

**Mémoire présenté le :**

**pour l'obtention du Diplôme Universitaire d'actuariat de l'ISFA  
et l'admission à l'Institut des Actuaires**

Par : Paul DEWAELE

Titre Tarification d'un contrat d'assurance emprunteur et modélisation de sa rentabilité par la  
génération d'un portefeuille fictif d'assurés

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

*Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus*

*Membre présents du jury de l'Institut des Actuaires*

*Membres présents du jury de l'ISFA*

*Nom :*

*Signature :*

*Directeur de mémoire en entreprise :*

*Nom :*

*Signature :*

*Invité :*

*Nom :*

*Signature :*

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

***Signature du responsable entreprise***

***Signature du candidat***

Mots clefs : assurance emprunteur, tarification, simulation, rentabilité, portefeuille fictif

## **Résumé**

Les évolutions du marché de l'assurance emprunteur nous amènent à nous interroger sur les processus de tarification qui étaient utilisés jusqu'à présent. Ceux-ci étaient coûteux en temps et destinés à la conception d'un produit ciblant une population aux caractéristiques générales. Le marché semble désormais pousser à la conception de produits spécialisés sur certaines populations spécifiques.

Afin de résoudre cette problématique, le présent mémoire présente une méthodologie de tarification se voulant rapide, adaptable à des populations de caractéristiques différentes. Néanmoins, cette méthode implique un processus itératif, dans le sens où les hypothèses utilisées dans celle-ci ne peuvent être calibrées de manière plus précise que dans un second temps. En effet, ces hypothèses concernent notamment la répartition démographique de la cible d'assurés, qui dépendra notamment de la compétitivité du tarif et de sa rentabilité sur les principaux segments de population visés.

Afin d'analyser la rentabilité du tarif, nous nous sommes intéressés à deux indicateurs qui sont la rentabilité technique et la marge brute d'ASP, afin d'intégrer la double vision ASP (courtier grossiste) et porteur de risque. Pour réaliser cette analyse, les outils à disposition n'étaient pas suffisants. En effet, nous sommes dans une situation particulière où les données à disposition sont partielles et dans le cas où elles sont présentes, elles sont globales. C'est pour cette raison que nous présentons une méthode de simulation d'un portefeuille fictif d'assurés, qui se base sur lois marginales des différentes variables qui représentent un portefeuille d'assurés et sur les dépendances entre celles-ci.

Enfin, nous réalisons une application pratique, grâce à certaines données démographiques externes au portefeuille ASP. Pour cela, nous présentons tout d'abord de manière précise le modèle de projection de la rentabilité utilisé, avant d'effectuer certains scénarios sur la granularité des modalités de différenciation tarifaire comme la segmentation Fumeur / Non-fumeur ou Présence d'un co-emprunteur / Pas de co-emprunteur.

Grâce aux tests présentés dans notre dernière partie, nous montrons de quelle manière il est possible d'optimiser la structure tarifaire suivant différents horizons de temps considérés et nous validons également l'ensemble de la méthode de tarification présentée précédemment.

Key words: loan insurance, pricing, simulation, profitability, fictive portfolio

## **Summary**

Developments in the loan insurance market question the pricing processes that have been used so far. These were until now expensive in time and intended for the design of a product targeting a population with general characteristics. The market now seems to be pushing for the design of specialized products for certain specific populations.

In order to solve this problem, this thesis presents a pricing methodology that is intended to be fast and adaptable to populations of different characteristics. Nevertheless, this method involves an iterative process, in the sense that the hypotheses used in it have to be calibrated more precisely in a second time. In fact, these assumptions concern in particular the demographic distribution of the intended insured target, which will depend in particular on the competitiveness of the pricing and its profitability on the main target population segments.

In order to analyze the profitability of the pricing, we are interested in two indicators: the technical profitability and the gross margin of ASP. This ensures that we have the dual vision ASP and risk-bearing partner that characterizes the specific role of ASP as a wholesale broker. To carry out this analysis, the available tools were not sufficient. Indeed, we are in a particular situation where the available data are partial and in the case where they are present, they are global. It is for this reason that we present a method of simulation of a fictitious portfolio of policyholders, which are based on marginal laws of the various variables that represent a portfolio of policyholders and the dependencies between them.

Finally, we realize a practical application, thanks to some demographic data external to the ASP portfolio. To do this, we first present in a precise way the profitability projection model used, before making some scenarios on the granularity of tariff differentiation methods such as Smoker / Non-smoker segmentation or Presence of a co-borrower / no co-borrower.

Thanks to the tests presented in our last part, we show how it is possible to optimize the pricing structure according to different time horizons and we also validate the entire pricing method presented above.

## **Remerciements**

Je tiens tout d'abord à remercier Sophie D'HONDT, responsable du service Actuariat Emprunteur d'APRIL Santé Prévoyance, pour son aide lors de la réalisation de ce mémoire ainsi que de sa confiance dans la réalisation des différents travaux menés au cours de mon année d'alternance.

Je remercie également Eugénie POYET pour son accompagnement, son expertise et ses conseils dans la réalisation de cette étude. Quelques mots ne suffiraient pas à la remercier pour tout ce qu'elle m'a apporté à un niveau professionnel, mais également dans sa manière toujours très pédagogique de le faire.

Enfin, je remercie toutes les autres personnes du service Conception Emprunteur, aussi bien en Actuariat qu'en Marketing, auprès de qui j'ai énormément appris et qui ont toujours rendu le travail à leur côté très inspirant.

Enfin, j'adresse mes remerciements à mon tuteur académique Pierre THEROND, pour son accompagnement et ses conseils très avisés tout au long de la réalisation de ce mémoire.

# Sommaire

Introduction générale .....	8
<b>1. Contexte.....</b>	<b>10</b>
<b>1.1. L'assurance emprunteur en France .....</b>	<b>10</b>
1.1.1. Objet de l'assurance emprunteur en France.....	10
1.1.2. Les garanties disponibles .....	10
1.1.3. Souscription d'un contrat d'assurance de prêts .....	13
1.1.4. Actualités règlementaires et mutations du marché.....	15
<b>1.2. APRIL et l'assurance emprunteur .....</b>	<b>16</b>
1.2.1. Présentation générale d'APRIL Santé Prévoyance .....	16
1.2.2. L'assurance emprunteur au sein d'ASP .....	17
<b>1.3. Tarification d'un contrat d'assurance emprunteur .....</b>	<b>17</b>
1.3.1. Méthodologie générale .....	18
1.3.2. Facteurs explicatifs de la sinistralité .....	18
1.3.3. Construction des taux de primes pour les garanties DC/PTIA et ITT/IPT.....	21
<b>2. Construction d'un outil de tarification automatisé.....</b>	<b>26</b>
<b>2.1. Présentation du processus de tarification envisagé .....</b>	<b>26</b>
2.1.1. Rappel des éléments disponibles .....	26
2.1.2. Problématiques liées à la structure des primes ADP APRIL 2017 .....	27
2.1.3. Principe de déformation des primes.....	27
2.1.4. Calcul de la prime commerciale.....	29
<b>2.2. Description de l'outil et souplesse de la structure tarifaire .....</b>	<b>31</b>
2.2.1. Fonctionnement général de l'outil .....	31
2.2.2. Présentation des Inputs .....	32

2.2.3. Définition des Outputs .....	36
2.3. Validation de la tarification.....	36
3. Construction d'un générateur de portefeuille fictif .....	39
3.1. Les limites de l'utilisation des model points .....	39
3.2. Traitement des bases de données .....	41
3.3. Analyse des caractéristiques du portefeuille actuel .....	45
3.3.1. Variables caractérisant les assurés.....	46
3.3.2. Variable caractérisant les prêts.....	47
3.3.3. Variables concernant les garanties de l'assurance de prêt.....	49
3.4. Statistiques bidimensionnelles et indicateurs de corrélations .....	50
3.5. Présentation de méthodes de simulation de portefeuille fictif d'assurés adaptées à notre problématique.....	55
3.5.1. Résumé théorique des différentes mesures de la dépendance entre variables de différentes structures.....	56
3.5.2. Méthode des distributions conditionnelles .....	58
3.5.3. Méthode retenue pour la simulation du portefeuille fictif d'assurés à partir de la spécification des marginales et de la structure de dépendance .....	59
4. Application : Analyse de la rentabilité à partir du portefeuille fictif.....	77
4.1. Outil de projection du compte de résultats .....	77
4.4.1. Rentabilité et compte de résultat .....	78
4.1.2. Choix des différents paramètres de projection .....	80
4.2. Utilisation du portefeuille fictif pour analyser les niveaux de marges techniques et brutes d'un tarif .....	83
4.3. Recherche de la structure tarifaire et de la structure des chargements optimales au regard de contraintes a priori .....	87
4.4. Comparaison de la méthode sur un portefeuille réel d'assurés.....	90

<b>Conclusion générale .....</b>	<b>93</b>
<b>Bibliographie.....</b>	<b>94</b>
<b>Annexes.....</b>	<b>96</b>
Annexe I : Formule de calcul des VAP assureur Décès / PTIA.....	96
Annexe II : Formule de calcul des VAP assureur Arrêt de travail.....	98
<b>Table des figures.....</b>	<b>100</b>
<b>Table des tableaux .....</b>	<b>100</b>

## **Introduction générale**

Lors de la souscription d'un prêt, l'assurance emprunteur est un contrat permettant de couvrir les capitaux sous risque en cas de décès, d'arrêt de travail (IPT/ITT) et parfois également en cas de perte d'emploi. Historiquement, l'assurance emprunteur est un marché très largement dominé par les réseaux bancaires, qui leur procure en plus de cela une rentabilité régulière et pérenne. Toutefois, depuis les récentes réglementations, et notamment la mise en application de la résiliation annuelle par l'amendement Bourquin, de nouveaux acteurs ont fait leur apparition sur ce marché. Le développement des comparateurs d'assureurs permet aux consommateurs de faire jouer pleinement la concurrence alors que la baisse prolongée des taux de crédits immobiliers a encouragé les ménages à emprunter ou renégocier leur prêt en cours. Le marché de l'assurance emprunteur est donc plus dynamique que jamais et il est de plus en plus nécessaire de sortir des produits plus rapidement pour répondre à ces évolutions.

Dans ce contexte et dans le cadre de sa stratégie de développement, l'année 2017 a été marquée chez ASP par le besoin de répondre à différents appels d'offres pour la conception de produits d'assurance emprunteur en gestion pour compte de tiers et par la nécessité de construire des produits visant à cibler plus spécifiquement certaines cibles secondaires des gammes en cours de commercialisation.

Ces appels d'offres et nouvelles gammes nécessitent une construction tarifaire plus rapide qu'usuellement avec un besoin de format de cotisation cible pouvant être différent du produit phare d'APRIL (moins de segmentation notamment) et une structure de population assurée attendue sur de tels produits potentiellement différente de celle observée dans le portefeuille stock issu des produits historiques d'April Santé Prévoyance. Par ailleurs, les informations disponibles sur la cible finale pour de tels appels d'offre ou pré-étude d'opportunité sont généralement peu détaillées (âge moyen, composition en catégories socio-professionnelles, etc.)

Suite à la refonte de sa gamme en 2016, ASP dispose de bases techniques de sinistralité d'expérience récentes, mais la tarification complète d'un contrat d'assurance emprunteur est complexe, avec de multiples critères de segmentation technique du risque non nécessairement retenus dans la structure finale du tarif commercial.

Le but de ce mémoire est donc de parvenir à un outil, permettant d'accélérer le processus de tarification en repartant des bases techniques disponibles et d'apprécier rapidement la rentabilité technique induite par les tarifs commerciaux retenus.

Nous détaillerons dans un premier temps le contexte ce qui permettra de mieux appréhender les différents détails de la problématique à résoudre. Puis, nous présenterons une méthodologie de tarification d'un contrat d'assurance emprunteur. Dans une troisième partie, nous verrons comment adapter cette méthodologie à notre problématique de gain de vitesse dans la tarification. Enfin, nous verrons comment nous pouvons modéliser la rentabilité de ce produit d'assurance par la génération d'un portefeuille d'assurés fictif.

## **1. Contexte**

Nous présentons ici April Santé Prévoyance ainsi que le marché de l'assurance emprunteur, afin de permettre une meilleure compréhension des différentes problématiques auxquelles répond ce mémoire.

### **1.1. L'assurance emprunteur en France**

#### **1.1.1. Objet de l'assurance emprunteur en France**

L'assurance de prêt est une assurance permettant aux banques de sécuriser le remboursement d'un emprunt en cas de décès ou d'arrêt de travail, qui est habituellement exigée par l'organisme prêteur lors de l'octroi d'un crédit. C'est une assurance temporaire car elle est limitée à la durée d'un prêt. Les principales garanties sont celles couvrant le décès, la perte totale et irréversible d'autonomie, l'arrêt de travail et la perte d'emploi.

En France, l'assurance emprunteur n'est pas obligatoire légalement. Néanmoins, elle est exigée par la banque lors de l'octroi d'un emprunt. Les garanties décès et PTIA sont notamment obligatoires dans tous les cas, alors que les garanties couvrant l'IPT et l'ITT sont demandées seulement pour un achat résidentiel. D'autres garanties comme l'IPP ou la perte d'emploi sont généralement facultatives.

#### **1.1.2. Les garanties disponibles**

L'assurance emprunteur est très limitée dans l'innovation au niveau des garanties. En effet, il est nécessaire de proposer des garanties au moins équivalentes aux critères proposés par le C.C.S.F. (cf. 1.3.3. Avis du Comité Consultatif du Secteur Financier (C.C.S.F.) sur l'équivalence du niveau de garantie en assurance emprunteur).

Nous présentons maintenant les garanties proposées habituellement dans un contrat d'assurance individuel de prêt. Les définitions peuvent varier légèrement entre les différents acteurs du marché, notamment les dates de fin de couverture, la limite des capitaux assurés, etc. Ces définitions peuvent notamment être retrouvées en détail dans les Conditions Générales (CG) de n'importe quel produit d'assurance emprunteur d'APRIL Santé Prévoyance.

## **Décès et Perte Totale Irréversible d'Autonomie (P.T.I.A.)**

En cas de décès de l'assuré et sous réserve des exclusions prévues par le contrat d'assurance, les bénéficiaires désignés reçoivent le Capital Restant Dû (CRD) au jour du décès, renseigné dans le Tableau d'Amortissement (TA) fourni par l'organisme prêteur, et dans la limite du montant garanti, qui correspond à la multiplication de la quotité garantie et du CRD.

La Perte Totale et Irréversible d'Autonomie (PTIA) consécutive à une maladie garantie ou à un accident garanti est assimilé à un décès. En cas de reconnaissance de l'état de PTIA d'un assuré, les bénéficiaires reçoivent le CRD, renseigné dans le tableau d'amortissement et dans la limite du montant garanti. Le capital restant dû est déterminé par référence au jour où la date de consolidation de la PTIA est fixée par expertise médicale.

Le paiement du capital met fin à l'assurance.

## **Arrêt de travail**

### **En cas d'Incapacité Temporaire Totale (I.T.T.) :**

Il est versé après expiration de la franchise souscrite à l'adhésion (i.e. à compter du 31<sup>ème</sup>, du 61<sup>ème</sup>, du 91<sup>ème</sup> ou du 181<sup>ème</sup> jour d'Incapacité Temporaire Totale), les arrérages de remboursements ou de loyer venant à l'échéance tels qu'ils sont mentionnés au tableau d'amortissement remis par l'organisme prêteur dans la limite du montant garanti.

Pour donner droit aux versements, l'incapacité doit entraîner une interruption réelle et complète des activités professionnelles de l'assuré ou des occupations de la vie quotidienne pour l'assuré n'exerçant pas de manière effective d'activité professionnelle au jour du sinistre.

En cas de mi-temps thérapeutique (i.e. reprise du travail à mi-temps pour raisons médicales), il sera versé pendant six mois maximum les arrérages de remboursement ou de loyer venant à échéance à hauteur de 50% du montant garanti sous réserve que l'assuré ait été indemnisé pendant au moins deux mois au titre de l'Incapacité Temporaire Totale.

### **En cas d'Invalidité Permanente Totale (I.P.T.) :**

La base de calcul des prestations est identique à celle de l'I.T.T. au jour du sinistre.

Pour qu'il y ait Invalidité Permanente Totale, l'assuré qui exerce de manière effective une activité professionnelle au jour de l'arrêt initial de son travail constaté médicalement doit

présenter une invalidité fonctionnelle physique ou mentale et conjointement une invalidité professionnelle.

Le taux d'invalidité permanente doit être supérieur ou égal à 66%.

### **Garantie Invalidité Spéciale Professions Médicales**

En complément de la garantie Décès/PTIA, cette garantie peut être souscrite uniquement pour les assurés exerçant une profession médicale (médecin, chirurgien, pharmacien, vétérinaire, ...). Consécutivement à une maladie garantie ou à un accident garanti, elle est assimilée au décès.

### **Garantie Invalidité Spécifique (G.I.S)**

Cette garantie est prévue par la convention AERAS (s'Assurer et Emprunter avec un Risque Aggravé de Santé). Elle ne peut être souscrite qu'en complément de la garantie Décès/PTIA.

Dans le cas d'un refus de la garantie IPT pour raisons médicales, elle est proposée sous réserve d'acceptation médicale.

En cas d'invalidité spécifique, il est versé les arrérages de remboursement ou de loyer tels qu'indiqués dans le tableau d'amortissement remis par l'organisme prêteur et dans la limite du montant garanti.

Une invalidité spécifique se caractérise par une incapacité professionnelle jointe à une incapacité fonctionnelle à au moins 70%.

### **Perte d'emploi**

En cas de licenciement par son employeur, il sera versé à l'organisme prêteur tout ou partie des mensualités de remboursement. Le montant de l'indemnisation est plafonné et sa durée est également limitée.

### **Option de rachats des exclusions de pathologies psychologiques et/ou dorsales**

L'Arrêt de travail lié à des pathologies psychologiques ou dorsales est assuré dans le contrat d'assurance emprunteur sous des conditions d'hospitalisation ou d'intervention chirurgicale

(30j. d'hospitalisation continue pour les pathologies psychologiques / intervention chirurgicale nécessitant une hospitalisation de 15j. continue pour les affections dorsales). Une option permet de racheter l'une des deux exclusions ou les deux exclusions, c'est-à-dire sans condition d'hospitalisation ou d'intervention chirurgicale.

### **Invalidité Permanente Partielle (I.P.P.)**

En complément de la garantie Arrêt de travail, la garantie Invalidité Permanente Partielle peut être souscrite par l'assuré. Si à la suite d'un accident ou d'une maladie garantie, l'assuré est reconnu en état d'I.P.P., il sera versé à hauteur de 50% les arrérages de remboursement venant à échéance tels qu'ils sont indiqués dans le tableau d'amortissement fourni par l'organisme prêteur.

Le taux d'invalidité permanente doit être compris entre 33% et 66%.

### **1.1.3. Souscription d'un contrat d'assurance de prêts**

Nous présentons ici les différentes caractéristiques de la souscription dans le cas d'un contrat d'assurance de prêts : les moyens de distribution possibles et la sélection des risques préalables à la souscription.

- **La distribution des contrats**

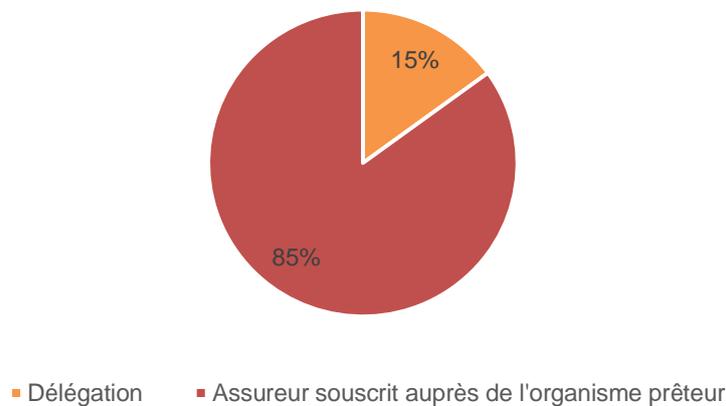
Le marché de l'assurance emprunteur est ultra-dominé par les bancassureurs qui sont porteurs d'offres de groupe et de certaines offres individuelles. De plus en plus d'acteurs alternatifs basés sur les offres numériques se développent.

Les acteurs indépendants de la délégation d'assurance peuvent être classés en 3 groupes distincts :

- Les courtiers historiques : April, Metlife, Cardif, AFI ESCA
- Le courtage en crédit : AXA, Generali, BPSIs...
- Les réseaux d'agent et de salarié

La part des délégataires sur le marché reste faible (12% des cotisations en 2016, soit 1,0 milliard d'euros, part identique en 2015 ; source FFA). La proportion des cotisations versées aux délégataires sur l'assurance emprunteur immobilier est légèrement plus élevée (15%, identique 2015 ; source FFA).

## Répartition des cotisations en assurance emprunteur (immobilier) en 2016 (source FFA)



- **La sélection des risques**

La convention s'Assurer et Emprunter avec un Risque Aggravé de Santé AERAS a pour rôle de faciliter l'accès au crédit et à une assurance aux personnes ayant de graves problèmes de santé. Les assureurs ont alors l'obligation de proposer une tarification spécifique aux candidats à l'assurance pouvant bénéficier de la convention.

Elle a également défini les questionnaires médicales que les assureurs pouvaient utiliser, répondu aux questions sur la confidentialité des données et introduit le « droit à l'oubli ».

En fonction de l'âge à l'adhésion ainsi que du total des capitaux assurés, une sélection des risques est effectuée à la souscription, qui peut conduire à des majorations, des exclusions de garanties ou un refus. Ces risques peuvent concerner les domaines suivants :

- Risques médicaux : exclusions de pathologies, ...
- Professions à risques : profession avec port d'armes à feu, ...
- Risques de séjours ou de résidence : l'assuré séjourne ou réside dans un pays considéré comme risqué
- Risques sportifs : exclusions de sports tels que le parachutisme ou le saut à l'élastique, ...

Nous ne rentrerons pas dans les détails de cette sélection des risques et notamment, de son impact lors de la tarification du produit, mais nous nous appuyerons plutôt sur des études internes si besoin.

#### **1.1.4. Actualités réglementaires et mutations du marché**

Historiquement, l'assurance de prêt était un produit d'assurance vendu dans la grande majorité par les banques directement. Toutefois, des récentes réglementations ont permis au marché de s'ouvrir un peu plus à la concurrence, tout en tentant de protéger le mieux possible le consommateur. Nous présentons les mesures réglementaires suivantes afin de mieux décrire les changements s'effectuant sur le marché de l'assurance emprunteur.

##### **► Loi Lagarde**

La loi Lagarde a été adoptée le 1er juillet 2010. Elle transpose une directive européenne sur les crédits à la consommation. La principale mesure de cette loi est la libre concurrence pour le choix de son assurance de prêt par l'emprunteur. L'article L. 312-9 du Code de la consommation précise également : « *Le prêteur ne peut pas refuser en garantie un autre contrat d'assurance dès lors que ce contrat présente un niveau de garantie équivalent au contrat d'assurance de groupe qu'il propose. Toute décision de refus doit être motivée.* »

##### **► Loi Hamon**

La loi Hamon a été adoptée le 17 mars 2014. La principale mesure pour l'assurance emprunteur mise en place par cette loi est la possibilité de changer l'assurance de prêt immobilier durant une période d'un an à compter de la signature du prêt, à condition que ce soit pour une offre concurrente présentant des garanties équivalentes à celles proposées par la banque. Aucun frais ne peut alors être facturé par l'organisme prêteur.

##### **► Avis du Comité Consultatif du Secteur Financier (C.C.S.F.) sur l'équivalence du niveau de garantie en assurance emprunteur**

Les banques se sont engagées, depuis le 1er mai 2015, à utiliser pour l'équivalence des garanties uniquement 26 critères dont 18 pour les garanties emprunteur et 8 pour les garanties chômage, critères ayant été défini le 13 janvier 2015 par le Comité Consultatif du Secteur Financier (C.C.S.F.).

Chaque banque détermine 11 critères sur ces 18 définissant les garanties minimales exigées. Ils sont obligatoirement communiqués lors de la négociation du prêt. La banque a

par la suite 10 jours ouvrés pour examiner la demande de changement de contrat et doit motiver son refus de délégation de façon explicite et datée.

## ► **L'amendement Bourquin : résiliation annuelle**

Signé le 1<sup>er</sup> mars 2017, l'amendement Bourquin permet, chaque année, de résilier son contrat d'assurance de prêt, puis de le substituer par un contrat alternatif. Cette disposition sera applicable de façon systématique aux offres de prêts émises à partir du 22 mars 2017. A compter du 1<sup>er</sup> janvier 2018, elle sera applicable à tous les contrats d'assurance en cours.

Ce droit est consacré dans l'article L113-12 du Code de la Consommation :

*« [...] l'assuré a le droit de résilier le contrat à l'expiration d'un délai d'un an, en envoyant une lettre recommandée à l'assureur au moins deux mois avant la date d'échéance. Ce droit appartient, dans les mêmes conditions, à l'assureur. Il peut être dérogé à cette règle pour les contrats individuels d'assurance maladie et pour la couverture des risques autres que ceux des particuliers. Le droit de résilier le contrat tous les ans doit être rappelé dans chaque police. Le délai de résiliation court à partir de la date figurant sur le cachet de la poste. [...] »*

La Fédération française bancaire (FBF) s'est opposé à cet amendement en l'attaquant tout d'abord devant le Parlement avant de parvenir à renvoyer devant le Conseil constitutionnel cette question. Celui-ci a rendu son avis le 12 janvier 2018 et validé l'amendement Bourquin introduisant définitivement la possibilité de résiliation annuelle des contrats d'assurance emprunteur.

Cet amendement a profondément bouleversé le marché et a poussé les acteurs de celui-ci, anciens comme nouveaux, à s'adapter aux changements à venir. C'est notamment à la lumière de cet amendement qu'on a pu voir un certain nombre d'assureurs et de mutuelles émettre des appels d'offres avant d'entrer sur le marché. APRIL Santé Prévoyance est bien armé pour répondre à ce genre d'opportunités, comme nous le verrons dans la partie suivante.

## **1.2. APRIL et l'assurance emprunteur**

### **1.2.1. Présentation générale d'APRIL Santé Prévoyance**

APRIL conçoit, gère et distribue des produits d'assurances pour les particuliers, les professionnels et les entreprises. Elle emploie 3 800 employés et elle est présente dans 33 pays.

Créée en 1988, APRIL Santé Prévoyance est la filiale originelle du groupe APRIL. Courtier grossiste, elle se situe donc entre l'assureur et le courtier. Elle gère et conçoit des produits d'assurance santé et prévoyance pour tous les profils d'assurés (jeunes ou seniors, professionnels ou particuliers, risques aggravés). APRIL s'occupe également de gestion pour compte de tiers (appelé AGLAE depuis 1998) et produit également des produits en marque blanche (conception/gestion d'un produit commercialisé sous la marque d'un partenaire).

La stratégie de distribution d'APRIL Santé Prévoyance est multicanale : elle s'appuie sur un réseau d'assureurs-conseils très développé réparti sur l'ensemble du territoire français, mais elle possède également son propre réseau d'agences (April Mon Assurance) et s'appuie de plus en plus sur les différents acteurs présents sur Internet (comparateurs d'assurance, etc.).

### **1.2.2. L'assurance emprunteur au sein d'ASP**

APRIL Santé Prévoyance est le leader du marché de la délégation en assurance de prêts. Le courtier grossiste lyonnais a produit un CA de 528M€ sur l'activité Santé – Prévoyance en 2016 d'après les communiqués financiers publiés.

Le Service Actuariat Assurance Emprunteur, au sein duquel j'ai effectué mon alternance et ce mémoire, fait partie de la direction Conception Assurance Emprunteur et s'occupe exclusivement des contrats d'assurance de prêts. Il est chargé de leur conception et notamment, de leur tarification en relation avec les autres services, du suivi des produits et, de la relation avec les différents partenaires assureurs et réassureurs.

Enfin, depuis 20 ans, APRIL Santé Prévoyance est un expert de l'assurance de prêts. Fort de cette expérience, reconnu pour son savoir-faire technique, ses qualités de gestionnaire et une connaissance du marché conséquente par le biais de son important réseau de courtiers, la direction Conception Assurance Emprunteur est également sollicitée pour accompagner des partenaires externes dans la conception d'offres adaptées.

### **1.3. Tarification d'un contrat d'assurance emprunteur**

La tarification d'un produit d'assurance de prêt est un processus long et complexe.

Dans le cadre de la refonte récente de la gamme emprunteur, en se basant sur le portefeuille d'assurés conséquent disponible, l'Actuariat Emprunteur d'ASP a effectué de nombreuses études de modélisation du risque emprunteur qui ont servi de base à la tarification du dernier produit ADP commercialisé par APRIL Santé Prévoyance depuis

février 2017, avec en particulier l'identification des critères influant sur la sinistralité et la quantification des écarts entre modalités des différentes variables.

Dans cette partie, nous présentons de façon synthétique la méthodologie générale de construction tarifaire d'un contrat d'assurance emprunteur, telle qu'elle a été appliquée pour le produit ADP APRIL 2017. Cette étape d'analyse a été primordiale afin de déterminer le niveau de conservation ou d'adaptation des études existantes à réaliser dans le cadre de l'outil souhaité.

### **1.3.1. Méthodologie générale**

Les différentes étapes suivies lors de la conception tarifaire d'un produit en assurance emprunteur sont les suivantes :

- **Construction des primes pures théoriques par critère**
  - Identification des facteurs explicatifs de la survenance en décès et de l'entrée en indemnisation d'arrêt de travail.
  - Construction des taux de mortalité et d'incidence en indemnisation d'arrêt de travail par critère (vision *Best Estimate*).
  - Définition de la loi de maintien en indemnisation d'arrêt de travail à utiliser, soit au travers d'abattements applicables aux lois de maintien en arrêt de travail règlementaires BCAC calibrés à partir de la sinistralité observée sur le portefeuille, soit d'une loi de maintien d'expérience.
  - Détermination des primes pures théoriques égalisant l'engagement de l'assureur.
  
- **Construction des primes techniques associées** à la structure tarifaire cible (liée aux contraintes paramétrage et aux partis pris commerciaux) permettant d'égaliser au mieux les VAP assurés (avant intégration de la marge attendue) aux VAP assureurs
- **Définition des différents coefficients de passage** par modalité de critère
- **Intégration d'un niveau de marge technique** cible attendu permettant de garantir un certain niveau de rentabilité à nos partenaires porteurs de risque
- **Intégration des chargements, taxes et différents financements**

### **1.3.2. Facteurs explicatifs de la sinistralité**

Nous nous intéresserons ici aux différents facteurs explicatifs de sinistralité qui fournissent la segmentation technique.

- **L'âge**

Le premier facteur évident de discrimination de la sinistralité, que ce soit en décès ou en arrêt de travail, est l'âge. L'âge est un des facteurs de discrimination les plus importants en assurance vie, la segmentation technique des contrats d'assurance emprunteur le prend donc directement en compte.

- **Le sexe**

Le sexe est un autre facteur de discrimination de la sinistralité : l'espérance de vie des femmes est en moyenne plus élevée celle des hommes de même âge. Néanmoins, depuis un arrêt de la cour de justice de l'Union Européenne rendu le 1<sup>er</sup> mars 2011, les assureurs n'ont plus le droit d'utiliser ce critère comme discriminant pour les tarifs. Les tables de mortalité utilisées sont donc unisexes.

- **La Catégorie Socio-Professionnelle**

Il existe des différences de sinistralité suivant la catégorie socio-professionnelle de l'individu. Par exemple, selon l'Insee : « les hommes cadres vivent en moyenne 6 ans de plus que les hommes ouvriers dans les conditions de mortalité 2009-2013 ». De manière grossière, plus le diplôme est élevé, plus l'espérance de vie l'est. Nous pouvons également retrouver la même tendance pour les arrêts de travail.

Les CSP de tarification sont la résultante d'informations relatives à la situation professionnelle et à la nature de la profession de l'individu.

Sur le marché de l'assurance emprunteur, les CSP sont déterminées à partir du libellé du statut professionnel (cadre, profession libérale, fonctionnaire, ...) auquel s'ajoutent parfois des informations issues de questions sur la nature du travail : secteur d'activité, manutention, travail en hauteur, kilométrage. Cependant, il est connu qu'à l'intérieur d'un même statut professionnel, il peut exister des hétérogénéités au niveau du risque, ce qui induit une mutualisation naturelle. Par exemple, les professions libérales médicales n'ont pas les mêmes sinistralités que des professions libérales comme les avocats.

Afin d'être le plus proche du risque, ADP APRIL 2017 a introduit une innovation à ce niveau en décidant de déterminer les CSP à partir non plus seulement du seul statut, mais du croisement de ce dernier avec le libellé exact de la profession.

La granularité choisie pour les CSP a ainsi été définie dans le but de répondre à un objectif de clientèle spécifique. L'un des objectifs de l'outil que nous souhaitons construire sera de pouvoir en tarifier de nouvelles en s'appuyant sur celles-ci, ce qui se justifie par le fait que les CSP d'APRIL 2017 ont été construites au plus proche du risque et que les lois Best Estimate associées sont déjà disponibles.

- **Statut Fumeur/Non-fumeur**

Les modélisations effectuées montrent que le statut Fumeur induit une sur sinistralité en incidence sur les garanties Décès (DC/PTIA) et Arrêt de travail (ITT/IPT), d'autant plus élevée que l'assuré est âgé.

Ce n'est pas une surprise que les fumeurs aient une surmortalité vis-à-vis des non-fumeurs. Toutefois, l'impact du tabagisme sur l'incidence en Arrêt de travail est un résultat moins intuitif. Certaines études semblent montrer un lien entre le tabagisme et la dépression ([LASSER et al., \[2000\]](#))

Nous pouvons imaginer que les assurés fumeurs soient plus fréquemment en Arrêt de travail du fait d'un état de santé dégradé ou que le statut Fumeur soit le reflet d'un comportement spécifique.

Cette variable est un facteur potentiel de sélection adverse : un produit ne présentant pas de segmentation Fumeur/Non-fumeur pourrait potentiellement capter un grand nombre de fumeurs ce qui aurait un impact négatif sur la mutualisation.

A contrario, un tarif construit à partir des majorations *Best Estimate* est assez dissuasif et peut conduire à écarts tarifaires entre deux personnes co-assurées difficilement présentables et justifiables commercialement.

- **Zone géographique**

Les modélisations effectuées sur le Décès et l'entrée en Indemnisation d'incapacité démontrent des différences de sinistralités sur certains départements en Arrêt de travail.

Certains départements présentent ainsi une sur sinistralité alors que d'autres montrent une sous sinistralité. Il est toutefois difficile d'interpréter plus profondément ces résultats sans une étude qualitative plus approfondie car l'effet provenant du département peut être la résultante d'autres variables exogènes non disponibles dans nos bases de données (taux de chômage, densité médicale, ...). Par exemple, [BEN HALIMA \[2011\]](#) explique la disparité des arrêts maladies entre départements.

De même que pour la segmentation effective des CSP ou la prise en compte des écarts de sinistralité entre fumeurs et non-fumeurs, l'intégration de ce facteur en critère tarifant définitif dépend également de considérations plus commerciales.

- **Présence d'un co-emprunteur**

Les modélisations effectuées tendent à indiquer que la présence d'un multi-emprunteur est neutre que ce soit sur la sinistralité en Décès ou la sinistralité en Arrêt de travail, alors même qu'elle est fréquemment intégrée dans les structures tarifaires des produits ADP.

L'intégration de ce critère à la structure tarifaire est alors un arbitrage commercial, afin d'être cohérent avec les pratiques du marché et dans la continuité des précédents produits APRIL qui la proposaient justement.

- **Objet du prêt et montant de capital assuré**

- **Décès / PTIA**

Les modélisations effectuées tendent à montrer que l'objet du prêt et le montant de capital assuré ne sont pas discriminants sur la sinistralité, ce qu'on peut expliquer soit car les informations sont réellement non significatives, soit car elles sont déjà captées dans les effets des autres facteurs (CSP, tranche d'âges, zone géographique).

- **Incapacité**

L'objet du prêt semble être informatif avec une sous-sinistralité en entrée en AT pour les assurés ayant un prêt de type professionnel. Pas d'apport de la tranche de capital assuré.

### **1.3.3. Construction des taux de primes pour les garanties DC/PTIA et ITT/IPT**

Nous aborderons ici succinctement les différentes étapes suivies et les hypothèses prises dans la construction des différentes lois « Décès » et « Arrêt de travail ». Pour plus de détails sur la théorie associée, le lecteur intéressé pourra se référer à l'ouvrage de [PLANCHET & THEROND \[2006\]](#).

- **Lois d'incidence empiriques**

Les lois empiriques utilisées pour la construction tarifaire d'ADP APRIL 2017 ont été construites en appliquant l'estimateur de Kaplan-Meier par risque x CSP x statut Fumeur/Non-fumeur, après exclusion des assurés avec majorations médicales, ainsi que pour l'arrêt de travail des assurés ayant souscrit l'une des options rachetant les exclusions sur le dos et le psy.

La couverture dos/psy est en effet une option rachetable en échange d'une majoration de la prime dont le calibrage a lieu par ailleurs.

L'exclusion des assurés ayant des majorations médicales revient à prendre l'hypothèse que les majorations définies par les tarificateurs sont *best estimate*. C'est une hypothèse forte qui n'a jamais été validée empiriquement.

- **Lissages**

Un premier ajustement théorique a été réalisé par transformation d'une loi de survie modèle (la loi d'expérience APRIL certifiée pour le décès / la loi moyenne toutes CSP confondues pour l'entrée en indemnisation d'ITT), avec minimisation de l'erreur quadratique entre les  $q_x$  empiriques et les  $q_x$  modélisés.

Un second ajustement a ensuite été effectué en retravaillant les lois obtenues de sorte à corriger les minorations / majorations trop fortes sur les âges sous-représentés, à intégrer les coefficients de marge et de prudence souhaitées et à obtenir un risque croissant avec la CSP.

Cette étude a été effectuée en modulant les majorations / minorations par âge appliquées aux  $q_x$  et  $i_x$  issus des lois modèles, à partir de l'analyse :

- des nombres de sinistres observés et modélisés par CSP x tranches d'âges ;
- de la croissance des taux de mortalité et d'entrée en fonction de l'âge ;
- de la croissance du risque avec la CSP.

- **Lois de maintien en arrêt de travail**

Les lois de maintien utilisées pour la tarification de l'ITT / IPT sont les tables BCAC (maintien en incapacité, passage en invalidité, maintien en invalidité), prolongées jusqu'aux âges limites d'indemnisation et abattues d'un coefficient calibré à partir de la sinistralité observée.

- **Taux de Primes pures en %CRD**

Les primes payées par l'assuré ont pour assiette le CRD au 1<sup>er</sup> janvier de chaque année, alors que la prestation payée en Décès / PTIA sera le CRD au moment du décès.

De même, le nombre maximal de mensualités payables en ITT / IPT est fonction du nombre de mensualités d'emprunt restantes. Il décroît donc chaque mois à l'intérieur d'une même année calendaire.

Les taux de primes pures sont donc calculés en prenant en compte ce facteur et en égalisant les VAP de chaque année.

Pour calculer les VAP, nous considérons tous les flux possibles, que nous actualisons financièrement, puis que nous pondérons par les probabilités.

Ainsi, dans le cas où la 1<sup>ère</sup> année d'assurance comporte 12 mois, le taux de prime annuel est donné par :

$$\pi_x^{risque,annee\_n} = \frac{12 \cdot VAP_{assureur}^{risque,annee\_n}}{\left( \sum_{k=12 \cdot (n-1)}^{12 \cdot n - 1} P_x^{(m)} \cdot \frac{1}{(1+i^{(m)})^k} \right) \cdot CRD_{12 \cdot (n-1)}}$$

Où :

$x$  : âge de l'assuré en début de période

$i$  : taux d'actualisation annuel

$i^{(m)}$  : taux d'actualisation mensuel

$$i^{(m)} = (1+i)^{1/12} - 1$$

$tx$  : taux d'intérêt de l'emprunt

$M_{k-1}$  : k-ième mensualité de l'emprunt,  $k \geq 1$

$CRD_k$  : le capital restant dû assuré après  $k$  mensualités

$$CRD_0 = capital\_assure\_initial, \quad CRD_k = CRD_{k-1} \cdot \left(1 + \frac{tx}{12}\right) - M_{k-1}, \quad k \geq 1$$

$q_x^{(m)}$  : probabilité mensuelle de survenance d'un sinistre décès / PTIA pour un assuré d'âge  $x$

$$q_x^{(m)} = 1 - (1 - q_x)^{1/12}$$

$r_x^{(m)}$  : probabilité mensuelle de chute (hors décès indemnisé) pour un assuré d'âge  $x$

$$r_x^{(m)} = 1 - (1 - r_x)^{1/12}$$

${}_{\frac{k}{12}}P_x^{(m)}$  : probabilité pour un assuré d'âge  $x$  d'être encore en portefeuille au k-ième mois d'emprunt

$${}_{\frac{k}{12}}P_x^{(m)} = {}_{\frac{k-1}{12}}P_x^{(m)} \cdot \left( 1 - q_{\left[ x + \frac{k-1}{12} \right]}^{(m)} \right) \cdot \left( 1 - r_{\left[ x + \frac{k-1}{12} \right]}^{(m)} \right)$$

L'ensemble des formules de calcul des VAP assureurs se trouve en annexe.

- **Egalisation des Valeurs Actuelles Probables : détermination des taux de primes techniques sous structure d'ADP AVRIL 2017**

PETAUTON [2004] montre qu'une des contraintes en tarification est l'égalisation des VAP concernant des primes (i.e. l'engagement de l'assuré) et des VAP concernant des prestations (i.e. l'engagement de l'assureur).

En effet, les primes pures théoriques sont fonction de facteurs liés à l'assuré et aux caractéristiques de son prêt :

- Age
- CSP
- Durée et taux d'emprunt
- Année d'ancienneté
- ...

En pratique, compte-tenu du nombre de combinaisons possibles et de la complexité associée, tous ces facteurs ne sont pas retenus dans la structure tarifaire finale avec leur granularité la plus fine (durée d'emprunt et ancienneté notamment).

Les taux de primes techniques réellement retenus pour la construction tarifaire finale (i.e les taux appelés pour calculer les primes commerciales nettes de taxes et de chargements) sont

ainsi déterminés en trouvant le meilleur compromis entre simplicité de structure tarifaire (pour des besoins de paramétrage dans les outils et d'appropriation / portage par la Gestion et le Commerce) et adéquation au risque.

Les approximations sont réalisées en minimisant les écarts entre les VAP assureurs et assurés.

Une nouvelle fois, le produit ADP APRIL 2017 possède une structure tarifaire spécifique qui a été pensée pour répondre au mieux aux contraintes et aux objectifs qui lui étaient liés. Les produits que nous chercherons à tarifer posséderont leurs propres contraintes et objectifs, et donc, une structure tarifaire a priori bien différente.

## **2. Construction d'un outil de tarification automatisé**

Dans le but d'être plus efficace dans la réponse aux appels d'offres, nous cherchons à créer un ou plusieurs outils qui permettraient de gagner en rapidité, mais également en souplesse face aux différentes cibles clients et contraintes associées à la conception d'une nouvelle gamme en propre ou en réponse à un appel d'offre.

L'objectif étant d'utiliser les bases techniques existantes et construites telles que présentées dans la partie 1, nous aborderons dans cette partie une méthodologie de tarification destinée à faire le meilleur usage possible des bases en notre possession, et ceci, afin de gagner du temps dans le processus de tarification. Notre processus devra être également être souple afin de s'adapter à des populations cibles aux caractéristiques pouvant être très différentes de celles du portefeuille ASP. Nous présenterons dans un premier temps de manière générale le principe de tarification envisagé avant d'entrer davantage dans les détails dans une deuxième partie.

### **2.1. Présentation du processus de tarification envisagé**

#### **2.1.1. Rappel des éléments disponibles**

Dans le cadre de la construction tarifaire d'un nouveau produit en réponse à un appel d'offre ou en propre, nous disposons de divers éléments issus de la conception d'ADP APRIL 2017 :

- Les taux de primes techniques *Best Estimate* pour les CSP ADP APRIL 2017, âge atteint et tranche de durée, pour les garanties Décès/PTIA et ITT/IPT.
- Les majorations techniques *Best Estimate* des différents critères discriminants de sinistralité étudiés (Statut Fumeur/Non-fumeur ; Statut marital ; Zone géographique ; Option de rachat dos/psy, ...)

Nous disposons également d'éléments plus ou moins détaillés sur la structure de population attendue.

La tarification est basée sur une population ayant des caractéristiques spécifiques, que nous nous attendons à retrouver lors de la phase de commercialisation du contrat. En d'autres termes, afin de considérer que la tarification est adaptée, le portefeuille d'assurés liés à ce contrat doit avoir des caractéristiques similaires à celles de la population sur laquelle se base la tarification. Néanmoins, après avoir tarifé le produit, nous pouvons effectuer une comparaison de ce tarif avec la concurrence ainsi qu'étudier sa rentabilité. Ces deux analyses peuvent alors nous indiquer soit que les hypothèses prises sur les caractéristiques

d'assurés sont fausses (si un concurrent propose un tarif plus attractif sur la cible par exemple) soit que le produit ne sera pas suffisamment rentable. Il est alors nécessaire d'être en mesure de simuler un portefeuille fictif d'assurés afin d'effectuer ces analyses, mais également afin de pouvoir ajuster le tarif et résoudre cette problématique s'apparentant à une référence circulaire.

### **2.1.2. Problématiques liées à la structure des primes ADP APRIL 2017**

Les taux de primes disponibles pour ADP APRIL 2017 suivent une certaine structure. Toutefois, ce produit est très segmenté, c'est un produit de spécialiste, destiné à une clientèle très précise. Afin de répondre à des appels d'offres, il nous semble adapté de concevoir un produit plus simple, avec notamment des CSP différentes.

Plusieurs raisons peuvent en effet pousser à choisir une structure différente :

- Le choix de la segmentation est avant tout une décision de cible clientèle et commerciale tarifaire car une segmentation moins forte implique nécessairement une plus grande mutualisation. Ainsi, si on réduit le nombre de CSP, on augmente la mutualisation avec le risque d'afficher un tarif moins attractif sur les cibles les moins sinistrées et inversement sur les cibles les plus risquées. De même, si on supprime la distinction Fumeur/Non-fumeur, on augmente le tarif Non-fumeur. Cela peut engendrer un phénomène d'anti-sélection si les tarifs des concurrents sont moins chers sur les profils moins risqués.
- La complexité du produit est également importante dans le sens où il est nécessaire que le produit soit aisément compris par les canaux de vente, afin d'éviter les erreurs et de maximiser leur efficacité.
- Le système informatique servant au paramétrage limite également les possibilités de tarification du produit d'assurance.

Nous allons voir dans quelles mesures nous pouvons créer un outil prenant en données d'entrée les éléments d'ADP APRIL 2017 pour parvenir à notre objectif.

### **2.1.3. Principe de déformation des primes**

Notons  $E$  l'ensemble des états,  $j_{ref}$  l'état de référence et  $j_k$ , les autres états, où  $k = 1, \dots, card(E) - 1$  et tel qu'on ait  $E = j_{ref} \cup (\cup_k j_k)$ . Par exemple, pour le statut Fumeur/Non-fumeur, on aura :  $E = \{Fumeur; Non-fumeur\}$ ;  $j_{ref} = Non-fumeur$ ;  $j_1 = Fumeur$ .

On note :

$x$  : **âge** de l'assuré en début de période

$\theta_x^k$  : le **poinds en capital assuré** des individus dans l'état  $j_k$  à l'âge  $x$

$M_x^k$  :  $M_x^k = \frac{\text{mortalité des assurés dans l'état } j_k \text{ à l'âge } x}{\text{mortalité des assurés dans l'état } j_{ref} \text{ à l'âge } x} - 1 = \text{ratio de mortalité} - 1$ ,

c'est-à-dire la **majoration technique** *best estimate*.

$\text{passage}_x^k$  : le coefficient de passage qu'on souhaite imposer entre le tarif de référence et le tarif de l'état  $j_k$ , c'est-à-dire la **majoration commerciale**.

Par exemple, si on prend  $j_{ref} = \text{« Non-Fumeur »}$  et  $j_1 = \text{« Fumeur »}$ , on peut observer  $M_x^1 = 50\%$  (i.e. la mortalité Fumeur est 1,5 fois plus élevée que la mortalité non-fumeur), il peut être décidé d'imposer un coefficient de passage pour le tarif final supérieur ou inférieur à ce ratio de mortalité (par exemple  $\text{passage}_x^k = 30\%$ ). De même, pour le statut marital, une réduction pour les assurés « En couple » peut être voulue, mais on n'observe pas de réelle sous-sinistralité sur cette population. Dans ce cas, la population en état « Isolé » sera majorée pour tenir compte d'un besoin de réduction commerciale de la population en état « Couple ».

On note également :

$\pi_x^{j_{ref}.init}$  : le taux de prime du tarif de référence initial

$\pi_x^{j_{ref}.déf}$  : le taux de prime du tarif de référence après transformation

Alors :

$$\pi_x^{j.déf} = \frac{\theta_x^{j_{ref}} + \sum_k \theta_x^k (1 + M_x^k)}{\theta_x^{j_{ref}} + \sum_k \theta_x^k (1 + \text{passage}_x^k)} \pi_x^{j_{ref}.init}$$

Cette relation ne sera pas démontrée formellement ici, mais nous vérifierons de manière empirique dans la dernière partie de ce mémoire qu'elle est valable. Intéressons-nous néanmoins à l'intuition se cachant derrière cette équation.

**Preuve** : On utilise la formule de décomposition des probabilités totales sur la probabilité de décès, sur les événements disjoints  $\{j_{ref}, j_1, \dots, j_{card(E)-1}\}$  :

$$\begin{aligned} q_x &= P[T \leq x] = P[T \leq x | j_{ref}] \cdot P[j_{ref}] + \sum_k P[T \leq x | j_k] \cdot P[j_k] \\ &= P[T \leq x | J] \cdot P[j_{ref}] + \sum_k P[T \leq x | j_k] \cdot (1 - P[j_{ref}]) \\ &= \alpha_j q_x^{j_{ref}.init} + (1 - \alpha_j) \sum_k q_x^{j_k.init} \end{aligned}$$

Cette propriété reste vraie après qu'on impose la majoration commerciale donc on a l'égalité :

$$\alpha_J q_x^{j_{ref}.init} + (1 - \alpha_J) \sum_k q_x^{j_k.init} = \alpha_J q_x^{j_{ref}.def} + (1 - \alpha_J) \sum_k q_x^{j_k.def}$$

Or, par définition :

$$q_x^{j_k.init} = (1 + M_x^k) q_x^{j_{ref}.init}$$

$$q_x^{j_k.def} = (1 + passage_x^k) q_x^{j_{ref}.def}$$

Donc :

$$\alpha_J q_x^{j_{ref}.init} + (1 - \alpha_J) \sum_k (1 + M_x^k) q_x^{j_{ref}.init} = \alpha_J q_x^{j_{ref}.def} + (1 - \alpha_J) \sum_k (1 + passage_x^k) q_x^{j_{ref}.def}$$

Finalement :

$$q_x^{j_{ref}.def} = \frac{\alpha_J + (1 - \alpha_J) \sum_k (1 + M_x^k)}{\alpha_J + (1 - \alpha_J) \sum_k (1 + passage_x^k)} q_x^{j_{ref}.init}$$

□

#### **2.1.4. Calcul de la prime commerciale**

La formule pour calculer le taux de prime commerciale est la suivante :

$$\text{Taux de prime commercial HT} = \text{Taux de prime technique} \times \frac{(1 + \text{marge})(1 + \text{financement})}{1 - \text{chargement pour frais de gestion} - \text{chargement apport} - \text{assureur}}$$

- Les chargements pour frais de gestion (APRIL Santé Prévoyance) sont destinés à couvrir les différents frais exposés pour assurer l'exploitation de l'entreprise. Ils peuvent être segmentés si on est en mesure de justifier cette segmentation. Par exemple, on peut imaginer qu'en raison de frais fixes, les adhérents qui assurent de plus petits capitaux nécessitent une plus grande part de leurs primes servant à couvrir les frais de gestion.
- Les chargements pour frais d'apport (ou frais d'acquisition) correspondent aux dépenses engagées dès qu'une police d'assurance est établie et qu'il est nécessaire de rémunérer un apporteur. Le montant de ces commissions est déterminé de sorte à

ce qu'il soit au moins égal à la valeur des commissions réellement versées aux apporteurs.

- Les marges techniques sont issues d'écart entre la mortalité observée et la mortalité prévue au tarif, segmentées par CSP et par âge atteint.
- Les frais d'assureur représentent la commission destinée au fonctionnement du partenaire assureur.
- Les frais de financement sont destinés à couvrir les différentes opérations commerciales et autres frais divers (honoraires médicaux,...).

Afin d'obtenir la prime TTC, il faut également ajouter à la prime commerciale les différentes taxes suivantes :

- Décès toutes cause : 0%
- Décès Accidentel : 9%
- PTIA / ITT / IPT / IPP / Invalidité spéciale professions médicales / GIS : 9%

Parmi ces divers éléments, nous disposons de deux leviers principaux afin d'effectuer des derniers ajustements sur les tarifs : les chargements pour frais de gestion et les marges techniques. En effet, les autres frais sont fixes, en dehors de notre capacité d'action (hormis les frais d'apport si on est en capacité de ne proposer que des commissions courtiers qui permettraient de les réduire).

Les chargements pour frais de gestion sont calculés lors de la conception du produit, mais les frais réellement supportés ne seront connus avec exactitude que bien plus tard et à condition de disposer d'une comptabilité claire et détaillée le permettant. Ils peuvent être réduits par des pratiques de gestion comme, par exemple, un ajustement de la sélection médicale en entrée (télé sélection, e-sélection).

En dehors des marges de prudence inhérentes aux tables de mortalité et d'entrée en indemnisation, nous devons estimer une marge technique attendue, en fonction de la rentabilité prévue, mais également des comparaisons concurrentielles.

L'outil présenté plus tôt permet également d'effectuer le passage de la prime technique à la prime commerciale. Les frais de gestion et les marges techniques sont définissables grâce à une structure assez libre. Cela nous sera très utile dans la dernière partie de ce mémoire lorsque nous chercherons à optimiser la structure tarifaire de ce produit.

## **2.2. Description de l'outil et souplesse de la structure tarifaire**

### **2.2.1. Fonctionnement général de l'outil**

L'outil envisagé prendra la forme d'un classeur Excel.

Il prendra en entrée des éléments de différentes natures.

- Des inputs issus des bases ASP, présentés précédemment dans ce mémoire, fixes *a priori*. L'outil permettra néanmoins la modification de ces données grâce à la possibilité de choisir un niveau de détail très fin, notamment en ce qui concerne les taux de primes vision *Best Estimate* âge par âge pour chacune des Catégories Socio-Professionnelles et les coefficients de sur-sinistralité et sous-sinistralité pour chaque modalité des critères de segmentation. Ainsi, si une analyse ultérieure permet d'affiner ces données, il sera toujours possible d'actualiser l'outil.
- Des données sur la population cible : répartition par âge pour chaque nouvelle Catégorie de tarification, répartition des capitaux assurés entre CSP APRIL 2017 pour chaque nouvelle Catégorie de tarification et le poids en capitaux assurés pour chaque modalité différente de la catégorie de référence (par exemple, fumeur, présence d'un co-emprunteur, ...)
- Des paramètres sur la structure tarifaire choisie pour le produit : critères de segmentation, majorations / minorations voulues (âge par âge ou par tranches d'âges), etc.

A partir de ces différents éléments et du principe présenté dans la partie précédente de ce mémoire, l'outil permet d'obtenir les taux de primes techniques pour la catégorie de référence, puis pour chacune des autres catégories à partir des coefficients de majoration et de minoration choisis dans la structure tarifaire.

Dans un second temps, à partir de paramètres sur les taux de chargements et sur la marge technique, l'outil calculera les taux de primes commerciales pour la catégorie de référence et nous obtiendrons de la même manière que précédemment les primes commerciales pour toutes les autres catégories.

## 2.2.2. Présentation des Inputs

- **Structure tarifaire cible paramétrable**

La structure tarifaire pourra être modulée aux niveaux suivants :

- De 1 à 10 Catégories Socio-Professionnelles discriminantes, avec un coefficient de passage (linéaire ou âge par âge) entre les taux de primes de la CSP de référence et les autres CSP. Elles seront définies par le poids en capital assuré de chacune des CSP d'ADP APRIL V4.
- Fumeur/Non-fumeur : possibilité de déterminer un coefficient de majoration linéaire ou âge par âge entre la référence, Non-fumeur, et l'état Fumeur.
- Zone géographique : possibilité de déterminer un coefficient de majoration linéaire ou âge par âge entre la référence, « Pas de réduction », et les états « Réduction 1 » et « Réduction 2 ».
- Présence d'un ou plusieurs co-emprunteurs : possibilité de déterminer un coefficient de majoration linéaire ou âge par âge entre la référence, « Non », et l'état « Oui ».
- Rachat des exclusions dos/psy en Arrêt de travail : possibilité de déterminer un coefficient de majoration linéaire ou âge par âge entre la référence, « Pas d'option », et l'état « Option rachat dos/psy ».
- IPP : possibilité de déterminer un coefficient de majoration linéaire ou âge par âge entre la référence, « Pas d'IPP », et l'état « IPP ».

<b>Nombre de CAT de tarification</b>	<b>3</b>
--------------------------------------	----------

### Coef. de passage entre CAT

Type	Linéaire
CAT1 --> CAT1	0%
CAT1 --> CAT2	10%
CAT1 --> CAT3	30%
CAT1 --> CAT4	0%
CAT1 --> CAT5	0%
CAT1 --> CAT6	0%
CAT1 --> CAT7	0%

CAT1 --> CAT8	0%
CAT1 --> CAT9	0%
CAT1 --> CAT10	0%

**coefficient passage**

	Type	Taux --> si linéaire
Co-emprunteur(s)	Linéaire	-10%
Fumeur	Linéaire	25%
Option	Linéaire	40%
IPP	Linéaire	40%
Géo	Linéaire	-40%

**Tableau 1 - Capture d'écran : Sélection de la structure tarifaire pour la garantie AT**

Il sera néanmoins impossible de changer d'autres modalités directement dans l'outil Excel. Ce sera notamment le cas de tout ce qui touche à la structure même des taux de prime *Best Estimate* ADP APRIL V4 (données en entrée). Ainsi :

- Les taux de primes sont exprimés en pourcentage du Capital Restant Dû et par âge atteint.
- Garantie Décès : pas de tranche de durées.
- Garantie Arrêt de Travail (I.P.T. /I.T.T.) : nous faisons le choix arbitraire de 3 tranches de durée qui sont (en années) : [0 ; 10[ ; [10 ; 20[ ; [20 ; 40].

Enfin, certaines modalités n'ont pas encore été envisagées et une amélioration de l'outil est possible :

- Les prêts In Fine n'ont pas été pris en compte.
- Les franchises n'ont pas été prises en compte (99% des franchises souscrites sont des franchises 90 jours, ce qui correspond à la référence prise dans notre étude).

• **Majoration / Minoration tarifaire par modalité de critère**

Les critères de segmentation sont des éléments fondamentaux d'un produit d'assurance emprunteur. Une structure plus légère permet une meilleure appréhension des commerciaux alors qu'un nombre plus élevé de critères implique que les commerciaux auront besoin de

davantage de temps et de mieux maîtriser le marché de l'assurance emprunteur pour être au meilleur niveau de vente.

Afin d'assurer la souplesse de la structure tarifaire du produit conçu par notre outil, nous avons laissé une liberté sur la définition de coefficients de majoration ou de minoration tarifaire pour chaque modalité de critère. Ainsi, il est possible de choisir un coefficient âge par âge ou linéaire (i.e. même pour tous les âges) pour chaque modalité de critère défini précédemment. Il est néanmoins nécessaire de garder en tête l'impact que peuvent avoir le choix de ces coefficients sur le tarif.

En effet, il y a pour chacun de ces coefficients un équivalent *Best Estimate* qui représente la vision la plus proche de la sur-sinistralité ou de la sous-sinistralité pour chaque modalité de critère. Afin de respecter le principe de déformation des primes que nous avons présenté précédemment, si nous choisissons de fixer un coefficient de majoration inférieur au coefficient de majoration *Best Estimate*, il en découle que le taux de prime de référence sera plus élevé. A contrario, si nous prenons un coefficient inférieur au coefficient *Best Estimate*, le taux de prime sera plus faible.

- **Répartition cible attendue par critère**

Le principe de déformation des primes que nous avons présentés précédemment nécessite de posséder des données sur la répartition cible attendue par critère.

Dans le cas d'un d'appel d'offres, le niveau d'informations fourni est très faible, voire nul pour certains informations propres à l'assurance emprunteur, comme le taux de souscription aux garanties, les quotités, etc. Lorsque l'information nous est fournie, elle correspond à des statistiques moyennes globales et elle n'est jamais au niveau de détail de ce que nous pouvons nous attendre à trouver dans les bases de données d'ASP. Hors, la méthodologie de tarification que nous proposons nécessite des données de répartition cible attendue par critère. Plus le niveau de détail est élevé sur ces informations, plus nous nous attendons à ce que la prime calculée soit proche du véritable niveau de risque.

De plus, si nous nous limitons à l'observation du portefeuille ASP afin de pallier aux lacunes dans les informations à notre disposition, il nous semble évident que cela fera apparaître des biais : la population cible pour laquelle le tarif est conçu n'est pas la même que celle d'APRIL. Il est de plus nécessaire de prendre en compte les dépendances entre chaque variable (une personne âgée emprunte statistiquement un capital plus élevé qu'une personne plus jeune, etc.).

C'est là tout l'intérêt d'avoir à disposition un portefeuille fictif à disposition : grâce à cela, nous pouvons calibrer ces inputs de manière plus précise et s'assurer d'avoir un tarif plus

proche de la réalité. La partie suivante de notre mémoire traitera précisément des différentes méthodes possibles pour générer un portefeuille fictif d'assurés et laquelle nous semble la plus adaptée à nos problématiques.

- **Chargements**

Les chargements inclus dans la prime commerciale sont destinés à couvrir les frais engagés par ASP au cours de la vie du contrat que ce soit pour les frais liés à la gestion des contrats (à la souscription et en cours de vie du contrat) ou les frais liés à la rémunération des apporteurs. Nous distinguons les chargements d'acquisition et les chargements de gestion. Tous les deux sont déterminés suivant un coefficient de proportionnalité à la prime technique.

La rémunération des commerciaux dépend du taux de commission choisi par ceux-ci. Nous ne cherchons pas à différencier ce chargement selon les caractéristiques des assurés. Notons néanmoins que sur le marché, il y a deux types de commissions qui représentent le gros du volume : la commission en 40/10 (40% de la prime HT en 1<sup>ère</sup> année, 40% de la prime HT les années suivantes) et la commission en 15/15 (15% de la prime HT sur toutes les années). Il est à noter que la commission en 15/15 est en réalité une commission 40/10 lissée sur 6 années. Pour faciliter notre étude, nous nous limiterons donc à une commission en 15/15 avec le chargement équivalent décidé pour les portefeuilles ASP.

Les frais de gestion peuvent être au contraire différenciés selon les caractéristiques des clients d'après le principe suivant : il y a une partie incompressible de frais fixe, à la souscription notamment, et, le chargement de gestion étant proportionnel à la prime technique, les contrats ayant des primes plus petites sont donc plus coûteux à ce niveau-là. Cependant, le niveau de chargement de gestion est en général donné à une maille globale.

De plus, il y a une partie de marge technique (parfois également appelé chargement de sécurité) qui est ajoutée à la prime technique. Cela permet à l'assureur de se couvrir face à la volatilité naturelle des sinistres. Il peut être décidé d'avoir un niveau de marge technique global ou suivant des cibles particulières (par exemple, un niveau de marge technique minimale par CSP ou par garantie).

Notons qu'il y a en général un chargement de sécurité « naturel », dans le sens où il est inclus directement dans les lois de mortalité et d'incidence en arrêt de travail lors de la construction de celles-ci.

Enfin, il est possible de répartir ces chargements et la marge technique grâce à des financements croisés suivant la cible d'assurés privilégiée. Ceci va nous permettre d'ajuster notre tarif pour atteindre les objectifs de cible et de rentabilité voulus.

### **2.2.3. Définition des Outputs**

Les Outputs de notre outil sont de deux natures : les primes techniques au format « commercial » et les véritables primes commerciales toutes taxes comprises.

Ces dernières dépendant notamment du paramétrage effectué sur les chargements et la marge technique. Ces deux variables peuvent être considérées dans une certaine limite comme des variables d'ajustement du tarif. En général, le niveau donné pour celles-ci est assez global. Nous pouvons chercher à le répartir grâce à des données de volume. Encore une fois, le portefeuille fictif d'assurés et les différentes analyses de concurrence et de rentabilité nous permettront d'apprécier le calibrage de ces paramètres afin d'optimiser la structure tarifaire de notre produit.

## **2.3. Validation de la tarification**

Maintenant que nous avons défini la méthodologie de tarification ainsi que le fonctionnement de l'outil associée à notre proposition de processus de tarification « rapide », nous nous proposons de vérifier la tarification. Pour cela, nous suivons la méthode suivante :

- Les engagements de l'assuré seront ventilés par exercice d'affectation des primes.
- Les engagements de l'assureur seront ventilés par exercice de survenance des sinistres.
- L'équilibre actuariel est alors établi par construction car la somme des primes probables actualisées en  $t=0$  est égale à la somme des sinistres probables actualisés en  $t=0$ .

Les formules de calcul des différents postes du compte de résultat diffèrent suivant que nous sommes dans le cas de la validation du modèle de tarification ou dans le cas de la constitution du compte de résultat.

De plus, une hypothèse importante que nous prenons est que le calcul des différents postes de notre compte de résultat sont réalisés de façon probable et donc théorique, et cela à



De plus, afin de faciliter ce travail, nous pouvons considérer avec un œil plus attentif en tout premier lieu les segments où nous savons qu'il y aura du volume. Par exemple, comme le montre le tableau ci-dessous, concernant les durées d'emprunts, nous nous focalisons sur les durées de 5, 7, 10, 12, 15, 17, 20, 25 et 30 ans, car ce sont ces durées qui sont les plus courantes dans la pratique.

### **3. Construction d'un générateur de portefeuille fictif**

Nous avons vu dans la partie précédente que disposer de statistiques moyennes globales peut s'avérer insuffisant pour tarifer correctement une offre. Cela est d'autant plus vrai pour apprécier convenablement la rentabilité du produit conçu.

L'analyse de celle-ci permet de savoir pour ASP si les niveaux de chargement et de marge intégrés en tarification sont suffisants pour garantir la rentabilité de l'offre tant pour ASP que pour ses porteurs de risque partenaires.

April Santé Prévoyance est un courtier grossiste et ne porte pas le risque. Pour s'occuper de cela, ASP s'entoure de partenaires assureurs et réassureurs. Les notions de rentabilité ne sont alors pas les mêmes pour ASP et les porteurs de risque. Afin de dépeindre la vision de ces deux parties, nous nous intéressons à définir deux notions de rentabilité, la rentabilité technique et la rentabilité ASP.

La rentabilité technique est la différence entre les primes techniques et les charges sinistres. Elle représentera notre indicateur de rentabilité pour les partenaires porteurs de risque.

La rentabilité ASP est une notion différente. En effet, la marge brute ASP correspond à la différence entre les commissions de gestion, d'apport et de développement que perçoit ASP d'un côté et les rétrocessions aux courtiers de l'autre. Nous définissons alors le taux de marge brute comme le ratio entre la marge brute ASP et la somme des cotisations HT perçues.

Ces deux notions de rentabilité dépendent de l'horizon auquel on les considère. Il est ainsi courant de regarder des horizons différents à 1 an, 4 ans, 7 ans ou à l'ultime. Il y a deux principales raisons à cela : l'allure des courbes de primes et le taux de chute.

Ainsi que le besoin d'apprécier ces notions à différents horizons de temps : en 1<sup>ère</sup> année, à 4 ans, 7 ans, à l'ultime....

#### **3.1. Les limites de l'utilisation des model points**

Dans le cas d'un produit pour lequel on prévoit que la cible clientèle sera similaire à celle présente dans notre portefeuille, on utilise les données issues de celui-ci. Le biais par rapport à la réalité sera alors, en théorie, négligeable. Les dépendances entre les variables semblent en effet bien captées.

Mais si le produit a pour objectif d'atteindre une cible différente que celle présente dans notre portefeuille, cela se complique

En particulier, le principal moyen utilisé jusqu'à aujourd'hui par l'équipe conception emprunteur d'ASP se montre limité dans la situation d'un appel d'offres, où les données sont trop globales et où on se retrouve alors dans la nécessité de s'appuyer uniquement sur les données du portefeuille stock ASP.

Le moyen en question consiste à résumer le portefeuille d'assurés cible à un model point. L'idée est de considérer des contrats agrégés dont les caractéristiques imitent celles du portefeuille complet. Le caractère partiel des données représentées épouse celui des données à disposition lors d'un appel d'offres : par exemple, la variable « Âge à la souscription de l'assuré » est découpée en tranches, de même pour la variable « Capital emprunté », etc.

Tr. Age (ans)	Tr. capital	Tr. Durée (mois)	Sexe	Co-emprunteur (s)	Cat_pro	Fumeur	Zone_geo	Moy. Age (ans)	Moy. Capital (€)	Moy. Durée (mois)	Nb. assurés
<25	<50k€	<50	Femme	NON	Cadre	NON	Reference	24	25000	36	1
<25	<50k€	<50	Femme	NON	Non cadre	NON	Reference	22	20001	48	2
<25	<50k€	<50	Homme	NON	Non cadre	NON	Reference	22,3	18433,2	40	3
<25	[50k€ ; 100k€[	<50	Femme	OUI	Non cadre	NON	Réduction	22	88000	12	1
<25	<50k€	[100 ; 150[	Femme	NON	Cadre	NON	Réduction	24	20000	120	1
<25	<50k€	[100 ; 150[	Homme	NON	Cadre	NON	Réduction	24	23686	143	1
<25	<50k€	[100 ; 150[	Homme	NON	Cadre	NON	Reference	22	24000	132	2
...	...	...	...	...	...	...	...	...	...	...	...

**Tableau 3 - Exemple de model point utilisée pour l'étude de rentabilité**

Cependant, dans de telles situations, on dépend des informations imaginées sur la cible a priori, qui ne concernent pas, en général, toutes les variables qui pourraient nous intéresser (il sera difficile d'évaluer la proportion de garanties souscrites). De plus, elles seront également très souvent partielles : nous aurons un capital emprunté moyen par exemple. Enfin, nous n'aurons pas d'informations relatives à la dépendance entre variables. Cela conduisait, jusqu'à ce mémoire, à supposer l'indépendance entre les variables, ce qui est en faux en pratique, comme nous le verrons après. Il est par exemple intuitif que, plus on est âgé, moins la durée d'emprunt aura tendance à être élevée.

Ces diverses approximations biaiseront forcément notre calcul de la rentabilité, et ce pour différentes raisons :

- La non-linéarité du tarif et de la sinistralité sur certains critères : âge, durée de prêt

- La vision approximative des provisions de primes qui sont mal résumées par un profil moyen

Nous proposons donc, dans la suite, une méthode permettant de prendre en compte la dépendance entre les variables, tout en permettant de changer facilement les données concernant les variables.

Nous cherchons ainsi à simuler un portefeuille fictif, c'est-à-dire à générer un nombre donné d'individus et l'observation de variables associées à ceux-ci, tout en prenant en compte la structure de dépendance entre ces variables, ainsi qu'en permettant, de manière plus ou moins libre, de moduler les lois marginales de celles-ci.

Avant de s'intéresser à la simulation proprement dite, nous aborderons tout d'abord le traitement des bases de données que nous avons à disposition et commencerons à étudier les caractéristiques du portefeuille, grâce à des statistiques descriptives.

De plus, la structure de dépendance entre les variables qui sera utilisée par la suite est déterminée sur le portefeuille construit lors de cette étape de traitement des données.

### **3.2. Traitement des bases de données**

Nous cherchons à créer une table de données qui regrouperaient certaines variables telle que les caractéristiques de notre portefeuille fictif simulé d'assurés seraient équivalentes à la simulation de données ayant la même structure que cette table.

Les données utilisées sont issues des bases de données de gestion d'APRIL Santé Prévoyance, à la date d'arrêté du 30 juin 2017.

- **Périmètre produit**

Nous n'avons pas d'a priori sur le périmètre produit, malgré les différences de segmentation entre ceux-ci comme la présence ou non d'une majoration Fumeur, d'une minoration couple, etc.

- **Périmètre temporel**

Au cours du temps, le marché de l'assurance emprunteur a connu des changements concernant les profils des assurés. Ce marché est directement impacté par le marché de l'emprunt, et notamment, de l'emprunt immobilier, qui est lui-même influence par la variation des taux d'intérêts et des prix immobiliers.

De plus, le portefeuille d'assurés APRIL que nous pouvons observer a également évolué. Les générations de produits qui se sont suivis n'ont pas ciblé les mêmes clients, les mêmes segmentations et même, les mêmes tarifs. Cependant, ce qui nous intéresse, est de pouvoir simuler un portefeuille moderne, c'est-à-dire qui reproduirait le profil des assurés récent.

A cause de cela, nous nous limités à un périmètre temporel récent.

Nous ne sommes pas en mesure de créer une variable « Année de production » car la variable « Date de création », qui se situe dans les bases Affaires, n'est pas fiable. Elle enregistre en réalité la date de la dernière modification sur l'affaire, ce qui se passe parfois après la date d'effet.

Notre filtre se basera donc sur la date d'effet, en plus de la date de création de l'affaire.

Périmètre temporel : année(date d'effet)  $\geq$  2015 et année(date de création affaire)  $\geq$  2015.

- **Filtre sur le motif de fin**

Le motif de fin correspond au motif de fin de l'affaire. Nous filtrons toutes les affaires ayant ce motif renseigné comme annulé, ainsi que les affaires en attente. Cela signifie que les affaires résiliées pour toutes causes sont conservées, ainsi que les affaires en cours.

- **Traitement des modifications en cours de prêt**

Nous possédons les informations concernant les modifications effectuées au cours de l'emprunt au capital, à la durée ou au taux d'intérêt (renégociation du prêt). Nous cherchons cependant à générer un portefeuille à la souscription, nous ne prendrons donc pas en compte ces informations.

- **Construction de la base de données « Affaires »**

Nous construisons dans un premier temps une base « Affaires » ayant comme variables :

- Le code affaire
- L'âge à la souscription
- Le sexe
- L'indicatrice de la présence d'un co-emprunteur (Non Renseigné pour certains produits, suivant si le produit prenait en compte cette information ou non) qui indique si l'affaire possède plus d'un emprunteur ou non
- Le libellé de la catégorie professionnelle
- L'état Fumeur ou Non-Fumeur (Non Renseigné pour certains produits)
- Le zonage géographique, au sens du produit ADP APRIL 2017
- Le type de prêt
- Le type de taux
- Le capital emprunté initial
- La quotité Décès/PTIA
- La quotité Arrêt de Travail
- La durée d'emprunt initiale
- Le taux d'emprunt initial

Elle possède 155 182 lignes.

Sur celles-ci :

4 311 lignes n'ont pas de garantie Décès renseignée car ce sont des affaires futures (i.e. leur date d'effet est postérieure à la date d'arrêté).

323 lignes n'ont pas de statut Fumeur renseigné.

- **Retraitement concernant l'indicatrice de présence de co-emprunteur(s)**

L'information concernant l'indicatrice de présence de co-emprunteur(s) est libellée « Non Renseigné » pour 2 276 lignes.

Certains produits n'ont pas de segmentation en cas de présence d'un ou de plusieurs co-emprunteurs. Ce critère n'étant pas discriminant, il n'est pas renseigné.

Pour les premiers produits incluant cette discrimination, le paramétrage n'a pas prévu de moyen permettant de renseigner l'information. Dans ce cas-ci, nous reconstruisons cette

information en attribuant la valeur « Oui » pour les codes adhérents où nous comptons plus d'une personne.

Finalement, le nombre de lignes pour lesquelles où la valeur de cette variable est égale à « Non renseigné » est de 1 727 lignes.

- **Construction de la base de données « Projet »**

A partir de cette base « Affaires », nous constituons notre base d'études, pour laquelle nous décidons de nous placer à une vision « projet ».

Nous définissons un projet comme étant une succession de prêts dont les dates d'effets et les dates de fin se succèdent et pour lesquels la date de création de l'affaire est la même. Par exemple, si un emprunteur A et son co-emprunteur B souscrivent des assurances pour les prêts suivants à la même date de création de l'affaire :

- Prêt à taux zéro, capital emprunté 1, durée d'emprunt 1, taux d'emprunt 1 et date de début : 01/01/2015 ; date de fin : 31/03/2016
- Prêt classique, capital emprunté 2, durée d'emprunt 2, taux d'emprunt 2 et date de début : 01/04/2016 ; ...

Alors, nous considérerons un projet pour chaque assuré où nous nous intéresserons aux informations suivantes :

Projet 1 : Informations de l'assuré A ; Somme des capitaux assurés ; Maximum de la durée d'emprunt ; Maximum du taux d'emprunt.

Projet 2 : Informations de l'assuré B ; Somme des capitaux assurés ; Maximum de la durée d'emprunt ; Maximum du taux d'emprunt.

- **Retraitements divers**

Certains retraitements arbitraires sont ensuite effectués :

- Nous pouvons observer un statut marital ou un statut Fumeur différent pour des prêts d'un même projet. Ceci peut être dû à une erreur de gestion. Nous imposons alors un statut indiquant la présence d'un co-emprunteur « Oui » et un statut Fumeur « Oui ».
- Nous nous intéressons à l'âge à la souscription le plus jeune.

- Concernant les informations liées au prêt, nous étudierons la somme des capitaux assurés en Décès et en Arrêt de travail de manière séparée. Nous prenons le maximum des taux d'intérêt observés ainsi que le maximum des durées d'emprunts observées sur les prêts du projet.

### **3.3. Analyse des caractéristiques du portefeuille actuel**

Dans cette partie, l'objectif est de présenter des statistiques permettant de mieux comprendre les variables en jeu et les dépendances entre elles.

Les différentes mesures de dépendance utilisées sont différenciées par la nature des variables auxquelles on les applique. Les natures des variables sont les suivantes :

- Quantitatives : les variables reflètent une notion de grandeur (âge, capital, durée, taux, ...). Elles peuvent être discrètes ou continues.
- Qualitatives : les variables reflètent des informations que l'on appréhende par des modalités. Nous pouvons également sous-diviser ce type de variable suivant la distinction suivante :
  - Catégorielles : les modalités de ces variables n'ont pas de sens quantitatif, il n'y a aucune relation d'ordre entre celles-ci. C'est le cas par exemple du sexe, de la CSP, de la situation géographique, etc. Ces variables peuvent résulter de la transformation d'une variable quantitative par un découpage en classes. Il existe alors une variable quantitative sous-jacente.
  - Ordinales : il existe une relation d'ordre entre les modalités de ces variables. Cependant, on ne peut pas quantifier la différence entre deux modalités successives. Ces variables peuvent résulter de la transformation d'une variable quantitative par attribution de rangs aux différentes valeurs de la distribution de la variable.

Les variables que nous étudions sont de différents types :

- Quantitatives continues : Capital, Taux.
- Quantitatives discrètes : Âge, Durée.
- Qualitatives catégorielles (à 2 modalités ou plus) : Sexe, Présence d'un co-emprunteur, Catégorie professionnelle, Zone géographique, Type de prêt, Type de taux, Garantie Arrêt de Travail.

La variable Taux est une variable particulière et sa modélisation nécessiterait une étude plus approfondie (modèle de taux, ...). Néanmoins, nous l'incluons dans un premier temps dans notre étude descriptive afin de voir quelles informations elle peut potentiellement apporter.

### **3.3.1. Variables caractérisant les assurés**

- **Statistiques descriptives unidimensionnelles**

L'individu moyen est résumé dans le tableau qui suit par les moyennes des variables quantitatives et le poids de la modalité la plus fréquente pour les variables qualitatives.

<b>Homme</b>	<b>Age moyen à la souscription</b>	<b>Avec co-emprunteur</b>	<b>Pas de réduction géo</b>	<b>Cadre</b>
55%	41 ans	75%	52%	55%

**Tableau 4 - « Photographie » de l'assuré moyen**

De ces données, la seule variable quantitative est la variable « Age à la souscription », dont nous calculons les quartiles.

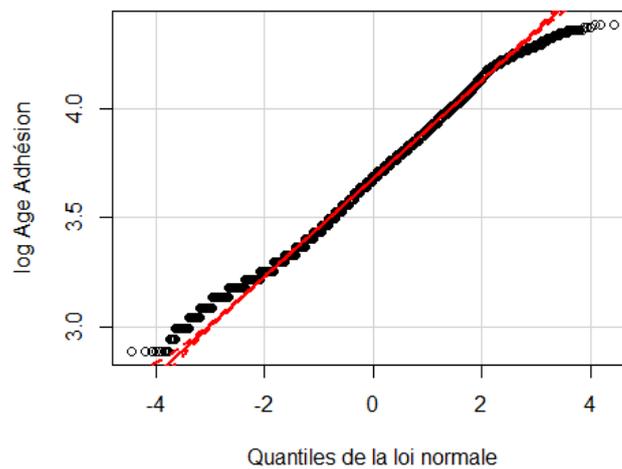
<i>Résumé</i>	<b>Age souscription</b>
Min.:	18 ans
1st Qu.:	34 ans
Median:	40 ans
Mean:	41 ans
3rd Qu.:	46 ans
Max.:	83 ans

**Tableau 5 - Description des variables quantitatives**

- **Ajustement de la loi marginale de l'âge à une loi paramétrique :**

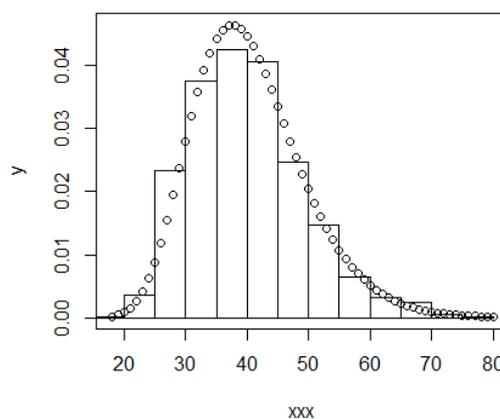
Nous considérons dans un premier temps la variable Age

D'après la forme de l'histogramme de la variable, ainsi que de la comparaison par le moyen d'un diagramme quantile-quantile, nous affirmons que la variable Age suit une loi log-normale.



**Figure 1-Diagramme Quantile-Quantile log(Age à l'adhésion)**

Nous estimons les paramètres des lois log-normales pour cette variable par la méthode du maximum de vraisemblance. Une vérification graphique nous montre alors que l'ajustement est satisfaisant.



**Figure 2-Ajustement log-normal sur l'âge**

### **3.3.2. Variable caractérisant les prêts**

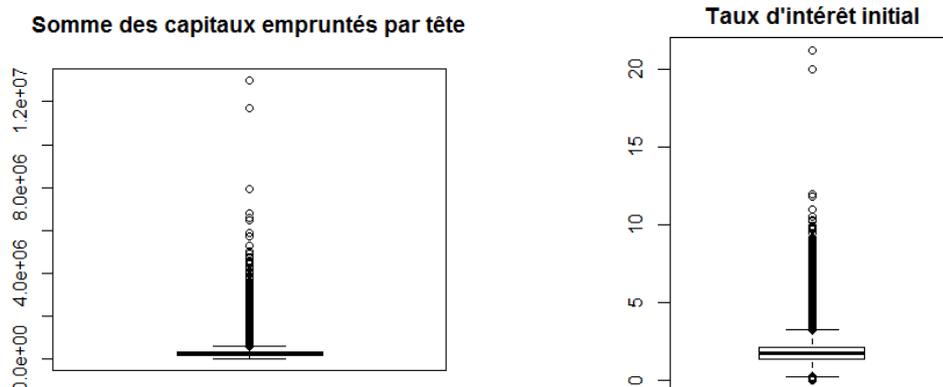
- **Statistiques descriptives unidimensionnelles**

Capital emprunté moyen	Durée d'emprunt moyen	Taux d'emprunt moyen
274 842,50 €	201 mois	1,84%

Résumé	Durée	Taux d'intérêt initial	$\Sigma$ capitaux empruntés
Min.:	3 mois	0,00%	3 700,00 €
1st Qu.:	144 mois	1,38%	140 500,00 €
Median:	204 mois	1,75%	210 000,00 €
Mean:	201 mois	1,84%	274 842,00 €
3rd Qu.:	240 mois	2,14%	321 869,00 €
Max.:	446 mois	15,2%	13 000 000,00 €

Les variables « Taux d'intérêt initial » et «  $\Sigma$  capitaux empruntés » présentent des valeurs extrêmes.

Les taux d'emprunts nuls concernent les prêts à taux zéro alors que les taux d'emprunts élevés concernent principalement les prêts à la consommation. De même, les capitaux empruntés faibles concernent majoritairement des prêts à la consommation. Les capitaux empruntés élevés correspondent à des emprunts exceptionnels.

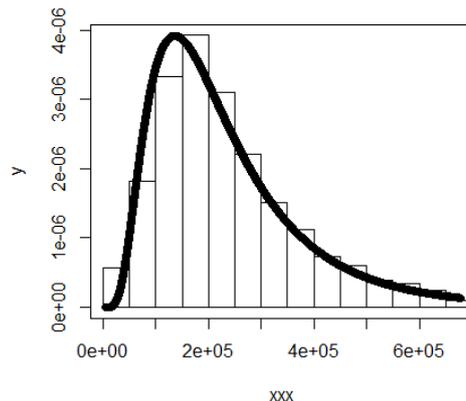


**Figure 3 - Boite à moustache des variables "Taux d'intérêt à l'adhésion" et "Capital emprunté initial"**

- **Ajustement de la loi marginale capital à une loi paramétrique**

Nous excluons de cette analyse les valeurs extrêmes de la variable Capital (i.e. nous gardons les valeurs inférieures ou égales au quantile à 95% de la variable).

D'après la forme de l'histogramme des variables, ainsi que de la comparaison par le moyen d'un diagramme quantile-quantile, nous affirmons que la variable Capital (hors valeurs extrêmes) suit une loi log-normale.



**Figure 4-Ajustement log-normal sur le capital (hors valeurs extrêmes)**

Concernant les variables Durée et Taux, il est plus compliqué d'ajuster une loi paramétrique.

- Les durées d'emprunt présentent une surpondération de certaines valeurs correspondant aux durées d'emprunts « classiques » : 10, 15, 20, 25 ans...
- Le taux d'emprunt dépend de différents critères dont la période de souscription du prêt mais également de caractéristiques de l'assuré. Il est malaisé de modéliser cette variable dont toute une théorie lui est consacrée. De plus, le taux d'emprunt n'est pas une variable fondamentale dans l'appréciation de la rentabilité. Nous proposerons néanmoins une approximation de cette variable dans la suite.

### **3.3.3. Variables concernant les garanties de l'assurance de prêt**

Certaines variables concernent les garanties souscrites par les assurés :

- Les quotités souscrites pour la garantie Décès et pour la garantie Arrêt de Travail,
- La proportion de souscription aux options de rachats des exclusions des pathologiques dorsales et/ou des pathologies psychologiques,

- La proportion de souscription à l'option d'Invalidité Permanente Partielle.

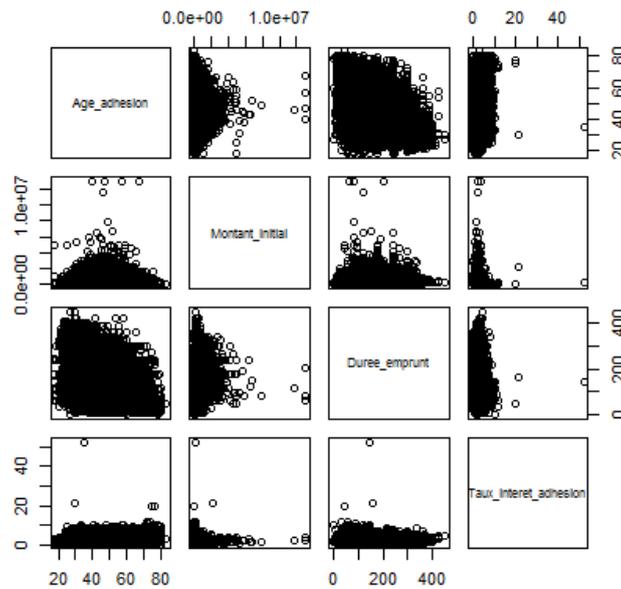
Il s'agit de variables binaires pour les indicatrices de garantie et de variables quantitatives valant entre 0 (pas de souscription à la garantie) et 1 pour les quotités.

### **3.4. Statistiques bidimensionnelles et indicateurs de corrélations**

Afin de compléter notre analyse du portefeuille actuel, nous nous intéressons maintenant aux statistiques bidimensionnelles, ce qui nous permettra d'appréhender la structure de dépendance entre les variables que nous étudions. Pour cela, nous utiliserons des statistiques usuelles (dans le sens où elles sont fréquemment utilisées dans des études de ce type, et en particulier, chez ASP), pour la raison qu'elles sont facilement calculables et compréhensibles, en plus de fournir des illustrations parlantes. Néanmoins, nous verrons que ces différents indicateurs ne sont pas les plus adaptés aux méthodes de simulation de variables multivariées et de leur structure de dépendance que nous pouvons trouver dans la littérature, méthodes notamment utilisées dans le domaine de la biostatistique. Cela nous conduira à présenter dans un second temps d'autres indicateurs de dépendance, pour ensuite introduire les méthodes de simulation que nous avons retenues afin de résoudre notre problématique.

- **Matrice de nuage de points**

Parmi les variables auxquelles nous nous intéressons, il y a en particulier quatre variables quantitatives : l'âge à la souscription, le montant d'emprunt initial, la durée d'emprunt et le taux d'intérêt à la souscription. Le dessin des nuages de point des variables deux à deux nous permet d'avoir une idée des dépendances qui existent entre elles.



**Figure 5-Matrice de nuage de points**

**Analyse :** Nous observons que les valeurs extrêmes des variables « capital » et « taux » déforment ces graphiques. De même, la variable « taux » possède une masse en 0 provenant des taux à prêt zéro qui déforme les nuages de points.

Il ne semble pas se dégager de structure de corrélation très nette.

- **Coefficients de corrélation de Pearson**

Nous vérifions les premières conclusions en calculant les coefficients de corrélation par la méthode de Pearson, en supprimant préalablement les valeurs extrêmes des variables « Capital » et « Taux » par un filtre grossier sur les valeurs supérieures au 95<sup>ème</sup>-percentile.

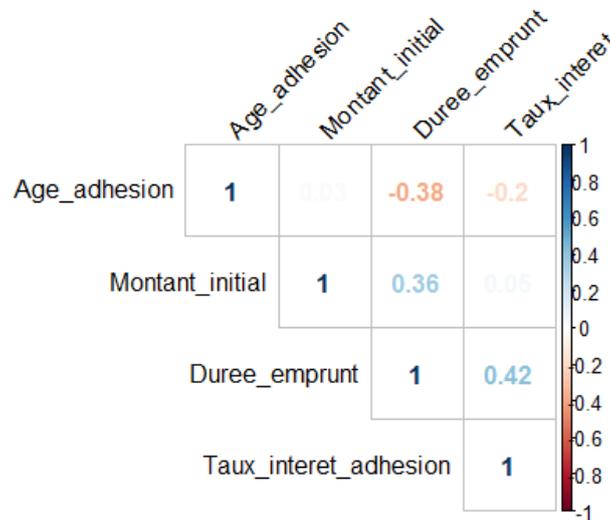
Le coefficient de Pearson permet de mesurer la corrélation linéaire entre deux variables x et y dont la version empirique est donnée par :

$$r_{xy} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_x s_y}$$

Où  $s_x$  et  $s_y$  sont les écarts-types de x et y. Le numérateur correspond à la covariance observée. Le coefficient de Pearson n'est cependant pas robuste car il est très sensible aux valeurs extrêmes des variables qui influencent la moyenne. De plus, il ne définit complètement la structure de dépendance entre des variables aléatoires que dans certains

cas particuliers, par exemple quand la distribution est une distribution multidimensionnelle normale.

Les résultats obtenus sont les suivants :



**Figure 6 - Matrice de corrélation de Pearson**

Analyse : Les corrélations positives entre le montant emprunté initial et la durée d'emprunt, ainsi qu'entre la durée d'emprunt et le taux d'intérêt sont cohérents : plus la durée d'emprunt est longue, plus nous observons de capital emprunté élevé ainsi que de taux d'intérêt élevé. De même, la corrélation négative entre la durée d'emprunt et l'âge à l'adhésion a du sens. Néanmoins, l'absence de corrélation entre l'âge à l'adhésion et le montant emprunté n'est pas ce à quoi nous nous attendons, et c'est pourquoi nous analysons maintenant les coefficients de corrélations partielles.

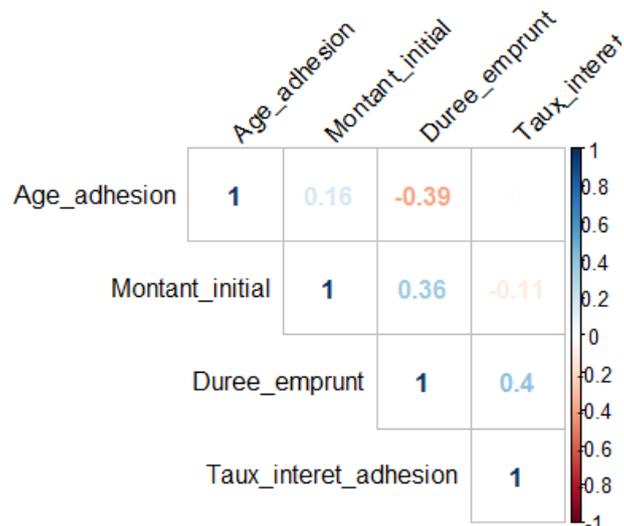
- **Coefficients de corrélations partielles**

Nous pouvons imaginer que lorsque nous calculons les coefficients de Pearson sur ces variables prises deux à deux, une troisième variable a un lien avec les deux variables d'intérêt. Dans ce cas, l'utilisation du coefficient de Pearson pour déterminer l'existence d'une relation numérique entre les variables n'est pas adaptée. On préférera utiliser les corrélations partielles qui mesurent l'association entre deux variables aléatoires après avoir neutralisé l'effet de toutes les autres variables aléatoires.

Ainsi, la corrélation partielle entre x et y sachant z est donnée par la formule suivante :

$$r_{xy|z} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{1 - r_{xz}^2} \sqrt{1 - r_{yz}^2}}$$

Voici les résultats obtenus :

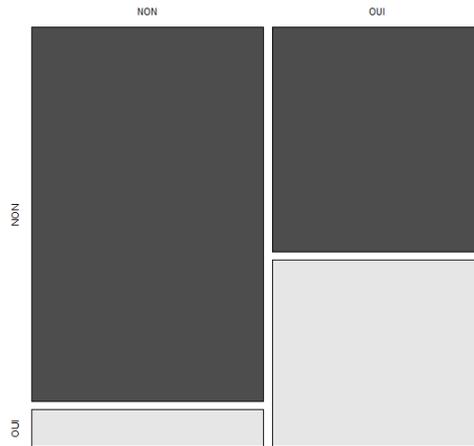


**Figure 7 - Matrice de corrélation partielle**

**Analyse :** Les coefficients obtenus nous semblent plus cohérents : Plus on est âgé, plus on est capable d'emprunter des capitaux élevés et moins la durée d'emprunt sera élevée. De même, plus on emprunte sur une durée longue, plus le capital emprunté a tendance à être important et moins les taux seront avantageux.

- **Graphes en mosaïque**

Concernant les variables qualitatives, nous visualisons la structure de dépendance au moyen de graphes en mosaïque.



**Figure 8 - Souscription IPP x Souscription Option de Rachat**

Cela nous permet d'observer qu'il semble plus probable de souscrire la garantie IPP en même temps qu'une option de rachat d'exclusions.

- **V de Cramer**

Pour la variable « Catégorie professionnelle », le graphe en mosaïque est difficilement lisible en raison du nombre de modalités de la variable. Afin de vérifier nos observations et de faciliter l'interprétation des résultats concernant la variable « Catégorie professionnelle », nous utilisons une mesure d'association entre deux variables qualitatives : le V de Cramer.

Le V de Cramer permet d'obtenir une mesure d'association comprise entre 0 (indépendance) et 1 (liaison fonctionnelle).

$$V = \sqrt{\frac{\chi^2}{\chi_{max}^2}} = \sqrt{\frac{\chi^2}{n[\min(l, c) - 1]}}$$

Où :  $\chi^2$  correspond à la statistique associée au test du Khi-deux.

n correspond au nombre total d'observations.

l correspond au nombre de lignes du tableau de contingence entre les deux variables qualitatives.

c correspond au nombre de colonnes du tableau de contingence entre les deux variables qualitatives.

Nous choisissons de créer des variables qualitatives à partir de la transformation des variables quantitatives par découpage en classes, afin de les introduire dans cette analyse. Pour cela, nous divisons en 4 classes d'effectifs égales par la méthode des quartiles les variables quantitatives Age, Durée, Capital et Taux.

Nous pouvons résumer les V de Cramer calculés entre les différentes variables dans le tableau suivant :

	Fumeur	Rabais couple	Sexe	Cat Pro	AT	Zone géo	Age	Durée	Capital	Taux
Fumeur	1	0,0	0,0	0,1	0,0	0,0	0,1	0,0	0,0	0,1
Rabais couple		1	0,1	0,2	0,2	0,0	0,2	0,1	0,2	0,0
Sexe			1	0,2	0,1	0,0	0,1	0,1	0,0	0,0
Cat Pro				1	0,4	0,2	0,2	0,2	0,2	0,1
AT					1	0,1	0,3	0,2	0,2	0,0
Zone géo						1	0,1	0,1	0,2	0,1
Age							1	0,2	0,1	0,1
Durée								1	0,2	0,2
Capital									1	0,0
Taux										1

**Tableau 6 - Matrice des V de Cramer**

Les associations mesurées semblent relativement faibles. Les variables Catégorie Professionnelle et Age semblent a priori être les variables ayant le plus d'impact.

Notons que cette méthode induit une perte de données, du fait de la non prise en compte de la variable quantitative sous-jacente aux variables « Âge à la souscription », « Montant de capital emprunté », « Durée d'emprunt » et « Taux d'emprunt ».

### **3.5. Présentation de méthodes de simulation de portefeuille fictif d'assurés adaptées à notre problématique**

L'objectif est de simuler un portefeuille d'individus, où les variables générées respecteront la structure de dépendance observée sur le portefeuille actuel. Les variables que nous cherchons à générer sont de structures différentes : quantitatives (continues et discrètes) et

qualitatives (nominales à 2 modalités ou à plus de 2 modalités). Ceci implique des moyens de mesurer la dépendance différents, mais également des méthodes de simulation différentes.

Dans cette partie, nous présenterons dans un premier temps certaines notions théoriques sur les mesures de dépendance qui seront utilisées par la suite dans la méthode retenue pour la simulation du portefeuille fictif d'assurés. Puis, nous verrons une méthode de simulation à partir des distributions conditionnelles dont la méthodologie à partir des model points présentée en introduction de cette partie est une dérivée. Enfin, nous présenterons la méthode que nous avons choisie pour notre outil de génération du portefeuille fictif d'assuré, qui est une méthode NORTA modifiée et que nous pouvons retrouver en particulier dans la littérature du domaine de la biostatistique, notamment avec le docteur Hakan DEMIRTAS.

### **3.5.1. Résumé théorique des différentes mesures de la dépendance entre variables de différentes structures**

Les indicateurs de dépendance que nous avons calculés auparavant nous donnent des illustrations des interactions entre les différentes variables. Néanmoins, en l'état, ils ne peuvent pas être intégrés dans une méthode de simulation reproduisant ces variables ainsi que leur structure de dépendance, notamment dans les méthodes utilisées dans le domaine de la biostatistique et dont la littérature est abondante. La partie qui suit est donc une partie préalable à celle où nous présentons deux méthodes de simulation de variables et de leur structure de dépendance.

Aucune des variables étudiées ne peut être considérée comme une variable ordinale. Nous ne présenterons donc pas de mesures pour celles-ci.

Les méthodes de mesure de la dépendance varient selon la nature des variables que l'on considère, c'est-à-dire si elles sont quantitatives ou qualitatives (binaires, ordinales, ...). Une explication plus détaillée est faite par [SAPORTA \[2006\]](#).

Pour deux variables quantitatives  $x$  et  $y$ , il est courant d'utiliser le coefficient de corrélation de Bravais-Pearson, définie par :

$$\rho(x, y) = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y}$$

Dont la version empirique est :

$$r = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_x s_y}$$

Où  $s_x$  et  $s_y$  sont les écarts-types de  $x$  et  $y$ . Le numérateur correspond à la covariance observée.

Le coefficient de corrélation de Pearson permet de mesurer la corrélation linéaire entre deux variables quantitatives. Il n'est cependant pas robuste car il est très sensible aux valeurs extrêmes des variables. De plus, le coefficient de Pearson ne définit complètement la structure de dépendance entre des variables aléatoires que dans certains cas particuliers, par exemple quand la distribution est une distribution multidimensionnelle normale.

Des cas particuliers du coefficient de corrélation de Pearson sont :

- La corrélation bisériale de point, que l'on rencontre lorsqu'une des deux variables est binaire et l'autre est continue. En attribuant des valeurs numériques distinctes à la variable binaire, on retrouve le coefficient de Pearson.
- Le coefficient Phi, qui mesure la corrélation entre deux variables binaires.

Dans le cas où l'on cherche à mesurer la dépendance entre deux variables qualitatives à plus de deux modalités, il n'est plus possible de considérer cette formulation. On utilise alors couramment le  $\chi^2$  de Pearson. Celui-ci mesure la distance entre la table de contingence observée et la table de contingence qu'on obtiendrait sous hypothèse d'indépendance.

Pour mesurer la dépendance entre une variable quantitative et une variable qualitative, on utilise le rapport de corrélation  $\eta_{Y|X}^2$  tel que :

$$\eta_{Y|X}^2 = \frac{V(E(Y|X))}{V(Y)}$$

Il est maximal si  $Y$  est lié fonctionnellement à  $X$ . Son interprétation est similaire à celle du coefficient de Pearson.

$\eta_{Y|X}^2$  peut être appliqué aussi bien lorsque  $X$  et  $Y$  sont quantitatives que lorsque  $Y$  est quantitative et  $X$  est qualitative à  $k$  modalités.

Enfin, dans le cadre de la méthode de simulation de portefeuille fictif d'assurés, nous présentons la corrélation latente qui est utilisée dans le cadre de variables binaires ou ordinales. Celle-ci possède une autre appellation considérée comme imprécise et obsolète

aujourd'hui qui est le coefficient tétrachorique de corrélation (ou coefficient polychorique dans le cas de variables à plus de deux modalités).

Dans le cas de deux variables dichotomiques, nous avons vu que le coefficient Phi était le coefficient usuel. Toutefois, si nous pouvons faire l'hypothèse que ces deux variables sont issues de la dichotomisation de deux variables inobservables, continues et normalement distribuées, alors il est plus correct d'utiliser la corrélation latente qui correspond à l'estimateur du maximum de vraisemblance de la corrélation entre ces deux variables inobservables.

Ceci est le cas lorsqu'on a, de manière artificielle, dichotomisé deux variables normales X et Y. Par exemple, on peut mesurer la taille d'un individu en centimètres et la traiter en variable continue, mais on peut également la dichotomiser artificiellement en classant les individus en « petits » ou en « grands ».

Dans la même famille, la corrélation bisériale est une corrélation entre une variable continue et une variable dont on suppose qu'elle est issue de la dichotomisation d'une variable normale. C'est un cas particulier de la corrélation polysériale qui est la corrélation latente inférée entre une variable continue et une variable ordinale.

Dans notre situation, nous avons les variables Sexe, Indicatrice Co-emprunteur(s), Indicatrice Arrêt de travail et Fumeur qui sont des variables catégorielles à deux modalités. Nous pouvons également transformer en variables catégorielles à deux modalités les variables Catégorie Professionnelle (Cadre/Non cadre) et Zone géographique (Réduction/Pas de réduction). Néanmoins, aucune de ces variables n'est issue d'une variable continue normalement distribuée qu'on ne pourrait pas observer.

### **3.5.2. Méthode des distributions conditionnelles**

Supposons qu'on connaît entièrement la distribution d'un vecteur aléatoire  $\mathbf{X}$ , i.e. pour  $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbf{R}^n$ , on connaît la valeur de la fonction de répartition (FdR) jointe :

$$F(\mathbf{x}) = P(\mathbf{X} \leq \mathbf{x}) = P(X_1 \leq x_1, \dots, X_n \leq x_n)$$

Ceci détermine toutes les corrélations.

Notons :

- Pour  $i = 1, \dots, n$ ,  $F_i(\bullet)$  la distribution marginale de  $X_i$
- Pour  $k = 2, \dots, n$ ,  $F_k(\bullet | X_1, \dots, X_{k-1})$  la distribution conditionnelle de  $X_k$  sachant  $X_1, \dots, X_{k-1}$

Alors, un algorithme pour générer  $\mathbf{X}$  est :

- 1) Générer  $X_1$  à partir de  $F_1$
- 2) Générer  $X_2$  à partir de  $F_2(\bullet | X_1)$
- 3) Générer  $X_3$  à partir de  $F_3(\bullet | X_1, X_2)$
- 
- 
- 
- n) Générer  $X_n$  à partir de  $F_n(\bullet | X_1, X_2, \dots, X_{n-1})$
- n+1) Retourner  $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$

Cette méthode très générale permet de générer entièrement la distribution du vecteur aléatoire  $\mathbf{X}$ . Elle est adaptée dans la situation où on désire générer un portefeuille ayant les mêmes caractéristiques que le portefeuille observée. Cependant, si nous désirons modifier les caractéristiques d'une marginale (par exemple : la répartition par âge, la proportion de cadres, ...) ou celles de la structure de dépendance, il devient alors nécessaire de reconstruire l'intégralité des données en entrées. Cette reconstruction est de plus impossible si nous ne possédons pas déjà un portefeuille ayant exactement les mêmes marginales ou la même structure de dépendance.

Cependant, nous pouvons imaginer une adaptation de cette méthode par le biais d'une méthode de classification.

### **3.5.3. Méthode retenue pour la simulation du portefeuille fictif d'assurés à partir de la spécification des marginales et de la structure de dépendance**

Nous cherchons à simuler une variable aléatoire multidimensionnelle dont on connaît les marginales, en capturant la dépendance. Le moyen habituel pour cela est de sélectionner une copule qui détermine la structure de dépendance entre ces marginales. Toutefois, si les dimensions sont élevées, il est plus confortable de choisir la copule normale.

### ► Cas gaussien

Dans ce cas particulier, il est équivalent de connaître les distributions marginales ainsi que toutes les covariances/corrélations ou de connaître entièrement la distribution jointe.

Notons :

4.  $\mu = (\mu_1, \dots, \mu_n)^T$  la moyenne du vecteur  $\mathbf{X}$
5.  $\Sigma = [\sigma_{ij}]_{n \times n}$  la matrice de covariance du vecteur  $\mathbf{X}$ , définie positive.

D'après le théorème de la décomposition de Cholesky, pour  $\Sigma$  matrice définie positive, il existe une unique matrice  $T$  triangulaire inférieure telle que :  $\Sigma = TT^T$ . La démonstration est rappelée en annexe.

L'algorithme pour générer  $\mathbf{X}$  est alors :

- 1) Générer  $Z_1, \dots, Z_n$  variables Indépendantes et Identiquement Distribuées suivant une loi normale centrée et réduite. On note  $Z = (Z_1, \dots, Z_n)^T$ .
- 2) Retourner  $X = \mu + TZ$

### ► Cas non gaussien

#### ▪ Variables continues

Dans le cas où l'on cherche à simuler des variables corrélées non gaussiennes, une méthode populaire est la méthode NORTA, introduite par [CARIO & NELSON \[1997\]](#).

L'hypothèse prise par la méthode NORTA est que la copule appropriée pour  $\mathbf{X}$  est proche d'une copule gaussienne, copule présentant une indépendance des queues.

Il est bon de remarquer qu'il n'y a pas de théorie statistique supportant l'utilisation d'une distribution NORTA. En réalité, il y a même peu de chance pour que la distribution sous-jacente soit de type NORTA. La philosophie de cette approche repose en fait sur la volonté de correspondre au mieux aux distributions marginales et à la matrice de corrélation plutôt que de chercher à considérer le véritable modèle, en supposant qu'un tel modèle existe.

Le modèle tel que décrit par [CARIO & NELSON \[1997\]](#) s'écrit de la manière qui suit.

Le but est de définir un vecteur aléatoire  $\mathbf{X}$  ayant les propriétés suivantes :

6.  $X_i \sim F_i$ ,  $i \in \llbracket 1, n \rrbracket$ , où  $F_i$  est une fonction de répartition arbitraire
7.  $\text{Corr}(X) = \Sigma_X$  où  $\Sigma_X$  est donnée

Nous représentons  $\mathbf{X}$  comme la transformation d'un vecteur normal multidimensionnel centré et réduit  $\mathbf{Z} = (Z_1, \dots, Z_n)^T$  avec une matrice de corrélation  $\Sigma_Z$ . Le vecteur NORTA  $\mathbf{X}$  est :

$$\mathbf{X} = \begin{pmatrix} F_{X_1}^{-1}[\Phi(Z_1)] \\ F_{X_2}^{-1}[\Phi(Z_2)] \\ \vdots \\ F_{X_k}^{-1}[\Phi(Z_k)] \end{pmatrix}$$

où  $\Phi$  est la FdR de la loi unidimensionnelle normale centrée et réduite et  $F_X^{-1}(u) = \inf\{x : F_X(x) \geq u\}$  est la fonction inverse de la FdR de  $\mathbf{X}$ .

La transformation  $F_{X_i}^{-1}[\Phi(\bullet)]$  assure que les  $X_i$  possède la FdR  $F_{X_i}$  désirée.

Le problème devient alors de choisir la matrice de corrélation  $\Sigma_Z$  qui donne la matrice de corrélation  $\Sigma_X$  voulue.

Une résolution numérique est généralement effectuée pour cela grâce aux relations entre  $\Sigma_X$  et  $\Sigma_Z$  explicitées par [CARIO ET NELSON \[1997\]](#). Il est cependant intéressant de noter que dans le cas de marginales continues,  $\Sigma_X \approx \Sigma_Z$  alors que dans le cas de marginales discrètes, ces deux matrices sont significativement différentes.

Enfin, l'algorithme proposée et implémentée dans la librairie `SimCorMultRes` par [TOULOU MIS \[2016\]](#) est le suivant :

- i) Déterminer une factorisation de Cholesky de  $\Sigma_Z$  tel que  $MM' = \Sigma_Z$  où  $M$  est une matrice triangulaire inférieure.
- ii) Générer  $W = (W_1, \dots, W_k)^T$  un vecteur aléatoire dont les composantes sont des variables de loi normale standard et i.i.d.
- iii) Calculer  $Z \leftarrow MW$ .
- iv) Retourner  $\mathbf{X}$  où  $X_i \leftarrow F_{X_i}^{-1}[\Phi(Z_i)]$ ,  $i = 1, \dots, k$ .
- v) Aller à l'étape ii).

Lorsqu'une variable discrète présente suffisamment de valeurs, nous la considérerons comme une variable continue, quitte à prendre l'arrondi à l'entier le plus proche de la donnée simulée *a posteriori*. Ainsi, nous considérons les variables « Age à l'adhésion » et « Durée d'emprunt » comme des variables continues pour lesquelles nous pouvons appliquer telle quelle la méthode NORTA.

- **Variables binaires**

La simulation de variables binaires corrélées est notamment explorée par [LEISCH & WEINGESSEL \[1998\]](#).

La distribution d'une variable binaire est un cas particulier d'une distribution discrète, ne pouvant prendre que deux valeurs (OUI/NON, Femme/Homme, etc.), que nous pouvons recoder par  $\{0 ; 1\}$ .

Pour A une variable binaire, nous notons  $p_A = P(A = 1)$  et  $q_A = 1 - p_A = P(A = 0)$ .

Un algorithme possible pour la simulation d'un vecteur binaire multidimensionnelle  $(A_1, \dots, A_n)$  est le suivant :

- i) Calculer  $\mu_i = \Phi^{-1}(p_{A_i})$ , où  $\Phi$  est la distribution normale standard.
- ii) Trouver une matrice de corrélation  $\Sigma_Z$  appropriée pour la distribution normale.
- iii) Simuler un échantillon d'un vecteur normal n-dimensionnel Z de moyenne  $\mu$  et de matrice de corrélation  $\Sigma_Z$ .
- iv) On retourne :  $a_i = 1$  si  $z_i > 0$  et  $a_i = 0$  si  $z_i < 0$ .

Nous utiliserons une version légèrement modifiée de cet algorithme afin de faciliter la génération simultanée de variables continues non normales telle que décrite précédemment.

Ainsi, nous n'effectuons pas l'étape i) et nous avons alors à l'étape iv) :  $a_i = 1$  si  $z_i \leq \Phi^{-1}(p_{A_i})$  et  $a_i = 0$  sinon.

#### ▪ Variables ordinales

La dichotomisation de la variable « Catégorie professionnelle » en deux modalités « Cadre » et « Non cadre » ne nous satisfait pas. Nous choisissons de garder trois modalités : « Cadre », « Profession libérale » et « Non cadre ».

La simulation d'une variable discrète  $X \in \{x_1, \dots, x_n\}$  avec  $F_X$  sa fonction de répartition est construite telle que :

$$\text{Si } Z < r_1 \rightarrow X = x_1$$

$$\text{Si } r_1 \leq Z \leq r_2 \rightarrow X = x_2$$

.

.

$$\text{Si } r_{n-1} \leq Z \rightarrow X = x_n$$

Où :  $r_i = \Phi^{-1}(F_X(x_i))$ ,  $\Phi$  étant la fonction de répartition de la loi normale standard.

## Algorithme

L'approche par la théorie des copules semble être la plus évidente. Néanmoins, il n'existe pas, à notre connaissance du moins, de copules multivariées conçues pour prendre en compte différents types de données simultanément. La méthode suivante peut néanmoins être vue comme une variante d'une méthode basée sur une copule multivariée gaussienne.

**Hypothèse importante :** Les variables qualitatives en jeu sont toutes catégoriques, i.e. elle possède deux ou plus catégories, mais il n'y a pas d'ordre intrinsèque aux catégories. Une variable ordinale est similaire à une variable catégorique, la différence étant donc qu'il y a un ordre clair aux différentes catégories d'une même variable. Dans la suite, nous assimilerons pourtant toutes les variables qualitatives à des variables ordinales, dont les longueurs des intervalles entre les valeurs sont toutes égales (ce qui permet des calculs statistiques). L'ordre que nous attribuons aux catégories sera arbitraire.

L'algorithme que nous proposons est conçu pour générer simultanément des variables binaires, ordinales et continues. Bien que la variable binaire soit un cas particulier d'une variable ordinale, nous présentons les étapes de manière séparée.

Le moteur de l'algorithme est de calculer la matrice de corrélation de la variable multivariée normale sous-jacente. Il suffira ensuite d'appliquer ce que nous avons vu précédemment sur la génération d'une variable normale multidimensionnelle.

Notons B, O, C et N les variables, respectivement, Binaires, Ordinales, Continues et Normales. Soit  $\Sigma$  la matrice de corrélation de Pearson. Sans perte de généralité, nous supposons que les variables sont ordonnées par type. Nous cherchons donc la matrice de corrélation,  $\Sigma_Z$ , du vecteur aléatoire multivarié normale sous-jacent.

$$\begin{pmatrix} BB & OB & CB \\ BO & OO & CB \\ BC & OC & CC \end{pmatrix}_{initial} \rightarrow \begin{pmatrix} BB & OB & CB \\ BO & OO & CB \\ BC & OC & CC \end{pmatrix}_Z$$

Les coefficients de cette matrice intermédiaire seront calculés paire par paire.

Les valeurs des paramètres sont désignées par :  $p$  pour les variables binaires et ordinales,  $(\mu, \sigma^2)$  pour les variables normales, F la fonction de répartition pour les variables continues. Ces données sont soit spécifiées soit estimés à partir d'un jeu de données qu'on cherche à imiter.

1. Vérifier que  $\Sigma$  est définie positive.
2. Trouver les bornes inférieures et supérieures de Fréchet pour toutes les paires par la méthode de classification de [DEMIRTAS & HEDEKER \[2011\]](#).

On trouve dans la littérature différents exemples qui montrent que les corrélations ne sont pas toujours comprises entre -1 et +1 dans la plupart des cas bivariés car

différentes bornes inférieures et/ou supérieures peuvent être imposées par les distributions marginales. Ces restrictions s'appliquent aussi bien aux variables discrètes qu'aux variables continues.

On note  $\pi(F, G)$  l'ensemble des fonctions de répartition  $H$  ayant pour fonctions de répartition marginales  $F$  et  $G$ . [HOEFFDING \[1940\]](#) et [FRECHET \[1951\]](#) ont prouvé que dans  $\pi(F, G)$ , il existe des FdR  $H_L$  et  $H_U$ , appelées bornes inférieures et bornes supérieures, ayant la corrélation minimum et la corrélation maximum. Pour tout  $(x, y) \in R^2$ ,  $H_L(x, y) = \max[F(x) + G(y) - 1, 0]$  et  $H_U(x, y) = \min[F(x), G(y)]$ . Pour tout  $H \in \pi(F, G)$  et pour tout  $(x, y) \in R^2$ ,  $H_L(x, y) \leq H(x, y) \leq H_U(x, y)$ . Si  $\delta_L, \delta_U$  et  $\delta$  désignent les coefficients de corrélation de Pearson pour  $H_L, H_U$  et  $H$ , respectivement, alors :  $\delta_L \leq \delta \leq \delta_U$ . Si  $V$  est uniforme dans  $[0, 1]$ , alors  $F^{-1}(V)$  et  $G^{-1}(V)$  sont comonotones,  $F^{-1}(V)$  et  $G^{-1}(1 - V)$  sont anti-comonotoniques.

En pratique, générer de manière indépendante  $X$  et  $Y$  avec un large nombre de points de données avant de les trier dans la même direction et dans la direction opposée donne une approximation des bornes, respectivement, inférieures et supérieures.

3. Pour les combinaisons B-B, trouver les corrélations tétrachoriques (avant la dichotomisation) étant donnée le coefficient phi (après la dichotomisation). Soient  $X_1$  et  $X_2$  des variables binaires telles que  $E(X_j) = p_j$  et  $Cor(X_1, X_2) = \delta_{12}$ , où  $p_j$  et  $\delta_{12}$  sont donnés. Soit  $\Phi[t_1, t_2, \rho_{12}]$  la FdR d'une variable aléatoire normale standard de coefficient de corrélation tétrachorique  $\rho_{12}$ .

Nous avons :  $\Phi[t_1, t_2, \rho_{12}] = \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} f(z_1, z_2, \rho_{12}) dz_1 dz_2$

Où  $f(z_1, z_2, \rho_{12}) = [2\pi(1 - \rho_{12}^2)]^{-1} \times \exp\left[-\frac{z_1^2 - 2\rho_{12}z_1z_2 + z_2^2}{2(1 - \rho_{12}^2)}\right]$ .

La relation entre  $\delta_{12}$  et  $\rho_{12}$  est donnée par l'équation suivante :

$$\Phi[z(p_1), z(p_2), \rho_{12}] = \delta_{12}(p_1q_1p_2q_2)^{\frac{1}{2}} + p_1p_2$$

Trouver  $\rho_{12}$  grâce à cette équation, où  $z(p_j)$  désigne le  $p_j^{ième}$  quantile de la distribution normale standard et  $q_j = 1 - p_j$ . Répéter cette étape pour toutes les paires B-B.

4. Pour les paires B-O et O-O, étant donné le coefficient phi ordinal (après la discrétisation), trouver la corrélation polychorique (avant la discrétisation).

On suppose que  $Z = (Z_1, Z_2) \sim N(0, \Delta_{z_1z_2})$ , où  $Z$  désigne la distribution normale standard de matrice de corrélation  $\Delta_{z_1z_2}$  de coefficients  $\delta_{z_1z_2}$ . Soit  $X = (X_1, X_2)$  la variable bivariée ordinaire correspondant à la discrétisation de la variable  $Z$  selon les quantiles normaux, étant données les proportions marginales, avec une matrice de

corrélation  $\Delta_{Z_1 Z_2}$ . La relation entre  $\delta_{Z_1 Z_2}$  et  $\delta_{X_1 X_2}$  peut être trouvée par l'algorithme suivant :

- a. Générer une variable normale standard bivariée de corrélation  $\delta_{Z_1 Z_2}^0$  où  $\delta_{Z_1 Z_2}^0 = \delta_{X_1 X_2}$  (Ici,  $\delta_{X_1 X_2}^0$  est la corrélation polychorique initiale).
  - b. Discrétiser  $Z_1$  et  $Z_2$ , en se basant sur les probabilités cumulées des distributions marginales  $F_1$  et  $F_2$  afin d'obtenir, respectivement,  $X_1$  et  $X_2$ .
  - c. Calculer  $\delta_{X_1 X_2}^1$  (Ici,  $\delta_{X_1 X_2}^1$  est le coefficient phi ordinal après la première itération).
  - d. Exécuter la boucle suivante : tant que  $|\delta_{X_1 X_2}^v - \delta_{X_1 X_2}| > \varepsilon$  et  $1 \leq v \leq v_{max}$  ( $v_{max}$  et  $\varepsilon$  sont, respectivement, le nombre maximum d'itérations et l'erreur absolue tolérée maximale, toutes les deux étant choisies par l'utilisateur) :
    - i. Modifier  $\delta_{Z_1 Z_2}^v$  par  $\delta_{Z_1 Z_2}^v = \delta_{Z_1 Z_2}^{v-1} g(v)$  où  $g(v) = \frac{\delta_{X_1 X_2}}{\delta_{X_1 X_2}^v}$ . Ici,  $g(v)$  sert de coefficient correcteur qui converge ultimement vers 1.
    - ii. Générer une variable normale bivariée avec  $\delta_{Z_1 Z_2}^v$  et calculer  $\delta_{X_1 X_2}^{v+1}$  après discrétisation.
5. Pour chaque combinaison B-N/O-N, trouver la corrélation bisériale/polysériale (avant discrétisation d'une des variables) étant donnée la corrélation de point bisériale/polysériale (post discrétisation) par la linéarité et la constance des arguments proposés par [DEMIRTAS & HEDEKER \[2016\]](#). Supposons que  $X$  et  $Y$  suivent une loi normale de corrélation  $\delta_{XY}$ . Sans perte de généralité, nous pouvons supposer que  $X$  et  $Y$  sont standardisées afin d'avoir une moyenne nulle et une variance unitaire. Soit  $X_D$  la variable binaire résultant de la séparation de  $X$ ,  $X_D = I(X \geq k)$ . Ainsi,  $E[X_D] = p$  et  $V[X_D] = pq$  où  $q = 1 - p$ . La corrélation entre  $X_D$  et  $X$ ,  $\delta_{X_D X}$ , peut être obtenue par :  $\delta_{X_D X} = \frac{Cov[X_D, X]}{\sqrt{V[X_D]V[X]}} = \frac{E[X_D X]}{\sqrt{pq}} = \frac{E[X|X \geq k]}{\sqrt{pq}}$ . Nous pouvons aussi exprimer la relation entre  $X$  et  $Y$  par la régression linéaire suivante :

$$Y = \delta_{XY}X + \varepsilon \quad (1)$$

Où  $\varepsilon$  est indépendant de  $X$  et  $Y$ , et suit une loi  $N(0, 1 - \delta_{XY}^2)$ . Si nous généralisons ceci à  $X$  et/ou  $Y$  non-normales (toutes deux centrées réduites), la même relation peut être admise identique à l'exception que la distribution de  $\varepsilon$  est une distribution non-normale. Tant que l'équation (1) est vérifiée, alors :

$$\begin{aligned} Cov[X_D, Y] &= Cov[X_D, \delta_{XY}X + \varepsilon] \\ &= Cov[X_D, \delta_{XY}X] + Cov[X_D, \varepsilon] \\ &= \delta_{XY}Cov[X_D, X] + Cov[X_D, \varepsilon] \end{aligned} \quad (2)$$

Comme  $\varepsilon$  est indépendant de  $X$ , il est aussi indépendant de n'importe quelle fonction déterministe de  $X$  telle que  $X_D$  et donc  $Cov[X_D, \varepsilon]$  sera égale à 0.

Comme  $E[X] = E[Y] = 0, V[X] = V[Y] = 1, Cov[X_D, Y] = \delta_{X_D Y} \sqrt{pq}$  et  $Cov[X, Y] = \delta_{XY}$ , nous parvenons à

$$\delta_{X_D Y} = \delta_{XY} \delta_{X_D X} \quad (3)$$

Dans le cas bivarié normal,  $\delta_{X_D X} = h/\sqrt{pq}$  où  $h$  est l'ordonnée de la courbe normale au point de dichotomisation. L'équation (3) indique que l'association linéaire entre  $X_D$  et  $Y$  est supposée être complètement expliquée par leur association mutuelle avec  $X$  (DEMIRTAS & HEDEKER [2016]). Le ratio  $\delta_{X_D Y} / \delta_{XY}$  est égal à  $\delta_{X_D X} = \frac{E[X_D X]}{\sqrt{pq}} = \frac{E[X|X \geq k]}{\sqrt{pq}}$ . Etant donné  $p$  et la distribution de  $(X, Y)$ , il est constant. Ces corrélations sont invariantes par translation de moyenne ou réduction/augmentation de variance,  $X$  et  $Y$  ne nécessitent pas d'être centrées et réduites. Dès que le ratio est trouvé, nous pouvons calculer la corrélation bisériale si la corrélation de point bisériale est spécifiée. Si  $X$  est ordinalisé afin d'obtenir  $X_O$ , l'idée fondamentale reste inchangée. Si les hypothèses des équations (1) et (3) sont vérifiées, la méthode est également applicable au cas ordinal dans le contexte de la relation entre la corrélation polysériale (pré-ordinalisation) et la corrélation de point polysériale (post-ordinalisation). Le moyen le plus facile de calculer  $\delta_{X_O X}$  est de générer  $X$  avec un large nombre de points, puis de l'ordinaliser afin d'obtenir  $X_O$  et enfin, de calculer la corrélation entre  $X_O$  et  $X$ .  $X$  peut suivre n'importe quelle distribution univariée continue. Cependant,  $X$  est supposée faire partie d'une variable multivariée normale avant la discrétisation.

6. Calculer les corrélations intermédiaires CC à l'aide de la formule donnée par DEMIRTAS et al. [2012], qui comprend les corrélations CC initiales et les coefficients de Fleishman pour les variables continues.

Fleishman caractérise les distributions de variables, dans des cas réelles, par leurs quatre premiers moments. Il a également présenté une procédure de correspondance des moments qui simule des distributions non normales. Cette procédure est basée sur la transformation polynomiale :  $Y = a + bZ + cZ^2 + dZ^3$  où  $Z$  suit une loi normale standard et  $Y$  est centrée et réduite. La distribution de  $Y$  dépend des valeurs des constantes  $a, b, c$  et  $d$  appelées coefficients de Fleishman.

Dans notre cas, nous n'utiliserons pas cette procédure pour simuler les variables continues non-normales car, même si les variables en sorties posséderaient les mêmes moments que les variables en entrée, nous nous retrouverions néanmoins

avec des valeurs négatives pour l'âge, la durée et le capital. Toutefois, cela nous donne une procédure de calcul des coefficients de corrélation intermédiaires CC.

7. Construire la matrice intermédiaire finale  $\Sigma_Z$  en utilisant les résultats précédents.
8. Générer un vecteur normal de moyenne 0 et de matrice de corrélation  $\Sigma_Z$ .
9. Construire les variables binaires, ordinales et continues comme présenté précédemment.

Un dernier problème se pose maintenant. Les statistiques descriptives unidimensionnelles de la partie 4.1.2. nous montraient qu'il était difficile d'ajuster aux lois marginales des lois paramétriques usuelles. Nous cherchons alors à estimer la fonction de répartition empirique pour ensuite calculer son inverse.

La fonction de répartition empirique  $F_n$  est une fonction en escalier attribuant la probabilité  $i/n$  à chacune des valeurs observées, où  $i$  est le nombre d'observations attaché à cette valeur.

Pour des observations  $(x_1, \dots, x_n)$ ,  $F_n$  est la proportion d'observations égales ou inférieurs à  $t$ , i.e.

$$F_n(t) = \frac{\text{Card}\{X_i ; X_i \leq t\}}{n} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i \leq t\}}$$

Afin de déterminer la fonction quantile à utiliser lors des simulations, nous cherchons alors  $y$  tel que :  $F_n(x) - y = 0$ . La fonction est ensuite estimée à l'aide d'interpolations linéaires (nous utilisons la fonction `proxfun` de la librairie `stats` de R).

Dans un premier temps, nous supprimons les valeurs extrêmes Capital à cause de l'impact qu'elles ont sur la corrélation linéaire et la déformation de la distribution. Pour cela, on ne garde que les lignes dont la variable Capital prend une valeur inférieure strictement au 95<sup>ème</sup>-percentile.

### Remarques sur cette méthode :

Pour le calcul de la matrice de corrélation intermédiaire, nous nous sommes appuyés sur les travaux de [DEMIRTAS et al. \[2012\]](#). Dans ceux-ci, on peut également trouver un moyen différent de simuler les variables continues non normales à l'aide des ordres polynomiaux de Fleishman.

### Inconvénients :

Nous utilisons les coefficients de corrélation de Pearson, qui ne sont pas adaptés à toutes les situations.

Nous faisons également l'hypothèse très forte que la matrice de corrélation que nous observons sur notre portefeuille est la même que celle du portefeuille que nous cherchons à simuler.

**Avantages :**

- L'exécution du code est très rapide.
- On simule 9 variables simultanément de manière très facile.
- Les inputs sont modifiables facilement, sans qu'il n'y ait de vérification préalables (sauf présence de valeurs extrêmes).

En pratique, sur R :

- La matrice de corrélation intermédiaire (`over.cor.mat`) est calculée à l'aide de la fonction `overall.corr.mat` de la librairie `OrdBinNonNor`, prenant comme arguments en entrée : le nombre de variables binaires, le nombre de variables continues non-normales, le vecteur des probabilités associées aux variables binaires, la matrice de corrélation linéaire que nous cherchons à reproduire et les coefficients de Fleishman (calculés grâce à la fonction `fleishman.coef` prenant en entrée le nombre de variables continues non-normales, leur skewness et leur kurtosis).
- La méthode NORTA est appliquée à l'aide de la fonction `rnorta` de la librairie `SimcorMultRes`, prenant comme arguments en entrée : la taille de l'échantillon que nous cherchons à simuler, la matrice de corrélation intermédiaire `over.cor.mat` et les fonctions inverses des différentes fonctions de répartition.
- On arrondit à l'entier le plus proche les variables Age et Durée.

**Vérification :**

Nous allons vérifier qu'un portefeuille simulé par cette méthode présente les mêmes caractéristiques de distributions et de dépendance que les données initiales.

Les variables continues considérées sont : l'âge à l'adhésion, le capital emprunté initial, la durée d'emprunt initiale, la quotité garantie Arrêt de travail.

Variables binaires : Sexe, Co-emprunteur(s), Fumeur, IPP, Option, Zone géographique.

Variable ordinale : Catégorie professionnelle (Cadre / Profession libérale / Non cadre).

Les probabilités associées à chaque catégorie des variables qualitatives correspondent à leur proportion dans le portefeuille observé. Nous calculons également la moyenne, la variance, le kurtosis et la skewness de chaque variable continue.

► **Les marginales :**

Dans un premier temps, nous regardons les distributions marginales.

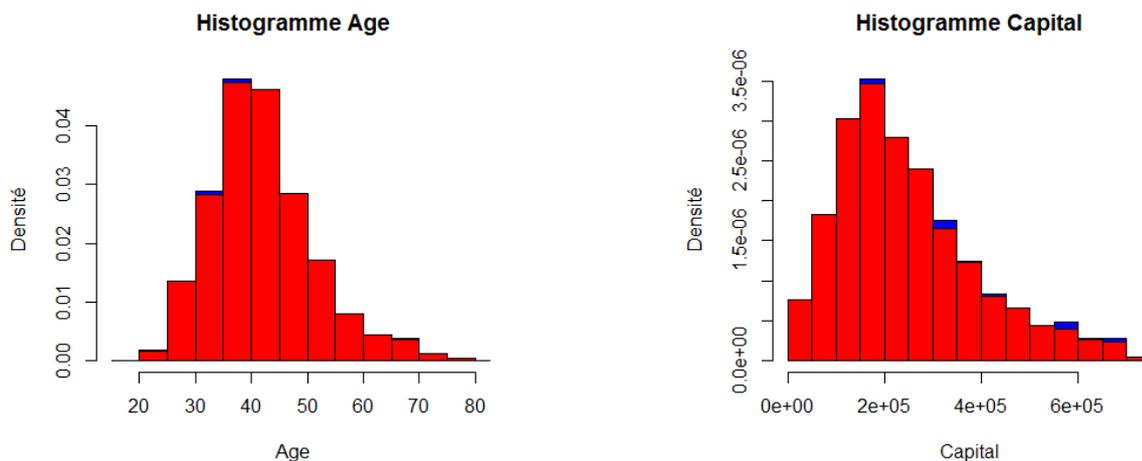
Concernant les variables qualitatives, une rapide comparaison des proportions pour chaque catégorie entre les données observées et les données simulées nous permet de vérifier la bonne conservation des marginales.

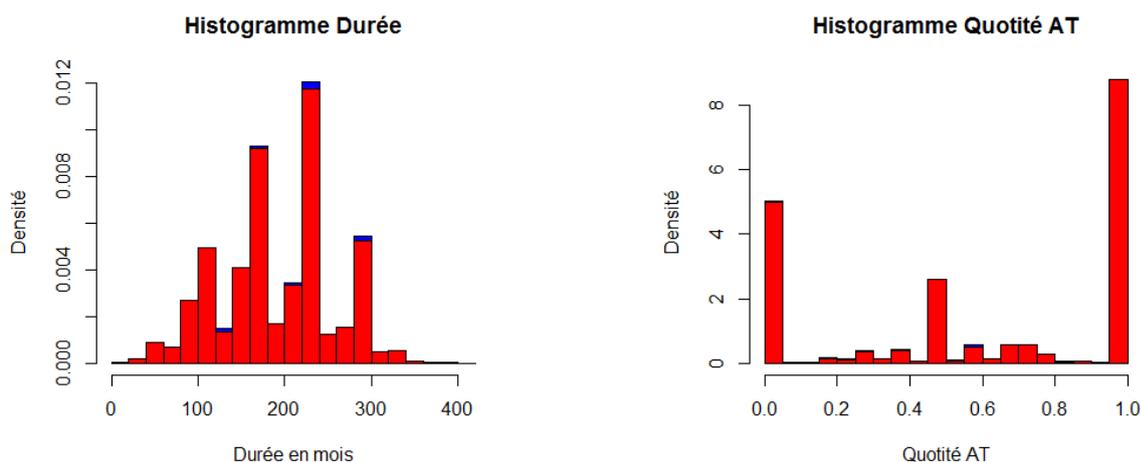
Nous vérifions rapidement en comparant les histogrammes que les variables en entrée et en sortie de l'algorithme présentent les mêmes distributions.

Comme nous pouvons le voir, les histogrammes des variables continues se superposent presque complètement, ce qui confirme que les marginales simulées sont proches des marginales en entrée de manière satisfaisante.

**En rouge :** l'histogramme des données observées.

**En bleu :** l'histogramme des données simulées.





**Figure 9 - Histogrammes données observées vs. simulées**

► La matrice de corrélation :

Nous vérifions ensuite que la matrice de corrélation soit proche de la matrice observée. Pour cela, nous décidons de comparer les erreurs absolues entre les coefficients de corrélation de Pearson calculés sur les données observées et sur les données simulées.

Sexe	Co-emprunteur(s)	Fumeur	Zone géo	Cat pro	IPP	Option	Age	Capital	Quotité AT	Durée
5,64E-02	6,49E-02	4,19E-02	6,14E-02	6,17E-02	1,45E-01	8,30E-02	9,63E-02	5,91E-02	1,77E-01	9,82E-02

**Tableau 7 - Somme par colonne des erreurs absolues entre matrice de corrélation sur données observées vs. données simulées**

Ces résultats nous satisfont. Les coefficients de corrélation restent dans le même ordre de grandeur. Des changements de signes peuvent apparaître lorsque le coefficient initial est proche de 0.

► **Statistiques bidimensionnelles**

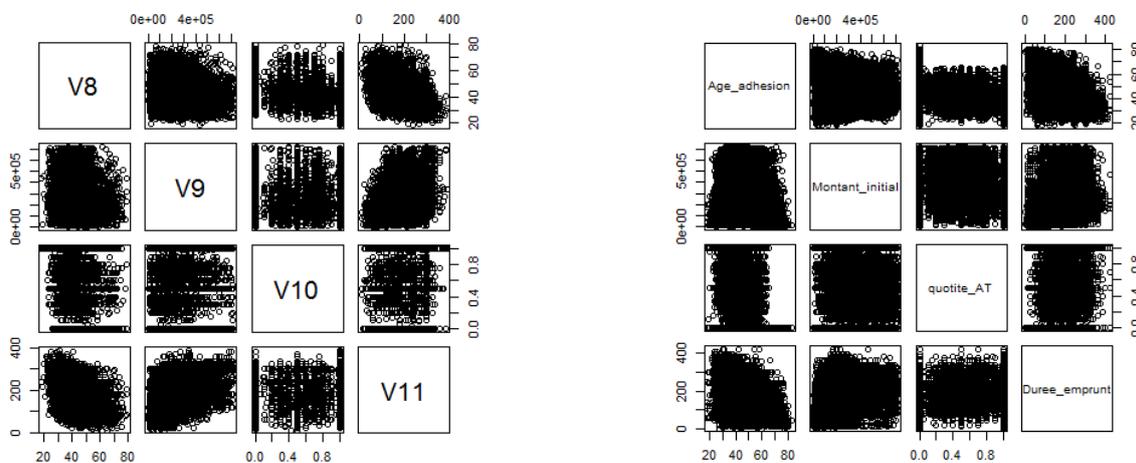
Les deux vérifications précédentes étaient, en théorie, gagnées d'avance. C'est sur l'accomplissement de ces deux derniers points que l'algorithme utilisé est basé. Toutefois, la corrélation linéaire ne semble pas adaptée aux variables rencontrées. Nous voulons vérifier que nous avons conservés la même structure de dépendance.

Pour cela, nous comparons les statistiques bidimensionnelles, et notamment, le V de Cramer.

Absolute Error	Sexe	Rabais couple	Fumeur	Zone_geo	cat_ret	IPP	Option	Age	Montant	quotité	Durée
Sexe											
Rabais_couple	0,00										
Fumeur	0,01	0,00									
Zone_geo	0,00	0,00	0,01								
cat_ret	0,00	0,07	0,00	0,01							
IPP	0,00	0,05	0,01	0,00	0,00						
Option	0,00	0,01	0,02	0,02	0,02	0,01					
Age	0,03	0,03	0,00	0,00	0,05	0,01	0,02				
Montant	0,02	0,02	0,01	0,01	0,01	0,00	0,03	0,02			
quotité	0,04	0,10	0,01	0,08	0,02	0,05	0,07	0,06	0,05		
Durée	0,01	0,00	0,01	0,01	0,02	0,01	0,00	0,00	0,05	0,01	

**Tableau 8 - Erreur absolue entre V de Cramer données observées et données simulées**

La variable continue quotité Arrêt de travail semble être la cause principale de déformation. Cela reste néanmoins acceptable globalement.



**Données simulées**

**Données observées**

**Tableau 9 - Comparaison variables continues observées et simulées**

Nous pouvons observer le même motif pour chaque paire de variables, à l'exception de celles concernées par la variable quotité Arrêt de travail.

Maintenant que nous sommes convaincus que notre simulation remplit bien les objectifs recherchés, nous nous intéressons à la modélisation de certaines variables que nous avons

négligés jusqu'ici. En effet, il est nécessaire de modéliser *a posteriori* la variable « Taux d'emprunt », mais également les valeurs extrêmes de la valeur « Capital emprunté ».

### ► Modélisation du taux d'intérêt

La modélisation du taux d'intérêt à l'adhésion est un sujet que nous avons décidé de traiter à part. Les raisons de ce choix sont en grande partie dues à la difficulté liée à celle-ci.

Le taux d'intérêt dépend de nombreuses variables. Nous possédons des informations sur certaines : la date d'octroi du prêt, sa durée, son type, le projet associé, l'âge de l'emprunteur, son lieu de résidence, sa catégorie professionnelle. D'autres de ces variables ne sont *a priori* connues que de l'organisme prêteur, comme la solvabilité de l'emprunteur (revenus, ...). Enfin, le taux d'intérêt est également dépendant des taux des OAT (Obligations Assimilables au Trésor) et de l'Euribor.

Concernant l'importance du taux d'intérêt dans l'évaluation de la rentabilité, l'analyse effectuée par [AUER \[2007\]](#) semble montrer que, dans notre cas (i.e. tarification sur le CRD), le compte de résultat est peu sensible au taux d'intérêt, relativement à d'autres variables.

Afin de se persuader du réalisme des taux d'intérêt observés sur notre portefeuille, nous les comparons avec des données issues de l'historique des taux d'intérêt [meilleurtaux.com](#).

**Traitement des données :** Nous nous plaçons en vision « affaires ». Seuls les prêts classiques sont conservés. Nous ancrons le taux d'intérêt à la date d'effet du prêt, l'information concernant la date de création n'étant pas fiable. Nous nous intéressons au taux d'intérêt moyen par trimestre et nous segmentons les prêts selon si l'assuré associé est cadre ou non-cadre. Nous ne prenons pas en compte la zone géographique de résidence de l'assuré, la segmentation que nous avons choisie étant différente de celle adoptée par [meilleurtaux.com](#).

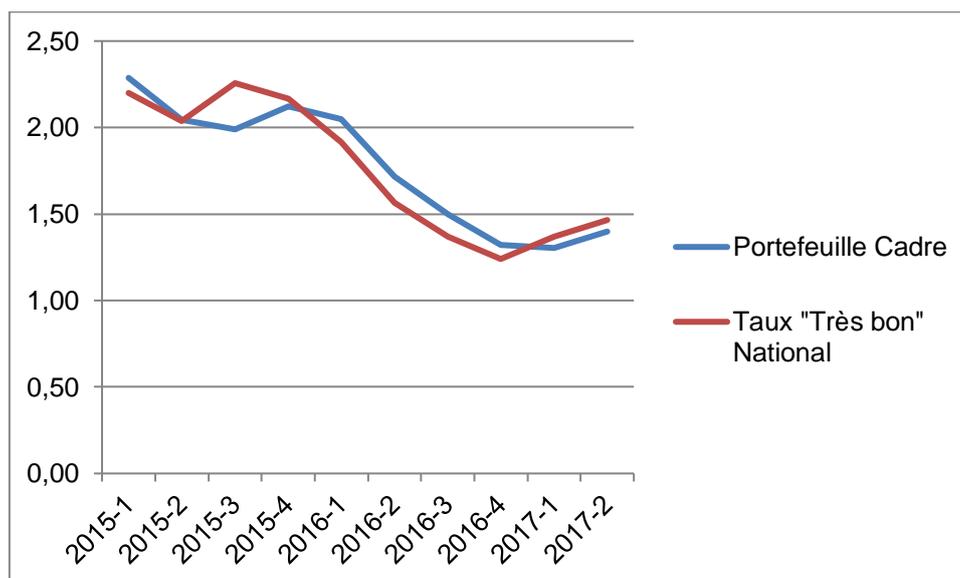
Cadre			
Date	15 ans	20 ans	25ans
2015-1	2,08%	2,29%	2,57%
2015-2	1,86%	2,05%	2,35%
2015-3	1,87%	1,99%	2,34%
2015-4	1,94%	2,12%	2,42%
2016-1	1,95%	2,05%	2,34%
2016-2	1,55%	1,72%	2,03%
2016-3	1,35%	1,50%	1,79%

2016-4	1,21%	1,32%	1,57%
2017-1	1,19%	1,30%	1,52%
2017-2	1,28%	1,40%	1,64%
2017-3	1,28%	1,41%	1,73%

**Tableau 10 - Taux d'intérêt moyen par trimestre pour les assurés Cadre**

Nous extrayons les informations de l'historique meilleurtaux.com. Nous avons accès aux taux « bon », « très bon » et « excellent ». Nous nous plaçons également au niveau « National ». Il apparaît que, pour ces trois durées, les courbes Cadre (resp. Non-cadre) sont proches, au sens de l'erreur quadratique, des courbes « très bon » (resp. « bon »).

Pour la modélisation du taux d'intérêt, on attribuera selon la catégorie professionnelle et la zone géographique, le dernier taux d'intérêt moyen renseigné dans notre base de données. Si c'est nécessaire, nous pourrions compléter ces informations par celles extraites de l'historique meilleurtaux.com.



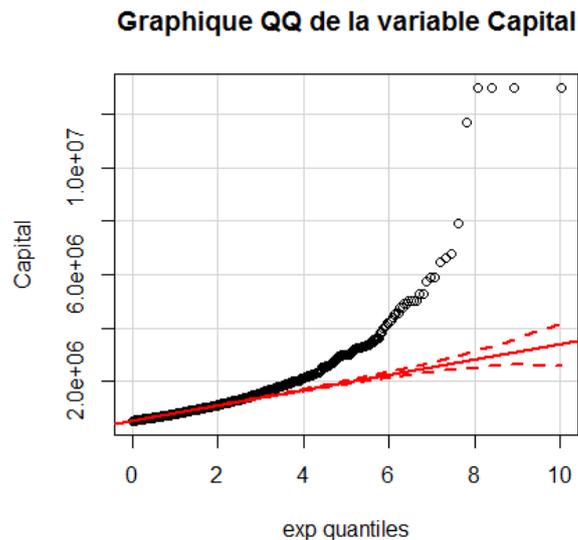
**Figure 10 - Taux durée 20 ans**

► **Modélisation des valeurs extrêmes pour la variable « Montant du capital emprunté »**

Nous avons décidé, dans un premier temps, de ne garder que les lignes dont la valeur de la variable Capital ne dépassait pas le 95-ème percentile. En effet, comme dit précédemment, le coefficient de corrélation de Pearson est très sensible aux valeurs extrêmes. Ainsi, en gardant les valeurs extrêmes, la méthode que nous avons utilisée donne des résultats peu

satisfaisants : la matrice de corrélation en sortie est significativement éloignée de la matrice en entrée.

De ce fait, la méthode de simulation présentée plus haut ne permet pas de simuler les capitaux « extrêmes ». Pour cela, nous ajoutons une étude des extrêmes de la variable capital, afin de pouvoir, si cela paraît nécessaire, ajouter cela en complément du portefeuille obtenu.



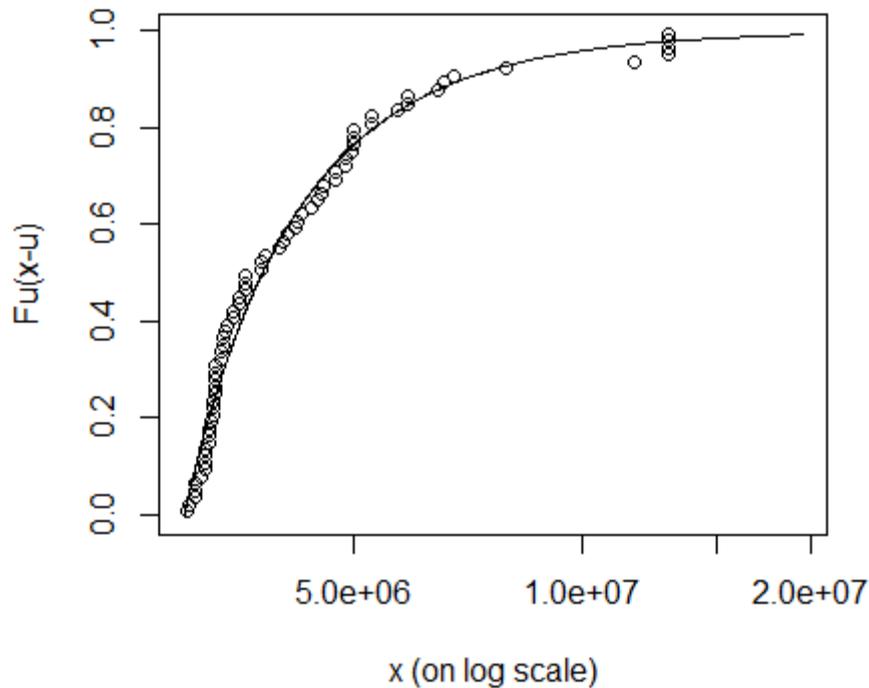
**Figure 11 - Graphique QQ de la variable Capital et de la loi exponentielle**

Nous allons utiliser des outils issus de la théorie des valeurs extrêmes afin de mieux comprendre le comportement de la variable Capital.

Tout d'abord, d'après le graphique quantile-quantile, la variable Capital est dans le domaine d'attraction de Fréchet.

Le graphique de l'espérance de vie résiduelle nous porte à choisir un seuil de 3 000 000. Le graphique de l'estimateur de Hill semble appuyer ce choix.

Finalement, nous estimons les paramètres de la GPD par maximisation de la vraisemblance.



**Figure 12 - Distribution des excès de la variable Capital**

Connaissant les paramètres de la GPD, il est alors très facile de simuler la variable Capital par méthode d'inversion. En effet :

$$G_{\gamma,\sigma,s}(x) = 1 - \left(1 + \frac{\gamma \cdot (x - s)}{\sigma}\right)^{-\frac{1}{\gamma}}$$

$$\Leftrightarrow s + \frac{\sigma}{\gamma} \cdot \left[ \left( -(G_{\gamma,\sigma,s}(x) - 1) \right)^{-\gamma} - 1 \right] = x$$

D'où finalement :  $x = G_{\gamma,\sigma,s}^{-1}(U) = s + \frac{\sigma}{\gamma} \cdot [U^{-\gamma} - 1]$ .

Afin de simuler une variable suivant une GPD de paramètre  $(\gamma, \sigma)$  et de seuil  $s$  :

- On simule une variable  $U$  de loi uniforme sur  $[0,1]$ .
- On applique à cette variable la transformation  $s + \frac{\sigma}{\gamma} \cdot [U^{-\gamma} - 1]$ .

Nous n'avons toutefois pas de méthode systématique à appliquer dans le cas d'un portefeuille dont les caractéristiques seraient différentes. Il faudra, au cas par cas, tenter de s'appuyer sur les informations apportées par notre étude afin d'estimer les capitaux empruntés extrêmes et la population concernée. Dans une approche prudente de l'estimation de la rentabilité, cela a d'ailleurs plus de sens de négliger ces capitaux extrêmes

qui ne représentent pas un risque pour la rentabilité technique ou la marge brute d'ASP, mais plutôt le contraire.

## **4. Application : Analyse de la rentabilité à partir du portefeuille fictif**

Notre objectif est ici d'obtenir une vision de la rentabilité pour le produit d'assurance de prêt que nous tarifons en suivant la méthodologie présentée précédemment dans ce mémoire.

L'absence de données statistiques a priori implique par construction que le risque associé à ce nouveau produit n'est pas maîtrisé. On peut cependant en apprécier le niveau de rentabilité en simulant dans le temps, grâce au compte de résultat, les flux techniques et financiers de ce produit.

Afin d'avoir une vision dans le temps de la rentabilité de ce produit sur les garanties couvrant le risque Décès et le risque Arrêt de travail, nous mettons en place un outil qui permet de projeter le Plan d'Affaires. Celui-ci est un moyen d'obtenir la projection de l'intégralité du compte de résultat, à un horizon de temps donné. De cette manière, nous pouvons estimer la globalité des flux qui entrent et qui sortent pour les années futures et donc, de mesurer la rentabilité du produit.

Dans un premier temps, nous définirons les différents éléments qui constituent le Plan d'Affaires dont notamment la construction d'un compte de résultat comptable et nous présenterons le modèle de projection développé par ASP.

Puis, nous évaluerons, à l'aide de notre outil de génération de portefeuille fictif d'assurés que nous avons présenté auparavant, la rentabilité du produit, après avoir préalablement vérifié la tarification du produit.

### **4.1. Outil de projection du compte de résultats**

Notre objectif ici est d'explicitier les différents éléments sous-jacents à la conception de notre outil de constitution du Plan d'Affaires, qui permet d'estimer les flux entrants et sortants des différents exercices et donc d'obtenir les comptes de résultat pour les garanties « Décès » et « Arrêt de travail ».

Notre outil se présentera en deux parties :

- Un tarificateur calculant les primes Hors Taxes pour les garanties « Décès » et « Arrêt de travail », suivant les taux de primes déterminés en amont, et calculant également les Valeurs Actuelles Probables aussi bien pour l'engagement de l'assureur que pour celui de l'assuré et aussi bien pour la garantie « Décès » que pour la garantie « Arrêt de travail ».

- Un outil projetant le Plan d’Affaires jusqu’à l’année d’extinction du dernier contrat d’assurance.

Il nous faudra décrire quels sont les flux que nous avons retenus pour la constitution des comptes de résultat.

#### **4.4.1. Rentabilité et compte de résultat**

Pour rappel, April Santé Prévoyance est un courtier grossiste et pour cette raison, la rentabilité de notre produit d’assurance de prêt peut être considérée sous deux visions différentes :

- La rentabilité technique du produit, qui impacte en premier lieu les partenaires assureurs et réassureurs d’April Santé Prévoyance, mais qui touche également directement April Santé Prévoyance suivant les engagements pris auprès de ces mêmes partenaires ;
- La marge brute d’April Santé Prévoyance.

La marge nette, i.e. la marge brute à laquelle est déduit les coûts qu’ASP ne refacture pas à ses partenaires comme la sélection médicale par exemple, est la vision la plus proche de la réalité. Néanmoins, en pratique, elle est du ressort du contrôle de gestion d’ASP car l’équipe de conception des produits emprunteurs ne modélisent pas les coûts. Afin de ne pas prendre des hypothèses sur des coûts que nous ne maîtrisons pas, nous nous intéressons donc à la marge brute d’ASP.

Nous chercherons à déterminer la rentabilité sous ces deux visions, sous la contrainte d’hypothèses que nous définirons. Ces hypothèses seront détaillées par la suite et permettront d’assurer la simplicité de l’outil, qui se veut avant tout être un outil permettant de vérification de la tarification, mais également de production d’éléments appuyant la viabilité du produit pour les différents acteurs impliqués dans la conception et la commercialisation de celui-ci.

##### **a) Rentabilité technique**

La rentabilité technique est impactée par les différents éléments que sont la tarification, la sinistralité et le provisionnement.

Afin d’apprécier la rentabilité technique du produit, il y a plusieurs postes à calculer :

- Les primes émises :
  - ▶ Provisions pour Primes Non Emises (PPNE)
  - ▶ Provisions pour Primes Non Acquises (PPNA)
  - ▶ Provisions pour Risques Croissants (PRC)
- Les sinistres
- Les provisions
  - ▶ Provisions pour Sinistres en Cours (PSC) : Cette provision constitue le moyen d'assurer le paiement des prestations en Arrêt de travail pour les sinistres qui se répartissent sur plusieurs mois ou années. Elle n'intervient pas dans le compte de résultat technique car nous nous intéressons aux versements futurs à la date zéro.
  - ▶ Provisions pour Sinistres A Payer (PSAP)

Nous prenons plusieurs hypothèses simplificatrices pour la constitution de ce compte de résultat :

- Les différents éléments qui constituent le compte le résultat sont considérés de façon probable et donc théorique.
- Les primes acquises sont égales aux primes émises. Nous ne modélisons pas les PPNE ou de PPNA.
- Nous ne modélisons pas les déclarations ou règlements tardifs de sinistres. Les PSAP sont inclus dans les charges sinistres.

Exercices de survenance	
Primes Hors Taxes	
<i>dont chargements</i>	
<i>dont Primes Pures</i>	(1)
Charges Sinistres	(2)
PRC au 01/01	(3)
PRC au 31/12	(4)
Variation de PRC	(5) = (3) - (4)
Intérêts sur prime	(6)
Intérêts sur sinistre	(7)
Intérêts sur provision	(8)
Produits financiers	(9) = (6) + (7) + (8)
Résultat technique	(11) = (1) + (2) + (5) + (9)

**Tableau 11 - Postes du compte de résultat**

Les différents éléments qui constituent le compte le résultat sont considérés de façon probable et donc théorique.

**b) Marge brute**

Certains coûts de gestion des adhésions et des sinistres ne sont pas refacturés aux partenaires comme, par exemple, ceux liés à la sélection médicale hors les honoraires médicaux. Ils entrent donc en jeu uniquement dans le compte de résultat d'April Santé Prévoyance.

<b>Commissions</b>	
<i>Commissions de gestion</i>	(1)
<i>Commissions d'apport</i>	(2)
<i>Commissions de courtage en réassurance et de développement</i>	(3)
<i>Commissions courtiers</i>	(4)
<b>Marge brute</b>	<b>= (1) + (2) + (3) – (4)</b>

**Tableau 12 - Définition de la marge nette ASP**

La marge brute correspond aux commissions perçues par April Santé Prévoyance pour la gestion et l'acquisition des contrats ainsi que d'une part des commissions de courtage en réassurance et de développement, nettes des rétrocessions des commissions des courtiers.

**4.1.2. Choix des différents paramètres de projection**

Un certain nombre de paramètres doivent être fixés dans l'objectif de projeter les comptes de résultat. Ces paramètres peuvent soit être d'origine réglementaire, soit être des paramètres d'expérience.

**Paramètres réglementaires :**

Le taux d'actualisation est constant et choisi en fonction du TME (Taux Moyen d'Emprunt). Les méthodes de calcul sont données par les articles du Code des Assurances A. 132-1-1 pour l'assurance vie et A. 331-22 pour l'assurance non vie.

<b>Date</b>	<b>TME mensuel</b>	<b>TME Moyen Semestriel (fixé au 01/n)</b>	<b>0,6</b>	<b>0,75</b>	<b>Taux technique 60%</b>	<b>Taux technique 75%</b>
-------------	--------------------	--	------------	-------------	-------------------------------	-------------------------------

					TME VIE	TME NON VIE
31/01/2017	0,910%	0,538%	0,323%	0,526%	0,250%	0,526%
28/02/2017	1,110%	0,688%	0,413%	0,540%	0,250%	0,540%
31/03/2017	1,100%	0,830%	0,498%	0,557%	0,250%	0,557%
30/04/2017	0,970%	0,932%	0,559%	0,573%	0,250%	0,573%
31/05/2017	0,870%	0,960%	0,576%	0,571%	0,250%	0,571%
30/06/2017	0,710%	0,945%	0,567%	0,555%	0,250%	0,555%

**Tableau 13 - Historique des TME**

Nous fixons le taux d'actualisation à 0,25%.

Concernant les lois de mortalité et des lois de maintien en indemnisation d'arrêt de travail, nous utilisons des tables d'expérience d'ASP, qui ont reçues la certification, et que nous ne présenterons pas pour des raisons de confidentialité.

Les lois d'incidence en Arrêt de travail, nous utilisons les lois utilisés pour la tarification.

**Paramètres d'expérience :**

- Les taux de primes pures best estimate en arrêt de travail sont construits à partir de celles décrites en 2.4. Elles correspondent à la moyenne des coefficients de chaque table pondérée par les proportions de CSP dans chaque CAT.
- Le taux de chute, au sens des chutes commerciales et celles liées aux rachats ou aux renégociations de prêts, est fixé dans un premier temps de façon arbitraire à 10% par an. C'est une vision simpliste de la réalité car le taux de chute est fonction de la durée et de l'ancienneté.

**Autres remarques sur les paramètres :**

- Dans le cas de l'ITT, nous ne considérons que des franchises égales à 90 jours, qui constituent 99% du volume des contrats.
- **Hypothèses de simplifications :**
  - i) Nous nous limitons à une seule génération d'emprunteurs.
  - ii) Il n'y a pas de nouvelles souscriptions d'un pas de projection temporel à un autre.
  - iii) La charge de sinistralité est vue globalement, i.e. nous n'isolons pas les Provisions pour Sinistres à Payer.

Il est nécessaire de calculer les quantités qui suivent à chaque mois. La plupart des notations peuvent être retrouvées en annexes.

- CRD
- Taux de mortalité :
  - ▶ Règlementaire (i.e. certifiée)  ${}_kq_x^{Reg}$ , dans le cas du calcul des PRC
  - ▶ D'expérience (utilisée pour la tarification)  ${}_kq_x^{BE}$ , dans le cas de la modélisation des sinistres *Best Estimate*
- Taux de chute
- Exposition =  $exposition_{k-1}(1 - \text{taux de mortalité}_{k-1})(1 - \text{taux de chute}_{k-1})$ 
  - ▶ Règlementaire  ${}_k p_x^{(m,Reg)}$  : calculée avec le taux de mortalité règlementaire
  - ▶ *Best Estimate*  ${}_k p_x^{(m,BE)}$  : calculée avec le taux de mortalité d'expérience
- Taux d'incidence en Arrêt de travail utilisée dans la tarification
- Table de maintien en arrêt de travail règlementaire (utilisée dans la tarification)
- $tx_{primes}$  les taux de primes

**Primes (formules identiques en Décès et en Arrêt de travail) :**

$$\text{Prime au } 31/12/(N+t) = \sum_{k=0}^{d-1} {}_k p_x^{(m,BE)} \times tx_{primes} \times CRD_{k-1}$$

**Commissions (formules identiques en Décès et en Arrêt de travail) :**

$$\text{Commissions apport au } 31/12/(N + t) = TxComm_{apport} \times \text{Primes au } 31/12/(N + T)$$

$$\text{Commissions gestion au } 31/12/(N + t) = TxComm_{gestion} \times \text{Primes au } 31/12/(N + T)$$

**Sinistres :**

Pour le décès :

$$\text{Sinistres au } 31/12/(N+t) = \sum_{k=0}^{d-1} {}_k p_x^{(m,BE)} \times q_{\left[x+\frac{k}{12}\right]}^{Reg} \times CRD_k$$

Pour l'arrêt de travail :

$$\text{Sinistres au } 31/12/(N+t) = \sum_{k=0}^{d-1} {}_k p_x^{(m,BE)} \times i_{\left[x+\frac{k}{12}\right]}^{(m)} \times Prov^{AT} \left( \left[ x + \frac{k - fr}{12} \right], fr, k, d - k \right)$$

Avec :

$$Prov^{AT} = \frac{\sum_{j=0}^{d-1} M_{k+j} \cdot l_{x,fr+j}^{AT} + \sum_{j=1}^{d-1} M_{k+j} \cdot l_{x,fr+j}^{AT}}{2l_{x,fr}^{AT}}$$

### **Provisions pour Risque Croissant :**

$$PRC = \max(VAP_{\text{assureur}} - VAP_{\text{assuré}}, 0)$$

La formule de calcul des Valeurs Actuelles Probables pour l'assureur (Décès et Arrêt de Travail) sont données en Annexe I.

$$VAP_{\text{assuré}} = \sum_{k=0}^{d-1} \frac{k}{12} p_x^{(m)} \times tx_{\text{primesDC}} \times CRD_{k-1} \times \frac{1}{(1 + i^{(m)})^k}$$

### **4.2. Utilisation du portefeuille fictif pour analyser les niveaux de marges techniques et brutes d'un tarif**

Afin d'analyser les niveaux de marges techniques d'un tarif, une possibilité est de comparer les VAP assureur et VAP assuré sur chaque sous-segment tarifaire. L'intérêt est de pouvoir visualiser les poches de rentabilité ou de perte. Néanmoins, nous pouvons dénombrer trois inconvénients à cette méthode :

- L'analyse devient rapidement difficile avec un nombre de critères et de modalités élevé
- Cela ne nous informe que très peu sur le niveau de rentabilité globale
- C'est une vision un peu puriste qui ne prend pas en compte les décalages de résultats comptables liés aux provisions de primes notamment.

Une manière d'adapter cette méthode est alors d'utiliser un portefeuille fictif d'assurés qui a les mêmes caractéristiques que celles utilisées pour la tarification. Cela permet principalement d'éliminer la première difficulté car nous ne sommes plus obligés de calculer les VAP pour toutes les combinaisons d'âges et de durée d'emprunt, sur chaque sous-segment tarifaire, mais nous pouvons limiter les calculs aux assurés du portefeuille fictif.

En effet, cette vérification peut permettre d'ajuster le tarif sur certains segments en jouant sur la variable « marge technique » définie comme :

$$\text{Taux de prime commercial HT} = \text{Taux de prime technique} \times \frac{(1 + \text{marge technique})(1 + \text{financement})}{1 - \text{chargement pour frais de gestion} - \text{chargement apport} - \text{assureur}}$$

L'avantage du portefeuille fictif est qu'il devient aisé d'analyser les écarts entre engagements sur un segment donné, compte-tenu de ses caractéristiques sur les autres modalités (ex : par tranche d'âges). Ainsi, si notre analyse montre que la méthode de tarification ne permet pas d'avoir des engagements qui se compensent partout, il sera possible d'ajuster le tarif en le majorant sur les segments qui le nécessitent et en le minorant sur d'autres.

Une première analyse montre, qu'au global, une marge de prudence est naturellement incluse dans le tarif, c'est-à-dire que les engagements de l'assuré sont supérieurs aux engagements de l'assureur.

### **Application.**

La comparaison des VAP nous permet donc d'avoir une idée de la marge technique du produit que nous tarifons. Cela permet la modulation des marges techniques par segment, et donc, la modulation des taux de primes commerciales.

Nous nous plaçons dans une situation réelle. Les données que nous utiliserons sont issues d'un appel d'offre.

Nous allons effectuer cette comparaison des VAP pour des taux de primes construits grâce à l'outil présenté en partie 3. Les différents taux de passage entre chaque critère discriminant sont fictifs car l'objectif est ici de décrire la démarche.

Nous construirons ensuite notre portefeuille fictif par le simulateur présenté précédemment.

#### ▪ **Présentation des données**

Nous supposons que nous disposons des données macro suivantes :

- ▶ Les assurés : répartition par tranche d'âge (20-25 ; 25-30 ; 35-40 ; 40-45 ; 45-50 ; 50-55 ; 55-60), répartition par secteur d'activité, répartition par statut, répartition par département
- ▶ Les prêts : le montant moyen
- ▶ Le nombre d'assurés potentiel

Pour affiner l'analyse et utiliser l'outil de tarification simplifié, nous avons besoin de compléter les données manquantes suivantes :

- ▶ La répartition de fumeurs
- ▶ La répartition Isolé/Couple
- ▶ La répartition par tranche de durée
- ▶ La répartition par tranche de capital
- ▶ Le taux de souscription de la garantie Arrêt de travail, des rachats des exclusions des pathologies dos/psy et de la garantie IPP

Excepté si l'organisme à l'origine de l'appel d'offre possède une expérience dans l'assurance de prêt, il sera courant de faire face à ce manquement dans les données qui concernent directement le prêt ou les garanties.

▪ **Présentation du tarif :**

<i>Compo.</i>	<b>CSP1</b>	<b>CSP2</b>	<b>CSP3</b>	<b>CSP4</b>
<b>CAT1</b>	60%	20%	20%	0%
<b>CAT2</b>	50%	20%	20%	10%
<b>CAT3</b>	10%	20%	40%	30%

**Tableau 14 - Composition des CAT du tarif exemple**

Les nouvelles catégories sont très proches des CSP moins risquées de l'ancien produit. Les CAT1 et CAT2 ont une composition similaire.

<b>DC</b>	
1 -> 2	30%
1 -> 3	50%
Co-emprunteur	-15%
Fumeur	50%

**Tableau 15 - Coefficients de passage du tarif Décès exemple**

Du fait de la proximité de composition entre les deux premières catégories, le passage entre les catégories 1 et 2 aura un effet déformateur important sur les taux de primes de référence. La même remarque peut être faite pour la minoration « Présence d'un co-emprunteur » qui, rappelons-le, est un critère de segmentation purement commercial.

AT	
1 -> 2	10%
1 -> 3	30%
Couple	-10%
Fumeur	25%
Co-emprunteur	40%
IPP	40%
Géo	-40%

**Tableau 16 - Coefficients de passage du tarif Arrêt de travail exemple**

Si nous imposons un coefficient de passage pas assez important par rapport à la sur sinistralité réellement observée, cela aura pour effet d'augmenter le tarif de référence, et vice versa.

▪ **Utilisation du simulateur construit :**

Pour rappel, le simulateur de portefeuille fictif que nous avons construit prend comme paramètres en entrée une matrice de corrélation linéaire et les lois marginales des variables que nous cherchons à simuler.

Concernant la matrice de corrélation, nous choisissons de prendre la même matrice que celle observée sur le portefeuille ASP.

Nous sommes en mesure de calculée une fonction de répartition pour la variable « Âge » grâce aux informations données. La variable « Capital » est modélisée suivant une loi log-normale. La variable « Durée » suit la même loi que celle observée sur le portefeuille ASP, centrée sur 180 mois. Si l'information n'est pas donnée pour une variable binaire, nous complétons par les données du portefeuille ASP.

- Calcul de la rentabilité dans le cas « de base », avec les caractéristiques décrites plus haut :

Nous présentons uniquement les résultats pour la garantie Décès.

Pour un portefeuille fictif de 1 000 assurés, nous modélisons le compte de résultat Décès à chaque exercice de survenance.

Exercices de survenance	Année 1	Année 2	Année 3	...
<b>Primes HT</b>	<b>192 911 €</b>	<b>193 676 €</b>	<b>193 367 €</b>	...
<i>Gestion</i>	50 157 €	50 356 €	50 276 €	
<i>Acquisition</i>	28 937 €	29 051 €	29 005 €	
<i>Commission de réassurance</i>	1 929 €	1 937 €	1 934 €	
<i>Primes techniques (1)</i>	111 888 €	112 332 €	112 153 €	
<b>Sinistres+Honoraires (2)</b>	113 058 €	113 606 €	113 374 €	...
<i>Charges Sinistres</i>	<b>111 129 €</b>	<b>111 670 €</b>	<b>111 440 €</b>	
<i>Honoraires</i>	1 929 €	1 937 €	1 934 €	
<b>PRC au 01/01/N (3)</b>	- €	54 510 €	51 011 €	
<b>PRC au 31/12/N (4)</b>	54 510 €	51 011 €	47 552 €	
<b>Variations de PRC (5)=(3)-(4)</b>	54 510 €	- 3 499 €	3 459 €	...
<b>Produits financiers (9)</b>	- €	132 €	123 €	...
<i>Intérêts sur PRC (8)</i>	- €	132 €	123 €	
<b>Résultat technique (11)=(1)-(2)-(5)+(9)</b>	- 55 680 €	2 357 €	2 361 €	...
<b>Résultat technique cumulé</b>	- <b>55 680 €</b>	- <b>53 323 €</b>	- <b>50 962 €</b>	...

La constitution des Provisions pour Risques Croissants en première année a un lourd impact négatif sur le résultat technique comme nous pouvons le constater.

#### **4.3. Recherche de la structure tarifaire et de la structure des chargements optimaux au regard de contraintes a priori**

Ce paragraphe a pour objectif de mettre en lumière des moyens d'optimiser le tarif en utilisant la méthodologie spécifique présentée précédemment dans ce mémoire.

L'optimisation du tarif peut s'entendre comme la recherche de la maximisation de la rentabilité (rentabilité technique et marge brute), la minimisation des écarts avec la concurrence (sur la cible du produit – par exemple, les cadres de 35 à 45 ans) ou une combinaison des deux (ce qui est recherché en pratique). Néanmoins, nous n'avons pas accès à un benchmark des tarifs proposés par la concurrence. Les seules informations à notre disposition dans la situation d'un appel d'offres concernent la segmentation utilisée par les concurrents et le fait que plus une segmentation est faible, plus un risque en particulier est attiré ou éloigné de la cible. Nous chercherons donc à trouver la structure tarifaire maximisant la rentabilité.

### **Tests d'impact de la modification des critères de segmentation ou de leur granularité :**

Afin de tester l'impact de la modification des critères de segmentation, nous proposons d'effectuer les comparaisons entre les rentabilités techniques et les marges brutes du tarif de base (présenté précédemment) et un tarif où nous altérons certains critères de segmentation, ce qui a pour conséquence de modifier le tarif de référence. L'objectif est de montrer de quelle manière il est possible d'utiliser les différents outils construits dans les décisions de construction de la structure tarifaire.

Cette partie est une ouverture et une ébauche de ce pour quoi le portefeuille fictif pourrait être utilisé avec l'ensemble des données réelles, ce que nous n'avons pas (notamment sur la cible d'assurés recherché, les arbitrages à faire, les contraintes opérationnelles, ...).

Dans la suite de cette partie, lorsque nous neutralisons une segmentation (fumeur, présence d'un co-emprunteur, ...), cela signifie que le coefficient de majoration dans l'outil de tarification est fixé à 0. Cela implique, par exemple dans le cas de la segmentation fumeur, une augmentation du tarif de référence.

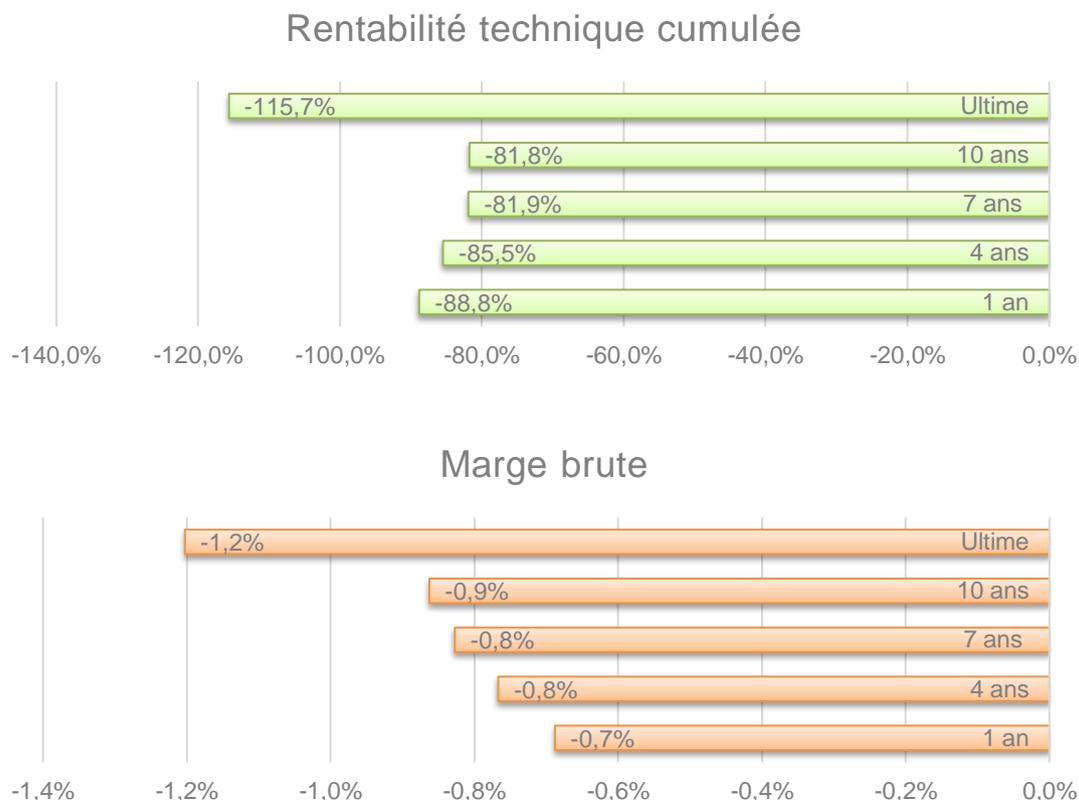
Nous ne touchons pas aux taux de chargements utilisés dans la tarification.

- Tarif sans segmentation Fumeur / Non-fumeur

Nous reprenons dans un premier temps le tarifateur présenté dans la deuxième partie de ce mémoire. Dans celui-ci, nous fixons le coefficient de majoration tarifaire entre le statut Fumeur et le statut Non-fumeur à 0%. Le tarif de référence est donc majoré. Nous ne touchons pas aux chargements.

Puis, nous calculons les postes du Plan d'Affaires comme présenté dans la partie précédente. Les rentabilités techniques et marges brutes à horizon 1 an, 4 ans, 7 ans, 10 ans et à l'ultime sont comparées avec celles obtenues dans le cas de base.

Les différences relatives dans le cas de la garantie Décès / PTIA sont présentées ci-dessous :



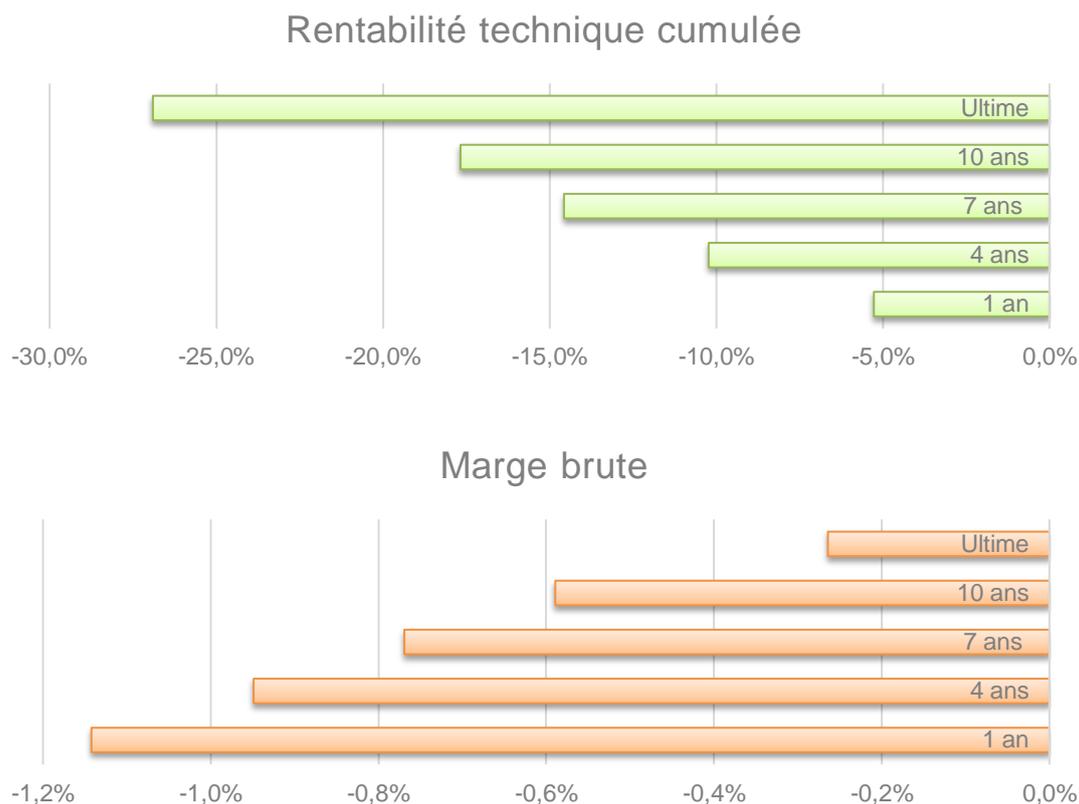
**Figure 13 - Différences de rentabilité relatives entre tarif de base et tarif sans segmentation Fumeur**

A l'ultime, nous pouvons conclure de la même manière pour les deux indicateurs de rentabilité : la suppression de la segmentation Fumeur impacte de manière négative la rentabilité du tarif. C'est également le cas aux autres horizons de temps à des degrés moindres.

- Tarif sans critère Présence d'un co-emprunteur

Nous suivons la même méthodologie que pour le tarif sans segmentation Fumeur, cette fois-ci en fixant à 0% le critère de minoration tarifaire appliqué dans le cas où il y a présence d'un co-emprunteur.

Les résultats obtenus dans le cas de la garantie Décès / PTIA sont présentés ci-dessous :



**Figure 14 - Différences de rentabilité relatives entre le tarif sans critère Présence d'un co-emprunteur et le tarif de base**

Les deux critères de rentabilité ont une évolution différente suivant les horizons de temps considérés.

#### **4.4. Comparaison de la méthode sur un portefeuille réel d'assurés**

Afin de valider la méthode que nous avons présentée, nous cherchons maintenant à étudier la sensibilité du tarif obtenu ainsi que de la rentabilité qui lui est associé par rapport à l'outil de simulation de portefeuille fictif. Autrement dit, nous cherchons à savoir quel est l'écart entre un tarif qui serait calculé à partir d'un portefeuille « réel », pour lequel toutes les informations sont connues et cela, de manière très précise, et un tarif calculé à partir d'un portefeuille aux caractéristiques « moyennes », comme cela est le cas dans ce genre d'appel d'offres (cf. partie précédente). Pour cela, nous suivons le processus suivant :

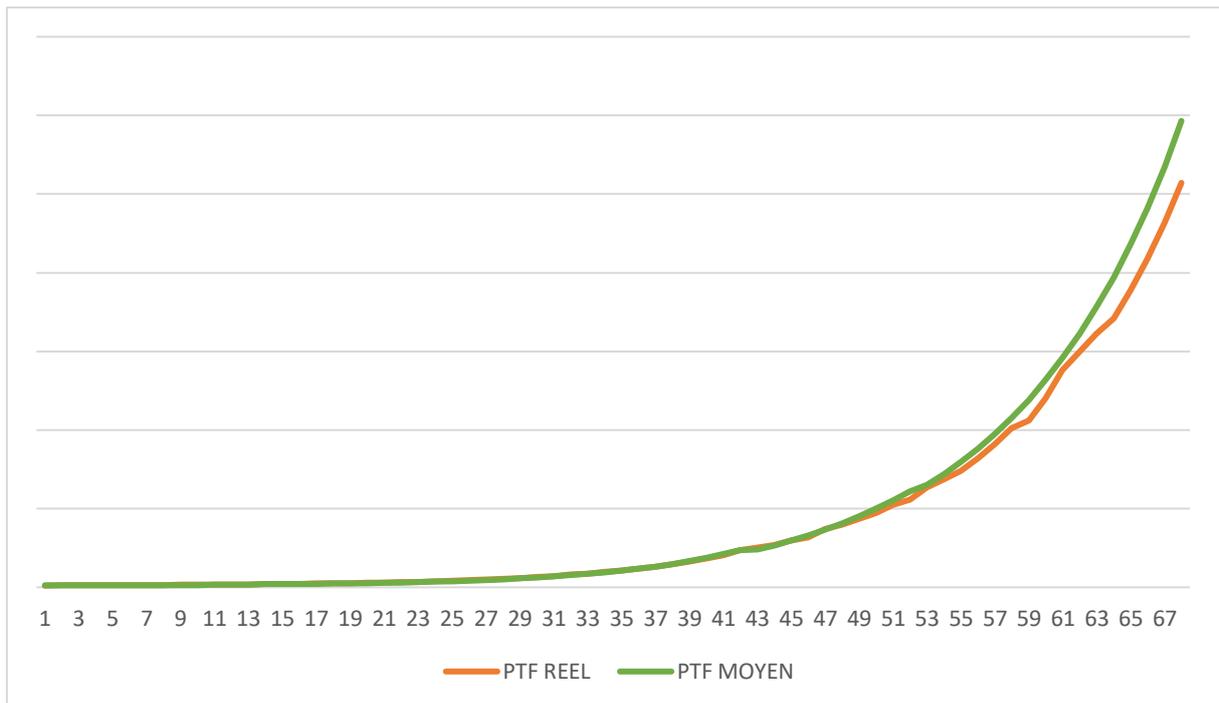
- 1) Nous prenons un portefeuille réel d'assurés dont nous connaissons toutes les informations de la manière la plus précise possible, âge par âge, pour tous les critères de segmentation.

- 2) Nous construisons à partir de ce portefeuille un portefeuille aux caractéristiques « moyennes », c'est-à-dire un portefeuille dont le niveau d'informations est aussi détaillé que pour celui de la partie précédente. A partir de ces informations, nous simulons un portefeuille fictif d'assurés, suivant la méthode présentée plus tôt. Pour être plus précis :
- ▶ La répartition âge par âge est prise constante par tranches d'âges (18-30 ; 31 – 39 ; 40 - 49 ; 50 – 59 ; 60 – 69 ; 70 – 85), égale à la moyenne de la répartition du portefeuille réel.
  - ▶ Les répartitions âge par âge des critères de discrimination Fumeur, Présence d'un co-emprunteur, Réduction géographique, Option, IPP, ... sont égales pour tous les âges à la répartition moyenne totale.
- 3) Pour chacun des portefeuilles, nous calculons le tarif grâce au tarificateur suivant la méthode de ce mémoire. Nous supposons dans un premier temps avoir les mêmes niveaux de chargements et de marge technique.
- 4) Nous analysons et comparons ces deux tarifs.

Deux facteurs vont donc jouer sur la sensibilité du tarif :

- Le manque d'informations à disposition,
- La volatilité des données générées par notre outil de simulation.

Nous commençons par analyser les deux tarifs générés sur la catégorie de référence, pour la garantie Décès / PTIA, afin de neutraliser en partie la volatilité induite par l'outil de simulation.



**Figure 15 - Comparaison des tarifs obtenus pour un portefeuille "réel" et un portefeuille "moyen"**

Ces premiers résultats montrent, dans ce cas, que :

- Les deux tarifs sont relativement proches. Il y a un écart sur les âges plus élevés dû à la volatilité des données disponibles sur ces âges.
- Le tarif calculé à partir des données moyennes est globalement plus élevé que celui calculé à partir du portefeuille complet. Cela ajoute une marge de prudence au tarif.

Globalement, ces informations ainsi que les tests de robustesse effectués dans la partie de simulation du portefeuille fictif semblent indiquer une méthode plutôt satisfaisante de tarifier rapidement dans un contexte d'informations partielles.

## **Conclusion générale**

Ce mémoire s'inscrit dans un contexte d'ouverture de plus en plus large à la concurrence sur le marché de l'assurance de prêts. De nouveaux acteurs cherchent à en profiter, parfois sans avoir les outils en interne pour y parvenir. Afin de remporter les appels d'offres découlant de ce besoin, il est nécessaire d'être en mesure de concevoir un tarif rapidement, tout en étant adapté à la population cible et de produire des arguments concrets prouvant les avantages de ce tarif.

Nous avons dans un premier lieu présenté une méthode de tarification, se voulant rapide et qui peut s'adapter à des portefeuilles aux caractéristiques très différentes, en s'appuyant notamment sur des travaux déjà effectués par ASP. Puis, nous avons cherché à modéliser la rentabilité de ce tarif, en s'aidant pour cela d'un outil permettant de générer un portefeuille d'assurés fictif, respectant les interactions entre les différentes variables ainsi que les caractéristiques à notre disposition lors d'un appel d'offre de ce type : moyenne, répartition par âge, composition cadre / non-cadre, ...

La première partie consacrée à la méthodologie de tarification est avant tout une avancée au niveau opérationnel : l'outil qui en découle apporte un gain de temps conséquent et une flexibilité intéressante aux équipes de Conception en Assurance emprunteur, pour produire un tarif très rapidement dans le cas d'un appel d'offres.

L'outil de génération de portefeuille d'assurés fictif apporte une réponse à la problématique de tarification et de calcul de la rentabilité en l'absence de données statistiques. Bien qu'imparfait car s'appuyant sur des certaines hypothèses restrictives sur la forme de la dépendance à modéliser (coefficients de corrélation de Pearson) mais également sur les données du portefeuille ASP qui peuvent être biaisés, il permet une meilleure prise en compte des spécificités de la cible visée et ainsi une tarification et une appréciation de la rentabilité affinées. Cela est également un moyen d'analyser l'impact d'une modification des critères de segmentation.

Enfin, l'outil permet également de tester l'impact sur les résultats de scénarios de déformation de la démographie produite par rapport à celle imaginée a priori au moment de la tarification, ce qui est très utile dans l'exercice de tarification et d'analyse de la rentabilité.

En complément, ce générateur de portefeuille fictif pourrait également être envisagé dans des cadres d'utilisation différents, et notamment, une mesure globale ajustée de la compétitivité tarifaire d'une nouvelle gamme, un meilleur calibrage de la prime moyenne attendue du new business ou une aide à la définition de cibles à valeur ajoutée qui serait à privilégier...

## **Bibliographie**

AUER E. [2007] *Modélisation d'un contrat emprunteur : Impacts et résultats*, Mémoire.

BEN HALIMA M. A., DEBRAND T., REGAERT C. [2011]. *Comprendre les disparités des arrêts maladie selon les départements*, Revue française d'économie, vol. xxvi, no. 4, 2011, pp. 121-159.

DEMIRTAS H, HEDEKER D, MERMELSTEIN RJ [2012]. *Simulation of massive public health data by power polynomials*. Statistics in medicine. 2012;31(27):10.1002/sim.5362. doi:10.1002/sim.5362.

DEMIRTAS H. [2014]. *Joint Generation of Binary and Nonnormal Continuous Data*. J Biom Biostat S12:001. doi: 10.4172/2155-6180.S12-001

DENUIT M., CHARPENTIER A. [2005] *Mathématiques de l'assurance non-vie Tome II*, Economica.

FABER M. H. [2003], *Risk and Safety*. Lectures Notes, SFIT.

LASSER, K., BOYD, J.W., WOOLHANDLER, S., HIMMELSTEIN, D.U., McCORMICK, D., BOR, D.H. [2000], *Smoking and mental illness: a population-based prevalence study*. Jama, 284(20):2606e10.

MINASSIAN D. [2013] *Tarification, modélisation et rentabilité d'un contrat emprunteur*, Mémoire.

PETAUTON P. [2004], *Théorie et Pratique de l'Assurance-Vie*. DUNOD 3ème édition.

PLANCHET F., THEROND P. [2006] *Modèles de durée : applications actuarielles* Economica.

SAPORTA G. [2006], *Probabilités, analyse des données et statistiques*, TECHNIP 2ème édition.

TALOURD F. [2003], *La rentabilité des contrats collectifs d'assurance emprunteur*, Mémoire.

[TOULOUMIS, A. \(2016\)](#). *Simulating Correlated Binary and Multinomial Responses under Marginal Model Specification: The SimCorMultRes Package*. *The R Journal* 8:2, 79-91.

## Annexes

### Annexe I : Formule de calcul des VAP assureur Décès / PTIA

#### • Décès / PTIA

Etant donnés :

$d$  : durée de l'emprunt en mois

$x$  : âge de l'assuré en début de période

$q_x$  : probabilité annuelle de survenance d'un sinistre décès / PTIA pour un assuré d'âge  $x$

$r_x$  : probabilité annuelle de chute (hors décès indemnisé) pour un assuré d'âge  $x$

$i$  : taux d'actualisation annuel

$tx$  : taux d'intérêt de l'emprunt

$M_{k-1}$  :  $k$ -ième mensualité de l'emprunt,  $k \geq 1$

$CRD_k$  : le capital restant dû assuré après  $k$  mensualités

$$CRD_0 = \text{capital\_assure\_initial}, \quad CRD_k = CRD_{k-1} \cdot \left(1 + \frac{tx}{12}\right) - M_{k-1}, \quad k \geq 1$$

On définit :

$q_x^{(m)}$  : probabilité mensuelle de survenance d'un sinistre décès / PTIA pour un assuré d'âge  $x$

$$q_x^{(m)} = 1 - (1 - q_x)^{1/12}$$

$r_x^{(m)}$  : probabilité mensuelle de chute (hors décès indemnisé) pour un assuré d'âge  $x$

$$r_x^{(m)} = 1 - (1 - r_x)^{1/12}$$

$i^{(m)}$  : taux d'actualisation mensuel

$$i^{(m)} = (1 + i)^{1/12} - 1$$

${}_{\frac{k}{12}}p_x^{(m)}$  : probabilité pour un assuré d'âge  $x$  d'être encore en portefeuille au k-ième mois d'emprunt

$${}_{\frac{k}{12}}p_x^{(m)} = {}_{\frac{k-1}{12}}p_x^{(m)} \cdot \left( 1 - q_{\left[ x + \frac{k-1}{12} \right]}^{(m)} \right) \cdot \left( 1 - r_{\left[ x + \frac{k-1}{12} \right]}^{(m)} \right)$$

Etant donnés ces notations, la valeur actuelle probable de l'engagement assureur en Décès / PTIA est donnée par :

$$VAP_{assureur}^{DC/PTIA} = \sum_{k=0}^{d-1} {}_{\frac{k}{12}}p_x^{(m)} \cdot q_{\left[ x + \frac{k}{12} \right]}^{(m)} \cdot CRD_k \cdot \frac{1}{\left( 1 + i^{(m)} \right)^k}$$

## Annexe II : Formule de calcul des VAP assureur Arrêt de travail

- **ITT**

En complément des notations précédentes, soit :

$i_x$  : probabilité annuelle d'entrée en indemnisation d'incapacité pour un assuré d'âge  $x$

$f$  : franchise d'incapacité

Et :

$i_x^{(m)}$  : probabilité mensuelle d'entrée en indemnisation d'incapacité pour un assuré d'âge  $x$

$$i_x^{(m)} = 1 - (1 - i_x)^{1/12}$$

Etant donnés ces notations, la valeur actuelle probable de l'engagement assureur en ITT est donnée par :

$$VAP_{assureur}^{ITT} = \sum_{k=fr}^{d-1} p_x^{(m)} \cdot i_{\lfloor x + \frac{k}{12} \rfloor}^{(m)} \cdot \text{Pr ov}^{Incap} \left( x + \frac{k-fr}{12}, fr, k, d-k \right) \cdot \frac{1}{(1+i^{(m)})^k}$$

Avec :

$$\text{Pr ov}^{incap} (x, fr, k, d) = \frac{\sum_{j=0}^{d-1} M_{k+j} \cdot l_{x, fr+j}^{incap} \cdot \frac{1}{(1+i^{(m)})^j} + \sum_{j=1}^{d-1} M_{k+j} \cdot l_{x, fr+j}^{incap} \cdot \frac{1}{(1+i^{(m)})^j}}{2 \cdot l_{x, fr}^{incap}}$$

Etant donnée une période de prêt  $[k ; k+1[$ , on suppose que les sorties d'indemnisation se font en fin de période.

Le premier terme du numérateur correspond alors à une entrée en indemnisation en début de période et le second à une entrée en fin de période.

- **IPT**

$$VAP_{assureur}^{IPT} = \sum_{k=0}^{d-1} p_x^{(m)} \cdot i_{\lfloor x + \frac{k}{12} \rfloor}^{(m)} \cdot \text{Pr ov}^{IVA} \left( x + \frac{k}{12}, k, d-k \right) \cdot \frac{1}{(1+i^{(m)})^k}$$

Avec :

$$Pr ov^{IVA}(x, k, d) = \frac{\sum_{j=0}^{d-1} S_{x,j} \cdot Pr ov^{Inval} \left( x + \frac{j}{12} \right], k + j, d - j) \cdot \frac{1}{(1+i^{(m)})^j} + \sum_{j=1}^{d-1} S_{x,j} \cdot Pr ov^{Inval} \left( x + \frac{j}{12} \right], k + j, d - j) \cdot \frac{1}{(1+i^{(m)})^j}}{2 \cdot l_{x,fr}^{incap}}$$

$Pr ov^{Inval}(x, k, d)$  = Provision d'invalidité en cours correspondant à une entrée en invalidité à l'âge x au cours du k-ième mois d'emprunt et sachant qu'il reste d mensualités d'emprunt.

$S_{x,j}$  = Nombre d'assurés qui, entrés en incapacité à l'âge x, passent en invalidité après une période d'incapacité de j mois.

- **Exonération des cotisations d'ITT/IPT en cas d'indemnisation en cours**

L'évaluation du coût de cette garantie est effectué en remplaçant la mensualité d'emprunt par la chronique des cotisations ITT / IPT dans les formules de provisions ci-dessus.

## Table des figures

Figure 1-Diagramme Quantile-Quantile log(Age à l'adhésion).....	47
Figure 2-Ajustement log-normal sur l'âge .....	47
Figure 3 - Boite à moustache des variables "Taux d'intérêt à l'adhésion" et "Capital emprunté initial" .....	48
Figure 4-Ajustement log-normal sur le capital (hors valeurs extrêmes) .....	49
Figure 5-Matrice de nuage de points .....	51
Figure 6 - Matrice de corrélation de Pearson.....	52
Figure 7 - Matrice de corrélation partielle .....	53
Figure 8 - Souscription IPP x Souscription Option de Rachat.....	54
Figure 9 - Histogrammes données observées vs. simulées.....	70
Figure 10 - Taux durée 20 ans .....	73
Figure 11 - Graphique QQ de la variable Capital et de la loi exponentielle.....	74
Figure 12 - Distribution des excès de la variable Capital.....	75
Figure 13 - Différences de rentabilité relatives entre tarif de base et tarif sans segmentation Fumeur .....	89
Figure 14 - Différences de rentabilité relatives entre le tarif sans critère Présence d'un co-emprunteur et le tarif de base .....	90

## Table des tableaux

Tableau 1 - Capture d'écran : Sélection de la structure tarifaire pour la garantie AT .....	33
Tableau 10 - Vérification de l'équilibre technique du tarif.....	37
Tableau 2 - Exemple de model point utilisée pour l'étude de rentabilité .....	40
Tableau 3 - « Photographie » de l'assuré moyen .....	46
Tableau 4 - Description des variables quantitatives .....	46
Tableau 5 - Matrice des V de Cramer.....	55

Tableau 6 - Somme par colonne des erreurs absolues entre matrice de corrélation sur données observées vs. données simulées .....	70
Tableau 7 - Erreur absolue entre V de Cramer données observées et données simulées .....	71
Tableau 8 - Comparaison variables continues observées et simulées .....	71
Tableau 9 - Taux d'intérêt moyen par trimestre pour les assurés Cadre.....	73
Tableau 10 - Postes du compte de résultat.....	79
Tableau 11 - Définition de la marge nette ASP .....	80
Tableau 12 - Historique des TME .....	81
Tableau 13 - Composition des CAT du tarif exemple.....	85
Tableau 14 - Coefficients de passage du tarif Décès exemple.....	85
Tableau 15 - Coefficients de passage du tarif Arrêt de travail exemple.....	86