaire de Ingram et al. (2014) de l'anglais en français notamment pour la zone francophone correspondant à la majorité des répondants. Une étude récente réalisée par V. Blum et al. (2019) [53] explore les liens triangulaires entre langage, reporting comptable et incertitude. Elle pose des questions sur les impacts liés à la compréhension et à l'interprétation des mots lors des traductions des termes des normes comptables. Ces auteurs se sont interrogés sur l'impact du choix des mots sur la propension des managers à décider, ainsi que l'appréciation/appréhension des termes d'incertitude que l'on trouve dans les normes comptables. Ces co-auteurs ont mené une expérimentation auprès de certains utilisateurs des normes comptables (experts comptables, directeurs financiers, contrôleurs de gestion, etc…). Cette expérience a consisté à interroger deux fois de suite les mêmes participants sur les interprétations des termes de vraisemblance utilisés dans les normes comptables internationales en deux langues, à savoir le français et l'anglais. L'objectif étant d'évaluer la variabilité des interprétations des termes dans un cadre où les caractéristiques individuelles et le contexte ont été fixés par les auteurs. Ils constatent que la traduction des termes de vraisemblance chez les mêmes individus présente une variabilité dans l'estimation des points et des étendues, ce qui entraîne des différences d'interprétation susceptibles d'entraver la comparabilité visée par les normes internationales. Mais une telle variabilité pourrait également être induite par l'ambiguïté des termes eux-mêmes. Avec un échantillon de contrôle d'une langue, nous démontrons que la variabilité persiste au sein des langues. Ils concluent que ces résultats suggèrent que des interprétations instables de langages d'incertitude existent à la fois dans et à l'intérieur des langues. Ces résultats révèlent les freins liés à l'interprétation de certains termes dans une langue et entre deux langues. Ce biais est difficile à éviter en pratique, mais suggère de bien valider la compréhension des termes auprès des participants afin d'améliorer la comparabilité et d'assurer la cohérence des décisions.

Des études complémentaires de profilage purement statistique pourraient être réalisées en continuité de cette démarche afin de comparer les résultats. Ce profilage statistique pourrait soit être réaliser par l'utilisation de l'Analyse en Composante Multiple pour définir les principaux axes factoriels, puis la mise en œuvre d'une classification de type Classification Ascendante Hiérarchique (CAH) pour la caractérisation des groupes homogènes obtenus. Une méthode prédictive de type CART pourrait aider à récupérer les règles d'affectation. L'approche alternative non supervisée du *Fuzzy-C-Means* (FCM) peut également être une piste possible d'investigation. L'approche non supervisée FCM devrait permettre d'obtenir une segmentation en groupe homogène plus fine, tout en gardant les spécificités géographiques. Elle intègre des probabilités d'affectation des individus dans une classe afin de mieux

capter les incertitudes dans les réponses données aux questions. Elle utilise directement les données brutes fournies en calculant les scores des catégories d'attitude face au risque comme somme algébrique des notes obtenues aux questions. Elle n'est donc pas soumise au risque d'erreur d'estimation de paramètres comme c'est le cas de l'approche paramétrique, donc plus facilement généralisable à n'importe quelle population.

Enfin, les résultats obtenus sont partiellement cohérents avec la théorie socioculturelle du risque et le principe de l'adaptation plurielle des individus face au risque. Toutefois, la robustesse de ces principaux résultats nécessite d'être contrôlée sur un panel de taille plus grande. Il convient également de poursuivre l'étude en complétant les variables sociodémographiques manquantes et en introduisant de nouvelles variables (socioéconomiques du pays ou de la zone géographique, niveau d'éducation, composition familiale, la part variable du salaire, etc…).

6. Conclusions

Nous avons retenu comme approche de référence, le modèle paramétrique de détection les profils d'attitude face au risque calibré de manière empirique et sur des avis d'experts, à partir d'un panel issu de la zone Amérique, proposé par Ingram et al. (2014). Nos travaux ont permis à la fois de généraliser ce modèle afin de l'appliquer à d'autres zones géographiques (Europe et Afrique) et de tester statistiquement sur ces données l'existence de liens d'association éventuels entre les profils d'attitude face au risque et certaines variables sociodémographiques (zone géographique, secteur d'activité, position occupée dans l'entreprise, le genre, l'ancienneté professionnelle). Nous avons constitué un panel et mené des sondages entre 2013 et 2015 auprès de professionnels des secteurs financiers (notamment en Banque et en Assurances).

Les résultats obtenus montrent qu'il existe une opposition de préférence d'attitude face au risque entre la zone Europe et la zone Amérique. La zone Europe est liée positivement mais avec une faible intensité avec les catégories de profils « *Managers* » et « *Pragmatists* » contrairement à la zone Amérique. La zone Amérique n'a aucun lien d'association avec les catégories de profils mixtes « *Conservators/Managers* » et « *Conservators/Pragmatists* », mais reste lié positivement avec les autres catégories de profils d'attitude face au risque contrairement à la zone Europe. Le panel de la zone Afrique composé de salarié d'une banque nationale a un lien d'association positif mais faible avec les catégories « *Conservators/Managers* » et « *Managers/Pragmatists* ». Nous notons que les liens d'association entre les positions occupées dans l'entreprise (respectivement le secteur d'activité) et les catégories de profils d'attitude face au risque sont positives et faibles, mais a contrario il n'existe aucun lien avec les variables sociodémographiques (genre et ancienneté professionnelle) pour les zones Europe et Afrique.

Enfin, ces résultats restent partiellement cohérents avec les conclusions de la théorie socioculturelle face au risque, notamment avec une différence de classification entre cette théorie et le modèle proposé pour la zone Europe, composée en majorité de francophones. Ceci nous amène à nous interroger sur le biais pouvant résulter de la traduction en français des termes du questionnaire conçu initialement en anglais. Cette étude pourrait également être complétée par des approches de classifications statistiques supervisées ou non supervisées (par exemple, le *Fuzzy-C-Means* basée sur la logique floue).

Ce travail pourrait servir d'étape préalable d'investigation des comportements face au risque, avant la mise en œuvre d'un programme d'expérimentation. Les résultats apportent également des éléments de discussion au sein des Comités de Risque et des Conseils des organisations du secteur financier afin de rendre plus agile leur programme *Entreprise Risk Management* (ERM). C'est ainsi que ces acteurs pourraient aligner leurs décisions de gestion des risques aux cycles économiques propres au contexte économique très mouvant ces dernières années.

Références

[1] Bru, B., Bru, M-F, Chung, K.L., (1999), Borel et la martingale de Saint-Pétersbourg, Revue d'histoire des mathématiques, p. 181-247.

[2] Morgenstern, O., Von Neumann, J., (1944), Theory of Games and Economic Behavior, PUP, 1re éd.

[3] Savage L.J., (1954), The Foundations of Statistics, New York: Wiley.

[4] Ellsberg D., (1961), Risk, ambiguity and the Savage axioms, The Quarterly Journal of Economics, 75, 643-669.

[5] Kahneman D., Tversky A., (1979), Prospect theory: an analysis of decision under risk, Econometrica, 47, 263-91.

[6] Allais, M., (1953), Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école américaine, Econometrica, 4, p. 503-546.

[7] Machina, M.J., (1982), Expected Utility Analysis without the Independence Axiom, Econometrica, Vol. 50, No. 2., pp. 277-323.

[8] Fishburn, P.C. (1988), Nonlinear Preference and Utility Theory. Baltimore, Md.: Johns Hopkins University Press.

[9] Campion, B., et al., (2013), Apports et limites de l'expérimentation comme moyen d'investigation des effets éducatifs des médias, ESSACHESS. Journal for Communication Studies, vol. 6, no. 1(11) / 2013 : 257-267, eISSN 1775-352X.

[10] Lemaine, G, & Lemaine, J.-M. (1969). Psychologie sociale et expérimentation. France : Mouton et Bordas.

[11] Matalon, B. (1995). La logique des plans d'expérience. In J.-F. Le Ny & M.-D. Gineste (Eds.), La psychologie : textes essentiels (pp. 48-60). Paris : Larousse. (Publication originale : (1969). In G. Lemaine & J.-M. Lemaine (Eds) Psychologie sociale et expérimentation. Paris : Mouton/Bordas).

[12] Delory, C. (2003). Guide pratique de la recherche en sciences humaines. Namur: Erasme.

[13] Adair, J. G. (1984). The Hawthorne effect: A reconsideration of the methodological artifact. Journal of Applied Psychology, 69(2), 334–345. doi :10.1037/0021-9010.69.2.334

[14] Bloch, H. et al. (1997). Dictionnaire fondamental de la psychologie. Paris : Larousse Bordas.

[15] Valade, Bernard. « Jean Stoetzel : théorie des opinions et psychosociologie de la communication », Hermès, La Revue, vol. 48, no. 2, 2007, pp. 72-74.

[16] Eagly, A. H., & Chaiken, S. (1993). The psychology of attitudes. Orlando, FL, US: Harcourt Brace Jovanovich College Publishers.

[17] Thomas, R., Alaphilippe, D. (1993) : Les attitudes. Paris.

[18] Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), Handbook of social psychology (pp.798–844). Worcester, MA : Clark University Press.

[19] Ebale Moneze, C. (2001). Le développement théorique de la psychologie sociale. Yaoundé : Presses Universitaires de Yaoundé.

[20] LaPiere, R. T. (1934). Attitude versus action. Social Forces, 13, 230-237

[21] Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioural responses to attitude objects. Journal of Social Issues, 25, 41–78.

[22] Mischel, W. (2013). Personality and assessment. Psychology Press.

[23] Rajecki, D. W. (1990). Attitudes. Sinauer Associates.

[24] Bickman, L. (1972). Environmental attitudes and actions. The Journal of social psychology, 87(2), 323-324.

[25] Hofstede, G. (2001). Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations, 2nd ed. Thousand Oaks, CA: Sage.

[26] Hofstede, G., & Minkov, M. (2010). Long-versus short-term orientation: new perspectives. Asia Pacific Business Review, 16(4), 493-504.

[27] Inglehart, R. (1997). Modernization and postmodernization: Cultural, economic and political change in 43 societies. Princeton, NJ: Princeton University Press.

[28] Inglehart, R., Basañez, M., & Moreno, A. (1998). Human Values and Beliefs: A Cross-Cultural Sourcebook, Ann Arbor.

[29] Schwartz, S. H. (2004). Mapping and interpreting cultural differences around the world. In H. Vinken, J. Soeters, & P. Ester (Eds.), Comparing cultures, dimensions of culture in a comparative perspective. Leiden, the Netherlands: Brill.

[30] Schwartz, S. H. (2006). A theory of cultural value orientations: Explication and applications. Comparative Sociology, 5, 137–182.

[31] Hsu, S.-Y., Woodside, A. G. & Marshall, R. (2013). Critical Tests of Multiple Theories of Cultures' Consequences: Comparing the Usefulness of Models by Hofstede, Inglehart and Baker, Schwartz, Steenkamp, as well as GDP and Distance for Explaining Overseas Tourism Behavior. Journal of Travel Research, 52 (6), 679-704

[32] Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. Policy sciences, 9(2), 127-152.

[33] Slovic, P., Fischhoff, B. & Lichtenstein, S. (1980). Facts and fears: Understanding perceived risk. In R. Schwing and W. A. Albers, Jr. (Eds.), Societal risk assessment: How safe is safe enough? (pp. 181-214). New York: Plenum Press. Reprinted in P. Slovic (Ed.), The perception of risk. London: Earthscan, 2001.

[34] Slovic. P. (1987). Perception of Risk. Science, 236 (4799), 280-285.

[35] Slovic, P., Flynn, J., Mertz, C.K., Poumadère, M., & Mays., C. (2000). Nuclear Power and the Public: A Comparative Study of Risk Perception in France and the United States. In O. Renn and R. Rohrmann (eds.). Cross-Cultural Risk Perception: A Survey of Empirical Studies. Boston, MA : Kluwer Academic Publishers.

[36] Mary Douglas (2001) De la souillure : essai sur les notions de pollution et de tabou. La découverte: Paris, p. 200.

[37] Douglas, M., Wildavsky, A. (1983). Risk and culture, an essay on the selection of technical and environmental dangers, Berkeley, University of California Press.

[38] Rohrmann, B. (2000). Cross-cultural studies on the perception and evaluation of hazards. In O. Renn & B. Rohrmann (Eds.), Cross-cultural risk perception: A survey of empirical studies (pp. 103-144). Dordrecht: Kluwer Academic Publishers.

[39] Boholm, A. (1998). Comparative studies of risk perception: a review of twenty years of research. Journal of risk research, 1(2), 135-163.

[40] Ingram D., Thompson E. (2010). « A Universe in Four Parts ». Wilmott Magazine, March 2010

[41] Ingram D., Thompson E., TAYLER (2010). « Eyes Wide Open: Towards Rational Adaptability ». Wilmott Magazine, July 2010

[42] Ingram D. (2009) « Four Seasons of Risk Management ». Actuary Magazine, December 2009

[43] Ingram D., Underwood A. (2010) « Full Spectrum of Risk Attitude ». Actuary Magazine, August 2010

[44] Ingram D., Underwood A. (2011) « The Human Dynamics of the Insurance Cycle and Implications for Insurers: An Introduction to the Theory of Plural Rationalities ». ERM Symposium/SOA Monograph, January 2010

[45] Ingram D., Thompson E. (2011) « Changing Seasons of Risk Attitudes ». Actuary Magazine, April 2011

[46] Ingram D., Thompson E. (2012) « What's Your Risk Attitude? ». Harvard Business Review Blog, June 2012

[47] Dake K. (1991), « Orienting Dispositions in the Perception of Risk: An Analysis of contemporary Worldviews and Cultural Biases »

[48] Aaron Wildavsky et Karl Dake (1990) Theories of Risk Perception: Who Fears What and Why? Daedalus, vol. 119(4), p. 41-60.

[49] Llosa, S. (1997). L'analyse de la contribution des éléments du service à la satisfaction : un modèle tétraclasse. Décisions marketing, 81-88.

[50] Tremblay, P., Beauregard, B. [2006], « Application du modèle tétraclasse aux résultats de sondage d'un organisme public : le cas de la régie des rentes du Québec »,

[51] Morineau, A. (1984). Note sur la caractérisation statistique d'une classe et les valeurs-tests. Bulletin Technique Centre Statistique Informatique Appliquées, 2(1-2), 20-27.

[52] Lebart, L., Piron, M., & Morineau, A. (2006). Statistique exploratoire multidimensionnelle : visualisation et inférences en fouilles de données.

[53] Véronique Blum & Pierre-Emmanuel Thérond & David Alexander & Emmanuel Laffort & Solvita Jancevska, 2019. "New developments in language issues in accounting regulation: likelihood terms and the certainty of uncertainty," Working Papers hal-01991845, HAL.

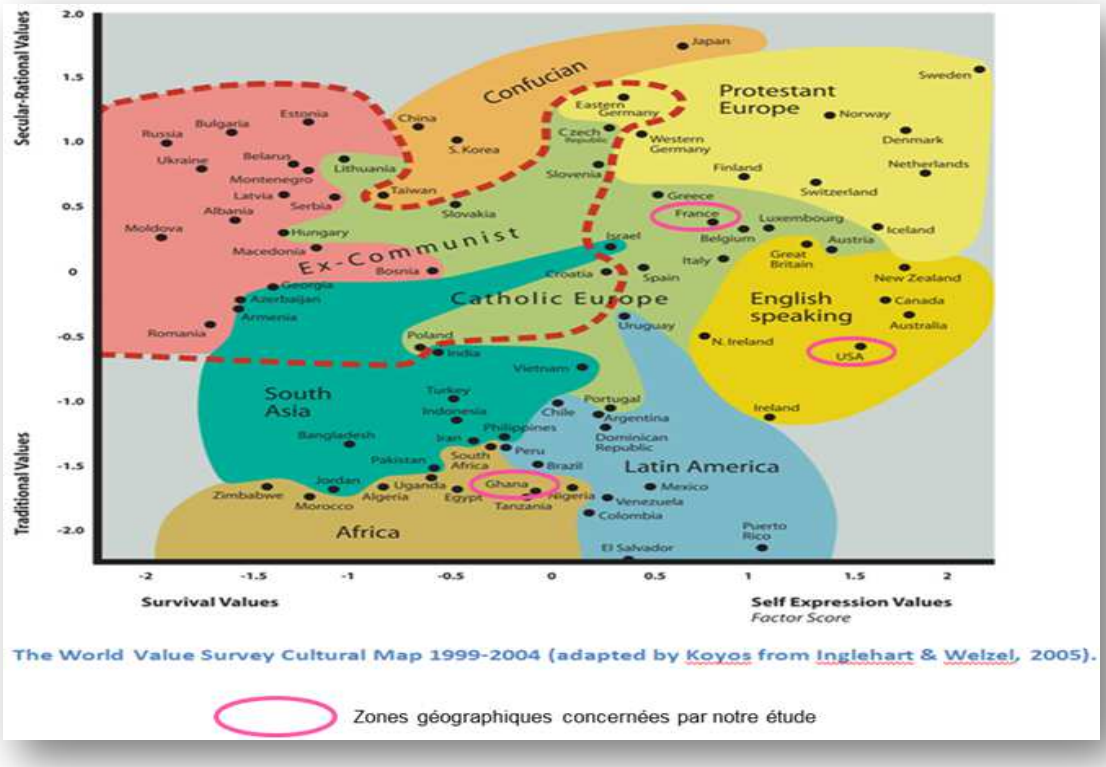
[54] Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Second Edition. Hillsdale, N.J, L. Erlbaum Associates.

Annexes

Annexe 1 :

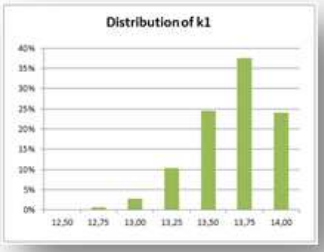
Risk Attitude	Questions
Conservators	Progressive income taxes are more fair than flat taxes.
Conservators	It's best to hire inexperienced people and train them ourselves.
Conservators	Fines for companies that break laws should be proportionate to their profits.
Conservators	Some things are too important to be decided by cost-benefit considerations.
Conservators	Everyone in this company should be paid enough to support a good standard of living.
Conservators	A firm should not take on any more risk than it absolutely must.
Conservators	Sometimes the firm must give a higher priority to security than to profits.
Conservators	If people in this company were treated more equally, then we would have fewer problems.
Conservators	The degree of inequality of income in the company is a problem.
Conservators	It is always important to keep the worst-case outcome in mind when making decisions.
Managers	There should be stricter regulation of our industry to protect the better-run companies from competition that is just cutting corners.
Managers	The company has the right to monitor employee emails, web browsing, and phone calls.
Managers	We should take maximum advantage of new technology to solve our firm's problems.
Managers	I expect meetings to start on time.
Managers	Management and the Board must clearly take responsibility for the safety of the firm.
Managers	One of the problems with running a company today is that too many employees challenge authority.
Managers	The company has a longstanding heritage that should be preserved.
Managers	One of the most important steps in decision making is to gather input and recommendations from experts on the key issues.
Managers	Our firm should have a code of conduct and strictly enforce it.
Managers	Employees should be given information on a need-to-know basis.
Maximizers	People who work the hardest are not always rewarded enough at some companies.
Maximizers	It is not vital for the company to worry about how equally people are compensated.
Maximizers	People get to the top as a result of hard work.
Maximizers	Sometimes the firm must give a higher priority to profits than to security.
Maximizers	Continued growth is important to the health of the company.
Maximizers	The most intelligent people should get the most responsibility.
Maximizers	People in leadership positions have earned better compensation and perks.
Maximizers	The country is best off if it gives companies the freedom to prosper.
Maximizers	In a fair system, people with more ability should earn more.
Maximizers	If everyone followed all of the rules and regulations to the letter, business would never get done.
Pragmatists	The firm is better off going it alone and not placing much trust in other organizations.
Pragmatists	Companies need to be highly adaptable to succeed in the long run.
Pragmatists	The key to business success is to keep your options open, staying away from large long term commitments.
Pragmatists	It is important to avoid getting too tied to any theory of business so that you are free to adjust when the situation changes.
Pragmatists	The treatment of companies by the regulators is often unfair.
Pragmatists	Sometimes it seems that competitors get ahead mostly because they have been lucky.
Pragmatists	We do not seem to have any influence in how things get done in our industry segment.
Pragmatists	A successful business manager must be comfortable dealing with ambiguity and uncertainty.
Pragmatists	It is not wise to call attention to others' violations of laws and regulations.
Pragmatists	It's usually not worthwhile to put much energy into long term planning, because you always have to change what you do to adapt to what's really going on.

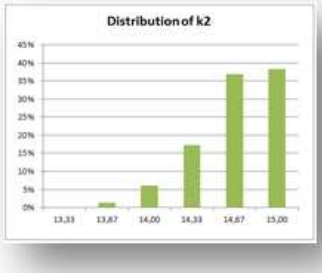
Annexe 2 : Cartographie des groupes socio-culturels (adapté par Koyos de Inglehart et Welzel)



Annexe 3 : calibrage du score de référence du modèle paramétrique par bootstrap

	k1	k2
Mean	13,67	14,68
Median	13,75	14,67
Q1	13,50	14,67
Q3	13,75	15,00
Variance	0,08	0,12
Standard deviation	0,29	0,34
CV	0,0212	0,0233

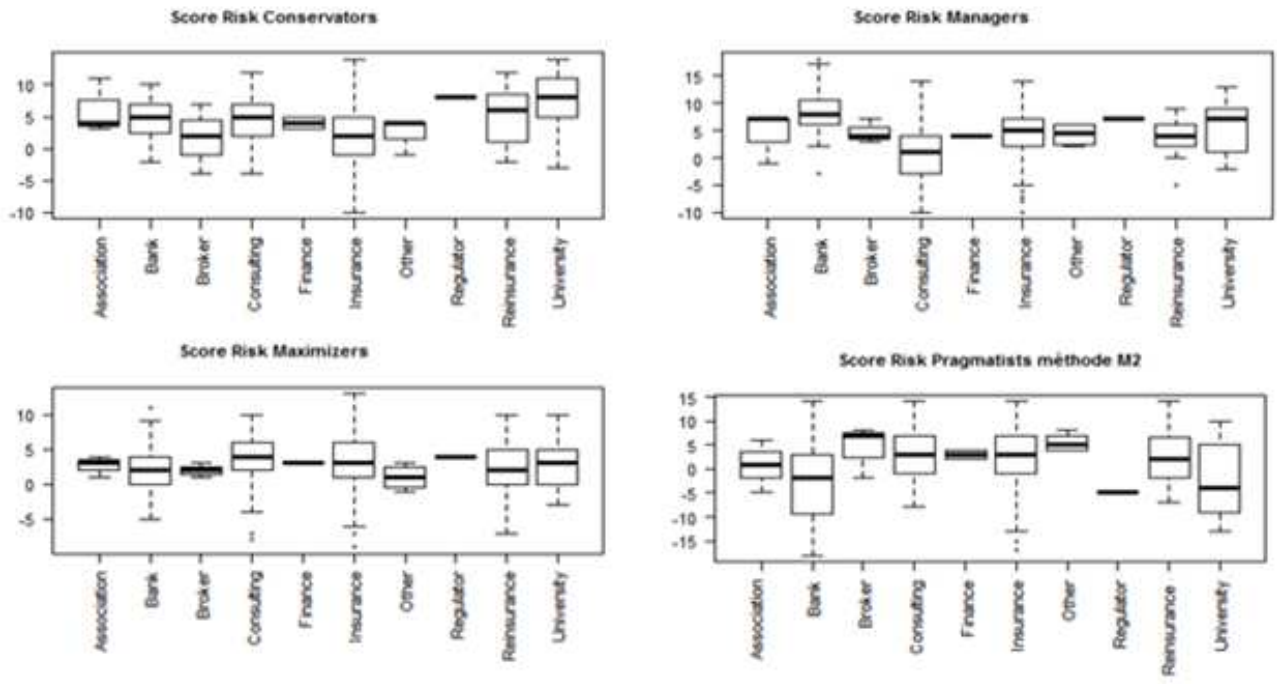




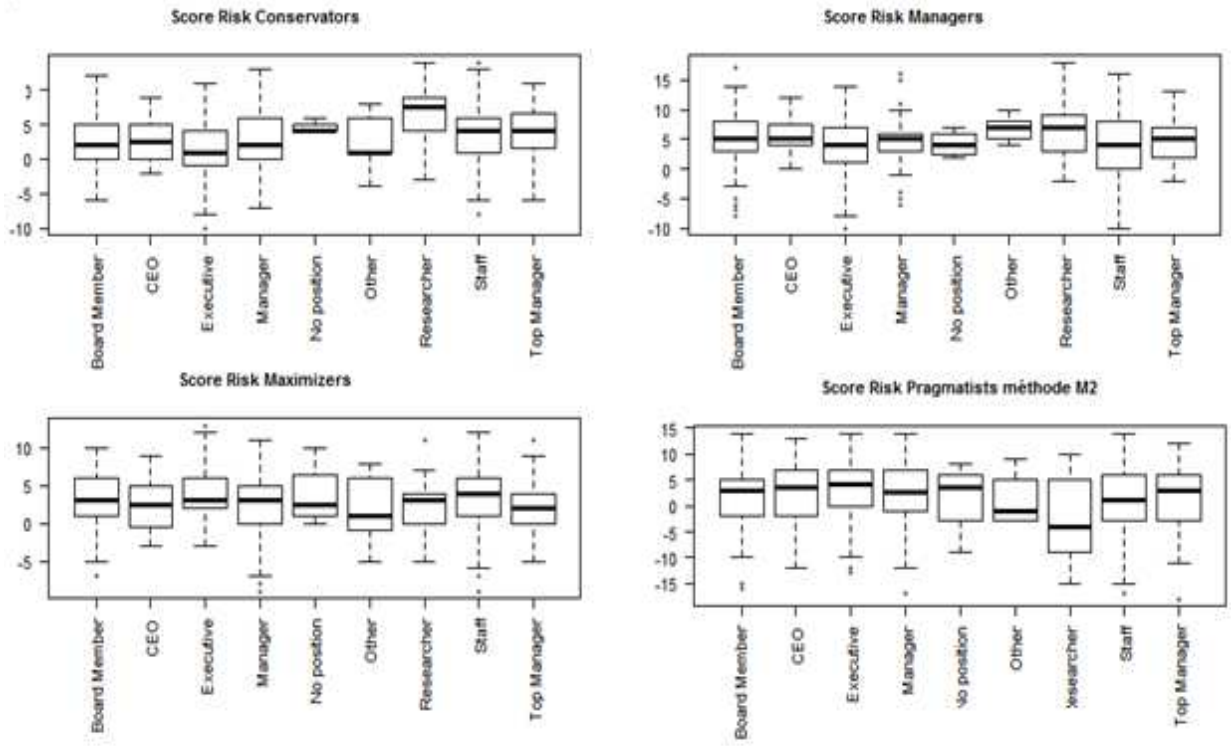
Annexe 4 : Scores moyens normalisés des catégories d'attitude face au risque par zone géographique

	Scores moyens par zone géographique			
Risk preference	ECHANTILLON	AMERIQUE	EUROPE	AFRIQUE
CONS	2,98	1,74	4,39	4,89
MGR	4,16	3,25	4,55	9,23
MAX	3,10	3,62	2,57	2,00
PRAG	2,16	3,15	1,64	- 3,06

Graphique 2 : Boite à moustache (moyenne et dispersion) des scores obtenus par les différents secteurs d'activité dans les 4 catégories d'attitude face au risque avec la méthode M2



Graphique 2 : Boite à moustache (moyenne et dispersion) des scores obtenus par les différentes positions professionnelles dans les 4 catégories d'attitude face au risque avec la méthode M2



Conclusion générale

CONCLUSION GENERALE

Les différents travaux présentés dans ce manuscrit s'inscrivent dans le cadre de la mesure et de la gestion des risques. L'objectif de ce travail a été, de présenter trois différentes thématiques traitant des problématiques d'évaluation des risques, notamment pour un assureur de personne. Le premier chapitre « Prédiction des paramètres de provisionnement individuel avec des méthodes d'apprentissage ensembliste en assurance non-vie » présente une méthodologie de modification des algorithmes d'apprentissage automatique ensembliste pour les adaptés aux données censurées (à droite) pour le provisionnement des sinistres en assurance non vie. Le second chapitre « Approche quantitative d'estimation de chocs des risques d'incidence et de maintien en assurance non-vie sous Solvabilité II », propose une approche de calibrage des paramètres spécifiques à l'entité (c'est-à-dire sur les données propres d'une compagnie d'assurance) dans le cadre des calculs de fonds propres économiques sous la norme Solvabilité II. Le troisième chapitre « *Long-Terme Care : Construction of an economic balance sheet and solvency capital requirement calculation in solvency II* » est une présentation de la démarche de calcul des fonds propres sous Solvabilité II du risque de dépendance. Nous y présentons les limites de la formule standard de cette norme et proposons des pistes d'amélioration. Enfin le quatrième et dernier chapitre est une « Etude empirique des attitudes face au risque dans le secteur de la banque et assurance ». Cette étude est une généralisation du modèle paramétrique de détection des préférences d'attitude face au risque des individus.

1. Prédiction des paramètres de provisionnement individuel avec des méthodes d'apprentissage ensembliste en assurance non vie

En assurance non-vie, pour la tarification des affaires ou le provisionnement des sinistres, il est d'usage d'avoir recourt à des modèles statistiques pour prédire les probabilités d'incidence et de maintien sur une période à partir de données de durée. Ces données sont souvent soumises au phénomène de censure à droite, caractérisant les individus partiellement observés durant la période d'estimation des probabilités. Ainsi, en matière de provisionnement non-vie, différentes méthodes classiques telles que le *Chain Ladder*, le Micro *Level Reserving* (GLM) et l'estimateur de Kaplan Meier sont couramment utilisées par les praticiens. Malgré leur facilité de mise en œuvre, elles présentent certaines limites et ne permettent pas d'utiliser toute la richesse des données individuelles disponibles pouvant porter du signal donc contribuer à expliquer le risque. Ces données peuvent être des données identitaires, des informations transactionnelles et relationnelles, des données démographiques, des variables contractuelles et sur l'état des assurés.

Récemment, le modèle d'arbres de régression censurées (*Tree base censored*) une adaptation du modèle CART a été proposée comme une alternative gérant à la fois les censures à droite et utilisant différentes variables explicatives disponibles afin d'extraire les informations pertinentes pour la prédiction. La principale limite de cette solution est son instabilité héritée du modèle CART. De plus les tests ont été réalisés sur des données simulées.

La valeur ajoutée principale de notre étude est double. D'une part, nous proposons de corriger l'instabilité de CART par l'utilisation de deux algorithmes d'apprentissage automatique (*Random Forest* et *Gradient Boosting*) fournissant des prédictions plus robustes et plus stables. Ces algorithmes ont été adaptés pour prendre en compte les données censurées à droite par l'introduction comme paramètre des poids de Kaplan-Meier (*Inverse Probability of Censoring Weighting*). Nous avons également adapté les métriques de performance pour tenir compte de la présence des censures à droite dans les données et de l'échelle de grandeur des variables à expliquer. D'autre part, l'illustration par des tests sur deux jeux de données réelles, montre que notre approche semble robuste et améliore les résultats. Nous

appliquons ces algorithmes aux données de portefeuilles d'assurance de prêts et de prévoyance collective d'une compagnie d'assurance de personnes pour estimer les durées de maintien et les charges sinistres. Les différentes méthodes proposées fournissent des résultats plus performants qu'à ceux des modèles classiques. Elles apportent des améliorations significatives à la méthode de Kaplan Meier, et permettent de faire du provisionnement tête par tête, en exploitant au mieux toutes les informations disponibles. Nous avons également constaté que les prédictions des risques en prévoyance collective pouvaient être améliorées en complétant les bases de l'assureur. Une manière d'enrichir les données de l'assureur est l'utilisation des données publiques (Données INSEE, Base DAMIR…). Les adaptations proposées pourraient s'adapter facilement à la prédiction des risques de maintien en dépendance et de prévoyance individuelle.

Enfin, notons que ces méthodes de Machine Learning ont également certaines limites. Elles ont besoin de quantité importante de données pour le calibrage des modèles. Elles produisent des résultats certes performants mais difficilement interprétables.

2. Approche quantitative d'estimation de chocs des risques d'incidence et de maintien en assurance non-vie sous Solvabilité II

L'approche proposée a servi à l'estimation des chocs des risques d'incidence et de maintien (ou de rétablissement) de la garantie arrêt de travail d'un contrat d'assurance de prêts (ou emprunteurs) d'une compagnie d'assurance vie. Ce portefeuille dispose de plus de 12 millions de têtes assurées et de générations de souscription dépassant 20 années d'ancienneté. L'une des principales difficultés de calibrage des paramètres des modèles retenus est la disponibilité de données de taille et de profondeur d'historique suffisantes. L'introduction d'une variante normalisée de la théorie de la crédibilité à variation limitée ou crédibilité américaine (TCVL) dans la modélisation permet d'estimer des niveaux de chocs spécifiques par ancienneté des polices du portefeuille. Les niveaux de chocs obtenus conduisent à des besoins de capitaux requis (SCR) en moyenne de l'ordre de 50% à 60% plus faibles que ceux estimés avec les paramètres forfaitaires de la formule standard. La mise en place d'une approche USP pourrait donc être source d'économie de besoin de capital requis dans certains cas.

Le cadre méthodologique proposé dans cet article pourrait servir de base de discussion et d'implémentation chez les compagnies disposant de données suffisantes et de qualité pour estimer les paramètres spécifiques des risques d'incidence et de maintien en assurance de personne. Ce cadre est également adaptable aux travaux de gestion des risques notamment lors des évaluations ORSA de la norme Solvabilité II. Ainsi, l'usage d'une modélisation très précise des risques permet de détecter de manière anticipée les dérives éventuelles.

L'introduction des facteurs de crédibilité est une approche prudente car permettant de tenir compte au moins partiellement de l'expérience de la compagnie, mais il reste à approfondir la problématique du choix des coefficients de crédibilité, à savoir, l'hypothèse de l'application des coefficients de crédibilité calibrés sur lois centrales à des paramètres de chocs. De plus, l'approche série temporelle nécessite de disposer d'historique conséquent pour son utilisation. Elle peut donc ne pas être adaptée dans la plupart des cas, ce qui conduirait à l'utilisation de l'approche modèle linéaire généralisé de type Log Normale qui fournit des paramètres un peu plus sévères.

3. Long-Terme Care: Construction of an economic balance sheet and solvency capital requirement calculation in solvency II

In this chapter, we have mainly detailed the methodology for assessing the capital requirement for the risk of underwriting the Long Term Care risk. Even if this chapter does not deal in depth with Market

SCR and operational SCR, one should supplement the issue risk consideration, by an assessment of financial and operational risks. Thus, the insurer will have a holistic assessment of the required capital to deal with the underlying risks of the Long Term Care coverage.

As mentioned in this chapter, the rules for calculating capital required under Solvency II, the prudential standard in force since January 1, 2016, have not been specifically designed for the Long Term Care risk. They seem to be poorly adapted to this risk, which poses a number of problems of coherence and measurement of the actual level of capital needed for this risk. Nevertheless, regulators offer the insurers the opportunity to deviate from these rules by using a Partial Internal Model. As mentioned above, designing PIM presents its own challenges. As a result, the majority of insurers subject to this standard are not satisfied with the treatment of the Long Term Care risk under Solvency II but are forced to apply the standard formula that generates a very high capital requirement.

Given the long duration of liabilities and the significant level of the associated SCR, it is indispensable for an insurer to hold a projection model taking into account management actions. These choices allow to take into account monitoring tools to be used in order to calculate the appropriate capital requirement. Thus, the integration of management actions is a means for the insurer to meet the prudential requirements by ensuring the continuity of a product with lifetime liabilities.

Finally, on February 28, 2018, the European Insurance and Occupational Pensions Authority (EIOPA) published new recommendations to amend the rules used to calculate the standard SCR formula. Unfortunately, these recommendations do not make any changes to the treatment of lifetime Long Term Care insurance benefits. However, it should be noted that a new standard formula revision is planned for 2020.

4. Etude empirique des attitudes face au risque dans le secteur de la banque et assurance

Nous avons retenu comme approche de référence, le modèle paramétrique de détection des profils d'attitude face au risque calibré de manière empirique et sur des avis d'experts, à partir d'un panel issu de la zone Amérique, proposé par Ingram et al. (2014). Nos travaux ont permis à la fois de généraliser ce modèle afin de l'appliquer à d'autres zones géographiques (Europe et Afrique) et de tester statistiquement sur ces données l'existence de liens d'association éventuels entre les profils d'attitude face au risque et certaines variables sociodémographiques (zone géographique, secteur d'activité, position occupée dans l'entreprise, le genre, l'ancienneté professionnelle). Nous avons constitué un panel et mené des sondages entre 2013 et 2015 via le site « *SurveyMonkey©* » auprès de professionnels des secteurs financiers (notamment en Banque et en Assurances).

Les résultats obtenus montrent qu'il existe une opposition de préférence d'attitude face au risque entre la zone Europe et la zone Amérique. La zone Europe est liée positivement mais avec une faible intensité avec les catégories de profils « *Managers* » et « *Pragmatists* » contrairement à la zone Amérique. La zone Amérique n'a aucun lien d'association avec les catégories de profils mixtes « *Conservators/Managers* » et « *Conservators/Pragmatists* », mais reste lié positivement avec les autres catégories de profils d'attitude face au risque contrairement à la zone Europe. Le panel de la zone Afrique composé de salarié d'une banque nationale a un lien d'association positif mais faible avec les catégories « *Conservators/Managers* » et « *Managers/Pragmatists* ». Nous notons que les liens d'association entre les positions occupées dans l'entreprise (respectivement le secteur d'activité) et les catégories de profils d'attitude face au risque sont positives et faibles, mais a contrario il n'existe aucun lien avec les variables sociodémographiques (genre et ancienneté professionnelle) pour les zones Europe et Afrique.

Enfin, ces résultats restent partiellement cohérents avec les conclusions de la théorie socioculturelle face au risque, notamment avec une différence de classification entre cette théorie et le modèle proposé pour la zone Europe, composée en majorité de francophones. Ceci nous amène à nous interroger

sur le biais pouvant résulter de la traduction en français des termes du questionnaire conçu initialement en anglais. Cette étude pourrait également être complétée par des approches de classifications statistiques supervisées ou non supervisées (par exemple, le *Fuzzy-C-Means* basée sur la logique floue).

Ce travail pourrait servir d'étape préalable d'investigation des comportements face au risque, avant la mise en œuvre d'un programme d'expérimentation. Les résultats apportent également des éléments de discussion au sein des Comités de Risque et des Conseils des organisations du secteur financier afin de rendre plus agile leur programme *Entreprise Risk Management* (ERM). C'est ainsi que ces acteurs pourraient aligner leurs décisions de gestion des risques aux cycles économiques propres au contexte économique très mouvant ces dernières années.

BIBLIOGRAPHIE GENERALE

Chapitre 1

Antonio, K., Plat, R. (2014). “Micro-level stochastic loss reserving for general insurance”. Scandinavian Actuarial Journal 2014/7, 649-669.

Astesan, E., (1938). « Les réserves techniques des sociétés d'assurances contre les accidents automobiles » ; Librairie générale de droit et de jurisprudence, 1938.

Arjas, E., (1989). “The claims reserving problem in non-life insurance: some structural ideas”. ASTIN Bulletin 19/2, 139-152.

Badescu, A.L., Lin, X.S., Tang, D., (2016). “A marked Cox model for the number of IBNR claims: theory”. Insurance: Mathematics & Economics 69, 29-37.

Badescu, A.L., Lin, X.S., Tang, D., (2016). “A marked Cox model for the number of IBNR claims: estimation and application”. Version March 14, 2016. SSRN Manuscript 2747223.

Bandyopadhyay, S., Wolfson, J., Vock, D.M., Vazquez-Benitez, G., Adomavicius, G., Elidrisi, M., Johnson, P.E., O’Connor, P.J., (2015). “Data mining for censored time-to-event data: a Bayesian network model for predicting cardiovascular risk from electronic health record data”, Data Min. Knowl. Disc. 29 (4) (2015) 1033–1069.

Bang, H., Tsiatis, A.A., (2000). “Estimating medical costs with censored data”, Biometrika 87 (2) 2000, 329 -343.

Baudry, M., Robert, C.Y. (2017). “Non parametric individual claim reserving in insurance”. Preprint.

Biganzoli, E., Boracchi, P., Mariani, L., Marubini, E., (1998). “Feed forward neural networks for the analysis of censored survival data: a partial logistic regression approach”, Stat. Med. 17 (10) 1998, 1169–1186.

Breiman, L., (2001). “Random Forests”. Machine Learning 45 (1): 5-32. doi :10.1023/A: 1010933404324

Breiman, L., (1996). Bagging predictors. Machine learning, 24 (2), 123-140.

Breiman, L., Friedman, J., Olshen, R., Stone, C., (1984). “Classification and regression trees”. Wadsworth Books, 358.

Blanco, R., Inza, I., Merino, M., Quiroga, J., Larrañaga, P., (2005). “Feature selection in Bayesian classifiers for the prognosis of survival of cirrhotic patients treated with TIPS”, J. Biomed. Inform. 38 (5) 2005, 376 -388.

Chen,T., Guestrin, C., (2016). “XGBoost : A Scalable Tree Boosting System”, Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’16, 2016, p. 785–794.

Chen, Y., Jia, Z., Mercola, D., & Xie, X. (2013). “A Gradient Boosting Algorithm for Survival Analysis via Direct Optimization of Concordance Index”. Computational and Mathematical Methods in Medicine, 2013, 873595. http://doi.org/10.1155/2013/873595

Cooney, M.T., Dudina, A.L., Graham, I.M., (2009). “Value and limitations of existing scores for the assessment of cardiovascular risk: a review for clinicians”, J. Am. Coll. Cardiol. 54 (14) 2009, 1209–1227.

Collins, G.S., Altman, D.G.,(2009). “An independent external validation and evaluation of QRISK cardiovascular risk prediction: a prospective open cohort study”, Br. Med. J. 339 (2009) b2584.

D’Agostino, R.B., Vasan, R.S., Pencina, M.J., Wolf, P.A., Cobain, M., Massaro, J.M., Kannel, W. B., (2008). “General cardiovascular risk profile for use in primary care: the Framingham heart study”, Circulation 118 (4) E86.

Davison, A. C., Hinkley, D. V., (1997). “Bootstrap Methods and Their Application”, Cambridge University Press, 28 octobre 1997. ISBN 0-521-57471-4.

Freund, Y., Schapire, R., (1996). “Experiments with a new boosting algorithm”. In Proceedings of the Thirteenth International Conference on Machine Learning.

Friedland, J., (2010). “Estimating Unpaid Claims Using Basic Techniques; Estimating Unpaid Claims Using Basic Techniques”.

Friedman, J-H., (2002). “Stochastic gradient boosting”, Computational Statisrics and Data Analysis 38.

Gregorutti, B., Michel, B., Saint-Pierre, O., (2014). « Corrélation et importance des variables dans les forêts aléatoires ».

Genuer, R., Poggi, J-M., Tuleau-Malot, C. (2012). « Variable selection using Random Forests”.

Goff, D.C., Lloyd-Jones, D.M., Bennett, G., Coady, S., D’Agostino, R.B., Gibbons, R., Greenland, P., Lackland, D.T., Levy, D., O’Donnell, C.J., Robinson, J.G., Schwartz, J.S., Shero, S. T., Smith, S.C., Sorlie, P., Stone, N.J., Wilson, P.W., (2013). “ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association task force on practice guidelines”, J. Am. Coll. Cardiol. 63 (25) (2014) 2935–2959.

Harej, B., Gächter, R., Jamal, S. (2017). “Individual claim development with machine learning”. ASTIN Report.

Hesselager, O., (1994). “A Markov model for loss reserving”. ASTIN Bulletin 1994, 24(2), 183-193.

Hoem, J.M., (1971), « Point Estimation of Forces of Transition in Demographic Models », Journal of the Royal Statistical Society. Series B (Methodological), vol. 33, n°2, pp. 275−289.

Hothorn, T., Lausen, B., Benner, A., Radespiel-Tröger, M., (2004). “Bagging survival trees”, Stat. Med. 23 (1) 77–91.

Hothorn, T., Hornik, K., Zeileis, A., (2006) « Unbiased Recursive Partitioning: A Conditional Inference Framework », Journal of Computational and Graphical Statistics, vol. 15, no 3, 2006, p. 651–674 (DOI 10.1198/106186006X133933, JSTOR 27594202)

Ibrahim, N.A., Abdul Kudus, A., Daud, I., Abu Bakar, M.R., (2008). “Decision tree for competing risks survival probability in breast cancer study”, Int. J. Biol. Med. Sci. 3 (1) 2008, 25–29.

Ishwaran, H., Kogalur, U.B., Blackstone, E.H., Lauer, M.S., (2008). “Random survival forests”, Ann. Appl. Stat. (2008) 841–860.

Jessen, A.H., Mikosch, T., Samorodnitsky, G., (2011). “Prediction of outstanding payments in a Poisson cluster model”. Scandinavian Actuarial Journal 2011/3, 214-237.

Jewell, W.S., (1989). “Predicting IBNYR events and delays I. Continuous time”. ASTIN Bulletin 1989, 19(1), 25-55.

Kalbfleisch, J.D., Prentice, R.L., (2002). “The Statistical Analysis of Failure Time Data”, Wiley, Hoboken, NJ, 2002.

Kaplan E.L., Meier P., (1958), « Nonparametric Estimation from Incomplete Observations », Journal of the American Statistical Association, vol. 53, n°282, pp. 457−481.

Kattan, M.W., Hess, K.R., Beck, J.R., (1998). “Experiments to determine whether recursive partitioning (CART) or an artificial neural network overcomes theoretical limitations of Cox proportional hazards regression”, Comput. Biomed. Res. 31 (5) (1998) 363–373.

Khan, F.M., Zubek, V.B., (2008). “Support vector regression for censored data (SVRc): a novel tool for survival analysis”, in: Eighth IEEE International Conference on Data Mining (ICDM 2008), IEEE, 2008, pp. 863–868.

Lemberger, P., Batty, M., Morel, M., Raffaëlli, JL., (2015). “Big Data et Machine Learning”, Dunod, 2015, pp 130-131.

Lin, Y.-K., Chen, H., Brown, R.A., Li, S.-H., Yang, H.-J., (2014). “Predictive analytics for chronic care: a time-to-event modeling framework using electronic health records”, Available at SSRN 2444025.

Lucas, P.J.F., van der Gaag, L.C., Abu-Hanna, A., (2004). “Bayesian networks in biomedicine and healthcare, Artif. Intell”. Med. 30 (3) 2004, 201–214.

Lopez, O. (2018). “A censored copula model for micro-level claim reserving”. HAL Id: hal-01706935.

Lopez, O., Milhaud, X., Thérond, P., (2016). “Tree-based censored regression with applications in insurance”, Electronic Journal of Statistics, (Oct. 2016) Volume 10 issue 2, pp.2685-2716.

Matheny, M., McPheeters, M.L., Glasser, A., Mercaldo, N., Weaver, R.B., Jerome, R.N., Walden, R., McKoy, J.N., Pritchett, J., Tsai, C., (2011). “Systematic review of cardiovascular disease risk assessment tools”, Tech. Rep., Agency for Healthcare Research and Quality (US), 2011.

Mack, T., (1993). “Distribution-free calculation of the standard error of chain ladder reserve estimates.” ASTIN Bulletin 23/2, 213-225.

Nisbet, R., Elder, J., Miner, G., (2009) “Handbook for Statistical Analysis And Data Mining”, Academic Press, Page 247 Edition 2009

Norberg, R. (1993). “Prediction of outstanding liabilities in non-life insurance”. ASTIN Bulletin 1993, 23(1), 95-115.

Pigeon, M., Antonio, K., Denuit, M. (2013). Individual loss reserving with the multivariate skew normal framework. ASTIN Bulletin 43/3, 399-428.

Ridker, P.M., Buring, Julie N., Rifai, E., Cook, N.R., (2007). “Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds risk score”, JAMA: J. Am. Med. Assoc. 297 (6) (2007) 611–619.

Ridker, P.M., Paynter, N.P., Rifai, N., Gaziano, J.M., Cook, N.R., (2008). “C-Reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds risk score for men”, Circulation 118 (18) (2008) S1145.

Ripley, B.D., Ripley, R.M., (2001). “Neural networks as statistical methods in survival analysis”, Clin. Appl. Artif. Neural Networks 2001, 237–255.

Rouviere, L., (2018). « Introduction aux méthodes d’agrégation : boosting, bagging et forêts aléatoires ». Illustrations avec R.

Shivaswamy, P.K., Chu, W., Jansche, M., (2007). “A support vector approach to censored targets”, in: Seventh IEEE International Conference on Data Mining, 2007. ICDM 2007, IEEE, 2007, pp. 655–660.

Sierra, B., Larranaga, P., (1998). “Predicting survival in malignant skill melanoma using Bayesian networks automatically induced by genetic algorithms: an empirical comparison between different approaches”, Artif. Intell. Med. 14 (1-2) 1998, 215 - 230.

Štajduhar, I., Dalbelo-Bašic´, B., Bogunovic´, N., (2009). “Impact of censoring on learning Bayesian networks in survival modelling”, Artif. Intell. Med. 47 (3) 2009, 199 - 217.

Štajduhar, I., Dalbelo-Bašic´, B., (2010). “Learning Bayesian networks from survival data using weighting censored instances”, J. Biomed. Inform. 43 (4) (2010) 613–622.

Stute, W. (1993). “Consistent estimation under random censorship when covariables are present”. Journal of Multivariate Analysis, 45(1), 89-103.

Stute, W. (1999). “Nonlinear censored regression”. Statistica Sinica, 1089-1102.

Strobl, C., Malley, J., Tutz, G., (2009). « An Introduction to Recursive Partitioning: Rationale, Application and Characteristics of Classification and Regression Trees, Bagging and Random Forests », Psychological Methods, vol. 14, no 4, 2009, p. 323–348 (DOI 10.1037/a0016973)

Tsiatis, A.A., (2006). “Semiparametric Theory and Missing Data”, Springer, New York, 2006.

Verrall, R.J., Wüthrich, M.V. (2016). “Understanding reporting delay in general insurance”. Risks 4/3, 25.

Vock, D. M., Wolfson, J., Bandyopadhyay, S., Adomavicius, G., Johnson, P. E., Vazquez-Benitez, G., & O’Connor, P. J. (2016).” Adapting machine learning techniques to censored time-to-event health record data: A general-purpose approach using inverse probability of censoring weighting”. Journal of biomedical informatics, 61, 119-131.

Wu, J., Roy, J., Stewart, W.F., (2010). “Prediction modeling using EHR data: challenges, strategies, and a comparison of machine learning approaches”, Med. Care 48 (6) 2010, S106–S113.

Wüthrich (2018), Neural Networks Applied to Chain-Ladder Reserving, European Actuarial Journal, December 2018, Volume 8, Issue 2, pp 407–436

Zarkadoulas, A., (2017). “Neural network algorithms for the development of individual losses”. MSc thesis, University of Lausanne.

Chapitre 2

ACP (2012). La revue de l’Autorité de contrôle prudentiel, octobre-novembre 2012

Aguir, N., (2012). Impact de la modélisation multi-états sur le SCR en prévoyance, Mémoire de fin d’étude présenté devant l’Institut de Science Financière et d’Assurances pour l’obtention du diplôme d’Actuaire de l’Université de Lyon le 26 Septembre 2012

Bailey, A. L., (1945). «A generalized theory of credibility», Proceedings of the Casualty Actuarial Society, vol. 32, p. 13–20.

Bailey, A. L., (1950). «Credibility procedures, Laplace’s generalization of Bayes’ rule and the combination of collateral knowledge with observed data», Proceedings of the Casualty Actuarial Society, vol. 37, p.7–23.

Bourdonnais, R., Terraza, M., (2010). Analyse des séries temporelles, 3ème édition Dunod

Box, G.E.P., Jenkins, G.M., (2015). Time Series Analysis: forecasting and control, Ed. Holden-Day

Brockwell, P.J., Davis, R.A., (1987). Time Series: Theory and Method, Ed. Springer-Verlag

Brockwell, P. J., DAVIS, R.A., (2003). Introduction to Time Series and Forecasting, 2nd Ed. Springer

Bühlmann, H., (1967). «Experience rating and credibility», ASTIN Bulletin, vol. 4, p. 199–207

Bühlmann, H., (1969). «Experience rating and credibility», ASTIN Bulletin,vol. 5, p. 157–165

Cambon, A., (2011). « Elaboration d’un modèle interne partiel concernant le risque de souscription non-vie pour tenir compte des spécificités d’une société spécialisée dans les branches longues », Mémoire présenté devant l'Institut de Statistique de l'Université Pierre et Marie Curie pour l'obtention du diplôme de statisticien mention actuariat

Canadian Institute of Actuaries, (2002). Commission des rapports financiers des compagnies d’assurances vie, Note éducative, mortalité prévue : polices canadiennes d’assurances vie individuelle avec tarification complète

Canadian Institute of Actuaries, (2014), Commission des rapports financiers des régimes de retraite, Note éducative révisée : sélection des hypothèses de mortalité aux fins des évaluations actuarielles des régimes de retraite

Cerchiara, R. R., Magatti, V., (2014). “Undertaking Specific Parameters or a Partial Internal Model under Solvency II?”, 30th International Congress of Actuaries, 30 March to 4 April 2014, Washington, D.C., USA

Ostrow, C.W. Jr., (1982). « Time Series Analysis : Regression Techniques. Serie : Quantitative Applications in the Social Sciences », Ed. Sage Publications

De Vylder, F., (1981). « Practical credibility theory with emphasis on parameter estimation », ASTIN Bulletin, vol. 12, p. 115–131

Doitteau, M., (2011). « La réglementation et la modélisation stochastique de l'incapacité », Mémoire de fin d'étude présenté devant l'Institut de Science Financière et d'Assurances pour l'obtention du diplôme d'Actuaire de l'Université de Lyon le 4 janvier 2011

EIOPA (2011). "Draft proposal for Implementing Technical Standard on Undertaking Specific Parameters: Methods", EIOPA-FinReq-11/023

EIOPA (2013). "Technical Specification on the Long Term Guarantee Assessment (Part 1)"

Goel, P. K., (1982). «On implications of credible means being exact bayesian», Scandinavian Actuarial Journal, p.41–46.

Goovaerts, M. J., Hoogstad, W. J., (1987). "Credibility Theory, no 4, Surveys of actuarial studies", Nationale-Nederlanden N.V., Netherlands

Goulet, V., (2010). « ACT 2008 Mathématiques actuarielles IARD II (Théorie de la crédibilité) », école d'actuariat université Laval

Guibert, Q., Juillard, M., Teuguia, O. N., Planchet, F., (2014). « Solvabilité Prospective en Assurance Méthodes quantitatives pour l'ORSA », Ed. Economica

Hachemeister, C. A., (1975). «Credibility for regression models with application to trend», dans Credibility, theory and applications, Proceedings of the Berkeley actuarial research conference on credibility, Academic Press, New York, p. 129–163

Jewell, W. S., (1974). «Credible means are exact bayesian for exponential families», ASTIN Bulletin, vol. 8, p.77–90

Mack, T., (1993). "Distribution-free calculation of the standard error of chain ladder reserve estimates", Astin Bulletin 23 (2), 213-225

Merz, M., Wüthrich, M., (2008). "Modelling the claims development result for solvency purposes", CAS E-Forum, Fall 2008, 542-568

Norberg, R., (1979). «The credibility approach to ratemaking», Scandinavian Actuarial Journal, vol. 1979, p. 181–221

Olympio, A., Ranaivozanany, V., Wittmer, S., (2012). « Gestion des risques d'entreprise : qualité des données, levier de pilotage stratégique », article congrès ERM AFIR Lyon juin 2012

Perrin, M., (2012). « Calibration des Undertaking Specific Parameters et leurs impacts sur les fonds propres », Mémoire présenté devant le jury de l'EURIA en vue de l'obtention du diplôme d'Actuaire EURIA et de l'admission à l'Institut des Actuaires le 28 septembre 2012

Planchet, F., Therond, P., (2011), « Modèles de durée, Applications actuarielles », Ed. Economica

Planchet, F., Leroy, G., Juillard, M., (2012). « Modèle standard, quelle utilisation des paramètres spécifiques à l'entité », La Tribune de l'assurance, avril 2012, n°168

Planchet, F., « Segmentation tarifaire et suivi du risque en incapacité », Journées d'étude IA / SACEI du 17 septembre 2009

Planchet, F., Thérond, P., Kamega, A., (2009), « Scénarios économiques en assurance modélisation et simulation », Ed. Economica

Planchet, F., Guibert, Q., Juillard, M., (2010), « Un cadre de référence pour un modèle interne partiel en assurance de personnes : application à un contrat de rentes viagères », Bulletin Français d'Actuariat, Institut des Actuaires, 2010, 10 (20), pp. 5-34. <hal-00530864>

Sanders, D.H., A.F. Murph, A.F., ENG, R.J., (1984). « Les statistiques : une approche nouvelle ». Ed. MacGraw Hill

Whitney, A. W., (1918). «The theory of experience rating», Proceedings of the Casualty Actuarial Society, vol. 4, p. 275–293.

Chapitre 3

CEIOPS, (2009). “Consultation Paper no.51-Draft Level 2 Advice on SCR Standard Formula-Counterparty Default Risk”

Courbage, C., Roudaut, N., (2008). “Empirical Evidence on Long-term Care Insurance Purchase in France”, The Geneva Papers on Risk and Insurance - Issues and Practice, October 2008, Volume 33, Issue 4, pp 645–658.

Croix, J.C., Planchet, F., Thérond, P.E., (2015). « Mortality: a statistical approach to detect model misspecification », Bulletin Français d’Actuariat, vol. 15, n°29.

DREES, (2018). « Étude 1046 - Les Français vivent plus longtemps, mais leur espérance de vie en bonne santé reste stable » (janv.2018)

EIOPA, (2015). « Règlement délégué n°2015-35 complétant la directive 2009/138/CE du Parlement européen et du Conseil sur l'accès aux activités de l'assurance et de la réassurance et leur exercice (solvabilité II) ».

Guibert, Q., Juillard, M., Planchet, F., (2010). « Un cadre de référence pour un modèle interne partiel en assurance de personnes », Bulletin Français d’Actuariat, vol.10,n°20.

Lusson, F., (2013). « L’équilibre actuariel de long terme en assurance dépendance en France », Revue d’Analyse Financière, n°47.

OCIRP, (2017). “Baromètre OCIRP Autonomie 2017”

Planchet, F., Tomas, J., (2014a). « Uncertainty on Survival Probabilities and Solvency Capital Requirement: Application to LTC Insurance », Scandinavian Actuarial Journal, doi: 10.1080/03461238.2014.925496.

Planchet, F., Tomas, J., (2013). « Multidimensional smoothing by adaptive local kernel-weighted log-likelihood with application to long-term care insurance », Insurance : Mathematics and Economics, Vol. 52, pp. 573–589. http://dx.doi.org/10.1016/j.insmatheco.2013.03.009

Sator, N., Sother, G., (2013). « Approche Solvabilité II et ERM du risque Dépendance », Proceedings of the AFIR Colloquium

Chapitre 4

Adair, J. G., (1984). “The Hawthorne effect: A reconsideration of the methodological artifact”. Journal of Applied Psychology, 69(2), 334–345. doi :10.1037/0021-9010.69.2.334

Allais, M., (1953). « Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école américaine », Econometrica, 4, 1953, p. 503-546.

Allport, G. W. (1935). « Attitudes ». In C. Murchison (Ed.), Handbook of social psychology (pp.798–844). Worcester, MA : Clark University Press.

Bickman, L. (1972). “Environmental attitudes and actions”. The Journal of social psychology, 87(2), 323-324.

Blum, V., Thérond, P-E., Alexander, D., Laffort, E., Jancevska, S., (2019). "New developments in language issues in accounting regulation: likelihood terms and the certainty of uncertainty," Working Papers hal-01991845, HAL.

Boholm, A. (1998). Comparative studies of risk perception: a review of twenty years of research. Journal of risk research, 1(2), 135-163.

Bru, B., Bru, M-F., Lai Chung, K., (1999). « Borel et la martingale de Saint-Pétersbourg », Revue d'histoire des mathématiques, 1999, p. 181-247

Bloch, H. et al. (1997). « Dictionnaire fondamental de la psychologie ». Paris : Larousse Bordas.

Campion, B., et al., (2013). « Apports et limites de l'expérimentation comme moyen d'investigation des effets éducatifs des médias », ESSACHESS. Journal for Communication Studies, vol. 6, no. 1(11) / 2013 : 257-267, eISSN 1775-352X

Cohen, J. (1988). "Statistical power analysis for the behavioral sciences". Second Edition. Hillsdale, N.J, L. Erlbaum Associates.

Dake, K., (1991), « Orienting Dispositions in the Perception of Risk: An Analysis of contemporary Worldviews and Cultural Biases »

Delory, C., (2003). « Guide pratique de la recherche en sciences humaines ». Namur: Erasme.

Douglas, M., Wildavsky, A., (1983). "Risk and culture, an essay on the selection of technical and environmental dangers", Berkeley, University of California Press.

Douglas, M., (2001). « De la souillure : essai sur les notions de pollution et de tabou ». La découverte: Paris, p. 200.

Eagly, A. H., & Chaiken, S., (1993). "The psychology of attitudes". Orlando, FL, US: Harcourt Brace Jovanovich College Publishers.

Ebale Moneze, C., (2001). « Le développement théorique de la psychologie sociale ». Yaoundé : Presses Universitaires de Yaoundé.

Ellsberg D., (1961). "Risk, ambiguity and the Savage axioms", The Quarterly Journal of Economics, 75, 643-669, 1961

Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B., (1978). "How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits". Policy sciences, 9(2), 127-152.

Fishburn, Peter C. (1988). "Nonlinear Preference and Utility Theory". Baltimore, Md.: Johns Hopkins University Press.

Hofstede, G., (2001). "Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations", 2nd ed. Thousand Oaks, CA: Sage.

Hofstede, G., & Minkov, M. (2010). "Long-versus short-term orientation: new perspectives". Asia Pacific Business Review, 16(4), 493-504.

Hsu, S.-Y., Woodside, A. G. & Marshall, R. (2013). "Critical Tests of Multiple Theories of Cultures' Consequences: Comparing the Usefulness of Models by Hofstede, Inglehart and Baker, Schwartz, Steenkamp, as well as GDP and Distance for Explaining Overseas Tourism Behavior". Journal of Travel Research, 52 (6), 679-704

Inglehart, R., (1997). "Modernization and postmodernization: Cultural", economic and political change in 43 societies. Princeton, NJ: Princeton University Press.

Inglehart, R., Basañez, M., & Moreno, A., (1998). "Human Values and Beliefs: A Cross-Cultural", Sourcebook, Ann Arbor.

Ingram D., Thompson E., (2010). « A Universe in Four Parts ». Wilmott Magazine, March 2010

Ingram D., Thompson E., Tayler, (2010). « Eyes Wide Open: Towards Rational Adaptability ». Wilmott Magazine, July 2010

Ingram D., (2009) « Four Seasons of Risk Management ». Actuary Magazine, December 2009

Ingram D., Underwood A., (2010) « Full Spectrum of Risk Attitude ». Actuary Magazine, August 2010

Ingram D., Underwood A., (2011) « The Human Dynamics of the Insurance Cycle and Implications for Insurers: An Introduction to the Theory of Plural Rationalities ». ERM Symposium/SOA Monograph, January 2010

Ingram D., Thompson E., (2011) « Changing Seasons of Risk Attitudes ». Actuary Magazine, April 2011

Ingram D., Thompson E., (2012) « What's Your Risk Attitude? ». Harvard Business Review Blog, June 2012

Kahneman D., Tversky A., "Prospect theory: an analysis of decision under risk", Econometrica, 47, 263-91, 1979

LaPiere, R. T. (1934). Attitude versus action. Social Forces, 13, 230-237

Lebart, L., Piron, M., & Morineau, A., (2006). « Statistique exploratoire multidimensionnelle : visualisation et inférences en fouilles de données ».

Lemaine, G, & Lemaine, J.-M., (1969). « Psychologie sociale et expérimentation ». France : Mouton et Bordas.

Llosa, S., (1997). « L'analyse de la contribution des éléments du service à la satisfaction : un modèle tétraclasse ». Décisions marketing, 81-88.

Mark J. Machina, (1982) "Expected Utility Analysis without the Independence Axiom", Econometrica, Vol. 50, No. 2. (Mar., 1982), pp. 277-323

Matalon, B., (1995). La logique des plans d'expérience. In J.-F. Le Ny & M.-D. Gineste (Eds.), La psychologie : textes essentiels (pp. 48-60). Paris : Larousse. (Publication originale : (1969). In G. Lemaine & J.-M. Lemaine (Eds) Psychologie sociale et expérimentation. Paris : Mouton/Bordas).

Mischel, W., (2013). Personality and assessment. Psychology Press.

Morineau, A., (1984). Note sur la caractérisation statistique d'une classe et les valeurs-tests. Bulletin Technique Centre Statistique Informatique Appliquées, 2(1-2), 20-27.

Morgenstern, O., Von Neumann, J., (1994). "Theory of Games and Economic Behavior", PUP, 1944, 1re éd.

Rajecki, D. W., (1990). « Attitudes ». Sinauer Associates.

Rohrmann, B., (2000). "Cross-cultural studies on the perception and evaluation of hazards". In O. Renn & B. Rohrmann (Eds.), "Cross-cultural risk perception: A survey of empirical studies" (pp. 103-144). Dordrecht: Kluwer Academic Publishers.

Savage L.J., (1954). "The Foundations of Statistics", New York: Wiley, 1954

Schwartz, S. H., (2004). "Mapping and interpreting cultural differences around the world". In H. Vinken, J. Soeters, & P. Ester (Eds.), Comparing cultures, dimensions of culture in a comparative perspective. Leiden, the Netherlands: Brill.

Schwartz, S. H., (2006). "A theory of cultural value orientations: Explication and applications". Comparative Sociology, 5, 137–182.

Schwing and W. A. Albers, Jr. (Eds.), (2001). "Societal risk assessment: How safe is safe enough?" (pp. 181-214). New York: Plenum Press. Reprinted in P.Slovic (Ed.), The perception of risk. London: Earthscan, 2001.

Slovic, P., Fischhoff, B. & Lichtenstein, S., (1980). "Facts and fears: Understanding perceived risk. In R".

Slovic, P., (1987). "Perception of Risk". Science, 236 (4799), 280-285.

Slovic, P., Flynn, J., Mertz, C.K., Poumadère, M., & Mays., C., (2000). "Nuclear Power and the Public: A Comparative Study of Risk Perception in France and the United States". In O. Renn and R. Rohrmann (eds.). Cross-Cultural Risk Perception: A Survey of Empirical Studies. Boston, MA : Kluwer Academic Publishers.

Thomas, R., Alaphilippe, D. (1993) : Les attitudes. Paris.

Tremblay, P., Beauregard, B. [2006], « Application du modèle tétraclasse aux résultats de sondage d'un organisme public : le cas de la régie des rentes du Québec ».

Valade, Bernard, (2007). « Jean Stoetzel : théorie des opinions et psychosociologie de la communication », Hermès, La Revue, vol. 48, no. 2, 2007, pp. 72-74.

Wicker, A. W., (1969). "Attitudes versus actions: The relationship of verbal and overt behavioural responses to attitude objects". Journal of Social Issues, 25, 41–78.

Wildavsky, A., Dake, K., (1990). "Theories of Risk Perception: Who Fears What and Why?". Daedalus, vol. 119(4), p. 41-60.