

Faculté des sciences

# Modélisation actuarielle de la plateforme Medi-share

Auteur : Enkack Tientcheu Jacques Juste

Promoteur : Michel Denuit

Lectrice : Cindy Courtois

Année académique 2022-2023



Institut de Statistique, Biostatistique et Actuariat

Modélisation actuarielle de la plate-forme Medi-share

Promoteur :

Prof. Denuit Michel  
UCLouvain

Mémoire présenté en vue de  
l'obtention du mastère  
en sciences actuarielles

par :

Enkack Tientcheu Jacques Juste

Louvain-La-Neuve  
Septembre 2023



## Résumé

L'évolution de la technologie ainsi que les changements comportementaux ont favorisés l'émergence d'une nouvelle forme d'assurance basée sur le partage de risque entre pairs avec ou sans l'intervention d'un (re-)assureur. Ce nouveau modèle de partage de risque offre une perspective nouvelle et révolutionnaire pour le secteur de l'assurance en transformant fondamentale la manière dont les risques sont gérés et repartis entre les participants. Ce travail propose la modélisation actuarielle d'une plate-forme d'assurance collaborative destinée au partage des coûts des soins médicaux entre les participants. Grâce à ses propriétés intéressantes, la conditional mean risk sharing a été utilisée pour définir la manière dont les contributions sont calculées dans le groupe, ce qui permet de garantir un partage aussi bien équitable que transparente et compréhensible pour tous les participants, puisque basé sur le concept de moyenne. La possibilité d'un recours à un assureur partenaire pour la couverture d'une partie du coût total des frais médicaux du groupe a été étudié, permettant ainsi de limiter l'exposition des participants aux fluctuations du coût total pouvant rendre la valeur des cotisations très élevés. Ce nouveau modèle d'assurance répond ainsi à l'un des principaux défis auxquels est confronté le secteur de l'assurance au 21<sup>e</sup> siècle à savoir : répondre aux besoins spécifiques des clients en leur apportant une couverture appropriée, répondant à leurs aspirations personnelles.

## Abstract

The evolution of technology and changes in behavior have favored the emergence of a new form of insurance based on peer-to-peer risk sharing, with or without the intervention of a (re-)insurer. This new risk-sharing model offers a new and revolutionary perspective for the insurance industry, fundamentally transforming the way risks are managed and distributed among participants. In this work, we propose the actuarial modeling of a collaborative insurance platform designed to share the costs of healthcare between participants. Thanks to its interesting properties, conditional mean risk sharing has been used to define the way contributions are calculated within the group, guaranteeing a sharing risk-sharing rule that is both fair, transparent and comprehensible to all participants, since it is based on the concept of averaging. We looked into the possibility of having a partner insurer cover part of the pool's total healthcare costs, thereby limiting participants' exposure to fluctuations in total costs that could make the value of contributions very high. This new insurance model thus responds to one of the key challenges facing the insurance industry in the 21st century : meeting customers' specific needs by providing them with appropriate coverage that meets their personal aspirations.

## Remerciement

En entamant la rédaction de ce mémoire, il est essentiel pour moi de reconnaître le soutien et la contribution précieuse de nombreuses personnes qui ont rendu cette réalisation possible.

Tout d'abord, je tiens à exprimer ma profonde gratitude envers mon promoteur, le Professeur *Michel Denuit*, dont les conseils éclairés, la patience et la disponibilité ont été une source d'inspiration tout au long de ce parcours. Votre expertise et votre soutien ont été des atouts majeurs dans la réalisation de ce mémoire.

Je tiens également à remercier sincèrement le corps enseignant de la LSBA pour leur apport et leur soutien tout au long de cette aventure intellectuelle. Leurs contributions ont apporté une dimension supplémentaire à mes recherches.

Enfin, ma reconnaissance s'étend à ma famille et à mes amis, dont le soutien moral et émotionnel a été un pilier solide pendant ces moments d'intense engagement académique.

Ce mémoire n'aurait pas pu être réalisé sans le concours de ces individus exceptionnels, et je leur adresse mes remerciements les plus sincères.

# Table des figures

2.5.1	distribution du nombre de sinistre $N_i$ et loi de probabilité des coûts $C_{ik}$ . . .	26
2.5.2	distribution du coût annuel $Z_i$ pour une femme (à droite) et un homme(à gauche)	27
2.5.3	distribution du coût total partagé $X_i$ pour une femme (à droite) et un homme(à gauche) avec $\delta = 4$ . . . . .	27
2.5.4	distribution du coût total $S_n = \sum_{i=1}^n X_i$ pour $n = 1000$ participants, $\delta = 4$ . .	28
2.5.5	Contribution $h_{i,n}^*(s) = E[X_i S = s]$ pour un homme et une femme et probabilités mis à l'échelle de $S_n$ . $n = 1500$ , $\delta = 4$ et 60% d'hommes. . . . .	29
2.5.6	$h_{i,n}^*(s)$ et $h_{i,n}^{reg}(s)$ pour un homme et une femme de 30 ans, $n = 1500$ . . . . .	30
2.5.7	Loi de probabilité des $h_{i,n}^*(S_n)$ en fonction de $n$ . . . . .	31
2.5.8	Convergence des cotisations $h_{i,n}^*(s)$ vers la primes pure pour un homme de 30 ans.	32
2.5.9	Convergence des cotisations $h_{i,n}^*(s)$ vers la primes pure pour une femme de 30 ans. . . . .	32
2.5.10	Loi de probabilité des $h_{i,n}^*(S_n)$ en fonction de $\delta$ , $n = 1500$ et $p = 0.6$ . . . . .	33
2.5.11	Influence de la franchise sur les cotisations, $n = 1500$ et $p = 0.6$ . . . . .	33
2.5.12	Influence de la proportion d'homme sur la distribution des cotisations $\mathbb{E}[X_i S_n]$ , $n = 1500$ et $\delta = 4$ . . . . .	34
2.5.13	Variabilité des cotisations pour un homme en fonction de la proportion d'homme dans le groupe, $n = 1500$ et $\delta = 4$ . . . . .	35
2.5.14	Variabilité des cotisations pour une femme en fonction de la proportion d'homme dans le groupe, $n = 1500$ et $\delta = 4$ . . . . .	35
3.6.1	Niveau des cotisation en fonction de l'âge et du sexe dans le groupe. . . . .	43
3.6.2	Somme des contributions dans le pool. . . . .	43
3.6.3	Loi de probabilité du cash-back $B_{i,n} : n = 1500, \theta_{SL} = 0.5, \beta = 0.85$ . . . . .	45
3.6.4	Effet de la rétention sur les cotisations des groupes 1,2 et 3. . . . .	46
3.6.5	Effet de la rétention sur les cotisations des groupes 4,5 et 6. . . . .	46
3.6.6	Effet de la rétention sur les cotisations des groupes 7,8 et 9. . . . .	47
3.6.7	Effet de la rétention sur les cotisations des groupes 10,11 et 12 sachant. . . .	47
3.6.8	Effet de la rétention sur les cotisations des groupes 13 et 14. . . . .	48
3.6.9	Variation de la rétention et effet sur les cotisations individuelles sachant que $S_n < \omega_n$ . . . . .	48

# Table des matières

<b>Résumé</b>	<b>i</b>
<b>Abstract</b>	<b>i</b>
<b>Remerciement</b>	<b>ii</b>
<b>Table des figures</b>	<b>iii</b>
<b>Introduction</b>	<b>1</b>
<b>1 Présentation de Medi-share et généralités</b>	<b>3</b>
1.1 Introduction . . . . .	3
1.2 Les différents types d'assurances collaboratives . . . . .	3
1.2.1 Modèle autonome ("self-governing model") . . . . .	4
1.2.2 Modèle de courtage : ("broker model") . . . . .	5
1.2.3 P2P carrier model . . . . .	6
1.3 Présentation de Medi-share . . . . .	6
1.3.1 Candidature : Conditions . . . . .	7
1.3.2 Franchise annuelle : AHP (annual household portion) . . . . .	7
1.3.3 Déclarer une facture . . . . .	7
1.3.4 Paiement des Factures admissibles . . . . .	7
1.3.5 Coûts . . . . .	8
1.3.6 Limitation de partage, stage d'attente . . . . .	8
1.3.7 Éligibilité des frais médicaux au partage . . . . .	8
1.3.8 Cas des membres de 65 ans et plus . . . . .	9
1.4 Modèle et Hypothèses de travail . . . . .	9
1.4.1 Hypothèses . . . . .	10
1.4.1.1 Règle d'adhésion . . . . .	10
1.4.1.2 Franchise Annuelle . . . . .	10
1.4.1.3 Stage d'attente . . . . .	10
1.4.1.4 Limitation de partage . . . . .	10
1.4.2 Modélisation . . . . .	11
1.5 Conclusion . . . . .	12

<b>2</b>	<b>Conditional mean risk sharing</b>	<b>13</b>
2.1	Introduction . . . . .	13
2.2	Partage de risque . . . . .	13
2.2.1	Définition du partage de risque . . . . .	14
2.2.2	Propriétés de la règle de partage . . . . .	15
2.2.2.1	Équité actuarielle . . . . .	15
2.2.2.2	Attractivité (willingness to join) . . . . .	15
2.2.2.3	Optimum de Pareto . . . . .	16
2.3	Conditional mean risk-sharing . . . . .	16
2.3.1	Loi des coûts individuels $Z_i$ . . . . .	18
2.3.2	Loi des coûts partagées $X_i$ . . . . .	18
2.3.3	Loi du coût total $S_n$ . . . . .	19
2.3.4	Formulation de la conditional mean risk sharing . . . . .	21
2.4	Approximation des cotisations en grand portefeuille . . . . .	22
2.4.1	Approximation linéaire de la règle de partage . . . . .	22
2.4.2	Lien entre la conditional mean risk sharing et l'approximation linéaire	23
2.4.3	Théorème central limite . . . . .	23
2.5	Application . . . . .	24
2.5.1	Application sur un sous-groupe . . . . .	24
2.5.1.1	Discrétisation des sévérités . . . . .	24
2.5.1.2	Coût total $S_n$ . . . . .	28
2.5.1.3	Contribution respective des Hommes et des Femmes . . . . .	29
2.5.2	Approximation linéaire de la règle de partage . . . . .	30
2.5.2.1	Sensibilité de la règle aux variations du nombre de participants	31
2.5.3	Influence de la franchise $\delta$ . . . . .	32
2.5.4	Influence de la proportion Homme/Femme dans le groupe . . . . .	32
2.6	Conclusion . . . . .	35
<b>3</b>	<b>Étude d'un système avec (re-)assureur</b>	<b>36</b>
3.1	Introduction . . . . .	36
3.2	Décomposition des contributions . . . . .	36
3.3	Définition de la rétention . . . . .	37
3.4	Contribution à la tranche inférieure et Bonus . . . . .	38
3.4.1	Définition de la tranche inférieure . . . . .	38
3.4.2	Calcul du cash-back . . . . .	39
3.5	Définition de la tranche supérieure . . . . .	40
3.6	Application . . . . .	41
3.6.1	Illustration de la décomposition en couche . . . . .	42
3.6.2	Effet de la rétention sur les cotisations dans la tranche inférieure . . .	44
3.6.3	Variation de la valeur de la rétention . . . . .	45
3.6.4	Conclusion . . . . .	49
	<b>Conclusion Générale</b>	<b>50</b>
<b>A</b>	<b>Code R Partie 1</b>	<b>i</b>
A.1	Chargement des packages . . . . .	i

A.2	Définition des paramètres . . . . .	i
A.2.1	Paramètres pour les Hommes . . . . .	i
A.2.2	Paramètres pour les Femmes . . . . .	i
A.2.3	Définition de la composition du portefeuille par âge . . . . .	ii
A.2.4	test . . . . .	ii
A.3	Distribution des fréquences et sévérités . . . . .	ii
A.3.1	Représentation pour un homme et une femme d'âge = 30 ans . . . . .	ii
A.3.2	Mise en évidence de la méthode de discrétisation <i>unbiased</i> . . . . .	iii
A.3.3	Représentation . . . . .	iii
A.4	Distribution du Coût total individuel $Z_i$ . . . . .	iv
A.4.1	Calcul des fonctions de répartitions pour tous les groupes d'âges . . . . .	iv
A.4.2	Calcul de la fonction de répartition pour un homme et pour une femme d'âge égale à 30 ans. . . . .	v
A.4.3	Représentation de la fonction de répartition pour un homme et une femme d'âge = 30 ans . . . . .	v
A.5	Distribution du coût partagé $X_i$ . . . . .	v
A.5.1	Représentation . . . . .	vi
A.6	Distribution de $X_i   X_i > 0$ . . . . .	vi
A.7	Calcul de $S_n$ dans le modèle collectif . . . . .	vii
A.7.1	Calcul des fonctions de répartition des sévérités nécessaires pour $S_n$ et de $S_n - X_i$ dans le modèle collectif . . . . .	vii
A.7.2	Calcul de la fonction de répartition de $\$S\_nS$ : Algorithme de Panjer . . . . .	ix
A.7.3	Test : Essais pour un Groupe composé d'hommes et de femmes âgés de 30 ans. . . . .	ix
A.7.4	Représentation . . . . .	ix
A.7.5	Représentation de la fonction de répartition de $S$ et $S - X_i$ . . . . .	x
A.8	Calcul Direct de la fonction de répartition de $S$ (modèle binomial composé) . . . . .	x
A.8.1	Coût total pour les Femmes . . . . .	x
A.8.2	Coût total pour les Femmes . . . . .	xi
A.8.3	Convolution coût total Hommes + Femmes ==> Coût total pour l'ensemble du groupe . . . . .	xi
A.9	Espérance Conditionnelle : CMRS . . . . .	xii
A.10	Approximation de la CMRS par le modèle de régression linéaire. . . . .	xiii
A.11	Étude de sensibilité de la CMRS . . . . .	xiv
A.11.1	Au nombre de participants $n$ . . . . .	xiv
A.11.1.1	Calcul des cotisations $h_{i,n}^*(s)$ . . . . .	xv
A.11.1.2	Loi de probabilité des $\mathbb{E}[X_i   S_n]$ . . . . .	xv
A.11.1.3	Représentation 1 : Homme Vs Femme . . . . .	xvi
A.11.1.4	Représentation 2 : contribution relatives de chaque groupe . . . . .	xvii
A.11.1.5	Représentation 3 : CMRS Vs Regression linéaire . . . . .	xvii
A.11.1.6	Représentation 4 : Loi de probabilité en fonction de $n$ . . . . .	xvii
A.11.1.7	Représentation 5 : Boxplot . . . . .	xviii
A.11.2	A la franchise . . . . .	xix
A.11.2.1	Cas d'un homme . . . . .	xix
A.11.2.2	Cas d'une femme . . . . .	xx
A.11.2.3	Loi de probabilité des cotisations par rapport à la franchise . . . . .	xxi

A.11.2.4	Représentation 1 : Loi de probabilité en fonction de $\delta$ . . . . .	xxii
A.11.2.5	Représentation 2 . . . . .	xxiii
A.11.3	A la proportion Homme/Femme . . . . .	xxiv
A.11.3.1	Influence de la proportion Homme/Femme sur la distribution des cotisations . . . . .	xxv
A.11.3.2	Représentation 1 : Loi de probabilité en fonction de $p$ . . . . .	xxv
A.11.3.3	Représentation 2 : Boxplot . . . . .	xxvi
<b>B</b>	<b>Code R Partie 2</b>	<b>xxvii</b>
B.1	Coût total individuel $Z_i$ . . . . .	xxvii
B.1.1	Calcul des fonctions $F_Z$ pour les hommes . . . . .	xxvii
B.1.2	Calcul des fonctions $F_Z$ pour les femmes . . . . .	xxvii
B.2	Distribution de $X_i$ . . . . .	xxvii
B.3	Distribution de $X_i   X_i > 0$ . . . . .	xxviii
B.4	Création d'un groupe . . . . .	xxviii
B.5	Modèle collectif . . . . .	xxix
B.5.1	Calcul des fonctions de répartition des sévérités nécessaires pour $S_n$ et de $S_n - X_i$ dans le modèle collectif . . . . .	xxix
B.5.2	Fonction de répartition de $S_n$ : Algorithme de Panjer . . . . .	xxx
B.5.3	Essais sur l'ensemble des participant . . . . .	xxx
B.6	Étude des Cotisations . . . . .	xxxi
B.6.1	Sensibilité au nombre de participants $n$ . . . . .	xxxi
B.6.1.1	Représentation 1 : Homme Vs Femme . . . . .	xxxii
B.6.1.2	Loi de probabilité des $\mathbb{E}[X_i   S_n]$ . . . . .	xxxii
B.6.1.3	Représentation 2 : Variabilité des cotisations en l'absence de (re-)assurance . . . . .	xxxiii
B.7	Application 1 : Décomposition pour un groupe de $n = 1000$ . . . . .	xxxiii
B.7.1	Loi du Cash-Back $B_n$ . . . . .	xxxiv
B.7.2	Représentation 3 : Loi de probabilité du cash-back $B_{i,n}$ . . . . .	xxxv
B.8	Variabilité des cotisations dans tranche inférieure $[0, \omega_n]$ . . . . .	xxxvi
B.8.1	Comparaison avec une variation du nombre de participant . . . . .	xxxvi
B.8.2	Représentation 4 : Variabilité des cotisations dans la tranche inférieure	xxxvii
B.8.3	Représentation 5 : Variabilité des cotisations dans la tranche supérieure	xxxvii
B.8.4	Représentation 5 : Box plot rétention + nombre de participants . . . . .	xxxviii
B.8.5	Variabilité par rapport au niveau de rétention $\omega$ . . . . .	xxxix
B.8.6	Représentation 6 : Effet de la variation de la valeur de la rétention . . . . .	xl



# Introduction

L'essor des nouvelles technologies au cours des dernières années, les évolutions comportementales ainsi que la croissance des demandes en matière de personnalisation des services ont générés des changements significatifs dans l'industrie en général et dans le secteur de l'assurance en particulier. Dans ce contexte, les assurances collaboratives, encore connus sous le nom d'assurances peer-to-peer (P2P), ont émergées comme une alternative novatrice aux assurances traditionnelles. Ce qui fait la particularité de ce nouveau système est leur approche décentralisée, où les participants se regroupent pour partager collectivement leurs risques. Ce nouveau modèle de partage de risques entre pairs offre une perspective nouvelle et révolutionnaire pour le secteur de l'assurance en transformant fondamentalement le système de gestion de risque ainsi que la répartition de ces derniers entre les participants.

L'assurance traditionnelle semble ne plus être aussi attrayante qu'auparavant, du fait d'une démutualisation progressive du secteur de l'assurance qui a rompu le lien avec le mécanisme ancestral d'entraides consistant à utiliser les contributions du plus grand nombre pour compenser les pertes de quelques-uns. Sous le couvert de l'assurance collaborative, on assiste à la renaissance de ce vieux concept sous une forme plus moderne et numérique, éliminant au passage l'intermédiation que requiert une assurance classique. La valeur des primes pour une couverture d'assurance classique peut parfois atteindre des sommes importantes en raison des exigences de solvabilité imposée aux assureurs par le régulateur et de ce fait du coût de la détention de ce capital de solvabilité, des marges bénéficiaires ainsi que des autres frais afférents à l'activité d'assurance. Il arrive souvent que des personnes à haut risque ne soient pas en mesure de trouver une couverture d'assurance adaptée à leur besoin à cause des primes qui sont parfois très élevées. C'est le cas par exemple en assurance maladies qui ne prennent pas souvent compte des affections antérieures. Outre l'aspect purement financier, il y a un volet social où le sentiment d'appartenance et la poursuite d'objectifs sociaux contribuent également à pousser les pairs à s'auto-assurer.

La tarification des produits d'assurance, et le mécanisme d'indemnisation généralement perçu comme étant peu clair et inaccessible par le consommateur lambda ont contribué à instaurer un climat de méfiance à l'égard du secteur de l'assurance. À l'inverse, l'assurance collaborative met en avant la clarté de ses règles de coopération pour regagner la confiance des consommateurs et faire revenir au cœur de l'activité le principe fondamental de mutualité à l'origine de l'assurance. Plus précisément, les participants de l'assurance collaborative peuvent s'accorder pour partager ensemble la première tranche du risque auquel ils font face, tandis que les grosses pertes sont quant à elles cédées soit à un assureur ou alors un (ré-)assureur.

En pratique, le pool conserve une partie des sinistres en dessous d'une certaine rétention et la partie au-delà est transférée à un (ré-)assureur par le biais d'un contrat de type stop-loss, limitant ainsi le montant des contributions versées par les participants. En fin de période s'il reste encore des fonds non utilisés, un cash-back peut être accordé aux participants ou alors versé au profit d'une œuvre de charité ou encore utilisé pour financer un projet commun auquel la communauté adhère. Le principal avantage d'un tel système est qu'il permet aux participants de prétendre à des rétentions beaucoup plus élevées que celles incluses dans les produits d'assurances classiques.

Il existe plusieurs modèles d'assurance collaborative à travers le monde, et chacun de ces modèles se distingue des autres principalement par le niveau de responsabilité que les participants assument, notamment en ce qui concerne la gouvernance et la couverture des sinistres auxquels le pool fait face. Nous essaierons ainsi de proposer dans ce travail, une modélisation actuarielle d'une plate-forme d'assurance collaborative qui pourrait potentiellement être lancée sur le marché Belge. Partant de Medi-share, une plate-forme de partage des coûts des soins de santé basée aux États-Unis d'Amérique, dont le modèle n'implique aucun calcul actuariel, il s'agit de proposer un modèle qui se base justement sur les calculs actuariels pour déterminer de quelle manière sont repartis les coûts entre les différents acteurs du système. On peut à ce propos se poser les questions suivantes : Comment répartir équitablement, de manière compréhensible et transparente les coûts mis en jeu entre les différents participants ? Dans l'éventualité où le pool ferait appel à un (re-)assureur pour la couverture d'une partie de leur frais médicaux, quelle en serait l'impact financier pour les participants ? La formule proposée est-elle viable, en comparaison à l'assurance classique ?

Pour essayer de répondre à ces questions, nous aborderons ce travail en trois étapes. - Tout d'abord au chapitre un, nous ferons une présentation des différents modèles d'assurance collaborative qui existent et plus précisément, nous présenterons Medi-share. - Au chapitre deux, nous présentons la conditional mean risk sharing, une règle de partage permettant de répartir équitablement les coût des soins de santé entre les participants. Un exemple numérique est proposé pour montrer comment les calculs sont fait. - Enfin, au chapitre trois, nous étudierons la possibilité d'un transfert d'une partie du coût total vers un (re-)assureur afin de limiter l'impact financier sur les participants du pool.

# Chapitre 1

## Présentation de Medi-share et généralités

### 1.1 Introduction

Alors que les économies collaboratives ont connu une forte croissance ces dernières années dans le monde et en Belgique en particulier, l'assurance collaborative encore connue sous le nom d'assurance pair à pair (Peer-to-Peer ou P2P) reste pour sa part très peu connue du public. Pourtant tout comme Airbnb ou encore Blablacar qui sont elles aussi des plates-formes d'économie collaboratives, l'assurance collaborative présente un réel atout pour séduire les consommateurs. Ce concept peut offrir de nombreux avantages aux consommateurs tels que des cotisations moins élevées, une meilleure protection des membres les plus à risque et une plus grande transparence dans la gestion des sinistres. Aux États Unis par exemple dans le domaine de la santé, il existe une multitude d'organismes qui par leur notoriété ont mis en place des systèmes similaires permettant à leurs membres de faire pot commun afin de partager entre eux les dépenses relatives à leurs soins médicaux. Plus connus sous le nom de HCSM (health care sharing ministry), il s'agit d'organismes regroupant des membres qui partagent les mêmes convictions religieuses et s'engagent à suivre les principes de santé basés sur la Bible. Dans le présent chapitre, nous nous focaliserons sur l'un de ces organismes notamment Medi-share, qui nous servira de point de départ pour proposer notre modèle actuarielle d'assurance collaborative pour le marché belge.

### 1.2 Les différents types d'assurances collaboratives

Bien qu'étant très peu médiatisé, voir même inconnue du grand public, l'assurance collaborative a elle aussi sa carte à jouer dans un contexte où l'économie collaborative est en pleine évolution. L'idée sous-jacente est de regrouper les individus qui mettent en commun leurs risques et s'assurent eux-mêmes ou alors auprès d'un (ré)assureur pour la tranche supérieure de leur risque.

Tout comme on peut partager sa voiture, sa maison, ou encore tout autre bien, on

pourrait très bien se demander pourquoi ne pas se mettre en commun pour se couvrir mutuellement contre un risque ? En suivant ce principe, on peut s'imaginer de nombreuses possibilités en termes d'assurance par exemple pour la santé, l'automobile, l'habitation, les soins vétérinaires, etc. De telles solutions ont déjà été implémenté en France avec par exemple Inspeer et Otherwise, ou encore en Allemagne avec Friendsurance.

Plusieurs modèles sont ainsi possibles en termes d'organisation pour les assurances P2P. La principale différence réside au niveau de la responsabilité des assurés en ce qui concerne la couverture de leurs sinistres. La plupart des plate-formes existantes, proposant une assurance collaborative fonctionnent suivant le modèle de courtier "broker model", mais on distingue également le modèle autonome "self-governing model" et enfin le "carrier model".

### 1.2.1 Modèle autonome ("self-governing model")

Le modèle est constitué d'une communauté qui s'auto-organise, s'auto-régule, partage les risques et gère toutes les fonctions d'assurance sans avoir recours à aucun intermédiaire ; chaque décision est prise par une procédure de vote. Les participants gardent un plus grand degré de contrôle sur la gestion et l'administration de la couverture. Dans ce modèle, les membres du groupe décident collectivement des règles et des procédures du groupe et assument également la responsabilité de la gestion et de l'administration de la couverture d'assurance. Le premier modèle de ce genre a été lancé en Chine en 2015 avec TongJuBao. Les groupes sont créés par la plate-forme numérique sur la base d'affinités ou d'intérêts communs. Les membres sont regroupés en fonction du type d'assurance (vie, santé, automobile), du statut social, de la profession, de la région de résidence, de la composition familiale et d'autres caractéristiques pour former des groupes de risque homogènes. Les membres entrent sans payer de commission initiale mais signent un contrat qui les oblige à contribuer ex-post aux pertes totales réalisées par le groupe. Ce modèle peut être vu comme de la mutualisation pure où le risque est entièrement partagé entre les participants sans recours à une assurance.

Dans le domaine de la finance et plus précisément de la finance décentralisée (DeFi), il existe des solutions similaires destinées à couvrir les risques sous-jacents liés à cette activité. Nexus Mutual est une plate-forme d'assurance qui offre aux institutions et aux particuliers la possibilité d'acheter des polices d'assurance qui protègent contre les risques uniques très souvent rencontrés dans les plate-formes de finance décentralisée (DeFi) et de finance centralisée (CeFi) tels que les défaillances des *smart contracts* : des programmes autonomes capables d'exécuter les conditions du contrat, rendant tout intermédiaire caduque ; les piratages et autres types risques. Contrairement aux compagnies d'assurance traditionnelles, Nexus Mutual est gérée par ses membres qui détiennent des jetons NXM, leur permettant de participer à la gouvernance de la plate-forme et de détenir des droits de vote sur les décisions importantes. Les membres peuvent par exemple acheter une couverture pour protéger leurs actifs numériques ; souscrire des polices d'assurance en tant que évaluateurs de risques ; examiner la validité des réclamations en tant que évaluateurs de réclamations. Les primes d'assurance sont déterminées par les membres de la communauté en fonction des risques évalués par les modèles d'analyse des risques. En cas de sinistre, les membres peuvent soumettre une demande de réclamation qui sera examinée et votée par les autres membres. Si la demande est approuvée, le paiement de l'indemnisation sera effectué en utilisant la crypto-monnaie Ether. La plate-forme vise à fournir une alternative décentralisée et transparente aux compagnies d'assurance traditionnelles, tout

en offrant des produits d'assurance mutuelle personnalisés et abordables pour les utilisateurs de la blockchain.

L'avantage du modèle autonome est qu'il permet un plus grand degré de flexibilité et de contrôle pour les membres du groupe. Le groupe peut décider collectivement des termes et conditions de la couverture, ainsi que des règles de traitement des demandes d'indemnisation, ce qui peut conduire à un processus plus transparent et équitable. En outre, ce modèle permet au groupe de prendre des décisions et de s'adapter aux changements de circonstances plus rapidement et plus efficacement qu'une compagnie d'assurance traditionnelle. Toutefois, les membres du groupe peuvent avoir différents niveaux d'expertise et de connaissances en matière d'assurance, ce qui peut entraîner des désaccords ou des malentendus. De plus, les membres du groupe assument une plus grande responsabilité dans la gestion et l'administration de la couverture, ce qui peut être long et complexe.

### 1.2.2 Modèle de courtage : ("broker model")

Ce modèle permet aux participants de se regrouper et de partager les risques et les coûts de leur couverture d'assurance à travers une plate-forme intermédiaire. La plate-forme qui fait office d'intermédiaire met en relation des groupes de personnes ayant des besoins d'assurance similaires. Ces groupes mettent ensuite en commun leurs ressources pour acheter collectivement une couverture d'assurance. Une partie des primes est placée dans un fond commun et l'autre partie est versée chez un assureur ou un réassureur. S'il arrive que le fond soit insuffisant pour couvrir les sinistres du groupe, alors l'assureur ou le réassureur se charge de payer l'excédent au-delà de la rétention du groupe. Si au contraire les sinistres sont moindres par rapport au niveau du fond, alors le surplus est redistribué au profit des membres du groupe.

Otherwise qui est l'un des pionniers de l'assurance collaborative sur le marché français fonctionne suivant ce modèle. La plate-forme met en relation des personnes présentant des profils de risque similaires et les organise en groupes de partage autour d'un même produit. Elle propose une gamme de produits comprenant une assurance santé complémentaire, une assurance auto et une assurance pour animaux. Les produits sont conçus pour être les plus flexibles possibles et personnalisables de sorte que les clients peuvent choisir la couverture qui convient le mieux à leurs besoins. Le système est entièrement digitalisé, ce qui permet de simplifier et d'améliorer l'expérience de l'assurance pour ses clients. De plus, en traitant directement les petits sinistres, elle vise à réduire les primes exigées pour des couvertures d'assurance traditionnelles. Les mécanismes coopératifs selon lesquels moins les sinistres réglés sont nombreux et plus la valeur du cashback redistribué aux membres est élevée, réduisent le phénomène de l'aléa moral, les conséquences de la sélection adverse, la fraude et les inefficacités

Ce modèle permet aux participants d'accéder à une couverture qu'ils ne pourraient potentiellement pas se permettre s'ils étaient seuls. En partageant leur franchise, les membres du groupe peuvent accéder à une couverture à un coût moindre que s'ils devaient l'acheter individuellement. Il permet également un processus de gestion des demandes d'indemnisation plus transparent et plus équitable.

### 1.2.3 P2P carrier model

C'est un modèle hybride qui combine les principes de l'assurance collaborative avec celles d'une compagnie d'assurance traditionnelle. Les mêmes petits groupes partagent leur risque en payant conjointement des primes et au lieu d'un simple transfert de risque à la compagnie d'assurance, les participants peuvent prendre une partie du risque global du groupe auquel ils ont choisi d'adhérer. En fin de période, s'il reste de l'argent non réclamé, les membres se partagent les fonds restants à titre de remboursement. La plate-forme digitale fonctionne comme une compagnie d'assurance agréée, souscrit les polices et gère les sinistres comme un assureur traditionnel.

Lancé aux États-Unis en 2015, Lemonade est un exemple type de cette approche. Par rapport aux assureurs traditionnels, Lemonade propose une relation client entièrement digitalisée et d'une efficacité impressionnante. Le modèle commercial de Lemonade se distingue de celle des compagnies d'assurance traditionnelles par l'utilisation de l'intelligence artificielle pour offrir des produits d'assurance plus rapides et plus abordables. Grâce à la plate-forme numérique, les assurés peuvent souscrire à la police choisie en quelques minutes en utilisant leur smartphone ou leur ordinateur et grâce aux chatbots mis en place, les délais de règlement des sinistres sont tout aussi courts. Les frais sont fixes et transparents ; les assurés partagent une partie du risque global du groupe auquel ils ont choisi d'adhérer et une fraction des primes non utilisées est reversée à des œuvres de charité choisie par les clients. Lemonade approfondit l'auto-apprentissage et les technologies de l'information pour créer des groupes homogènes, satisfaire les clients, favoriser l'efficacité et tirer parti de l'approche traditionnelle de l'assurance et de sa capacité à supporter les risques.

Ce modèle combine les avantages de l'assurance P2P, tels que des coûts réduits, des traitements plus transparents et équitables des sinistres, avec la structure et la stabilité d'une compagnie d'assurance traditionnelle. En effet, l'assureur fournit des garanties et des capacités de réassurance supplémentaires pour les risques couverts par la plate-forme. Il fournit également une expertise en souscription, en gestion des sinistres et en réglementation, ce qui peut aider la plate-forme P2P à fonctionner plus efficacement. Cela offre un plus grand niveau de sécurité et de fiabilité aux assurés.

## 1.3 Présentation de Medi-share

Medi-share est une plate-forme d'assurance collaborative basée en Amérique, c'est l'un des programmes de l'organisme Christian Care Ministry dont le rôle est d'organiser le partage des coûts des soins de santé entre ses différents membres. Elle fait partie de ce qu'on appelle les HCSM "Health Care Sharing Ministry", qui sont des organismes dont les membres ont en commun des même croyances religieuses ou éthiques et s'accordent pour partager ensemble les coûts de leurs frais médicaux .

Tout comme dans le cas d'une assurance traditionnelle, les membres paient des cotisations mensuelles et doivent payer de leur poche un montant annuel prédéfini avant de voir leurs frais médicaux soumis au partage. Toutefois, il ne s'agit pas réellement d'une assurance, puisqu'il n'existe aucune obligation légale tenant l'organisme aussi bien que ses membres à payer pour les frais médicaux de l'un de ses membres. Contrairement aux autres types d'assurance

peer-to-peer, il existe un unique groupe qui englobe l'ensemble des membres. Les membres participent à la création et à l'amendement des règles du groupe par vote soit directement soit par l'intermédiaire d'un comité constitué de membres mandatés à cet effet.

### **1.3.1 Candidature : Conditions**

Medi-share calque essentiellement son fonctionnement sur des principes bibliques, de ce fait tout postulant âgé d'au moins 18 ans doit attester de sa foi, ce qui peut parfois passer par un entretien avec un leader de l'église. L'une des principales conditions d'acceptation au sein du groupe est la profession de foi que chaque membre adulte doit dire à l'adhésion. Il existe entre autres des conditions qui exigent aux membres d'observer un style de vie sain bannissant toute consommation de drogue ou d'alcool aussi bien que la cigarette, ou avoir arrêté au moins douze mois avant la candidature.

### **1.3.2 Franchise annuelle : AHP (annual household portion)**

À l'adhésion, les nouveaux membres doivent choisir un niveau de AHP (annual household portion). L'AHP fonctionne comme une franchise annuelle dans le cas d'une assurance traditionnelle, chaque membre paie tout de sa poche jusqu'à ce qu'il atteigne cette franchise annuelle après quoi Medi-share paie tout le reste et pour le reste de l'année. Chaque membre ouvre par la suite un compte au sein d'un organisme bancaire, le America's Christian Credit Union (ACCU), qui est le partenaire financier de Medi-share. Ce compte bancaire est destiné à assurer le transfert des fonds d'une part entre les membres qui déposent chaque mois leur cotisation et les autres membres ayant fait des réclamations.

### **1.3.3 Déclarer une facture**

Lorsqu'un participant ou un membre de sa famille a besoin de soins, celui-ci rencontre son médecin ou alors un médecin agréé de Medi-share, et c'est celui-ci qui se charge de déclarer la facture à Medi-share pour le partage si elle est jugée admissible. Cependant les membres qui rencontrent leur propre médecin pour les soins, si ceux-ci ne font pas partis de la liste de médecins agréés, le montant de la facture sera limité au coût standard du service sollicité, et le montant en excès restera à la charge du participant. Pour les prestations dans un hôpital non agréé, le participant est responsable d'un montant supplémentaire de 20% de sa facture, mais qui n'est pas comptabilisé pour atteindre le AHP.

### **1.3.4 Paiement des Factures admissibles**

Si le montant de la AHP a été atteint au moment de la soumission de la facture, Medi-share se charge de la négociation d'éventuelles réductions et ensuite paie l'intégralité de la facture à partir des cotisations mensuelles des autres membres qui ont fait le partage. Cependant si le montant AHP n'a pas été atteint, Medi-share se charge juste de la négociation des réductions et ensuite il revient au participant de payer le pourcentage de la facture qui lui revient. Ce pourcentage de la facture que paie le participant lui permet d'atteindre son niveau de AHP sur l'année, et une fois atteint, la facture est effectivement soumise au partage pour remboursement.

### 1.3.5 Coûts

A la souscription, les membres choisissent un niveau de la AHP compris entre \$3000 à \$12000. Le montant des cotisations mensuelles dépend cependant de la composition de la famille, de l'âge, du statut matrimonial, du montant de AHP choisis. Ces cotisations tiennent lieu de primes mensuelles, bien que n'en n'étant pas et servent à couvrir les coûts des soins de santé des autres membres dans le besoin. Pour un niveau de AHP de \$3000, le montant mensuel de cotisations est d'environ \$284 pour une personne seule, \$506 pour un couple sans enfant, et \$898 pour une famille. Au niveau de AHP de \$6000, on a environ \$210 pour une personne seule, \$377 pour un couple et \$685 pour une famille. Avec un AHP de \$9000, on aura à peu près \$170 de cotisations mensuelles pour une personne seule, \$299 pour un couple et \$523 pour une famille. Et en fin pour un AHP de \$12000, on aura environ \$107 pour une personne seule, \$211 pour un couple et \$362 pour une famille. Ces montants sont approximatifs car peuvent varier en fonction de l'âge des participants, ainsi que du nombre de personnes dans le ménage .

### 1.3.6 Limitation de partage, stage d'attente

Les membres bénéficient du partage des factures médicales sans limite annuelle et à vie, toutefois, il existe quelques exceptions pour les conditions préexistantes, la maternité, les accidents de moto, et pendant le premier mois d'adhésion.

- Premier mois : les membres peuvent recevoir remboursements d'un montant maximum de \$50000 correspondant aux frais médicaux soumis au partage pendant le premier mois d'adhésion
- Conditions préexistantes : les frais médicaux relatifs au diagnostic ou au traitement d'une maladie préexistante dont le diagnostic a été fait dans les 36 mois précédant l'adhésion peuvent bénéficier du partage pour un montant maximum de \$100000 par année si le membre a contribué aux partages pendant 36 mois d'affilées, ou alors pour un montant maximum de \$500000 si le membre a contribué pendant 60 mois consécutifs
- une limite de \$125000 est appliquée pour une seule grossesse à condition que le couple ait contribué au partage depuis le mois de conception jusqu'au mois de l'accouchement sans interruption.
- Si un membre est victime d'un accident de moto, il peut bénéficier selon les conditions d'un montant maximum de \$100000 sur une période de 12 mois.

Pour des raisons de simplicité, nous ne tiendrons pas compte de toutes ces limitations dans le cadre de ce travail.

### 1.3.7 Éligibilité des frais médicaux au partage

L'éligibilité d'une facture n'est déterminée qu'une fois le soin effectué, sauf dans les cas où le soin a été pré-notifié. L'éligibilité est déterminée à partir des informations sur le style de vie, ainsi qu'à partir d'un historique médical couvrant au moins 36 mois avant l'adhésion suivant la nature des soins relatifs à la facture en question.

### 1.3.8 Cas des membres de 65 ans et plus

Les candidats de 65 ans ou plus ne sont pas éligible au programme Medi-share, toutefois il existe un autre programme spécialement conçu pour eux appelé Senior Assist. Lorsqu'un membre atteint 65 ans il peut choisir de continuer avec Medi-share ou faire la transition vers senior assist si celui-ci est préalablement couvert d'une assurance hospitalisation (medicare partie A) ou d'une assurance maladie (medicare partie B). Medicare est le système d'assurance santé géré par le gouvernement fédéral des états unis. Ceux n'ayant aucune couverture medicare (partie A et B) peuvent rester dans Medi-share.

Les membres atteignant 65 ans et qui décident de rester dans Medi-share doivent continuer à payer leur contribution. Si le membre a une assurance medicare, alors Medi-share sera secondaire, dans ce cas le partage des facture éligibles encourues pendant ou après le premier jour du mois où le membre a atteint ses 65 ans sera basée sur la différence entre les frais admissibles par medicare et le montant effectivement versé par medicare.

## 1.4 Modèle et Hypothèses de travail

En Belgique, l'assurance maladie obligatoire relève de la sécurité sociale qui prend en charge une part importante du coût des hospitalisations. Toutefois, sa couverture n'est pas complète dans la mesure où elle laisse le plus souvent une part assez importante des frais à la charge des patients, par exemple si ceux-ci ont opté pour une chambre individuelle, les coûts peuvent être encore plus élevés en raison de la politique des suppléments d'honoraires pratiquée par certains hôpitaux belges.

Les prestations proposées par les assureurs sont très souvent indemnitaires. Les frais liés à une hospitalisation pour cause de maladie, d'accident ou à la suite d'un accouchement peuvent être élevés, dans ce cas, l'assureur s'engage à prendre en charge une partie ou l'entièreté de la facture d'hôpital après intervention de l'assurance maladie obligatoire. Pour faire face au coût souvent important et imprévu d'un séjour à l'hôpital, le patient peut souscrire (pour lui-même et, s'il le souhaite, aussi pour les membres de son ménage) une assurance hospitalisation complémentaire auprès d'un assureur privé ou d'une mutualité. L'assurance hospitalisation est souvent souscrite individuellement, mais il arrive que l'employeur couvre ses employés (et, s'ils le désirent, les membres de leur ménage) au moyen d'une assurance collective.

Afin de réduire les coûts à la charge des patients, plusieurs polices d'assurances privée sont nées, couvrant ainsi totalement ou partiellement les frais laissés à la charge de l'assuré en cas d'hospitalisation. Dès lors, nous voulons proposer dans ce travail un modèle actuariel pour concevoir une assurance collaborative destinée à couvrir les coûts supplémentaires de l'assurance maladie obligatoire laissés à la charge des assurés en faisant appel au principe de la mutualisation. Le but étant de proposer une alternative beaucoup moins coûteuse qu'une assurance traditionnelle, mais également dont le mécanisme de calcul des cotisations sont beaucoup plus transparents à l'égard des assurés.

L'idée est à partir du fonctionnement de Medi-share, de proposer une formule similaire qui pourrait être lancé sur le marché Belge, mais qui utilise les calculs actuariels pour déterminer

la manière dont le coût total des frais médicaux de l'ensemble peut être répartis entre tous les participants et la nécessité de recourir à un réassureur pour les cas où les cotisations sont trop volatiles d'une année à une autre. Nous considérerons pour cela quelques hypothèses simplifiées inspiré de celles de Medi-share et que nous énoncerons plus bas.

### **1.4.1 Hypothèses**

#### **1.4.1.1 Règle d'adhésion**

On suppose pour ce qui est de l'adhésion au système que les conditions préexistantes ne sont pas prises en compte. En effet, leur prise en compte implique un traitement particulier des cas où l'hospitalisation  $y$  est liée, ce qui peuvent rendre les calculs plus complexes. Il n'y a pas de distinction faites sur les causes d'hospitalisation, exception faite des conditions préexistantes qui sont d'office exclues. De plus nous n'imposerons pas de restrictions quant à l'âge à l'entrée également.

#### **1.4.1.2 Franchise Annuelle**

Le AHP joue le rôle de la franchise annuelle en dessous de laquelle chaque participant est tenu de payer de sa poche pour ses frais médicaux. Il s'agit d'un moyen de responsabiliser les participant en diminuant l'aléa moral. Les participants auront la possibilité d'opter pour un niveau de franchise parmi une liste préétablie. Cependant, nous considérerons dans un premier temps que tous les participants optent pour un niveau de franchise unique et plus tard, nous pourrions éventuellement étendre les niveaux possibles de franchise.

#### **1.4.1.3 Stage d'attente**

Il n'existe pas de stage d'attente dans le fonctionnement de Medi-share, mais plutôt une limitation de paiement pour les factures soumises pendant le premier mois d'adhésion. Il pourrait s'avérer plus intéressant et plus prudent d'instaurer un stage d'attente durant lequel aucune prestation n'est accordée à l'adhérent et donc aucune facture soumise au partage par exemple pendant le premier mois suivant l'adhésion. Ceci permettrait d'éviter des cas de fraude et de contrer l'antisélection. Cependant la prise en compte du stage d'attente n'a pas un impact explicite dans la modélisation du coût annuel, mais on pourrait très bien s'imaginer que celui-ci influe sur le nombre de factures déclarées dans l'année.

#### **1.4.1.4 Limitation de partage**

Nous n'imposerons pas de limitation au niveau des montants partagés par les participants puisqu'il n'y a pas de distinction entre les causes d'entrée en hospitalisation pouvant nécessiter un traitement différencié des factures liées à certaines causes spécifiques. Pour les femmes par exemple, on supposera que les hospitalisations considérées peuvent indifféremment être liées à une grossesse ou toutes autres causes autres que la grossesse. La prise en compte des causes d'hospitalisations à un impact à la fois sur le nombre moyen d'hospitalisation et aussi sur le coût moyen de cette hospitalisation. Pour les femmes, ces paramètres contiennent globalement les hospitalisations pour cause de grossesse et les autres

### 1.4.2 Modélisation

Afin de modéliser le fonctionnement de la plate-forme, nous allons considérer un ensemble de  $n$  individus que nous allons repérer par l'indice  $i$ ,  $i = 1, 2, \dots, n$  et qui ensemble forment un pool afin de partager entre eux leurs frais médicaux. Notons  $Z_i$  la variable aléatoire non négative représentant le montant total annuel des frais médicaux que le participant  $i$  subit durant l'année. Les variables  $Z_i$ , représentant la sinistralité par participant sont telle que  $P(Z_i = 0) > 0$ .

Nous adopterons l'approche de l'assurance dommage pour représenter le coût total par participant. En supposant que la sinistralité suit une loi Poisson composée avec des sévérités de lois Gamma, on peut écrire :

$$Z_i = \sum_{k=1}^{N_i} C_{ik} \quad \text{avec} \quad N_i \sim Poi(\lambda_i) \quad \text{et} \quad C_{ik} \sim Gamma(\alpha_i, \beta_i), i = 1, 2, \dots, n \quad (1.4.1)$$

Toutes les variables aléatoires sont supposées indépendantes, avec :

$N_i$  = le nombre d'évènement affectant le participant  $i$

$C_{ik}$  = le coût du  $k$ -ème évènement affectant le participant  $i$ .

On pourrait très bien envisager un système où on isolerait certains types d'hospitalisations et dans ce cas-là, le coût total par participant se décomposera comme suit

$$Z_i = \sum_{k=1}^{N_{i1}} C_{ik}^{(1)} + \sum_{k=1}^{N_{i2}} C_{ik}^{(2)} + \dots \quad (1.4.2)$$

Où  $N_{i1}, N_{i2}, \dots$  sont des variables aléatoires représentant les nombres d'hospitalisation relative à chaque type considéré et les  $C_{ik}^{(j)}$  les coûts correspondants, on suppose que toutes ces variables aléatoires sont indépendantes.

La décomposition (1.4.2) est idéale si on souhaite accorder un traitement particulier à certains types d'hospitalisation telles que les grossesses, les hospitalisations relatives à une affection préexistante ou encore les hospitalisations consécutives à un accident de moto. Toutefois, dans le présent travail, nous considérerons qu'il n'existe pas de différence entre les causes d'hospitalisation.

Plus précisément, si on note  $\delta_i$  la valeur de la franchise annuelle choisie par le participant  $i$  alors, la responsabilité du groupe n'est engagée que lorsque le coût total des soins de ce participant durant l'année entière excède la franchise choisie et dans ce cas le groupe paie pour la tranche supérieure. Le montant soumis au partage par le participant  $i$  est donné par :

$$X_i = (Z_i - \delta_i)_+ = \max\{Z_i - \delta_i, 0\} \quad (1.4.3)$$

Où  $\delta_i$  représente la franchise annuelle.

## 1.5 Conclusion

Dans ce chapitre, nous avons passé en revue les différents types d'assurance collaboratives qui existent sur le marché ainsi que leur mode de fonctionnement, il s'agit notamment du modèle auto-gouverné ou modèle autonome, du modèle de courtage où le groupe transfère une partie du risque à un assureur, et le "carrier model" qui combine les principes d'une assurance collaborative à celle d'une compagnie d'assurance classique. Nous avons également présenté Medi-share qui est le point de départ de notre modélisation. Un modèle actuariel de ce fait a été proposé sur base d'hypothèses simplificatrices que nous avons considérées. Le modèle de Medi-share tel que défini ne repose sur aucun principe actuariel pour le calcul des cotisations que chaque participant doit payer. Cependant à fin de présenter un système plus transparent vis-à-vis des membres, nous nous proposons de déterminer dans la suite à travers les calculs actuariels une méthode d'attribution des coûts entre les membres qui soit plus transparente, mais aussi équitable pour tous.

## Chapitre 2

# Conditional mean risk sharing

### 2.1 Introduction

Dans un système d'assurance collaborative, le partage de risque est un principe fondamentale dont le but est de définir de quelle manière seront répartis les coûts et les pertes associés aux sinistres entre tous les participants du groupe. Il existe plusieurs manières différentes d'opérer ce partage au sein d'une assurance collaborative, mais cela dépend des règles et des mécanismes spécifiques mis en place par le groupe ainsi que des objectifs visés. Dans ce chapitre nous présentons la conditional mean risk sharing définie par DENUIT et DHAENE 2012 qui présente l'avantage d'être transparente et est relativement facile à communiquer aux participants car repose sur le concept familier de la moyenne. Sous cette règle, chaque participant paie comme contribution au pool, l'espérance conditionnelle des pertes qu'il a subies, compte tenu de la perte totale subie par l'ensemble du pool. Nous étudierons également le comportement asymptotique de cette règle en grand portefeuille à fin d'en extraire un équivalent asymptotique de la conditional mean risk sharing.

### 2.2 Partage de risque

En situation de partage de risque, chaque participant du pool reçoit une compensation du pool pour ses frais médicaux, et en retour ce dernier paie une cotisation au groupe de sorte que l'ensemble des cotisations reçues par le groupe soit égal au coût total des frais médicaux de l'ensemble.

Considérons le pool formé des  $n$  individus numérotés  $i = 1, 2, 3, \dots, n$  soumis chacun à une perte dont la valeur est représentée par la variable aléatoire  $Z_i$  et donnant le montant total des frais médicaux dû au participant  $i$  durant l'année. Soit  $X_i$  le montant effectivement mutualisé par le participant  $i$  après application de la franchise annuelle. Les variables aléatoires  $X_i$  sont supposées non négatives, indépendantes et d'espérances finies ;

$$\mu_i = \mathbb{E}[X_i] > 0 \quad \text{et} \quad \sigma_i^2 = \text{Var}[X_i] > 0 \quad (2.2.1)$$

En tenant compte de la franchise annuelle  $\delta_i$  propre à chaque participant, le montant soumis

au partage par le participant  $i$  se réduit à  $\max\{Z_i - \delta_i, 0\}$ . On définit  $X_i = \max\{Z_i - \delta_i, 0\}$  et de ce fait le montant total des frais médicaux au sein du groupe en fin de période est :

$$S_n = \sum_{i=1}^n \max\{Z_i - \delta_i, 0\} = \sum_{i=1}^n X_i \quad (2.2.2)$$

Avec  $Z_i$  tel que défini à l'équation (1.4.1).

En l'absence d'un assureur ou encore d'un pool, chaque individu endosse la charge que représente le coût de ses soins de santé alors qu'en présence d'un pool, les participants peuvent décider de partager leur facture ce qui peut être bénéfique. L'objectif du partage de risque tel que défini par DENUIT, DHAENE et ROBERT 2022 est de définir une façon optimale de répartir le coût total  $S_n$  observé en fin de période entre tous les participant du groupe.

### 2.2.1 Définition du partage de risque

**Définition 2.2.1** (Partage de risque). Soit le pool  $\mathbf{X}_n = (X_1, X_2, \dots, X_n)$ . Le partage de risque est un processus à deux étapes, consistant dans un premier temps à l'instant 0, en une réallocation des pertes du pool vers un vecteur  $\mathbf{h}_n(\mathbf{X}_n)$  de même dimension et défini par :

$$\mathbf{h}_n(\mathbf{X}_n) = (h_{1,n}(\mathbf{X}_n), h_{2,n}(\mathbf{X}_n), \dots, h_{n,n}(\mathbf{X}_n)) \quad (2.2.3)$$

avec  $h_{i,n}(\mathbf{X}_n) \geq 0$ ,  $i = 1, 2, \dots, n$  et vérifiant la propriété d'allocation totale ( la somme des pertes réallouées doit correspondre à la somme des pertes observées) :

$$\sum_i^n h_{i,n}(\mathbf{X}_n) = \sum_i^n X_i \quad (2.2.4)$$

Dans un second temps, en fin de période, Lorsque la réalisation  $\mathbf{x}_n = (x_1, x_2, \dots, x_n)$  de  $\mathbf{X}_n$  est observée, chaque participant du pool reçoit la réalisation  $x_i$  de son coût total  $X_i$  et en retour contribue la réalisation  $h_{i,n}(\mathbf{x}_n)$  de la réallocation de sa perte  $h_{i,n}(\mathbf{X}_n)$

**Définition 2.2.2** (Règle de partage). La règle de partage  $\mathbf{h}_n$  est une correspondance qui transforme un pool  $\mathbf{X}_n$  en un autre vecteur aléatoire  $\mathbf{h}_n(\mathbf{X}_n)$  de même dimension pour laquelle les composantes  $h_{i,n}(\mathbf{X}_n)$  sont toutes non-négatives et satisfont la condition de réallocation totale (2.2.4).

Dans la suite on notera plus simplement  $\mathbf{h}_n(\mathbf{X}_n) = \mathbf{h}_n(S_n)$ . Ceci définit les règles de partage "agrégées", qui ne dépendent pas expressément de  $X_1, X_2, \dots, X_n$ , mais seulement de leur somme  $S_n$ .

Soit  $x_i$  une réalisation de  $X_i$  et soit  $s = \sum_{i=1}^n x_i$  une réalisation de  $S_n$ . Un cas particulier de partage de risque est le cas d'un pool où tous les  $X_i$  sont indépendant et identiquement distribués, dans ce cas, une loi raisonnable de partage du pool peut être définie par la règle  $h_{i,n}(s) = s/n$  pour  $i = 1, 2, 3 \dots, n$ , ce qui impliquerait que tous les participant d'un tel pool paieront un même montant. Si par contre les  $X_i$  sont non homogènes, alors cette règle pourrait s'avérer inefficace dans la mesure où par exemple deux participant subissant des coûts  $X_1$  et  $X_2$  et tels que  $\mathbb{E}[X_1] = \mathbb{E}[X_2]$  et  $Var[X_1] < Var[X_2]$ , devront payer la même cotisation, ce

qui n'est pas équitable d'un point de vue actuarielle. Par équité actuarielle, on entend que la contribution moyenne d'un participant soit égale au coût moyen des soins que ce participant réclame au groupe. L'équité actuarielle est entre autres l'une des caractéristiques principales exigée de la règle de partage. Nous devons donc définir une règle qui prenne en compte cet aspect qui plus est un principe phare de l'assurance collaborative.

## 2.2.2 Propriétés de la règle de partage

L'un des objectifs du partage de risque est de répartir les coûts de manière équitable et efficace entre les différents participants. Elle doit de ce fait inciter les parties concernées à prendre des mesures conformes à l'objectif global de gestion du risque dans le groupe. La règle doit donc respecter des exigences qui pour certaines s'inspirent directement des principes de calculs des primes ou des mesures de risque, alors que d'autres sont adaptées aux règles de partage de risques.

Nous reprenons ci-après quelques-unes de ces propriétés dont une liste plus complète a été proposée et analysée par DENUIT, DHAENE et ROBERT 2022.

### 2.2.2.1 Équité actuarielle

Cette propriété signifie que la règle se doit d'être juste et équitable pour tous les participants, en veillant à attribuer à chacun une contribution appropriée. Par définition, une règle de partage équitable est telle qu'en moyenne, aucun participant ne fasse gain ou perte du partage. Autrement dit, la contribution moyenne du participant  $i$  dans le pool doit être égale au coût moyen de ses soins telle qu'il aurait payé s'il était seul, déduction faite de la franchise. Cela se traduit par

$$\mathbb{E}[h_{i,n}(S_n)] = \mathbb{E}[X_i] \quad i = 1, 2, 3, \dots, n \quad (2.2.5)$$

### 2.2.2.2 Attractivité (willingness to join)

Cette propriété permet d'exprimer la volonté d'un participant à adhérer au groupe. En réalité, la règle se doit d'être telle qu'un individu  $i$  préfère payer les cotisations  $h_{i,n}(S_n)$  plutôt que de subir seul le coût  $X_i$  de ses soins. Cependant, même si la règle n'est pas optimale pour tous les participants (dans le sens où elle ne prend que partiellement les préférences des participants en compte), elle se doit d'être attractive pour tous les agents de la même classe telle que par exemple la classe des agents risquophobes.

La préférence dont il est question ici est définie au sens de l'ordre convexe ( $\preceq_{Cx}$ ). Rappelons que pour un agent soumis à deux risques  $X$  et  $Y$ ,  $X$  est préféré à  $Y$  ( $X \preceq_{Cx} Y$ ) si l'on a simultanément  $\mathbb{E}[u(\kappa - X)] \geq \mathbb{E}[u(\kappa - Y)]$  et  $\mathbb{E}[X] = \mathbb{E}[Y]$  pour toute fonction d'utilité  $u$  concave et non-décroissante et une richesse initiale  $\kappa$ . Ceci revient à dire  $X \preceq_{Cx} Y \iff \mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)]$  pour toute fonction convexe  $g$  telle que les espérances existent.

La règle de partage  $h_{i,n}(S_n)$  est attractive si

$$h_{i,n}(S_n) \preceq_{Cx} X_i \quad i = 1, 2, 3, \dots, n \quad (2.2.6)$$

L'inégalité stochastique  $X \preceq_{Cx} Y$  signifie de façon intuitive que  $X$  et  $Y$  ont la même "taille" (puisque  $\mathbb{E}[X] = \mathbb{E}[Y]$ ), mais que  $X$  est moins volatile que  $Y$ . Par exemple  $Y$  a une variance qui est plus grande que celle de  $X$ .

### 2.2.2.3 Optimum de Pareto

Un partage de risque  $(h_{1,n}^*(\mathbf{X}_n), h_{2,n}^*(\mathbf{X}_n), \dots, h_{n,n}^*(\mathbf{X}_n))$  est Pareto optimal s'il n'existe aucun autre partage de risque  $(h_{1,n}(\mathbf{X}_n), h_{2,n}(\mathbf{X}_n), \dots, h_{n,n}(\mathbf{X}_n))$  pour laquelle l'inégalité

$$h_{i,n}(\mathbf{X}_n) \preceq_{Cx} h_{i,n}^*(\mathbf{X}_n) \quad (2.2.7)$$

soit vraie pour tout  $i = 1, 2, \dots, n$  avec au moins une de ces inégalités convexes qui soit stricte. Par conséquent, un partage de risque Pareto optimal est une allocation sans alternative, c'est à dire que tous les participants sont dans un état tel qu'il soit impossible de strictement améliorer la situation de l'un des participants (au sens de l'ordre convexe) sans toutefois détériorer celle d'un autre.

Dans le cas particulier où le vecteur aléatoire  $\mathbf{X}_n$  est comonotonique, alors le partage de risque  $(h_{1,n}(\mathbf{X}_n), h_{2,n}(\mathbf{X}_n), \dots, h_{n,n}(\mathbf{X}_n))$  est Pareto optimal. Rappelons que  $\mathbf{X}_n = (X_1, X_2, \dots, X_n)$  est comonotonique s'il existe une variable aléatoire  $Z$  et des fonctions  $g_i$  strictement croissantes telles que  $\mathbf{X}_n$  est identiquement distribué à  $(g_1(Z), g_2(Z), \dots, g_n(Z))$ .

## 2.3 Conditional mean risk-sharing

Dans le cas où les  $X_i$  sont inhomogènes, La conditional mean risk sharing est une meilleure alternative à la règle de partage homogène  $h_{i,n}(s) = s/n$ . Elle bénéficie de plusieurs propriétés théoriques intéressantes telles que par exemple l'équité actuarielles, l'attractivité (voir DENUIT, DHAENE et ROBERT 2022 ), et se révèle de ce fait particulièrement efficace pour le calcul des cotisations dans le cadre des assurances collaboratives.

La conditional mean risk sharing est parmi toutes les règles de partage du coût total  $S_n$ , celle qui minimise l'écart moyen quadratique entre les coûts individuels partagés  $X_i$  et le montant des cotisations  $h_{i,n}(S_n)$ , c'est à dire

$$\mathbb{E}[(X_i - h_{i,n}^*(S_n))^2] = \min_{h(\cdot): Var[h(S_n)] < \infty} \mathbb{E}[(X_i - h(S_n))^2] \quad (2.3.1)$$

La solution à ce problème est donnée par :

$$h_{i,n}^*(S_n) = \mathbb{E}[X_i | S_n] \quad i = 1, 2, 3, \dots, n \quad (2.3.2)$$

Autrement dit, la contribution de chaque participant  $i$  au coût total  $S_n$  est égale à la valeur moyenne du coût total des frais médicaux qu'il ramène dans le groupe, sachant le coût total des factures de l'ensemble du groupe. Cette règle permet une répartition complète du coût total entre tous les participant, puisqu'on a bien

$$\sum_{i=1}^n h_{i,n}^*(S_n) = \sum_{i=1}^n \mathbb{E}[X_i | S_n] = S_n$$

Connaissant la réalisation  $s$  de  $S_n$ , la conditional mean risk sharing de l'équation (2.3.2) s'écrit

$$h_{i,n}^*(s) = \mathbb{E}[X_i | S_n = s] \quad (2.3.3)$$

La règle est aussi équitable car en moyenne la contribution de chaque participant est égale au coût moyen des frais médicaux que celui-ci ramène dans le pool, on a

$$\mathbb{E}[h_{i,n}^*(S_n)] = \mathbb{E}[\mathbb{E}[X_i | S_n]] = \mu_i \quad i = 1, 2, 3, \dots, n$$

DENUIT et DHAENE 2012 montrent que pour les participants du pool il existe toujours un mécanisme de partage des risques mutuellement bénéfique (par rapport à l'ordre convexe ( $\preceq_{Cx}$ )). Et si les participants recourent à la conditional mean risk sharing, alors ils améliorent leur situation au sens de l'ordre convexe, c'est-à-dire qu'ils augmentent leurs utilités espérées respectives (en supposant qu'ils soient averses au risque). Autrement dit

$$h_{i,n}^*(S_n) \preceq_{Cx} X_i \quad i = 1, 2, 3, \dots, n$$

Cette relation peut également traduire l'attractivité du partage par la conditional mean risk sharing. En effet, si on considère une fonction convexe  $g$ , alors on a par application de l'inégalité de Jensen :

$$\begin{aligned} \mathbb{E}[g(h_{i,n}^*(S_n))] &= \mathbb{E}[g(\mathbb{E}[X_i | S_n])] \\ &\leq \mathbb{E}[\mathbb{E}[g(X_i) | S_n]] = \mathbb{E}[g(X_i)] \quad i = 1, 2, 3, \dots, n \end{aligned}$$

Par définition de l'ordre convexe, on déduit bien que  $h_{i,n}^*(S_n) \preceq_{Cx} X_i$ . La conditional mean risk sharing est une alternative intéressante à l'assurance conventionnelle avec des primes fixes mais à condition que la volatilité des cotisations  $h_{i,n}^*(S_n)$  reste en accord avec l'appétit au risque du participant  $i$ . De plus, en restant dans le cadre de la théorie de l'utilité, DENUIT et ROBERT 2021b montrent que l'augmentation du nombre de participants est toujours considérée comme bénéfique par tous les agents économiques averses au risque, quelle que soit la distribution des pertes qu'ils apportent au pool, tant que celles-ci sont mutuellement indépendantes et que le partage s'effectue selon (2.3.2) :

$$h_{i,n+1}^*(S_{n+1}) \preceq_{Cx} h_{i,n}^*(S_n), \quad \text{pour } i \in \{1, 2, \dots, n\} \quad (2.3.4)$$

De plus, si  $h_{1,n}^*(S_n), h_{2,n}^*(S_n), \dots, h_{n,n}^*(S_n)$  sont comonotoniques, c'est à dire si  $s \mapsto h_{i,n}^*(s) = \mathbb{E}[X_i | S_n = s]$  est non-décroissante pour tout  $i = 1, 2, \dots, n$ , alors la conditional mean risk sharing est Pareto optimale. La croissance des fonctions  $h_{i,n}^*(s)$  est une exigence raisonnable dans le cas de l'assurance collaborative car elle garantit que les participants auront un intérêt à ce que les pertes totales du groupe soient aussi faibles que possible au risque de voir les cotisations grimper.

### 2.3.1 Loi des coûts individuels $Z_i$

Reprenons l'expression du coût total donné par (1.4.1) avec tous les  $C_{ik} > 0$ . Supposons que les variables aléatoires  $C_{ik}$  prennent des valeurs discrètes, c'est à dire dans  $\{1, 2, 3, \dots\}$ . C'est généralement le cas dans les calculs actuariels où les coût sont exprimés en unité monétaire après les avoir discrétisé.

La loi de Poisson composée donnant les coûts  $Z_i$  peut être obtenue aisément à l'aide de l'algorithme de Panjer. En effet, soit  $C_i$  une variable aléatoire identiquement distribuée à  $C_{ik}$ . La fonction de répartition de  $Z_i$  est donnée par

$$F_{Z_i}(k) = P[Z_i \leq k]$$

$$F_{Z_i}(k) = \sum_{j=0}^k g_{ij}, \quad k \geq 0. \quad (2.3.5)$$

Où les  $g_{ij} = P[Z_i = j]$  sont calculés de façon récursive grâce à l'algorithme de Panjer. Partant de  $P[Z_i = 0] = \exp(-\lambda_i)$ , l'algorithme s'écrit

$$P[Z_i = j] = \frac{\lambda_i}{j} \sum_{l=1}^j l P[C_i = l] P[Z_i = j - l], \quad j = 1, 2, 3 \dots \quad (2.3.6)$$

Nous utiliserons dans la suite le package `actuar` de R, qui permet d'obtenir aisément la distribution d'une variable de Poisson-composée grâce à la fonction `aggregateDist()`. L'application du calcul de la fonction de répartition de  $Z_i$  par l'algorithme de Panjer peut être trouvée dans le code en annexe à la section A.4.

### 2.3.2 Loi des coûts partagées $X_i$

Après application de la franchise, chaque participant rapporte au pool  $X_i = \max\{Z_i - \delta_i, 0\}$ , où  $\delta_i$  désigne la franchise annuelle du participant  $i$  on a :

$$X_i = \begin{cases} Z_i - \delta_i & \text{si } Z_i > \delta_i \\ 0 & \text{si } Z_i \leq \delta_i \end{cases} \quad (2.3.7)$$

Le coût des frais médicaux soumis par le participant  $i$  peut encore s'exprimer sous la forme

$$X_i = Y_i I_i \quad \text{avec} \quad Y_i =_d Z_i - \delta_i | Z_i > \delta_i \quad (2.3.8)$$

Où  $I_i$  vaut 1 si le montant des frais médicaux du participant  $i$  est supérieure à sa franchise et 0 sinon ; et  $Y_i$  est la valeur des frais médicaux du participant  $i$  à la charge du groupe, après déduction de la franchise et sachant que le total de ses frais médicaux est au-delà de la franchise. Les variables  $I_1, I_2, \dots, I_n, Y_1, Y_2, \dots, Y_n$  sont supposées mutuellement indépendantes. On a

$$X_i = \begin{cases} Y_i & \text{avec probabilité } q_i = P[Z_i > \delta_i] = P[X_i > 0] \\ 0 & \text{avec probabilité } p_i = P[Z_i \leq \delta_i] \end{cases} \quad (2.3.9)$$

Connaissant la loi de  $Z_i$  donnée à l'équation (2.3.5), on obtient aisément l'expression de la fonction de répartition des  $X_i$

$$F_{X_i}(k) = F_{Z_i}(k + \delta_i) \quad k \in \{0, 1, 2, \dots\} \quad (2.3.10)$$

### 2.3.3 Loi du coût total $S_n$

La fonction de répartition de  $X_i$  est  $F_{X_i}(x) = P[X_i \leq k]$  et on a  $F_{X_i}(0) = p_i < 1$ . Notons  $B_i(k) = P[X_i \leq k | X_i > 0]$  ( donc  $B_i$  est la fonction de répartition de  $Y_i$  avec  $Y_i =_d [X_i | X_i > 0]$ ), on a

$$B_i(k) = \frac{F_{X_i}(k) - F_{X_i}(0)}{1 - F_{X_i}(0)}, \quad k \geq 0 \quad (2.3.11)$$

En tenant compte de la relation (2.3.10), on peut réécrire (2.3.11) en fonction de la fonction de répartition de  $Z_i$

$$B_i(k) = \frac{F_{Z_i}(k + \delta_i) - F_{Z_i}(\delta_i)}{1 - F_{Z_i}(\delta_i)}, \quad k \geq 0 \quad (2.3.12)$$

Soit  $F_{X_i}^{*(n)}$  le produit de convolution d'ordre  $n$  de  $F_{X_i}$  défini par

$$F_{X_i}^{*(n)}(k) = P[X_1 + X_2 + \dots + X_n \leq k]$$

Partant de la condition initiale  $F_{X_i}^{*(1)} = F_{X_i}(k)$ , on déduit de manière exacte la fonction de répartition  $F$  de  $S_n$  qui est donnée par

$$F(k) = P[S_n \leq k] = \sum_{n=0}^{\infty} P[N = n] F_{X_i}^{*(n)}(k), \quad k \geq 0. \quad (2.3.13)$$

Cependant le calcul de  $F$  selon cette approche est assez difficile car le calcul intermédiaire des produits de convolution ainsi que la somme sur  $n$ , rendent cette formule très lourde et gourmande en temps de calcul. Le passage au modèle collectif se révèle nécessaire et va de ce fait nous permettre de contourner ce problème.

En utilisant l'approximation du modèle collectif pour le coût total, on peut approcher  $S_n$  par

$$S_n^c = \sum_{k=1}^N Y_k \sim \mathcal{CPoi}(\lambda_{\bullet}^c, L) \quad (2.3.14)$$

Avec  $N \sim \text{Pois}(\lambda_{\bullet}^c)$  et  $Y_1, Y_2, \dots$  des variables aléatoires indépendantes et identiquement distribuées à  $Y$  tels que,

$$\lambda_{\bullet}^c = \sum_{i=1}^n \lambda_i^c \quad \text{et} \quad L(k) = P[Y \leq k] = \frac{1}{\lambda_{\bullet}^c} \sum_{i=1}^n \lambda_i^c B_i(k) \quad (2.3.15)$$

La loi de  $S_n^c$  pourra par la suite être déterminée numériquement en utilisant l'algorithme de Panjer. La mise en œuvre du modèle collectif requiert de faire un choix approprié pour les paramètres  $\lambda_i^c$ . Il existe de ce fait deux standards pour choisir ces paramètres.

**Choix 1 :**  $\lambda_i^c = q_i$

Ce choix permet d'avoir l'égalité entre le nombre moyen de factures qu'un participant déclare dans le modèle individuel et le nombre moyen de factures déclarées dans le modèle collectif, c'est à dire

$$\sum_{i=1}^n q_i = \sum_{i=1}^n \lambda_i^c = \lambda_{\bullet}^c = \mathbb{E}[N]$$

Avec le choix  $\lambda_i^c = q_i$ , le coût total moyen des factures du groupe est le même dans les deux modèles, en effet,

$$\begin{aligned} \mathbb{E}[S_n] &= \sum_{i=1}^n \mathbb{E}[X_i] \\ &= \sum_{i=1}^n q_i \mathbb{E}[Y_i] = \mathbb{E}[S_n^c] \end{aligned}$$

En revanche, le modèle collectif surestime la variance du coût total car

$$\begin{aligned} \text{Var}[S_n] &= \sum_{i=1}^n \text{Var}[X_i] \\ &= \sum_{i=1}^n \left( \mathbb{E}[\text{Var}[X_i|I_i]] + \mathbb{E}[\text{Var}[X_i|I_i]] \right) \\ &= \sum_{i=1}^n \left( q_i \text{Var}[Y_i] + q_i(1 - q_i)(\mathbb{E}[Y_i])^2 \right) \\ &= \sum_{i=1}^n \left( q_i \mathbb{E}[Y_i^2] - q_i^2 (\mathbb{E}[Y_i])^2 \right) \\ &\leq \sum_{i=1}^n q_i \mathbb{E}[Y_i^2] = \text{Var}[S_n^c] \end{aligned}$$

Avec ce choix de paramètres, le modèle collectif est donc prudent car sous ce modèle on travaille avec un coût total ayant une plus grande variabilité par rapport à celui du modèle individuel. En plus de l'inégalité des variances, on peut montrer que  $S_n \leq_{Cx} S_n^c$ . Ce choix est approprié pour le calcul de primes stop-loss.

**Choix 2 :**  $\lambda_i^c = -\ln(1 - q_i)$

Avec ce choix, on s'assure que la probabilité qu'un participant ne déclare pas de facture est la même dans les deux modèles, on a

$$\begin{aligned}\lambda_i^c = -\ln(1 - q_i) &\iff \exp(-\lambda_i^c) = 1 - q_i \\ &\iff P[N_i = 0] = P[X_i = 0]\end{aligned}$$

Il en est de même au niveau de l'ensemble du groupe où  $P[S_n = 0] = P[S_n^c = 0]$ . De plus  $\mathbb{E}[S_n] < \mathbb{E}[S_n^c]$  et  $Var[S_n] \leq Var[S_n^c]$ .

Ce deuxième choix est prudent et plus adapté quand on veut évaluer la solvabilité ou alors un capital sous risque.

Puisque l'on veut que la règle de partage respecte l'équité actuarielle, il est donc important que l'approximation que nous choisissons pour le coût total  $S_n$  et donc le choix des  $\lambda_i^c$  permette une conservation de la moyenne au niveau de l'ensemble du groupe c'est à dire  $\mathbb{E}[S_n] = \mathbb{E}[S_n^c]$ . Pour atteindre ce résultat, le choix propice pour les  $\lambda_i^c$  est donc le premier c'est à dire  $\lambda_i^c = q_i$  pour  $i \in \{1, 2, \dots, n\}$ .

De la même manière, la variable  $\sum_{j \neq i} X_j = S_n - X_i$  peut être approchée dans le modèle collectif par  $\sum_{k=1}^M T_{ik}$  avec  $M \sim Poisson(\lambda_\bullet - \lambda_i)$  et  $T_{i1}, T_{i2}, \dots$  étant des variables aléatoires indépendantes et identiquement distribuée à  $T_i$ , avec

$$P[T_i = k] = \sum_{j \neq i} \frac{q_j}{\lambda_\bullet^c - q_i} B_j(k), \quad k = 0, 1, \dots \quad (2.3.16)$$

Nous sommes à présent en mesure de calculer toutes les fonctions de répartition des variables aléatoires nécessaires au calcul des cotisations des différents participants. Le calcul des  $\lambda_i$  ainsi que celles des fonctions de répartition de  $S_n$  et  $S_n - X_i$  sont données dans le code en annexes à la section B.5.1.

### 2.3.4 Formulation de la conditional mean risk sharing

Revenons à la règle de partage (2.3.3), on a clairement  $\mathbb{E}[X_i | S_n = 0] = 0$  et pour  $s \geq 0$ , on a

$$\begin{aligned}\mathbb{E}[X_i | S_n = s] &= \sum_{k=1}^s k P[X_i = k | S_n = s] \\ &= \sum_{k=1}^s k \frac{P[X_i = k] P[\sum_{j \neq i} X_j = s - k]}{P[S_n = s]}\end{aligned}$$

Or pour  $k > 0$ , on sait que

$$P[X_i = k] = P[Z_i = k + \delta_i]$$

On en déduit donc que

$$\mathbb{E}[X_i | S_n = s] = \sum_{k=1}^s k \frac{P[Z_i = k + \delta_i] P[\sum_{j \neq i} X_j = s - k]}{P[S_n = s]} \quad (2.3.17)$$

Nous disposons dès lors, de tous les ingrédients nécessaires au calcul de la contribution de chaque participant au coût total des frais médicaux observé en fin d'année. Dans la suite, nous calculerons numériquement toutes ces fonctions de répartitions à fin d'appliquer la règle sur un groupe composé de  $n$  individus, le code R pour ce calcul est donné à l'annexe A à la section A.9.

## 2.4 Approximation des cotisations en grand portefeuille

L'idée de l'approximation en grand portefeuille est de déterminer comment fluctuent les cotisations individuelles autour de la prime pure lorsque le groupe devient suffisamment grand.

### 2.4.1 Approximation linéaire de la règle de partage

Dans la conception d'un système reconnu et fiable, il est important que la règle de partage représentée par les fonctions  $h_{i,n}$  soit à la fois acceptable et transparente. De ce fait, les règles de partages linéaires de la forme

$$h_{i,n}^{lin}(S_n) = \mathbb{E}[X_i] + a_{i,n}(S_n - \mathbb{E}[S_n]), \quad i = 1, 2, \dots, n \quad \text{avec} \quad \sum_{i=1}^n a_{i,n} = 1$$

sont particulièrement intéressantes. Cette règle satisfait la propriété de réallocation totale (2.2.4) et est également équitable puisque  $\mathbb{E}[h_{i,n}^{lin}(S_n)] = \mathbb{E}[X_i]$ . Sous cette règle, les membres du groupe acceptent de payer la prime pure  $\mathbb{E}[X_i]$  et de se partager l'écart du coût total  $S_n$  par rapport à sa moyenne  $\mathbb{E}[S_n]$ , en proportion de  $a_{i,n}$ .

L'approximation linéaire de la règle de partage est celle qui minimise entre toutes les règles linéaires  $h_{i,n}^{lin}$ , l'écart quadratique moyen entre les coûts individuels  $X_i$  et les cotisations par participant  $h_{i,n}(S_n)$ , c'est à dire

$$\mathbb{E}\left[\left(X_i - h_{i,n}^{reg}(S_n)\right)^2\right] = \min_a \mathbb{E}\left[\left(X_i - \mathbb{E}[X_i] - a(S_n - \mathbb{E}[S_n])\right)^2\right]$$

La solution à ce problème est

$$a = \frac{Cov[X_i, S_n]}{Var[S_n]} = \frac{\sigma_i^2}{\sum_{j=1}^n \sigma_j^2} = a_{i,n}^{reg}$$

et la règle correspondante est donnée par

$$h_{i,n}^{reg}(S_n) = \mu_i + \frac{\sigma_i^2}{\sum_{j=1}^n \sigma_j^2} (S_n - \mathbb{E}[S_n]), \quad i = 1, 2, \dots, n \quad (2.4.1)$$

Sous cette nouvelle règle, les membres du groupe répartissent l'écart de  $S_n$  par rapport à la prime pure totale  $\mathbb{E}[S_n]$  en fonction de la volatilité relative du risque qu'ils apportent au pool. La seule différence entre  $h_{i,n}^{reg}$  et  $h_{i,n}^*$  réside dans la classe de règle de partage considérée : en effet, si les participants préfèrent que la règle de partage soit linéaire, alors  $h_{i,n}^{reg}$  est la plus appropriée tandis que si ceux-ci autorisent à ce que la règle soit aussi non linéaire, alors la conditional mean risk sharing (2.3.3) sera la plus adaptée. Dans tous les deux cas, les deux règles minimisent l'écart quadratique moyen entre  $X_i$  et les cotisations.

## 2.4.2 Lien entre la conditional mean risk sharing et l'approximation linéaire

Une sorte de loi des grands nombres a été établie pour la conditional mean risk sharing afin d'étudier les fluctuations des cotisations individuelles lorsque la taille du pool devient suffisamment large. En effet, DENUIT et ROBERT 2021a montrent sous certaines conditions, que la prime pure individuelle est obtenue à la limite lorsque le nombre de participants dans le groupe tend vers l'infini aussi bien pour la conditional mean risk sharing que pour la règle de partage linéaire, c'est à dire

$$P\left[\lim_{n \rightarrow \infty} h_{i,n}^*(S_n) = \mu_i\right] = 1 \quad \text{et} \quad P\left[\lim_{n \rightarrow \infty} h_{i,n}^{reg}(S_n) = \mu_i\right] = 1 \quad (2.4.2)$$

Le modèle collectif permet d'obtenir une loi pour  $S_n$ , mais pour la répartition des coûts entre les individus, on a besoin de calculer  $\mathbb{E}[X_i|S_n]$ . Cependant, les variables aléatoires intervenant dans la formule (2.3.17) sont bien celles du modèle individuel qui pourtant n'interviennent plus dans l'écriture de  $S_n^c$  à l'équation (2.3.14). L'approximation en grand portefeuille est importante à ce niveau car elle permet de calculer la contribution de chaque participant en se basant uniquement sur la loi de  $S_n$  que l'on a approchée par celle de  $S_n^c$ .

## 2.4.3 Théorème central limite

Si les  $X_1, X_2, \dots, X_n$  sont absolument continues, ou suivent des lois de poisson avec des sévérités absolument continues. Sous certaines conditions telles que énoncées dans DENUIT et ROBERT 2021a, on a :

$$\frac{\tau_n}{\sigma_i^2} \begin{pmatrix} h_{i,n}^{reg}(S_n) - \mathbb{E}[X_i] \\ h_{i,n}^*(S_n) - \mathbb{E}[X_i] \end{pmatrix} \xrightarrow{d} \text{Normal} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right) \quad (2.4.3)$$

avec  $\tau_n^2 = \text{Var}[S_n] = \sum_{j=1}^n \sigma_j^2$ .

Cette propriété montre que l'approximation linéaire et la conditional mean risk sharing sont asymptotiquement équivalentes, en d'autres termes, les fluctuations des cotisations individuelles autour de la prime pure sont à peu près identiques pour  $h_{i,n}^*$  et  $h_{i,n}^{reg}$ . Puisque  $h_{i,n}^*$  et  $h_{i,n}^{reg}$  convergent tous les deux vers  $\mathbb{E}[X_i]$  avec la probabilité 1, alors on conclut que  $h_{i,n}^{reg}$  fournit une bonne approximation de  $h_{i,n}^*$  lorsque la taille du pool est suffisamment grande.

## 2.5 Application

Considérons un groupe formé de  $n$  individus partageant ensemble le coût de leurs soins médicaux. Le coût total des frais médicaux d'un individu  $i \in \{1, 2, 3, \dots, n\}$  en fin de d'année est noté  $X_i$  et le niveau de franchise qui lui est appliqué est  $\delta_i$ .

Pour un produit type d'assurance hospitalisation, les coûts des soins médicaux varient fortement en fonction des paramètres tels que l'âge et le sexe de l'assuré et il en va de même pour le nombre moyen d'hospitalisation qui est différent selon qu'il s'agisse d'un homme ou d'une femme, cette différence est d'autant plus grande que l'âge est avancé. On fait l'hypothèse simpliste qu'il n'existe pas de différence entre les causes d'entrée en hospitalisation.

Les données utilisées à titre d'illustration dans ce chapitre proviennent du cours d'assurance des personnes dispensé à l'UCLouvain par Denuit Michel. Les tableaux 2.5.1 et 2.5.2 reprennent les valeurs des paramètres que nous utiliserons dans la suite pour les illustrations numériques.

Age	20	30	40	50	60	70	80
$\lambda_i$	0.033	0.040	0.067	0.120	0.200	0.305	0.483
$\mathbb{E}[C_{ik}]$	84	100	104	104	108	115	134
$\alpha_i$	0.546	0.546	0.546	0.546	0.546	0.546	0.546
$\beta_i(\times 10^2)$	0.65	0.546	0.525	0.525	0.51	0.475	0.407

TABLE 2.5.1 – Paramètres des lois pour les hommes.

Age	20	30	40	50	60	70	80
$\lambda_i$	0.078	0.220	0.111	0.130	0.167	0.246	0.380
$\mathbb{E}[C_{ik}]$	79	90	96	104	115	133	158
$\alpha_i$	0.754	0.754	0.754	0.754	0.754	0.754	0.754
$\beta_i(\times 10^2)$	0.954	0.838	0.785	0.725	0.656	0.567	0.477

TABLE 2.5.2 – Paramètres des lois pour les femmes.

### 2.5.1 Application sur un sous-groupe

Dans un premier temps, nous considérerons un groupe de  $n$  individu comprenant uniquement des hommes et des femmes tous âgés de 30 ans et dont les paramètres sont repris dans le tableau ci-dessus avec une proportion  $\theta$  d'hommes par rapport aux femmes.

Pour ce travail, nous considérons dans un premier temps que tous les participants sont soumis à un même et unique niveau de franchise, c'est à dire  $\delta_i = \delta$ . Les données des tableaux 2.5.1 et 2.5.2 sont normalisés à 100 pour un homme de 30 ans. La franchise  $\delta$  est fixée par rapport à ce niveau de référence et représente le coût moyen des soins hospitaliers pour un homme de 30 ans. Nous travaillerons essentiellement avec  $\delta = 0.040 \times 100 = 4$ .

#### 2.5.1.1 Discrétisation des sévérités

Le coût total par individu  $X_i = \sum_{k=1}^{N_i} C_{ik}$  est de loi poisson composée avec  $N_i \sim \mathcal{Poi}(\lambda_i)$  avec les sévérités de loi  $\text{Gamma}(\alpha_i, \beta_i)$ . Cependant l'algorithme de Panjer requiert des sévérités

de valeurs discrètes, c'est à dire à valeurs dans  $\{1, 2, 3, \dots\}$ . À partir des paramètres des lois *Gamma* des tableaux 2.5.1 et 2.5.2, l'une des premières étapes pour nos illustrations consiste en la discrétisation des lois *Gamma* des sévérités.

Si on note  $F_{i,k}(x)$  la fonction de répartition à discrétiser sur un certain intervalle  $[a, b]$ ,  $\mathbb{E}[X \wedge x] = \mathbb{E}[\min\{X, x\}]$  et  $f_{i,k}(x)$  la masse de probabilité en  $x$ , alors la discrétisation de  $F_{i,k}$  peut se faire par différente méthode, on peut par exemple citer :

1. Discrétisation par différence progressive :

$$f_{i,k}(x) = F_{i,k}(x + h) - F_{i,k}(x)$$

Pour  $x = a, a + h, \dots, b - h$ . La fonction de répartition calculée par cette méthode est toujours au-dessus de la fonction de répartition réelle.

2. Discrétisation par différence régressive :

$$f_{i,k}(x) = \begin{cases} F_{i,k}(a), & x = a \\ F_{i,k}(x) - F_{i,k}(x - h), & x = a + h, \dots, b. \end{cases}$$

La fonction discrétisée est toujours en dessous de la fonction de répartition réelle.

3. Discrétisation par la méthode du point médian :

$$f_{i,k}(x) = \begin{cases} F_{i,k}(a + h/2), & x = a \\ F_{i,k}(x + h/2) - F_{i,k}(x - h/2), & x = a + h, \dots, b - h. \end{cases}$$

La fonction exacte passe exactement à mi-chemin entre les pas de la fonction de répartition discrétisée.

4. Discrétisation sans biais ou par correspondance du premier moment :

$$f_{i,k}(x) = \begin{cases} \frac{\mathbb{E}[X \wedge a] - \mathbb{E}[X \wedge (a+h)]}{h} + 1 - F_{i,k}(a), & x = a \\ \frac{2\mathbb{E}[X \wedge x] - \mathbb{E}[X \wedge (x-h)] - \mathbb{E}[X \wedge (x+h)]}{h}, & a < x < b \\ \frac{\mathbb{E}[X \wedge a] - \mathbb{E}[X \wedge (b-h)]}{h} - 1 + F_{i,k}(b), & x = b \end{cases}$$

La discrétisation de la fonction de répartition par cette méthode permet que la fonction de répartition discrétisée et la fonction de répartition réelle aient la même probabilité totale et la même espérance sur l'intervalle  $[a, b]$ .

Pour rester cohérent avec les calculs des sections précédentes qui conservaient la moyenne inchangée, c'est cette dernière méthode de discrétisation que nous utiliserons dans la suite. Ces méthodes de discrétisation sont développées et implémentées par GOULET 2008 dans le package R `actuar`, qui contient également de nombreuses fonctions très intéressantes et pratiques pour les calculs actuariels. Le code R relatif à la discrétisation des fonctions de répartition des sévérités peut être retrouvée en annexe à la section A.3.2.

Tout d'abord, nous avons représenté à la figure 2.5.1, les lois de probabilités des composantes fréquences et sévérités des hospitalisations à partir des paramètres retenus, pour les hommes et pour les femmes. Sur les deux sous-figures de droite, la courbe en rouge représente la fonction densité de probabilité exacte de la *Gamma*, tandis que les traits verticaux eux représentent

la masse de probabilité en  $k$ , obtenue après discrétisation de la fonction de répartition. Les données utilisées suggèrent qu'il est plus probable pour un homme de 30 ans au courant d'une année de n'avoir aucune hospitalisation comparée à une femme du même âge.

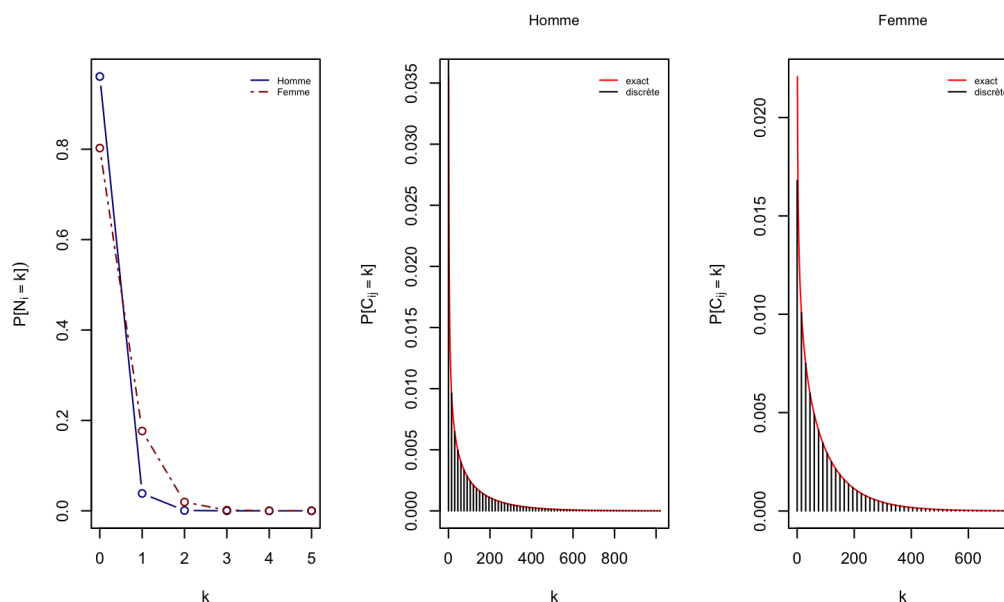


FIGURE 2.5.1 – distribution du nombre de sinistre  $N_i$  et loi de probabilité des coûts  $C_{ik}$

Cependant pour au moins une hospitalisation, la probabilité est plus élevée chez les femmes contrairement aux hommes. Cela peut se justifier, si l'on suppose qu'à cette âge les femmes ont plus souvent tendance à être hospitalisées par rapport aux hommes pour des raisons en lien avec la grossesse et l'accouchement. Notons cependant que les taux d'hospitalisations peuvent aussi être influencés par les facteurs comme le statut socio-économique et le lieu de résidence, qui peuvent varier entre les hommes et les femmes. Par conséquent il convient de dire que ces différences observées entre les sexes ne peuvent pas toujours être attribuées à des différences biologiques entre les sexes. On pourrait dès lors s'attendre à ce que cette différence induise des niveaux de cotisations plus élevés chez les femmes puisque celle-ci sont susceptible de déclarer plus de factures au groupe.

Les distributions des variables aléatoires  $Z_i$  et  $X_i$  sont respectivement représentées aux figures 2.5.2 et 2.5.3. Les points apparaissant sur les courbes de la figures 2.5.3 représentent masses de probabilités en zéro, c'est à dire les probabilités que le participant en question ne réclame aucun remboursement au groupe car dans ce cas le coût total de ses soins médicaux n'auront pas été au-delà de la franchise. Notons que ces masses de probabilité en zéro sont plus petite à la figure 2.5.2, comparé à celles de la figure 2.5.3.

Ainsi, pour un même niveau de franchise donné, il est plus probable que les frais médicaux d'un homme ne dépassent pas la franchise durant toute l'année par rapport à une femme où cela est moins probable.

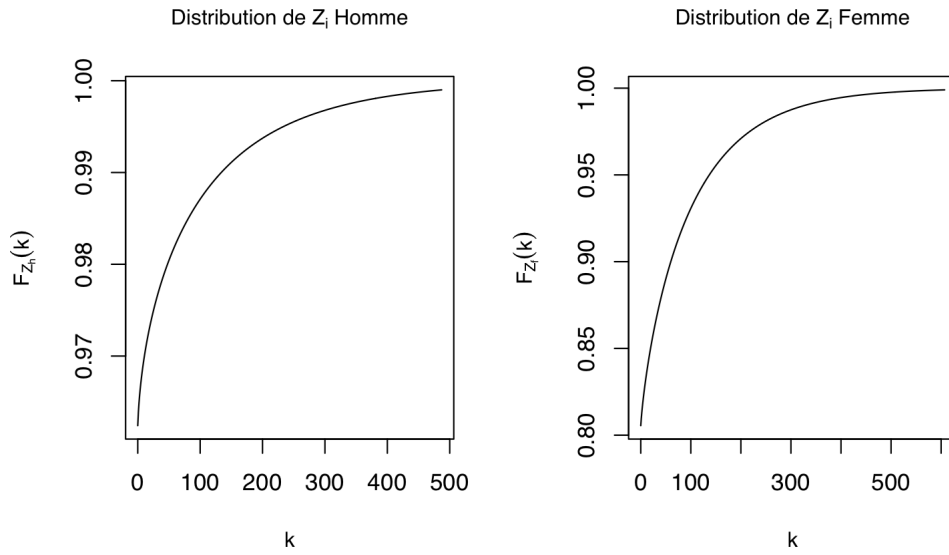


FIGURE 2.5.2 – distribution du coût annuel  $Z_i$  pour une femme (à droite) et un homme(à gauche)

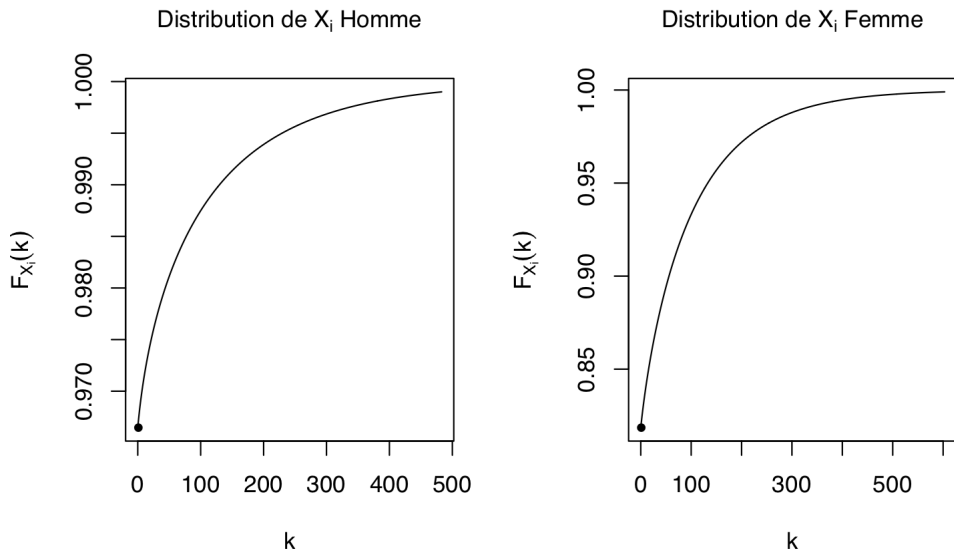


FIGURE 2.5.3 – distribution du coût total partagé  $X_i$  pour une femme (à droite) et un homme(à gauche) avec  $\delta = 4$

### 2.5.1.2 Coût total $S_n$

Pour le calcul exact de la loi de  $S_n$  au niveau de l'ensemble du groupe, on a recours à la loi Binomiale composée pour modéliser le coût total  $S_n$ . Ainsi nous avons tout d'abord scindé notre groupe en deux sous-groupes : celui des hommes et celui des femmes. Notons  $n_h$  et  $n_f$  sont respectivement les nombre d'hommes et de femmes dans le groupe, alors on écrit

$$S_n = S^h + S^f = \sum_{i=1}^{n_h} X_i^h + \sum_{i=1}^{n_f} X_i^f, \quad n = n_h + n_f \quad (2.5.1)$$

avec

$$S^h = \sum_{i=1}^{n_h} X_i^h =_d \sum_{i=1}^{N_h} Y_i^h \quad \text{où } N_h \sim \mathcal{Bin}(n_h, q_i^h) \quad \text{et } q_i^h = P[X_i^h > 0] \quad (2.5.2)$$

et de même nous avons,

$$S^f =_d \sum_{i=1}^{N_f} Y_i^f \quad \text{où } N_f \sim \mathcal{Bin}(n_f, q_i^f) \quad \text{et } q_i^f = P[X_i^f > 0] \quad (2.5.3)$$

Les fonctions de répartition de  $S^f$  et  $S^h$  sont calculées grâce à l'algorithme de Panjer et la fonction de répartition du coût total est quant à elle obtenue par produit de convolution des fonctions de répartition de  $S^f$  et  $S^h$ , soit  $F_{S_n} = F_{S^h} \star F_{S^f}$ . Nous avons implémenté le calcul exact de la fonction de répartition de  $S_n$  dans le code en annexe à la section A.8.

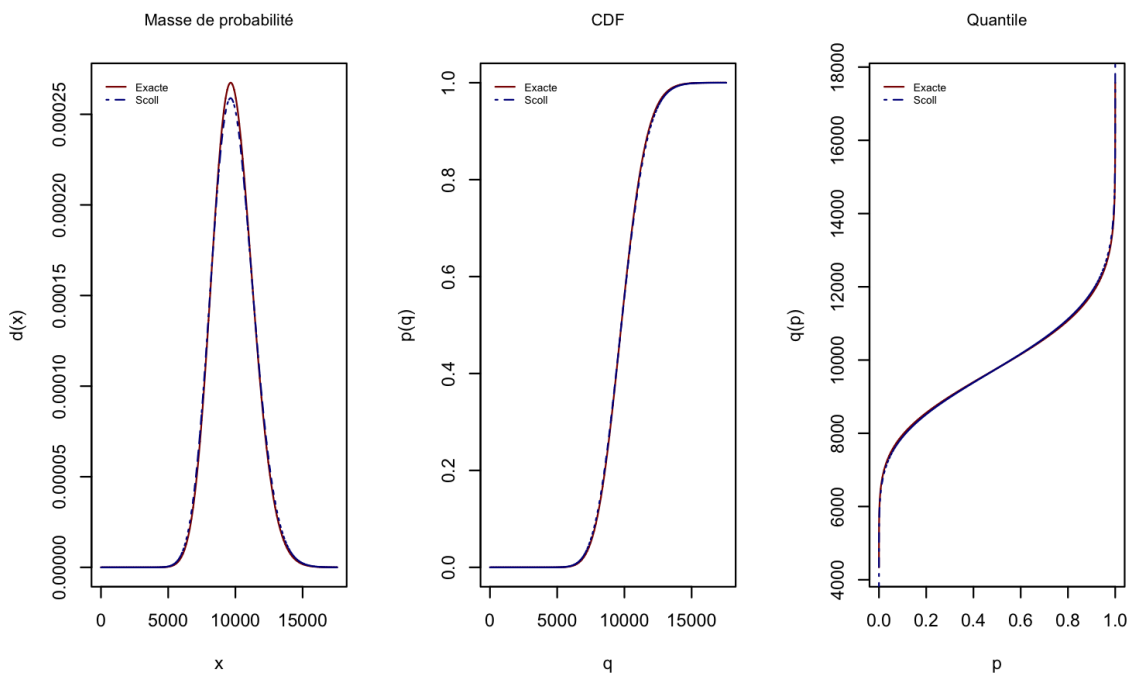


FIGURE 2.5.4 – distribution du coût total  $S_n = \sum_{i=1}^n X_i$  pour  $n = 1000$  participants,  $\delta = 4$ .

La distribution de  $S_n$  est représentée à la figure 2.5.4 pour  $n = 1000$  participants, nous avons pour cette figure considérée un niveau de franchise  $\delta = 4$ , ce qui représente le coût moyen des frais médicaux pour un homme de 30 ans, et une proportion Homme/Femme égale à 0.6. La figure présente également les résultats obtenus par le calcul exact de la distribution de  $S_n$  ainsi que celui obtenu par l'approximation du modèle collectif. On peut voir que le modèle collectif fournit une approximation convenable de la loi du coût total  $S_n$ , en effet, les courbes obtenues par ces deux approches se confondent presque parfaitement, ce qui permet de reproduire la loi de probabilité de  $S_n$ . Aux valeurs extrêmes, l'approximation par le modèle collectif surestime les probabilités des coûts des sinistres, en revanche il sous-estime ces probabilités lorsqu'on est autour de la moyenne, ce qui en fait une approximation prudente dans la mesure où il rend plus probable les événements extrêmes. Pour la suite de ce travail, nous continuerons nos calculs en considérant l'approximation du modèle collectif pour le coût total  $S_n$ .

### 2.5.1.3 Contribution respective des Hommes et des Femmes

L'application de la conditional mean risk sharing (2.3.3) au groupe est représentée à la figure 2.5.5, on y voit l'évolution des cotisations des membres du groupe en fonction de la réalisation  $s$  du coût total  $S_n$  observé en fin de période. Les cotisations  $h_{i,n}^*(s)$  sont strictement croissante ce qui implique une augmentation des cotisations avec le coût total. On peut donc en déduire que la règle est Pareto optimale, dans la mesure où une augmentation du coût total implique nécessairement une augmentation de la valeur des cotisations pour tous les membres, qui auraient plutôt intérêt à adopter un comportement visant à réduire la valeur du coût total. Du fait d'une fréquence moyenne des hospitalisations  $\lambda_i$  plus élevée chez les femmes, on constate comme on pouvait s'y attendre que la valeur des cotisations est plus élevée pour les femmes comparées à celles des hommes du même âge.

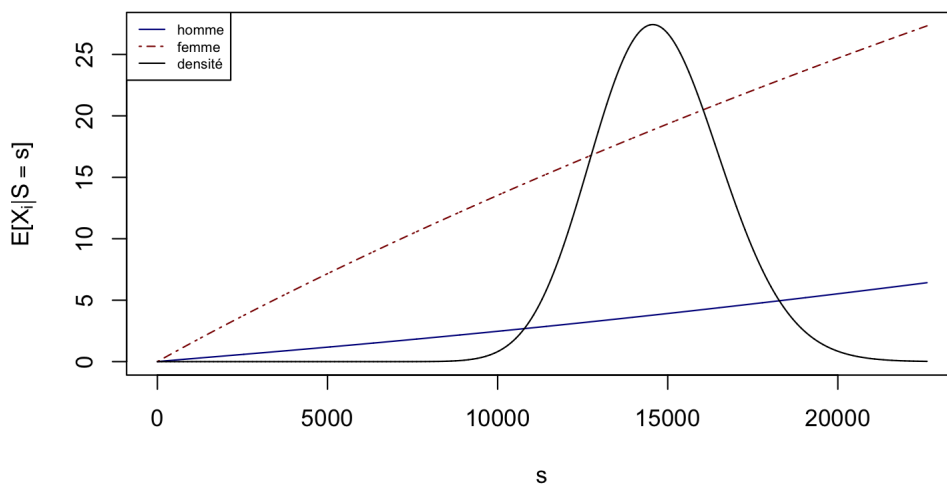


FIGURE 2.5.5 – Contribution  $h_{i,n}^*(s) = E[X_i | S = s]$  pour un homme et une femme et probabilités mis à l'échelle de  $S_n$ .  $n = 1500$ ,  $\delta = 4$  et 60% d'hommes.

Sur cette figure, on peut également voir les valeurs de  $S_n$  les plus probables pour le groupe de  $n = 1500$  participants, et les valeurs les plus probables des cotisations dans ces cas. Par exemple, quand  $s$  varie de  $s = 10000$  à  $s = 20000$ , on s'attend à des cotisations des hommes qui varient entre 2.470 et 5.532, et celles des femmes entre 13.532 et 24.695.

## 2.5.2 Approximation linéaire de la règle de partage

Pour appliquer le calcul des cotisations par l'approximation linéaires (2.4.1), il est tout d'abord nécessaire de calculer l'espérance et la variance de  $X_i$  qui sont nécessaire à l'application de cette formule. Le package `distr` de CAMPHAUSEN et al. 2013 nous fournis des fonctions assez intéressantes nous permettant d'effectuer aisément ces calculs.

Nous avons représenté à la figure 2.5.6, les valeurs des cotisations calculées avec la CMRS ainsi que celles obtenues grâce à la règle de partage linéaire. Le code R permettant d'aboutir à ce résultat est donnée en annexe à la section A.11.1. L'approximation linéaire permet plus ou moins de retrouver la tendance globale de l'évolution avec la réalisation  $s$  du coût totale. Dans la zones où les valeurs de  $S_n$  sont les plus probables, l'approximation linéaire est convenable et se confond bien à la conditional mean risk sharing. Ceci n'est plus le cas lorsqu'on s'éloigne vers la queue de la distribution de  $S_n$ , les écarts observés peuvent en effet être dû aux fait que pour ces valeurs, les probabilités sont très faibles et proches de zéro ce qui peut causer des instabilités dans les calculs numériques.

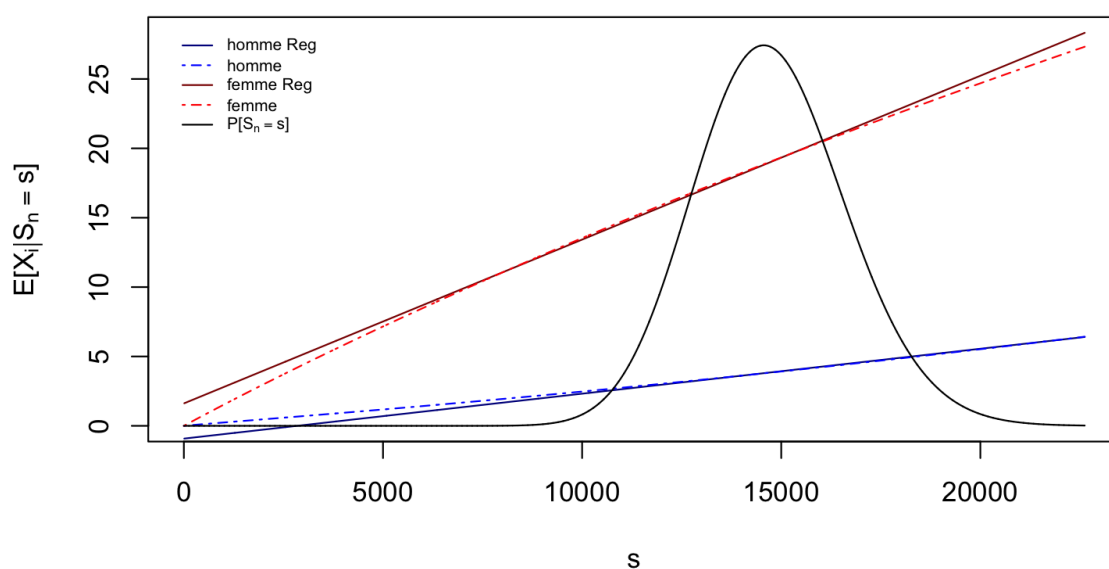


FIGURE 2.5.6 –  $h_{i,n}^*(s)$  et  $h_{i,n}^{reg}(s)$  pour un homme et une femme de 30 ans,  $n = 1500$ .

### 2.5.2.1 Sensibilité de la règle aux variations du nombre de participants

A la figure 2.5.7, nous avons représenté la loi de probabilité des cotisations  $h_{i,n}^*(S_n)$  en fonction du nombre de participant, on retrouve bien le comportement asymptotique 2.4.3. Il y a une réduction des probabilités des valeurs extrêmes avec des cotisations qui se recentrent de plus en plus autour de la prime pure  $\mathbb{E}[X_i]$  lorsque  $n$  augmente.

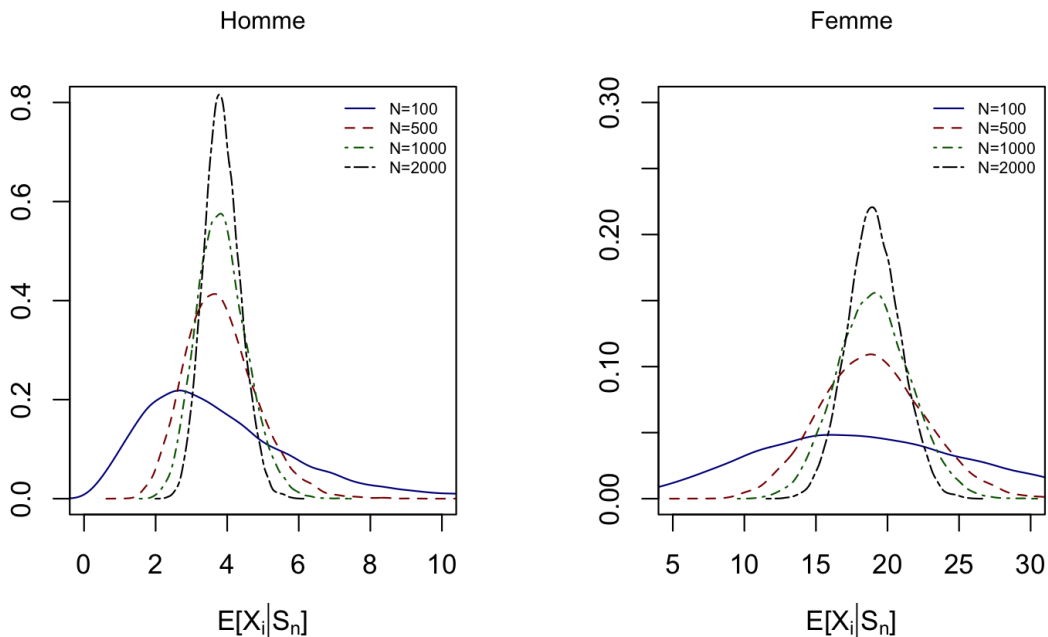


FIGURE 2.5.7 – Loi de probabilité des  $h_{i,n}^*(S_n)$  en fonction de  $n$ .

La figure 2.5.8 permet de mieux se rendre compte de l'impact du nombre de participants sur le niveau général des cotisations pour un homme. La ligne rouge représente la prime pure  $\mathbb{E}[X_i]$ . On peut constater comme c'est le cas, que la variance des cotisations décroît progressivement quand  $n$  augmente et tend à stabiliser les cotisations autour de la prime pure. Le même constat est fait lorsqu'on regarde l'évolution des cotisations pour les femmes dans le groupe à la figure 2.5.8. En général, la contribution des individus converge vers la prime pure individuelle lorsque le nombre de participants devient suffisamment grand comme le suggère (2.4.2). Lorsque le nombre de participant est faible, le système pourrait être peu attractif, dans la mesure où d'une année à une autre, les coûts peuvent changer fortement, cependant lorsque le pool est suffisamment grand, on observe comme une certaine stabilité au niveau des cotisations à payer.

En effet, Lorsque le nombre de participants est suffisamment grand, les effets du pooling deviennent plus importants. Cela signifie que les pertes  $X_i$  se compensent mutuellement à mesure que le nombre de participants augmente, ce qui réduit l'incertitude globale sur les cotisations. La conséquence est que, les cotisations nécessaires pour couvrir ces pertes

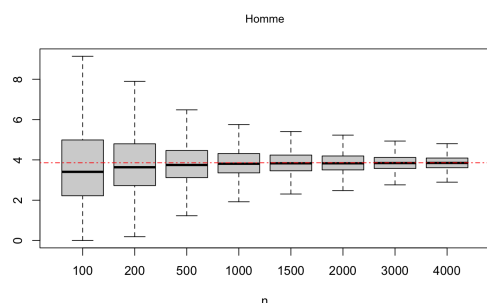


FIGURE 2.5.8 – Convergence des cotisations  $h_{i,n}^*(s)$  vers la primes pure pour un homme de 30 ans.

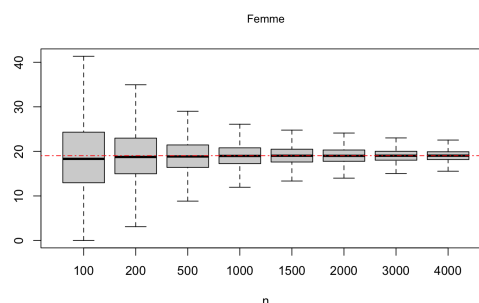


FIGURE 2.5.9 – Convergence des cotisations  $h_{i,n}^*(s)$  vers la primes pure pour une femme de 30 ans.

deviennent plus prévisibles et tendent à se stabiliser autour de la prime pure. Ce résultat suggère donc que dans un système d'assurance collaborative avec un grand nombre de participants, les cotisations peuvent être ajustées de manière plus précise en fonction des pertes subies par les participants du groupe, ce qui aura pour effet de contribuer à rendre le système plus équitable d'un point de vue actuariel.

### 2.5.3 Influence de la franchise $\delta$

En général, les montants des cotisations sont fortement influencées par la franchise choisie par le participant. Une franchise élevée est le plus souvent associée à des cotisations moins élevées, car dans ce cas c'est le participant qui prend en charge une part plus importante de ses frais médicaux. En revanche, pour une franchise plus basse, les cotisations associées sont plus élevées, car c'est le groupe qui dans ce cas prend en charge une plus grande partie des frais médicaux.

Cependant le constat est tout autre ici, on constate en effet à la figure 2.5.10 qu'un changement du niveau de franchise a très peu d'impact sur le niveau de cotisation de cet individu. Cela pourrait s'expliquer par le fait que dans le groupe, les pertes individuelles sont réparties entre tous les membres et par conséquent, même si la franchise d'un participant est modifiée, l'effet sur sa cotisation pourrait être atténué par le nombre de participants dans le groupe. Ce constat est aussi visible à la figure 2.5.11, où en faisant varier la franchise de  $\delta = 0$  à  $\delta = 20$ , comme c'est le cas ici, on constate que cela n'a qu'un très faible impact sur le niveau des cotisations. Sur cette base, on pourrait faire la recommandation de n'autoriser dans le système qu'un seul niveau de franchise identique pour tous les participants.

### 2.5.4 Influence de la proportion Homme/Femme dans le groupe

La solidarité est une caractéristique clé de l'assurance collaborative, car dans ce type de système, les membres se soutiennent mutuellement et partagent leurs coûts. Cela signifie que les membres avec le moins de factures supportent une partie des coûts des membres qui ont plus de factures. De plus, l'augmentation de la proportion des membres les plus coûteux dans le groupe pourrait induire l'augmentation de la volatilité des cotisations au sein du groupe. La figure 2.5.12 montre l'effet d'une augmentation de la proportion d'homme sur la distribution

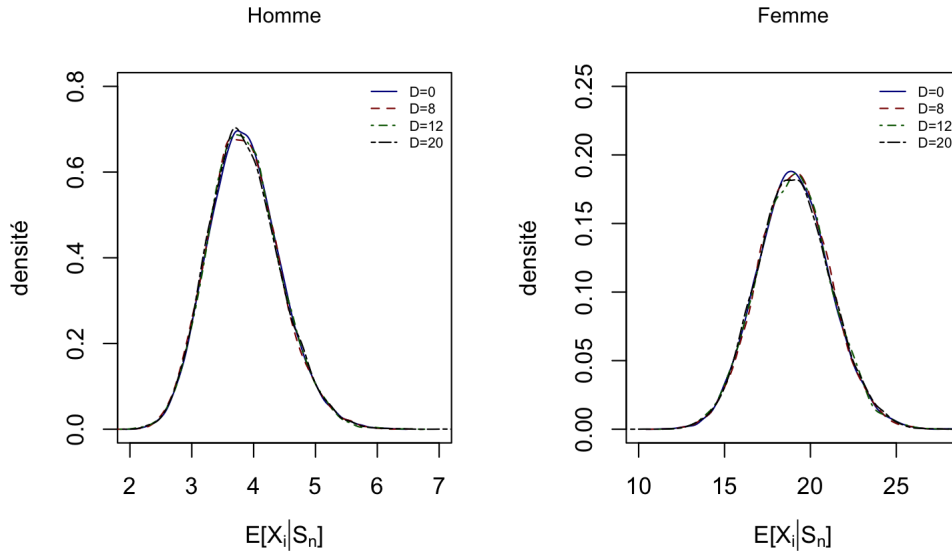


FIGURE 2.5.10 – Loi de probabilité des  $h_{i,n}^*(S_n)$  en fonction de  $\delta$ ,  $n = 1500$  et  $p = 0.6$ .

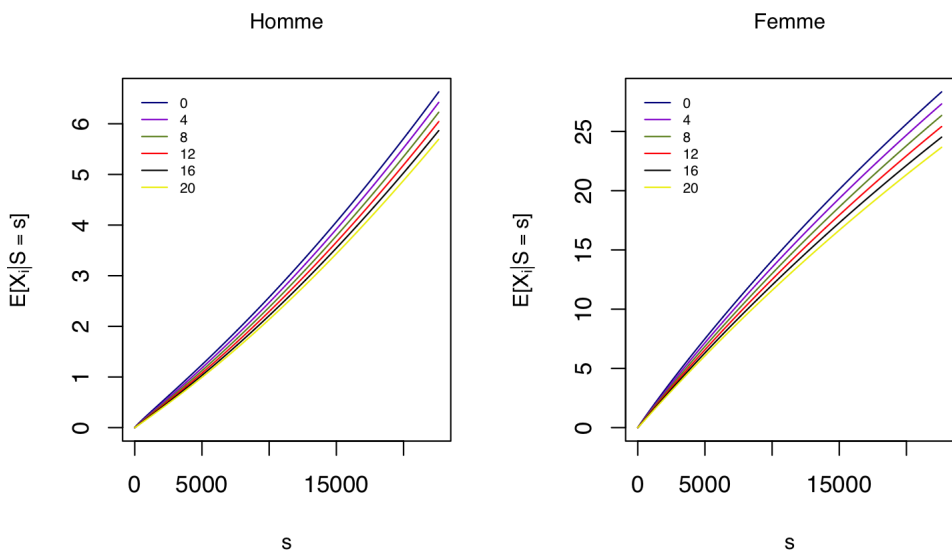


FIGURE 2.5.11 – Influence de la franchise sur les cotisations,  $n = 1500$  et  $p = 0.6$ .

des cotisations. Chacune des courbes présente déjà l'allure d'une normale centrée autour de la prime pure, cependant l'augmentation de la proportion d'homme induit de la volatilité dans les cotisations, rendant ainsi les valeurs extrêmes des cotisations plus probables.

L'effet de l'augmentation de la proportion d'homme sur la variabilité des cotisations est aussi visible sur les figures 2.5.13 et 2.5.14. Le nombre de participant dans ce cas étant fixé à  $n = 1500$ , une augmentation de la proportion d'homme implique une diminution de celle des femmes. Les variations des comportements de santé entre les hommes et les femmes peuvent être à l'origine de cette influence sur la distribution des cotisations. En effet, si les hommes ont plus souvent tendance à présenter des pertes plus volatile ou imprévisibles, cela peut contribuer à rendre plus volatile les cotisations dans un groupe avec une proportion élevée d'hommes. C'est ce qui est le cas ici où on a une fréquence d'hospitalisation pour les hommes qui est beaucoup plus faible par rapport à celle des femmes (voir 2.5.1 et 2.5.1), ce qui rend les hospitalisation rares et plus coûteux quand ils se produisent pour les hommes.

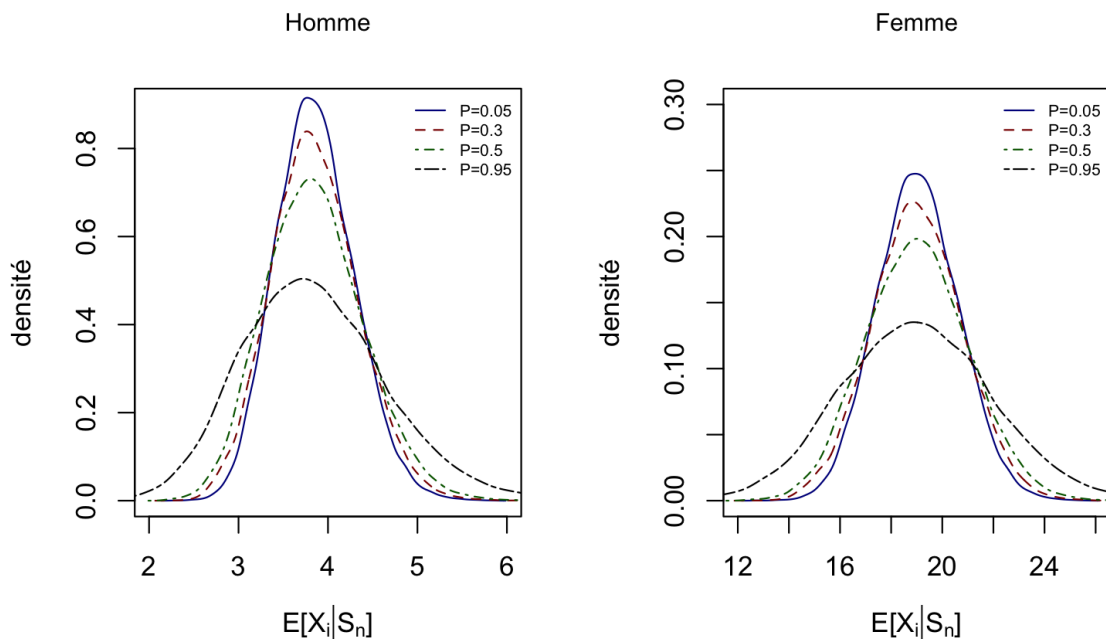


FIGURE 2.5.12 – Influence de la proportion d'homme sur la distribution des cotisations  $\mathbb{E}[X_i|S_n]$ ,  $n = 1500$  et  $\delta = 4$ .

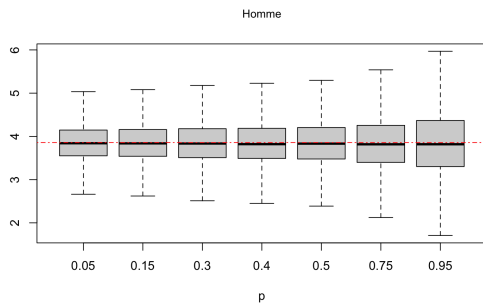


FIGURE 2.5.13 – Variabilité des cotisations pour un homme en fonction de la proportion d’homme dans le groupe,  $n = 1500$  et  $\delta = 4$ .

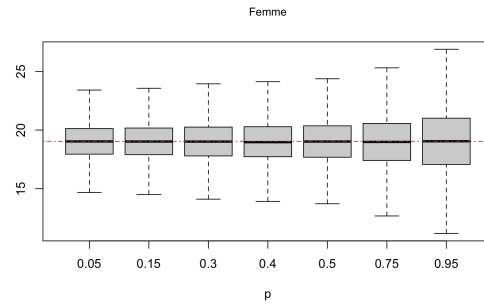


FIGURE 2.5.14 – Variabilité des cotisations pour une femme en fonction de la proportion d’homme dans le groupe,  $n = 1500$  et  $\delta = 4$ .

## 2.6 Conclusion

Dans cette partie, nous avons présenté la conditional mean risk sharing ainsi que quelques-unes des propriétés importantes pouvant rendre cette règle appropriée pour le calcul des cotisations dans le cadre des assurances collaboratives. (i) c’est une règle qui est équitable en ce sens qu’aucun participant ne peut en moyenne faire gain ou perte du partage. (ii) C’est une règle attractive qui pourrait inciter des participants risquophobes à adhérer au système afin de remplacer le coût de leur facture individuelle par le montant des cotisations calculés sous cette règle. (iii) Elle est Pareto-optimale. L’application de la conditional mean risk sharing sur un exemple réduit nous a permis de mettre en évidence le caractère croissant du montant des cotisations avec le coût total des soins de l’ensemble du groupe. En effet ce résultat est important dans la mesure où elle incite les participants à adopter des comportements visant à réduire le coût total en fin de période.

Sur notre exemple, nous avons également mis en évidence l’influence de certains paramètres importants du système sur le niveau et la stabilité des cotisations dans le groupe. Notamment le nombre  $n$  de participants qui s’est révélé exercer une influence significative sur la stabilité des cotisations dans le groupe. Il en va de même pour la composition du groupe qui pouvait rendre plus volatile les cotisations selon qu’il était composé d’une plus grande proportion d’hommes. La franchise n’a cependant pas eu autant d’influence sur les cotisations contrairement à ce qu’on aurait pu penser, nous avons de ce fait décidé d’opter pour la suite de ce travail pour un système avec une franchise unique pour tous les participants. Nous appliquerons la conditional mean risk sharing dans la suite, sur un modèle plus réaliste où on a une population plus complète. Nous étudierons également la possibilité de recourir à un assureur afin de rendre les cotisations plus stables d’une année à une autre quand celles-ci sont très variables.

## Chapitre 3

# Étude d'un système avec (re-)assureur

### 3.1 Introduction

Dans le chapitre précédent, nous avons décrit une approche permettant une répartition équitable du coût total des soins entre tous les participants. Cependant le fait que le montant des cotisations  $h_{i,n}^*(s)$  n'est connu qu'en fin de période fait en sorte que cette approche exige un grand niveau de confiance entre les participants. En effet, il pourrait arriver qu'en fin de période, certains participants soient dans l'incapacité de verser leur contribution ou alors qu'ils ne souhaitent tout simplement pas le faire. De plus, d'une année à une autre, les cotisations peuvent fortement varier exposant ainsi les participants à verser des cotisations fortement élevées. Théoriquement, la valeur des cotisations versées en fin de période sont illimitées, cependant la capacité du groupe à pouvoir supporter de tels coûts est quant à elle restreinte. Il n'est donc pas anodin de penser à la possibilité de recourir à un (re-)assureur pour supporter une partie des factures de l'ensemble. C'est généralement ce qui est fait en pratique dans les assurances collaboratives.

Dans ce chapitre, nous étudierons la possibilité pour le groupe de transférer une partie du risque à un (re-)assureur afin de construire un système suffisamment flexible qui va transformer les contributions aléatoires versées en fin de période en des paiements fixes en début de période avec à l'appui un mécanisme de cash-back et une couverture d'assurance.

### 3.2 Décomposition des contributions

La contribution  $h_{i,n}^*(s)$  que chaque participant doit verser en fin de période est à présent remplacée par la combinaison :

- d'une provision  $\pi_{i,n}$  versée en début de période et qui dépend à la fois de la distribution de  $X_i$  et de celle de  $S_n$ .

- d'un cash-back  $B_{i,n}$  distribué en fin de période en cas de situation favorable. Par situation favorable, on entend des scénarios pour lesquelles la valeur des contributions effectivement collectée est supérieure à la valeur du coût total des factures déclarées.

Si le pool opte pour une couverture par un contrat stop-loss avec une rétention  $\omega_n$ , alors le pool aura à sa charge

$$S_n^{P2P} = \min\{S_n, \omega_n\} \quad \text{où} \quad S_n = \sum_{i=1}^n X_i. \quad (3.2.1)$$

Le pool prend à sa charge la première tranche  $[0, \omega_n]$  du coût total  $S_n$  des frais médicaux de l'ensemble, tandis que la tranche supérieure  $[\omega_n, +\infty[$  est transférée vers un (re-)assureur. Ainsi le montant maximum payé par le participant  $i$  sera dans tout les cas limitée à  $\pi_{i,n}$  qui se décompose comme suit :

$$\pi_{i,n} = \pi_{i,n}^{P2P} + \pi_{i,n}^{SL} \quad (3.2.2)$$

Où  $\pi_{i,n}^{P2P}$  est la part du coût total payé par le participant  $i$  et servant à la couverture de la première tranche  $[0, \omega_n]$ .  $\pi_{i,n}^{SL}$  est la contribution au prix de la couverture par la stop-loss(tranche  $[\omega_n, +\infty[$ ).

En faisant intervenir un (re-)assureur ici, on fait une migration du modèle auto-gouverné où le risque est entièrement supporté par le pool vers le modèle de courtage où il y a un transfert de risque vers un (re-)assureur. Le modèle de courtage engendre néanmoins des coûts supplémentaires du fait que l'une des parties des primes d'assurances perçus est utilisée pour la rémunération des actionnaires qui détiennent la société d'assurance. Ces frais rendent le modèle de courtage beaucoup moins attrayant par rapport au modèle auto-gouverné. Cependant, ils garantissent une plus grande probabilité que les membres auront leurs factures remboursées. De plus, en ayant recours à une couverture d'assurance, le pool bénéficie de l'expertise de l'assureur en matière de traitement et de règlement de sinistre.

### 3.3 Définition de la rétention

Un élément essentiel lors de la mise en œuvre d'un tel système est le choix de la rétention  $\omega_n$  du groupe, c'est en effet important pour trouver un équilibre entre le transfert de risque à l'assureur et la responsabilité assumée en interne par le groupe. Plusieurs considérations stratégiques et financières doivent être prises en compte pour déterminer la capacité  $\omega_n$  du groupe.

La valeur de la rétention  $\omega_n$  peut par exemple être choisie de façon à ce que la probabilité  $P[S_n < \omega_n]$  soit la plus grande possible. Il s'agit de la probabilité pour qu'un surplus  $S_n - \omega_n$  soit accordé en fin de période aux participants. La valeur de  $\omega_n$  a un impact autant sur la valeur de  $\pi_{i,n}^{P2P}$  que sur celle de  $\pi_{i,n}^{SL}$ . Elle doit donc être choisie de façon à ce que la participation au système soit suffisamment attrayante pour inciter les nouveaux membres à payer  $\pi_{i,n}$  plutôt que la prime réclamée par un assureur pour une couverture d'assurance classique.

Soit  $\beta$  la probabilité cible pour laquelle un surplus est obtenue en fin de période. La rétention  $\omega_n$  correspondante à cette probabilité  $\beta$  est donnée par

$$\beta = F_{S_n}(\omega_n) \Leftrightarrow \omega_n = F_{S_n}^{-1}(\beta) \quad (3.3.1)$$

Le niveau de rétention choisie est défini comme le quantile de la loi de  $S_n$  correspondant à la probabilité  $\beta$ .

## 3.4 Contribution à la tranche inférieure et Bonus

### 3.4.1 Définition de la tranche inférieure

Le niveau de rétention  $\omega_n$  du pool pouvant être déterminé selon les objectifs du groupe, reste à présent à définir comment seront calculées les contributions  $\pi_{i,n}^{P2P}$  des participants pour que celui-ci puisse être en mesure d'amortir cette partie du coût total constituant la tranche inférieure.

On suppose que le groupe est suffisamment large et que les fonctions  $s \mapsto \mathbb{E}[X_i | S_n = s]$  sont non-décroissantes pour tout  $i$ , ce qui est bien le cas ici, puisque les résultats que nous avons obtenus au chapitre précédent montraient bien que les cotisations étaient des fonctions croissantes de la réalisation  $s$  du coût total. Cette exigence nous garantit que le partage soit comonotonique et de ce fait Pareto-optimal.

Définissons les quantités  $\omega_{1,n}, \omega_{2,n}, \dots, \omega_{n,n}$  tels que

$$\omega_{i,n} = F_{h_{i,n}^*(S_n)}^{-1}(F_{S_n}(\omega_n)) = \mathbb{E}[X_i | S_n = \omega_n] \quad (3.4.1)$$

$\omega_{i,n}$  est le niveau de rétention individuelle qu'il faut définir pour être en mesure d'accorder un bonus aux participants en fin de période avec une probabilité  $\beta$ . C'est la contribution maximale que peut verser un participant pour supporter la couverture de la tranche inférieure. Par définition, on a

$$\pi_{i,n}^{P2P} = \omega_{i,n} = h_{i,n}^*(\omega_n) = \mathbb{E}[X_i | S_n = \omega_n] \quad (3.4.2)$$

Par la propriété d'allocation totale de la conditional mean risk sharing (2.2.4), on vérifie bien que l'ensemble des rétentions individuelles  $\omega_{i,n}$  est suffisant pour couvrir la rétention  $\omega_n$  au niveau du groupe. En effet,

$$\omega_{i,n} = h_{i,n}^*(\omega_n) \Rightarrow \sum_{i=1}^n \omega_{i,n} = \omega_n$$

Puisque les fonctions  $h_{i,n}^*$  sont des fonctions continues et strictement croissantes, alors on a

$$\begin{aligned} S_n \geq \omega_n &\Leftrightarrow h_{i,n}^*(S_n) \geq h_{i,n}^*(\omega_n) \\ &\Leftrightarrow h_{i,n}^*(S_n) \geq \omega_{i,n} = F_{h_{i,n}^*(S_n)}^{-1}(F_{S_n}(\omega_n)) \quad i = 1, 2, \dots, n, \end{aligned}$$

Et donc on peut écrire

$$\begin{aligned} (S_n - \omega_n)_+ &= \left( \sum_{i=1}^n \left( h_{i,n}^*(S_n) - F_{h_{i,n}^*(S_n)}^{-1}(F_{S_n}(\omega_n)) \right) \right)_+ \\ &= \sum_{i=1}^n \left( h_{i,n}^*(S_n) - F_{h_{i,n}^*(S_n)}^{-1}(F_{S_n}(\omega_n)) \right)_+ \\ &= \sum_{i=1}^n \left( h_{i,n}^*(S_n) - \omega_{i,n} \right)_+ \end{aligned} \quad (3.4.3)$$

Un raisonnement analogue peut être adoptée pour décomposer le surplus  $(\omega_i - S_n)_+$  en des variables aléatoires  $(\omega_{i,n} - h_{i,n}^*(S_n))_+$  et de cette façon, on peut déterminer comment répartir le surplus lorsqu'il y en a un, ainsi que la contribution de chaque participant pour couvrir  $S_n^{P2P}$  telle que définie à l'équation (3.2.1) car en effet,

$$\begin{aligned} S_n^{P2P} &= \min\{S_n, \omega_n\} = \omega_n - (\omega - S_n)_+ \\ &= \sum_{i=1}^n \left( \omega_{i,n} - (\omega_{i,n} - h_{i,n}^*(S_n))_+ \right) \\ &= \sum_{i=1}^n \min\{h_{i,n}^*(S_n), \omega_{i,n}\} \end{aligned} \quad (3.4.4)$$

Ainsi chaque participant devra payer une contribution égale à  $\min\{h_{i,n}^*(S_n), \omega_{i,n}\}$  pour couvrir  $S_n^{P2P}$ .

### 3.4.2 Calcul du cash-back

Si un surplus  $B_n = (\omega_n - S_n)_+$  est disponible en fin de période, alors le cash-back sera redistribué aux participants ou encore au profit d'une cause à laquelle ils adhèrent. Si on note  $B_{i,n}$  le cash-back redistribué au participant  $i$ , alors on aura

$$B_{i,n} = \begin{cases} 0 & \text{si } S_n \geq \omega_n = \sum_{i=1}^n \pi_{i,n}^{P2P} \\ \pi_{i,n}^{P2P} - h_{i,n}^*(S_n) & \text{si } S_n < \omega_n \end{cases} \quad (3.4.5)$$

Le cash-back suis une distribution zéro-augmenté avec une probabilité  $P[B_{i,n} = 0] = P[S_n \geq \omega_n]$ . Étant donné la relation (3.4.2), et d'après les propriétés des fonctions  $h_{i,n}^*$ , on sait que pour tout  $s$  tel que  $s \leq \omega_n$ , la contribution finale payé par le participant  $i$  sera toujours plus

petite ou égale à  $\pi_{i,n}^{P2P}$ , c'est à dire

$$h_{i,n}^*(s) \leq \pi_{i,n}^{P2P} = h_{i,n}^*(\omega_n) \quad \text{pour tout } s \leq \omega_n \Rightarrow B_{i,n} \geq 0 \quad \forall i$$

De sorte qu'aucune contribution supplémentaire n'est réclamée en fin de période au participant  $i$ . Si  $S_n < \omega_n$ , alors le participant  $i$  reçoit un cash-back égal à  $B_{i,n} = \omega_{i,n} - h_{i,n}^*(S_n)$ . Chaque participant doit alors payer en début de période une contribution maximale  $\omega_{i,n}$  à laquelle il est exposé et le mécanisme de cash-back opère en fin de période à fin de redistribuer le surplus entre tous les membres du groupe.

### 3.5 Définition de la tranche supérieure

En passant par un (re-)assureur partenaire pour la couverture de la tranche supérieur du coût total  $S_n$ , le groupe transfère la partie  $[\omega_n, +\infty[$  de  $S_n$  à l'assureur. Le coût de cette couverture doit ainsi être répartie entre tous les participants. La répartition de cette prime au sein du pool peut être réalisée de différentes manières en fonction des règles et méthodes spécifiques adoptées dans le groupe.

Plusieurs approches peuvent être envisagées dans cet optique. En effet, la prime peut être répartie soit de façon égale entre tous les participant, de sorte qu'indépendamment de la distribution des  $X_i$ , tous les participants paient la même quantité. Soit on pourrait également envisager une répartition proportionnelle où chaque participant contribue une proportion fixe de la prime tout comme c'est par exemple le cas de la règle de partage proportionnelle. Ou encore, la prime payée par chaque participant pour la couverture de la tranche supérieure pourrait être déterminée en fonction de ce que chaque participant aura apporté comme réclamation à la tranche supérieure du coût total  $S_n$  en termes de factures déclarées. Puisque nous nous proposons de construire un système qui reste le plus équitable possible d'un point de vue actuarielle, c'est cette dernière approche qui sera adoptée ici pour répartir le coût de la couverture stop-loss au sein du groupe et entre tous les participants.

Supposons la forme suivante pour la prime stop-loss

$$\pi_n^{SL} = (1 + \theta_{SL}) \mathbb{E}[(S_n - \omega_n)_+] \quad (3.5.1)$$

où  $\theta_{SL}$  correspond aux différent frais à appliquer pour obtenir la prime réclamée par l'assureur pour la couverture de la tranche supérieure  $[\omega_n, +\infty[$ .

À l'aide de la relation (3.4.3), nous sommes en mesure de définir de quelle façon la prime d'assurance vas être redistribuée entre tous les participants. La prime totale est donnée par

$$\pi_n^{SL} = (1 + \theta_{SL}) \sum_{i=1}^n \mathbb{E}[(h_{i,n}^*(S_n) - \omega_{i,n})_+] \quad (3.5.2)$$

Chaque participant devra de ce fait payer

$$\pi_{i,n}^{SL} = (1 + \theta_{SL}) \mathbb{E}[(h_{i,n}^*(S_n) - \omega_{i,n})_+] \quad (3.5.3)$$

pour couvrir la tranche au-dessus de la capacité du groupe.

En récapitulatif, l'opération d'assurance sous ce modèle peut être décomposée de la manière suivante : chaque participant

- effectue un paiement initial déterministe  $\pi_{i,n}^{P2P} = \omega_{i,n}$ , destinée à la couverture de la tranche inférieure
- paie une contribution déterministe  $\pi_{i,n}^{SL}$  pour la couverture de la tranche supérieure incluant des frais  $\theta_{SL}$

$$\pi_{i,n}^{SL} = \mathbb{E}[(h_{i,n}^*(S_n) - \omega_{i,n})_+],$$

- reçoit en fin de période un cash-back stochastique  $B_{i,n}$

$$B_{i,n} = \left( \pi_{i,n}^{P2P} - h_{i,n}^*(S_n) \right) \mathbb{I}[S_n \leq \omega_n] = \left( \pi_{i,n}^{P2P} - h_{i,n}^*(S_n) \right)_+,$$

- et en échange, reçoit le remboursement de ses frais médicaux  $X_i$ .

Le fait que l'assureur charge des frais dans la prime stop-loss fait en sorte que l'on ne peut plus espérer du système qu'il soit équitable dans son ensemble comme c'était le cas dans la première partie. Cependant on voudrait que l'équité actuarielle soit satisfaite si on met de côté ces chargements. Par équité on voudrait que la contribution moyenne payée par le participant  $i$  soit égale au montant des frais médicaux qu'il ramène dans le pool. En effet, nous avons en moyenne

$$\mathbb{E} \left[ X_i - \omega_{i,n} + \left( \pi_{i,n}^{P2P} - h_{i,n}^*(S_n) \right)_+ - \mathbb{E} \left[ (h_{i,n}^*(S_n) - \omega_{i,n})_+ \right] \right] = \mathbb{E} \left[ X_i - h_{i,n}^*(S_n) \right] = 0$$

pour  $i = 1, 2, \dots, n$ . Noter que nous n'avons pas tenu compte des effets de l'actualisation ici car nous travaillons sous taux d'intérêts nuls comme c'est généralement le cas en non-vie avec des opérations à court termes.

### 3.6 Application

Considérons de nouveau un groupe formé de  $n$  individus qui font pot commun afin de partager le coût de leurs soins médicaux. Chaque participant  $i$ ,  $i \in 1, 2, 3, \dots, n$ , réclame au pool le remboursement de ses soins médicaux  $X_i$  après déduction de sa franchise  $\delta_i = \delta$ .

Dans cette partie, nous considérons la franchise comme étant fixe pour tous les participant et est égale au coût moyen des soins médicaux pour un homme de 30 ans. Le groupe est formé de participant de toutes les tranches d'âge telles que repris dans les tableaux 2.5.1 et 2.5.2 avec les paramètres correspondant. En complément à ces paramètres, on suppose la composition tu tableau 3.6.1 pour l'ensemble du groupe.

La composition du pool telle que proposée au tableau 3.6.1 est inspirée de la composition de la population Belge par tranche d'âge pour l'année 2022. Au sein du pool nous repérons les participants par des groupes en fonction de l'âge et du sexe. Les groupes sont numérotés de 1

Âge	20	30	40	50	60	70	80
	Grp 1	Grp 2	Grp 3	Grp 4	Grp 5	Grp 6	Grp 7
Homme	0.079	0.085	0.085	0.090	0.077	0.051	0.023
	Grp 8	Grp 9	Grp 10	Grp 11	Grp 12	Grp 13	Grp 14
Femme	0.078	0.085	0.084	0.088	0.080	0.060	0.035

TABLE 3.6.1 – Composition du groupe par âge en proportion de  $n$ .

à 14. Les groupes 1 à 7 représentent respectivement les hommes âgés de 20 à 80 ans tandis que les groupes 8 à 14 représentent respectivement les femmes âgées de 20 à 80 ans.

Les coûts des soins médicaux augmentent généralement avec l'âge, car les individus plus âgés ont tendance à avoir des besoins médicaux plus fréquents et plus coûteux. En conséquence, pour maintenir un partage équitable, les personnes plus âgées peuvent être amenées à contribuer davantage pour couvrir ces coûts plus élevés. Outre les différences liées à l'âge, les coûts des soins médicaux peuvent différer entre hommes et femmes en raison de facteurs biologiques, de comportements liés à la santé, ou de besoins médicaux spécifiques (comme les soins liés à la maternité). Tout ceci explique les différences que l'on peut observer au niveau des valeurs des cotisations telles que le montre les figures 3.6.1. Dans le groupe ce sont les hommes et femmes de 80 ans (Grp 8 et 14) qui paient les contributions les plus élevées. Ceci s'explique au vu des paramètres que nous avons utilisé dans nos calculs, le coût moyen des hospitalisations étant plus élevées à ces âges. En général les contributions des femmes sont plus grandes que celles des hommes du même âge sauf pour les âges extrêmes où les hommes paient plus.

La conditionnai mean risk sharing est une règle qui en théorie satisfait la condition d'allocation totale 2.2.4, c'est à dire qu'en théorie la somme des cotisations collectées est suffisante pour couvrir le coût total des factures déclarées dans le groupe en fin de période. Cependant lorsqu'on fait les calculs, on constate des écarts. En effet il y a des valeurs de  $S_n$  total pour lesquelles  $\sum h_{i,n}^*(s) > s$  et des valeurs pour lesquelles on observe le contraire comme le montre la figure 3.6.2. Des écarts que l'on peut considérer comme négligeable comparés aux valeurs du coût total mis en jeu. Des écarts plus importants sont observés dans les queues de la distributions de  $S_n$  là où les probabilités sont très faibles ce qui peut être à l'origine des instabilités numériques. La portion du code permettant d'obtenir ce résultat est reportée en annexe à la section B.6.

### 3.6.1 Illustration de la décomposition en couche

Dans un premier exemple pour illustrer la décomposition en couche, et pour se faire une idée des montants mis en jeu, nous avons effectué les calculs pour un groupe composé de  $n = 1500$  participants avec les proportions du tableau 3.6.1. Nous avons supposé une valeur des frais  $\theta_{SL}$  à appliquer à la prime stop-loss égale à  $\theta_{SL} = 0.5$ . Nous avons également considéré comme valeur  $\beta = 0.85$ .

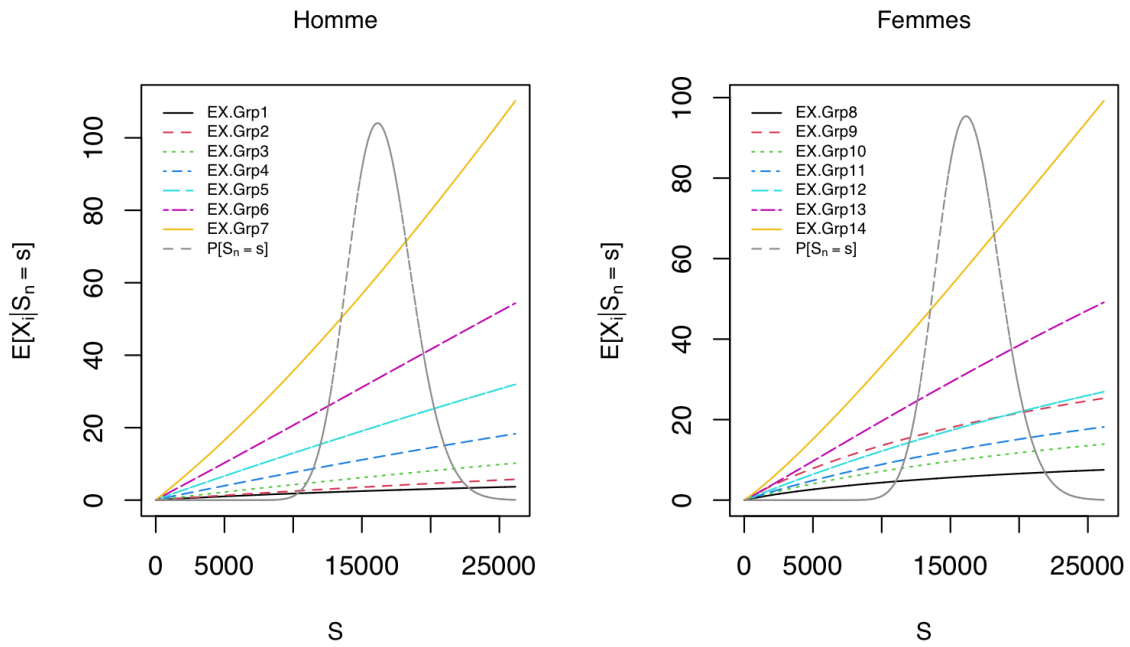


FIGURE 3.6.1 – Niveau des cotisation en fonction de l'âge et du sexe dans le groupe.

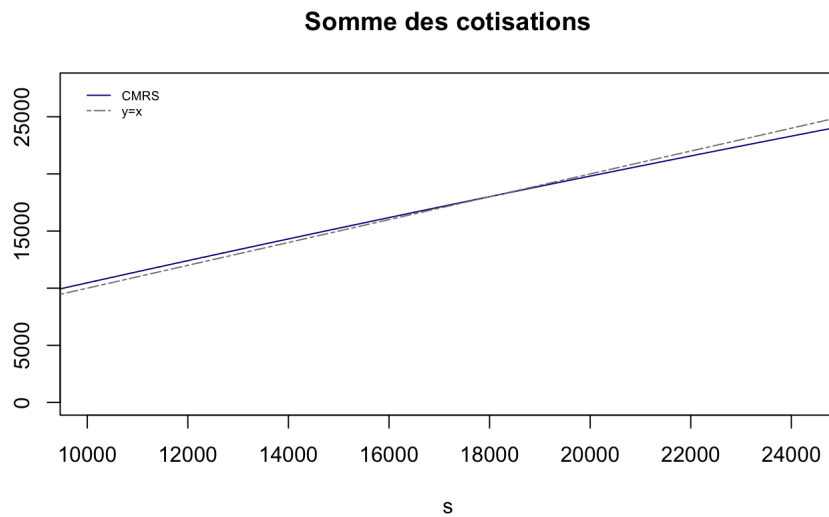


FIGURE 3.6.2 – Somme des contributions dans le pool.

Tout calcul fait, l'ensemble des résultats de nos calculs est consigné dans le tableau 3.6.2. On constate que les coûts sont plutôt représentatifs du profil de chaque participant. On retrouve là une des caractéristiques de la conditional mean risk sharing, qui reste d'application tout au long du processus de calcul des primes des différents participants dans le pool, chaque participant contribue en fonction de ce qu'il réclame en moyenne au groupe. En effet, on voit bien que les primes  $\pi_{i,n}$  payées ne sont pas démesurées par rapport au coût moyen des frais médicaux  $\mathbb{E}[X_i]$  que chaque participant aurait payé s'il était resté seul.

	Âge	Sexe	$n_i$	$\omega_i$	$\pi_{i,n}^{SL}$	$\pi_n$	$\mathbb{E}[B_{i,n}]$	$\mathbb{E}[X_i]$
Grp1	20	M	119	2.939	0.032	2.972	0.306	2.653
Grp2	30	M	128	4.349	0.058	4.407	0.530	3.855
Grp3	40	M	128	7.626	0.106	7.733	0.964	6.729
Grp4	50	M	135	13.674	0.190	13.864	1.732	12.063
Grp5	60	M	116	23.594	0.339	23.933	3.069	20.741
Grp6	70	M	77	39.111	0.610	39.721	5.423	34.081
Grp7	80	M	32	74.235	1.388	75.623	11.782	63.353
Grp8	20	F	117	6.367	0.051	6.419	0.521	5.879
Grp9	30	F	128	20.844	0.193	21.037	1.930	19.037
Grp10	40	F	127	11.295	0.113	11.408	1.108	10.258
Grp11	50	F	132	14.479	0.158	14.637	1.529	13.050
Grp12	60	F	120	20.851	0.254	21.105	2.413	18.600
Grp13	70	F	90	36.325	0.523	36.849	4.795	31.866
Grp14	80	F	51	68.604	1.208	69.812	10.553	58.832

TABLE 3.6.2 – Récapitulatif des résultats de la décomposition en couche :  $n = 1500$ ,  $\theta_{SL} = 0.5$ ,  $\beta = 0.85$ .

À la figure 3.6.3, nous avons représenté la loi de probabilité du cash-back pour les différents groupes du pool. Cette loi est typique de la distribution zéro-augmentée, avec une forte masse de probabilité en zéro. La masse de probabilité en zéro est d'autant plus grande pour les profils à faible risque, c'est à dire les profils pour lesquels les coûts moyens des frais médicaux sont peu élevés. La probabilité en zéro se réduit lorsqu'on se rapproche des profils pour lesquels  $\mathbb{E}[X_i]$  est grand, il devient donc plus probable pour ces profils de recevoir un cash-back en fin de période. Ceci pourra s'expliquer par le fait que ces profils contribuent les plus au coût total, car ayant une rétention plus élevée et donc lorsqu'un surplus est disponible, ils sont ceux qui sont susceptibles d'en bénéficier les plus. L'application de la décomposition en couche à été faite grâce au code de l'annexe B.7

### 3.6.2 Effet de la rétention sur les cotisations dans la tranche inférieure

Lorsque dans le système on transfère une partie du risque vers un (re-)assureur, on peut s'attendre à une réduction de la volatilité des cotisations. Le transfert de risque vers l'assureur permet de réduire l'exposition directe du groupe aux fluctuations des valeurs du coût total observée en fin de période.

Les figures 3.6.4 à 3.6.8 présentent l'effet de la rétention ainsi que celle du nombre de participant sur la variabilité des cotisations dans le groupe sachant que le coût total est

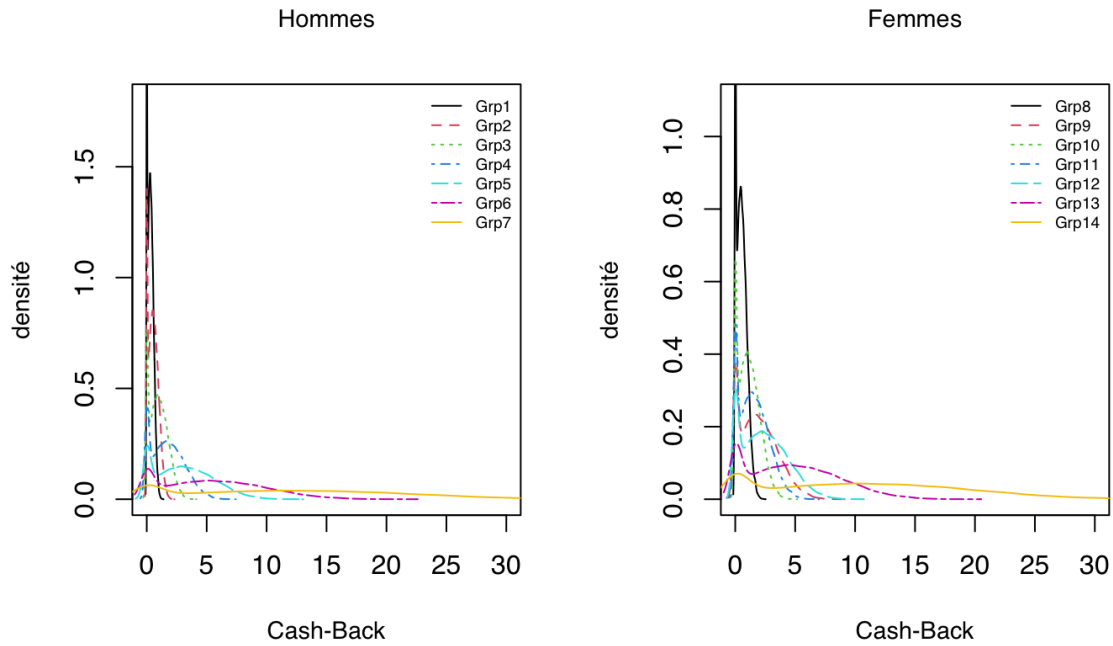


FIGURE 3.6.3 – Loi de probabilité du cash-back  $B_{i,n}$  :  $n = 1500$ ,  $\theta_{SL} = 0.5$ ,  $\beta = 0.85$ .

inférieur à la rétention  $\omega_n$  du groupe ( $S_n < \omega_n$ ). Nous avons représenté sur chacune de ces figures la variabilité des cotisations servant à la couverture de la tranche inférieure  $S_n^{P2P}$  et à l'entièreté du coût total  $S_n$ . Pour un nombre de participant fixe, lorsque le coût total est inférieur à la rétention du groupe, on obtient des cotisations beaucoup moins variables comparé à une situation où le groupe ne ferait pas recours à un (re-)assureur.

En général, l'assureur prend en charge une partie des pertes potentielles du groupe, ce qui diminue le risque financier supporté par les membres du groupe lui-même. L'assureur assume une part des frais excédant la rétention du groupe, ce qui réduit l'impact financier des valeurs exceptionnelles du coût total pour le groupe. En conséquence, la rétention conservée par le groupe après le transfert de risque devient plus prévisible et stable. Les variations potentielles des contributions dues aux fluctuations aléatoires du coût total sont de ce fait atténuées.

### 3.6.3 Variation de la valeur de la rétention

Lorsqu'on fait varier le niveau de rétention  $\omega_n$  du groupe à travers la probabilité  $\beta$  avec laquelle un bonus est accordé aux participants du groupe, on constate que la volatilité des cotisations augmente progressivement et tend à se stabiliser à un certain niveau comme on peut le voir à la figure 3.6.9 où nous avons représenté la variabilité des cotisations dans la tranche inférieure pour un homme et une femme de 30 ans dans le pool. Nous avons choisi de faire une représentation juste pour ces deux tranches d'âges car l'effet décrit ici est le même pour tous les autres groupes cités précédemment.

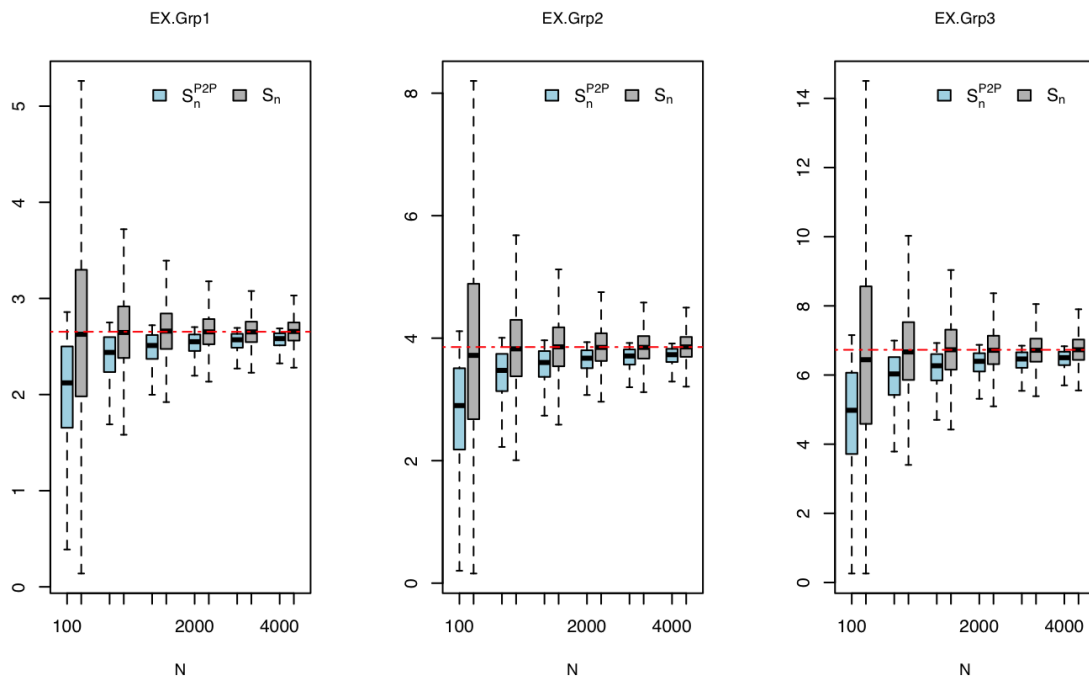


FIGURE 3.6.4 – Effet de la rétention sur les cotisations des groupes 1,2 et 3.

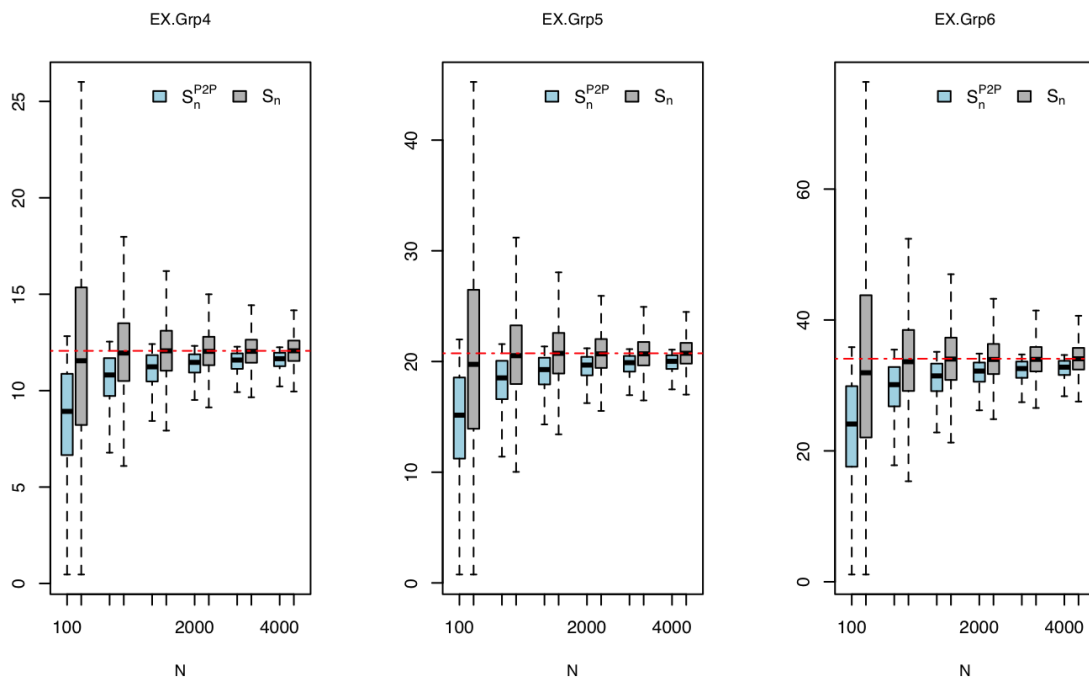


FIGURE 3.6.5 – Effet de la rétention sur les cotisations des groupes 4,5 et 6.

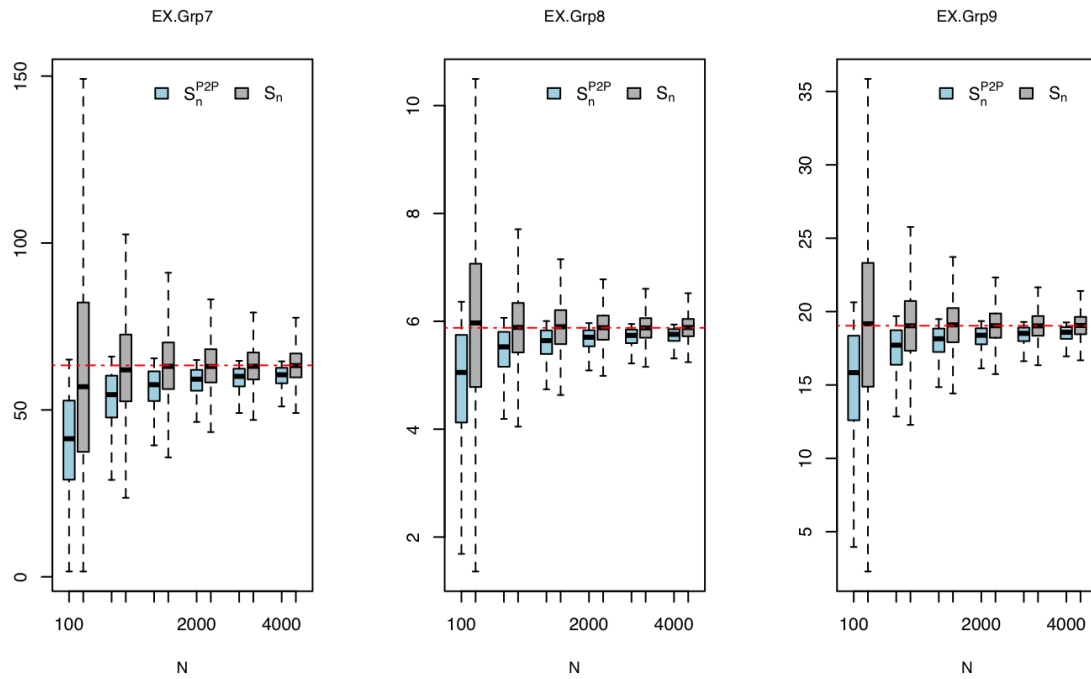


FIGURE 3.6.6 – Effet de la rétention sur les cotisations des groupes 7,8 et 9.

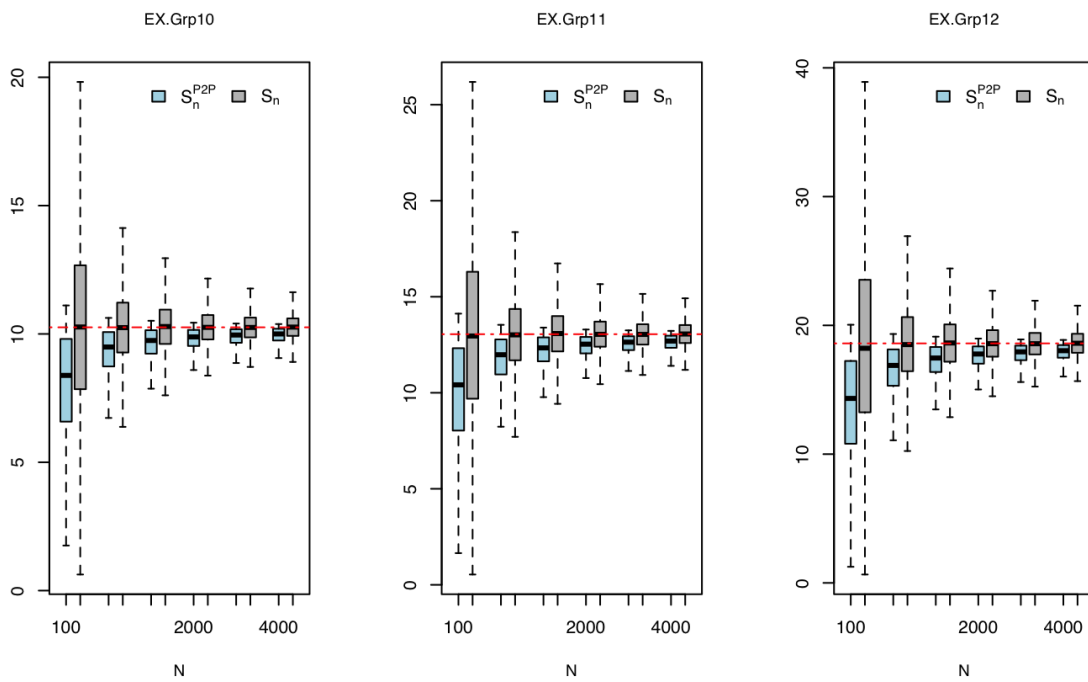


FIGURE 3.6.7 – Effet de la rétention sur les cotisations des groupes 10,11 et 12 sachant.

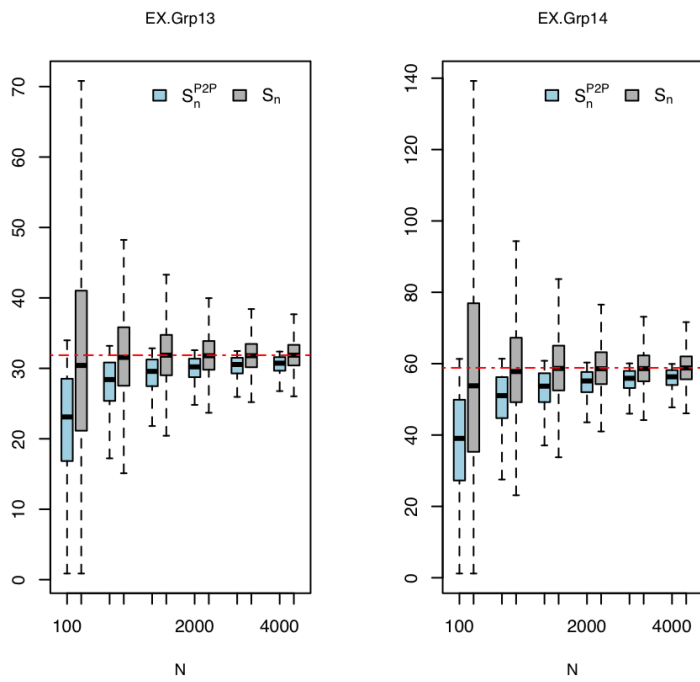


FIGURE 3.6.8 – Effet de la rétention sur les cotisations des groupes 13 et 14.

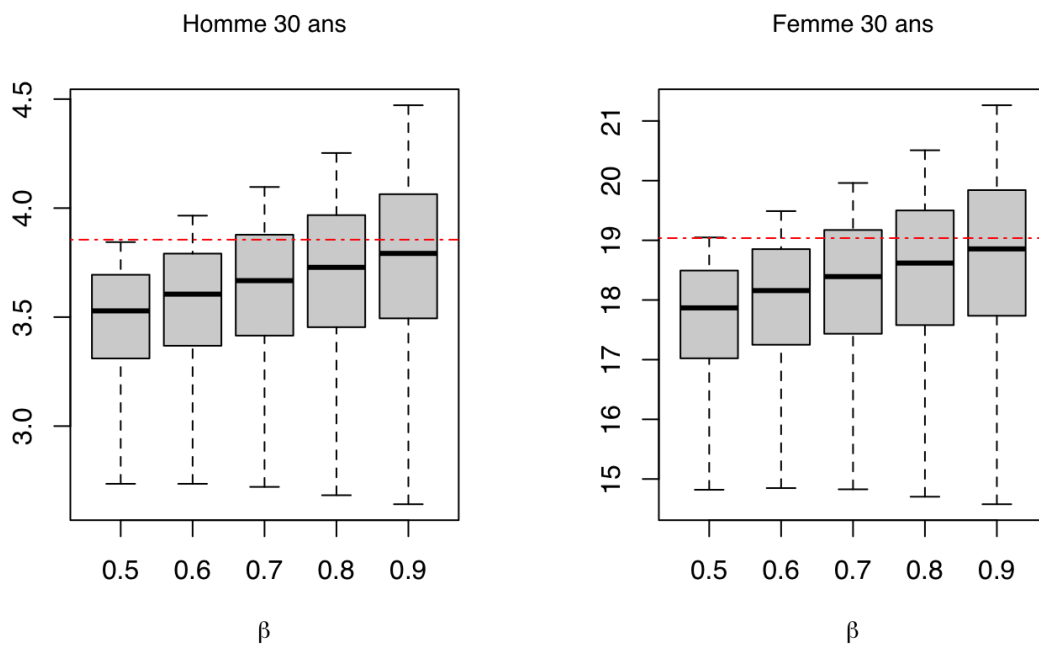


FIGURE 3.6.9 – Variation de la rétention et effet sur les cotisations individuelles sachant que  $S_n < \omega_n$ .

Lorsque le groupe augmente sa rétention, il prend en charge une proportion de plus en plus importante des coûts potentiels sans faire appel à l'assureur. Cela signifie que les fluctuations des pertes sont davantage ressenties par le groupe, entraînant ainsi une augmentation de la volatilité des cotisations dans la tranche inférieure conservée par le groupe. Cette volatilité est liée aux incertitudes inhérentes aux coûts des soins médicaux et aux variations aléatoires dans le montant des demandes de remboursement.

Toutefois, à mesure que la rétention du groupe augmente, le montant total des coûts des soins de santé couvert par l'assureur diminue, car le groupe assume une plus grande part de ces coûts. Cela peut avoir pour effet de réduire l'impact global des sinistres sur l'assureur, rendant ainsi la part du risque qu'il assume plus stable et prévisible. En conséquence, la volatilité des cotisations dans la partie couverte par l'assureur tend à l'inverse à diminuer progressivement.

### 3.6.4 Conclusion

Dans ce chapitre, nous avons étudié la possibilité pour le groupe de transférer une partie du coût total des soins médicaux à un (re-)assureur, ceci afin de réduire la volatilité dans la valeur des cotisations. Lorsque le groupe opte pour une couverture d'assurance par un contrat de type stop-loss ils définissent un certain niveau de rétention selon des objectifs qu'il se fixe (par exemple en terme de redistribution de cash-back en fin de période).

L'opération d'assurance pour un participant en présence du contrat stop-loss consiste dans ce cas en le Paiement en début de période d'une provision fixe pour supporter la couverture de la tranche inférieure du coût total, plus une prime fixe pour la couverture d'assurance et en fin de période, le participant reçoit un cash-back stochastiques ainsi que le remboursement du montant de ses frais médicaux. De cette façon il est garanti qu'en fin de période, tous les participants verront leurs frais médicaux être intégralement remboursés. Ce qui pourrait ne pas être le cas si le pool ne recourt pas à la couverture supplémentaire.

Un autre résultat que nous avons obtenu ici concerne la volatilité des cotisations lorsque le groupe transfère une partie du coût total à un (re-)assureur. En effet, on constate que les cotisations sont beaucoup moins volatiles que lorsque le groupe doit supporter l'entièreté du coût total. Cela permet de garantir une plus grande stabilité dans les montants à charge des participants et aussi limiter les valeurs possibles de ceux-ci à des montants raisonnables.

# Conclusion Générale

Notre étude a permis d'explorer les différents aspects de ce modèle novateur d'assurance basé sur le partage des coûts de soins de santé entre pairs. Nous avons présenté les différentes formes d'assurances collaboratives existantes, mettant en évidence leur diversité et leur adaptabilité aux besoins spécifiques des assurés dans différents domaines. Ce modèle prometteur offre une alternative aux assurances traditionnelles en favorisant la personnalisation des offres et en renforçant l'engagement des assurés au sein du pool.

L'étude de la conditional mean risk sharing en tant que règle de partage a révélé son potentiel pour établir une répartition équitable du coût total des soins médicaux au sein du pool en tenant compte au niveau individuel de ce que chaque participant apporte comme réclamation dans le pool. Elle pourrait être intuitivement attrayante car basée sur un concept familier de moyenne en plus d'être facile à mettre en œuvre. La transparence est totale puisque les participants peuvent être informés dès l'adhésion du mécanisme de remboursement en fonction du coût total observé en fin de période. Des paramètres clés du système se sont révélés être significativement importants pour contrôler la stabilité des cotisations dans le pool, notamment le nombre de participant ainsi que la composition du pool. Le groupe peut ainsi se baser sur ces paramètres pour atteindre des objectifs spécifiques en ce qui concerne la stabilité des cotisations.

Le coût total des frais médicaux peut en principe atteindre des valeurs illimitées, et de ce fait les cotisations peuvent très fortement varier d'une année à une autre, en plus le paiement des cotisations est généralement faite en fin d'année lorsqu'on connaît la valeur du coût total. Il pourrait donc arriver que certains participants soient dans l'incapacité de payer leur cotisations ou ne souhaitent tout simplement pas le faire. De ce fait le groupe pourrait remplacer les cotisations par des primes payées en début de période en recourant à une couverture d'assurance pour la partie excédentaire de leur frais. Ce transfert de risque vers un réassureur offre une opportunité pour l'assurance collaborative de mieux gérer les fluctuations liées aux coûts des soins médicaux. En effet, cela permet de réduire la volatilité des cotisations dans le pool, tout en assurant une stabilité financière plus prévisible. Cette approche permet de remplacer les cotisations variables du système sans assureur par une opération à prime fixe avec la possibilité de redistribution d'un cash back en fin de période aux participants.

Notons que le système présenté ici est conçu pour n'autoriser qu'un seul niveau de franchise annuelle pour tous les participants, ce qui n'offre pas autant de flexibilité que ce que l'on

voudrait aux participants. On pourrait donc envisager un Système pour lequel les participants ont le choix entre différents niveaux de franchise annuelle à l'adhésion. Nous avons fait l'hypothèse au début de ce travail de ne pas faire de distinctions entre les causes d'entrée en hospitalisations. Cependant, il pourrait aussi s'avérer intéressant de les distinguer car cela peut jouer un rôle important pour d'une part estimer le coût total et donc les cotisations versées.

Les assurances collaboratives représentent une voie prometteuse pour l'avenir de l'industrie de l'assurance. Grâce à leur flexibilité, leur capacité à répondre aux besoins spécifiques des participants, et leur approche innovante de partage des risques. Ces systèmes peuvent contribuer à créer un écosystème assurantiel plus transparent, équitable et résilient. Il convient cependant de rester attentif à certains éléments essentiels pour la viabilité du système tels que : –le choix de la rétention lorsqu'il y a recours à un (re-)assureur. En effet, pour assurer la stabilité financière du groupe, il est essentiel de maintenir une rétention adéquate. Une rétention insuffisante pourrait conduire à une dépendance excessive à l'assureur ou à une augmentation des primes pour les membres. –La taille du groupe. La taille du groupe pourrait également être un élément critique pour la viabilité du système car un groupe trop petit pourrait ne pas fournir une diversification suffisante, tandis qu'un groupe trop grand peut rendre la coordination et la communication plus complexes. – Il convient de noter également que les aspects réglementaires, technologiques et opérationnels sont essentiels pour assurer la conformité et la pérennité d'un tel système car opère différemment des assureurs traditionnels.

# Bibliographie

- DENUIT, Michel et Jan DHAENE (sept. 2012). “Convex order and comonotonic conditional mean risk sharing”. In : *Insurance : Mathematics and Economics* 51.2, p. 265-270.
- DENUIT, Michel, Jan DHAENE et Christian Y. ROBERT (sept. 2022). “Risk-sharing rules and their properties, with applications to peer-to-peer insurance”. In : *Journal of Risk and Insurance* 89.3, p. 615-667.
- DENUIT, Michel et Christian Y. ROBERT (jan. 2021a). “From risk sharing to pure premium for a large number of heterogeneous losses”. In : *Insurance : Mathematics and Economics* 96, p. 116-126.
- (juin 2021b). “Risk sharing under the dominant peer-to-peer property and casualty insurance business models”. In : *Risk Management and Insurance Review* 24.2, p. 181-205.
- GOULET, Vincent (2008). “actuar : An R Package for Actuarial Science”. In : *Journal of Statistical Software* 25.7.

## Webographie et autre

- BIBLEREASONS (2022). *Medi-Share Cost*. <https://urlz.fr/jRwl> [En ligne ; Page disponible le 21-Nov-2022].
- CAMPHAUSEN, Florian et al. (jan. 2013). *R Package distr : Object oriented implementation of distributions*.
- US\_LEGISLATIVE\_COUNSEL (2010). *compilation of patient protection and affordable care act*. ACA Section 1501/5000A(d)(2)(B), page 148. <https://urlz.fr/jRwi>.

# Annexe A

## Code R Partie 1

### A.1 Chargement des packages

```
Requirepackages<-c("actuar","distrEx","distr","latex2exp","ggplot2","gridExtra",
  "dplyr","plyr","plotrix","rgdal","rgeos","xtable","mapproj",
  "scales","readr","questionr","labelled","ggpubr","purrr",
  "conflicted","reshape2")

for (i in Requirepackages) {
  if(!require(i,character.only = T)) install.packages(i)
  require(i,character.only = T)
}
conflict_prefer("r","distr")
conflict_prefer("d","distr")
conflict_prefer("p","distr")
conflict_prefer("q","distr")
```

### A.2 Définition des paramètres

#### A.2.1 Paramètres pour les Hommes

Définition des paramètres des lois pour les hommes (voir tableau 2.5.1).

```
Param_M=data.frame(Age = seq(20,80,10),
  Lambda = c(0.033, 0.040, 0.067, 0.120, 0.200, 0.305, 0.483),
  MeanSev = c(84, 100, 104, 104, 108, 115, 134),
  Alpha = rep(0.546,7),
  Beta = c(0.0065, 0.00546, 0.00525, 0.00525, 0.0051,
    0.00475,0.00407))

row.names(Param_M)<-as.character(Param_M$Age)
```

#### A.2.2 Paramètres pour les Femmes

Définition des paramètres des lois pour les femmes (voir tableau 2.5.2).

```
Param_F=data.frame(Age = seq(20,80,10),
                   Lambda = c(0.078, 0.220, 0.111, 0.130, 0.167, 0.246, 0.380),
                   MeanSev = c(79, 90, 96, 104, 115, 133, 158),
                   Alpha = rep(0.754,7),
                   Beta = c(0.00954, 0.00838, 0.00785, 0.00725, 0.00656,
                             0.00567, 0.00477))
row.names(Param_F)<-as.character(Param_F$Age)
```

### A.2.3 Définition de la composition du portefeuille par âge

La composition du groupe telle que donnée ici est issu de la composition de la population belge par âge. Dans ce cas nous avons considéré l'ensemble de la population dont l'âge était compris entre 20 et 80 ans et représenté chaque proportion en fonction du total.

```
Param_M$Prop<-c(0.079,0.085,0.085,0.090,0.077,0.051,0.023)
Param_F$Prop<-c(0.078,0.085,0.084,0.088,0.080,0.060,0.035)

prop.h<-sum(Param_M$Prop)
prop.f<-sum(Param_F$Prop)
```

La fonction suivante permet de retourner le paramètre `param` pour un participant d'âge `Age` et de sexe `Gender`.

```
feat<-function(param=c("Lambda","MeanSev","Alpha","Beta"),Gender=c("M","F"),Age){
  I<-ifelse(Gender=="M", yes = 1, no = 0)
  res<-I*Param_M[as.character(Age),param] + (1-I)*Param_F[as.character(Age),param]
  return(res)
}
```

### A.2.4 test

On définit ici et pour la suite le niveau de franchise  $\delta_i$  pris pour un niveau de référence qui est défini comme étant le coût moyen des soins médicaux pour un homme de 30 ans.

```
agetest=30

# franchise annuelle commune à tous les participants
delta_i<-feat("Lambda","M",agetest)*feat("MeanSev","M",agetest)
cat("delta_i=",delta_i,"\n")

# proportion d'hommes
feat("Lambda","M",agetest)
```

## A.3 Distribution des fréquences et sévérités

### A.3.1 Représentation pour un homme et une femme d'âge = 30 ans

Représentation de la loi probabilité pour les fréquences des hospitalisations et celle des coûts des hospitalisations.

```
# Distribution des fréquences
par(mfrow=c(1,2), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
u=seq(0,5,1)
plot(u,dpois(u,feat("Lambda","M",agetest)),type="b",lty=1,xlab="k",
     ylab=list(TeX('$\\P[N_i=k]$',col="darkblue"))
```

```

lines(u,dpois(u,feat("Lambda","F",agetest)),type="b",lty=4,col="darkred")
legend("topright",inset = 0.02,cex = 0.6,legend = c("Homme","Femme"),
      col = c("darkblue","darkred"),lty = c(1,4),box.lty=0)

# Distribution des sévérités
up1 <- qgamma(0.999,shape=feat("Alpha","M",agetest),rate = feat("Beta","M",agetest))
up2 <- qgamma(0.999,shape=feat("Alpha","F",agetest),rate = feat("Beta","F",agetest))
v1=seq(0,up1,1)
v2=seq(0,up2,1)
plot(v1,dgamma(v1,shape=feat("Alpha","M",agetest),rate = feat("Beta","M",agetest)),
     type="l",xlab="x",ylab=list(TeX('$f_{Ci}(x)$')),col="darkblue")
lines(v2,dgamma(v2,shape=feat("Alpha","F",agetest),rate = feat("Beta","F",agetest)),
     type="l",lty=4,col="darkred")
legend("topright",inset=0.02,cex = 0.6,legend = c("homme","femme"),
      col=c("darkblue","darkred"),lty=c(1,4),box.lty=0)

```

### A.3.2 Mise en évidence de la méthode de discrétisation *unbiased*

Discrétisation de la fonction de répartition des sévérités  $C_{ik}$  sur un intervalle  $[0, b_i]$ , où la borne supérieure  $b_i$  est le quantile d'ordre 99,99% de la loi  $\text{Gamma}(\alpha_i, \beta_i)$ .

```

## Discrétisation des sévérités
a=0
b=up1
step=1
x <- seq(a, b, step)
discr_H <- discretize(pgamma(x, shape = feat("Alpha","M",agetest),
                             rate = feat("Beta","M",agetest)),
                    method = "unbiased",
                    lev = levgamma(x, shape = feat("Alpha","M",agetest),
                                   rate = feat("Beta","M",agetest)),
                    from = a, to = b, step = step)
H<-data.frame(x=x,H=discr_H)

a=0
b=up2
step=1
x <- seq(a, b, step)
discr_F <- discretize(pgamma(x, shape = feat("Alpha","F",agetest),
                             rate = feat("Beta","F",agetest)),
                    method = "unbiased",
                    lev = levgamma(x, shape = feat("Alpha","F",agetest),
                                   rate = feat("Beta","F",agetest)),
                    from = a, to = b, step = step)
F<-data.frame(x=x,F=discr_F)

```

### A.3.3 Représentation

Représentation des masses de probabilités pour les coûts des hospitalisations des hommes et des femmes.

```

## les deux lignes suivantes permettent de conserver les valeur des uniquement
## les valeurs des fonctions discrétisées pour un pas égal à pas
pas=12
H<-subset(H,x %% pas==0)

```

```

F<-subset(F,x %% pas==0)
# Distribution des fréquences
par(mfrow=c(1,3), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
u=seq(0,5,1)

plot(u,dpois(u,feat("Lambda","M",agetest)),type="b",lty=1,xlab="k",
     ylab=list(TeX('$\\P[N_i=k]$',)),col="darkblue")
lines(u,dpois(u,feat("Lambda","F",agetest)),type="b",lty=4,col="darkred")
legend("topright",inset = 0.02,cex = 0.6,legend = c("Homme","Femme"),
      col = c("darkblue","darkred"),lty = c(1,4),box.lty=0)

# Distribution des fréquences

v=seq(0,up1,1)
plot(v,dgamma(v,shape=feat("Alpha","M",agetest),rate = feat("Beta","M",agetest)),
     type="l",xlab="k",ylab=list(TeX('$P[C_{ij}=k]$',)),main="Homme",col="red")
lines(H$x,H$H,type="h",col="black")
legend("topright",inset = 0.02,cex = 0.6,legend = c("exact","discrète"),
      col=c("red","black"),lty=c(1,1),box.lty=0)

v=seq(0,up2,1)
plot(v,dgamma(v,shape=feat("Alpha","F",agetest),rate = feat("Beta","F",agetest)),
     type="l",xlab="k",ylab=list(TeX('$P[C_{ij}=k]$',)),main="Femme",col="red")
lines(F$x,F$F,type="h",col="black")
legend("topright",inset = 0.02,cex = 0.6,legend = c("exact","discrète"),
      col=c("red","black"),lty=c(1,1),box.lty=0)

```

## A.4 Distribution du Coût total individuel $Z_i$

### A.4.1 Calcul des fonctions de répartitions pour tous les groupes d'âges

Le Calcul suivant est celui de la fonction de répartition  $F_{Z_i}(\cdot)$  des variables aléatoires  $Z_i$ . Les fonctions sont obtenues grâce à l'algorithme de Panjer dont l'implémentation à déjà été réalisée via la fonction `aggregateDist()`. Il suffit juste de renseigner la loi des sévérités que nous avons préalablement discrétisé et ainsi que le paramètre de la loi de poisson.

```

FZ<-function(Age,Gender){
  ## Fonction de calcul de la fonction de répartition du coût total individuel
  ## en fonction de l'âge et du sexe du participant
  ## Entrée:
  ## Age: l'âge du participant
  ## Gender: Sexe du participant
  ## Sortie:
  ## Retourne la fonction de répartition complète de la variable  $X_i$ 
  a=0
  b=10*qgamma(0.999,shape=feat("Alpha",Gender,Age),rate = feat("Beta",Gender,Age))
  step=1
  x <- seq(a, b, step)
  ## discrétisation de la fonction de répartition de la gamma entre a et b
  discr <- discretize(pgamma(x, shape = feat("Alpha",Gender,Age),
    rate = feat("Beta",Gender,Age)),
    method = "unbiased",
    lev = levgamma(x, shape = feat("Alpha",Gender,Age),
    rate = feat("Beta",Gender,Age)),

```

```

        from = a, to = b, step = step)
### Aggregate distribution
Fz<- aggregateDist("recursive", model.freq = "poisson", model.sev = discr,
                  lambda = feat("Lambda",Gender,Age), maxit = 100000)
Fz<- DiscreteDistribution(supp = knots(Fz), prob = diff(Fz))
return(Fz)
}

```

#### A.4.2 Calcul de la fonction de répartition pour un homme et pour une femme d'âge égale à 30 ans.

Le code suivant permet dans un premier temps de calculer la fonction de répartition du coût total individuel  $Z_i$  aussi bien pour les hommes que pour les femmes et dans un second temps de transformer le résultat obtenu en une distribution discrète grâce à la fonction `DiscreteDistribution()` de `distr`.

```

## Calcul de la fonction pour les hommes
FZ.H<-FZ(agetest, "M")

## Calcul des fonction pour les femmes
FZ.F<-FZ(agetest, "F")

q(FZ.H)(0.999)

```

#### A.4.3 Représentation de la fonction de répartition pour un homme et une femme d'âge = 30 ans

Représentation des fonctions de répartition du coût total individuel  $Z_i$ .

```

par(mfrow=c(1,2), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
u1=seq(0,q(FZ.H)(0.999),1)
plot(u1,p(FZ.H)(u1), main = list(TeX('Distribution de $Z_i$ Homme')), xlab="k",
    sub = "",ylab = list(TeX('$F_{Z_h}(k)$')),xlim=c(0,q(FZ.H)(0.999)),type= "l")

u2=seq(0,q(FZ.F)(0.999),1)
plot(u2,p(FZ.F)(u2),main = list(TeX('Distribution de $Z_i$ Femme')), sub = "",xlab="k",
    ylab = list(TeX('$F_{Z_f}(k)$')),xlim=c(0,q(FZ.F)(0.999)),type= "l")

rm(u1,u2)

```

### A.5 Distribution du coût partagé $X_i$

Pour rappel, la fonction de répartition  $F_{X_i}(x)$  en un point est obtenue grâce à la relation de la formule (2.3.10).

```

FX <- function(x, Age, Gender, Delta){
  fx<-FZ(Age=Age, Gender = Gender)(x+Delta)
  return(fx)
}

```

Pour la suite nous utiliserons le package `distr` afin convertir la fonction `FX` en un objet reconnu par ce package et ainsi reconstruire la distribution complète de la variable aléatoire  $X$ . Ceci sera fait à l'aide de la fonction `DiscreteDistribution` du package. L'avantage que cela offre est qu'il permet

aisément d'obtenir la densité, la fonction de répartition ainsi que la fonction quantile aisément. Les moments d'ordre 1 et deux sont eux obtenus plus facilement avec cette approche.

Le support `suppest` pris de 0 à 2000 pour être certain d'avoir complètement la masse de probabilité, sinon la fonction `DiscreteDistribution` ne pourra pas fonctionner correctement.

```
supp <- seq(0,2000,1)

# Pour les hommes
probM <- FX(supp,agetest,"M",delta_i) # calcul de la fonction de répartition
pM0 <-FX(0,agetest,"M",delta_i) # masse de proba en 0
Cplt.M <- data.frame(supp = supp, prob = c(pM0,diff(probM)))

## Conversion de la distribution de X en une distribution discrète
FX.H <- DiscreteDistribution(supp = Cplt.M$supp, prob = Cplt.M$prob)

## Pour les femmes
probF <- FX(supp,agetest,"F",delta_i)
pF0 <-FX(0,agetest,"F",delta_i)
Cplt.F <- data.frame(supp = supp, prob = c(pF0,diff(probF)))

## Conversion de la distribution de X en une distribution discrète
FX.F <- DiscreteDistribution(supp = Cplt.F$supp, prob = Cplt.F$prob)

rm(probM,pM0,Cplt.M,probF,pF0,Cplt.F)
```

### A.5.1 Représentation

Représentation des fonctions de répartition des coûts partagés  $X_i$ .

```
par(mfrow=c(1,2), family = "sans", font.lab = 2, font.axis = 1, cex.lab=1,
    cex.main=1, font.main=1,font.lab=1,font.axis=1)
u1 <-seq(0,q(FX.H)(0.999),1)
plot(u1,p(FX.H)(u1),type="l",
     main = list(TeX('Distribution de $X_i$ Homme')), sub = "",xlab="k",
     ylab = list(TeX('$F_{X_i}(k)$'))
points(p(FX.H)(0),type = "p",pch=20)

u2 <-seq(0,q(FX.F)(0.999),1)
plot(u2,p(FX.F)(u2),type="l",
     main = list(TeX('Distribution de $X_i$ Femme')), sub = "",xlab="k",
     ylab = list(TeX('$F_{X_i}(k)$'))
points(p(FX.F)(0),type = "p",pch=20)
```

## A.6 Distribution de $X_i|X_i > 0$

Calcul de la fonction de répartition de la variable  $Y_i = X_i|X_i > 0$ . Voir la formule à l'équation (2.3.11).

```
BX<-function(x, Gender){
  I <- ifelse(Gender=="M",1,0)
  a<-(p(FX.H)(x)-p(FX.H)(0))/(1-p(FX.H)(0))
  b<-(p(FX.F)(x)-p(FX.F)(0))/(1-p(FX.F)(0))

  res <- a*I+(1-I)*b
  return(res)
```

```
}

```

## A.7 Calcul de $S_n$ dans le modèle collectif

### A.7.1 Calcul des fonctions de répartition des sévérités nécessaires pour $S_n$ et de $S_n - X_i$ dans le modèle collectif

LA fonction suivante permet de calculer la fonction de répartition de la formule (2.3.15) et pour un participant donné la fonction de répartition ses sévérités pour  $S_n - X_i$  donc la formule est donnée à l'équation (2.3.16). La fonction de répartition est calculée pour chacun des sous-groupes auquel  $X_i$  peut appartenir. Par exemple si on considère un groupe composé d'hommes et de femmes tous âgés de 30 ans, dans ce cas nous aurons deux sous-groupes : celui des hommes de 30 ans et celui des femmes de 30 ans. un sous-groupe est déterminé par l'âge et le sexe des participants.

Si l'un des membres de ces sous-groupes opte pour une franchise différente de celle du reste du groupe, la fonction LX() suivante permet de calculer les fonctions de répartitions dans ce cas là. Ce cas de figure nous sera utile plus tard pour pouvoir évaluer l'influence de la franchise sur les cotisations lorsque un seul membre dans le groupe opte pour un niveau de franchise annuelle différente, dans ce cas il faut préciser la franchise `delta` de tous les autres participants ainsi que la franchise `delta.I` du participant en question.

```
LX <- function(x,delta,delta.I,Gend.I,n,p){
  # x= valeur de x pour laquelle on calcule L(x)=P[Y'=x] ou L'(x)=P[T=x]
  # delta= franchise
  # delta.I= franchise de l'individu d'intérêt
  # n= nombre d'individus dans le PF
  # p=proportion Homme
  if(Gend.I == "M"){
    nh <- n*p-1 # Nombre d'homme ayant la franchise delta
    nf <- n*(1-p) # Nombre de Femme ayant la franchise delta

    ##Calcul des lambda_i Homme
    q_H<-1-p(FZ.H)(delta)

    ## Calcul des lambda_i Femme
    q_F<-1-p(FZ.F)(delta)

    ## Calcul du lambda_i si I est un homme
    q_I <-1-p(FZ.H)(delta.I)

    ## Lambda total
    lambda_dot <- nh*q_H + nf*q_F + q_I

    ##### Calcul de la fonction de répartition L(x) (voir modèle collectif)
    Val1 <- nh*q_H*BX(x,"M") + q_I*BX(x,"M") + nf*q_F*BX(x,"F")
    Value1= Val1/lambda_dot

    ##### Suppression d'un homme: fonction L'(x)
    ValH = (nh-1)*q_H*BX(x,"M") + q_I*BX(x,"M") + nf*q_F*BX(x,"F")
    ValueH= ValH/(lambda_dot-q_H)

    ##### Suppression d'une femme: fonction L'(x)
    ValF = nh*q_H*BX(x,"M") + q_I*BX(x,"M")+ (nf-1)*q_F*BX(x,"F")
    ValueF= ValF/(lambda_dot-q_F)
  }
}
```

```

#### Suppression de l'individu I: fonction L'(x)
ValI = nh*q_H*BX(x,"M") + nf*q_F*BX(x,"F")
ValueI= ValI/(lambda_dot-q_I)

### Résulta
reslt<-list(LX = ValueI,LXh = ValueH, LXf = ValueF,LXI = ValueI,
           Lda = lambda_dot, LdaH = lambda_dot-q_H, LdaF = lambda_dot-q_F, LdaI= lambda_dot-q_I)
return(reslt)

}else if (Gend.I == "F"){
  nh <- n*p # Nombre d'homme ayant la franchise delta
  nf <- n*(1-p)-1 # Nombre de Femme ayant la franchise delta

  ##Calcul des lambda_i Homme
  q_H<-1-p(FZ.H)(delta)

  ## Calcul des lambda_i Femme
  q_F<-1-p(FZ.F)(delta)

  ## Calcul du lambda_i si I est une femme
  q_I <-1-p(FZ.F)(delta.I)

  ## Lambda total
  lambda_dot <- nh*q_H + nf*q_F + q_I

  #### Calcul de la fonction de répartition L(x) (voir modèle collectif)
  ValI <- nh*q_H*BX(x,"M") + q_I*BX(x,"F") + nf*q_F*BX(x,"F")
  ValueI= ValI/lambda_dot

  #### Suppression d'un homme: fonction L'(x)
  ValH = (nh-1)*q_H*BX(x,"M") + q_I*BX(x,"F") + nf*q_F*BX(x,"F")
  ValueH= ValH/(lambda_dot-q_H)

  #### Suppression d'une femme: fonction L'(x)
  ValF = nh*q_H*BX(x,"M") + q_I*BX(x,"F")+ (nf-1)*q_F*BX(x,"F")
  ValueF= ValF/(lambda_dot-q_F)

  #### Suppression de l'individu I: fonction L'(x)
  ValI = nh*q_H*BX(x,"M") + nf*q_F*BX(x,"F")
  ValueI= ValI/(lambda_dot-q_I)

  ### Résulta
  reslt<-list(LX = ValueI,LXh = ValueH, LXf = ValueF,LXI = ValueI,
           Lda = lambda_dot, LdaH = lambda_dot-q_H, LdaF = lambda_dot-q_F,
           LdaI= lambda_dot-q_I)
  return(reslt)
}else{
  message("Le Gend.I doit etre 'M' ou 'F' ")
}
}

```

### A.7.2 Calcul de la fonction de répartition de $S_n$ : Algorithme de Panjer

### A.7.3 Test : Essais pour un Groupe composé d'hommes et de femmes âgés de 30 ans.

Dans cet exemple on considère que tout les participants ont la même franchise  $\delta_i = \delta = \$4$ .

```

timer<- Sys.time()
# Le nombre de participant dans le groupe
n=1000

# franchise annuelle comune à tous les participants
delta_i <- feat("Lambda", "M", agetest)*feat("MeanSev", "M", agetest)

# proportion d'hommes
p = 0.6

u<-seq(0, 4000, 1)
## Construction des fonctions de répartitions pour les sévérités
F.all<-LX(u, delta = delta_i, delta.I = delta_i,"M", n, p)

## Fonction de répartition de S
FS <- aggregateDist("recursive", model.freq = "poisson", lambda = F.all$Lda,
                    model.sev = diff(F.all$LX), maxit = 100000)
FS <- DiscreteDistribution(supp = knots(FS), prob = diff(FS))

## Fonction de répartition de S-Xi où Xi est un homme
SXH<- aggregateDist("recursive", model.freq = "poisson", lambda = F.all$LdaH,
                    model.sev =diff(F.all$LXh), maxit = 100000)
SXH <- DiscreteDistribution(supp = knots(SXH), prob = diff(SXH))

## Fonction de répartition de S-Xi où Xi est une femme
SXF<- aggregateDist("recursive", model.freq = "poisson", lambda = F.all$LdaF,
                    model.sev =diff(F.all$LXf), maxit = 100000)
SXF <- DiscreteDistribution(supp = knots(SXF), prob = diff(SXF))
print(Sys.time()-timer)

```

### A.7.4 Représentation

Représentation de la fonction de répartition du coût total  $S_n$ , ainsi que de sa fonction masse de probabilité.

```

plf<-q(FS)(0.999)
u<-seq(0,plf,1)
v<-seq(0,1,1/plf)
### Représentation de Fs
par(mfrow=c(1,3), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)

plot(u,d(FS)(u),type="l",xlab="x",ylab="d(x)", main="Masse de probabilité")

plot(u,p(FS)(u),type="l",xlab="q",ylab="p(q)",main="CDF")

plot(v,q(FS)(v),type="l",xlab="p",ylab="q(p)",main="Quantile")

```

### A.7.5 Représentation de la fonction de répartition de $S$ et $S - X_i$

Représentation des fonctions de répartition de  $S_n$  et de  $S_n - X_i$ .

```
### Représentation de FS
par(mfrow=c(1,3), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)

plot(u,d(FS)(u),type="l",xlab="x",ylab="d(x)", main="Masse de probabilité")
lines(u,d(SXH)(u),lty=2, col="darkblue")
lines(u,d(SXF)(u),lty=4, col="darkred")

plot(u,p(FS)(u),type="l",xlab="q",ylab="p(q)",main="CDF")
lines(u,p(SXH)(u),lty=2, col="darkblue")
lines(u,p(SXF)(u),lty=4, col="darkred")

plot(v,q(FS)(v),type="l",xlab="p",ylab="q(p)",main="Quantile")
lines(v,q(SXH)(v),lty=2, col="darkblue")
lines(v,q(SXF)(v),lty=4, col="darkred")
```

## A.8 Calcul Direct de la fonction de répartition de $S$ (modèle binomial composé)

Dans cette partie nous refaisons le calcul de la fonction de répartition de  $S_n$  en faisant un calcul direct par le modèle binomial composé. l'idée est de scinder le coût total  $S_n$  en un coût total  $S_n^h$  pour la partie du groupe composé uniquement des hommes et un coût total  $S_n^f$  pour l'autre partie du groupe composé uniquement des femmes.

### A.8.1 Coût total pour les Femmes

On calcule le coût total  $S_n^h$  des hommes en supposant que ce dernier se calcule comme à l'équation (2.5.2).

```
timer<- Sys.time()
# Le nombre de participant dans le groupe
n=1000

# franchise annuelle commune à tous les participants
delta_i<-feat("Lambda","M",agetest)*feat("MeanSev","M",agetest)

# proportion d'hommes
p = 0.6

# Pr[Xi>0]
qh<-1-p(FZ.H)(delta_i)

# Pr[Xi|Xi>0]

Bx_H<- BX(u,Gender="M")

FH_bin <- aggregateDist("recursive", model.freq = "binom",
                        model.sev = diff(Bx_H), size=n*p,
                        prob =qh ,maxit = 100000)

plf<-quantile(FH_bin,1)
u<-seq(0,plf,1)
```

```
v<-seq(0,1,1/plf)

#### Représentation
par(mfrow=c(1,3), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
     cex.main=0.9, font.main=1,font.lab=1,font.axis=1)

plot(u[1:plf],diff(FH_bin(u)),type="l",xlab="x",ylab="d(x)", main="Masse de probabilité")

plot(u,FH_bin(u),type="l",xlab="q",ylab="p(q)",main="CDF")

plot(v,quantile(FH_bin,v),type="l",xlab="p",ylab="q(p)",main="Quantile")
```

### A.8.2 Coût total pour les Femmes

De même, le calcul de la fonction de répartition  $S_n^f$  se calcule comme à l'équation (2.5.3).

```
# Pr[Xi>0]
qf<-1-p(FZ.F)(delta_i)

# Pr[Xi|Xi>0]
Bx_F<- BX(u,Gender="F")

FF_bin <- aggregateDist("recursive", model.freq = "binom",
                        model.sev = diff(Bx_F), size=(1-p)*n,
                        prob =qf ,maxit = 100000)

plf<-quantile(FF_bin ,1)
u<-seq(0,plf,1)
v<-seq(0,1,1/plf)
#### Représentation de FS

par(mfrow=c(1,3), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
     cex.main=0.9, font.main=1,font.lab=1,font.axis=1)

plot(u[1:plf],diff(FF_bin (u)),type="l",xlab="x",ylab="d(x)",
     main="Masse de probabilité")

plot(u,FF_bin (u),type="l",xlab="q",ylab="p(q)",main="CDF")

plot(v,quantile(FF_bin,v),type="l",xlab="p",ylab="q(p)",main="Quantile")
```

### A.8.3 Convolution coût total Hommes + Femmes ==> Coût total pour l'ensemble du groupe

Une fois les deux fonctions de répartitions calculées, la fonction de répartition du coût total  $S_n$  s'obtient en calculant le produit de convolution de  $S_n^h$  et  $S_n^f$ ,  $S_n = S_n^h \star S_n^f$ .

```
plf1<-quantile(FH_bin,1)
u<-seq(0,plf1,1)
HT<-DiscreteDistribution(supp = u[1:plf1], prob = diff(FH_bin(u)))

plf2<-quantile(FF_bin,1)
u<-seq(0,plf2,1)
FT<-DiscreteDistribution(supp = u[1:plf2], prob = diff(FF_bin(u)))
```

```

### Calcul du produit de convolution des coûts totaux HT et FT
timer<-Sys.time()
Stot<- HT + FT
print(Sys.time()-timer)

plf<-q(Stot)(1)
u<-seq(0,plf,1)
v<-seq(0,1,1/plf)
Sc.vs.S<-data.frame(s = u, Sc = d(FS)(u), S = d(Stot)(u))

#### Représentation
par(mfrow=c(1,3), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)

plot(u,d(Stot)(u),type="l",xlab="x",ylab="d(x)",
     main="Masse de probabilité",col="darkred")
lines(u,d(FS)(u),lty=4,col="darkblue")
legend("topleft", inset = 0.02,cex = 0.6,legend =c("Exacte","Scoll"),
      col=c("darkred","darkblue"),lty=c(1,4),box.lty=0)

plot(u,p(Stot)(u),type="l",xlab="q",ylab="p(q)",main="CDF",col="darkred")
lines(u,p(FS)(u),lty=4,col="darkblue")
legend("topleft", inset = 0.02,cex = 0.6,legend =c("Exacte","Scoll"),
      col=c("darkred","darkblue"),lty=c(1,4),box.lty=0)

plot(v,q(Stot)(v),type="l",xlab="p",ylab="q(p)",main="Quantile",col="darkred")
lines(v,q(FS)(v),lty=4,col="darkblue")
legend("topleft", inset = 0.02,cex = 0.6,legend =c("Exacte","Scoll"),
      col=c("darkred","darkblue"),lty=c(1,4),box.lty=0)

```

## A.9 Espérance Conditionnelle : CMRS

La fonction CondExp suivante permet d'implémenter la formule (2.3.17) du calcul de l'espérance conditionnelle.

```

CondExp<-function(s,delt,Fs1,SXH1,SXF1){
  ## Fonction de calcul de l'espérance conditionnelle de XH ou XF sachant S=s
  ## s: réalisation de S
  ## Fs1: Fonction de répartition de S
  ## SXH1: Fonction de répartition de S-XH
  ## SXF1: Fonction de répartition de S-XF
  ## Retourne E[XH|S=s] et E[XF|S=s]
  if(s==0){
    ph=0
    pf=0
  } else{
    ps=p(Fs1)(s)-p(Fs1)(s-1) #Pr[S=s]
    ## hommes
    px=p(FZ.H)(1:s+delt)-p(FZ.H)(1:s+delt-1) # Pr[Xi=k] k allant de 1 à s
    psx=p(SXH1)(s-1:s)-p(SXH1)(s-1:s-1) # Pr[S-Xi=s-k]
    ph=sum((1:s)*px*psx/ps) # E[Xi|S=s]

    ## femmes
    py=p(FZ.F)(1:s+delt)-p(FZ.F)(1:s+delt-1)
    psy=p(SXF1)(s-1:s)-p(SXF1)(s-1:s-1)
  }
}

```

```

    pf=sum((1:s)*py*psy/ps)
  }
  return(list(EH=ph,EF=pf))
}
timer<-Sys.time()

A<-CondExp(s=q(FS)(0.999),delt=4,FS,SXH,SXF)

Sys.time()-timer
A
(n*p*A$EH+(1-p)*n*A$EF)/q(FS)(0.999)
rm(A)

```

## A.10 Approximation de la CMRS par le modèle de régression linéaire.

Avec la fonction `Exp.Reg`, on applique le calcul de la règle de partage linéaire (2.4.1).

```

Exp.Reg<-function(s,delt,FS,n,p){
  ## Fonction de calcul de l'espérance conditionnelle de XH ou XF sachant S=s
  ## s: réalisation de S
  ## nh: Nombre d'hommes
  ## nf: Nombre de femmes
  ## FS: Fonction de répartition de SF
  ## Retourne E[XH|S=s] et E[XF|S=s]
  mean.H <- E(FX.H) # espérance de Xi
  var.H <- distr::sd(FX.H)^2 # variance de Xi
  mean.F <- E(FX.F)
  var.F <- distr::sd(FX.F)^2
  var.S <- distr::sd(FS)^2
  # Calcul du partage par la regression linéaire

  nh <- n*p
  nf <- n*(1-p)
  #A_H.reg<- var.H/(nh*var.H + nf*var.F)
  #A_F.reg<- var.F/(nh*var.H + nf*var.F)

  A_H.reg<- var.H/var.S
  A_F.reg<- var.F/var.S

  ## hommes
  ph = (mean.H + A_H.reg*(s-E(FS)))*ifelse(s==0,0,1)

  # femmes
  pf = (mean.F + A_F.reg*(s-E(FS)))*ifelse(s==0,0,1)

  return(list(EH=ph,EF=pf))
}
B<-Exp.Reg(s=q(FS)(0.999),delt=delta_i,FS,1500,0.6)
B

```

## A.11 Étude de sensibilité de la CMRS

### A.11.1 Au nombre de participants $n$

Ce bloc permet de calculer la fonction de répartition de  $S_n$  pour les différentes valeurs que prennent  $n$ . l'ensemble des fonctions de répartitions calculées est stockée dans une liste dont on se servira pour calculer l'espérance conditionnelle plus tard.

Pour le calcul de la CMRS en fonction des valeurs de  $S_n$ , nous faisons varier  $s$  entre 0 et  $s^* = F_{S_n}^{-1}(0.9999)$ . Puisque la distribution de  $S_n$  est discrète, le calcul de la fonction de répartition par la fonction `\verb|aggregateDist()` fait en sorte que au delà de la valeur ( $s > F_{S_n}^{-1}(-1)$ ), on a toujours  $F_{S_n}(s) = 1$  et dans ce cas là,  $P[S_n = s] = F_{S_n}(s) - F_{S_n}(s - 1) = 0$ , ceci rend impossible le calcul de la CMRS pour des valeurs de  $s$  plus grand que  $F_{S_n}^{-1}(-1)$  puisque cette probabilité intervient dans le calcul de la CMRS au dénominateur. On pourrait résoudre ce problème en remplaçant l'argument `\verb|maxit|` de `\verb|aggregateDist()` par `\verb|tol|` qui vas préciser la tolérance pour le calcul de la fonction de répartition, c'est à dire de combien la fonction de répartition calculée doit s'éloigner au maximum de 1 avant l'arrêt des récursions. Cependant cela ne semble pas très bien fonctionner dans notre cas, raison pour laquelle nous limitons le calcul de l'espérance à certaine valeurs de  $s$  pour un groupe donné. Le même constat est fait pour le calcul des fonctions de répartitions de  $S_n - X_i$ , ainsi pour éviter tout biais dans les calculs, nous limitons la plage de valeur de  $s$  à  $\min(F_{S_n}^{-1}(1), F_{S_n - X_i}^{-1}(1))$ .

```
n <- c(100,200,500,1000,1500,2000,3000,4000)
#n<-c(100,200,500,1000,1500,2000)
delta_i <- feat("Lambda","M",agetest)*100 # franchise annuelle
p<- 0.6
Fpart.S.n <- list()
Fpart.SXH.n <- list()
Fpart.SXF.n <- list()
Quant_S.n <- rep(NA,length(n))
Num_part <- data.frame(n=n,Nh=rep(NA,length(n)),Nf=rep(NA,length(n)))
timer <- Sys.time()
for (i in 1:length(n)) {

  U <- seq(0,4000,1)
  ### Fonctions de répartitions
  fc <- LX(U, delta = delta_i, delta.I = delta_i,"M", n[i], p)
  ## S
  Fc <- aggregateDist("recursive", model.freq = "poisson",
                      model.sev = diff(fc$LX), lambda = fc$Lda, maxit = 100000)
  Fc <- DiscreteDistribution(supp = knots(Fc), prob = diff(Fc))
  ## S-Xh
  Fch<- aggregateDist("recursive", model.freq = "poisson",
                      model.sev = diff(fc$LXh), lambda = fc$LdaH, maxit = 100000)
  Fch <- DiscreteDistribution(supp = knots(Fch), prob = diff(Fch))
  ## S-Xf
  Fcf <- aggregateDist("recursive", model.freq = "poisson",
                      model.sev = diff(fc$LXf), lambda = fc$LdaF, maxit = 100000)
  Fcf <- DiscreteDistribution(supp = knots(Fcf), prob = diff(Fcf))

  Quant_S.n[i] <- min(q(Fc)(0.9999),q(Fch)(0.9999),q(Fcf)(0.9999))
  Fpart.S.n[[length(Fpart.S.n)+1]] <- Fc
  Fpart.SXH.n[[length(Fpart.SXH.n)+1]] <- Fch
  Fpart.SXF.n[[length(Fpart.SXF.n)+1]] <- Fcf
  rm(fc,Fc,Fch,Fcf)
}
print(Sys.time()-timer)
```

```
names(Fpart.S.n) <- paste("N", as.character(n), sep = "")
names(Fpart.SXH.n) <- paste("N", as.character(n), sep = "")
names(Fpart.SXF.n) <- paste("N", as.character(n), sep = "")
```

#### A.11.1.1 Calcul des cotisations $h_{i,n}^*(s)$

Ce block permet de calculer la valeur des cotisations suivant les différentes valeurs de  $n$ .

```
timer<-Sys.time()
n <- c(100,200,500,1000,1500,2000,3000,4000)
# franchise annuelle commune à tous les participants
delta_i <- feat("Lambda","M",agetest)*feat("MeanSev","M",agetest)

# proportion d'hommes
p = 0.6

CMRS.NH <- list()
CMRS.NF <- list()
timer <- Sys.time()
for (i in 1:length(n)) {
  S <- seq(0,Quant_S.n[i],10)
  S.Quant <- S/Quant_S.n[i]

  ESP_h <- data.frame(S = S, S.Quant=S.Quant,
                     EXP= rep(NA,length(S)), EXP.reg = rep(NA,length(S)))
  ESP_f <- data.frame(S = S, S.Quant=S.Quant,
                     EXP = rep(NA,length(S)), EXP.reg = rep(NA,length(S)))
  for (j in 1:length(S)) {
    #conditional mean risk sharing
    A <- CondExp(s=S[j], delt=delta_i, Fpart.S.n[[i]],
                Fpart.SXH.n[[i]], Fpart.SXF.n[[i]])
    ESP_h$EXP[j] <- A$EH
    ESP_f$EXP[j] <- A$EF

    #Approximation linéaire
    B <- Exp.Reg(s=S[j],delt=delta_i,Fpart.S.n[[i]],n[i],p)
    ESP_h$EXP.reg[j] <- B$EH
    ESP_f$EXP.reg[j] <- B$EF
  }

  CMRS.NH[[length(CMRS.NH)+1]] <- ESP_h
  CMRS.NF[[length(CMRS.NF)+1]] <- ESP_f
  rm(S,ESP_h,ESP_f,A,B)
}
print(Sys.time() - timer)
names(CMRS.NH) <- paste("N", as.character(n), sep = "")
names(CMRS.NF) <- paste("N", as.character(n), sep = "")
#save(CMRS.NH,file="CMRS_Nh.Rdata")
#save(CMRS.NF,file="CMRS_Nf.Rdata")
```

#### A.11.1.2 Loi de probabilité des $\mathbb{E}[X_i|S_n]$

Ce block permet pour chaque valeur de  $n$ , de générer un vecteur aléatoire  $S$  de longueur long et contenant les valeurs des réalisations  $s$  de  $S_n$ . chaque valeur  $s$  est générée à partir de la fonction de répartition correspondante à la valeur de  $n$ . Le but étant de trouver les fonctions densités de probabilités des cotisations.

```

timer<-Sys.time()
n <- c(100,200,500,1000,1500,2000,3000,4000)
N.ind <- c(1,2,4,6) # inndice des valeurs de n dans le vecteur n ci-dessus
# franchise annuelle commune à tous les participants
delta_i <- feat("Lambda","M",agetest)*feat("MeanSev","M",agetest)

# proportion d'hommes
p = 0.6

# nombre de valeurs de s à générer
long <- 15000

Law.N <- list()

timer <- Sys.time()
for (i in 1:length(n)) {
  S <- r(Fpart.S.n[[i]])(long)

  ESP <- data.frame(s=S, Exp.H=rep(NA,long), Exp.F=rep(NA,long),
                    Prob = p(Fpart.S.n[[i]])( S))

  for (j in 1:length(S)) {
    #conditional mean risk sharing
    A <- CondExp(s=S[j], delt=delta_i, Fpart.S.n[[i]],
                Fpart.SXH.n[[i]], Fpart.SXF.n[[i]])
    ESP$Exp.H[j] <- A$EH
    ESP$Exp.F[j] <- A$EF
  }

  Law.N[[length(Law.N)+1]] <- ESP

  rm(S,ESP,A)
}
print(Sys.time() - timer)
names(Law.N) <- paste("N", as.character(n), sep = "")
save(Law.N,file="CMRS_Law.Rdata")

```

### A.11.1.3 Représentation 1 : Homme Vs Femme

Comparaison des valeurs des cotisations entre celles des hommes et celle des femmes.

```

### Hommes vs femme n=1500
tr <- which(n==1500)
plf <- Quant_S.n[tr]
u <- seq(0, plf,1)
par(family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
plot(CMRS.NH[[tr]]$S, CMRS.NH[[tr]]$EXP,
     ylim = c(0,max(CMRS.NH[[tr]]$EXP, CMRS.NF[[tr]]$EXP)),
     type = "l", xlab = "s", ylab = list(TeX('$E[X_i|S=s]$')), col = "darkblue")
lines(CMRS.NF[[5]]$S,CMRS.NF[[tr]]$EXP, type = "l",lty=4,col = "darkred")
lines(u, d(Fpart.S.n[[tr]])(u)*(1.3e5), type = "l",xlab = "x",ylab = "d(x)",
      col="black")
legend("topleft", cex = 0.6,legend = c("homme","femme","densité"),
      col = c("darkblue","darkred","black"),lty=c(1,4),box.lty=0,inset = 0.02)

```

#### A.11.1.4 Représentation 2 : contribution relatives de chaque groupe

Représentation de contributions relatives des hommes et des femmes par rapport au coût total observé.

```
### Hommes vs femme n=1500
tr <- which(n==1500)
p = 0.6
nh <- p*1500
nf <- 1500-nh
Rel.Exp <- data.frame(S=CMRS.NH[[tr]]$, Hoe=nh*CMRS.NH[[tr]]$EXP/CMRS.NH[[tr]]$,
                      Fem=nf*CMRS.NF[[tr]]$EXP/CMRS.NF[[tr]]$)
Rel.Exp[1,] <- 0
Rel.Exp$tot <- Rel.Exp$Hoe+Rel.Exp$Fem

Rel.Exp <- subset(Rel.Exp,S %>% 1500==0)

par(family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
     cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
plot(Rel.Exp$S[-1],Rel.Exp$Hoe[-1],ylim=c(0,max(Rel.Exp$Hoe,Rel.Exp$Fem)),
     type="l",xlab="s",ylab=list(TeX('$E[X_i|S_n=s]$')),col="darkblue")
lines(Rel.Exp$S[-1],Rel.Exp$Fem[-1],type="l",col="red")
```

#### A.11.1.5 Représentation 3 : CMRS Vs Regression linéaire

Comparaison de la conditional mean risk sharing (2.3.17) et de la règle de partage linéaire (2.4.1).

```
n<-c(100,200,500,1000,1500,2000,3000,4000)
### Hommes vs femme n=1500
tr<-which(n==1500)
plf<-Quant_S.n[tr]
u<-seq(0,plf,1)

par(family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
     cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
plot(CMRS.NH[[tr]]$S[-1],CMRS.NH[[tr]]$EXP.reg[-1],
     ylim=c(0,max(CMRS.NH[[tr]]$EXP.reg,CMRS.NF[[tr]]$EXP.reg)),
     type="l",xlab="s",ylab=list(TeX('$E[X_i|S_n=s]$')),col="darkblue")

lines(CMRS.NH[[tr]]$S,CMRS.NH[[tr]]$EXP,type="l",lty=4,col="blue")
lines(CMRS.NF[[tr]]$S[-1],CMRS.NF[[tr]]$EXP.reg[-1],type="l",col="darkred")
lines(CMRS.NF[[tr]]$S,CMRS.NF[[tr]]$EXP,type="l",lty=4,col="red")
lines(u,d(Fpart.S.n[[tr]])(u)*(1.3e5),
      type="l",xlab="x",ylab="d(x)",col="black")
legend("topleft",inset = 0.02,cex = 0.6,
      legend =c("homme Reg","homme","femme Reg","femme",TeX('$P[S_n=s]$')),
      col=c("darkblue","blue","darkred","red","black"),lty=c(1,4),box.lty=0)
```

#### A.11.1.6 Représentation 4 : Loi de probabilité en fonction de $n$

Ici nous représentons la fonction densité de probabilité de  $h_{i,n}^*(S_n)$  en fonction des différentes valeurs de  $n$ .

```
timer <- Sys.time()
pdf.N.H <- list()
pdf.N.F <- list()
for (i in 1:length(n)) {
  den.H <- density(Law.N[[i]]$Exp.H)
```

```

den.F <- density(Law.N[[i]]$Exp.F)
A <- data.frame(val=den.H$x, prob=den.H$y)
B <- data.frame(val=den.F$x, prob=den.F$y)
pdf.N.H[[length(pdf.N.H)+1]] <- A
pdf.N.F[[length(pdf.N.F)+1]] <- B
rm(den.H, den.F, A, B)
}
print(Sys.time()-timer)
names(pdf.N.H) <- paste("N", as.character(n), sep = "")
names(pdf.N.F) <- paste("N", as.character(n), sep = "")

par(mfrow=c(1,2),font.main=1,cex.main = 0.9, cex.sub = 0.8)

plot(pdf.N.H[[1]]$val, pdf.N.H[[1]]$prob, type = "l", lty = 1,
     col = "darkblue", ylim = c(0,0.8), xlim = c(0,10),
     xlab = list(TeX('$E[X_i|S_n]$')), ylab = "densité",main = "Homme")
lines(pdf.N.H[[3]]$val, pdf.N.H[[3]]$prob, type = "l", lty = 2,
      col = "darkred")
lines(pdf.N.H[[4]]$val, pdf.N.H[[4]]$prob, type = "l", lty = 4,
      col = "darkgreen")
lines(pdf.N.H[[6]]$val, pdf.N.H[[6]]$prob, type = "l", lty = 6,
      col = "black")
legend("topright", cex = 0.6,legend = c("N=100","N=500","N=1000","N=2000"),
      col = c("darkblue","darkred","darkgreen","black"),lty=c(1,2,4,6),
      box.lty=0,inset = 0.005)

plot(pdf.N.F[[1]]$val, pdf.N.F[[1]]$prob, type = "l", lty = 1,
     col = "darkblue", ylim = c(0,0.3), xlim = c(5,30),
     xlab = list(TeX('$E[X_i|S_n]$')), ylab = "densité",main = "Femme")
lines(pdf.N.F[[3]]$val, pdf.N.F[[3]]$prob, type = "l", lty = 2,
      col = "darkred")
lines(pdf.N.F[[4]]$val, pdf.N.F[[4]]$prob, type = "l", lty = 4,
      col = "darkgreen")
lines(pdf.N.F[[6]]$val, pdf.N.F[[6]]$prob, type = "l", lty = 6,
      col = "black")
legend("topright", cex = 0.6,legend = c("N=100","N=500","N=1000","N=2000"),
      col = c("darkblue","darkred","darkgreen","black"),lty=c(1,2,4,6),
      box.lty=0,inset = 0.005)

```

### A.11.1.7 Représentation 5 : Boxplot

Mise en évidence de la variabilité des cotisations autour de la prime pure.

```

n<-c(100,200,500,1000,1500,2000,3000,4000)
Bx.pltH<-list()
Bx.pltF<-list()
for (i in 1:length(Law.N)) {
  Bx.pltH[[length(Bx.pltH)+1]]<-Law.N[[i]]$Exp.H
  Bx.pltF[[length(Bx.pltF)+1]]<-Law.N[[i]]$Exp.F
}
names(Bx.pltH)<-as.character(n)
names(Bx.pltF)<-as.character(n)
Bx.pltH.long<-melt(do.call(rbind,Bx.pltH))
Bx.pltF.long<-melt(do.call(rbind,Bx.pltF))

par(font.main=1,cex.main = 0.9, cex.sub = 0.8)

```

```

boxplot(value~Var1,Bx.pltH.long,xlab ="n",ylab = "", main = "Homme")
  abline(h=E(FX.H),lwd = 1,lty=4,col="red")

boxplot(value~Var1,Bx.pltF.long,xlab ="n",ylab = "", main = "Femme")
  abline(h=E(FX.F),lwd = 1,lty=4,col="red")

```

## A.11.2 A la franchise

Dans cette partie, nous étudierons l'influence d'un changement de franchise d'un participant en particulier sur ses cotisations, la franchise des autres membres du groupe restant la même.

### A.11.2.1 Cas d'un homme

On suppose ici qu'il y a un seul homme dont on fait varier la franchise dans le groupe, tout en laissant la valeur de la franchise pour le reste du groupe inchangée.

```

n<-1500
# franchise annuelle du participant d'intérêt
delta<-seq(0,20,4)
# franchise annuelle commune à tous les autres participants
delta_i<-feat("Lambda","M",agetest)*feat("MeanSev","M",agetest)
p<- 0.6

Fpart.SH.D<-list()
Fpart.SXH.D<-list()
Fpart.SXF.D<-list()
Quant_SH.D<-rep(NA,length(delta))

timer<-Sys.time()
for (i in 1:length(delta)) {
  U <- seq(0,4000,1)
  ### Fonctions de répartitions
  fc <- LX(U, delta = delta_i, delta.I = delta[i],"M", n, p)
  ## S
  Fc <- aggregateDist("recursive", model.freq = "poisson",
                      model.sev = diff(fc$LX), lambda = fc$Lda, maxit = 100000)
  Fc <- DiscreteDistribution(supp = knots(Fc), prob = diff(Fc))
  ## S-XI
  Fch<- aggregateDist("recursive", model.freq = "poisson",
                      model.sev = diff(fc$LXI), lambda = fc$LdaI, maxit = 100000)
  Fch <- DiscreteDistribution(supp = knots(Fch), prob = diff(Fch))
  ## S-Xf
  Fcf<- aggregateDist("recursive", model.freq = "poisson",
                      model.sev = diff(fc$LXf), lambda = fc$LdaF, maxit = 100000)
  Fcf <- DiscreteDistribution(supp = knots(Fcf), prob = diff(Fcf))

  Quant_SH.D[i]<-min(q(Fc)(0.9999),q(Fch)(0.9999),q(Fcf)(0.9999))

  Fpart.SH.D[[length(Fpart.SH.D)+1]]<-Fc
  Fpart.SXH.D[[length(Fpart.SXH.D)+1]]<-Fch
  Fpart.SXF.D[[length(Fpart.SXF.D)+1]]<-Fcf
  rm(fc,Fc,Fch,Fcf)
}
print(Sys.time()-timer)
names(Fpart.SH.D)<-paste("D", as.character(delta), sep = "")
names(Fpart.SXH.D)<-paste("D", as.character(delta), sep = "")

```

```

names(Fpart.SXF.D)<-paste("D", as.character(delta), sep = "")

timer<-Sys.time()
n<-1500
# franchise annuelle du participant d'intérêt
delta<-seq(0,20,4)
# franchise annuelle commune à tous les autres participants
delta_i<-feat("Lambda","M",agetest)*feat("MeanSev","M",agetest)
# proportion d'hommes
p = 0.6

CMRS.DH<-list()
timer<-Sys.time()
for (i in 1:length(delta)) {
  S <- seq(0,Quant_SH.D[i],10)
  SH.Quant<-S/Quant_SH.D[i]

  ESP_h<- data.frame(S = S, S.Quant=SH.Quant,
                    EXP= rep(NA,length(S)))
  for (j in 1:length(S)) {
    #conditional mean risk sharing
    A<-CondExp(s=S[j],delt=delta[i],Fpart.SH.D[[i]],
              Fpart.SXH.D[[i]],Fpart.SXF.D[[i]])
    ESP_h$EXP[j]<- A$EH
  }

  CMRS.DH[[length(CMRS.DH)+1]]<-ESP_h
  rm(S,ESP_h,A)
}
print(Sys.time()-timer)
names(CMRS.DH)<-paste("D", as.character(delta), sep = "")
save(CMRS.DH,file="CMRS_Dh.Rdata")

```

```

n<-1500
# franchise annuelle du participant d'intérêt
delta<-seq(0,20,4)
# franchise annuelle commune à tous les autres participants
delta_i<-feat("Lambda","M",agetest)*feat("MeanSev","M",agetest)
p<- 0.6

Fpart.SF.D<-list()
Fpart.XH.D<-list()
Fpart.XF.D<-list()
Quant_SF.D<-rep(NA,length(delta))

timer<-Sys.time()
for (i in 1:length(delta)) {
  U <- seq(0,4000,1)
  ### Fonctions de répartition
  fc <- LX(U, delta = delta_i, delta.I = delta[i], "F", n, p)
  ## S
  Fc <- aggregateDist("recursive", model.freq = "poisson",
                    model.sev = diff(fc$LX), lambda = fc$Lda, maxit = 100000)

```

```

Fc <- DiscreteDistribution(supp = knots(Fc), prob = diff(Fc))
## S-Xh
Fch<- aggregateDist("recursive", model.freq = "poisson",
                    model.sev = diff(fc$LXh), lambda = fc$LdaH, maxit = 10000)
Fch <- DiscreteDistribution(supp = knots(Fch), prob = diff(Fch))
## S-XI
Fcf<- aggregateDist("recursive", model.freq = "poisson",
                    model.sev = diff(fc$LXI), lambda = fc$LdaI, maxit = 10000)
Fcf <- DiscreteDistribution(supp = knots(Fcf), prob = diff(Fcf))

Quant_SF.D[i]<-min(q(Fc)(0.9999),q(Fch)(0.9999),q(Fcf)(0.9999))
Fpart_SF.D[[length(Fpart_SF.D)+1]]<-Fc
Fpart_XH.D[[length(Fpart_XH.D)+1]]<-Fch
Fpart_XF.D[[length(Fpart_XF.D)+1]]<-Fcf
rm(fc,Fc,Fch,Fcf)
}
print(Sys.time()-timer)
names(Fpart_SF.D)<-paste("D", as.character(delta), sep = "")
names(Fpart_XH.D)<-paste("D", as.character(delta), sep = "")
names(Fpart_XF.D)<-paste("D", as.character(delta), sep = "")

timer<-Sys.time()
n<-1500
# franchise annuelle du participant d'intérêt
delta<-seq(0,20,4)
# franchise annuelle commune à tous les autres participants
delta_i<-feat("Lambda","M",agetest)*feat("MeanSev","M",agetest)
# proportion d'hommes
p = 0.6

CMRS.DF<-list()
timer<-Sys.time()
for (i in 1:length(delta)) {
  S <- seq(0,Quant_SF.D[i],10)
  SF.Quant<-S/Quant_SF.D[i]

  ESP_f<- data.frame(S = S, S.Quant=SF.Quant,
                    EXP= rep(NA,length(S)))
  for (j in 1:length(S)) {
    #conditional mean risk sharing
    A<-CondExp(s=S[j],delt=delta[i],Fpart.SH.D[[i]],
              Fpart.XH.D[[i]],Fpart.XF.D[[i]])
    ESP_f$EXP[j]<- A$EF
  }

  CMRS.DF[[length(CMRS.DF)+1]]<-ESP_f
  rm(S,ESP_f,A)
}
print(Sys.time()-timer)
names(CMRS.DF)<-paste("D", as.character(delta), sep = "")
save(CMRS.DF,file="CMRS_Df.Rdata")

```

### A.11.2.3 Loi de probabilité des cotisations par rapport à la franchise

Ce block permet pour chaque valeur de  $\delta$  du participant d'intérêt, de générer un vecteur aléatoire  $S$  de longueur `long` et contenant les valeurs des réalisations  $s$  de  $S_n$ . chaque valeur  $s$  est générée à partir

de la fonction de répartition correspondante à la valeur de  $\delta$ .

```

timer<-Sys.time()
n<-1500
# franchise annuelle du participant d'intérêt
delta<-seq(0,20,4)
# franchise annuelle commune à tous les autres participants
delta_i<-feat("Lambda","M",agetest)*feat("MeanSev","M",agetest)
p<- 0.6
# nombre de valeurs de s à générer
long <- 15000

Law.D <- list()

timer <- Sys.time()
for (i in 1:length(delta)) {
  SH <- r(Fpart.SH.D[[i]])(long)
  SF <- r(Fpart.SF.D[[i]])(long)

  ESP <- data.frame(S.H=SH,S.F=SF, Exp.H=rep(NA,long), Exp.F=rep(NA,long))

  for (j in 1:long) {
    #conditional mean risk sharing
    A <- CondExp(s=SH[j], delt=delta_i, Fpart.SH.D[[i]],
                Fpart.SXH.D[[i]], Fpart.SXH.D[[i]])
    B <- CondExp(s=SF[j], delt=delta_i, Fpart.SF.D[[i]],
                Fpart.XF.D[[i]], Fpart.XF.D[[i]])
    ESP$Exp.H[j] <- A$EH
    ESP$Exp.F[j] <- B$EF
  }

  Law.D[[length(Law.D)+1]] <- ESP

  rm(SH,SF,ESP,A,B)
}
print(Sys.time() - timer)
names(Law.D) <- paste("D", as.character(delta), sep = "")
save(Law.D,file="CMRS_Law.D.Rdata")

```

#### A.11.2.4 Représentation 1 : Loi de probabilité en fonction de $\delta$

Ici nous représentons la fonction densité de probabilité de  $h_{i,n}^*(S_n)$  en fonction des différentes valeurs de  $\delta$ .

```

timer <- Sys.time()
pdf.D.H <- list()
pdf.D.F <- list()
for (i in 1:length(delta)) {
  den.H <- density(Law.D[[i]]$Exp.H)
  den.F <- density(Law.D[[i]]$Exp.F)
  A <- data.frame(val=den.H$x, prob=den.H$y)
  B <- data.frame(val=den.F$x, prob=den.F$y)
  pdf.D.H[[length(pdf.D.H)+1]] <- A
  pdf.D.F[[length(pdf.D.F)+1]] <- B
  rm(den.H, den.F, A, B)
}
print(Sys.time()-timer)

```

```

names(pdf.D.H ) <- paste("D", as.character(delta), sep = "")
names(pdf.D.F ) <- paste("D", as.character(delta), sep = "")

par(mfrow=c(1,2),font.main=1,cex.main = 0.9, cex.sub = 0.8)

plot(pdf.D.H[[1]]$val, pdf.D.H[[1]]$prob, type = "l", lty = 1,
     col = "darkblue", ylim = c(0,0.8), xlim = c(2,7),
     xlab = list(TeX('$E[X_i|S_n]$')), ylab = "densité",main = "Homme")
lines(pdf.D.H[[3]]$val, pdf.D.H[[3]]$prob, type = "l", lty = 2,
      col = "darkred")
lines(pdf.D.H[[4]]$val, pdf.D.H[[4]]$prob, type = "l", lty = 4,
      col = "darkgreen")
lines(pdf.D.H[[6]]$val, pdf.D.H[[6]]$prob, type = "l", lty = 6,
      col = "black")
legend("topright", cex = 0.6,legend = c("D=0","D=8","D=12","D=20"),
      col = c("darkblue","darkred","darkgreen","black"),lty=c(1,2,4,6),
      box.lty=0,inset = 0.005)

plot(pdf.D.F[[1]]$val, pdf.D.F[[1]]$prob, type = "l", lty = 1,
     col = "darkblue", ylim = c(0,0.25), xlim = c(10,28),
     xlab = list(TeX('$E[X_i|S_n]$')), ylab = "densité",main = "Femme")
lines(pdf.D.F[[3]]$val, pdf.D.F[[3]]$prob, type = "l", lty = 2,
      col = "darkred")
lines(pdf.D.F[[4]]$val, pdf.D.F[[4]]$prob, type = "l", lty = 4,
      col = "darkgreen")
lines(pdf.D.F[[6]]$val, pdf.D.F[[6]]$prob, type = "l", lty = 6,
      col = "black")
legend("topright", cex = 0.6,legend = c("D=0","D=8","D=12","D=20"),
      col = c("darkblue","darkred","darkgreen","black"),lty=c(1,2,4,6),
      box.lty=0,inset = 0.005)

```

### A.11.2.5 Représentation 2

Influence de la franchise sur les cotisations.

```

par(mfrow = c(1,2) , family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
### Hommes fonction de delta
plot(CMRS.DH[[1]]$S,CMRS.DH[[1]]$EXP,
     xlim=c(0,max(CMRS.DH[[6]]$S,0)),
     ylim=c(0,max(CMRS.DH[[1]]$EXP,0)),
     type="l",xlab="s",main = "Homme",
     ylab=list(TeX('$E[X_i|S=s]$')),col="darkblue")
lines(CMRS.DH[[2]]$S,CMRS.DH[[2]]$EXP,type="l",col="darkviolet")
lines(CMRS.DH[[3]]$S,CMRS.DH[[3]]$EXP,type="l",col="olivedrab")
lines(CMRS.DH[[4]]$S,CMRS.DH[[4]]$EXP,type="l",col="red")
lines(CMRS.DH[[5]]$S,CMRS.DH[[5]]$EXP,type="l",col="black")
lines(CMRS.DH[[6]]$S,CMRS.DH[[6]]$EXP,type="l",col="yellow2")
legend("topleft",inset = 0.02,cex = 0.6,legend = as.character(delta),
      lty= rep(1,7),col=c("darkblue","darkviolet","olivedrab","red",
      "black","yellow2"),box.lty=0)
### Femmes fonction de delta
plot(CMRS.DF[[1]]$S,CMRS.DF[[1]]$EXP,
     xlim=c(0,max(CMRS.DF[[6]]$S,0)),
     ylim=c(0,max(CMRS.DF[[1]]$EXP,0)),
     type="l",xlab="s",main = "Femme",

```

```

      ylab=list(TeX('$E[X_i|S=s]$')),col="darkblue")
lines(CMRS.DF[[2]]$S,CMRS.DF[[2]]$EXP,type="l",col="darkviolet")
lines(CMRS.DF[[3]]$S,CMRS.DF[[3]]$EXP,type="l",col="olivedrab")
lines(CMRS.DF[[4]]$S,CMRS.DF[[4]]$EXP,type="l",col="red")
lines(CMRS.DF[[5]]$S,CMRS.DF[[5]]$EXP,type="l",col="black")
lines(CMRS.DF[[6]]$S,CMRS.DF[[6]]$EXP,type="l",col="yellow2")
legend("topleft",inset = 0.02,cex = 0.6,legend = as.character(delta),
      lty= rep(1,7),col=c("darkblue","darkviolet","olivedrab","red",
      "black","yellow2"),box.lty=0)

```

### A.11.3 A la proportion Homme/Femme

Dans cette partie nous étudions l'effet de la composition du groupe sur le niveau des cotisations.

```

timer<-Sys.time()
n<-1500
# franchise annuelle commune à tous les participants
delta_i<-feat("Lambda","M",agetest)*feat("MeanSev","M",agetest)

# proportion d'hommes
p = seq(0.05,0.95,0.05)

Fpart.S.P<-list()
Fpart.SXH.P<-list()
Fpart.SXF.P<-list()
Quant_S.P<-rep(NA,length(p))

timer<-Sys.time()
for (i in 1:length(p)) {

  U <- seq(0,4000,1)
  ### Fonctions de répartition
  fc<-LX(U, delta = delta_i, delta.I = delta_i,"M", n, p[i])
  ## S
  Fc <- aggregateDist("recursive", model.freq = "poisson",
    model.sev = diff(fc$LX), lambda = fc$Lda, maxit = 10000)
  Fc <- DiscreteDistribution(supp = knots(Fc), prob = diff(Fc))
  ## S-Xh
  Fch<- aggregateDist("recursive", model.freq = "poisson",
    model.sev = diff(fc$LXh), lambda = fc$LdaH, maxit = 10000)
  Fch <- DiscreteDistribution(supp = knots(Fch), prob = diff(Fch))
  ## S-Xf
  Fcf<- aggregateDist("recursive", model.freq = "poisson",
    model.sev = diff(fc$LXf), lambda = fc$LdaF, maxit = 10000)
  Fcf <- DiscreteDistribution(supp = knots(Fcf), prob = diff(Fcf))

  Quant_S.P[i]<-min(q(Fc)(0.9999),q(Fch)(0.9999),q(Fcf)(0.9999))
  Fpart.S.P[[length(Fpart.S.P)+1]]<-Fc
  Fpart.SXH.P[[length(Fpart.SXH.P)+1]]<-Fch
  Fpart.SXF.P[[length(Fpart.SXF.P)+1]]<-Fcf
  rm(fc,Fc,Fch,Fcf)
}
print(Sys.time()-timer)
names(Fpart.S.P)<-paste("P", as.character(p), sep = "")
names(Fpart.SXH.P)<-paste("P", as.character(p), sep = "")
names(Fpart.SXF.P)<-paste("P", as.character(p), sep = "")

```

### A.11.3.1 Influence de la proportion Homme/Femme sur la distribution des cotisations

Dans cette partie nous étudions l'effet de la composition du groupe sur le niveau des cotisations.

```

timer<-Sys.time()
n<-1500
# franchise annuelle commune à tous les participants
delta_i<-feat("Lambda","M",agetest)*feat("MeanSev","M",agetest)

# proportion d'hommes
p = seq(0.05,0.95,0.05)

# nombre de valeurs de s à générer
long <- 15000

Law.P <- list()

timer <- Sys.time()
for (i in 1:length(p)) {
  S <- r(Fpart.S.P[[i]])(long)

  ESP <- data.frame(s=S, Exp.H=rep(NA,long), Exp.F=rep(NA,long),
                    Prob = p(Fpart.S.P[[i]])( S))

  for (j in 1:length(S)) {
    #conditional mean risk sharing
    A <- CondExp(s=S[j], delt=delta_i, Fpart.S.P[[i]],
                Fpart.SXH.P[[i]], Fpart.SXF.P[[i]])
    ESP$Exp.H[j] <- A$EH
    ESP$Exp.F[j] <- A$EF
  }

  Law.P[[length(Law.P)+1]] <- ESP

  rm(S,ESP,A)
}
print(Sys.time() - timer)
names(Law.P) <- paste("P", as.character(p), sep = "")
save(Law.P,file="CMRS_LawP.Rdata")

```

### A.11.3.2 Représentation 1 : Loi de probabilité en fonction de $p$

Ici nous représentons la fonction densité de probabilité de  $h_{i,n}^*(S_n)$  en fonction des différentes valeurs de  $p$ .

```

timer <- Sys.time()
pdf.P.H <- list()
pdf.P.F <- list()
for (i in 1:length(p)) {
  den.H <- density(Law.P[[i]]$Exp.H)
  den.F <- density(Law.P[[i]]$Exp.F)
  A <- data.frame(val=den.H$x, prob=den.H$y)
  B <- data.frame(val=den.F$x, prob=den.F$y)
  pdf.P.H[[length(pdf.P.H)+1]] <- A
  pdf.P.F[[length(pdf.P.F)+1]] <- B
  rm(den.H, den.F, A, B)
}

```

```

}
print(Sys.time()-timer)
names(pdf.P.H ) <- paste("P", as.character(p), sep = "")
names(pdf.P.F ) <- paste("P", as.character(p), sep = "")

par(mfrow=c(1,2),font.main=1,cex.main = 0.9, cex.sub = 0.8)

plot(pdf.P.H[[1]]$val, pdf.P.H[[1]]$prob, type = "l", lty = 1,
     col = "darkblue", ylim = c(0,0.9), xlim = c(2,6),
     xlab = list(TeX('$E[X_i|S_n]$')), ylab = "densité",main = "Homme")
lines(pdf.P.H[[6]]$val, pdf.P.H[[6]]$prob, type = "l", lty = 2,
     col = "darkred")
lines(pdf.P.H[[10]]$val, pdf.P.H[[10]]$prob, type = "l", lty = 4,
     col = "darkgreen")
lines(pdf.P.H[[19]]$val, pdf.P.H[[19]]$prob, type = "l", lty = 6,
     col = "black")
legend("topright", cex = 0.6,legend = c("P=0.05","P=0.3","P=0.5","P=0.95"),
     col = c("darkblue","darkred","darkgreen","black"),lty=c(1,2,4,6),
     box.lty=0,inset = 0.005)

plot(pdf.P.F[[1]]$val, pdf.P.F[[1]]$prob, type = "l", lty = 1,
     col = "darkblue", ylim = c(0,0.3), xlim = c(12,26),
     xlab = list(TeX('$E[X_i|S_n]$')), ylab = "densité",main = "Femme")
lines(pdf.P.F[[6]]$val, pdf.P.F[[6]]$prob, type = "l", lty = 2,
     col = "darkred")
lines(pdf.P.F[[10]]$val, pdf.P.F[[10]]$prob, type = "l", lty = 4,
     col = "darkgreen")
lines(pdf.P.F[[19]]$val, pdf.P.F[[19]]$prob, type = "l", lty = 6,
     col = "black")
legend("topright", cex = 0.6,legend = c("P=0.05","P=0.3","P=0.5","P=0.95"),
     col = c("darkblue","darkred","darkgreen","black"),lty=c(1,2,4,6),
     box.lty=0,inset = 0.005)

```

### A.11.3.3 Représentation 2 : Boxplot

Mise en évidence de la variabilité des cotisations autour de la prime pure.

```

P.draw <- c(1,3,6,8,10,15,19)
Box.pltH<-list()
Box.pltF<-list()
for (i in P.draw) {
  Box.pltH[[length(Box.pltH)+1]]<-Law.P[[i]]$Exp.H
  Box.pltF[[length(Box.pltF)+1]]<-Law.P[[i]]$Exp.F
}
names(Box.pltH)<-as.character(p[P.draw])
names(Box.pltF)<-as.character(p[P.draw])
Box.pltH.long<-melt(do.call(rbind,Box.pltH))
Box.pltF.long<-melt(do.call(rbind,Box.pltF))

par(font.main=1,cex.main = 0.9, cex.sub = 0.8)
boxplot(value~Var1,Box.pltH.long,xlab = "p",ylab = "", main = "Homme")
abline(h=E(FX.H),lwd = 1,lty=4,col="red")

boxplot(value~Var1,Box.pltF.long,xlab = "p",ylab = "", main = "Femme")
abline(h=E(FX.F),lwd = 1,lty=4,col="red")

```

# Annexe B

## Code R Partie 2

### B.1 Coût total individuel $Z_i$

#### B.1.1 Calcul des fonctions $F_Z$ pour les hommes

Le calcul suivant, permet grâce à la fonction FZ précédente de construire la fonction de répartition de tous les hommes à tous les âges considérés.

```
FZ.H<-list()
for (Age in Param_M$Age) {
  #A <- FZ(Age, "M")
  FZ.H[[length(FZ.H)+1]] <- FZ(Age, "M")
}
names(FZ.H)<-paste("H", as.character(Param_M$Age), sep = "")
```

#### B.1.2 Calcul des fonctions $F_Z$ pour les femmes

Cette fonction fait pareil que la précédente, mais pour le cas des femmes cette fois-ci.

```
FZ.F<-list()
for (Age in Param_F$Age) {
  #A <- FZ(Age, "F")
  FZ.F[[length(FZ.F)+1]] <- FZ(Age, "F")
}
names(FZ.F) <- paste("F", as.character(Param_F$Age), sep = "")
## Stockage de toutes les fonctions de répartitions dans une liste pour être
## utilisées plus tard.
F.all<-c(FZ.H,FZ.F)
```

### B.2 Distribution de $X_i$

Conversion de la fonction FX en une distribution discrète par la fonction DiscreteDistribution.

```
supp <- seq(0,4000,1)

# Pour les hommes
FX.H <- list()
for (i in 1:length(Param_M$Age)) {
```

```

age <- Param_M$Age[i]
probM <- FX(supp,age,Gender = "M",delta_i) # calcul de la fonction de répartition
pM0 <- FX(0,age,Gender = "M",delta_i) # masse de proba en 0
Cplt.M <- data.frame(supp = supp, prob = c(pM0,diff(probM)))

## Conversion de la distribution de X en une distribution discrète
FX.H[[length(FX.H)+1]] <- DiscreteDistribution(supp = Cplt.M$supp, prob = Cplt.M$prob)

rm(age, probM, pM0,Cplt.M)
}
names(FX.H)<-paste("H", as.character(Param_M$Age), sep = "")

# Pour les hommes
FX.F <- list()
for (i in 1:length(Param_F$Age)) {
  age <- Param_F$Age[i]
  probF <- FX(supp,age,"F",delta_i) # calcul de la fonction de répartition
  pF0 <- FX(0,age,"F",delta_i) # masse de proba en 0
  Cplt.F <- data.frame(supp = supp, prob = c(pF0,diff(probF)))

## Conversion de la distribution de X en une distribution discrète
FX.F[[length(FX.F)+1]] <- DiscreteDistribution(supp = Cplt.F$supp, prob = Cplt.F$prob)

rm(age, probF, pF0,Cplt.F)
}
names(FX.F)<-paste("F", as.character(Param_F$Age), sep = "")

FX.all<-c(FX.H,FX.F)

```

### B.3 Distribution de $X_i|X_i > 0$

Calcul de la fonction de répartition  $B_i$  de l'équation 2.3.11 pour tous les groupes d'hommes et de femmes dans le pool.

```

BX2<-function(x, Age, Gender, Delta){
  G <- ifelse(Gender=="M", "H", "F")
  nom <- paste(G,as.character(Age),sep = "")
  a <- p(F.all[[nom]])(x+Delta)
  b <- p(F.all[[nom]])(Delta)
  res <- (a-b)/(1-b)
  return(res)
}

```

### B.4 Création d'un groupe

La fonction Grp() permet de représenter dans un dataframe la composition du portefeuille par âge et par sexe. En entrée il faut avoir la composition du pool par âge et par sexe.

```

GRP<-function(Age=30,prop.H=0.49,prop.F=0.51,n=100){
  ## Fonction de création d'un groupe de participants
  ## Entrée:
  ## Age: une valeur unique ou vecteur contenant les âges observés.
  ## prop.H: la proportion d'hommes de chaque âge dans le groupe.
  ## prop.F: proportion de femme
  ## prop.H et prop.F ont la même longueur que Âge

```

```

## n: nombre de participants dans le groupe
## Sortie: retourne un dataframe contenant la proportion des individus
L <- length(Age)
NH<-prop.H*n
NH <- c(ceiling(NH[1:(L-1)]),sum(NH)-sum(ceiling(NH[1:(L-1)])))
NF<-prop.F*n
NF <- c(ceiling(NF[1:(L-1)]),sum(NF)-sum(ceiling(NF[1:(L-1)])))
N<-c(NH,NF)
Hoe<-data.frame(Age=Age,Gender="M",Ni=NH)
Fem<-data.frame(Age=Age,Gender="F",Ni=NF)
Grp<-rbind(Hoe,Fem)
return(Grp)
}

test<-GRP(Age=Param_M$Age,prop.H=Param_M$Prop,prop.F=Param_F$Prop,n=1500)
#print(test)

```

## B.5 Modèle collectif

### B.5.1 Calcul des fonctions de répartition des sévérités nécessaires pour $S_n$ et de $S_n - X_i$ dans le modèle collectif

La fonction LX2() est une version adaptée de la fonction LX() qui permet de prendre en compte tous les groupes du tableau 3.6.1. Pour chaque  $X_i$  appartenant à l'un de ces groupes, on calcule la fonction de répartition de  $S_n - X_i$  donnée à l'équation (2.3.16).

```

LX2 <- function(x, Grp, delta){
  ## Permet pour un groupe donné, de:
  ## - Construire la fonction de répartition des sévérités pour le calcul de la
  ##   fonction de répartition de S dans le modèle collectif--> LX;
  ## - Construire la fonction de répartition des sévérités pour le calcul de la
  ##   fonction de répartition de S-X dans le modèle collectif--> LLX;
  ## Entré:
  ## x: le point au quel est calculé la fonction de répartition.
  ## Grp: composition du porte-feuille
  ## delta: franchise applicable à tous les participants
  tmp <- Grp
  tmp$delta <- delta

  # Calcul des qi pour chaque groupe
  for (i in 1:nrow(tmp)) {
    G <- ifelse(tmp$Gender[i]=="M","H","F")
    nom <- paste(G,as.character(tmp$Age[i]),sep = "")
    tmp$qi[i] <- 1-p(F.all[[nom]])(tmp$delta[i])
  }
  long <- nrow(tmp)
  # Calcul de lambda totale
  Lda.dot<-sum(tmp$Ni*tmp$qi)

  # Calcul de la fonction de répartition LX (voir modèle collectif )
  val1 <- 0
  for (i in 1:nrow(tmp)) {
    val1 <- val1 + tmp$qi[i]*tmp$Ni[i]*BX2(x,tmp$Age[i],tmp$Gender[i],tmp$delta[i])
  }
  value1 <- val1/Lda.dot

```

```

# Calcul de la fonction de répartition de L'(x) pour le calcul de S-X_i
# Le calcul est fait en supposant successivement que X_i appartient aux
# différents groupes
Lda.X <- list()
LLX <- list()
for (i in 1:nrow(tmp)) {
  lda <- Lda.dot-tmp$qi[i]
  Lda.X[[length(Lda.X)+1]] <- lda
  val2 <- 0
  for (j in 1:nrow(tmp)) {
    t <- ifelse(i==j,1,0)
    val2 <- val2 + tmp$qi[j]*(tmp$Ni[j]-t)*BX2(x,
                                                tmp$Age[j],tmp$Gender[j],tmp$delta[j])
  }
  LLX[[length(LLX)+1]] <- val2/lda
}

names(Lda.X) <- paste("Grp", as.character(seq(1:long)), sep = "")
names(LLX) <- paste("Grp", as.character(seq(1:long)), sep = "")
row.names(tmp) <- paste("Grp", as.character(seq(1:long)), sep = "")
return(list(LX=value1,Lda=Lda.dot,Grp=tmp,LLX=LLX,Lda.X=Lda.X))
}

```

### B.5.2 Fonction de répartition de $S_n$ : Algorithme de Panjer

La fonction AGGRe() suivante permet grâce qu résultat de la fonction précédente de calculer les fonctions de répartition de  $S_n$ , de  $S_n - X_i$  pour un  $X_i$  appartenant à tous les sous-groupes.

```

AGGRe <- fonction(FS,maxit){
  S <- aggregateDist("recursive", model.freq = "poisson", model.sev = diff(FS$LLX),
                    lambda = FS$Lda, maxit = maxit)
  S <- DiscreteDistribution(supp = knots(S), prob = diff(S))
  LLX <- FS[["LLX"]]
  LdaX <- FS[["Lda.X"]]
  SX <- list()
  for (i in 1:length(LLX)) {
    FSX <- aggregateDist("recursive", model.freq = "poisson",
                        model.sev =diff(LLX[[i]],lambda = LdaX[[i]], maxit = maxit)
    FSX <- DiscreteDistribution(supp = knots(FSX), prob = diff(FSX))
    SX[[length(SX)+1]]<-FSX
  }
  names(SX)<-paste("SX",names(FS$LLX), sep = ".")
  return(c(S = S, SX))
}

```

### B.5.3 Essais sur l'ensemble des participant

Dans cet exemple on considère que tous les participants ont la même franchise  $\delta_i = \delta = 4$ . Nous construisons les fonctions de répartitions de  $S_n$  et de  $S_n - X_i$  pour tous les participants avec la composition telles que définie plus haut au tableau 3.6.1.

```

timer<- Sys.time()
# Le nombre de participant dans le groupe
n=2000

#Création d'un groupe d'homme et de femme d'âge égal à 30 ans

```

```

grp<-GRP(Age = Param_M$Age, prop.H = Param_M$Prop,
         prop.F = Param_F$Prop, n=n)

u<-seq(0,4000,1)
## Construction des fonctions de répartition pour les sévérités
FS<-LX2(u, grp, delta=delta_i)
Grp.N<-FS$Grp
Fpart<-AGGre(FS,maxit = 100000)

## Fonction de répartition de S
Fs <-Fpart[["S"]]
print(Sys.time()-timer)

```

## B.6 Étude des Cotisations

### B.6.1 Sensibilité au nombre de participants $n$

Ce bloc permet de calculer l'espérance conditionnelle pour différentes valeurs du nombre de participants  $n$ .

```

timer<-Sys.time()
# nombre de participants
n<-c(100,500,1000,2000,3000,4000)

CMRS.N<-list()
func.repart.N<-list()
Quant_S.n<-rep(NA,length(n))
Group.comp<-list()
for (i in 1:length(n)) {
  #Création d'un groupe d'homme et de femme d'âge égal à 30 ans
  group <- GRP(Age = Param_M$Age, prop.H = Param_M$Prop,
              prop.F = Param_F$Prop, n=n[i])
  U <- seq(0,4000,1)
  FS <- LX2(U, group, delta=delta_i)
  Group.N <- FS$Grp
  Fpart.all <- AGGre(FS, maxit = 100000)

  Quant_S.n[i] <- q(Fpart.all[["S"]])(0.9999999)
  S <- seq(0,Quant_S.n[i],10)

  ESP <- data.frame (S=S ,S.Prob = d(Fpart.all[["S"]])(S))
  A <- data.frame()
  for (j in 1:length(S)) {
    B <- CondExp(s = S[j] , Group.N, Fpart.all)
    A <- rbind(A,as.data.frame(B)[1,1:length(B)])
  }
  ESP<-cbind(ESP,A)
  CMRS.N[[length(CMRS.N)+1]]<-ESP
  func.repart.N[[length(func.repart.N)+1]]<-Fpart.all
  Group.comp[[length(Group.comp)+1]] <- Group.N
  rm(group,U,FS,Group.N,Fpart.all,S,ESP,A,B)
}
names(CMRS.N)<-paste("N", as.character(n), sep = "")
names(func.repart.N)<-paste("N", as.character(n), sep = "")
names(Group.comp)<-paste("N", as.character(n), sep = "")

```

```
save(CMRS.N,file="CMRS_N.Rdata")
print(Sys.time()-timer)
```

### B.6.1.1 Représentation 1 : Homme Vs Femme

Représentation des fonction  $h_{i,n}^*(s)$  pour les différents groupes du pool.

```
n<-c(100,500,1000,2000,3000,4000)
### Hommes vs femme n=1500
tr=which(n==1000)
plf <- Quant_S.n[tr]
u <- seq(0, plf,1)
toplot <- CMRS.N[[tr]]

par(mfrow=c(1,2),family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)

matplot(toplot$$,toplot[3:9],type = "l", lty=1:7, col = 1:7,
        xlabel="S",ylab=list(TeX('$E[X_i|S_n=s]$')), main = "Homme")
lines(u,d(func.repart.N[[tr]][["S"]])(u)*(6.0e5),type="l",lty=8,col=8)
legend("topleft",inset = 0.02,cex = 0.6,legend =c(colnames(toplot)[3:9],
    TeX('$P[S_n=s]$')),col=1:8,lty=1:8,box.lty=0)

matplot(toplot$$,toplot[10:16],type = "l", lty=1:7, col = 1:7,
        xlabel="S",ylab=list(TeX('$E[X_i|S_n=s]$')), main = "Femmes")
lines(u,d(func.repart.N[[tr]][["S"]])(u)*(5.5e5),type="l",lty=8,col=8)
legend("topleft",inset = 0.02,cex = 0.6,legend =c(colnames(toplot)[10:16],
    TeX('$P[S_n=s]$')),col=1:8,lty=1:8,box.lty=0)
```

Vérification de la propriété d'allocation totale de la conditional mean risk sharing.

```
### Vérification somme des cotisations
Sum.Cotis <- CMRS.N[["N1000"]]
#Sum.Cotis$Sum <- colSums(Sum.Cotis[,3:16]*Group.comp[["N1000"]][1:14,3])
Sum.Cotis$sum <- 0
for (i in 1:nrow(Sum.Cotis)) {
  Sum.Cotis$sum[i] <- sum(Sum.Cotis[i,3:16]*Group.comp[["N1000"]][1:14,3])
}
#Sum.Cotis <- subset(Cotisations, S %% 100 == 0)
up <- distr::q(func.repart.N[["N1000"]])$S(10^(-3))
dow <- distr::q(func.repart.N[["N1000"]])$S(1-10^(-3))
plot(Sum.Cotis$$,Sum.Cotis$sum, type = "l", lty=1, col = "darkblue",
    xlabel = "s", ylab = "", main = "Somme des cotisations",
    xlim=c(up,dow))
abline(0,1,lty=6, col = gray(0.5))
legend("topleft", inset = 0.02,cex = 0.6,legend =c("CMRS","y=x"),
    col=c("darkblue",gray(0.5)),lty=c(1,6),box.lty=0)

rm(up,dow)
```

### B.6.1.2 Loi de probabilité des $\mathbb{E}[X_i|S_n]$

Ce block permet pour chaque valeur de  $n$ , de générer un vecteur aléatoire  $S$  de longueur  $\backslash\text{verb}\backslash\text{long}$  et contenant les valeurs des réalisations  $s$  de  $S_n$ . chaque valeur  $s$  est générée à partir de la fonction de répartition correspondante à la valeur de  $n$ . Le but étant de trouver les fonction densités de probabilités des cotisations.

### B.6.1.3 Représentation 2 : Variabilité des cotisations en l'absence de (re-)assurance

Mise en évidence de la variabilité des cotisations autour de la prime pure.

```
#nombre de participants
n<-c(100,500,1000,2000,3000,4000)
Box.plot <- list()
for (j in 2:ncol(Law.N[[1]])) {
  Bx.plt<-list()
  for (i in 1:length(Law.N)) {
    Bx.plt[[length(Bx.plt)+1]]<-Law.N[[i]][[j]]
  }
  names(Bx.plt)<-as.character(n)
  Bx.plt.long<-melt(do.call(rbind,Bx.plt))
  Box.plot[[length(Box.plot)+1]] <- Bx.plt.long
}
names(Box.plot) <- colnames(Law.N[[1]][2:ncol(Law.N[[1]])])

par(mfrow=c(1,3), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
for (i in 1:length(Box.plot)) {
  boxplot(value~Var1,Box.plot[[i]],xlab="",ylab="",
          main = names(Box.plot)[i],
          outline = FALSE)
  abline(h=E(FX.all[[i]]),lwd = 1,ltty=4,col="red")
}
```

## B.7 Application 1 : Décomposition pour un groupe de $n = 1000$

Dans cette première illustration, nous faisons l'application de la décomposition en tranche pour un groupe de  $n = 1000$  participants.

```
timer <- Sys.time()
## Vecteur des probabilités bêta pour octroyer des bonus en fin de période
beta <- 0.85

## Chargement de sécurité pour la stop-loss
theta.SL <- 0.5

#Création d'un groupe de n=1000 avec tous les classes d'âges
Grp.layer <- Group.comp$N1000

## Fonction de répartition de Sn n=1000
Fpart <- func.repart.N$N1000

## retentions
W.group <- distr::q(Fpart[["S"]])(beta) #= 27344
Wi <- CondExp(s=W.group,group = Grp.layer, Fpart) ## rétention individuelles
Grp.layer$Wi <- t(as.data.frame(Wi))[1:length(Wi)]

## Valeurs E[Xi|S=s] simulées précédemment
CMRS.layer <- Law.N[["N1000"]]

print(Sys.time()-timer)
```

```

timer <- Sys.time()
## Vecteur contenant les valeurs de min{S,W}
Charge.P2P <- CMRS.layer
Charge.P2P[1:nrow(CMRS.layer),2:ncol(CMRS.layer)] <- 0

## Vecteur contenant les valeurs de la prime SL
PI.SL <- CMRS.layer
PI.SL[1:nrow(CMRS.layer),2:ncol(CMRS.layer)] <- 0

## Vecteur contenant les valeurs du cash-back
BI.n <- CMRS.layer
BI.n[1:nrow(CMRS.layer),2:ncol(CMRS.layer)] <- 0

Grp.layer$EXi <- rep(NA,nrow(Grp.layer))
for (i in 2:ncol(CMRS.layer)) {
  #E[Xi]
  Grp.layer$EXi[i-1] <- E(FX.all[[i-1]])

  #S_P2P=min{h,wi}
  Charge.P2P[,i] <- ifelse(
    test = CMRS.layer[,i] < Grp.layer$Wi[i-1],
    yes = CMRS.layer[,i],
    no = Grp.layer$Wi[i-1])

  # Calcul de la prime stop-loss
  PI.SL[,i] <- ifelse(
    test = CMRS.layer[,i] - Grp.layer$Wi[i-1] > 0,
    yes = CMRS.layer[,i] - Grp.layer$Wi[i-1],
    no = 0
  )
  # Cash-back
  BI.n[,i] <- ifelse(
    test = Grp.layer$Wi[i-1]-CMRS.layer[,i] > 0,
    yes = Grp.layer$Wi[i-1]-CMRS.layer[,i],
    no = 0
  )
}
print(Sys.time()-timer)
Grp.layer$prime.SL <- (1+theta.SL)*colMeans(PI.SL[,-1],na.rm = TRUE)
Grp.layer$Prime <- Grp.layer$prime.SL + Grp.layer$Wi
Grp.layer$Ex.Bi <- colMeans(BI.n[,-1],na.rm = TRUE)
Grp.table <- Grp.layer[,-c(4,5)]

## génère le script latex pour le tableau Grp.table
#print( xtable(Grp.table,type="latex",digits = 3))

```

### B.7.1 Loi du Cash-Back $B_n$

Un résultat similaire à 2.3.10 et à 2.3.11 peut être obtenue pour la loi de  $B_n$  et celle de  $B_n|B_n > 0$ . Le block suivant permet à partir de la loi de  $S_n$  de déterminer celle du cash-back.

```

# fonction de répartition du cash-back pour l'ensemble du groupe
F_Bi <- function(x,FS,omega){
  func <-( p(FS)(omega)-p(FS)(omega-x) )/p(FS)(omega)
  return(func)
}

```

```

supp=seq(0,W.group,1)
prob=diff(F_Bi(supp,Fpart[["S"]],W.group))
cash.back <- DiscreteDistribution(supp=supp[-1],prob=prob)
rm(supp,prob)

plf<-q(cash.back)(0.999)
u<-seq(1,plf,1)
v<-seq(0,1,1/plf)
### Représentation de Fs
par(mfrow=c(1,3),font.main=1,cex.main = 0.9, cex.sub = 0.8)

plot(u,d(cash.back)(u),type="l",xlab="x",ylab="d(x)", main="Density")

plot(u,p(cash.back)(u),type="l",xlab="q",ylab="p(q)",main="CDF")

plot(v,q(cash.back)(v),type="l",xlab="p",ylab="q(p)",main="Quantile")

```

### B.7.2 Représentation 3 : Loi de probabilité du cash-back $B_{i,n}$ .

Représentation de la loi de probabilité du cash-back  $B_{i,n}$  pour tous les groupes de participants.

```

BI.n <- na.omit(BI.n)
pdf.Cash <- list()
for (i in colnames(BI.n[-1])) {
  den.Bi <- density(BI.n[,i])
  A <- data.frame(val=den.Bi$x, prob=den.Bi$y)
  pdf.Cash[[length(pdf.Cash)+1]] <- A
  rm(den.Bi, A)
}
par(mfrow=c(1,2),family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
plot(pdf.Cash[[1]]$val, pdf.Cash[[1]]$prob, type = "l", lty = 1,
     col = 1, ylim = c(0,1.8), xlim = c(0,30),
     xlab = "Cash-Back", ylab = "densité", main = "Hommes")
for (i in 2:7) {
  lines(pdf.Cash[[i]]$val, pdf.Cash[[i]]$prob, type = "l", lty = i,
        col = i)
}
legend("topright", cex = 0.6,legend =rownames(Grp.layer)[1:7],col = 1:7,lty=1:7,
      box.lty=0,inset = 0.005)

plot(pdf.Cash[[8]]$val, pdf.Cash[[8]]$prob, type = "l", lty = 1,
     col = 1, ylim = c(0,1.1), xlim = c(0,30),
     xlab = "Cash-Back", ylab = "densité", main = "Femmes")
for (i in 9:length(pdf.Cash)) {
  lines(pdf.Cash[[i]]$val, pdf.Cash[[i]]$prob, type = "l", lty = (i-7),
        col = (i-7))
}
legend("topright", cex = 0.6,legend =rownames(Grp.layer)[8:14],col=1:8,lty= 1:8,
      box.lty=0,inset = 0.005)

```

## B.8 Variabilité des cotisations dans tranche inférieure $[0, \omega_n]$

### B.8.1 Comparaison avec une variation du nombre de participant

Etude de la variabilité dans la tranche inférieure en fonction du nombre de participants.

```

timer <- Sys.time()
#nombre de participants.
n<-c(100,500,1000,2000,3000,4000)

## Vecteur des probabilités beta pour octroyer des bonus en fin de période
beta <- 0.60

## Chargement de sécurité pour la stop-loss
theta.SL <- 0.5

down.layer.N <- list()
up.layer.N <- list()
SL.prem <- data.frame(Grp=seq(1,14,1))
for (i in 1:length(n)) {
  #Création d'un groupe avec tous les classes d'âges
  Grp.Layer <- Group.comp[[i]]

  ## Fonction de répartition de Sn n=1500
  Fpart <- func.repart.N[[i]]

  ## rétention du groupe
  W.group <- distr::q(Fpart[["S"]])(beta)

  ## rétention individuelles Wi
  Wi <- CondExp(s=W.group,group = Grp.Layer, Fpart)
  Grp.Layer$Wi <- t(as.data.frame(Wi))[1:length(Wi)]

  Group.comp[[i]]$Wi <- Grp.Layer$Wi
  ## Récupération des valeurs de E[Xi|S] pour les différentes valeurs de n
  CMRS.Layer <- Law.N[[i]]

  ## Vecteur contenant les valeurs de min{E[Xi|S],Wi}
  S.P2P <- CMRS.Layer
  S.P2P[1:nrow(CMRS.Layer),2:ncol(CMRS.Layer)] <- 0

  ## Vecteur contenant les valeurs de min{E[Xi|S],Wi}
  S.UP <- CMRS.Layer
  S.UP[1:nrow(CMRS.Layer),2:ncol(CMRS.Layer)] <- 0
  prem <- data.frame(SL=rep(NA,nrow(Grp.Layer)))
  for (i in 2:ncol(CMRS.Layer)) {
    for (j in 1:nrow(CMRS.Layer)) {
      #tranche inférieure: S.P2P=min{h,wi}
      S.P2P[j,i] <- ifelse(
        test = CMRS.Layer[j,i] < Grp.Layer$Wi[i-1],
        yes = CMRS.Layer[j,i],
        no = Grp.Layer$Wi[i-1])

      # tranche supérieure S.UP=S-min{S,wm}
      S.UP[j,i] <- ifelse(
        test = CMRS.Layer[j,i] - Grp.Layer$Wi[i-1] > 0,

```

```

    yes = CMRS.Layer[j,i] - Grp.Layer $Wi[i-1],
    no = 0)
  }
  prem[i-1,1] <- (1+theta.SL)*mean(S.UP[,i])
}
SL.prem <- cbind(SL.prem,pre,Grp.Layer$Wi)
down.layer.N[[length(down.layer.N)+1]] <- S.P2P
up.layer.N[[length(up.layer.N)+1]] <- S.UP
rm(Grp.Layer,W.group,Wi,S.P2P,S.UP)
}
colnames(SL.prem)[1] <- "Grp"
names(down.layer.N)<-paste("N", as.character(n), sep = "")
names(up.layer.N)<-paste("N", as.character(n), sep = "")
print(Sys.time()-timer)

```

### B.8.2 Représentation 4 : Variabilité des cotisations dans la tranche inférieure

Mise en évidence de la variabilité des cotisations autours de la prime pure dans la tranche inférieure.

```

#nombre de participants
n<-c(100,500,1000,2000,3000,4000)
Layer.plot <- list()
for (j in 2:ncol(down.layer.N[[3]])) {
  Bx_plt<-list()
  for (i in 1:length(down.layer.N)) {
    tmp <- down.layer.N[[i]][[j]]
    tmp <- subset(tmp,tmp < Group.comp[[i]]$Wi[j-1])
    Bx_plt[[length(Bx_plt)+1]]<-tmp
  }
  names(Bx_plt)<-as.character(n)
  Bx_plt_long<-melt(do.call(rbind,Bx_plt))
  Layer.plot[[length(Layer.plot)+1]] <- Bx_plt_long
}
rm(Bx_plt_long,Bx_plt)
names(Layer.plot) <- colnames(down.layer.N[[1]])[2:ncol(down.layer.N[[1]])]

par(mfrow=c(1,3), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
    cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
for (i in 1:length(Layer.plot)) {
  boxplot(value~Var1,Layer.plot[[i]],xlab="",ylab="",
    main = names(Layer.plot)[i],
    outline = FALSE)
  abline(h=E(FX.all[[i]]),lwd = 1,ltty=4,col="red")
}

```

### B.8.3 Représentation 5 : Variabilité des cotisations dans la tranche supérieure

Variabilité des cotisations dans la tranche supérieure.

```

#nombre de participants
n<-c(100,500,1000,2000,3000,4000)
Layer.Up <- list()
for (j in 2:ncol(up.layer.N[[3]])) {
  Bx_plt<-list()

```

```

for (i in 1:length(up.layer.N)) {
  tmp <- up.layer.N[[i]][[j]]
  tmp <- subset(tmp,tmp > 0)
  Bx_plt[[length(Bx_plt)+1]]<-tmp
  rm(tmp)
}
names(Bx_plt)<-as.character(n)
Bx_plt_long<-melt(do.call(rbind,Bx_plt))
Layer.Up[[length(Layer.Up)+1]] <- Bx_plt_long
}
rm(Bx_plt_long,Bx_plt)
names(Layer.Up) <- colnames(up.layer.N[[1]])[2:ncol(up.layer.N[[1]])]

par(mfrow=c(1,3), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
     cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
for (i in 1:length(Layer.Up)) {
  boxplot(value~Var1,Layer.Up[[i]],xlab="",ylab="",
          main = names(Layer.Up)[i],
          outline = FALSE)
  #abline(h=E(FX.all[[i]]),lwd = 1,lty=4,col="red")
}

```

#### B.8.4 Représentation 5 : Box plot rétention + nombre de participants

variabilité des cotisations en fonction du nombre de participants pour une valeur de  $\beta$  fixée.

```

n<-c(100,500,1000,2000,3000,4000)
Box.Vs <- list()
for (j in 2:ncol(down.layer.N[[1]])) {
  Bx.vs<-list()
  for (i in 1:length(down.layer.N)) {
    tmp <- down.layer.N[[i]][[j]]
    tmp <- subset(tmp,tmp < Group.comp[[i]]$Wi[j-1])
    tmp <- as.data.frame(tmp)
    tmp$Hin <- Law.N[[i]][[j]][1:nrow(tmp)]
    tmp$N <- n[i]
    colnames(tmp) <- c("SP2P","Hin","N")
    Bx.vs[[length(Bx.vs)+1]]<-tmp
  }
  names(Bx.vs)<-as.character(n)
  Bx.vs_long<-melt(do.call(rbind,Bx.vs),id="N")
  Bx.vs_long$N <- as.factor(Bx.vs_long$N)
  Box.Vs[[length(Box.Vs)+1]] <- Bx.vs_long
}
names(Box.Vs) <- colnames(down.layer.N[[1]])[2:ncol(down.layer.N[[1]])]

par(mfrow=c(1,3), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
     cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
for (i in 1:length(Layer.plot)) {
  boxplot(value~variable+N,Box.Vs[[i]],xlab="N",ylab="",
          at = c(1:2, 4:5, 7:8, 10:11, 13:14, 16:17),
          main = names(Layer.plot)[i], col = c("lightblue","gray"),
          names = c("100","","500","","","2000","","","4000",""),
          outline = FALSE)
  abline(h=E(FX.all[[i]]),lwd = 1,lty=4,col="red")
  legend("topright",box.lty=0,inset = 0.02,

```

```

    fill = c("lightblue","gray"),
    legend = c(TeX('$^{P2P}_n$'),TeX('$_n$')), horiz = T)
}

```

### B.8.5 Variabilité par rapport au niveau de rétention $\omega$

Etude de la variabilité dans la tranche inférieure en fonction du nombre de participants.

```

timer <- Sys.time()
#nombre de participants.
n <- 1000

## Vecteur des probabilités beta pour octroyer des bonus en fin de période
beta.vec <- seq(0.5,0.95,0.1)

## Chargement de sécurité pour la stop-loss
theta.SL <- 0.5

down.layer.B <- list()
Up.layer.B <- list()
Grp.comp.B <- list()
for (i in 1:length(beta.vec)) {
  #Création d'un groupe avec tous les classes d'âges
  Grp.Layer <- Group.comp$N1000

  ## Fonction de répartition de Sn n=1500
  Fpart <- func.repart.N$N1000

  ## rétention du groupe
  W.group <- distr::q(Fpart[["S"]])(beta.vec[i])

  ## rétention individuelles Wi
  Wi <- CondExp(s=W.group,group = Grp.Layer, Fpart)
  Grp.Layer$Wi <- t(as.data.frame(Wi))[1:length(Wi)]
  Grp.comp.B[[length(Grp.comp.B)+1]] <- Grp.Layer
  ## Récupération des valeurs de E[Xi|S] pour les différentes valeurs de n
  CMRS.Layer <- Law.N$N1000

  ## Vecteur contenant les valeurs de min{E[Xi|S],Wi}
  S.P2P <- CMRS.Layer
  S.P2P[1:nrow(CMRS.Layer),2:ncol(CMRS.Layer)] <- 0
  S.UP <- CMRS.layer
  S.UP[1:nrow(CMRS.Layer),2:ncol(CMRS.Layer)] <- 0
  for (i in 2:ncol(CMRS.Layer)) {
    for (j in 1:nrow(CMRS.Layer)) {
      #S.P2P=min{h,wi}
      S.P2P[j,i] <- ifelse(
        test = CMRS.Layer[j,i] < Grp.Layer$Wi[i-1],
        yes = CMRS.Layer[j,i],
        no = Grp.Layer$Wi[i-1])

      # tranche supérieure S.UP=S-min{S,wn}
      S.UP[j,i] <- ifelse(
        test = CMRS.Layer[j,i] - Grp.Layer$Wi[i-1] > 0,
        yes = CMRS.Layer[j,i] - Grp.Layer $Wi[i-1],
        no = 0)
    }
  }
}

```

```

}
}
down.layer.B[[length(down.layer.B)+1]] <- S.P2P
Up.layer.B[[length(Up.layer.B)+1]] <- S.UP
rm(Grp.Layer,W.group,Wi,S.P2P,S.UP)
}
names(down.layer.B)<-paste("B", as.character(beta.vec), sep = "")
names(Up.layer.B)<-paste("B", as.character(beta.vec), sep = "")
print(Sys.time()-timer)

```

### B.8.6 Représentation 6 : Effet de la variation de la valeur de la retention

Mise en évidence de la variabilité des cotisations autours de la prime pure.

```

Layer.plot.B <- list()
for (j in 2:ncol(down.layer.B[[1]])) {
  Bx_plt<-list()
  for (i in 1:length(down.layer.B)) {
    tmp <- down.layer.B[[i]][[j]]
    tmp <- subset(tmp,tmp < Grp.comp.B[[i]]$Wi[j-1])
    Bx_plt[[length(Bx_plt)+1]]<-tmp
    rm(tmp)
  }
  names(Bx_plt)<-as.character(beta.vec)
  Bx_plt_long<-melt(do.call(rbind,Bx_plt))
  Layer.plot.B[[length(Layer.plot.B)+1]] <- Bx_plt_long
}
rm(Bx_plt_long,Bx_plt)
names(Layer.plot.B) <- colnames(down.layer.B[[1]])[2:ncol(down.layer.B[[1]])]

par(mfrow=c(1,3), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
     cex.main=0.9, font.main=1,font.lab=1,font.axis=1)
for (i in 1:length(Layer.plot.B)) {
  boxplot(value~Var1,Layer.plot.B[[i]],xlab = "",ylab = "",
          main = names(Layer.plot.B)[i],
          outline = FALSE)
  #abline(h=E(FX.all[[i]]),lwd = 1,lty=4,col="red")
}

par(mfrow=c(1,2), family = "sans", font.lab = 2, font.axis = 1, cex.lab=0.9,
     cex.main=0.9, font.main=1,font.lab=1,font.axis=1)

boxplot(value~Var1,Layer.plot.B[[2]],xlab =TeX('$\\beta$'),ylab = "",
        main = "Homme 30 ans",
        outline = FALSE)
abline(h=E(FX.all[[2]]),lwd = 1,lty=4,col="red")
boxplot(value~Var1,Layer.plot.B[[9]],xlab =TeX('$\\beta$'),ylab = "",
        main = "Femme 30 ans",
        outline = FALSE)
abline(h=E(FX.all[[9]]),lwd = 1,lty=4,col="red")

```

UNIVERSITÉ CATHOLIQUE DE LOUVAIN  
Faculté des sciences

Place des sciences, 2 bte L6.06.01, 1348 Louvain-la-Neuve, Belgique | [www.uclouvain.be/sc](http://www.uclouvain.be/sc)