

Mémoire présenté devant le jury de l'EURIA en vue de l'obtention du
Diplôme d'Actuaire EURIA
et de l'admission à l'Institut des Actuaires

le 15 Décembre 2020

Par : Lariosse KOUAKOU

Titre : Réseaux et assurance : Optimisation d'un modèle de franchise collaborative dans le cadre d'un contrat d'assurance IARD

Confidentialité : Non

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

Membres présents du jury de l'Institut des Actuaires :

Dominique ABGRALL

Signature :

Florence PICARD

Signature :

Sonia GUELOU

Signature :

Hélène GIBELLO

Signature :

Membres présents du jury de l'EURIA :

Franck VERMET

Signature :

Entreprise :

Université du Québec à Montréal (UQAM)

Signature :

Directeur de mémoire en entreprise :

Arthur CHARPENTIER

Signature :

Autorisation de publication et de mise en ligne sur un site de diffusion de documents actuariels

(après expiration de l'éventuel délai de confidentialité)

Signature du responsable entreprise :

Signature du candidat :

Résumé

Les graphes sont de plus en plus utilisés dans diverses disciplines incluant la biologie, l'épidémiologie, les sciences sociales, l'économie ou la finance, pour la représentation, la compréhension et la modélisation des phénomènes d'interactions et de communautés. Ce mémoire d'actuariat se sert de modèles de graphe de la théorie des réseaux pour la valorisation de produits d'assurance IARD basés sur des franchises collectives, inspiré par le produit dit de « franchise collaborative » initié en 2014 par la startup française *Inspeer*. Ce produit consiste, en effet, pour un assuré à mutualiser ses franchises d'assurance dommage (automobile, habitation, moto) avec ses proches et ce, sans changer de contrat d'assurance. Ce système permet donc à l'assuré de : 1) diminuer le coût de son assurance classique en augmentant sa franchise d'assurance et 2) faire couvrir sa franchise d'assurance par la communauté des assurés. Cependant, étant donné que l'assuré s'engage à contribuer de manière équivalente à la franchise de ses relations, il porte une partie du risque en cas de sinistre d'un grand nombre de ces derniers. De ce fait, les couvertures et les contributions des assurés constituant les variables d'intérêt du produit de franchise collaborative, sont fortement dépendantes de la structure du réseau de la communauté des assurés.

L'objectif de ce mémoire consiste donc à analyser l'impact de la structure du réseau des assurés sur ces variables d'intérêt et à mettre en place efficacement des mécanismes permettant l'optimisation du produit, c'est-à-dire : garantir la couverture des charges (franchises et dépenses¹) annuelles d'un nombre maximum d'assurés tout en minimisant les contributions individuelles des assurés à la couverture des charges de leurs relations.

Pour ce faire, dans un premier temps, sur la base d'une communauté fictive d'assurés, des études d'optimisation ont été respectivement effectuées sur les paramètres du produit, à savoir le montant des engagements réciproques et le modèle de graphe social représentant le réseau des connexions entre les assurés de la communauté. Dans un second temps, il a été mis en place un produit de franchise collaborative, d'une part, sur la base du produit optimisé dit « scénario central » prenant en entrée les paramètres optimaux issus des études d'optimisation du produit et d'une étude de sensibilité testant ces paramètres selon différentes sensibilités, afin d'évaluer leurs impacts sur les variables d'intérêt du produit. Et d'autre part, sur la base d'applications à des données réelles, notamment des données de généalogie et de réseaux sociaux tels que *Facebook* afin d'analyser comment ces types spécifiques de connexions entre les assurés peuvent influer sur ces variables d'intérêt.

Mots clés :

Assurance collaborative, Franchise collaborative, Réseaux, Graphes, Couverture de charges, Contribution aux charges.

1. dépense = min(franchise, coût sinistres).

Abstract

Graphs are increasingly used in various disciplines such as biology, epidemiology, social sciences, economics or finance, for the representation, understanding and modeling of the interactions and communities phenomena. This actuarial thesis uses graph models of network theory for the valuation of P&C insurance products based on collective deductibles inspired by the product called « peer-to-peer deductible », initiated in 2014 by the French startup *Insppeer*. This product consist for an policyholder to mutualise his deductible insurance of damages (auto, home, bike) with his relatives and that without changing insurance contract. So, this system helps the policyholder firstly decrease the cost of his traditional insurance by increasing his deductible and secondly cover his deductible insurance by the community of policyholders. However, given that the policyholder is committed to contributing equivalently to his relations deductibles, he carries a part of risk whether a great deal of his relationship is claimed. Thus, the coverages and contributions of the policyholders that make the interest variables of the peer-to-peer deductible product, are strongly dependent on the structure of the policyholders network.

The objective of this thesis consists in analysing the impact of the policyholders network structure on those interest variables and set efficiently mechanisms that permit to optimize the product, that is to say to guarantee annual charges (deductibles and expenses²) coverage of a maximal number of policyholders as well as the minimization of the individual contributions of the policyholders to their relations charges.

To do this, firstly, based on a fictional community of policyholders, optimization studies have respectively been done on product's parameters such as the amount of the reciprocal contributions and the model of social graph representing the network of connections between the policyholders. Secondly, there has been set a peer-to-peer deductible product on the basis of the optimized product called « central scenario » taking as entry, the optimal parameters from the optimization studies of the product and a study of the sensitivity assessing those parameters according to different sensitivities in order to evaluate their impact on the interest variables, on the one hand. On the other hand, based on applications to real data namely genealogical data and social networks such as *Facebook* so as to analyse how those specific types of connections between the policyholders can impact on those interest variables.

Keywords :

Peer-to-peer insurance, Peer-to-peer deductible, Networks, Graphs, Coverage of charges, Contribution to charges.

2. expense = min(deductible, claims cost).

Note de synthèse

Contexte et problématique

Ce mémoire s'intéresse à la valorisation de produits d'assurance IARD basés sur des franchises collectives, inspiré par le produit dit de « franchise collaborative » initié en 2014 par la startup française *Inspeer*, consistant pour un assuré à mutualiser ses franchises d'assurance dommage (automobile, habitation, moto) avec ses proches et ce, sans changer de contrat d'assurance. Ce système permet donc à l'assuré de : 1) diminuer le coût de son assurance classique en augmentant sa franchise d'assurance et 2) faire couvrir sa franchise d'assurance par la communauté des assurés. Cependant, étant donné que l'assuré s'engage à contribuer de manière équivalente à la franchise de ses relations, il porte une partie du risque en cas de sinistre d'un grand nombre de ces derniers. De plus, la startup ne collectait les contributions (ou cotisations) des assurés qu'en cas de sinistre de leurs relations. Ce qui engendrait indéniablement un risque de contrepartie des assurés. Aussi, la startup a dû déposer le bilan, en 2018, faute de capitaux et de souscriptions suffisants pour financer son développement.

Par ailleurs, force est de constater que, les couvertures et les contributions des assurés constituant les variables d'intérêt du produit de franchise collaborative, sont fortement dépendantes de la structure du réseau de la communauté des assurés.

C'est dans ce contexte d'analyse de l'impact de la structure du réseau des assurés sur ces variables d'intérêt, que s'inscrit ce mémoire d'actuariat dont l'objectif principal consiste, avec un peu plus de recul depuis 2018, à mettre en place efficacement des mécanismes permettant l'optimisation du produit, c'est-à-dire : garantir la couverture des charges (franchises et dépenses³) annuelles d'un nombre maximum d'assurés tout en minimisant les contributions individuelles des assurés à la couverture des charges de leurs relations.

Méthodologie

Afin de mener à bien cet objectif d'optimisation du produit de franchise collaborative, la démarche suivante a été adoptée.

- ❖ Tout d'abord, une formalisation mathématique du produit de franchise collaborative a été effectuée et une définition de ses variables d'intérêt.
- ❖ Ensuite, sur la base d'une communauté fictive d'assurés, des études d'optimisation ont été respectivement effectuées sur les paramètres du produit de franchise collaborative, à savoir le montant des engagements réciproques et le modèle de graphe social représentant le réseau des connexions entre les assurés de la communauté.
- ❖ Enfin, un produit de franchise collaborative a été mis en place, d'une part, sur la base du produit optimisé dit « scénario central » prenant en entrée les paramètres optimaux issus des études d'optimisation du produit et d'une étude de sensibilité testant ces paramètres selon différentes sensibilités afin d'évaluer leurs impacts sur les variables d'intérêt du produit. Et d'autre part, sur la base d'applications à des données réelles, notamment des données

3. dépense = $\min(\text{franchise}, \text{coût sinistres})$.

de généalogie et de réseaux sociaux tels que *Facebook* afin d'analyser comment ces types spécifiques de connexions entre les assurés peuvent influer sur ces variables d'intérêt.

Formalisation mathématique du produit

Tout d'abord, le produit de franchise collaborative comporte deux paramètres, à savoir :

- ❖ **Le graphe social** : représentant le réseau des connexions entre les $n \geq 2$ membres de la communauté des assurés et caractérisé par sa matrice d'adjacence $R = (R_{ij})_{1 \leq i,j \leq n}$ telle que :

$$\forall \text{ assurés } i, j = 1 : n, R_{ij} = \begin{cases} 1 & , \text{ si les assurés } i \text{ et } j \text{ sont en relation} \\ 0 & , \text{ sinon.} \end{cases}$$

- ❖ **Les engagements réciproques** : $E = (E_{ij})_{1 \leq i,j \leq n}$, nécessaires à la couverture des franchises des assurés où E_{ij} représente l'engagement de l'assuré i à la couverture de la franchise F_j de sa relation j .

Ensuite, les charges annuelles des assurés de la communauté sont les franchises annuelles $(F_j)_{1 \leq j \leq n}$ et les dépenses annuelles $(D_j)_{1 \leq j \leq n}$ où la dépense individuelle D_j de l'assuré $j = 1 : n$ dans l'année, est définie par :

$$D_j = \min(S_j, F_j) ,$$

avec S_j , représentant le coût de sinistre de l'assuré $j = 1 : n$ dans l'année.

Enfin, les variables d'intérêt du produit de franchise collaborative, notamment les variables de couverture des charges et de contribution aux charges de la communauté des assurés, sont définies comme suit :

- ❖ **Variables de couverture** : $(Couv_f_j)_{1 \leq j \leq n}$ et $(Couv_d_j)_{1 \leq j \leq n}$ représentent respectivement les variables de couvertures individuelles des franchises et des dépenses annuelles des assurés de la communauté, telles que :

$$\forall \text{ assuré } j = 1 : n, Cov_f_j = \sum_{i=1}^n E_{ij} \quad \text{et} \quad Cov_d_j = \sum_{i=1}^n C_{ij} ,$$

avec E_{ij} , représentant l'engagement de l'assuré i à la couverture de franchise de sa relation j et C_{ij} , représentant la contribution de l'assuré i à la couverture de dépense de sa relation j , telle que :

$$\forall \text{ assurés } i, j = 1 : n, C_{ij} = \min \left(\frac{S_j}{F_j} \times E_{ij}, E_{ij} \right) = \begin{cases} E_{ij} & , \text{ si } F_j \leq S_j \\ \frac{S_j}{F_j} \times E_{ij} & , \text{ sinon.} \end{cases}$$

- ❖ **Variables de contribution** : $(Cont_f_i)_{1 \leq i \leq n}$ et $(Cont_d_i)_{1 \leq i \leq n}$ représentent respectivement les variables de contributions individuelles des assurés aux franchises et aux dépenses annuelles de la communauté, telles que :

$$\forall \text{ assuré } i = 1 : n, Cont_f_i = \sum_{j=1}^n E_{ij} \quad \text{et} \quad Cont_d_i = \sum_{j=1}^n C_{ij} .$$

Optimisation du produit

L'objectif principal du mémoire consistant en l'optimisation du produit de franchise collaborative, c'est-à-dire, en la mise en place efficace de mécanismes sur les paramètres du produit afin d'optimiser ses variables d'intérêt, les travaux ont été orientés via les deux interrogations suivantes :

Question (Q1) : Comment fixer le montant des engagements réciproques afin d'optimiser le produit de franchise collaborative ?

Question (Q2) : Quel(s) modèle(s) de graphe social permet(tent) d'optimiser le produit de franchise collaborative ?

(Q1). Optimisation des engagements réciproques

Dans un premier temps, dans le but d'obtenir un montant d'engagement réciproque **équivalent** entre deux assurés en relation, les hypothèses dites de base (**HDB**) suivantes, ont été émises :

(HDB1) Franchise : On suppose que le montant de franchise individuel est constant et identique pour tous les assurés de la communauté, i.e. :

$$\forall \text{ assuré } j = 1 : n, F_j = F ;$$

(HDB2) Engagements réciproques : On suppose que l'engagement E_{ij} de l'assuré i à la couverture de franchise de sa relation j est fonction de l'engagement moyen \bar{E} et de la matrice d'adjacence R du graphe social de la communauté des assurés, tel que :

$$\forall \text{ assurés } i, j = 1 : n, E_{ij} = \bar{E} \times R_{ij} = \begin{cases} \frac{\bar{F}}{\bar{d}}, & \text{si les assurés } i \text{ et } j \text{ sont en relation} \\ 0, & \text{sinon.} \end{cases}$$

où $\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n R_{ij}$, représente le degré moyen des assurés de la communauté (i.e. le nombre moyen de relations par assuré).

Cependant, sous les hypothèses (**HDB**), les charges des assurés moins populaires (i.e. possédant un nombre d'amis inférieur à la moyenne) ne sont pas couvertes par les contributions de leurs relations. En effet, si le degré d_j d'un assuré $j = 1 : n$, est inférieur au degré moyen \bar{d} de la communauté (i.e. $d_j < \bar{d}$), alors les charges annuelles de cet assuré ne sont pas couvertes, i.e. :

$$\text{Couv_}f_j = \sum_{i=1}^n E_{ij} = \frac{F}{\bar{d}} \times d_j < F_j \quad \text{et} \quad \text{Couv_}d_j = \sum_{i=1}^n C_{ij} = \frac{D_j}{\bar{d}} \times d_j < D_j .$$

On constate donc une éventuelle disparité des couvertures individuelles des assurés en fonction de la variance de la distribution des degrés des assurés. De ce fait, une étude de couverture a été effectuée sous les hypothèses (**HDB**), afin de sélectionner le modèle de graphe social permettant d'optimiser le produit de franchise collaborative, par une optimisation des couvertures individuelles des charges annuelles des assurés.

Par ailleurs, pour pallier à ce problème de disparité des couvertures individuelles, sachant que le **système couverture-contribution est en moyenne équilibré** (i.e. les contributions mutuelles des assurés financent les couvertures mutuelles de leurs charges), l'assureur pourrait réallouer les surplus de couverture observés chez les assurés les plus populaires vers les assurés les moins populaires dont les charges ne sont pas couvertes. De ce fait, on suppose dans un second temps, que le montant des engagements réciproques entre les assurés, est cette fois, fonction des degrés respectifs des assurés et non du degré moyen de la communauté, au travers des hypothèses de couverture optimale (**HCO**) :

(HCO1) Franchise : Comme pour l'hypothèse (**HDB1**), on suppose également que le montant de franchise individuel est constant et identique pour tous les assurés de la communauté, i.e. :

$$\forall \text{ assuré } j = 1 : n, F_j = F ;$$

(HCO2) Engagements réciproques : On suppose que l'engagement E_{ij} de l'assuré i à la couverture de franchise F_j de sa relation j est fonction de la franchise F_j et du degré d_j de l'assuré j et de la matrice d'adjacence R du graphe social de la communauté, tel que :

$$\forall \text{ assurés } i, j = 1 : n, E_{ij} = \frac{F_j}{d_j} \times R_{ij} = \begin{cases} \frac{F_j}{d_j} & , \text{ si les assurés } i \text{ et } j \text{ sont en relation} \\ 0 & , \text{ sinon.} \end{cases}$$

On constate que les hypothèses **(HCO)** permettent de couvrir "parfaitement" les charges individuelles des assurés de la communauté, i.e. : \forall assuré $j = 1 : n$,

$$Couv_f_j = \sum_{i=1}^n E_{ij} = F_j \quad \text{et} \quad Cov_d_j = \sum_{i=1}^n C_{ij} = D_j .$$

Cependant, on constate également que les contributions individuelles des assurés sont fonction des degrés de leurs relations. De ce fait, une étude de contribution a été effectuée sous les hypothèses **(HCO)** afin de sélectionner le modèle de graphe social permettant d'optimiser le produit de franchise collaborative, par une minimisation des contributions individuelles (ou cotisations) des assurés.

(Q2). Optimisation du graphe social

Sur la base des charges annuelles des assurés de la communauté fictive **(CF)** présentée dans la section 5.1, une étude de couverture et une étude de contribution ont été respectivement effectuées sous les hypothèses **(HDB)** et **(HCO)** afin de sélectionner le modèle de graphe social permettant d'optimiser les distributions des variables d'intérêt du produit de franchise collaborative. Aussi, afin d'analyser l'impact du graphe social sur ces distributions et prendre en compte les différents scénarios possibles de connexions entre les assurés, quatre modèles classiques de graphe, à savoir les modèles à **clusters homogènes**, d'**Erdős-Rényi**, de **Watts-Strogatz** et de **Barabási-Albert** (présentés dans la section 2.4), ont été utilisés pour la calibration du graphe social de la communauté **CF**.

❖ **Étude de couverture sous (HDB) :** Les couvertures individuelles des charges (franchises et dépenses) annuelles des assurés de la communauté **CF** ont été obtenues à l'aide de la méthode de Monte-Carlo et pour 1.000 simulations de scénarios de graphe social pour chacun des trois modèles de graphe aléatoire (Erdős-Rényi, Watts-Strogatz et Barabási-Albert), sachant que le scénario du modèle à clusters homogènes est déterministe.

Aussi, afin d'analyser et de comparer la performance des couvertures obtenues pour chacun de ces quatre modèles de graphe, on a introduit les indicateurs de couverture suivants : le **taux de couverture** (mesurant la proportion d'assurés dont les charges sont couvertes pour un modèle de graphe donné), la **couverture moyenne** (représentant la moyenne de la distribution de couverture obtenue pour un modèle de graphe donné) et la **volatilité de couverture** (mesurant la dispersion des valeurs de la distribution de couverture obtenue pour un modèle de graphe donné). Le Tableau 1 récapitule les indicateurs de couverture des charges calculés d'après leurs définitions sur la base des charges de la communauté **CF** et des couvertures individuelles de ces charges obtenues sous les hypothèses **(HDB)**, pour chaque modèle de graphe.

Il ressort de cette étude de couverture que sous les hypothèses **(HDB)**, les modèles à clusters homogènes et de Watts-Strogatz permettent d'optimiser les couvertures individuelles des charges annuelles de la communauté **CF**, en présentant les meilleures performances sur les indicateurs de couverture, soit avec des taux de couverture de franchises respectifs de 54% et de 49%.

❖ **Étude de contribution sous (HCO) :** De même que pour l'étude de couverture, afin d'analyser et de comparer la performance des contributions individuelles des assurés obtenues sous les hypothèses **(HCO)**, pour chacun des quatre modèles de graphe étudiés, on a introduit les indicateurs de contribution suivants : le **taux de surcontribution** (mesurant la proportion

d'assurés contribuant au-delà de leurs charges, pour un modèle de graphe donné), la **contribution moyenne** (représentant la moyenne de la distribution de contribution obtenue pour un modèle de graphe donné) et la **volatilité de contribution** (mesurant la dispersion des valeurs de la distribution de contribution obtenue pour un modèle de graphe donné). Le Tableau 2 résume les indicateurs de contribution aux charges calculés d'après leurs définitions sur la base des charges de la communauté CF et des contributions individuelles aux charges des assurés obtenues sous les hypothèses (**HCO**), pour chaque modèle de graphe.

Il ressort de cette étude de contribution que sous les hypothèses (**HCO**), le modèle à clusters homogènes permet d'optimiser les contributions individuelles des assurés aux charges de la communauté CF, en présentant les meilleures performances sur les indicateurs de contribution, soit avec un taux de surcontribution aux franchises de 0% (i.e. les assurés contribuent au maximum pour le même montant que leurs franchises individuelles).

Indicateurs de couverture	Franchises ($F = 1.000 \text{ €}$)			Dépenses ($\bar{D} = 108 \text{ €}$)		
	Moy	Vol	Taux	Moy	Vol	Taux
Modèles						
Clusters homogènes	1.000	362,40	54	105,20	309,63	93
Erdős-Rényi	999,84	21,99	48	108,24	301,09	94
Watts-Strogatz	1.000,20	19,43	49	108,07	300,46	95
Barabási-Albert	995,51	2.847,20	4	88,31	764,73	88,4

TABLE 1 – Tableau récapitulatif : Indicateurs de couverture par modèle de graphe. Taux = taux de couverture (en %), Moy = couverture moyenne (en €) et Vol = volatilité ou écart-type de couverture (en €).

Indicateurs de contribution	Franchises ($F = 1.000 \text{ €}$)			Dépenses ($\bar{D} = 108 \text{ €}$)		
	Moy	Vol	Taux	Moy	Vol	Taux
Modèles						
Clusters homogènes	1.000	0	0	108	74,64	79,5
Erdős-Rényi	999,83	21,69	47	107,7	7,36	88,2
Watts-Strogatz	999,22	18,47	46	108,4	17,95	88,2
Barabási-Albert	1.004,67	5.544,44	4	108,3	615,26	88,4

TABLE 2 – Tableau récapitulatif : Indicateurs de contribution par modèle de graphe. Taux = taux de surcontribution (en %), Moy = contribution moyenne (en €) et Vol = volatilité ou écart-type de contribution (en €).

Mise en place du produit

Les études de couverture et de contribution effectuées ont permis de montrer que les paramètres permettant d'optimiser les variables d'intérêt (en particulier, les couvertures et les contributions individuelles des assurés) du produit de franchise collaborative sont respectivement : les hypothèses (**HCO**) pour la calibration du montant des engagements réciproques entre les assurés et le modèle à **clusters homogènes** pour la calibration du graphe social de la communauté des assurés.

Ces paramètres optimaux ont donc constitué les entrées du produit optimisé (i.e. le scénario central) et une étude de sensibilité a permis de tester ces paramètres selon différentes sensibilités afin d'évaluer leurs impacts sur les sorties suivantes du produit : la **couverture moyenne** des charges et la **contribution maximale** des assurés aux charges de la communauté CF. Le Tableau 3 récapitule les paramètres associés aux trois variantes du produit de franchise collaborative établies.

Paramètres \ Variantes	Scénario central	Sensibilité engagements	Sensibilité graphe social
Engagements réciproques	Hypothèses (HCO)	Hypothèses (HDB)	Hypothèses (HCO)
Graphe social	Clusters homogènes	Clusters homogènes	Erdős-Rényi Watts-Strogatz Barabási-Albert

TABLE 3 – Tableau récapitulatif : Variantes du produit de franchise collaborative établies

Les résultats obtenus pour les sorties du produit associées aux franchises de la communauté CF sont les suivants. Pour une franchise moyenne de 1.000 €, les trois variantes du produit fournissent une couverture moyenne de franchises de 1.000 ± 0 €. Ce qui peut s'expliquer par le fait que le **système de couverture des franchises est équilibré en moyenne** et coïncide avec le montant de franchise moyenne. Cependant, en comparaison aux 1.000 € de contribution maximale aux franchises obtenus pour le scénario central, on obtient des augmentations de 49,1%, de 75,6%, de 64,8% et de 4.498,9% respectivement pour la sensibilité engagements et pour les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert, constituant la sensibilité graphe social.

Conclusion

Ces trois variantes du produit de franchise collaborative établies, combinées à des applications à des données réelles, à savoir des données de généalogie (i.e. connexions par familles) et de réseaux sociaux (e.g. *Facebook*) ont permis de confirmer l'optimalité du scénario central et d'analyser les impacts des paramètres (engagements réciproques et graphe social) sur les sorties du produit. Ainsi, le scénario central constitue la variante optimisée du produit de franchise collaborative qu'un assureur serait susceptible de proposer à une éventuelle communauté d'assurés, comme suit :

- 1) **Le graphe social** : calibré via le modèle à **clusters homogènes** suivant les deux cas envisageables :

- ✓ soit la connexion est **libre**, c'est-à-dire, les assurés peuvent former librement leurs groupes ou réseaux de relations avec la contrainte qu'un assuré ne peut appartenir qu'à un unique groupe indépendamment des autres groupes (e.g. entre amis, proches ou familles) ;
- ✓ soit la connexion est **imposée** par l'assureur qui peut alors mettre les assurés dans des groupes homogènes indépendants où les assurés appartenant à un même groupe ont des caractéristiques les plus similaires (e.g. âge ou profil de risque).

- 2) **Les engagements réciproques** : fixés suivant les hypothèses (**HCO**) afin non seulement de garantir une couverture optimale des charges, mais également combinées au point 1) de minimiser les contributions individuelles des assurés aux charges de leurs relations.

En effet, les hypothèses (**HCO**) combinées au point 1) permettent d'obtenir un engagement réciproque entre deux assurés appartenant à un même groupe, fonction du degré du groupe de telle sorte que plus ce degré augmente (i.e. le nombre d'assurés du groupe augmente), plus cet engagement diminue. Ce qui pourrait être un **argument commercial** afin d'inciter les assurés à agrandir leur réseau de relations et permettrait ainsi de les fidéliser tout en accroissant le nombre de nouvelles souscriptions.

Mots clés :

Assurance collaborative, Franchise collaborative, Réseaux, Graphes, Couverture de charges, Contribution aux charges.

Executive summary

Context and problematic

This thesis is interested in valorising P&C insurance products based on collective deductibles inspired by the product called « peer-to-peer deductible », initiated in 2014 by the French startup *Inspeer*, that consist for an policyholder to mutualise his deductible insurance of damages (auto, home, bike) with his relatives and this without changing insurance contract. So, this system helps the policyholder firstly decrease the cost of his traditional insurance by increasing his deductible and secondly cover his deductible insurance by the community of policyholders. However, given that the policyholder is committed to contributing equivalently to his relations deductibles, he carries a part of risk whether a great deal of his relationship is claimed. In addition, the startup only collected contributions from policyholders in the event of a claim in their relationship. That undeniably created a risk of default by policyholders. Also, the startup had to file for bankruptcy in 2018, for lack of sufficient capital and subscriptions to finance its development.

Moreover, it must be noted that, the coverages and contributions of the policyholders that make the interest variables of the peer-to-peer deductible product, are strongly dependent on the structure of the policyholders network.

It is in that analysis of the impact of the policyholders network structure on those interest variables that this actuarial thesis is interested in whose main objective consists in, with a little more hindsight since 2018, set efficiently mechanisms that permit to optimize the product, that is to say to guarantee annual charges (deductibles and expenses⁴) coverage of a maximal number of policyholders as well as the minimization of the individual contributions of the policyholders to their relations charges.

Methodology

In order to carry out this objective of optimization of the peer-to-peer deductible product, the following step has been adopted.

- ❖ First, a mathematical formalization of the product has been made and a definition of its interest variables.
- ❖ Then, based on a fictional community of policyholders, optimization studies have respectively been done on product's parameters such as the amount of the reciprocal contributions and the model of social graph representing the network of connections between the policyholders.
- ❖ Finally, a peer-to-peer deductible product have been set on the basis of the optimized product called « central scenario » taking as entry, the optimal parameters from the optimization studies of the product and a study of the sensitivity assessing those parameters according to different sensitivities in order to evaluate their impact on the interest variables, on the one hand. On the other hand, based on applications to real data namely genealogical data and social networks such as *Facebook* so as to analyse how those specific types of connections between the policyholders can impact on those interest variables.

4. expense = min(deductible, claims cost).

Mathematical formalization of the product

First, the peer-to-peer deductible product include two parameters namely :

- ❖ **Social graph :** representing network of connections between the $n \geq 2$ policyholders of the community and described by his adjacency matrix $R = (R_{ij})_{1 \leq i,j \leq n}$ such as :

$$\forall \text{ insured } i, j = 1 : n, R_{ij} = \begin{cases} 1 & , \text{ if the insured } i \text{ and } j \text{ are in relation} \\ 0 & , \text{ otherwise.} \end{cases}$$

- ❖ **Reciprocal contributions :** $E = (E_{ij})_{1 \leq i,j \leq n}$, necessary to cover the insured deductibles where E_{ij} represents the contribution of the insured i to cover the deductible F_j of his relation j .

Then, the annual charges of insured's community are the annual deductibles $(F_j)_{1 \leq j \leq n}$ and the annual expenses $(D_j)_{1 \leq j \leq n}$ where the individual expense D_j of the insured $j = 1 : n$ in the year, is defined by :

$$D_j = \min(S_j, F_j),$$

with S_j , representing the claims cost of the insured $j = 1 : n$ in the year.

Finally, the interest variables of the peer-to-peer deductible product, namely the charges coverage variables and contribution variables to charges of the community's insured, are defined by :

- ❖ **Coverage variables :** $(Couv_f_j)_{1 \leq j \leq n}$ and $(Couv_d_j)_{1 \leq j \leq n}$ respectively represent individual coverages variables of the deductibles and the expenses of the insured, such as :

$$\forall \text{ insured } j = 1 : n, Cov_f_j = \sum_{i=1}^n E_{ij} \quad \text{and} \quad Cov_d_j = \sum_{i=1}^n C_{ij},$$

with E_{ij} , representing contribution of the insured i to cover the deductible of his relation j and C_{ij} , representing contribution of the insured i to cover the expense D_j of his relation j , such as :

$$\forall \text{ insured } i, j = 1 : n, C_{ij} = \min\left(\frac{S_j}{F_j} \times E_{ij}, E_{ij}\right) = \begin{cases} E_{ij} & , \text{ if } F_j \leq S_j \\ \frac{S_j}{F_j} \times E_{ij} & , \text{ otherwise.} \end{cases}$$

- ❖ **Contribution variables :** $(Cont_f_i)_{1 \leq i \leq n}$ and $(Cont_d_i)_{1 \leq i \leq n}$ respectively represent individual contributions variables to the deductibles and the expenses of the insured, such as :

$$\forall \text{ insured } i = 1 : n, Cont_f_i = \sum_{j=1}^n E_{ij} \quad \text{and} \quad Cont_d_i = \sum_{j=1}^n C_{ij}.$$

Optimization of the product

The main objective of this thesis consists in optimizing the peer-to-peer deductible product, that is to say to set efficiently mechanisms on the parameters of the product so as to optimize its interest variables. To that end, reflections have been made on those parameters through these two following questions :

Question (Q1) : How to set the amount of reciprocal contributions in order to optimize the peer-to-peer deductible product ?

Question (Q2) : What model of social graph does permit to optimize the peer-to-peer deductible product ?

(Q1). Optimization of reciprocal contributions

First, in order to get an **equivalent** reciprocal contribution between two policyholders in relation, the following basis-hypotheses (**HDB**) « Hypothèses De Base » have been suggested :

(HDB1) Deductible : We assume that the amount of individual deductible is constant and identical for all the policyholders of the community, i.e. :

$$\forall \text{ insured } j = 1 : n, F_j = F ;$$

(HDB2) Reciprocal contributions : We assume that the contribution E_{ij} of the insured i to cover the deductible of his relation j is function of mean contribution \bar{E} and the adjacency matrix R of social graph of the insured's community, such as :

$$\forall \text{ insured } i, j = 1 : n, E_{ij} = \bar{E} \times R_{ij} = \begin{cases} \frac{\bar{F}}{\bar{d}} & , \text{ if the insured } i \text{ and } j \text{ are in relation} \\ 0 & , \text{ otherwise.} \end{cases}$$

where $\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n R_{ij}$, represents the mean degree of the insured's community (i.e. mean number of relations per insured).

However, under the (**HDB**) assumptions, the charges of less popular policyholders (i.e. having a lower number of relations than the average) are not covered by their relationship contributions. Indeed, if the degree d_j of an insured $j = 1 : n$, is less than the mean degree \bar{d} of the community (i.e. $d_j < \bar{d}$), then the annual charges of this insured are not covered, i.e. :

$$Couv_f_j = \sum_{i=1}^n E_{ij} = \frac{F}{\bar{d}} \times d_j < F_j \quad \text{and} \quad Couv_d_j = \sum_{i=1}^n C_{ij} = \frac{D_j}{\bar{d}} \times d_j < D_j .$$

There is therefore a possible disparity in the individual coverages of policyholders depending on the volatility of the policyholders degrees distribution. So, a coverage study was carried out under the (**HDB**) assumptions, in order to select the social graph model to optimize the peer-to-peer deductible product, by optimizing the individual coverages of the insured's annual charges.

Moreover, to resolve this problem of disparity in the individual coverages, knowing that the **coverage-contribution system is on average balanced** (i.e. mutual contributions of insured finance mutual coverages of their charges), the insurer could reallocate the excess cover observed among the most popular policyholders to the less popular policyholders, whose charges are not covered. Therefore, we assume secondly, that the amount of reciprocal contributions between the insured, now, is function of the insured's respective degrees and not of community's mean degree, through the hypotheses of optimal coverage (**HCO**) « Hypothèses de Couverture Optimale » :

(HCO1) Deductible : Like (**HDB1**) hypothesis, we assume that the amount of individual deductible is constant and identical for all the policyholders of the community, i.e. :

$$\forall \text{ insured } j = 1 : n, F_j = F ;$$

(HCO2) Reciprocal contributions : We assume that the contribution E_{ij} of the insured i to cover the deductible F_j of his relation j is function of the deductible F_j and the degree d_j of the insured j and of the adjacency matrix R of social graph of the insured's community, such as :

$$\forall \text{ insured } i, j = 1 : n, E_{ij} = \frac{F_j}{d_j} \times R_{ij} = \begin{cases} \frac{F_j}{d_j} & , \text{ if the insured } i \text{ and } j \text{ are in relation} \\ 0 & , \text{ otherwise.} \end{cases}$$

It can be seen that the **(HCO)** hypotheses make it possible to cover "perfectly" the individual charges of the insured members of the community, i.e. : \forall insured $j = 1 : n$,

$$\text{Couv_}f_j = \sum_{i=1}^n E_{ij} = F_j \quad \text{and} \quad \text{Couv_}d_j = \sum_{i=1}^n C_{ij} = D_j .$$

However, we also note that the individual contributions of the insured depend on the degrees of their relationship. Therefore, a contribution study was carried out under the **(HCO)** hypotheses in order to select the social graph model making it possible to optimize the peer-to-peer deductible product, by minimizing the individual contributions of policyholders.

(Q2). Optimization of social graph

On the basis of annual charges of the policyholders of fictional community **(CF)** shown in section 5.1, a coverage study and a contribution study have respectively been made according to the **(HDB)** and **(HCO)** hypotheses so as to select the model of social graph allowing the optimization of the distributions of interest variables of the peer-to-peer deductible product. Also, in order to analyse the impact of the social graph on those distributions and to take into account the different possible scenarios of connections among the policyholders, four traditional models of graph namely **homogeneous clusters**, **Erdos Rényi**, **Watts-Strogatz** and **Barabási-Albert** models (shown in section 2.4) have been used to calibrate the social graph of the community **CF**.

❖ **Coverage study under (HDB)** : The individual coverages of the insured's annual charges have been obtained through the Monte-Carlo method with 1.000 simulations of social graph scenarios for each random graph model (Erdős-Rényi, Watts-Strogatz et Barabási-Albert) knowing that the scenario of the homogeneous clusters model is deterministic.

Also in order to analyse and compare the performance of the coverage got for each of those four graph models, we bring in the following coverage indicators : the **coverage rate** (measuring the proportion of the policyholders whose charges are covered for a given graph model), the **mean coverage** (representing the average of the coverage distribution obtained for a given graph model) and the **coverage volatility** (measuring the dispersion of the values of the coverage distribution obtained for a given graph model). The Table 4 summarizes the coverage indicators of the charges calculated from their definitions on the basis **CF** community's charges and individual coverages of those charges got under the **(HDB)** hypotheses, for each graph model.

We notice from this coverage study that under the **(HDB)** hypotheses, homogeneous clusters and Watts-Strogatz models permit to optimize individual coverage of annual charges of community **CF** showing the best performances on the coverage indicators, with respectively 54% and 49% as rates of deductibles coverage.

❖ **Contribution study under (HCO)** : Similarly to the coverage study, in order to analyse and compare the performance of the insured's individual contributions obtained under the **(HCO)** hypotheses, for each of the four graphs models studied, following contribution indicators are introduced : the **over-contribution rate** (measuring the proportion of the policyholders who contribute beyond their charges for a given graph model), the **mean contribution** (representing the average of the contribution distribution obtained for a given graph model) and the **contribution volatility** (measuring the dispersion of the values of the contribution distribution obtained for a given graph model). The Table 5 summarizes the contribution indicators to the charges calculated from their definitions on the basis **CF** community charges and individual contributions to charges of insured got under the **(HCO)** hypotheses, for each graph model.

We notice from this contribution study that under the **(HCO)** hypotheses, the homogeneous clusters model helps to optimize the individual contributions of the policyholders to the charges of community **CF** by showing the best performances on the contribution indicators, with 0% as over-contribution rate to deductibles (i.e. the policyholders exactly contribute to the same amount as their individual deductibles).

Coverage indicators	Deductibles ($F = 1.000 \text{ €}$)			Expenses ($\bar{D} = 108 \text{ €}$)		
Models	Mean	Vol	Rate	Mean	Vol	Rate
Homogeneous clusters	1.000	362,40	54	105,20	309,63	93
Erdős-Rényi	999,84	21,99	48	108,24	301,09	94
Watts-Strogatz	1.000,20	19,43	49	108,07	300,46	95
Barabási-Albert	995,51	2.847,20	4	88,31	764,73	88,4

TABLE 4 – Summary table : Coverage indicators by graph model. Rate = coverage rate (in %), Mean = mean coverage (in €) and Vol = volatility or standard-deviation of coverage (in €).

Contribution indicators	Deductibles ($F = 1.000 \text{ €}$)			Expenses ($\bar{D} = 108 \text{ €}$)		
Models	Mean	Vol	Rate	Mean	Vol	Rate
Homogeneous clusters	1.000	0	0	108	74,64	79,5
Erdős-Rényi	999,83	21,69	47	107,7	7,36	88,2
Watts-Strogatz	999,22	18,47	46	108,4	17,95	88,2
Barabási-Albert	1.004,67	5.544,44	4	108,3	615,26	88,4

TABLE 5 – Summary table : Contribution indicators by graph model. Rate = over-contribution rate (in %), Mean = mean contribution (in €) and Vol = volatility or standard-deviation of contribution (in €).

Setting of the product

The coverage and contribution studies have helped to show that the parameters permitting to optimize the interest variables (particularly, the individual coverages and contributions of the policyholders) of the peer-to-peer deductible product are respectively : the **(HCO)** hypotheses for the calibration of the amount of reciprocal contributions between the policyholders and the **homogeneous clusters** model for the calibration of the social graph of the policyholders community.

Those optimal parameters have made up the inputs of the optimized product (i.e. central scenario) and a sensitivity study helped to test those parameters through different sensitivities so as to assess their impacts on the following outputs of the product : the **mean coverage** of the charges and the **maximum contribution** of policyholders to **CF** community's charges. The Table 6 summarizes the parameters associated with the three variants of the peer-to-peer deductible product established.

Results obtained from the outputs of the product associated with the **CF** community's deductibles are the following. For an mean deductible of 1.000 €, the three variants provide an average coverage of $1.000 \pm 0 \text{ €}$. This can be explained by the fact that the **deductibles coverage system is balanced in average** and coincides with the amount of mean deductible. However in comparison with the 1.000 € of maximum contribution to deductibles obtained for the central scenario, we get increases of 49.1%, of 75.6%, of 64.8% and of 4,498.9% respectively for sensitivity contributions and for the models by Erdos-Rényi, Watts-Strogatz and Barabási-Albert that make up the sensitivity social graph.

Parameters \ Variants	Central scenario	Sensitivity contributions	Sensitivity social graph
Reciprocal contributions	Hypotheses (HCO)	Hypotheses (HDB)	Hypotheses (HCO)
Social graph	Homogeneous clusters	Homogeneous clusters	Erdős-Rényi Watts-Strogatz Barabási-Albert

TABLE 6 – Summary table : Variants of the peer-to-peer deductible product established

Conclusion

These three variants of the peer-to-peer deductible product established, combined with applications to real data namely genealogical data (i.e. connections per family) and social networks (e.g. *Facebook*) have contributed to confirm the optimality of central scenario and analyse the impacts of parameters (reciprocal contributions and social graph) on the outputs of the product. Thus, the central scenario is the optimized variant of the peer-to-peer deductible product that an insurer would likely suggest to a possible community of policyholders as follows :

1) The social graph : calibrated through the **homogeneous clusters** model following the two possible cases :

- ✓ either the connection is **free**, that is to say the insured can freely create their groups or network of relations with the constraint that a policyholder can belong only to a unique group independently of others (e.g. among friends, relatives, or families)
- ✓ or the connection is **imposed** by the insurer who can then put the policyholders in homogeneous and independent groups where the policyholders belong to a same group have the most similar characteristics (e.g. age or risk profile)

2) The reciprocal contributions : set following the **(HCO)** hypotheses so as to guarantee both an optimal coverage of charges and minimize the individual contributions of the policyholders to their relations charges. Indeed, the **(HCO)** hypotheses with **homogeneous clusters** model help to get a reciprocal contribution between two policyholders from a same group, function of the group degree to the point that the more that degree increases, the more that contribution decreases. That could be a **marketing argument** in order to encourage policyholders to expand their network of relationships and thus build loyalty while increasing the number of new subscriptions.

Keywords :

Peer-to-peer insurance, Peer-to-peer deductible, Networks, Graphs, Coverage of charges, Contribution to charges.

Remerciements

Au terme de la rédaction de ce mémoire d'actuariat, je tiens à exprimer toute ma gratitude à toutes les personnes ayant contribuées de près ou de loin au bon déroulement de cette expérience professionnelle.

Tout d'abord, je tiens à remercier M. Arthur CHARPENTIER, mon directeur de mémoire, pour ses conseils judicieux et la confiance qu'il m'a accordée en me permettant de travailler sur ce sujet de recherche et développement qui fut très enrichissant.

Ensuite, je dis également merci à M. Franck VERMET, mon tuteur académique, pour son suivi régulier, sa disponibilité et ses remarques pertinentes qui ont orientées mes travaux tout le long de la rédaction de ce mémoire.

Aussi, j'adresse mes sincères remerciements à tous les professeurs et intervenants de l'EURIA pour la qualité de leurs cours prodigués et leurs contributions au bon déroulement de cette expérience.

Enfin, à toutes les personnes qui par leurs conseils et leurs critiques ont guidées mes réflexions pour la bonne réalisation de ce mémoire, qui est le couronnement de plusieurs années d'études, je souhaiterais vous dire un seul mot : MERCI !

Table des matières

Résumé	i
Abstract	ii
Note de synthèse	iii
Executive summary	ix
Remerciements	xv
Introduction	1
I Concepts Fondamentaux	3
1 Marché de l'assurance collaborative	4
1.1 Principe de la méthode	4
1.2 Différents modèles du marché	5
1.2.1 L'achat groupé d'assurance	5
1.2.2 La franchise collaborative	5
1.2.3 La micro-mutuelle collective	6
1.2.4 L'autoassurance	6
1.3 Historique et principaux acteurs du marché	6
1.3.1 Generali (France, 2009)	6
1.3.2 Friendsurance (Allemagne, 2010)	7
1.3.3 Bought By Many (Royaume-Uni, 2012)	7
1.3.4 Guevara (Royaume-Uni, 2014)	8
1.3.5 Inspeer (France, 2014)	8
1.3.6 Lemonade (USA, 2015)	9
1.3.7 Groupago (Belgique, 2015)	10
1.3.8 Wizzas (France, 2015)	10
1.3.9 Otherwise (France, 2016)	11
1.3.10 Wecover (France, 2017)	12
1.3.11 Besure (Canada, 2017)	12
1.3.12 Teamrella (Russie, 2017)	13
1.4 Contexte et problématique	13
1.5 Méthodologie	13
2 Théorie des réseaux	15
2.1 Introduction	15
2.2 Réseaux et graphes	18
2.3 Caractéristiques des graphes	19

2.4	Modèles classiques de graphe	20
2.4.1	Le modèle à clusters homogènes	21
2.4.2	Le modèle d'Erdős-Rényi	22
2.4.3	Le modèle du petit monde de Watts-Strogatz	23
2.4.4	Le modèle d'attachement préférentiel de Barabási-Albert	24
2.5	Conclusion	26
	Synthèse de la partie I	27
	II Optimisation du produit de franchise collaborative	28
3	Introduction au produit de franchise collaborative	29
3.1	Vocabulaire et terminologie	29
3.1.1	Vocabulaire de l'assurance	29
3.1.2	Vocabulaire spécifique au produit de franchise collaborative	31
3.2	Introduction au produit	32
3.3	Formalisation mathématique du produit	33
3.3.1	Premières définitions	33
3.3.2	Définitions des variables d'intérêt	34
3.4	Pistes de réflexion	36
4	Étude théorique des engagements réciproques optimaux	37
4.1	Hypothèses de base (HDB)	37
4.1.1	Hypothèses	37
4.1.2	Remarques	38
4.1.3	Résultats théoriques	39
4.1.4	Analyse des résultats	41
4.2	Hypothèses de couverture optimale (HCO)	42
4.2.1	Hypothèses	42
4.2.2	Remarques	42
4.2.3	Résultats théoriques	43
4.2.4	Analyse des résultats	45
4.3	Synthèse hypothèses HDB vs HCO	46
5	Étude pratique du modèle de graphe social optimal	50
5.1	Présentation de la communauté fictive	50
5.1.1	Description de la communauté fictive	50
5.1.2	Analyse des charges de la communauté fictive	51
5.2	Calibration des paramètres	53
5.2.1	Calibration du graphe social	53
5.2.2	Calibration des engagements réciproques	58
5.3	Programme de calcul PC1	59
5.3.1	Les entrées du programme PC1	59
5.3.2	Les sorties du programme PC1	59
5.3.3	Le fonctionnement du programme PC1	59
5.4	Étude de couverture	60
5.4.1	Analyse de la couverture en fonction du degré	60
5.4.2	Sélection de modèle de graphe	65
5.5	Étude de contribution	69
5.5.1	Analyse de la contribution en fonction du degré	70
5.5.2	Sélection de modèle de graphe	75

Synthèse de la partie II	80
III Mise en place d'un produit de franchise collaborative	81
6 Aspect technique du produit de franchise collaborative	82
6.1 Introduction	82
6.2 Programme de calcul PC2	83
6.2.1 Les entrées du programme PC2	83
6.2.2 Les sorties du programme PC2	83
6.2.3 Le fonctionnement du programme PC2	83
6.3 Scénario central	84
6.4 Étude de sensibilité	86
6.4.1 Sensibilité engagements	86
6.4.2 Sensibilité graphe social	89
6.5 Conclusion	99
7 Applications à des données réelles	100
7.1 Données de Généalogie (DG)	100
7.1.1 Description des données DG	100
7.1.2 Propriétés du graphe social de la communauté DG	101
7.1.3 Sorties associées aux franchises de la communauté DG	102
7.2 Données de Facebook (DF)	104
7.2.1 Description des données DF	104
7.2.2 Propriétés du graphe social de la communauté DF	104
7.2.3 Sorties associées aux franchises de la communauté DF	106
Synthèse de la partie III	108
Conclusion	109
Annexes	i
Annexe A Données de Généalogie (DG)	i
Annexe B Rappels théoriques	iii
B.1 Réseau sans échelle et loi de puissance	iii
B.2 Classification Ascendante Hiérarchique (CAH)	iv
Bibliographie	vi

Introduction

Ces dernières années, l'explosion du numérique (*digital* en anglais) et le contexte socio-économique assez tendu favorisent une montée en puissance de l'économie collaborative. En effet, neuf Français sur dix ont déjà vendu, acheté, loué ou mis à disposition des biens ou des services à d'autres particuliers via Internet⁵, grâce à l'émergence de plateformes comme *Blablacar*, *Leboncoin* ou encore *Airbnb*. Le secteur de l'assurance n'est pas en reste avec notamment, depuis 2009, plusieurs initiatives autour de l'assurance collaborative à travers le monde, telles que *Friendsurance* en Allemagne depuis 2010 et *Lemonade* aux USA depuis 2015. En France, le site *Inspeer* avait lancé en 2014, le premier produit d'assurance collaborative basé sur des franchises collectives dit « franchise collaborative », permettant de mutualiser ses franchises d'assurance dommage (automobile, habitation, moto) avec ses proches et ce, sans changer de contrat d'assurance. L'assuré demande à ses proches de contribuer à sa franchise et il s'engage à contribuer de manière équivalente à la franchise de ces derniers en cas de sinistre.

Ce système permet donc à l'assuré de : diminuer le coût de son assurance classique en augmentant sa franchise d'assurance⁶ et faire couvrir sa franchise d'assurance par la communauté des assurés. Pour l'assureur, ce système permet de réduire les coûts de gestion des petits sinistres et la fraude (cette dernière étant estimée en 2018, à environ 1 demi-milliard d'euros pour l'assurance IARD en France⁷). Cependant, étant donné que l'assuré s'engage à contribuer de manière équivalente à la franchise de ses relations, il porte une partie du risque en cas de sinistre d'un grand nombre de ces derniers. Pour pallier à ce risque, Dubois [2015] propose une assurance complémentaire dit « assurance stoploss » afin de limiter la contribution d'un assuré en cas de sinistre de ses relations. En effet, force est de constater que, les couvertures et les contributions des assurés constituant les variables d'intérêt du produit de franchise collaborative, sont fortement dépendantes de la structure du réseau de la communauté des assurés.

L'objectif de ce mémoire consiste donc à analyser l'impact de la structure du réseau des assurés sur ces variables d'intérêt et à mettre en place efficacement des mécanismes permettant l'optimisation du produit, c'est-à-dire : garantir la couverture des charges (franchises et dépenses⁸) annuelles d'un nombre maximum d'assurés tout en minimisant les contributions individuelles des assurés à la couverture des charges de leurs relations. Pour ce faire, ce mémoire s'articule en trois parties :

- ❖ **La première partie :** consistera à introduire les **Concepts Fondamentaux** nécessaires à la compréhension de la suite des travaux, notamment le chapitre 1 présentera le marché mondial de l'assurance collaborative, notamment l'état de l'art, les différents modèles (dont le modèle de franchise collaborative à l'étude) et les principaux acteurs historiques du marché. Aussi, il fera un rappel du contexte de l'étude, de la problématique actuarielle traitée et de la méthodologie adoptée. Enfin, le chapitre 2 permettra l'introduction de

5. source : Cnc [2016].

6. d'après la relation : prime = $\mathbb{E}[\max(\text{coût sinistres} - \text{franchise}, 0)]$.

7. source : <https://www.alfa.asso.fr>.

8. dépense = min(franchise, coût sinistres).

modèles classiques de graphe de la théorie des réseaux, tels que les modèles à **clusters homogènes**, d'**Erdős-Rényi**, de **Watts-Strogatz** et de **Barabási-Albert**. En effet, ces modèles de graphe permettront de prendre en compte les différents scénarios possibles de connexions entre les membres d'une communauté d'assurés et d'analyser *in fine*, l'impact de la structure du réseau des assurés sur les variables d'intérêt du produit de franchise collaborative.

- ❖ **La deuxième partie :** sera consacrée à l'[Optimisation du produit de franchise collaborative](#). En effet, le chapitre [3](#) consistera à introduire, formaliser et définir les variables d'intérêt du produit de franchise collaborative, notamment la couverture des charges et la contribution aux charges des membres de la communauté des assurés. Aussi, les chapitres [4](#) et [5](#) traiteront respectivement de l'optimisation des paramètres (i.e. le montant des engagements réciproques ^{[9](#)} et le modèle de graphe social ^{[10](#)}) du produit, en cherchant à comprendre : 1) comment fixer le montant des engagements réciproques entre les assurés et 2) quel(s) modèle(s) de graphe social permet(tent) d'optimiser le produit de franchise collaborative.
- ❖ **La troisième partie :** sera dédiée à la [Mise en place d'un produit de franchise collaborative](#) d'une part au chapitre [6](#), sur la base du produit optimisé dit « scénario central » prenant en entrée les paramètres optimaux issus des études d'optimisation du produit et d'une étude de sensibilité testant ces paramètres selon différentes sensibilités afin d'évaluer leurs impacts sur les variables d'intérêt du produit. Et d'autre part au chapitre [7](#), sur la base d'applications à des données réelles, notamment des données de généalogie et de réseaux sociaux tels que *Facebook* afin d'analyser comment ces types spécifiques de connexions entre les assurés peuvent influer sur ces variables d'intérêt.

9. Un engagement réciproque est un montant fixe défini par deux assurés en relation, servant à contribuer réciproquement à la couverture de leurs franchises en cas de sinistre.

10. Le graphe social représente le graphe du réseau des relations entre les assurés d'une communauté.

Première partie

Concepts Fondamentaux

Chapitre 1

Marché de l'assurance collaborative

Ce chapitre a pour objectif de présenter le contexte dans lequel s'inscrit ce mémoire et le marché mondial de l'**assurance collaborative** (ou **assurance P2P**¹). En effet, il consistera d'une part, à faire l'état de l'art du secteur de l'assurance collaborative, notamment le principe de la méthode, les différents modèles et les principaux acteurs historiques du marché. Et d'autre part, à introduire le contexte de l'étude, la problématique actuarielle traitée et la méthodologie mise en place pour répondre à cette problématique.

1.1 Principe de la méthode

L'assurance collaborative repose sur un principe simple : **regrouper des personnes et mettre leurs risques en commun pour qu'elles s'assurent mutuellement.** Et ce principe peut s'appliquer à tous les niveaux : santé, automobile, habitation, biens matériels, animaux, etc.

Pour que la mutualisation des risques fonctionne, il faut que les assurés forment une communauté possédant des comportements et besoins, les plus similaires possibles. Les personnes sont regroupées selon leurs profils ou selon leurs choix. Par exemple pour l'automobile, on peut regrouper les conducteurs expérimentés et pour la santé, on peut regrouper les personnes âgées souhaitant une complémentaire santé plus couvrante.

Le phénomène est d'autant plus évident qu'il s'accommode parfaitement avec le climat socio-économique actuel porté par les deux tendances suivantes :

- ❖ **Le contexte économique** : Le budget assurance des familles ne cesse de croître alors que la situation économique réduit le portefeuille disponible des ménages ;
- ❖ **L'explosion du digital** : Les réseaux sociaux permettent de partager rapidement les bons plans et les pratiques collaboratives entre particuliers.

De plus, l'assurance est un domaine que les consommateurs estiment flou, avec beaucoup trop d'intermédiaires, trop d'offres peu lisibles et pas toujours conçues selon leurs besoins réels. Le Tableau 1.1 suivant résume les intérêts et les limites de l'assurance collaborative.

1. assurance P2P (assurance Peer-to-Peer ou Pair-à-Pair) : concept d'assurance basé sur la réciprocité entre assurés. Il s'agit de "mettre en réseau" des personnes ayant le même besoin d'assurance.

Intérêts	Limites
Assureur : <ul style="list-style-type: none"> -Diminution du risque de fraude -Diminution de l'aléa moral -Diminution de l'antisélection -Réduction des coûts de gestion -Réduction des coûts d'acquisition -Augmentation de la valeur client 	Assureur : <ul style="list-style-type: none"> -Uniquement risques IARD -Uniquement risques relativement petits
Assurés : <ul style="list-style-type: none"> -Meilleure expérience client -Couverture des besoins peu ou mal couverts -Réduction des primes -Indemnisation plus efficace et plus rapide -Possibilité de cashback 	Assurés : <ul style="list-style-type: none"> -Risque de contrepartie -Risque de surcontribution

TABLE 1.1 – Intérêts et limites de l’assurance collaborative

1.2 Différents modèles du marché

Sur le marché actuel de l’assurance collaborative, quatre grandes tendances, plus ou moins ambitieuses, se profilent. Elles sont toutes tournées vers un mode d’achat en nombre, avec une participation relative des acteurs classiques de l’assurance : l'**achat groupé d’assurance**, la **franchise collaborative**, la **micro-mutuelle collective** et l'**autoassurance**.

1.2.1 L’achat groupé d’assurance

❖ **Principe :** L’achat groupé d’assurance peut être sous deux formes.

- ✓ Soit la création d’une communauté qui permettra de négocier les prix d’un contrat d’assurance traditionnel. Dans ce cas, c’est le nombre de personnes qui est important : “*L’union fait la force*” ;
- ✓ Soit la création d’une communauté avec des besoins spécifiques en vue de développer un produit adapté et unique (e.g. les sportifs, les diabétiques, etc.). Dans ce cas, le produit est dédié à une communauté bien précise et faible.

❖ **Acteurs :** Bought By Many (Royaume-Uni, 2012), Groupago (Belgique, 2015), Wizzas.com (France, 2015).

1.2.2 La franchise collaborative

❖ **Principe :** Mutualiser sa franchise d’assurance dommage (automobile, habitation, moto) avec ses proches et ce, sans changer de contrat d’assurance. L’assuré demande à ses proches de contribuer à sa franchise et il s’engage à contribuer de manière équivalente à la franchise de ces derniers en cas de sinistre.

Ce principe permet donc pour un assuré, de réduire le coût de son assurance classique en augmentant sa franchise et en cas de sinistre, cette franchise est supportée financièrement par la communauté.

❖ **Acteurs :** Inspeer (France, 2014).

1.2.3 La micro-mutuelle collective

❖ **Principe :** L'internaute choisit une offre d'assurance traditionnelle (proposée notamment sur le site), il se construit une communauté (100 personnes en moyenne), de préférence des proches, de la famille ou avec des intérêts communs.

La communauté au complet versera une cotisation qui sera divisée en deux parties, 80 % de la somme servira à un fonds de garantie commun (le *float*) pour régler les petits sinistres et les 20 % restants seront utilisés pour souscrire une assurance classique pour les risques lourds uniquement.

Le concept est ainsi doublement intéressant pour les assurés :

- ✓ Les petits sinistres sont souvent mal couverts par les compagnies car des franchises s'appliquent, les démarches sont fastidieuses pour des petites sommes. Gérer ces soucis du quotidien au sein même de la cohorte en piochant la somme nécessaire dans le *float*, permet plus de souplesse et le reliquat est *in fine* redistribué aux assurés en fin d'année sous forme de *cashback*. Ce qui pousse également à une conduite responsable des assurés ;
- ✓ Les sinistres plus importants restent couverts par une compagnie traditionnelle afin de faire face à des dépenses plus lourdes.

❖ **Acteurs :** Friendsurance (Allemagne, 2010), Guevara (Royaume-Uni, 2014), Otherwise (France, 2016), Wecover (France, 2017), Besure (Canada, 2017).

1.2.4 L'autoassurance

❖ **Principe :** Ce modèle est le stade ultime de l'assurance collaborative où on supprime tous les intermédiaires pour s'autoassurer.

Ce principe permet donc de payer sa cotisation en cryptomonnaie notamment en bitcoins pour enlever l'intermédiaire bancaire.

Pour le moment, ce niveau d'assurance collaborative permet de couvrir seulement les petits risques, mais ce concept reste encore un peu flou pour les usagers. Comme toute innovation, il faut un temps d'adaptation et les futurs usagers passeront à ce niveau, quand il sera mieux développé et plus médiatisé.

❖ **Acteurs :** Peercover.com (USA, 2013), Teambrella (Russie, 2017).

1.3 Historique et principaux acteurs du marché

Depuis quelques années, on compte une trentaine d'initiatives autour de l'assurance collaborative à travers le monde. Toutes proposent un service essentiellement *digital* (i.e. en ligne) et ont pour objectif de réduire le coût de l'assurance. Ces startups ont également toutes à cœur de challenger les acteurs historiques du secteur, en améliorant significativement l'expérience client.

Voici une liste non exhaustive et chronologique des principaux acteurs du marché mondial de l'assurance collaborative, sur la base des informations disponibles lors de la rédaction de ce mémoire.

1.3.1 Generali (France, 2009)

En 2009, l'assureur italien **Generali**, au travers de sa filiale française *Equité*, était à l'initiative de la plateforme *kontournous.fr*. Cette plateforme initiée en France, proposait une offre d'**achat groupé** d'assurance **automobile**.



Le principe de la plateforme était d'assurer à prix avantageux une « tribu » de 2 à 15 personnes (familles, amis, ou potes). Pour cela, l'internaute devait créer une tribu à sa convenance, ou rejoindre une tribu existante. Il souscrivait ensuite individuellement un contrat auto adapté à ses propres besoins.

C'est le simple fait de souscrire ou de renouveler son contrat qui permettait à la tribu de cumuler des points. Ainsi, plus il y avait de potes au sein d'une tribu, plus la tribu disposait de points. Et ces points accumulés permettaient de réduire la franchise d'un membre du groupe en cas de sinistre.

Cependant, la plateforme a dû mettre la clé sous la porte 18 mois après son lancement, faute de souscriptions.

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2009	France	Achat groupé	Automobile
Fin : 2011			

TABLE 1.2 – Tabeau récapitulatif Generali (Kontournous.fr)

1.3.2 Friendsurance (Allemagne, 2010)

Lancé en **2010** en **Allemagne**, **Friendsurance** est un pionnier de l'assurance collaborative de type **micro-mutuelle collective**, proposant une offre d'assurance sur la **RC**, l'**habitation**, la **protection juridique** et les **biens électroniques**.



Le principe de Friendsurance est de financer les petits sinistres par un pot commun constitué par les membres d'une communauté et de laisser aux assureurs leur rôle traditionnel sur les gros sinistres. Aussi, le courtier d'assurance propose de rembourser une partie de la prime aux assurés : moins il y a de sinistres au sein de la communauté, plus le montant de ce « *cash-back* » sera important, jusqu'à 50% de la prime versée, ce montant augmentant également avec le nombre de membres au sein d'une communauté.

En **2017**, le pionnier de l'assurance P2P dans le monde étend son activité en **Australie** avec l'offre **cash-back bike insurance**, dédiée à la protection des vélos et de leurs conducteurs.

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2010	Allemagne	Micro-mutuelle	Habitation, RC, PJ et Biens
Expansion : 2017	Australie	collective	Vélos et conducteurs

TABLE 1.3 – Tabeau récapitulatif Friendsurance

1.3.3 Bought By Many (Royaume-Uni, 2012)

Lancée en **2012** au **Royaume-Uni**, la startup **Bought By Many** est une plateforme d'assurance collaborative de type **achat groupé**, proposant une offre d'assurance pour les **animaux de compagnie**.



Depuis sa création, l'entreprise a permis à des milliers de clients d'obtenir des couvertures d'assurance sur mesure pour leurs animaux de compagnie.

En 2017, la startup a mis sur pied une offre d'assurance singulière pour animaux dont les polices sont essentiellement basées sur les commentaires de plus de 40.000 propriétaires d'animaux. Cette offre lancée en partenariat avec l'assureur britannique *Great Lakes Insurance SE*, était une innovation sur ce marché d'assurance de niche.

Ce qui a permis à la startup d'obtenir en 2019, le prix « *Most Trusted Pet Insurance Provider* » de meilleur fournisseur d'assurance d'animaux de compagnie aux **Moneywise Customer Service Awards**².

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2012	Royaume-Uni	Achat groupé	Animaux de compagnie

TABLE 1.4 – Tabeau récapitulatif Bought By Many

1.3.4 Guevara (Royaume-Uni, 2014)

Lancée en **2013**, au **Royaume-Uni**, initialement sous le nom de *jFloat*, la startup **Guevara** (renommée en **2014**) était une plate-forme d'assurance collaborative de type **micro-mutuelle collective**, proposant une offre d'assurance **automobile**.



En clair, la communauté constituait un fonds appelé « protection pool » en versant une prime d'entrée incluant une prime pour Guevara. Les années suivantes, les assurés versaient une prime récurrente en fonction de la sinistralité du groupe, l'objectif étant de reconstituer le « protection pool » à son niveau initial. Les assurés pouvaient ainsi bénéficier jusqu'à 80% de réduction s'il n'y avait pas de sinistre. En cas de grand sinistre, si le « protection pool » ne suffisait pas pour indemniser les assurés, Guevara utilisait ce qu'on appellait le fonds Waterloo pour compléter l'indemnisation.

Cependant, l'aventure ne durera que pendant trois années d'activité pour la jeune pousse. Faute de capitaux, la startup ne pouvait plus financer son développement et a fermé ses portes en **2016**.

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2014	Royaume-Uni	Micro-mutuelle	Automobile
Fin : 2016		collective	

TABLE 1.5 – Tabeau récapitulatif Guevara

1.3.5 Inspeer (France, 2014)

Co-fondé par Louis de Broglie et Emmanuelle Mury en **Février 2014**, **Inspeer** est une entreprise pionnière de l'assurance collaborative en **France**. En **Mars 2015**, elle a lancé le premier service de **rachat de franchise** collaboratif.



2. Source : <https://boughtbymany.com>.

La plateforme permettait notamment de mutualiser ses franchises d'assurance **dommage** (automobile, habitation, moto) avec ses proches et ce, sans changer de contrat d'assurance. L'assuré demande à ses proches de contribuer à sa franchise et ce dernier s'engage à contribuer de manière équivalente à la franchise de ceux-ci en cas de sinistre. Les contributions aux franchises ne sont collectées qu'en cas de sinistre et Inspeer prélève une commission fixe de 10% à ce moment.

Ce système permettait donc : 1) pour un assuré, de diminuer le coût de son assurance en augmentant sa franchise d'assurance³ et de faire couvrir sa franchise d'assurance par la communauté des assurés et 2) pour un assureur, de réduire les coûts de gestion des petits sinistres et la fraude.

Par ailleurs, en **2017**, la startup s'était associée à *Generali* pour proposer une formule d'assurance collaborative de type **micro-mutuelle collective**, adaptée aux **véhicules électriques** : location de batterie, couverture des câbles, assistance panne sèche électrique, etc.

Ainsi, une partie des cotisations des assurés était versée à l'assureur pour les gros sinistres alors qu'une autre partie était versée dans un fonds collaboratif dédié aux assurés pour les petits sinistres. Et, s'il y avait moins de sinistres que tarifié initialement par l'assureur, l'argent restant sur le fonds collaboratif était rendu à la fin de l'année aux assurés. Aussi, les assurés pouvaient suivre les sinistres impactant le fonds tout le long de l'année pour suivre leur cagnotte.

Le système avait donc de multiples vertus, cependant la startup a dû déposé le bilan en **Juillet 2018**, faute de fonds suffisants et n'ayant pas trouvé son public sur cette niche dédiée aux véhicules électriques.

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2014	France	Franchise collaborative	Auto, moto, habitation
Générali : 2017		Micro-mutuelle	Véhicules électriques
Fin : 2018			

TABLE 1.6 – Tabeau récapitulatif Inspeer

1.3.6 Lemonade (USA, 2015)

Lancée aux **USA** (New-York) en **Avril 2015** notamment par Daniel Schreiber et Shai Winger, **Lemonade** est une startup de l'assurance collaborative de type **micro-mutuelle collective**, avec une licence d'assureur.



Elle propose une formule d'assurance **habitation**, avec une expérience client 100% digitale via une application, de la souscription au règlement des sinistres.

Lemonade a ainsi déployé un outil de tchat basé sur l'intelligence artificielle (IA), capable de résoudre des demandes complexes. Son robot *Maya* gère aussi de façon instantanée des processus multiples, par exemple : en cas de déménagement, la résiliation de l'ancienne police, l'évaluation du risque à la nouvelle adresse et la fixation d'un prix adapté.

Par ailleurs, la startup se rémunère en prélevant une commission fixe du montant de la prime et reverse à des œuvres caritatives la partie des cotisations qui n'aurait pas été utilisée dans

3. D'après la relation : prime = $\mathbb{E}[\max(\text{coût sinistres} - \text{franchise}, 0)]$.

l'année pour payer des sinistres.

Avec plus de 500.000 assurés, l'entreprise est soutenue par de puissants financiers. En effet, après avoir levé 60 millions de dollars en 2016 auprès d'investisseurs comme *General Catalyst*, *GV (Google Ventures)*, *Thrive Capital* et *Tusk Ventures*, elle a réalisé une nouvelle levée de fonds de 120 millions de dollars, fin 2017 auprès du géant japonais *Softbank Group*. Elle compte également *AXA* et *Allianz* dans son tour de table.

En **2019**, l'entreprise new-yorkaise avait fait un autre tour de table de 300 millions de dollars mené par le japonais *SoftBank*. Ce qui valorisait son capital à plus de 2 milliards de dollars, contre 500 millions de dollars lors de la précédente levée et a favorisé le financement de son expansion à l'Europe, notamment en **Allemagne** en collaboration avec *AXA*.

Depuis **2020**, Lemonade propose une formule d'assurance pour les **animaux de compagnie**, dès une cotisation de 10 \$ par mois. Aussi, l'entreprise fait une réduction de 10% sur cette cotisation, si le client possède également un contrat d'assurance habitation avec elle.

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2015	USA	Micro-mutuelle	Habitation
Expansion : 2019	Allemagne	collective	
Depuis : 2020			Animaux de compagnie

TABLE 1.7 – Tabeau récapitulatif Lemonade

1.3.7 Groupago (Belgique, 2015)

Lancée en **2015**, **Groupago** est la première plateforme d'assurance collaborative **belge d'achat groupé**, qui fournit de l'assurance **automobile** et de l'assurance **habitation**.



La plateforme se base sur le concept de solidarité qui permet à ses clients de profiter d'avantages non négligeables en matière d'assurance.

Concrètement, elle propose des achats groupés pour les formules d'assurance **auto** et **habitation**. Ce qui offre l'opportunité incontestable de s'allier en famille, avec des amis et des milliers d'autres personnes dans le même besoin d'assurance afin de bénéficier des priviléges de groupe et des prix défiant toute concurrence.

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2015	Belgique	Achat groupé	Automobile Habitation

TABLE 1.8 – Tabeau récapitulatif Groupago

1.3.8 Wizzas (France, 2015)

Lancée en **2015**, **Wizzas** est la première plateforme digitale **d'achats groupés** en assurance en **France**. L'assurtech allie la puissance du digital et la force du groupe.



Au sein de communautés, elle définit avec ses clients leurs besoins en assurance et en services associés. Elle se charge ensuite de négocier avec les assureurs traditionnels. Cette approche nouvelle et unique permet à ses clients d'obtenir les meilleures garanties et services aux meilleures conditions.

Wizzas propose cinq formules d'assurance en fonction des communautés : assurance des vélos à assistance électrique commercialisés par *Moustaches Bikes*, assurance responsabilité du dirigeant, assurance pour golfeurs, assurance cyber sécurité, assurance pour EDPM/NVEI⁴.

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2015	France	Achat groupé	Vélo Moustache RC Dirigeant Golfeurs EDPM/VVEI

TABLE 1.9 – Tabeau récapitulatif Wizzas

1.3.9 Otherwise (France, 2016)

Lancée en 2016, l'assurtech française **Otherwise**, spécialisée dans l'assurance collaborative de type micro-mutuelle collective, proposait une formule de mutuelle santé avec le concours de la SAM^a *Thélem assurances*.

a. SAM : Société d'Assurance Mutuelle.



Aussi, une formule d'assurance animaux (chiens, chats) avec l'aide de l'*Equité*, filiale d'assurance dommages de *Generali France* dédiée aux grands comptes.

L'année suivant son lancement (i.e. en 2017), Otherwise avait levé 1,6 millions d'euros. Puis en Juin 2019, la startup avait récidivé et fait encore mieux avec une levée de fonds d'un montant de 2 millions d'euros. Depuis, l'assurtech propose une formule d'assurance automobile ainsi qu'une d'assurance responsabilité civile professionnelle (RC Pro), en partenariat avec la filiale *Altima* de la *MAIF*.

Enfin, notons que selon ses responsables, la jeune pousse a redistribué aux assurés qui ont joué le jeu « entre 7% et 25% des cotisations hors taxe », durant ses deux premières années d'exercice. Et, la quasi-totalité des clients ont eu droit à un bonus, « preuve que ce modèle est vertueux ».

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2016	France	Micro-mutuelle collective	Santé Animaux Automobile RC Pro

TABLE 1.10 – Tabeau récapitulatif Otherwise

4. EDPM : Engins de Déplacement Personnels Motorisés / NVEI : Nouveaux Véhicules Electriques Individuels.

1.3.10 Wecover (France, 2017)

Wecover était une plateforme d'assurance collaborative **française** de type **micro-mutuelle collective**, proposant une offre d'assurance **automobile** 100% digital, avec trois niveaux de garanties (tiers, tiers + et tous risques).



La startup fut lancée en **Février 2017** après une levée de fonds de 500.000 €. L'objectif de Wecover était de proposer à des conducteurs présentant un bon comportement de conduite, une offre qui leur permettait de récupérer une partie de leur cotisation.

Grâce à une application mobile mise à disposition par Wecover, les clients avaient la possibilité de perfectionner leurs habitudes au volant. Regroupés en fonction de leur mode de vie, les clients pouvaient récupérer jusqu'à 25% de leur prime en l'absence d'accident, sous forme de *cashback*.

Cependant, l'aventure s'est terminée en **Octobre 2019** pour la startup. Son portefeuille de clients a été transféré chez un nouveau porteur de risques pendant l'été 2019. En effet, les fonds levés en 2017 ne suffisaient plus pour financer le développement de l'entreprise et les souscriptions n'ont pas suivi.

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2017	France	Micro-mutuelle	Automobile
Fin : 2019		collective	

TABLE 1.11 – Tabeau récapitulatif Wecover

1.3.11 Besure (Canada, 2017)

Besure est une plateforme d'assurance collaborative **canadienne** de type **micro-mutuelle collective**, qui fournit de l'assurance de **dommages** pour certains biens.



Ces biens peuvent être des appareils électroniques, des vélos, des skis et des téléphones intelligents. La plateforme vise à agir comme un « filet de protection sociale pour ceux qui sont laissés de côté par l'industrie de l'assurance ».

Elle affirme offrir la paix de l'esprit pour de petites réclamations, par exemple la perte d'un vélo ou d'une guitare.

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2017	Canada	Micro-mutuelle	Certains biens
		collective	Animaux

TABLE 1.12 – Tabeau récapitulatif Besure

1.3.12 Teambrella (Russie, 2017)

Teambrella est une plateforme d'assurance collaborative **russe** de type **autoassurance** qui utilise la technologie blockchain et la cryptomonnaie. Lancée en **Octobre 2017**, elle est également accessible aux Pays-Bas et en Amérique du Sud.



Elle décrit ses services comme « n'étant pas de l'assurance, quelque chose de mieux ». Les membres s'engagent à former des groupes uniquement pour de l'assurance **automobile**, de l'assurance pour les **animaux** et de l'assurance pour certains **biens personnels** (cellulaires, vélos et ordinateurs). Ils s'engagent à ne pas se regrouper pour de l'assurance de personnes.

Dates	Pays	Modèle d'assurance P2P	Produits offerts
Début : 2017	Russie	Autoassurance	Automobile
	Pays-Bas		Animaux
	Amérique du Sud		Biens personnels

TABLE 1.13 – Tableau récapitulatif Teambrella

1.4 Contexte et problématique

Des quatre modèles rencontrés sur le marché de l'assurance collaborative, notre intérêt se porte sur le modèle de type « franchise collaborative », initié par *Inspeer* en 2014. En effet, la startup ne collectait les contributions (ou cotisations) des assurés à la couverture de franchises de leurs relations qu'en cas de sinistre. Ce qui engendrait indéniablement un risque de contrepartie des assurés. Aussi, la startup a dû déposer le bilan, en 2018, faute de capitaux et de souscriptions suffisantes pour financer son développement.

Par ailleurs, force est de constater que, les couvertures et les contributions des assurés constituant les variables d'intérêt du produit de franchise collaborative, sont fortement dépendantes de la structure du réseau de la communauté des assurés.

C'est dans ce contexte d'analyse de l'impact de la structure du réseau des assurés sur ces variables d'intérêt, que s'inscrit ce mémoire d'actuariat dont l'objectif principal consiste, avec un peu plus de recul depuis 2018, à mettre en place efficacement des mécanismes permettant l'optimisation du produit, c'est-à-dire : garantir la couverture des charges (franchises et dépenses⁵) annuelles d'un nombre maximum d'assurés tout en minimisant les contributions individuelles des assurés à la couverture des charges de leurs relations.

1.5 Méthodologie

Afin de mener à bien cet objectif d'optimisation du produit de franchise collaborative, la démarche suivante sera adoptée.

- ❖ **Premièrement**, le chapitre 2 introduira des notions et des modèles classiques de graphe de la théorie des réseaux, qui permettront de prendre en compte les différents scénarios possibles de connexions entre les membres d'une communauté d'assurés et d'analyser *in fine*, l'impact de la structure du réseau des assurés sur les variables d'intérêt du produit de franchise collaborative.

5. dépense = min(franchise, coût sinistres).

- ❖ **Deuxièmement**, au chapitre 3, une formalisation mathématique du produit de franchise collaborative sera effectuée et une définition de ses variables d'intérêt.
- ❖ **Troisièmement**, sur la base d'une communauté fictive d'assurés, des études d'optimisation seront respectivement effectuées aux chapitres 4 et 5, sur les paramètres du produit de franchise collaborative, à savoir le montant des engagements réciproques et le modèle de graphe social représentant le réseau des connexions entre les assurés de la communauté.
- ❖ **Quatrièmement**, un produit de franchise collaborative sera mis en place, d'une part au chapitre 6, sur la base du produit optimisé dit « scénario central » prenant en entrée les paramètres optimaux issus des études d'optimisation du produit et d'une étude de sensibilité testant ces paramètres selon différentes sensibilités afin d'évaluer leurs impacts sur les variables d'intérêt du produit. Et d'autre part au chapitre 7, sur la base d'applications à des données réelles, notamment des données de généalogie et de réseaux sociaux tels que *Facebook* afin d'analyser comment ces types spécifiques de connexions entre les assurés peuvent influer sur ces variables d'intérêt.

Chapitre 2

Théorie des réseaux

2.1 Introduction

L'objectif de ce chapitre consiste en un rappel et une présentation de notions et terminologie de la théorie des réseaux, nécessaires à la compréhension de la suite du mémoire.

Désignant un ensemble d'éléments interconnectés, la notion de réseau a un caractère polysémique dans le langage courant. En fonction du domaine, cette notion peut être employée pour désigner soit une forme organisationnelle ou structurelle, soit un système de relations inter-individuelles, inter-organisationnelles ou personnelles.

La **théorie des réseaux** regroupe l'ensemble des méthodes et des outils visant à la compréhension des phénomènes d'interactions et de communautés observés dans diverses disciplines telles que : les sciences sociales, l'économie, la finance, la biologie, l'informatique et bien d'autres. Aussi, en fonction des domaines, on rencontre plusieurs types de réseaux tels que : les réseaux sociaux (e.g. Sociogramme de Moreno¹, Figure 2.1), les réseaux de disciplines (e.g. Figure 2.2), les réseaux biologiques (e.g. Réseau métabolique, Figure 2.3), les réseaux de textes (e.g. Figure 2.4), les réseaux de pages (e.g. *World Wide Web*), les réseaux internet, les réseaux de citations, les réseaux de collaborations, les réseaux routiers, les réseaux de communication, etc.

Par ailleurs, les graphes constituent un outil permettant d'une part la formalisation mathématique et d'autre part la visualisation des réseaux. Ce chapitre présente donc dans un premier temps, comment les graphes peuvent être utilisés pour la caractérisation des réseaux. Dans un second temps, un rappel est effectué à propos des caractéristiques (mesures, propriétés et représentations) et des modèles classiques de graphe étudiés dans le cadre de ce mémoire. En effet, ces modèles de graphe permettront de prendre en compte les différents scénarios possibles de connexions entre les membres d'une communauté d'assurés et d'analyser *in fine*, l'impact de la structure du réseau des assurés sur les variables d'intérêt du produit de franchise collaborative à l'étude.

1. Jacob Levy Moreno (1889-1974) est un médecin américain d'origine roumaine, pionnier de la recherche sociométrique selon une approche sociologique des phénomènes sociaux mais également selon une approche psychiatrique en initiant la psychothérapie de groupe (1932). Source : https://fr.wikipedia.org/wiki/Jacob_Levy_Moreno.

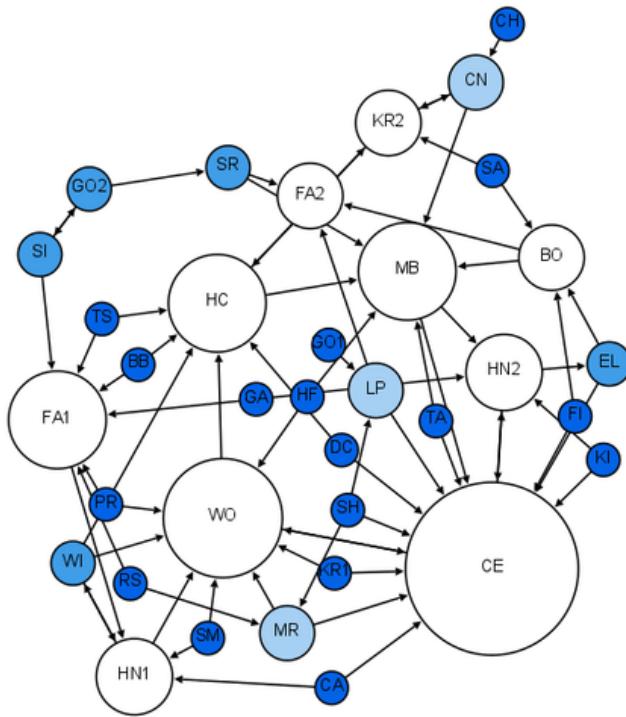


FIGURE 2.1 – Sociogramme de Moreno représentant des relations sociales dans une classe de 1^{ère} année du primaire. Source : Grandjean [2015].

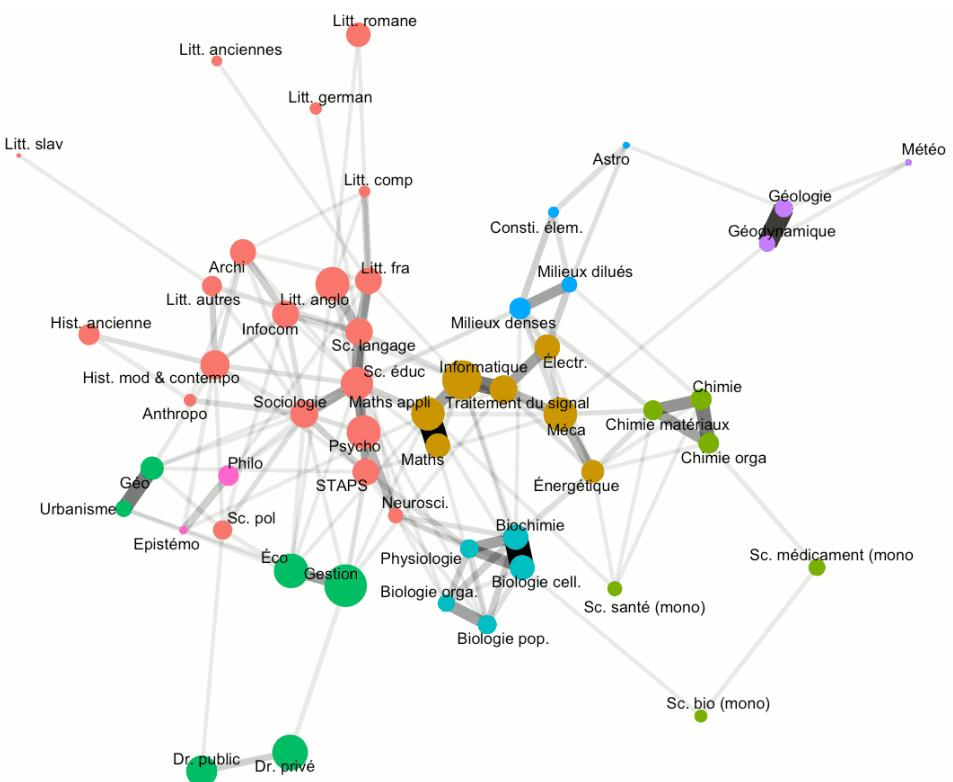


FIGURE 2.2 – Réseau de disciplines. Les couleurs représentent les disciplines ayant plutôt tendance à se partager des postes. Source : Coulmont [2020].

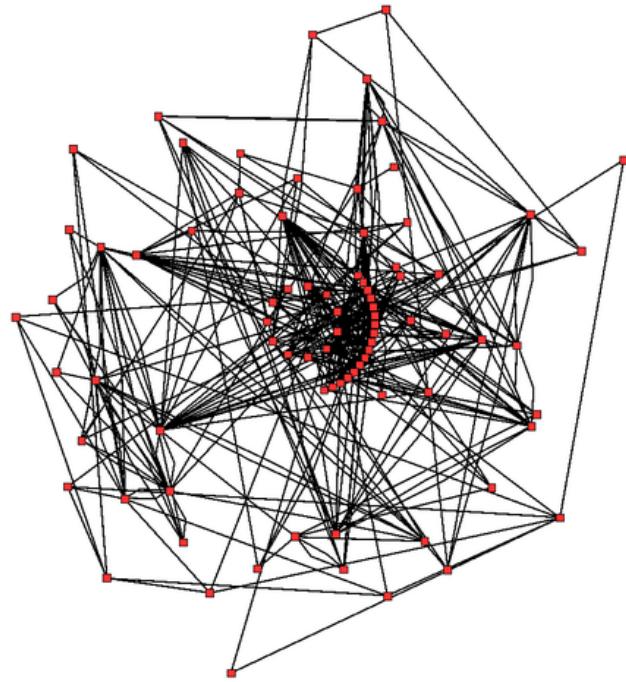


FIGURE 2.3 – Réseau métabolique montrant les liens entre 43 enzymes et 40 métabolites qui interagissent avec le cycle de Krebs d'Arabidopsis. Source : Vickers [2007].

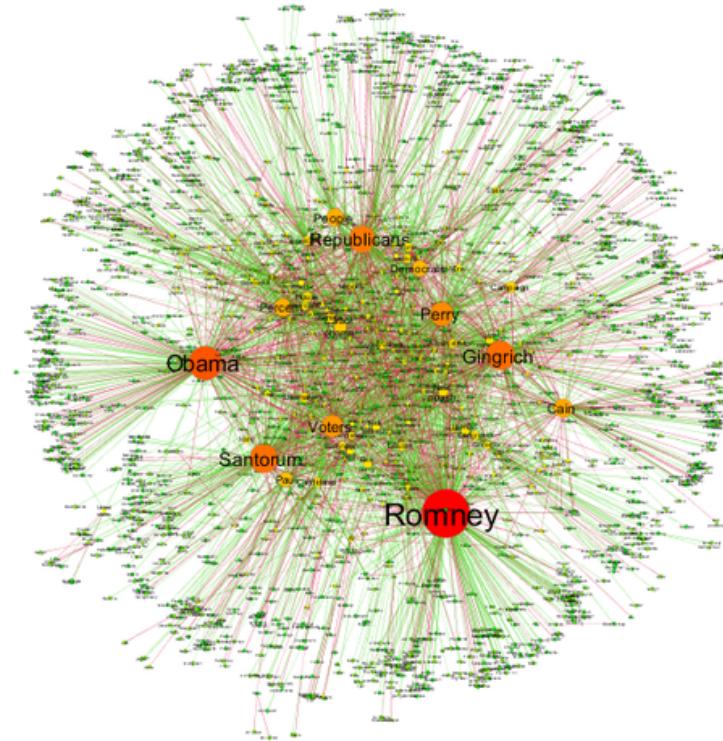


FIGURE 2.4 – Réseau textuel autour des élections américaines de 2012. Les nœuds indiquent les textes (noms des candidats et des électeurs), les liens vont d'un sujet à l'autre, la couleur exprime une relation de soutien ou d'opposition. Source : Sudhahar et al. [2015].

2.2 Réseaux et graphes

L'étude des réseaux a émergé dans diverses disciplines afin de permettre l'analyse des données relationnelles complexes. Le plus ancien document connu dans le domaine de l'étude des graphes constitue les travaux de Euler [1735], concernant le problème des sept ponts de Königsberg (cf. Figure 2.5, source : Wikipédia²). La description mathématique par Euler des sommets et des arêtes fut le fondement de la théorie des graphes, une branche des mathématiques qui étudie les propriétés des relations dyadiques (i.e. de pair à pair) d'une structure en réseau.

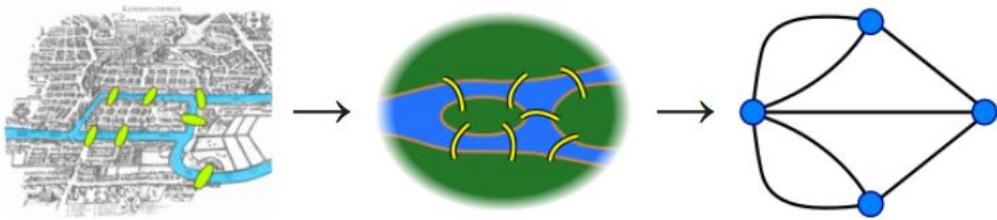


FIGURE 2.5 – Visualisations du problème des sept ponts de Königsberg résolu par Euler : Recherche d'un chemin passant une fois et une seule par chacun des sept ponts de la ville.

Intéressons-nous dans un premier temps, à la définition d'un graphe comme objet mathématique permettant de modéliser un réseau.

Définition 1 (Graphe fini, ordre et taille d'un graphe).

- ❖ Un **graphe fini** non orienté (resp. orienté) G est la donnée d'un couple $G = (S, A)$ tel que :
 - ✓ $S = \{s_1, s_2, \dots, s_n\}$ est un ensemble fini dont les éléments sont appelés **sommets** ou **nœuds** ;
 - ✓ $A = \{a_1, a_2, \dots, a_m\}$ est un ensemble fini de couples non ordonnés (resp. ordonnés) de sommets dont les éléments sont appelés **arêtes** (resp. **arcs**).
- ❖ On appelle **ordre** du graphe G , le nombre d'éléments n de l'ensemble S .
- ❖ On appelle **taille** du graphe G , le nombre d'éléments m de l'ensemble A .

Dans la suite du mémoire, sauf mention contraire, on s'intéressera essentiellement aux graphes non orientés. En effet, dans le chapitre 3, le graphe social représentant le réseau des connexions entre les membres de la communauté des assurés, sera modélisé par un graphe non orienté où deux assurés en connexion sont reliés par une arête.

Introduisons dans un second temps, les notions de graphe partiel et de sous-graphe associées à un graphe.

Définition 2 (Graphe partiel et sous-graphe d'un graphe).

Soit $G = (S, A)$ un graphe non orienté.

- ❖ Le graphe $G' = (S, A')$ est un **graphe partiel** de G , si A' est inclus dans A . Autrement dit, on obtient G' en enlevant une ou plusieurs arêtes au graphe G .
- ❖ Pour un sous-ensemble de sommets S_0 inclus dans S , le **sous-graphe** de G induit par S_0 , est le graphe $G|_{S_0} = (S_0, A(S_0))$ dont l'ensemble des sommets est S_0 et l'ensemble des arêtes $A(S_0)$ est formé de toutes les arêtes de G ayant leurs deux extrémités dans S_0 . Autrement dit, on obtient $G|_{S_0}$ en enlevant un ou plusieurs sommets au graphe G , ainsi que toutes les arêtes incidentes à ces sommets.

2. Source : https://fr.wikipedia.org/wiki/Theorie_des_graphes.

Les notions basiques de graphe étant introduites, la section suivante est consacrée au rappel de quelques notions permettant la caractérisation des graphes telles que des notions de mesures, de propriétés et de représentations des graphes.

2.3 Caractéristiques des graphes

Cette section consiste à introduire quelques notions évoquées dans la suite du mémoire, permettant de caractériser les graphes, notamment des notions associées aux mesures, aux propriétés et aux représentations des graphes.

Tout d'abord, on énonce les définitions de quelques mesures permettant d'analyser les propriétés et les caractéristiques d'un graphe, notamment : le degré moyen, le coefficient de clustering et la longueur de chemin moyenne d'un graphe.

Définition 3 (Degré moyen, coefficient de clustering, longueur de chemin).

Soit $G = (S, A)$ un graphe non orienté d'ordre n et de taille m .

- ❖ Le **degré** d_s d'un sommet (ou noeud) $s \in S$ du graphe G correspond au nombre d'arêtes qui lui sont connectées et prend les valeurs de 0 à $n - 1$ (en supposant qu'un sommet n'est pas connecté à lui même par une arête).
- ❖ Le **degré moyen** \bar{d} du graphe G correspond au nombre moyen d'arêtes connectées à un sommet quelconque $s \in S$ et vérifie :

$$\bar{d} = \frac{2m}{n} \quad (2.1)$$

- ❖ Le **coefficient de clustering** (ou coefficient de regroupement) du graphe G est une mesure du regroupement des sommets du graphe G . Plus précisément, ce coefficient est la probabilité que deux noeuds soient connectés sachant qu'ils ont un noeud voisin en commun.
Un coefficient de clustering élevé pour un réseau est une caractéristique de la propriété de "**petit monde**" ou de "**cohésion sociale**".
- ❖ Un **chemin** entre deux sommets du graphe G correspond à une collection d'arêtes reliant ces deux noeuds.
- ❖ La **longueur d'un chemin** entre deux sommets du graphe G correspond au nombre d'arêtes composant le chemin entre ces deux noeuds.
- ❖ Le **plus court chemin** entre deux sommets du graphe G correspond au chemin de longueur minimale reliant ces deux noeuds.
- ❖ La **longueur de chemin moyenne** du graphe G correspond à la valeur moyenne des longueurs des plus courts chemins entre toutes les paires de noeuds du graphe.
Autrement dit, cette valeur représente en moyenne, le nombre d'étapes (ou d'arêtes) à franchir pour passer d'un noeud du graphe G à un autre.
On dit que le graphe G est un **petit monde**, si sa longueur de chemin moyenne est une fonction logarithmique de son ordre n .

Les définitions des mesures essentielles ayant été rappelées, énonçons maintenant quelques propriétés relatives à des graphes particuliers.

Définition 4 (Graphe simple, multi-graphe, graphe régulier et complet, clique).

Soit $G = (S, A)$ un graphe non orienté d'ordre n .

- ❖ Une **boucle** est une arête reliant un sommet $s \in S$ du graphe G à lui-même.

- ❖ Le graphe G est dit **simple** s'il ne comporte pas de boucle, et s'il ne comporte pas plus d'une arête entre deux sommets $s_i, s_j \in S$, avec $i, j = 1 : n$.

- ❖ Le graphe G est un **multi-graphe** s'il n'est pas simple. Dans ce cas, A n'est plus un ensemble mais un multi-ensemble d'arêtes.

Dans la suite du mémoire, on se restreindra aux graphes simples sachant qu'un assuré ne peut être en relation avec lui-même (i.e. pas de boucles) et deux assurés sont soit en relation ou pas (i.e. pas de relations multiples entre deux assurés).

- ❖ Le graphe G est dit **régulier** si tous les sommets de S ont le même degré. Si ce degré est k , on dit que le graphe G est k -régulier.

- ❖ Supposons en plus que le graphe G est simple.

Alors G est dit **complet** s'il comporte une arête $\{s_i, s_j\} \in A$ pour toute paire de sommets différents $s_i, s_j \in S$, avec $i, j = 1 : n$. On note K_n un graphe complet d'ordre n .

En particulier, le graphe K_n est $(n - 1)$ -régulier.

- ❖ Soit $S_0 \subset S$ un sous-ensemble non vide de sommets de S .

Le sous-graphe $G|_{S_0} = (S_0, A(S_0))$ de G induit par S_0 est une **clique** s'il est complet, c'est-à-dire, si tous les sommets de S_0 sont connectés deux à deux.

Enfin, intéressons-nous à la représentation des graphes. Il existe dans la littérature plusieurs outils permettant de représenter un graphe. Dans ce mémoire, on évoque uniquement la caractérisation d'un graphe par sa matrice d'adjacence.

Définition 5 (Matrice d'adjacence d'un graphe).

Soit $G = (S, A)$ un graphe non orienté d'ordre n . On suppose sans perte de généralité que les sommets de S sont numérotés de 1 à n , c'est-à-dire, $S = \{1, 2, \dots, n\}$.

La matrice d'adjacence R du graphe G correspond à une matrice booléenne de taille $n \times n$ telle que :

$$R_{ij} = \begin{cases} 1 & , \text{ si } \{i, j\} \in A \\ 0 & , \text{ sinon.} \end{cases}$$

Remarques : Partant de la définition de la matrice d'adjacence R du graphe G , on a :

- ✓ Le degré d_i d'un sommet $i \in S$ s'écrit : $d_i = \sum_{j=1}^n R_{ij}$;
- ✓ Le degré moyen \bar{d} du graphe G s'écrit : $\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n R_{ij}$.

Les caractéristiques des graphes ayant été évoquées, la section suivante est dédiée à la présentation des modèles classiques de graphe de la théorie des réseaux, étudiés dans le cadre de ce mémoire.

2.4 Modèles classiques de graphe

Ce mémoire portant sur le produit de franchise collaborative dont l'un des paramètres constitue le graphe social, représentant le réseau des connexions entre les assurés ; afin de prendre en compte les différents scénarios possibles de connexions entre les assurés, on s'intéresse à quatre modèles classiques de graphe de la théorie des réseaux : le **modèle à clusters homogènes**, le **modèle d'Erdős-Rényi**, le **modèle du petit monde de Watts-Strogatz** et le **modèle d'attachement préférentiel de Barabási-Albert**.

2.4.1 Le modèle à clusters homogènes

On désigne par "cluster homogène" un graphe régulier, c'est-à-dire un graphe dont tous les sommets ont le même degré. Sachant qu'un graphe complet est en particulier régulier, dans la suite du mémoire, le modèle à clusters homogènes simulant le graphe social de la communauté des assurés sera calibré via des cliques (i.e. des sous-graphes complets). Par exemple, sur la Figure 2.6, le graphe social de la communauté d'ordre 12 est composé de trois cliques (ou clusters homogènes) de tailles³ respectives 3, 4 et 5.

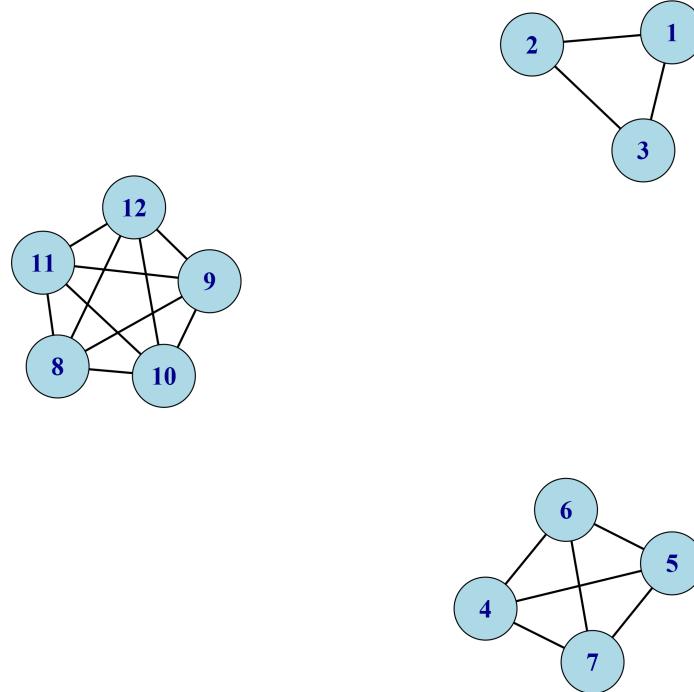


FIGURE 2.6 – Graphe social : Exemple clusters homogènes

On constate que le cluster ou clique de taille 3 est 2-régulier, c'est-à-dire chaque membre du cluster possède 2 relations. Aussi, le cluster de taille 4 est 3-régulier, c'est-à-dire chaque membre du cluster possède 3 relations. Enfin, le cluster de taille 5 est 4-régulier, c'est-à-dire chaque membre du cluster possède 4 relations.

Par ailleurs, la calibration du graphe social via le modèle à clusters homogènes permet de prendre en compte les scénarios où les assurés se connecteraient par groupes indépendants (e.g. entre amis, familles ou personnes ayant le même besoin d'assurance). Cette calibration effectuée sous le logiciel R et présentée dans la section 5.2, permet d'obtenir une distribution de degrés des assurés homogène par clusters.

Aussi, dans le cas où la connexion entre les assurés serait libre et sans contrainte, il convient de prendre en compte tous les scénarios de connexions possibles. Pour ce faire, on s'intéresse à trois modèles de génération de graphe aléatoire, notamment : les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert.

3. Dans la suite du mémoire, la taille d'un cluster est assimilée à l'ordre de ce cluster, c'est-à-dire au nombre de sommets composant ce cluster.

2.4.2 Le modèle d'Erdős-Rényi

Le modèle d'Erdős-Rényi est parfois appelé "le" graphe aléatoire au sens où, dans ce modèle, la topologie du réseau est tirée uniformément au hasard. Ce modèle permet donc de simuler aléatoirement un graphe non orienté et simple. Dans la littérature, le modèle d'Erdős-Rényi fait référence à deux modèles formellement différents, mais étroitement liés asymptotiquement : le modèle uniforme à nombre d'arêtes m fixé, noté $G(n, m)$ et le modèle binomial où la probabilité p de connexion entre deux noeuds quelconques est fixée, noté $G(n, p)$, avec n l'ordre du graphe.

- ❖ **Modèle uniforme :** Dans ce modèle, le graphe aléatoire $G(n, m)$ est tiré uniformément dans l'ensemble des graphes non orientés et simples d'ordre n (i.e. à n noeuds) et de taille m (i.e. à m arêtes). Le choix d'un tel graphe correspond au choix de ses paires de noeuds connectés, or on dénombre $N = \binom{n}{2} = \frac{n(n-1)}{2}$ arêtes possibles dans un graphe d'ordre n . Le nombre possible de graphes d'ordre n et de taille m est donc $\binom{N}{m}$. Ainsi, la probabilité d'obtenir (uniformément) un graphe G dans l'ensemble des graphes $G(n, m)$ d'ordre n et de taille m est :

$$\mathbb{P}(G(n, m) = G) = \frac{1}{\binom{N}{m}}$$

La probabilité qu'une arête soit présente est identique pour toute paire de noeuds et vaut $\frac{m}{N}$, comme rappelée dans Channarond [2013]. Ce modèle a été introduit par Paul Erdős et Alfréd Rényi en 1959 et a fait l'objet de plusieurs articles de 1959 à 1968, dont l'article fondateur Erdős and Rényi [1959], ou encore Erdős and Rényi [1961a] et Erdős and Rényi [1961b].

Ce modèle théorique a été très étudié du point de vue des Probabilités, mais il reste peu pratique à utiliser puisque généralement dans les applications le nombre d'arêtes m n'est pas déterministe.

- ❖ **Modèle binomial :** Dans ce modèle noté $G(n, p)$, on tire pour tous noeuds $i, j \in S = \{1, 2, \dots, n\}$, tels que $i < j$, une variable de Bernoulli $R_{ij} \sim \mathcal{B}(p)$ de paramètre $p \in]0, 1[$, les variables étant mutuellement indépendantes. Le graphe aléatoire $G(n, p)$ est représenté par la matrice d'adjacence symétrique $R = (R_{ij})_{\{i,j \in S\}}$. Chacune des $N = \binom{n}{2} = \frac{n(n-1)}{2}$ arêtes possibles est présente avec probabilité p , absente avec probabilité $1 - p$, cela indépendamment du statut des autres arêtes. Ainsi, le nombre m_p d'arêtes de $G(n, p)$ prenant les valeurs de 0 à N , suit la loi binomiale $\mathcal{B}(N, p)$ de paramètres N et p . De même, la loi des degrés des noeuds est binomiale. En effet, pour tout noeud $i \in S$, on a :

$$d_i = \sum_{\substack{j=1 \\ j \neq i}}^n R_{ij} \sim \mathcal{B}(n-1, p) \quad (2.2)$$

Le modèle $G(n, p)$ avec $p = 0,5$ a été en particulier discuté dans Erdős [1947].

C'est au modèle binomial qu'on donne communément aujourd'hui le nom d'Erdős-Rényi et c'est ce modèle qui sera étudié dans la suite du mémoire. La Figure 2.7 donne un exemple des étapes pour la génération d'un graphe aléatoire via le modèle d'Erdős-Rényi.

Par ailleurs, la calibration du graphe social de la communauté des assurés via le modèle d'Erdős-Rényi effectuée sous le logiciel R et présentée dans la section 5.2, permet d'obtenir une distribution de degrés des assurés relativement homogène. Aussi, pour la calibration, le degré moyen \bar{d} de la communauté des assurés étant fixée, la probabilité p de connexion entre deux assurés est déterminée d'après la relation 2.2, comme suit : \forall assuré $i \in S$,

$$\begin{aligned}\mathbb{E}(d_i) = \bar{d} &\iff (n-1) \times p \stackrel{\text{(d'après (2.2))}}{=} \bar{d} \\ &\iff p = \frac{\bar{d}}{n-1}\end{aligned}\tag{2.3}$$

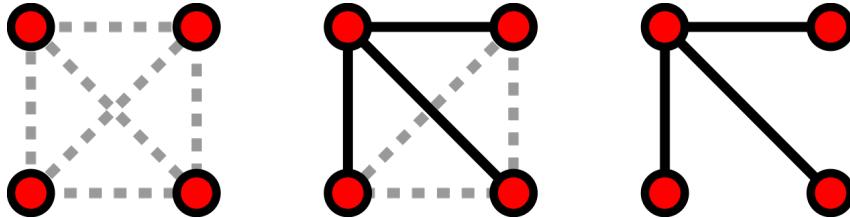


FIGURE 2.7 – Illustration de la génération d'un graphe aléatoire $G(n, p)$ par le modèle d'Erdős–Rényi. Pour chaque arête du graphe complet formé par les $n = 4$ noeuds, un nombre aléatoire est généré et comparé à la probabilité p donnée. Si le nombre aléatoire est inférieur à p , une arête est formée sur le modèle. Source : Jmcatania [2011].

2.4.3 Le modèle du petit monde de Watts-Strogatz

L'étude des graphes aléatoires remonte aux travaux de Erdős and Rényi [1959] sur le modèle portant leur nom, qui fournit un modèle simple avec beaucoup d'applications. Cependant, ce modèle ne vérifie pas la propriété de **petit monde** (cf. Définition 3) typique des réseaux réels. Par exemple, les réseaux sociaux possèdent la propriété de petit monde : dans la majorité des cas, deux noeuds (e.g. deux assurés) peuvent être reliés par un très petit nombre d'amis intermédiaires.

Afin de prendre en compte la propriété de petit monde, caractéristique des réseaux réels, pour la calibration du graphe social de la communauté des assurés, il importe de s'intéresser au modèle de génération de graphe aléatoire proposé par Watts and Strogatz [1998], via un algorithme permettant de capter la propriété de petit monde.

Concrètement, étant donnés le nombre n de noeuds, le degré moyen $\bar{d} = 2k$ supposé pair et vérifiant la relation $n \gg \bar{d} \gg \ln(n) \gg 1$, ainsi qu'un paramètre $p \in]0, 1[$, l'algorithme construit un graphe non orienté à $m = \frac{n\bar{d}}{2}$ arêtes (d'après la relation 2.1) de la façon suivante :

- ❖ **Étape 1 :** On part d'un graphe \bar{d} -régulier : chacun des n noeuds est connecté à \bar{d} voisins, en raison de $k = \bar{d}/2$ de chaque côté ;
- ❖ **Étape 2 :** Pour chaque noeud et chacune de ses arêtes, on reconnecte chaque arête en changeant sa destination avec probabilité p , en évitant de dupliquer des arêtes existantes (i.e. arêtes multiples) et de relier un noeud à lui-même (i.e. boucle).

La Figure 2.8 illustre l'impact de la probabilité de reconnexion $p \in [0, 1]$ sur le processus de reconnexion des arêtes partant d'un graphe régulier (*regular graph* : $p = 0$) à un graphe petit monde de Watts-Strogatz (*small-world graph* : $p \in]0, 1[$) ou à un graphe aléatoire d'Erdős–Rényi (*random graph* : $p = 1$).

Par ailleurs, la calibration du graphe social de la communauté des assurés via le modèle de Watts-Strogatz effectuée sous le logiciel R et présentée dans la section 5.2, permet d'obtenir une distribution des degrés des assurés plus homogène que celle obtenue via le modèle d'Erdős–Rényi.

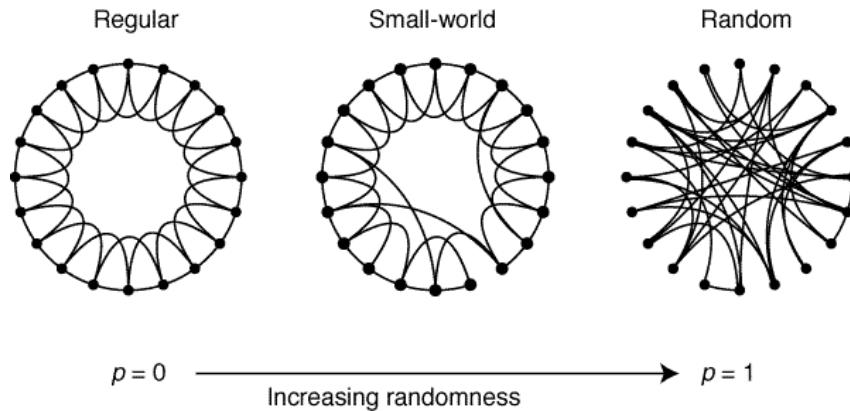


FIGURE 2.8 – Illustration du processus de reconnexion des arêtes en fonction de la probabilité de reconnexion $p \in [0, 1]$: Graphe régulier de départ (*Regular* : $p = 0$), Modèle de Watts-Strogatz (*Small-world* : $p \in]0, 1[$) et Modèle d'Erdős-Rényi (*Random* : $p=1$). Source : Watts and Strogatz [1998].

2.4.4 Le modèle d'attachement préférentiel de Barabási-Albert

La distribution de degrés issue du modèle d'Erdős-Rényi converge vers une loi de Poisson⁴, plutôt que vers une loi de puissance observée dans les réseaux réels **sans échelle**⁵. La plupart des systèmes naturels ou humains observés, tel que l'internet, le *World Wide Web*, les réseaux de citations et certains réseaux sociaux font partie de la catégorie des réseaux sans échelle (au moins approximativement).

Afin de prendre en compte la propriété de réseau sans échelle, caractéristique des réseaux réels, pour la calibration du graphe social de la communauté des assurés, il importe de s'intéresser au modèle de génération de graphe aléatoire proposé par Barabási and Albert [1999], via un algorithme permettant de capter la propriété de réseau sans échelle, à l'aide d'un mécanisme **d'attachement préférentiel**.

En l'occurrence, le principe d'attachement préférentiel consiste à relier le nouveau noeud de préférence aux noeuds déjà les plus populaires (i.e. ceux de plus grand degré) expliquant ainsi la formation de quelques noeuds concentrateurs avec un degré inhabituellement élevé par rapport aux autres noeuds du réseau et la loi de puissance des degrés. Par exemple, dans le *World Wide Web*, les nouvelles pages se lient préférentiellement à de tels noeuds concentrateurs, c'est-à-dire des sites très bien connus tels que *Google*, plutôt que vers les pages moins connues du public. Concrètement, l'algorithme proposé par Barabási and Albert [1999] fonctionne comme suit :

- ❖ **Étape 1 :** Le réseau commence par un premier réseau complètement connecté de m_0 noeuds ;
- ❖ **Étape 2 :** De nouveaux noeuds sont ajoutés au réseau. Chaque noeud est connecté à $m \leq m_0$ noeuds existants, avec une probabilité proportionnelle au nombre de liens que ces noeuds existants ont déjà. Formellement, la probabilité p_i qu'un nouveau noeud se connecte à un

4. Approximation de la loi binomiale par la loi de Poisson : si X est une variable aléatoire suivant la loi binomiale $\mathcal{B}(n, p)$ avec $n \geq 30$, $p \leq 0,1$ et $np \leq 15$, on peut approximer la loi de X par la loi de Poisson de paramètre np .

5. Un réseau sans échelle (ou réseau invariant d'échelle, ou encore *scale-free network* en anglais) est un réseau dont les degrés suivent une loi de puissance. c'est-à-dire la proportion $P(k)$ d'assurés de degré k est telle que : $P(k) \sim k^{-\gamma}$, où le paramètre $\gamma > 0$ est appelé exposant d'invariance d'échelle.

nœud $i \in S$ existant est :

$$p_i = \frac{d_i}{\sum_{j \in S} d_j},$$

où d_i représente de degré du nœud $i \in S$ et la somme est effectuée sur tous les nœuds existants $j \in S$.

La Figure 2.9 représente trois graphes de même ordre générés via le modèle de Barabási-Albert et l'impact du paramètre d'attachement m sur la distribution de degrés des nœuds de chaque graphe.

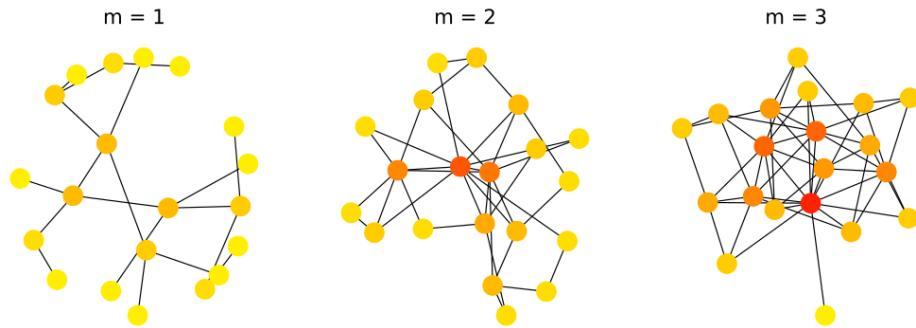


FIGURE 2.9 – Impact du paramètre d'attachement m sur la distribution de degrés des graphes générés via le modèle de Barabási-Albert. Chaque graphe possède 20 nœuds et le paramètre d'attachement m indiqué. La couleur de chaque noeud est déterminée par son degré (une couleur plus foncée désigne un nœud plus populaire). Source : HeMath [2017].

Par ailleurs, la calibration du graphe social de la communauté des assurés via le modèle de Barabási-Albert effectuée sous le logiciel R et présentée dans la section 5.2, permet d'obtenir une distribution de degrés relativement hétérogène et de loi de puissance, contrairement à la distribution de degrés relativement homogène et de loi de Poisson obtenue via le modèle d'Erdős-Rényi.

La Figure 2.10 illustre un exemple de fonctions de survie des distributions de degrés issues de réseaux obtenus respectivement via les modèles d'Erdős-Rényi et de Barabási-Albert. On constate, d'après la Figure 2.10 (*Power distribution*), que le réseau obtenu via le modèle de Barabási-Albert est sans échelle, c'est-à-dire la distribution de degrés est de loi de puissance. En effet, on montre que le logarithme de la fonction de survie des degrés d'un réseau sans échelle est linéaire en fonction du logarithme des degrés (cf. Réseau sans échelle et loi de puissance en Annexe B).

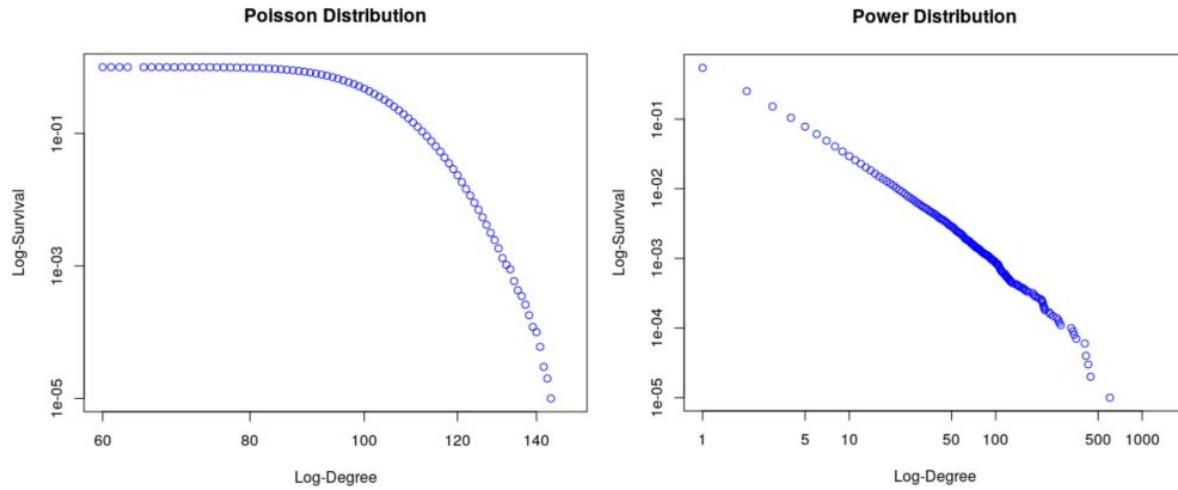


FIGURE 2.10 – Comparaison des fonctions de survie des distributions de degrés issues des modèles d’Erdős–Rényi et de Barabási-Albert en échelle log-log : Modèle d’Erdős–Rényi (*Poisson distribution*), Modèle de Barabási-Albert (*Power distribution*). Source : Charpentier [2017].

2.5 Conclusion

Ce chapitre a consisté d’une part, à présenter des notions de la théorie des réseaux abordées dans le cadre de ce mémoire et à introduire les graphes qui constituent un outil mathématique essentiel non seulement pour la formalisation et la caractérisation, mais également pour la représentation des réseaux. D’autre part, quatre modèles classiques de graphe ont été présentés, notamment le modèle à **clusters homogènes**, le modèle d’**Erdős-Rényi**, le modèle du petit monde de **Watts-Strogatz** et le modèle d’attachement préférentiel de **Barabási-Albert**. En effet, ces modèles de graphe serviront dans les chapitres 5 et 6, pour la calibration du graphe social représentant le réseau des connexions entre les membres de la communauté des assurés.

Synthèse de la partie I

Cette première partie du mémoire a permis de présenter au chapitre 1, le marché mondial de l'assurance collaborative, notamment l'état de l'art, les différents modèles (dont le modèle de franchise collaborative à l'étude) et les principaux acteurs historiques du marché. Aussi, de définir le contexte de notre étude, la problématique actuarielle traitée et la méthodologie adoptée. Enfin, le chapitre 2 a consisté à introduire des modèles classiques de graphe de la théorie des réseaux, tels que les modèles à **clusters homogènes**, d'**Erdős-Rényi**, de **Watts-Strogatz** et de **Barabási-Albert**. En effet, ces modèles de graphe permettront de prendre en compte les différents scénarios possibles de connexions entre les membres d'une communauté d'assurés et d'analyser *in fine*, l'impact de la structure du réseau des assurés sur les variables d'intérêt du produit de franchise collaborative.

L'objectif principal du mémoire consistant à mettre en place efficacement des mécanismes permettant l'optimisation du produit de franchise collaborative, la deuxième partie du mémoire consistera dans un premier temps, au chapitre 3, à formaliser mathématiquement et à définir les variables d'intérêt du produit. Et dans un second temps, aux chapitres 4 et 5, à optimiser respectivement les paramètres du produit en cherchant à comprendre : 1) comment fixer le montant des engagements réciproques entre les assurés et 2) quel(s) modèle(s) de graphe social permet(tent) d'optimiser le produit de franchise collaborative.

Deuxième partie

Optimisation du produit de franchise collaborative

Chapitre 3

Introduction au produit de franchise collaborative

Ce chapitre fait tout d'abord, une présentation de termes de l'assurance en général et du produit de franchise collaborative en particulier, utiles à la compréhension de la suite du mémoire. Ensuite, une introduction et une formalisation mathématique du produit de franchise collaborative sont effectuées. Enfin, il (ce chapitre) s'achève par une énumération de pistes de réflexions dans l'objectif d'optimiser le produit de franchise collaborative. Par ailleurs, il s'appuie essentiellement sur des notions de graphes évoquées dans le chapitre 2 dédié à la [Théorie des réseaux](#).

3.1 Vocabulaire et terminologie

Cette section fait dans un premier temps, un rappel de termes classiques de l'assurance utiles à la compréhension de la suite du mémoire et dans un second temps, une introduction de termes spécifiques au produit de franchise collaborative à l'étude dans le cadre de ce mémoire.

3.1.1 Vocabulaire de l'assurance

Cette sous-section consiste à rappeler quelques notions classiques de l'assurance nécessaires pour la présentation du produit spécifique d'assurance à l'étude.

L'assurance

L'assurance peut se définir comme un système qui permet de prémunir une partie (l'**assuré**), moyennant le paiement d'une **prime** ou **cotisation**, contre les conséquences financières et économiques liées à la survenance d'un **risque** (événement aléatoire) particulier, par le versement en cas de réalisation du risque (**sinistre**), d'une **prestation** de la part d'une autre partie (l'**assureur**).

Le risque

C'est un événement dommageable (**futur**, **aléatoire** et **involontaire**) contre lequel l'assuré cherche à se prémunir : l'incendie, le vol, le dégât des eaux, l'accident, la maladie, le décès, etc.

La prime ou cotisation

La prime est la contribution que verse le **souscripteur** (ou l'**assuré**) à l'**assureur** en échange de la garantie qui lui est accordée de se voir indemnisé (selon les conditions contractuelles) en cas de la réalisation du risque pour lequel il s'est assuré. Elle est payable à l'avance et son montant

est fonction de plusieurs paramètres : le niveau de garantie souhaité, l'objet assuré, la probabilité de survenance du risque.

La prestation

C'est le versement d'une somme d'argent par l'assureur au **bénéficiaire** (soit l'assuré, le souscripteur, un tiers ou l'ayant droit) en cas de réalisation du risque.

La franchise

C'est la **somme restant à la charge** de l'assuré (donc non indemnisée par l'assureur) en cas de réalisation du risque.

Les franchises ont différents rôles en fonction de l'objectif visé par l'assureur. Elles permettent de :

- ❖ Supprimer les **petits sinistres** qui coûtent plus cher en frais de gestion qu'en indemnisation ;
- ❖ Diminuer l'**aléa moral** et responsabiliser les assurés, en les faisant participer à la prise en charge du coût de leur sinistre ;
- ❖ Diminuer le risque de ruine de l'assureur, en se protégeant contre les **sinistres exceptionnels** (gros sinistres).

Il existe plusieurs formules de franchise sur le marché dont :

- ❖ **Franchise proportionnelle** : L'assuré conserve à sa charge une part proportionnelle du coût de sinistre.

Cette formule réduit l'**aléa moral** en intéressant l'assuré à la non réalisation du risque, mais n'empêche pas la **gestion des petits sinistres** et le **paiement des gros sinistres**.
Cette franchise se retrouve sur les **frais de santé**.

- ❖ **Franchise relative ou atteinte** : L'assureur n'intervient pas pour les sinistres dont le coût est inférieur à la franchise et indemnise totalement les sinistres qui dépassent le montant de la franchise.

Cette formule supprime la **gestion des petits sinistres**, mais augmente l'**aléa moral** et le **risque de fraude**, puisque l'assuré aura tendance à majorer le coût des petits sinistres afin de se les voir totalement pris en charge.

Cette franchise se retrouve sur les garanties **accidents corporels**.

- ❖ **Franchise absolue ou déduite** : L'assureur n'intervient pas pour les sinistres dont le coût est inférieur à la franchise et paie le surplus du montant indemnisable au-delà de la franchise.

Cette formule supprime la **gestion des petits sinistres**, réduit en partie l'**aléa moral**, mais n'empêche pas le **paiement des gros sinistres**.

Cette franchise se retrouve le plus couramment sur les contrats d'**assurance automobile**.

- ❖ **Franchise plafond** : L'assureur intervient pour les sinistres dont le coût est inférieur à la franchise (**plafond**) et le surplus du coût de sinistre reste à la charge de l'assuré.

Cette formule n'empêche pas la **gestion des petits sinistres**, diminue la **valeur client**, mais réduit l'**aléa moral** et la prime de l'assuré, aussi limite le **paiement des gros sinistres**.

Cette franchise se retrouve systématiquement sur les contrats **RC** et **MRH**.

La mutualisation

C'est l'un des principes fondateurs de l'assurance, qui consiste à utiliser les **primes de la communauté des assurés** pour payer les **sinistres de la communauté des assurés** au titre du même exercice.

On parle aussi de gestion des primes par **répartition**, mode de gestion collectif des **assurances non vie**¹, en opposition au mode de gestion par **capitalisation** (mode de gestion individuel où les primes de l'assuré servent à lui délivrer une prestation au moment de la survenance du risque), utilisé dans les **assurances vie**².

3.1.2 Vocabulaire spécifique au produit de franchise collaborative

Pour rappel, le produit de franchise collaborative constitue l'objet de cette étude. Cette sous-section fait donc, une introduction de termes spécifiques à ce produit d'assurance.

Franchise collaborative

C'est le modèle d'assurance collaboratif distribué sur une plateforme (généralement en ligne), consistant pour un assuré à mutualiser sa franchise d'assurance dommage (auto, habitation, moto) avec ses proches. En d'autres termes, à faire couvrir sa franchise par ses proches, en cas de sinistre et vice-versa.

Communauté des assurés

La communauté des assurés correspond à l'ensemble des assurés ayant les mêmes besoins d'assurance et ayant souscrit un contrat via la plateforme distribuant un produit de franchise collaborative.

Relation

Une relation existe entre deux assurés de la plateforme s'ils sont en lien direct au sein de la communauté des assurés. On dit aussi qu'ils sont **amis**.

Graphe social

Le graphe social représente le graphe du réseau des relations entre les membres de la communauté des assurés.

Degré

Le degré d'un assuré représente son nombre de relations dans la communauté des assurés.

Dépense individuelle

La dépense individuelle associée au sinistre d'un assuré est le minimum entre le montant de sa franchise et le coût du sinistre. Pour des raisons de simplification, on suppose dans ce mémoire,

1. Les assurances « non vie » : assurances de biens, assurances de responsabilité et assurances santé.
2. Les assurances « vie » : vie, décès, épargne, retraite, ...

qu'un assuré ne peut avoir qu'au plus un sinistre dans l'année.

Charges individuelles

Les charges individuelles d'un assuré sont sa franchise individuelle et sa dépense individuelle dans l'année.

Engagement réciproque

Un engagement réciproque est un montant fixe défini par deux assurés en relation, servant à contribuer réciproquement à la couverture de leurs franchises individuelles en cas de sinistre. Ce montant sera défini dans ce mémoire comme fonction du degré des assurés de la communauté.

Contribution réciproque

La contribution réciproque aux charges est un montant fixe défini par deux assurés en relation, servant à contribuer réciproquement à la couverture de leurs charges (franchises et dépenses) annuelles. En particulier, la contribution réciproque aux franchises entre deux assurés correspond à leur engagement réciproque pour la couverture de leurs franchises individuelles.

Contribution individuelle

La contribution individuelle aux charges d'un assuré correspond à la somme de ses contributions réciproques aux charges de ses relations. En d'autres termes, c'est la cotisation annuelle de cet assuré pour la couverture des charges annuelles de la communauté des assurés.

Couverture individuelle

La couverture individuelle des charges d'un assuré correspond à la somme des contributions réciproques aux charges de ses relations.

3.2 Introduction au produit

Le produit de franchise collaborative initié en 2014 par la startup française *Inspeer*, consiste pour un assuré à mutualiser ses franchises d'assurance dommage (automobile, habitation, moto) avec ses proches ou ses relations et ce, sans changer de contrat d'assurance. En cas de sinistre, la communauté des assurés contribue au paiement de la franchise : la couverture de franchise est mutuelle, chacun est soutenu à tour de rôle et reçoit le montant sur lequel ses proches partenaires se sont engagés. Ce système permet donc pour un assuré de : 1) diminuer le coût de son assurance en augmentant sa franchise d'assurance³ et 2) faire couvrir sa franchise d'assurance par la communauté des assurés.

Cependant, étant donné que l'assuré s'engage à contribuer de manière équivalente à la franchise de ses relations, il porte une partie du risque en cas de sinistre d'un grand nombre de ces derniers. Pour pallier à ce risque, Dubois [2015] propose une assurance complémentaire dit « assurance stoploss » afin de limiter la contribution d'un assuré en cas de sinistre de ses relations. En effet, force est de constater que, les couvertures et les contributions des assurés constituant les variables d'intérêt du produit de franchise collaborative, sont fortement dépendantes de la structure du réseau de la communauté des assurés.

3. D'après la relation : prime = $\mathbb{E} [\max(\text{coût sinistres} - \text{franchise}, 0)]$.

L'objectif de ce mémoire consiste donc à analyser l'impact de la structure du réseau des assurés sur ces variables d'intérêt et à mettre en place efficacement des mécanismes permettant l'optimisation du produit, c'est-à-dire : 1) garantir la couverture des charges (franchises et dépenses) annuelles d'un nombre maximum d'assurés et 2) minimiser les contributions individuelles des assurés à la couverture des charges annuelles de leurs relations.

3.3 Formalisation mathématique du produit

Dans cette section, une formalisation mathématique du produit de franchise collaborative est effectuée partant des premières définitions jusqu'aux définitions des variables d'intérêt du produit, notamment les variables de couverture des charges et de contribution aux charges des assurés.

3.3.1 Premières définitions

La communauté des assurés est représentée par son graphe social $G = (A, R)$, qui est un graphe non orienté représentant le réseau des relations entre les assurés, tel que :

- ❖ $A = (A_i)_{1 \leq i \leq n}$ représente l'ensemble des n ($n \geq 2$) sommets du graphe G (i.e. les assurés) ;
- ❖ $R = (R_{ij})_{1 \leq i,j \leq n}$ représente l'ensemble des arêtes du graphe G (i.e. les relations entre les assurés), assimilé à la matrice d'adjacence de ce graphe de telle sorte que :

$$R_{ij} = \begin{cases} 1 & , \text{ si les assurés } i \text{ et } j \text{ sont en relation} \\ 0 & , \text{ sinon.} \end{cases}$$

On note également par :

- ❖ $d = (d_i)_{1 \leq i \leq n}$, l'ensemble des degrés des n sommets du graphe G (i.e. le nombre de relations de chaque assuré) avec :

$$\forall \text{ assuré } i = 1 : n, \quad d_i = \sum_{j=1}^n R_{ij} ;$$

- ❖ $F = (F_i)_{1 \leq i \leq n}$, l'ensemble des franchises associées aux contrats d'assurance individuels des n assurés ;
- ❖ $D = (D_i)_{1 \leq i \leq n}$, l'ensemble des dépenses individuelles des n assurés dans l'année où la dépense individuelle D_i de l'assuré $i = 1 : n$, est fonction de la franchise individuelle F_i associée au contrat de cet assuré et du coût S_i de ses sinistres dans l'année, telle que :

$$D_i = \min(S_i, F_i) ;$$

- ❖ $E = (E_{ij})_{1 \leq i,j \leq n}$, l'ensemble des engagements réciproques entre les assurés, nécessaires à la couverture de leurs franchises où E_{ij} représente l'engagement de l'assuré i envers sa relation j ;
- ❖ $C = (C_{ij})_{1 \leq i,j \leq n}$, l'ensemble des contributions réciproques aux dépenses entre les assurés, nécessaires à la couverture de leurs dépenses annuelles où la contribution C_{ij} de l'assuré i à la couverture de dépense de sa relation j est définie par :

$$\forall \text{ assurés } i, j = 1 : n, \quad C_{ij} = \min\left(\frac{S_j}{F_j} \times E_{ij}, E_{ij}\right) = \begin{cases} \frac{E_{ij}}{F_j} & , \text{ si } F_j \leq S_j \\ \frac{S_j}{F_j} \times E_{ij} & , \text{ sinon.} \end{cases} \quad (3.1)$$

Partant de ces notations, on définit également par :

- ❖ $\bar{F} = \frac{1}{n} \sum_{i=1}^n F_i$, la **franchise moyenne** associée aux n contrats des assurés de la communauté dans l'année ;
- ❖ $\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n R_{ij}$, le **degré moyen** des n sommets du graphe G (i.e. le nombre moyen de relations par assuré) ;
- ❖ $\bar{E} = \frac{\bar{F}}{\bar{d}}$, l'**engagement moyen** d'un assuré pour sa relation dans l'année.

3.3.2 Définitions des variables d'intérêt

Dans cette sous-section, on définit les variables d'intérêt du produit de franchise collaborative, notamment les variables de couverture des charges et de contribution aux charges de la communauté des assurés.

On définit donc dans un premier temps les charges annuelles de la communauté.

Définition 6 (Charges annuelles de la communauté).

- ❖ On appelle **charges individuelles** d'un assuré $j = 1 : n$, la franchise individuelle F_j et la dépense individuelle D_j de cet assuré j dans l'année telle que :

$$D_j = \min(S_j, F_j) ,$$

avec S_j , représentant le coût de sinistres de l'assuré $j = 1 : n$ dans l'année ;

- ❖ On appelle **charges mutuelles** de la communauté des assurés, la franchise mutuelle F^m et la dépense mutuelle D^m de la communauté dans l'année, telles que :

$$F^m = \sum_{j=1}^n F_j \quad \text{et} \quad D^m = \sum_{j=1}^n D_j .$$

Les charges annuelles de la communauté des assurés étant définies, on peut introduire la définition des variables de couverture de ces charges.

Définition 7 (Variables de couverture).

- ❖ Soit $(Couv_f_j)_{1 \leq j \leq n}$, les n variables de **couvertures individuelles des franchises** annuelles des assurés de la communauté définies par :

$$\forall \text{ assuré } j = 1 : n, \ Cov_f_j = \sum_{i=1}^n E_{ij} ,$$

où E_{ij} représente l'engagement de l'assuré i à la couverture de franchise de sa relation j .

- ❖ Soit $(Couv_d_j)_{1 \leq j \leq n}$, les n variables de **couvertures individuelles des dépenses** annuelles des assurés de la communauté définies par :

$$\forall \text{ assuré } j = 1 : n, \ Cov_d_j = \sum_{i=1}^n C_{ij} ,$$

où C_{ij} représente la contribution de l'assuré i à la couverture de dépense de sa relation j .

- ❖ Soit $Couv_f^m$, la variable de **couverture mutuelle des franchises** annuelles de la communauté définie par :

$$Couv_f^m = \sum_{j=1}^n Cov_f_j = \sum_{j=1}^n \sum_{i=1}^n E_{ij} .$$

- ❖ Soit $Couv_d^m$, la variable de **couverture mutuelle des dépenses** annuelles de la communauté définie par :

$$Couv_d^m = \sum_{j=1}^n Cov_d_j = \sum_{j=1}^n \sum_{i=1}^n C_{ij} .$$

On introduit également la définition des variables de contribution aux charges de la communauté des assurés.

Définition 8 (Variables de contribution).

- ❖ Soit $(Cont_f_i)_{1 \leq i \leq n}$, les n variables de **contributions individuelles aux franchises** annuelles des assurés de la communauté définies par :

$$\forall \text{ assuré } i = 1 : n, Cont_f_i = \sum_{j=1}^n E_{ij} ,$$

où E_{ij} représente l'engagement de l'assuré i à la couverture de franchise de sa relation j .

- ❖ Soit $(Cont_d_i)_{1 \leq i \leq n}$, les n variables de **contributions individuelles aux dépenses** annuelles des assurés de la communauté définies par :

$$\forall \text{ assuré } i = 1 : n, Cont_d_i = \sum_{j=1}^n C_{ij} ,$$

où C_{ij} représente la contribution de l'assuré i à la couverture de dépense de sa relation j .

- ❖ Soit $Cont_f^m$, la variable de **contribution mutuelle aux franchises** annuelles de la communauté définie par :

$$Cont_f^m = \sum_{i=1}^n Cont_f_i = \sum_{i=1}^n \sum_{j=1}^n E_{ij} .$$

- ❖ Soit $Cont_d^m$, la variable de **contribution mutuelle aux dépenses** annuelles de la communauté définie par :

$$Cont_d^m = \sum_{i=1}^n Cont_d_i = \sum_{i=1}^n \sum_{j=1}^n C_{ij} .$$

Remarque :

D'après la définition des variables de **contribution mutuelle aux franchises** et de **couverture mutuelle des franchises**, on a :

$$\begin{aligned}
 Cont_f^m &= \sum_{i=1}^n \sum_{j=1}^n E_{ij} \\
 &= \sum_{j=1}^n \sum_{i=1}^n E_{ij} \\
 Cont_f^m &= Couv_f^m
 \end{aligned} \tag{3.2}$$

De même, d'après la définition des variables de **contribution mutuelle aux dépenses** et de **couverture mutuelle des dépenses**, on a :

$$\begin{aligned}
 Cont_d^m &= \sum_{i=1}^n \sum_{j=1}^n C_{ij} \\
 &= \sum_{j=1}^n \sum_{i=1}^n C_{ij} \\
 Cont_d^m &= Couv_d^m
 \end{aligned} \tag{3.3}$$

C'est-à-dire, les contributions mutuelles des assurés financent les couvertures mutuelles des charges de la communauté. Ce qui permet d'obtenir en moyenne, l'**équilibre du système couverture-contribution**. Dans la suite du mémoire, on cherchera à comprendre comment fixer les engagements réciproques entre les assurés de telle sorte que les couvertures mutuelles couvrent (i.e. excèdent) les charges mutuelles.

3.4 Pistes de réflexion

Pour rappel, l'objectif principal de ce mémoire consiste en l'optimisation du produit de franchise collaborative. C'est-à-dire, mettre en place efficacement des mécanismes sur les paramètres (montant des **engagements réciproques** et modèle de **graphe social**) du produit de franchise collaborative, afin de :

- 1) Garantir la **couverture des charges** (franchises et dépenses) annuelles d'un nombre maximum d'assurés ;
- 2) Minimiser les **contributions individuelles** des assurés à la couverture des charges annuelles de la communauté.

Pour ce faire, la suite du mémoire consistera à porter les réflexions sur ces paramètres, au travers des deux interrogations suivantes :

Question (Q1) : Comment fixer le montant des engagements réciproques afin d'optimiser le produit de franchise collaborative ?

Question (Q2) : Quel(s) modèle(s) de graphe social permet(tent) d'optimiser le produit de franchise collaborative ?

Le chapitre 4 sera consacré à apporter des réponses à la question (Q1) par une approche théorique au moyen d'hypothèses sur le montant des engagements réciproques optimaux. Et le chapitre 5 traitera de la question (Q2), au travers d'une étude pratique de couverture et de contribution par la mise en place d'indicateurs qui permettront la sélection du modèle de graphe social optimal.

Chapitre 4

Étude théorique des engagements réciproques optimaux

Ce chapitre a pour objectif d'apporter des éléments de réponses à la question (**Q1**), évoquée dans la section [3.4](#), concernant le choix du montant des engagements réciproques permettant d'optimiser le produit de franchise collaborative. Pour ce faire, deux hypothèses sont émises sur la manière de fixer le montant des engagements réciproques en fonction de différentes motivations : la première hypothèse est motivée par le souhait d'avoir un montant d'engagement réciproque équivalent entre deux assurés en relation et la seconde hypothèse permet d'obtenir une couverture optimale des charges de la communauté. Aussi, ce chapitre s'achève par un exemple illustratif permettant d'appréhender, à une échelle réduite, le produit de franchise collaborative et en particulier, l'impact des hypothèses concernant les engagements réciproques sur les distributions des variables d'intérêt du produit de franchise collaborative.

4.1 Hypothèses de base (HDB)

Cette section consiste, tout d'abord, en une présentation des hypothèses dites de base (**HDB**) et de leur motivation, ensuite en une énumération des remarques et premières conséquences associées à ces hypothèses. Enfin, un énoncé et une analyse des résultats théoriques obtenus sous ces hypothèses sont effectués concernant les distributions des variables d'intérêt du produit de franchise collaborative.

4.1.1 Hypothèses

Pour rappel, un engagement réciproque entre deux assurés en relation est un montant fixe défini par ces assurés, servant à contribuer réciproquement à la couverture de leurs franchises individuelles en cas de sinistre. Dans le but d'obtenir un montant d'engagement réciproque **équivalent** entre deux assurés en relation, les hypothèses dites de base (**HDB**) suivantes sur la franchise et les **engagements réciproques** entre les assurés, sont émises :

(HDB1) Franchise : On suppose que le montant de franchise individuel est constant et identique pour tous les assurés de la communauté, i.e. :

$$\forall \text{ assuré } j = 1 : n, F_j = F ;$$

(HDB2) Engagements réciproques : On suppose que l'engagement E_{ij} de l'assuré i à la couverture de franchise de sa relation j est fonction de l'engagement moyen \bar{E} et de la matrice d'adjacence R du graphe social de la communauté des assurés, tel que :

$$\forall \text{ assurés } i, j = 1 : n, E_{ij} = \bar{E} \times R_{ij} = \begin{cases} \frac{\bar{F}}{d}, & \text{si les assurés } i \text{ et } j \text{ sont en relation} \\ 0, & \text{sinon.} \end{cases}$$

4.1.2 Remarques

Sous les conditions (**HDB**), on a les premières remarques suivantes :

- 1) La matrice $E = (E_{ij})_{1 \leq i,j \leq n}$ des engagements réciproques entre les assurés est **symétrique**, i.e. : \forall assurés $i, j = 1 : n$, tels que $i \neq j$,

$$\begin{aligned} E_{ij} &\stackrel{\text{(d'après (HDB2))}}{=} \frac{\bar{F}}{\bar{d}} \times R_{ij} \\ E_{ij} &\stackrel{\text{(d'après (HDB1))}}{=} \frac{F}{\bar{d}} \times R_{ij} \\ &\stackrel{\text{(par symétrie de } R)}{=} \frac{F}{\bar{d}} \times R_{ji} \\ E_{ij} &= E_{ji} \end{aligned} \quad (4.1)$$

En d'autres termes, deux assurés en relation, s'engagent à contribuer réciproquement pour un montant **équivalent** à la couverture de leurs franchises respectives.

- 2) La matrice $C = (C_{ij})_{1 \leq i,j \leq n}$ des contributions réciproques aux dépenses entre les assurés est **non symétrique**, i.e. : \forall assurés $i, j = 1 : n$, tels que $i \neq j$,

$$\begin{aligned} C_{ij} &\stackrel{\text{(d'après (3.1))}}{=} \min \left(\frac{S_j}{F_j} \times E_{ij}, E_{ij} \right) \\ &\stackrel{\text{(d'après (HDB))}}{=} \min \left(\frac{S_j}{\bar{F}} \times \frac{F}{\bar{d}}, \frac{F}{\bar{d}} \right) R_{ij} \\ &= \min \left(\frac{S_j}{\bar{d}}, \frac{F}{\bar{d}} \right) R_{ij} \\ &= \frac{\min(S_j, F)}{\bar{d}} \times R_{ij} \\ C_{ij} &= \frac{D_j}{\bar{d}} \times R_{ij} \\ C_{ij} &\neq \frac{D_i}{\bar{d}} \times R_{ji} = C_{ji} \end{aligned} \quad (4.2)$$

En d'autres termes, la contribution d'un assuré à la couverture de dépense de sa relation dépend de cette dépense.

- 3) La **symétrie** de la matrice E des engagements réciproques entre les assurés entraîne l'égalité entre les **couvertures individuelles des franchises** et les **contributions individuelles aux franchises**. En effet : \forall assuré $j = 1 : n$,

$$\begin{aligned} \text{Couv_f}_j &= \sum_{i=1}^n E_{ij} \\ &\stackrel{\text{(par symétrie de } E)}{=} \sum_{i=1}^n E_{ji} \\ \text{Couv_f}_j &= \text{Cont_f}_j \end{aligned} \quad (4.3)$$

En d'autres termes, la contribution individuelle aux franchises (ou cotisation annuelle) d'un assuré coïncide avec la couverture de sa franchise annuelle.

- 4) La remarque précédente sur l'égalité entre les **couvertures individuelles des franchises** et les **contributions individuelles aux franchises** n'est plus valable pour la couverture des dépenses individuelles car la matrice C des contributions réciproques aux dépenses est **non symétrique**. En effet : \forall assuré $j = 1 : n$,

$$\begin{aligned} \text{Couv_}d_j &= \sum_{i=1}^n C_{ij} \\ \text{Couv_}d_j &\stackrel{\text{(par non symetrie de } C\text{)}}{\neq} \sum_{i=1}^n C_{ji} = \text{Cont_}d_j \end{aligned} \quad (4.4)$$

En d'autres termes, la contribution individuelle aux dépenses d'un assuré est indépendante de sa couverture de dépense annuelle.

4.1.3 Résultats théoriques

Partant des hypothèses (**HDB**), on peut énoncer les résultats théoriques suivants concernant les distributions des variables d'intérêt du produit de franchise collaborative.

On s'intéresse dans un premier temps, à la distribution des variables de couverture des charges annuelles des assurés.

Couverture des charges

Sous les conditions (**HDB**), on obtient les résultats suivants :

- ❖ **Franchises individuelles** : Pour tout assuré $j = 1 : n$, sa couverture individuelle $\text{Couv_}f_j$ de franchise annuelle s'écrit :

$$\begin{aligned} \text{Couv_}f_j &= \sum_{i=1}^n E_{ij} \\ &\stackrel{\text{(d'après (4.1))}}{=} \sum_{i=1}^n \frac{F}{\bar{d}} \times R_{ij} \\ &= \frac{F}{\bar{d}} \sum_{i=1}^n R_{ij} \\ \text{Couv_}f_j &= \frac{F}{\bar{d}} \times d_j \end{aligned} \quad (4.5)$$

- ❖ **Dépenses individuelles** : Pour tout assuré $j = 1 : n$, sa couverture individuelle $\text{Couv_}d_j$ de dépense annuelle s'écrit :

$$\begin{aligned} \text{Couv_}d_j &= \sum_{i=1}^n C_{ij} \\ &\stackrel{\text{(d'après (4.2))}}{=} \sum_{i=1}^n \frac{D_j}{\bar{d}} \times R_{ij} \\ &= \frac{D_j}{\bar{d}} \sum_{i=1}^n R_{ij} \\ \text{Couv_}d_j &= \frac{D_j}{\bar{d}} \times d_j \end{aligned} \quad (4.6)$$

- ❖ **Franchise mutuelle :** La couverture mutuelle $Couv_f^m$ des franchises annuelles de la communauté s'écrit :

$$\begin{aligned}
 Cov_f^m &= \sum_{j=1}^n Cov_f_j \\
 &\stackrel{(d'apr\grave{e}s (4.5))}{=} \sum_{j=1}^n \frac{F}{\bar{d}} \times d_j \\
 &= \frac{F}{\bar{d}} \sum_{j=1}^n d_j \\
 &\stackrel{(par d\acute{e}finition de \bar{d})}{=} \frac{F}{\bar{d}} \times n \times \bar{d} \\
 &= n \times F \\
 &\stackrel{(d'apr\grave{e}s (HDB1))}{=} \sum_{j=1}^n F_j \\
 Cov_f^m &= F^m
 \end{aligned} \tag{4.7}$$

- ❖ **Dépense mutuelle :** La couverture mutuelle $Couv_d^m$ des dépenses annuelles de la communauté s'écrit :

$$\begin{aligned}
 Cov_d^m &= \sum_{j=1}^n Cov_d_j \\
 Cov_d^m &\stackrel{(d'apr\grave{e}s (4.6))}{=} \sum_{j=1}^n \frac{D_j}{\bar{d}} \times d_j
 \end{aligned} \tag{4.8}$$

$$Cov_d^m \neq \sum_{j=1}^n D_j = D^m \tag{4.9}$$

Intéressons-nous maintenant, à la distribution des variables de contribution aux charges annuelles des assurés.

Contribution aux charges

Sous les conditions (**HDB**), on obtient les résultats suivants :

- ❖ **Franchises individuelles :** Pour tout assuré $i = 1 : n$, sa contribution individuelle $Cont_f_i$ aux franchises annuelles de la communauté s'écrit :

$$\begin{aligned}
 Cont_f_i &\stackrel{(d'apr\grave{e}s (4.3))}{=} Cov_f_i \\
 Cont_f_i &\stackrel{(d'apr\grave{e}s (4.5))}{=} \frac{F}{\bar{d}} \times d_i
 \end{aligned} \tag{4.10}$$

- ❖ **Dépenses individuelles :** Pour tout assuré $i = 1 : n$, sa contribution individuelle $Cont_d_i$ aux dépenses annuelles de la communauté s'écrit :

$$\begin{aligned} Cont_d_i &= \sum_{j=1}^n C_{ij} \\ Cont_d_i &\stackrel{(d'après (4.2))}{=} \sum_{j=1}^n \frac{D_j}{\bar{d}} \times R_{ij} \end{aligned} \quad (4.11)$$

- ❖ **Franchise mutuelle :** La contribution mutuelle $Cont_f^m$ aux franchises annuelles de la communauté s'écrit :

$$\begin{aligned} Cont_f^m &\stackrel{(d'après (3.2))}{=} Couv_f^m \\ Cont_f^m &\stackrel{(d'après (4.7))}{=} F^m \end{aligned} \quad (4.12)$$

- ❖ **Dépense mutuelle :** La contribution mutuelle $Cont_d^m$ aux dépenses annuelles de la communauté s'écrit :

$$\begin{aligned} Cont_d^m &\stackrel{(d'après (3.3))}{=} Couv_d^m \\ Cont_d^m &\stackrel{(d'après (4.8))}{=} \sum_{j=1}^n \frac{D_j}{\bar{d}} \times d_j \end{aligned} \quad (4.13)$$

4.1.4 Analyse des résultats

Dans cette sous-section, il est effectué une analyse des résultats théoriques précédents, concernant les distributions des variables d'intérêt du produit de franchise collaborative.

Analyse de la couverture des charges mutuelles

D'après la relation 4.7, on constate que la couverture mutuelle des franchises coïncide avec la franchise mutuelle. C'est-à-dire, d'après la relation 3.2, la contribution mutuelle aux franchises des assurés permet de couvrir la franchise mutuelle de la communauté.

Cependant les relations 3.3 et 4.9 permettent de conclure que la contribution mutuelle aux dépenses des assurés ne couvre pas nécessairement la dépense mutuelle de la communauté.

Analyse de la couverture des charges individuelles

D'après les relations 4.5 et 4.6, si le degré d_j d'un assuré $j = 1 : n$, est inférieur au degré moyen \bar{d} de la communauté (i.e. $d_j < \bar{d}$), alors les charges annuelles de cet assuré ne sont pas couvertes. En effet :

$$Couv_f_j \stackrel{(4.5)}{=} \frac{F}{\bar{d}} \times d_j < F_j \quad \text{et} \quad Couv_d_j \stackrel{(4.6)}{=} \frac{D_j}{\bar{d}} \times d_j < D_j .$$

En d'autres termes, les charges des assurés moins populaires (i.e. possédant un nombre d'amis inférieur à la moyenne) ne sont pas couvertes par les contributions de leurs relations.

Pour les mêmes causes, les assurés les plus populaires (i.e. possédant un nombre d'amis supérieur à la moyenne) ont leurs charges largement couvertes par les contributions de leurs relations.

On constate donc une éventuelle disparité des couvertures individuelles des charges en fonction de la variance de la distribution des degrés des assurés. De ce fait, une étude de couverture sera effectuée dans le chapitre 5 sous les hypothèses (**HDB**), afin de sélectionner le modèle de graphe social permettant d'optimiser le produit de franchise collaborative, par une optimisation des couvertures individuelles des charges annuelles des assurés.

Par ailleurs, pour pallier à ce problème de disparité des couvertures individuelles, sachant que le **système couverture-contribution est en moyenne équilibré**, l'assureur peut donc réallouer les surplus de couverture observés chez les assurés les plus populaires vers les assurés les moins populaires dont les charges ne sont pas couvertes.

La section suivante consistera à proposer un mécanisme de réallocation de ces surplus de couverture, en supposant que, le montant des engagements réciproques entre les assurés, est cette fois, fonction des degrés respectifs des assurés et non du degré moyen de la communauté, comme c'est le cas sous les conditions (**HDB**).

4.2 Hypothèses de couverture optimale (HCO)

Dans cette section, d'autres hypothèses sont émises sur le montant des engagements réciproques entre les assurés, dans le but d'optimiser la couverture des charges annuelles des assurés. On pourra également analyser les conséquences de ces nouvelles hypothèses sur la distribution des variables de contribution aux charges annuelles des assurés.

4.2.1 Hypothèses

Sous les conditions (**HDB**), on a pu constater que les charges annuelles des assurés les moins populaires ne sont pas "couvertes", tandis que celles des assurés les plus populaires pourraient être "trop couvertes" en fonction de la distribution des degrés des assurés. Pour pallier à cette disparité des couvertures individuelles, on émet les hypothèses de couverture optimale (**HCO**) suivantes :

(HCO1) Franchise : Comme pour l'hypothèse (**HDB1**), on suppose également que le montant de franchise individuel est constant et identique pour tous les assurés de la communauté, i.e. :

$$\forall \text{ assuré } j = 1 : n, F_j = F ;$$

(HCO2) Engagements réciproques : On suppose que l'engagement E_{ij} de l'assuré i à la couverture de franchise F_j de sa relation j est fonction de la franchise F_j , du degré d_j de l'assuré j et de la matrice d'adjacence R du graphe social de la communauté, tel que :

$$\forall \text{ assurés } i, j = 1 : n, E_{ij} = \frac{F_j}{d_j} \times R_{ij} = \begin{cases} \frac{F_j}{d_j}, & \text{si les assurés } i \text{ et } j \text{ sont en relation} \\ 0, & \text{sinon.} \end{cases}$$

4.2.2 Remarques

Sous les conditions (**HCO**), on a les premières remarques suivantes :

- 1) La matrice $E = (E_{ij})_{1 \leq i, j \leq n}$ des engagements réciproques entre les assurés est **non symétrique**, i.e. : \forall assurés $i, j = 1 : n$, tels que $i \neq j$,

$$\begin{aligned}
 E_{ij} &\stackrel{\text{(d'après (HCO2))}}{=} \frac{F_j}{d_j} \times R_{ij} \\
 E_{ij} &\stackrel{\text{(d'après (HCO1))}}{=} \frac{F}{d_j} \times R_{ij} \\
 E_{ij} &\neq \frac{F}{d_i} \times R_{ji} = E_{ji}
 \end{aligned} \tag{4.14}$$

En d'autres termes, la contribution d'un assuré à la couverture de franchise annuelle de sa relation dépend du degré de sa relation. Plus ce degré est élevé (i.e assuré populaire), plus le montant de l'engagement est faible.

- 2) La matrice $C = (C_{ij})_{1 \leq i, j \leq n}$ des contributions réciproques aux dépenses entre les assurés est **non symétrique**, i.e. : \forall assurés $i, j = 1 : n$, tels que $i \neq j$,

$$\begin{aligned}
 C_{ij} &\stackrel{\text{(d'après (3.1))}}{=} \min \left(\frac{S_j}{F_j} \times E_{ij}, E_{ij} \right) \\
 &\stackrel{\text{(d'après (HCO))}}{=} \min \left(\frac{S_j}{F} \times \frac{F}{d_j}, \frac{F}{d_j} \right) R_{ij} \\
 &= \min \left(\frac{S_j}{d_j}, \frac{F}{d_j} \right) R_{ij} \\
 C_{ij} &= \frac{D_j}{d_j} \times R_{ij} \\
 C_{ij} &\neq \frac{D_i}{d_i} \times R_{ji} = C_{ji}
 \end{aligned} \tag{4.15}$$

En d'autres termes, la contribution d'un assuré à la couverture de dépense annuelle de sa relation dépend du degré et de la dépense de sa relation.

4.2.3 Résultats théoriques

Partant des hypothèses (**HCO**), on peut énoncer les résultats théoriques suivants concernant les distributions des variables d'intérêt du produit de franchise collaborative.

On s'intéresse dans un premier temps, à la distribution des variables de couverture des charges annuelles des assurés.

Couverture des charges

Sous les conditions (**HCO**), on obtient les résultats suivants :

- ❖ **Franchises individuelles** : Pour tout assuré $j = 1 : n$, sa couverture individuelle $Couv_f_j$ de franchise annuelle s'écrit :

$$\begin{aligned}
Couv_f_j &= \sum_{i=1}^n E_{ij} \\
&\stackrel{\text{(d'après (4.14))}}{=} \sum_{i=1}^n \frac{F}{d_j} \times R_{ij} \\
&= \frac{F}{d_j} \sum_{i=1}^n R_{ij} \\
&= \frac{F}{d'_j} \times d'_j = F \\
Couv_f_j &\stackrel{\text{(d'après (HCO1))}}{=} F_j
\end{aligned} \tag{4.16}$$

❖ **Dépenses individuelles :** Pour tout assuré $j = 1 : n$, sa couverture individuelle $Couv_d_j$ de dépense annuelle s'écrit :

$$\begin{aligned}
Couv_d_j &= \sum_{i=1}^n C_{ij} \\
&\stackrel{\text{(d'après (4.15))}}{=} \sum_{i=1}^n \frac{D_j}{d_j} \times R_{ij} \\
&= \frac{D_j}{d_j} \sum_{i=1}^n R_{ij} \\
&= \frac{D_j}{d'_j} \times d'_j \\
Couv_d_j &= D_j
\end{aligned} \tag{4.17}$$

❖ **Franchise mutuelle :** La couverture mutuelle $Couv_f^m$ des franchises annuelles de la communauté s'écrit :

$$\begin{aligned}
Couv_f^m &= \sum_{j=1}^n Couv_f_j \\
&\stackrel{\text{(d'après (4.16))}}{=} \sum_{j=1}^n F_j \\
Couv_f^m &= F^m
\end{aligned} \tag{4.18}$$

❖ **Dépense mutuelle :** La couverture mutuelle $Couv_d^m$ des dépenses annuelles de la communauté s'écrit :

$$\begin{aligned}
Couv_d^m &= \sum_{j=1}^n Couv_d_j \\
Couv_d^m &\stackrel{\text{(d'après (4.17))}}{=} \sum_{j=1}^n D_j \\
Couv_d^m &= D^m
\end{aligned} \tag{4.19}$$

Intéressons-nous maintenant, à la distribution des variables de contribution aux charges annuelles des assurés.

Contribution aux charges

Sous les conditions (**HCO**), on obtient les résultats suivants :

- ❖ **Franchises individuelles** : Pour tout assuré $i = 1 : n$, sa contribution individuelle $Cont_f_i$ aux franchises annuelles de la communauté s'écrit :

$$\begin{aligned} Cont_f_i &= \sum_{j=1}^n E_{ij} \\ Cont_f_i &\stackrel{\text{(d'après (4.14))}}{=} \sum_{j=1}^n \frac{F}{d_j} \times R_{ij} \end{aligned} \quad (4.20)$$

- ❖ **Dépenses individuelles** : Pour tout assuré $i = 1 : n$, sa contribution individuelle $Cont_d_i$ aux dépenses annuelles de la communauté s'écrit :

$$\begin{aligned} Cont_d_i &= \sum_{j=1}^n C_{ij} \\ Cont_d_i &\stackrel{\text{(d'après (4.15))}}{=} \sum_{j=1}^n \frac{D_j}{d_j} \times R_{ij} \end{aligned} \quad (4.21)$$

- ❖ **Franchise mutuelle** : La contribution mutuelle $Cont_f^m$ aux franchises annuelles de la communauté s'écrit :

$$\begin{aligned} Cont_f^m &\stackrel{\text{(d'après (3.2))}}{=} Couv_f^m \\ Cont_f^m &\stackrel{\text{(d'après (4.18))}}{=} F^m \end{aligned} \quad (4.22)$$

- ❖ **Dépense mutuelle** : La contribution mutuelle $Cont_d^m$ aux dépenses annuelles de la communauté s'écrit :

$$\begin{aligned} Cont_d^m &\stackrel{\text{(d'après (3.3))}}{=} Couv_d^m \\ Cont_d^m &\stackrel{\text{(d'après (4.19))}}{=} D^m \end{aligned} \quad (4.23)$$

4.2.4 Analyse des résultats

Dans cette sous-section, on fait une analyse des résultats théoriques précédents, concernant les distributions des variables d'intérêt du produit de franchise collaborative.

Analyse de la couverture des charges mutuelles

D'après les relations 4.18, 4.19, 4.22 et 4.23, on peut conclure que les contributions mutuelles des assurés finançant les couvertures mutuelles, permettent de couvrir les charges mutuelles de la communauté.

Analyse de la couverture des charges individuelles

Les relations 4.16 et 4.17 permettent de conclure que les charges individuelles des assurés de la communauté sont "parfaitement" couvertes.

Cependant, d'après les relations 4.20 et 4.21, on constate que les contributions individuelles des assurés sont fonction des degrés de leurs relations. De ce fait, une étude de contribution sera effectuée dans le chapitre suivant sous les hypothèses (**HCO**) afin de sélectionner le modèle de graphe social permettant d'optimiser le produit de franchise collaborative, par une minimisation des contributions individuelles (ou cotisations) des assurés.

4.3 Synthèse hypothèses HDB vs HCO

Cette section qui s'inspire d'un exemple illustratif, a un objectif double :

- 1) Présentation du produit de franchise collaborative à une échelle réduite, afin de mieux appréhender les caractéristiques (franchises, coûts de sinistres et degrés des assurés) de la communauté ;
- 2) Analyse à une échelle réduite de l'impact des hypothèses (**HDB**) et (**HCO**) concernant les engagements réciproques entre les assurés sur les distributions des variables d'intérêt (variables de couverture et de contribution) du produit de franchise collaborative.

Exemple illustratif : Énoncé

On considère une communauté de $n = 8$ assurés où chaque assuré possède un contrat d'assurance IARD (e.g. automobile). Cette communauté intègre une plateforme de franchise collaborative dans le but de couvrir mutuellement les franchises associées à leurs contrats d'assurance respectifs. Les caractéristiques (franchises, coûts de sinistres annuels, dépenses¹ annuelles et degrés des assurés) de la communauté sont connues et résumées dans le Tableau 4.1.

Assurés \ Caract.	Franchises	Coûts des sinistres	Dépenses	Degrés
1	400	100	100	4
2	400	500	400	4
3	400	0	0	3
4	400	50	50	3
5	400	400	400	3
6	400	100	100	3
7	400	600	400	3
8	400	0	0	3
Mutuel	3.200	1.750	1.450	26

TABLE 4.1 – Exemple illustratif : Caractéristiques de la communauté

Par ailleurs, la Figure 4.1 représente le graphe social des relations entre les assurés de la communauté.

1. dépense = min(franchise, coût sinistres).

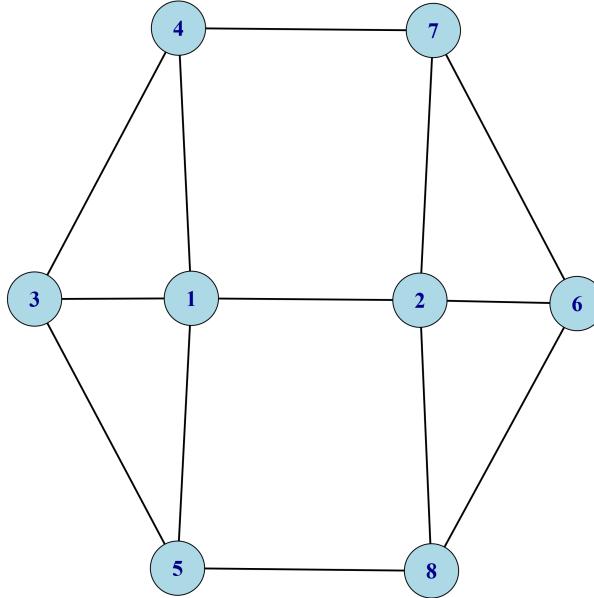


FIGURE 4.1 – Exemple illustratif : Graphe social

Résultats matriciels

Partant des caractéristiques (franchises, dépenses annuelles, graphe social et degrés des assurés) de la communauté, on peut déduire la **matrice d'adjacence** du graphe social, la **matrice des engagements réciproques** entre les assurés et la **matrice des contributions réciproques** aux dépenses. En effet, par définition, la détermination des distributions des variables d'intérêt du produit de franchise collaborative, nécessite premièrement, le calcul des matrices des engagements réciproques et des contributions réciproques aux dépenses entre les assurés.

❖ **Matrice d'adjacence** : La matrice d'adjacence R du graphe social de la communauté des assurés s'écrit :

$$R = \left(\begin{array}{cccccccc} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \end{array} \right) \left. \begin{array}{l} \text{Degrés des assurés} \\ d_i = \sum_{j=1}^8 R_{ij}, \forall \text{ assuré } i = 1 : 8 . \end{array} \right\}$$

Aussi, le degré moyen \bar{d} (i.e. le nombre moyen d'amis par assuré) de la communauté est :

$$\bar{d} = \frac{1}{8} \sum_{i=1}^8 d_i \stackrel{\text{(A.N.)}}{=} \frac{26}{8} = 3,25 .$$

❖ **Matrice des engagements réciproques** : La matrice E des engagements réciproques, sous les hypothèses (**HDB**) s'écrit :

$$E = \left(\begin{array}{cccc|ccccc} 0 & 123 & 123 & 123 & 123 & 0 & 0 & 0 \\ 123 & 0 & 0 & 0 & 0 & 123 & 123 & 123 \\ 123 & 0 & 0 & 123 & 123 & 0 & 0 & 0 \\ \hline 123 & 0 & 123 & 0 & 0 & 0 & 123 & 0 \\ 123 & 0 & 123 & 0 & 0 & 0 & 0 & 123 \\ 0 & 123 & 0 & 0 & 0 & 0 & 123 & 123 \\ 0 & 123 & 0 & 123 & 0 & 123 & 0 & 0 \\ 0 & 123 & 0 & 0 & 123 & 123 & 0 & 0 \end{array} \right) \left. \begin{array}{l} \\ \\ \\ \underbrace{\qquad\qquad\qquad}_{\text{Couverte des franchises}} \\ \\ \\ \end{array} \right\}$$

Contribution aux franchises

$$Cont_f_i = \sum_{j=1}^8 E_{ij}, \forall \text{ assuré } i$$

Couverture des franchises

$$Couv_f_j = \sum_{i=1}^8 E_{ij}, \forall \text{ assuré } j = 1 : 8$$

avec $E_{ij} = \frac{F}{d} \times R_{ij} \stackrel{\text{A.N.}}{=} \frac{400}{3.25} \times R_{ij} = 123 \times R_{ij}, \forall \text{ assurés } i, j = 1 : 8.$

De même, on détermine la matrice des engagements réciproques sous les hypothèses (**HCO**), pour le calcul des variables d'intérêt sous les conditions (**HDB**) et (**HCO**). Les Tableaux 4.2 et 4.3 suivants résument les distributions obtenues sous ces hypothèses.

❖ **Matrice des contributions réciproques :** La matrice C des contributions réciproques aux dépenses, sous les hypothèses (**HDB**) s'écrit :

$$C = \left(\begin{array}{cccc|ccccc} 0 & 123 & 0 & 15 & 123 & 0 & 0 & 0 \\ 31 & 0 & 0 & 0 & 0 & 31 & 123 & 0 \\ 31 & 0 & 0 & 15 & 123 & 0 & 0 & 0 \\ \hline 31 & 0 & 0 & 0 & 0 & 0 & 123 & 0 \\ 31 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 123 & 0 & 0 & 0 & 0 & 123 & 0 \\ 0 & 123 & 0 & 15 & 0 & 31 & 0 & 0 \\ 0 & 123 & 0 & 0 & 123 & 31 & 0 & 0 \end{array} \right) \left. \begin{array}{l} \\ \\ \\ \underbrace{\qquad\qquad\qquad}_{\text{Couverte des dépenses}} \\ \\ \\ \end{array} \right\}$$

Contribution aux dépenses

$$Cont_d_i = \sum_{j=1}^8 C_{ij}, \forall \text{ assuré } i$$

Couverture des dépenses

$$Couv_d_j = \sum_{i=1}^8 C_{ij}, \forall \text{ assuré } j = 1 : 8$$

avec $C_{ij} = \frac{D_j}{d} \times R_{ij}, \forall \text{ assurés } i, j = 1 : 8.$

De même, on détermine la matrice des contributions réciproques aux dépenses sous les hypothèses (**HCO**), pour le calcul des variables d'intérêt sous les conditions (**HDB**) et (**HCO**). Les Tableaux 4.2 et 4.3 suivants résument les distributions obtenues sous ces hypothèses.

Distributions des variables d'intérêt

Les Tableaux 4.2 et 4.3 résument respectivement les distributions des variables de couverture des charges et de contribution aux charges pour la communauté de l'exemple illustratif, suivant les hypothèses (**HDB**) et (**HCO**) sur les engagements réciproques entre les assurés.

Distribution de couverture	Couverture des franchises (en €)			Couverture des dépenses (en €)		
Assurés (degrés)	Franchises	(HDB)	(HCO)	Dépenses	(HDB)	(HCO)
1 (4)	400	492	400	100	123	100
2 (4)	400	492	400	400	492	400
3 (3)	400	369	400	0	0	0
4 (3)	400	369	400	50	46	50
5 (3)	400	369	400	400	369	400
6 (3)	400	369	400	100	92	100
7 (3)	400	369	400	400	369	400
8 (3)	400	369	400	0	0	0
Mutuel (26)	3.200	3.200	3.200	1.450	1.492	1.450

TABLE 4.2 – Exemple illustratif : Distributions de couverture des charges

Distribution de contribution	Contribution aux franchises (en €)			Contribution aux dépenses (en €)		
Assurés (degrés)	Franchises	(HDB)	(HCO)	Dépenses	(HDB)	(HCO)
1 (4)	400	492	500	100	262	250
2 (4)	400	492	500	400	185	192
3 (3)	400	369	367	0	169	175
4 (3)	400	369	367	50	154	158
5 (3)	400	369	367	400	31	25
6 (3)	400	369	367	100	246	233
7 (3)	400	369	367	400	169	150
8 (3)	400	369	367	0	277	267
Mutuel (26)	3.200	3.200	3.200	1.450	1.492	1.450

TABLE 4.3 – Exemple illustratif : Distributions de contribution aux charges

Analyse des distributions des variables d'intérêt

Les distributions de couverture et de contribution obtenues dans le cadre de l'exemple illustratif viennent confirmer les résultats théoriques énoncés plus haut concernant les distributions des variables d'intérêt sous les hypothèses **(HDB)** et **(HCO)**. En effet, on constate que sous ces hypothèses, les charges mutuelles de la communauté sont couvertes par les contributions mutuelles des assurés.

Cependant, notons que sous les hypothèses **(HDB)**, les charges individuelles des assurés numéro 3 à 8 (de degré 3, inférieur à la moyenne qui vaut 3,25) ne sont pas couvertes. Tandis que les charges individuelles des assurés numéro 1 et 2 (de degré 4, supérieur à la moyenne qui vaut 3,25) sont largement couvertes. C'est-à-dire, la distribution des degrés des assurés a une influence sur les couvertures individuelles des assurés. De ce fait, une étude de couverture sera effectuée dans le chapitre suivant sous les hypothèses **(HDB)** afin de sélectionner le modèle de graphe social permettant d'optimiser les couvertures individuelles des charges des assurés.

Par ailleurs, sous les hypothèses **(HCO)**, on constate que les charges individuelles des assurés sont "parfaitement" couvertes. Cependant, les contributions individuelles des assurés dépendent non seulement des charges mais également de la distribution des degrés de leurs relations respectives. De ce fait, une étude de contribution sera effectuée dans le chapitre suivant sous les hypothèses **(HCO)** afin de sélectionner le modèle de graphe social permettant d'optimiser (i.e. minimiser) les contributions individuelles des assurés.

Chapitre 5

Étude pratique du modèle de graphe social optimal

L'objectif de ce chapitre consiste à apporter des éléments de réponse à la question (**Q2**), évoquée dans la section [3.4](#), concernant le modèle de graphe social permettant d'optimiser le produit de franchise collaborative. Pour ce faire, partant d'une communauté fictive d'assurés et des hypothèses (**HDB**) et (**HCO**) sur les engagements réciproques entre les assurés, une étude de couverture et une étude de contribution sont respectivement effectuées sous les hypothèses (**HDB**) et (**HCO**) afin de sélectionner le modèle de graphe social permettant d'optimiser les couvertures individuelles des charges annuelles des assurés et de minimiser les contributions individuelles aux charges annuelles des assurés. Toutefois, ce chapitre fait dans un premier temps, une présentation des caractéristiques de la communauté fictive, de la calibration des paramètres (graphe social et engagements réciproques) et du programme implémenté sous le logiciel R pour le calcul des distributions des variables d'intérêt du produit de franchise collaborative.

5.1 Présentation de la communauté fictive

5.1.1 Description de la communauté fictive

La communauté fictive (**CF**) servant de portefeuille d'étude dans le cadre de ce chapitre est constituée d'assurés possédant chacun un contrat d'assurance automobile de 1^{ère} catégorie ¹ avec une garantie tous risques. De plus, la communauté **CF** est supposée faire partie intégrante d'une plateforme de franchise collaborative, dans le but de couvrir mutuellement les franchises associées aux contrats respectifs des assurés en cas de sinistre. Par ailleurs, rappelons que pour que la mutualisation des risques fonctionne, il faut que les assurés forment une communauté possédant des comportements et besoins les plus similaires possibles. On suppose donc que le portefeuille **CF** à l'étude possède les caractéristiques suivantes :

- ❖ **Coûts des sinistres** : On suppose que les assurés de la communauté **CF** ont le même profil de risque, avec une loi de sinistres suivant la sinistralité moyenne observée en 2018 sur le marché de l'assurance automobile en France. En effet, selon la FFA ², pour l'année 2018, la fréquence moyenne de sinistres des contrats auto (avec la garantie tous risques) a été de 8,7%, pour un coût moyen de sinistres de 1.565 €.

Partant de ces chiffres officiels, le **nombre de sinistres** dans l'année d'un assuré est simulé par une loi de Bernoulli $\mathcal{B}(p = 1/10)$, en supposant qu'un assuré ne peut avoir qu'au plus un sinistre dans l'année, avec une probabilité de **1/10**.

1. Véhicules 4 roues à moteur de P.T.A.C. (Poids Total Autorisé en Charge) inférieur à 3,5 tonnes hors flottes : voitures particulières et camionnettes.

2. FFA : Fédération Française d'Assurance.

Aussi, le **coût de sinistres** dans l'année d'un assuré est simulé par une loi Gamma $\mathcal{G}(shape = 15, 65 ; rate = 0, 01)$, de telle sorte que :

$$\mathbb{E}[\mathcal{G}(shape = 15, 65 ; rate = 0, 01)] = \frac{shape}{rate} = 1.565 ,$$

qui est le coût moyen de sinistres de l'année 2018 rappelé plus haut ;

- ❖ **Fanchises** : Les franchises annuelles des assurés de la plateforme sont supposées identiques et fixées à **1.000 €**.

Le choix de ce montant de franchise est motivé par le coût moyen de sinistres qui est de 1.565 €, de telle sorte, à ce que l'assureur rembourse en moyenne les 1/3 du coût de sinistre annuel d'un assuré et les 2/3 restant à la charge de l'assuré (et donc de la communauté). En effet, si une variable aléatoire X suit une loi Gamma $\mathcal{G}(shape = 15, 65 ; rate = 0, 01)$, alors on a le résultat suivant :

$$\mathbb{E}[\max(X - 1.000, 0)] \approx \frac{1}{3} \times \mathbb{E}[X]$$

Ce montant assez important de reste à charge de l'assuré permet donc à un assuré de la plateforme, de réduire sa prime classique d'assurance auto et de faire couvrir sa franchise (i.e. son reste à charge) par la communauté des assurés ;

- ❖ **Ordre et degré moyen** : On suppose que le portefeuille **CF** est constitué de $n = 493$ assurés. Cet ordre permet de regrouper les assurés dans des "clusters homogènes" de tailles ³ allant de 3 à 31 assurés (cf. section 5.2, pour plus de détails sur la calibration du graphe social de la communauté **CF**). Par ailleurs, ce choix permet d'obtenir des degrés allant de 2 à 30 relations, en raison de : 2 relations pour chacun des assurés composant le cluster homogène de taille 3, 3 relations pour chacun des assurés composant le cluster homogène de taille 4, et ainsi de suite jusqu'à 30 relations pour chacun des assurés composant le cluster homogène de taille 31.

Ce qui conduit à un degré moyen de $\bar{d} = 20,12 \approx 20$ relations (i.e. chaque assuré de la communauté **CF** possède en moyenne 20 relations). Ainsi, un assuré contribue en moyenne, à hauteur de 50 € pour la couverture des 1.000 € de franchise de sa relation. Ce montant d'engagement réciproque moyen de 50 € correspond donc en une prise de position marketing : il est assez faible pour ne pas dissuader les assurés de rejoindre la plateforme et il permet d'entrer en relation avec un nombre raisonnable d'assurés, soit 20 relations en moyenne.

5.1.2 Analyse des charges de la communauté fictive

Partant des hypothèses émises sur le montant des franchises et la loi de sinistres des assurés de la communauté **CF**, les charges (franchises, coûts des sinistres et dépenses⁴) annuelles des assurés sont déterminées par simulation. Sachant que la suite des travaux consiste à étudier la couverture de ces charges annuelles, il importe avant tout, d'analyser les distributions de charges obtenues. La boîte à moustaches (Figure 5.1) présente les distributions des charges annuelles de la communauté **CF**.

- ❖ **Franchises** : D'après l'hypothèse de constance du montant des franchises individuelles, on constate que la franchise moyenne et les trois quartiles (premier, deuxième et troisième quartiles) de la distribution des franchises sont confondus en 1.000 € (cf. Figure 5.1 (a)) ;

3. La taille d'un cluster correspond au nombre de membres de ce cluster.

4. dépense = min(franchise, coût sinistres).

- ❖ **Coûts des sinistres :** Après calcul des coûts de sinistres annuels par simulation, on obtient 88,2% d'assurés non sinistrés (i.e. avec des coûts de sinistres nuls), soit 435 sur les 493 assurés. Ce qui explique le fait que les trois quartiles de la distribution des coûts de sinistres, sont confondus en 0 € (cf. Figure 5.1 (b)). Aussi, le minimum des coûts de sinistres non nuls est de 504 € et le coût de sinistres le plus élevé est de 2.407 € avec 8,1% des sinistres ayant un coût supérieur à 1.000 €, soit 40 assurés sur 493. Enfin, les 18 assurés restants sur les 493, soit 3,7%, ont des coûts de sinistres compris entre 504 € et 1.000 €, pour un coût moyen de sinistres estimé à 137 € ;
- ❖ **Dépenses :** A l'instar de la distribution des coûts de sinistres et d'après la définition des dépenses annuelles, 88,2% des assurés ont des dépenses nulles. Pour cette même raison, les trois quartiles de la distribution des dépenses annuelles, sont confondus en 0 € (cf. Figure 5.1 (c)). Aussi, de même que pour les coûts de sinistres, le minimum des dépenses non nulles est de 504 € et le maximum des dépenses correspond au 1.000 € de franchise individuelle. Enfin, 3,7% des dépenses annuelles sont comprises entre 504 € et 1.000 €, 8,1% des dépenses annuelles valent exactement 1.000 € et la dépense annuelle moyenne de la communauté CF est de 108 €.

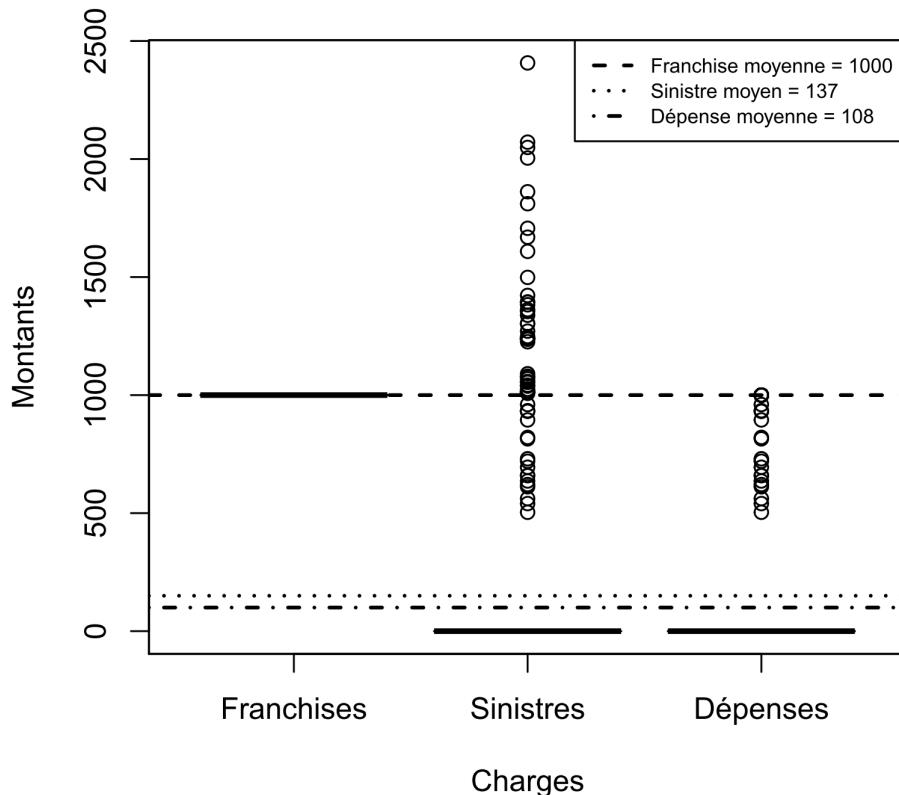


FIGURE 5.1 – Boîte à moustaches des charges de la communauté CF, obtenues par simulation : *Franchises* (a), *Coûts des sinistres* (b), *Dépenses* (c).

L'analyse effectuée ci-dessus sur la répartition des distributions des charges annuelles de la communauté CF, se résume par les diagrammes en secteurs de ces charges (cf. Figure 5.2).

Notons que dans la suite des travaux, les charges prises en compte pour le calcul des couvertures et des contributions sont : les **franchises** et les **dépenses** annuelles des assurés de la communauté CF.

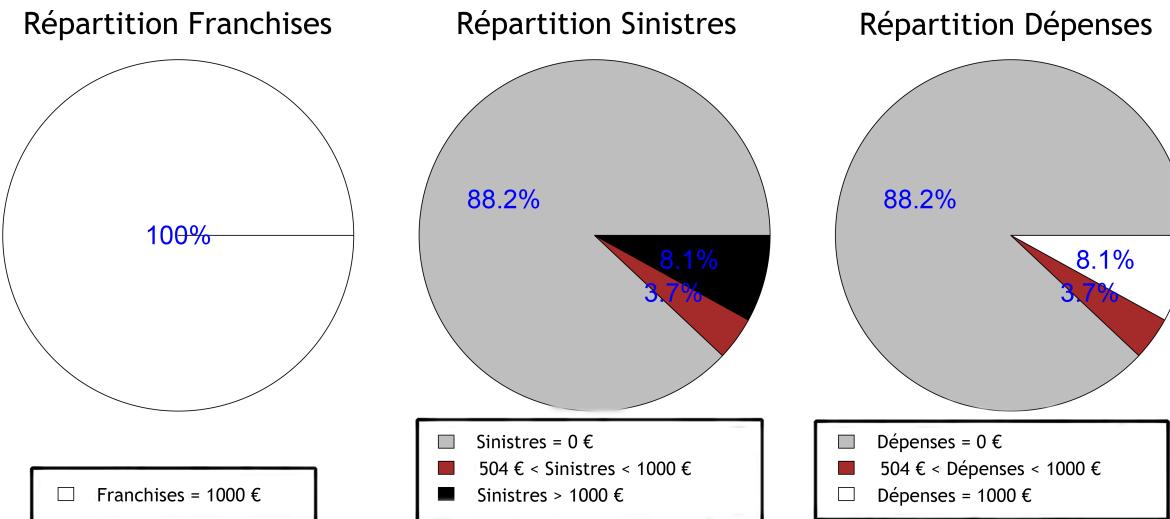


FIGURE 5.2 – Diagrammes en secteurs des charges de la communauté **CF**, obtenues par simulation : *Franchises* (a), *Coûts des sinistres* (b), *Dépenses* (c).

5.2 Calibration des paramètres

Pour rappel, les paramètres du produit de franchise collaborative étudié dans le cadre de ce mémoire sont : le montant des **engagements réciproques** entre les assurés et le modèle de **graphe social**. L'objectif de cette section est de présenter la calibration de ces paramètres qui font partie des entrées du programme R (détailé dans la section suivante) de calcul des distributions des variables d'intérêt (couverture et contribution) du produit de franchise collaborative.

5.2.1 Calibration du graphe social

Afin d'analyser l'impact du graphe social sur les variables d'intérêt du produit et prendre en compte les différents scénarios possibles de connexions entre les assurés, quatre modèles classiques de graphe de la théorie des réseaux (cf. section 2.4, pour plus de détails sur ces modèles) sont utilisés pour la calibration du graphe social représentant le réseau des relations entre les assurés de la communauté **CF**. Aussi, afin de pouvoir comparer les distributions de degrés obtenues pour chacun de ces modèles de graphe et le degré moyen de la communauté **CF** étant supposé dans la sous-section 5.1.1, égale à $\bar{d} = 20$ relations (i.e. en moyenne, chaque assuré possède 20 relations), les paramètres des différents modèles de graphe sont choisis de telle sorte à simuler des graphes de degré moyen proche de celui de la communauté **CF**.

Le modèle à clusters homogènes

La calibration du graphe social de la communauté **CF** via un modèle à "clusters homogènes" permet de prendre en compte les scénarios où :

- ✓ soit la connexion est **imposée** par l'assureur et celui-ci peut décider de rassembler les assurés dans des groupes indépendants avec des caractéristiques les plus similaires possibles entre les assurés appartenant à un même groupe (e.g. l'âge ou le profil de risque) ;
- ✓ soit la connexion est **libre** et les assurés peuvent se regrouper entre proches (parents ou amis), avec la contrainte selon laquelle : un assuré ne peut appartenir à plus d'un groupe.

Aussi, dans le but d'obtenir comme degré moyen $\bar{d} = 20$ et d'analyser l'impact du degré d'un assuré sur sa couverture de charges et sa contribution aux charges, les 493 membres de la com-

munauté **CF** sont répartis dans 29 clusters homogènes composés de 3 à 31 assurés, soit de degrés respectifs de 2 à 30 tels que : les k assurés composant le cluster de taille k possèdent chacun $(k - 1)$ relations, k variant de 3 à 31.

Sous le logiciel R, la fonction `make_full_graph` du package **igraph**, permettant la simulation d'un graphe complet est utilisée pour la calibration du graphe social de la communauté **CF** via le modèle à clusters homogènes, avec les lignes de code suivantes :

```

1 > library(igraph)
2 > size_max <- 31 #Taille maximale
3 > g_ch <- make_full_graph(size_max) #Graphe complet d'ordre size_max
4 > order <- size_max - 1 #Initialisation
5 > while(order>2){
6 >   g_full <- make_full_graph(order) #Graphe complet d'ordre order
7 >   g_ch <- g_ch + g_full
8 >   order <- order - 1 }
9 > degs_ch <- degree(g_ch) #Degrés des assurés
10 > hist(degs_ch,freq=F,col="yellow",xlab="Degrés",ylab="Densité",main="")
11 > lines(density(degs_ch),lwd=2,col="red") #Fonction de densité

```

La Figure 5.3 représente l'histogramme de la distribution des degrés des assurés de la communauté **CF** via le modèle à clusters homogènes⁵.

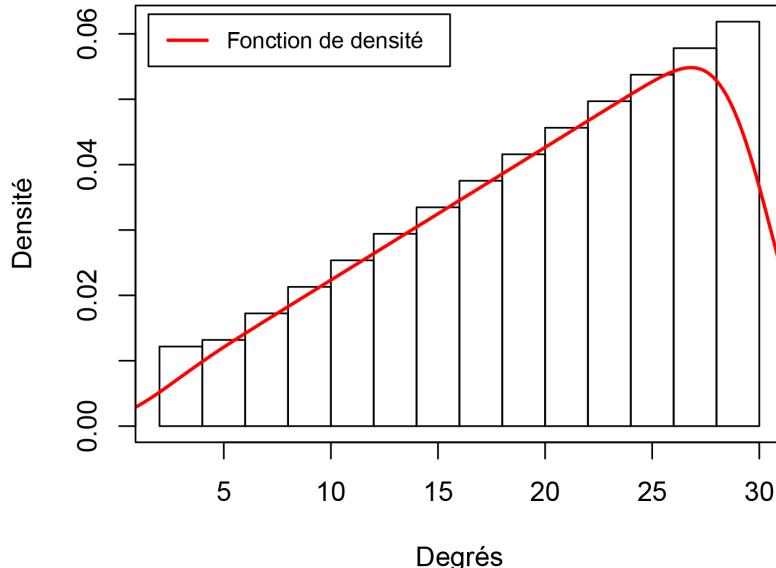


FIGURE 5.3 – Histogramme distribution de degrés : Modèle à clusters homogènes

Dans le cas où la connexion entre les assurés est **libre et sans contrainte** (i.e. les assurés peuvent se connecter selon leur bon vouloir, sans contrainte), il importe de prendre en compte tous les scénarios de connexions possibles. Pour ce faire, on s'intéresse à trois modèles classiques de graphe aléatoire de la théorie des réseaux présentés dans la section 2.4, à savoir :

5. Note : Dans toute la suite du mémoire, la décroissance finale de l'estimateur de la densité est dûe à l'utilisation d'un noyau symétrique (par défaut) dans la fonction `density` du logiciel R, alors que la distribution est à support compact (il y a un biais multiplicatif de l'ordre de 1/2 à la limite du support).

modèle d’Erdős-Rényi, le modèle de Watts-Strogatz et le modèle de Barabási-Albert.

Le modèle d’Erdős-Rényi

Le graphe social de la communauté **CF** est calibré via le modèle d’Erdős-Rényi sous la forme d’un graphe aléatoire $G = G(n, p)$ avec :

- ✓ $n = 493$, l’ordre du graphe (i.e. le nombre d’assurés) ;
- ✓ $p = \bar{d}/(n - 1) = 0.04$, la probabilité d’avoir une relation entre deux assurés, déterminée d’après la relation 2.3.

Sous le logiciel R, la fonction `erdos.renyi.game` du package **igraph** permet la simulation d’un tel graphe avec les lignes de code suivantes :

```
1 > library(igraph)
2 > set.seed(30)
3 > g_er<- erdos.renyi.game(493, 0.04)
4 > degs_er <- degree(g_er) #Degrés des assurés
5 > hist(degs_er,freq=F,col="green",xlab="Degrés",ylab="Densité",main="")
6 > lines(density(degs_er),lwd=2,col="red") #Fonction de densité
```

Le modèle d’Erdős-Rényi permet d’obtenir une distribution de degrés des assurés de la communauté **CF**, relativement homogène comme le montre la Figure 5.4.

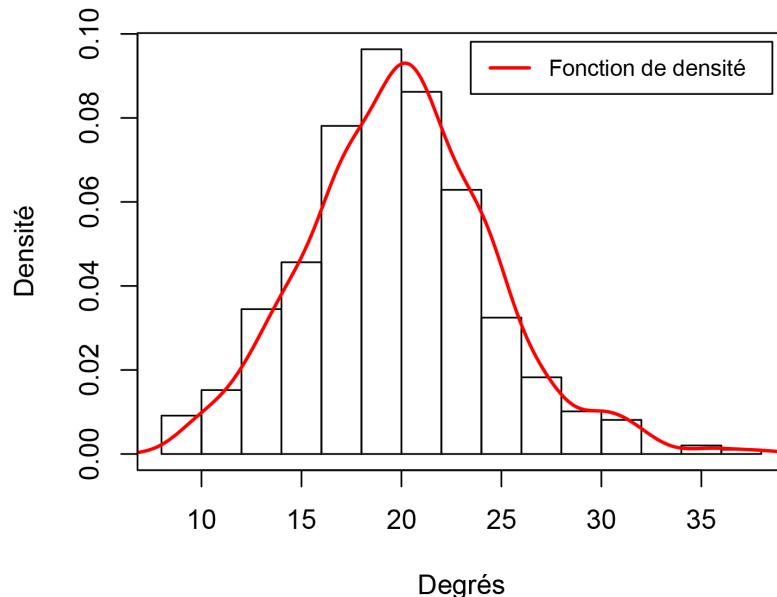


FIGURE 5.4 – Histogramme distribution de degrés : Modèle d’Erdős-Rényi

Le modèle du petit monde de Watts-Strogatz

La calibration du graphe social de la communauté **CF** est également effectuée via le modèle de Watts-Strogatz sous la forme d’un graphe aléatoire $G = G(\text{dim}, \text{size}, \text{nei}, p)$ avec :

- ✓ $\text{dim} = 1$, la dimension du réseau initial ;

- ✓ $\text{size} = 493$, la taille (ou l'ordre) du réseau le long de chaque dimension ;
- ✓ $\text{nei} = \bar{d}/2 = 10$, le nombre de voisins avec lesquels les assurés du réseau sont connectés à chaque étape de l'algorithme ;
- ✓ $p = 0,5$, la probabilité de reconstitution des relations (choisie arbitrairement dans $]0,1[$).

Par ailleurs, la fonction `watts.strogatz.game` du package **igraph** de R, permet la simulation d'un tel graphe comme-suit :

```
1 > library(igraph)
2 > set.seed(30)
3 > g_ws <- watts.strogatz.game(1, 493, 10, 0.5)
4 > degs_ws <- degree(g_ws) #Degrés des assurés
5 > hist(degs_ws,freq=F,col="lightblue",xlab="Degrés",ylab="Densité",main="")
6 > lines(density(degs_ws),lwd=2,col="red") #Fonction de densité
```

Le modèle de Watts-Strogatz permet d'obtenir une distribution de degrés plus homogène comme l'illustre la Figure 5.5.

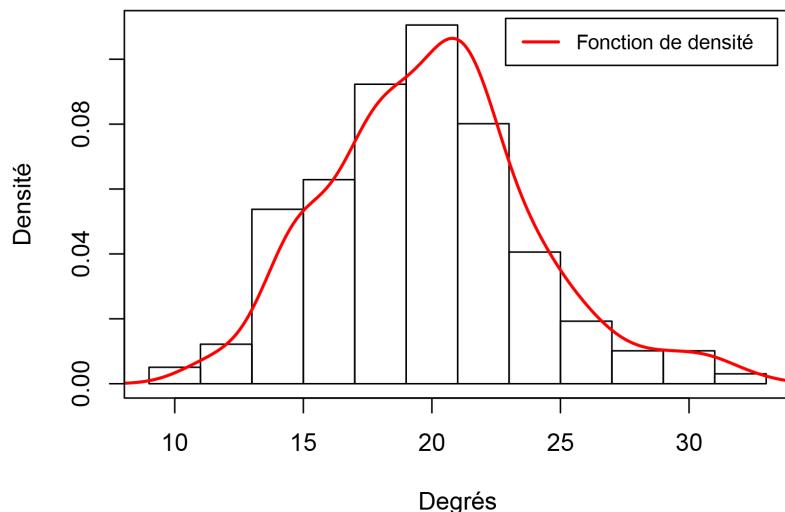


FIGURE 5.5 – Histogramme distribution de degrés : Modèle de Watts-Strogatz

Le modèle d'attachement préférentiel de Barabási-Albert

Le graphe social de la communauté **CF** est également calibré via le modèle de Barabási-Albert tel que $G = G(n, \text{power}, m, \text{directed} = \text{False}, \text{algorithm} = \text{"psumtree"})$ avec :

- ✓ $n = 493$, l'ordre du graphe (i.e. le nombre d'assurés) ;
- ✓ $\text{power} = 3$, la puissance ou le paramètre de l'attachement préférentiel ;
- ✓ $m = \bar{d}/2 = 10$, le nombre de relations se nouant à chaque étape de l'algorithme ;
- ✓ $\text{directed} = \text{False}$, permet d'obtenir un graphe non orienté ;
- ✓ $\text{algorithm} = \text{"psumtree"}$, algorithme d'implémentation sélectionné.

Sous le logiciel R, la fonction `barabasi.game` du package **igraph** permet la simulation d'un tel graphe avec les lignes de code suivantes :

```

1 > library(igraph)
2 > set.seed(30)
3 > g_ba <- barabasi.game(493, 3, 10, directed = False, algorithm = "psumtree")
4 > degs_ba <- degree(g_ba) #Degrés des assurés
5 > hist(degs_ba,freq=F,col="gray",xlab="Degrés",ylab="Densité",main="")
6 > lines(density(degs_ba),lwd=2,col="red") #Fonction de densité

```

Par ailleurs, le modèle de Barabási-Albert permet d'obtenir une distribution de degrés relativement hétérogène comme le montre la Figure 5.6.

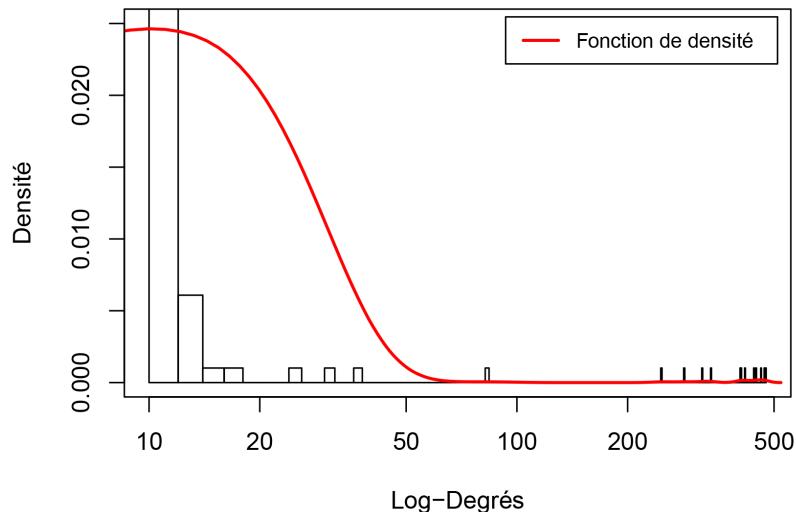


FIGURE 5.6 – Histogramme distribution de degrés en échelle log : Modèle de Barabási-Albert

Récapitulatif des distributions de degrés

La calibration du graphe social de la communauté **CF** via les quatre modèles de graphe, présentée ci-haut, permet de prendre en compte les différents scénarios possibles de connexions entre les assurés et d'avoir des distributions de degrés de variances relativement distinctes. Aussi, le Tableau 5.1 récapitule les quatre distributions de degrés des assurés, obtenues via chaque modèle de graphe et représentées respectivement par les Figures 5.3, 5.4, 5.5 et 5.6.

On constate que les modèles de Watts-Strogatz et d'Erdős-Rényi permettent d'obtenir des distributions de degrés "relativement homogènes", c'est-à-dire de variances minimales, soit respectivement de 16,21 et de 21,37. Aussi, le modèle à clusters homogènes permet quant à lui d'obtenir une distribution de degrés "homogène par clusters", soit de variance 53,15. Enfin, le modèle de Barabási-Albert conduit à une distribution de degrés "relativement hétérogène", c'est-à-dire de variance maximale, soit de 3.630,06.

Le diagramme en bâtons comparatif des distributions de degrés (Figure 5.7) illustre les différences de dispersion obtenues par la calibration du graphe social de la communauté **CF** via ces quatre modèles de graphe.

Modèles	Clusters homogènes	Erdős-Rényi	Watts-Strogatz	Barabási-Albert
Minimum	2	9	10	10
Maximum	30	38	33	475
Moyenne	20,12	20,02	20	19,78
Variance	53,15	21,37	16,21	3.630,06

TABLE 5.1 – Tableau récapitulatif : Distributions de degrés par modèle de graphe

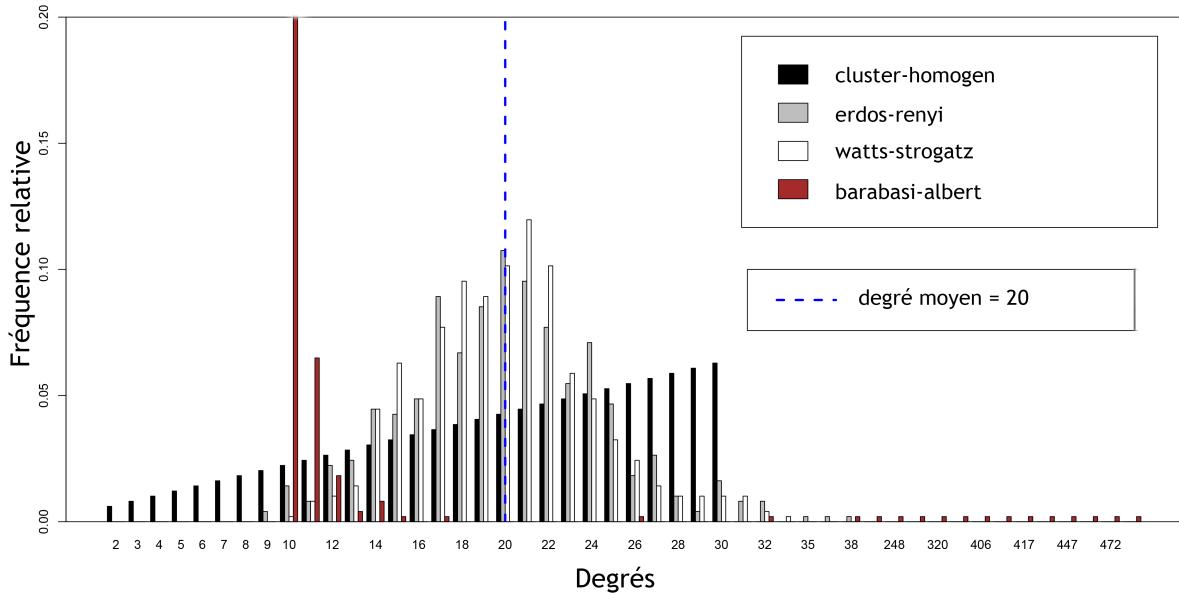


FIGURE 5.7 – Diagramme en bâtons : Distributions de degrés par modèle de graphe

Pour rappel, notons que ces distributions des degrés des assurés, obtenues sur la base d'une simulation pour chaque modèle de graphe, ne sont pas généralisables. Toutefois, elles permettront dans les sections 5.4 et 5.5, d'analyser l'impact du degré d'un assuré sur respectivement sa couverture de charges et sa contribution aux charges, pour chacun des quatre modèles de graphe étudiés.

5.2.2 Calibration des engagements réciproques

D'après la définition des engagements réciproques entre les assurés, leur calibration nécessite avant tout, la détermination de la matrice d'adjacence du graphe social de la communauté des assurés.

De ce fait, partant de la calibration du graphe social (présentée dans la sous-section 5.2.1), on déduit la matrice d'adjacence R qui est une matrice de taille $n \times n$, avec $n = 493$, le nombre d'assurés de la communauté CF et telle que :

$$\forall \text{ assurés } i, j = 1 : 493, R_{ij} = \begin{cases} 1 & , \text{ si les assurés } i \text{ et } j \text{ sont en relation} \\ 0 & , \text{ sinon.} \end{cases}$$

Sous le logiciel R, la fonction `get.adjacency` du package **igraph**, prenant un graphe en argument, permet d'obtenir la matrice d'adjacence de ce graphe.

Ensuite, on déduit la matrice E des engagements réciproques entre les assurés, sous les hypothèses **(HDB)** et **(HCO)** telle que :

(HDB) : $E_{ij}^{\text{hdb}} = \frac{F}{d} \times R_{ij} = 50 \times R_{ij}$, avec $F = 1.000 \text{ €}$, la franchise commune et $d = 20$, le degré moyen de la communauté **CF** ;

(HCO) : $E_{ij}^{\text{hco}} = \frac{F}{d_j} \times R_{ij}$, avec $F = 1.000 \text{ €}$, la franchise commune de la communauté **CF** et $d_j = \sum_{i=1}^n R_{ij}$, le degré de l'assuré $j = 1 : 493$.

5.3 Programme de calcul PC1

Cette section présente le programme **PC1** implémenté sous le logiciel R pour le calcul des distributions des variables d'intérêt du produit de franchise collaborative, notamment les variables de couverture des charges et de contribution aux charges des assurés de la communauté fictive **CF**.

5.3.1 Les entrées du programme PC1

Les entrées du programme de calcul **PC1** sont : les **charges** annuelles de la communauté **CF** (présentées dans la section 5.1) et les **paramètres** du produit de franchise collaborative (présentés dans la section 5.2).

5.3.2 Les sorties du programme PC1

Les sorties du programme de calcul **PC1** sont les distributions des **couvertures individuelles** des charges et des **contributions individuelles** aux charges des assurés de la communauté **CF**.

5.3.3 Le fonctionnement du programme PC1

Les étapes du programme de calcul **PC1** s'établissent comme-suit :

- ❖ **Étape 1** : Partant des entrées (charges annuelles de la communauté **CF**, modèles de graphe social et hypothèses sur les engagements réciproques) du programme, les matrices R d'adjacence, E des engagements réciproques et C des contributions réciproques aux dépenses entre les assurés, sont respectivement déterminées d'après leurs expressions mathématiques, pour chaque hypothèse sur le modèle de graphe social et sur les engagements réciproques ;
- ❖ **Étape 2** : La connaissance des matrices des engagements réciproques et des contributions réciproques aux dépenses entre les assurés, permet de déduire les distributions des couvertures individuelles des charges et des contributions individuelles aux charges, selon leurs expressions mathématiques ;
- ❖ **Étape 3** : Pour le calcul des indicateurs (cf. sections 5.4 et 5.5) permettant la comparaison et la sélection de graphes, on effectue 1.000 simulations de scénarios de graphe social pour chacun des trois modèles de graphe aléatoire (le scénario du modèle à clusters homogènes étant déterministe et présenté en 5.2.1). Pour chaque scénario, les couvertures individuelles et les contributions individuelles des assurés, sont calculées d'après leurs expressions mathématiques.

Afin d'obtenir les résultats pour chaque assuré, une moyenne sur l'ensemble des scénarios est effectuée, d'après la méthode de Monte-Carlo. On obtient donc une distribution des couvertures individuelles des charges et une distribution des contributions individuelles aux charges pour chacun des quatre modèles de graphe étudiés.

Enfin, les indicateurs sont calculés sur la base des distributions obtenues en sortie du programme.

5.4 Étude de couverture

L'objectif de cette section consiste à présenter et à analyser des sorties du programme de calcul **PC1**, notamment les distributions des couvertures individuelles des charges annuelles des assurés de la communauté **CF**, calculées sous les hypothèses (**HDB**) et pour chacun des quatre modèles de graphe étudiés.

En effet, l'étude effectuée au chapitre 4 a permis de montrer que les hypothèses (**HDB**) sur les engagements réciproques entre les assurés, peuvent entraîner des disparités de couverture individuelle en fonction de la variance de la distribution des degrés des assurés. De ce fait, sur la base des distributions de couverture obtenues en sortie du programme **PC1** pour chacun des quatre modèles de graphe étudiés et sous les hypothèses (**HDB**), une étude de couverture est effectuée afin d'évaluer l'impact du modèle de graphe social (via les degrés des assurés) sur les couvertures individuelles des charges annuelles des assurés de la communauté **CF**.

Cette étude de couverture consiste dans un premier temps, à analyser la couverture des charges d'un assuré en fonction de son degré pour chacun des quatre modèles de graphe et dans un second temps, à comparer et sélectionner le modèle de graphe permettant d'optimiser les couvertures individuelles des charges des membres du portefeuille **CF**.

5.4.1 Analyse de la couverture en fonction du degré

Dans cette sous-section, il s'agit d'une part de présenter les couvertures individuelles de charges annuelles obtenues pour chaque modèle de graphe (sous les hypothèses (**HDB**)), sur la base des distributions de degrés présentées en 5.2.1 (Calibration du graphe social) et d'autre part d'analyser l'impact du degré d'un assuré sur les couvertures de ses charges annuelles.

Distributions des couvertures individuelles sous (**HDB**)

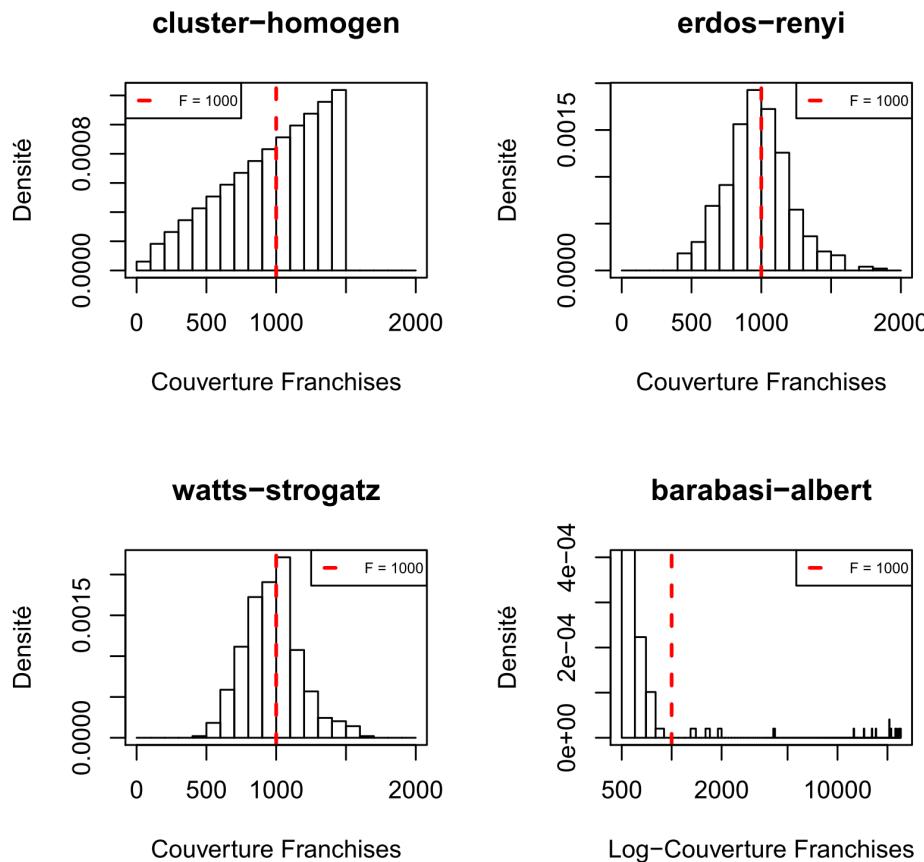
Partant des distributions de degrés présentés en 5.2.1 et obtenues par calibration du graphe social de la communauté **CF** via chacun des quatre modèles de graphe étudiés, on déduit les couvertures individuelles des charges (franchises et dépenses) annuelles des assurés sous les hypothèses (**HDB**), d'après leurs expressions mathématiques (cf. relations 4.5 et 4.6), pour chaque modèle de graphe.

❖ **Couverture des franchises :** Le Tableau 5.2 et la Figure 5.8 représentent respectivement le tableau récapitulatif et les histogrammes des distributions de couverture des franchises obtenues sous les hypothèses (**HDB**), pour chacun des quatre modèles de graphe étudiés.

La couverture individuelle de franchise obtenue, via le modèle à clusters homogènes, varie de 99,42 € à 1.491,23 € avec une valeur moyenne de $1.000 \pm 362,40$ € (cf. Tableau 5.2, 2^{ème} colonne et Figure 5.8 (a)). Le modèle d'Erdős-Rényi permet quant à lui, d'obtenir une couverture individuelle de franchise comprise entre 449,50 € et 1.898,10 € avec une valeur moyenne de $1.000 \pm 230,90$ € (cf. Tableau 5.2, 3^{ème} colonne et Figure 5.8 (b)). Aussi, la couverture individuelle de franchise obtenue, via le modèle de Watts-Strogatz, part de 500 € à 1.650 € avec une valeur moyenne de $1.000 \pm 201,34$ € (cf. Tableau 5.2, 4^{ème} colonne et Figure 5.8 (c)). Enfin, on obtient une couverture individuelle de franchise allant de 505,60 € à 24.017,90 € avec une valeur moyenne de $1.000 \pm 3.046,48$ €, pour le modèle de Barabási-Albert (cf. Tableau 5.2, 5^{ème} colonne et Figure 5.8 (d)).

Modèles	Clusters homogènes	Erdős-Rényi	Watts-Strogatz	Barabási-Albert
Min. (en €)	99,42	449,50	500	505,60
Max. (en €)	1.491,23	1.898,10	1.650	24.017,90
Moy. (en €)	1.000	1.000	1.000	1.000
Sd (en €)	362,40	230,90	201,34	3.046,48

TABLE 5.2 – Tableau : Distributions de couverture des franchises par modèle de graphe

FIGURE 5.8 – Histogrammes des distributions de couverture des franchises par modèle de graphe (F représente la franchise individuelle des assurés de la communauté **CF**) : *Clusters homogènes* (a), *Erdős-Rényi* (b), *Watts-Strogatz* (c) et *Barabási-Albert* en échelle log (d).

En particulier, on constate que quelque soit le modèle de graphe, la couverture moyenne de franchise coïncide avec les 1.000 € de franchise moyenne de la communauté **CF**. Ceci peut s'expliquer par la relation 4.7 exprimant l'égalité entre la couverture mutuelle (i.e. somme des couvertures individuelles) des franchises et la franchise mutuelle (i.e. somme des franchises individuelles) de la communauté **CF**, sous les hypothèses (**HDB**). Ce qui permet d'avoir également l'égalité entre les valeurs moyennes. Aussi, sur les quatre modèles de graphe, la couverture de franchise minimale évaluée à 99,42 €, est observée via la calibration du graphe social par le modèle à clusters homogènes. En effet, ce montant de couverture correspond à celui des assurés composant le cluster de taille minimale (i.e. composé de 3 assurés), soit de degré 2 par assuré (cf. Tableau 5.1). Sous les conditions (**HDB**), l'engagement réciproque étant de $1.000/20,12 = 49,70$ €, on obtient donc une couverture de 99,42 € pour un assuré possédant 2 relations. Cependant, la couverture de franchise maximale étant de 24.017,90 €, est observée via le modèle de Barabási-Albert. En effet, l'assuré le plus populaire (i.e assuré ayant un degré élevé) via ce modèle, possède

475 relations (cf. Tableau 5.1), soit une couverture de franchise de $475 \times 1.000 / 19,78 = 24.017,90$ € (d'après la relation 4.5).

Par ailleurs, les écart-types des distributions de couverture (cf. Tableau 5.2) obtenus pour chacun des quatre modèles permettent de conclure que les distributions de couverture conservent relativement les mêmes dispersions que les distributions de degrés correspondantes, c'est-à-dire "homogène par cluster" pour le modèle à clusters homogènes, "relativement homogène" pour les modèles d'Erdős-Rényi et de Watts-Strogatz et "relativement hétérogène" pour le modèle de Barabási-Albert. Pour ces mêmes raisons, on constate que les histogrammes des distributions de couverture des franchises (cf. Figure 5.8), obtenus pour chaque modèle de graphe conservent relativement les mêmes formes que les histogrammes des distributions de degrés correspondantes.

Cette influence du degré d'un assuré sur la couverture de sa franchise peut s'expliquer par la relation 4.5, exprimant linéairement la couverture de franchise d'un assuré en fonction de son degré, sous les hypothèses (**HDB**). Cette formule fermée existant entre la couverture de franchise et le degré d'un assuré, est illustrée dans le point suivant (cf. Couverture en fonction du degré).

❖ **Couverture des dépenses :** Le Tableau 5.3 et la Figure 5.9 représentent respectivement le tableau récapitulatif et les histogrammes des distributions de couverture des dépenses pour chacun des quatre modèles de graphe étudiés.

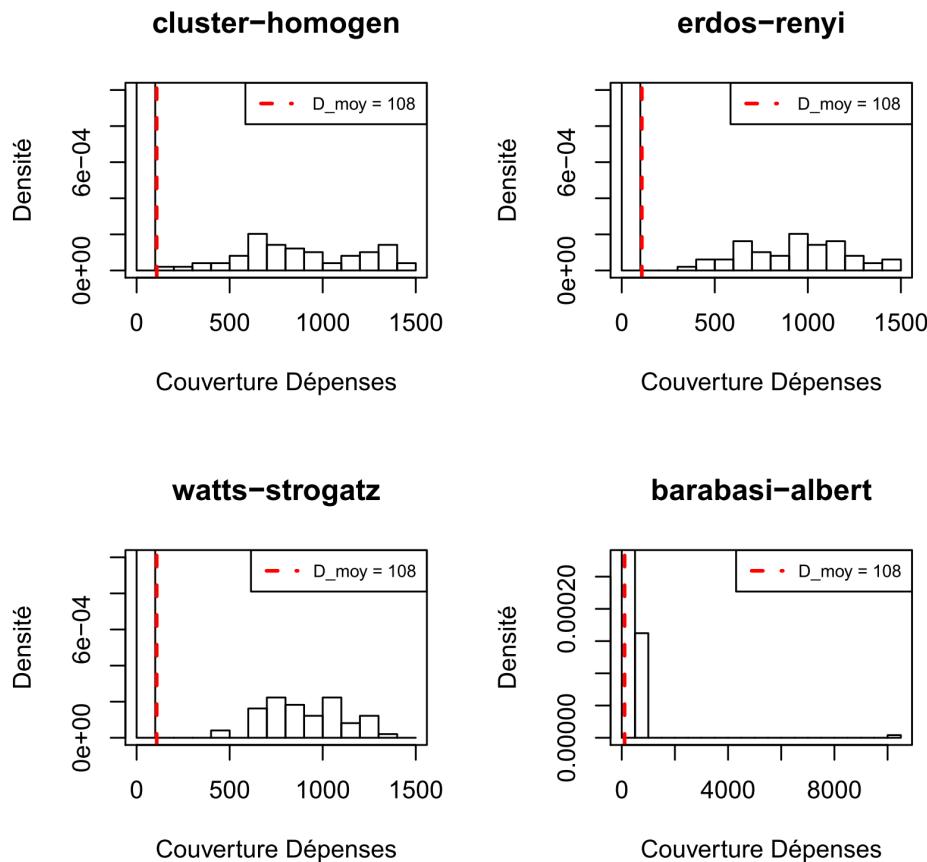
La couverture individuelle de dépense obtenue, via le modèle à clusters homogènes, varie de 0 € à 1.491,23 € avec une valeur moyenne de $105,20 \pm 309,63$ € (cf. Tableau 5.3, 2^{ème} colonne et Figure 5.9 (a)). Le modèle d'Erdős-Rényi permet quant à lui, d'obtenir une couverture individuelle de dépense comprise entre 0 € et 1.498,50 € avec une valeur moyenne de $110,9 \pm 318,91$ € (cf. Tableau 5.3, 3^{ème} colonne et Figure 5.9 (b)). Aussi, la couverture individuelle de dépense obtenue, via le modèle de Watts-Strogatz, part de 0 € à 1.350 € avec une valeur moyenne de $109,10 \pm 307,74$ € (cf. Tableau 5.3, 4^{ème} colonne et Figure 5.9 (c)). Enfin, on obtient une couverture individuelle de dépense allant de 0 € à 10.293,71 € avec une valeur moyenne de $75,45 \pm 486,05$ €, pour le modèle de Barabási-Albert (cf. Tableau 5.3, 5^{ème} colonne et Figure 5.9 (d)).

Ensuite, l'allure (i.e. la forte fréquence en 0) des histogrammes des distributions de couverture des dépenses (cf. Figure 5.9) peut s'expliquer par le fait que 88,2% des assurés n'ont pas "déclarés" de sinistres dans l'année, donc ont des dépenses annuelles nulles, d'où ont des couvertures de dépenses nulles. Aussi, comme annoncé par la relation 4.9, on constate que les couvertures moyennes de dépenses étant de 105,20 €, de 110,90 €, de 109,10 € et de 75,45 €, respectivement pour les modèles à clusters homogènes, d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert (cf. Tableau 5.3), ne coïncident pas nécessairement avec la dépense moyenne de la communauté **CF** qui est de 108 €.

En particulier, sur les quatre modèles de graphe, le modèle de Barabási-Albert fournit la distribution de couverture de dépenses la plus volatile avec un écart-type de 486,05 € et une couverture maximum de 10.293,71 €, tandis que le modèle de Watts-Strogatz enregistre la distribution de couverture de dépenses la moins volatile avec un écart-type de 307,74 € et une couverture maximum de 1.350 €, soit des diminutions respectives de 37% et de 87% par rapport au modèle de Barabási-Albert.

Modèles	Clusters homogènes	Erdős-Rényi	Watts-Strogatz	Barabási-Albert
Min. (en €)	0	0	0	0
Max. (en €)	1.491,2	1.498,5	1.350	10.293,71
Moy. (en €)	105,2	110,9	109,1	75,45
Sd (en €)	309,63	318,91	307,74	486,05

TABLE 5.3 – Tableau : Distributions de couverture des dépenses par modèle de graphe

FIGURE 5.9 – Histogrammes des distributions de couverture des dépenses par modèle de graphe (D_{moy} représente la dépense moyenne des assurés de la communauté **CF**) : *Clusters homogènes* (a), *Erdős-Rényi* (b), *Watts-Strogatz* (c) et *Barabási-Albert* (d).

Enfin, après avoir présenté les distributions de couverture des charges obtenues pour chaque modèle de graphe, sous les conditions (**HDB**), le point suivant consiste à analyser l'évolution de ces couvertures individuelles en fonction des distributions de degrés issues de la calibration du graphe social via les quatre modèles de graphe étudiés.

Couverture en fonction du degré

La Figure 5.10 représente l'évolution de la couverture de franchise d'un assuré en fonction de son degré, pour chaque modèle de graphe. On constate que la Figure 5.10 illustre clairement la relation 4.5, exprimant linéairement la couverture de franchise d'un assuré en fonction de son degré sous les hypothèses (**HDB**), telle que :

$$y = \frac{F}{\bar{d}} \times x = 50 \times x \quad (5.1)$$

avec x et y représentant respectivement le degré et la couverture individuelle de franchise d'un assuré, $F = 1.000 \text{ €}$ et $\bar{d} = 20$ représentant respectivement la franchise commune et le degré moyen de la communauté **CF**.

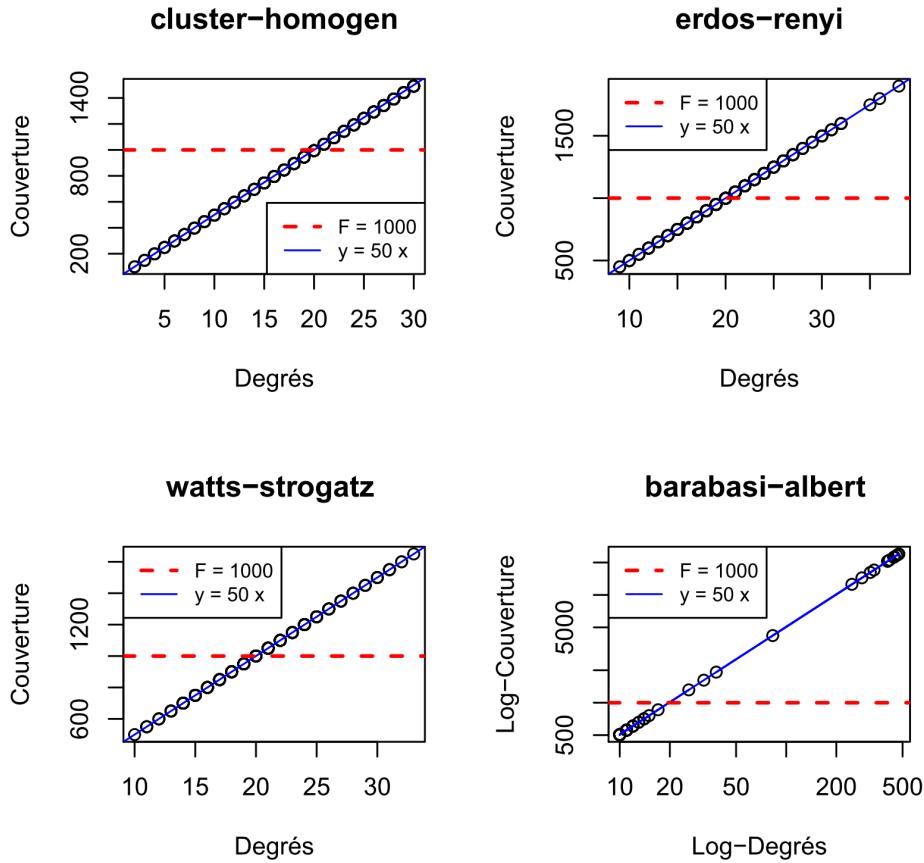


FIGURE 5.10 – Couverture de franchise en fonction du degré par modèle de graphe : *Clusters homogènes* (a), *Erdős-Rényi* (b), *Watts-Strogatz* (c) et *Barabási-Albert* en échelle log-log (d).

Aussi, la disparité de couverture individuelle annoncée sous les hypothèses (**HDB**) est illustrée par le fait que les assurés ayant un degré inférieur au degré moyen $\bar{d} = 20$ de la communauté **CF**, ont des couvertures individuelles de franchises en dessous du niveau de franchise individuelle de 1.000 € représenté par la bande "rouge" sur la Figure 5.10. En effet, sur les 493 assurés, le nombre d'assurés ayant un degré inférieur à la moyenne est de 228 soit 46%, de 275 soit 56%, de 273 soit 55% et de 477 soit 97% respectivement pour les modèles à clusters homogènes, d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert.

Par ailleurs, d'après la relation 4.6, la couverture de dépense d'un assuré est non seulement fonction de son degré mais également de cette dépense. Il n'est donc pas possible d'obtenir comme pour la couverture de franchise, une formule fermée exprimant la couverture de dépense d'un assuré uniquement en fonction de son degré.

Toutefois, rappelons que cette étude de couverture effectuée, en fonction des degrés des assurés se base essentiellement sur les distributions de degrés résumées dans le Tableau 5.1 et obtenues par simulation d'un scénario de graphe social pour chaque modèle de graphe. Dans la sous-section suivante, afin d'obtenir des résultats plus robustes et généralisables pour les couvertures individuelles et d'évaluer la performance des couvertures issues de chaque modèle de graphe,

la méthode de Monte-Carlo est utilisée avec 1.000 simulations de scénarios de graphe social pour chacun des trois modèles de graphe aléatoire, sachant que le scénario du modèle à clusters homogènes est déterministe et est présenté en 5.2.1.

5.4.2 Sélection de modèle de graphe

Cette sous-section présente les couvertures individuelles des charges obtenues en sortie du programme **PC1** pour chaque modèle de graphe et sous les hypothèses (**HDB**). Aussi, afin d'analyser et de comparer la performance des couvertures obtenues pour chacun de ces quatre modèles de graphe, on introduit les indicateurs de couverture suivants.

Définition 9 (Indicateurs de couverture).

- ❖ **Le ratio de couverture (RC)** : qui représente pour un assuré, le rapport entre sa couverture individuelle et sa charge individuelle (non-nulle), tel que :

$$\forall \text{ assuré } j = 1 : n, \quad RC_j = \frac{Couv_j}{Ch_j},$$

où $Couv_j$ représente la couverture de franchise ou de dépense et Ch_j représente respectivement la franchise ou la dépense (non-nulle) de l'assuré $j = 1 : n$.

Pour tout assuré $j = 1 : n$, si son ratio de couverture vérifie : $RC_j \geq 1$, alors : $Couv_j \geq Ch_j$. On dit que la charge de cet assuré est "**couverte**".

- ❖ **Le taux de couverture (TC)** : qui mesure la proportion d'assurés dont les charges sont "couvertes" pour un modèle de graphe donné tel que :

$$TC = \frac{1}{m} \sum_{j=1}^m \mathbb{1}_{\{RC_j \geq 1\}},$$

pour les $1 \leq m \leq n$ charges Ch_j non-nulles, ou de manière générale, en considérant les n charges individuelles des assurés :

$$TC = \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{Couv_j \geq Ch_j\}}$$

Cet indicateur compris dans l'intervalle $[0, 1]$ est à maximiser.

- ❖ **La couverture moyenne** : qui représente la moyenne de la distribution de couverture obtenue pour un modèle de graphe donné. Cet indicateur est à maximiser.
- ❖ **La volatilité de couverture** : qui mesure la dispersion des valeurs de la distribution de couverture obtenue pour un modèle de graphe donné. Cet indicateur évalué à l'aide de l'**écart-type**, est à minimiser.

Pour rappel, pour des raisons de temps de calcul, les distributions de couverture présentées dans cette sous-section, ont été obtenues sur la base de 1.000 simulations de scénario de graphe social pour les trois modèles de graphe aléatoire et du scénario déterministe du modèle à clusters homogènes, présenté en 5.2.1.

L'objectif consiste à déterminer les indicateurs de couverture issus de ces différentes distributions, afin de comparer les performances de couverture obtenues pour chaque modèle de graphe et *in fine*, de sélectionner le modèle permettant d'optimiser les couvertures individuelles des charges des assurés de la communauté **CF**. Pour ce faire, on s'intéresse dans un premier temps,

aux ratios de couverture associés aux charges non-nulles et dans un second temps, aux autres indicateurs de couverture (taux de couverture, couverture moyenne et volatilité de couverture) associés à l'ensemble des charges du portefeuille **CF**.

Distributions des ratios de couverture

Pour rappel, le ratio de couverture de charge d'un assuré consiste au rapport entre sa couverture individuelle et sa charge individuelle (non-nulle). De ce fait, partant des charges annuelles de la communauté **CF** et des distributions de couverture de charges obtenues en sortie du programme **PC1**, on déduit les ratios de couverture des charges de chaque assuré. On dispose donc de distributions de ratios de couverture respectivement associées aux franchises et aux dépenses non-nulles des assurés, pour chaque modèle de graphe.

❖ **Ratios de couverture des franchises** : La Figure 5.11 représente les fonctions de densité des ratios de couverture des franchises des assurés pour chaque modèle de graphe.

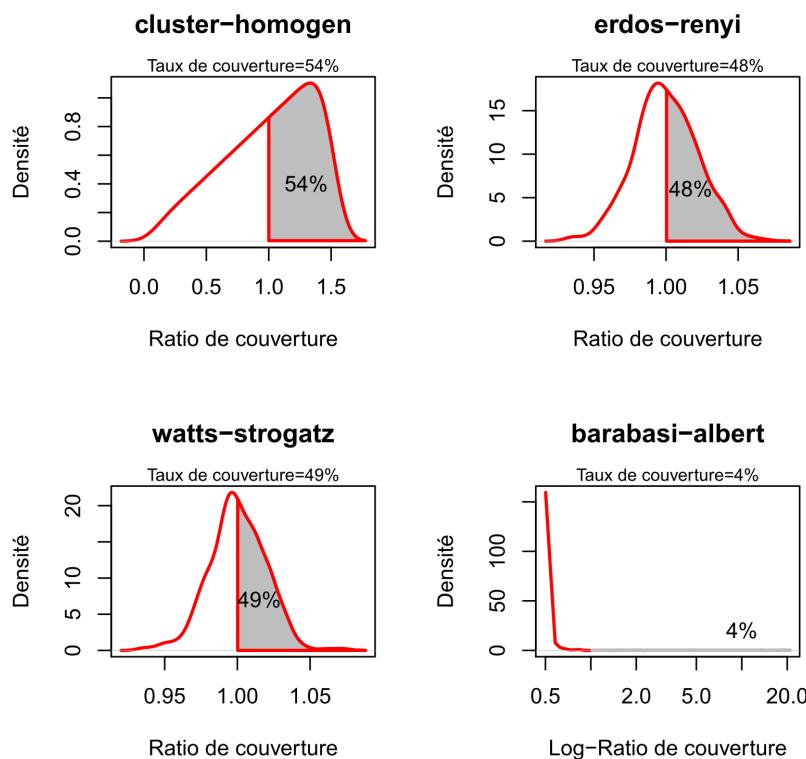


FIGURE 5.11 – Fonctions de densité des ratios de couverture des franchises par modèle de graphe : *Clusters homogènes* (a), *Erdős-Rényi* (b), *Watts-Strogatz* (c) et *Barabási-Albert* en échelle log (d).

Sur la Figure 5.11, les aires des parties colorées en "gris" représentent les taux de couverture de franchises (i.e. le pourcentage d'assurés ayant leurs franchises couvertes), respectivement pour les modèles à clusters homogènes, d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert. On constate par exemple que, le taux de couverture le plus faible étant de 4% (i.e. seulement 18 assurés sur 493 ont leurs franchises couvertes), est obtenu via le modèle de Barabási-Albert. Ce qui explique le pic de la fonction densité au voisinage de 1 (cf. Figure 5.11 (d)), sachant que 96% des ratios individuels de couverture sont inférieures à 1. Tandis que, le modèle à clusters homogènes fournit le taux de couverture le plus élevé, estimé à 54%, c'est-à-dire, les franchises de 265 assurés sont couvertes via ce modèle, soit environ 15 fois plus d'assurés couverts par rapport

au modèle de Barabási-Albert.

❖ **Ratios de couverture des dépenses :** La Figure 5.12 représente les fonctions de densité des ratios de couverture des dépenses non-nulles des assurés pour chaque modèle de graphe. En effet, d'après la définition du ratio de couverture, il prend en compte uniquement les charges non-nulles, c'est-à-dire les dépenses des 58 assurés sinistrés, soit 11,8% des 493 assurés du portefeuille **CF**.

De même que sur la Figure 5.11 relative aux franchises, sur la Figure 5.12 relative aux dépenses non-nulles des assurés, les aires des parties colorées en "gris" représentent les taux de couverture des dépenses non-nulles (i.e. le pourcentage d'assurés sinistrés ayant leurs dépenses couvertes), respectivement pour les modèles à clusters homogènes, d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert. Par exemple, les modèles d'Erdős-Rényi et de Watts-Strogatz fournissent les meilleurs taux de couverture, soit respectivement de 52% (i.e. 30 assurés sur les 58 assurés sinistrés ont leurs dépenses couvertes) et de 57% (i.e. 33 assurés sur les 58 assurés sinistrés ont leurs dépenses couvertes).

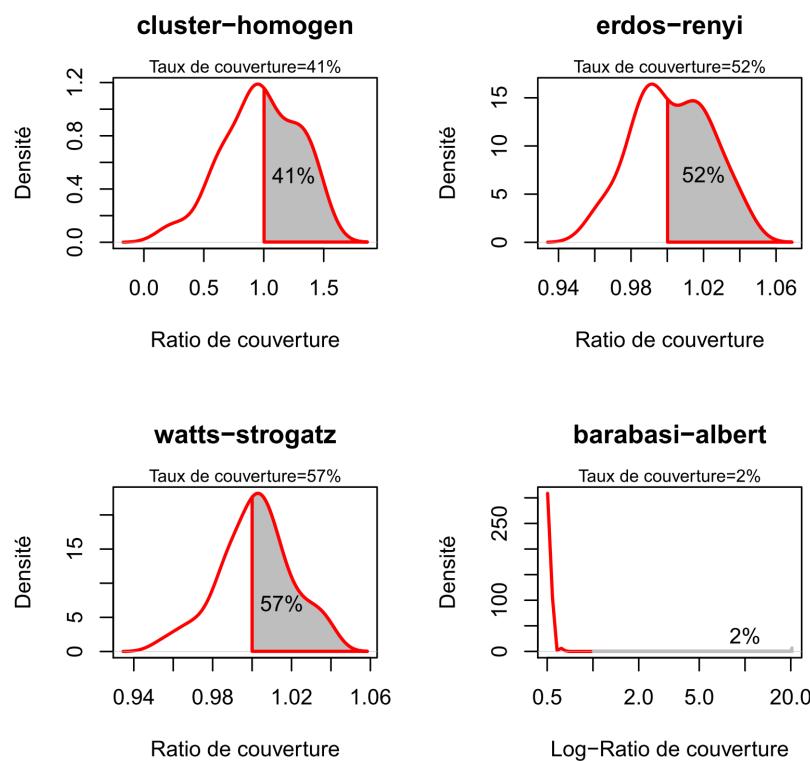


FIGURE 5.12 – Fonctions de densité : Ratios de couverture des dépenses par modèle de graphe : *Clusters homogènes* (a), *Erdős-Rényi* (b), *Watts-Strogatz* (c) et *Barabási-Albert* en échelle log (d).

Rappelons que ces ratios de couverture de dépenses ont été obtenus en considérant uniquement les 58 membres de la communauté **CF** possédant des dépenses annuelles non-nulles. Le point suivant fait un récapitulatif des indicateurs de couverture calculés en prenant en compte toutes les charges annuelles des assurés du portefeuille **CF**.

Récapitulatif des indicateurs de couverture

Le Tableau 5.4 résume les indicateurs de couverture des charges calculés d'après leurs définitions sur la base des charges de la communauté **CF** et des couvertures individuelles obtenues en sortie du programme **PC1** pour chaque modèle de graphe. En effet, ces indicateurs consistant en la **couverture moyenne** (notée "Moy." dans le Tableau 5.4), la **volatilité de couverture** (notée "Vol." dans le Tableau 5.4) et le **taux de couverture** (noté "Taux" dans le Tableau 5.4) sont calculés pour chacune des distributions de couverture obtenue par modèle de graphe et sous les hypothèses (**HDB**).

Indicateurs de couverture	Franchises ($F = 1.000 \text{ €}$)			Dépenses ($\bar{D} = 108 \text{ €}$)		
	Modèles	Moy	Vol	Taux	Moy	Vol
Clusters homogènes	1.000	362,40	54	105,20	309,63	93
Erdős-Rényi	999,84	21,99	48	108,24	301,09	94
Watts-Strogatz	1.000,20	19,43	49	108,07	300,46	95
Barabási-Albert	995,51	2.847,20	4	88,31	764,73	88,4

TABLE 5.4 – Tableau récapitulatif : Indicateurs de couverture par modèle de graphe. Taux = taux de couverture (en %), Moy = couverture moyenne (en €) et Vol = volatilité ou écart-type de couverture (en €).

Les distributions des couvertures individuelles des charges (franchises et dépenses) annuelles des assurés de la communauté **CF**, obtenues en sortie du programme **PC1** sous les hypothèses (**HDB**) et sur la base de 1.000 simulations de scénarios de graphe social via les trois modèles de graphe aléatoire et du scénario déterministe de graphe social via le modèle à clusters homogènes, s'établissent comme-suit :

- ❖ **Couverture des franchises** : Comme présenté en 5.4.1, le scénario déterministe de graphe social simulé via le modèle à clusters homogènes entraîne une couverture individuelle de franchise variant de 99,42 € à 1.491,23 € avec une couverture moyenne de $1.000 \pm 362,40 \text{ €}$ et un taux de couverture de 54%. En effet, via ce modèle de graphe et sous les hypothèses (**HDB**), les franchises de 265 assurés sur les 493 sont couvertes (cf. Tableau 5.4, 3^{ème} ligne et volet Fanchises). Le modèle d'Erdős-Rényi permet quant à lui, d'obtenir une couverture individuelle de franchise comprise entre 933,60 € et 1.068,70 € avec une couverture moyenne de $999,84 \pm 21,99 \text{ €}$ et un taux de couverture de 48% (i.e. 236 assurés sur les 493 ont leurs franchises couvertes). Aussi, la couverture individuelle de franchise obtenue, via le modèle de Watts-Strogatz, part de 934 € à 1.074 € avec une couverture moyenne de $1.000,20 \pm 19,43 \text{ €}$ et un taux de couverture de 49% (i.e. 244 assurés sur les 493 ont leurs franchises couvertes). Enfin, on obtient une couverture individuelle de franchise allant de 505,60 € à 20.818,30 € avec une couverture moyenne de $995,51 \pm 2.847,20 \text{ €}$ et un taux de couverture de 4%, pour le modèle de Barabási-Albert (i.e. seulement 18 assurés sur les 493 ont leurs franchises couvertes) ;
- ❖ **Couverture des dépenses** : Comme présenté en 5.4.1, le scénario déterministe de graphe social simulé via le modèle à clusters homogènes entraîne une couverture individuelle de dépense comprise entre 0 € et 1.491,23 € avec une couverture moyenne de $105,20 \pm 309,63 \text{ €}$ et un taux de couverture de 93%. En effet, via ce modèle de graphe et sous les hypothèses (**HDB**), les dépenses de 459 assurés sur les 493 sont couvertes (cf. Tableau 5.4, 3^{ème} ligne et volet Dépenses). Le modèle d'Erdős-Rényi permet quant à lui, d'obtenir une couverture individuelle de dépense variant de 0 € à 1.041,30 € avec une couverture moyenne de $108,24 \pm 301,09 \text{ €}$ et un taux de couverture de 94% (i.e. 465 assurés sur les 493 ont leurs dépenses couvertes). Aussi, la couverture individuelle de dépense obtenue, via le modèle de Watts-Strogatz, part de 0 € à 1.039 € avec une couverture moyenne de $108,07 \pm 300,46 \text{ €}$ et un taux de couverture de 95% (i.e. 468 assurés sur les 493 ont leurs dépenses couvertes).

Enfin, on obtient une couverture individuelle de dépense allant de 0 € à 16.690,34 € avec une couverture moyenne de $88,31 \pm 764,73$ € et un taux de couverture de 88,4%, pour le modèle de Barabási-Albert (i.e. 436 assurés sur l'ensemble des 493 assurés ont leurs dépenses couvertes, soit les 435 assurés avec des dépenses nulles et 1 assuré avec une dépense non-nulle). Remarquons que les taux de couverture et les volatilités de couverture associées aux dépenses sont relativement élevés car 88,2% des assurés de la communauté **CF** n'ont pas déclarés de sinistres donc ont des dépenses annuelles nulles, d'où des couvertures de dépenses nulles. Ainsi, ces dépenses nulles étant supposées "couvertes" augmentent le taux de couverture et la volatilité des distributions de couverture.

Au regard de cette étude de couverture, il ressort que sous les hypothèses (**HDB**), les modèles à clusters homogènes et de Watts-Strogatz permettent d'optimiser les couvertures individuelles des charges annuelles de la communauté **CF**, en présentant les meilleures performances sur les indicateurs de couverture. En effet, le modèle à clusters homogènes conduit à une couverture moyenne des franchises coïncidant avec la franchise moyenne de la communauté **CF** pour un taux de couverture de franchises estimé à 54%, sachant que les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert, fournissent respectivement des taux de couverture de franchises de 48%, de 49% et de 4%. Aussi, le modèle de Watts-Strogatz permet d'obtenir une couverture moyenne des dépenses "couvrant" la dépense moyenne de la communauté **CF** avec un taux de couverture de dépenses estimé à 95%, contre respectivement 93%, 94% et 88,4% pour les modèles à clusters homogènes, d'Erdős-Rényi et de Barabási-Albert.

Par ailleurs, sachant que sous les hypothèses (**HDB**), le modèle de graphe le plus performant au sens du taux de couverture, notamment le modèle à clusters homogènes fournit un taux de couverture de franchises de 54% (i.e. seulement 265 assurés sur les 493 ont leurs franchises couvertes) et l'objectif du mémoire étant d'optimiser l'offre de franchise collaborative afin d'assurer une couverture au plus grand nombre, il importe de s'intéresser aux hypothèses (**HCO**) sur les engagements réciproques entre les assurés. En effet, l'étude théorique effectuée dans la section [4.2](#) a permis de montrer que sous les hypothèses (**HCO**), les charges individuelles des assurés sont "parfaitement couvertes", c'est-à-dire on obtient un taux de couverture des charges de 100% quelque soit le modèle de graphe considéré.

Cependant, sous ces conditions, d'après les relations [4.20](#) et [4.21](#), la contribution individuelle d'un assuré dépend non seulement des charges mais également des degrés de ses relations. De ce fait, une étude de contribution est effectuée dans la section suivante sous les hypothèses (**HCO**) afin de sélectionner le modèle de graphe social permettant d'optimiser (i.e. minimiser) les contributions individuelles des assurés, tout en assurant la couverture des charges de l'intégralité du portefeuille **CF**.

5.5 Étude de contribution

L'objectif de cette section consiste à présenter et à analyser des sorties du programme de calcul **PC1**, notamment les distributions des contributions individuelles des assurés aux charges annuelles de la communauté **CF**, calculées sous les hypothèses (**HCO**) et pour chacun des quatre modèles de graphe étudiés.

En effet, l'étude de couverture effectuée dans la section [5.4](#), a permis de constater que sous les hypothèses (**HDB**), le modèle de graphe présentant le meilleur taux de couverture de franchises est le modèle à clusters homogènes avec un taux de couverture de 54%, c'est-à-dire seulement 265 assurés sur 493 ont leurs franchises couvertes. Ainsi, dans le but d'optimiser le produit de franchise collaborative en assurant la couverture des charges (franchises et dépenses) annuelles

d'un nombre maximum d'assurés, on suppose que les engagements réciproques entre les assurés, sont fixés suivant les hypothèses (**HCO**). En effet, l'étude théorique effectuée dans la section 4.2 montre que sous ces conditions, on obtient un taux de couverture des charges de 100% quelque soit le modèle de graphe considéré.

Cependant, d'après les relations 4.20 et 4.21, ces hypothèses permettent d'exprimer les contributions individuelles aux charges d'un assuré en fonction des charges et des degrés de ses relations. De ce fait, une étude de contribution est effectuée sous les hypothèses (**HCO**) afin d'évaluer l'impact du modèle de graphe social (via les degrés des assurés) sur les contributions individuelles des assurés aux charges annuelles de la communauté **CF**.

Cette étude de contribution consiste dans un premier temps, à analyser la contribution aux charges d'un assuré en fonction de son degré pour chacun des quatre modèles de graphe et dans un second temps, à comparer et sélectionner le modèle de graphe permettant d'optimiser les contributions individuelles aux charges des membres du portefeuille **CF**.

5.5.1 Analyse de la contribution en fonction du degré

Cette sous-section présente d'une part les contributions individuelles aux charges annuelles obtenues pour chaque modèle de graphe (sous les hypothèses (**HCO**)), sur la base des distributions de degrés présentées en 5.2.1 (Calibration du graphe social) et d'autre part l'analyse de l'impact du degré d'un assuré sur ses contributions individuelles aux charges annuelles de ses relations.

Distributions des contributions individuelles sous (**HCO**)

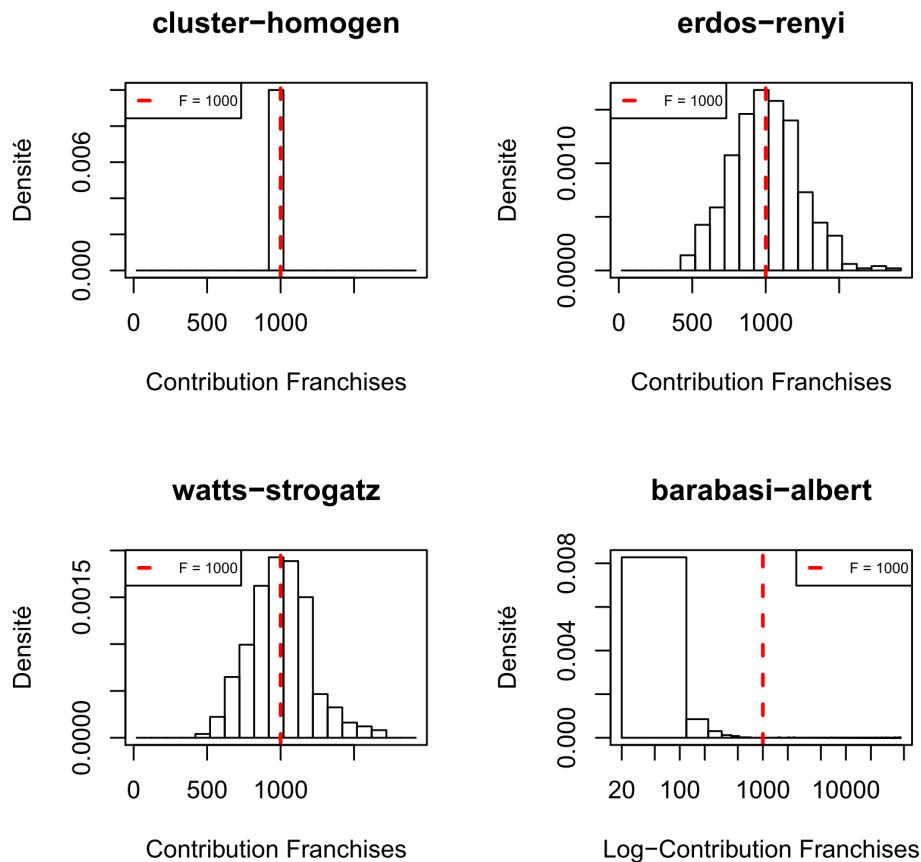
Partant des distributions de degrés présentés en 5.2.1 et obtenues par calibration du graphe social de la communauté **CF** via chacun des quatre modèles de graphe étudiés, on déduit les contributions individuelles aux charges (franchises et dépenses) annuelles des assurés sous les hypothèses (**HCO**), d'après leurs expressions mathématiques (cf. relations 4.20 et 4.21), pour chaque modèle de graphe.

❖ **Contribution aux franchises :** Le Tableau 5.5 et la Figure 5.13 représentent respectivement le tableau récapitulatif et les histogrammes des distributions de contribution aux franchises obtenues sous les hypothèses (**HCO**), pour chacun des quatre modèles de graphe étudiés.

La contribution individuelle aux franchises obtenue, via le modèle à clusters homogènes, est identiquement égale au montant de franchise individuelle avec une valeur moyenne de 1.000 ± 0 € (cf. Tableau 5.5, 2^{ème} colonne et Figure 5.13 (a)). En effet, via ce modèle de graphe, les assurés composant le cluster de taille 3 (par exemple) possèdent chacun 2 relations, d'où sous les hypothèses (**HCO**), l'engagement réciproque entre les assurés de ce cluster vaut $1.000/2 = 500$ €. Ainsi, chaque assuré contribue à hauteur de $500 \times 2 = 1.000$ €, pour la couverture des franchises de ses relations. Un raisonnement analogue permet d'avoir une contribution individuelle aux franchises, égale à $(k - 1) \times 1.000/(k - 1) = 1.000$ €, pour chacun des k assurés composant le cluster de taille $k = 3 : 31$. Aussi, la contribution individuelle aux franchises obtenue, via le modèle d'Erdős-Rényi, varie de 423,90 € à 1.887,20 € avec une valeur moyenne de $1.000 \pm 237,48$ € (cf. Tableau 5.5, 3^{ème} colonne et Figure 5.13 (b)). Le modèle de Watts-Strogatz permet quant à lui, d'obtenir une contribution individuelle aux franchises comprise entre 509,20 € et 1.673 € avec une valeur moyenne de $1.000 \pm 209,36$ € (cf. Tableau 5.5, 4^{ème} colonne et Figure 5.13 (c)). Enfin, on obtient une contribution individuelle aux franchises variant de 24,30 € à 45.565,92 € avec une valeur moyenne de $1.000 \pm 5.903,98$ €, pour le modèle de Barabási-Albert (cf. Tableau 5.5, 5^{ème} colonne et Figure 5.13 (d)).

Modèles	Clusters homogènes	Erdős-Rényi	Watts-Strogatz	Barabási-Albert
Min. (en €)	1.000	423,90	509,20	24,30
Max. (en €)	1.000	1.887,20	1.673	45.565,92
Moy. (en €)	1.000	1.000	1.000	1.000
Sd (en €)	0	237,48	209,36	5.903,98

TABLE 5.5 – Tableau : Distributions de contribution aux franchises par modèle de graphe

FIGURE 5.13 – Histogrammes des distributions de contribution aux franchises par modèle de graphe (F représente la franchise individuelle des assurés de la communauté **CF**) : : *Clusters homogènes* (a), *Erdős-Rényi* (b), *Watts-Strogatz* (c) et *Barabási-Albert* en échelle log (d).

En particulier, on constate que quelque soit le modèle de graphe, la contribution moyenne aux franchises coïncide avec les 1.000 € de franchise moyenne de la communauté **CF**. Ce qui peut s'expliquer par la relation 4.22 exprimant l'égalité entre la contribution mutuelle (i.e. somme des contributions individuelles) aux franchises et la franchise mutuelle (i.e. somme des franchises individuelles) de la communauté **CF**, sous les hypothèses (HCO). Ce qui permet d'avoir également l'égalité entre les valeurs moyennes.

Par ailleurs, sur les quatre modèles de graphe, le modèle de Barabási-Albert fournit les plus petite et grande contributions aux franchises évaluées respectivement à 24,30 € et 45.565,92 € avec une volatilité de 5.903,98 €. Ce qui peut s'expliquer par l'hétérogénéité de la distribution de degrés des assurés, issue de la calibration du graphe social de la communauté **CF** via ce modèle de graphe aléatoire. Cette influence du degré d'un assuré sur sa contribution aux franchises, pour chaque modèle de graphe, est étudiée dans la suite (cf. Contribution en fonction du degré).

❖ **Contribution aux dépenses :** Le Tableau 5.6 et la Figure 5.14 représentent respectivement le tableau récapitulatif et les histogrammes des distributions de contribution aux dépenses pour chacun des quatre modèles de graphe étudiés.

Modèles	Clusters homogènes	Erdős-Rényi	Watts-Strogatz	Barabási-Albert
Min. (en €)	0	0	0	0
Max. (en €)	307,69	363,11	363,30	4.984,27
Moy. (en €)	108	108	108	108
Sd (en €)	74,64	74,86	72,85	647,77

TABLE 5.6 – Tableau : Distributions de contribution aux dépenses par modèle de graphe

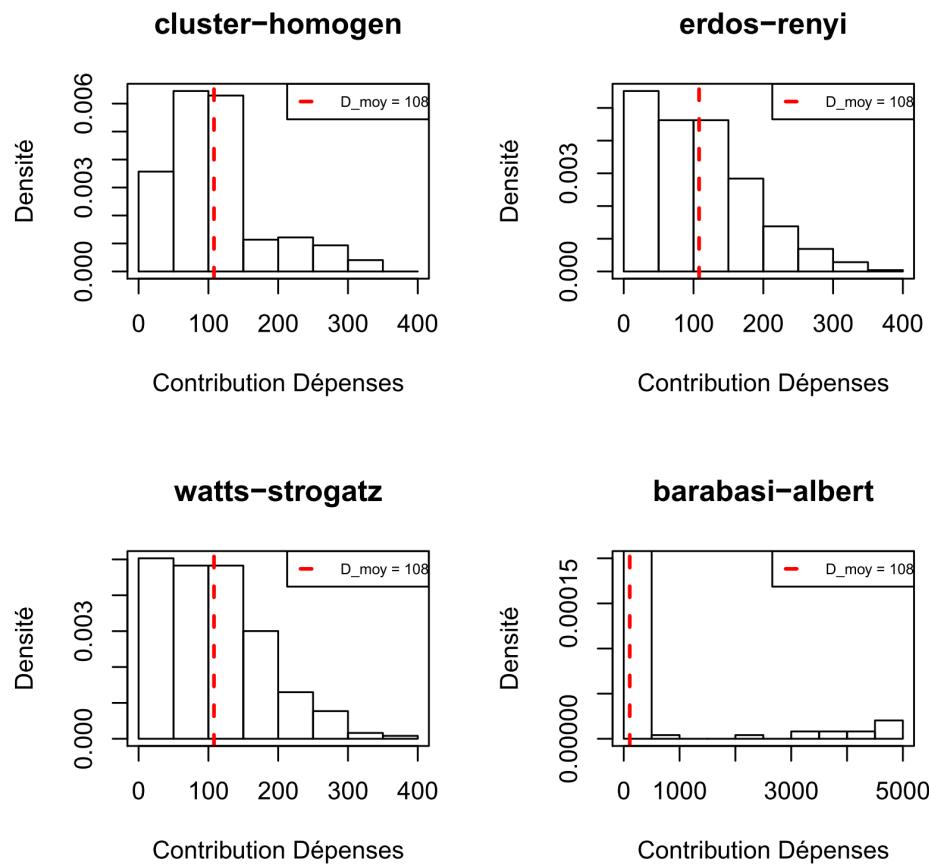


FIGURE 5.14 – Histogrammes des distributions de contribution aux dépenses par modèle de graphe (D_{moy} représente la dépense moyenne des assurés de la communauté CF) : *Clusters homogènes* (a), *Erdős-Rényi* (b), *Watts-Strogatz* (c) et *Barabási-Albert* (d).

La contribution individuelle aux dépenses obtenue, via le modèle à clusters homogènes, est comprise entre 0 € à 307,69 € avec une valeur moyenne de $108 \pm 74,64$ € (cf. Tableau 5.6, 2^{ème} colonne et Figure 5.14 (a)). Aussi, la contribution individuelle aux dépenses obtenue, via le modèle d’Erdős-Rényi, varie de 0 € à 363,11 € avec une valeur moyenne de $108 \pm 74,86$ € (cf. Tableau 5.6, 3^{ème} colonne et Figure 5.14 (b)). Le modèle de Watts-Strogatz permet quant à lui, d’obtenir une contribution individuelle aux dépenses allant de 0 € et 363,30 € avec une valeur moyenne de $108 \pm 72,85$ € (cf. Tableau 5.6, 4^{ème} colonne et Figure 5.14 (c)). Enfin, on

obtient une contribution individuelle aux dépenses variant de 0 € à 4.984,27 € avec une valeur moyenne de $108 \pm 647,77$ €, pour le modèle de Barabási-Albert (cf. Tableau 5.6, 5^{ème} colonne et Figure 5.14 (d)).

On constate, en particulier, que quelque soit le modèle de graphe, la contribution moyenne aux dépenses coïncide avec les 108 € de dépense moyenne de la communauté **CF**. Ce qui s'explique par la relation 4.23 exprimant l'égalité entre la contribution mutuelle (i.e. somme des contributions individuelles) aux dépenses et la dépense mutuelle (i.e. somme des dépenses individuelles) de la communauté **CF**, sous les hypothèses (**HCO**). Ce qui permet d'avoir également l'égalité entre les valeurs moyennes.

Par ailleurs, la contribution minimale de 0 €, constatée via les quatre modèles de graphe, peut s'expliquer par le fait que les simulations effectuées pour chacun de ces modèles, fournissent des assurés uniquement connectés à des assurés non sinistrés donc ayant des dépenses annuelles nulles. En particulier, le modèle à clusters homogènes fournit la contribution maximale la plus faible, estimée à 307,69 € et le modèle de Barabási-Albert, la contribution maximale la plus forte, estimée à 4.984,27 €, soit 16 fois plus élevée par rapport au modèle à clusters homogènes.

Enfin, après avoir présenté les distributions de contribution aux charges obtenues pour chaque modèle de graphe, sous les conditions (**HCO**), le point suivant consiste à analyser l'évolution de ces contributions individuelles des assurés en fonction des distributions de degrés issues de la calibration du graphe social via les quatre modèles de graphe étudiés.

Contribution en fonction du degré

La Figure 5.15 représente l'évolution de la contribution aux franchises d'un assuré en fonction de son degré, pour chaque modèle de graphe.

- ❖ **Modèle à clusters homogènes :** Comme annoncé plus-haut et illustré par la Figure 5.15 (a), quelque soit le degré d'un assuré, sa contribution individuelle aux franchises via le modèle à clusters homogènes, est constante et égale au montant de franchise individuelle, soit :

$$y = 1.000 ,$$

avec y représentant la contribution individuelle aux franchises d'un assuré de la communauté **CF** ;

- ❖ **Modèle d'Erdős-Rényi :** La contribution individuelle aux franchises en fonction du degré, via le modèle d'Erdős-Rényi, est représentée sur la Figure 5.15 (b). Une regression linéaire de la contribution en fonction du degré, permet d'obtenir les résultats suivants : le R^2 ajusté égal à 0,948 et l'hypothèse de nullité de la pente du modèle est rejetée à 99,9%. Ainsi, la contribution individuelle aux franchises peut être calculée selon l'équation linéaire suivante :

$$y = 50,01 \times x - 1,33 ,$$

avec x et y représentant respectivement le degré et la contribution aux franchises d'un assuré de la communauté **CF** ;

- ❖ **Modèle de Watts-Strogatz :** La Figure 5.15 (c) représente la contribution individuelle aux franchises en fonction du degré, via le modèle de Watts-Strogatz. Un ajustement linéaire de la contribution en fonction du degré donne les résultats suivants : le R^2 ajusté égal à 0,952 et l'hypothèse de nullité de la pente du modèle est rejetée à 99,9%. Ainsi, la contribution individuelle aux franchises peut être calculée selon la formule fermée suivante :

$$y = 50,74 \times x - 14,74 ,$$

avec x et y représentant respectivement le degré et la contribution aux franchises d'un assuré de la communauté CF ;

- ❖ **Modèle de Barabási-Albert** : La contribution individuelle aux franchises en fonction du degré, via le modèle de Barabási-Albert, est représentée sur la Figure 5.15 (d). Par une regression linéaire de la contribution en fonction du degré, on obtient les résultats suivants : le R^2 ajusté égal à 0,999 et l'hypothèse de nullité des coefficients du modèle est rejetée à 99,9%. Ainsi, l'équation linéaire suivante permet le calcul de la contribution individuelle aux franchises d'un assuré en fonction de son degré :

$$y = 97,99 \times x - 937,89 ,$$

avec x et y représentant respectivement le degré et la contribution aux franchises d'un assuré de la communauté CF.

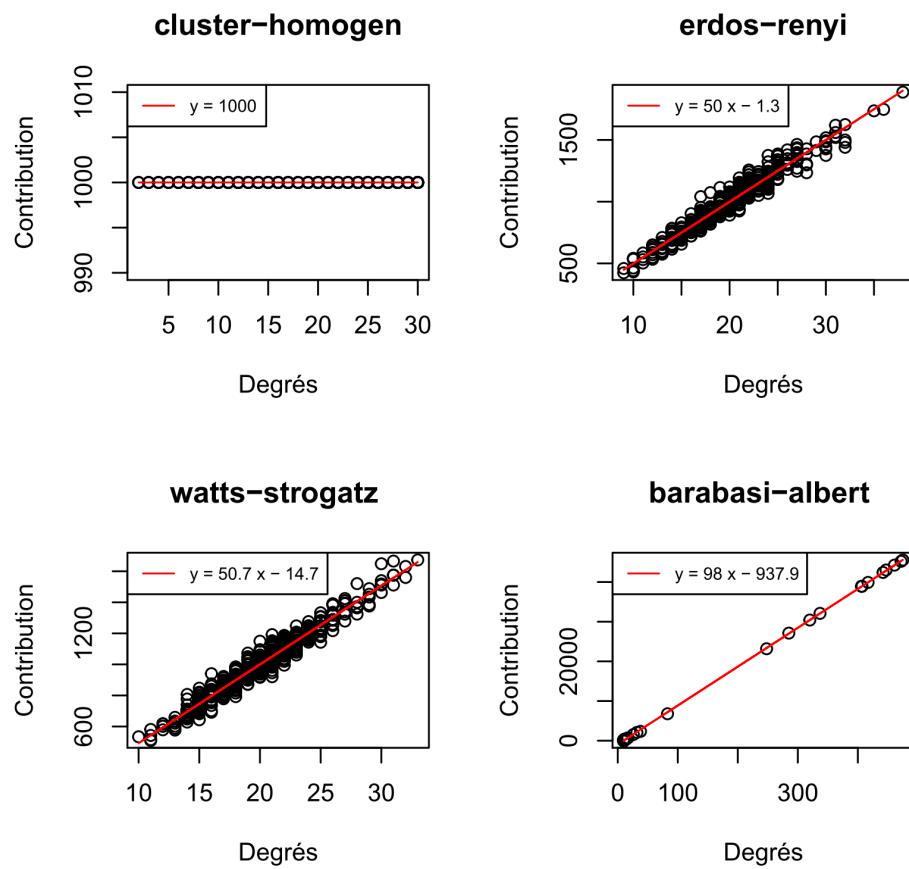


FIGURE 5.15 – Contribution aux franchises en fonction du degré par modèle de graphe : *Clusters homogènes* (a), *Erdős-Rényi* (b), *Watts-Strogatz* (c) et *Barabási-Albert* (d).

Par ailleurs, d'après la relation 4.6, la contribution aux dépenses des assurés est non seulement fonction des degrés mais également des dépenses de leurs relations. Il n'est donc pas possible d'avoir comme pour la contribution aux franchises, une formule fermée exprimant la contribution aux dépenses d'un assuré uniquement en fonction de son degré.

Toutefois, rappelons que cette étude de contribution aux charges en fonction des degrés des assurés se base essentiellement sur les distributions de degrés résumées dans le Tableau 5.1 et

obtenues par simulation d'un scénario de graphe social via chaque modèle de graphe. Dans la sous-section suivante, afin d'obtenir des résultats plus robustes et généralisables pour les contributions individuelles et afin d'évaluer et comparer la performance des contributions issues de chaque modèle de graphe, la méthode de Monte-Carlo est utilisée avec 1.000 simulations de scénarios de graphe social pour chacun des trois modèles de graphe aléatoire, sachant que le scénario du modèle à clusters homogènes est déterministe et est présenté en 5.2.1.

5.5.2 Sélection de modèle de graphe

Dans cette sous-section, il est fait une présentation des contributions individuelles aux charges obtenues en sortie du programme **PC1** sous les hypothèses (**HCO**), pour chaque modèle de graphe. Aussi, afin d'analyser et de comparer la performance des contributions obtenues pour chacun de ces quatre modèles de graphe, on introduit les indicateurs de contribution suivants.

Définition 10 (Indicateurs de contribution).

- ❖ **Le ratio de contribution (RCN)** : qui représente pour un assuré, le rapport entre sa charge individuelle et sa contribution individuelle (non-nulle), en analogie avec le ratio **S/P** tel que :

$$\forall \text{ assuré } i = 1 : n, \quad RCN_i = \frac{Ch_i}{Cont_i},$$

où $Cont_i$ représente la contribution (non-nulle) aux franchises ou aux dépenses et Ch_i représente respectivement la franchise ou la dépense de l'assuré i .

Pour tout assuré i , si son ratio de contribution vérifie : $RCN_i < 1$, alors : $Ch_i < Cont_i$.

On dit que cet assuré "**surcontribute**", c'est-à-dire contribue au-delà de sa charge.

- ❖ **Le taux de surcontribution (TSC)** : qui mesure la proportion d'assurés contribuant au-delà de leurs charges, pour un modèle de graphe donné tel que :

$$TSC = \frac{1}{l} \sum_{i=1}^l \mathbf{1}_{\{RCN_i < 1\}},$$

pour les $1 \leq l \leq n$ contributions $Cont_i$ non-nulles, ou de manière générale, en considérant les n contributions individuelles des assurés :

$$TSC = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{Ch_i < Cont_i\}}$$

Cet indicateur compris dans l'intervalle $[0, 1]$ est à minimiser.

- ❖ **La contribution moyenne** : qui représente la moyenne de la distribution de contribution obtenue pour un modèle de graphe donné. Cet indicateur est à maximiser.
- ❖ **La volatilité de contribution** : qui mesure la dispersion des valeurs de la distribution de contribution obtenue pour un modèle de graphe donné. Cet indicateur évalué à l'aide de l'**écart-type**, est à minimiser.

Pour rappel, pour des raisons de temps de calcul, les distributions de contribution présentées dans cette sous-section, ont été obtenues sur la base de 1.000 simulations de scénario de graphe social pour les trois modèles de graphe aléatoire et du scénario déterministe du modèle à clusters homogènes, présenté en 5.2.1.

L'objectif consiste à déterminer les indicateurs de contribution issus de ces différentes distributions, afin de comparer les performances de contribution obtenues pour chaque modèle de graphe

et *in fine*, de sélectionner le modèle permettant d'optimiser les contributions individuelles aux charges annuelles des assurés de la communauté **CF**. Pour ce faire, on s'intéresse dans un premier temps, aux ratios de contributions associés aux contributions non-nulles et dans un second temps, aux autres indicateurs de contribution (taux de surcontribution, contribution moyenne et volatilité de contribution) associés à l'ensemble des contributions individuelles de tout le portefeuille **CF**.

Distributions des ratios de contribution

Pour rappel, le ratio de contribution aux charges d'un assuré consiste au rapport entre sa charge individuelle et sa contribution individuelle (non-nulle). De ce fait, partant des charges annuelles de la communauté **CF** et des distributions de contribution aux charges obtenues en sortie du programme **PC1**, on déduit les ratios de contribution aux charges pour les assurés ayant des contributions non-nulles. On dispose donc de deux distributions de ratios de contribution associées respectivement aux franchises et aux dépenses des assurés, pour chaque modèle de graphe.

❖ **Ratios de contribution aux franchises** : La Figure 5.16 représente les fonctions de densité des ratios de contribution aux franchises des assurés pour chaque modèle de graphe.

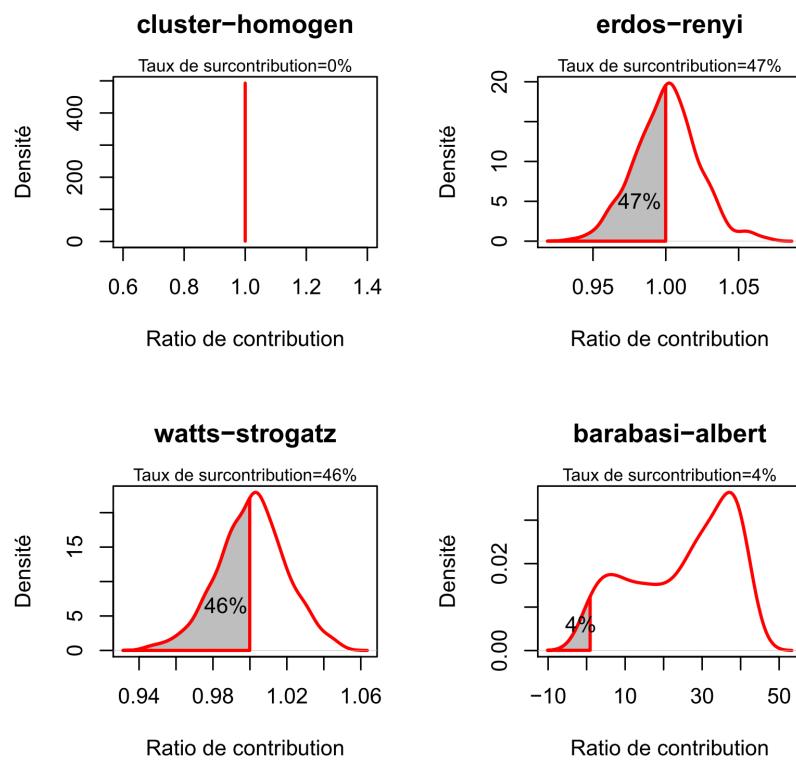


FIGURE 5.16 – Fonctions de densité des ratios de contribution aux franchises par modèle de graphe : *Clusters homogènes* (a), *Erdős-Rényi* (b), *Watts-Strogatz* (c) et *Barabási-Albert* (d).

Sur la Figure 5.16, les aires des parties colorées en "gris" représentent les taux de surcontribution aux franchises (i.e. le pourcentage d'assurés contribuant au-delà de leurs franchises), respectivement pour les modèles à clusters homogènes, d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert. On constate par exemple que, le modèle à clusters homogènes fournit le taux de surcontribution le plus faible et donc le plus performant, soit de 0%. En effet, comme expliqué en 5.5.1, ce modèle de graphe permet d'obtenir sous les conditions (**HCO**), une contribution

individuelle aux franchises identiquement égale aux 1.000 € de franchise individuelle, soit un ratio de contribution égal à 1 pour chacun des 493 assurés de la communauté CF (cf. Figure 5.16 (a)). Aussi, les modèles de Barabási-Albert, de Watts-Strogatz et d'Erdős-Rényi fournissent respectivement dans l'ordre croissant, des taux de surcontribution aux franchises de 4%, de 46% et de 47%.

✿ **Ratios de contribution aux dépenses :** La Figure 5.17 représente les fonctions de densité des ratios de contribution aux dépenses associés aux contributions non-nulles des assurés, pour chaque modèle de graphe. En effet, d'après la définition du ratio de contribution, il prend en compte uniquement les contributions non-nulles des assurés, soit 444 via le modèle à clusters homogènes et 493 via les trois modèles de graphe aléatoire. En effet, le modèle à clusters homogènes permet d'avoir 49 assurés ayant des contributions nulles, étant uniquement en relation avec des assurés ayant des dépenses annuelles nulles.

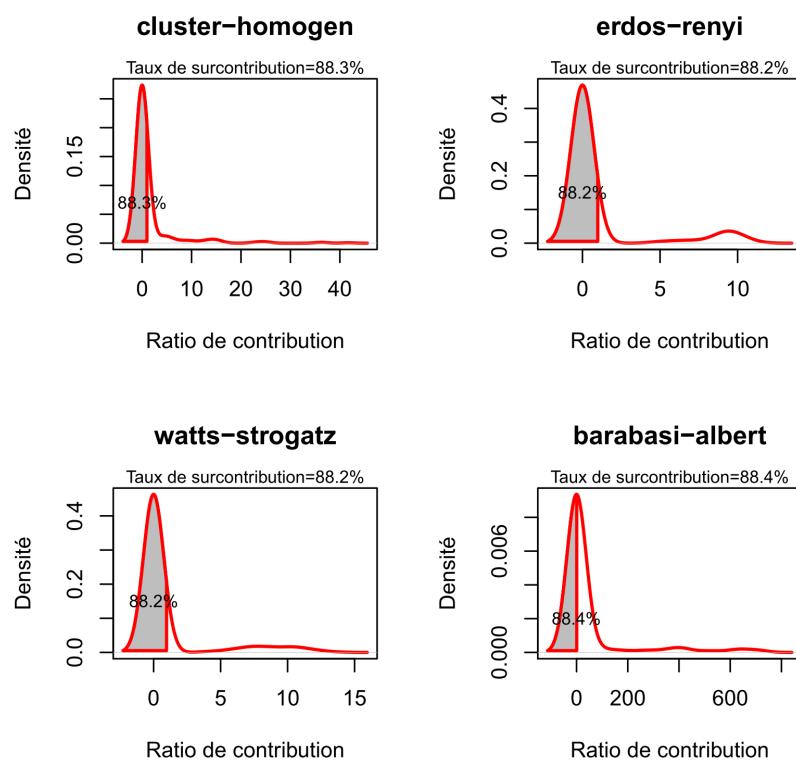


FIGURE 5.17 – Fonctions de densité des ratios de contribution aux dépenses par modèle de graphe : *Clusters homogènes* (a), *Erdős-Rényi* (b), *Watts-Strogatz* (c) et *Barabási-Albert* (d).

De même que sur la Figure 5.16 relative aux franchises, sur la Figure 5.17 relative aux dépenses des assurés, les aires des parties colorées en "gris" représentent les taux de surcontribution aux dépenses (i.e. le pourcentage d'assurés contribuant au-delà de leurs dépenses), respectivement pour les modèles à clusters homogènes, d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert. Par exemple, les modèles d'Erdős-Rényi et de Watts-Strogatz fournissent le meilleur taux de surcontribution qui est de 88,2% (i.e. 435 assurés sur 493 contribuent au-delà de leurs dépenses). En effet, ces 435 assurés correspondent aux assurés non sinistrés, c'est-à-dire ayant des dépenses nulles. Les modèles à clusters homogènes et de Barabási-Albert quant à eux, donnent respectivement des taux de surcontribution de 88,3% (i.e. 392 assurés sur les 444 assurés ayant des contributions non-nulles via ce modèle, contribuent au-delà de leurs dépenses) et de 88,4% (i.e. 436 assurés sur l'ensemble des 493 assurés ayant des contributions non-nulles via ce modèle,

contribuent au-delà de leurs dépenses, soit les 435 assurés avec des dépenses nulles et 1 assuré avec une dépense non-nulle).

Rappelons que ces ratios de contribution aux dépenses ont été obtenus en considérant uniquement les assurés possédant des contributions non-nulles, soit 444 assurés via le modèle à clusters homogènes et les 493 assurés via les trois modèles de graphe aléatoire. Le point suivant fait un récapitulatif des indicateurs de contribution calculés en prenant en compte toutes les 493 contributions individuelles des assurés du portefeuille **CF**.

Récapitulatif des indicateurs de contribution

Le Tableau 5.7 résume les indicateurs de contribution aux charges calculés d'après leurs définitions sur la base des charges de la communauté **CF** et des contributions individuelles obtenues en sortie du programme **PC1** pour chaque modèle de graphe. En effet, ces indicateurs consistant en la **contribution moyenne** (notée "Moy." dans le Tableau 5.7), la **volatilité de contribution** (notée "Vol." dans le Tableau 5.7) et le **taux de surcontribution** (noté "Taux" dans le Tableau 5.7) sont calculés pour chacune des distributions de contribution obtenue par modèle de graphe et sous les hypothèses (**HCO**).

Indicateurs de contribution	Franchises ($F = 1.000 \text{ €}$)			Dépenses ($\bar{D} = 108 \text{ €}$)		
	Moy	Vol	Taux	Moy	Vol	Taux
Modèles						
Clusters homogènes	1.000	0	0	108	74,64	79,5
Erdős-Rényi	999,83	21,69	47	107,7	7,36	88,2
Watts-Strogatz	999,22	18,47	46	108,4	17,95	88,2
Barabási-Albert	1.004,67	5.544,44	4	108,3	615,26	88,4

TABLE 5.7 – Tableau récapitulatif : Indicateurs de contribution par modèle de graphe. Taux = taux de surcontribution (en %), Moy = contribution moyenne (en €) et Vol = volatilité ou écart-type de contribution (en €).

Les distributions des contributions individuelles des assurés aux franchises et aux dépenses de la communauté **CF**, obtenues en sortie du programme **PC1** sous les hypothèses (**HCO**) et sur la base de 1.000 simulations de scénarios de graphe social via les trois modèles de graphe aléatoire et du scénario déterministe de graphe social via le modèle à clusters homogènes, s'établissent comme-suivi :

- ❖ **Contribution aux franchises** : Comme présenté en 5.5.1, le scénario déterministe de graphe social simulé via le modèle à clusters homogènes entraîne une contribution individuelle aux franchises identiquement égale au montant de franchise individuelle, c'est-à-dire avec une contribution moyenne de $1.000 \pm 0 \text{ €}$ et un taux de surcontribution de 0%. En effet, via ce modèle de graphe et sous les hypothèses (**HCO**), les assurés contribuent exactement au niveau des 1.000 € correspondant au montant de leur franchise individuelle (cf. Tableau 5.7, 3^{ème} ligne et volet Franchises). Aussi, la contribution individuelle aux franchises obtenue, via le modèle d'Erdős-Rényi, varie de 934,1 € à 1.069,7 € avec une contribution moyenne de $999,83 \pm 21,69 \text{ €}$ et un taux de surcontribution de 47% (i.e. 234 assurés sur les 493 ont des contributions individuelles supérieures à leur franchise). Le modèle de Watts-Strogatz permet quant à lui, d'obtenir une contribution individuelle aux franchises comprise entre 952,5 € et 1.058,2 € avec une contribution moyenne de $999,22 \pm 18,47 \text{ €}$ et un taux de surcontribution de 46% (i.e. 225 assurés sur les 493 ont des contributions individuelles supérieures à leur franchise). Enfin, on obtient une contribution individuelle aux franchises variant de 23,17 € à 38.033,71 € avec une contribution moyenne de $1.004,67 \pm 5.544,44 \text{ €}$

et un taux de surcontribution de 4%, pour le modèle de Barabási-Albert (i.e. 18 assurés sur les 493 ont des contributions individuelles supérieures à leur franchise) ;

- ❖ **Contribution aux dépenses :** Comme présenté en 5.5.1, le scénario déterministe de graphe social simulé via le modèle à clusters homogènes fournit une contribution individuelle aux dépenses comprise entre 0 € et 307,69 € avec une contribution moyenne de $108 \pm 74,64$ € et un taux de surcontribution de 79,5% (cf. Tableau 5.7, 3^{ème} ligne et volet Dépenses). En effet, via ce modèle de graphe, on obtient 392 assurés sur les 493 assurés avec des contributions individuelles non-nulles et des dépenses nulles. Aussi, via le modèle d'Erdős-Rényi, la contribution individuelle aux dépenses obtenue varie de 83,35 € à 128,83 € avec une contribution moyenne de $107,7 \pm 7,36$ € et un taux de surcontribution de 88,2% (i.e. seulement les 435 assurés sur 493 possédant des dépenses nulles contribuent au-delà de leurs dépenses nulles). Le modèle de Watts-Strogatz permet quant à lui, d'obtenir une contribution individuelle aux dépenses allant de 68,06 € et 171,35 € avec une contribution moyenne de $108,4 \pm 17,95$ € et un taux de surcontribution de 88,2% (i.e. seulement les 435 assurés sur 493 possédant des dépenses nulles contribuent au-delà de leurs dépenses nulles). Enfin, on obtient une contribution individuelle aux dépenses variant de 1,30 € à 4.223,52 € avec une contribution moyenne de $108,3 \pm 615,26$ € et un taux de surcontribution de 88,4%, pour le modèle de Barabási-Albert (i.e. 436 assurés sur l'ensemble des 493 assurés contribuent au-delà de leurs dépenses, soit les 435 assurés avec des dépenses nulles et 1 assuré avec une dépense non-nulle).

Au regard de cette étude de contribution, il ressort que sous les hypothèses (**HCO**), le modèle à clusters homogènes permet d'optimiser les contributions individuelles des assurés aux charges de la communauté **CF**, en présentant les meilleures performances sur les indicateurs de contribution. En effet, ce modèle de graphe permet d'obtenir une contribution individuelle aux franchises exactement égale au montant de franchise individuelle de la communauté **CF**, soit un taux de surcontribution aux franchises de 0%, sachant que les modèles de Barabási-Albert, de Watts-Strogatz et d'Erdős-Rényi fournissent respectivement des taux de surcontribution aux franchises de 4%, de 46% et de 47%. Aussi, le modèle à clusters homogènes conduit à une contribution moyenne aux dépenses coïncidant avec la dépense moyenne de la communauté **CF** pour un taux de surcontribution aux dépenses estimé à 79,5%, sachant que les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert, fournissent respectivement des taux de surcontribution aux dépenses de 88,2%, de 88,2% et de 88,4%.

Ainsi, le modèle à clusters homogènes est le plus performant des quatre modèles de graphe étudiés, au sens du taux de surcontribution. De ce fait, ce modèle serait le plus à même d'optimiser le produit de franchise collaborative par une optimisation des contributions individuelles des assurés sous les hypothèses (**HCO**).

Synthèse de la partie II

Cette deuxième partie du mémoire a consisté dans un premier temps, au chapitre 3, à introduire et à formaliser mathématiquement le produit de franchise collaborative. Puis, dans un second temps, à comprendre comment fixer les paramètres (i.e. le modèle de graphe social et les engagements réciproques entre les assurés) afin d'optimiser le produit de franchise collaborative. Pour ce faire, l'étude théorique effectuée au chapitre 4 a montré que les hypothèses (**HCO**) sur les engagements réciproques, permettent d'optimiser les couvertures individuelles des charges annuelles des assurés. Aussi, l'étude de contribution effectuée au chapitre 5 sous les conditions (**HCO**), a sélectionné le modèle à **clusters homogènes** comme étant le modèle de graphe permettant d'optimiser les contributions individuelles des assurés aux charges annuelles de la communauté **CF**.

De ce fait, dans la troisième et dernière partie du mémoire dédiée à la mise en place d'un produit de franchise collaborative, ces paramètres optimaux, notamment les hypothèses (**HCO**) et le modèle à **clusters homogènes** serviront d'entrée pour le « scénario central » (i.e. le produit optimisé). Puis, une étude de sensibilité permettra de prendre en compte les autres hypothèses sur les paramètres afin d'analyser leur impact sur les variables d'intérêt du produit de franchise collaborative (cf. chapitre 6). Aussi, au chapitre 7, des applications à des données réelles, notamment des données de généalogie et de réseaux sociaux tels que *Facebook* seront effectuées afin d'analyser comment ces types spécifiques de connexions entre les assurés peuvent influer sur ces variables d'intérêt.

Troisième partie

Mise en place d'un produit de franchise collaborative

Chapitre 6

Aspect technique du produit de franchise collaborative

6.1 Introduction

Ce chapitre est consacré à la mise en place d'un produit de franchise collaborative dans le cadre d'un contrat d'assurance IARD (e.g. automobile). En particulier, on s'intéresse à l'aspect technique du produit, c'est-à-dire partant des charges annuelles de la communauté **CF** présentée dans la section 5.1, il s'agit de déterminer la couverture moyenne des charges et la contribution maximale aux charges d'un assuré quelconque de la communauté. Ainsi, les résultats obtenus dans ce chapitre sont étroitement liés aux hypothèses et aux caractéristiques de la communauté **CF**. Toutefois, la méthodologie de calcul reste généralisable et applicable au processus de tarification "classique" d'un assureur non-vie. Aussi, la prime annuelle commerciale d'un assuré le cas échéant, peut être déduite de sa contribution individuelle (ou cotisation) annuelle obtenue dans le cadre technique de ce chapitre, par l'ajout d'un pourcentage de majoration (e.g. 10% en général) pour la prise en compte des frais de gestion, d'acquisition et d'administration.

Par ailleurs, sachant que les variables d'intérêt du produit de franchise collaborative sont les variables de couverture et de contribution définies respectivement en Définitions 7 et 8, et que la couverture moyenne des charges annuelles d'un assuré coïncide avec sa contribution moyenne annuelle aux charges de la communauté d'après les relations 3.2 et 3.3, les sorties du programme de calcul **PC2** implémenté sous le logiciel R, dans le cadre de ce chapitre sont respectivement : la **couverture moyenne** des charges et la **contribution maximale** aux charges des membres de la communauté **CF** dans l'année.

Ainsi, ce chapitre présente tout d'abord, le fonctionnement du programme de calcul **PC2**, ensuite les sorties obtenues via le produit optimisé dit **scénario central** prenant en entrée les paramètres optimaux issus des études d'optimisation effectuées aux chapitres 4 et 5. Enfin, une **étude de sensibilité** permettra de tester ces paramètres selon différentes sensibilités afin de mettre en exergue leurs impacts sur les sorties du programme **PC2** pour ces différentes variantes du produit de franchise collaborative. Le Tableau 6.1 résume les paramètres associés aux différentes variantes du produit de franchise collaborative, établies dans ce chapitre.

Paramètres \ Variantes	Scénario central	Sensibilité engagements	Sensibilité graphe social
Engagements réciproques	Hypothèses (HCO)	Hypothèses (HDB)	Hypothèses (HCO)
Graphe social	Clusters homogènes	Clusters homogènes	Erdős-Rényi Watts-Strogatz Barabási-Albert

TABLE 6.1 – Tableau récapitulatif : Variantes du produit de franchise collaborative établies

6.2 Programme de calcul PC2

Cette section présente le programme **PC2** implémenté sous le logiciel R pour le calcul des distributions des variables d'intérêt du produit de franchise collaborative, notamment les variables de couverture des charges et de contribution aux charges des assurés de la communauté fictive **CF**.

6.2.1 Les entrées du programme PC2

De même que pour le programme **PC1**, les entrées du programme de calcul **PC2** sont : les **charges** annuelles de la communauté **CF** (présentées dans la section 5.1) et les **paramètres** du produit de franchise collaborative (présentés dans la section 5.2).

6.2.2 Les sorties du programme PC2

Les sorties du programme de calcul **PC2** sont : la **couverture moyenne** des charges et la **contribution maximale** aux charges des membres de la communauté **CF**.

6.2.3 Le fonctionnement du programme PC2

Pour chacune des trois variantes (**scénario central**, **sensibilité engagements** et **sensibilité graphe social**) du produit de franchise collaborative établies, les étapes du programme de calcul **PC2** s'établissent comme-suit.

Scénario central

- ❖ **Étape 1 :** Le graphe social de la communauté **CF** est calibré via le modèle à clusters homogènes. Ensuite, on déduit la matrice d'adjacence R de la communauté décrivant les relations entre les assurés telle que $R_{ij} = 1$, si les assurés i et j sont en relation et 0 sinon. Enfin, partant des charges annuelles de la communauté et des hypothèses (HCO) sur les engagements réciproques, on déduit les matrices E des engagements réciproques et C des contributions réciproques aux dépenses entre les assurés, d'après leurs expressions mathématiques ;
- ❖ **Étape 2 :** La connaissance des matrices des engagements réciproques et des contributions réciproques aux dépenses entre les assurés, permet de déduire les distributions des couvertures individuelles des charges et des contributions individuelles aux charges, selon leurs expressions mathématiques ;
- ❖ **Étape 3 :** La dernière étape consiste à prendre la moyenne sur chacune des deux distributions de couvertures individuelles des charges (franchises et dépenses) afin d'obtenir les couvertures moyennes des charges annuelles de la communauté et les maximum sur chacune des deux distributions de contributions individuelles aux charges (franchises et dépenses)

afin d'obtenir les contributions maximales aux charges annuelles de la communauté.

Sensibilité engagements

Les étapes du programme **PC2** pour le calcul des sorties associées à la sensibilité engagements sont similaires aux étapes du scénario central, excepté les engagements réciproques qui sont calibrés suivant les hypothèses (**HDB**).

Sensibilité graphe social

- ❖ **Étape 1 :** Le graphe social de la communauté **CF** est calibré via respectivement les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert. Ensuite, on déduit la matrice d'adjacence R de la communauté décrivant les relations entre les assurés telle que $R_{ij} = 1$, si les assurés i et j sont en relation et 0 sinon. Enfin, partant des charges annuelles de la communauté **CF** et des hypothèses (**HCO**) sur les engagements réciproques, on déduit les matrices E des engagements réciproques et C des contributions réciproques aux dépenses entre les assurés, d'après leurs expressions mathématiques, pour chacun des trois modèles de graphe aléatoire ;
- ❖ **Étape 2 :** La connaissance des matrices des engagements réciproques et des contributions réciproques aux dépenses entre les assurés, permet de déduire les distributions des couvertures individuelles des charges et des contributions individuelles aux charges, selon leurs expressions mathématiques, pour chacun des trois modèles de graphe aléatoire ;
- ❖ **Étape 3 :** Dans le but d'analyser l'impact du degré moyen sur la couverture moyenne et la contribution maximale, on fait varier le degré moyen de la communauté **CF** de 15 à 100 relations. Pour chaque valeur du degré moyen, on effectue 1.000 simulations de scénarios de graphe social via chacun des trois modèles de graphe aléatoire. Pour chaque scénario, on déduit les couvertures moyennes des charges (franchises et dépenses) et les contributions maximales aux charges, en prenant respectivement la moyenne sur les distributions des couvertures individuelles et le maximum sur les distributions des contributions individuelles obtenues à l'étape 2.

Afin d'obtenir les sorties pour chaque degré moyen variant de 15 à 100 relations, une moyenne sur l'ensemble des scénarios est effectuée, d'après la méthode de Monte-Carlo. Enfin, des intervalles de confiance à 95% des valeurs estimées sont calculés afin d'évaluer la précision des estimations en fonction du degré moyen de la communauté **CF**.

6.3 Scénario central

Cette section a pour objectif de présenter les sorties (i.e. la couverture moyenne et la contribution maximale) du programme **PC2** pour la variante optimisée du produit de franchise collaborative dit **scénario central**, prenant en entrée les paramètres optimaux issus des études effectuées aux chapitres 4 et 5, relatives respectivement à la recherche des engagements réciproques optimaux et du modèle de graphe social optimal permettant d'optimiser le produit de franchise collaborative.

En effet, ces études ont respectivement montré que les hypothèses (**HCO**) permettent d'optimiser les couvertures individuelles des charges des assurés avec un taux de couverture des charges de 100% (i.e. tous les assurés ont leurs charges couvertes quelque soit le modèle de graphe considéré) et le modèle à **clusters homogènes** permet d'optimiser les contributions individuelles aux charges des assurés avec un taux de surcontribution aux franchises de 0%, sous les conditions

(HCO) (i.e. les assurés contribuent exactement pour le même montant que leurs franchises individuelles).

Ainsi les Figures 6.1 (a) et 6.2 (a) représentent les sorties (couverture moyenne en "vert" et contribution maximale en "orange") du scénario central respectivement associées aux franchises et aux dépenses de la communauté **CF**, pour un degré moyen de $\bar{d} = 20, 12$ (degré moyen obtenu par la calibration du graphe social via le modèle à clusters homogènes). Les sorties pour les autres variantes (sensibilité engagements et sensibilité graphe social) du produit de franchise collaborative seront complétées au fil de l'étude de sensibilité.

On constate sur la Figure 6.1 (a), que pour une franchise moyenne (ou individuelle) de 1.000 €, le scénario central permet d'obtenir une couverture moyenne des franchises de 1.000 ± 0 € et une contribution maximale aux franchises de 1.000 ± 0 € pour un assuré quelconque de la communauté **CF**. C'est-à-dire, pour cette variante du produit de franchise collaborative, les assurés contribuent exactement pour le même montant que leurs franchises individuelles. Il n'existe donc pas de risque d'insuffisance de couverture (i.e. franchise non couverte), ni de risque de surcontribution (i.e. contribution au-delà de la franchise) pour les assurés.

Par ailleurs, d'après la Figure 6.2 (a), pour une dépense moyenne de 108 €, ce produit optimisé fournit une couverture moyenne des dépenses de 108 ± 0 € et une contribution maximale aux dépenses de 308 ± 0 € pour un assuré quelconque de la communauté **CF**. En effet, le scénario central prenant en entrée le montant des engagements réciproques calibré suivant les hypothèses **(HCO)**, permet d'assurer une couverture optimale des dépenses annuelles des assurés. Cependant, la contribution maximale d'un assuré aux dépenses étant de 308 € pour une dépense moyenne de 108 €, on obtient donc en moyenne une "sur-contribution" environ 3 fois plus élevée que la dépense moyenne de la communauté **CF**.

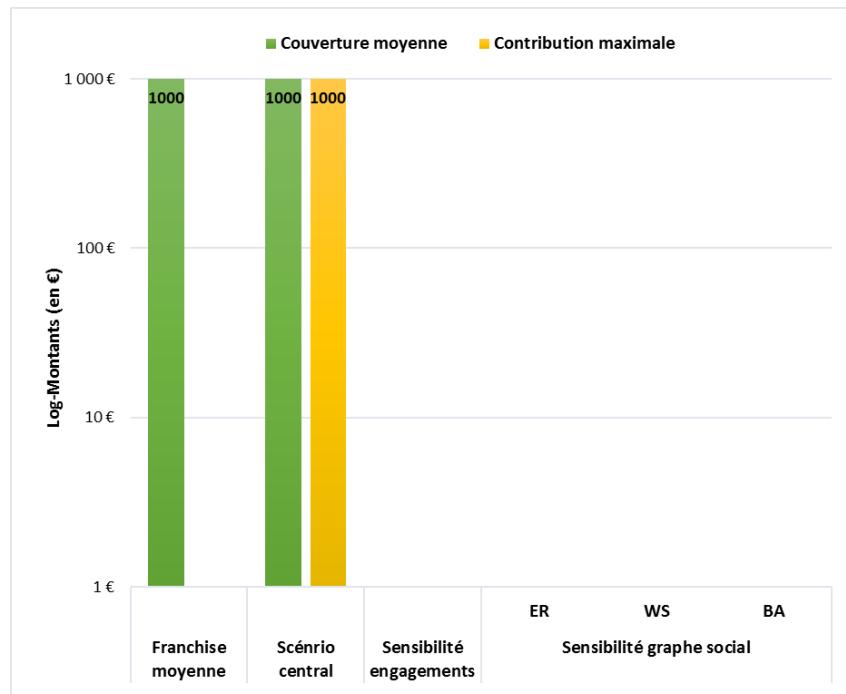


FIGURE 6.1 – Sorties (couverture moyenne et contribution maximale en échelle log) associées aux franchises par variante du produit de franchise collaborative : *Scénario central* (a), *Sensibilité engagements* (b), *Sensibilité graphe social* (c).

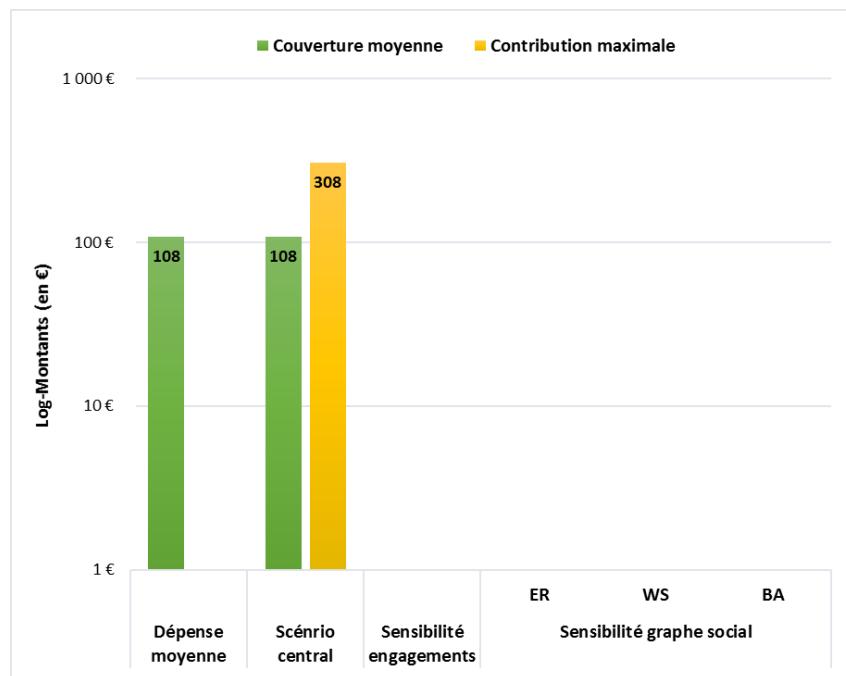


FIGURE 6.2 – Sorties (couverture moyenne et contribution maximale en échelle log) associées aux dépenses par variante du produit de franchise collaborative : *Scénario central* (a), *Sensibilité engagements* (b), *Sensibilité graphe social* (c).

6.4 Étude de sensibilité

Le scénario central présenté dans la section précédente représente la variante optimisée du produit de franchise collaborative, c'est-à-dire, prenant en entrée les paramètres permettant d'optimiser le produit, à savoir les hypothèses (**HCO**) pour la détermination du montant des **engagements réciproques** entre les assurés et le modèle à **clusters homogènes** pour la calibration du **graphe social** de la communauté **CF**.

Afin de mesurer l'impact de chaque paramètre sur les sorties du produit de franchise collaborative, notamment la couverture moyenne des charges et la contribution maximale aux charges des assurés de la communauté **CF**, deux autres variantes du produit sont mises en place via des études de sensibilité : **sensibilité engagements** et **sensibilité graphe social**.

6.4.1 Sensibilité engagements

Cette étude vise à étudier l'impact des **engagements réciproques** sur la couverture moyenne des charges et la contribution maximale aux charges des assurés. Ainsi, le graphe social reste identique à celui du scénario central, c'est-à-dire calibré via le modèle à clusters homogènes. Cependant, à la différence du scénario central, dans lequel le montant des engagements réciproques est fixé suivant les hypothèses (**HCO**), ce montant est maintenant fixé suivant les hypothèses (**HDB**).

Pour rappel, les hypothèses (**HDB**) exprimant les engagements réciproques en fonction du degré moyen de la communauté, permettent d'obtenir un montant d'engagement réciproque **équivalent** entre deux assurés en relation. Aussi, d'après la relation 4.7, sous ces conditions la couverture mutuelle des franchises (i.e. somme des couvertures individuelles) coïncide avec la franchise mutuelle (i.e. somme des franchises individuelles). C'est-à-dire, la contribution mutuelle des assurés (i.e. somme des contributions individuelles) aux franchises permet de couvrir

la franchise mutuelle de la communauté. Cependant, ces conditions d'une part, n'assurent pas nécessairement la couverture de la dépense mutuelle (i.e. somme des dépenses individuelles) par la contribution mutuelle des assurés aux dépenses de la communauté, d'après la relation 4.9 et d'autre part, sont susceptibles d'entraîner des disparités dans les couvertures individuelles des charges (franchises et dépenses) en fonction de la dispersion de la distribution des degrés des assurés, d'après les relations 4.5 et 4.6.

Les Figures 6.3 (b) et 6.4 (b) représentent les sorties de la sensibilité engagements respectivement associées aux franchises et aux dépenses de la communauté **CF**, pour un degré moyen de $\bar{d} = 20, 12$ (degré moyen obtenu par la calibration du graphe social via le modèle à clusters homogènes). Les résultats pour l'étude de sensibilité graphe social seront complétés en 6.4.2.

On constate sur la Figure 6.3 (b), que pour une franchise moyenne (ou individuelle) de 1.000 €, la sensibilité engagements fournit une couverture moyenne des franchises de 1.000 ± 0 €, comme le scénario central et une contribution maximale aux franchises de 1.491 ± 0 € pour un assuré quelconque de la communauté **CF**. En effet, l'égalité entre la couverture mutuelle des franchises et la franchise mutuelle de la communauté évoquée plus-haut entraîne également l'égalité entre les valeurs moyennes respectives. Aussi, sous les conditions (**HDB**), le montant d'engagement réciproque entre deux assurés en relation vaut : $F/\bar{d} = 1.000/20, 12 = 49, 70$ €. Ainsi, la contribution maximale aux franchises de 1.491 € constatée, correspond à la contribution aux franchises des assurés composant le cluster de taille maximale (i.e. composé de 31 assurés). En effet, chacun des 31 assurés constituant ce cluster possède 30 relations, soit une contribution individuelle aux franchises de $30 \times 49, 70 = 1.491$ €. D'où une augmentation de 49,1% par rapport à la contribution maximale aux franchises de 1.000 € obtenue pour le scénario central.

Par ailleurs, d'après la Figure 6.4 (b), pour une dépense moyenne de 108 €, la sensibilité engagements fournit une couverture moyenne des dépenses de 105 ± 0 € et une contribution maximale aux dépenses de 257 ± 0 € pour un assuré quelconque de la communauté **CF**. En effet, comme évoqué plus-haut, sous les hypothèses (**HDB**), la couverture mutuelle des dépenses ne "couvre" pas nécessairement la dépense mutuelle de la communauté. Ce qui est également vrai pour les valeurs moyennes. On constate donc une diminution d'environ 3% par rapport à la couverture moyenne de 108 € obtenue pour le scénario central. Aussi, la contribution maximale d'un assuré aux dépenses pour cette variante, étant de 257 €, on observe donc une diminution de 16,6% par rapport au 308 € obtenue pour le scénario central. Ainsi, en comparaison au produit optimisé (i.e. le scénario central), cette variante du produit de franchise collaborative (i.e. la sensibilité engagements) permet de réduire la contribution maximale aux dépenses des assurés, cependant elle ne garantit pas la couverture de la dépense moyenne de la communauté **CF**.

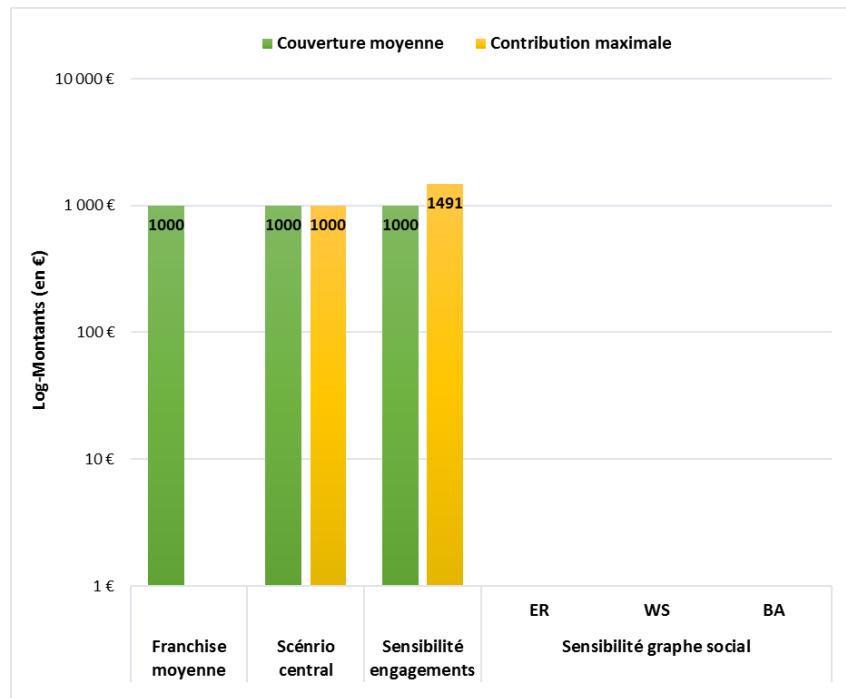


FIGURE 6.3 – Sorties (couverture moyenne et contribution maximale en échelle log) associées aux franchises par variante du produit de franchise collaborative : *Scénario central* (a), *Sensibilité engagements* (b), *Sensibilité graphe social* (c).

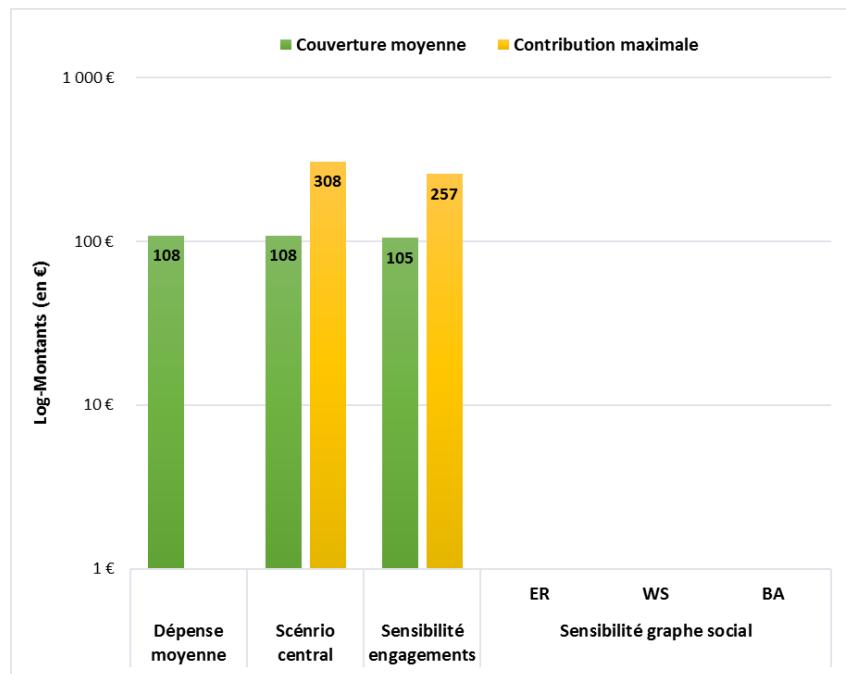


FIGURE 6.4 – Sorties (couverture moyenne et contribution maximale en échelle log) associées aux dépenses par variante du produit de franchise collaborative : *Scénario central* (a), *Sensibilité engagements* (b), *Sensibilité graphe social* (c).

6.4.2 Sensibilité graphe social

Cette étude vise à étudier l'impact du **graphe social** sur la couverture moyenne des charges et la contribution maximale aux charges des assurés. Ainsi, le montant des engagements réciproques reste identique à celui du scénario central, c'est-à-dire déterminé suivant les hypothèses (**HCO**). Cependant, à la différence du scénario central, dans lequel le graphe social de la communauté **CF** est calibré via le modèle à clusters homogènes, le graphe social est maintenant calibré via les trois modèles de graphe aléatoire étudiés : le modèle d'**Erdős-Rényi**, le modèle de **Watts-Strogatz** et le modèle de **Barabási-Albert**.

Pour rappel, ces modèles de graphe aléatoire permettent de prendre en compte les différents scénarios possibles dans le cas où la connexion entre les assurés serait **libre et sans contrainte**, c'est-à-dire les assurés peuvent se mettre en relation selon leur bon vouloir et peuvent appartenir à plusieurs groupes contrairement au modèle à clusters homogènes où les clusters sont indépendants.

Par ailleurs, par construction, la calibration du graphe social de la communauté **CF** via le scénario déterministe du modèle à clusters homogènes présenté en 5.2.1, fournit un degré moyen fixe de $\bar{d} = 20,12 \approx 20$ relations pour un ordre de $n = 493$ assurés. Ainsi, dans le but de pouvoir comparer les sorties des scénarios prenant en entrée les modèles de graphe aléatoire aux sorties du scénario central, les paramètres des modèles de graphe aléatoire sont choisis dans un premier temps, de telle sorte à simuler des graphes de degré moyen proche de $\bar{d} = 20$ relations.

Aussi, contrairement à la calibration via le modèle à clusters homogènes où le degré moyen de la communauté est fixe, lors de la calibration du graphe social via les modèles de graphe aléatoire, ce degré moyen peut être modifié. Ainsi, afin d'analyser l'impact du degré moyen sur les sorties de la variante sensibilité graphe social, le degré moyen de la communauté **CF** est dans un second temps, varié de 15 à 100 relations, soit 86 observations. Ce choix est motivé par le fait que pour un degré moyen en-dessous de $\bar{d} = 15$ avec $n = 493$, on obtient des distributions de degrés non "graphiques"¹ et pour un degré moyen au-dessus de $\bar{d} = 100$, on s'éloigne des scénarios réalistes sachant que l'un des principaux intérêts du produit de franchise collaborative est de responsabiliser les assurés en les connectant à un nombre raisonnable de relations.

Sensibilité graphe social VS scénario central

Les Figures 6.5 (c) et 6.6 (c) représentent les sorties (couverture moyenne en "vert" et contribution maximale en "orange") de la sensibilité graphe social respectivement associées aux franchises et aux dépenses de la communauté **CF**, pour un degré moyen proche de $\bar{d} = 20$ relations via chacun des trois modèles de graphe aléatoire, à savoir les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert.

On constate sur la Figure 6.5 (c), que pour une franchise moyenne (ou individuelle) de 1.000 €, la sensibilité graphe social fournit des couvertures moyennes de franchises de 1.000 ± 0 €, respectivement pour chacun des trois modèles de graphe aléatoire et des contributions maximales aux franchises de 1.756 ± 107 €, de 1.648 ± 83 € et de 45.989 ± 532 €, respectivement pour les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert. En effet, les hypothèses (**HCO**) permettent d'optimiser la couverture des charges des assurés, soit avec une couverture individuelle de franchise correspondant aux 1.000 € de franchise individuelle. D'où la couverture moyenne de franchises de 1.000 € constatée pour ces deux variantes (scénario central et sensibilité graphe social) du produit de franchise collaborative prenant en entrées des engagements

1. Une suite finie $d = (d_1, d_2, \dots, d_n)$ d'entiers positifs est dite "graphique" s'il existe un graphe d'ordre n dont la suite des degrés des sommets est d .

réciproques calibrés suivant les conditions (**HCO**). Aussi, en comparaison aux 1000 € de contribution maximale aux franchises obtenues pour le scénario central, on observe respectivement des augmentations de 75,6%, de 64,8% et de 4.498,9% pour les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert.

Par ailleurs, d'après la Figure 6.6 (c), pour une dépense moyenne de 108 €, la sensibilité graphe social fournit des couvertures moyennes de dépenses de 108 ± 0 €, respectivement pour chacun des trois modèles de graphe aléatoire et des contributions maximales aux franchises de 386 ± 43 €, de 381 ± 39 € et de 5.155 ± 61 €, respectivement pour les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert. En effet, les hypothèses (**HCO**) permettent d'optimiser la couverture des charges de la communauté en fournissant des couvertures individuelles de dépenses correspondant aux dépenses respectives des assurés. Ce qui explique la couverture moyenne de dépenses de 108 € constatée pour ces deux variantes (scénario central et sensibilité graphe social) du produit de franchise collaborative prenant en entrée des engagements réciproques calibrés suivant les conditions (**HCO**). Aussi, en comparaison aux 308 € de contribution maximale aux dépenses obtenus pour le scénario central, on observe respectivement des augmentations de 25,3%, de 23,7% et de 1.573,7% pour les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert.

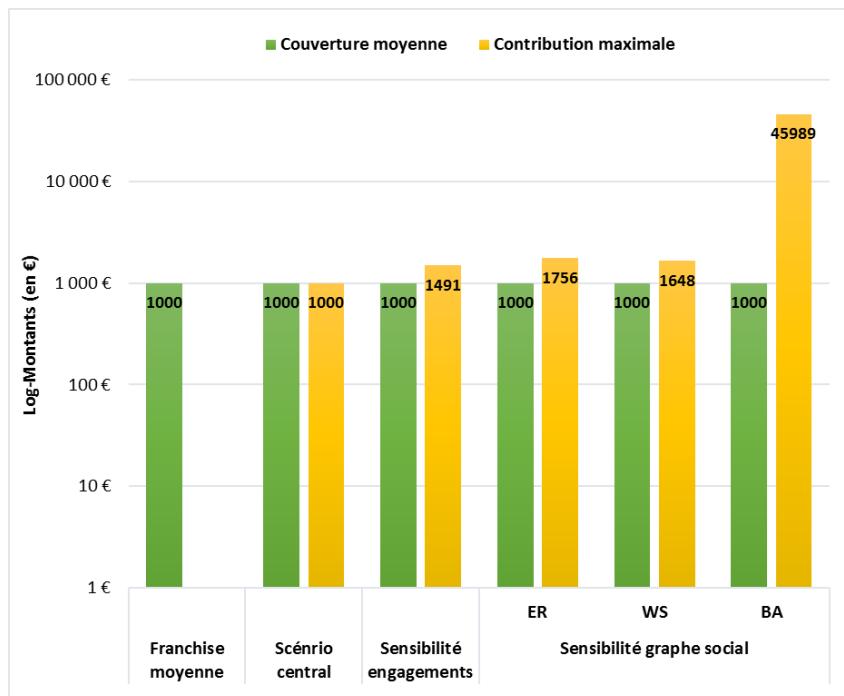


FIGURE 6.5 – Sorties (couverture moyenne et contribution maximale en échelle log) associées aux franchises par variante du produit de franchise collaborative : *Scénario central* (a), *Sensibilité engagements* (b), *Sensibilité graphe social* (c).

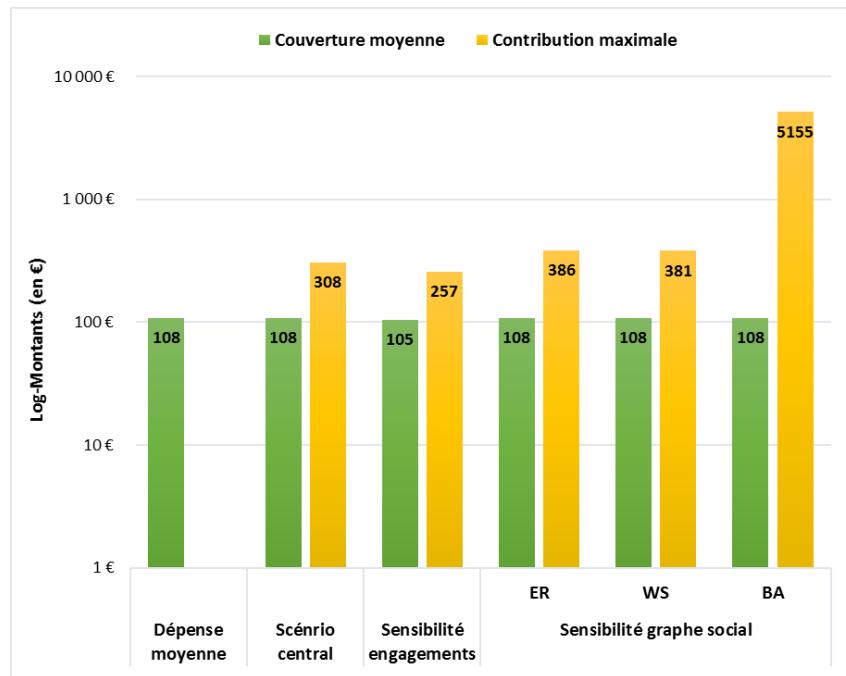


FIGURE 6.6 – Sorties (couverture moyenne et contribution maximale en échelle log) associées aux dépenses par variante du produit de franchise collaborative : *Scénario central* (a), *Sensibilité engagements* (b), *Sensibilité graphe social* (c).

Sorties de la sensibilité graphe social en fonction du degré moyen

Ce point consiste à présenter les sorties (i.e. la couverture moyenne et la contribution maximale) du programme **PC2** pour la variante sensibilité graphe social du produit de franchise collaborative, prenant en entrée les hypothèses (**HCO**) pour la détermination du montant des engagements réciproques entre les assurés et les trois modèles de graphe aléatoire étudiés pour la calibration du graphe social de la communauté **CF**.

Pour rappel, afin d'analyser l'impact du degré moyen de la communauté sur ces sorties, le degré moyen de la communauté **CF** est varié de 15 à 100 relations. Et les sorties pour chaque valeur du degré moyen, sont obtenues en prenant la moyenne sur 1.000 simulations de scénarios de graphe social via chacun des trois modèles de graphe aléatoire, par application de la méthode de Monte-Carlo (cf. section 6.2 pour plus de détails sur le fonctionnement du programme de calcul **PC2**). Aussi, sachant que sous les hypothèses (**HCO**), les couvertures moyennes des charges (franchises et dépenses) des assurés coïncident avec les charges moyennes respectives (i.e. 1000 € pour la franchise moyenne et 108 € pour la dépense moyenne) de la communauté **CF**, des intervalles de confiance à 95% sont déterminés uniquement pour la contribution maximale aux charges des assurés en fonction du degré moyen et du modèle de graphe aléatoire, sous la forme :

$$IC(0,95) = \bar{X}_N \pm e_N , \text{ avec } e_N = 1,96 \times \frac{\hat{\sigma}_N}{\sqrt{N}} ,$$

où pour chaque degré moyen variant de 15 à 100 relations, on a :

- ✓ $N = 1.000$, représente le nombre de simulations via un modèle de graphe aléatoire donné ;
- ✓ e_N , représente la marge d'erreur d'estimation de la moyenne théorique de la contribution maximale pour un niveau de confiance à 95% et pour N simulations ;
- ✓ \bar{X}_N , représente la moyenne empirique sur l'ensemble des N observations de contribution maximale associées aux N simulations ;

- ✓ $\hat{\sigma}_N$, représente l'écart-type empirique sur l'ensemble des N observations de contribution maximale associées aux N simulations.
- ❖ **Résultats associés aux franchises :** La Figure 6.7 et le Tableau 6.2 représentent respectivement l'évolution des sorties (couverture moyenne et contribution maximale) associées aux franchises et les intervalles de confiance à 95% pour les contributions maximales aux franchises de la variante sensibilité graphe social en fonction du degré moyen de la communauté CF.

D'après la Figure 6.7, on constate que quelque soit le modèle de graphe aléatoire, la couverture moyenne des franchises est constante et coïncide avec les 1.000 € de franchise moyenne de la communauté CF. Ceci peut s'expliquer par la relation 4.18, exprimant l'égalité entre la couverture mutuelle des franchises (i.e. la somme des couvertures individuelles) et la franchise mutuelle (i.e. la somme des franchises individuelles), sous les conditions (**HCO**). Ainsi, la couverture moyenne des franchises des assurés ne dépend pas du degré moyen de la communauté des assurés, pour la variante sensibilité graphe social du produit de franchise collaborative.

En revanche, on constate que la contribution maximale aux franchises décroît de 1.889 € à 1.283 €, de 1.810 € à 1.256 € et de 67.023 € à 7.217 €, respectivement pour les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert. Ainsi, la contribution maximale aux franchises est une fonction décroissante du degré moyen. Le point suivant consistera à expliciter cette fonction par une formule fermée exprimant la contribution maximale aux franchises des assurés en fonction du degré moyen de la communauté.

Par ailleurs, quelque soit le degré moyen de la communauté, on constate que la contribution maximale aux franchises obtenue via le modèle de Watts-Strogatz reste en-dessous de celle obtenue via le modèle d'Erdős-Rényi, qui elle-même reste en-dessous de la contribution maximale aux franchises obtenue via le modèle de Barabási-Albert. Ceci peut s'expliquer par le fait que de ces trois modèles de graphe aléatoire, le modèle de Watts-Strogatz fournit une distribution de degrés des assurés plus homogène et donc des contributions individuelles aux franchises moins volatiles et moins extrêmes.

D'après le Tableau 6.2, pour $N = 1.000$ simulations et un niveau de confiance à 95%, on obtient une marge d'erreur relativement faible et diminuant lorsque le degré moyen de la communauté augmente : en effet, elle décroît de 7,16 € à 2,14 €, de 6,99 € à 2,01 € et de 49,77 € à 8,27 €, respectivement pour les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert. Par exemple, pour un degré moyen de 20 relations, on obtient avec un niveau de confiance à 95%, une contribution maximale aux franchises de $1.755,83 \pm 6,63$ €, de $1.648,38 \pm 5,13$ € et de $45.989,29 \pm 33,01$ €, respectivement pour chacun des trois modèles de graphe aléatoire.

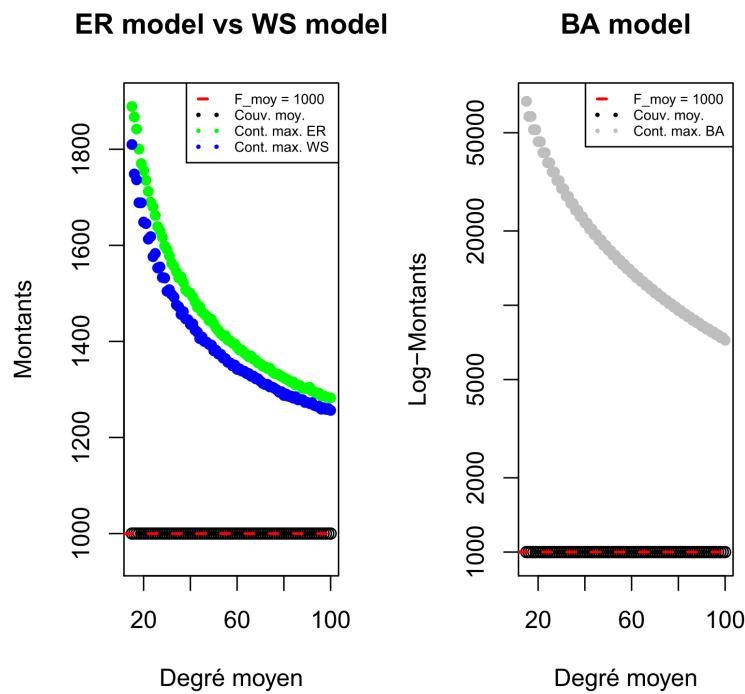


FIGURE 6.7 – Sorties (couverture moyenne et contribution maximale) de la variante sensibilité graphe social associées aux franchises en fonction du degré moyen (F_moy représente la franchise moyenne des assurés de la communauté **CF**) : à gauche (échelle linéaire : *Erdős-Rényi* (a), *Watts-Strogatz* (b)) et à droite (échelle logarithmique : *Barabási-Albert* (c)).

Intervalle de confiance	Erdős-Rényi		Watts-Strogatz		Barabási-Albert	
	Degrés moyens	Moyenne	Marge	Moyenne	Marge	Moyenne
15	1.888,98	7,16	1.809,91	6,99	67.023,27	49,77
20	1.755,83	6,63	1.648,38	5,13	45.989,29	33,01
25	1.662,17	5,71	1.583,17	5,06	37.799,89	26,74
30	1.591,28	4,67	1.504,37	3,91	29.680,56	21,16
35	1.532,40	4,33	1.472,06	3,61	25.847,78	19,37
40	1.499,99	3,98	1.435,28	3,42	21.552,53	16,72
45	1.461,05	3,66	1.409,09	3,29	19.334,21	15,44
50	1.436,77	3,66	1.380,61	2,98	16.694,13	14,16
55	1.412,24	3,19	1.364,28	2,85	15.265,02	13,66
60	1.388,22	3,15	1.341,75	2,75	13.489,11	12,69
65	1.367,75	2,86	1.333,18	2,64	12.483,64	11,57
70	1.353,09	2,75	1.316,98	2,50	11.199,50	11,44
75	1.337,43	2,61	1.305,11	2,44	10.470,08	10,67
80	1.324,49	2,57	1.287,42	2,33	9.510,53	10,22
85	1.309,03	2,46	1.284,93	2,21	8.971,61	9,92
90	1.302,90	2,42	1.272,96	2,19	8.232,08	9,86
95	1.292,66	2,30	1.265,25	2,11	7.805,40	8,29
100	1.282,55	2,14	1.256,42	2,01	7.217,33	8,27

TABLE 6.2 – Intervalles de confiance à 95% : Contribution maximale (en €) aux franchises de la variante sensibilité graphe social en fonction du degré moyen.

❖ **Résultats associés aux dépenses :** La Figure 6.8 et le Tableau 6.3 représentent respectivement l'évolution des sorties (couverture moyenne et contribution maximale) associées aux dépenses et les intervalles de confiance à 95% pour les contributions maximales aux dépenses de la variante sensibilité graphe social en fonction du degré moyen de la communauté **CF**.

Comme pour les franchises, on constate d'après la Figure 6.8, que quelque soit le modèle de graphe aléatoire, la couverture moyenne des dépenses est constante et coïncide avec les 108 € de dépense moyenne de la communauté **CF**. Ce qui s'explique également par la relation 4.19, exprimant l'égalité entre la couverture mutuelle des dépenses (i.e. la somme des couvertures individuelles) et la dépense mutuelle (i.e. la somme des dépenses individuelles), sous les conditions (**HCO**). Ainsi, à l'instar de la couverture moyenne des franchises, celle associée aux dépenses des assurés ne dépend pas non plus du degré moyen de la communauté, pour la variante sensibilité graphe social du produit de franchise collaborative.

Aussi, on constate que la contribution maximale aux dépenses décroît de 449,70 € à 201,20 €, de 457,90 € à 201,80 € et de 7.420,90 € à 857,10 €, respectivement pour les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert. Ainsi, la contribution maximale aux dépenses est une fonction décroissante du degré moyen. Le point suivant consistera à expliciter cette fonction par une formule fermée exprimant la contribution maximale aux dépenses des assurés en fonction du degré moyen de la communauté.

Par ailleurs, quelque soit le degré moyen de la communauté, on constate que la contribution maximale aux dépenses obtenue via le modèle d'Erdős-Rényi est sensiblement confondue à celle obtenue via le modèle de Watts-Strogatz. Et tous deux restent toujours en-dessous de la contribution maximale aux dépenses obtenue via le modèle de Barabási-Albert. Ceci peut s'expliquer par le fait que d'une part, la contribution aux dépenses est étroitement liée aux dépenses des assurés et d'autre part, les modèles d'Erdős-Rényi et de Watts-Strogatz fournissent des distributions de degrés relativement homogènes et donc des contributions individuelles aux dépenses moins volatiles et moins extrêmes par rapport au modèle de Barabási-Albert qui permet d'avoir une distribution de degrés plus hétérogène.

D'après le Tableau 6.3, pour $N = 1.000$ simulations et un niveau de confiance à 95%, on obtient une marge d'erreur relativement faible et diminuant lorsque le degré moyen de la communauté augmente : en effet, elle décroît de 3,43 € à 0,72 €, de 3,79 € à 0,80 € et de 5,42 € à 1,25 €, respectivement pour les modèles d'Erdős-Rényi, de Watts-Strogatz et de Barabási-Albert. Par exemple, pour un degré moyen de 100 relations, on obtient avec un niveau de confiance à 95%, une contribution maximale aux dépenses de $201,19 \pm 0,72$ €, de $201,82 \pm 0,80$ € et de $857,07 \pm 1,25$ €, respectivement pour chacun des trois modèles de graphe aléatoire.

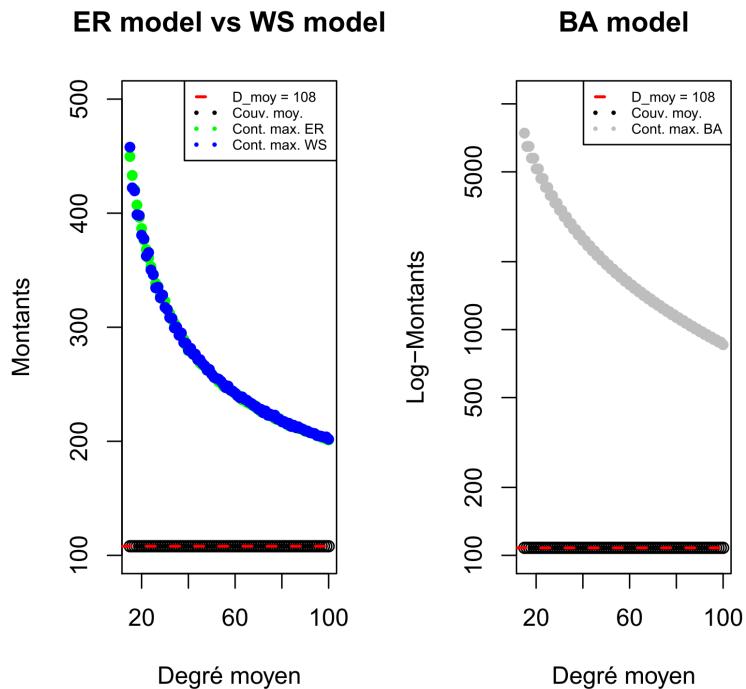


FIGURE 6.8 – Sorties (couverture moyenne et contribution maximale) de la variante sensibilité graphe social associées aux dépenses en fonction du degré moyen (D_{moy} représente la dépense moyenne des assurés de la communauté **CF**) : à gauche (échelle linéaire : *Erdős-Rényi* (a), *Watts-Strogatz* (b)) et à droite (échelle logarithmique : *Barabási-Albert* (c)).

Intervalle de confiance	Erdős-Rényi		Watts-Strogatz		Barabási-Albert	
	Degrés moyens	Moyenne	Marge	Moyenne	Marge	Moyenne
15	449,66	3,43	457,86	3,79	7.420,92	5,42
20	386,55	2,66	380,69	2,40	5.155,44	3,81
25	346,56	2,27	346,05	2,23	4.268,34	3,51
30	322,83	1,99	317,44	1,86	3.380,75	2,89
35	298,68	1,66	300,20	1,73	2.954,76	2,69
40	279,31	1,47	279,89	1,57	2.482,83	2,44
45	268,67	1,38	271,64	1,42	2.234,19	2,29
50	258,51	1,32	258,93	1,27	1.938,93	2,08
55	248,21	1,23	249,95	1,24	1.779,10	1,96
60	241,99	1,19	242,34	1,14	1.578,93	1,88
65	234,16	1,08	236,07	1,06	1.466,47	1,85
70	228,90	1,03	228,01	0,98	1.320,76	1,67
75	222,26	0,96	223,80	0,93	1.237,42	1,62
80	217,08	0,86	217,13	0,89	1.129,98	1,58
85	213,01	0,86	214,12	0,86	1.064,62	1,56
90	208,74	0,82	209,23	0,82	978,81	1,46
95	204,93	0,75	205,06	0,81	928,18	1,39
100	201,19	0,72	201,82	0,80	857,07	1,25

TABLE 6.3 – Intervalles de confiance à 95% : Contribution maximale (en €) aux dépenses de la variante sensibilité graphe social en fonction du degré moyen.

Régression polynomiale : Contribution maximale en fonction du degré moyen

Dans le point précédent consacré à la présentation des sorties du programme **PC2** pour la variante sensibilité graphe social en fonction du degré moyen de la communauté **CF**, on a pu constater que la contribution maximale aux charges des assurés est une fonction décroissante du degré moyen, quelque soit le modèle de graphe aléatoire. L'objectif de ce point consiste à expliciter cette fonction par une formule fermée permettant de déterminer la contribution maximale des assurés aux charges (franchises et dépenses) de la communauté **CF** en fonction du degré moyen de cette communauté. Pour ce faire, on effectue une regression polynomiale² de la contribution maximale expliquée par le degré moyen, respectivement pour les contributions aux franchises et aux dépenses de la communauté **CF**, obtenues pour chaque modèle de graphe.

❖ **Contribution maximale aux franchises :** La Figure 6.9 représente l'évolution de la contribution maximale d'un assuré aux franchises de ses relations en fonction du degré moyen de la communauté **CF**, pour chaque modèle de graphe.

✓ **Modèle à clusters homogènes :** Par construction du modèle à clusters homogènes, il fournit un degré moyen fixe de $\bar{d} = 20,12 \approx 20$ relations. De plus, sous les hypothèses (**HCO**), ce modèle fournit une contribution individuelle aux franchises des assurés exactement égale aux 1.000 € de franchise individuelle. Ainsi, comme illustré par la Figure 6.9 (a), quelque soit le degré moyen de la communauté (quitte à modifier l'ordre de la communauté, i.e. le nombre d'assurés), la contribution maximale d'un assuré aux franchises de ses relations, via le modèle à clusters homogènes, est constante et égale au montant de franchise individuelle, soit :

$$y = 1.000 ,$$

avec y représentant la contribution maximale d'un assuré aux franchises de la communauté **CF** ;

✓ **Modèle d'Erdős-Rényi :** La contribution maximale d'un assuré aux franchises de ses relations en fonction du degré moyen de la communauté **CF**, via le modèle d'Erdős-Rényi, est représentée par la Figure 6.9 (b). Une regression polynomiale de degré 5 de la contribution maximale en fonction du degré moyen, permet d'obtenir les résultats suivants : le R^2 ajusté égal à 0,9996 et l'hypothèse de nullité des coefficients du modèle est rejetée à 99,9%. Ainsi, la contribution maximale aux franchises peut être calculée selon l'équation suivante sous la forme $y = P_{er}(x)$, où P_{er} est un polynôme de degré 5 :

$$y = -6,3 \cdot 10^{-7} \times x^5 + 2,2 \cdot 10^{-4} \times x^4 - 2,9 \cdot 10^{-2} \times x^3 + 1,9 \times x^2 - 73,7 \times x + 2.636 ,$$

avec x et y représentant respectivement le degré moyen et la contribution maximale d'un assuré aux franchises de la communauté **CF** ;

✓ **Modèle de Watts-Strogatz :** La Figure 6.9 (c) représente la contribution maximale d'un assuré aux franchises de ses relations en fonction du degré moyen de la communauté **CF**, via le modèle de Watts-Strogatz. Un ajustement polynomiale de degré 5 de la contribution maximale en fonction du degré moyen, donne les résultats suivants : le R^2 ajusté égal à 0,998 et l'hypothèse de nullité des coefficients du modèle est rejetée à 99,9%. Ainsi, la contribution maximale aux franchises peut être calculée selon la formule fermée suivante sous la forme $y = P_{ws}(x)$, où P_{ws} est un polynôme de degré 5 :

$$y = -7,1 \cdot 10^{-7} \times x^5 + 2,3 \cdot 10^{-4} \times x^4 - 3,1 \cdot 10^{-2} \times x^3 + 2,1 \times x^2 - 73,2 \times x + 2.511 ,$$

avec x et y représentant respectivement le degré moyen et la contribution maximale d'un assuré aux franchises de la communauté **CF** ;

2. La régression polynomiale est une analyse statistique qui décrit la variation d'une variable aléatoire expliquée à partir d'une fonction polynomiale d'une variable aléatoire explicative.

✓ **Modèle de Barabási-Albert** : La contribution maximale d'un assuré aux franchises de ses relations en fonction du degré moyen de la communauté **CF**, via le modèle de Barabási-Albert, est représentée sur la Figure 6.9 (d). Par une regression polynomiale de degré 5 de la contribution maximale en fonction du degré moyen, on obtient les résultats suivants : le R^2 ajusté égal à 0,9973 et l'hypothèse de nullité des coefficients du modèle est rejetée à 99,9%. Ainsi, l'équation suivante sous la forme $y = P_{ba}(x)$, où P_{ba} est un polynôme de degré 5, permet le calcul de la contribution maximale aux franchises d'un assuré en fonction du degré moyen de la communauté **CF** :

$$y = -1,1 \cdot 10^{-4} \times x^5 + 3,6 \cdot 10^{-2} \times x^4 - 4,7 \times x^3 + 302,8 \times x^2 - 1.019 \times x + 16.250 ,$$

avec x et y représentant respectivement le degré moyen et la contribution maximale d'un assuré aux franchises de la communauté **CF**.

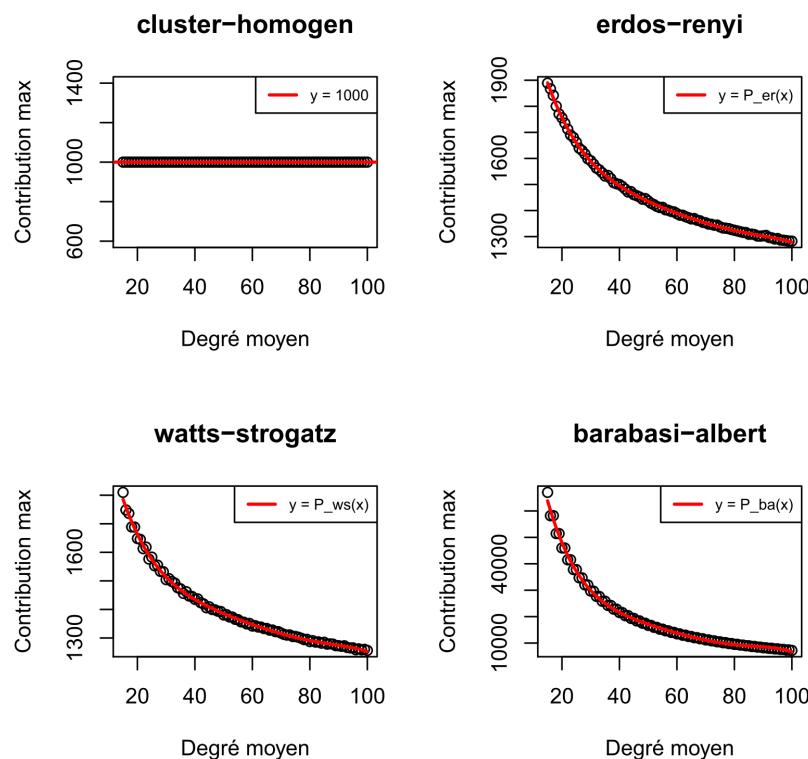


FIGURE 6.9 – Contribution maximale aux franchises en fonction du degré moyen, sous les hypothèses (**HCO**) et par modèle de graphe : *Clusters homogènes* (a), *Erdős-Rényi* (b), *Watts-Strogatz* (c) et *Barabási-Albert* (d).

❖ **Contribution maximale aux dépenses :** La Figure 6.10 représente l'évolution de la contribution maximale d'un assuré aux dépenses de ses relations en fonction du degré moyen de la communauté **CF**, pour chaque modèle de graphe excepté le modèle à clusters homogènes. En effet, le calcul de contribution maximale aux dépenses étant fonction des dépenses des 493 assurés de la communauté **CF** et le modèle à clusters homogènes fournissant un degré moyen fixe d'environ 20 relations, modifier le degré moyen de la communauté, revient à modifier l'ordre (i.e. le nombre d'assurés) de la communauté et donc à ne pas prendre en compte l'intégralité des dépenses des 493 membres du portefeuille **CF**.

✓ **Modèle d'Erdős-Rényi :** La contribution maximale d'un assuré aux dépenses de ses relations en fonction du degré moyen de la communauté **CF**, via le modèle d'Erdős-Rényi, est représentée par la Figure 6.10 (a). Une regression polynomiale de degré 5 de la contribution maximale en fonction du degré moyen, permet d'obtenir les résultats suivants : le R^2 ajusté égal à 0,9995 et l'hypothèse de nullité des coefficients du modèle est rejetée à 99,9%. Ainsi, la contribution maximale aux dépenses peut être calculée selon l'équation suivante sous la forme $y = P_{er}(x)$, où P_{er} est un polynôme de degré 5 :

$$y = -3,2 \cdot 10^{-7} \times x^5 + 1,1 \cdot 10^{-4} \times x^4 - 1,4 \cdot 10^{-2} \times x^3 + 0,9 \times x^2 - 33,4 \times x + 776,6 ,$$

avec x et y représentant respectivement le degré moyen et la contribution maximale d'un assuré aux dépenses de la communauté **CF** ;

✓ **Modèle de Watts-Strogatz :** La Figure 6.10 (b) représente la contribution maximale d'un assuré aux dépenses de ses relations en fonction du degré moyen de la communauté **CF**, via le modèle de Watts-Strogatz. Un ajustement polynomiale de degré 5 de la contribution maximale en fonction du degré moyen, donne les résultats suivants : le R^2 ajusté égal à 0,9978 et l'hypothèse de nullité des coefficients du modèle est rejetée à 99,9%. Ainsi, la contribution maximale aux dépenses peut être calculée selon la formule fermée suivante sous la forme $y = P_{ws}(x)$, où P_{ws} est un polynôme de degré 5 :

$$y = -3,5 \cdot 10^{-7} \times x^5 + 1,2 \cdot 10^{-4} \times x^4 - 1,5 \cdot 10^{-2} \times x^3 + 1,01 \times x^2 - 35 \times x + 787,4 ,$$

avec x et y représentant respectivement le degré moyen et la contribution maximale d'un assuré aux dépenses de la communauté **CF** ;

✓ **Modèle de Barabási-Albert :** La contribution maximale d'un assuré aux dépenses de ses relations en fonction du degré moyen de la communauté **CF**, via le modèle de Barabási-Albert, est représentée sur la Figure 6.10 (c). Par une regression polynomiale de degré 5 de la contribution maximale en fonction du degré moyen, on obtient les résultats suivants : le R^2 ajusté égal à 0,9974 et l'hypothèse de nullité des coefficients du modèle est rejetée à 99,9%. Ainsi, l'équation suivante sous la forme $y = P_{ba}(x)$, où P_{ba} est un polynôme de degré 5, permet le calcul de la contribution maximale aux dépenses d'un assuré en fonction du degré moyen de la communauté **CF** :

$$y = -1,1 \cdot 10^{-5} \times x^5 + 3,8 \cdot 10^{-3} \times x^4 - 0,5 \times x^3 + 32,2 \times x^2 - 1.090 \times x + 17.660 ,$$

avec x et y représentant respectivement le degré moyen et la contribution maximale d'un assuré aux dépenses de la communauté **CF**.

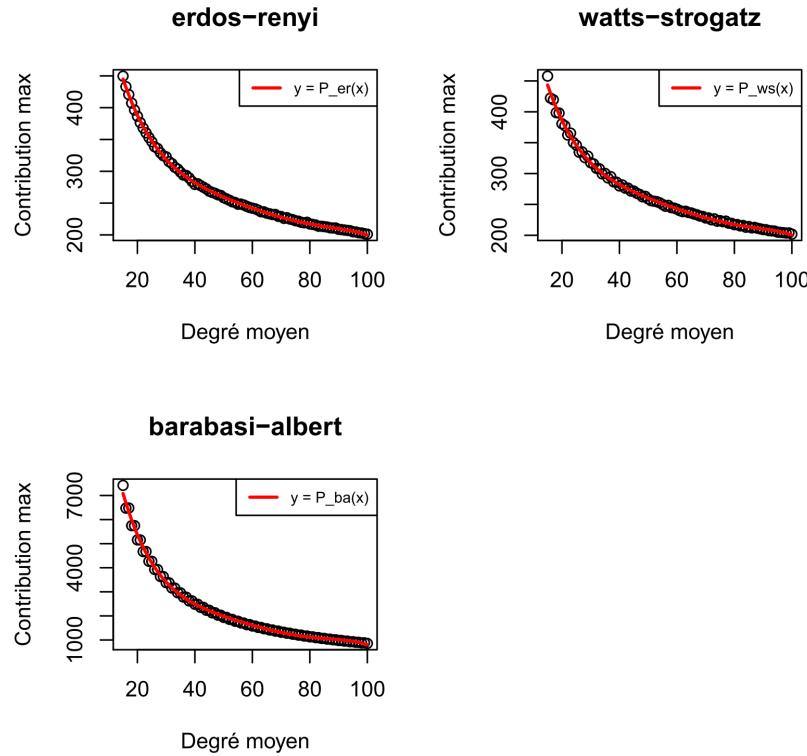


FIGURE 6.10 – Contribution maximale aux dépenses en fonction du degré moyen, sous les hypothèses (**HCO**) et par modèle de graphe aléatoire : *Erdős-Rényi* (a), *Watts-Strogatz* (b) et *Barabási-Albert* (c).

6.5 Conclusion

Partant des charges annuelles associées aux contrats d’assurance auto respectifs des assurés de la communauté **CF**, ce chapitre a consisté à établir trois variantes du produit de franchise collaborative, notamment le **scénario central** permettant d’optimiser la couverture moyenne des charges et la contribution maximale aux charges des assurés de la communauté, la **sensibilité engagements** et la **sensibilité graphe social** permettant respectivement d’évaluer les impacts du montant des engagements réciproques et du modèle de graphe social sur ces sorties du produit de franchise collaborative. Aussi, la variante sensibilité graphe social a permis, à l’aide d’une regression polynômiale, d’établir une formule fermée permettant de déterminer la contribution maximale d’un assuré aux charges annuelles de ses relations en fonction du degré moyen de la communauté **CF**.

Par ailleurs, le chapitre suivant consistera à appliquer la méthodologie de calcul du programme **PC2** à des données réelles, c’est-à-dire connaissant le graphe social décrivant le réseau des connexions entre les membres des communautés respectives, il s’agira d’analyser l’influence des engagements réciproques sur les sorties du produit de franchise collaborative.

Chapitre 7

Applications à des données réelles

L'objectif de ce chapitre consiste en des applications de la méthodologie de calcul des distributions des variables d'intérêt (couverture et contribution) du produit de franchise collaborative via le programme **PC2** présenté dans la section 6.2, à des données réelles notamment des Données de Généalogie (**DG**) et des Données de réseaux sociaux tels que Facebook (**DF**). En effet, les membres des communautés respectives **DG** et **DF** sont supposés détenir chacun un contrat d'assurance auto et appartenir à une plateforme de franchise collaborative dans le but de couvrir mutuellement les franchises associées à leurs contrats d'assurance respectifs, en cas de sinistre.

Aussi, sachant que les paramètres du produit de franchise collaborative sont le modèle de graphe social et le montant de engagements réciproques entre les assurés, et connaissant les graphes sociaux représentant les connexions entre les membres respectifs des communautés **DG** et **DF**, ce chapitre consiste à mettre en exergue l'influence des engagements réciproques sur les variables d'intérêt du produit de franchise collaborative pour chacune de ces deux communautés d'assurés. Ainsi, les charges annuelles considérées dans le cadre de ce chapitre sont essentiellement les franchises individuelles fixées à 1.000 €, associées aux contrats d'assurance auto respectifs des assurés de chacune des communautés **DG** et **DF**.

7.1 Données de Généalogie (**DG**)

Cette section fait tout d'abord, une description des Données de Généalogie (**DG**), ensuite une analyse des propriétés du graphe social de la communauté **DG**. Enfin, une présentation des sorties (i.e. couverture moyenne et contribution maximale) associées aux franchises des assurés de la communauté **DG**, est effectuée en fonction des hypothèses (**HDB**) et (**HCO**) sur les engagements réciproques.

7.1.1 Description des données **DG**

Les Données de Généalogie (**DG**) étudiées dans cette section sont issues du site web Geneanet¹. Il s'agit de données collectées par des amateurs de généalogie à la recherche de leurs ancêtres. Les assurés du service peuvent construire leur arbre généalogique en renseignant, de manière plus ou moins détaillée, les informations glanées au cours de leurs recherches.

La base de données **DG** a été récueillie au format CSV. Il s'agit de 19 fichiers CSV correspondants à 19 familles de cousins. Pour chaque famille, on a une personne, ses frères et sœurs, des cousins, les cousins des cousins, etc. L'idée est de supposer que les personnes s'assurent mutuellement entre cousins de même famille. Le Tableau 1 (cf. Annexe A) résume le nombre

1. <https://www.geneanet.org/>.

d'observations (lignes) initiales et restantes après traitements (i.e. après suppression des doublons et des boucles²) pour chacune des 19 familles de la base **DG**. Aussi, le Tableau 2 (cf. Annexe A) récapitule les différentes variables (colonnes) présentes initialement dans la base **DG** et les variables conservées pour la suite des travaux.

Afin de déterminer le graphe social puis les degrés des assurés de la communauté **DG** composée de 19 familles ou clusters, seulement les variables *id_personne_num* et *id_family_member*, représentant respectivement l'identifiant d'une personne et l'identifiant de sa relation (i.e. son cousin ou son *sibling*³), sont conservées sur l'ensemble des 14 variables initiales de la base **DG** (cf. Tableau 2 en Annexe A). En effet, ces deux variables décrivant les relations entre les cousins de même famille, suffisent pour déterminer le graphe social de la communauté **DG** et *in fine* déduire les degrés des assurés de la communauté.

7.1.2 Propriétés du graphe social de la communauté DG

Partant de la base initiale **DG** et des traitements effectués ayant permis de conserver les observations (lignes) et les variables (colonnes) représentant les liens entre les membres de chaque famille, on déduit le graphe social de la communauté **DG** composée de 19 familles ou clusters (cf. Figure 7.1).

On obtient des clusters (ou familles) de tailles allant de 2 à 34 membres, pour des degrés moyens allant de 1 à 14 relations. Aussi, l'ordre du graphe social de la communauté **DG** est de 190 assurés pour un degré moyen d'environ 8 relations. Par ailleurs, on constate que trois familles ne fournissent pas des clusters homogènes, c'est-à-dire les membres respectifs de ces familles n'ont pas le même nombre de relations. Ceci peut s'expliquer par le fait que des frères consanguins ou utérins (i.e uniquement de même père ou de même mère, respectivement) peuvent avoir des cousins différents d'où des relations distinctes.

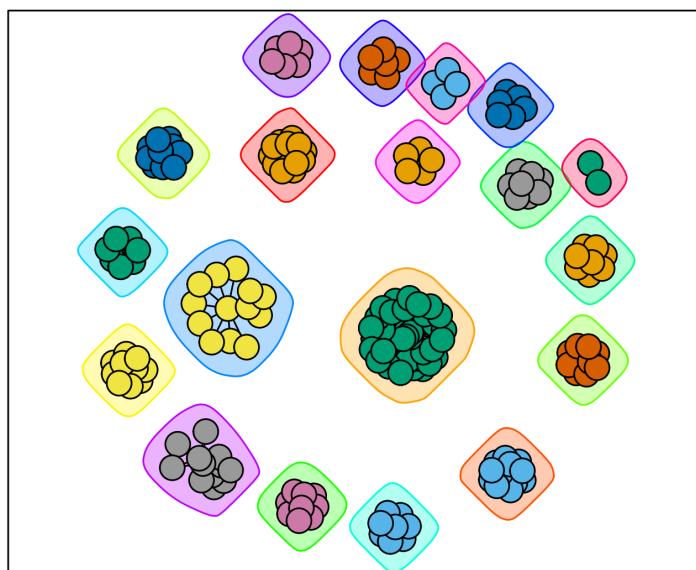


FIGURE 7.1 – Graphe social de la communauté DG

Dans le but d'obtenir un graphe social pour la communauté **DG**, suivant le modèle à clusters homogènes permettant d'optimiser le produit de franchise collaborative, des retraitements sont

2. On a une boucle lorsque l'identifiant d'une personne coïncide avec l'identifiant de sa relation.
3. *sibling* : désigne la fratrie (frères et soeurs).

effectués sur les connexions des assurés de ces trois familles à clusters non homogènes afin d'obtenir des sous-graphes complets (i.e. des cliques). Ce qui permet d'avoir des clusters homogènes (i.e. des familles) de degrés moyens allant de 1 à 33 relations et un degré moyen d'environ 13 relations pour l'ensemble de la communauté **DG**. La Figure 7.2 représente les tailles et les degrés moyens respectifs des familles de la communauté **DG**, avant et après les retraitements en clusters homogènes.

On constate que les retraitements en clusters homogènes permettent de faire passer le degré moyen de la communauté **DG** de 8 à 13 relations, soit une augmentation de 62,5% par rapport à l'état initial (i.e. avant retraitements). Ce qui permettrait donc, sous les hypothèses (**HDB**), de diminuer le montant des engagements réciproques et donc des contributions individuelles aux franchises, d'après leurs expressions mathématiques.

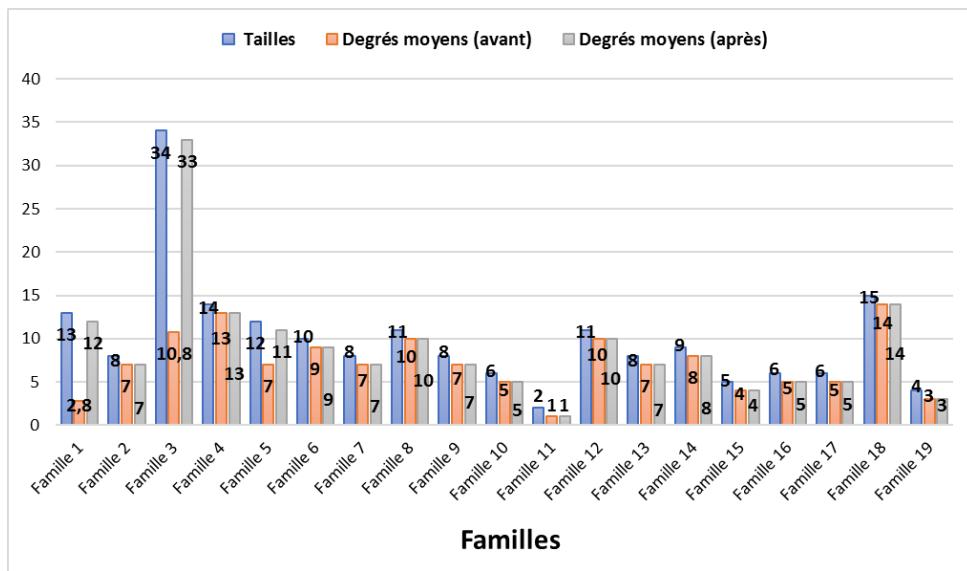


FIGURE 7.2 – Tailles et degrés moyens par famille de la communauté **DG**, avant et après les retraitements en clusters homogènes.

7.1.3 Sorties associées aux franchises de la communauté DG

Cette sous-section consiste à présenter les sorties (i.e. couverture moyenne et contribution maximale) associées aux franchises des assurés de la communauté **DG**, avant et après les retraitements en clusters homogènes. Partant des graphes sociaux respectifs de la communauté **DG** avant et après les retraitements, on déduit les distributions des degrés des assurés pour chacun des deux cas (avant et après retraitements).

La Figure 7.3 représente les sorties associées aux franchises des assurés de la communauté **DG** avant et après les retraitements en clusters homogènes, sous respectivement les hypothèses (**HDB**) et (**HCO**) concernant les engagements réciproques entre les assurés.

D'après la Figure 7.3, on constate que les hypothèses (**HDB**) et (**HCO**) permettent tous deux, d'obtenir une couverture moyenne de franchise coïncidant avec les 1.000 € de franchise individuelle. c'est-à-dire, en moyenne un assuré de la communauté possède une couverture de franchise équivalente au montant de sa franchise individuelle. Ceci peut s'expliquer par les relations 4.7 et 4.18 exprimant l'égalité entre la couverture mutuelle de franchise (i.e. la somme des couvertures individuelles) et la franchise mutuelle (i.e. la somme des franchises individuelles) de la

communauté, respectivement sous les conditions (**HDB**) et (**HCO**). Aussi, sous les hypothèses (**HDB**), les retraitements en clusters homogènes permettent comme annoncé au point 7.1.2, de diminuer la contribution maximale d'un assuré de 3.899 € à 2.476 €, soit une réduction de 36,5% par rapport à l'état initial (i.e. avant retraitements). De même, sous les hypothèses (**HCO**), les retraitements en clusters homogènes permettent de diminuer la contribution maximale d'un assuré de 9.000 € à 1.000 €, soit 9 fois moins coûteux par rapport à l'état initial. Ainsi, le modèle à clusters homogènes permet de minimiser la contribution maximale d'un assuré aux franchises de ses relations quelles que soient les hypothèses sur les engagements réciproques entre les assurés.

Par ailleurs, suivant la calibration du graphe social de la communauté **DG** via le modèle à clusters homogènes, on obtient des contributions maximales aux franchises de 2.476 € et de 1.000 €, respectivement sous les conditions (**HDB**) et (**HCO**), soit une économie de 1.476 € sous les hypothèses (**HCO**). Ce qui vient confirmer l'optimalité du **scénario central** pour le produit de franchise collaborative, c'est-à-dire, le scénario prenant en entrée les paramètres suivants : le modèle de graphe calibré via le modèle à **clusters homogènes** et les engagements réciproques fixés suivant les hypothèses (**HCO**).

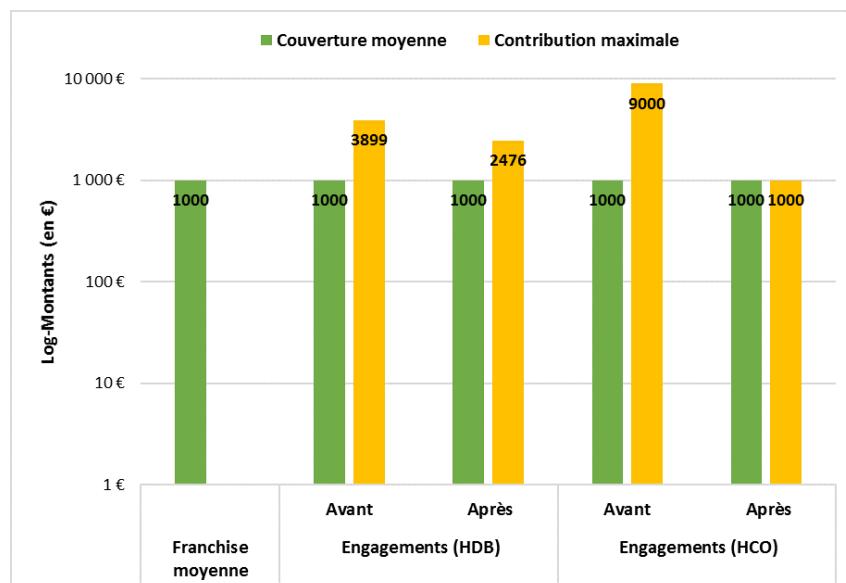


FIGURE 7.3 – Sorties (couverture moyenne et contribution maximale en échelle log) associées aux franchises de la communauté **DG**, avant et après les retraitements en clusters homogènes sous les hypothèses (**HDB**) et (**HCO**) concernant les engagements réciproques : Engagements (**HDB**) (a), Engagements (**HCO**) (b).

7.2 Données de Facebook (DF)

Cette section fait tout d'abord, une description des Données de Facebook (**DF**), ensuite une analyse des propriétés du graphe social de la communauté **DF**. Enfin, elle s'achève par une présentation des sorties (i.e. couverture moyenne et contribution maximale) associées aux franchises des assurés de la communauté **DF** en fonction des hypothèses (**HDB**) et (**HCO**) sur les engagements réciproques.

7.2.1 Description des données DF

Les Données de Facebook (**DF**) étudiées dans cette section sont issues des travaux de Leskovec and Krevl [2014]. Ces travaux consistent en une collection de plus de 50 jeux de données de grands réseaux (composés de dizaines de milliers de nœuds et d'arêtes) de la ville de Stanford aux USA. Ce sont entre autres, des réseaux de types : réseaux sociaux, graphiques web, réseaux routiers, réseaux internet, réseaux de citations, réseaux de collaborations et réseaux de communication. Dans le cadre de ce mémoire, on s'intéresse aux réseaux sociaux de type de *Facebook* téléchargeables via le lien <http://snap.stanford.edu/data/facebook.tar.gz>. Ce jeu de données est constitué de 10 communautés (ou réseaux) d'amis anonymisés sur *Facebook*, soit 4.039 nœuds (assurés) pour 88.234 arêtes (relations).

L'objectif est de supposer que des amis d'une même communauté peuvent s'assurer mutuellement. Pour ce faire, on s'intéresse à l'une des 10 communautés qui sera nommée "communauté **DF**" dans la suite, pour faire allusion à la communauté étudiée, issue des Données de Facebook. Ainsi, la communauté **DF** est composée de 333 assurés pour 2.519 relations. Le choix de cette communauté peut être motivé par l'ordre du graphe social obtenu qui est de 333 assurés. Ce qui constitue un ordre assez réaliste et en adéquation avec l'un des principaux principes du produit de franchise collaborative, consistant à responsabiliser les assurés en les connectant à un nombre raisonnable de relations.

7.2.2 Propriétés du graphe social de la communauté DF

Partant de la base initiale **DF** et des traitements effectués ayant permis l'extraction de la communauté **DF**, on déduit le graphe social puis la distribution des degrés des assurés de la communauté **DF**. La Figure 7.4 représente la distribution des degrés des assurés de la communauté **DF**.

On obtient une distribution de degrés variant de 1 à 77 relations avec une valeur moyenne d'environ 15 ± 15 relations. Aussi, d'après l'histogramme de la distribution des degrés (cf. Figure 7.4 (a)), on constate que 67% des assurés ont un degré inférieur au degré moyen d'environ 15 relations de la communauté **DF**, soit 223 sur les 333 assurés et seulement 8,7% des assurés ont un degré supérieur à 40 relations, soit 29 assurés sur les 333. Ce qui justifie la relative hétérogénéité de la distribution des degrés des assurés de la communauté **DF**.

De plus, la Figure 7.4 (b) représentant la fonction de survie des degrés en fonction des degrés des assurés (en échelle log-log), permet de conclure que le réseau formé par la communauté **DF** est approximativement sans échelle⁴, comme le réseau obtenu via le modèle de Barabási-Albert. En effet, on montre que le logarithme de la fonction de survie des degrés d'un réseau sans échelle est linéaire en fonction du logarithme des degrés (cf. Réseau sans échelle et loi de puissance en Annexe B).

4. Un réseau sans échelle (ou réseau invariant d'échelle, ou encore *scale-free network* en anglais) est un réseau dont les degrés suivent une loi de puissance. c'est-à-dire la proportion $P(k)$ d'assurés de degré k est telle que : $P(k) \sim k^{-\gamma}$, où le paramètre $\gamma > 0$ est appelé exposant d'invariance d'échelle.

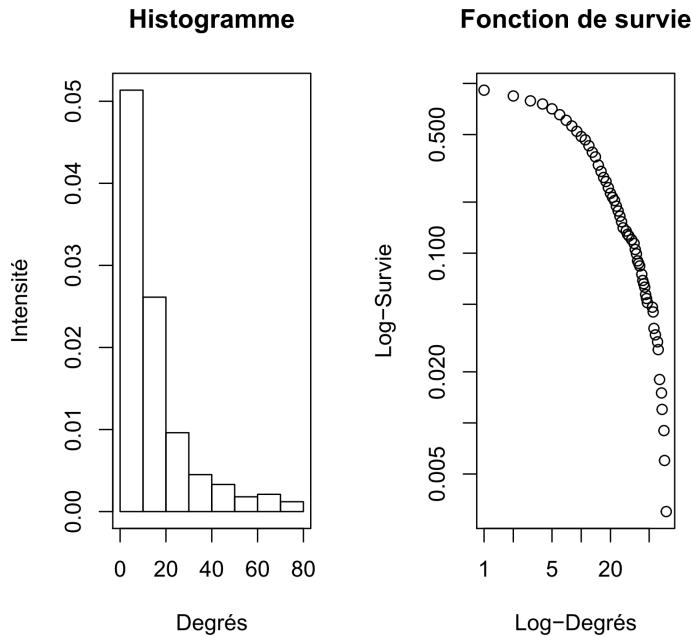


FIGURE 7.4 – Distribution des degrés des assurés de la communauté **DF** : Histogramme en échelle linéaire (a), Fonction de survie en échelle log-log (b).

Par ailleurs, dans le but d'obtenir un graphe social pour la communauté **DF**, calibré suivant le modèle à clusters homogènes permettant d'optimiser le produit de franchise collaborative, on effectue dans un premier temps un clustering (i.e. un regroupement des assurés en classes ou clusters) de la communauté **DF**. Pour ce faire, on utilise une approche d'optimisation rapide et gloutonne⁵ proposée par Clauset et al. [2004] sous la forme d'un algorithme de Classification Ascendante Hiérarchique (CAH)⁶ et implémentée via la fonction `fastgreedy.community` du package **igraph** du logiciel R. La Figure 7.5 représente le clustering obtenu pour la communauté **DF**, soit douze clusters de tailles respectives dans l'ordre croissant : 2, 2, 2, 3, 8, 10, 18, 24, 39, 52, 52, et 121.

Cependant, on constate que les membres de chaque cluster n'ont pas nécessairement le même nombre de relations. Ainsi, on procède dans un second temps, à des retraitements des connexions entre les assurés de chacune des douze classes, afin d'obtenir des sous-graphes complets (i.e. des cliques) et donc des clusters homogènes.

Enfin, notons que les retraitements en clusters homogènes permettent de faire passer le degré moyen de la communauté **DF** de 15 à 67 relations, soit un degré moyen environ 5 fois plus élevé que celui de l'état initial (i.e. avant retraitements). Ce qui permettrait donc, sous les hypothèses (**HDB**), de diminuer le montant des engagements réciproques et donc des contributions individuelles aux franchises, d'après leurs expressions mathématiques.

5. Un algorithme glouton (*greedy algorithm* en anglais) est un algorithme qui suit le principe de faire à chaque étape un choix optimum local sans revenir sur ses décisions, en espérant obtenir un résultat optimum global. Suivant les problèmes, cet algorithme ne fournit pas de garantie d'optimalité : on parle d'heuristique glouton, c'est-à-dire, obtention d'une solution réalisable non nécessairement optimale.

6. CAH. : Ensemble de méthodes de clustering détaillé en [Annexe B](#).

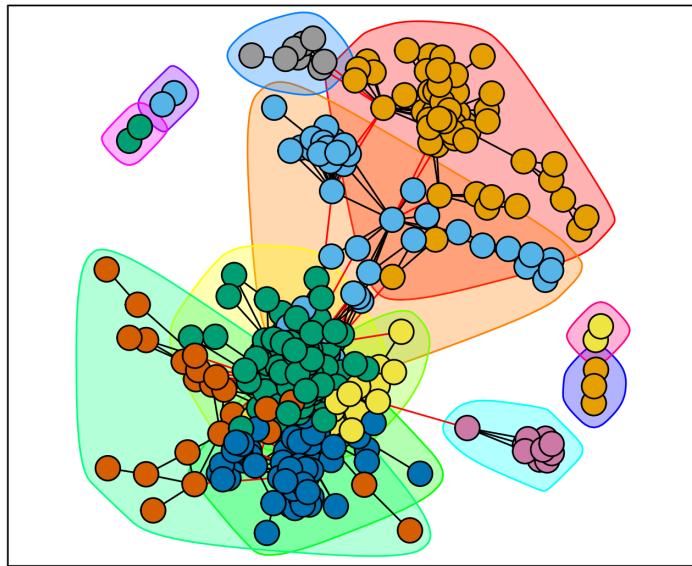


FIGURE 7.5 – Graphe social et clustering de la communauté **DF**

7.2.3 Sorties associées aux franchises de la communauté DF

Cette sous-section consiste à présenter les sorties (i.e. couverture moyenne et contribution maximale) associées aux franchises des assurés de la communauté **DF**, avant et après les retraitements en clusters homogènes. Partant des graphes sociaux respectifs de la communauté **DF**, avant et après les retraitements, on déduit les distributions des degrés des assurés pour chacun des deux cas (avant et après retraitements).

La Figure 7.6 représente les sorties associées aux franchises des assurés de la communauté **DF** avant et après les retraitements en clusters homogènes, sous respectivement les hypothèses (**HDB**) et (**HCO**) concernant les engagements réciproques entre les assurés.

D'après la Figure 7.6, on constate que les hypothèses (**HDB**) et (**HCO**) permettent tous deux, d'obtenir une couverture moyenne de franchise coïncidant avec les 1.000 € de franchise individuelle. c'est-à-dire, en moyenne un assuré de la communauté possède une couverture de franchise équivalente au montant de sa franchise individuelle. Ceci peut s'expliquer par les relations 4.7 et 4.18 exprimant l'égalité entre la couverture mutuelle de franchise (i.e. la somme des couvertures individuelles) et la franchise mutuelle (i.e. la somme des franchises individuelles) de la communauté, respectivement sous les conditions (**HDB**) et (**HCO**). Aussi, sous les hypothèses (**HDB**), les retraitements en clusters homogènes permettent comme annoncé au point 7.2.2, de diminuer la contribution maximale d'un assuré de 5.089 € à 1.790 €, soit une contribution maximale environ 3 fois moins coûteuse que celle de l'état initial (i.e. avant retraitements). De même, sous les hypothèses (**HCO**), les retraitements en clusters homogènes permettent de diminuer la contribution maximale d'un assuré de 4.961 € à 1.000 €, soit environ 5 fois moins coûteux par rapport à l'état initial. Ainsi, le modèle à clusters homogènes permet de minimiser la contribution maximale d'un assuré aux franchises de ses relations quelles que soient les hypothèses sur les engagements réciproques entre les assurés.

Par ailleurs, suivant la calibration du graphe social de la communauté **DF** via le modèle à clusters homogènes, on obtient des contributions maximales aux franchises de 1.790 € et de 1.000 €, respectivement sous les conditions (**HDB**) et (**HCO**), soit une économie de 790 € sous les hypothèses (**HCO**). Ce qui vient encore confirmer l'optimalité du **scénario central** pour

le produit de franchise collaborative, c'est-à-dire le scénario prenant en entrée les paramètres suivants : le modèle de graphe calibré via le modèle à **clusters homogènes** et les engagements réciproques fixés suivant les hypothèses (**HCO**).

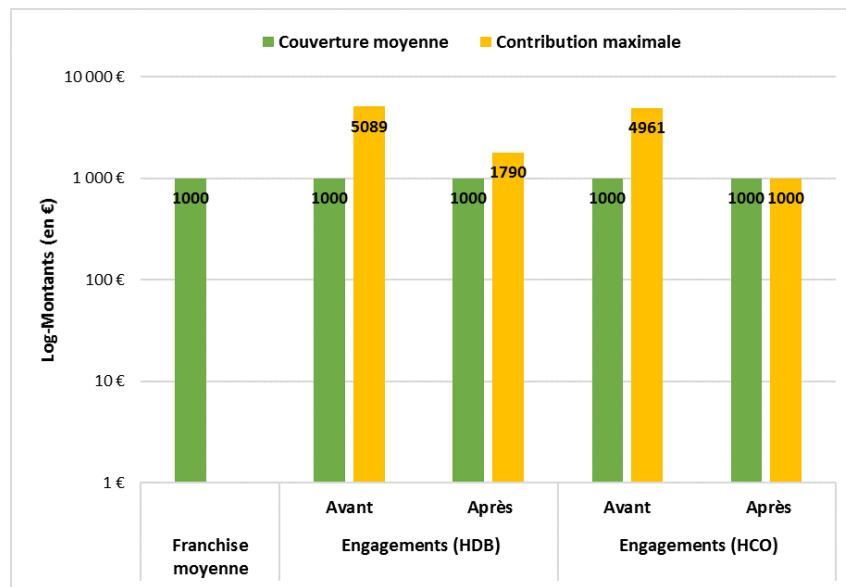


FIGURE 7.6 – Sorties (couverture moyenne et contribution maximale en échelle log) associées aux franchises de la communauté **DF**, avant et après les retraitements en clusters homogènes sous les hypothèses (**HDB**) et (**HCO**) concernant les engagements réciproques : Engagements (**HDB**) (a), Engagements (**HCO**) (b).

Synthèse de la partie III

Cette troisième et dernière partie du mémoire a consisté en la mise en place d'un produit de franchise collaborative pour différentes communautés dans le cadre d'un contrat d'assurance auto. Dans un premier temps, le chapitre 6 a consisté à établir trois variantes du produit de franchise collaborative pour la couverture des charges annuelles de la communauté fictive **CF**, notamment le **scénario central** permettant d'optimiser les sorties du programme de calcul **PC2** (i.e. la couverture moyenne des charges et la contribution maximale aux charges des assurés), la **sensibilité engagements** et la **sensibilité graphe social** permettant respectivement d'évaluer les impacts du montant des engagements réciproques et du modèle de graphe social sur les sorties du produit de franchise collaborative.

Dans un second temps, le chapitre 7 a été consacré à appliquer la méthodologie de calcul du programme **PC2** à des données réelles, à savoir des Données de Généalogie (**DG**) et des Données de Facebook (**DF**). L'étude de couverture des franchises annuelles des membres respectifs des communautés **DG** et **DF**, supposés s'assurer mutuellement, a permis de déterminer les sorties associées aux franchises des assurés de chacune de ces deux communautés.

Par ailleurs, cette étude a contribué à confirmer l'optimalité du scénario central sachant qu'il garantit en moyenne la couverture des charges des assurés et permet de limiter la contribution maximale d'un assuré aux charges de ses relations. Ainsi, le scénario central constitue la variante optimisée du produit de franchise collaborative qu'un assureur serait susceptible de proposer à une éventuelle communauté d'assurés, comme suit :

1) Le graphe social : calibré via le modèle à **clusters homogènes** suivant les deux cas envisageables :

- ✓ soit la connexion est **libre**, c'est-à-dire, les assurés peuvent former librement leurs groupes ou réseaux de relations avec la contrainte qu'un assuré ne peut appartenir qu'à un unique groupe indépendamment des autres groupes (e.g. entre amis, proches ou familles) ;
- ✓ soit la connexion est **imposée** par l'assureur qui peut alors mettre les assurés dans des groupes homogènes indépendants où les assurés appartenant à un même groupe ont des caractéristiques les plus similaires (e.g. âge ou profil de risque).

2) Les engagements réciproques : fixés suivant les hypothèses de couverture optimale (**HCO**) afin non seulement d'assurer une couverture optimale des charges, mais également combinées au point 1) de minimiser les contributions individuelles des assurés aux charges de leurs relations.

En effet, les hypothèses (**HCO**) combinées au point 1) permettent d'obtenir un engagement réciproque entre deux assurés appartenant à un même groupe fonction du degré du groupe de telle sorte que plus ce degré augmente (i.e. le nombre d'assurés du groupe augmente), plus cet engagement diminue. Ce qui pourrait être un **argument commercial** afin d'inciter les assurés à agrandir leur réseau de relations et permettrait ainsi de les fidéliser tout en accroissant le nombre de nouvelles souscriptions.

Conclusion

La problématique actuarielle traitée dans le cadre de ce mémoire consiste en l’analyse de l’impact de la structure du réseau des assurés sur les variables d’intérêt (couverture et contribution) du produit de franchise collaborative et en la mise en place efficace de mécanismes permettant l’optimisation du produit, c’est-à-dire, pour une communauté d’assurés : garantir la couverture des charges (franchises et dépenses⁷) annuelles d’un nombre maximum d’assurés tout en minimisant les contributions individuelles des assurés à la couverture des charges de leurs relations.

Pour ce faire, dans un premier temps sur la base d’une communauté fictive d’assurés, des études d’optimisation ont été respectivement effectuées sur les paramètres du produit de franchise collaborative, notamment le montant des engagements réciproques et le modèle de graphe social représentant le réseau des connexions entre les assurés de la communauté. Ainsi, d’une part, l’étude sur les engagements réciproques optimaux basée sur les hypothèses (**HDB**)⁸ et (**HCO**)⁹, a montré que les conditions (**HCO**) permettent d’optimiser les couvertures individuelles des charges annuelles des assurés avec un taux de couverture des charges de 100% (i.e. tous les assurés ont leurs charges couvertes). D’autre part, l’étude sur le modèle de graphe social optimal basée sur quatre modèles classiques de graphe de la théorie des réseaux, à savoir les modèles à **clusters homogènes**, d’**Erdős-Rényi**, de **Watts-Strogatz** et de **Barabási-Albert**, a sélectionné le modèle à clusters homogènes comme étant le modèle de graphe permettant d’optimiser sous les hypothèses (**HCO**), les contributions individuelles aux charges des assurés avec un taux de surcontribution aux franchises de 0% (i.e. la contribution maximale d’un assuré correspond au montant de sa franchise individuelle).

Dans un second temps, il a été mis en place un produit de franchise collaborative, d’une part, sur la base du produit optimisé dit **scénario central** prenant en entrée les paramètres optimaux issus des études d’optimisation du produit et d’une **étude de sensibilité** testant ces paramètres selon différentes sensibilités (i.e. dans le cas des scénarios défavorables) afin d’évaluer leurs impacts sur les sorties du produit, notamment la couverture moyenne des charges et la contribution maximale des assurés aux charges de la communauté fictive. Et d’autre part, sur la base d’applications à des données réelles, à savoir des données de généalogie et de réseaux sociaux tels que *Facebook* afin d’analyser comment ces types spécifiques de connexion entre les assurés peuvent influer sur les sorties du produit de franchise collaborative.

Par ailleurs, ces applications ont contribué à confirmer l’optimalité du scénario central sachant qu’il garantit en moyenne la couverture des charges des assurés et permet de limiter la contribution maximale d’un assuré aux charges de ses relations. Ainsi, le scénario central constitue la variante optimisée du produit de franchise collaborative qu’un assureur serait susceptible de proposer à une éventuelle communauté d’assurés, comme suit :

7. dépense = $\min(\text{franchise}, \text{coût sinistres})$.

8. Hypothèses dites de base (**HDB**) : fixent le montant des engagements réciproques comme fonction du degré moyen de la communauté (cf. section 4.1).

9. Hypothèses de couverture optimale (**HCO**) : fixent le montant des engagements réciproques comme fonction des degrés respectifs des assurés de la communauté (cf. section 4.2).

1) Le graphe social : calibré via le modèle à **clusters homogènes** suivant les deux cas envisageables :

- ✓ soit la connexion est **libre**, c'est-à-dire, les assurés peuvent former librement leurs groupes ou réseaux de relations avec la contrainte qu'un assuré ne peut appartenir qu'à un unique groupe indépendamment des autres groupes (e.g. entre amis, proches ou familles) ;
- ✓ soit la connexion est **imposée** par l'assureur qui peut alors mettre les assurés dans des groupes homogènes indépendants où les assurés appartenant à un même groupe ont des caractéristiques les plus similaires (e.g. âge ou profil de risque).

2) Les engagements réciproques : fixés suivant les hypothèses (**HCO**) afin non seulement de garantir une couverture optimale des charges, mais également combinées au point **1**) de minimiser les contributions individuelles des assurés aux charges de leurs relations.

En effet, les hypothèses (**HCO**) combinées au point **1**) permettent d'obtenir un engagement réciproque entre deux assurés appartenant à un même groupe, fonction du degré du groupe de telle sorte que plus ce degré augmente (i.e. le nombre d'assurés du groupe augmente), plus cet engagement diminue. Ce qui pourrait être un **argument commercial** afin d'inciter les assurés à agrandir leur réseau de relations et permettrait ainsi de les fidéliser tout en accroissant le nombre de nouvelles souscriptions.

En outre, sachant que ce mémoire s'est principalement intéressé à l'aspect technique du produit de franchise collaborative, c'est-à-dire à la modélisation actuarielle du produit et la mise en place efficace de mécanismes permettant l'optimisation de ses paramètres, sur la base d'hypothèses portant sur la sinistralité des assurés de la communauté fictive ; des travaux futurs pourraient s'intéresser à l'aspect commercial du produit par la mise en place du scénario central (i.e. le produit optimisé) sur la base d'un portefeuille réel de sinistres d'un assureur non-vie, combiné au mécanisme de *cashback* associé au modèle de micro-mutuelle collective (cf. sous-section 1.2.3), comme suit :

- 1) Premièrement**, une tarification *a priori* sur la base du portefeuille réel, des contributions individuelles ou cotisations des assurés, encaissées en début d'exercice contrairement au modèle d'*Inspeer* qui percevait les contributions des assurés lors de la réalisation d'un sinistre. Ce qui engendrait indéniablement un risque de contrepartie des assurés ;
- 2) Deuxièmement**, une redistribution *a posteriori* d'une partie des cotisations encaissées sous forme de *cashback* en l'absence de sinistre dans l'année, afin de responsabiliser les assurés et améliorer par ailleurs, l'expérience client.

Table des figures

2.1	Sociogramme de Moreno	16
2.2	Réseau de disciplines	16
2.3	Réseau métabolique	17
2.4	Réseau textuel autour des élections américaines de 2012	17
2.5	Visualisations du problème des sept ponts de Königsberg résolu par Euler	18
2.6	Graphe social : Exemple clusters homogènes	21
2.7	Illustration de la génération d'un graphe aléatoire $G(n, p)$ par le modèle d'Erdős–Rényi	23
2.8	Illustration du processus de reconnexion des arêtes en fonction de la probabilité de reconnexion p	24
2.9	Impact du paramètre d'attachement m sur la distribution de degrés des graphes générés via le modèle de Barabási-Albert	25
2.10	Comparaison des fonctions de survie des distributions de degrés issues des modèles d'Erdős–Rényi et de Barabási-Albert	26
4.1	Exemple illustratif : Graphe social	47
5.1	Boîte à moustaches des charges de la communauté CF	52
5.2	Diagrammes en secteurs des charges de la communauté CF	53
5.3	Histogramme distribution de degrés : Modèle à clusters homogènes	54
5.4	Histogramme distribution de degrés : Modèle d'Erdős-Rényi	55
5.5	Histogramme distribution de degrés : Modèle de Watts-Strogatz	56
5.6	Histogramme distribution de degrés en échelle log : Modèle de Barabási-Albert	57
5.7	Diagramme en bâtons : Distributions de degrés par modèle de graphe	58
5.8	Histogrammes des distributions de couverture des franchises par modèle de graphe	61
5.9	Histogrammes des distributions de couverture des dépenses par modèle de graphe	63
5.10	Couverture de franchise en fonction du degré par modèle de graphe	64
5.11	Fonctions de densité des ratios de couverture des franchises par modèle de graphe	66
5.12	Fonctions de densité des ratios de couverture des dépenses par modèle de graphe	67
5.13	Histogrammes des distributions de contribution aux franchises par modèle de graphe	71
5.14	Histogrammes des distributions de contribution aux dépenses par modèle de graphe	72
5.15	Contribution aux franchises en fonction du degré par modèle de graphe	74
5.16	Fonctions de densité des ratios de contribution aux franchises par modèle de graphe	76
5.17	Fonctions de densité des ratios de contribution aux dépenses par modèle de graphe	77
6.1	Sorties associées aux franchises par variante du produit	85
6.2	Sorties associées aux dépenses par variante du produit	86
6.3	Sorties associées aux franchises par variante du produit	88
6.4	Sorties associées aux dépenses par variante du produit	88
6.5	Sorties associées aux franchises par variante du produit	90
6.6	Sorties associées aux dépenses par variante du produit	91
6.7	Sorties de la sensibilité graphe social associées aux franchises en fonction du degré moyen	93

6.8	Sorties de la sensibilité graphe social associées aux dépenses en fonction du degré moyen	95
6.9	Contribution maximale aux franchises en fonction du degré moyen	97
6.10	Contribution maximale aux dépenses en fonction du degré moyen	99
7.1	Graphe social de la communauté DG	101
7.2	Tailles et degrés moyens par famille de la communauté DG	102
7.3	Sorties associées aux franchises de la communauté DG	103
7.4	Distribution des degrés des assurés de la communauté DF	105
7.5	Graphe social et clustering de la communauté DF	106
7.6	Sorties associées aux franchises de la communauté DF	107

Liste des tableaux

1	Tableau récapitulatif : Indicateurs de couverture par modèle de graphe	vii
2	Tableau récapitulatif : Indicateurs de contribution par modèle de graphe	vii
3	Tableau récapitulatif : Variantes du produit de franchise collaborative établies . .	viii
4	Summary table : Coverage indicators by graph model	xiii
5	Summary table : Contribution indicators by graph model	xiii
6	Summary table : Variants of the peer-to-peer deductible product established	xiv
1.1	Intérêts et limites de l'assurance collaborative	5
1.2	Tableau récapitulatif Generali	7
1.3	Tableau récapitulatif Friendsurance	7
1.4	Tableau récapitulatif Bought By Many	8
1.5	Tableau récapitulatif Guevara	8
1.6	Tableau récapitulatif Inspeer	9
1.7	Tableau récapitulatif Lemonade	10
1.8	Tableau récapitulatif Groupago	10
1.9	Tableau récapitulatif Wizzas	11
1.10	Tableau récapitulatif Otherwise	11
1.11	Tableau récapitulatif Wecover	12
1.12	Tableau récapitulatif Besure	12
1.13	Tableau récapitulatif Teambrella	13
4.1	Exemple illustratif : Caractéristiques de la communauté	46
4.2	Exemple illustratif : Distributions de couverture des charges	49
4.3	Exemple illustratif : Distributions de contribution aux charges	49
5.1	Tableau récapitulatif : Distributions de degrés par modèle de graphe	58
5.2	Tableau : Distributions de couverture des franchises par modèle de graphe	61
5.3	Tableau : Distributions de couverture des dépenses par modèle de graphe	63
5.4	Tableau récapitulatif : Indicateurs de couverture par modèle de graphe	68
5.5	Tableau : Distributions de contribution aux franchises par modèle de graphe . . .	71
5.6	Tableau : Distributions de contribution aux dépenses par modèle de graphe . . .	72
5.7	Tableau récapitulatif : Indicateurs de contribution par modèle de graphe	78
6.1	Tableau récapitulatif : Variantes du produit de franchise collaborative établies . .	83
6.2	Intervalles de confiance : Contribution maximale aux franchises	93
6.3	Intervalles de confiance : Contribution maximale aux dépenses	95
1	Tableau récapitulatif : Nombre d'observations par famille de la base DG	i
2	Tableau récapitulatif : Variables de la base DG	ii

Annexes

Annexe A

Données de Généalogie (DG)

Cette annexe consiste en un complément des descriptions des Données de Généalogie (**DG**) présentées dans la section 7.1. Le Tableau 1 résume le nombre d'observations (lignes) initiales et restantes après traitements (i.e. après suppression des doublons et des boucles¹⁰) pour chacune des 19 familles de la base **DG**.

Lignes Familles	Nombre d'observations initiales	Nombre de doublons supprimées	Nombre de boucles supprimées	Nombre d'observations restantes
Famille 1	82	49	5	28
Famille 2	177	113	8	56
Famille 3	492	218	13	261
Famille 4	561	365	14	182
Famille 5	206	119	9	78
Famille 6	209	109	10	90
Famille 7	137	73	8	56
Famille 8	165	44	11	110
Famille 9	155	91	8	56
Famille 10	119	83	6	30
Famille 11	15	11	2	2
Famille 12	273	152	11	110
Famille 13	207	143	8	56
Famille 14	173	92	9	72
Famille 15	65	40	5	20
Famille 16	77	41	6	30
Famille 17	96	60	6	30
Famille 18	350	125	15	210
Famille 19	35	19	4	12
Total	3594	1947	158	1489

TABLE 1 – Tableau récapitulatif : Nombre d'observations par famille de la base **DG**

10. On a une boucle lorsque l'identifiant d'une personne coïncide avec l'identifiant de sa relation.

Le Tableau 2 récapitule les différentes variables (colonnes) présentes initialement dans la base **DG** et les variables conservées pour l'étude sur la communauté **DG**.

Nomenclature des variables	Description des variables	Variables conservées
id_personne_num	identifiant de la personne	✓
id_family_member	identifiant de la personne liée	✓
connection	type de connexion entre les deux personnes (“sibling” ou “cousins”)	✗
side	NA pour les “sibling” et “mother” ou “father” si le cousin est du côté de la mère ou du père, respectivement	✗
nom	nom de famille de la personne	✗
prenoms	prénoms de la personne	✗
sexe	sexe de la personne (1 pour homme, 2 pour femme)	✗
date_N	date de naissance de la personne au format YYYYMMDD	✗
date_D	date de décès de la personne au format YYYYMMDD	✗
nom_family_member	nom de famille de la personne liée	✗
prenoms_family_member	prénoms de la personne liée	✗
sexe_family_member	sexe de la personne liée (1 pour homme, 2 pour femme)	✗
date_N_family_member	date de naissance de la personne liée au format YYYYMMDD	✗
date_D_family_member	date de décès de la personne liée au format YYYYMMDD	✗

TABLE 2 – Tableau récapitulatif : Variables de la base **DG** (variables conservées : ✓ et variables non conservées : ✗).

Annexe B

Rappels théoriques

Cette annexe consiste en un rappel de notions théoriques abordées dans ce mémoire, notamment les notions de réseaux sans échelle et de Classification Ascendante Hiérarchique (CAH).

B.1 Réseau sans échelle et loi de puissance

Un réseau sans échelle (ou réseau invariant d'échelle, ou encore *scale-free network* en anglais) est un réseau dont la distribution des degrés est une loi de puissance. C'est-à-dire la proportion $f(k)$ d'assurés de degré k est telle que : $f(k) \sim k^{-\gamma}$, où le paramètre $\gamma > 0$ est appelé exposant d'invariance d'échelle.

Graphiquement, on reconnaît un réseau sans échelle (ou approximativement sans échelle) si on observe une tendance linéaire dans le tracé de la fonction de survie ou de la distribution des degrés de ce réseau en échelle logarithmique. En effet, le logarithme de la fonction de survie $S(k)$ de la distribution des degrés $f(k)$ d'un réseau sans échelle est linéaire en fonction du logarithme des degrés $k \in \mathbb{R}^+ : \forall t \geq 0$ tel que $k \in [t; +\infty[$,

$$\begin{aligned}
 S(t) &= \int_t^{+\infty} f(k) dk \\
 S(t) &= \int_t^{+\infty} k^{-\gamma} dk \\
 &= \frac{1}{-\gamma + 1} [k^{-\gamma+1}]_t^{+\infty} \\
 &= \frac{1}{-\gamma + 1} [0 - t^{1-\gamma}] , \text{ si } \gamma > 1 \\
 S(t) &= \frac{1}{\gamma - 1} t^{1-\gamma} , \text{ si } \gamma > 1
 \end{aligned} \tag{1}$$

D'où, en passant aux logarithmes dans la relation 1, on obtient : $\forall \gamma > 1$,

$$\begin{aligned}
 \log(S(t)) &= (1 - \gamma) \log(t) - \log(\gamma - 1) \\
 \log(S(t)) &= a \log(t) + b ,
 \end{aligned} \tag{2}$$

avec $a = 1 - \gamma < 0$ et $b = -\log(\gamma - 1) \in \mathbb{R}$.

D'après la relation 2, on a bien la linéarité (décroissante car le coefficient directeur $a < 0$) du logarithme de la fonction de survie des degrés par rapport au logarithme des degrés. \square

B.2 Classification Ascendante Hiérarchique (CAH)

La méthode de Classification Ascendante Hiérarchique (CAH) est une technique très populaire de regroupement procédant par fusion successive d'ensembles de points appelés clusters. Intuitivement, au début de l'algorithme on crée autant de classes que d'individus puis, ces classes sont regroupées deux à deux jusqu'à l'obtention d'une seule classe suivant un critère de ressemblance. L'algorithme CAH est le suivant :

Entrée : Variable aléatoire X de densité f ou table de données X

Code de calcul (Φ) :

- Calcul de la table des distances [1] entre individus. Chaque individu constitue un cluster ;
- **Répéter** : Détecter les 2 clusters les plus proches et les agréger [2] ;
- **Jusqu'à** : tous les individus forment un seul cluster.

Sortie : $Y = \Phi(X)$ définie par un indicateur de partitionnement

[1] La distance d entre deux individus a et b d'une population de taille p peut être :

- ✓ euclidienne : $d^2(a, b) = \sum_{j=1}^p (x_j(a) - x_j(b))^2$;
- ✓ euclidienne pondérée par l'inverse de la variance : $d^2(a, b) = \sum_{j=1}^p \frac{1}{\sigma_j^2} (x_j(a) - x_j(b))^2$;
- ✓ cosinus : $d(a, b) = 1 - \cos(a, b)$.

[2] La distance d entre deux groupes G_1 et G_2 peut être définie par :

- ✓ single linkage ou saut minimum : $d(G_1, G_2) = \min(a, b)$, $a \in G_1$ et $b \in G_2$;
- ✓ complete linkage ou saut maximum : $d(G_1, G_2) = \max(a, b)$, $a \in G_1$ et $b \in G_2$;
- ✓ distance de Ward : $d^2(G_1, G_2) = \frac{n_1 n_2}{n_1 + n_2} d^2(B_1, B_2)$, où n_i et B_i représentent respectivement la taille et le barycentre du groupe G_i , $i = 1, 2$.

La formule de Ward détermine la distance entre deux groupes à partir de la distance pondérée entre les barycentres.

Bibliographie

- Cnc. Avis du Conseil national de la consommation sur les plateformes numériques collaboratives, 2016. https://www.economie.gouv.fr/files/files/directions_services/cnc/avis/2016/avis-cnc-plateformes-numeriques-collaboratives-janvier2016.pdf.
- U. Dubois. Tarification d'un stoploss en assurance IARD dans le cadre d'un modèle d'assurance collaboratif. Master's thesis, EURIA, 2015. <https://www.institutdesactuaires.com/se-documenter/memoire-d-actuariat-38?id=cf446a09e9eaaff84fba27ec16d0c2b5>.
- M. Grandjean. Social network analysis and visualization : Moreno's sociograms revisited, 2015.
- B. Coulmont. L'espace de la multidisciplinarité, 2020. ISSN 2269-1960. <http://coulmont.com/blog/2020/02/27/lespace-de-la-multidisciplinarite>.
- T. Vickers, 2007. https://commons.wikimedia.org/wiki/File:A_thaliana_metabolic_network.png.
- S. Sudhahar, GA. Veltri, and N. Cristianini. Automated analysis of the us presidential elections using big data and network analysis. *Big Data & Society*, 2(1), 2015. ISSN 1-28.
- L. Euler. The solution of a problem relating to the geometry of position. *Commentarii academiae scientiarum Petropolitanae*, 8(128-140), 1735. ISSN 2572-6803. <https://scholarlycommons.pacific.edu/euler-works/53>.
- A. Channarond. *Recherche de structure dans un graphe aléatoire : modèles à espace latent*. PhD thesis, Université Paris Sud - Paris XI, 2013. <https://tel.archives-ouvertes.fr/tel-01157186>.
- P. Erdős and A. Rényi. On Random Graphs I. *Publicationes Mathematicae Debrecen*, 6 :290–297, 1959.
- P. Erdős and A. Rényi. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, 12(1) :261–267, 1961a.
- P. Erdős and A. Rényi. On the evolution of random graphs. *Bull. Inst. Internat. Statist*, 38(4) : 343–347, 1961b.
- P. Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4) :292–294, April 1947. <https://projecteuclid.org:443/euclid.bams/1183510596>.
- Jmcatania, Dec 2011. <https://commons.wikimedia.org/w/index.php?curid=85075052>.
- D. Watts and S. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 398 :440–442, 1998. <https://doi.org/10.1038/30918>.
- A. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439) : 509–512, 1999.
- HeMath, Nov 2017. <https://commons.wikimedia.org/w/index.php?curid=64122479>.

- A. Charpentier. Introduction to Graphs & Networks. Note de cours, Master Statistique & Économétrie - Université Rennes 1, 2017.
- J. Leskovec and A. Krevl. SNAP Datasets : Stanford large network dataset collection, June 2014. <http://snap.stanford.edu/data>.
- A. Clauset, M. E. J. Newman, and C. Moore. Finding community structure in very large networks. *Physical Review E*, 70(6), Dec 2004. ISSN 1550-2376. <http://dx.doi.org/10.1103/PhysRevE.70.066111>.
- A. Horvath, Jan 2009. <https://commons.wikimedia.org/w/index.php?curid=5815171>.
- E.D. Kolaczyk and G Csárdi. *Statistical Analysis of Network Data with R*. Springer Verlag New York, 2014.
- M. Grandjean. La connaissance est un réseau. *Les Cahiers du Numérique*, 10(3), 2014. ISSN 37-54. doi : 10.3166/LCN.10.3.37-54.